

Robust Ground Plane Tracking in Cluttered Environments from Egocentric Stereo Vision

Tobias Schwarze and Martin Lauer¹

Abstract—Estimating the ground plane is often one of the first steps in geometric reasoning processes as it offers easily accessible context knowledge. Especially unconstrained platforms that capture video from egocentric viewpoints can benefit from such knowledge in various ways. A key requirement here is keeping orientation, which can be greatly achieved by keeping track of the ground. We present an approach to keep track of the ground plane in cluttered inner-urban environments using stereo vision in real-time. We fuse a planar model fit in low-resolution disparity data with the direction of the vertical vanishing point. Our experiments show how this effectively decreases the error of plane attitude estimation compared to classic least-squares fitting and allows to track the plane with camera configurations in which the ground is not visible. We evaluate the approach using ground-truth from an inertial measurement unit and demonstrate long-term stability on a dataset of challenging inner city scenes.

I. INTRODUCTION

A fundamental requirement for any robotic system which aims at perceiving and understanding unknown environments is the ability to acquire information about the coarse scene structure. Being able to distinguish salient foreground objects from navigable background space constitutes one of the challenging tasks for such systems. Navigable space can often be seen as a continuous surface in a limited area around the point of view. This especially holds in structured environments. Consequently, the ground surface is often expressed through geometric models. These representations offer simple but effective context knowledge which can be incorporated into subsequent reasoning steps, as e.g. demonstrated in [1]. In this work we treat the measuring and tracking of such model from a free moving stereo camera system.

In mobile robotics plane parameter tracking is most frequently treated from mono-vision by tracking sparse features. The homography induced by the surface seen from different viewpoints is usually estimated to this end [2] and tracked with particle [3] or Kalman filters [4]. Sometimes the interest is not the plane itself but the camera motion [5], [6].

Using stereo vision or depth data, ground surface estimation is often found as preprocessing step towards free-space estimation or obstacle detection, where regions of interest are created based on the elevation over ground.

Surface modelling is an important building block in the context of autonomous vehicles to estimate drivable free-space. A popular method was introduced by Labayrade et al.



Fig. 1: Two typical inner-urban street scenes.

using the v-disparity [7], a disparity row histogram in which planar surfaces can be obtained by 2D line fitting. In its plain form it requires to compensate the camera roll beforehand, which makes it difficult to apply in unconstrained camera setups. When vehicle fixed camera attitude is given, the necessary adoptions to compensate changes in roll and pitch over time are comparably small, but often more expressive geometric models like splines [8] or clothoids [9] are fitted here due to the larger range of interest.

In more unconstrained setups planar surfaces or ground planes are usually fitted in the disparity domain [10], [11], or can also be estimated directly from stereo images [12]. When the camera motion is small enough, the last known plane parameters are a sufficient starting point to find the new parameters by refinement [13]. In a similar setting to ours, [14] assumes small camera roll and uses the v-disparity domain to find ground support points for a subsequent least-squares plane fit. All these approaches still require at least a portion of the ground to be visible in every estimation step. Strong camera motion is not taken into consideration, tracking plane parameters only succeeds as long as inter-frame camera pose changes are small enough and the ground is still the dominant plane.

In this work we consider vision systems that perceive the environment from an egocentric viewpoint. In our case this is a calibrated stereo camera mounted close to a person's head. There are various challenges of such unconstrained platforms that existing approaches usually do not deal with: First, camera motion is too strong to use the last known estimate as initial guess for parameter optimization as in [13], [12]. Secondly, a bad camera elevation angle often leads to invisibility of the ground for many consecutive frames. Approaches that require the ground to be the dominant plane cannot deal with such situation and also fail in cases of total occlusion due to objects and scene clutter. Our goal is to estimate and track the ground from a dense but low-resolution stereo disparity image under these conditions.

¹The authors are with the Institute of Measurement and Control Systems, Karlsruhe Institute of Technology, Karlsruhe, Germany tobias.schwarze@kit.edu

The next section gives an overview of the components we use to achieve this goal. In Section III we explain our filter framework which fuses all information into the ground plane estimate. We evaluate the approach in Section IV on datasets of cluttered inner-urban scenes with ground-truth measurements from an inertial measurement unit.

II. SYSTEM

We model the ground as a planar surface, which appears to be a valid assumption in our scenarios, considering a limited range of interest of around 15m. To obtain a parameter prediction with strong camera movement we estimate the camera motion by visual odometry. We then use two complementary measurements to correct the prediction. First, we find an optimized plane using a robust least-squares fit in disparity data. Secondly, we estimate the vertical vanishing direction, which coincides with the normal vector of the (non-inclined) ground plane in a calibrated camera. This second measurement allows us to carry out a correction without the ground plane being actually visible. The idea is probably most similar to [15], here the vanishing directions are used to constrain a homography estimation from omnidirectional images. The individual steps are detailed in the following.

A. Visual Odometry

The process of estimating the motion of a camera solely from image data is commonly referred to as visual odometry (VO). The general goal is to find the transformation in all six degrees of freedom that relates the camera poses of frame $k-1$ and k . Various standalone implementations exist, an overview can be found in [16].

In this work we employ libViso2 [17]; the frame to frame transformation is provided as a translation vector \mathbf{t} and a 3x3 rotation matrix R which can be written as an affine transformation $T = \begin{bmatrix} R & \mathbf{t} \\ 0^T & 1 \end{bmatrix}$. Since this estimation is done incrementally it is inevitably subject to drift.

B. Iterative Least-Squares plane fitting

We represent a plane in $uv\delta$ space as

$$\alpha u_n + \beta v_n + \gamma - \delta = 0 \quad (1)$$

using normalized image coordinates $u_n = \frac{u-c_u}{f}$, $v_n = \frac{v-c_v}{f}$, with camera focal length f and principal point (c_u, c_v) .

The initial ground plane is found using the RANSAC scheme by repeatedly sampling planes through 3 random points. A plane is evaluated by counting the support points with point-to-plane distance $|\alpha u_n + \beta v_n + \gamma - \delta|$ smaller than a disparity margin ϵ around the plane.

Having obtained an initial solution we can optimize the parameters using a robust iterative least-squares estimator. The set of $uv\delta$ plane support points is selected by validating the plane equation $|\alpha u_n + \beta v_n + \gamma - \delta| \leq \epsilon$. The optimized $uv\delta$ -plane parameters are then result of

$$\begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = (H^T H)^{-1} H^T \mathbf{y} \quad (2)$$

with H being the measurement matrix $[u_{n_{1..n}} \ v_{n_{1..n}} \ 1]$ and \mathbf{y} the corresponding disparity measurements. We apply the least-squares estimation a few iterations until no considerable update in the parameters remains.

C. Vanishing point estimation

Given a calibrated camera system, lines in 3D-space are projected to lines in the image plane. The intersection of projected parallel space lines is known as vanishing point.

In the calibrated camera case, each vanishing point defines a direction vector originating from the camera principal point c_u, c_v . If we assume a non-inclined ground plane, the vertical vanishing direction coincides with the normal vector of the ground plane.

We estimate this vanishing point following the approach of Tardif [18]. Based on a Canny edge detector, edge crossings are suppressed and connected points extracted by flood fill region growing. The edge hypotheses are split into straight segments. The remaining segments longer than 30px are fitted with a line model and constitute the edge list ξ .

From ξ we seek the subset of edges ξ_{vert} , which support the expected direction of the vertical vanishing point. We initialize this direction with the current ground plane normal vector \mathbf{n} . Once the vanishing point is initialized we predict its direction using the visual odometry rotation R between frame k and $k-1$ as

$$\mathbf{n}_{VP,k}^- = R \mathbf{n}_{VP,k-1} \quad (3)$$

To evaluate the support of an edge ξ_j for a given vanishing point \mathbf{n}_{VP} we define its error as the orthogonal distance $D_j(\mathbf{n}_{VP}, \xi_j)$ of one of the line endpoints to the line connecting the vanishing point with the edge centroid [18]. All edges with $D < \epsilon_{VP}$ are considered when updating the vanishing point through minimizing

$$\mathbf{n}_{VP,k}^+ = \min_{\mathbf{n}_{VP}} \sum_{\xi_j \in \xi_{\text{vert}}} D_j(\mathbf{n}_{VP}, \xi_j) \quad (4)$$

Internally we represent the vanishing direction \mathbf{n}_{VP} with spherical coordinates (θ, φ) and also perform minimization in this domain.

III. FILTER

The aim of our filter framework is to fuse all measurements described in the previous section. Note that these input features are selected in such a way that they provide stochastically independent and complementary information about the ground plane. A recursive estimator is desired to enable access to the current best estimate and retain real-time capabilities. A Kalman filter fulfills these requirements and is chosen here. However, we need to fuse measurements in camera coordinates (plane fit) with measurements in world coordinates (vanishing direction, egomotion T), the transformation between which is non-linear. Hence, we apply an extended Kalman filter (EKF) with state transition and observation model

$$\mathbf{x}_k = f(\mathbf{x}_{k-1}, \mathbf{u}_k, w_k), \quad (5)$$

$$\mathbf{z}_k = h(\mathbf{x}_k, v_k), \quad (6)$$

$\mathbf{x}_k \in \mathbb{R}^n$ being state of the system at time t_k with associated measurements $\mathbf{z}_k \in \mathbb{R}^m$ and $f(\cdot)$ and $h(\cdot)$ being the non-linear process and measurement functions with w_k and v_k the process and measurement noises.

Our system state $\mathbf{x} = [\theta, \varphi, d]^T$ consists of the ground plane normal \mathbf{n} in spherical coordinates θ, φ and the orthogonal camera to plane distance d in Euclidean XYZ space.

Prediction We use the visual odometry transformation T between last frame $k-1$ and current frame k to obtain the predicted plane parameters for the current view. An XYZ -plane $\mathbf{p}_{k-1} = \begin{bmatrix} \mathbf{n} \\ d \end{bmatrix}$ transforms via

$$\mathbf{p}_k = (T^{-1})^T \mathbf{p}_{k-1}, \quad (7)$$

where

$$(T^{-1})^T = \begin{bmatrix} R & 0 \\ (-R^T \mathbf{t})^T & 1 \end{bmatrix}. \quad (8)$$

The according process function $f(\cdot)$ uses T as control input and is calculated as

$$\mathbf{x}_k^- = f([\theta, \varphi, d]^T, T) \quad (9)$$

$$= g_{sph} \left((T^{-1})^T g_{sph}^{-1}([\theta, \varphi, d]^T) \right) \quad (10)$$

where $g_{sph}(\cdot)$ transforms the Euclidean plane representation into spherical coordinates

$$\begin{bmatrix} \theta \\ \varphi \\ d \end{bmatrix} = g_{sph} \left(\begin{bmatrix} \mathbf{n} \\ d \end{bmatrix} \right) = \begin{bmatrix} \arccos(n_z) \\ \text{atan2}(n_y, n_x) \\ d \end{bmatrix} \quad (11)$$

Measurement prediction The current state prediction \mathbf{x}_k^- is used as starting point for the iterative least-squares estimation as described in section II-B. This requires to transform \mathbf{x} to normalized $uv\delta$ plane parameters $[\alpha, \beta, \gamma]$ via

$$\begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = -\frac{Bf}{d} \mathbf{n} = g_{uv\delta} \left(\begin{bmatrix} \theta \\ \varphi \\ d \end{bmatrix} \right) = -\frac{Bf}{d} \begin{bmatrix} \sin(\theta) \cos(\varphi) \\ \sin(\theta) \sin(\varphi) \\ \cos(\theta) \end{bmatrix} \quad (12)$$

with stereo baseline B and camera focal length f .

Correction From (2) we obtain a measurement $\mathbf{m}_{LS} = [\alpha, \beta, \gamma]^T$. The vanishing point estimation according to Section II-C results in a measurement $\mathbf{m}_{VP} = [\theta, \varphi]^T$.

Using $g_{uv\delta}(\mathbf{x}_k^-)$ as filter measurement prediction we apply a correction step with \mathbf{m}_{LS} . The vanishing direction measurement \mathbf{m}_{VP} is directly applied as additional correction step. Alternatively both can be combined in one step. As long as the camera elevation w.r.t. to the ground plane stays below a fixed angle (in our setup 75° , generally depending on the vertical camera opening angle) the ground plane is still visible and a least-squares ground plane measurement can be obtained. In this case we correct the state using both measurements. Whenever the camera is tilted further away we only consider \mathbf{m}_{VP} to correct the ground plane attitude.

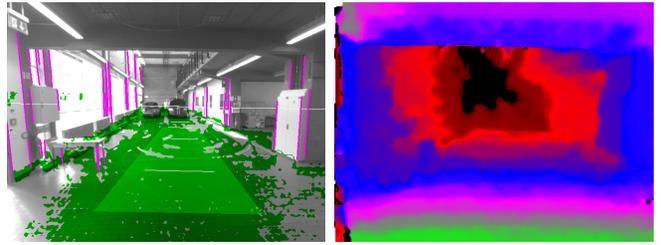


Fig. 2: Measurement setup for parameter variance estimation. Left: plane support points are coloured green, edges supporting the vertical vanishing point are highlighted pink. Right: Corresponding disparity image at half-resolution.

We validate the correction by setting a validation gate around $e^2 = \mathbf{r}^T S^{-1} \mathbf{r}$ with update residual \mathbf{r} and measurement prediction covariance S .

IV. EXPERIMENTS

Our experimental setup consists of a calibrated stereo rig with a baseline of around 18 cm and wide-angle lenses of 3.5 mm focal length mounted on a helmet. Having real-time applicability on a mobile platform in mind we capture images with 640×480 px at 30 fps. We use libViso2 [17] to estimate the camera motion and acquire the disparity with off-the-shelf estimators like the OpenCV semi-global matching (SGBM) at half resolution (320×240 px). Visual odometry and disparity estimation are done parallel with around 15 fps, the subsequent vanishing point estimation and least-squares fitting are carried out with the same rate on an i7 dual-core notebook with 2.4GHz. All experiments are done in real-time, i.e. frames are skipped when processing is slower than the capture rate of 30 fps.

A. Process and Measurement noise

In order to parameterize the process and measurement noise w_k and ϵ_k of the EKF we empirically estimate the measurement variances of the least-squares parameter fit for α, β, γ , the vanishing point direction $\theta_{VP}, \varphi_{VP}$ and the uncertainty in estimated camera motion in a set of experiments.

For the $uv\delta$ -plane fitting we re-estimate the plane parameters 1000 times from a fixed camera position (see Figure 2), using the same initial solution for the support point selection. In the same manner we obtain variances for $\theta_{VP}, \varphi_{VP}$ by re-evaluating (4) from an initial solution.

For a $uv\delta$ -plane fitted in a half-resolution SGBM disparity image the parameters vary with

$$\sigma_\alpha = 0.038 \quad \sigma_\beta = 0.045 \quad \sigma_\gamma = 0.016.$$

The vanishing point direction varies with

$$\sigma_{\theta_{VP}} = 0.0007 \quad \sigma_{\varphi_{VP}} = 0.0004.$$

The process noise needs to cover the uncertainty in parameter prediction using the estimated camera motion. While there is no straight forward uncertainty estimation for each single visual odometry estimation, we approximate its standard deviation of rotation and translation by evaluating short sequences on the KITTI [19] odometry dataset and compare it to groundtruth data. We find

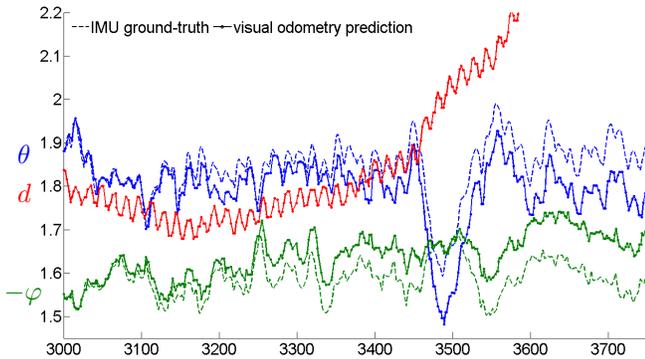


Fig. 3: Ground plane parameters (θ, φ, d) predicted by visual odometry without correcting drift. Dashed lines are the IMU ground truth, solid lines the predicted parameters.

$$\sigma_R = 0.00073 \quad \sigma_t = 0.046.$$

Using process noise distributed with σ_R for θ and φ and σ_t for the distance parameter d puts much weight onto the prediction. In situations, in which the actual uncertainty is larger this quickly leads to tracking-loss, since the drifting prediction cannot be corrected fast enough. Due to our unconstrained platform the accumulated odometry error adds up much higher than in the KITTI benchmark used for evaluation, we therefore increase the process noise to $\sigma_R = 0.001$ and $\sigma_t = 0.05$.

B. Evaluation against IMU Ground-Truth

We mounted an inertial measurement unit (IMU) close to the camera setup in order to evaluate the accuracy of the ground plane attitude (θ and φ) estimation. The IMU used is an Xsens MTi-300 which fuses the sensor measurements internally into a global orientation in form of a rotation matrix R_{IMU} . From R_{IMU} we can extract the vector of gravity, which corresponds to our estimated ground plane normal. As long as the unit is exposed only to short translational accelerations as in our case, the gravity referenced IMU readings can be considered free of drift and are precise enough to be used as basis for comparison here. The IMU is externally calibrated to the left camera through a rotation M , which allows to determine the vector of gravity in camera coordinates \mathbf{n}_{IMU} as the third row of MR_{IMU} . We apply (11) to convert it to the spherical representation.

We captured a dataset of 4740 frames at 30 fps along with the IMU orientation for every camera frame. Some example frames are shown in Figure 6.

Filter Prediction: Figure 3 shows the parameter prediction using visual odometry. After initializing the ground plane using RANSAC in frame 3000 we predict the parameters but do not carry out any correction. The parameter drift becomes obvious after a few seconds. The drift in attitude leads to a strong drift in the distance parameter d , which is caused by the straight horizontal translation that is applied to the camera. If this drift is not corrected, the support point selection for least-squares fitting quickly becomes erroneous and eventually the plane track gets lost.

TABLE I: Absolute angular deviation in degrees from IMU ground-truth for different disparity estimators compared to conventional least-squares plane fitting as baseline.

	θ	φ
SGBM ¹ LS baseline	1.25 ± 1.04	2.04 ± 1.21
Filter LS + VP		
SGBM	0.64 ± 0.54	0.53 ± 0.69
libElas ²	0.50 ± 0.40	0.52 ± 0.52
libToast2 [20]	0.82 ± 0.62	0.54 ± 0.65
libToast2 w/o subpix	0.81 ± 0.65	0.53 ± 0.59

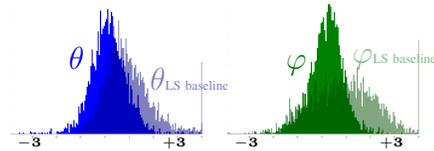


Fig. 4: Distribution of angular error in θ and φ with and without inclusion of vanishing direction. The roll parameter φ benefits most, since objects in lateral direction often disturb the least-squares fit.

Filter Correction: For each filter update we measure the absolute deviation in degrees from the IMU ground truth for both attitude parameters. We evaluate the tracking over the whole sequence of 4740 frames with different disparity estimators, Table I shows the mean and standard deviation. As baseline for comparison we use the mere least-squares plane fit on the predicted plane as filter measurement, this corresponds to conventional ground plane fitting methods as in [10], [11] or [14]. The results after adding the vertical vanishing direction are listed in the box below. Both attitude parameters benefit significantly from the additional measurement, particularly the roll parameter φ becomes more accurate. Figure 4 shows the distributions of parameter deviations, where this also becomes apparent. This effect can be explained by different facts. First, the ground plane is often occluded by cars or buildings in lateral direction which disturb the least-squares fit due to imprecise support point selection. Secondly, the planar surface assumption often does not hold perfectly, e.g., when walking on a slightly elevated pavement, resulting in tilted measurements. See Figure 5 for an example. This section of the dataset lasts around 1000 frames. The absolute errors here are 1.57° for θ and 3.08° for φ with the baseline approach and decrease to 0.73° for θ and

¹<http://opencv.org/>

²<http://www.cvlibs.net/software/libelas/>

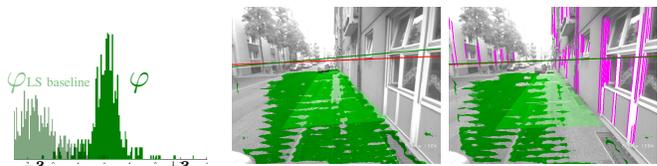


Fig. 5: Effect of planar surface model violation with and without vanishing direction measurement. Ground plane support points are coloured green and the plane is overlaid schematically. The IMU ground-truth virtual horizon is drawn in red, the measured virtual horizon in green. The according error distribution for a sequence of 1000 frames around the depicted scenario is shown on the left.

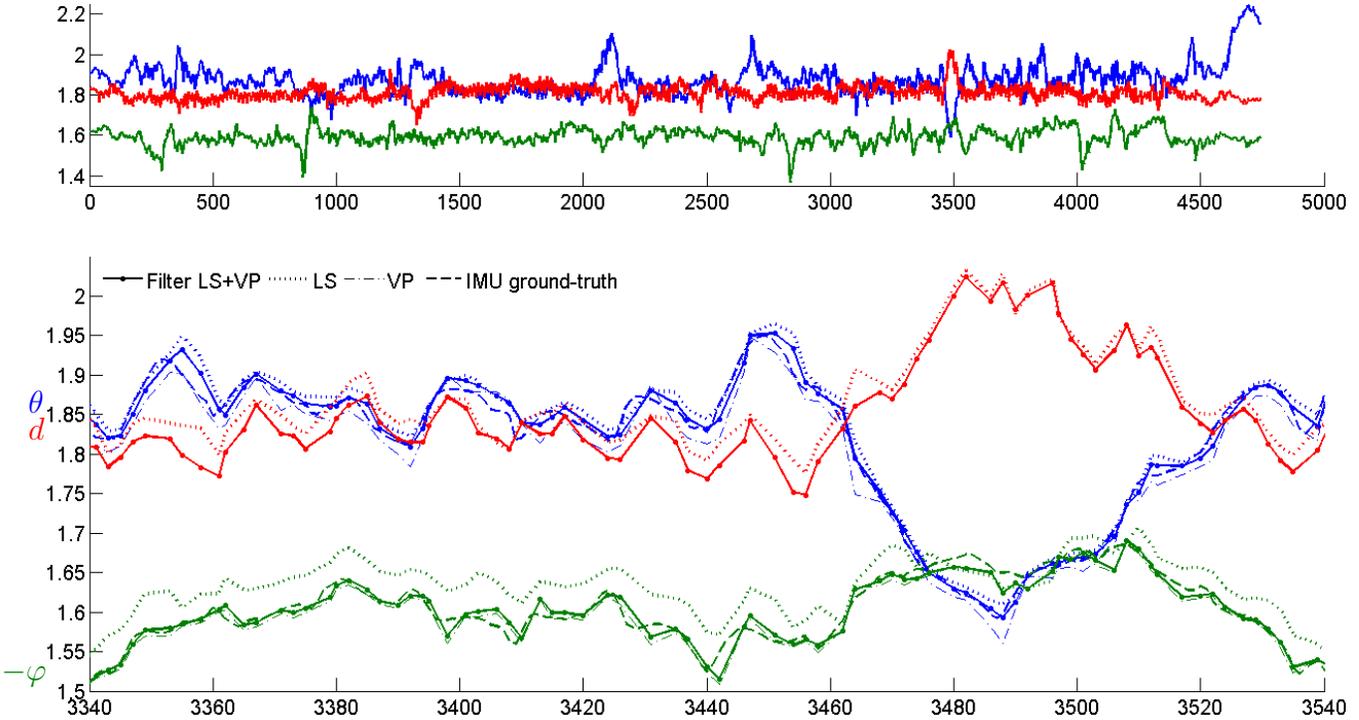


Fig. 6: Ground plane parameters (θ, φ, d) for the evaluation dataset corrected with our filter combining least-squares ground plane fit (LS) and vanishing direction (VP). The top row contains exemplary frames from the sequence.

0.41° for φ after adding the vanishing direction. Finally, the filtered parameters for the whole dataset are plotted in Figure 6.

C. City dataset

To test the long-term stability we ran the ground plane estimation on a dataset consisting of 45 minutes walking through inner-urban scenes, which was mostly captured on narrow side-walks between house facades and cars, but also contains some vast spaces with persons and objects frequently occluding the free view onto the ground. The lighting conditions were challenging with considerably over- and underexposed image areas that lack disparity measurements. Figures 7, 8 and 9 contain some example shots. The ap-

proach enables the system to keep track of the ground plane throughout the whole sequence (except for major violations of the continuous, non-inclined surface assumption, as e.g. on stairways).

The dataset we used for quantitative comparison in Section IV-B does not contain scenes with total occlusion or the ground plane out of view. This is different here and leads to various situations in which the baseline least-squares approach loses track. In the right scene of Figure 8 a passing car blocks the view onto the street for a couple of frames, the disparity clutter causes the plane to drift away. The vanishing direction effectively helps to keep the correct attitude in such cases. Figure 9 demonstrates a sequence with the ground out of view due to bad camera inclination. The drift in prediction

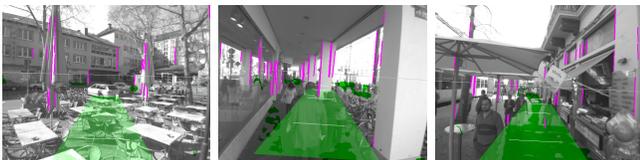


Fig. 7: Example shots from the city dataset.

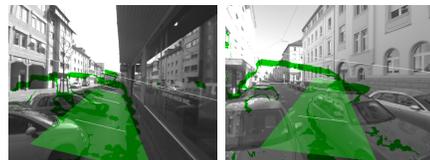


Fig. 8: Typical failure examples without vanishing point correction.



Fig. 9: Sequence with ground plane not in free view due to camera inclination. Only the vanishing direction is considered here to correct the plane parameters.

during this situation is too strong to remeasure the plane successfully afterwards again, when the vanishing direction is not considered. Our filter approach dismisses the erroneous plane fit and uses only the vanishing direction in such case.

V. CONCLUSION

In this work we present an approach to track a planar ground plane from an unconstrained stereo camera platform in inner-urban environments. The proposed framework is able to deal with challenging situations in which existing approaches fail. These situations include strong camera motion and frequent invisibility of the ground due to bad camera orientation or heavy occlusions.

We propose to take visual odometry into the tracking loop in order to obtain a parameter prediction for the current estimate. To correct the prediction drift we combine two complementary measurements: A traditional least-squares plane fit in disparity data when the ground is visible, augmented by the direction of vertical scene structures to provide a measurement of the ground plane normal vector. Note that the direction of vertical scene structures does not depend on the visibility of the ground and can therefore be extracted even if the ground is occluded or out of sight. Moreover, even if the ground is visible and a least-squares fit can be calculated the additional feature of vertical scene structures provides a second, stochastically independent cue for estimation of the ground plane and thus, improves the ground plane estimation accuracy. On a sequence of cluttered inner-urban street scenes we underpin this by comparing our results quantitatively to ground-truth taken from an inertial measurement unit and successfully test the long term stability on a 45 min sequence without losing track of the ground.

Current work focuses on extending the framework to unstructured environments and situations which violate the non-inclined and continuous ground surface assumption, as for instance in case of drop-offs or stairways.

ACKNOWLEDGMENT

The work was supported by the German Federal Ministry of Education and Research within the project OIWOB. The authors would like to thank the "Karlsruhe School of Optics and Photonics" for supporting this work.

REFERENCES

[1] B. Leibe, N. Cornelis, K. Cornelis, and L. Van Gool, "Dynamic 3d scene analysis from a moving vehicle," in *Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on*, June 2007, pp. 1–8.

[2] J. Zhou and B. Li, "Homography-based ground detection for a mobile robot platform using a single camera," in *Robotics and Automation, 2006. ICRA 2006. Proceedings 2006 IEEE International Conference on*, May 2006, pp. 4100–4105.

[3] A. Linarth, M. Brucker, and E. Angelopoulou, "Robust ground plane estimation based on particle filters," in *Intelligent Transportation Systems, 2009. ITSC '09. 12th International IEEE Conference on*, Oct 2009, pp. 1–7.

[4] V. Chari and C. Jawahar, "Multiple plane tracking using unscented kalman filter," in *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*, Oct 2010, pp. 2914–2919.

[5] G. Simon, A. Fitzgibbon, and A. Zisserman, "Markerless tracking using planar structures in the scene," in *Augmented Reality, 2000. (ISAR 2000). Proceedings. IEEE and ACM International Symposium on*, 2000, pp. 120–128.

[6] D. Cobzas, M. Jägersand, and P. Sturm, "3D SSD tracking with estimated 3D planes," *Image and Vision Computing*, vol. 27, no. 1, pp. 69–79, Jan. 2009.

[7] 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, June 2002, pp. 646 – 651 vol.2.

[8] A. Wedel, H. Badino, C. Rabe, H. Loose, U. Franke, and D. Cremers, "B-spline modeling of road surfaces with an application to free-space estimation," *Trans. Intell. Transport. Sys.*, vol. 10, no. 4, pp. 572–583, Dec. 2009.

[9] S. Nedeveschi, R. Danescu, D. Frentiu, T. Marita, F. Oniga, C. Pocol, T. Graf, and R. Schmidt, "High accuracy stereovision approach for obstacle detection on non-planar roads," in *IEEE Intelligent Engineering Systems (INES, 2004)*, pp. 211–216.

[10] S. Se and M. Brady, "Ground plane estimation, error analysis and applications," *Robotics and Autonomous Systems*, vol. 39, no. 2, pp. 59–71, 2002.

[11] N. Chumerin and M. M. Van Hulle, "Ground Plane Estimation Based on Dense Stereo Disparity," in *The Fifth International Conference on Neural Networks and Artificial Intelligence (ICNNAI)*, Minsk, Belarus, May 2008, pp. 209–213.

[12] J. Corso, D. Burschka, and G. Hager, "Direct plane tracking in stereo images for mobile navigation," in *Robotics and Automation, 2003. Proceedings. ICRA '03. IEEE International Conference on*, vol. 1, Sept 2003, pp. 875–880 vol.1.

[13] P. Lombardi, M. Zanin, and S. Messelodi, "Unified stereovision for ground, road, and obstacle detection," in *Intelligent Vehicles Symposium, 2005. Proceedings. IEEE*, June 2005, pp. 783–788.

[14] T.-S. Leung and G. Medioni, "Visual navigation aid for the blind in dynamic environments," in *Computer Vision and Pattern Recognition Workshops (CVPRW), 2014 IEEE Conference on*, June 2014, pp. 579–586.

[15] J. C. Bazin, P. Laffont, I. Kweon, C. Demonceaux, and P. Vasseur, "An original approach for automatic plane extraction by omnidirectional vision," in *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems, October 18-22, 2010, Taipei, Taiwan, 2010*, pp. 752–758.

[16] D. Scaramuzza and F. Fraundorfer, "Visual odometry [tutorial]," *Robotics Automation Magazine, IEEE*, vol. 18, no. 4, pp. 80–92, Dec 2011.

[17] A. Geiger, J. Ziegler, and C. Stillner, "Stereoscan: Dense 3d reconstruction in real-time," in *IEEE Intelligent Vehicles Symposium*, Baden-Baden, Germany, June 2011.

[18] J.-P. Tardif, "Non-iterative approach for fast and accurate vanishing point detection," in *Computer Vision, 2009 IEEE 12th International Conference on*, Sept 2009, pp. 1250–1257.

[19] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? the kitti vision benchmark suite," in *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012.

[20] B. Ranft and T. Strauß, "Modeling arbitrarily oriented slanted planes for efficient stereo vision based on block matching," in *Intelligent Transportation Systems, 17th International IEEE Conference on*, Qingdao, China, Oct 2014.