

Accuracy Analysis of Surface Normal Reconstruction in Stereo Vision

Hannes Harms¹, Johannes Beck¹, Julius Ziegler² and Christoph Stiller¹

Abstract—Estimating surface normals is an important task in computer vision, e.g. in surface reconstruction, registration and object detection. In stereo vision, the error of depth reconstruction increases quadratically with distance. This makes estimation of surface normals an especially demanding task. In this paper, we analyze how error propagates from noisy disparity data to the orientation of the estimated surface normal. Firstly, we derive a transformation for normals between disparity space and world coordinates. Afterwards, the propagation of disparity noise is analyzed by means of a Monte Carlo method. Normal reconstruction at a pixel position requires to consider a certain neighborhood of the pixel. The extent of this neighborhood affects the reconstruction error. Our method allows to determine the optimal neighborhood size required to achieve a pre specified deviation of the angular reconstruction error, defined by a confidence interval. We show that the reconstruction error only depends on the distance of the surface point to the camera, the pixel distance to the principal point in the image plane and the angle at which the viewing ray intersects the surface.

I. INTRODUCTION

Surface normal estimation from range images [1] is a well known problem in surface reconstruction [2] [3] [4], registration [5] and object detection [6]. It forms a basis for environment perception in driver assistance systems, e.g. in the identification of drivable space and objects. Stereo vision offers a challenging field for normal estimation, since the reconstruction error grows quadratically with distance.

Previous work in stereo vision includes linear error propagation by means of first order Taylor expansion [7] [8]. The influence of sensor parameters on stereo quantization errors is surveyed in [9]. An approach for recovering 3d surface orientation and discontinuities from stereo disparity was presented in [10]. [11] analyzed error propagation for the estimation of line and planar surface orientations, comparing methods based on line features and absolute correspondence of points. To our best knowledge, there exists no work that analyzes the effects of disparity noise on the orientation of surface normals, so that an appropriate patch size at a surface point can be chosen to guarantee a certain orientation accuracy of the estimated normal.

We provide an analysis of the expected angular error distribution by error propagation. The main contribution of this paper is the evaluation of the expected angular error distributions regarding the image position, the distance

¹Hannes Harms, Johannes Beck and Christoph Stiller are with Institute of Measurement and Control Systems, Karlsruhe Institute of Technology (KIT), 76131 Karlsruhe, Germany {harms, johannes.beck, stiller}@kit.edu

²Julius Ziegler is with FZI Research Center for Information Technology, Mobile Perception Systems, 76131 Karlsruhe, Germany ziegler@fzi.de

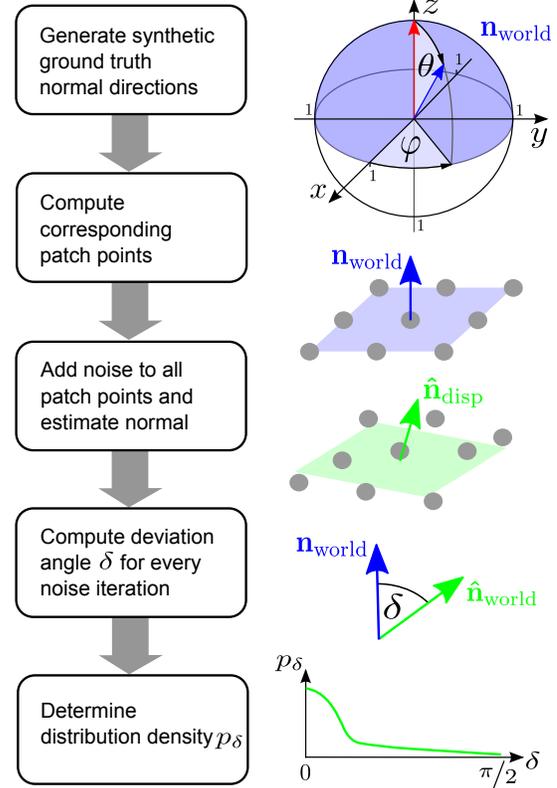


Fig. 1: Main steps of the Monte Carlo based error propagation for all analyzed surface points.

and the orientation of the normal. Moreover, we present a transformation of normals between disparity space and world coordinates. Regarding real world applications, we present a method that allows to determine an optimal patch size to estimate the surface normal for a given angular deviation, defined by a confidence interval, at a certain distance.

This paper is structured as follows: Section II introduces a transformation for surface normals between disparity space and world coordinates. Section III provides details about the selected Monte Carlo method based error propagation. Three simulation scenarios are proposed, that are compared during error propagation. Section IV evaluates experiments in 2d and 3d of these scenarios and identifies parameters for the final application. In Section V a general solution is presented to determine an optimal patch size for a surface normal by taking observations of the error propagation into account. We close this paper with conclusions and an outlook.

II. NORMAL TRANSFORMATION

Dense stereo matching for a rectified stereo camera produces a disparity image, assigning one disparity d to every pixel position (u, v) of the (left) input image. Hence, the surface is firstly represented as a set of samples $\mathbf{p} = (u, v, d)^\top$ in *disparity space*. For each of these samples, the 3d position $\mathbf{x} = (x, y, z)^\top$ in the camera coordinate system (*world space*) can be reconstructed by the function $r: \mathbb{R}^3 \rightarrow \mathbb{R}^3$,

$$\mathbf{x} = r(\mathbf{p}) = \frac{b}{d} \begin{pmatrix} u - c_u \\ v - c_v \\ f \end{pmatrix}, \quad (1)$$

where the focal length f and the principal point position $(c_u, c_v)^\top$ are the calibrated pinhole parameters of the identical single cameras, while b is the baseline length of the stereo system.

An important, well known property of r is that it is plane preserving, *i.e.* the four world points $r(\mathbf{p}_0)$, $r(\mathbf{p}_1)$, $r(\mathbf{p}_2)$ and $r(\mathbf{p}_3)$ are coplanar, if and only if \mathbf{p}_0 , \mathbf{p}_1 , \mathbf{p}_2 and \mathbf{p}_3 are coplanar in disparity space [7]. Because of this property, we can at first compute a surface normal in disparity space, and subsequently transform it into world space. This is beneficial for two reasons: Firstly, for normal computation, it is necessary to establish a set of neighborhood samples around every surface point. These samples are then used to locally fit a plane. The normal vector of the local plane approximates the normal vector of the surface. In disparity space, a neighborhood can be easily chosen by picking adjacent pixels in the disparity image. Secondly, a better approximation of the surface normal is obtained if the local plane fit is performed in disparity space instead of world space. As described in [12], disparity measurements can approximately be modeled as normally distributed. The transformation from disparity space to world coordinates is non-linear. That results in a distribution of the noisy points in world space which is not Gaussian.

Let the surface point $\mathbf{p}_0 = (u_0, v_0, d_0)^\top$ and the matching normal vector $\mathbf{n}_{\text{disp}} = (n_u, n_v, n_d)^\top$ represent a plane in disparity space. We now derive the normal vector $\mathbf{n}_{\text{world}} = (n_x, n_y, n_z)^\top$ at the reconstructed surface point $\mathbf{x}_0 = r(\mathbf{p}_0)$ in world space. We pick the two auxiliary points

$$\mathbf{p}_1 = \mathbf{p}_0 + \begin{pmatrix} -n_v \\ n_u \\ 0 \end{pmatrix} \quad (2)$$

and

$$\mathbf{p}_2 = \mathbf{p}_0 + \begin{pmatrix} 0 \\ n_d \\ -n_v \end{pmatrix}. \quad (3)$$

It can be easily seen that \mathbf{p}_0 , \mathbf{p}_1 and \mathbf{p}_2 are within a local plane in disparity space which interpolates \mathbf{p}_0 and has the normal vector \mathbf{n}_{disp} . We now transform \mathbf{p}_0 , \mathbf{p}_1 and \mathbf{p}_2 through r to yield the three world points

$$\mathbf{x}_i = r(\mathbf{p}_i), i \in \{0, 1, 2\}. \quad (4)$$

Since r is plane preserving, \mathbf{x}_0 , \mathbf{x}_1 and \mathbf{x}_2 are now on the plane in world space that maps all points of the plane

represented in disparity space by \mathbf{p}_0 and \mathbf{n}_{disp} . A normal vector in world space can now be yielded by the cross product

$$\mathbf{n}_{\text{world}} = (\mathbf{x}_1 - \mathbf{x}_0) \times (\mathbf{x}_2 - \mathbf{x}_0). \quad (5)$$

For later reference, we denote the function that transforms an arbitrary surface point $\mathbf{p} = (u, v, d)^\top$ and matching normal vector $\mathbf{n}_{\text{disp}} = (n_u, n_v, n_d)^\top$ in disparity space into a normal vector $\mathbf{n}_{\text{world}}$ in world space as $f: \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^3$. Inserting (1) - (4) into (5), we yield

$$\mathbf{n}_{\text{world}} = f(\mathbf{p}, \mathbf{n}_{\text{disp}}) = \begin{pmatrix} fn_u \\ fn_v \\ (c_u - u)n_u + (c_v - v)n_v - dn_d \end{pmatrix}. \quad (6)$$

Inversely, an arbitrary surface point $\mathbf{x} = (x, y, z)^\top$ and matching normal vector $\mathbf{n}_{\text{world}} = (n_x, n_y, n_z)^\top$ in world space can be transformed back into disparity space by the function g ,

$$\mathbf{n}_{\text{disp}} = g(\mathbf{x}, \mathbf{n}_{\text{world}}) = \begin{pmatrix} bn_x \\ bn_y \\ -xn_x - yn_y - zn_z \end{pmatrix}. \quad (7)$$

III. ERROR PROPAGATION

We use the Monte Carlo method [13] to propagate disparity noise through the normal reconstruction process. This allows us to observe effects that cannot be detected by linear error propagation.

We will now briefly explain how the Monte Carlo method is used for error propagation in general. Let the random variable X with the distribution density p_X represent some noisy observation. We are interested in the distribution that is obtained when propagating the observation through an arbitrary function h , *i.e.* in the distribution of the random variable $Y = h(X)$. We draw a set of N samples $S_X = \{x_0, \dots, x_{N-1}\}$ from the distribution of X , $x_k \sim p_X$. By propagating the samples through h , we obtain a sample set $S_Y = \{y_0, \dots, y_{N-1}\}$ with $y_k = h(x_k)$. The set S_Y approximates the distribution of Y . Hence, statistic properties, like mean μ_Y and standard deviation σ_Y , can be derived from it.

In our application, the function under consideration is the complete normal reconstruction process. It is composed by a stage that computes a normal vector in disparity space, based on a given set of neighborhood pixels in the disparity image, and the function for transforming the normal into world space, f from equation (6).

In this paper, ground truth surface points and normals are generated for different image positions and depths. Each surface point is associated with a set of preset normal directions. Every ground truth normal $\mathbf{n}_{\text{world}}$ is assigned to a ground truth patch of points. These points are transformed to disparity space. Disparity noise is added to all ground truth patch points in a noise step, which we apply N times. After each noisy iteration, we estimate the surface normal $\hat{\mathbf{n}}_{\text{disp}}$ of the noisy points, transform it to $\hat{\mathbf{n}}_{\text{world}}$ and compute the deviation angle

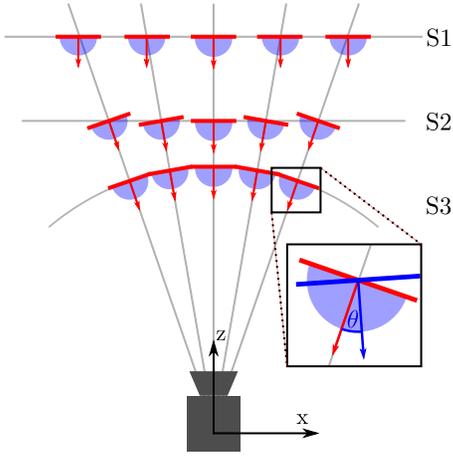


Fig. 2: Selection of simulation scenarios to generate synthetic ground truth surface normals $\mathbf{n}_{\text{world}}$ and corresponding patches.

$$\delta = \arccos \frac{\mathbf{n}_{\text{world}} \cdot \hat{\mathbf{n}}_{\text{world}}}{\|\mathbf{n}_{\text{world}}\| \|\hat{\mathbf{n}}_{\text{world}}\|} \quad (8)$$

to the ground truth normal in world coordinates. Finally, all deviation angles of one patch are represented by the distribution density p_δ , see Fig. 1.

For every surface point that is evaluated, multiple ground truth normal vectors with different orientations are generated, which are sampled around a central normal, that is highlighted in red in Fig. 1. In the 2d case, we will denote the angle difference towards the central normal as θ . In 3d, this will be complemented by a second angle, φ .

The noisy observations are the disparities measured at every pixel in the image. Note that the noise in disparity data can be assumed to be stationary over the disparity image, and that, in particular, it does not depend on the distance of the surface point from the camera. We assume that disparity measurements are normally distributed with a standard derivation of σ_d . Image coordinates u, v of every pixel are not subject to any noisy measurement process, and can be assumed to be known exactly.

The set of surface points and matching normals to be considered in the analysis are based on three different scenarios. The scenarios are illustrated schematically in Fig. 2. The central patches with corresponding normals are highlighted in red, and the sampling range of the additional normal directions is indicated in blue.

Scenario 1 (S1) compares planar patches that are located at identical depth and which have the central normal vector oriented along the camera z-axis.

Scenario 2 (S2) compares planar patches that are located at identical depth with the central normal vector oriented along the cameras viewing rays.

Scenario 3 (S3) compares planar patches that are located at identical distance from the camera with the central normal oriented along the cameras viewing rays.

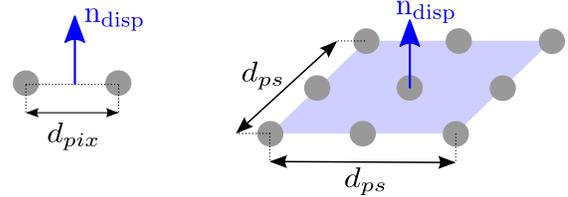


Fig. 3: Surface patches and corresponding normals \mathbf{n}_{disp} in 2d (left) and 3d (right) used in this paper.

These scenarios are designed to expose the dependency of the angular reconstruction error on specific geometric conditions, as will be demonstrated in the next section.

IV. PARAMETER IDENTIFICATION THROUGH SIMULATIONS

In this section the characteristic effects during error propagation are searched, to identify parameters for our goal, the determination of an optimal (i.e. minimal) neighborhood patch size to estimate a surface normal with a preset angular accuracy, which is later defined by a 95% confidence interval.

We expect the following effects observing the distribution density p_δ , which we state by three hypotheses:

- *H1*: σ_δ increases with increasing depth due to depth noise characteristics.
- *H2*: σ_δ decreases the more orthogonal a normal is w.r.t. the camera z-axis, because the normal direction is less influenced by the disparity noise.
- *H3*: σ_δ decreases the farther apart the points for normal estimation are, because the noise influence diminishes.

The first parameter chosen is the distance of the surface point, since the error of depth reconstruction increases quadratically with distance and must be considered (*H1*). To identify other parameters, all experiments are executed at a constant distance of 10 meters, as defined in Sec. III for each scenario.

In the beginning, further parameters are searched by analyzing the distribution density p_δ for different surface points and normal directions by the mean μ_δ and the standard deviation σ_δ . Later, the confidence interval width γ_δ is used to interpret p_δ . γ_δ represents the maximum deviation angle of the 95% confidence interval of p_δ , so that 95% of all estimated normals $\hat{\mathbf{n}}_{\text{world}}$ have an angle deviation $|\delta|$ smaller than γ_δ .

We divide our experiments in a part observing relevant effects in 2d before showing results in 3d. Note that the angle deviation δ is limited from -90° up to 90° in 2d, because the sign of the deviation can be taken into account. In 3d, δ is the dihedral angle and hence, is limited from 0° up to 90° . Dense stereo matching is lately moving towards sub pixel accurate disparity maps [14] [15] [16]. We model uncertainties of the disparity as a bias free normal distribution with $\sigma_d = 0.1$ pixels. Furthermore, realistic camera parameters are used from the KITTI data set [17] with $f \approx 722$ pixels, $c_u \approx 609$ pixels, $c_v \approx 173$ pixels and $b \approx 0.54$ meters. The image resolution considered is 1217×345 pixels.

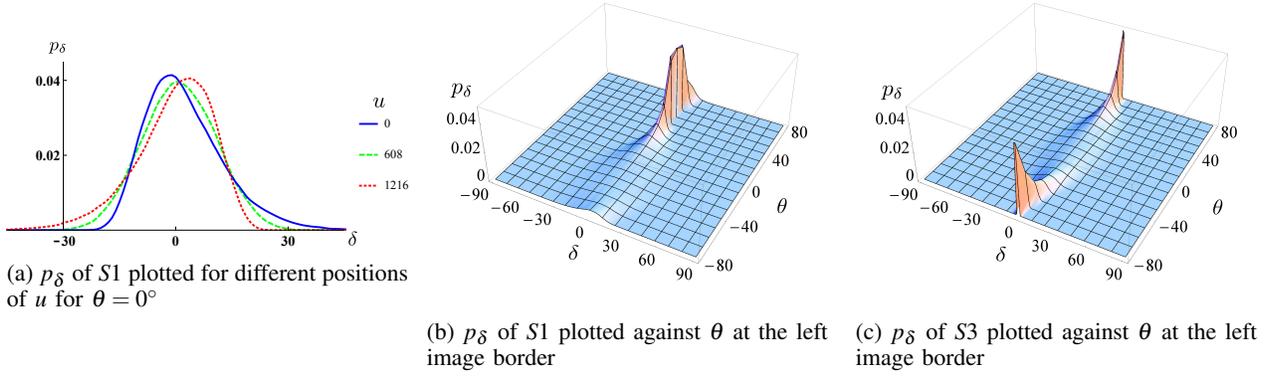


Fig. 4: p_δ plotted against u and θ

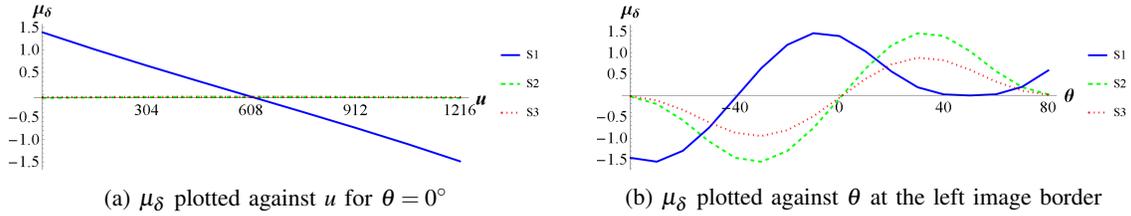


Fig. 5: μ_δ plotted against u and θ

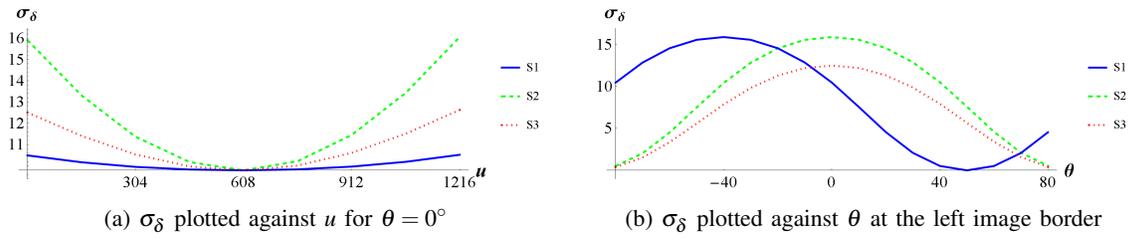


Fig. 6: σ_δ plotted against u and θ

A. Accuracy analysis in 2d

Firstly, we observe general effects for δ by setting the world coordinate $y = 0$ and observe two points on the camera xz -plane. The normal \mathbf{n}_{disp} of the two points $\mathbf{p}_1 = (u_1, c_v, d_1)^\top$ and $\mathbf{p}_2 = (u_2, c_v, d_2)^\top$ can be computed directly to:

$$\mathbf{n}_{\text{disp}} = \begin{pmatrix} d_1 - d_2 \\ 0 \\ u_2 - u_1 \end{pmatrix} \quad (9)$$

and transformed to $\mathbf{n}_{\text{world}}$ by equation (6), considering the mean point of \mathbf{p}_1 and \mathbf{p}_2 .

The following evaluations to compare S1, S2 and S3 are executed for a constant pixel distance $d_{\text{pix}} = 15$ pixels between the observed point pair (drafted in fig. 3). We add disparity noise $N = 10^6$ times separately on both ground truth points in disparity space, determine $\hat{\mathbf{n}}_{\text{world}}$ after each noise iteration and observe the distribution density p_δ by varying:

- the u -coordinate of the observed surface point over the image width from left to right border of the image and
- the direction of the surface normal by the angle θ from -80° to 80° of every surface point.

At first, we analyze the effects on p_δ that appear when the ground truth patch is shifted from the left to the right

image border. Figure 4a shows p_δ for S1 plotted for different image coordinates u for the normal direction $\theta = 0^\circ$. It can be seen, that p_δ does not match a Gaussian distribution at the image borders. In Fig. 5a, the mean μ_δ of S1 is symmetric w.r.t. the principal point and shows an approximately linear behavior over the image width. Whereas μ_δ of S2 and S3 are approximately zero over the image width. When observing the standard deviation σ_δ in Fig. 6a, all scenarios are axially symmetrical and increase with increasing pixel distance w.r.t. the principal point.

Secondly, we evaluate p_δ when varying θ (i.e. the direction) of the ground truth normal. To point out the differences of the compared scenarios, p_δ is plotted against θ at the left image border. Fig. 4b and Fig. 4c show p_δ for S1 and S3. It is obvious that p_δ of S3 is symmetric w.r.t. normal direction $\theta = 0^\circ$, while p_δ of S1 is not symmetric. μ_δ of S2 and S3 show point symmetry w.r.t. the direction $\theta = 0^\circ$ of the central normal in contrary to S1, as can be seen in Fig. 5b. S2 and S3 are axially symmetrical w.r.t. the direction $\theta = 0^\circ$ of the central normal in σ_δ and decrease the larger θ grows. No symmetric characteristics can be found for σ_δ of S1 in Fig. 6b.

Furthermore we diagnose the correlation of the image

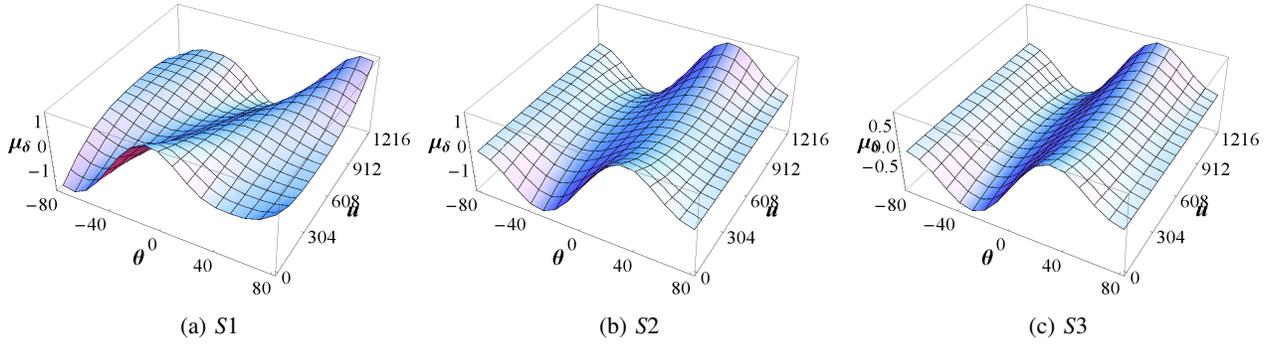


Fig. 7: μ_δ plotted against u and θ

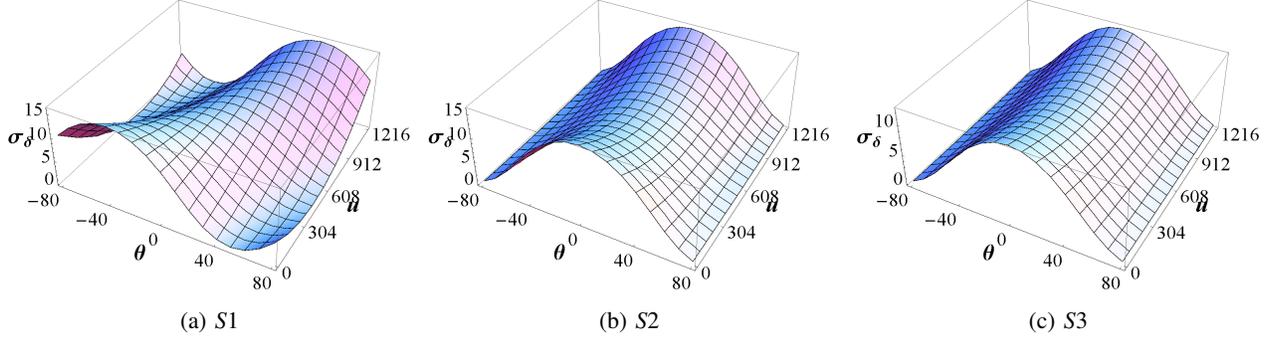


Fig. 8: σ_δ plotted against u and θ

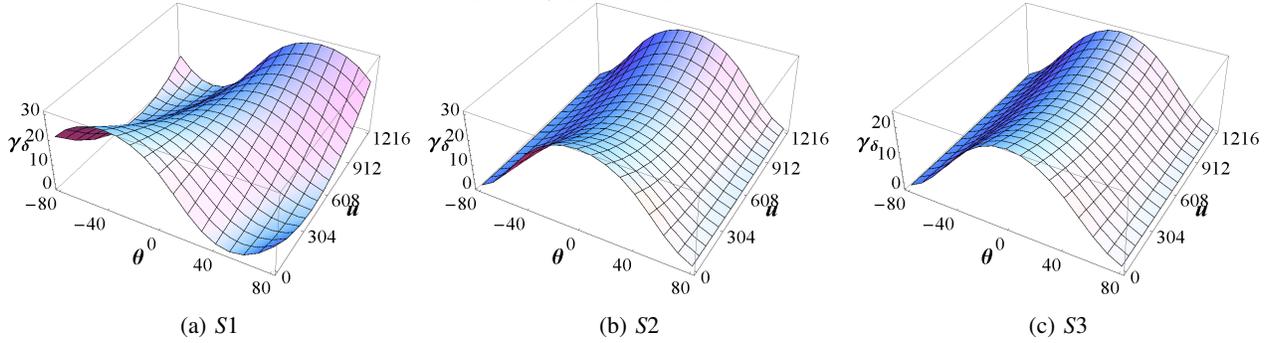


Fig. 9: γ_δ plotted against u and θ

position u and the angle θ , with regard of the distributions of μ_δ and σ_δ . Fig. 7 and 8 show the results for all three scenarios. While symmetry effects are not observable for S1, we found symmetric distributions for μ_δ and σ_δ w.r.t. the principal point for both scenarios S2 and S3. Regarding S2 and S3, a general characteristic is, that μ_δ and σ_δ increase from the principal point to the image borders.

S2 and S3 align the central normal of the observed patch in direction of the camera's viewing ray. That results in symmetric distributions of μ_δ and σ_δ when observing the presented 2d effects. We observed the discovered symmetry effects as well, when d_{pix} and the distance to the observed surface point are varied. The difference between S2 and S3 is caused by the smaller influence of depth noise for S3, due to d_{Euc} . That results in smaller μ_δ and σ_δ compared to S2.

Since p_δ is biased and not a Gaussian distribution, σ_δ does not correspond to the angle deviation of the ground truth normal. To get a characteristic value for this angle deviation, we evaluate p_δ by the angle deviation γ_δ , as introduced at

the beginning of this section. Fig. 9 shows the distribution of γ_δ plotted against θ and the image width u . Maximum values of γ_δ can be found for ground truth normals oriented along the camera's viewing ray (i.e. for $\theta = 0^\circ$ for S2 and S3). Further experiments have confirmed that observation. This means that the disparity depth noise has the biggest effect on the accuracy, if the surface patch is aligned orthogonal to the camera's viewing ray and not, as we expected in H2, if the surface patch is aligned orthogonal to the camera's z-axis. Fig. 9b and Fig. 9c show as well, that the expected orientation error decreases the larger $|\theta|$ is chosen.

The shown effects on the xy -plane can be transferred to the yz -plane, since the reconstruction of u and v from disparity space to x and y in world coordinates is identical, as can be seen in equation (1). In the following, the symmetric distributions of γ_δ in 2d w.r.t. the principal point on the image plane and w.r.t. the normal direction $\theta = 0^\circ$ are evaluated closer in 3d, to identify d_{pp} as another parameter.

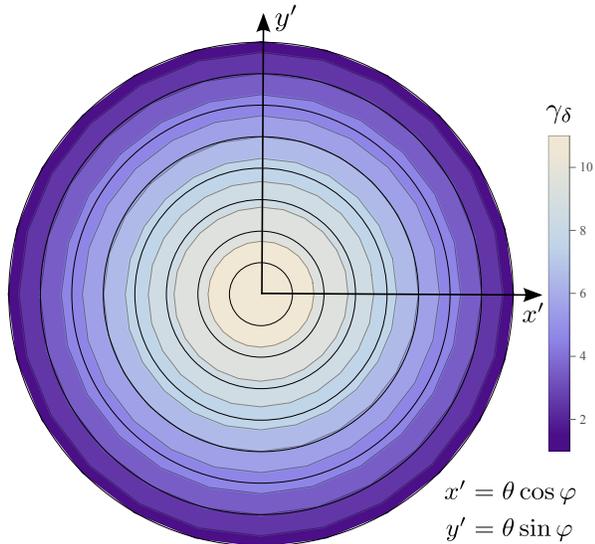


Fig. 10: Polar coordinate plot of the distribution of γ_δ for all synthetic generated ground truth normal orientations using S3. A 3d surface point with the distance $d_{Euc} = 10$ meters at image position $\mathbf{p}_E = (732, 51)$ pixels (see Fig. 11) is chosen. The viewing direction is aligned with the center normal. The black circles represent steps for θ of 10° from the inner circle ($\theta = 10^\circ$) to the outer circle ($\theta = 80^\circ$).

B. Accuracy analysis in 3d

The accuracy analysis in 3d is observed for ground truth patches of nine points. The patch size is defined as quadratic on the image plane with side length d_{ps} in pixels, see Fig. 3. The simulation in 3d uses normals generated by spherical coordinates w.r.t. central normal with steps for θ of 10° and for φ of 15° , as described in Sec. III. The nine points of every patch are transformed into disparity space, noise is added to the disparity of all points before the surface normal is estimated in disparity space and reconstructed in world coordinates, see Fig. 1. To analyze further symmetry effects for γ_δ in 3d, all experiments are based on S3.

We use a principal components analysis (PCA) in disparity space to determine $\hat{\mathbf{n}}_{disp}$. The eigenvector with the smallest eigenvalue is taken as the surface normal. The mean point of a patch in disparity space is used additional to $\hat{\mathbf{n}}_{disp}$ to determine $\hat{\mathbf{n}}_{world}$ as introduced in Sec. II.

The characteristics of the distribution of γ_δ are evaluated for different orientations of the surface normals in 3d. Fig. 10 shows a polar coordinate plot of the distribution of γ_δ for all sampled normal directions of the hemisphere referring to S3. The plot evaluates the image position $\mathbf{p}_E = (732, 51)$ as can be seen in Fig. 11 in a distance $d_{Euc} = 10$ meters. It can be seen that the maximum value of γ_δ can be found for $\theta = 0^\circ$. Moreover, we evaluated a set of image positions, which all showed the maximum value for γ_δ if the normal orientation is aligned to the camera's viewing ray, i.e. $\theta = 0^\circ$ for S3. Since the direction of the surface normal is not known before the estimation of $\hat{\mathbf{n}}_{world}$, we consider the normal direction that is least accurate in the following (i.e. $\theta = 0^\circ$ for S3).

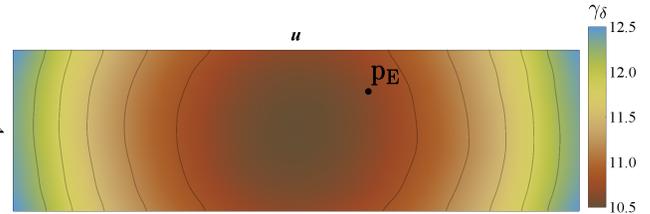


Fig. 11: γ_δ ($\theta = 0^\circ$) of S3 plotted on the image plane with constant Euclidian distance $d_{Euc} = 10$ meters. The position \mathbf{p}_E is evaluated in Fig. 10.

The distribution of γ_δ is evaluated for different image positions (u, v) , to prove the symmetric behavior of γ_δ w.r.t. the distance d_{pp} to the principal point, found during the 2d experiments. Fig. 11 shows that the distribution of γ_δ in the image plane is indeed point symmetrical to the principal point. Therefore we chose d_{pp} as the second parameter to determine the optimal patch size of a surface normal.

The two parameters identified are the distance to the surface point d_{Euc} in meters and the distance to the principal point d_{pp} in pixels. In practice, these two parameters can be determined directly by the information that is given by the disparity map.

V. DETERMINING THE OPTIMAL PATCH SIZE

The previous section illustrated which parameters can be used to generalize the distribution densities p_δ for normals with varying directions and surface points. In the following we use S3 and focus on the behavior of the Euclidean distance d_{Euc} to the surface point in world coordinates, the distance d_{pp} in pixels to the principal point and the patch size d_{ps} when choosing a constant γ_δ .

This interrelationship is shown for $\gamma_\delta = 10^\circ$ in Fig. 12a and 12b, when considering nine patch points to estimate $\hat{\mathbf{n}}_{world}$. We compare this method by considering all points for the estimation of $\hat{\mathbf{n}}_{world}$ that lie in between the current patch boundaries. The result is shown in Fig. 12c and 12d for $\gamma_\delta = 1^\circ$. Normals can be estimated more accurate when considering all patch points and the larger the patch size is chosen (as expected by H3). Note the scale of the d_{ps} axis between Fig. 12a and Fig. 12c.

Note that Fig. 12 is generated from points that are obtained by analyzing the distribution of γ_δ for $\theta = 0^\circ$ (see Fig. 9c) for different d_{Euc} and d_{ps} . Due to interpolation effects, small discontinuities can appear in Fig. 12.

Considering d_{Euc} and d_{pp} , one can determine d_{ps} to fulfill the predefined angle deviation γ_δ . For a constant d_{pp} , d_{ps} increases with increasing distance d_{Euc} . As already observed in Sec. IV, d_{ps} also increases with increasing d_{pp} at constant distance d_{Euc} .

VI. CONCLUSION AND OUTLOOK

In this contribution, we introduced a transformation of surface normals between disparity space and world coordinates. We presented a Monte Carlo method based error propagation for the estimation of surface normals by considering

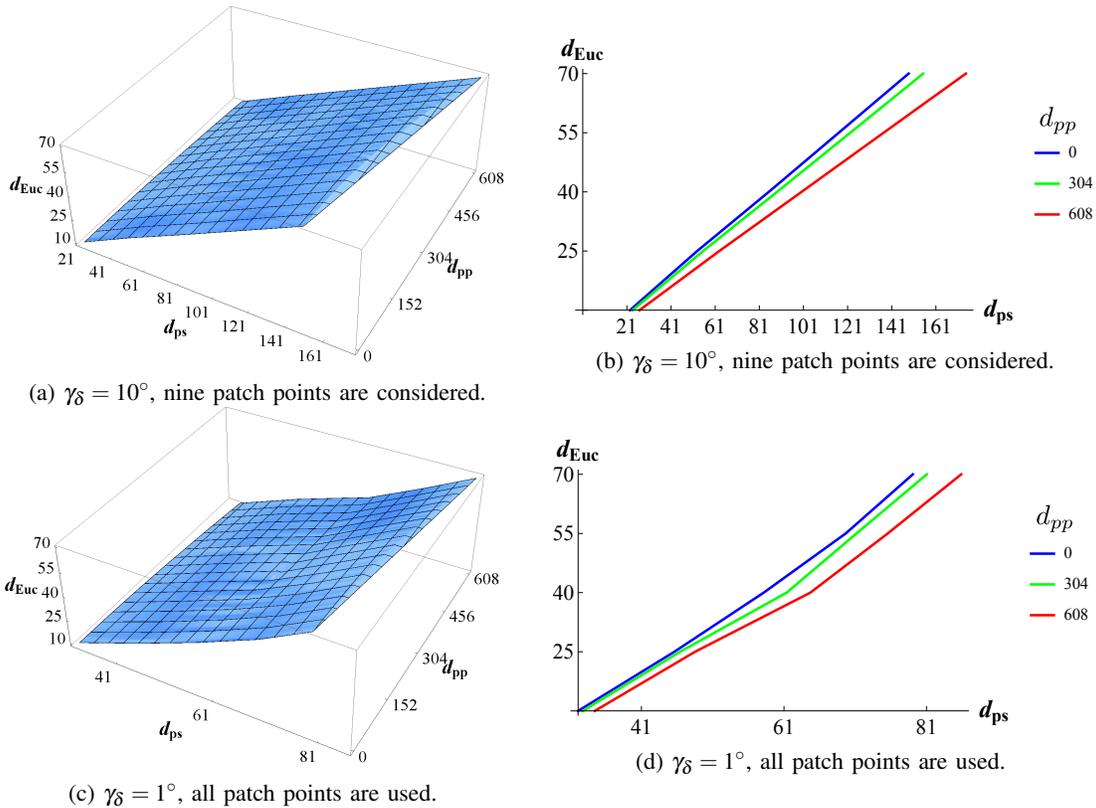


Fig. 12: Relation between the distance d_{Euc} in meters to the 3d surface point and the distance to the principal point d_{pp} in pixels to determine a patch size d_{ps} . (a) and (b) show the relation for $\gamma_\delta = 10^\circ$ considering nine patch points to estimate the normal $\hat{\mathbf{n}}_{world}$. (c) and (d) show the relation for $\gamma_\delta = 1^\circ$ using all patch points to estimate the normal $\hat{\mathbf{n}}_{world}$.

disparity depth noise. A method was proposed that allows to determine an optimal patch size for normal estimation for a pre specified angular deviation of the angular reconstruction error. We showed that a suitable value for the patch size can be determined by considering only two quantities: the pixel distance to the principal point and the Euclidean distance to the 3d surface point.

We think that further work can extend the proposed analysis by considering noise effects for calibration parameters. Furthermore, the comparison with other error propagation methods can be evaluated, e.g. Taylor series and Unscented transform. We will focus the future research on environment perception applications, considering the knowledge of accurate normal estimation we gathered in this work.

REFERENCES

- [1] H. Badino, D. Huber, Y. Park, and T. Kanade, "Fast and accurate computation of surface normals from range images," in *Robotics and Automation (ICRA)*, 2011. IEEE, 2011, pp. 3084–3091.
- [2] W. E. Lorensen and H. E. Cline, "Marching cubes: A high resolution 3d surface construction algorithm," in *ACM Siggraph Computer Graphics*, vol. 21, no. 4. ACM, 1987, pp. 163–169.
- [3] M. Kazhdan, M. Bolitho, and H. Hoppe, "Poisson surface reconstruction," in *Proceedings of the fourth Eurographics symposium on Geometry processing*, 2006.
- [4] N. Amenta, M. Bern, and M. Kamvysselis, "A new voronoi-based surface reconstruction algorithm," in *Proceedings of the 25th annual conference on Computer graphics and interactive techniques*. ACM, 1998, pp. 415–421.
- [5] J. Salvi, C. Matabosch, D. Fofi, and J. Forest, "A review of recent range image registration methods with accuracy evaluation," *Image and Vision Computing*, vol. 25, no. 5, pp. 578–596, 2007.
- [6] A. E. Johnson and M. Hebert, "Using spin images for efficient object recognition in cluttered 3d scenes," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 21, no. 5, pp. 433–449, 1999.
- [7] O. Faugeras, *Three dimensional computer vision: A geometric viewpoint*. the MIT Press, 1993.
- [8] D. Murray and J. J. Little, "Environment modeling with stereo vision," in *Intelligent Robots and Systems (IROS). Proceedings.*, vol. 3. IEEE, 2004, pp. 3116–3122.
- [9] D. F. Llorca, M. A. Sotelo, I. Parra, M. Ocaña, and L. M. Bergasa, "Error analysis in a stereo vision-based pedestrian detection sensor for collision avoidance applications," *Sensors*, vol. 10, no. 4, pp. 3741–3758, 2010.
- [10] R. P. Wildes, "Direct recovery of three-dimensional scene geometry from binocular stereo disparity," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 8, pp. 761–774, 1991.
- [11] L. B. Wolff and T. E. Boult, "Using line correspondence stereo to measure surface orientation," in *IJCAI*, 1989, pp. 1655–1660.
- [12] G. Sibley, L. Matthies, and G. Sukhatme, *Bias reduction and filter convergence for long range stereo*. Springer, 2007.
- [13] S. Geman and D. Geman, "Stochastic relaxation, gibbs distributions, and the bayesian restoration of images," *Pattern Analysis and Machine Intelligence*, no. 6, pp. 721–741, 1984.
- [14] A. Geiger, M. Roser, and R. Urtasun, "Efficient large-scale stereo matching," in *Asian Conference on Computer Vision (ACCV)*, 2010.
- [15] H. Hirschmuller, "Accurate and efficient stereo processing by semi-global matching and mutual information," in *Computer Vision and Pattern Recognition (CVPR)*, vol. 2. IEEE, 2005, pp. 807–814.
- [16] K. Yamaguchi, D. McAllester, and R. Urtasun, "Robust monocular epipolar flow estimation."
- [17] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, "Vision meets robotics: The kitti dataset," *International Journal of Robotics Research (IJRR)*, 2013.