Representing Feature Location Uncertainties in Spherical Images

Bernd Krolla
DFKI - German Research Center for Artificial Intelligence
bernd.krolla@dfki.de

Gabriele Bleser
DFKI - German Research Center for Artificial Intelligence
gabriele.bleser@dfki.de

Yan Cui
DFKI - German Research Center for Artificial Intelligence
yan.cui@dfki.de

Didier Stricker
DFKI - German Research Center for Artificial Intelligence
didier.stricker@dfki.de

ABSTRACT

Pose uncertainty estimation of calibrated cameras is a common task in the field of computer vision and uses location uncertainties of image features. For spherical cameras, those uncertainties cannot be optimally described using conventional latitude-longitude representation. Increasing distortions close to the poles of the spherical coordinate system prevent a suitable description through Gaussians.
To overcome this limitation, we present a consistent location uncertainty representation for spherical image features: Our approach is based on normal vectors in Cartesian space and applicable to any kind of camera with convex projection surfaces, such as catadioptric and spherical systems. We compare its performance against latitude-longitude representation by estimating camera pose uncertainties through first order error propagation in a weighted least squares pose estimation scenario. Our experiments on synthetic and real data show that the proposed approach delivers consistent results outperforming conventional latitude-longitude representation.

Keywords

Spherical imaging, feature extraction, uncertainty visualization, camera pose optimization

1 INTRODUCTION

Spherical imaging experienced increasing attention in the recent past: Microsofts Streetside project as well as Googles Street View project [ADF+10] exemplify the usability of spherical images for large scale applications. The availability of omnidirectional cameras like SpheronVRs SceneCam [VR], Weiss AGs Civetta [Wei] or the Ladybug camera from Point Grey Research also corroborate the interest on spherical imaging.

Since those cameras provide the largest possible field of view (FOV), their usage avoids problems, known from perspective imaging: Visual SLAM can be performed without loosing image features caused by camera rotation in combination with a limited FOV as studied by Gutierrez et al. [GRMG11]. Furthermore provide omnidirectional cameras due to the extended FOV more extractable features for spherical Structure from Motion (SfM) algorithms as proposed by Pagani et al. [PS11] and Torii et al. [TIO05].

The propagation of image feature uncertainties can augment spherical SfM and SLAM algorithms and improves for example outlier detection or surface fitting in a subsequent reconstruction process.

Since feature uncertainty description in perspective images can be handled based on Gaussian error approximation in the Cartesian coordinate system [BBS07, ZGS+09], this approach is not applicable for spherical images. A spherical image is hereby considered as a mapping from a given three-dimensional environment through the camera center onto a unit sphere. As this sphere is a two-dimensional manifold, two parameters are sufficient to identify any point unambiguously.

Building on that, spherical cameras typically provide images in an unwrapped form of the widely used latitude-longitude representation relying on (θ, ϕ)-coordinates with $0 \leq \theta \leq \pi$ and $0 \leq \phi \leq 2\pi$ (Figure 1(b) and 2(a)).

Problem: Since the description of image feature uncertainties implies the neighborhood of the features, local distortions as they occur close to the poles of the spherical coordinate system, affect the uncertainty description. A Gaussian distributed uncertainty on a sphere is therefore not adequately describable by a single Gaussian in (θ, ϕ)-coordinates (Figure 1(a) and 1(b)). Due to the discontinuity of the longitude-coordinate at the poles, location uncertainties spreading over the image pole cannot be represented at all.

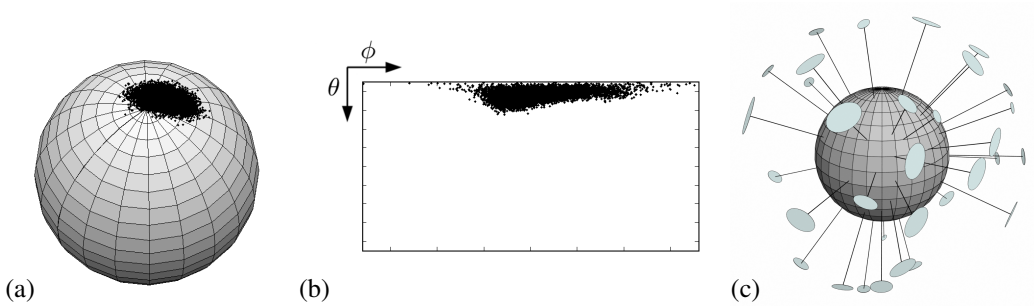


Figure 1: The Monte-Carlo based transformation from a Gaussian distribution on a sphere (a) to latitude-longitude representation (b) illustrates, that the resulting distribution cannot be properly described by a Gaussian. We propose therefore to quantify feature uncertainties through 3D uncertainties attached to normal vectors (c). Compared to latitude-longitude representation, this approach describes uncertainties consistently throughout the spherical image.

Contribution: In this work, we introduce an uncertainty representation based on normal vectors (Figure 1(c)) to overcome the previously mentioned limitations. Our proposed normal vector representation is not only limited to full spherical images. Spherical or catadioptric cameras producing a locally varying pixel-per-angle ratio as examined by Streckel and Koch [SK05] can be handled by replacing the sphere with ellipsoidal shapes.

Related work: In the field of computer vision and SfM much effort has been put into the deduction of uncertainty propagation techniques within the last years [CF05, CZZF97, DLLP11, Har98, HZ00], including approaches like first order error propagation [BBS07] or sparse grids [JU04]. Even though the listed work covers exclusively perspective imaging.

The approach taken in this work to avoid given singularities of latitude-longitude representation can to a certain extend be considered to be analogue to the concept of quaternions: Stuelpnagel shows in [Stu64] that all three-parameter representations of rotations are highly nonlinear and possess singularities in their description in three dimensional space. In order to prevent the occurrence of these singularities, the usage of quaternions introduces an additional dimension. Our usage of the normal vector exploits the same principle of introducing an additional third dimension to avoid singularities in feature uncertainty representation on a sphere.

The concept of normal vectors for position representation is furthermore not limited to spherical image feature representation but is also used in different context: Gade proposes for example a method for global navigation based on normal vectors [Gad10].

In the course of this work location uncertainties of image features are assumed to be Gaussian distributed in image coordinate system. Zeisl et al. [ZGS+09] show that this assumption is valid for the uncertainty characterization of SIFT [Low04] and SURF [BTVG06] descriptors. For other state of the art descriptors, such as CARD [AY11], BRISK [LCS11], DAISY [TLF10] or ORB [RRKB11] the appropriate feature uncertainties can be deduced under consideration of their particular manner of feature extraction.

Outline: The extraction of feature uncertainties from spherical images is described in section 2. The following section 3 outlines the proposed uncertainty description. Section 4 contains a comparison of our method with the conventional latitude-longitude representation by performing uncertainty propagation and calculating camera pose uncertainties for different scenarios. Section 5 summarizes the obtained results. The work is concluded in section 6. A video introducing the problem statement and the proposed approach is available in [Kro13].

2 FEATURE UNCERTAINTY EXTRACTION

In this section we propose a method to extract feature uncertainties from spherical images. This was achieved by decomposing a given spherical image into virtual perspective images. The positions of the appendant virtual cameras were set to the center of the spherical camera and their viewing directions were uniformly distributed throughout the sphere (Figure 2).

In the next step, features were extracted within the resulting perspective images and transformed to latitude-longitude representation of the spherical coordinate system. The performed calculations are hereby equivalent to transformations between world coordinate system and perspective camera coordinate system as outlined in Hartley et al. [HZ00], assuming the spherical camera aligned with world coordinates. The transformation between Cartesian and spherical coordinates is given in [KK00].

Within the perspective images, the uncertainty distribution of the extracted features was obtained based on the

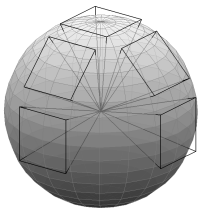
(a)  (b)  (c) 

Figure 2: Spherical image (14'000 × 7'000 pixel) of the Mogao cave number 322 in China. A set of 11 spherical HDR-images was acquired to perform pose uncertainty estimation and 3D reconstruction (a). Illustration of the decomposition of a spherical image into perspective ones (b). Resulting perspective images, in which the features are extracted (c). Note, that those cut-outs have to overlap to assure that all image data of the sphere was considered.

method introduced by Zeisl et al. [ZGS+09]. The distributions are represented by a 2D covariance matrix such as

$$\sigma_{2D} = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} \\ \sigma_{yx} & \sigma_{yy} \end{pmatrix}. \tag{1}$$

When describing these uncertainties in latitude-longitude representation of the spherical image, their essential characteristic of being Gaussian distributed is lost (Figure 1(b)). Approximating the resulting distribution by a Gaussian is inaccurate and other descriptions, such as mixture of Gaussians complicate further propagation of the obtained uncertainties.

A representation, which allows an appropriate description of the extracted uncertainties as Gaussians will therefore be presented in the following section.

3 UNCERTAINTY DESCRIPTION

This section introduces the proposed uncertainty representation of spherical images. Instead of representing a given feature in spherical coordinates through latitude-longitude representation, we propose to use a normal vector $\mathbf{n}$. This vector, described in the 3D Cartesian camera coordinate system is characterized by its perpendicular alignment to the spheres surface and its unit constraint. The previously obtained 2D uncertainties of the image features as specified in equation 1 can then be expressed by assigning an appropriate uncertainty to the normal vector $\mathbf{n}$.

The requirement of this method is a globally strictly convex and differentiable shape, the image is represented on. This assures that each point on the shape is uniquely identifiable by the normal emanating from the surface at that position. The considered spherical camera as well as other catadioptric camera models meet this requirement.

The position of an image feature $\mathbf{P}$ on the sphere, available in latitude-longitude representation as (θ_P, ϕ_P)-coordinates, can simply be transformed into the corresponding normal vector description $\mathbf{n}_P$ by

$$\mathbf{n}_P = \begin{pmatrix} \sin(\theta_P)\sin(\phi_P) \\ \sin(\theta_P)\cos(\phi_P) \\ \cos(\theta_P) \end{pmatrix}, \tag{2}$$

whereat the condition $\|\mathbf{n}_P\| = 1$ holds.

Any location uncertainty of the 2D point, such as extracted SIFT-location uncertainty, also has to be transferred from the two-dimensional representation in the local perspective images to the 3D normal vector representation. The 3D uncertainty of the normal vector is initialized as

$$\sigma_{\mathbf{n}} = \begin{pmatrix} \sigma_{2D} & \mathbf{0}_2 \\ \mathbf{0}_2^T & \sigma_\varepsilon \end{pmatrix} = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} & 0 \\ \sigma_{yx} & \sigma_{yy} & 0 \\ 0 & 0 & \sigma_\varepsilon \end{pmatrix} \tag{3}$$

by considering equation 1 for the feature uncertainty in 2D. This equation adds a third dimension to the uncertainty representation. When it comes to the implementation of this equation, it is necessary to ensure numerical stability by choosing $\sigma_{\mathbf{n}}$ to be positive-definite through setting $\sigma_{\mathbf{n}_{33}} = \sigma_\varepsilon > 0$. The value of σ_ε was chosen to

$$\sigma_\varepsilon \ll \min(\lambda_1, \lambda_2), \tag{4}$$

where λ_1 and λ_2 represent the eigenvalues of σ_{2D}. Figure 3(a) shows a geometrical visualization of $\sigma_{\mathbf{n}}$.

The following calculations outline the correct alignment of the uncertainty distribution $\sigma_{\mathbf{n}}$ with the accordant $\mathbf{n}$-vector, splitting up into two consecutive steps. Firstly, $\sigma_{\mathbf{n}}$ is converted to comply with the location of the image feature on the sphere as illustrated in figure 3(b). Afterwards, the correct orientation of the uncertainty is ensured as shown in figure 3(c).

The first step is described through the rotation vector $\mathbf{r}_A$, conforming to the constraint

$$\mathbf{n}_P = R(\mathbf{r}_A)\,\hat{\mathbf{e}}_z. \tag{5}$$

$R(\cdot)$ represents hereby a rotation matrix based on axis angle by applying the Rodrigues formula [ZF92]. The vector $\mathbf{r}_A$ is obtained through

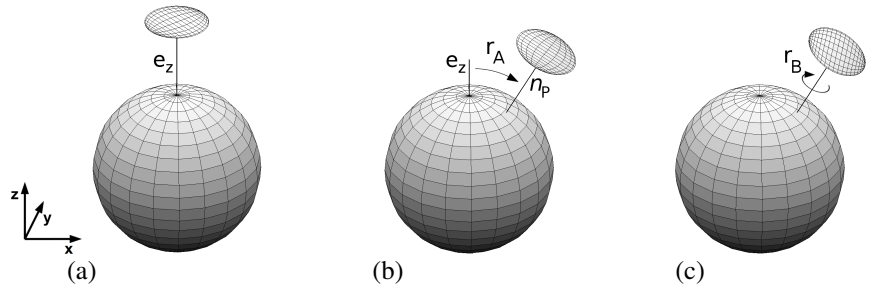


Figure 3: Uncertainty representation with the proposed normal-vector method: An ellipsoid representing the 1σ (68% confidence) neighbourhood of the extracted feature uncertainty is by default initialized at the image pole along the $\hat{e}_z$-axis (a). Then the uncertainty distribution is aligned with its corresponding location by applying $\mathbf{r}_A$ (b). Finally, the rotation $\mathbf{r}_B$ is applied to ensure the orientation of the originally extracted uncertainties (c).

$$\mathbf{r}_A = \cos^{-1}(\hat{\mathbf{e}}_z \cdot \mathbf{n}_P) \frac{\hat{\mathbf{e}}_z \times \mathbf{n}_P}{\|\hat{\mathbf{e}}_z \times \mathbf{n}_P\|}. \tag{6}$$

The geometrical alignment of $\mathbf{r}_A$ is visualized in figure 3(b) by the dashed line, oriented perpendicular to $\hat{e}_z$ and $\mathbf{n}_P$. The term $\frac{1}{\|\hat{\mathbf{e}}_z \times \mathbf{n}_P\|}$ normalizes the vector $\hat{\mathbf{e}}_z \times \mathbf{n}_P$ and $\cos^{-1}(\hat{\mathbf{e}}_z \cdot \mathbf{n}_P)$ assigns the length to the vector, which corresponds to the rotation angle. Finally, a further rotation with $\mathbf{n}_P$ as rotation axis is applied trough

$$\mathbf{r}_B = \cos^{-1}\left(\frac{\mathrm{sgn}(\mathbf{n}_{P_y}) \cdot \mathbf{n}_{P_x}}{\|\mathbf{n}_{P_{x,y}}\|}\right) \cdot \mathbf{n}_P. \tag{7}$$

Based on this rotation, it is assured that location uncertainties of features are oriented in the way they were extracted from the perspective image (Figure 3(c)). $\mathbf{n}_{P_x}$ determines the x-element of the vector $\mathbf{n}_P$. $\mathbf{n}_{P_{x,y}}$ represents the vector $(\mathbf{n}_{P_x}\ \mathbf{n}_{P_y})^\top$. The term $\mathrm{sgn}(\mathbf{n}_{P_y})$ identifies the correct quadrant for the evaluation of $\cos^{-1}(\cdot)$.

4 APPLICATION: ESTIMATION OF CAMERA POSE UNCERTAINTY

To evaluate the capability of our approach, we estimated the pose uncertainty of a spherical camera based on our proposed representation and compared the results against those obtained with latitude-longitude representation. The camera pose uncertainty is calculated based on correspondences between features of the spherical image and 3D points in the world coordinate system (wcs). Without loss of generality, we choose the wcs and the camera coordinate system (ccs) of the spherical camera to be congruent. This simplifies the transition of a 3D point $\mathbf{M}^w$ in wcs to $\mathbf{M}^c$ in ccs given through

$$\mathbf{M}^c = \mathbf{R}^{cw}\mathbf{M}^w + \mathbf{t}^{cw} \tag{8}$$

to an identity operation $\mathbf{M}^c = \mathbf{M}^w$, whereas the superscript $[\cdot]^{cw}$ identifies the direction of transformation from wcs to ccs. Additionally $\mathbf{R}$ and $\mathbf{t}$ identify the rotation and translation of wcs against ccs.

The image points $\mathbf{M}^s$ of the spherical image are finally obtained by projecting all points $\mathbf{M}^c$ through the central projection point of the camera, given as $\mathbf{t}^{cw}$, onto a unit sphere. Based on this step, 2D-3D correspondences between all points $\mathbf{M}^s$ in 2D latitude-longitude representation and $\mathbf{M}^w$ in wcs are established.

Our evaluations have been performed on synthetical and real data. For synthetical evaluations, n 3D points in wcs as well as the camera model and the image features were generated with Matlab®. For evaluation with real data, sphercial images of calibrated cameras were chosen: The camera pose had previously been calibrared using an SfM-approach for spherical images and was optimized by minimizing the squared reprojection error between image points and 3D points, projected into the image. Software such as the *Sparse Bundle Adjustment Package* provided by Lourakis and Argyros [LA09] perform this task efficiently. For the minimization of the reprojection error, the measurement function h was chosen according to Pagani and Stricker [PS11] as

$$\sum_{j=1}^{n} h(\mathbf{M}^{s_j}, \mathbf{M}^{c_j}) = \sum_{j=1}^{n} \cos^{-1}\left(\frac{(\mathbf{M}^{s_j})^T \mathbf{M}^{c_j}}{\|\mathbf{M}^{c_j}\|}\right), \tag{9}$$

which corresponds to the geodesic distance between the 2D-3D correspondences. The point $\mathbf{M}^{s_j}$ denotes hereby the jth image point on the spherical image, represented as 3D unit vector, $\mathbf{M}^{c_j}$ identifies the jth 3D point in ccs. $\left(\frac{(\mathbf{M}^{s_j})^T \mathbf{M}^{c_j}}{\|\mathbf{M}^{c_j}\|}\right)$ represents finally the normalized scalar product between $\mathbf{M}^{s_j}$ and $\mathbf{M}^{c_j}$.

For the calculation of the pose uncertainties, we assume an optimization to be performed based on equation 9. Since such an optimization does not deliver perfect agreement of image features and backprojected

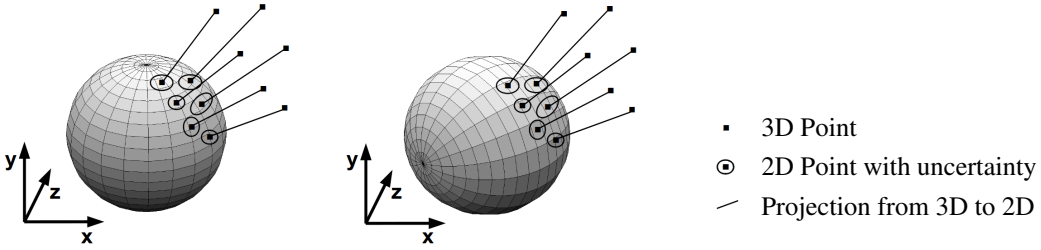


Figure 4: Two omnidirectional cameras in identical environments. 3D points are projected into the spherical images and uncertainties are attached to them. Based on their location on the sphere, their accuracy of description varies when using latitude-longitude representation, depending on their latitude-longitude position.

3D points for real data, this fact was also considered in our simulations. Image features where firstly generated through backprojection of 3D points. After applying Gaussian noise to those 3D points, they were once again backprojected creating deviation between backprojected 3D points and image features.

The uncertainty of the camera pose for simulated and real data was then obtained by performing a first order error propagation with respect to the covariance matrices of $\mathbf{t}^{cw}$ and $\mathbf{R}^{cw}$. Without loss of generality, in the current work evaluations are limited to the translation covariance of the camera to proof the reliability of the proposed representations.

The derivation of the needed Jacobian matrices from the measurement function h for performing camera pose uncertainty estimation is outlined below and implies the following preconditions: Uncertainties of all error-prone measurements, such as image points $\mathbf{M}^s$ on the sphere and 3D world points $\mathbf{M}^w$ are assumed to be Gaussian distributed. The subsequent uncertainty propagation is furthermore restricted to linear propagation methods. It is assumed, that the camera pose was optimized towards its environment by minimizing the reprojection error of 2D3D-correspondences between image points $\mathbf{M}^{s_j}$ and world points $\mathbf{M}^{w_j}$, which have to be transferred to the camera coordinate system (*ccs*) by applying equation 8. This optimization step can reliably be performed by software-packages based on the Levenberg-Marquardt-algorithm, such as SBA [LA09]. For quantifying the reprojection error between the 2D3D correspondences, the geodesic error as measurement function h is used (Equation 9).

After the optimization step, the uncertainty of the camera pose σ_{pose} is calculated using the retrieved results as linearization point. Bleser et al. [BBS07] present those calculations for perspective cameras. Since we limit our calculations to the translational component of the pose uncertainty, we obtain for spherical cameras and n 2D3D-correspondences

$$\sigma_t \approx \left(\sum_{j=1}^{n} \left(\frac{\partial h}{\partial \mathbf{t}_j^{cw}} \right)^{\top} P_j^{-1} \frac{\partial h}{\partial \mathbf{t}_j^{cw}} \right)^{-1}. \quad (10)$$

For $\partial h / \partial \mathbf{t}^{cw_j}$ we obtain

$$\frac{\partial h}{\partial \mathbf{t}^{cw_j}} = \frac{1}{\xi_j} \left(\frac{1}{\kappa_j} \mathbf{M}^{c_j} \left((\mathbf{M}^{s_j})^{\top} \mathbf{M}^{w_j} \right) - \frac{\mathbf{M}^{s_j}}{\|\mathbf{M}^{w_j}\|} \right), \quad (11)$$

with

$$\kappa_j = \left((\mathbf{M}_x^{w_j})^2 + (\mathbf{M}_y^{w_j})^2 + (\mathbf{M}_z^{w_j})^2 \right)^{3/2} \quad (12)$$

and

$$\xi_j = \sqrt{1 - \left(\frac{(\mathbf{M}^{s_j})^{\top} \mathbf{M}^{w_j}}{\|\mathbf{M}^{w_j}\|} \right)^2} \quad (13)$$

since $\frac{\partial \cos^{-1}(\alpha)}{\partial \alpha} = -\frac{1}{\sqrt{1-\alpha^2}}$.
P_j is approximated as

$$P_j \approx \left(\frac{\partial h}{\partial \mathbf{M}^{w_j}} \frac{\partial h}{\partial \mathbf{M}^{s_j}} \right) \begin{pmatrix} \sigma_{\mathbf{M}^{w_j}} & 0 \\ 0 & \sigma_{\mathbf{M}^{s_j}} \end{pmatrix} \begin{pmatrix} \frac{\partial h}{\partial \mathbf{M}^{w_j}} \\ \frac{\partial h}{\partial \mathbf{M}^{s_j}} \end{pmatrix}. \quad (14)$$

$\sigma_{\mathbf{M}^{w_j}}$ and $\sigma_{\mathbf{M}^{s_j}}$ describe the uncertainties of the world and image points as Gaussians. The matrix $\partial h / \partial \mathbf{M}^{w_j}$ represents the Jacobian of h with respect to $\mathbf{M}^{w_j}$ and $\partial h / \partial \mathbf{M}^{s_j}$ the equivalent with respect to $\mathbf{M}^{s_j}$:

$$\frac{\partial h}{\partial \mathbf{M}^{w_j}} = \frac{1}{\xi_j \kappa_j} ((\mathbf{M}^{s_j})^{\top} \mathbf{M}^{w_j}) \left((\mathbf{R}^{cw})^{\top} \mathbf{M}^{w_j} \right) - \frac{1}{\xi_j} \frac{(\mathbf{R}^{cw})^{\top} \mathbf{M}^{s_j}}{\|\mathbf{M}^{w_j}\|} \quad (15)$$

$$\frac{\partial h}{\partial \mathbf{M}^{s_j}} = -\xi_j \|\mathbf{M}^{w_j}\| \cdot \mathbf{M}^{c_j} \quad (16)$$

Based on this, the resulting pose uncertainties of the spherical cameras were evaluated for the common latitude-longitude representation as well as for the normal-vector representation. The results for the calculated pose uncertainties will be detailed in the next section.

5 EVALUATION AND RESULTS

This section presents results of pose uncertainty estimation for spherical cameras in several scenarios.

Since omnidirectional cameras are able to register image features in arbitrary directions, the calculated pose uncertainty can be expected to be independent from the camera orientation $\mathbf{R}^{cw}$, when considering $\mathbf{t}^{cw}$ to be constant. This is not the case, when choosing latitude-longitude representation for the uncertainty calculation (Figure 4).

The reason for this is illustrated in figure 1(b), exemplifying that a Gaussian distribution on a sphere cannot be adequately described by a Gaussian near the image poles. This leads to the conclusion, that the (θ,ϕ)-representation is not valid to reliably approximate error distributions with Gaussians.

To highlight this, we reduced the distribution of artificially generated image features in our evaluations with synthetic data to a very limited solid angle of the spherical image, emphasizing the influence of changing camera orientation for (θ,ϕ)-representation. The parameter σ_ε specified in equation 4 was chosen as $\sigma_\varepsilon = 10^{-10} \cdot \min(\lambda_1,\lambda_2)$ throughout the evaluations. This choice of σ_ε is justified by evaluations showing negligible impact to the results with $\sigma_\varepsilon < 10^{-8}$.

Figure 5(a-e) visualizes the resulting pose uncertainty for rotation around the $\hat{e}_x$-axis and the $\hat{e}_z$-axis for synthetic data: The evaluation was based on a distribution of $n = 20$ 3D points $\mathbf{M}^w$. The coordinates of those points as well as the vector $\mathbf{t}^{cw}$ were kept constant in wcs throughout the evaluation process. Only the orientation of the omnidirectional camera was rotated in steps of 10° by varying $\mathbf{R}^{cw}$ accordingly. With each iteration step, the distance of the points towards the pole changed.

Due to the distortion in latitude-longitude representation, the image feature uncertainties enter differently into the pose calculations. Results of pose uncertainty calculation at the image poles differ up to 27% from the uncertainty obtained at the image equator. Reason for this disparity is an overestimation of uncertainties at the poles when using latitude-longitude-representation. The resulting twofold symmetry of the pose uncertainty is explained by the fact, that the point distribution passes both poles during the evaluation steps. Throughout the whole evaluation process, the pose uncertainty calculated based on the normal vector representation is unchanged and complies with the expectation of obtaining constant results.

Our evaluation with real data was based on two datasets: The first dataset contains a set of 10 images of the Mogao cave number 322 in China. The second dataset consists of 183 images taken of a street crossing in Berlin (Oberwallstraße, Jägerstraße). All images were taken as high-dynamic range (HDR) images and have a resolution of 14'000 $\times$ 7'000 pixel. Therefore up to 108'000 image features were extracted in a single image and considered for the calculation of the camera pose uncertainty.

Since the extracted features were distributed uniformly throughout the whole image, the camera pose uncertainty in conventional latitude-longitude representation varied less compared to the synthetic scenario. When using standard low-dynamic range images, a non-uniformal distribution of features throughout the images is more likely, leading to higher disparities between the presented approaches. Figure 5(f) shows a typical result for a spherical image.

6 CONCLUSION AND FUTURE WORK

A method for uncertainty representation of spherical image features based on normal vectors has been proposed and compared to the commonly used latitude-longitude representation. It was shown, that this method - in contrast to latitude-longitude representation - is able to reliably represent image features throughout the whole sphere, since the used normal vector concept is not subject to distortions, discontinuities or nonlinear effects at the image poles.

Basing the pose uncertainty estimation of spherical cameras on the proposed approach improves its reliability and makes the quality of its results comparable to perspective estimates, such as [BBS07].

The proposed method is furthermore expected to be robust against up and down-sampling of the images, as long as a consistent uncertainty extraction of the image features is assured for different scales (e.g. by using scale-invariant SIFT [Low04], SURF [BTVG06] or ORB [RRKB11] features). A detailed evaluation has to be considered as future work.

Also the use of the proposed uncertainty representation, e.g. as additional information for surface fitting, methods for uncertainty propagation through subsequent reconstruction algorithms such as sparse reconstruction and pointcloud generation will be part of future work.

7 ACKNOWLEDGEMENTS

This work was partially funded by the BMBF-projects CAPTURE (01IW09001) and DENSITY (01IW12001).

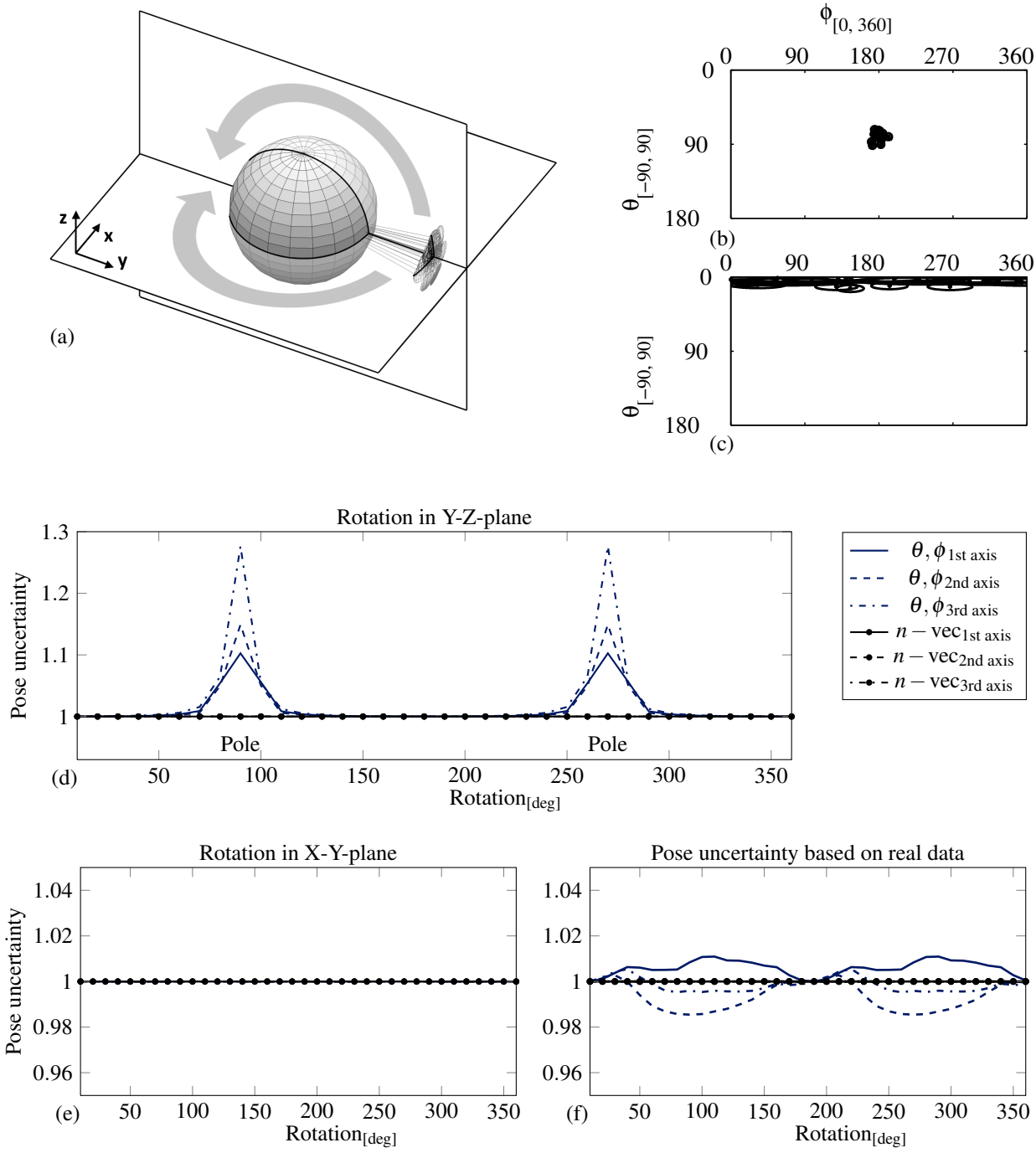


Figure 5: Subfigure (a) shows a synthetic point distribution M^s represented as normal vectors and two exemplarily evaluated rotation planes, the vectors were rotated in. Point uncertainties in latitude-longitude representation are visualized, when being located at the equator (b) and at the image poles (c). Subfigure (d) and (e) show the resulting camera pose uncertainties, expressed through the eigenvalues of the camera pose uncertainty matrices. These eigenvalues furthermore represent the expansion of the uncertainty ellipsoids of the cameras along the three principle axis (1st, 2nd, 3rd axis) and were normalized to the first evaluation step. Subfigure (f) finally illustrates a typical pose uncertainty distribution for a real spherical image with 101778 extracted features. Note that within the figures (d)-(f) all normal vector graphs (*abbrev.: n-vec*) overlap in a single plotted line.

A REFERENCES

[ADF+10] D. Anguelov, C. Dulong, D. Filip, C. Frueh, S. Lafon, R. Lyon, A. Ogale, L. Vincent, and J. Weaver. Google street view: capturing the world at street level. *IEEE Computer*, 43(6):32–38, 2010.

[AY11] M. Ambai and Y. Yoshida. Card: Compact and real-time descriptors. ICCV, 2011.

[BBS07] Gabriele Bleser, Mario Becker, and Didier Stricker. Real-time vision-based tracking and reconstruction. *Journal of Real-Time Image Processing*, 2:161–175, 2007.

[BTVG06] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. Surf: Speeded up robust features. *ECCV*, 2006.

[CF05] Eduardo Bayro Corrochano and Wolfgang Förstner. Uncertainty and projective geometry. In *Handbook of Geometric Computing*, pages 493–534. Springer Berlin Heidelberg, 2005.

[CZZF97] G. Csurka, C. Zeller, Z. Zhang, and O.D. Faugeras. Characterizing the uncertainty of the fundamental matrix. *Computer Vision and Image Understanding*, 68(1):18–36, 1997.

[DLLP11] G. Di Leo, C. Liguori, and A. Paolillo. Covariance propagation for the uncertainty estimation in stereo vision. *Instrumentation and Measurement, IEEE Transactions on*, 60(5):1664–1673, 2011.

[Gad10] K Gade. A non-singular horizontal position representation. *Cambridge Univ Press*, 2010.

[GRMG11] D. Gutierrez, A. Rituerto, JMM Montiel, and JJ Guerrero. Adapting a real-time monocular visual slam from conventional to omnidirectional cameras. *OMNIVIS*, 2011.

[Har98] R.M. Haralick. Propagating covariance in computer vision. *Performance Characterization in Computer Vision*, 1:95–115, 1998.

[HZ00] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.

[JU04] S.J. Julier and J.K. Uhlmann. Unscented filtering and nonlinear estimation. *Proceedings of the IEEE*, 92(3):401–422, 2004.

[KK00] G.A. Korn and T.M. Korn. *Mathematical handbook for scientists and engineers: definitions, theorems, and formulas for reference and review*. Dover Pubns, 2000.

[Kro13] B. Krolla. `http://www.youtube.com/watch?v=0oOr0V4u_e4`, 2013.

[LA09] M.I. A. Lourakis and A.A. Argyros. Sba: A software package for generic sparse bundle adjustment. *ACM Trans. Math. Software*, 36(1):1–30, 2009.

[LCS11] S. Leutenegger, M. Chli, and R.Y. Siegwart. Brisk: Binary robust invariant scalable keypoints. *ICCV*, 2011.

[Low04] D.G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.

[PS11] A. Pagani and D. Stricker. Structure from motion using full spherical panoramic cameras. *OMNIVIS2011*, 2011.

[RRKB11] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski. Orb: An efficient alternative to sift or surf. *ICCV*, 2011.

[SK05] B. Streckel and R. Koch. Lens model selection for visual tracking. *Pattern Recognition*, pages 41–48, 2005.

[Stu64] J. Stuelpnagel. On the parametrization of the three-dimensional rotation group. *SIAM review*, 6(4):422–430, 1964.

[TIO05] A. Torii, A. Imiya, and N. Ohnishi. Two- and three-view geometry for spherical cameras. In *Proc. Workshop Omnidirect. Vis*, pages 1–8. Citeseer, 2005.

[TLF10] E. Tola, V. Lepetit, and P. Fua. Daisy: An efficient dense descriptor applied to wide-baseline stereo. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 32(5):815–830, 2010.

[VR] Spheron VR. SceneCam® HDR http://www.spheron.com/?id=107.

[Wei] AG Weiss. Civetta camera http://www.weiss-ag.org/business-divisions/360-technology.

[ZF92] Zhengyou Zhang and Olivier Faugeras. *Three D-Dynamic Scene Analysis: A Stereo Based Approach*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 1992.

[ZGS+09] B. Zeisl, P. Georgel, F. Schweiger, E. Steinbach, and N. Navab. Estimation of location uncertainty for scale invariant feature points. BMVC, 2009.