Improved Inverse-Depth Parameterization for Monocular Simultaneous Localization and Mapping

Evren Imre, Marie-Odile Berger, Nicolas Noury

HAL Id: inria-00429318

https://inria.hal.science/inria-00429318

Submitted on 31 Mar 2010

Improved Inverse-Depth Parameterization for Monocular Simultaneous Localization and Mapping

E. Imre, M.-O. Berger, N. Noury

Abstract— Inverse-depth parameterization can successfully deal with the feature initialization problem in monocular simultaneous localization and mapping applications. However, it is redundant, and when multiple landmarks are initialized from the same image, it fails to enforce the "common origin" constraint. The authors propose two new variants that addresses both of these issues. The experimental results indicate that the proposed approach achieves a better performance at a lower computational cost.

I. INTRODUCTION

THE monocular simultaneous localization and mapping (SLAM) problem involves the estimation of the location of a set of 3D landmarks in an unknown environment (mapping), in order to compute the pose of a sensor platform within this environment (localization), via the photometric measurements of these landmarks acquired by a camera, i.e. the 2D images [3]. The best-known solutions [3]-[6] utilize either an extended Kalman filter (EKF) [1], or a particle filter [2] (with the notable exception of [7]), and point landmarks or small planar patches [4] to describe the map. The estimation procedure aims to maximize the SLAM posterior, which is defined as [8],

$$p_{SLAM} = p(\mathbf{x_k}, \mathbf{M_k} \mid \mathbf{Z_{0:k}}, \mathbf{x_0}), \qquad (1)$$

over the pose and the landmarks. In (1), $\mathbf{x_k}$ and $\mathbf{M_k}$ denote the pose and the map estimates at time instant k, respectively. $\mathbf{Z_{0:k}}$ stands for all past and current measurements, and $\mathbf{x_0}$, the initial pose estimate. The maximization involves projecting $\mathbf{M_k}$ to the measurement space, i.e. the image plane of a camera at the estimated pose (prediction), and backprojecting the difference between the predicted and the actual measurements to the state space, to update $\mathbf{x_k}$ and $\mathbf{M_k}$ (correction).

The above scheme is common to all mainstream SLAM algorithms. What sets apart the monocular SLAM is the landmark initialization stage: Since a camera is a bearing-only sensor, a single measurement can constrain the position of the corresponding landmark only to a ray originating from the camera center (the projection ray). A

straightforward solution to obtain the depth along this ray is performing the initialization from two consecutive measurements via triangulation. However, in low parallax case, which occurs when a distant landmark is observed, or the camera is moving slowly, triangulation is not reliable. Tracking the image feature between two frames with a sufficiently wide baseline may still render it a viable strategy, unless the camera has a narrow field-of-view. However, it is wasteful in terms of information, as the intervening measurements between the two frames are not incorporated into the state estimate.

One method to cope with this problem is employing an auxiliary process, as proposed in [10]: A set of depth hypotheses is obtained by sampling along the projection ray. For each hypothesis, an associated measurement search region is computed from the uncertainty of the ray and the camera, and an individual matching is performed. Each hypothesis is scored via a monotonous function of the distance between the predicted and the actual measurement. When the match-score procedure is repeated for a number of frames, the cumulative score function collapses to a single peak, where it is safe to approximate the landmark position uncertainty with a Gaussian. This method is capable of handling any form of uncertainty. One limitation is that only a finite number of hypotheses can be evaluated, as the search space for the depth must be discrete and bounded, a shortcoming which is alleviated in [5], by adjusting the sampling density so that the potential modes of the cumulative score function are examined at a higher resolution. However, the measurements obtained during the initialization do not contribute to the main SLAM process; hence, the underutilization problem remains unaddressed.

Another alternative is retaining the Gaussian model, but changing the underlying parameterization. The *inverse-depth parameterization* (IDP), proposed in [6], is inspired by the observation that, the uncertainty of inverse of depth can be approximated very well by a Gaussian at low parallax, hence can be estimated by an EKF successfully.. Moreover, IDP landmarks can be used to update the pose estimates, both in particle filter [6] and EKF-based systems [9]; therefore, the available information is fully utilized. This eliminates the need for a separate initialization process: The map is represented with a hybrid set of regular 3-parameter (Cartesian coordinates) and IDP landmarks, with each landmark type having its own measurement function, but still operating seamlessly within the same EKF. Since only a single hypothesis is evaluated, it is

computationally more efficient than the hypothesis testing method described above. Besides, it does not suffer from any resolution or search range problems: The IDP can express every point from infinity up to the camera center. The only downside is that an IDP landmark is represented by 6 parameters, whereas the minimal representation of a landmark in 3D space requires only 3 parameters. Hence, the IDP is employed as an intermediate stage, to be upgraded to the minimal parameterization, as soon as the uncertainty of the latter becomes approximately Gaussian.

In this work, the authors study the generation of multiple IDP landmarks from a single frame. This is a case in which, in the original IDP, the constraint that all such landmarks share a common origin is not enforced. Its practical significance stems from the fact that, a SLAM algorithm gathers information by observing the landmarks, and the sooner it captures the salient features of the scene, the more information it can incorporate into the estimation process. Multiple landmark generation improves the responsiveness to new salient structures, hence, the quality of the results. The authors propose two variations of the IDP, and apply them in an EKF-SLAM system to demonstrate that it is possible to achieve gains in both computational cost, and accuracy by respecting the "common origin" constraint.

The paper is organized as follows: In the next section, an EKF-SLAM system that exclusively utilizes IDP landmarks is described. In Section 3, the shortcomings of the IDP in multiple-initialization case are discussed, and two alternatives to the original parameterization are proposed. An experimental comparison of the original and the proposed parameterizations are presented in Section 4. Finally, Section 5 concludes the paper.

II. EKF-SLAM WITH INVERSE-DEPTH PARAMETERIZATION

In the system implemented for this study, the state vector of the EKF is composed of two parts: $\mathbf{x_k}$, the camera parameters, and $\mathbf{M_k}$, the set of IDP landmarks that comprise the map. $\mathbf{x_k}$ is defined as

$$\mathbf{x_k} = \begin{bmatrix} \mathbf{c_k} & \mathbf{q_k} & \mathbf{t_k} & \mathbf{w_k} \end{bmatrix}^T, \quad (2)$$

where $\mathbf{c_k}$ and $\mathbf{q_k}$ denote the position of the camera center in 3D world coordinates, and, its orientation in quaternions, respectively. $\mathbf{t_k}$ and $\mathbf{w_k}$ are the translational and rotational velocity, which describe the constant velocity model employed in the system. The definition of $\mathbf{M_k}$ is

$$\mathbf{M_k} = \begin{bmatrix} \mathbf{L_{1k}} & \cdots & \mathbf{L_{Nk}} \end{bmatrix}^T, \quad (3)$$

where $\mathbf{L_{ik}}$ is the ith landmark, and N is the number of landmarks at the time instant k. Each IDP landmark is described by 6 parameters as [9]

$$\mathbf{L_{ik}} = \begin{bmatrix} \mathbf{c_{ik}^0} & \Theta_{ik}\left(\mathbf{u_{ik}^0}, \mathbf{q_{ik}^0}\right) & \Phi_{ik}\left(\mathbf{u_{ik}^0}, \mathbf{q_{ik}^0}\right) & \lambda_{ik} \end{bmatrix}^T. \quad (4)$$

In the above equation, $\mathbf{c_{ik}^0}$ is the coordinates of the camera where the landmark is first observed, Θ_{ik} and Φ_{ik} are the azimuth and the elevation of the projection ray in the world reference frame, and λ_{ik} is the inverse of the distance (depth) of the landmark from $\mathbf{c_{ik}^0}$, along the projection ray. Both Θ_{ik} and Φ_{ik} are functions of the first 2D image measurement pertaining to the landmark, and the orientation of the camera which made this measurement, namely, $\mathbf{u_{ik}^0}$ and $\mathbf{q_{ik}^0}$. Their dependence on k signifies that, any update to the angles implicitly means an update to the first orientation and measurement.

The IDP is related to the minimal 3D point parameterization through the expression

$$\mathbf{L^{3D}} = \mathbf{c^0} + \frac{1}{\lambda}\begin{bmatrix} \cos \Phi \sin \Theta & -\sin \Phi & \cos \Phi \cos \Theta \end{bmatrix}^T, \quad (5)$$

and the measurement equation is simply the projection of $\mathbf{L^{3D}_{ik}}$ to the image plane of the camera [11], i.e.

$$\begin{bmatrix} h_{ik}^x & h_{ik}^y & h_{ik}^z \end{bmatrix}^T = \mathbf{R}(-\mathbf{q_k})[\mathbf{I} \quad -\mathbf{c_k}]\begin{bmatrix} \mathbf{L^{3D}_{ik}} \\ 1 \end{bmatrix},$$

$$\mathbf{u_{ik}^p} = \begin{bmatrix} v_x - \alpha_x \dfrac{h_{ik}^x}{h_{ik}^z} & v_y - \alpha_y \dfrac{h_{ik}^y}{h_{ik}^z} \end{bmatrix}^T \quad (6)$$

where $(v_x; v_y)$ denote the principal point of the camera and $(\alpha_x; \alpha_y)$, the focal length-related scale factors. $\mathbf{u_{ik}^p}$ is the predicted image measurement for the landmark $\mathbf{L_{ik}}$. $\mathbf{R(q)}$ maps the quaternion $\mathbf{q}$ to the corresponding rotation matrix.

The EKF-SLAM used in this work is very similar to the one described in [9]. The inputs are the initial pose and a small number of initial landmarks, known with high accuracy, to set the correct scale for the map. For a short period, the algorithm operates in the initialization mode, i.e., only uses the initial landmark set for updating the state, to ensure that the velocities converge to the correct range. New landmarks are added using FAST [12], when the number of observations in an image falls below a certain number (T_{min}), and in regions where the number of observations is less than the average. For matching, normalized cross correlation is utilized [11]. A landmark is deleted when it is missed for a number of consecutive frames, or when the ratio of the number of observations to the number of attempts is below a certain threshold.

III. IMPROVED INVERSE DEPTH PARAMETERIZATION

A. Problem: IDP and Multiple Landmark Initialization

Equation (4) implies that the IDP actually encodes the pose estimate of the camera that first observed the landmark (i.e. $\mathbf{c_{ik}^0}$ and $\mathbf{q_{ik}^0}$), the image coordinates of the first observation ($\mathbf{u_{ik}^0}$) and the inverse depth. A *cluster* is a set of landmarks initialized from the same image. All members of a cluster

share the same $\mathbf{c^0}$ and $\mathbf{q^0}$. The "common origin" constraint enforces that this commonality is preserved throughout the entire estimation process. However, the IDP encodes these parameters separately for each individual landmark, i.e., keeps multiple representations of the $\mathbf{c^0}$-$\mathbf{q^0}$ pair of the cluster, with two important consequences:

- When the pose uncertainty (which is closely related to the initial covariance of $\mathbf{c^0}$, Θ and Φ) is much smaller than the measurement noise, the pose related parameters remain very close to their initial values, hence, each other, and the multiple copies become redundant. Such redundancy increases the size of the state vector and the computational cost.
- When the pose uncertainty is sufficiently large to allow significant updates, the copies may diverge. Such violation of the common-origin constraint deteriorates the performance.

The solution lies in factoring out the common pose parameters, and encoding the clusters, instead of the individual landmarks, as discussed next.

B. IDP1- "The Quick"

A straightforward application of the above idea is defining a *super-landmark* (IDP1), which is composed of a single $\mathbf{c^0}$, and the azimuth and elevation angles, and the inverse depth values of the member landmarks, i.e.,

$$\mathbf{L}^{\mathbf{Q}}_{\mathbf{ik}} = \begin{bmatrix} \mathbf{c}^{\mathbf{0}}_{\mathbf{ik}} & \mathbf{l}^{\mathbf{1}}_{\mathbf{ik}} & \cdots & \mathbf{l}^{\mathbf{M}}_{\mathbf{ik}} \end{bmatrix}^{T}, \qquad (7)$$

where M is the number of landmarks in the ith super-landmark at time instant k. $\mathbf{l}^{\mathbf{j}}_{\mathbf{ik}}$ contains the remaining parameters of the landmark it corresponds to, i.e.

$$\mathbf{l}^{\mathbf{j}}_{\mathbf{ik}} = \begin{bmatrix} \Theta^{0j}_{ik}\left(\mathbf{u}^{\mathbf{0j}}_{\mathbf{ik}}, \mathbf{q}^{\mathbf{0j}}_{\mathbf{ik}}\right) & \Phi^{0j}_{ik}\left(\mathbf{u}^{\mathbf{0j}}_{\mathbf{ik}}, \mathbf{q}^{\mathbf{0j}}_{\mathbf{ik}}\right) & \lambda^{j}_{ik} \end{bmatrix}. \qquad (8)$$

M decreases over time, as the members of a super-landmark are lost. When M=1, the super-landmark is identical to an IDP landmark, and when M=0, it is deleted.

A super-landmark encodes a cluster of size M with $3M+3$ parameters, as opposed to the original IDP, which needs $6M$ parameters. Since the operations performed on a super-landmark is identical to those applied to an IDP landmark, a decrease in the computational cost is guaranteed when M>1. It also partially enforces the common-origin constraint, therefore is expected to be more accurate.

C. IDP2- "The Strict"

A logical extension of the IDP1 is the removal of the duplicate orientation components, i.e.

$$\begin{aligned} \mathbf{L}^{\mathbf{S}}_{\mathbf{ik}} &= \begin{bmatrix} \mathbf{c}^{\mathbf{0}}_{\mathbf{ik}} & \mathbf{q}^{\mathbf{0}}_{\mathbf{ik}} & \mathbf{l}^{\mathbf{1}}_{\mathbf{ik}} & \cdots & \mathbf{l}^{\mathbf{M}}_{\mathbf{ik}} \end{bmatrix}^{T} \\ \mathbf{l}^{\mathbf{j}}_{\mathbf{ik}} &= \begin{bmatrix} \mathbf{u}^{\mathbf{0j}}_{\mathbf{ik}} & \lambda^{j}_{ik} \end{bmatrix} \end{aligned}. \qquad (9)$$

This variant, dubbed IDP2, is linked to the minimal 3D point parameterization as

Fig. 1. *Top left:* The bed used in the experiment to produce the ground truth trajectory. *Top right:* The initial image for *S1* and *S2*. *Bottom row:* The final images of *S1* and S2, respectively.

$$\mathbf{u}^{\mathbf{0j}} = \begin{bmatrix} u^{0j}_x & u^{0j}_y \end{bmatrix}$$

$$\mathbf{n} = \mathbf{R}(\mathbf{q}^{\mathbf{0}}) \begin{bmatrix} \dfrac{u^{0j}_x - \nu_x}{\alpha_x} & \dfrac{u^{0j}_y - \nu_y}{\alpha_y} & 1 \end{bmatrix}^{T}. \qquad (10)$$

$$\mathbf{L}^{\mathbf{3D}j} = \mathbf{c}^{\mathbf{0}} + \dfrac{1}{\lambda^j} \dfrac{\mathbf{n}}{\|\mathbf{n}\|}$$

Equations (6) and (10) constitute the measurement equation.

IDP2 achieves a complete separation of the pose and the first image measurement, and strictly enforces the common-origin constraint. However, this comes at a cost of reduced computational efficiency, as a cluster of size M is represented with $3M+7$ parameters, and after each state update, all quaternions must be normalized.

D. Related Work

Very recently, two more inverse depth parameterizations, which offer an accuracy vs. efficiency trade-off, were reported. In [14], the trade-off involves, in effect, setting the uncertainty of $\mathbf{c^0}$ to 0, to obtain a 4-component representation. On the other hand, [15] is a multi-landmark parameterization which can encode each member of a cluster by a single parameter, under the assumption that the initial measurement, $\mathbf{u^0}$, is not subject to any error.

IV. EXPERIMENTAL RESULTS

The performance of the proposed parameterizations is studied via a pose estimation task. For this purpose, a bed, which can translate and rotate a camera in two axes with a positional and angular precision of 0.48 mm and 0.001°, respectively, is used to acquire the images of an indoor scene, with the dimensions 5x2x3 meters. In the sequence *S1*, the camera moves on a horizontal trajectory of 50 cm, with constant translational and zero angular velocity. In *S2*, the movement is on a 65-cm oblique path, with rotation, and obeys the constant angular and translational velocity model. *S1d* and *S2d* are derived from *S1* and *S2*,

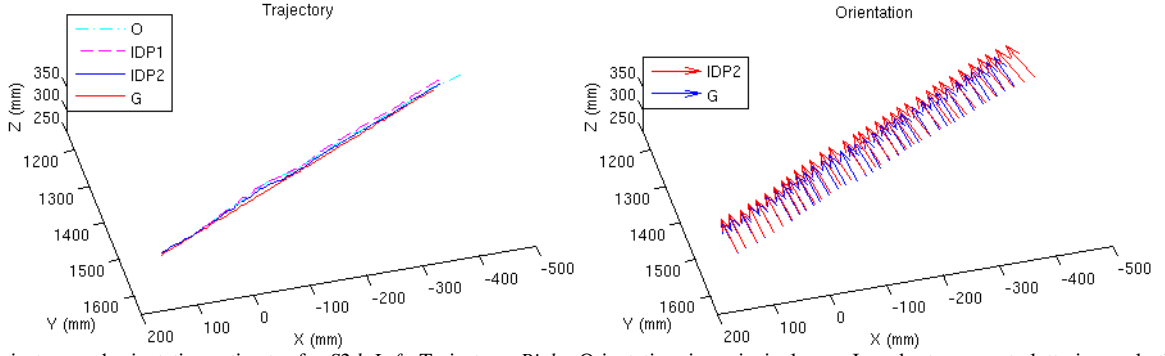| Parameter	D1			D2			D1r			D2r		
	IDP	IDP T1	IDP T2	IDP	IDP T1	IDP T2	IDP	IDP T1	IDP T2	IDP	IDP T1	IDP T2
Relative Execution Time (IDP)	1	0.36	0.71	1	0.27	0.43	1	0.39	0.71	1	0.37	0.62
Median # Measurements for Convergence	27	27	28	27	27	29	14	14	15	15	16	16
Maximum # Measurements for Convergence	55	58	58	75	75	75	119	74	44	78	87	43

Fig. 2: Trajectory and orientation estimates for *S2d*. *Left:* Trajectory. *Right:* Orientation, i.e, principal axes. In order to prevent cluttering, only the IDP2 case is plotted, and the trajectory is downsampled by 4.
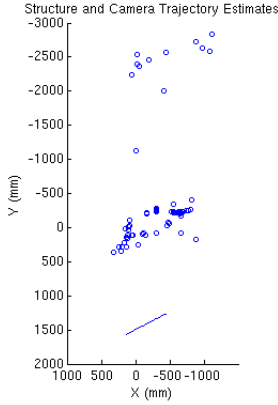

Fig. 3. Top view of the trajectory and structure estimate for *S2d*.in IDP2 case. Blue circles indicate the estimated landmarks. Those closer to the camera belong to the table, whereas the farther ones are located on the wall. The figure is representative of the IDP and IDP1 cases.

respectively, by discarding 2/3 of the images randomly to obtain a nonconstant-velocity motion. Figure 1 depicts the setup, and the initial and final images of the sequences.

The pose estimation task involves the computation of the position and the orientation of the camera from the sequences by using the EKF-SLAM scheme described in Section 2, with the map representation alternatives described in Sections 2 and 3. The maps utilize exclusively the original IDP, or one of the proposed variants. For all experiments, T_{min} is set to 30. New landmarks are added in batches of 5. The recovered maps are similar, but not identical, as each parameterization leads to a different trajectory estimate, and the lifetime of landmarks are linked closely to the pose estimates.

The following criteria are utilized for the evaluation:

- **Position error:** The Euclidean distance between the estimated and the actual camera path, in millimeters.
- **Orientation error:** The angle between the the estimated and the actual normals to the image plane

(i.e., the principal axis vectors), in degrees.

- **Relative execution time:** Execution time with respect to that of the IDP. The adoption of this criterion as a measure of the computational load is due to the fact that the alternative, the absolute execution time, also depends on the hardware, and the maturity and efficiency of the code, factors that are not within the scope of this paper.
- **# Measurements for convergence:** *Linearity index*, proposed in [9], is a measure of the Gaussianity of the depth uncertainty of an IDP landmark. Convergence occurs when the linearity index of a landmark falls below 0.1, an empirically determined threshold [13], and at that point, it is safe to switch to the 3-parameter, minimal parameterization. The lower the number of measurements necessary for convergence, the better, as this implies a faster convergence rate.

The results are presented in Table I and Figures 2-5.

The first row of Table I indicates that both IDP1 and IDP2 fulfill their promise of reduced computational cost. Especially, IDP1 achieves a 250-300% increase in processing speed. The gains provided by IDP2 are more modest in comparison, as discussed in Section III-C, but still significant. In [9], it is maintained that a hybrid map that uses both the IDP and minimally parameterized landmarks can achieve 30 frames per second (fps). Since a map represented only with IDP1 landmarks can already be processed at 15-25 fps when T_{min}=30, it can reasonably be expected that in a hybrid map application, both IDP1 and IDP2 can go beyond 30 fps.

The the trajectory and the map estimates in Figure 2 and Figure 3 for *S2d*, the most challenging sequence, suggest that the gain in computational cost comes without a compromise in performance. On the contrary, in Figure 4, it
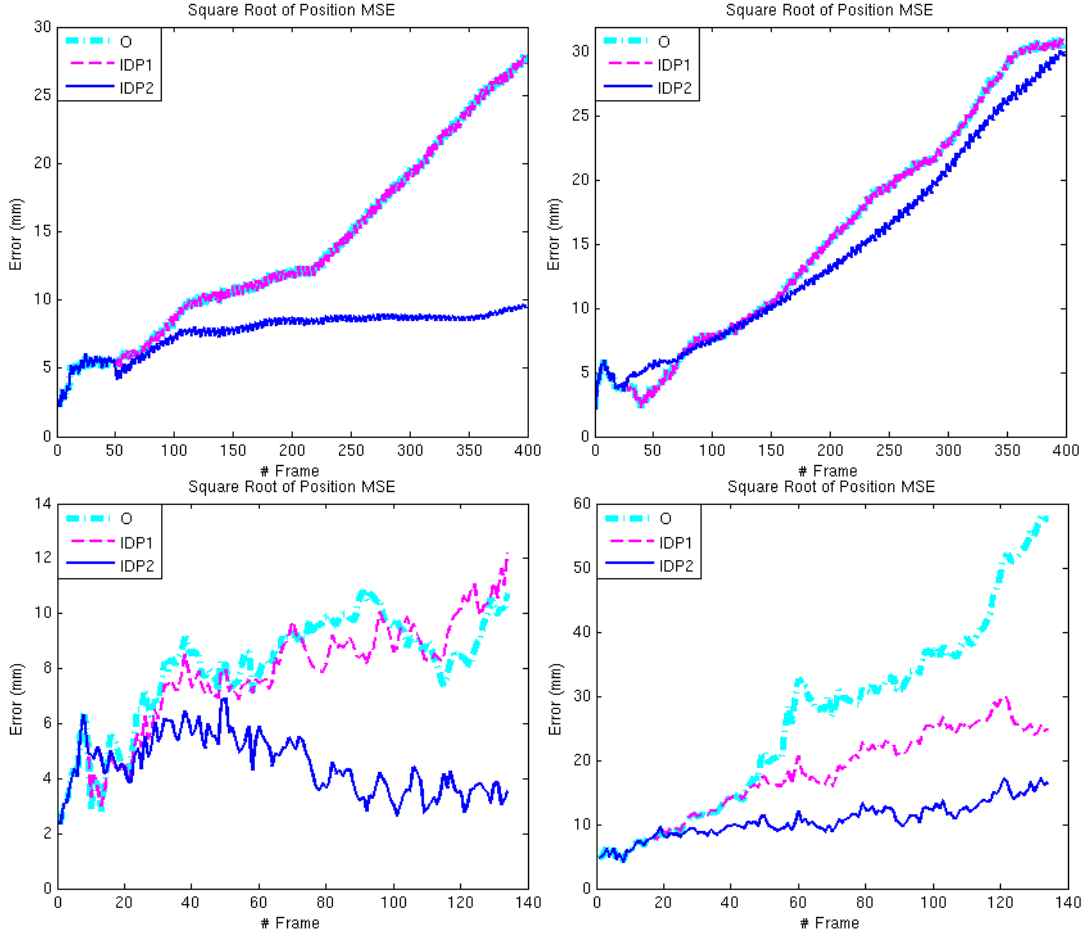
Fig. 4. Euclidean distance between the estimated and the actual trajectory. *Top-left: S1. Top-right: S2. Bottom-left: S1d. Bottom-right: S2d.*

is observed that, the IDP2 consistently outperforms the IDP and the IDP1, owing to employing a superior parameterization that enforces the common-origin constraint. However, Figure 4 also hosts some unexpected results: In *S1* and *S2*, the performances of the IDP and IDP1 are almost identical. A closer study reveals the cause as the low pose uncertainty case, described in Section III-A: The initial pose and map estimates are known with very high accuracy, and the measurement covariance is much lower than the initial uncertainty of the velocity parameters. Therefore, the angular and translational velocities are very responsive to the updates, and given that the actual motion perfectly satisfies the constant-velocity model, their covariances are reduced rapidly. Since the state vector contains the same information, the predicted measurements, the measurement errors, and the state updates, hence the next estimate of the state vector, remain the same. *S1d* and *S2d* further support this explanation. The less predictable camera motion leads to a higher pose uncertainty, and in this case, the IDP1 performs better. Such a phenomenon is not observed in the IDP2, as it has a considerably different state vector, i.e. a landmark is described by 10 parameters, instead of 6, hence, traces a different path through the solution space.

As for the orientation estimation, Figure 5 indicates that

the error is already small, especially in *S1* and *S1d*, so any performance comparison is not meaningful. The increasing error observed in *S2* and *S2d* in all parameterizations is a byproduct of the accumulation of the linearization errors, and is expected in the absence of a loop closure.

The convergence rate for various sequences and parameterizations are presented in Table I. In all cases, the median rate is an indicator of convergence for the landmarks on the table, as they vastly outnumber the landmarks on the wall. On the other hand, the maximum rate always belongs to a wall landmark, as it takes longer for farther landmarks to achieve the necessary parallax.

The only cases where there is a difference between the parameterizations are the wall landmarks in *S1d* and *S2d*. In such landmarks, the utilization of the common-origin constraint is important, as unlike the closer landmarks, the constraints provided by parallax are weak. This statement also holds for *S1* and *S2*, but both are 3 times longer than *S1d* and *S2d*, hence, the weakness of the individual parallax constraints are offset by their sheer numbers.

V. CONCLUSION

The IDP is introduced in [6] to the landmark initialization problem in monocular SLAM as an effective tradeoff
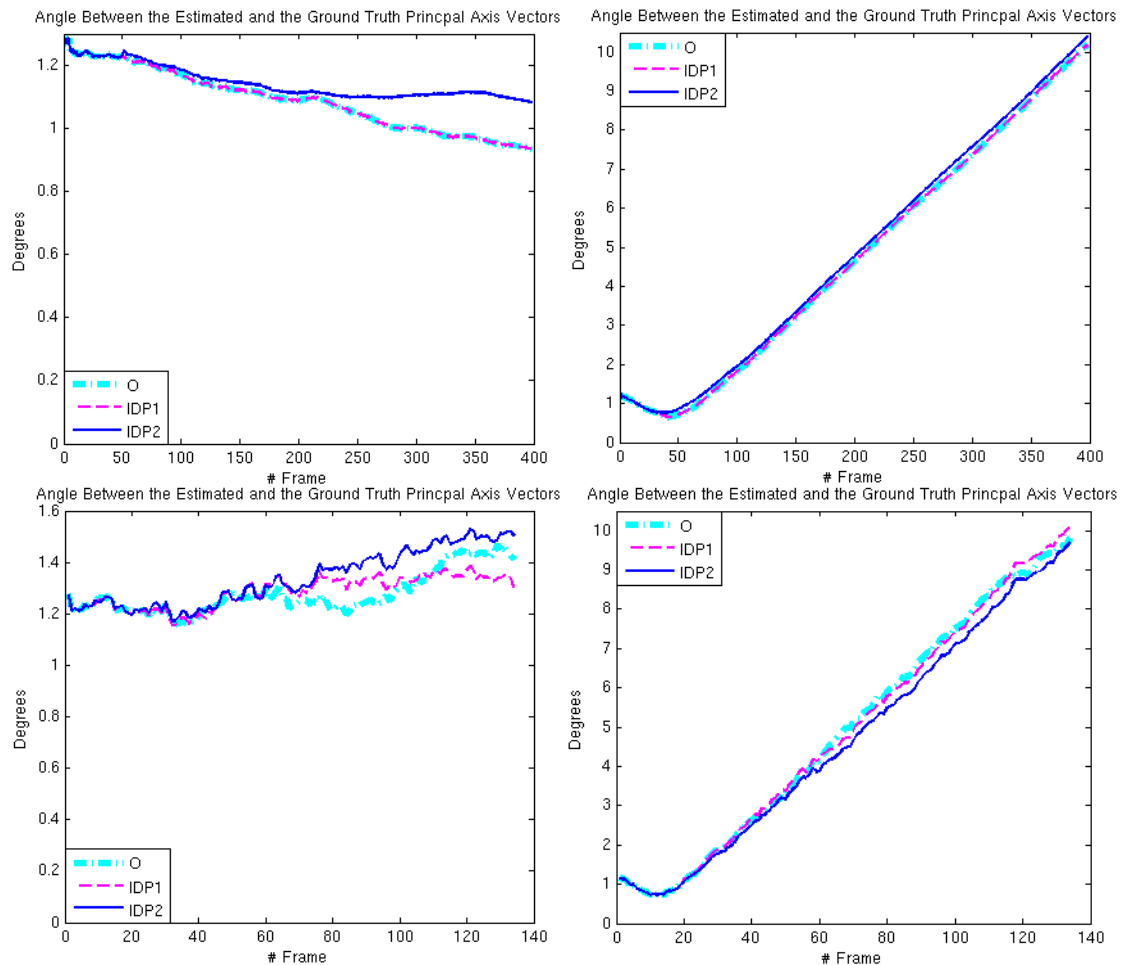
Fig. 5. The angle between the estimated and the ground truth principal axis. *Top-left: S1. Top-right: S2. Bottom-left: S1d. Bottom-right: S2d.*

between computational efficiency and information utilization efficiency. However, when multiple landmarks are initialized, a case with practical importance, the IDP fails to enforce the common-origin constraint. The authors propose two alternatives in order to alleviate the efficiency and performance hazard this condition causes, and experimentally demonstrate that, the IDP2 is superior to the IDP both in computational cost and in performance, whereas the IDP1 delivers a similar performance to the IDP at a much lower computational burden.

REFERENCES

[1] R. C. Smith, P. Cheeseman, "On the Representation and Estimation of Spatial Uncertainty," *International Journal of Robotics Research*, vol. 5, pp. 56-68, Dec. 1986.

[2] M. Montemerlo, "FastSLAM: A Factored Solution to the Simultaneous Localization and Mapping," Ph. D. dissertation, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, 2003.

[3] A. J. Davison, I. D. Reid, N. D. Molton, O. Stasse, "MonoSLAM: Real-Time Single Camera SLAM," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 29, pp. 1052-1067, Jun. 2007.

[4] H. Jin, P. Favaro, S.Soatto, "A Semi-Direct Approach to Structure from Motion," *The Visual Computer,* vol. 19, pp.377-394, Oct. 2003.

[5] M. Pupilli, A. Calway, "Real-Time Camera Tracking Using a Particle Filter," *16th British Machine Vision Conf.*, Oxford, United Kingdom, Sep. 5-8, 2005, vol.1, pp. 519-528.

[6] E. Eade, T. Drummond, "Scalable Monocular SLAM," *2006 IEEE Conf. Computer Vision and Pattern Recognition,* New York, NY, Jun 17-22, vol. 1, pp. 469-476.

[7] D. Nistèr, O. Narodistsky, J. Bergen, "Visual Odometry," 2004 IEEE Conf. Computer Vision and Pattern Recognition, Washington, DC, Jun. 27-Jul. 2, vol. 1, pp. 652-659.

[8] H. Durrant Whyte, T. Bailey, "Simultaneous Localization and Mapping: Part I, " *IEEE Robotics and Automation Magazine,* vol. 13, pp.99-110, Jun. 2006.

[9] J. M. M Montiel, J. Civera, A. J. Davison, "Unified Inverse Depth Parameterization for Monocular SLAM,"*Robotics Science and Systems 2006,* Philadelphia, PA, Aug. 16-19, 2006, Available: www.roboticsproceedings.org/rss02/.

[10] A. J. Davison, "Real-Time Simultaneous Localization and Mapping with a Single Camera," *9th Int. Conf. Computer Vision,* Nice, France, Oct. 13-16, 2003, vol. 2, pp.1403-1410.

[11] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge UK: University Press, 2003.

[12] E. Rosten, T. Drummond, "Fusing Points and Lines for High Performance Tracking," *10th Int. Conf. Computer Vision*, Beijing, China, Oct. 15-21, 2005, vol. 2, pp. 1508-1515.

[13] J. Civera, A. J. Davison, J. M. M. Montiel, "Inverse Depth to Depth Conversion for Monocular SLAM," *2007 IEEE Int. Conf. Robotics and Automation*, Rome, Italy, Apr.10-14, vol. 1, pp.2778-2783.

[14] D. Marzorati, D. Matteucci, D. Migliore, D. G. Sorrenti, "Monocular SLAM with Inverse Scaling Parameterization," *19th British Machine Vision Conf.*, Leeds, UK, Sept. 1-4, 2008.

[15] T. Pietzsch, "Efficient Feature Parameterization for Visual SLAM Using Inverse-Depth Bundles," *19th British Machine Vision Conf.*, Leeds, UK, Sept. 1-4, 2008.