university
of milano
bicocca

DISCo

QD quaderni

department of informatics, systems and communication

Emerging Paradigms in Informatics, Systems and Communication

Carlo Batini and Raimondo Schettini

research report n.1
june 2009

Starrylink Editrice

Title*: Emerging Paradigms in Informatics, Systems and Communication*
QD Quaderni - Department of informatics, systems and communication
Research Report n.1, June 2009
Editors: Carlo Batini and Raimondo Schettini

Thesis and Research
www.starrylink.it
editrice@starrylink.it

ISSN 1828-3357
ISBN 978-88-96225-21-9

Printed by Laser Copy Center (MI), June 2009

university
of milano
bicocca

UNIVERSITÀ DEGLI STUDI DI MILANO BICOCCA

disco **QD** quaderni

department of informatics, systems and communication

Emerging Paradigms in Informatics, Systems and Communication

Carlo Batini and Raimondo Schettini

research report n. 1
june 2009

Starrylink Editrice

Inverse Scaling Parametrization for Visual SLAM

D. Marzorati[1], M. Matteucci[2], D. Migliore, and D. G. Sorrenti[1]

[1] DISCo, Università di Milano-Bicocca,
Viale Sarca, 336 - 20126 Milano, Italy
[2] Politecnico di Milano, Dipartimento di Elettronica e Informazione,
via Ponzio 34/5, 20133, Milano, Italy

Abstract. We propose a parametrization alternative to the known Unified Inverse Depth, called Inverse Scaling, that also allows an un-delayed features initialization, but reduced the number of needed parameters and simplifies the measurement model. Experiments in simulation and real data demonstrate that the use of Inverse Scaling improves the performance of the monocular EKF SLAM filter.

1 Introduction

One of the most relevant problems encountered in mobile robotics is to obtain a reliable map of the robot workspace. The difficulties in this task come from the impossibility of an a-priori compilation of a map of every environment a robot might be located in, detailed enough for performing mobile robotics tasks, e.g., navigation. The problem lays in handling the inherent inaccuracies of both the sensing and the motion systems, to obtain an accurate map. In order to map a realistic environment it is necessary, for a mobile robot, to move, i.e., to explore, while incrementally building a consistent map of the environment. This requires the simultaneous determination of its location within the map. In the literature different solutions to this problem are known, and most of them are from the probabilistic Simultaneous Localization And Mapping (SLAM hereafter) subfield, see [8]. The Extended Kalman Filter (EKF) is one of the most used technique to solve this problem, and its usage was formalized by Durrant Whyte et al. [4]. Strong point, and major drawback, of this approach is the assumption of Gaussianity of the error distribution on the measures. On one hand, this assumption reduces the difficulty of the process; on the other, it can lead to inconsistencies due to the non-linearities of the problem, which make non-realistic the Gaussianity assumption.

Multi-mega-pixel cameras, very small in size, and lightweight, carefully engineered for low power consumption, and quite inexpensive, are of widespread usage in today mobile phones. In the last years, many works on visual-SLAM have been presented, e.g., [3], [2], [6], with some impressive results. The richness

Quaderni del Dipartimento
University of Milano – Bicocca ©, 1-2009

DIPARTIMENTO DI INFORMATICA SISTEMISTICA E COMUNICAZIONI

of the vision output, i.e., its usability for different tasks, as well as the considerations about cost, size, weight, and power consumption, are our main motivations for a vision-only SLAM approach. This research field has also applicability outside robotics, like in video-surveillance, automotive, wearable system etc.

In this paper we propose a novel parametrization for vision data, to properly model its uncertainty. It is called Inverse Scaling parametrization and has the twofold goal of better uncertainty modeling and improved EKF linearization. In the next section we shortly introduce our novel parametrization. Then, in Section 3, our proposal is validated on both simulated and real data and the results are compared with the state of the art parametrization proposed in [7] for the monocular SLAM problem.

2 The Inverse Scaling Parametrization

SLAM systems based on vision generate data affected by uncertainties that strongly depend on the observer-to-feature distance and thus they can not be modeled with a Gaussian propagation approach, i.e., the basic Extended Kalman Filtering assumptions. Montiel et al. [7] faced this issue by changing the parametrization of the features in the filter, adopting an Inverse Depth parametrization to make the uncertainty distribution, in the new parameter space, more similar to a Gaussian (see [1] for more details) and to allow a reduction in the non linearity of the measurement equation.

We propose the "Inverse Scaling parametrization", in order to reduce the non-linearity of the measurement equation and the space required for storing each feature. Let consider Figure 1 and the points in homogeneous coordinates $\mathbf{X} = (X, Y, Z, 1)$, $\mathbf{X_1} = (X', Y', Z', 1)$ and $\mathbf{X_2} = (X'', Y'', Z'', 1)$. They can be interpreted as the vertices of three similar triangles. In the homogeneous space, being α_1 and α_2 two scale factors that turn respectively $\mathbf{X}$ in $\mathbf{X}_1$ and $\mathbf{X}_2$, we can always rewrite the similarity by introducing the inverse of the scale, as follows:

$$\mathbf{X} = \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \equiv \begin{pmatrix} X' \\ Y' \\ Z' \\ 1/\alpha_1 \end{pmatrix} \equiv \begin{pmatrix} X'' \\ Y'' \\ Z'' \\ 1/\alpha_2 \end{pmatrix}. \tag{1}$$

If we know the 2D image point $\mathbf{x} = (u, v)$, i.e., the projection of $\mathbf{X}$ on the image plane, being it the vertex of another similar triangle, we can write:

$$\mathbf{X} = \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \equiv \begin{pmatrix} u \\ v \\ f \\ 1/\alpha \end{pmatrix} \equiv \begin{pmatrix} u \\ v \\ f \\ \omega \end{pmatrix}, \tag{2}$$

where f is the focal length of the camera and α is the scale factor for this new triangle. We also renamed $1/\alpha$ as ω since this will be the variable we are

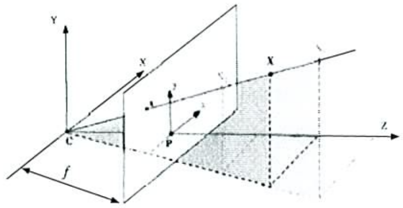


Fig. 1. Pinhole camera model and the intuition behind the Inverse Scaling parametrization.

interested in, for modeling the uncertainty; we name it *Inverse Scaling*. This reasoning shows that we can represent the whole set of points laying on the a projection ray of a 3D world point with four parameters; the fourth parameter, i.e., the scale parameter α or its inverse ω, acts as line coordinate on the projection ray. We can thus define the position of a 3D point using:

$$\begin{pmatrix} X_i \\ Y_i \\ Z_i \end{pmatrix} = \frac{1}{\omega_i} \begin{pmatrix} u_i \\ v_i \\ f_i \end{pmatrix} . \tag{3}$$

Equation 3 uses non-normalized homogeneous coordinates and a possibly different focal length for each feature. Although this might appear too general, as it could be simplified because we do need only 3 coordinates to represent a 3D point, by doing as we propose we can model the skewed uncertainty distribution of the visual data simply by using a Gaussian noise model for ω_i. In this way we obtain the desired accuracy of the cartesian distribution on the left hand side of Equation 3 by using only Gaussian distributions in the Inverse Scaling space of the right hand side of Equation 3. This uncertainty model resembles the Inverse Depth approach; however, we claim that our parametrization simplifies the measurement equation of such state-of-the-art parametrization, while reducing the space required for each feature.

3 Experimental Results

In this section we perform a quantitative and comparative validation of our proposal for the Monocular SLAM problem in a simulated environment, while a real implementation is used to evaluate qualitatively the effectiveness of the approach. To verify if a better uncertainty modeling leads to a better Monocular SLAM results (somehow confirming the results in [5]), we compared, in a 2D simulated rectangular environment (point features are equally distributed along the environment borders), on one hand a monocular EKF SLAM system based on Inverse Depth parametrization, and on the other hand, the same monocular EKF SLAM, but based on the Inverse Scaling parametrization. Our Inverse Depth implementation is an adapted version of [7] to the particular case of 3DoF simulation, using direct cosine matrices instead of quaternions for rotations. The data association has been performed manually so that estimates are comparable and the main aspect to be considered is uncertainty modeling. We assume obvious to the reader that the differences between the two approaches in a real setting, i.e., without relying on the correct data association, can only be amplified, since better uncertainty modeling can only improve data association.

In Figure 2, we have the plots of both the error in pose estimation, respectively for x, y and ϑ during the robot path and the final map estimation. The path is a simple circle with the camera always looking outside, i.e., toward the borders of the environment. As it can be easily noticed the variance of the robot pose estimate (blue lines placed at $\pm 3\sigma$) is underestimated for the Inverse Depth parametrization while this is not the case for the Inverse Scaling parametrization. This underestimation leads to filter inconsistency around step 600. As it is visible also in the map reconstruction, inconsistencies in the features position are located especially in the bottom right angle of the map.

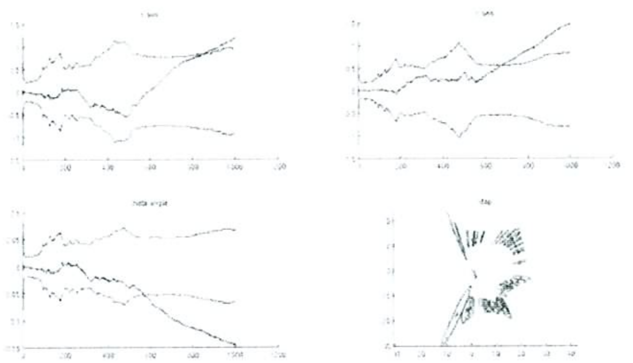
We have also tested our parametrization in a real indoor context. In Figure 3 some frames are reported, taken using a 320x240 BW camera at 30Hz. The camera was hand held and moved; the figure shows the results of the estimation process using the proposed monocular approach: the top-left image shows the camera image (here we have in red the prediction for the features, and in blue the matched ones); the top-right image presents the estimated map with the uncertainty ellipsoids, the bottom-left shows the camera position with its uncertainty ellipsoid, and the bottom-right features the camera trajectory. All the views are from the above, but the system uses a full 6DoF pose representation.

Conclusions

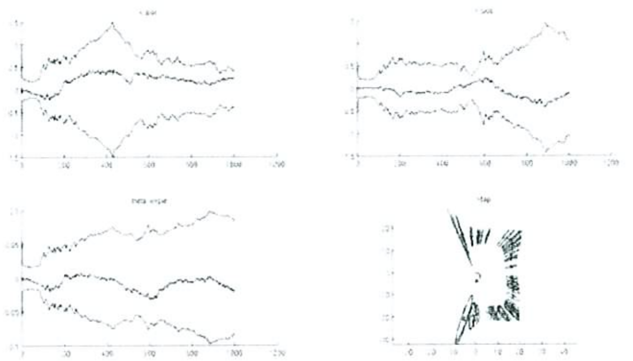
We proposed the Inverse Scaling Parametrization, and compared with the state-of-the-art Unified Inverse Depth Parametrization by performing experiments both in simulation and real. The use of Inverse Scaling improves the performance of the monocular EKF SLAM filter, when compared with the Unified Inverse Depth.

Acknowledgment

This work has been supported by the EEC contract FP6-045144 (RAWSEEDS).



(a)



(b)

Fig. 2. Map reconstruction and error in robot localization (x,y,ϑ): (top) using Inverse Depth parametrization, (bottom) using Inverse Scaling parametrization. In red the error w.r.t. the ground truth, in blue $\pm 3\sigma$.

References

1. Javier Civera, Andrew J. Davison, and J. M. M. Montiel. Inverse depth to depth conversion for monocular slam. In *ICRA*, pages 2778–2783, 2007.
2. A. Davison. Real-time simultaneous localisation and mapping with a single camera. In *Proc. IEEE Int. Conf. Computer Vision*, pages 1403–1410, 2003.
3. A. J. Davison, I. D. Reid, N.D. Molton, and O. Stasse. MonoSLAM: Real-time single camera SLAM. *IEEE Trans. on PAMI*, (6), 2007.
4. J.J. Leonard and H.F. Durrant-Whyte. Simultaneous map building and localisation for an autonomous mobile robot. *Proc. IEEE Conf. Intell. Robots Syst. (IROS)*, 1991.

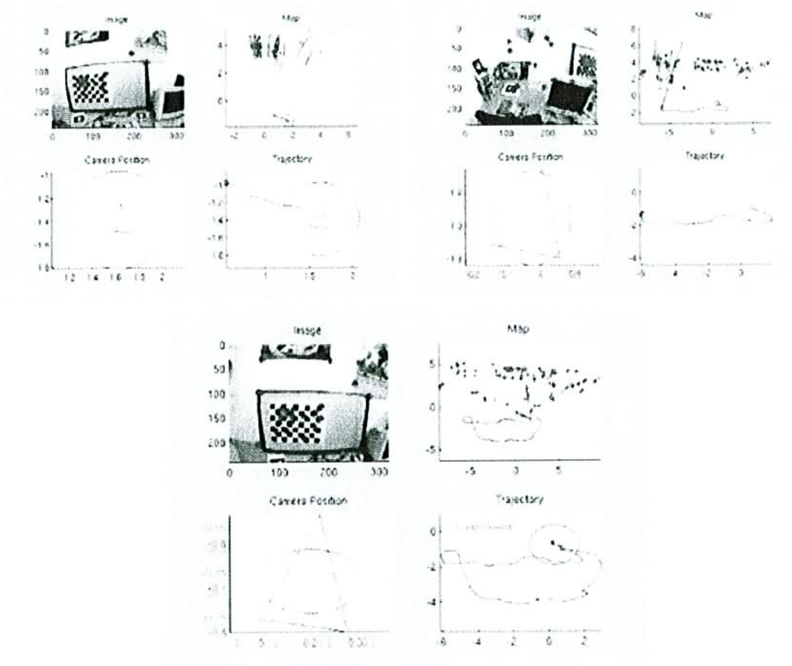


Fig. 3. Map reconstruction using Inverse Scaling parametrization in a real indoor environment.

5. D. Marzorati, M. Matteucci, and D. G. Sorrenti. Particle-based sensor modeling for 3d-vision slam. In *Proc. of IEEE Int. Conf. on Robotics and Automation*, 2007.
6. D. Marzorati, D. G. Sorrenti, and M. Matteucci. 3d-6dof hierarchical slam with trinocular vision. In *Proc. of ECMR Conf.*, 2005.
7. J. Montiel, J. Civera, and A. Davison. Unified inverse depth parametrization for monocular slam. In *Proc. of Robotics: Science and Systems*, 2006.
8. S. Thrun, W. Burgard, and D. Fox. *Probabilistic Robotics*. The MIT Press, September 2005.