

M. Schenderlein / K. Debes / A. Koenig / H.-M. Gross

Appearance-based Visual Localisation in Outdoor Environments with an Omnidirectional Camera

1 Introduction

After achieving many goals in indoor robot navigation the focus of research is shifting to outdoor applications. This poses a new challenge since outdoor environments do not only lack the presence of mainly straight and close in range walls as can be found in indoor office environments but also come up with more dynamic lighting conditions. Thus localising the robot in its environment becomes a greater challenge. On the other hand robots can have the advantage of an additional sensory input - the Global Positioning System (GPS) receiver - which measures a global position of the robot and is not applicable indoors. Though it has a bounded error, the GPS position estimate can not be exclusively used for robot localisation since it is prone to jumps in precision and might include an offset over longer periods of time because of multipath reflections from buildings or trees. Therefore we use a visual localisation system which was made robust to changes in illumination and sensor occlusion. Our robot platform is a three-wheeled vehicle equipped with an omnidirectional camera as well as a low-cost GPS receiver, a laser range finder, sonar sensors and an inertial odometry measurement unit (see Figure 2(a)).

2 Related Work

This section provides a brief review of the literature in visual outdoor localisation. The proposed approaches can be divided into ones dealing with relative localization and ones dealing with absolute localisation. Relative localisation means deriving the current pose of the robot from the previous one by measuring or estimating the relative pose change. Optical flow techniques on monocular images have the drawback of a flat world assumption in order to get reliable movement estimations [5]. This is overcome by 3D pose estimation from tracking features (e.g. corners) in an image stream [14]. Since such a calculation needs a lot of computational power and moreover requires the camera platform to be in motion, the current focus of research in those so-called visual odometry techniques is the tracking of 3D points derived from stereo camera images. Candidate points can be found by corners [14] or SIFT features [2]. The other group of approaches deals with absolute localisation which means deriving the robots current pose from a topological visual map. Such a map consists of reference locations at which low dimensional features from belonging reference images are stored. Colour histograms are used as feature vectors describing the

colour distribution in the image and are combined with nearest neighbour searches to derive the possible location of the query image [3]. The image pixel intensity gradients pose a feature which is robust to illumination changes [4]. Local image structure is also examined by integral invariant features [16] or SIFT features [17], [16]. Features that go towards object recognition are region based object classification [12] and edges of buildings [7]. Simultaneous Localisation and Mapping (SLAM) in the visual domain in large scale outdoor environments is yet tackled by only a few approaches. Pose estimation by means of a stereo camera and loop closing detection with the help of gradient features is done in [11] to solve the SLAM problem. Separate loop closing detection using SIFT features is done in [13]. Another SLAM approach is the application of a relaxation algorithm on the topological map based on image feature and odometry relations [1]. The visual odometry techniques for relative robot localisation have the advantage to be independent from a map and can therefore localise the robot immediately. Their drawback is the fact that they have no global reference and thus accumulate the estimation error. By contrast using a visual map provides this global framework. Such a map has to be build beforehand which is not always possible though. SLAM approaches are meant to solve this drawback and are therefore a promising future technique for visual outdoor localisation.

3 Approach

We propose the use of appearance-based features from panoramic images in combination with a probabilistic framework as a solution to the localisation problem. The full localisation process consists of two phases, a mapping phase and a localisation phase, which are lined out below.

First, in the mapping phase, a topological map of visual representations of the environment is built by manually acquiring the image data of the environment in conjunction with the associated pose $\underline{x} = (x, y, \varphi)$ of the robot. This kind of map stores the information in a few distinct reference places, which is especially useful in large outdoor environments. The uncertainty in position of the robot caused by erroneous odometry information is integrated out by means of a Bayesian filter [6] combining the odometry information and the GPS measurements. For further details see [15]. The mapping process presented in this work still utilises the GPS sensor despite its mentioned shortcomings. This is due to ongoing research in this field. The accuracy of acquired maps therefore has to be judged by a human inspector. Inappropriate maps possibly have to be rejected and acquired again.

For fast and noise-reduced computation the image data is stored as low dimensional representations. At this point the use of the omnidirectional camera has the advantage that only one image per reference position is needed to gather visual information from all view directions. Several test runs showed that structure-based image features are the best choice

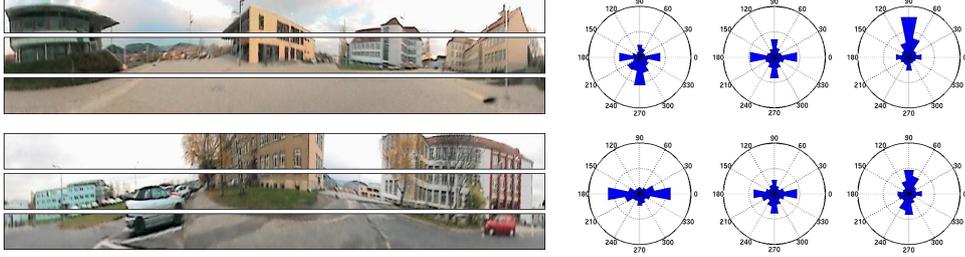


Figure 1: Two omnidirectional scenes of the campus divided into three horizontal segments and the corresponding 16-bin gradient orientation histograms left to right for the top to bottom segments.

considering invariance to different illumination settings and general capability to distinguish adjacent as well as relocate mapped places. Hence we decided to use the gradient orientation histogram as image feature for our purpose. To calculate the feature vector first the omnidirectional image is converted into a panoramic image. Of that panoramic image the approximate first order gradients G_x, G_y are calculated from the pixel intensities by means of Sobel edge filtering. Each pixel contributes to the histogram by its gradient orientation. We found a number of 16 orientation classes to give the best results. To overcome the need of a threshold for very low gradients they are weighted by their magnitude $M = \sqrt{G_x^2 + G_y^2}$ as in [10]. For the feature vector \underline{h} to be more distinctive several histograms are computed from horizontal image segments (see Figure 1) and concatenated. We found a number of six segments to give the best results. Two feature vectors m and n are compared by calculating the statistical χ^2 distances

$$d_{\chi^2}^{(i)}(\underline{m}^{(i)}, \underline{n}^{(i)}) = \sum_{k=1}^b \frac{(\underline{m}_k^{(i)} - \underline{n}_k^{(i)})^2}{(\underline{m}_k^{(i)} + \underline{n}_k^{(i)})} \quad (1)$$

for all segment histograms indexed by i with b histogram bins. In order to be robust to partial image occlusions by e.g. persons, only the three smallest distances $d_{\chi^2}^{(i)}$ are summed up to the overall distance d_{χ^2} .

In the localisation phase the robot needs to estimate its pose \underline{x} by comparing the current visual sensory input with the acquired map. The visual localisation algorithm applies a Bayesian filter variant, namely a standard Monte-Carlo-Localisation (MCL) method [6], for estimating the robots pose via the current sensor input of the camera. This is accomplished by approximating the probability distribution of the pose by a number of discrete samples. Each of those samples is moved by a motion model derived from the robot vehicle dynamics given the last movement of the robot. To adapt the distribution according to the sensor input, for every sample the closest reference feature vector \underline{h}_r at the samples position in the map is compared to the feature vector \underline{h}_c calculated of the current sensor input. In the resampling step samples whose distance $d_{\chi^2}(\underline{h}_r, \underline{h}_c)$ is small are more likely to be kept and

duplicated as opposed to the remaining ones which are likely to be erased. Eventually the pose estimation is calculated as the mean of the pose distribution of all samples.

4 Results

To examine the performance of our localisation system we measured the precision at measuring points arranged as markers at known positions on the ground. Their relative positions were determined by tape measure and best fitted to the GPS measurements. Two places on the campus of the TU Ilmenau with an area of 60 m×18 m and 90 m×35 m (see Fig. 2(b)) were investigated. First a topological visual map was build at each of those two places on a sunny as well as a cloudy day. We found a minimum distance of the maps reference points of 2 m to be sufficient. Also in each of the four cases a test run of about 1.5 km length was performed which included several loops. In those test runs we did not stick to the paths used in the mapping stage, but arbitrarily passed the measuring points. At those points we were able to calculate the localisation error. Since our robot is not equipped with a compass we used the GPS as best available ground truth for the robots heading direction. The robots pose was estimated for all pairs of test runs and maps at each place. The results are lined out in table 1.

Test run	Map	Position				Heading			
		Place A		Place B		Place A		Place B	
		μ	σ	μ	σ	μ	σ	μ	σ
sunny	sunny	2.77	1.62	3.14	1.73	18.07	14.29	10.30	7.55
cloudy	cloudy	3.48	1.77	3.62	2.45	13.79	10.86	12.40	9.57
sunny	cloudy	3.86	2.77	3.95	2.29	20.60	15.28	12.11	9.11
cloudy	sunny	3.65	3.24	4.43	2.56	14.92	13.04	12.35	12.47

Table 1: Mean difference and standard deviation of the position estimation in meters and the heading direction estimation in degrees.



Figure 2: (a) The outdoor robot MILVA. (b) Places A and B on the campus of the TU Ilmenau where the performance examinations were carried out (magenta regions).

The results show that there is no big difference in estimation precision between the test runs performed on maps with equal illumination conditions and those with different conditions.

This means the localisation system is robust to changing environment illumination caused by different day times and weather situations. The gradient orientation histogram is robust up to 50% of image occlusion. The GPS/odometry integration achieved a mean position estimation error of 2.54 m and a mean heading direction estimation error of 10.89 degrees. Figure 3 shows estimated trajectories of the visual MCL and the GPS/odometry integration.

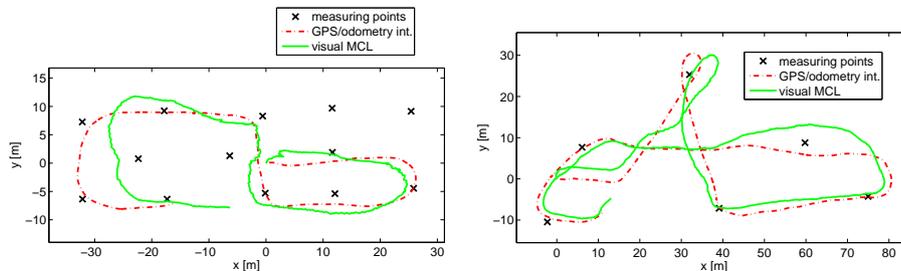


Figure 3: Parts of the estimated trajectories of the visual MCL and the GPS/odometry integration.

5 Conclusion

We presented a visual localisation framework able to localise our robot on the campus of the TU Ilmenau. It is robust to changes in environment illumination and image occlusions. At the moment its localisation precision is lower than the one of the GPS system used for map building. Considering the precision of the visual system is only as good as the map it relies on, the results are promising though. Moreover has the visual localisation the advantage that the map serves as a stable observation reference, with extreme image distortions causing only a temporal uncertainty of all particles which would shift the belief to the odometry. By contrast jumps in the GPS data which are possible for a longer period of time cause the pose estimation to drift to the erroneous measurements. This especially bears problems in the context of robot navigation. As mentioned above for map building this way of pose estimation is still useful since maps with extreme pose outliers can be identified manually after the mapping process and reacquired. Future investigations will have to tackle a gain in heading direction estimation precision which is likely to improve the overall performance as well.

References

- [1] H. Andreasson, T. Duckett and A. Lilienthal, “Mini-SLAM: Minimalistic Visual SLAM in Large-Scale Environments Based on a New Interpretation of Image Similarity”, *Proc. of the 2007 IEEE International Conference on Robotics and Automation (ICRA 2007)*, pp. 4096-4101, 2007
- [2] T. D. Barfoot, “Online Visual Motion Estimation using FastSLAM with SIFT Features”, *Proc. of the International Conference on Robotics and Intelligent Systems (IROS 2005)*, pp. 579-585, 2005

- [3] P. Blaer and P. K. Allen, "Topological mobile robot localization using fast vision techniques", *Proc. of the IEEE International Conference on Robotics and Automation (ICRA 2002)*, Vol. 1, pp. 1031-1036, 2002
- [4] D. C. Bradley, R. Patel, N. Vandapel and S. M. Thayer, "Real-Time Image-Based Topological Localization in Large Outdoor Environments", *Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2005)*, pp. 3062-3069, 2005
- [5] P. Corke, D. Strelow and S. Singh, "Omnidirectional visual odometry for a planetary rover", *Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2004)*, Vol. 4, pp. 4007-4012, 2004
- [6] F. Dellaert, D. Fox, W. Burgard and S. Thrun, "Monte Carlo localization for mobile robots", *Proc. of the IEEE International Conference on Robotics and Automation (ICRA 1999)*, Vol. 2, pp. 1322-1328, 1999
- [7] A. Georgiev and P. K. Allen, "Localization Methods for a Mobile Robot in Urban Environments", *IEEE Transactions on Robotics*, Vol. 20(5), pp. 851-864, 2004
- [8] H.-M. Gross, A. Koenig, Ch. Schroeter and H.-J. Boehme, "Omnivision-based Probabilistic Self-localization for a Mobile Shopping Assistant Continued", *Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003)*, pp. 1505-1511, 2003
- [9] H.-M. Gross and A. Koenig, "Robust Omniview-based Probabilistic Self-Localization for Mobile Robots in Large Maze-like Environments", *Proc. of the 17th Int. Conference on Pattern Recognition (ICPR 2004)*, Vol. 3, pp. 266-269, 2004
- [10] J. Kosecka and F. Li, "Vision based topological markov localization", *Proc. of the IEEE International Conference on Robotics and Automation (ICRA 2004)*, Vol. 2, pp. 1481-1486, 2004
- [11] T. Lemaire and S. Lacroix, "SLAM with Panoramic Vision", *Journal of Field Robotics (Special Issue on SLAM in the Field)*, Vol. 24(1-2), pp. 91-111, 2007
- [12] H. Morita, M. Hild, J. Miura and Y. Shirai, "Panoramic View-Based Navigation in Outdoor Environments Based on Support Vector Learning", *Proc. of the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2006)*, pp. 2302-2307, 2006
- [13] P. Newman, D. Cole and K. Ho, "Outdoor SLAM using visual appearance and laser ranging", *Proc. of the IEEE International Conference on Robotics and Automation (ICRA 2006)*, pp. 1180-1187, 2007
- [14] D. Nister, O. Naroditsky and J. Bergen, "Visual Odometry for Ground Vehicle Applications", *Journal of Field Robotics*, Vol. 23(1), pp. 3-20, 2006
- [15] M. Schenderlein, "Outdoor-Roboternavigation und Lokalisation in interaktiv erstellten topologischen Graphen mit Omnidivision-Attributierung unter Einsatz probabilistischer Zustandsschaetzer", *Diploma Thesis*, 2007
- [16] C. Weiss, A. Masselli, H. Tamimi and A. Zell, "Fast Outdoor Robot Localization Using Integral Invariants", *Proc. of the 5th International Conference on Computer Vision Systems (ICVS 2007)*, 2007
- [17] W. Zhang And J. Kosecka, "Image Based Localization in Urban Environments", *Third International Symposium on 3D Data Processing, Visualization and Transmission*, pp. 33-40, 2006

Author Information:

M. Schenderlein

Dr. K. Debes, Dipl.-Inf. A. Koenig, Prof. Dr. H.-M. Gross

Department of Neuroinformatics and Cognitive Robotics

Technische Universitaet Ilmenau

Fakultaet fuer Informatik und Automatisierung

G.-Kirchhoff-Str. 2 (Bionikgebäude)

98693 Ilmenau

Tel: +49-(0)3677-69-2770 (K. Debes), Fax: +49-(0)3677-69-1665

E-mail: klaus.debes@tu-ilmenau.de