On the Robustness of Visual Homing under Landmark Uncertainty [1]

Derik Schroeter and Paul Newman
Robotics Research Group, Dept. Engineering Science,
Oxford University, Parks Road, Oxford, OX1 3PJ, United Kingdom
{ds,pnewman}@robots.ox.ac.uk

Abstract. Although a large body of literature exists on methods for landmark-based navigation of mobile robots, little has been said about their robustness in the presence of changing and wrong landmark correspondences as well as erroneous compass measurements. To this end, we investigate the practical implications of different compass-dependent approaches to *2D Bearing-Only Navigation (2D BON)* under such conditions, and we suggest measures to quantify their robustness. The analysis is based on the notion of the *Homing Vector Field (HVF)*, and carried out in simulation, since complete HVFs can not be determined otherwise. However, HVFs provide a generic framework that enables the sound comparison of different 2D BON methods. In addition, we propose a method that is based on the approximation of the gradient of an implicit objective function, which constitutes a gradient descent approach. We present practical results that show the applicability of our approach using natural visual landmarks, omni-directional vision and a compass.

Keywords. visual homing, topological navigation, omnivision

1. Introduction

Enabling mobile platforms to autonomously navigate around their environment is still one of the major challenges in mobile robotics. Although much progress has been made for constrained settings, indoors [4] and outdoors [13,23], the problem remains widely unsolved for more complex scenarios. It is an open question, for example, what kind of perceptual clues are adequate for local navigation, i.e. when moving from one place to another along a given path. For many indoor environments, 2D laser range data seems to be sufficient, although additional sensors are necessary for robust obstacle detection [6]. In outdoors, position measurements from GPS have been successfully utilised together with image and laser range data [13,23]. However, relying primarily on GPS is delicate, since it is not always available, and multi-path GPS signals cause wrong measurements with high confidence. This is particularly true for urban environments, where in addition, the low precision of standard GPS impedes navigation in confined places.

As a consequence, much effort has been put into developing algorithms that solely rely on image data from cameras, e.g. [2,5,8,10,14,15,16]. A common approach, also deployed in the work presented here, is to associate locations in the environment with place descriptors. Based on correspondences between the current and a reference view, the task is to enable the robot to move towards the reference place. For mobile robot navigation the positioning problem is commonly considered to be 2D. That means, the related algorithms determine 2D directions, which when followed, move the robot towards a reference place [2,10,12,15]. A large body of literature covers the interaction of perception and control as well as the stability of the respective control laws [10,18,24]. In [9] Kak gives a comprehensive survey of vision-based navigation and map building. However, a detailed discussion of motion control is beyond the scope of this paper. Here we assume a control algorithm that adjusts the orientation of a mobile robot given a homing direction,

[1] The work reported in this paper was funded by the Systems Engineering for Autonomous Systems (SEAS) Defence Technology Centre (DTC) established by the UK Ministry of Defence.

and that under perfect circumstances it would converge for all the considered methods[2]. The robot moves with a constant translational velocity.

Our work is motivated by the idea of non-metric topological navigation, that is, enabling mobile robots to perform place traversal in dynamic urban environments using visual clues. Assuming distances between places to be larger than a metre, induces the problem of wide baseline matching, which causes landmarks to be unreliable and time-varying. In addition, dynamic objects frequently cause landmarks to be occluded, which together makes establishing large numbers of reliable landmark correspondences a difficult task. To this end, we investigate approaches to *2D Bearing-Only Navigation*, which potentially work with very few (<20) and changing landmark correspondences. The respective algorithms use only the 2D bearings of landmarks and a compass. They were mainly inspired by studies of insect behaviour, e.g. bees [7,21] and ants [19]. In a seminal paper, Cartwright and Collett [7] proposed the "snapshot" model, where a single-line 360 degree view describes the reference pose. Different schemes have been investigated as to how the two views can be compared to give the homing direction, e.g. [10,12,15,20,22]. In [17] Mallot and Franz give a comprehensive survey.

A large body of established literature deals with the robust estimation of the relative pose between two or more cameras by means of multiple view geometry [11]. Some recent approaches utilise the epipolar geometry between the current and the reference view either by means of the position of the epipoles in the images [16] or by determining the relative pose directly [14]. However, our empirical experience was that to obtain the performance we sought, geometric methods required more reliable view-to-view correspondences than frequently available. We decided to investigate, if we could replace or augment our visual geometry implementation with something that entirely eschewed the metric point of view and demanded fewer stable multi-view feature correspondences.

In this work we present a comparison of different compass-dependent approaches to *2D Bearing-Only Navigation* with respect to realistic settings (Sec. 4). This is to investigate the practical implications of such methods, and we suggest measures to quantify their robustness. The analysis is based on the notion of the *Homing Vector Field (HVF)*, which can only be fully evaluated in simulation. However, to the best of our knowledge, such an analysis has not been published before. A brief survey of related methods is given in Sec. 2. In addition, we propose a method that is based on the approximation of the gradient of an implicit objective function (Sec. 3). We present preliminary practical results that show the applicability of our approach in Sec. 5. Finally, Sec. 6 summarises the conclusions and discusses directions of future work.

2. Brief Survey of Methods for 2D-Bearing-Only Navigation

Following the navigation hierarchy as presented by Mallot and Franz [17], only methods of the category *Guidance* are considered in this work. It is important to note at this point, that although their survey is impressively comprehensive, its focus is on the implementation of biologically inspired methods rather than their robustness. In the majority of cases, implementations on mobile robots were used to show the plausibility of the suggested models using artificial landmarks [7,10,15]. Little has been said about the robustness of these approaches in the presence of changing correspondences due to occlusions and dynamic changes, wrong correspondences as well as compass measurement errors. However, if we want to apply these methods in real scenarios, where a robot moves autonomously in an urban environment, robustness in the above sense is a crucial measure for the applicability of any of the given approaches.

[2]Note that stability proofs exist for several such control algorithms and homing behaviours under perfect conditions, see for example [10,18,24]. Our focus, however, is the robustness under landmark uncertainty.

Bekris etal. [2] proposed a purely geometric approach and presented results using a real robot in indoor environments. Although the method is attractive because of not requiring a compass, the paper suggests that it would not work in the presence of changing or wrong correspondences. Other approaches incorporate the size of landmarks, e.g. [3], which assume that objects are approximately cylindrical and their left and right boundaries can be detected. However, these methods are not applicable to the scenario investigated in this paper, which considers 2D point landmarks. Chen et al. [8], propose a qualitative navigation scheme following the teach-and-replay paradigm. Since this approach does not explicitly generate a homing direction, it does not fit into the class of methods investigated here. The following approaches to *2D Bearing-Only Navigation* will be considered in this work:

Online-ALV. The *Average Landmark Vector* (ALV) [15] is computed from the landmark vectors v_{lm}^i that point from the robot pose to landmark i. Given N landmarks, the ALV a is defined as the averaged sum over all landmark vectors: $a = \frac{1}{N}\sum_{i=1}^{N} v_{lm}^i$. The only value stored to describe the reference pose p_{ref} is the respective a_{ref}. For new poses p_{new} in the vicinity of p_{ref} the corresponding a_{new} is computed as above. The homing direction v_{hom} is then given as the difference of the two ALVs: $v_{hom} = a_{new} - a_{ref}$. The ALV model is based on the notion that all landmarks that were visible from p_{ref} are also visible from p_{new}, which implies perfect landmark matching. Not surprisingly, the classic ALV approach performs poorly or not at all in the presence of changing or wrong correspondences. Therefore it is excluded from the following comparison, and a modified online-version of this approach is suggested and used instead. For *Online-ALV* all landmark vectors (or their bearings) are stored together with descriptors to represent p_{ref}. The descriptors are used to establish landmark correspondences between p_{ref} and p_{new}. The homing vector is determined as before, except that both a_{ref} and a_{new} are explicitly computed using only the current correspondences.

Predictive Search Based on the Equal-Distance-Assumption - PredSearchEDA [10]. Assuming that all landmarks have the same distance to p_{new}, new bearings as observed from a pose close to p_{new} can be predicted. This prediction is then compared to the bearings stored for p_{ref}. Practically this means, that the range of possible directions, i.e. 0-360 degree is discretised, and bearings are predicted for small displacement (10-40 cm) in those directions. The sum of bearing differences for landmark correspondences (or the sum of squares thereof) gives a measure of how well the predicted views match with the reference view. The direction that minimises this measure gives the homing vector.

3. Visual Homing using the Direct Gradient Approximation

Considering Fig. 1 (left), the angles α_T^i denote the bearings for a particular landmark i as observed from the reference or target position (subscript T). Accordingly, α_R^i refers to the bearing for the same landmark viewed from the current robot position. Given a particular landmark configuration and target pose, the bearings α_T^i are constant. The goal is to move a robot such that the bearing differences of corresponding landmarks are minimised. Consequently, a natural choice for the respective objective function is the sum of the squared differences between these bearings for all N landmarks:

$$F(\alpha_R^i, i = 1..N) = \sum_{i=1}^{N} (\alpha_T^i - \alpha_R^i)^2 \tag{1}$$

Since in practice this function can only be evaluated for the current position, it can be seen as an implicit objective function that is underlying the problem of *2D-Bearing-Only Navigation* in its above formulation. Fig. 1 (right) depicts an example. Assuming that F is convex within the vicinity of the reference pose p_{ref}, following the gradient of F leads

the robot towards p_{ref}. Let (x_R, y_R) and (x_L^i, y_L^i) be the positions of the robot and the i'th landmark, respectively. Then α_R^i can be determined as follows:

$$\alpha_R^i = \arctan(\frac{y_L^i - y_R}{x_L^i - x_R}) \quad (x_L^i, y_L^i) = const, i = 1..N$$

Thus, for a fixed scenario, the implicit objective function F in Cartesian coordinates only depends on the position of the robot:

$$F(x_R, y_R) = \sum_{i=1}^{N} \left[\alpha_T^i - \arctan\left(\frac{y_L^i - y_R}{x_L^i - x_R}\right) \right]^2 \tag{2}$$

The Jacobian of the objective function $F(x_R, y_R)$ from Eq. 2 is given by:

$$\begin{pmatrix} \frac{\partial F(x_R,y_R)}{\partial x_R} \\ \frac{\partial F(x_R,y_R)}{\partial y_R} \end{pmatrix} = f(\alpha_R^i, i = 1..N) = \sum_{i=1}^{N} \left[\frac{2\Delta\alpha^i}{d_{RL}^i} \cdot \begin{pmatrix} -\sin(\alpha_R^i) \\ \cos(\alpha_R^i) \end{pmatrix} \right] \quad \Delta\alpha^i = \alpha_T^i - \alpha_R^i \tag{3}$$

Since the distances d_{RL}^i between the current and the reference pose are unknown, the gradient of F at a certain position (x_R, y_R) can only be approximated. A common approach [10], also applied here, is to assume that all landmarks have the same distance to the current position, i.e. $d_{RL}^i = const \ \forall i$. This is also referred to as the *Equal-Distance-Assumption (EDA)*. We suggest three variants of how to incorporate the angle difference term in Eq. 3:

Angle Weighting (*Gradient-AW*): Uses Eq. 3 directly, but without the $2/d_{RL}^i$ term. This approximation is the closest to the analytical gradient after applying EDA.

No Weighting (*Gradient-NW*): Instead of $\Delta\alpha^i$ only its sign is considered. The rationale is that wrong correspondences are likely to cause large angle differences, thus, biasing the overall homing vector towards a wrong direction. Note that this approximation would also arise, if the sum of squares in F is replaced with the sum of absolute differences.

Kernel Weighting (*Gradient-KW*): Applies a simple non-linear kernel to determine weights w_i from the angle differences $\Delta\alpha^i$, which are then used to weight the vector contributions. The kernel (Eq. 4) maps $\Delta\alpha^i$ to $[0, 1]$, where for $\Delta\alpha^i = k$ the resulting weight is 0.5. In our evaluation k was set to 2 degrees. This variant presents a trade-off between the other two. It still incorporates weights that are proportional to $\Delta\alpha^i$, but it limits the impact of large angle differences from wrong correspondences. The kernel function is given by:

$$w_i = \frac{(\Delta\alpha^i)^2}{k^2 + (\Delta\alpha^i)^2} \qquad \Delta\alpha^i = \alpha_T^i - \alpha_R^i \tag{4}$$

Note that the methods suggested here are very similar to *PredSearchEDA* (Sec. 2). In fact, *PredSearchEDA* presents a numerical approximation of the gradient, whereas the gradient-based approaches give a direct mathematical approximation. Both assume equal distances of landmarks to the current position. However, for *PredSearchEDA* this distance has to be specified explicitly, and should somehow relate to the real settings, e.g. the expected mean. In contrast, for the gradient-based approach the actual distances are irrelevant, if they are assumed to be the same for all landmarks, see Eq. 3. In addition, for *PredSearchEDA* a small displacement has to be specified, which defines the position of the predicted view relative to the current pose. In practice, computing the approximation of the gradient from the true Jacobian is computationally more efficient (more than factor 250) than searching for the best predicted view. In addition, the time needed for *PredSearchEDA* depends linearly on the chosen angle discretisation. However, using a one-degree resolution for *PredSearchEDA*, it can be calculated in a few milliseconds on a state-of-the-art machine.

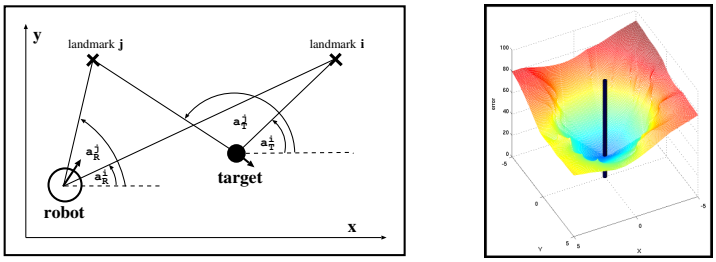


Figure 1. Left: The problem of *2D Bearing-Only Navigation*. **Right:** Example of an objective function (Eq. 1)

4. Comparison & Discussion Based on Simulation

All methods considered here produce a homing vector for the current position given landmark correspondences and bearings according to Sec. 2 and 3. In practice, at a specific point in time only the current homing vector can be determined. However, in a simulated environment, the positions of all landmarks are available in a global frame of reference, thus the homing vector can be computed for all possible robot poses. Note that the homing vector for the reference pose is not well defined. The next section describes the simulation setup. The results of the comparison are presented and discussed in Sec. 4.2.

4.1. Simulation Setup

Our analysis is based on the notion of the *Homing Vector Field (HVF)*, which is a discretisation of the space around the reference pose, where each position in the grid gets assigned the respective homing vector. Without loss of generality the reference pose is chosen to be the origin of the coordinate system. In realistic settings, at different, even very close, points in time vastly differing sets of landmarks can be observed, which in general causes the corresponding HVFs to differ. If these variations are small over time, despite wrong correspondences and compass error, the robot would still be able to approach the reference pose to some precision[3]. Thus, the performance of a given method strongly depends on the stability of the underlying HVFs. Fig. 2 shows a scenario where N landmarks are uniformly distributed around the robot, i.e. from 0 to 360 degree, with distances D_i sampled uniformly from a given interval (D_{min}, D_{max}). Although the landmark distributions differ significantly, the HVFs look very similar. To assess the difference between two HVFs, two landmark distributions are randomly generated, and the respective HVFs are computed on a given grid around the reference pose. The difference is measured as the normalised sum of the differences of two homing vectors, i.e. their angles, referring to the same position in the grid:

$$\Delta(hvf_1, hvf_2) = \frac{1}{N_x N_y} \cdot \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} |\Theta(hvf_1(i,j)) - \Theta(hvf_2(i,j)| \tag{5}$$

Where (i, j) are indexes that refer to discrete (x, y) positions in the grid. (N_x, N_y) are the number of grid points in x- and y-direction defined by the size of the area and the resolution. The values of Eq. 5 present the average angular differences of the two HVFs. The angle-function Θ is simply given by:

$$\Theta(hvf(i,j)) = \arctan\left(\frac{v_y(i,j)}{v_x(i,j)}\right) \tag{6}$$

[3]We assume here that a robust control algorithm exists for the case of perfect conditions.

(v_x, v_y) are the components of the respective homing vectors. In addition, each HVF has been compared to an "ideal" HVF using the above measure. Here "ideal" means that all homing vectors point directly towards the reference pose. If in this case the distribution of difference values for a particular setting is well below 90 degree, following any of the homing vectors will move the robot closer to the reference pose. In other words, this measure relates to the convexity of a given HVF. A third measure considers the distribution of target poses as determined from a single HVF for different settings. In particular, we consider the distance of the deduced target pose to the ground truth. The number of landmarks is fixed to 25. While this measure gives additional indications of the convexity of the HVFs, it also shows to which precision the target pose can potentially be reached. For the results presented here, the grid size is 10x10 metres with a resolution of 0.2 metres, thus, N_x=N_y=51. The distance interval for the landmarks is: D_{min}=5 metres, D_{max}=20 metres. 400 Simulations have been carried out for all combinations of: number of landmarks = {5,10,25,50,75,100}, variance of compass noise = {0,2,4,7,10,13,16} degree, percentage of wrong correspondences = {0,10,15,20,25,30,40,50} percent.

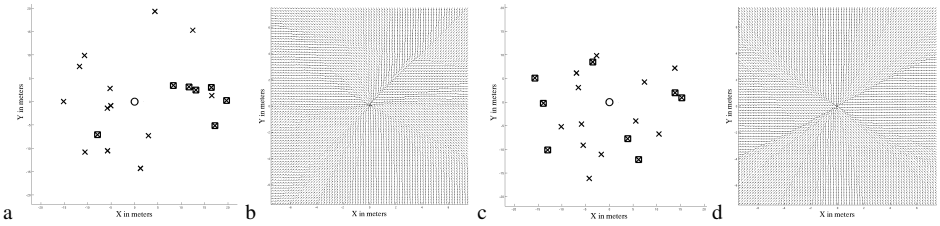


Figure 2. Two example scenarios **(a,b)** and **(c,d)** using the *Gradient-NW* method. 20 landmarks, 30 percent wrong correspondences. **a/c:** Landmark Distribution, **b/d:** *Homing Vector Field (HVF)*. Crosses refer to valid and square markers to wrong landmark correspondences. The circle in the centre marks the reference pose. Note that although the landmark distributions differ significantly, the HVFs look very similar.

4.2. Comparison and Discussion

In case of perfect conditions, i.e. static and correct correspondences and no compass error, all considered methods perform equally well. That means, they potentially steer the robot arbitrarily close to the reference pose. However, in practice, wrong feature correspondences and erroneous compass measurements can not be entirely avoided. The y-axis in Fig. 3 and 4 refers to the mean of the average angular difference in degree of two HVFs computed from two different random landmark distributions and the comparison of one such HVF to an "ideal" HVF, respectively. The vertical lines in all graphs in this section depict the standard deviation.

The first comparison considers only changing correspondence sets over time, but no compass error or wrong correspondences. Fig. 3 shows that in this case the difference between two HVFs becomes smaller with increasing number of landmark correspondences. This tendency holds true despite wrong correspondences and compass error. The comparisons with the "ideal" HVF as described in the previous section give similar results (not shown here). All angle differences are well below 90 degrees, which indicates the convexity of the underlying objective function. Together this implies that the robustness with respect to changing correspondences is comparable for all considered methods.

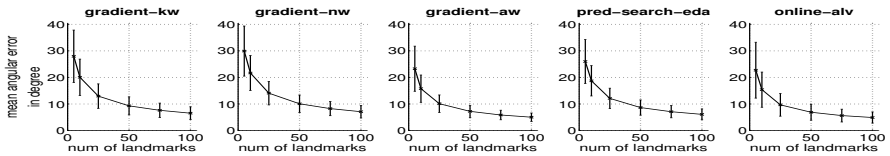


Figure 3. Pairwise comparison of HVFs for randomly generated landmark distributions - no compass error, perfect landmark correspondences.

Fig. 4-A depicts the results when introducing a zero-mean Gaussian compass error. In this case *Online-ALV* outperforms all other methods, with the *Gradient-AW* scheme being almost as good. However, the performance of the other three methods is clearly inferior in the presence of compass errors. Fig. 4-B shows the results when incorporating wrong correspondences, and assuming perfect compass measurements. It is clearly visible that the *Online-ALV* and *Gradient-AW* methods perform poorly in this case. For the latter this is due to the weighting of vector contributions per landmark pair, where some of the weights stem from wrong angle difference. The *PredSearchEDA* and *Gradient-NW* methods show less degradation in this case. For *Gradient-NW* this can be attributed to the vector contributions themselves being less affected by wrong correspondences, see Eq. eq:jac-obj-func. In fact, as long as the sign of the angle differences are correct, the respective vector contributions are perfectly valid. Note that the difference between *Gradient-KW* and *Gradient-NW* is rather small, suggesting that the kernel-weighted scheme does not provide a much better approximation of the gradient. When considering both compass error and wrong landmark correspondences, see Fig. 4-C, it becomes clear that the influence of wrong correspondences dominates the performance. In this case, *Gradient-NW* and *PredSearchEDA* outperform all other methods.

Fig. 5 shows the distributions of distances between the target pose implied by a single HVF to the true target pose. It increases with increasing compass error and percentage of wrong correspondences. Similar to the findings above the robustness is dominated by the percentage of wrong correspondences. Note that the *Gradient-NW* and *PredSearchEDA* methods show only minor degradation in the presence of wrong landmark correspondences.

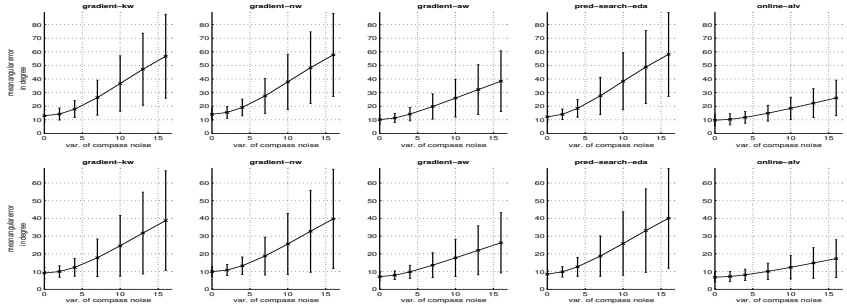
5. Experimental Results

Following the results of the previous section, we have implemented and tested the *Gradient-NW* method on our research platform "Homer" (Fig. 6, left). A crucial part of such a navigation system is the robust detection and matching of points-of-interest (POIs) preferably with a 360 degree field of view. To this end we utilise an omni-directional sensor[4], the Fast-Hessian POI detector and SURF descriptors [1]. Since in our case the camera does not experience rotation around its optical axis, the rotation invariance of the SURF descriptors has been 'switched off', which leads to more robust feature correlation. In addition, as suggested in [1], the sign of the Laplacian from the POI detection is used to speed up the matching process and improve robustness. The matching is performed using cross-correlation of the descriptor vectors. The Ladybug2 sensor comprises six cameras. Five are arranged in a ring around the vertical axis, and one is looking upwards (currently not used). The navigation system runs at about 6 frames per second on an off-the-shelf PC, which equates to about 30 images (384x512) per second.

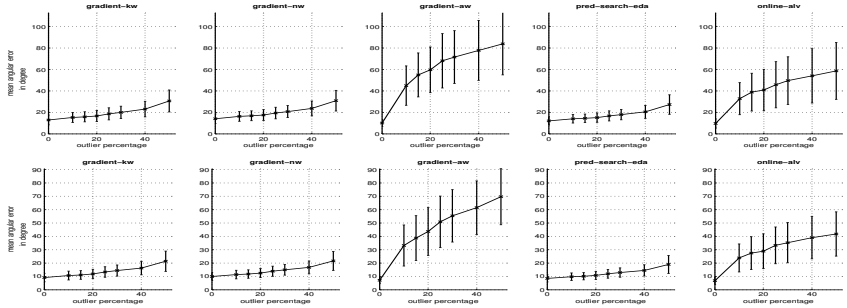
As mentioned earlier, the implemented homing behaviour provides a homing direction, which when followed, moves the robot towards the reference pose, as can be seen in Fig. 6 (right). The robot performs autonomous continuous motion using a simple controller that adapts the orientation of the robot according to the homing vector. The translational velocity is constant at about 0.2 m/s. To capture the trajectories, a 2D laser range finder was placed such that it observes the whole scenario. Afterwards, segmentation and tracking techniques have been applied to estimate the robot positions as shown in Fig. 6 (right). The resulting paths are defined by the shape of the implicit objective function F, which in turn depends on the landmark distribution. Given a large number of correct landmark correspondences, e.g. more than fifty equally distributed around the robot, the path would converge towards a straight line. In practice, however, less than twelve corre-

[4] Ladybug2 from Point Grey Research Inc.

A: variable compass error variance in degree (x-axis), perfect correspondences.



B: no compass error, variable percentage of wrong correspondences (x-axis).



C: 10 degree compass error variance, variable percentage of wrong correspondences (x-axis).

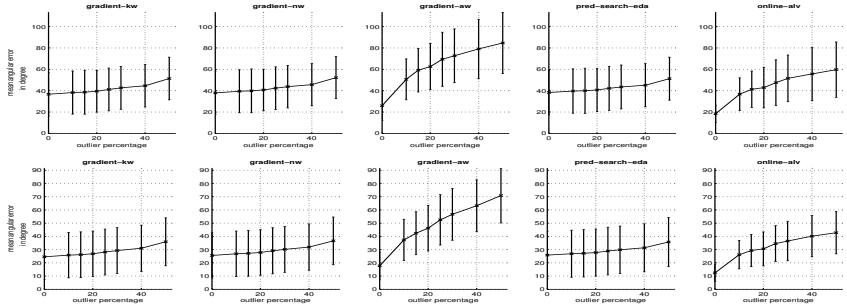


Figure 4. Comparison of HVFs for randomly generated landmark distributions (25 landmarks). The top rows for the three scenarios (A,B,C) refer to the pairwise comparison, the bottom rows to the comparison with the "ideal" HVF. The y-axis in all graphs refers to the average angular difference of the HVFs.

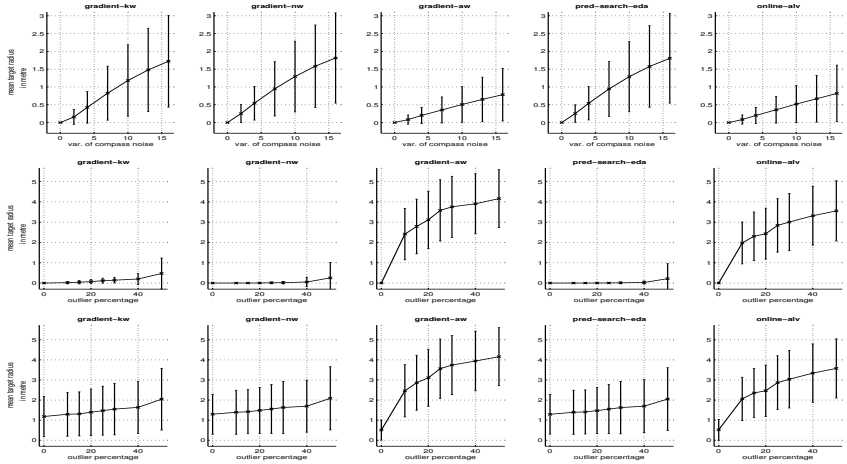


Figure 5. Distance distributions of the target pose (wrt ground truth) as determined from a single HVF (25 landmarks). **Top:** variable compass error, perfect correspondences. **Middle** and **Bottom:** variable percentage of wrong correspondences, no compass error and 10 degree compass error variance, respectively.

spondences were established in average, including varying percentages of wrong correspondences, changing correspondence sets and unsymmetric landmark distributions. As a consequence F changes over time, eventually leading the robot slightly away from the target before returning. It should be noted here that despite these difficult settings the robot successfully approaches the reference pose (centre of the pink circle) from different start poses and with different initial orientation. The maximum final distance to the reference pose is about 0.55 metre (the radius of the circle).

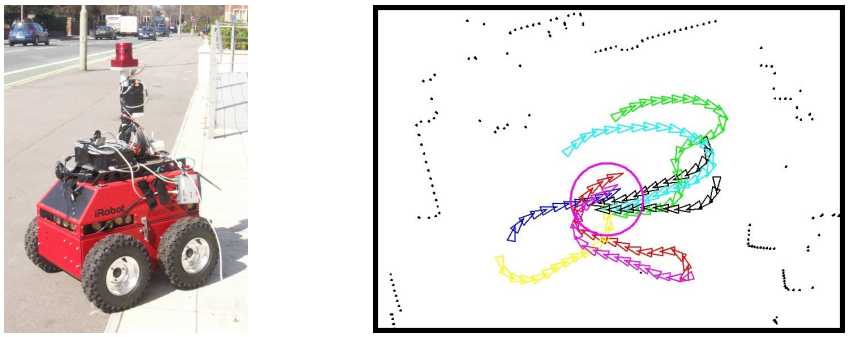


Figure 6. **Left**: Mobile Platform "Homer" - modified ATRV Junior from iRobot Corporation. The Ladybug2 omni-vision sensor is visible on the top. In addition, an Xsens inertial measurement unit and a 3.2 GHz Pentium 4 (1 GB RAM) form the main components of the presented navigation system. **Right:** Resulting trajectories from running our 2D Bearing-Only Navigation system (*Gradient-NW* method) on our mobile platform. The maximum final distance to the reference pose is about 0.55 metre (radius of the circle).

6. Conclusions & Future Work

In this paper we investigated the robustness of landmark-based, compass-dependent *Visual Homing* methods with respect to changing and wrong landmark correspondences as well as erroneous compass measurements. The considered algorithms derive homing directions from landmark correspondences. Our analysis is based on the notion of *Homing Vector Fields (HVF)*, which provide a generic framework for the comparison of such methods. Three measures have been proposed. (1) The pairwise comparison of HVFs reveals the stability of a given method over time. (2) The comparison with an 'ideal' HVF indicates the degree of convexity that can be expected. (3) The target distribution relates to convexity as well as reachability. Our analysis showed that the robustness of all methods is dominated by the presence and amount of wrong correspondences. While the *Online-ALV* and *Gradient-AW* methods are more robust, if only erroneous compass measurements are present, their performance degrades drastically in the presence of wrong correspondences. *PredSearchEDA* and *Gradient-NW* clearly outperform the other methods in the context of both wrong correspondences and compass error. In comparison to *PredSearchEDA*, the gradient approximations proposed in this paper are based on less strict assumptions. They are simpler to implement and faster to compute. Experiments on our mobile platform show the applicability of our approach using natural visual landmarks, omni-directional vision and a compass.

Given the results of the comparison, it is evident that keeping the compass error within certain bounds, e.g. by means of state estimation techniques, would improve the robustness of all the considered approaches. The same holds for the number of wrong correspondences, which requires elaborate schemes for outlier rejection and robust feature matching. A different route of continuing research is to devise methods that exhibit the same generality and robustness as the approaches considered here, yet are entirely independent of compass measurements. Our main focus, however, is to embed these methods

into a non-metric place traversal scheme that effectively constitutes appearance-based topological navigation. Eventually, the homing algorithms are at some point merely used to provide navigational clues that are to be combined with constraints from collision avoidance and dedicated navigation behaviours, e.g. path following.

References

[1] H. Bay, T. Tuytelaars, and L. V. Gool. SURF: Speeded Up Robust Features. In *Proc. of the 9th European Conference on Computer Vision (ECCV)*, 2006.

[2] K. Bekris, A. Argyros, and L. Kavraki. Angle-Based Methods for Mobile Robot Navigation:Reaching the Entire Plane. In *Proc. of IEEE Intl. Conf. on Robotics and Automation (ICRA)*, 2004.

[3] J. Bonora and D. Gallardo. Visual Homing Navigation With Two Landmarks: The Balanced Proportional Triangulation Method. In *Proc. of IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, 2006.

[4] J. Buhmann, W. Burgard, A. Cremers, D. Fox, T. Hofmann, F. Schneider, J. Strikos, and S. Thrun. The Mobile Robot RHINO. *AI Magazine*, 16(2):31–38, 1995.

[5] D. Burschka and G. Hager. Vision-based control of mobile robots. In *Proc. of IEEE Intl. Conf. on Robotics and Automation (ICRA)*, 2001.

[6] D. Burschka and G. Hager. Stereo-Based Obstacle Avoidance in Indoor Environments with Active Sensor Re-Calibration. In *Proc. of IEEE Intl. Conf. on Robotics and Automation (ICRA)*, 2002.

[7] B. Cartwright and T. Collett. Landmark learning in bees. *Journal of Comparative Physiology*, 151(4):521–543, 1983.

[8] Z. Chen and S. Birchfield. Qualitative Vision-Based Mobile Robot Navigation. In *Proc. of IEEE Intl. Conf. on Robotics and Automation (ICRA)*, 2006.

[9] G. DeSouza and A. Kak. Vision for Mobile Robot Navigation. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, volume 24, Feb. 2002.

[10] M. Franz, B. Schölkopf, H. Mallot, and H. Bülthoff. Where did I take this snapshot? Scene-based homing by image matching. *Biological Cybernetics*, 79(3):191–202, 1998.

[11] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521623049, 2000.

[12] J. Hong, X. Tan, B. Pinette, R. Weiss, and E. Riseman. Image-based homing. In *Proc. of the IEEE Intl. Conference on Robotics and Automation (ICRA)*, pages 620–625, 1991.

[13] K. Konolige, M. Agrawal, R. Bolles, C. Cowan, M. Fischler, and B. Gerkey. Outdoor Mapping and Navigation using Stereo Vision. In *Proc. of the Intl. Symposium on Experimental Robotics (ISER)*, 2006.

[14] J. Kosecka and X. Yang. Global Localization and Relative Positioning based on Scale-invariant features. In *Proc. of the 17th Intl. Conference on Pattern Recognition (ICPR)*, pages 319–322, 2004.

[15] D. Lambrinos, R. Möller, R. Pfeifer, and R. Wehner. Landmark Navigation without Snapshots: the Average Landmark Vector Model. In *Proc. Neurobiol. Conf. Göttingen*. Georg Thieme Verlag, 1998.

[16] G. López-Nicolás, C. Sagüés, J. Guerrero, D. Kragic, and P. Jensfelt. Nonholonomic epipolar visual servoing. In *Proc. of the IEEE Intl. Conference on Robotics and Automation (ICRA)*, 2006.

[17] H. Mallot and M. Franz. Biological Approaches to Spatial Representation - A Survey. In *Proc. of the 16th Intl. Joint Conference on Artificial Intelligence (IJCAI)*. Morgan Kaufmann, 1999.

[18] R. Möller. Insect visual homing strategies in a robot with analog processing. *Journal Biological Cybernetics*, 83/3:231–243, 2000.

[19] R. Möller. Modeling the landmark navigation behavior of the desert ant Cataglyphis. Technical Report IFI-AI-00.24, Artificial Intelligence Lab, Dept. Computer Science, University of Zurich, 2000.

[20] H. Neven and G. Schöner. Dynamics parametrically controlled by image correlations organize robot navigation. *Biological Cybernetics*, 75(4):293–307, 1996.

[21] A. Si, M. Srinivasan, and S. Zhang. Honeybee navigation: properties of the visually driven 'odometer'. *Journal of Experimental Biology*, 206(4):1265–1273, 2003.

[22] W. Stürzl and H. Mallot. Efficient visual homing based on Fourier transformed panoramic images. *Robotics and Autonomous Systems*, 54:300–313, 2006.

[23] S. Thrun, M. Montemerlo, H. Dahlkamp, D. Stavens, A. Aron, J. Diebel, P. Fong, J. Gale, M. Halpenny, G. Hoffmann, K. Lau, C. Oakley, M. Palatucci, V. Pratt, P. Stang, S. Strohband, C. Dupont, L.-E. Jendrossek, C. Koelen, C. Markey, C. Rummel, J. van Niekerk, E. Jensen, P. Alessandrini, G. Bradski, B. Davies, S. Ettinger, A. Kaehler, A. Nefian, and P. Mahoney. Winning the DARPA Grand Challenge. *Journal of Field Robotics*, 2006. accepted for publication.

[24] R. Wei, R. Mahony, and D. Austin. A Bearing-Only Control Law for Stable Docking of Unicycles. In *Proc. of IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, 2003.