4-2017

An Approach to Robust Homing with Stereovision

Fuqiang Fu
Fordham University

Damian Lyons
Fordham University, dlyons@fordham.edu

Robust Homing with Stereovision

Fuqiang Fang and Damian M. Lyons
Robotics & Computer Vision Laboratory
Department of Computer and Information Science
Fordham University NY 10458 USA
{ffang3,dlyons}@fordham.edu

ABSTRACT

Visual Homing is a bioinspired approach to robot navigation which can be fast and uses few assumptions. However, visual homing in a cluttered and unstructured outdoor environment offers several challenges to homing methods that have been developed for primarily indoor environments. One issue is that any current image during homing may be tilted with respect to the home image. The second is that moving through a cluttered scene during homing may cause obstacles to interfere between the home scene and location and the current scene and location. In this paper, we introduce a robust method to improve a previous developed Homing with Stereo Vision (HSV) method for visual homing. HSV adds stereo information to the image information typically used in homing resulting in improved performance. The Robust Homing with Stereo Vision (RHSV) algorithm is modified to deal with current images taken at arbitrary pitch and roll values, and to handle homing and navigation through occluding obstacles. The results for several trials comparing HSV and RHSV are presented and the future direction of this work outlined.

Keywords: visual homing, robotics, stereo vision, cluttered environment, navigation, autonomy

1.INTRODUCTION

One of the primary tasks of a mobile robot is to navigate its environment and several approaches have been developed to address this. One effective approach is for the robot to use SLAM (Thrun, Burgard, & Fox, 2005) to construct a metric map of its surroundings and to plan and then follow explicit paths (Dudek & Jenkin, 2000). This approach is not always appropriate for navigating large-area outdoor terrains due to the size requirements of a metric map, so in such cases a topological map (Dudek, Jenkin, Milios, & Wilkes, 1991), representing locations and their connectivity rather than metric details, may be more appropriate. While local navigation around a topological map location might be accomplished with the aid of a local metric map, global navigation between topological map locations can employ homing to the sensory signature of the location (Stelzer, Mair, & Suppa, 2014). Visual homing moves a robot from a start location to a previously recorded *home* location without the benefit of a map. It uses information calculated from comparing a visual image taken at a home location with an image taken at the robot's current location to generate motion commands that navigate the robot from its current location back to the location at which the home image was taken. To be effective for global navigation between landmarks, a visual homing algorithm needs to be robust with respect to navigating in unstructured outdoor terrain, including uneven surfaces and occluding objects. Because visual homing relies on comparing a home and current image, uneven terrain will result in the images being acquired/stored at different platform roll and/or pitch angles, and subsequent degradation of homing performance. Furthermore, because the home location needs to be visible for navigation to occur, intervening and occluding objects can also degrade performance.

In this paper, we take a step towards making the HSV (Homing with Stereovision (Nirmal & Lyons, 2015)) homing algorithm robust to both challenges so that it can be effectively used for global navigation in outdoor unstructured terrain. The algorithm is modified to efficiently include information from an inclinometer in image detection and comparison, and an approach to avoiding obstacles while reducing their effect on image comparisons is presented. The modified HSV algorithm, RHSV (Robust Homing with Stereovision), was implemented on a Pioneer 3-AT robot equipped with Point Gray BumbleBee2 stereo-camera and its performance was compared to HSV on a series of laboratory trials.

The remainder of the paper is organized as follows. Section 2 reviews the prior literature in this area. Section 3 briefly reviews the HSV algorithm of (Nirmal & Lyons, 2015), and then presents the RHSV algorithm modifications for uneven and cluttered terrain. Section 4 presents the performance experiments and results. These results are discussed in Section 5 and next steps considered.

2.PRIOR WORK

Visual homing is a robot navigation technique originally inspired by animal navigation studies (Cartweight & Collet, 1983) (Vardy, 2005). Since it does not require or generate a metric map or a 3D reconstruction, it has found application to visual navigation with a topological map (Dudek & Jenkin, 2000), a map which just represents places and the connectivity between them. Visual homing sidesteps the memory requirement and sensor uncertainty issues associated with maintaining a metric map, important in particular for very wide area navigation. Homing can be applied to short-range navigation with the home location in view but distant, and also in long-range navigation with the ultimate destination not in view but reached by a series of short-range navigations between places in a topological map (Argyros, Bekris, Orphanoudakis, & Kavraki, 2005) (Stelzer, Mair, & Suppa, 2014). The burgeoning area of service robots is another application area for visual homing, where robots need to operate in unstructured terrain (Moller, Krzykawski, Gerstmayr-Hillen, Fleer, & de Jong, 2013). Other recent applications for robots demand the ability to deal with unstructured and unknown terrain, such as exploration on planetary surfaces, disaster sites, and underground tunnels (Stelzer, Mair, & Suppa, 2014).

It is common to categorize visual homing methods based on the techniques they use to compare current and home images to generate the home vector. The two main categories are *holistic* and *feature-based* matching approaches. One example of the first category is the *warping method*, which warps the current image to look like the home image and calculates the movement vectors that correspond to these changes (Franz, Scholkopf, Mallot, & Bulthoff, 1998) (Zhu, Liu, & Cai, 2015). Unlike holistic methods, feature based methods identify visual landmarks. *Average landmark vector (ALV)* (Lambrinos, Moller, Labhart, Pfeifer, & Wehner, 2000) (Ramisa, Goldhoom, Aldavert, Toledo, & de Mantaras, 2011) and related methods calculate a unit motion vector by differencing the average landmark vectors in home and current images. *Correspondence methods* (Churchill & Vardy, 2008) (Pons, Huhner, Dahmen, & Mallot, 2007) (Zhu, Liu, & Cai, 2015) use features such as SIFT to identify the correspondence between current and home images to identify a motion vector. By identifying correspondences and calculating corresponding angular differences, the need for a global compass in ALV like methods (Tron & Daniilidis, 2014) is eliminated. Homing with stereovision (HSV) (Nirmal & Lyons, 2015), on which this paper is based, is a correspondence-based approach.

In bearing-only visual homing, the robot matches images, extracts the relative bearings of landmarks, and moves some distance in the direction calculated from the bearing measurements (Tron & Daniilidis, 2014). It has been shown that homing performance can be improved by adding depth information: Some researchers (Churchill & Vardy, 2008) (Liu, Pradlier, Pomerleau, & Siegwert, 2012) have used SIFT scale information (from feature matching) to estimate not only bearing but also the distance to travel in that direction. Nirmal & Lyons (Nirmal & Lyons, 2015) modified the SIFT scale based homing algorithm of (Churchill & Vardy, 2008) to leverage depth information from a stereo-camera and showed that this improved homing performance even further.

As discussed above, an important use of visual homing is in unstructured, wide area navigation – for example, traversing across outdoor, natural landscapes, remote planetary surfaces, and so forth. For these, effective feature-based matching requires knowing (or assuming) some orientation information about the home and current image, and most feature-based homing approaches assume a planar, 2D problem – that is, that the horizon is the same (usually horizontal) in both home and current images (Lim & Barnes, 2009) (Ohnishi & Imiya, 2013) (Zhu, Liu, & Cai, 2015). Many of these approaches are evaluated in indoor terrain, e.g., the TU Bielefeld visual homing database (Vardy & Moller, 2005), which makes it easy to maintain this assumption. In fact, we argue (Lyons, Barriage, & Del Signore, 2017) that outdoor terrain is a more appropriate performance evaluation for visual homing – given its strengths – and we are developing a set of stereo RGBD databases for outdoor homing (see http://goo.gl/h3pU7Q but this is a work in progress). In this paper we will look at the consequence of relaxing this 2D assumption and propose an approach for homing on uneven terrain.

A robot moving in unstructured terrain could encounter obstacles in its path. Visual homing is only affected by obstacles in that they occlude a part of the current image and/or in that they are a source of mismatches for features in the home image, and many homing algorithms ignore obstacles (Szenher, 2005). In ALV, obstacle occlusions have been shown to produce a limited form of obstacle avoidance (Hafner, 2004). In a correspondence-based approach however, these will not reliably result in the homing vector being directed around the obstacle! Argos et al. (Argyros, Bekris, Orphanoudakis, & Kavraki, 2005) in their work on long-range visual homing propose (but do not implement) an integration with reactive obstacle avoidance to handle this. Steltzer (Stelzer, Mair, & Suppa, 2014) proposes a hybrid approach for visual homing in cluttered terrain: metric mapping and navigation is carried out in the area immediately around the robot – handling uneven terrain and obstacles, and visual homing is carried out using a bearing-angle map data structure, LT-Map, for long distance navigation. Ohnishi and Imiya (Ohnishi & Imiya, 2013) present a navigation approach for a robot operating on a

ground plane with obstacles using an appearance based visual potential to avoid obstacles. However, the 'home' point is known geometrically, used to add an attractive potential to the image potential field and used to draw the robot to the goal, while optic flow is used to produce repulsive potentials that drive the object from obstacles. Liu et al. (Liu, Pradalier, Pomerleau, & Siegwart, 2012) integrate stereo-camera based obstacle avoidance and an omnidirectional camera for bearing-based homing as separate subsystems, using a cost map to constrain local navigation. Their approach is the most similar to what we will propose in using two interacting subsystems – in our case the stereovision-based homing and ultrasound-based obstacle avoidance. However, they propose the two subsystems interact only with respect to navigation, whereas obstacle occlusion can also affect the quality of the homing vector (Hafner, 2004). A key novelty in our approach is that it will also address this issue.

3.APPROACH

In this section, we present our approach, Robust Homing with Stereovision (RHSV). We begin with a brief introduction to the Homing with Stereovision (HSV) algorithm and then present our approach to the two key issues that we have argued need to be addressed: Homing on uneven terrain and homing in cluttered environments.

3.1 Homing with Stereovision(HSV)

Nirmal & Lyons (Nirmal & Lyons, 2015) present an approach to using the depth information available from a stereo-camera to improve the performance of a feature-based homing algorithm develop by Churchill and Vardy (Churchill & Vardy, 2012). Called Homing with Stereovision (HSV), this is the basis for our work in this paper.

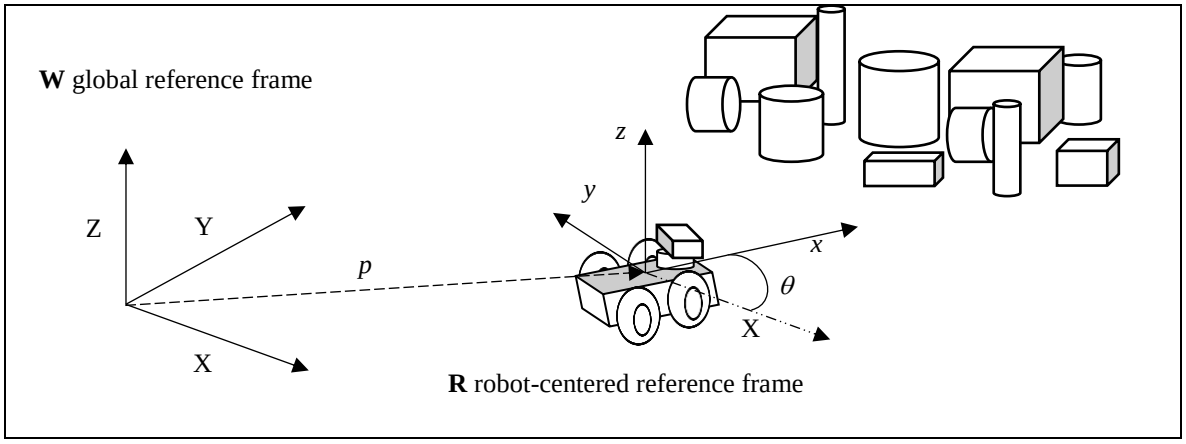

Figure 1: Summary of coordinate frames for an HSV homing mission.

Let the pose of the robot in some global reference frame **W** be $p = (X, Y) \in R^2$ (Figure 1) and orientation $\theta \in \{-\pi, \ldots, \pi\}$. The robot takes a *home image H* at the (goal) home pose (p_h, θ_h) and this image is stored. For HSV, the home image consists of both a visual image and a depth image, $H = (I_{Hv}, I_{Hd})$. The images are mutually aligned, that is, the visual pixel $I_{Hv}(u,v)$ is from a point in the scene with coordinates in a robot-centered reference frame **R**, $I_{Hd}(u, v) = (x, y, z) \in R^3$.

The robot moves from the home pose (p_h, θ_h), typically as part of an exploration or service mission, and at some later point needs to return to this pose. A homing mission therefore begins at an arbitrary start position p_s at which the home location must be, at least partially and however distantly, in view. At this, and each successive position in the homing mission, the robot takes a current image $C = (I_{Cv}, I_{Cd})$ which includes both the visual image and depth images. The current and home images are compared, and from that comparison a *homing vector* Θ calculated:

$$(\Theta_a, \Theta_d) = Compare(C, H) \quad (1)$$

This vector describes how much the robot must be rotated Θ_a and then translated Θ_d, relative to its current location, to reach the home pose (p_h, θ_h). The detail of *Compare* is the heart of the HSV algorithm and we expand on this now.

Homing algorithms typically assume an omnidirectional camera, so that the home location can be seen from any orientation. A key difference between HSV and many other homing approaches is its use of a common stereo-camera, the Point Gray BumbleBee2 with a wide-angle 3.8 mm lens, which at 66° still has a relative small horizontal field of view (FOV) when compared to an omnidirectional camera. To address this limited FOV, Nirmal & Lyons propose that each

home and current image is composed not of a single visual and depth image from the stereo camera, but by a composite image made by panning the stereo-camera across the scene.

The current image C (and home image H), for both the visual and depth image components, is captured as a sequence of 5 concatenated images: one directly in front of the robot, two images clockwise and two images counterclockwise from the front. The five images are simply concatenated, and not stitched or otherwise transformed; according to Nirmal & Lyons, this is faster and the image overlap improves the number of matches found. The height N_v of the resultant image is the same as a single image but the width N_u is five times larger. Each image has a 66º FOV and there is a 15º overlap between images. They calculate the FOV of the resultant composite visual and depth images as 5×66 - 4×15 = 270º, giving a wide, but not omnidirectional FOV.

The *compare* function in the HSV algorithm extracts SIFT features from the C and H visual images and produces a list of matched image feature locations

$$M = \{(c_m, h_m) \mid m = 0, \ldots N, c_m, h_m \in \{ 0 \ldots N_u\text{-}1, 0 \ldots N_v\text{-}1 \} \} \tag{2}$$

where c_m is the image coordinates of the SIFT feature in the current visual image I_{Cv}, and h_m the image coordinates in the home visual image I_{Hv}, for each matched SIFT feature *m*. Figure 2 (from (Lyons, Barriage, & Del Signore, 2017)) shows an example visual image comparison. The top strip of 5 images is the composite home visual image. The bottom strip of 5 images is the composite current visual image. They happen to be the same in this case; that is the robot is at the home location. Recall that the depth images and visual images are aligned so that the visual pixel in the home image $I_{Hv}(c_m)$ is from a point in the scene with robot-centered coordinates $I_{Hd}(c_m) = (x,y,z)$, for each matched feature $m = 0, \ldots N$.

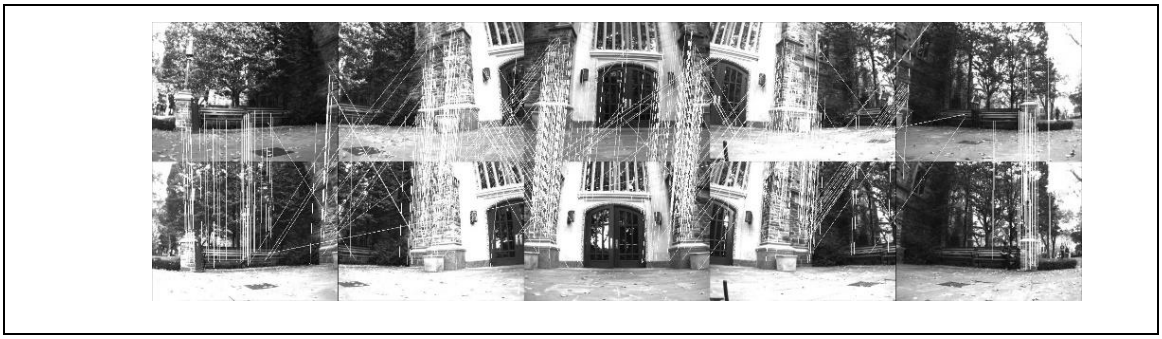

Figure 2: Example wide field of view SIFT matching of home image I_{Hv} (top) and current image I_{Cv} (bottom) with lines between matched features (from (Lyons, Barriage, & Del Signore, 2017))

The notation $c_{m,q}$ for $q \in \{x, y, z\}$ refers to the q coordinate (*with respect to* **R**) of matched feature *m* from the current depth image; $h_{m,q}$ is defined similarly for the home image. The robot-centered coordinate frame **R** is shown in Figure 1, with *x* pointing forwards and *y* to the left. The rotation component of the homing vector is calculated as the average angle difference between current and home depth images for this feature point in the *xy* plane:

$$\Theta_a = \frac{1}{|M|} \sum_{m \in M} \left(tan^{-1} \frac{h_{m,y}}{h_{m,x}} - tan^{-1} \frac{c_{m,y}}{c_{m,x}} \right) \tag{3}$$

Nirmal & Lyons calculate the distance component of the homing vector as the average difference of *x coordinates*. This produces small motions when features are at the edge of the field of view, so we have modified the measure here to be the difference in distances to the feature in home and current image, and where $sx(.)$ is the sign function on the *x* axis:

$$\Theta_d = \frac{1}{|M|} \sum_{m \in M} sx(h_m)(|h_m| - |c_m|) \tag{4}$$

$$sx(h_m) = \begin{cases} +1 & if\ h_{m,x} \geq 0 \\ -1 & else \end{cases}$$

Nirmal & Lyons use four performance measurements — angular error, positional error, number of steps, and return ratio – to evaluate the performance of HSV on a series of homing trials using a Pioneer 3AT robot. They favorably compare HSV performance to that of the *HiSS* algorithm (Churchill & Vardy, 2008), an algorithm that that uses SIFT scale information to approximate depth. They also present a convergence proof for the HSV algorithm under the assumptions of error-free SIFT matching and commanded robot motion. The HSV algorithm is shown in Algorithm 1 below.

HSV(H, Θ_{tol} , D_{tol}, g_θ , g_D)

1. *H = Capture(5)*
2. *Do:*
3. *C = Capture(5)*
4. *(Θ_a , Θ_d) = Compare(C,H)*
5. *rotate by g_a Θ_a & translate by g_d Θ_d*
6. *While ($|\Theta_a| > \Theta_{tol}$ & $|\Theta_d| > D_{tol}$)*

Algorithm 1: Homing with Stereovision Algorithm

The g_a , g_d parameters in the HSV algorithm are motion gains in the half open interval (0,1]; the Θ_{tol} , D_{tol} parameters are tolerance parameters to terminate homing. In the next subsection we address the challenge of modifying HSV to work on uneven terrain.

3.2 Homing in Uneven Terrain

On uneven terrain, the homing problem of Figure 1 is changed, as shown in Figure 3(a), and the robot coordinate frame **R** may become inclined **R**'. The angle that the (robot) *x* axis makes with the (global) X axis is the *yaw* angle and is simply the orientation θ of the robot and so we set this aside. The rotation of **R'** around the *x* axis is the *roll* angle α – a side-to-side inclination of the robot. The rotation around the *y* axis is the *pitch* angle β – a front to back inclination of the robot.

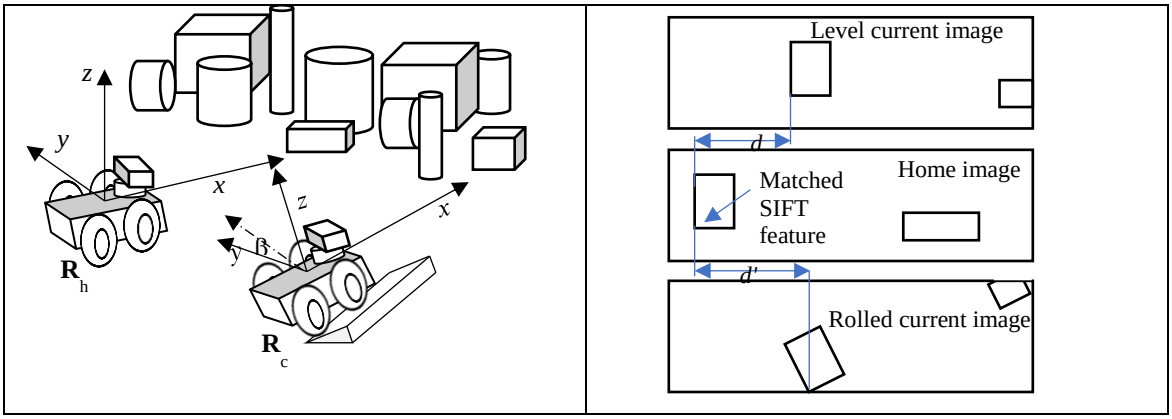

Figure 3: (a) Summary of coordinate frames on uneven terrain for an HSV homing mission. (b) Homing without knowledge of roll angle showing erroneous matched feature distance *d'* compared to correct distance *d<d'*.

The issue with homing on uneven terrain can be seen with the following example. Assume the robot has taken a home image on level terrain, i.e., with respect to $\mathbf{R}_h$, and then, later, attempts to begin homing to this location starting from a position on uneven terrain. For simplicity, let us assume only a non-zero roll angle. As show in Figure 3(b), SIFT feature matching will produce a set of matched features in home and (rolled) current image. However, when eq. (3) is applied to calculate the difference in angles in the robot coordinate frame $\mathbf{R}_c$, the (rolled with respect to $\mathbf{R}_h$) feature coordinates will be offset by a factor of *sin(β)* and *cos(β)*, leading to an underestimation of the angle. While a non-zero roll angle during homing might in the best case affect a few homing steps that happen to be on uneven terrain, and perhaps zero out over time, a non-zero roll angle when the home image is taken will produce a chronic effect.

In a similar fashion, a non-zero pitch angle will affect distance estimation, eq. (4).

To address this problem, we equip our homing platform with a TCM2 integrated compass, inclinometer, magnetometer and thermometer (from ActiveMedia Robotics) as shown in Figure 4. The roll and pitch angles can be read from the TCM2 unit at any time, and in particular, they can be read when the home image is stored and when each current image is stored. Let $\mathbf{R_h}$ be the coordinate frame in which the home image was taken, and $\mathbf{R_c}$ be the coordinate frame in which the current image is taken. On even terrain, these are related by the home vector Θ which translates and orients the robot back to its home location. On uneven terrain, the combined rotation of the roll and pitch at home and at current needs to be accounted for as well. This combined roll and pitch rotation is written as *A* in the following:

$$\boldsymbol{R}_h = \boldsymbol{R}_c \cdot A \cdot \Theta \tag{5}$$

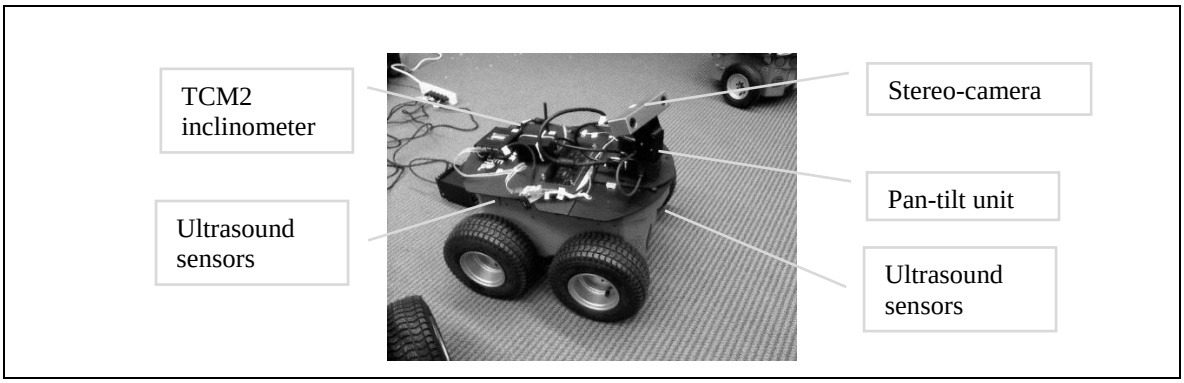

Figure 4: Pioneer 3AT with Bumbleb2 stereo-camera on Dynamic Perception Pan-Tilt unit and TCM2 inclinometer.

The homing algorithm iteratively approximates Θ. To calculate A, we need four parameters from the TCM2: the current roll angle of robot α_c , the current pitch angle of robot β_c, the home roll angle of robot α_h and the home pitch robot β_h.

The transformation matrix $RP(\alpha, \beta)$ can be calculated by:

$$RP(\alpha,\beta) = \begin{bmatrix} \cos(\beta) & -\sin(\beta) & 0 \\ \cos(\alpha)\sin(\beta) & \cos(\alpha)\cos(\beta) & -\sin(\alpha) \\ \sin(\alpha)\sin(\beta) & \sin(\alpha)\cos(\beta) & \cos(\alpha) \end{bmatrix} \tag{6}$$

And in that case, *A* can be calculated as:

$$A = RP(\alpha_c,\beta_c) \cdot RP(\alpha_h,\beta_h)^{-1} \tag{7}$$

Therefore, to compensate for uneven terrain, it is only necessary to correct each c_m in eq. (3) by rotating it by *A*. In Section 4, performance results for homing on uneven terrain are presented for the case in which A is used and in which it is not.

3.3 Homing in a Cluttered Space

When homing in an unstructured environment, there may possibly be obstacles to motion and visibility occluding the home location (Figure 5). Calculation of a correct home vector by HSV can be accomplished as long as sufficient feature points in the home image can be correctly matched in the current image. Thus, visual occlusion is not the major challenge.

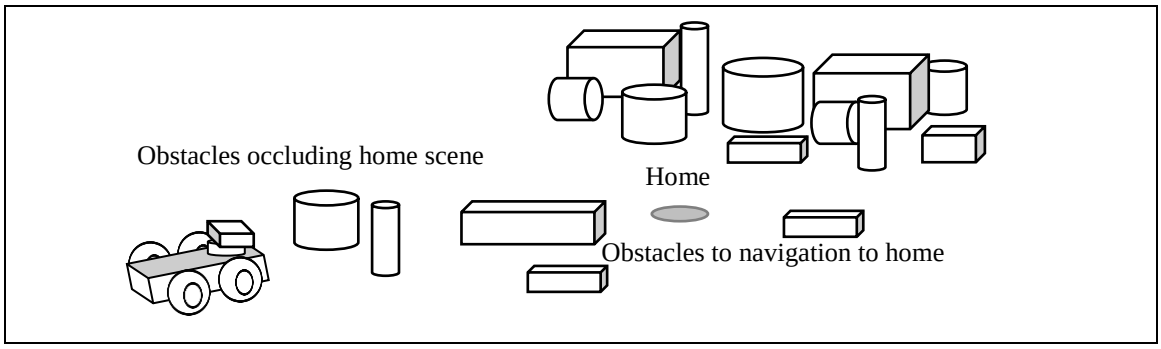

Figure 5: Homing in a cluttered environment.

The first challenge is to prevent features on obstacles from mismatching occluded home features and thereby generating poor homing vectors. A robot moving in cluttered area may find that a large portion of its FOV is of nearby obstacles. These occlude the more distant parts of the scene that may have yielded features in the home image. Occasionally the best match for such occluded features may then be features on the obstacles. The error introduced into the calculation of the home vector in eqs. (3) & (4) can drive the robot away from the home location and may even drive it towards collisions. To address this challenge, any matched features close to the robot are eliminated. A filtered match set M' is calculated as follows, where d_c is the closest feature distance allowed:

$$M' = \{ (c_m, h_m) \mid (c_m, h_m) \in M \;\&\; |c_m| \geq d_c \} \tag{8}$$

When the robot begins the homing mission, and we will refer to this as time step $t=0$, we will assume that the home location is at its most distant; we set d_c to an a-priori determined maximum obstacle distance D_c. Once the first homing vector is calculated, that is for time steps $t>0$, we have an estimate of how far the home location is from the current location, Θ_d, and the value of d_c is set based on this value. As the robot nears its home location, successive Θ_d values decrease and M' approaches the original M, achieving the result of allowing all the home features to match no matter their distance.

$$d_c = \begin{cases} D_c & t = 0 \\ g_c \Theta_d & t > 0 \end{cases} \qquad g_c \in (0,1] \tag{9}$$

The second challenge is to prevent collision in the event that the home vector does direct the robot towards an obstacle lying between its location and the home location. Our approach is to use a modified version of the Vector Field Histogram (VFH) algorithm (Borenstein & Koren, 1991). VFH constructs its histogram using a map of the active region, the region immediately around the robot. While homing does not construct or need a map, in fact the current depth image I_{Cd} contains readings from which a map could be constructed. We chose not to do this for two reasons: Since a principal advantage of homing is that it does *not* use a map, forcing it to include one (and the probabilistic reasoning to make the map more dependable) seems to weaken the advantage of using homing. Secondly, I_{Cd} is only updated sporadically, i.e., at each homing step, and seems to be a poor candidate for an obstacle avoidance sensor.

Instead we choose to directly use the depth readings from the 16 ultrasound sensors around the body of the Pioneer 3-AT robot to populate a polar histogram similar to that used in VFH. Let $s(k)$, $k \in 1 \ldots 16$, be the current depth reading from sonar sensor k, and let $a(k)$ be the angle that the central axis of sonar sensor k makes with the x axis of **R.** Let γ be the half-angle beam width for all sonar sensors. We define the polar function $h : \{-\pi, \ldots, \pi\} \rightarrow Z^+$ as follows, where D_0 is the maximum distance at which obstacles are considered:

$$h(i) = \begin{cases} s(k) - D_0 & s(k) < D_0 \; and \; i \in \{a(k) - \gamma, \ldots, a(k) + \gamma\} \\ 0 & else \end{cases} \tag{10}$$

The selection of valleys and obstacle avoidance heading is then calculated from h in a manner similar to that in VFH and we compensate for lack of information about the robot size by insisting that any navigable valley be at least $\frac{\pi}{4}$ wide and there be at least $\frac{\pi}{8}$ free (i.e., h value below the valley floor) on any side of the robot heading.

Homing is integrated with this simplified VFH obstacle avoidance method as follows: The homing vector Θ provides the goal direction and distance, and a heading is continually selected that avoids obstacles and is closest to Θ and followed until the point is reached or Θ_d has been travelled[1] (in the case that an obstacle directly occupies the goal location).

The RHSV algorithm is summarized in Algorithm 2.

RHSV(H, Θ_{tol} , D_{tol}, g_θ , g_D , g_c , D_0)

1. *H = Capture(5)*
2. *Collect* α_h β_h
3. $d_c = D_c$
4. *Do:*
5. *C = Capture(5)*
6. *Collect* α_c β_c
7. $A = RP(\alpha_c, \beta_c) \cdot RP(\alpha_h, \beta_h)^{-1}$
8. *(* Θ_a , Θ_d *) = Compare(C,H,A)*
9. *Repeat*
10. *Construct polar histogram h using* D_0
11. *Travel collision free heading closest to* $g_a \Theta_a$

[1] In fact, $g_d \Theta_d$ but we omit the gains from the Algorithm 1 here for clarity of presentation.

12. *Until distance travelled* $\geq g_d \Theta_d$
13. $d_c = g_c \Theta_d$
14. *While* ($|\Theta_d| > \Theta_{tol}$ & $| \Theta_d | > D_{tol}$)

Algorithm 2: Robust Homing with Stereovision Algorithm

4. RESULTS

Experiments were conducted to evaluate the performance of the RHSV algorithm. The platform in each case was the Pioneer 3-AT robot shown in Figure 4, equipped with BumbleBee2 stereo-camera on a Directed Perception pan-tilt base, a TCM2 inclinometer and 16 ultrasound sensors distributed around its circumference. The RHSV program was written in C++ using the Aria programming toolkit from ActiveMedia Robotics. All the experimentation was done in an indoor laboratory setting, with uneven terrain and obstacles simulated by placing material under the robot wheels or in front of the robot. Evaluations were conducted in a 4m × 4m square open area of the lab, with a level carpeted floor and indoor fluorescent lighting. Two performance evaluations experiments were carried out: an uneven terrain homing performance evaluation, and a cluttered space homing performance evaluation. In each case these were compared to the performance of the HSV algorithm. The same set of gain and tolerance values were used throughout and are shown in Table 4.1

Table 4.1: RHSV Gain and Tolerance values

Parameter	Value
Θ_{tol}	15^0
D_{tol}	250 mm
g_θ	0.9
g_D	0.9
g_c	0.5
D_0	1500 mm

4.1 Uneven terrain homing performance

The uneven terrain performance evaluation was carried out by rolling the robot by approximately 20^o during the collection of the home image, moving the robot to an arbitrary start location, and allowing homing to proceed on level terrain. Inclining the home image and allowing homing on level terrain is significant easier to carry out than a level homing image and inclining the platform during homing – the incline only needs to be done once, and it affects every homing step, since the (inclined) home image needs to be compared with the current image at every step.

Table 4.2: Performance results for homing on uneven terrain

Homing Algorithm	Home Image	Average homing error (mm)	Variance in homing error (mm^2)
HSV	No Roll	562	51,596
HSV	Rolled 20^o	1,558	666,856
RHSV	Rolled 20^o	763	310,788

Homing was first conducted on level, uncluttered terrain in an indoor laboratory using HSV for 10 trials. Notice that in level terrain with no obstacles, HSV and RHSV are the same algorithm. Homing to an inclined home image was then repeated 20 times; 10 times using HSV and 10 times using RHSV. The homing error – defined as the distance from the location at which the home image was taken to the location at which homing terminated – was measured in each case. Under the assumption that the error could be reasonably approximated as normal, images statistics were generated for HSV and RHSV. These statistics are shown in Table 4.2. The null hypothesis that HSV performs the same on even and uneven terrain was evaluated using Welch's t-test. The t-test statistic $0.005185 < 0.05$ indicated that the null hypothesis can be rejected at a 95% significance level. Clearly homing performance is reduced in the presence of uneven terrain.

The null hypothesis that RHSV and HSV have the same performance for homing in uneven terrain was tested for again using Welch's t-test. The t-test statistic 0.028355 <0.05 indicated that the null hypothesis can be rejected at a 95% significance level. The performance on uneven terrain is therefore significantly improved by the modifications we have introduced to HSV to make RHSV.

4.2 Cluttered space homing performance

To evaluate the performance of RHSV in cluttered terrain, a series of obstacles was placed between the home location and the start location (Figure 6). The obstacles were positioned in an interleaved pattern between the area of start locations and the (much smaller) area of goal locations to ensure that no clear path or view existed from start to goal. The obstacles were hardboard screens painted with a textured pattern to present strong candidates for SIFT features (and hence promote erroneous home feature matching).

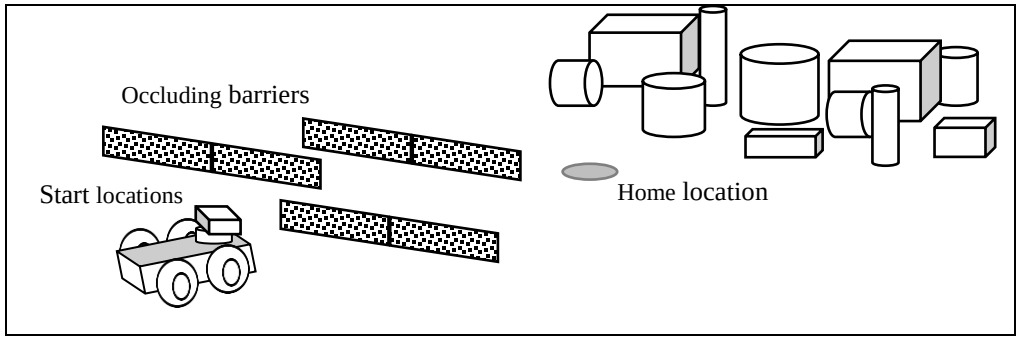

Figure 6: Scenario for evaluating RHSV performance in cluttered terrain.

The home location was collected, and the robot then navigated by hand to an arbitrary start location within the area of start locations. The robot odometry was zeroed at the start of homing, and odometry estimates logged for the duration of the homing mission. Homing was conducted 20 times; 10 times with RHSV and 10 times with HSV. The homing error, defined as before, was collected for each. Again, under the assumption that the error could be reasonably approximated as normal, images statistics were generated for HSV and RHSV. These statistics are shown in Table 4.3.

Figure 7 shows the odometry data collected from the performance trials: Figure 7(a) shows the RHSV homing tracks and Figure 7(b) shows the HSV homing tracks. Note that every track appears to have the same start location; this is because odometry is zeroed at the start of homing. Thus, although the tracks are not shown in correct relation to one another, the two graphs are the same scale, so track shape and length can be reasonably compared.

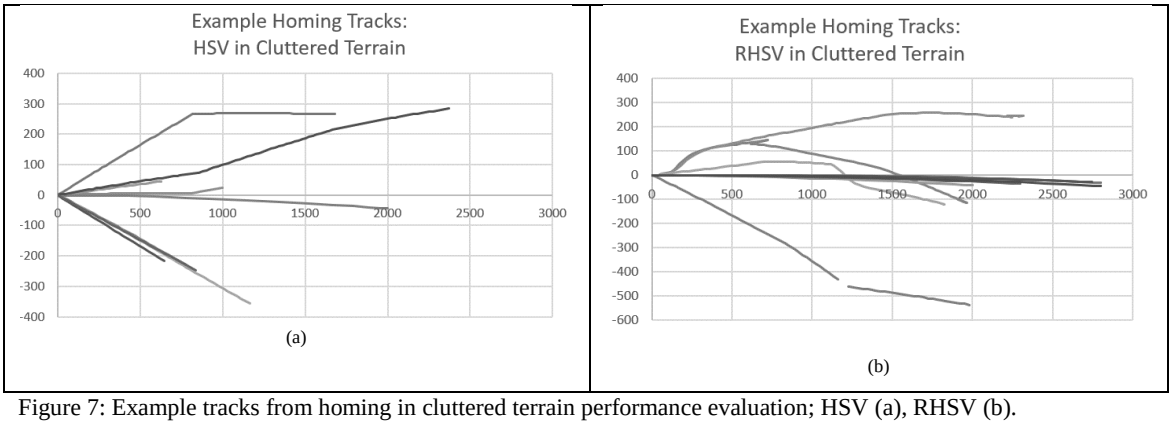

Figure 7: Example tracks from homing in cluttered terrain performance evaluation; HSV (a), RHSV (b).

The average homing error for HSV in cluttered terrain is just under four and a half times as large as that for RHSV. Indeed for many of the homing trials HSV simply collided with a barrier, sooner or later in the track. Some very few of the HSV trials did make it to the home location. All RHSV trials made it to the home location without any collisions.

Table 4.3: Performance results for homing in cluttered terrain

Homing Algorithm	Average homing error (mm)	Variance in homing error (mm^2)
HSV	1918	455,625
RHSV	427	42,436

The null hypothesis, that RHSV and HSV have the same performance for homing in cluttered terrain, was tested for using Welch's t-test. The t-test statistic 0.000171 <0.05 indicated that the null hypothesis was rejected at a 95% significance level.

Note that the RHSV performance in cluttered terrain (Table 4.3) appears to be better than HSV in uncluttered terrain (Table 4.2) – though not significantly so (t-test statistic=0.217061). We discuss this in the next section.

5. DISCUSSION

Visual homing is a navigation technique that does not require metric map information and is therefore useful for wide area navigation (Stelzer, Mair, & Suppa, 2014) where constructing and maintaining a large area metric map would be expensive or infeasible. Nonetheless, many visual homing researchers restrict their attention and experimentation to indoor, even terrain. In this paper we argue that to capitalize on the strengths of visual homing it should be developed for use in unstructured large area terrain and we take one step in this direction by extending the Homing with Stereovision (HSV) (Nirmal & Lyons, 2015) to handle two challenges in homing in unstructured, large area terrain: uneven terrain and cluttered terrain. We extend HSV to a more robust form to handle these challenges, Robust homing with Stereovision (RHSV), and we present performance evaluation results to show our extensions significantly improve homing performance in these two cases.

We evaluate homing on even terrain by measuring the performance of HSV versus RHSV in the case that the home image is inclined and all current images are not. RHSV reduces the mean homing error by more than 50% and this improvement can be shown to be statistically significant at α=0.05. However, this performance test examines the response to a tilt in the roll angle in the home image. It does not include any pitch disturbance testing, or any testing of roll or pitch disturbance during homing. Because pitch compensation is done in the same manner as roll compensation, we argue that our test is sufficient, and that while testing a pitch disturbance would increase confidence in our approach, we don't expect it would yield any new insights. Similarly, since the roll disturbance between $\mathbf{R}_c$ and $\mathbf{R}_h$ is relative; while testing a roll (or pitch) disturbance during homing would increase confidence in our approach, we again don't expect it would yield any new insights.

Homing on cluttered terrain is evaluated by a series of homing runs through interleaved, textured barriers which are designed to both block navigation, occlude the home scene and provide a distraction to SIFT matching. The homing error recorded from RHSV in this case is just under four and a half times that of HSV, a difference which is shown to be significant at α=0.05. A key novelty of our approach is not just the integration of obstacle avoidance (which others have done, e.g., (Liu, Pradlier, Pomerleau, & Siegwert, 2012)) but the relationship between homing distance calculation and SIFT feature filtering to avoid mismatches, eq. (9). However, these trials only test mission completion – they do not measure the specifics of how many features were mismatched in RHSV vs HSV due to the barriers, which could yield valuable new insights.

It is noted in the results that the homing error in cluttered terrain (Table 4.3) appears (non-significantly) to be better than HSV in uncluttered terrain. We believe the reason for this result is the 'shepherding' effect of the interleaved barriers shown in Figure 6. Any complete mission through these barriers ends up funneled by the obstacle avoidance behavior into the vicinity of the home location. In future work, we propose to control for this effect by allowing for multiple exit points from the barriers.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the efforts of Juan Ruiz in assisting with the completion of RHSV debugging and performance evaluations tests and Zhihen Huang for his assistance robot platform maintenance and upgrading.

6. REFERENCES

Argyros, A., Bekris, K., Orphanoudakis, S., & Kavraki, L. (2005). Robot Homing by Exploiting Panoramic Vision. *Autonomous Robots*(19), 7-25.

Borenstein, J., & Koren, Y. (1991). The vector field histogram-fast obstacle avoidance for mobilerobots. *7*(3).

Cartweight, B., & Collet, T. (1983). Landmark learning in bees. *Journal of Comparative Physiology*(151), 521-543.

Churchill, D., & Vardy, A. (2008). Homing in scale space. *IEEE/RSJ Conf. on Intelligent Robots and Systems (IROS).*

Churchill, D., & Vardy, A. (2012). An orientation invarient visual homing algorithm. *Journal of Intelligent and Robotics Systems, 17*(1), 3-29.

Dudek, G., & Jenkin, M. (2000). *Computational Principles of Mobile Robotics.* Cambridge: Cambridge University Press.

Dudek, G., Jenkin, M., Milios, E., & Wilkes, D. (1991). Robotic exploration as graph construction. *IEEE Trans. on Rob. & Aut., 7*(6).

Franz, M., Scholkopf, B., Mallot, M., & Bulthoff, H. (1998). Where did I take that snapshot? Scene-based homing by image matching. *Biological Cybernetics*(79), 191-202.

Hafner, V. (2004). Agent-Environment Interaction in Visual Homing. In P. R. Iida F., *Embodied Artificial Intelligence, Lecture Notes in Computer Science, vol 3139.* . Berlin, Heidelberg: Springer.

Lambrinos, D., Moller, R., Labhart, T., Pfeifer, R., & Wehner, R. (2000). Mobile robot employing insect strategies for navigation. *Robotics and Autonomous Systems*(30), 39-64.

Lim, J., & Barnes, N. (2009). Robust Visual Homing with Landmark Angles. *Robotics Science and Systems V.* Seattle .

Liu, M., Pradalier, C., Pomerleau, F., & Siegwart, R. (2012). The role of homing in visual topological navigation. *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS).*

Liu, M., Pradlier, C., Pomerleau, F., & Siegwert, R. (2012). Scale-only visual Homing from an Omnidirectional Camera. *IEEE Int. Conf. Robotics & Automation.*

Lyons, D., Barriage, B., & Del Signore, L. (2017). Effect of Field of View on Stereovision-based Visual Homing. *IEEE International Conference on Tools with AI* . Boston MA.

Moller, R., Krzykawski, M., Gerstmayr-Hillen, L., Fleer, D., & de Jong, J. (2013). Cleaning robot navigation using panoramic views and particle clouds as landmarks. *Robotics & Autnomous Systems*(Cleaning robot navigation using panoramic views and particle clouds as landmarks.), 1415-1439.

Nirmal, P., & Lyons, D. (2015). Homing With Stereovision. *Robotica , 34*(12).

Ohnishi, N., & Imiya, A. (2013). Appearance-based navigation and homing for autonomous mobile robot. (31).

Pons, J., Huhner, W., Dahmen, J., & Mallot, H. (2007). Vision-based robot homing in dynamic environments. *13th IASTED Int. Conference n Robotics and Applications.*

Ramisa, A., Goldhoom, A., Aldavert, D., Toledo, R., & de Mantaras, R. (2011). Combining invariant features and the ALV homing method for autonomous robot navigation based on panoramas. *Journal of Intelligent Robot. Sys,., 64*, 625-649.

Stelzer, A., Mair, E., & Suppa, M. (2014). Trail-Map: A scalable landmark data structure for biologically inspired range-free navigation. *IEEE Int. Conf. Robotics & Biomimetics.* Bali, Indonesia.

Szenher, M. (2005). *Navigation by Image-based Visual Homing.* Edinburgh, UK: PhD. Thesis School of Informatics, University of Edinburgh.

Thrun, S., Burgard, W., & Fox, D. (2005). *Probabilistic Robotics.* Cambridge MA: MIT Press.

Tron, R., & Daniilidis, K. (2014). An Optimization Approach to Bearing-only Visual Homing with Applications to a 2-D Unicycle Model. *IEEE Int. Conf. on Robotics & Aut.*

Vardy, A. (2005). *Biologically Plausible Methods for Robot Visual Homing.* Ph.D. Thesis. School of Computer Science, Carleton University.

Vardy, A., & Moller, R. (2005). Biologically plausible visual homing methods based on optical flow techniques. *Connection Science*(17).

Zhu, Q., Liu, C., & Cai, C. (2015). A Novel Robot Visual Homing Method Based on SIFT Features. *Sensors, 15*, 26063-26084.