EXPLOITING PANORAMIC VISION FOR BEARING-ONLY ROBOT HOMING

Kostas E. Bekris[1], Antonis A. Argyros[2], Lydia E. Kavraki[1]

[1]*Computer Science Department, Rice University,*
Houston, TX, 77005, USA
{bekris,kavraki}@cs.rice.edu

[2]*Institute of Computer Science, Foundation for Research and Technology - Hellas (FORTH),*
Heraklion, Crete, Greece
argyros@ics.forth.gr

Abstract   Omni-directional vision allows for the development of techniques for mobile robot navigation that have minimum perceptual requirements. In this work, we focus on robot navigation algorithms that do not require range information or metric maps of the environment. More specifically, we present a homing strategy that enables a robot to return to its home position after executing a long path. The proposed strategy relies on measuring the angle between pairs of features extracted from panoramic images, which can be achieved accurately and robustly. In the heart of the proposed homing strategy lies a novel, local control law that enables a robot to reach any position on the plane by exploiting the bearings of at least three landmarks of unknown position, without making assumptions regarding the robot's orientation and without making use of a compass. This control law is the result of the unification of two other local control laws which guide the robot by monitoring the bearing of landmarks and which are able to reach complementary sets of goal positions on the plane. Long-range homing is then realized through the systematic application of the unified control law between automatically extracted milestone positions connecting the robot's current position to the home position. Experimental results, conducted both in a simulated environment and on a robotic platform equipped with a panoramic camera validate the employed local control laws as well as the overall homing strategy. Moreover, they show that panoramic vision can assist in simplifying the perceptual processes required to support robust and accurate homing behaviors.

1.   Introduction

Vision-based robot navigation is an important application of computer vision techniques and tools. Many approaches to this problem either assume the

existence of a geometric model of the environment [Kosaka and Pan, 1995] or the capability of constructing an environmental map [Davison and Murray, 2002]. In this context, the problem of navigation is reduced to the problem of reconstructing the workspace, computing the robot's pose therein and planning the motion of the robot between desired positions. Probabilistic methods [Thrun, 2000] have been developed in robotics that deal with this problem, which is usually referred to as the simultaneous localization and mapping (SLAM) problem.

Catadioptric sensors have been proposed as suitable sensors for robot navigation. A panoramic field of view is advantageous for the achievement of robotic navigational tasks in the same way that a wide field of view facilitates the navigational tasks of various biological organisms such as insects and arthropods [Srinivasan et al., 1997]. Many robotic systems that use panoramic cameras employ a methodology similar to the one employed in conventional camera systems. Adorni et al. discuss stereo omnidirectional vision and its advantages for robot navigation [Adorni et al., 2003]. Correlation techniques have been used to find the most similar pre-stored panoramic image to the current one [Aihara et al., 1998]. Winters et al. [Winters et al., 2000] qualitatively localize the robot from panoramic data and employ visual path following along a pre-specified trajectory in image coordinates.

Panoramic cameras, however, offer the possibility of supporting navigational tasks without requiring range estimation or a localization approach in the strict sense. Methods that rely on primitive perceptual information regarding the environment are of great importance to robot navigation because they pose minimal requirements on a-priori knowledge regarding the environment, on careful system calibration and, therefore, have better chances to result in efficient and robust robot behaviors. This category includes robot navigation techniques that mainly exploit angular information on image-based features that constitute visual landmarks. Several such methods exist for addressing a specific navigation problem, the problem of homing [Hong et al., 1991]. Homing amounts to computing a path that returns a robot to a pre-visited "home" position (see Figure 1). One of the first biologically-inspired methods for visual homing was based on the "snapshot model" [Cartwright and Collett, 1983]. A snapshot represents a sequence of landmarks labeled by their compass bearing as seen from a position in the environment. According to this model, the robot knows the difference in pose between the start and the goal and uses this information to match the landmarks between the two snapshots and to compute its path. There have been several implementations of snapshot-based techniques on real mobile robots. Some of the implemented methods rely on the assumption that the robot has constant orientation or can make use of a compass [Lambrinos et al., 2000, Moller, 2000]. These approaches are not able to support robot homing for any combination of goal (home) snapshot, current
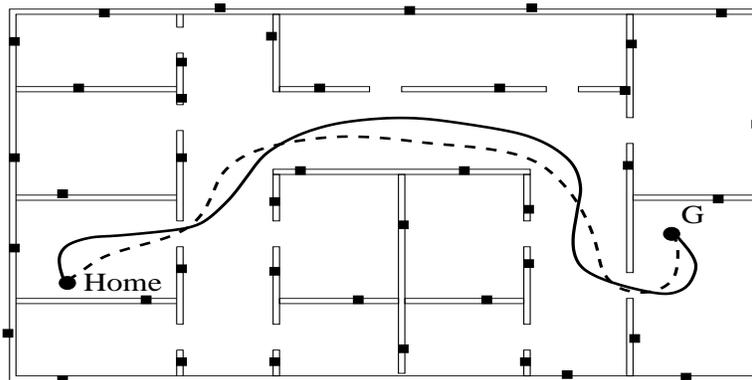
Figure 1. The robot acquires a snapshot of the environment at home position. Then, it wanders in its environment (solid line) and, at some position G homing is initiated so as to return to home (dashed line) by making use of the landmarks available in the workspace (small black rectangles).

position and landmark configuration. Furthermore, the conditions under which the related control laws are successful are not straightforward and cannot be directly inferred from the visual information available at the current and the goal snapshots.

In this work, we present a complete long-range homing strategy for a robot equipped with a panoramic camera. The robot does not have to be aware of its position and orientation and does not have to reconstruct the scene. At the core of the proposed strategy lies a snapshot-based local control law [Argyros et al., 2001], which was later further studied and extended [Bekris et al., 2004]. The advantage of this particular local control law is that it can guide a robot between two positions provided that three landmarks can be extracted and corresponded in the panoramas acquired at these two positions. This implies that there is no inherent control-related issue that restricts the set of position pairs that the algorithm can accommodate. Constraints are only related to difficulties in corresponding features in images acquired from different viewpoints.

Establishing feature correspondences in images acquired from adjacent viewpoints is a relatively easy problem. Thus, short-range homing (i.e., homing that starts at a position close to home) can be achieved by directly applying the proposed local control law as it is described in [Argyros et al., 2005]. In the case of long-range homing (i.e., homing that starts at a position far from home), prominent features are greatly displaced and/or occluded, and the correspondence problem becomes much more difficult to solve [Lourakis et al., 2003]. Therefore, control laws based on the comparison of two snapshot are only local in nature and they cannot support long-range homing. To overcome this problem, the proposed method decomposes homing into a series of simpler

navigational tasks, each of which can be implemented based on the proposed local control law. More precisely, long-range homing is achieved by automatically decomposing the path between the current robot position and the home position with the aid of a set of milestone positions. The selection process guarantees that pairs of milestone positions view at least three common landmarks. The local control law can then be used to move the robot between consecutive milestone positions. The overall mechanism leads the robot to the home position through the sequential application of the control law. Note that using only the basic control law to move between adjacent milestone positions leads to a more conservative selection of such intermediate goals. With the introduction of the complementary control law [Bekris et al., 2004] and its unification with the basic one, the only constraints on the selection of the milestone positions are due to landmark visibility.

The proposed method for robot homing has been implemented and extensively tested on a robotic platform equipped with a panoramic camera in a real indoor office environment. Different kinds of visual features have been employed and tested as alternative landmarks to the proposed homing strategy. In all experiments the home position could be achieved with high accuracy after a long journey during which the robot performed complex maneuvers. There was no modification of the environment in order to facilitate the robot's homing task. The proposed method can efficiently achieve homing as long as enough features exist in the world. Homing will fail only if three robust features cannot be extracted and tracked at any time.

Our approach of robot navigation is similar to that of purposive vision [Aloimonos, 1993]. We use information specific to our problem which is probably not general enough to support many other navigational tasks. We derive partial representations of the environment by employing retinal motion-based quantities which, although sufficient for the accomplishment of the task at hand do not allow for the reconstruction of the state of the robot. Similar findings have been reported for other robotic tasks such as robot centering in the middle of corridors [Argyros et al., 2004].

The rest of the work is organized as follows. Section 2 focuses on the local control strategy that enables a robot to move between adjacent positions provided that a correspondence between at least three features has been established in panoramas acquired at these positions. Section 3 describes our approach on how to automatically decompose a long-range homing task into a series of short-range navigation tasks each of which can be implemented through the proposed local control law. In section 4 we present alternative panoramic image features that can be used to perceptually support the homing strategy. Extensive experimental results from implementations of the proposed homing strategy on a robotic platform are provided in section 5. Moreover, the benefits stemming of the use of panoramic cameras compared to conventional

ones are described in section 6. The work concludes in section 7 with a brief discussion on the key contributions of this work.

2.   Control law

In the following, the robot is abstracted as a point on the 2D plane. The objective of the local control law is to use angular information related to features extracted in panoramic images in order to calculate a motion vector $\overrightarrow{M}$ that, when updated over time, drives the robot to a pre-visited goal position. A *snapshot* of the workspace from a configuration $P \in (\mathbb{R}^2 \times S^1)$, corresponds both to the sequence of visible landmarks and the bearings with which the landmarks are visible from P. The current and the goal position of the robot, together with the corresponding snapshots, will be denoted as A and T, respectively.

Basic control law

We will first consider the case of two landmarks L_i and L_j. The *angular separations* $\theta_{ij}, \theta'_{ij} \in [0, 2\pi)$ correspond to the angles between L_i and L_j as measured at A and T respectively. If $\Delta\theta_{ij} = \theta'_{ij} - \theta_{ij}$ is positive, then the robot views the two landmarks from position T with a greater angle than from position A. The robot will move in a direction that increases the angle θ_{ij}. If $0 \le \theta_{ij} \le \pi$ and $\Delta\theta_{ij} \ge 0$, the robot should move closer to the landmarks. All
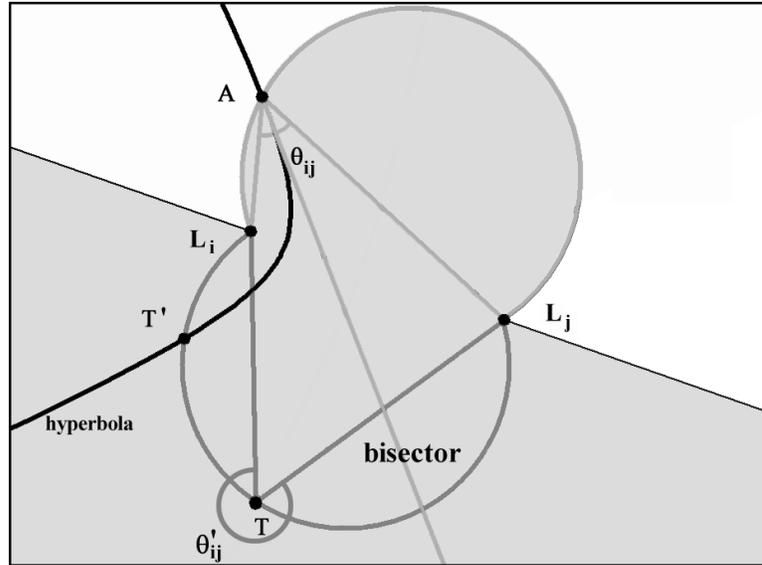

Figure 2.   The definition of the motion vector for two landmarks.

directions that are in the interior of the angle between vectors $\overrightarrow{AL_i}$ and $\overrightarrow{AL_j}$ will move the robot to a new position with greater θ_{ij} including the direction of the angle bisector $\overrightarrow{\delta_{ij}}$. Similarly, when $\theta_{ij} \geq \pi$, moving on the direction of $\overrightarrow{\delta_{ij}}$ increases θ_{ij}. When $\Delta\theta_{ij}$ is negative, the robot should follow the inverse of $\overrightarrow{\delta_{ij}}$. A motion vector that has the above properties and has magnitude that is a continuous function over the entire plane is given by the following equation:

$$\overrightarrow{M_{ij}} = \begin{cases} \Delta\theta_{ij} \cdot \overrightarrow{\delta_{ij}}, & \text{if } -\pi \leq \Delta\theta_{ij} \leq \pi \\ (2\pi - \Delta\theta_{ij}) \cdot \overrightarrow{\delta_{ij}}, & \text{if } \Delta\theta_{ij} > \pi \\ (-2\pi - \Delta\theta_{ij}) \cdot \overrightarrow{\delta_{ij}}, & \text{if } \Delta\theta_{ij} < -\pi. \end{cases} \tag{1}$$

If the robot moves according to the motion vector $\overrightarrow{M_{ij}}$ as this is described in Eq.(1), it is guaranteed to reach the point of intersection of the circular arc $(L_i T L_j)$ and the branch of the hyperbola that goes through A and has points L_i and L_j as foci. An example of such a point is T' in Figure 2. If a third landmark, L_k, exists in the environment, then every position T is constrained to lie on two more circular arcs. A partial motion vector $\overrightarrow{M_{ij}}$ is then defined for each possible pair of different landmarks L_i and L_j. By taking the vector sum of all these vectors the resultant motion vector $\overrightarrow{M}$ is produced. Figure 3 gives an example where $\overrightarrow{M_{ki}}$ and $\overrightarrow{M_{jk}}$ have the same direction as the bisector vectors. $\overrightarrow{M_{ij}}$ is opposite to $\overrightarrow{\delta_{ij}}$ because $\Delta\theta_{ij}$ is negative. The control law can be summarized in the equation

$$\overrightarrow{M} = \overrightarrow{M_{ij}} + \overrightarrow{M_{jk}} + \overrightarrow{M_{ki}}, \tag{2}$$

where the component vectors are defined in Eq.(1). Note that when the robot reaches the goal position, it is guaranteed to remain there because at that point the magnitude of the global motion vector $\overrightarrow{M}$ is equal to zero.

In order to determine the reachability set of this basic control law, i.e., the set of points of the plane that can be reached by employing it in a particular configuration of three landmarks, we ran extensive experiments using a simulator as computed by detailed simulations. The gray area in Figure 4(a) shows the reachability area of the basic control law. The sets of points that are always reachable, independently of the robot's start position, are summarized below:

- The interior $\hat{C}$ of the circle defined by L_1, L_2 and L_3.

- The union $\hat{H}$ of all sets H_j. A set H_j is the intersection of two half-planes. The first half-plane is defined by line $(L_i L_j)$ and does not include landmark L_k, while the second is defined by the line $L_j L_k$ and does not include landmark L_i, where $k \neq i \neq j \neq k$. In Figure 4(b) the white area outside the circle defined by the three landmarks corresponds to the set $\hat{H}$.

Complementary control law

We now present the complementary control law, that reaches the positions that are unreachable by the basic law. As in the case of the basic control law, the complementary control law exploits the bearings of three landmarks.

We first define the π-*difference* of an angular separation θ_{ij} to correspond to $|\pi - \theta_{ij}|$. Points on the line segment $(L_i L_j)$ will have π-difference of θ_{ij} equal to zero. The *nearest landmark pair* (NLP) to the goal is the pair of landmarks $(L_i L_j)$, that has the minimum π-*difference*. The corresponding motion vector will be called the *nearest motion vector* (NMV). From the study of the basic control law, it can be shown that for an unreachable point T, the dominating component vector is the NMV. The robot follows a curve that is close to the hyperbola with the NLP landmarks L_i and L_j as the foci, until it approaches the circular arc $(L_i T L_j)$. Close to the arc, the NMV stops dominating, because $\Delta\theta_{ij}$ approaches zero. If the goal position is located at the intersection of the curve and the arc $(L_i T L_j)$, then the robot reaches the goal. Otherwise, the robot reaches the arc and follows the opposite direction from the goal. Notice that the robot can easily detect which landmark pairs do not correspond to the NLP. When the robot is close to the circular arc defined by the NLP, those two vectors guide the robot away from the goal.

In order to come up with a control law that reaches the complementary set of points to that of the basic control law, the two component motion vectors that are not the NMV vectors should be inverted. The gray area in Figure 4(b) shows the reachability set of this new law.
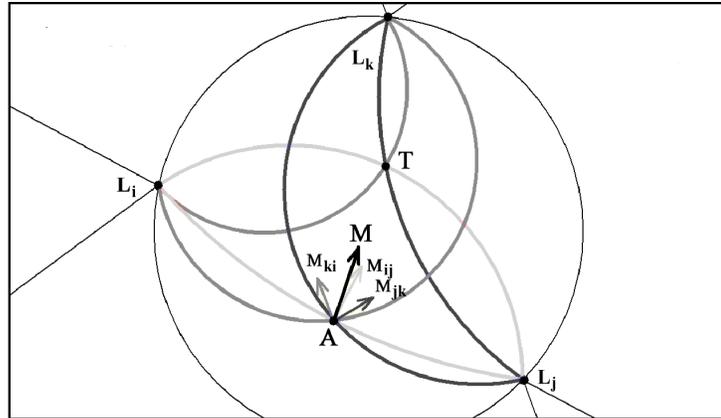

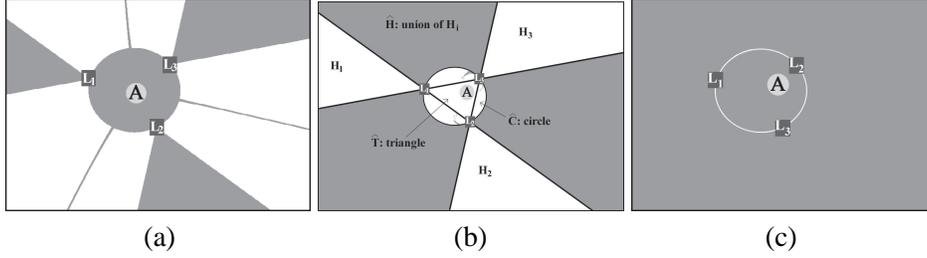*Figure 3.*   The definition of the motion vector for three landmarks.

Figure 4. Simulation results. The robot's initial position is point A and three landmarks L_1, L_2, L_3 exist in the scene. Every point is painted gray if it constitutes a reachable destination by employing (a) the basic control law, (b) the complementary law or (c) the unified control law.

The unification of the two local control laws

In this section we show how to unify the two control laws that have complementary reachability areas in a single law with a reachability area that equals the entire plane. The previous discussion suggests that in order to decide which is the appropriate algorithm to use, the robot must distinguish whether the goal is located in the set $\hat{C}$ or in the set $\hat{H}$ so as to use the basic control law or whether it is located somewhere in the rest of the plane and the complementary law must be used. Deciding whether a snapshot has been taken from the interior of the circle of the landmarks based only on angular information is impossible. Nevertheless, the robot can always move towards the goal by employing the basic algorithm and, while moving, it can collect information regarding the goal snapshot. Based on a set of geometric criteria it is possible to infer whether the basic algorithm was the right choice or if the robot should switch to the complementary law. The geometric criteria consider only the bearings of the landmarks and in one case their rate of change.

For the description of the geometric criteria, we will denote the interior of the landmark's triangle as $\hat{T}$ and the circumscribed circle of two landmarks and the goal as a landmark-goal circle. If the landmarks that correspond to a landmark-goal circle belong to the NLP pair then the circle is called the NLP landmark-goal circle. The geometric criteria that can be used to infer which control law to use based on angular measurements are the following:

1. $\mathbb{T} \in \hat{\mathbb{T}}$? The goal snapshot T is in the set $\hat{T}$ if and only if $\theta'_{ij} < \pi$, $\forall i, j \in [1, 3]$, where L_i and L_j are consecutive landmarks as they are seen from T.

2. $T \in \hat{H}$ and $A \in \hat{T}$? The goal snapshot T is in the set $\hat{H}$ if and only if T can see the landmarks with a different order than A does when A is in $\hat{T}$.

3. $T \notin \hat{T}$ and A on opposite half-plane defined by NLP pair? The robot will then enter $\hat{T}$. If it is going to exit $\hat{T}$ then:
   If the last landmark-goal circle intersected by the robot before leaving $\hat{T}$ is the NLP circle then: $T \notin \hat{C}$.

4. A is on the NLP landmark-goal circle? The goal T is reachable by the basic control law if the non-NLP differences in angular separation are decreasing when the robot has reached the NLP landmark-goal circle.

The overall algorithm that is used for the navigation of the robot is described in Algorithm 1. The robot can be in three possible states: UNCERTAIN, BASIC and COMPLEMENTARY. When in BASIC the robot moves according to the basic control law and when in COMPLEMENTARY the complementary control law is applied.

Algorithm 1   Unified Control Law

---

status = UNCERTAIN;
repeat

  if status is UNCERTAIN **then**
    if $T \in \hat{T}$ **then**
      status = BASIC;
    else if $T \in \hat{H}$ and $A \in \hat{T}$ **then**
      status = BASIC;
    else if $T \notin \hat{T}$ and A on opposite half-plane defined by NLP pair **then**
      if last landmark-goal circle intersected before leaving $\hat{T}$ is the NLP circle **then**
        status = COMPLEMENTARY
      else
        status = BASIC
      end if
    else if A is on the NLP landmark-goal circle **then**
      if the non-NLP differences in angular separation are increasing **then**
        status = COMPLEMENTARY
      end if
    end if
  end if

  if status is BASIC or status is UNCERTAIN **then**
    compute motion vector M with Basic Control Law
  else
    compute motion vector M with Complementary Control Law
  end if
  move according to M
until current snapshot A and goal snapshot T are similar

The initial state is the UNCERTAIN one. The robot is applying the basic control law, but also continuously monitors whether any of the above geometric

conditions have been met. If the goal is located in the interior of the landmark's triangle then the unified algorithm will immediately switch to BASIC. The second criterion can be checked if the robot enters the landmarks' triangle while the third one only upon exiting this triangle. The last criterion is used only if none of the previous ones has given any information and the robot has reached the NLP landmark-goal circle. At this point, the robot can switch behavior by tracking the change in angular separations. These criteria guarantee that the appropriate control law will be used, regardless of the location of the goal.

3.   The strategy for long-range homing

The presented unified local control law may support homing when the latter is initiated from a position close to home. However, in the case that home is far apart from the position where homing is initiated, it may be the case that these two positions do not share any visual feature in common and, therefore, the unified local control strategy cannot support homing. In the following, we propose a memory-based extension to the local control law which enables it to support such a type of long range homing.

The proposed approach operates as follows. Initially the robot detects features in the view acquired at its home position. As it departs from this position, it continuously tracks these features in subsequent panoramic frames. During its course, some of the initially selected features may not be visible anymore while other, new features may appear in the robot's field of view. In the first case the system "drops" the features from subsequent tracking. In the second case, features start being tracked. This way, the system builds an internal "visual memory" where information regarding the "life-cycle" of features is stored.

A graphical illustration of this type of memory is provided in Figure 5. The vertical axis in this figure corresponds to all the features that have been identified and tracked during the journey of the robot from its home position to the current position G. The horizontal dimension corresponds to time. Each of the horizontal black lines corresponds to the life cycle of a certain feature. In the particular example of Figure 5, the home position and position G do not share any common feature and, therefore, the local control law presented in section 2 cannot be employed to directly support homing. In order to alleviate this problem, milestone positions (MPs) are introduced. Being at the end position G, the method first decides how far the robot can go towards home based on the extracted and tracked features. A position with these characteristics is denoted as MP_1 in Figure 5. Achieving MP_1 from the goal position is feasible (by definition) by employing features F_5, F_6 and F_7 in the proposed local control law. The algorithm proceeds in a similar manner to define the next MP
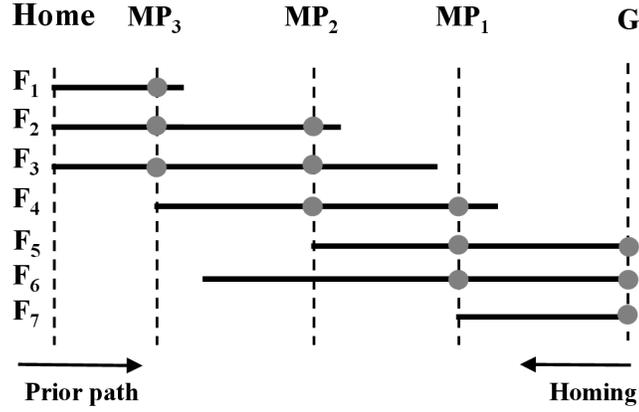
Figure 5.   Graphical illustration of the memory used in long-range homing.

towards home. The procedure terminates when the last MP achieved coincides with the home position.

The local control law of section 2 guarantees the achievement of a target position but not necessarily the achievement of the orientation with which the robot has previously visited this position. This is because it takes into account the differences of the bearings of features and not the bearings themselves. This poses a problem in the process of switching from the features that drove the robot to a certain MP to the features that will drive the robot to the next MP. This problem is solved as follows. Assume that the robot has originally visited a milestone position P with a certain orientation and that during homing it arrives at position P' where P' denotes position P, visited under a different orientation. Suppose that the robot arrived at P' via features F_1, F_2, ..., F_n. The bearings of these features as observed from position P are $A_p(F_1)$, $A_p(F_2)$, $\cdots$, $A_p(F_n)$ and the bearings of the same features as observed from P' are $A_{P'}(F_1)$, $A_{P'}(F_2)$, $\cdots$, $A_{P'}(F_n)$. Then, it holds that

$$A_P(F_i) - A_{P'}(F_i) = \phi, \ \forall i, 1 \leq i \leq n,$$

where ϕ is constant and equal to the difference in the robot orientation at P and P'. This is because panoramic images that have been acquired at the same location but under a different orientation differ by a constant rotational factor ϕ. Since both $A_P(F_i)$ and $A_{P'}(F_i)$ are known, ϕ can be calculated. Theoretically, one feature suffices for the computation of ϕ. Practically, for robustness purposes, all tracked (and therefore corresponded) features should contribute to the estimation of ϕ. Errors can be due to the inaccuracies in the feature tracking process and/or due to the non-perfect achievement of P during homing.

For the above reasons, ϕ is computed as:

$$\phi = median\{A_P(F_i) - A_{P'}(F_i)\}, 1 \leq i \leq n.$$

Having an estimation of the angular shift ϕ between the panoramas acquired at P and P', it is possible to start a new homing procedure. The retinal coordinates of all features detected during the visit of P can be predicted based on the angular displacement ϕ. Feature selection is then applied to small windows centered at the predicted locations. This calculation results in registering all features acquired at P and P' which permits the identification of a new MP and the continuation of the homing procedure. Moreover, if the robot has already arrived at the home position it can align its orientation with the original one by rotating according to the computed angle ϕ.

An important implementation decision is the selection of the number of features that should be corresponded between two consecutive MPs. Although three features suffice more features can be used, if available. The advantage of considering more than three corresponded features is that reaching MPs (and consequently reaching the home position) becomes more accurate because feature-tracking errors are smoothed-out. However, as the number of features increases, the number of MPs also increases because it is less probable for a large number of features to "survive" for a long period. In a sense, the homing scheme becomes more conservative and it is decomposed into a larger number of safer, shorter and more accurate reactive navigation sessions. Specific implementation choices are discussed in the experimental results section of this work.

4.   Extracting and tracking landmarks

The proposed bearing-only homing strategy assumes that three landmarks can be detected and corresponded in panoramic images acquired at different robot positions and that the bearings of these features can be measured. Two different types of features have been employed in different experiments, namely image corners and centroids of color blobs.

Image corners

One way to achieve feature correspondence is through the detection and tracking of image corners. More specifically, we have employed the KLT tracking algorithm [Shi and Tomasi, 1993]. KLT starts by identifying characteristic image features, which it then tracks in a series of images. The KLT corner detection and tracking is not applied directly on the panoramic images provided by a panoramic camera (e.g., the image of Figure 7) but on the cylindrical image resulting by unfolding such an image using a polar-to-Cartesian transformation [Argyros et al., 2004] (see for example the image in Figure

6). In the resulting cylindrical image, the full 360° field of view is mapped on the horizontal image dimension. Once a corner feature F is detected and tracked in a sequence of such images, its bearing $AP(F)$ can be computed as $AP(F) = 2\pi x_F/D$ where x is the x-coordinate of feature F and D is the width of this panoramic image in pixels.
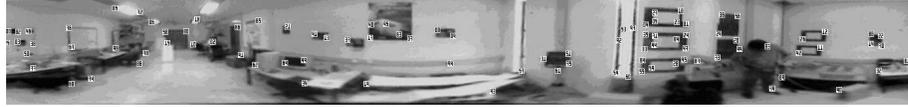

Figure 6. Cylindrical panoramic view of the workspace from the home position that the robot is approaching in Fig. 13. The features extracted and tracked at this panorama are also shown as numbered rectangles.

Centroids of colored blobs

The detection and tracking of landmarks can also be accomplished with the aid of a blob tracker [Argyros and Lourakis, 2004]. Although originally developed for tracking skin-colored regions, this tracker may track multiple colored objects in images acquired by a possibly moving camera. The method encompasses a collection of techniques that enable the modeling and detection of colored objects and their temporal association in image sequences. In the employed tracker, colored objects are detected with a Bayesian classifier which is bootstrapped with a small set of training data. A color model is learned through an off-line procedure that permits the avoidance of much of the burden involved in the process of generating training data. Moreover, the employed tracker adapts the learned color model based on the recent history of tracked objects. Thus, without relying on complex models, is able to robustly and efficiently detect colored objects even in the case of changing illumination conditions. Tracking in time is performed by employing a novel technique that can cope with multiple hypotheses which occur when a time-varying number of objects move in complex trajectories and occlude each other in the field of view of a moving camera.

For the purposes of the experiments of this work, the employed tracker has been trained with color distributions corresponding to three colored posters (Figure 7). These posters are detected and subsequently tracked in the panoramic images acquired during a navigation session. A byproduct of the tracking process is the coordinate (x_{F_i}, y_{F_i}) of the centroid of each tracked landmark F_i. Then, assuming that the center of the panoramic image is (x_p, y_p), the bearing of landmark F_i can easily be computed as $\tan^{-1}\left(\frac{y_p - y_{F_i}}{x_p - x_{F_i}}\right)$. Landmarks that appear natural in indoor environments, such as office doors and desks, have also been successfully employed in our homing experiments.
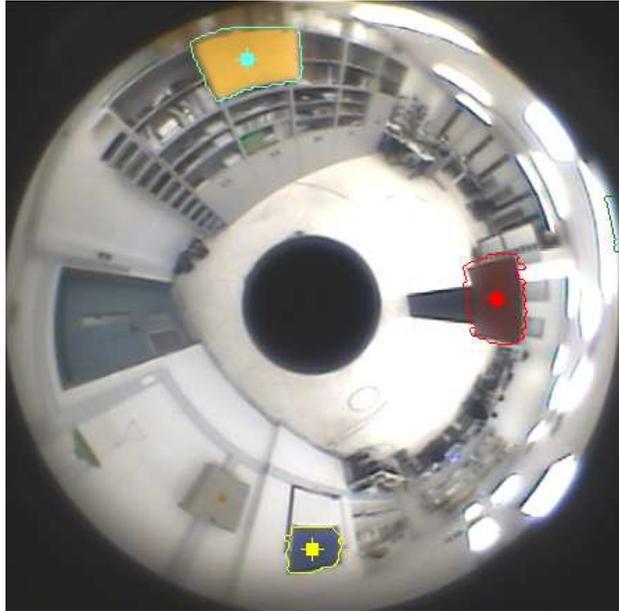
Figure 7. Sample panoramic image with extracted landmarks. Small squares represent the position of the detected and tracked landmarks. The contour of each detected landmark is also shown.

5.   Experiments

A series of experiments have been conducted in order to assess qualitatively and quantitatively the performance of the proposed homing scheme.

Verifying the local control laws

Towards verifying the developed local control strategies, a simulator has been built which allows the design of 2D environments populated with landmarks. The simulator was used to visualize the path of a simulated robot as the latter moves according to the proposed local control laws. Examples of such paths as computed by the simulator can be seen in Figure 8. Additionally, the simulator proved very useful in visualizing and verifying the shape of the reachability areas for the basic, the complementary and the unified local control laws.

Although simulations provide very useful information regarding the expected performance of the proposed local control laws, it is only experiments employing real robots in real environments that can actually test the performance of the proposed navigational strategy. For this reason, another series of experiments employ an I-Robot, B21R robot equipped with a Neuronics,
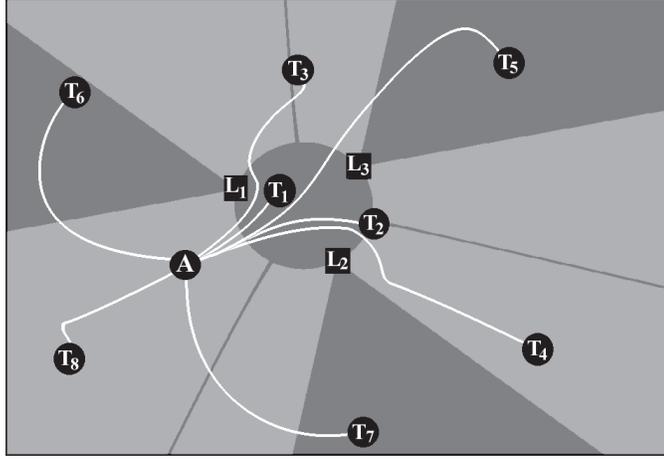
Figure 8. Paths computed by the unified local control law. The reachability sets of the basic and the complementary control laws are shown as dark and light gray regions, respectively.

V-cam360 panoramic camera in a typical laboratory environment. Figure 9(a) illustrates the setting where the reported experiments were conducted. As it can be seen in the figure, three distinctive colored panels were used as landmarks. Landmarks were detected and tracked in the panoramic images acquired by the robot using the method described in section 6. The floor of the workspace was divided into the sets $\hat{C}$, $\hat{H}$ and the rest of the plane for the particular landmark configuration that was used. It should be stressed out that this was done only to visually verify that the conducted experiments were in agreement with the results from simulations. The workspace also contains six marked positions. Figure 9(b) shows a rough drawing of the robot's workspace where the sets $\hat{C}$, $\hat{H}$ as well as the marked positions are shown. Note that these six positions are representative of robot positions of interest to the proposed navigation algorithm, since $A \in \hat{T}$, $F \in \hat{C} - \hat{T}$, $C, D \in \hat{H}$ and B, E are positions in the rest of the plane.

In order to assess the accuracy of the tracking mechanism in providing the true bearings of the detected and tracked landmarks, the robot was placed in various positions in its workspace and was issued a variety of constant rotational velocities (0.075 rad/sec, 0.150 rad/sec). Since this corresponds to a pure rotational motion of the panoramic camera, it was expected for the tracker to report landmark positions changing at a constant rate, corresponding to the angular velocity of the robot. For all conducted experiments the accuracy in estimating the bearing was less than 0.1 degrees per frame, with a standard deviation of less than 0.2.
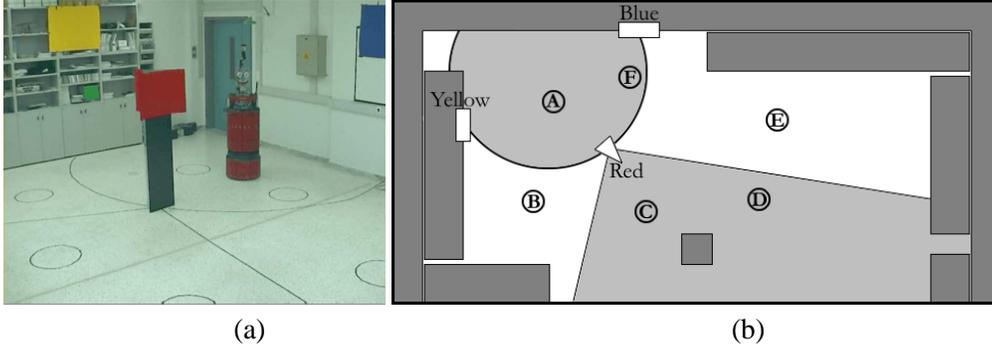
Figure 9. The environment where the experiments were conducted.

A first experiment was designed so as to provide evidence regarding the reachability sets of the three control strategies (basic, complementary and unified). For this reason, each algorithm has been tested for various start and goal positions (3 different starting positions $\times$ 3 different types of starting positions $\times$ 3 different goal positions $\times$ three algorithms). The table in Figure 10 summarizes the results of the 81 runs by providing the accuracy in reaching goal positions, measured in centimeters.

The main conclusions that can be drawn from this table are the following:

- The basic control law fails to reach certain goal positions, independently of the starting position. The reachability set is in agreement with simulation results.

- The complementary control law fails to reach certain goal positions, independently of the starting position. The reachability set is in agreement with simulation results.

Algorithm		Basic Law			Complementary			Combination		
Attempt	Positions	A	C	E	A	C	E	A	C	E
1st	Initial point	3.5	3.0	Fail	Fail	Fail	4.5	1.0	4.5	5.5
2nd	in $\hat{C}$	2.0	1.0	Fail	Fail	Fail	5.5	2.0	3.5	8.5
3rd		0.0	1.5	Fail	Fail	Fail	4.0	4.0	3.0	3.0
1st	Initial point	3.5	11.5	Fail	Fail	Fail	6.0	2.0	9.0	1.5
2nd	in $\hat{H}$	1.5	1.5	Fail	Fail	Fail	2.5	3.5	3.0	6.5
3rd		2.5	2.0	Fail	Fail	Fail	8.5	2.0	3.0	3.5
1st	Initial point	2.0	2.0	Fail	Fail	Fail	2.5	1.5	2.0	2.0
2nd	not in $\hat{C}$	4.0	0.0	Fail	Fail	Fail	9.0	3.5	2.0	5.5
3rd	or , $\hat{H}$	0.5	5.5	Fail	Fail	Fail	3.0	1.5	3.5	8.0

Figure 10. Experiments testing the reachability area and the accuracy of the proposed local control laws.

- The unified control law reaches all goal positions.

- The accuracy in reaching a goal position is very high for all control laws.

To further assess the accuracy of the unified algorithm in reaching a goal position, as well as the mechanisms that the algorithm employs to switch between the complementary and the basic control law, the unified control law was employed 30 times to reach each of the 6 marked positions, resulting in 180 different runs. Figure 11 shows the results of the experiments and summarizes them by providing the mean error and the standard deviation of the error in achieving each position. As it can be verified from Figure 11, the accuracy of the unified law in reaching a goal position is very high as it is in the order of a very few centimeters for all goal positions.

Position:	A	B	E	D	F	C
Mean Val.	1.45	4.65	3.22	2.55	2.28	2.85
St. Dev.	1.13	2.10	1.96	1.35	1.22	1.41

Figure 11. Accuracy of the proposed local control laws in reaching a desired position (distance from actual position, in centimeters)

Additional experiments have been carried out for different landmark configurations, including the special case of collinear landmarks. It is important to note that except from different landmark configurations, different landmarks have been also used. These landmarks were not specially made features such as the colored panels but corresponded to objects that already existed in the laboratory (e.g. the door that can be seen in Figure 9(a), the surface of an office desk, a pile of boxes, etc). The algorithm was also successful in the case that a human was moving in the environment occasionally occluding the landmarks for a number of frames. The tracker was able to recapture the landmark as soon as it reappeared in the robot's visual field. Finally, if the robot's motion towards the goal was interrupted by another process, such as manual control of the robot, the algorithm was able to continue guiding the robot as soon as the interrupting process completed. Sample representative videos from such experiments can be found in http://www.ics.forth.gr/cvrl/demos. In all the above cases the accuracy in reaching the goal position was comparable to the results reported in Figures 10 and 11.

Verifying the strategy for long-range homing

Besides verifying the proposed local control strategy in isolation, further experiments have been carried out to assess the accuracy of the full, long-range navigation scheme. Figure 12 gives an approximate layout of the robot's workspace and starting position in a representative long-range homing experiment. The robot leaves its home position and after executing a predetermined

set of motion commands, reaches position G, covering a distance of approximately eight meters. Then, homing is initiated, and three MPs are automatically defined. The robot sequentially reaches these MPs to eventually reach the home position. Note that the properties of the local control strategy applied to reaching successive MPs are such that the homing path is not identical to the prior path. During this experiment, the robot has been acquiring panoramic views and processing them on-line. Image preprocessing involved unfolding of the original panoramic images and Gaussian smoothing (σ=1.4). The resulting images were then fed to the KLT corner tracker to extract features as described in section 4. Potential features were searched in 7x7 windows over the whole image. The robots maximum translational velocity was 4.0 cm/sec and its maximum rotational velocity was 3 deg/sec. These speed limits depend on the image acquisition and processing frame rate and are set to guarantee small inter-frame feature displacements which, in turn, guarantee robust feature tracking performance. The 100 strongest features were tracked at each time. After the execution of the initial path, three MPs were automatically defined by the algorithm so as to guarantee that at least 80 features would be constantly available during homing.

Figure 13 shows snapshots of the homing experiment as the robot reaches the home position. Figure 6 shows the visual input to the homing algorithm after image acquisition, unfolding and the application of the KLT tracker. The tracked features are superimposed on the image. It must be emphasized that although the homing experiment has been carried out in a single room, the appearance of the environment changes substantially between home position and position G. As it can be observed, the robot has achieved the home position with high accuracy (the robot in Figure 13(c) covers exactly the circular mark on the ground).
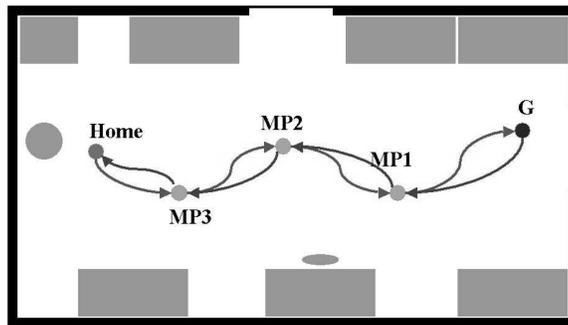

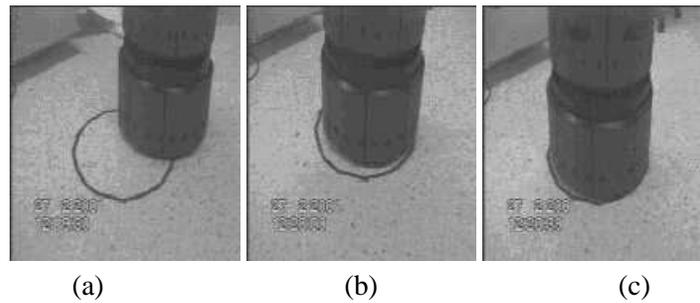*Figure 12.*   Workspace layout of a representative long-range homing experiment.

(a)   (b)   (c)

Figure 13.   Snapshots of the long-range homing experiment, as the robot approaches home.

6.   Advantages of panoramic vision for bearing-only navigation

A major advantage of panoramic vision for navigation is that by exploiting such cameras, a robot can observe most of its surroundings without the need for elaborate, human-like gaze control. An alternative would be to use perspective cameras and alter their gaze direction via pan-tilt platforms, manipulator arms or spherical parallel manipulators. Another alternative would be to use a multi-camera system in which cameras jointly provide a wide field of view. Both alternatives, however, may present significant mechanical, perceptual and control challenges. Thus, panoramic cameras, which offer the possibility to switch the looking direction effortlessly and instantaneously, emerge as an advantageous solution.

Besides the practical problems arising when navigational tasks have to be supported by conventional cameras, panoramic vision is also important because the accuracy in reaching a goal position depends on the spatial arrangement of features around the target position. To illustrate this, assume a panoramic view that captures 360 degrees of the environment in a typical 640×480 image. The dimensions of the unfolded panoramic images produced by such panoramas are 1394×163, which means that each pixel represents 0.258 degrees of the visual field. If the accuracy of landmark localization is 3 pixels, the accuracy of measuring a bearing of a feature is 0.775 degrees or 0.0135 radians. This implies that the accuracy in determining the angular extent of a pair of features is 0.027 radians, or, equivalently, that all positions in space that view pair of features within the above bounds cannot be distinguished. Figure 14 shows results from related simulation experiments. In Figure 14(a), a simulated robot, equipped with a panoramic camera, observes the features in its environment with the accuracy indicated above. Then the set of all positions that the robot would stop by the proposed control strategy are shown in the figure in dark gray color. It is evident that all such positions are quite close to the true robot location. Figure 14(b) shows a similar experiment but involves

a robot that is equipped with a conventional camera with limited field of view that observes three features. Because of the limited field of view, features do not surround the robot. Due to this fact, the uncertainty in determining the true robot location has increased substantially although that the accuracy in measuring each landmark's direction is higher.
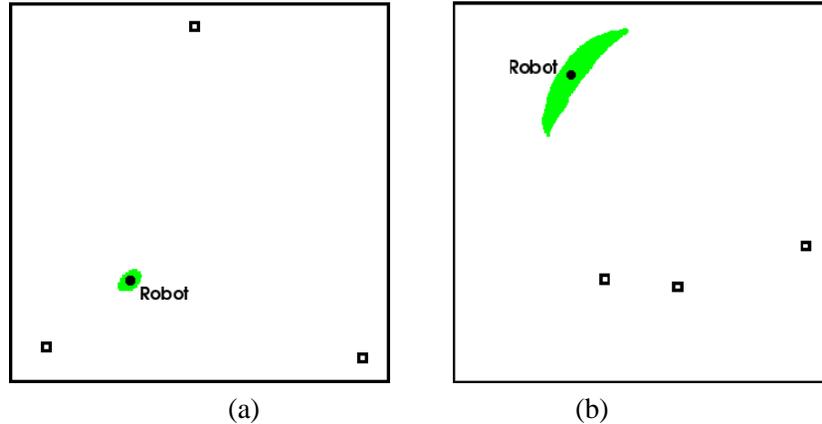

| (a) | (b) |

Figure 14.   Influence of the arrangement of features on the accuracy of reaching a desired position. The darker area represents the uncertainty in position due to the error in feature localization (a) for a panoramic camera and (b) for a 60° f.o.v. conventional camera, and the corresponding landmark configuration.

In current implementations of panoramic cameras, however, the omnidirectional field of view is achieved at the expense of low resolution, in the sense of low visual acuity. This reduced acuity could be a significant problem for tasks like fine manipulation. For navigation tasks, however, it seems that acuity could be sacrificed in favor of a wide field of view. For example, the estimation of 3D motion is facilitated by a wide field of view, because this removes the ambiguities inherent in this process when a narrow field of view is used [Fermuller and Aloimonos, 2000]. As an example, in the experiment of Figure 14(b), the camera captures 60 degrees of the visual field in a 640×480 image. Thus, each pixel represents 0.094 degrees of the visual field and the accuracy of measuring a bearing of a feature is 0.282 degrees or 0.005 radians. Consequently, accuracy in determining the angular extend of a pair of features is 0.01 radians, which is almost three times better compared to the accuracy of the panoramic camera. Still, the accuracy in determining the goal position is larger in the case of panoramic camera.

7.   Discussion

This work has shown that panoramic vision is suitable for the implementation of bearing-only robot navigation techniques. These techniques are able

to accurately achieve a goal position as long as the visual input is able to provide angular measurements without having to reconstruct the robot's state in the workspace. Compared to the existing approaches to robot homing, the proposed strategy has a number of attractive properties. The requirement for an external compass is no longer necessary. The proposed local control strategy does not require the definition of two types of motion vectors (tangential and centrifugal), as in the original "snapshot model" [Cartwright and Collett, 1983] and, therefore, the definition of motion vectors is simplified. We have extended the capabilities of the local control law strategy so that the entire plane is reachable as long as the features are visible by the robot while executing homing. This fact greatly simplifies the use of the proposed local strategy as a building block for implementing long-range homing strategies. In this work we have also presented one such long-range homing algorithm that builds a memory of visited positions during an exploration step. By successively applying the local control strategy between snapshots stored in memory the robot can return to any of the positions it has visited in the past. Last, but certainly not least, it has been shown that panoramic vision can be critically important in such navigation tasks because a wide field of view corresponds to greater accuracy in the achievement of the goal position compared to the increased resolution that pinhole cameras offer. Both the local control laws and the long-range strategy have been validated in a series of experiments which have shown that homing can be achieved with a remarkable accuracy, despite the fact that primitive visual information is employed in simple mechanisms.

References

[Adorni et al., 2003] Adorni, G., Mordonini, M., and Sgorbissa, A. (2003). Omnidirectional Stereo Systems for Robot Navigation. In *Proc. of IEEE Workshop on Omnidirectional Vision and Camera Networks, Omnivis-03*, pages 79–89, Madison, Wisconsin.

[Aihara et al., 1998] Aihara, N., Iwasa, H., Yokoya, N., and Takemura, H. (1998). Memory-based Self Localizaton using Omnidirectional Images. In *Proc. of the Fourteenth Inernational Conference on Pattern Recognition, ICPR-98*, volume 2, pages 1799–1803, Brisbane, Australia.

[Aloimonos, 1993] Aloimonos, Y. (1993). *Active Perception*. Lawrence Erlbaum Assoc.

[Argyros et al., 2005] Argyros, A. A., Bekris, K. E., Orphanoudakis, S. C., and Kavraki, L. E. (2005). Robot Homing by Exploiting Panoramic Vision. *Autonomous Robots*, 19(1):7–25.

[Argyros et al., 2001] Argyros, A. A., Bekris, K. E., and Orphanoudakis, S. E. (2001). Robot Homing based on Corner Tracking in a Sequence of a Panoramic Images. In *CVPR*, volume 2, pages 3–10, Kauai, Hawaii.

[Argyros and Lourakis, 2004] Argyros, A.A. and Lourakis, M.I.A. (2004). Real-time Tracking of Skin-colored Regions by a Potentially Moving Camera. In *Proc. of European Conference on Computer Vision, ECCV-04*, volume 3, pages 368–379, Prague, Czech Republic. Springer-Verlag.

[Argyros et al., 2004] Argyros, A.A., Tsakiris, D.P., and Groyer, C. (2004). Bio-mimetic Centering Behavior: Mobile Robots with Panoramic Sensors. *IEEE Robotics and Automation Magazine, special issue on Panoramic Robotics*, pages 21–33.

[Bekris et al., 2004] Bekris, K. E., Argyros, A. A., and Kavraki, L. E. (2004). New Methods for Reaching the Entire Plane with Angle-Based Navigation. In *Proc. of the IEEE International Conference on Robotics and Automation (ICRA 2004)*, pages 2373–2378, IEEE press, New Orleans, LA.

[Cartwright and Collett, 1983] Cartwright, B. A. and Collett, T. S (1983). Landmark learning in bees: Experiments and models. *Computational Physiology*, 151:521–543.

[Davison and Murray, 2002] Davison, A. J. and Murray, D. W. (2002). Simultaneous Localization and Map-Building Using Active Vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7):865–880.

[Fermuller and Aloimonos, 2000] Fermuller, C. and Aloimonos, Y. (2000). Geometry of Eye Design: Biology and Technology. In R. Klette, T.S. Huang, G.L. Gimelfarb, editor, *Proc. of the tenth International Workshop on Theoretical Foundations of Computer Vision: Multi-Image Analysis (also appeared in Lecture Notes in Computer Science, Vol. 2032)*, pages 22–38. Springer.

[Hong et al., 1991] Hong, J., Tan, X., Pinette, B., Weiss, R., and Riseman, E. M. (1991). Image-based Homing. In *Proc. of the 1991 IEEE International Conference on Robotics and Automation, ICRA-91*, pages 620–625, Sacramento, CA, USA.

[Kosaka and Pan, 1995] Kosaka, A. and Pan, J. (1995). Purdue Experiments in Model-based Vision for Hallway Navigation. In *Proc. of Workshop on Vision for Robots, IROS-95*, pages 87–96, Pittsburg, PA.

[Lambrinos et al., 2000] Lambrinos, D., Moller, R., Labhart, T., Pfeifer, R., and Wehner, R. (2000). A mobile robot employing insect strategies for navigation. *Robotics and Autonomous Systems*, 30:39–64.

[Lourakis et al., 2003] Lourakis, M.I.A., Tzourbakis, S., Argyros, A.A., and S.C., Orphanoudakis (2003). Feature Transfer and Matching in Disparate Views through the Use of Plane Homographies. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(2):271–276.

[Moller, 2000] Moller, R. (2000). Insect Visual Homing Strategies in a robot with Analog Processing. *Biological Cybernetics, special issue: Navigation in Biological and Artificial Systems*, 83-3:231–243.

[Shi and Tomasi, 1993] Shi, J. and Tomasi, C. (1993). Good features to track. Technical Report TR-93-1399, Department of Computer Science - Cornell University.

[Srinivasan et al., 1997] Srinivasan, M., Weber, K., and Venkatesh, S. (1997). *From Living Eyes to Seeing Machines*. Oxford University Press.

[Thrun, 2000] Thrun, S. (2000). Probabilistic Algorithms in Robotics. *AI Magazine*, 21(4):93–109.

[Winters et al., 2000] Winters, N., Gaspar, J., Lacey, G., and Santos-Victor, J. (2000). Omnidirectional Vision for Robot Navigation. In *Proc. of the IEEE Workshop on Omnidirectional Vision 2000, Omnivis-00*, pages 21–28, Hilton Head Island, SC, USA.