

Article

Local Homing Navigation Based on the Moment Model for Landmark Distribution and Features

Changmin Lee [†] and DaeEun Kim ^{*,†}

School of Electrical and Electronic Engineering, Yonsei University, Seoul 03722, Korea; lcmin@yonsei.ac.kr

* Correspondence: daeeun@yonsei.ac.kr; Tel.: +82-2-2123-5879

† Current address: 50 Yonsei-ro, Seodaemun-gu, Seoul 03722, Korea.

Received: 27 September 2017; Accepted: 14 November 2017; Published: 17 November 2017

Abstract: For local homing navigation, an agent is supposed to return home based on the surrounding environmental information. According to the snapshot model, the home snapshot and the current view are compared to determine the homing direction. In this paper, we propose a novel homing navigation method using the moment model. The suggested moment model also follows the snapshot theory to compare the home snapshot and the current view, but the moment model defines a moment of landmark inertia as the sum of the product of the feature of the landmark particle with the square of its distance. The method thus uses range values of landmarks in the surrounding view and the visual features. The center of the moment can be estimated as the reference point, which is the unique convergence point in the moment potential from any view. The homing vector can easily be extracted from the centers of the moment measured at the current position and the home location. The method effectively guides homing direction in real environments, as well as in the simulation environment. In this paper, we take a holistic approach to use all pixels in the panoramic image as landmarks and use the RGB color intensity for the visual features in the moment model in which a set of three moment functions is encoded to determine the homing vector. We also tested visual homing or the moment model with only visual features, but the suggested moment model with both the visual feature and the landmark distance shows superior performance. We demonstrate homing performance with various methods classified by the status of the feature, the distance and the coordinate alignment.

Keywords: local visual navigation; moment model; landmark vector; range sensor; homing navigation

1. Introduction

Navigation is a process of monitoring and controlling the movement of an agent from one place to another. Many navigation systems have their goal positions that the agent is supposed to reach. Recent navigation studies show diverse forms of moving vehicles from simple wheeled robots [1] to a wide range of underwater vehicles [2], unmanned aerial vehicles [3] and spacecraft [4] with various sensors [5,6]. Furthermore, there have been techniques using various sensors like vision [7], inertia [8] and RFID [9], which represent a great deal of engineering achievement.

Different from the above engineering approaches, various animals demonstrate remarkable homing capacity, and their navigation system robustly works in real environments. Local visual homing based on a snapshot model [10] is inspired by insect navigation. An agent is supposed to return to the nest using visual cues or landmarks. The snapshot model uses only a pair of snapshot images at the nest and at the current position. The difference of landmark positions in the two images can be used to derive information about the relative location difference or homing direction. The angular difference in landmark position can greatly contribute to decisions about homing. Honeybees can find the homing direction by reducing the differences of the angular distribution of visual landmarks observed in a pair of snapshots [10].

Desert ants (*Cataglyphis fortis*) are known to use odometry and visual information for their navigation. The path integration with odometry information is related to calculating the accurate home location [11–15]. Cumulative errors in path integration can be compensated by visual cues including the skyline and polarized information, as well as the surrounding landmark information [16–18]. There are many other examples with visual cues [19–22], and also, their navigation involves many types of sensors including vision, olfactory, auditory, odometry and magnetic sensors. [23–26]

There have been many local visual homing techniques. We can largely divide these into two parts using depth or intensity information [27] as shown in Figure 1. Furthermore, intensity-based approaches can be divided into holistic methods and correspondence methods. Correspondence methods match extracted features in the images and thus need a complex algorithm to process feature extraction and feature matching. If there is a reference compass available, the one-to-one correspondence matching can be omitted. Holistic methods try to match the whole pixel information in the two images without extracting landmark features or classifying the features. Holistic methods have relatively low complexity.

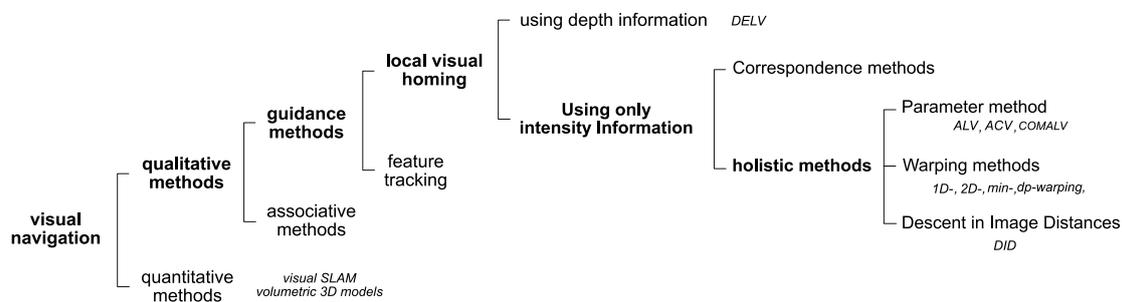


Figure 1. Overview of visual navigation. DELV, Distance Estimated Landmark Vector; ALV, Average Landmark Vector; ACV, Average Correctional Vector.

Landmark vector models to represent the surrounding environment have been studied, and the models often use angular positions of landmarks without landmark matching, if a reference compass such as a light or magnetic compass is available and the two snapshots can be aligned with the reference coordinate. The Average Landmark Vector (ALV) model [28] is a typical example of the landmark vector model covering the omnidirectional view. Even if it is assumed that all the landmarks on the retinal image have equal distances, the whole distribution of landmarks can be simply represented as the average landmark vector, each of which has a unit length with its angular position. According to the snapshot model [10], the model compares two ALVs obtained from the home snapshot and the current view. The difference between the two vectors can estimate the homing direction from the current position. The ALV model can be combined with visual feature detection [29]. The Average Correctional Vector (ACV) [30,31] is a variation of the ALV model. This model uses the amount of angular differences as the length of the landmark vector. Another method, the Distance Estimated Landmark Vector (DELV) model [32,33], suggests encoding distance in the landmark vectors and provides better estimation of the homing vector at the current position by localizing the current position in a reference map. The snapshot matching method can be combined with optic flow [34], and it has been applied to aircraft trying to estimate the current location.

In the holistic methods, there have been two ways to handle the whole image pixels, the image distance model and the warping model. The Descent in Image Distance (DID) method [35] uses the all pixels in a pair of images to calculate the image distance. The pixel difference between the snapshot image at a given position and the home snapshot can roughly estimate the relative distance between the two positions. If a snapshot among candidate snapshots at different positions is close to the home snapshot in terms of the image distance, it is assumed that the direction to the snapshot is close to the homing direction. More advanced models following the idea have been studied [36,37].

Homing in scale space [38] uses correspondences between SIFT features and analyzes the resulting flow field to determine the movement direction. It is mathematically justified by another study [39]. Another view-based homing method based on the Image Coordinates Extrapolation (ICE) algorithm has been tested [40].

In contrast to the above image distance models, there have been warping methods to calculate all possible matchings between all pixels in a pair of images as another holistic approach. In the one-dimensional warping model [41–43], all possible changes of pixels along the horizontal line are calculated under the assumption that landmarks have equal distances. The homing direction can be estimated by searching for the smallest difference in a particular angle between the candidate image and the home reference image. There are also advanced warping models including the 2D-warping and min-warping model [44–46] that apply a variation of alignment angle estimation for the environment without a reference compass. There have been other variations of the warping model [47,48]. Recently, a method with various visual features like SURF has been compared with the holistic approach for robotic experiment [49]. Generally, the holistic methods show robust performance, but need high computing time.

The image warping methods normally have an equal distance assumption for landmarks. However, the visual information is easily changed by position, distance, luminous sources, shades or other environmental factors. They can produce large homing error depending on the environmental situation. The distance information has been very often neglected in the holistic approaches, as well as landmark vector models, although it can contribute to reading the surrounding environmental information. Recently, it was shown that the depth information greatly improves the homing performance [33,50]. In this paper, we suggest a moment model to combine the distance information with visual features or image pixel information.

Another issue in local visual homing is related to an alignment problem of two snapshot images. If there is no reference compass, the orientations of the current view and the home need to be aligned together. That is, one visual feature in the current view should correspond to a visual feature in the home snapshot, since the visual feature points to the same landmark object. A solution to match the two different orientations is to calculate the image differences between the home snapshot and the rotated image of the current snapshot and find the rotation angle of the current view with the minimal image difference, which is called the visual compass approach [35]. Another approach is the landmark arrangement method [51] in which a set of visual landmarks in the current view are re-mapped to visual landmarks one by one in the reference coordinate, and a circular shift of the landmarks is applied repeatedly to find the best matching of the visual landmarks in the two orientations by checking if a set of the resultant homing vectors starting from each landmark has converged into one point with small variance. In our experiments, we will test the above two approaches for the alignment of the two snapshot images, if there is no reference compass.

In the complex cluttered environment, the Simultaneous Localization and Mapping (SLAM) method has been popular. SLAM often uses a laser sensor for distance information to build a map for the environment. Interestingly, a moment model called Elevation Moment of Inertia (EMOI) has been suggested to imitate a physical quantity, moment of inertia [52]. The model characterizes the environmental landscape, the surrounding range value and height information as a scalar value called EMOI. Inspired by the model, we suggest a new type of moment function to cover various environmental information, which can read the landmark distribution from the environment with two components, the distance to landmarks and the visual feature of landmarks.

The main contribution of our work is to suggest a new type of moment potential to guide homing navigation and prove the convergence of the moment model to a unique reference point. The homing navigation follows the snapshot theory to compare two snapshots to determine the homing direction. We provide a homing vector estimation based on the reference point in the moment model for a pair of snapshots. Furthermore, it is shown that the moment model can encode the landmark distribution and features. Our approach can be extended to a moment model with multiple features, which can

produce multiple reference points for robust homing performance. The combinational model with distance and visual features shows better homing performance than the distance information alone or the visual features alone. We demonstrate robotic homing experiments with the moment model and various methods.

2. Method

2.1. Robot Platform and the Environment

We test robotic experiments in a $6\text{ m} \times 6\text{ m}$ room with several objects including dresser, drawers, a trash can, a plant, windows and walls. We use i-Robot Roomba for a mobile robot with two wheels, which is connected to a laptop computer for control. Here, the robot platform can read a panoramic image of the surrounding environment through an omnidirectional camera in Figure 2, which consists of a Logitech Webcam E3500, a metal ball for a reflection mirror and the acryl support for mounting. The robot can also be mounted with a HOKUYO laser sensor URG-04 LX model shown in Figure 2. The laser sensor can cover a range of 240 degrees in angular space with 0.36 degree resolution. To obtain an omnidirectional distance image, two shots from the laser sensor are needed. In this way, we can collect both an omnidirectional color image and depth information. Figure 3 shows an example of the reconstructed image and depth map for the surrounding environment (red \times marks show sensor readings from the laser sensor).

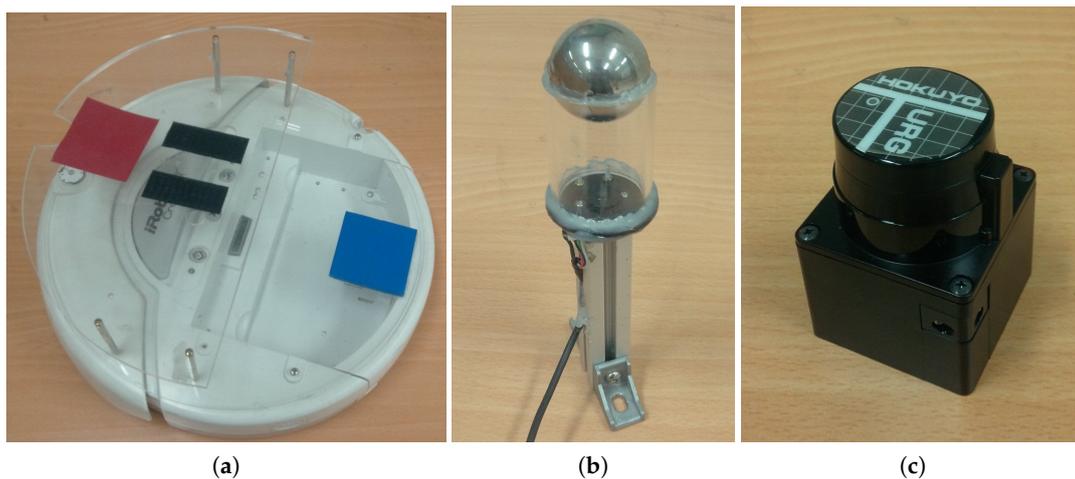


Figure 2. Mobile robot platform: (a) i-Robot ROOMBA mobile; (b) omnidirectional camera; (c) HOKUYO laser sensor.



Figure 3. Cont.

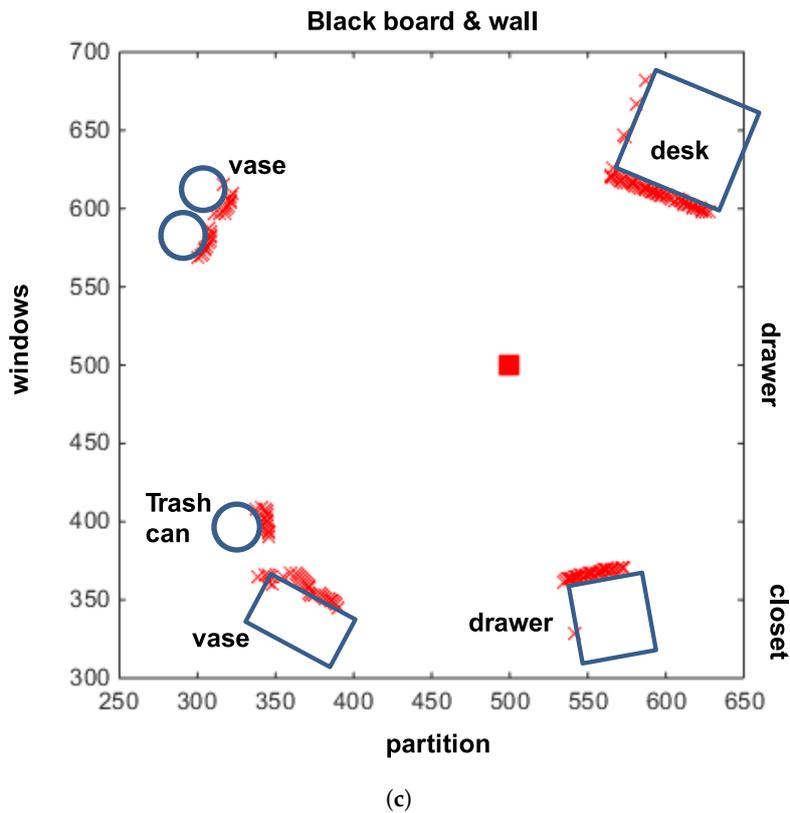


Figure 3. Experimental environment *env0*: (a) a view of the environment; (b) omnidirectional image from the camera; (c) distance map from the range sensor; the red square indicates the home position (500,500).

An omni-directional image that the robot takes has 640×480 RGB pixels. It is converted into a panoramic image, 720×120 pixels (with 0.5 degree angular resolution), with a uniform size of pixels for each angular position, while the omnidirectional image has a relatively small number of pixels near the observation point. The panoramic form can easily access the pixel in terms of the angular position and distance, the angular position in the x -axis and distance from the center of the image in the y -axis. Figure 4a,b shows examples of panoramic images obtained at two different positions. Figure 4c,d shows range data from laser sensor readings corresponding to the panoramic images. The snapshot images or snapshot distance maps have a similar landscape, but they are distorted depending on the position.



Figure 4. *Cont.*

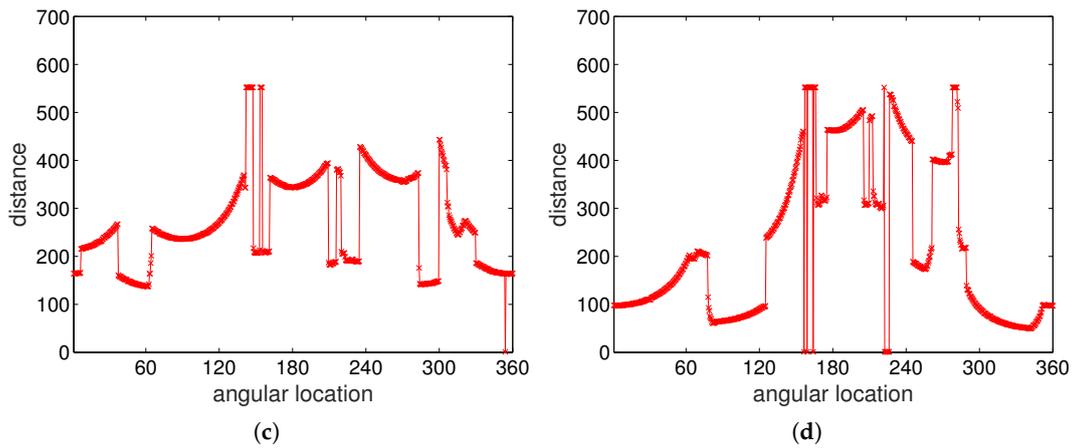


Figure 4. Omnidirectional image and distance map: (a,b) panoramic images at (50,50) and (62,54), respectively; (c,d) omnidirectional depth information at at (50,50) and (62,54), respectively.

2.2. Moment Model for Landmark Distribution

In physics, the moment of inertia is a property of an area that reflects how its points are distributed. By analogy, the moment is defined as a distribution of point measurements in our navigation model.

For a given set of landmarks in the environment, we analyze the landmark distribution as a combination of their positions and features. The color intensity or height of landmarks can be feature candidates. A landmark is defined as a natural feature in the world environment, which is observable even at a far distance. All the landmarks are projected into the image plane, and the snapshot view includes a collection of landmarks. Often, an object is represented with a cluster of pixels in the image through the image segmentation process. Without any object feature extraction, each pixel in the image view can be regarded as a landmark, and then, the feature extraction process can be omitted. In real environments, the color pixels in the surrounding view (omnidirectional view) are regarded as landmarks with the color intensity, as well as the range information. Even the background at a far distance is represented as a set of landmarks. If salient landmarks are identified from the background, only those landmarks may be used in our moment model.

The color of the visual cue is the feature used in this paper. We define the moment measure M as follows:

$$M = \sum_{i=1}^N r_i^2 C_i = \sum_{i=1}^N ((x - a_i)^2 + (y - b_i)^2) C_i \quad (1)$$

where there are N landmarks, r_i is the range value of the i -th landmark, that is the distance from the current location (x, y) to the landmark location (a_i, b_i) , and C_i is the feature value, for example the color intensity of the i -th landmark.

The above measure is similar to the moment of inertia in physics, $\int r^2 dm$. We can also see this measure as a potential function built with a set of landmarks. From that, we can find the gradient as the first derivative of the potential function as follows:

$$\nabla M = \left(\frac{dM}{dx}, \frac{dM}{dy} \right) = (M_x, M_y) = \sum_{i=1}^N (2(x - a_i) C_i, 2(y - b_i) C_i) \quad (2)$$

where this gradient vector indicates the change of the potential function corresponding to the current position (x, y) . To find the minimum convergence point with the gradient, we calculate the determinant of the Hessian matrix given below:

$$H = \begin{pmatrix} \frac{d^2M}{dx^2} & \frac{d^2M}{dxdy} \\ \frac{d^2M}{dxdy} & \frac{d^2M}{dy^2} \end{pmatrix} = \begin{pmatrix} M_{xx} & M_{xy} \\ M_{yx} & M_{yy} \end{pmatrix} = \sum_{i=1}^N \begin{pmatrix} 2C_i & 0 \\ 0 & 2C_i \end{pmatrix} \quad (3)$$

where this Hessian matrix is produced from the second-order differential of the moment equation. M_{xx} is a second-order differential with respect to x and M_{yy} with respect to y . M_{xy} is equal to M_{yx} , and they are zero.

The determinant of the Hessian matrix is calculated as:

$$\det(J) = \left(\sum_{i=1}^N 2C_i \right) \times \left(\sum_{i=1}^N 2C_i \right) = \left(\sum_{i=1}^N 2C_i \right)^2 > 0 \quad (4)$$

We assume that each feature value (C_i) is positive. The sign of the second derivatives of potential function is positive as shown below:

$$M_{xx} = M_{yy} = \sum_{i=1}^N 2C_i > 0 \quad (5)$$

From the above property, there is only one global convergence (minimum potential) point with its gradient zero, and the determinant of the Hessian matrix is positive. Let (X, Y) be the convergence point. Then:

$$\nabla M(X, Y) = \sum_{i=1}^N (2(X - a_i)C_i, 2(Y - b_i)C_i) = 0 \quad (6)$$

The position (X, Y) is calculated as:

$$X = \sum_{i=1}^N a_i C_i / \sum_{i=1}^N C_i, \quad Y = \sum_{i=1}^N b_i C_i / \sum_{i=1}^N C_i \quad (7)$$

where the convergence point (X, Y) is the weighted average of landmark positions with respect to the landmark features, that is the center of the landmark distribution. The moment measure based on features of landmarks has unique convergence point (X, Y) , regardless of any current position (x, y) .

Thus, we argue that if there is no environmental change or no occlusion observed as the robot moves, then we can find the same convergence point in spite of any movement or any change in angular position. To guarantee the unique convergence point, the feature value should be positive. In our experiments, the landmark characteristics are defined as the height of landmarks or color intensity, which is positive. The moment measure is an index of landmark distribution, and its center of distribution can easily be estimated as an invariant feature, which will be useful in homing navigation. In Figure 5, the surface of the potential function is convex-shaped, and the unique convergence point is available. Various types of feature (C_i) values are available, and the convergence point can change depending on the features.

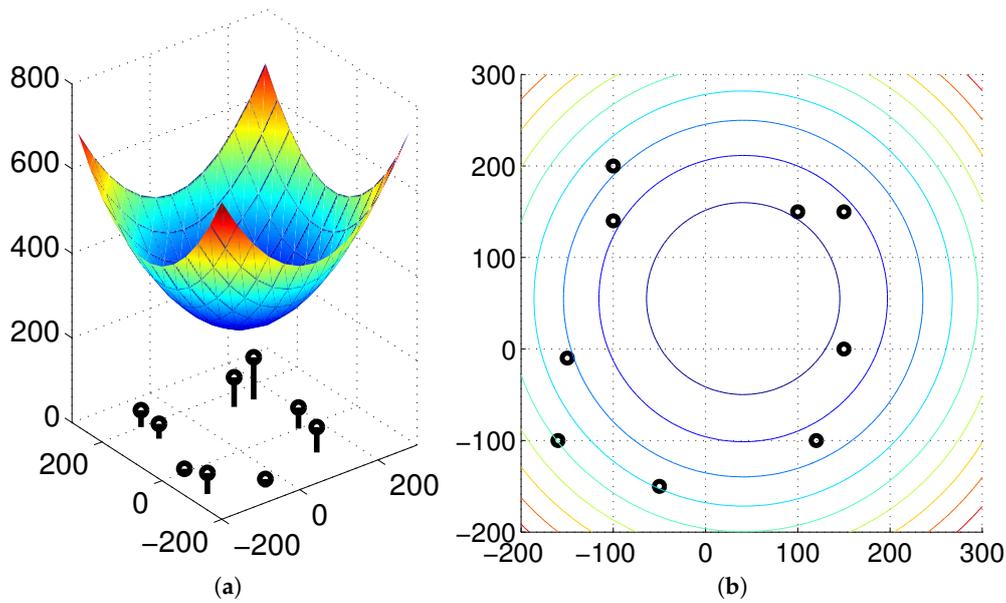


Figure 5. Moment measure as a potential function where the landmark height is regarded as a feature value: (a) potential function, (b) contour line.

2.3. Homing Vector Using the Moment Model

We introduce how to estimate the homing vector using the moment function. We assume there is a reference compass available. Each landmark has the feature value and range information. An agent can observe a distribution of landmarks at a given position. Assume the same landmarks are observed at any position in the environment. We take the above global convergence point as the reference point to estimate the homing vector.

If there are N landmarks observed at the current position $P = (x, y)$, their relative distance $r_i = \|(a_i - x, b_i - y)\|$ and the feature value C_i are measured for $i = 1, \dots, N$, where $(a_i - x, b_i - y)$ is the estimated landmark position in the coordinate with origin at the current position. The reference point vector $R = (X_c, Y_c)$ can be calculated by Equation (7). In a similar way, N landmarks are observed at the home position $\bar{H} = (H_x, H_y)$. Their relative distance $r'_i = \|(a'_i - H_x, b'_i - H_y)\|$ and the feature value C'_i are measured for $i = 1, \dots, N$, where $(a'_i - H_x, b'_i - H_y)$ is the estimated landmark position with origin at the home position \bar{H} . The reference point vector $R' = (X_h, Y_h)$ can be calculated by Equation (7) again at the home position.

Then, we find the relation for homing vector \vec{H} :

$$\vec{H} = \bar{H} - P \simeq R - R', \quad (8)$$

since we assume that the same reference point is estimated irrespective of any observation point, that is the two vectors R and R' should end at the same reference point, starting from the different positions, the current position and the home location (a little deviation of the reference points may be observed by noisy sensor readings or landmark occlusions).

At an arbitrary position $P = (x, y)$, a mobile robot has information of the relative distance and the visual features with a laser sensor and a vision camera. Equation (7) has absolute coordinate representation, and so, we evaluate the convergence point in the coordinate with origin at the observation point.

$$X_c = \sum_{i=1}^N (a_i - x)C_i / \sum_{i=1}^N C_i, \quad Y_c = \sum_{i=1}^N (b_i - y)C_i / \sum_{i=1}^N C_i \quad (9)$$

where $P = (x, y)$ is the current observation point and $(a_i - x, b_i - y)$ is the relative distance of the i -th landmark in the current view. Similarly, the convergence point can be evaluated in the home coordinate as follows:

$$X_h = \frac{\sum_{i=1}^N (a'_i - H_x) C'_i}{\sum_{i=1}^N C'_i}, \quad Y_h = \frac{\sum_{i=1}^N (b'_i - H_y) C'_i}{\sum_{i=1}^N C'_i} \quad (10)$$

where (H_x, H_y) is the home location. Then, the difference of the two reference points measured at two observation points (the home location and the current position) is given by:

$$R - R' = (X_c, Y_c) - (X_h, Y_h) \quad (11)$$

$$= \left(\frac{\sum (a_i - x) C_i}{\sum C_i}, \frac{\sum (a_i - y) C_i}{\sum C_i} \right) - \left(\frac{\sum (a'_i - H_x) C'_i}{\sum C'_i}, \frac{\sum (a'_i - H_y) C'_i}{\sum C'_i} \right) \quad (12)$$

$$\simeq \left(\frac{\sum (a_i - x) C_i}{\sum C_i}, \frac{\sum (a_i - y) C_i}{\sum C_i} \right) - \left(\frac{\sum (a_i - H_x) C_i}{\sum C_i}, \frac{\sum (a_i - H_y) C_i}{\sum C_i} \right) \quad (13)$$

$$= \left(\frac{\sum (H_x - x) C_i}{\sum C_i}, \frac{\sum (H_y - y) C_i}{\sum C_i} \right) \quad (14)$$

$$= (H_x - x, H_y - y) = \vec{H} - P \quad (15)$$

where it is assumed that the same landmarks and the same visual features are observed at any position, $(a_i, b_i) = (a'_i, b'_i)$, $C_i = C'_i$ for $i = 1, \dots, N$. Hence, the homing vector \vec{H} can be estimated by the above property,

$$\vec{H} = (X_c, Y_c) - (X_h, Y_h) \quad (16)$$

Each position defines its own reference map, but there exists a unique convergence point that is the same position regardless of any coordinate. By the convexity of the moment potential function, the minimal potential point can be reached from any position. We provided a proof that the homing vector calculated by the above model can reach the home position from any position, if the environment is isotropic, that is all landmarks and their features are invariantly observed at any position.

2.4. Moment Model with Multiple Features

If there are multiple features available for landmarks, then we can build a separate moment model for each feature. The set of moment models will lead to independent reference points, but we can assume that the distribution of each feature in the environment will be almost equal for any measured position if the environment is isotropic, that is the majority of landmarks are commonly observed in the environment. The homing vector for each feature can be voted together, which can help estimate homing direction more accurately.

We can test the moment model with RGB color intensities, three visual features for each pixel. The image colors provide three different features, red, green and blue color intensity for each pixel. The landmark feature C_i can thus have three components. The moment measure for each feature, red, blue and green intensity, respectively, is defined as follows:

$$M_R = \sum_{i=1}^N r_i^2 R_i = \sum_{i=1}^N ((x - a_i)^2 + (y - b_i)^2) R_i \quad (17)$$

$$M_G = \sum_{i=1}^N r_i^2 G_i = \sum_{i=1}^N ((x - a_i)^2 + (y - b_i)^2) G_i \quad (18)$$

$$M_B = \sum_{i=1}^N r_i^2 B_i = \sum_{i=1}^N ((x - a_i)^2 + (y - b_i)^2) B_i \quad (19)$$

where (a_i, b_i) for $i = 1, \dots, N$ are the landmark position with respect to the current position $P = (x, y)$ and (R_i, B_i, G_i) are the color intensity for the i -th landmark.

Then, the above three measures lead to three reference points at a given position $P = (x, y)$, using Equation (9).

$$(X_R, Y_R) = \left(\frac{\sum (a_i - x)R_i}{\sum R_i}, \frac{\sum (b_i - y)R_i}{\sum R_i} \right), \quad (20)$$

$$(X_G, Y_G) = \left(\frac{\sum (a_i - x)G_i}{\sum G_i}, \frac{\sum (b_i - y)G_i}{\sum G_i} \right), \quad (21)$$

$$(X_B, Y_B) = \left(\frac{\sum (a_i - x)B_i}{\sum B_i}, \frac{\sum (b_i - y)B_i}{\sum B_i} \right), \quad (22)$$

Three reference points can be determined both at the current position and at the home location. The difference of the reference points can estimate the homing direction.

The home vector \vec{H} via the three reference points using the color intensities can be derived as a combinational form,

$$\vec{H} = (\vec{H}_R + \vec{H}_G + \vec{H}_B) / 3 \quad (23)$$

where \vec{H}_R is calculated with red color intensity, \vec{H}_G with green and \vec{H}_B with blue, using Equation (16).

As shown above, the color intensity of pixels can be applied to the moment model with multiple features. We can extend the moment measure into that with various visual features. The visual feature C_i allows any characteristics of landmarks, and also, multiple features can derive multiple homing vectors. The sum of the homing vector for each feature can be effective on noisy feature readings. RGB color space can be converted into another space, for example HSV space, and each feature can make separate homing vectors. Furthermore, to handle noisy sensor readings, we can allow a cut-off threshold for a feature value, and some feature can be set to $C_i = 0$. This has the effect of choosing a set of landmarks in the omnidirectional view, instead of using the whole pixels. If $C_i = 0$ or $C_i = 1$ with the range value $r_i = 1$, the moment model becomes similar to the ALV model [28]. If $C_i = 0$ or $C_i = 1$ with continuous range value $r_i > 0$, the model is similar to the DELV model [33].

2.5. Comparison with Other Methods

To compare our moment model with other conventional approaches, we consider possible combinations of the process components. The components are related to what kind of features will be used, whether the range sensor is available or the distance can be estimated in the visual image and what kind of coordinate alignment process will be applied.

In the moment model, we can allow variable features in C_i . In our experiments, we will mostly use the RGB color intensity as the feature value. To compare variable visual features and no visual feature, we can test $C_i = 1$ for equal color intensity. In addition, the moment model requires the distance information to calculate the centroid. Since there is no distance information of landmarks in the image set, we estimated the distance using the ground line in the image, which was used for the moment model. The ground line is the boundary line between the floor and an object in the panoramic image. After blurring the whole image, a moving mask (4×2 pixels) to detect the horizontal edge was applied, and the number of pixels between the horizontal line and the detected ground line was counted to estimate the distance; more pixels counted indicate a larger distance to the ground or the landmark. We discriminate the methods with the ground-distance estimation in the visual image and those with the laser sensor readings for distance.

Without a reference compass, the snapshot images at two positions are not aligned. Their image coordinates may be different. We need to make an effort to find the rotation angle of a given coordinate to be matched to another coordinate. One of the well-known algorithms for alignment is the visual compass [35]. The home snapshot is taken as a reference image, and the current view is shifted a step angle until the home image and the shifted image have the minimum difference. The shifted angle is the right angle in the coordinate alignment.

Another alignment algorithm called landmark rearrangement is available [51]. A set of landmarks observed in the home coordinate should match another set of landmarks observed at the current position, if it is assumed that the environment is isotropic. That is, two sets of landmarks at the home location and at the current position should correspond to each other. If two coordinates (orientations) are not aligned, then the two sets of landmarks should be compared for one-to-one correspondence by a rotating shift of one landmark at a time until they closely match each other. The landmark rearrangement method first draws landmark vectors from the home location to each landmark and then adds the opposite of landmark vectors from the current position to each landmark by one-to-one mapping. Then, the sum of the resulting vectors can be converged to a point, if there is no matching error. If the two coordinates are not aligned and landmarks have a mismatch, the end points of the resulting vectors have a large variance. We use this variance criterion to make the two coordinates aligned by following the landmark rearrangement [51]. We need to apply a rotational shift of landmarks in a given coordinate to update a series of landmarks.

For visual homing navigation, the Descent in Image Distance (DID) method [35] can be applied to calculate homing direction, which uses snapshots near the current position and the home snapshot. We take a variation of the DID with multiple (three) images as reference images and the snapshot at the current location in order to determine the homing direction. Using the property that the absolute image difference between a pair of snapshots is proportional to the distance, the snapshot at the current position is compared with three snapshots (including the home snapshot) near the home location, and then, the image differences can determine the homing direction by the relative ratio of the image distance between each of the three snapshots and the current image. This DID method uses only visual images, and the relative difference of the in-between images can guide the homing direction without any depth sensor.

Table 1 shows various methods classified by the feature selection, the range sensor or distance estimation and the alignment process. With the reference compass, no alignment process is required, and six methods are available. Without a reference compass, eight methods are listed (for $C_i = 1$, no visual feature is observed and only landmark rearrangement will be tested, since the visual compass is not applicable).

Table 1. Various methods classified by the feature, the range sensor and the alignment process.

| | Method | Feature | Range | Alignment Method | Class |
|---------------------------------|--------|-------------------------------|------------------------------|------------------------|------------------------|
| with reference compass | DID | color intensity | no sensor | aligned by compass | <i>I</i> |
| | moment | color intensity | equal distance ($r_i = 1$) | aligned by compass | <i>II</i> |
| | moment | equal intensity ($C_i = 1$) | ground-line estimation | aligned by compass | <i>III</i> |
| | moment | color intensity | ground-line estimation | aligned by compass | <i>IV</i> |
| | moment | equal intensity ($C_i = 1$) | laser sensor | aligned by compass | <i>V</i> |
| | moment | color intensity | laser sensor | aligned by compass | <i>VI</i> |
| without reference compass | DID | color intensity | no sensor | visual compass | <i>I_v</i> |
| | moment | color intensity | equal distance ($r_i = 1$) | visual compass | <i>II_v</i> |
| | moment | equal intensity ($C_i = 1$) | ground-line estimation | landmark rearrangement | <i>III_r</i> |
| | moment | color intensity | ground-line estimation | visual compass | <i>IV_v</i> |
| | moment | color intensity | ground-line estimation | landmark rearrangement | <i>IV_r</i> |
| | moment | equal intensity ($C_i = 1$) | laser sensor | landmark rearrangement | <i>V_r</i> |
| | moment | color intensity | laser sensor | visual compass | <i>VI_v</i> |
| | moment | color intensity | laser sensor | landmark rearrangement | <i>VI_r</i> |

3. Experimental Results

We defined the moment function as a property to characterize the landmark distribution consisting of the distance and visual features. Initially, we will see the homing performance with the model in the simulation environments. Then, we will test the moment model for homing navigation in the indoor environments. We also investigate the model under various conditions to see the effect of the visual features, the landmark distance and the coordinate alignment (orientation alignment).

3.1. Simulation Environment

Initially, we test our approach in the simulation environment consisting of a set of salient landmarks instead of pixel-wise landmarks. In the environment, it is assumed that an agent can observe extrusive objects discriminated from the background area at a far distance, and those objects are regarded as landmarks. Furthermore, the agent sees all the landmarks at any position. A set of landmarks (discrete landmarks) is given in an arena, and each landmark has its height. We assume that the distance to each landmark is obtained with a laser sensor. In Figure 6, we can observe the unique convergence point with a feature value of height for each environment. The home location may be different from the convergence point (with the minimum potential), but the moment model guides the homing direction well. In the simulation environment, a mobile agent makes no obstacle avoidance behavior and only moves towards the target position, assuming all the object landmarks are visible at any position. Figure 6 shows only homing vectors calculated at a grid position based on the moment model by ignoring the obstacles.

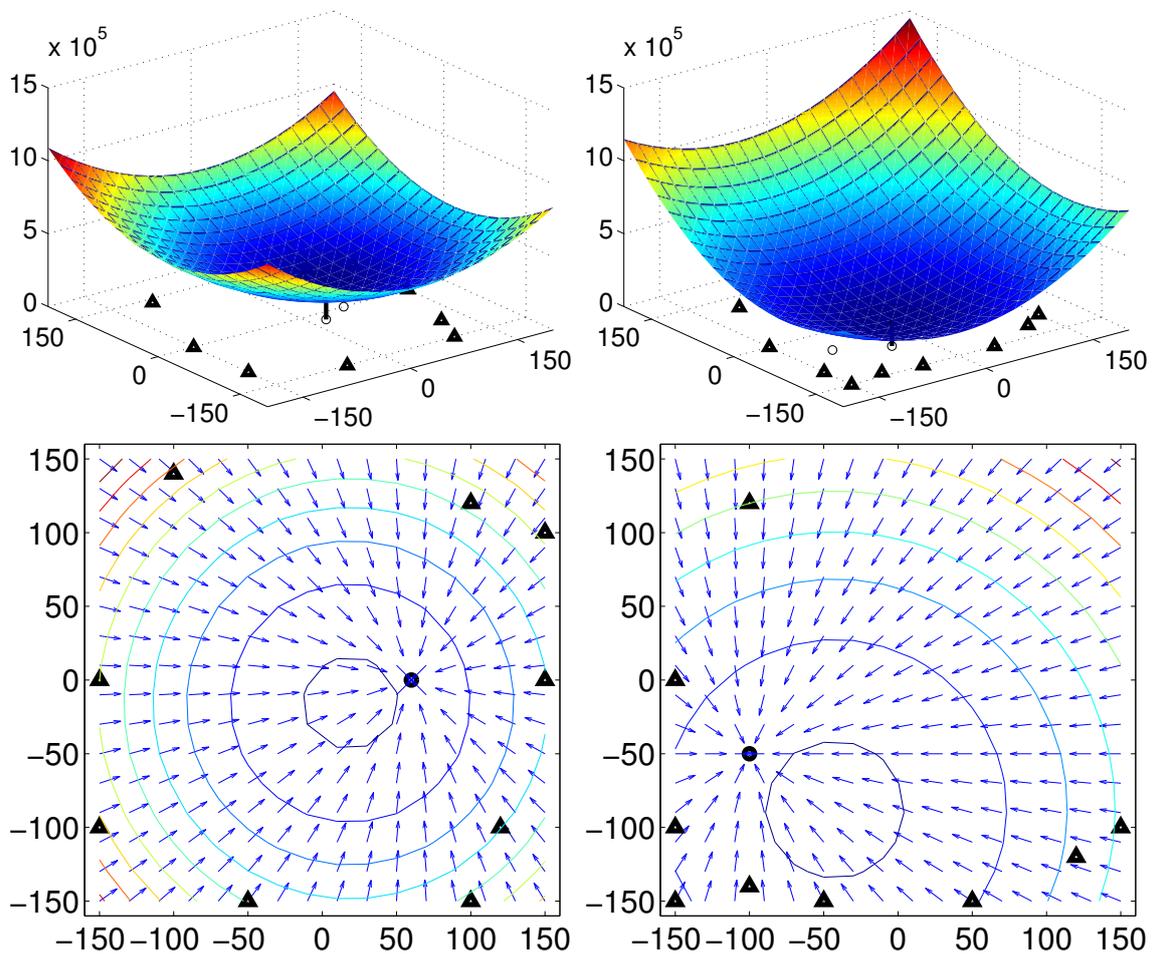


Figure 6. Test with two different simulation environments; the first row shows the moment potential, and the second row shows vector maps including the homing vector on the contour plot (triangles: landmark positions, black dot: home location).

Figure 6 shows that the moment potential has a convex shape and it has the minimal peak point. Even if the landmarks are not uniformly distributed, the moment measure has the minimal reference point. A pair of positions including the home location can share the same reference point using the

relative distances to a set of landmarks and the landmark features. In the simulation environment, there is no error to estimate the reference point, and the homing direction at each position is very accurate.

3.2. Moment Model in Real Environments with Depth Sensor Readings

We investigate the possibility of the moment model in real environments. The moment measure allows various features for landmarks in its calculation. In our experiments, a holistic approach over the landmark distribution is used, and the omnidirectional depth information in combination with visual features are encoded in the moment value.

Initially, we measured the moment with RGB color intensities in the panoramic image. The holistic view takes all the pixels as landmarks in the environment. The information about 720 pixels near the horizontal line and the corresponding depth determines a moment potential. The contour map of the moment potential is displayed in Figure 7. The range sensor readings and color intensities can change depending on the observation point. Figure 7 shows that the minimal potential positions closely match each other in the real environment, although there is a little deviation of the minimal potential position (reference point) when it is calculated at two different observation points. The measurement error or a few occlusions of landmarks can induce a small deviation of the ideal reference point, thus causing homing error potentially.

To see the effect of the two components, the color intensity and the distance of landmarks, we tested the moment model with the unit distance $r_i = 1$ or the unit feature $C_i = 1$. In the first test, only color intensity is available for landmarks under the equal distance assumption (unit distance $r_i = 1$). Figure 8a shows the result, and relatively large homing errors are observed. This implies that landmarks only with color intensity have difficulty in representing the surrounding environment. The distribution of RGB colors changes depending on the observation point, which will highly influence the estimation of the reference point, as well as the homing direction. In contrast, another test with continuous-ranged distance, but no color intensity shows much better homing performance; see Figure 8b. The depth information of landmarks greatly contributes to the estimation of the homing direction.

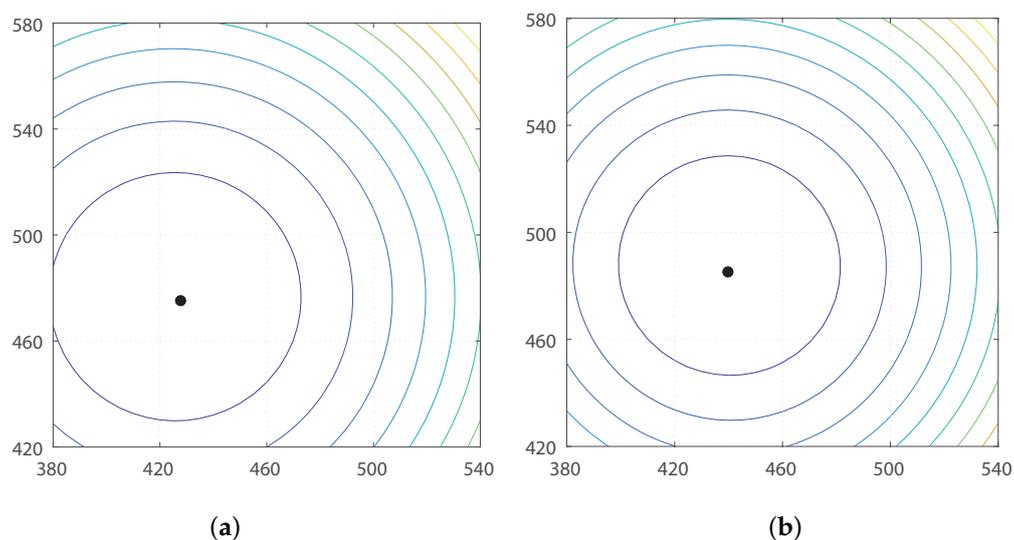


Figure 7. Contour plots at two different positions in the indoor environment *env0* (Figure 3). (a) Contour of moment potential at (460,460); (b) contour of moment potential at home position (500,500); black dots indicate the minimum potential points.

Figure 9 shows vector maps using both the distance information and color intensity when a reference compass is available. The moment model with both features shows good homing

performance at any position in the environment, mostly better than the moment model only with distance. Thus, the moment model, a combination model of distance and visual features, is a more promising approach to read the landmark distribution or characteristics in the environment.

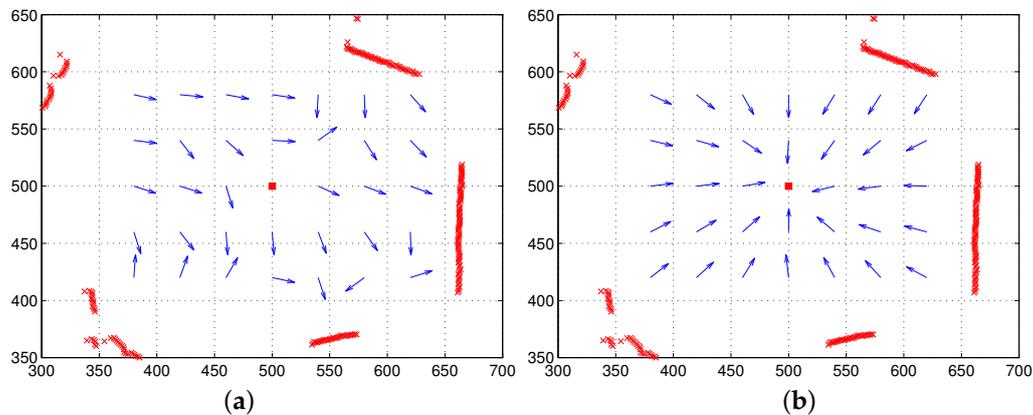


Figure 8. Homing performance only with color intensity or only with range data (reference compass available): (a) only with vision, but the unit distance ($r_i = 1$) (all landmarks have equal distances); (b) with only range data, but the unit feature ($C_i = 1$) (all landmarks have the same visual feature or same color intensity) (the red box at (500,500) indicates home, and arrows show the homing direction at each point).

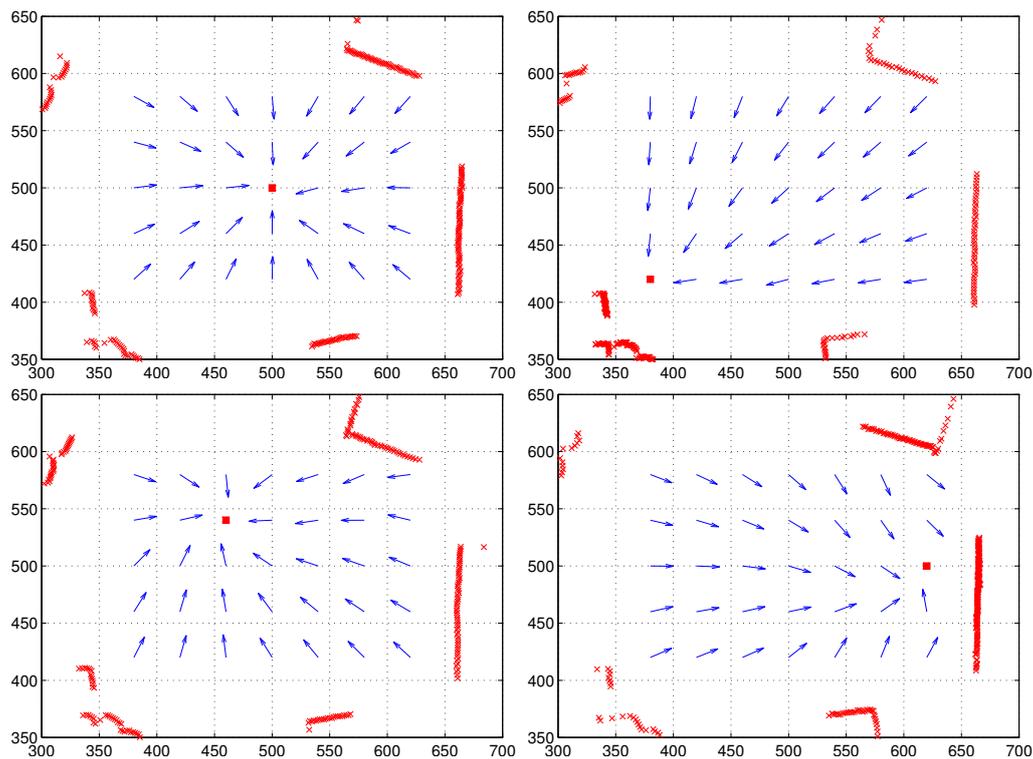


Figure 9. Homing performance with both range data and color intensity (reference compass available); the red box indicates the home position, and the arrows show the homing direction at each point.

Instead of the RGB space, we can apply one color, for example red-colored intensity, to the moment measure. In this case, a single reference point is available to derive the homing vector. Furthermore, the HSV space for one pixel can produce another visual feature, and the moment with the HSV intensity can derive different reference points; but it also produces successful homing performance.

To discretize the continuous-ranged attribute value, a cut-off threshold can be used; if a feature value is greater than the threshold, we set $C_i = 1$, otherwise $C_i = 0$. In this case, the moment measure is like extracting a special feature of landmarks, by collecting all the landmarks with $C_i = 1$. Then, the moment model can be converted to the ALV or the DELV model depending on the distance information (unit distance or continuous-ranged distance). Figure 10 shows the homing performance result depending on various patterns of the feature C_i . Even one color intensity together with depth information provides reasonable homing performance. More features or a different representation of the visual feature like the HSV space can estimate homing directions well. A choice of special features satisfying a given condition can also find homing directions; see Figure 10d.

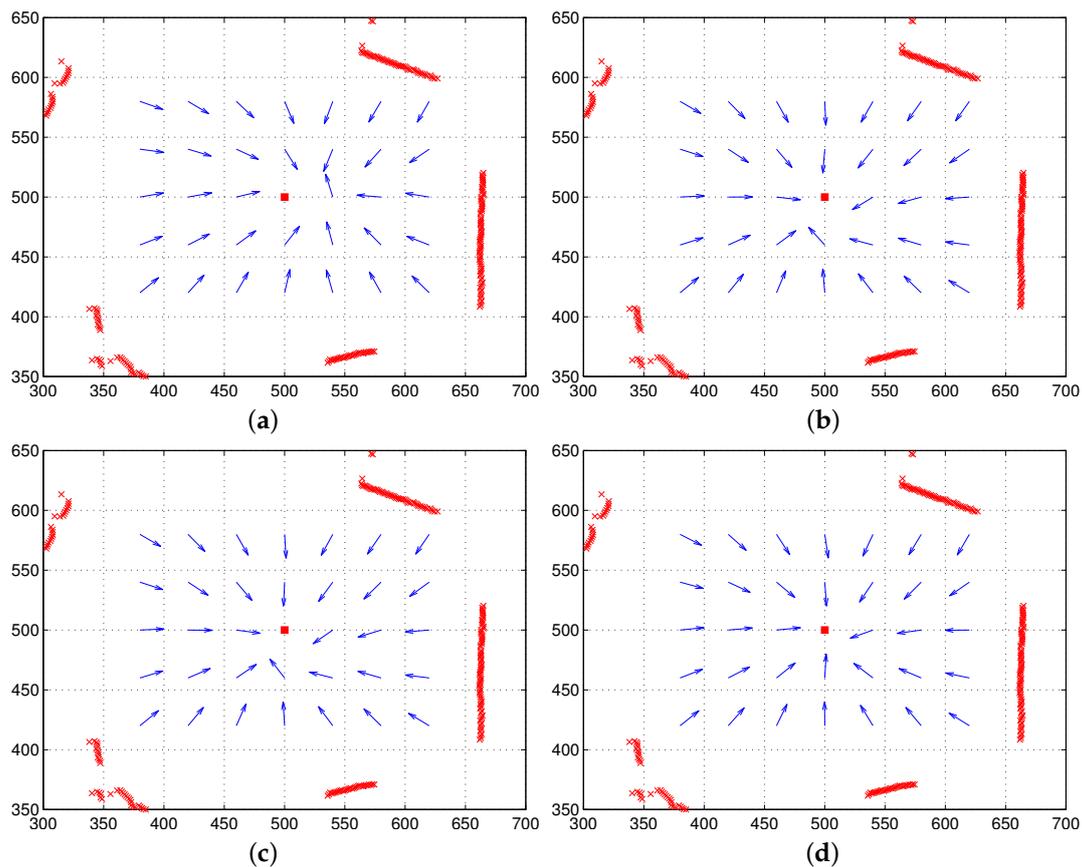


Figure 10. Homing performance with various feature values: (a) C_i is the blue color intensity; (b) C_i is the red color intensity; (c) C_i is the V intensity in HSV space; (d) $C_i = 1$ or 0 depending on the discretized condition with HSV (if $\sqrt{H^2 + S^2 + V^2} > 50$, $C_i = 1$, otherwise $C_i = 0$); the red box at (500,500) indicates home, and arrows show the homing direction at each point.

If there is no reference compass available, we need to align the coordinates for the snapshot images obtained at two positions. One of popular approaches is the visual compass approach by the pixel matching process. Another approach is the landmark rearrangement method. The methods were applied to the above environment. The moment model needs this alignment process a priori. Figure 11 shows the results. With the visual compass, there are some large angular errors at the right side corner, while the landmark rearrangement method shows relatively small angular errors towards home at all points.

To compare the moment model and the DID (Descent in Image Distance) method, we demonstrated the homing results with the DID method in Figure 12. The DID method uses only visual images to estimate the homing direction. Despite the limitation, it can estimate the homing directions reasonably. It seems that whether the DID method has a reference compass or not has no large impact on the

homing performance. Overall, the DID method has larger angular errors on average than the moment model, and this is due to lack of depth information of landmarks.

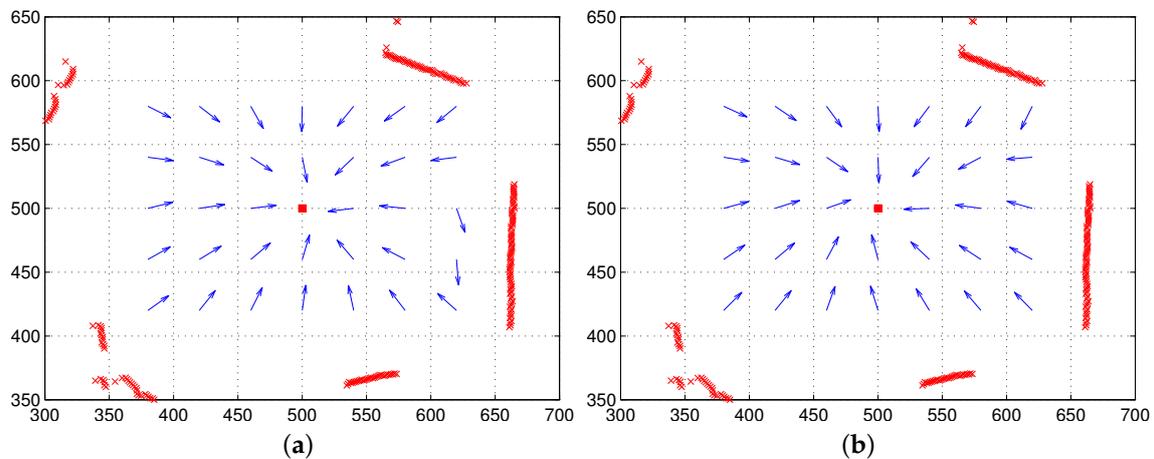


Figure 11. Homing performance of the moment model without a reference compass; both range data and color intensity are used, and the coordinate alignment process is applied: (a) visual compass; (b) landmark rearrangement method.

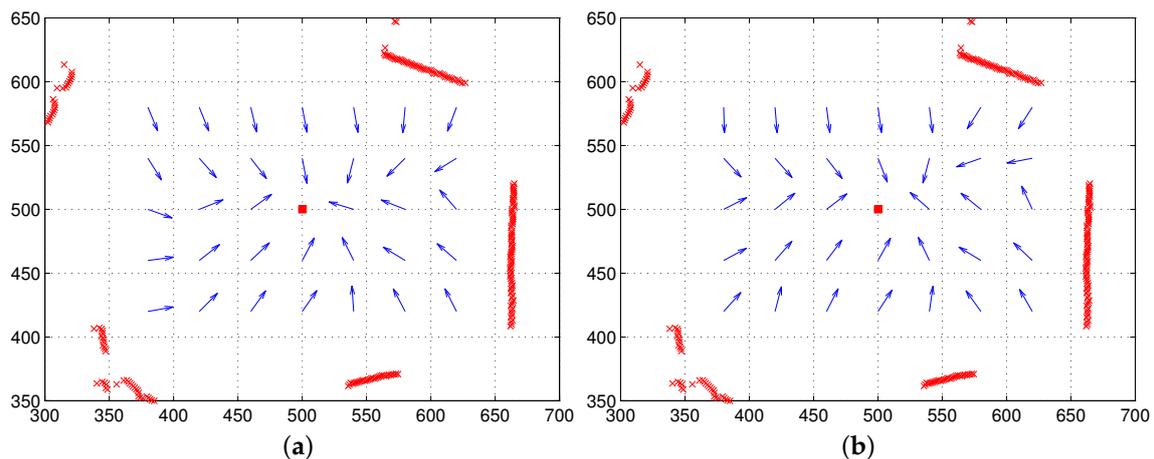


Figure 12. Homing performance with the DID approach (a) with a reference compass; (b) visual compass without a reference compass; the red box indicates the home position, and arrows show the homing direction at each point.

3.3. Moment Model in a Real Environment with Ground-Line Distance

We applied our moment model to another environment, one of Vardy's dataset called 'a1original' [36,53], which includes many panoramic snapshots, but without a distance map. The omnidirectional images were collected for the indoor environment; the arena size is $2.7 \text{ m} \times 4.3 \text{ m}$, and there are 170 points with 10-cm regular intervals for snapshots.

Figure 13 shows the homing performance results with a reference compass. Similar to the previous results with our lab environment *env0* shown in Figure 3, the moment model only with color intensity shows the worst performance. The model with ground-line distance shows very successful performance. The model with both distance and color intensity together shows a little more improved performance. Some positions have angular deviation in the direction to the goal position, which is related to the error in distance estimation. Without the laser sensor, only the visual image can determine the homing direction effectively. It can be inferred that the distance parameter in the moment model greatly influences the homing performance.

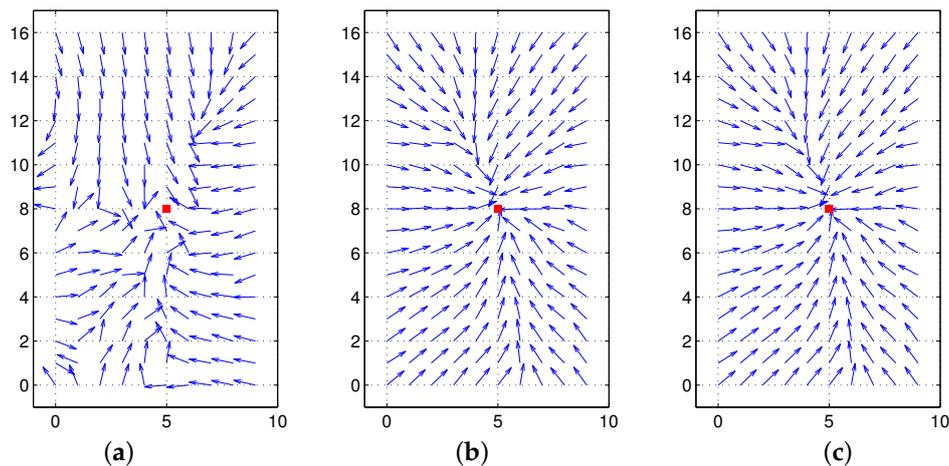


Figure 13. Homing performance for Vardy's image environment with a reference compass: (a) moment model only with color intensity; (b) moment model only with estimated distance; (c) moment model with estimated distance and color intensity; distance is estimated using the ground line in the image without the laser sensor (the arrow indicates the homing direction).

Figure 14 shows the performance depending on the coordinate alignment methods. Even without a reference compass, similar patterns of homing performance were observed. The moment model only with color intensity is insufficient to guide homing, but the model with the ground-line distance shows much better homing results. Furthermore, the landmark rearrangement method seems better than the visual compass method; see Figure 14c,d. Even when we change the home locations, the moment model with landmark rearrangement robustly works well for homing, as shown in Figure 15. With only visual snapshots, we are successful at estimating the distance to landmarks, as well as adjusting the coordinate alignment and deciding the homing direction ultimately.

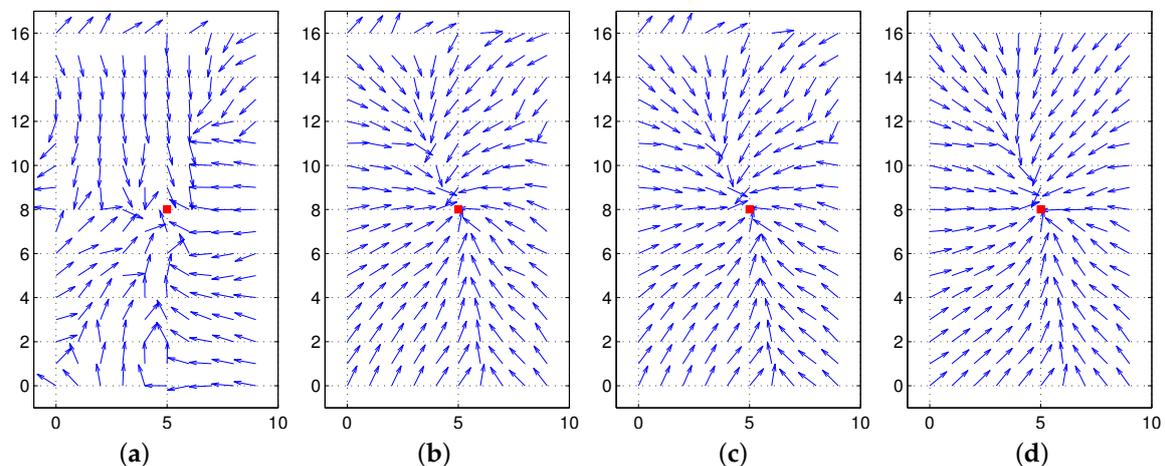


Figure 14. Homing performance for Vardy's image environment without a reference compass: (a) visual compass and moment model only with color intensity; (b) visual compass and moment model only with estimated distance; (c) visual compass and moment model with both estimated distance and color intensity; (d) landmark rearrangement method and moment model with both estimated distance and color intensity; distance is estimated using the ground line in the image without a laser sensor (arrow indicates the homing direction).

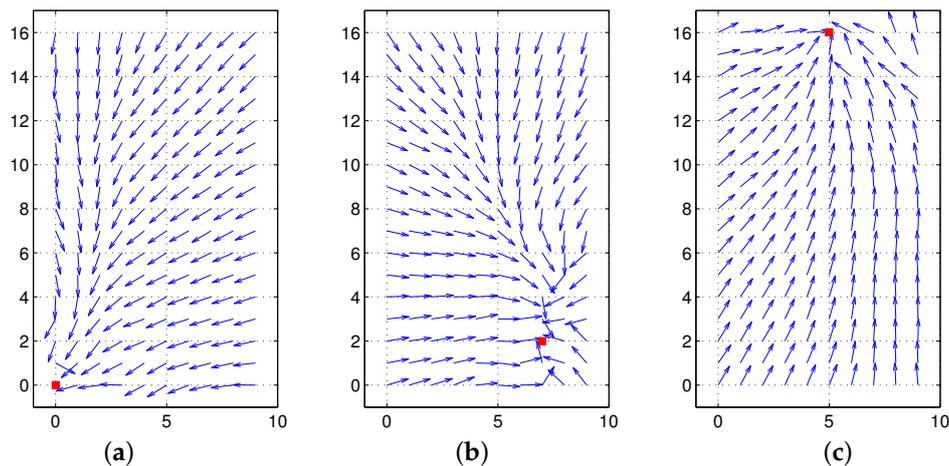


Figure 15. Homing performance for Vardy's image set without a reference compass; the landmark rearrangement method to align the coordinate and the moment model with estimated distance and color intensity are used: (a) home location (0,0); (b) home location (7,2); (c) home location (5,16).

3.4. Comparison of Homing Performances with Various Methods

We measured angular errors for homing directions with various methods listed in Table 1. We assume that the desired homing direction is the direct path from the current position to the home location. In the moment model, we encoded the depth information of landmarks and the visual features into a moment potential. We also tested the coordinate (orientation) alignment for the environment without a reference compass. The snapshot image at the home location becomes warped at the current position, and it influences the landmark distribution and the moment characteristics.

The methods *I* and *II* use only visual features to decide the homing direction, and they have large angular errors. The moment model without depth information has shortcomings, compared to the DID method (the method *I*), as we see the performance for the methods *I* and *II*. However, in the moment model, the depth information of landmarks by a laser sensor or the estimation of the ground-line distance in the visual image significantly improves the homing performance; see the methods *III–VI*. The same result can be found in both the indoor environment *env0* and Vardy's dataset. Furthermore, if there is a real measurement of the distances to landmarks, it can further improve the performance as shown in Figure 16a.

In the moment model, the orientation alignment is required for the environment without a reference compass. We found that the landmark rearrangement is quite effective in aligning the coordinates (orientations). In many cases, the average performance with the landmark rearrangement is better than that of the visual compass, although not statistically significant. The methods (*IV* and *VI*) with a combination of two components, landmark distance and visual feature, have lower error performance, when compared to the methods with only distance (*III* and *V*). The DID method is based on the visual images and their image difference, and the above experiments imply that visual information alone seems insufficient to guide homing. The DID is significantly worse in homing performance than the moment model with the visual feature and the depth information.

Tables 2 and 3 show angular errors for homing direction and 95% confidence intervals (assuming a *t*-distribution). Mostly, the errors are within 45 degrees. The method *II* with the visual feature alone shows a relatively large portion of angular errors greater than 45 degrees in the two different environments. The results are consistent with the performance shown in Figure 16. Especially, the DID method shows somewhat poor performance in Vardy's image set without a reference compass.

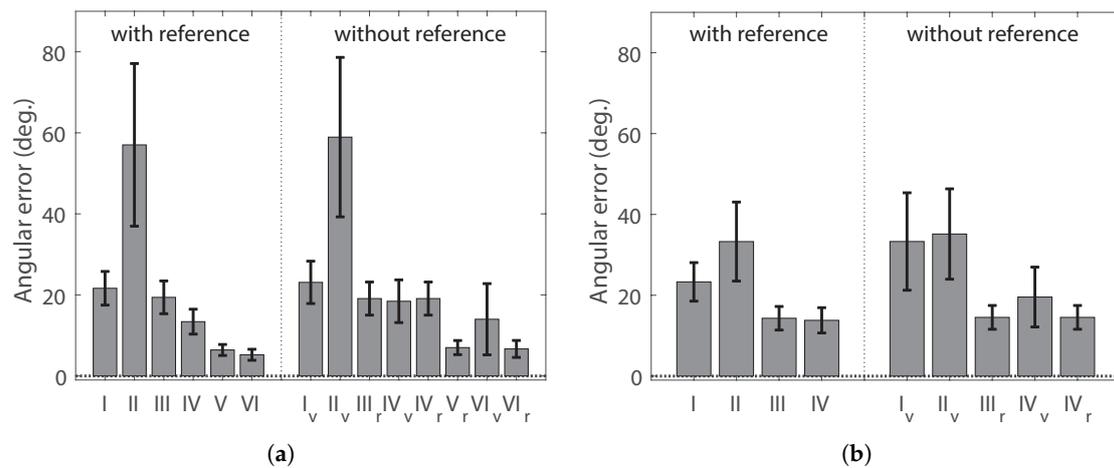


Figure 16. Homing performance with a reference compass and without a reference compass; the x-axis indicates testing methods described in Table 1: (a) angular errors in the indoor environment *env0*; (b) angular errors in Vardy's image set; error bars indicate 95% confidence intervals by assuming a *t*-distribution.

Table 2. Angular errors with various methods for the indoor environment *env0*; testing methods are described in Table 1 (*N* is the number of testing points, μ the average of angular errors, σ the 95% confidence interval (assuming a *t*-distribution)).

| Method | | error $\mu(\pm\sigma)$ deg. | N | $0 \leq \varepsilon_{\theta} < 45^{\circ}$ | $45 \leq \varepsilon_{\theta} < 90^{\circ}$ | $90 \leq \varepsilon_{\theta} < 180^{\circ}$ |
|---------------------------------|---------|-----------------------------|----|--|---|--|
| with reference compass | I | 21 (± 4.2) | 34 | 97.1% | 2.9% | 0% |
| | II | 57 (± 20.1) | 34 | 67.6% | 5.9% | 26.5% |
| | III | 19.4 (± 4.0) | 34 | 97.1% | 2.9% | 0% |
| | IV | 13.4 (± 3.1) | 34 | 100% | 0% | 0% |
| | V | 6.4 (± 1.4) | 34 | 100% | 0% | 0% |
| | VI | 5.7 (± 1.4) | 34 | 100% | 0% | 0% |
| without reference compass | I_v | 23 (± 5.2) | 34 | 94.1% | 5.9% | 0% |
| | II_v | 59 (± 19.7) | 34 | 58.8% | 14.7% | 26.5% |
| | III_r | 19.1 (± 4.1) | 34 | 97.1% | 2.9% | 0% |
| | IV_v | 18.4 (± 5.3) | 34 | 94.1% | 5.9% | 0% |
| | IV_r | 19.1 (± 4.1) | 34 | 97.1% | 2.9% | 0% |
| | V_r | 7 (± 1.7) | 34 | 100% | 0% | 0% |
| | VI_v | 14 (± 8.8) | 34 | 94.1% | 0% | 5.9% |
| | VI_r | 7 (± 2.1) | 34 | 100% | 0% | 0% |

Table 3. Angular errors with various methods for Vardy's image set; testing methods are described in Table 1 (*N* is the number of testing points, μ the average of angular errors and σ the 95% confidence interval).

| Method | | error $\mu(\pm\sigma)$ deg. | N | $0 \leq \varepsilon_{\theta} < 45^{\circ}$ | $45 \leq \varepsilon_{\theta} < 90^{\circ}$ | $90 \leq \varepsilon_{\theta} < 180^{\circ}$ |
|---------------------------------|---------|-----------------------------|-----|--|---|--|
| with reference compass | I | 23.31 (± 4.8) | 169 | 97.1% | 2.9% | 0% |
| | II | 33.28 (± 9.8) | 169 | 72.19% | 24.26% | 3.55% |
| | III | 14.31 (± 2.9) | 169 | 100% | 0% | 0% |
| | IV | 13.8 (± 3.1) | 169 | 100% | 0% | 0% |
| without reference compass | I_v | 33.30 (± 12.1) | 169 | 79.88% | 11.83% | 8.29% |
| | II_v | 35.15 (± 11.2) | 169 | 69.82% | 23.08% | 7.1% |
| | III_r | 14.53 (± 2.9) | 169 | 100% | 0% | 0% |
| | IV_v | 19.56 (± 7.4) | 169 | 96.45% | 0% | 3.55% |
| | IV_r | 14.53 (± 2.9) | 169 | 100% | 0% | 0% |

4. Discussion

In the current approach, we used a holistic approach to take all the pixels as landmarks. Each pixel produces a landmark vector with its own angular position, instead of an object in the real environment. If an object is close to the observer, it creates many vectors, while a small number of vectors is assigned for an object far away from the observer. Summing these landmark vectors may be different from an object representation in the real environment. It is possible that the image segmentation or clustering process can help with identifying objects from the image. Then, the corresponding object features and the distances of objects in the environment can be encoded with the moment model. In other words, a sophisticated feature extraction algorithm can be combined with the suggested moment model. We need further work to check if this feature extraction approach will be better than the holistic approach. A basic assumption in this moment model is that invariant features can be observed at any position, and a collection of the features can represent the environment well, while its centroid can be localized as a reference point.

In our experiments, the RGB color intensity in the image view can be changed depending on the observation point. The visual images can be affected by the illumination, the angle and position of the light source, as well as glint on the floor. If these changes are intense, then the mobile robot cannot estimate the homing direction accurately. Thus, the robust features available or a good measurement apparatus will be helpful for homing accuracy. We believe that the range sensor has high accuracy of measuring the distance, but the visual image has relatively noisy sensor readings for pixel-wise landmarks. An ensemble of the two measurements seems to compensate one for the other. Possibly alternative features would be helpful to improve the homing performance. An invariant feature such as landmark height rather than the color intensity can be supportive to read the environmental information better, which can play the role of a milestone to guide homing.

The suggested moment function models a distribution of landmarks with their landmark characteristics. If we take a model with $M = \sum C_i/r_i$, the potential has a high peak at the landmark position, which can be used as a collision avoidance model or landmark search model. The potential value will rapidly decay to zero, if the measuring position becomes far from the landmark. If a mobile robot senses a potential value greater than a threshold, then it can take the behavior of avoiding the landmark or moving towards the landmark. This is another possible application of the moment model.

In our local homing navigation, we assume that landmarks are commonly observed at any position. The suggested moment approach is based on the snapshot model, which works only in the isotropic environment, and an agent can search for the target position as the home position, starting from the current position. Our local navigation approach can be extended to a wide range of navigation in complex cluttered environments with many occlusions of landmarks. Multiple reference images as milestones can guide the right direction to the goal position [37], even if the goal position is far way. That is, local homing navigation can be applied to each waypoint. A sequence of waypoint searches may lead to the final goal position. For future work, we will test this approach in cluttered environments to require complex homing paths.

5. Conclusions

In this paper, we suggest a new navigation method based on the moment model to characterize the landmark distribution and features. The moment model is inspired by the moment of inertia in physics, and it sees how the landmarks and their features are distributed in a given environment. Here, the landmark features and the relative distance are encoded in the moment function. The moment model allows multiple features, which can help estimate the homing direction more robustly. We proved that the landmark distribution has a unique minimum peak of moment potential, and it can be a reference point as an invariant feature at any position, which corresponds to the center of mass in physics. In the moment model, the homing vector is calculated based on this reference point.

In the experiments, we used the RGB color intensity in visual images as the features in the moment model. The distance was measured by a laser sensor, but without the depth sensor, the

ground-line distance estimated in the panoramic image was tested as alternative distance information. The two components, the depth information and the features, highly contribute to successful homing performance, which is distinguished from other approaches' results based on vision images. Especially, the depth information cannot be neglected for good homing performance.

Our approach is a holistic approach to use all the pixels as landmarks. We assumed that the environment is isotropic, that is the majority of landmarks is commonly observed in view at any position in the arena. If the mobile robot moves far away from the home location, the environment landscape will change greatly, and many occlusions occur, which may violate the isotropic assumption. In that case, the homing performance may degrade to a great extent, since the centroid of the moment value greatly migrates to another point. Thus, the moment model is appropriate for local homing navigation. We need further study to handle the problem or to cover a long-range path problem.

In the moment model, we simply applied the RGB color intensity as the visual feature. However, more sophisticated features, such as invariant features, can be developed in the image, which will further improve the homing performance. For future work, we will find what kind of visual features or other features will be useful in the moment model. Our model can be extended to various characteristics that are not easily changeable by the change of measuring positions. If the invariant property of the feature is preserved, the feature can be encoded in the moment model. Thus, various forms of the moment model can be produced.

Acknowledgments: This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (Ministry of Education, Science and Technology (MEST)) (No. 2017R1A2B4011455, 2014R1A2A1A11053839)

Author Contributions: Changmin Lee performed the experiments and analyzed the data, and DaeEun Kim provided suggestions for the paper. The manuscript was written by Changmin Lee and revised by DaeEun Kim.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Widyotriatmo, A.; Hong, K.S. Navigation function-based control of multiple wheeled vehicles. *IEEE Trans. Ind. Electron.* **2011**, *58*, 1896–1906.
2. Zhang, T.; Huang, S.; Liu, D.; Shi, L.; Zhou, C.; Xiong, R. A method of state estimation for underwater vehicle navigation around a cylindrical structure. In Proceedings of the 2016 IEEE 11th Conference on Industrial Electronics and Applications (ICIEA), Hefei, China, 5–7 June 2016; pp. 101–106.
3. Zhu, B.; Zaini, A.H.B.; Xie, L. Distributed guidance for interception by using multiple rotary-wing unmanned aerial vehicles. *IEEE Trans. Ind. Electron.* **2017**, *64*, 5648–5656.
4. Huang, X.; Yan, Y. Output Feedback Control of Underactuated Spacecraft Hovering in Circular Orbit With Radial or In-Track Controller Failure. *IEEE Trans. Ind. Electron.* **2016**, *63*, 5569–5581.
5. Lyshevski, S.E. Signal processing in cyber-physical MEMS sensors: Inertial measurement and navigation systems. *IEEE Trans. Ind. Electron.* **2017**, *64*, 9618–9625.
6. Delgado, F.J.; Quero, J.M.; Garcia, J.; Tarrida, C.L.; Ortega, P.R.; Bermejo, S. Accurate and wide-field-of-view MEMS-based sun sensor for industrial applications. *IEEE Trans. Ind. Electron.* **2012**, *59*, 4871–4880.
7. Gilg, A.; Schmidt, G. Landmark-oriented visual navigation of a mobile robot. *IEEE Trans. Ind. Electron.* **1994**, *41*, 392–397.
8. Wang, W.; Xie, G. Online high-precision probabilistic localization of robotic fish using visual and inertial cues. *IEEE Trans. Ind. Electron.* **2015**, *62*, 1113–1124.
9. Park, S.; Hashimoto, S. Autonomous mobile robot navigation using passive RFID in indoor environment. *IEEE Trans. Ind. Electron.* **2009**, *56*, 2366–2373.
10. Cartwright, B.; Collett, T. Landmark learning in bees. *J. Comp. Physiol. A* **1983**, *151*, 521–543.
11. Darwin, C. Origin of certain instincts. *Nature* **1873**, *7*, 417–418.
12. Kimchi, T.; Etienne, A.; Terkel, J. A subterranean mammal uses the magnetic compass for path integration. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 1105.
13. Vickerstaff, R.; Merkle, T. Path integration mediated systematic search: A Bayesian model. *J. Theor. Biol.* **2012**, *307*, 1–19.

14. Wintergerst, S.; Ronacher, B. Discrimination of inclined path segments by the desert ant *Cataglyphis fortis*. *J. Comp. Physiol. A* **2012**, *198*, 363–373.
15. Etienne, A.; Jeffery, K. Path integration in mammals. *Hippocampus* **2004**, *14*, 180–192.
16. Reid, S.; Narendra, A.; Hemmi, J.; Zeil, J. Polarised skylight and the landmark panorama provide night-active bull ants with compass information during route following. *J. Exp. Biol.* **2011**, *214*, 363–370.
17. Basten, K.; Mallot, H. Simulated visual homing in desert ant natural environments: efficiency of skyline cues. *Biol. Cybern.* **2010**, *102*, 413–425.
18. Wystrach, A.; Beugnon, G.; Cheng, K. Ants might use different view-matching strategies on and off the route. *J. Exp. Biol.* **2012**, *215*, 44–55.
19. Kirchner, W.; Braun, U. Dancing honey bees indicate the location of food sources using path integration rather than cognitive maps. *Anim. Behav.* **1994**, *48*, 1437–1441.
20. Garm, A.; Oskarsson, M.; Nilsson, D. Box jellyfish use terrestrial visual cues for navigation. *Curr. Biol.* **2011**, doi:10.1016/j.cub.2011.03.054.
21. Zeil, J.; Hemmi, J. The visual ecology of fiddler crabs. *J. Comp. Physiol. A* **2006**, *192*, 1–25.
22. Garm, A.; Nilsson, D.E. Visual navigation in starfish: first evidence for the use of vision and eyes in starfish. *Proc. R. Soc. B* **2014**, *281*, 20133011.
23. Collett, M.; Collett, T. How do insects use path integration for their navigation? *Biol. Cybern.* **2000**, *83*, 245–259.
24. Ugolini, A.; Borgioli, G.; Galanti, G.; Mercatelli, L.; Hariyama, T. Photoresponses of the Compound Eye of the Sandhopper *Talitrus saltator* (Crustacea, Amphipoda) in the Ultraviolet-Blue Range. *Biol. Bull.* **2010**, *219*, 72–79.
25. Steck, K.; Knaden, M.; Hansson, B. Do desert ants smell the scenery in stereo? *Anim. Behav.* **2010**, *79*, 939–945.
26. Gould, J.L. Animal Navigation: Birds Have Magnetic Maps. *Curr. Biol.* **2015**, *25*, R836–R838.
27. Möller, R. Local visual homing by warping of two-dimensional images. *Robot. Auton. Syst.* **2009**, *57*, 87–101.
28. Lambrinos, D.; Moller, R.; Labhart, T.; Pfeifer, R.; Wehner, R. A mobile robot employing insect strategies for navigation. *Robot. Auton. Syst.* **2000**, *30*, 39–64.
29. Ramisa, A.; Goldhoorn, A.; Aldavert, D.; Toledo, R.; de Mantaras, R. Combining Invariant Features and the ALV Homing Method for Autonomous Robot Navigation Based on Panoramas. *J. Intell. Robot. Syst.* **2011**, *64*, 625–649.
30. Smith, L.; Philippides, A.; Graham, P.; Baddeley, B.; Husbands, P. Linked local navigation for visual route guidance. *Adapt. Behav.* **2007**, *15*, 257–271.
31. Weber, K.; Venkatesh, S.; Srinivasan, M. Insect-inspired robotic homing. *Adapt. Behav.* **1999**, *7*, 65–97.
32. Yu, S.E.; Kim, D. Landmark vectors with quantized distance information for homing navigation. *Adapt. Behav.* **2011**, *19*, 121–141.
33. Yu, S.; Lee, C.; Kim, D. Analyzing the effect of landmark vectors in homing navigation. *Adapt. Behav.* **2012**, *20*, 337–359.
34. Denuelle, A.; Strydom, R.; Srinivasan, M.V. Snapshot-based control of UAS hover in outdoor environments. In Proceedings of the 2015 IEEE International Conference on Robotics and Biomimetics (ROBIO), Zhuhai, China, 6–9 December 2015; pp. 1278–1284.
35. Zeil, J.; Hofmann, M.; Chahl, J. Catchment areas of panoramic snapshots in outdoor scenes. *J. Opt. Soc. Am. A* **2003**, *20*, 450–469.
36. Möller, R.; Vardy, A.; Kreft, S.; Ruwisch, S. Visual homing in environments with anisotropic landmark distribution. *Auton. Robot.* **2007**, *23*, 231–245.
37. Vardy, A. Long-range visual homing. In Proceedings of the IEEE International Conference on Robotics and Biomimetics, Kunming, China, 17–20 December 2006; pp. 220–226.
38. Churchill, D.; Vardy, A. Homing in scale space. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, Nice, France, 22–26 September 2008; pp. 1307–1312.
39. Churchill, D.; Vardy, A. An orientation invariant visual homing algorithm. *J. Intell. Robot. Syst.* **2013**, *77*, 3–29.
40. Denuelle, A.; Thurrowgood, S.; Kendoul, F.; Srinivasan, M.V. A view-based method for local homing of unmanned rotorcraft. In Proceedings of the 2015 6th International Conference on Automation, Robotics and Applications (ICARA), Queenstown, New Zealand, 17–19 February 2015; pp. 443–449.

41. Franz, M.; Schölkopf, B.; Mallot, H.; Bühlhoff, H. Where did I take that snapshot? Scene-based homing by image matching. *Biol. Cybern.* **1998**, *79*, 191–202.
42. Franz, M. Minimalistic Visual Navigation = Minimalistische Visuelle Navigation. Ph.D. Thesis, Universität Tübingen, Tübingen, Germany, 1999.
43. Hong, J.; Tan, X.; Pinette, B.; Weiss, R.; Riseman, E. Image-based homing. *IEEE Control Syst. Mag.* **1992**, *12*, 38–45.
44. Möller, R.; Krzykawski, M.; Gerstmayr, L. Three 2D-warping schemes for visual robot navigation. *Auton. Robot.* **2010**, *29*, 253–291.
45. Labrosse, F. Short and long-range visual navigation using warped panoramic images. *Robot. Auton. Syst.* **2007**, *55*, 675–684.
46. Stürzl, W.; Möller, R. An insect-inspired active vision approach for orientation estimation with panoramic images. In *Bio-Inspired Modeling of Cognitive Tasks*; Springer: Berlin/Heidelberg, Germany, 2007; pp. 61–70.
47. Möller, R. *A SIMD Implementation of the MinWarping Method for Local Visual Homing*; Technical Report; Computer Engineering, Faculty of Technology, Bielefeld University: Bielefeld, Germany, 2016.
48. Möller, R. *Column Distance Measures and their Effect on Illumination Tolerance in MinWarping*; Technical Report; University of Bielefeld, Faculty of Technology, Computer Engineering Group: Bielefeld, Germany, 2016.
49. Fler, D.; Möller, R. Comparing holistic and feature-based visual methods for estimating the relative pose of mobile robots. *Robot. Auton. Syst.* **2017**, *89*, 51–74.
50. Lee, C.; Yu, S.E.; Kim, D. Landmark-Based Homing Navigation Using Omnidirectional Depth Information. *Sensors* **2017**, *17*, 1928.
51. Yu, S.E.; Kim, D. Image-based homing navigation with landmark arrangement matching. *Inf. Sci.* **2011**, *181*, 3427–3442.
52. Kwon, T.; Song, J. A new feature commonly observed from air and ground for outdoor localization with elevation map built by aerial mapping system. *J. Field Robot.* **2011**, *28*, 227–240.
53. Vardy, A.; Möller, R. Biologically plausible visual homing methods based on optical flow techniques. *Connect. Sci.* **2005**, *17*, 47–89.



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).