

Article

# A Novel Robot Visual Homing Method Based on SIFT Features

Qidan Zhu, Chuanjia Liu \* and Chengtao Cai

College of Automation, Harbin Engineering University, Harbin 150001, China;

E-Mails: zhuqidan@hrbeu.edu.cn (Q.Z.); caichengtao@hrbeu.edu.cn (C.C.)

\* Author to whom correspondence should be addressed; E-Mail: b211040007@sina.com;  
Tel./Fax: +86-451-8258-9656.

Academic Editor: Vittorio M. N. Passaro

Received: 2 July 2015 / Accepted: 9 October 2015 / Published: 14 October 2015

---

**Abstract:** Warping is an effective visual homing method for robot local navigation. However, the performance of the warping method can be greatly influenced by the changes of the environment in a real scene, thus resulting in lower accuracy. In order to solve the above problem and to get higher homing precision, a novel robot visual homing algorithm is proposed by combining SIFT (scale-invariant feature transform) features with the warping method. The algorithm is novel in using SIFT features as landmarks instead of the pixels in the horizon region of the panoramic image. In addition, to further improve the matching accuracy of landmarks in the homing algorithm, a novel mismatching elimination algorithm, based on the distribution characteristics of landmarks in the catadioptric panoramic image, is proposed. Experiments on image databases and on a real scene confirm the effectiveness of the proposed method.

**Keywords:** visual homing; mismatching elimination; robot navigation; dynamic environment; catadioptric panoramic image

---

## 1. Introduction

The local navigation methods based on visual information (known as local visual homing) have attracted much attention in the mobile robot field [1–3]. Inspired by biological navigation, local visual homing provides the ability to guide the robot to return to a goal position only by relying on visual information [4–6]. It can calculate a homing direction by comparing the differences between the current image and the goal image, and in that homing direction, the robot can move from the current

position to the goal position [7–10]. Note that neither the current image nor the goal image directly provide the depth information of the environment. By combining with the environmental topological map, the local visual homing algorithm can divide the complex large-scale navigation problem into a series of local navigation problems, which are easier to solve [11,12]. In the topological framework, the algorithm is used to solve navigation problems between adjacent connected nodes. Most previous research on visual homing was carried out under the assumption that environments are static. However, changes of the environment (objects, illumination) often happen in a real scene. Despite recent advances, the tolerance against changes in the environment is still widely unresolved [12,13]. So far, most studies of visual homing are limited to indoor environments. If the problem mentioned above can be resolved effectively, it will be beneficial to implement visual homing in more challenging outdoor environments [14,15].

Most local visual homing algorithms can be divided into the following four categories [16–18]. The first class is correspondence-based homing methods. They use feature extraction and matching to set up a set of correspondence vectors between the current image and the goal image. These vectors can be transformed into movement directions, and an overall homing vector can be obtained by combining these movement directions [19–22]. The second class is the DID (descent in image distances) methods. In these methods, the homing can be achieved by calculating the gradient descent direction in the image distance between the current image and the goal image [14,15,23,24]. The third class is the method based on the ALV (average landmark vector) model. ALV is a unit representation vector of a certain position. The homing direction can be acquired by the subtraction of ALV between the current position and the goal position [25–27]. The last class is a robust method, which is known as warping, despite the large amount of calculation. This method supposes that the robot performs virtual movement at the goal position according to the three motion parameters. These parameters separately describe the direction, distance and rotation of the robot movement from the goal position to its current position. The goal image is distorted according to the motion parameters, after which it is compared to the current image. The optimal parameter set can be found when the differences between the two images are minimal. The homing vector can be yielded by the optimal parameter set [13,28–30].

In the above four classes of visual homing methods, DID and ALV fail to reach the performance of the warping method [13,20]. Although the homing performance of some correspondence-based methods is superior, most of them need an external compass to adjust the current image and goal image to the same horizontal orientation in order to compute the homing vector [30]. However, warping does not need an external compass to align the orientations between the two images, as the change of orientation is one of the search parameters in warping. In conclusion, warping is an attractive visual homing method. Despite the above advantages, some problems of warping still remain to be solved. Firstly, the homing performance of warping is influenced greatly by the change of environment (objects, illumination), as it determines the homing vector according to the differences of the gray value between corresponding pixels of the horizon region in the current image and the goal image. Secondly, it is difficult to balance the homing accuracy and the amount of calculation. The homing accuracy of the warping method partly depends on the search interval of the parameter space. The smaller the search interval is, the higher the homing accuracy is. However, at the same time, the computation will also increase significantly. Although Möller *et al.* have improved the warping

method, extending it to operate directly on two-dimensional images [30] and relaxing the equal distance assumption [13], there is no effective solution to the above two problems.

Motivated by the above problems, we propose a novel visual homing algorithm by combining the SIFT (scale-invariant feature transform) features with the warping method. Compared with the warping method, the proposed homing algorithm uses the SIFT features instead of the pixels of the horizon region in warping as landmarks, and its advantages can be stated as follows: (1) by using the good stability of SIFT features in translation, rotation, scaling, occlusion and illumination [31], the novel algorithm is more robust to the influence caused by the changes of the environment and has a better homing accuracy; (2) the proposed homing algorithm can resolve the conflict between the homing accuracy and the amount of calculation, because it can obtain the homing vector by solving a system of ternary equations in place of doing the exhaustive search in the parameter space. In addition, unlike most visual homing algorithms that need to unwarp the initial panoramic images, the proposed homing algorithm directly operates on the initial catadioptric panoramic images. As a result, the additional amount of calculation is reduced. It is according to the angle change of landmarks between the current image and the goal image that our homing algorithm computes the homing direction, so the matching accuracy of landmarks is crucial. Because of the complex imaging relations of catadioptric panoramic images, most existing mismatching elimination algorithms cannot directly operate on the initial panoramic images [32–34]. For the above reasons, in order to further improve the matching accuracy of landmarks, a novel mismatching elimination algorithm is proposed on the basis of the distribution characteristics of landmarks in the catadioptric panoramic image.

The remainder of the paper is organized as follows: Section 2 describes the design of the proposed homing algorithm. The selection of landmarks and the design of the proposed mismatching elimination algorithm are introduced in Section 3. Section 3 also shows the overview of the proposed homing method. The results and analysis of the experiments are mainly presented in Section 4. Section 5 draws conclusions and points out the focus of future research.

## 2. Homing Algorithm

For the derivation of the warping method, refer to [28]. Similar to the warping method, our homing algorithm is also based on the equal distance assumption, which assumes that all of the surrounding landmarks have identical distances from the goal position. Although this assumption is usually violated, Franz *et al.* have proven that the error due to this assumption decreases when the robot approaches the goal [28].

The geometric relations of the homing algorithm are shown in Figure 1. The goal position and the current position of the robot are indicated separately by H and C. L is a landmark in the scene. The distance from L and H is denoted by  $r$ . The initial orientation of the robot at position H is shown by  $O_H$ . The robot moves away from H in direction  $\alpha$  relative to its initial orientation  $O_H$  and covers a distance  $d$  to C. After the movement, the initial orientation changes by  $\psi$ . The new orientation of the robot is shown by  $O_C$ . The dashed arrow at position C indicates the old orientation. In order to facilitate the derivation, we assume that the landmark L corresponds to some feature in the real scene. The angle between landmark L and orientation  $O_H$  is denoted by  $\theta$ . As the robot moves, the angle  $\theta$

changes to  $\theta'$ , that is the angle between landmark L and new orientation  $O_C$ . Applying the law of sines to the triangle LHC, we can obtain the following equation:

$$\frac{\sin(\psi + \theta' - \theta)}{d} = \frac{\sin(\pi - [\psi + \theta' - \alpha])}{r} \tag{1}$$

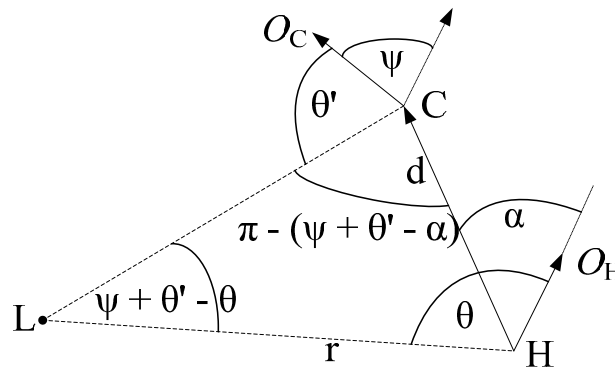


Figure 1. Derivation of the proposed homing algorithm.

Equation (1) can be rearranged as follows:

$$\sin(\psi + \theta' - \theta) = \rho \sin(\psi + \theta' - \alpha) \tag{2}$$

where  $\rho = d/r$ . There are three unknown parameters ( $\rho, \psi, \alpha$ ) in Equation (2).  $\psi$  and  $\alpha$  are invariable for all of the landmarks in a certain scene. Suppose there are three landmarks L1, L2 and L3 in the scene. The angles between these landmarks and the orientation of the robot are separately denoted by  $\theta_1, \theta_2$  and  $\theta_3$  in the goal image. In the current image, the corresponding angles are denoted by  $\theta'_1, \theta'_2$  and  $\theta'_3$ . Substitute  $(\theta_1, \theta'_1), (\theta_2, \theta'_2)$  and  $(\theta_3, \theta'_3)$  into Equation (2), and we can obtain the following equations:

$$\begin{cases} \sin(\theta'_1 - \theta_1 + \psi) = \rho_1 \sin(\theta'_1 - \alpha + \psi) \\ \sin(\theta'_2 - \theta_2 + \psi) = \rho_2 \sin(\theta'_2 - \alpha + \psi) \\ \sin(\theta'_3 - \theta_3 + \psi) = \rho_3 \sin(\theta'_3 - \alpha + \psi) \end{cases} \tag{3}$$

Before Equation (3) can be applied for homing, two problems have to be explained:

1. The distance  $d$  is identical for all of the landmarks in the scene when the robot arrives at current position C from goal position H. According to the equal distance assumption, the distance  $r$  is also identical for all of the landmarks. Based on the above two conditions, there exist the relation  $\rho_1 = \rho_2 = \rho_3$ .

2. The angles  $\theta_1, \theta_2, \theta_3$  and  $\theta'_1, \theta'_2, \theta'_3$  can be worked out if the positions of landmarks L1, L2 and L3 in the panoramic image are known.

By solving Equations (3), we can get a parameter set  $(\rho, \psi, \alpha)$ . The homing direction  $\beta$  relative to the new orientation of the robot can be computed as follows:

$$\beta = \pi + \alpha - \psi \tag{4}$$

It can be seen from Equations (3) to (4) that the homing direction can be determined by only three landmarks. As the number of landmarks is usually more than three in the real scene, a number of

parameter sets, such as  $(\rho_1, \psi_1, \alpha_1), (\rho_2, \psi_2, \alpha_2) \dots (\rho_n, \psi_n, \alpha_n)$ , can be obtained by solving Equation (3). In order to get the optimal parameter  $\hat{\alpha}$ , we first give the definition of the sum of squares of deviations:

$$\alpha_{SSD} = \sum_{i=1}^n \text{diff}(\alpha_i - \bar{\alpha})^2 \quad (5)$$

where  $n$  is the number of parameter sets,  $\bar{\alpha}$  changes from  $-\pi$  to  $\pi$  and  $\text{diff}(\alpha_i - \bar{\alpha})$  yields the difference between two angles, which is defined as:

$$\text{diff}(\alpha_i - \bar{\alpha}) = \begin{cases} 2\pi - |\alpha_i - \bar{\alpha}|, & |\alpha_i - \bar{\alpha}| \geq \pi \\ |\alpha_i - \bar{\alpha}|, & |\alpha_i - \bar{\alpha}| < \pi \end{cases} \quad (6)$$

According to Equation (5), we acquire  $\hat{\alpha} = \bar{\alpha}_{SSD\min}$ , and  $\bar{\alpha}_{SSD\min}$  is the value of  $\bar{\alpha}$  when  $\alpha_{SSD}$  is minimum. The calculation of  $\hat{\psi}$  is the same as that of  $\hat{\alpha}$ . The final homing direction  $\hat{\beta}$  can be determined as follows:

$$\hat{\beta} = \pi + \hat{\alpha} - \hat{\psi} \quad (7)$$

### 3. Landmark Optimization and Overview of the Proposed Method

For the purpose of guiding the robot to return to the goal position, all of the visual homing algorithms need to acquire accurate and reliable information from the image. It can be seen from Section 2 that the calculation precision of the proposed homing algorithm mainly depends on the landmarks extracted from the image, so both the selection of reliable landmarks and the matching accuracy are crucial to the homing performance.

#### 3.1. Landmark Selection

In this section, we will make the qualitative analysis of the landmark selection. Gillner *et al.* [35] suggested the principle of landmark selection, which includes the following three points: (1) uniqueness: landmarks in the scene must be unique and can be identified clearly; (2) reliability: the reliability of landmarks refers to the stable visibility, which means that the landmarks should be always detected every time the robot travels to the same position; and (3) relevance: the relevance of landmarks is defined as their importance in determining the homing direction.

According to the above three principles, we choose SIFT features in the image as landmarks. Firstly, as SIFT features are highly distinctive and can be correctly matched with high probability against large features [31], they can meet the uniqueness condition of landmarks. Secondly, as SIFT features are invariant to image translation, rotation, scale and are stable with respect to the changes in illumination [31], the reliability condition can be satisfied. At last, as the angles between SIFT features and the orientations of the robot can be computed by the localization of SIFT features, we can get the homing direction by utilizing Equations (3)–(7). The relevance condition can be satisfied.

#### 3.2. Mismatching Elimination

In order to obtain a larger field of view and a faster processing speed, the proposed homing algorithm directly extracts SIFT features from the initial catadioptric panoramic images as landmarks.

Although the SIFT features have good performance, there still exist some mismatching points in the real scene. According to Section 2, the mismatching of the landmarks can generate the wrong corresponding angle pair  $(\theta, \theta')$ , which will affect the homing precision and even cause the failure of homing. To solve the above problems, we present two constraints according to the distribution characteristics of landmarks in the initial panoramic images and propose a novel mismatching elimination algorithm based on these constraints.

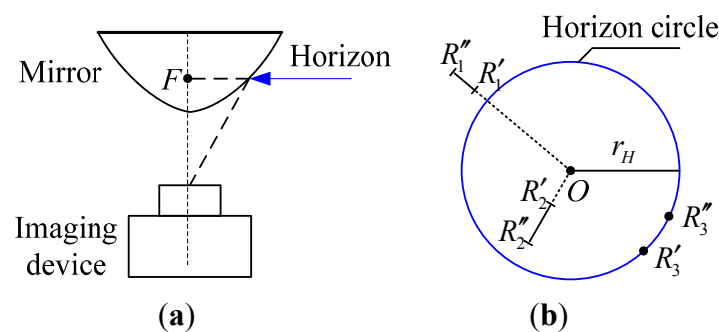
### 3.2.1. Two Distribution Constraints

The panoramic images used in this paper are all generated by the catadioptric panoramic imaging system based on the hyperbolic mirror. Before presenting the mismatching elimination algorithm, we firstly introduce two distribution constraints of the landmarks as follows:

Distribution Constraint 1: As shown in Figure 2a, in the catadioptric panoramic imaging system based on the hyperbolic mirror, a curved mirror is used to form the image of the surroundings. Under the curved mirror is located the camera, whose optic axis points to the mirror. The curved mirror works together with the camera to collect the panoramic image. Ideally, in the initial panoramic image, there exists a circle whose center is just at the center of the image. For landmarks located at the same horizontal level with focus  $F$  of the curved mirror, their projection must be on that circle. As shown in Figure 2b, we call this circle the horizon circle in this paper. According to the characteristics of the catadioptric panoramic imaging system, suppose the robot moves on a plane; the landmarks in the scene can be divided into three categories: (1) the landmarks are located higher than focus  $F$ ; their imaging radius range is always outside the horizon circle, shown as  $R'_1R''_1$ ; (2) the landmarks are located lower than focus  $F$ ; their imaging radius range is always within the horizon circle, shown as  $R'_2R''_2$ ; and (3) the landmarks are located at the same horizontal level with focus  $F$ . As  $R'_3(R''_3)$  shows, the imaging radius range of these landmarks is always on the circle as the robot travels. In this paper, we employ  $\gamma_{IO}(L)$  to determine the position relationship between the landmark  $L$  and the horizon circle in quantization, which can be defined as follows:

$$\gamma_{IO}(L) = \begin{cases} -1, & r_L < r_H \\ 0, & r_L = r_H \\ 1, & r_L > r_H \end{cases} \quad (8)$$

where  $r_L$  is the imaging radius of the landmark and  $r_H$  is the radius of the horizon circle.



**Figure 2.** Distribution Constraint 1. (a) The formation of the horizon circle; (b) the distribution constraint of the landmarks based on the horizon circle.

According to the above principle, suppose  $L_T$  is a landmark to be tested in the goal image, and  $\tilde{L}_T$  denotes its matching landmark in the current image. If the pair of landmarks  $(L_T, \tilde{L}_T)$  matches correctly, there will exist the relation  $\gamma_{IO}(L_T) = \gamma_{IO}(\tilde{L}_T)$ . In other words, if  $\gamma_{IO}(L_T) \neq \gamma_{IO}(\tilde{L}_T)$ ,  $(L_T, \tilde{L}_T)$  is a pair of mismatching landmarks.

Distribution Constraint 2: As shown in Figure 3, the left picture shows the goal image, and the right picture shows the current image.  $O$  is the center of the image.  $L_1, L_2$  and  $L_3$  are three landmarks in the goal image.  $\tilde{L}_1, \tilde{L}_2$  and  $\tilde{L}_3$  are their matching landmarks in the current image, respectively.  $(L_1, \tilde{L}_1)$  is a pair of landmarks to be tested.  $(L_2, \tilde{L}_2)$  and  $(L_3, \tilde{L}_3)$  are used as reference landmark pairs. The two directed lines radiate from the center  $O$  and point to  $L_1$  and  $\tilde{L}_1$ , based on which both the goal image and the current image can be divided into the left half plane and the right half plane. As the robot moves from the goal position to the current position, the reference landmark pairs  $(L_2, \tilde{L}_2)$  and  $(L_3, \tilde{L}_3)$  will separately appear in the same half plane of the two images on the condition that  $(L_1, \tilde{L}_1)$  are matching correctly. Conversely, if  $(L_1, \tilde{L}_1)$  is a pair of mismatching landmarks,  $(L_2, \tilde{L}_2)$  and  $(L_3, \tilde{L}_3)$  may randomly appear in any position of the two images.

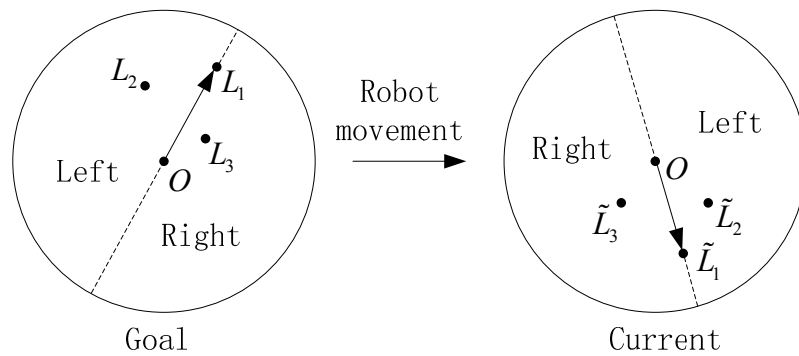


Figure 3. Distribution Constraint 2.

Suppose the coordinates of  $O, L_1$  and  $L_2$  are  $(x_o, y_o), (x_t, y_t)$  and  $(x_r, y_r)$ , respectively. The equation of the directed line  $OL_1$  is  $Ax + By + C = 0$ , where  $A = y_t - y_o, B = x_o - x_t$  and  $C = x_t y_o - x_o y_t$ . The positional relation of the reference landmark  $L_2$  can be determined as follows:

$$D_R = Ax_R + By_R + C \tag{9}$$

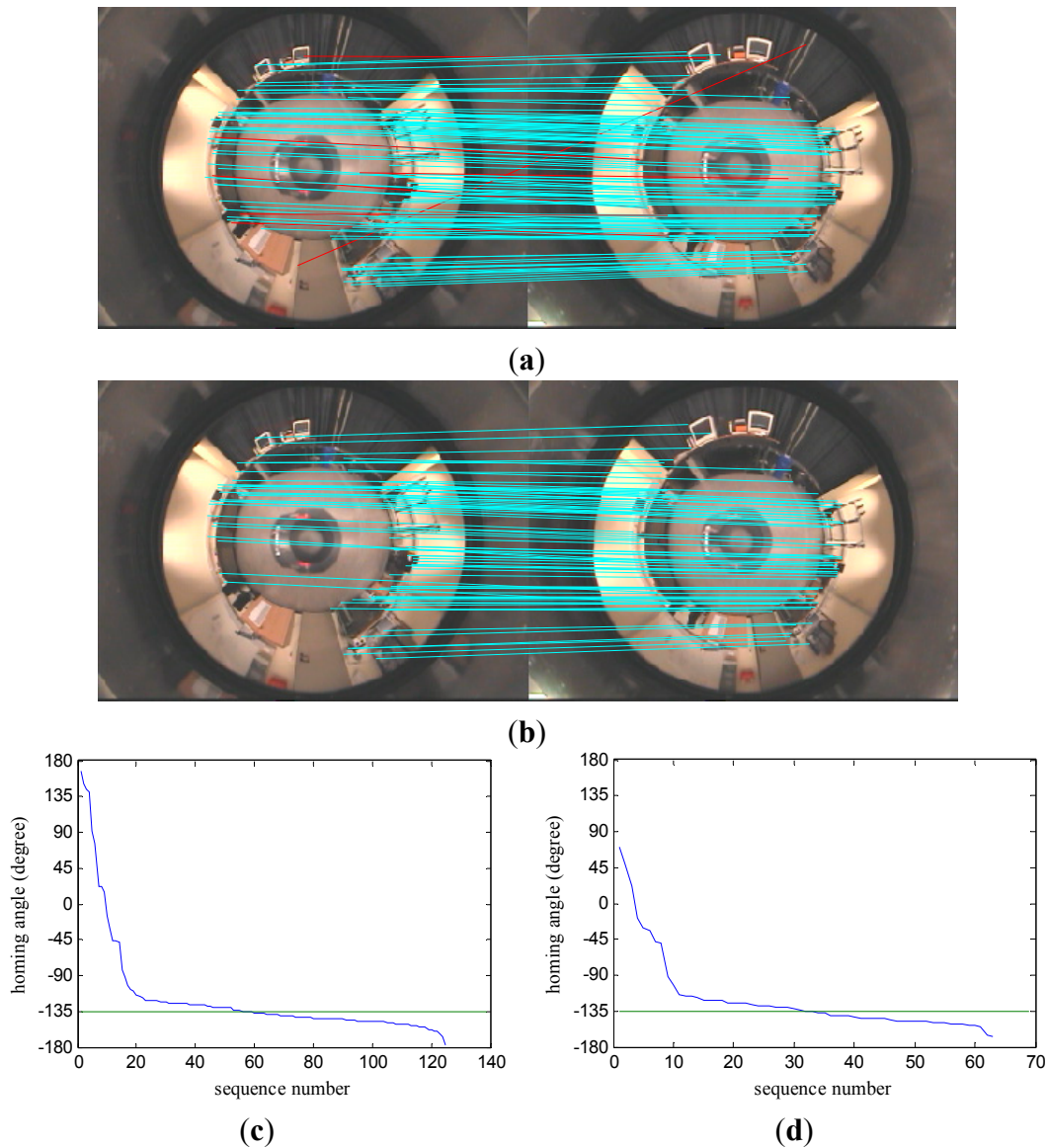
In Equation (9): (1) if  $D_R < 0$ ,  $L_2$  is located in the left half plane; (2) if  $D_R > 0$ ,  $L_2$  appears in the right half plane; (3) if  $D_R = 0$ ,  $L_2$  is on the line  $OL_1$ . We employ  $\gamma_{LR}(L)$  to indicate the positional relation of a reference landmark in quantization, which is defined as:

$$\gamma_{LR}(L) = \begin{cases} -1, & D_R < 0 \\ 0, & D_R = 0 \\ 1, & D_R > 0 \end{cases} \tag{10}$$

### 3.2.2. Mismatching Elimination Algorithm

Suppose  $S_1$  is the initial set of matching landmarks, which are extracted from the goal image  $I_H$  and the current image  $I_C$ . The matching landmark pair  $(L_T, \tilde{L}_T) \in S_1$  is the one to be tested, where  $L_T$  and  $\tilde{L}_T$

are landmarks extracted separately from  $I_H$  and  $I_C$ . The proposed mismatching elimination algorithm mainly includes two phases, which can be presented as follows:



**Figure 4.** The performance of the proposed mismatching elimination algorithm. **(a)** The matching of landmarks before mismatching elimination; **(b)** the matching of landmarks after mismatching elimination; **(c,d)** the distribution of homing angles computed by the proposed homing algorithm before and after mismatching elimination.

The first phase: According to Constraint 1, we firstly calculate  $\gamma_{IO}(L_T)$  and  $\gamma_{IO}(\tilde{L}_T)$  for each pair  $(L_T, \tilde{L}_T) \in S_1$ . If  $\gamma_{IO}(L_T) \neq \gamma_{IO}(\tilde{L}_T)$ , the corresponding landmark pair will be discarded from set  $S_1$ ; if  $\gamma_{IO}(L_T) = \gamma_{IO}(\tilde{L}_T)$ , the corresponding landmark pair will be reserved. After the preliminary filtration of set  $S_1$ , the remaining matching landmark pairs make up set  $S_2$ .

The second phase: For each matching landmark pair  $(L_T, \tilde{L}_T) \in S_2$ , we firstly calculate the distance between  $L_T$  and other landmarks in  $I_H$  and choose  $n_R$  landmarks that are nearest to  $L_T$  as reference landmarks to constitute set  $S_{RH}$ , then find their matching landmarks in  $I_C$  to constitute set  $S_{RC}$ . The correctness of each matching landmark pair to be examined can be evaluated by the reference

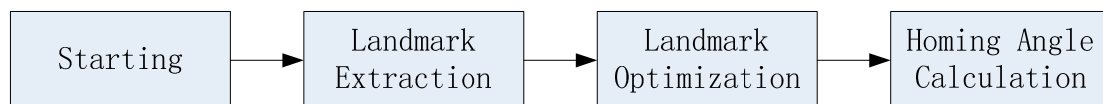


landmark pairs in set  $S_{RH}$  and set  $S_{RC}$ , and the result will be registered in  $V_m$ , whose initial value is zero. According to Constraint 2, we separately calculate  $\gamma_{LR}(L_R)$  and  $\gamma_{LR}(\tilde{L}_R)$  for each reference landmark  $L_R \in S_{RH}$  and its matching landmark  $\tilde{L}_R \in S_{RC}$ . If  $\gamma_{LR}(L_R) = \gamma_{LR}(\tilde{L}_R)$ , the current reference landmark pair supports the correctness of the matching pair to be examined, and the value of  $V_m$  increases by one; if  $\gamma_{LR}(L_R) \neq \gamma_{LR}(\tilde{L}_R)$ , the matching pair to be examined is regarded as the wrong one, and the value of  $V_m$  does not change. Suppose the threshold is  $V_{TH}$ , a matching pair  $m = (L_T, \tilde{L}_T) \in S_2$  can be directly judged by the value of its  $V_m$ . If  $V_m \geq V_{TH}$ ,  $m$  is a correct matching pair and needs to be reserved; if  $V_m < V_{TH}$ ,  $m$  is regarded as a mismatching pair and will be removed from set  $S_2$ . The remaining matching pairs will be used as the final landmarks and constitute set  $S_3$ .

In order to evaluate the performance of the proposed mismatching elimination algorithm, the panoramic image databases provided by Bielefeld University [19] are adopted. The image databases will also be used in the subsequent homing experiments, and the details will be introduced in Section 4.1. Figure 4 shows the performance of the proposed mismatching elimination algorithm. In Figure 4a, the red lines indicate the obvious mismatching landmarks. It can be seen from Figure 4b that these mismatching pairs are eliminated effectively by the proposed algorithm. In Figure 4c,d, green lines denote the ideal homing angle. After eliminating the mismatching landmarks, the distribution of the homing angles calculated by the proposed homing algorithm is much closer to the ideal homing angle, as shown in Figure 4c,d. Experimental results show that the proposed algorithm in this section can effectively eliminate the mismatching landmark pairs and further improve the computational accuracy of the proposed homing algorithm.

### 3.3. Overview of the Proposed Visual Homing Method

As shown in Figure 5, the procedure of our homing method mainly includes three steps as follows:



**Figure 5.** Flow diagram of the proposed homing method.

(1) Landmark extraction: We extract SIFT features from the goal image  $I_H$  and the current image  $I_C$ , then match the features to form the initial landmark set  $S_1$ .

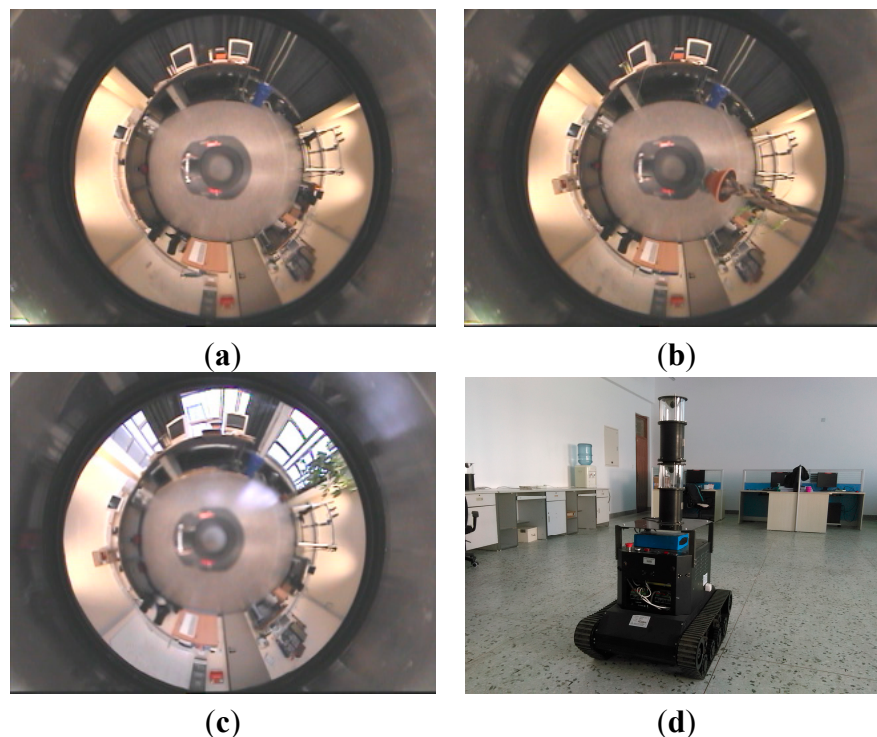
(2) Landmark optimization: The primary purpose of this step is to eliminate the mismatching landmark pairs in set  $S_1$ . The proposed algorithm in this paper solves this problem in a coarse-to-fine hierarchical way. First of all, according to the landmark distribution Constraint 1, the matching landmark pairs in set  $S_1$  are filtered preliminarily by making full use of the distribution characteristic between the landmarks and the horizon circle. The remainder in set  $S_1$  form set  $S_2$ . After that, with the help of the landmark distribution Constraint 2, the mismatching pairs in set  $S_2$  are further removed on the basis of the relative distribution relations of the landmarks in an initial panoramic image. The final landmark set  $S_3$  consists of the remaining landmark pairs in set  $S_2$ .

(3) Homing angle calculation: According to Section 2, we can get an angle pair  $(\theta, \theta')$  for each pair  $(L, \tilde{L}) \in S_3$ . The final homing direction  $\hat{\beta}$  can be worked out based on Equations (3) to (7).

## 4. Experiments

### 4.1. Image Databases and Robot Platform

The panoramic image databases used in this paper were collected by the Computer Department of Bielefeld University. These databases were widely applied to the tests of robot visual homing algorithms. The panoramic images in the databases were all collected by the catadioptric imaging system in different scenes. The capture grid of the above image databases measured  $2.7 \times 4.8$  m, and images were collected uniformly spaced 0.3-m apart. The resolution of these images is  $752 \times 564$ . The images acquired in three different scenes of original, arboreal and day were used for the experiments. Three panoramic sample images are shown in Figure 6a to c, and the corresponding scene conditions can be introduced as follows: (a) original means a common room with doors and windows shut, and its overhead fluorescent light bars were on; (b) arboreal represents a tall plant that was added in the capture grid; and (c) day means images that were collected with curtains open in full daylight. The databases and their further details can be accessed through [36].



**Figure 6.** Panoramic sample images and robot platform. (a–c) The samples of three image databases: original, arboreal and day; (d) the robot platform for experiments in the real scene.

In this paper, a tracked mobile robot was used in the experiments in the real scene, as shown in Figure 6d. The panoramic imaging system mounted on the top of the robot was composed of a hyperbolic mirror and a camera with a resolution of  $1024 \times 768$ . Although there were two sets of panoramic imaging equipment on the robot, we just used the one below in the experiments. Image processing and robot movement were controlled by an onboard computer (Pentium (R), 2 GHz).

#### 4.2. Parameter Settings for Experiments

The parameter settings in the experiments are shown in Table 1. In this paper, we employ the SIFT features of the image as landmarks. On the premise of accurate results, we modified several parameters recommended by Lowe in [31] to get more SIFT features. Firstly, the number of scale layers  $S$  where SIFT features are extracted is increased from 3 to 5, which can both increase the total number of extracted landmarks and maintain a reasonable time of execution. Secondly, the response threshold of extreme points  $T_{DOG}$  in the difference of Gaussian images is decreased from  $0.08/S$  to  $0.04/S$  in order to get more features from areas of low contrast, as indoor environments often contain such areas.

**Table 1.** Parameters for the experiments.

Parameters	Value	Parameters	Value
$S$	5	$\alpha_w$	[0, 355]/72/5
$T_{DOG}$	$0.04/S$	$n_R$	5
$\rho_w$	[0, 0.95]/20/0.05	$V_{TH}$	4
$\psi_w$	[0, 355]/72/5		

In Table 1, the parameter format of  $\rho_w$ ,  $\psi_w$  and  $\alpha_w$  is shown as search range/search steps/resolution. In order to give full play to the performance of the warping method in the experiments, we adopted the parameters recommended by Möller in [30]. The search range of  $\rho_w$  is [0, 0.95], and there are 20 search steps. The settings of  $\psi_w$  and  $\alpha_w$  are the same, whose search range is [0, 355], and there are 72 search steps. In addition, according to the requirements of the proposed homing algorithm, the number of the reference landmark pairs  $n_R$  and the threshold  $V_{TH}$  for eliminating the mismatching landmarks were separately set to 5 and 4.

We separately took 100 pairs of images from the three image databases randomly. These image pairs were used as the goal image and the current image. Based on the parameter settings in Table 1, we determined the average computation time for two methods (2 GHz Pentium (R), MATLAB R2007b), as shown in Table 2. It can be seen that the average computation time of the warping method for each homing test was about 40% more than that of our method. We consider the settings in Table 1 reasonable enough for the comparison of homing performance.

**Table 2.** Average computation time for the two methods.

Method	Average Computation Time (s)		
	<i>original</i>	<i>arboreal</i>	<i>day</i>
Warping	21.051	18.643	19.541
Proposed	14.559	12.638	13.854

#### 4.3. Performance Metrics

According to the previous presentation, the goal position and the current position are denoted separately by H and C in the experiments. In this paper, three metrics known as angular error (AE), average homeward component (AHC) and return ratio (RR) are adopted to evaluate the performance of the proposed method.

For a pair of H and C, suppose the homing angle computed by the homing algorithm is indicated by  $\beta_{\text{homing}}$ .  $\beta_{\text{ideal}}$  represents the ideal homing angle, which directly points from C to H. The angular error can be determined as:

$$AE(H,C) = \text{diff}(\beta_{\text{homing}} - \beta_{\text{ideal}}) \quad (11)$$

where the function of  $\text{diff}()$  refers to Equation (6).

The average homeward component, used in the homing experiments frequently, is the evaluation criterion, which can measure both the validity and the angular deviation of the computed homing angles. As long as the value of AHC always stays above zero, the robot travels nearer to the goal position; the closer the value of AHC is to 1, the closer the movement direction of the robot gets to the ideal homing direction. Based on the angular error, the average homeward component can be defined as follows:

$$AHC(n) = \cos\left(\frac{1}{n} \sum_{i=1}^n AE(H_i, C_i)\right) \quad (12)$$

where  $n$  indicates the number of different pairs (H, C) selected in the experiment scene. In the experiments of AHC, we separately took 100 pairs of images from the image databases randomly according to the distance between C and H, which ranges from 30 to 390 cm in steps of 30 cm.

The last performance metric is the return ratio [18,19], which is defined as the successful percentage of homing trials. The return ratio can be computed by carrying out simulated homing trials on the capture grid of the image databases. A dummy robot is placed at all integer grid positions and allowed to move according to the homing vectors, which have been pre-computed with the two methods. The robot moves at a step of  $r_h$ , which is generally determined by the ratio between the actual step length of the robot and the sampling interval of the images. The value of  $r_h$  was set to 0.8 in the trials.  $\beta_h(x,y)$  denotes the homing angle pre-computed at each position in the capture grid, and the motion direction of the robot is determined by  $\beta_h(x,y)$ , whose position is closest to the current position. A trial is considered successful if the robot can reach a place within a given distance threshold of H from C. The threshold was set to 0.5. The result of a homing trial at position C can be evaluated as follows:

Step 1: The robot moves a step according to the corresponding  $\beta_h(x,y)$ .

Step 2: If the following two cases happen, jump to Step 4.

Case 1: The robot arrives at the goal position H.

Case 2: The robot travels a distance longer than half of the perimeter of the capture grid.

Step 3: Continue to perform Step 1.

Step 4: If Case 1 happens, the homing trial is successful; if Case 1 does not happen and Case 2 happens, the trial has failed.

We define  $\lambda(H,C)$  as a binary evaluation function with a value of 1 for successful homing and 0 for unsuccessful homing. The return ratio is determined as:

$$RR(H) = \frac{1}{n} \sum_{i=1}^n \lambda(H, C_i) \quad (13)$$

where  $n$  indicates the number of different current positions selected in the experiment scene.

#### 4.4. Homing Experiments on Image Databases

The image databases introduced in Section 4.1 were used to perform the experiments. The experiment environment of the scene in this section was divided into two classes: (1) static environment: the surroundings of the scene remain static during the experiment; (2) dynamic environment: the objects or illumination in the scene changes during the experiment. According to different experiment environments, three different groups of experiments were conducted: (1) Group 1: the main goal is to test the homing performance of the proposed method under static conditions; both current images and goal images were selected from the database original, and this experiment condition was represented by original-original; (2) Group 2: the main goal is to test the homing performance of the proposed method when the objects in the scene change; the changing of objects was simulated by using cross-database experiments, *i.e.*, the current images were taken from the database original, and the goal images were taken from the database arboreal; the experiment condition was represented by original-arboreal; (3) Group 3: the main goal is to test the homing performance of the proposed method when the illumination of the scene changes; similar to the second group of experiments, the current images were taken from the database original and the goal images were taken from the database day, in order to simulate the changing of illumination; the experiment condition was represented by original-day.

Figure 7 shows the homing vector fields for the warping method and the proposed method. (1, 4), (7, 13) and (5, 9) in the capture grid were selected as goal positions, and the corresponding experiment conditions were original-original, original-arboreal and original-day, respectively. In Figure 7, the current positions are marked by blue squares, and the goal positions are marked by red squares. The homing direction is denoted by the line put out from the blue square. Figure 8 shows the AE results for two homing methods. The experiment conditions and goal positions are the same as the settings of Figure 7. The gray scale of each grid indicates the AE value of the corresponding (x, y) position in the capture grid of the database. Its changing from black to white represents the value ranging from 0 to the maximum of AE computed by the two methods. For observing, we change the threshold of white to 100 in the experiments. From Figures 7 to 8, we can draw the following conclusions: Most of the homing directions computed by our method are more accurate than those computed by warping method. The AE of our homing method is lower than that of the warping method on the whole.

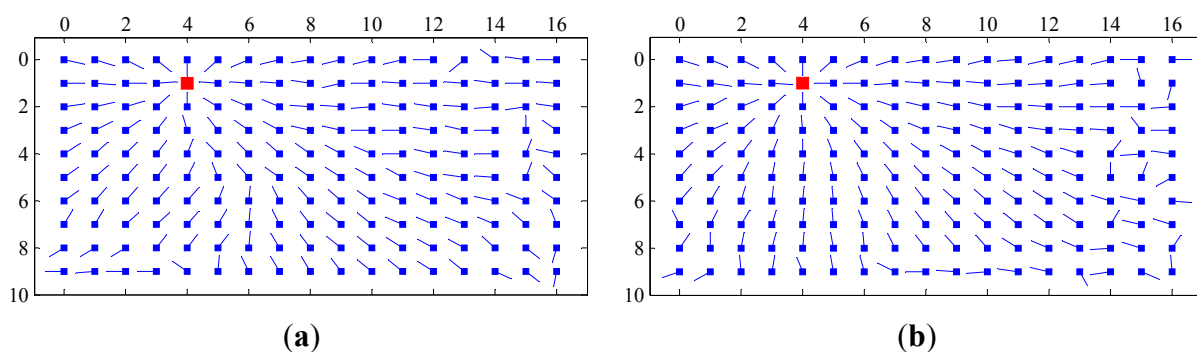
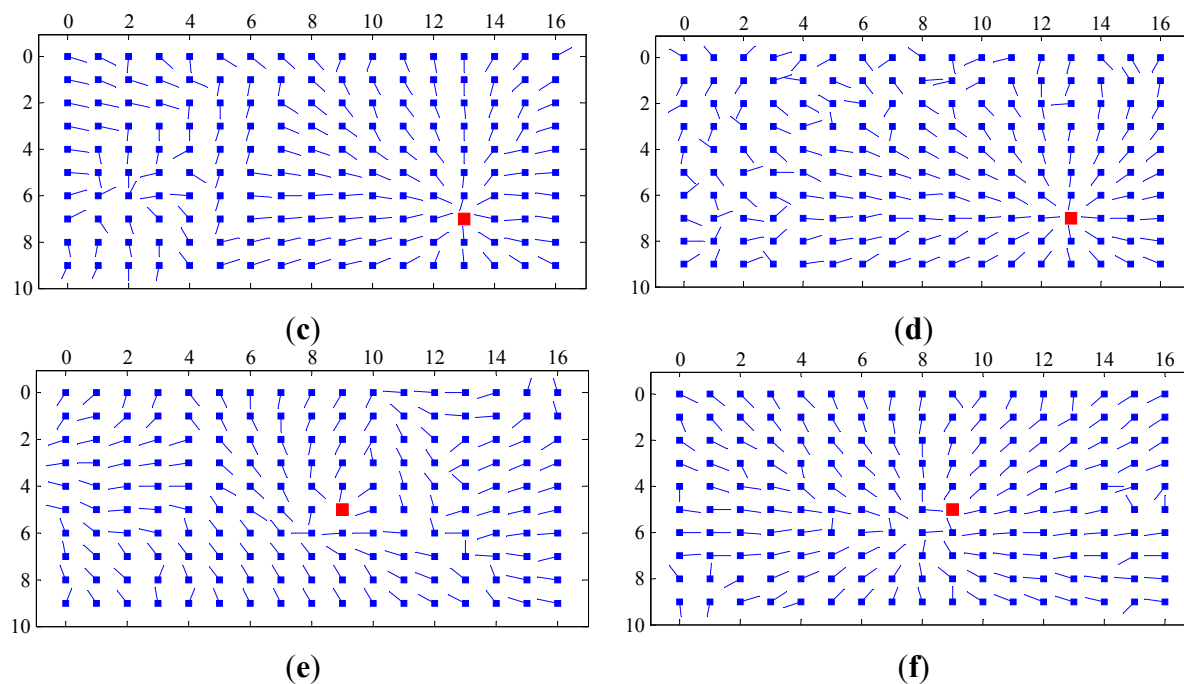
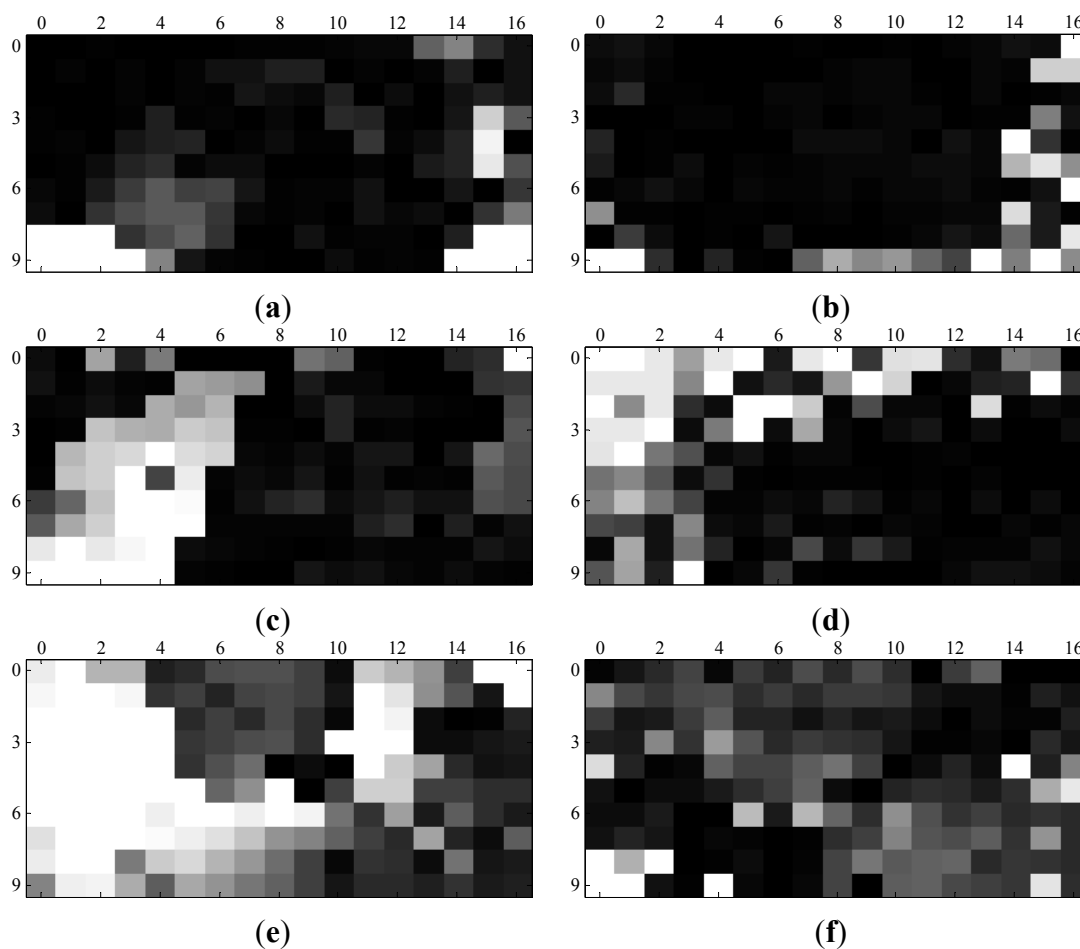


Figure 7. Cont.



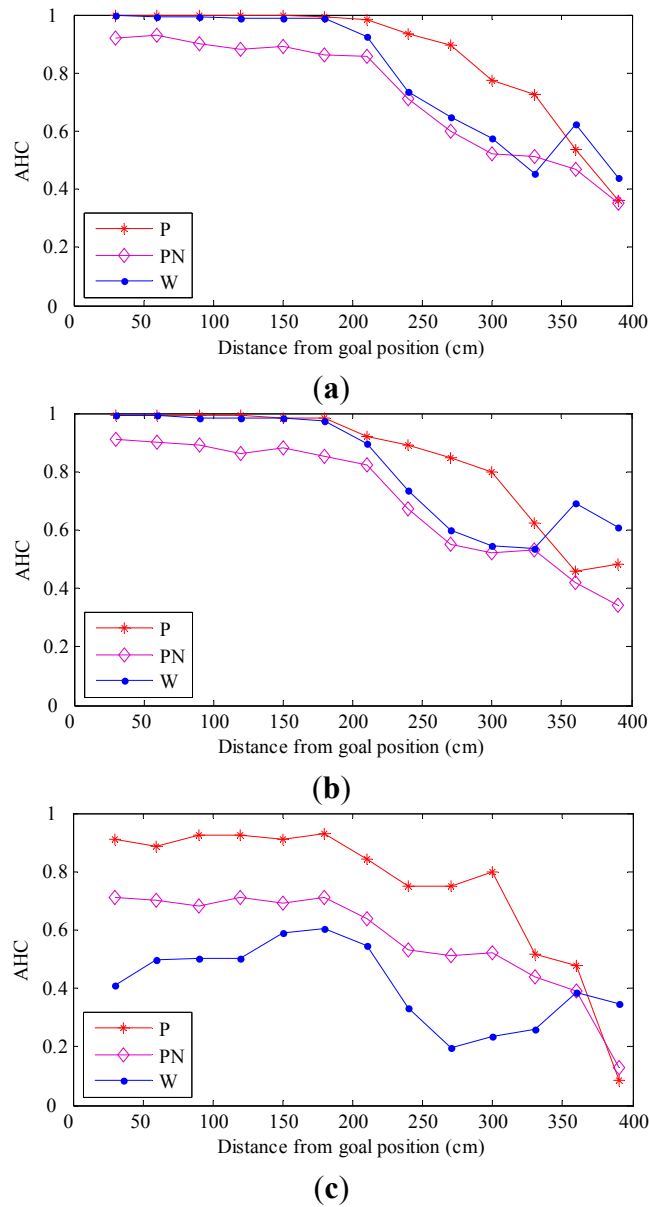
**Figure 7.** Homing vector fields. (a,c,e) The homing vectors generated by the warping method; (b,d,f) the homing vectors generated by the proposed homing method.



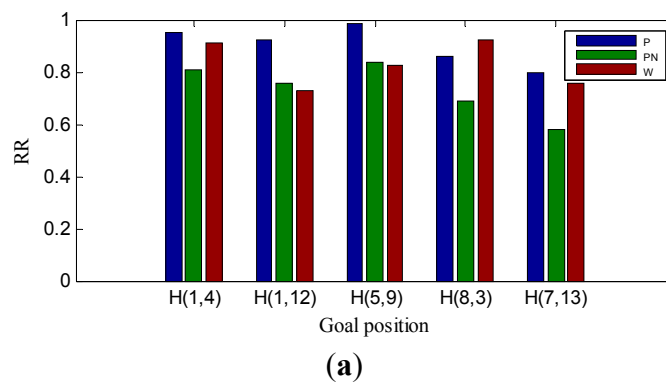
**Figure 8.** Angular error (AE) results. (a,c,e) The homing angular errors generated by warping method; (b,d,f) the homing angular errors generated by the proposed homing method.

Figure 9 shows the distribution of AHC according to the distance between C and H. The corresponding experiment conditions for Figure 9a–c were original-original, original-arboreal and original-day, respectively. From Figure 9, it can be seen from the changing trends of “P” and “PN” that the mismatching elimination step can effectively improve the homing performance. It also can be seen intuitively from the changing trends of “P” and “W” that the AHC of our homing method is superior to that of the warping method on the whole. In the three groups of experiments, the AHC of two methods is always above 0, which indicates that both methods have the ability to guide the robot to approach the goal position gradually. From Figure 9a,b, although the change of a tall plant in the experiment scene leads to a slight decrease in AHC for two methods, our homing method still performs better. As shown in Figure 9a,c, when the illumination of the experiment scene changes, the performance of the warping method drops dramatically, while the performance of our homing method drops slightly. In conclusion, the results show that our homing method has better robustness to the changes of the environment. In Figure 9, we can see an interesting phenomenon: when the distance between C and H is approximately within the range 30 to 330 cm, the AHC of our method is closer to 1 than that of the warping method, which shows that the average AE of our homing method is smaller; when the distance between C and H is approximately within the range 360 to 390 cm, the AHC of our homing method is smaller than that of the warping method, which indicates that the average AE of our homing method is higher. The main reasons for this situation are as follows: (1) when the distance between C and H is short (30 to 330 cm), there are more correct matching landmarks due to minor differences between the two images; consequently, the proposed mismatching elimination algorithm can effectively remove the mismatching landmarks, and our homing algorithm can get higher calculation accuracy; (2) when the distance between C and H is far (360 to 390 cm), the number of matching landmarks is smaller, and among them, there are more mismatching landmarks. For this reason, the proposed algorithm cannot effectively eliminate the mismatching landmarks, which makes the calculation accuracy of our homing algorithm decrease significantly.

Figure 10 shows the RR for the two methods under the three experiment conditions. “P” represents the proposed homing method; “PN” represents the proposed homing method without mismatching elimination step; and “W” represents the warping method. Trying to avoid the influence of randomness, we chose (1, 4), (1, 12), (5, 9), (8, 3) and (7, 13) in the capture grid of the image databases as goal positions, which are uniformly distributed in the experiment scene. From Figure 10, it can be seen from the bar chart of “P” and “PN” that the mismatching elimination step can effectively improve the RR of our homing method. There are 15 different goal positions tested; compared to the warping method, our method performs better at 13 goal positions. From Figure 10a,b, although the RR for the two methods drops slightly with the changing of a tall plant in the experiment scene, our homing method still performs better. The performance for the two methods is greatly influenced when the illumination changes in the experiment scene, as shown in Figure 10a,c. Especially for the warping method, the homing performance declines dramatically. The results of RR turn out to be the same as those of AHC: compared to the warping method, our homing method has better robustness to the changes of the environment in the scene.

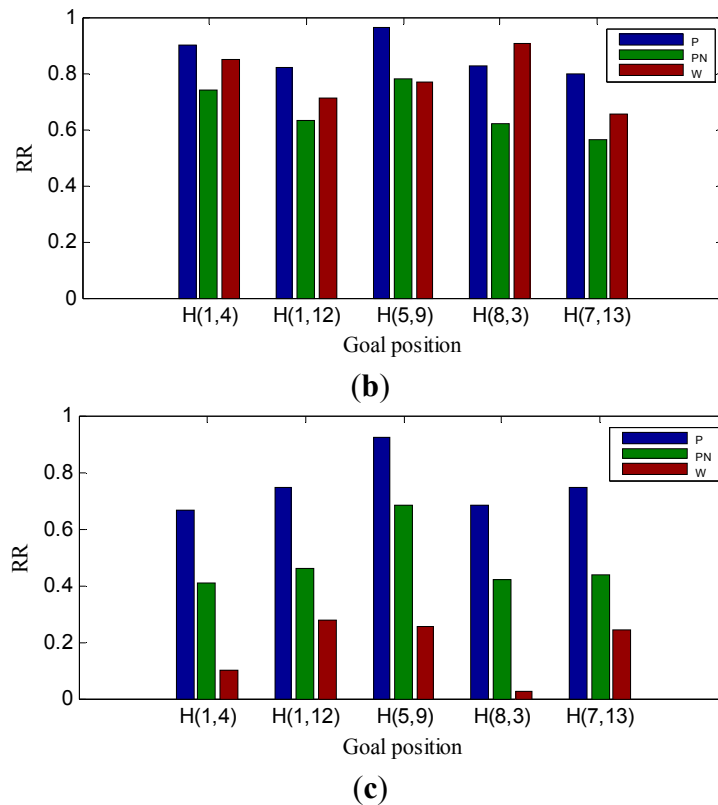


**Figure 9.** Average homeward component (AHC) results. (a–c) The distribution of AHC under the experiment conditions: original-original, original-arboreal and original-day. P, proposed homing method; PN, proposed homing method without a mismatching elimination step; W, warping method.



**Figure 10. Cont.**





**Figure 10.** Return ratio (RR) results. (a–c) The RR for five goal positions under the experiment conditions: original-original, original-arboreal and original-day.

#### 4.5. Homing Trials in a Real Scene

In order to further evaluate the performance of our method in practice, experiments were conducted in a real scene. We selected the intelligent robot lab of Harbin Engineering University as the experiment scene. The surroundings of the trial area are shown in Figure 11. The tracked mobile robot introduced in Section 4.1 was used in the trials. We randomly chose four goal positions, which were uniformly distributed in the trial area. Five current positions for each goal position spaced evenly throughout the area were selected for tests. The robot took the panoramic image at its current position, compared it to the goal image stored in memory and computed the homing angles separately by the proposed method and the warping method. After that, in the homing direction, the robot moved a step at the fixed length of 25 cm. In the trials, the above process would repeat until the robot reached a place within the range of 30 cm of the goal position or its movement distance was more than half of the circumference of the trial area, which was when the robot moves no more than 43 steps. If the robot can arrive at the goal position within the prescribed steps, the homing is successful; otherwise, the homing has failed. Because most stopping criteria based on image information are likely to lead the robot to oscillate around the goal position, the robot was manually stopped once it reached the goal area.

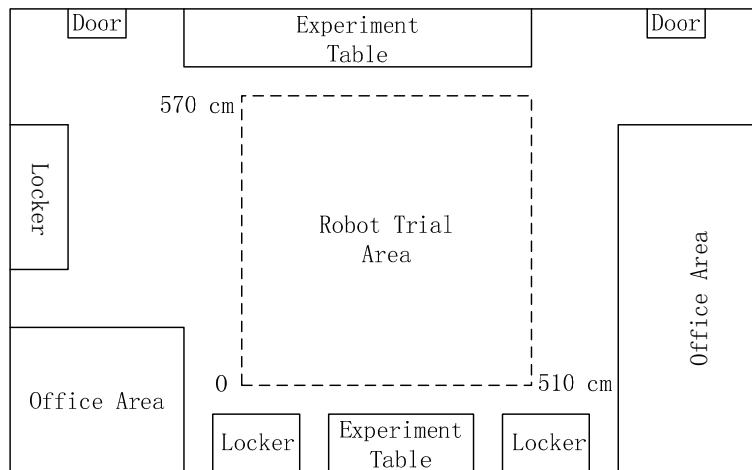
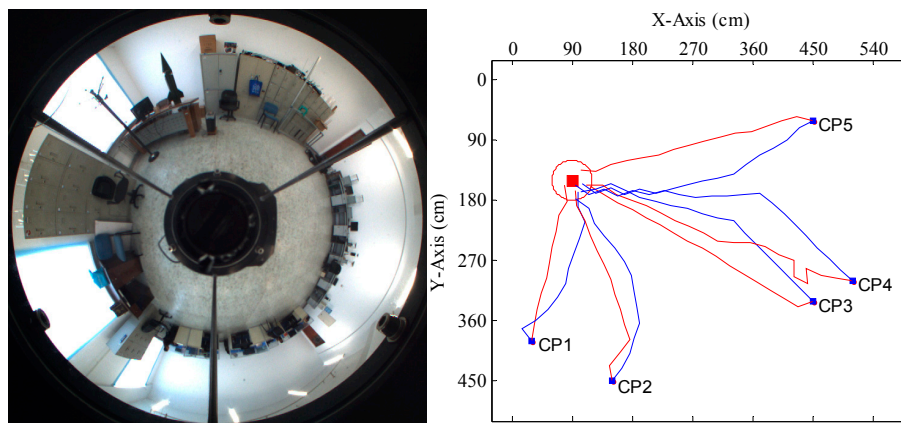


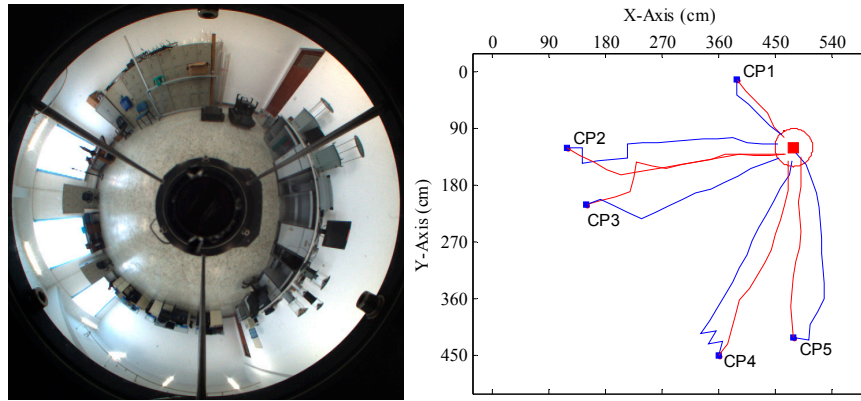
Figure 11. Robot trial environment in the real scene.

Figures 12–15 show the trajectories of the robot for the two methods, with tables listing the statistics of the number of homing steps  $N$  and the average angular errors  $\sigma(^{\circ})$  in each trial. The goal area is indicated by the red circle. “CP” represents the current position. Red lines represent the trajectories of the proposed method. Blue lines represent those of the warping method. As shown in Figures 12–15, most homing trajectories for the warping method are more curved than those for our method. In total, 20 homing trials were carried out, 17 of which suggest that the average angular errors for our method are smaller; the number of total homing steps for the warping method is 312, while for our method, that number is only 292. Both the trajectories and statistics indicate that the homing angles computed by our method are more accurate, and the movement distance is shorter.



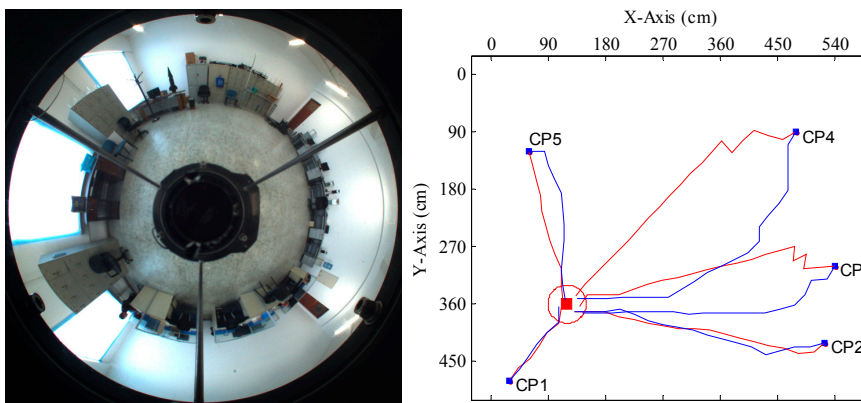
Method	CP1		CP2		CP3		CP4		CP5	
	$N$	$\sigma$	$N$	$\sigma$	$N$	$\sigma$	$N$	$\sigma$	$N$	$\sigma$
Warping	11	24.54	13	21.12	17	15.03	20	19.14	16	18.78
Proposed	10	6.34	13	11.57	16	5.81	20	15.43	15	6.47

Figure 12. Robot homing Trial 1. Top left: the panorama of the goal position; top right: the homing trajectories for five different current positions; the table below: the number of homing steps and the average angular error for each current position. CP, current position.



Method	CP1		CP2		CP3		CP4		CP5	
	<i>N</i>	$\sigma$	<i>N</i>	$\sigma$	<i>N</i>	$\sigma$	<i>N</i>	$\sigma$	<i>N</i>	$\sigma$
Warping	5	9.66	16	15.23	14	11.79	17	19.09	14	20.19
Proposed	5	6.47	15	11.42	15	14.92	14	9.29	12	8.31

Figure 13. Robot homing Trial 2.



Method	CP1		CP2		CP3		CP4		CP5	
	<i>N</i>	$\sigma$	<i>N</i>	$\sigma$	<i>N</i>	$\sigma$	<i>N</i>	$\sigma$	<i>N</i>	$\sigma$
Warping	6	8.01	17	10.65	18	14.79	21	23.82	11	13.89
Proposed	6	8.13	17	8.07	20	16.24	20	11.89	10	5.10

Figure 14. Robot homing Trial 3.

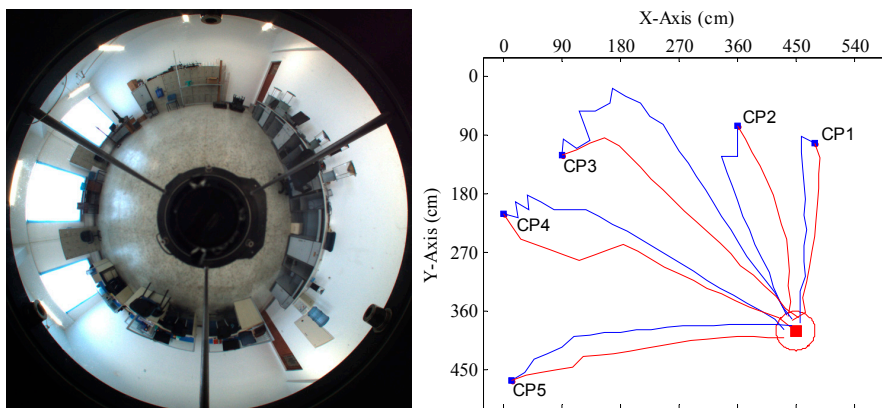


Figure 15. Cont.

Method	CP1		CP2		CP3		CP4		CP5	
	$N$	$\sigma$	$N$	$\sigma$	$N$	$\sigma$	$N$	$\sigma$	$N$	$\sigma$
Warping	13	15.13	14	12.93	27	31.89	23	15.99	19	12.14
Proposed	12	9.50	13	9.89	20	15.22	21	11.32	18	5.29

**Figure 15.** Robot homing Trial 4.

## 5. Conclusions

This paper proposes a novel method to solve the problem of local visual homing. The method is composed of a novel visual homing algorithm and a novel mismatching elimination algorithm. The former is inspired by the warping method, and the latter is based on the distribution characteristics of landmarks in the initial panoramic image. Compared to the warping method, the proposed homing method improves the homing accuracy effectively and has better robustness to the changes in the environment. Experiments on image databases and in a real scene confirm the improved performance.

For the visual homing algorithms based on landmarks, a reduction in the number of landmarks can effectively reduce the amount of computation, while the homing performance might be affected. In the future, we will focus on how to reduce the number of landmarks effectively on the premise of guaranteeing the homing precision.

## Acknowledgments

The authors are very grateful to the reviewers and editors for their valuable comments and hard work. With their help, this paper has been improved significantly. This work is partially supported by the National Natural Science Foundation of China (61175089, 61203255, 51409053).

## Author Contributions

The work presented in this paper was accomplished in a concerted effort by all authors. Qidan Zhu and Chuanjia Liu conceived of the study, designed the homing method and prepared the manuscript. Chuanjia Liu and Chengtao Cai performed the experiments and analyzed the data. All authors commented on the manuscript and approved the final version.

## Conflicts of Interest

The authors declare no conflict of interest.

## References

1. López-Nicolás, G.; Guerrero, J.J.; Sagüés, C. Multiple homographies with omnidirectional vision for robot homing. *Robot. Auton. Syst.* **2010**, *58*, 773–783.
2. Ohnishi, N.; Imiya, A. Appearance-based navigation and homing for autonomous mobile robot. *Image Vis. Comput.* **2013**, *31*, 511–532.
3. Aranda, M.; López-Nicolás, G.; Sagüés, C. Angle-based homing from a reference image set using the 1D trifocal tensor. *Auton. Robot.* **2013**, *34*, 73–91.

4. Labrosse, F. Short and long-range visual navigation using warped panoramic images. *Robot. Auton. Syst.* **2007**, *55*, 675–684.
5. Baddeley, B.; Graham, P.; Philippides, A.; Husbands, P. Holistic visual encoding of ant-like routes: Navigation without waypoints. *Adapt. Behav.* **2011**, *19*, 3–15.
6. Yu, S.E.; Lee, C.; Kim, D. Analyzing the effect of landmark vectors in homing navigation. *Adapt. Behav.* **2012**, *20*, 337–359.
7. Fu, Y.; Hsiang, T.R. A fast robot homing approach using sparse image waypoints. *Image Vis. Comput.* **2012**, *30*, 109–121.
8. Liu, M.; Pradalier, C.; Pomerleau, F.; Siegwart, R. Scale-only visual homing from an omnidirectional camera. In Proceedings of the IEEE International Conference on Robotics and Automation, St Paul, MN, USA, 14–18 May 2012; pp. 3944–3949.
9. Möller, R.; Krzykowski, M.; Gerstmayr-Hillen, L.; Horst, M.; Flier, D.; de Jong, J. Cleaning robot navigation using panoramic views and particle clouds as landmarks. *Robot. Auton. Syst.* **2013**, *61*, 1415–1439.
10. Liu, M.; Pradalier, C.; Siegwart, R. Visual homing from scale with an uncalibrated omnidirectional camera. *IEEE Trans. Robot.* **2013**, *29*, 1353–1365.
11. Guzel, M.S.; Bicker, R. A behaviour-based architecture for mapless navigation using vision. *Int. J. Adv. Robot. Syst.* **2012**, *9*, 18:1–18:13.
12. Qidan, Z.; Xue, L.; Chengtao, C. Feature optimization for long-range visual homing in changing environments. *Sensors* **2014**, *14*, 3342–3361.
13. Möller, R.; Krzykowski, M.; Gerstmayr, L. Three 2D-warping schemes for visual robot navigation. *Auton. Robot.* **2010**, *29*, 253–291.
14. Zeil, J.; Hofmann, M.I.; Chahl, J.S. Catchment areas of panoramic snapshots in outdoor scenes. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* **2003**, *20*, 450–469.
15. Sturzl, W.; Zeil, J. Depth, contrast and view-based homing in outdoor scenes. *Biol. Cybern.* **2007**, *96*, 519–531.
16. Arena, P.; De Fiore, S.; Fortuna, L.; Nicolosi, L.; Patane, L.; Vagliasindi, G. Visual homing: experimental results on an autonomous robot. In Proceedings of the European Conference on Circuit Theory and Design, Univ Sevilla, Seville, Spain, 26–30 August 2007; pp. 304–307.
17. Möller, R. A model of ant navigation based on visual prediction. *J. Theor. Biol.* **2012**, *305*, 118–130.
18. Churchill, D.; Vardy, A. An orientation invariant visual homing algorithm. *J. Intell. Robot. Syst.* **2013**, *71*, 3–29.
19. Vardy, A.; Möller, R. Biologically plausible visual homing methods based on optical flow techniques. *Connect. Sci.* **2005**, *17*, 47–89.
20. Briggs, A.J.; Detweiler, C.; Li, Y.; Mullen, P.C.; Scharstein, D. Matching scale-space features in 1D panoramas. *Comput. Vis. Image Underst.* **2006**, *103*, 184–195.
21. Loizou, S.G.; Kumar, V. Biologically inspired bearing-only navigation and tracking. In Proceedings of the IEEE Conference on Decision and Control, New Orleans, LA, USA, 12–14 December 2007; pp. 6121–6126.
22. Liu, M.; Pradalier, C.; Chen, Q.J.; Siegwart, R. A bearing-only 2D/3D-homing method under a visual servoing framework. In Proceedings of the IEEE International Conference on Robotics and Automation, Anchorage, AK, USA, 3–8 May 2010; pp. 4062–4067.

23. Möller, R.; Vardy, A. Local visual homing by matched-filter descent in image distances. *Biol. Cybern.* **2006**, *95*, 413–430.
24. Möller, R.; Vardy, A.; Kreft, S.; Ruwisch, S. Visual homing in environments with anisotropic landmark distribution. *Auton. Robot.* **2007**, *23*, 231–245.
25. Lambrinos, D.; Möller, R.; Labhart, T.; Pfeifer, R.; Wehner, R. A mobile robot employing insect strategies for navigation. *Robot. Auton. Syst.* **2000**, *30*, 39–64.
26. Basten, K.; Mallot, H.A. Simulated visual homing in desert ant natural environments: Efficiency of skyline cues. *Biol. Cybern.* **2010**, *102*, 413–425.
27. Ramisa, A.; Goldhoom, A.; Aldavert, D.; Toledo, R.; de Mantaras, R.L. Combining invariant features and the ALV homing method for autonomous robot navigation based on panoramas. *J. Intell. Robot. Syst.* **2011**, *64*, 625–649.
28. Franz, M.O.; Schölkopf, B.; Mallot, H.A.; Bülthoff, H.H. Where did I take that snapshot? Scene-based homing by image matching. *Biol. Cybern.* **1998**, *79*, 191–202.
29. Sturzl, W.; Mallot, H.A. Efficient visual homing based on Fourier transformed panoramic images. *Robot. Auton. Syst.* **2006**, *54*, 300–313.
30. Möller, R. Local visual homing by warping of two-dimensional images. *Robot. Auton. Syst.* **2009**, *57*, 87–101.
31. Lowe, D.G. Distinctive image features from scale invariant keypoints. *Int. J. Comput. Vis.* **2004**, *60*, 91–110.
32. Fischler, M.A.; Bolles, R.C. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM* **1981**, *24*, 381–395.
33. Wei, Z.; Xu, W.S.; Yu, Y.L. Area harmony dominating rectification method for SIFT image matching. In Proceedings of the IEEE Conference on Electronic Measurement and Instruments, Xi'an, China, 16–18 August 2007; pp. 935–939.
34. Wang, C.; Ma, K.K. Bipartite graph-based mismatch removal for wide-baseline image matching. *J. Vis. Commun. Image Represent.* **2014**, *25*, 1416–1424.
35. Gillner, S.; Weiß, A.M.; Mallot, H.A. Visual homing in the absence of feature-based landmark information. *Cognition* **2008**, *109*, 105–122.
36. Panoramic Image Database. Available online: <http://www.ti.uni-bielefeld.de/html/research/avardy/index.html> (accessed on 22 September 2014).