Preprints of the 19th World Congress
The International Federation of Automatic Control
Cape Town, South Africa. August 24-29, 2014

Improved Feature Distribution for Robot Homing

Qidan Zhu. Xue Liu. Chengtao Cai

College of Automation, Harbin Engineering University, Harbin, China (Tel: +86-451-8258-8940; e-mail: zhuqidan@hrbeu.edu.cn, liuxue_keyan@hrbeu.edu.cn, caichengtao@hrbeu.edu.cn)

Abstract: Visual homing which is behavior-based can make use of the natural landmarks to achieve homing in the unstructured environment. The local features, represented by the SIFT (Scale Invariant Feature Transform), have a variety of invariance and are quite suitable as the natural landmarks. Many visual homing methods can exhibit high precision only when the landmarks meet to isotropic distribution. However, most of the local feature extraction algorithms have not taken into account the uniformity of the feature distribution. Based on the initial SIFT, an improved feature extraction algorithm called UD-SIFT (Uniform Distribution-SIFT) is obtained which can improve the uniformity. The visual homing experiments are carried out indoors and in the corridor, using the ADV (Average Displacement Vector) and ALV (Average Landmark Vector) methods which are both based on the panoramic vision. The results turn out that the UD-SIFT algorithm has improved the feature distribution and the robot homing precision.

Keywords: visual homing; uniform distribution; SIFT feature; panoramic vision.

1. INTRODUCTION

Autonomous navigation is a challenging task in robotics. Some insects own incredible abilities of navigation. Inspired by this, some researchers have started studying the navigation methods that are based on insect behaviours. Among these methods, the visual homing methods have attracted much attention due to their low computation complexity and memory resources. In visual homing, there have been two major theories on the description of localization information: template hypothesis and parameter hypothesis (Moller 2001). Between them, the template hypothesis, which is mainly based on the snapshot model (Cartwright & Collett 1983), has a great influence. The snapshot model has been employed in the control of Micro Air Vehicle (Garratt & Lambert 2013) and robot. Hong et al. conducted homing experiments based this principle (Hong & Tan 1992) and came up with the ADV (Average Displacement Vector) method (Franz & Schölkopf 1998). The representative model of parameter hypothesis is the ALV (Average Landmark Vector) method (Lambrinos & Möller 2000), where the agent localizations are denoted by two dimensional vectors.

In both template and parameter hypothesises, there is a need of feature extraction from surrounding environments. During the verification experiments, artificial landmarks were widely used. In order to get around the limitation of use of artificial landmarks in un-structured environments, many researchers extract local invariant features as natural landmarks (Argyros & Bekris 2005). The SIFT (Scale Invariant Feature Transform) features (Lowe 2004) were widely employed in visual navigations due to their invariance to transform, rotation, changes in illumination and scale, and occlusion (Churchill & Vardy 2008). Ramisa et al. conducted homing experiments using ALV method by extracting SIFT features, and drew a conclusion that the homing precision of ALV method was affected by the uniformity of feature distribution (Ramisa & Goldhoorn 2011). In the ALV method, there is no specific requirement on isotropic landmark distribution. However, due to the lack of depth information, the ALV method implies an equal-distance assumption (Yu & Lee 2012). In addition, the ADV method is directly built on the equal distance assumption (Lambrinos & Möller 2000). Though SIFT features have good invariance, there is no control on the distribution. The majority of versions of SIFT algorithms have focused on the phase of feature description, and only a few have improved the feature extraction (Lingua, & Marenchino 2009, Sedaghta & Mokhtarzade 2011).

In this paper, we extend the conventional feature extraction algorithm SIFT to UD-SIFT (Uniform Distribution-SIFT). Compared with the SIFT, the proposal is capable of acquiring uniformly distributed features to improve the homing precision. The effectiveness of this algorithm was verified indoors and in the corridor, using the ADV and ALV methods.

2. ROBOT HOMING METHODS

For various visual homing methods, the homing precision increases with the uniformity of feature distribution. To identify the improvement of the proposed UD-SIFT algorithm, the ADV and ALV methods, which are influenced greatly by the feature distribution, are adopted in this study.

2.1 Average Displacement Vector

The landmark equal distance assumption is a key in many visual homing methods, and ADV is directly based on this. According to this assumption, the landmarks have the same distance from the agent location. What's more, the probability of observing the landmarks is assumed to be

identical, regardless of the agent location. With regard to all displacement vectors, the vector components orthogonal to the direction of the movement can cancel each other when all displacement vectors are added up under the circumstance.

When the robot moves from the home position to the current position, the landmark bearings L_i change from θ_{H}^i to θ_{C}^i (Fig. 1). Then the homing direction β_L^i is computed according to (1). The homing vector $\vec{M}$ and direction $\vec{\delta}$ are obtained by adding up all the features' homing direction. The error η between the actual homing direction and ideal homing direction will decrease as the robot gets closer to the home position. In addition, this method needs the assist of angle sensor, or the initial bearing remains unchanged.

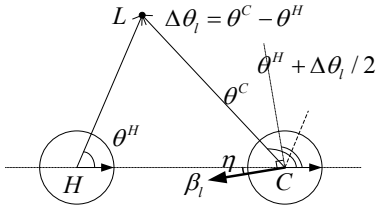


Fig. 1. Computation of the homing direction.

$$\beta_l = \begin{cases} \theta^H + \Delta\theta_l / 2 + \pi / 2 \ \ (\Delta\theta_l > 0) \\ \theta^H + \Delta\theta_l / 2 - \pi / 2 \ \ (\Delta\theta_l < 0) \end{cases} \tag{1}$$

$$\begin{cases} M_x = \sum_l \cos(\beta_l) \\ M_y = \sum_l \sin(\beta_l) \end{cases} \tag{2}$$

$$\vec{M} = (M_x, M_y) \tag{3}$$

$$\vec{\delta} = \frac{\vec{M}}{\| \vec{M} \|} \tag{4}$$

2.2 Average Landmark Vector

The landmark vectors $\vec{LV}_1$ and $\vec{LV}_2$ are unit vectors from robot position to landmark l_1 and l_2, respectively (Fig. 2). The average of all the landmark vectors is named as the average landmark vector. It is obvious that this vector can represent the features of different positions. The homing vector $\vec{M}$ can be computed according to (5) by comparing the average landmark vector $\vec{ALV}_C$ and $\vec{ALV}_H$ in current and home positions.

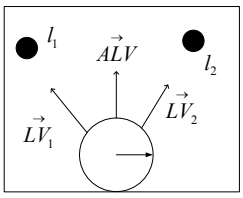


Fig. 2. Model of the ALV.

$$\vec{M} = \vec{ALV}_C - \vec{ALV}_H, \ \ \vec{ALV} = \frac{1}{n}\sum_{i=1}^{n} \vec{LV}_i \tag{5}$$

3. EXTRACTION OF THE UD-SIFT FEATURE

In the initial SIFT algorithm, the Gaussian pyramid is obtained by sub-sampling the original image and using a Gaussian kernel with different scales. The images of difference-of-Gaussian pyramid are differences between adjacent Gaussian images. The locations and scales of features are acquired by searching for local maximums. During the extraction phase, there is no control on the feature distribution. In the case that the distribution of features is uneven, the homing performance will turn down.

Because of the smoothing characteristics of the Gaussian kernel used in generating the pyramid, the image details and the number of features tend to decrease with the increase of Gaussian scales. Thus the number of features in each scale layer of each octave is set in an inverse proportion to the corresponding Gaussian scale. Supposing that the number of key-points extracted by the standard SURF algorithm in the scale layer (s) of the octave (o) is N_{os}, and the corresponding Gaussian scale is σ_{os}, N_{os} is inversely proportional to σ_{os}. If the required number of features is N and the proportion of N_{os} is P_{os}, then

$$\begin{cases} P_{os} = N_{os} / N, \ 1 \le o \le O, 2 \le s \le S-1 \\ \sum_{o=1}^{O}\sum_{s=2}^{S-1} P_{os} = 1 \end{cases} \tag{6}$$

During the construction of Gaussian pyramid, the Gaussian scale in the same scale layer of different octaves is identical. To discern the difference of feature number among different scale layers, the number of extracted features in each scale layer is determined by (7).

$$\begin{cases} \frac{N_{os}}{N_{o2}} = \frac{P_{os}}{P_{o2}} = \frac{\sigma_{o2}}{\sigma_{os}} \times \frac{2e^{-(o-1)}}{1+e^{-(o-1)}}, 2 \le s \le S-1 \\ N_{o2} = N_{(o-1)(S-1)} / 2, 2 \le o \le O \end{cases} \tag{7}$$

Then the distribution of features in each scale layer is considered. In consideration of the use of panoramic images, each scale layer can be divided into regular sector rings along the radial and circumferential directions, as shown in Fig. 3. The number of sector rings in the scale layer (s) of the octave (o) is n_{os}, and the N_{os} features should be evenly distributed in these sector rings. Supposing that the initial number of key-points in the ith sector ring is m_i, the average contrast is M_{con}^i, and the spatial divergence factor is D_{dist}^i, the required number N_{os}^i is computed according to (10). E_j is entropy of a circular region around the feature, and q_l is the histogram-driven probability within the region. By the control of feature number in each scale layer and sector ring, uniformly distributed features are obtained.

$$\begin{cases} D_{\text{dist}}^{i} = \sqrt{\dfrac{\sum_{j=1}^{m_i}(x_j-\overline{x})^2+\sum_{j=1}^{m_i}(y_j-\overline{y})^2}{m_i}} \\ (\overline{x},\overline{y}) = (\dfrac{\sum_{j=1}^{m_i}E_j x_j}{\sum_{j=1}^{m_i}E_j}, \dfrac{\sum_{j=1}^{m_i}E_j y_j}{\sum_{j=1}^{m_i}E_j}) \end{cases} \tag{8}$$

$$E_j = -\sum_{l} q_l \log_2 q_l, \quad j = 1,2,\cdots m_i \tag{9}$$

$$\begin{cases} N_{os}^{i} = N_{os}[\dfrac{W_{\text{num}} m_i}{\sum_{i=1}^{n_{os}} m_i} + \dfrac{W_{\text{dist}} D_{\text{dist}}^{i}}{\sum_{i=1}^{n_{os}} D_{\text{dist}}^{i}} + \\ \qquad \dfrac{(1-W_{\text{num}}-W_{\text{dist}})M_{\text{con}}^{i}}{\sum_{i=1}^{n_{os}} M_{\text{con}}^{i}}] \\ i = 1,2,\cdots n_{os}, \ 2 \le s \le S-1, 1 \le o \le O \end{cases} \tag{10}$$

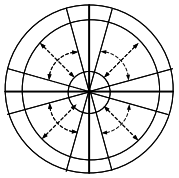


Fig. 3. Division of the DoG image.

In accordance with the aforementioned procedure, the UD-SIFT feature extraction algorithm is as follows.

(1) The N_{init} initial key-points are extracted by the SIFT algorithm, and those with low contrast and strong responses along edges are eliminated. Too many points increase the computing time. On the contrary, the low number may give rise to false matches. So it is necessary to determine the required number N according to different environments.

(2) Compute N_{os} according to (6) and (7).

(3) Divide each scale layer into n_{os} sector rings and compute the required number N_{os}^{i}. In each ring, $2\times N_{os}^{i}$ features with the highest entropy are reserved. And then N_{os}^{i} features with the highest contrast are selected. In the cases where it is not possible to select the required number of features in each ring, the problem of insufficient features will be addressed by a proportional increase of feature number in other rings.

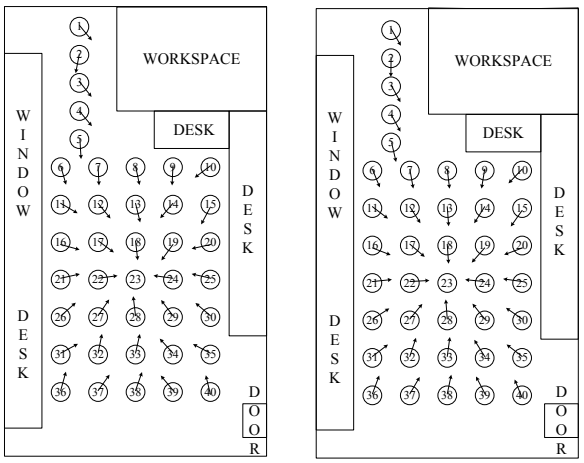
4. EXPERIMENTS

Our experiments were performed with a crawler mobile robot (Fig. 4), which was equipped with a digital compass. A panoramic vision system with a 360° visual field, composed of a mirror with a diameter of 100 mm and a camera with a resolution of 800 × 800, was mounted on top of the robot. When the robot moved indoors, N is a third of N_{init}. When the robot moved in the corridor, N is two thirds of N_{init}. The difference-of-Gaussian pyramid had four octaves, and each octave had five scale layers, that is to say, $O=4$ and $S=5$. There were $5\times5\times4$ sector rings in each scale layer of the first octave, and the others were $4\times4\times4$, $3\times3\times4$ and $2\times2\times4$, respectively. The radius of the circular region used to compute the entropy in (9) was set at $4.5\times\sigma_{os}$. What's more, W_{num} and W_{dist} were experimentally set at 0.2 and 0.3, respectively. In addition, the compass was used to compute the bearing of the robot.

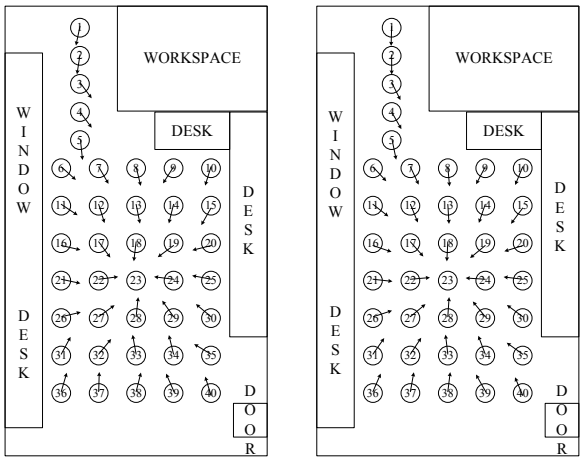
Fig. 4. Experimental system.

4.1 Homing Indoors

Forty locations in a room of 10m×5m were selected to capture panoramic images (Fig. 5). The number in the circle was the location number, and the location with number of 23 was the home position. Moreover, the direction of arrow was the homing direction by calculation. Fig. 6(a)-6(d) were the homing errors using the initial SIFT algorithm and UD-SIFT algorithm with ADV and ALV methods, respectively. Obviously, the homing errors were relatively small when the UD-SIFT algorithm was applied. This can be explained by the improvement of feature distribution.

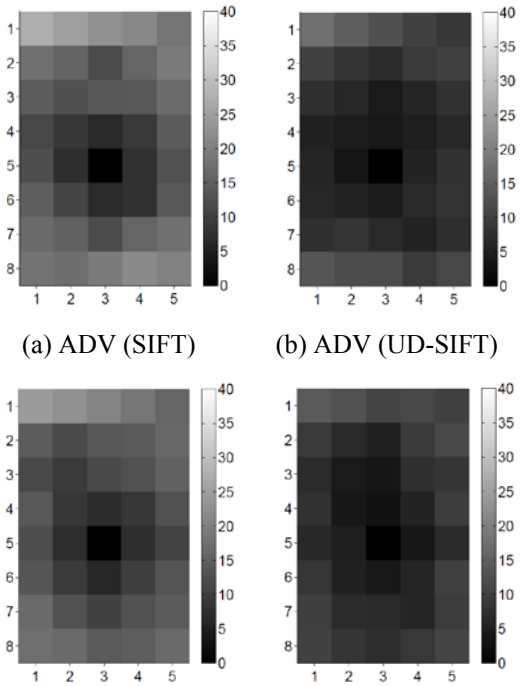


(a) ADV (SIFT) (b) ADV (UD-SIFT)



(c) ALV (SIFT) (d) ALV (UD-SIFT)

Fig. 5. Vector maps of homing indoors.



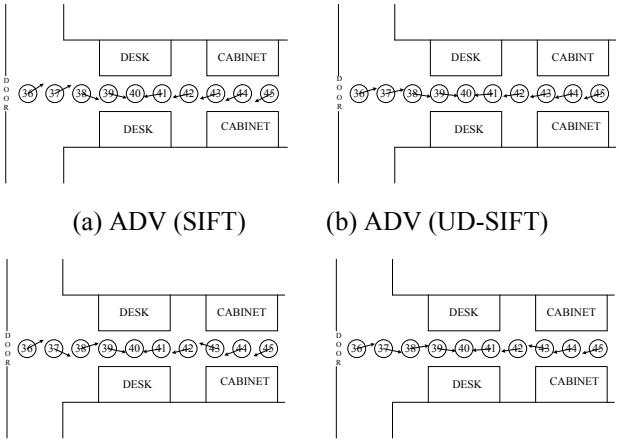
(a) ADV (SIFT) (b) ADV (UD-SIFT)

(c) ALV (SIFT) (d) ALV (UD-SIFT)

Fig. 6. Homing errors indoors.

4.2 Homing in the Corridor

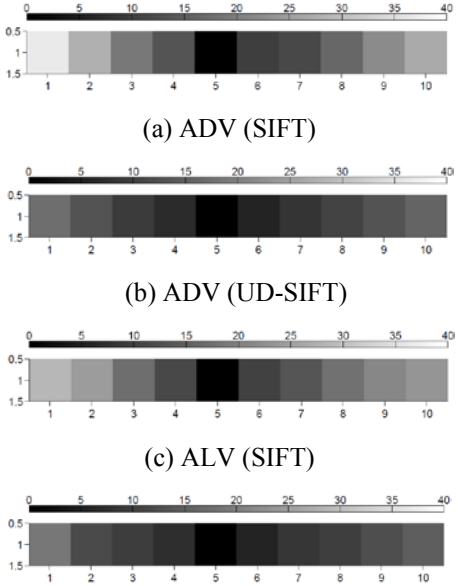
Ten locations in the corridor of 15m×3m were selected to capture panoramic images (Fig. 7). The location with number of 40 was the home position. Fig. 7 shows the vector maps of homing using different methods, and Fig. 8 corresponding to the homing errors. Though the homing errors in the corridor were relatively high, the UD-SIFT algorithm indeed improves the homing precision.



(a) ADV (SIFT) (b) ADV (UD-SIFT)

(c) ALV (SIFT) (d) ALV (UD-SIFT)

Fig. 7. Vector maps of homing in the corridor.



(a) ADV (SIFT)

(b) ADV (UD-SIFT)

(c) ALV (SIFT)

(d) ALV (UD-SIFT)

Fig. 8. Homing errors in the corridor.

4.3 Analysis of Experimental Results

The experimental results show that the UD-SIFT algorithm has improved the homing performance of both ADV and ALV methods. The average homing errors obtained through the above experiments are presented in Table 1. The increase of homing precision is credited to the improvement of feature distribution. In addition, the experimental results further validate that the ADV and ALV methods will perform better with uniformly distributed features.

Table 1. Average homing errors (°)

Environment	ADV		ALV	
	SIFT	UD-SIFT	SIFT	UD-SIFT
Indoor	35.4	21.1	31.2	17.6
Corridor	56.3	36.9	53.7	33.4

The better results were obtained with the panoramas captured indoors, compared with the results obtained in the corridor. When the robot moves in the region with high length-width ratio, the imaging positions on the panorama get closer to each other as the increase of distances between features and the robot (Fig. 9), thus the uniformity of feature distribution is worse.

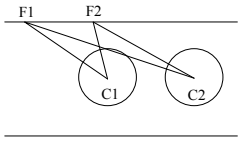


Fig. 9. Feature projection in the panorama.

5. CONCLUSION

Autonomous navigation is the critical technology for mobile robot to realize intelligence. Visual homing, which is inspired by insect navigation, is an important research direction due to the advantages of low storage and simple computation. When the robot performs tasks in un-structured environments, it often needs to extract local invariant features as nature landmarks. What is more, the distribution of features always affects accuracy. In order to increase the uniformity of feature distribution, an improved feature extraction algorithm UD-SIFT is proposed. The experimental results show that the UD-SIFT algorithm does significantly improve the feature distribution and homing accuracy. However, the number of extracted features is manually determined in advance. The robot can succeed in homing with only a few features. The restriction on number of features without decrease in homing accuracy will be the focus of future research.

REFERENCES

Moller, R. (2001). Do insects use templates or parameters for landmark navigation. *Journal of the Oretical Biology*, 210(1), 33–45.

Cartwright, B.A. and Collett, T.S. (1983). Landmark learning in bees: experiments and models. *Journal of Comparative Physiology*, 151(4), 521–543.

Garratt, M.A., Lambert, A.J., and Teimoori, H. (2013). Design of a 3D snapshot based visual flight control system using a single camera in hover. *Autonomous Robots*, 34(1), 19–34.

Hong, J.W., Tan, X.N., Pinette, B., Weiss, R., and Riseman, E.m. (1992). Image-based homing. *IEEE Control Systems Magazine*, 12(1), 38-45.

Franz, M.O., Schölkopf, B., Mallot, H.A., and Bülthoff, H.H. (1998). Where did I take that snapshot? scene-based homing by image matching. *Biological Cybernetics*, 79(3), 191–202.

Lambrinos, D., Möller, R., Labhart, T., and Pfeifer, R. (2000). A mobile robot employing insect strategies for navigation. *Robotics and Autonomous Systems*, 30(1), 39–64.

Argyros, A.A., Bekris, K.E., and Orphanoudakis, S.C. (2005). Robot homing by exploiting panoramic vision. *Autonomous Robots*, 19(1), 7–25.

Lowe, D.G. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2), 91–110.

Churchill, D. and Vardy, A. (2008). Homing in scale space. *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Nice, France, pp. 1307-1312.

Ramisa, A., Goldhoorn, A., Aldavert, D., and Toledo R. (2011) Combining invariant features and the ALV homing method for autonomous robot navigation based on panoramas. *Journal of Intelligent & Robotic Systems*, 64(3/4), 625–649.

Yu, S., Lee, C., and Kim, D. (2012). Analyzing the effect of landmark vectors in homing navigation. *Adaptive Behavior*, 20(5), 337-359.

Lingua, A., Marenchino, D., and Nex, F. (2009). Performance analysis of the SIFT operator for automatic feature extraction and matching in photogrammetric applications. *Sensors*, 9(5), 3745–3766.

Sedaghat, A., Mokhtarzade, M., and Ebadi, H. (2011). Uniform robust scale-invariant feature matching for optical remote sensing images. *IEEE Transtractions on Geoscience and Remote Sensing*, 49(11), 4516–4527.