Development of an Environmentally Adaptable Autonomous Mobile Robot

Iwata Naokazu
Graduate School of Engineering, Kyushu Institute of Technology
Kitakyushu, 804-8550, Fukuoka, Japan

Joo Kooi Tan
Faculty of Engineering, Kyushu Institute of Technology
Kitakyushu, 804-8550, Fukuoka, Japan
E-mail: etheltan@cntl.kyutech.ac.jp

Abstract

To realize a service-robot system for supporting a human life, we propose techniques for estimating self-position of a robot using local features of images, segmenting an image for finding a movable area, and also planning a route for finding the destination. In the route planning, junctions are labelled numbers for the robot to travel in order. The developed mobile robot travels to the destination employing the information on the estimated road region, its self-position and the planned route.

Keywords: Autonomous Mobile Robot, Oriented FAST and Rotated BRIEF, Bag of Features, K-means++

1. Introduction

In recent years, robot technology has been widely used not only in industrial robots, but also as service robots for supporting a human life. One of the main backgrounds is the technological innovation of both hardware and software such as sensor technology, artificial intelligence technology, and data processing technology which are necessary for robots. In particular, much attention has been focused on service robots for supporting a human life. These kinds of robots are expected to be used in a daily life, and necessary to adapt flexibly to various usage, scene, or even to environmental change. This paper proposes an autonomous mobile robot system that adapts to various environments and travel routes.

Robot self-localization methods of previous work include using sensors such as GPS, LiDAR[1,2], markers[3], and template matching[4]. However, these methods have the problem that they depend on the sensor characteristics, limit in setting the location of markers, and affected by the weather.

In this paper, we propose a self-localization method using local features based on a RGB-D camera. Since the developed robot needs to move autonomously in various environments, Oriented FAST and Rotated BRIEF (ORB)[5] are used for local features. The ORB procedure is high speed without affected by scaling, rotation and illumination. In order to reduce the local features calculation cost, we use Bag of Features (BoF)[6] that represents an image by a feature vector However, for developing an autonomous mobile robot, it is necessary to find an area where the robot can travel. In this study, we define that an initial area is a movable road region of the robot. We segment the image into areas and find a road region/area which is a similar area as the initial robot area. For travelling autonomously to a destination, the robot must plan a route from the current location to the destination. In this study, for the route planning, junctions, corners and a destination are labelled numbers for the robot to travel in order. The developed mobile robot travels to the destination employing the information on the estimated road region, its self-position and the planned route.

2. Self-Position Estimation of a Robot

In this section we describe a robot's self-position estimation method based on local features. Since many local features are extracted from a single image, comparison and verification of each local feature, furthermore estimating the position based on the results of similarity or coincidence of each local feature increases the calculation cost. Therefore, in the proposed method, local features are extracted from the image, and the obtained local features are converted to a histogram representing the frequency of appearance of local features using Bag of Features. The robot's self-position is estimated from the similarity of the histogram.

2.1. *Bag of Features*

Bag of Features (BoF) represents one image as one feature vector. The process of BoF is shown as follows.
(i) Extracting local features from images.
(ii) Creating Visual Words by clustering local features.
(iii) Creating a histogram of image based on Visual Words.

(i) *Extracting local features*

We extract local features using Oriented FAST and Rotated BRIEF (ORB). **Fig. 1** shows an example of feature points extracted by ORB. The points surrounded by red circles in Fig. 1 are extracted feature points. Here, let us suppose that $\boldsymbol{F}_j(j = 1,2,\ldots,m)$ local feature points are extracted from an image.

(ii) *Clustering the local features*

We apply K-means++[7] to cluster the local features $\boldsymbol{F}_j$. The K-means++ is an improved version of the K-means method[8] It can cluster the local features independently on the initial cluster allocation.

(iii) *Creating a histogram*

After having clustered the local features, the center of each cluster $i(i = 1,\ldots,n)$ is defined as Visual Words (VWs) $\boldsymbol{V}_i(i = 1,2,\ldots,n)$ The nearest VW is searched from each local feature $\boldsymbol{F}_j(j = 1,2,\ldots,m)$, and a histogram is created by voting to the VWs by considering

Fig. 1 Example of extracted feature points from ORB

a score . The score is obtained by the following equation to determine which VWs is closer to each local feature.

$$s = 1 - \frac{1}{2}\left\| \frac{\boldsymbol{V}_i}{\|\boldsymbol{V}_i\|} - \frac{\boldsymbol{F}_j}{\|\boldsymbol{F}_j\|} \right\| \tag{1}$$

Here, $\boldsymbol{V}_i(i = 1,2,\ldots,n)$ is the VWs obtained by clustering and $\boldsymbol{F}_j$ is the local feature of the image obtained by ORB. By voting $\boldsymbol{F}_j$s to the VWs with the smallest score s, a histogram representing the frequency of the appearance of local features is created. The histogram is a feature vector that represents an image.

2.2. *Self-position Estimation of the Robot*

In order to estimate the robot's self-position, the histograms of local features are created using the BoF described in Section 2.1 from an input image and key frame images such as junctions. Next, the similarity is calculated between the input image and the key frame images using the histograms. The robot's self-position is estimated based on the similarity. The procedure of the robot's self-position estimation is shown below.

(i) The appearance frequency histograms H_{in} and H_{key} of local features are created by BoF from the input image and the key frame image, respectively.
The similarity (S_{SAD}) is calculated between H_{in} and H_{key}. It using Sum of Absolute Difference (SAD) defined by

$$S_{SAD} = \sum_{n=0}^{N-1} \left| H_{in} - H_{key} \right|_n \tag{2}$$

where N is the number of classes of the VWs.

(ii) To make the self-position estimation more accurate, the sum of the similarities of the past M input images is used. If it is less than or equal to a threshold T_s, it is judged that the robot is at the position where the key frame indicates.

3. Road Region Estimation

Road region is estimated using Graph Based Segmentation[10], which is one of the image segmentation methods.

3.1. *Graph Based Segmentation*

Graph Based Segmentation (GBS) is one of the renowned image segmentation methods, which combines pixels with similar characteristic pixel values into multiple regions. The algorithm of GBS is shown below.

(i) Smoothing the input image $I(x,y)$ using the following equation.

$$L(x,y,\sigma) = G(x,y,\sigma) \cdot I(x,y) \tag{3}$$

$$G(x,y,\sigma) = \frac{1}{2\pi\sigma}\exp\left(-\frac{x^2+y^2}{2}\right) \quad (4)$$

Here σ is the standard deviation, $G(x,y,\sigma)$ is the Gaussian function, and $L(x,y,\sigma)$ is a smoothed image.

(ii) Creating a graph with the node v_i representing each pixel of the smoothed image and the edge e_q connecting adjacent pixels in the image. Using the illuminance difference between the pixels connected by an edge, the weight of the edge $w(e_q)$ is calculated by the following equation.

$$\omega(e_q) = \sqrt{(R_i - R_j)^2 + (G_i - G_j)^2 + (B_i - B_j)^2} \quad (5)$$

where, $R_\#$, $G_\#$, and $B_\#$ are the red, green and blue values of the node.

(iii) Assignment of a separate region $S_q (q = 1, \ldots, m)$ to each node where m is the total number of pixels in the image.

(iv) When nodes v_i and v_j connected by edge $e_q(q = 1, \ldots m)$ satisfy the following equation, the region S_i and S_j are combined.

$$\omega(e_q) \le \min\left(\max_{e \in E_i}\{\omega(e)\} + \frac{c}{|S_i|}, \max_{e \in E_j}\{\omega(e)\} + \frac{c}{|S_j|}\right) \quad (6)$$

Here, c is a predetermined fixed value, $|S_\#|$ is the number of nodes constituting the region $S_\#$, and $E_\#$ is a set of edges connecting nodes in the region $S_\#$.

For all edges, the node is divided into several regions by determining the region connection by Equation (6). If, with a certain region, the number of nodes is less than or equal to a threshold S_{th}, the region is merged with an adjacent region with which an edge connecting the two regions has the smallest weight among adjacent regions.

3.2. *Road Region Estimation*

The input image is segmented by GBS described in Section 3.1. In the proposed method, the rectangle region surrounded by a red frame as shown in **Fig. 2** is defined as the robot's foot candidate region. The maximum region which contains the foot candidate region is regarded as the road region. If the area of the provisionally estimated road region A_R is smaller than a threshold A_{th}, the latest road region in the past frames is chosen as the current

Fig. 3 Example of the robot's foot region

Fig. 2 Example of estimation of a road region

road region. **Fig. 3** shows an example of the road region. The gray area is the estimated road region.

4. Algorithm on Autonomous Movement

First, the robot estimates its own position using local features obtained from input images. Second, the road region is estimated from the input image by GBS. Third, it is examined whether the estimated robot's self-position is at a key frame position in the database or not. If yes, the key frame is examined whether it is the destination or not. If yes, the robot stops. If the key frame is not the destination, the robot's direction of travel is changed to the direction determined in the route setting. If the robot's self-position is not at a key frame position in the database, the estimated road region is divided into three sub-regions in the lower part of the image. We calculate the area of each sub-region. If the area is small, it is judged that there are few areas to advance ahead of the robot, which is dangerous, and the robot stops temporarily. If the areas of the sub-regions are larger than a certain threshold, the robot moves in the direction of the sub- region having the maximum area by finely adjusting its direction.

5. Experiment

5.1. *Experimental Environment*

The robot moves autonomously in an outdoor environment. The robot travels along two types of travel routes, route A and route B, 5 times each. The performance of the robot is evaluated by if it has reached the destination successfully. If an emergency stop is necessary due to the passage of a vehicle during the run, the experiment is interrupted and is resumed after the vehicle has passed. **Table 1** shows the parameters related to this experiment. The forward speed of the robot is 0.2 [m/s] and the rotational speed is 0.2 [rad/s].

5.2. *Experimental Results*

Table 2 shows the results of the experiment. In Table 2, '○' indicates that the robot reached the destination successfully, and '×' indicates failure. The processing time was 65.4 [ms/frame]. The travel of the robot was successful 3 times

Table. 1 Parameters used in the experiment

No. dimensions of ORB	32
No. features	1000
No. clusters	500
No. learning images	1

Table. 2 Experimental results

	1	2	3	4	5
route A	○			○	○
route B		○	○	○	○

out of 5 with route A, whereas it was successful 4 times out of 5 with route B. As for the weather, the success cases were 2 times out of 4 when it was fine, whereas they were 5 times out of 6 when cloudy.

6. Discussion

In the performed experiment, as shown in Table 2, the destination could be reached in many cases regardless of the travel route and the weather. This is thought to be due to the high accuracy of self-position estimation and road region estimation. However, in the third experiment on route A, the destination could not be reached. This is because the estimation of the road region was incomplete. When the shadow is in the road region, the part that is originally the same region is divided into some regions like the situation in the white circles as shown in **Fig. 4**. Moreover, in this study, the road region was determined using the maximum area's region among the divided sub-regions, so the estimation of the road region was incomplete. More robust road region estimation may be possible by using an algorithm that does not depend on color and includes a process of removing shadows.

7. Conclusion

In this paper, we proposed an environmentally adaptable autonomous mobile robot system. In this system, the robot's self-position estimation based on the local feature of the image and the road region estimation using the image segmentation were performed. In the experiment on autonomous movement, it was possible to reach the destination 3 times out of 5 times with the travel route A and 4 times out of 5 times with the travel route B. By these

Fig. 4 Example of a shaded road region

results, the effectiveness of the proposed autonomous mobile robot system was confirmed. In addition, it was possible to reach the destination 2 times out of 4 times when it was fine, and 5 times out of 6 times when it was cloudy. This also confirms the effectiveness of the proposed system for various weather conditions.
In the system, only local features of input images are used for robot's self-position estimation. However, it is considered that self-position estimation can be performed with high accuracy by using distance information in addition to local features. Moreover, because the road region was estimated by color, there were cases where the estimation of the road region was incomplete. Therefore, it is considered that the road region can be estimated in various environments by devising a method that does not depend on color.

8. References

1. Masatoshi Shinohara, S. A. Rahok, Kazumichi Inoue, Koichi Ozaki : "Development of Localization Method Using Magnetic Sensor and LIDAR", *Transactions of the Society of Instrument and Control Engineers*, Vol.49, No.8, 795-801, 2013.
2. Hisashi Date, Shinya Ohkawa, Yoshihiro Takita, Jun Kikuchi : "High Precision Localization of Mobile Robot Using KIDAR Intensity of Surface", *The Japan Society of Mechanical Engineers*, Vol.79, No.806, pp.100-109, 2013.
3. J. Wendlerr, S. Bruggert, H. Burkhard, H. Myritz, "Fault-tolerant self localization by case-based reasoning", *RoboCup 2000: Robot Soccer World Cup IV*, pp,259-268, 2001.
4. Kosei Demura, Yu Nakagawa : "A Monte Carlo Localization based on Template Matching using an Omni-directional Camera", Journal of the Robotics Society of Japan, Vol.27, No.2, pp.249-257, 2009.
5. E. Rublee, V. Rabaud, K. Konolige, G. R. Bradski : "ORB: An efficient alternative to SIFT or SURF", *Proceedings of IEEE International Conference on Computer Vision*, pp.2564-2571, 2011.
6. G. Csurka : "Visual categorization with bags of keypoints", *Proceedings of ECCV Learning in Computer Vision*, pp.1-22, 2004.
7. D. Arthur, S. Vassilivitskii : "k-means++: The advantage of careful seeding", *Proceedings of the 18 Annual ACM-SIAM Symposium on Discrete Algorithm*, pp.1027-1035, 2007.
8. P. Agarwak, N. Mustafa, "k-Means Projective Clustering", *PODS '04 Proceedings of the twenty-third ACM DIGMOD-SIGACT-SIGART symposium on Principles of the database systems*, pp.155-165, 2004
9. P. Felzenswalb, D. Huttenlocher : "Efficient graph-based image segmentation", *International Journal of Computer Vision*, Vol.59, pp.167-181, 2004.
10. M. Calonder, V. Lepetit, C. Strecha, P. Fua, "BRIEF: Binary Robust Independent Elementary Features", *11th European Conference on Computer Vision*, pp.1-14, 2010.