Paper:

Robust global localization using laser reflectivity

DongXiang Zhang*, Ryo Kurazume*, Yumi Iwashita*, Tsutomu Hasegawa*

*Graduate School of Information Science and Electrical Engineering, Kyushu University
744 Motooka, Nishi-ku, Fukuoka 819-0395, Japan
E-mail: zhang@irvs.ait.kyushu-u.ac.jp

Global localization, which determines an accurate global position without prior knowledge, is a fundamental requirement for a mobile robot. Map-based global localization gives a precise position by comparing a provided geometric map and current sensory data. Although 3D range data is preferable for 6D global localization in terms of accuracy and reliability, comparison with large 3D data is quite time-consuming. On the other hand, appearance-based global localization, which determines the global position by comparing a captured image with recorded ones, is simple and suitable for real-time processing. However, this technique does not work in the dark or in an environment in which the lighting conditions change remarkably. We herein propose a two-step strategy, which combines map-based global localization and appearance-based global localization. Instead of camera images, which are used for appearance-based global localization, we use reflectance images, which are captured by a laser range finder as a byproduct of range sensing. The proposed method consists of two steps: i) reflectance images acquired by a laser range finder are used for rough estimation of global position based on the bag-of-features (BoF) technique, and ii) precise global positions are determined automatically by the iterative closest points (ICP) algorithm for 3D range data. The effectiveness of the proposed technique is demonstrated through experiments in real environments.

Keywords: Global localization, Appearance-based localization, Map-based localization, Laser range finder, Reflectance image

1. INTRODUCTION

In numerous practical applications, the external environment around a robot is unpredictable, unstructured, and uncontrolled. For example, in an area that has been struck by a strong earthquake or by a mine disaster, previously navigable areas may be blocked by heaps of rubble or collapsed walls, and the geometric structure will differ from the original structure. In order for a robot to efficiently accomplish search and rescue tasks in this type of unknown and unpredictable environment, accurate map creation and localization are fundamental requirements.

Global localization, which is the problem of determining an accurate position in a global coordinate system using surrounding features given a map but has no prior knowledge, is also a basic requirement for mobile robots. A number of global localization techniques have been proposed [11], and these techniques can be classified into two groups, i.e., map-based global localization and appearance-based global localization.

Map-based global localization is a commonly used, traditional technique that determines a global position using a provided geometric map and sensory data[1]. The map can consist of either two-dimensional (2D) or three-dimensional (3D) structures. The problem is to find the best position at which the observed geometric features match those in the provided geometric map. From the point of view of accuracy, comparison of 3D range data captured by a range sensor and a pre-constructed 3D map is preferable because this will enable precise 6D (position and attitude) global localization. However, the comparison of large 3D data is quite time-consuming.

On the other hand, appearance-based global localization is a technique that determines a global position using camera images[14],[15],[16]. Many camera images are recorded under natural light or indoor illumination in the environment in which the robot operates, and global localization is performed by finding the best match using a newly-captured image and stored images. Appearance-based global localization is simple and suitable for real-time processing. Moreover, appearance-based global localization is very similar to the method used by human beings. However, appearance-based global positioning encounters a critical problem in dark environments or in environments in which the lighting condition changes dramatically.

We herein propose a two-step strategy that combines map-based global localization and appearance-based global localization. Instead of camera images, which are used for the appearance-based global localization, we use reflectance images, which are captured by a laser range finder as a byproduct of range sensing. As a result of the characteristics of the reflectance image, which is not subject to significant variation of the external illumination conditions, the proposed technique is useful even in the dark or in an environment under severe lighting conditions. Furthermore, fast and precise localization can be performed by comparing a few 3D range

images, which are selected based on the similarity of the reflectance images. The proposed two-step strategy is as follows: i) reflectance images acquired by a laser range finder are used for rough estimation of a global position based on the bag-of-features (BoF) technique, and ii) precise global positions are automatically determined by the iterative closest points algorithm for 3D range data.

The remainder of the present study is organized as follows. After a brief introduction of related research in Section 2, Section 3 introduces the cooperative positioning system on which the proposed method is based. The proposed two-step strategy is described in detail in Sections 4 and 5, and experimental results are presented in Section 6.

2. Related research

In the robotics community, several appearance-based localization systems have been developed. These systems differ primarily in the features used to match images and can be categorized into two major types. The first uses local features extracted from small regions that are robust to variations and occlusions. Vertical lines features and their line support region (LSR) have been used for this purpose [17]. The second type uses global features extracted from the entire image. The first 15 Fourier components have been used as a compact representation of omnidirectional images [18]. In a previous study [20], the matching of the local features named Modified SIFT (MSIFT) algorithm was used, and compares MSIFT with three other features, including one global feature and two local features.

The BoF technique is a popular technique that combines the advantages of local and global features for efficient representation of raw images captured by a camera. The BoF technique first extracts local features of an image and then constructs a histogram using the extracted local features as a representation of the global features. The BoF technique is successfully applied to the SLAM problem. In [2], the loop closure detection, which is a problem to find a location where the robot previously visited, was solved by the BoF technique in a probabilistic manner. However, these methods assume that the illumination condition does not change so severely. Therefore, the travel distance of the robot must be short enough so that the lighting condition will not change drastically. In another study, the experimental results were not evaluated explicitly under significant changes in illumination [3]. These methods will fail in the dark or in environments in which the lighting is dramatically changed.

Panoramic vision has become prevalent in recent years and has several advantages compared to conventional camera images. Omnidirectional cameras have been used in numerous studies on robotic self-localization [18][17][21][20]. Another approach that does not use omnidirectional images extracts a visibility region σ_M for each image M. The visibility region σ_M corresponds to all positions in a given metric map of the environment from which the closest object in image M in the direction of the optical axis is visible. This method can be considered to be an expansion of panoramic vision to a conventional camera.

A laser range finder, which measures the distance from the sensor to the surroundings, is a popular device for robot localization, map creation, and 3D modeling. When range is measured by a laser range finder, the reflectivity, which indicates the strength of the reflected laser, can be obtained as a byproduct of range data. Note that all of the pixels in the range image have corresponding reflectance values. In other words, the range image and the reflectance image are precisely and fundamentally aligned. In addition, since the reflectance image is not subject to any extreme variations in the external illumination conditions, stable reflectance images can be obtained even at night.

The proposed technique uses a panoramic reflectance image instead of a regular camera image. By applying the BoF technique for a reflectance image that corresponds to 3D range data, the global localization using 3D range data and 2D images is achieved efficiently. Hara et al.[4] utilizes reflection intensity from a 1D laser range finder for localization in 2D space. On the other hand, our technique utilizes 2D reflectance images obtained by a 3D laser range finder for 3D localization. This technique is deterministic and so does not rely on the Markov assumption, in which the future state of the robot, given the present and past states, depends on only the present state [13]. Moreover, this technique does not encounter the kidnapped robot problem [12], because its current state depends purely on the current sensor reading.

The proposed technique is similar to that described in [21] in that it uses pyramidal matching kernels to obtain a topological localization and then uses the most similar image found in the image database and the 1D trifocal tensor estimated from three view feature matches to achieve accurate metric localization. In addition to the illumination requirements, the proposed technique would fail in non-planar locations because of the use of a 1D trifocal tensor. In Section 6, we demonstrate the capability of the proposed technique in various environments.

3. Development of a 3D global map by the cooperative positioning system

3.1. Laser-based environmental modeling by multiple mobile robots

For map-based global localization, an environmental map must be created and provided beforehand. For precise 3D mapping of the environment around a robot, we have proposed an efficient and precise system called CPS-SLAM [5], which can construct a rather accurate large-scale 3D map by means of a laser range finder and multiple robots based on technology used for geographical surveying. This technique is used as the basis of the urban search and rescue (USAR) robot [6].

Figure 1 shows the fifth CPS-SLAM model, called

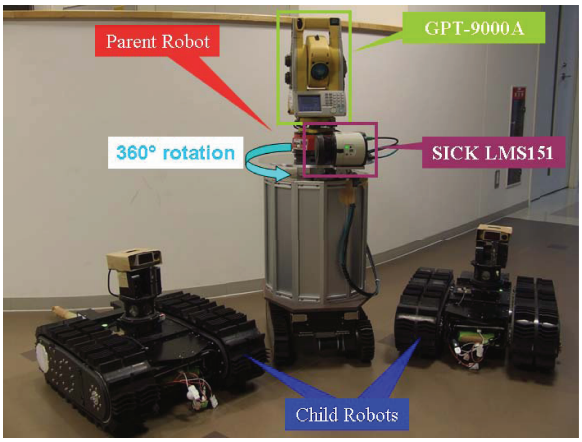


Fig. 1. Three-dimensional modeling robots, CPS-V

Table 1. Laser range finder (SICK LMS151)

Measuring range	$50[m]$
Field of view	$270[\circ]$
Precision	$\pm 30[mm]$
Angular resolution	$0.25[\circ]$

CPS-V. This system consists of one parent robot and two child robots. The parent robot is equipped with a highly precise laser range finder (GPT-9000A, TOPCON LTD), a 2D laser range finder (SICK LMS151), and a three-axis attitude sensor. The two child robots are equipped with corner cubes. The GPT-9000A and corner cubes are used cooperatively for self-positioning, as shown in Fig.2. The LMS151 (Table 1) placed on a rotating table acquires two-dimensional slit-like range data, which are vertical to the ground. This sensor can capture reflectance data at the same time. Therefore, by rotating the table around the vertical axis for 360° while scanning with a 2D laser range finder, 3D range data and a 2D reflectance image are acquired. The number of pixels on a reflectance image is exactly the same as the number of 3D points in the range data, i.e., there exists a one-to-one mapping relationship between 2D pixels of the reflectance image and 3D points of the local 3D map.

3.2. Three-dimensional global map

The process of mapping the entire field is displayed in Fig.3. In each location, the parent robot collects a local 3D map and its measurement position based on the relative observation between the parent and child robots. Eventually, all of the local 3D maps are aligned into a global 3D map using the measurement position information. Additional details about CPS-based simultaneous localization and mapping (CPS-SLAM) can be found in [5].

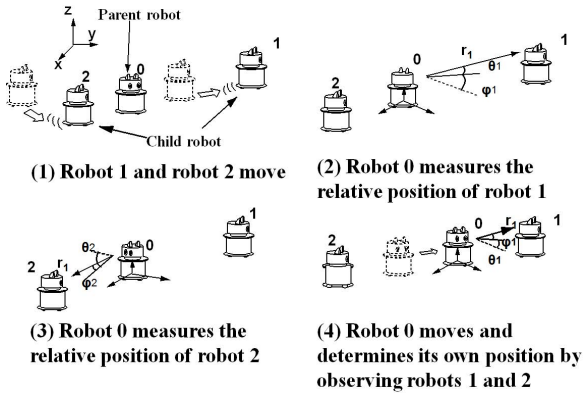


Fig. 2. Cooperative positioning system (CPS)

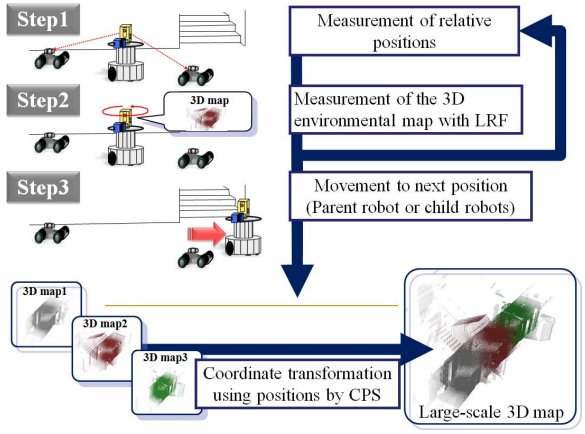


Fig. 3. Construction of a large-scale 3D map by the CPS

3.3. Reflectance images

As mentioned above, reflectance images are captured as a byproduct of range sensing. Two examples of a reflectance image and its corresponding 3D data acquired by CPS-V are shown in Fig.4. Note that each reflectance image contains the position information for the location at which the image was captured. These images are used for appearance-based global localization in the proposed two-step strategy.

3.4. Image retrieval using the BoF technique

When the global 3D map of the target field is constructed, a Kd-tree structure storing BoF representations of reflectance images is also constructed at the same time. Reflectance images are represented as histograms of occurrence of the visual words in an image. First, regions in feature space are mapped to visual words by clustering all SURF [7] or SIFT [8] features extracted from recorded images into representative words using k-means clustering, and the words are stored using a Kd-tree structure. Using these words as the x-axis, we quantize each feature in a reflectance image to its approximate nearest word by searching the Kd-tree, and all of the recorded reflectance images are represented as statistics of words (histograms).

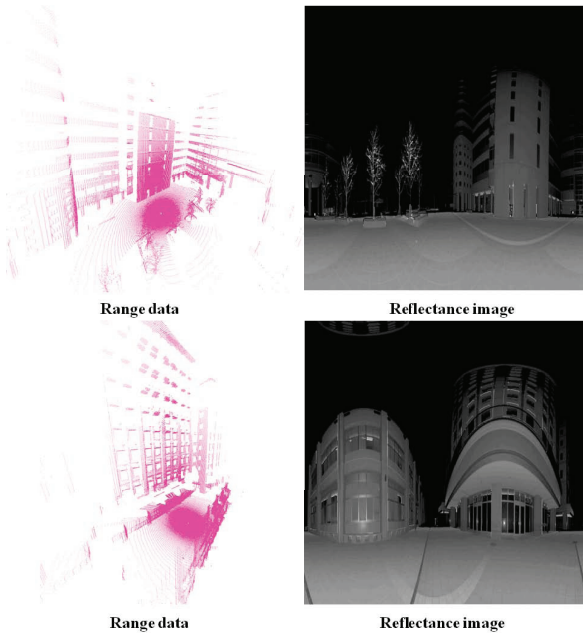


Fig. 4. Reflectance and range images

Finally, the histograms of all recorded images are stored using a Kd-tree structure. A newly-captured image is also represented as a histogram, and M images that best match the newly-captured image are retrieved by quantizing the histogram to its nearest M histograms.

4. Two-step strategy combining appearance-based and map-based global localization

This section presents the two-step strategy for precise localization using a 3D map. First, we need to create a global map as a training dataset. As explained in Section 3, the CPS robots move in the environment and construct a 3D global map. At the same time, the parent robot collects reflectance images at each measurement position. Then, all of reflectance images are represented using the BoF technique, and a training dataset is created. Finally, the dataset of all of the BoF representations is stored in a Kd-tree, which is efficient for information retrieval.

For global localization, a new robot that is equipped with a 3D range sensor, such as CPS-V, collects local 3D data and 2D reflectance images (test data). In the first step, we retrieve initial location candidates by comparing stored reflectance images (training dataset) and captured reflectance images (test data) using the BoF technique and a Kd-tree. We then apply a 3D geometric constraint in order to extract true feature pairs and run automatic ICP in the second step. As a result of the first step, fast and precise localization can be performed in the second step by comparing a few 3D range images, which are selected based on the similarity of the reflectance images in the first step.

We hereinafter denote variables related to training data as Tr_m or D_{tr} and variables related to test data as Te_n or D_{te}, where D indicates the 3D distance between two points in a local 3D map. Moreover, $Train\_i.ref$ and $Train\_i.pts$ represent the i^{th} reflectance image and the local 3D map, respectively, in the training dataset, and $Test\_j.ref$ and $Test\_j.pts$ represent the j^{th} reflectance image and the local 3D map, respectively, in the test dataset.

In the following, the proposed two-step strategy is described in detail.

4.1. First step: initial localization by BoF using 2D reflectance images

All $Test\_j.ref$ are converted into BoF representations, and the M best matches are searched in the Kd-tree that was constructed from the training dataset. The M $Train\_i.ref$ and $Train\_i.pts$ are then selected as M candidates for the position of the robot in the 3D global map.

4.2. Second step: precise localization by automatic ICP using 3D data

With M candidate positions, automatic ICP [9][10], which consists of two processes is applied to remove incorrect candidates. However, before applying automatic ICP, 3D geometric constraints are used to remove outliers for using RANSAC, as follows:

1 Rough alignment with RANSAC

   a Find the corresponding features between $Test\_j.ref$ and $Train\_i.ref$.

   b Get the 3D coordinates of corresponding features using $Test\_j.pts$ and $Train\_i.pts$, which correspond to $Test\_j.ref$ and $Train\_i.ref$, respectively.

   c Remove outliers by 3D geometric constraints. This process will be explained in Section 5.

   d 3D transformation between $Test\_j.pts$ and $Train\_i.pts$ is estimated by RANSAC.

   e Align $Test\_j.pts$ to $Train\_i.pts$.

2 Precise alignment with ICP

   a Run ICP[9] using $Test\_j.pts$ and $Train\_i.pts$, which are already roughly aligned.

As a result of ICP, two metrics are defined for evaluating the accuracy of the alignment between $Test\_j.pts$ and $Train\_i.pts$. One is the alignment ratio, and the other is the average error. First, we set a threshold of maximum distance between a pair of 3D points in $Test\_j.pts$ and $Train\_i.pts$. Suppose $Test\_j.pts$ has K points in total, N of which have corresponding points in $Train\_i.pts$. The alignment ratio is defined as N/K. In these N pairs of points, the sum of errors between each pair of points divided by N is defined as the average error.

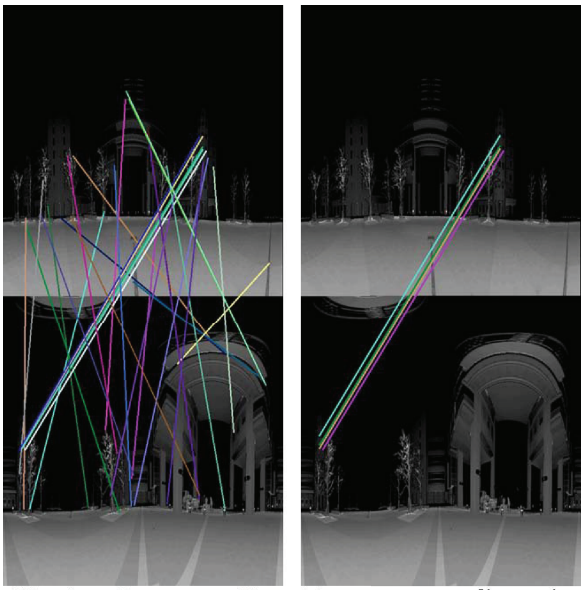


Fig. 5. Corresponding features between $Test\_17.ref$ and $Train\_22.ref$. The left-hand image shows all of the correspondences, and the right-hand image shows the correct correspondences extracted by voting algorithm.

5. Outlier removal by 3D geometric constraints

Unlike the color and undistorted images used in other studies, gray and distorted reflectance images do not contain much information. Therefore, a number of false pairs of matching features (POMFs) between two reflectance images will be extracted. In some cases, the number of MFs is larger than the number of true POMFs. As shown in Fig.5, there are only four true pairs of matching features between $Test\_17.ref$ and $Train\_22.ref$, and their actual locations are shown in Fig.8. Since each feature on the 2D reflectance images has a unique 3D position on the 3D local map, geometric constraints, such as the distance and the normal vector of a surface can be used to extract true POMFs. We propose a voting algorithm to maintain the true POMFs and remove the false POMFs using 3D geometric information. This process corresponds to Step 1-c) in Section 4-B.

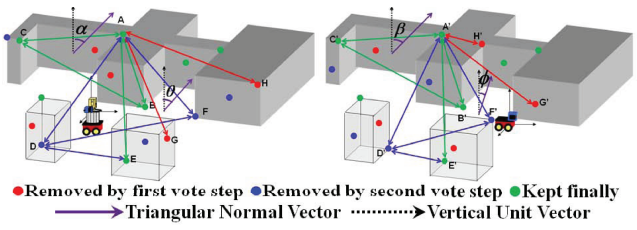


Fig. 6. 3D geometric constraint

The first step of the voting algorithm is outlier removal by comparing the 3D distance between two POMFs (see Algorithm 1). Here, D_{tr} is the 3D distance between Tr_m and Tr_n, and D_{te} is the 3D distance between Te_m and Te_n. If the error between D_{tr} and D_{te} is less than D_{thresh}, then the scores of two of the POMFs are increased by 1. If the final score of a POMF is less than $\eta \times factor1$, then this POMF is removed as an outlier. In Fig.6, $|AG| \neq |A'G'|$, $|AH| \neq |A'H'|$, and the red points are removed during this step.

The second step is outlier removal by comparing the edges and normal vectors of triangles in 3D space. This is based on a self-evident theorem. Namely, given any three points in 3D space, independent of the $(0,0,0)^T$ of a local 3D map, the lengths of the three edges are constant. In addition, the angle between a normal vector and an unit vector vertical to the ground $(0,0,1)^T$ is also constant. In Fig.6, α, β, θ, and ϕ are the angles between the normal vector of triangles and the unit vector $(0,0,1)^T$. Here, $\triangle ADE$, $\triangle A'D'E'$, $\triangle ADF$, and $\triangle A'D'F'$ have AA' and DD' as common POMFs. We also have $|AD| \neq |A'D'|$, $|DE| \neq |D'E'|$, and $|\theta - \phi| > ANG_{thresh}$. In addition, $\triangle ABC \cong \triangle A'B'C'$, and $\alpha = \beta$. If the final scores of POMFs $DD' and FF'$ are smaller than $\omega \times factor2$, then these POMFs are removed. Thus, POMFs AA', BB', CC' and EE' remain. The blue points are removed during this step. Then, the remaining green points are the inputs of RANSAC.

Using Algorithm 1, the outliers of POMFs between reflectance images can be removed effectively. A number of false candidate features will be excluded by this voting algorithm. In Fig.5, 4 true corresponding pairs are extracted from 32 pairs of corresponding features successfully. Since a small number of reliable POMFs remain, this can reduce the retrieval time. The effectiveness of the proposed method is demonstrated in Section 6.

6. Experiment

Two experiments are conducted in order to verify the performance of the proposed two-step strategy. The paths for the experiments are shown in Fig.7. The path for the first experiment is shown in Fig.8, where a number of buildings exist along the path. The other experimental path, which is along a long and slope road, is shown in Fig.9. Only a few trees and poles exist along the roadway. Example reflectance images for the second experiment are shown in Fig.10. In the first experiment, a number of buildings that are good references are near the path of the robot. On the other hand, in the second experiment, the buildings are relatively far from the robot, and the reflectance images are made up of primarily the ground, trees, and poles. Intuitively, trees and poles are not good references, since they are similar and the leaves and branches of trees move due to the wind.

6.1. Experiment in the area containing buildings

The size of reflectance images is 590×569 pixels. The CPS-V robots move and stop at 58 different locations in

Algorithm 1 Voting algorithm

Assume η POMFs are extracted from $Train\_i.ref$ and $Test\_j.ref$
for $0 \leq m < \eta$ **do**
  Get two 3D points (Tr_m and Te_m) from the m^{th} feature.
  for $m+1 \leq n < \eta$ **do**
    Get another two 3D points (Tr_n and Te_n) from the n^{th} feature

$$if\,(|D_{tr} - D_{te}| < D_{thresh}) \quad (1)$$

$$\{score1[m]++;$$
$$score1[n]++;\}$$

  end for
end for
for $0 \leq m < \eta$ **do**

$$if\,(Score1\,[m] < \eta \times factor1) \quad (0 < factor1 < 1) \quad (2)$$

Delete the m^{th} pair
end for
Assume ω POMFs remain after above vote step
for $0 \leq iteration < \omega \times N$ **do**
  Randomly select three POMFs $\omega \times N$ (N is a pre-defined iteration number) times to form two triangles in 3D space ($Tr\Delta$ and $Te\Delta$).

$$if\left(\left|Tr\Delta_{edg} - Te\Delta_{edg}\right| < D\Delta_{thresh}\right) \quad \ldots \quad (3)$$
$$if\,(|Tr_{DEG} - Te_{DEG}| < ANG_{thresh}) \quad \ldots \quad (4)$$

$$\{score2[m]++;$$
$$score2[n]++;$$
$$score2[k]++;\}$$

end for
for $0 \leq m < \omega$ **do**

$$if\,\{score2\,[m] < \omega \times factor2\} \quad (0 < factor2 < 1) \quad (5)$$

Delete the m^{th} pair
end for

the experimental area indicated by the red path in Fig.8, i.e., 58 data ($Train\_i.ref$ and $Train\_i.pts$) are stored as the training dataset. On the other hand, test data ($Test\_j.ref$ and $Test\_j.pts$) are collected at 29 locations as indicated by the three colored paths in Fig.8. The three colors indicate data collected at different times from different starting points, so that the global coordinates of the data are different (see [5]).

6.1.1. Results of position estimation after the first step (coarse estimation)

The results of location estimation after the first step are listed in Table 2. In this experiment, M in Section 4.1 is set to be 5, i.e., five candidates for each $Test\_j.ref$ are retrieved from the training dataset by the Kd-tree. The experimental results show that all of the positions of the robot are included in selected five candidate locations. Especially in 25 positions, the positions of the robot are correctly estimated as the first candidate. In two positions, the second candidates are the actual locations. The remaining two positions are also correctly estimated as the third and fifth candidates, respectively.

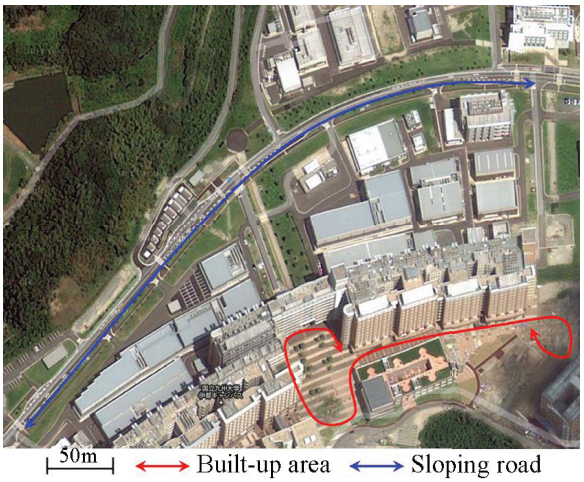


Fig. 7. Experimental areas

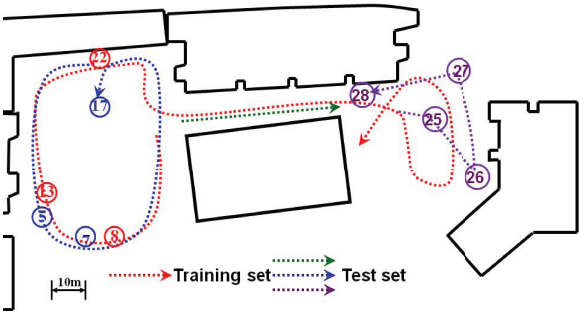


Fig. 8. Experimental area containing buildings

Table 2. Correctness of position estimation after the first step in the area containing buildings

No.	1st	2nd	3rd	4th	5th	Total	Correctness ratio
Correct localization	25	2	1	0	1	29	100%

6.1.2. Results of position estimation after the second step (precise estimation)

Initially, $D_{threshold}$ in Algorithm 1-(1) is set to be 3 [m] according to the attributes of LMS151 shown in Table 1, $factor1$ in (2) is $1/3$, N is $\omega/2$, $D\Delta_{thresh}$ of (4) is 1 [m], and factor2 of (5) is $1/3$. The threshold of the average error for terminating ICP is set to be 0.02 [m] after 40

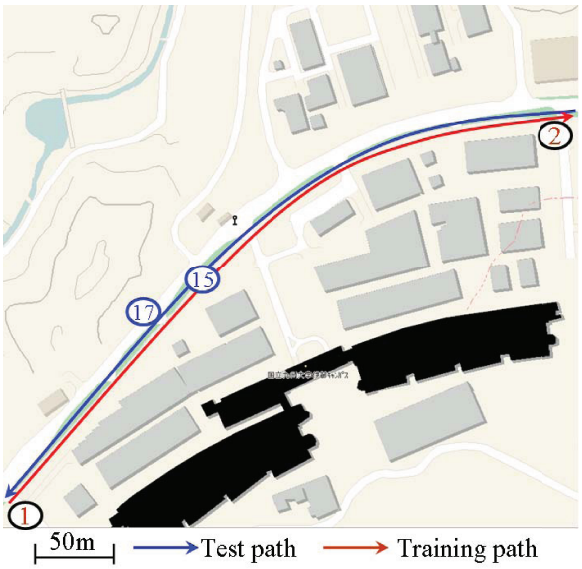


Fig. 9. Experimental area along a road. Here (1) indicates the start of the training path and the end of the test path, and (2) indicates the start of the test path and the end of the training path.

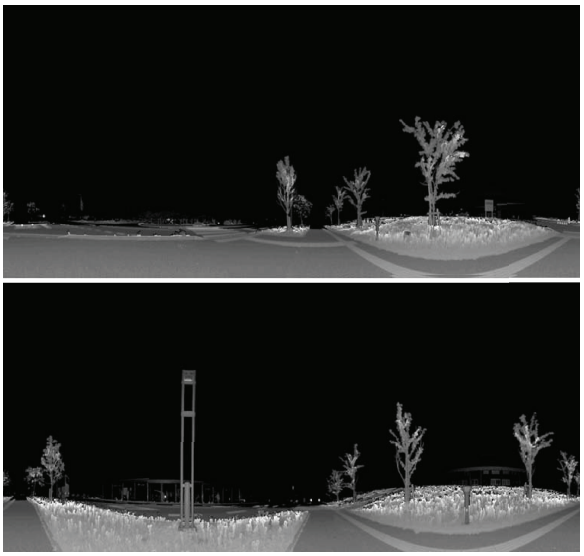


Fig. 10. Reflectance images collected along the road. Top: $Test\_15.ref$, and bottom: $Test\_17.ref$

iterations of ICP. Table 3 shows the results of the precise position estimation. Five positions are not estimated correctly due to the failure in the voting algorithm, i.e., the number of true 3D POMFs between $Train\_i.pts$ and $Test\_j.pts$ is quite small. One false localization is excluded by the large average error of ICP. The total number of true positives is 23, so that the recall is $23/29 \approx 79.3\%$, and no false positives occur. Figure 11 shows the alignment results between $Test\_5.pts$ and $Train\_13.pts$ and between $Test\_7.pts$ and $Train\_8.pts$ after RANSAC and automatic ICP. Table 4 shows the average errors of their results. These pairs are correctly selected in Step 1, and

Table 3. Correctness of position estimation after the second step in the area containing buildings

Total	29
Excluded by voting algorithm	5
Excluded by large ICP error	1
True positive	23

their relative locations are shown in Fig.8.

In the proposed method, the most time-consuming part is ICP. The incorrect candidates can be excluded by the voting algorithm. Since M is set to be 5 in the experiment, the ICP should be executed $29 \times 5 = 145$ times if the voting algorithm is not applied. On the other hand, using the voting algorithm, ICP is executed only 37 times, i.e., $37/29 \approx 1.28$ times for each test data.

Table 4. Average error of the two examples in Fig.11

Example	Coa_alig	Pre_alig
$Train\_13.pts$ and $Test\_5.pts$	27.68 [mm]	11.19 [mm]
$Train\_8.pts$ and $Test\_7.pts$	5.85 [mm]	4.88 [mm]

6.2. Experiment along the road

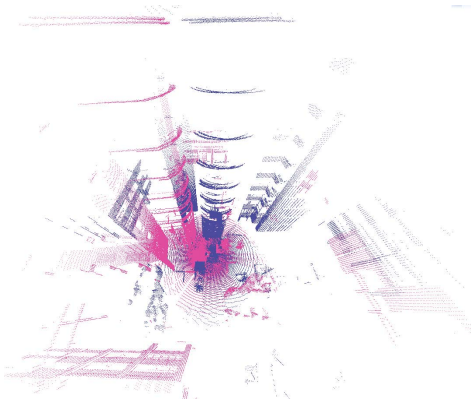
Since many buildings alongside the road are almost out of the range of the LMS151, to make them more prominent in the reflectance images as soon as possible, reflectance images with dimensions of 530×1134 are collected. A total of 33 training data ($Train\_i.ref$ and $Train\_i.pts$) and 28 test data ($Test\_j.ref$ and $Test\_j.pts$) are stored. Since the reflectance images are panoramic, the training data and test data are collected in the converse direction along the same path as shown in Fig.9.

6.2.1. Results of position estimation after the first step (coarse estimation)

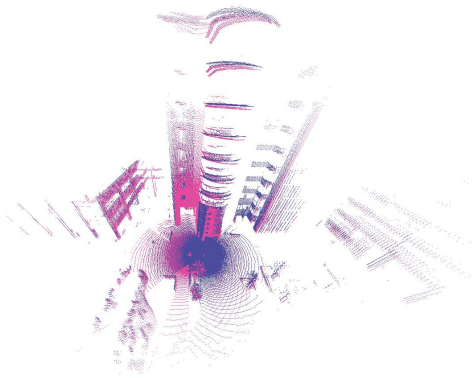
The results of location estimation after the first step are listed in Table 2. In Section 4.1, M is set to be 5 as before. The results are shown in Table 5. A total of 27 positions of the robot are correctly estimated. In 16 positions, the robot position is correctly estimated as the first candidate. In eight positions, the second candidates are the actual locations. Three positions are correctly estimated as the third candidates. Only one position is falsely estimated, so that the correctness ratio is $27/28 \approx 96.4\%$.

6.2.2. Results of position estimation after the second step (precise estimation)

In this experiment, since the robots move along sloping ground and several rigid references, such as buildings, are located far from the robots, a number of the parameters in Algorithm 1 are set to be relatively flexible. For example, $factor1$ in (2) is 0.1, and ANG_{thresh} of Eq. (4) is set

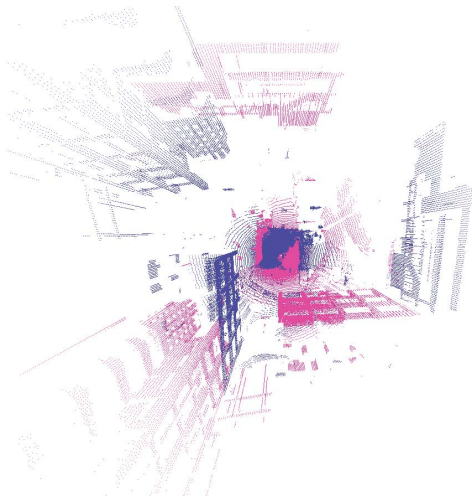


(a) Original maps of *Train_13.pts* and *Test_5.pts*

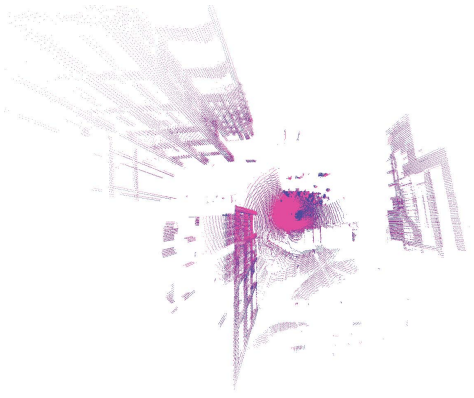


(b) Coarse alignment by RANSAC between *Train_13.pts* and *Test_5.pts*

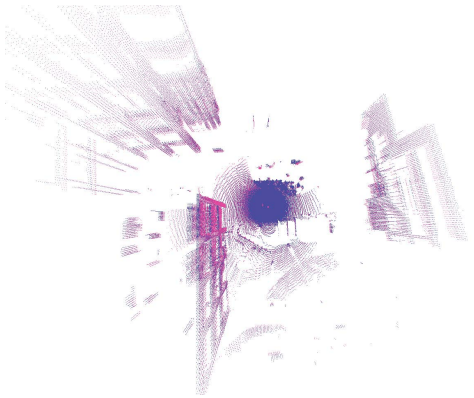
(c) Precise alignment by ICP between *Train_13.pts* and *Test_5.pts*



(d) Original maps of *Train_8.pts* and *Test_7.pts*



(e) Coarse alignment by RANSAC between *Train_8.pts* and *Test_7.pts*

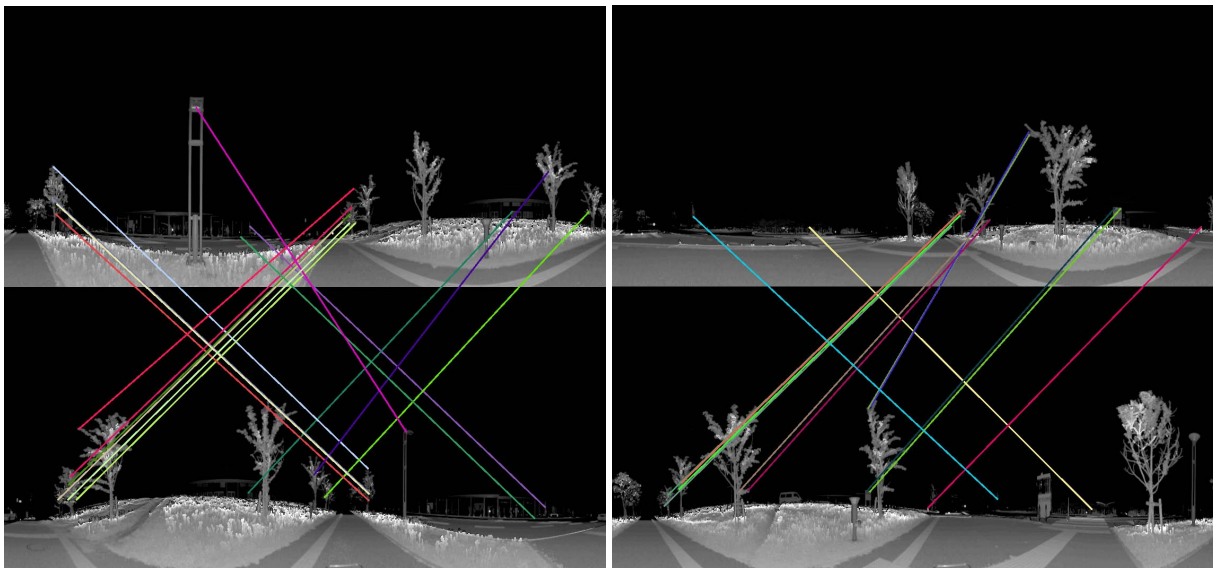


(f) Precise alignment by ICP between *Train_8.pts* and *Test_7.pts*

Fig. 11. Alignment results between the 5^{th} test data and the 13^{th} training data and between the 7^{th} test data and the 8^{th} training data

to be 20 degrees. In order to avoid the influence of road, which has a relatively homogeneous texture, we do not consider triangles that have a normal vector that is vertical to the ground. Table 6 shows the results of the precise position estimation. Nine positions are not estimated correctly due to a failure in the voting algorithm. The total number of true positives is 19, so that $19/28 \approx 67.9\%$ recall is achieved, and no false positives occur. Figure 12 shows the inliers from RANSAC in a 2D manner for the two examples in Fig.10. Several correct POMFs are obtained from branches and leaves of trees and lampposts. The ICP is executed only 21 times, i.e., $21/28 \approx 0.75$ times for each test datum.

The proposed method is robust for the presence of moving objects such as pedestrians or cars. Fig.13 shows the correct POMFs on 2D reflectance images including sev-



(a) Inliers of RANSAC between $Train\_15.ref$ and $Test\_15.ref$ (b) Inliers of RANSAC between $Train\_13.ref$ and $Test\_17.ref$

Fig. 12. Matching features between the 15^{th} test data and 15^{th} training data and between 17^{th} test data and 13^{th} training data of the experiment along the sloping road

Table 5. Correctness of position estimation after the first step along the road

No.	1st	2nd	3rd	4th	5th	Total	Correctness ratio
Correct localization	16	8	3	0	0	27	96.4%

Table 6. Correctness of position estimation after the second step along the road

Total	28
Excluded by voting algorithm	9
Excluded by large ICP error	0
True positive	19

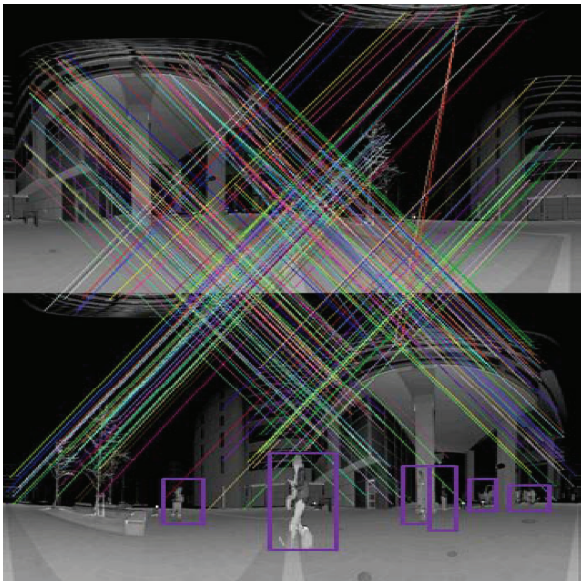


Fig. 13. The rectangles in the training image show the moving objects which are not in the test image.

eral pedestrians and a car in the bottom image. Fig.14 shows different cars parked in almost the same place but don't make any difference. It is clear that the proposed method using BoF and outlier removal hardly be influenced by these disturbances.

6.3. Parameters setting

In fact, four non-coplanar 3D true POMFs are enough for the calculation of a rigid transformation. In algorithm 1, parameters D_{thresh}, $D\Delta_{thresh}$, and ANG_{thresh} are set mainly depending on the attribute of LMS151 (see Table 1). As an extreme example, if only the right candidate correspondences in the results of BoF are processed by automatic ICP, $factor1$ and $factor2$ are set to be small to keep true POMFs. Consequently, all of right candidate correspondences can be correctly aligned by automatic ICP. On the other hand, small $factor1$ and $factor2$ will cause false initial candidates for automatic ICP. This would not only be costly(since ICP must be repeated many times), but also lead to false final results. If $factor1$ and $factor2$ are set relatively to be high, the number of false-positives can be decreased and ICP would be repeated few times. However, the final success ratio would be low since the chance for obtaining correct POMFs becomes low.

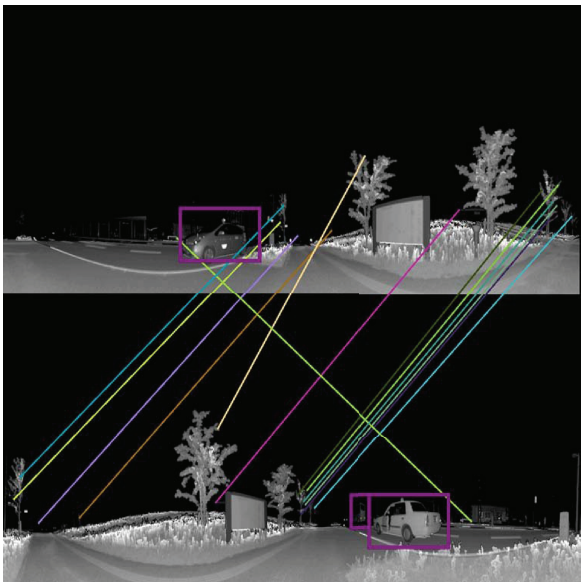


Fig. 14. The rectangles show different cars parked in same place which don't influence the results.

7. Conclusion

We proposed and demonstrated a two-step strategy for global localization of a mobile robot. Appearance-based global localization and map-based global localization are combined to improve the performance of correct position estimation. The reflectance image, which is obtained as a byproduct of range sensing and is independent of the variation of the illumination condition, is used for appearance-based global localization in the first step. Precise map-based global localization by ICP is applied using 3D local maps, which are selected by the first step. In order to improve the performance of the 2nd step, the voting algorithm based on the 3D geometric constraints and RANSAC-based course position estimation process are proposed. The effectiveness of the proposed technique is demonstrated through experiments in real environments.

Acknowledgements

The present study was supported in part by a Grant-in-Aid for Scientific Research (B) (23360115).

References:

[1] D. Filliat, J.-A. Meyer, "Map-based navigation in mobile robots: I. A review of localization strategies", *Cognitive Systems Research*, vol. 4, no. 4, pp. 243-282, 2003.

[2] A. Angeli, D. Filliat, S. Doncieux, and J.A. Meyer, "A Fast and Incremental Method for Loop-Closure Detection Using Bags of Visual Words", *IEEE Transactions on Robotics*, vol. 24, no.5, pp. 1027-1037, 2008.

[3] M. Cummins, P. Newman, "FAB-MAP: Probabilistic Localization and Mapping in the Space of Appearance", *International Journal of Robotics Research*, vol. 27, no. 6, pp. 647-665, 2008.

[4] Y. Hara, H. Kawata, A. Ohya, S. Yuta, "Mobile Robot Localization and Mapping by Scan Matching using Laser Reflection Intensity of the SOKUIKI Sensor", *IEEE Industrial Electronics, IECON*, art. no. 4153430, pp. 3018-3023, 2006

[5] R. Kurazume, S. Hirose, "An Experimental Study of a Cooperative Positioning System", *Journal of Autonomous Robots*, vol. 8, no. 1, pp. 43-52, 2000.

[6] M. Guarnieri, R. Kurazume, H. Masuda, "HELIOS system: A team of tracked robots for special urban search and rescue operations", *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS*, art. no. 5354452, pp. 2795-2800, 2009.

[7] H. Bay, A. Ess, T. Tuytelaars, "Speeded-Up Robust Features (SURF)", *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346-359, 2008.

[8] D. Lowe, "Distinctive image feature from scale-invariant keypoint", *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91-110, 2004.

[9] P. J. Besl, N. D. McKay, "A method for registration of 3-D shapes", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 239-256, 1992.

[10] C. S. Chen, Y. P. Hung, "RANSAC-Based DARCES: A new approach to fast automatic registration of partially overlapping range images" , *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 11, pp. 1229-1234, 1999.

[11] F. Bonin-Font, A. Ortiz and G. Oliver, "Visual Navigation for Mobile Robots: A Survey", *Journal of Intelligent and Robotic Systems: Theory and Applications*, vol. 53, no. 3, pp.263-296, 2008.

[12] S. Thrun, D. Fox, W. Bugard and F. Dellaert, "Robust Monte Carlo localization for mobile robots", *Artificial Intelligence*, vol. 128, pp. 99-141, 2001.

[13] S. Thrun, "Probabilistic algorithms in robotics", *AI Magazine*, vol. 21, no. 4, pp. 93-109, 2000.

[14] Y. Matsumoto, M. Inaba, H. Inoue, "Visual navigation using view-sequenced route representation", *Proceedings - IEEE International Conference on Robotics and Automation*, vol. 1, pp. 83-88, 1996.

[15] S. D. Jones, C. Andresen and J. L. Crowley, "Appearance based processes for visual navigation", *IEEE International Conference on Intelligent Robots and Systems*, vol. 2, pp. 551-557, 1997.

[16] T. Ohno, A. Ohya, S. Yuta, "Autonomous navigation for mobile robots referring pre-recorded image sequence", *IEEE International Conference on Intelligent Robots and Systems*, vol. 2, pp. 672-679, 1996.

[17] A.C. Murillo, C. Sagüés, J.J. Guerrero, T. Goedemé, T. Tuytelaars, L. Van Gool, "From omnidirectional images to hierarchical localization", *Robotics and Autonomous Systems*, vol. 55, no. 5, pp. 372-382, 2007.

[18] E. Menegatti, M. Zoccarato, E. Pagello, H. Ishiguro, "Image-based Monte Carlo localisation with omnidirectional images", *Robotics and Autonomous Systems*, vol. 48, no. 1, pp. 17-30, 2004.

[19] J. Wolf, W. Burgard, H. Burkhardt, "Robust vision-based localization for mobile robots using an image retrieval system based on invariant features", *IEEE International Conference on Robotics and Automation*, vol. 1, pp. 359-365, 2002.

[20] H. Andreasson, A. Treptow, T. Duckett, "Self-localization in non-stationary environments using omni-directional vision", *Robotics and Autonomous Systems*, vol. 55, no. 7, pp. 541-551, 2007.

[21] A.C. Murillo, J.J. Guerrero, C. Sagüés, "SURF features for efficient robot localization with omnidirectional images", *IEEE International Conference on Robotics and Automation*, art. no. 4209695, pp. 3901-3907, 2007.

Name:
Your Name

Affiliation:
Your Institute

Address:
Address of Your Institute

Brief Biographical History:
Your History

Main Works:
• Your Works

Membership in Learned Societies:
• Your Learned Societies