Appearance and map-based global localization using laser reflectivity

DongXiang Zhang, Ryo Kurazume, Yumi Iwashita, Tsutomu Hasegawa

Abstract-Global localization is a fundamental ability to recognize the accurate global position for a mobile robot in a revisited environment. The map-based global localization gives a precise position by calculating an accurate transformation, but the comparison with large 3D data is quite timeconsuming. The appearance-based global localization which determines the global position by image retrieval techniques with similar structures is real-time. However, this technique needs external illumination constraint and does not work in the dark extremely. This paper proposes a combination of the map-based global localization and the appearance-based global localization. Instead of camera images used for the appearancebased global localization, we utilize reflectance images which are taken as a byproduct of range sensing by a laser range finder. The proposed method not only detects previously visited scenes but also estimate relative poses precisely. The effectiveness of the proposed technique is demonstrated through experiments in real environments.

I. INTRODUCTION

In some practical robotic tasks, the perceptual information from external environment is unpredictable, unstructured and uncontrolled. For instance, in an area struck by a strong earthquake or by a mine disaster, the geometrical structure differs from the original one because of grounds blocked by a heap of rubble or collapsed walls. To accomplish a search and rescue task efficiently in such an uncontrolled and unpredictable environment, accurate mapping and localization are fundamental capabilities.

Global localization, is a definition to accurately localize a robot's position in a global coordinate system using surrounding features only given a map without any other prior knowledge. Plenty of global localization techniques have been proposed so far [11]. Abundant works have carried out to attack this problem based on camera images captured in constrained illumination environment, i.e. appearance-based localization. other works utilizing laser range finder, i.e. mapbased localization mainly focused on 2D range data and can't estimate accurate relative poses.

The map-based global localization[1] is defined to find a best position where the observed geometrical features match the ones in the provided geometrical map. Though 3D range data is preferable for 6D global localization in terms of accuracy and reliability, the comparison between 3D range data captured by a range sensor and pre-constructed 3D map is quite time-consuming. On the other hand, the appearance-based global localization using camera images[14],[15],[16]

is simple and suitable for a real-time processing. However, the appearance-based global positioning can't work in dark or an environment where lighting condition changes severely since illumination constraints is necessary for robust appearance-based localization.

This paper proposes a two-steps strategy which combines the map-based global localization and the appearance-based global localization. By using reflectance image which are taken as a byproduct of range sensing by a laser range finder instead, the proposed technique is useful even in the dark or an environment under severe lighting condition thanks to the characteristic of the reflectance image which is not subject to any severe variants of external illumination conditions. Furthermore, fast and precise localization can be done by comparing a few 3D range images which are selected based on the similarity of the reflectance images. The proposed two-steps strategy is as follows: i) reflectance images retrieval system for rough estimation of a global position based on Bag-of-feature technique, and ii) precise global position is determined by Iterative Closest Points algorithm for 3D range data automatically.

The remainder of this paper is organized as follows: after brief introduction of related works in section II, section III simply introduces cooperative positioning system which the proposed method is based on. The proposed two-steps strategy is presented in detail in section IV and V, and the experimental results are shown in section VI.

II. RELATED WORK

Image-based visual SLAM (simultaneous localization and mapping) has been reported in several literatures such as [2] and [3]. The Bag-of-Features (BoF) is a popular technique for efficient representation of a raw image captured by a camera. Non-false-positive is achieved in [2] by combining BoF and the probabilistic calculation. However, [2] requires a constant illumination condition. Therefore the travel distance of the robot has to be short enough so that the lighting condition will not change drastically. In [3], the authors didn't evaluate the experimental results explicitly under large changes of illumination. Extremely, both of them would fail in the dark or the environment where the lighting is severely changed.

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 we measure range 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

Ryo Kurazume, Yumi Iwashita and Tsutomu Hasegawa are with Faculties of the Graduate School of Information Science and Electrical Engineering, Kyushu University, Japan

DongXiang Zhang is with Phd candidate of the Graduate School of Information Science and Electrical Engineering, Kyushu University, Japan, zhang@irvs.ait.kyushu-u.ac.jp

values. In other words, the range image and the reflectance image are precisely and fundamentally aligned. In addition, the reflectance image is not subject to any severe variants of external illumination conditions and thus we can obtain stable reflectance images even at night.

The proposed technique utilizes the reflectance image instead of a regular camera image. By applying BoF for a reflectance image which corresponds to 3D range data, the global localization using 3D range data and 2D images is achieved efficiently. [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 proposed technique is deterministic, so it doesn't rely on Markov assumption, in which the future state of the robot, given the present and the past states, depends only on the present state[13]. It also doesn't encounter "kidnapped robot problem"[12], because its current state purely depends on current sensor reading.

III. 3D MAP CREATION BASED ON COOPERATIVE POSITIONING SYSTEM

For the map-based global localization, an environmental map has to be created and provided beforehand. For precise 3D mapping of an environment around a robot, we have been proposing an efficient and precise system named CPS-SLAM [5], which can construct a quite accurate large-scale 3D map by a laser range finder and multiple robots based on the technology for geographical surveying. This technique has been used for the basis of Urban Search and Rescue robot(USAR)[6].



Fig. 1. The 3D modeling robots, CPS-V

TABLE I					
SICK LMS151					

Measuring range	50[m]
Field of View	270[0]
Precision	$\pm 30[mm]$
Angular resolution	0.25[0]

Figure 1 shows the fifth CPS-SLAM model named 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 3-axes attitude sensor. On the other hand, two child robots are equipped with corner cubes. The GPT-9000A and corner cubes are used for selfpositioning cooperatively as shown in Fig.2. The LMS151 (Table I) 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. Thus by rotating the table around the vertical axis for 360° while scanning with the 2D laser range finder, 3D range data and a 2D reflectance image are acquired. The number of pixels on a reflectance image is exactly same as the number of 3D points in range data, i.e. there is one-to-one mapping relationship between 2D pixels of the reflectance image and 3D points of the local 3D map.



(3) Robot 0 measures the relative position of robot 2

(4) Robot 0 moves and determines its own position by observing robots 1 and 2

Fig. 2. Cooperative Positioning System, CPS



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

A. 3D 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 parent and child robots. In the end, all



Fig. 4. Reflectance and range images

of local 3D maps are aligned into a global 3D map using measurement position information. More details about CPS-based simultaneous localization and mapping (CPS-SLAM) can be found in [5].

B. Reflectance images

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

C. Image retrieval using Bag-of-features

When the global 3D map of the target field is constructed, a Kd-tree structure storing Bag-of-features(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. Firstly, some 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, the words are stored using Kd-tree structure. With these words as x-axis, quantize each feature in a reflectance image to its approximate nearest word by searching in Kdtree, all of recorded reflectance images are represented as statistics of words(histograms). Finally, the histograms of all recorded images are stored also using Kd-tree structure. A newly-captured image is also represented as a histogram and the best *M* matching images for it are retrieved by quantizing the histogram to its nearest M histograms.

IV. TWO-STEPS STRATEGY COMBINING APPEARANCE-BASED AND MAP-BASED GLOBAL LOCALIZATION

This section presents the two-steps strategy for precise localization using a 3D map. Firstly, we need to create a global map as a training dataset. As explained in Section III, 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 bag-of-features (BoF) technique and the training dataset is created. Finally, the dataset of all the BoF representations is stored in Kd-tree which is efficient for information retrieval.

For global localization, a new robot which is equipped with a 3D range sensor like CPS-V, is placed and collects local 3D data and 2D reflectance images (test data). In the 1st step, we retrieve some candidates of initial location by comparing stored reflectance images (training dataset) and captured reflectance images (test data) using BoF and Kdtree. Then we apply 3D geometrical constraint to extract true feature pairs and run automatic ICP in the 2nd step. Thanks to the 1st step, fast and precise localization can be done in the 2nd step by comparing a few 3D range images which are selected based on the similarity of the reflectance images in the 1st step.

Hereinafter, we denote variables related with training data as " Tr_m " or " D_{tr} ", and variables related with test data as " Te_n " or " D_{te} ". D means the 3D distance between two points in a local 3D map. *Train_i.ref* and *Train_i.pts* represent the i^{th} reflectance image and local 3D map in training dataset. *Test_j.ref* and *Test_j.pts* represent the j^{th} reflectance image and local 3D map in test dataset.

We will explain the proposed two-steps strategy in more detail below. The entire process is shown in Fig.5



Fig. 5. Combination of appearance-based localization and map-based localization.

A. 1st step: initial localization by Bag-of-features using 2D reflectance images

All $Test_j.ref$ are converted into bag-of-features (BoF) representations, and searched the best M matches in Kd-tree previously constructed from the training dataset. Then the M $Train_i.ref$ and $Train_i.pts$ are selected as M candidates of the position of the robot in the 3D global map.

B. 2nd step: precise localization by Automatic ICP using 3D data

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

- 1) Rough alignment with RANSAC
 - a) Find corresponding features between *Test_j.ref* and *Train_i.ref*.
 - b) Get the 3D coordinates of corresponding features by *Test_j.pts* and *Train_i.pts* which correspond to *Test_j.ref* and *Train_i.ref* respectively.
 - c) Remove outliers by 3D geometrical constraints. This process will be explained in Section V.
 - 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 aligned roughly.

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 "alignment ratio" and another is "average error". Firstly, 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 and N of them can be found corresponding points in $Train_i.pts$. The "alignment ratio" is defined by N/K. In those N pairs of points, the sum of errors between every pair of points divided by N is defined as "average error".

V. OUTLIER REMOVAL BY 3D GEOMETRICAL CONSTRAINTS

Unlike color and undistorted images used in other works, gray and distorted reflectance images don't contain much information. Therefore many false POMFs (Pair Of Matching Features) between two reflectance images will be extracted. In some cases, the number of false POMFs is larger than true POMFs. As shown in Fig.6, there are only 4 true pairs of matching features between *Test_17.ref* and *Train_22.ref*, their actual locations are shown in Fig.8. Since all the features on the 2D reflectance images have their own 3D positions in the 3D local map, geometric constraints such as distance and a normal vector of a surface can be used for extracting true POMFs. We propose a voting algorithm to keep the true POMFs and remove the false POMFs by using 3D geometrical information. This process corresponds to the step 1-c) in Section IV-B.

The 1st step of the voting algorithm is the outliers removal by comparing the 3D distance between two POMFs (see Algorithm 1). 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 below D_{thresh} , the scores of two POMFs are increased by 1. In the end if the score of POMF is below $\eta \times factor1$, this POMF is removed as an



32 pairs of corresponding 4 true corresponding pairs features in total

Fig. 6. Corresponding features between *Test_*17*.ref* and *Train_*22*.ref*. Left image shows all correspondences, Right image shows the correct correspondences.



Fig. 7. 3D geometric constraint

outlier. In Fig.7, $|AG| \neq |A'G'|$, $|AH| \neq |A'H'|$, and the red points are removed by this step.

The 2nd step is the outliers removal by comparing edges and normal vectors of triangles in 3D space. This is based on a self-evident theorem: Given any three points in 3D space, no matter where the $(0,0,0)^T$ of a local 3D map is, the length of three edges are constant. In addition, the angle between a normal vector and a unit vector vertical to the ground $(0,0,1)^T$ is also constant. In Fig.7, α,β,θ , and ϕ are the angles between the normal vector of triangles and the unit vector $(0,0,1)^T$. $\triangle ADE$, $\triangle A'D'E'$, $\triangle ADF$, and $\triangle A'D'F'$ have common POMFs AA', DD'. $|AD| \neq |A'D'|$, $|DE| \neq |D'E'|$ and $|\theta - \phi| > ANG_{thresh}$; Another pair of triangles $\triangle ABC \cong \triangle A'B'C'$, $\alpha = \beta$. In the end, if the scores of POMFs DD', FF' are smaller than $\omega \times factor2$, they are removed. POMFs AA', BB', CC', EE' are kept. The blue points are removed by this step. The green points are finally kept as the input of RANSAC.

With the Algorithm 1, the outliers of POMFs between reflectance images can be removed effectively. Many false candidate features will be excluded by this voting algorithm. Since a small number of reliable POMFs are remained, this

Algorithm 1 Voting algorithm

Suppose η POMFs extracted from *Train_i.ref* and *Test_j.ref* for $0 \le m < \eta$ do Get two 3D points (*Tr_m* and *Te_m*) from *m*th feature.

for $m+1 \le n < \eta$ do

Get another two 3D points $(Tr_n \text{ and } Te_n)$ from n^{th} feature

$$if (|D_{tr} - D_{te}| < D_{thresh})$$

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

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

$$(1)$$

end for end for for $0 \le m < \eta$ do

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

Delete the m^{th} pair

end for

Suppose ω POMFs remained from above vote step

for $0 \leq iteration < \boldsymbol{\omega} \times N$ do

Randomly select 3 POMFs for $\omega \times N$ (N is a pre-defined iteration number) times to form two triangles in 3D space $(Tr\Delta \text{ and } Te\Delta)$.

$$if\left(\left|Tr\Delta_{edg} - Te\Delta_{edg}\right| < D\Delta_{thresh}\right) \tag{3}$$

$$if(|Tr_{DEG} - Te_{DEG}| < ANG_{thresh})$$

$$\tag{4}$$

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

end for for $0 \le m < \omega$ do

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



Training set

Fig. 8. Experimental field

that is, 5 candidates for each *Test_i.ref* are retrieved from the training dataset by Kd-tree. All the positions of the robot are correctly estimated and included in 5 candidate locations. In 25 positions, the positions of the robot are correctly estimated as the first candidate. In 2 positions, the second candidates are the actual locations. The remaining 2 positions are also correctly estimated as third and fifth candidates, respectively.

TABLE II

CORRECTNESS OF POSITION ESTIMATION AFTER 1ST STEP

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

TABLE III

CORRECTNESS OF POSITION ESTIMATION AFTER 2ND STEP

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

B. Results of position estimation after 2nd step (precise estimation)

At first, D_{thresh} in Algorithm 1-(1) is set to be 3 [m] according to the attribute of LMS151 shown in Table I, 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 iterations of ICP. Table III shows the results of the precise position estimation. 5 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 large average error of ICP. True positive is 23 in total, so $23/29 \approx 79.3\%$ recall is achieved and no-false-positive is obtained. Figure 10 shows the alignment results between Test_3.pts and Train_16.pts and results between Test_21.pts and Train_36.pts after RANSAC and automatic ICP. Table IV shows the average errors of their results. These pairs are correctly selected in Step 1. Their relative locations are displayed in Fig.8.

can save retrieval time. The effectiveness of our method will be shown in section VI.

VI. EXPERIMENT

An experiment is conducted to verify the proposed twosteps strategy. The experimental field is shown in Fig.8. The size of every reflectance image is 590×569 . The CPS-V robots move and stop at different locations in experimental fields for 58 times as the red path shown in Fig.8, i.e. 58 data (*Train_i.ref* and *Train_i.pts*) are stored as training dataset. On the other hand, test data (*Test_j.ref* and *Test_j.pts*) are collected in 29 locations as the three colored paths shown in Fig.8.

A. Results of position estimation after 1st step (coarse estimation)

The results of location estimation after 1st step are listed in Table II. In this experiment, M in Section IV-A is set to be 5,

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 for $29 \times 5 = 145$ times if the voting algorithm is not applied. On the other hand, with the voting algorithm, ICP is executed for only 37 times, that is, $37/29 \approx 1.28$ times for each test data.

The proposed method is robust for the presence of moving objects such as pedestrians or cars. Fig.9 shows the correct POMFs on 2D reflectance images including several pedestrians and a car in the bottom image. It is clear that the proposed method using BoF and outlier removal hardly be influenced by these disturbances.



Fig. 9. The rectangles in the training image (lower) show the moving objects which are not in the test image (upper).

TABLE IV Average error of the two examples in Fig.10

Example	Coarse alignment	Precise alignment
Train_16.pts and Test_3.pts	60.17[<i>mm</i>]	30.86[<i>mm</i>]
Train_36.pts and Test_21.pts	43.34[<i>mm</i>]	25.03[<i>mm</i>]

VII. CONCLUSION

This paper proposes and demonstrates the two-steps strategy for global localization of a mobile robot. The appearance-based global localization and the map-based global localization are combined for improving the performance of correct position estimation. The reflectance image which is taken as a byproduct of range sensing and invariant to the change of illumination condition is utilized for the appearance-based global localization in the 1st step. Then the precise map-based global localization by ICP is applied using 3D local maps which are selected by the 1st step. To improve the performance of the 2nd step, the voting algorithm based on the 3D geometrical constraints and RANSAC-based course position estimation process are proposed. The effectiveness of the proposed technique is demonstrated through experiments in real environments.

VIII. ACKNOWLEDGMENTS

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

REFERENCES

- 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] H. Yoshitaka, K. Hirohiko, O. Akihisa, "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 viewsequenced 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.



(b) Coarse alignment by RANSAC between *Train_16.pts* and *Test_3.pts*



(c) Precise alignment by ICP between *Train_16.pts* (d) Original maps of *Train_36.pts* and *Test_21.pts* and *Test_3.pts*



(e) Coarse alignment by RANSAC between *Train_36.pts* (f) Precise alignment by ICP between *Train_36.pts* and and *Test_21.pts*

Fig. 10. Alignment Results between the 16th test data and 3th training data, 21th test data and 36th training data

(a) Original maps of Train_16.pts and Test_3.pts