Evaluation and Improvement of a Relocalization Method for Camera Tracking in Nuclear Power Plants

Kimura Taro[1], Tokumaru Hiroki[2], Ishii Hirotake[3], Shimoda Hiroshi[4] and Kouda Yuya[5]

1. Graduate School of Energy Science, Kyoto University, Yoshida Honmachi, Sakyo-ku, Kyoto 606-8501, Japan (kimura@ei.energy.kyoto-u.ac.jp)
2. Graduate School of Energy Science, Kyoto University, Yoshida Honmachi, Sakyo-ku, Kyoto 606-8501, Japan (tokumaru@ei.energy.kyoto-u.ac.jp)
3. Graduate School of Energy Science, Kyoto University, Yoshida Honmachi, Sakyo-ku, Kyoto 606-8501, Japan (hirotake@ei.energy.kyoto-u.ac.jp)
4. Graduate School of Energy Science, Kyoto University, Yoshida Honmachi, Sakyo-ku, Kyoto 606-8501, Japan (shimoda@ei.energy.kyoto-u.ac.jp)
5. Fugen Decommissioning Engineering Center, Japan Atomic Energy Agency, Tsuruga, Fukui, 914-8510, Japan (kouda.yuya@jaea.go.jp)

Abstract: Augmented reality (AR) can improve the safety and efficiency of maintenance and decommissioning work in nuclear power plants (NPP). Accuracy and speed is required for relocalization methods in order to use AR in NPP. However, the usefulness of the existing methods is not yet evaluated for this application. Thus, in the present research, the representative relocalization methods such as Randomized Fern and FAB-MAP are evaluated and an improved method is proposed. The evaluation results show that Randomized Fern showed relatively faster and more accurate retrieval of camera poses than that of FAB-MAP. However, when the density of keyframes stored in the database decreased for making Randomized Fern applied to a larger environment, the accuracy of the retrieved poses decreased. Therefore, the authors proposed a method that selects the keyframes similar to an input image from a sparse database and searches more appropriate frame around the selected in a dense database. By increasing the keyframes in the dense database, the proposed method is able to prevent errors caused by insufficiency of keyframes in the sparse database. The proposed method was evaluated using the NPP datasets. Consequently, it was able to suppress the degradation of accuracy even when the density of the keyframes were decreased. In the future, evaluation of the proposed method with large datasets that include every part of NPP is required. It is also necessary to evaluate the proposed method in changing environment where workers conduct maintenance or dismantling work of NPP.

Keyword: maintenance support; coarse-to-fine approach; RGB-D camera; simulation; simulation; nuclear energy; augmented reality

1 Introduction

The accident at Fukushima nuclear power plant after the Great East Japan Earthquake remarkably damaged public trust in the safety. After that, many nuclear power plants (NPP) in Japan have been stopped and decided to be decommissioned. The problem when conducting the decommissioning work is that it is necessary to work accurately by manual with step-by-step procedure.

Augmented reality (AR) can improve the safety and efficiency of the decommissioning work in NPP. For utilization of AR in NPP, real time camera tracking is indispensable. The camera tracking is a technique to estimate the position and orientation of a camera using image processing and geometric calculation. In a tracking method using natural features, on the assumption that changes of the position and orientation of a camera and changes of the appearance of images are small in sequential frames, the improvement of speed and stability is aimed by using estimation results of the position and orientation of a camera in the previous frame. However, when the estimation results in the previous frame are not available because of rapid movement of a camera and so on, re-estimation of position and orientation of the camera is necessary. The technique to re-estimate the position and orientation of the camera is called a relocalization. Accuracy and speed is required for relocalization methods in order to use AR in NPP. However, the performance of the existing relocalization methods used in NPP is not yet evaluated. Thus, the existing relocalization methods need to be evaluated in NPP. The purpose of this research is to evaluate the existing methods and propose an improved method that solves problems of the existing methods.

2 Related works

AR is a technology enriching the real world with digital information. For example, digital information such as 3D computer graphics, text information is overlapped with images taken by cameras. Digital data like this enrich our perception and help to understand real world.

For overlapping digital information with images of real world in the right position, position and orientation of a camera with respect to environment taken by a camera are necessary. A technique for the estimation of camera position and orientation (camera pose) is called tracking. There are a lot of tracking methods using different technologies [1][2]. For example, there are technologies using gyroscope sensor, acceleration sensor, magnetic sensor, ultrasonic sensor, GPS and vision sensor such as a camera.

However, methods available in the environment such as NPP are limited. Gyroscope and acceleration sensor have accumulated errors. NPP is large and has complicated objects such as pipes, valves, tanks. In such complex environment, ultrasonic sensor is useless because accurate reflected waves are not obtainable. Magnetic sensor can't be used in environment that has large metal objects. GPS is useless in indoor environment.

Thus, appearance-based method using a camera is desirable in NPP. In this method, camera poses are calculated by using artificial markers put in the environment or natural features such as lines and points existing in the environment. However, method using artificial markers requires users to set markers in the environment and to measure the positions of markers. Therefore, the method using artificial markers is not suited to be used in NPP. Thus, method utilizing natural features [3] is desirable.

In most feature-based methods, a camera pose is estimated by minimizing difference between 2D position of a natural feature point detected in a camera image, and projected point of a corresponding point whose 3D position is known onto the same camera image.

On the assumption that a change of moving camera pose is small in the sequence images, a current camera pose is calculated by solving an optimization problem that minimizes errors defined in each tracking methods by setting the initial value to the previous frame's camera pose. When this assumption is broken because of an erratic motion, the solution of the optimization problem may not converge to an optimal solution (true camera pose). If the error after the minimization is bigger than a threshold, the tracking system judges that tracking fails. This situation is called a lost. When a lost happens, the technique that re-estimates a current camera pose in the different way is necessary and the technique is called a relocalization.

In relocalization methods, at first, the database that consists of an image and a camera pose where the image is taken is created while the tracking successes. In many cases, the captured images are encoded to reduce the amount of data and process time. When the tracking lost happens, in other words, when a current camera pose is lost, the current input camera image is compared with images stored in the database and the image similar with the input image is searched from the database. If two camera poses are similar, two images are also similar. Therefore, we can assume that when an image similar with the current input image is found in the database, the camera pose of the similar image is close to the current camera pose. Thus, the system tries to resume tracking by solving error minimization problem again by setting an initial value to the camera pose paired with the similar image.

There are few researches that evaluate the performance of feature-based tracking and relocalization methods in NPP. Therefore, the performance of the existing relocalization methods is unclear. Generally the appearance-based relocalization methods that use camera images have some problems. (1) : The camera poses can be retrieved only from the poses stored in the database, in other words, the camera can be relocalized only to the places where the camera visited previously. (2) : The system recognizes different scenes as same place when the scenes are similar apparently. (3) : In a large environment, database becomes large and process time becomes long.

3 Evaluation of the existing methods

3.1 Methods evaluated in this research

In the present research, Randomized Fern [4] and FAB-MAP [5] are evaluated as the representative existing relocalization methods.

Randomized Fern is a method that encodes RGB-D images as binary codes and searches for an apparently similar image quickly by searching a similar binary code. In this method, pixels are randomly selected from an image. A fern is defined as a set of consecutive nodes. A fern generates a binary code block from RGB-D pixel values by comparing the values with thresholds which are also decided randomly. Thus, whole image is represented as short binary code. Based on those representations, the block-wise hamming distance (BlockHD) is introduced as an efficient frame dissimilarity measure. The BlockHD calculates the number of differing code blocks between two images and is normalized. For every incoming new frame, minimum BlockHD with respect to previously stored frames is calculated. Only input images whose minimum BlockHD is larger than the threshold are stored in database as new keyframes. In the relocalization, a keyframe with minimum BlockHD against input image is selected.

FAB-MAP is an approach to appearance-based place recognition and can be applied to the relocalization. FAB-MAP compares an input image with images stored in database and determines the probability that the input image is taken in same location where images in database are taken. In this method, a camera image is represented as a set of features by using visual bag-of-words models. These features are called words and a collection of features is called vocabulary. FAB-MAP requires training data (e.g. a collection of images taken in a similar but not identical environment) to make a visual vocabulary, by using a Chow-Liu tree representation of feature likelihood. In FAB-MAP, all input images for the database are stored in the database. A frame that has highest probability of same location is selected from the database in the relocalization.

3.2 Evaluation method

Recovery rate is used as an evaluation indicator of relocalization methods in this research. The recovery rate is the percentage of recovery frames that succeed the relocalization with respect to the total number of frames in input dataset. This indicator is used to evaluate how accurately relocalization methods can estimate camera pose.

In this research, if the translation error between estimated camera pose and true camera pose (ground truth) is within 0.1 meters and rotation error is within 5 degrees, the relocalization is considered as a success. These threshold values are decided because most existing tracking system can converge to optimal solution even with these errors.

By using this estimated pose as the initial value, the tracking system solves optimization problem and tries to restart tracking. If the errors are within the thresholds, we assume that the problem definitely converges to optimal solution and tracking is success

We also use process time as an evaluation indicator. Given an input image, a relocalization system searches for similar frame in the database and return camera pose paired with this frame as estimated pose. The time that this process needs is the process time per frame. The process time is calculated as a mean of process time per frame of all input dataset frames.

In this research, we changed number of ferns and a threshold of BlockHD in Randomized Fern and vocabularies in FAB-MAP by trials and evaluated methods. Randomized Fern uses BlockHD when making database and when relocalizing. At these times, the threshold of BlockHD is set to same value. Vocabularies used in this research are shown in table 1. The performance of FAB-MAP depends on vocabularies.

Table 1 Vocabulary used for FAB-MAP

	Default vocabulary	Plant vocabulary
Acquisition method	Attached in OpenCV 2.4.9	Made in the water purification room
Number of words	545	7,661
Size	679KB	9,533KB

3.3 Datasets

The ground truth of camera poses is necessary to evaluate the accuracy of relocalization methods. However, appropriate datasets in NPP for the evaluation are not available. Thus, in this research the pairs of an image in nuclear power plant and a camera pose where the image is taken are created and used for the evaluation of relocalization methods. For the following discussions, the set of these pairs is called a dataset.

Two types of dataset are required for the evaluation of the relocalization methods evaluated in this research. The first one is for building the database. The second one is used as input data in the relocalization. Camera poses obtained with the other methods are used as ground truth to evaluate the accuracy of the relocalization methods.

We utilized KinectFusion [6] for obtaining the ground truth of camera the poses. Only when the camera moved very slowly, KinectFusion can create a 3D environmental model from depth images captured by a depth camera. In this process, KinectFusion estimates the camera poses of the depth camera simultaneously. We used these estimated values as the ground truth. KinectFusion requires users only a depth camera and easy to use even in nuclear power plant.

But the problem exists that the ground truth obtained with KinectFusion has some accumulating errors. The longer the range of the camera trajectory is, the larger the error of camera poses obtained by Kinect Fusion is. We limited the maximum range of one camera trajectory up to 3 meters. (The maximum range should be longer to evaluate the relocalization methods more properly. We plan to increase the range in the future.) We repeated the capture several times along with the different trajectory (we mainly varied the height of the trajectory), and combined the results to obtain a large dataset. In order to combine the results into single dataset, the coordinate systems must be properly aligned. We placed two tripods in the capturing environment as the starting point (Check point A) and the end point (Check point B) of the trajectory and tried to set the initial and final pose of the camera for every trajectory identical. It is difficult to set the camera pose in every trajectory to the strictly same pose, but we can assume that the camera poses are almost same relative to the environment compared to the accumulating error caused by Kinect Fusion. Data capturing area and captured environment are shown in Fig.1. We captured images along with the arrows.

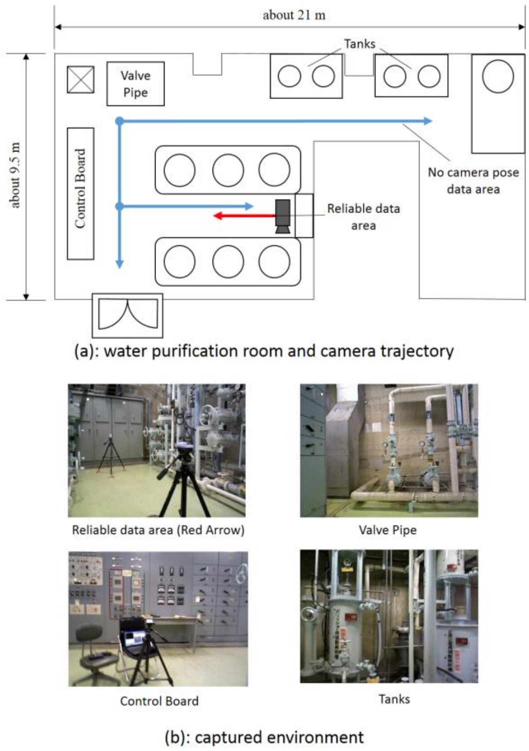


Fig.1 Capturing data in water purification room

We captured images in the water purification room in JAEA Fugen Decommissioning Engineering Center and created datasets. We used ASUS Xtion Pro live and capturing frame rate was 30 fps. We used kinfu included in PCL library Ver.1.7.1.

We captured images along with a red trajectory in Fig.1 and created the ground truth by KinectFusion. The datasets with ground truth are called the datasets of reliable area. The water purification room is much larger than the area that KinectFusion can produce reliable results. But we want to evaluate the ability to search for a similar image from the entire room database. Then, datasets with no camera pose data were also made by using images of the entire room (blue arrows in Fig.1). Only datasets with ground truth are used as input data in relocalization methods but the database was built using the images for the entire room. If the relocalization method returns the frame of the reliable data area, the pose with which the frame was captured is used as the estimated pose. If the frame without camera pose is returned, it is deemed that the retrieved pose is not in the reliable data area and the relocalization failed. The trajectories of datasets of the reliable data area are very narrow. However, we can properly evaluate the performance of the retrieval at least in the horizontal direction of trajectories.

The number of the frames in the datasets for evaluation is shown in table 2. Datasets of the reliable data area consist of seven trajectory datasets as shown in table 3. One of these datasets is used as input dataset. The rest of datasets are used as datasets for making database for the relocalization. There are seven pattern of combination. We used result of a mean of seven dataset patterns. When making the database for the relocalization, one frame per five frames in datasets for the database was used for the reduction of the time of the relocalization. We confirmed that this reduction doesn’t affect the results of the relocalization,

Table 2 The number of frames of datasets for evaluation

Dataset with ground truth	Dataset with no camera pose data	Total number of frames
9,461	111,272	120,733

Table 3 Datasets of the reliable data area

	Number of frames
Trajectory 1	1,243
Trajectory 2	1,237
Trajectory 3	1,158
Trajectory 4	1,279
Trajectory 5	1,293
Trajectory 6	1,684
Trajectory 7	1,567

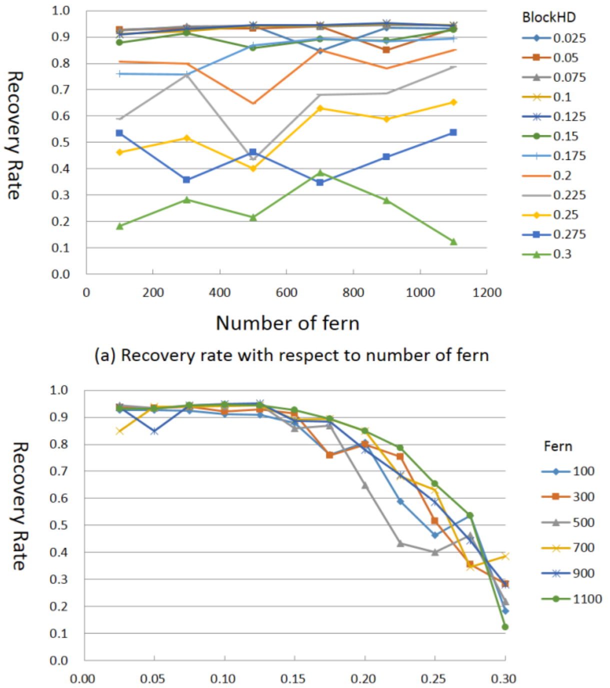
3.4 PC used in the evaluation

PC used in this research is as follows. The processor is Intel Core i7-4770S 3.10 GHz (single core was used). Memory is DDR3 PC3-12800 8GB × 2 (total 16GB).

3.5 Evaluation results

3.5.1 Recovery rate

The recovery rate of Randomized Fern is shown in Fig.2. Absolute value of the recovery rate doesn’t matter because the rate depends on environment. When the threshold of BlockHD is lower than 0.15, recovery rate is high and stable. The higher threshold of Block HD is, the lower recovery rate is. This is because as shown in Fig.3 the number of keyframes in database is few when Block HD is high. When the number of keyframes is few, estimated pose can’t retrieve to the correct pose. Regardless of the number of fern, recovery rate is not stable and sometimes decreases significantly. In Randomized Fern, the thresholds for encoding images are decided randomly and binary code of images changes by trials. Therefore, BlockHD between images also changes. This is why recovery rate is sometimes unstable. This changes the number of keyframes and contents of keyframes and changes the result. It is considered that the change of keyframes affects the result when threshold of BlockHD is high and the number of keyframes is few. As described above, if the threshold of BlockHD is lower than 0.15, high and stable recovery rate is achieved regardless of the number of fern.



(a) Recovery rate with respect to number of fern

(b) Process time with respect to threshold of BlockHD

Fig.2 Recovery rate of Randomized Fern

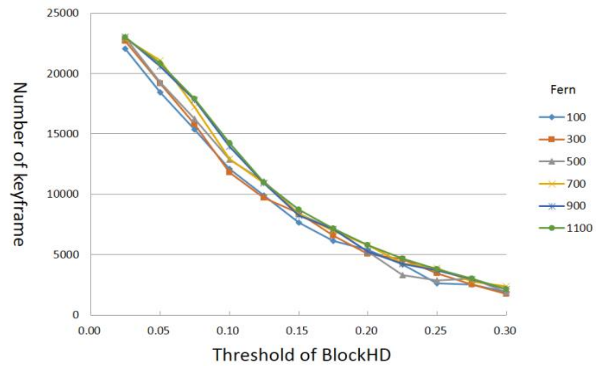


Fig.3 Number of keyframes with respect to threshold of BlockHD

The recovery rate of FAB-MAP is shown in Table 4. The result with the plant vocabulary is higher than the result with the default vocabulary. In FAB-MAP, images are represented as combination of words (feature points) in vocabulary. Therefore, FAB-MAP can distinguish images accurately in environment that has many word feature points. On the other hand, FAB-MAP can't distinguish images in environment that does not have any feature points in the vocabulary. Therefore, with the plant vocabulary FAB-MAP can judge the degree of similarity of images more accurately than with the default vocabulary. This is why recovery rate with the plant vocabulary is higher than that with the default vocabulary. Using the plant vocabulary we can increase the recovery rate of FAB-MAP. Therefore, we use the plant vocabulary in FAB-MAP in the following.

Table 4 Results of FAB-MAP

	Default vocabulary	Plant vocabulary
Recovery rate	0.564	0.784
Process time (ms)	289.04	737.247
Dispersion of Process time (ms^2)	80.68	414.579
Size of testdata (KB)	85,371	853,202
Memory Usage (KB)	86,733	962,378

3.5.2 Process time

The process time is shown in Fig.4. In Randomized Fern, process time tends to increase with respect to the number of fern. Because BlockHD is calculated with respect to binary codes of all frames in database, the process time per frame is increasing in proportion to the number of fern. On the other hand, when threshold of BlockHD is high, the process time decreases exponentially. This is because the number of keyframes in the database decreases exponentially.

The process time in FAB-MAP is shown in Table 4. Process time with the plant vocabulary is longer than that with the default vocabulary. The plant vocabulary has more words than the default vocabulary. Therefore, FAB-MAP needs more time to calculate.

When AR is applied to the dismantling work support, it is desirable that process time is less than about 1000 ms. Process time in Randomized Fern is about few ms. Thus, Randomized Fern is useful for practical use in terms of time. Process time in FAB-MAP is also less than 1000 ms in the slowest condition. Unlike computer we use in this research, mobile computer assumed to be used for supporting dismantling works will take longer time. It is probably difficult to use in large environment like NPP.

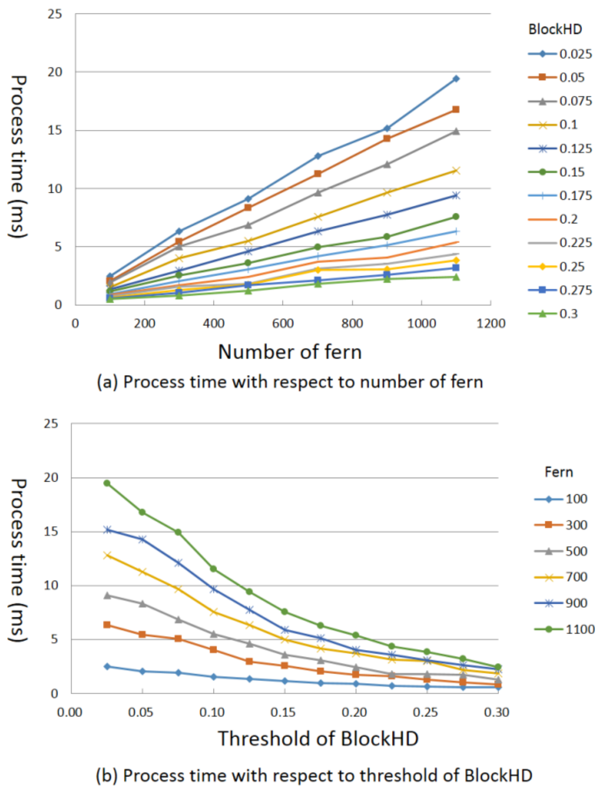


Fig.4 Process time of Randomized Fern

3.5.3 Problems of the existing methods

According to the above results, Randomized Fern showed high recovery rate in the dataset made for NPP. However, the performance of recovery in Randomized Fern highly depends on the number of keyframes. The more densely keyframes are collected, the higher the success rate of relocalization will be. But it is difficult that the relocalization system converges to correct camera pose when spatial density of keyframes are low. This is because the relocalization system can return only keyframes stored in the database. In this research, the dataset includes only a part of a nuclear power plant. When practical supporting works are assumed, relocalization methods are required to be capable of retrieving in larger area. Then, it is expected that the number of keyframes remarkably increases and the process time increases. As the usability of AR is considered, real-time processing is required. However, when the threshold of BlockHD is set high to reduce the number of keyframes for process acceleration, the recovery rate significantly decreases as shown in Fig.2. When the number of keyframes decreases, the randomness of Randomized Fern makes the results unstable. In this research, the range in which estimated camera pose converges to optimal solution is assumed that translation error is within 0.1 meters and rotation error is within 5 degrees. However, actual tracking systems and environments may require more strict conditions. Then, Randomized Fern requires the much more density of keyframes to show a sufficient performance of the recovery rate.

4 Proposal and evaluation of 2-step Randomized Fern

4.1 Proposal of 2-step Randomized Fern

We propose a method that tackles the problems of existing methods mentioned above. Our approach is based on Randomized Fern and a simple coarse-fine searching. In the proposed method, dense and sparse database are created. In the relocalization, the proposed method selects several candidates of similar images from the sparse database and searches for most similar frame from the frames around each candidate in the dense database. If only one candidate is selected, the randomness may cause the change of the result. However, if the number of candidates is too many, frames to be searched will be too many. Therefore, in this research the number of candidate is set to five. Five most similar frames are selected. Then, using these frames, tracking system tries to restart tracking.

By using the sparse database, time for searching can decrease, on the other hand spatial density of keyframes is kept by using the dense database. In a large environment, the dense database is large. However, the sparse database makes searching fast. The dense database can keep the recovery rate high. Searching keyframes around the candidate also mitigates the effect of randomness of Randomized Fern. When the number of the candidates is few, wrong selection caused by the change of a binary code from the randomness can affect the result of the relocalization. Increasing the number of the candidates decreases the effect of wrong selection because the probability of selecting a wrong candidate decreases.

In the following, the proposed method is explained in detail. First, when creating the dense database, a threshold of BlockHD of the dense database is set low. Then, the proposed system calculates BlockHD of the dense database between an input image and all frames in the dense database. The input images that have the BlockHD more than the threshold, in other words, images that are dissimilar to the dense database frames and new, are only collected as keyframes. Keyframes of the dense database are collected densely. When the input image is collected, the proposed method proceeds to the next step. Next, BlockHD of the sparse database between the input image and all frames in the sparse database is calculated. The threshold of BlockHD of the sparse database is set higher than that of the dense database. The sparse database collects only frames that fulfill more strict condition. Therefore, the number of sparse keyframes is small and appearance of frames of the sparse database is much different each other. In the relocalization, firstly the system searches for five frames that have smallest BlockHD against the input image in the sparse database. Then, BlockHDs between input and 100 frames around the five selected frames (50 frames before and 49 frames after each candidate and candidate itself) in the dense database are calculated, which means that only at most 500 frames are searched in total. Finally, system selects a frame that has smallest BlockHD against input image from each 100 frames. Thus, five keyframes are selected. Five camera poses that are paired with the five keyframes are used as estimated camera poses. Using these poses, tracking is tried to restart. The tracking system tries five poses in turn as an initial value. When restarting succeeds, this trial is stopped. Appropriate values of parameters (the number of candidate and the number of frames around candidate) depend on frame rate, movement and speed of camera, complexity of environment and so on. Thus, we need to consider optimal settings of parameters in the future.

In a large environment, the number of keyframes in the sparse database is relatively small compared with that of the original Randomized Fern. Therefore, process time is kept fast. The system can search from much more candidates than Randomized Fern and can retrieve to much more keyframes.

On the other hand, the proposed method requires two datasets and uses larger memory. In addition, images apparently similar with each other sometimes can't be distinguished. In environment that doesn't change apparently, this method also doesn't work appropriately.

4.2 Evaluation results

4.2.1 Evaluation method

Evaluation method for 2-step Randomized Fern is same as that for the existing methods. The number of ferns is set to 500. Evaluation indicators are the recovery rate and process time.

4.2.2 Recovery rate

Recovery rate of the proposed method and the original Randomized Fern is shown in Fig.5. In 2-step Randomized Fern, recovery rate is almost same as that of the original Randomized Fern when the threshold of BlockHD of the sparse database is low. Even when the threshold of BlockHD of the sparse database is high, the recovery rate of the proposed method is kept high. As far as the threshold of BlockHD of the dense database is set from 0 to 0.5, the threshold doesn't affect recovery rate so much. The dense database can compensate the effect of the reduction of keyframes in the sparse database.

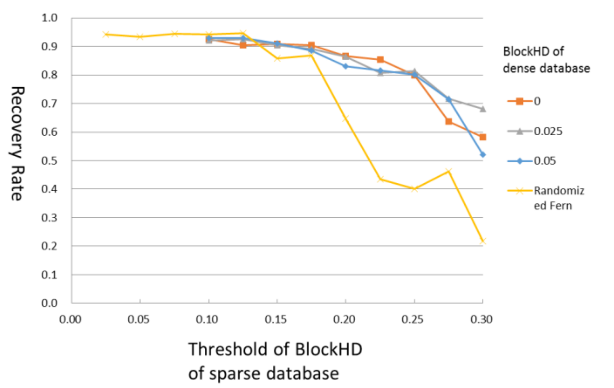


Fig.5 Recovery rate of 2-step Randomized Fern

4.2.3 Process time

Process time of 2-step Randomized Fern and the original Randomized Fern is shown in Fig.6. When the threshold of BlockHD of the sparse database is high, the number of keyframes in the database becomes small. Therefore, the process time tends to decrease in proportion to the threshold of BlockHD of the sparse database. Change of the threshold of BlockHD of the dense database doesn't affect the process time so much. As supporting maintenance and decommissioning work are assumed, it is desirable that process time is less than about 1000 ms. Therefore, it seems that the proposed method can be used practically in a large environment like NPP. In the condition of this research, the proposed method takes more time than the original Randomized Fern because of 100 frames searching in the dense database. However, if the environment is much larger, the threshold of BlockHD of the original Randomized Fern will be required to be lower for keeping the recovery rate of the original method high. Therefore, in case of the much larger environment, the process of the proposed method will be faster than that of the original Randomized Fern.

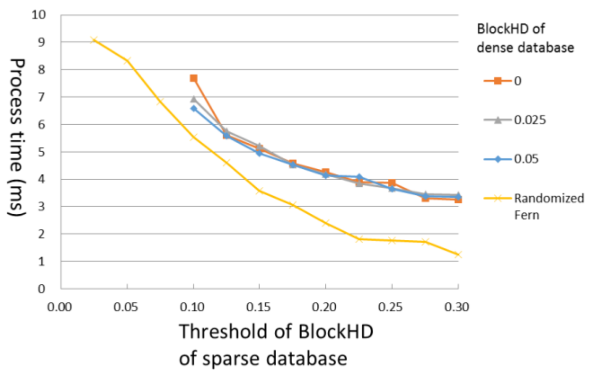


Fig.6 Process time of 2-step Randomized Fern

5 Conclusion

This study aims to evaluate the existing relocalization methods and to propose an improved method in NPP.

We evaluated Randomized Fern and FAB-MAP using a dataset which includes the images of water purification room in JAEA Fugen Decommissioning Engineering Center.

The results showed that Randomized Fern showed relatively faster and more accurate retrieval of camera poses than that of FAB-MAP. However, in

Randomized Fern, when the number of keyframes stored in the database decreased, the accuracy of the retrieved poses also decreased. This may cause a problem when using in NPP as it is difficult to store all keyframes in a database without thinning of the keyframes if a low performance mobile computer is used.

We proposed 2step-Randomized Fern which used both sparse and dense database. The sparse database can accelerate searching and the dense database can keep the density of keyframes.

In 2-step Randomized Fern, recovery rate was kept high even if the number of keyframes in the sparse database decreased. Maximum recovery rate was almost same as that of Randomize Fern with optimal setting. Process time of the proposed method was longer than that of Randomize Fern but was practical.

In this research, KinectFusion was used for obtaining the ground truth. However, the accuracy of the ground truth in large environment was not enough. Thus, the dataset with reliable ground truth was limited. The performance of relocalization depends on the environment. Therefore, the results of evaluation may change for the different environment.

In the future, evaluation of the proposed method with large datasets that include every part of NPP is necessary. It is also necessary to evaluate the proposed method in changing environment where workers conduct maintenance or dismantling work of NPP.

References

[1] J. Newman, D. Ingram, and A. Hopper, Augmented Reality in a Wide Area Sentient Environment, Proceedings of the International Symposium on Mixed and Augmented Reality(ISMAR) 2011, (2001) pp.77-86.

[2] H. Kato, M. Billinghurst, Marker Tracking and HMD Calibration for a Video-based Augmented Reality, 2nd IEEE and ACM International Workshop, (1999) pp.85-94.

[3] H. Wuest, F. Vial and D. Stricker. Adaptive line tracking with multiple hypotheses for augmented reality, Proceedings of the International Symposium on Mixed and Augmented Reality(ISMAR)2005,(2005)pp. 62-69,

[4] B. Glocker, S. Izadi, J. Shotton, and A. Criminisi, Real-Time RGB-D Camera Relocalization, IEEE, Science and Technology Proceedings, 1(4) (2013) pp.173-179.

[5] M. Cummins and P. Newman, FAB-MAP: Probabilistic Localization and Mapping in the Space of Appearance, The International Journal of Robotics Research 2008, 27(6) (2008) pp.647-665.

[6] R. A. Newcombe, O. Hilliges, D. Molyneaux, D. Kim, A. J. Davison, P. Kohli, J. Shotton, S. Hodges, and A. Fitzgibbon, KinectFusion: Real-Time Dense Surface Mapping and Tracking, IEEE International Symposium on Mixed and Augmented Reality 2011, Science and Technolgy Proceedings, (2011) pp.26 -29.