The 2009 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 11-15, 2009 St. Louis, USA

# View-Sequece Based Indoor/Outdoor Navigation Robust to Illumination Changes

Yoichiro Yamagi, Junichi Ido, Kentaro Takemura, Yoshio Matsumoto, Jun Takamatsu, and Tsukasa Ogasawara

*Abstract*— **We propose a view-based indoor/outdoor navigation method as an extension of the view-sequence navigation. The original view-sequence navigation method uses the template matching method with normalized correlation for localization. Because the matching method is sensitive to local illumination changes, it is only used for indoor environment. In this paper, we propose to adopt the accumulated block matching method to improve robustness against locally changing illumination, in which a template is split into small patches and matched by maximizing the average of the normalized correlations of all the patches. We also propose a localization criterion which helps the robot decide its motion. Our experimental results demonstrate that the proposed methods can be applied to both indoor and outdoor environments.**

## I. INTRODUCTION

The advancement of mobile robot technology gives birth to various kinds of helpful service robots such as surveillance robot and museum guide robot [2]. They strongly attract our attention because of their effectiveness. Although it is not a problem that a specific-purpose robot as described above can only move around restricted interesting areas, multi-purpose robots should have the ability to seamlessly move in both indoor and outdoor environments.

In this paper, we propose a robot system for seamless indoor/outdoor navigation. The proposed system does not distinguish the way of navigation for the type of environment, which simplifies the implementation of the system. We improved the view sequence navigation [6], which is only applicable to the indoor environment. We provided the method with robustness for the illumination changes which often occur in the outdoor environment.

The rest of this paper is organized as follows. Section 2 introduces related work. Section 3 describes the method to achieve the robustness for the illumination changes. Section 4 shows the comparison between the original view-sequence navigation method and the proposed method. We also show the experimental result of indoor/outdoor navigation. Section 5 concludes this paper and describes the future work.

## II. RELATED WORK

Research methods for mobile robot navigation are roughly classified into two types based on the use of specific devices and/or visual tags (denoted as *landmark*). As examples of the methods which use landmarks, Takeuchi *et al*. [10] proposed to use the QR code and Kulyukin *et al*. [5] proposed to use the RFID tag for the navigation. However, these methods require lots of work to set up the landmarks and limit the environment where the landmarks are available.

In contrast, methods that use existing landmarks have been proposed. For example, the navigation using the global positioning system (GPS) was proposed (*e.g*., [1]). Since the position estimated by the GPS is usually not accurate enough for navigation, methods that fuse with odometry [7] and with both odometry and laser range sensor [4] were proposed to improve the accuracy of the estimation. However, the GPS is only available in the outdoor environment. Yoshida *et al*. [12] proposed to use Braille blocks. But the blocks are sparse and therefore the method limits the applicable environment.

As examples of methods which do not use landmarks, Matsumoto *et al*. proposed a view sequence navigation method [6], which helps the robot navigate by adjusting the difference between the currently observed view and the corresponding view which were recorded in advance. This method uses a template matching method to compare the recorded and the current images. The method is very sensitive to occlusions and illumination changes. This drawback limits the method to the indoor environment. In order to solve the occlusion problem, the use of ceiling images was proposed [11]. However, problem on illumination changes still remains.

Katsura *et al*. [3] proposed the view sequence method to compare the images, not directly, but based on pre-learned visual features, such as features of sky, tree, building, and artificial materials. As a result, this method is robust to illumination changes. The visual features used in this method usually appear in the outdoor environment, not in the indoor environment. They did not show the effectiveness of the method for the indoor environment.

In this paper, we improve the view sequence navigation method proposed by Matsumoto *et al*. [6]. Our method has two advantages: it enables the easy implementation and it does not rely on landmarks. By providing the method with the robustness to illumination changes, we aim at achieving seamless indoor/outdoor navigation.

## III. VIEW SEQUENCE NAVIGATION ROBUST TO ILLUMINATION CHANGES

The view sequence navigation method [6] first records camera images while moving the mobile robot along the target path, for teaching the view sequence. Next, it moves the robot to adjust the difference between the current view

Y. Yamagi, K. Takemura, J. Takamatsu, and T. Ogasawara are with Graduate School of Information Science, Nara Institute of Science and Technology, Japan {yoichiro-y, junichi-i, kenta-ta, j-taka, ogasawar}@is.naist.jp

J. Ido and Y. Matsumoto are with National Institute of Advanced Industrial Science and Technology, Japan {ido-j, yoshio.matsumoto}@aist.go.jp

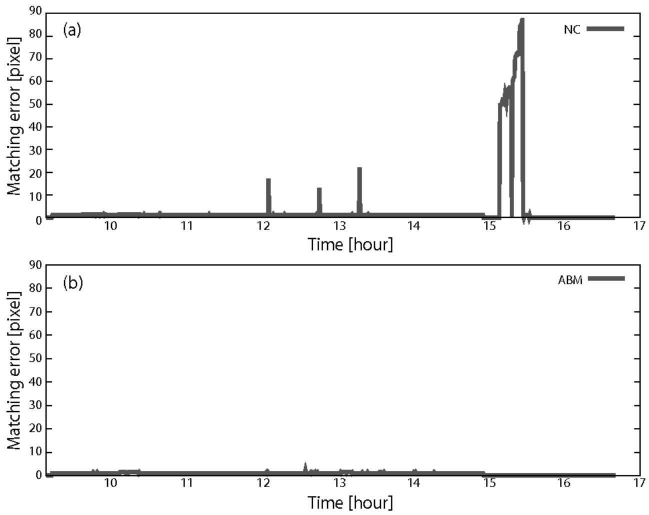



Fig. 1. Matching errors of the template matching method (upper) and the ABM method (lower). The horizontal axis indicates the time when the image was captured and the vertical axis indicates the error. The errors in the ABM method are smaller than the ones in the template matching method.

sequence and the recorded one. For navigation, the following three issues should be considered:

1) How to calculate the difference between the current view and the given target view
2) How to select the appropriate target view from the view sequence
3) How to decide the reaction of the robot in order to adjust the view difference

Note that the order of the views in the view sequence helps to select the target view. This means that the algorithm only needs to decide whether to keep the current target view or to change it to the next recorded view; a robot is firstly located on the start position and therefore the first target view is known.

### A. Accumulated block matching method

The original view sequence navigation uses a template matching method with normalized correlation. Due to the drawbacks of the matching method, the local illumination changes affect the navigation. To solve this problem, we propose to use an accumulated block matching (ABM) method [8] in place of the template matching method. The ABM method splits the template into several small patches[1]. Matching score is defined as the average of the normalized correlations in all the patches, while they move together during matching. The ABM method can correctly match the template, even in the case of partial occlusions, by maximizing the score. Because the local illumination changes can be regarded as occlusion, it is expected that the use of the ABM method improves the matching accuracy.

[1]Saji *et al*. proposed the method for adaptively splitting the template and for appropriately changing the number of patches [9]. This method improves matching stability, but requires much computation. In this method, we fix the splitting method.

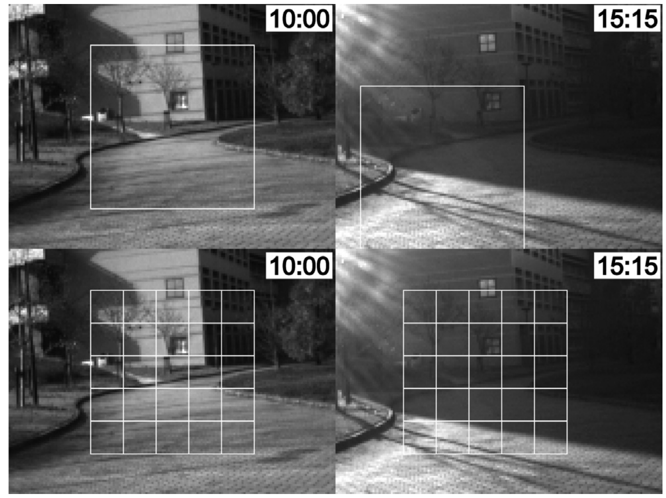



Fig. 2. The results of the template matching method (upper row) and of the ABM method (lower row). Rectangles at the left column correspond to the templates and ones at the right column show the matching results. The severe illumination change affects the result of the template matching method; the matched position is far from the center, *i.e*., ground truth. In contrast, the ABM method correctly matches the template.

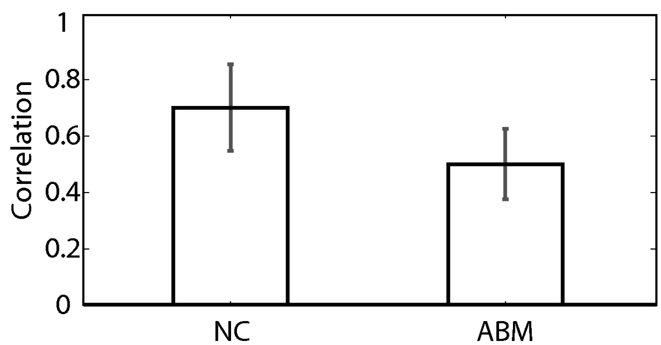



Fig. 3. Averages of the normalized correlations in the template matching method (left) and the matching scores in the ABM method (right). Although the ranges of the two values are the same, the average in the ABM method is smaller.

To investigate the robustness to illumination changes, we captured 840 outdoor images of a fixed camera from 10:00 to 16:00 every 30 seconds and then applied the template matching method and the ABM method to the images. The image captured at 10:00 was used as the template.

Fig. 1 shows the magnitude of the displacement vector $(x_{error}, y_{error})$ of matching position, *i.e*., $\sqrt{x_{error}^2 + y_{error}^2}$. Since the camera position was also fixed, it is preferable that the matching position is fixed, *i.e*., the magnitude equals zero. These graphs prove that matching by the ABM method is stabler than by the template matching method.

Fig. 2 shows the matching results at around 15:00. Severe illumination change affects the result of the template matching method. However, the ABM method correctly matches the template.

Fig. 3 shows the averages of the normalized correlation and the ABM's matching scores. Although the range of

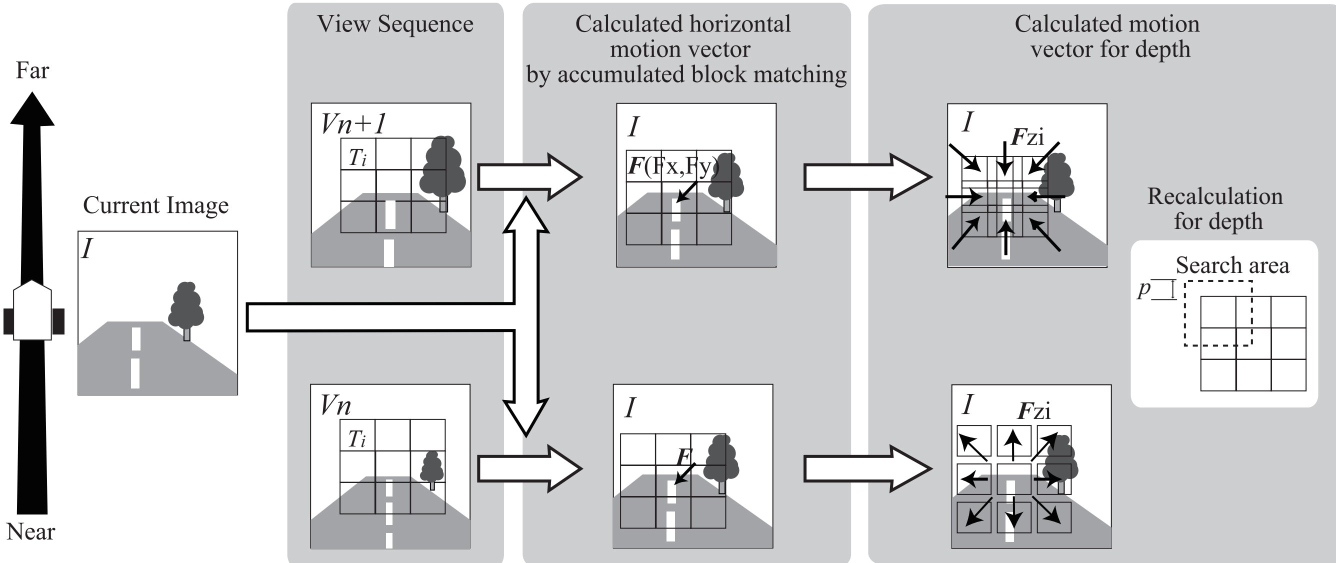



Fig. 4. Calculation of two types of image distances. The ABM method first matches the current image $I$ to the target view $V_n$ and the next view $V_{n+1}$. We define the displacement of the matching position along horizontal direction $F_x$ as *the horizontal distance of the image*. After matching by the ABM method, the template matching method with the normalized correlation searches the more precise matching independently in each of the small patches. We define the average of the magnitudes of the displacement $\mathbf{F}_{zi}$ as *the depth distance of the image*. These two criteria are employed for selecting the target view, deciding the robot motion, and generating the view sequences.

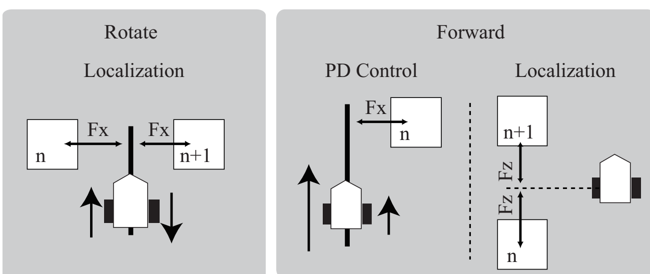



Fig. 5. Robot control and localization. Localization in rotation is achieved by the horizontal distance of the image. Localization in moving forward is achieved by the depth distance of the image. In the case of moving forward, the robot simultaneously adjusts the steering based on the horizontal distance.

each value is the same, *i.e.*, from $-1$ to $+1$, the average of the ABM's matching score is smaller. Due to this affect, it becomes inaccurate to select the target view by the score. Thus, we designed novel criteria for the selection.

### *B. Selection metrics between two images*

For appropriately selecting the target view from the view sequence, we define two types of criteria. Fig. 4 shows the overview of the method for calculating them. The ABM method first matches the currently captured image $I$ and the target view $V_n$. The central area of the target view is used as the template. Thus, the displacement of the matching position from the center, $\boldsymbol{F}$, is assumed as the difference of the directions between the target path and the current path. We define the horizontal element of the displacement, $F_x$, as *the horizontal distance of the image*.

After matching by the ABM method, the template matching method with normalized correlation searches the more precise matching independently in each of the $m$ small patches, $T_i$ $(i = 1, 2, \ldots, m)$. As the result, the displacements $\boldsymbol{F_{zi}}$ of all the patches are obtained. We define $F_z$ in Eq. (1) as *the depth distance of the image*.

$$F_z = \frac{1}{s} \sum_{i \in K} |\mathbf{F_{z}}_i|. \tag{1}$$

$K$ is the set which includes all the small patches whose normalized correlation is greater than zero and $s$ is the number of elements in the set $K$. To improve the stability in calculating the depth distance, we use the values which are obtained by applying a low-pass filter (simple moving average in this paper) to $F_z$.

### *C. View sequence using ABM method and image distances*

As shown in Fig. 5, robot motions in indoor/outdoor environments consist of the following two types:

- rotation (*e.g.*, pivoting) in a narrow area, and
- forward motion with small direction change.

In the case of rotation, we employ the horizontal distance of the image, $F_x$, for changing the target view. From the definition, the distance corresponds to the difference of the directions between the recorded view and the current view. Let the target view and the next view be $V_n$ and $V_{n+1}$, respectively. The target view is changed just when $F_{x(n)} > F_{x(n+1)}$, where $F_{x(n)}$ and $F_{x(n+1)}$ are the horizontal distances for the two views, $V_n$ and $V_{n+1}$, respectively.

In the case of moving forward, we employ the depth distance of the image, $F_z$, for changing the target view. In detail, the target view is changed just when $F_{z(n)} > F_{z(n+1)}$,

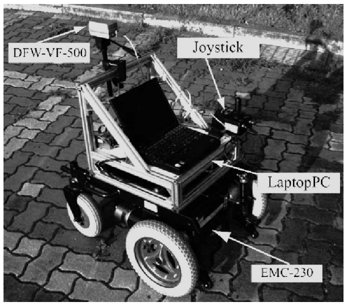



Fig. 6. Mobile Robot (EMC-230)

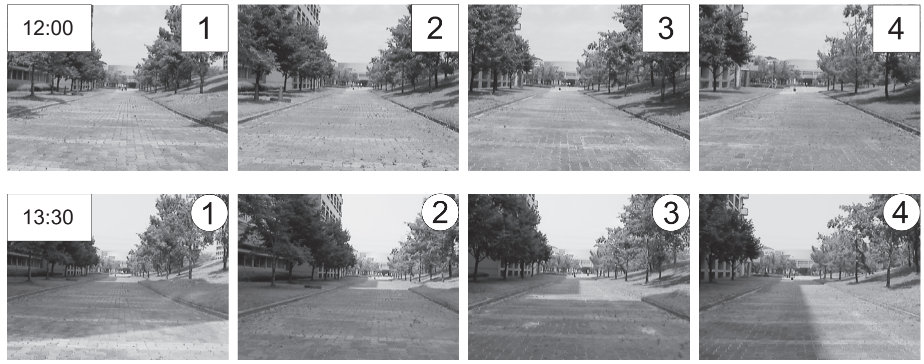



Fig. 7. Recorded image sequences at 12:00 (upper) and at 13:30 (lower)

where $F_{z(n)}$ and $F_{z(n+1)}$ are the depth distances for the two views, $V_n$ and $V_{n+1}$, respectively.

Consider the case where the robot moves before the position where the view $V_n$ was captured. As the robot moves forward, $F_{z(n)}$ and $F_{z(n+1)}$ decrease together while $F_{z(n)} < F_{z(n+1)}$ holds. When a robot approaches the position where the view $V_n$ was captured, the value of $F_{z(n)}$ is minimized. After that, $F_{z(n)}$ increases and $F_{z(n+1)}$ decreases, as the robot moves away. Thus, the target view is changed when $F_{z(n)} > F_{z(n+1)}$. Note that in teaching the view sequences the horizontal distance $F_x$ for rotating and the depth distance $F_z$ for moving forward, are used in place of the normalized correlation.

The rotation can be achieved by continuously rotating until approaching the goal view. In the case of moving forward, the robot needs to control the steering to adjust the difference between the current view and the target view. The steering angle is decided from the horizontal distance $F_{x(n)}$. This algorithm can move the robot along a moderately curved path.

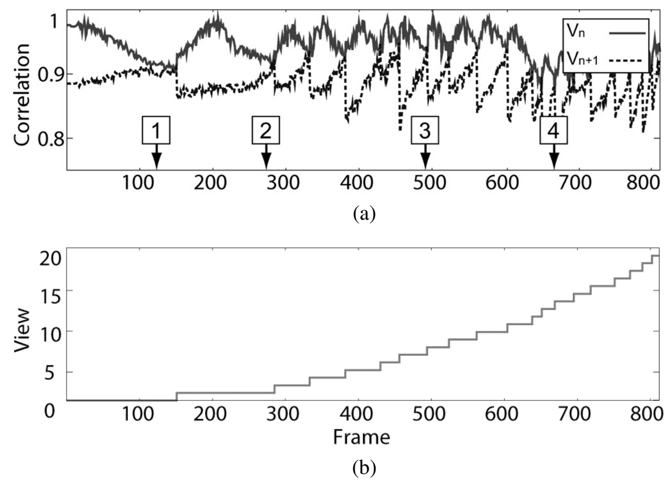



Fig. 8. (a) Normalized correlations of the target view $V_n$ and the next view $V_{n+1}$ to the current image, and (b) the selected view ID in each frame at 12:00. The horizontal axis indicates the frame number of the image which is used as the current image. The vertical axes indicate normalized correlation in the upper graph and view ID in the lower graph. The target views are appropriately changed.

## IV. Experiments

### A. Mobile robot

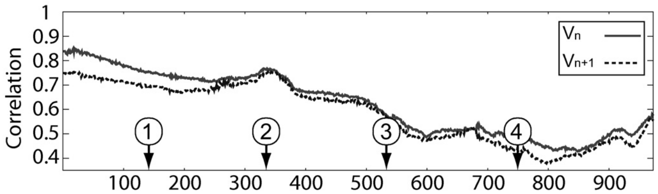
In this experiment, an electric wheelchair (Imasen Engineering Corporation: EMC-230) was used as the mobile robot (see Fig. 6). The robot is controlled through the USB I/O port (Technowave Ltd.: USBM3069F) and is equipped with an IEEE 1394 camera (SONY Corporation: DFW-VL-500) to capture gray-scale images.

### B. Comparison to original view sequence method

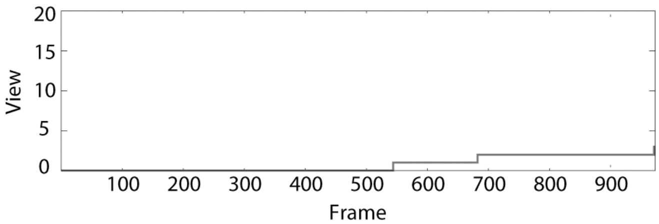
To verify the effectiveness of the proposed method, we compared the proposed method to the original view sequence method. Success in the localization can be determined if the target views are appropriately changed. We employed a 100 [m] outdoor straight path in our campus for the comparison. Fig. 7 shows several views of the path. As can be seen from this figure, illumination changes due to time differences dramatically alter the images.

The size of the input view image was $80 \times 60$ [pixel]. The image was obtained by down-sampling the captured gray-scale images whose size was $640 \times 480$ [pixel]. We manually moved the robot twice at 12:00 and once at 13:30 while recording the images. The view sequence was generated from the images recorded firstly at 12:00. The threshold for sparsely sampling the images for generating the view sequence was manually adjusted in order that the number of views in the original method was similar to the number in the proposed method. In detail, we set the threshold for the normalized correlation in the original method to be $0.9$ and the threshold for the depth distance to be $1.8$ [pixel]. In the proposed method, the number of small patches in the ABM method was 25 and the size of the search area after the ABM method was $5 \times 5$ [pixel]. The simple moving average with 20 frames was used as low-pass filter for the depth distance of the image. 22 views in the original method and 24 views in the proposed method were obtained.

We verified if the target views were correctly changed when using the images recorded second at 12:00 and recorded at 13:30 as the current views. Figs. 8 and 9 show the results of the original method. Figs. 10 and 11 show the results of the proposed method. The numbers in these figures

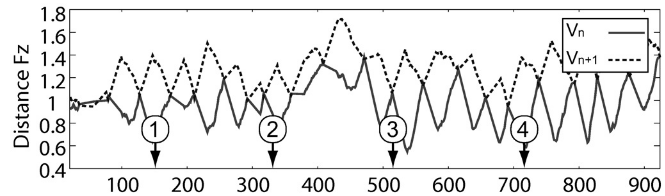



(a)

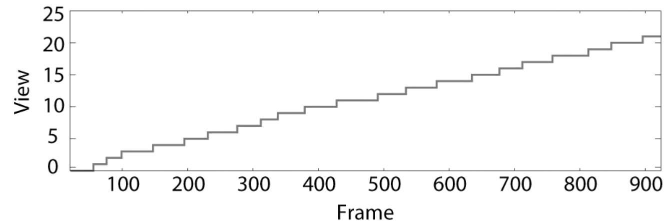



(b)

Fig. 9. (a) Normalized correlations of the target view $V_n$ and the next view $V_{n+1}$ to the current image, and (b) the selected view ID in each frame at 13:30. Despite arriving at the goal, the selected view is not the final one.

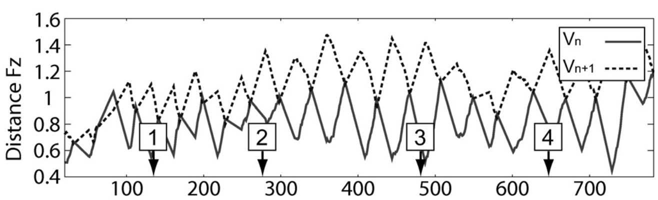



(a)

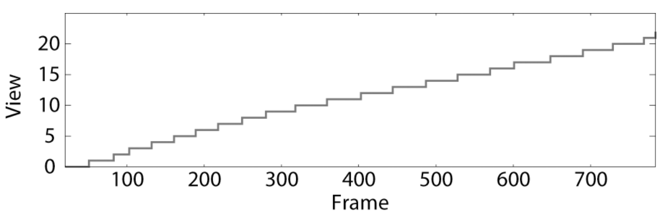



(b)

Fig. 11. (a) Depth distances of the target view $V_n$ and the next view $V_{n+1}$ to the current image, and (b) the selected view ID in each frame at 13:30. Although the original navigation method could not appropriately change the target views, the proposed method can do.




(a)




(b)

Fig. 10. (a) Depth distances of the target view $V_n$ and the next view $V_{n+1}$ to the current image, and (b) the selected view ID in each frame at 12:00. The target views are appropriately changed.

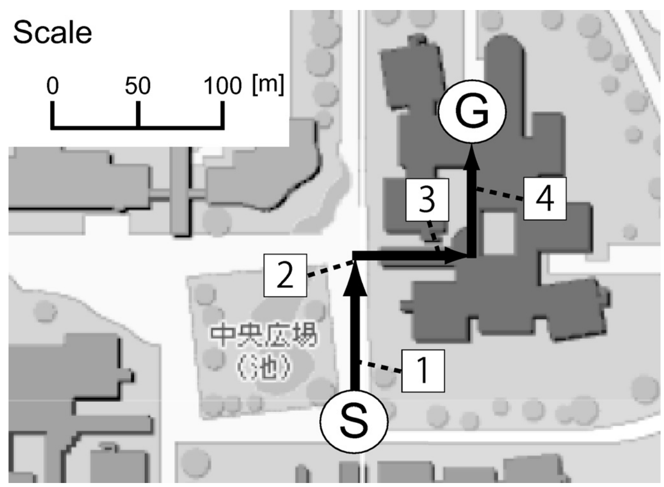



Fig. 12. Map of indoor/outdoor environments and robot navigation path. $S$ and $G$ denote the start and the goal and the arrows indicate the path.

correspond to the location IDs in Fig. 7.

Fig. 8 (a) shows the normalized correlations of the target view $V_n$ and the next view $V_{n+1}$ to the current image $I$ in the original method. As the robot approaches the position where the view $V_{n+1}$ was captured, the correlation of the view $V_n$ decreases and that of the view $V_{n+1}$ increases. The target view is appropriately changed as shown in Fig. 8 (b); the final view in the view sequence is selected at the goal.

As shown in Fig. 10 (a), the depth distance of the view $V_n$ increases and the distance of the view $V_{n+1}$ decreases in the proposed method as the robot approaches the position. The target view is appropriately changed as shown in Fig. 10 (b).

The original method could not appropriately change the target view for the images recorded at 13:30 as shown in Fig. 9. The normalized correlations of the view $V_n$ and the view $V_{n+1}$ decrease together. However, the proposed method can appropriately change the view as shown in Fig. 11. We verified the effectiveness of the proposed method using the images recorded at 16:00. From the results, we conclude that the proposed method is more robust for illumination changes than the original method.

### C. Indoor/outdoor navigation

We conducted experiments for navigation outdoor and indoor (the first floor in the building of information science) in our campus. The distance covered was about 300 [m] and the weather was fine. Although there is an automatic door at the entrance of the building, the door was turned off and left open during the experiment. We set the thresholds for the depth and the horizontal distances to be 1.2 [pixel], and 10 [pixel], respectively. The view sequence, which

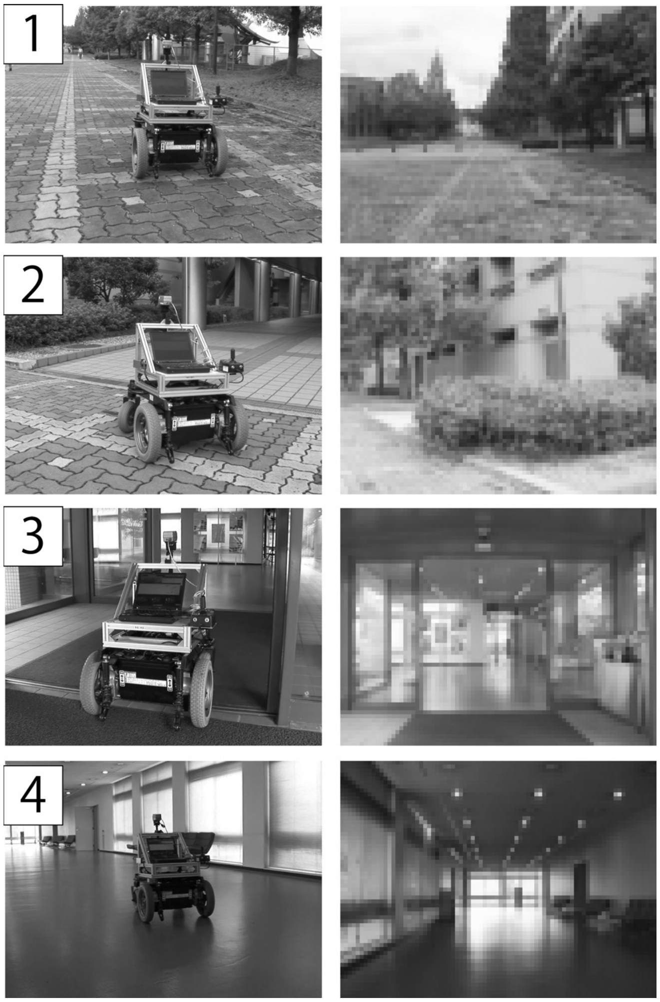



Fig. 13. Snapshots during the navigation (left) and the corresponding views from the robot (right)

includes 95 views, was generated from the images recorded at 9:00. The navigation was performed at 12:00. $S$ and $G$ in Fig. 12 denote the start and the goal, and the arrows indicate the path. Fig. 13 shows snapshots during the navigation and the corresponding views from the robot. The numbers in this figure correspond to the location IDs in Fig. 12. Through the navigation, we verified that the proposed method appropriately changed the target views independent of the indoor/outdoor environments and succeeded in the navigation including rotating motion. The proposed method achieved indoor/outdoor navigation without using tricks that depend on the type of navigation environment.

## V. Conclusion

In this paper, we proposed a view sequence navigation method which can adapt to the environment with illumination changes for achieving seamless indoor/outdoor navigation. The original navigation method changes the target view based on the normalized correlation of the template matching method. However, due to the intrinsic drawbacks of the matching method, it is difficult to apply the navigation method to the outdoor environment where illumination condition changes every moment. To solve this problem, we first proposed to use the accumulated block matching (ABM) method, which is robust against occlusions. Next, we defined novel criteria for changing the target view, *i.e.*, the horizontal and the depth distances of the images. Then, we actually designed the navigation method using them. The method navigates the robot by moving forward and rotating. We demonstrated that the proposed method was able to achieve indoor/outdoor navigation without using tricks that depend on the type of navigation environment.

The robustness to illumination changes causes erroneous navigation of the robot in the case where obstacles exist on the path, (*e.g.*, accident caused by hitting an obstacle). The reason is that the proposed method does not consider whether small changes on the image are caused by illumination changes or by obstacles. To realize safe and reliable navigation requires us to implement the algorithm for detecting obstacles from images or to use additional types of sensors, such as a laser range sensor.

## References

[1] E. Abbot and D. Powell. Land-vehicle navigation using gps. In *Proc. of the IEEE*, pages 145–162, 1999.

[2] W. Burgard, A. B. Cermers, D. Fox, D. Hahnel, G. Lakemeyer, D. Schulz, and W. Steiner. Experiences with an interactive museum tour-guide robot. *Artificial Intelligence*, 114:1–2, 1999.

[3] H. Katsura, J. Miura, M. Hild, and Y. Shirai. A view-based outdoor navigation using object recognition robust to changes of weather and seasons. In *Proc. of IEEE Int. Conf. on Intelligent Robots and Systems*, pages 2974–2979, 2003.

[4] S.-H. Kim, C.-W. Roh, S.-C. Kang, and M.-Y. Park. Outdoor navigation of a mobile robot using differential gps and curve detection. In *Proc. of IEEE Int. Conf. on Robotics and Automation*, pages 3414–3419, 2007.

[5] V. Kulyukin, C. Gharpure, J. Nicolson, and S. Pavithran. RFID in robot-assisted indoor navigation for the visually impaired. In *Proc. of IEEE Int. Conf. on Intelligent Robots and Systems*, pages 1979–1984, 2004.

[6] Y. Matsumoto, M. Inaba, and H. Inoue. Visual navigaion using view-sequenced route representation. In *Proc. of IEEE Int. Conf. on Robotics and Automation*, pages 83–88, 1996.

[7] K. Ohno, T. Tsubouchi, B. Shigematsu, S. Maeyama, and S. Yuta. Outdoor navigation of a mobile robot between buildings based on dgps and odometry data fusion. In *Proc. of IEEE Int. Conf. on Robotics and Automation*, pages 1978–1984, 2003.

[8] F. Saitoh. Robust image matching for occlusion using vote by block matching. *IEICE Trans. on Information and Systems*, J84-D-2(10):2270–2279, 2001.

[9] H. Saji and K. Mitami. Template matching by using variable size block division. *IEICE Trans. on Information and Systems*, J88-D-2(2):450–455, 2005.

[10] K. Takeuchi, J. Ota, K. Ikeda, Y. Aiyama, and T. Arai. Mobile robot navigation using artificial landmarks. *Trans. of the Japan Society of Mechanical Engneers. C*, 66(647):2239–2246, 2000.

[11] S. Thrun, M. Beetz, M. Bennewitz, W. Burgard, A. B. Cremers, F. Dellaert, D. Fox, D. Hähnel, C. Rosenberg, N. Roy, J. Schulte, and D. Schulz. Probabilistic algorithms and the interactive museum tour-guide robot minerva. *Int. J. of Robotics Research*, 19(11):972–999, 2000.

[12] T. Yoshida, A. Ohya, and S. Yuta. Autonomous mobile robot navigation using braille blocks in outdoor environment. *J. of the Robotics Society of Japan*, 22(4):469–477, 2004.