INDOOR-OUTDOOR SEAMLESS POSITIONING AND IMAGE-BASED CAMERA POSITION ADJUSTMENT FOR INFRASTRUCTURE INSPECTION UAV

Kazuha Saito[1], Masafumi Nakagawa[1] and Yusuke Kawasaki[2], Masaaki Takebayashi[2], Hirotoshi Kurashige[2], Shozo Nishimura[2], Masafumi Miwa[3]

[1] Shibaura Institute of Technology, 3-7-5 Toyosu, Koto-ku, Tokyo, 135-8548, Japan
[2] Keisoku Research Consultant CO., 1665-1 Fukuda, Higashi-ku, Hiroshima-shi, Hiroshima, 732-0029, Japan
[3] Tokushima University, 2-1 Minamijousanjima-cho,Tokushima, 770-8506,Japan
Email: ah18037@shibaura-it.ac.jp

KEY WORDS: UAV, GNSS, Visual Odometry, Stereo Camera, Indoor–Outdoor Seamless Positioning.

ABSTRACT: Unmanned aerial vehicles (UAVs) have been introduced recently for the inspection of structures such as bridges and dams because they enable safe and low-cost inspections from a variety of viewpoints and angles. In particular, high-resolution images taken from gimbal-mounted cameras on the UAVs can be used to detect physical changes in the structures. However, autonomous UAV flight is usually made possible by using global navigation satellite system (GNSS) positioning, implying that seamless indoor–outdoor flight cannot be realized in non-GNSS environments, such as the interior subspace of bridges. To be able to use inspection imagery, accurate measurement of each shooting position and direction is required, but any errors in positioning and angle measurement will cause misalignment when combining image sequences. Therefore, an indoor–outdoor seamless positioning system for UAVs is required, in addition to an automatic adjustment system for the angle of view to mitigate the misalignment between images taken in sequence. In this study, we develop an indoor–outdoor seamless positioning system that integrates real-time kinematic GNSS positioning and visual odometry (VO). We also propose a method for correcting the accumulated errors caused by VO and a method for calculating position correction values from captured images for automatic adjustment of the angle of view. We have evaluated the proposed methodology via walking measurements and simulation experiments using a digital camera.

1. INTRODUCTION

The integrity of social infrastructure is important for the safety of the public and for socioeconomic activity. This implies appropriate management and maintenance after the initial construction. However, infrastructure inspection work is usually time-consuming and labor-intensive because it involves many inspections over an extensive area. To address this problem, UAVs have been adopted for use as efficient 3D measurement platforms for the inspection of bridges and dams because they are able to inspect such structures from various angles using gimbal-mounted cameras. However, global navigation satellite system (GNSS) positioning is difficult in spaces under structures such as bridges where the GNSS signal may be poor or even absent. In infrastructure inspection work, it is necessary to capture high-resolution images while the UAV is in flight, together with the shooting position and orientation of each image. However, when sequences of images are combined, misalignments occur if there are errors in the shooting position and angle measurements. Therefore, we first need an indoor–outdoor seamless positioning system that enables UAVS to inspect structures such as bridges, dams, and tunnels in environments where GNSS positioning is inadequate. We also need a camera positioning system that can mitigate any misalignments between images taken at different times. In this study, we used real-time kinematic (RTK)-GNSS receiver and antenna as an outdoor positioning system. RTK-GNSS is a positioning methodology to correct the gap in positioning information of single-positioning by acquiring GNSS position information at two points, BASE and ROVER, and using the difference between them as correction information. We also perform visual odometry (VO) processing using an IMU stereo camera as an indoor positioning system. VO is an aspect of simultaneous localization and mapping (SLAM) processing. We develop an indoor–outdoor seamless positioning system that integrates RTK-GNSS positioning and VO that enables UAVs to fly across GNSS and non-GNSS positioning environments while capturing seamless image sequences.

First, we propose a methodology to address the cumulative error problem caused by VO and a methodology to calculate positional corrections from captured images. Next, we provide an overview of our experiments. In our walking measurement experiment for indoor and outdoor spaces, we investigate whether fast switching between positioning modes is possible by using the positioning mode type of NMEA data acquired by RTK-GNSS positioning. (“NMEA” refers to the National Marine Electronics Association, USA, which developed this standardized data format.) We then summarize and discuss our acquired data and experimental results. We investigated whether our methodology could compensate for cumulative errors by improving the relocalization process in VO. We also investigated whether the positioning accuracy indoors depends on the camera direction and time variations within the image sequence and give an overview of these experimental results. In addition, we investigated whether our methodology could detect feature points by using the speeded-up robust features (SURF) algorithm in a camera simulation experiment with the corrected captured positions.

2. METHODOLOGY

2.1 Indoor-outdoor seamless positioning system

The methodology proposed in this study is outlined in Figure 1. First, we can determine the GNSS and non-GNSS positioning environments by checking the positioning status described in the NMEA data acquired by RTK-GNSS positioning. Next, we start taking measurements from the point where the RTK-GNSS positioning acquires status = Fix and switch the positioning mode according to the determined indoor–outdoor space. If RTK-GNSS positioning status = Fix, we receive positional data from the RTK-GNSS receiver system and perform a transformation of these WGS84 coordinates to plane Cartesian coordinates. ("WGS84" refers to the World Geodetic System (1984) standard for such data.) If the RTK-GNSS positioning status = Float, then the VO position, velocity, and azimuth data are acquired. Whenever we move from an outdoor environment to an indoor environment, we also generate rotation and translation data using the last acquired status = Fix positional data as the origin and apply the VO results. Then, we regenerate the NMEA data from the position, velocity, and azimuth data by correcting the applied results via the adjustment control for the accumulation errors.

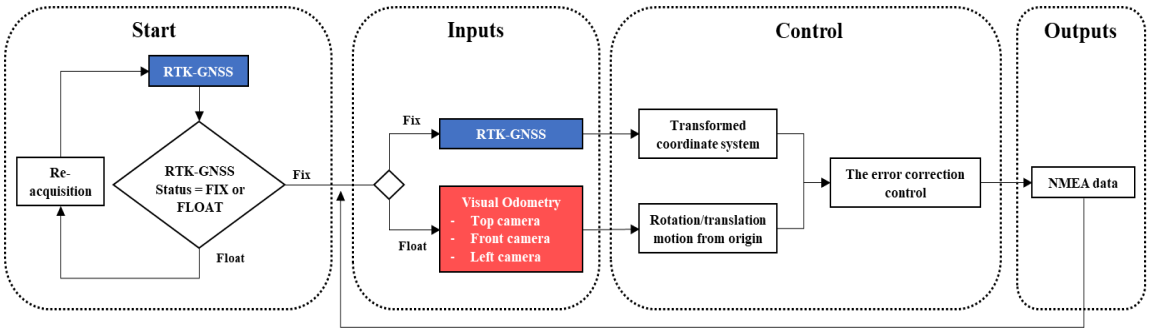

Figure 1. Indoor-outdoor seamless positioning system methodology

2.1.1 Changing the positioning mode

We designed a method to achieve seamless transition between an indoor positioning system that uses VO and an outdoor positioning system that uses RTK-GNSS positioning. The proposed methodology uses the RTK-GNSS positioning status, such as RTK-Fix, RTK-Float, and no GNSS signals, to determine the positioning modes, as shown in Figure 2. When a UAV moves from a GNSS environment to a non-GNSS environment, the position and orientation are first obtained by RTK-GNSS positioning with RTK-Fix status. The status then changes from RTK-Fix to RTK-Float for the non-GNSS environment. During the status change, VO takes an initial position and orientation from the GNSS positioning results, and positioning is continued via VO. When the UAV moves from a non-GNSS environment to a GNSS environment, the position and orientation is obtained by RTK-GNSS positioning according to the timing of the status change from RTK-Float to RTK-Fix.

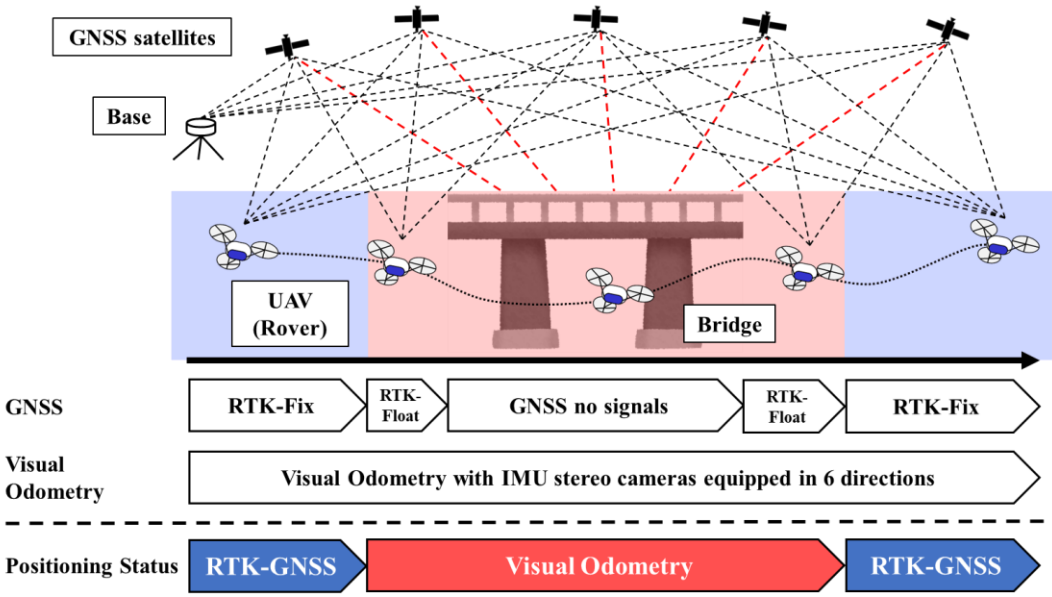

Figure 2. Changing positioning mode

2.1.2 Correction control of accumulation errors

We now describe our proposed method for correcting accumulation errors. An accumulation error is a technical problem that occurs when switching the RTK-GNSS positioning status between GNSS and non-GNSS environments, where the closing differences do not match. This is caused by an accumulation of image matching errors in VO. Generally, this problem can be solved by loop closure. Loop closure refers to the process of identifying and closing loops in the path or trajectory of a mobile robot or camera system as it moves through an environment. However, error correction by loop closure causes rapid position changes during UAV flight, as the UAV moves from the erroneous position to the corrected position at high speed. Therefore, this approach is not suitable for real-time and stable UAV flight, and in the worst-case scenario, the UAV may become uncontrollable and crash. In this study, we detect results above 6–12 km/h centered on the estimated position at which errors have accumulated and calculate the corrected position. After the correction amount is calculated, the acquired data is corrected step by step to avoid drastic changes in position and orientation when the RTK-GNSS positioning status is switched. In this way, we can provide real-time stability in the UAV flight.

2.2 Image-based camera position adjustment system

Figure 3 shows the methodology proposed in this study for adjusting the camera position. Our methodology provides a reproduction of the same captured angle of view for the captured object by adjusting the camera position through image matching between the first captured image and the second captured image. First, the original image and the corresponding image taken 2 m from the object are used as input data. Next, the positions of the corresponding feature points are obtained from each of the two images by detecting and extracting feature points using the SURF algorithm. Finally, the corresponding images are superimposed on the original image using the matched correspondence points (without outliers), and the camera position is adjusted to reproduce the captured angle of view. In this study, we assume that, at a shooting position of 2 m from the object, the image will involve an error of approximately 10–20 pixels at most, considering the rotation of the gimbal camera. We also apply the proposed methodology to features accelerated segment test (FAST), scale invariant feature transform (SIFT), and binary robust invariant scalable keypoints (BRISK) feature detection for comparison with SURF.

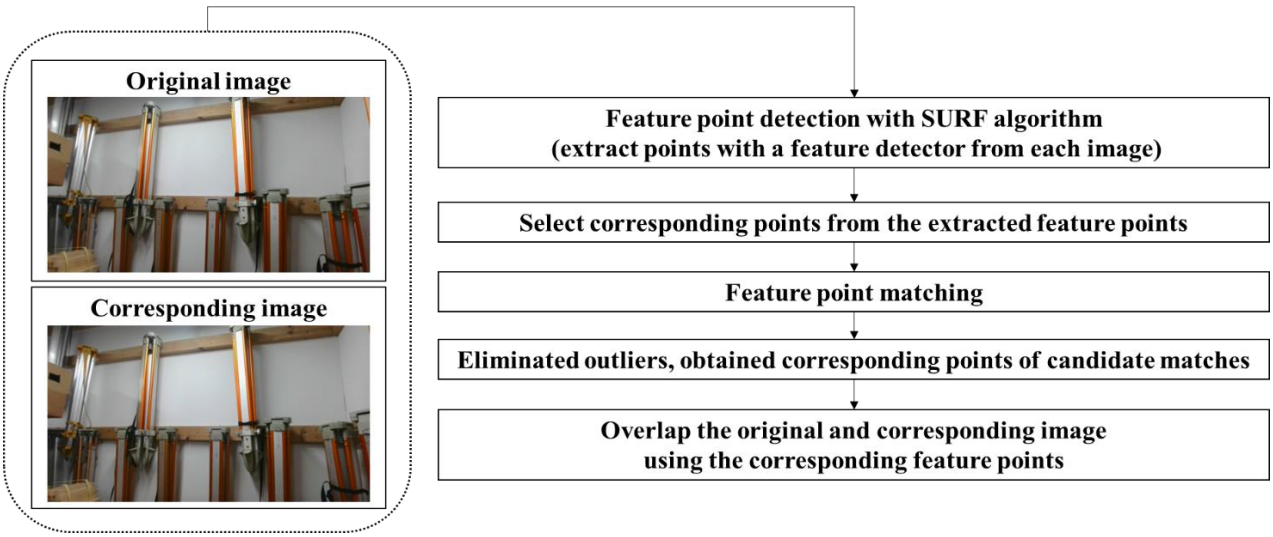

Figure 3. Image-based camera-position adjustment system methodology

3. EXPERIMENTS

3.1 Walking measurement experiment

In this study, we conducted two types of experiments to validate our proposed methodology. First, we conducted a walking measurement experiment to verify the positioning mode switching between RTK-GNSS positioning and VO in indoor and outdoor environments using RTK-GNSS positioning status (see Figure 4). We used a 2D robotic “rover” to simulate the 3D UAV. We set up a study path through both indoor and outdoor environments to investigate how the camera images of the indoor environment changed at Door 1 and Door 2. The RTK-GNSS receiver and antenna comprised a ZED-F9P (u-blox) and GNSS antenna, as shown in Figure 5. We also equipped the rover with three inertial measurement unit (IMU) stereo cameras (front, left, and top) to cover the three directions, as shown in Figure 6. We selected the RealSense T265 (Intel) as the IMU stereo camera. The RTK-GNSS positioning and VO positional data were sampled at a rate of 5 Hz.

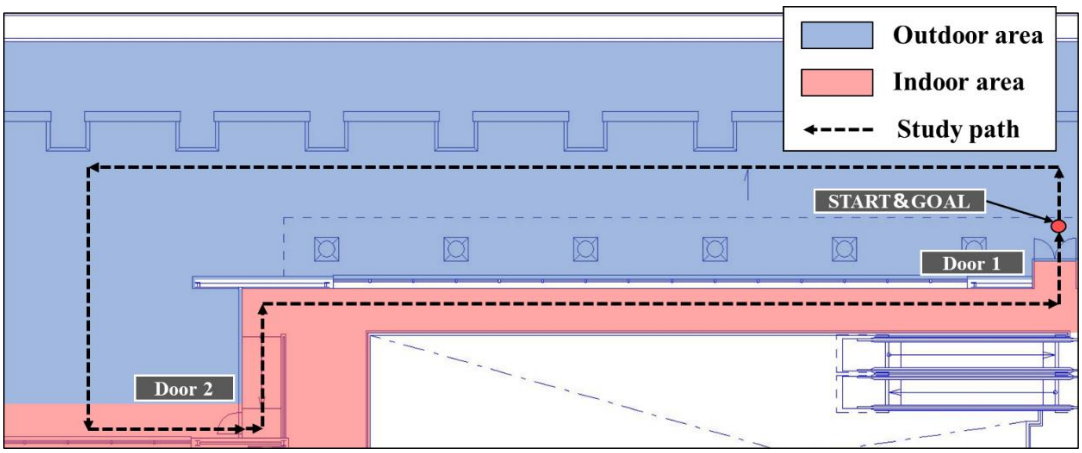

Figure 4. Study path

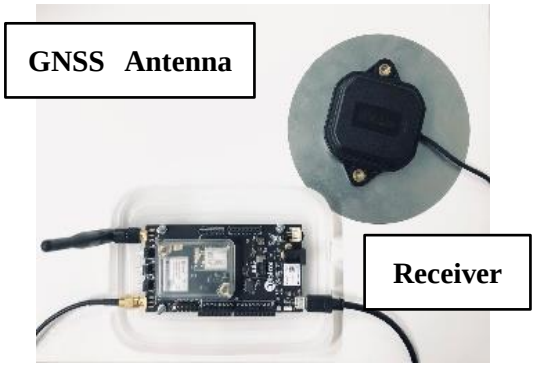

Figure 5. RTK-GNSS receiver and antenna

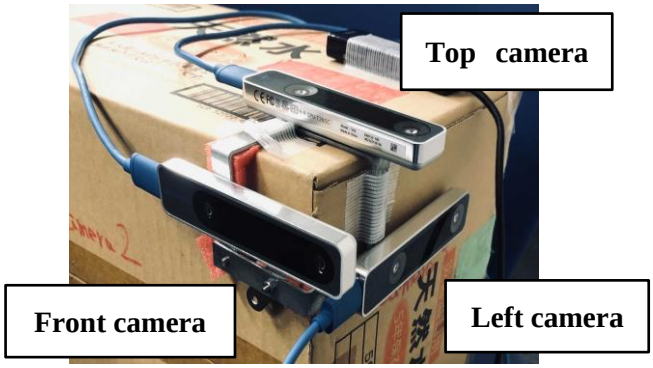

Figure 6. IMU stereo cameras

3.2 Simulated camera position experiment

We conducted a simulated camera-position experiment to evaluate our methodology using a digital camera and the camera-position adjustment system based on the captured images proposed in this study. We assumed that the required posture accuracy in the indoor–outdoor seamless positioning environment proposed in this study would be about ±3 degrees. We used a digital camera (DSC-HX60V, SONY) with a resolution of approximately 21.1 megapixels as our experimental equipment, as shown in Figure 7. The camera's angle of view was fixed at 50 degrees.

Figure 7. Appearance of DSC-HX60V (SONY) and study area

4. RESULTS

4.1 Walking measurement experiment results

In the walking measurement experiment, we integrated the acquired data between RTK-GNSS positioning and VO to output good trajectory data and NMEA data. Figure 8 shows the trajectory data output using our proposed methodology. Figure 9 shows the regenerated NMEA data. We confirmed that the position and orientation estimation using the camera following the flight direction represented the most similar trajectory to that of the actual trajectory. Therefore, the main camera selection must adapt to the flight directions if the VO performance of the UAV flight is to be improved.

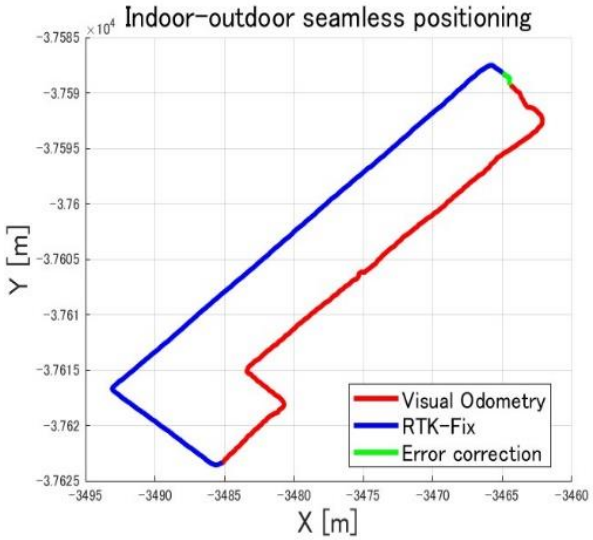

Figure 8. Trajectory data (2D plot)

Figure 9. NMEA data (Google map)

4.2 Simulated camera position experiment results

Figure 10 shows the results of detecting the corresponding points from two camera images and the results of translating and superimposing the corresponding images onto the original image. We confirmed that there were some outliers among the detected correspondence points and calculated the average value by removing the outliers and extracting only the valid vectors. We compared the camera images before and after applying the proposed methodology and confirmed that the camera position can be adjusted by superimposing the corresponding feature points.

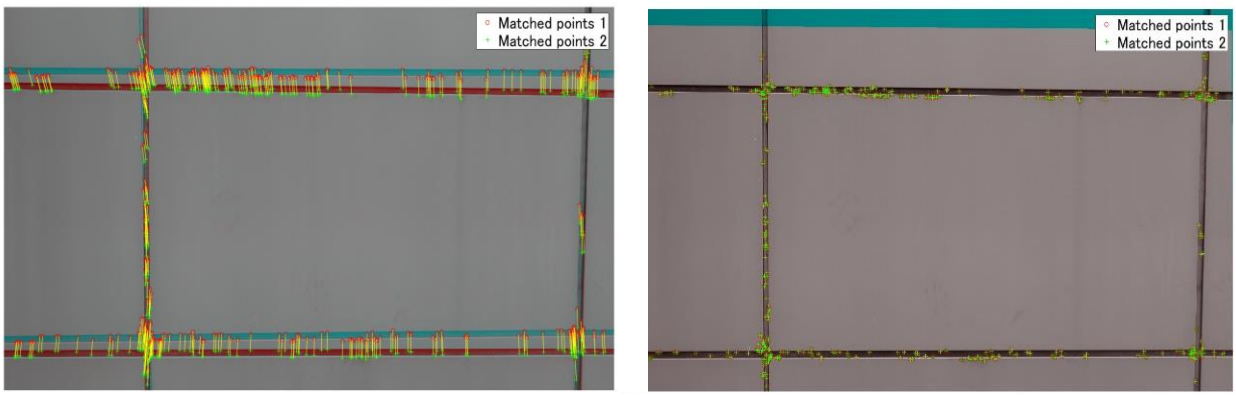

Figure 10. Camera image processing results
(Left image: correspondence point detection, right image: correspondence point superimposition)

5. DISCUSSION

5.1 Walking measurement experiment discussion

In this study, we focused on the performance improvement of VO with multidirectional image acquisition using IMU stereo cameras. We determined the center position of a pseudo-circle from the median of the data acquired by the IMU stereo cameras (top, front, and left) and calculated the average value of the positional data within the circle to output the corrected data. In repeated walking measurement experiments, we confirmed that the radius of the circle should be approximately 3–5 cm to remove outliers and to output stable positional data.

We confirmed that the use of VO depends on the number of errors of closure (Table 1). Errors of closure refer to the difference between the start and end points of the walking measurement experiment. The camera directions and distances from the camera to objects are related to scene changes in sequenced camera images. Therefore, rapid turns and image acquisition from closed points will affect the accuracy of VO. Compared with the front camera, the left camera and the top camera captured more sky and window glass, making it difficult to detect feature points. In this case, we set the front camera to be the camera in the movement direction, which resulted in the smallest error in the closed position. We can therefore confirm that the camera direction is one of the factors in improving measurement accuracy for position estimation.

Table 1. Error of closure and ratio of closure

	Error of closure	Ratio of closure	X error [m]	Y error [m]
Corrected data	0.993	1/101	0.937	-0.329
Acquired data (front)	1.090	1/92	1.072	-0.195
Acquired data (left)	1.360	1/74	-1.284	-0.449
Acquired data (top)	6.132	1/16	-4.568	-4.091

We compared the position changes before and after the application of the correction control of the accumulated error when switching positioning modes, as shown in Figure 11. For the pre-correction results, an accumulated error in position and orientation estimation was observed, and a different position was estimated from that for the study path. We also confirmed that the position changes at each frame increased because of the sampling-rate difference between RTK-GNSS positioning and VO, when the positioning mode switches from a non-GNSS environment to a GNSS environment. This confirms that loop closure was caused by incorrect tracking because of a significant reduction in the number of correspondence points tracked from the feature checkout. We applied the proposed methodology to detect movement results above 6–12 km/h in a non-GNSS environment and confirmed that the gradual correction at the loop closure point avoids abrupt position changes for each frame. We also confirmed that a threshold value below one meter is required, from consideration of the accuracy of position and posture estimation and the average moving speed. In this study, we note that the sampling rate of indoor–outdoor seamless positioning should be adjustable for VO to conduct feature point detection by camera-position estimation.

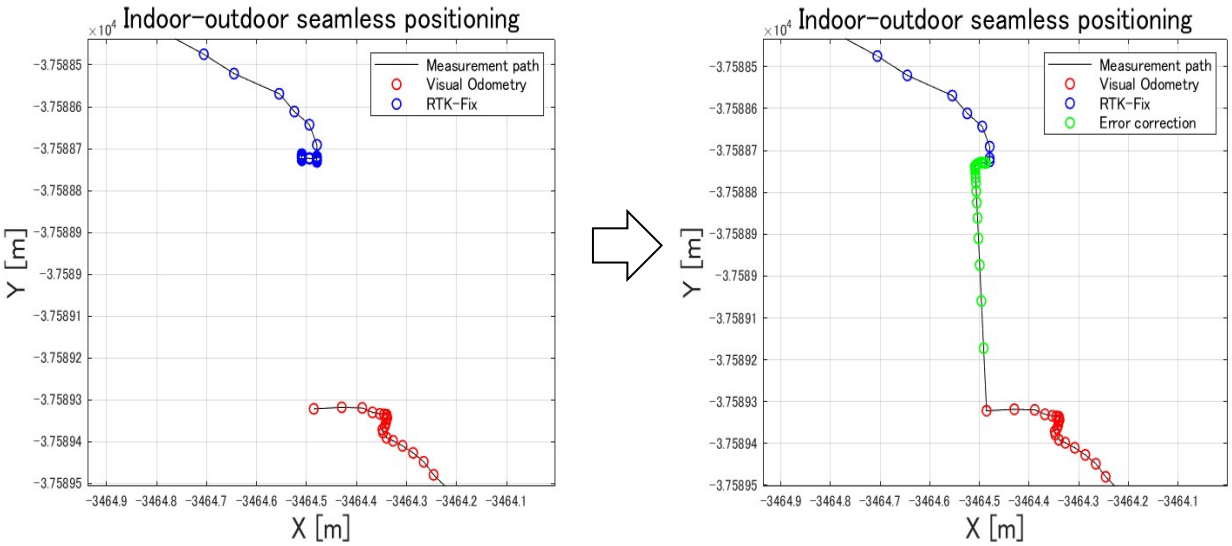

Figure 11. Correction control when switching positioning modes
(Left image: readjusted data without the proposed methodology,
right image: readjusted data with the proposed methodology)

We also conducted measurements along the previous experiment's study path to improve the accuracy of position and posture estimation in indoor positioning based on the relative relationship between the camera direction and the direction of travel. First, we performed a gait measurement by aligning the camera direction with the direction of motion (Setup 1). Next, gait measurement was performed with the camera direction not aligned with the direction of travel but fixed at a particular direction (Setup 2). The relocalization error for Setup 1 was 1.473 m, and that for Setup 2 was 0.875 m. We could therefore confirm the superiority of Setting 2 with respect to positioning accuracy. This also shows that highly accurate indoor positioning can be achieved by fixing the camera direction and constantly switching the main camera to match the direction of travel.

5.2 Simulated camera position experiment discussion

We evaluated the number of feature points detected on the concrete wall using representative algorithms for feature inspection such as FAST, SURF, SIFT, and BRISK. The feature points, corresponding points, processing time, and matching rate for the algorithms are shown in Table 2. Comparing the algorithms, we confirmed that the SURF algorithm was the most capable of detecting feature points for concrete walls, in terms of feature point detection, extraction, and processing time. For the camera-position simulation experiment, both images were taken from the front of the object, so

the proposed methodology was reapplied to the corresponding image and the original image taken from different angles of view. As a result, we confirmed that our methodology cannot accurately superimpose the camera images if the feature characteristics change because of changes in the background of the object, because our methodology uses the SURF algorithm for feature point detection.

Table 2. Comparison of detection of feature points (for each algorithm)

	Feature points		Corresponding points		Processing time [s]	Matching rate [%]
	Original image	Corresponding image	Including outliers	Without outliers		
FAST	203	265	24	18	1.55	75.0
SURF	8102	9233	969	376	2.52	39.4
SIFT	1164	1433	116	77	5.21	66.4
BRISK	6154	7585	125	77	5.36	61.6

Next, we evaluated the feature point detection parameters and their matching rates to improve the matching rate for low-texture surfaces. The initial value for the feature point detection parameter of the SURF algorithm was set to 100, and we evaluated the proposed methodology by changing this value. We confirmed that when the threshold for feature points is increased, the number of detected feature points in the camera image decreases and the matching rate increases. Therefore, we note that, when using the SURF algorithm for feature point detection in infrastructure inspection work, it is possible to achieve highly accurate feature point detection by setting appropriate parameters for the object.

6. CONCLUSION

In this study, we have proposed a methodology for correcting the accumulated errors caused by VO and a methodology for adjusting the camera position based on the captured images. We have investigated the proposed methodology via walking measurements and simulated experiments using a digital camera and verified the validity of our approach. In our future work, we intend to develop a system to utilize high-resolution images taken by UAVs equipped with the proposed methodology as building-information-modeling and civil-information-modeling data.

ACKNOWLEDGMENTS

This research was supported by the METI Monozukuri R&D Support Grant Program for SMEs Grant Number JPJ005698.

REFERENCES

Christopher, J, H., Eric, C., 2008. Evolution of the Global Navigation Satellite System (GNSS). *IEEE*, 18(9), pp.1-16.

Mostafa, M., Shady, Z., Adel, M., Naser, E, S. and Abu, S., 2018. Radar and Visual odometry Integrated System Aided Navigation for UAVS in GNSS Denied Environment. *Sensors*, 18(9), pp.13-27.

Shaojie, S., Yash, M., Nathan, M., Vijay, K., 2014. Multi-sensor fusion for robust autonomous flight in indoor and outdoor environments with a rotorcraft MAV. *IEEE*, pp.1-8.

Ke, S., Kartik, M., Bernd, P., Michael, W., Sikang, L., and Yash, M., 2018. Robust Stereo Visual Inertial Odometry for Fast Autonomous Flight, *IEEE*, pp.965-972.

Faragher, R, M., Harle, R, K., 2013. SmartSLAM - An Efficient Smartphone Indoor Positioning System Exploiting Machine Learning and Opportunistic Sensing. *Nashville*, pp.1006-1019.

Amin, B., Mohammadreza, A, O., Anahid, B., 2018. Particle Filter and Finite Impulse Response Filter Fusion and Hector SLAM to Improve the Performance of Robot Positioning. *Journal of Robotics*, pp.5-8.

Zhang, H., Hu, Q., 2011. Fast image matching based-on improved SURF algorithm. *IEEE*, pp.1-4.

Nabeel, Y, K., Brendan, M., Geoff, W., 2011. SIFT and SURF Performance Evaluation against Various Image Deformations on Benchmark Dataset. *IEEE*, pp1-6.