Riverbed Depth Measurement Method Combining Visual SLAM-Based Camera Pose and Ultrasonic Slant Range

Norihito Kishi[1]

[1]Civil Engineering Research Institute for Cold Region, Japan
kishi-n22aa@ceri.go.jp

Abstract –

Heavy rains and flooding due to global warming are becoming more frequent, increasing the risk of scouring and riverbed deformation. Low-cost and simple techniques can be effective for detecting these damages at an early stage, however, water currents, turbidity, and waves make surveys difficult. Therefore, we devised a method to measure riverbed depth with spatial distribution by combining the device pose estimated by Visual SLAM with the slant range measured using underwater ultrasonic. This paper describes the results of land-based tests aim to integrate pose data and slant range data.

Keywords –

Riverbed Depth; Visual SLAM; Camera Pose; Ultrasonic; Slant Range

1 Introduction

In recent years, the frequency of heavy rainfall events has increased due to global warming in Japan [1]. Heavy rainfall that causes river flooding increases the risk of scouring around bridge piers and riverbed deformation, which is a global problem [2]. To mitigate these damages, it is important to understand the current conditions, and various survey techniques have been developed. For example, ALB (Airborne LiDAR Bathymetry) can efficiently survey wide areas with high precision, but it faces issues such as susceptibility to turbid water, the need for skilled operation, and high equipment costs [3]. A method that combines a remote-controlled boat and sonar, as in [4], can safely survey wide areas, but there are issues such as limited maneuverability in fast-flow environments. Many small, inexpensive remotely operated vehicles have been sold, and their use for infrastructure inspections has become more widespread, but underwater surveys in rivers with currents and turbidity remain difficult. A simpler and cheaper method is a float-type sonar with GNSS. In Japan, the Ministry of Land, Infrastructure, Transport and Tourism is working on applying it to disaster surveys, and its low cost and ease of use compared to other technologies are considered factors in its adoption. Furthermore, as an advanced version of this, research has been conducted on tethering the sonar using an unmanned aerial vehicle for bathymetry observation of inland water bodies [5]. We have also used a similar sonar and confirmed that it can measure water depth with an error of a few centimeters in flatwater. However, there are issues such as inaccurate depth measurement caused by surface waves, and poor GNSS reception near bridges or other structures.

For these reasons, we considered that a survey technique with "precision and data density comparable to float-type sonar, inexpensive, simple, and unaffected by water currents and waves" would be useful. Therefore, we devised a method to measure riverbed depth with spatial distribution by combining device pose estimated by Visual SLAM (Simultaneous Localization and Mapping) with slant range measured by underwater ultrasonic [6]. This paper describes the results of land-based tests that aim to integrate trajectory (a time-series of poses) and slant range data.

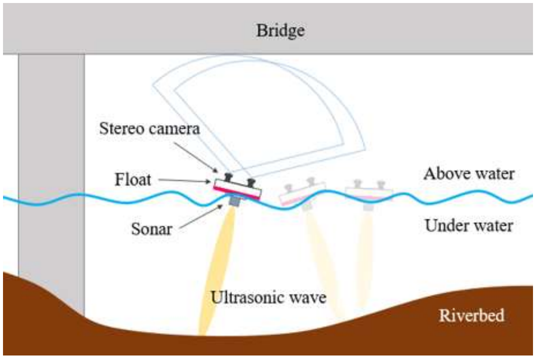

Figure 1. The schematic of proposed method

2 The Overview of Proposed Method

As mentioned above, the sway of device caused by water currents and waves makes surveys difficult. Therefore, we devised a method to measure riverbed depth spatially by swinging the ultrasonic beam transmitted from the sonar in multiple directions, taking advantage of the device's sway, and combining the device pose with underwater slant range data. A schematic of the proposed method is shown in Figure 1.

The measurement device is expected to have a simple configuration consisting of a stereo camera mounted on the top of the float, a SBC (Single Board Computer) inside the float, and a sonar attached to the bottom of the float. When surveying areas near structures such as under bridges, GNSS reception may be poor, but the structure is likely to be within the camera's field of view. Figure 2 shows an example of the underside of a bridge, which contains many feature points. For this reason, we considered that by specifically targeting environments containing structures that provide sufficient visual features, the camera pose obtained by Visual SLAM could serve as the device's pose.

Figure 2. An example of the underside of a bridge

Figure 3. The measuring unit

Table 1. Devices

Device	Specification
SBC	Raspberry Pi 5, 8GB
Stereo camera	ELP-USB3DGS1200P01-H120, angle of view 120 deg
Ultrasonic distance sensor	HC-SR04

3 Measuring Unit and Analysis

3.1 Measuring Unit

The measuring unit and devices are shown in Figure 3 and Table 1. The camera was fixed to the top of the SBC with its optical axis facing vertically upward, intended to capture images of the underside of the bridge. The ultrasonic distance sensor was fixed to the underside of the SBC with its axis facing vertically downward, intended to measure the riverbed depth. Since this was a land-based test to verify the feasibility of the method, we used an airborne ultrasonic distance sensor instead of an underwater sonar.

3.2 Measurement and Analysis

The SBC was used for data acquisition only, and the SBC and smartphone were connected via Wi-Fi, allowing the measurement application to be launched and operated on the smartphone. The data were post-processed on another laptop PC. The laptop PC ran Ubuntu 20.04, and ORB-SLAM3 [7] was used for Visual SLAM. The stereo_euroc program included in ORB-SLAM3 was used to perform the SLAM processing, and the depth was calculated using formula (1).

$$depth = camera\_pos + R_q \cdot sensor\_dir \cdot sr \quad (1)$$

where
camera_pos is the vector of camera position,
R_q is the rotation matrix obtained from the camera orientation quaternion $[q_x, q_y, q_z, q_w]$,
sensor_dir is the vector of the ultrasonic sensor relative to the camera coordinate system,
as shown in equation (2), (3) and (4), respectively, and *sr* is the slant range.

$$camera\_pos = \begin{bmatrix} p_x \\ p_y \\ p_z \end{bmatrix} \quad (2)$$

$$R_q = \begin{bmatrix} 1-2(q_y{}^2+q_z{}^2) & 2(q_xq_y-q_zq_w) & 2(q_xq_z+q_yq_w) \\ 2(q_xq_y+q_zq_w) & 1-2(q_x{}^2+q_z{}^2) & 2(q_yq_z-q_xq_w) \\ 2(q_xq_z-q_yq_w) & 2(q_yq_z+q_xq_w) & 1-2(q_x{}^2+q_y{}^2) \end{bmatrix} \quad (3)$$

$$sensor\_dir = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} \quad (4)$$

For camera calibration, OpenCV's pinhole camera model [8] was used.

4 Position Accuracy Validation

4.1 Experimental Conditions and Setup

In order to evaluate this measuring unit, we conducted simple tests. The schematic of the experimental setup is shown in Figure 4. The measuring unit was moved between Y = 0.00 m and Y = 1.00 m along the rail while recording images and slant range data. The ultrasonic distance sensor transmitted vertically downward so that the distance to the floor were measured. The distance between the sensor and the floor measured with a steel tape, was 0.77 m.

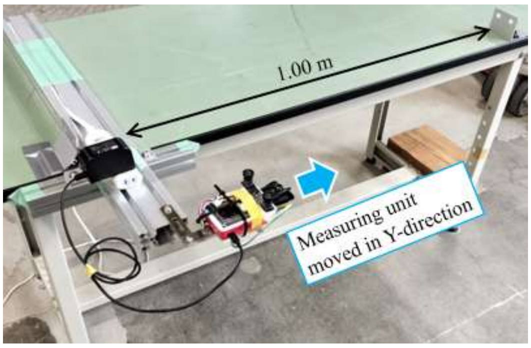

Figure 4. The experimental setup

4.2 Experimental Result

The position estimated by ORB-SLAM3 is shown as the red line in Figure 5. The actual traveling distance in the Y direction was 1.00 m, while the estimated distance was 0.898 m. The actual change in the Z direction was 0.00 m, whereas the estimated change was 0.036 m at Y = 0.898. The average absolute trajectory error of Stereo SLAM on the ORB-SLAM3 EuRoC dataset is reported to be 0.084 m [7], and our results in the Y direction in this experiment were similar. On the other hand, the estimated Z value tends to increase with increasing distance from Y = 0.00. This was considered to be due to the use of the pinhole camera model for camera calibration. Therefore, the estimated Z value was corrected using formulas (5) and (6). Coefficients a, b, c, d, e, and f were calculated using the least squares method. In this validation test, a, c, and d were fixed to zero because the measuring unit could not move in the X direction. The corrected Z value is shown as the blue line, with Z = 0.004 at Y = 0.898.

$$Z_{fit} = ax^2 + by^2 + cxy + dx + ey + f \quad (5)$$

$$Z_{corrected} = Z_{estimated} - Z_{fit} \quad (6)$$

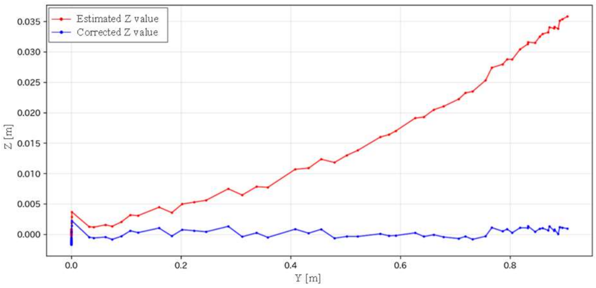

Figure 5. Changes of Z value

5 Depth Measurement

5.1 Experimental Conditions and Setup

This measuring unit was set on a flatbed cart, and depth data (the distance between the ultrasonic distance sensor and the floor) was measured while the cart was moving in a laboratory. The experimental setup is shown in Figure 6. The measuring unit is expected to be used while floating on the water surface with waves. Therefore, we built a stand using an aluminum frame on a cart and suspended the measuring unit from the stand with polyester string and rubber bands to allow it to swing more easily. Marker lines were drawn on the floor of the laboratory as shown in Figure 7, and measurements were taken by moving the cart through points A to F in order, with the measuring unit passing over the mark lines. The floor was made of concrete and was flat but had small irregularities, and while the cart was moving, the measuring unit shook by a few centimeters. The resolution of the input images was 1280 × 720 pixels, with a frame rate of 10 fps. The distance from the bottom of the ultrasonic distance sensor to the floor was measured with a surveying staff and was 1.20 m.

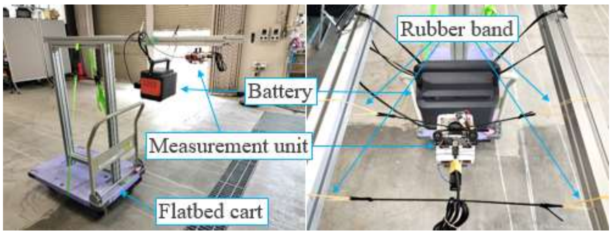

Figure 6. The experimental setup

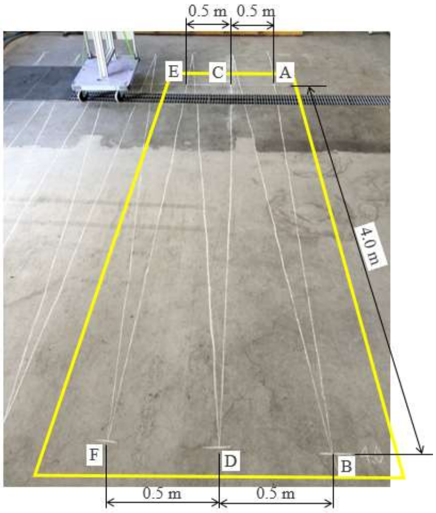

Figure 7. Marker lines

5.2 Experimental Result

The results are shown in Figure 8. The blue lines represent the corrected trajectory of the measurement unit, the red lines represent the slant range from the ultrasonic distance sensor to the floor surface, and the green dots indicate the depth calculated using formula (1). The contour lines generated from the depth data are also displayed. The points A' and F' are the positions of the measuring unit when the cart started and stopped moving, respectively, and points B' to E' are the positions of the measuring unit when the cart changed direction. Each point corresponds to point A to F in Figure 7. There are loop-shaped sections of the trajectory near points C' and E', which occurred when the cart changed direction by turning its wheels, and they closely represent the actual movement. The minor fluctuations in the trajectories represent the sway of the measuring unit. The Z value at point F', the farthest point from the starting point, was 0.01 m, and the distortion was significantly reduced through the correction process. The depth data had an average of -1.17 m, a minimum of -1.21 m, a maximum of -1.14 m, and a standard deviation of -0.01 m.

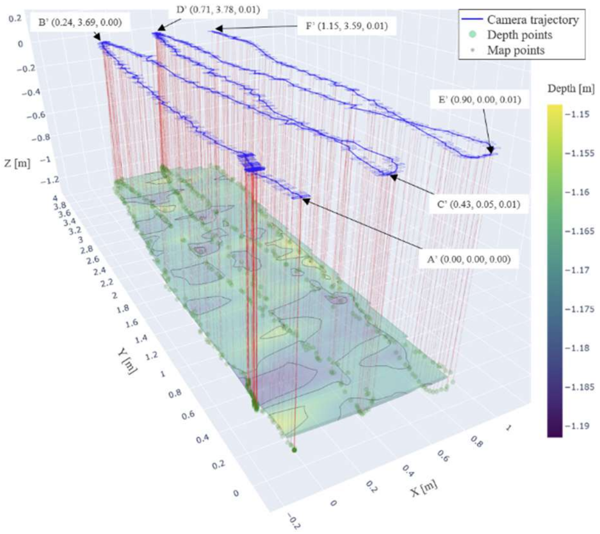

Figure 8. The trajectory and depth points

6 Conclusion

We devised a method to measure the riverbed depth around bridges and other structures in spatially by combining the pose data obtained through Visual SLAM with the underwater slant range data, and verified its feasibility through land-based tests. The result showed that even in images taken when the measuring device is shaking by a few centimeters, pose estimation by SLAM works, and depth can also be calculated from the slant range. Regarding horizontal accuracy, the error was approximately 0.1 m; however, since these results were obtained under limited indoor conditions, further verification is required to assess the generalization of the drift correction over a wider area and under actual bridges. Furthermore, by correcting the distortion of the SLAM output data using a two-variable quadratic polynomial, it became possible to estimate depth on the order of centimeters. However, since factors contributing to errors in the depth data (e.g., pose error, range noise, and mounting geometry) and the effects of actual underwater conditions (e.g., water waves and bottom return characteristics) remain to be investigated, we plan to conduct underwater tests using a underwater sonar to verify these factors. The sampling density along the trajectory provided by Visual SLAM is on the order of several centimeters. However, the effective spatial resolution of the depth map is constrained by the sonar's footprint. For instance, a 7-degree beam width used for a float-type sonar with GNSS, results in a footprint of 0.25 m at a depth of 2 m and 0.61 m at 5 m. Since the SLAM-based sampling density is finer than these footprint sizes, the system's overall horizontal resolution is dominated by the sonar’s beam width.

References

[1] Ministry of Education, Culture, Sports, Science and Technology, Japan Meteorological Agency. Climate Change in Japan 2025. On-line: https://www.data.jma.go.jp/cpdinfo/ccj/2025/pdf/cc2025_gaiyo_en.pdf, Accessed: 04/12/2025.

[2] G. Zhang, Y. Liu, J. Liu, S. Lan and J. Yang. Causes and statistical characteristics of bridge failures: A review. Journal of Traffic and Transportation Engineering (English Edition), 9(3): 388–406, 2022.

[3] R. Tsubaki, T. Moramarco, G. Maniatis, J. L. Coz and M. Muste. Advances in river measurement technologies. New Waves in Hydrometry, Vol.1, 36-59, 2025.

[4] R. Shuchman and G. Meadows. BathyBoat. On-line: https://www.mtu.edu/mtri/research/project-areas/environmental/water/bathyboat/factsheet.pdf, Accessed: 10/12/2025.

[5] F. Bandini, D. Olesen, J. Jakobsen, C. M. M. Kittel, S. Wang, M. Garcia and P. Bauer-Gottwein. Bathymetry observations of inland water bodies using a tethered single-beam sonar controlled by an Unmanned Aerial Vehicle, Hydrology and Earth System Sciences, 22, 4165–4181, 2018.

[6] N. Kishi. Study of Riverbed Depth Measurement Method by Integrating Camera Posture Data Obtained by Visual SLAM and Slant Range Data Obtained Underwater, The 23th Symposium on Construction Robotics in Japan, P1-27, 2025 (in Japanese).

[7] C. Campos, R. Elvira, J. J. Gómez Rodríguez, J. M. M. Montiel and J. D. Tardós. ORB-SLAM3: An Accurate Open-Source Library for Visual, Visual–Inertial, and Multimap SLAM, IEEE TRANSACTIONS ON ROBOTICS, VOL.37, NO.6, pp. 1874 – 1890, 2021.

[8] OpenCV. Camera Calibration. On-line: https://docs.opencv.org/4.x/dc/dbb/tutorial_py_calibration.html, Accessed: 10/12/2025.