A new covariance intersection based integrated SLAM framework for 3D outdoor agricultural applications

Hann-Gyoo Kim,[1] Hea-Min Lee,[2] and Seung-Hwan Lee[1,✉]

[1] *Kumoh National Institute of Technology, Gumi, South Korea*

[2] *IT Application Research Center, Jeonbuk Regional Branch, Korea Electronic Technologies Institute, Jeon-Ju, Jeonbuk, South Korea*

✉ E-mail: leesh@kumoh.ac.kr

This letter introduces a novel integrated framework for simultaneous localization and mapping (SLAM) tailored for general agricultural applications. The framework combines a cutting-edge SLAM method, LIO-SAM, with covariance intersection for sensor fusion. Agricultural robots often operate in unstructured environments with sparse feature points and encounter repeated similar information, such as trees. Therefore, a fusion framework based on 3D SLAM augmented with additional information, such as feature-independent GPS data, becomes essential. This study proposes an integrated SLAM framework by introducing a convergence strategy based on covariance analysis, incorporating a state-of-the-art 3D SLAM technique. The convergence methods, namely "dynamic weight assignment" and "winner takes all", are presented alternatively, tailored to seamlessly integrate with the proposed framework. Evaluations using a public dataset and an experiment demonstrate the effectiveness of the approach through numerical analysis and visual representation. The results illustrate that this method surpasses conventional approaches in accurately estimating the robot's position. In the future, this research will focus on automating crop cultivation and harvesting by integrating the proposed system with robot arm control.

Introduction: Agriculture is rapidly transforming into a technology-intensive sector, thanks to advancements in 3D sensor technology [1]. Particularly, significant progress is being made through the integration of robotic technology with outdoor simultaneous localization and mapping (SLAM) technologies [2]. Notably, 3D LiDAR-based SLAM techniques have recently advanced, with noteworthy contributions from algorithms like LOAM, as described in reference [3]. LOAM effectively extracts features from point clouds, ensuring the accuracy of geometric constraints remains challenging. In such cases, solving non-linear optimization often results in divergent outcomes [4]. Subsequent enhancements, such as LeGO-LOAM, incorporate point cloud clustering and ground segmentation to improve registration [5]. Additionally, in reference [6], LIO-SAM integrates LiDAR and inertial measurements, enabling real-time motion estimation and precise map-building by integrating various measurement types in a factor graph framework.

However, outdoor agricultural environments pose unique challenges due to their unstructured nature, sparse feature points, and repeated, similar information [2]. In these environments, 3D LiDAR information can lead to matching failures caused by inconsistent features. Repeated matching failures result in divergent estimation outcomes, characterized by a significant increase in covariance. To address this challenge, GPS, providing absolute location information, becomes a crucial sensor.

Hence, this study introduces a fusion structure that integrates LIO-SAM with GPS information. While LIO-SAM-based approaches have included GPS information in their factor graph [6, 7], it is difficult to utilize covariance appropriately. In contrast, the proposed approach employs covariance intersection (CI) as a covariance-based convergence technique to combine the two sets of information in the back-end stage. It determines the most appropriate weight for fusion at each stage, which ensures consistent and accurate estimation of the robot's position even in agricultural environments.

The contributions of this study are summarized as follows:

1. Two CI-based fusion approaches are formulated that are alternatively performed and combined with LIO-SAM to demonstrate powerful 3D SLAM performance.
2. The proposed approach, based on covariance minimization, effectively manages the uncertainty of the robot's position by keeping it low.
3. Tests using a public dataset and the self-designed experiment clearly demonstrate the outstanding performance of the proposed approach.

Covariance intersection based 3D LiDAR SLAM with GPS—Winner takes all: In the general SLAM mechanism, estimated states and acquired measurements each have their respective covariances. In our approach, $P_{LIO,t}$ represents the covariance of the estimated robot pose using 3D LiDAR with an inertial measurement unit (IMU), while $P_{GPS,t}$ signifies the covariance of the GPS measurements at time t. Both covariances are updated according to the rules of CI. Similarly, $x_{LIO,t}$ stands for the robot pose estimated at time t through LIO-SAM optimization. Additionally, $x_{GPS,t}$ denotes the GPS measurement obtained at the same time. The fused covariance and robot states are calculated using CI as follows:

$$P_{F,t} = \left(\omega(P_{LIO,t})^{-1} + (1-\omega)(P_{GPS,t})^{-1}\right)^{-1}, \tag{1}$$

$$\begin{aligned} x_{F,t} = P_{F,t}(\omega(P_{LIO,t})^{-1}x_{LIO,t} \\ + (1-\omega)(P_{GPS,t})^{-1}x_{GPS,t}), \end{aligned} \tag{2}$$

where $P_{F,t}$ represents the fused covariance. x_F represents the fused robot state. The parameter ω signifies the weight assigned to the fusion of estimated results. The quality of the results varies based on the value of ω. Here, we adopt two specific rules for fusion: the "winner takes all" approach and the dynamic weight algorithm outlined in reference [8]. They are combined alternatively.

Here, the "winner takes all" fusion is referred to as WTA fusion. It is implemented by setting ω to either 1 or 0. The rule is defined as follows:

$$w = \begin{cases} 1, & \text{if } \mathrm{tr}(P_{LIO,t}) \le \mathrm{tr}(P_{GPS,t}) \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

where $\mathrm{tr}(\cdot)$ represents the trace of a matrix. The WTA fusion rule computes P_F by selecting the covariance matrix with a relatively smaller trace. Choosing information with low uncertainty is a rational approach. Consequently, $x_{F,t}$ is determined as the state corresponding to the previously selected covariance matrix among the two states $x_{LIO,t}$ and $x_{GPS,t}$.

Covariance intersection based 3D LiDAR SLAM with GPS—Dynamic weight algorithm: Here, the dynamic weight algorithm is referred to as DWA fusion outlined in reference [8]. DWA fusion serves as the primary fusion algorithm. However, if DWA fails to produce a weight within the range of 0 to 1, WTA is employed instead of DWA. The DWA process is elaborated in detail below. First off, a diagonal matrix with the eigenvalues of $P_{LIO,t}$ on its diagonal and corresponding eigenvector should be computed. Those can be represented by D_{LIO} and V_{LIO}, respectively. Now a matrix, T_{LIO}, to transform $P_{LIO,t}$ can be computed as follows:

$$T_{LIO} = \left(\sqrt{D_{LIO}}\right)^{-1}(V_{LIO})^T, \tag{4}$$

$$P'_{GPS} = T_{LIO}P_{GPS}(T_{LIO})^T, \tag{5}$$

where P'_{GPS} is the transformed matrix of P_{GPS}. P'_{GPS} is decomposed into its eigenvector, V'_{GPS} and eigenvalue matrix, D'_{GPS}. As a diagonal matrix, $\tilde{d}$ is defined as the inverse of D'_{GPS}. Based on it, $\tilde{d}$ is represented as follows:

$$\tilde{d}_i = \frac{\bar{d}_i}{1-\bar{d}_i}, \qquad \text{for} \quad 1 \le i \le n \tag{6}$$

where n is the number of elements in the fused state.

A diagonal matrix, a, can be computed using the multiplication of V'_{GPS}, D_{LIO} and $\left(V'_{GPS}\right)^T$. Lastly, p and q for the construction of a weight matrix, W, are computed as follows:

$$p = \frac{a_1\tilde{d}_2(1+\tilde{d}_1) + a_2\tilde{d}_1(1+\tilde{d}_2)}{a_1(1+\tilde{d}_1) + a_2(1+\tilde{d}_2)}, \tag{7}$$

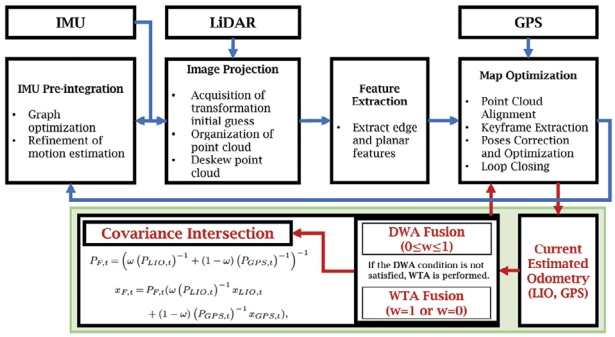


Fig. 1 *Proposed 3D simultaneous localization and mapping (SLAM) framework with LIO-SAM and covariance intersection (CI)-based fusion. In our proposed 3D SLAM framework, we integrate LIO-SAM with CI-based fusion. Utilizing the LIO-SAM framework as the foundation, LIO-SAM odometry is fused with GPS odometry using DWA or WTA. These two fusion strategies operate alternatively to calculate an appropriate weight*

$$q = \frac{a_1 \tilde{d}_2^2 (1 + \tilde{d}_1) + a_2 (\tilde{d}_1)^2 (1 + \tilde{d}_2)}{a_1 (1 + \tilde{d}_1) + a_2 (1 + \tilde{d}_2)}, \tag{8}$$

where a_1 and a_2 are the first and second diagonal elements of a, respectively. Likewise, $\tilde{d}_1$ and $\tilde{d}_2$ are the first and second diagonal elements of $\tilde{d}$, respectively. Finally, we can obtain the weight W of each state as follows:

$$W = \begin{bmatrix} -p + \sqrt{p^2 - q} & 0 \\ 0 & -p + \sqrt{p^2 - q} \end{bmatrix}, \tag{9}$$

where each diagonal element of W should be involved in [0, 1]. If the aforementioned condition is not met, for instance, if q exceeds p^2, then the rules of WTA are applied instead.

Proposed SLAM framework: Figure 1 illustrates the proposed integrated SLAM framework employing the fusion mechanism based on CI. The fundamental framework employed here is the LIO-SAM approach. Its cloud information undergoes processing through modules: "imageProjection", "featureExtraction", and "mapOptimization". In the "mapOptimization" step, mapping accuracy is improved by aligning point clouds, estimating LiDAR odometry for precise robot tracking, and refining motion trajectories through graph optimization. Additionally, the "imuPreintegration" module refines motion estimation using IMU and LiDAR odometry data, ensuring accurate mapping by removing point cloud distortions and enhancing initial value estimation. The LIO-SAM odometry is fused using two fusion strategies, namely DWA and WTA, specifically when GPS odometry is acquired. Although DWA and WTA represent different fusion methods, WTA is employed complementarily in the DWA-based fusion process.

The GPS cycle operates at a rate 10 times slower than the LIO-SAM odometry cycle. Therefore, if GPS data is not received, it is fused with a weight ($w = 1$). Additionally, it is essential to note that the coordinate systems of GPS and LIO-SAM differ. Consequently, a coordinate system transformation matrix is necessary between them. To calculate this transformation matrix, singular value decomposition-based optimization [9] is performed using the initial three consecutive pairs of each set of information. From the optimal $R_{rotation}$ and t_{trans} matrices derived from these results, the transformation matrix ($T_{GPS,LIO}$) is obtained as follows:

$$T_{GPS,LIO} = \begin{bmatrix} R_{rotation} & t_{trans} \\ \mathbf{0} & 1 \end{bmatrix}, \tag{10}$$

where $T_{GPS,LIO}$ is the 3x3 matrix. According to $T_{GPS,LIO}$, the state and covariance of original GPS information for fusion are transformed as follows.:

$$x_{GPS,t} = T_{GPS,LIO} x_{OGPS,t}, \tag{11}$$

(a) (b)

Fig. 2 *Public dataset. (a) The designed robot and equipped sensors. (b) The robot trajectory. S and E represent the start and end points, respectively. More details can be found in reference [10]*

$$P_{GPS,t} = T_{GPS,LIO} P_{OGPS,t} T_{GPS,LIO}^{-1}, \tag{12}$$

where $x_{OGPS,t}$ and $x_{OGPS,t}$ are 3x1 state vector and 3x3 covariance matrix regarding original GPS information, respectively. The results are transformed state and covariance. $x_{GPS,t}$ is a 3x1 state vector, but because the last element is 1, it was employed as a 2x1 vector in the formulation.

Public dataset: The CitrusFarm dataset [10] is a comprehensive agricultural robotics dataset that covers various aspects of farming. It includes both multi-spectral images and navigational sensor data, making it valuable for tasks such as localization, mapping, and crop monitoring. This dataset was gathered during the summer of 2023 by a wheeled mobile robot at the agricultural experimental station of the university of California, Riverside, which is shown in Figure 2. It includes data from wheel odometry, LiDAR, IMU, and GPS-RTK(Real Time Kinematic), all measured along the robot's trajectory depicted in Figure 2b.

Performance evaluation of public dataset: In this dataset, the robot's position obtained from RTK GPS is set to the true value. Therefore, to demonstrate the fusion performance of GPS and SLAM methods, which is an advantage of this study, we additionally created noisy GPS data with noise added to the RTK GPS and used it for fusion. The equation is as follows:

$$x_{GPS,t} = x_{GPS,t} + v, \tag{13}$$

where v is zero-mean white noise with covariance matrix $R = \sigma_v^2 * I$.

For quantitative performance evaluation, the norm of a vector is calculated as the difference between $x_{true,last}$ obtained from the final position of RTK GPS and $x_{est,last}$ estimated from each method, which is as follows:

$$E_{pose} = \sqrt{(x_{true,last} - x_{est,last})^T (x_{true,last} - x_{est,last})}, \tag{14}$$

where $x_{est,last}$ can be either $x_{LIO,last}$ resulted from the LIO-SAM or $x_{F,last}$ obtained by the proposed method.

Since we proposed the covariance fusion-based 3D SLAM framework, $P_{F,t}$ is one of the evaluation factors. It is important to demonstrate that the proposed approach exhibits lower covariance when compared to the covariances of GPS and LIO-SAM. It is clear that tr($P_{F,t}$) remains consistently lower than tr($P_{LIO,t}$) and tr($P_{GPS,t}$).

Tests and analysis: Tests were conducted using the dataset. Figure 3 shows the robot trajectories estimated by LIO-SAM, GPS odometer, LIO-SAM with the WTA fusion, and the proposed method. The proposed method outperforms other methods by fusing each step consistently. This integration is highlighted further in Figure 4. The results of the covariance traces were compared over time. The fused trace of covariance of the proposed approach denoted as tr($P_{F,t}$), is significantly lower than those of other methods. This is because the proposed method is a covariance-based fusion technique. Table 1 provides a comparison of the final position error, E_{pose}. Our proposed CI-based fusion SLAM approach shows the minimum error compared to conventional methods. This implies that the estimated and fused location information is converging closer to the true value by reducing the covariance. Table 2 shows the average processing time per cycle in the CitrusFarm dataset.

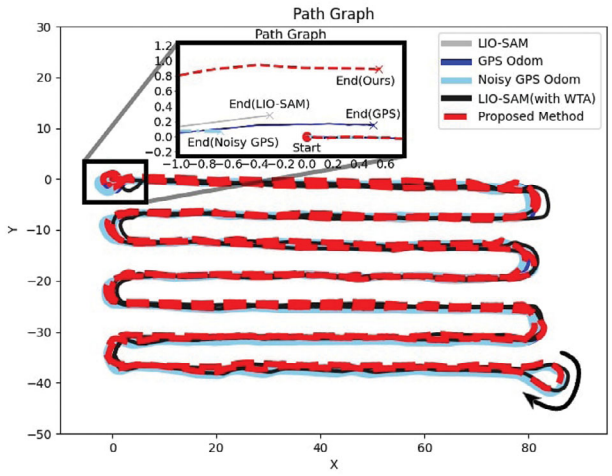


Fig. 3 *Estimated robot poses in CitrusFarm dataset. This figure shows the robot trajectories estimated by LIO-SAM, GPS odometry, LIO-SAM with the WTA fusion, and the proposed method. The proposed method outperforms other methods because it integrates GPS information while fusing each step consistently*

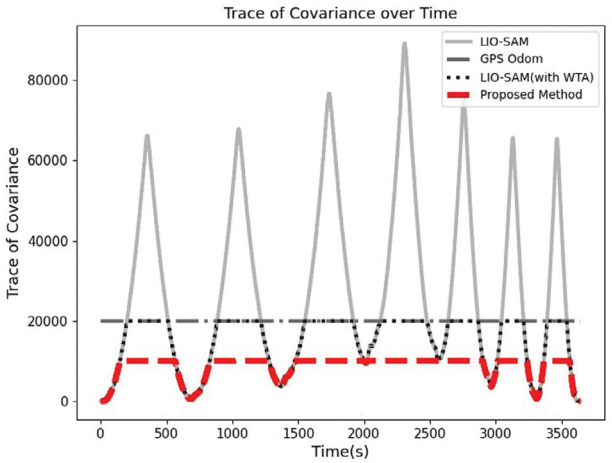


Fig. 4 *Trace of covariance over time in CitrusFarm dataset. The traces of covariances obtained from each method are depicted and compared over time. The fused trace of covariance of the proposed approach, denoted as* $\mathrm{tr}(P_{F,t})$*, is significantly lower than those of other methods*

Table 1. Final position error in CitrusFarm dataset (unit: meter)

	$x_{true,last}$	$x_{est,last}$	E_{pose}
Noisy GPS	0.5078, 0.1512	-0.6643, 0.07008	1.1749
LIO-SAM (without GPS)	0.5078, 0.1512	-0.2861, 0.2797	0.8042
LIO-SAM (with GPS)	0.5078, 0.1512	2.2642, 0.471551	1.7854
LIO-SAM (with WTA)	0.5078, 0.1512	3.745, 0.3356	3.2424
Proposed method	0.5078, 0.1512	0.5502, 0.8882	0.7382

Table 2. Average processing time per cycle (unit: ms)

	Average processing time
LIO-SAM (without GPS)	139.14
LIO-SAM (with GPS)	141.01
LIO-SAM (with WTA)	141.08
Proposed method	142.18

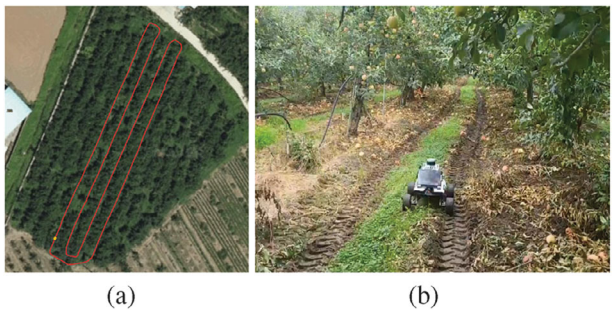


Fig. 5 *Experimental environment. Panel (a) depicts the top view of the experimental environment. Red lines indicate the guided robot trajectory. Panel (b) shows the robot under experimentation and its surrounding environment*

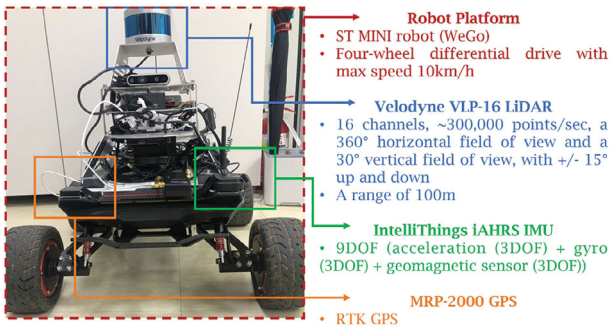
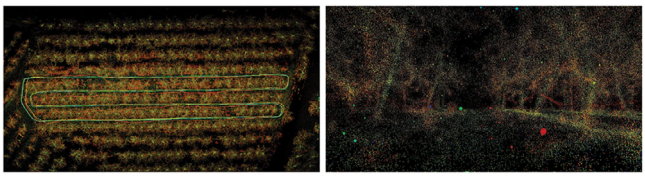


Fig. 6 *Robot platform. The robot platform is equipped with sensors including a 3D LiDAR sensor, an inertial measurement unit (IMU) sensor, and a GPS sensor*

Our approach results in a slight increase in computational time attributed to matrix operations. Nevertheless, this increment in computational load is minimal, accounting for about 2% of the overall runtime.

Experimental setup and performance evaluation: An experiment was also conducted in an orchard located in Goesan-gun, Chungcheongbuk-do, South Korea as shown in Figure 5. Figure 5 shows a top view of the environment where the experiment was conducted and a photo of the internal orchard. The robot moved along the fruit trees, and at this time, it ran in shape such as Π. The collected data is publicly available on [11] and called the orchard dataset. Figure 6 shows Wego's ST MINI robot, which has four-wheel differential drive and independent suspension and can reach speeds of up to 10 km/h. The dimensions of the robot are 612 × 580 × 245 mm. Sensors onboard the robot include a 3D LiDAR sensor, an IMU sensor, and a GPS sensor. The model of the equipped LiDAR sensor is Velodyne VLP-16, which is one of the popular outdoor 3D LiDAR sensors. It supports a range of 100 m, 16 channels, up to 300,000 points/s, a 360° horizontal field of view, and a 30° vertical field of view with +/- 15° up and down. IntelliThings iAHRS was the model of IMU sensor utilized, which has a 9-Degree of Freedom (DoF) consisting of an accelerometer, gyroscope, and magnetometer. As the GPS sensor, MRP-2000 Real-time kinematic GPS was used. These sensors have been used to effectively perform outdoor 3D SLAM. For performance evaluation in the experiment, the same performance evaluation formula described in Equation (14) is employed. Since the robot's arrival position is controlled to be the same as the origin, $x_{true,last}$ becomes [0,0].

Experiment and analysis: Experiments were conducted in the aforementioned environment. The sensor data acquired by the robot were employed in both LIO-SAM and the proposed method. Figure 7 displays the estimated test environment from different perspectives. Particularly, Figure 7a is detailed in Figure 8, showing the robot trajectories estimated by LIO-SAM, GPS odometer, LIO-SAM with the WTA fusion, and the proposed method. The proposed method outperforms other methods because it integrates GPS information while fusing each step consistently. This integration is highlighted further in Figure 9, where the results of the covariance traces were compared over time. The fused trace of



(a) (b)

Fig. 7 *Estimated 3D environment. Panel (a) shows the top view of the estimated environment, featuring the estimated robot trajectory represented by the green line. Panel (b) shows a front view of the estimated environment*

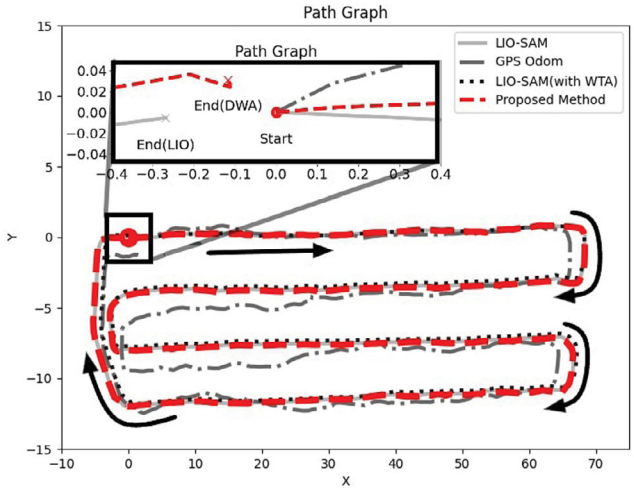


Fig. 8 *Estimated robot poses in orchard dataset. This figure shows the robot trajectories estimated by LIO-SAM, GPS odometry, LIO-SAM with the WTA fusion, and the proposed method. The proposed method outperforms other methods because it integrates GPS information while fusing each step consistently*

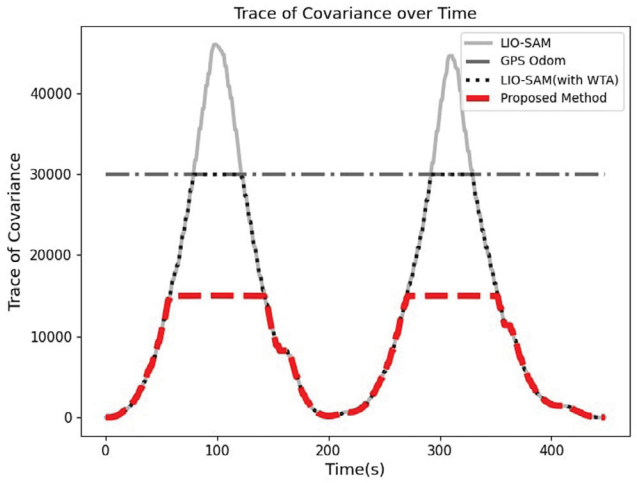


Fig. 9 *Trace of covariance over time. The traces of covariances obtained from each method are depicted and compared over time. The fused trace of covariance of the proposed approach, denoted as* $\mathrm{tr}(P_{F,t})$*, is significantly lower than those of other methods*

covariance of the proposed approach, denoted as $\mathrm{tr}(P_{F,t})$, is significantly lower than those of other methods. In other words, the robot's trajectory does not deviate beyond the uncertainties of the GPS and LIO-SAM. Table 3 provides a comparison of the final position error, E_{pose}. It can be observed that the proposed CI-based fusion SLAM approach shows improvements compared to conventional methods. This implies that the estimated and fused location information is converging closer to the true value due to the reduction in covariance. Table 4 presents the average processing time per cycle within the orchard dataset. Similar to the CitrusFarm dataset, our method slightly increases the computation

Table 3. Final position error (unit: m)

	$x_{true,last}$	$x_{est,last}$	E_{pose}
LIO-SAM (without GPS)	0.0, 0.0	0.2619, 0.0082	0.2620
LIO-SAM (with GPS)	0.0, 0.0	-0.7223, -0.2645	0.7692
LIO-SAM (with WTA)	0.0, 0.0	1.0665, 0.2978	1.1073
GPS odometry	0.0, 0.0	1.50315, 0.67912	1.6494
Proposed method	0.0, 0.0	-0.1165, 0.03222	0.1209

Table 4. Average processing time per cycle (unit: ms)

	Average processing time
LIO-SAM (without GPS)	135.36
LIO-SAM (with GPS)	137.97
LIO-SAM (with WTA)	139.08
Proposed method	139.26

time due to matrix operations. However, this increase in computational load is negligible.

Conclusion: This letter introduces an integrated SLAM framework tailored for challenging agricultural environments. By combining the cutting-edge technique known as LIO-SAM and supplementing it with feature-independent GPS data, our approach presents an alternative convergence strategy based on DWA and WTA. Through comprehensive numerical and graphical evaluation, the proposed approach exhibited a high level of accuracy in estimating the robot's position in both the dataset and outdoor experiment. In the future, our research aims to revolutionize agricultural practices by seamlessly integrating this system with robot arm control by paving the way for automated crop cultivation and harvesting.

Author contributions: **HannGyoo Kim**: Investigation; methodology; software; validation; visualization; writing—original draft. **Hea—Min Lee**: Conceptualization; funding acquisition; resources; writing—review and editing. **SeungHwan Lee**: Methodology; project administration; supervision; writing—review and editing.

Acknowledgements: This work was supported by the Korea Institute of Planning and Evaluation for Technology in Food, Agriculture and Forestry (IPET) through Open Field Smart Agriculture Technology Short-term Advancement Program, funded by the Ministry of Agriculture, Food and Rural Affairs (122032-03).

Conflict of interest statement: The authors declare no conflicts of interest.

Data availability statement: The data that support the findings of this study are openly available at http://gofile.me/5hn5O/8oO2ULJMn, reference number [11].

Received: *27 October 2023* Accepted: *8 February 2024*
doi: 10.1049/ell2.13206

References

1 Xie, D., Chen, L., Liu, L., Chen, L., Wang, H.: Actuators and sensors for application in agricultural robots: a review. *Machines* **10**(10), 913, 1–31 (2022). https://doi.org/10.3390/machines10100913

2 Ding, H., Zhang, B., Zhou, J., Yan, Y., Tian, G., Gu, B.: Recent developments and applications of simultaneous localization and mapping in agriculture. *J. Field Robot.* **39**(6), 956–983 (2022). https://doi.org/10.1002/rob.22077
3 Zhang, J., Singh, S.: LOAM: Lidar odometry and mapping in real-time. In: *Robotics: Science and Systems*, pp. 1–9, Berkeley, USA (2014). https://doi.org/10.15607/RSS.2014.X.007
4 Chen, S., Zhou, B., Jiang, C., Xue, W., Li, Q.: A LiDAR/visual SLAM backend with loop closure detection and graph optimization. *Remote Sens.* **13**(14), 2720 (2021). https://doi.org/10.3390/rs13142720
5 Shan, T., Englot, B.: LeGO-LOAM: lightweight and ground-optimized lidar odometry and mapping on variable terrain. In: *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 4758–4765. IEEE, Madrid, Spain (2018). https://doi.org/10.1109/IROS.2018.8594299
6 Shan, T., Englot, B., Meyers, D., Wang, W., Ratti, C., Rus, D.: LIO-SAM: tightly-coupled lidar inertial odometry via smoothing and mapping. In: *2020 IEEE/RSJ international conference on intelligent robots and systems (IROS)*, pp. 5135–5142. IEEE, Las Vegas, NV (2020). https://doi.org/10.1109/IROS45743.2020.9341176
7 Li, D., Sun, B., Liu, R., Xue, R.: Tightly coupled 3D lidar inertial SLAM for ground robot. *Electronics* **12**(7), 1649, 1–13 (2023). https://doi.org/10.3390/electronics12071649
8 Reinhardt, M., Noack, B., Hanebeck, U.D.: Closed-form optimization of covariance intersection for low-dimensional matrices. In: *2012 15th International Conference on Information Fusion*, pp. 1891–1896. IEEE, Singapore (2012)
9 Sorkine-Hornung, O., Rabinovich, M.: Least-Squares Rigid Motion Using SVD. In: Technical note, pp. 1–5. ETH Zurich, Department of Computer Science (2017)
10 Citrus Fram Dataset, https://ucr-robotics.github.io/Citrus-Farm-Dataset. Accessed 24 April 2023
11 Our Dataset, http://gofile.me/5hn5O/8oO2ULJMn, Accessed 24 April 2023