

## 3차원 포즈 추정을 위한 시각 관성 주행 거리 측정

이보은, 고낙용<sup>†</sup>

## Visual Inertial Odometry for 3-Dimensional Pose Estimation

Boeun Lee<sup>ORCID</sup>, Nak Yong Ko<sup>ORCID</sup>

Department of Electronics Engineering, Interdisciplinary Program in IT-Bio Convergence System, Chosun University, Gwangju 61452, Korea

### ABSTRACT

Real-time localization is essential for autonomous driving of robots. This paper presents the implementation and a performance analysis of a localization algorithm. To estimate the position and attitude of a robot, a visual inertial odometry (VIO) algorithm based on a multi-state constraint Kalman filter is used. The sensors employed in this study are a stereo camera and an inertial measurement unit (IMU). The performance is analyzed through experiments using three different camera view directions: floor-view, front-view, and ceiling-view. The number of detected features also affects navigation performance. Even if the number of recognized feature points is large, performance degrades if the correspondence between feature points is not accurately identified. The results show that VIO improves navigation performance even with low-cost sensors, thus facilitating map building as well as autonomous navigation.

**Keywords:** VIO, MSCKF, pose estimation, calibration

**주요어:** 시각 관성 주행 거리 측정, 다중 상태 제약 칼만 필터, 포즈 추정, 보정

### 1. 서론

최근 자율주행 기술의 발전으로 자율주행 로봇 및 자동차에 대한 관심이 높아지고, 연구 개발도 활발하게 이루어지고 있다. 국내 기업뿐 아니라 구글, 엔비디아, 인텔 등에서도 연구 개발이 꾸준히 이루어지고 있다 (Jung & Ryu 2019). 이러한 자율주행 기술은 특히 로봇 산업 시장 확대에 따라 로봇 분야의 수요가 높아지고 있다.

자율주행 로봇은 우주 또는 해저 지역, 군사 지역, 그리고 재난 및 재해 지역 등의 인간이 작업을 직접 수행하기 어렵고 부적합한 곳에서 업무를 대신 수행하기 위해 연구 및 개발되어왔다. 초기의 로봇은 인간이 직접 조작해야 했지만, 현재는 Global Positioning System (GPS), 관성 측정 장치(inertial measurement unit, IMU), 그리고 카메라 등 로봇에 설치할 수 있는 센서의 개발로 인간의 개입 없이도 로봇 스스로가 판단하고 업무를 수행할

수 있게 되었다 (Lee & Choe 2013).

로봇이 자율주행을 하기 위해서는 로봇 자신의 위치 및 자세 파악이 필수적이다. 이를 위해 비전 센서를 사용하는 경우 일반적으로 점, 선 혹은 점과 선을 모두 사용한 특징 추출을 통해 위치 및 자세를 추정한다. 이 경우 대부분 카메라 렌즈가 정면을 향하도록 고정하여 실험 및 시뮬레이션을 진행한다 (Zhang et al. 2020, Gurturk et al. 2021, Chen et al. 2023). 비전 관련 추정 연구에는 대표적으로 시각 관성 주행 거리 측정(visual inertial odometry, VIO), 그리고 시각 관성 항법 시스템(visual inertial navigation system) 등이 있다.

Mourikis & Roumeliotis (2007)는 실시간 vision-aided 관성 항법을 위한 확장 칼만 필터(extended Kalman filter) 기반 알고리즘을 제시한다. 알고리즘의 성능은 정면 주행 장면을 기록해 놓은 KITTI 데이터를 통해 검증하였다. Kottas & Roumeliotis (2013)는 관성 측정과 선의 시각 관찰을 기반으로 3차원 이동체의 상태를 추정하는 알고리즘을 제안한다. 이 연구도 카메라가

Received Aug 13, 2024 Revised Sep 26, 2024 Accepted Oct 25, 2024

<sup>†</sup>Corresponding Author E-mail: nyko@chosun.ac.kr



Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

정면을 바라본 실험을 통해 알고리즘을 검증하였다. 마지막으로 Yang et al. (2019)는 3차원 선 삼각측량에 대한 모션 분석과 점과 선을 이용한 모노 카메라 VIO 알고리즘을 제안한다. 제안한 알고리즘을 Euroc Mav 데이터를 통해 검증하였고, 이는 드론을 사용하여 기록한 데이터로 공중에서 아래를 비스듬히 보는 이미지로 이루어져 있다.

기존의 제시된 비전 관련 연구에서 사용한 위치 및 자세 추정 은 대부분이 카메라를 정면을 향하게 하고 실험 및 시뮬레이션을 진행한다. 하지만, 카메라 렌즈가 정면을 향한 VIO는 수중에서 작업하는 지형 탐지 로봇 등에 적합하지 않을 수 있다는 한계점이 존재한다. 수중에서 지형을 탐지하는 로봇의 경우 카메라를 정면을 향하도록 하면 어두워 정확한 탐지가 어려울 수 있다. 이 경우에는 카메라를 바닥을 향하도록 하여 탐지하는 것이 정확도 향상에 도움을 줄 수 있다.

이와 같이 비전 관련 연구의 대다수는 카메라가 향하는 방향에 대한 자세한 설명은 생략되어 있고, 이를 중요한 요소로 여기지 않는다. 이는 카메라가 정면을 향하는 일반적인 방식보다 방향을 바꿨을 때의 알고리즘 성능이 높을 수 있다는 가능성이 존재한다는 것을 보여준다. 따라서, 카메라가 향한 방향이 연구 결과에 미치는 영향에 대한 분석이 필수적이다.

본 논문은 스테레오 카메라와 IMU를 사용하여 자율주행을 위한 VIO 시스템을 Linux의 Ubuntu 환경에서 ROS를 사용하여 구현하여 이동체의 위치 및 자세를 추정을 다룬다. VIO는 GPS를 사용하지 않기 때문에, GPS 측정이 어려운 실내 환경이나 신호가 원활하게 잡히지 않는 환경에서도 사용 가능하다. 또한, 카메라 렌즈 방향을 정면뿐만 아니라 천장 방향과 바닥 방향을 향한 실험을 진행하여 MATLAB을 통해 결과 및 성능을 분석하였다.

본 논문의 구성은 다음과 같다. 2장에서는 카메라와 IMU의 보정에 대해 설명한다. 3장에서는 다중 상태 제약 결합 필터를 사용한 시각 관성 주행 거리 측정 시스템을 소개한다. 4장에서는 실험을 통해 제안하는 실험 방법의 성능을 평가한다. 마지막으로 5장에서는 연구 결과를 정리하고 결론을 맺는다.

## 2. 센서 보정

### 2.1 스테레오 카메라 보정

카메라 보정은 오랜 기간동안 중요한 주제였다. 컴퓨터 비전 분야에서 측정의 더 높은 정확도의 필요성이 증가함에 따라, 컴퓨터 비전 연구가 지속적으로 진행되어왔다 (Weng et al. 1990). 영상의 움직임, 차원 측정, 그리고 입체 투시 깊이 등과 같은 정량적 측정이 동반되는 분야에 카메라의 정확한 보정은 특히 중요하다. 또한, 서로 다른 각 카메라 간의 기하학적 관계를 추정하는 것도 보정의 중요한 측면이다. 이러한 관계를 특징짓는 매개변수를 외부 매개변수라고 한다. 기하학적인 측정의 경우 카메라의 왜곡(distortion)이 주된 문제가 되며, 이는 영상 평면에서 영상 포인트의 위치와 관련이 있지만, 영상의 품질과는 직접적인 관련이 없다 (Weng et al. 1992). 따라서, 카메라 보정은 중요한 과제가 되

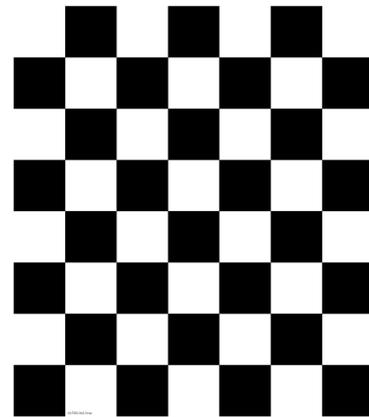


Fig. 1. Checker board pattern.

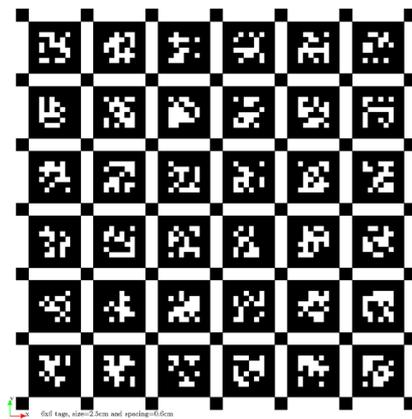


Fig. 2. April grid pattern.

었다.

각 카메라의 내부 및 외부 파라미터를 추정하는 것이 스테레오 카메라를 보정하는 목적이다. 내부 파라미터는 이미지 스케일링 계수, 초점거리, 이미지 중심 및 렌즈 왜곡 계수를 포함한 카메라 광학 내부 특성을 나타내고, 외부 파라미터는 좌표계에 대한 카메라의 위치와 방향을 나타낸다 (Memon & Khan 2001).

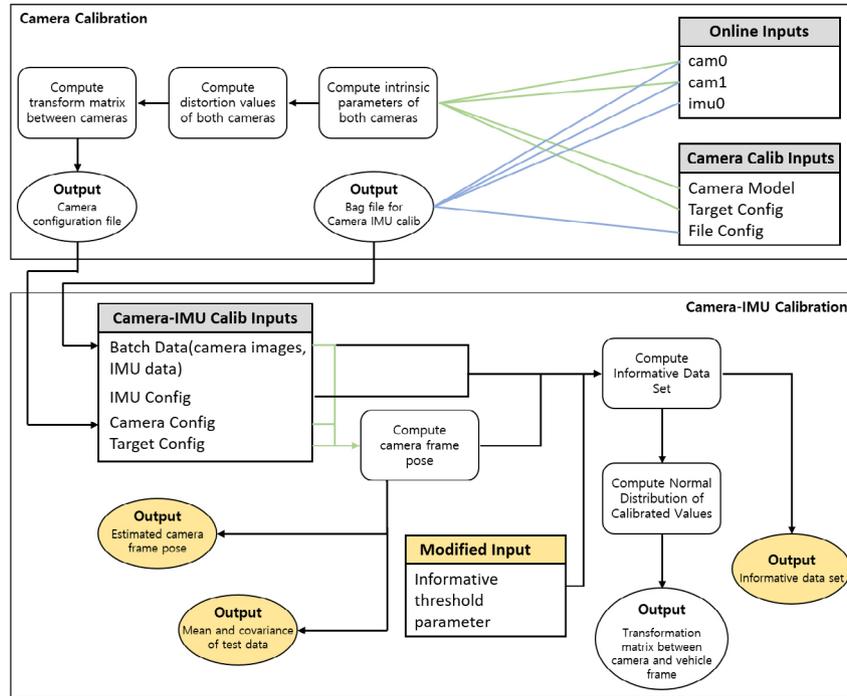
카메라를 보정하기 위해서는 주로 Figs. 1과 2와 같은 체커보드 패턴 또는 april grid 패턴을 사용한다. 특히, 체커보드는 지난 20년 동안 보정에 가장 많이 사용되었다 (Maye et al. 2013). 스테레오 카메라를 상하좌우로 회전하며 움직이면서 저장한 데이터 파일을 Kalibr 툴박스를 사용하여 카메라의 내부 및 외부 매개변수를 추정한다 (Rehder et al. 2016). 추정 결과는 Table 1과 같다.

### 2.2 스테레오 카메라와 IMU의 보정

카메라와 IMU 사이의 보정은 카메라의 시각적 타겟과 IMU의 중력을 사용하여 두 센서가 모두 수직 방향을 관찰하도록 하여 추정할 수 있다. Fig. 3은 스테레오 카메라와 IMU의 보정 과정을 나타낸다. 카메라와 IMU 보정은 두 센서 사이의 상대적인 변환, 회전 및 시간 오프셋을 결정하기 위해 진행한다. 두 센서 사이의 보정에 대해서는 Jennings & Eshghi (2019)에 자세하게 기술되어

**Table 1.** Intrinsic and extrinsic parameters of stereo camera.

Cameras	Intrinsic and extrinsic parameter	Values			
cam0	Distortion coefficients	[-0.4206	0.1731	$-1.5014 \times 10^{-5}$	-0.0017]
	Intrinsics vector	[483.45	484.82	315.56	247366]
cam1	Transformation matrix	[0.9999	-0.0004	-0.0152	-0.1202]
		[0.0006	0.9999	0.0098	-1.9093]
		[0.0152	-0.0098	0.9998	-0.0014]
	[0	0	0	1]	
Distortion coefficients	[-0.425	0.1725	-0.0005	-0.0019]	
Intrinsics vector	[485.32	486.22	326.98	235.87]	



**Fig. 3.** Calibration block diagram of stereo camera and IMU.

**Table 2.** Transformation matrix between cameras and IMU.

Cameras	Transformation matrix				
cam0	Intrinsics vector $T_{cam0, imu}$	[0.0058	-0.9999	0.0068	-0.0037]
		[0.0037	-0.0068	-0.9999	0.0001]
		[0.9999	0.0058	0.0036	-0.0004]
		[0	0	0	1]
cam1	Transformation matrix $T_{cam1, imu}$	[-0.0095	-0.9999	0.0072	-0.1165]
		[0.0135	-0.0073	-0.9999	0.0001]
		[0.9999	-0.0094	0.1357	-0.0018]
		[0	0	0	1]
cam1	Transformation matrix $T_{cam1, cam0}$	[0.9999	-0.0004	-0.0152	-0.1202]
		[0.0055	0.9999	0.0098	-0.0001]
		[0.0152	-0.0098	0.9998	-0.0015]
		[0	0	0	1]

있다. Table 2는 카메라와 IMU 사이의 보정 결과로 얻은 변환 행렬을 나타낸다.

여기서,  $T_{cam0, imu}$ 는 IMU 프레임을 cam0 프레임으로 변환하는 변환 행렬,  $T_{cam1, imu}$ 는 IMU 프레임을 cam1 프레임으로 변환하는 변환 행렬, 그리고  $T_{cam1, cam0}$ 는 cam0 프레임을 cam1 프레임으로 변환하는 변환 행렬을 나타낸다.

### 3. 다중 상태 제약 칼만 필터를 사용한 시각 관성 주행 거리 측정

다중 상태 제약 칼만 필터(multi-state constraint Kalman filter, MSCKF)를 사용한 VIO를 통하여 3차원 포즈를 추정한다. 3차원 포즈에는 위치 ( $x, y, z$ )와 자세 ( $\phi, \theta, \psi$ )가 포함된다.

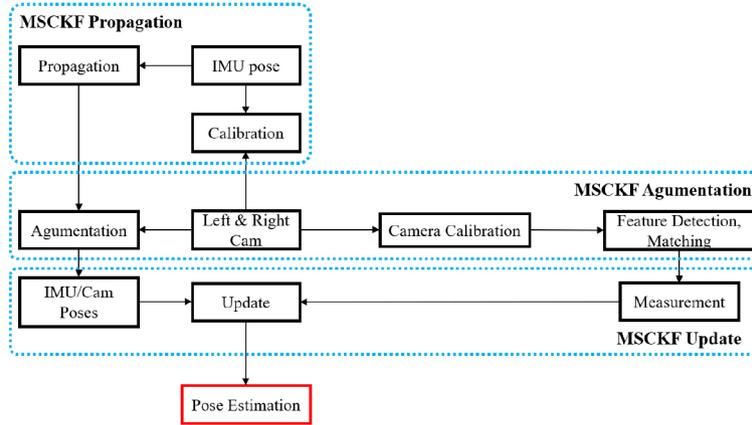


Fig. 4. VIO block diagram.

MSCKF는 고정 슬라이딩 윈도우를 사용하여 3D 움직임을 추적하는 확장 칼만 필터를 기본으로 사용하는 알고리즘이다. 특히, MSCKF는 simultaneous localization and mapping (SLAM) 기반 알고리즘과 비교하여 일관성 있고 정밀도가 높으며 계산의 복잡도가 낮다 (Do et al. 2021).

일반적인 스테레오 MSCKF 알고리즘은 Fig. 4와 같이 나타낼 수 있다. 카메라를 통한 이미지에서 특징을 추출하고, 스테레오 카메라의 각 카메라 이미지 사이에 매칭을 진행한다. 상태는 왼쪽 카메라의 포즈 파라미터로 증강되고, 두 카메라 사이의 보정 파라미터를 고려하여 오른쪽 카메라의 포즈가 왼쪽 카메라에서 계산된다. 마지막으로, 상태는 각 시간 단계에서 두 카메라에서 얻은 측정값에 의해 업데이트된다.

MSCKF의 IMU 상태 벡터, 각각의 카메라 오차 상태 벡터, 그리고 전체 오차 상태 벡터는 Eqs. (1-3)으로 정의한다.

$$\mathbf{x}_{IMU} = \begin{bmatrix} {}^I_G \mathbf{q}^T & \mathbf{b}_g^T & {}^G \mathbf{v}_I^T & \mathbf{b}_a^T & {}^G \mathbf{p}_I^T & {}^I_C \mathbf{q}^T & {}^I \mathbf{p}_C^T \end{bmatrix} \quad (1)$$

$$\tilde{\mathbf{x}}_{C_i} = \begin{bmatrix} c_i \tilde{\theta}^T & {}^G \tilde{\mathbf{p}}_{C_i}^T \end{bmatrix} \quad (2)$$

$$\tilde{\mathbf{x}} = \begin{bmatrix} \tilde{\mathbf{x}}_{IMU}^T & \tilde{\mathbf{x}}_{C_1}^T & \tilde{\mathbf{x}}_{C_2}^T & \dots & \tilde{\mathbf{x}}_{C_N}^T \end{bmatrix}^T \quad (3)$$

여기서,  ${}^I_G \mathbf{q}^T$ 는 global 프레임 G에서 IMU 프레임 I로의 회전을 나타내고,  $\mathbf{b}_g^T$ 와  $\mathbf{b}_a^T$ 는 각각 IMU의 자이로스코프와 가속도계에서 측정된 각속도와 바이어스를,  ${}^G \mathbf{v}_I^T$ 와  ${}^G \mathbf{p}_I^T$ 는 각각 G에 대한 IMU 속도와 위치를 나타낸다. 또한,  ${}^I_C \mathbf{q}^T$ 와  ${}^I \mathbf{p}_C^T$ 는 카메라 프레임 C에서 I로의 회전과 이동을 나타낸다.

연속 시간에서 IMU 상태의 process model은 Eq. (4)로 정의된다.

$$\begin{aligned} {}^I_G \dot{\mathbf{q}} &= \frac{1}{2} \Omega(\hat{\omega}) {}^I_G \hat{\mathbf{q}}, & \dot{\mathbf{b}}_g &= \mathbf{0}_{3 \times 1} \\ {}^G \dot{\mathbf{v}} &= C \left( {}^I_G \hat{\mathbf{q}} \right)^T \hat{\mathbf{a}} + {}^G \mathbf{g} \\ \dot{\mathbf{b}}_a &= \mathbf{0}_{3 \times 1}, & {}^G \dot{\mathbf{p}}_I^T &= {}^G \dot{\mathbf{v}} \\ {}^I_C \dot{\mathbf{q}} &= \mathbf{0}_{3 \times 1}, & {}^I \dot{\mathbf{p}}_C &= \mathbf{0}_{3 \times 1} \end{aligned} \quad (4)$$

여기서,  $\hat{\omega}$ 와  $\hat{\mathbf{a}}$ 는 각각 각속도와 가속도에 대한 3×1의 IMU 측정값이며 Eqs. (5-6)으로 정의되고, C는 쿼터니언(quaternion)을 회전 행렬로 변환하는 함수이다.  $\Omega(\hat{\omega})$ 는 Eq. (7)로 정의되고,  $[\hat{\omega}_\times]$ 는  $\hat{\omega}$ 의 반대칭 행렬(skew-symmetric matrix)을 나타낸다. Eq. (4)에 기반한 IMU 오차 상태에 대한 process model은 Sun et al. (2018)에 자세하게 기술되어 있다.

$$\hat{\omega} = \omega_m - \hat{\mathbf{b}}_g \quad (5)$$

$$\hat{\mathbf{a}} = \mathbf{a}_m - \hat{\mathbf{b}}_a \quad (6)$$

$$\Omega(\hat{\omega}) = \begin{bmatrix} -[\hat{\omega}_\times] & \omega \\ \omega^T & \mathbf{0} \end{bmatrix} \quad (7)$$

스테레오 카메라의 measurement model  $\mathbf{z}_j^i$ 는 Eq. (8)로 정의된다.

$$\mathbf{z}_j^i = \begin{bmatrix} u_{i,1}^j \\ v_{i,1}^j \\ u_{i,2}^j \\ v_{i,2}^j \end{bmatrix} = \begin{bmatrix} 1 & \mathbf{0}_{2 \times 2} \\ \mathbf{0}_{2 \times 2} & 1 \end{bmatrix} \begin{bmatrix} c_{i,1} X_j \\ c_{i,1} Y_j \\ c_{i,2} X_j \\ c_{i,2} Y_j \end{bmatrix} \quad (8)$$

여기서,  $k \in \{1, 2\}$ 일 때  $({}^{c_{i,k}} X_j, {}^{c_{i,k}} Y_j, {}^{c_{i,k}} X_j)^T$ 는 특징점의 위치를 나타낸다. 각 카메라에서의 특징점의 위치는 Eqs. (9-10)으로 나타난다.

$${}^{c_{i,1}} \mathbf{p}_j = \begin{bmatrix} c_{i,1} X_j \\ c_{i,1} Y_j \\ c_{i,1} Y_j \end{bmatrix} = C \left( {}^{c_{i,1}}_G \mathbf{q} \right) \left( {}^G \mathbf{p}_j - {}^G \mathbf{p}_{C_{i,1}} \right) \quad (9)$$

$${}^{c_{i,2}} \mathbf{p}_j = \begin{bmatrix} c_{i,2} X_j \\ c_{i,2} Y_j \\ c_{i,2} Y_j \end{bmatrix} = C \left( {}^{c_{i,2}}_G \mathbf{q} \right) \left( {}^G \mathbf{p}_j - {}^G \mathbf{p}_{C_{i,2}} \right) \quad (10)$$

Measurement model을 현재 추정값에서 선형화하면 측정 잔차(residual)는 Eq. (11)과 같다.

$$\mathbf{r}_i^j = \mathbf{z}_i^j - \hat{\mathbf{z}}_i^j = \mathbf{H}_{C_i}^j \tilde{\mathbf{x}}_{C_i} + \mathbf{H}_{f_i}^j \hat{\mathbf{p}}_j + \mathbf{n}_i^j \quad (11)$$

$$\mathbf{H}_{C_i}^j = \frac{\partial \mathbf{z}_i^j}{\partial \mathbf{x}_{C_i}^j} \cdot \frac{\partial \mathbf{z}_i^j}{\partial \mathbf{x}_{C_i}^j} + \frac{\partial \mathbf{z}_i^j}{\partial \mathbf{p}_j} \cdot \frac{\partial \mathbf{z}_i^j}{\partial \mathbf{p}_j} \quad (12)$$

$$\mathbf{H}_{f_i}^j = \frac{\partial \mathbf{z}_i^j}{\partial \mathbf{p}_j} \cdot \frac{\partial \mathbf{z}_i^j}{\partial \mathbf{p}_j} + \frac{\partial \mathbf{z}_i^j}{\partial \mathbf{p}_j} \cdot \frac{\partial \mathbf{z}_i^j}{\partial \mathbf{p}_j} \quad (13)$$

$$\frac{\partial \mathbf{z}_i^j}{\partial \mathbf{x}_{C_i}^j} = \frac{1}{c_{i,1} \hat{Z}_j} = \begin{bmatrix} 1 & 0 & -\frac{c_{i,1} \hat{X}_j}{c_{i,1} \hat{Z}_j} \\ 0 & 1 & -\frac{c_{i,1} \hat{Y}_j}{c_{i,1} \hat{Z}_j} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (14)$$

$$\frac{\partial \mathbf{z}_i^j}{\partial \mathbf{x}_{C_i}^j} = \frac{1}{c_{i,2} \hat{Z}_j} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & -\frac{c_{i,2} \hat{X}_j}{c_{i,2} \hat{Z}_j} \\ 0 & 1 & -\frac{c_{i,2} \hat{Y}_j}{c_{i,2} \hat{Z}_j} \end{bmatrix} \quad (15)$$

$$\frac{\partial \mathbf{z}_i^j}{\partial \mathbf{x}_{C_i}^j} = \left( \begin{bmatrix} c_{i,1} \hat{\mathbf{p}}_{j^*} \end{bmatrix} \right) - C \left( \begin{bmatrix} c_{i,1} \hat{\mathbf{q}} \end{bmatrix} \right) \quad (16)$$

$$\frac{\partial \mathbf{z}_i^j}{\partial \mathbf{p}_j} = C \left( \begin{bmatrix} c_{i,1} \hat{\mathbf{q}} \end{bmatrix} \right) \quad (17)$$

$$\frac{\partial \mathbf{z}_i^j}{\partial \mathbf{x}_{C_i}^j} = C \left( \begin{bmatrix} c_{i,1} \mathbf{q} \\ c_{i,2} \mathbf{q} \end{bmatrix} \right)^T \left( \begin{bmatrix} c_{i,1} \hat{\mathbf{p}}_{j^*} \end{bmatrix} \right) - C \left( \begin{bmatrix} c_{i,1} \hat{\mathbf{q}} \end{bmatrix} \right) \quad (18)$$

$$\frac{\partial \mathbf{z}_i^j}{\partial \mathbf{p}_j} = C \left( \begin{bmatrix} c_{i,1} \mathbf{q} \\ c_{i,2} \mathbf{q} \end{bmatrix} \right)^T C \left( \begin{bmatrix} c_{i,1} \hat{\mathbf{q}} \end{bmatrix} \right) \quad (19)$$

여기서,  $\mathbf{n}_i^j$ 는 measurement noise,  $\mathbf{H}_{C_i}^j$ 와  $\mathbf{H}_{f_i}^j$ 는 measurement 자코비안을 나타낸다.

Mourikis & Roumeliotis (2007)에서  $\mathbf{p}_j$ 는 계산 시 카메라 포즈를 사용하기 때문에  $\mathbf{p}_j$ 의 불확실성은 그 상태에서 카메라 포즈와 관련이 있다는 것을 알 수 있다. 잔차에  $\mathbf{p}_j$ 의 불확실성이 영향을 미치지 않도록 Eq. (11)의 잔차는  $\mathbf{H}_{f_i}^j$ 의 null 공간인  $\mathbf{V}$ 에 투영된다. Eq. (20)은 투영된 잔차를 나타낸다. 잔차는 Mourikis & Roumeliotis (2007)에 따라 Eq. (21)로 다시 쓸 수 있다.

$$\mathbf{r}_o^j = \mathbf{V}^T \mathbf{r}^j = \mathbf{V}^T \mathbf{H}_x^j \tilde{\mathbf{x}} + \mathbf{V}^T \mathbf{n}^j = \mathbf{H}_{x,o}^j \tilde{\mathbf{x}} + \mathbf{n}_o^j \quad (20)$$

$$\mathbf{r}_n = \mathbf{Q}_1^T \mathbf{r}_o = \mathbf{T}_H \tilde{\mathbf{x}} + \mathbf{n}_n \quad (21)$$

$$\mathbf{H}_x = \begin{bmatrix} \mathbf{Q}_1 & \mathbf{Q}_2 \\ \mathbf{T}_H \\ \mathbf{0} \end{bmatrix} \quad (22)$$

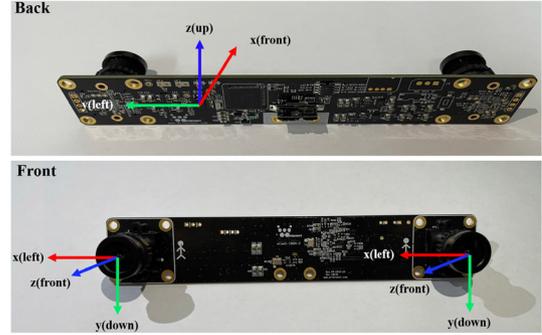


Fig. 5. Coordinate of camera and IMU.



Fig. 6. Experiment environment.

$$\mathbf{r}_o = \begin{bmatrix} \mathbf{Q}_1 & \mathbf{Q}_2 \end{bmatrix} \begin{bmatrix} \mathbf{T}_H \\ \mathbf{0} \end{bmatrix} \tilde{\mathbf{x}} + \mathbf{n}_o \quad (23)$$

$$\mathbf{n}_n = \mathbf{Q}_1^T \mathbf{n}_o \quad (24)$$

여기서,  $\mathbf{Q}_1$ 과  $\mathbf{Q}_2$ 는 각각  $\mathbf{H}_x$ 의 영공간의 기저를 형성하는 열로 구성된 단위 행렬이며,  $\mathbf{T}_H$ 는 상삼각 행렬이다.

마지막으로, 칼만 이득과 공분산 행렬 업데이트 수식은 Eqs. (25-26)과 같다.

$$\mathbf{K} = \mathbf{P} \mathbf{T}_H^T \left( \mathbf{T}_H \mathbf{P} \mathbf{T}_H^T + \mathbf{R}_n \right)^{-1} \quad (25)$$

$$\mathbf{P}_{k+1|k+1} = \left( \mathbf{I}_\xi - \mathbf{K} \mathbf{T}_H \right) \mathbf{P}_{k+1|k} \left( \mathbf{I}_\xi - \mathbf{K} \mathbf{T}_H \right)^T + \mathbf{K} \mathbf{R}_n \mathbf{K}^T \quad (26)$$

여기서,  $\xi$ 는 공분산 행렬의 차원을 나타낸다. 자세한 수식과 그에 대한 설명은 Sun et al. (2018)와 Mourikis & Roumeliotis (2007)에 자세하게 기술되어 있다.

#### 4. 실험 및 결과

본 연구에서 사용한 센서는 위드로봇의 oCamS-ICGN-U 모델의 스테레오 카메라이다. 이는 내장 IMU를 포함한다. Fig. 5는 카메라와 IMU의 좌표계를 나타낸다. 각 카메라의 좌표계는 일치하고, 오른쪽이 x축, 아래쪽이 y축, 그리고 앞쪽이 z축 방향이다. IMU의 좌표계는 앞쪽이 x축, 왼쪽이 y축, 그리고 위쪽이 z축 방향이다.

실험은 조선대학교 IT융합대학 공유교육실습실에서 진행하였다. 실험 환경은 Fig. 6과 같고 트랙 내부는 3 m × 3 m, 외부는 5 m × 5 m, 그리고 트랙의 폭은 1 m이다. 또한, 바닥과 천장에는 VIO

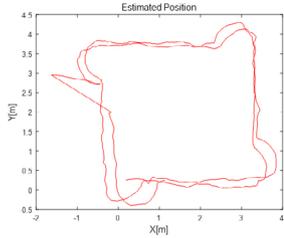


Fig. 7. Estimated position through VIO for the floor-view.

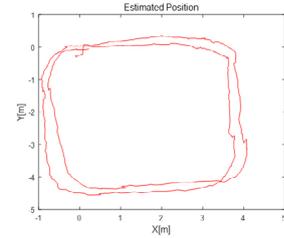


Fig. 9. Estimated position through VIO for the front-view.

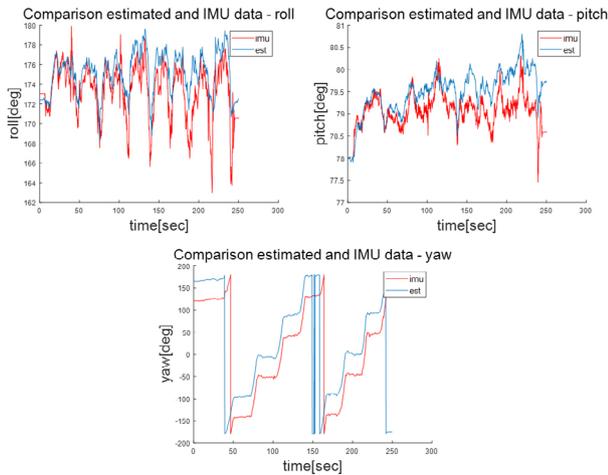


Fig. 8. Estimated attitude through VIO for the floor-view.

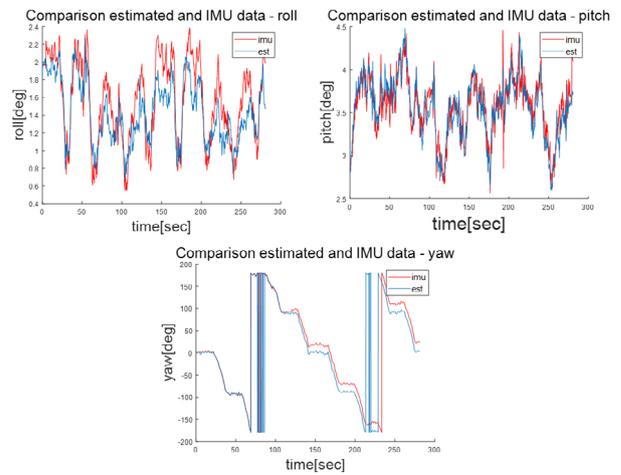


Fig. 10. Estimated attitude through VIO for the front-view.

를 위해 특징점이 필요하여 다각형의 랜드마크를 부착하였다.

#### 4.1 카메라가 바닥을 향한 경우

카메라가 바닥을 향하도록 고정한 후 실험실 트랙을 2바퀴 도는 실험을 진행하였다. 실험은 카메라가 빛의 영향을 많이 받기 때문에 블라인드를 내리고 불을 끈 상태로 진행하였다. Figs. 7과 8은 VIO를 통해 추정된 위치와 자세를 나타낸다. 추정된 위치로 확인한 경로는 약 4 m × 4 m로 출력되었다. VIO는 GPS 데이터를 사용하지 않는 추정 방법이기 때문에 위치에 대한 참값이 존재하지 않는다. 또한, 직접 센서를 부착한 카트를 밀어 실험을 진행하였으므로 정확한 성능을 확인할 수는 없다. 따라서, 본 연구에서 정확도는 추정된 경로와 실제 실험을 진행한 트랙을 비교하여 판단하였다. Fig. 7을 보면 추정된 경로와 실제 트랙을 비교해 보았을 때 거의 비슷하다는 것을 알 수 있다. 그래프 중간에 튀는 현상은 실험을 진행하며 특징점이 없어졌을 때 생기는 드리프트(drift)로 인해 발생한다. 또한, 자세의 경우 pitch가 90°가 되어 짐벌락(gimbal lock) 현상이 발생하여 추정값과 실제 IMU 측정값 사이의 오차가 약 50°로 크게 발생하였다. 오차가 약 50°로 유지되지만, 자세의 패턴은 추정값과 IMU 측정값이 거의 일치하는 것을 볼 수 있다.

#### 4.2 카메라가 정면을 향한 경우

카메라가 정면을 향한 실험도 마찬가지로 블라인드를 내리고

조명을 끈 상태에서 진행하였다. Figs. 9와 10은 VIO를 통해 추정된 위치와 자세를 나타낸다. 추정된 경로는 약 4.3 m × 4.5 m로 나타나고, 자세는 약 20°의 오차를 유지한다. 이는 카메라의 성능 및 컴퓨터 성능과 관련된다. 데이터를 저장할 때 생기는 끊김 현상으로 인해 정확도가 떨어지게 된다. 또한, 외벽에 막혀 특징점이 추출되지 않는 경우 등에도 드리프트가 발생하여 정확도가 떨어질 수 있다. 하지만, 자세의 패턴은 추정값과 측정값이 거의 일치하는 것을 볼 수 있다.

#### 4.3 카메라가 천장을 향한 경우

카메라가 천장을 향한 실험은 카메라에 빛이 정면으로 들어가지 않게 하기 위해 블라인드를 내리고 조명을 끈 상태로 진행하였다. 이 경우 또한 pitch가 -90°가 되어 짐벌락 현상이 발생한다. Figs. 11과 12는 VIO를 통해 추정된 위치와 자세를 나타낸다. 추정된 경로는 약 4 m × 4 m로 트랙 크기와 거의 일치하고, 자세 추정값은 약 15°의 오차를 보인다. 오차가 발생하는 이유는 앞서 기술한 것과 같다.

#### 4.4 성능 비교

각 카메라 방향에 대한 실험의 성능을 비교하기 위해 각각 7번의 실험을 진행하였다. Fig. 13은 카메라가 바닥을 향한 경우 7번의 실험에 대한 z축 값의 통계를 나타낸다. 실험을 진행할 때 카메라의 높이는 고정되어 있어 이론상 z축 값은 0 m이다. 하지만,

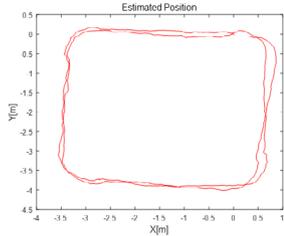


Fig. 11. Estimated position through VIO toward for the ceiling-view.

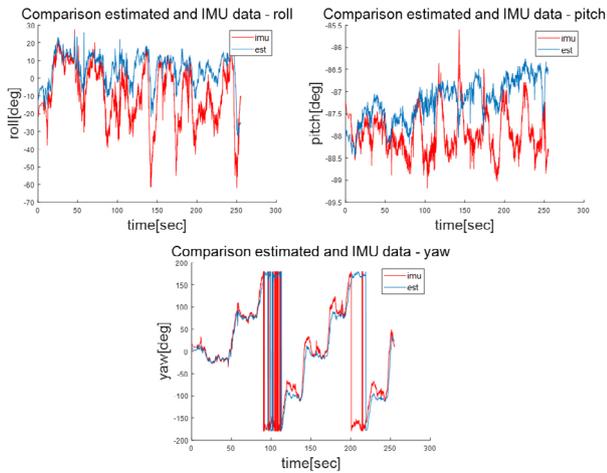


Fig. 12. Estimated attitude through VIO for the ceiling-view.

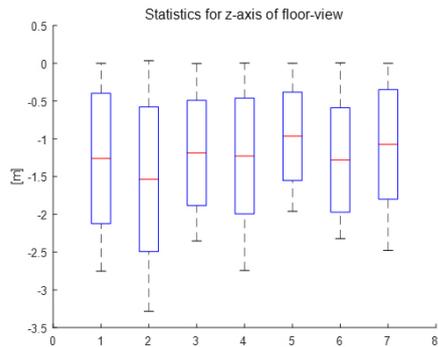


Fig. 13. Z-axis statistics for the floor-view.

7번의 실험에 대한 z축 값의 평균은 -1.2181 m로 큰 오차가 발생한 것을 볼 수 있다. 또한, 평균 최대, 최솟값은 각각 0 m와 -2.55 m이고, 평균 표준편차는 0.7552 m이다.

그래프에서 빨간 실선, 파란 직사각형, 그리고 검은 가로 실선은 각각 평균, 평균±표준편차, 그리고 최대, 최솟값을 의미한다. 카메라가 바닥을 향한 경우 발생한 오차는 트랙 바닥에 빛이 반사되어 특징점을 추출하는 데 실패하여 측정값이 MSCKF에 입력되지 않아 시간 업데이트만 진행하고, 측정 업데이트는 진행되지 않는다. 시간이 지남에 따라 공분산이 커지고 오차가 증가하여 드리프트가 발생하게 되고 이로 인해 오차가 누적되어 점점 커지게 된다.

Fig. 14는 카메라가 정면을 향한 경우 7번의 실험에 대한 z축 값의 통계를 나타낸다. 이 경우도 카메라의 높이가 고정되어 있

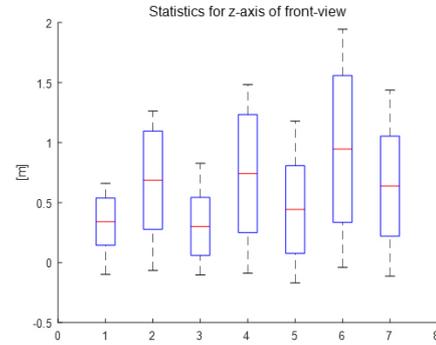


Fig. 14. Z-axis statistics for the front-view.

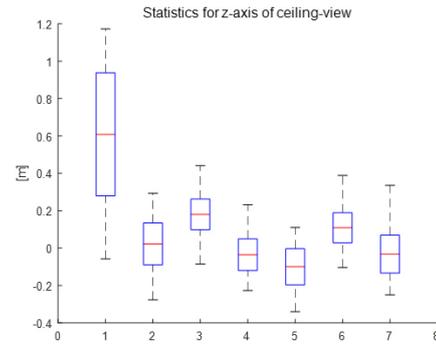


Fig. 15. Z-axis statistics for the ceiling-view.

기 때문에 z축 값은 0 m이다. 하지만, 7번의 실험에 대한 z축 값의 평균은 0.5853 m로 카메라가 정면을 향한 경우보다 더 큰 오차가 발생하였다. 또한, 평균 최대, 최솟값은 각각 1.26 m와 -0.10 m이고, 평균 표준편차는 0.3906 m이다.

카메라가 정면을 향한 경우 발생한 오차는 실험을 진행한 트랙에서 코너를 돌아갈 때 카메라가 벽에 막혀 특징점을 추출할 수 없어 발생한다. 이로 인해 드리프트가 발생하면서 오차가 점점 커지게 된다.

Fig. 15는 카메라가 천장을 향한 경우 7번의 실험에 대한 z축 값의 통계를 나타낸다. 이 경우도 마찬가지로 이론상 z축 값은 0 m이다. 하지만, 7번의 실험에 대한 z축 값의 평균은 0.1076 m로 앞의 두 경우보다 오차가 작은 것을 볼 수 있다. 또한, 평균 최대, 최솟값은 각각 0.42 m와 -0.19 m이고, 평균 표준편차는 0.1267 m이다. 앞의 두 경우에 비해 가장 좋은 성능을 보인다. 이와 같은 결과가 도출된 이유는 천장이 빛 반사가 상대적으로 적게 발생하고, 벽이 없어 특징점 추출에 실패하는 경우가 상대적으로 적기 때문이다. 그럼에도 드리프트가 발생하는데, 이는 카메라가 천장을 향하게 되면 카메라 렌즈로 빛이 직접적으로 들어와 발생하게 된다.

Figs. 13-15를 전체적으로 살펴보면, 카메라가 바닥, 정면, 그리고 천장을 향한 경우 평균은 각각 -1.2181 m, 0.5853 m, 그리고 0.1076 m로 0 m에 가장 가까운 카메라가 천장을 향한 경우 오차가 제일 적은 반면에 정면을 향한 경우 오차가 가장 큰 것을 볼 수 있다. 각 경우에 대한 최댓값과 최솟값의 차이는 각각 2.56 m, 1.35 m, 그리고 0.62 m로 카메라가 천장을 향한 경우 가장 작을 것을

알 수 있다. 마지막으로, 각 경우에 대한 표준편차의 평균은 각각 0.7552 m, 0.3906 m, 그리고 0.1267 m로 카메라가 천장을 향한 경우가 가장 적다. 따라서, 카메라가 천장을 향한 경우가 오차와 편차가 가장 작을 것을 알 수 있다.

VIO의 성능은 특징점의 개수 등에 따라 성능이 결정될 수 있다. 카메라가 바닥과 정면을 향한 경우 랜드마크의 개수가 많았음에도 VIO 성능은 카메라가 천장을 향한 경우 가장 좋게 나타났다. 이 현상은 data association과 관련이 있는데, 이는 현재 추출된 특징점을 이전 시간 이미지에서 추출한 특징점이 일치하는지 확인하는 것이다. 이 과정에서 특징점이 과하게 많으면 이전 시간에 추출된 특징점이 현재 추출된 특징점과 같은 것인지 확인하는 과정에서 정확하게 판단하지 못할 가능성이 커지게 된다. 또한, 특징점이 밀집되어 있으면 정확하지 않은 판단을 할 수 있다.

## 5. 결론

본 논문은 카메라가 다른 방향을 향하도록 하였을 때, 카메라와 IMU를 사용한 VIO를 사용한 위치 및 자세 추정의 성능을 비교하였다. 기존의 VIO는 일반적으로 카메라를 정면을 향하도록 고정하고 실험을 진행하고, 카메라가 향한 방향은 중요한 요소가 아니었다. 하지만, 본 논문은 카메라가 바닥, 정면, 그리고 천장을 향하도록 하는 실험을 진행하여 각 경우에 대한 특성을 파악하였고, 성능을 분석하였다. 카메라가 천장을 향하는 경우 가장 적은 오차를 보였고, 가장 좋은 성능을 갖는 것을 확인하였다.

향후에는, VIO 알고리즘에 짐벌락을 일으킬 수 있는 Euler 각 대신 쿼터니언(quaternion)을 사용하여 실험을 진행할 예정이다. 짐벌락을 극복한다면, 카메라 방향에 따른 VIO 성능을 본 연구 결과보다 정확한 비교 및 분석이 가능할 것으로 판단된다. 또한, VIO와 LiDAR를 동시에 사용하는 연구를 진행하여 알고리즘의 정확도를 올리는 연구를 수행할 예정이다. 카메라만 사용하는 것보다 LiDAR를 함께 사용하면 각각의 단점을 보완하여 더 나은 성능을 기대할 수 있을 것이다. 또한, VIO를 빛을 제어할 수 없는 실외에서도 진행하여 성능을 비교할 예정이다.

## ACKNOWLEDGMENTS

이 논문은 조선대학교 학술연구비의 지원을 받아 연구되었음 (2024년).

## AUTHOR CONTRIBUTIONS

Conceptualization, B.Lee and N.Y.Ko; methodology, B.Lee and N.Y.Ko; software, B.Lee and N.Y.Ko; validation, B.Lee and N.Y.Ko; formal analysis, B.Lee and N.Y.Ko; investigation, B.Lee and N.Y.Ko; resources, B.Lee and N.Y.Ko; data curation, B.Lee and N.Y.Ko; writing—original draft preparation, B.Lee; writing—review and editing,

N.Y.Ko; visualization, B.Lee; supervision, N.Y.Ko; project administration, N.Y.Ko; funding acquisition, N.Y.Ko.

## CONFLICTS OF INTEREST

The authors declare no conflict of interest.

## REFERENCES

- Chen, P., Guan, W., & Lu, P. 2023, ESVIDEO: Event-Based Stereo Visual Inertial Odometry, *IEEE Robotics and Automation Letters*, 8, 3661-3668. <https://doi.org/10.1109/LRA.2023.3269950>
- Do, H., V. Kim, Y., H., Kwon, Y., S., Kang, S., H., & Kim, H., J., et al. 2021, An Adaptive Approach based on Multi-State Constraint Kalman Filter for UAVs, *21st International Conference on Control, Automation and Systems*, 12-15 October 2021, Jeju, Korea, pp.481-485. <https://doi.org/10.23919/ICCAS52745.2021.9649897>
- Gurturk, M., Yusefi, A., Aslan, M. F., Soyacan, M., & Durdu, A., et al. 2021, The YTU dataset and recurrent neural network based visual-inertial odometry, *Measurement*, 184. <https://doi.org/10.1016/j.measurement.2021.109878>
- Jennings, S. & Eshghi, S. 2019, Stereo Camera-IMU Calibration, *ResearchGate*. <https://doi.org/10.13140/RG.2.2.23257.65123>
- Jung, G.-W. & Ryu, J.-H. 2019, Fusion of a Particle Filter and an Extended Kalman Filter for Real-time High-Precision Localization of Autonomous Vehicles, *Journal of Institute of Control, Robotics and Systems*, 25, 526-533. <https://doi.org/10.5302/J.ICROS.2019.19.0044>
- Kottas, D. G. & Roumeliotis, S. I. 2013, Efficient and Consistent Vision-aided Inertial Navigation using Line Observations, *2013 IEEE International Conference on Robotics and Automation*, 06-10 May 2013, Karlsruhe, Germany, pp.1540-1547. <https://doi.org/10.1109/ICRA.2013.6630775>
- Lee, J.-H. & Choe, J.-K. 2013, A Study on Self-localization for Autonomous Mobile Robot, *Journal of KIIT*, 12, 29-34. <https://doi.org/10.14801/kiitr.2013.11.12.29>
- Maye, J., Furgale, P., & Siegwart, R. 2013, Self-supervised Calibration for Robotic Systems, *2013 IEEE Intelligent Vehicles Symposium*, 23-26 June 2013, Gold Coast, QLD, Australia, pp.473-480. <https://doi.org/10.1109/IVS.2013.6629513>
- Memon, Q. & Khan, S. 2001, Camera Calibration and Three-dimensional World Reconstruction of Stereo-vision Using Neural Networks, *International Journal of Systems Science*, 32, 1155-1159. <https://doi.org/10.1080/00207720010024276>
- Mourikis, A. I., & Roumeliotis, S. I. 2007, A Multi-State Constraint

Kalman Filter for Vision-aided Inertial Navigation, 2007 IEEE International Conference on Robotics and Automation, 10-14 April 2007, Rome, Italy, pp.3565-3572. <https://doi.org/10.1109/ROBOT.2007.364024>

Rehder, J., Nikolic, J., Schneider, T., Hinzmann, T., & Siegwart, R. 2016, Extending kalibr: Calibrating the Extrinsic of Multiple IMUs and of Individual Axes, 2016 IEEE International Conference on Robotics and Automation, 16-21 May 2016, Stockholm, Sweden, pp.4304-4311. <https://doi.org/10.1109/ICRA.2016.7487628>

Sun, K., Mohta, K., Pfrommer, B., Watterson, M., Liu, S., et al. 2018 Robust Stereo Visual Inertial Odometry for Fast Autonomous Flight, IEEE Robotics and Automation Letters, 3, 965-972. <https://doi.org/10.1109/LRA.2018.2793349>

Weng, J., Cohen, P., & Herniou, M. 1990, Calibration of Stereo Cameras Using a Non-linear Distortion Model, Proceedings 10th International Conference on Pattern Recognition, 16-21 June 1990, Atlantic City, NJ, USA, pp.246-253. <https://doi.org/10.1109/ICPR.1990.118105>

Weng, J., Cohen, P., & Herniou, M. 1992, Camera Calibration with Distortion Models and Accuracy Evaluation, IEEE Transactions on Pattern Analysis and Machine Intelligence, 14, 965-980. <https://doi.org/10.1109/34.159901>

Yang, Y., Geneva, P., Eckenhoff, K., & Huang, G. 2019, Visual-Inertial Odometry with Point and Line Features, 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems, 03-08 November 2019, Macau, China, pp.2447-2454. <https://doi.org/10.1109/IROS40897.2019.8967905>

Zhang, C., Chen, L., & Yuan, S. 2020, ST-VIO: Visual-Inertial Odometry Combined with Image Segmentation and Tracking, IEEE Transactions on Instrumentation and Measurement, 69, 8562-8570. <https://doi.org/10.1109/TIM.2020.2989877>

Gwangju, Korea, since 1992. He has worked as a Visiting Research Scientist with The Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, USA, from 1997 to 1998 and from 2004 to 2005. He is involved in projects developing unmanned aerial vehicles and underwater vehicles, especially in the development of navigation algorithms. His research interests include navigation, collision avoidance, and motion planning in robotics. Prof. Ko is a member of IEEE Robotics and Automation Society and Oceanic Engineering Society and a member of the Institute of Control, Robotics and Systems, Korea.



**Boeun Lee** received the B.S. and M.S. degrees from the Department of Electronic Engineering, Chosun University, Gwangju, Korea, in 2021 and 2022, respectively. She is currently pursuing the Ph.D. degree at the same university. Her research interests include navigation, visual inertial odometry, and control.



**Nak Yong Ko** received the B.S., M.S., and Ph.D. degrees from the Department of Control and Instrumentation Engineering, Seoul National University, Seoul, Korea, in 1985, 1987, and 1993, respectively. He has been a Professor with the Department of Electronic Engineering, Chosun University,

