

소실점 정보의 Loss 함수를 이용한 특징선 기반 SLAM

Line-Based SLAM Using Vanishing Point Measurements
Loss Function임 현 준¹, 명 현[†]Hyunjun Lim¹, Hyun Myung[†]

Abstract: In this paper, a novel line-based simultaneous localization and mapping (SLAM) using a loss function of vanishing point measurements is proposed. In general, the Huber norm is used as a loss function for point and line features in feature-based SLAM. The proposed loss function of vanishing point measurements is based on the unit sphere model. Because the point and line feature measurements define the reprojection error in the image plane as a residual, linear loss functions such as the Huber norm is used. However, the typical loss functions are not suitable for vanishing point measurements with unbounded problems. To tackle this problem, we propose a loss function for vanishing point measurements. The proposed loss function is based on unit sphere model. Finally, we prove the validity of the loss function for vanishing point through experiments on a public dataset.

Keywords: Visual SLAM, Line Feature, Vanishing Point

1. 서 론

Global Navigation Satellite System (GNSS) 정보를 이용할 수 없는 실내외 환경에서 로봇의 위치 인식을 위해 동시적 위치추정 및 지도작성(SLAM) 연구가 활발히 진행되었다. 그 중에서도, 카메라를 사용하는 visual SLAM 알고리즘은 특징점을 기반으로 하는 연구가 주로 진행되었다. 하지만, 특징점은 복도와 같이 특징점이 적은 환경에서 잘 추출되지 못한다. 또한, 특징점은 희소하기 때문에 작성된 지도가 구조적 특성을 나타내기 어렵다^[1].

특징점의 단점을 보완하기 위하여, 특징선을 추가로 사용하는 방법들이 제안되었다. 특징선은 복도와 같은 환경에서 중요한 정보로 사용될 수 있다. 또한, 선은 여러 개의 점으로 이루어져 있기 때문에, 빛이 변화하는 환경에서도 강인하다.

마지막으로, 특징선은 구조적 특징을 가지고 있기 때문에, 3차원 지도작성을 통해 구조물에 대한 정보를 확인할 수 있다^[2].

특징선을 visual SLAM에서 사용하기 위하여, 기존 연구에서는 선 측정 모델을 선의 재투영 오차로 정의했다^[3]. 하지만, 3D 선은 4 자유도를 가지기 때문에, 2 자유도를 가지는 선 측정 모델만으로는 관측 가능성을 보장하지 못한다. 이를 해결하기 위해, 우리의 이전 연구에서는 소실점 측정 모델을 제안했다^[4]. 소실점 측정 모델은 카메라 이미지에서 관측된 2D 특징선들을 통해 관측할 수 있는 소실점과, 3차원 특징선으로부터 계산할 수 있는 소실점과의 차이를 잔차로 정의하여 사용한다.

Visual SLAM의 관측 모델들은 이상점을 제거하기 위하여 손실 함수를 사용한다. 일반적인 특징점과 특징선의 재투영 모델의 경우, Huber나 Cauchy 손실 함수를 사용한다. 하지만, 소실점 측정 모델을 일반적인 정규화 이미지 평면의 잔차로 정의하는 경우, 무한대의 값을 가지는 문제가 발생한다. 이는, 최적화 과정에서 특징선의 변수가 변하면서 잔차가 급격하게 변화하는 현상이 발생할 수 있다. UV-SLAM^[4]에서는 소실점 측정 모델의 손실 함수에 대한 분석이 없었다. 이 문제를 해결하기 위해, 본 논문에서는 소실점 측정 모델에 맞는 새로운 손실 함수를 제안한다.

Received : May. 10. 2023; Accepted : Jun. 7. 2023

※ This work was financially supported in part by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No. 2021-0-00230, development of real · virtual environmental analysis based adaptive interaction technology). The students are supported by BK21 FOUR

1. Ph.D Candidate, Electrical Engineering, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea (tp02134@kaist.ac.kr)

† Professor, Corresponding author: Electrical Engineering, KAIST, Daejeon, Korea (hmyung@kaist.ac.kr)

2. 관련 연구

2.1 특징선 기반 SLAM

특징선을 visual SLAM에 사용하는 많은 연구들이 있다. Bartoli et al.^[5]은 특징선을 visual SLAM에서 사용하기 위해서 3차원 표현 방법으로 Plücker coordinates와 orthonormal representation을 제안했다. 이 표현법들을 기반으로, 선 측정 모델은 점 측정 모델과 같이 정의되었다. 선 측정 모델은 추정된 3차원 선을 재투영하여 새롭게 관측된 특징선과의 차이를 잔차로 계산했다. 대부분의 특징선 기반 알고리즘들은 특징점 기반 알고리즘을 기반으로 특징선을 추가했다. 필터링 기반 알고리즘^[6,7]들은 주로 MSCKF^[8]를 기반으로 개발되었다. 이와는 다르게, 최적화 기반의 알고리즘^[14, 9-11]들은 ORB-SLAM^[12]과 VINS-Mono^[13]를 기반으로 개발되었다.

2.2 Manhattan 혹은 Atlanta world 가정을 사용하는 특징선 기반 SLAM

특징선과 함께, 소실점을 visual SLAM에 사용하는 연구들이 있다. 먼저, Kim et al.^[14]은 소실점을 이용하여 카메라의 자세를 추정하였다. 또한, 카메라의 자세를 추정한 후 위치를 추정하는 비결합 방법을 제안한 연구들이 있었다^[15-17]. 이에 더하여, Manhattan 혹은 Atlanta world 가정법을 사용하여 소실점을 visual SLAM에 적용한 방법들이 있다^[18,19]. 이 방법들은 구조적 규칙성을 가지고 있는 특징선만을 사용하기 위해 새로운 2-DoF의 선 표현법을 사용했다. 하지만, 이러한 방법들은 구조적인 선들만을 사용하기 때문에 가정이 성립하는 실내에서만 사용이 가능하다는 단점이 있다.

2.3 소실점 측정 모델

Visual SLAM에서 소실점 측정 모델을 정의하여 사용한 방법들이 있다. 이 연구에서는 소실점을 통해 평행한 선들끼리 구별한다^[20]. 그 다음, 평행하는 선들이 한 평면에 있어야 한다는 조건과 외적이 0이 되어야 한다는 조건을 사용하여 잔차를 구성했다. 하지만, 해당 논문에서는 소실점을 평행한 선들을 찾는 방법으로만 사용하고, 소실점이 아닌 추정된 3차원 선들 간의 오차로만 잔차를 정의했다. 따라서, 최적화 전의 3차원 선이 부정확한 경우 정확한 선 지도 결과를 얻기 어렵다. 소실점을 최적화 과정에서 측정 모델로 직접 사용한 연구가 있다^[21]. 이 연구에서는 소실점 측정 모델을 제안했으며, 무한한 잔차가 발생하는 문제를 해결하기 위한 손실 함수 또한 제안했다. 하지만, 이 논문에서 제안한 손실 함수는 소실점 측정 모델에 적절하지 못하다는 단점이 있다.

3. 제안 방법

3.1 프레임워크

제안 방법의 전체적인 구조는 [Fig. 1]과 같다. 제안 방법은 VINS-Mono를 기반으로 하여 IMU와 특징점 모델을 사용한다. 특징점은 Shi-Tomasi^[22] 방법을 통해 추출되고, KLT tracker^[23]를 사용하여 추적한다. 이에 더하여, IMU 측정 모델은 pre-integration 방법^[24]을 통해 정의한다. 마지막으로, 최적화 기법을 사용하기 위해서 Schur complement^[25]를 통한 marginalization을 사용한다.

이와 같은 VINS-Mono 기반의 단안 visual SLAM에 특징선을 추가하기 위하여, 추출 및 매칭을 위해 각각 LSD^[26]와 LBD^[27]를 사용한다. 관측된 특징선은 삼각법을 통해 3차원 특징선으로 만들 수 있다. 3차원 특징점은 (x, y, z) 의 형태로 표현할 수 있는 것에 반해, 3차원 특징선은 표현하는 방법이 필요하다. 우리는 Bartoli et al.^[5]에서 사용한 Plücker coordinates와 orthonormal representation을 사용한다. Plücker coordinates는 3차원 특징선을 표현할 수 있는 직관적인 방법으로 다음과 같다:

$$\mathbf{L}(\mathbf{n}, \mathbf{d})^T \in \mathbb{R}^6, \quad (1)$$

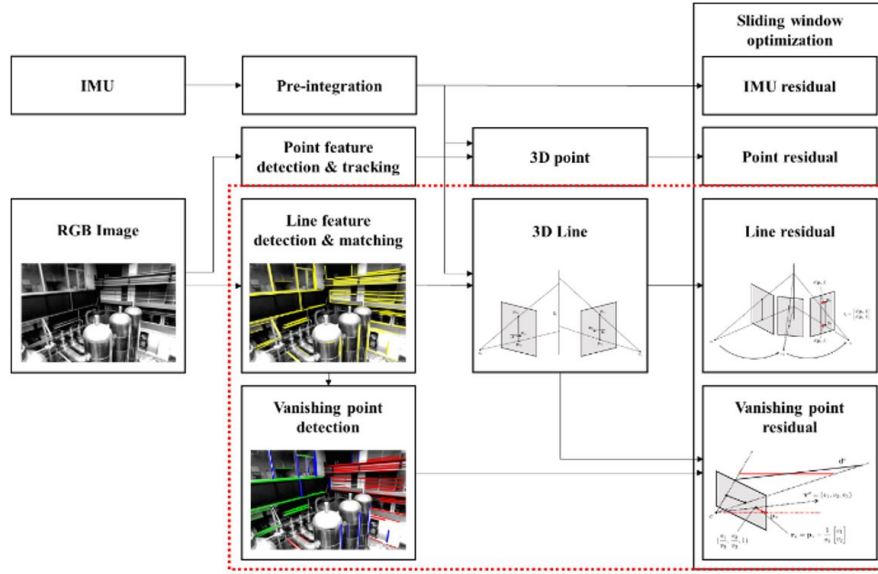
이때 \mathbf{n} 과 \mathbf{d} 는 각각 수직과 방향 벡터를 표현한다. Plücker coordinates는 triangulation 및 재투영 과정에서 사용한다. 하지만, 3차원 특징선이 4-DoF인 것에 반해, Plücker coordinates로 표현된 3차원 특징선은 6-DoF다. 따라서, 최적화 과정에서 over-parameterization이 발생한다. 이 문제를 해결하기 위해, 최적화 과정에서는 4-DoF를 가지는 orthonormal representation을 사용하고, 다음과 같다:

$$\mathbf{o} = [\psi, \phi], \quad (2)$$

이때 ψ 와 ϕ 는 각각 카메라 좌표계를 기준으로 하는 3차원 특징선 자세의 Euler 각과 3차원 특징선과 카메라 원점까지의 최소 거리를 의미한다.

우리가 제안하는 방법은 특징선과 함께 특징선에서 얻을 수 있는 소실점을 사용한다. 소실점은 특징선이 구조적 규칙성을 가지고 있는지 판단할 수 있는 정보를 가지고 있다. 우리는 소실점 추출을 위해 J-linkage^[28]를 사용한다. 이 방법은 잡음과 이상점이 존재하는 상황에서 여러가지 후보군을 찾을 수 있는 방법이다. 이를 통해, Manhattan world 가정 없이 더 많은 개수의 소실점을 찾을 수 있다는 장점이 있다.

이 논문에서 $(\cdot)^w$, $(\cdot)^c$, 그리고 $(\cdot)^b$ 는 각각 world, camera, 그리고 body 좌표계를 의미한다. 이에 더하여 $(\cdot)_c^b$ 는 body에서 camera 좌표계로의 변환행렬을 의미한다. 이 시스템에서 사용



[Fig. 1] The proposed algorithm's framework is depicted in a block diagram, where newly added blocks are represented by dashed boxes. Upon receiving an RGB image, the algorithm performs line feature detection and matching, followed by vanishing point detection and line clustering. The triangulation process creates 3D lines, and residuals of lines and vanishing points are defined. Eventually, SLAM results can be achieved by employing sliding window optimization

하는 상태 벡터는 다음과 같다:

$$\begin{aligned} \chi = & [\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{I-1}, \\ & \lambda_0, \lambda_1, \dots, \lambda_{J-1}, \\ & \mathbf{o}_0, \mathbf{o}_1, \dots, \mathbf{o}_{K-1}], \\ \mathbf{x}_i = & [\mathbf{p}_i^w, \mathbf{q}_i^w, \mathbf{v}_i^w, \mathbf{b}_a, \mathbf{b}_g], i \in [0, I-1], \\ \mathbf{o}_k = & [\psi_k, \phi_k], k \in [0, K-1], \end{aligned} \quad (3)$$

이때 χ 는 전체 상태 벡터, 그리고 \mathbf{x}_i 는 위치, 자세, 속도, 그리고 bias들로 이루어진 i 번째 sliding window의 body 상태 벡터를 의미한다. 이에 더하여, 전체 상태 벡터는 $\lambda_j, j \in [0, J-1]$ 로 표현된 특징점의 역 깊이를 포함하고 있다. 이 논문에서는 \mathbf{o} 로 표현된 특징선의 orthonormal representation이 새롭게 추가되었다. I, J , 그리고 K 는 각각 sliding window, 특징점, 그리고 특징선의 숫자를 의미한다.

식 (3)에서 정의한 상태 벡터를 사용한 최적화 과정의 비용 함수는 다음과 같다:

$$\begin{aligned} \min_{\chi} \{ & \|\mathbf{r}_0 - \mathbf{J}_0 \chi\|^2 \\ & + \sum_{i \in B} \|\mathbf{r}_I(\mathbf{z}_{b_{i+1}}^i, \chi)\|^2_{\Sigma_{b_{i+1}}^i} + \sum_{(i,j) \in P} \rho_p \|\mathbf{r}_p(\mathbf{z}_{p_j}^i, \chi)\|^2_{\Sigma_{p_j}^i} \\ & + \sum_{(i,k) \in L} \rho_l \|\mathbf{r}_l(\mathbf{z}_{l_k}^i, \chi)\|^2_{\Sigma_{l_k}^i} + \sum_{(i,k) \in V} \rho_v \|\mathbf{r}_v(\mathbf{z}_{v_k}^i, \chi)\|^2_{\Sigma_{v_k}^i} \}, \end{aligned} \quad (4)$$

이때 $\mathbf{r}_0, \mathbf{r}_I, \mathbf{r}_p, \mathbf{r}_l$, and \mathbf{r}_v 는 각각 marginalization, IMU, 특징점,

특징선, 그리고 소실점 측정 모델에 대한 잔차를 의미한다.

$\mathbf{z}_{b_{i+1}}^i, \mathbf{z}_{p_j}^i, \mathbf{z}_{l_k}^i$, 그리고 $\mathbf{z}_{v_k}^i$ 는 각각 IMU, 특징점, 특징선, 그리고 소실점에 대한 관측을 의미한다. B, P, L , 그리고 V 는 각각 sliding window 내 IMU pre-integration, 특징점, 특징선, 그리고 소실점 측정에 대한 집합을 의미한다. $\Sigma_{b_{i+1}}^i, \Sigma_{p_j}^i, \Sigma_{l_k}^i$, 그리고 $\Sigma_{v_k}^i$ 는 각각 IMU, 특징점, 특징선, 그리고 소실점 측정에 대한 공분산 행렬을 의미한다. ρ_p, ρ_l, ρ_v 는 각각 특징점, 특징선, 그리고 소실점의 손실 함수를 의미한다. 최적화 과정에서는 Ceres Solver^[29]를 사용한다.

3.2 특징선 측정 모델

특징선 측정 모델을 정의하기 위해서, Plücker coordinates로 정의된 3차원 특징선의 재투영은 다음과 같다:

$$\begin{aligned} \mathbf{l}^c = & \begin{bmatrix} l_1 \\ l_2 \\ l_3 \end{bmatrix} = \mathbf{K}' \mathbf{n}^c = f_x f_y (\mathbf{K}^{-1})^\top \mathbf{n}^c \\ = & \begin{bmatrix} f_y & 0 & 0 \\ 0 & f_x & 0 \\ -f_y c_x & -f_x c_y & f_x f_y \end{bmatrix} \mathbf{n}^c = \mathbf{n}^c, \end{aligned} \quad (5)$$

이때 $\mathbf{l}^c, \mathbf{K}'$, 그리고 \mathbf{K} 는 각각 재투영된 특징선, 특징선의 투영 행렬, 그리고 카메라의 내부 파라미터를 의미한다. (f_x, f_y) 와 (c_x, c_y) 는 각각 카메라 이미지의 초점거리와 주점을 의미

한다. 우리가 제안한 알고리즘에서는 정규화된 이미지를 사용하기 때문에, \mathbf{K} 와 \mathbf{K}' 는 단위행렬과 같다. 결과적으로, 재투영된 특징선은 3차원 특징선의 수직벡터와 동일하다.

재투영된 특징선을 추정값으로 사용하여 정의한 특징선 추정 모델의 잔차는 다음과 같다:

$$\mathbf{r}_l = \begin{bmatrix} d(\mathbf{p}_s, \mathbf{l}^c) \\ d(\mathbf{p}_e, \mathbf{l}^c) \end{bmatrix}, \quad (6)$$

이때

$$d(\mathbf{p}, \mathbf{l}^c) = \frac{\mathbf{p}^\top \mathbf{l}^c}{l_d}, l_d = \sqrt{l_1^2 + l_2^2}, \quad (7)$$

$$\mathbf{p}_s = (u_s, v_s, 1), \mathbf{p}_e = (u_e, v_e, 1),$$

그리고 \mathbf{r}_l 과 d 는 각각 특징선의 잔차와 점과 직선 사이의 거리를 의미한다. \mathbf{p}_s 와 \mathbf{p}_e 는 각각 이미지에서 관측된 특징선의 양 끝 점을 의미한다.

특징선 모델의 잔차에 상응하는 자코비안 행렬을 계산하기 위해 body 상태의 극소 변화 $\delta \mathbf{x}$ 와 orthonormal representation로 표현된 특징선의 변화 $\delta \mathbf{o}$ 로 표현하면 다음과 같다:

$$\mathbf{J}_l = \frac{\partial \mathbf{r}_l}{\partial \mathbf{l}^c} \frac{\partial \mathbf{l}^c}{\partial \mathbf{L}^c} \begin{bmatrix} \frac{\partial \mathbf{L}^c}{\partial \delta \mathbf{x}} & \frac{\partial \mathbf{L}^c}{\partial \mathbf{L}^w} \frac{\partial \mathbf{L}^w}{\partial \delta \mathbf{o}} \end{bmatrix}, \quad (8)$$

이때

$$\frac{\partial \mathbf{r}_l}{\partial \mathbf{l}^c} = \begin{bmatrix} \frac{-l_1(\mathbf{p}_s^\top \mathbf{l}^c)}{l_d^3} + \frac{u_s}{l_d} & \frac{-l_2(\mathbf{p}_s^\top \mathbf{l}^c)}{l_d^3} + \frac{v_s}{l_d} & \frac{1}{l_d} \\ \frac{-l_1(\mathbf{p}_e^\top \mathbf{l}^c)}{l_d^3} + \frac{u_e}{l_d} & \frac{-l_2(\mathbf{p}_e^\top \mathbf{l}^c)}{l_d^3} + \frac{v_e}{l_d} & \frac{1}{l_d} \end{bmatrix}_{2 \times 3}, \quad (9)$$

$$\frac{\partial \mathbf{l}^c}{\partial \mathbf{L}^c} = [\mathbf{K}' \quad \mathbf{0}_{3 \times 3}]_{3 \times 6},$$

$$\frac{\partial \mathbf{L}^c}{\partial \delta \mathbf{x}} =$$

$$(\tau_c^b)^{-1} \begin{bmatrix} (\mathbf{R}_b^w)^\top [\mathbf{d}^w]_\times & [(\mathbf{R}_b^w)^\top (\mathbf{n}^w + [\mathbf{d}^w]_\times \mathbf{p}_b^w)]_\times \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{6 \times 3} & \mathbf{0}_{6 \times 3} & \mathbf{0}_{6 \times 3} \end{bmatrix}_{6 \times 15},$$

$$\frac{\partial \mathbf{L}^c}{\partial \mathbf{L}^w} \frac{\partial \mathbf{L}^w}{\partial \delta \mathbf{o}} =$$

$$(\tau_c^b)^{-1} \begin{bmatrix} \mathbf{0}_{3 \times 1} & -w_1 \mathbf{u}_3 & w_1 \mathbf{u}_2 & -w_2 \mathbf{u}_1 \\ w_2 \mathbf{u}_3 & \mathbf{0}_{3 \times 1} & -w_2 \mathbf{u}_1 & w_1 \mathbf{u}_2 \end{bmatrix}_{6 \times 4},$$

$$\mathbf{U} = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3] = \begin{bmatrix} \frac{\mathbf{n}}{\|\mathbf{n}\|} & \frac{\mathbf{d}}{\|\mathbf{d}\|} & \frac{\mathbf{n} \times \mathbf{d}}{\|\mathbf{n} \times \mathbf{d}\|} \end{bmatrix}, \quad (10)$$

$$\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \frac{1}{\sqrt{\|\mathbf{n}\|^2 + \|\mathbf{d}\|^2}} \begin{bmatrix} \|\mathbf{n}\| \\ \|\mathbf{d}\| \end{bmatrix},$$

그리고 τ_c^b 는 Plücker coordinates에서 카메라 좌표계에서 body 좌표계로의 변환 행렬이다.

3.3 소실점 측정 모델

이 논문에서는 특징선에서 얻을 수 있는 소실점을 관측된 정보로 사용한다. 특징선 측정 모델과 비슷하게, 3차원 특징선으로부터 소실점을 추정한다. 소실점을 추정하기 위해서, 3차원 특징선 위 한 점을 동차 좌표계로 표현하면 다음과 같다:

$$\mathbf{V}(t) = \mathbf{V}_0 + t\mathbf{D} = \begin{bmatrix} x_0 + td_1 \\ y_0 + td_2 \\ z_0 + td_3 \\ 1 \end{bmatrix}, t \in (0, \infty), \quad (11)$$

이때

$$\mathbf{V}_0 = [x_0 \quad y_0 \quad z_0 \quad 1]^\top, \quad (12)$$

$$\mathbf{D} = [\mathbf{d}^\top, 0]^\top = [d_1, d_2, d_3, 0]^\top,$$

여기서 \mathbf{V}_0 는 3차원 특징선 위의 한 점을 의미한다. 소실점의 정의처럼 3차원 특징선 위 무한대의 점을 투영시키면 다음과 같다:

$$\mathbf{v}^c = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \lim_{t \rightarrow \infty} \mathbf{P}(\mathbf{V}_0 + t\mathbf{D}) = \mathbf{K}\mathbf{d}^c = \mathbf{d}^c, \quad (13)$$

이때 $\mathbf{P} = \mathbf{K}[\mathbf{I} \mid \mathbf{0}]$ 는 카메라 투영 행렬이다. 마찬가지로, 우리가 제안하는 알고리즘에서는 정규화된 이미지를 사용하기 때문에, 추정된 소실점은 3차원 특징선의 방향벡터와 같다. 최종적으로, 소실점 측정 모델에 대한 잔차는 다음과 같다:

$$\mathbf{r}_v = \mathbf{P}_v - \frac{1}{v_3} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}, \quad (14)$$

이때, \mathbf{r}_v 과 \mathbf{P}_v 는 각각 소실점 측정 모델의 잔차와 이미지에서 관측된 특징선으로부터 계산한 소실점 관측 정보다.

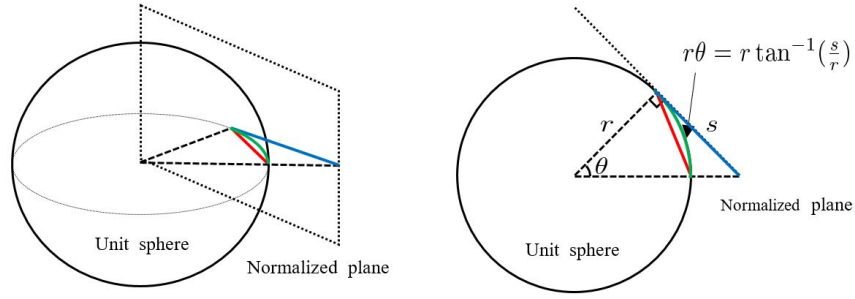
소실점 모델의 잔차에 상응하는 자코비안 행렬을 $\delta \mathbf{x}$ 와 $\delta \mathbf{o}$ 를 통해 표현하면 다음과 같다:

$$\mathbf{J}_v = \frac{\partial \mathbf{r}_v}{\partial \mathbf{v}^c} \frac{\partial \mathbf{v}^c}{\partial \mathbf{L}^c} \begin{bmatrix} \frac{\partial \mathbf{L}^c}{\partial \delta \mathbf{x}} & \frac{\partial \mathbf{L}^c}{\partial \mathbf{L}^w} \frac{\partial \mathbf{L}^w}{\partial \delta \mathbf{o}} \end{bmatrix} \quad (15)$$

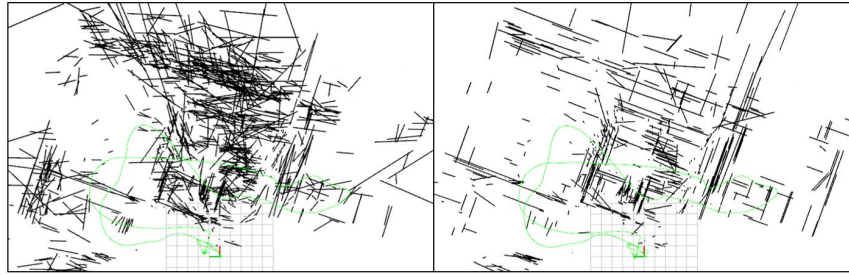
이때

$$\frac{\partial \mathbf{r}_v}{\partial \mathbf{v}^c} = \begin{bmatrix} -\frac{1}{v_3} & 0 & \frac{v_1}{v_3^2} \\ 0 & -\frac{1}{v_3} & \frac{v_2}{v_3^2} \end{bmatrix}_{2 \times 3}, \quad (16)$$

$$\frac{\partial \mathbf{v}^c}{\partial \mathbf{L}^c} = [\mathbf{0}_{3 \times 3} \quad \mathbf{K}]_{3 \times 6}.$$



[Fig. 2] Illustration of the residuals from two different views: a perspective view and a top view. The green arc represents the inverse tangent loss function proposed in this paper



[Fig. 3] The top views of mapping results of ALVIO (left) and proposed (right) for MH_04_difficult in the EuRoC datasets

제안하는 소실점 관측 모델은 일반적인 정규화 이미지 평면이 아닌, 단위 구에서의 잔차로 정의하고 [Fig. 2]와 같다. 또한, 이를 식으로 표현하면 다음과 같다:

$$\|\mathbf{r}\| = r\theta = r \tan^{-1}\left(\frac{s}{r}\right) = \tan^{-1}s, \quad (17)$$

이때 \mathbf{r} , r , s , 그리고 θ 는 각각 단위 구에서의 잔차, 단위 구의 반지름, 정규화 이미지 평면에서의 잔차, 그리고 소실점 관측과 추정 사이의 각도 차이를 의미한다. 최종적으로, 우리는 소실점 측정 모델의 손실 함수로 역탄젠트 함수를 사용한다.

4. 실험 방법 및 결과

실험은 32GB RAM이 장착된 Intel Core i7-9700K 프로세서 PC에서 수행되었다. 우리가 제안하는 소실점 측정 모델의 손실 함수가 타당함을 입증하기 위해, 영상 관측 오도메트리의 성능 평가에서 주로 사용되는 EuRoC dataset^[30]을 사용했다. 이 데이터셋은 조명, 텍스처 및 MAV 속도와 같은 다양한 요인을 포함하고 있다. 비교 알고리즘으로는 우리가 제안하는 방법이 기반으로 하는 VINS-Mono를 선정했다. 또한, 특징선을 사용하는 알고리즘인 PL-VINS, ALVIO^[2], 그리고 우리가 이전에 제안했던 알고리즘^[3]과 비교했다. 우리는 위치 정확도 비교를 위해 rpg trajectory evaluation tool^[31]을 사용했다.

[Table 1] Translational RMSE without loop closing for the EuRoC dataset (Unit: m)

Translation RMSE	VINS-Mono	PL-VINS	ALVIO	Previous work	Proposed
MH_01_easy	0.159	0.164	0.148	0.142	0.139
MH_02_easy	0.14	0.174	0.136	0.126	0.094
MH_03_medium	0.225	0.187	0.209	0.198	0.189
MH_04_difficult	0.408	0.335	0.389	0.301	0.261
MH_05_difficult	0.312	0.347	0.317	0.293	0.188
V1_01_easy	0.094	0.071	0.085	0.087	0.067
V1_02_medium	0.115	0.086	0.075	0.072	0.07
V1_03_difficult	0.203	0.152	0.2	0.156	0.109
V2_01_easy	0.099	0.09	0.094	0.098	0.085
V2_02_medium	0.161	0.12	0.133	0.103	0.112
V2_03_difficult	0.341	0.278	0.288	0.277	0.213

[Table 1]은 EuRoC dataset에서 각 알고리즘의 평균 제공근 오차에 대한 표다. 제안한 방법은 VINS-Mono, PL-VINS, ALVIO, 그리고 우리가 이전에 제안했던 알고리즘^[3]보다 약 32.3%, 23.8%, 26.4%, 그리고 17.6% 더 작은 평균 제공근 오차를 보였다. 이는 다른 최신 알고리즘과 다르게, 소실점 측정 모델 및 적절한 손실 함수를 사용했기 때문에 더 좋은 결과를 얻을 수 있었다.

[Fig. 3]은 EuRoC dataset의 MH_04_difficult에서 ALVIO와

우리가 제안한 알고리즘의 특징선 지도 작성 결과다. ALVIO의 경우 특징선들이 부정확한 것을 확인할 수 있다. 이와는 다르게, 우리가 제안한 알고리즘에선 소실점 측정 모델을 사용하기 때문에, 3차원 특징선들이 구조적 규칙성을 가지며 정확하게 지도 작성이 되는 것을 확인할 수 있다.

5. 결 론

이 논문에서는 소실점 측정 모델을 위한 손실 함수를 제안하여 특징선 기반 SLAM의 정확도 및 지도 작성의 정확도를 향상시켰다. 추후에는, 추정된 위치 및 3차원 특징점 및 특징선 정보를 통해 3차원 지도를 재구성하는 연구를 진행할 예정이다.

References

- [1] Y. He, J. Zhao, Y. Guo, W. He, and K. Yuan, "PL-VIO: Tightly coupled monocular visual-inertial odometry using point and line features," *Sensors*, vol. 18, no. 4, pp. 1159, Apr., 2018, [Online], <https://www.mdpi.com/1424-8220/18/4/1159>.
- [2] K. Jung, Y. Kim, H. Lim, and H. Myung, "ALVIO: Adaptive line and point feature-based visual inertial odometry for robust localization in indoor environments," *International Conference on Robot Intelligence Technology and Applications (RiTA)*, Singapore, 2021, pp. 171-184, [Online], https://link.springer.com/chapter/10.1007/978-981-16-4803-8_19.
- [3] H. Lim, Y. Kim, K. Jung, S. Hu, and H. Myung, "Avoiding Degeneracy for Monocular Visual SLAM with Point and Line Features," *IEEE International Conference on Robotics and Automation (ICRA)*, Xi'an, China, 2021, DOI: 10.1109/ICRA48506.2021.9560911.
- [4] H. Lim, J. Jeon, and H. Myung, "UV-SLAM: Unconstrained Line-Based SLAM Using Vanishing Points for Structural Mapping," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 1518-1525, Apr., 2022, DOI: 10.1109/LRA.2022.3140816.
- [5] A. Bartoli and P. Sturm, "Structure-from-motion using lines: Representation, triangulation, and bundle adjustment," *Computer Vision and Image Understanding*, vol. 100, no. 3, pp. 416-441, Dec., 2005, DOI: 10.1016/j.cviu.2005.06.001.
- [6] F. Zheng, G. Tsai, Z. Zhang, S. Liu, C.-C. Chu, and H. Hu, "Trifo-VIO: Robust and Efficient Stereo Visual Inertial Odometry Using Points and Lines," *IEEE International Workshop on Intelligent Robots and Systems (IROS)*, Madrid, Spain, 2018, DOI: 10.1109/IROS.2018.8594354.
- [7] Y. Yang, P. Geneva, K. Ekenhoff, and G. Huang, "Visual-Inertial Odometry with Point and Line Features," *IEEE International Workshop on Intelligent Robots and Systems (IROS)*, Macau, China, 2019, DOI: 10.1109/IROS40897.2019.8967905.
- [8] A. I. Mourikis and S. I. Roumeliotis, "A Multi-State Constraint Kalman Filter for Vision-aided Inertial Navigation," *IEEE International Conference on Robotics and Automation (ICRA)*, Rome, Italy, pp. 3565-3572, 2007, DOI: 10.1109/ROBOT.2007.364024.
- [9] A. Pumarola, A. Vakhtov, A. Agudo, A. Sanfeliu, and F. Moreno-Noguer, "PL-SLAM: Real-time Monocular Visual SLAM with Points and Lines," *IEEE International Conference on Robotics and Automation (ICRA)*, Singapore, pp. 4503-4508, 2017, DOI: 10.1109/ICRA.2017.7989522.
- [10] X. Zuo, X. Xie, Y. Liu, and G. Huang, "Robust visual SLAM with Point and Line Features," *IEEE International Workshop on Intelligent Robots and Systems (IROS)*, Vancouver, Canada, pp. 1775-1782, 2017, DOI: 10.1109/IROS.2017.8205991.
- [11] S. J. Lee and S. S. Hwang, "Elaborate Monocular Point and Line SLAM with Robust Initialization," *International Conference on Computer Vision (ICCV)*, Seoul, Korea, pp. 1121-1129, 2019, DOI: 10.1109/ICCV.2019.00121.
- [12] R. Mur-Artal, J. M. M. Montiel, and J. D. Tardos, "ORB-SLAM: A Versatile and Accurate Monocular SLAM System," *IEEE Transactions on Robotics*, vol. 31, no. 5, pp. 1147-1163, Aug., 2015, DOI: 10.1109/TRO.2015.2463671.
- [13] T. Qin, P. Li, and S. Shen, "VINS-mono: A Robust and Versatile Monocular Visual-Inertial State Estimator," *IEEE Transactions on Robotics*, vol. 34, no. 4, pp. 1004-1020, Aug., 2018, DOI: 10.1109/TRO.2018.2853729.
- [14] P. Kim, B. Coltin, and H. J. Kim, "Low-Drift Visual Odometry in Structured Environments by Decoupling Rotational and Translational Motion," *IEEE International Conference on Robotics and Automation (ICRA)*, Brisbane, Australia, 2018, DOI: 10.1109/ICRA.2018.8463207.
- [15] Y. Li, N. Brasch, Y. Wang, N. Navab, and F. Tombari, "Structure-SLAM: Low-Drift Monocular SLAM in Indoor Environments," *IEEE Robotics and Automation Letters*, vol. 5, no. 4, pp. 6583-6590, Oct., 2020, DOI: 10.1109/LRA.2020.3015456.
- [16] P. Kim, B. Coltin, and H. J. Kim, "Indoor RGB-D compass from a single line and plane," *Conference on Computer Vision and Pattern Recognition (CVPR)*, Salt Lake City, USA, 2018, DOI: 10.1109/CVPR.2018.00491.
- [17] R. Yunus, Y. Li, and F. Tombari, "ManhattanSLAM: Robust Planar Tracking and Mapping Leveraging Mixture of Manhattan Frames," *IEEE International Conference on Robotics and Automation (ICRA)*, Xi'an, China, pp. 6687-6693, 2021, DOI: 10.1109/ICRA48506.2021.9562030.
- [18] H. Zhou, D. Zou, L. Pei, R. Ying, P. Liu, and W. Yu, "StructSLAM: Visual SLAM with Building Structure Lines," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 4, pp. 1364-1375, Apr., 2015, DOI: 10.1109/TVT.2015.2388780.
- [19] D. Zou, Y. Wu, L. Pei, H. Ling, and W. Yu, "StructVIO: Visual-Inertial Odometry With Structural Regularity of Man-Made Environments," *IEEE Transactions on Robotics*, vol. 35, no. 4, pp. 999-1013, Aug., 2019, DOI: 10.1109/TRO.2019.2915140.
- [20] J. Lee and S.-Y. Park, "PLF-VINS: Real-time monocular visual-inertial SLAM with point-line fusion and parallel-line fusion,"

IEEE Robotics and Automation Letters, vol. 6, no. 4, pp. 7033-7040, Oct., 2021, DOI: 10.1109/LRA.2021.3095518.

- [21] J. Ma, X. Wang, Y. He, X. Mei, and J. Zhao, "Line-based stereo SLAM by Junction Matching and Vanishing Point Alignment," *IEEE Access*, vol. 7, pp. 181800-181811, Dec., 2019, DOI: 10.1109/ACCESS.2019.2960282.
- [22] J. Shi and Tomasi, "Good features to track," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Seattle, USA, pp. 593-600, 1994, DOI: 10.1109/CVPR.1994.323794.
- [23] B. D. Lucas and T. Kanade, "An Iterative Image Registration Technique with an Application to Stereo Vision," *International Joint Conference on Artificial Intelligence*, Vancouver, Canada, pp. 674-679, 1981, [Online], <https://onlinelibrary.wiley.com/doi/10.1002/rob.20360>.
- [24] C. Forster, L. Carlone, F. Dellaert, and D. Scaramuzza, "On-Manifold Preintegration for Real-Time Visual-Inertial Odometry," *IEEE Transactions on Robotics*, vol. 33, no. 1, pp. 1-21, Feb., 2017, DOI: 10.1109/TRO.2016.2597321.
- [25] G. Sibley, L. Matthies, and G. Sukhatme, "Sliding window filter with application to planetary landing," *Journal of Field Robotics*, vol. 27, no. 5, pp. 587-608, Aug., 2010, DOI: 10.1002/rob.20360.
- [26] R. G. Von Gioi, J. Jakubowicz, J.-M. Morel, and G. Randall, "LSD: A Fast Line Segment Detector with a False Detection Control," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 4, pp. 722-732, Apr., 2008, DOI: 10.1109/TPAMI.2008.300.
- [27] L. Zhang and R. Koch, "An efficient and robust line segment matching approach based on LBD descriptor and pairwise geometric consistency," *Journal of Visual Communication and Image Representation*, vol. 24, no. 7, pp. 794-805, Oct., 2013, DOI: 10.1016/j.jvcir.2013.05.006.
- [28] R. Toldo and A. Fusiello, "Robust Multiple Structures Estimation with J-Linkage," *European Conference on Computer Vision (ECCV)*, Berlin, Germany, pp. 537-547, 2008, DOI: 10.1007/978-3-540-88682-2_41.
- [29] S. Agarwal, K. Mierle, and Others, "Ceres solver," [Online], <http://ceres-solver.org>, Accessed: May 1, 2023.
- [30] M. Burri, J. Nikolic, P. Gohl, T. Schneider, J. Rehder, S. Omari, M. W. Achtelik, and R. Siegwart, "The EuRoC micro aerial vehicle datasets," *The International Journal of Robotics Research*, vol. 35, no. 10, pp. 1157-1163, Jan., 2016, DOI: 10.1177/0278364915620033.
- [31] Z. Zhang and D. Scaramuzza, "A Tutorial on Quantitative Trajectory Evaluation for Visual (-Inertial) Odometry," *IEEE International Workshop on Intelligent Robots and Systems (IROS)*, Madrid, Spain, pp. 7244-7251, 2018, DOI: 10.1109/IROS.2018.8593941.

임 현 준



2018 한국과학기술원 기계공학과(학사)
2020 한국과학기술원 전기및전자공학부 (석사)
2020~ 한국과학기술원 전기및전자공학부 (박사)

관심분야: Visual SLAM, Robotics

명 현



1992 한국과학기술원 전자전산학부(공학사)
1994 한국과학기술원 전기및전자공학과 (공학석사)
1998 한국과학기술원 전기및전자공학과 (공학박사)
1998~2002 한국전자통신연구원(ETRI) 선임 연구원

2002~2003 (주)이머시스 연구소장/기술이사
2003~2008 삼성전자종합기술원 전문연구원
2008~2018 한국과학기술원 건설및환경공학과 교수
2018~2021 한국과학기술원 로봇공학학제 책임교수
2019~현재 한국과학기술원 전기및전자공학부 교수

관심분야: Autonomous Robot/Drone/Car Navigation, SLAM, Deep Learning, Machine Learning