Journal of Korea Robotics Society (2023) 18(3):260-270
https://doi.org/10.7746/jkros.2023.18.3.260 ISSN: 1975-6291 / eISSN: 2287-3961

열화상 이미지 히스토그램의 가우시안 혼합 모델 근사를 통한 열화상-관성 센서 오도메트리

Infrared Visual Inertial Odometry via Gaussian Mixture Model Approximation of Thermal Image Histogram

신 재 호[1] · 전 명 환[2] · 김 아 영[†]
Jaeho Shin[1], Myung-Hwan Jeon[2], Ayoung Kim[†]

Abstract: We introduce a novel Visual Inertial Odometry (VIO) algorithm designed to improve the performance of thermal-inertial odometry. Thermal infrared image, though advantageous for feature extraction in low-light conditions, typically suffers from a high noise level and significant information loss during the 8-bit conversion. Our algorithm overcomes these limitations by approximating a 14-bit raw pixel histogram into a Gaussian mixture model. The conversion method effectively emphasizes image regions where texture for visual tracking is abundant while reduces unnecessary background information. We incorporate the robust learning-based feature extraction and matching methods, SuperPoint and SuperGlue, and zero velocity detection module to further reduce the uncertainty of visual odometry. Tested across various datasets, the proposed algorithm shows improved performance compared to other state-of-the-art VIO algorithms, paving the way for robust thermal-inertial odometry.

Keywords: Thermal-Inertial Odometry, SuperGlue, AGC, SLAM

1. 서 론

러시아-우크라이나 전쟁을 비롯하여 현대전의 가장 중요한 특징으로 전장의 무인화를 꼽을 수 있다. 이미 무인지상차량(Unmanned Ground Vehicle, UGV), 무인 항공기(Unmanned Aerial Vehicle, UAV)는 인간을 대신하여 여러 전장에 투입되어 물자 수송, 폭격 등의 중요한 임무를 수행 중이다. 무인 로봇이 주어진 임무를 수행하기 위해선 센서 데이터로부터 주변 환경을 인식하고, 자신의 정확한 위치를 파악하는 것이 중요하다.

Received : May. 25. 2023; Revised : Jun. 30. 2023; Accepted : Jul. 22. 2023
※ This study is a part of the research project, “Development of core machinery technologies for autonomous operation and manufacturing (NK230G)”, which has been supported by a grant from National Research Council of Science & Technology under the R&D Program of Ministry of Science, ICT and Future Planning
1. Master Student, Mechanical Engineering, Seoul National University, Seoul, Korea (leah100@snu.ac.kr)
2. Post Doctoral Research Fellow, Institute of Advanced Machines and Design, Seoul National University, Seoul, Korea (myunghwan.jeon@snu.ac.kr)
† Associate Professor, Corresponding author: Mechanical Engineering, Seoul National University, Seoul, Korea (ayoungk@snu.ac.kr)

로봇이 미지의 환경에서 자신의 위치를 추정하고 지도를 작성하는 기술을 동시적 위치추정 및 지도작성(SLAM, Simultaneous Localization And Mapping)라 하며, 카메라로부터 촬영된 이미지를 데이터로 활용하는 기술을 visual SLAM라 한다. 일반적인 RGB 카메라는 오직 가시광선 영역의 빛을 포착하기에 밤, 숲과 같이 시야가 제한적인 환경에서 무인기가 SLAM을 통해 위치를 추정하는 데에 많은 어려움이 따른다. 이에 비해 열화상 카메라는 적외선 영역의 빛을 포착하여 인간의 눈으로 쉽게 감지할 수 없는 어두운 물체를 인식하는 것에 적합하며, 이러한 장점은 언급한 환경에서 열화상 카메라로 주변 환경을 인지하는 로봇의 임무 수행을 용이하게 만든다.

그러나 열화상 이미지는 해상도가 낮고, 밝은 부분과 어두운 부분의 대비가 크지 않으며 Signal to Noise Ratio (SNR)이 낮다는 단점이 있다. 이는 센서가 촬영할 수 있는 온도의 범위가 일상 생활에서 관측되는 물체들의 온도 범위에 비해 훨씬 넓고, 오직 적외선 파장 대의 빛을 관측함으로써 가시광선 영역의 텍스처 정보가 소실되기 때문이다. 또한 14-bit의 열화상 이미지를 여러 컴퓨터 비전 알고리즘에 적용하여 분석하기 위해서는 이를 8-bit 데이터로 변환하는 과정이 필요하다.

대표적인 방법은 현재 프레임의 최대, 최소 픽셀 값을 사용하여 8-bit 선형 표준화를 적용하는 Automatic Gain Control (AGC)이다. 그러나 이러한 방법은 온도가 매우 높거나 낮은 물체가 관측될 때, 시퀀스 간 급격한 밝기의 변화가 발생한다는 문제점이 있다. 이를 극복하기 위해 긴 시간대를 촬영한 이미지 시퀀스의 히스토그램 분포를 토대로 표준화의 범위를 설정하는 방법이 있다. 그러나 이는 사전에 시퀀스에 대한 정보가 필요할 뿐 아니라, 경험에 의한 파라미터의 선정과정을 필요로 한다. 또한 열화상 카메라는 카메라 자체의 온도로 인해 축적된 노이즈를 제거하는 비균일 보정(NUC, Nonuniformity Correction)을 반드시 거친다. 이 과정 중에는 열화상 이미지 촬영이 중단되며, 보정 전후로 픽셀 값의 변화가 발생하여 연속된 프레임 간에 데이터의 일관성을 잃게 된다. 이러한 문제점들은 SLAM 알고리즘에 필수적인 correspondence matching, place recognition과 같은 여러 모듈의 성능을 저해하는 요인이 된다. 본 논문에서는 이를 극복한 새로운 이미지 표준화 방법과 딥러닝 기반 특징점 추출 및 매칭 알고리즘인 SuperPoint[1], SuperGlue[2]를 활용한 Visual-Inertial Odometry (VIO) 알고리즘을 제안한다. 제안하는 표준화 방법은 시퀀스의 히스토그램을 기반으로 텍스처 정보가 풍부한 영역을 강조하고, 노이즈가 큰 배경 영역은 최소화하기에 visual tracking에 최적화된 8-bit 이미지를 생성할 수 있다. 또한 움직임의 uncertainty를 감소하는 방법으로써 IMU와 열화상 이미지의 disparity 측정값을 결합한 정확한 zero velocity detection 알고리즘을 이용한다.

본 논문에서 제안하는 알고리즘의 논점은 다음과 같다.

- 14-bit 열화상 데이터 시퀀스의 히스토그램에 multiple Gaussian fitting을 적용하여 표준화된 8-bit 이미지를 얻는 과정을 새롭게 제안하였고, 이로부터 경험적인 파라미터의 선정 없이 텍스처 정보 손실을 최소화한 8-bit 열화상 이미지를 얻을 수 있다.
- 모든 VIO 모듈의 특징점과 descriptor 추출, 매칭 방법으로 SuperPoint 및 SuperGlue를 사용한다. 이로부터 기존의 특징점 알고리즘 (ORB[3], SIFT[4], SURF[5])에 비해 노이즈가 풍부한 열화상 이미지에서도 정확하고 강건한 특징점 매칭 성능을 확보하였다.
- IMU 측정값을 고려한 픽셀 disparity의 평균을 계산하여 정확하게 정지 상태를 감지하였고, 이를 통해 동적 상황에서도 멈춰 있는 시스템의 state uncertainty를 zero velocity update로 감소시켰다.
- STheReo Valley Morning 시퀀스에서 4.905%p 개선된 RMSE를 포함하여 4개의 데이터셋에서 state-of-the-art VIO 알고리즘인 VINS-Mono에 비해 개선된 결과를 얻었다.

2. 선행 연구 조사

2.1 Thermal Visual Odometry

최근 주변 환경의 조명 변화에 강건한 열화상 이미지의 장점을 활용하여 visual odometry를 수행하는 여러 알고리즘들이 제안되었다. Borges et al.[6]은 NUC에 의한 tracking 성능 저하를 보완하는 NUC trigger manager를 적용하여 열화상 이미지에 적용 가능한 semi-dense optical flow를 수행하였다. Vidas et al.[7]는 프레임의 최솟값과 최댓값의 평균을 중심으로 픽셀 값의 차이가 256 이내에 위치한 픽셀들을 표준화한 후, GFTT detector와 optical flow를 통해 움직임을 추정하였다. Nilsson et al.[8]는 Harris corner detector를 통하여 특징점을 검출한 뒤, 주위 패치의 normalized cross correlation을 최소화하는 6 DOF 움직임을 계산하였다. 그러나 이러한 방법들은 표준화된 이미지의 픽셀 값을 직접적으로 비교하고 brightness consistency 가정을 사용하기 때문에 이미지의 표준화 방법에 크게 의존하며 노이즈에 강건하지 않다. 이러한 문제점들을 극복하고자 [9,10]은 표준화를 적용하지 않은 14-bit 이미지의 radiometric error를 최소화하였고, [9,11]은 IMU의 측정값을 prior로 설정하여 특징점 tracking을 수행하였다. 또한 동일한 장면에 대한 가시광선 영역과 적외선 영역의 정보를 모두 활용하고자 DWT (Discrete Wavelet Transform)로 두 이미지를 합성하거나 [12] 각각의 이미지에서 추출한 특징점을 함께 사용하는 시도가 있었다[13-15]. 이 방법들은 특징점 tracking에 IMU, RGB 카메라와 같은 추가적인 센서를 이용한 것으로, 열화상 이미지만을 사용하여 매칭을 수행하는 본 연구의 방법과 달리 센서 데이터를 결합하는 추가적인 과정이 필요하다. Mouats et al.[16]는 열화상 이미지에서 여러 특징점 추출 알고리즘의 성능을 매칭 점수를 통해 비교하고 연속된 시퀀스에서 왼쪽과 오른쪽 이미지의 reprojection error의 합을 동시에 최소화하는 pose를 추정하였으며, 평균적으로 4% 이하의 travelled error를 얻었다. 그러나 위 연구를 수행한 직선 경로가 많고 정형화된 환경과 다르게 정형화되지 않은 환경에서 [3-5]와 같이 RGB 이미지에 적용되는 알고리즘은 안정적인 성능을 보여주지 못한다.

2.2 Learning-based Visual Odometry

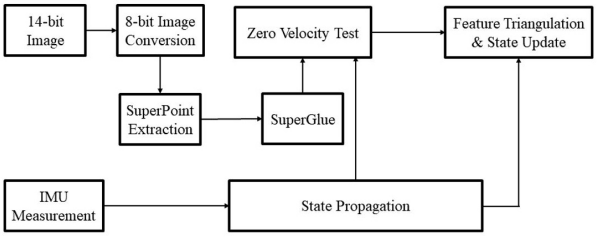
Deep learning을 visual odometry에 적용하여 기존의 방법이 지니는 여러 한계점을 극복하고자 하는 연구가 활발히 진행되고 있다. 그 중 한 예로, end-to-end learning의 방법으로 카메라의 움직임을 추정하는 여러 네트워크들이 제안되었다[17-20]. 그러나 이는 실제 데이터셋에 적용하였을 때 경쟁력 있는 성능을 보여주지 못했으며, 지도 작성이 불가능하고 학습된 환경

에 대해서만 카메라의 위치 추정이 가능하다는 단점이 있다. 특히 이미지의 텍스쳐 정보와 특징점이 적은 열화상 카메라의 경우, 움직임 추정에 이러한 방법들을 적용하기 불가능하다. 한편, 신경망을 기반으로 특징점을 추출하여 이미지의 품질에 대한 강건함을 확보하고자 하는 시도들이 있었다[1,21-23]. Li et al.[24], Tang et al.[25]은 이러한 장점을 활용하여 ORB-SLAM2[26]의 특징점 추출 방법을 학습 기반의 네트워크로 대체하여 위치 추정 성능을 개선하였다. Liang et al.[27]는 이미지의 visual saliency를 계산하여 semantic 정보가 높은 영역의 점을 추출하도록 Direct Sparse Odometry (DSO)[28]를 개선하여 이미지 그래디언트가 낮은 환경에서도 안정적으로 특징점을 추출하여 위치 추정을 수행하였다.

3. 연구 방법

3.1 System Overview

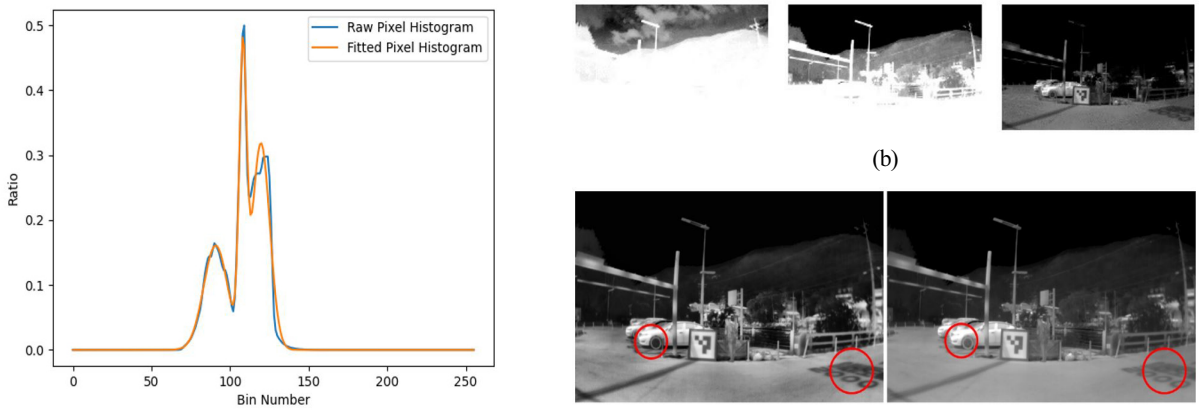
전체적인 시스템에 대한 diagram은 [Fig. 1]와 같다. 100 Hz로 취득 가능한 IMU 데이터를 전달받아 IMU motion model에 의하여 state와 covariance를 propagate 하는 데에 사용한다. 14-bit 열화상 이미지는 본 논문에서 제시한 전처리 방법을 통해 8-bit 이미지로 변환 후, SuperPoint 특징점을 추출, 이전 프레임과 SuperGlue 알고리즘으로 이를 매칭한다. 매칭된 픽셀쌍과 propagated IMU state로부터 시스템의 정지 여부를 검증한 뒤 특정 횟수 이상 tracking된 특징점은 triangulation을 통해 3D 좌표를 구하고, zero velocity test 결과에 맞게 measurement model을 설정하여 시스템의 state update를 진행한다.



[Fig. 1] Block diagram illustrating pipeline of proposed algorithm. SuperGlue outdoor weight is used to match feature points of 8-bit converted thermal images

3.2 AGC by Multiple Gaussian Fitting

전체 시퀀스에 대한 히스토그램을 나타낸 결과는 [Fig. 2]의 (a)와 같으며, 이로부터 피크점이 위치하여 근방에 많은 픽셀이 분포하는 총 3개의 지점을 확인하였다. 각 지점 근방의 픽셀들이 이미지에서 나타내는 영역을 직접 확인하기 위해 먼저 전체 히스토그램을 각 피크점을 중심으로 하는 총 3개의 Gaussian의 합으로 근사하였다. 이때, i번째 Gaussian의 ($i=1,2,3$) 평균을 m_i, 표준편차를 σ_i라 하고, $[m_i-k_i\sigma_i, m_i+k_i\sigma_i]$에 위치한 픽셀들을 표준화한 결과, [Fig. 2]의 (b)와 같은 세 장의 8-bit 이미지를 얻었다(k_i는 상수). 첫번째 정점을 기준으로 이미지를 표준화한 결과, 해당 영역은 주로 하늘과 구름에 해당하는 픽셀, 두번째 정점은 산과 같은 뒷부분의 배경을 포함한 픽셀, 마지막 정점은 visual odometry에 필요한 텍스처 정보가 풍부한 물체들을 포함한 픽셀이 주로 분포함을 확인하였다. 위의 결과로부터 한 장의 열화상 이미지를 세 개의 영역으로 분류할 수 있음을 확인하였으며, 각 이미지가 포함하고 있는 semantic 정보에 따라 각기 다른 가중치 α_i를 두어 아래와 같이 최종 이미지 I_{AGC}를 얻었다. 세 장의 이미지를 합한 뒤엔 CLAHE를 적용하여 이미지의 전체적인 대비를 높였다.



(a)

(b)

(c)

[Fig. 2] Proposed multiple Gaussian fitting-based thermal infrared image conversion. (a) Histogram of 14-bit raw image pixels with three peaks and fitted result by Gaussians. (b) Dominant region of each peak from the image. (c) Converted 8-bit image by proposed method (left), and clipping method (right)

$$I_{AGC} = clahe(\Sigma_{i=1}^{3}\alpha_i I_i) \tag{1}$$

첫번째 이미지에 주로 나타나는 구름에서 추출된 특징점은 먼 거리에 위치하여 시스템이 크게 움직일 때에도 disparity가 0에 가까워 위치 추정에 부정적인 영향을 미친다. 따라서 첫번째 이미지의 가중치를 0으로, 세번째 이미지의 가중치를 가장 높게 설정하여 visual odometry에 적합한 8-bit 이미지를 얻을 수 있다. 제안한 방법을 적용하였을 때, 단일 구간을 임의로 설정하여 표준화하는 기존의 clipping 방법에 비해 배경의 잡음이 제거되고 차량과 같은 물체의 대비는 강조된 것을 확인할 수 있다.

3.3 Multi-State Constraint Kalman Filter

본 논문에서는 IMU 측정값과 열화상 이미지에서 추출한 특징점 매칭 결과를 필터링 기반의 방법인 Multi-Sate Constraint Kalman Filter (MSCKF)[29]를 통해 결합한다. MSCKF는 기존 Extended Kalman Filter (EKF) 기반의 VIO와 다르게 Gaussian으로 표현되는 특징점들의 3D 좌표 대신 이전 시간 대의 IMU pose를 state에 추가하여 특징점의 개수에 선형적인 계산 복잡도를 가지며, 큰 스케일의 환경에서도 정밀한 pose 추정이 가능하다. 본 논문에서 사용할 시간 t_k에서의 IMU state 표현은 아래와 같다:

$$X_{I_k} = \left[q_{I_kG}^T,\ p_{I_KG}^T,\ v_{I_kG}^T,\ b_r^T,\ b_a^T\right] \tag{2}$$

각 문자는 순서대로 global 좌표계 $\{G\}$에 대한 IMU rotation의 unit quaternion, 위치, 속도 그리고 각속도와 선 가속도의 bias에 해당한다. 이로부터 시스템의 전체 state X_k는 다음과 같다.

$$X_k = \left[X_{I_k}^T,\ X_C\right] \tag{3}$$

$$X_C = \left[X_{T_{k-1}}^T,\ \cdots,\ X_{T_{k-n+1}}^T\right] \tag{4}$$

X_C는 위치와 unit quaternion으로 표현되는 sliding window 내 과거 $n-1$개의 camera state이다.

3.3.1 IMU Propagation

두 시간 t_k, t_{k+1} 사이 Δt 동안 IMU kinematics에 의해 얻어지는 t_{k+1}에서의 state estimation은 다음과 같다. 이 때, $\widehat{(\cdot)}_{k+1|k}$는 IMU kinematic 모델에 의해 추정된 t_{k+1}에서의 state, $\widehat{(\cdot)}_{k|k}$는 measurement model에 의해 update된 t_k에서의 state이다. 또한 ω_m은 IMU 각속도 측정값, a_m은 IMU 가속도 측정값, $\hat{R}_{I_\tau}$는 IMU orientation의 rotation matrix 표현이다.

$$\begin{aligned}
\hat{q}_{k+1|k} &= \exp\left(\frac{1}{2}\begin{bmatrix} -[\omega_m]_\times & \omega_m \\ -\omega_m^T & 0 \end{bmatrix}\Delta t\right)\hat{q}_{k|k} \\
\hat{v}_{k+1|k} &= \hat{v}_{k|k} - g\Delta t + \int_{t_k}^{t_{k+1}} \hat{R}_{WI_\tau}\left(a_m(\tau) - \hat{b}_a(\tau)\right)d\tau \\
\hat{p}_{k+1|k} &= \hat{p}_{k|k} + \int_{t_k}^{t_{k+1}} \hat{v}\,(\tau)d\tau \\
\hat{b}_{r,k+1|k} &= \hat{b}_{r,k|k} \\
\hat{b}_{a,k+1|k} &= \hat{b}_{a,k|k}
\end{aligned} \tag{5}$$

이때, $[\cdot]_\times$는 3차원 벡터로부터 반대칭 행렬을 정의하는 연산자로, 아래와 같이 정의된다.

$$\omega = \left[\omega_x, \omega_y, \omega_z\right]^T,\ [\omega]_\times = \begin{bmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{bmatrix} \tag{6}$$

본 연구에서는 4th order Runge-Kutta method를 통해 위 적분을 계산하며, 각 state의 perturbation 모델을 위 식에 대입하면 state Jacobian $\Phi(t_{k+1}, t_k)$, 노이즈 n의 Jacobian G_k에 대하여 다음과 같은 꼴의 error state propagation 식을 얻는다[30].

$$\tilde{X}_{I_{k+1}} = \Phi(t_{k+1}, t_k)\tilde{X}_{I_k} + G_k n \tag{7}$$

위 관계식과 사전에 알고 있는 state covariance $P_{k|k}$, 노이즈 covariance $Q_{k|k}$로부터 아래의 covariance propagation을 얻는다.

$$P_{k+1|k} = \Phi(t_{k+1}, t_k)P_{k|k}\Phi(t_{k+1}, t_k)^T + G_k Q_{k|k} G_k \tag{8}$$

3.3.2 Nullspace Projected Measurement Model

MSCKF는 하나의 특징점에 대하여 여러 프레임에서의 픽셀 측정값을 쌓은 뒤, triangulation을 통해 기준 좌표계에 대한 3차원 좌표를 얻기에 tracking 단계에 놓여있는 특징점은 위치와 covariance를 알 수 없다. 또한, 특징점의 3D 좌표를 state variable로써 포함하지 않기 때문에, 이에 무관한 measurement model로 state를 update해야 한다. 일반적인 EKF VIO에서 IMU의 pose X, 특징점의 위치 p, 측정 노이즈 n_i, projection model $h(X,p)$로부터 계산되는 i번째 측정값 z_i의 값은 다음과 같다.

$$z_i = h(X,p) + n_i \tag{9}$$

실제로 얻어진 측정값 z와 위 식을 통해 계산한 추정값 $\hat{z}$의 차이로부터 계산된 측정 residual $\tilde{z}$를 pose와 특징점의 추정값에 대하여 선형화하면 최종적으로 pose residual $\tilde{X}$, 특징점 residual $\tilde{p}$에 대한 아래의 식을 얻는다.

$$\tilde{z}_i = H_x\tilde{X} + H_p\tilde{p}_i + n_i \tag{10}$$

이때, H_x, H_p는 각각 pose X와 특징점 p에 대한 h의 Jacobian이다. 한편, 특징점 p에 대한 prior covariance와 추정값을 알 수 없으므로 H_p를 QR 분해 후 얻어진 Q_2의 전치행렬을 양 변에 곱하여 $\tilde{p}$에 무관한 아래의 새로운 measurement residual model을 얻고 이로부터 state를 update할 수 있다.

$$H_p = [Q_1 \;\; Q_2]\begin{bmatrix} R_1 \\ 0 \end{bmatrix} \tag{11}$$

$$\tilde{z}_{o,i} = H_{o,x}\tilde{X} + n_{o,i} \tag{12}$$

$$\tilde{z}_{o,i} = Q_2^T\tilde{z}_i,\; H_{o,x} = Q_2^T H_x,\; n_{o,i} = Q_2^T n_i \tag{13}$$

3.4 Zero Velocity Detector

만일 여러 센서 정보로부터 시스템의 정지 상태를 확실하게 감지할 수 있다면 선 가속도와 각속도의 참값을 0으로 설정 후, IMU measurement model을 filter의 measurement model로 설정하여 state uncertainty를 줄일 수 있을 것이다. 이를 zero velocity update라 하고, 시스템의 정지 여부를 판단하는 방법은 크게 IMU의 선 가속도, 각속도 측정값을 통해 판단하는 방법과 특징점 사이의 disparity를 통해 판단하는 방법이 있다. OpenVINS[31]는 아래의 두 방법으로 이를 구현하였다.

먼저 IMU 측정값의 참값을 0으로 두었을 때 residual $\tilde{z}$는 다음과 같다.

$$\tilde{z} = \begin{bmatrix} -(a_m - b_a - R_{IW}g - n_a) \\ -(\omega_m - b_r - n_r) \end{bmatrix} \tag{14}$$

이때, n_a, n_r은 각각 IMU 선 가속도 측정값, 각속도 측정값에 대한 노이즈이다. 이에 대한 IMU state (rotation, bias)의 Jacobian H와 이를 선형화한 결과는 아래와 같다.

$$H = \begin{bmatrix} \frac{\partial\tilde{z}}{\partial R_{IW}} & \frac{\partial\tilde{z}}{\partial b_a} & \frac{\partial\tilde{z}}{\partial b_g} \end{bmatrix} = [-[R_{IW}g]_\times \;\; -I_{3\times3} \;\; -I_{3\times3}] \tag{15}$$

$$\tilde{z} = H\,[R_{IW} \;\; b_a \;\; b_g]^T + [n_a \;\; n_r]^T \tag{16}$$

IMU state의 covariance을 P, 측정값의 노이즈 covariance를 R라 하면, 적절한 파라미터 ϵ과 χ^2을 설정하여 아래와 같은 chi-squared test를 통해 zero velocity 가정의 타당성을 검정할 수 있다.

$$\tilde{z}^T(HPH^T + \epsilon R)^{-1}\tilde{z} < \chi^2 \tag{17}$$

그러나 이 방법은 시스템이 등속도로 움직이는 경우에도 선 가속도와 각속도의 측정값이 0에 가까워 정지 상태와 구분할 수 없다는 문제점이 있다.

다음으로 아래와 같이 tracking된 픽셀 쌍의 disparity norm의 평균을 계산하여 threshold α와 비교 후, 시스템의 정지 여부를 판단하는 방법이 있다.

$$\frac{1}{N}\Sigma_{i=0}^{N} \parallel u_1 - u_0 \parallel_2 < \alpha \tag{18}$$

이때, N은 두 프레임 사이 tracking된 특징점의 개수, u_0, u_1은 각 프레임에서 특징점의 픽셀 좌표이다. 이 방법 역시 주위에 움직이는 물체가 많은 경우, 정지 상태와 움직이는 상태를 구분하기 힘들다는 문제점이 있다. 두 방법을 병행하여 사용할 때에도 카메라 시야에 잡히는 모든 물체가 시스템과 동일한 속도로 움직이는 경우에는 시스템의 운동 상태를 결정할 수 없게 된다. 따라서 본 연구에서는 IMU의 측정값으로부터 계산된 시스템의 state 변화를 고려하였을 때, 신뢰성 있는 disparity를 보이는 픽셀 쌍에 대해서만 disparity의 평균을 계산하는 새로운 방법을 제안한다. Disparity를 통해 시스템의 정지 여부를 판단할 때에 움직이거나 너무 먼 거리에 위치한 물체에 놓인 픽셀은 평균을 계산하는 과정에 포함하지 않아야 한다.

먼저, 시간 t_0, t_1에서 정지한 물체 위에 놓인 동일한 특징점을 관측하였다는 가정하에 아래와 같은 projection geometry가 성립한다.

$$z_{c_i}u_i = KP_{c_i} = K(R_{c_iW}P_w + t_{c_iw}) \quad (i = 0,1) \tag{19}$$

이때, z_{c_i}는 시간 t_i의 카메라 좌표계에서 측정한 특징점의 depth, K는 카메라의 intrinsic matrix, P_w는 기준 좌표계에 대한 특징점의 3D 좌표, R_{c_iw}, t_{c_iw}는 기준 좌표계에 대한 카메라의 pose이다. $i = 0$인 경우의 식 (19)을 P_w에 대하여 정리하면 아래와 같다.

$$P_w = z_{c_0}R_{c_0w}^{-1}K^{-1}u_0 - R_{c_0w}^{-1}t_{c_0w} \tag{20}$$

식 (20)을 시간 t_1에서의 식 (19)에 대입하면 아래와 같은 관계식을 얻는다.

$$u_1 = \frac{z_{c_0}}{z_{c_1}}KR_{c_1c_0}K^{-1}u_0 + \frac{1}{z_{c_1}}K(t_{c_1w} - R_{c_1c_0}t_{c_0w}) \tag{21}$$

두 프레임 간 시간 차이 $t_1 - t_0 \ll 1$라는 가정하에 각각의 좌표계에서 측정한 depth 값의 차이는 매우 적으므로

$\frac{z_{c_0}}{z_{c_1}} \approx 1$라 가정할 수 있으며, 최종적으로 아래의 식이 성립한다.

$$u_1 - KR_{c_1c_0}K^{-1}u_0 = \frac{1}{z_{c_1}}K(t_{c_1w} - R_{c_1c_0}t_{c_0w}) \quad (22)$$

IMU propagation을 통해 얻은 두 좌표계 간의 rotation $R_{c_1c_0}$와 두 프레임의 기준 좌표계에 대한 위치 t_{c_1w}, t_{c_0w}, 특징점 매칭을 통해 얻은 두 픽셀 좌표 u_1, u_0를 통해 오른쪽 항의 z_{c_1}를 제외한 값을 모두 계산할 수 있다. 따라서 좌변과 우변의 벡터는 평행하므로 cosine similarity를 계산 후, 특정 threshold θ_{thresh}와 비교를 통해 정지 물체 관측에 대한 가정의 타당성을 확인할 수 있다.

$$Sim(u_1 - KR_{c_1c_0}K^{-1}u_0, K(t_{c_1w} - R_{c_1c_0}t_{c_0w})) < \theta_{thresh} \quad (23)$$

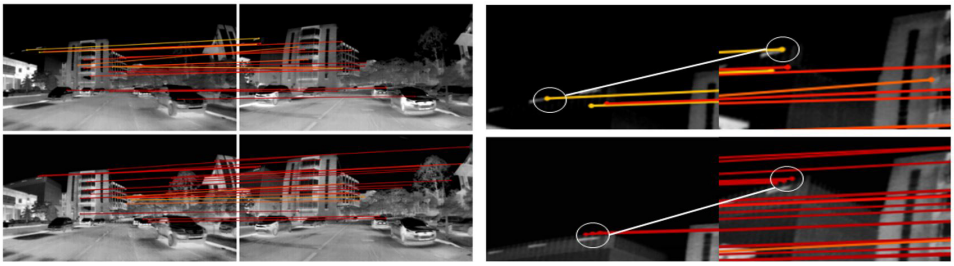
만일 위와 같은 가정을 통해 정지 상태를 확인하게 되면, 식 (14)을 measurement model로 하여 state를 update한다.

4. 실험 및 결과

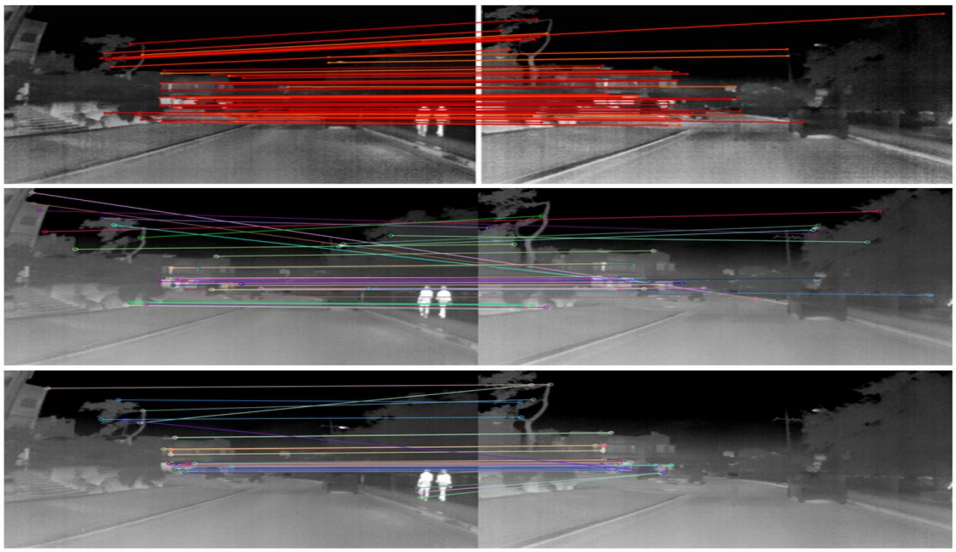
4.1 데이터셋 구성 및 평가 지표

평가에 사용한 데이터셋은 STheReo Valley Morning (1011.287 m), Evening (1006.383 m) Sequence, 그리고 같은 센서 구성으로 학교 캠퍼스 환경 내 동일한 장소의 낮(697.46 m)과 밤(702.648 m)을 촬영한 SNU Sequence로 구성되어 있다. 두 데이터셋은 차량에 설치된 두 스테레오 열화상 카메라(FLIR, fA-65, 10 Hz)와 IMU (Xsens, MTi-300, 100 Hz)로부터 얻어졌으며, 본 연구에서는 단일 카메라의 이미지만을 사용했다. Valley Sequence는 긴 직진 구간과 급격한 회전으로 이루어져 있으며, SNU Sequence는 짧은 거리에 비해 많은 회전 구간으로 이루어져 있다. 특히 SNU Evening Sequence는 초반 좌회전 교차로에서 정지 후, 많은 동적 물체를 촬영하기 때문에 이를 통해 제안된 zero velocity detection 알고리즘의 유효성을 검증하기 적합하다.

제안된 8-bit 열화상 이미지 변환 알고리즘의 유효성은



(a)



(b)

[Fig. 3] Comparison of feature matching performance between various algorithms. (a) Clipping method (top) and proposed method (bottom) with SuperGlue. (b) Proposed method with SuperGlue (top), SIFT (medium), and ORB (bottom)

clipping 방법을 적용했을 때와의 매칭 inlier 수 평균과 Absolute Trajectory Error (ATE)의 비교를 통해 이루어졌다. 또한 추가적으로 두 개의 이미지에 각각의 방법을 적용하였을 때 SuperGlue 매칭 성능을 정성적으로 분석하였다. 제안한 알고리즘의 위치 추정 성능은 특징점 매칭 방법으로 LK optical flow와 기존 disparity 기반의 zero velocity detection을 이용한 OpenVINS[31], 단안 카메라와 IMU를 사용하는 state-of-the-art VIO 알고리즘인 VINS-Mono[32]와의 translational, rotational error의 비교를 통해 검증하였으며, 두 알고리즘에서 14-bit 이미지는 clipping 방법으로 변환하였다.

4.2 Image quality evaluation

두 방법의 매칭 결과를 이미지를 통해 직접 비교한 결과, 제안된 변환 방법을 적용하였을 때, 잘못된 매칭이 감소할 뿐만 아니라 이미지의 전영역에서 더 많은 특징점을 추출하여 올바르게 매칭함을 [Fig. 3]의 (a)를 통해 확인하였다. 추가적으로, Visual SLAM에서 많이 사용되는 특징점 추출 알고리즘인 SIFT, ORB과의 성능 비교를 통해 제안한 방법이 기존의 방법에 비해서도 우수한 결과를 나타냄을 보았다. Tracking 단계에서 매칭된 특징점들의 RANSAC inlier 개수를 평균한 결과와 ATE는 [Table 1], [Table 2]와 같다.

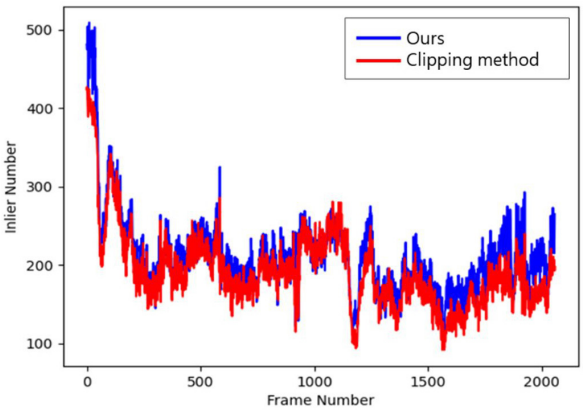
비교 결과, 제안된 알고리즘이 기존의 clipping 기법에 비해 모든 시퀀스에서 평균적으로 16.83% 높은 매칭 inlier를 보였다. 즉, 히스토그램 fitting을 통한 8-bit 변환 방법이 단일 구간으로 표준화하는 기존의 방법에 비해 노이즈와 정보 손실에 강건한 열화상 이미지 변환 방법임을 확인하였다. 이는 특정 시점이 아닌 시퀀스의 전 시간 대에 걸쳐 이루어졌음을 [Fig. 4]를 통해 확인할 수 있다. 이러한 이미지 품질 개선 효과는 [Table 2]의 RMSE 비교 결과에서 볼 수 있듯이 VIO 알고리즘을 통한 위치 추정의 정확도 향상으로 이어졌다.

[Table 1] Matching inlier numbers of two preprocessing methods

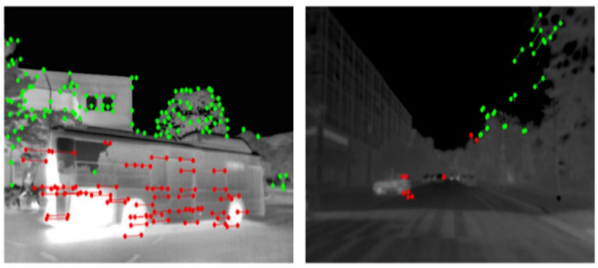
Sequence	Proposed	Clipping
Valley morning	211.26	192.46
Valley evening	92.74	79.74
SNU afternoon	285.72	264.76
SNU evening	290.76	218.08

[Table 2] RMSE comparison of two preprocessing methods

Sequence	Proposed	Clipping
Valley morning	6.04	7.33
Valley evening	11.68	23.85
SNU afternoon	5.47	9.41
SNU evening	6.77	12.76



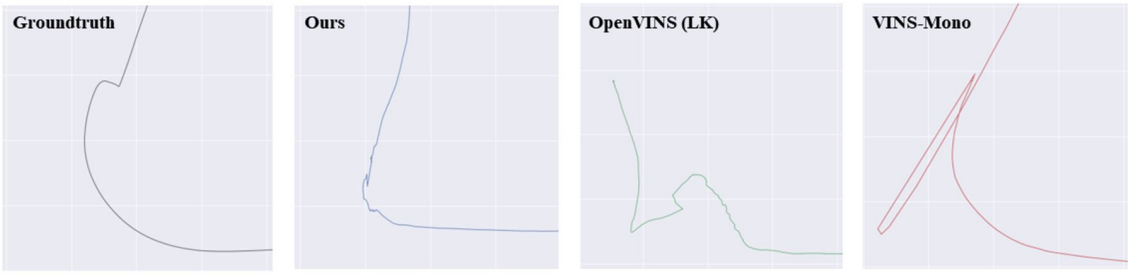
[Fig. 4] Comparison of matching inliers between proposed preprocessing method and existing clipping method in Valley Morning sequence



[Fig. 5] Visualized results of the proposed pixel rejection algorithm. Pixels on dynamic or far objects are colored red

4.3 Zero velocity detection 평가

주어진 데이터셋 내에서 차량의 정지 시점에 대해 판단할 수 있는 정량적인 기준을 얻기 어려운 관계로 본 논문에서는 정성적인 방법으로 제시된 zero velocity detection 알고리즘의 성능을 평가하였다. [Fig. 5]는 IMU 측정값으로 얻은 움직임에 부합하는 픽셀들의 tracking 결과를 나타낸 것이다. 왼쪽 이미지는 교차로에서 시스템이 정지하였을 때, 빠른 속도로 움직이는 버스 위의 픽셀을, 오른쪽 이미지는 움직이는 차량과 원거리 물체 위 픽셀을 구분하여 나타낸 것이다. 이로부터 제안된 알고리즘이 동적물체 위의 픽셀과 정지하지 않은 상황임에도 disparity의 평균을 낮추는 원거리의 픽셀을 효과적으로 제거하였음을 확인할 수 있다. 특히 왼쪽의 이미지는 SNU Evening 시퀀스의 교차로에서 차량이 잠시 정지한 상황으로, 기존의 zero velocity detection 알고리즘에서는 움직이는 버스 위의 픽셀에 의해 [Fig. 6]와 같이 grountruth에서 크게 어긋난



[Fig. 6] Effect of dynamic objects for pose estimation of each algorithm. Proposed zero velocity detection method effectively relieves unstable estimation by forcing zero-value measurement

[Table 3] Translational RMSE in datasets (meters)

Sequence	Ours	OpenVINS (LK)	VINS-Mono
Valley morning	**6.04**	10.6	55.64
Valley evening	**11.68**	20.84	93.95
SNU afternoon	**5.47**	8.58	34.97
SNU evening	**6.77**	26.76	28.67

[Table 4] Rotational RMSE in datasets (degrees per a meter)

Sequence	Ours	OpenVINS (LK)	VINS-Mono
Valley morning	0.26	0.265	3.26
Valley evening	0.35	0.37	6.79
SNU afternoon	3.28	3.18	4.56
SNU evening	3.20	3.19	4.6

trajectory가 계산된 반면, 본 논문의 detection 방법은 불과 cm 단위의 떨림만이 존재하였다.

4.4 경로 오차 분석

4.4.1 STheReo Valley Sequence

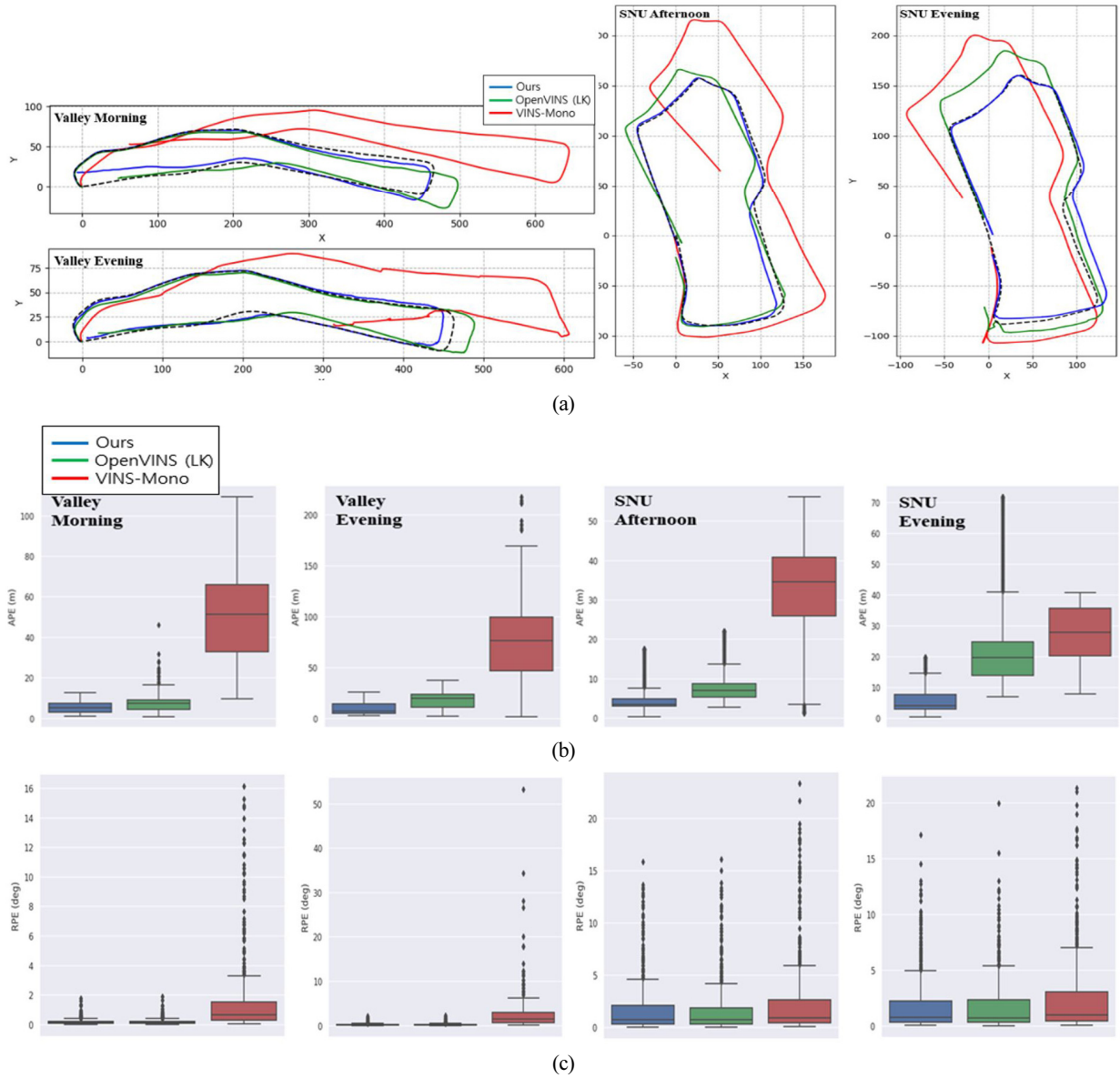
Valley Sequence는 출발 이후 차량 기준 두 번의 급격한 우회전 후 약 500 m의 직진, 유턴을 거쳐 다시 원점으로 돌아오는 trajectory를 보인다. 따라서 알고리즘의 직진 구간에서의 스케일 추정, 큰 각도로 회전 시 orientation 추정에 대한 성능을 검증하기 용이하다. 실험 결과, OpenVINS는 직진 구간에서의 스케일 추정에 실패하여 실제보다 x 방향으로 긴 trajectory가 계산된 것을 [Fig. 7]의 (a)에서 볼 수 있다. 또한 VINS-Mono는 추가적으로 회전을 포함한 모든 구간에서 orientation을 정확히 계산하는 데 실패하여 trajectory가 전반적으로 어긋난 모습을 보인다. 반면, 제안한 알고리즘은 낮 시퀀스와 밤 시퀀스 모두에서 개선된 translational error와 rotational error를 보임을 [Table 3]과 [Table 4]에서 확인할 수 있다. 특히, 낮에 비해 이미지의 품질이 낮고 특징점의 수가 적은 밤 시퀀스에서도 제안한 알고리즘은 강건한 성능을 나타낸다.

4.4.2 SNU Sequence

SNU Sequence는 Valley Sequence에 비해 전체 길이는 짧지만 교내 환경에서 주행하여 많은 회전과 동적물체를 포함하고 있는 것이 특징이다. 차량의 주행 속도가 동적물체의 움직임에 비해 빠르지 않기에 해당 시퀀스를 통해 동적 환경에서 위치 추정 알고리즘의 강건함을 확인할 수 있다. 실험 결과, 비교군의 두 알고리즘은 스케일과 orientation이 정확하지 않아 groundtruth와 불일치한 trajectory를 보였다. 이는 Valley Sequence에서와 마찬가지로 제안한 알고리즘에 비해 전처리된 이미지의 품질이 낮고 특징점 매칭이 부정확하여 회전 시 orientation과 직진 시 이동 거리를 추정하는 것에 실패했기 때문이다. 특히 Valley Evening 시퀀스의 교차로에서 정지 후 동적 물체 위 픽셀 비율이 큰 상황에서 제안된 zero velocity update 모듈이 유효함을 전체적인 trajectory와 RMSE를 통해 재 차 확인할 수 있다.

5. 결 론

본 논문에서 제안한 알고리즘은 multiple Gaussian fitting을 통한 14-bit 이미지 변환, SuperPoint와 SuperGlue를 통한 강건한 특징점 매칭, 원거리 또는 동적 물체 위에 위치한 픽셀의 disparity를 계산에서 제외하는 zero velocity detection 모듈을 통해 열화상-IMU odometry의 성능을 향상시켰다. 알고리즘의 위치 추정 성능은 두 장소의 낮과 밤을 촬영한 4개의 데이터셋에서 state-of-the-art 알고리즘과의 RMSE의 비교를 통해 검증하였다. 제안한 열화상 이미지 전처리 방법은 기존의 clipping 기반의 방법에 비해 특징점 매칭 성능을 눈에 띄게 개선하였으며 이는 매칭 inlier의 수 변화와 알고리즘 적용 전후 RMSE 비교를 통해 확인하였다. 또한, zero velocity update 모듈의 disparity 계산에 사용할 픽셀을 IMU motion 기반으로 선정하여 동적 상황에서도 강건한 zero velocity detection을 구현하였다. 로봇이 RGB 카메라를 사용하기 어려운 제한적 환경에서도 본 논문에서 제안한 알고리즘을 적용하여 강건한 위치 추정을 수행할 수 있을 것으로 기대된다. 추후 SuperPoint, SuperGlue 모델 경량화, 열화상 카메라의 개수를 확장 및 스테



(a)

(b)

(c)

[Fig. 7] Results of the dataset experiment with comparison against OpenVINS (LK Optical flow) and VINS-Mono. (a) Trajectory of the algorithms for each sequence. Groundtruth is obtained through INSPVA. (b) Boxplot of the ATE of translation for each algorithm. (c) Boxplot of the relative rotation error per a meter

레오 정보 취득을 통한 특징점 위치 초기화 방법 개선 등을 통해 알고리즘의 성능을 향상시킬 수 있을 것이다.

References

[1] L. Landrieu and M. Simonovsky, "Large-Scale Point Cloud Semantic Segmentation with Superpoint Graphs," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, pp. 4558-4567, 2018, DOI: 10.1109/CVPR.2018.00479.

[2] P.-E. Sarlin, D. DeTone, T. Malisiewicz, and A. Rabinovich, "Superglue: Learning feature matching with graph neural networks," *IEEE Conference on Computer Vision and Pattern Recognition*, Seattle, WA, USA, pp. 4938-4947, 2020, DOI: 10.1109/CVPR 42600.2020.00499.

[3] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: An efficient alternative to SIFT or SURF," *IEEE International Conference on Computer Vision (ICCV)*, pp. 2564-2571, 2011.

[4] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, pp. 91-110, Nov., 2004, DOI: 10.1023/B:VISI.0000029664.99615.94.

[5] H. Bay, A. Ess, T. Tuytelaars, and L. V. Gool, "Speeded-up robust features (SURF)," *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346-359, Jun., 2008, DOI: 10.1016/

j.cviu.2007.09.014.

[6] P. V. K. Borges and S. Vidas, "Practical infrared visual odometry," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 8, pp. 2205-2213, Aug., 2016, DOI: 10.1109/TITS.2016.2515625.

[7] S. Vidas and S. Sridharan, "Hand-held monocular slam in thermal-infrared," *2012 12th International Conference on Control Automation Robotics & Vision (ICARCV)*, Guangzhou, China, pp. 859-864, 2012, DOI: 10.1109/ICARCV.2012.6485270.

[8] E. Nilsson, "An optimization based approach to visual odometry using infrared images," 2010, [Online] https://www.diva-portal.org/smash/record.jsf?pid=diva2%3A351934&dswid=4224, Accessed: Sept. 24, 2010.

[9] S. Khattak, C. Papachristos, and K. Alexis, "Keyframe-based thermal-inertial odometry," *Journal of Field Robotics*, vol. 37, no. 4, pp. 552-579, Jun., 2020, DOI: 10.1002/rob.21932.

[10] Y.-S. Shin and A. Kim, "Sparse depth enhanced direct thermal-infrared SLAM beyond the visible spectrum," *IEEE Robotics and Automation Letters*, vol. 4, no. 3, pp. 2918-2925, Jul., 2019, DOI: 10.1109/LRA.2019.2923381.

[11] S. Zhao, P. Wang, H. Zhang, Z. Fang, and S. Scherer, "TP-TIO: A robust thermal-inertial odometry with deep thermalpoint," *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Las Vegas, NV, USA, pp. 4505-4512, 2020, DOI: 10.1109/IROS45743.2020.9341716.

[12] J. Poujol, C. A. Aguilera, E. Danos, B. X. Vintimilla, R. Toledo, and A. D. Sappa, "A visible-thermal fusion based monocular visual odometry," *Robot 2015: Second Iberian Robotics Conference: Advances in Robotics*, vol. 417, pp. 517-528, 2015, DOI: 10.1007/978-3-319-27146-0_40.

[13] T. Mouats, N. Aouf, A. D. Sappa, C. Aguilera, and R. Toledo, "Multispectral Stereo Odometry," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 3, pp. 1210-1224, Jun., 2015, DOI: 10.1109/TITS.2014.2354731.

[14] L. Chen, L. Sun, T. Yang, L. Fan, K. Huang, and Z. Xuanyuan, "Rgb-t slam: A flexible slam framework by combining appearance and thermal information," *2017 IEEE International Conference on Robotics and Automation (ICRA)*, Singapore, pp. 5682-5687, 2017, DOI: 10.1109/ICRA.2017.7989668.

[15] S. Khattak, C. Papachristos, and K. Alexis, "Visual-thermal landmarks and inertial fusion for navigation in degraded visual environments," *2019 IEEE Aerospace Conference*, Big Sky, MT, USA, pp. 1-9, 2019, DOI: 10.1109/AERO.2019.8741787.

[16] T. Mouats, N. Aouf, L. Chermak, and M. A. Richardson, "Thermal stereo odometry for UAVs," *IEEE Sensors Journal*, vol. 15, no. 11, pp. 6335-6347, Nov., 2015, DOI: 10.1109/JSEN.2015.2456337.

[17] S. Wang, R. Clark, H. Wen, and N. Trigoni, "DeepVO: Towards end-to-end visual odometry with deep Recurrent Convolutional Neural Networks," *2017 IEEE International Conference on Robotics and Automation (ICRA)*, Singapore, pp. 2043-2050, 2017, DOI: 10.1109/ICRA.2017.7989236.

[18] A. Kendall, M. Grimes, and R. Cipolla, "PoseNet: A Convolutional Network for Real-Time 6-DOF Camera Relocalization," *2015 IEEE International Conference on Computer Vision (ICCV)*, Santiago, Chile, pp. 2938-2946, 2015, DOI: 10.1109/ICCV.2015.336.

[19] I. Melekhov, J. Ylioinas, J. Kannala, and E. Rahtu, "Relative Camera Pose Estimation Using Convolutional Neural Networks," *Advanced Concepts for Intelligent Vision Systems*, pp. 675-687, 2017, DOI: 10.1007/978-3-319-70353-4_57.

[20] D. DeTone, T. Malisiewicz, and A. Rabinovich, "Deep Image Homography Estimation," *arXiv:1606.03798*, 2016, DOI: 10.48550/arXiv.1606.03798.

[21] E. Simo-Serra, E. Trulls, L. Ferraz, I. Kokkinos, P. Fua, and F. Moreno-Noguer, "Discriminative Learning of Deep Convolutional Feature Point Descriptors," *2015 IEEE International Conference on Computer Vision (ICCV)*, Santiago, Chile, pp. 118-126, 2015, DOI: 10.1109/ICCV.2015.22.

[22] K. M. Yi, E. Trulls, V. Lepetit, and P. Fua, "Lift: Learned invariant feature transform," *European Conference on Computer Vision (ECCV)*, pp. 467-483, Sept., 2016, DOI: 10.1007/978-3-319-46466-4_28.

[23] H. Noh, A. Araujo, J. Sim, T. Weyand, and B. Han, "Large-scale image retrieval with attentive deep local features," *2017 IEEE International Conference on Computer Vision (ICCV)*, Venice, Italy, pp. 3456-3465, 2017, DOI: 10.1109/ICCV.2017.374.

[24] D. Li, X. Shi, Q. Long, S. Liu, W. Yang, F. Wang, Q. Wei, and F. Qiao, "DXSLAM: A robust and efficient visual SLAM system with deep features," *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Las Vegas, NV, USA, pp. 4958-4965, 2020, DOI: 10.1109/IROS45743.2020.9340907.

[25] J. Tang, L. Ericson, J. Folkesson, and P. Jensfelt, "GCNv2: Efficient correspondence prediction for real-time SLAM," *IEEE Robotics and Automation Letters*, vol. 4, no. 4, pp. 3505-3512, Oct., 2019, DOI: 10.1109/LRA.2019.2927954.

[26] R. Mur-Artal and J. D. Tardós, "ORB-SLAM2: An Open-Source SLAM System for Monocular, Stereo, and RGB-D Cameras," *IEEE Transactions on Robotics*, vol. 33, no. 5, pp. 1255-1262, Oct., 2017, DOI: 10.1109/TRO.2017.2705103.

[27] H.-J. Liang, N. J. Sanket, C. Fermüller, and Y. Aloimonos, "SalientDSO: Bringing Attention to Direct Sparse Odometry," *IEEE Transactions on Automation Science and Engineering*, vol. 16, no. 4, pp. 1619-1626, Oct., 2019, DOI: 10.1109/TASE.2019.2900980.

[28] J. Engel, V. Koltun, and D. Cremers, "Direct Sparse Odometry," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 3, pp. 611-625, Mar., 2018, DOI: 10.1109/TPAMI.2017.2658577.

[29] A. I. Mourikis and S. I. Roumeliotis, "A Multi-State Constraint Kalman Filter for Vision-aided Inertial Navigation," *2007 IEEE International Conference on Robotics and Automation*, Rome, Italy, pp. 3565-3572, 2007, DOI: 10.1109/ROBOT.2007.364024.

[30] N. Trawny and S. I. Roumeliotis, "Indirect Kalman Filter for 3D Attitude Estimation," *University of Minnesota*, Minneapolis, USA, Rep. 2005-002, Mar., 2005, [Online] https://www.semanticscholar.

org/paper/Indirect-Kalman-Filter-for-3-D-Attitude-Estimation-Trawny-Roumeliotis/56ea520b4a9a7a718c490f392f275f5f8feb2886.

[31] P. Geneva, K. Eckenhoff, W. Lee, Y. Yang, and G. Huang, "OpenVINS: A Research Platform for Visual-Inertial Estimation," *2020 IEEE International Conference on Robotics and Automation (ICRA)*, Paris, France, pp. 4666-4672, 2020, DOI: 10.1109/ICRA40945.2020.9196524.

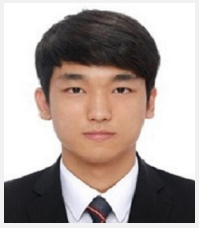
[32] T. Qin, P. Li, and S. Shen, "Vins-mono: A robust and versatile monocular visual-inertial state estimator," *IEEE Transactions on Robotics*, vol. 34, no. 4, pp. 1004-1020, Aug., 2018, DOI: 10.1109/TRO.2018.2853729.

신 재 호

2023 서울대학교 기계공학과(학사)

2023~현재 서울대학교 기계공학과 석사과정

관심분야: Visual Inertial SLAM, Thermal-infrared SLAM



전 명 환

2017 광운대학교 로봇학부(학사)

2020 KAIST 로봇공학제(석사)

2023 KAIST 로봇공학제(박사)

2023~현재 정밀기계설계공동연구소 박사후 연구원

관심분야: Robot vision, Pose Estimation

김 아 영

2005 서울대학교 기계항공공학과(공학사)

2007 서울대학교 기계항공공학전공(공학석사)

2012 미시간대학교 기계공학전공(공학박사)

2014~2021 한국과학기술원 건설및환경공학과 부교수

2021~현재 서울대학교 공과대학 기계공학부 부교수

관심분야: 영상기반 SLAM