

비주얼 오도메트리의 특징점 검출 및 매칭 기법의 발전과 성능 분석

Advancements and Performance Analysis of Feature Detection and Matching Methods in Visual Odometry

최지훈¹, 강창호², 김선영^{3,*}

(Ji Hun Choi¹, Chang Ho Kang², and Sun Young Kim^{3,*})

¹Department of Artificial Intelligence, Sejong University

²Department of AI Robotics, Sejong University

³School of Mechanical Engineering, Kunsan National University

Abstract: This study presents a comprehensive performance analysis of feature detectors and descriptors in visual odometry (VO) based on four key metrics: absolute trajectory error (ATE), relative pose error (RPE), cumulative drift, and frames per second (FPS). Five algorithm-based methods (SIFT, ROOTSIFT, ORB, KAZE, and AKAZE) and five AI-based methods (SuperPoint, DISK, ALIKED, D2NET, and Xfeat) are evaluated using two matching strategies: Brute-Force and LIGHTGLUE. Performance assessment is conducted on the KITTI dataset, demonstrating that KAZE achieves the best results in ATE and cumulative drift, while ALIKED outperforms in RPE and FPS. These findings offer valuable insights for selecting the optimal feature detection and description techniques in real-world VO applications.

Keywords: visual odometry, feature detection, feature description, deep learning, KITTI dataset

1. 서론

현재 위치를 추정할 수 있는 오도메트리(odometry) 기술은 로봇, 자율주행, AR, VR 등 다양한 분야에서 핵심적인 역할을 한다. 특히 LiDAR나 GPS 기반 방법은 높은 정확도를 제공하지만, 고가의 장비를 필요하고 날씨, 물체의 반사율 등 환경적 제약에 영향을 받는 한계가 있다[1]. 이에 반해, 카메라를 사용하는 비주얼 오도메트리(visual odometry, VO)는 비교적 저렴한 비용으로 위치를 추정할 수 있는 대안으로 주목받고 있다. VO는 깊이 정보를 센서로 직접 얻을 수 없는 상황에서 이미지를 기반으로 특징점을 추출하고 매칭한 후, 추가적인 기하학적 분석을 수행하여 위치를 계산하는 과정을 통해 작동한다. 이러한 특징 덕분에 VO는 다양한 환경에서 활용될 수 있으며, 특징점을 검출한 후, 디스크립터(descriptor)를 생성하고 매칭 및 기하학적 계산 단계를 수행한다. 각 단계의 개선이 정확도, 처리 속도와 같은 핵심 성능 지표에 직접적인 영향을 미친다[2]. VO의 단계 중 특징점 검출, 디스크립터 및 매칭은 지속적으로 발전해 왔으며, 최근 딥러닝 기술이 도입되면서 기존 알고리즘 기반 접근법을 뛰어넘는 새로운 가능성이 열리고 있다. 인공지능 기반 기술은 기존 알고리즘이 놓치는 특징점을 탐지할 수 있을 뿐만 아니라, 조명 변화나 텍스처가 부족한 환경에서도 강건한 성능을 발휘하는 장점을 갖고 있다[3]. 그러나 인공지능 기반 기법들은 학습하는 과정에 있어 과적합(overfitting)

문제로 인해 특정 이미지에 대한 의존성이 높아질 수 있으며, 복잡하고 무거운 모델 구조로 인해 실시간 처리에 어려움을 겪는 문제가 존재한다. 따라서 비인공지능 기법과 인공지능 기법의 장단점을 통합적으로 분석하고, 이를 조합하여 비교하는 연구가 필요하다.

본 연구에서는 기존 알고리즘 기법과 인공지능 기법을 조합하여 VO의 특징점 추출 및 매칭 과정을 분석하고자 한다. 또한 기존 연구에서 주로 활용되던 절대 궤적 오차(absolute trajectory error, ATE)기반 평가 방식[1,4]에서 한 단계 더 나아가, 상대 위치 오차를 측정하는 RPE (Relative Pose Error), 각 프레임별 누적 오차를 평가하는 Drift, 그리고 처리 속도를 나타내는 FPS (Frames Per Second)를 포함한 성능 평가 지표를 도입하여 기존 연구와의 차별성을 확보하고자 한다. 이를 통해 실제 주행 데이터셋 KITTI Dataset을 사용하여 각 기법의 강점과 약점을 명확히 분석하고, 각 기법이 알고리즘, 인공지능 분야에서 좋은 성능을 가지는지에 대한 근거를 제공하고자 한다.

본 논문의 구성은 다음과 같다. 2장에서는 VO 파이프라인의 각 단계를 설명하고, 이에 단계별 매칭되는 다양한 기법들을 3장에서 소개한다. 이후, 성능 평가 지표를 활용하여 VO 기법들의 성능을 4장에서 비교하고, 이를 바탕으로 효과적인 기법을 분석하며 향후 연구 방향을 제시하고 5장에서 결론을 맺는다.

* Corresponding Author

Manuscript received December 12, 2024; revised February 14, 2025; accepted February 17, 2025

최지훈: 세종대학교 인공지능학부 학사과정(magon8696@naver.com, ORCID[®] 0009-0001-4893-0624)

강창호: 세종대학교 AI로봇학과 조교수(kcguri@sejong.ac.kr, ORCID[®] 0000-0002-9899-3076)

김선영: 국립군산대학교 기계공학부 부교수(sykim77@kunsan.ac.kr, ORCID[®] 0000-0002-6189-0268)

※ 본 연구는 과학기술정보통신부의 신진연구지원 사업(No. 2021R1C1009219)에서 지원을 받아 수행되었음.

※ 상기 논문은 2024년도 ICROS-KROS 호남제주 그랜드 합동 학술대회에서 초안이 발표되었음.

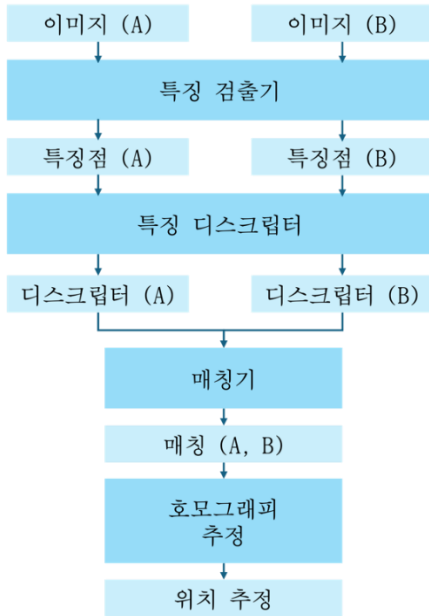


그림 1. VO 파이프라인.

Fig. 1. VO Pipeline.

II. VO 파이프라인

VO는 연속된 이미지 시퀀스(sequence)를 활용하여 카메라의 움직임을 추정하는 기술이다. VO 파이프라인은 그림 1과 같이 4단계로 구성된다[2]. 특징점 검출(feature point detection) 단계에서는 특징 검출기(detector)를 이용하여 코너(corner), 엣지(edge), 블롭(blob) 등 식별에 용이하고 추적 가능한 특징점들을 검출한다. 앞선 과정에 가져온 특징점들을 특징 디스크립터(feature descriptor)를 이용하여 검출된 특징점 주변의 이미지를 벡터로 변환한다. 기법에 따라 디스크립터를 이진 벡터, 실수 벡터로 변환하며, 벡터 변환 전 정규화나 그리드(grid) 계산 등의 전처리 과정을 수행할 수 있다[5]. 특징 매칭 단계에서는 연속된 프레임 간의 특징점의 대응점을 찾는다. 본 연구에서는 Brute-Force Matcher [6]와 2023년에 SUPERGLUE를 경량화하여 만든 LIGHTGLUE [7]를 사용하였다. 호모그래피 추정(homography estimation) 단계에서는 매칭된 특징점들을 기하학적으로 분석하며, 본 연구에서는 노이즈 및 이상치에 강건한 RANSAC을 사용하였다. RANSAC은 Inlier와 Outlier를 구분하여 Inlier에 해당하는 특징점을 이용해 위치를 추정함으로써 오매칭의 영향을 최소화하였다[8].

본 논문에서는 이러한 파이프라인에서 비인공지능 기법과 인공지능 기법들을 조합하여 활용하였으며, KITTI 데이터셋을 이용하여 성능 분석하고 비교하였다.

III. 특징점 검출 및 매칭 기술방식

1. 특징점 검출 비인공지능 기법

1.1 SIFT

SIFT는 2004년에 제안된 특징점 검출 방법으로, 스케일 스페이스에서 가우시안 블러 차이(difference of Gaussian)를 기반으로 특징점을 탐지한다[9]. 안정적인 키포인트들을 얻기 위해서 잠재적인 키포인트에서 정확한 위치를 보정하는

태일러 급수를 활용한다. 추가적으로, 이미지 회전에 대한 강건성을 유지하기 위해 주변 영역의 그래디언트 방향과 크기를 계산한 후, 가장 빈도가 높은 방향을 주 방향으로 설정한다. 위 단계에서 키포인트를 추출한 후, 16×16 픽셀 영역을 4×4 의 서브 영역으로 나눈다. 이후, 각 영역에서 8방향 그래디언트 히스토그램을 생성하여 128차원의 디스크립터를 구성한다. 이런 과정들을 통해 SIFT 기법은 입력 이미지의 크기와 회전, 주변 조명 변화에 강건한 특징점을 검출한다.

1.2 ROOTSIFT

ROOTSIFT는 2012년에 기존 SIFT에서 디스크립터의 비교 방식을 변경하여 성능을 향상시켰다[10]. 기존 SIFT의 특징점 추출 방식을 유지하면서, ROOTSIFT는 디스크립터 비교 방식에서 L1 정규화 후 제곱근을 적용하여 히스토그램 유사도를 개선하는 방식이다. 이는 추가적인 연산이 적어 부담이 없으며 동일한 차원을 유지하여 비교 방식을 개선하였다. 위와 같은 디스크립터는 특징점들을 확률분포로 해석하고 유사도를 계산하기 위해 유클리드(Euclidean) 거리로 계산되어 기존 SIFT 보다 조명, 잡음, 크기 변화에도 강건한 성능을 제공한다.

1.3 ORB

ORB는 2011년에 제안된 기법으로, FAST 검출기의 방향성 부족 문제를 해결하기 위해 특징점의 방향 정보를 추가하고, BRIEF 디스크립터를 개선하여 회전불변성을 확보한다[11]. FAST 검출기에서 방향 정보를 제공하지 않는 문제를 해결하기 위해 검출된 코너 주변 패치에서 밝기 중심을 찾아 그래디언트 강도를 기반으로 특징점 방향을 설정한다. 이후, 계산된 방향 정보를 활용하여 특징점을 정렬한 후, BRIEF 디스크립터를 회전 취약성을 보완한 rBRIEF 디스크립터를 적용하여 이진 특징 벡터를 생성한다. 이는 실수 기반 특징 벡터에 비해 매칭 속도를 빠르게 하는 장점을 가진다. 또한, ORB는 특징 벡터의 균일성을 향상시키기 위해 해밍 거리(Hamming distance)를 기반으로 매칭을 수행하여 계산 효율성을 높인다.

1.4 KAZE

KAZE는 2012년에 제안된 방법으로 비선형 확산 필터링(nonlinear diffusion filtering)을 기반으로 스케일 공간을 구성하여 높은 반복성과 변별성을 가진 특징점 검출 기법이다[12]. 특징점 검출에서는 Scharr 필터를 활용하여 1차 및 2차 미분을 근사화한다. 이 필터는 Sobel 필터의 단점을 개선하여, 커널의 중심에서 멀어질수록 엣지 방향성이 약해지는 현상을 보완함으로써 회전에 강인한 성능을 제공한다. 디스크립터 생성에서는 SIFT보다 성능이 향상된 SURF 기법을 기반으로 발전한 M-SURF (Modified Speeded-Up Robust Features) 방식을 활용한다. M-SURF는 박스 필터와 적분을 사용하여 처리 속도를 높이는 장점이 있다. 또한, KAZE는 비선형 확산 필터 기반 스케일 공간에 적합하도록 특징점 계산 방식을 조정하였다.

1.5 AKAZE

AKAZE는 2013년에 제안된 기법으로, 2012년 KAZE에서 비선형 확산 필터와 스케일 공간 계산의 높은 복잡도를 개선하기 위해 개발되었다[13]. 비선형 확산 필터의 연산 복잡도를

줄이기 위해 시간 간격을 불균등하게 조정하는 방식인 FED (FAST-Explicit Diffusion)와 개선된 M-LDB (Local Difference Binary) 디스크립터를 도입하였다. FED는 명시적과 반암시적 방법의 장점을 결합하여 박스 필터 분해 기반으로 안정적이고 효율적인 비선형 확산 필터링을 가능하게 한다. 또한, M-LDB는 기존 LDB 디스크립터를 개선해 표현력과 스케일 불변성을 강화하였으며, 수평·수직 미분 값을 비교해 이진 벡터를 생성한다. 특히 기존 2비트 비교 방식 대신 3비트 비교 방식을 도입하여 디스크립터 성능을 향상시켰다.

2. 특징점 검출 인공지능 기법

2.1 SuperPoint

SuperPoint는 2018년에 제안된 딥러닝 기반 특징점 검출 기법으로, 자기 지도 학습과 새로운 데이터 증강 기법인 호모그래픽 변환 적응(homographic adaptation)을 통해 특징점 검출의 반복성과 도메인 적응 성능을 크게 개선했다[14]. 네트워크 구조는 이미지 전체를 단일 순방향 패스로 처리할 수 있는 완전 컨볼루션 방식으로 설계되었다. VGG 컨볼루션 네트워크 기반의 인코더는 입력 이미지를 저차원으로 압축하고, 각 픽셀의 특징점 확률을 예측하여 소프트맥스를 통해 점수 맵을 생성한다. 이 점수 맵을 기반으로, 디스크립터 디코더(decoder)는 L2 정규화된 고정 길이 벡터 디스크립터를 생성한다.

호모그래픽 변환 적응은 실제 데이터셋에서 대량의 수작업인 레이블링(labeling) 없이도 훈련 데이터를 생성할 수 있도록 도와주는 기법이다. 입력 이미지에 임의의 다양한 변환을 적용하고, 변환된 이미지에서 특징점을 감지한 후, 변환된 위치를 원본 이미지로 다시 매핑하여 정답 데이터를 만드는 데이터 증강 기법이다.

2.2 D2NET

D2NET은 2019년에 기존의 특징점 검출 후 디스크립터 생성 방식이 아니라 단일 CNN을 통해 검출과 디스크립터를 동시에 생성하는 방식을 설계한 기법이다[15]. 입력 이미지를 단일 CNN으로 특징 맵을 추출하고, 다중 해상도 처리 방법으로 각 스케일의 특징 맵을 통합하여 크기 변화에 강건한 성능을 만들며 각 픽셀에서 추출된 특징 벡터에 L2정규화를 적용하여 디스크립터를 생성한다. 이로 인해 기존 방식에 비해 낮과 밤의 조명 변화에 대해 높은 재현성과 정확도를 유지하는 장점이 있다.

2.3 DISK

DISK는 2020년에 제안된 딥러닝 기반 특징 검출 및 강화 학습을 활용하여 생성된 기법으로, 기존 CNN 기반 접근 방식의 한계를 극복하기 위해 설계되었다[16]. U-NET 기반 네트워크를 사용하여 특징점 후보를 추출한 후, 그리드 단위로 소프트맥스를 적용하여 특징점 확률을 계산하고, L2 정규화된 디스크립터 정보를 상태(state)로 사용한다. 행동(action)은 각 그리드 셀에서 선택된 특징점 위치를 결정하는 것으로 설정되며, 보상(reward)은 두 이미지 간 특징 매칭 결과에 따라 부여된다. 매칭이 성공할 경우 긍정적인 보상을, 매칭이 불가능한 경우 부정적인 보상을 부여하여 정확한 선택을 유도하는 강화학습 기법을 사용한다. DISK는 CNN을 활용해 특징점 검출하고 디스크립터 생성한 후, 매칭 확률

분포를 정책으로 설정하여 정책경사법(policy gradient)으로 최적의 특징점 검출 방법을 학습한다.

2.4 ALIKED

ALIKED는 2023년에 제안된 딥러닝 기법으로 SDDH (Sparse Deformable Descriptor Head)를 활용하여 효율적이며 표현력이 뛰어난 디스크립터를 생성하는 특징이 있다[17]. SDDH는 이미지의 조밀한 디스크립터 대신 검출된 키포인트에서만 디스크립터를 추출하며, 각 키포인트 주변에서 변형 가능한 위치를 학습한다. 이를 통해 지오메트릭(geometric) 불변 디스크립터를 생성하여 불필요한 계산을 줄이면서 표현력을 강화한다. 또한, 다중 스케일 특징 추출과 특징 결합으로 이미 스케일 불변성을 가지고 서브 픽셀 단위를 검출하는 DKD (Differentiable Keypoint Detection)를 사용하여 서브 픽셀 단위의 정밀한 키포인트 검출이 가능하다.

2.5 Xfeat

Xfeat는 2024년에 제안된 경량 딥러닝 기반 로컬 특징 추출 및 매칭 기법으로, 제한된 환경에서도 계산 비용과 해상도를 효율적으로 동작할 수 있도록 설계되었다[18]. 초기 레이어의 채널 수를 최소화한 뒤, 점진적으로 증가시키면서 해상도를 줄이는 방식으로 네트워크의 표현력을 유지한다. 또한, 각 로컬 디스크립터의 매칭 신뢰도를 평가하는 맵을 생성하여 매칭 성능을 보장한다. Xfeat는 Coarse-to-Fine 전략을 적용하여 계산량을 줄이면서도 매칭 성능을 극대화한다. 먼저 낮은 해상도에서 대략적인 매칭을 수행해 전반적인 구조 관계를 파악하고, 이 정보를 기반으로 가장 가까운 디스크립터 쌍을 매칭하며 다음 단계로 고해상도에서 정밀한 매칭을 수행하여 매칭 정확성을 향상시켰다.

3. 특징점 매칭 기법

3.1 Brute-Force

Brute-Force (BF)는 무차별 대입이라는 이름으로 가장 단순하면서도 직관적인 특징점 매칭 방법이다[6]. 특징점 디스크립터 간의 유사도를 측정하기 위해 유클리드 거리와 해밍 거리를 사용한다. 유클리드 거리는 실수형 디스크립터, 해밍 거리는 이진 디스크립터에서 주로 사용된다. 계산된 거리 값 중 가장 작은 값을 갖는 디스크립터 쌍을 매칭 결과로 선택한다.

3.2 LIGHTGLUE

LIGHTGLUE는 2023년에 제안된 특징 매칭 네트워크로, 기존 SUPERGLUE의 복잡한 이미지 매칭, 높은 계산 비용 및 학습 난이도 등 단점을 개선하고자 설계되었다[7]. 기존 단점을 개선하고자 매칭 난이도 따라 계산량을 동적으로 조정하는 적응형 깊이(adaptive depth) 기법을 사용하여 초기 레이어에서 빠르게 결과를 도출하고 어려운 매칭에서는 추가적인 계산을 수행하여 정확도를 높이고 계산 비용을 낮춘다. 또한 매칭 가능성 필터링을 통해 검색 공간을 줄이고 불가능한 매칭점을 제거하는 과정을 추가하였다. 이러한 최적화를 덕분에, SUPERGLUE에 비해 학습 속도와 GPU 리소스가 제한된 환경에서 높은 성능을 보일 수 있었다.

IV. 실험 결과 및 분석

1. 실험 환경

본 연구에서는 다양한 특징점 검출 및 매칭 기법의 성능을

표 1. KITTI 오도메트리 데이터셋(흑백).

Table 1. KITTI odometry dataset (grayscale).

경로	해상도	Frame
00	W: 1241, H: 376	4540
01		1100
02		4660
03	W: 1242, H: 375	800
04	W: 1226, H: 370	270
05		2760
06		1100
07		1100
08		4070
09		1590
10		1200

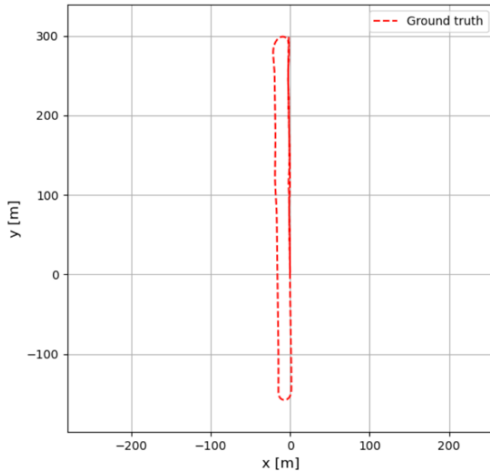


그림 2. KITTI 데이터셋 6번 경로 맵.

Fig. 2. KITTI dataset route map for sequence 6.

분석하기 위해 자율주행 연구에서 널리 활용되는 KITTI 데이터셋을 사용하였다. 특히 오도메트리 성능 분석을 위해 흑백 스테레오 이미지 데이터셋을 활용하였다. 데이터셋의 세부 구성은 표 1에 제시되어 있다.

성능 분석을 위해 데이터셋 중 6번 경로를 선택하였다. 그림 2와 같이 6번 경로의 시각화에서 확인할 수 있듯이, 이 경로는 코너와 직선 구간이 적절히 포함되어 있어 VO 성능 분석에 적합하다. 이를 통해 코너에서는 성능 평가가 가능하며, 직선 구간에서는 일정한 이동을 확인할 수 있다는 점에서 선택하였다.

2. 성능 평가 방법

각 기법의 성능을 분석하기 위해서 경로의 프레임별 실제 위치 (x, y, z) 와 모델의 예측 위치 (x, y, z) 를 가져와 아래의 성능 평가 방법에 사용하였다. 이때 P^{est} 는 예측 위치 (x, y, z) 좌표, P^{gt} 는 실제 위치 (x, y, z) 좌표, N 은 각 경로의 프레임 개수, $Time_{total}$ 은 모델 실행 총 시간으로 정의된다.

2.1 ATE

ATE는 예측한 경로와 실제 경로의 두 경로 사이의 거리를 계산하여 일반적인 주행 기록 평가에 전반적인 경로 예측 오차율을 계산할 수 있다. 모든 시간대에서 발생된 오차

계산식은 오차의 이상치에 민감도가 상대적으로 낮은 평균 제곱근 오차(root mean square error, RMSE)를 사용한다.

$$ATE = \sqrt{\frac{1}{N} \sum_{i=1}^N \|P_i^{est} - P_i^{gt}\|^2} \quad (1)$$

2.2 RPE

RPE는 경로를 시간대 쌍으로 나누어 실제 경로와 예측한 경로의 상대적인 오차를 계산할 수 있는 주행 기록 평가에 사용되는 오차이다. ATE보다 상대적인 정확도를 평가하기 때문에 부분적인 움직임을 정확하게 추적했는지 확인하는데 유용하다.

$$RPE = \frac{1}{N-1} \sum_{i=1}^{N-1} \|(P_{i+1}^{est} - P_i^{est}) - (P_{i+1}^{gt} - P_i^{gt})\| \quad (2)$$

2.3 Drift

Drift는 상대적인 움직임의 연속적인 계산으로 인해 발생하는 오차를 분석하는 방법으로 점점 쌓여서 최종적으로는 실제 위치와 추정된 위치 간의 차이를 평가하는데 사용하기에 장기적인 성능평가를 하는데 사용된다.

$$Drift = \sqrt{E_x^2 + E_y^2 + E_z^2} \quad (3)$$

$$E_x = P_x^{est} - P_x^{gt}, E_y = P_y^{est} - P_y^{gt}, E_z = P_z^{est} - P_z^{gt} \quad (4)$$

본 연구에서는 위 Drift식을 사용하여 전체 프레임의 Drift 값들을 합쳐 누적 Drift를 계산하였다. 이를 통해 각 기법들 간 장기적인 성능 평가를 직관적으로 비교할 수 있도록 하였다.

2.4 FPS

FPS는 초당 프레임의 수를 의미하며 초당 해당 기법들이 몇 프레임을 처리하는 성능을 평가하기 위해서 사용되는 방법으로 초당 프레임 수가 높을수록 그 기법들은 실시간성이 우수한 VO기법이라고 판단된다.

$$FPS = \frac{N}{Time_{total}} \quad (5)$$

3. 성능 분석

표 2는 KITTI 데이터셋의 6번 경로를 대상으로 ATE, RPE, Drift, FPS 결과를 비교하여 각 기법의 성능을 평가한 결과를 제시한 것이다. 이와 더불어, 그림 3은 각 기법이 6번 경로에서

표 2. 기법에 따른 성능 비교.

Table 2. Performance comparison based on techniques.

기법	ATE	RPE	Drift	FPS
SIFT	6.964	0.0401	5567.17	10.6
ROOTSIFT	6.256	0.0380	5207.29	9.2
ORB	4.310	0.0319	3732.60	11.8
KAZE	2.506	0.0221	2545.19	0.9
AKAZE	4.629	0.0368	4195.55	12.5
SuperPoint	4.609	0.0290	4258.94	6.1
DISK	7.424	0.0405	6213.22	4.4
ALIKED	2.951	0.0218	2888.98	14.1
D2NET	5.267	0.0440	3828.55	2.6
Xfeat	11.35	0.0702	8723.92	3.1

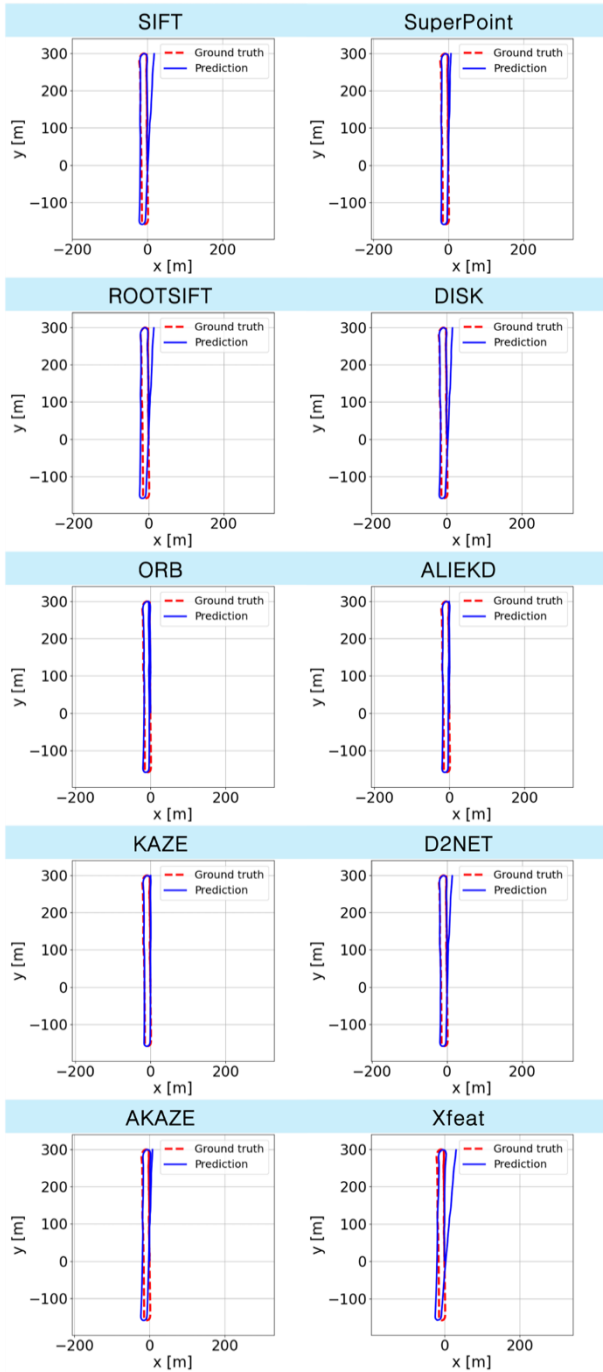


그림 3. 각 기법 실제 경로와 예측 경로 비교 시각화.

Fig. 3. Visualization of actual and predicted path comparison for each technique.

생성한 실제 경로와 예측 경로를 시각화하여 기법 간의 경로 추정 성능 차이를 직관적으로 보여준다. 파란색 선은 예측된 궤적을, 빨간색 점선은 실제 궤적을 나타내며, 이를 통해 특징점 추출 및 매칭 능력의 차이를 명확히 확인할 수 있다. KAZE는 ATE에서 2.506, Drift에서 2545.19로 가장 우수한 성능을 보였으며, ALIKED는 FPS에서 14.1, RPE에서 0.0218로 가장 뛰어난 결과를 기록했다. 반면, Xfeat은 ATE에서 11.35, RPE에서 0.0702, Drift에서 8723.92로 가장 낮은 성능을 보였고, FPS에서는 KAZE가 0.9로 가장 낮은 결과를 기록했다.

전체적인 분포 분석 결과, 알고리즘 기법은 인공지능 기법에 비해 안정적인 성능을 보였다. 이는 오랜 기간 다양한 기법이 추가되고 기존 기법들이 최적화되면서 다양한 환경에서 신뢰성을 입증해 온 점에 기인한 것으로 보인다. 반면, 인공지능 기법은 비교적 최근에 개발된 분야로, 전체적으로 안정적인 결과를 보이지는 않았으나, 최신 기법에서는 우수한 성과를 나타내는 경향을 보였다. 하지만 비교 기법 중 가장 최근에 개발된 Xfeat은 ATE, PRE, 누적 Drift 평가에서 가장 낮은 성능을 기록했다.

최종 실험 결과, KAZE 기법은 ATE와 누적 Drift 항목에서 가장 우수한 성능을 보였으나 가장 낮은 FPS를 보였다. 반면, ALIKED 기법은 RPE와 FPS 측면에서 가장 뛰어난 결과를 나타내었다. 특히 ALIKED 기법은 본 연구에서 사용된 기법들 중 인공지능 기반 기법임에도 불구하고 상대적으로 실시간 성능을 확보하였고 기존 알고리즘 기반 기법들의 격차를 빠르게 줄이고 있음을 보여주고 있다. 그리고 ALIKED는 성능 분석에서 상위권을 차지하여 인공지능 기반 기법들이 성능 측면에서도 지속적으로 발전하고 있음을 확인할 수 있다.

V. 결론

본 연구에서는 VO 파이프라인의 특징점 검출과 특징점 매칭 기법을 비교하였고 KITTI 데이터셋을 활용하여 성능을 분석하였다. 분석 결과, 전통적인 알고리즘 기반 조합들은 전반적으로 안정적이고 우수한 성능을 보여주었는데, 이러한 결과가 나타난 이유는 오랜 기간을 거쳐 다양한 필터와 실수 처리기법에서 이진처리기법처럼 다양한 접근 방법을 통해 최적화되어 다양한 환경에서 신뢰성을 입증해왔기 때문이다. 반면 인공지능 기반 조합은 전반적인 안정성은 부족했으나, 특정 기법에서는 알고리즘 기반 기법을 뛰어넘는 성능을 보였다. 이는 향후 인공지능 기법 연구에서는 안정성을 개선하여 다양한 환경에서 좋은 결과를 보여줄 수 있을 거란 가능성을 제시한다. 특히, ALIKED 기법은 본 연구에서 가장 빠른 결과값을 보이며 대부분의 성능 분석에서 상위 분포하여 좋은 성능을 가져왔다. 이 연구를 통해, 기존 알고리즘 기반 기법과 인공지능 기법의 장단점을 비교 분석하고 향후 인공지능 기반 VO 시스템의 발전 가능성을 제시하였다.

REFERENCES

- [1] G. W. Kim and J. Y. Kang, "3D space object and road detection for autonomous vehicles using monocular camera images and deep learning algorithms," *Journal of Institute of Control, Robotics and Systems (in Korean)*, vol. 30, no. 7, pp. 677-684, 2024.
- [2] A. R. Sahili, S. Hassan, S. Sakhrich, J. Mounsef, N. Maalouf, B. Arain, and T. Taha, "A survey of visual SLAM methods," *IEEE Access*, vol. 11, pp. 139643-139677, 2023.
- [3] J. H. Lee and J. Y. Kang, "3D multi-object tracking using instance segmentation of stereo camera images in autonomous vehicles," *Journal of Institute of Control, Robotics and Systems (in Korean)*, vol. 30, no. 12, pp. 1389-1397, 2024.
- [4] J. H. Choi, C. H. Kang, and S. Y. Kim, "A comparative study of AI and Non-AI feature detection and description methods for

visual odometry,” *Proc. of 2024 ICROS-KROS Honam-Jeju Regional Conference (in Korean)*, pp. 109-110, Kunsan, Korea, Dec. 2024.

- [5] S. Cerri, “Deep descriptors for object recognition,” Politecnico di Milano, Department of Electronics, Information and Bioengineering, Master’s thesis, 2017.
- [6] OpenCV Team, “OpenCV Documentation-Brute-Force Matcher (BFMatcher),” *OpenCV*, 2024. [Online]. Available: <https://docs.opencv.org/>
- [7] P. Lindenberger, P.-E. Sarlin, and M. Pollefeys, “LightGlue: Local feature matching at light speed,” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4567-4577, 2023.
- [8] M. A. Fischler and R. C. Bolles, “Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography,” *Communications of the ACM*, vol. 24, no. 6, pp. 381-395, 1981.
- [9] D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91-110, 2004.
- [10] R. Arandjelovic and A. Zisserman, “Three things everyone should know to improve object retrieval,” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2911-2918, 2012.
- [11] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, “ORB: An efficient alternative to SIFT or SURF,” *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, IEEE, pp. 2564-2571, 2011.
- [12] P. F. Alcantarilla, A. Bartoli, and A. J. Davison, “KAZE features,” *Proceedings of the European Conference on Computer Vision (ECCV)*, Springer, vol. 7577, pp. 214-227, 2012.
- [13] P. F. Alcantarilla, J. Nuevo, and A. Bartoli, “Fast explicit diffusion for accelerated features in nonlinear scale spaces,” *British Machine Vision Conference (BMVC)*, pp. 13.1-13.11, 2013.
- [14] D. DeTone, T. Malisiewicz, and A. Rabinovich, “SuperPoint: Self-supervised interest point detection and description,” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 224-236, 2018.
- [15] M. Dusmanu, I. Rocco, T. Pajdla, M. Pollefeys, J. Sivic, A. Torii, and T. Sattler, “D2-Net: A trainable CNN for joint description and detection of local features,” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 8092-8101, 2019.
- [16] M. J. Tyszkiewicz, P. Fua, and E. Trulls, “DISK: Learning local features with policy gradient,” *Proceedings of the Neural Information Processing Systems (NeurIPS)*, pp. 21078-21089, 2020.
- [17] X. Zhao, X. Wu, W. Chen, P. C. Y. Chen, Q. Xu, and Z. Li, “ALIKED: A lighter keypoint and descriptor extraction network via deformable transformation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, vol. 45, no. 6, pp. 4235-4248, 2023.
- [18] G. Potje, F. Cadar, A. Araujo, R. Martins, and E. R. Nascimento, “XFeat: Accelerated features for lightweight image matching,” *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 12345-12355, 2024.



최지훈

2024년~현재 세종대학교 인공지능학과 학사과정 재학 중. 관심분야는 자율주행, 딥러닝, SLAM, Odometry, NeRF, 컴퓨터 비전, 3D재구성.



강창호

2009년 세종대학교 기계항공우주공학과 졸업. 2016년 서울대학교 공과대학 기계항공공학부 대학원 졸업(공학박사). 2016년~2018년 BK21 플러스 서울대학교 융합지식기반 창조형 기계항공인재양성 사업단 박사후 연구원. 2018년~2019년 고려대학교 공학연구소 연구교수. 2019년~2024년 국립금오공과대학교 기계시스템공학과(항공기계전자융합전공) 조교수. 2024년~현재 세종대학교 AI로봇학과 조교수. 관심분야는 항법 시스템, 딥러닝, 디지털 신호 처리, 비선형 필터링, 다중 목표물 추적 기법.



김선영

1999년 국민대학교 전자공학과 졸업. 2015년 서울대학교 공과대학 기계항공공학부 대학원 졸업(공학석사). 2019년 동 대학원 졸업(공학박사). 2019년~2020년 세종대학교 지능기전공학부 박사후 연구원. 2020년~2023년 국립군산대학교 기계공학부 조교수. 2023년~현재 국립군산대학교 기계공학부 부교수. 관심분야는 항법 시스템, 필터링, 위치추정, 자율주행, 딥러닝, GNSS 전파교란 신호 검출 및 완화 기법, 다중 센서 융합, 다중 목표물 추적, 목표물 검출 및 식별.