



Optical Flow와 Optical Filter 기반 하이브리드 다중 객체 추적 기법

한 건 희¹⁾ · 기 석 철²⁾

충북대학교 스마트자동차공학과 대학원¹⁾ · 충북대학교 지능로봇공학과²⁾

A Hybrid Approach Multi-Object Tracking with Optical Flow and Kalman Filter

Geonhee Han¹⁾ · Seok-Cheol Kee^{*2)}

¹⁾Department of Smart Car Engineering, Graduate School of Chungbuk National University, Chungbuk 28644, Korea

²⁾Department of Intelligent Systems and Robotics, Chungbuk National University, Chungbuk 28644, Korea

(Received 12 March 2025 / Revised 19 May 2025 / Accepted 19 May 2025)

Abstract : Multi-Object Tracking(MOT) is a critical task in computer vision with applications that range from autonomous driving to intelligent surveillance systems. However, achieving reliable tracking remains challenging, particularly in scenarios involving frequent occlusions and complex interactions. This paper proposes an enhanced MOT that combines the predictive power of Kalman filters with the robustness of optical flow for accurate tracking. The framework adopts a two-stage matching strategy based on detection confidence to effectively reduce ID Switches and improve tracking stability. Experiments conducted on the MOT17 dataset demonstrate improvements in tracking metrics, including MOTA and IDF1, across diverse scenarios. Particularly, the framework achieves reduced ID Switches, while maintaining high accuracy and consistency even under complex environments. This study addresses key limitations in existing MOT systems by leveraging a dynamic fusion of Kalman filter predictions and optical flow adjustments, thus enabling robust tracking in real-world applications.

Key words : Multi object tracking(다중객체추적), Computer vision(컴퓨터 비전), Autonomous vehicle(자율주행차), Kalman filter(칼만필터), Intelligent surveillance system(지능형 감시 시스템)

1. 서론

다중 객체 추적(Multi-Object Tracking, MOT)은 연속적인 비디오 시퀀스에서 다수의 객체를 실시간으로 식별하고, 동일 객체에 대해 일관된 ID를 유지하며 지속적으로 위치를 추적하는 기술로, 컴퓨터 비전 분야에서 핵심적인 역할을 수행하고 있다. 특히, 다중 객체 추적 기술은 자율주행, 지능형 교통 시스템(Intelligent Transportation Systems, ITS), 비디오 감시, 스포츠 분석, 군중 모니터링 등 다양한 응용 분야에서 필수적인 요소로 활용된다.^{1,2)} 이 중에서도 자율주행 분야에서 다중 객체 추적은 핵심 안전 기술 중 하나로, 주행 환경 내 다양한 객체(차량, 보행자, 자전거, 오토바이 등)를 정확하게 인식하고 추적하는 역할을 담당하며, 이를 통해 차량 제어, 충돌 회피, 경로 계획 등과 같은 의사결정 과정에 직접적인 영향을 미

친다.³⁾

자율주행 환경은 실시간성이 중요한 특성을 가지며, 복잡한 도심 환경, 다수의 객체가 동시에 존재하는 상황, 빠르게 변화하는 장면, 잦은 폐색(Occlusion) 등으로 인해 다중 객체 추적 기술에 상당한 도전 과제를 제시한다.⁴⁾ 예를 들어, 교차로에서의 신호 대기 상황에서는 다수의 보행자와 차량이 밀집하게 되며, 이 과정에서 특정 객체가 일시적으로 가려지거나 다른 객체와 겹치는 현상이 빈번하게 발생한다. 이러한 상황에서도 자율주행 시스템은 개별 객체를 안정적으로 추적하고, 과거 상태를 기반으로 이후 움직임을 예측하여 안전한 주행을 보장해야 한다. 따라서, 폐색이나 복잡한 상호작용이 존재하는 환경에서도 신뢰성 있는 추적 성능을 유지하는 기술은 자율주행의 안전성과 직결된다고 할 수 있다.

*Corresponding author, E-mail: sckee@chungbuk.ac.kr

¹⁾This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium provided the original work is properly cited.

기존의 다중 객체 추적 기법들⁵⁾은 Kalman filter 기반의 선형 운동 예측 모델을 사용하거나, 위치 정보에 의존한 단순 매칭 방식을 적용하는 기법이 대부분이었다. 대표적인 MOT 알고리즘인 SORT(Simple Online and Realtime Tracking)⁶⁾는 Kalman filter를 이용해 객체의 위치를 예측하고, Hungarian 알고리즘을 통해 프레임 간 객체를 매칭하는 구조로 설계되었다. SORT는 계산 효율성과 실시간 처리 측면에서 강점이 있으나, 복잡한 환경이나 다수의 객체가 존재하는 상황, 특히 폐색이 빈번하게 발생하는 도심 환경에서는 안정적인 성능을 유지하는데 한계가 있다. 이러한 문제를 해결하기 위해, Deep SORT⁷⁾와 같은 방식에서는 객체의 외형 정보(Appearance feature) 기반의 시각적 특징을 통해 매칭 정확도를 높이고 폐색 이후에도 재식별을 가능하게 하였다. 그러나, 외형 정보에 의존할 경우 유사한 외형을 가진 객체 간 오인식이 발생할 수 있고, 계산 비용이 증가하여 자율주행과 같은 실시간 환경에서는 적용에 제한이 존재한다.

이후 ByteTrack⁸⁾과 같은 기법이 등장하여, 낮은 신뢰도의 탐지 결과까지 활용함으로써 더 많은 객체를 안정적으로 추적하는 방법이 제안되었다. ByteTrack은 탐지 신뢰도를 기준으로 High confidence와 Low confidence 객체를 구분한 뒤, 1차와 2차 매칭 과정을 통해 누락되는 객체를 최소화하고, 추적 연속성을 높이는 방식으로 성능을 향상시켰다. 그러나 Kalman filter를 기반으로 한 위치 예측이 여전히 선형 운동 모델에 한정되기 때문에, 급격한 방향 전환이나 비선형 움직임을 보이는 객체에 대해서는 추적 성능이 저하될 수 있다. 또한, 복잡한 도심 환경과 같이 다수의 객체가 존재하며, 상호작용하는 상황에서는 객체 간 ID가 혼선되는 ID Switch 현상이 발생하기 쉬우며, 이는 자율주행 시스템의 안전성과 직결된 문제로 작용할 수 있다.

이에 따라, 본 연구에서는 Kalman filter 기반의 선형 예측을 보완하기 위해 Optical flow 기법을 결합한 하이브리드 접근법을 제안한다. Optical flow는 연속된 프레임 간 픽셀 이동량을 계산하여 객체의 실제 이동 궤적을 반영하는데 효과적이며, 폐색 직전의 움직임을 고려해 폐색 이후에도 객체의 위치를 신뢰성 있게 예측할 수 있는 장점을 가진다. 특히, Lucas-Kanade⁹⁾와 같은 경량화된 Optical flow 기법을 활용하면, 실시간성을 요구하는 자율주행 환경에서도 적용이 용이하다. 본 연구는 이러한 Optical flow의 보강적 역할을 통해 Kalman filter의 선형 모델 한계를 극복하고, 자율주행 차량이 복잡한 교통 환경에서도 안정적이고 신뢰성 있는 다중 객체 추적 성능을 달성할 수 있도록 하는 것을 목표로 한다.

2. 연구방법

2.1 시스템 구성

다중 객체 추적 알고리즘은 객체를 검출하는 Detector¹⁰⁾와 검출된 결과를 바탕으로 객체를 추적하는 Tracker로 구분된다.

본 연구에서는 Detector로 YOLOX¹¹⁾를 활용하였다. Tracker는 Kalman filter와 Optical flow를 결합한 위치 예측 기법을 적용하여 보다 정교한 추적을 수행하도록 설계되었다. 특히, 객체 탐지 신뢰도(Confidence score)를 활용하여 매칭 단계를 고신뢰 객체와 저신뢰 객체로 구분하는 이중 매칭 기법을 적용하였다. 이는 고신뢰 객체를 먼저 매칭한 후, 잔여 객체에 대해 추가적인 매칭을 수행함으로써 트랙 관리의 정밀도를 향상시킨다. 이러한 과정은 Fig. 1에 나타낸 전체 프로세스를 따른다.

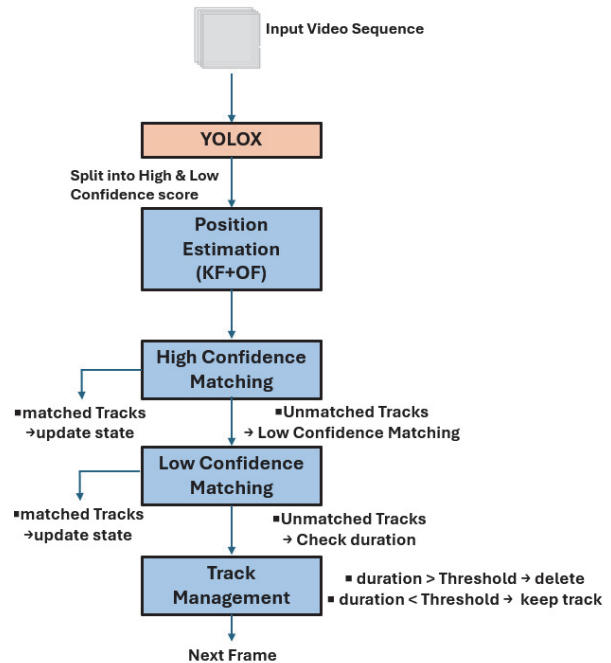


Fig. 1 Flowchart of the proposed method

2.2 객체 검출

YOLOX는 다양한 객체에 대해 우수한 검출 성능을 제공하는 모델로, 본 연구에서는 이를 활용하여 객체의 바운딩 박스 좌표와 신뢰도(Confidence score)를 추출하였다. Detector에서 검출된 객체 정보는 Tracker에 전달되며, 이를 기반으로 객체의 연속적인 추적이 수행된다.

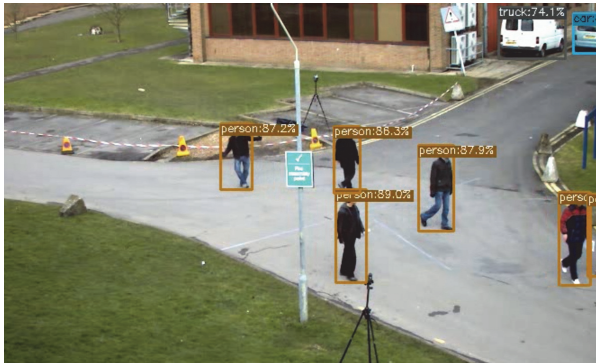


Fig. 2 YOLOX detection results

2.3 객체 추적

2.3.1 Kalman filter와 Optical flow를 이용한 위치 예측

제안하는 기법에서 Kalman filter는 객체의 위치(Position), 속도(Velocity), 가로세로 비율(Aspect ratio), 높이(Height)를 포함하는 상태 벡터를 업데이트하며, Optical flow는 프레임 간 이동량을 계산하여 이를 보정함으로써 보다 정확한 위치 예측을 가능하게 한다.

1) 상태 벡터와 공분산 행렬 초기화

객체 추적을 위해 Kalman filter의 상태 벡터 X 와 공분산 행렬 P 를 초기화 한다. 초기 공분산 행렬 P 은 객체의 위치 불확실성을 나타내며, 각 초기 추정 값에 대한 오차를 반영하여 설정한다.

$$X = \begin{bmatrix} x \\ y \\ a \\ h \\ v_x \\ v_y \\ v_a \\ v_h \end{bmatrix} \quad (1)$$

- x, y : 객체 중심의 위치
- a : 종횡비(aspect ratio)
- h : 높이
- v_x, v_y : 속도 성분
- v_a, v_h : 종횡비 및 높이의 변화율

2) Kalman filter 예측

Kalman filter는 초기화된 상태 벡터와 공분산 행렬을 기반으로 위치를 예측한다.

$$X_t = F \cdot X_{t-1} \quad (2)$$

- X_t : 현재 시간 t 에서 Kalman filter로 예측된 상태 벡터
- F : 상태 전이 행렬, 객체의 상태(위치와 속도)변화 모델을 정의
- X_{t-1} : 이전 시간 $t-1$ 에서 업데이트된 상태 벡터

$$P_t^- = F \cdot P_{t-1}^- \cdot F + Q \quad (3)$$

- P_t^- : 현재 시간 t 의 예측 공분산 행렬
- P_{t-1}^- : 이전 시간 $t-1$ 에서 업데이트된 공분산 행렬
- Q : 프로세스 노이즈 공분산 행렬. 시스템의 불확실성을 반영

3) Optical flow 기반 이동 보정

Kalman filter의 선형 예측 오차를 보완하기 위해 Optical flow를 활용하여 프레임 간 이동량을 계산하고, 이를 Kalman filter의 예측 과정에 반영한다. Optical flow는 Lucas-Kanade 방식을 적용하여 객체의 프레임 간 변위를 추정한다.

$$I_x u + I_y v = -I_t \quad (4)$$

- I_x, I_y : x, y 방향에 대한 기울기
- I_t : 시간에 따른 밝기 변화
- (u, v) : Optical flow 벡터 (픽셀 이동량)

4) 잔차 기반 가중치 조정

제안 기법은 Kalman filter의 선형 예측 위치를 기준으로, Optical flow 기반 이동량을 가중 결합하여 최종 위치를 계산한다. 이때, 두 예측값 간의 차이인 잔차(Residual)의 크기를 기준으로 Optical flow 보정량의 반영 비율을 조절하는 융합 계수 α 를 도입한다. 잔차가 작을수록 Optical flow와 Kalman filter의 예측이 일치한다고 판단하여 Optical flow 보정의 비중을 낮추고, 잔차가 클수록 Kalman filter의 단일 예측을 보정하기 위해 Optical flow의 반영 비율을 높인다. 이처럼 α 는 Kalman filter의 보정량의 조절 계수로 해석할 수 있으며, 최종 위치는 Kalman filter 예측 위치를 기준으로 Optical flow 이동 보정을 가중 결합하여 도출된다. Threshold 값은 Optical flow 기반 Displacement의 평균 분포(약 40 ~ 60픽셀)를 고려하여 실험적으로 50으로 설정하였다. 이 값은 잔차가 작을 경우 Optical flow 보정을 억제하고, 잔차가 커질 경우 보정을 더 강하게 반영하는 기준점으로 작용한다.

$$residual = \| flow_pred - kalman_pred \| \quad (5)$$

flow pred : Optical flow를 이용하여 계산된 프레임 간 이동량에 따른 예측 위치
kalman pred : Kalman filter로 계산된 선형 위치 예측

$$\alpha = 1 - \frac{\text{residual}}{\text{threshold}}, \beta = 1 - \alpha \quad (6)$$

α : Kalman filter 결과에 대한 조절 가중치
 β : Optical flow의 결과에 대한 조절 가중치
 residual : Optical flow와 Kalman filter의 예측 간 차이
 threshold : 잔차를 정규화하기 위한 임계값. Optical flow와 Kalman filter의 조절 가중치 간 균형을 조정

$$X_{final} = \alpha \cdot X_i + \beta \cdot (X_i + \Delta X) \quad (7)$$

X_{final} : Kalman filter와 Optical flow를 결합하여 계산된 최종 위치
 X_i : Kalman filter로 계산된 선형 위치 예측치
 ΔX : Optical flow로 계산된 프레임 간 이동량
 α, β : Kalman filter와 Optical flow 방식의 가중치

2.4 탐지 신뢰도에 따른 매칭

탐지 신뢰도(Score)에 따라 매칭을 두 번에 걸쳐 수행하여, 높은 신뢰도 탐지 결과는 우선 매칭하고, 남은 트랙을 대상으로 낮은 신뢰도 결과와의 매칭을 시도한다.

(1) 높은 신뢰도 탐지 결과의 IoU 계산

IoU(Intersection over Union)는 객체 탐지 및 다중 객체 추적에서 예측된 바운딩 박스와 실제 정답(Ground truth) 박스 간의 일치 정도를 수치적으로 평가하기 위한 대표적인 지표이다. 이는 두 박스의 교집합 영역(Intersection)과 합집합 영역(Union)의 비율로 정의된다. 제안 방식에서는 높은 신뢰도 탐지 결과를 기준으로 예측된 위치와 탐지된 위치 간의 유사도를 IoU로 계산한다.

$$IoU = \frac{|A \cap B|}{|A \cup B|} \quad (8)$$

A : 예측된 박스의 위치
 B : 탐지된 박스의 위치

(2) Hungarian 알고리즘을 통한 최적 매칭

각 예측 객체와 탐지 객체 간의 IoU를 기반으로 매칭 비용 행렬 C 를 정의한다. 헝가리안 알고리즘을 적용하

여 최소 비용 매칭을 수행하고, 매칭된 객체는 상태를 업데이트한다. 여기서 IoU_{ij} 는 I번째 트랙 예측값과 j번째 탐지 결과 사이의 IoU 기반 매칭 비용이다.

$$C_{ij} = 1 - IoU_{ij} \quad (9)$$

(3) 새로운 객체 활성화
 매칭되지 않은 높은 신뢰도 탐지 결과는 새로운 객체로 활성화하여 추적을 시작한다.

(4) 낮은 신뢰도 탐지 결과와의 2차 매칭
 높은 신뢰도 탐지 결과와 매칭되지 않은 트랙을 대상으로, 낮은 신뢰도의 탐지 결과와의 매칭을 시도한다. 이는 기존 트랙이 소실된 경우 낮은 신뢰도의 탐지 결과를 활용하여 다시 활성화하기 위함이다.

(5) 잔여 트랙과 낮은 신뢰도 탐지 결과 간 매칭
 잔여 트랙과 낮은 신뢰도 탐지 결과 간의 IoU를 계산하고, 비용 행렬을 생성하여 헝가리안 알고리즘을 통해 매칭을 수행한다.

(6) 재활성화 및 소실 트랙 관리
 매칭된 트랙은 상태를 갱신하고, 트랙을 재활성화한다. 낮은 신뢰도 탐지 결과와 매칭되지 않은 트랙은 소실 상태로 관리되며 소실 상태가 지속될 경우 최종적으로 제거된다.

3. 실험 구성

3장에서는 제안하는 연구의 실험에 사용된 데이터셋과 평가 지표를 설명한다. 이를 통해 본 연구에서 제안한 알고리즘의 성능을 평가하기 위한 기준과 실험 조건을 명확히 설정하고, 실험의 객관성을 확보하고자 한다.

3.1 데이터셋 : MOT 17

실험에 사용한 데이터셋은 MOT17¹²⁾으로 다중 객체 추적 분야에서 널리 사용되는 벤치마크 데이터셋이다. MOT17 데이터셋은 보행자 추적을 위한 일련의 비디오 시퀀스를 포함하며, 각 프레임마다 객체의 바운딩 박스, ID, 객체 상태 정보 등이 주어진다.

본 데이터셋의 각 시퀀스는 군중 밀도, 조명변화, 객체의 복잡한 움직임, 폐쇄 등 다양한 시나리오를 포함하고 있어 추적 알고리즘의 강건성 등을 평가하기에 적합하다.

또한, 동일한 비디오 시퀀스에 대해 FRCNN(Faster

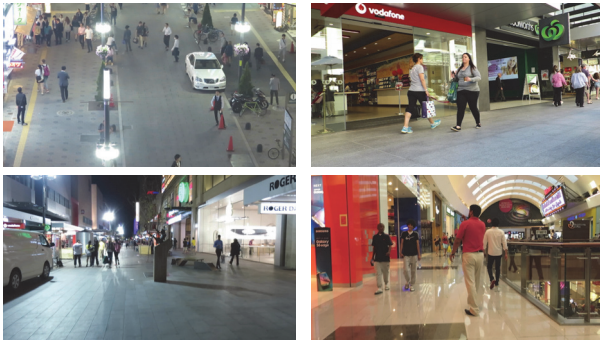


Fig. 3 MOT17 video sequences

R-CNN), DPM(Deformable Parts Model), SDP(Scale-dependant Pooling) 기법으로 주석(Annotation)된 데이터를 제공하여 다양한 탐지 기법에 대한 추적 성능의 영향을 비교할 수 있다.

3.2 성능 지표

본 연구에서 실험에 사용한 주요 지표는 다음과 같다.

- MOTA(Multi-Object Tracking Accuracy): 추적 정확도를 나타내는 지표로, FP(False Positives), FN(False Negatives), IDsw(ID switches)를 포함하여 정의된다. MOTA가 높을수록 추적의 전반적인 정확도가 높음을 의미한다.

$$MOTA = 1 - \frac{FN + FP + IDsw}{GT} \quad (10)$$

- MOTP(Multi-Object Tracking Precision): 추적된 객체의 위치 정확도를 나타내는 지표로, 추적 바운딩 박스와 실제 객체 바운딩 박스 간의 평균 거리를 기반으로 한다. 낮은 MOTP는 추적된 객체의 위치가 실제 위치와 매우 유사하다는 것을 의미한다. 여기서 $d_{i,t}$ 는 시간 t에서 객체 i에 대한 예측 위치와 실제 위치 간의 거리, c_t 는 시간 t에 추적된 객체의 총 수이다.

$$MOTP = \frac{\sum_{i,t} d_{i,t}}{\sum_t c_t} \quad (11)$$

- IDF1(ID F1 Score): 추적된 객체의 ID 일관성을 나타내는 지표로, 추적과 실제 객체 간의 ID 매칭 정확도를 평가한다. IDF1이 높을수록 추적 ID의 일관성이 높음을 의미한다. 여기서 IDTP(ID True Positive)는 올바른 ID로 정확히 추적한 프레임 수를, IDFP(ID False Positive)는 잘못된 ID로 추적한 프레임 수

를, IDFN(ID False Negative)는 ID를 아예 식별하지 못한 프레임수를 뜻한다.

$$IDF1 = \frac{2 \cdot IDTP}{2 \cdot IDTP + IDFP + IDFN} \quad (12)$$

- FP(False Positive): 잘못된 추적 결과로, 실제 객체가 없는 위치에 추적된 객체가 발생한 경우를 의미한다.
- FN(False Negative): 탐지 실패를 의미하며, 실제 객체가 존재함에도 추적되지 않은 경우를 나타낸다.
- IDsw(ID Switch): 추적 과정에서 동일한 객체에 부여된 ID가 바뀌는 횟수를 나타내며, 낮은 IDsw는 동일한 객체에 대해 추적 ID가 일관되게 유지됨을 의미한다.

이와 같은 지표들은 MOT 성능 평가의 주요 요소로, 본 연구의 실험에서도 각 지표를 기준으로 추적 성능을 평가한다.

4. 실험 결과

4장에서는 제안한 다중 객체 추적 기법의 성능을 검증하기 위해 MOT17 데이터셋을 활용한 실험 결과를 제시한다. 실험은 MOTA, MOTP, IDF1, IDsw와 같은 주요 지표를 기준으로 평가되었으며, Detector별 및 개별 비디오 시퀀스별 성능을 통해 제안 방식의 전반적인 성능과 일관성을 확인하였다. 각 결과는 제안 방식이 기존 기법보다 유의미한 성능 개선을 이루었는지를 분석한다.

Fig. 4는 ByteTrack 알고리즘의 추적 결과를 나타낸다. 해당 프레임에서 ID 24번, 27번 객체는 일시적으로 장애물에 의해 가려지는 폐색 현상이 발생하며, 이로 인해 ByteTrack 알고리즘은 ID 27번 객체를 지속적으로 추적하지 못하고 새로운 ID 41번을 할당하였을 뿐만 아니라 두 객체간의 ID가 서로 바뀌었다. 이는 기존 ByteTrack이 폐색 상황에서 객체의 연속성을 충분히 유지하지 못하는 한계를 보여준다.

한편, Fig. 5는 본 연구에서 제안하는 기법을 적용한 추적 결과를 보여준다. 제안된 방법은 각 객체의 바운딩 박스 중심점을 기준으로 Optical flow를 계산하고, 이를 Kalman filter와 결합하여 보다 강건한 추적 성능을 달성하도록 설계되었다. Fig. 5에서 ID 27번, 24번 객체간 폐색이 발생함에도 불구하고, 기존 ID를 유지하며 정상적으로 추적이 이루어지고 있음을 확인할 수 있다.

또한, 객체가 서로 교차하면서 겹치는 상황에서도 ID



Fig. 4 Tracking results of ByteTrack



Fig. 5 Tracking results of proposed method

가 변경되지 않고 지속적으로 유지되는데, 이는 Optical flow를 활용한 보정이 Kalman filter의 예측을 효과적으로 보완한 결과로 해석할 수 있다. 특히, 전방 장애물로 인해 객체의 일부가 가려지는 경우에도 ID switch 없이 연속적인 추적 궤적(Trajectory)이 유지되고 있음을 확인할 수 있다.

본 연구에서는 MOT17 데이터셋을 기반으로 기존 다중 객체 추적 기법과 제안된 방식의 성능을 비교하여 평가를 수행하였다. Table 1은 Kalman filter 계열의 대표적인 알고리즘과의 성능 비교를 나타낸다. 이를 통해 제안된 방식이 기존 기법 대비 얼마나 높은 성능과 안정성을 보이는지 검증하고자 하였다.

Table 1의 결과에서 제안된 방식은 기존 Kalman filter 기반의 주요 Motion 계열 기법들과 비교하여 MOTA와 IDF1에서 가장 높은 성능을 기록하였다. 특히, IDsw 수치가 가장 낮은 수준을 유지했으며, 이는 Optical flow와 Kalman filter의 결합이 효과적으로 작용했음을 의미한다.

Table 1 Comparison of the proposed method with Kalman filter-based motion models

Method	Similarity	MOTA ↑	IDF1 ↑	ID sw ↓
SORT	Motion(K)	74.6	76.9	291
DeepSORT	Motion(K) + Re-ID	75.4	77.2	239
MOTDT	Motion(K) + Re-ID	75.8	77.6	273
ByteTrack	Motion(K)	76.6	79.3	159
Ours	Motion(K)+ Optical flow	76.8	79.4	152

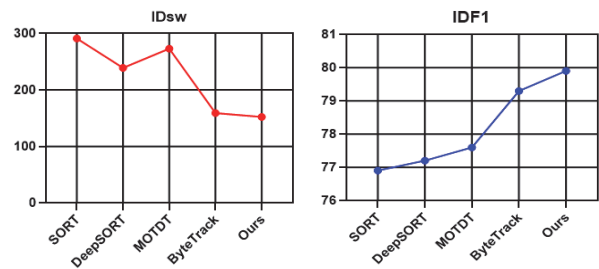


Fig. 6 Performance comparison between Kalman filter-based motion models and proposed method

다. 이러한 성능 향상을 통해 빠른 움직임이나 복잡한 이동 패턴에서도 추적의 일관성을 유지할 수 있었으며, 다중 객체 간의 겹침이나 상호작용이 많은 환경에서도 우수한 성능을 제공할 가능성을 보여준다.

Table 2는 Detector별로 기존 ByteTrack과 제안된 방식을 비교한 결과를 나타내며, 다양한 detector 환경에서의 성능 변화를 분석하여 제안 방식의 안정성과 일관성을 평가하고자 하였다.

Table 2의 결과에서 FRCNN, DPM, SDP detector 모두에서 제안 방식이 기존 ByteTrack 방식 대비 우수한 성능을 기록하였다. 특히 IDsw 수치가 전반적으로 감소한 것은 폐쇄이 빈번한 복잡한 환경에서도 제안 방식이 기존 방식보다 우수한 ID 일관성을 유지할 수 있음을 의미한다. 이를 통해 제안 방식이 다양한 detector 환경에서도 강건한 추적 성능을 제공할 수 있는 가능성을 보여준다.

Table 3은 MOT17 데이터셋의 개별 비디오 시퀀스에서 제안 방식과 기존 ByteTrack 방식을 비교한 결과를 나타내며, 다양한 시퀀스에서 제안 방식이 일관된 성능을 유지할 수 있는지 평가하였다. 본 실험은 제안 방식의 유연성과 견고성을 검증하기 위한 목적을 가진다.

Table 3의 결과에서 MOTP와 MOTA 성능이 전반적으로 향상된 점은 Optical flow와 Kalman filter의 결합을 통해 객

Table 2 Performance comparison of the proposed method and ByteTrack across different object detectors on MOT17 dataset

DATA	Method	MOTP ↓	MOTA ↑	IDF1 ↑	IDsw ↓
MOT17 FRCNN	BYTE	0.123	90.0	83.1	428
	Ours	0.121	90.2	83.3	391
MOT17 DPM	BYTE	0.123	90.0	82.9	431
	Ours	0.122	90.2	82.8	398
MOT17 SDP	BYTE	0.123	90.0	82.8	431
	Ours	0.121	90.2	83.2	393

Table 3 Evaluation of the proposed method and ByteTrack on individual MOT17 sequences

DATA	Method	MOTP ↓	MOTA ↑	IDF1 ↑	IDsw ↓
MOT17-02 FRCNN	BYTE	0.153	81.3	68.0	104
	Ours	0.152	81.5	66.9	108
MOT17-04 FRCNN	BYTE	0.081	98.3	96.1	10
	Ours	0.079	98.4	96.3	8
MOT17-05 FRCNN	BYTE	0.153	81.9	68.3	83
	Ours	0.145	82.4	71.4	71
MOT17-09 FRCNN	BYTE	0.138	87.7	68.3	31
	Ours	0.131	88.7	71.4	24
MOT17-10 FRCNN	BYTE	0.189	79.0	63.3	138
	Ours	0.190	79.4	64.5	125
MOT17-11 FRCNN	BYTE	0.124	90.7	84.5	29
	Ours	0.115	90.9	87.9	20
MOT17-13 FRCNN	BYTE	0.169	87.6	80.8	33
	Ours	0.167	87.5	79.5	35

체 위치 예측의 정확성이 높아졌음을 의미한다. 특히, MOT17-05, MOT17-09, MOT17-11 등의 시퀀스에서 전반적인 수치가 두드러지게 개선되었으며, 이는 제안 방식이 객체의 위치를 보다 정밀하게 예측할 수 있음을 보여준다.

또한, IDsw 수치의 대폭 감소는 제안 방식이 객체의 ID를 보다 일관되게 유지할 수 있음을 증명하며, 이는 다중 객체 간의 복잡한 상호작용이 존재하는 환경에서도 안정적인 추적 성능을 제공할 수 있음을 의미한다.

Fig. 8은 Kalman filter 기반의 ByteTrack 알고리즘과 비교하여, 제안 방식이 시퀀스별로 얼마나 효과적으로 FP 및 FN를 감소시켰는지를 개선율(Improvement)(%)로 시각화한 결과이며, 개선율은 다음과 같이 계산된다.

$$Improvement (\%) = \frac{BYTE - Ours}{BYTE} \times 100 \quad (13)$$

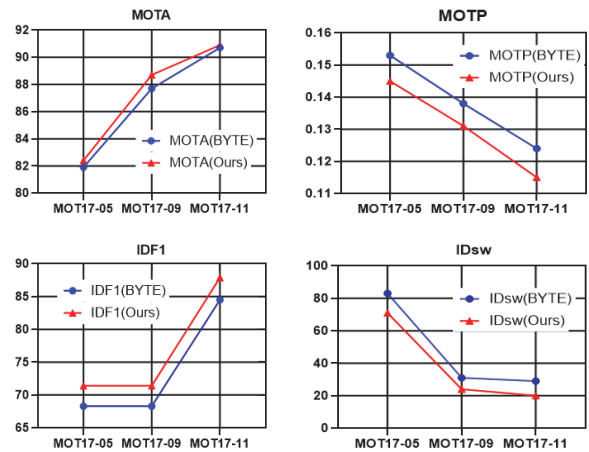


Fig. 7 Performance comparison between BYTE and proposed method

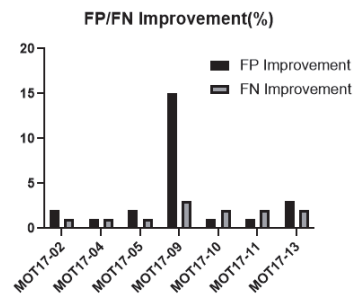


Fig. 8 Comparison of false positive and false negative reduction between BYTE and proposed method

그래프에서 확인할 수 있듯, 제안 방식은 대부분의 시퀀스에서 FP 및 FN을 모두 유의미하게 감소시켰다. 이는 기존 Kalman filter 기반 예측만을 사용한 ByteTrack 방식에 비해, Optical flow 기반 이동 보정 기법이 더욱 정밀한 위치 예측을 가능하게 했음을 나타낸다. 특히 객체 간의 겹침(Occlusion)이나 급격한 움직임이 포함된 시퀀스에서도 ID 일관성 유지와 위치 추정 정확도가 향상되며, 최대 16%에 달하는 성능 개선 효과를 확인할 수 있었다.

5. 결론

본 연구에서는 Optical flow와 Kalman 필터를 결합하여 다중 객체 추적 성능을 개선하는 기법을 제안하였다. MOT17 데이터셋을 활용한 실험 결과, 제안 방식은 기존 기법 대비 MOTA, IDF1 등의 성능 지표에서 우수한 결과를 보였으며, IDsw 감소를 통해 객체의 ID 일관성을 효과적으로 유지할 수 있음을 입증하였으나, 일부 한계점이 존재하며, 이를 개선하기 위한 후속 연구를 고려할 수 있다.

첫째, Optical flow의 성능을 강화하기 위한 연구가 요구된다. 본 연구에서는 Lucas-Kanade 방식을 적용하여 프레임 간 객체 이동을 보정하였으나, 복잡한 상호작용, 비선형적 움직임, 빠른 속도 변화가 발생하는 환경에서는 한계가 발생할 수 있다. 이를 해결하기 위해 RAFT (Recurrent All-Pairs Field Transforms)¹³⁾나 FlowNet¹⁴⁾과 같은 딥러닝 기반 Optical flow 기법을 도입함으로써, 다양한 움직임 패턴에 대한 정밀한 보정이 가능한 시스템으로 확장할 필요가 있다.

둘째, 폐색 상황에 대한 대응력을 강화하는 방안에 대한 고려가 필요하다. Optical flow 기반 이동 보정만으로는 장시간 가려짐이 지속되는 환경에서 안정적인 추적 성능을 유지하기 어렵다. 따라서 Occlusion 지속 시간, 움직임 패턴 등을 고려한 적응적 예측 보강 기법을 도입하거나, 트랙의 신뢰도를 동적으로 평가하고 관리하는 방법을 통해 폐색 상황에서도 견고한 추적이 가능하도록 개선할 수 있을 것이다.

셋째, 실제 자율주행 환경과 같은 실시간 처리 요구에 대응하기 위한 연산 최적화 연구가 필요하다. 제안 기법을 고속으로 수행하기 위해 모델 경량화, 병렬 처리 구조 설계, 불필요한 계산 절차의 최소화 등을 통해 연산 효율성을 극대화하고, 낮은 지연 시간과 적은 자원으로 안정적으로 동작할 수 있는 다중 객체 추적 시스템 개발이 요구된다.

이와 같은 후속 연구를 통해, 다양한 환경과 복잡한 상황에서도 안정적이고 강건한 다중 객체 추적 성능을 확보할 수 있을 것으로 기대되며, 자율주행, 스마트 시티, 지능형 영상 감시 등 실제 응용 분야에서의 활용성을 더욱 높일 수 있을 것으로 사료된다.

후 기

이 논문은 산업통상자원부의 ‘자율주행기술개발혁신사업’으로 지원을 받아 수행된 연구 결과입니다(20014476, 혼합현실기반자율주행부품및시스템평가기술개발).

References

- 1) S. Kim, Y. Kim, H. Jeon, D. Kum and K. Lee, “Autonomous Driving Technology Trend And Future Outlook: Powered by Artificial Intelligence,” Transactions of KSAE, Vol.30, No.10, pp.819–830, 2022.
- 2) R. Ravindran, M. J. Santora and M. M. Jamali, “Multi-object Detection and Tracking, Based on DNN, for Autonomous Vehicles: A Review,” IEEE Sensors Journal, Vol.21, No.5, pp.5668–5677, 2021.
- 3) S. Pal, S. K. Roy and D. P. Dogra, “Deep Learning in Multi-object Detection and Tracking: State of the Art,” The Visual Computer, Vol.37, pp.1619–1637, 2021.
- 4) M. A. Hossain and D. Choi, “Multiple Object Tracking in Robotic Applications: Trends and Challenges,” Applied Sciences, Vol.12, No.19, pp.1–26, 2022.
- 5) Z. Wang, L. Xie and Q. Xu, “Multiple Object Tracking: A Literature Review,” Artificial Intelligence Review, Vol.54, No.1, pp.1–32, 2021.
- 6) A. Bewley, Z. Ge, L. Ott, F. Ramos and B. Upcroft, “Simple Online and Realtime Tracking,” Proceedings of the IEEE International Conference on Image Processing(ICIP), pp.3464–3468, 2016.
- 7) N. Wojke, A. Bewley and D. Paulus, “Simple Online and Realtime Tracking with a Deep Association Metric,” Proceedings of the IEEE International Conference on Image Processing (ICIP), pp.3645–3649, 2017.
- 8) Y. Zhang, P. Sun, Y. Jiang, D. Yu, F. Weng, Z. Yuan, P. Luo, W. Liu and X. Wang, “ByteTrack: Multi-object Tracking by Associating Every Detection Box,” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition(CVPR), pp.198–207, 2021.
- 9) B. D. Lucas and T. Kanade, “An Iterative Image Registration Technique with an Application to Stereo Vision,” Proceedings of Image Understanding Workshop, pp.121–130, 1981.
- 10) S. Hwang, J. Choi, H. Park and J. Kim, “A Study on the Implementation of Camera Sensor for Object Detection Based on Deep Learning and Distance Estimation,” Transactions of KSAE, Vol.30, No.8, pp.659–665, 2022.
- 11) Z. Ge, S. Liu, F. Wang, Z. Li and J. Sun, “YOLOX: Exceeding YOLO Series in 2021,” arXiv preprint arXiv:2107.08430, 2021.
- 12) P. Dendorfer, A. Osep, A. Milan, K. Schindler, D. Cremers, I. Reid, S. Roth and L. Leal-Taixé, “MOT Challenge: A Benchmark for Single-camera Multiple Target Tracking,” arXiv preprint arXiv:2010.07548, 2020.
- 13) Z. Teed and J. Deng, “RAFT: Recurrent All-pairs Field Transforms for Optical Flow,” arXiv preprint arXiv:2003.12039, 2020.
- 14) P. Fischer, A. Dosovitskiy, E. Ilg, P. Häusser, C. Hazırbaş, V. Golkov, P. van der Smagt, D. Cremers and T. Brox, “FlowNet: Learning Optical Flow with Convolutional Networks,” arXiv preprint arXiv:1504.06852, 2015.