

CenterTrack-EKF: 확장된 칼만 필터를 이용한 개선된 다중 객체 추적

(CenterTrack-EKF: Improved Multi Object Tracking with Extended Kalman Filter)

양현성*, 심준보*, 정세훈**

(Hyun-Sung Yang, Chun-Bo Sim, Se-Hoon Jung)

요약

객체 궤적 모델링은 다중 객체 추적(Multi Object Tracking, MOT)의 주요 과제다. CenterTrack은 객체 중심 위치를 추적하는 Heatmap 기반의 방법으로 이를 해결하고자 했다. 하지만 복잡한 움직임과 비선형성을 가진 객체를 추적할 때 제한적인 성능을 보였다. 우리는 CenterTrack의 성능 저하 요인을 보행자의 동적 움직임으로 간주하여 확장된 칼만 필터(Extended Kalman Filter, EKF)를 CenterTrack에 통합했다. 우리가 제안하는 방법의 우수성을 입증하기 위해 기존 칼만 필터(Kalman Filter, KF)와 무향 칼만 필터(Unscented Kalman Filter, UKF)를 CenterTrack에 적용 후 다양한 데이터셋에 비교 평가했다. 실험결과, EKF를 CenterTrack에 통합했을 때 73.7% MOTA(Multiple Object Tracking Accuracy)를 달성하며 CenterTrack에 가장 적합한 필터임을 확인했다.

■ 중심어 : 다중객체추적 ; 칼만필터 ; 확장된 칼만필터 ; CenterTrack

Abstract

Multi-Object trajectory modeling is a major challenge in MOT. CenterTrack tried to solve this problem with a Heatmap-based method that tracks the object center position. However, it showed limited performance when tracking objects with complex movements and nonlinearities. Considering the degradation factor of CenterTrack as the dynamic movement of pedestrians, we integrated the EKF into CenterTrack. To demonstrate the superiority of our proposed method, we applied the existing KF and UKF to CenterTrack and compared and evaluated it on various datasets. The experimental results confirmed that when EKF was integrated into CenterTrack, it achieved 73.7% MOTA, making it the most suitable filter for CenterTrack.

■ keywords : Multi Object Tracking ; Kalman Filter ; Extended Kalman Filter ; CenterTrack

I. 서론

다중 객체 추적(Multi Object Tracking, MOT)은 컴퓨터 비전 분야에서 지속해서 연구되고 있는 과제 중 하나로, 연속적인 비디오 프레임에서 관심 객체 궤적 예측을 목표로 한다. MOT는 자율주행 차량, 지능형 감시 시스템, 증강현실 등 다양한 실생활 애플리케이션에서 중추적인 역할을 담당하고 있다. 그러나 이러한 시스템들은 동적이고 복잡한 환경에서 작동해야 하므로 객체의 가려짐, 조명 변화, 급격한 움직임 등 고려해야 할 요소들이 많다. 따라서 MOT는 복잡한 상황에서도 안정적이고 정확한 추적 성능 달성을 목표로

연구되고 있다. 최근 MOT 연구[1-3]는 딥러닝 기술의 발전에 힘입어 괄목할 만한 성과를 거두고 있다. 특히 TBD(Tracking-by-Detection) 패러다임은 객체 검출과 추적을 분리하여 처리함으로써 MOT 문제를 효과적으로 해결할 수 있는 방법론으로 주목받고 있다. 그러나 기존 TBD 기반 MOT 알고리즘은 Anchor-based 방식의 객체 검출기를 사용하기 때문에 방대한 양의 Bounding Box를 처리해야 한다. 이는 실시간 추적이 요구되는 애플리케이션에 적용하기 어려운 한계로 작용한다. 또한, 고정된 Anchor Box의 사용은 객체의 스케일이나 종횡비 변화에 민감하게 반응하므로 추적 성능 저하를 야기할 수 있다. 이러한 문

* 정회원, 순천대학교 It-bio융합시스템전공

** 정회원, 순천대학교 컴퓨터공학과

이 논문은 2023년 순천대학교 학술연구비(과제번호: 2023-0293) 공모과제로 연구되었음

접수일자 : 2024년 04월 11일

수정일자 : 2024년 05월 02일

게재확정일 : 2024년 05월 07일

교신저자 : 정세훈 e-mail : shjung@scnu.ac.kr

제를 해결하기 위해 Anchor-free 방식의 객체 검출 알고리즘이 제안됐다.

CenterNet[4]은 객체의 중심점을 기반으로 Heatmap을 예측함으로써 Anchor Box의 복잡한 설계 과정을 제거하고, 연산량을 크게 감소시킬 수 있다. 이를 기반으로 한 CenterTrack[5]은 CenterNet의 검출 결과를 프레임 간 매칭하여 MOT 문제에 적용함으로써, 실시간성과 정확성을 모두 만족시키는 우수한 추적 성능을 보여줬다. 그러나 CenterTrack은 프레임 간 객체의 이동을 선형 모션 모델로 가정하기 때문에, 급격한 속도 변화나 방향 전환이 발생하는 비선형적 움직임을 효과적으로 예측하는 데 한계가 있다. MOT에서 이러한 비선형성을 고려하지 않을 경우, 추적 정확성과 강인성이 크게 저하될 수 있다. 따라서 MOT가 실제 환경에 적용되기 위해서는 비선형 움직임을 적절히 모델링할 수 있는 방법론의 도입이 필수적이다.

이에 본 연구에서는 CenterTrack에 EKF(Extended Kalman Filter)[6]를 결합하여, 비선형 시스템에서의 객체 추적 문제를 해결하고자 한다. EKF는 비선형 시스템을 국부적으로 선형화하여 근사하는 방법으로, 객체의 위치와 속도를 동시에 추정할 수 있는 강력한 도구다. 우리는 CenterTrack의 검출 결과를 EKF의 측정값으로 사용하고, 객체의 동적 특성을 상태 벡터로 모델링하여 프레임 간 궤적을 예측한다. 이를 통해 급변하는 객체의 움직임에도 강인한 추적이 가능했다.

제안하는 방법의 우수성을 입증하기 위해 우리는 다양한 벤치마크 데이터셋에 대한 광범위한 실험을 수행했다. 실험결과, EKF를 적용한 CenterTrack은 기존 방법 대비 MOTA(Multiple Object Tracking Accuracy)와 IDF1(Identification F1-score) 등의 주요 성능 지표에서 유의미한 향상을 보였다. 이는 본 연구에서 제안한 비선형 모델링 기법이 MOT 알고리즘의 실용성을 높이는 데 기여할 수 있음을 시사한다.

본 연구는 실시간 MOT 알고리즘에 비선형 상

태 추정 기법을 접목함으로써, 정확하고 신뢰할 수 있는 MOT를 가능케 하는 새로운 방법론을 제시한다는 점에서 의의가 있다.

II. 관련 연구

1. 비선형 움직임 처리

비선형 움직임 처리는 MOT에서 중요한 문제 중 하나다. 주로 객체의 움직임은 객체의 속도나 방향이 급작스럽게 변화하거나 외부 요인에 의해 영향을 받기 때문에 비선형적인 특성을 보인다. 비선형 움직임을 다루기 위한 다양한 방법 중 KF(Kalman Filter, KF)[7]는 선형 시스템에서 상태를 추정하는 데 사용되지만 비선형 시스템에서는 한계가 있다. 이를 극복하기 위해 EKF가 제안됐다. EKF는 비선형 시스템을 현재 상태 근처에서 선형화하여 KF와 유사한 방식으로 상태를 추정한다. 또한, 테일러 급수 전개를 이용하여 비선형 함수를 선형 근사하는 방법에 기초한다. EKF는 비선형 시스템을 다룰 수 있다는 장점이 있지만 몇 가지 한계점도 존재한다. 첫째, 선형화 과정에서 발생하는 근사화 오차로 인해 추정 성능이 저하될 수 있다. 둘째, 초기 추정치가 실제 값과 크게 다른 경우 필터가 발산할 위험이 있다. 이러한 한계를 극복하기 위해 UKF(Unscented Kalman Filter)[8]나 파티클 필터(Particle Filter)[9]와 같은 발전된 형태의 비선형 필터들이 제안됐다. 이들은 EKF보다 정확한 비선형 근사를 제공하지만, 계산 복잡도가 증가한다는 단점이 있다.

2. Multi Object Tracking

MOT는 주로 Convolutional Neural Network와 Transformer 기반의 연구들로 진전되고 있다. [10]은 YOLOv5를 Anchor-free 방식으로 수정하여 최신 추적 패러다임과 동등한 수준으로 개선함으로써 기존 Anchor-based 방식의 MOT 가능성을 제시했다. [11]은 전통적인 Anchor-based 방식을 Anchor-free 방식으로 대체하고 Re-Identification

기능을 통합함으로써 복잡성 감소 및 처리 속도를 향상했다. CenterTrack은 객체 중심점 기반의 추적 알고리즘으로 객체의 위치 및 속도 벡터를 활용하여 프레임 간 객체 궤적을 예측한다. 이를 위해 CenterTrack은 객체 중심점 좌표와 속도 벡터를 예측하는 신경망을 사용한다. 신경망은 이전 프레임과 현재 프레임의 특징 맵을 입력받아 객체의 이동을 예측한다. 이러한 방식은 전통적인 Hungarian 알고리즘에 기반한 데이터 연결 방식과는 다른 접근법이다. 그러나 CenterTrack은 주로 정적인 환경에 최적화된 모델이기 때문에 동적인 움직임과 급격한 변화가 있는 상황에서는 추적 성능이 저하될 수 있다는 한계가 있다.

III. 본 론

우리는 Anchor-free Detection-based 객체 추적의 한계를 인지하고 이를 개선하기 위해 EKF를 적용한 CenterTrack 방식을 제안했다.

$$F = \begin{bmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (1)$$

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \quad (2)$$

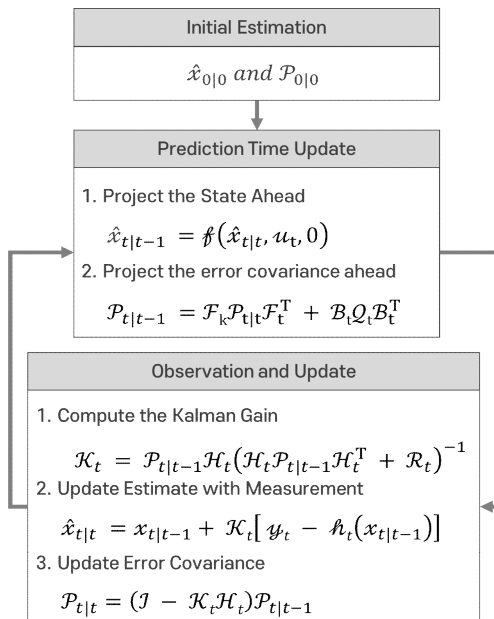


그림 1. EKF 프로세스

기존 CenterTrack이 프레임 내 국부적인 정보만을 활용하여 객체의 움직임을 예측했다면, 제안하는 방법은 연속된 프레임 간의 전역적인 정보를 활용하여 비선형 움직임을 모델링하고자 했다. 상태 전이 행렬 F 는 식 (1)과 같이 정의되며 객체가 등속 운동을 한다고 가정한다. Δt 는 프레임 간 시간 간격을 의미한다. 측정 행렬 H 는 식 (2)와 같이 정의되고, 측정치가 객체의 위치만을 반영하고 속도는 직접 측정되지 않음을 의미한다. 논문에서 사용되는 [6]의 프로세스는 그림 1과 같다. 제안하는 방법은 선형 가정을 기반으로 하고 비선형 동역학을 포착하기 위해 Jacobian 행렬을 이용하여 선형 근사화 방식을 사용한다. F 는 시간 단계의 위치와 속도를 결합함으로써 미래 상태 예측에 사용된다. 따라서 EKF는 측정된 위치 정보를 바탕으로 객체의 위치와 속도를 추정할 수 있다. 예측 단계는 이전 시점 $t-1$ 의 상태 추정치 $\hat{x}_{t-1|t-1}$ 와 오차 공분산 $P_{t-1|t-1}$ 를 바탕으로 현재 시점 t 의 상태를 예측한다. 상태 전이 행렬 F_t 를 이용하여 이전 상태 추정치를 현재 시점으로 투영한 예측값 $\hat{x}_{t|t-1}$ 와 프로세스 노이즈 공분산 Q_t 를 고려하여 예측된 오차 공분산 $P_{t|t-1}$ 을 계산한다. 업데이트 단계에서는 예측값 $\hat{x}_{t|t-1}$ 와 실제 측정치 z_t 의 잔차를 이용해 칼만 이득(Kalman Gain) K_t 를 계산한다. K_t 는 예측 오차 공분산 $P_{t|t-1}$ 과 측정 노이즈 공분산 R_t 의 비율을 고려하여 계산되며 최적의 가중치를 결정하는 역할을 한다. 측정 노이즈 공분산 R_t 는 센서의 노이즈 특성을 나타내는 파라미터로, 측정값의 불확실성을 정량화한다.

업데이트는 K_t 를 바탕으로 상태 추정치 $\hat{x}_{t|t}$ 와 오차 공분산 $P_{t|t}$ 를 계산한다. 업데이트된 상태 추정치는 예측값과 측정값의 잔차에 K_t 를 곱한 값을 예측값에 더함으로써 획득되며, 이를 통해 측정값의 정보를 반영하여 상태 추정치의 정확성을 높인다. 오차 공분산의 업데이트는 K_t 와 H_t 를 사용하여 수행되며 예측된 오차 공분산에서 K_t 와 측정 행렬의 곱을 제거한 값으로 계산된다. 이는 측정값의 정보가 상태 추정에 반영됨에 따라 추정의 불확실성이 감소

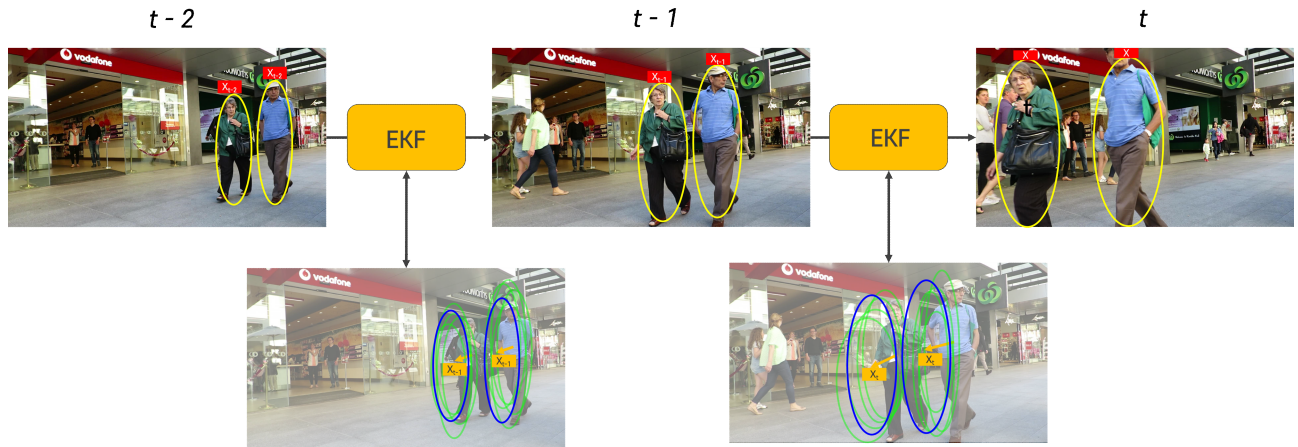


그림 3. EKF를 통한 객체 초기화 과정

함을 의미한다.

한편, 초기 추정치 $\hat{x}_{0|0}$ 와 초기 공분산 $P_{0|0}$ 의 설정도 중요하다. 초기값은 필터가 수렴하는 속도와 안정성에 영향을 줄 수 있기 때문이다. 일반적으로 초기 추정치는 첫 번째 측정치를 이용하여 설정하고, 초기 공분산은 추정치의 불확실성을 반영하여 비교적 큰 값으로 설정한다. 우리는 이러한 EKF의 예측 및 업데이트 과정을 통해 비선형 시스템을 선형화하여 복잡하고 동적인 환경에서도 객체의 상태를 효과적으로 추정할 수 있음을 확인했다. 또한, 프로세스 노이즈 공분산 Q_t 와 측정 노이즈 공분산 R_t 등의 필터 설정을 적절히 조정함으로써 MOT에 최적화된 EKF 기반 모델을 구축하고, 그 유효성을 입증하고자 했다.

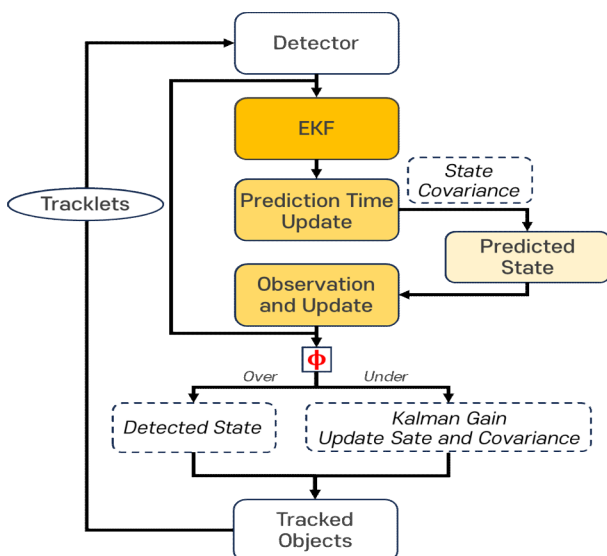


그림 2. 제안하는 방법의 블록 다이어그램

그림 2는 제안하는 방법의 전체 구조를 나타낸 블록 다이어그램으로 [3]의 Matching Cascade를 기반으로 하되, 객체 상태 추정 과정에 EKF를 도입했다. Detector에서 검출된 객체 정보는 EKF의 Observation으로 사용되고, Time Update와 Observation and Update를 통해 정확한 상태 추정값을 얻는다. 기존 CenterTrack은 프레임 간 위치 변화를 기반으로 추적하기 때문에 데이터 연속성이 끊기거나 객체가 일시적으로 사라질 경우 추적 성능이 저하된다. 제안하는 방법은 EKF를 사용하여 추적 상태를 연속적으로 업데이트하고, 프로세스 및 측정 노이즈를 통제하여 추적 정확성을 높인다. EKF를 통한 예측은 추적 대상의 현재 상태와 이전 상태 간의 일치성 ϕ 을 기준으로 수행된다. ϕ 은 현재 프레임에서 검출된 객체 중심 좌표와 EKF에 의해 예측된 객체 중심 좌표 사이의 거리로 정의된다. ϕ 가 임계값보다 작으면 EKF를 통한 추정값을 사용하고, 그렇지 않으면 이전 프레임의 중앙 위치를 사용한다. 최종적으로 추적 객체는 Tracklet으로 피드백된다.

그림 3은 EKF를 CenterTrack에 적용하여 객체 초기화 및 시간에 따른 추적 정보를 업데이트하는 과정을 나타낸 것이다. 각 프레임에서 노란색 타원은 Detector가 검출한 측정치를 의미한다. 파란색 타원은 EKF를 통해 생성된 추정치로써 추적 대상의 위치를 시간에 따라 예측하고 현재 측정치와 비교하여 보정된 것이다. 녹색 타원은 오차 공분산을 나타내며 추적의 불확실성 정도를 의미한다. 화살표는 추적 대

상의 상태가 시간에 따라 EKF를 통해 업데이트되는 과정을 보여주며, 모델 예측과 실제 측정치 간의 차이를 조정하는 역할을 한다. 각 시점의 상태 변수 x 는 예측과 측정치를 결합하여 업데이트된다. 이를 통해 객체의 위치, 속도, 가속도 등의 상태 정보가 지속적으로 추정되고 보정된다. EKF는 이전 시점의 상태 추정치와 현재 시점의 측정치를 바탕으로 K 를 계산하여 상태 변수와 오차 공분산을 업데이트한다. 이 과정에서 측정치와 예측치 간의 불일치를 최소화하는 방향으로 상태 변수가 조정되며, 오차 공분산은 추정의 불확실성을 나타내는 지표로 활용된다.

객체 초기화는 추적 대상이 최초 등장할 때 이전 프레임 관측치를 바탕으로 초기 추정치를 생성한다. 생성된 추정치는 추적 대상 위치 및 속도 등 동역학적 정보를 포함하기 때문에 예측 정확성을 높이고 객체 재등장까지 추적 안정성을 보장할 수 있다.

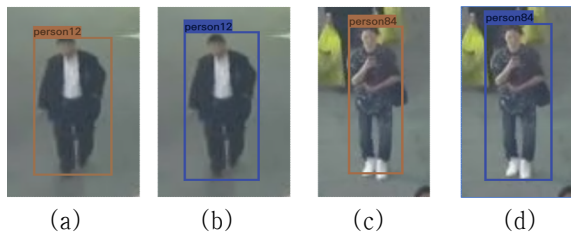


그림 4. EKF를 통한 객체 초기화 결과

그림 4의 (a)와 (c)는 CenterTrack의 객체 초기화고, (b)와 (d)는 EKF를 적용한 객체 초기화다. EKF는 보행자의 동역학적 움직임과 현재 객체의 중앙 위치를 사용하기 때문에 EKF 적용 시 CenterTrack보다 초기화가 정확하다.

Reg Weighted L1 Loss : (3)

$$\frac{\sum |pred \times mask - target \times mask|}{\sum mask + 1e - 4}$$

본 연구는 추적 대상 위치에 중점을 두고 배경 정보로 인한 오차를 최소화하기 위해 식 (3)을 이용했다. 식 (3)은 추적 대상 위치에 가중치를 부여하고 추적 마스크를 활용하여 계산했다. L1을

이용한 오차 계산은 전체 배치에 대한 손실을 최소화하고, 정규화 과정은 모델의 과대적합을 방지하여 추적 성능을 향상했다.

IV. 실험

1. 데이터셋 및 성능평가지표

우리는 보다 정확한 비교를 위해 [5]에서 사용한 데이터셋을 동일하게 사용했다.

MOT16[12]은 총 14개의 비디오 시퀀스로 구성되고, 7개의 훈련 세트와 7개의 테스트 세트로 구성된다. MOT16은 밀집된 보행자가 있는 장면에서 고정된 카메라를 이용해 데이터를 수집했고, 보행자에 대한 주석만 포함하며 프레임 속도는 25~30FPS(Frame Per Second)다. MOT16은 공식 테스트 세트를 제공하지 않기 때문에 훈련 세트에서 일부 분할해 테스트에 이용했다.

MOT17은 MOT16과 마찬가지로 총 14개의 비디오 시퀀스를 포함하며, 7개의 훈련 세트와 7개의 테스트 세트로 구성된다. 데이터는 고정된 카메라를 이용하여 고밀도 인구가 분포하는 곳에서 수집됐다. 주석 처리 및 평가는 보행자로 한정했다. 비디오 프레임 속도는 25~30FPS다. MOT17 역시 공식 테스트 세트를 제공하지 않기 때문에 훈련 세트를 분할하여 일부는 훈련에 사용하고, 나머지는 테스트에 사용했다.

KITTI(Karlsruhe Institute of Technology and Toyota Technological Institute) tracking[13]은 21개의 훈련용 시퀀스와 29개의 테스트용 시퀀스로 구성된다. 데이터는 도로 교통 상황을 촬영한 이동 차량에 설치된 카메라로부터 수집됐다. KITTI tracking은 자동차, 보행자, 자전거에 대한 Bounding Box 주석을 제공하고 공식 평가는 자동차에 집중돼있다. 비디오 프레임 속도는 10FPS다. KITTI tracking은 공식적인 검출 데이터를 제공하지 않기 때문에 사전 훈련된 객체 검출 모델을 이용하여 보행자를 포함한 다양한 객체들을 검출한다. 마찬가지로 전체 훈련용 시퀀스를 훈련 세트와 테스트 세트로 분할 해 실험했다.

표 1. MOT17에서 하이퍼파라미터 ϕ 최적화 결과

ϕ	Recall	Precision	MT	PT	FP	FN	MOTP	MOTA	IDF1
0.1	75.0%	98.4%	52.8%	35.7%	1.2%	25.0%	0.116%	73.0%	71.3%
0.3	75.3%	98.5%	53.4%	34.8%	1.2%	24.7%	0.111%	73.4%	72.2%
0.5	75.5%	98.5%	54.6%	35.1%	1.2%	24.5%	0.111%	73.7%	72.1%
0.7	74.2%	98.8%	50.1%	36.3%	0.9%	25.8%	0.103%	72.7%	71.7%
1.0	75.2%	98.2%	52.5%	36.0%	1.4%	24.8%	0.117%	73.2%	71.5%

본 연구에서는 다중 객체 추적 알고리즘의 성능을 평가하기 위해 다양한 평가지표를 사용했다.

MOTA는 추적 알고리즘의 종합적인 성능을 나타내는 지표로, 객체를 놓치는 경우 FN(False Negative), 객체가 존재하지 않는 위치에서 검출이 발생하는 경우 FP(False Positive), 객체의 ID가 바뀌는 경우 IDs를 모두 고려하여 계산된다. 이 세 가지 요소는 추적 알고리즘의 정확성과 안정성에 직접적인 영향을 미치기 때문에 MOTA는 알고리즘의 실제 활용 가능성을 가늠하는 데 유용한 지표로 활용된다. 구체적으로, MOTA가 높다는 것은 추적 알고리즘이 프레임 내에서 객체를 정확하게 검출하고 객체의 궤적을 안정적으로 유지하며 ID Switch 등의 오류 최소화를 의미한다.

IDF1은 객체 식별 성능에 초점을 맞춘 평가지표로 추적 과정에서 각 객체에 부여된 ID가 얼마나 일관되게 유지되는지를 평가한다. 객체 식별은 여러 객체가 상호작용하며 겹치거나 가려지는 복잡한 상황에서 추적 알고리즘의 견고성을 판단하는 역할을 한다. IDF1이 높으면 추적 알고리즘이 각 객체의 고유한 특성을 잘 파악하여 프레임 간 객체의 일관성을 유지할 수 있음을 나타낸다. 또한, 객체 간 상호작용이 빈번한 실제 환경에서 추적 결과를 신뢰할 수 있는 근거가 된다.

그 외 지표들은 다음과 같다. Recall은 실제 객체 중 올바르게 검출된 객체 비율을 나타내고, Precision은 검출된 객체 중 실제 객체 비율을 나타낸다. 이 둘의 조화 평균인 F1-score는 객체 검출의 정확성과 재현율을 종합적으로 평가한다. MT(Mostly Tracked)는 추적 대상의 80% 이상 성공적으로 추적된 객체

비율을 의미하고, ML(Mostly Lost)은 20% 미만 추적된 객체 비율을 의미한다. PT(Partially Tracked)는 20% 이상 80% 미만 추적된 객체 비율을 나타낸다. MOTP는 추적 결과와 실제 객체 위치 간 평균 거리를 나타내어 추적 정밀도를 평가한다.

2. 실험결과

제안하는 방법은 CenterTrack을 기반으로 작동하기 때문에 CenterTrack과 동일한 하이퍼파라미터 설정을 유지했다. 한편, EKF의 도입으로 인한 모델 복잡도 증가와 비선형성 학습의 필요성 때문에 Epoch만 140으로 다르게 설정하여 훈련했다. 우리의 실험은 2020년 8월 21일에 공개된 CenterTrack 실험결과와 비교했다. CPU는 Intel i7-13700KF를 사용했고 GPU는 NVIDIA RTX A6000 2개를 사용했다. 표 1은 하이퍼파라미터 ϕ 의 최적값을 찾기 위한 MOT17 실험결과다. ϕ 의 범위는 0.1부터 1.0까지 설정했고, 각 설정에 대한 추적 성능을 비교했다. 실험결과, ϕ 가 0.5일 때 가장 높은 MOTA를 보였고 IDF1은 다른 값들과 비교할 때 상대적으로 높았다. 따라서 본 연구는 하이퍼파라미터 ϕ 를 0.5로 설정하여 모든 실험을 진행했다. 우리는 KF와 관련된 필터를 CenterTrack에 적용할 때 초기 추정치의 불안정성이 중심 좌표 추정의 오차 범위를 증가시킨다고 판단했다. 따라서 중심 좌표 x, y 가 모두 ϕ 의 오차 범위를 초과하지 않는 조건에서 실험했다. KF를 적용한 성능 지표들은 다른 방법들보다 우세한 결과를 보였다. 한편, ID 관련 지표의 낮은 성능은 실시간 처리에 중점을 둔 온라인 추적 알고리즘의 한계로 볼 수 있다.

표 2. MOT17에서 제안하는 방법과 CenterTrack의 성능 비교

Method	Recall	Precision	MT	PT	FP	FN	MOTP	MOTA	IDF1
Center Track	74.7%	98.4%	52.8%	35.4%	1.2%	25.3%	0.110%	72.8%	71.9%
KF	74.7%	98.6%	54.3%	34.8%	1.1%	25.3%	0.103%	72.9%	72.2%
UKF	75.0%	98.5%	54.0%	35.4%	1.1%	25.0%	0.104%	73.2%	72.1%
EKF	75.5%	98.5%	54.6%	35.1%	1.2%	24.5%	0.111%	73.7%	72.1%

표 3. MOT16에서 제안하는 방법과 CenterTrack의 성능 비교

Method	Recall	Precision	MT	PT	FP	FN	MOTP	MOTA	IDF1
Center Track	75.3%	98.3%	54.3%	36.4%	1.3%	24.7%	0.110%	73.3%	72.4%
KF	74.9%	98.4%	53.1%	37.7%	1.2%	25.1%	0.109%	73.1%	72.4%
UKF	75.3%	98.2%	53.4%	36.4%	1.4%	24.7%	0.110%	73.2%	72.2%
EKF	75.6%	98.4%	53.4%	37.3%	1.2%	24.4%	0.111%	73.7%	72.8%

표 4. KITTI tracking에서 제안하는 방법과 CenterTrack의 성능 비교

Method	MOTA	Recall	Precision	F1-score	MT	ML	PT
Center Track	66.4%	74.6%	91.3%	82.1%	46.4%	13.1%	40.4%
UKF	67.1%	74.7%	91.9%	82.4%	46.4%	13.1%	40.4%
KF	64.6%	74.6%	89.5%	81.4%	46.4%	15.4%	38.1%
EKF	67.4%	75.0%	91.9%	82.6%	44.0%	13.1%	42.8%

KF 적용은 CenterTrack과 동일한 Recall을 보였지만, MOTA와 MOTP에서 약간의 우위를 보였다. 추적 대상의 시간에 따른 상태 추정과 관측 데이터와의 오차 최소화 과정에 기인한 결과로 판단된다.

UKF 적용은 EKF와 비교하여 1% 미만의 성능 차이를 보였다. 이는 trial에 따른 편차로 볼 수 있기 때문에 우리는 EKF에 집중하여 성능을 분석하고자 한다.

EKF 적용은 75.5% Recall과 98.5% Precision을 기록했다. 이는 비선형 모델을 선형화하여 처리하는 능력이 추적 과정에서 발생할 수 있는 오차를 감쇠시킬 수 있음을 보여준다. 또한, 추적 대상의 움직임을 선형화하는 과정은 추적 정확성 향상에 기여함으로써

다른 방법들보다 높은 73.7% MOTA를 달성했다. EKF는 매 시점 선형화 지점을 변경하며 추정하므로 비선형성으로 인한 추정 오차를 일부 상쇄할 수 있기 때문에 다른 방법들 대비 우수한 성능을 보였다. 따라서, EKF 특성은 추적 성능 향상에 주요한 역할을 한 것으로 분석된다.

표 3과 4는 각각 MOT16과 KITTI tracking에 대한 결과다. 표 3에서 CenterTrack은 75.3% Recall, 98.3% Precision, 73.3% MOTA를 보였다. EKF 적용은 75.6% Recall, 98.4% Precision, 73.7% MOTA로 다른 방법보다 우수한 성능을 보였다. 표 4에서 EKF 적용은 67.4% MOTA와



그림 5. 제안하는 방법과 CenterTrack의 MOT 17 추적 결과 비교: (a) CenterTrack 추적 결과, (b) EKF를 CenterTrack에 적용한 추적 결과

82.6% F1-score를 달성하며 CenterTrack 대비 각각 1.0%와 0.5%의 성능 향상을 보였다. 논문에서 사용하는 EKF는 선형 가정을 기반으로 하면서 비선형성을 선형화된 모델로 근사할 수 있기 때문에 다른 방법들 대비 KITTI tracking 데이터셋에서 우수한 성능을 보인 것으로 해석된다. 하지만, EKF는 선형화 과정에서 고차항을 무시하기 때문에 강한 비선형성이 있는 경우 추정 성능이 저하되어 다른 방법 대비 PT가 높았다.

표 5는 MOT17 데이터셋에 대해 제안하는 방법과 기존 객체 추적 모델의 MOTA, IDF1 성능을 비교한 결과다. 제안하는 방법은 ByteTrack을 제외한 다른 모델들에 비해 MOTA와 IDF1에서 경쟁력 있는 결과를 보였다. 이러한 이유는 EKF의 비선형 상태 추정 능력이 복잡한 추적 환경에서도 객체의 동적 움직임을 효과적으로 모델링할 수 있음을 시사한다. 한편, 제안하는 방법의 MOTA와 IDF1이 ByteTrack보다 저조한 이유는 다음과 같다. ByteTrack은 연속된 프레임 간의 검출 결과를 결합하는 이중 연결 알고리즘을 통해 단절된 트랙을 효과적으로 재결합하고 폐색과 검출 실패에도 견고한 추적을 유지한다.

반면, CenterTrack은 검출 결과를 단일 단계로 처리하기 때문에 빠른 움직임이나 장기간의 폐색에 취약하여 추적 연속성을 유지하기 어렵다. 제안하는 방법은 CenterTrack의 기존 파이프라인을 기반으로 작동하기 때문에 Matching 알고리즘 차이를 고려하면 MOTA와 IDF1 측면에서

표 5. MOT17에서 Private Protocol에 따른 EKF CenterTrack과 비교 모델 간의 성능 비교

Other Method	MOTA	IDF1
CenterTrack[5]	72.8%	71.9%
Tracktor++[14]	53.5%	52.3%
TraDes[15]	69.1%	63.9%
QDTrack[16]	68.7%	66.3%
MOTR[17]	73.4%	68.6%
FairMOT[18]	73.7%	72.3%
ByteTrack[19]	80.3%	77.3%
Proposed	73.7%	72.1%

도 우수한 성능이라는 것을 알 수 있다.

그림 5는 MOT17 데이터셋에 제안하는 방법의 검출 및 추적 능력을 시각적으로 보여준다. (a)와 (b)는 기존 CenterTrack과 EKF를 적용한 CenterTrack의 결과로, (b)에서 객체 초기화가 정확하게 이루어짐을 보여준다. 객체 초기화가 개선된 이유는 상태 추정 시 발생하는 오차를 감소시켜 초기 추적 단계의 신뢰도를 높이기 때문이다.

V. 결 론

우리는 CenterTrack의 객체 추적 성능 향상을 위해 EKF를 도입하여 복잡한 환경에서 객체의 비선형적 움직임을 효과적으로 추정하는 방법을 제안했다. 우리는 CenterTrack의 객체 검출 결과를 EKF의 측정값으로 사용하고, 객체의 위치와 속도를 상태 벡터로 설정하여 객체 궤적을 예측했다. EKF는

비선형성을 국부적으로 선형화함으로써, 객체 상태를 예측하고 업데이트하는 과정에서 비선형성을 처리할 수 있었다. 실험결과, EKF를 적용한 CenterTrack이 기존 방법보다 MOTA, MOTP, IDF1 등의 주요 평가 지표에서 우수한 성능을 보였다. 이는 EKF가 객체의 비선형적 움직임을 효과적으로 모델링하고 측정 노이즈와 예측 노이즈를 적절히 처리할 수 있음을 입증한다.

제안하는 방법은 MOT17, MOT16, KITTI tracking 등 다양한 벤치마크 데이터셋에서 경쟁력 있는 성능을 보였다. 이는 제안하는 방법이 실제 환경에서 발생할 수 있는 다양한 추적 상황에 효과적으로 대응할 수 있음을 시사한다. 한편, 제안하는 방법은 ByteTrack과 같은 최신 MOT 알고리즘에 비해 MOTA와 IDF1 측면에서 다소 미흡한 성능을 보였다. 이는 CenterTrack의 단일 단계 매칭 알고리즘이 갖는 한계로 인해 빠른 움직임이나 장기간 폐색에 취약하기 때문이다. 향후 이러한 한계를 극복하기 위해 다중 매칭 알고리즘을 도입하거나, EKF 외에도 다양한 비선형 필터링 기법을 MOT 문제에 적용하여 성능을 개선할 계획이다.

REFERENCES

- [1] Xufeng Lin, Chang-Tsun Li, Victor Sanchez and Carsten Maple, "On the detection-to-track association for online multi-object tracking," arXiv preprint, <https://doi.org/10.48550/arXiv.2107.00500>, 2021.
- [2] Kaer Huang and et al., "ReIDTrack: Multi-Object Track and Segmentation Without Motion," arXiv preprint, <https://doi.org/10.48550/arXiv.2308.01622>, 2023.
- [3] Chen Xuesong and et al., "TrajectoryFormer: 3D Object Tracking Transformer with Predictive Trajectory Hypotheses," arXiv preprint, <https://doi.org/10.48550/arXiv.2306.05888>, 2023.
- [4] Kaiwen Duan and et al., "Centernet: Keypoint triplets for object detection," *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 6569-6578, 2019.
- [5] Xingyi Zhou, Vladlen Koltun and Philipp Krähenbühl, "Tracking objects as points," *European Conference on Computer Vision Springer International Publishing*, pp. 474-490, 2020.
- [6] RIBEIRO and Maria Isabel, "Kalman and Extended Kalman Filters: Concept, Derivation and Properties," *Institute for Systems and Robotics*, Vol. 43, No. 46, pp. 3736-3741, 2004.
- [7] Simon Dan, "Kalman filtering," *Embedded systems programming*, Vol. 14, No. 6, pp. 72-79, 2001.
- [8] Wan Eric and Van Der Merwe Rudolph, "The unscented Kalman filter for nonlinear estimation," *Proceedings of the IEEE adaptive systems for signal processing*, pp. 153-158, 2000.
- [9] Djuric Petar M and et al., "Particle filtering," *IEEE signal processing magazine*, Vol. 5, No. 20, pp. 19-38, 2003.
- [10] Jie Cao, Jianxun Zhang, Bowen Li, Linfeng Gao and Jie Zhang, "RetinaMOT: rethinking anchor-free YOLOv5 for online multiple object tracking," *Complex & Intelligent Systems*, pp. 1-19, 2023.
- [11] Shuting He, Hao Luo, Pichao Wang, Fan Wang, Hao Li and Wei Jiang, "TransReID: Transformer-based Object Re-Identification," arXiv preprint, <https://doi.org/10.48550/arXiv.2102.04378>, 2021.
- [12] Anton Milan, Laura Leal-Taixe, Ian Reid, Stefan Roth and Konrad Schindler, "MOT16: A benchmark for multi-object tracking," arXiv preprint, <https://doi.org/10.48550/arXiv.1603.00831>, 2016.
- [13] Geiger Andreas, Philip Lenz and Raquel Urtasun, "Are we ready for autonomous driving? the kitti vision benchmark suite," *IEEE conference on computer vision and pattern recognition*, pp. 3354-3361, 2012.
- [14] Bergmann Philipp, Tim Meinhardt and Laura Leal-Taixe, "Tracking without bells and whistles," *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 941-951, 2019.
- [15] Jialian Wu and et al., "Track To Detect and Segment: An Online Multi-Object Tracker," *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 12352-12361, 2021.
- [16] Pang Jiangmiao and et al., "Quasi-dense similarity learning for multiple object tracking," *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021.
- [17] Fangao Zeng and et al., "MOTR: End-to-End Multiple-Object Tracking with Transformer," *European Conference on Computer Vision*, pp. 659-675, 2022.
- [18] Yifu Zhang, Chunyu Wang, Xinggang Wang, Wenjun Zeng and Wenyu Liu, "Fairmot: On the fairness of detection and re-identification in multiple object tracking," *International Journal of Computer Vision*, pp. 3069-3087, 2021.
- [19] Zhang Yifu and et al, "Bytetrack: Multi-object tracking by associating every detection box," *European Conference on Computer Vision*, pp. 1-21, 2022.

저 자 소 개

**양현성(준회원)**

2022년 순천대학교 멀티미디어공학과
졸업(공학사)

2022년~현재 순천대학교 스마트융합
학부 멀티미디어공학전공
석박사통합과정

<관심분야 : 컴퓨터 비전, 딥러닝, 다중객체추적>

**심춘보(정회원)**

1996년 전북대학교 컴퓨터공학과
졸업(공학사)

1998년 전북대학교 컴퓨터공학과
졸업(공학석사)

2003년 전북대학교 컴퓨터공학과
졸업(공학박사)

2005년~현재 순천대학교 인공지능공학부 교수
<관심분야 : 빅데이터, 블록체인, 딥러닝, 생성모델,
자연어 처리, 강화학습>

**정세훈(정회원)**

2012년 순천대학교 멀티미디어공
학과 졸업(공학석사)

2017년 순천대학교 멀티미디어공
학과 졸업(공학박사)

2018년 영산대학교 빅데이터융합
전공 조교수

2020년 안동대학교 창의융합학부

조교수

2022년~현재 순천대학교 컴퓨터공학과 조교수
<관심분야 : 소프트웨어공학, 강화학습, 블록체인, 딥러닝,
데이터 마이닝, 빅데이터 분석 및 예측>