

## LSTM과 칼만 필터를 활용한 폐색에 강인한 다중 객체 추적 알고리즘

\*이재훈 박경수 김병조

Zuminternet 부설연구소

\*ljh0128@zuminternet.com

## Robust Multiple Object Tracking Algorithm for Occlusion Handling using LSTM and Kalman Filter

Jaehoon Lee Gyoungsoo Park Byeongjo Kim

Zuminternet Research Institute

## 요약

비디오 영상으로부터 객체를 추적하는 문제에 있어서 폐색은 오늘날까지도 해결해야하는 문제 중 하나다. 폐색이란 영상 속 찾고자 하는 객체가 이전 프레임에서는 존재했지만 특정 프레임에서는 전경 혹은 다른 객체에 의해 가려져 모습이 보이지 않는 것을 의미한다. 폐색이 나타난 상황에서 해당 객체를 추적하기 위해서는 이전 프레임까지 추적된 정보를 바탕으로 영상에 다시 객체가 나타날 때까지 위치를 잘 예측해야 한다.

본 논문은 비디오 영상의 폐색 환경에 강인한 다중 객체 추적 알고리즘을 제시한다. 이를 위해 딥러닝 기반의 LSTM 구조를 활용하여 객체의 형태 정보를 학습하고 칼만 필터를 이용해 객체의 속도 정보를 학습한다. 두 정보를 조합하여 폐색이 발생하였을 때 객체의 형태와 위치를 예측하여 영상 속에 객체가 다시 등장하더라도 추적 성능을 최대화 한다.

## 1. 서론

다중 객체 추적 알고리즘은 자율주행[1]이나 영상 보안, CCTV[2] 등 다양한 분야에서 적용될 수 있는 기술로 오늘날까지 계속 발전해오고 있다. 다중 객체 추적에 있어서 해결해야하는 문제점들 중 하나는 바로 폐색이다. 폐색이란 특정 대상이 다른 물체에 의해 가려져 일부 또는 전체가 보이지 않는 것을 말한다. 카메라는 3차원의 실세계를 2차원의 영상으로 표현하기 때문에 객체가 전경이나 다른 객체에 의해 폐색이 발생할 경우 잠시 영상 속에서 사라지게 된다. 만일 영상 속에서 찾고자 하는 객체에 폐색이 발생할 경우 이전에 찾았던 정보를 바탕으로 다시 모습이 보일 때까지 예측하여 찾아내야 한다.

본 논문은 다음과 같이 구성된다. 먼저, LSTM 기반의 객체의 형태 정보 모델을 제시한다. 둘째, 칼만 필터 기반의 객체의 속도 정보 모델을 제시한다. 셋째, 두 정보 모델을 조합하여 폐색에 강인한 다중 객체 추적 알고리즘을 제시한다. 마지막으로, 제시한 알고리즘의 정성적 평가와 결론을 도출한다.

## 2. 제안하는 알고리즘

비디오 영상 속 객체 추적에 있어서 이전 프레임까지의 정보를 바탕으로 현재 프레임의 위치를 예측할 때 객체의 형태 정보를 아는 것은 중요하다. 보통 객체의 형태 정보에 있어서 HoG 특징점[3]과 같은 기울기 정보나 히스토그램 특징점[4]과 같은 색상 정보를 사용한다. 하지만 이들의 문제점은 영상 속 객체의 모습이 변한다는 것을 처리하지 못하는 것이다. 3차원의 실세계를 2차원의 카메라 영상으로 표현하기 때문에 회

전이나 원근에 의해 객체가 영상에 나타나는 모습이 달라질 수밖에 없다. 따라서 이전 프레임들의 정보를 통해 객체의 모습을 학습할 필요가 있고 현재 프레임에 객체의 변화될 모습을 예측하는 것이 필요하다. 이를 위해 본 논문에서는 LSTM 기반의 Re3 추적 알고리즘[5]을 활용하여 객체의 형태 정보를 모델링하였다. Re3 추적기는 딥러닝 기반의 다중 객체 추적 알고리즘으로 합성곱 신경망(CNN)을 이용해 객체의 형태 정보를 추출할 수 있는 필터를 자동으로 학습한다. 그리고 Long Short-Term Memory(LSTM) 구조를 통해 이전 입력 영상으로부터 추출된 정보를 기억하여 일련의 프레임 동안 나타나는 객체의 형태 정보를 학습할 수 있다. 따라서 찾고자 하는 객체의 모습이 변하거나 폐색이 발생해 일부 또는 전체가 가려지더라도 이전 프레임에서 학습한 객체의 형태 정보를 바탕으로 예측을 할 수 있게 된다.

하지만 단순히 객체의 형태 정보만을 이용해 다중 객체 추적을 수행할 경우 문제가 발생할 수 있다. 예를 들어, 두 객체가 유사한 모습을 하고 있을 경우 같은 객체로 판단할 수 있고 객체가 전경에 의해 긴 시간 동안 가려질 경우 학습된 형태 정보가 사라져 추적에 실패할 수가 있다. 따라서 이를 처리하기 위해 본 논문에서는 Sort 추적 알고리즘[6]을 이용하여 객체의 속도 정보를 함께 활용하였다. Sort 추적 알고리즘은 선형 운동을 하는 물체를 가정하여 과거의 정보와 새로운 측정값을 통해 객체의 위치를 학습하고 예측하는 칼만 필터[7]를 기반으로 객체의 속도 정보를 모델링한다. 그리고 다중 객체 추적을 위해 헝가리안 알고리즘을 사용하여 객체들의 예측된 위치 정보와 실제 관측된 위치 정보 사이에 최선의 매칭을 수행한다.

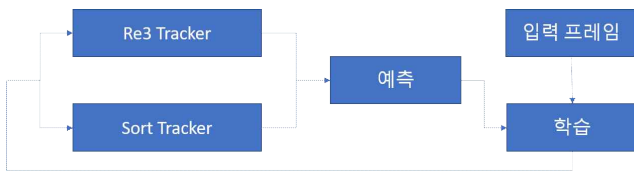


그림 1. 다중 객체 추적 알고리즘 흐름도

그림 1은 본 논문에서 제안하는 다중 객체 추적 알고리즘의 흐름도이다. 비디오 영상으로부터 각 영상 프레임이 입력으로 들어오고 객체의 형태 정보를 모델링하는 Re3 추적기와 속도 정보를 모델링하는 Sort 추적기가 독립적으로 수행되어 찾고자 하는 객체를 학습한다. 그렇게 학습된 정보를 바탕으로 현재 프레임에서 객체들의 위치를 예측하고 실측값과 비교하여 다음 프레임에서의 예측을 위해 각각 만들어진 객체들의 형태와 속도 정보 모델을 다시 학습한다.

### 3. 실험

본 논문에서는 과학기술정보통신부에서 주관한 2019 인공지능 그랜드 챌린지 트랙 1번 데이터셋을 이용해 실험하였다. 2019 인공지능 그랜드 챌린지는 지진, 화재와 같은 대형 복합 재난 상황에 신속한 인명 구조를 위한 인공지능 기술 경진 대회로, 이 중 트랙 1번은 드론 영상으로부터 다중 객체 추적을 수행하는 문제이다. 그림 2와 같이 드론 영상은 동적 카메라 환경에서 촬영되어 객체의 크기가 수시로 변화하며 폐색

이 빈번하게 나타난다. 드론으로 촬영된 비디오 영상으로부터 검출된 객체들에 대해 본 논문에서 제시하는 알고리즘을 이용하여 다중 객체 추적을 수행하였다. 비교를 위해 다중 객체 추적 대회인 Multiple Object Tracking Challenge에 소개된 V-IOU tracker[8]를 사용해 평가하였다. V-IOU tracker는 객체의 검출된 위치를 기반으로 추적하는 IOU tracker에 KCF[9]와 같이 시각 정보를 활용하는 시각 추적기를 조합한 다중 객체 추적 알고리즘이다.

그림 3과 그림 4는 비디오 영상에 대해 V-IOU tracker와 본 논문에서 제시한 추적 알고리즘의 결과를 비교한 것이다. 해당 영상에서 중심부에 이동하는 사람이 신호등에 의해 잠시 가려져 폐색이 발생한다. 시각 정보만을 활용하는 V-IOU tracker의 경우 폐색이 발생하면 추적 대상의 정보를 잃어버리게 되어 객체의 위치를 예측하지 못한다. 따라서 다음 프레임에서 동일한 물체가 나타나더라도 새로운 객체로 인식해 추적 성능이 나빠지게 된다. 하지만 본 논문에서 제시하는 추적 알고리즘을 이용하면 이와 같은 폐색 상황을 처리할 수 있다. 그림 4와 같이 Sort 추적기를 활용하여 이전의 프레임까지 학습된 객체의 속도와 위치 정보를 기반으로 예측하기 때문에 신호등에 객체가 가려지더라도 잃어버리지 않고 객체를 찾을 수 있다. 또한 Re3 추적기를 활용하여 객체가 다시 나타날 경우 이전 프레임들로부터 학습했던 객체의 형태 정보를 바탕으로 예측이 가능하기 때문에 동일한 물체임을 판단할 수 있다.

그림 5와 그림 6은 전경이 아닌 비슷한 모습과 색상을 가진 물체 사이에 폐색이 발생하는 비디오 영상으로 두 알고리즘의 결과를 분석한 것이다. V-IOU tracker의 경우 비슷한 모습의 물체가 가까이 있어 폐색



그림 2 샘플 데이터셋

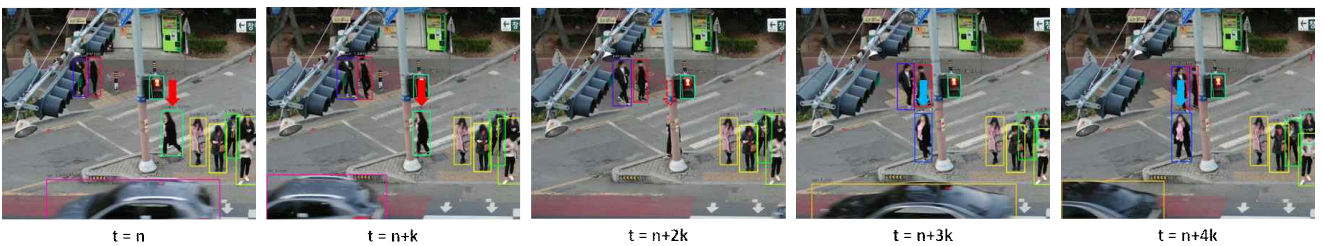


그림 3 V-IOU tracker

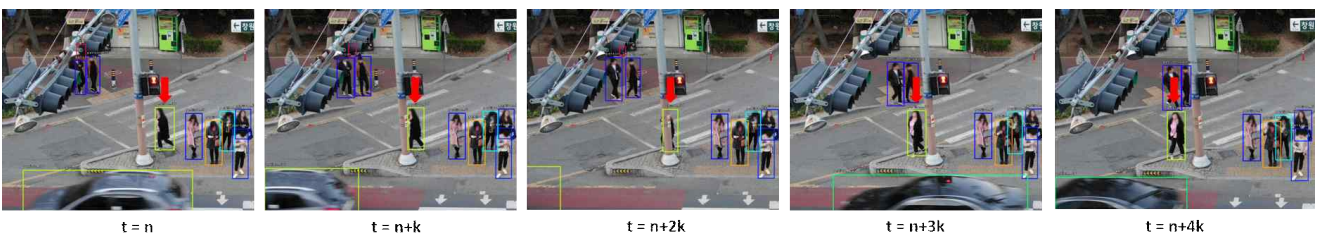


그림 4. Ours

이 발생할 경우 두 객체를 구별하지 못하는 것을 확인할 수 있다. 또 같은 색상의 물체가 가까이 지나가면 서로 추적하던 객체가 바뀌는 스위칭 오류가 발생하는 것도 확인할 수 있다. 이는 V-IOU tracker로 학습되는 시각 정보가 물체의 모습이 변화하는 것을 잘 고려하지 못한다는 것을 의미한다. 본 논문에서는 LSTM을 활용하여 객체의 모습이 변하더라도 이에 강인한 형태 모델을 학습하기 때문에 그림 6과 같이 구별을 잘 하는 것을 볼 수 있다. 또한 칼만 필터를 이용한 속도 정보를 조합하기 때문에 모습이 비슷한 물체가 서로 마주치거나 지나가더라도 기존에 추적하던 객체를 잘 따라가는 것을 확인할 수 있다.

#### 4. 결론

본 논문에서는 폐색에 강인한 다중 객체 추적 알고리즘을 제안한다. 찾고자 하는 객체가 전경에 의해 가려지거나 비슷한 모습의 다른 객체에 의해 가려지는 경우 해당 객체의 위치를 잘 예측하는 것이 중요하다. 이를 위해 영상의 이전 프레임들로부터 객체의 형태 정보를 학습한다. LSTM 구조를 활용하여 객체의 모습이 변형되는 것에 강인한 형태 정보를 모델링할 수 있다. 이를 활용할 경우 객체가 전경에 가려지더라도 이전 프레임들로부터 학습된 객체의 형태 정보를 바탕으로 예측이 가능하다. 그리고 프레임들로부터 검출된 객체의 위치 정보를 바탕으로 속도 정보를 학습한다. 칼만 필터를 활용하여 추적하고자 하는 객체들의 속도 정보를 모델링해, 유사한 모습의 객체가 가까이 지나가더라도 더 정확한 추적을 수행할 수 있다.

제안하는 다중 객체 추적 알고리즘을 이용하여 과학기술정보통신부에서 주관한 2019 인공지능 그랜드 챌린지의 다중 객체 추적 대회에서 2등을 수상하였다. 본 대회의 문제 환경은 비디오 영상 500개에 대하여 사람과 특정 사물들(자동차, 오토바이, 자전거, 소화기, 소화전)의 개수를 제한 시간 안에 도출하는 것이었다. 평가 기준은 실제 정답 수와 예측한 객체 수와의 유클리디안 거리 값이고 사람의 경우 가중치 5배를 두고 계산하였다.

#### 감사의 글

이 논문은 2020년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No.IITP20190017620021001, 드론 영상에서의 다중 객체 추적을 위한 딥러닝 기반 영상 처리 기술 개발)

#### 참고 문헌

- [1] Lee, Jae-Jun, and Jee-Hwan Ryu. "LIDAR based Multi-object Tracking Algorithm." Proceedings of the Korea Information Processing Society Conference. Korea Information Processing Society, 2015.
- [2] Shim, Young-Bin, and Hwa-Jin Park. "A Study on the Surveillance System of Multiple Object's Dangerous Behaviors." Journal of Digital Contents Society 14.4 (2013): 455-462.
- [3] Bilinski, Piotr, François Bremond, and Mohamed Becha Kaaniche. "Multiple object tracking with occlusions using HOG descriptors and multi resolution images." (2009): 36-36.

- [4] Comaniciu, Dorin, Visvanathan Ramesh, and Peter Meer. "Real-time tracking of non-rigid objects using mean shift." Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No. PR00662). Vol. 2. IEEE, 2000.

- [5] Farhadi, Daniel Gordon Ali, and Dieter Fox. "Re 3: Real-Time Recurrent Regression Networks for Visual Tracking of Generic Objects." IEEE Robot. Autom. Lett. 3.2 (2018): 788-795.

- [6] Bewley, Alex, et al. "Simple online and realtime tracking." 2016 IEEE International Conference on Image Processing (ICIP). IEEE, 2016.

- [7] Welch, Greg, and Gary Bishop. "An introduction to the Kalman filter." (1995): 41-95.

- [8] Bochinski, Erik, Tobias Senst, and Thomas Sikora. "Extending IOU based multi-object tracking by visual information." 2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS). IEEE, 2018.

- [9] Henriques, João F., et al. "High-speed tracking with kernelized correlation filters." IEEE transactions on pattern analysis and machine intelligence 37.3 (2014): 583-596.

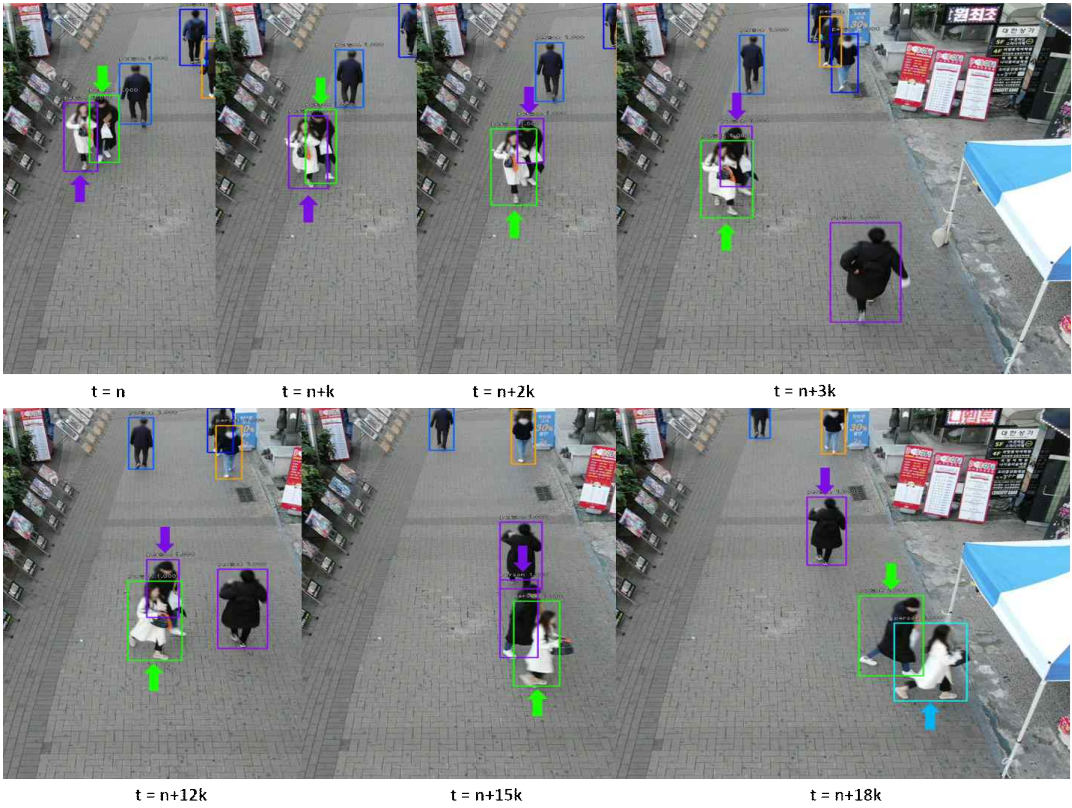


그림 5. V-IOU tracker

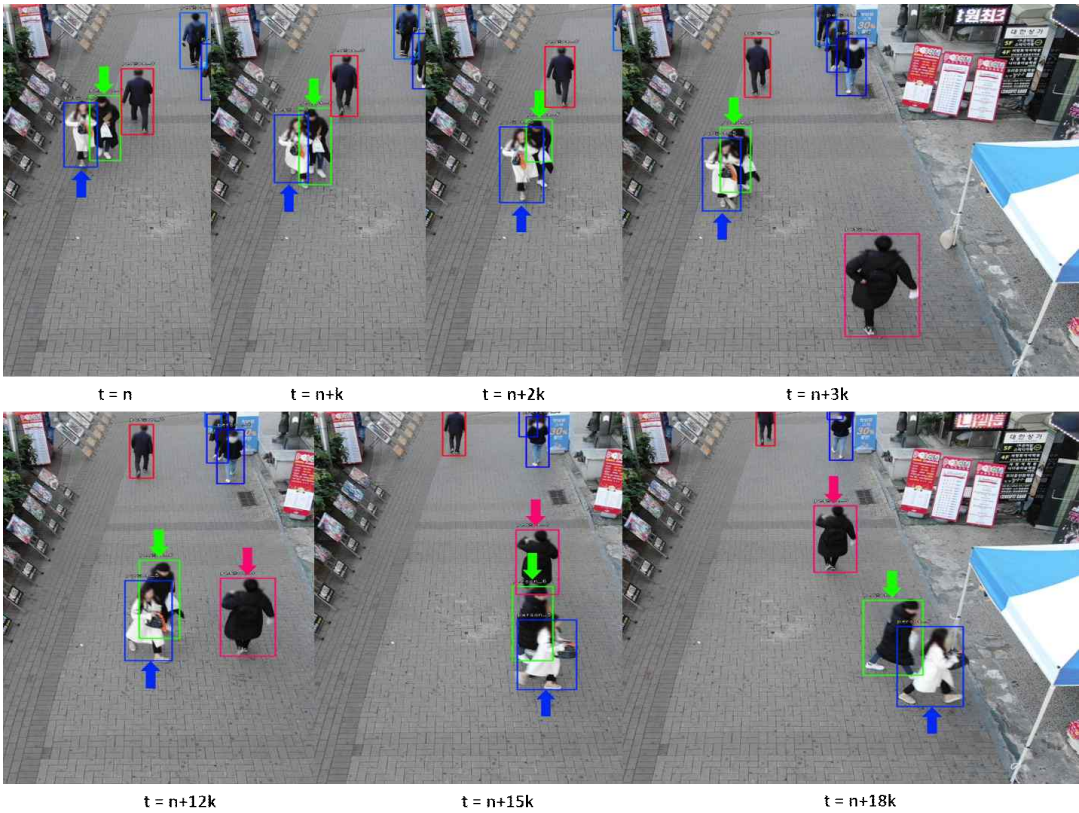


그림 6. Ours