



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0068254
(43) 공개일자 2024년05월17일

(51) 국제특허분류(Int. Cl.)

G06T 7/246 (2017.01) G06T 7/11 (2017.01)
G06V 10/40 (2022.01) G06V 10/74 (2022.01)

(52) CPC특허분류

G06T 7/246 (2017.01)
G06T 7/11 (2017.01)

(21) 출원번호 10-2022-0149454

(22) 출원일자 2022년11월10일

심사청구일자 2022년11월10일

(71) 출원인

인하대학교 산학협력단

인천광역시 미추홀구 인하로 100(용현동, 인하대학교)

(72) 발명자

배승환

인천광역시 미추홀구 인하로 100(용현동)

이성호

인천광역시 미추홀구 경인로 67, (송의동, The 미추홀퍼스트)

박대현

인천광역시 미추홀구 용정공원로83번길 49, 1266호(용현동)

(74) 대리인

양성보

전체 청구항 수 : 총 8 항

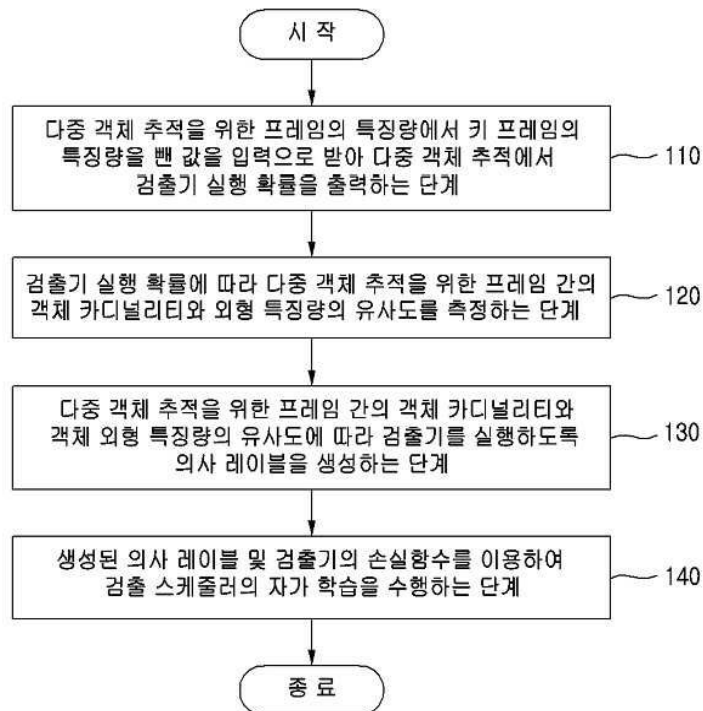
(54) 발명의 명칭 영상 컨텍스트 정보를 기반한 효율적인 다중 객체 추적 동작 메커니즘 학습 방법 및 시스템

(57) 요약

제안하는 영상 컨텍스트 정보를 기반한 효율적인 다중 객체 추적 동작 메커니즘 학습 방법 및 시스템이 제시된다. 본 발명에서 제안하는 영상 컨텍스트 정보를 기반한 효율적인 다중 객체 추적 동작 메커니즘 학습 방법은 다중 객체 추적을 위한 프레임의 특징량에서 키 프레임의 특징량을 뺀 값을 입력으로 받아 다중 객체 추적에서 검출기 실행 확률을 출력하는 단계

(뒷면에 계속)

대표도 - 도1



에서의 검출기 실행 확률을 출력하는 단계, 상기 검출기 실행 확률에 따라 유사도 측정부가 다중 객체 추적을 위한 프레임 간의 객체 카디널리티와 객체 외형 특징량의 유사도를 측정하는 단계, 의사 레이블 생성부가 상기 다중 객체 추적을 위한 프레임 간의 객체 카디널리티와 객체 외형 특징량의 유사도에 따라 검출기를 실행하도록 의사 레이블을 생성하는 단계 및 검출 스케줄러가 상기 생성된 의사 레이블 및 상기 검출기의 손실함수를 이용하여 검출 기반 추적 방법 또는 검출 없는 추적 방법 중 선택하여 검출 스케줄러의 자가 학습을 수행하는 단계를 포함한다.

(52) CPC특허분류

G06V 10/40 (2023.08)

G06V 10/761 (2023.08)

G06T 2207/20081 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711154594
과제번호	2022R1C1C1009208
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	우수신진연구
연구과제명	[Ezbaro] 대규모 확장 가능한 On-Vision AI를 위한 E3-AI 네트워크와 협업 학습 연

구

기 여 율	70/100
과제수행기관명	인하대학교 산학협력단
연구기간	2022.03.01 ~ 2023.02.28

이 발명을 지원한 국가연구개발사업

과제고유번호	1711160586
과제번호	2022-0-00448-001
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	방송통신산업기술개발사업
연구과제명	[Ezbaro] 인간처럼 회상이 가능한 인공 신경망 지속학습 플랫폼 개발(1차년도)

기 여 율	10/100
과제수행기관명	인하대학교 산학협력단
연구기간	2022.04.01 ~ 2022.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호	1711152759
과제번호	2020-0-01389-003
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	인공지능융합연구센터지원사업(국고)
연구과제명	[Ezbaro][지자체] 인공지능융합연구센터지원(3차년도)

기 여 율	10/100
과제수행기관명	인하대학교 산학협력단
연구기간	2022.01.01 ~ 2022.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호	1711174118
과제번호	RS-2022-00155915
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	정보통신인재양성(교육운영수익)
연구과제명	[Ezbaro][정부+민간] 인공지능융합혁신인재양성(1차년도)

기 여 율	10/100
과제수행기관명	인하대학교 산학협력단
연구기간	2022.07.01 ~ 2022.12.31

명세서

청구범위

청구항 1

다중 객체 추적을 위한 프레임의 특징량에서 키 프레임의 특징량을 뺀 값을 입력으로 받아 다중 객체 추적에서의 검출기 실행 확률을 출력하는 단계;

상기 검출기 실행 확률에 따라 유사도 측정부가 다중 객체 추적을 위한 프레임 간의 객체 카디널리티와 객체 외형 특징량의 유사도를 측정하는 단계;

의사 레이블 생성부가 상기 다중 객체 추적을 위한 프레임 간의 객체 카디널리티와 객체 외형 특징량의 유사도에 따라 검출기를 실행하도록 의사 레이블을 생성하는 단계; 및

검출 스케줄러가 상기 생성된 의사 레이블 및 상기 검출기의 손실함수를 이용하여 검출 기반 추적 방법 또는 검출 없는 추적 방법 중 선택하여 검출 스케줄러의 자가 학습을 수행하는 단계

를 포함하는 다중 객체 추적 동작 메커니즘 학습 방법.

청구항 2

제1항에 있어서,

상기 검출기 실행 확률에 따라 유사도 측정부가 다중 객체 추적을 위한 프레임 간의 객체 카디널리티와 객체 외형 특징량의 유사도를 측정하는 단계는,

다중 객체 추적을 위한 해당 프레임에서 검출기 실행을 스킵할 경우의 추적 트랙과 검출기를 실행할 경우의 추적 트랙 간의 객체 카디널리티를 비교하여 상기 객체 카디널리티의 유사도를 측정하는

다중 객체 추적 동작 메커니즘 학습 방법.

청구항 3

제1항에 있어서,

상기 검출기 실행 확률에 따라 유사도 측정부가 다중 객체 추적을 위한 프레임 간의 객체 카디널리티와 객체 외형 특징량의 유사도를 측정하는 단계는,

상기 검출기 실행을 스킵할 경우의 추적 트랙의 바운딩 박스와 상기 검출기를 실행할 경우의 추적 트랙의 바운딩 박스로 구성된 이중 그래프를 정의하고,

상기 검출기 실행을 스킵할 경우의 추적 트랙에 대한 각 노드는 상기 검출기를 실행할 경우의 추적 트랙에 대한 모든 노드와 연결되고,

상기 연결된 노드 간의 IoU(Intersection of Union) 점수를 사용하여 엣지 가중치(edge weight)를 평가하고, 평균 IoU 점수를 계산함으로써 상기 두 바운딩 박스 쌍 간의 객체 위치정보를 이용하여 객체 외형 특징량의 유사도를 측정하는

다중 객체 추적 동작 메커니즘 학습 방법.

청구항 4

제1항에 있어서,

상기 의사 레이블 생성부가 상기 다중 객체 추적을 위한 프레임 간의 객체 카디널리티와 객체 외형 특징량의 유사도에 따라 검출기를 실행하도록 의사 레이블을 생성하는 단계는,

객체 카디널리티와 객체 외형 특징량의 유사도의 곱을 미리 정해진 의사 레이블 임계값과 비교하여, 상기 비교 결과에 따라 검출 없는 추적 또는 검출 기반 추적을 실행하도록 유도하는 의사 레이블을 생성하고, 상기 생성된

의사 레이블을 통해 검출 스케줄러를 학습하기 위한 손실함수를 계산하는
다중 객체 추적 동작 메커니즘 학습 방법.

청구항 5

다중 객체 추적을 위한 프레임의 특징량에서 k 프레임의 특징량을 뺀 값을 입력으로 받아 다중 객체 추적에서의 검출기 실행 확률을 출력하는 검출기;

상기 검출기 실행 확률에 따라 다중 객체 추적을 위한 프레임 간의 객체 카디널리티와 객체 외형 특징량의 유사도를 측정하는 유사도 측정부;

상기 다중 객체 추적을 위한 프레임 간의 객체 카디널리티와 객체 외형 특징량의 유사도에 따라 검출기를 실행하도록 의사 레이블을 생성하는 의사 레이블 생성부; 및

상기 생성된 의사 레이블 및 상기 검출기의 손실함수를 이용하여 검출 기반 추적 방법 또는 검출 없는 추적 방법 중 선택하여 자가 학습을 수행하는 검출 스케줄러

를 포함하는 다중 객체 추적 동작 메커니즘 학습 시스템.

청구항 6

제5항에 있어서,

상기 유사도 측정부는,

다중 객체 추적을 위한 해당 프레임에서 검출기 실행을 스킵할 경우의 추적 트랙과 검출기를 실행할 경우의 추적 트랙 간의 객체 카디널리티를 비교하여 상기 객체 카디널리티의 유사도를 측정하는

다중 객체 추적 동작 메커니즘 학습 시스템.

청구항 7

제5항에 있어서,

상기 유사도 측정부는,

상기 검출기 실행을 스킵할 경우의 추적 트랙의 바운딩 박스와 상기 검출기를 실행할 경우의 추적 트랙의 바운딩 박스로 구성된 이중 그래프를 정의하고,

상기 검출기 실행을 스킵할 경우의 추적 트랙에 대한 각 노드는 상기 검출기를 실행할 경우의 추적 트랙에 대한 모든 노드와 연결되고,

상기 연결된 노드 간의 IoU(Intersection of Union) 점수를 사용하여 엣지 가중치(edge weight)를 평가하고, 평균 IoU 점수를 계산함으로써 상기 두 바운딩 박스 쌍 간의 객체 위치정보를 이용하여 객체 외형 특징량의 유사도를 측정하는

다중 객체 추적 동작 메커니즘 학습 시스템.

청구항 8

제5항에 있어서,

상기 의사 레이블 생성부는,

객체 카디널리티와 객체 외형 특징량의 유사도의 곱을 미리 정해진 의사 레이블 임계값과 비교하여, 상기 비교 결과에 따라 검출 없는 추적 또는 검출 기반 추적을 실행하도록 유도하는 의사 레이블을 생성하고, 상기 생성된 의사 레이블을 통해 검출 스케줄러를 학습하기 위한 손실함수를 계산하는

다중 객체 추적 동작 메커니즘 학습 시스템.

발명의 설명

기술 분야

[0001] 본 발명은 영상 컨텍스트 정보를 기반한 효율적인 다중 객체 추적 동작 메커니즘 학습 방법 및 시스템에 관한 것이다.

배경 기술

[0002] 검출 기반 추적(Tracking-by-detection)은 먼저 검출기를 사용하여 주어진 이미지에서 객체 위치를 결정한 다음, 연속 프레임 간에 검출 결과를 연결하여 식별 가능한 ID를 갖는 트랙을 생성하는 추적 방법이다. 검출 기반 추적 방법은 SDE(Seperated Detection and Embedding)와 JDE (Joint Detection and Embedding)로 구분된다. 두 방법의 주요 차이점은 검출 작업과 연관 작업이 동일 네트워크에서 처리하는지에 대한 여부로 설명할 수 있다. SDE는 검출 및 연관 작업이 독립적으로 학습된 서로 다른 네트워크를 사용한다. 한편, JDE는 검출 및 연관 헤더를 백본 네트워크에 연결하여 동시에 학습한다. 따라서 JDE는 저수준(Low-level)의 특징량을 공유하여 추적 속도를 향상한다. 그러나 두 방법 모두 전체 프레임에 대해 검출기를 적용하고 있어 검출 단계는 여전히 전반적인 MOT 연산 비용에 크게 영향을 미친다.

[0003] MOT 전체 속도를 향상시키기 위해 검출기 실행을 스킵 하는 방법이 고려될 수 있다. 이때 고정 프레임 간격을 갖는 검출기 실행 스킵 방법은 그 단순성 때문에 일반적으로 사용된다. 하지만 객체 카디널리티 오류 및 객체 위치정보 오류로 인한 추적 객체 손실을 고려하지 않고 검출을 결정하는 키 프레임을 선택하기 때문에 MOTA를 저하시킬 가능성이 있다. 이를 해결하기 위해 일부 적응형 샘플링 방법은 객체 영역(object region)과 영역 특징(region feature)을 비교하여 키 프레임을 선택한다. 예를 들어, [1]은 객체 크기 및 연속 프레임 간의 모션 차이를 고려하여 키 프레임을 결정한다. [2, 3]은 이 문제를 강화 학습 문제로 규정하고, 객체 위치정보 유사도를 바탕으로 키 프레임을 결정한다. 하지만 이런 방법들은 프레임 간의 객체 카디널리티 변동을 고려하지 않는다.

[0004] 자가 학습[4, 5]은 레이블이 없는 데이터로 모델을 학습하기 위한 의사 레이블을 생성한다. 자가 학습을 사용하여 MOT 모델의 일반화 능력을 향상하기 위한 몇 가지 연구가 있다. 종래기술에서는 사람 검출 데이터 셋을 사용하여 재식별 (Re-ID)에 대한 일반화 방법을 개선한다. 또 다른 종래기술에서는 공간 상관 학습(Spatial-Correlation Learning)의 콘스트레인(Constraint)을 적용하기 위해 자가 학습 손실 함수를 활용한다. 또 다른 종래기술은 자가 학습을 사용하지 않고 Re-ID 모델을 학습시키기 위해 SORT를 사용하여 의사 레이블(pseudo label)을 생성 한다.

[0005] 다중 객체 추적(Multi-Object Tracking)의 주요 패러다임인 검출 기반 추적(Tracking-by-detection)은 시간적 흐름에 따라 검출 객체 간의 지역/전역적 연관(Association)을 수행해 트랙을 생성한다. 따라서 고성능 검출기를 사용할 경우 객체 위치에 의한 불확실성을 감소하게 하여 다중 객체 추적 정확도(Multi-Object Tracking Accuracy; MOTA)를 높일 수 있다. 하지만 검출 기반 추적에서 검출기를 실행하기 위한 연산 비용이 연관을 위한 연산 비용보다 훨씬 높기 때문에 전체 추적 속도가 하락하는 문제가 발생한다.

선행기술문헌

비특허문헌

- [0006] (비특허문헌 0001) [1] Chen, Kai, et al., "Optimizing video object detection via a scale-time lattice.", Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 7814-7823, Jun. 2018, DOI: <https://doi.org/10.1109/CVPR.2018.00815>
- (비특허문헌 0002) [2] Luo, Hao, et al., "Detect or track: Towards cost-effective video object detection/tracking.", Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 33. No. 01, pp. 880 3-8810, Jan. 2019, DOI: <https://doi.org/10.1609/aaai.v33i01.3 3018803>
- (비특허문헌 0003) [3] Yao, Chun-Han, et al., "Video object detection via object-level temporal aggregation.", European conference on computer vision, pp. 160-177, Nov. 2020, DOI: https://doi.org/10.1007 /978-3-030-58568-6_10
- (비특허문헌 0004) [4] Chen, Ting, et al., "Self-supervised gans via auxiliary rotation loss.", Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 12154-12163, Jun. 2019, DOI: 10.1109/CVPR.2019.01243
- (비특허문헌 0005) [5] Noroozi, Mehdi, et al., "Boosting self-supervised learning via knowledge

transfer.", Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 9359-9367, Jun. 2018, DOI: <https://doi.org/10.1109/CVPR.2018.00975>

발명의 내용

해결하려는 과제

[0007] 본 발명이 이루고자 하는 기술적 과제는 실시간 및 고성능 추적을 위한 최적의 검출 기반 추적 메커니즘을 결정할 수 있는 검출 스케줄러 제안하고, 자가 학습 방법을 통해 검출 스케줄러 학습을 위한 검출기 실행 여부를 나타내는 의사 레이블(Pseudo Label) 및 객체 위치정보(Localization) 차이를 통해 새로운 검출 스케줄러 손실 함수를 제공하는데 있다.

과제의 해결 수단

[0008] 일 측면에 있어서, 본 발명에서 제안하는 영상 컨텍스트 정보를 기반한 효율적인 다중 객체 추적 동작 메커니즘 학습 방법은 다중 객체 추적을 위한 프레임의 특징량에서 키 프레임의 특징량을 뺀 값을 입력으로 받아 다중 객체 추적에서의 검출기 실행 확률을 출력하는 단계, 상기 검출기 실행 확률에 따라 유사도 측정부가 다중 객체 추적을 위한 프레임 간의 객체 카디널리티와 객체 외형 특징량의 유사도를 측정하는 단계, 의사 레이블 생성부가 상기 다중 객체 추적을 위한 프레임 간의 객체 카디널리티와 객체 외형 특징량의 유사도에 따라 검출기를 실행하도록 의사 레이블을 생성하는 단계 및 검출 스케줄러가 상기 생성된 의사 레이블 및 상기 검출기의 손실함수를 이용하여 검출 기반 추적 방법 또는 검출 없는 추적 방법 중 선택하여 검출 스케줄러의 자가 학습을 수행하는 단계를 포함한다.

[0009] 상기 검출기 실행 확률에 따라 유사도 측정부가 다중 객체 추적을 위한 프레임 간의 객체 카디널리티와 객체 외형 특징량의 유사도를 측정하는 단계는 다중 객체 추적을 위한 해당 프레임에서 검출기 실행을 스킵할 경우의 추적 트랙과 검출기를 실행할 경우의 추적 트랙 간의 객체 카디널리티를 비교하여 상기 객체 카디널리티의 유사도를 측정한다.

[0010] 상기 검출기 실행 확률에 따라 유사도 측정부가 다중 객체 추적을 위한 프레임 간의 객체 카디널리티와 객체 외형 특징량의 유사도를 측정하는 단계는 상기 검출기 실행을 스킵할 경우의 추적 트랙의 바운딩 박스와 상기 검출기를 실행할 경우의 추적 트랙의 바운딩 박스로 구성된 이중 그래프를 정의하고, 상기 검출기 실행을 스킵할 경우의 추적 트랙에 대한 각 노드는 상기 검출기를 실행할 경우의 추적 트랙에 대한 모든 노드와 연결되고, 상기 연결된 노드 간의 IoU(Intersection of Union) 점수를 사용하여 엣지 가중치(edge weight)를 평가하고, 평균 IoU 점수를 계산함으로써 상기 두 바운딩 박스 쌍 간의 객체 위치정보를 이용하여 객체 외형 특징량의 유사도를 측정한다.

[0011] 상기 의사 레이블 생성부가 상기 다중 객체 추적을 위한 프레임 간의 객체 카디널리티와 객체 외형 특징량의 유사도에 따라 검출기를 실행하도록 의사 레이블을 생성하는 단계는 객체 카디널리티와 객체 외형 특징량의 유사도의 값을 미리 정해진 의사 레이블 임계값과 비교하여, 상기 비교 결과에 따라 검출 없는 추적 또는 검출 기반 추적을 실행하도록 유도하는 의사 레이블을 생성하고, 상기 생성된 의사 레이블을 통해 검출 스케줄러를 학습하기 위한 손실함수를 계산한다.

[0012] 또 다른 일 측면에 있어서, 본 발명에서 제안하는 영상 컨텍스트 정보를 기반한 효율적인 다중 객체 추적 동작 메커니즘 학습 시스템은 다중 객체 추적을 위한 프레임의 특징량에서 키 프레임의 특징량을 뺀 값을 입력으로 받아 다중 객체 추적에서의 검출기 실행 확률을 출력하는 검출기, 상기 검출기 실행 확률에 따라 다중 객체 추적을 위한 프레임 간의 객체 카디널리티와 객체 외형 특징량의 유사도를 측정하는 유사도 측정부, 상기 다중 객체 추적을 위한 프레임 간의 객체 카디널리티와 객체 외형 특징량의 유사도에 따라 검출기를 실행하도록 의사 레이블을 생성하는 의사 레이블 생성부 및 상기 생성된 의사 레이블 및 상기 검출기의 손실함수를 이용하여 검출 기반 추적 방법 또는 검출 없는 추적 방법 중 선택하여 자가 학습을 수행하는 검출 스케줄러를 포함한다.

발명의 효과

[0013] 본 발명의 실시예들에 따른 검출 스케줄러의 자가 학습을 통해 현재 프레임과 키 프레임 간의 컨볼루션 특징량의 차이를 측정하고, 온라인 다중 객체 추적이 실행되는 동안, MOT에 대한 최적의 검출 기반 추적 메커니즘을

결정할 수 있다. 또한, 검출기 실행 여부를 나타내는 의사 레이블(Pseudo Label)과 검출기 실행 유무에 따른 다중 객체 추적 결과 기반 객체 카디널리티(Cardinality) 및 객체 위치정보(Localization) 차이를 통해 새로운 검출 스케줄러 손실 함수를 통해 검출 스케줄러의 자가 학습을 수행할 수 있다.

도면의 간단한 설명

- [0014] 도 1은 본 발명의 일 실시예에 따른 영상 컨텍스트 정보를 기반한 효율적인 다중 객체 추적 동작 메커니즘 학습 방법을 설명하기 위한 흐름도이다.
- 도 2는 본 발명의 일 실시예에 따른 영상 컨텍스트 정보를 기반한 효율적인 다중 객체 추적 동작 메커니즘 학습 시스템의 구성을 나타내는 도면이다.
- 도 3은 본 발명의 일 실시예에 따른 검출 스케줄링 네트워크를 나타내는 도면이다.
- 도 4는 본 발명의 일 실시예에 따른 데이터셋에 대한 정성적 결과를 나타내는 도면이다.

발명을 실시하기 위한 구체적인 내용

- [0015] 본 발명에서는 높은 연산 비용이 필요한 검출기를 매 프레임마다 실행할 필요 없이 추적을 진행할 수 있는 방법을 제시한다. 본 발명은 인접 프레임 간의 유사한 컨텍스트(context)를 가진 특정 프레임에서 객체 카디널리티(cardinality)와 객체 외형(Appearance)의 변화가 적다는 점을 이용한다. 이를 통해 검출 없이 모션 모델로 추적을 수행할 수 있는 프레임을 결정하여, 검출 기반 연관 없이 객체를 추적할 수 있는 모션을 학습할 수 있다. 결과적으로 검출기 실행 없이 모션 모델로 추적을 수행함으로써 전체 추적 속도를 향상할 수 있다. 이를 위해 본 발명에서는 일정 간격(즉, 프레임 도메인 내 균일 샘플링)마다 검출기 실행 스킵(skip)을 적용할 수 있다. 그러나 이 간격은 동일한 시퀀스 내에서도 다를 수 있기 때문에 최적 간격을 찾는 것은 어려운 일이다. 예를 들어, 객체의 모션(Motion) 변화가 증가하거나 추적 품질이 낮아질 때 MOTA(Multi-Object Tracking Accuracy)를 유지하기 위해서는 검출기를 더 빈번히 실행함으로써 간격을 더 좁혀야 한다.
- [0016] 따라서, 본 발명에서는 실시간 다중 객체 추적을 수행하기 위해 감소를 최소화하며 전체 추적 속도를 향상할 수 있는 최적의 검출 시점(다시 말해, 키 프레임)을 찾는 것을 목표로 한다. 이를 위해 일부 최근 연구는 이전 프레임과 현재 프레임 간의 추적 결과의 바운딩 박스(bounding box) 쌍(pair)을 서로 비교하고, 바운딩 박스 간의 일치 정도가 낮을 경우 검출기를 실행한다. 그러나 트랙 초기화 및 종료는 검출 기반 연관 없이 수행될 수 없으므로 추적 품질뿐만 아니라 객체 카디널리티 또한 고려되어야 한다. 따라서, 추적 객체의 바운딩 박스 또는 외형 특징량만을 비교하는 것만으로는 두 변동 사항을 모두 측정할 수 없다.
- [0017] 이를 해결하기 위해 본 발명에서는 현재 프레임과 키 프레임 간의 컨볼루션 특징량의 차이를 측정하는 방법을 제안한다. 또한, 온라인 다중 객체 추적이 실행되는 동안, MOT에 대한 최적의 검출 기반 추적 메커니즘을 결정할 수 있는 검출 스케줄러를 제안한다. 그러나 검출 스케줄러에 대한 GT가 없기 때문에 지도 학습(Supervised learning)을 사용할 수 없다. 또한, 이에 대한 GT를 만드는 것은 추적기와 검출기의 성능에 의존하기 때문에 무의미할 수 있다. 따라서, 본 발명에서는 자가 학습을 통해 검출 스케줄러를 학습한다. 또한, 본 발명에서는 검출기 실행 여부를 나타내는 의사 레이블(Pseudo Label)을 제안한다. 검출기 실행 유무에 따른 다중 객체 추적 결과 기반 객체 카디널리티(Cardinality) 및 객체 위치정보(Localization) 차이를 통해 새로운 검출 스케줄러 손실 함수를 제안한다. 본 발명에서는 손실 함수를 최소화함으로써 MOTA를 최대한 유지하면서 MOT 전체 속도를 높이는 최적의 검출 기반 추적 메커니즘을 결정할 수 있다. 이 메커니즘은 검출 기반 추적 방법의 정확성에 따라 적응적으로 결정될 수 있다. 결과적으로 본 발명의 주요 기여는 다음과 같다.
- [0018] 첫째로 실시간 및 고성능 추적을 위한 최적의 검출 기반 추적 메커니즘을 결정할 수 있는 검출 스케줄러 제안하고, 둘째로 자가 학습 방법을 통해 검출 스케줄러 학습을 위한 의사 레이블 및 검출 스케줄러 손실 함수를 제안한다. 베이스 라인(Baseline)과 비교해 MOTA를 5.8% 줄이는 것만으로 속도를 47.0% 향상시켰다. 또한, 본 발명에서는 광범위한 접제 실험과 MOT 벤치마크 데이터 셋에서 최신 MOT 모델 간의 성능 비교를 제공한다. 제안하는 검출 스케줄러는 단일 NVIDIA Titan Xp를 사용하여 MOT16 테스트 셋에서 MOTA 74.5%, 20.6Hz 속도를 달성한다. 이하, 본 발명의 실시 예를 첨부된 도면을 참조하여 상세하게 설명한다.

- [0020] 도 1은 본 발명의 일 실시예에 따른 영상 컨텍스트 정보를 기반한 효율적인 다중 객체 추적 동작 메커니즘 학습 방법을 설명하기 위한 흐름도이다.

- [0021] 본 발명에서는 실시간 고성능 다중 객체 추적을 수행하기 위해 최적의 TBD(Tracking-by-Detection) 메커니즘을 결정할 수 있는 검출 스케줄러(Detector Scheduler)를 제안한다. 검출 스케줄러는 서로 다른 프레임 간의 특징량 차이를 측정하는 것으로 검출기 실행 여부를 결정하여 전체 추적 속도를 향상시킨다. 하지만, 검출 스케줄러의 학습에 필요한 GT(Ground Truth) 생성이 어렵기 때문에 본 발명에서는 검출 스케줄러를 추적 결과만을 통해 학습 가능한 자가 학습 방법을 제안한다. 제안하는 자가 학습 방법은 프레임 간의 객체 카디널리티와 객체 외형 특징량의 유사도가 커질 때 검출기를 실행할 수 있도록 의사 레이블을 생성하고 제안된 손실함수를 통해 검출 스케줄러를 학습한다.
- [0022] 제안하는 영상 컨텍스트 정보를 기반한 효율적인 다중 객체 추적 동작 메커니즘 학습 방법은 다중 객체 추적을 위한 프레임의 특징량에서 키 프레임의 특징량을 뺀 값을 입력으로 받아 다중 객체 추적에서의 검출기 실행 확률을 출력하는 단계(110), 상기 검출기 실행 확률에 따라 유사도 측정부가 다중 객체 추적을 위한 프레임 간의 객체 카디널리티와 객체 외형 특징량의 유사도를 측정하는 단계(120), 의사 레이블 생성부가 상기 다중 객체 추적을 위한 프레임 간의 객체 카디널리티와 객체 외형 특징량의 유사도에 따라 검출기를 실행하도록 의사 레이블을 생성하는 단계(130) 및 검출 스케줄러가 상기 생성된 의사 레이블 및 상기 검출기의 손실함수를 이용하여 검출 기반 추적 방법 또는 검출 없는 추적 방법 중 선택하여 검출 스케줄러의 자가 학습을 수행하는 단계(140)를 포함한다.
- [0023] 단계(110)에서, 다중 객체 추적을 위한 프레임의 특징량에서 키 프레임의 특징량을 뺀 값을 입력으로 받아 다중 객체 추적에서의 검출기 실행 확률을 출력한다.
- [0024] 단계(120)에서, 상기 검출기 실행 확률에 따라 유사도 측정부가 다중 객체 추적을 위한 프레임 간의 객체 카디널리티와 객체 외형 특징량의 유사도를 측정한다.
- [0025] 다중 객체 추적을 위한 해당 프레임에서 검출기 실행을 스킵할 경우의 추적 트랙과 검출기를 실행할 경우의 추적 트랙 간의 객체 카디널리티를 비교하여 상기 객체 카디널리티의 유사도를 측정한다.
- [0026] 상기 검출기 실행을 스킵할 경우의 추적 트랙의 바운딩 박스와 상기 검출기를 실행할 경우의 추적 트랙의 바운딩 박스로 구성된 이중 그래프를 정의한다.
- [0027] 상기 검출기 실행을 스킵할 경우의 추적 트랙에 대한 각 노드는 상기 검출기를 실행할 경우의 추적 트랙에 대한 모든 노드와 연결된다. 상기 연결된 노드 간의 IoU(Intersection of Union) 점수를 사용하여 엣지 가중치(edge weight)를 평가하고, 평균 IoU 점수를 계산함으로써 상기 두 바운딩 박스 쌍 간의 객체 위치정보를 이용하여 객체 외형 특징량의 유사도를 측정한다.
- [0028] 단계(130)에서, 상기 의사 레이블 생성부가 상기 다중 객체 추적을 위한 프레임 간의 객체 카디널리티와 객체 외형 특징량의 유사도에 따라 검출기를 실행하도록 의사 레이블을 생성한다.
- [0029] 객체 카디널리티와 객체 외형 특징량의 유사도의 곱을 미리 정해진 의사 레이블 임계값과 비교하여, 상기 비교 결과에 따라 검출 없는 추적 또는 검출 기반 추적을 실행하도록 유도하는 의사 레이블을 생성한다. 상기 생성된 의사 레이블을 통해 검출 스케줄러를 학습하기 위한 손실함수를 계산한다.
- [0030] 단계(140)에서, 검출 스케줄러가 상기 생성된 의사 레이블 및 상기 검출기의 손실함수를 이용하여 검출 기반 추적 방법 또는 검출 없는 추적 방법 중 선택하여 검출 스케줄러의 자가 학습을 수행한다.
- [0032] 도 2는 본 발명의 일 실시예에 따른 영상 컨텍스트 정보를 기반한 효율적인 다중 객체 추적 동작 메커니즘 학습 시스템의 구성을 나타내는 도면이다.
- [0033] 본 실시예에 따른 다중 객체 추적 동작 메커니즘 학습 시스템(200)은 프로세서(210), 버스(220), 네트워크 인터페이스(230), 메모리(240) 및 데이터베이스(250)를 포함할 수 있다. 메모리(240)는 운영체제(241) 및 다중 객체 추적 동작 메커니즘 학습 루틴(242)을 포함할 수 있다. 프로세서(210)는 검출부(211), 유사도 측정부(212), 의사 레이블 생성부(213) 및 검출 스케줄러(214)를 포함할 수 있다. 다른 실시예들에서 다중 객체 추적 동작 메커니즘 학습 시스템(200)은 도 2의 구성요소들보다 더 많은 구성요소들을 포함할 수도 있다. 그러나, 대부분의 종래기술적 구성요소들을 명확하게 도시할 필요성은 없다. 예를 들어, 다중 객체 추적 동작 메커니즘 학습 시스템(200)은 디스플레이나 트랜시버(transceiver)와 같은 다른 구성요소들을 포함할 수도 있다.
- [0034] 메모리(240)는 컴퓨터에서 판독 가능한 기록 매체로서, RAM(random access memory), ROM(read only memory) 및

디스크 드라이브와 같은 비소멸성 대용량 기록장치(permanent mass storage device)를 포함할 수 있다. 또한, 메모리(240)에는 운영체제(241)와 다중 객체 추적 동작 메커니즘 학습 루틴(242)을 위한 프로그램 코드가 저장될 수 있다. 이러한 소프트웨어 구성요소들은 드라이브 메커니즘(drive mechanism, 미도시)을 이용하여 메모리(240)와는 별도의 컴퓨터에서 판독 가능한 기록 매체로부터 로딩될 수 있다. 이러한 별도의 컴퓨터에서 판독 가능한 기록 매체는 플로피 드라이브, 디스크, 테이프, DVD/CD-ROM 드라이브, 메모리 카드 등의 컴퓨터에서 판독 가능한 기록 매체(미도시)를 포함할 수 있다. 다른 실시예에서 소프트웨어 구성요소들은 컴퓨터에서 판독 가능한 기록 매체가 아닌 네트워크 인터페이스(230)를 통해 메모리(240)에 로딩될 수도 있다.

- [0035] 버스(220)는 다중 객체 추적 동작 메커니즘 학습 시스템(200)의 구성요소들간의 통신 및 데이터 전송을 가능하게 할 수 있다. 버스(220)는 고속 시리얼 버스(high-speed serial bus), 병렬 버스(parallel bus), SAN(Storage Area Network) 및/또는 다른 적절한 통신 기술을 이용하여 구성될 수 있다.
- [0036] 네트워크 인터페이스(230)는 다중 객체 추적 동작 메커니즘 학습 시스템(200)을 컴퓨터 네트워크에 연결하기 위한 컴퓨터 하드웨어 구성요소일 수 있다. 네트워크 인터페이스(230)는 다중 객체 추적 동작 메커니즘 학습 시스템(200)을 무선 또는 유선 커넥션을 통해 컴퓨터 네트워크에 연결시킬 수 있다.
- [0037] 데이터베이스(250)는 영상 컨텍스트 정보를 기반한 효율적인 다중 객체 추적 동작 메커니즘 학습을 위해 필요한 모든 정보를 저장 및 유지하는 역할을 할 수 있다. 도 2에서는 다중 객체 추적 동작 메커니즘 학습 시스템(200)의 내부에 데이터베이스(250)를 구축하여 포함하는 것으로 도시하고 있으나, 이에 한정되는 것은 아니며 시스템 구현 방식이나 환경 등에 따라 생략될 수 있고 혹은 전체 또는 일부의 데이터베이스가 별개의 다른 시스템 상에 구축된 외부 데이터베이스로서 존재하는 것 또한 가능하다.
- [0038] 프로세서(210)는 기본적인 산술, 로직 및 다중 객체 추적 동작 메커니즘 학습 시스템(200)의 입출력 연산을 수행함으로써, 컴퓨터 프로그램의 명령을 처리하도록 구성될 수 있다. 명령은 메모리(240) 또는 네트워크 인터페이스(230)에 의해, 그리고 버스(220)를 통해 프로세서(210)로 제공될 수 있다. 프로세서(210)는 검출부(211), 유사도 측정부(212), 의사 레이블 생성부(213) 및 검출 스케줄러(214)를 위한 프로그램 코드를 실행하도록 구성될 수 있다. 이러한 프로그램 코드는 메모리(240)와 같은 기록 장치에 저장될 수 있다.
- [0039] 검출부(211), 유사도 측정부(212), 의사 레이블 생성부(213) 및 검출 스케줄러(214)는 도 1의 단계들(110~140)을 수행하기 위해 구성될 수 있다.
- [0040] 다중 객체 추적 동작 메커니즘 학습 시스템(200)은 검출부(211), 유사도 측정부(212), 의사 레이블 생성부(213) 및 검출 스케줄러(214)를 포함할 수 있다.
- [0041] 검출부(211)는 다중 객체 추적을 위한 프레임의 특징량에서 키 프레임의 특징량을 뺀 값을 입력으로 받아 다중 객체 추적에서의 검출기 실행 확률을 출력한다.
- [0042] 유사도 측정부(212)는 상기 검출기 실행 확률에 따라 다중 객체 추적을 위한 프레임 간의 객체 카디널리티와 객체 외형 특징량의 유사도를 측정한다.
- [0043] 유사도 측정부(212)는 다중 객체 추적을 위한 해당 프레임에서 검출기 실행을 스킵할 경우의 추적 트랙과 검출기를 실행할 경우의 추적 트랙 간의 객체 카디널리티를 비교하여 상기 객체 카디널리티의 유사도를 측정한다.
- [0044] 유사도 측정부(212)는 상기 검출기 실행을 스킵할 경우의 추적 트랙의 바운딩 박스와 상기 검출기를 실행할 경우의 추적 트랙의 바운딩 박스로 구성된 이중 그래프를 정의한다.
- [0045] 유사도 측정부(212)는 상기 검출기 실행을 스킵할 경우의 추적 트랙에 대한 각 노드는 상기 검출기를 실행할 경우의 추적 트랙에 대한 모든 노드와 연결된다. 상기 연결된 노드 간의 IoU(Intersection of Union) 점수를 사용하여 엣지 가중치(edge weight)를 평가하고, 평균 IoU 점수를 계산함으로써 상기 두 바운딩 박스 쌍 간의 객체 위치정보를 이용하여 객체 외형 특징량의 유사도를 측정한다.
- [0046] 의사 레이블 생성부(213)는 상기 다중 객체 추적을 위한 프레임 간의 객체 카디널리티와 객체 외형 특징량의 유사도에 따라 검출기를 실행하도록 의사 레이블을 생성한다.
- [0047] 의사 레이블 생성부(213)는 객체 카디널리티와 객체 외형 특징량의 유사도의 값을 미리 정해진 의사 레이블 임계값과 비교하여, 상기 비교 결과에 따라 검출 없는 추적 또는 검출 기반 추적을 실행하도록 유도하는 의사 레이블을 생성한다. 상기 생성된 의사 레이블을 통해 검출 스케줄러를 학습하기 위한 손실함수를 계산한다.
- [0048] 검출 스케줄러(214)는 상기 생성된 의사 레이블 및 상기 검출기의 손실함수를 이용하여 검출 기반 추적 방법 또

는 검출 없는 추적 방법 중 선택하여 검출 스케줄러의 자가 학습을 수행한다.

[0050] 도 3은 본 발명의 일 실시예에 따른 검출 스케줄링 네트워크를 나타내는 도면이다.

[0051] 도 3을 참조하여 본 발명의 실시예에 따른 MOT에 대한 검출 스케줄러의 검출 스케줄링 네트워크에 대해 기술한다. 검출 스케줄러의 각 구성 요소들인 검출 스케줄링 네트워크(Detector Scheduling Network), 의사 레이블 및 검출기 손실함수, 자가학습에 대해 더욱 상세히 설명한다.

[0052] 본 발명의 실시예에 따른 검출 스케줄러(Detector Scheduler)의 학습 목표는 다중 객체 추적에서 검출기 실행 여부를 결정하는 것이다. 이를 위해 현재 t 번째 프레임의 특징량 $\mathbf{F}_t$ 와 k_e 프레임의 특징량 $\mathbf{F}_{k_e}$ 을 뺀 특징량 $\mathbf{P}_t \in \mathbb{R}^{C \times H \times W}$ 을 통해, 두 특징량 간의 차이를 계산한다. 이 때 $\mathbf{P}_t$ 는 다음 과 같다.

$$\mathbf{P}_t = \mathbf{F}_t - \mathbf{F}_{k_e} \quad (1)$$

[0054] 검출 스케줄링 네트워크는 $\mathbf{P}_t$ 를 입력으로 하여 검출기 실행 확률(detector scheduling probability) $f(s_t)$ 을 출력한다. 이 때, $f(s_t)$ 가 낮은 경우, 프레임에서 검출기를 실행하도록 유도한다. 구체적으로 $\mathbf{P}_t$ 를 3×3 ConvBlocks에 입력한 다음, 1×1 ConvBlocks에 차례대로 입력한다. 여기서 각 Blocks은 컨볼루션, 배치 정규화 및 Leaky ReLU 활성화 함수($\alpha=0.01$) 레이어들을 포함한다. 3×3 및 컨볼루션 레이어의 출력 채널 크기는 각각 C 및 1이다. 연속 컨볼루션 연산 후 특징량을 $H \times W$ 차원을 갖는 특징 벡터로 평탄화(Flatten)한 다음 64, 64, 1개의 뉴런으로 3개의 Fully Connected Layers를 적용하여 s_t 를 출력한다. 마지막으로, 시그모이드 함수 $f(\cdot)$ 로 $f(s_t)$ 를 계산할 수 있다.

[0055] 본 발명의 실시예에 따른 검출 스케줄러는 다음과 같은 근거에 기반하여 자가 학습 방법으로 학습된다. 첫째로, MOT에 대한 검출 기반 추적 메커니즘 학습을 위한 사용 가능한 공개된 GT 또는 정의된 가이드라인이 없다. 둘째로, 추적기와 검출기 간의 성능 의존성이 강하기 때문에 GT를 생성하는 것이 어렵다. 정확도와 속도는 트레이드-오프(trade-off) 관계이기 때문에 특정 성능에 맞는 GT를 생성하는 것은 까다로운 일이다. 따라서 본 발명에서는 k_e 프레임과 현재 프레임의 추적 결과 간의 비유사성을 측정하여 검출기 실행 여부를 나타낼 수 있는 의사 레이블 생성 방법을 제안한다. 구체적으로 본 발명에서는 두 프레임 간의 객체 카디널리티 및 객체 위치 정보 오류를 측정한다. 결과적으로 본 발명에서 제안하는 자가 학습 방법은 추적 결과에 따라 적응적으로 의사 레이블을 생성할 수 있다.

[0056] N 프레임의 시퀀스가 주어질 때, 프레임 I_t 에서 검출기 D를 실행함으로써 검출 결과 $\mathbf{D}(I_t) = D_t = \{\mathbf{d}_t^i, y_t^i\}_{i=1}^{|D_t|}$ 를 생성한다. 여기서 d 는 객체 i 에 대한 바운딩 박스이고, $|D_t|$ 는 t 번째 프레임에서 탐지된 객체 수다. 본 발명에서는 현재 프레임인 t 번째 프레임에서 $T_t = \{\hat{\mathbf{d}}_t^j, \hat{y}_t^j\}_{j=1}^{|T_t|}$ 를 평가하기 위해 트랙 $T_t^{(TD)} = \mathbf{T}\{T_{t-1}, D_t\}$ 를 위한 온라인 검출 기반 추적 문제를 공식화한다. 이 때 T는 추적기, $\hat{\mathbf{d}}_t^j$ 와 $\hat{y}_t^j$ 는 각각 트랙 j 에 대한 추적 결과와 추적 식별 레이블(track identity label)이다. $|T_t|$ 는 t 번째 프레임에서의 트랙들의 개수이다.

[0057] 본 발명에서는 객체 모션을 예측하기 위해 칼만 필터를 사용한다. t 번째 프레임에서 검출기 실행을 스킵하면 추적 문제를 $T_t^{(TO)} = \mathbf{T}(T_{t-1})$ 로 공식화할 수 있다. 일반적으로, 이 문제는 프레임 $t-1$ 까지의 추적 결과 T_{t-1} 를 기반으로 한 예측값을 통해 해결될 수 있다. 대부분의 경우 검출을 수행하여 추적하는 $T_t^{(TD)}$ 의 추적

품질이 $T_t^{(TO)}$ 보다 높다. 하지만 $T_t^{(TD)}$ 에 대한 추적 연산 비용은 $T_t^{(TO)}$ 에 비해 높기 때문에 MOT 정확도를 최소한으로 줄이면서 $T_t^{(TO)}$ 의 수를 늘리는 것을 목표로 한다. 전체 프레임에서 최적의 $T_t^{(TO)}$ 를 찾기 위해 $T_t^{(TO)} \approx T_t^{(TD)}$ 일 경우, t 번째 프레임에서 검출 없이 추적이 가능하다는 가정을 세운다. 따라서 객체 카디널리티와 객체 위치정보 측면에서 $T_t^{(TO)}$ 의 추적 결과가 $T_t^{(TD)}$ 의 추적 결과와 유사하다면 본 가정은 성립될 수 있다. 따라서 최적의 $T_t^{(TO)}$ 를 찾는 문제는 전체 프레임에서 $T_t^{(TO)} \approx T_t^{(TD)}$ 를 찾는 문제로 변환할 수 있다. 이를 위해 $T_t^{(TO)} \approx T_t^{(TD)}$ 를 추정하기 위해 두 방법 간의 객체 카디널리티와 객체 위치정보 유사성을 정의한다.

[0058] 본 발명의 실시예에 따른 객체 카디널리티 유사성 측정 방법은 앞서 설명된 바와 같이, $T_t^{(TD)}$ 와 $T_t^{(TO)}$ 간의 객체 카디널리티가 추적 품질의 척도로 고려되어야 한다. 따라서, 본 발명에서는 $T_t^{(TD)}$ 와 $T_t^{(TO)}$ 사이의 객체 카디널리티를 비교하고, 객체 카디널리티 유사성 척도 S_{card} 를 다음과 같이 정의할 수 있다.

$$S_{card}(T_t^{(TO)}, T_t^{(TD)}) = \min(-\frac{1}{e} \cdot \ln(1 - R(T_t^{(TO)}, T_t^{(TD)})), 1) \quad (2)$$

[0060] $R(T_t^{(TO)}, T_t^{(TD)})$ 는 $T_t^{(TD)}$ 와 $T_t^{(TO)}$ 사이의 객체 카디널리티의 비율이다. 이를 $[0, 1]$ 사이의 값으로 정규화하기 위해, 다음과 같이 계산할 수 있다.

$$R(T_t^{(TO)}, T_t^{(TD)}) = \min(\frac{T_t^{(TD)}}{T_t^{(TO)}}, \frac{T_t^{(TO)}}{T_t^{(TD)}}) \quad (3)$$

[0062] 식(2)에서 $S_{card} = 1$ 인 경우, 두 추적 방법이 같은 개수의 트랙을 생성했음을 나타낸다. 즉, t 번째 프레임에서 트랙 초기화 및 종료의 일어난지 않음을 의미한다. 따라서 S_{card} 가 높을수록 $T_t^{(TO)}$ 의 추적 품질이 높으며, 아니라면 추적 품질이 낮다는 것을 의미한다.

[0063] 객체 위치정보 유사성 측정 방법은 $T_t^{(TD)}$ 와 $T_t^{(TO)}$ 사이의 바운딩 박스 $\hat{\mathbf{d}}_t$ 사이의 비유사성을 평가하기 위해, 본 발명에서는 객체 위치정보 유사성 척도 S_{loc} 를 정의한다. 두 방법 간의 바운딩 박스 쌍이 주어지면, IoU(Interest of Union)를 사용하여 두 바운딩 박스 쌍 간의 유사성을 평가할 수 있다. 따라서, 본 발명에서는 S_{loc} 을 다음과 같이 정의할 수 있다.

$$S_{loc}(T_t^{(TD)}, T_t^{(TO)}) = \frac{1}{T_t^*} \sum_{(i_k, j_k)_{k=1}^{|T_t^*|}} IOU(\hat{\mathbf{b}}_t^{(TD), i_k}, \hat{\mathbf{b}}_t^{(TO), j_k}) \quad (4)$$

[0065] 여기서 $|T_t^*|$ 는 $T_t^{(TD)}$ 와 $T_t^{(TO)}$ 간에 매칭된 트랙 쌍에 대한 카디널리티다. 그리고 (i_k, j_k) 간의 관련성을 결정하기 위해 본 발명에서는 $T_t^{(TD)}$ 와 $T_t^{(TO)}$ 의 바운딩 박스로 구성된 이중 그래프(bigraph)를 정의한다. 그리고 $T_t^{(TD)}$ 에 대한 각 노드는 $T_t^{(TO)}$ 에 대한 모든 노드와 연결된다. 이때 엣지 가중치(edge weight)는 연결

된 노드 간의 IoU 점수를 사용하여 평가한다. 이 그래프의 최대 가중치 일치 쌍은 헝가리안 알고리즘에 의해 최적으로 결정될 수 있다. 그런 다음 일치된 쌍의 평균 IoU 점수를 계산하여 S_{loc} 을 평가한다.

[0066] 본 발명의 실시예에 따른 검출 스케줄러를 학습하기 위해, 의사 GT 레이블 $G = \{G_t\}_{t=1}^N$ 을 생성한다. 학습 데이터 N개 프레임의 시퀀스가 주어지면 온라인 추적기 T와 검출기 D를 실행하여 각 프레임에서 의사 레이블 (pseudo label)을 생성한다. 이를 통해 매 프레임마다 $T_t^{(TD)}$ 와 $T_t^{(TO)}$ 를 생성할 수 있다. t번째 프레임의 각 레이블 G_t 는 이진 레이블이며, 검출 없는 추적 방법(TO=1) 또는 검출 기반 추적 방법(TD=0)을 나타낸다. 그리고 $T_t^{(TD)}$ 와 $T_t^{(TO)}$ 를 비교하여 S_{card} 와 S_{loc} 의 유사성을 평가한다. 따라서 G_t 를 다음과 같이 정의할 수 있다.

$$G_t = \begin{cases} 1, & \text{s.t. } S_{card} \cdot S_{loc} \geq \theta_{ps} \\ 0, & \text{s.t. } S_{card} \cdot S_{loc} < \theta_{ps} \end{cases} \quad (5)$$

[0068] 이때 θ_{ps} 는 의사 레이블 임계값(threshold)이다. 이때, $S_{card} \cdot S_{loc} \geq \theta_{ps}$ 일 경우 추적 방법 $T_t^{(TD)}$ 와 $T_t^{(TO)}$ 의 추적 품질은 비슷하다. 따라서 해당 프레임에서는 검출 없는 추적 방법 $T_t^{(TO)}$ 를 유도한다. 반대의 경우, 객체 위치정보 혹은 객체 카디널리티 오류로 인해 추적 품질이 저하된 경우이므로, 검출 기반 추적 방법 $T_t^{(TO)}$ 을 유도한다. 따라서 $G_t = 1$ 일 때는 t번째 프레임에서 추적 속도를 향상할 수 있으며, $G_t = 0$ 일 때는 검출 기반 연관을 통해 MOTA를 향상시킬 수 있다. 본 발명에서는 G_t 를 통해 검출 스케줄러를 학습하기 위한 손실함수 $L_{decision}$ 을 제안한다.

$$L_{decision} = -G_t \log(f(s_t)) - (1 - G_t) \log(1 - f(s_t)) \quad (6)$$

[0070] 본 발명의 실시예에 따른 검출 스케줄러를 구현하기 위해 MOT 데이터 셋에서 최신(state-of-the-art) 모델의 정확도와 속도를 나타내는 FairMOT를 베이스라인(baseline)으로 사용한다. 따라서 본 발명에서는 도 3에 제안된 네트워크와 베이스라인을 결합한다. 각 프레임의 기존 트랙과 연관되지 않은 검출 결과가 주어질 때, IoU 연관 방법을 통해 t-1번째 프레임의 트랙과 t번째 프레임의 검출 결과를 연관하여 트랙 초기화를 수행한다. 그리고 $f(s_t)$ 을 추론하여 각 프레임의 추적 동작을 결정한다. 만약 $f(s_t) \leq \theta_{det}$ 라면 검출 기반 추적(TD) 방법을 수행한다. 따라서 검출기를 사용해 검출 결과를 도출한 뒤 기존 트랙과 연관시킨다. 반대로 $f(s_t) > \theta_{det}$ 일 때, 검출 없는 추적(TO) 방법을 수행하며, 칼만 필터로 모션을 예측해 트랙을 도출한다. 또한, 트랙 종료를 위해 트랙 신뢰도(track confidence)를 정의하고 연관된 검출 결과의 분류 점수(classification score)로 초기화한다. 그리고 검출 없는 추적이 발생할 때, 스케일링 계수(본 발명의 실시예에 따른 실험에서는 0.9로 설정)로 트랙 신뢰도를 감소시킨다. 긍정 오류(False Positive)를 줄이기 위해 추적 신뢰도가 0.5 미만일 때 트랙을 종료한다. 전체 추적 속도를 더욱 높이기 위해, 다음과 같이 $f(s_t) \geq \theta_{det}$ 일 때, 검출 스케줄러를 사용하지 않고 $f(s_t)$ 를 t+1번째 프레임으로 전파하여 $f(s_{t+1})$ 를 계산한다.

$$f(s_{t+1}) = \begin{cases} \xi \cdot f(s_t), & \text{s.t. } f(s_t) \geq \theta_{det} \\ DM(F_k, F_t), & \text{s.t. } f(s_t) < \theta_{det} \end{cases} \quad (7)$$

[0072] 여기서, ξ 는 0.83으로 설정한다. $f(s_t) < \theta_{det}$ 일 때 검출 스케줄러를 실행하고, $f(s_t) \geq \theta_{det}$ 일 때 객체의 모션 모델만으로 객체를 추적한다. 이 예측 기반 추적은 $T_t^{(TO)} \approx T_t^{(TD)}$ 일 때 실행할 수 있으며,

$f(s_t) \geq \theta_{\text{det}}$ 일 때 $T_t^{(TO)} \approx T_t^{(TD)}$ 라고 고려한다.

[0074] 도 4는 본 발명의 일 실시예에 따른 데이터셋에 대한 정성적 결과를 나타내는 도면이다.

[0075] 본 발명의 실시예에 따른 검출 스케줄러의 효과를 증명하기 위해 SOTA 모델 간의 비교와 광범위한 절제 실험을 수행 한다.

[0076] 실험에 사용할 데이터 셋은 베이스라인이 MOTChallenge 데이터 셋으로 학습하였기 때문에 본 모델 또한 이 데이터 셋으로 학습한다. 본 발명의 실시예에 따른 검출 스케줄러를 학습하기 위해 MOT17 학습 데이터 셋을 사용한다. 이 학습 데이터 셋은 14-30Hz 프레임의 동적/정적 카메라에서 캡처된 7개의 시퀀스로 구성된다. SOTA 모델 간의 비교 평가를 위해 MOT16, MOT17 및 MOT20 테스트 데이터 셋에서 검출 스케줄러를 실행하고 챌린지 서버에서 성능을 평가한다. 절제 실험(ablation study)의 경우 MOT15 학습 데이터 셋을 검증 데이터 셋으로 활용한다. 단, 학습에 사용되었던 중복된 시퀀스는 사용하지 않는다. 이때 중복되어 절제 실험에 사용되지 않는 시퀀스는 Venice-2, ADL-Rundle-8, ADLRundle-6, ETH-Pedcross2이다.

[0077] 평가 메트릭(Evaluation Metric)은 MOTChallenge에서 일반적으로 사용하는 메트릭을 사용한다. Multi Object Tracking Accuracy(MOTA $\uparrow$), ID F1 score (IDF1 $\uparrow$), False Positive (FP $\downarrow$), False Negative (FN $\downarrow$), ID Switches (ID $\uparrow$), 다중 객체 추적 속도(Hz $\downarrow$). 여기서 $\uparrow$ 와 $\downarrow$ 는 각각 더 높거나 낮을수록 더 우수한 MOT 결과임을 나타낸다. 또한, 검출기 실행 수를 총 프레임 수로 나누어 검출기 실행 비율(Detection Run; DR)을 평가한다. 앞서 설명된 바와 같이, 검출기 실행 비율은 검출 기반 추적 방법의 정확도와 속도에 가장 큰 영향을 미친다. 일반적으로 DR 이 증가할수록 정확도는 높아지지만 속도는 느려진다.

[0078] 앞서 언급한 바와 같이, 본 발명에서는 FairMOT를 베이스라인으로 사용한다. 앞서 제안한 검출 스케줄링 네트워크를 베이스라인에 결합하여 검출 스케줄러를 구현한다. 절제 실험의 경우, 먼저 동일한 학습 설정으로 MOT17 학습 데이터 셋에서 베이스라인을 학습한다. 그 후, 학습된 베이스라인을 프리징(freezing)하고, 식 6의 L_{decision} 를 최소화하는 것으로 검출 스케줄러를 학습한다. 또한, MOTChallenge 테스트 데이터 셋에서 평가하기 위해, MIX 데이터 셋에서 학습된 베이스라인을 프리징하고, 스크래치로 우리의 모델을 학습 한다. 3장 4절에서 의사 GT를 생성하기 위해, 사전 학습된 베이스라인을 사용하여 $T_t^{(TD)}$ 와 $T_t^{(TO)}$ 를 생성한다. 본 발명에서는 추적 품질 측정을 사용하여 두 방법 간의 $S_{\text{card}} \cdot S_{\text{loc}}$ 을 계산하고, 식 5를 사용하여 G_t 를 생성한다. 또한 $\beta_1 = 0.9$ 인 Adam 옵티마이저를 사용한다. 검출 스케줄러는 서로 다른 시퀀스의 6 개의 이미지를 포함한 미니 배치로 30 epoch 동안 추적기를 학습한다. 본 발명에서는 학습률(Learning Rate)을 $5e-5$ 로 설정하고, 20개의 epoch에서 0.1의 비율로 decay한다. 검출 스케줄러는 i7-8700K CPU와 단일 Titan Xp 로 구성된 PC에서 검출 스케줄러를 실행한다.

[0079] MOTA와 추적 속도 간의 트레이드-오프를 조사하기 위해, 일정 간격으로 검출기 수행 스킵을 적용한다. 따라서 각 추적기는 일정 간격으로 검출기 실행을 스킵한다. #4TD-#1TO (DR=80%), #3TD-#1TO (DR=75%), #2TD-#1TO (DR=66.6%), 서로 다른 3가지의 간격을 적용한다. 여기서 #TD와 #TO는 각 각 검출 기반 추적 방법과 검출 없는 추적 방법을 적용하는 프레임의 수다. 검출 스케줄러의 경우, θ_{det} 를 변경하는 것으로 3가지 검출기 실행 스킵 체계와 유사하도록 검출 스케줄러의 DR을 조정한다.

[0080] 표 1은 검출 스케줄러와 베이스라인 간의 자세한 비교 결과를 나타낸다.

[0081] <표 1>

	DR	MOTA $\uparrow$	FP $\downarrow$	FN $\downarrow$	Hz $\uparrow$
Baseline	100%	72.9%	1,409	3,204	19.3
	80.0%	57.4%	1,068	6,210	23.5
	75.0%	53.5%	930	7,005	24.9
	66.6%	46.9%	801	8,253	27.9
Detector Scheduler (Ours)	100%	72.9%	1,409	3,204	18.9
	80.3%	69.2% [3.7% $\downarrow$]	1,881	3,307	23.1 [22.2% $\uparrow$]
	74.1%	68.1% [4.8% $\downarrow$]	1,924	3,523	25.1 [32.8% $\uparrow$]
	66.3%	67.1% [5.8% $\downarrow$]	1,852	3,753	27.8 [47.0% $\uparrow$]

[0082]

[0083] 그리고 DR=100%일 때(즉, 모든 프레임에서 검출기를 실행) 두 추적기가 동일한 MOTA를 달성한다는 것을 확인할 수 있다. 하지만 검출 스케줄러와 베이스라인 간의 정확도 차이는 DR이 감소할수록 더 커진다. 구체적으로 DR=100%의 베이스라인과 비교하여 DR=66%의 검출 스케줄러는 MOTA를 5.8%를 줄이는 것으로 속도를 47.0% 향상할 수 있다. 그러나 DR=66%의 베이스라인은 DR=100%를 사용할 때와 비교하여 MOTA를 26%까지 저하시킨다. 유사한 DR을 가진 다른 추적기와 비교하여, 본 발명의 검출 스케줄러는 더 나은 MOTA와 속도를 보여준다. 이러한 결과로부터, 본 발명의 제안하는 방법이 전체 추적 속도를 높이는 동시에 MOTA 감소를 최소화하는 데 매우 효과적이라는 것을 보여준다.

[0084] 최신 추적기 모델과 비교 평가를 위해 MOT16 및 MOT17 테스트 셋에서 검출 스케줄러를 평가한다. 비교를 위해, 본 발명에서는 전용 검출기(Private Detector)를 사용하는 온라인 추적 방법과 비교한다. 본 발명에서는 이때 θ_{det} 을 0.5로 설정했다. 표 2에 나타난 바와 같이, 검출 스케줄러는 다른 모델보다 더 높은 MOTA를 보여주면서 놀라운 속도를 달성한다.

[0085] <표 2>

	Tracker	MOTA ↑	IDF1 ↑	FP ↓	FN ↓	IDs ↓	Hz ↑
MOT16	QDTrack [21]	69.8%	67.1%	9,861	44,050	1,097	20.3
	CTracker [22]	67.6%	57.2%	8,934	48,305	1,897	6.8
	LM_CNN [23]	67.4%	61.2%	10,109	48,435	981	1.7
	HTA [24]	62.4%	64.2%	19,071	47,839	1,619	19.7
	TraDeS [25]	70.1%	64.7%	8,091	45,210	1,144	22.3
	Tube_TK [26]	64.0%	59.4%	10,962	53,626	1,117	1.0
	GSDT [27]	74.5%	68.1%	8,913	36,428	1,229	1.6
	CenterTrack [28]	69.6%	60.7%	10,458	42,805	2,124	17.5
	SOTMOT [29]	72.1%	72.3%	14,344	34,784	1,681	16.0
	Ours	74.5%	72.3%	8,614	36,856	1,016	20.6
MOT17	QDTrack [21]	68.7%	66.3%	26,589	146,643	3,378	20.3
	CTracker [22]	66.6%	49.0%	22,284	160,491	5,529	6.8
	TraDeS [25]	69.1%	63.9%	20,892	150,060	3,555	17.5
	Tube_TK [26]	63.0%	58.6%	27,060	177,483	4,137	3.0
	GSDT [27]	73.2%	66.5%	26,397	120,666	3,891	4.9
	CenterTrack [28]	67.8%	64.7%	18,498	160,332	3,039	17.5
	PermaTrack [30]	73.8%	68.9%	28,998	115,104	3,699	11.9
	Ours	73.0%	71.3%	23,622	125,664	3,114	20.6

[0086]

[0087] 이 비교를 통해, 제안하는 방법이 추적 정확도를 유지하면서 전체 추적 속도 향상에 실제로 효과적이라는 것을 입증한다.

[0088] 본 발명에서는 실시간으로 작동 가능한 고성능 다중 객체 추적을 위해 최적의 검출 기반 추적 메커니즘을 결정할 수 있는 검출 스케줄러를 제안한다. 제안된 방법에서는 의사 레이블과 검출 스케줄러 손실 함수 기반으로 검출 스케줄러를 학습한다. 결과적으로, 검출 유무에 따른 추적 차이를 줄임으로써 자가 학습 방법을 통해 검출 스케줄러를 학습시킬 수 있다. 또한, 광범위한 절제 실험과 SOTA 모델 간의 비교를 통해, 본 발명에서 제안하는 방법이 MOTA의 손실을 최소화하면서 전체 추적 속도를 높이는 데 유리하다는 것을 검증하였다.

[0090] 이상에서 설명된 장치는 하드웨어 구성요소, 소프트웨어 구성요소, 및/또는 하드웨어 구성요소 및 소프트웨어 구성요소의 조합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치 및 구성요소는, 예를 들어, 프로세서, 콘트롤러, ALU(arithmetic logic unit), 디지털 신호 프로세서(digital signal processor), 마이크로컴퓨터, FPGA(field programmable gate array), PLU(programmable logic unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 하나 이상의 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(OS) 및 상기 운영 체제 상에서 수행되는 하나 이상의 소프트웨어 애플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리 요소(processing element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알 수 있다. 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 콘트롤러를 포함할 수 있다. 또한, 병렬 프로세서(parallel processor)와 같은, 다른 처리 구성(processing configuration)도 가능하다.

[0091] 소프트웨어는 컴퓨터 프로그램(computer program), 코드(code), 명령(instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로

(collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성요소(component), 물리적 장치, 가상 장치(virtual equipment), 컴퓨터 저장 매체 또는 장치에 구체화(embody)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.

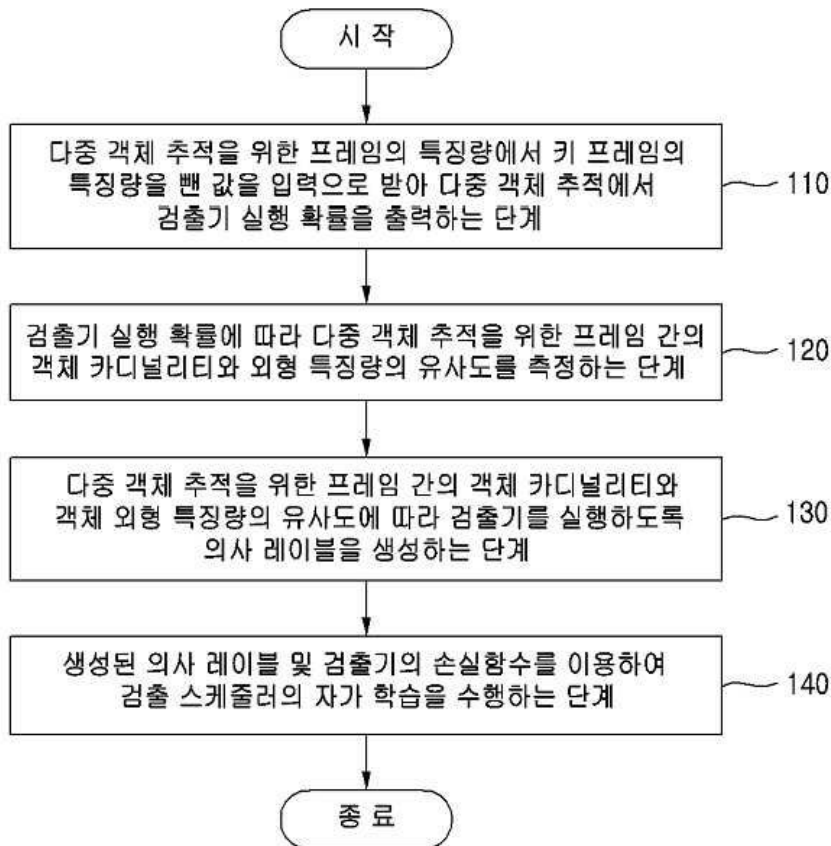
[0092] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다.

[0093] 이상과 같이 실시예들이 비록 한정된 실시예와 도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 상기의 기재로부터 다양한 수정 및 변형이 가능하다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다.

[0094] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

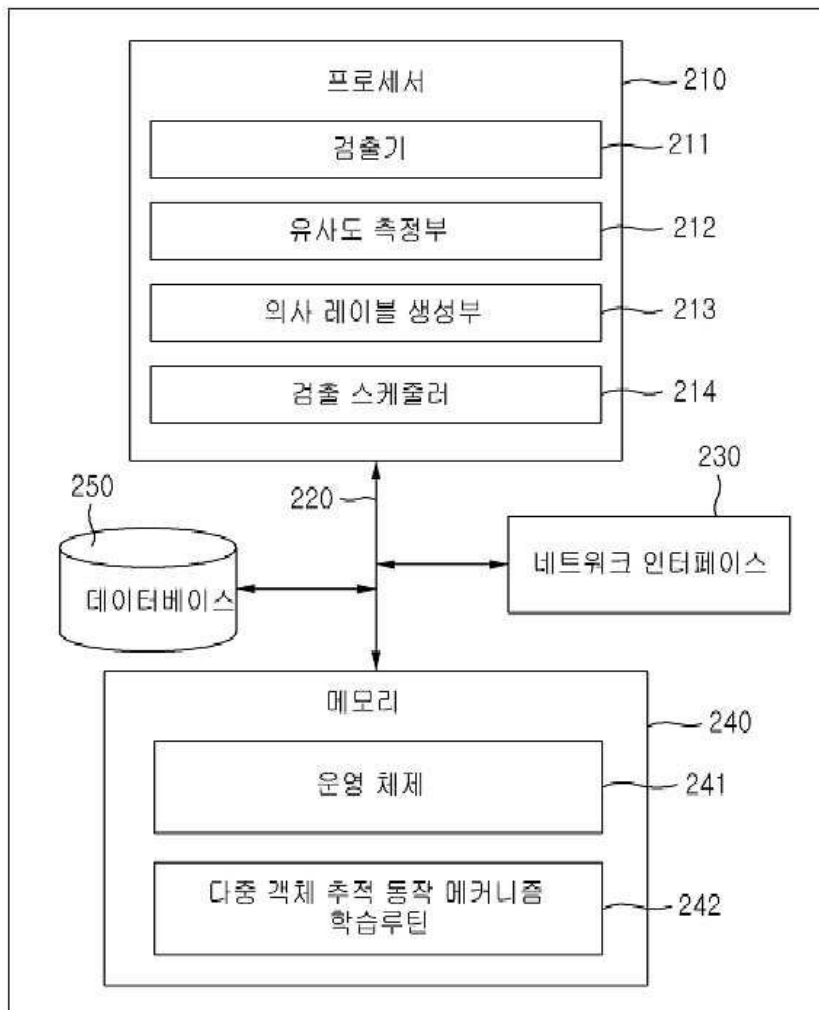
도면

도면1

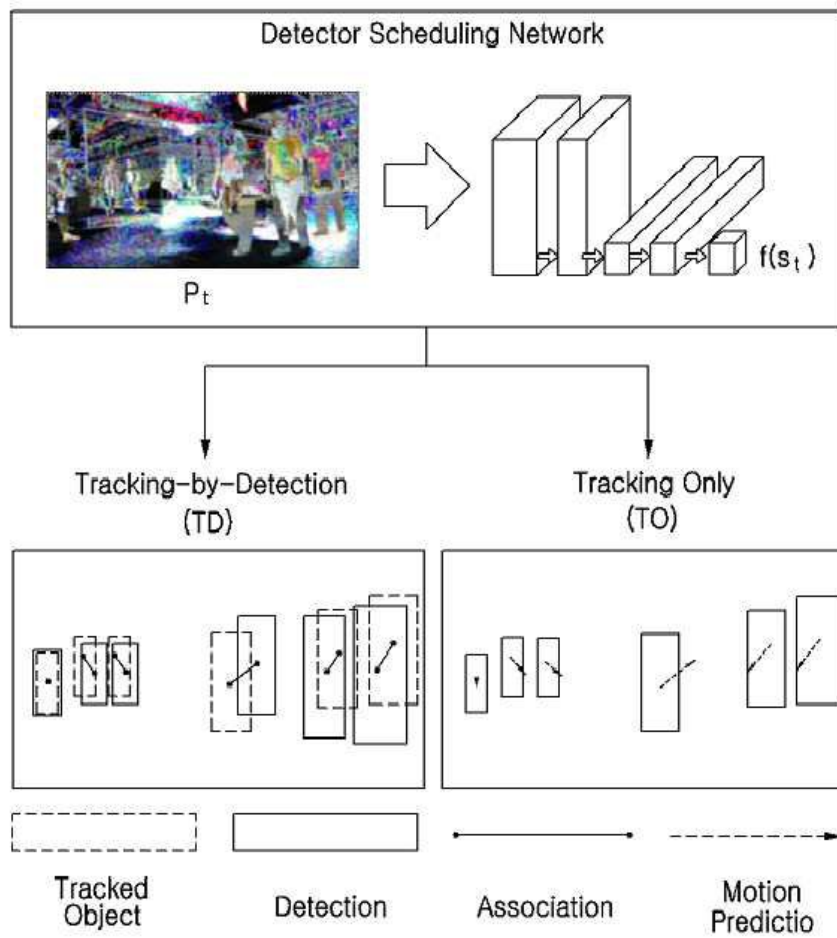


도면2

200



도면3



도면4

