



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2023년04월06일
(11) 등록번호 10-2519367
(24) 등록일자 2023년04월04일

(51) 국제특허분류(Int. Cl.)
G06T 7/246 (2017.01) G06F 18/00 (2023.01)
G06N 3/08 (2023.01) G06T 7/11 (2017.01)
G06T 7/292 (2017.01) H04N 7/18 (2023.01)

(52) CPC특허분류
G06T 7/246 (2017.01)
G06N 3/08 (2023.01)

(21) 출원번호 10-2020-0155092

(22) 출원일자 2020년11월19일

심사청구일자 2020년11월19일

(65) 공개번호 10-2022-0068373

(43) 공개일자 2022년05월26일

(56) 선행기술조사문헌
Debaditya Acharya 등, Real-time detection and tracking of pedestiran in CCTV images using a deep convolutional neural network, CEUR Workshop Proceedings. (2017.04.)*
(뒷면에 계속)

(73) 특허권자
연세대학교 산학협력단
서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)

(72) 발명자
이상윤
서울특별시 서초구 청계산로7길 43, 504동 1302호(신원동, 서초포레스타5단지)

정태영
서울특별시 서대문구 연희로10길 24-10, 101호(연희동)
(뒷면에 계속)

(74) 대리인
민영준

전체 청구항 수 : 총 9 항

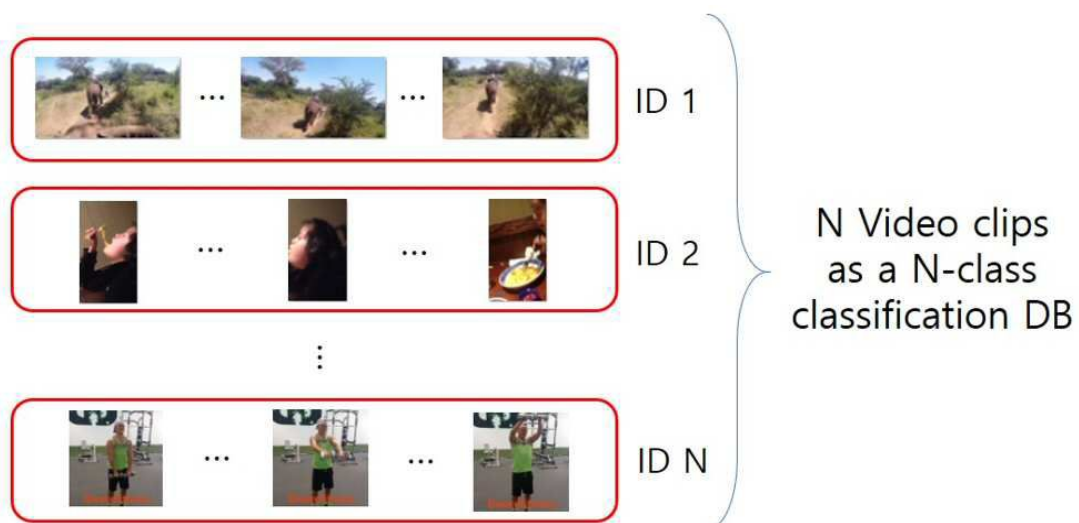
심사관 : 이재원

(54) 발명의 명칭 CCTV 환경에서의 보행자 추적 장치 및 방법

(57) 요약

본 발명은 기지정된 방식에 따라 인가된 입력 영상의 다수의 프레임 각각에서 보행자가 포함된 영역을 추출하여 패치를 획득하는 보행자 검출부, 미리 학습된 인공 신경망으로 구현되어, 학습된 방식에 따라 패치의 특징을 추출하여 패치 특징맵을 획득하는 특징 추출 네트워크 및 다수 프레임의 영상 중 서로 다른 프레임 영상에서 획득된 패치 특징맵을 서로 비교하여 대응하는 패치 특징맵을 탐색하여 보행자를 추적하는 추적부를 포함하고, 특징 추출 네트워크는 학습 시에 보행자의 포함 여부에 무관하게 획득된 학습 영상에 식별자가 부여된 학습 데이터를 인가받아 학습 데이터의 다수의 프레임의 특징을 추출하여 학습 특징맵을 획득하고, 별도로 구비된 학습부가 학습 특징맵을 분류하여 계산되는 손실을 역전파하여 학습되어, 학습 데이터로 보행자들에 대해 미리 레이블된 대량의 패치를 필요로 하지 않는 보행자 추적 장치 및 방법을 제공할 수 있다.

대표도 - 도7



(52) CPC특허분류

G06T 7/11 (2017.01)
G06T 7/292 (2017.01)
G06V 40/10 (2022.01)
H04N 7/181 (2013.01)
G06T 2207/20081 (2013.01)
G06T 2207/20084 (2013.01)
G06T 2210/12 (2013.01)

(72) 발명자

조명아

경기도 성남시 분당구 판교로 421, 204동 403호(야
탑동, 탑마을대우아파트)

조수환

서울특별시 양천구 목동동로 350, 510동 304호(목
동, 목동신시가지아파트5단지)

(56) 선행기술조사문헌

Hyeonseob Nam 등, Learning Multi-Domain
Convolutional Neural Networks for Visual
Tracking, arXiv:1510.07945v2 (2016.01.06.)*
JP2015167017 A
JP2016219004 A
KR1020160132731 A
KR1020180042254 A

*는 심사관에 의하여 인용된 문헌

이 발명을 지원한 국가연구개발사업

과제고유번호	1711119679
과제번호	2018M3E3A1057289
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	복합인지기술개발사업
연구과제명	이중 CCTV 영상에서의 딥러닝 기반 실종자 초동 신원확인 및 추적 시스템 (2
단계)(1/3)	
기 여 율	1/1
과제수행기관명	연세대학교 산학협력단
연구기간	2020.04.23 ~ 2020.12.31

명세서

청구범위

청구항 1

기 지정된 방식에 따라 인가된 입력 영상의 다수의 프레임 각각에서 보행자가 포함된 영역을 추출하여 패치를 획득하는 보행자 검출부;

미리 학습된 인공 신경망으로 구현되어, 학습된 방식에 따라 상기 패치의 특징을 추출하여 패치 특징맵을 획득하는 특징 추출 네트워크; 및

다수 프레임의 영상 중 서로 다른 프레임 영상에서 획득된 패치 특징맵을 서로 비교하여 대응하는 패치 특징맵을 탐색하여 보행자를 추적하는 추적부를 포함하고,

상기 특징 추출 네트워크는

학습 시에 보행자의 포함 여부에 무관하게 획득된 학습 영상에 식별자가 부여된 학습 데이터를 인가받아 상기 학습 데이터의 다수의 프레임의 특징을 추출하여 학습 특징맵을 획득하고, 별도로 구비된 학습부가 상기 학습 특징맵을 분류하여 계산되는 손실을 역전파하여 학습되고,

상기 학습 데이터는

다수의 카메라에서 획득된 다수의 영상 각각에서 기 지정된 시간 길이의 연속되는 프레임을 추출하여 획득되는 상기 학습 영상에 상기 식별자를 부여하여 획득되며,

상기 학습부는

학습 시에 상기 특징 추출 네트워크와 함께 학습이 수행되어, 학습되는 방식에 따라 상기 특징 추출 네트워크에서 획득된 상기 학습 특징맵을 분류하는 영상 분류부; 및

상기 학습 특징맵의 분류 결과와 상기 학습 데이터에 부여된 식별자를 비교하여 기 지정된 방식으로 손실을 계산하고, 계산된 손실을 역전파하여 상기 특징 추출 네트워크와 상기 영상 분류부를 학습시키는 손실 역전파부를 포함하는 보행자 추적 장치.

청구항 2

삭제

청구항 3

제1항에 있어서, 상기 학습 영상은

카메라의 회전 또는 이동 등의 변화 없이 고정된 위치를 촬영하여 획득된 영상인 보행자 추적 장치.

청구항 4

삭제

청구항 5

제1항에 있어서, 상기 보행자 검출부는

미리 학습된 방식에 따라 상기 입력 영상의 각 프레임에서 보행자가 포함된 영역을 추정하여, 보행자가 포함된 영역을 나타내는 경계 박스의 위치 및 크기를 설정하는 경계 박스 설정부; 및

상기 입력 영상에서 설정된 경계 박스에 대응하는 영역을 추출하여 상기 패치를 획득하는 패치 획득부를 포함하

는 보행자 추적 장치.

청구항 6

제1항에 있어서, 상기 추적부는

기지정된 기간 동안의 다수의 프레임 각각에 대응하는 패치 특징맵이 인가되면, 인가된 패치 특징맵을 저장하고, 현재 인가된 패치 특징 맵과 저장된 패치 특징맵들 사이의 유사도를 분석하여, 동일 보행자에 대한 패치 특징맵들을 판별하여 매칭시키고, 매칭된 패치 특징맵들의 각 프레임에서의 위치를 확인하여 보행자를 추적하는 보행자 추적 장치.

청구항 7

보행자 검출부, 특징 추출 네트워크 및 추적부를 포함하는 보행자 추적 장치의 보행자 추적 방법에 있어서,

학습 시에 인공 신경망으로 구현되는 상기 특징 추출 네트워크가 보행자의 포함 여부에 무관하게 획득된 학습 영상에 식별자가 부여된 학습 데이터를 인가받아 상기 학습 데이터의 다수의 프레임의 특징을 추출하여 학습 특징맵을 획득하면, 획득된 상기 학습 특징맵을 분류하여 계산되는 손실을 역전파하여 상기 특징 추출 네트워크를 학습시키는 단계;

상기 특징 추출 네트워크에 대한 학습이 종료된 이후, 상기 보행자 검출부가 기지정된 방식에 따라 인가된 입력 영상의 다수의 프레임 각각에서 보행자가 포함된 영역을 추출하여 패치를 획득하는 단계;

상기 특징 추출 네트워크가 학습된 방식에 따라 상기 패치의 특징을 추출하여 패치 특징맵을 획득하는 단계; 및
상기 추적부가 다수 프레임의 영상 중 서로 다른 프레임 영상에서 획득된 패치 특징맵을 서로 비교하여 대응하는 패치 특징맵을 탐색하여 보행자를 추적하는 단계를 포함하되,

상기 학습 데이터는

다수의 카메라에서 획득된 다수의 영상 각각에서 기지정된 시간 길이의 연속되는 프레임을 추출하여 획득되는 상기 학습 영상에 상기 식별자를 부여하여 획득되고,

상기 학습시키는 단계는

학습 시에 상기 특징 추출 네트워크와 함께 학습이 수행되는 인공 신경망을 이용하여, 학습되는 방식에 따라 상기 특징 추출 네트워크에서 획득된 상기 학습 특징맵을 분류하는 단계; 및

상기 학습 특징맵의 분류 결과와 상기 학습 데이터에 부여된 식별자를 비교하여 기지정된 방식으로 손실을 계산하고, 계산된 손실을 역전파하는 단계를 포함하는 보행자 추적 방법.

청구항 8

삭제

청구항 9

제7항에 있어서, 상기 학습 영상은

카메라의 회전 또는 이동 등의 변화 없이 고정된 위치를 촬영하여 획득된 영상인 보행자 추적 방법.

청구항 10

삭제

청구항 11

제7항에 있어서, 상기 패치를 획득하는 단계는

미리 학습된 방식에 따라 상기 입력 영상의 각 프레임에서 보행자가 포함된 영역을 추정하는 단계;

추정된 보행자가 포함된 영역을 나타내는 경계 박스의 위치 및 크기를 설정하는 단계; 및

상기 입력 영상에서 설정된 경계 박스에 대응하는 영역을 추출하여 상기 패치를 획득하는 단계를 포함하는 보행자 추적 방법.

청구항 12

제11항에 있어서, 상기 보행자를 추적하는 단계는

기지정된 기간 동안의 다수의 프레임 각각에 대응하는 패치 특징맵이 인가되면, 인가된 패치 특징맵을 저장하는 단계;

현재 인가된 패치 특징 맵과 저장된 패치 특징맵들 사이의 유사도를 분석하여, 동일 보행자에 대한 패치 특징맵들을 판별하여 매칭시키는 단계; 및

매칭된 패치 특징맵들의 각 프레임에서의 위치를 확인하여 보행자를 추적하는 단계를 포함하는 보행자 추적 방법.

청구항 13

제7항에 따른 보행자 추적 방법을 실행하기 위한 프로그램 명령어가 기록된 기록 매체.

발명의 설명

기술 분야

[0001] 본 발명은 CCTV 환경에서의 보행자 추적 장치 및 방법에 관한 것으로, 자기 지도 학습 기반 CCTV 환경에서의 보행자 추적 장치 및 방법에 관한 것이다.

배경 기술

[0002] 다중 객체 추적(Multi-object Tracking) 분야는 단일 카메라에서 촬영된 여러 대상을 동시에 추적하는 기술을 다루는 분야로서, 차량과 같은 다양한 객체를 추적하는 것을 목표로 하는 연구도 있으나, 주로 CCTV 영상에 기반한 보행자들을 추적하는 기술이 많이 연구되고 있다.

[0003] 이런 추적 분야에서 주로 사용되는 접근법은 검출 기반 추적(Tracking by Detection) 이라고 불리는 패러다임이다. 이는 특정 객체를 검출하는 객체 검출기를 활용하여 영상의 다수의 프레임 각각에서 대상 객체들을 검출하고, 검출된 대상 객체를 프레임의 시간에 따라 연결하는 접근법이다. 따라서 검출 결과로 각 프레임에서 추출된 다수의 객체들 중 서로 대응하는 객체들을 정확하게 매칭시키는 것이 추적 알고리즘 성능의 핵심이 된다.

[0004] 현재 다양한 검출 기반 추적 기법이 제시되었으며, 이중 대표적 기법으로 다수의 프레임 각각에서 다수의 객체 각각이 검출된 영역을 나타내는 객체 패치의 특징을 추출하여 서로 매칭하는 방법이 있다.

[0005] 도 1은 영상에서 객체를 검출하여 패치를 획득하는 일 예를 나타낸다.

[0006] 패치 기반 객체 추적 기법에서는 영상에서 검출 대상 객체가 포함된 경계 박스를 검출하고, 검출된 경계 박스를 객체 패치로 획득한다. 도 1에서는 보라색 박스가 검출된 경계 박스를 나타낸다. 그리고 영상에서 검출된 경계 박스 각각이 객체 패치로 획득한다.

[0007] 객체 패치가 획득되면, 미리 학습된 특징 추출 네트워크를 이용하여 획득된 패치의 특징을 추출한다. 이후 서로 다른 프레임에서 획득된 다수의 패치에 각각에서 추출된 특징을 서로 비교하여 유사한 특징을 갖는 패치끼리 매칭함으로써, 다수의 프레임에서 특정 객체가 검출된 영역의 변화를 분석하여 객체의 이동 경로를 추적할 수 있다. 이에 도 1에서는 추적 대상 객체인 보행자의 이동이 빨간색 선으로 표시되었다.

[0008] 이러한 패치 기반 객체 추적 기법의 경우, 다수의 대상 객체가 포함된 영상에서도 각 객체에 대한 패치로부터 추출된 특징을 비교하여 객체에 대한 추적을 수행하므로, 다수의 객체 추적에도 용이하게 이용될 수 있다.

[0009] 도 2 및 도 3은 영상에서 객체를 검출하여 획득된 패치를 기반으로 추출된 특징을 비교하여 보행자를 추적하는 개념을 설명하기 위한 도면이다.

[0010] 도 2 및 도 3에서 (a) 및 (b)는 영상의 서로 다른 2개의 프레임을 도시하였으며, 2개의 프레임에서 2명의 보행

자에 대한 패치가 녹색과 보라색 경계 박스로 획득된 경우를 나타낸다. 도 2의 (c) 및 (d)에 나타난 바와 같이, 패치 기반 객체 추적 기법에서는 영상에 포함된 다수의 보행자를 검출하여 각 보행자에 대한 패치가 획득되면, 미리 학습된 방식에 따라 획득된 패치로부터 특징을 추출한다. 그리고 각 프레임에서 획득된 다수의 패치들에서 추출된 특징을 서로 비교하여, 대응하는 패치를 탐색한다. 그리고 대응하는 보행자에 대한 패치가 탐색되면 도 3과 같이 (c)와 (f) 패치를 서로 매칭시키고, (d)와 (e) 패치를 서로 매칭시켜 각 프레임에서의 보행자 위치를 추적할 수 있다.

[0011] 이러한 패치 기반 객체 추적 기법은 우수한 성능을 나타낼 수 있으나, 이를 위해서는 먼저 객체 패치로부터 특징을 추출하는 특징 추출 네트워크가 충분하게 학습되어야만 한다.

[0012] 도 4는 패치로부터 특징을 추출하는 특징 추출 네트워크를 학습시키기 위한 학습 데이터의 일 예를 나타낸다.

[0013] 상기한 바와 같이, 패치 기반 객체 추적 기법에서는 특징 추출 네트워크가 충분하게 학습되어야만 하며, 입력된 객체 패치로부터 특징을 추출하는 특징 추출 네트워크를 학습시키기 위해서는 학습 대상 객체가 미리 레이블링된 대량의 학습 데이터가 필요하다. 일 예로 도 4의 우측에 도시된 바와 같이, 학습 데이터는 다수의 프레임에서 획득된 패치들에서 동일한 객체에 대한 패치들이 동일한 식별자(ID)로 레이블되어 특징 추출 네트워크가 동일한 객체에 대한 패치임을 인식할 수 있어야 한다. 그러나 학습 데이터는 일반적으로 다수의 패치에 대해 사람이 수작업으로 식별자를 레이블링하여 획득해야 하므로, 대량의 학습데이터를 획득하는 것은 용이하지 않다. 비록 영상이 1분 분량의 짧은 영상일지라도 30 fps 기준으로 1800개의 대량의 프레임이 포함되며, 1800개의 프레임에서 추출되는 패치들 각각에 대응하는 식별자를 레이블링하는 것은 매우 어렵다.

[0014] 도 5는 다수의 보행자가 포함된 영상 프레임의 일 예를 나타낸다.

[0015] 상기한 바와 같이, 짧은 시간 분량의 영상에서도 대량의 프레임으로 인해 패치를 수작업으로 레이블링하는 것은 매우 어려우나, 영상에서는 도 5와 같이, 추적 대상 객체인 보행자가 다수로 포함된 영상 또한 존재한다. 이렇게 다수의 보행자가 포함된 영상에서 추출된 패치들을 레이블링하여 학습 데이터를 획득하는 것은 현실적으로 불가능하다.

[0016] 따라서 기존의 패치 기반 객체 추적 기법에서는 적은 양의 학습 데이터로만 학습되어 요구되는 성능을 나타내지 못하는 문제가 빈번하게 발생하였다.

선행기술문헌

특허문헌

[0017] (특허문헌 0001) 한국 등록 특허 제10-1941994호 (2019.01.18 등록)

발명의 내용

해결하려는 과제

[0018] 본 발명의 목적은 레이블된 대량의 학습 데이터없이 자기 지도 학습 방식으로 학습되어 특징을 추출하여 보행자를 추적할 수 있는 보행자 추적 장치 및 방법을 제공하는데 있다.

[0019] 본 발명의 다른 목적은 시간적으로 연속된 다수의 프레임을 기반으로 특징 추출 네트워크가 학습되어 보행자의 형상 변화에도 강건하게 보행자를 추적할 수 있는 보행자 추적 장치 및 방법을 제공하는데 있다.

과제의 해결 수단

[0020] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른 보행자 추적 장치는 기지정된 방식에 따라 인가된 입력 영상의 다수의 프레임 각각에서 보행자가 포함된 영역을 추출하여 패치를 획득하는 보행자 검출부; 미리 학습된 인공 신경망으로 구현되어, 학습된 방식에 따라 상기 패치의 특징을 추출하여 패치 특징맵을 획득하는 특징 추출 네트워크; 및 다수 프레임의 영상 중 서로 다른 프레임 영상에서 획득된 패치 특징맵을 서로 비교하여 대응하는 패치 특징맵을 탐색하여 보행자를 추적하는 추적부를 포함하고, 상기 특징 추출 네트워크는 학습 시에 보행자의 포함 여부에 무관하게 획득된 학습 영상에 식별자가 부여된 학습 데이터를 인가받아 상기 학습 데이터의 다수의 프레임의 특징을 추출하여 학습 특징맵을 획득하고, 별도로 구비된 학습부가 상기 학습 특징맵을 분류하

여 계산되는 손실을 역전파하여 학습된다.

- [0021] 상기 학습 데이터는 다수의 카메라에서 획득된 다수의 영상 각각에서 기지정된 시간 길이의 연속되는 프레임을 추출하여 획득되는 상기 학습 영상에 상기 식별자를 부여하여 획득될 수 있다.
- [0022] 상기 학습 영상은 카메라의 회전 또는 이동 등의 변화 없이 고정된 위치를 촬영하여 획득될 수 있다.
- [0023] 상기 학습부는 학습 시에 상기 특징 추출 네트워크와 함께 학습이 수행되어, 학습되는 방식에 따라 상기 특징 추출 네트워크에서 획득된 상기 학습 특징맵을 분류하는 영상 분류부; 및 상기 학습 특징맵의 분류 결과와 상기 학습 데이터에 부여된 식별자를 비교하여 기지정된 방식으로 손실을 계산하고, 계산된 손실을 역전파하여 상기 특징 추출 네트워크와 상기 영상 분류부를 학습시키는 손실 역전파부를 포함할 수 있다.
- [0024] 상기 보행자 검출부는 미리 학습된 방식에 따라 상기 입력 영상의 각 프레임에서 보행자가 포함된 영역을 추정하여, 보행자가 포함된 영역을 나타내는 경계 박스의 위치 및 크기를 설정하는 경계 박스 설정부; 및 상기 입력 영상에서 설정된 경계 박스에 대응하는 영역을 추출하여 상기 패치를 획득하는 패치 획득부를 포함할 수 있다.
- [0025] 상기 추적부는 기지정된 기간 동안의 다수의 프레임 각각에 대응하는 패치 특징맵이 인가되면, 인가된 패치 특징맵을 저장하고, 현재 인가된 패치 특징 맵과 저장된 패치 특징맵들 사이의 유사도를 분석하여, 동일 보행자에 대한 패치 특징맵들을 판별하여 매칭시키고, 매칭된 패치 특징맵들의 각 프레임에서의 위치를 확인하여 보행자를 추적할 수 있다.
- [0026] 상기 목적을 달성하기 위한 본 발명의 다른 실시예에 따른 보행자 추적 방법은 보행자 검출부, 특징 추출 네트워크 및 추적부를 포함하는 보행자 추적 장치의 보행자 추적 방법에 있어서,
- [0027] 학습 시에 인공 신경망으로 구현되는 상기 특징 추출 네트워크가 보행자의 포함 여부에 무관하게 획득된 학습 영상에 식별자가 부여된 학습 데이터를 인가받아 상기 학습 데이터의 다수의 프레임의 특징을 추출하여 학습 특징맵을 획득하면, 획득된 상기 학습 특징맵을 분류하여 계산되는 손실을 역전파하여 상기 특징 추출 네트워크를 학습시키는 단계; 상기 특징 추출 네트워크에 대한 학습이 종료된 이후, 상기 보행자 검출부가 기지정된 방식에 따라 인가된 입력 영상의 다수의 프레임 각각에서 보행자가 포함된 영역을 추출하여 패치를 획득하는 단계; 상기 특징 추출 네트워크가 학습된 방식에 따라 상기 패치의 특징을 추출하여 패치 특징맵을 획득하는 단계; 및 상기 추적부가 다수 프레임의 영상 중 서로 다른 프레임 영상에서 획득된 패치 특징맵을 서로 비교하여 대응하는 패치 특징맵을 탐색하여 보행자를 추적하는 단계를 포함한다.

발명의 효과

- [0028] 따라서, 본 발명의 실시예에 따른 보행자 추적 장치 및 방법은 각 카메라에서 촬영된 기지정된 구간 동안의 다수의 프레임들을 학습 데이터로 이용하여 특징 추출 네트워크를 학습시켜, 학습 데이터로 보행자들에 대해 미리 레이블된 대량의 패치를 필요로 하지 않아 학습이 용이하다. 또한 연속하는 다수의 프레임을 학습 데이터로 이용함에 따라 보행자의 형상 변화에도 강건하게 특징을 추출할 수 있을 뿐만 아니라 배경 환경 변화 등에서도 정확하게 보행자의 특징을 추출하여 보행자 추적 성능을 향상시킬 수 있다.

도면의 간단한 설명

- [0029] 도 1은 영상에서 객체를 검출하여 패치를 획득하는 일 예를 나타낸다.
- 도 2 및 도 3은 영상에서 객체를 검출하여 획득된 패치를 기반으로 추출된 특징을 비교하여 보행자를 추적하는 개념을 설명하기 위한 도면이다.
- 도 4는 패치로부터 특징을 추출하는 특징 추출 네트워크를 학습시키기 위한 학습 데이터의 일 예를 나타낸다.
- 도 5는 다수의 보행자가 포함된 영상 프레임의 일 예를 나타낸다.
- 도 6은 본 발명의 일 실시예에 따른 보행자 추적 장치의 개략적 구조를 나타낸다.
- 도 7은 도 6의 특징 추출 네트워크를 학습시키기 위한 학습 데이터의 일 예를 나타낸다.
- 도 8은 도 7의 학습 데이터를 이용하여 도 6의 특징 추출 네트워크를 자기 지도 학습 방식에 따라 학습시키는 학습부의 일 예를 나타낸다.
- 도 9는 본 발명의 일 실시예에 따른 보행자 추적 방법을 나타낸다.

발명을 실시하기 위한 구체적인 내용

- [0030] 본 발명과 본 발명의 동작상의 이점 및 본 발명의 실시에 의하여 달성되는 목적을 충분히 이해하기 위해서는 본 발명의 바람직한 실시예를 예시하는 첨부 도면 및 첨부 도면에 기재된 내용을 참조하여야만 한다.
- [0031] 이하, 첨부한 도면을 참조하여 본 발명의 바람직한 실시예를 설명함으로써, 본 발명을 상세히 설명한다. 그러나, 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며, 설명하는 실시예에 한정되는 것이 아니다. 그리고, 본 발명을 명확하게 설명하기 위하여 설명과 관계없는 부분은 생략되며, 도면의 동일한 참조부호는 동일한 부재임을 나타낸다.
- [0032] 명세서 전체에서, 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라, 다른 구성요소를 더 포함할 수 있는 것을 의미한다. 또한, 명세서에 기재된 "...부", "...기", "모듈", "블록" 등의 용어는 적어도 하나의 기능이나 동작을 처리하는 단위를 의미하며, 이는 하드웨어나 소프트웨어 또는 하드웨어 및 소프트웨어의 결합으로 구현될 수 있다.
- [0033] 도 6은 본 발명의 일 실시예에 따른 보행자 추적 장치의 개략적 구조를 나타낸다.
- [0034] 도 6을 참조하면, 본 실시예에 따른 보행자 추적 장치는 영상 획득부(110), 보행자 검출부(120), 특징 추출 네트워크(130) 및 추적부(140)를 포함할 수 있다.
- [0035] 영상 획득부(100)는 보행자를 검출하여 추적 또는 재식별하기 위한 입력 영상을 획득한다. 영상 획득부(100)는 CCTV나 카메라 등과 같이 영상 획득 수단으로 구현될 수 있으며, 영상 획득 수단에서 획득된 영상을 저장하는 저장 장치 등으로 구현될 수 있다. 여기서 영상 획득부(100)는 연속하는 다수 프레임을 포함하는 입력 영상을 획득할 수 있다.
- [0036] 보행자 검출부(120)는 영상 획득부(110)에서 획득된 입력 영상을 인가받고, 미리 학습된 방식에 따라 입력 영상의 다수의 프레임 각각에서 보행자가 포함된 영역을 검출하여 패치를 획득한다.
- [0037] 보행자 검출부(120)는 경계 박스 설정부(121) 및 패치 획득부(122)를 포함할 수 있다. 우선 경계 박스 설정부(121)는 로우 레벨 특징맵을 인가받고, 미리 학습된 패턴 추정 방식에 따라 로우 레벨 특징맵에서 보행자가 포함된 영역을 추정하여, 경계 박스의 위치 및 크기를 설정한다.
- [0038] 경계 박스 설정부(121)는 영상 획득부(100)가 획득한 입력 영상에서 보행자를 검출하고, 검출된 보행자가 포함되는 영역의 크기를 추정하여 대응하는 크기의 경계 박스를 설정할 수 있다. 도 1 및 도 2를 참조하면, 경계 박스는 직사각형 형태로 설정될 수 있으나, 경계 박스의 형태는 다양하게 조절될 수 있다. 그리고 경계 박스 설정부(121)는 도 1 내지 도 5에 도시된 바와 같이, 입력 영상에 다수의 보행자가 포함된 경우, 다수의 보행자 각각을 개별적으로 검출하여 다수의 경계 박스를 설정할 수 있다. 이때 경계 박스 설정부(121)는 검출된 다수의 보행자에 대응하는 영역을 각각 추정하여 다양한 크기의 다수의 경계 박스를 설정할 수 있다.
- [0039] 또한 경우에 따라서 경계 박스 설정부(121)는 더욱 정확하게 보행자를 추정하여 경계 박스를 설정할 수 있도록 추적부(140)가 보행자를 추적한 결과를 인가받아 영상에서 보행자를 판별하는데 이용할 수 있다. 이는 도 3에서 보라색 및 녹색 화살표로 표시된 바와 같이, 이전 획득된 영상을 이용하여 추적 중이던 보행자의 경우, 보행자의 이동 방향이 미리 예측될 수 있음에 따라 더욱 정확하게 보행자를 판별하고, 경계 박스를 설정할 수 있기 때문이다.
- [0040] 경계 박스 설정부(121)는 미리 학습된 인공 신경망 등으로 구현될 수 있으며, 이외에도 영상의 프레임에서 지정된 객체를 검출하는 다양한 기법이 공지되어 있으므로 여기서는 상세하게 설명하지 않는다.
- [0041] 한편, 경계 박스 설정부(121)에 의해 경계 박스가 설정되면, 패치 획득부(122)는 경계 박스에 의해 설정된 영역을 추출하여 패치를 획득한다. 이때 상기한 바와 같이, 경계 박스 설정부(121)가 입력 영상의 다수의 프레임 각각에서 다수의 경계 박스를 설정한 경우, 패치 획득부(122)는 각 프레임에서 다수의 패치를 추출하여 획득할 수 있다.
- [0042] 특징 추출 네트워크(130)는 미리 학습된 인공 신경망으로 구현되어 보행자 검출부(120)에서 획득된 패치를 인가받고 학습된 방식에 따라 패치의 특징을 추출하여 패치 특징맵을 획득한다.
- [0043] 상기한 바와 같이, 기존의 보행자 검출 장치의 경우 특징 추출 네트워크(130)는 보행자 검출부(120)에 의해 획득된 패치 각각 대해 객체 식별자가 레이블링된 대량의 학습 데이터에 의해 학습이 수행되어야 한다. 그러나

대량의 학습 데이터를 획득하기 어려우므로, 본 실시예에서 특징 추출 네트워크(130)는 패치에 객체 식별자가 레이블링된 대량의 학습 데이터를 이용한 지도 학습(supervised learning) 방식이 아니라 자기 지도 학습 방식(Self-supervised learning)에 따라 학습이 수행된다.

- [0044] 자기 지도 학습 방식(Self-supervised learning)에 따라 학습되는 특징 추출 네트워크(130)에 대한 상세한 설명은 후술하도록 한다.
- [0045] 추적부(140)는 특징 추출 네트워크(130)에서 획득된 패치 특징맵을 인가받고, 이전 다른 프레임으로부터 획득된 패치 특징맵과 유사도를 비교하여, 보행자를 추적한다. 상기한 바와 같이, 보행자 추적 장치에서 영상 획득부(110)는 시간적으로 연속하는 다수 프레임의 영상을 획득할 수 있다. 이에 추적부(140)는 각 프레임 영상에서 검출된 패치에 대해 특징 추출 네트워크(130)가 추출한 패치 특징맵 사이의 유사도를 분석하여 동일 보행자에 대한 패치 특징맵인지를 판별한다. 즉 추적부(140)는 다수 프레임 영상에서 이전 프레임에서 각각에서 검출된 패치로부터 추출된 패치 특징맵과 현재 프레임에서 추출된 패치 특징맵 사이의 유사도를 분석하여 동일 보행자에 대한 패치 특징맵인지를 판별할 수 있다.
- [0046] 그리고 추적부(140)는 연속하는 다수의 프레임에서 동일한 보행자에 대한 패치 특징맵을 계속적으로 판별하고 서로 매칭시킴으로써, 도 3과 같이, 영상에서 검출된 각 보행자의 위치 변화를 추적하여 표시할 수 있다.
- [0047] 이때, 입력 영상의 다수 프레임에서 특정 보행자가 다른 객체나 장애물 등에 의해 일시적으로 가려지거나 형상의 변화가 크게 나타나는 경우가 발생할 수도 있다. 이와 같이, 보행자가 일시적으로 사라지거나 형상이 크게 변화하는 경우에도, 이후 프레임에서 다시 정상적으로 대응하는 보행자를 추적할 수 있도록 추적부(140)는 특징 추출 네트워크(130)에서 이전 인가된 다수의 패치 특징맵을 누적하여 저장할 수도 있다. 즉 추적부(140)는 직전 프레임에서 추출된 패치 특징맵에서 현재 탐색해야 하는 패치 특징맵에 대응하는 패치 특징맵을 탐색하지 못하는 경우, 누적 저장된 그 이전 기지정된 기간 동안의 프레임에서 추출된 패치 특징맵과 유사도 비교를 통해 탐색할 수도 있다.
- [0048] 또한 경우에 따라서 추적부(140)는 경계 박스 설정부(121)와 마찬가지로 이전 추적된 보행자의 이동 방향 또는 이동 경로를 더 반영하여 패치 특징맵들 사이의 유사도를 분석할 수도 있다.
- [0049] 도 7은 도 6의 특징 추출 네트워크를 학습시키기 위한 학습 데이터의 일 예를 나타내고, 도 8은 도 7의 학습 데이터를 이용하여 도 6의 특징 추출 네트워크를 자기 지도 학습 방식에 따라 학습시키는 학습부의 일 예를 나타낸다.
- [0050] 자기 지도 학습 방식은 비지도 학습(Unsupervised learning) 방식의 일종으로 고안된 학습 방식으로 비록 네트워크의 실제 목적과 다른 용도로 획득된 데이터일지라도 수작업에 의한 레이블링 없이 획득될 수 있는 구실 업무(pretext tasks)를 설계하여 학습 데이터로 획득하고, 획득된 학습 데이터를 이용하여 네트워크를 학습시킨 후, 학습된 네트워크를 실제 목적에 적용하는 방식이다.
- [0051] 일 예로 기존의 자기 지도 학습 방식에서는 획득된 영상에서 임의의 영역을 마스킹하여 여백으로 치환하여 치환된 영상의 여백을 채우도록 학습시킨 후, 학습된 네트워크를 객체 인식이나 객체 탐지에 이용하는 경우가 있다. 다른 예로는 영상을 임의의 각도로 회전시킨 후, 회전 각도를 추정하도록 학습시킨 후, 학습된 네트워크를 객체 인식이나 객체 탐지에 이용하는 경우 등이 있다. 이러한 자기 지도 학습은 객체 인식이나 객체 탐지와 무관하게 여백의 픽셀값을 추정하거나 회전 각도를 추정하는 방식으로 학습되지만, 오히려 다양한 환경에서 다양한 형태로 획득된 입력 영상에서도 객체 인식이나 객체 탐지를 수행할 수 있는 성능을 나타낼 수도 있다.
- [0052] 이러한 자기 지도 학습 방식의 특성을 고려하여 본 실시예에서도 용이하게 획득 가능한 영상을 이용하여 특징 추출 네트워크(130)를 학습시키고, 학습된 특징 추출 네트워크(130)를 이용하여 보행자 검출부(120)에 의해 획득된 패치의 특징을 효과적으로 추출할 수 있도록 한다.
- [0053] 우선 기존의 특징 추출 네트워크(130)를 학습시키기 위한 학습 데이터에 요구되는 조건을 살펴보면 3가지로 요약될 수 있다. 먼저 기존의 특징 추출 네트워크(130)는 보행자 검출부(120)에 의해 획득된 패치를 인가받아 특징을 검출하는 구조이므로, 학습 데이터는 보행자가 포함된 경계 박스 형태의 패치이어야 한다. 두 번째로 시간적으로 연속된 다수의 프레임에서 추출된 동일 보행자에 대한 다수의 패치이어야 한다. 마지막 세번째는 동일 보행자에 대한 패치는 하나의 카메라에서 촬영된 영상의 프레임들에서 추출된 패치이어야 한다.
- [0054] 그에 반해 도 7을 참조하면, 본 실시예에서는 영상에서 별도로 추출된 패치가 아니라 개별 카메라 각각에서 획득되는 영상 자체를 식별자(ID 1, ID 2, ... ID N)로 구분되는 패치 그룹으로 고려하여 학습 데이터로 획득한다.

즉 개별 영상 각각을 서로 다른 식별자로 구분되고 확장된 크기를 갖는 패치 그룹으로 고려하여 학습 데이터로 이용한다.

[0055] 이때, 식별자로 구분되는 각 영상에는 추적 대상인 보행자의 포함될 수 있으나, 포함되지 않아도 무방하다. 즉 학습 데이터가 보행자가 포함된 영상에 한정되지 않고, 다양한 환경에서 획득된 영상을 학습 데이터로 이용할 수 있다.

[0056] 하나의 영상에서는 유사한 외형을 가진 연속하는 다수의 프레임으로 구성되며, 보행자가 포함된 패치를 별도로 추출하지 않아도 무방하므로 용이하게 대량의 데이터를 획득할 수 있다. 그리고 용이하게 획득 가능한 대량의 데이터를 학습 데이터를 이용하여 학습을 수행하는 경우, 특징 추출 네트워크(130)가 보행자뿐만 아니라 다양한 객체에 대한 식별 성능이 향상되어 보행자가 포함된 패치만으로 학습된 경우보다 오히려 우수한 특징 추출 성능을 나타낼 수 있다.

[0057] 여기서 연속하는 다수의 프레임으로 구성되는 영상을 학습 데이터로 이용하는 것은 특징 추출 네트워크(130)가 시간의 흐름에 따른 객체 또는 배경의 변화에도 강건하게 특징을 추출할 수 있도록 학습시키기 위해서이다.

[0058] 다만 영상의 길이가 길어지게 되면, 영상의 변화가 너무 크게 나타나게 되어 오히려 학습이 어려워질 수 있으므로, 각 카메라에서 획득된 영상에서 기지정된 길이(예를 들면 10초)에 대응하는 프레임만을 비디오 클립으로 추출하여 개별 식별자를 부여할 수 있다. 또한 영상의 변화가 크게 나타나지 않도록, 동일한 식별자가 부여되는 프레임은 동일한 카메라에서 획득된 영상의 프레임이어야 하며, 카메라의 회전이나 이동없이 특정 위치를 촬영하여 획득된 영상의 프레임이어야 한다. 즉 카메라의 상태 변경으로 인한 변화가 특징으로 추출되지 않아야 한다.

[0059] 도 7과 같이, 다수의 카메라 각각에서 획득된 기지정된 길이의 영상에 서로 다른 식별자를 부여하여 학습 데이터가 획득되면, 도 8에 도시된 바와 같이, 학습부(200)가 특징 추출 네트워크(130)를 학습시킬 수 있다.

[0060] 학습부(200)는 획득된 학습 데이터를 이용하여, 특징 추출 네트워크(130)와 함께 학습될 수 있다. 학습부(200)는 영상 분류부(210)와 손실 역전파부(220)를 포함할 수 있다.

[0061] 영상 분류부(210)는 특징 추출 네트워크(130)가 연속하는 다수의 프레임으로 구성된 기지정된 길이의 영상으로 획득된 학습 데이터를 인가받아 특징을 추출하면, 추출된 특징을 기반으로 인가된 학습 데이터의 식별자를 구분한다. 손실 역전파부(220)는 영상 분류부(210)에서 구분된 식별자와 학습 데이터의 식별자를 비교하여 기지정된 방식으로 손실을 계산하고, 계산된 손실을 역전파함으로써, 특징 추출 네트워크(130)와 영상 분류부(210)를 함께 학습시킨다.

[0062] 그리고 특징 추출 네트워크(130)에 대한 학습이 완료되면, 학습된 특징 추출 네트워크(130)를 도 6의 보행자 추적 장치에 적용한다. 여기서 학습된 특징 추출 네트워크(130)를 보행자 추적 장치에 적용하는 것은 학습된 특징 추출 네트워크(130)를 도 6과 같이 보행자 추적 장치에 배치하여 적용할 수도 있으나, 학습된 특징 추출 네트워크의 학습 가중치를 보행자 추적 장치에 배치된 특징 추출 네트워크(130)에 그대로 반영하여 적용할 수도 있다.

[0063] 비록 특징 추출 네트워크(130)가 자기 지도 학습 방식에 따라 보행자의 포함 여부에 무관한 영상을 학습 데이터로 이용하여 학습되지만, 대량의 학습 데이터를 기반으로 다양한 영상에 대한 특징을 정확하게 추출하도록 학습이 수행될 수 있으므로, 학습되어 보행자 추적 장치에 적용된 특징 추출 네트워크(130)는 보행자 검출부(120)에서 보행자가 포함된 패치로부터 각 보행자의 특징 또한 정확하게 구분하여 추출할 수 있다.

[0064] 즉 영상의 각 프레임에서 보행자가 포함된 영역을 검출하여 패치를 획득하고, 득된 패치 각각의 보행자를 식별하여 레이블링하는 작업없이, 다수의 카메라 각각에서 획득된 영상을 그대로 학습 데이터로 이용하여 특징 추출 네트워크(130)를 학습시킬 수 있다. 따라서 학습 데이터를 매우 용이하게 획득할 수 있을 뿐만 아니라, 대량의 학습 데이터를 기반으로 특징 추출 네트워크(130)를 학습시켜 매우 정확하게 보행자들에 대한 특징을 추출할 수 있도록 할 수 있다.

[0065] 그리고 특징 추출 네트워크(130)의 특징 추출 성능 향상은 추적부(140)가 영상에 포함된 각 보행자를 더욱 용이하게 추적할 수 있도록 하므로, 결과적으로 보행자 추적 장치의 성능을 향상시킬 수 있다.

[0066] 도 9는 본 발명의 일 실시예에 따른 보행자 추적 방법을 나타낸다.

[0067] 도 6 내지 도 8을 참조하여, 도 9의 보행자 추적 방법을 설명하면, 우선 패치를 인가받아 특징을 추출하여 패치

특징맵을 획득하는 인공 신경망인 특징 추출 네트워크(130)를 학습시킨다(S10).

- [0068] 특징 추출 네트워크(130)를 학습시키는 단계에서는 우선 학습용 영상을 획득한다(S11). 여기서 학습용 영상은 특정 조건에 따라 별도로 획득되는 영상이 아니라 다수의 카메라에서 획득된 일반적인 영상을 의미한다. 다만 학습용 영상은 기지정된 시간 길이의 연속되는 프레임으로 구성되는 영상일 수 있다.
- [0069] 그리고 획득된 학습용 영상 각각에 대해 식별자를 부여하여 학습 데이터를 획득한다(S12). 여기서 식별자는 획득된 각 영상을 구분하기 위해 부여되는 식별자로서, 영상을 구분할 수 있는 형태이면 어떠한 형식이든 무관하다. 일 예로 획득된 영상의 파일명이나 획득 시간 등이 식별자로 이용될 수도 있다. 또는 각 영상에 대해 별도로 부여되는 숫자나 문자 형식의 식별자일 수도 있다.
- [0070] 따라서 본 실시예에서는 단지 식별자로 구분 가능한 영상을 학습 데이터로 이용하므로, 영상에서 보행자를 검출하고, 검출된 보행자 영역에 해당하는 경계 박스를 설정하여 패치를 획득한 후, 획득된 패치에 포함된 보행자별로 식별자를 구분하여 레이블링하여 획득되는 기존의 학습 데이터에 비해 매우 용이하게 획득될 수 있다. 만일 영상의 파일명 등을 식별자로 이용하는 경우, 획득된 영상에 추가적인 작업을 수행하지 않고 그대로 학습데이터로 이용할 수도 있다.
- [0071] 학습 데이터가 획득되면, 획득된 학습 데이터를 학습 대상이 되는 특징 추출 네트워크(130)에 인가하고, 특징 추출 네트워크(130)는 학습되는 방식에 따라 인가된 학습 데이터에 포함된 다수의 프레임 각각의 특징을 추출하여 학습 특징맵을 획득한다(S13).
- [0072] 그리고 인공 신경망으로 구현되는 영상 분류부(210)는 특징 추출 네트워크(130)에서 획득된 학습 특징맵을 인가받아 학습되는 방식에 따라 분류하여 출력한다(S14). 영상 분류부(210)가 학습 특징맵을 분류하면, 손실 역전과부(220)가 영상 분류부(210)에서 구분된 식별자와 학습 데이터의 식별자를 비교하여 기지정된 방식으로 손실을 계산하고, 계산된 손실을 역전과함으로써, 특징 추출 네트워크(130)와 영상 분류부(210)를 함께 학습시킨다(S15).
- [0073] 그리고 학습 종료 여부를 판별한다(S16). 여기서 학습 종료는 특징 추출 네트워크(130)와 영상 분류부(210)가 기지정된 횟수 이상 반복 학습되거나, 계산된 손실이 기지정된 문턱값 이하이면 학습을 종료하는 것으로 판별할 수 있다.
- [0074] 만일 학습을 종료하지 않는 것으로 판단되면, 다시 학습 데이터를 인가받아 학습 데이터의 특징을 추출한다(S13). 그러나 학습을 종료하는 것으로 판단되면, 학습된 특징 추출 네트워크(130)를 보행자 추적 장치에 적용한다(S16).
- [0075] 학습 단계에서 학습된 특징 추출 네트워크(130)가 보행자 추적 장치에 적용되면, 보행자 추적을 수행한다(S20).
- [0076] 보행자 추적 단계에서는 우선 학습된 특징 추출 네트워크(130)가 적용된 보행자 추적 장치는 보행자를 검출해야 하는 입력 영상을 획득한다(S21). 여기서 영상은 CCTV나 카메라 등의 영상 획득 수단을 이용하여 획득되거나, 미리 획득되어 저장된 영상일 수 있으며, 연속하여 획득되는 다수 프레임 영상일 수 있다.
- [0077] 그리고 미리 학습된 방식에 따라 획득된 입력 영상의 각 프레임에서 보행자를 검출하고, 보행자가 검출된 영역의 경계 박스의 위치 및 크기를 설정한다(S22). 여기서 경계 박스는 영상에 포함된 보행자 각각에 대해 설정될 수 있다. 경계 박스가 설정되면, 입력 영상의 각 프레임에서 경계 박스에 대응하는 영역을 추출하여 패치를 획득한다(S23).
- [0078] 이후, 학습 단계에서 일반 영상을 학습 데이터로 이용하여 학습된 특징 추출 네트워크(130)가 획득된 패치 각각에 대해 미리 학습된 방식에 따라 보행자의 특징을 추출하여 패치 특징맵을 획득한다(S24).
- [0079] 입력 영상의 각 프레임에서 추출된 패치에 대한 패치 특징맵이 획득되면, 이전 기지정된 기간 동안의 프레임에서 추출된 패치에 대한 패치 특징맵과의 유사도를 비교하여 서로 대응하는 패치 특징맵을 탐색한다(S25). 그리고 탐색된 패치 특징맵 각각의 대응하는 프레임에서의 위치를 확인하여 각 프레임에서의 위치를 판별함으로써 보행자를 추적한다(S26).
- [0080] 본 발명에 따른 방법은 컴퓨터에서 실행시키기 위한 매체에 저장된 컴퓨터 프로그램으로 구현될 수 있다. 여기서 컴퓨터 판독가능 매체는 컴퓨터에 의해 액세스될 수 있는 임의의 가용 매체일 수 있고, 또한 컴퓨터 저장 매체를 모두 포함할 수 있다. 컴퓨터 저장 매체는 컴퓨터 판독가능 명령어, 데이터 구조, 프로그램 모듈 또는 기타 데이터와 같은 정보의 저장을 위한 임의의 방법 또는 기술로 구현된 휘발성 및 비휘발성, 분리형 및 비분

리형 매체를 모두 포함하며, ROM(판독 전용 메모리), RAM(랜덤 액세스 메모리), CD(컴팩트 디스크)-ROM, DVD(디지털 비디오 디스크)-ROM, 자기 테이프, 플로피 디스크, 광데이터 저장장치 등을 포함할 수 있다.

[0081] 본 발명은 도면에 도시된 실시예를 참고로 설명되었으나 이는 예시적인 것에 불과하며, 본 기술 분야의 통상의 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 타 실시예가 가능하다는 점을 이해할 것이다.

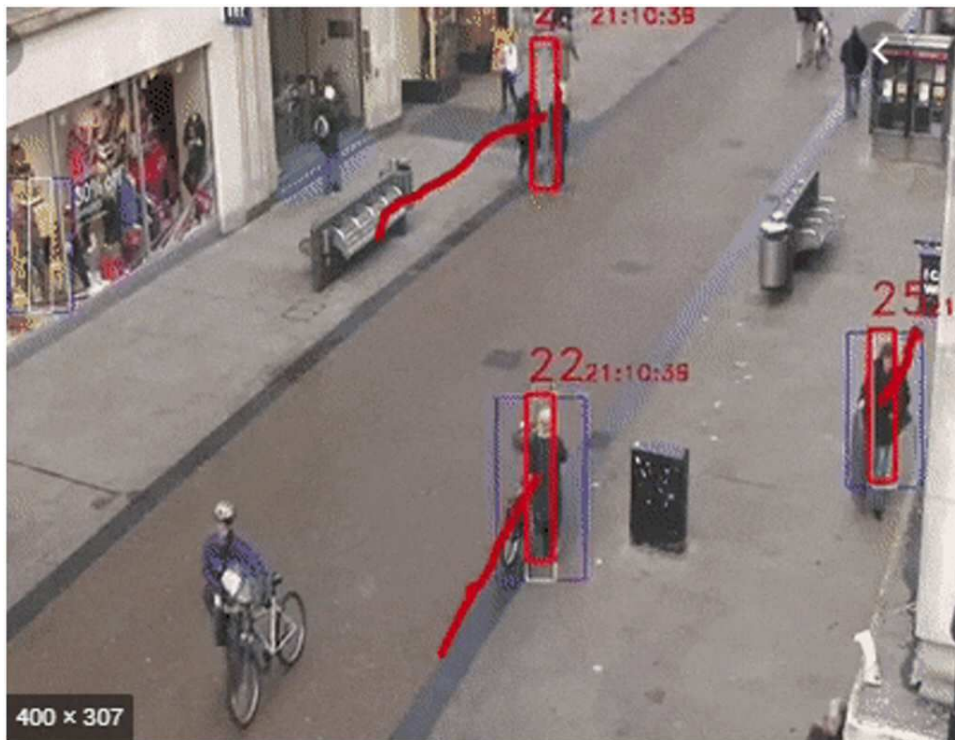
[0082] 따라서, 본 발명의 진정한 기술적 보호 범위는 첨부된 청구범위의 기술적 사상에 의해 정해져야 할 것이다.

부호의 설명

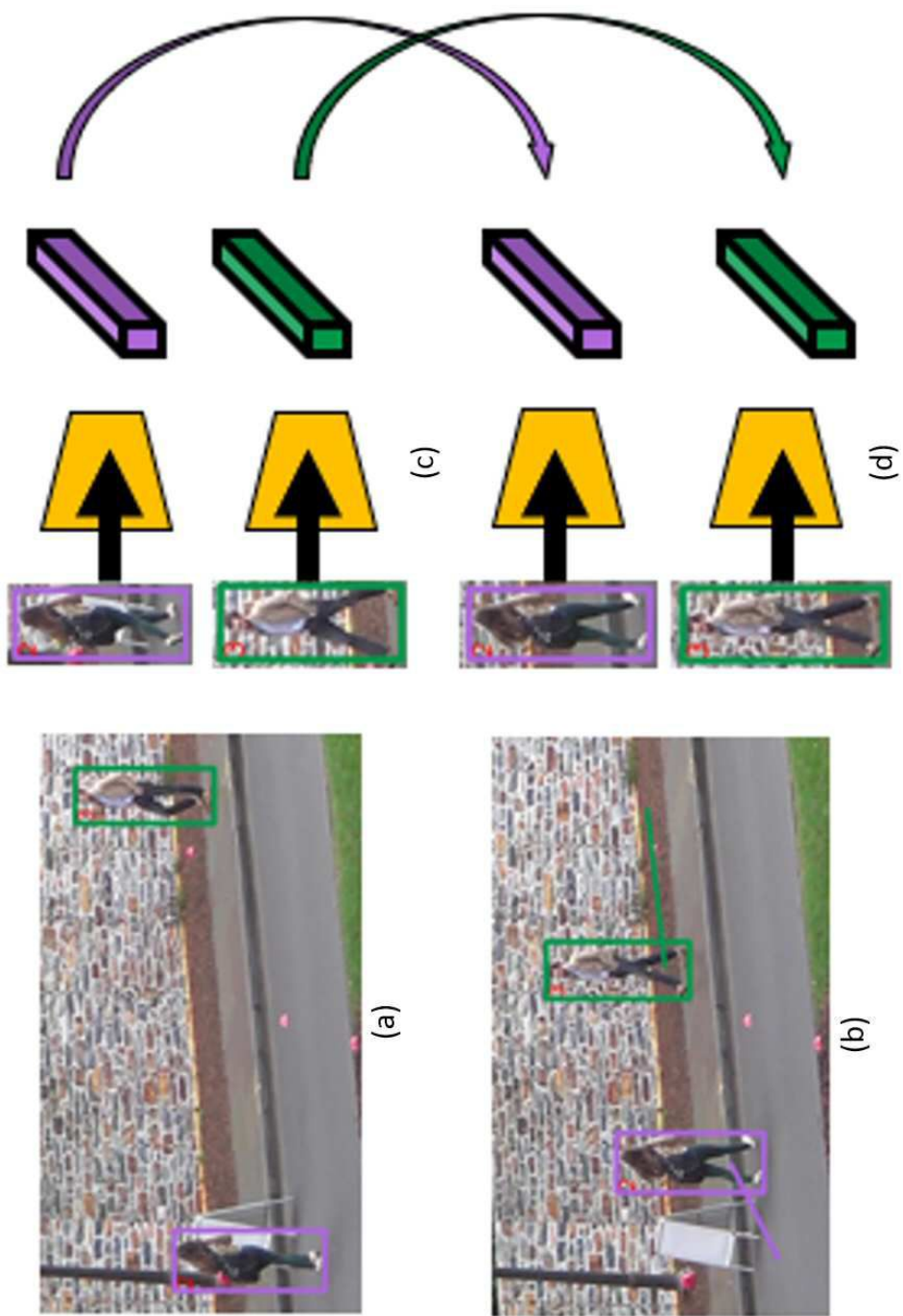
- [0083]
- | | |
|-----------------|--------------|
| 110: 영상 획득부 | 120: 보행자 검출부 |
| 121: 경계 박스 설정부 | 122: 패치 획득부 |
| 130: 특징 추출 네트워크 | 140: 추적부 |
| 200: 학습부 | 210: 영상 분류부 |
| 220: 손실 역전파부 | |

도면

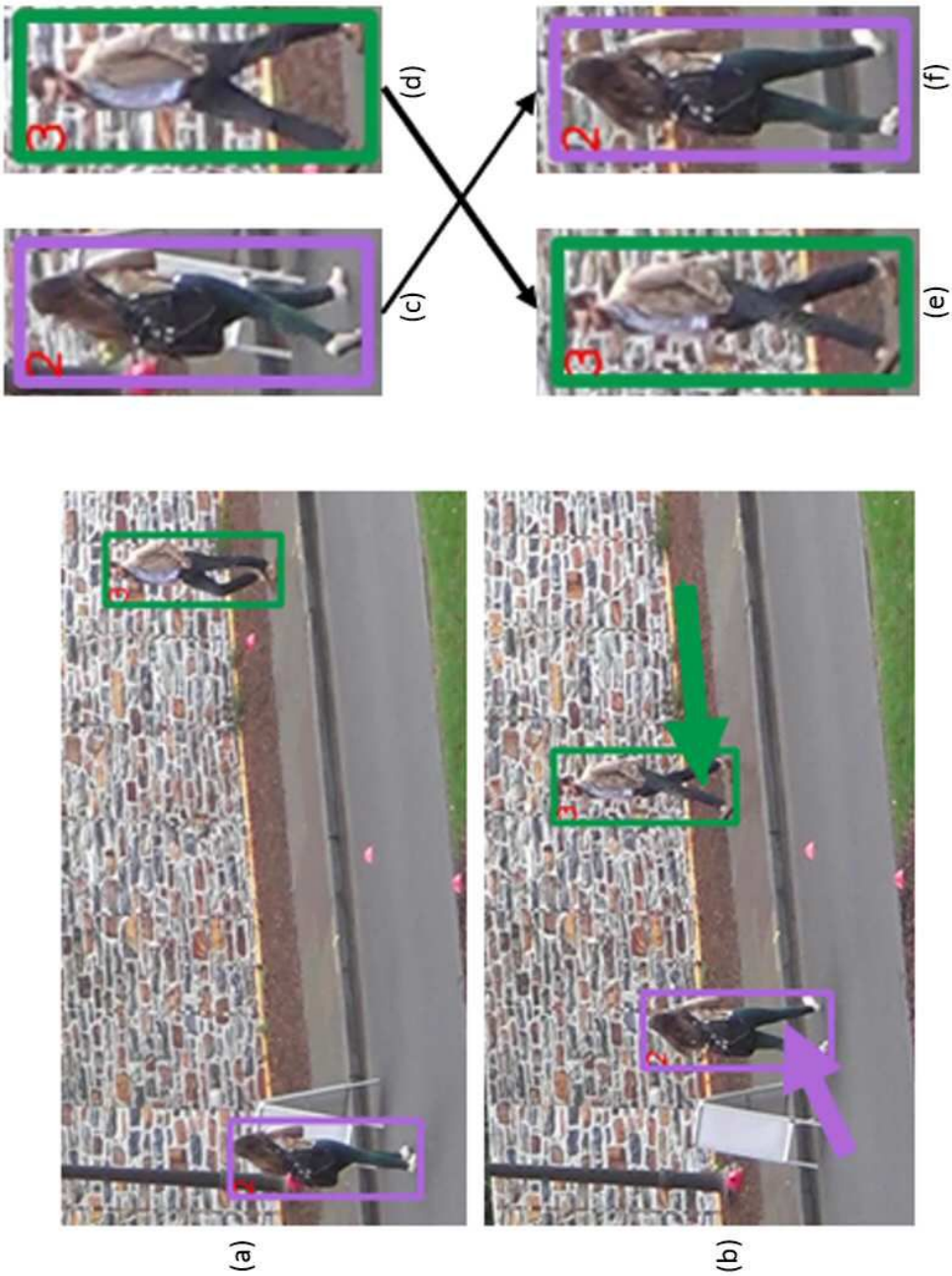
도면1



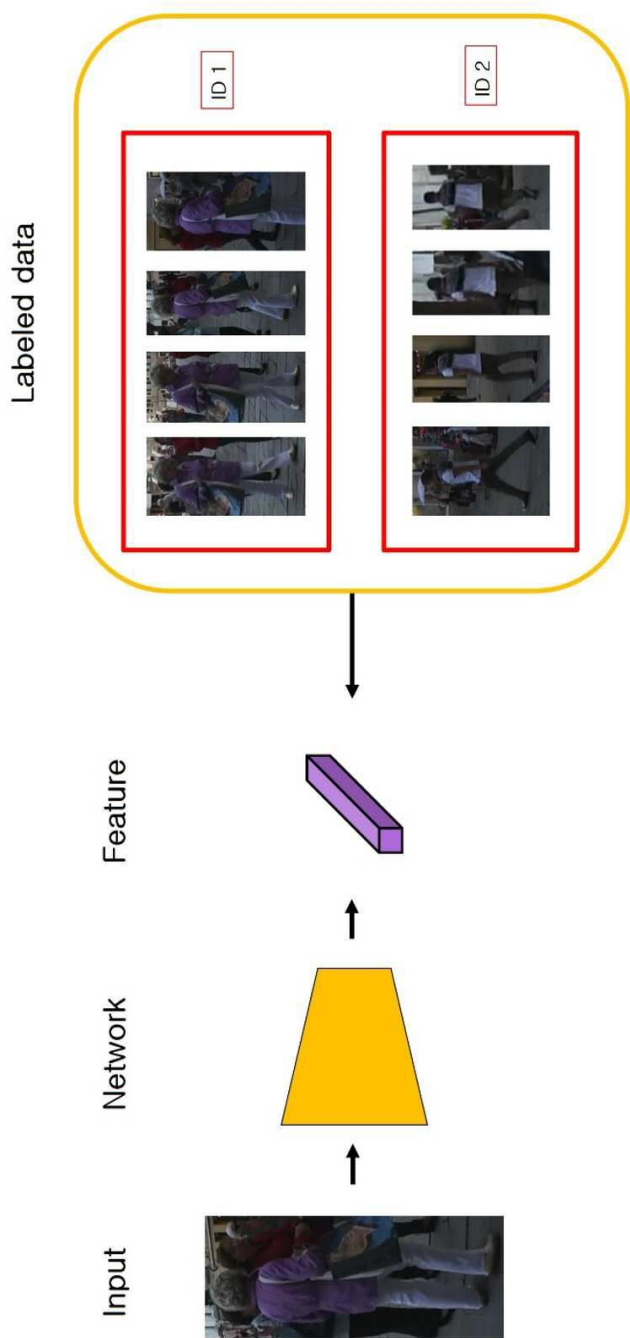
도면2



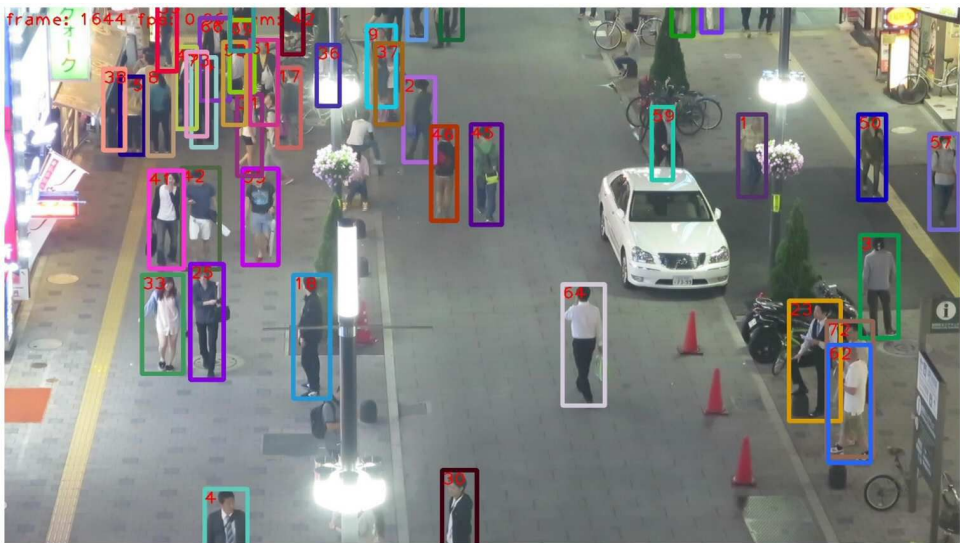
도면3



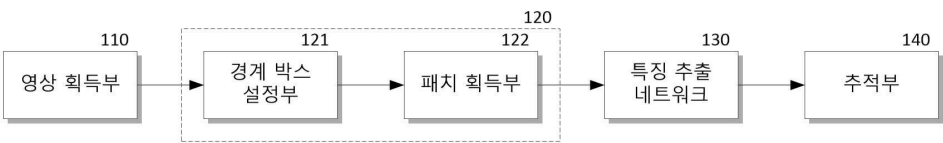
도면4



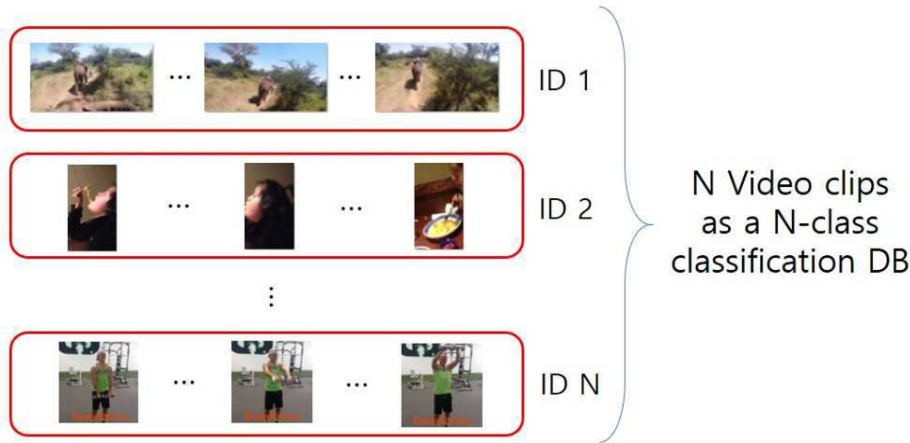
도면5



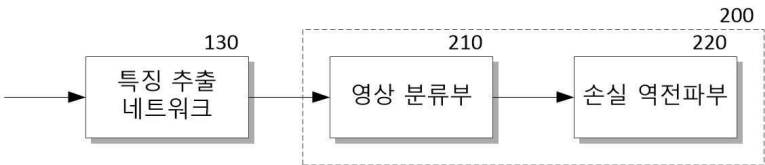
도면6



도면7



도면8



도면9

