

(19) 대한민국특허청(KR)  
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0063544

(43) 공개일자 2022년05월17일

(51) 국제특허분류(Int. Cl.)

G06T 7/20 (2017.01) G06N 3/08 (2006.01)

G06T 7/70 (2017.01)

(52) CPC특허분류

G06T 7/20 (2013.01)

G06N 3/08 (2013.01)

(21) 출원번호 10-2020-0149533

(22) 출원일자 2020년11월10일

심사청구일자 2020년11월10일

(71) 출원인

연세대학교 산학협력단

서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)

(72) 발명자

김은태

서울특별시 서대문구 연세로 50, 제3공학관 C607호(신촌동, 연세대학교)

이영조

서울특별시 서대문구 연세로 50, 제3공학관 C607호(신촌동, 연세대학교)

(뒷면에 계속)

(74) 대리인

특허법인우인

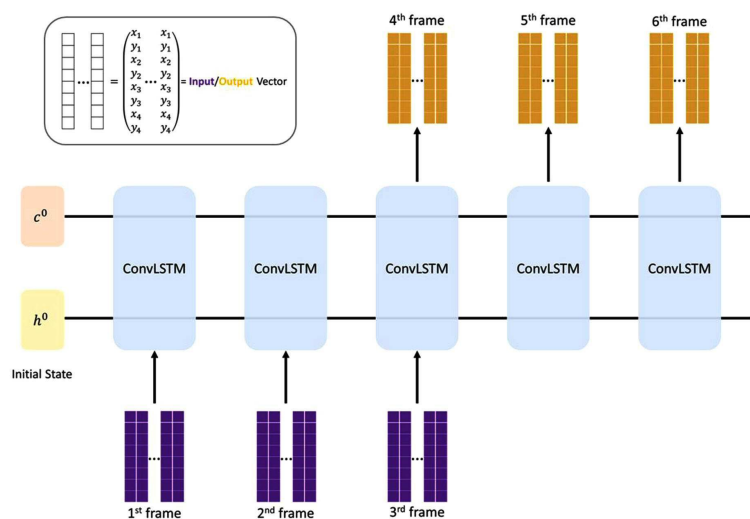
전체 청구항 수 : 총 4 항

(54) 발명의 명칭 다중 객체의 박스-레벨 위치 정보만을 이용한 박스-레벨 다중 객체 동선 예측 장치

## (57) 요약

본 발명의 바람직한 실시예에 따른 다중 객체의 박스-레벨 위치 정보만을 이용한 박스-레벨 다중 객체 동선 예측 장치는, 장단기 기억 신경망(long short term memory, LSTM) 구조를 통해 다중 객체들의 박스-레벨 위치 정보만을 입력으로 받아 미래의 다중 객체 동선을 박스-레벨로 예측함으로써, 가벼운 연산을 통해 다중 객체의 움직임을 빠르게 예측할 수 있고, 또한, 다중 객체의 박스-레벨 위치 정보만을 이용하여 다중 객체의 위치를 예측함으로써, 움직임을 예측할 객체가 다른 환경들에 의해 가려진 상황이나 객체들끼리 겹치는 상황에서도 다중 객체의 움직임을 예측할 수 있다.

대표도 - 도3



(52) CPC특허분류

**G06T 7/70** (2017.01)  
*G06T 2207/20081* (2013.01)  
*G06T 2207/20084* (2013.01)  
*G06T 2210/12* (2013.01)

**현준혁**

서울특별시 서대문구 연세로 50, 제3공학관 C607  
 호(신촌동, 연세대학교)

(72) 발명자

**성홍제**

서울특별시 서대문구 연세로 50, 제3공학관 C607  
 호(신촌동, 연세대학교)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711116308
과제번호	2017M3C4A706937022
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	원천기술개발사업
연구과제명	딥러닝 기반 의미론적 상황 이해 원천기술 연구 (2단계)(2/2)
기 여 율	1/1
과제수행기관명	연세대학교 산학협력단
연구기간	2020.04.01 ~ 2020.12.31

---

## 명세서

### 청구범위

#### 청구항 1

학습 데이터의 다중 객체에 대한 박스-레벨 위치 정보만을 이용하여, 장단기 기억 신경망(LSTM)으로 이루어지는 다중 객체 위치 예측 모델을 학습하는 학습부; 및

미리 설정된 개수의 연속된 과거 프레임들에 대한 다중 객체의 박스-레벨 위치 정보들을 상기 다중 객체 위치 예측 모델에 입력하고, 상기 다중 객체 위치 예측 모델로부터 출력되는 미리 설정된 개수의 연속된 미래 프레임들에 대한 다중 객체의 박스-레벨 위치 정보들을 기반으로, 다중 객체의 미래의 동선을 예측하는 예측부;

를 포함하는 다중 객체의 박스-레벨 위치 정보만을 이용한 박스-레벨 다중 객체 동선 예측 장치.

#### 청구항 2

제1항에서,

상기 박스-레벨 객체 위치 정보는,

객체를 둘러싸는 바운딩 박스(bounding box)의 네 모서리의 좌표값으로 이루어지는,

다중 객체의 박스-레벨 위치 정보만을 이용한 박스-레벨 다중 객체 동선 예측 장치.

#### 청구항 3

제2항에서,

상기 학습부는,

상기 학습 데이터에 포함된 학습 동영상에 대하여, 프레임에 대한 다중 객체의 박스-레벨 위치 정보들을 이용하여 다중 객체의 위치를 미리 설정된 최대 객체 개수x8 형태의 행렬에 채우고, 행렬의 나머지 부분을 0으로 패딩하여 프레임에 대한 행렬을 획득하고, 획득한 행렬을 이용하여 상기 다중 객체 위치 예측 모델을 학습하는,

다중 객체의 박스-레벨 위치 정보만을 이용한 박스-레벨 다중 객체 동선 예측 장치.

#### 청구항 4

제2항에서,

상기 예측부는,

미리 설정된 개수의 연속된 과거 프레임들 각각에 대하여, 과거 프레임에 대한 다중 객체의 박스-레벨 위치 정보들을 이용하여 다중 객체의 위치를 미리 설정된 최대 객체 개수x8 형태의 행렬에 채우고, 행렬의 나머지 부분을 0으로 패딩하여 과거 프레임에 대한 행렬을 획득하고, 미리 설정된 개수의 연속된 과거 프레임들 각각에 대응되는 복수개의 행렬을 상기 다중 객체 위치 예측 모델에 입력하고, 상기 다중 객체 위치 예측 모델로부터 출력되는 복수개의 행렬을 이용하여 미리 설정된 개수의 연속된 미래 프레임들에 대한 다중 객체의 박스-레벨 위치 정보들을 획득하는,

다중 객체의 박스-레벨 위치 정보만을 이용한 박스-레벨 다중 객체 동선 예측 장치.

## 발명의 설명

### 기술 분야

본 발명은 다중 객체의 박스-레벨 위치 정보만을 이용한 박스-레벨 다중 객체 동선 예측 장치에 관한 것으로서, 더욱 상세하게는 다중 객체의 과거 위치를 기반으로 다중 객체의 미래 위치를 예측하는, 장치에 관한 것이다.

[0001]

## 배경 기술

- [0002] 다중 객체의 움직임을 분석하는 다중 객체 추적(Multiple Object Tracking, MOT)이나 경로 예측(Trajectory Prediction)을 하는 종래 기술들은 많은 실행 시간이 소요된다. 이는 움직임을 분석할 여러 객체들 사이의 관계뿐만 아니라 객체를 제외한 배경과 객체들 사이의 관계까지 고려하여 픽셀-레벨(pixel-level)별 움직임을 분석하기 때문이다. 이러한 다중 객체의 움직임을 분석하는 업무는 보통 자율 주행 자동차나 로봇과 같은 기계 장치의 자율적 운동에서 사고 예방을 위한 장치에 많이 쓰이고 실시간에 가까울수록 기계 장치의 운동이 자연스럽게 더 빨라 질 수 있다.
- [0003] 최근 기계 장치들의 자율적 움직임 변화를 살펴보면, 먼저, 카메라를 통해 주변을 둘러본다. 그후, 장치 내의 알고리즘들을 통해 주변의 변화를 예측한다. 예측된 변화에서 충돌과 같은 문제 상황이 발생할 수 있다면 최대한 발생하지 않는 방향으로 기계 장치에게 동작 명령을 내릴 것이고, 이 명령에 따라 기계 장치는 움직일 것이다. 이때, 주변을 둘러보는 카메라는 대부분 1초에 60프레임을 찍을 수 있을 정도로 빠르다. 동작 명령 또한 대부분 전기적 신호로 움직이기 때문에 사고가 예상되는 상황에서 빠르게 기계 장치의 움직임을 바꿀 수 있다. 기계 장치의 움직임을 부자연스럽게 만들고 움직임의 속도를 낮추는 부분은 바로 주변의 변화 상황을 예측하는 알고리즘의 연산이다. 그렇기 때문에 알고리즘 자체의 연산이 가벼워지면 기계 장치의 상황에 대한 반응성이 빨라 질 수 있다.
- [0004] 또한, 종래의 기술들은 이미지를 기반으로 다중 객체들의 움직임을 분석하기 때문에 객체들끼리 가려진 상황에서 기존 객체를 그대로 따라가는데 어려움이 있다.

## 발명의 내용

### 해결하려는 과제

- [0005] 본 발명이 이루고자 하는 목적은, 시계열 데이터의 미래 예측에 많이 사용되는 장단기 기억 신경망(long short term memory, LSTM) 구조를 통해 다중 객체들의 박스-레벨 위치 정보만을 입력으로 받아 미래의 다중 객체 동선을 박스-레벨로 예측하는, 다중 객체의 박스-레벨 위치 정보만을 이용한 박스-레벨 다중 객체 동선 예측 장치를 제공하는 데 있다.
- [0006] 본 발명의 명시되지 않은 또 다른 목적들은 하기의 상세한 설명 및 그 효과로부터 용이하게 추론할 수 있는 범위 내에서 추가적으로 고려될 수 있다.

### 과제의 해결 수단

- [0007] 상기의 목적을 달성하기 위한 본 발명의 바람직한 실시예에 따른 다중 객체의 박스-레벨 위치 정보만을 이용한 박스-레벨 다중 객체 동선 예측 장치는, 학습 데이터의 다중 객체에 대한 박스-레벨 위치 정보만을 이용하여, 장단기 기억 신경망(LSTM)으로 이루어지는 다중 객체 위치 예측 모델을 학습하는 학습부; 및 미리 설정된 개수의 연속된 과거 프레임들에 대한 다중 객체의 박스-레벨 위치 정보들을 상기 다중 객체 위치 예측 모델에 입력하고, 상기 다중 객체 위치 예측 모델로부터 출력되는 미리 설정된 개수의 연속된 미래 프레임들에 대한 다중 객체의 박스-레벨 위치 정보들을 기반으로, 다중 객체의 미래의 동선을 예측하는 예측부;를 포함한다.
- [0008] 여기서, 상기 박스-레벨 객체 위치 정보는, 객체를 둘러싸는 바운딩 박스(bounding box)의 네 모서리의 좌표값으로 이루어질 수 있다.
- [0009] 여기서, 상기 학습부는, 상기 학습 데이터에 포함된 학습 동영상에 대하여, 프레임에 대한 다중 객체의 박스-레벨 위치 정보들을 이용하여 다중 객체의 위치를 미리 설정된 최대 객체 개수x8 형태의 행렬에 채우고, 행렬의 나머지 부분을 0으로 패딩하여 프레임에 대한 행렬을 획득하고, 획득한 행렬을 이용하여 상기 다중 객체 위치 예측 모델을 학습할 수 있다.
- [0010] 여기서, 상기 예측부는, 미리 설정된 개수의 연속된 과거 프레임들 각각에 대하여, 과거 프레임에 대한 다중 객체의 박스-레벨 위치 정보들을 이용하여 다중 객체의 위치를 미리 설정된 최대 객체 개수x8 형태의 행렬에 채우고, 행렬의 나머지 부분을 0으로 패딩하여 과거 프레임에 대한 행렬을 획득하고, 미리 설정된 개수의 연속된 과거 프레임들 각각에 대응되는 복수개의 행렬을 상기 다중 객체 위치 예측 모델에 입력하고, 상기 다중 객체 위치 예측 모델로부터 출력되는 복수개의 행렬을 이용하여 미리 설정된 개수의 연속된 미래 프레임들에 대한 다중 객체의 박스-레벨 위치 정보들을 획득할 수 있다.

## 발명의 효과

- [0011] 본 발명의 바람직한 실시예에 따른 다중 객체의 박스-레벨 위치 정보만을 이용한 박스-레벨 다중 객체 동선 예측 장치에 의하면, 장단기 기억 신경망(long short term memory, LSTM) 구조를 통해 다중 객체들의 박스-레벨 위치 정보만을 입력으로 받아 미래의 다중 객체 동선을 박스-레벨로 예측함으로써, 가벼운 연산을 통해 다중 객체의 움직임을 빠르게 예측할 수 있다.
- [0012] 또한, 다중 객체의 박스-레벨 위치 정보만을 이용하여 다중 객체의 위치를 예측함으로써, 움직임을 예측할 객체가 다른 환경들에 의해 가려진 상황이나 객체들끼리 겹치는 상황에서도 다중 객체의 움직임을 예측할 수 있다.
- [0013] 본 발명의 효과들은 이상에서 언급한 효과들로 제한되지 않으며, 언급되지 않은 또 다른 효과들은 아래의 기재로부터 통상의 기술자에게 명확하게 이해될 수 있을 것이다.

## 도면의 간단한 설명

- [0014] 도 1은 본 발명의 바람직한 실시예에 따른 다중 객체의 박스-레벨 위치 정보만을 이용한 박스-레벨 다중 객체 동선 예측 장치를 설명하기 위한 블록도이다.
- 도 2는 본 발명의 바람직한 실시예에 따른 박스-레벨 객체 위치를 설명하기 위한 도면으로, 도 2의 (a)는 종래의 바운딩 박스의 위치를 나타내는 방법이고, 도 2의 (b)는 본 발명에 따른 바운딩 박스의 위치를 나타내는 방법이다.
- 도 3은 본 발명의 바람직한 실시예에 따른 다중 객체 위치 예측 모델의 일례를 설명하기 위한 도면이다.
- 도 4는 본 발명의 바람직한 실시예에 따른 다중 객체의 박스-레벨 위치 정보만을 이용한 박스-레벨 다중 객체 동선 예측 장치의 성능을 설명하기 위한 도면이다.

## 발명을 실시하기 위한 구체적인 내용

- [0015] 이하, 첨부된 도면을 참조하여 본 발명의 실시 예를 상세히 설명한다. 본 발명의 이점 및 특징, 그리고 그것들을 달성하는 방법은 첨부되는 도면과 함께 상세하게 후술되어 있는 실시 예들을 참조하면 명확해질 것이다. 그러나 본 발명은 이하에서 개시되는 실시예들에 한정되는 것이 아니라 서로 다른 다양한 형태로 구현될 수 있으며, 단지 본 실시예들은 본 발명의 개시가 완전하도록 하고, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 발명의 범주를 완전하게 알려주기 위해 제공되는 것이며, 본 발명은 청구항의 범주에 의해 정의될 뿐이다. 명세서 전체에 걸쳐 동일 참조 부호는 동일 구성 요소를 지칭한다.
- [0016] 다른 정의가 없다면, 본 명세서에서 사용되는 모든 용어(기술 및 과학적 용어를 포함)는 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 공통적으로 이해될 수 있는 의미로 사용될 수 있을 것이다. 또 일반적으로 사용되는 사전에 정의되어 있는 용어들은 명백하게 특별히 정의되어 있지 않는 한 이상적으로 또는 과도하게 해석되지 않는다.
- [0017] 본 명세서에서 "제1", "제2" 등의 용어는 하나의 구성요소를 다른 구성요소로부터 구별하기 위한 것으로, 이들 용어들에 의해 권리범위가 한정되어서는 아니 된다. 예를 들어, 제1 구성요소는 제2 구성요소로 명명될 수 있고, 유사하게 제2 구성요소도 제1 구성요소로 명명될 수 있다.
- [0018] 본 명세서에서 각 단계들에 있어 식별부호(예를 들어, a, b, c 등)는 설명의 편의를 위하여 사용되는 것으로 식별부호는 각 단계들의 순서를 설명하는 것이 아니며, 각 단계들은 문맥상 명백하게 특정 순서를 기재하지 않는 이상 명기된 순서와 다르게 일어날 수 있다. 즉, 각 단계들은 명기된 순서와 동일하게 일어날 수도 있고 실질적으로 동시에 수행될 수도 있으며 반대의 순서대로 수행될 수도 있다.
- [0019] 본 명세서에서, "가진다", "가질 수 있다", "포함한다" 또는 "포함할 수 있다"등의 표현은 해당 특징(예: 수치, 기능, 동작, 또는 부품 등의 구성요소)의 존재를 가리키며, 추가적인 특징의 존재를 배제하지 않는다.
- [0020] 또한, 본 명세서에 기재된 '~부'라는 용어는 소프트웨어 또는 FPGA(field-programmable gate array) 또는 ASIC과 같은 하드웨어 구성요소를 의미하며, '~부'는 어떤 역할들을 수행한다. 그렇지만 '~부'는 소프트웨어 또는 하드웨어에 한정되는 의미는 아니다. '~부'는 어드레싱할 수 있는 저장 매체에 있도록 구성될 수도 있고 하나 또는 그 이상의 프로세서들을 재생시키도록 구성될 수도 있다. 따라서, 일 예로서 '~부'는 소프트웨어 구성요소들, 객체지향 소프트웨어 구성요소들, 클래스 구성요소들 및 태스크 구성요소들과 같은 구성요소들과, 프로세스

들, 함수들, 속성들, 프로시저들, 서브루틴들, 프로그램 코드의 세그먼트들, 드라이버들, 펌웨어, 마이크로코드, 회로, 데이터 구조들 및 변수들을 포함한다. 구성요소들과 '~부'들 안에서 제공되는 기능은 더 작은 수의 구성요소들 및 '~부'들로 결합되거나 추가적인 구성요소들과 '~부'들로 더 분리될 수 있다.

- [0022] 이하에서 첨부한 도면을 참조하여 본 발명에 따른 다중 객체의 박스-레벨 위치 정보만을 이용한 박스-레벨 다중 객체 동선 예측 장치의 바람직한 실시예에 대해 상세하게 설명한다.
- [0024] 먼저, 도 1 내지 도 3을 참조하여 본 발명의 바람직한 실시예에 따른 다중 객체의 박스-레벨 위치 정보만을 이용한 박스-레벨 다중 객체 동선 예측 장치에 대하여 설명한다.
- [0025] 도 1은 본 발명의 바람직한 실시예에 따른 다중 객체의 박스-레벨 위치 정보만을 이용한 박스-레벨 다중 객체 동선 예측 장치를 설명하기 위한 블록도이고, 도 2는 본 발명의 바람직한 실시예에 따른 박스-레벨 객체 위치를 설명하기 위한 도면으로, 도 2의 (a)는 종래의 바운딩 박스의 위치를 나타내는 방법이고, 도 2의 (b)는 본 발명에 따른 바운딩 박스의 위치를 나타내는 방법이며, 도 3은 본 발명의 바람직한 실시예에 따른 다중 객체 위치 예측 모델의 일례를 설명하기 위한 도면이다.
- [0026] 도 1을 참조하면, 본 발명의 바람직한 실시예에 따른 다중 객체의 박스-레벨 위치 정보만을 이용한 박스-레벨 다중 객체 동선 예측 장치(100)는 시계열 데이터의 미래 예측에 많이 사용되는 장단기 기억 신경망(long short term memory, LSTM) 구조를 통해 다중 객체들의 박스-레벨 위치 정보만을 입력으로 받아 미래의 다중 객체 동선을 박스-레벨로 예측한다.
- [0027] 여기서, 박스-레벨 객체 위치 정보는 객체를 둘러싸는 바운딩 박스(bounding box)의 네 모서리의 좌표값으로 이루어질 수 있다. 즉, 본 발명은 도 2의 (a)에 도시된 바와 같이, 객체를 둘러싸는 바운딩 박스의 중심 좌표(x, y), 높이(h) 및 너비(w)로 이루어지는 정보인 "(x, y, w, h)"를 박스-레벨 객체 위치로 이용하지 않고, 도 2의 (b)에 도시된 바와 같이, 객체를 둘러싸는 바운딩 박스의 네 모서리의 (x, y) 좌표값이 순서대로 이루어지는 정보인 "(x<sub>1</sub>, y<sub>1</sub>, x<sub>2</sub>, y<sub>2</sub>, x<sub>3</sub>, y<sub>3</sub>, x<sub>4</sub>, y<sub>4</sub>)"를 박스-레벨 객체 위치로 이용한다. 이에 따라, 축 방향으로 정렬된 경계 상자(Axis-Aligned Bounding Box, AABB)뿐만 아니라 방향성이 있는 경계 상자(Oriented Bounding Box, OBB)의 검출에도 사용될 수 있다.
- [0029] 이를 위해, 박스-레벨 다중 객체 동선 예측 장치(100)는 학습부(110) 및 예측부(130)를 포함할 수 있다.
- [0030] 학습부(110)는 학습 데이터의 다중 객체에 대한 박스-레벨 위치 정보만을 이용하여, 장단기 기억 신경망(LSTM)으로 이루어지는 다중 객체 위치 예측 모델을 학습한다.
- [0031] 이때, 학습부(110)는 학습 데이터에 포함된 학습 동영상에 대하여, 프레임에 대한 다중 객체의 박스-레벨 위치 정보들을 이용하여 다중 객체의 위치를 미리 설정된 최대 객체 개수x8 형태의 행렬에 채우고, 행렬의 나머지 부분을 0으로 패딩하여 프레임에 대한 행렬을 획득하고, 획득한 행렬을 이용하여 다중 객체 위치 예측 모델을 학습할 수 있다. 여기서, 미리 설정된 최대 객체 개수는 동영상의 프레임들 중에서 프레임 단위로 검출할 수 있는 객체의 최대 개수보다 큰 값을 가지며, 예컨대, 1,000개일 수 있다.
- [0032] 예컨대, 학습 동영상 내의 프레임별 객체의 개수는 일정하지 않기 때문에, 프레임별 최대 객체 개수(예컨대, 1,000개 등)를 미리 설정하고, 프레임에 대한 다중 객체의 박스-레벨 위치 정보들을 이용하여 미리 설정된 최대 객체 개수x8 형태의 행렬에 검출된 다중 객체의 해당 위치에 값을 채운 후, 행렬의 나머지 부분, 즉 (미리 설정된 최대 객체 개수 - 검출된 다중 객체 개수)x8 형태의 행렬에 0으로 패딩한다. 최종적으로 다중 객체 위치 예측 모델에 미리 설정된 최대 객체 개수x8의 행렬이 입력 벡터로 들어가고, 출력 벡터 역시 미리 설정된 최대 객체 개수x8의 행렬이 나오게 된다.
- [0034] 예측부(130)는 학습부(110)를 통해 학습된 다중 객체 위치 예측 모델을 통해, 입력된 동영상을 기반으로 다중 객체의 미래 위치를 예측한다.
- [0035] 즉, 예측부(130)는 미리 설정된 개수(예컨대, 3개 등)의 연속된 과거 프레임들에 대한 다중 객체의 박스-레벨

위치 정보들을 다중 객체 위치 예측 모델에 입력할 수 있다. 종래의 다중 객체 위치 예측 방법은 장단기 기억 신경망(LSTM)의 입력 벡터로 동영상 중 프레임 일부와 각각의 프레임에 대한 박스 정보(또는 박스 정보를 이용하여 원본 이미지에서 잘라낸 패치들)를 이용한다. 이에 반면, 본 발명은 입력 벡터로 이미지 대신 다중 객체에 대한 박스-레벨 위치 정보를 이용한다.

[0036] 그리고, 예측부(130)는 다중 객체 위치 예측 모델로부터 출력되는 미리 설정된 개수(예컨대, 3개 등)의 연속된 미래 프레임들에 대한 다중 객체의 박스-레벨 위치 정보들을 기반으로, 다중 객체의 미래의 동선을 예측할 수 있다.

[0037] 이때, 예측부(130)는 미리 설정된 개수의 연속된 과거 프레임들 각각에 대하여, 과거 프레임에 대한 다중 객체의 박스-레벨 위치 정보들을 이용하여 다중 객체의 위치를 미리 설정된 최대 객체 개수 $\times$ 8 형태의 행렬에 채우고, 행렬의 나머지 부분을 0으로 패딩하여 과거 프레임에 대한 행렬을 획득하고, 미리 설정된 개수의 연속된 과거 프레임들 각각에 대응되는 복수개의 행렬을 다중 객체 위치 예측 모델에 입력하고, 다중 객체 위치 예측 모델로부터 출력되는 복수개의 행렬을 이용하여 미리 설정된 개수의 연속된 미래 프레임들에 대한 다중 객체의 박스-레벨 위치 정보들을 획득할 수 있다. 여기서, 미리 설정된 최대 객체 개수는 동영상의 프레임들 중에서 프레임 단위로 검출할 수 있는 객체의 최대 개수보다 큰 값을 가지며, 예컨대, 1,000개일 수 있다.

[0038] 예컨대, 도 3에 도시된 바와 같이, 본 발명에 따른 다중 객체 위치 예측 모델은 입력 벡터와 출력 벡터로 이미지 대신에 다중 객체에 대한 박스-레벨 위치 정보들을 이용한다. 다중 객체 위치 예측 모델은 과거의 연속된 3개의 프레임에 대한 다중 객체의 박스-레벨 위치 정보들을 입력받고, 이후의 연속된 3개의 프레임에 대한 다중 객체의 박스-레벨 위치 정보들을 예측한다. 즉, 동영상 내의 프레임별 객체의 개수는 일정하지 않기 때문에, 프레임별 최대 객체 개수(예컨대, 1,000개 등)를 미리 설정하고, 프레임에 대한 다중 객체의 박스-레벨 위치 정보들을 이용하여 미리 설정된 최대 객체 개수 $\times$ 8 형태의 행렬에 검출된 다중 객체의 해당 위치에 값을 채운 후, 행렬의 나머지 부분, 즉 (미리 설정된 최대 객체 개수 - 검출된 다중 객체 개수) $\times$ 8 형태의 행렬에 0으로 패딩한다. 최종적으로 다중 객체 위치 예측 모델에 미리 설정된 최대 객체 개수 $\times$ 8의 행렬이 입력 벡터로 들어가고, 출력 벡터 역시 미리 설정된 최대 객체 개수 $\times$ 8의 행렬이 나오게 된다.

[0041] 그러면, 도 4를 참조하여 본 발명의 바람직한 실시예에 따른 다중 객체의 박스-레벨 위치 정보만을 이용한 박스-레벨 다중 객체 동선 예측 장치의 성능에 대하여 설명한다.

[0042] 도 4는 본 발명의 바람직한 실시예에 따른 다중 객체의 박스-레벨 위치 정보만을 이용한 박스-레벨 다중 객체 동선 예측 장치의 성능을 설명하기 위한 도면이다.

[0043] 본 발명에 따른 다중 객체의 박스-레벨 위치 정보만을 이용한 박스-레벨 다중 객체 동선 예측 장치의 성능을 실험하기 위해, 다중 객체들의 위치에 대한 박스 정보를 가지고 있는 MOT16 데이터 셋(Milan, A., et al. "MOT16: A benchmark for multi-object tracking. arXiv 2016." arXiv preprint arXiv:1603.00831 9. 참고)을 이용하여 실험을 진행하였다.

[0044] 도 4에 도시된 위의 세 열은 다중 객체 위치 예측 모델에 입력으로 들어간 프레임에 대한 다중 객체의 박스 정보를 모두 표현한 이미지이고, 도 4에 도시된 아래의 세 열은 다중 객체 위치 예측 모델로부터 출력으로 나온 프레임별 다중 객체의 박스 정보를 모두 표현한 이미지이다. 위아래 모두 왼쪽부터 시간순으로 표현되어 있다.

[0045] 정답 이미지는 예측된 박스 위치와 많이 중복되어 도 4에 도시하지 않았으나, 본 발명에 따라 예측된 박스의 위치가 정답 위치와 잘 맞는다는 것을 확인할 수 있었다. 또한, 해당 데이터는 많은 데이터 셋 중 가시성이 좋은 데이터 셋을 표현한 것인데 다른 영상의 경우 객체가 너무 많아서 박스를 도면으로 표시하지는 못하였지만, 이 경우에도 박스의 위치는 준수한 정확도를 보였으며 다른 객체에 의해 가려지는 상황에서도 꾸준한 정확도를 보이는 것을 확인할 수 있었다.

[0048] 아울러, 아래의 [표 1]은 입력으로 사용되는 프레임의 수를 조절하면서 그 이후 세 프레임에 대한 다중 객체의 움직임을 예측했을 때, 이후 세 프레임에 대한 예측에 걸리는 시간을 측정한 것이다.

표 1

방식(입력 프레임 개수)	소요 시간(ms)
2개	0.037
3개	0.040
4개	0.054
5개	0.076

[0049]

[0050]

일반적으로 30 fps를 실시간 동작으로 이야기한다. 즉, 실시간 예측을 위해선 세 프레임에 대해 적어도 33ms 안에 연산이 이뤄져야 한다는 것이다. [표 1]의 값들은 대략 이 값의 1/1000에 해당한다. 따라서, 본 발명은 더 빠른 속도로 카메라 입력이 들어와도 다 처리할 수 있다. 본 발명은 이미지를 직접적으로 이용하는 종래 방식들에 비해 정확도는 다소 떨어질 수는 있지만, 더 순간적인 대응을 할 수 있다는 점에서, 실질적으로 예측 기술이 응용되는데에는 더 큰 역할을 할 수 있다.

[0053]

이상에서 설명한 본 발명의 실시예를 구성하는 모든 구성요소들이 하나로 결합하거나 결합하여 동작하는 것으로 기재되어 있다고 해서, 본 발명이 반드시 이러한 실시예에 한정되는 것은 아니다. 즉, 본 발명의 목적 범위 안에서라면, 그 모든 구성요소들이 하나 이상으로 선택적으로 결합하여 동작할 수도 있다. 또한, 그 모든 구성요소들이 각각 하나의 독립적인 하드웨어로 구현될 수 있지만, 각 구성요소들의 그 일부 또는 전부가 선택적으로 조합되어 하나 또는 복수개의 하드웨어에서 조합된 일부 또는 전부의 기능을 수행하는 프로그램 모듈을 갖는 컴퓨터 프로그램으로서 구현될 수도 있다. 또한, 이와 같은 컴퓨터 프로그램은 USB 메모리, CD 디스크, 플래쉬 메모리 등과 같은 컴퓨터가 읽을 수 있는 기록 매체(Computer Readable Media)에 저장되어 컴퓨터에 의하여 읽혀지고 실행됨으로써, 본 발명의 실시예를 구현할 수 있다. 컴퓨터 프로그램의 기록 매체로서는 자기기록매체, 광 기록매체 등이 포함될 수 있다.

[0054]

이상의 설명은 본 발명의 기술 사상을 예시적으로 설명한 것에 불과한 것으로서, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자라면 본 발명의 본질적인 특성에서 벗어나지 않는 범위 내에서 다양한 수정, 변경 및 치환이 가능할 것이다. 따라서, 본 발명에 개시된 실시예 및 첨부된 도면들은 본 발명의 기술 사상을 한정하기 위한 것이 아니라 설명하기 위한 것이고, 이러한 실시예 및 첨부된 도면에 의하여 본 발명의 기술 사상의 범위가 한정되는 것은 아니다. 본 발명의 보호 범위는 아래의 청구범위에 의하여 해석되어야 하며, 그와 동등한 범위 내에 있는 모든 기술 사상은 본 발명의 권리범위에 포함되는 것으로 해석되어야 할 것이다.

### 부호의 설명

[0055]

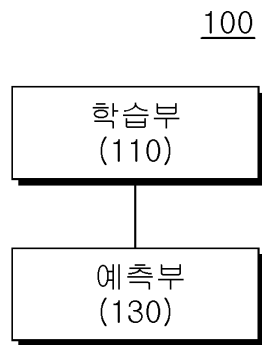
100 : 박스-레벨 다중 객체 동선 예측 장치,

110 : 학습부,

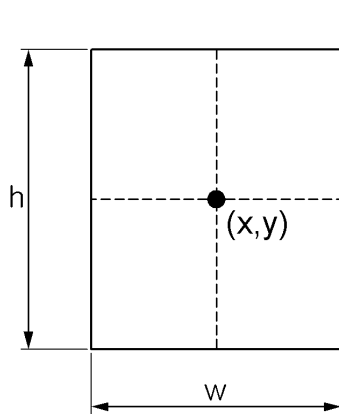
130 : 예측부

도면

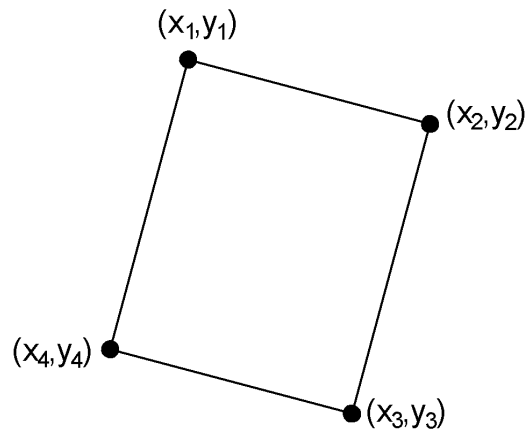
도면1



도면2

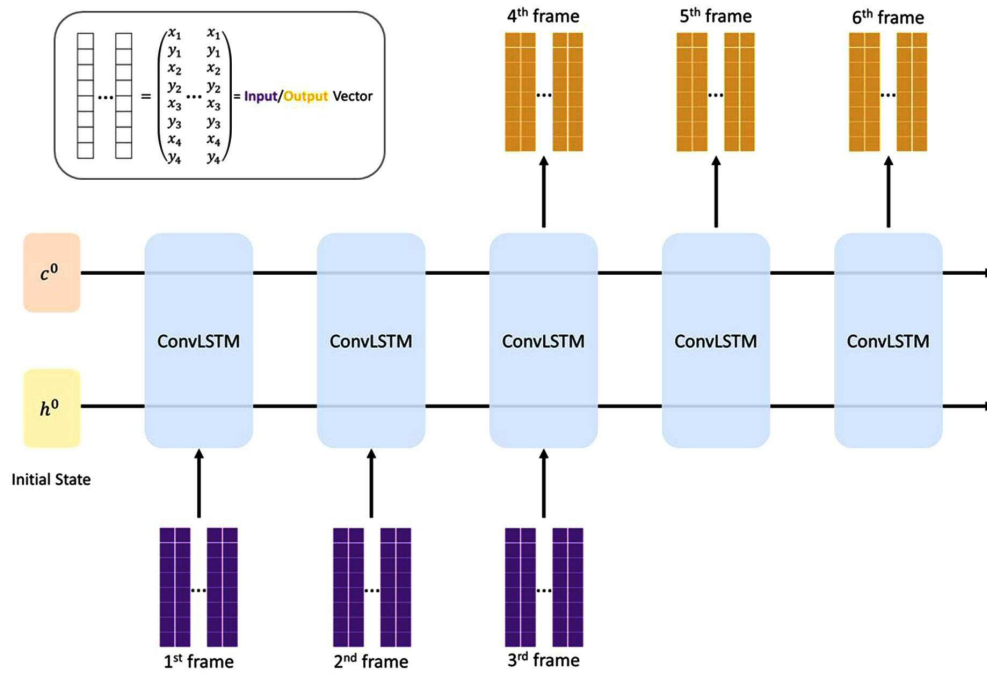


(a)  $(x, y, w, h)$



(b)  $(x_1, y_1, x_2, y_2, x_3, y_3, x_4, y_4)$

도면3



도면4

