



(19) 대한민국특허청(KR)  
(12) 공개특허공보(A)

(11) 공개번호 10-2020-0062686  
(43) 공개일자 2020년06월04일

(51) 국제특허분류(Int. Cl.)  
G06T 7/20 (2017.01) G06T 7/11 (2017.01)  
G06T 7/194 (2017.01)  
(52) CPC특허분류  
G06T 7/20 (2013.01)  
G06T 7/11 (2017.01)  
(21) 출원번호 10-2018-0148370  
(22) 출원일자 2018년11월27일  
심사청구일자 2018년11월27일

(71) 출원인  
연세대학교 산학협력단  
서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)  
(72) 발명자  
이상윤  
서울특별시 서초구 청두곶14길 6, 202호(방배동)  
김우진  
서울특별시 마포구 백범로1길 83, 505호(노고산동, 신촌다울노블리움)  
(74) 대리인  
민영준

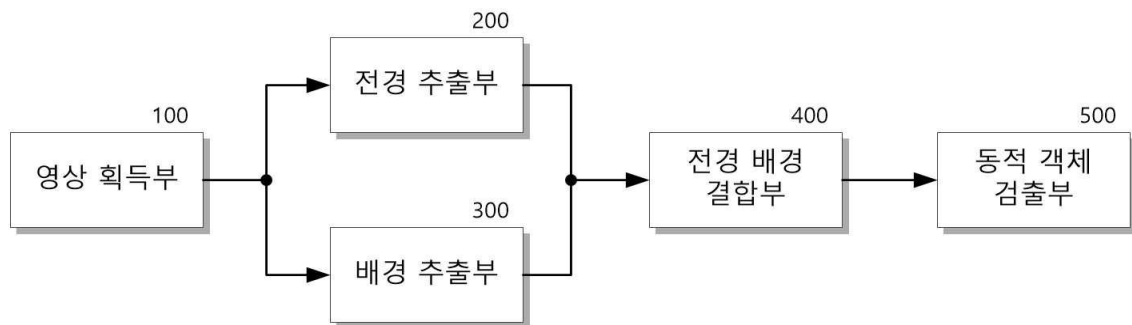
전체 청구항 수 : 총 10 항

(54) 발명의 명칭 동적 객체 검출 장치 및 방법

(57) 요약

본 발명은 단일 영상에서 연속하는 기지정된 개수의 다수 프레임을 인가받아 미리 학습된 패턴 추정 방식에 따라 각 프레임에 대해 독립적으로 인코딩하여 다수의 전경 특징 맵을 생성하며, 다수의 전경 특징 맵을 결합하여 디코딩하여 전경 영상을 생성하는 전경 추출부, 다수 프레임을 인가받아 미리 학습된 패턴 추정 방식에 따라 다수 프레임을 결합하고 인코딩하여 배경 특징 맵을 생성하며, 배경 특징 맵을 디코딩하여 배경 영상을 생성하는 배경 추출부, 전경 영상과 배경 영상을 결합하여, 동적 객체 추정 영상을 생성하는 전경 배경 결합부 및 동적 객체 추정 영상을 인가받고, 미리 학습된 패턴 추정 방식에 따라 동적 객체 영상을 생성하는 동적 객체 검출부를 포함하는 동적 객체 검출 장치 및 방법을 제공할 수 있다.

대표도 - 도1



(52) CPC특허분류

**G06T 7/194** (2017.01)

G06T 2207/20081 (2013.01)

(72) 발명자

**황상원**

서울특별시 강서구 공항대로45길 68-6, 106동 605호(등촌동, 서광등촌마을아파트)

**이준협**

서울특별시 서대문구 연희로10길 29-5, 205호(연희동)

이 발명을 지원한 국가연구개발사업

과제고유번호 1711036311

부처명 과학기술정보통신부

연구관리전문기관 정보통신기술진흥센터(NIPA산하)

연구사업명 정보통신방송연구개발사업

연구과제명 [이지바로]스마트카 다중 센서와 딥러닝을 이용한 초정밀 내추럴 3D 뷰 생성 기술 개발(창조씨앗형 2단계)(2/2단계)(2/5연차)

기 여 율 1/1

주관기관 연세대학교 산학협력단

연구기간 2018.01.01 ~ 2018.12.31

## 명세서

### 청구범위

#### 청구항 1

단일 영상에서 연속하는 기지정된 개수의 다수 프레임을 인가받아 미리 학습된 패턴 추정 방식에 따라 각 프레임에 대해 독립적으로 인코딩하여 다수의 전경 특징 맵을 생성하며, 상기 다수의 전경 특징 맵을 결합하여 디코딩하여 전경 영상을 생성하는 전경 추출부;

상기 다수 프레임을 인가받아 미리 학습된 패턴 추정 방식에 따라 상기 다수 프레임을 결합하고 인코딩하여 배경 특징 맵을 생성하며, 상기 배경 특징 맵을 디코딩하여 배경 영상을 생성하는 배경 추출부;

상기 전경 영상과 배경 영상을 결합하여, 동적 객체 추정 영상을 생성하는 전경 배경 결합부; 및

상기 동적 객체 추정 영상을 인가받고, 미리 학습된 패턴 추정 방식에 따라 동적 객체 영상을 생성하는 동적 객체 검출부;를 포함하는 동적 객체 검출 장치.

#### 청구항 2

제1 항에 있어서, 상기 전경 추출부는

상기 다수 프레임 중 대응하는 프레임을 인가받고, 미리 학습된 패턴 추정 방식에 따라 전경 특징을 추출하여, 각각 전경 특징 맵을 생성하는 다수의 전경 인코더;

상기 다수의 전경 특징 맵을 기지정된 방식으로 결합하여 통합 특징 맵을 생성하는 특징 결합부; 및

상기 통합 특징 맵을 인가받고, 미리 학습된 방식으로 상기 통합 특징 맵의 전경 특징을 디코딩하여 상기 전경 영상을 획득하는 전경 디코더;를 포함하는 동적 객체 검출 장치.

#### 청구항 3

제2 항에 있어서, 상기 배경 추출부는

상기 다수 프레임을 인가받아 결합하고, 미리 학습된 패턴 추정 방식에 따라 결합된 다수의 프레임을 인코딩하여 상기 배경 특징 맵을 생성하는 배경 인코더; 및

상기 배경 특징 맵의 배경 특징을 디코딩하여 상기 배경 영상을 생성하는 배경 디코더;를 포함하는 동적 객체 검출 장치.

#### 청구항 4

제3 항에 있어서, 상기 전경 추출부는

검증된 전경 영상을 갖는 학습 영상을 이용하여 학습되고,

상기 배경 추출부는

상기 학습 영상에 포함된 상기 전경 영상을 반전한 반전 전경 영상을 이용하여 학습된 동적 객체 검출 장치.

#### 청구항 5

제1 항에 있어서, 상기 전경 배경 결합부는

상기 배경 영상을 반전하고, 반전된 배경 영상을 상기 전경 영상과 논리곱하여 상기 동적 객체 추정 영상을 생성하는 동적 객체 검출 장치.

#### 청구항 6

제1 항에 있어서, 상기 동적 객체 검출부는

적대적 학습(Adversarial learning) 방식으로 학습된 GAN(Generative Adversarial Network)으로 구현되는 동적

객체 검출 장치.

#### 청구항 7

미리 학습된 패턴 추정 방식에 따라 단일 영상에서 연속하는 기지정된 개수의 다수 프레임 각각에 대해 독립적으로 인코딩하여 다수의 전경 특징 맵을 생성하며, 상기 다수의 전경 특징 맵을 결합하여 디코딩하여 전경 영상을 생성하는 단계;

미리 학습된 패턴 추정 방식에 따라 상기 다수 프레임을 결합하고 인코딩하여 배경 특징 맵을 생성하며, 상기 배경 특징 맵을 디코딩하여 배경 영상을 생성하는 단계;

상기 전경 영상과 배경 영상을 결합하여, 동적 객체 추정 영상을 생성하는 단계; 및

미리 학습된 패턴 추정 방식에 따라 상기 동적 객체 추정 영상으로부터 동적 객체 영상을 생성하는 단계; 를 포함하는 동적 객체 검출 방법.

#### 청구항 8

제7 항에 있어서, 상기 전경 영상을 생성하는 단계는

미리 학습된 패턴 추정 방식에 따라 상기 다수 프레임 중 대응하는 프레임 각각에서 전경 특징을 추출하여, 다수의 전경 특징 맵을 생성하는 단계;

상기 다수의 전경 특징 맵을 기지정된 방식으로 결합하여 통합 특징 맵을 생성하는 단계; 및

미리 학습된 방식으로 상기 통합 특징 맵의 전경 특징을 디코딩하여 상기 전경 영상을 획득하는 단계; 를 포함하는 동적 객체 검출 방법.

#### 청구항 9

제8 항에 있어서, 상기 배경 영상을 생성하는 단계는

미리 학습된 패턴 추정 방식에 따라 결합된 다수 프레임을 인코딩하여 상기 배경 특징 맵을 생성하는 단계; 및

상기 배경 특징 맵의 배경 특징을 디코딩하여 상기 배경 영상을 생성하는 단계; 를 포함하는 동적 객체 검출 방법.

#### 청구항 10

제7 항에 있어서, 상기 동적 객체 추정 영상을 생성하는 단계는

상기 배경 영상을 반전하고, 반전된 배경 영상을 상기 전경 영상과 논리곱하여 상기 동적 객체 추정 영상을 생성하는 동적 객체 검출 방법.

### 발명의 설명

#### 기술 분야

[0001] 본 발명은 동적 객체 검출 장치 및 방법에 관한 것으로, 비 고정식 카메라 환경에서 획득된 단일 영상으로부터 동적 객체를 검출할 수 있는 동적 객체 검출 장치 및 방법에 관한 것이다.

#### 배경 기술

[0002] 비 고정식 카메라 환경 하에서의 동적 객체 검출 기술은 카메라의 움직임에 따라 배경이 변화하기 때문에 실제로 움직이는 물체를 검출하기 매우 어렵다.

[0003] 기존에는 카메라에서 획득되는 영상의 평균, 분산을 이용한 가우시안 모델을 배경 모델로 사용하고 영상 간의 투영변환(homography)을 계산하여 물체의 이동 방향을 예측하여 동적 객체 검출하였다.

[0004] 그러나 배경을 여러 모델로 가정하는 경우, 모델의 의존성이 높기 때문에 날씨 변화나 카메라의 움직임이 클 경우 검출 결과가 부정확하게 된다는 문제가 있다. 특히 영상의 평균과 분산을 이용하여 배경을 가우시안 모델로 가정할 경우, 조도변화나 전경 픽셀의 잡음에 의해 배경모델이 손상되어 동적 객체의 정확한 검출이 어렵다.

또한 영상 간의 투영변환을 계산하는 방법은 2차원 평면 환경(2D planar scene)에서만 계산 될 수 있다는 한계로 인해 실제 상황인 3차원 환경에서의 투영변환 계산 불가능하다.

[0005] 이에 최근에는 딥 러닝 네트워크(deep learning network)를 이용하여 단일 영상으로부터 이미지와 배경모델을 비교하여, 움직임이 있는 영역을 검출하는 방법이 제안되었으나, 카메라의 움직임이나 날씨변화가 역동적인(dynamic)환경일때 배경 모델링이 불확실해지면서 정적 객체와 동적 객체의 구별 검출이 어려워지고, 눈이나 비와 같은 날씨변화로 인한 가림현상(Occlusion)이 일어날 시 관심 물체의 영역이 순간적으로 가려지기 때문에 부정확하게 구별되는 현상 발생된다.

[0006] 한편, 딥 러닝 네트워크에 옵티컬 플로우(optical flow)를 적용하는 방안도 제안된 바 있으나, 옵티컬 플로우는 대량의 연산을 요구할 뿐만 아니라, 검출되는 객체 영역이 실제보다 크게 예측되므로, 객체를 정확하게 검출하지 못한다는 한계가 있다.

## 선행기술문헌

### 특허문헌

[0007] (특허문헌 0001) 한국 등록 특허 제10-1675492호 (2016.11.07 등록)

## 발명의 내용

### 해결하려는 과제

[0008] 본 발명의 목적은 단일 영상으로부터 전경과 배경을 각각 획득하고, 획득된 전경과 배경으로부터 동적 객체를 정확하게 검출할 수 있는 동적 객체 검출 장치 및 방법을 제공하는데 있다.

[0009] 본 발명의 다른 목적은 배경, 조도 및 날씨 변화에 강인하고 카메라의 움직임이 크거나 객체가 빠르게 움직일 때도 동적 객체를 정확하게 검출할 수 있는 동적 객체 검출 장치 및 방법을 제공하는데 있다.

### 과제의 해결 수단

[0010] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른 동적 객체 검출 장치는 단일 영상에서 연속하는 기지정된 개수의 다수 프레임을 인가받아 미리 학습된 패턴 추정 방식에 따라 각 프레임에 대해 독립적으로 인코딩하여 다수의 전경 특징 맵을 생성하며, 상기 다수의 전경 특징 맵을 결합하여 디코딩하여 전경 영상을 생성하는 전경 추출부; 상기 다수 프레임을 인가받아 미리 학습된 패턴 추정 방식에 따라 상기 다수 프레임을 결합하고 인코딩하여 배경 특징 맵을 생성하며, 상기 배경 특징 맵을 디코딩하여 배경 영상을 생성하는 배경 추출부; 상기 전경 영상과 배경 영상을 결합하여, 동적 객체 추정 영상을 생성하는 전경 배경 결합부; 및 상기 동적 객체 추정 영상을 인가받고, 미리 학습된 패턴 추정 방식에 따라 동적 객체 영상을 생성하는 동적 객체 검출부;를 포함한다.

[0011] 상기 전경 추출부는 상기 다수 프레임 중 대응하는 프레임을 인가받고, 미리 학습된 패턴 추정 방식에 따라 전경 특징을 추출하여, 각각 전경 특징 맵을 생성하는 다수의 전경 인코더; 상기 다수의 전경 특징 맵을 기지정된 방식으로 결합하여 통합 특징 맵을 생성하는 특징 결합부; 및 상기 통합 특징 맵을 인가받고, 미리 학습된 방식으로 상기 통합 특징 맵의 전경 특징을 디코딩하여 상기 전경 영상을 획득하는 전경 디코더;를 포함할 수 있다.

[0012] 상기 배경 추출부는 상기 다수 프레임을 인가받아 결합하고, 미리 학습된 패턴 추정 방식에 따라 결합된 다수의 프레임을 인코딩하여 상기 배경 특징 맵을 생성하는 배경 인코더; 및 상기 배경 특징 맵의 배경 특징을 디코딩하여 상기 배경 영상을 생성하는 배경 디코더;를 포함할 수 있다.

[0013] 상기 전경 추출부는 검증된 전경 영상을 갖는 학습 영상을 이용하여 학습되고, 상기 배경 추출부는 상기 학습 영상에 포함된 상기 전경 영상을 반전한 반전 전경 영상을 이용하여 학습될 수 있다.

[0014] 상기 전경 배경 결합부는 상기 배경 영상을 반전하고, 반전된 배경 영상을 상기 전경 영상과 논리곱하여 상기 동적 객체 추정 영상을 생성할 수 있다.

[0015] 상기 동적 객체 검출부는 적대적 학습 방식으로 학습된 GAN(Generative Adversarial Network)으로 구현될 수

있다.

[0016] 상기 목적을 달성하기 위한 본 발명의 다른 실시예에 따른 동적 객체 검출 방법은 미리 학습된 패턴 추정 방식에 따라 단일 영상에서 연속하는 기지정된 개수의 다수 프레임 각각에 대해 독립적으로 인코딩하여 다수의 전경 특징 맵을 생성하며, 상기 다수의 전경 특징 맵을 결합하여 디코딩하여 전경 영상을 생성하는 단계; 미리 학습된 패턴 추정 방식에 따라 상기 다수 프레임을 결합하고 인코딩하여 배경 특징 맵을 생성하며, 상기 배경 특징 맵을 디코딩하여 배경 영상을 생성하는 단계; 상기 전경 영상과 배경 영상을 결합하여, 동적 객체 추정 영상을 생성하는 단계; 및 미리 학습된 패턴 추정 방식에 따라 상기 동적 객체 추정 영상으로부터 동적 객체 영상을 생성하는 단계; 를 포함한다.

### 발명의 효과

[0017] 따라서, 본 발명의 실시예에 따른 동적 객체 검출 장치 및 방법은 각각 서로 다른 방식으로 미리 학습된 다수의 인공 신경망을 이용하여 연속 프레임의 단일 영상으로부터 전경과 배경을 구분하여 각각 획득하고, 획득된 전경과 배경을 다시 미리 학습된 인공 신경망에 입력하여 배경, 조도 및 날씨 변화에 강인하고 카메라의 움직임이 크거나 객체가 빠르게 움직일 때도 동적 객체를 정확하게 검출할 수 있다.

### 도면의 간단한 설명

[0018] 도1 은 본 발명의 일실시예에 따른 동적 객체 검출 장치의 개략적 구조를 나타낸다.  
 도2 는 도1 의 전경 추출부 구성의 일예를 나타낸다.  
 도3 은 도1 의 배경 추출부 구성의 일예를 나타낸다.  
 도4 는 본 발명의 일실시예에 따른 동적 객체 검출 방법을 나타낸다.  
 도5 는 입력 영상과 전경 추출부에서 추출된 전경 영상 및 배경 추출부에서 추출된 배경 영상의 일예를 나타낸다.  
 도6 은 본 발명의 동적 객체 검출 장치 및 방법에 의해 검출된 동적 객체 영상과 검증 영상을 비교하여 나타낸다.

### 발명을 실시하기 위한 구체적인 내용

[0019] 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.

[0020] 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 부재를 나타낸다.

[0021] 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.

[0022] 도1 은 본 발명의 일실시예에 따른 동적 객체 검출 장치의 개략적 구조를 나타낸다.

[0023] 도1 을 참조하면, 본 실시예에 따른 동적 객체 검출 장치는 영상 획득부(100), 전경 추출부(200), 배경 추출부(300), 전경 배경 결합부(400) 및 동적 객체 검출부(500)를 포함한다.

[0024] 영상 획득부(100)는 다수의 연속하는 프레임으로 구성되는 단일 영상을 획득한다. 여기서 단일 영상은 일예로 RGB 영상일 수 있으며, 고정식 또는 비 고정식 카메라에서 획득된 영상일 수 있다. 영상 획득부(100)는 카메라, 저장 장치 및 통신 장치로 구현될 수 있으나, 이에 한정되지 않는다.

[0025] 전경 추출부(200)는 미리 학습된 인공 신경망(Artificial Neural Network)으로 구현되어, 영상 획득부(100)에서 획득된 다수 프레임의 단일 영상에서 2개의 프레임( $t$ ,  $t+1$ )을 인가받아 전경 영상을 획득한다. 이때 전경 추출부(200)는 미리 학습된 패턴 추정 방식에 따라 2개의 프레임( $t$ ,  $t+1$ ) 각각에 대해 독립적으로 인코딩하여

각 프레임의 전경에 대한 특징을 추출하여 전경 특징 맵을 생성하고, 각 프레임에서 생성된 전경 특징 맵을 병합하여 디코딩함으로써, 전경 영상을 획득한다.

- [0026] 여기서 전경 추출부(200)는 전경 영상에 대한 검증 자료를 이용하여 학습될 수 있다.
- [0027] 한편 배경 추출부(300)는 전경 추출부(200)와 상이한 구조의 미리 학습된 인공 신경망으로 구현되어, 영상 획득부(100)에서 획득된 다수 프레임의 단일 영상에서 2개의 프레임(t, t+1)을 인가받아 배경 영상을 획득한다. 그러나 배경 추출부(300)는 전경 추출부(200)와 달리 2개의 프레임(t, t+1)을 함께 인코딩하여 배경에 대한 특징을 추출하여 배경 특징 맵을 생성하고, 생성된 배경 특징 맵을 디코딩함으로써, 배경 영상을 획득한다.
- [0028] 여기서 배경 추출부(300)는 학습 과정에서 학습 데이터의 의도하지 않은 배경을 학습함으로써, 배경 추출 성능이 저하되는 것을 방지하기 위해, 전경 영상에 대한 검증 자료를 반전한 반전 검증 자료를 이용하여 학습될 수 있다. 즉 배경 추출부(300)는 학습 데이터의 배경을 이용하여 학습되는 것이 아니라, 검증된 전경 영상을 반전하는 방식으로 전경을 제외하도록 학습되어, 배경을 직접 학습하는 방식보다 정확하게 배경을 추출하도록 학습될 수 있다.
- [0029] 전경 배경 결합부(400)는 전경 추출부(200)에서 획득된 전경 영상과 배경 추출부(300)에서 획득된 배경 영상을 인가받고, 배경 영상을 반전하고 합성하여 동적 객체 추정 영상을 생성한다. 전경 배경 결합부(400)는 일예로 배경 영상을 반전하고, 반전된 배경 영상을 전경 영상과 논리 곱하여 노이즈를 제거함으로써 동적 객체 추정 영상을 생성할 수 있다.
- [0030] 동적 객체 검출부(500)는 미리 학습된 인공 신경망으로 구현되어 전경 배경 결합부(400)에서 생성된 동적 객체 추정 영상을 인가받아 동적 객체 추정 영상에서 동적 객체를 검출한다. 동적 객체 검출부(500)는 일예로 비지도 학습 방식으로 학습되는 인공 신경망의 하나인 GAN(Generative Adversarial Network)로 구현될 수 있다. GAN은 입력 영상의 픽셀값을 확률 분포 모델로 모델링하여 영상을 추정하는 인공 신경망이다. 본 실시예에서 동적 객체 검출부(500)는 검증 자료(ground truth)를 기반으로 적대적 학습(Adversarial learning)된 인공 신경망일 수 있다.
- [0031] 동적 객체 검출부(500)는 0 또는 1의 이진 분류 방식으로 동적 객체를 검출하기 때문에 크로스 엔트로피(cross entropy)를 사용하여, 검증 자료와의 오차에 따라 인공 신경망의 가중치(weight)와 바이어스(bias)를 역전파(backpropagation)하여 학습될 수 있다. 동적 객체 검출부(500)는 수학적 1에 따라 계산되는 비용 함수(Cost Function)(H(p,q))가 낮아지도록 학습될 수 있다.

## 수학적식 1

$$H(p, q) = - \sum_i p_i \log q_i = - y \log \hat{y} - (1 - y) \log(1 - \hat{y})$$

- [0032]
- [0033] 여기서  $p_i$ 는 검증 자료(ground truth)의 픽셀 값을 나타내고,  $q_i$ 는 동적 객체 검출부(500)에서 판별한 픽셀 값이며,  $p_i$ 와  $q_i$ 는 모두 0 또는 1의 값을 갖는다.
- [0034] 도2 는 도1 의 전경 추출부 구성의 일예를 나타낸다.
- [0035] 도2 에서 (a)는 전경 추출부(200)의 구성을 나타내고, (b)는 (a)의 구성에 따라 구현되는 인공 신경망의 구조의 예를 나타낸다.
- [0036] 도2 를 참조하면, 전경 추출부(200)는 제1 전경 인코더(210), 제2 전경 인코더(220), 특징 결합부(230) 및 전경 디코더(240)를 포함한다.
- [0037] 제1 전경 인코더(210)와 제2 전경 인코더(220)는 영상 획득부(100)에서 획득된 다수 프레임의 단일 영상에서 연속하는 2개의 프레임 중 대응하는 프레임을 인가받아 각각 전경 특징 맵을 추출한다.
- [0038] 예를 들면, 제1 전경 인코더(210)는 단일 영상에서 t번째 프레임(t)을 인가받아 제1 전경 특징 맵을 추출하고, 제2 전경 인코더(220)는 t+1번째 프레임(t+1)을 인가받아 제2 전경 특징 맵을 추출한다.
- [0039] 제1 및 제2 전경 인코더(210, 220)는 각각 다양한 물체에 대한 데이터가 포함된 ImageNet에서 미리 학습된 인공 신경망으로 구현될 수 있다. 이때, 제1 및 제2 전경 인코더(210, 220)는 움직임이 가능한 객체(예를 들면, 차



량, 사람, 동물)에 대해서만 학습하고, 움직임이 거의 나타나지 않는 객체(예를 들면, 나무, 건물, 도로)에 대해서는 학습하지 않음으로써, 동적 객체에 대한 전경 특징 맵을 추출할 수 있도록 학습될 수 있다.

- [0040] 제1 및 제2 전경 인코더(210, 220)는 일예로 VGG16으로 구현될 수 있으나, 동적 물체 검출에서는 고해상도와 물체의 위치 정보도 중요하기 때문에, 제1 및 제2 전경 인코더(210, 220)는 다수개의 컨볼루션 레이어(Convolutional Layer)와 3 개의 풀리 커넥티드 레이어(Fully-Connected Layer)를 포함하는 VGG16 네트워크 전체를 이용하지 않고, 풀리 커넥티드 레이어를 제외한 1 ~ 4개의 컨볼루션 레이어를 포함하여 구성될 수 있다.
- [0041] 여기서 제1 및 제2 전경 인코더(210, 220)는 동일 구조와 동일 가중치를 갖는 쌍(siamese) 신경망으로 구현되어, 동일한 방식으로 특징을 추출할 수 있다.
- [0042] 특징 결합부(230)는 제1 및 제2 전경 인코더(210, 220)에서 추출된 제1 및 제2 특징 맵을 결합한다. 특징 결합부(230)는 제1 및 제2 특징 맵을 연결(concatenate)한 이후, 컨볼루션 연산을 수행함으로써, 통합 특징 맵을 생성한다.
- [0043] 전경 디코더(240) 또한 미리 학습된 인공 신경망으로 구현되어, 통합 특징맵을 인가받아 전경 영상을 생성한다. 전경 디코더(240)는 일예로 다수의 디컨볼루션(Deconvolution) 레이어를 포함하는 신경망으로 구현될 수 있다. 본 실시예에서 전경 디코더(240)는 제1 및 제2 특징 맵의 특징이 포함된 통합 특징맵을 인가받음으로써, 제1 및 제2 전경 인코더(210, 220)가 추출하는 특징 정보를 참조하여, 전경 영상을 생성한다.
- [0044] 그리고 전경 디코더(240)는 최종적으로 활성화 함수(active function)로서 시그모이드(sigmoid) 함수를 이용함으로써, 전경 영상의 각 픽셀 값의 범위가 0 ~ 1이 되도록 제한한다.
- [0045] 도3 은 도1 의 배경 추출부 구성의 일예를 나타낸다.
- [0046] 도3 에서도 (a)는 배경 추출부(300)의 구성을 나타내고, (b)는 (a)의 구성에 따라 구현되는 인공 신경망의 구조의 예를 나타낸다.
- [0047] 도3 을 참조하면, 배경 추출부(300)는 배경 인코더(310) 및 배경 디코더(320)를 포함한다.
- [0048] 배경 인코더(310)는 영상 획득부(100)에서 획득된 다수 프레임의 단일 영상에서 연속하는 2개의 프레임( $t$ ,  $t+1$ )을 연결(concatenate)하여 인가받아, 미리 학습된 패턴 추정 방식에 따라 2개의 프레임( $t$ ,  $t+1$ )에서 배경 특징 맵을 추출한다. 여기서 배경 인코더(310)는 배경의 개략적인 정보(Low level information)가 필요하기 때문에, 전경 인코더(210, 220)에 비해 적은 수의 레이어를 갖도록 구성될 수 있으며, 일예로 컨볼루션 신경망(CNN)으로 구현될 수 있다.
- [0049] 그리고 배경 디코더(320)는 배경 인코더(310)로부터 배경 특징 맵을 인가받아 디코딩하여 배경 영상을 생성한다. 배경 디코더(320) 또한 전경 디코더(240)와 유사하게 마지막 레이어에 활성화 함수로서 시그모이드(sigmoid) 함수를 이용함으로써, 배경 영상의 각 픽셀 값의 범위가 0 ~ 1이 되도록 제한한다.
- [0050] 상기한 바와 같이, 배경 추출부(300)의 배경 인코더(310)와 배경 디코더(320)는 검증된 정경 영상을 반전한 반전 전경 영상을 이용하여 학습될 수 있다.
- [0051] 상기에서는 전경 추출부(200)와 배경 추출부(300)가 각각 연속하는 2개의 프레임( $t$ ,  $t+1$ )을 인가받아 전경 영상 및 배경 영상을 획득하는 것으로 설명하였으나, 본 발명은 이에 한정되지 않는다. 즉 전경 추출부(200)와 배경 추출부(300)는 기지정된 다수의 프레임을 인가받아 전경 영상 및 배경 영상을 획득하도록 구성될 수도 있다.
- [0052] 도4 는 본 발명의 일실시예에 따른 동적 객체 검출 방법을 나타낸다.
- [0053] 도1 내지 도3 을 참조하여, 도4 의 동적 객체 검출 방법을 설명하면, 동적 객체 검출 장치는 우선 영상 획득부(100)가 동적 객체를 검출하고자 하는 영상을 획득한다(S10). 여기서 획득되는 영상은 다수의 프레임을 갖는 단일 영상으로 RGB 영상일 수도 있으나 이에 한정되지 않는다. 또한 고정식 또는 비 고정식 카메라에서 획득된 영상일 수 있다.
- [0054] 전경 추출부(200)는 획득된 단일 영상에서 연속하는 기지정된 개수의 프레임 각각에 대해 독립적으로 인코딩하여 전경 특징을 추출함으로써, 전경 특징 맵을 생성하고, 생성된 전경 특징 맵을 디코딩하여 전경 영상을 획득한다(S20).
- [0055] 전경 추출부(200)가 전경 영상을 획득하는 과정을 상세하게 설명하면, 전경 추출부(200)의 다수의 전경 인코더(210, 220)는 각각 연속하는 프레임 중 대응하는 프레임을 인가받고, 미리 학습된 패턴 추정 방식에 따라 전경



특징을 추출하여 다수의 전경 특징 맵을 생성한다(S21).

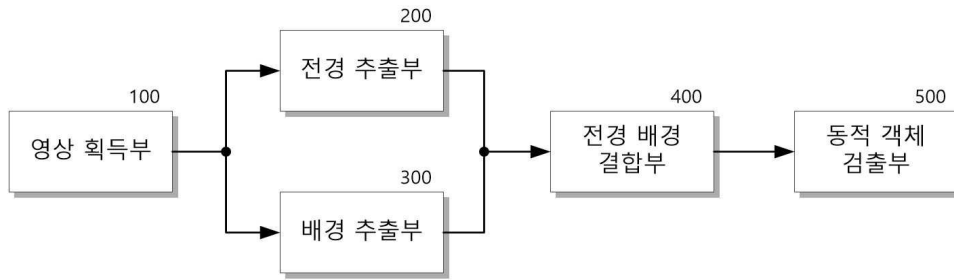
- [0056] 그리고 특징 결합부(230)는 다수의 전경 인코더(210, 220)에서 생성된 다수의 전경 특징 맵을 기지정된 방식으로 결합하여 통합 특징 맵을 생성한다(S22).
- [0057] 전경 디코더(240)는 통합 특징 맵을 인가받고 미리 학습된 방식에 따라 통합 특징 맵의 전경 특징을 디코딩함으로써, 전경 영상을 획득한다(S23).
- [0058] 한편, 전경 추출부(200)와 별도로 배경 추출부(300)는 연속하는 기지정된 개수의 프레임을 연결하여 인가받아 인코딩하여 배경 특징을 추출함으로써 배경 특징 맵을 생성하고, 생성된 배경 특징 맵을 디코딩하여 배경 영상을 획득한다(S30).
- [0059] 배경 영상을 획득하는 단계(S30)는 우선 배경 인코더(310)가 연속하는 기지정된 개수의 프레임을 연결하여 인가받고, 미리 학습된 패턴 추정 방식에 따라 연속하는 프레임으로부터 배경 특징을 추출하여 배경 특징 맵을 생성한다(S31). 그리고 배경 디코더(320)는 생성된 배경 특징 맵을 미리 학습된 패턴 추정 방식에 따라 디코딩하여 배경 영상을 획득한다(S32).
- [0060] 전경 영상과 배경 영상이 획득되면, 전경 배경 결합부(400)가 전경 영상과 배경 영상을 결합하여 동적 객체 추정 영상을 생성한다(S40). 여기서 전경 배경 결합부(400)는 일예로 배경 영상을 반전하고, 논리곱하여 동적 객체 추정 영상을 생성할 수 있다.
- [0061] 한편, 동적 객체 검출부(500)는 동적 객체 추정 영상을 인가받고, 미리 학습된 패턴 추정 방식에 따라 동적 객체 추정 영상으로부터 동적 객체를 검출한다.
- [0062] 도5 는 입력 영상과 전경 추출부에서 추출된 전경 영상 및 배경 추출부에서 추출된 배경 영상의 일예를 나타내고, 도6 은 본 발명의 동적 객체 검출 장치 및 방법에 의해 검출된 동적 객체 영상과 검증 영상을 비교하여 나타낸다.
- [0063] 도5 에서 (a)는 입력 영상의 특정 프레임을 나타내며, (b)는 전경 추출부(200)에서 추출된 전경 영상을 나타내고, (c)는 배경 추출부(300)에서 추출된 배경 영상을 나타낸다. 그리고 도6 에서 (a)는 본 실시예에 따른 동적 객체 검출 장치에서 검출한 동적 객체 검출 영상을 나타내고, (b)는 검증 자료를 나타낸다.
- [0064] 도5 의 (b) 및 (c)를 (a)의 입력 영상과 비교하면, 전경 추출부(200)와 배경 추출부(300)는 각각 입력 영상 상에서 전경과 배경을 추출하고 있으나, 조도 및 날씨 변화 및 주변 환경에서의 움직임 요소 등에 의해 객체가 아닌 영역에 대해서도 전경으로 식별하는 영역이 있음을 알 수 있다.
- [0065] 그러나 도6 의 (a)에 도시된 바와 같이, 전경 영상과 배경 영상을 결합하여 생성된 동적 객체 추정 영상에서 동적 객체를 검출한 동적 객체 영상을 도5 의 (b) 및 (c)와 비교하면 매우 정확하게 동적 객체만을 검출하였음을 알 수 있다. 그리고 이러한 동적 객체 영상은 (b)에 도시된 검증 자료와 비교하여도 객체 검출 성능이 매우 우수함을 확인할 수 있다.
- [0066] 본 발명에 따른 방법은 컴퓨터에서 실행 시키기 위한 매체에 저장된 컴퓨터 프로그램으로 구현될 수 있다. 여기서 컴퓨터 판독가능 매체는 컴퓨터에 의해 액세스 될 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 비분리형 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.
- [0067] 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.
- [0068] 따라서, 본 발명의 진정한 기술적 보호 범위는 첨부된 청구범위의 기술적 사상에 의해 정해져야 할 것이다.

### 부호의 설명

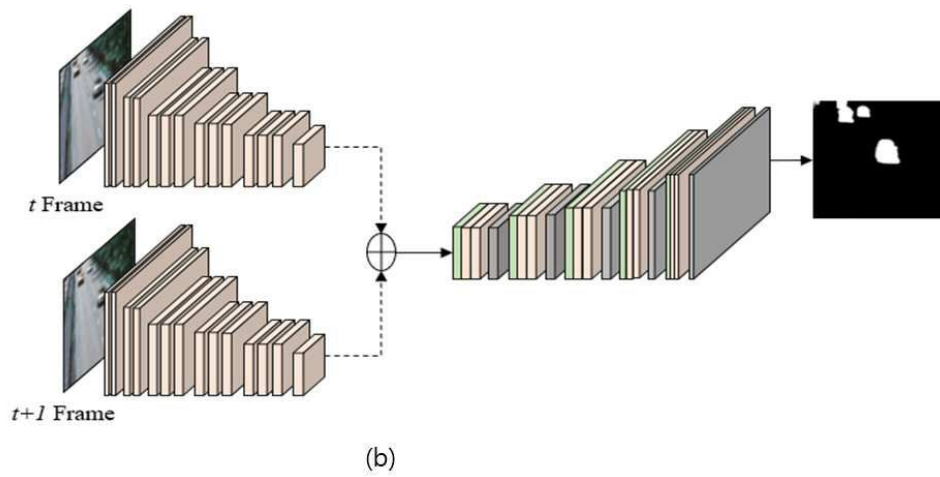
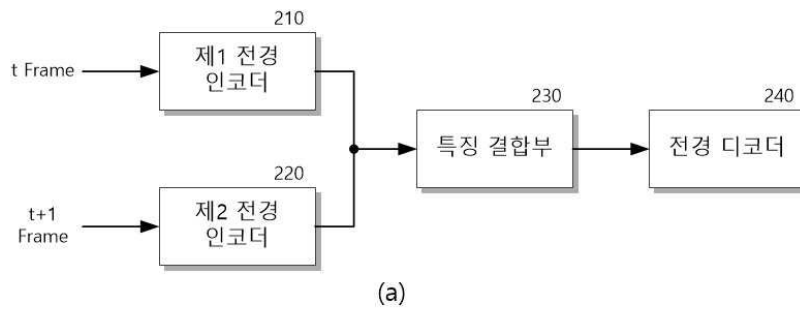
- [0069] 100: 영상 획득부 200: 전경 추출부  
300: 배경 추출부 400: 전경 배경 결합부  
500: 동적 객체 검출부

도면

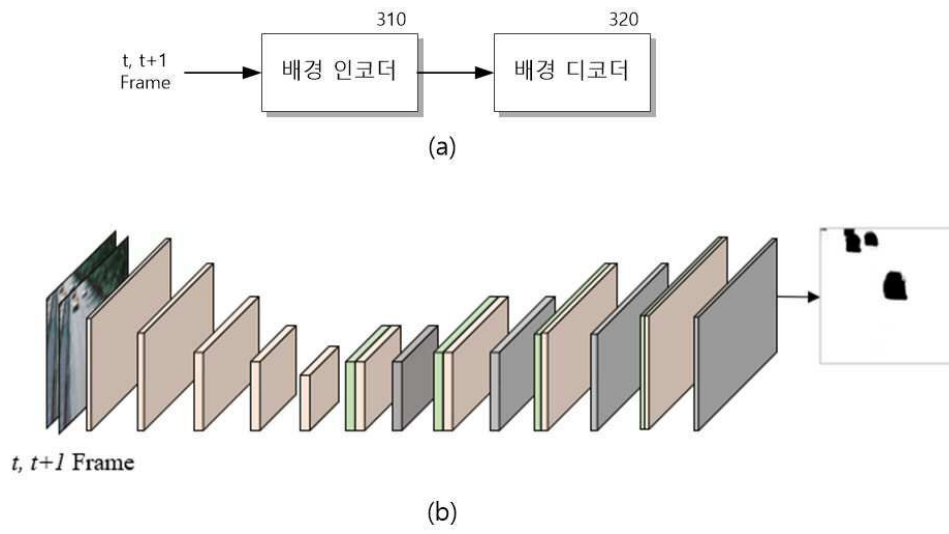
도면1



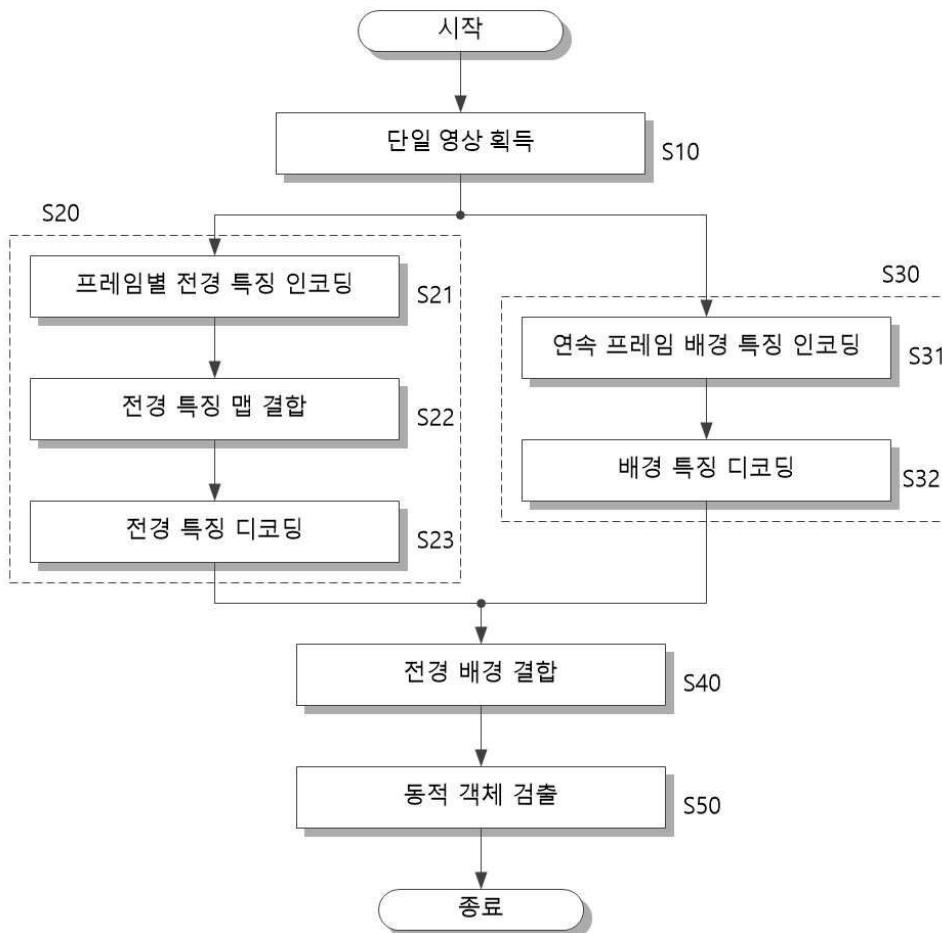
도면2



도면3



도면4



도면5

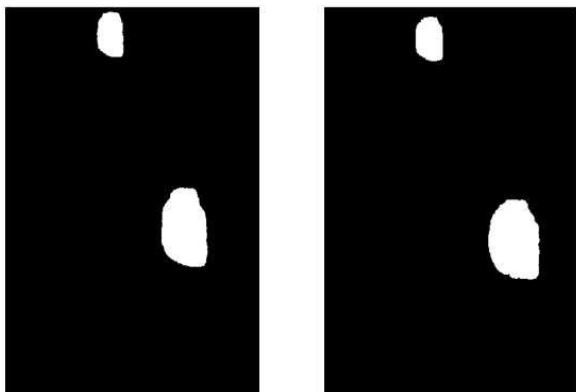
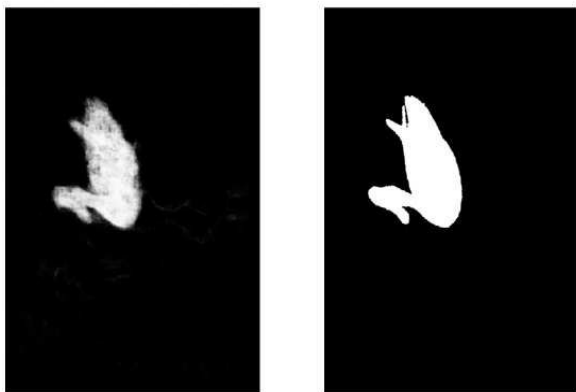
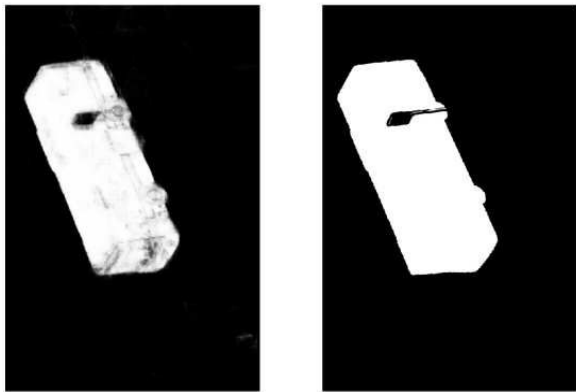


(a)

(b)

(c)

도면6



(a)

(b)