



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2022년02월07일

(11) 등록번호 10-2359982

(24) 등록일자 2022년02월03일

(51) 국제특허분류(Int. Cl.)  
G06T 7/292 (2017.01) G06T 7/246 (2017.01)

(52) CPC특허분류  
G06T 7/292 (2017.01)  
G06T 7/248 (2017.01)

(21) 출원번호 10-2020-0047605

(22) 출원일자 2020년04월20일

심사청구일자 2020년04월20일

(65) 공개번호 10-2021-0129503

(43) 공개일자 2021년10월28일

(56) 선행기술조사문헌

MD.MAKLACHUR RAHMAN 등, Efficient Visual Tracking With Stacked Channel-Spatial Attention Learning, IEEE Access.(2020.03.27.)\*

(뒷면에 계속)

(73) 특허권자

연세대학교 산학협력단

서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)

(72) 발명자

이상윤

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

이현성

서울특별시 강서구 양천로 595, 101동 1202호(염창동, 엠창금호타운아파트)

(뒷면에 계속)

(74) 대리인

민영준

전체 청구항 수 : 총 18 항

심사관 : 이재원

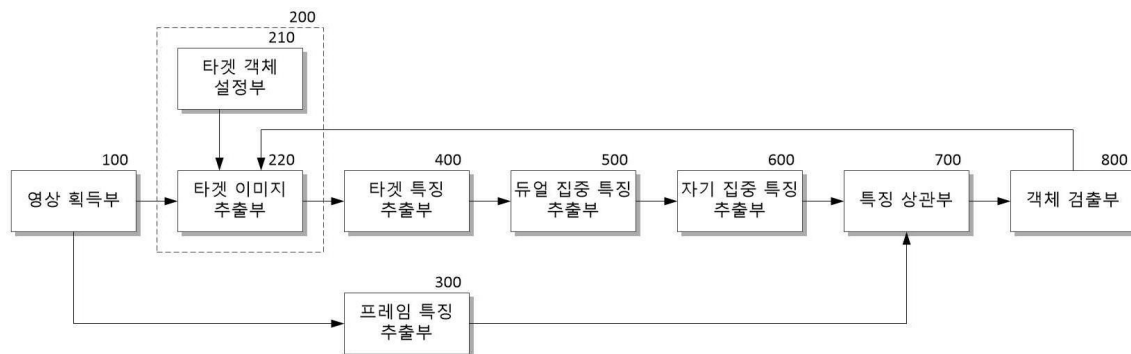
(54) 발명의 명칭 자기 집중 방식 객체 추적 장치 및 방법

### (57) 요약

본 발명은 영상의 다수 프레임 이미지 중 초기 프레임 이미지 또는 이전 타겟 객체가 추출된 프레임 이미지에서 타겟 객체 영역을 추출하여 타겟 이미지를 획득하는 타겟 이미지 획득부, 다수 프레임 이미지 중 타겟 객체가 검출되어야 하는 현재 프레임 이미지에 대해 미리 학습된 패턴 추정 방식에 따라 특징을 추출하여 프레임 특징맵을

(뒷면에 계속)

### 대표도



획득하는 프레임 특징 추출부, 타겟 이미지에 대해 미리 학습된 패턴 추정 방식에 따라 특징을 추출하여 타겟 특징맵을 획득하는 타겟 특징 추출부, 타겟 특징맵에서 채널 및 공간적 중요도에 따라 가중치를 가중하여 가중 특징맵을 획득하는 듀얼 집중 특징 추출부, 이전까지 획득된 가중 특징맵을 중요도에 따라 가중 누적하고 특징 추출하여 기준 특징맵을 획득하는 자기 집중 특징 추출부, 프레임 특징맵과 기준 특징맵에 대한 상관 연산을 수행하여 스코어 맵을 획득하는 특징 상관부 및 스코어 맵을 기반으로 현재 프레임 이미지에서의 타겟 객체의 위치를 판별하는 객체 검출부를 포함하여, 타겟 객체의 외형 변화에도 강인하고 정확하게 객체를 추적할 수 있는 객체 추적 장치 및 방법을 제공할 수 있다.

(52) CPC특허분류

G06T 2207/20081 (2013.01)

(72) 발명자

**정태영**

서울특별시 서대문구 연희로10길 24-10, 101호(연희동)

**조명아**

경기도 성남시 분당구 판교로 421, 204동 403호(야탑동, 탐마울대우아파트)

(56) 선행기술조사문헌

Zheng Zhu 등, End-to-end Flow Correlation Tracking with Spatial-temporal Attention, arXiv:1711.01124v4.(2018.02.27.)\*

Qiang Wang 등, Learning Attention: Residual Attentional Network for High Performance Online Visual Tracking, IEEE CVF.(2018.06.18.)

Mohamed H. Abdelpakey 등, DensSiam: End-to-End Densely-Siamese Network with Self-Attention Model for Object Tracking, arXiv:1809.02714v1.(2018.09.07.)

Mohamed H. Abdelpakey 등, DomainSiam: Domain-Aware Siamese Network for Visual Object Tracking, arXiv:1908.07905v1.(2019.08.21.)

\*는 심사관에 의하여 인용된 문헌

이 발명을 지원한 국가연구개발사업

과제고유번호

2019070970

과제번호

NRF-2018M3E3A1057289

부처명

과학기술정보통신부

과제관리(전문)기관명

한국연구재단

연구사업명

복합인지기술개발사업

연구과제명

(2세부)이중 CCTV 영상에서의 딥러닝 기반 실종자 초동 신원확인 및 추적 시스템

(1단계)(2/2)

기 여 율

1/1

과제수행기관명

연세대학교 산학협력단

연구기간

2019.04.23 ~ 2020.04.22

## 명세서

### 청구범위

#### 청구항 1

영상의 다수 프레임 이미지 중 초기 프레임 이미지 또는 이전 타겟 객체가 추출된 프레임 이미지에서 타겟 객체 영역을 추출하여 타겟 이미지를 획득하는 타겟 이미지 획득부;

상기 다수 프레임 이미지 중 타겟 객체가 검출되어야 하는 현재 프레임 이미지에 대해 미리 학습된 패턴 추정 방식에 따라 특징을 추출하여 프레임 특징맵을 획득하는 프레임 특징 추출부;

상기 타겟 이미지에 대해 미리 학습된 패턴 추정 방식에 따라 특징을 추출하여 타겟 특징맵을 획득하는 타겟 특징 추출부;

상기 타겟 특징맵에서 채널 및 공간적 중요도에 따라 가중치를 가중하여 가중 특징맵을 획득하는 듀얼 집중 특징 추출부;

이전까지 획득된 가중 특징맵을 중요도에 따라 가중 누적하고 특징 추출하여 기준 특징맵을 획득하는 자기 집중 특징 추출부;

상기 프레임 특징맵과 상기 기준 특징맵에 대한 상관 연산을 수행하여 스코어 맵을 획득하는 특징 상관부; 및

상기 스코어 맵을 기반으로 현재 프레임 이미지에서의 타겟 객체의 위치를 판별하는 객체 검출부를 포함하되,

상기 자기 집중 특징 추출부는

미리 학습된 패턴 추정 방식에 따라 상기 가중 특징맵에서 타겟 객체의 특징을 키 벡터로 추출하는 키 벡터 획득부;

미리 학습된 패턴 추정 방식에 따라 상기 가중 특징맵에서 타겟 객체의 표현값을 값 벡터로 추출하는 값 벡터 획득부;

이전 획득된 기준 특징맵을 인가받고, 인가된 이전 시점 기준 특징맵에 대해 미리 학습된 패턴 추정 방식에 따라 이전 시점까지 가중 누적 추출된 타겟 객체의 특징을 나타내는 쿼리 벡터를 추출하는 쿼리 벡터 획득부; 및

상기 키 벡터와 상기 쿼리 벡터 사이의 관계에 따른 가중치를 상기 값 벡터에 가중하여 상기 기준 특징맵을 획득하는 자기 집중부를 포함하는 객체 검출 장치.

#### 청구항 2

제1항에 있어서, 상기 듀얼 집중 특징 추출부는

상기 타겟 특징맵에서 채널 축 방향으로 각 채널별 중요도를 나타내는 채널 집중 특징을 추정하여 채널 집중맵을 추출하고, 추출된 채널 집중맵을 타겟 특징맵에 가중하여 채널 집중 특징맵을 획득하는 채널 집중 특징 추출부; 및

상기 채널 집중 특징맵에서 높이와 폭 방향의 2차원 공간 위치에 따른 공간적 중요도를 나타내는 공간 집중 특징을 추정하여 공간 집중맵을 추출하고, 추출된 공간 집중맵을 상기 채널 집중 특징맵에 가중하여 상기 가중 특징맵을 획득하는 공간 집중 특징 추출부를 포함하는 객체 검출 장치.

#### 청구항 3

제2항에 있어서, 상기 채널 집중 특징 추출부는

미리 학습된 패턴 추정 방식에 따라 상기 타겟 특징맵의 채널 축 방향으로의 다수의 채널 벡터 각각의 중요도를 추출하여 채널 집중맵을 획득하는 채널 집중 모듈; 및

상기 타겟 특징맵에 대해 상기 채널 집중맵을 가중하여 상기 채널 집중 특징맵을 획득하는 채널 가중부를 포함

하는 객체 검출 장치.

#### 청구항 4

제3항에 있어서, 상기 채널 집중 모듈은

상기 타겟 특징맵의 채널 축 방향으로의 다수의 채널 벡터 각각에 대해 맥스 풀링 및 평균값 풀링을 수행하는 풀링부;

2가지 풀링 결과를 기반으로 학습된 패턴 추정 방식에 따라 각 2가지 방식으로 풀링된 채널 벡터에서 채널별 중요도를 추출하는 다중 퍼셉트론부; 및

추출된 2가지 채널별 중요도를 서로 가산한 후, 기지정된 활성화 함수로 활성화하여 채널 집중맵을 획득하는 활성화부를 포함하는 객체 검출 장치.

#### 청구항 5

제3항에 있어서, 상기 공간 집중 특징 추출부는

미리 학습된 패턴 추정 방식에 따라 상기 채널 집중 특징맵의 높이와 폭 방향의 2차원 공간 벡터에서 공간 위치별 중요도를 추출하여 공간 집중맵을 획득하는 채널 집중 모듈; 및

상기 채널 집중 특징맵에 대해 상기 공간 집중맵을 가중하여 상기 가중 특징맵을 획득하는 채널 가중부를 포함하는 객체 검출 장치.

#### 청구항 6

제5항에 있어서, 상기 채널 집중 모듈은

상기 채널 집중 특징맵에서 높이와 폭 방향의 2차원 공간 벡터단위로 맥스 풀링 및 평균값 풀링을 수행하는 풀링부;

2가지 풀링 결과를 결합하고, 결합된 풀링 결과에 대해 학습된 패턴 추정 방식에 따라 공간 특징을 추출하는 공간 특징 추출부; 및

추출된 공간 특징을 기지정된 활성화 함수로 활성화하여 공간 집중맵을 획득하는 활성화부를 포함하는 객체 검출 장치.

#### 청구항 7

삭제

#### 청구항 8

제1항에 있어서, 상기 자기 집중부는

상기 키 벡터(K)와 상기 쿼리 벡터(Q)에 따른 가중치를 수학식 1

$$f_t = \text{Softmax}(QK^T)V$$

에 따라 상기 값 벡터(V)에 가중하여 상기 기준 특징맵( $f_t$ )을 획득하는 객체 검출 장치.

#### 청구항 9

제1항에 있어서, 상기 타겟 이미지 획득부는

다수 프레임 이미지 중 초기 프레임 이미지에서는 사용자 명령에 따라 타겟 객체 영역을 설정하거나 다른 장치에서 획득하여 인가된 타겟 이미지와 기지정된 방식으로 비교하여 유사한 것으로 판별되는 영역을 타겟 객체 영역으로 설정하여 상기 타겟 이미지를 획득하는 객체 검출 장치.

#### 청구항 10

제1항에 있어서, 상기 타겟 이미지 획득부는

다수 프레임 이미지 중 나머지 프레임 이미지에서는 상기 객체 검출부에서 검출된 타겟 객체 위치에 따라 타겟 객체 영역을 설정하여 상기 타겟 이미지를 획득하는 객체 검출 장치.

#### 청구항 11

영상의 다수 프레임 이미지 중 초기 프레임 이미지 또는 이전 타겟 객체가 추출된 프레임 이미지에서 타겟 객체 영역을 추출하여 타겟 이미지를 획득하는 단계;

상기 다수 프레임 이미지 중 타겟 객체가 검출되어야 하는 현재 프레임 이미지에 대해 미리 학습된 패턴 추정 방식에 따라 특징을 추출하여 프레임 특징맵을 획득하는 단계;

상기 타겟 이미지에 대해 미리 학습된 패턴 추정 방식에 따라 특징을 추출하여 타겟 특징맵을 획득하는 단계;

상기 타겟 특징맵에서 채널 및 공간적 중요도에 따라 가중치를 가중하여 가중 특징맵을 획득하는 단계;

이전까지 획득된 가중 특징맵을 중요도에 따라 가중 누적하고 특징 추출하여 기준 특징맵을 획득하는 단계;

상기 프레임 특징맵과 상기 기준 특징맵에 대한 상관 연산을 수행하여 스코어 맵을 획득하는 단계; 및

상기 스코어 맵을 기반으로 현재 프레임 이미지에서의 타겟 객체의 위치를 판별하는 단계를 포함하되,

상기 기준 특징맵을 획득하는 단계는

미리 학습된 패턴 추정 방식에 따라 상기 가중 특징맵에서 타겟 객체의 특징을 키 벡터로 추출하는 단계;

미리 학습된 패턴 추정 방식에 따라 상기 가중 특징맵에서 타겟 객체의 표현값을 값 벡터로 추출하는 단계;

이전 획득된 기준 특징맵을 인가받고, 인가된 이전 시점 기준 특징맵에 대해 미리 학습된 패턴 추정 방식에 따라 이전 시점까지 가중 누적 추출된 타겟 객체의 특징을 나타내는 쿼리 벡터를 추출하는 단계; 및

상기 기준 특징맵을 획득하기 위해 상기 키 벡터와 상기 쿼리 벡터 사이의 관계에 따른 가중치를 상기 값 벡터에 가중하는 단계를 포함하는 객체 검출 방법.

#### 청구항 12

제11항에 있어서, 상기 가중 특징맵을 획득하는 단계는

상기 타겟 특징맵에서 채널 축 방향으로 각 채널별 중요도를 나타내는 채널 집중 특징을 추정하여 채널 집중맵을 추출하고, 추출된 채널 집중맵을 타겟 특징맵에 가중하여 채널 집중 특징맵을 획득하는 단계; 및

상기 채널 집중 특징맵에서 높이와 폭 방향의 2차원 공간 위치에 따른 공간적 중요도를 나타내는 공간 집중 특징을 추정하여 공간 집중맵을 추출하고, 상기 가중 특징맵을 획득하기 위해 추출된 공간 집중맵을 상기 채널 집중 특징맵에 가중하는 단계를 포함하는 객체 검출 방법.

#### 청구항 13

제12항에 있어서, 상기 채널 집중 특징맵을 획득하는 단계는

미리 학습된 패턴 추정 방식에 따라 상기 타겟 특징맵의 채널 축 방향으로의 다수의 채널 벡터 각각의 중요도를 추출하여 채널 집중맵을 획득하는 단계; 및

상기 타겟 특징맵에 대해 상기 채널 집중맵을 가중하여 상기 채널 집중 특징맵을 획득하는 단계를 포함하는 객체 검출 방법.

#### 청구항 14

제13항에 있어서, 상기 채널 집중맵을 획득하는 단계는

상기 타겟 특징맵의 채널 축 방향으로의 다수의 채널 벡터 각각에 대해 맥스 풀링 및 평균값 풀링을 수행하는 단계;

2가지 폴링 결과를 기반으로 학습된 패턴 추정 방식에 따라 각 2가지 방식으로 폴링된 채널 벡터에서 채널별 중요도를 추출하는 단계; 및

추출된 2가지 채널별 중요도를 서로 가산한 후, 기지정된 활성화 함수로 활성화하여 채널 집중맵을 획득하는 단계를 포함하는 객체 검출 방법.

#### 청구항 15

제13항에 있어서, 상기 채널 집중 특징맵에 가중하는 단계는

미리 학습된 패턴 추정 방식에 따라 상기 채널 집중 특징맵의 높이와 폭 방향의 2차원 공간 벡터에서 공간 위치별 중요도를 추출하여 공간 집중맵을 획득하는 단계; 및

상기 채널 집중 특징맵에 대해 상기 공간 집중맵을 가중하여 상기 가중 특징맵을 획득하는 단계를 포함하는 객체 검출 방법.

#### 청구항 16

제15항에 있어서, 상기 공간 집중맵을 획득하는 단계는

상기 채널 집중 특징맵에서 높이와 폭 방향의 2차원 공간 벡터단위로 맥스 폴링 및 평균값 폴링을 수행하는 단계;

2가지 폴링 결과를 결합하고, 결합된 폴링 결과에 대해 학습된 패턴 추정 방식에 따라 공간 특징을 추출하는 단계; 및

추출된 공간 특징을 기지정된 활성화 함수로 활성화하여 공간 집중맵을 획득하는 단계를 포함하는 객체 검출 방법.

#### 청구항 17

삭제

#### 청구항 18

제11항에 있어서, 상기 값 벡터에 가중하는 단계는

상기 키 벡터(K)와 상기 쿼리 벡터(Q)에 따른 가중치를 수학적 식 1

$$f_t = \text{Softmax}(QK^T)V$$

에 따라 상기 값 벡터(V)에 가중하여 상기 기준 특징맵( $f_t$ )을 획득하는 객체 검출 방법.

#### 청구항 19

제11항에 있어서, 상기 타겟 이미지를 획득하는 단계는

다수 프레임 이미지 중 초기 프레임 이미지에서는 사용자 명령에 따라 타겟 객체 영역을 설정하거나 다른 방법에서 획득하여 인가된 타겟 이미지와 기지정된 방식으로 비교하여 유사한 것으로 판별되는 영역을 타겟 객체 영역으로 설정하여 상기 타겟 이미지를 획득하는 객체 검출 방법.

#### 청구항 20

제11항에 있어서, 상기 타겟 이미지를 획득하는 단계는

다수 프레임 이미지 중 나머지 프레임 이미지에서는 이전 타겟 객체의 위치를 판별하는 단계에서 검출된 타겟 객체 위치에 따라 타겟 객체 영역을 설정하여 상기 타겟 이미지를 획득하는 객체 검출 방법.

**발명의 설명**

**기술 분야**

[0001] 본 발명은 객체 추적 장치 및 방법에 관한 것으로, 강인하게 단일 객체를 추적하기 위한 자기 집중 방식 객체 추적 장치 및 방법에 관한 것이다.

## 배경 기술

[0002] 단일 객체 추적 기술은 영상 내에서 임의의 객체에 대한 위치 정보를 나타내는 박스가 추적 대상인 타겟 객체로 설정되면, 영상이 진행됨에 따라 타겟 객체의 위치를 지속적으로 추적해 나가는 기술로서, 자율주행, 드론, 산업 현장 감시, CCTV 상에서의 추적, 군수 분야 등 많은 분야들에 사용되고 있다.

[0003] 그리고 최근에는 객체 추적 기술에 딥러닝 기법이 주로 이용되고 있다. 기존의 객체 추적 기술의 경우, 다수 프레임 이미지의 영상에서 최초 프레임 이미지에 포함된 특정 객체가 타겟 객체로 설정되면, 설정된 타겟 객체 영역으로부터 기준 특징을 추출하고, 이후 매 프레임 이미지에서 기준 특징과 가장 비슷한 특징이 있는 위치를 지속적으로 탐색하는 방식으로 타겟 객체에 대한 추적이 수행된다. 그러나 이러한 방식은 최초 프레임 이미지에서 주어지는 타겟 객체 영역에 대한 단일 프레임 이미지만을 기준으로 이후 다수 프레임 이미지에 대한 객체 추적을 진행하므로, 시간에 따라 각 프레임 이미지에서 타겟 객체의 외형이 변화할 경우 추적 정확도가 급격히 하락하는 문제점이 존재한다.

## 선행기술문헌

### 특허문헌

[0004] (특허문헌 0001) 한국 등록 특허 제10-1852958호 (2018.04.23 등록)

## 발명의 내용

### 해결하려는 과제

[0005] 본 발명의 목적은 타겟 객체의 외형 변화에도 강인하게 객체를 추적할 수 있는 객체 추적 장치 및 방법을 제공하는데 있다.

[0006] 본 발명의 다른 목적은 영상에서 시간의 변화에 따른 타겟 객체의 외형 변화에 대응하여, 다수 프레임 이미지의 타겟 객체의 특징에서 불필요한 특징을 억제하면서 누적하여 강인한 기준 특징을 획득함으로써, 정확하게 객체를 추적할 수 있는 객체 추적 장치 및 방법을 제공하는데 있다.

### 과제의 해결 수단

[0007] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른 객체 추적 장치는 영상의 다수 프레임 이미지 중 초기 프레임 이미지 또는 이전 타겟 객체가 추출된 프레임 이미지에서 타겟 객체 영역을 추출하여 타겟 이미지를 획득하는 타겟 이미지 획득부; 상기 다수 프레임 이미지 중 타겟 객체가 검출되어야 하는 현재 프레임 이미지에 대해 미리 학습된 패턴 추정 방식에 따라 특징을 추출하여 프레임 특징맵을 획득하는 프레임 특징 추출부; 상기 타겟 이미지에 대해 미리 학습된 패턴 추정 방식에 따라 특징을 추출하여 타겟 특징맵을 획득하는 타겟 특징 추출부; 상기 타겟 특징맵에서 채널 및 공간적 중요도에 따라 가중치를 가중하여 가중 특징맵을 획득하는 듀얼 집중 특징 추출부; 이전까지 획득된 가중 특징맵을 중요도에 따라 가중 누적하고 특징 추출하여 기준 특징맵을 획득하는 자기 집중 특징 추출부; 상기 프레임 특징맵과 상기 기준 특징맵에 대한 상관 연산을 수행하여 스코어맵을 획득하는 특징 상관부; 및 상기 스코어맵을 기반으로 현재 프레임 이미지에서의 타겟 객체의 위치를 판별하는 객체 검출부를 포함한다.

[0008] 상기 듀얼 집중 특징 추출부는 상기 타겟 특징맵에서 채널 축 방향으로 각 채널별 중요도를 나타내는 채널 집중 특징을 추정하여 채널 집중맵을 추출하고, 추출된 채널 집중맵을 타겟 특징맵에 가중하여 채널 집중 특징맵을 획득하는 채널 집중 특징 추출부; 및 상기 채널 집중 특징맵에서 높이와 폭 방향의 2차원 공간 위치에 따른 공간적 중요도를 나타내는 공간 집중 특징을 추정하여 공간 집중맵을 추출하고, 추출된 공간 집중맵을 상기 채널 집중 특징맵에 가중하여 상기 가중 특징맵을 획득하는 공간 집중 특징 추출부를 포함할 수 있다.

[0009] 상기 채널 집중 특징 추출부는 미리 학습된 패턴 추정 방식에 따라 상기 타겟 특징맵의 채널 축 방향으로의 다수의 채널 벡터 각각의 중요도를 추출하여 채널 집중맵을 획득하는 채널 집중 모듈; 및 상기 타겟 특징맵에 대

해 상기 채널 집중맵을 가중하여 상기 채널 집중 특징맵을 획득하는 채널 가중부를 포함할 수 있다.

[0010] 상기 채널 집중 모듈은 상기 타겟 특징맵의 채널 축 방향으로의 다수의 채널 벡터 각각에 대해 맥스 풀링 및 평균값 풀링을 수행하는 풀링부; 2가지 풀링 결과를 기반으로 학습된 패턴 추정 방식에 따라 각 2가지 방식으로 풀링된 채널 벡터에서 채널별 중요도를 추출하는 다중 퍼셉트론부; 및 추출된 2가지 채널별 중요도를 서로 가산한 후, 기지정된 활성화 함수로 활성화하여 채널 집중맵을 획득하는 활성화부를 포함할 수 있다.

[0011] 상기 공간 집중 특징 추출부는 미리 학습된 패턴 추정 방식에 따라 상기 채널 집중 특징맵의 높이와 폭 방향의 2차원 공간 벡터에서 공간 위치별 중요도를 추출하여 공간 집중맵을 획득하는 채널 집중 모듈; 및 상기 채널 집중 특징맵에 대해 상기 공간 집중맵을 가중하여 상기 가중 특징맵을 획득하는 채널 가중부를 포함할 수 있다.

[0012] 상기 채널 집중 모듈은 상기 채널 집중 특징맵에서 높이와 폭 방향의 2차원 공간 벡터단위로 맥스 풀링 및 평균값 풀링을 수행하는 풀링부; 2가지 풀링 결과를 결합하고, 결합된 풀링 결과에 대해 학습된 패턴 추정 방식에 따라 공간 특징을 추출하는 공간 특징 추출부; 및 추출된 공간 특징을 기지정된 활성화 함수로 활성화하여 공간 집중맵을 획득하는 활성화부를 포함할 수 있다.

[0013] 상기 자기 집중 특징 추출부는 미리 학습된 패턴 추정 방식에 따라 상기 가중 특징맵에서 타겟 객체의 특징을 키 벡터로 추출하는 키 벡터 획득부; 미리 학습된 패턴 추정 방식에 따라 상기 가중 특징맵에서 타겟 객체의 표현값을 값 벡터로 추출하는 값 벡터 획득부; 이전 획득된 기준 특징맵을 인가받고, 인가된 이전 시점 기준 특징맵에 대해 미리 학습된 패턴 추정 방식에 따라 이전 시점까지 가중 누적 추출된 타겟 객체의 특징을 나타내는 쿼리 벡터를 추출하는 쿼리 벡터 획득부; 및 상기 키 벡터와 상기 쿼리 벡터 사이의 관계에 따른 가중치를 상기 값 벡터에 가중하여 상기 기준 특징맵을 획득하는 자기 집중부를 포함할 수 있다.

[0014] 상기 타겟 이미지 획득부는 다수 프레임 이미지 중 초기 프레임 이미지에서는 사용자 명령에 따라 타겟 객체 영역을 설정하거나 다른 장치에서 획득하여 인가된 타겟 이미지와 기지정된 방식으로 비교하여 유사한 것으로 판별되는 영역을 타겟 객체 영역으로 설정하여 상기 타겟 이미지를 획득할 수 있다.

[0015] 상기 타겟 이미지 획득부는 다수 프레임 이미지 중 나머지 프레임 이미지에서는 상기 객체 검출부에서 검출된 타겟 객체 위치에 따라 타겟 객체 영역을 설정하여 상기 타겟 이미지를 획득할 수 있다.

[0016] 상기 다른 목적을 달성하기 위한 본 발명의 다른 실시예에 따른 객체 추적 방법은 영상의 다수 프레임 이미지 중 초기 프레임 이미지 또는 이전 타겟 객체가 추출된 프레임 이미지에서 타겟 객체 영역을 추출하여 타겟 이미지를 획득하는 단계; 상기 다수 프레임 이미지 중 타겟 객체가 검출되어야 하는 현재 프레임 이미지에 대해 미리 학습된 패턴 추정 방식에 따라 특징을 추출하여 프레임 특징맵을 획득하는 단계; 상기 타겟 이미지에 대해 미리 학습된 패턴 추정 방식에 따라 특징을 추출하여 타겟 특징맵을 획득하는 단계; 상기 타겟 특징맵에서 채널 및 공간적 중요도에 따라 가중치를 가중하여 가중 특징맵을 획득하는 단계; 이전까지 획득된 가중 특징맵을 중요도에 따라 가중 누적하고 특징 추출하여 기준 특징맵을 획득하는 단계; 상기 프레임 특징맵과 상기 기준 특징맵에 대한 상관 연산을 수행하여 스코어 맵을 획득하는 단계; 및 상기 스코어 맵을 기반으로 현재 프레임 이미지에서의 타겟 객체의 위치를 판별하는 단계를 포함한다.

## 발명의 효과

[0017] 따라서, 본 발명의 실시예에 따른 객체 추적 장치 및 방법은 다수 프레임 이미지의 영상에서 타겟 객체가 효과적으로 추출된 프레임 이미지와 추출되지 않은 프레임 이미지를 구분하고, 이전 프레임 이미지에서 추출된 타겟 객체의 특징에서 불필요한 특징을 억제하면서 누적하여 기준 특징을 획득함으로써, 타겟 객체의 외형 변화에도 강인하고 정확하게 객체를 추적할 수 있다.

## 도면의 간단한 설명

[0018] 도 1은 본 발명의 일 실시예에 따른 객체 추적 장치의 개략적 구조를 나타낸다.

도 2는 도 1의 객체 추적 장치의 전체적인 동작을 설명하기 위한 도면이다.

도 3 및 도 4는 도 1의 듀얼 집중 특징 추출부의 구성과 동작을 설명하기 위한 도면이다.

도 5는 도 1의 자기 집중 특징 추출부의 구성과 동작을 설명하기 위한 도면이다.

도 6은 본 발명의 일 실시예에 따른 객체 추적 방법을 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 부재임을 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] 도 1은 본 발명의 일 실시예에 따른 객체 추적 장치의 개략적 구조를 나타내고, 도 2는 도 1의 객체 추적 장치의 전체적인 동작을 설명하기 위한 도면이다.
- [0023] 도 1 및 도 2를 참조하면, 본 실시예에 따른 객체 추적 장치는 영상 획득부(100), 타겟 이미지 획득부(200), 프레임 특징 추출부(300), 타겟 특징 추출부(400), 듀얼 집중 특징 추출부(500), 자기 집중 특징 추출부(600), 특징 상관부(700) 및 객체 검출부(800)를 포함한다.
- [0024] 영상 획득부(100)는 다수 프레임 이미지( $I_0 \sim I_n$ )로 구성된 영상을 획득한다. 영상 획득부(100)는 카메라 또는 CCTV와 같이 이미지 센서를 이용하여 영상을 획득할 수 있는 영상 획득 모듈로 구현되거나, 미리 획득된 영상이 저장된 메모리 모듈 등으로 구현될 수 있다. 또한 경우에 따라서는 영상을 획득한 다른 장치로부터 네트워크를 통해 영상을 전송받는 통신 모듈로 구현될 수도 있다.
- [0025] 타겟 이미지 획득부(200)는 영상 획득부(100)에서 획득된 영상의 각 프레임 이미지( $I_0 \sim I_n$ )에서 타겟 객체가 포함된 영역을 타겟 이미지로 추출하여 획득한다. 타겟 이미지 획득부(200)는 다수 프레임 이미지( $I_0 \sim I_n$ ) 중 최초 프레임 이미지( $I_0$ )에 대해서는 타겟이 되는 객체가 포함된 타겟 객체 영역이 별도로 설정되어 타겟 이미지를 획득하고, 나머지 프레임 이미지( $I_1 \sim I_n$ )에 대해서는 객체 검출부(800)에 의해 검출된 타겟 객체 위치에 따라 타겟 객체 영역이 선택되어 타겟 이미지를 획득한다.
- [0026] 타겟 이미지 획득부(200)는 타겟 객체 설정부(210)와 타겟 이미지 추출부(220)를 포함할 수 있다. 타겟 객체 설정부(210)는 영상의 다수 프레임 이미지( $I_0 \sim I_n$ ) 중 초기 프레임 이미지( $I_0$ )에서 타겟 객체 영역을 설정한다. 이때 타겟 객체 설정부(210)는 사용자 명령에 응답하여 타겟 객체 영역을 설정할 수도 있으나, 초기 프레임 이미지( $I_0$ )를 영상 획득부(100)가 아닌 다른 장치에서 획득하여 인가된 타겟 이미지와 기지정된 방식으로 비교하여, 초기 프레임 이미지( $I_0$ )에서 유사한 것으로 판별되는 영역을 타겟 객체 영역으로 설정할 수도 있다. 이와 같이, 다른 장치에서 획득된 타겟 이미지와 비교하여 초기 프레임 이미지( $I_0$ )의 유사 영역을 판별하는 기술을 재식별(Re-identification) 기술이라 하며, 이는 공지된 기술이므로 여기서는 상세하게 설명하지 않는다.
- [0027] 타겟 이미지 추출부(220)는 우선 초기 프레임 이미지( $I_0$ )에서 타겟 객체 설정부(210)에 의해 설정된 타겟 객체 영역을 타겟 이미지로 추출한다. 그리고 타겟 이미지 추출부(220)는 이후 나머지 프레임 이미지( $I_1 \sim I_n$ ) 각각에 대해서는 객체 검출부(800)에 의해 검출된 타겟 객체 위치에 따라 타겟 객체 영역을 추출하여 타겟 이미지( $TI_0 \sim TI_n$ )를 획득한다. 여기서 타겟 이미지 추출부(220)가 추출하는 타겟 이미지는 다음 프레임 이미지에서 타겟 객체를 검출하기 위한 기준 특징을 추출하기 위해 획득되는 이미지이므로, 타겟 이미지 추출부(220)는 타겟 객체가 검출되어야 하는 현재 프레임 이미지( $I_t$ )가 아니라, 이전 이미 타겟 객체가 추출된 이전 프레임 이미지( $I_{t-1}$ )에서 타겟 이미지( $TI_{t-1}$ )를 추출한다.
- [0028] 즉 본 실시예에서 타겟 이미지 추출부(220)는 영상을 구성하는 다수 프레임 이미지( $I_0 \sim I_n$ ) 중 초기 프레임 이미지( $I_0$ )에서는 타겟 객체 설정부(210)에 의해 미리 지정된 방식으로 타겟 이미지( $TI_0$ )를 추출하여 획득하는 반

면, 나머지 프레임 이미지( $I_1 \sim I_n$ )에 대해서는 본 실시예에 따른 객체 검출 기법에 따라 검출된 객체 위치를 기반하여 타겟 이미지( $TI_0 \sim TI_n$ )를 추출한다.

[0029] 프레임 특징 추출부(300)는 영상 획득부(100)에서 획득된 영상의 다수 프레임 이미지( $I_0 \sim I_n$ ) 중 현재 타겟 객체가 검출되어야 하는 현재 프레임 이미지( $I_t$ )를 인가받고, 인가된 프레임 이미지( $I_t$ )에 대해 미리 학습된 패턴 추정 방식에 따라 특징을 추출하여, 프레임 특징맵( $F_t$ )을 획득한다. 그리고 타겟 특징 추출부(400)는 타겟 이미지 추출부(220)에서 추출된 타겟 이미지( $TI_{t-1}$ )에 대해 미리 학습된 패턴 추정 방식에 따라 특징을 추출하여, 타겟 특징맵( $F_{t-1}$ )을 획득한다.

[0030] 도 2에 도시된 바와 같이, 타겟 이미지 추출부(220)는 이미 타겟 객체가 검출된 이전 프레임 이미지( $I_0 \sim I_{t-1}$ )에서 추출된 타겟 이미지( $TI_0 \sim TI_{t-1}$ )에 대해 특징을 추출하여 타겟 특징맵( $F_0 \sim F_{t-1}$ )을 획득하고, 프레임 특징 추출부(300)가 현재 프레임 이미지( $I_t$ )로부터 타겟 특징맵( $F_t$ )을 획득하는 것을 알 수 있다. 도 2에서는 타겟 이미지 추출부(220)가 다수의 타겟 이미지( $TI_0 \sim TI_{t-1}$ )로부터 다수의 타겟 특징맵( $F_0 \sim F_{t-1}$ )을 병렬로 추출하는 것처럼 도시하였으나, 이는 후술하는 자기 집중 특징 추출부(600)의 동작을 설명하기 위한 것으로, 타겟 이미지 추출부(220)는 직전 하나의 프레임 이미지( $I_{t-1}$ )의 타겟 이미지( $TI_{t-1}$ )로부터 타겟 특징맵( $F_{t-1}$ )을 획득한다.

[0031] 여기서 프레임 특징 추출부(300)와 타겟 특징 추출부(400)는 패턴 추정 방식이 미리 학습된 CNN 등과 같은 인공 신경망으로 구현될 뿐만 아니라, 동일 구조를 갖고 동일하게 학습되어 동일한 가중치를 갖는 삼 신경망(Siamese network)으로 구현된다. 프레임 특징 추출부(300)와 타겟 특징 추출부(400)가 삼 신경망으로 구현되는 것은, 프레임 특징 추출부(300)와 타겟 특징 추출부(400)가 각각 현재 프레임 이미지( $I_t$ )와 이전 프레임 이미지( $I_{t-1}$ )의 일부 영역으로 추출된 타겟 이미지에 대해 동일한 방식으로 특징을 추출하여 특징맵( $F_t, F_{t-1}$ )을 획득함으로써 이후 특징맵 사이의 연관성을 신뢰도 있게 비교할 수 있도록 하기 위함이다.

[0032] 여기서 특징맵( $F_t, F_{t-1}$ )을 획득하는 프레임 특징 추출부(300)와 타겟 특징 추출부(400)의 동작은 함수( $\phi$ )로 표현될 수 있다.

[0033] 듀얼 집중 특징 추출부(500)는 타겟 특징 추출부(400)에서 추출된 타겟 특징맵( $F_{t-1}$ )을 인가받고, 인가된 타겟 특징맵( $F_{t-1}$ )에서 미리 학습된 패턴 추정 방식에 따라 각 채널별 중요도와 공간적 중요도를 추정하고, 추정된 채널별 중요도와 공간적 중요도를 타겟 특징맵( $F_{t-1}$ )에 가중하여 가중 특징맵( $F'_{t-1}$ )을 획득한다.

[0034] 도 3 및 도 4는 도 1의 듀얼 집중 특징 추출부의 구성과 동작을 설명하기 위한 도면이다.

[0035] 도 3에서 (a)는 듀얼 집중 특징 추출부의 구성을 나타내고, (b)는 듀얼 집중 특징 추출부의 개략적 동작을 나타낸다.

[0036] 도 3의 (a) 및 (b)를 참조하면, 듀얼 집중 특징 추출부(500)는 채널 집중 특징 추출부(510)와 공간 집중 특징 추출부(520)를 포함할 수 있다. 채널 집중 특징 추출부(510)는 타겟 특징맵( $F_{t-1}$ )에서 미리 학습된 패턴 추정 방식에 따라 각 채널별 중요도를 나타내는 채널 집중맵( $M_c$ )을 추정하고, 추정된 채널 집중맵( $M_c$ )을 타겟 특징맵( $F_{t-1}$ )에 가중함으로써, 채널 가중 특징맵을 획득한다.

[0037] 그리고 공간 집중 특징 추출부(520)는 채널 집중 특징 추출부(510)에서 획득된 채널 가중 특징맵을 인가받고, 인가된 채널 가중 특징맵에서 미리 학습된 패턴 추정 방식에 따라 각 공간적 위치 중요도를 나타내는 공간 집중맵( $M_s$ )을 추정하고, 추정된 공간 집중맵( $M_s$ )을 채널 가중 특징맵에 가중함으로써, 가중 특징맵( $F'_{t-1}$ )을 획득한다.

[0038] 도 4의 (a)는 듀얼 집중 특징 추출부(500)에서 채널 집중 특징 추출부(510)와 공간 집중 특징 추출부(520)의 개략적 구조를 나타내고, (b)는 (a)에 도시된 채널 집중 모듈(Channel Attention Module)(511)의 상세 구성을 나타내며, (c)는 (a)에 도시된 공간 집중 모듈(Spatial Attention Module)(521)의 상세 구성을 나타낸다.

[0039] 도 4의 (a)를 참조하면, 채널 집중 특징 추출부(510)는 채널 집중 모듈(511)과 채널 가중부(512)를 포함하고, 공간 집중 특징 추출부(520)는 공간 집중 모듈(521)과 공간 가중부(522)를 포함할 수 있다. (b)에 도시된 바와

같이, 채널 집중 모듈(511)은 인가되는 타겟 특징맵( $F_{t-1}$ )에서 학습된 패턴 추정 방식에 따라 채널 축 방향으로 중요 영역에 높은 가중치를 부여하고 중요하지 않은 영역에는 낮은 가중치를 부여하여, 이후 채널 축 방향에서 중요 영역에 더 많은 관심이 집중되도록 한다.

[0040] 타겟 특징 추출부(400)가 높이(H)와 폭(W)을 갖는 2차원의 타겟 이미지( $TI_{t-1}$ )에 대해 특징을 추출하는 과정에서 채널(C)이 더 추가되어 타겟 특징맵( $F_{t-1}$ )은 높이(H)와 폭(W) 및 채널(C)을 축으로 갖는 3차원 구조( $F_{t-1} \in \mathbb{R}^{C \times H \times W}$ )를 갖는다. 여기서 채널(C)은 2차원의 타겟 이미지( $TI_{t-1}$ )의 각 픽셀에 대해 서로 다른 특징을 추출한 결과이지만 픽셀에 따라 모든 채널(C) 방향의 값이 중요한 가치를 갖지는 않는다. 이에 채널 집중 모듈(511)은 타겟 특징맵( $F_{t-1}$ )의 채널 축 방향에서 중요 영역과 비중요 영역을 나타내는 채널 집중맵( $M_c$ )을 추정한다.

[0041] 도 4의 (b)를 참조하면 채널 집중 모듈(511)은 풀링부, 다중 퍼셉트론(Multi-Layer Perceptron: MLP)부 및 활성화부를 포함할 수 있다. 풀링부는 인가되는 입력 특징인 타겟 특징맵( $F_{t-1}$ )의 채널 축 방향으로의 다수의 채널 벡터 각각에 대해 맥스 풀링(MaxPool) 및 평균값 풀링(AvgPool)을 수행하고, 다중 퍼셉트론부는 2가지 풀링 결과를 기반으로 학습된 패턴 추정 방식에 따라 각 2가지 방식으로 풀링된 채널 벡터에서 채널별 중요도를 추출한다. 그리고 추출된 2가지 채널별 중요도를 서로 가산한 후, 활성화부가 기지정된 활성화 함수(Activation Function)로 활성화함으로써, 채널 집중맵( $M_c \in \mathbb{R}^{C \times 1 \times 1}$ )을 획득할 수 있다. 즉 타겟 특징맵( $F_{t-1}$ )의 다수의 채널 벡터의 최대값과 평균값에 따른 특징을 기반으로 채널 집중맵( $M_c$ )을 획득할 수 있다. 여기서 활성화 함수로는 일 예로 시그모이드(Sigmoid) 함수가 이용될 수 있다.

[0042] 채널 집중 모듈(511)에서 채널 집중맵( $M_c$ )이 획득되면, 채널 가중부(512)는 타겟 특징맵( $F_{t-1}$ )의 각 채널 벡터에 획득된 채널 집중맵( $M_c$ )을 곱하여 채널 집중 특징맵( $CF_{t-1} \in \mathbb{R}^{C \times H \times W}$ )을 획득한다.

[0043] 도 4의 (c)를 참조하면, 공간 집중 모듈(521)은 풀링부와 공간 특징 추출부 및 활성화부를 포함할 수 있다. 풀링부는 채널 집중 모듈(511)에서 획득된 채널 집중 특징맵( $CF_{t-1}$ )에서 높이(H)와 폭(W)의 2차원 공간 벡터단위로 맥스 풀링(MaxPool) 및 평균값 풀링(AvgPool)을 수행하고, 공간 특징 추출부는 2가지 풀링 결과를 결합하고, 결합된 풀링 결과에 대해 학습된 패턴 추정 방식에 따라 특징을 추출한 후, 기지정된 활성화 함수로 활성화함으로써, 공간 집중맵( $M_s$ )을 획득할 수 있다. 여기서도 활성화 함수로는 일 예로 시그모이드(Sigmoid) 함수가 이용될 수 있다.

[0044] 여기서 공간 집중 모듈(521)이 공간적으로 중요한 영역과 중요하지 않은 영역을 구분하여 2차원 공간 상에서 서로 다른 가중치를 갖는 공간 집중맵( $M_s$ )을 획득하는 것은 타겟 이미지( $TI_{t-1}$ )에서 실제 객체에 해당하는 영역과 배경을 구분하고, 객체에 해당하는 영역에서도 객체의 특성이 반영된 영역을 강조할 수 있도록 하기 위함이다.

[0045] 비록 타겟 이미지( $TI_{t-1}$ )가 프레임 이미지에서 타겟 객체가 포함된 영역을 추출한 이미지일지라도, 타겟 이미지( $TI_{t-1}$ )에는 객체 영역뿐만 아니라 배경 영역이 함께 포함된다. 따라서 객체를 정확하게 추적할 수 있도록 하기 위해서는 타겟 이미지( $TI_{t-1}$ )에서 추출되는 특징에서 배경 보다는 객체 영역이 더 강조되는 되어야 한다. 또한 객체 영역에서도 객체의 특성을 정확하게 나타내는 영역과 특성을 나타내기 어려운 영역이 존재하며, 이에 객체의 특성을 정확하게 나타낼 수 있는 영역이 강조되는 것이 바람직하다. 이에 공간 집중 모듈(521)은 타겟 이미지( $TI_{t-1}$ )에서 객체의 특성이 나타나는 영역을 보다 강조되고 배경은 영향을 미치지 못하도록 공간 집중맵( $M_s$ )을 획득한다.

[0046] 공간 집중 모듈(521)에서 공간 집중맵( $M_s$ )이 획득되면, 공간 가중부(522)는 채널 집중 특징맵( $CF_{t-1}$ )과 획득된 공간 집중맵( $M_s$ )을 곱하여 채널 및 공간적 위치에 따른 가중치가 적용된 가중 특징맵( $F'_{t-1}$ )을 획득한다.

[0047] 자기 집중 특징 추출부(600)는 듀얼 집중 특징 추출부(500)에서 획득된 가중 특징맵( $F'_{t-1}$ )을 인가받고, 이전 인가된 가중 특징맵( $F'_0 \sim F'_{t-2}$ )과의 누적 가중 특징을 추출하여 기준 특징맵( $f_t$ )을 획득한다. 여기서 자기 집중 특징 추출부(600)는 자기 집중(Self-attention) 방식에 따라 시간의 흐름에 따라 인가되는 가중 특징맵( $F'_0 \sim F'_{t-1}$ )의 특징을 누적하되, 대응하는 프레임 이미지( $I_0 \sim I_{t-1}$ )의 중요도에 따른 가중치를 가중하여 기준 특징맵

( $f_t$ )을 획득한다.

[0048] 도 5는 도 1의 자기 집중 특징 추출부의 구성과 동작을 설명하기 위한 도면이다.

[0049] 도 5에서 (a)는 자기 집중 특징 추출부(600)의 구성을 나타내고, (b)는 자기 집중 특징 추출부(600)의 동작을 나타낸다. 도 5를 참조하면, 자기 집중 특징 추출부(600)는 키 벡터 획득부(610), 값 벡터 획득부(620), 쿼리 벡터 획득부(630) 및 자기 집중부(640)를 포함할 수 있다.

[0050] 키 벡터 획득부(610)와 값 벡터 획득부(620)는 각각 가중 특징맵( $F'_{t-1}$ )을 인가받고, 인가된 가중 특징맵( $F'_{t-1}$ )에 대해 미리 학습된 패턴 추정 방식에 따라 각 시간대별 시점에서 타겟 객체의 특징을 나타내는 키 벡터(K)와 각 시간대 별 타겟 객체 그 자체를 표현하는 값 벡터(V)를 추출한다.

[0051] 상기한 바와 같이, 가중 특징맵( $F'_{t-1}$ )은 이전 시점( $t-1$ )에서의 프레임 이미지( $I_{t-1}$ )를 기반으로 추출된 특징맵이며, 타겟 객체의 특징뿐만 아니라 타겟 객체의 표현값이 포함된다. 이에 키 벡터 획득부(610)는 인가된 가중 특징맵( $F'_{t-1}$ )에서 타겟 객체의 특징을 키 벡터(K)로 추출하고, 값 벡터 획득부(620)는 가중 특징맵( $F'_{t-1}$ )에서 타겟 객체의 표현값을 값 벡터(K)로 추출한다. 여기서 키 벡터 획득부(610)와 값 벡터 획득부(620)는 일 예로서 서로 다른 가중치를 갖도록 학습된  $1 \times 1$  크기의 컨볼루션 레이어로 구현될 수 있다.

[0052] 쿼리 획득부(630)는 자기 집중부(640)에서 이전 획득된 기준 특징맵( $f_{t-1}$ )을 인가받고, 인가된 이전 시점 기준 특징맵( $f_{t-1}$ )에 대해 키 벡터 획득부(610)와 유사하게 미리 학습된 패턴 추정 방식에 따라 이전 시점( $t-2$ )까지 가중 누적 추출된 타겟 객체의 특징을 나타내는 쿼리 벡터(Q)를 추출한다. 여기서 쿼리 벡터(Q)는 키 벡터(K)와의 비교를 위해 추출되는 특징 벡터이다. 그리고 쿼리 획득부(630)는 키 벡터 획득부(610)에 대응하여 학습에 의한 가중치가 적용된  $1 \times 1$  크기의 컨볼루션 레이어로 구현될 수 있다.

[0053] 자기 집중부(640)는 키 벡터 획득부(610)와 값 벡터 획득부(620) 및 쿼리 획득부(630)에서 획득된 키 벡터(K)와 값 벡터(V) 및 쿼리 벡터(Q)를 인가받고, 키 벡터(K)와 쿼리 벡터(Q) 사이의 관계에 따른 가중치를 값 벡터(V)에 가중하여 기준 특징맵( $f_t$ )을 획득한다.

[0054] 자기 집중부(640)는 수학적 식 1에 따라 기준 특징맵( $f_t$ )을 획득할 수 있다.

### 수학적 식 1

$$f_t = \text{Softmax}(QK^T)V$$

[0055]

[0056] 여기서  $K_t$ 는 키 벡터(K)의 전치 행렬을 나타낸다.

[0057] 즉 자기 집중부(640)는 이전까지의 가중 특징맵( $F'_0 \sim F'_{t-2}$ )의 중요도에 따라 가중 누적 추출된 타겟 객체의 특징을 나타내는 쿼리 벡터(Q)에 대비한 가중 특징맵( $F'_{t-1}$ )의 특징을 나타내는 키 벡터(K)의 중요도를 계산하여 값 벡터(V)에 가중함으로써, 기준 특징맵( $f_t$ )을 획득한다.

[0058] 여기서 초기 기준 특징맵( $f_0$ )은 초기 타겟 이미지( $TI_0$ )로부터 타겟 특징 추출부(400)가 획득한 타겟 특징맵( $F_0$ )이 사용될 수 있다.

[0059] 상기한 바와 같이, 자기 집중 특징 추출부(600)가 계속적으로 인가되는 가중 특징맵( $F'_0 \sim F'_{t-2}$ )의 중요도에 따라 값 벡터(V)를 가중하여 기준 특징맵( $f_t$ )을 획득하는 것은 다수 프레임( $I_0 \sim I_{t-1}$ )에서 타겟 객체에 대응하여 획득된 타겟 이미지( $TI_0 \sim TI_{t-1}$ )가 불완전할 수 있기 때문이다. 다수 프레임( $I_0 \sim I_n$ )으로 구성되는 영상에서는 일부 프레임에서 타겟 객체 영역에 폐색(occlusion) 등이 발생되어 타겟 객체가 불명확하게 나타나는 경우가 발생할 수 있다. 이런 경우, 특정 시점에서는 타겟 이미지( $TI_{t-1}$ )로부터 타겟 특징맵( $F_{t-1}$ )을 추정하고, 가중 특징맵( $F'_{t-1}$ )을 획득할 지라도, 해당 시점에서의 가중 특징맵( $F'_{t-1}$ )에는 타겟 객체의 특징이 제대로 반영되지 않을 수 있다.



- [0060] 이에 자기 집중 특징 추출부(600)는 타겟 객체의 특징이 정확하게 반영되지 않은 가중 특징맵이 적게 반영되고, 타겟 객체의 특징이 제대로 반영된 가중 특징맵은 크게 반영되도록 다수의 가중 특징맵을 가중 누적하여 기준 특징맵( $f_t$ )을 획득한다. 즉 이전 다수의 프레임 이미지( $I_0 \sim I_{t-1}$ )에 대응하는 다수의 타겟 이미지( $TI_0 \sim TI_{t-1}$ ) 중 타겟 객체의 특징이 잘 나타난 타겟 이미지에 더 큰 가중치를 부여하여 누적하므로, 타겟 이미지의 중요 특징이 더욱 명확하게 표현되도록 한다.
- [0061] 특징 상관부(700)는 프레임 특징 추출부(300)에서 획득된 프레임 특징맵( $F_t$ )과 자기 집중 특징 추출부(600)에서 획득된 기준 특징맵( $f_t$ )을 인가받고, 프레임 특징맵( $F_t$ )과 기준 특징맵( $f_t$ )에 대한 상관 연산을 수행하여, 프레임 특징맵( $F_t$ )과 기준 특징맵( $f_t$ ) 사이의 위치별 유사도를 나타내는 스코어 맵을 획득한다.
- [0062] 객체 검출부(800)는 특징 상관부(700)에서 획득된 스코어 맵에서 가장 높은 최고 스코어의 위치를 탐색하고, 탐색된 위치를 기반으로 현재 프레임 이미지( $I_t$ )에서 타겟 객체의 위치를 판별한다. 이때 객체 검출부(800)는 미리 지정된 경계 박스의 크기에 따라 판별된 위치를 기준으로 기지정된 크기를 타겟 객체 영역으로 지정하여 타겟 이미지 추출부(220)로 전달함으로써, 타겟 이미지 추출부(220)가 프레임 이미지( $I_t$ )에서 타겟 이미지( $TI_t$ )를 추출하도록 할 수 있다.
- [0063] 상기한 바와 같이, 본 실시예에 따른 객체 추적 장치에서는 이전 프레임 이미지( $I_0 \sim I_{t-1}$ )에 대응하여 추출된 타겟 이미지( $TI_0 \sim TI_{t-1}$ )에서 특징을 추출하여 타겟 특징맵( $F_0 \sim F_{t-1}$ )을 획득하고, 타겟 특징맵( $F_0 \sim F_{t-1}$ )의 채널 및 공간적 중요도에 따른 가중치를 가중하여 타겟 특징맵( $F_0 \sim F_{t-1}$ )에서 주요 특징이 강조된 가중 특징맵( $F'_0 \sim F'_{t-1}$ )을 획득하며, 획득된 가중 특징맵( $F'_0 \sim F'_{t-1}$ )에서 시점별(프레임별) 중요도를 가중하고 누적하여 기준 특징맵( $f_t$ )을 획득한다. 그리고 기준 특징맵( $f_t$ )과 현재 프레임 이미지에서 추출된 프레임 특징맵( $F_t$ )을 비교하여 객체를 추적한다.
- [0064] 따라서 다수 프레임 이미지를 포함하는 영상에서 시간에 따라 타겟 객체의 외형이 변화(포즈 변화, 조명 변화, 회전, 변신, 외부적 요인으로 인한 변화)하거나 일부 프레임 이미지에서 타겟 객체 영역에 폐색이 발생되더라도 기준 특징맵( $f_t$ )이 타겟 객체의 특징을 명확하게 표현할 수 있어 강건하게 타겟 객체를 추적할 수 있다.
- [0065] 도 6은 본 발명의 일 실시예에 따른 객체 추적 방법을 나타낸다.
- [0066] 도 1 내지 도 5를 참조하여, 도 6의 객체 추적 방법을 설명하면, 먼저 타겟 객체가 추적되어야 할 다수 프레임 이미지( $I_0 \sim I_n$ )로 구성된 영상을 획득한다(S10). 영상이 획득되면, 획득된 영상에서 다수 프레임 이미지( $I_0 \sim I_n$ ) 중 초기 프레임 이미지( $I_0$ )에서 타겟 객체 영역을 설정한다(S20).
- [0067] 그리고 다수 프레임 이미지( $I_0 \sim I_n$ ) 중 타겟 객체가 검출되어야 하는 현재 프레임 이미지( $I_t$ )에 대해 미리 학습된 패턴 추정 방식에 따라 특징을 추출하여, 프레임 특징맵( $F_t$ )을 획득한다(S30).
- [0068] 이와 함께 초기 프레임 이미지( $I_0$ )에서 비설정된 타겟 객체 영역 또는 이전 프레임 이미지( $I_{t-1}$ )에서 검출된 타겟 객체 영역에 따라 타겟 이미지( $TI_{t-1}$ )를 추출한다(S40). 타겟 이미지( $TI_{t-1}$ )가 추출되면, 추출된 타겟 이미지( $TI_{t-1}$ )에 대해 프레임 특징맵( $F_t$ )을 추출하는 방식과 동일하게 학습된 패턴 추정 방식에 따라 특징을 추출하여 타겟 특징맵( $F_{t-1}$ )을 획득한다.
- [0069] 타겟 특징맵( $F_{t-1}$ )이 획득되면, 획득된 타겟 특징맵( $F_{t-1}$ )에서 채널 및 공간적 중요도에 따라 가중치를 가중하여 가중 특징맵( $F'_{t-1}$ )을 획득한다(S60). 가중 특징맵( $F'_{t-1}$ )을 획득하는 단계(S60)에서는 우선, 타겟 특징맵( $F_{t-1}$ )에서 채널 축 방향으로 각 채널별 중요도를 나타내는 채널 집중 특징을 추정하여 채널 집중맵( $M_c$ )을 추출하고, 추출된 채널 집중맵( $M_c$ )을 타겟 특징맵( $F_{t-1}$ )에 가중하여 채널 집중 특징맵( $CF_{t-1}$ )을 획득한다(S61). 그리고 채널 집중 특징맵( $CF_{t-1}$ )에서 높이(H)와 폭(W)의 2차원 공간 위치에 따른 공간적 중요도를 나타내는 공간 집중 특징을 추정하여 공간 집중맵( $M_s$ )을 추출하고, 추출된 공간 집중맵( $M_s$ )을 채널 집중 특징맵( $CF_{t-1}$ )에 가중하여 가중 특징맵( $F'_{t-1}$ )을 획득한다(S62).

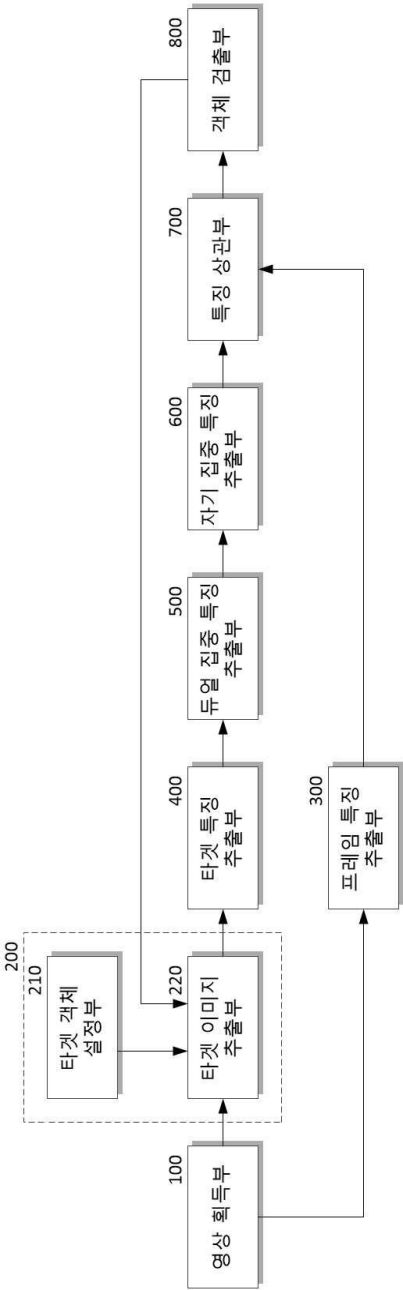
- [0070] 가중 특징맵( $F'_{t-1}$ )이 획득되면, 현재까지 획득된 가중 특징맵( $F'_0 \sim F'_{t-1}$ )을 가중 누적하여 특징 추출된 기준 특징맵( $f_t$ )을 획득한다(S70).
- [0071] 여기서 기준 특징맵( $f_t$ )을 획득하는 단계(S70)는 먼저, 획득된 가중 특징맵( $F'_{t-1}$ )으로부터 각 시간대별 시점에 서 타겟 객체의 특징을 나타내는 키 벡터(K)와 각 시간대 별 타겟 객체 그 자체를 표현하는 값 벡터(V)를 추출 하고, 이전 획득된 기준 특징맵( $f_{t-1}$ )으로부터, 이전 시점( $t-2$ )까지 가중 누적 추출된 타겟 객체의 특징을 나타 내는 쿼리 벡터(Q)를 추출한다(S71). 그리고 추출된 키 벡터(K)와 쿼리 벡터(Q) 사이의 관계에 따른 가중치를 값 벡터(V)에 가중하여 기준 특징맵( $f_t$ )을 획득한다(S72).
- [0072] 기준 특징맵( $f_t$ )이 획득되면, 프레임 특징맵( $F_t$ )과 기준 특징맵( $f_t$ )에 대한 상관 연산을 수행하여 스코어 맵을 획득한다(S80). 이후 획득된 스코어 맵에서 가장 높은 최고 스코어의 위치를 탐색하고, 탐색된 위치를 기반으로 현재 프레임 이미지( $I_t$ )에서 타겟 객체의 위치를 판별한다(S90).
- [0073] 본 발명에 따른 방법은 컴퓨터에서 실행시키기 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 비분 리형 매체를 모두 포함하며, ROM(판독 전용 메모리), RAM(랜덤 액세스 메모리), CD(컴팩트 디스크)-ROM, DVD(디 지털 비디오 디스크)-ROM, 자기 테이프, 플로피 디스크, 광데이터 저장장치 등을 포함할 수 있다.
- [0074] 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.
- [0075] 따라서, 본 발명의 진정한 기술적 보호 범위는 첨부된 청구범위의 기술적 사상에 의해 정해져야 할 것이다.

## 부호의 설명

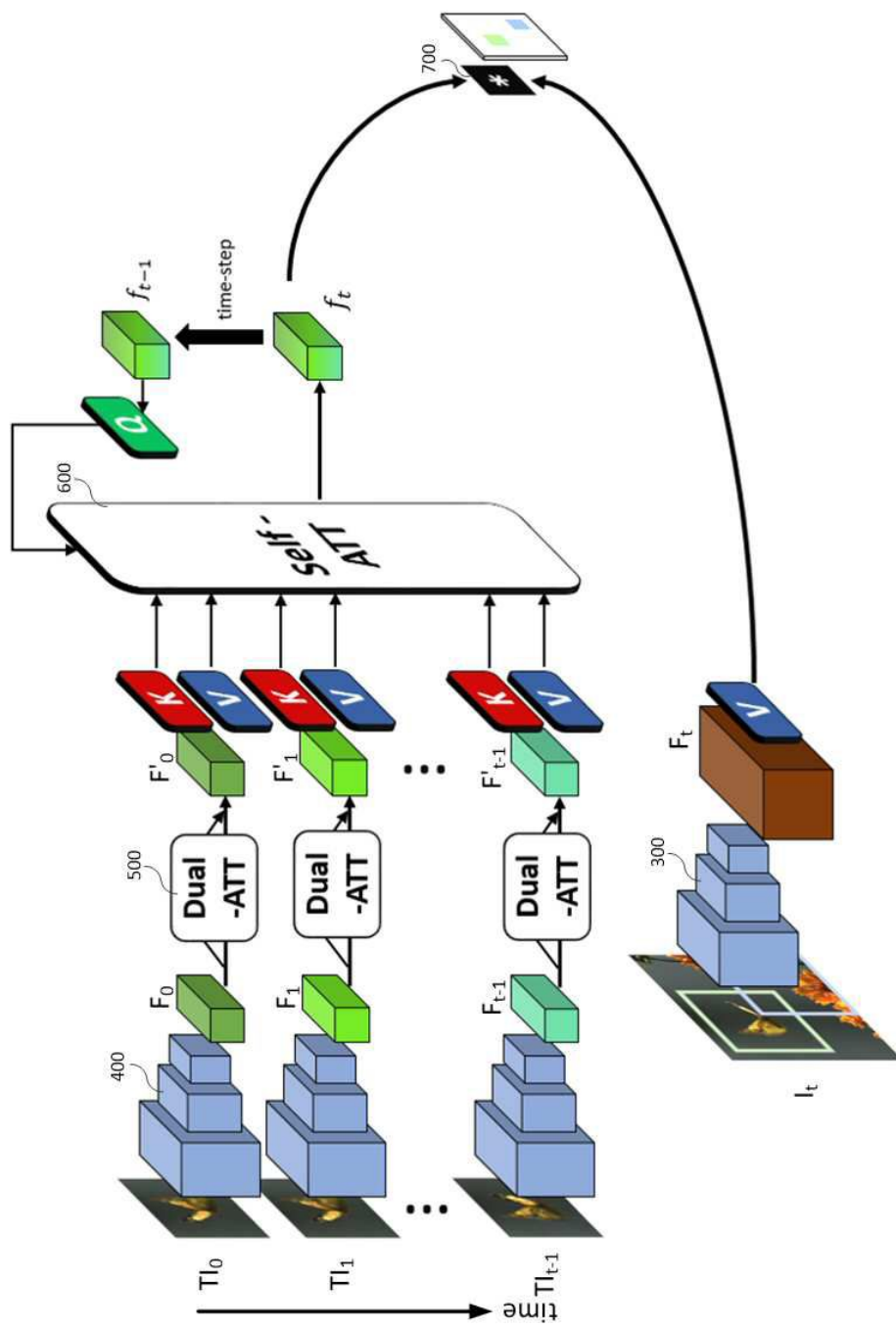
- |        |                   |                   |
|--------|-------------------|-------------------|
| [0076] | 100: 영상 획득부       | 200: 타겟 이미지 획득부   |
|        | 300: 프레임 특징 추출부   | 400: 타겟 특징 추출부    |
|        | 500: 듀얼 집중 특징 추출부 | 600: 자기 집중 특징 추출부 |
|        | 700: 특징 상관부       | 800: 객체 검출부       |
|        | 510: 채널 집중 특징 추출부 | 520: 공간 집중 특징 추출부 |
|        | 610: 키 벡터 획득부     | 620: 값 벡터 획득부     |
|        | 630: 쿼리 벡터 획득부    | 640: 가지 집중부       |

도면

도면1

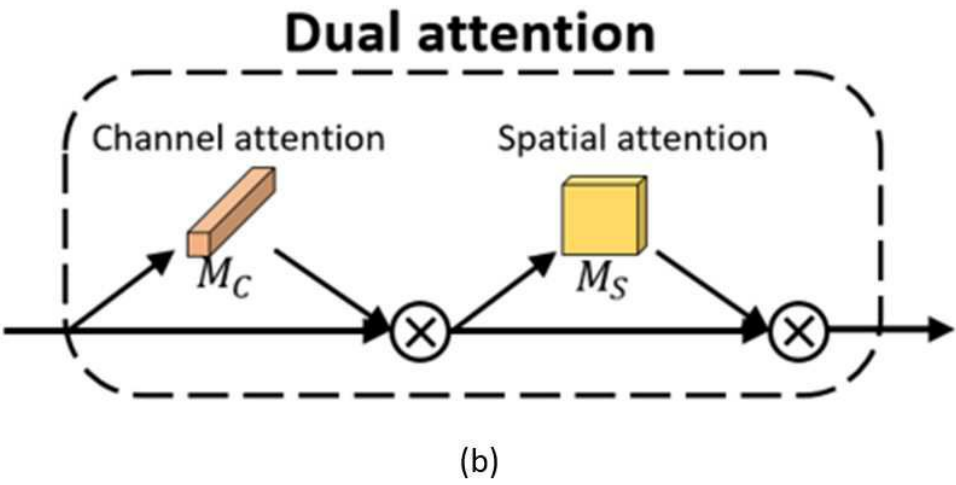
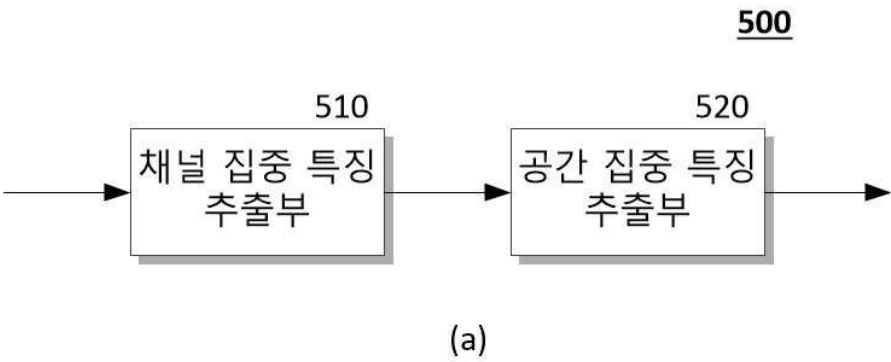


도면2

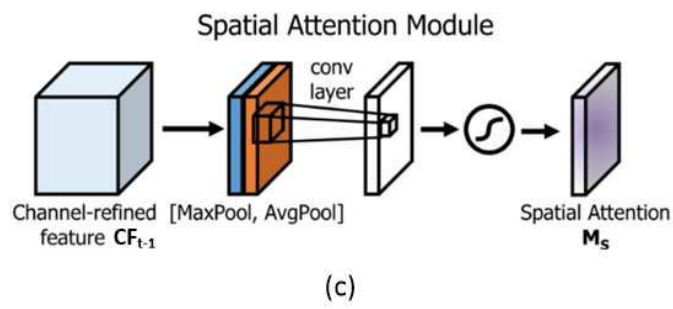
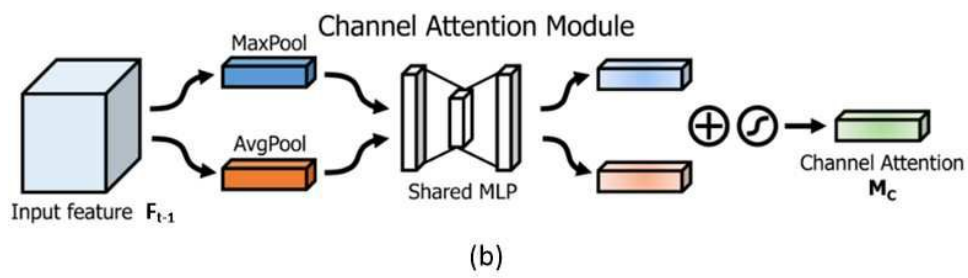
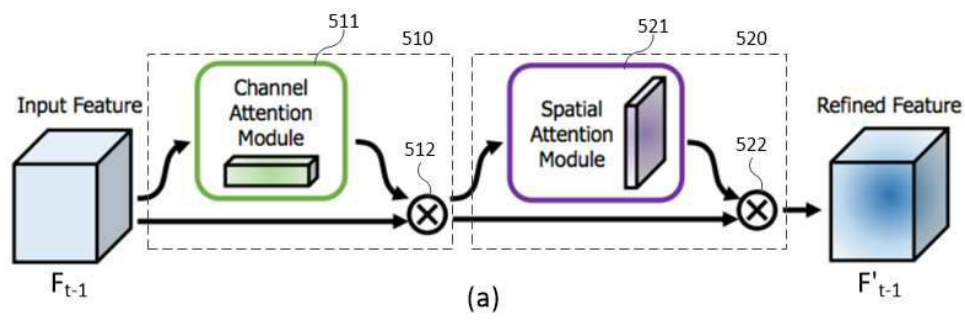




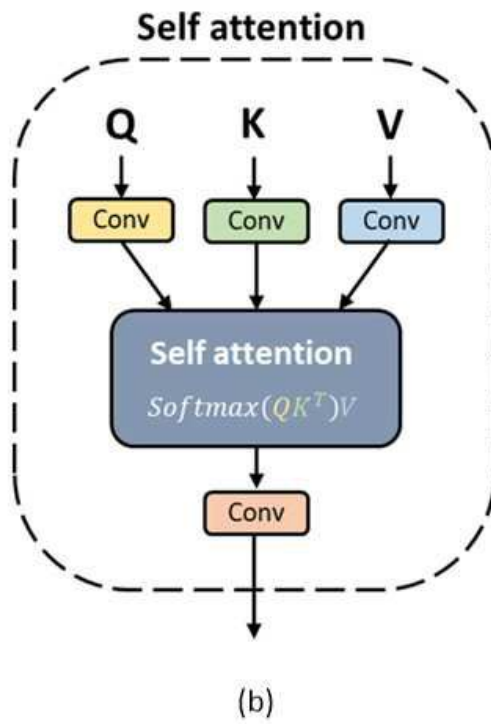
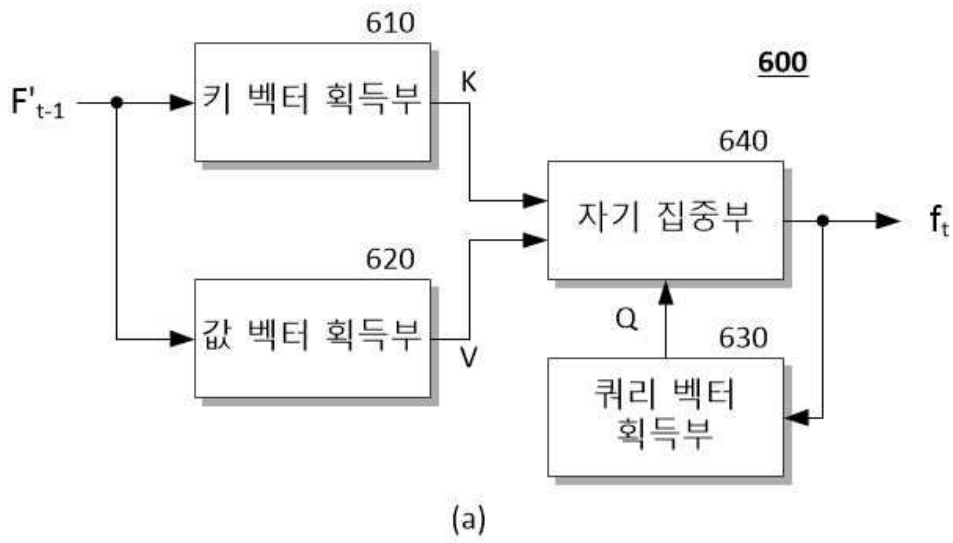
도면3



도면4



도면5



도면6

