



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2022년08월18일

(11) 등록번호 10-2433631

(24) 등록일자 2022년08월12일

(51) 국제특허분류(Int. Cl.)
G06V 10/10 (2022.01) G06N 3/04 (2006.01)
G06N 3/08 (2006.01) G06T 3/00 (2019.01)
G06T 5/00 (2019.01) G06V 10/40 (2022.01)

(52) CPC특허분류
G06V 10/10 (2022.01)
G06N 3/04 (2013.01)

(21) 출원번호 10-2020-0109800

(22) 출원일자 2020년08월31일

심사청구일자 2020년08월31일

(65) 공개번호 10-2022-0028604

(43) 공개일자 2022년03월08일

(56) 선행기술조사문헌

KR101995294 B1*

*는 심사관에 의하여 인용된 문헌

(73) 특허권자

연세대학교 산학협력단

서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)

(72) 발명자

손광훈

서울특별시 서대문구 연세로 50, 연세대학교 제3공학관 C129호(신촌동)

김한재

서울특별시 서대문구 연세로 50, 연세대학교 제3공학관 C129호(신촌동)

(74) 대리인

민영준

전체 청구항 수 : 총 18 항

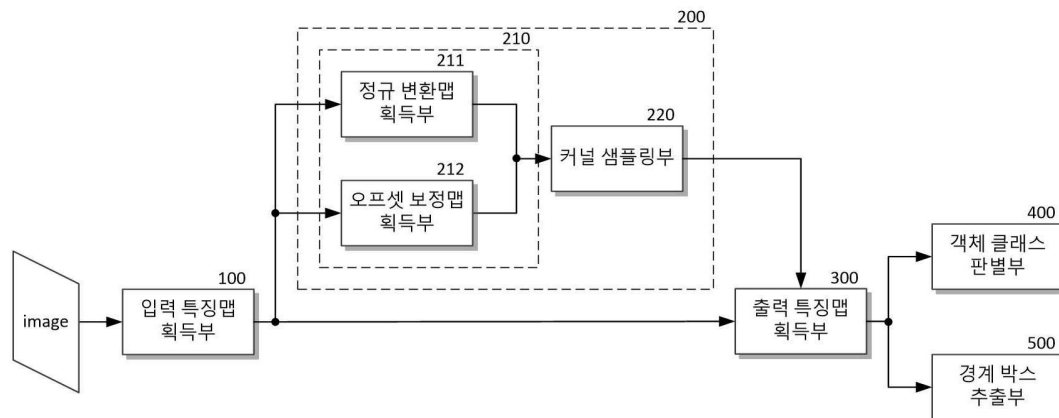
심사관 : 전한철

(54) 발명의 명칭 동적 커널 기반 객체 검출 장치 및 방법

(57) 요약

본 발명은 입력 특징맵에서 객체 특징맵을 추출하지 않는 단일 스테이지 방식에서 객체의 구조적 특징에 대응하여 가중치가 동적으로 변화되는 동적 커널을 이용하여 객체를 판별하고 경계 박스를 추정함으로써, 픽셀이 갖는 객체 정보의 누락없이 다양한 크기 및 형상을 갖는 객체를 정확하게 판별하고 경계 박스를 추정할 수 있는 객체 검출 장치 및 방법을 제공할 수 있다.

대표도



(52) CPC특허분류

G06N 3/08 (2013.01)
G06T 3/0093 (2013.01)
G06T 5/001 (2013.01)
G06V 10/40 (2022.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711094090
과제번호	2018M3E3A1057289
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	복합인지기술개발사업
연구과제명 (2단계)(1/3)	(2세부)이중 CCTV 영상에서의 딥러닝 기반 실종자 초동 신원확인 및 추적 시스템
기 여 율	1/1
과제수행기관명	연세대학교 산학협력단
연구기간	2020.04.23 ~ 2020.12.31

명세서

청구범위

청구항 1

패턴 추정 방식이 미리 학습된 인공 신경망으로 구현되어, 인가된 입력 이미지의 특징을 추출하여 입력 특징맵을 획득하는 입력 특징맵 획득부;

패턴 추정 방식이 미리 학습된 인공 신경망으로 구현되어, 상기 입력 특징맵에서 객체 검출을 위한 특징을 추출하기 위해 학습에 의해 미리 획득된 메인 커널에서 샘플링될 기지정된 개수의 위치 정보가 미리 지정된 원소로 포함된 디폴트 파라미터를 기지정된 다수의 스케일 파라미터 중 하나의 스케일 파라미터에 따라 스케일 변환하고, 스케일 변환된 정규 파라미터의 원소들이 나타내는 위치 정보 각각의 오프셋을 보정하여 상기 입력 특징맵의 다수의 픽셀 각각에 대응하는 변환 파라미터를 획득하고, 획득된 상기 변환 파라미터를 기반으로 상기 메인 커널을 기지정된 방식으로 샘플링하여, 상기 입력 특징맵의 다수의 픽셀 각각에 따라 적응적으로 가변되는 다수의 동적 커널을 추출하는 동적 커널 추출부;

상기 입력 특징맵과 상기 입력 특징맵의 각 픽셀에 대응하는 다수의 동적 커널을 기지정된 방식으로 연산하여 출력 특징맵을 획득하는 출력 특징맵 획득부; 및

패턴 추정 방식이 미리 학습된 인공 신경망으로 구현되어, 상기 출력 특징맵의 패턴을 추정하여 상기 입력 이미지에서 객체가 포함된 영역을 나타내는 경계 박스와 상기 경계 박스에 포함된 객체를 판별하는 객체 및 경계 박스 추출부를 포함하되,

상기 동적 커널 추출부는

미리 학습된 패턴 추정 방식에 따라 상기 입력 특징맵의 패턴을 추정하여, 상기 디폴트 파라미터를 상기 입력 특징맵의 다수의 픽셀 각각에 대응하는 패턴으로 스케일 변환 및 오프셋 보정하기 위한 변환 파라미터맵을 획득하는 변환 파라미터 추출부; 및

상기 변환 파라미터맵으로부터 상기 입력 특징맵의 다수의 픽셀 각각에 대응하여 상기 디폴트 파라미터의 각 원소가 나타내는 위치 정보를 스케일 파라미터와 오프셋에 따라 변경한 다수의 변환 파라미터를 획득하고, 획득된 다수의 변환 파라미터에 따라 미리 획득되어 저장된 상기 메인 커널을 샘플링하여 상기 입력 특징맵의 다수의 픽셀 각각에 대응하는 다수의 동적 커널을 추출하는 커널 샘플링부를 포함하는 객체 검출 장치.

청구항 2

삭제

청구항 3

제1항에 있어서, 상기 변환 파라미터 추출부는

상기 입력 특징맵을 인가받고, 미리 학습된 패턴 추정 방식에 따라 상기 입력 특징맵의 패턴을 추정하여, 상기 입력 특징맵의 각 픽셀에 대응하는 객체의 크기가 반영되는 스케일 파라미터를 추출하기 위한 정규 변환맵을 획득하는 정규 변환맵 획득부; 및

상기 입력 특징맵 또는 상기 정규 변환맵 중 적어도 하나를 인가받고, 미리 학습된 패턴 추정 방식에 따라 상기 입력 특징맵의 각 픽셀에 대응하는 객체의 형상이 반영되는 오프셋을 추출하기 위한 오프셋 보정맵을 획득하는 오프셋 보정맵 획득부를 포함하는 객체 검출 장치.

청구항 4

제3항에 있어서, 상기 정규 변환맵 획득부는

상기 디폴트 파라미터의 각 원소가 나타내는 위치 정보에 따른 거리를 각각 서로 다른 지정된 스케일로 변환하

기 위해 미리 설정되는 기지정된 개수의 스케일 파라미터에 대응하는 크기로 상기 정규 변환맵을 획득하고,
상기 오프셋 보정맵 획득부는

상기 동적 커널의 크기에 대응하여 미리 지정되는 상기 디폴트 파라미터의 원소 개수에 대응하는 크기로 상기 오프셋 보정맵을 획득하는 객체 검출 장치.

청구항 5

제4항에 있어서, 상기 정규 변환맵과 상기 오프셋 보정맵 각각은

상기 스케일 파라미터와 상기 오프셋을 상기 메인 커널의 행방향과 열방향에 따른 x축 방향과 y축 방향의 위치 정보를 구분하여 스케일 변환 및 오프셋 보정할 수 있도록 스케일 파라미터 개수의 2배에 대응하는 크기와 상기 동적 커널의 크기의 2배에 대응하는 크기로 획득되는 객체 검출 장치.

청구항 6

제5항에 있어서, 상기 스케일 파라미터는

상기 디폴트 파라미터를 스케일 변환한 정규 변환 파라미터의 각 원소가 나타내는 위치 정보가 상기 메인 커널의 크기를 초과하지 않는 값으로 설정되는 객체 검출 장치.

청구항 7

제3항에 있어서, 상기 변환 파라미터 추출부는

상기 정규 변환맵과 상기 오프셋 보정맵을 결합하여 상기 변환 파라미터맵을 출력하는 객체 검출 장치.

청구항 8

제1항에 있어서, 상기 커널 샘플링부는

상기 입력 특징맵의 다수의 픽셀 각각에 대응하는 변환 파라미터의 원소들이 나타내는 위치 정보에 따라 상기 메인 커널의 대응하는 위치의 가중치를 샘플링하여 상기 동적 커널을 추출하는 객체 검출 장치.

청구항 9

제3항에 있어서, 상기 객체 검출 장치는

학습 시에 이미지에 포함된 객체에 대한 객체 클래스와 경계 박스가 미리 주석된 다수의 학습 데이터가 입력 이미지로 인가받아 객체와 경계 박스를 판별하며, 판별된 객체와 경계 박스를 주석된 객체 클래스 및 경계 박스와 비교하여 계산된 손실을 역전파하여 상기 메인 커널을 획득하고, 상기 정규 변환맵 획득부와 상기 오프셋 보정맵 획득부를 학습시키는 객체 검출 장치.

청구항 10

제1항에 있어서, 상기 객체 및 경계 박스 추출부는

패턴 추정 방식이 미리 학습된 인공 신경망으로 구현되어, 상기 출력 특징맵의 패턴을 추정하여 상기 출력 특징맵의 객체 영역을 판별하고, 판별된 객체 영역을 상기 입력 이미지에 대응하는 영역으로 회귀시켜 상기 경계 박스를 추출하는 경계 박스 추출부; 및

기지정된 개수의 객체 클래스별로 구분되어 패턴 추정 방식이 미리 학습된 인공 신경망으로 구현되어, 상기 출력 특징맵의 패턴을 추정하여 상기 객체 영역에 포함된 객체를 판별하는 객체 클래스 판별부를 포함하는 객체 검출 장치.

청구항 11

패턴 추정 방식이 미리 학습된 인공 신경망을 이용하여, 입력 이미지의 특징을 추출하여 입력 특징맵을 획득하는 단계;

패턴 추정 방식이 미리 학습된 인공 신경망을 이용하여, 상기 입력 특징맵에서 객체 검출을 위한 특징을 추출하기 위해 학습에 의해 미리 획득된 메인 커널에서 샘플링될 기지정된 개수의 위치 정보가 미리 지정된 원소로 포

함된 디폴트 파라미터를 기지정된 다수의 스케일 파라미터 중 하나의 스케일 파라미터에 따라 스케일 변환하고, 스케일 변환된 정규 파라미터의 원소들이 나타내는 위치 정보 각각의 오프셋을 보정하여 상기 입력 특징맵의 다수의 픽셀 각각에 대응하는 변환 파라미터를 획득하는 단계;

획득된 상기 변환 파라미터를 기반으로 상기 메인 커널을 기지정된 방식으로 샘플링하여, 상기 입력 특징맵의 다수의 픽셀 각각에 따라 적응적으로 가변되는 다수의 동적 커널을 추출하는 단계;

상기 입력 특징맵과 상기 입력 특징맵의 각 픽셀에 대응하는 다수의 동적 커널을 기지정된 방식으로 연산하여 출력 특징맵을 획득하는 단계; 및

패턴 추정 방식이 미리 학습된 인공 신경망을 이용하여, 상기 출력 특징맵의 패턴을 추정하여 상기 입력 이미지에서 객체가 포함된 영역을 나타내는 경계 박스와 상기 경계 박스에 포함된 객체를 판별하는 단계를 포함하되,

상기 변환 파라미터를 획득하는 단계는

미리 학습된 패턴 추정 방식에 따라 상기 입력 특징맵의 패턴을 추정하여, 상기 디폴트 파라미터를 상기 입력 특징맵의 다수의 픽셀 각각에 대응하는 패턴으로 스케일 변환 및 오프셋 보정하기 위한 변환 파라미터맵을 획득하는 단계; 및

상기 변환 파라미터맵으로부터 상기 입력 특징맵의 다수의 픽셀 각각에 대응하여 상기 디폴트 파라미터의 각 원소가 나타내는 위치 정보를 스케일 파라미터와 오프셋에 따라 변경한 다수의 변환 파라미터를 획득하는 단계를 포함하는 객체 검출 방법.

청구항 12

삭제

청구항 13

제11항에 있어서, 상기 변환 파라미터맵을 획득하는 단계는

미리 학습된 패턴 추정 방식에 따라 상기 입력 특징맵의 패턴을 추정하여, 상기 입력 특징맵의 각 픽셀에 대응하는 객체의 크기가 반영되는 스케일 파라미터를 추출하기 위한 정규 변환맵을 획득하는 단계; 및

상기 입력 특징맵 또는 상기 정규 변환맵 중 적어도 하나를 인가받고, 미리 학습된 패턴 추정 방식에 따라 상기 입력 특징맵의 각 픽셀에 대응하는 객체의 형상이 반영되는 오프셋을 추출하기 위한 오프셋 보정맵을 획득하는 단계를 포함하는 객체 검출 방법.

청구항 14

제13항에 있어서, 상기 정규 변환맵을 획득하는 단계는

상기 디폴트 파라미터의 각 원소가 나타내는 위치 정보에 따른 거리를 각각 서로 다른 지정된 스케일로 변환하기 위해 미리 설정되는 기지정된 개수의 스케일 파라미터에 대응하는 크기로 상기 정규 변환맵을 획득하고,

상기 오프셋 보정맵을 획득하는 단계는

상기 동적 커널의 크기에 대응하여 미리 지정되는 상기 디폴트 파라미터의 원소 개수에 대응하는 크기로 상기 오프셋 보정맵을 획득하는 객체 검출 방법.

청구항 15

제14항에 있어서, 상기 정규 변환맵과 상기 오프셋 보정맵 각각은

상기 스케일 파라미터와 상기 오프셋을 상기 메인 커널의 행방향과 열방향에 따른 x축 방향과 y축 방향의 위치 정보를 구분하여 스케일 변환 및 오프셋 보정할 수 있도록 스케일 파라미터 개수의 2배에 대응하는 크기와 상기 동적 커널의 크기의 2배에 대응하는 크기로 획득되는 객체 검출 방법.

청구항 16

제15항에 있어서, 상기 스케일 파라미터는

상기 디폴트 파라미터를 스케일 변환한 정규 변환 파라미터의 각 원소가 나타내는 위치 정보가 상기 메인 커널의 크기를 초과하지 않는 값으로 설정되는 객체 검출 방법.

청구항 17

제13항에 있어서, 상기 변환 파라미터맵을 획득하는 단계는

상기 정규 변환맵과 상기 오프셋 보정맵을 결합하여 상기 변환 파라미터맵을 출력하는 단계를 더 포함하는 객체 검출 방법.

청구항 18

제11항에 있어서, 상기 동적 커널을 추출하는 단계는

상기 입력 특징맵의 다수의 픽셀 각각에 대응하는 변환 파라미터의 원소들이 나타내는 위치 정보에 따라 미리 획득되어 저장된 상기 메인 커널의 대응하는 위치의 가중치를 샘플링하여 상기 입력 특징맵의 다수의 픽셀 각각에 대응하는 다수의 동적 커널을 추출하는 객체 검출 방법.

청구항 19

제13항에 있어서, 상기 객체 검출 방법은

이미지에 포함된 객체에 대한 객체 클래스와 경계 박스가 미리 주석된 다수의 학습 데이터가 입력 이미지로 인가받아 객체와 경계 박스를 판별하며, 판별된 객체와 경계 박스를 주석된 객체 클래스 및 경계 박스와 비교하여 계산된 손실을 역전파하여 학습을 수행하는 단계를 더 포함하는 객체 검출 방법.

청구항 20

제11항에 있어서, 상기 객체 및 경계 박스 추출부는

패턴 추정 방식이 미리 학습된 인공 신경망으로 구현되어, 상기 출력 특징맵의 패턴을 추정하여 상기 출력 특징맵의 객체 영역을 판별하고, 판별된 객체 영역을 상기 입력 이미지에 대응하는 영역으로 회귀시켜 상기 경계 박스를 추출하는 경계 박스 추출부; 및

기지정된 개수의 객체 클래스별로 구분되어 패턴 추정 방식이 미리 학습된 인공 신경망으로 구현되어, 상기 출력 특징맵의 패턴을 추정하여 상기 객체 영역에 포함된 객체를 판별하는 객체 클래스 판별부를 포함하는 객체 검출 방법.

발명의 설명

기술 분야

[0001] 본 발명은 객체 검출 장치 및 방법에 관한 것으로, 객체에 따라 동적으로 가변되는 동적 커널을 이용하여 객체를 검출하는 동적 커널 기반 객체 검출 장치 및 방법에 관한 것이다.

배경 기술

[0002] 최근 이미지에 포함된 객체를 색별하고 객체가 포함된 위치를 검출하는 객체 검출 기법에는 인공 신경망이 주로 이용되고 있다. 기존에 인공 신경망을 이용하여 객체를 검출하는 기법으로는 단일 스테이지 객체 검출 기법과 다중 스테이지 객체 검출 기법이 이용되고 있다.

[0003] 우선 단일 스테이지 객체 검출 기법과 다중 스테이지 객체 검출 기법은 우선 입력된 이미지에서 특징을 추출하여 획득된 입력 특징맵을 획득한다.

[0004] 이후 다중 스테이지 객체 검출 기법에서는 제1 스테이지가 입력 특징맵에서 적어도 하나의 객체 후보 검출한다. 그리고 제2 스테이지는 검출된 각 객체 후보에 대한 적어도 하나의 경계 박스(bounding box)를 추정하고, 추정된 경계 박스에 의해 설정된 관심 영역이 기지정된 크기가 되도록 관심 영역 풀링(ROI Pooling) 등을 수행하여 객체 특징맵을 획득한다.

- [0005] 이때 다중 스테이지 객체 검출 기법에서는 객체 후보를 검출하기 위한 제1 스테이지를 위한 제1 스테이지 커널과 객체 경계 박스를 추정하기 위한 제2 스테이지 커널이 별도로 구비된다. 또한 추가적인 스테이지 커널이 더 포함되는 경우도 있다.
- [0006] 반면 단일 스테이지 객체 검출 기법에서는 다중 스테이지 객체 검출 기법과 달리 제1 스테이지와 제2 스테이지를 구분하지 않고, 단일 스테이지의 커널을 이용하여 객체 후보를 검출함과 동시에 객체 경계 박스를 추정한다.
- [0007] 그리고 단일 스테이지 객체 검출 기법과 다중 스테이지 객체 검출 기법은 공통적으로 검출된 객체 후보의 객체를 판별하고, 경계 박스를 재설정(refinement)하여 입력 이미지의 영역으로 회귀시킨다.
- [0008] 상기한 바와 같이 다중 스테이지 검출 기법에서는 객체 후보 검출과 경계 상자를 각각 구분하여 추출하므로, 추출된 객체 특징맵에 객체 크기 정보가 포함된다. 따라서 이미지에 포함된 객체의 크기가 매우 다양한 경우에도 객체와 경계 박스를 정확하게 판별하여 검출할 수 있다. 다만 다중 스테이지 검출 기법은 입력 특징맵에서 추정된 경계 박스에 따라 객체 특징맵을 추출하고 추출된 객체 특징맵을 기반으로 객체를 판별하므로, 객체 특징맵 추출에 오차가 포함되는 경우, 입력 특징맵에서 높은 의미론적 객체 정보가 포함된 픽셀이 누락될 수 있다.
- [0009] 반면, 단일 스테이지 객체 검출 기법은 객체 후보를 검출함과 동시에 객체 경계 박스를 추정하기 위한 커널이 구분되지 않으므로, 픽셀의 갖는 객체 정보가 누락될 가능성은 낮다. 반면, 객체의 크기와 형상과 같은 객체의 구조적 특징이 반영되는 객체 특징맵을 추출하지 않고, 지정된 크기의 단일 커널을 이용하므로, 객체의 구조적 특징이 정확하게 고려되지 않아 따라 객체와 경계 박스를 검출하는 성능에 한계가 있다.
- [0010] 이에 다양한 객체의 구조적 특징에도 정확하게 객체를 판별하고 경계 박스를 검출할 수 있는 인공 신경망이 요구되고 있다.

선행기술문헌

특허문헌

- [0011] (특허문헌 0001) 한국 공개 특허 제10- 2020-0002770호 (2020.01.08 공개)

발명의 내용

해결하려는 과제

- [0012] 본 발명의 목적은 다양한 크기 및 형상을 갖는 객체의 구조적 특징을 고려하여 객체를 판별하고 경계 박스를 추정할 수 있는 객체 검출 장치 및 방법을 제공하는데 있다.
- [0013] 본 발명의 다른 목적은 객체의 구조적 특징에 대응하여 동적으로 변화되는 가중치를 갖는 동적 커널을 이용하여 정확하게 객체를 판별하고 경계 박스를 추정하는 객체 검출 장치 및 방법을 제공하는데 있다.

과제의 해결 수단

- [0014] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른 객체 검출 장치는 패턴 추정 방식이 미리 학습된 인공 신경망으로 구현되어, 인가된 입력 이미지의 특징을 추출하여 입력 특징맵을 획득하는 입력 특징맵 획득부; 패턴 추정 방식이 미리 학습된 인공 신경망으로 구현되어, 상기 입력 특징맵에서 객체 검출을 위한 특징을 추출하기 위해 학습에 의해 미리 획득된 메인 커널에서 샘플링될 기지정된 개수의 위치 정보가 미리 지정된 원소로 포함된 디폴트 파라미터를 기지정된 다수의 스케일 파라미터 중 하나의 스케일 파라미터에 따라 스케일 변환하고, 스케일 변환된 정규 파라미터의 원소들이 나타내는 위치 정보 각각의 오프셋을 보정하여 상기 입력 특징맵의 다수의 픽셀 각각에 대응하는 변환 파라미터를 획득하고, 획득된 상기 변환 파라미터를 기반으로 상기 메인 커널을 기지정된 방식으로 샘플링하여, 상기 입력 특징맵의 다수의 픽셀 각각에 따라 적응적으로 가변되는 다수의 동적 커널을 추출하는 동적 커널 추출부; 상기 입력 특징맵과 상기 입력 특징맵의 각 픽셀에 대응하는 다수의 동적 커널을 기지정된 방식으로 연산하여 출력 특징맵을 획득하는 출력 특징맵 획득부; 및 패턴 추정 방식이 미리 학습된 인공 신경망으로 구현되어, 상기 출력 특징맵의 패턴을 추정하여 상기 입력 이미지에서 객체가 포함된 영역을 나타내는 경계 박스와 상기 경계 박스에 포함된 객체를 판별하는 객체 및 경계 박스 추출부를 포함한다.

- [0015] 상기 동적 커널 추출부는 미리 학습된 패턴 추정 방식에 따라 상기 입력 특징맵의 패턴을 추정하여, 상기 디폴트 파라미터를 상기 입력 특징맵의 다수의 픽셀 각각에 대응하는 패턴으로 스케일 변환 및 오프셋 보정하기 위한 변환 파라미터맵을 획득하는 변환 파라미터 추출부; 및 상기 변환 파라미터맵으로부터 상기 입력 특징맵의 다수의 픽셀 각각에 대응하여 상기 디폴트 파라미터의 각 원소가 나타내는 위치 정보를 스케일 파라미터와 오프셋에 따라 변경한 다수의 변환 파라미터를 획득하고, 획득된 다수의 변환 파라미터에 따라 미리 획득되어 저장된 상기 메인 커널을 샘플링하여 상기 입력 특징맵의 다수의 픽셀 각각에 대응하는 다수의 동적 커널을 추출하는 커널 샘플링부를 포함할 수 있다.
- [0016] 상기 변환 파라미터 추출부는 상기 입력 특징맵을 인가받고, 미리 학습된 패턴 추정 방식에 따라 상기 입력 특징맵의 패턴을 추정하여, 상기 입력 특징맵의 각 픽셀에 대응하는 객체의 크기가 반영되는 스케일 파라미터를 추출하기 위한 정규 변환맵을 획득하는 정규 변환맵 획득부; 및 상기 입력 특징맵 또는 상기 정규 변환맵 중 적어도 하나를 인가받고, 미리 학습된 패턴 추정 방식에 따라 상기 입력 특징맵의 각 픽셀에 대응하는 객체의 형상이 반영되는 오프셋을 추출하기 위한 오프셋 보정맵을 획득하는 오프셋 보정맵 획득부를 포함할 수 있다.
- [0017] 상기 정규 변환맵 획득부는 상기 디폴트 파라미터의 각 원소가 나타내는 위치 정보에 따른 거리를 각각 서로 다른 지정된 스케일로 변환하기 위해 미리 설정되는 기지정된 개수의 스케일 파라미터에 대응하는 크기로 상기 정규 변환맵을 획득하고, 상기 오프셋 보정맵 획득부는 상기 동적 커널의 크기에 대응하여 미리 지정되는 상기 디폴트 파라미터의 원소 개수에 대응하는 크기로 상기 오프셋 보정맵을 획득할 수 있다.
- [0018] 상기 정규 변환맵과 상기 오프셋 보정맵 각각은 상기 스케일 파라미터와 상기 오프셋을 상기 메인 커널의 행방향과 열방향에 따른 x축 방향과 y축 방향의 위치 정보를 구분하여 스케일 변환 및 오프셋 보정할 수 있도록 스케일 파라미터 개수의 2배에 대응하는 크기와 상기 동적 커널의 크기의 2배에 대응하는 크기로 획득될 수 있다.
- [0019] 상기 스케일 파라미터는 상기 디폴트 파라미터를 스케일 변환한 정규 변환 파라미터의 각 원소가 나타내는 위치 정보가 상기 메인 커널의 크기를 초과하지 않는 값으로 설정될 수 있다.
- [0020] 상기 변환 파라미터 추출부는 상기 정규 변환맵과 상기 오프셋 보정맵을 결합하여 상기 변환 파라미터맵을 출력할 수 있다.
- [0021] 상기 커널 샘플링부는 상기 입력 특징맵의 다수의 픽셀 각각에 대응하는 변환 파라미터의 원소들이 나타내는 위치 정보에 따라 상기 메인 커널의 대응하는 위치의 가중치를 샘플링하여 상기 동적 커널을 추출할 수 있다.
- [0022] 상기 객체 검출 장치는 학습 시에 이미지에 포함된 객체에 대한 객체 클래스와 경계 박스가 미리 주석된 다수의 학습 데이터가 입력 이미지로 인가받아 객체와 경계 박스를 판별하며, 판별된 객체와 경계 박스를 주석된 객체 클래스 및 경계 박스와 비교하여 계산된 손실을 역전파하여 상기 메인 커널을 획득하고, 상기 정규 변환맵 획득부와 상기 오프셋 보정맵 획득부를 학습시킬 수 있다.
- [0023] 상기 목적을 달성하기 위한 본 발명의 다른 실시예에 따른 객체 검출 방법은 패턴 추정 방식이 미리 학습된 인공 신경망을 이용하여, 입력 이미지의 특징을 추출하여 입력 특징맵을 획득하는 단계; 패턴 추정 방식이 미리 학습된 인공 신경망을 이용하여, 상기 입력 특징맵에서 객체 검출을 위한 특징을 추출하기 위해 학습에 의해 미리 획득된 메인 커널에서 샘플링된 기지정된 개수의 위치 정보가 미리 지정된 원소로 포함된 디폴트 파라미터를 기지정된 다수의 스케일 파라미터 중 하나의 스케일 파라미터에 따라 스케일 변환하고, 스케일 변환된 정규 파라미터의 원소들이 나타내는 위치 정보 각각의 오프셋을 보정하여 상기 입력 특징맵의 다수의 픽셀 각각에 대응하는 변환 파라미터를 획득하는 단계; 획득된 상기 변환 파라미터를 기반으로 상기 메인 커널을 기지정된 방식으로 샘플링하여, 상기 입력 특징맵의 다수의 픽셀 각각에 따라 적응적으로 가변되는 다수의 동적 커널을 추출하는 단계; 상기 입력 특징맵과 상기 동적 커널의 각 픽셀에 대응하는 다수의 동적 커널을 기지정된 방식으로 연산하여 출력 특징맵을 획득하는 단계; 및 패턴 추정 방식이 미리 학습된 인공 신경망을 이용하여, 상기 출력 특징맵의 패턴을 추정하여 상기 입력 이미지에서 객체가 포함된 영역을 나타내는 경계 박스와 상기 경계 박스에 포함된 객체를 판별하는 단계를 포함한다.

발명의 효과

- [0024] 따라서, 본 발명의 실시예에 따른 객체 검출 장치 및 방법은 입력 특징맵에서 객체 특징맵을 추출하지 않는 단일 스테이지 방식에서 객체의 구조적 특징에 대응하여 가중치가 동적으로 변화되는 동적 커널을 이용하여 객체를 판별하고 경계 박스를 추정함으로써, 픽셀이 갖는 객체 정보의 누락없이 다양한 크기 및 형상을 갖는 객체를 정확하게 판별하고 경계 박스를 추정할 수 있다.

도면의 간단한 설명

- [0025] 도 1은 본 발명의 일 실시예에 따른 동적 객체 기반 객체 검출 장치의 개략적 구조를 나타낸다.
- 도 2는 도 1의 객체 기반 객체 검출 장치의 전체적인 동작의 개념을 설명하기 위한 도면이다.
- 도 3은 도 1의 변환 파라미터 추출부가 메인 커널에서 샘플링될 커널을 추출하는 개념을 설명하기 위한 도면이다.
- 도 4는 이미지 내의 객체 별 크기 및 형상의 차이와 동적 커널의 개념을 설명하기 위한 도면이다.
- 도 5는 도 1의 샘플링부에서 각 객체의 픽셀별로 샘플링되어 추출된 동적 커널과 이에 따라 추출된 경계 박스의 일 예를 나타낸다.
- 도 6은 본 발명의 일 실시예에 따른 동적 객체 기반 객체 검출 방법을 나타낸다.

발명을 실시하기 위한 구체적인 내용

- [0026] 본 발명과 본 발명의 동작상의 이점 및 본 발명의 실시예에 의하여 달성되는 목적을 충분히 이해하기 위해서는 본 발명의 바람직한 실시예를 예시하는 첨부 도면 및 첨부 도면에 기재된 내용을 참조하여야만 한다.
- [0027] 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.
- [0028] 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.
- [0029] 도 1은 본 발명의 일 실시예에 따른 동적 객체 기반 객체 검출 장치의 개략적 구조를 나타내고, 도 2는 도 1의 객체 기반 객체 검출 장치의 전체적인 동작의 개념을 설명하기 위한 도면이며, 도 3은 도 1의 변환 파라미터 추출부가 메인 커널에서 샘플링될 커널을 추출하는 개념을 설명하기 위한 도면이다. 그리고 도 4는 이미지 내의 객체 별 크기 및 형상의 차이와 동적 커널의 개념을 설명하기 위한 도면이다.
- [0030] 도 1을 참조하면, 본 실시예에 따른 객체 검출 장치는 입력 특징맵 획득부(100), 동적 커널 추출부(200), 출력 특징맵 획득부(300), 객체 클래스 판별부(400) 및 경계 박스 추출부(500)를 포함할 수 있다.
- [0031] 우선 입력 특징맵 획득부(100)는 입력 이미지를 인가받고, 미리 학습된 패턴 추정 방식에 따라 인가된 입력 이미지의 패턴을 추정하여 입력 특징맵(F_1)을 획득한다.
- [0032] 여기서 입력 이미지는 검출되어야 할 적어도 하나의 객체가 포함된 이미지일 수 있다. 이때 입력 이미지는 가로 세로 크기 중 적어도 하나가 기지정된 크기(일 예로 가로 세로 중 작은 방향 크기가 800픽셀)를 갖도록 크기 변환된 이미지일 수 있다.
- [0033] 입력 특징맵 획득부(100)는 패턴 추정 방식이 미리 학습된 인공 신경망으로 구현되어 입력 특징맵(F_1)을 획득할 수 있다. 입력 특징맵 획득부(100)는 일 예로 입력 이미지내의 객체의 다양한 스케일에도 강건하게 객체 검출을 위한 특징을 추출할 수 있는 것으로 알려진 특징 피라미드 네트워크(Feature Pyramid Network: 이하 FPN)로 구현될 수 있다. FPN은 입력 이미지로부터 상향 방식으로 인코딩함으로써 점차 크기를 축소하면서 특징맵을 추출하고, 축소된 특징맵을 다시 디코딩하여 크기를 확대하면서 특징맵을 추출하며, 대응하는 크기의 특징맵 사이를 스킵 커넥션을 통해 연결하여 특징을 결합함으로써, 여러 스케일에서의 특징이 반영된 입력 특징맵(F_1)을 획득할 수 있다.
- [0034] 여기서는 일 예로 입력 특징맵 획득부(100)가 FPN으로 구현되는 것으로 가정하여 설명하지만, 다양한 스케일에서 객체를 검출하기 위한 입력 특징맵(F_1)을 추출하는 다양한 인공 신경망이 공개되어 있으므로, 입력 특징맵 획득부(100)는 다른 인공 신경망으로 구현될 수도 있다.

- [0035] 입력 특징맵 획득부(100)에서 획득된 입력 특징맵(F_I)은 도 2에 도시된 바와 같이, $H \times W \times C$ 의 크기로 획득될 수 있다. 여기서 H , W , C 는 각각 특징맵의 높이, 폭 및 채널 수를 의미한다.
- [0036] 동적 커널 추출부(200)는 패턴 추정 방식이 미리 학습된 인공 신경망을 포함하여, 입력 특징맵(F_I)의 다수의 픽셀 각각에 대해 적응적으로 가변되는 동적 커널을 획득한다.
- [0037] 동적 커널 추출부(200)는 입력 특징맵 획득부(100)에서 획득된 입력 특징맵(F_I)의 다수의 픽셀 각각에 따라 기존의 단일 스테이지 객체 검출 기법에서 이용되는 메인 커널(W_g)을 서로 다르게 샘플링하여 각 픽셀에 대응하여 가변되는 동적 커널을 추출한다. 즉 동적 커널 추출부(200)는 입력 이미지의 특징을 추출한 입력 특징맵(F_I)의 픽셀에 각각에 따라 커널을 서로 다르게 획득한다.
- [0038] 여기서 동적 커널 추출부(200)가 입력 특징맵(F_I)의 각 픽셀에 대해 서로 다른 동적 커널을 획득하는 것은 입력 특징맵(F_I)의 각 픽셀이 나타내는 객체에 대한 특징 정보가 서로 상이하기 때문이다.
- [0039] 이에 본 실시예에서 동적 커널 추출부(200)는 각 픽셀이 나타내는 객체의 크기 및 형상에 따른 구조적 특징을 반영하여 메인 커널(W_g)을 구성하는 다수의 가중치를 샘플링하여 추출하여 동적 커널을 획득함으로써, 동적 커널에 각 픽셀에 포함된 객체 특징이 반영되도록 한다. 즉 동적 커널을 픽셀에 따라 적응적으로 가변하여 획득한다.
- [0040] 이때 메인 커널(W_g)에서 샘플링되어 획득되는 동적 크기는 기존의 단일 스테이지 객체 검출 기법에 이용되는 메인 커널(W_g)보다 작은 크기로 획득된다. 일 예로 메인 커널(W_g)이 9×9 의 크기를 갖는 경우, 동적 커널은 3×3 의 크기로 획득될 수 있다. 이는 커널의 크기를 줄여 연산량을 저감하기 위해서이다. 비록 동적 커널의 크기가 메인 커널(W_g)의 크기보다 작아질지라도, 동적 커널은 픽셀에 따른 객체의 구조적 특징이 미리 반영되어 있으므로, 작은 크기로도 객체의 특징을 더욱 정확하게 추출할 수 있도록 한다.
- [0041] 동적 커널 추출부(200)는 도 1에 도시된 바와 같이, 변환 파라미터 추출부(210) 및 커널 샘플링부(220)를 포함할 수 있다.
- [0042] 변환 파라미터 추출부(210)는 패턴 추정 방식이 미리 학습된 인공 신경망으로 구현되어, 입력 특징맵(F_I)의 각 픽셀에 대응하여 메인 커널(W_g)로부터 동적 커널을 추출하기 위한 변환 파라미터맵을 추출한다. 변환 파라미터 추출부(210)는 학습된 패턴 추정 방식에 따라 입력 특징맵(F_I)의 각 픽셀 패턴을 추정함으로써, 변환 파라미터맵을 추출한다.
- [0043] 여기서 변환 파라미터맵은 입력 특징맵(F_I)의 각 픽셀이 나타내는 객체의 구조적 특징에 대응하여 메인 커널(W_g)을 동적 커널로 변환하기 위해 획득되는 특징맵이다. 일 예로 상기한 바와 같이, 메인 커널(W_g)이 9×9 의 크기를 갖고 동적 커널이 3×3 의 크기로 획득되는 경우, 변환 파라미터 추출부(210)는 9×9 크기의 메인 커널(W_g)로부터 3×3 크기의 동적 커널을 획득하기 위한 샘플링 기준이 되는 변환 파라미터맵을 입력 특징맵(F_I)의 각 픽셀이 나타내는 객체의 구조적 특징을 기반으로 추출한다.
- [0044] 변환 파라미터 추출부(210)는 정규 변환맵 획득부(211)와 오프셋 보정맵 획득부(212)를 포함할 수 있다. 정규 변환맵 획득부(211)와 오프셋 보정맵 획득부(212)는 각각 패턴 추정 방식이 미리 학습된 인공 신경망으로 구현될 수 있다. 그리고 정규 변환맵 획득부(211)는 미리 지정된 디폴트 파라미터를 입력 특징맵(F_I)의 다수의 픽셀 각각에 대응하여 기지정된 정규 크기로 변경하여 정규 파라미터맵을 획득하기 위한 정규 변환맵을 획득한다.
- [0045] 오프셋 보정맵 획득부(212)는 입력 특징맵(F_I)의 다수의 픽셀 각각에 대응하여 정규 변환맵 획득부(211)에서 획득된 정규 파라미터의 오프셋을 보정하여 변환 파라미터를 추출하기 위한 오프셋 보정맵을 획득한다.
- [0046] 도 3을 참조하여, 정규 변환맵 획득부(211)와 오프셋 보정맵 획득부(212) 각각의 동작을 설명하면, 우선 본 실시예에 따른 변환 파라미터 추출부(210)에는 메인 커널(W_g)을 샘플링하기 위한 디폴트 파라미터(default parameter)(g)가 미리 지정될 수 있다. 여기서 디폴트 파라미터(g)는 (a)에 도시된 바와 같이, 메인 커널(W_g)에서 샘플링될 위치 정보가 기지정된 개수의 원소로 포함된 파라미터이다. 즉 디폴트 파라미터는 메인 커널(W

g)의 샘플링될 가중치를 지정하는 위치 정보의 집합으로 구성될 수 있다. 디폴트 파라미터(g)는 동적 커널의 크기에 대응하는 n 개의 샘플링 포인트 위치를 원소로 포함한다.

[0047] 이때 디폴트 파라미터(g)는 입력 특징맵(F_I)에 포함된 객체의 특징이 반영되지 않은 파라미터이므로, 모든 픽셀에 동일하게 적용되는 파라미터로서 메인 커널(W_g)의 전체적인 특징을 균등하게 샘플링할 수 있도록 설정된다. 따라서 디폴트 파라미터(g)는 (a)에서와 같이, 동적 커널의 크기에 대응하는 개수로 균등하게 분할되는 메인 커널(W_g)의 각 영역을 대표하는 기지정된 위치 좌표들의 조합으로 설정될 수 있다.

[0048] 여기서는 동적 커널의 크기가 3×3 인 것으로 가정하였으므로, 디폴트 파라미터(g)는 $n(= 9)$ 개의 디폴트 위치 좌표의 조합으로 구성될 수 있다. 또한 메인 커널(W_g)이 9×9 크기 가중치 행렬 형태로 구성되므로, 균등 분할된 메인 커널(W_g)의 각 영역을 대표하는 위치 좌표의 조합으로 구성되는 디폴트 파라미터(g) 또한 (a)에 도시된 바와 같이 서로 균등한 간격으로 이격된 사각 형상의 9개의 위치 좌표로 구성될 수 있다.

[0049] 상기와 같이 메인 커널(W_g)이 9×9 크기를 갖는 가중치 행렬로 구성되는 경우, 중앙에 위치한 가중치를 원점 좌표(0,0)라 하고 행방향 및 열방향을 각각 x 축 및 y 축으로 볼 때, 메인 커널(W_g)은 x 축 방향 및 y 축 방향으로 각각 $-4 \sim 4$ 의 크기를 갖는 것으로 볼 수 있다. 이때, 디폴트 파라미터(g)의 9개의 위치 좌표 중 하나는 중앙에 위치하는 원점 좌표(0,0)를 가질 수 있다. 그리고 나머지 8개의 위치 좌표는 도 3의 (a)에 도시된 바와 같이, 일 예로 각각 $((-2.5, 2.5), (0, 2.5), (2.5, 2.5), (-2.5, 0), (0, 2.5), (-2.5, -2.5), (0, -2.5), (2.5, -2.5))$ 로 설정될 수 있다. 디폴트 파라미터(g)의 각 위치 좌표는 설정에 따라 조절될 수 있으나, 정규 변환맵 획득부(211)와 오프셋 보정맵 획득부(212)에 의해 디폴트 파라미터(g)의 각 위치 좌표가 가변될 때, 메인 커널(W_g)의 크기 범위에서 초과되지 않도록 설정되어야 한다.

[0050] 정규 변환맵 획득부(211)는 입력 특징맵(F_I)의 각 픽셀의 패턴에 따라 디폴트 파라미터(g)를 정규화된 방식으로 변환하기 위한 정규 변환맵을 획득한다. 여기서 정규 변환맵 획득부(211)는 미리 학습된 패턴 추정 방식에 따라 입력 특징맵(F_I)의 각 픽셀(i)에 대한 패턴과 기지정된 다수의 스케일 파라미터(S)들 각각 사이의 관계를 추정하여 정규 변환맵을 생성한다. 여기서 정규 변환맵은 입력 특징맵(F_I)의 픽셀(i) 각각에 가장 적합한 스케일 파라미터(S)를 선택하기 위해 획득되는 특징맵이고, 스케일 파라미터(S)는 선택된 스케일 파라미터(S)에 따라 스케일 변환된 정규 파라미터($S \cdot g$)가 나타내는 위치 좌표가 메인 커널(W_g)의 크기를 초과하지 않을 뿐만 아니라, 메인 커널(W_g)의 특정 위치의 단일 가중치만을 지정하지 않도록, 미리 지정될 수 있다. 여기서는 일 예로 스케일 파라미터(S)는 $\{0.5, 0.75, 1, 1.25, 1.5\}$ 의 5개로 설정된 것으로 가정한다.

[0051] 그리고 정규 변환맵 획득부(211)는 디폴트 파라미터(g)의 스케일 변환시에 x 축 방향과 y 축 방향을 구분하여 별도로 스케일 변환될 수 있도록 기지정된 s 개 (여기서는 일 예로 5개)의 스케일 파라미터(S) 각각이 x 축 방향의 x 스케일 파라미터($S_x = \{0.5, 0.75, 1, 1.25, 1.5\}$)와 y 축 방향의 y 스케일 파라미터($S_y = \{0.5, 0.75, 1, 1.25, 1.5\}$)를 구분될 수 있도록 한다. 이는 객체의 다양한 형상 구조에 대응할 수 있도록 하기 위함이다. 즉 디폴트 파라미터(g)가 정사각형의 형상 구조를 갖더라도, 디폴트 파라미터(g)가 스케일 파라미터(S)에 의해 변환된 정규 파라미터($S \cdot g$)는 (b)에 도시된 바와 같이 직사각형의 형상 구조를 갖는 위치 정보의 원소로 지정될 수 있다.

[0052] 따라서 정규 변환맵 획득부(211)에서 획득되는 정규 변환맵은 입력 특징맵(F_I)의 높이와 폭, 그리고 x 스케일 파라미터와 y 스케일 파라미터 각각에 대응할 수 있도록 $H \times W \times 2(s)$ 의 크기로 획득될 수 있다. 상기에서 x 스케일 파라미터와 y 스케일 파라미터 각각이 5개씩으로 지정되는 것으로 가정하였으므로, 정규 변환맵은 $H \times W \times 10$ 의 크기로 획득될 수 있다.

[0053] 오프셋 보정맵 획득부(212)는 입력 특징맵(F_I)의 픽셀 패턴에 따라 정규 변환맵 획득부(211)에서 획득된 정규 변환맵을 기반으로 정규 파라미터($S \cdot g$)의 각 위치 정보를 오프셋값(Δg)만큼 보정하기 위한 오프셋 보정맵을 추출한다. 즉 오프셋 보정맵은 정규 파라미터($S \cdot g$)를 변환 파라미터($S \cdot g + \Delta g$)로 변환하기 위해 획득될 수 있다.

[0054] 상기한 정규 변환맵 획득부(211)는 입력 특징맵의 각 픽셀이 나타내는 객체의 개략적 크기에 대응하는 구조의 정규 파라미터($S \cdot g$)를 획득할 수 있는 스케일 파라미터를 선택하기 위한 정규 변환맵을 획득한다. 그러나 정

규 변환맵을 기반으로 디폴트 파라미터(g)를 변환하여 획득되는 정규 파라미터($S \cdot g$)는 객체의 형상에 대한 정보를 나타내지 못한다.

[0055] 이에 오프셋 보정맵 획득부(212)는 오프셋 보정맵을 추출하여 정규 파라미터($S \cdot g$)의 각 위치 좌표에 대해 (c)에 도시된 바와 같이, 개별적인 오프셋(Δg)을 추가로 적용할 수 있도록 함으로서, 변환 파라미터($S \cdot g + \Delta g$)의 각 위치 정보가 객체 형상에 대응하는 패턴을 갖도록 할 수 있다. 이때 오프셋(Δg) 또한 x축 방향과 y축 방향으로 구분되어 x 오프셋(Δg_x)과 y 오프셋(Δg_y)으로 각각 적용될 수 있다. 그리고 오프셋(Δg)은 정규 파라미터($S \cdot g$)의 각 위치 좌표에 개별적으로 적용되어야 한다.

[0056] 따라서 오프셋 보정맵 획득부(212)는 동적 커널의 가중치 개수, 즉 디폴트 파라미터의 원소 개수에 대응하는 개수의 x 오프셋(Δg_x)과 y 오프셋(Δg_y)을 획득하기 위한 오프셋 보정맵을 추출한다. 따라서 오프셋 보정맵은 $H \times W \times 2(n)$ 크기로 추출될 수 있다.

[0057] 여기서 오프셋(Δg)은 정규 변환맵 획득부(211)에 의한 스케일 변환만으로 표현할 수 없는 객체의 형상에 따른 패턴을 표현하기 위한 값이므로, 오프셋(Δg)의 과조한 변화로 객체 크기를 표현한 정규 파라미터($S \cdot g$)의 구조가 심각하게 왜곡되지 않도록 크기가 기지정된 범위(예를 들면 $-1 \sim 1$) 이내로 제한될 수 있다.

[0058] 변환 파라미터 추출부(210)는 정규 변환맵 획득부(211)에서 획득된 $H \times W \times 2(s)$ 크기의 정규 변환맵과 오프셋 보정맵 획득부(212)에서 획득된 $H \times W \times 2(n)$ 크기의 오프셋 보정맵을 결합하여 $H \times W \times 2(s+n)$ 크기의 변환 파라미터맵을 커널 샘플링부(220)로 전달한다.

[0059] 상기에서는 일 예로 변환 파라미터 추출부(210)의 정규 변환맵 획득부(211)와 오프셋 보정맵 획득부(212)가 병렬로 연결되어 개별적으로 정규 변환맵과 오프셋 보정맵을 획득하는 것으로 도시하였으나, 정규 변환맵 획득부(211)와 오프셋 보정맵 획득부(212)는 직렬로 연결될 수도 있다. 이 경우, 오프셋 보정맵 획득부(212)는 정규 변환맵 획득부(211)에서 출력되는 정규 변환맵을 인가받아 오프셋 보정맵이 결합된 변환 파라미터맵을 출력할 수도 있다.

[0060] 커널 샘플링부(220)는 학습에 의해 미리 획득된 메인 커널(W_g)이 저장되어 변환 파라미터 추출부(210)로부터 변환 파라미터맵이 인가되면, 메인 커널(W_g)을 구성하는 다수의 가중치를 인가된 변환 파라미터맵에 따라 지정된 방식으로 샘플링하여 추출함으로써, 입력 특징맵(F_I)의 각 픽셀(i)에 대응하는 동적 커널을 획득한다.

[0061] 이때 커널 샘플링부(220)는 $H \times W \times 2(s+n)$ 크기의 변환 파라미터맵에서 $H \times W$ 평면의 픽셀(i) 각각에 대한 채널 방향의 $2(s+n)$ 크기의 벡터를 기반으로 디폴트 파라미터(g)를 정규 변환할 스케일 파라미터(S)를 선택하고, 오프셋(Δg)을 판별한다. 그리고 선택된 스케일 파라미터(S)와 오프셋(Δg)을 기반으로 입력 특징맵(F_I)의 각 픽셀(i)에 대응하는 변환 파라미터($\tilde{g}^{(i)}$)를 수학적 식 1과 같이 획득할 수 있다.

수학적 식 1

[0062]
$$\tilde{g}^{(i)} = S(i) \cdot g + \Delta g(i)$$

[0063] 입력 특징맵(F_I)의 각 픽셀(i)에 대응하는 변환 파라미터($\tilde{g}^{(i)}$)가 획득되면, 커널 샘플링부(220)는 획득된 픽셀(i)별 변환 파라미터($\tilde{g}^{(i)}$)에 따라 메인 커널(W_g)에 대해 서로 다르게 가중치를 샘플링하여 추출함으로써, 픽셀(i)별 동적 커널을 획득한다. 여기서 획득되는 픽셀(i)별 동적 커널은 $W_g(\tilde{g}^{(i)})$ 로 표현될 수 있다.

[0064] 커널 샘플링부(220)가 변환 파라미터($\tilde{g}^{(i)}$)에 따라 메인 커널(W_g)을 샘플링하여 동적 커널($W_g(\tilde{g}^{(i)})$)을 획득하는 방식은 미리 지정될 수 있다. 일 예로 커널 샘플링부(220)는 메인 커널(W_g)에서 변환 파라미터($\tilde{g}^{(i)}$)의 다수의 위치 정보 각각 가장 인접한 다수의 가중치를 추출하여 동적 커널($W_g(\tilde{g}^{(i)})$)로 획득할 수 있다. 또는 커널 샘플링부(220)는 변환 파라미터($\tilde{g}^{(i)}$)에 따라 메인 커널(W_g)의 특정 위치의 가중치를 추출하여 동적 커널($W_g(\tilde{g}^{(i)})$)로 획득할 수 있다.

플러부(220)는 메인 커널(W_g)에서 변환 파라미터($\tilde{g}^{(i)}$)의 다수의 위치 정보 각각에 따라 주변 가중치들을 위치 정보에 대응하는 비율로 결합하여 결합된 가중치를 동적 커널($W_g(\tilde{g}^{(i)})$)의 원소로 획득할 수도 있다. 이외에도 커널 샘플링부(220)는 기지정된 다양한 샘플링 방식에 따라 메인 커널(W_g)로부터 동적 커널($W_g(\tilde{g}^{(i)})$)의 원소가 되는 가중치를 추출할 수 있다.

[0065] 도 4는 입력 이미지의 일 예로 야구 경기의 한 장면으로 타자와 포수 및 심판이 포함된 이미지를 나타내었다. 도 4에 도시된 바와 같이, 입력 이미지에는 다수의 객체가 동시에 포함될 수 있으며, 입력 이미지에 포함된 각 객체의 크기 및 형태 등은 서로 상이하다. 특히 타자와 포수를 비교하는 경우, 입력 이미지에서 타자는 서있는 자세이므로 세로 방향으로 길고 큰 형태를 갖는 반면, 포수는 앉아있는 자세이므로 작은 크기로 밀집된 형태를 갖는다. 즉 입력 이미지에는 각 객체의 크기 및 형상이 서로 상이하게 나타날 수 있다.

[0066] 이와 같이 각 객체의 크기 및 형상이 서로 상이하게 나타남에도 불구하고 기존의 단일 스테이지 객체 검출 기법에서는 (a)에 도시된 바와 같이 디폴트 파라미터(g)의 경우 객체의 구조적 특징에 무관하게 모든 픽셀에 대해 동일한 커널을 샘플링하여 커널을 생성한다. 즉 크고 긴 형태를 갖는 타자를 나타내는 객체에 해당하는 픽셀과 작고 밀집된 형태를 갖는 포수를 나타내는 객체에 해당하는 픽셀에 대해 동일한 커널을 적용하였다. 이 경우, 커널은 객체의 구조적 특징을 고려하지 않고 입력 특징맵(F_I)에 대한 연산을 수행하게 되어 상기한 바와 같이, 다양한 크기 및 형상을 갖는 객체의 특징을 정확하게 추출하지 못하게 된다.

[0067] 그에 반해 본 실시예에 따른 객체 검출 장치에서는 (b)에 도시된 바와 같이 변환 파라미터($\tilde{g}^{(i)}$)는 객체의 크기 및 형상에 따라 메인 커널(W_g)을 다르게 샘플링하여 픽셀(i)별로 객체의 구조적 특징이 반영된 서로 다른 동적 커널($W_g(\tilde{g}^{(i)})$)을 생성할 수 있도록 한다.

[0068] 본 실시예에서 동적 커널 추출부(200)가 입력 특징맵(F_I)의 각 픽셀에 대응하는 동적 커널을 개별적으로 별도로 생성하지 않고, 메인 커널(W_g)을 샘플링을 하여 각 픽셀에 대응하는 동적 커널($W_g(\tilde{g}^{(i)})$)을 생성하는 것은 동적 커널($W_g(\tilde{g}^{(i)})$)을 생성하기 위한 파라미터의 수를 줄여 연산량을 저감하기 위해서이다.

[0069] 출력 특징맵 획득부(300)는 동적 커널 추출부(200)에서 입력 특징맵(F_I)의 다수의 픽셀 각각을 각 픽셀에 대응하여 적응적으로 획득된 동적 커널($W_g(\tilde{g}^{(i)})$)과 기지정된 연산을 수행하여 출력 특징맵(F_O)을 획득한다. 출력 특징맵 획득부(300)는 일 예로 수학식 2와 같이 입력 특징맵(F_I)의 각 픽셀을 픽셀 위치에 따라 서로 다르게 적응적으로 획득된 동적 커널($W_g(\tilde{g}^{(i)})$)과 컨볼루션 연산을 수행하여 출력 특징맵(F_O)을 획득할 수 있다.

수학식 2

$$F_O(i) = \sum_n W_g(\tilde{g}_n(i)) \cdot F_I(i + g_n)$$

[0070]

[0071] 즉 출력 특징맵 획득부(300)는 입력 특징맵(F_I)의 각 픽셀에 대해 객체의 구조적 특징이 반영된 동적 커널을 적용하여 획득된 출력 특징맵(F_O)을 획득한다. 따라서 출력 특징맵(F_O)은 각 픽셀에 대응하는 객체의 구조적 특징이 포함된 특징맵으로 볼 수 있다. 여기서 출력 특징맵(F_O)은 입력 특징맵(F_I)과 마찬가지로 $H \times W \times C$ 의 크기로 획득될 수 있다.

[0072] 객체 클래스 판별부(400)와 경계 박스 추출부(500)는 각각 출력 특징맵 획득부(300)에서 획득된 출력 특징맵(F_O)을 기반으로 입력 이미지에 포함된 객체를 판별하고, 입력 이미지에서 각 객체가 포함된 영역을 나타내는 경계 박스를 추출한다.

- [0073] 객체 클래스 판별부(400)와 경계 박스 추출부(500)는 기존의 단일 스테이지 객체 검출 기법에서와 동일한 방식으로 출력 특징맵(F_0)으로부터 포함된 객체를 판별하고, 경계 박스를 추출할 수 있다.
- [0074] 일 예로 경계 박스 추출부(500)는 패턴 추정 방식이 미리 학습된 인공 신경망으로 구현되어, 학습된 패턴 추정 방식에 따라 출력 특징맵(F_0)에서 객체 특징을 나타내는 픽셀을 판별하여 객체 영역을 추정한다. 이때, 경계 박스 추출부(500)는 각 픽셀이 나타내는 객체 특징을 구분하여 판별함으로써, 입력 이미지 내에 다수의 객체가 포함된 경우에도 각 객체 영역을 구분하여 추정할 수 있다. 그리고 추정된 객체 영역을 입력 이미지에 대응하는 좌표 형태로 회귀시켜 경계 박스를 추출할 수 있다. 이때, 경계 박스 추출부(500)는 추정된 각 객체 영역으로부터 경계 박스를 추출하기 위한 기준이 되도록 미리 지정된 다수의 앵커(Anchor)를 기반으로, 앵커의 크기를 변형하는 방식으로 각 객체의 크기 및 형상에 적합한 경계 박스를 추출할 수 있다.
- [0075] 한편, 일 예로서 객체 클래스 판별부(400) 또한 패턴 추정 방식이 미리 학습된 인공 신경망으로 구현되어, $H \times W \times C$ 크기의 출력 특징맵(F_0)에서 채널(C)축 방향의 픽셀 특징 벡터를 추출하고, 추출된 픽셀 특징 벡터에 대해 기지정된 연산을 수행함으로써 다수의 픽셀 특징 벡터 각각이 기지정된 다수의 객체 클래스 각각에 대응하는 확률을 나타내는 스코어 맵(score map)을 생성할 수 있다. 이 때, 객체 클래스 판별부(400)는 경계 박스 추출부(500)에서 추정된 각 객체 영역에 포함된 픽셀을 구분하여 스코어 맵(score map)을 생성할 수 있으며, 각 객체 영역별로 구분되어 생성된 스코어 맵(score map)을 기반으로 객체 영역에 포함된 객체를 판별할 수 있다.
- [0076] 다른 예로 객체 클래스 판별부(400)와 경계 박스 추출부(500)는 도 2에 도시된 바와 같이, 다수의 서브 네트워크를 포함하여 구성될 수 있다. 경계 박스 추출부(500)는 미리 학습된 패턴 추정 방식에 따라 기지정된 개수의 앵커 각각의 4개의 모서리 위치를 결정하기 위한 경계 박스 특징맵을 추출할 수 있다. 여기서 경계 박스 특징맵은 다수의 앵커 각각에서 4개의 모서리 위치 각각을 구분하여 추정하도록 구분되어 획득될 수 있다. 이에 앵커의 개수가 A개라면, 경계 박스 특징맵은 $H \times W \times 4A$ 개로 획득될 수 있다. 객체 클래스 판별부(400)는 $H \times W \times 4A$ 개의 경계 박스 특징맵으로부터 추정된 객체 영역을 입력 이미지에 대응하는 좌표 형태로 회귀시켜 경계 박스를 추출할 수 있다.
- [0077] 그리고 객체 클래스 판별부(400)는 각각 미리 지정된 객체를 식별하기 위한 K개의 클래스 서브 네트워크를 포함할 수 있다. 이 경우 K개의 클래스 서브 네트워크는 출력 특징맵(F_0)에서 대응하는 객체가 포함되어 있는지 여부를 나타내는 K개의 객체 클래스 특징맵을 출력할 수 있다. 다만 객체 클래스 판별부(400)는 객체 클래스 판별부(400)에서 앵커별로 지정되는 경계 박스 영역에 대한 객체를 구분하여 검출해야 하므로, 객체 클래스 특징맵은 $H \times W \times KA$ 개로 획득될 수 있다. 그리고 $H \times W \times KA$ 개의 객체 클래스 특징맵으로부터 각 객체 영역에 대한 객체를 판별할 수 있다.
- [0078] 객체 클래스 판별부(400)와 경계 박스 추출부(500)가 출력 특징맵(F_0)으로부터 포함된 객체를 판별하고, 경계 박스를 추출하는 방식은 이외에도 다양하게 공개되어 있으므로 여기서는 상세하게 설명하지 않는다.
- [0079] 상기에서는 설명의 편의를 위하여, 객체 클래스 판별부(400)와 경계 박스 추출부(500)를 별도로 구분하였으나, 객체 클래스 판별부(400)와 경계 박스 추출부(500)는 객체 및 경계 박스 추출부로 통합될 수 있다.
- [0080] 상기한 본 실시예에 따른 객체 검출 장치에서 정규 변환맵 획득부(211)와 오프셋 보정맵 획득부(212)는 인공 신경망으로 구현되므로 미리 학습되어야 한다. 또한 커널 샘플링부(220)에 저장되는 메인 커널(W_g) 또한 학습에 의해 획득될 수 있다. 이에 본 실시예에 따른 객체 검출 장치는 이미지에 포함된 객체의 클래스와 경계 박스가 미리 주석된 다수의 학습 데이터를 기반으로 객체 검출 장치를 학습시킬 수 있다. 즉 학습 데이터를 입력 데이터로 인가받아 객체 클래스 판별부(400)에서 판별된 객체 클래스와 학습 데이터에 주석된 객체 클래스 사이의 차이를 판별하고, 경계 박스 추출부(500)에서 획득된 경계 박스와 학습 데이터에 주석된 객체 클래스 사이의 오차를 분석하여 손실을 계산하고, 계산된 손실이 저감되도록 역전파함으로써, 정규 변환맵 획득부(211)와 오프셋 보정맵 획득부(212)를 학습시키고, 메인 커널(W_g)을 획득할 수 있다. 이때, 인공 신경망으로 구현되는 입력 특징맵 획득부(100) 또한 함께 학습될 수 있다. 또한 상기에서는 메인 커널(W_g)이 커널 샘플링부(220)에 저장되는 것으로 설명하였으나, 메인 커널(W_g)은 변환 파라미터 추출부(210)에 포함될 수도 있으며, 메인 커널(W_g)은 동적 커널 추출부(200)와 별도로 기존의 단일 스테이지 객체 검출 기법에 따라 학습되어 획득될 수도 있다.
- [0081] 도 5는 도 1의 샘플링부에서 각 객체의 픽셀별로 샘플링되어 추출된 동적 커널과 이에 따라 추출된 경계 박스의 일 예를 나타낸다.

- [0082] 도 5를 참조하면, 본 실시예에 따른 객체 검출 장치는 (a)와 같이 마이크와 운전자가 결합된 객체가 포함된 입력 이미지에서 (b) 내지 (e)의 각 픽셀에 대해, 우측의 (b) 내지 (e)에 도시된 바와 같이, 객체의 구조적 특징이 반영된 서로 다른 동적 커널을 추출할 수 있다.
- [0083] 또한 (f)와 같이 버스가 포함된 입력 이미지에서 (g) 내지 (j)의 각 픽셀에 대해서도 우측의 (g) 내지 (j)에 도시된 바와 같이, 객체의 구조적 특징이 반영된 서로 다른 동적 커널을 추출할 수 있다.
- [0084] 이와 같이 본 실시예에 따른 객체 검출 장치는 각 픽셀에 대해 서로 다른 동적 커널을 적용함에 따라 단일 스테이지 객체 검출 기법과 마찬가지로 객체 특징맵을 별도로 추출하지 않고서도, (a) 및 (f)에 도시된 바와 같이 정확한 경계 박스를 검출할 수 있다. 또한 객체 특징맵을 별도로 추출하여 객체를 검출하지 않으므로, 객체 특징맵 추출에 따른 픽셀의 객체 정보 누락을 방지하여 더욱 정확하게 객체를 검출할 수 있다.
- [0085] 도 6은 본 발명의 일 실시예에 따른 동적 객체 기반 객체 검출 방법을 나타낸다.
- [0086] 도 1 내지 도 5를 참조하여, 도 6의 동적 객체 기반 객체 검출 방법을 설명하면, 우선 검출해야 할 객체가 포함된 입력 이미지를 획득한다(S11). 그리고 패턴 추정 방식이 미리 학습된 인공 신경망을 이용하여, 입력 이미지의 특징을 추출하여 입력 특징맵(F_I)을 획득한다(S12). 여기서 입력 특징맵(F_I)의 각 픽셀(i)은 입력 이미지의 대응하는 픽셀과 주변 픽셀의 패턴에 따른 특징을 추출한 픽셀값을 가진다. 즉 입력 이미지에서 대응하는 객체의 특징 정보가 포함된다.
- [0087] 입력 특징맵(F_I)이 획득되면, 입력 특징맵(F_I)에서 객체 검출을 위한 특징을 추출하기 위해 학습에 의해 미리 획득된 메인 커널(W_g)에서 샘플링될 위치 정보가 기지정된 개수의 원소로 포함된 디폴트 파라미터(g)를 기지정된 다수의 스케일 파라미터(S) 중 하나의 스케일 파라미터에 따라 스케일 변환하여 정규 파라미터($S \cdot g$)를 획득하기 위한 정규 변환맵을 패턴 추정 방식이 미리 학습된 인공 신경망을 이용하여 추출한다(S13).
- [0088] 여기서 스케일 파라미터(S)는 디폴트 파라미터(g)의 원소들이 나타내는 위치 정보 사이의 거리를 스케일 파라미터(S)에 대응하는 비율로 변경하여, 메인 커널(W_g)에서 객체의 크기에 대응하는 가중치를 샘플링할 수 있도록 하는 정규 파라미터($S \cdot g$)를 획득하기 위한 파라미터이다.
- [0089] 이와 함께 입력 특징맵(F_I) 또는 정규 변환맵으로부터, 디폴트 파라미터(g)가 스케일 파라미터(S)에 대응하는 비율로 변경된 정규 파라미터($S \cdot g$)의 원소들이 나타내는 위치 정보 각각의 오프셋을 개별적으로 보정하기 위한 오프셋 보정맵을 패턴 추정 방식이 미리 학습된 인공 신경망을 이용하여 추출한다(S13).
- [0090] 여기서 오프셋(Δg)은 정규 파라미터($S \cdot g$)의 원소들이 나타내는 위치 정보 각각을 객체의 형상에 대응하는 위치 정보로 변경하여, 메인 커널(W_g)에서 객체의 형상에 대응하는 가중치를 샘플링할 수 있도록 하는 변환 파라미터($S \cdot g + \Delta g$)를 획득하기 위한 파라미터이다.
- [0091] 입력 특징맵(F_I)으로부터 정규 변환맵과 오프셋 보정맵이 획득되면, 정규 변환맵과 오프셋 보정맵이 결합된 변환 파라미터맵을 획득한다(S15).
- [0092] 그리고 획득된 변환 파라미터맵으로부터 메인 커널(W_g)을 입력 특징맵(F_I)의 다수의 픽셀(i) 각각에 대응하여 적응적으로 샘플링하기 위한 변환 파라미터($\tilde{g}(i) = S \cdot g(i) + \Delta g(i)$)를 추출한다(S16).
- [0093] 입력 특징맵(F_I)의 각 픽셀(i)에 대응하는 변환 파라미터($\tilde{g}(i)$)가 획득되면, 획득된 변환 파라미터를 기반으로 메인 커널(W_g)을 샘플링하여, 입력 특징맵(F_I)의 각 픽셀(i)에 대응하는 동적 커널($W_g(\tilde{g}(i))$)을 획득한다(S17).
- [0094] 그리고 패턴 추정 방식이 미리 학습된 인공 신경망을 이용하여 입력 특징맵(F_I)과 획득된 픽셀별 동적 커널($W_g(\tilde{g}(i))$)에 대해 기지정된 연산을 수행하여 출력 특징맵을 획득한다(S18). 여기서 출력 특징맵은 입력 특징맵(F_I)의 각 픽셀이 나타내는 객체의 구조적 특징이 반영된 동적 커널($W_g(\tilde{g}(i))$)을 이용하여 입력 특징맵(F_I)의 특징을 추출함으로써 획득되는 특징맵으로서, 객체의 분류 클래스뿐만 아니라 객체의 구조적인 특징이 잘 반영된 특징맵으로 볼 수 있다.

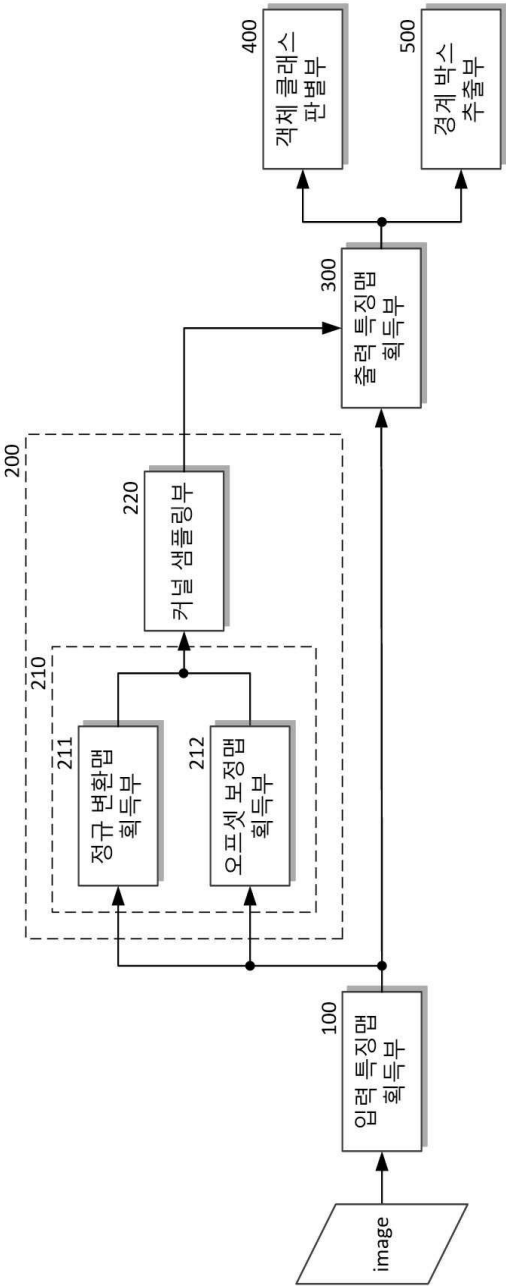
- [0095] 이에 패턴 추정 방식이 미리 학습된 인공 신경망을 이용하여 출력 특징맵에서 객체 영역을 추출하고, 추출된 객체 영역을 입력 이미지에 대응하는 형태로 회귀시킴으로써, 입력 이미지에서 객체가 포함된 영역을 나타내는 경계 박스를 추출하고, 각 객체 영역에서의 출력 특징맵의 패턴을 추정함으로써 각 객체 영역에 포함된 객체의 클래스를 판별할 수 있다(S19). 즉 입력 이미지에서 객체가 포함된 영역을 경계 박스로 추출하고, 각 경계 박스에 포함된 객체를 판별할 수 있다.
- [0096] 본 발명에 따른 방법은 컴퓨터에서 실행시키기 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 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.
- [0097] 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.
- [0098] 따라서, 본 발명의 진정한 기술적 보호 범위는 첨부된 청구범위의 기술적 사상에 의해 정해져야 할 것이다.

부호의 설명

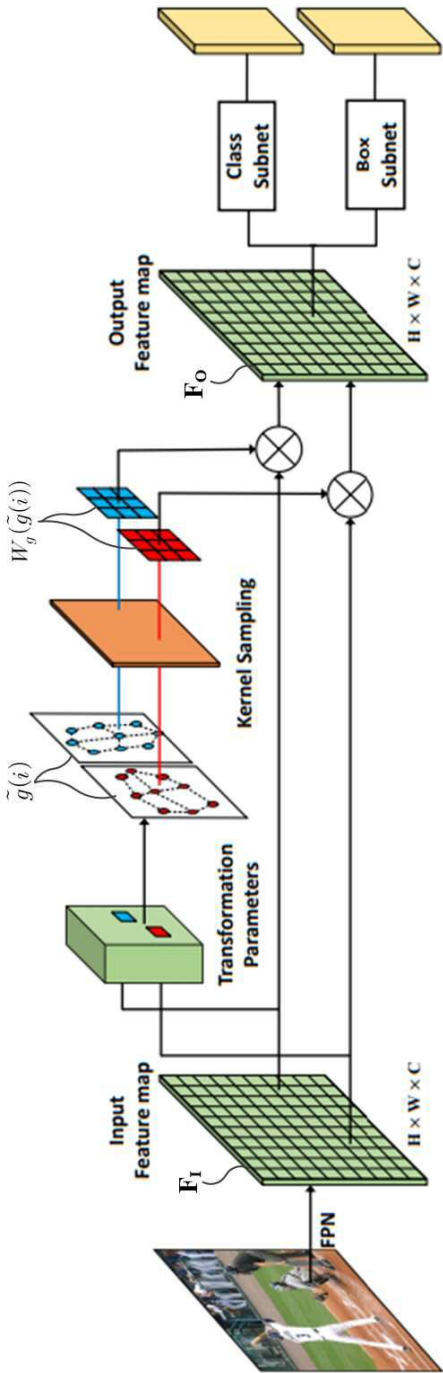
- | | | |
|--------|------------------|-----------------|
| [0099] | 100: 입력 특징맵 획득부 | 200: 동적 커널 추출부 |
| | 210: 변환 파라미터 추출부 | 211: 정규 변환맵 획득부 |
| | 212: 오프셋 보정맵 획득부 | 220: 커널 샘플링부 |
| | 300: 출력 특징맵 획득부 | 400: 객체 클래스 판별부 |
| | 500: 경계 박스 추출부 | |

도면

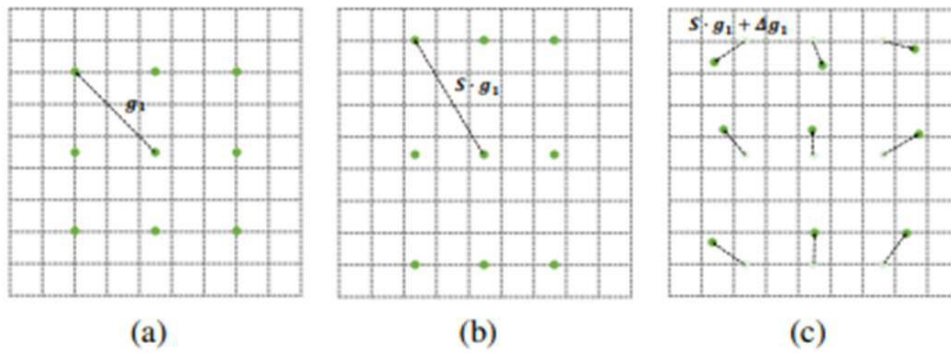
도면1



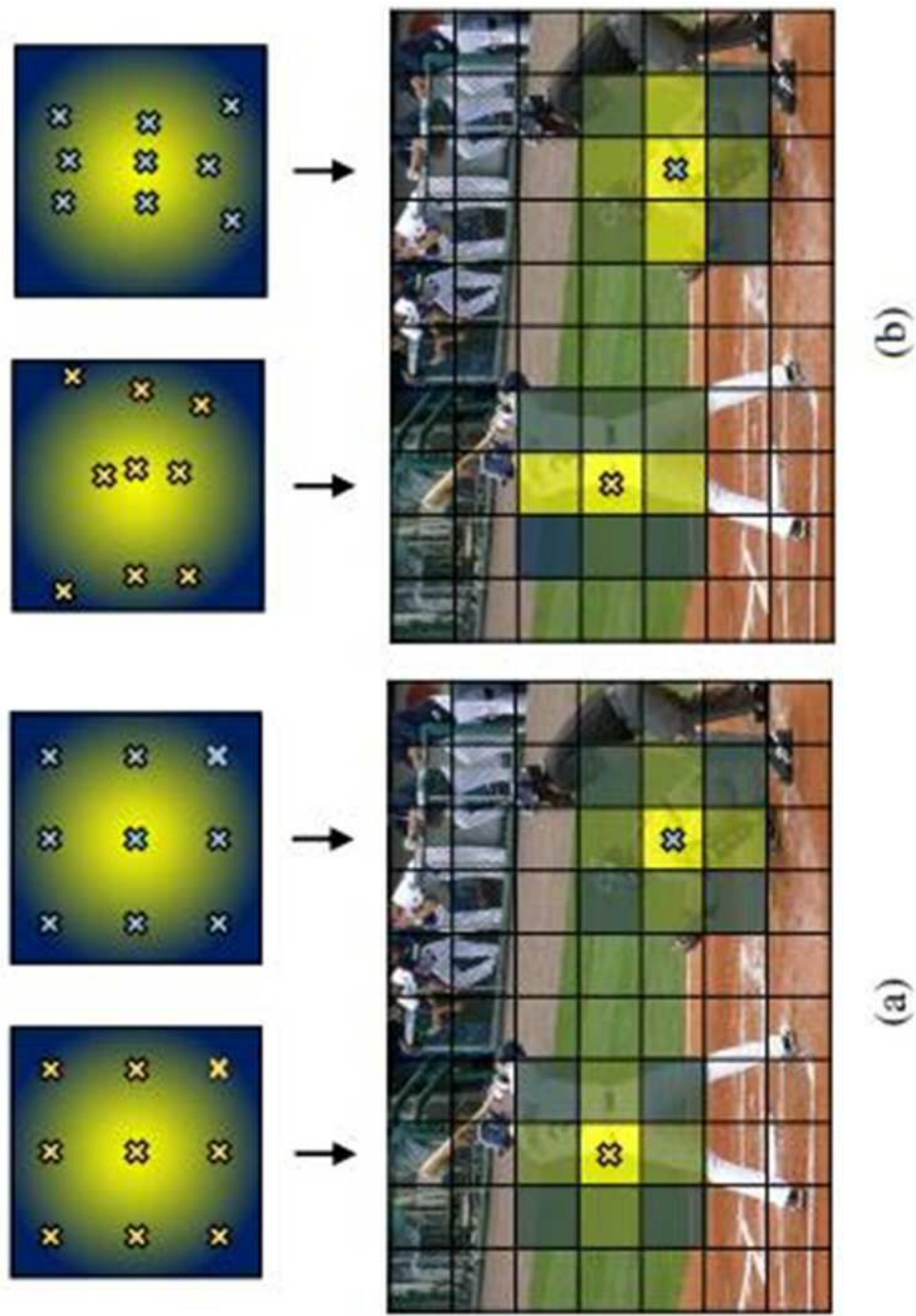
도면2



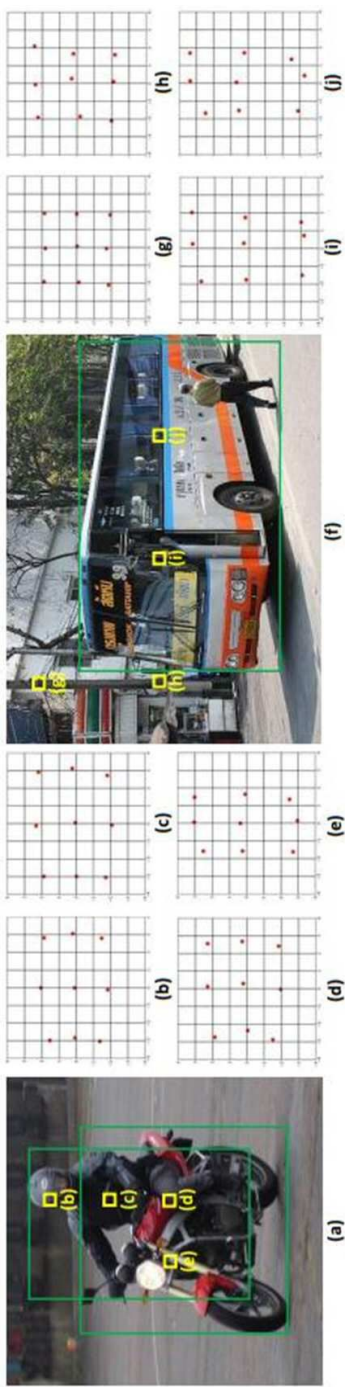
도면3



도면4



도면5



도면6

