(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2023-0051036
(43) 공개일자 2023년04월17일

(51) 국제특허분류(Int. Cl.)
G06T 7/11 (2017.01) *G06N 3/04* (2023.01)
G06N 3/08 (2023.01) *G06V 10/40* (2022.01)
(52) CPC특허분류
G06T 7/11 (2017.01)
G06N 3/045 (2023.01)
(21) 출원번호 10-2022-0006514
(22) 출원일자 2022년01월17일
심사청구일자 2022년01월17일
(30) 우선권주장
1020210134298 2021년10월08일 대한민국(KR)

(71) 출원인
고려대학교 산학협력단
서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)
(72) 발명자
고성제
서울특별시 서초구 방배로 270, 나-304호 (방배동, 삼호아파트)
조성진
경기도 성남시 수정구 위례중앙로 216, 5313동 1102호 (창곡동, 위례 호반베르디움)
이소열
충청남도 공주시 관골1길 31-8,106동 1106호(신관동, 삼환나우빌)
(74) 대리인
윤귀상

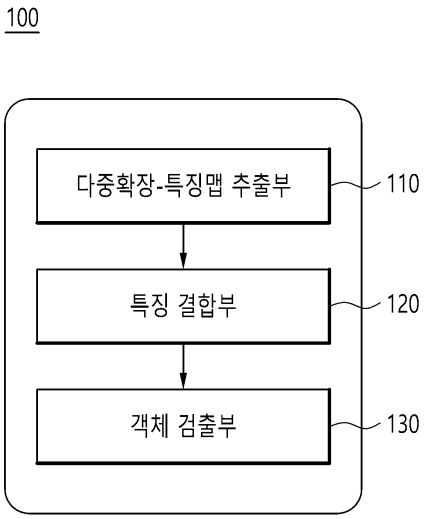
전체 청구항 수 : 총 11 항

(54) 발명의 명칭 **관심객체 검출장치, 방법 및 이를 위한 컴퓨터 판독가능 프로그램**

(57) 요 약

관심객체 검출장치, 방법 및 이를 위한 컴퓨터 판독가능 프로그램이 개시된다. 본 발명에 따른 관심객체 검출장치는 입력 이미지 상에서 적어도 하나 이상의 앵커 박스에 기반하여 관심 객체를 검출하는 관심객체 검출장치로서, 상기 입력 이미지로부터 상기 앵커 박스의 형상에 기초한 복수의 다중확장- 특징맵을 추출하는 다중확장-특징맵 추출부, 각각의 다중확장-특징맵에서 추출된 특징들을 결합하여 결합 특징맵을 생성하는 특징결합부 및 상기 결합 특징맵으로부터 오프셋(offset)을 예측하여 상기 관심 객체를 검출하는 객체검출부를 포함한다. 따라서, 상술한 본 발명에 따르면, 입력 이미지에서 앵커 박스의 다양한 형상에 기초한 다중확장-특징맵들을 추출하고, 이들의 특징들을 상호 결합한 결합 특징맵으로부터 관심 객체를 검출하므로 객체의 형상에 보다 강인하게 객체 검출이 수행될 수 있다.

대 표 도 - 도3



(52) CPC특허분류
G06N 3/08 (2023.01)
G06V 10/40 (2023.01)
G06T 2210/12 (2013.01)

이 발명을 지원한 국가연구개발사업
과제고유번호 1711126469
과제번호 2014-3-00077-008
부처명 과학기술정보통신부
과제관리(전문)기관명 정보통신기획평가원
연구사업명 혁신성장동력프로젝트(R&D)(과기정통부)
연구과제명 (딥뷰-2세부) 대규모 실시간 비디오 분석에 의한 전역적 다중 관심객체 추적 및 상황 예측 기술 개발
기 여 율 1/1
과제수행기관명 광주과학기술원
연구기간 2021.01.01 ~ 2021.12.31

명 세 서

청구범위

청구항 1

입력 이미지 상에서 적어도 하나 이상의 앵커 박스에 기반하여 관심 객체를 검출하는 관심객체 검출장치로서,

상기 입력 이미지로부터 상기 앵커 박스의 형상에 기초한 복수의 다중확장-특징맵을 추출하는 다중확장-특징맵 추출부;

각각의 다중확장-특징맵에서 추출된 특징들을 결합하여 결합 특징맵을 생성하는 특징결합부; 및

상기 결합 특징맵으로부터 오프셋(offset)을 예측하여 상기 관심 객체를 검출하는 객체검출부를 포함하는, 관심객체 검출장치.

청구항 2

제 1 항에 있어서,

상기 다중확장-특징맵 추출부는,

상기 앵커박스의 형상에 대응하는 비율을 갖는 확장된 컨볼루션(convolution) 연산을 수행하여 상기 다중확장-특징맵을 추출하는 것인, 관심객체 검출장치.

청구항 3

제 2 항에 있어서,

상기 다중확장-특징맵 추출부는,

상기 앵커박스의 종횡비에 대응하는 복수의 컨볼루션 커널(convolution kernel)을 적용한 상기 확장된 컨볼루션 연산을 수행하는 것인, 관심객체 검출장치.

청구항 4

제 3 항에 있어서,

상기 특징결합부는,

상기 복수의 다중확장-특징맵에 $1^{\times}1$ 컨볼루션 연산을 적용하여 출력 특징들을 추출하고,

상기 출력 특징들을 분할하고 분할된 출력 특징들을 상호 결합하여 복수의 결합 특징맵을 생성하는 것인, 관심객체 검출장치.

청구항 5

제 4 항에 있어서,

상기 객체 검출부는,

상기 복수의 결합 특징맵 각각에 고정된 크기의 컨볼루션 커널을 적용하여 컨볼루션 연산을 수행함으로써 상기 오프셋을 예측하는 것인, 관심객체 검출장치.

청구항 6

관심객체 검출장치에서의 입력 이미지 상에서 적어도 하나 이상의 앵커 박스에 기반하여 관심 객체를 검출하는 방법으로서,

상기 입력 이미지로부터 상기 앵커 박스의 형상에 기초한 복수의 다중확장- 특징맵을 추출하는 단계;

각각의 다중확장-특징맵에서 추출된 특징들을 결합하여 결합 특징맵을 생성하는 단계; 및

상기 결합 특징맵으로부터 오프셋(offset)을 예측하여 상기 관심 객체를 검출하는 단계를 포함하는, 관심객체 검출방법.

청구항 7

제 6 항에 있어서,

상기 다중확장-특징맵을 추출하는 단계는,

상기 앵커박스의 형상에 대응하는 비율을 갖는 확장된 컨볼루션(convolution) 연산을 수행하여 상기 다중확장-특징맵을 추출하는 것인, 관심객체 검출방법.

청구항 8

제 7 항에 있어서,

상기 다중확장-특징맵을 추출하는 단계는,

상기 앵커박스의 종횡비에 대응하는 복수의 컨볼루션 커널(convolution kernel)을 적용한 상기 확장된 컨볼루션 연산을 수행하는 것인, 관심객체 검출방법.

청구항 9

제 8 항에 있어서,

상기 결합 특징맵을 생성하는 단계는,

상기 복수의 다중확장-특징맵에 1×1 컨볼루션 연산을 적용하여 출력 특징들을 추출하고,

상기 출력 특징들을 분할하고 분할된 출력 특징들을 상호 결합하여 복수의 결합 특징맵을 생성하는 것인, 관심객체 검출방법.

청구항 10

제 9 항에 있어서,

상기 관심 객체를 검출하는 단계는,

상기 복수의 결합 특징맵 각각에 고정된 크기의 컨볼루션 커널을 적용하여 컨볼루션 연산을 수행함으로써 상기 오프셋을 예측하는 것인, 관심객체 검출방법.

청구항 11

제 6 항 내지 제 10 항 중 어느 한 항에 따른 관심객체 검출방법을 실행하도록 구성된, 컴퓨터로 판독가능한 기록매체에 저장된 컴퓨터 판독가능한 프로그램.

발명의 설명

기 술 분 야

[0001] 본 발명은 관심객체 검출장치, 방법 및 이를 위한 컴퓨터 판독가능 프로그램에 관한 것으로서, 보다 상세하게는 입력 이미지 상에서 앵커 박스에 기반하여 관심객체를 검출하는 관심객체 검출장치, 방법 및 이를 위한 컴퓨터 판독가능 프로그램이다.

배 경 기 술

[0003] 객체 검출(object detection) 분야는 컴퓨터 비전 분야에서 가장 활발히 연구되고 있는 분야 중 하나이다.

[0004] 객체 검출이란 영상 내에서 찾고자 하는 객체의 종류를 자동으로 구분하고, 구분된 객체의 위치와 영역을 자동

으로 바운딩 박스(bounding box)로 표현하는 것을 말한다.

[0005] 최근에 이러한 객체 검출 기술이 자율주행과 보안 감시 기술이 성장함에 따라 그 중요성이 대두되고 있다.

[0006] 특히, 자율주행에서의 객체 검출기술은 주행을 하면서 도로에서 보이는 자동차 또는 보행자 등을 자동으로 검출하여 운전자에게 도로상태, 교통정보 등의 정보를 제공한다. 이러한 장점 때문에, 주행환경에서의 객체 검출은 자율주행 자동차, 자동 감지 시스템 등 다양한 어플리케이션으로의 활용 가능성이 높아 관련 시장 규모가 빠르게 확대되고 있다.

[0007] 따라서, 보다 정확하게 객체를 검출할 수 있는 객체 검출 기술이 필요한 실정이다.

선행기술문헌

특허문헌

[0008] (특허문헌 0001) 한국 등록특허공보 제10-1480065호

발명의 내용

해결하려는 과제

[0009] 본 발명의 일 측면에 따르면, 입력 이미지로부터 앵커박스의 형상에 기초한 다중확장-특징맵을 추출하고, 이렇게 추출된 특징들을 결합하여 결합 특징맵을 생성하며 이러한 결합 특징맵으로부터 오프셋을 예측하여 관심객체를 검출하는 관심객체 검출장치, 방법 및 이를 위한 컴퓨터 판독가능 프로그램을 제공하는데 그 목적이 있다.

과제의 해결 수단

[0010] 본 발명의 일 실시예에 따른 관심객체 검출장치는 입력 이미지 상에서 적어도 하나 이상의 앵커 박스에 기반하여 관심 객체를 검출하는 관심객체 검출장치로서, 상기 입력 이미지로부터 상기 앵커 박스의 형상에 기초한 복수의 다중확장- 특징맵을 추출하는 다중확장-특징맵 추출부, 각각의 다중확장-특징맵에서 추출된 특징들을 결합하여 결합 특징맵을 생성하는 특징결합부 및 상기 결합 특징맵으로부터 오프셋(offset)을 예측하여 상기 관심객체를 검출하는 객체검출부를 포함한다.

[0011] 한편, 다중확장-특징맵 추출부는, 상기 앵커박스의 형상에 대응하는 비율을 갖는 확장된 컨볼루션(convolution) 연산을 수행하여 상기 다중확장-특징맵을 추출하는 것일 수 있다.

[0012] 또한, 다중확장-특징맵 추출부는, 상기 앵커박스의 종횡비에 대응하는 복수의 컨볼루션 커널(convolution kernel)을 적용한 상기 확장된 컨볼루션 연산을 수행하는 것일 수 있다.

[0013] 또한, 특징결합부는, 상기 복수의 다중확장-특징맵에 1×1 컨볼루션 연산을 적용하여 출력 특징들을 추출하고, 상기 출력 특징들을 분할하고 분할된 출력 특징들을 상호 결합하여 복수의 결합 특징맵을 생성하는 것일 수 있다.

[0014] 또한, 객체 검출부는, 상기 복수의 결합 특징맵 각각에 고정된 크기의 컨볼루션 커널을 적용하여 컨볼루션 연산을 수행함으로써 상기 오프셋을 예측하는 것일 수 있다.

[0015] 본 발명의 다른 실시예에 따른 관심객체 검출방법은 관심객체 검출장치에서의 입력 이미지 상에서 적어도 하나 이상의 앵커 박스에 기반하여 관심 객체를 검출하는 방법으로서, 상기 입력 이미지로부터 상기 앵커 박스의 형상에 기초한 복수의 다중확장- 특징맵을 추출하는 단계, 각각의 다중확장-특징맵에서 추출된 특징들을 결합하여 결합 특징맵을 생성하는 단계 및 상기 결합 특징맵으로부터 오프셋(offset)을 예측하여 상기 관심 객체를 검출하는 단계를 포함한다.

[0016] 한편, 다중확장-특징맵을 추출하는 단계는, 상기 앵커박스의 형상에 대응하는 비율을 갖는 확장된 컨볼루션(convolution) 연산을 수행하여 상기 다중확장-특징맵을 추출하는 것일 수 있다.

[0017] 또한, 다중확장-특징맵을 추출하는 단계는, 상기 앵커박스의 종횡비에 대응하는 복수의 컨볼루션 커널(convolution kernel)을 적용한 상기 확장된 컨볼루션 연산을 수행하는 것일 수 있다.

[0018] 또한, 결합 특징맵을 생성하는 단계는, 상기 복수의 다중확장-특징맵에 1×1 컨볼루션 연산을 적용하여 출력 특징들을 추출하고, 상기 출력 특징들을 분할하고 분할된 출력 특징들을 상호 결합하여 복수의 결합 특징맵을 생성하는 것일 수 있다.

[0019] 또한, 관심 객체를 검출하는 단계는, 상기 복수의 결합 특징맵 각각에 고정된 크기의 컨볼루션 커널을 적용하여 컨볼루션 연산을 수행함으로써 상기 오프셋을 예측하는 것일 수 있다.

[0020] 또한, 본 발명의 또 다른 실시예는 관심객체 검출방법을 실행하도록 구성된, 컴퓨터로 판독가능한 기록매체에 저장된 컴퓨터 판독가능한 프로그램을 포함할 수 있다.

발명의 효과

[0021] 상술한 본 발명에 따르면, 입력 이미지에서 앵커 박스의 다양한 형상에 기초한 다중확장-특징맵들을 추출하고 이들의 특징들을 상호 결합한 결합 특징맵으로부터 관심 객체를 검출하므로, 객체의 형상에 보다 강인하게 객체를 검출할 수 있다.

[0022] 따라서, 기존의 관심객체 검출장치에서 검출하지 못한 객체에 대해서도 검출이 가능하므로, 관심객체 검출장치의 성능이 큰 폭으로 향상될 수 있다.

도면의 간단한 설명

[0023] 도 1 및 도 2 는 종래의 관심 객체를 검출하는 장치에서의 관심 객체를 검출하는 방법을 설명하기 위한 도면이다.

도 3 은 본 발명의 일 실시예에 따른 관심객체 검출장치의 구성을 도시한 도면이다.

도 4 는 도 3 에 도시된 관심객체 검출장치에서의 관심객체 검출기작을 모식화한 도면이다.

도 5 는 도 3 에 도시된 다중확장-특징맵 추출부에서의 다중확장-특징맵을 추출하는 과정을 모식화한 도면이다.

도 6 은 도 3 에 도시된 특징 결합부에서의 결합 특징맵을 생성하는 과정을 모식화한 도면이다.

도 7 은 도 3 에 도시된 객체 검출부에서의 오프셋을 예측하는 과정을 모식화한 도면이다.

도 8 은 본 발명의 다른 실시예에 따른 관심객체 검출방법을 도시한 순서도이다.

도 9 는 도 8 에 따른 관심객체 검출방법과 종래의 방법에 따라 관심객체를 검출한 결과의 일 예를 도시한 도면이다.

발명을 실시하기 위한 구체적인 내용

[0024] 후술하는 본 발명에 대한 상세한 설명은, 본 발명이 실시될 수 있는 특정 실시예를 예시로서 도시하는 첨부 도면을 참조한다. 이들 실시예는 당업자가 본 발명을 실시할 수 있기에 충분하도록 상세히 설명된다. 본 발명의 다양한 실시예는 서로 다르지만 상호 배타적일 필요는 없음이 이해되어야 한다. 예를 들어, 여기에 기재되어 있는 특정 형상, 구조 및 특성은 일 실시예와 관련하여 본 발명의 정신 및 범위를 벗어나지 않으면서 다른 실시예로 구현될 수 있다. 또한, 각각의 개시된 실시예 내의 개별 구성요소의 위치 또는 배치는 본 발명의 정신 및 범위를 벗어나지 않으면서 변경될 수 있음이 이해되어야 한다. 따라서, 후술하는 상세한 설명은 한정적인 의미로서 취하려는 것이 아니며, 본 발명의 범위는, 적절하게 설명된다면, 그 청구항들이 주장하는 것과 균등한 모든 범위와 더불어 첨부된 청구항에 의해서만 한정된다. 도면에서 유사한 참조부호는 여러 측면에 걸쳐서 동일하거나 유사한 기능을 지칭한다.

[0025] 이하, 도면들을 참조하여 본 발명의 바람직한 실시예들을 보다 상세하게 설명하기로 한다.

[0026] 도 1 및 도 2 는 관심 객체를 검출하는 장치에서의 관심 객체를 검출하는 방법을 설명하기 위한 도면이다.

[0027] 관심객체 검출장치는 입력이미지 상에서 관심객체를 실시간으로 검출하는 장치로서, 예컨대 SSD(Single Shot multi-box detector) 또는 Refine detector 일 수 있다.

[0028] 이러한 관심객체 검출장치는 도 1 에서와 같이, 다수의 검출헤드(Detection Head)를 구비하며, 각 검출헤드에서는 앵커 박스(Anchor box)의 크기와 실제 객체의 크기(ground truth box)와 차이에 기초하여 관심객체를 검출한

다. 이를 위해, 각 검출헤드는 도 2 에서와 같이, 입력 특징맵(Feature map)에 일괄적으로 3×3 컨볼루션 연산을 적용하여 최종적인 오프셋을 예측(Offset prediction)한다.

[0029] 한편, 이러한 관심객체 검출장치에서 사용하는 앵커박스는 입력 이미지 내에 사전에 존재하는 박스로서 다양한 형상, 즉, 다양한 크기와 종횡비를 갖는다. 예컨대, 300×300 의 입력크기를 갖는 SSD 의 경우 8700 개의 앵커 박스가 사전에 존재할 수 있다. 따라서, 앵커박스의 다양한 형상을 고려하여 관심객체를 검출할 필요성이 있다.

[0030] 이하에서는 이러한 앵커박스의 다양한 형상에 기초하여 관심객체를 검출하는 장치에 대해 설명한다.

[0031] 도 3 은 본 발명의 일 실시예에 따른 관심객체 검출장치의 구성을 도시한 도면이고, 도 4 는 도 3 에 도시된 관심객체 검출장치에서의 관심객체 검출기작을 모식화한 도면이며, 도 5 는 도 3 에 도시된 다중확장-특징맵 추출부에서의 다중확장-특징맵을 추출하는 과정을 모식화한 도면이고, 도 6 은 도 3 에 도시된 특징 결합부에서의 결합 특징맵을 생성하는 과정을 모식화한 도면이며, 도 7 은 도 3 에 도시된 객체 검출부에서의 오프셋을 예측하는 과정을 모식화한 도면이다.

[0032] 본 실시예에 따른 관심객체 검출장치(100)는 입력 이미지 상에서 적어도 하나 이상의 앵커 박스에 기반하여 관심 객체를 검출하는 관심객체 검출장치(100)로서, 다중확장-특징맵 추출부(110), 특징결합부(120) 및 객체검출부(130)를 포함한다.

[0033] 다중확장-특징맵 추출부(110)는 입력 이미지로부터 앵커 박스의 형상에 기초한 복수의 다중확장- 특징맵(multi-dilated feature maps)을 추출한다.

[0034] 다중확장-특징맵 추출부(110)는 앵커박스의 형상에 대응하는 비율을 갖는 확장된 컨볼루션(convolution) 연산을 수행하여 상기 다중확장-특징맵을 추출할 수 있다.

[0035] 보다 구체적으로, 다중확장-특징맵 추출부(110)는, 앵커박스의 종횡비에 대응하는 복수의 컨볼루션 커널(convolution kernel)을 적용한 확장된 컨볼루션 연산을 수행할 수 있다.

[0036] 도 5 를 참조하면, 다중확장-특징맵 추출부(110)는 1:1, 1:2, 2:1 의 종횡비를 갖는 앵커박스에 대응하는 컨볼루션 커널을 선택하고, 상기 선택된 컨볼루션 커널과 확장된 컨볼루션 연산을 수행하여 복수의 다중확장-특징맵(m_1, m_2, m_3, m_4)을 생성할 수 있다. 먼저, 다중확장-특징맵 추출부(110)는 1:1 의 종횡비를 갖는 앵커박스(빨간색 앵커박스)의 형상을 고려한 컨볼루션 커널(d=(1,1)) 과 입력 특징맵을 3×3 컨볼루션 연산하여 제 1 다중확장-특징맵(m_1)을 추출한다. 또한, 다중확장-특징맵 추출부는 1:2 의 종횡비를 갖는 앵커박스(노란색 앵커박스)의 형상을 고려한 컨볼루션 커널(d=(1,2)) 과 입력 특징맵을 3×3 컨볼루션 연산하여 제 2 다중확장-특징맵(m_2)을 추출하며, 2:1 의 종횡비를 갖는 앵커박스(보라색 앵커박스)의 형상을 고려한 컨볼루션 커널(d=(2,1)) 과 입력 특징맵을 3×3 컨볼루션 연산하여 제 3 다중확장-특징맵(m_3)을 추출한다. 또한, 다중확장-특징맵 추출부는 와이드-스케일(wide-scale) 객체의 특징을 보완하기 위해 컨볼루션 커널(d=(2,2)) 과 입력 특징맵을 3×3 컨볼루션 연산하여 제 4 다중확장-특징맵(m_4)을 추출한다.

[0037] 한편, 도 5 에서는 1:1, 1:2, 2:1 의 종횡비를 갖는 앵커박스의 형상을 고려하여 다중확장-특징맵(m_1, m_2, m_3, m_4)을 추출하였으나, 이는 일 예일 뿐, 다양한 종횡비를 갖는 앵커박스의 형상을 고려하여 다중확장-특징맵을 추출할 수 있음은 물론이다.

[0038] 특징결합부(120)는 각각의 다중확장-특징맵에서 추출된 특징들을 결합하여 결합 특징맵을 생성한다.

[0039] 보다 구체적으로, 특징결합부(120)는, 복수의 다중확장-특징맵에 1×1 컨볼루션 연산을 적용하여 감소된 채널의 출력 특징들을 추출하고, 상기 출력 특징들을 분할하고 미리 설정된 크기의 결합 채널로 결합하여 복수의 결합 특징맵을 생성하는 것일 수 있다.

[0040] 도 6 을 참조하면, 먼저, 특징결합부(120)는 각 다중확장-특징맵(m_1, m_2, m_3, m_4)에 1×1 컨볼루션 연산을 적용하여 출력 특징들(F_1, F_2, F_3, F_4)을 추출한다. 보다 구체적으로, 특징결합부(120)는 제 1 다중확장-특징맵(m_1)에 1×1 컨볼루션 연산을 적용하여 256 채널의 제 1 출력특징($F_1 = \{ f_1^1, f_2^1, f_3^1, f_4^1 \}$)을 추출하고, 제 2 다중

확장-특징맵(m_2)에 1×1 컨볼루션 연산을 적용하여 128 채널의 제 2 출력특징(F_2={ f_1^2, f_2^2})을 추출한다. 또한, 특징결합부(120)는 제 3 다중확장-특징맵(m_3)에 1×1 컨볼루션 연산을 적용하여 128 채널의 제 3 출력특징(F_3 ={f_1^3, f_2^3 })을 추출하고, 제 4 다중확장-특징맵(m_4)에 1×1 컨볼루션 연산을 적용하여 256 채널의 제 4 출력특징(F_4 ={$f_1^4, f_2^4, f_3^4, f_4^4$ })을 추출한다.

[0041] 또한, 특징결합부(120)는 이렇게 추출된 출력 특징들(F_1 , F_2 , F_3 , F_4)을 분할하고 일정크기의 사이즈를 갖는 결합채널로 결합하여 복수의 결합 특징맵(G_1, G_2, G_3)을 생성한다. 즉, 특징결합부(120)는 서로 다른 형상을 갖는 앵커박스로부터의 출력특징들이 상호 결합된 결합 특징맵(G_i)을 생성하는 것이다. 이러한 결합 특징맵(G_i) 은 아래 수학식 1 과 같이 정의될 수 있다.

[0042] [수학식 1]

[0043] $$G_i = \begin{cases} \{f_1^1, f_2^1, f_1^{N+1}, f_2^{N+1}\}, & i = 1, \\ \{f_1^i, f_2^i, f_{i+1}^1, f_{i+1}^{N+1}\}, & i = 2,3 \end{cases}$$

[0044] 여기서, N 은 앵커박스의 개수이다. 한편, 본 실시예에서는 3 개의 앵커박스를 고려한바 N 은 3이며, 이에 한정되지 않는다.

[0045] 즉, 제 1 출력특징들(F_1) 과 제 4 출력특징들(F_4)은 분할되어 제 1 결합특징맵(G_1), 제 2 결합특징맵(G_2) 및 제 3 결합특징맵(G_3)의 구성성분이 되고, 제 2 출력특징들(F_2) 과 제 3 출력특징들(F_3)은 분할된 제 1 출력특징들(F_1) 과 제 4 출력특징들(F_4) 중 적어도 하나 이상과 결합하여 각각 제 2 결합특징맵(G_2) 및 제 3 결합특징맵(G_3)으로 생성되는 것이다.

[0046] 객체검출부(130)는 결합 특징맵으로부터 오프셋(offset)을 예측하여 관심 객체를 검출한다.

[0047] 이를 위해, 객체검출부(130)는 복수의 결합 특징맵 각각에 고정된 크기의 컨볼루션 커널을 적용하여 3×3 컨볼루션 연산을 수행함으로써 오프셋을 예측할 수 있다. 즉, 객체검출부(130)는 서로 다른 형상의 앵커박스에 기초한 출력특징들이 상호 결합된 결합 특징맵 각각에 동일한 컨볼루션 커널을 적용하여 최종 오프셋을 예측하고 이를 통해 관심객체를 검출하는 것이다.

[0048] 도 8 은 본 발명의 다른 실시예에 따른 관심객체 검출방법을 도시한 순서도이다.

[0049] 본 실시예에 따른 관심객체 검출방법은 관심객체 검출장치에서의 입력 이미지 상에서 적어도 하나 이상의 앵커박스에 기반하여 관심 객체를 검출하는 방법으로서, 다중확장-특징맵 추출부에서 상기 입력 이미지로부터 상기 앵커 박스의 형상에 기초한 복수의 다중확장-특징맵을 추출하는 단계(S10), 특징결합부에서 각각의 다중확장-특징맵에서 추출된 특징들을 결합하여 결합 특징맵을 생성하는 단계(S20) 및 객체검출부에서 상기 결합 특징맵으로부터 오프셋(offset)을 예측하여 상기 관심 객체를 검출하는 단계(S30)를 포함한다.

[0050] 한편, 다중확장-특징맵을 추출하는 단계(S10)는, 상기 앵커박스의 형상에 대응하는 비율을 갖는 확장된 컨볼루션(convolution) 연산을 수행하여 상기 다중확장-특징맵을 추출하는 것일 수 있다. 또한, 다중확장-특징맵을 추출하는 단계(S10)는, 상기 앵커박스의 종횡비에 대응하는 복수의 컨볼루션 커널(convolution kernel)을 적용한 상기 확장된 컨볼루션 연산을 수행하는 것일 수 있다.

[0051] 또한, 결합 특징맵을 생성하는 단계(S20)는, 상기 복수의 다중확장-특징맵에 1×1 컨볼루션 연산을 적용하여 출력 특징들을 추출하고, 상기 출력 특징들을 분할하고 상기 출력 특징들을 분할하고 분할된 출력 특징들을 상호 결합하여 복수의 결합 특징맵을 생성하는 것일 수 있다.

[0052] 또한, 관심 객체를 검출하는 단계(S30)는, 상기 복수의 결합 특징맵 각각에 고정된 크기의 컨볼루션 커널을 적용하여 컨볼루션 연산을 수행함으로써 상기 오프셋 값을 산출하는 것일 수 있다.

[0053] 그 밖의 설명은 관심객체 검출장치와 동일한 바 이에 대한 설명은 도 3 내지 도 7 에 대한 설명으로 대체한다.

[0054] 도 9 는 종래의 방법에 따라 관심객체를 검출한 결과(도 9 의 (a)) 와 본 발명의 관심객체 검출방법에 따라 관

심객체를 검출한 결과(도 9 의 (b)) 를 도시한 도면으로서, 도 9 를 참조하면, 본 발명에 따라 관심객체를 검출하면 기존 기법에서 검출하지 못한 객체들이 다양하게 검출되는 모습을 확인할 수 있다.

[0055] 즉, 상술한 본 발명에 따르면, 입력 이미지에서 앵커 박스의 다양한 형상에 기초한 다중확장-특징맵들을 추출하고, 이들의 특징들을 상호 결합한 결합 특징맵으로부터 관심 객체를 검출하므로 객체의 형상에 보다 강인하게 객체 검출이 수행될 수 있다. 따라서, 기존의 관심객체 검출장치에서 검출하지 못한 객체에 대해서도 검출이 가능하므로, 관심객체 검출장치의 성능이 큰 폭으로 향상될 수 있다.

[0056] 이상에서 설명한 실시예들에 따른 관심객체 검출방법에 의한 동작은 적어도 부분적으로 컴퓨터 프로그램으로 구현되고 컴퓨터로 읽을 수 있는 기록매체에 기록될 수 있다. 실시예들에 따른 관심객체 검출동작을 구현하기 위한 프로그램이 기록되고 컴퓨터가 읽을 수 있는 기록매체는 컴퓨터에 의하여 읽혀질 수 있는 데이터가 저장되는 모든 종류의 기록장치를 포함한다. 컴퓨터가 읽을 수 있는 기록매체의 예로는 ROM, RAM, CD-ROM, 자기 테이프, 플로피디스크, 광 데이터 저장장치 등이 있다. 또한 컴퓨터가 읽을 수 있는 기록매체는 네트워크로 연결된 컴퓨터 시스템에 분산되어, 분산 방식으로 컴퓨터가 읽을 수 있는 코드가 저장되고 실행될 수도 있다. 또한, 본 실시예를 구현하기 위한 기능적인 프로그램, 코드 및 코드 세그먼트(segment)들은 본 실시예가 속하는 기술 분야의 통상의 기술자에 의해 용이하게 이해될 수 있을 것이다.

[0057] 이상에서는 실시예들을 참조하여 설명하였지만, 해당 기술 분야의 숙련된 당업자는 하기의 특허 청구범위에 기재된 본 발명의 사상 및 영역으로부터 벗어나지 않는 범위 내에서 본 발명을 다양하게 수정 및 변경시킬 수 있음을 이해할 수 있을 것이다.

부호의 설명

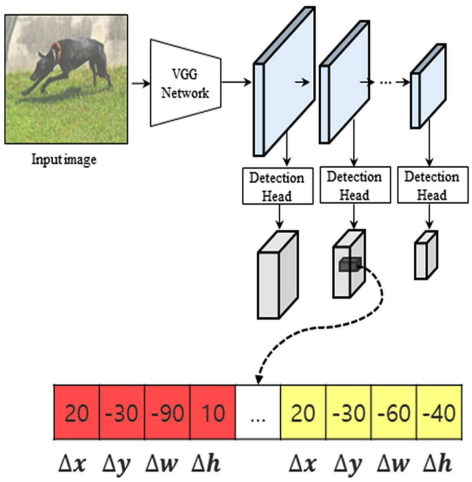
[0058] 110: 다중확장-특징맵 추출부

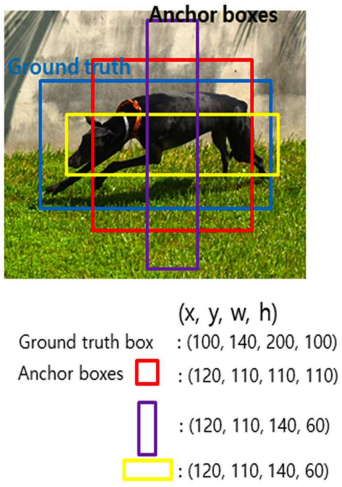
120: 특징결합부

130: 객체검출부

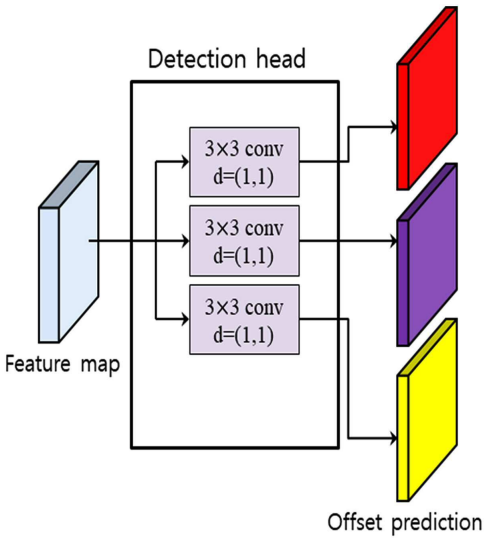
도면

도면1

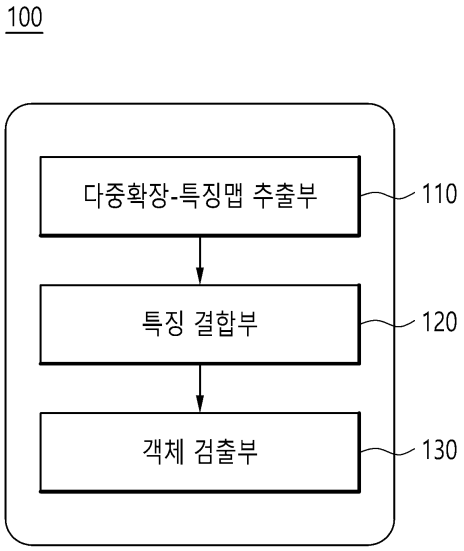




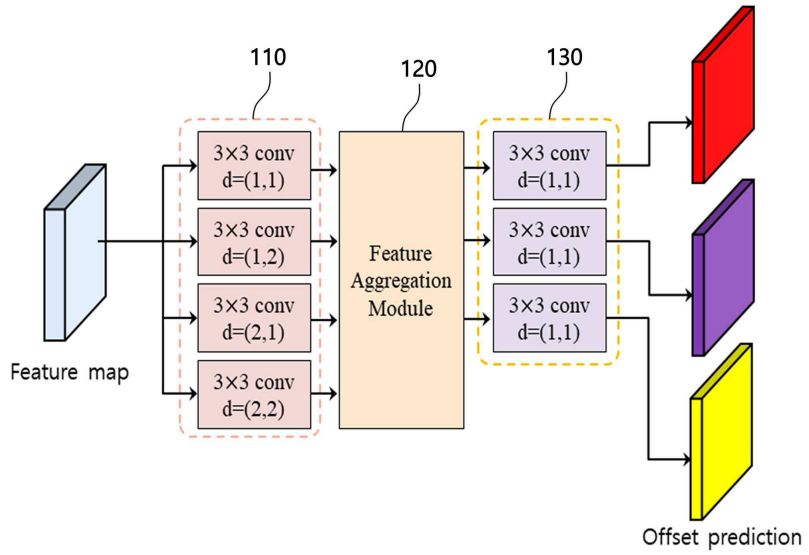
도면2


Detection head
3×3 conv
d=(1,1)
3×3 conv
d=(1,1)
3×3 conv
d=(1,1)
Feature map
Offset prediction

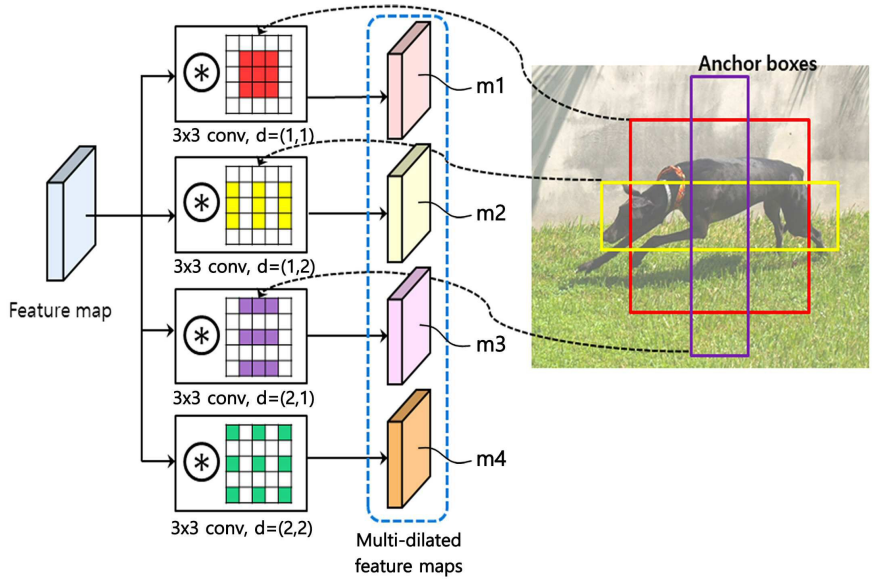
도면3


100
다중확장-특징맵 추출부
110
특징 결합부
120
객체 검출부
130

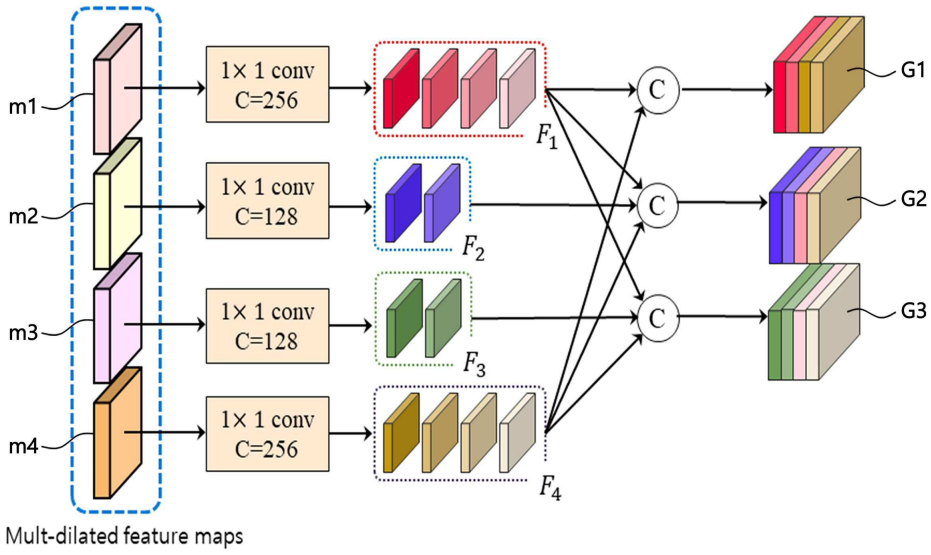
도면4


110
120
130
3×3 conv d=(1,1)
3×3 conv d=(1,2)
3×3 conv d=(2,1)
3×3 conv d=(2,2)
Feature Aggregation Module
3×3 conv d=(1,1)
3×3 conv d=(1,1)
3×3 conv d=(1,1)
Feature map
Offset prediction

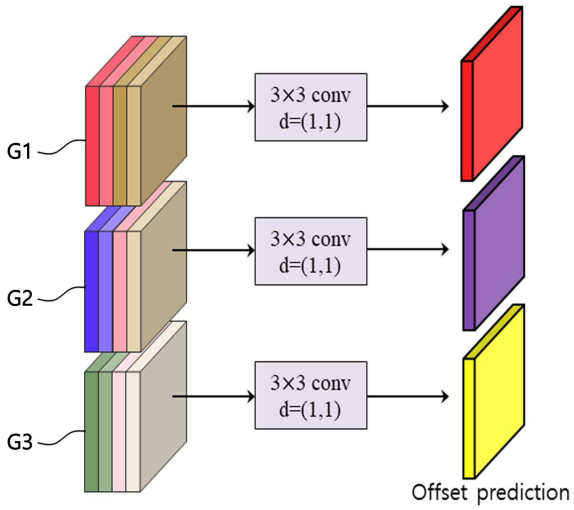
도면5


Anchor boxes
m1
m2
m3
m4
3x3 conv, d=(1,1)
3x3 conv, d=(1,2)
3x3 conv, d=(2,1)
3x3 conv, d=(2,2)
Feature map
Multi-dilated feature maps

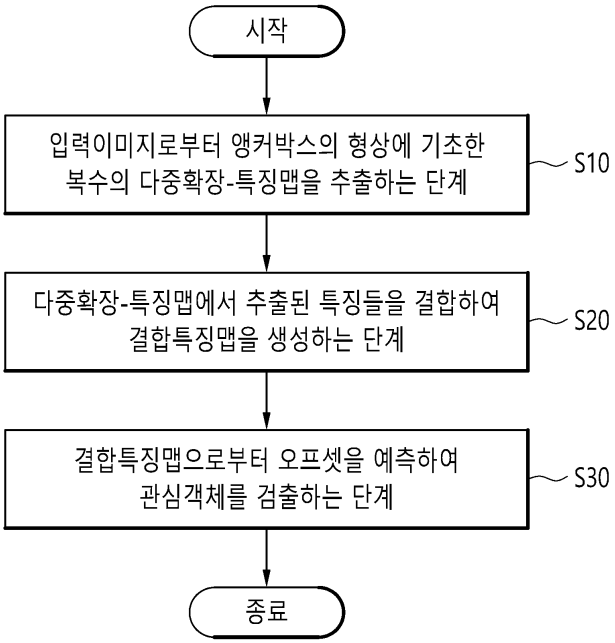
도면6


m1
m2
m3
m4
1× 1 conv C=256
1× 1 conv C=128
1× 1 conv C=128
1× 1 conv C=256
F_1
F_2
F_3
F_4
C
C
C
G1
G2
G3
Mult-dilated feature maps

도면7


G1
G2
G3
3×3 conv d=(1,1)
3×3 conv d=(1,1)
3×3 conv d=(1,1)
Offset prediction

도면8


시작
입력이미지로부터 앵커박스의 형상에 기초한
복수의 다중확장-특징맵을 추출하는 단계
S10
다중확장-특징맵에서 추출된 특징들을 결합하여
결합특징맵을 생성하는 단계
S20
결합특징맵으로부터 오프셋을 예측하여
관심객체를 검출하는 단계
S30
종료

도면9


(a)
aeroplane: 0.98
bus: 0.87
sofa: 0.99
dog: 0.98
sofa: 0.97
person: 0.98
(b)
aeroplane: 1.00
bus: 0.95
bus: 0.85
car: 0.66
sofa: 1.00
sofa: 0.89
sofa: 0.77
person: 0.94