

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)(11) 공개번호 10-2022-0021154
(43) 공개일자 2022년02월22일

(51) 국제특허분류(Int. Cl.)
G06N 3/063 (2006.01) G06N 3/04 (2006.01)
G06N 3/08 (2006.01)
(52) CPC특허분류
G06N 3/063 (2013.01)
G06N 3/04 (2013.01)
(21) 출원번호 10-2020-0101612
(22) 출원일자 2020년08월13일
심사청구일자 2020년08월13일

(71) 출원인
연세대학교 산학협력단
서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)
(72) 발명자
손광훈
서울특별시 서대문구 연세로 50, 연세대학교 제3공학관 C129호(신촌동)
정성훈
서울특별시 서대문구 연세로 50, 연세대학교 제3공학관 C129호(신촌동)
(74) 대리인
민영준

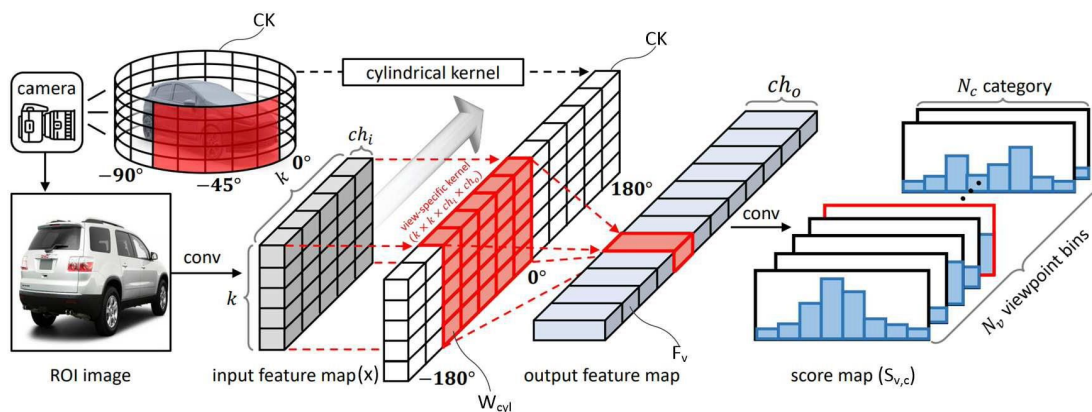
전체 청구항 수 : 총 20 항

(54) 발명의 명칭 원통형 컨볼루션 네트워크 연산 장치와 이를 이용한 객체 인식 및 시점 추정장치 및 방법

(57) 요약

본 발명은 기지정된 각도 단위 간격에 대응하는 다수의 가중치 열(column)이 360도 단위로 주기적으로 반복되는 구조를 갖는 원통형 커널을 포함하고, 기지정된 크기의 입력 이미지가 인가되면, 상기 입력 이미지와 상기 원통형 커널의 다수의 가중치 열 중 기지정된 크기의 각도 범위에 대응하는 개수의 가중치 열을 컨볼루션 연산을 수행하여, 각 각도 단위 간격에 따른 다수의 시점별 특징 벡터를 추출하는 원통형 컨볼루션 네트워크 연산 장치를 이용하여 이미지의 객체 인식 및 시점을 판별하도록 함으로써, 객체를 정확하게 식별하여 인식할 수 있을 뿐만 아니라 객체가 촬영된 시점을 정확하게 판별할 수 있는 원통형 컨볼루션 네트워크와 이를 이용한 객체 인식 및 시점 추정장치 및 방법을 제공할 수 있다.

대표도



(52) CPC특허분류

G06N 3/08 (2013.01)

G06V 10/32 (2022.01)

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

기 지정된 각도 단위 간격에 대응하는 다수의 가중치 열(column)이 360도 단위로 주기적으로 반복되는 구조를 갖는 원통형 커널을 포함하고,

기 지정된 크기의 입력 이미지가 인가되면, 상기 입력 이미지와 상기 원통형 커널의 다수의 가중치 열 중 기 지정된 크기의 각도 범위에 대응하는 개수의 가중치 열을 컨볼루션 연산을 수행하여, 각 각도 단위 간격에 따른 다수의 시점별 특징 벡터를 추출하는 원통형 컨볼루션 네트워크 연산 장치.

청구항 2

제1항에 있어서, 상기 원통형 커널은

미리 지정된 객체의 여러 시점별 특징이 추출되도록, 학습 시에 상기 객체에 대해 기 지정된 각도 단위 간격 각각 대응하는 시점에서 촬영된 다수의 이미지를 이용하여 상기 다수의 가중치 열의 가중치가 획득된 원통형 컨볼루션 네트워크 연산 장치.

청구항 3

제1항에 있어서, 상기 원통형 커널은

대응하는 각도에 따라 360도 주기로 반복되는 구조의 상기 다수의 가중치 열 중 기 지정된 각도에 대응하는 가중치 열이 단절되어 펼쳐진 2차원 구조의 가중치 어레이로 저장되는 원통형 컨볼루션 네트워크 연산 장치.

청구항 4

제3항에 있어서, 상기 원통형 커널은

각 각도에 대응하는 가중치 열을 기준으로 인접한 기 지정된 개수의 가중치 열이 결합된 다수의 가중치 어레이로 저장되는 원통형 컨볼루션 네트워크 연산 장치.

청구항 5

제3항에 있어서, 상기 원통형 컨볼루션 네트워크 연산 장치는

상기 원통형 커널에서 기 지정된 기준 각도에 대응하는 가중치 열로부터 상기 입력 이미지를 적어도 일측 방향으로 열 단위로 이동시키면서 컨볼루션 연산을 수행하며, 상기 입력 이미지가 2차원 구조로 펼쳐진 원통형 커널의 일측단으로 이동되면, 360도 주기로 반복되는 구조의 가중치 어레이의 타측단과 컨볼루션 연산을 수행하는 원통형 컨볼루션 네트워크 연산 장치.

청구항 6

제3항에 있어서, 상기 가중치 어레이는

360도 주기로 반복되는 구조에 따라 양측단에 서로 상대 측단의 인접한 기 지정된 개수의 가중치 열이 추가로 배치되는 원통형 컨볼루션 네트워크 연산 장치.

청구항 7

적어도 하나의 객체가 포함된 입력 이미지를 인가받아, 상기 입력 이미지 내에서 적어도 하나의 객체 각각의 영역인 객체 영역 특징맵을 추출하는 전처리부;

기 지정된 각도 단위 간격에 대응하는 다수의 가중치 열이 360도 단위로 주기적으로 반복되는 구조를 갖는 원통형 커널을 포함하여, 상기 객체 영역 특징맵과 상기 원통형 커널의 다수의 가중치 열 중 기 지정된 크기의 각도 범위에 대응하는 개수의 가중치 열을 컨볼루션 연산을 수행하여, 각 각도 단위 간격에 따른 다수의 시점별 특징

벡터를 추출하는 원통형 커널부;

다수의 객체 클래스 및 다수의 객체별 시점의 패턴이 학습된 인공 신경망을 이용하여 상기 다수의 시점별 특징 벡터 각각에 대한 객체 클래스별 신뢰도 스코어를 추출하여 다수의 스코어 맵을 획득하는 스코어 맵 획득부; 및 상기 다수의 시점별 특징 벡터 각각에 대응하여 시점별로 획득된 상기 다수의 스코어 맵을 분석하여 객체 영역에 포함된 객체의 클래스와 시점을 판별하는 객체 및 시점 판별부를 포함하는 객체 인식 및 시점 추정장치.

청구항 8

제7항에 있어서, 상기 원통형 커널은

대응하는 각도에 따라 360도 주기로 반복되는 구조의 상기 다수의 가중치 열 중 기지정된 각도에 대응하는 가중치 열이 단절되어 펼쳐진 2차원 구조의 가중치 어레이로 저장되는 객체 인식 및 시점 추정장치.

청구항 9

제7항에 있어서, 상기 전처리부는

패턴 추정 방식이 미리 학습된 인공 신경망으로 구현되어 상기 입력 이미지의 특징을 추출하여 특징맵을 획득하는 특징맵 추출부;

패턴 추정 방식이 미리 학습된 인공 신경망으로 구현되어 상기 특징맵에서 객체를 검출하고, 검출된 객체 각각이 포함된 영역을 경계 박스 형태로 추출하여 객체 위치 정보를 획득하는 객체 검출부; 및

검출된 객체 각각에 대한 객체 위치 정보를 인가받아 상기 특징맵에서 각 객체별 영역을 기지정된 크기로 정규화하는 관심 특징 추출부를 포함하는 객체 인식 및 시점 추정장치.

청구항 10

제9항에 있어서, 상기 객체 및 시점 판별부는

상기 다수의 스코어 맵에 대해 소프트 맥스 연산을 수행하여 상기 다수의 스코어 맵을 정규화함으로써, 각 클래스 및 시점별 시점 확률 분포를 획득하고, 상기 다수의 스코어 맵 각각에 대응하는 시점 확률 분포를 가중 누적하여 최종 클래스 스코어를 획득하여 객체를 인식하는 객체 인식 및 시점 추정장치.

청구항 11

제10항에 있어서, 상기 객체 및 시점 판별부는

상기 최종 클래스 스코어(S_c)를 수학식

$$S_c = \sum_{v=1}^{N_v} S_{v,c} \cdot P_{v,c}$$

(여기서 N_v 는 원통형 커널의 시점 개수로 가중치 열 개수이며, $S_{v,c}$ 는 스코어 맵이며, $P_{v,c}$ 는 클래스 및 시점별 시점 확률 분포로서 소프트 맥스 함수(σ)를 이용하여 ($P_{v,c} = \sigma(S_{v,c})$)로 획득된다.)

에 따라 획득하고,

상기 시점 확률 분포와 함께 상기 원통형 커널의 다수의 가중치 열 각각에 지정된 시점 각도(i_v)에 시누소이달(sinusoidal) 함수를 적용하여 객체의 시점(θ_c)을 수학식

$$\theta_c = \text{atan2} \left(\sum_{v=1}^{N_v} P_{v,c} \sin(i_v), \sum_{v=1}^{N_v} P_{v,c} \cos(i_v) \right)$$

에 따라 판별하는 객체 인식 및 시점 추정장치.

청구항 12

제9항에 있어서, 상기 객체 인식 및 시점 추정장치는

패턴 추정 방식이 미리 학습된 인공 신경망을 이용하여 판별된 객체 클래스와 시점에 대응하여 상기 전처리부에서 획득된 객체 위치 정보를 조절하여 경계 박스를 회귀시키는 객체 위치 회귀부를 더 포함하는 객체 인식 및 시점 추정장치.

청구항 13

제12항에 있어서, 상기 객체 인식 및 시점 추정장치는

학습을 수행하기 위한 학습부를 더 포함하고,

상기 학습부는

객체 클래스(c)와 경계 상자 정보(t) 및 시점(θ)이 미리 지정된 이미지를 학습 데이터로 획득하여, 이미지를 상기 입력 이미지로 인가하고, 인가된 입력 이미지로부터 판별된 객체 클래스(\hat{c})와 시점($\hat{\theta}$) 및 경계 상자 정보(\hat{t})를 학습 데이터에 지정된 객체 클래스(c)와 경계 상자 정보(t) 및 시점(θ)과 비교하여 손실을 계산 및 역전파하는 객체 인식 및 시점 추정장치.

청구항 14

적어도 하나의 객체가 포함된 입력 이미지를 인가받아, 상기 입력 이미지 내에서 적어도 하나의 객체 각각의 영역인 객체 영역 특징맵을 추출하는 단계;

기지정된 각도 단위 간격에 대응하는 다수의 가중치 열이 360도 단위로 주기적으로 반복되는 구조를 갖는 원통형 커널을 포함하여, 상기 객체 영역 특징맵과 상기 원통형 커널의 다수의 가중치 열 중 기지정된 크기의 각도 범위에 대응하는 개수의 가중치 열을 컨볼루션 연산을 수행하여, 각 각도 단위 간격에 따른 다수의 시점별 특징 벡터를 추출하는 단계;

다수의 객체 클래스 및 다수의 객체별 시점의 패턴이 학습된 인공 신경망을 이용하여 상기 다수의 시점별 특징 벡터 각각에 대한 객체 클래스별 신뢰도 스코어를 추출하여 다수의 스코어 맵을 획득하는 단계; 및

상기 다수의 시점별 특징 벡터 각각에 대응하여 시점별로 획득된 상기 다수의 스코어 맵을 분석하여 객체 영역에 포함된 객체의 클래스와 시점을 판별하는 단계를 포함하는 객체 인식 및 시점 추정방법.

청구항 15

제14항에 있어서, 상기 원통형 커널은

대응하는 각도에 따라 360도 주기로 반복되는 구조의 상기 다수의 가중치 열 중 기지정된 각도에 대응하는 가중치 열이 단절되어 펼쳐진 2차원 구조의 가중치 어레이로 저장되는 객체 인식 및 시점 추정방법.

청구항 16

제14항에 있어서, 상기 객체 영역 특징맵을 추출하는 단계는

패턴 추정 방식이 미리 학습된 인공 신경망으로 구현되어 상기 입력 이미지의 특징을 추출하여 특징맵을 획득하는 단계;

패턴 추정 방식이 미리 학습된 인공 신경망으로 구현되어 상기 특징맵에서 객체를 검출하고, 검출된 객체 각각이 포함된 영역을 경계 박스 형태로 추출하여 객체 위치 정보를 획득하는 단계; 및

검출된 객체 각각에 대한 객체 위치 정보를 인가받아 상기 특징맵에서 각 객체별 영역을 기지정된 크기로 정규화하는 단계를 포함하는 객체 인식 및 시점 추정방법.

청구항 17

제16항에 있어서, 상기 객체의 클래스와 시점을 판별하는 단계는

상기 다수의 스코어 맵에 대해 소프트 맥스 연산을 수행하여 상기 다수의 스코어 맵을 정규화함으로써, 각 클레

스 및 시점별 시점 확률 분포를 획득하는 단계;

상기 다수의 스코어 맵 각각에 대응하는 시점 확률 분포를 가중 누적하여 최종 클래스 스코어를 획득하여 객체를 인식하는 단계; 및

상기 시점 확률 분포와 함께 상기 원통형 커널의 다수의 가중치 열 각각에 지정된 시점 각도에 시누소이달 (sinusoidal) 함수를 적용하여 객체의 시점을 판별하는 단계를 포함하는 객체 인식 및 시점 추정방법.

청구항 18

제17항에 있어서, 상기 객체를 인식하는 단계는

상기 최종 클래스 스코어(S_c)를 수학식

$$S_c = \sum_{v=1}^{N_v} S_{v,c} \cdot P_{v,c}$$

(여기서 N_v 는 원통형 커널의 시점 개수로 가중치 열 개수이며, $S_{v,c}$ 는 스코어 맵이며, $P_{v,c}$ 는 클래스 및 시점별 시점 확률 분포로서 소프트 맥스 함수(σ)를 이용하여 ($P_{v,c} = \sigma(S_{v,c})$)로 획득된다.)

에 따라 획득하고,

상기 객체의 시점을 판별하는 단계는

상기 객체의 시점(θ_c)을 수학식

$$\theta_c = \text{atan2} \left(\sum_{v=1}^{N_v} P_{v,c} \sin(i_v), \sum_{v=1}^{N_v} P_{v,c} \cos(i_v) \right)$$

(여기서 i_v 는 다수의 가중치 열 각각에 지정된 시점 각도)

에 따라 판별하는 객체 인식 및 시점 추정방법.

청구항 19

제16항에 있어서, 상기 객체 인식 및 시점 추정방법은

패턴 추정 방식이 미리 학습된 인공 신경망을 이용하여 판별된 객체 클래스와 시점에 대응하여 상기 객체 영역 특징맵을 추출하는 단계에서 획득된 객체 위치 정보를 조절하여 경계 박스를 회귀시키는 단계를 더 포함하는 객체 인식 및 시점 추정방법.

청구항 20

제19항에 있어서, 상기 객체 인식 및 시점 추정방법은

학습 단계를 더 포함하고,

상기 학습 단계는

객체 클래스(c)와 경계 상자 정보(t) 및 시점(θ)이 미리 지정된 이미지를 학습 데이터로 획득하여, 이미지를 상기 입력 이미지로 인가하는 단계; 및

인가된 입력 이미지로부터 판별된 객체 클래스(\hat{c})와 시점($\hat{\theta}$) 및 경계 상자 정보(\hat{t})를 학습 데이터에 지정된 객체 클래스(c)와 경계 상자 정보(t) 및 시점(θ)과 비교하여 손실을 계산 및 역전파하는 단계를 포함하는 객체 인식 및 시점 추정방법.

발명의 설명

기술 분야

[0001] 본 발명은 인공 네트워크 연산 장치에 관한 것으로, 원통형 컨볼루션 네트워크 연산장치와 이를 이용한 객체 인식 및 시점 추정장치 및 방법에 관한 것이다.

배경 기술

[0002] 최근 다양한 분야에서 인공 네트워크가 적용되고 있으며, 이 중에서 컨볼루션 네트워크(Convolutional Neural Network: CNN) 이미지 분류(image classification), 시맨틱 세그멘테이션(semantic segmentation), 객체 검출(object detection), 인스턴스 세그멘테이션(instance segmentation)과 같은 시각적 인식 분야에서 우수한 성능을 나타내는 것으로 잘 알려져 있다.

[0003] 컨볼루션 네트워크가 객체의 기하학적 변형에도 객체를 정확하게 분류 또는 인식하기 위해서는 대량의 학습 데이터에 의한 학습이 필요하다. 그러나 컨볼루션 네트워크는 객체에 대해 특정 방향 및 거리에서 획득된 2차원 평면 이미지들을 이용하여 학습이 수행되므로, 객체의 스케일 변화나 관측 시점 변화 및 각 객체별 형상 변화와 같은 다양한 조건들로 인해 객체에 대한 분류 또는 인식 성능이 제한적이다.

[0004] 이에 이미지에서의 객체의 다양한 기하학적 변형에 대응하기 위해 여러 기법들이 제안된 바 있다. 일 예로 아핀 변환(affine transformation)과 같은 공간 변환 모델을 이용하거나 컨볼루션 연산에서 샘플링 위치의 오프셋을 학습하여 입력 데이터를 변환하는 기법 등이 있다.

[0005] 그러나 기존의 기법에서는 2D 공간 상에서 기하학적 변형을 처리하므로, 3D 공간 상에서의 다양한 시점에서 획득된 이미지의 객체를 정확하게 분류 또는 인식하기에는 한계가 있다. 또한 경우에 따라서는 객체에 대한 시점이 판별되어야 하는 경우도 있다. 일 예로 최근 자율 주행 차량 등은 전방 카메라뿐만 아니라 후방 카메라 및 측방 카메라 등을 구비하여 차량 주변 전방향에서의 이미지를 획득하고, 획득된 이미지로부터 주변 객체를 인식하도록 한다. 이때 차량과 주변 객체 사이의 상관 관계를 정확하게 판별하기 위해서는 객체가 촬영된 시점을 확인할 수 있어야 한다. 즉 차량 주변에서 동일 거리에 위치한 다른 차량이 인식되더라도 전면이 촬영된 차량의 경우에는 접근 중인 상태로 판단할 수 있는 반면, 후면이 촬영된 차량의 경우에는 멀어지는 상태로 판단할 수 있으므로 더 정확한 상태를 판별할 수 있다. 그러나 상기한 바와 같이 기존의 컨볼루션 네트워크는 객체 인식 성능에 제약이 있을 뿐만 아니라, 객체가 촬영된 시점을 판별하기 어렵다는 한계가 있다.

[0006] 따라서 객체 인식 성능의 향상과 함께 객체에 대한 이미지가 획득된 방위각을 정확하게 추정할 수 있는 새로운 네트워크에 대한 요구가 증가하고 있다.

선행기술문헌

특허문헌

[0007] (특허문헌 0001) 한국 등록 특허 제10-2029852호 (2019.10.01 등록)

발명의 내용

해결하려는 과제

[0008] 본 발명의 목적은 객체 인식 성능을 대폭 향상시킬 수 있는 원통형 컨볼루션 네트워크 연산 장치를 제공하는데 있다.

[0009] 본 발명의 다른 목적은 객체 인식 성능뿐만 아니라 이미지에서 객체가 촬영된 시점을 정확하게 추정할 수 있는 원통형 컨볼루션 네트워크 연산 장치를 제공하는데 있다.

과제의 해결 수단

[0010] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른 원통형 컨볼루션 네트워크 연산 장치는 기지정된 각도 단위 간격에 대응하는 다수의 가중치 열(column)이 360도 단위로 주기적으로 반복되는 구조를 갖는 원통형 커널을 포함하고, 기지정된 크기의 입력 이미지가 인가되면, 상기 입력 이미지와 상기 원통형 커널의 다수의 가중치 열 중 기지정된 크기의 각도 범위에 대응하는 개수의 가중치 열을 컨볼루션 연산을 수행하여, 각 각도 단위 간

격에 따른 다수의 시점별 특징 벡터를 추출할 수 있다.

- [0011] 상기 원통형 커널은 미리 지정된 객체의 여러 시점별 특징이 추출되도록, 학습 시에 상기 객체에 대해 기지정된 각도 단위 간격 각각 대응하는 시점에서 촬영된 다수의 이미지를 이용하여 상기 다수의 가중치 열의 가중치가 획득될 수 있다.
- [0012] 상기 원통형 커널은 대응하는 각도에 따라 360도 주기로 반복되는 구조의 상기 다수의 가중치 열 중 기지정된 각도에 대응하는 가중치 열이 단절되어 펼쳐진 2차원 구조의 가중치 어레이로 저장될 수 있다.
- [0013] 상기 원통형 커널은 각 각도에 대응하는 가중치 열을 기준으로 인접한 기지정된 개수의 가중치 열이 결합된 다수의 가중치 어레이로 저장될 수 있다.
- [0014] 상기 원통형 컨볼루션 네트워크 연산 장치는 상기 원통형 커널에서 기지정된 기준 각도에 대응하는 가중치 열로부터 상기 입력 이미지를 적어도 일측 방향으로 열 단위로 이동시키면서 컨볼루션 연산을 수행하며, 상기 입력 이미지가 2차원 구조로 펼쳐진 원통형 커널의 일측단으로 이동되면, 360도 주기로 반복되는 구조의 가중치 어레이의 타측단과 컨볼루션 연산을 수행할 수 있다.
- [0015] 상기 가중치 어레이는 360도 주기로 반복되는 구조에 따라 양측단에 서로 상대 측단의 인접한 기지정된 개수의 가중치 열이 추가로 배치될 수 있다.
- [0016] 상기 목적을 달성하기 위한 본 발명의 다른 실시예에 따른 객체 인식 및 시점 추정 장치는 적어도 하나의 객체가 포함된 입력 이미지를 인가받아, 상기 입력 이미지 내에서 적어도 하나의 객체 각각의 영역인 객체 영역 특징맵을 추출하는 전처리부; 기지정된 각도 단위 간격에 대응하는 다수의 가중치 열이 360도 단위로 주기적으로 반복되는 구조를 갖는 원통형 커널을 포함하여, 상기 객체 영역 특징맵과 상기 원통형 커널의 다수의 가중치 열 중 기지정된 크기의 각도 범위에 대응하는 개수의 가중치 열을 컨볼루션 연산을 수행하여, 각 각도 단위 간격에 따른 다수의 시점별 특징 벡터를 추출하는 원통형 커널부; 다수의 객체 클래스 및 다수의 객체별 시점의 패턴이 학습된 인공 신경망을 이용하여 상기 다수의 시점별 특징 벡터 각각에 대한 객체 클래스별 신뢰도 스코어를 추출하여 다수의 스코어 맵을 획득하는 스코어 맵 획득부; 및 상기 다수의 시점별 특징 벡터 각각에 대응하여 시점별로 획득된 상기 다수의 스코어 맵을 분석하여 객체 영역에 포함된 객체의 클래스와 시점을 판별하는 객체 및 시점 판별부를 포함한다.
- [0017] 상기 목적을 달성하기 위한 본 발명의 다른 실시예에 따른 객체 인식 및 시점 추정 방법은 적어도 하나의 객체가 포함된 입력 이미지를 인가받아, 상기 입력 이미지 내에서 적어도 하나의 객체 각각의 영역인 객체 영역 특징맵을 추출하는 단계; 기지정된 각도 단위 간격에 대응하는 다수의 가중치 열이 360도 단위로 주기적으로 반복되는 구조를 갖는 원통형 커널을 포함하여, 상기 객체 영역 특징맵과 상기 원통형 커널의 다수의 가중치 열 중 기지정된 크기의 각도 범위에 대응하는 개수의 가중치 열을 컨볼루션 연산을 수행하여, 각 각도 단위 간격에 따른 다수의 시점별 특징 벡터를 추출하는 단계; 다수의 객체 클래스 및 다수의 객체별 시점의 패턴이 학습된 인공 신경망을 이용하여 상기 다수의 시점별 특징 벡터 각각에 대한 객체 클래스별 신뢰도 스코어를 추출하여 다수의 스코어 맵을 획득하는 단계; 및 상기 다수의 시점별 특징 벡터 각각에 대응하여 시점별로 획득된 상기 다수의 스코어 맵을 분석하여 객체 영역에 포함된 객체의 클래스와 시점을 판별하는 단계를 포함한다.

발명의 효과

- [0018] 따라서, 본 발명의 실시예에 따른 원통형 컨볼루션 네트워크 연산 장치와 이를 이용한 객체 인식 및 시점 추정 장치 및 방법은 기지정된 각도 단위로 전방향에서 획득된 객체 이미지에 따라 원통형 구조로 구성된 컨볼루션 커널을 포함하여 다수의 시점별 특징 벡터를 추출하고, 다수의 시점별 특징 벡터를 기반으로 스코어 맵을 생성함으로써, 객체 클래스와 객체가 촬영된 시점을 정확하게 판별할 수 있다. 또한 스케일 변화에도 강건하며 이미지에서 객체 영역을 정확하게 설정할 수 있다.

도면의 간단한 설명

- [0019] 도 1은 본 발명의 일 실시예에 따른 원통형 컨볼루션 네트워크 연산 장치를 설명하기 위한 도면이다.
- 도 2는 도 1의 원통형 컨볼루션 네트워크가 적용된 객체 인식 및 시점 추정장치의 개략적 구성을 나타낸다.
- 도 3은 원통형 컨볼루션 네트워크에 기반하여 객체 및 시점을 추정하는 개념을 설명하기 위한 도면이다.
- 도 4는 본 발명의 일 실시예에 따른 객체 인식 및 시점 추정방법을 나타낸다.

도 5 및 도 6은 본 실시예에 따른 객체 인식 및 시점 추정방법을 이용하여 객체 인식 및 시점 추정 결과의 일 예를 나타낸다.

발명을 실시하기 위한 구체적인 내용

- [0020] 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.
- [0021] 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.
- [0022] 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.
- [0023] 도 1은 본 발명의 일 실시예에 따른 원통형 컨볼루션 네트워크 연산 장치를 설명하기 위한 도면이다.
- [0024] 본 실시예에 따른 원통형 컨볼루션 네트워크 연산 장치는 도 1에 나타난 바와 같이, 원통형 커널(CK)을 포함한다. 기존의 컨볼루션 네트워크에서는 커널이 2차원 평면 형태로 구성되어 있는 반면, 본 실시예의 원통형 컨볼루션 네트워크에서는 커널이 2차원 평면이 원통형으로 말려있는 형태로 구성된다. 따라서 기존의 컨볼루션 커널의 경우 2차원 평면 형태로 구성됨에 따라 양측단이 단절된 형태로 구성되지만, 원통형 커널(CK)에서는 360도의 주기로 반복되어 로테이션되는 평면 구조로 고려될 수 있다.
- [0025] 그리고 원통형 커널(CK)에서 특정 열(column)이 기준 각도(0°)로 설정되는 경우, 기준 각도(0°)는 동시에 1회 로테이션된 각도(360°)로 볼 수도 있다. 즉 원통형 커널(CK)에서는 양측단의 개념이 존재하지 않고 360° 로테이션되는 구조를 갖게 된다. 여기서는 원통형 커널(CK)의 기준 각도(0°)를 중심으로 일측 방향(예를 들면 시계 방향)이 양의 각도 방향으로 가정하고, 타측 방향(예를 들면 반시계 방향)이 음의 각도 방향 가정하여, 원통형 커널(CK)의 원소, 즉 커널 가중치의 위치를 표시하는 것으로 설정한다. 이 경우, 원통형 커널(CK)에서 기준 각도(0°)에 대향하는 열의 각도는 180° 또는 -180° 가 되며, 180° 와 -180° 는 원통형 커널(CK)에서 동일한 원소 열, 즉 커널 가중치 열을 나타내게 된다. 이는 원통형 커널(CK)이 360도의 주기로 반복되는 로테이션되 로테이션 가능함을 나타낸다.
- [0026] 그리고 원통형 커널(CK)의 각 원소인 커널 가중치는 원통형 커널(CK) 내에 특정 객체가 위치한 경우, 열 단위로 각 커널 가중치의 위치에서 대응하는 시점에서의 객체 특징을 추출할 수 있도록 미리 학습으로 획득될 수 있다.
- [0027] 원통형 커널(CK)의 학습은 지정된 객체에 대해 미리 지정된 각도 단위(여기서는 일 예로 15°)에서 획득된 다수의 객체 이미지를 학습 데이터로 이용하여 수행될 수 있다. 즉 특정 객체에 대해 전방위 시점에서 각 시점별로 획득된 객체 이미지를 학습 데이터로 이용하여 원통형 커널(CK)의 커널 가중치가 열 단위로 획득될 수 있다.
- [0028] 다만, 특정 시점에서 객체를 촬영하여 획득되더라도 객체 이미지는 일정 각도 범위의 객체 이미지로 획득된다. 따라서 비록 지정된 각도에서 획득된 다수의 객체 이미지일지라도 해당 각도뿐만 아니라 인접한 일정 범위의 각도 영역에 대응하는 다른 열의 가중치를 업데이트하기 위해 함께 이용될 수 있다. 따라서 학습 데이터가 전 방위의 모든 시점에서의 객체 이미지를 포함하지 않더라도, 원통형 커널(CK)의 가중치 열들은 학습 데이터로부터 학습될 수 있다.
- [0029] 기존의 컨볼루션 네트워크에서는 일반적으로 객체의 특징을 객체에 대한 시점에 대한 구분이 없이 전체적인 특징으로 추출하기 위해 이용되었다. 따라서 여러 다양한 시점에서 획득된 객체 이미지의 공통적 특징을 추출할 수 있도록 학습되어야 했다. 이에 컨볼루션 네트워크의 경우, 컨볼루션 커널의 학습 시에도 다양한 시점에서 획득된 객체 이미지를 이미지 획득 시점에 대한 구분없이 학습에 이용하였으며, 이로 인해 컨볼루션 네트워크가 객체 인식 장치 등에 적용되는 경우, 객체의 시점별 상이한 특징에 의해 객체 인식 장치의 객체 인식 성능에 한계를 가져오는 주요 요인으로 작용한다.
- [0030] 일 예로 도 1에 도시된 바와 같이, 차량에 대한 객체 인식을 수행하는 경우, 전면을 기준으로 45° 시점에서 촬영된 차량 이미지와 -135° 시점에서 촬영된 차량 이미지의 경우, 매우 상이한 패턴을 갖게 된다. 그럼에도 불

구하고, 기존의 컨볼루션 네트워크에서는 촬영 시점에 대한 구분없이 모두 동일한 차량이라는 특징을 추출하도록 컨볼루션 커널이 학습된다. 그러므로 45° 시점과 -135° 시점에서 촬영된 차량 이미지에서 공통되는 특징을 추출하도록 가중치가 설정된다. 이는 컨볼루션 네트워크가 차량의 특징을 정확하게 추출하기 어렵게 하는 요인이 된다. 뿐만 아니라 인가된 이미지에 포함된 차량이 촬영된 시점을 판별할 수 없다는 한계가 있다.

- [0031] 그에 반해, 본 실시예에 따른 원통형 커널(CK)의 경우, 객체에 대해 기지정된 시점별로 촬영된 학습 이미지를 기반으로 학습되어, 각 위치별 가중치가 객체의 대응하는 시점에서의 특징을 추출할 수 있다. 즉 시점을 특정한 시점별 특징을 추출할 수 있으며, 도 1의 경우, 45° 시점과 -135° 시점에서의 차량 특징을 구분하여 추출할 수 있다. 따라서 객체 인식 장치가 원통형 컨볼루션 네트워크에 기반한 연산 장치를 이용하는 경우, 이미지에 포함된 객체를 매우 정확하게 식별하여 인식할 수 있을 뿐만 아니라 이미지가 촬영된 시점을 정확하게 판별할 수 있다는 장점이 있다.
- [0032] 도 2는 도 1의 원통형 컨볼루션 네트워크가 적용된 객체 인식 및 시점 추정장치의 개략적 구성을 나타내고, 도 3은 원통형 컨볼루션 네트워크에 기반하여 객체 및 시점을 추정하는 개념을 설명하기 위한 도면이다.
- [0033] 도 2를 참조하면, 본 실시예에 따른 객체 인식 및 시점 추정장치는 이미지 획득부(110), 전처리부(120), 원통형 커널부(130), 스코어 맵 획득부(140), 객체 및 시점 판별부(150) 및 객체 위치 회귀부(160)를 포함할 수 있다.
- [0034] 이미지 획득부(110)는 적어도 하나의 객체가 포함된 이미지를 획득한다. 이미지 획득부(110)는 식별되어야 하는 적어도 하나의 객체가 포함된 이미지를 획득하기 위한 구성으로, 객체를 촬영하는 카메라 등의 이미지 획득 수단으로 구현되거나, 이전 획득된 이미지가 저장된 저장 수단 또는 외부 장치가 획득한 이미지를 인가받는 통신 수단 등으로 구현될 수 있다.
- [0035] 전처리부(120)는 이미지 획득부(110)에서 획득된 이미지로부터 객체에 해당하는 관심 영역(Region of Interest: ROI)인 객체 영역을 추출한다. 전처리부(120)는 기지정된 방식에 따라 이미지에서 객체 영역을 추출할 수도 있으나, 여기서는 전처리부(120)가 인공 신경망으로 구현되어 미리 학습된 패턴 추정 방식에 따라 객체 영역에 대한 특징맵을 추출하는 것으로 가정한다.
- [0036] 전처리부(120)는 특징맵 추출부(121), 객체 검출부(122) 및 관심 특징 추출부(123)를 포함할 수 있다. 특징맵 추출부(121)는 이미지 획득부(110)에서 획득된 이미지를 인가받고, 미리 학습된 패턴 추정 방식에 따라 인가된 이미지의 특징을 추출하여 특징맵을 획득한다. 특징맵 추출부(121)는 일 예로 컨볼루션 네트워크(CNN)로 구현될 수 있으며, 인가된 이미지로부터 $H \times W$ 의 크기로 ch_i 채널 수를 갖는 특징맵을 획득할 수 있다.
- [0037] 객체 검출부(122)는 특징맵 추출부(121)에서 획득된 특징맵을 인가받고, 미리 학습된 패턴 추정 방식에 따라 특징맵에서 적어도 하나의 객체 후보를 검출한다. 객체 검출부(122)는 일 예로 영역 제안 네트워크(Region Proposal Network: 이하 RPN)로 구현될 수 있다. RPN으로 구현되는 객체 검출부(122)는 인된 특징맵의 패턴으로부터 이미지에 포함된 객체 후보 개수(N)와 경계 박스(Bounding Box) 형태로 설정되는 객체 후보별 객체 위치 정보(t)를 추출할 수 있다. 여기서 객체 위치 정보(t)는 경계 박스의 4개의 모서리 위치에 대한 좌표값의 집합으로 구성될 수 있다.
- [0038] 관심 특징 추출부(123)는 특징맵 추출부(121)에서 획득된 특징맵을 인가받고, 객체 검출부(122)에서 추정된 객체 후보 개수(N) 및 객체 위치 정보(t)를 인가받아, 특징맵에서 각 객체 후보에 대한 관심 영역이 기지정된 크기(여기서는 일 예로 $k \times k$)가 되도록 기지정된 방식으로 관심 영역 풀링(ROI Pooling)하여 객체 영역 특징맵(x)을 추출한다. 여기서 관심 영역 풀링은 특징맵에서 각 객체에 대한 관심 영역의 크기를 기지정된 크기로 정규화함으로써, 후술하는 원통형 커널부(130)에서 객체에 대한 특징을 용이하게 추출할 수 있도록 한다. 관심 영역 풀링은 다양한 방식으로 수행될 수 있으며, 일 예로 샘플링 기법에 따라 수행될 수 있다.
- [0039] 이미지를 인가받아 이미지에 포함된 적어도 하나의 객체 각각 대한 객체 영역을 추출하는 전처리부(120)의 구성 및 각 구성의 동작은 기존의 객체 검출 장치에서도 적용되는 방식이므로 여기서는 상세한 설명을 생략한다.
- [0040] 원통형 커널부(130)는 미리 학습되어 커널 가중치가 적용된 원통형 커널(CK)을 포함하는 원통형 컨볼루션 네트워크로 구성된다. 원통형 커널부(130)는 전처리부(120)가 이미지로부터 추출한 적어도 하나의 객체에 대한 관심 영역 이미지인 객체 영역 특징맵(x)을 인가받고, 인가된 객체 영역 특징맵(x)을 원통형 커널(CK)의 기준 위치로부터 열 단위로 슬라이드 이동하며 컨볼루션 연산하여 출력 특징맵을 획득한다.
- [0041] 원통형 커널부(130)의 원통형 커널(CK)은 원통형 커널(CK)의 높이는 기지정된 크기로 ROI 풀링되어 입력되는 객체 영역 특징맵(x)의 높이(k)에 대응하여 k개의 행을 포함할 수 있다. 그리고 원통형 커널(CK)은 상기한 바와

같이, 기지정된 각도 단위로 객체 주변에서의 시점이 변화하면서 객체를 촬영하여 획득된 다수의 시점별 가중치 열을 포함하여 구성된다. 따라서 원통형 커널(CK)은 높이(k)에 대응하는 행 개수와 시점 개수(viewpoint bins)(N_v)에 대응하는 열 개수를 갖는 크기($k \times N_v$)의 커널로 구성될 수 있다.

[0042] 상기한 바와 같이, 원통형 커널(CK)이 15° 각도 단위로 객체를 촬영한 다수의 학습 데이터로 학습된 경우를 가정하면, 원통형 커널(CK)은 회전 각도 방향으로 1회전하는 동안 $24(360 / 15 = 24)$ 개의 시점 개수(N_v)에 대응하는 가중치 열을 포함하도록 구성될 수 있다.

[0043] 다만, 본 실시예에 따른 원통형 커널(CK)은 도 1에 도시된 바와 같이, 원통형으로 구성되는 것으로 볼 수 있으나, 실제 원통형 컨볼루션 네트워크에서의 연산은 소프트웨어적으로 각종 연산기에 의해 수행되며, 원통형 커널(CK) 또한 메모리에 저장된다. 따라서 원통형 커널(CK)은 개념적으로는 도 1에 도시된 바와 같이, 원통형으로 구성되지만 실제 메모리에는 도 3에 도시된 바와 같이, 원통이 펼쳐진 2차원 평면 형태의 가중치 어레이로 구성되어 저장될 수 있다. 원통형 커널(CK)이 2차원 평면 형태로 펼쳐진 가중치 어레이 형태로 구성되는 경우, 펼쳐진 원통형 커널(CK)의 양측단은 서로 연결되는 것으로 볼 수 있다. 즉 원통형 커널(CK)은 360° 주기로 반복되도록 구성된 가중치 어레이로 볼 수 있다.

[0044] 따라서 $k \times k$ 크기의 객체 영역 특징맵(x)을 원통형 커널(CK)을 따라 이동하며 컨볼루션 연산을 수행하는 경우, 펼쳐진 원통형 커널(CK)의 양측단에서는 $k \times k$ 크기의 객체 영역 특징맵(x)의 일부에 대응하는 가중치가 존재하지 않는다. 그러나 원통형 커널(CK)의 개념에 따라 타단의 가중치를 이용하여 컨볼루션 연산을 수행할 수 있다. 일 예로 도 3에 도시된 바와 같이, 5×5 크기의 객체 영역 특징맵(x)을 슬라이드하면서 펼쳐진 원통형 커널(CK)과 컨볼루션을 수행하는 경우, 객체 영역 특징맵(x)이 원통형 커널(CK)의 180° 방향의 일측단으로 슬라이딩되면, 원통형 커널(CK)에는 5×5 크기의 객체 영역 특징맵(x)의 일측 2열과 연산되어야 하는 가중치가 존재하지 않는다. 이 경우, 본 실시예의 원통형 커널부(130)는 객체 영역 특징맵(x)의 일측 2열을 원통형 커널(CK)에서 -180° 방향의 타측단의 2열에 배치된 가중치와 연산하여 컨볼루션 연산을 수행할 수 있다. 즉 180° 일측단 이후, 165° 와 150° 에 해당하는 가중치를 객체 영역 특징맵(x)의 일측 2열과 연산할 수 있다.

[0045] 또한 여기서는 원통형 커널(CK)이 단일 커널인 것으로 도시하였으나, 원통형 커널(CK)은 각각 대응하는 각도가 지정되고 객체 영역 특징맵(x)의 열 개수에 대응하는 열 개수를 갖는 다수의 가중치 어레이로 구성될 수도 있다. 일 예로 원통형 커널(CK)의 다수의 가중치 열이 서로 15° 각도 단위 간격을 나타내는 경우, 다수의 가중치 어레이 각각은 5개의 열로 구성되는 객체 영역 특징맵(x)에 대응하여 0° ~ 75° 범위의 각도에 대응하는 가중치 어레이와 15~ 90위의 각도에 대응하는 가중치 어레이 등과 같이 각도 범위 단위로 5개의 가중치 열을 포함하는 다수의 가중치 어레이로 구성될 수 있다. 이 경우에도, 원통형 커널(CK)의 특징에 따라 각 가중치 어레이의 가중치 열은 360° 주기로 반복되는 가중치를 갖도록 구성될 수 있다. 이와 같이, 원통형 커널(CK)이 각 각도에 대응하는 다수의 가중치 어레이로 구성되는 경우, 객체 영역 특징맵(x)은 다수의 가중치 어레이 각각과 컨볼루션 연산될 수 있다.

[0046] 그리고 원통형 커널(CK)은 객체 영역 특징맵(x)의 채널 수(ch_i)와 출력되는 출력 특징맵의 채널 수(ch_o)에 대응하는 채널 수($ch_i \times ch_o$)를 가질 수 있다.

[0047] 따라서 원통형 커널(CK)은 객체 영역 특징맵(x)의 높이(k)와 시점 개수(N_v)에 따른 크기($k \times N_v$) 및 채널 수($ch_i \times ch_o$)를 갖도록 구성될 수 있다.

[0048] 다만, 경우에 따라서는 펼쳐진 원통형 커널(CK)의 양측단에 각각 타단측의 가중치를 추가로 배치함으로써, 로테이션되는 구조의 원통형 커널(CK)에 대한 연산을 더욱 단순화할 수 있다. 즉 180° 방향의 일측단에 추가적으로 -180° 방향의 타측단 방향에 배치된 기지정된 개수의 열을 추가 배치하고, -180° 방향의 타측단에 180° 방향의 일측단 방향에 배치된 기지정된 개수의 열을 추가 배치함으로써, 원통형 커널(CK)에 대해 객체 영역 특징맵(x)이 슬라이딩 이동되어 로테이션되는 연산을 단순화할 수 있다.

[0049] 상기한 바와 같이 원통형 커널부(130)에서는 객체 영역 특징맵(x)이 원통형 커널(CK)을 따라 슬라이딩하며 이동되어 컨볼루션 연산이 수행된다. 따라서 원통형 커널부(130)의 원통형 커널(CK)은 특정 시점에 따른 시점별 컨볼루션 커널(W_v)의 조합으로 볼 수 있으며, 원통형 커널부(130)에서 수행되는 연산은 도 3에 도시된 바와 같이, 객체 영역 특징맵(x)과 다수의 시점 개수(N_v) 중 특정 시점 위치(v)에서의 가중치(W_{vyl}) 사이의 컨볼루션 연산으로 볼 수 있다. 따라서 객체 영역 특징맵(x)과 원통형 커널(CK) 사이의 컨볼루션 연산에 따라 획득되는 시점별

특징 벡터(F_v)는 수학식 1과 같이 표현될 수 있다.

수학식 1

$$F_v = \sum_{p \in \mathcal{R}} W_v(p) \cdot x(p) = \sum_{p \in \mathcal{R}} W_{cyl.}(p + o_v) \cdot x(p)$$

[0050]

[0051]

여기서 \mathcal{R} 은 객체 영역 특징맵(x)의 크기에 대응하는 $k \times k$ 크기의 윈도우를 나타내고, p 는 윈도우 \mathcal{R} 내의 위치를 나타내며, o_v 는 각 시점(v)에 대한 원통형 커널 가중치($W_{cyl.}$)의 오프셋을 나타낸다.

[0052]

원통형 커널부(130)는 시점 개수(N_v)에 대응하는 횡수로 이동하며 시점별 특징 벡터(F_v)를 추출하여, 다수의 시점별 특징 벡터(F_v)로 구성되는 출력 특징맵을 획득한다. 상기에서 시점 개수(N_v)가 24개인 것으로 가정하였으므로, 출력 특징맵은 24개의 시점별 특징 벡터(F_v)를 포함하여 구성될 수 있다. 그리고 다수의 시점별 특징 벡터(F_v) 각각은 출력 채널 길이(ch_o)의 벡터로 획득될 수 있다.

[0053]

스코어 맵 획득부(140)는 원통형 커널부(130)로부터 출력 특징맵을 인가받고, 출력 특징맵의 다수의 시점별 특징 벡터(F_v)가 기지정된 다수의 객체 클래스(c) 및 시점(v) 각각에 대해 대응하는 수준을 추정하여 다수의 스코어 맵(score map)($S_{v,c}$)을 생성한다. 스코어 맵 획득부(140)는 객체 클래스(c)별 가중치가 미리 학습된 인공 신경망으로 구현되어, 출력 특징맵의 다수의 시점별 특징 벡터(F_v)에 대해 기지정된 연산을 수행함으로써, 다수의 시점별 특징 벡터(F_v) 각각이 특정 시점(v)의 객체 클래스(c)에 대응하는 수준, 즉 확률을 나타내는 스코어 맵($S_{v,c}$)을 획득한다. 스코어 맵 획득부(140)는 일 예로 다수의 시점별 특징 벡터(F_v)에 대해 컨볼루션 연산을 수행하는 다수의 컨볼루션 커널로 구현될 수도 있으나, 경우에 따라서는 다수의 완전 연결 레이어(Fully connected layer)로 구성될 수도 있다.

[0054]

도 3을 참조하면, 다수의 스코어 맵($S_{v,c}$)은 시점 개수(N_v)에 대응하는 시점별 특징 벡터(F_v)의 개수에 따라 시점 개수(N_v)만큼 획득될 수 있으며, 다수의 스코어 맵($S_{v,c}$) 각각에는 기지정된 객체 클래스(c)의 개수(N_c)에 대응하는 개수의 객체 클래스 확률값이 포함될 수 있다. 도 3에서는 일 예로 객체 클래스(c)의 개수(N_c)가 8개인 것으로 가정하여, 다수의 스코어 맵($S_{v,c}$) 각각에 8개의 객체 클래스 확률값이 바 그래프 형태로 표현되었으나, 객체 클래스(c)의 개수(N_c)는 다양하게 조절될 수 있다.

[0055]

즉 다수의 스코어 맵($S_{v,c}$) 각각은 특정 시점(v)에서 시점별 특징 벡터(F_v)가 기지정된 다수의 객체 클래스(c) 각각에 대응하는 확률을 나타낸다.

[0056]

객체 및 시점 판별부(150)는 스코어 맵 획득부(140)에서 각 시점(v) 별로 획득된 다수의 스코어 맵($S_{v,c}$)으로부터 이미지에 포함된 객체의 클래스(c)와 객체가 촬영된 시점(v)을 판별한다.

[0057]

상기한 바와 같이, 다수의 스코어 맵($S_{v,c}$) 각각이 특정 시점(v)에서 시점별 특징 벡터(F_v)가 기지정된 다수의 객체 클래스(c)에 대응하는 확률을 나타내므로, 객체 및 시점 판별부(150)는 다수의 스코어 맵($S_{v,c}$)에서 최대 확률값을 갖는 객체 클래스(c)와 스코어 맵($S_{v,c}$)을 탐색함으로써 이미지에 포함된 객체에 해당하는 객체 클래스(c)로 객체를 인식하고 스코어 맵($S_{v,c}$)으로부터 시점(v)을 판별할 수 있다.

[0058]

객체 및 시점 판별부(150)는 우선 다수의 스코어 맵($S_{v,c}$)에 대해 소프트 맥스 연산(σ)을 수행하여 스코어 맵($S_{v,c}$)을 정규화함으로써 시점 확률 분포($P_{v,c} = \sigma(S_{v,c})$)를 획득한다.

[0059]

그리고 수학식 2에 따라 각 시점별로 획득된 다수의 스코어 맵($S_{v,c}$)으로부터 최종 클래스 스코어(S_c)를 계산하여 획득한다.

수학식 2

$$S_c = \sum_{v=1}^{N_v} S_{v,c} \cdot P_{v,c}$$

[0060]

[0061]

다만, 다수의 스코어 맵($S_{v,c}$)은 시점 개수(N_v)에 따라 이산된 시점에서의 확률값으로 획득되게 된다. 다수의 스코어 맵($S_{v,c}$)이 이산된 시점에서의 확률값으로 구성되므로, 객체 및 시점 판별부(150)는 기지정된 다수의 시점(v) 사이의 시점에서 촬영된 객체의 시점을 정확하게 예측할 수 없다. 즉 지정된 N_v 개의 시점으로만 판별하게 될 수 있다. 이에 이미지에 포함된 객체의 정확한 시점을 추정하기 위해 객체 및 시점 판별부(150)는 시누소이달 소프트-아규맥스(sinusoidal soft-argmax) 함수를 적용하여 정확한 객체 시점을 추정한다.

[0062]

객체 및 시점 판별부(150)는 도 3에 도시된 바와 같은 각 시점 각도(i_v)에 대한 시누소이달 함수를 적용하여 시점을 추정한다. 여기서는 원통형 커널(CK)이 15° 단위로 N_v 개의 시점을 포함하는 것으로 가정하였으므로, 각 시점 각도(i_v)는 ($\{-180^\circ, -165^\circ, \dots, 0^\circ, 15^\circ, \dots, 180^\circ\}$, 여기서 -180° 와 180° 는 동일)이다. 그리고 수학식 2의 시점 확률 분포($P_{v,c} = \sigma(S_{v,c})$)와 시누소이달 함수를 사용하여 연속 시점 구간에서의 객체의 정확한 시점(θ_c)을 수학식 3에 따라 추정할 수 있다.

수학식 3

$$\theta_c = \text{atan2} \left(\sum_{v=1}^{N_v} P_{v,c} \sin(i_v), \sum_{v=1}^{N_v} P_{v,c} \cos(i_v) \right)$$

[0063]

[0064]

여기서 객체 및 시점 판별부(150)가 시누소이달 함수를 사용하여 다수의 스코어 맵($S_{v,c}$)의 이산된 시점을 연속 시점 구간으로 변환하여 추정하는 것은 객체의 정확한 시점을 추정할 뿐만 아니라 미분 가능한 구조로 변환하여 객체 인식 및 시점 추정장치의 학습이 가능하도록 하기 위함이다.

[0065]

객체 위치 회귀부(160)는 판별된 객체 클래스(c)와 시점(θ_c)에 따라 객체의 더욱 정확한 위치를 추정하기 위해 경계 박스를 회귀시킨다. 객체 위치 회귀부(160) 또한 인공 신경망으로 구현될 수 있으며, 일 예로 컨볼루션 네트워크로 구현될 수 있다. 경계 위치 회귀부(160)는 미리 학습된 패턴 추정 방식에 따라 다수의 시점별 특징 벡터(F_v)로부터 객체 클래스(c) 및 시점(θ_c)에 대응하는 크기의 경계 상자 오프셋($t_{v,c}$)을 생성하여 전처리부(120)의 객체 검출부(122)가 추정한 객체 위치 정보(t)를 인코딩하여 객체 영역을 정확하게 지정하도록 경계 박스를 회귀시킨다.

[0066]

학습부(200)는 객체 인식 및 시점 추정장치의 학습 시에 구비되어, 객체 인식 및 시점 추정장치를 학습시킨다. 객체 인식 및 시점 추정장치의 학습 시 객체 클래스(c)와 경계 상자 정보(t) 및 시점(θ)이 미리 지정된 이미지가 학습 데이터로 이용된다.

[0067]

학습부(200)는 학습 데이터의 이미지를 이미지 획득부(110)로 전달하고, 객체 및 시점 판별부(150)에서 판별된 객체 클래스(\hat{c})와 시점($\hat{\theta}$) 및 객체 위치 회귀부(160)에서 획득된 경계 상자 정보(\hat{t})를 인가받아 손실(L)을 수학식 4에 따라 계산하고, 계산된 손실(L)을 역전파하여 객체 인식 및 시점 추정장치를 학습시킨다.

수학식 4

$$L = L_{\text{cls}}(c, \hat{c}) + [\hat{c} \geq 1] \{ L_{\text{reg}}(t_{v,c}, \hat{t}) + [\hat{\theta} \neq \emptyset] L_{\text{view}}(\theta_c, \hat{\theta}) \}$$

[0068]

- [0069] 수학적 식 4에서 $[\]$ 는 진실이면 1을 거짓이면 0을 출력하는 아이버슨(Iverson)함수이고, 객체 클래스(\hat{C})가 0인 경우, 경계 상자 손실(L_{reg})과 시점 손실(L_{view})은 무시된다. 시점이 포함($\hat{\theta} \neq 0$)된 학습 데이터를 이용하는 경우, 시점 손실(L_{view})이 적용될 수 있으나, 시점이 포함되지 않은 학습 데이터를 이용하는 경우에, 시점 손실(L_{view})은 생략될 수 있다. 그리고 클래스 손실(L_{cls})은 크로스 엔트로피 기법에 따라 계산될 수 있으며, 경계 상자 손실(L_{reg})과 시점 손실(L_{view})은 L1 놈 함수로 계산될 수 있다.
- [0070] 추가적으로 학습부(200)는 좌우 대칭인 객체 클래스를 원통형 커널(CK)에 학습시키는 경우, 원통형 커널(CK)의 가중치(W_{cyl})에 추가적인 제약 조건을 설정할 수 있다. 좌우 대칭인 객체 클래스를 학습시키는 경우, 원통형 커널(CK)의 가중치(W_{cyl})는 전면에 해당하는 기준 각도(0°)에서의 전면 가중치(W_{front})와 후면에 해당하는 각도(180°)에서의 후면 가중치(W_{rear})와 좌 각도 범위 및 우 각도 범위에 해당하는 좌 사이드 가중치(W_{left}) 및 우 사이드 가중치(W_{right})로 구분될 수 있다.
- [0071] 그리고 객체가 좌우 대칭인 경우 좌 사이드 가중치(W_{left})와 우 사이드 가중치(W_{right})는 서로 수평 플립된 가중치를 가져야 하므로, 우 사이드 가중치(W_{right})는 좌 사이드 가중치(W_{left})에 플립 함수(h)를 적용하여 $W_{right} = h(W_{left})$ 로 획득될 수 있다. 즉 학습부(200)는 학습 대상인 객체가 좌우 대칭인 경우, 좌 사이드 가중치(W_{left})와 우 사이드 가중치(W_{right}) 중 하나만 학습으로 획득하고 나머지는 플립 함수를 이용하여 획득할 수 있다. 이는 학습부(200)의 메모리 용량을 크게 줄일 수 있을 뿐만 아니라, 연산량을 줄여 효율적으로 객체 인식 및 시점 추정장치를 학습시킬 수 있다.
- [0072] 도 4는 본 발명의 일 실시예에 따른 객체 인식 및 시점 추정방법을 나타낸다.
- [0073] 도 1 내지 도 3을 참조하여, 도 4의 객체 인식 및 시점 추정방법을 설명하면, 우선 인식되어야 하는 적어도 하나의 객체가 포함된 이미지를 획득한다(S11). 그리고 획득된 이미지에 대해 기지정된 전처리를 수행한다(S12). 전처리 단계에서는 우선 컨볼루션 네트워크를 이용하여 획득된 이미지의 패턴으로부터 특징을 추출하여 특징맵을 생성한다. 그리고 영역 제안 네트워크(RPN)등을 이용하여 특징맵에서 객체 후보 개수(N)와 경계 박스 형태로 설정되는 객체 후보별 객체 위치 정보(t)를 추출한다. 또한 객체 후보별 객체 위치 정보(t) 각각을 인공 신경망을 이용하여 ROI 풀링하여 정규화함으로써 객체 영역 특징맵(x)을 추출한다.
- [0074] 이미지를 전처리하여 이미지에 포함된 각 객체에 대한 객체 영역 특징맵(x)이 추출되면, 객체에 대한 가중치가 시점 각도별로 열 단위로 설정되고 360도 단위로 반복되는 원통형 커널에 객체 영역 특징맵(x)을 슬라이드하면서 컨볼루션 연산을 수행하여 수학적 식 1에 따라 획득되는 다수의 시점별 특징 벡터(F_v)로 구성된 출력 특징맵을 획득한다(S13). 그리고 인공 신경망을 이용하여 다수의 시점별 특징 벡터(F_v) 각각이 각 시점(v)의 다수의 객체 클래스(c)에 대응하는 확률을 나타내는 다수의 스코어 맵($S_{v,c}$)을 획득한다(S14).
- [0075] 다수의 스코어 맵($S_{v,c}$)이 획득되면, 다수의 스코어 맵($S_{v,c}$)에서 대응하는 객체 클래스(c)들에 대한 스코어의 합에 따라 객체 클래스를 판별한다(S15). 이때, 다수의 스코어 맵($S_{v,c}$)에 대해 소프트 맥스 연산(σ)을 수행하여 스코어 맵($S_{v,c}$)을 정규화함으로써 시점 확률 분포($P_{v,c}$)를 획득하고, 획득된 시점 확률 분포($P_{v,c}$)를 계산하고, 계산된 시점 확률 분포($P_{v,c}$)와 스코어 맵($S_{v,c}$)을 이용하여 수학적 식 2에 따라 최종 클래스 스코어(S_c)를 계산하여 객체의 클래스(c)를 판별할 수 있다.
- [0076] 이후 다수의 스코어 맵($S_{v,c}$)과 판별된 객체의 클래스(c)에 따라 이미지에 포함된 객체가 촬영된 시점(θ_c)을 판별한다(S16). 이때, 객체 시점(θ_c)은 시점 확률 분포($P_{v,c}$)와 함께 학습 시에 지정된 시점 각도(i_v)에 시누소이달 함수를 적용하여 수학적 식 3에 따라 계산될 수 있다. 그리고 인공 신경망을 이용하여 판별된 클래스(c)와 객체 시점(θ_c)에 대응하도록 객체 위치 정보(t)의 객체 경계 박스를 조절한다(S17).
- [0077] 이때, 이미지에 포함된 객체가 다수개인 경우, 이미지에 포함된 모든 객체에 대한 클래스(c)와 시점(θ_c) 및 객체 경계 박스가 추출될 때까지 원통형 커널 연산 단계(S13)부터 반복하여 수행할 수 있다.
- [0078] 이미지 내의 특정 객체에 대한 클래스(c)와 시점(θ_c)이 판별되고, 객체 경계 박스가 조절되면, 현재 상태가 학

습 상태인지 판별한다(S18). 만일 학습 상태인 것으로 판별되면, 획득된 클래스(c)와 시점(θ_c) 및 경계 박스를 학습 데이터로 획득된 이미지에 미리 지정된 클래스(c)와 시점(θ_c) 및 경계 박스와 비교하여 손실(L)을 수학적 4에 따라 계산하고, 계산된 손실을 역전파하여 학습을 수행한다(S19). 이때 학습은 기지정된 횟수로 반복되거나 손실(L)이 기지정된 기준 손실 이하가 될 때까지 반복될 수 있다.

[0079] 도 5 및 도 6은 본 실시예에 따른 객체 인식 및 시점 추정방법을 이용하여 객체 인식 및 시점 추정 결과의 일 예를 나타낸다.

[0080] 도 5는 Pascal3D+ 데이터 셋에 대한 객체 인식 및 시점 추정 결과를 나타내고, 도 6은 KITTI 데이터 셋에 대한 객체 인식 및 시점 추정 결과를 나타낸다.

[0081] 도 5에 도시된 바와 같이, 본 실시예에 따른 객체 인식 및 시점 추정방법의 경우, 다양한 시점에서 촬영된 다양한 객체에 대해 정확하게 검출하여 객체에 대한 경계 박스를 설정할 수 있을 뿐만 아니라, 객체가 촬영된 시점을 정확하게 추정할 수 있다. 또한 도 6을 참조하면, 본 실시예에 따른 객체 인식 및 시점 추정방법은 서로 다른 스케일의 다수의 객체가 포함된 이미지에서도 정확하게 객체를 검출하여 경계 박스를 설정할 수 있을 뿐만 아니라, 각 객체에 대한 시점을 정확하게 추정할 수 있음을 알 수 있다.

[0082] 본 발명에 따른 방법은 컴퓨터에서 실행시키기 위한 매체에 저장된 컴퓨터 프로그램으로 구현될 수 있다. 여기서 컴퓨터 판독가능 매체는 컴퓨터에 의해 액세스될 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.

[0083] 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.

[0084] 따라서, 본 발명의 진정한 기술적 보호 범위는 첨부된 청구범위의 기술적 사상에 의해 정해져야 할 것이다.

부호의 설명

[0085] 110: 이미지 획득부 120: 전처리부

121: 특징맵 추출부 122: 객체 검출부

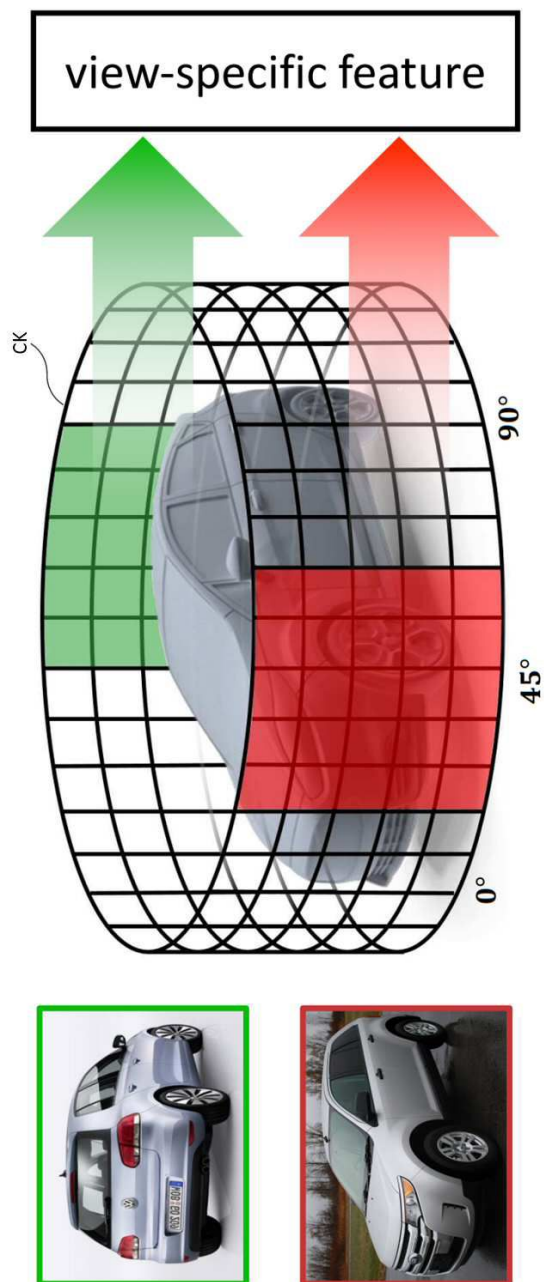
123: 관심 특징 추출부 130: 원통형 커널부

140: 스코어 맵 획득부 150: 객체 및 시점 판별부

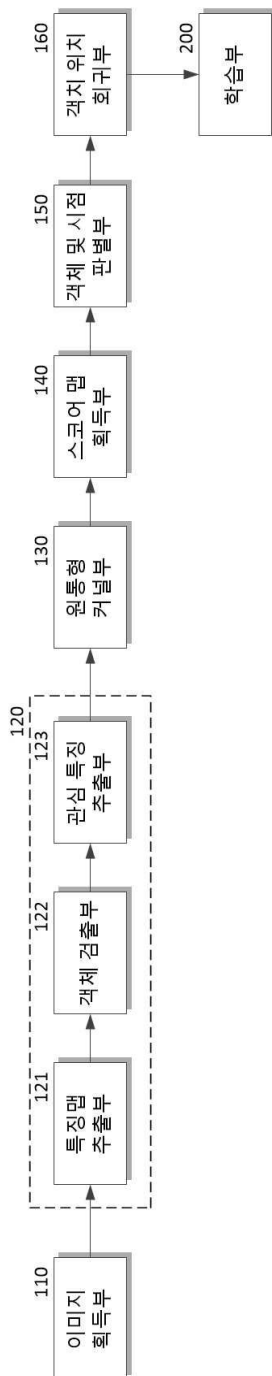
160: 객체 위치 회귀부 200: 학습부

도면

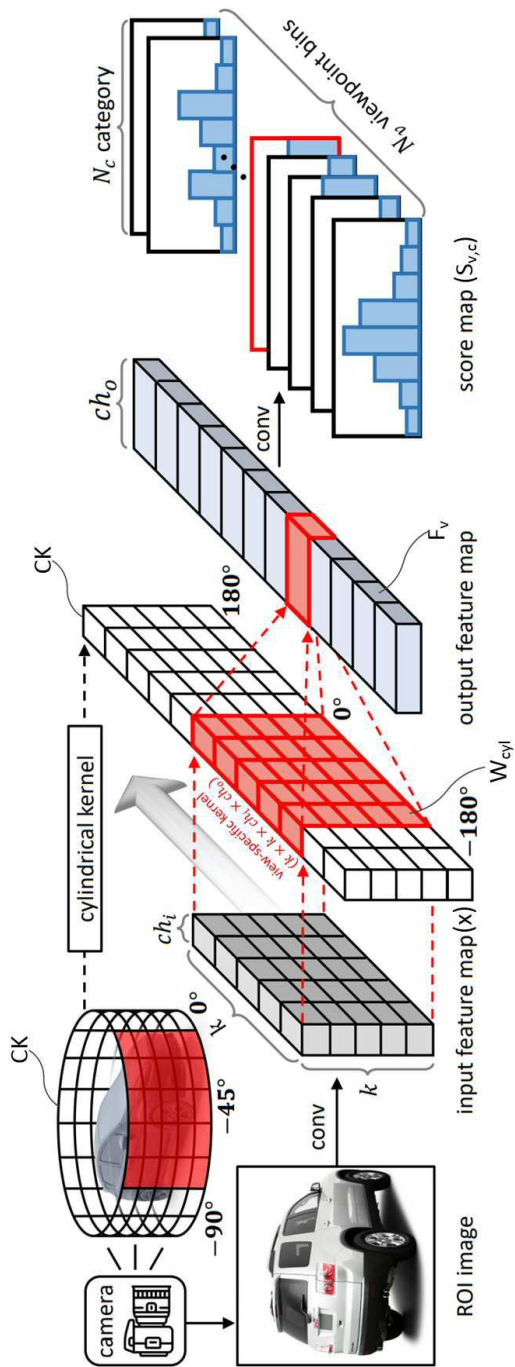
도면1



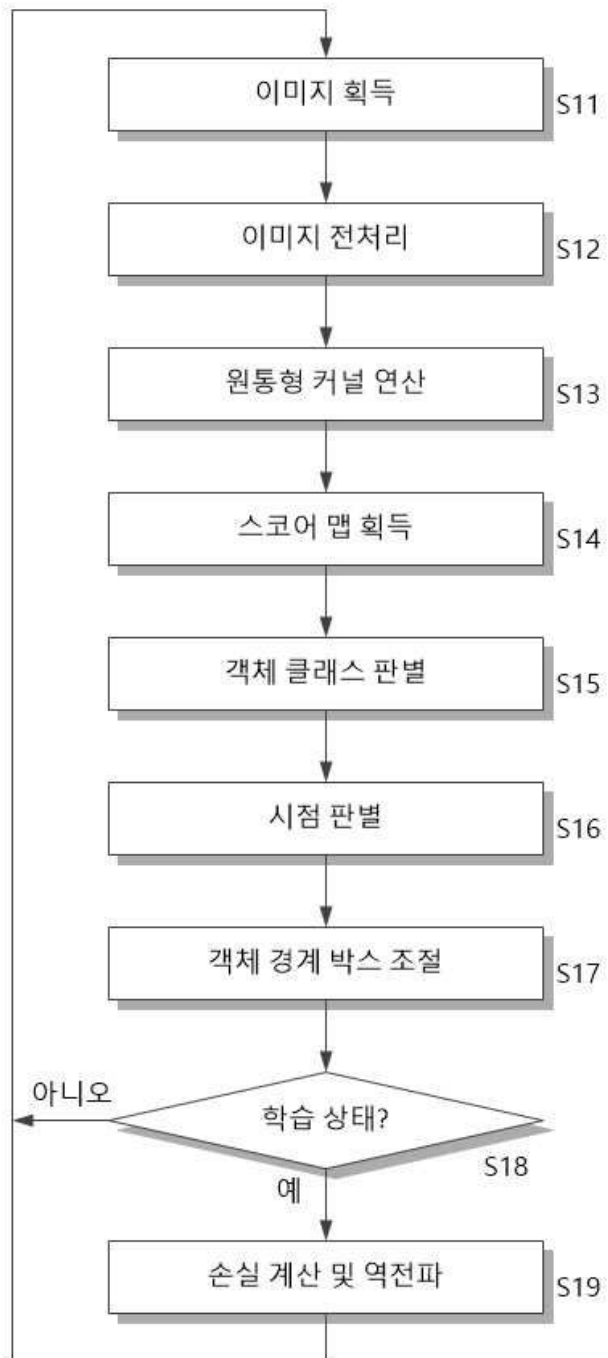
도면2



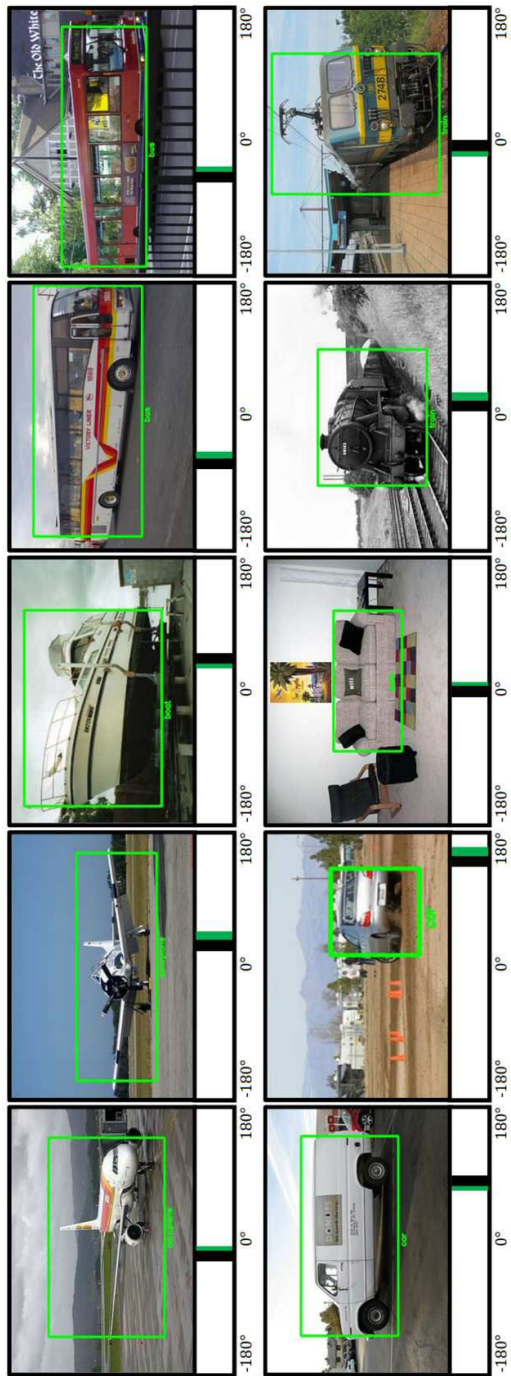
도면3



도면4



도면5



도면6

