(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2021-0103717
(43) 공개일자 2021년08월24일

(51) 국제특허분류(Int. Cl.)
H04N 9/67 (2006.01) *G06N 20/00* (2019.01)
H04N 1/60 (2006.01)
(52) CPC특허분류
H04N 9/67 (2013.01)
G06N 20/00 (2021.08)
(21) 출원번호 10-2020-0018234
(22) 출원일자 2020년02월14일
심사청구일자 2020년02월14일

(71) 출원인
세종대학교산학협력단
서울특별시 광진구 능동로 209 (군자동, 세종대학교)
(72) 발명자
최유경
서울특별시 광진구 능동로 209, 세종대학교 지능기전공학부(군자동)
한대찬
서울특별시 광진구 능동로 209, 세종대학교 지능기전공학부(군자동)
(뒷면에 계속)
(74) 대리인
송인호, 최관락

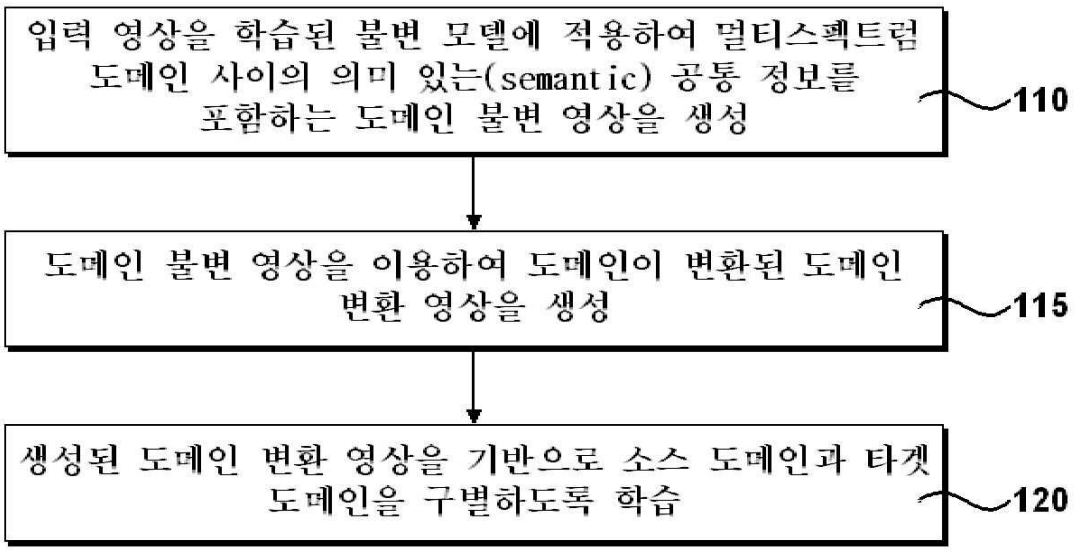
전체 청구항 수 : 총 17 항

(54) 발명의 명칭 **다중 스펙트럼 영상 변환 방법 및 그 장치**

(57) 요 약

다중 스펙트럼 영상 변환 방법 및 그 장치가 개시된다. 입력 영상을 학습된 불변 모델에 적용하여 멀티스펙트럼 도메인 사이의 의미 있는(semantic) 공통 정보를 포함하는 도메인 불변 영상을 생성하는 단계; 및 상기 도메인 불변 영상을 이용하여 도메인이 변환된 도메인 변환 영상을 생성하는 단계를 포함하되, 상기 불변 모델은 도메인 불변 영상을 이용하여 소스 도메인의 영상을 복원하도록 학습될 수 있다.

대 표 도 - 도1



(52) CPC특허분류
H04N 1/6016 (2013.01)
(72) 발명자
황유진
서울특별시 광진구 능동로 209, 세종대학교 지능기전공학부(군자동)
김남일
경기도 용인시 수지구 동천로153번길 7, 1103동 1401호(동천동, 한빛마을 래미안 이스트팰리스 1단지)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711084754
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	개인기초연구(과기정통부)(R&D)
연구과제명	센서 이상 상황에 강인한 멀티 센서 기반 보행자 검출 기법
기 여 율	1/1
과제수행기관명	세종대학교 산학협력단
연구기간	2018.09.01 ~ 2020.08.31

명 세 서

청구범위

청구항 1

입력 영상을 학습된 불변 모델에 적용하여 멀티스펙트럼 도메인 사이의 의미 있는(semantic) 공통 정보를 포함하는 도메인 불변 영상을 생성하는 단계; 및

상기 도메인 불변 영상을 이용하여 도메인이 변환된 도메인 변환 영상을 생성하는 단계를 포함하되,

상기 불변 모델은 도메인 불변 영상을 이용하여 소스 도메인의 영상을 복원하도록 학습되는 것을 특징으로 하는 다중 스펙트럼 영상 변환 방법.

청구항 2

제1 항에 있어서,

상기 불변 모델은,

상기 멀티스펙트럼 도메인의 영상으로부터 공통 정보를 포함하는 특징맵을 추출한 후 평균 및 분산을 이용하여 상기 특징맵을 정규화하고, 활성화함수(ReLu)를 이용하여 음수 변수를 잘라내어 상기 도메인 불변 영상을 생성하는 것을 특징으로 하는 다중 스펙트럼 영상 변환 방법.

청구항 3

제1 항에 있어서,

상기 불변 모델은,

상기 도메인 불변 영상이 상기 멀티스펙트럼 도메인 각각의 영상으로 복구하는데 유용한 시맨틱 정보를 포함하도록 순환 손실을 고려하여 상기 도메인 불변 영상을 생성하는 것을 특징으로 하는 다중 스펙트럼 영상 변환 방법.

청구항 4

제1 항에 있어서,

상기 불변 모델은,

상기 멀티스펙트럼 도메인의 영상에서의 불변 영상 구별을 최소화하기 위해 시맨틱 정보를 유지하도록 픽셀-단위 손실을 고려하여 상기 도메인 불변 영상을 생성하는 것을 특징으로 하는 다중 스펙트럼 영상 변환 방법.

청구항 5

제1 항에 있어서,

상기 불변 모델은,

단일 스케일 구조적 유사성 인덱스(SSIM) 손실을 반영한 일관성 제약을 고려하여 상기 도메인 불변 영상을 생성하는 것을 특징으로 하는 다중 스펙트럼 영상 변환 방법.

청구항 6

제1 항에 있어서,

상기 불변 모델은,

삼중항 손실(triplet loss)을 고려한 중간 특징(intermediate feature)을 이용하여 상기 도메인 불변 영상을 생성하는 것을 특징으로 하는 다중 스펙트럼 영상 변환 방법.

청구항 7

제1 항에 있어서,

상기 불변 모델은,

목적 함수를 만족하도록 상기 멀티스펙트럼 도메인의 영상을 이용하여 학습되되,

상기 목적 함수는 하기 수학식에 의해 계산되는 것을 특징으로 하는 다중 스펙트럼 영상 변환 방법.

$$L(A) = \{L_{adv}(A) + L_{cycle}(A)\} + \{\lambda_0 L_{domain}(A) + \lambda_1 L_{disc}(A) + \lambda_2 L_{divs}(A)\}$$

여기서, A는 도메인을 나타내고, $L_{adv}(A)$는 적대적 손실(adversarial loss) 함수를 나타내며, $L_{cycle}(A)$는 순환 손실 함수를 나타내며, $L_{domain}(A)$는 픽셀-단위 손실 함수를 나타내고, $L_{disc}(A)$는 일관성 제약 함수를 나타내고, $L_{divs}(A)$는 다양성 제약 함수를 나타내며, λ_0, λ_1, λ_2는 각각 가중치를 나타냄.

청구항 8

제7 항에 있어서,

상기 불변 모델은,

자기 복원을 이용한 영상 변환법 기반의 판별 모델과 생성 모델 사이의 미니맥스 게임(minimax game)에 기초한 하기 수학식을 이용하여 상기 목적함수를 최적화하는 것을 특징으로 하는 다중 스펙트럼 영상 변환 방법.

$$\mathbb{M}^* = \min_{\mathbb{M},\mathbb{G}} \max_{\mathbb{D}} L(A)$$

여기서, M은 불변 모델을 나타내고, G는 생성 모델을 나타내며, D는 판별 모델을 나타냄.

청구항 9

제1 항 내지 제8 항 중 어느 하나의 항에 따른 방법을 수행하기 위한 프로그램 코드를 기록한 컴퓨터로 판독 가능한 기록매체 제품.

청구항 10

입력 영상을 학습된 불변 모델에 적용하여 멀티스펙트럼 도메인 사이의 의미 있는(semantic) 공통 정보를 포함하는 도메인 불변 영상을 생성하는 불변 모델;

상기 도메인 불변 영상을 이용하여 도메인이 변환된 도메인 변환 영상을 생성하는 생성 모델; 및

상기 도메인 변환 영상을 이용하여 입력 도메인의 영상을 복구하도록 학습하는 판별 모델을 포함하는 다중 스펙트럼 영상 변환 장치.

청구항 11

제10 항에 있어서,

상기 멀티스펙트럼 도메인의 영상으로부터 공통 정보를 포함하는 특징맵을 추출한 후 평균 및 분산을 이용하여 상기 특징맵을 정규화하고, 활성화함수(ReLu)를 이용하여 음수 변수를 잘라내어 상기 도메인 불변 영상을 생성하는 것을 특징으로 하는 다중 스펙트럼 영상 변환 장치.

청구항 12

제10 항에 있어서,

상기 불변 모델은,

상기 도메인 불변 영상이 상기 멀티스펙트럼 도메인 각각의 영상으로 복구하는데 유용한 시맨틱 정보를 포함하도록 순환 손실을 고려하여 상기 도메인 불변 영상을 생성하는 것을 특징으로 하는 다중 스펙트럼 영상 변환 장치.

청구항 13

제10 항에 있어서,

상기 불변 모델은,

상기 멀티스펙트럼 도메인의 영상에서의 불변 영상 구별을 최소화하기 위해 시맨틱 정보를 유지하도록 픽셀-단위 손실을 고려하여 상기 도메인 불변 영상을 생성하는 것을 특징으로 하는 다중 스펙트럼 영상 변환 장치.

청구항 14

제10 항에 있어서,

상기 불변 모델은,

단일 스케일 구조적 유사성 인덱스(SSIM) 손실을 반영한 일관성 제약을 고려하여 상기 도메인 불변 영상을 생성하는 것을 특징으로 하는 다중 스펙트럼 영상 변환 장치.

청구항 15

제10 항에 있어서,

상기 불변 모델은,

삼중항 손실(triplet loss)을 고려한 중간 특징(intermediate feature)을 이용하여 상기 도메인 불변 영상을 생성하는 것을 특징으로 하는 다중 스펙트럼 영상 변환 장치.

청구항 16

멀티스펙트럼 도메인의 영상을 이용하여 상기 멀티스펙트럼 도메인 사이의 의미 있는 공통 정보를 포함하는 도메인 불변 영상을 생성하도록 불변 모델을 학습하는 인코더; 및

입력 영상을 상기 학습된 불변 모델에 적용하여 도메인 불변 영상을 생성하고, 상기 생성된 도메인 불변 영상을 이용하여 도메인 변환된 도메인 변환 영상을 생성하는 디코더를 포함하는 다중 스펙트럼 영상 변환 장치.

청구항 17

제16 항에 있어서,

상기 인코더 및 상기 디코더는 각각,

멀티스펙트럼 도메인 사이의 의미 있는(semantic) 공통 정보를 포함하는 도메인 불변 영상을 생성하는 불변 모델;

상기 도메인 불변 영상을 이용하여 도메인이 변환된 도메인 변환 영상을 생성하는 생성 모델; 및

상기 도메인 변환 영상을 이용하여 입력 도메인의 영상을 복구하도록 학습하는 판별 모델을 포함하는 것을 특징으로 하는 다중 스펙트럼 영상 변환 장치.

발명의 설명

기 술 분 야

[0001] 본 발명은 다중 스펙트럼 영상 변환 방법 및 그 장치에 관한 것이다.

배 경 기 술

[0003] RGB 영상 센서는 장소 인식, 물체 인식 및 자율 주행과 같은 다양한 잠재적 응용 분야를 통해 지난 수 십년 동안 컴퓨터 비전 및 로봇 공학 분야에서 핵심적인 구성 요소로 자리잡고 있다. 종래의 방법은 대부분 색상 정보에만 초점을 맞추기 때문에 조명 변화 조건에서 작동하지 않을 수 있다. 열 화상 센서가 이에 대한 대안을 제공해준다. 열 화상 센서는 물체의 방사선을 포착하기 때문에 조명 조건의 변화와 상관없이 선명한 물체 및 구조 실루엣을 제공하나 RGB 영상의 시각적 디테일은 손실될 수 있다. 다중 스펙트럼 영상이 이를 본질적으로 보완할 수 있다.

[0004] 최근 다중 스펙트럼 관련 연구가 많은 학문적 관심을 받았으며, RGB 및 열 영상이 포함된 여러 데이터 세트가 컴퓨터 비전 및 로봇 공학 커뮤니티에 도입되었다. 영상 변환에서 모델은 한 도메인을 다른 도메인으로 변환하도록 학습된다. RGB-to-RGB 영상 변환 작업과 달리 다중 스펙트럼(RGB-Thermal) 영상 변환에는 근본적인 문제가 있다. RGB 영상 및 열 영상은 가장자리, 실루엣과 같은 일부 정보를 공유하더라도 영상 촬영 원리로 인해 시각적 세부 사항(detail)은 다르다. 예를 들어, 열 영상은 운전자 차량에서 복사열을 캡쳐할 수 있으나, RGB 영상의 텍스처 및 색상 정보는 캡쳐할 수 없다.

발명의 내용

해결하려는 과제

[0006] 본 발명은 RGB 영상과 열 영상 사이의 도메인 격차를 줄일 수 있는 새로운 다중 스펙트럼 영상 변환 방법 및 그 장치를 제공하기 위한 것이다.

[0007] 또한, 본 발명은 RGB 및 열 도메인 사이의 공통 정보를 포함하는 다중 스펙트럼 불변 영상을 추출할 수 있는 다중 스펙트럼 영상 변환 방법 및 그 장치를 제공하기 위한 것이다.

과제의 해결 수단

[0009] 본 발명의 일 측면에 따르면, RGB 영상과 열 영상 사이의 도메인 격차를 줄일 수 있는 새로운 다중 스펙트럼 영상 변환 방법이 제공된다.

[0010] 본 발명의 일 실시예에 따르면, 입력 영상을 학습된 불변 모델에 적용하여 멀티스펙트럼 도메인 사이의 의미 있는(semantic) 공통 정보를 포함하는 도메인 불변 영상을 생성하는 단계; 및 상기 도메인 불변 영상을 이용하여 도메인이 변환된 도메인 변환 영상을 생성하는 단계를 포함하되, 상기 불변 모델은 도메인 불변 영상을 이용하여 입력 도메인의 영상을 복원하도록 학습되는 것을 특징으로 하는 다중 스펙트럼 영상 변환 방법이 제공될 수 있다.

[0011] 상기 불변 모델은, 상기 멀티스펙트럼 도메인의 영상으로부터 공통 정보를 포함하는 특징맵을 추출한 후 평균 및 분산을 이용하여 상기 특징맵을 정규화하고, 활성화함수(ReLu)를 이용하여 음수 변수를 잘라내어 상기 도메인 불변 영상을 생성할 수 있다.

[0012] 상기 불변 모델은,

[0013] 상기 도메인 불변 영상이 상기 멀티스펙트럼 도메인 각각의 영상으로 복구하는데 유용한 시맨틱 정보를 포함하도록 순환 손실을 고려하여 상기 도메인 불변 영상을 생성할 수 있다.

[0014] 상기 불변 모델은, 상기 멀티스펙트럼 도메인의 영상에서의 불변 영상 구별을 최소화하기 위해 시맨틱 정보를 유지하도록 픽셀-단위 손실을 고려하여 상기 도메인 불변 영상을 생성할 수 있다.

[0015] 상기 불변 모델은, 단일 스케일 구조적 유사성 인덱스(SSIM) 손실을 반영한 일관성 제약을 고려하여 상기 도메인 불변 영상을 생성할 수 있다.

[0016] 상기 불변 모델은, 삼중항 손실(triplet loss)을 고려한 중간 특징(intermediate feature)을 이용하여 상기 도메인 불변 영상을 생성할 수 있다.

[0017] 상기 불변 모델은, 목적 함수를 만족하도록 상기 멀티스펙트럼 도메인의 영상을 이용하여 학습되되, 상기 목적 함수는 하기 수학식에 의해 계산되되,

[0018]
$$L(A) = \{L_{adv}(A) + L_{cycle}(A)\} + \{\lambda_0 L_{domain}(A) + \lambda_1 L_{disc}(A) + \lambda_2 L_{divs}(A)\}$$

[0019] 여기서, A는 도메인을 나타내고, $L_{adv}(A)$ 는 적대적 손실(adversarial loss) 함수를 나타내며, $L_{cycle}(A)$는 순환 손실 함수를 나타내며, $L_{domain}(A)$ 는 픽셀-단위 손실 함수를 나타내고, $L_{disc}(A)$는 일관성 제약 함수를 나타내고, $L_{divs}(A)$ 는 다양성 제약 함수를 나타내며, λ_0, λ_1 , λ_2 는 각각 가중치를 나타낸다.

[0020] 상기 불변 모델은, 자기 복제를 이용한 영상 변환법 기반의 판별 모델과 생성 모델 사이의 미니맥스 게임(minimax game)에 기초한 하기 수학식을 이용하여 상기 목적함수를 최적화하되,

[0021]
$$\mathbb{M}^* = \min_{\mathbb{M},\mathbb{G}} \max_{\mathbb{D}} L(A)$$

[0022] 여기서, M은 불변 모델을 나타내고, G는 생성 모델을 나타내며, D는 판별 모델을 나타낸다.

[0024] 본 발명의 다른 측면에 따르면, RGB 영상과 열 영상 사이의 도메인 격차를 줄일 수 있는 새로운 다중 스펙트럼 영상 변환 장치가 제공된다.

[0025] 본 발명의 일 실시예에 따르면, 입력 영상을 학습된 불변 모델에 적용하여 멀티스펙트럼 도메인 사이의 의미 있는(semantic) 공통 정보를 포함하는 도메인 불변 영상을 생성하는 불변 모델; 상기 도메인 불변 영상을 이용하여 도메인이 변환된 도메인 변환 영상을 생성하는 생성 모델; 및 상기 도메인 변환 영상을 이용하여 입력 도메인의 영상을 복구하도록 학습하는 판별 모델을 포함하는 다중 스펙트럼 영상 변환 장치가 제공될 수 있다.

[0026] 본 발명의 다른 실시예에 따르면, 멀티스펙트럼 도메인의 영상을 이용하여 상기 멀티스펙트럼 도메인 사이의 의미 있는 공통 정보를 포함하는 도메인 불변 영상을 생성하도록 불변 모델을 학습하는 인코더; 및 입력 영상을

상기 학습된 불변 모델에 적용하여 도메인 불변 영상을 생성하고, 상기 생성된 도메인 불변 영상을 이용하여 도메인 변환된 도메인 변환 영상을 생성하는 디코더를 포함하는 다중 스펙트럼 영상 변환 장치가 제공될 수 있다.

[0027] 상기 인코더 및 상기 디코더는 각각,

[0028] 멀티스펙트럼 도메인 사이의 공통 정보를 포함하는 도메인 불변 영상을 생성하는 불변 모델; 상기 도메인 불변 영상을 이용하여 도메인이 변환된 도메인 변환 영상을 생성하는 생성 모델; 및 상기 도메인 변환 영상을 이용하여 입력 도메인의 영상을 복구하도록 학습하는 판별 모델을 포함할 수 있다.

발명의 효과

[0030] 본 발명의 일 실시예에 따른 다중 스펙트럼 영상 변환 방법 및 그 장치를 제공함으로써, RGB 영상과 열 영상 사이의 도메인 격차를 줄일 수 있는 새로운 다중 스펙트럼 영상 변환이 가능케 할 수 있다.

[0031] 또한, 본 발명은 RGB 및 열 도메인 사이의 공통 정보를 포함하는 다중 스펙트럼 불변 영상을 추출할 수도 있다.

도면의 간단한 설명

[0033] 도 1은 본 발명의 일 실시예에 따른 다중 스펙트럼 영상 변환 방법을 나타낸 순서도.

도 2는 본 발명의 일 실시예에 따른 RGB 및 열 영상에 기초한 도메인 불면 영상을 예시한 도면.

도 3은 본 발명의 일 실시예에 따른 다중 스펙트럼 영상 변환 장치의 구성을 개략적으로 도시한 블록도.

도 4는 종래와 본 발명의 일 실시예에 따른 영상 변환 결과에 기반한 다중 스펙트럼 위치 인식 결과를 비교한 도면.

도 5는 다중 스펙트럼 시맨틱 세그멘테이션 작업을 통해 도메인 불변 영상을 평가한 결과를 나타낸 도면.

도 6은 종래와 본 발명의 일 실시예에 따른 방법에 따른 감독되지 않은 도메인 적응에 대한 실험 결과를 도시한 도면.

도 7은 입력 영상과 MNIST 및 USPS의 생성된 영상을 각각 예시한 도면.

발명을 실시하기 위한 구체적인 내용

[0034] 본 명세서에서 사용되는 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 명세서에서, "구성된다" 또는 "포함한다" 등의 용어는 명세서상에 기재된 여러 구성 요소들, 또는 여러 단계들을 반드시 모두 포함하는 것으로 해석되지 않아야 하며, 그 중 일부 구성 요소들 또는 일부 단계들은 포함되지 않을 수도 있고, 또는 추가적인 구성 요소 또는 단계들을 더 포함할 수 있는 것으로 해석되어야 한다. 또한, 명세서에 기재된 "...부", "모듈" 등의 용어는 적어도 하나의 기능이나 동작을 처리하는 단위를 의미하며, 이는 하드웨어 또는 소프트웨어로 구현되거나 하드웨어와 소프트웨어의 결합으로 구현될 수 있다.

[0035] 본 발명은 위치 인식을 위한 RGB 및 열 도메인 사이의 공통 정보를 포함하는 다중 스펙트럼 불변 영상을 생성하기 위한 것이다. 이하에서는 이해와 설명의 편의를 도모하기 위해 X는 영상으로 정의하기로 하며, Y는 도메인 불변 영상으로 정의하기로 한다. 또한, F는 중간 특징맵으로 정의하기로 한다.

[0036] 또한, 이하의 설명에서 표기법 AB는 A 도메인에서 B 도메인으로 변환되거나 BA는 그 반대를 의미한다. 이하에서는 이해와 설명의 편의를 도모하기 위해 도메인 A를 중심으로 공식화하여 설명하기로 한다.

[0037] 또한, 다중 스펙트럼 도메인에서 소스 도메인은 RGB 도메인 및 열 영상 도메인 중 어느 하나이며, 타겟 도메인은 RGB 도메인 및 열 영상 도메인 중 다른 하나로 확장 해석되어야 할 것이다.

[0038] 이하, 첨부된 도면들을 참조하여 본 발명의 실시예를 상세히 설명한다.

[0040] 도 1은 본 발명의 일 실시예에 따른 다중 스펙트럼 영상 변환 방법을 나타낸 순서도이며, 도 2는 본 발명의 일 실시예에 따른 RGB 및 열 영상에 기초한 도메인 불면 영상을 예시한 도면이다. 이하에서 설명되는 다중 스펙트

럼 불변 영상은 기본적으로 자기 복원을 이용한 영상 변환법을 기반으로 한다. 따라서, 하기에서 별도의 설명이 없더라도, 본 발명의 일 실시예에 따른 다중 스펙트럼 영상 변환 장치(100)의 각각의 구성은 자기 복원을 이용한 영상 변환법의 각각의 기능을 기본적으로 포함하는 것으로 확장 해석되어야 할 것이다. 자기 복원을 이용한 영상 변환법의 기본적인 동작은 당업자에게는 자명한 사항이므로 이에 대한 별도의 설명은 생략하기로 한다.

[0041] 단계 110에서 다중 스펙트럼 영상 변환 장치(100)는 입력 영상을 학습된 불변 모델에 적용하여 멀티스펙트럼 도메인 사이의 의미 있는(semantic) 공통 정보를 포함하는 도메인 불변 영상을 생성한다.

[0042] 이해와 설명의 편의를 도모하기 위해, 본 발명의 일 실시예에 따른 불변 모델이 도메인 불변 영상을 생성하는 방법에 대해 보다 상세히 설명하기로 한다.

[0043] 본 발명의 일 실시예에 따르면, 기본적으로 다중 스펙트럼 영상 변환 장치는 소스 도메인과 타겟 도메인간의 의미 있는 공통 정보를 포함하는 도메인 불변 영상을 생성할 수 있다. 이를 위해, 본 발명의 일 실시예에 따른 불변 모델은 다중 스펙트럼(멀티스펙트럼 도메인) 사이의 시맨틱 정보를 포함하는 도메인 불변 영상을 생성하도록 학습될 수 있다. 도메인 불변 영상은 두 도메인으로의 복구(recovery)가 가능하도록 시맨틱 정보를 유지하도록 생성될 수 있다.

[0044] 본 발명의 일 실시예에 따르면 입력 영상은 RGB 도메인 및 열 영상 도메인의 영상이며, 두 도메인(즉, RGB 도메인 및 열 영상 도메인) 사이의 공통 정보를 포함하는 도메인 불변 영상일 수 있다.

[0045] 이하에서, 소스 도메인(제1 도메인)은 RGB 도메인 및 열 영상 도메인 중 어느 하나일 수 있으며, 타겟 도메인 또는 제2 도메인은 RGB 도메인 및 열 영상 도메인 중 다른 하나일 수 있다.

[0046] 본 발명의 일 실시예에 따르면 다중 스펙트럼 영상 변환 장치(100)는 불변 모델을 포함한다. 불변 모델은 두 도메인에서 공통 정보를 추출하여 도메인 불변 영상을 생성할 수 있다.

[0047] 도 2를 참조하면, 도 2의 왼쪽에서부터 오른쪽으로 각각, RGB 영상, 열 영상, RGB 불변 영상, 열 불변 영상을 나타낸다.

[0048] 도 2에서 보여지는 바와 같이, 본 발명의 일 실시예에 따른 불변 모델은 두 도메인에서 공통 정보를 추출하여 도메인 불변 영상을 생성할 수 있으며, 도메인 불변 영상은 소스 도메인에서 타겟 도메인으로 변환된 영상을 다시 소스 도메인으로 변환할 수 있도록 의미 있는 공통 정보를 포함하도록 생성될 수 있다.

[0049] 이하에서는 불변 모델에서 도메인 불변 영상을 생성하는 방법에 대해 보다 상세히 설명하기로 한다.

[0050] 불변 모델은 다중 스펙트럼 도메인의 영상을 입력받아 시맨틱 정보를 포함하는 도메인 불변 영상을 추출하기 위한 수단이다.

[0051] 예를 들어, 불변 모델은 RGB 영상을 입력받아 시맨틱 정보를 포함하는 도메인 불변 영상을 추출(생성)할 수 있다.

[0052] 보다 상세히 설명하면, 도메인 불면 모델은 두 도메인(예를 들어, RGB 도메인 및 열 영상 도메인)으로부터 공통 정보를 추출하기 위한 도메인 불변 영상을 생성할 수 있다.

[0053] 즉, 불변 모델의 입력은 RGB 영상 및 열 영상이며, 출력은 도 2에서 보여지는 바와 같이, 다중 스펙트럼 설정하에 두 이미지의 시맨틱 정보를 포함하는 도메인 불변 영상이다.

[0054] 본 발명의 일 실시예에 따르면, 불변 모델은 생성 모델(G_A, G_B)와 유사하게 잔차 블록(residual block)으로 구성된다. 다만, 불변 모델은 두가지 차이점을 가지고 있다. 첫번째, 불변 모델은 다운 샘플링을 통한 미묘한 정보의 손실과 업 샘플링 레이어의 체크보드 패턴 아티팩트와 같은 아티팩트를 방지하기 위해 불변 모델에서 특징맵의 크기를 조정하지 않는다

[0055] 또한, 본 발명의 일 실시예에 따르면, 불변 모델은 출력 레이어 전에 자기 흡수 모듈로 구성될 수 있다. 도메인 불변 영상에서 RGB 영상의 패턴 및 텍스쳐 (texture)와 같은 중복된 상세 정보(detail)를 피하기 위해, 불변 모델은 화이트닝(whitening)과 같은 평균 및 분산을 이용하여 특징맵을 정규화한 후 ReLu를 이용하여 음수 변수를 잘라낼 수 있다.

[0056] 즉, 불변 모델은 수학식 1을 이용하여 불필요한 상세 정보(detail) 없이 RGB 영상과 열 영상으로부터 공통 정보를 포함하는 특징맵을 효과적으로 정규화할 수 있다.

수학식 1

[0057] $$F_{norm}(A) = \max\left\{\frac{Y_A - \mu_A}{\sqrt{\sigma_A^2 + \epsilon}}, 0\right\}$$

[0058] 또한, 본 발명의 일 실시예에 따르면, 불변 모델은 두가지 유형의 일관성 손실을 기반으로 학습될 수 있다.

[0059] 즉, 도메인 불변 영상으로부터 각 도메인 이미지를 복구(recover)할 수 있다는 것을 보증하기 위해, 본 발명의 일 실시예에 따른 불변 모델은 순환 일관성 제약을 적용한다.

[0060] 수학식 2의 순환 손실(cyclic loss)는 도메인 불변 영상이 RGB 영상과 열 영상 각각으로 복구하는데 유용한 시맨틱 정보를 포함할 수 있다.

[0061] 본 발명의 일 실시예에서는 소스 도메인(예를 들어, 도메인 A)와 타겟 도메인(예를 들어, 도메인 B)이 모두 동일한 순환 손실을 적용하는 것을 가정하여 이를 중심으로 설명하기로 한다.

[0062] 순환 손실은 수학식 2와 같이 계산될 수 있다.

수학식 2

[0063] $$L_{cycle}(A) = \|\mathbb{G}_A(\mathbb{M}(\mathbb{G}_B(\mathbb{M}(X_A)))) - X_A\|_1$$

[0064] 본 발명의 일 실시예에 따르면, 불변 모델은 RGB 영상과 열 영상으로부터의 출력이 동일한 불변 정보를 가지도록 학습될 수 있다. 따라서, 본 발명의 일 실시예에 따른 불변 모델은 다중 스펙트럼 도메인에서 불변 영상의 구별을 최소화하도록 수학식 3과 같이 픽셀-단위 손실(pixel-wise lose)이 적용될 수 있다.

[0065] 상술한 순환 손실과 픽셀-단위 손실에 기반한 일관성 제약은 모달리티 영상을 재구성하고 다중 스펙트럼 영상간의 차이를 최소화하기 위한 시맨틱 정보를 유지하도록 할 수 있다.

[0066] 따라서, 도메인 불변 영상은 단지 시맨틱 정보만을 가지는 것이 아니라 다중 스펙트럼 영상의 공유된 정보를 가질 수 있다.

[0067] 픽셀-단위 손실은 수학식 3과 같이 계산될 수 있다.

수학식 3

[0068] $$L_{domain}(A) = \|\mathbb{M}(\mathbb{G}_B(\mathbb{M}(X_A))) - \mathbb{M}(X_A)\|_1$$

[0069] 상술한 일관성 제약 조건은 불변 영상이 다중 스펙트럼 영상을 생성하도록 할 수 있다.

[0070] 또한, 위치 인식에 알맞은 적용 범위를 획득하기 위해 특별한 제약이 필요하다. 본 발명의 일 실시예에서는 두가지 유형의 제약을 판별성(discriminability) 및 다양성(diversity)으로 정의하기로 한다.

[0071] 본 발명의 일 실시예에서는 판별성 제약으로, 이미지 프로세싱 알고리즘을 평가하는 척도로 널리 이용되는 구조적 유사성 인덱스(SSIM: structural similarity index, 이하 SSIM이라 칭하기로 함)를 이용할 수 있다. 픽셀-단위 척도와 비교하여 SSIM은 로컬 구조의 변화를 설명하는 영상을 평가할 수 있다.

[0072] 위치 인식에서 글로벌 디스크립터(descriptor)는 로컬 디스크립터의 세트에 의해 계산되므로, 로컬리티의 유사성은 보다 나은 특징 표현을 위해 매우 중요하다.

[0073] 따라서, 본 발명의 일 실시예에서는 가우시안 대신 3 x 3 블록 필터를 이용하여 수학식 4와 같은 단일 스케일

SSIM 손실을 이용할 수 있다.

수학식 4

[0074] $$L_{disc}(A) = \frac{1}{N}\sum_{(i,j)}\frac{(1 - SSIM(Y_A, Y_{AB}))}{2}$$

[0075] 여기서, $Y_A = \mathbb{M}(X_A), Y_{AB} = \mathbb{M}(\mathbb{G}_B(\mathbb{M}(X_A)))$ 이다.

[0076] 모드 축소는 GAN을 기반으로 하는 방법에서는 까다로운 문제이며, 생성기가 축소되어 제한된 종류의 출력을 생성하게 된다. 만일 도메인 불변 영상이 축소되는 경우, 축소된 도메인 불변 영상으로 인해 위치 인식 성능이 저하될 수 있다.

[0077] 또한, 자율 주행의 많은 온-로드 시나리오에서는 차선, 도로, 중앙 소실점 또는 주변 차량과 같은 모양이 유사하므로 정확한 위치를 추정하는데 다양성이 매우 중요하다.

[0078] 이러한 이슈를 처리하기 위해, 본 발명의 일 실시예에서는 수학식 5와 같은 삼중항 손실을 이용한다. 여기서, F_A 로서 앵커 입력은 양의 F_{AB} 및 음의 F_B 와 비교되며, 그 반대도 마찬가지이다.

[0079] 일관성 및 다양성 제약 조건의 충돌을 방지하기 위해 본 발명의 일 실시예에 따른 불변 모델은 중간 특징(intermediate feature)이 이용될 수 있다.

수학식 5

[0080] $$L_{divs}(A) = \max\{d(F_A, F_{AB}) - d(F_A, F_B) + \alpha, 0\}$$

[0081] 여기서, $\alpha = 1,\ F_A = M_L(X_A), Y_A = M_U(M_L(X_A))$ 이다. M_L과 M_U는 입력 레이어와 출력 레이어 사이에서 정확하게 절반으로 분리된다.

[0082] 상술한 제약들을 통해 도메인 불변 영상을 차별화되고 다양하게 만들어줌으로써, 본 발명의 일 실시예에 따르면, 간접적으로 영상이 특정되도록 할 수 있다. 또한, 본 발명의 일 실시예에 따르면 기하학적 속성을 명시적으로 적용하지 않으며, 자동적으로 위치 인식이 가능하도록 할 수 있다.

[0083] 즉, 불변 모델은 하기의 목적 함수를 만족하도록 학습될 수 있다. 목적함수는 수학식 6과 같이 정리될 수 있다.

수학식 6

[0084] $$L(A) = \{L_{adv}(A) + L_{cycle}(A)\} + \{\lambda_0 L_{domain}(A) + \lambda_1 L_{disc}(A) + \lambda_2 L_{divs}(A)\}$$

[0085] 수학식 6에서 첫번째 중괄호는 자기 복원을 이용한 영상 변환법의 목적함수를 나타내고, L_{adv}는 적대적 손실(adversarial loss)을 나타낸다. 두번째 중괄호는 다중 스펙트럼 도메인 불변 영상을 생성하기 위한 목적함수를 나타낸다. 또한, λ_0, λ_1 및 λ_2 는 1, 0.1, 0.1로 각각 설정된다. 자기 복원을 이용한 영상 변환법으로부터 상속에 기반하여 목적함수를 최적화하면 최소-최대 문제를 해결할 수 있다.

[0086] 이를 수학식으로 나타내면, 수학식 7과 같다.

수학식 7

[0087] $$\mathbb{M}^* = \min_{\mathbb{M},G} \max_{\mathbb{D}} L(A)$$

[0088] 불변 모델을 학습한 후에, 본 발명의 일 실시예에서는 $\mathbb{M}$ 만을 이용하여 전처리 단계로 장소 인식을 위한 도메인 불변 영상을 생성할 수 있다.

[0089] 단계 115에서 다중 스펙트럼 영상 변환 장치(100)는 도메인 불변 영상을 이용하여 도메인이 변환된 도메인 변환 영상을 생성한다. 이미 전술한 바와 같이, 다중 스펙트럼 영상 변환 장치(100)는 자기 복원을 이용한 영상 변환법의 생성 모델(generator)를 포함한다. 해당 생성 모델은 소스 도메인의 영상을 타겟 도메인으로 변환된 도메인 변환 영상을 생성할 수 있다. 기본적은 동작은 자기 복원을 이용한 영상 변환법의 동작과 동일하므로 중복되는 설명은 생략하기로 한다. 즉, 다중 스펙트럼 영상 변환 장치(100)는 소스 도메인(예를 들어, RGB 도메인)의 영상을 기초로 학습된 불변 모델에서 생성된 도메인 불변 영상이 추출되면, 이를 이용하여 타겟 도메인(예를 들어, 열 영상 도메인)으로 변환된 도메인 변환 영상을 생성할 수 있다.

[0090] 단계 120에서 다중 스펙트럼 영상 변환 장치(100)는 생성된 도메인 변환 영상을 기반으로 소스 도메인과 타겟 도메인을 구별하도록 학습한다. 즉, 다중 스펙트럼 영상 변환 장치(100)는 판별 모델(discriminator)을 포함하며, 기본적인 동작은 자기 복원을 이용한 영상 변환법과 동일하므로 중복되는 설명은 생략하기로 한다.

[0092] 도 3은 본 발명의 일 실시예에 따른 다중 스펙트럼 영상 변환 장치의 구성을 개략적으로 도시한 블록도이며, 도 4는 종래와 본 발명의 일 실시예에 따른 영상 변환 결과에 기반한 다중 스펙트럼 위치 인식 결과를 비교한 도면이고, 도 5는 다중 스펙트럼 시맨틱 세그멘테이션 작업을 통해 도메인 불변 영상을 평가한 결과를 나타낸 도면이며, 도 6은 종래와 본 발명의 일 실시예에 따른 방법에 따른 감독되지 않은 도메인 적응에 대한 실험 결과를 도시한 도면이고, 도 7은 입력 영상과 MNIST 및 USPS의 생성된 영상을 각각 예시한 도면이다.

[0093] 도 3을 참조하면, 본 발명의 일 실시예에 따른 다중 스펙트럼 영상 변환 장치(100)는 인코더(310) 및 디코더(320)를 포함하여 구성된다.

[0094] 인코더(310) 및 디코더(320)는 각각 불변 모델을 가지는 생성 모델(G_A, G_B)과 판별 모델(D_A, D_B)을 포함하여 구성된다.

[0095] 각각의 구성의 동작 과정은 동일하므로, 하나의 구성에 대해서만 설명하기로 한다.

[0096] 불변 모델은 도 1에서 설명한 바와 동일하므로 중복되는 설명은 생략하기로 한다.

[0097] 생성 모델(G_A, G_B)은 소스 도메인의 도메인 불변 영상을 이용하여 타겟 도메인으로 변환된 도메인 변환 영상을 생성하도록 학습될 수 있다.

[0098] 판별 모델(D_A, D_B)은 생성 모델(G_A, G_B)에서 생성된 가상의 도메인 변환 영상과 실제 소스 도메인의 영상을 구별하도록 학습된다.

[0099] 생성 모델(G_A, G_B) 및 판별 모델(D_A, D_B)의 기본적인 동작은 자기 복원을 이용한 영상 변환법과 동일하므로 중복되는 설명은 생략하기로 한다.

[0100] 다만, 본 발명의 일 실시예에 따른 생성 모델(G_A, G_B) 및 판별 모델(D_A, D_B)은 종래의 자기 복원을 이용한 영상 변환법과는 상이하게 불변 모델을 포함하며, 해당 불변 모델은 두 도메인으로의 복구가 가능하도록 시맨틱 정보를 포함하는 도메인 불변 영상을 생성하도록 학습될 수 있다. 이에 대해서는 이미 전술한 바와 동일하므로 중복되는 설명은 생략하기로 한다.

[0101] 도 4는 종래와 본 발명의 일 실시예에 따른 영상 변환 결과에 기반한 다중 스펙트럼 위치 인식 결과를 비교한 것이다.

[0102] 도 4에서 보여지는 바와 같이, 본 발명의 일 실시예에 따른 영상 변환 결과에 기반한 결과가 종래와 비교하여

성능이 크게 개선된 것을 알 수 있다. 이는 종래의 방법은 속성이 서로 상이한 RGB 도메인과 열 영상 도메인의 세부 사항들을 복구하지 못하기 때문에 본 발명에 비해 성능이 낮은 것을 알 수 있다.

[0103] 결과적으로 본 발명의 일 실시예에 따른 다중 스펙트럼 도메인이 차별적이고 다양한 도메인 불변 영상을 생성하는데 효과적인 것을 알 수 있다.

[0104] 도 5는 다중 스펙트럼 시맨틱 세그멘테이션 작업을 통해 도메인 불변 영상을 평가한 결과를 도시한 것이다.

[0105] 본 발명의 일 실시예에서는 운전 중 일반적으로 발생하는 8가지 클래스의 장애물을 가지는 1569 픽셀 단위로 라벨링된 RGB 및 열 영상으로 구성된 다중 스펙트럼 시맨틱 세그멘테이션 데이터 셋을 사용했다. 소스 도메인으로 RGB 영상을 사용하고, 페어인 열 영상을 타겟 도메인으로 사용하였다.

[0106] 위치 인식 실험과 유사하게 데이터 세트의 학습 샘플을 이용하여 인코더를 학습한 후 불변 특징 영상을 추출한 후 이를 이용하여 시맨틱 세그멘테이션 모델을 학습하였다. 모든 구성은 클래스별 픽셀 정확도와 mIoU(mean Intersection-over-Union) 척도를 평가하였다.

[0107] 유효성 확인을 위해 각각 MFNET과 SegNet의 수정된 버전으로 두가지 유형의 세그멘테이션 모델을 테스트하였다.

[0108] 도 5에서 보여지는 바와 같이, 본 발명이 종래에 비해 mIoU 척도에서 명확하게 개선된 것을 알 수 잇다. 위치 인식 실험 결과 동일한 학습 절차 및 설정을 사용하더라도 본 발명이 가장 일반적인 클래스를 조정하는데 매우 효과적이며 영상 변환 기준 방법보다 더 낳은 평균 정확도를 보여주는 것을 알 수 있다.

[0109] 도 6은 종래와 본 발명의 일 실시예에 따른 방법에 따른 감독되지 않은 도메인 적응에 대한 실험 결과를 도시한 도면이다. 도 6에서 보여지는 바와 같이, 본 발명의 일 실시예에 따른 방법이 종래에 비해 개선된 성능을 보여주는 것을 알 수 있다. 도 7은 입력 영상과 MNIST 및 USPS의 생성된 영상을 각각 예시한 도면으로, 모델이 숫자를 성공적으로 강조 표시하고, 두 영상에서 공통의 시맨틱 영역을 추출할 수 있는 것을 알 수 있다.

[0111] 본 발명의 실시 예에 따른 장치 및 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 컴퓨터 판독 가능 매체에 기록되는 프로그램 명령은 본 발명을 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 분야 통상의 기술자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media) 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다.

[0112] 상술한 하드웨어 장치는 본 발명의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.

[0113] 이제까지 본 발명에 대하여 그 실시 예들을 중심으로 살펴보았다. 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자는 본 발명이 본 발명의 본질적인 특성에서 벗어나지 않는 범위에서 변형된 형태로 구현될 수 있음을 이해할 수 있을 것이다. 그러므로 개시된 실시 예들은 한정적인 관점이 아니라 설명적인 관점에서 고려되어야 한다. 본 발명의 범위는 전술한 설명이 아니라 특허청구범위에 나타나 있으며, 그와 동등한 범위 내에 있는 모든 차이점은 본 발명에 포함된 것으로 해석되어야 할 것이다.

부호의 설명

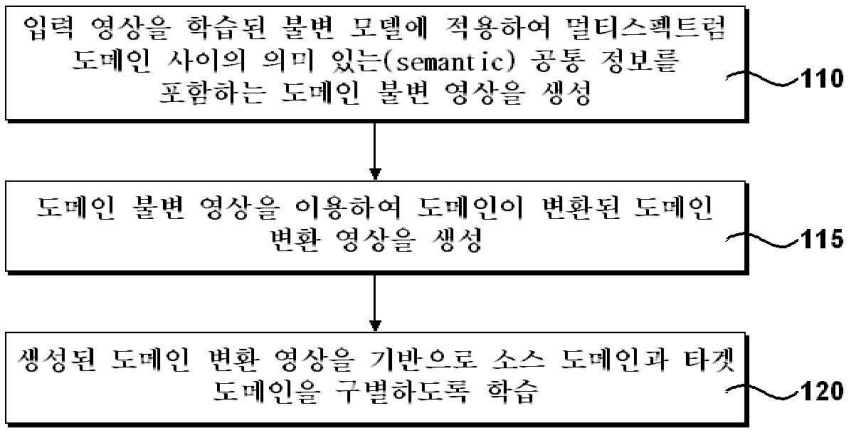
[0115] 100: 다중 스펙트럼 영상 변환 장치

310: 인코더

320: 디코더

도면

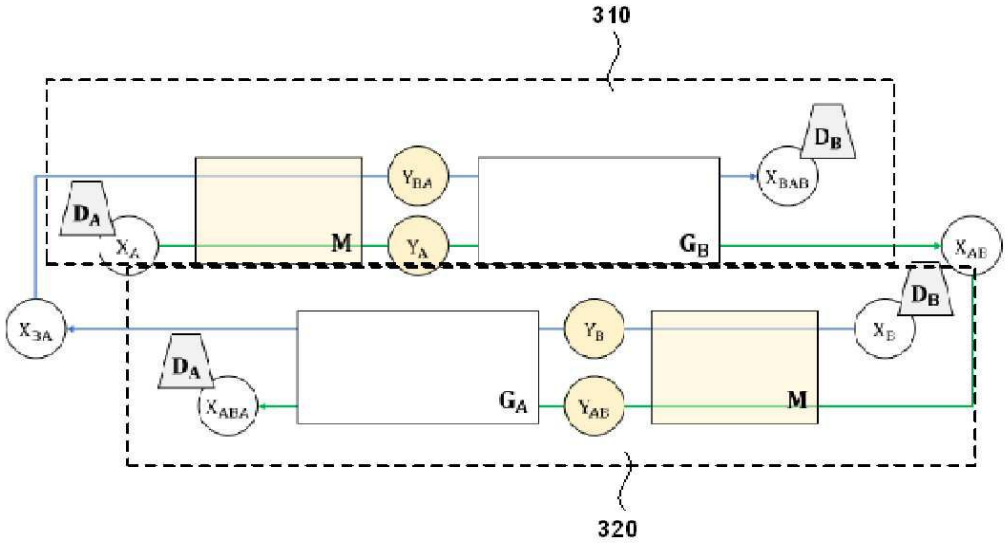
도면1


입력 영상을 학습된 불변 모델에 적용하여 멀티스펙트럼 도메인 사이의 의미 있는(semantic) 공통 정보를 포함하는 도메인 불변 영상을 생성
110
도메인 불변 영상을 이용하여 도메인이 변환된 도메인 변환 영상을 생성
115
생성된 도메인 변환 영상을 기반으로 소스 도메인과 타겟 도메인을 구별하도록 학습
120

도면2



도면3



도면4

Methods	0.25m	0.50m	1.00m	2.00m	3.00m	5.00m	10.0m	0.25m	0.50m	1.00m	2.00m	3.00m	5.00m	10.0m
	Top-1							Top-5						
DSIFT+BOW [22]	0.6	1.5	3.7	5.4	7.9	10.8	16.3	1.7	4.8	11.6	14.4	17.7	21.8	29.6
DenseVLAD [20]	0.6	1.6	4.3	6.1	8.1	10.1	12.6	1.8	5.1	12.9	15.5	18.1	21.2	26.2
NetVLAD [21]	0.2	0.5	1.2	1.9	2.8	4.2	7.1	1.1	2.8	6.6	10.9	17.0	25.8	44.9
CycleGAN [15]	0.2	0.6	1.2	2.1	2.8	2.8	3.5	0.8	1.9	4.1	5.0	6.2	7.7	9.8
Pix2pixHD [13]	0.5	1.6	3.8	5.5	8.0	10.9	14.6	2.1	5.7	13.5	16.4	19.9	23.4	28.6
Excavate [13]	0.1	0.4	1.3	2.1	2.8	4.0	5.4	1.1	2.1	5.1	6.3	7.6	9.0	11.0
Ours	1.4	5.3	12.2	18.3	24.7	30.6	37.4	5.0	15.2	36.3	42.5	48.4	53.0	57.6

도면5

Methods	unlabeld	car	pedestrian	bike	curve	car stop	guardrail	color cone	bump	class avg.	mIoU
	MFNet [27]										
Source Only	93.2	7.2	8.5	0.	0.	0.	0.	0.	0.	12.1	5.4
CycleGAN [15]	95.6	11.0	17.1	33.4	0.	0.	0.	0.	0.	17.4	6.1
Ours	93.5	24.8	13.7	3.7	26.9	0.	0.	0.	52.0	23.9	21.9
Target Only	95.7	67.6	78.0	58.7	64.4	10.0	0.	0.	36.4	45.7	39.4
	SegNet [32]										
Source Only	92.9	7.8	24.6	0.	0.	3.2	0.	0.	0.	14.3	10.2
CycleGAN [15]	93.2	28.6	44.6	0.	0.	6.6	0.	0.	0.	19.2	10.7
Ours	93.4	43.6	29.8	0.	11.0	0.	0.	0.	42.2	24.4	26.3
Target Only	96.4	72.5	81.3	56.9	58.6	0.	0.	0.	3.6	41.0	36.2

도면6

Source	MNIST	USPS
Target	USPS	MNIST
Source only	96.3	76.9
SE[42]	**97.3**	98.1
DCRN[22]	73.7	91.8
G2A[43]	90.8	92.5
ADDA[44]	90.1	89.4
SBADA-GAN[45]	97.6	95.4
CyCADA[46]	95.7	94.8
Ours	**94.9**	**98.7**
Target only	97.8	99.6

도면7

