



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2020년06월26일
(11) 등록번호 10-2127153
(24) 등록일자 2020년06월22일

- (51) 국제특허분류(Int. Cl.)
G06T 7/50 (2017.01) G06N 20/00 (2019.01)
G06T 5/00 (2019.01)
- (52) CPC특허분류
G06T 7/50 (2017.01)
G06N 20/00 (2019.01)
- (21) 출원번호 10-2020-0043096
(22) 출원일자 2020년04월09일
심사청구일자 2020년04월09일
- (56) 선행기술조사문헌
"generative adversarial networks for depth map estimation from RGB video"
"analyzing modular cnn architectures for joint depth prediction and semantic segmentation"
"geometry-aware symmetric domain adaptation for monocular depth estimation"

- (73) 특허권자
한밭대학교 산학협력단
대전광역시 유성구 동서대로 125 (덕명동)
- (72) 발명자
이승호
대전광역시 유성구 엑스포로 448 엑스포아파트 402동 402호
박동훈
충청북도 청주시 흥덕구 대농로 55 두산위브지웰 시티2차 203동 2904호
- (74) 대리인
이은철, 이우영

전체 청구항 수 : 총 13 항

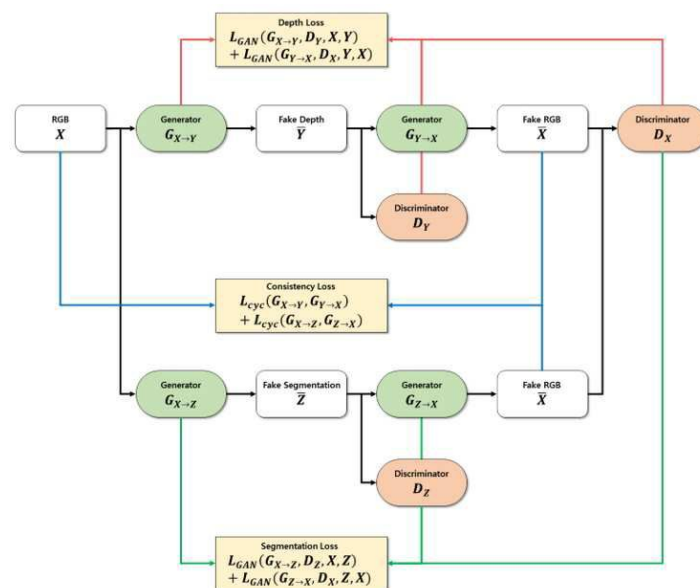
심사관 : 박금옥

(54) 발명의 명칭 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 방법 및 시스템

(57) 요약

본 발명은 특수 장비나 카메라를 이용하지 않고 사이클 GAN과 세그멘테이션을 통해 단일 영상만을 사용하여 영상의 깊이 정보를 추정하는 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 방법 및 시스템에 관한 것이다. 본 발명의 실시 예에 따른 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 방법은 표준 데이터베이스의 입력 RGB 영상(뒷면에 계속)

대표도 - 도3



에 대하여 생성자를 이용하여 깊이 정보와 세그멘테이션 영상 정보를 생성하는 단계(S10)와, 생성된 깊이 정보와 세그멘테이션 영상 정보를 이용하여 RGB 영상을 복원하는 단계(S20) 및 생성된 깊이 정보와 세그멘테이션 영상 정보 및 복원된 RGB 영상을 표준 데이터베이스와 각각 비교 판별하여 손실 및 판별 확률을 계산하는 단계(S30)를 포함한다. 또한, 계산된 결과값을 토대로 각각의 손실 및 판별 확률 값이 미리 설정된 기준 수렴값을 만족하는지 판단하는 단계(S40)와, 판단 결과를 토대로 각각의 손실 및 판별자의 판별 확률 값이 미리 설정된 기준 수렴값에 수렴되도록 학습을 조정하고, 상기 (S10) 단계 내지 (S40) 단계를 반복 수행하는 단계(S50)와, 상기 (S10) 단계 내지 (S50) 단계를 통해 생성된 생성자를 이용하여 RGB 영상에 대한 깊이 정보를 추정하는 단계(S60)를 포함한다.

(52) CPC특허분류

G06T 5/001 (2013.01)

G06T 2207/20081 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호 1711107408

부처명 과학기술정보통신부

연구관리전문기관 한국연구재단

연구사업명 개인기초연구(과기정통부)(R&D)

연구과제명 효율적인 시설물 관리를 위한 3D 맵의 자동 생성 및 AR 솔루션 개발

기 여 율 1/1

주관기관 한밭대학교

연구기간 2020.03.01 ~ 2021.02.28

공지예외적용 : 있음

명세서

청구범위

청구항 1

단일 영상만을 사용하여 영상의 깊이 정보를 추정하는 깊이 추정 방법에 있어서,

표준 데이터베이스의 입력 RGB 영상 X 에 대하여 생성자(Generator) $G_{X \rightarrow Y}$ 와 생성자(Generator) $G_{X \rightarrow Z}$ 을 이용하여 깊이 정보와 세그멘테이션 영상 정보를 생성하는 단계(S10);

생성된 깊이 정보와 세그멘테이션 영상 정보를 이용하여 RGB 영상을 복원하는 단계(S20);

생성된 깊이 정보와 세그멘테이션 영상 정보 및 복원된 RGB 영상을 표준 데이터베이스와 각각 판별하여 비교하고, 각각에 대한 손실(Loss) 및 판별 확률을 계산하는 단계(S30);

계산된 결과값을 토대로 각각의 손실 및 판별자(Discriminator)의 판별 확률 값이 미리 설정된 기준 수렴값을 만족하는지 판단하는 단계(S40);

판단 결과를 토대로 손실 및 판별 확률 값이 미리 설정된 기준 수렴값을 만족하지 않는 경우 각각의 손실 및 판별자의 판별 확률 값이 미리 설정된 기준 수렴값에 수렴되도록 학습을 조정하고, 상기 (S10) 단계 내지 (S40) 단계를 반복 수행하는 단계(S50); 및

상기 (S10) 단계 내지 (S50) 단계를 통해 생성된 생성자(Generator) $G_{X \rightarrow Y}$ 를 이용하여 RGB 데이터의 입력 RGB 영상에 대한 깊이 정보를 추정하는 단계(S60)를 포함하는 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 방법.

청구항 2

제1항에 있어서,

상기 (S10) 단계 내지 (S50) 단계를 통해 생성된 생성자(Generator) $G_{X \rightarrow Z}$ 을 이용하여 상기 RGB 데이터의 입력 RGB 영상에 대한 세그멘테이션 영상 정보를 추정하는 단계(S70)를 더 포함하는 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 방법.

청구항 3

제1항에 있어서,

상기 표준 데이터베이스는 RGB 영상 정보와 깊이 정보 및 세그멘테이션 정보를 포함하는 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 방법.

청구항 4

제1항에 있어서,

상기 표준 데이터베이스는 NYU Depth Dataset V2인 것을 특징으로 하는 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 방법.

청구항 5

제1항에 있어서,

상기 입력 RGB 영상에 대한 깊이 정보를 추정하는 단계(S60)에서 상기 RGB 데이터는 RGB 영상 정보만을 포함하는 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 방법.

청구항 6

제1항에 있어서,

상기 손실(Loss) 및 판별 확률을 계산하는 단계(S30)는 사이클 GAN의 목적 함수를 통해 손실 및 판별 확률 결과값의 수치를 계산하는 것을 특징으로 하는 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 방법.

청구항 7

제6항에 있어서,

상기 목적 함수는 사이클 GAN(Generative Adversarial Network)의 적대적인 손실 함수(Adversarial Loss Function)와 사이클 일관성 손실 함수(Cycle-Consistency Loss Function)로 구성되는 것을 특징으로 하는 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 방법.

청구항 8

제1항에 있어서,

상기 (S10) 단계 내지 (S40) 단계를 반복 수행하는 단계(S50)에서 손실(Loss)은 0에 수렴하고, 판별자의 판별 확률 값은 50%에 수렴하도록 학습을 조정하는 것을 특징으로 하는 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 방법.

청구항 9

제1항에 있어서,

상기 손실(Loss) 및 판별 확률을 계산하는 단계(S30) 이후에,

동일한 입력 RGB 영상 X 에 대하여 서로 다른 출력값을 갖지만 후에 이를 다시 RGB 영상 $\bar{X}$ 로 복원하였을 때 원본 RGB 영상 X 와 복원된 RGB 영상 $\bar{X}$ 를 비교하여 원본 RGB 영상 X 의 형상을 유지하면서 깊이 정보와 세그멘테이션 영상을 생성할 수 있도록 유도하는 사이클 일관성 손실(Cycle-Consistency Loss) 단계를 더 포함하는 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 방법.

청구항 10

제9항에 있어서,

상기 사이클 일관성 손실(Cycle-Consistency Loss) 단계는,

상기 생성자와 판별자의 미니맥스(Minimax) 결과를 반영(Back-propagation)하여 학습을 진행하기 위해 상기 (S10) 단계 내지 (S30) 단계를 통해 생성된 생성자(Generator) $G_{x \rightarrow y}$ 와, 생성자(Generator) $G_{x \rightarrow z}$ 을 피드백하여 깊이 정보와 세그멘테이션 영상 정보를 추정하는 것을 특징으로 하는 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 방법.

청구항 11

단일 영상만을 사용하여 영상의 깊이 정보를 추정하는 깊이 추정 시스템에 있어서,

표준 데이터베이스의 RGB 영상을 입력받고, 생성자(Generator) $G_{x \rightarrow y}$ 를 이용하여 깊이 정보를 생성하며, 생성자(Generator) $G_{x \rightarrow z}$ 을 이용하여 세그멘테이션 영상 정보를 생성하고, 생성된 깊이 정보와 세그멘테이션 영상 정보를 이용하여 RGB 영상을 복원하며, 사이클 GAN의 목적 함수를 통해 학습을 수행하는 영상정보 학습부;

상기 영상정보 학습부에서 생성된 깊이 정보와 세그멘테이션 영상 정보 및 복원된 RGB 영상을 표준 데이터베이스와 각각 판별하여 비교하고, 각각에 대한 손실(Loss) 및 판별 확률을 계산하는 연산부;

상기 연산부에서 계산된 결과값을 토대로 각각의 손실 및 판별자(Discriminator)의 판별 확률 값이 미리 설정된 기준 수렴값을 만족하는지 판단하는 판단부;

RGB 영상을 입력받는 영상 입력부; 및

상기 영상정보 학습부에서 학습이 완료된 상기 생성자(Generator) $G_{x \rightarrow z}$ 와, 생성자(Generator) $G_{x \rightarrow y}$ 를 이용하여 영상 입력부에서 입력받은 RGB 영상에 대한 깊이 정보를 추정하는 영상정보 추정부를 포함하는 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 시스템.

청구항 12

제11항에 있어서,

상기 영상정보 학습부는 RGB 영상 정보에서 세그멘테이션 정보로 변환하기 위하여 생성자(Generator) $G_{x \rightarrow z}$ 에 의해 입력 RGB 영상 X에 대한 세그멘테이션 정보를 획득하여 해당 정보를 깊이 정보 추정에 활용하고,

상기 생성자(Generator) $G_{x \rightarrow y}$ 에 의해 입력 RGB 영상 X에 대한 깊이 정보를 획득하여 해당 정보를 세그멘테이션 정보 추정에 활용하는 것을 특징으로 하는 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 시스템.

청구항 13

제11항에 있어서,

상기 판단부는 판단 결과를 토대로 손실 및 판별 확률 값이 미리 설정된 기준 수렴값을 만족하지 않는 경우 각각의 손실 및 판별자의 판별 확률 값이 미리 설정된 기준 수렴값에 수렴되도록 학습을 조정하고, 상기 영상정보 학습부에서 재학습이 수행되도록 피드백하는 것을 특징으로 하는 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 시스템.

발명의 설명

기술 분야

[0001] 본 발명은 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 방법 및 시스템에 관한 것으로서, 더욱 상세하게는 특수 장비나 카메라를 이용하지 않고 사이클 GAN과 세그멘테이션을 통해 단일 영상만을 사용하여 영상의 깊이 정보를 추정하는 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 방법 및 시스템에 관한 것이다.

배경 기술

[0003] 영상처리 분야에서 3차원 정보란 영상의 시각적 정보 이외의 깊이(Depth), 규모(Scale) 등의 공간감적 정보가 포함된 정보를 나타낸다. 4차 산업혁명을 시작으로 VR, AR 및 자율주행 분야 등에서 이러한 3차원 정보는 없어서 안 되는 필수적인 정보이며, 이를 보다 정확하고 빠른 시간 내 획득하고 계산할 수 있는 기술 등을 요구하고

있다.

[0004] 예를 들어, 증강현실(AR) 분야에서는 실제 환경 위에 가상의 환경을 덮어씌워 사용자에게 부가적인 정보를 보강하여 제공한다. 컴퓨터 그래픽으로 만들어진 가상환경이 실제 환경과 자연스럽게 오버랩되어 사용자에게 보다 몰입감 있는 서비스를 제공할 수 있다. 이러한 기술들은 카메라를 통해 들어오는 시각적인 정보에 3차원 정보가 결합된 형태로 구성되어야 자연스러운 형태의 가상 환경을 구축할 수 있다.

[0005] 따라서 이러한 3차원 정보를 얻기 위하여 레이더나 초음파 및 레이저 센서 등이 개발되어 왔으며 아울러 특수 카메라나 스테레오 카메라들을 통한 3차원 촬영 방법 등이 제시되어 왔다.

[0006] 하지만, 종래의 3차원 정보를 얻기 위해서는 특수 장비나 카메라, 레이더, 초음파 및 센서 등을 사용함으로써 3차원 정보를 추출하기 위한 비용이 높고, 자료를 쉽게 구할 수 없는 문제가 있다.

선행기술문헌

특허문헌

[0008] (특허문헌 0001) 대한민국 등록특허 제10-1650702호(2016년 08월 24일 공고)

발명의 내용

해결하려는 과제

[0009] 따라서, 본 발명은 종래의 단점을 해결하기 위한 것으로서, 특수 장비나 카메라, 레이더, 초음파 및 센서를 이용하지 않고 단일 카메라만을 사용하여 저렴하게 3차원 영상을 추출하고자 하는데 그 목적이 있다. 또한, 3차원 영상 정보를 생성하기 위한 자료를 용이하게 구할 수 있도록 하는데 그 목적이 있다. 또한, 깊이 정보를 추정하는 과정에서 발생하는 데이터 불균형 문제를 해결하고자 하는데 그 목적이 있다.

과제의 해결 수단

[0011] 이러한 기술적 과제를 이루기 위한 본 발명의 일 측면에 따른 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 방법은 표준 데이터베이스의 입력 RGB 영상 X 에 대하여 생성자(Generator) $G_{x \rightarrow y}$ 와 생성자(Generator) $G_{x \rightarrow z}$ 를 이용하여 깊이 정보와 세그멘테이션 영상 정보를 생성하는 단계(S10)와, 생성된 깊이 정보와 세그멘테이션 영상 정보를 이용하여 RGB 영상을 복원하는 단계(S20) 및 생성된 깊이 정보와 세그멘테이션 영상 정보 및 복원된 RGB 영상을 표준 데이터베이스와 각각 판별하여 비교하고, 각각에 대한 손실(Loss) 및 판별 확률을 계산하는 단계(S30)를 포함한다.

[0012] 또한, 계산된 결과값을 토대로 각각의 손실 및 판별자(Discriminator)의 판별 확률 값이 미리 설정된 기준 수렴값을 만족하는지 판단하는 단계(S40)와, 판단 결과를 토대로 손실 및 판별 확률 값이 미리 설정된 기준 수렴값을 만족하지 않는 경우 각각의 손실 및 판별자의 판별 확률 값이 미리 설정된 기준 수렴값에 수렴되도록 학습을 조정하고, 상기 (S10) 단계 내지 (S40) 단계를 반복 수행하는 단계(S50)를 포함한다.

[0013] 또한, 상기 (S10) 단계 내지 (S50) 단계를 통해 생성된 생성자(Generator) $G_{x \rightarrow y}$ 를 이용하여 RGB 데이터의 입력 RGB 영상에 대한 깊이 정보를 추정하는 단계(S60)와, 상기 (S10) 단계 내지 (S50) 단계를 통해 생성된 생성자(Generator) $G_{x \rightarrow z}$ 를 이용하여 상기 RGB 데이터의 입력 RGB 영상에 대한 세그멘테이션 영상 정보를 추정하는 단계(S70)를 포함한다.

[0014] 또한, 본 발명의 다른 측면에 따른 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 시스템은 영상정보 학습부, 연산부, 판단부, 데이터베이스, 영상 입력부 및 영상정보 추정부를 포함한다. 이때, 상기 데이터베이스는 표준 데이터베이스를 포함한다.

[0015] 또한, 상기 영상정보 학습부는 표준 데이터베이스의 RGB 영상을 입력받고, 생성자(Generator) $G_{x \rightarrow y}$ 를 이용하여 깊이 정보를 생성하며, 생성자(Generator) $G_{x \rightarrow z}$ 를 이용하여 세그멘테이션 영상 정보를 생성하고, 생성된

깊이 정보와 세그멘테이션 영상 정보를 이용하여 RGB 영상을 복원하며, 사이클 GAN의 목적 함수를 통해 학습을 수행한다.

[0016] 또한, 상기 연산부는 영상정보 학습부에서 생성된 깊이 정보와 세그멘테이션 영상 정보 및 복원된 RGB 영상을 표준 데이터베이스와 각각 판별하여 비교하고, 각각에 대한 손실(Loss) 및 판별 확률을 계산한다. 또한, 상기 판단부는 연산부에서 계산된 결과값을 토대로 각각의 손실 및 판별자(Discriminator)의 판별 확률 값이 미리 설정된 기준 수렴값을 만족하는지 판단한다.

[0017] 또한, 상기 영상 입력부는 RGB 영상을 입력받는다. 또한, 영상정보 추정부는 영상정보 학습부에서 학습이 완료된 상기 생성자(Generator) $G_{x \rightarrow z}$ 와, 생성자(Generator) $G_{x \rightarrow y}$ 를 이용하여 영상 입력부에서 입력받은 RGB 영상에 대한 깊이 정보를 추정한다.

발명의 효과

[0019] 이상에서 설명한 바와 같이, 본 발명에 따른 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 방법 및 시스템은 특수 장비나 카메라, 레이더, 초음파 및 센서를 이용하지 않고 단일 영상만을 사용하여 3차원 정보를 생성함으로써 저렴하게 3차원 영상을 추출할 수 있는 효과가 있다. 또한, 확장성이 높아 스테레오 이미지, 광학 흐름 기법 또는 포인트 클라우드와 같은 다른 정보를 사용할 수 없는 경우에도 3차원 정보를 생성할 수 있고, 3차원 정보를 추출하기 위한 장비의 소형화에 유리한 효과가 있다.

[0020] 또한, 단일 영상을 사용하여 영상의 깊이 정보를 추정할 수 있어 3차원 영상 정보를 생성하기 위한 자료를 용이하게 구할 수 있다. 또한, 세그멘테이션(Segmentation)을 토대로 깊이 정보를 추정하는 과정에서 발생하는 데이터 불균형 문제를 시각적으로 표시하고 상대적으로 큰 특징에 묻혀 소실되는 작은 특징들을 부각시켜 해결할 수 있는 효과가 있다.

도면의 간단한 설명

[0022] 도 1은 종래의 깊이 추정 과정에서 발생하는 문제점을 나타내는 도면이다.
 도 2는 본 발명의 실시 예에 따른 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 시스템을 나타내는 구성도이다.
 도 3은 본 발명의 실시 예에 따른 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 방법을 나타내는 개념도이다.
 도 4는 본 발명의 실시 예에 따른 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 방법을 나타내는 흐름도이다.
 도 5는 본 발명의 실시 예에 따른 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 방법을 나타내는 순서도이다.
 도 6은 본 발명의 실시 예에 따른 단일 영상의 깊이 정보를 추정하는 방법의 동작 순서를 나타내는 도면이다.
 도 7은 본 발명의 실시 예에 따른 세그멘테이션 추정 과정을 나타내는 도면이다.
 도 8은 본 발명의 실시 예에 따른 깊이 추정 과정을 나타내는 도면이다.
 도 9는 생성자(Generator)의 생성 분포 및 판별자(Discriminator)의 판별 확률을 나타내는 도면이다.
 도 10은 본 발명의 실시 예에 따른 사이클 일관성 손실(Cycle-Consistency Loss)을 나타내는 도면이다.
 도 11은 본 발명의 실시 예에 따른 실행단계의 깊이 정보 추정 단계를 나타내는 도면이다.
 도 12는 본 발명의 실시 예에 따른 실행단계의 세그멘테이션 정보 추정 단계를 나타내는 도면이다.
 도 13a 및 도 13b는 깊이 정보 추정 과정에서 세그멘테이션 과정을 사용하기 전과 후를 비교하여 나타내는 도면이다.
 도 14는 본 발명의 실시 예에 따른 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 방법의 평가 절차를 나타내는 도면이다.

발명을 실시하기 위한 구체적인 내용

[0023] 아래에서는 첨부한 도면을 참고로 하여 본 발명의 실시 예에 대하여 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 상세히 설명한다. 그러나 본 발명은 여러 가지 상이한 형태로 구현

될 수 있으며 여기에서 설명하는 실시 예에 한정되지 않는다. 그리고 도면에서 본 발명을 명확하게 설명하기 위해서 설명과 관계없는 부분은 생략하였으며, 명세서 전체를 통하여 유사한 부분에 대해서는 유사한 도면부호를 붙였다.

- [0024] 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" 등의 용어는 적어도 하나의 기능이나 동작을 처리하는 단위를 의미하며, 이는 하드웨어나 또는 소프트웨어 또는 하드웨어 및 소프트웨어의 결합으로 구현될 수 있다.
- [0025] 이하, 첨부된 도면을 참조하여 본 발명의 바람직한 실시 예를 설명함으로써, 본 발명을 상세히 설명한다.
- [0026] 각 도면에 제시된 동일한 참조 부호는 동일한 부재를 나타낸다.
- [0027] 도 1은 종래의 깊이 추정 과정에서 발생하는 문제점을 나타내는 도면이고, 도 2는 본 발명의 실시 예에 따른 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 시스템(10)을 나타내는 구성도이다. 즉, 도 1은 영상에 대한 종래의 깊이(Depth) 추정 과정에서 깊이 정보가 모호하게 나오는 문제점을 나타내는 도면이다.
- [0028] GAN(Generative Adversarial Network)은 적대적(Adversarial)으로 생성(Generative)하는 네트워크(Network)의 의미로서 비지도 학습 기반의 생성모델(Unsupervised generative model)을 나타낸다. 이는 서로 상대적인 특성을 가진 2개의 신경망이 서로 경쟁하여 상승효과를 나타내게 된다.
- [0029] 상기 GAN은 각각 데이터 인스턴스를 생성하는 생성자(Generator)와 데이터의 진위 여부를 판단하는 판별자(Discriminator)를 포함한다. 여기에서, 생성자는 Zero-mean Gaussian으로 생성되는 랜덤 노이즈(Random Noise) z 를 입력받아 실제 데이터 분포와 유사한 페이크 데이터(Fake data)를 생성한다.
- [0030] 이와는 대조적으로 상기 판별자는 생성자가 생성하는 데이터가 페이크 데이터인지 트레이닝 데이터셋의 데이터인지 구별하여 각각에 대한 확률을 나타낸다. 따라서, 판별자는 실수할 확률을 낮추고자 동작하고, 생성자는 판별자가 실수할 확률을 높이고자 동작하게 되는데 이를 미니맥스 문제(Minimax Problem)라 한다.
- [0031] 본 발명의 실시 예에 따른 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 방법 및 시스템(10)은 종래의 특수 장비나 여러 장의 영상 등을 필요로 하는 종래의 깊이 정보 추정 방식을 개선함으로써 사이클 GAN(Cycle Generative Adversarial Network)과 세그멘테이션(Segmentation)을 사용하여 깊이 정보를 추정할 수 있다.
- [0032] 일반적으로 RGB 영상에서 학습을 통해 깊이(Depth) 정보를 추정할 때, 도 1에서 도시된 바와 같이 학습 데이터 간 데이터 불균형을 이유로 비교적 덜 학습된 특징에 대해서는 깊이 정보가 모호하게 나오거나 아예 큰 특징에 묻히게 되는 페이딩(Fading) 등의 문제점들이 발생한다.
- [0033] 따라서, 본 발명의 실시 예에 따른 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 방법 및 시스템(10)은 종래의 깊이 정보를 추정하는 과정에서 발생하는 데이터 불균형 문제를 시각적으로 표시하고 상대적으로 큰 특징에 묻혀 소실되는 작은 특징들을 부각시키려는 목적으로 세그멘테이션(Segmentation)을 도입하여 해결하고자 한다.
- [0034] 본 발명의 실시 예에 따른 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 시스템(10)은 영상정보 학습부(100), 연산부(200), 판단부(300), 데이터베이스(400), 영상 입력부(500) 및 영상정보 추정부(600)를 포함할 수 있다. 이때, 데이터베이스(400)는 표준 데이터베이스(410)를 포함한다.
- [0035] 영상정보 학습부(100)는 표준 데이터베이스(410)의 RGB 영상을 입력받고, 생성자(Generator) $G_{x \rightarrow y}$ 를 이용하여 깊이 정보를 추정하며, 생성자(Generator) $G_{x \rightarrow z}$ 를 이용하여 세그멘테이션 영상 정보를 추정한다.
- [0036] 이때, 영상정보 학습부(100)는 RGB 영상 정보에서 세그멘테이션 정보로 변환하기 위하여 생성자(Generator) $G_{x \rightarrow z}$ 에 의해 입력 RGB 영상 X 에 대한 세그멘테이션 정보를 획득하여 해당 정보를 깊이 정보 추정에 활용한다.
- [0037] 또한, 영상정보 학습부(100)는 생성자(Generator) $G_{x \rightarrow y}$ 에 의해 입력 RGB 영상 X 에 대한 깊이 정보를 획득하여 해당 정보를 세그멘테이션 정보 추정에 활용한다. 또한, 영상정보 학습부(100)는 생성된 깊이 정보와 세그멘테이션 영상 정보를 이용하여 RGB 영상을 복원한다. 또한, 영상정보 학습부(100)는 사이클 GAN의 목적 함수를 통해 생성자(Generator) $G_{x \rightarrow z}$ 와, 생성자(Generator) $G_{x \rightarrow y}$ 에 대한 학습을 수행한다.
- [0038] 연산부(200)는 생성된 깊이 정보와 세그멘테이션 영상 정보 및 복원된 RGB 영상을 표준 데이터베이스(410)와 각

각 판별하여 비교하고, 각각에 대한 손실(Loss) 및 판별 확률을 계산한다. 이때, 연산부(200)는 사이클 GAN의 목적 함수를 통해 손실 및 판별 확률 결과값의 수치를 계산한다.

- [0039] 판단부(300)는 연산부(200)에서 계산된 결과값을 토대로 각각의 손실 및 판별자(Discriminator)의 판별 확률 값이 미리 설정된 기준 수렴값을 만족하는지 판단한다.
- [0040] 또한, 판단부(300)는 판단 결과를 토대로 손실 및 판별 확률 값이 미리 설정된 기준 수렴값을 만족하지 않는 경우 각각의 손실 및 판별자의 판별 확률 값이 미리 설정된 기준 수렴값에 수렴되도록 학습을 조정하고, 영상정보 학습부(100)에 피드백하여 재학습 또는 깊이 정보와 세그멘테이션 영상 정보를 다시 추정할 수 있도록 유도한다. 즉, 판단부(300)는 영상정보 학습부(100)에서 재학습이 수행되도록 조정된 결과를 영상정보 학습부(100)에 피드백한다.
- [0041] 데이터베이스(400)는 영상정보 학습부(100)에서 학습을 수행하기 위한 표준 데이터베이스(410)를 포함한다. 즉, 영상정보 학습부(100)는 데이터베이스(400)로부터 표준 데이터베이스(410)를 입력받아 깊이 정보와 세그멘테이션 영상 정보를 추정한다.
- [0042] 이때, 표준 데이터베이스(410)는 RGB 영상 정보와 깊이 정보 및 세그멘테이션 정보를 포함한다. 또한, 표준 데이터베이스(410)는 NYU Depth Dataset V2가 사용될 수 있다.
- [0043] 또한, 데이터베이스(400)는 판단부(300)의 판단 기준이 되는 기준 수렴값을 저장한다. 또한, 데이터베이스(400)는 영상정보 학습부(100)에서 깊이 정보 및 세그멘테이션 정보를 추정하는 생성자(Generator) $G_{x \rightarrow z}$ 와, 생성자(Generator) $G_{x \rightarrow y}$ 데이터를 저장한다.
- [0044] 영상 입력부(500)는 RGB 영상을 입력받는다. 또한, 영상정보 추정부(600)는 학습과정이 완료된 상기 생성자(Generator) $G_{x \rightarrow z}$ 와, 생성자(Generator) $G_{x \rightarrow y}$ 를 이용하여 영상 입력부(500)에서 입력받은 RGB 영상에 대한 깊이 정보 또는 세그멘테이션 정보를 추정한다.
- [0045] 도 3은 본 발명의 실시 예에 따른 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 방법을 나타내는 개념도이고, 도 4는 본 발명의 실시 예에 따른 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 방법을 나타내는 흐름도이며, 도 5는 본 발명의 실시 예에 따른 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 방법을 나타내는 순서도이다.
- [0046] 본 발명의 실시 예에 따른 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 방법은 도 4에서 도시된 바와 같이 학습단계와 실행단계를 포함한다. 상기 학습단계에서는 세그멘테이션과 이를 이용한 깊이 추정 방법을 학습한다. 또한, 상기 학습단계에서는 목적 함수와 판별 확률을 계산하여 학습을 조정한다.
- [0047] 상기 실행단계에서는 학습단계의 학습결과를 토대로 RGB 영상정보만을 이용하여 깊이 정보를 추정한다. 이때, 실행단계는 상기 학습단계에서 깊이 정보를 학습하던 과정에서 사용했던 생성자(Generator)를 사용하여 깊이 정보를 추정한다.
- [0048] 본 발명의 실시 예에 따른 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 방법의 학습단계는 표준 데이터베이스(410)의 입력 RGB 영상 X에 대하여 생성자(Generator) $G_{x \rightarrow y}$ 와 생성자(Generator) $G_{x \rightarrow z}$ 를 이용하여 깊이 정보와 세그멘테이션 영상 정보를 생성하는 단계(S10)와, 생성된 깊이 정보와 세그멘테이션 영상 정보를 이용하여 RGB 영상을 복원하는 단계(S20) 및 생성된 깊이 정보와 세그멘테이션 영상 정보 및 복원된 RGB 영상을 표준 데이터베이스(410)와 각각 판별하여 비교하고, 각각에 대한 손실(Loss) 및 판별 확률을 계산하는 단계(S30)를 포함할 수 있다.
- [0049] 또한, 계산된 결과값을 토대로 각각의 손실 및 판별자(Discriminator)의 판별 확률 값이 미리 설정된 기준 수렴값을 만족하는지 판단하는 단계(S40)와, 판단 결과 손실 및 판별 확률 값이 미리 설정된 기준 수렴값을 만족하지 않는 경우 각각의 손실 및 판별자의 판별 확률 값이 미리 설정된 기준 수렴값에 수렴되도록 학습을 조정하고, 상기 (S10) 단계 내지 (S40) 단계를 반복 수행하는 단계(S50)를 포함할 수 있다.
- [0050] 도 6은 본 발명의 실시 예에 따른 사이클 GAN과 세그멘테이션을 사용하여 단일 영상의 깊이 정보를 추정하는 방법의 동작 순서를 나타내는 도면이고, 도 7은 본 발명의 실시 예에 따른 세그멘테이션 추정 과정을 나타내는 도면이며, 도 8은 본 발명의 실시 예에 따른 깊이 추정 과정을 나타내는 도면이다.
- [0051] 상기 깊이 정보와 세그멘테이션 영상 정보를 생성하는 단계(S10)는 표준 데이터베이스(410)의 입력 RGB 영상을

토대로 깊이 정보와 세그멘테이션 정보를 추정하고, 목적 함수를 통해 학습률을 계산한다. 또한, 각 도메인(Domain)에 따른 사이클 일관성 손실(Cycle-Consistency Loss)들의 결합이 깊이 정보 추정의 성능에 미치는 영향을 평가할 수 있다.

[0052] 도 7 및 도 8에서 도시된 바와 같이 상기 깊이 정보와 세그멘테이션 영상 정보를 생성하는 단계(S10)에서 세그멘테이션 정보를 추정하는 세그멘테이션 네트워크 구조와, 깊이 정보를 추정하는 깊이 네트워크 구조는 동일한 구조를 가지며, 각 생성자(Generator)와 판별자(Discriminator)의 역할만 변경된다.

[0053] 상기 2개 네트워크의 동작 순서는 다음과 같다. 먼저, 표준 데이터베이스(410)의 세그멘테이션 정보를 통해 깊이 정보에 대한 힌트를 제공할 수 있다. 또한, RGB 영상 정보에서 세그멘테이션 정보로 변환하기 위하여 생성자(Generator) $G_{X \rightarrow Z}$ 에 의해 입력 RGB 영상 X에 대한 세그멘테이션 정보를 획득하여 해당 정보를 깊이 정보 추정에 활용할 수 있다.

[0054] 또한, 도 8과 같이 마찬가지로 생성자(Generator) $G_{X \rightarrow Y}$ 에 의해 입력 RGB 영상 X에 대한 깊이 정보를 획득하여 해당 정보를 세그멘테이션 정보 추정에 활용할 수 있다. 이때, 2개 네트워크의 생성자는 도 4와 같이 상기 손실(Loss) 및 판별 확률을 계산하는 단계(S30)에서 피드백 되어 RGB 영상을 통해 깊이 정보 및 세그멘테이션 정보를 추정할 수 있도록 변환된다.

[0055] 상기 손실(Loss) 및 판별 확률을 계산하는 단계(S30)는 사이클 GAN의 목적 함수를 통해 손실 및 판별 확률 결과값의 수치를 계산한다. 여기에서, 상기 목적 함수는 사이클 GAN(Generative Adversarial Network)의 적대적인 손실 함수(Adversarial Loss Function)와 사이클 일관성 손실 함수(Cycle-Consistency Loss Function)로 구성될 수 있다.

[0056] 또한, 상기 적대적인 손실 함수(Adversarial Loss Function)는 생성자(Generator)와 판별자(Discriminator)의 미니맥스(Minimax) 결과에 따라 학습을 진행한다. 상기 적대적인 손실(Adversarial Loss)에서 생성자(Generator)는 데이터의 표준 분포를 모방하고, 판별자(Discriminator)는 이에 따른 판별 확률을 계산한다.

[0057] 즉, 상기 적대적인 손실 함수(Adversarial Loss Function)는 상기 깊이 정보와 세그멘테이션 영상 정보를 생성하는 단계(S10)에서 계산된 목적 함수 중 적대적인 손실(Adversarial Loss)만을 계산한다. 해당 단계에서는 깊이 및 세그멘테이션 상호간의 교점이 존재하지 않으므로 서로 독립적으로 깊이 정보 및 세그멘테이션 정보 추정을 진행한다.

[0058] 따라서, 상기 손실(Loss) 및 판별 확률을 계산하는 단계(S30)는 상기 깊이 정보와 세그멘테이션 영상 정보를 생성하는 단계(S10) 및 RGB 영상을 복원하는 단계(S20)에서 추정된 깊이 영상과 세그멘테이션 영상 및 복원된 RGB 영상을 판별자(Discriminator)를 통해 판별하여 각각에 대한 확률을 계산할 수 있다.

[0059] 이때, 본 발명에 따른 네트워크의 목적 함수는 아래의 [수학식 1]과 같이 나타낼 수 있다.

[0060] [수학식 1]

$$\begin{aligned} L(G_{X \rightarrow Y}, G_{Y \rightarrow X}, D_X, D_Y) = & L_{GAN}(G_{X \rightarrow Y}, D_Y, X, Y) + L_{GAN}(G_{Y \rightarrow X}, D_X, Y, X) \\ & + L_{GAN}(G_{X \rightarrow Z}, D_Z, X, Z) + L_{GAN}(G_{Z \rightarrow X}, D_X, Z, X) \\ & + \lambda_1 L_{Cyc}(G_{X \rightarrow Y}, G_{Y \rightarrow X}) + \lambda_2 L_{Cyc}(G_{X \rightarrow Z}, G_{Z \rightarrow X}) \end{aligned}$$

[0061]

[0062] 여기에서, D는 각 입력에 대한 판별자의 판별 확률을 나타내고, G는 입력에 대한 생성자의 데이터 공간 매핑을 나타내며, λ 는 가중치 부여에 사용되는 하이퍼 파라미터를 나타낸다.

[0063] 또한, 상기 [수학식 1]은 사이클 GAN(Generative Adversarial Network)의 적대적인 손실 함수(Adversarial Loss Function)와 사이클 일관성 손실 함수(Cycle-Consistency Loss Function)로 구성된다.

[0064] 상기 적대적인 손실 함수(Adversarial Loss Function)는 RGB 영상 정보에서 세그멘테이션 정보와 깊이 정보를 추정하는 과정에서 연산 가능하며, 아래의 [수학식 2] 및 [수학식 3]과 같이 나타낼 수 있다.

[0065] [수학식 2]

$$L_{GAN}(G_{X \rightarrow Y}, D_Y, X, Y) = E_{y \sim P_{data}(y)}[\log D_Y(y)] + E_{x \sim P_{data}(x)}[\log(1 - D_Y(G_{X \rightarrow Y}(x)))]$$

[0067] [수학식 3]

$$L_{GAN}(G_{X \rightarrow Z}, D_Z, X, Z) = E_{z \sim P_{data}(z)}[\log D_Z(z)] + E_{x \sim P_{data}(x)}[\log(1 - D_Z(G_{X \rightarrow Z}(x)))]$$

[0069] 여기에서, E는 해당 분포에 대한 기댓값을 나타내고, P_i는 i에 대한 확률 분포를 나타낸다. 도 5에서 도시된 바와 같이 생성자(Generator) $G_{X \rightarrow Y}$ 가 생성한 페이크 데이터 $\bar{Y}$ 를 판별자 D_Y 가 판별하고, 생성자(Generator) $G_{X \rightarrow Z}$ 가 생성한 페이크 데이터 $\bar{Z}$ 를 판별자 D_Z 가 판별하면서 학습을 진행한다.

[0070] 상기 적대적인 손실 함수(Adversarial Loss Function)의 핵심은 GAN을 통해 생성된 분포를 실제 분포로 매핑하는 것이다. 따라서, 상기 적대적인 손실 함수(Adversarial Loss Function)는 생성자와 판별자의 미니맥스(Minimax) 결과에 따라 학습이 진행되며, 상기 생성자는 실제 분포와 완벽히 유사한 분포를 생성할 수 있다. 또한, 이에 따른 판별자의 판별 확률은 50%에 수렴하게 된다.

[0071] 도 9는 생성자(Generator)의 생성 분포 및 판별자(Discriminator)의 판별 확률을 나타내는 도면이다. 즉, 도 9는 학습단계에서 GAN의 학습이 진행됨에 따른 생성자의 생성 분포와 판별자의 판별 확률이 변동되는 과정을 나타내는 그래프이다.

[0072] 도 9에서 검은색 점선은 실제 데이터 분포(Discriminator distribution)를 나타내고, 녹색 실선은 생성자의 생성 분포(Generative distribution)를 나타내며, 파란 점선은 판별자의 판별 확률(Discriminator distribution)을 나타낸다. 여기에서, 두 분포의 거리가 좁다는 것은 해당 분포가 매우 유사하다는 것을 의미하며, 이는 판별자가 쉽게 판별하지 못함을 나타낸다.

[0073] 따라서, 판별자의 판별 확률은 낮게(min), 생성자의 생성 분포 유사도를 높게(max) 학습을 진행하여 생성자의 생성 분포가 실제 데이터 분포와 매우 유사하도록 학습을 진행한다.

[0074] 또한, 재건 손실(Reconstruction Loss)을 추가함으로써 종래의 CNN 기반 학습 방법에서 사용하던 손실(Loss)을 결합하여 생성자의 생성 분포가 목표(Target)의 표준 분포를 학습할 수 있도록 유도한다. 상기 재건 손실(Reconstruction Loss)은 아래의 [수학식 4]와 같이 나타낼 수 있다.

[0075] [수학식 4]

$$L_{Rec}(G_{X \rightarrow Y}) = E_{x \sim P_{data}(x), y \sim P_{data}(y)}[\|y - G_{X \rightarrow Y}(x)\|_1]$$

[0077] 상기 손실(Loss) 및 판별 확률을 계산하는 단계(S30) 이후에, 동일한 입력 RGB 영상 X에 대하여 서로 다른 출력 값을 갖지만 후에 이를 다시 RGB 영상 $\bar{X}$ 로 복원하였을 때 원본 RGB 영상 X와 복원된 RGB 영상 $\bar{X}$ 를 비교하여 원본 RGB 영상 X의 형상을 유지하면서 깊이 정보와 세그멘테이션 영상을 생성할 수 있도록 유도하는 사이클 일관성 손실(Cycle-Consistency Loss) 단계(S31)를 더 포함할 수 있다.

[0078] 상기 사이클 일관성 손실(Cycle-Consistency Loss) 계산 단계(S31)는 2개의 목적 함수와 별개로 추정된 두 영상 정보를 다시 원래의 영상정보로 복원하였을 때의 유사도를 평가한다. 상기 사이클 일관성 손실(Cycle-Consistency Loss) 계산 단계(S31)는 생성자가 각 도메인(Domain)의 형태를 유지한 채 도메인으로의 변환을 시도하게끔 유도하는 역할을 한다.

[0079] 본 발명의 실시 예에 따른 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 방법은 세그멘테이션(Segmentation)과 깊이(Depth)의 2가지 도메인(Domain)으로 구성되는 사이클 GAN을 사용할 수 있다. 따라서, 상기 사이클 일관성 손실(Cycle-Consistency Loss)은 2가지 손실(Loss)의 합으로 구성되는 아래의 [수학식 5]와 같이 나타낼 수 있다.

[0080] [수학식 5]

$$L_{cyc}(G_{X \rightarrow Y}, G_{Y \rightarrow X}) + L_{cyc}(G_{X \rightarrow Z}, G_{Z \rightarrow X}) = \\ E_{x \sim P_{data}(x)} [\| G_{Y \rightarrow X}(G_{X \rightarrow Y}(x)) - x \|_1] + E_{y \sim P_{data}(y)} [\| G_{X \rightarrow Y}(G_{Y \rightarrow X}(y)) - y \|_1] \\ + E_{x \sim P_{data}(x)} [\| G_{Z \rightarrow X}(G_{X \rightarrow Z}(x)) - x \|_1] + E_{z \sim P_{data}(z)} [\| G_{X \rightarrow Z}(G_{Z \rightarrow X}(z)) - z \|_1]$$

[0081]

[0082] 즉, 복원을 통해 생긴 오차 값을 사이클 일관성 손실(Cycle-Consistency Loss)로 설정한다. 따라서, 복원이 잘 이루어진다면 손실(Loss) 함수는 낮아진다.

[0083] 도 10은 본 발명의 실시 예에 따른 사이클 GAN 모델의 사이클 일관성 손실(Cycle-Consistency Loss)을 나타내는 도면이다. 도 10에서 도시된 바와 같이 상기 사이클 일관성 손실(Cycle-Consistency Loss)을 살펴보면 RGB 영상으로 다시 복원이 진행될 때, 생성자는 깊이 정보로부터의 복원뿐만 아니라 세그멘테이션 정보로부터의 복원까지 고려해야함을 나타낸다.

[0084] 따라서, 본 발명의 실시 예에 따른 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 방법은 깊이 정보만을 고려하여 복원하는 종래의 제약 방식에 세그멘테이션 정보를 추가함으로써 더욱 객체 분류화된 깊이 정보의 생성이 가능하다. 마찬가지로, 세그멘테이션 정보의 복원을 수행할 경우에도 깊이 정보를 고려한 세그멘테이션 정보의 복원이 수행됨으로써 배경 분리 등에 대한 상승효과를 나타낼 수 있다.

[0085] 또한, 상기 사이클 일관성 손실(Cycle-Consistency Loss)에 L1 손실(Loss)을 통해 모델 가중치의 L1 놈(norm)(가중치 각 요소 절대값의 합)에 대해 패널티를 부과하여 정규화를 진행한다. L1 손실(Loss)은 L2 손실(Loss)에 비해 상대적으로 강건(Robust)하고, 불안정한 솔루션 문제(Unstable solution problem)에 강인하다.

[0086] 상기 L1 손실(Loss)은 아래의 [수학식 6]과 같이 나타낼 수 있다.

[0087] [수학식 6]

$$d_1(p, q) = \| p - q \|_1 = \sum_{i=1}^n |p_i - q_i|, \text{ where } p = (p_1, p_2, \dots, p_n), q = (q_1, q_2, \dots, q_n)$$

[0088]

[0089] 이와 같이, 본 발명의 실시 예에 따른 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 방법은 상기 학습단계를 통해 적대적인 손실(Adversarial Loss)과 사이클 일관성 손실(Cycle-Consistency Loss)을 계산하고, 이를 기반으로 학습이 완료된 후 최종적으로 페이크 깊이(Fake Depth) 정보 $\bar{Y}$ 를 추정할 수 있다.

[0090] 즉, 상기 학습단계를 통해 실제 깊이 정보와 미리 설정된 유사도를 가지고 유사하도록 학습이 완료되면 사이클 GAN의 학습 과정을 종료한다.

[0091] 도 11은 본 발명의 실시 예에 따른 실행단계의 깊이 정보 추정 단계를 나타내는 도면이고, 도 12는 본 발명의 실시 예에 따른 실행단계의 세그멘테이션 정보 추정 단계를 나타내는 도면이다.

[0092] 본 발명의 실시 예에 따른 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 방법의 실행단계는 상기 (S10) 단계 내지 (S50) 단계를 통해 생성된 생성자(Generator) $G_{x \rightarrow y}$ 를 이용하여 RGB 데이터의 입력 RGB 영상에 대한 깊이 정보를 추정하는 단계(S60)와, 상기 (S10) 단계 내지 (S50) 단계를 통해 생성된 생성자(Generator) $G_{x \rightarrow z}$ 를 이용하여 상기 RGB 데이터의 입력 RGB 영상에 대한 세그멘테이션 영상 정보를 추정하는 단계(S70)를 포함할 수 있다.

[0093] 상기 깊이 정보를 추정하는 단계(S60)는 도 11과 같이 학습단계에서 생성된 생성자(Generator) $G_{x \rightarrow y}$ 를 사용하여 깊이 정보를 추정(Estimation)한다.

[0094] 즉, 상기 깊이 정보를 추정하는 단계(S60)는 실제 깊이 영상과 유사한 RGB 영상의 깊이 영상을 생성하기 위해 상기 학습단계에서 생성된 생성자(Generator) $G_{x \rightarrow y}$ 를 사용하여 깊이 정보 추정(Depth estimation)을 수행한다.

- [0095] 또한, 상기 세그멘테이션 영상 정보를 추정하는 단계(S70)는 도 12와 같이 학습단계에서 생성된 생성자(Generator) $G_{x \rightarrow z}$ 를 사용하여 세그멘테이션 정보를 추정한다. 즉, 상기 세그멘테이션 영상 정보를 추정하는 단계(S70)는 상기 학습단계에서 생성된 생성자(Generator) $G_{x \rightarrow z}$ 를 사용하여 RGB 영상을 세그멘테이션 정보로 변환한다.
- [0096] 도 13a 및 도 13b는 깊이 정보 추정 과정에서 세그멘테이션 과정을 사용하기 전과 후를 비교하여 나타내는 도면이다. 즉, 도 13a는 깊이 정보 추정 과정에서 세그멘테이션 과정을 사용하지 않고 추정된 영상의 불확실성을 나타내는 도면이고, 도 13b는 세그멘테이션 과정을 추가하여 입력 영상에 대한 깊이 정보를 추정한 결과를 나타내는 도면이다.
- [0097] 종래의 깊이 정보 추정 과정에 세그멘테이션 과정을 추가하는 이유는 도 13b에서 도시된 바와 같이 입력 영상에 대한 깊이 정보 추정 결과에 대하여 깊이 정보의 불확실성 문제를 해결하기 위한 것이다.
- [0098] 종래의 깊이 정보 추정 과정을 통해 완벽하게 깊이 정보를 추정할 수는 없으므로 도 13b와 같이 세그멘테이션 추정 과정을 통해 생성되는 여러 변수를 복합적으로 적용하여 깊이 정보의 불확실성을 줄일 수 있다.
- [0099] 이와 같이 본 발명의 실시 예에 따른 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 방법 및 시스템(10)은 여러 변수를 통해 성능을 향상시키는 멀티태스킹 학습(multi-task learning) 기법을 통해 깊이 정보 추정 결과를 향상시킬 수 있다.
- [0100] 도 14는 본 발명의 실시 예에 따른 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 방법의 평가 절차를 나타내는 도면이다. 도 14와 같이 본 발명의 실시 예에 따른 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 방법의 신뢰성 평가를 위해 공개된 표준 데이터베이스인 NYU Depth Dataset V2를 이용하여 평가를 수행한다.
- [0101] 상기 NYU Depth Dataset V2 데이터베이스는 마이크로소프트사의 Kinect v1 모델을 사용하여 촬영된 다양한 실내 장면 중심의 비디오 시퀀스 데이터를 제공한다. 또한, 상기 NYU Depth Dataset V2 데이터베이스는 Labeled Dataset을 통해 RGB 영상에 대한 깊이 정보와 세그멘테이션 정보를 제공한다.
- [0102] 도 14의 Segmentation 및 Depth 추정 단계(S110)는 표준 데이터베이스(410)의 입력 RGB 영상 X에 대하여 생성자(Generator) $G_{x \rightarrow r}$ 와 생성자(Generator) $G_{x \rightarrow z}$ 을 이용하여 깊이 정보와 세그멘테이션 영상 정보를 생성하는 단계(S10)를 나타낸다.
- [0103] 즉, 상기 Segmentation 및 Depth 추정 단계(S110)에서는 깊이 정보와 세그멘테이션 정보를 추정함과 동시에 목적 함수를 통해 학습률을 계산한다. 또한, 본 발명의 각 도메인(Domain)에 따른 사이클 일관성 손실(Cycle-Consistency Loss)들의 결합이 성능에 미치는 영향을 평가한다.
- [0104] 또한, 도 14의 Adversarial Loss 계산 단계(S130)는 생성된 깊이 정보와 세그멘테이션 영상 정보 및 복원된 RGB 영상을 표준 데이터베이스(410)와 각각 판별하여 비교하고, 각각에 대한 손실(Loss) 및 판별 확률을 계산하는 단계(S30)를 나타낸다.
- [0105] 즉, 상기 Adversarial Loss 계산 단계(S130)에서는 상기 Segmentation 및 Depth 추정 단계(S110)에서 계산된 목적 함수 중 적대적인 손실(Adversarial Loss)만을 계산한다. 또한, 상기 Adversarial Loss 계산 단계(S130)에서는 깊이 정보 및 세그멘테이션 정보 상호 간의 교점이 존재하지 않으므로 서로 독립적으로 깊이 및 세그멘테이션 추정을 진행한다.
- [0106] 이때, 상기 Segmentation 및 Depth 추정 단계(S110) 이후에 생성된 깊이 정보와 세그멘테이션 영상 정보를 이용하여 RGB 영상을 복원하는 단계(S120)를 더 포함할 수 있다.
- [0107] 또한, 도 14의 Cycle-Consistency Loss 계산 단계(S131)는 동일한 입력 RGB 영상 X에 대하여 서로 다른 출력값을 갖지만 후에 이를 다시 RGB 영상 $\bar{X}$ 로 복원하였을 때 원본 RGB 영상 X와 복원된 RGB 영상 $\bar{X}$ 를 비교하여 원본 RGB 영상 X의 형상을 유지하면서 깊이 정보와 세그멘테이션 영상을 생성할 수 있도록 유도하는 사이클 일관성 손실(Cycle-Consistency Loss) 단계(S31)를 나타낸다.
- [0108] 즉, 상기 Cycle-Consistency Loss 계산 단계(S131)에서는 복원된 RGB 영상을 원본 RGB 영상과 비교하여 사이클 일관성 손실(Cycle-Consistency Loss)을 계산한 뒤 각 깊이 정보 및 세그멘테이션 정보의 생성자에 패널티를 부여한다. 이러한 과정을 통해 각 생성자는 깊이 정보와 세그멘테이션 정보의 복원까지 고려하여 깊이 및 세그멘

테이션 추정을 진행하게 된다.

또한, 도 14의 Depth 및 Segmentation 평가 단계(S140)는 계산된 결과값을 토대로 각각의 손실 및 판별자(Discriminator)의 판별 확률 값이 미리 설정된 기준 수렴값을 만족하는지 판단하는 단계(S40)를 나타낸다.

상기 Depth 및 Segmentation 평가 단계(S140)에서는 생성된 결과물에 대하여 수치적 에러를 측정하는 평균 제곱근 오차(RMSE)의 변형인 RMSLE를 측정하여 생성된 깊이 정보를 평가한다. 상기 RMSLE는 아래의 [수학식 7]과 같이 나타낼 수 있다.

[수학식 7]

$$RMSLE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\log(p_i + 1) - (\log(a_i + 1)))^2}$$

여기에서, 상기 RMSLE를 계산하기 위해 필요한 P_i 와 a_i 는 0부터 1 사이의 값으로 정규화되어 입력된다. 상기 RMSLE 비용함수는 주로 과대평가 된 항목보다 과소평가 된 항목에 패널티를 주기 위해 사용되며 정답에 대한 오류를 숫자로 나타낸 값으로 값이 클수록 오차가 크다.

아래의 [표 1]은 상기 NYU Depth Dataset V2 데이터베이스를 토대로 본 발명의 실시 예에 따른 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 방법을 다른 기법들과 비교 평가한 결과를 나타낸다.

[표 1] 본 발명의 실시 예에 따른 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 방법과 다른 기법들의 NYU Depth Dataset V2에 대한 비교 결과

Method	RMSLE
Isola	0.283
Pilzer	0.240
AdaDepth	0.225
본 발명에서 제안하는 기법	0.220

상기 [표 1]에 나타난 바와 같이 본 발명의 실시 예에 따른 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 방법의 RMSLE 값이 0.220으로 다른 기법들보다 낮은 수치를 나타낸다. 즉, 상기 RMSLE 값은 수치가 낮을수록 우수한 깊이 추정 방법을 나타내므로, 다른 기법들보다 본 발명의 실시 예에 따른 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 방법이 높은 유사도를 나타내는 것을 확인할 수 있다.

이와 같이 본 발명의 실시 예에 따른 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 방법 및 시스템(10)은 종래의 3차원 정보를 얻기 위해 특수 장비나 센서 등을 사용해야 했던 방식에 비해 단일 카메라만을 사용하여 3차원 정보를 생성할 수 있어 보다 저렴하며, 확장성이 높고 소형화에 유리하며, 무엇보다도 단일 영상으로 구성된 자료를 구하기 쉽다는 장점이 있다.

또한, 본 발명의 실시 예에 따른 사이클 GAN과 세그멘테이션을 사용한 깊이 추정 방법 및 시스템(10)은 깊이 정보를 추정함과 동시에 깊이 정보의 정밀성을 향상할 수 있다. 또한, 종래의 입력 영상에 대한 깊이 정보를 추정하는 과정에서 발생하는 데이터 불균형 문제를 시각적으로 표시하고, 상대적으로 큰 특징에 묻혀 소실되는 작은 특징들을 세그멘테이션을 도입하여 해결할 수 있다. 즉, 입력 영상의 깊이 정보 추정 결과에 대하여 깊이 정보의 불확실성 문제를 해결할 수 있다.

이상으로 본 발명에 관한 바람직한 실시 예를 설명하였으나, 본 발명은 상기 실시 예에 한정되지 아니하며, 본 발명의 실시 예로부터 당해 발명이 속하는 기술분야에서 통상의 지식을 가진 자에 의한 용이하게 변경되어 균등하다고 인정되는 범위의 모든 변경을 포함한다.

부호의 설명

10 : 깊이 추정 시스템 100 : 영상정보 학습부
200 : 역산부 300 : 판단부

400 : 데이터베이스

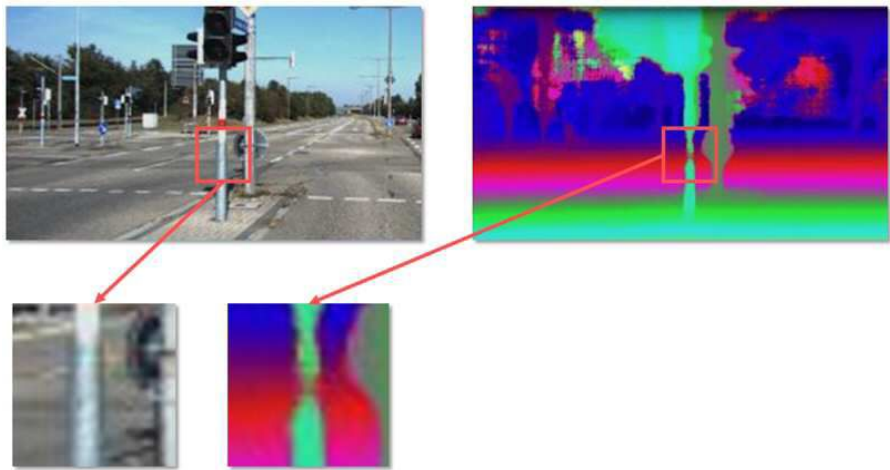
410 : 표준 데이터베이스

500 : 영상 입력부

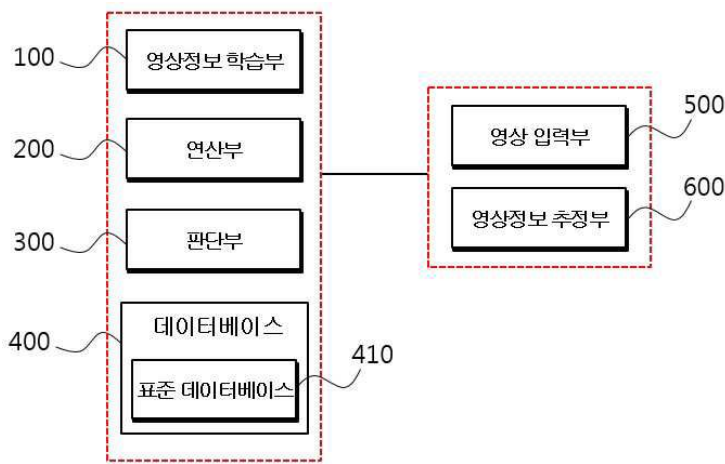
600 : 영상정보 추정부

도면

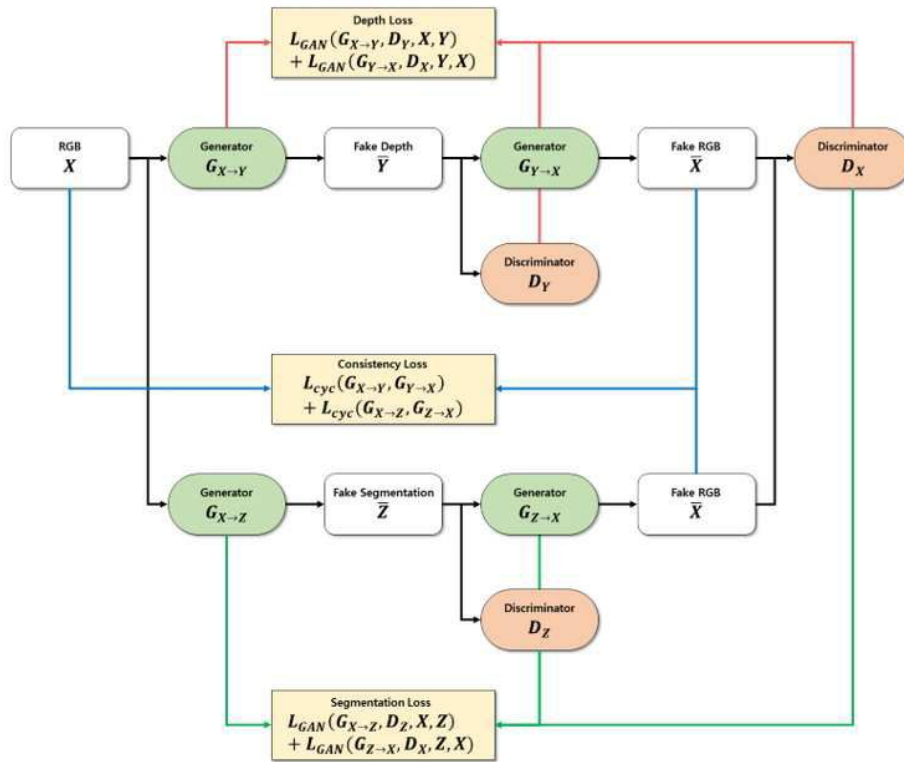
도면1



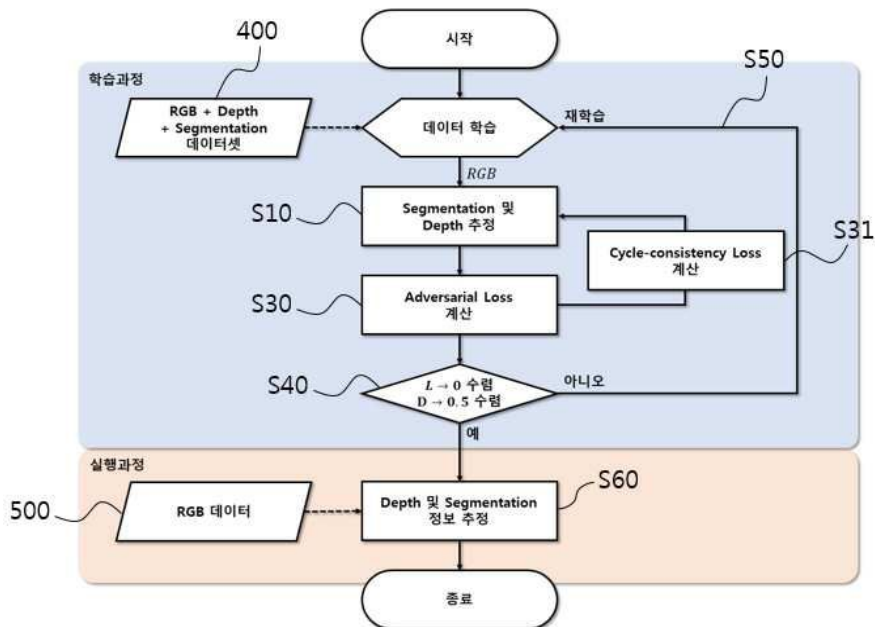
도면2



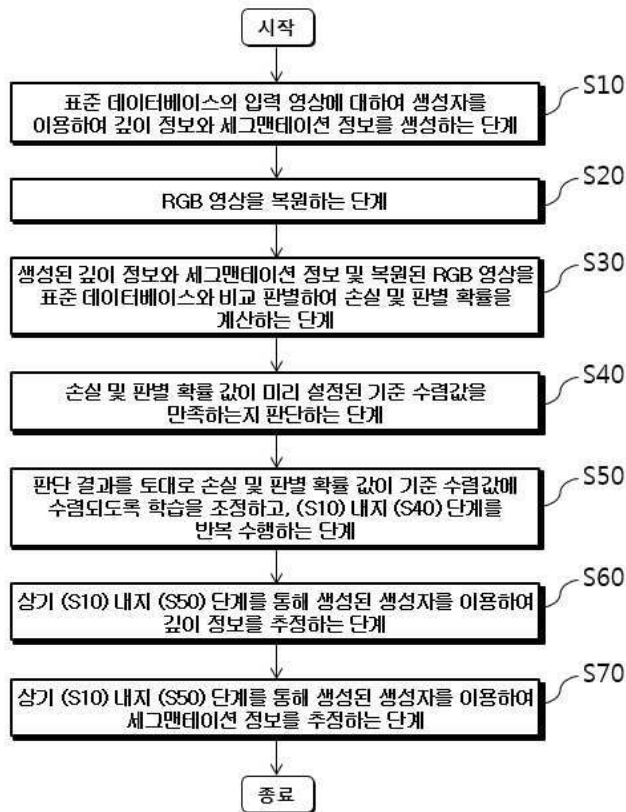
도면3



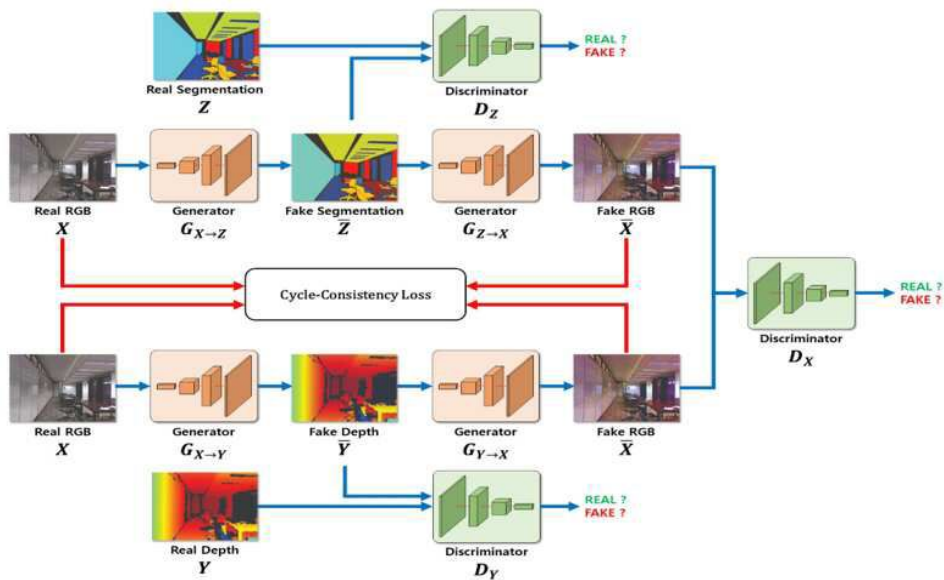
도면4



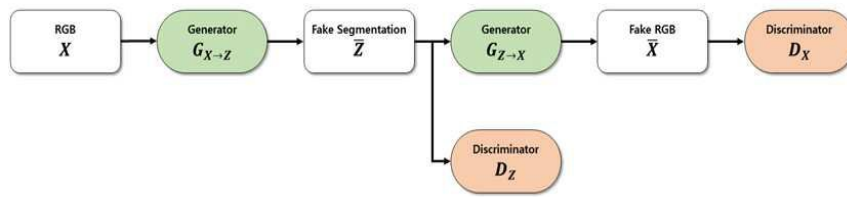
도면5



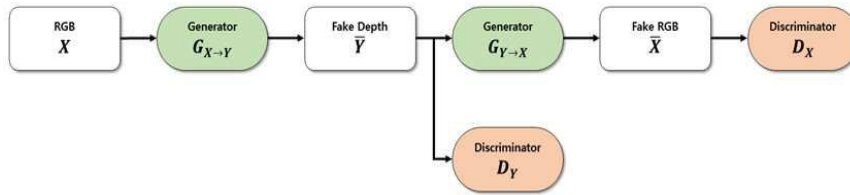
도면6



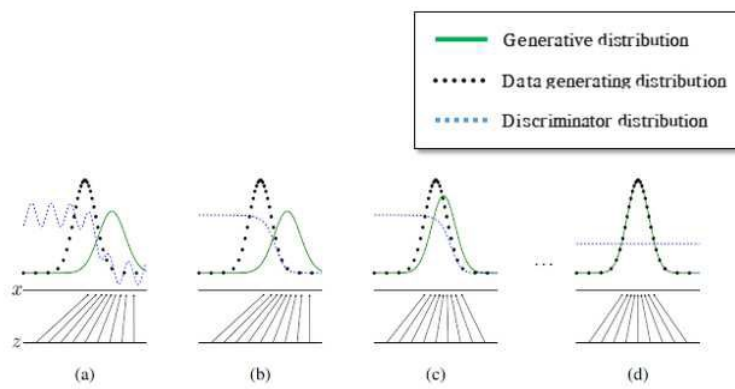
도면7



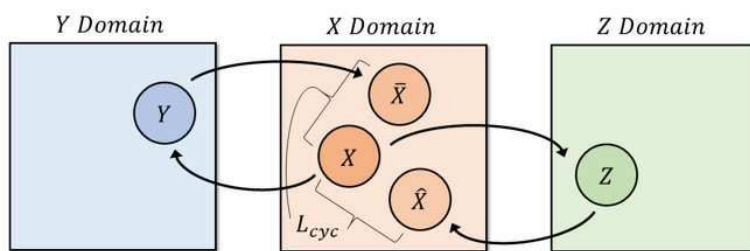
도면8



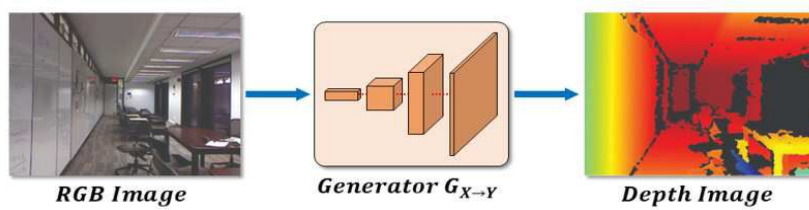
도면9



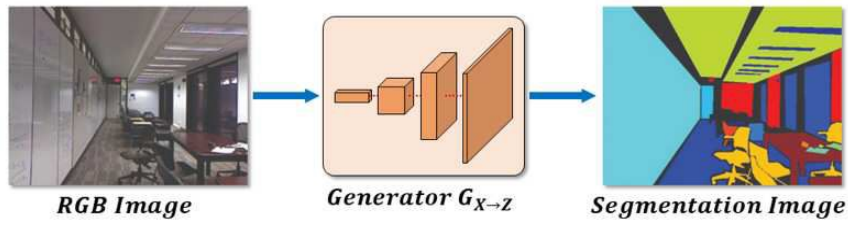
도면10



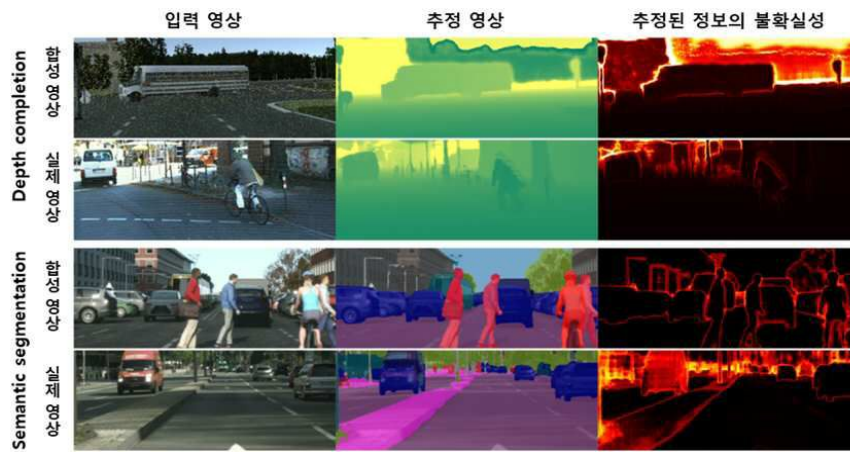
도면11



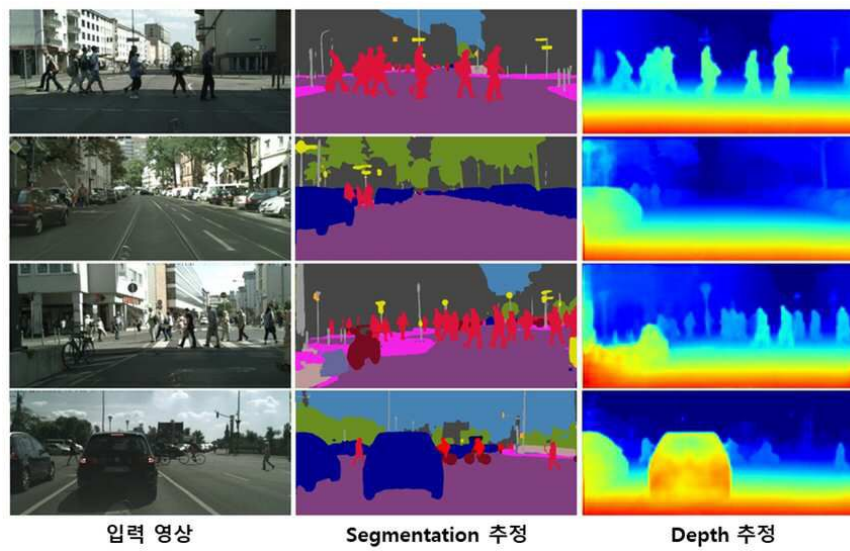
도면12



도면13a



도면13b



도면14

