

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)(11) 공개번호 10-2022-0057060
(43) 공개일자 2022년05월09일

(51) 국제특허분류(Int. Cl.)

G06T 17/05 (2011.01) G01S 17/894 (2020.01)

G06N 3/08 (2006.01) G06T 15/00 (2006.01)

G06T 3/40 (2006.01) G06T 7/11 (2017.01)

G06T 7/593 (2017.01)

(52) CPC특허분류

G06T 17/05 (2013.01)

G01S 17/894 (2022.01)

(21) 출원번호 10-2020-0141887

(22) 출원일자 2020년10월29일

심사청구일자 2020년10월29일

(71) 출원인

연세대학교 산학협력단

서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)

(72) 발명자

이상윤

서울특별시 서초구 청계산로7길 43, 504동 1302호(신원동, 서초포레스타5단지)

이준협

제주특별자치도 제주시 흥운길 57, 602호(용담이동, 용연엔타운아파트)

(뒷면에 계속)

(74) 대리인

민영준

전체 청구항 수 : 총 20 항

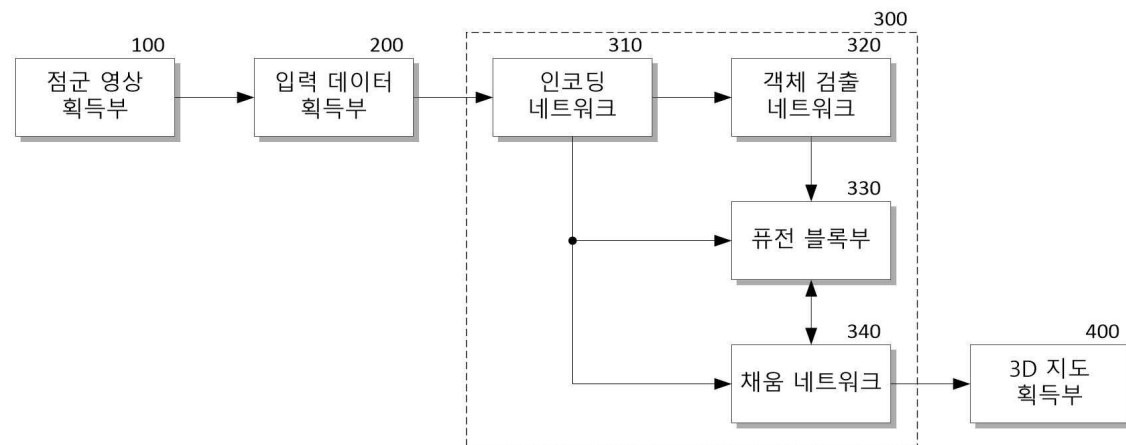
(54) 발명의 명칭 라이다 깊이 이미지 채움 장치 및 방법

(57) 요약

본 발명은 라이다를 이용하여 획득된 3차원 점군 이미지가 인가되면, 기지정된 방식으로 2차원 이미지로 변환하여 입력 데이터를 획득하는 입력 데이터 획득부, 미리 학습된 인공 신경망으로 구현되어, 입력 데이터를 학습된 방식에 따라 인코딩하여 특징맵을 추출하는 과정을 반복하여 획득되는 다수의 특징맵 중 최종 인코딩된 특징맵인

(뒷면에 계속)

대표도 - 도3



인코딩 특징맵을 학습된 방식에 따라 디코딩하여 동적 객체 영역을 나타내는 이진 마스크를 획득하고, 확률적 구조 유사도에 따라 다수의 특징맵 중 대응하는 특징맵에서 이진 마스크에 대응하는 영역을 이전 획득된 채움 특징맵의 영역으로 대체하여 관심맵을 획득하며, 획득된 관심맵을 미리 학습된 방식으로 디코딩하여 채움 특징맵을 반복적으로 획득하는 배경 채움 네트워크 및 최종 획득된 채움 특징맵을 채움 이미지로 인가받고, 기지정된 방식으로 3차원 이미지로 변환하는 3D 지도 획득부를 포함하여, 동적 객체 영역이 자연스러운 배경으로 대체되어 자율 주행에 이용 가능한 수준의 고품질의 3D 지도를 획득할 수 있는 라이다 깊이 이미지 채움 장치 및 방법을 제공할 수 있다.

(52) CPC특허분류

G06N 3/08 (2013.01)
G06T 15/00 (2013.01)
G06T 3/40 (2013.01)
G06T 7/11 (2017.01)
G06T 7/593 (2017.01)
G06T 2207/10028 (2013.01)
G06T 2207/20081 (2013.01)
G06T 2207/20084 (2013.01)

(72) 발명자

황상원

서울특별시 강서구 강서로56나길 110, 104동 702호(등촌동, 부영아파트)

김우진

서울특별시 강남구 남부순환로 2803, 105동 802호(도곡동, 삼성래미안아파트)

이도균

서울특별시 성동구 금호로 140, 108동 1804호(금호동1가, 이편한세상금호파크힐스)

우성민

서울특별시 서대문구 연세로11길 29, 201호(창천동)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711103201
과제번호	2016-0-00197-005
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원(한국연구재단부설)
연구사업명	정보통신방송연구개발사업
연구과제명	스마트카 다중 센서와 딥러닝을 이용한 초정밀 내추릴 3D 뷰 생성 기술 개발 (창조
씨앗형 2단계)(4/5)	
기 여 율	1/1
과제수행기관명	연세대학교 산학협력단
연구기간	2020.01.01 ~ 2020.12.31

명세서

청구범위

청구항 1

라이다를 이용하여 획득된 3차원 점군 이미지가 인가되면, 기지정된 방식으로 2차원 이미지로 변환하여 입력 데이터를 획득하는 입력 데이터 획득부;

미리 학습된 인공 신경망으로 구현되어, 상기 입력 데이터를 학습된 방식에 따라 인코딩하여 특징맵을 추출하는 과정을 반복하여 획득되는 다수의 특징맵 중 최종 인코딩된 특징맵인 인코딩 특징맵을 학습된 방식에 따라 디코딩하여 동적 객체 영역을 나타내는 이진 마스크를 획득하고, 확률적 구조 유사도에 따라 상기 다수의 특징맵 중 대응하는 특징맵에서 상기 이진 마스크에 대응하는 영역을 이전 획득된 채움 특징맵의 영역으로 대체하여 관심맵을 획득하며, 획득된 관심맵을 미리 학습된 방식으로 디코딩하여 채움 특징맵을 반복적으로 획득하는 배경 채움 네트워크; 및

최종 획득된 채움 특징맵을 채움 이미지로 인가받고, 기지정된 방식으로 3차원 이미지로 변환하는 3D 지도 획득부를 포함하는 라이다 깊이 이미지 채움 장치.

청구항 2

제1항에 있어서, 상기 배경 채움 네트워크는

각각 입력 데이터 또는 이전 배치된 컨볼루션 레이어에서 출력되는 특징맵을 인가받아 인코딩하여 특징맵을 출력하는 다수의 컨볼루션 레이어를 포함하여 인코딩 네트워크;

상기 인코딩 특징맵을 인가받아 학습된 방식에 따라 디코딩하여, 상기 입력 데이터에서 학습에 의해 지정된 동적 객체가 위치하는 영역을 나타내는 상기 이진 마스크를 획득하는 객체 검출 네트워크;

상기 인코딩 특징맵 또는 이전 획득된 관심맵을 인가받아 학습된 방식으로 디코딩하여 채움 특징맵을 획득하는 채움 네트워크; 및

상기 다수의 특징맵 중 대응하는 특징맵에서 상기 이진 마스크에 대응하는 영역을 이전 획득된 채움 특징맵의 영역으로 대체하여 퓨전 특징맵을 획득하고, 대응하는 특징맵과 획득된 퓨전 특징맵 사이의 확률적 구조 유사도를 계산하여 상기 퓨전 특징맵에 기지정된 방식으로 가중하여 상기 관심맵을 획득하는 다수의 퓨전 블록을 구비하는 퓨전 블록부를 포함하는 라이다 깊이 이미지 채움 장치.

청구항 3

제2항에 있어서, 상기 객체 검출 네트워크는

상기 이진 마스크를 상기 인코딩 네트워크에서 획득되는 다수의 특징맵 각각에 대응하는 크기로 스케일링하여 다수의 스케일 이진 마스크를 획득하는 라이다 깊이 이미지 채움 장치.

청구항 4

제3항에 있어서, 상기 다수의 퓨전 블록 각각은

상기 다수의 특징맵 중 대응하는 특징맵과 다수의 스케일 이진 마스크 중 대응하는 스케일 이진 마스크 및 상기 채움 네트워크에서 이전 획득된 채움 특징맵을 인가받아, 상기 대응하는 특징맵에서 대응하는 스케일 이진 마스크에 의해 지정된 영역을 이전 획득된 채움 특징맵의 영역으로 대체하여 퓨전 특징맵을 획득하는 퓨전부;

SSIM(structural similarity index measurement)의 구조 유사도 평가 방식에 따라 대응하는 특징맵과 상기 퓨전 특징맵 사이의 구조 유사도를 판별하고, 확률적 분포를 추가 반영하여 확률적 구조 유사도를 계산하는 구조 유사도 계산부; 및

상기 퓨전 특징맵에 계산된 확률적 구조 유사도를 가중하여 상기 관심맵을 획득하는 관심맵 획득부를 포함하는 라이다 깊이 이미지 채움 장치.

청구항 5

제4항에 있어서, 상기 퓨전부는

상기 퓨전 특징맵(F_f^n)을 수학식

$$F_f^n = F_e^n \cdot (\mathbb{1} - \tilde{m}) + F_i^n \cdot \tilde{m}$$

(여기서 F_e^n 는 특징맵이고, \tilde{m} 는 스케일 이진 마스크이며, F_i^n 는 채움 특징맵이다.)

에 따라 획득하는 라이다 깊이 이미지 채움 장치.

청구항 6

제5항에 있어서, 상기 구조 유사도 계산부는

상기 구조 유사도(S)를 수학식

$$s(k, q) = \frac{\sigma_{kq} + c_3}{\sigma_k \sigma_q + c_3}$$

(여기서 k 는 특징맵(F_e^n)이고, q 는 퓨전 특징맵(F_f^n)이며, σ 는 분산이다. 그리고 c_3 는 기지정된 상수이다.)

에 따라 계산하고,

상기 확률적 구조 유사도를 수학식

$$S^n = \text{softmax}(s(k, q)^n)$$

에 따라 계산하는 라이다 깊이 이미지 채움 장치.

청구항 7

제6항에 있어서, 상기 관심맵 획득부는

상기 관심맵(C^n)을 수학식

$$C^n = (\mathbb{1} + \lambda_{atn} S^n) \cdot F_f^n, \quad \lambda_{atn} \in [0, 1]$$

(여기서 λ_{atn} 은 관심 가중치이다.)

에 따라 획득하는 라이다 깊이 이미지 채움 장치.

청구항 8

제1항에 있어서, 상기 입력 데이터 획득부는

3D 점군 이미지로부터 깊이값, x축 값, y축 값, z 축 값 및 세기 값에 따른 5개의 2차원 거리 이미지를 획득하고, 획득된 5개의 2차원 거리 이미지를 채널축 방향으로 결합하여 상기 입력 데이터를 획득하는 라이다 깊이 이미지 채움 장치.

청구항 9

제4항에 있어서, 상기 라이다 깊이 이미지 채움 장치는

학습 시에 인공 신경망으로 구현되는 상기 배경 채움 네트워크를 학습시키기 위한 학습부를 더 포함하고,

상기 학습부는

동적 객체 영역이 미리 설정되어 배경으로 채워진 학습 데이터 셋을 입력 데이터로 상기 배경 채움 네트워크에 인가하여, 상기 객체 검출 네트워크에서 출력되는 이진 마스크와 설정된 동적 객체 영역 사이의 차이로 발생하

는 마스크 손실과 상기 배경 채움 네트워크에서 출력되는 채움 특징맵 사이의 차이로 발생하는 채움 손실의 합으로 총 손실을 계산하고, 계산된 총 손실을 역전파하여 학습을 수행하는 라이다 깊이 이미지 채움 장치.

청구항 10

제4항에 있어서, 상기 라이다 깊이 이미지 채움 장치는

학습 시에 인공 신경망으로 구현되는 상기 배경 채움 네트워크를 학습시키기 위한 학습부를 더 포함하고,

상기 학습부는

상기 배경 채움 네트워크에서 출력되는 채움 특징맵이 리얼 이미지인지 페이크 이미지인지 판별하도록 학습되는 판별자를 더 포함하여, 상기 배경 채움 네트워크인 생성자와 상기 판별자를 생성적 적대 네트워크 방식으로 학습시키는 라이다 깊이 이미지 채움 장치.

청구항 11

라이다를 이용하여 획득된 3차원 점군 이미지가 인가되면, 기지정된 방식으로 2차원 이미지로 변환하여 입력 데이터를 획득하는 단계;

미리 학습된 인공 신경망을 이용하여, 상기 입력 데이터를 학습된 방식에 따라 인코딩하여 특징맵을 추출하는 과정을 반복하여 다수의 특징맵을 획득하는 단계;

미리 학습된 인공 신경망을 이용하여, 상기 다수의 특징맵 중 최종 인코딩된 특징맵인 인코딩 특징맵을 학습된 방식에 따라 디코딩하여 동적 객체 영역을 나타내는 이진 마스크를 획득하는 단계;

미리 학습된 인공 신경망을 이용하여, 확률적 구조 유사도에 따라 상기 다수의 특징맵 중 대응하는 특징맵에서 상기 이진 마스크에 대응하는 영역을 이전 획득된 채움 특징맵의 영역으로 대체하여 관심맵을 획득하고, 획득된 관심맵을 미리 학습된 방식으로 디코딩하여 채움 특징맵을 획득하는 단계; 및

최종 획득된 채움 특징맵을 채움 이미지로 인가받고, 기지정된 방식으로 3차원 이미지로 변환하는 단계를 포함하는 라이다 깊이 이미지 채움 방법.

청구항 12

제11항에 있어서, 상기 다수의 특징맵을 획득하는 단계는

다수의 컨볼루션 레이어가 포함된 인공 신경망인 인코딩 네트워크를 이용하여 입력 데이터 또는 이전 배치된 컨볼루션 레이어에서 출력되는 특징맵을 인가받아 인코딩하여 다수의 특징맵을 출력하는 라이다 깊이 이미지 채움 방법.

청구항 13

제12항에 있어서, 상기 이진 마스크를 획득하는 단계는

상기 인코딩 특징맵을 학습된 방식에 따라 디코딩하여, 상기 입력 데이터에서 학습에 의해 지정된 동적 객체가 위치하는 영역을 나타내는 상기 이진 마스크를 획득하는 단계; 및

상기 이진 마스크를 상기 다수의 특징맵 각각에 대응하는 크기로 스케일링하여 다수의 스케일 이진 마스크를 획득하는 단계를 포함하는 라이다 깊이 이미지 채움 방법.

청구항 14

제13항에 있어서, 상기 채움 특징맵을 획득하는 단계는

학습된 방식으로 인코딩 특징맵 또는 이전 획득된 관심맵을 디코딩하여 상기 채움 특징맵을 획득하는 단계; 및

상기 다수의 특징맵 중 대응하는 특징맵에서 상기 이진 마스크에 대응하는 영역을 이전 획득된 채움 특징맵의 영역으로 대체하여 퓨전 특징맵을 획득하는 단계; 및

대응하는 특징맵과 획득된 퓨전 특징맵 사이의 확률적 구조 유사도를 계산하여 상기 퓨전 특징맵에 기지정된 방식으로 가중하여 상기 관심맵을 획득하는 단계를 포함하는 라이다 깊이 이미지 채움 방법.

청구항 15

제14항에 있어서, 상기 관심맵을 획득하는 단계는

SSIM(structural similarity index measurement)의 구조 유사도 평가 방식에 따라 대응하는 특징맵과 상기 퓨전 특징맵 사이의 구조 유사도를 판별하고, 상기 구조 유사도에 확률적 분포를 추가 반영하여 확률적 구조 유사도를 계산하는 단계; 및

상기 관심맵을 획득하기 위해 상기 퓨전 특징맵에 계산된 확률적 구조 유사도를 가중하는 단계를 포함하는 라이다 깊이 이미지 채움 방법.

청구항 16

제15항에 있어서, 상기 퓨전 특징맵을 획득하는 단계는

상기 퓨전 특징맵(F_f^n)을 수학적식

$$F_f^n = F_e^n \cdot (1 - \tilde{m}) + F_i^n \cdot \tilde{m}$$

(여기서 F_e^n 는 특징맵이고, \tilde{m} 는 스케일 이진 마스크이며, F_i^n 는 채움 특징맵이다.)

에 따라 획득하는 라이다 깊이 이미지 채움 방법.

청구항 17

제16항에 있어서, 상기 확률적 구조 유사도를 계산하는 단계는

상기 구조 유사도(S)를 수학적식

$$s(k, q) = \frac{\sigma_{kq} + c_3}{\sigma_k \sigma_q + c_3}$$

(여기서 k는 특징맵(F_e^n)이고, q는 퓨전 특징맵(F_f^n)이며, σ 는 분산이다. 그리고 c_3 는 기지정된 상수이다.)

에 따라 계산하는 단계; 및

상기 확률적 구조 유사도를 수학적식

$$S^n = \text{softmax}(s(k, q)^n)$$

에 따라 계산하는 단계를 포함하는 라이다 깊이 이미지 채움 방법.

청구항 18

제17항에 있어서, 상기 확률적 구조 유사도를 가중하는 단계는

상기 관심맵(C^n)을 수학적식

$$C^n = (1 + \lambda_{atn} S^n) \cdot F_f^n, \quad \lambda_{atn} \in [0, 1]$$

(여기서 λ_{atn} 은 관심 가중치이다.)

에 따라 획득하는 라이다 깊이 이미지 채움 방법.

청구항 19

제14항에 있어서, 상기 라이다 깊이 이미지 채움 방법은

인공 신경망을 학습시키기 위한 학습 단계를 더 포함하고,

상기 학습 단계는

동적 객체 영역이 미리 설정되어 배경으로 채워진 학습 데이터 셋을 입력 데이터로 상기 배경 채움 네트워크에 인가하여, 상기 객체 검출 네트워크에서 출력되는 이진 마스크와 설정된 동적 객체 영역 사이의 차이로 발생하는 마스크 손실을 계산하는 단계;

상기 배경 채움 네트워크에서 출력되는 채움 특징맵 사이의 차이로 발생하는 채움 손실을 계산하는 단계; 및

상기 마스크 손실과 상기 채움 손실의 합으로 총 손실을 계산하고, 계산된 총 손실을 역전파하는 단계를 포함하는 라이다 깊이 이미지 채움 방법.

청구항 20

제14항에 있어서, 상기 라이다 깊이 이미지 채움 방법은

인공 신경망을 학습시키기 위한 학습 단계를 더 포함하고,

상기 학습 단계는

상기 채움 특징맵이 리얼 이미지인지 페이크 이미지인지 판별하는 단계;

상기 채움 특징맵에 대한 판별 결과에 기초하여 생성적 적대 네트워크 방식으로 생성 손실과 적대적 손실을 각각 계산하는 단계; 및

계산된 생성 손실과 적대적 손실의 합을 총 손실로 역전파하는 단계를 포함하는 라이다 깊이 이미지 채움 방법.

발명의 설명

기술 분야

[0001] 본 발명은 라이다 깊이 이미지 채움 장치 및 방법에 관한 것으로, 라이다 깊이 이미지 채움 장치 및 방법에 관한 것이다.

배경 기술

[0002] 자율주행차에는 비주얼 카메라, 레이더, LiDAR, GNSS 등 다양한 센서가 부착돼 차량에 추가되고 있으며, 자율주행을 위한 정밀한 3D 지도를 만드는 데 있어서, 이러한 센서들을 활용하여 많은 종류의 연구 분야에서 연구되어 오고 있다.

[0003] 완전한 자율주행을 위해서는 주행 장면의 정확한 3D 지도를 만드는 것이 필수적이다. 그러나 움직이는 물체나 변화하는 환경으로 인하여 정적이고 정확한 3D 지도를 만드는 것은 어렵다. 기존에 3D 고밀도(High Density: HD) 지도를 취득하기 위한 대표적인 방식은 고성능의 3D 측정 장비를 탑재하여 동일한 공간을 여러 번 이동하며 3D 점군(point cloud)으로 구성되는 3D 점군 이미지를 취득하는 것이다. 취득된 3D 이미지에는 다양한 동적 객체(차량, 보행자 등)가 함께 포함되어 있는 경우가 대부분이므로, 정적 3D 지도를 만들기 위해서는 주변의 동적 객체들을 제거해야 한다. 그리고 동적 객체가 제거된 영역에 대해 동적 객체가 존재하지 않을 때의 정보를 맵핑(mapping)시켜 빈 공간을 채워야 한다.

[0004] 그러나 동적 객체가 빈번하게 위치하는 영역의 경우, 다수로 반복 촬영할지라도 동적 객체가 반복해서 촬영되어 동적 객체를 정상적으로 제거하지 못하는 경우가 빈번하게 발생할 뿐만 아니라, 제거하더라도 빈공간을 채우지 못하는 경우가 발생하게 된다.

[0005] 도 1은 기존의 방식으로 획득된 3D 점군 이미지의 일 예를 나타낸다.

[0006] 도 1의 (a) 내지 (c)에서 사각 박스 또는 타원으로 표시된 바와 같이, 기존의 방식으로 획득된 3D 점군 이미지의 경우 여러 번 반복하여 촬영될지라도, 다수의 차량이 빈번하게 이동하는 도심의 도로 등과 같은 영역에서는 서로 다른 동적 객체가 계속적으로 해당 위치에 존재하여 중복하여 누적 감지됨으로써, 동적 객체로 인식되지 않아 제거되지 않게 된다. 따라서 3D 지도에서도 동적 객체가 배경으로 인식되어 포함되는 오류가 발생하게 된다.

[0007] 도 2는 3D 점군 이미지로부터 기존의 인공 신경망을 이용하여 추출되는 특징점의 일 예를 나타낸다.

[0008] 도 2에서(a)는 3D 점군 이미지를 나타내고, (b)와 (c)는 각각 (a)의 3D 점군 이미지에서 기존의 인공 신경망을 이용하여 추출된 특징점과 수작업 기반 ISS(Intrinsic shape signatures) 방식으로 추출된 특징점을 나타낸다. 도 2의 (b) 및 (c)를 살펴보면, 인공 신경망을 이용하거나 수작업으로 특징점을 추출하는 경우에도, 배경과 동적 객체의 구분이 용이하지 않아 동적 객체에서 특징점이 검출됨을 알 수 있다. 따라서 이와 같은 기법을 활용한 3D 점군 들을 매칭하는 경우, 오차가 발생하게 되어 정확하게 동적 객체를 제거할 수 없다는 한계가 있다.

[0009] 또한 수작업에 기반하여 동적 객체를 제거하고, 배경의 정보로 매칭하는 방법은 많은 시간이 소요되는 작업이며, 도심 내에서 건물 또는 조형의 변화가 있다면 다시 이미지를 취득해야 하는 번거로움이 존재한다.

선행기술문헌

특허문헌

[0010] (특허문헌 0001) 한국 공개 특허 제10-2018-0134563호 (2018.12.19 공개)

발명의 내용

해결하려는 과제

[0011] 본 발명의 목적은 라이다를 이용하여 획득된 3D 점군 이미지에서 정확하게 동적 객체 영역을 검출하고, 검출된 동적 객체 영역을 배경으로 채워 동적 객체가 제거된 3D 지도를 획득할 수 있는 라이다 깊이 이미지 채움 장치 및 방법을 제공하는데 있다.

[0012] 본 발명의 다른 목적은 검출된 동적 객체 영역을 주변의 구조적 특성을 반영하여 배경을 채움으로써 동적 객체 영역을 이질감없이 배경으로 채워 자연스러운 3D 지도를 획득할 수 있는 라이다 깊이 이미지 채움 장치 및 방법을 제공하는데 있다.

과제의 해결 수단

[0013] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른 라이다 깊이 이미지 채움 장치는 라이다를 이용하여 획득된 3차원 점군 이미지가 인가되면, 기지정된 방식으로 2차원 이미지로 변환하여 입력 데이터를 획득하는 입력 데이터 획득부; 미리 학습된 인공 신경망으로 구현되어, 상기 입력 데이터를 학습된 방식에 따라 인코딩하여 특징맵을 추출하는 과정을 반복하여 획득되는 다수의 특징맵 중 최종 인코딩된 특징맵인 인코딩 특징맵을 학습된 방식에 따라 디코딩하여 동적 객체 영역을 나타내는 이진 마스크를 획득하고, 확률적 구조 유사도에 따라 상기 다수의 특징맵 중 대응하는 특징맵에서 상기 이진 마스크에 대응하는 영역을 이전 획득된 채움 특징맵의 영역으로 대체하여 관심맵을 획득하며, 획득된 관심맵을 미리 학습된 방식으로 디코딩하여 채움 특징맵을 반복적으로 획득하는 배경 채움 네트워크; 및 최종 획득된 채움 특징맵을 채움 이미지로 인가받고, 기지정된 방식으로 3차원 이미지로 변환하는 3D 지도 획득부를 포함한다.

[0014] 상기 배경 채움 네트워크는 각각 입력 데이터 또는 이전 배치된 컨볼루션 레이어에서 출력되는 특징맵을 인가받아 인코딩하여 특징맵을 출력하는 다수의 컨볼루션 레이어를 포함하여 인코딩 네트워크; 상기 인코딩 특징맵을 인가받아 학습된 방식에 따라 디코딩하여, 상기 입력 데이터에서 학습에 의해 지정된 동적 객체가 위치하는 영역을 나타내는 상기 이진 마스크를 획득하는 객체 검출 네트워크; 상기 인코딩 특징맵 또는 이전 획득된 관심맵을 인가받아 학습된 방식으로 디코딩하여 채움 특징맵을 획득하는 채움 네트워크; 및 상기 다수의 특징맵 중 대응하는 특징맵에서 상기 이진 마스크에 대응하는 영역을 이전 획득된 채움 특징맵의 영역으로 대체하여 퓨전 특징맵을 획득하고, 대응하는 특징맵과 획득된 퓨전 특징맵 사이의 확률적 구조 유사도를 계산하여 상기 퓨전 특징맵에 기지정된 방식으로 가중하여 상기 관심맵을 획득하는 다수의 퓨전 블록을 구비하는 퓨전 블록부를 포함할 수 있다.

[0015] 상기 객체 검출 네트워크는 상기 이진 마스크를 상기 인코딩 네트워크에서 획득되는 다수의 특징맵 각각에 대응하는 크기로 스케일링하여 다수의 스케일 이진 마스크를 획득할 수 있다.

[0016] 상기 다수의 퓨전 블록 각각은 상기 다수의 특징맵 중 대응하는 특징맵과 다수의 스케일 이진 마스크 중 대응하는 스케일 이진 마스크 및 상기 채움 네트워크에서 이전 획득된 채움 특징맵을 인가받아, 상기 대응하는 특징맵에서 대응하는 스케일 이진 마스크에 의해 지정된 영역을 이전 획득된 채움 특징맵의 영역으로 대체하여 퓨전

특징맵을 획득하는 퓨전부; SSIM(structural similarity index measurement)의 구조 유사도 평가 방식에 따라 대응하는 특징맵과 상기 퓨전 특징맵 사이의 구조 유사도를 판별하고, 확률적 분포를 추가 반영하여 확률적 구조 유사도를 계산하는 구조 유사도 계산부; 및 상기 퓨전 특징맵에 계산된 확률적 구조 유사도를 가중하여 상기 관심맵을 획득하는 관심맵 획득부를 포함할 수 있다.

상기 퓨전부는 상기 퓨전 특징맵(F_f^n)을 수학식

$$F_f^n = F_e^n \cdot (1 - \tilde{m}) + F_i^n \cdot \tilde{m}$$

(여기서 F_e^n 는 특징맵이고, \tilde{m} 는 스케일 이진 마스크이며, F_i^n 는 채움 특징맵이다.)에 따라 획득할 수 있다.

상기 구조 유사도 계산부는 상기 구조 유사도(S)를 수학식

$$s(k, q) = \frac{\sigma_{kq} + c_3}{\sigma_k \sigma_q + c_3}$$

(여기서 k 는 특징맵(F_e^n)이고, q 는 퓨전 특징맵(F_f^n)이며, σ 는 분산이다. 그리고 c_3 는 기지정된 상수이다.)에 따라 계산하고,

상기 확률적 구조 유사도를 수학식

$$S^n = \text{softmax}(s(k, q)^n)$$

에 따라 계산할 수 있다.

상기 관심맵 획득부는 상기 관심맵(C^n)을 수학식

$$C^n = (1 + \lambda_{atn} S^n) \cdot F_f^n, \quad \lambda_{atn} \in [0, 1]$$

(여기서 λ_{atn} 은 관심 가중치이다.)

에 따라 획득할 수 있다.

상기 입력 데이터 획득부는 3D 점군 이미지로부터 깊이값, x축 값, y축 값, z 축 값 및 세기 값에 따른 5개의 2차원 거리 이미지를 획득하고, 획득된 5개의 2차원 거리 이미지를 채널축 방향으로 결합하여 상기 입력 데이터를 획득할 수 있다.

상기 목적을 달성하기 위한 본 발명의 다른 실시예에 따른 라이다 깊이 이미지 채움 방법은 라이다를 이용하여 획득된 3차원 점군 이미지가 인가되면, 기지정된 방식으로 2차원 이미지로 변환하여 입력 데이터를 획득하는 단계; 미리 학습된 인공 신경망을 이용하여, 상기 입력 데이터를 학습된 방식에 따라 인코딩하여 특징맵을 추출하는 과정을 반복하여 다수의 특징맵을 획득하는 단계; 미리 학습된 인공 신경망을 이용하여, 상기 다수의 특징맵 중 최종 인코딩된 특징맵인 인코딩 특징맵을 학습된 방식에 따라 디코딩하여 동적 객체 영역을 나타내는 이진 마스크를 획득하는 단계; 미리 학습된 인공 신경망을 이용하여, 확률적 구조 유사도에 따라 상기 다수의 특징맵 중 대응하는 특징맵에서 상기 이진 마스크에 대응하는 영역을 이전 획득된 채움 특징맵의 영역으로 대체하여 관심맵을 획득하고, 획득된 관심맵을 미리 학습된 방식으로 디코딩하여 채움 특징맵을 획득하는 단계; 및 최종 획득된 채움 특징맵을 채움 이미지로 인가받고, 기지정된 방식으로 3차원 이미지로 변환하는 단계를 포함한다.

발명의 효과

따라서, 본 발명의 실시예에 따른 라이다 깊이 이미지 채움 장치 및 방법은 라이다를 이용하여 획득된 3D 점군 이미지에서 정확하게 동적 객체 영역을 검출하고, 검출된 동적 객체 영역을 주변의 구조적 특성이 확률 분포에 기반하여 반영된 배경으로 채움으로써, 동적 객체 영역이 자연스러운 배경으로 대체되어 자율 주행에 이용 가능한 수준의 고품질의 3D 지도를 획득할 수 있다. 또한 생성적 적대 신경망 기법으로 학습이 수행될 수 있어, 학습 데이터 셋 없이도 학습이 가능할 뿐만 아니라, 다른 인공 신경망을 학습시키기 위한 학습 데이터 셋으로 이용 가능한 3D 지도를 제공할 수 있다.

도면의 간단한 설명

- [0033] 도 1은 기존의 방식으로 획득된 3D 점군 이미지의 일 예를 나타낸다.
- 도 2는 3D 점군 이미지로부터 기존의 인공 신경망을 이용하여 추출되는 특징점의 일 예를 나타낸다.
- 도 3은 본 발명의 일 실시예에 따른 라이다 깊이 이미지 채움 장치의 개략적 구조를 나타낸다.
- 도 4는 도 3의 입력 데이터 획득부에서 획득되는 입력 데이터의 일 예를 나타낸다.
- 도 5는 도 3의 객체 보상 네트워크의 상세 구조의 일 예를 설명하기 위한 도면이다.
- 도 6은 도 5의 퓨전 블록의 구성의 일 예를 나타낸다.
- 도 7은 본 실시예에 따른 라이다 깊이 이미지 채움 장치를 이용하여 획득되는 3D 지도 이미지의 일 예를 나타낸다.
- 도 8은 본 발명의 일 실시예에 따른 라이다 깊이 이미지 채움 방법을 나타낸다.

발명을 실시하기 위한 구체적인 내용

- [0034] 본 발명과 본 발명의 동작상의 이점 및 본 발명의 실시예에 의하여 달성되는 목적을 충분히 이해하기 위해서는 본 발명의 바람직한 실시예를 예시하는 첨부 도면 및 첨부 도면에 기재된 내용을 참조하여야만 한다.
- [0035] 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.
- [0036] 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.
- [0037] 도 3은 본 발명의 일 실시예에 따른 라이다 깊이 이미지 채움 장치의 개략적 구조를 나타내고, 도 4는 도 3의 입력 데이터 획득부에서 획득되는 입력 데이터의 일 예를 나타낸다.
- [0038] 도 3을 참조하면, 본 실시예에 따른 라이다 깊이 이미지 채움 장치는 점군 이미지 획득부(100), 입력 데이터 획득부(200), 배경 채움 네트워크(300) 및 3D 지도 획득부(400)를 포함할 수 있다.
- [0039] 점군 이미지 획득부(100)는 동적 객체를 고려하지 않고 취득된 3D 점군 이미지를 획득한다. 점군 이미지 획득부(100)는 라이다로 구현되어 3D 점군 이미지를 직접 획득하거나, 라이다에서 획득된 3D 점군 이미지를 인가받는 통신 모듈 또는 미리 획득된 3D 점군 이미지가 저장된 저장 모듈 등으로 구현될 수 있다. 그리고 라이다는 기저정된 높이 범위에서 기저정된 회전 범위를 회전하면서 주변을 스캔하여 3D 점군 이미지를 획득될 수 있다. 일 예로 라이다가 기저정된 높이 범위에서 64개의 스캔 라인 각각을 따라 360도 회전하는 동안 2,048번 샘플링하는 경우, 라이다는 $2,048 \times 64$ 개의 샘플링 픽셀로 구성되는 3D 점군 이미지를 획득할 수 있다. 이때, 획득된 3D 점군 이미지는 각 점에 대한 x, y 및 z축 방향의 위치 정보와 깊이(depth) 정보 및 세기(Intensity) 정보가 포함될 수 있다.
- [0040] 입력 데이터 획득부(200)는 점군 이미지 획득부(100)에서 획득된 3D 점군 이미지를 2차원의 거리 이미지(Range Image)로 변환한다. 입력 데이터 획득부(100)는 3D 점군 이미지의 각 점에 대해, 수직각(θ) 및 2에 따라 구좌표계의 수직각(θ)과 수평각(ϕ)으로 변환하여 2차원 공간 상에 투영함으로써, 2차원 거리 이미지를 획득할 수 있다.

수학적식 1

[0041]
$$\theta = \text{atan}(z/\sqrt{x^2 + y^2})$$

수학식 2

$$\phi = \text{atan}(y/x)$$

[0042]

[0043]

이때 입력 데이터 획득부(200)는 라이다 스캔 규격에 대응하여 거리 이미지를 정규화하여 획득할 수 있다. 즉 2차원의 거리 이미지는 360도 전방향을 스캔하는 라이다 센서로부터 $2,048 \times 64$ 개의 픽셀을 갖는 크기로 획득될 수 있다. 다만 입력 데이터 획득부(200)는 라이다로부터 3D 점군 이미지가 360도 전방향에 대해 획득될지라도 기지정된 각도 범위(여기서는 일 예로 90도)에 대해서만 추출하여 입력 데이터로 획득할 수도 있다. 이 경우, 입력 데이터는 일 예로 512×64 의 픽셀을 가질 수 있다.

[0044]

또한 입력 데이터 획득부(200)는 도 4에 도시된 바와 같이, 2차원 거리 이미지를 깊이(depth)값, x축 값, y축값, z 축값 및 세기(Intensity)값에 따른 5개로 구분하여 입력 데이터로서 획득할 수 있다. 도 4에서 (a) 내지 (e)는 각각 깊이값, x축 값, y축값, z 축값 및 세기 값에 대한 2차원 거리 이미지를 나타낸다. 그리고 입력 데이터 획득부(200)는 5개의 거리 이미지를 채널 축 방향으로 결합하여 입력 데이터를 획득할 수 있다. 따라서 입력 데이터는 일 예로 $512 \times 64 \times 5$ 의 크기로 획득될 수 있다. 그리고 획득된 입력 데이터는 배경 채움 네트워크(300)에 인가된다.

[0045]

인공 신경망으로 구현되는 배경 채움 네트워크(300)는 입력 데이터를 인가받고, 미리 학습된 방식에 따라 인가된 입력 데이터에서 동적 객체를 검출하고, 검출된 동적 객체 영역을 배경과의 구조 유사도에 기반하여 학습된 방식에 따라 소거하고 주변 배경으로 채워 배경 채움 이미지를 획득한다.

[0046]

배경 채움 네트워크(300)는 인코딩 네트워크(310)와 객체 검출 네트워크(320), 퓨전 블록부(330) 및 채움 네트워크(340)를 포함할 수 있다. 여기서 인코딩 네트워크(310)와 객체 검출 네트워크(320) 및 채움 네트워크(340)는 각각 미리 학습된 인공 신경망으로 구현될 수 있다.

[0047]

도 5는 도 3의 객체 보상 네트워크의 상세 구조의 일 예를 설명하기 위한 도면이다.

[0048]

도 5를 참조하여 도 1의 배경 채움 네트워크(300)를 설명하면, 우선 인코딩 네트워크(310)는 입력 데이터를 인가받고, 학습된 방식에 따라 인가된 입력 데이터의 특징을 추정하여 인코딩 특징맵을 획득한다. 인코딩 네트워크(310)는 도 5에 도시된 바와 같이, 순차적으로 연결된 다수의 컨볼루션 레이어를 포함하여 구성될 수 있다. 도 5에서는 인코딩 네트워크(310)가 5개의 컨볼루션 레이어로 구성되는 경우를 예로 들어 도시하였으나, 컨볼루션 레이어의 개수 및 각각의 크기는 다양하게 조절될 수 있다. 다만 도 5에서와 같이, 다수의 컨볼루션 레이어는 제1 컨볼루션 레이어가 5×5 크기의 컨볼루션 레이어로 구성되는 반면, 나머지 컨볼루션 레이어는 3×3 크기의 컨볼루션 레이어로 구성될 수 있다. 즉 인코딩 네트워크(310)의 다수의 컨볼루션 레이어 중 제1 컨볼루션 레이어를 제외한 나머지 컨볼루션 레이어는 동일한 크기의 컨볼루션 레이어로 구성되는 것이 바람직하다.

[0049]

순차 연결된 5개의 컨볼루션 레이어는 각각 입력 데이터 또는 이전 컨볼루션 레이어에서 출력되는 특징맵(F_e^{n-1})(여기서 n은 자연수)을 인가받아 특징점을 추정하여 특징맵(F_e^n)을 획득하고, 획득된 특징맵(F_e^n)을 다음 컨볼루션 레이어로 전달하며, 마지막으로 컨볼루션 레이어에서 출력되는 특징맵을 인코딩 특징맵(F_e)이라 한다. 따라서 인코딩 네트워크(310)는 다수의 컨볼루션 레이어를 통해 입력 네트워크의 특징을 정확하게 추출하여 인코딩 특징맵(F_e)을 획득한다.

[0050]

객체 검출 네트워크(320)는 인코딩 네트워크(310)에서 획득된 인코딩 특징맵(F_e)을 인가받고, 미리 학습된 방식에 따라 인가된 특징맵(F_e)에서 동적 객체 영역을 추정하여 동적 객체 영역을 나타내는 이진 마스크(m)를 추출한다. 객체 검출 네트워크(320)는 동적 객체에 대한 패턴이 미리 학습되어, 인가된 특징맵에서 동적 객체가 포함된 것으로 추정되는 영역을 추출한다. 여기서 추출되는 영역은 후술하는 퓨전 블록부(330)가 인가된 특징맵(F_e^n)에서 동적 객체를 소거하기 위한 영역을 설정하기 위해 이용된다.

[0051]

객체 검출 네트워크(320) 또한 도 5에 도시된 바와 같이, 순차 연결된 다수의 컨볼루션 레이어를 포함하여 구성될 수 있으며, 포함되는 컨볼루션 레이어의 개수는 인코딩 네트워크(310)에 포함된 컨볼루션 레이어의 개수와 검출하고자 하는 동적 객체의 종류 수 등에 따라 다양하게 조절될 수 있다. 다만 인코딩 네트워크(310)와 객체

검출 네트워크(320)의 다수의 컨볼루션 레이어를 거치면서 전달되는 각 컨볼루션 레이어에서 출력되는 특징맵(F_e^n)의 2차원 크기가 점차로 줄어들게 된다. 이에 객체 검출 네트워크(320)는 다수의 컨볼루션 레이어를 거치면서 점차로 감소하는 특징맵(F_e^n)의 2차원 크기를 증가시키기 위해 다수의 업샘플링 레이어를 더 포함할 수 있다. 다수의 업샘플링 레이어는 각각 다수의 컨볼루션 레이어에서 기지정된 개수 컨볼루션 레이어 이후에 삽입 배치 되어 기지정된 방식 또는 학습된 방식에 따라 이전 레이어에서 인가되는 특징맵(F_e^n)의 2차원 크기를 확장하여 출력한다. 이때, 다수의 업샘플링 레이어는 일 예로 인가되는 특징맵의 채널축 방향의 값을 기지정된 방식으로 결합하여 2차원 크기로 확장시킬 수 있다.

[0052] 한편, 퓨전 블록부(330)는 인코딩 네트워크(310)의 다수의 컨볼루션 레이어에서 출력되는 다수의 특징맵(F_e^n)과 객체 검출 네트워크(320)에서 인가되는 이진 마스크(\tilde{m}) 및 채움 네트워크(340)의 다수의 컨볼루션 레이어에서 출력되는 채움 특징맵(F_i^n)을 인가받아, 기지정된 방식으로 결합하여 관심맵(attention map)(C^n)을 획득한다.

[0053] 도 4에 도시된 바와 같이, 퓨전 블록부(330)는 다수의 퓨전 블록(FB1 ~ FB3)을 포함할 수 있으며, 다수의 퓨전 블록(FB1 ~ FB3)은 각각 채움 네트워크(340)의 다수의 컨볼루션 레이어 사이에 배치될 수 있다.

[0054] 즉 도 3에서는 비록 설명의 편의를 위해, 퓨전 블록부(330)와 채움 네트워크(340)를 구분하여 도시하였으나, 퓨전 블록부(330)의 다수의 퓨전 블록(FB1 ~ FB3)은 도 4에 도시된 바와 같이, 채움 네트워크(340)와 결합되어 구성될 수 있다.

[0055] 퓨전 블록부(330)의 다수의 퓨전 블록(FB1 ~ FB3) 각각은 다수의 특징맵($F_e^1 \sim F_e^3$) 중 대응하는 특징맵(F_e^n)과 대응하는 이진 마스크(m^n) 및 다수의 채움 특징맵($F_i^1 \sim F_i^3$) 중 대응하는 채움 특징맵(F_i^n)을 인가받아 결합하여, 퓨전 특징맵(F_f^n)을 획득하고, 인가된 특징맵(F_e^n)과 획득된 퓨전 특징맵(F_f^n) 사이의 구조 유사도에 따라 대응하는 관심맵(C^n)을 획득할 수 있다.

[0056] 여기서 대응하는 이진 마스크(m^n)는 객체 검출 네트워크(320)에서 추출된 이진 마스크(m)를 퓨전 블록(FBn)에 대응하도록 스케일 축소된 이진 마스크를 나타낸다. 상기한 바와 같이, 객체 검출 네트워크(320)에서 추출된 이진 마스크(m)는 인코딩 특징맵(F_e)에 비해 2차원 크기가 확장된 크기로 추출될 수 있다. 따라서 퓨전 블록(FBn)에 인가되는 대응하는 특징맵(F_e^n)과도 2차원 크기가 매칭되지 않게 된다. 이는 퓨전 블록(FBn)이 이진 마스크(m)를 이용하여 특징맵(F_e^n)을 마스크하는 경우, 특징맵(F_e^n)의 동적 객체 영역이 아닌 영역을 마스크하게 되는 결과를 초래하게 되어 오류를 발생시킬 수 있다.

[0057] 이에 객체 검출 네트워크(320)는 최종 획득된 이진 마스크(m)를 다수의 특징맵(F_e^n) 각각에 대응하는 크기로 스케일 축소 변환하여 다수의 스케일 이진 마스크($\tilde{m} = [m^1, m^2, \dots, m^n]$)를 획득할 수 있다. 그리고 퓨전 블록부(330)의 다수의 퓨전 블록(FBn) 각각은 다수의 스케일 이진 마스크(\tilde{m}) 중 대응하는 스케일 이진 마스크(m^n)를 인가받아 특징맵(F_e^n)을 마스크할 수 있다.

[0058] 도 6은 도 5의 퓨전 블록의 구성의 일 예를 나타낸다.

[0059] 도 6을 참조하면, 다수의 퓨전 블록(FB1 ~ FB3) 각각은 특징 퓨전부(331), 구조 유사도 계산부(332) 및 관심맵 획득부(333)를 포함할 수 있다.

[0060] n번째 퓨전 블록(FBn)에서 특징 퓨전부(331)는 인코딩 네트워크(310)의 대응하는 컨볼루션 레이어로부터 특징맵(F_e^n)을 인가받고, 객체 검출 네트워크(320)로부터 다수의 스케일 이진 마스크(\tilde{m}) 중 대응하는 스케일 이진 마스크(m^n)를 인가받으며, 채움 네트워크(340) 중 대응하는 컨볼루션 레이어로부터 채움 특징맵(F_i^n)을

인가받는다. 그리고 인가된 특징맵(F_e^n)과 스케일 이진 마스크(m^n) 및 채움 특징맵(F_i^n)으로부터 수학식 3에 따라 퓨전 특징맵(F_f^n)을 획득한다.

수학식 3

$$F_f^n = F_e^n \cdot (1 - \tilde{m}) + F_i^n \cdot \tilde{m}$$

수학식 3에 따르면, 특징 퓨전부(331)는 스케일 이진 마스크(m^n)에 해당하지 않는 영역에 대해서는 인가된 특징맵(F_e^n)을 유지하고, 스케일 이진 마스크에 해당하는 영역은 채움 특징맵(F_i^n)의 대응하는 영역으로 채워 퓨전 특징맵(F_f^n)을 획득한다. 즉 특징 퓨전부(331)는 특징맵(F_e^n)에서 스케일 이진 마스크(m^n)에 의해 지정되는 영역을 채움 특징맵(F_i^n)으로 대체하여 퓨전 특징맵(F_f^n)을 합성한다.

그리고 획득된 퓨전 특징맵(F_f^n)을 구조 유사도 계산부(332)와 관심맵 획득부(333)로 전달한다.

구조 유사도 계산부(332)는 특징 퓨전부(331)로 인가된 특징맵(F_e^n)과 특징 퓨전부(331)에서 획득된 퓨전 특징맵(F_f^n)을 인가받아, 구조적 관점에서 인가된 특징맵(F_e^n)과 퓨전 특징맵(F_f^n) 사이의 유사도를 판별한다.

구조 유사도 계산부(332)는 구조적 관심 모듈(Structural Attention Module: SAM)으로 구현될 수 있으며, 구조적 관심 모듈은 이미지에서 화질 평가 목적으로 주로 이용되는 SSIM(structural similarity index measurement)의 구조적 요소(Structural Component)를 이용하여, 특징맵(F_e^n)과 퓨전 특징맵(F_f^n) 사이의 구조 유사도를 판별한다.

SSIM은 수학식 4와 같이, 밝기(Luminance)(l), 대조도(Contrast)(c) 및 구조(Structure)(s)의 관점에서 두 이미지(k, q) 사이의 유사도를 비교하기 위해 제안된 기법이다.

수학식 4

$$SSIM\{k, q\} = l(k, q)^\alpha c(k, q)^\beta s(k, q)^\gamma$$

(여기서 α, β, γ 는 각각 가중치이다.)

그리고 수학식 4의 밝기(l), 대조도(c) 및 구조(s) 각각의 유사도는 수학식 5에 따라 계산될 수 있다.

수학식 5

$$l(k, q) = \frac{2\mu_k\mu_q + c_1}{\mu_k^2 + \mu_q^2 + c_1}$$

$$c(k, q) = \frac{2\sigma_k\sigma_q + c_2}{\sigma_k^2 + \sigma_q^2 + c_2}$$

$$s(k, q) = \frac{\sigma_{kq} + c_3}{\sigma_k\sigma_q + c_3}$$

(여기서 k, q 는 비교 대상 이미지들이고, μ 는 평균(mean)이며, σ 는 분산(variance)이다. 그리고 c_1, c_2 및 c_3 는 기지정된 상수이다.)

[0072] 다만 본 실시예에서 구조 유사도 계산부(332)는 상기한 바와 같이, 수학적 5의 SSIM의 세가지 요소(l , c , s) 중 구조적 요소(s)만을 이용하여 두 이미지 사이의 유사도를 평가한다. 여기서 구조적 요소가 비교되는 두 개의 이미지는 특징 퓨전부(331)의 입력 이미지와 출력 이미지인 특징맵($k = F_e^n$)과 퓨전 특징맵($q = F_f^n$)이다. 즉 구조 유사도 계산부(332)는 특징 퓨전부(331)의 입력 이미지와 출력 이미지 사이의 구조 유사도를 판별한다.

[0073] 다만 본 실시예에서 구조 유사도 계산부(332)는 구조 유사도에 대해 확률적 분포를 추가로 고려하기 위해 소프트맥스(softmax) 함수를 수학적 6과 같이 추가로 반영하여 확률적 구조 유사도(S^n)를 판별한다.

수학적 6

[0074]
$$S^n = \text{softmax}(s(k, q)^n)$$

[0075] 관심맵 획득부(333)는 특징 퓨전부(331)에서 획득된 퓨전 특징맵(F_f^n)과 구조 유사도 계산부(332)에서 획득된 확률적 구조 유사도(S^n)로부터 관심맵(C^n)을 수학적 7에 따라 획득한다.

수학적 7

[0076]
$$C^n = (1 + \lambda_{atn} S^n) \cdot F_f^n, \quad \lambda_{atn} \in [0, 1]$$

[0077] (여기서 λ_{atn} 은 관심 가중치이다.)

[0078] 수학적 7에 따르면, 관심맵 획득부(333)는 퓨전 특징맵(F_f^n)에 관심 가중치(λ_{atn})에 비례하는 확률적 구조 유사도(S^n)를 가중하여 관심맵(C^n)을 획득한다.

[0079] 이때, 관심맵 획득부(333)는 도시하지 않았으나 디콘볼루션 레이어를 더 포함하여 관심맵(C^n)의 2차원 크기를 확장하여 출력할 수 있다.

[0080] 즉 확률적 구조 유사도(S^n)가 높을수록 특징맵(F_e^n)과 채움 특징맵(F_i^n)이 퓨전되어 획득된 퓨전 특징맵(F_f^n)이 강조되어 관심맵(C^n)을 획득한다.

[0081] 채움 네트워크(340)는 다수의 컨볼루션 레이어를 포함하여 구성되며, 이중 제1 컨볼루션 레이어는 인코딩 특징맵(F_e)을 인가받아 학습된 방식으로 디코딩하여 초기 채움 특징맵(F_i^1)을 획득한다. 그리고 나머지 컨볼루션 레이어는 각각 퓨전 블록부(330)의 다수의 퓨전 블록 중 이전 배치된 대응하는 퓨전 블록(FB_{n-1})으로부터 관심맵(C^{n-1})을 인가받고, 미리 학습된 방식에 따라 인가된 관심맵(C^{n-1})을 디코딩하여, 관심맵(C^{n-1})에서 특징맵(F_e^{n-1})과 채움 특징맵(F_i^{n-1})이 퓨전된 영역이 스무스하게 변화하도록 변형하여 채움 특징맵(F_i^n)을 출력한다.

[0082] 이때 채움 네트워크(340)는 다수의 컨볼루션 레이어 중 인코딩 특징맵(F_e)이 인가되는 제1 컨볼루션 레이어와 제1 퓨전 블록($FB1$) 사이에 배치되어 채움 특징맵(F_i^1)의 2차원 크기를 특징맵(F_e^1)의 크기로 확장하는 업샘플링 레이어를 더 포함하여, 초기 채움 특징맵(F_i^1)의 2차원 크기를 조절하여 출력할 수 있다.

[0083] 3D 지도 획득부(400)는 채움 네트워크(340)에서 최종 출력되는 채움 특징맵(F_i^n)인 채움 이미지(\tilde{I})를 인가받아, 기지정된 방식으로 3차원으로 변환하여 3D 지도 이미지를 획득한다. 여기서 3D 지도 획득부(400)는 일 예로 입력 데이터 획득부(200)의 2차원 이미지 획득 방식에 대항하는 방식으로 3차원 변환하여 3D 지도 이미지를 획득할 수 있다.

[0084] 상기한 라이다 깊이 이미지 채움 장치를 이용하기 위해서는 인공 신경망으로 구현되는 배경 채움 네트워크(300)가 미리 학습되어야 한다. 이에 본 실시예에 따른 라이다 깊이 이미지 채움 장치는 배경 채움 네트워크(300)의 학습 시에 결합되는 학습부(미도시)를 더 포함할 수 있다.

[0085] 학습부는 동적 객체 영역이 미리 설정된 학습 데이터를 이용하여 학습을 수행할 수 있다. 학습부는 우선 배경 채움 네트워크(300)의 객체 검출 네트워크(320)에 대한 마스크 손실(L_m)을 수학식 8에 따라 획득할 수 있다.

수학식 8

$$L_m = \lambda_m \left\| \tilde{m} - m_{gt} \right\|_1$$

[0087] (여기서 m_{gt} 는 동적 객체 영역에 대한 진리값을 나타내고, $\left\| \cdot \right\|_1$ 은 L1-norm 함수이며, λ_m)은 마스크 손실 가중치이다.

[0088] 그리고 학습부는 퓨전 블록부(330) 및 채움 네트워크(340)에 대한 채움 손실(L_i)을 수학식 8에 따라 획득할 수 있다.

수학식 9

$$L_i = \lambda_i \left\| \tilde{I} - I_{gt} \right\|_1$$

[0090] (여기서 \tilde{I} 는 채움 네트워크(340)에서 동적 객체가 배경으로 채워진 최종 채움 이미지이고, I_{gt} 는 배경 진리값을 나타낸다.)

[0091] 만일 학습 데이터가 대량으로 준비된 경우, 학습부는 수학식 8 및 9의 마스크 손실(L_m)과 채움 손실(L_i)의 합으로 총 손실(L_{total})을 계산하고, 계산된 총 손실(L_{total})을 역전파하여 배경 채움 네트워크(300)를 학습시킬 수 있다. 그러나 동적 객체가 포함된 3D 이미지에서 동적 객체가 지정된 학습 데이터의 경우, 상대적으로 용이하게 수집할 수 있는 반면, 동적 객체가 제거되어 배경으로 채워진 3D 이미지로 구성된 학습 데이터 셋을 대량으로 획득하는 것은 현실적으로 용이하지 않다.

[0092] 이에 본 실시예의 학습부는 생성적 적대 네트워크(Generative adversarial network: 이하 GAN) 기법을 적용하여 학습 데이터 셋 없이도 학습이 가능하도록 할 수 있다.

[0093] GAN 기법을 적용하기 위해, 학습부는 배경 채움 네트워크(300)를 생성자(Generator)로 하고, 별도의 판별자(discriminator)를 추가로 구비한다. 판별자는 생성자인 배경 채움 네트워크(300)와 함께 학습되어 획득되는 채움 이미지(\tilde{I})이 리얼 이미지인지 페이크 이미지인지 판별하도록 학습되는 반면, 생성자인 배경 채움 네트워크(300)는 획득되는 채움 이미지(\tilde{I})에 대해 판별자가 리얼 이미지로 인식하도록 학습된다.

[0094] 이에 학습부는 생성자인 배경 채움 네트워크(300)의 생성 손실(L_G)은 채움 이미지(\tilde{I})에 대한 판별자의 판별 결과($D(\tilde{I})$)에 기초하여 수학식 10으로 계산하는 한편, 적대적 손실(L_D)을 수학식 11에 따라 계산한다.

수학식 10

$$\mathcal{L}_G = -\mathbb{E}_{\mathbf{I} \sim P_I(\mathbf{I})}[D(\tilde{\mathbf{I}})]$$

수학식 11

$$\mathcal{L}_D = \mathbb{E}_{\mathbf{I} \sim P_{Data}} [\min(0, -1 + D(\mathbf{I}))] - \mathbb{E}_{\mathbf{x} \sim P_I(\mathbf{I})} [\min(0, -1 - D(\tilde{\mathbf{I}}))]$$

[0096]

[0097]

[0098]

[0099]

[0100]

[0101]

[0102]

[0103]

[0104]

[0105]

[0106]

[0107]

여기서 E는 데이터 분포에 대한 기대값을 획득하는 함수이다.

결과적으로 학습부는 학습 데이터 없이 수학식 10 및 11의 생성 손실(L_G)과 적대적 손실(L_D)의 합을 총 손실(L_{total})로 획득하여 학습을 수행하거나, 수학식 8의 마스크 손실(L_m)과 수학식 10 및 11의 생성 손실(L_G)과 적대적 손실(L_D)의 합을 총 손실(L_{total})로 획득하여 학습을 수행할 수도 있다.

그리고 수학식 8의 마스크 손실(L_m)과 수학식 10 및 11의 생성 손실(L_G)과 적대적 손실(L_D)의 합을 총 손실(L_{total})로 획득하여 학습을 수행되는 경우, 학습이 완료된 본 실시예에 따른 라이다 깊이 이미지 채움 장치는 다른 장치의 학습을 위한 학습 데이터 획득 장치로 이용될 수도 있다.

도 7은 본 실시예에 따른 라이다 깊이 이미지 채움 장치를 이용하여 획득되는 3D 지도 이미지의 일 예를 나타낸다.

도 7에서 (a)는 라이다 깊이 이미지 채움 장치의 입력 이미지로서, 동적 객체가 존재하는 3D 지도 이미지를 나타내고, (b)는 본 실시예에 따라 입력 이미지로부터 획득된 3D 지도 이미지를 나타내며, (c)는 퓨전 블록부(330)의 퓨전 블록(FBn)이 구조 유사도를 고려하지 않고, 특징맵(F_e^n)과 채움 특징맵(F_1^n)이 퓨전한 경우에 획득되는 3D 지도 이미지를 나타낸다. (b) 및 (c)에서 보라색 선은 동적 객체를 대체한 배경의 진리값을 나타내고, 녹색 선이 획득된 3D 지도 이미지의 배경을 나타낸다.

도 7의 (a)에 도시된 바와 같이, 입력 이미지의 경우, 동적 객체에 의해 배경이 나타나지 않는 영역이 크게 존재하지만, (b) 및 (c)의 경우, 동적 객체 영역이 채워졌음을 알 수 있다. 다만 (c)의 경우, 진리값인 보라색 라인과 획득된 녹색 라인 사이의 차이가 크게 발생하는데 반해, 확률적 구조 유사도를 고려하한 본 실시예의 경우, 진리값과 유사하게 획득됨을 알 수 있다. 즉 매우 정확하게 동적 객체를 배경으로 대체할 수 있다.

도 8은 본 발명의 일 실시예에 따른 라이다 깊이 이미지 채움 방법을 나타낸다.

도 3 내지 도 6을 참조하여, 도 8의 라이다 깊이 이미지 채움 방법을 설명하면, 우선, 동적 객체를 고려하지 않고 라이다에서 취득된 깊이 이미지인 3D 점군 이미지를 획득한다(S11). 그리고 획득된 3D 점군 이미지를 기지정된 방식으로 2차원의 거리 이미지로 변환하여 입력 데이터를 획득한다(S12). 이때, 3D 점군 이미지를 깊이값, x축 값, y축 값, z 축 값 및 세기 값 각각에 대한 5개의 2차원 거리 이미지로 변환하고, 5개의 거리 이미지를 채널 축 방향으로 결합하여 입력 데이터를 획득할 수 있다.

그리고 입력 이미지를 미리 학습된 인공 신경망으로 구현되는 인코딩 네트워크(310)에 입력하여 학습된 방식에 따라 인코딩함으로써, 인코딩 특징맵(F_e)을 획득한다(S13). 여기서 인코딩 네트워크는 순차적으로 연결된 다수의 컨볼루션 레이어로 구성되고, 각 컨볼루션 레이어 각각은 인가되는 입력 이미지로부터 단계적으로 특징맵(F_e^n)을 추출하고, 최종 컨볼루션 레이어는 추출된 특징맵을 인코딩 특징맵(F_e)으로서 출력한다.

인코딩 특징맵(F_e)이 획득되면, 미리 학습된 인공 신경망으로 구현되는 객체 검출 네트워크(320)는 인코딩 특징맵(F_e)을 인가받아 학습된 방식으로 디코딩하여, 입력 이미지에서 동적 객체에 해당하는 영역을 나타내는 이진 마스크(m)를 획득한다(S14). 여기서 객체 검출 네트워크(320) 또한 다수의 컨볼루션 레이어를 포함하여, 순차적 디코딩을 수행할 수 있으며, 다수의 컨볼루션 레이어의 컨볼루션 연산에 의해 축소되는 특징맵의 2차원 크기를 확장하기 위해 기지정된 간격으로 배치되는 다수의 업샘플링 레이어를 더 포함할 수 있다.

그리고 획득된 이진 마스크(m)를 다수의 특징맵(F_e^n) 각각에 대응하는 크기로 스케일 축소 변환하여 다수의 스케일 이진 마스크($\tilde{m} = [m^1, m^2, \dots, m^n]$)를 획득할 수 있다.

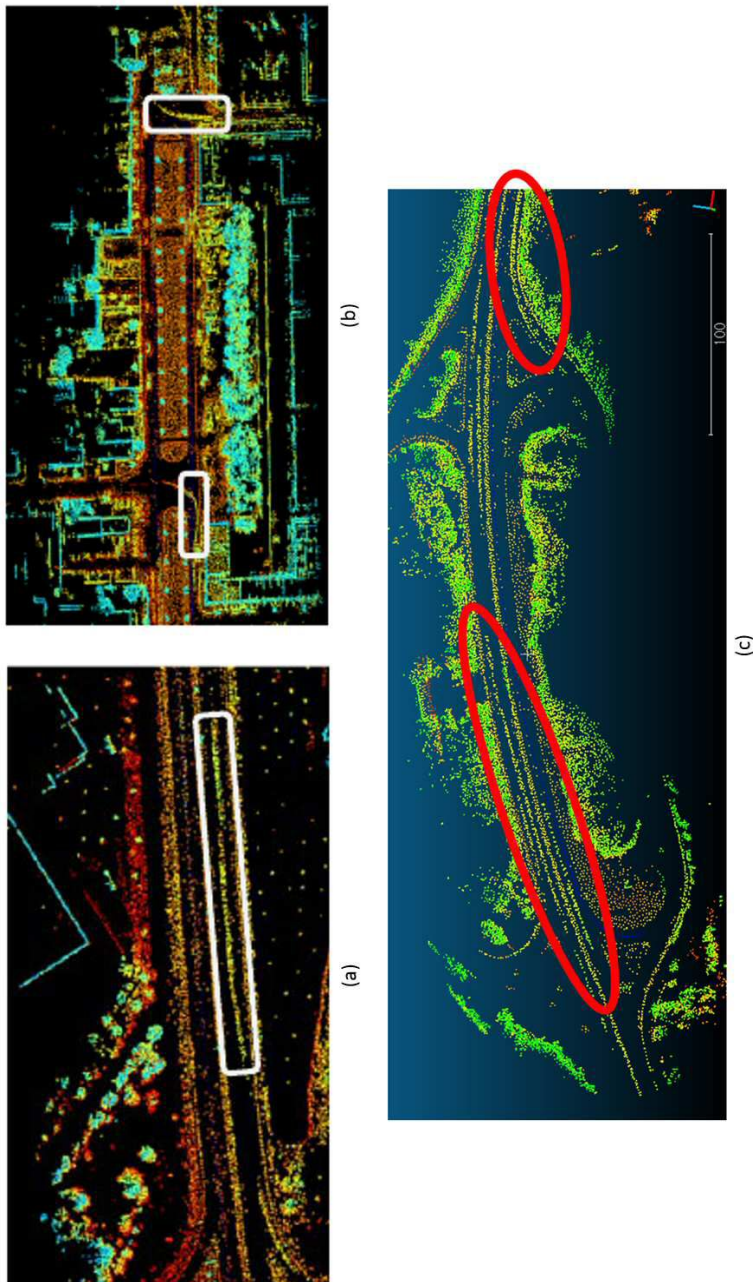
- [0108] 다수의 스케일 이진 마스크(\tilde{m})가 획득되면, 미리 학습된 인공 신경망으로 구현되는 채움 네트워크(340)의 다수의 컨볼루션 레이어 중 제1 컨볼루션 레이어가 학습된 방식에 따라 인코딩 특징맵(F_e)을 디코딩하여 초기 채움 특징맵(F_i^1)을 획득한다(S15).
- [0109] 초기 채움 특징맵(F_i^1)이 획득되면, 다수의 특징맵(F_e^n) 중 초기 채움 특징맵(F_i^1)에 대응하는 특징맵(F_e^1)에서 다수의 스케일 이진 마스크(\tilde{m}) 중 대응하는 스케일 이진 마스크(m^1)에 따라 동적 객체 영역을 마스킹하고, 마스킹된 영역을 초기 채움 특징맵(F_i^1)의 대응하는 영역인 배경을 수학적 식 3에 따라 대입하여 채움으로써 퓨전 특징맵(F_f^1)을 획득한다(S16).
- [0110] 그리고 특징맵(F_e^1)과 퓨전 특징맵(F_f^1) 사이의 확률적 구조 유사도(S^1)를 수학적 식 6에 따라 계산한다(S17). 이후, 수학적 식 7에 따라 계산된 확률적 구조 유사도(S^n)를 퓨전 특징맵(F_f^1)에 가중하여 관심맵(C^1)을 획득한다.
- [0111] 관심맵(C^1)이 획득되면, 채움 네트워크(340)의 다수의 컨볼루션 레이어 중 다음 컨볼루션 레이어가 학습된 방식에 따라 관심맵(C^1)을 인가받아 디코딩하여 채움 특징맵(F_i^2)을 획득한다(S19). 채움 특징맵(F_i^2)이 획득되면 배경 채움 횟수가 기지정된 횟수 보다 작은지 판별한다(S20). 여기서 기지정된 횟수는 채움 네트워크(340)의 컨볼루션 레이어의 개수에 대응한다. 배경 채움 횟수가 기지정된 횟수에 도달하지 않았으면, 획득된 채움 특징맵(F_i^2)을 이용하여 대응하는 특징맵(F_e^2)의 스케일 이진 마스크(m^2)에 해당하는 영역을 다시 채운다(S16).
- [0112] 그러나 기지정된 횟수에 도달한 것으로 판별되면, 획득된 채움 특징맵(F_i^n)을 채움 이미지(\tilde{I})로서 기지정된 방식으로 3차원으로 변환하여 3D 지도 이미지를 획득한다(S21).
- [0113] 본 발명에 따른 방법은 컴퓨터에서 실행시키기 위한 매체에 저장된 컴퓨터 프로그램으로 구현될 수 있다. 여기서 컴퓨터 판독가능 매체는 컴퓨터에 의해 액세스 될 수 있는 임의의 가용 매체일 수 있고, 또한 컴퓨터 저장 매체를 모두 포함할 수 있다. 컴퓨터 저장 매체는 컴퓨터 판독가능 명령어, 데이터 구조, 프로그램 모듈 또는 기타 데이터와 같은 정보의 저장을 위한 임의의 방법 또는 기술로 구현된 휘발성 및 비휘발성, 분리형 및 비분리형 매체를 모두 포함하며, ROM(판독 전용 메모리), RAM(랜덤 액세스 메모리), CD(컴팩트 디스크)-ROM, DVD(디지털 비디오 디스크)-ROM, 자기 테이프, 플로피 디스크, 광데이터 저장장치 등을 포함할 수 있다.
- [0114] 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.
- [0115] 따라서, 본 발명의 진정한 기술적 보호 범위는 첨부된 청구범위의 기술적 사상에 의해 정해져야 할 것이다.

부호의 설명

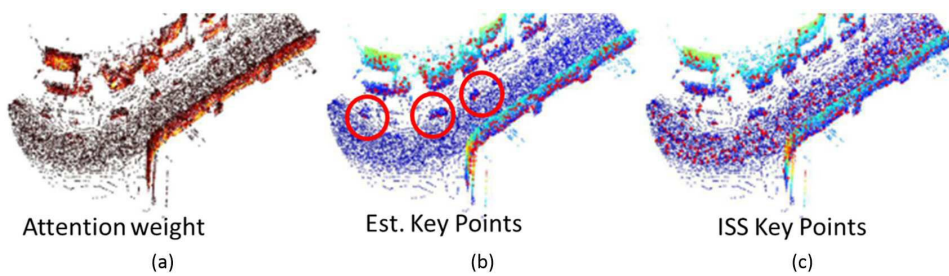
- [0116] 100: 점군 이미지 획득부 200: 입력 데이터 획득부
300: 배경 채움 네트워크 310: 인코딩 네트워크
320: 객체 검출 네트워크 330: 퓨전 블록부
FB: 퓨전 블록 331: 특징 퓨전부
332: 구조 유사도 계산부 333: 관심맵 획득부
340: 채움 네트워크 400: 3D 지도 획득부

도면

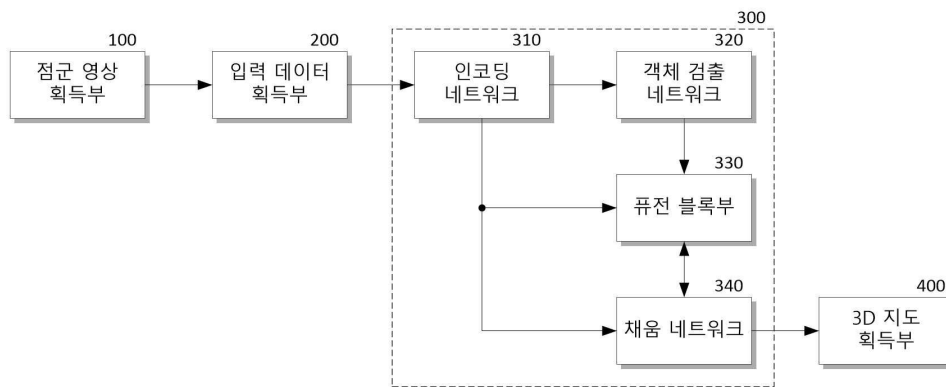
도면1



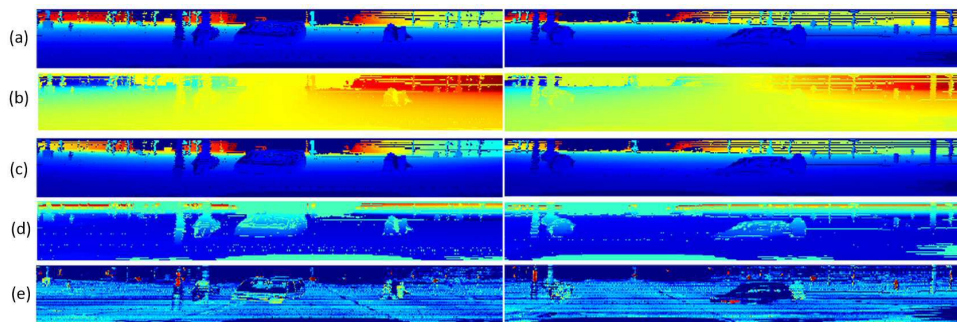
도면2



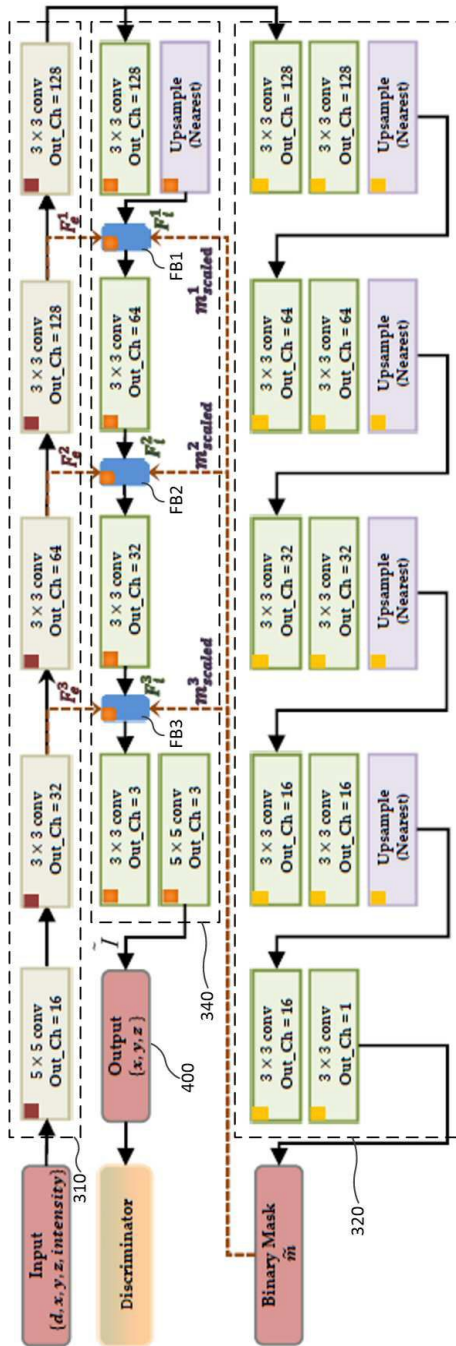
도면3



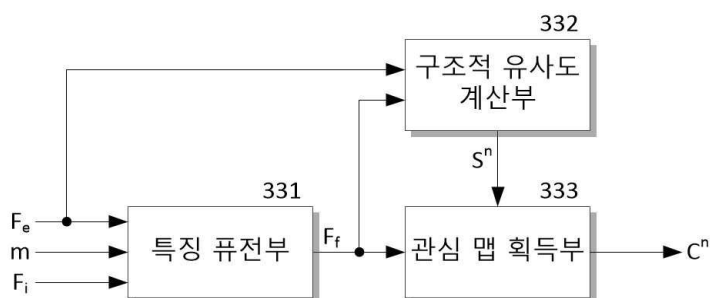
도면4



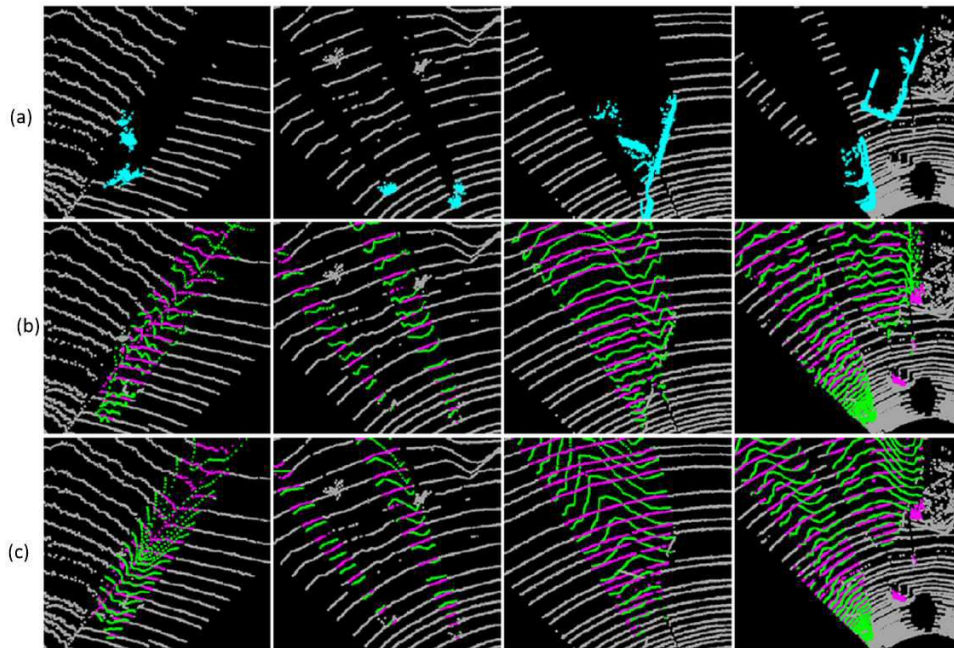
도면5



도면6



도면7



도면8

