



(19) 대한민국특허청(KR)  
(12) 등록특허공보(B1)

(45) 공고일자 2021년03월18일  
(11) 등록번호 10-2229861  
(24) 등록일자 2021년03월15일

(51) 국제특허분류(Int. Cl.)  
H04N 13/128 (2018.01) H04N 13/15 (2018.01)  
H04N 13/161 (2018.01)  
(52) CPC특허분류  
H04N 13/128 (2018.05)  
H04N 13/15 (2018.05)  
(21) 출원번호 10-2019-0132701  
(22) 출원일자 2019년10월24일  
심사청구일자 2019년10월24일  
(56) 선행기술조사문헌  
논문1\*  
KR1020170028749 A\*  
\*는 심사관에 의하여 인용된 문헌

(73) 특허권자  
연세대학교 산학협력단  
서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)  
(72) 발명자  
손광훈  
서울특별시 서대문구 연세로 50, 연세대학교 제3공학관 C125호(신촌동)  
박기홍  
서울특별시 서대문구 연세로 50, 연세대학교 제3공학관 C129호(신촌동)  
(74) 대리인  
민영준

전체 청구항 수 : 총 16 항

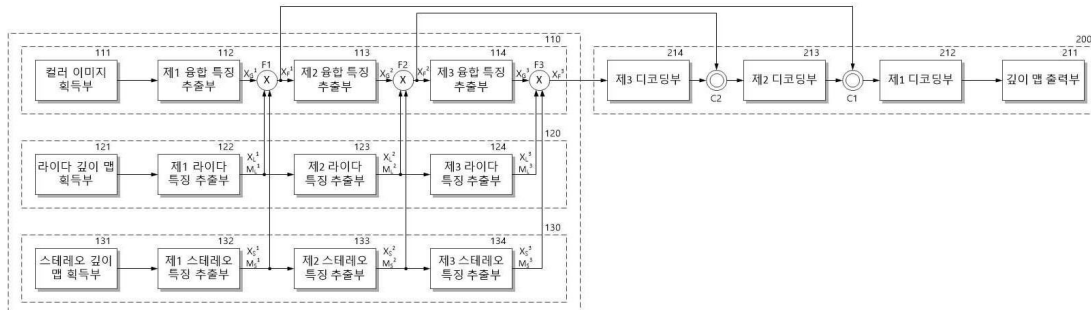
심사관 : 옥윤철

(54) 발명의 명칭 저채널 라이다와 스테레오 카메라를 이용한 깊이 추정 장치 및 방법

(57) 요약

본 발명은 카메라에서 획득된 컬러 정보를 가이드로 저채널 라이다에서 획득된 희소 밀도의 깊이 정보와 스테레오 카메라에서 획득된 부정확한 깊이 정보를 융합하여 저비용의 저채널 라이다로도 고밀도로 정확한 깊이 맵을 생성할 수 있는 깊이 추정 장치 및 방법을 제공할 수 있다.

대표도



(52) CPC특허분류

**H04N 13/161** (2018.05)

**H04N 13/257** (2018.05)

**H04N 13/271** (2018.05)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711080916
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)(3/5)	
기 여 율	1/1
과제수행기관명	연세대학교 산학협력단
연구기간	2019.01.01 ~ 2019.12.31

---

## 명세서

### 청구범위

#### 청구항 1

라이다에서 획득된 깊이 데이터로부터 라이다 입력 레이블을 획득하고, 미리 학습된 패턴 추정 방식에 따라 상기 라이다 입력 레이블에 대해 단계적으로 인코딩하여 라이다 깊이 특징을 추출하는 라이다 깊이 특징 추출부;

스테레오 카메라에서 획득된 디스패리티 맵으로부터 스테레오 입력 레이블을 획득하고, 미리 학습된 패턴 추정 방식에 따라 상기 스테레오 입력 레이블에 대해 단계적으로 인코딩하여 스테레오 깊이 특징을 추출하는 스테레오 깊이 특징 추출부;

상기 스테레오 카메라에서 획득된 컬러 이미지에서 미리 학습된 패턴 추정 방식에 따라 추출되는 가이드 특징과 대응하는 단계에서 추출된 상기 라이다 깊이 특징과 상기 스테레오 깊이 특징을 미리 학습된 방식에 따라 융합하여 융합 특징을 획득하고, 단계적으로 획득된 융합 특징을 인코딩하여 가이드 특징을 단계적으로 추출하는 컬러 가이드 특징 추출부; 및

상기 컬러 가이드 특징 추출부에서 최종 추출된 융합 특징을 인가받아 미리 학습된 특징 복원 방식에 따라 디코딩하여 복원 특징을 획득하고, 획득된 복원 특징을 단계적으로 추출되는 융합 특징 중 대응하는 융합 특징과 결합하여 결합 복원 특징을 획득하며, 획득된 결합 복원 특징으로부터 다시 복원 특징을 획득하여 최종 획득된 복원 특징을 깊이 맵으로 출력하는 디코더를 포함하되,

상기 라이다 깊이 특징 추출부는

상기 라이다에서 획득된 3D 저밀도 깊이 데이터를 2D 이미지 좌표 상에 투영하여 라이다 입력 레이블을 획득하는 라이다 깊이 맵 획득부; 및

다단 구조로 연결되어 상기 라이다 입력 레이블 또는 이전 단계에서 획득된 라이다 깊이 특징을 학습에 의해 획득되는 마스크와 결합하고 미리 학습된 패턴 추정 방식에 따라 인코딩하여 라이다 깊이 특징을 획득하는 다수의 라이다 특징 추출부를 포함하는 깊이 추정 장치.

#### 청구항 2

삭제

#### 청구항 3

제1 항에 있어서, 상기 스테레오 깊이 특징 추출부는

상기 스테레오 카메라에서 획득되는 디스패리티 맵을 인가받아 상기 라이다의 스캔 영역과 매칭시켜 스테레오 깊이 맵을 획득하고, 상기 스테레오 깊이 맵의 밀도를 상기 라이다 입력 레이블에 대응하는 밀도로 변환하여 상기 스테레오 입력 레이블을 획득하는 스테레오 깊이 맵 획득부; 및

다단 구조로 연결되어 상기 스테레오 입력 레이블 또는 이전 단계에서 획득된 스테레오 깊이 특징을 학습에 의해 획득되는 마스크와 결합하고 미리 학습된 패턴 추정 방식에 따라 인코딩하여 스테레오 깊이 특징을 획득하는 다수의 스테레오 특징 추출부를 포함하는 깊이 추정 장치.

#### 청구항 4

제3 항에 있어서, 상기 다수의 라이다 특징 추출부와 상기 다수의 스테레오 특징 추출부 각각은

이전 단계에서 획득된 라이다 깊이 특징 또는 스테레오 깊이 특징을 마스크와 결합하고 부분 컨볼루션 연산을 수행하여 라이다 깊이 특징 또는 스테레오 깊이 특징을 추출하는 깊이 추정 장치.

#### 청구항 5

제3 항에 있어서, 상기 컬러 가이드 특징 추출부는

컬러 이미지 또는 이전 단계에서 획득된 융합 특징을 인가받아 미리 학습된 패턴 방식에 따라 가이드 특징을 추출하는 다수의 융합 특징 추출부; 및

상기 다수의 융합 특징 추출부 각각에서 추출된 가이드 특징을 상기 다수의 라이다 특징 추출부 중 대응하는 라이다 특징 추출부에서 인가되는 라이다 깊이 특징 및 상기 다수의 스테레오 특징 추출부 중 대응하는 스테레오 특징 추출부에서 인가되는 스테레오 깊이 특징과 미리 학습된 방식에 따라 융합하여 융합 특징을 획득하는 다수의 특징 융합부를 포함하는 깊이 추정 장치.

#### 청구항 6

제5 항에 있어서, 상기 다수의 특징 융합부는

대응하는 라이다 특징 추출부와 대응하는 스테레오 특징 추출부 각각 이용된 마스크를 함께 인가받고, 학습에 의해 결정된 가중치에 따라 라이다 깊이 특징과 스테레오 깊이 특징 및 마스크를 융합하여 상기 융합 특징을 출력하는 깊이 추정 장치.

#### 청구항 7

제5 항에 있어서, 상기 다수의 융합 특징 추출부 각각은

인가된 융합 특징에 대한 특징을 추출하고, 추출된 특징을 인가된 융합 특징과 다시 결합하여 상기 가이드 특징을 출력하도록 잔류 네트워크 구조를 갖는 깊이 추정 장치.

#### 청구항 8

제5 항에 있어서, 상기 디코더는

상기 다수의 융합 특징 추출부의 역순에 대응하여 구성되고, 상기 컬러 가이드 특징 추출부에서 최종 추출된 융합 특징 또는 결합 복원 특징을 인가받아 미리 학습된 특징 복원 방식에 따라 디코딩하여 복원 특징을 출력하는 다수의 디코딩부; 및

상기 다수의 디코딩부 중 대응하는 디코딩부에서 출력되는 복원 특징과 대응하는 융합 특징 추출부에서 전달되는 융합 특징을 결합하여 결합 복원 특징을 출력하는 다수의 결합부를 포함하는 깊이 추정 장치.

#### 청구항 9

제1 항에 있어서, 상기 라이다에서 획득된 깊이 데이터는

상기 라이다의 다수의 채널이 가능한 수준에서 수평 방향을 스캔하여 획득된 깊이 데이터인 깊이 추정 장치.

#### 청구항 10

제1 항에 있어서, 상기 깊이 추정 장치는

학습시에 깊이 값이 미리 측정된 2차원 이미지를 학습용 데이터로 인가받고, 상기 디코더에서 출력되는 깊이 맵의 깊이 값과 차이를  $L_2$  놈(Norm)으로 손실로 계산하여 역전파하는 깊이 추정 장치.

#### 청구항 11

라이다에서 획득된 깊이 데이터로부터 라이다 입력 레이블을 획득하고, 미리 학습된 패턴 추정 방식에 따라 상기 라이다 입력 레이블에 대해 단계적으로 인코딩하여 라이다 깊이 특징을 추출하는 단계;

스테레오 카메라에서 획득된 디스패리티 맵으로부터 스테레오 입력 레이블을 획득하고, 미리 학습된 패턴 추정 방식에 따라 상기 스테레오 입력 레이블에 대해 단계적으로 인코딩하여 스테레오 깊이 특징을 추출하는 단계;

상기 스테레오 카메라에서 획득된 컬러 이미지에서 미리 학습된 패턴 추정 방식에 따라 추출되는 가이드 특징과 대응하는 단계에서 추출된 상기 라이다 깊이 특징과 상기 스테레오 깊이 특징을 미리 학습된 방식에 따라 융합하여 융합 특징을 획득하고, 단계적으로 획득된 융합 특징을 인코딩하여 가이드 특징을 추출하는 단계; 및

최종 추출된 융합 특징을 인가받아 미리 학습된 특징 복원 방식에 따라 디코딩하여 복원 특징을 획득하고, 획득

된 복원 특징을 단계적으로 추출되는 융합 특징 중 대응하는 융합 특징과 결합하여 결합 복원 특징을 획득하며, 획득된 결합 복원 특징으로부터 다시 복원 특징을 획득하여 최종 획득된 복원 특징을 깊이 맵으로 출력하는 단계를 포함하되,

상기 라이다 깊이 특징을 추출하는 단계는

상기 라이다에서 획득된 3D 저밀도 깊이 데이터를 2D 이미지 좌표 상에 투영하여 라이다 입력 레이블을 획득하는 단계; 및

상기 라이다 입력 레이블 또는 이전 획득된 라이다 깊이 특징을 학습에 의해 획득되는 마스크와 결합하고, 미리 학습된 패턴 추정 방식에 따라 인코딩하여 라이다 깊이 특징을 반복하여 획득하는 단계를 포함하는 깊이 추정 방법.

## 청구항 12

삭제

## 청구항 13

제11 항에 있어서, 상기 스테레오 깊이 특징을 추출하는 단계는

상기 스테레오 카메라에서 획득되는 디스패리티 맵을 인가받아 상기 라이다의 스캔 영역과 매칭시켜 스테레오 깊이 맵을 획득하고, 상기 스테레오 깊이 맵의 밀도를 상기 라이다 입력 레이블에 대응하는 밀도로 변환하여 상기 스테레오 입력 레이블을 획득하는 단계; 및

상기 스테레오 입력 레이블 또는 이전 획득된 스테레오 깊이 특징을 학습에 의해 획득되는 마스크와 결합하고 미리 학습된 패턴 추정 방식에 따라 인코딩하여 스테레오 깊이 특징을 반복하여 획득하는 단계를 포함하는 깊이 추정 방법.

## 청구항 14

제13 항에 있어서, 상기 가이드 특징을 추출하는 단계는

컬러 이미지 또는 이전 획득된 융합 특징을 인가받아 미리 학습된 패턴 방식에 따라 가이드 특징을 반복적으로 추출하는 단계; 및

추출된 가이드 특징을 반복하여 추출된 다수의 라이다 깊이 특징 중 대응하는 라이다 깊이 특징 및 반복하여 추출된 다수의 스테레오 깊이 특징 중 대응하는 스테레오 깊이 특징과 미리 학습된 방식에 따라 융합하여 융합 특징을 획득하는 단계를 포함하는 깊이 추정 방법.

## 청구항 15

제14 항에 있어서, 상기 융합 특징을 획득하는 단계는

대응하는 라이다 깊이 특징과 대응하는 스테레오 깊이 특징 추출시에 이용된 마스크를 함께 인가받고, 학습에 의해 결정된 가중치에 따라 라이다 깊이 특징과 스테레오 깊이 특징 및 마스크를 융합하여 상기 융합 특징을 출력하는 깊이 추정 방법.

## 청구항 16

제14 항에 있어서, 상기 깊이 맵으로 출력하는 단계는

상기 가이드 특징을 추출하는 단계에서 최종 추출된 융합 특징 또는 결합 복원 특징을 인가받아 미리 학습된 특징 복원 방식에 따라 디코딩하여 복원 특징을 출력하는 단계; 및

상기 복원 특징과 다수의 융합 특징 중 대응하는 융합 특징을 결합하여 결합 복원 특징을 출력하는 단계; 를 포함하는 깊이 추정 방법.

## 청구항 17

제11 항에 있어서, 상기 라이다에서 획득된 깊이 데이터는

상기 라이다의 다수의 채널이 가능한 수준에서 수평 방향을 스캔하여 획득된 깊이 데이터인 깊이 추정 방법.

## 청구항 18

제11 항에 있어서, 상기 깊이 추정 방법은

학습 단계를 더 포함하고,

상기 학습 단계는

깊이 값이 미리 측정된 2차원 이미지를 학습용 데이터로 인가받고, 출력되는 깊이 맵의 깊이 값과 차이를  $L_2$  norm (Norm)으로 손실로 계산하여 역전파하는 깊이 추정 방법.

## 발명의 설명

### 기술 분야

[0001] 본 발명은 깊이 추정 장치 및 방법에 관한 것으로, 저채널 라이다와 스테레오 카메라를 이용한 깊이 추정 장치 및 방법에 관한 것이다.

### 배경 기술

[0002] 장면의 3D 기하학적 구성을 인식하는 것은 자율 주행 차, 모바일 로봇, 위치 파악과 매핑, 장애물 회피 및 경로 계획 및 3D 재구성과 같은 많은 로봇 공학 및 컴퓨터 비전 응용 분야에서 수많은 작업에 필수적이다.

[0003] 현재 장면의 신뢰할 수 있는 깊이 정보를 추정하기 위해서는 RGB-D 센서 또는 3D 라이다(LiDAR)와 같은 능동형 3D 스캐너 또는 스테레오 이미지에 대한 패시브 매칭 알고리즘 등이 이용되고 있다.

[0004] 실외 환경에서는 RGB-D 센서의 경우 조명에 의해 정확도가 떨어지고 제한된 감지 범위를 제공하기 때문에, 센티미터 정도의 오차로 매우 정확한 깊이 정보를 제공할 수 있는 라이다가 주로 이용되고 있다.

[0005] 도 1은 라이다의 밀도와 정확도 사이의 관계를 나타낸 도면이다.

[0006] 도 1에서 (a)는 64 채널의 고채널 라이다에서 스캔된 깊이 맵을 나타내고, (b)는 16 채널의 저채널 라이다에서 스캔된 깊이 맵을 나타내며, (c)는 라이다가 (b)의 저채널 라이다와 동일한 밀도로 균일 분산 스캔을 수행한 경우에 획득되는 깊이 맵을 나타낸다.

[0007] 도 1의 (a)와 (b)를 비교하면, (a)는 64 채널을 이용하여 4.47%의 입력 밀도로 깊이 정보가 획득되어 평균 제곱근 오차(RMSE)는 856인데 반해 (b)는 16 채널을 이용하여 1.16%의 밀도로 깊이 정보가 획득되어 평균 제곱근 오차(RMSE)는 1573이다. 즉 (a)에 도시된 고채널 라이다에서 스캔된 고밀도의 깊이 맵이 (b)의 저밀도 깊이 맵에 비해 확연히 높은 정확도의 깊이 맵을 제공할 수 있음을 알 수 있다. 그러나 라이다는 스캔 밀도가 높아질수록 라이다의 비용이 기하급수적으로 증가되며, 이에 일반적으로 낮은 밀도의 저채널 라이다가 이용되고 있으나, 저채널 라이다는 희박한 스캔 밀도로 인해 장면에서 일부 객체가 스캔되지 않는 객체 누락이 발생할 수 있는 문제가 있다. 특히 대부분의 라이다는 채널 수에 대응하는 라인 단위로 스캔을 수행함에 따라 수직 방향의 희소성이 크게 나타남을 알 수 있다.

[0008] 또한 (b)와 (c)를 비교하면, 동일한 밀도(1.16%)로 스캔이 수행되더라도, 분산 스캔된 (c)의 깊이 맵의 평균 제곱근 오차(RMSE)가 1181로 (b)에 비해 더욱 정확한 깊이 맵을 제공할 수 있음을 알 수 있다. 그러나 상기한 바와 같이, 라이다는 하드웨어의 한계로 인해 (a) 및 (b)에 도시된 바와 같이 라인 단위로 스캔을 수행하며, (c)와 같은 분산 스캔을 수행할 수 없다. 즉 비용을 저감하기 위해 저채널 라이다를 이용하는 경우, 정확한 깊이 정보를 획득하기 어렵다는 한계가 있다.

[0009] 또한 라이다는 장면을 이해하는데 유용한 색상 정보를 제공하지 못한다는 한계가 있다.

[0010] 한편 스테레오 카메라에서 획득되는 스테레오 이미지에 대해 매칭 알고리즘을 적용하여 깊이 정보를 획득하는 경우, 요구되는 고밀도로 깊이 정보와 색상 정보를 함께 획득할 수 있다. 그러나 스테레오 이미지 매칭 알고

리즘으로 획득되는 깊이 정보는 카메라의 위치로부터 거리가 멀어질수록 제공에 비례하여 증가하게 되므로, 현재는 대략 20m 이내까지 깊이를 측정하도록 권장되고 있는 실정이다.

- [0011] 이에 저채널 라이다와 스테레오 카메라를 융합하여 라이다에서 획득되는 저밀도의 최소 깊이 정보를 스테레오 카메라에서 획득되는 고밀도의 깊이 정보로 보완하고자 하는 다양한 융합 기법이 기존에도 제안된 바 있으나, 실외 환경에서는 신뢰할 수 있는 깊이 정보를 제공하지 못하거나, 라이다의 깊이 정보가 없는 영역에서 정확도가 크게 저하되는 한계가 있어, 정확한 깊이 정보 오차를 획득할 수 있는 융합 기법이 계속적으로 요구되고 있다.

## 선행기술문헌

### 특허문헌

- [0012] (특허문헌 0001) 한국 등록 특허 제10-1976290호 (2019.04.30 등록)

## 발명의 내용

### 해결하려는 과제

- [0013] 본 발명의 목적은 저비용의 저채널 라이다와 스테레오 카메라를 이용하여 정확한 깊이 정보를 융합하여 할 수 있는 정확한 깊이 정보를 제공할 수 있는 깊이 추정 장치 및 방법을 제공하는데 있다.

- [0014] 본 발명의 다른 목적은 저채널 라이다에서 획득된 최소 밀도의 깊이 정보와 스테레오 카메라에서 획득된 부정확한 깊이 정보를 컬러 정보를 가이드로 하여 융합함으로써, 정확한 고밀도 깊이 정보를 제공할 수 있는 깊이 추정 장치 및 방법을 제공하는데 있다.

### 과제의 해결 수단

- [0015] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른 깊이 추정 장치는 라이다에서 획득된 깊이 데이터로부터 라이다 입력 레이블을 획득하고, 미리 학습된 패턴 추정 방식에 따라 상기 라이다 입력 레이블에 대해 단계적으로 인코딩하여 라이다 깊이 특징을 추출하는 라이다 깊이 특징 추출부; 스테레오 카메라에서 획득된 디스패리티 맵으로부터 스테레오 입력 레이블을 획득하고, 미리 학습된 패턴 추정 방식에 따라 상기 스테레오 입력 레이블에 대해 단계적으로 인코딩하여 스테레오 깊이 특징을 추출하는 스테레오 깊이 특징 추출부; 상기 스테레오 카메라에서 획득된 컬러 이미지에서 미리 학습된 패턴 추정 방식에 따라 추출되는 가이드 특징과 대응하는 단계에서 추출된 상기 라이다 깊이 특징과 상기 스테레오 깊이 특징을 미리 학습된 방식에 따라 융합하여 융합 특징을 획득하고, 단계적으로 획득된 융합 특징을 인코딩하여 가이드 특징을 단계적으로 추출하는 컬러 가이드 특징 추출부; 및 상기 컬러 가이드 특징 추출부에서 최종 추출된 융합 특징을 인가받아 미리 학습된 특징 복원 방식에 따라 디코딩하여 복원 특징을 획득하고, 획득된 복원 특징을 단계적으로 추출되는 융합 특징 중 대응하는 융합 특징과 결합하여 결합 복원 특징을 획득하며, 획득된 결합 복원 특징으로부터 다시 복원 특징을 획득하여 최종 획득된 복원 특징을 깊이 맵으로 출력하는 디코더를 포함한다.

- [0016] 상기 라이다 깊이 특징 추출부는 상기 라이다에서 획득된 3D 저밀도 깊이 데이터를 2D 이미지 좌표 상에 투영하여 라이다 입력 레이블을 획득하는 라이다 깊이 맵 획득부; 및 다단 구조로 연결되어 상기 라이다 입력 레이블 또는 이전 단계에서 획득된 라이다 깊이 특징을 학습에 의해 획득되는 마스크와 결합하고 미리 학습된 패턴 추정 방식에 따라 인코딩하여 라이다 깊이 특징을 획득하는 다수의 라이다 특징 추출부를 포함할 수 있다.

- [0017] 상기 스테레오 깊이 특징 추출부는 상기 스테레오 카메라에서 획득되는 디스패리티 맵을 인가받아 상기 라이다의 스캔 영역과 매칭시켜 스테레오 깊이 맵을 획득하고, 상기 스테레오 깊이 맵의 밀도를 상기 라이다 입력 레이블에 대응하는 밀도로 변환하여 상기 스테레오 입력 레이블을 획득하는 스테레오 깊이 맵 획득부; 및 다단 구조로 연결되어 상기 스테레오 입력 레이블 또는 이전 단계에서 획득된 스테레오 깊이 특징을 학습에 의해 획득되는 마스크와 결합하고 미리 학습된 패턴 추정 방식에 따라 인코딩하여 스테레오 깊이 특징을 획득하는 다수의 스테레오 특징 추출부를 포함할 수 있다.

- [0018] 상기 컬러 가이드 특징 추출부는 컬러 이미지 또는 이전 단계에서 획득된 융합 특징을 인가받아 미리 학습된 패턴 방식에 따라 가이드 특징을 추출하는 다수의 융합 특징 추출부; 및 상기 다수의 융합 특징 추출부 각각에서 추

출된 가이드 특징을 상기 다수의 라이다 특징 추출부 중 대응하는 라이다 특징 추출부에서 인가되는 라이다 깊이 특징 및 상기 다수의 스테레오 특징 추출부 중 대응하는 스테레오 특징 추출부에서 인가되는 스테레오 깊이 특징과 미리 학습된 방식에 따라 융합하여 융합 특징을 획득하는 다수의 특징 융합부를 포함할 수 있다.

[0019] 상기 다수의 특징 융합부는 대응하는 라이다 특징 추출부와 대응하는 스테레오 특징 추출부 각각 이용된 마스크를 함께 인가받고, 학습에 의해 결정된 가중치에 따라 라이다 깊이 특징과 스테레오 깊이 특징 및 마스크를 융합하여 상기 융합 특징을 출력할 수 있다.

[0020] 상기 디코더는 상기 다수의 융합 특징 추출부의 역순에 대응하여 구성되고, 상기 컬러 가이드 특징 추출부에서 최종 추출된 융합 특징 또는 결합 복원 특징을 인가받아 미리 학습된 특징 복원 방식에 따라 디코딩하여 복원 특징을 출력하는 다수의 디코딩부; 및 상기 다수의 디코딩부 중 대응하는 디코딩부에서 출력되는 복원 특징과 대응하는 융합 특징 추출부에서 전달되는 융합 특징을 결합하여 결합 복원 특징을 출력하는 다수의 결합부를 포함할 수 있다.

[0021] 상기 목적을 달성하기 위한 본 발명의 다른 실시예에 따른 깊이 추정 방법은 라이다에서 획득된 깊이 데이터로부터 라이다 입력 레이블을 획득하고, 미리 학습된 패턴 추정 방식에 따라 상기 라이다 입력 레이블에 대해 단계적으로 인코딩하여 라이다 깊이 특징을 추출하는 단계; 스테레오 카메라에서 획득된 디스패리티 맵으로부터 스테레오 입력 레이블을 획득하고, 미리 학습된 패턴 추정 방식에 따라 상기 스테레오 입력 레이블에 대해 단계적으로 인코딩하여 스테레오 깊이 특징을 추출하는 단계; 상기 스테레오 카메라에서 획득된 컬러 이미지에서 미리 학습된 패턴 추정 방식에 따라 추출되는 가이드 특징과 대응하는 단계에서 추출된 상기 라이다 깊이 특징과 상기 스테레오 깊이 특징을 미리 학습된 방식에 따라 융합하여 융합 특징을 획득하고, 단계적으로 획득된 융합 특징을 인코딩하여 가이드 특징을 추출하는 단계; 및 최종 추출된 융합 특징을 인가받아 미리 학습된 특징 복원 방식에 따라 디코딩하여 복원 특징을 획득하고, 획득된 복원 특징을 단계적으로 추출되는 융합 특징 중 대응하는 융합 특징과 결합하여 결합 복원 특징을 획득하며, 획득된 결합 복원 특징으로부터 다시 복원 특징을 획득하여 최종 획득된 복원 특징을 깊이 맵으로 출력하는 단계를 포함한다.

### 발명의 효과

[0022] 따라서, 본 발명의 실시예에 따른 저채널 라이다와 스테레오 카메라를 이용한 깊이 추정 장치 및 방법은 카메라에서 획득된 컬러 정보를 가이드로 저채널 라이다에서 획득된 회소 밀도의 깊이 정보와 스테레오 카메라에서 획득된 부정확한 깊이 정보를 융합하여 저비용의 저채널 라이다로도 고밀도로 정확한 깊이 맵을 생성할 수 있다.

### 도면의 간단한 설명

[0023] 도 1은 라이다의 밀도와 정확도 사이의 관계를 나타낸 도면이다.  
 도 2는 본 발명의 일 실시예에 따른 깊이 추정 장치의 개략적 구조를 나타낸다.  
 도 2는 본 실시예에 따른 깊이 추정 장치가 라이다와 스테레오 카메라에서 획득된 깊이 맵을 상호 보완하는 개념을 설명하기 위한 도면이다.  
 도 3은 라이다 깊이 맵에서 채널별로 측정된 평균 거리와 스테레오 카메라를 이용하여 감지된 거리에 따른 오차 오차를 나타낸다.  
 도 4는 본 발명의 일 실시예에 따른 깊이 추정 장치의 개략적 구조를 나타낸다.  
 도 5는 도 4의 깊이 추정 장치의 동작을 설명하기 위한 도면이다.  
 도 6은 도 4의 융합 특징 추출부의 동작을 설명하기 위한 도면이다.  
 도 7은 벤치마크 데이터를 이용하여, 본 실시예에 따른 깊이 추정 장치의 성능을 측정한 결과를 나타낸다.  
 도 8은 라이다의 채널 수와 스캔 방향에 따라 획득되는 깊이 맵의 차이를 설명하기 위한 도면이다.  
 도 9는 라이다의 채널 수에 따른 오차를 시뮬레이션한 결과를 나타낸다.  
 도 10은 본 발명의 일 실시예에 따른 깊이 추정 방법을 나타낸다.

### 발명을 실시하기 위한 구체적인 내용

[0024] 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.

- [0025] 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.
- [0026] 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.
- [0027] 도 2는 본 실시예에 따른 깊이 추정 장치가 라이다와 스테레오 카메라에서 획득된 깊이 맵을 상호 보완하는 개념을 설명하기 위한 도면이고, 도 3은 라이다 깊이 맵에서 채널별로 측정된 평균 거리와 스테레오 카메라를 이용하여 감지된 거리에 따른 오차 오차를 나타낸다.
- [0028] 도 2에서 (a)는 라이다에서 채널별 라인으로 측정된 깊이 맵으로 채널에 따라 색상을 다르게 표현하였다. 여기서 라이다의 다수의 채널 중 위쪽에 배치된 채널부터 순차적으로 식별자(ID)를 할당하였으며, 초록색이 낮은 식별자(ID)의 채널을 나타내고, 초록색이 낮은 붉은 색이 높은 식별자(ID)의 채널을 나타낸다. 차량 등에 장착되는 대부분의 라이다에서는 센서 각도에 따라 구면 형상(spherical shape)으로 레이저 광이 방출되므로, 레이저 광을 수평 방향으로 방출하는 채널, 즉 낮은 식별자(ID)를 갖는 채널에서 측정되는 깊이가 일반적으로 더 크게 나타난다.
- [0029] (b)는 스테레오 카메라를 이용하여 획득된 깊이 오차 맵으로 실제 거리와의 오차의 크기에 따라 색상을 다르게 표현하였다. (b)에서 파란색은 오차가 작음을 의미하고, 붉은 색으로 갈수록 오차가 큼을 의미한다.
- [0030] 도 3에서 (a)는 도 2의 (a)에 도시된 바와 같이 라이다에서 측정된 각 채널별로 측정된 거리의 평균을 나타내고, (b)는 도 2의 (b)에 깊이 오차 맵에서 거리에 따른 평균 오차를 나타낸다. 도 2와 도 3의 (b)를 참조하면 스테레오 카메라에서는 거리에 따라 오차가 매우 크게 증가됨을 알 수 있다.
- [0031] 즉 도 2의 (a), (b)와 도 3의 (a), (b)를 살펴보면, 라이다는 거리가 멀어지더라도 정확한 깊이 측정이 가능하지만, 라이다의 채널 수에 따라 측정이 되지 않는 영역의 비율이 매우 높다. 반면, 스테레오 카메라는 장면의 대부분의 영역에 대해 측정이 수행될 수 있으나, 거리가 멀어질수록 측정 오차가 크게 증가한다.
- [0032] 이에 본 실시예에서는 라이다와 스테레오 카메라의 상호 보완적인 특성을 고려하여, 저채널의 라이다가 상대적으로 먼거리를 측정하게 되는 수평 방향을 집중하여 스캔하도록 하고, 나머지 영역의 깊이 정보는 (c)와 같이 스테레오 카메라에서 획득된 데이터를 이용하여 보완함으로써 (d)에 도시된 바와 같이 정확한 깊이 맵을 획득할 수 있도록 한다. 특히 본 실시예에서는 컬러 이미지의 색상 정보를 라이다가 획득한 깊이 정보와 스테레오 카메라가 획득한 깊이 정보를 융합하기 위한 가이드로서 이용함으로써, 정확한 깊이 맵을 추출할 수 있도록 한다.
- [0033] 도 4는 본 발명의 일 실시예에 따른 깊이 추정 장치의 개략적 구조를 나타내고, 도 5는 도 4의 깊이 추정 장치의 동작을 설명하기 위한 도면이며, 도 6은 도 4의 융합 특징 추출부의 동작을 설명하기 위한 도면이다.
- [0034] 도 4 및 도 5를 참조하면, 본 실시예에 따른 깊이 추정 장치는 저채널 라이다에서 획득된 깊이 데이터와 스테레오 카메라에서 획득된 디스패리티 맵 및 컬러 이미지를 기반으로 미리 학습된 패턴 추정 방식에 따라 유효 영역의 깊이 특징을 추출하는 인코더(100)와 인코더(100)에서 추출된 특징을 미리 학습된 방식에 따라 디코딩하여 비유효 영역의 깊이 정보를 보완한 깊이 맵을 생성하는 디코더(200)를 포함할 수 있다.
- [0035] 그리고 인코더(100)는 컬러 가이드 특징 추출부(110), 라이다 깊이 특징 추출부(120) 및 스테레오 깊이 특징 추출부(130)를 포함할 수 있다.
- [0036] 라이다 깊이 특징 추출부(120)는 저채널 라이다가 획득한 저밀도의 깊이 데이터 인가받고, 인가된 저밀도의 깊이 데이터에서 유효 영역의 깊이 특징을 추출한다. 라이다 깊이 특징 추출부(120)는 라이다에서 획득된 저밀도 깊이 데이터로부터 저밀도 깊이 맵인 라이다 입력 레이블( $S_L$ )을 획득하는 라이다 깊이 맵 획득부(121)와 다단 구조로 연결되어 라이다 입력 레이블( $S_L$ )에서 유효 영역의 깊이 특징을 추출하는 다수의 라이다 특징 추출부(122 ~ 124)를 포함할 수 있다.
- [0037] 라이다 깊이 맵 획득부(121)는 저채널 라이다에서 획득된 3D 월드 좌표계 값으로 획득되는 저밀도의 깊이 데이

터를 스테레오 깊이 맵 및 컬러 이미지에 매칭시킬 수 있도록 2D 이미지 좌표로 변환하여, 라이다 입력 레이블( $S_L$ )을 획득한다. 여기서 저채널 라이다는 가능한 수평 방향에서의 깊이 정보를 획득하도록 레이저의 방사 방향이 조절될 수 있다. 이는 상기한 바와 같이, 스테레오 카메라의 경우 먼 거리에 대한 깊이 측정에 대한 오차가 매우 크기 때문에 라이다가 적은 수의 채널로도 가급적 먼 거리의 깊이 값을 정확하게 측정할 수 있도록 하기 위해서이다.

[0038] 라이다 입력 레이블( $S_L$ )은 저밀도의 깊이 데이터의 3D 라이다 포인트 클라우드( $d_L$ )를 수학적 1과 같이 캘리브레이션 매트릭스(calibration matrix)( $H$ ) 및 고유 카메라 메트릭(intrinsic camera metric)( $P$ )에 기반하여, 스테레오 카메라에서 획득되는 한 쌍의 스테레오 이미지 중 하나(여기서는 일예로 좌이미지)( $I_L$ )인 2D 이미지 좌표 상에 투영함으로써 획득될 수 있다.

### 수학적 1

$$[u, v, 1]^T = P H [x, y, z, 1]^T$$

[0039]

[0040] 여기서 ( $x, y, z$ )와 ( $u, v$ )는 각각 3D 월드 좌표(world coordinate)에서 라이다 포인트와 2D 이미지 좌표에서 대응하는 픽셀의 위치를 나타낸다.

[0041] 다수의 라이다 특징 추출부(122 ~ 124)는 도 5에 도시된 바와 같이, 라이다 깊이 맵 획득부(121)에서 획득한 라이다 입력 레이블( $S_L$ ) 또는 이전 단계의 라이다 특징 추출부(122 ~ 123)에서 획득된 라이다 깊이 특징에 대응하는 마스크를 결합하고 학습된 패턴 추정 방식에 따라 인코딩하여 유효한 깊이 특징을 추출한다. 여기서 라이다 입력 레이블( $S_L$ ) 또는 라이다 깊이 특징에 대응하는 마스크를 결합하는 것은 저밀도 깊이 맵인 라이다 입력 레이블( $S_L$ )이 최소한 유효값을 가지기 때문에, 유효값을 갖는 영역의 깊이 특징이 추출되는 반면 비유효한 영역의 값은 깊이 특징으로 추출되지 않도록 하기 위함이다.

[0042] 한편, 스테레오 깊이 특징 추출부(130)는 스테레오 카메라에서 획득되는 디스패리티 맵( $D_S$ )의 깊이 특징을 추출한다. 스테레오 깊이 특징 추출부(130)는 스테레오 카메라에서 획득된 디스패리티 맵( $D_S$ )을 인가받고, 인가된 디스패리티 맵( $D_S$ )을 카메라의 촬영 특성을 고려하여 정렬하여 스테레오 깊이 맵( $\hat{S}_S$ )을 획득하고, 획득된 스테레오 깊이 맵( $\hat{S}_S$ )의 밀도를 라이다 입력 레이블( $S_L$ )에 대응하는 밀도를 갖도록 변환하여 저밀도 스테레오 깊이 맵인 스테레오 입력 레이블( $S_S$ )을 획득하는 스테레오 깊이 맵 획득부(131)와 다단 구조로 연결되어 스테레오 입력 레이블( $S_S$ )에서 유효 영역의 깊이 특징을 추출하는 다수의 스테레오 특징 추출부(132 ~ 134)를 포함할 수 있다.

[0043] 스테레오 깊이 맵 획득부(131)는 스테레오 카메라에서 좌영상 및 우영상 사이의 차이를 기반으로 획득한 디스패리티 맵( $D_S$ )에 대해 스테레오 카메라의 수평 초점 거리( $f_u$ ) 및 기준선( $b_s$ )을 적용하여 수학적 2에 따라 스테레오 깊이 맵( $\hat{S}_S$ )을 획득한다. 이는 스테레오 카메라에서 획득되는 디스패리티 맵( $D_S$ )을 라이다가 스캔한 영역과 매칭시키기 위해서이다.

### 수학적 2

$$\hat{S}_S(u, v) = b_s f_u / D_S(u, v)$$

[0044]

[0045] 다만 스테레오 깊이 맵( $\hat{S}_S$ )은 저밀도 깊이 맵인 라이다 입력 레이블( $S_L$ )에 비해 고밀도 깊이 정보를 포함하고 있으므로, 이후 깊이 정보의 융합 과정에서 스테레오 깊이 맵( $\hat{S}_S$ )의 고밀도 정보는 라이다 입력 레이블( $S_L$ )보

다 크게 영향을 미칠 수 있다. 이에 스테레오 깊이 맵 획득부(131)는 랜덤하게 생성되는 초기 마스크( $\hat{M}_S$ )를 수학적 식 3과 같이 스테레오 깊이 맵( $\hat{S}_S$ )에 적용하여 라이다 입력 레이블( $S_L$ )에 대응하는 밀도를 갖는 스테레오 입력 레이블( $S_S$ )을 획득한다.

### 수학적 식 3

$$S_S = \hat{M}_S \cdot \hat{S}_S$$

여기서  $\cdot$ 는 원소별 곱셈 연산자이다.

다수의 스테레오 특징 추출부(132 ~ 134)는 스테레오 깊이 맵 획득부(131)에서 획득한 스테레오 입력 레이블( $S_S$ ) 또는 이전 단계의 스테레오 특징 추출부(132 ~ 133)에서 획득된 스테레오 깊이 특징에 대응하는 마스크를 결합하여 유효한 깊이 특징을 추출한다. 여기서 스테레오 입력 레이블( $S_S$ ) 또는 스테레오 깊이 특징에 대응하는 마스크를 결합하는 것은 라이다 입력 레이블( $S_L$ )에 대응하여 저밀도로 변환된 스테레오 입력 레이블( $S_S$ )이 최소한 유효값을 가지기 때문에, 유효값을 갖는 영역의 깊이 특징을 추출하기 위해서이다.

여기서 상기한 다수의 라이다 특징 추출부(122 ~ 124)와 다수의 스테레오 특징 추출부(132 ~ 134)는 인공 신경망으로서, 특히 각각 부분 컨볼루션(Partial Convolution: PC) 레이어로 구현될 수 있다.

부분 컨볼루션은 학습시에 각 레이어의 가중치뿐만 아니라 마스크를 함께 학습시켜, 유효하지 않은 픽셀 영역의 영향을 최소화할 수 있도록 구성되는 컨볼루션 레이어로서 수학적 식 4에 따라 부분 컨볼루션 필터링 동작을 수행하도록 구성될 수 있다.

### 수학적 식 4

$$\varphi(u, v) = \frac{\sum_{\mu, \nu=-k}^k m(u + \mu, v + \nu) x(u + \mu, v + \nu) w(\mu, \nu)}{\sum_{\mu, \nu=-k}^k m(u + \mu, v + \nu)} + b$$

여기서  $X$ 는 픽셀( $u, v$ )별 특성값(깊이값)( $x$ )을 원소로 갖는 부분 컨볼루션 레이어의 입력인 2차원 행렬로서 라이다 입력 레이블( $S_L$ ) 또는 라이다 깊이 특징이거나, 스테레오 입력 레이블( $S_S$ ) 또는 스테레오 깊이 특징이다. 그리고  $W$ 는 가중치( $w$ )를 원소로 갖는 가중치 행렬이고,  $b$ 는 바이어스(bias) 값이다. 한편  $M$ 은  $m$ 을 원소로 갖는 마스크 행렬이다. 라이다 입력 레이블( $S_L$ ) 또는 라이다 깊이 특징과 스테레오 입력 레이블( $S_S$ ) 또는 스테레오 깊이 특징에 대응하도록 학습되는 마스크 행렬( $M$ )의 마스크 원소( $m$ )는 수학적 식 5에 따라 업데이트 될 수 있다.

### 수학적 식 5

$$m'(u, v) = \begin{cases} 1 & \text{if } \sum_{\mu, \nu=-k}^k m(u + \mu, v + \nu) > 0 \\ 0 & \text{otherwise} \end{cases}$$

수학적 식 5에 따르면, 마스크 행렬( $M$ )의 원소( $m$ )는 주변 원소의 값과의 합에 따라 1 또는 0으로 설정될 수 있다.

다만 수학적 식 5에 따라 마스크 행렬( $M$ )의 위치( $u, v$ )에서의 마스크 원소( $m(u, v)$ )가 0으로 업데이트 되면, 수학적 식 4에서 0으로 나누어지는 오류가 발생되지 않도록, 부분 컨볼루션 레이어의 출력(패( $u, v$ )))을 0 벡터로 설정한다.

- [0056] 여기서 제1 및 제2 특징 추출부(122, 123, 132, 133)을 제외한 나머지 특징 추출부(124, 134)는 2의 스트라이드(stride)를 갖도록 설정될 수 있다.
- [0057] 도 4에서는 설명의 편의를 위하여 라이다 깊이 특징 추출부(120)와 스테레오 깊이 특징 추출부(130)가 각각 3개의 특징 추출부(122 ~ 124, 132 ~ 134)를 포함하는 것으로 도시하였으나, 특징 추출부의 개수는 다양하게 조절될 수 있다. 이에 도 5에서는 라이다 깊이 특징 추출부(120)와 스테레오 깊이 특징 추출부(130)가 각각 5개의 특징 추출부를 포함하는 것으로 가정하여 도시하였다.
- [0058] 다만, 라이다 깊이 특징 추출부(120)와 스테레오 깊이 특징 추출부(130)는 동일한 개수( $N_0$ )의 특징 추출부를 포함해야 하며, 컬러 가이드 특징 추출부(110)의 융합 특징 추출부(112 ~ 124) 및 디코더(200)의 디코딩부(212 ~ 214)의 개수 또한 동일하게 포함되어야 한다.
- [0059] 한편, 라이다 깊이 특징 추출부(120)의 다수의 라이다 특징 추출부(122 ~ 124) 및 스테레오 깊이 특징 추출부(130)의 다수의 스테레오 특징 추출부(132 ~ 134) 각각은 추출된 깊이 특징을 다음 단의 특징 추출부로 전달할 뿐만 아니라, 컬러 가이드 특징 추출부(110)의 대응하는 특징 융합부(F1 ~ F3)로도 전달한다.
- [0060] 라이다 특징 추출부(122 ~ 124)에서 출력되는 라이다 중간 특징( $X_L^o$ )은 수학적 4에 따라  $X_L^o = \Phi_L^o(X_L^{o-1}, M_L^{o-1})$ 로 표현될 수 있으며, 스테레오 특징 추출부(132 ~ 134)에서 출력되는 스테레오 중간 특징( $X_S^o$ ) 또한 유사하게 표현될 수 있다. 여기서 o는 특징 추출부(122 ~ 124, 132 ~ 134)의 인덱스를 나타내며, 라이다 중간 특징( $X_L^o$ )과 스테레오 중간 특징( $X_S^o$ ) 각각의 원소는  $x_L^o$  및  $x_S^o$ 로 표현될 수 있다.
- [0061] 또한 라이다 깊이 특징 추출부(120)의 다수의 라이다 특징 추출부(122 ~ 124) 및 스테레오 깊이 특징 추출부(130)의 다수의 스테레오 특징 추출부(132 ~ 134) 각각은 중간 특징( $X_L^o, X_S^o$ )과 함께 대응하는 마스크( $M_L, M_S$ )를 컬러 가이드 특징 추출부(110)의 대응하는 특징 융합부(F1 ~ F3)로도 전달한다.
- [0062] 도 5에 도시된 바와 같이, 라이다 특징 추출부와 스테레오 특징 추출부의 개수가 각각 5개인 경우, 라이다 특징 추출부와 스테레오 특징 추출부는 표 1과 같은 부분 컨볼루션 레이어로 구현될 수 있다.

표 1

Input	Module	Kernel	Channel	Stride	Output
LiDAR depth network $\Phi_L$					
$S_L, M_L^0$	PC, LN, LRelu	$3 \times 3$	1/16	1	$X_L^1, M_L^1$
$X_L^1, M_L^1$	PC, LN, LRelu	$3 \times 3$	16/16	1	$X_L^2, M_L^2$
$X_L^2, M_L^2$	PC, LN, LRelu	$3 \times 3$	16/32	2	$X_L^3, M_L^3$
$X_L^3, M_L^3$	PC, LN, LRelu	$3 \times 3$	32/64	2	$X_L^4, M_L^4$
$X_L^4, M_L^4$	PC, LN, LRelu	$3 \times 3$	64/128	2	$X_L^5, M_L^5$
Stereo depth network $\Phi_S$					
$S_S, M_S^0$	PC, LN, LRelu	$3 \times 3$	1/16	1	$X_S^1, M_S^1$
$X_S^1, M_S^1$	PC, LN, LRelu	$3 \times 3$	16/16	1	$X_S^2, M_S^2$
$X_S^2, M_S^2$	PC, LN, LRelu	$3 \times 3$	16/32	2	$X_S^3, M_S^3$
$X_S^3, M_S^3$	PC, LN, LRelu	$3 \times 3$	32/64	2	$X_S^4, M_S^4$
$X_S^4, M_S^4$	PC, LN, LRelu	$3 \times 3$	64/128	2	$X_S^5, M_S^5$

[0063]

[0064]

[0065]

표 1에서 LN은 레이어 정규화 구성을 나타내고, LReLU는 누설 정류 선형 유닛을 의미한다.

컬러 가이드 특징 추출부(110)는 스테레오 카메라에서 획득된 컬러 이미지를 인가받아 컬러 이미지의 컬러 정보를 기반으로 라이다 깊이 특징 추출부(120)에서 추출되는 라이다 깊이 특징과 스테레오 깊이 특징 추출부(130)

에서 추출되는 스테레오 깊이 특징을 융합하여 융합 특징을 추출한다.

[0066] 컬러 정보를 라이다 깊이 특징과 스테레오 깊이 특징을 융합하고자 하는 경우 다수의 픽셀 중 어느 픽셀의 깊이 정보가 중요한지 고려될 필요가 있으며, 이에 컬러 가이드 특징 추출부(110)는 컬러 정보를 라이다 깊이 특징과 스테레오 깊이 특징의 융합시에 선택적으로 융합되도록 하여, 최적의 깊이 특징을 추출하도록 한다.

[0067] 컬러 가이드 특징 추출부(110)는 컬러 이미지 획득부(111)와 다수의 융합 특징 추출부(112 ~ 114) 및 다수의 특징 융합부(F1 ~ F3)를 포함할 수 있다. 컬러 이미지 획득부(111)는 스테레오 카메라가 획득한 컬러 이미지인 좌 이미지( $I_L$ )와 우 이미지( $I_R$ ) 중 하나를 획득한다. 여기서는 일예로 좌 이미지( $I_L$ )를 획득하는 것으로 가정하여 설명하지만, 우 이미지( $I_R$ )를 획득하여도 무방하다.

[0068] 다수의 융합 특징 추출부(112 ~ 114)는 컬러 이미지 또는 이전단의 특징 융합부(F1 ~ F3)에서 출력되는 융합 특징을 인가받아 미리 학습된 패턴 방식에 따라 가이드 특징을 추출한다.

[0069] 여기서 다수의 융합 특징 추출부(112 ~ 114)는 개수가 증가되더라도 라이다 깊이 특징과 스테레오 깊이 특징의 융합을 가이드를 위해 이용되는 컬러 이미지의 특징을 유지하고 정확하게 학습될 수 있도록, 도 5에 도시된 바와 같이, 입력되는 융합 특징을 스킵 커넥션(skip connection)을 통해 출력되는 가이드 특징과 다시 결합(concatenation)하여 출력하는 잔류 네트워크(residual network: 이하 ResNet) 구조로 구성될 수 있다.

[0070] 다수의 특징 융합부(F1 ~ F3)는 대응하는 융합 특징 추출부(112 ~ 114)에서 추출된 가이드 특징과 라이다 깊이 특징 추출부(120)의 대응하는 라이다 특징 추출부(122 ~ 124)에서 인가되는 라이다 깊이 특징 및 스테레오 깊이 특징 추출부(130)의 대응하는 스테레오 특징 추출부(132 ~ 134)에서 인가되는 스테레오 깊이 특징을 학습된 방식으로 선택적 융합하여 융합 특징을 출력한다.

[0071] 이때 다수의 특징 융합부(F1 ~ F3)는 다수의 라이다 특징 추출부(122 ~ 124) 및 다수의 스테레오 특징 추출부(132 ~ 134)로부터 깊이 특징과 함께 대응하는 라이다 마스크( $M_L$ ) 및 스테레오 마스크( $M_S$ )도 함께 인가받고, 깊이 특징과 마스크를 함께 융합하여 융합 특징을 획득할 수 있다.

[0072] 도 4 및 도 5에서는 설명의 편의를 위해, 다수의 특징 융합부(F1 ~ F3)를 다수의 융합 특징 추출부(112 ~ 114)와 별도로 도시하였으나, 다수의 특징 융합부(F1 ~ F3)는 다수의 융합 특징 추출부(112 ~ 114)에 포함되어 구성될 수 있다.

[0073] 다수의 특징 융합부(F1 ~ F3)는 도 6에 도시된 바와 같이, 대응하는 융합 특징 추출부(112 ~ 114)에서 추출된 가이드 특징( $X_G^o$ )과 대응하는 라이다 특징 추출부(122 ~ 124)에서 인가되는 라이다 깊이 특징( $X_L^o$ ), 라이다 마스크( $M_L^o$ ) 및 다수의 스테레오 특징 추출부(132 ~ 134)에서 인가되는 스테레오 깊이 특징( $X_S^o$ ), 스테레오 마스크( $M_S^o$ )를 선택적으로 융합하여 융합 특징( $X_F^o$ )을 출력한다.

[0074] 특징 융합부(F1 ~ F3)는 수식식 6에 따라 특징( $X_G^o$ ,  $X_L^o$ ,  $X_S^o$ )들과 마스크들( $M_L^o$ ,  $M_S^o$ )을 융합하여 융합 특징( $X_F^o$ )을 출력할 수 있다.

## 수식식 6

$$\dot{x}_F^o = \frac{\dot{x}_G^o \dot{w}_G^o + \dot{m}_L^o \dot{x}_L^o \dot{w}_L^o + \dot{m}_S^o \dot{x}_S^o \dot{w}_S^o}{1 + \dot{m}_L^o + \dot{m}_S^o} + b$$

[0075]

[0076] 여기서  $\dot{x}_F^o$ 는 픽셀(u,v)에서의 융합 특징이고,  $\dot{x}_G^o$ ,  $\dot{x}_L^o$ ,  $\dot{x}_S^o$ 는 각각 픽셀(u,v)에 대한 가이드 특징과 라이다 깊이 특징 및 스테레오 깊이 특징이며,  $\dot{m}_L^o$ ,  $\dot{m}_S^o$ 는 각각 픽셀(u,v)에 대한 라이다 마스크 및 스테레오 마스크를 나타낸다. 그리고  $\dot{w}_G^o$ ,  $\dot{w}_L^o$ ,  $\dot{w}_S^o$ 는 각각 픽셀(u,v)에 대한 가이드 가중치와 라이다 가중치 및 스테레오



가중치이다.

- [0077] 수학식 6에서  $\dot{w}_L^o$ ,  $\dot{w}_S^o$ 는 도 6에 도시된 바와 같이, 라이다 깊이 특징 및 스테레오 깊이 특징에 대한 가중치 ( $\dot{w}_{XL}^o$ ,  $\dot{w}_{XS}^o$ )와 라이다 마스크 및 스테레오 마스크에 대한 가중치( $\dot{w}_{ML}^o$ ,  $\dot{w}_{MS}^o$ )로 구분될 수 있으나, 여기서는 통합된 가중치로 표현하였다.
- [0078] 본 실시예에서 인코더(100)의 라이다 깊이 특징 추출부(120) 및 스테레오 깊이 특징 추출부(130)가 각각 독립적으로 라이다 깊이 맵의 깊이 특징과 스테레오 깊이 맵의 특징을 다단 구조로 추출하고, 각 단에서 추출된 깊이 특징을 별도로 다단 구조로 구성된 컬러 가이드 특징 추출부(110)의 대응하는 단으로 전달하여 융합되도록 하는 것은 라이다를 통해 획득되는 라이다 깊이 맵과 스테레오 카메라를 통해 획득되는 스테레오 깊이 맵의 깊이 특징을 가능한 유지하면서도 색상 정보를 가이드로 하여 각 단계별 반복 융합되도록 하여 깊이 특징이 정확하게 추출될 수 있도록 하기 위함이다.
- [0079] 디코더(200)는 인코더(100)에서 획득된 융합 특징( $X_F^o$ )을 인가받아 디코딩하여 깊이 맵을 획득한다. 디코더(200)는 다수의 디코딩부(212 ~ 214)와 다수의 결합부(C1, C2) 및 깊이 맵 출력부(211)를 포함할 수 있다.
- [0080] 다수의 디코딩부(212 ~ 214)는 각각 다수의 융합 특징 추출부(112 ~ 114)의 역순에 대응하여 구성되고, 학습된 특징 복원 방식에 따라 인코더(100)에서 인가되는 융합 특징( $X_F^o$ ) 또는 결합부(C1, C2)에서 출력되는 결합 복원 특징을 순차적으로 디코딩하여 복원 특징을 출력한다.
- [0081] 다수의 결합부(C1, C2)는 대응하는 디코딩부(213, 214)에서 출력되는 복원 특징을 대응하는 특징 융합부(F2, F3)에서 출력되는 융합 특징과 결합하여 다음 단의 디코딩부(212, 213)로 결합 복원 특징을 출력한다. 다수의 결합부(C1, C2)는 대응하는 특징 융합부(F2, F3)에서 출력되는 융합 특징을 스킵 커넥션을 통해 인가받아 복원 특징과 결합하여 출력함으로써, 디코딩부(212 ~ 214)는 깊이 특징이 추출되기 이전의 정보를 함께 고려하여 정확한 복원을 수행할 수 있게 된다.
- [0082] 깊이 맵 출력부(211)는 다수의 디코딩부(212 ~ 214)에 의해 복원된 깊이 맵을 출력한다. 여기서 복원된 깊이 맵은 최소한 저밀도의 라이다 깊이 정보와 스테레오 깊이 정보를 융합하고 컬러 이미지의 색상 정보를 기반으로 고밀도로 보완한 깊이 맵으로 출력된다.
- [0083] 도 5에 도시된 바와 같이, 융합 특징 추출부와 디코딩부의 개수가 각각 5개인 경우, 융합 특징 추출부와 디코딩부는 표 2와 같이 ResNet 블록 및 컨볼루션 레이어로 구현될 수 있다.

## 표 2

Input	Module	Kernel	Channel	Stride	Output
Encoder of Color Guidance Network $\Phi_G$					
$I_l$	Conv, BN, Relu	$3 \times 3$	3/32	1	$X_G^1$
$X_G^1, X_L^1, M_L^1, X_S^1, M_S^1$	SF	$1 \times 1$	64/64	1	$X_F^1$
$X_F^1$	ResBlock	$3 \times 3$	64/64	1	$X_G^{2a}$
$X_G^{2a}$	ResBlock	$3 \times 3$	64/64	1	$X_G^{2b}$
$X_G^{2b}$	ResBlock	$3 \times 3$	64/64	1	$X_G^2$
$X_G^2, X_L^2, M_L^2, X_S^2, M_S^2$	SF	$1 \times 1$	96/64	1	$X_F^2$
$X_F^2$	ResBlock	$3 \times 3$	64/64	2	$X_G^{3a}$
$X_G^{3a}$	ResBlock	$3 \times 3$	64/64	1	$X_G^{3b}$
$X_G^{3b}$	ResBlock	$3 \times 3$	64/64	1	$X_G^{3c}$
$X_G^{3c}$	ResBlock	$3 \times 3$	64/64	1	$X_G^3$
$X_G^3, X_L^3, M_L^3, X_S^3, M_S^3$	SF	$1 \times 1$	192/128	1	$X_F^3$
$X_F^3$	ResBlock	$3 \times 3$	128/128	2	$X_G^{4a}$
$X_G^{4a}$	ResBlock	$3 \times 3$	128/128	1	$X_G^{4b}$
$X_G^{4b}$	ResBlock	$3 \times 3$	128/128	1	$X_G^{4c}$
$X_G^{4c}$	ResBlock	$3 \times 3$	128/128	1	$X_G^{4d}$
$X_G^{4d}$	ResBlock	$3 \times 3$	128/128	1	$X_G^{4e}$
$X_G^{4e}$	ResBlock	$3 \times 3$	128/128	1	$X_G^4$
$X_G^4, X_L^4, M_L^4, X_S^4, M_S^4$	SF	$1 \times 1$	384/256	1	$X_F^4$
$X_F^4$	ResBlock	$3 \times 3$	256/256	2	$X_G^{5a}$
$X_G^{5a}$	ResBlock	$3 \times 3$	256/256	1	$X_G^{5b}$
$X_G^{5b}$	ResBlock	$3 \times 3$	256/256	1	$X_G^5$
$X_G^5, X_L^5, M_L^5, X_S^5, M_S^5$	SF	$1 \times 1$	748/512	1	$X_F^5$
$X_F^5$	Conv, BN, Relu	$3 \times 3$	512/512	1	$X_F^5$
Decoder of Color Guidance Network $\Phi_G$					
$X_F^{5a}$	Conv, BN, Relu	$3 \times 3$	512/512	2	$X_D^4$
$X_D^4, X_F^4$	Concat	—	768/768	1	$X_D^{4a}$
$X_D^{4a}$	Conv, BN, Relu	$1 \times 1$	768/128	1	$X_D^{4b}$
$X_D^{4b}$	Conv, BN, Relu	$3 \times 3$	128/128	2	$X_D^3$
$X_D^3, X_F^3$	Concat	—	256/256	1	$X_D^{3a}$
$X_D^{3a}$	Conv, BN, Relu	$1 \times 1$	256/64	1	$X_D^{3b}$
$X_D^{3b}$	Conv, BN, Relu	$3 \times 3$	64/64	2	$X_D^2$
$X_D^2, X_F^2$	Concat	—	128/128	1	$X_D^{2a}$
$X_D^{2a}$	Conv, BN, Relu	$1 \times 1$	128/32	1	$X_D^{2b}$
$X_D^{2b}$	Conv, BN, Relu	$3 \times 3$	32/32	2	$X_D^1$
$X_D^1, X_F^1$	Concat	—	96/96	1	$X_D^{1a}$
$X_D^{1a}$	Conv, BN, Relu	$1 \times 1$	96/16	1	$X_D^{1b}$
$X_D^{1b}$	Conv	$1 \times 1$	16/1	1	$S_G$

[0084]

[0085]

한편 도 4에 도시된 본 실시예의 깊이 추정 장치는 인공 신경망으로 구현되므로, 학습이 수행되어야 하며, 학습은 미리 깊이 값이 측정되어 레이블된 컬러 이미지를 학습용 데이터로 이용하여 수행될 수 있다. 그리고 학습시에 손실은 깊이 맵 출력부(211)에서 출력되는 깊이 맵의 깊이 값( $S_G$ )과 학습용 데이터의 깊이값( $S_V$ ) 사이의 차이를  $L_2$  놈(Norm)으로 계산하여 수학적식 7과 같이 획득될 수 있다.

## 수학적식 7

$$\mathcal{L}_{L_2} = \sum_{p \in \Omega(S_V)} |S_G(p) - S_V(p)|_2$$

[0086]

[0087]

여기서  $p$ 는 깊이 맵에서의 좌표( $u, v$ )를 나타내고,  $\Omega(S_V)$ 는 유효 깊이값을 포함한 좌표 집합을 나타낸다.  $\Omega(S_V)$ 는 현재 제공될 수 있는 학습용 데이터에도 대부분 누락된 값이 있으므로, 유효한 위치에서의 손실만을 평가하도록 하기 위해서이다.

- [0088] 수학적식 7에 따라 손실이 계산되면, 계산된 손실을 디코더(200) 및 인코더(100)로 역전파하여 학습을 수행한다.
- [0089] 도 7은 벤치마크 데이터를 이용하여, 본 실시예에 따른 깊이 추정 장치의 성능을 측정한 결과를 나타낸다.
- [0090] 도 7에서 KITTI(Karlsruhe Institute of Technology and Toyota Technological Institute) 벤치마크 데이터를 이용하였으며, (a)는 입력되는 컬러 이미지를 나타내고, (b)는 저채널 라이다에서 획득된 깊이 데이터를 나타내고, (c)는 다수의 스테레오 특징 추출부(132 ~ 134)가 추출된 스테레오 특징을 각각 대응하는 특징 융합부(F1 ~ F3)로 전송하지고 깊이 정보를 추출한 결과를 나타내며, (d)는 디코더(200)의 다수의 결합부(C1, C2)가 복원 특징을 융합 특징과 결합하지 않는 경우를 나타낸다. 그리고 (e)는 본 실시예에 따라 획득된 깊이 맵을 나타내며, (f)는 고채널 라이다 등으로 실제 깊이가 측정된 깊이 맵을 나타낸다.
- [0091] 도 7의 (b)에 도시된 바와 같이, 본 실시예에서는 저채널 라이다의 모든 채널이 가능한 원거리를 측정할 수 있도록 수평 방향에 집중하여 스캔하도록 구성하였으며, 저채널에서 라이다에서 스캔되지 않은 영역에 대해서는 스테레오 카메라에서 획득된 깊이 정보와 컬러 정보를 이용하여 보완되도록 하였다.
- [0092] 표 3은 도 7의 결과를 정량적으로 평가한 결과를 나타낸다.

표 3

Methods	RMSE (mm)	MAE (mm)
Sparse-to-Dense [38] w/8ch.	1985.51	1057.54
Ours wo/Section III-B	1487.93	688.39
Ours wo/(3)	1450.75	681.19
Ours wo/ $\Phi_S$	1504.60	802.61
Ours wo/ $\Phi_D$	2309.74	1115.19
Ours wo/DS	1424.51	680.54
Ours	<b>1368.67</b>	<b>651.96</b>

- [0093]
- [0094] 표 3은 종래의 깊이 추정 기법과 본 실시예에서 라이다 깊이 맵 획득부(121)가 라이다 입력 레이블( $S_L$ ) 또는 스테레오 입력 레이블( $S_S$ )을 획득하지 않는 경우, 그리고 다수의 스테레오 특징 추출부(132 ~ 134)가 추출된 스테레오 특징을 각각 대응하는 특징 융합부(F1 ~ F3)로 전송하지 않거나 디코더(200)의 다수의 결합부(C1, C2)가 복원 특징을 융합 특징과 결합하지 않는 경우 및 다수의 특징 융합부(F1 ~ F3)가 가이드 특징과 라이다 깊이 특징 및 스테레오 깊이 특징을 선택적 융합하지 않는 경우 각각을 비교하여 성능을 평가한 결과를 나타낸다.
- [0095] 도 7 및 표 3에 도시된 바와 같이, 본 실시예에 따른 깊이 추정 장치는 평균제곱근오차(Root Mean Square Error: RMSE) 및 평균절대값오차(Mean Absolute Error: MAE)가 다른 경우에 비해 낮음을 알 수 있다. 즉 정확한 깊이를 측정할 수 있음을 알 수 있다.
- [0096] 도 8은 라이다의 채널 수와 스캔 방향에 따라 획득되는 깊이 맵의 차이를 설명하기 위한 도면이다.
- [0097] 도 8에서 (a)는 컬러 이미지와 실측된 깊이 맵을 나타내며, (b)는 고채널 라이다에서 획득된 라이다 깊이 데이터를 본 실시예에 적용한 결과를 나타낸다.
- [0098] 그리고 (c)는 저채널 라이다의 각 채널이 서로 균등 각도로 분산되어 스캔을 수행한 라이다 깊이 데이터와 이를 적용하여 획득되는 깊이 맵을 나타내고, (d) 내지 (f)는 저채널 라이다가 집중 스캔을 수행하는 스캔 방향에 따른 깊이 데이터와 이를 적용하여 획득되는 깊이 맵을 나타낸다.
- [0099] 도 8에서 (f)에 도시된 바와 같이 저채널 라이다를 수평 방향에 집중하여 스캔하는 경우, (c)와 같이 분산 스캔을 수행하는 경우나 (f) 지면 방향으로 집중하여 스캔을 수행하는 경우에 비해, 매우 정확도가 높아져서 (b)에 도시된 고채널 라이다를 이용하는 경우와 유사한 결과를 도출할 수 있음을 알 수 있다.
- [0100] 도 9는 라이다의 채널 수에 따른 오차를 시뮬레이션한 결과를 나타낸다.





F1 ~ F3: 특징 융합부

120: 라이다 깊이 특징 추출부

121: 라이다 깊이 맵 획득부

122 ~ 124: 라이다 특징 추출부

130: 스테레오 깊이 특징 추출부

131: 스테레오 깊이 맵 획득부

132 ~ 134: 스테레오 특징 추출부

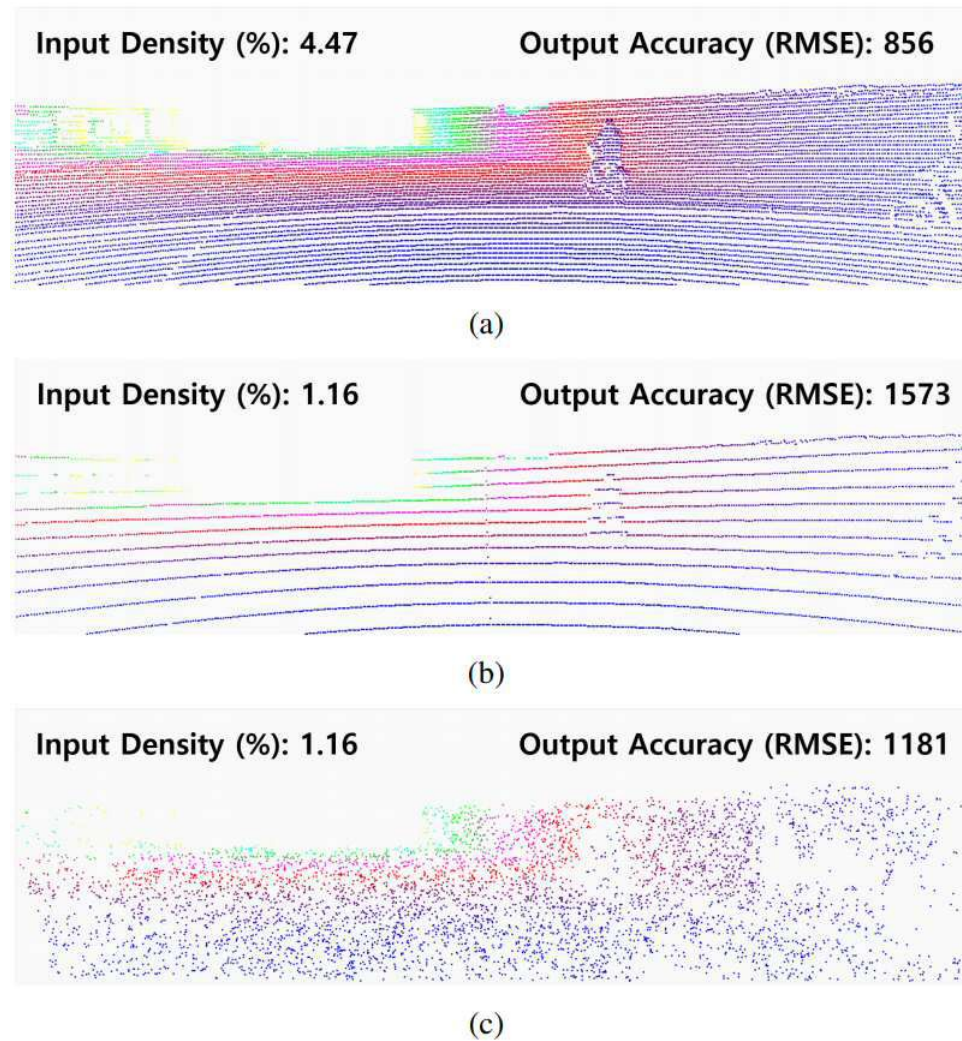
212 ~ 214: 디코딩부

C1, C2: 결합부

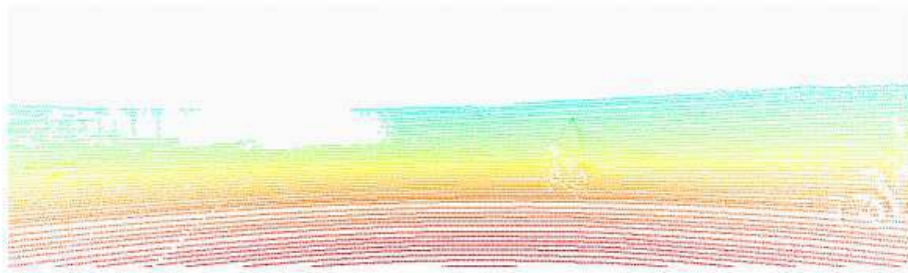
211: 깊이 맵 출력부

도면

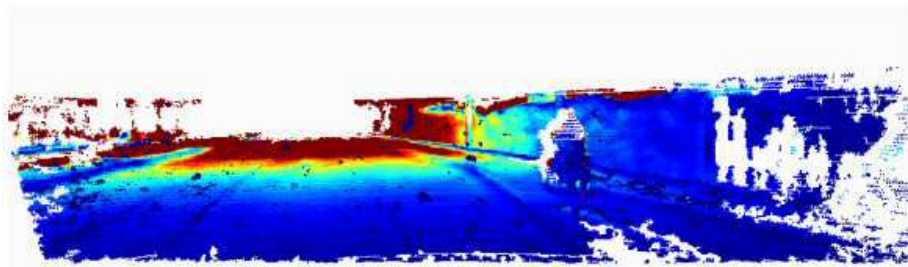
도면1



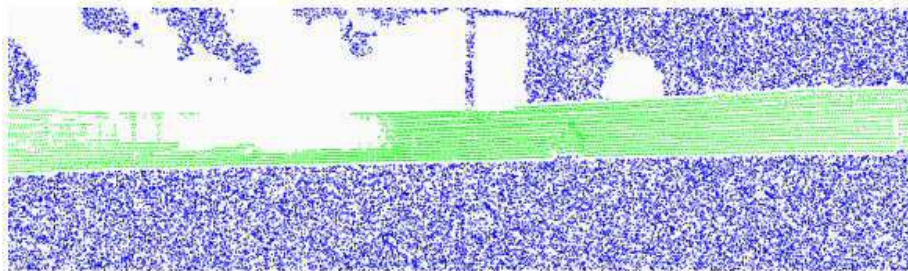
도면2



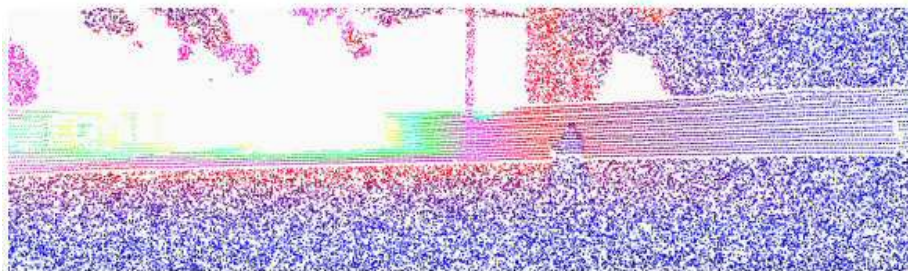
(a)



(b)



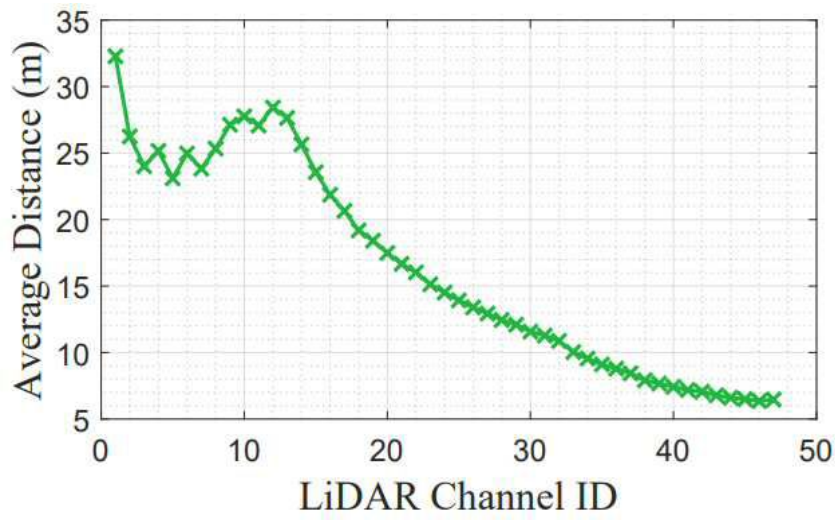
(c)



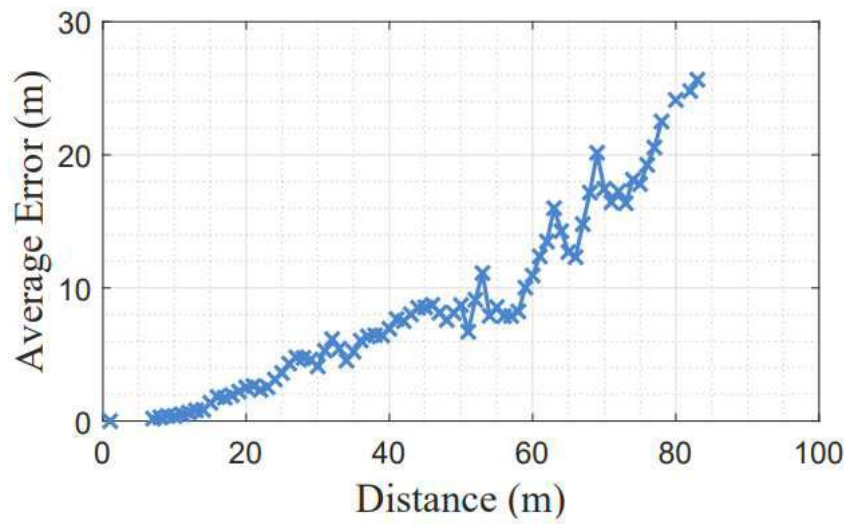
(d)



도면3

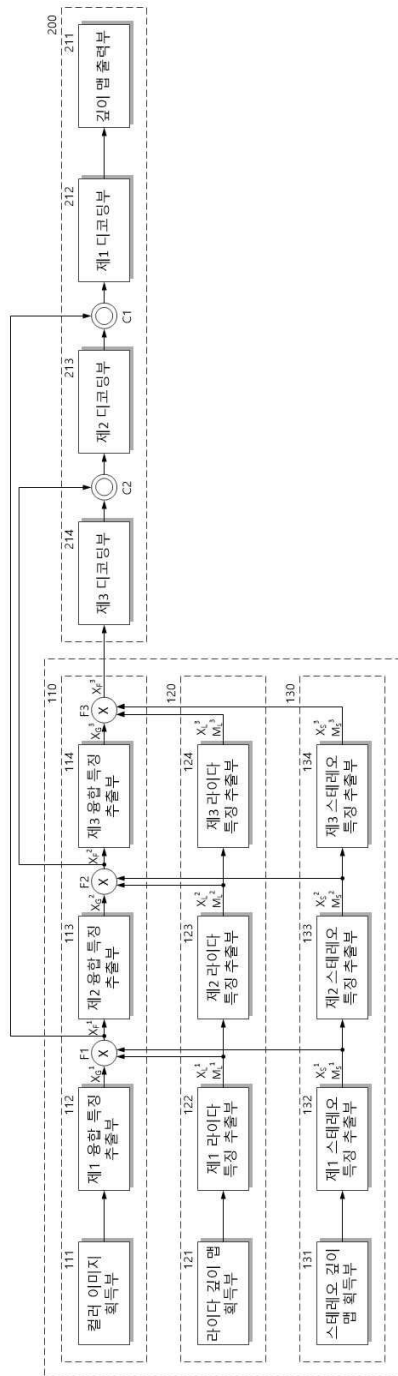


(a)

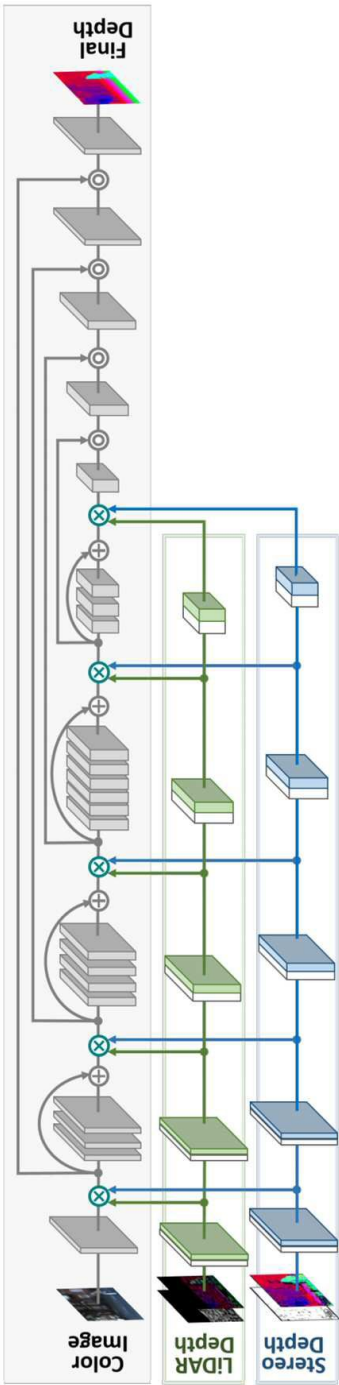


(b)

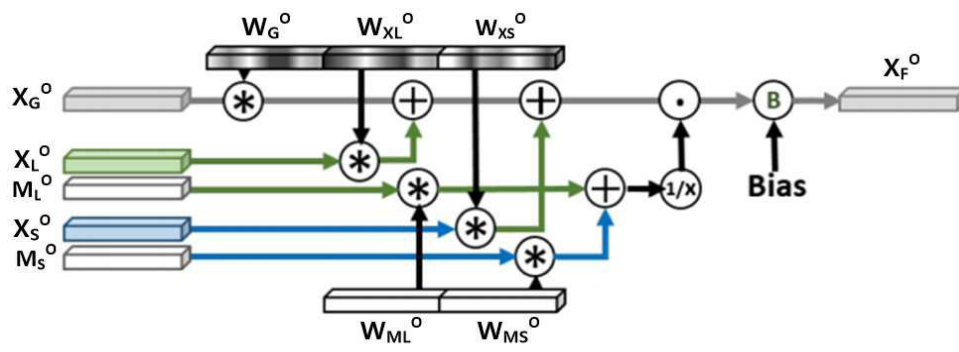
도면4



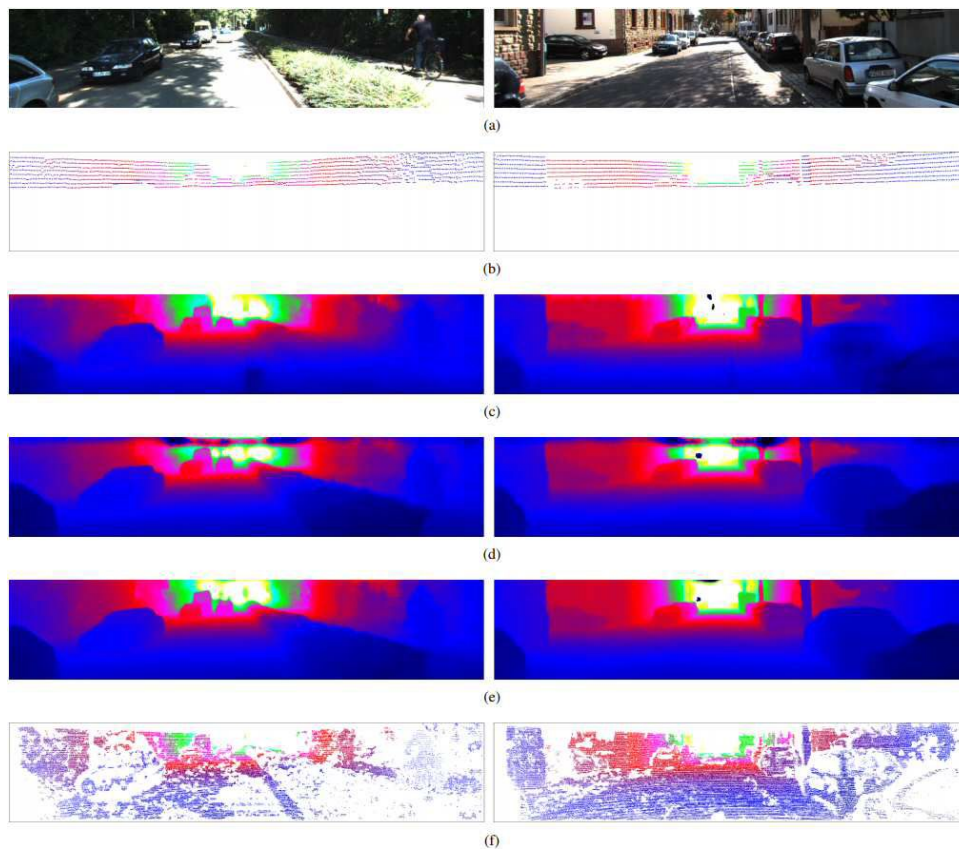
도면5



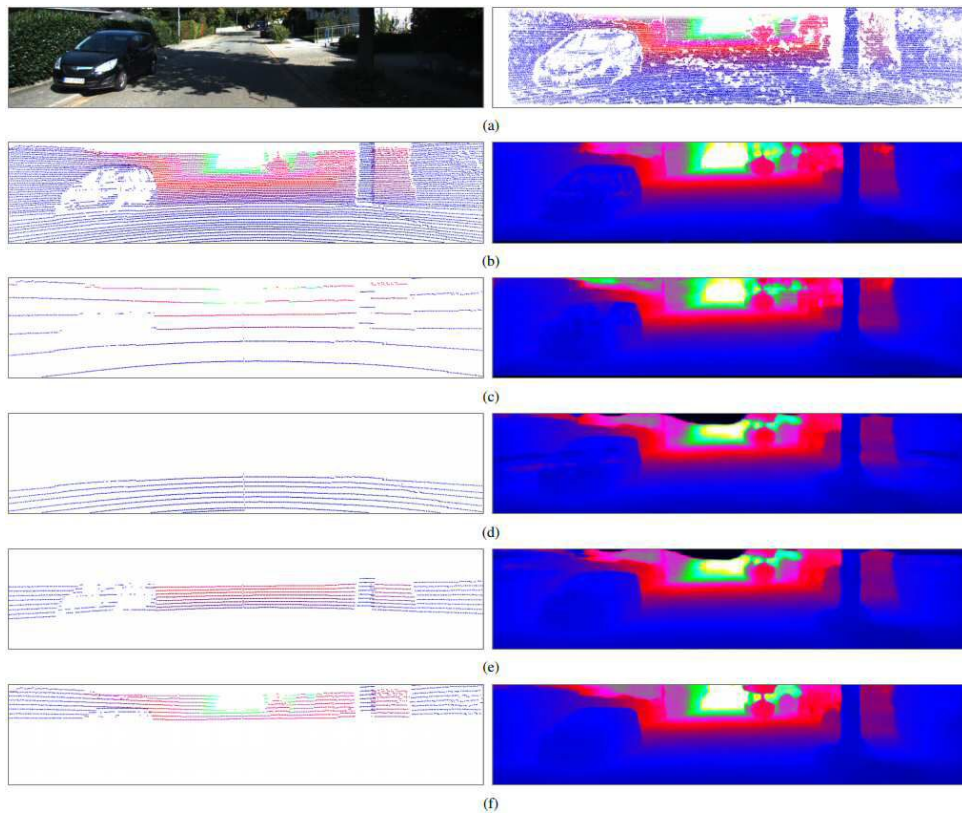
도면6



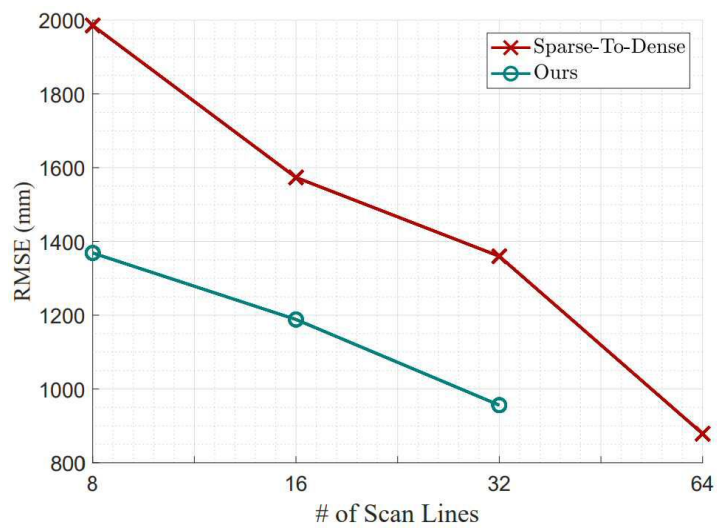
도면7



도면8



도면9





도면10

