



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2021년10월07일
(11) 등록번호 10-2310789
(24) 등록일자 2021년10월01일

(51) 국제특허분류(Int. Cl.)
G06T 7/70 (2017.01) G06T 3/00 (2019.01)
G06T 7/11 (2017.01) G06T 7/20 (2017.01)
G06T 7/50 (2017.01)
(52) CPC특허분류
G06T 7/70 (2017.01)
G06T 3/0093 (2013.01)
(21) 출원번호 10-2019-0141508
(22) 출원일자 2019년11월07일
심사청구일자 2019년11월07일
(65) 공개번호 10-2021-0055260
(43) 공개일자 2021년05월17일
(56) 선행기술조사문헌
Domain Transfer for 3D Pose Estimation from
Color Images Without Manual Annotations, ACCV
2018: Computer Vision(2019.05.26)*
(뒷면에 계속)

(73) 특허권자
연세대학교 산학협력단
서울특별시 서대문구 연세로 50 (신촌동, 연세대
학교)
(72) 발명자
이상윤
서울특별시 서초구 서초중앙로24길 43, 101동
1103호(서초동, 유원서초아파트)
황상원
서울특별시 강서구 공항대로45길 68-6, 106동 60
5호(등촌동, 서광등촌마을아파트)
(뒷면에 계속)
(74) 대리인
민영준

전체 청구항 수 : 총 18 항

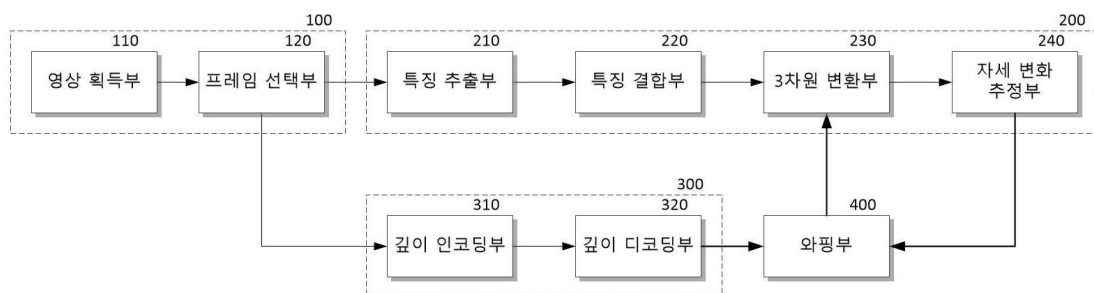
심사관 : 김기호

(54) 발명의 명칭 다양한 시점에서 연속적으로 획득된 영상을 활용한 카메라 자세 및 깊이 추정 장치 및 방법

(57) 요약

본 발명은 다양한 시점에서 연속적으로 획득된 다수 프레임의 영상에서 키 프레임과 참조 프레임을 선택하는 프레임 제공부, 키 프레임과 참조 프레임 중 적어도 하나를 인가받아 미리 학습된 패턴 추정 방식 및 패턴 복원 방식에 따라 인가된 프레임에 대응하는 깊이 맵을 획득하는 깊이 판별부 및 키 프레임과 참조 프레임에서 미리 학습된 패턴 추정 방식에 따라 특징을 추출하여 획득되는 키 특징맵과 참조 특징맵을 결합하여 결합 특징맵을 획득하고, 결합 특징맵에 3차원 좌표가 포함되도록 획득된 깊이 맵에 기반하여 생성된 3차원 좌표계 특징맵을 결합하여 3차원 특징맵을 획득하며, 3차원 특징맵으로부터 미리 학습된 패턴 추정 방식에 따라 키 프레임으로부터 참조 프레임으로의 카메라 자세 변화를 추정하는 자세 판별부를 포함하는 카메라 자세 및 깊이 추정 장치 및 방법을 제공할 수 있다.

대표도



(52) CPC특허분류

G06T 7/11 (2017.01)
G06T 7/20 (2013.01)
G06T 7/50 (2017.01)
G06T 2207/20081 (2013.01)
G06T 2207/30244 (2013.01)

(72) 발명자

김우진

서울특별시 마포구 백범로1길 83, 505호(노고산동, 신촌다올노블리움)

이준협

서울특별시 서대문구 연희로10길 29-5, 205호(연희동)

(56) 선행기술조사문헌

FastDepth: Fast Monocular Depth Estimation on Embedded Systems, Computer Science_Computer Vision and Pattern Recognition(2019.03.08.)*
 Siamese Convolutional Neural Network for Sub-millimeter-accurate Camera Pose Estimation and Visual Servoing, Computer Science_Robotics (2019.03.12.)*

KR1020100104088 A

KR1020180019329 A

KR1020190062794 A

*는 심사관에 의하여 인용된 문헌

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

다양한 시점에서 연속적으로 획득된 다수 프레임의 영상에서 키 프레임과 참조 프레임을 선택하는 프레임 제공부;

상기 키 프레임과 상기 참조 프레임 중 적어도 하나를 인가받아 미리 학습된 패턴 추정 방식 및 패턴 복원 방식에 따라 인가된 프레임에 대응하는 깊이 맵을 획득하는 깊이 판별부; 및

상기 키 프레임과 상기 참조 프레임에서 미리 학습된 패턴 추정 방식에 따라 특징을 추출하여 획득되는 키 특징맵과 참조 특징맵을 결합하여 결합 특징맵을 획득하고, 상기 결합 특징맵에 3차원 좌표가 포함되도록 획득된 깊이 맵에 기반하여 생성된 3차원 좌표계 특징맵을 결합하여 3차원 특징맵을 획득하며, 상기 3차원 특징맵으로부터 미리 학습된 패턴 추정 방식에 따라 상기 키 프레임으로부터 상기 참조 프레임으로의 카메라 자세 변화를 추정하는 자세 판별부를 포함하되,

상기 프레임 제공부는

다양한 시점에서 연속적으로 획득된 다수 프레임의 영상을 획득하는 영상 획득부; 및

상기 다수 프레임의 영상에서 이전 선택된 키 프레임과 카메라 자세 또는 깊이 차이 중 적어도 하나가 기지정된 기준값 이상으로 판별된 참조 프레임을 기준으로 이전 프레임을 상기 키 프레임으로 선택하고, 이후 프레임을 상기 참조 프레임으로 선택하는 프레임 선택부를 포함하는 카메라 자세 및 깊이 추정 장치.

청구항 2

제1 항에 있어서, 상기 자세 판별부는

상기 키 프레임과 상기 참조 프레임을 인가받고, 미리 학습된 패턴 추정 방식에 따라 상기 키 프레임과 상기 참조 프레임 각각의 특징을 추출하여, 상기 키 특징맵과 상기 참조 특징맵을 획득하는 특징 추출부;

상기 키 특징맵과 상기 참조 특징맵을 인가받아 결합하여 상기 결합 특징맵을 획득하는 특징 결합부;

지정된 패턴을 갖는 2차원 좌표계 특징맵에 상기 깊이 판별부에서 획득된 깊이 맵을 카메라 내부 파라미터에 따라 변형하여 가중함으로써 획득되는 상기 3차원 좌표계 특징맵을 상기 결합 특징맵에 결합하여 3차원 특징맵을 획득하는 3차원 변환부; 및

상기 3차원 특징맵을 인가받아 상기 키 프레임과 상기 참조 프레임 사이의 카메라의 회전 변화와 이동 변화를 각각 구분하여 추정하고, 추정된 카메라의 회전 변화와 이동 변화로부터 카메라 자세 변화를 판별하는 자세 변화 추정부를 포함하는 카메라 자세 및 깊이 추정 장치.

청구항 3

제2 항에 있어서, 상기 특징 추출부는

상기 키 프레임을 인가받아 상기 키 특징맵을 획득하는 키 프레임 특징 추출부; 및

상기 키 프레임 특징 추출부와 동일한 신경망으로 구성되고 동일하게 학습되어 동일한 가중치를 가지며 상기 참조 프레임을 인가받아 상기 참조 특징맵을 획득하는 참조 프레임 특징 추출부를 포함하는 카메라 자세 및 깊이 추정 장치.

청구항 4

제2 항에 있어서, 상기 3차원 변환부는

상기 결합 특징맵과 대응하는 크기를 갖고 x축 방향의 좌표값을 설정하기 위해 미리 지정된 패턴값을 갖는 제1

채널과 y축 방향의 좌표값을 설정하기 위해 미리 지정된 패턴값을 갖는 제2 채널 및 미리 지정된 값을 갖는 동작 좌표계인 제3 채널이 결합된 행렬에 상기 카메라 내부 파라미터로서 미리 획득되는 캘리브레이션 행렬의 역행렬과 상기 깊이 맵을 곱하여 상기 3차원 좌표계 특징맵을 획득하는 카메라 자세 및 깊이 추정 장치.

청구항 5

제2 항에 있어서, 상기 자세 변화 추정부는

상기 3차원 특징맵을 인가받고, 미리 학습된 패턴 추정 방식에 따라 상기 키 프레임으로부터 상기 참조 프레임으로의 카메라의 회전 변화량을 나타내는 추정하는 회전 변환 추출부;

상기 3차원 특징맵을 인가받고, 미리 학습된 패턴 추정 방식에 따라 상기 키 프레임으로부터 상기 참조 프레임으로의 카메라의 평행 이동량을 추정하는 이동 벡터 추출부; 및

상기 회전 변화량과 상기 평행 이동량을 기반으로 상기 키 프레임으로부터 상기 참조 프레임으로의 카메라 자세 변화를 판별하는 카메라 자세 획득부를 포함하는 카메라 자세 및 깊이 추정 장치.

청구항 6

제5 항에 있어서, 상기 회전 변환 추출부는

상기 3차원 특징맵으로부터 상기 카메라의 회전 변화량을 x, y, z 축 방향 각각에서의 회전을 나타내도록 지정된 크기를 갖는 회전 행렬(Rotation Matrix)로 획득하는 카메라 자세 및 깊이 추정 장치.

청구항 7

제6 항에 있어서, 상기 회전 변환 추출부는

상기 3차원 특징맵으로부터 상기 카메라의 회전 변화량을 3개의 오일러 각(Euler Angle)(ψ , ϕ , θ)으로 추정하고, 추정된 3개의 오일러 각을 기지정된 방식에 따라 상기 회전 행렬로 변환하는 카메라 자세 및 깊이 추정 장치.

청구항 8

삭제

청구항 9

제1 항에 있어서, 상기 카메라 자세 및 깊이 추정 장치는

상기 깊이 판별부에서 획득된 깊이 맵과 상기 자세 판별부에서 획득된 카메라 자세 변화를 이용하여 상기 키 프레임 또는 상기 참조 프레임 중 하나를 다른 프레임에 대응하도록 와핑시키고, 와핑된 프레임을 다른 프레임과 비교하여 깊이 맵의 깊이 정보를 보정하는 와핑부를 더 포함하는 카메라 자세 및 깊이 추정 장치.

청구항 10

제9 항에 있어서, 상기 카메라 자세 및 깊이 추정 장치는

학습 시에 상기 와핑부에서 와핑된 프레임과 다른 프레임 사이의 차이를 오차로 획득하고, 획득된 오차를 상기 깊이 판별부 및 상기 자세 판별부로 역전파하여 학습시키는 학습부를 더 포함하는 카메라 자세 및 깊이 추정 장치.

청구항 11

다양한 시점에서 연속적으로 획득된 다수 프레임의 영상에서 키 프레임과 참조 프레임을 선택하는 단계;

상기 키 프레임과 상기 참조 프레임 중 적어도 하나를 인가받아 미리 학습된 패턴 추정 방식 및 패턴 복원 방식에 따라 인가된 프레임에 대응하는 깊이 맵을 획득하는 단계; 및

상기 키 프레임과 상기 참조 프레임에서 미리 학습된 패턴 추정 방식에 따라 특징을 추출하여 획득되는 키 특징맵과 참조 특징맵을 결합하여 결합 특징맵을 획득하고, 상기 결합 특징맵에 3차원 좌표가 포함되도록 획득된 깊이 맵에 기반하여 생성된 3차원 좌표계 특징맵을 결합하여 3차원 특징맵을 획득하며, 상기 3차원 특징맵으로부터

터 미리 학습된 패턴 추정 방식에 따라 상기 키 프레임으로부터 상기 참조 프레임으로의 카메라 자세 변화를 추정하는 단계를 포함하되,

상기 프레임을 선택하는 단계는

다양한 시점에서 연속적으로 획득된 다수 프레임의 영상을 획득하는 단계; 및

상기 다수 프레임의 영상에서 이전 선택된 키 프레임과 카메라 자세 또는 깊이 차이 중 적어도 하나가 기지정된 기준값 이상으로 판별된 참조 프레임을 기준으로 이전 프레임을 상기 키 프레임으로 선택하고, 이후 프레임을 상기 참조 프레임으로 선택하는 단계를 포함하는 카메라 자세 및 깊이 추정 방법.

청구항 12

제11 항에 있어서, 상기 자세 변화를 추정하는 단계는

상기 키 프레임과 상기 참조 프레임을 인가받고, 미리 학습된 패턴 추정 방식에 따라 상기 키 프레임과 상기 참조 프레임 각각의 특징을 추출하여, 상기 키 특징맵과 상기 참조 특징맵을 획득하는 단계;

상기 키 특징맵과 상기 참조 특징맵을 인가받아 결합하여 상기 결합 특징맵을 획득하는 단계;

지정된 패턴을 갖는 2차원 좌표계 특징맵에 상기 깊이 판별부에서 획득된 깊이 맵을 카메라 내부 파라미터에 따라 변형하여 가중함으로써 획득되는 상기 3차원 좌표계 특징맵을 상기 결합 특징맵에 결합하여 3차원 특징맵을 획득하는 단계; 및

상기 3차원 특징맵을 인가받아 상기 키 프레임과 상기 참조 프레임 사이의 카메라의 회전 변화와 이동 변화를 각각 구분하여 추정하고, 추정된 카메라의 회전 변화와 이동 변화로부터 카메라 자세 변화를 판별하는 단계를 포함하는 카메라 자세 및 깊이 추정 방법.

청구항 13

제12 항에 있어서, 상기 참조 특징맵을 획득하는 단계는

동일 구조를 갖고 동일하게 학습되어 동일한 가중치를 갖는 삼 신경망을 이용하여 상기 키 프레임에 대응하는 상기 키 특징맵과 상기 참조 프레임에 대응하는 상기 참조 특징맵을 획득하는 카메라 자세 및 깊이 추정 방법.

청구항 14

제12 항에 있어서, 상기 3차원 특징맵을 획득하는 단계는

상기 결합 특징맵과 대응하는 크기를 갖고 x축 방향의 좌표값을 설정하기 위해 미리 지정된 패턴값을 갖는 제1 채널과 y축 방향의 좌표값을 설정하기 위해 미리 지정된 패턴값을 갖는 제2 채널 및 미리 지정된 값을 갖는 동차 좌표계인 제3 채널이 결합된 행렬에 상기 카메라 내부 파라미터로서 미리 획득되는 캘리브레이션 행렬의 역행렬과 상기 깊이 맵을 곱하여 상기 3차원 좌표계 특징맵을 획득하는 카메라 자세 및 깊이 추정 방법.

청구항 15

제12 항에 있어서, 상기 자세 변화를 추정하는 단계는

상기 3차원 특징맵을 인가받고, 미리 학습된 패턴 추정 방식에 따라 상기 키 프레임으로부터 상기 참조 프레임으로의 카메라의 회전 변화량을 나타내는 추정하는 단계;

상기 3차원 특징맵을 인가받고, 미리 학습된 패턴 추정 방식에 따라 상기 키 프레임으로부터 상기 참조 프레임으로의 카메라의 평행 이동량을 추정하는 단계; 및

상기 회전 변화량과 상기 평행 이동량을 기반으로 상기 키 프레임으로부터 상기 참조 프레임으로의 카메라 자세 변화를 판별하는 단계를 포함하는 카메라 자세 및 깊이 추정 방법.

청구항 16

제15 항에 있어서, 상기 회전 변화량을 나타내는 추정하는 단계는

상기 3차원 특징맵으로부터 상기 카메라의 회전 변화량을 x, y, z 축 방향 각각에서의 회전을 나타내도록 기지정된 크기를 갖는 회전 행렬(Rotation Matrix)로 획득하는 카메라 자세 및 깊이 추정 방법.

청구항 17

제16 항에 있어서, 상기 회전 변화량을 나타내는 추정하는 단계는

상기 3차원 특징맵으로부터 상기 카메라의 회전 변화량을 3개의 오일러 각(Euler Angle)(ψ , ϕ , θ)으로 추정하는 단계; 및

추정된 3개의 오일러 각을 기지정된 방식에 따라 상기 회전 행렬로 변환하는 단계를 포함하는 카메라 자세 및 깊이 추정 방법.

청구항 18

삭제

청구항 19

제11 항에 있어서, 상기 카메라 자세 및 깊이 추정 방법은

상기 깊이 맵과 상기 카메라 자세 변화를 이용하여 상기 키 프레임 또는 상기 참조 프레임 중 하나를 다른 프레임에 대응하도록 와핑시키는 단계;

와핑된 프레임을 다른 프레임과 비교하여 깊이 맵의 깊이 정보를 보정하는 단계를 더 포함하는 카메라 자세 및 깊이 추정 방법.

청구항 20

제19 항에 있어서, 상기 카메라 자세 및 깊이 추정 방법은

와핑된 프레임과 다른 프레임 사이의 차이를 오차로 획득하고, 획득된 오차를 역전파하여 학습시키는 단계를 더 포함하는 카메라 자세 및 깊이 추정 방법.

발명의 설명

기술 분야

[0001] 본 발명은 카메라 자세 및 깊이 추정 장치 및 방법에 관한 것으로, 다양한 시점에서 연속적으로 획득된 영상을 활용한 카메라 자세 및 깊이 추정 장치 및 방법에 관한 것이다.

배경 기술

[0002] 깊이 영상을 얻기 위해서는 주로 고가의 라이다 센서를 이용하거나, 동시에 취득된 스테레오 영상(좌영상과 우영상)에서 대응쌍을 탐색하고, 탐색된 대응쌍에 기반하여 디스패리티 맵(disparity map)을 획득한 후, 알고 있는 스테레오 카메라 사이의 거리(baseline)를 통해 실제 깊이 값을 얻는다. 라이다의 높은 비용으로 인해 스테레오 영상으로부터 깊이 영상을 획득하는 방식이 일반적으로 이용되고 있다.

[0003] 스테레오 영상으로부터 거리 정보, 즉 깊이 영상을 획득하는 방법은 두 카메라의 파라미터를 추정하는 카메라 칼리브레이션 단계와 두 카메라에서 획득된 영상 상에서 서로 대응되는 위치를 찾는 스테레오 매칭 단계를 포함하여 수행되며, 이때 깊이(z)는 수식식 $z = (b \times f) / d$ 에 따라 획득된다.

[0004] 여기서 b는 스테레오 카메라 사이의 거리(baseline), f는 카메라의 내부 파라미터인 초점거리(focal Length), d는 디스패리티(disparity)이다. 카메라 사이의 거리(b)와 초점거리(f)는 고정된 상수 값을 갖기 때문에 깊이값(z)은 디스패리티(d)에 따라 계산되어 깊이 영상이 획득될 수 있다. 그러나 디스패리티 (d)만을 이용하여 깊이값(z)이 계산되므로 다양한 깊이 영상의 범위(Depth of field)를 얻지 못한다는 한계점이 존재한다.

[0005] 뿐만 아니라 좌영상과 우영상만을 가지고 깊이값을 예측하기 때문에 후처리가 요구되는 문제가 있다.

[0006] 이러한 문제를 해결하기 위해 동시에 다수의 시점에서 획득된 영상을 활용하는 기법이 제안된 바 있으나, 이 기법을 적용하기 위해서는 각 영상을 획득하는 카메라의 위치 및 자세를 정확하게 알아야 한다는 문제가 있다.

선행기술문헌

특허문헌

[0007] (특허문헌 0001) 한국 공개 특허 제10-2019-0032532호 (2019.03.27 공개)

발명의 내용

해결하려는 과제

[0008] 본 발명의 목적은 다양한 시점에서 연속적으로 획득된 영상으로부터 카메라의 자세를 정확하게 추정할 수 있는 카메라 자세 및 깊이 추정 장치 및 방법을 제공하는데 있다.

[0009] 본 발명의 다른 목적은 추정된 카메라 자세를 기반으로 다양한 시점에서 연속적으로 획득된 영상을 와핑하여 비교함으로써, 정확한 깊이 영상을 획득할 수 있는 카메라 자세 및 깊이 추정 장치 및 방법을 제공하는데 있다.

[0010] 본 발명의 또 다른 목적은 다양한 시점에서 연속적으로 획득된 영상에서 키프레임을 선택하고, 선택된 키프레임과 참조 영상 사이의 차이를 비교하여 오차가 누적되지 않는 카메라 자세 및 깊이 정보를 획득할 수 있는 카메라 자세 및 깊이 추정 장치 및 방법을 제공하는데 있다.

과제의 해결 수단

[0011] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른 카메라 자세 및 깊이 추정 장치는 다양한 시점에서 연속적으로 획득된 다수 프레임의 영상에서 키프레임과 참조 프레임을 선택하는 프레임 제공부; 상기 키프레임과 상기 참조 프레임 중 적어도 하나를 인가받아 미리 학습된 패턴 추정 방식 및 패턴 복원 방식에 따라 인가된 프레임에 대응하는 깊이 맵을 획득하는 깊이 판별부; 및 상기 키프레임과 상기 참조 프레임에서 미리 학습된 패턴 추정 방식에 따라 특징을 추출하여 획득되는 키프레임 특징맵과 참조 특징맵을 결합하여 결합 특징맵을 획득하고, 상기 결합 특징맵에 3차원 좌표가 포함되도록 획득된 깊이 맵에 기반하여 생성된 3차원 좌표계 특징맵을 결합하여 3차원 특징맵을 획득하며, 상기 3차원 특징맵으로부터 미리 학습된 패턴 추정 방식에 따라 상기 키프레임으로부터 상기 참조 프레임으로의 카메라 자세 변화를 추정하는 자세 판별부를 포함한다.

[0012] 상기 자세 판별부는 상기 키프레임과 상기 참조 프레임을 인가받고, 미리 학습된 패턴 추정 방식에 따라 상기 키프레임과 상기 참조 프레임 각각의 특징을 추출하여, 상기 키프레임 특징맵과 상기 참조 특징맵을 획득하는 특징 추출부; 상기 키프레임 특징맵과 상기 참조 특징맵을 인가받아 결합하여 상기 결합 특징맵을 획득하는 특징 결합부; 지정된 패턴을 갖는 2차원 좌표계 특징맵에 상기 깊이 판별부에서 획득된 깊이 맵을 카메라 내부 파라미터에 따라 변형하여 가중함으로써 획득되는 상기 3차원 좌표계 특징맵을 상기 결합 특징맵에 결합하여 3차원 특징맵을 획득하는 3차원 변환부; 및 상기 3차원 특징맵을 인가받아 상기 키프레임과 상기 참조 프레임 사이의 카메라의 회전 변화와 이동 변화를 각각 구분하여 추정하고, 추정된 카메라의 회전 변화와 이동 변화로부터 카메라 자세 변화를 판별하는 자세 변화 추정부를 포함할 수 있다.

[0013] 상기 특징 추출부는 상기 키프레임을 인가받아 상기 키프레임 특징맵을 획득하는 키프레임 특징 추출부; 및 상기 키프레임 특징 추출부와 동일한 신경망으로 구성되고 동일하게 학습되어 동일한 가중치를 가지며 상기 참조 프레임을 인가받아 상기 참조 특징맵을 획득하는 참조 프레임 특징 추출부를 포함할 수 있다.

[0014] 상기 3차원 변환부는 상기 결합 특징맵과 대응하는 크기를 갖고 x축 방향의 좌표값을 설정하기 위해 미리 지정된 패턴값을 갖는 제1 채널과 y축 방향의 좌표값을 설정하기 위해 미리 지정된 패턴값을 갖는 제2 채널 및 미리 지정된 값을 갖는 동차 좌표계인 제3 채널이 결합된 행렬에 상기 카메라 내부 파라미터로서 미리 획득되는 캘리브레이션 행렬의 역행렬과 상기 깊이 맵을 곱하여 상기 3차원 좌표계 특징맵을 획득할 수 있다.

[0015] 상기 자세 변화 추정부는 상기 3차원 특징맵을 인가받고, 미리 학습된 패턴 추정 방식에 따라 상기 키프레임으로부터 상기 참조 프레임으로의 카메라의 회전 변화량을 나타내는 추정하는 회전 변환 추출부; 상기 3차원 특징맵을 인가받고, 미리 학습된 패턴 추정 방식에 따라 상기 키프레임으로부터 상기 참조 프레임으로의 카메라의 평행 이동량을 추정하는 이동 벡터 추출부; 및 상기 회전 변화량과 상기 평행 이동량을 기반으로 상기 키프레임으로부터 상기 참조 프레임으로의 카메라 자세 변화를 판별하는 카메라 자세 획득부를 포함할 수 있다.

[0016] 상기 회전 변환 추출부는 상기 3차원 특징맵으로부터 상기 카메라의 회전 변화량을 x, y, z 축 방향 각각에서의

회전을 나타내도록 기지정된 크기를 갖는 회전 행렬(Rotation Matrix)로 획득할 수 있다.

- [0017] 상기 회전 변환 추출부는 상기 3차원 특징맵으로부터 상기 카메라의 회전 변화량을 3개의 오일러 각(Euler Angle)(ψ , ϕ , θ)으로 추정하고, 추정된 3개의 오일러 각을 기지정된 방식에 따라 상기 회전 행렬로 변환할 수 있다.
- [0018] 상기 프레임 제공부는 다양한 시점에서 연속적으로 획득된 다수 프레임의 영상을 획득하는 영상 획득부; 및 상기 다수 프레임의 영상에서 이전 선택된 키 프레임과 카메라 자세 또는 깊이 차이 중 적어도 하나가 기지정된 기준값 이상으로 판별된 참조 프레임을 기준으로 이전 프레임을 상기 키 프레임으로 선택하고, 이후 프레임을 상기 참조 프레임으로 선택하는 프레임 선택부를 포함할 수 있다.
- [0019] 상기 카메라 자세 및 깊이 추정 장치는 상기 깊이를 판별부에서 획득된 깊이 맵과 상기 자세 판별부에서 획득된 카메라 자세 변화를 이용하여 상기 키 프레임 또는 상기 참조 프레임 중 하나를 다른 프레임에 대응하도록 와핑시키고, 와핑된 프레임을 다른 프레임과 비교하여 깊이 맵의 깊이 정보를 보정하는 와핑부를 더 포함할 수 있다.
- [0020] 상기 카메라 자세 및 깊이 추정 장치는 학습 시에 상기 와핑부에서 와핑된 프레임과 다른 프레임 사이의 차이를 오차로 획득하고, 획득된 오차를 상기 깊이를 판별부 및 상기 자세 판별부로 역전파하여 학습시키는 학습부를 더 포함할 수 있다.
- [0021] 상기 목적을 달성하기 위한 본 발명의 다른 실시예에 따른 카메라 자세 및 깊이 추정 방법은 다양한 시점에서 연속적으로 획득된 다수 프레임의 영상에서 키 프레임과 참조 프레임을 선택하는 단계; 상기 키 프레임과 상기 참조 프레임 중 적어도 하나를 인가받아 미리 학습된 패턴 추정 방식 및 패턴 복원 방식에 따라 인가된 프레임에 대응하는 깊이 맵을 획득하는 단계; 및 상기 키 프레임과 상기 참조 프레임에서 미리 학습된 패턴 추정 방식에 따라 특징을 추출하여 획득되는 키 특징맵과 참조 특징맵을 결합하여 결합 특징맵을 획득하고, 상기 결합 특징맵에 3차원 좌표가 포함되도록 획득된 깊이 맵에 기반하여 생성된 3차원 좌표계 특징맵을 결합하여 3차원 특징맵을 획득하며, 상기 3차원 특징맵으로부터 미리 학습된 패턴 추정 방식에 따라 상기 키 프레임으로부터 상기 참조 프레임으로의 카메라 자세 변화를 추정하는 단계를 포함한다.

발명의 효과

- [0022] 따라서, 본 발명의 실시예에 따른 카메라 자세 및 깊이 추정 장치 및 방법은 다양한 시점에서 연속적으로 획득된 다수 프레임의 영상에서 키 프레임과 참조 프레임을 선택하고, 선택된 키 프레임과 참조 프레임 사이의 카메라 자세 변화와 키 프레임과 참조 프레임 중 적어도 하나의 깊이 맵을 획득하고, 획득된 카메라 자세 변화와 깊이 맵을 기반으로 키 프레임과 참조 프레임 중 하나를 다른 하나에 대응하도록 와핑시켜 비교하여 학습시킴으로써, 카메라 자세와 깊이를 모두 정확하게 추정할 수 있도록 학습시킬 수 있다. 특히 키 프레임(key)과 참조 프레임(ref)에 대한 특징맵에 3차원 좌표계 특징맵을 결합하여 카메라 자세 변화를 추정하도록 함으로써, 정확하게 카메라 자세 변화를 추정할 수 있다. 또한 카메라 자세 변화 중 회전 변화를 추정하기 위해 오일러 각을 추정하도록 학습시킴으로써, 더욱 용이하고 정확하게 카메라 자세 변화를 추정할 수 있도록 한다. 뿐만 아니라 키 프레임을 이전 선택된 키 프레임과 카메라 자세 또는 깊이 차이 중 적어도 하나가 기지정된 기준값 이상으로 판별된 참조 프레임(ref)을 기준으로 직전 또는 직후 프레임 중 하나를 키 프레임(key)으로 선택함으로써, 오차가 누적되는 것을 최소화할 수 있다.

도면의 간단한 설명

- [0023] 도 1은 본 발명의 일 실시예에 따른 카메라 자세 및 깊이 추정 장치의 개략적 구조를 나타낸다.
- 도 2는 도 1의 프레임 선택부의 키 프레임 선택 여부에 따른 카메라 자세 판별 방식의 차이를 나타낸다.
- 도 3은 연속되는 프레임에 의한 오차의 누적을 설명하기 위한 도면이다.
- 도 4는 도 1의 카메라 자세 판별부의 상세 구조를 나타낸다.
- 도 5 및 도 6은 도 4의 3차원 변환 특징 추출부가 2차원 특징맵에 3차원 좌표 특징맵을 추가하는 개념을 설명하기 위한 도면이다.
- 도 7은 키 프레임과 참조 프레임, 키 프레임에 대한 깊이 맵 및 참조 프레임의 일예를 나타낸다.

도 8은 본 발명의 일 실시예에 따른 카메라 자세 및 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 1은 본 발명의 일 실시예에 따른 카메라 자세 및 깊이 추정 장치의 개략적 구조를 나타내고, 도 2는 도 1의 프레임 선택부의 키 프레임 선택 여부에 따른 카메라 자세 판별 방식의 차이를 나타내며, 도 3은 연속되는 프레임에 의한 오차의 누적을 설명하기 위한 도면이다. 그리고 도 4는 도 1의 카메라 자세 판별부의 상세 구조를 나타내고, 도 5 및 도 6은 도 4의 3차원 변환 특징 추출부가 2차원 특징맵에 3차원 좌표 특징맵을 추가하는 개념을 설명하기 위한 도면이며, 도 7은 키 프레임과 참조 프레임, 키 프레임에 대한 깊이 맵 및 참조 프레임의 일예를 나타낸다.
- [0028] 도 1을 참조하면, 본 실시예에 따른 카메라 자세 및 깊이 추정 장치는 다양한 시점의 연속 영상에서 키 프레임과 참조 프레임을 선택하여 제공하는 프레임 제공부(100), 프레임 제공부(100)에서 제공된 키 프레임과 참조 프레임 사이의 카메라의 자세 변화를 판별하는 카메라 자세 판별부(200), 깊이 정보를 획득하는 깊이 판별부(300) 및 카메라 자세 판별부(200)와 깊이 판별부(300)에서 판별된 카메라 자세와 깊이 정보를 기반으로 참조 프레임을 키 프레임에 대응하는 영상으로 와핑하는 와핑부(400)를 포함할 수 있다.
- [0029] 우선 프레임 제공부(100)는 영상 획득부(110)와 프레임 선택부(120)를 포함할 수 있다.
- [0030] 영상 획득부(110)는 카메라 장치가 이동하면서 연속적으로 취득한 서로 다른 시점에서 획득한다. 영상 획득부(110)는 카메라 장치로 구현될 수도 있으며, 카메라에서 획득된 영상을 저장하는 저장 장치 또는 유무선 통신망을 통해 영상을 전달받는 통신 장치 등으로 구현될 수 있다.
- [0031] 본 실시예에서 영상 획득부(110)가 획득하는 영상을 스테레오 카메라를 이용한 스테레오 영상이 아닌 단일 카메라에서 연속적으로 취득된 영상이며, 특히 단일 카메라의 이동에 따라 시점이 변화되어 획득된 영상일 수 있다.
- [0032] 프레임 선택부(120)는 영상 획득부(110)가 획득한 다양한 시점에서 연속적으로 획득된 다수 프레임의 영상을 키 프레임(key)과 참조 프레임(ref)으로 구분한다.
- [0033] 프레임 선택부(120)는 다수 프레임 영상에서 키 프레임(key)을 선택하고, 키 프레임(key)을 제외한 나머지 프레임 중 하나의 프레임을 참조 프레임(ref)으로 선택한다.
- [0034] 프레임 선택부(120)는 우선 다수 프레임의 영상에서 키 프레임(key)을 선택한다. 프레임 선택부(120)는 다수 프레임의 영상 중 최초 프레임 또는 기지정된 프레임을 초기 키 프레임(key)으로 선택할 수 있으며, 이후에는 이전 선택된 키 프레임과 카메라 자세 또는 깊이 차이 중 적어도 하나가 기지정된 기준값 이상으로 판별된 참조 프레임(ref)을 기준으로 직전 또는 직후 프레임 중 하나를 키 프레임(key)으로 선택할 수 있다. 일례로 프레임 선택부(120)는 와핑부(400)에서 참조 프레임(ref)을 키 프레임(key)에 대응하도록 변환하는 와핑 레벨에 기반하여 키 프레임을 선택할 수 있다.
- [0035] 최근에는 RGB 영상 또는 옵티컬 플로우(Optical Flow)와 같은 영상들에서 각 영상들 사이의 관계를 인공 신경망을 이용하여 예측하는 기법이 제안되었다. 그러나 기존에는 도 2의 (a)에 도시된 바와 같이, 연속되는 다수 프레임의 영상에서 서로 인접한 프레임 사이의 차이만을 비교하여 카메라 자세를 추정하였다. 이 경우, 도 3의 (a)와 같이 모든 프레임이 정상 촬영되면, 카메라 자세 또한 큰 오차가 발생되지 않고 정상적으로 추정될 수 있다. 그러나 다수 프레임 중 특정 프레임에 정상 촬영된 도 3의 (b)에 도시된 바와 같이 빛 번짐 등의 큰 변화가 발생되면, 해당 프레임에 대한 카메라 자세 추정에 오차가 발생된다. 그리고 발생된 오차는 이후 프레임에서의

카메라 자세 추정 시에도 누적하여 반영되므로 도 3의 (c)에 도시된 바와 같이 계속 증가되는 문제가 발생한다. 도 3의 (c)는 도 2의 (a)와 같이 연속되는 프레임 사이의 차이를 기반으로 추정된 카메라 자세(est)를 실제 카메라 자세(gt)와 비교하여 나타내었으며, 도 3의 (c)에 도시된 바와 같이, 연속되는 프레임 사이의 차이를 기반으로 카메라 자세를 추정하는 경우, 오차가 누적되어 점차로 증가됨을 알 수 있다.

- [0036] 이에 본 실시예에서는 누적되는 오차를 저감시킬 수 있도록, 도 2의 (b)에 도시된 바와 같이, 프레임 선택부(120)가 다수 프레임 영상 중 키 프레임(key)을 선택하고, 선택된 키 프레임(key)을 기준으로 이후 획득된 다수 프레임에서의 카메라 변화를 추정한다. 따라서 선택된 키 프레임들에 의해 추정된 카메라 자세에 오차는 누적될 수 있으나, 연속하여 선택된 키 프레임들 사이에서 획득된 적어도 하나의 참조 프레임에서 발생하는 오차는 다음 키 프레임(key)이 선택되면 전달되지 않으므로 누적되지 않는다. 즉 누적되는 오차를 저감시킬 수 있다.
- [0037] 여기서 프레임 선택부(120)가 다수 프레임 영상에서 단일 키 프레임(key)을 이용하지 않고, 반복적으로 키 프레임(key)을 선택하는 것은 참조 프레임(ref)과 키 프레임(key) 사이의 차이가 커지게 되면, 빛 번짐과 같은 비정상적 조건에 의한 오차가 발생되지 않더라도, 카메라의 이동에 의해 키 프레임(key)과 참조 프레임(ref) 사이의 변화가 커지게 되어 카메라 자세 및 깊이 추정이 어렵게 되기 때문이다. 또한 프레임 선택부(120)가 이전 선택된 키 프레임(key)과 카메라 자세 또는 깊이 차이가 기설정된 기준값 이상으로 판별된 참조 프레임(ref)의 직전 또는 직후 프레임을 키 프레임(key)으로 선택하는 것은 큰 오차가 발생된 프레임의 이전 프레임 또는 이후 프레임을 기반으로 하여 안정적으로 이후 프레임에서의 카메라 자세를 추정할 수 있도록 하기 위해서이다. 즉 선택되는 키 프레임들 사이에서 누적될 수 있는 오차를 최소화하기 위해서이다.
- [0038] 프레임 선택부(120)는 키 프레임(key)이 선택되면 선택된 키 프레임(key)과 선택된 키 프레임(key) 이후의 프레임 중 하나의 프레임을 참조 프레임(ref)으로 선택하여 카메라 자세 판별부(200)로 전달한다.
- [0039] 카메라 자세 판별부(200)는 프레임 선택부(120)로부터 키 프레임(key)과 참조 프레임(ref)을 인가받고, 키 프레임(key)을 기반으로 참조 프레임(ref)에서 변화된 카메라의 자세를 추정한다. 즉 키 프레임(key)을 촬영한 시점의 카메라의 자세를 기준으로 참조 프레임(ref)을 촬영한 시점의 카메라의 자세 변화를 추정한다.
- [0040] 카메라 자세 판별부(200)는 특징 추출부(210), 특징 결합부(220), 3차원 변환부(230) 및 자세 변화 추정부(240)를 포함한다.
- [0041] 특징 추출부(210)는 프레임 선택부(120)에서 선택된 키 프레임(key)과 참조 프레임을 인가받고, 미리 학습된 패턴 추정 방식에 따라 인가된 키 프레임(key)과 참조 프레임 각각의 특징을 추출하여 키 특징맵과 참조 특징맵을 획득한다.
- [0042] 도 4를 참조하면, 특징 추출부(210)는 키 프레임(key)을 인가받아 키 프레임(key)의 특징을 추출하는 키 프레임 특징 추출부(211)와 참조 프레임(ref)을 인가받아 참조 프레임(ref)의 특징을 추출하는 참조 프레임 특징 추출부(212)를 포함할 수 있다.
- [0043] 여기서 키 프레임 특징 추출부(211)와 참조 프레임 특징 추출부(212)는 동일 구조와 동일 가중치를 갖는 삼 신경망(siamese netwrok)으로 구현될 수 있다. 즉 키 프레임 특징 추출부(211)와 참조 프레임 특징 추출부(212)는 동일한 패턴 추정 방식에 따라 키 프레임(key)과 참조 프레임(ref)에서 특징을 추출하여 키 특징맵과 참조 특징맵을 획득할 수 있다.
- [0044] 키 프레임 특징 추출부(211)와 참조 프레임 특징 추출부(212) 각각은 일예로 컨볼루션 신경망(Convolutional Neural Networks: CNN)으로 구현될 수 있다.
- [0045] 특징 결합부(220)는 특징 추출부(211)에서 추출된 키 특징맵과 참조 특징맵을 결합(concatenate)하여 결합 특징맵을 3차원 변환부(230)로 출력한다.
- [0046] 3차원 변환부(230)는 특징 결합부(220)에서 인가되는 결합 특징맵에 3차원 좌표 정보를 추가하여 3차원 특징맵으로 변환한다.
- [0047] 특징 추출부(210)에서 추출되는 키 특징맵과 참조 특징맵은 단지 키 프레임(key)과 참조 프레임(ref)의 특징을 추출하여 표현할 뿐 좌표계 정보를 제공하지 않는다. 따라서 키 특징맵과 참조 특징맵이 결합된 결합 특징맵 또한 깊이 정보를 포함한 좌표 정보를 제공하지 않는다.
- [0048] 그러나 카메라 자세는 2차원이 아닌 3차원 정보를 추정하는 것이므로, 결합 특징맵으로부터 카메라 자세 변화를 그대로 추정하는 것은 매우 어렵다. 이에 본 실시예에서 3차원 변환부(230)는 결합 특징맵에 3차원 좌표 정보

를 나타내는 3차원 좌표계 특징맵을 추가한다. 즉 결합 특징맵에 3차원 좌표계 특징맵을 추가함으로써, 결합 특징맵에 3차원의 월드 좌표계가 포함될 수 있도록 변환한다. 이때, 결합 특징맵에 3차원의 월드 좌표계를 포함하기 위해서는 카메라 내부 파라미터와 각 픽셀에 대응하는 깊이 정보가 필요하다.

[0049] 카메라 내부 파라미터는 초점 거리(f_x , f_y)와 광학축과 영상 평면(image plane)이 만나는 픽셀 좌표인 주점(principal point)(x_0 , y_0)로 구성되며, 카메라 내부 파라미터는 미리 획득된다. 따라서 결합 특징맵의 각 픽셀에 대응하는 깊이(Z)를 알고 있다면 결합 특징맵으로부터 3차원의 월드 좌표계를 추출할 수 있다. 이에 3차원 변환부(230)는 결합 특징맵을 인가받아 3차원 특징맵으로 변환한다.

[0050] 3차원 변환부(230)는 깊이 판별부(300) 또는 와평부(400)로부터 키 프레임에 대해 획득된 키 깊이 정보를 기반으로 3차원 좌표계 특징맵을 획득하고, 획득된 3차원 좌표계 특징맵을 결합 특징맵에 결합하여 3차원 특징맵을 획득한다.

[0051] 이하에서는 도 5 및 도 6을 참조하여 3차원 변환부(230)가 3차원 좌표계 특징맵을 추가하는 방법을 설명한다.

[0052] 도 5에서 (a)는 2차원의 픽셀 좌표계와 3차원의 월드 좌표계 사이의 관계를 나타내고, (b)는 카메라 내부 파라미터와 각 픽셀의 깊이 값에 따른 카메라 좌표계와 월드 좌표계 사이의 관계를 나타낸다.

[0053] 도 5를 참조하면, 카메라 내부 파라미터는 초점 거리(f_x , f_y)와 광학축과 영상 평면(image plane)이 만나는 픽셀 좌표인 주점(principal point)(x_0 , y_0)로 구성되며, 카메라 내부 파라미터는 미리 획득된다. 2차원 영상의 각 픽셀에 대응하는 깊이(Z)를 알고 있다면, 수식식1에 따라 2차원 픽셀 좌표계(x , y)를 카메라를 중심으로 하는 카메라 좌표계(X_c , Y_c , Z_c)로 변환할 수 있다.

수학식 1

$$X_c = \frac{Z}{f_x} \times (x - x_0), \quad Y_c = \frac{Z}{f_y} \times (y - y_0)$$

$$X_c = \frac{Z_c}{f_x} \times (x - x_0), \quad Y_c = \frac{Z_c}{f_y} \times (y - y_0)$$

[0056] 그리고 카메라 내부 파라미터(Camera Intrinsic Parameter)는 캘리브레이션을 통해 캘리브레이션 행렬(K)로서 수학식 2의 형태로 미리 획득된다.

수학식 2

$$K = \begin{bmatrix} f_x & 0 & x_0 \\ 0 & f_y & y_0 \\ 0 & 0 & 1 \end{bmatrix}$$

[0058] 도 6에서 (a)는 결합 특징맵(fm)에 3차원 좌표계 특징맵(cof)이 추가 결합되는 방식을 나타내고, (b)는 3차원 좌표계 특징맵(cof)의 구성의 일례를 나타낸다. 도 6을 참조하면, 결합 특징맵(fm)에 추가 결합되는 3차원 좌표계 특징맵(cof)은 3개의 채널(ch1 ~ ch3)로 구성될 수 있다. 2차원 특징맵은 기지정된 크기의 행렬 형태로 인가되며, 3차원 좌표계 특징맵(cof)의 3개의 채널(ch1 ~ ch3)은 2차원 특징맵에 대응하는 형태로 구성될 수 있다. (b)에 도시된 바와 같이, 3개의 채널(ch1 ~ ch3) 중 제1 채널(ch1)은 x축 좌표계를 나타내는 특징맵으로 x축 방향으로 증가되는 값을 갖는 행렬로 구성될 수 있으며, 제2 채널(ch2)은 y축 좌표계를 나타내는 특징맵으로 y축 방향으로 증가되는 값을 갖는 행렬로 구성될 수 있다. 그리고 제3 채널(ch3)은 동차 좌표계(homogeneous coordinate)를 나타내는 행렬로서 일례로 모든 원소의 값이 1로 설정될 수 있다.

[0059] 다만, 3개의 채널(ch1 ~ ch3)에서는 z축 좌표계를 표현할 수 없다. 이에 3차원 변환부는 키 프레임(key)에 대

한 깊이 정보(D)를 캘리브레이션 행렬(K)의 역 행렬(K^{-1})과 곱하고, 3개의 채널(ch1 ~ ch3)에 곱하여 3차원 좌표계 특징맵을 획득할 수 있다. 즉 3차원 좌표계 특징맵(cof)의 3개의 채널 중 제1 및 제2 채널(ch1, ch2)는 결합 특징맵(fm)에 2차원 좌표계를 부가하기 위한 2차원 좌표계 특징맵으로 볼 수 있다. 따라서 본 실시예에서는 2차원 좌표계 특징맵에 동차 좌표계 특징맵에 해당하는 채널을 추가한 후, 카메라 내부 파라미터의 특성이 반영된 깊이 맵을 z축 좌표로 부가하여 3차원 좌표계 특징맵(cof)을 획득할 수 있다.

- [0060] 여기서 키 프레임(key)에 대한 깊이 정보는 깊이 판별부(300) 또는 와핑부(400)로부터 인가받아 획득할 수 있다.
- [0061] 키 프레임(key)에 대한 깊이 정보는 깊이 판별부(300)가 직접 키 프레임(key)을 인가받아 깊이 맵을 획득하여 전달할 수 있다. 그러나 상기한 바와 같이, 본 실시예에서 키 프레임(key)은 프레임 선택부에 의해 카메라 자세 정보와 깊이 정보가 획득된 참조 프레임(ref) 중 이전 선택된 키 프레임(key) 사이의 차이가 기지정된 기준 값 이상인 프레임의 직전 또는 직후 프레임이 선택될 수 있다. 즉 키 프레임(key)에 대한 깊이 정보는 미리 획득될 수 있으며, 이에 미리 획득된 키 프레임(key)에 대한 깊이 정보가 3차원 변환부(230)에 제공될 수 있다.
- [0062] 즉 본 실시예의 3차원 변환부(230)는 키 프레임(key)에 대해 이전 추정된 깊이 정보를 기반으로 3차원 좌표계 특징맵을 결합 특징맵에 추가하여 3차원으로 변환함으로써 카메라 자세 변화를 더 용이하게 추정할 수 있도록 한다.
- [0063] 자세 변화 추정부(240)는 3차원 변환부(230)로부터 3차원 특징맵을 인가받고, 미리 학습된 패턴 추정 방식에 따라 3차원 특징맵에서 카메라의 자세 변화를 추정한다. 상기한 바와 같이, 3차원 특징맵은 키 특징맵과 참조 특징맵 및 키 프레임(key)에 대한 좌표계 정보가 결합된 특징맵이다. 따라서 회전 변환 추출부(241)는 3차원 특징맵을 인가받아 키 프레임(key)에 대응하는 카메라로부터 참조 프레임(ref)에 대응하는 카메라로의 카메라 위치 및 자세 변화를 추정한다.
- [0064] 자세 변화 추정부(240)는 회전 변환 추출부(241), 이동 벡터 추출부(242), 역 오일러 변환부(243) 및 카메라 자세 획득부(244)를 포함할 수 있다.
- [0065] 카메라 자세 변화는 카메라의 회전 변화와 위치 변화로 구분될 수 있으며, 이에 자세 변화 추정부(240)에서 회전 변환 추출부(241)는 3차원 특징맵을 인가받고 미리 학습된 패턴 추정 방식에 따라 키 프레임(key)과 참조 프레임(ref) 사이에서 카메라의 회전 변화를 추정하고, 이동 벡터 추출부(242)는 카메라의 위치 변화를 추정한다.
- [0066] 회전 변환 추출부(241)는 카메라의 회전 변화를 오일러 각(Euler Angle)을 원소로 갖는 행렬 형태로 추정할 수 있다. 일반적으로 3차원 좌표계에서 회전 변환은 x, y, z 축 방향 각각에서의 회전을 고려하여 3 X 3 크기의 회전 행렬(Rotation Matrix)로 표현된다. 그러나 3 X 3 크기의 회전 행렬은 9개의 원소를 가짐에 따라, 회전 변환 추출부(241)는 카메라의 회전 변화를 9개의 값으로 추정해야 한다. 그러나 인공 신경망을 이용할 지라도 9개의 값을 추정하는 네트워크를 구성하기 어려울 뿐만 아니라 제약 조건이 많다. 또한 학습이 용이하지 않다는 문제가 있다.
- [0067] 이에 본 실시예에 따른 회전 변환 추출부(241)는 카메라의 회전 변화를 용이하게 추정할 수 있도록 3차원 특징맵으로부터 카메라 회전 변화를 오일러 각의 표현 형태로 추정할 수 있다.
- [0068] 회전 행렬(R)로부터 오일러 각으로의 변환은 수학식 3에 따라 수행될 수 있다.

수학식 3

$$R = \begin{bmatrix} r_{00} & r_{01} & r_{02} \\ r_{10} & r_{11} & r_{12} \\ r_{20} & r_{21} & r_{22} \end{bmatrix}$$

$$\phi_z \psi = \tan^{-1} \frac{-r_{12}}{r_{11}}$$

$$\phi_y \theta = \tan^{-1} \frac{r_{13}}{\sqrt{r_{33}^2 + r_{23}^2}}$$

$$\phi_x \phi = \tan^{-1} \frac{-r_{12}}{r_{33}}$$

[0069]

[0070]

오일러 각은 3개 각도(ψ , ϕ , θ)의 조합으로 3차원 좌표에서의 회전 변화를 표현하도록 제안된 표현법이므로, 회전 변환 추출부(241)가 3차원 특징맵으로부터 오일러 각을 추정하도록 학습된 경우, 3개의 각도만을 추정하면 되므로, 용이하게 카메라의 회전 변화를 추정할 수 있다.

[0071]

만일 회전 변환 추출부(241)가 카메라 회전 변화를 오일러 각의 형태로 추정하도록 학습된 경우, 회전 변환 추출부(241)는 오일러 각을 다시 회전 행렬로 역 변환하는 역 오일러 변환부(미도시)를 더 포함할 수 있다.

[0072]

역 오일러 각 변환은 수학식 4에 따라 수행될 수 있다.

수학식 4

$$R = R_z(\psi)R_y(\theta)R_x(\phi)$$

$$R_z(\psi) = \begin{bmatrix} \cos(\psi) & -\sin(\psi) & 0 \\ \sin(\psi) & \cos(\psi) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$R_y(\theta) = \begin{bmatrix} \cos(\theta) & 0 & \sin(\theta) \\ 0 & 1 & 0 \\ -\sin(\theta) & 0 & \cos(\theta) \end{bmatrix}$$

$$R_x(\phi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\phi) & -\sin(\phi) \\ 0 & \sin(\phi) & \cos(\phi) \end{bmatrix}$$

[0073]

[0074]

한편, 이동 벡터 추출부(242)는 3차원 특징맵으로부터 3차원 좌표에서 x, y, z 축 방향으로의 카메라의 평행 이동량을 추정한다. 이동 벡터 추출부(242)는 카메라의 이동을 x, y, z 축 방향에서 추정하므로 3개의 원소(t_x , t_y , t_z)를 갖는 평행 이동 벡터를 획득할 수 있다.

[0075]

카메라 자세 획득부(244)는 회전 변환 추출부(241)에서 획득된 회전 행렬과 이동 벡터 추출부(242)에서 추정된 평행 이동 벡터로부터 카메라의 회전 및 이동 변화를 분석하여 키 프레임(key)에 대비한 참조 프레임에서의 카메라 자세 변화값을 획득한다.

[0076]

한편, 깊이 판별부(300)는 프레임 선택부(120)로부터 키 프레임(key)을 인가받고, 인가된 키 프레임(key)의 깊이 정보를 추정하여 키 프레임 깊이 맵을 획득한다. 깊이 판별부(300)는 깊이 인코딩부(310)와 깊이 디코딩부(320)를 포함할 수 있다.

[0077] 깊이 인코딩부(310)는 미리 학습된 패턴 추정 방식에 따라 인가된 키 프레임(key)을 인코딩한다. 여기서 깊이 인코딩부(310)는 특징 추출부(210)의 키 프레임 특징 추출부(211) 및 참조 프레임 특징 추출부(212)와 유사하게 인가된 키 프레임(key)의 특징을 추출하지만, 특징 추출부(210)와 다른 특징을 추출하도록 학습된다. 그리고 깊이 디코딩부(320)는 깊이 인코딩부(310)에서 인코딩된 키 프레임(key)을 인가받고, 미리 학습된 패턴 복원 방식에 따라 인코딩된 키 프레임(key)을 디코딩하여 키 프레임(key)의 각 픽셀의 깊이 값을 나타내는 키 프레임 깊이 맵을 획득한다.

[0078] 와핑부(400)는 자세 변화 추정부(240)에서 획득된 카메라 자세 변화값과 깊이 판별부(300)에서 획득된 키 프레임(key)에 대한 깊이 맵을 이용하여 참조 프레임(ref)을 키 프레임(key)의 시점으로 와핑할 수 있다.

[0079] 키 프레임(key)의 이미지 좌표계(p_K)와 참조 프레임(ref)의 이미지 좌표계(p_R) 사이의 관계는 수학식 5로 나타낸다.

수학식 5

$$p_R \sim K T_{K \rightarrow R} D_K K^{-1} p_K$$

[0081] 여기서 \sim 는 비동차 좌표계 표현을 나타내고, $T_{K \rightarrow R}$ 은 키 프레임(key)의 이미지 좌표계(p_K)와 참조 프레임(ref)의 이미지 좌표계(p_R)로의 와핑을 나타내며, K 는 카메라 내부 파라미터를 나타내는 캘리브레이션 행렬을 나타내고, D_K 는 키 프레임(key)에 대한 깊이 맵을 나타낸다.

[0082] 도 7에서 (a)는 키 프레임 영상(I_K)을 나타내고, (b)는 참조 프레임 영상(I_R)을 나타내며, (c)는 깊이 판별부(300)에서 추정된 키 프레임 깊이 맵(D_K)을 나타낸다. 그리고 (d)는 와핑된 참조 프레임 영상(\tilde{I}_R)을 나타낸다.

[0083] 도 7에 도시된 바와 같이, 수학식 5를 적용하여 참조 프레임(ref)을 키 프레임(key)으로 와핑시킨 영상(\tilde{I}_R)과 키 프레임 영상(I_K)은 수학식 6과 같이 유사하게 표현되어야 한다.

수학식 6

$$\tilde{I}_R(p_K) = I_R(p_R) = \sum w^{ij} I_R(p_R^{ij}) = I_K(p_K)$$

$$\tilde{I}_R(p_K) \approx I_K(p_K)$$

[0085] 따라서 와핑된 참조 영상(\tilde{I}_R)을 기반으로 깊이 인코딩부(310)에서 획득된 키 프레임 깊이 맵을 보정할 수 있다. 그리고 와핑부(400)는 보정된 키 프레임 깊이 맵을 참조 프레임으로 와핑하여 참조 프레임(ref)에 대한 참조 프레임 깊이 맵을 획득할 수 있다.

[0086] 한편 본 실시예에 따른 카메라 자세 및 깊이 추정 장치는 카메라 자세 판별부(200)와 깊이 판별부(300)를 학습시키기 위한 학습부(미도시)를 더 포함할 수 있다.

[0087] 학습부는 와핑 참조 영상(\tilde{I}_R)과 키 프레임 영상(I_K) 사이의 차이($\tilde{I}_R(p_K) - I_K(p_K)$)를 손실 함수로 정의하고, 손실 함수의 값인 오차가 기지정된 기준 오차 이하가 되도록 오차를 역전파함으로써, 카메라 자세 판별부(200)와 깊이 판별부(300)를 학습시킬 수 있다.

[0088] 상기에서는 깊이 판별부(300)가 키 프레임(key)에 대한 키 프레임 깊이 맵(D_K)을 추정하는 것으로 설명하였으나, 학습이 완료된 깊이 판별부(300)는 참조 프레임(ref)을 인가받고, 참조 프레임(ref)으로부터 직

접 참조 프레임 깊이 맵(D_R)을 추정하도록 구성될 수 있다.

- [0089] 즉 깊이 판별부(300)는 학습 시에는 학습을 위하여 키 프레임(key)에 대한 키 프레임 깊이 맵(D_K)을 추정하는 반면, 카메라 자세 및 깊이 추정 장치가 실제 사용되는 경우에는 참조 프레임(ref)에 대한 참조 프레임 깊이 맵(D_R)을 추정할 수 있다. 그리고 이 경우에는 키 프레임(key)에 대한 키 프레임 깊이 맵(D_K)으로부터 참조 프레임(ref)에 대한 깊이 맵(D_R)을 획득할 필요가 없으므로 와핑부(400)가 생략될 수 있다.
- [0090] 또한 학습 시에도 깊이 판별부(300)가 참조 프레임(ref)에 대한 참조 프레임 깊이 맵(D_R)을 획득하고, 와핑부(400)가 키 프레임(key)을 참조 프레임(ref)으로 와핑시키도록 하여 오차를 획득할 수도 있다.
- [0091] 도 8은 본 발명의 일 실시예에 따른 카메라 자세 및 깊이 추정 방법을 나타낸다.
- [0092] 도 1 내지 도 7을 참조하여, 도 8의 카메라 자세 및 깊이 추정 방법을 설명하면, 우선 다양한 시점에서 연속적으로 획득된 다수 프레임의 영상에서 키 프레임(key)과 참조 프레임(ref)을 선택한다(S10). 여기서 키 프레임(key)은 다수 프레임의 영상 중 최초 프레임 또는 기지정된 프레임이 선택되거나, 이전 선택된 키 프레임과 카메라 자세 또는 깊이 차이 중 적어도 하나가 기지정된 기준값 이상으로 판별된 참조 프레임(ref)을 기준으로 직전 또는 직후 프레임 중 하나가 키 프레임(key)으로 선택될 수 있다.
- [0093] 이후 미리 학습된 패턴 추정 방식에 따라 인가된 키 프레임(key)과 참조 프레임(ref) 각각의 특징을 추출하여 키 특징맵과 참조 특징맵을 획득한다(S20). 키 특징맵과 참조 특징맵이 획득되면, 획득된 키 특징맵과 참조 특징맵을 결합하여 결합 특징맵을 획득한다(S30).
- [0094] 한편 결합 특징맵을 획득하는 과정과 별개로 키 프레임(key)과 참조 프레임(ref) 중 적어도 하나를 인가받아 미리 학습된 패턴 추정 방식 및 패턴 복원 방식에 따라 인가된 프레임(key)을 인코딩 및 디코딩하여 인가된 프레임에 대한 깊이 맵을 추정한다(S40). 깊이 맵 추정은 일례로 학습 시에는 키 프레임(key)에 대한 키 깊이 맵이 추정될 수 있고, 실제 운용시에는 참조 프레임(ref)에 대한 참조 깊이 맵이 추정될 수 있으나, 이에 한정되지 않는다.
- [0095] 그리고 추정된 깊이 맵을 기반으로 생성되는 3차원 좌표계 특징맵을 획득된 결합 특징맵에 결합하여 결합 특징맵을 3차원 변환함으로써 3차원 특징맵을 획득한다(S50).
- [0096] 3차원 특징맵이 획득되면, 획득된 3차원 특징맵으로부터 미리 학습된 패턴 추정 방식에 따라 카메라 회전 변환량을 추정한다(S60). 여기서 회전 변환량은 3×3 크기의 회전 행렬로 직접 추정될 수도 있으나, 본 실시예에서는 3개의 오일러 각(ψ , ϕ , θ)으로 추정된 후, 수학식 4에 따른 역 오일러 각 변환을 통해 회전 변환 행렬로 획득될 수도 있다.
- [0097] 한편 카메라 회전 변환량 추정과 별도로, 미리 학습된 패턴 추정 방식에 따라 3차원 특징맵으로부터 카메라 이동을 추정하여 카메라 이동 벡터를 획득한다(S70).
- [0098] 카메라 회전 변환량을 나타내는 회전 변환 행렬과 카메라의 x, y, z 축 방향의 평행 이동을 나타내는 카메라 이동 벡터가 획득되면, 획득된 회전 변환 행렬과 카메라 이동 벡터를 기반으로 키 프레임(key)에서의 카메라 자세에 대비한 참조 프레임(ref)에서의 카메라 자세 변화량을 판별한다(S80).
- [0099] 한편, 학습 시에는 프레임 영상 와핑 단계(S90)와 오차 계산 및 역전과 단계(S100)가 더 포함될 수 있다.
- [0100] 프레임 영상 와핑 단계(S90)에서는 깊이 맵 추정 단계(S40)에서 추정된 키 프레임(key)에 대한 키 깊이 맵과 판별된 카메라 자세 변화량을 기반으로 참조 프레임(ref)을 키 프레임(key)에 대응하도록 와핑시킨다. 그리고 오차 계산 및 역전과 단계(S100)에서는 와핑된 참조 프레임과 키 프레임 사이의 차이를 오차로서 계산하고, 계산된 오차를 카메라 자세 판별부(200)와 깊이 판별부(300)로 역전과하여 학습을 수행한다.
- [0101] 학습은 오차가 기지정된 기준 오차 이하가 되거나 기지정된 횟수에 도달할 때까지 반복하여 수행될 수 있다.
- [0102] 결과적으로 본 실시예에 따른 카메라 자세 및 깊이 추정 장치 및 방법은 다양한 시점에서 연속적으로 획득된 다수 프레임의 영상에서 키 프레임(key)과 참조 프레임(ref)을 선택하고, 선택된 키 프레임(key)과 참조 프레임(ref) 사이의 카메라 자세 변화와 키 프레임(key)과 참조 프레임(ref) 중 적어도 하나의 깊이 맵을 획득하고, 획득된 카메라 자세 변화와 깊이 맵을 기반으로 키 프레임(key)과 참조 프레임(ref) 중 하나를 다른 하나에 대응하도록 와핑시켜 비교하여 학습시킴으로써, 카메라 자세와 깊이를 모두 정확하게 추정할 수 있도록 학습시킬 수 있다. 특히 키 프레임(key)과 참조 프레임(ref)에 대한 특징맵에 3차원 좌표계 특징맵을 결합하여 카메라

자세 변화를 추정하도록 함으로써, 정확하게 카메라 자세 변화를 추정할 수 있다. 또한 카메라 자세 변화 중 회전 변화를 추정하기 위해 오일러 각을 추정하도록 학습시킴으로써, 더욱 용이하고 정확하게 카메라 자세 변화를 추정할 수 있도록 한다. 뿐만 아니라 키 프레임을 이전 선택된 키 프레임과 카메라 자세 또는 깊이 차이 중 적어도 하나가 기지정된 기준값 이상으로 판별된 참조 프레임(ref)을 기준으로 직전 또는 직후 프레임 중 하나가 키 프레임(key)으로 선택함으로써, 오차가 누적되는 것을 최소화할 수 있다.

[0103] 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.

[0104] 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.

[0105] 따라서, 본 발명의 진정한 기술적 보호 범위는 첨부된 청구범위의 기술적 사상에 의해 정해져야 할 것이다.

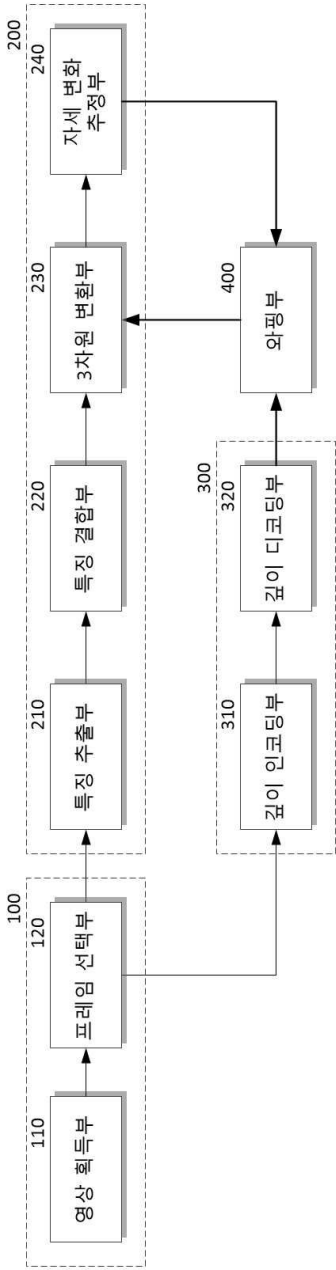
부호의 설명

[0106]

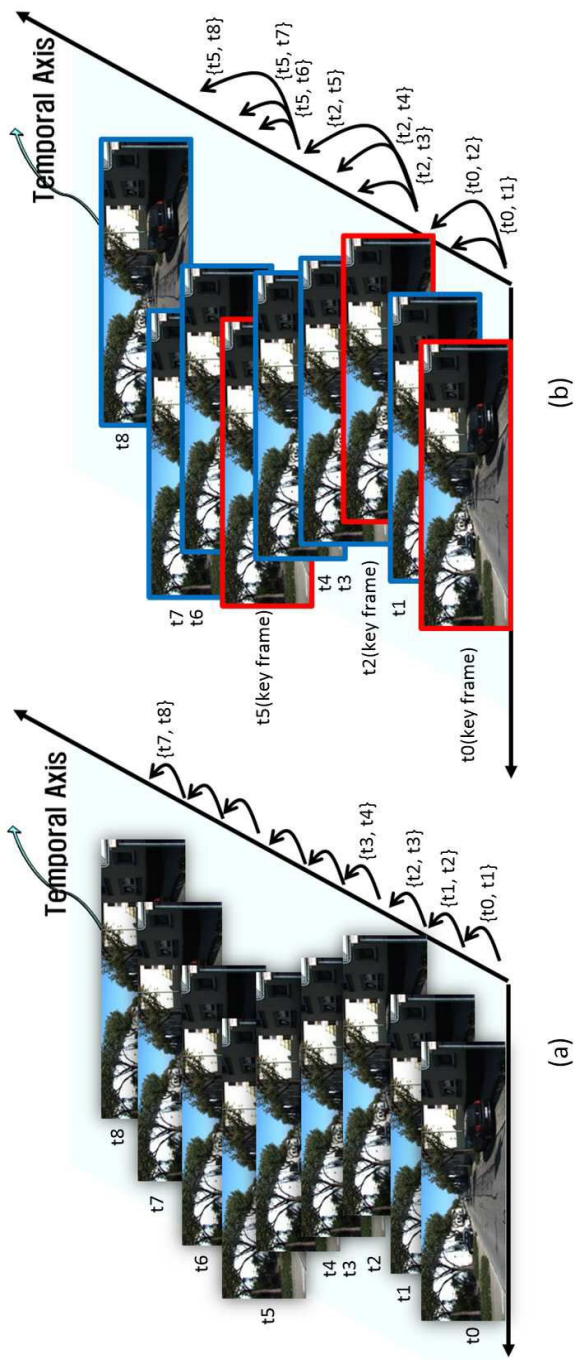
100: 프레임 제공부	110: 영상 획득부
120: 프레임 선택부	200: 자세 판별부
210: 특징 추출부	211: 키 프레임 특징 추출부
212: 참조 프레임 특징 추출부	220: 특징 결합부
230: 3차원 변환부	240: 자세 변화 추정부
241: 회전 변환 추출부	242: 이동 벡터 추출부
243: 카메라 자세 획득부	300: 깊이 판별부
310: 깊이 인코딩부	320: 깊이 디코딩부
400: 와핑부	

도면

도면1



도면2



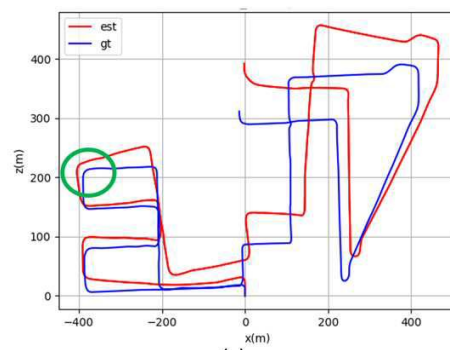
도면3



(a)

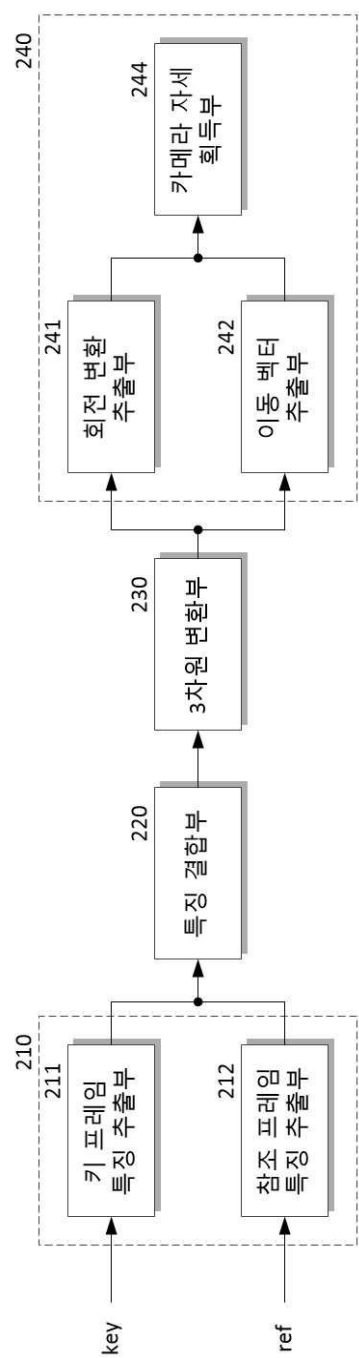


(b)

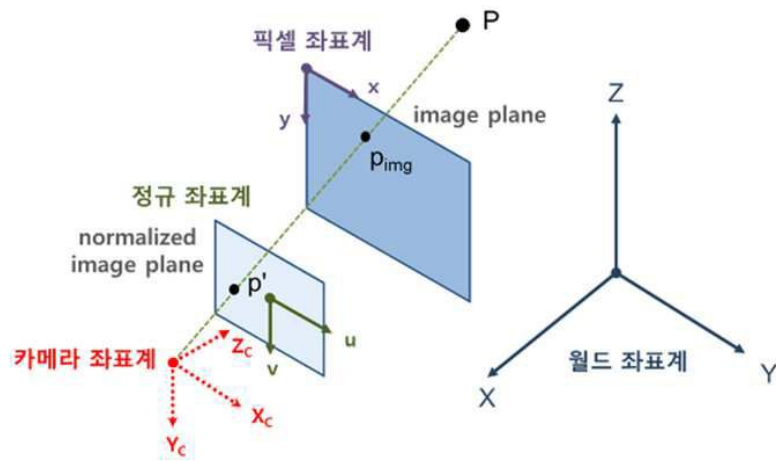


(c)

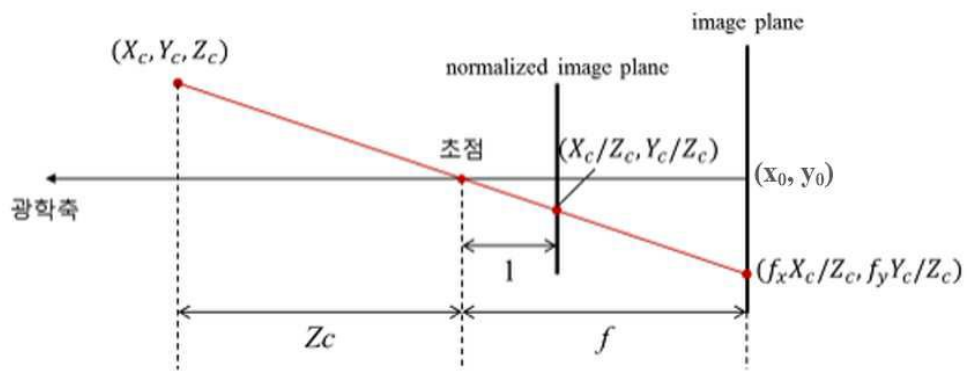
도면4



도면5

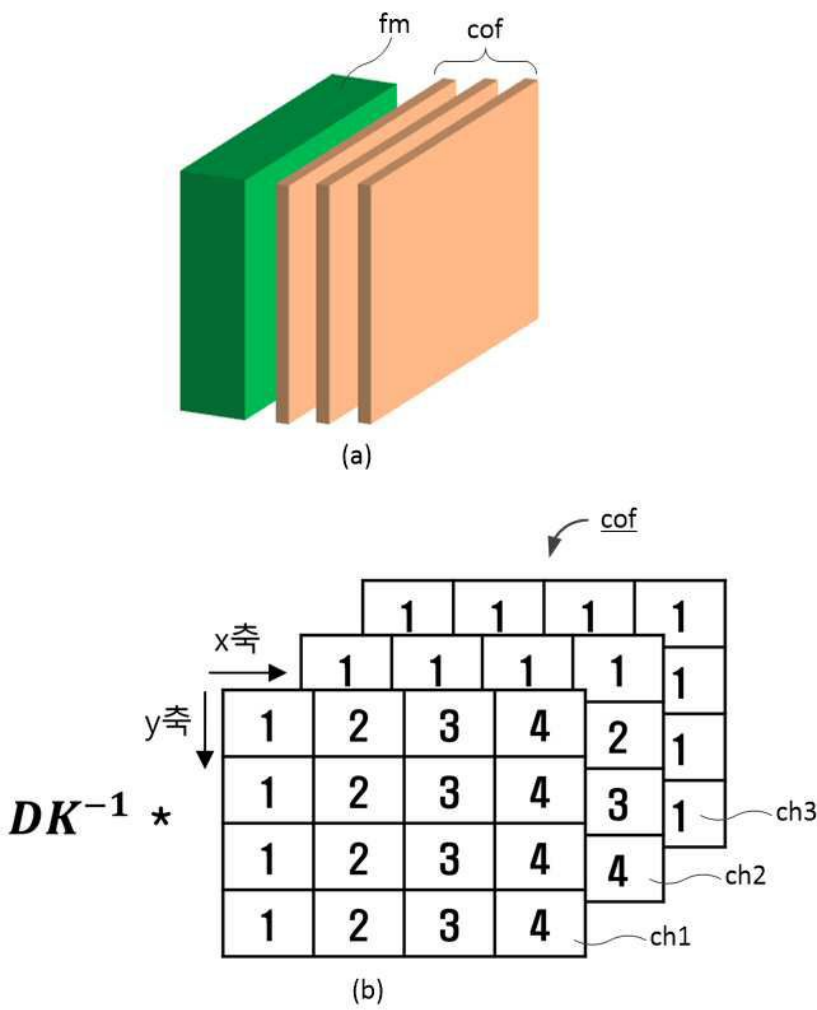


(a)



(b)

도면6



도면7



도면8

