



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2024년07월11일
(11) 등록번호 10-2684771
(24) 등록일자 2024년07월09일

(51) 국제특허분류(Int. Cl.)
 G06T 5/50 (2024.01) G06N 3/08 (2023.01)
 G06N 3/0895 (2023.01) G06T 3/00 (2024.01)
 G06T 5/00 (2024.01) G06T 7/11 (2017.01)
 G06T 7/593 (2017.01)

(52) CPC특허분류
 G06T 5/50 (2024.01)
 G06N 3/084 (2023.01)

(21) 출원번호 10-2023-0035579
 (22) 출원일자 2023년03월20일
 심사청구일자 2023년03월20일

(56) 선행기술조사문헌
 Jiyang Yu ET AL, Learning Video Stabilization Using Optical Flow, 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Page 8156-8164, June 2020(2020.06.13.) 1부.*
 Marwane Hariat ET AL, Rebalancing gradient to improve self-supervised co-training of depth, odometry and optical flow predictions, 2023 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), Page 1267-1276, 3-7 Jan. 2023(2023.01.03.) 1부.*
 Yuhua Chen ET AL, Self-supervised Learning with Geometric Constraints in Monocular Video Connecting Flow, Depth, and Camera, 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Page 7062-7071, 27 Oct. - 2 Nov. 2019(2019.10.27.) 1부.*
 KR1020210055260 A
 *는 심사관에 의하여 인용된 문헌

(73) 특허권자
 연세대학교 산학협력단
 서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)

(72) 발명자
 이상윤
 서울특별시 서초구 청두곳14길 6, 202호(방배동)
 이승훈
 경기도 의왕시 내손로 14, 220동 702호(내손동, 인덕원 센트럴 자이)
 (뒷면에 계속)

(74) 대리인
 민영준

전체 청구항 수 : 총 14 항

심사관 : 김광식

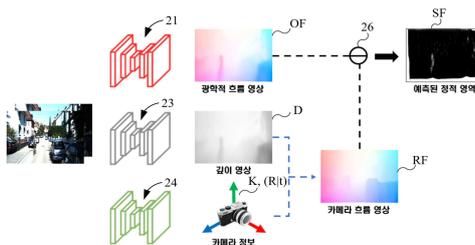
(54) 발명의 명칭 영상을 안정화하는 장치 및 방법

(57) 요약

본 개시는 영상의 다수의 프레임을 인가받아 신경망 연산하여 옵티컬 플로우와 3차원 정보를 획득하고, 상기 3차원 정보에 따라 획득되는 프레임간 카메라의 움직임에 따른 픽셀의 이동을 나타내는 리지드 플로우와 상기 옵티컬 플로우를 비교하여 구분되는 동적 영역과 정적 영역에 따라 정적 플로우를 획득하며, 상기 정적 플로우와 상

(뒷면에 계속)

대표도 - 도2



기 다수의 프레임을 인가받아 신경망 연산을 수행하여 영상을 보정하여, 영상을 효과적으로 안정화할 수 있는 장치 및 방법을 제공한다.

(52) CPC특허분류

G06N 3/0895 (2023.01)

G06T 3/06 (2024.01)

G06T 5/00 (2024.01)

G06T 7/11 (2017.01)

G06T 7/593 (2017.01)

G06T 2207/20081 (2013.01)

G06T 2207/20084 (2013.01)

(72) 발명자

김우진

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

우성민

서울특별시 서대문구 연세로7안길 35, 501호(창천동)

명세서

청구범위

청구항 1

메모리; 및 상기 메모리에 저장된 신경망 모델에 따른 동작의 적어도 일부를 실행하는 프로세서를 포함하여 영상을 안정화하는 장치로서,

상기 프로세서는

영상의 다수의 프레임을 인가받아 신경망 연산하여 옵티컬 플로우와 3차원 정보를 획득하고,

상기 3차원 정보에 따라 획득되는 프레임간 카메라의 움직임에 따른 픽셀의 이동을 나타내는 리지드 플로우와 상기 옵티컬 플로우를 비교하여 구분되는 동적 영역과 정적 영역에 따라 정적 플로우를 획득하며,

상기 정적 플로우와 상기 다수의 프레임을 인가받아 신경망 연산을 수행하여 영상을 보정하고,

상기 프로세서는

상기 장치에 포함된 신경망 모델에 대한 학습 시에 이전 영상과 보정된 영상에서 획득된 상기 옵티컬 플로우 사이의 차이와 상기 리지드 플로우 사이의 차이에 따른 손실을 계산하여 역전파하고, 보정된 영상을 다시 입력하여 보정하는 재귀적 자기 지도 학습을 수행하는 장치.

청구항 2

제1항에 있어서, 상기 프로세서는

상기 옵티컬 플로우와 상기 리지드 플로우에서 서로 대응하는 픽셀에 대한 벡터의 차이에 따라 동적 객체가 포함된 상기 동적 영역을 검출하는 장치.

청구항 3

제1항에 있어서, 상기 프로세서는

상기 옵티컬 플로우 또는 상기 리지드 플로우 중 하나에서 상기 동적 영역에 포함된 벡터를 제거하여 상기 정적 플로우를 획득하는 장치.

청구항 4

제1항에 있어서, 상기 프로세서는

상기 3차원 정보로서 획득된 깊이 맵과 카메라 내부 정보와 외부 정보를 기반으로 각 픽셀을 3차원 투영하고,

3차원 공간 상에서 위치 이동시킨 후 다시 2차원 투영 변환하며,

2차원 투영된 픽셀로의 위치 변화를 기반으로 상기 리지드 플로우를 획득하는 장치.

청구항 5

제1항에 있어서, 상기 프로세서는

상기 3차원 정보로서 획득된 깊이 맵과 카메라의 내부 정보와 외부 정보 중 상기 내부 정보를 기반으로 각 픽셀의 픽셀 좌표를 카메라 좌표로 변환하고,

상기 카메라 좌표로 변환된 픽셀의 좌표에 상기 깊이 맵의 깊이값을 적용하여 3차원 공간에 투영하는 장치.

청구항 6

제5항에 있어서, 상기 프로세서는

3차원 공간에 투영된 픽셀 좌표를 프레임간 카메라의 움직임에 나타내는 상기 외부 정보를 이용하여 3차원 공간 상에서 이동시키고,

3차원 공간에서 이동된 픽셀 좌표를 상기 카메라 내부 정보에 따라 다시 2차원 공간에 투영 변환하는 장치.

청구항 7

삭제

청구항 8

제1항에 있어서, 상기 프로세서는

상기 다수의 프레임에 대해 신경망 연산을 수행하여 노이즈를 제거한 후, 상기 옵티컬 플로우와 상기 3차원 정보를 획득하는 장치.

청구항 9

메모리에 저장된 신경망 모델에 따른 동작의 적어도 일부를 실행하는 프로세서에 의해 수행되는 영상을 안정화하는 방법으로서,

영상의 다수의 프레임을 인가받아 신경망 연산하여 옵티컬 플로우와 3차원 정보를 획득하는 단계;

상기 3차원 정보에 따라 획득되는 프레임간 카메라의 움직임에 따른 픽셀의 이동을 나타내는 리지드 플로우와 상기 옵티컬 플로우를 비교하여 구분되는 동적 영역과 정적 영역에 따라 정적 플로우를 획득하는 단계; 및

상기 정적 플로우와 상기 다수의 프레임을 인가받아 신경망 연산을 수행하여 영상을 보정하는 단계를 포함하되,

상기 신경망 모델에 대한 학습 시에 이전 영상과 보정된 영상에서 획득된 상기 옵티컬 플로우 사이의 차이와 상기 리지드 플로우 사이의 차이에 따른 손실을 계산하여 역전파하고, 보정된 영상을 다시 입력하여 보정하는 재귀적 자기 지도 학습을 수행하는 방법.

청구항 10

제9항에 있어서, 상기 정적 플로우를 획득하는 단계는

상기 옵티컬 플로우와 상기 리지드 플로우에서 서로 대응하는 픽셀에 대한 벡터의 차이에 따라 동적 객체가 포함된 상기 동적 영역을 검출하는 방법.

청구항 11

제9항에 있어서, 상기 정적 플로우를 획득하는 단계는

상기 옵티컬 플로우 또는 상기 리지드 플로우 중 하나에서 상기 동적 영역에 포함된 벡터를 제거하여 상기 정적 플로우를 획득하는 방법.

청구항 12

제9항에 있어서, 상기 정적 플로우를 획득하는 단계는

상기 3차원 정보로서 획득된 깊이 맵과 카메라 내부 정보와 외부 정보를 기반으로 각 픽셀을 3차원 투영하고,

3차원 공간 상에서 위치 이동시킨 후 다시 2차원 투영 변환하며,

2차원 투영된 픽셀로의 위치 변화를 기반으로 상기 리지드 플로우를 획득하는 방법.

청구항 13

제9항에 있어서, 상기 정적 플로우를 획득하는 단계는

상기 3차원 정보로서 획득된 깊이 맵과 카메라의 내부 정보와 외부 정보 중 상기 내부 정보를 기반으로 각 픽셀의 픽셀 좌표를 카메라 좌표로 변환하고,

상기 카메라 좌표로 변환된 픽셀의 좌표에 상기 깊이 맵의 깊이값을 적용하여 3차원 공간에 투영하는 방법.

청구항 14

제13항에 있어서, 상기 정적 플로우를 획득하는 단계는

3차원 공간에 투영된 픽셀 좌표를 프레임간 카메라의 움직임을 나타내는 상기 외부 정보를 이용하여 3차원 공간 상에서 이동시키고,

3차원 공간에서 이동된 픽셀 좌표를 상기 카메라 내부 정보에 따라 다시 2차원 공간에 투영 변환하는 방법.

청구항 15

삭제

청구항 16

제9항에 있어서, 상기 3차원 정보를 획득하는 단계는

상기 다수의 프레임에 대해 신경망 연산을 수행하여 노이즈를 제거한 후, 상기 옵티컬 플로우와 상기 3차원 정보를 획득하는 방법.

발명의 설명

기술 분야

[0001] 본 개시는 영상을 안정화하는 장치 및 방법에 관한 것으로, 영상에서 3차원 정보를 추출하여 2차원 영상을 안정화하는 장치 및 방법에 관한 것이다.

배경 기술

[0002] 카메라를 이용하여 획득되는 2차원 영상에는 흔들림이 자주 발생한다. 영상의 흔들림은 영상이 획득되는 동안 카메라 자체가 안정적인 자세를 유지하지 못하고 움직이는 경우에도 발생하지만, 영상에 포함되는 객체 이동, 즉 동적 객체에 의해서도 발생하게 된다.

[0003] 카메라의 움직임에 의한 흔들림은 획득된 영상 전체에서 균일한 방향 및 크기로 발생되므로 보상이 용이하다. 그에 반해, 동적 객체에 의한 흔들림은 주변 영역과 무관하게 객체의 이동 방향과 속도에 따라 일부 영역에 발생될 뿐만 아니라, 다수의 객체가 서로 상이한 방향 및 속도로 이동할 수도 있다. 즉 영상에서 각 동적 객체에 해당하는 영역별로 서로 상이하게 보상을 수행해야 하므로 보상이 상대적으로 용이하지 않다. 또한 영상에서 카메라의 움직임에 의한 흔들림과 동적 객체의 움직임에 의한 흔들림이 함께 나타나는 경우도 자주 발생하게 된다.

[0004] 기존에도 흔들림이 발생된 영상을 안정화하기 위한 영상 안정화(Video Stabilization) 기법이 다양하게 제안된 바 있다. 그리고 기존의 대표적인 영상 안정화 기법인 옵티컬 플로우(Optical Flow) 기법은 각 픽셀을 이동시켜 흔들림에 대한 보상을 수행한다. 하지만 영상은 3차원 정보를 2D에 투영한 결과이므로, 2차원 정보인 옵티컬 플로우만으로 카메라의 움직임과 동적 객체의 움직임에 의한 흔들림을 모두 보정하여 영상을 안정화시키는 데는 한계가 있다.

발명의 내용

해결하려는 과제

[0005] 본 개시의 목적은 영상에서 2차원 정보인 옵티컬 플로우뿐만 아니라 3차원 정보를 추출하고, 추출된 옵티컬 플로우와 3차원 정보를 융합하여 영상을 보상함으로써 영상을 효과적으로 안정화할 수 있는 장치 및 방법을 제공하는 데 있다.

과제의 해결 수단

[0006] 본 개시의 일 실시 예에 따른, 영상을 안정화하는 장치는 메모리; 및 상기 메모리에 저장된 신경망 모델에 따른 동작의 적어도 일부를 실행하는 프로세서를 포함하여 영상을 안정화하는 장치로서, 상기 프로세서는 영상의 다수의 프레임을 인가받아 신경망 연산하여 옵티컬 플로우와 3차원 정보를 획득하고, 상기 3차원 정보에 따라 획득

득되는 프레임간 카메라의 움직임에 따른 픽셀의 이동을 나타내는 리지드 플로우와 상기 옵티컬 플로우를 비교하여 구분되는 동적 영역과 정적 영역에 따라 정적 플로우를 획득하며, 상기 정적 플로우와 상기 다수의 프레임을 인가받아 신경망 연산을 수행하여 영상을 보정한다.

- [0007] 상기 프로세서는 상기 옵티컬 플로우와 상기 리지드 플로우에서 서로 대응하는 픽셀에 대한 벡터의 차이에 따라 동적 객체가 포함된 상기 동적 영역을 검출할 수 있다.
- [0008] 상기 프로세서는 상기 옵티컬 플로우 또는 상기 리지드 플로우 중 하나에서 상기 동적 영역에 포함된 벡터를 제거하여 상기 정적 플로우를 획득할 수 있다.
- [0009] 상기 프로세서는 상기 3차원 정보로서 획득된 깊이 맵과 카메라 내부 정보와 외부 정보를 기반으로 각 픽셀을 3차원 투영하고, 3차원 공간 상에서 위치 이동시킨 후 다시 2차원 투영 변환하며, 2차원 투영된 픽셀로의 위치 변화를 기반으로 상기 리지드 플로우를 획득할 수 있다.
- [0010] 상기 프로세서는 상기 3차원 정보로서 획득된 깊이 맵과 카메라의 내부 정보와 외부 정보 중 상기 내부 정보를 기반으로 각 픽셀의 픽셀 좌표를 카메라 좌표로 변환하고, 상기 카메라 좌표로 변환된 픽셀의 좌표에 상기 깊이 맵의 깊이값을 적용하여 3차원 공간에 투영할 수 있다.
- [0011] 상기 프로세서는 3차원 공간에 투영된 픽셀 좌표를 프레임간 카메라의 움직임을 나타내는 상기 외부 정보를 이용하여 3차원 공간 상에서 이동시키고, 3차원 공간에서 이동된 픽셀 좌표를 상기 카메라 내부 정보에 따라 다시 2차원 공간에 투영 변환할 수 있다.
- [0012] 상기 프로세서는 상기 장치에 포함된 신경망 모델에 대한 학습 시에 이전 영상과 보정된 영상에서 획득된 상기 옵티컬 플로우 사이의 차이와 상기 리지드 플로우 사이의 차이에 따른 손실을 계산하여 역전파하고, 보정된 영상을 다시 입력하여 보정하는 재귀적 자기 지도 학습을 수행할 수 있다.
- [0013] 상기 프로세서는 상기 다수의 프레임에 대해 신경망 연산을 수행하여 노이즈를 제거한 후, 상기 옵티컬 플로우와 상기 3차원 정보를 획득할 수 있다.
- [0014] 본 개시의 일 실시 예에 따른, 영상을 안정화하는 방법은 영상의 다수의 프레임을 인가받아 신경망 연산하여 옵티컬 플로우와 3차원 정보를 획득하는 단계; 상기 3차원 정보에 따라 획득되는 프레임간 카메라의 움직임에 따른 픽셀의 이동을 나타내는 리지드 플로우와 상기 옵티컬 플로우를 비교하여 구분되는 동적 영역과 정적 영역에 따라 정적 플로우를 획득하는 단계; 및 상기 정적 플로우와 상기 다수의 프레임을 인가받아 신경망 연산을 수행하여 영상을 보정하는 단계를 포함한다.

발명의 효과

- [0015] 상술한 바와 같이 본 개시의 영상을 안정화하는 장치 및 방법은 영상으로부터 2차원 정보인 옵티컬 플로우뿐만 아니라 3차원 정보를 추출하고, 추출된 옵티컬 플로우와 3차원 정보를 융합하여 보상을 수행함으로써, 영상 안정화 성능을 크게 향상시킬수 있다.

도면의 간단한 설명

- [0016] 도 1은 일 실시예에 따른 영상을 안정화하는 장치를 동작에 따라 구분한 구성을 나타낸다.
- 도 2는 도 1의 정적 영역 추출 모듈의 동작을 설명하기 위한 도면이다.
- 도 3은 도 1의 정적 영역 추출 모듈의 동작이 카메라 움직임을 추정하는 방식을 설명하기 위한 동작을 나타낸다.
- 도 4은 도 1의 영상을 안정화하는 장치의 학습 방식을 설명하기 위한 도면이다.
- 도 5는 일 실시예에 따른 영상을 안정화하는 방법을 나타낸다.
- 도 6은 일 실시예에 따른 컴퓨팅 장치를 포함하는 컴퓨팅 환경을 설명하기 위한 도면이다.

발명을 실시하기 위한 구체적인 내용

- [0017] 이하, 도면을 참조하여 본 개시의 실시예에 따른 구체적인 실시형태를 설명하기로 한다. 이하의 상세한 설명은 본 명세서에서 기술된 방법, 장치 및/또는 시스템에 대한 포괄적인 이해를 돕기 위해 제공된다. 그러나 이는

예시에 불과하며 본 발명은 이에 제한되지 않는다.

- [0018] 본 개시의 실시예들을 설명함에 있어서, 본 발명과 관련된 공지기술에 대한 구체적인 설명이 실시예의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명을 생략하기로 한다. 그리고, 후술되는 용어들은 본 발명에서의 기능을 고려하여 정의된 용어들로서 이는 사용자, 운용자의 의도 또는 관례 등에 따라 달라질 수 있다. 그러므로 그 정의는 본 명세서 전반에 걸친 내용을 토대로 내려져야 할 것이다. 상세한 설명에서 사용되는 용어는 단지 일 실시예들을 기술하기 위한 것이며, 결코 제한적이어서는 안 된다. 명확하게 달리 사용되지 않는 한, 단수 형태의 표현은 복수 형태의 의미를 포함한다. 본 설명에서, "포함" 또는 "구비"와 같은 표현은 어떤 특성들, 숫자들, 단계들, 동작들, 요소들, 이들의 일부 또는 조합을 가리키기 위한 것이며, 기술된 것 이외에 하나 또는 그 이상의 다른 특성, 숫자, 단계, 동작, 요소, 이들의 일부 또는 조합의 존재 또는 가능성을 배제하도록 해석되어서는 안 된다. 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.
- [0019] 도 1은 일 실시예 따른 영상을 안정화하는 장치를 동작에 따라 구분한 구성을 나타내고, 도 2는 도 1의 정적 영역 추출 모듈의 동작을 설명하기 위한 도면이며, 도 3은 도 1의 정적 영역 추출 모듈의 동작이 카메라 움직임을 추정하는 방식을 설명하기 위한 동작을 나타낸다.
- [0020] 도 1 내지 도 3을 참조하면, 본 개시의 실시예에 따른 영상을 안정화하는 장치는 영상 획득 모듈(10), 정적 플로우 획득 모듈(20) 및 영상 보정 모듈(30)을 포함할 수 있다.
- [0021] 영상 획득 모듈(10)은 연속하는 다수의 프레임으로 구성된 영상을 획득한다. 영상 획득 모듈(10)은 CCTV나 카메라 등과 같은 영상 획득 장치로 구현되어 영상을 획득할 수 있으며, 다수의 영상이 저장된 데이터베이스(미도시)로 구현되거나 외부의 장치와 네트워크를 통해 이미지를 인가받는 통신 모듈 등으로 구현되어 영상을 획득할 수도 있다.
- [0022] 정적 플로우 획득 모듈(20)은 영상 획득 모듈(10)에서 획득된 영상으로부터 영상의 흐림을 보상하기 위한 2차원 정보 및 3차원 정보를 추출하고, 추출된 2차원 정보 및 3차원 정보를 융합하여 동적 객체 영역이 제거된 정적 영역에 대한 정적 플로우를 획득한다.
- [0023] 정적 플로우 획득 모듈(20)은 2차원 기반 움직임 정보인 옵티컬 플로우(OF)를 획득하는 옵티컬 플로우 획득 모듈(21)과 3차원 기반 움직임 정보인 리지드 플로우(RF)를 획득하는 3차원 플로우 획득 모듈(22) 및 옵티컬 플로우(OF)와 리지드 플로우(Rigid flow)(RF) 사이의 차이를 기반으로 동적 영역에 대한 마스크를 생성하는 정적 영역 검출 모듈(26)을 포함할 수 있다.
- [0024] 옵티컬 플로우 획득 모듈(21)은 영상으로부터 2차원 정보로서 옵티컬 플로우(OF)를 획득하고, 3차원 플로우 획득 모듈(22)은 3차원 정보로서 깊이 맵(D)과 카메라 정보(K, (R|t))를 추출하여, 추출된 깊이 맵(D)과 카메라 정보(K, (R|t))를 기반으로 카메라의 움직임에 따른 영상의 각 프레임간 변화를 나타내는 리지드 플로우(RF)(또는 카메라 흐름 영상이라 함)를 획득한다. 그리고 정적 영역 검출 모듈(26)은 획득된 옵티컬 플로우(OF)와 리지드 플로우(RF) 사이의 차이를 분석하여 동적 영역과 정적 영역을 구분하고, 구분된 동적 영역과 정적 영역 중 동적 영역이 제외된 정적 영역에 대한 옵티컬 플로우(또는 리지드 플로우)를 추출하여 정적 플로우(SF)를 획득한다.
- [0025] 이후 영상 보정 모듈(30)은 영상 획득 모듈(10)에서 획득된 영상에 대해 정적 플로우 획득 모듈(20)에서 획득된 정적 플로우(SF)를 기반으로 보정을 수행하여 영상의 흐림을 보상한다.
- [0026] 구체적으로 정적 플로우 획득 모듈(20)에서 옵티컬 플로우 획득 모듈(21)은 영상의 다수의 프레임에서 서로 대응하는 픽셀을 검출하고, 검출된 픽셀들의 프레임에 따른 위치 변화를 나타내는 벡터를 추출함으로써 옵티컬 플로우(OF)를 획득한다. 옵티컬 플로우 획득 모듈(21)은 둘 이상의 다수의 프레임을 인가받아 신경망 연산하여 옵티컬 플로우(OF)(광학적 흐름 영상이라고도 함)를 획득하는 신경망 모델로 구현될 수 있다. 프레임에 따른 픽셀 위치 변화로부터 옵티컬 플로우를 추출하는 신경망 모델이 이미 공개되어 있으므로, 기존에 공개된 신경망 모델을 이용하여 구현될 수도 있다.
- [0027] 한편, 3차원 플로우 획득 모듈(22)은 3차원 정보로서 깊이 맵(D)을 획득하는 깊이 맵 획득 모듈(23)과 카메라 정보를 획득하는 카메라 정보 추출 모듈(24) 및 리지드 플로우 생성 모듈(25)을 포함할 수 있다. 깊이 맵 획득 모듈(23)은 다수의 프레임을 입력으로 인가받아 신경망 연산하여 다수의 프레임에서 각 픽셀에 대한 깊이 정보를 추출하여 깊이 맵(D)(깊이 영상이라고도 함)을 획득한다. 기존에는 깊이 맵을 획득하기 위해서는 스테레오

카메라를 이용해야만 하였으나, 최근에는 신경망 기술의 발전으로 인해 단안 카메라에서 획득된 영상의 다수의 프레임으로부터도 깊이 맵(D)을 획득할 수 있다. 단안 카메라에서 획득된 영상에서 깊이 맵(D)을 획득하는 신경망 모델도 공개되어 있으므로, 깊이 맵 획득 모듈(23) 또한 공개된 신경망 모델을 이용하여 구현될 수도 있다.

[0028] 한편, 카메라 정보 추출 모듈(24) 또한 획득된 영상의 다수의 프레임을 인가받아 신경망 연산을 수행하여 카메라 정보를 획득한다. 카메라 정보에는 내부 정보(또는 내부 파라미터라고 함)(K)와 외부 정보(또는 외부 파라미터라고 함)(R|t)가 포함될 수 있다. 카메라 내부 정보(K)는 카메라 내부에 설정되는 파라미터로서, 카메라 좌표와 픽셀 좌표(또는 영상 좌표라고 함) 사이의 변환 관계를 나타낸다. 카메라 내부 정보(K)에는 카메라의 초점 거리(focal length), 뒤틀림 계수(skew coefficient) 및 주점(principal point)이 포함된 행렬로 표현될 수 있다. 한편, 카메라 외부 정보(R|t)는 카메라 자체의 움직임을 나타내는 정보로서, 3차원 공간 상에서 카메라의 움직임을 나타낸다. 카메라 외부 정보(R|t)는 카메라의 회전(rotation, R) 및 평행 이동(translation, t)을 포함하는 행렬로 표현될 수 있다.

[0029] 다수의 프레임을 갖는 영상으로부터 카메라의 내부 정보(K)와 외부 정보(R|t)를 획득하는 신경망 모델이 최근 공개되었다. 따라서 공개된 신경망을 이용하여 카메라 내부 정보(K)와 외부 정보(R|t)를 획득할 수 있다. 여기서는 이해의 편의를 위하여 카메라 정보 추출 모듈(24)이 카메라 내부 정보(K)와 외부 정보(R|t)를 함께 획득하는 것으로 도시하였으나, 카메라 정보 추출 모듈(24)은 카메라 내부 정보(K)를 획득하기 위한 신경망 모델과 외부 정보(R|t)를 획득하기 위한 신경망 모델을 구분하여 각각 별도로 구비할 수도 있다.

[0030] 리지드 플로우 생성 모듈(25)은 깊이맵 획득 모듈(23)에서 획득된 깊이 맵(D)과 카메라 정보 추출 모듈(24)에서 획득된 카메라 내부 정보(K) 및 카메라 외부 정보(R|t)를 기반으로 프레임 사이의 카메라 움직임에 따른 픽셀의 위치 변화를 나타내는 리지드 플로우(RF)를 획득한다. 도 3에 도시된 바와 같이, 리지드 플로우 생성 모듈(25)은 우선 영상의 프레임의 각 픽셀에 대한 픽셀 좌표(p(x,y))를 카메라 내부 정보(K)의 역행렬(K^{-1})을 이용하여 카메라 좌표로 변환한다. 일반적으로 카메라 내부 정보(K)는 카메라의 초점 거리, 뒤틀림 계수 및 주점이 포함되는 3×3 행렬 형식으로 획득된다. 도 3에서는 카메라 내부 정보(K)의 역행렬(K^{-1})에서 f_x 와 f_y 는 각각 x축 및 y축 방향에서의 픽셀 단위 초점 거리를 나타내고, $w/2$, $h/2$ 는 주점의 위치로서 여기서는 주점이 이미지의 중심인 경우를 가정하였다. 그리고 뒤틀림 계수는 0인 경우를 가정하였다.

[0031] 그리고 카메라 좌표로 변환된 각 픽셀에 깊이 맵(D)에 따른 깊이 좌표값(D(p(x,y)))을 추가로 부여하여 3차원 좌표로 변환한다. 이후 카메라 외부 정보(R|t)를 이용하여 3차원 좌표로 변환된 각 픽셀의 좌표를 변환함으로써, 각 픽셀의 좌표(p(x,y))가 프레임에 따른 카메라의 움직임이 반영된 좌표로 변환되도록 한다. 여기서 카메라 외부 정보(R|t)는 도 3에 도시된 바와 같이, 3D에서의 카메라 회전을 나타내는 9개의 회전 파라미터($R_{11} \sim R_{13}$, $R_{21} \sim R_{23}$, $R_{31} \sim R_{33}$)와 3D에서의 평행 이동을 나타내는 3개의 이동 파라미터($t_1 \sim t_3$)를 포함하는 4×3 행렬 형식으로 획득될 수 있다.

[0032] 각 픽셀의 픽셀 좌표(p(x,y))가 이미 카메라 좌표로 변환되고, 깊이 좌표값(D(p(x,y)))이 추가되어 3차원 좌표로 변환되었으므로, 카메라 외부 정보(R|t)에 에 따라 추가 변환을 수행하게 되면, 특정 프레임에서의 픽셀(p(x,y))의 3차원 좌표는 다음 프레임으로의 카메라의 움직임에 대응하여 이동하게 된다.

[0033] 카메라 외부 정보(R|t)에 의해 각 픽셀의 3차원 좌표가 이동하도록 변환되면, 이동 변환된 프레임의 3차원 좌표에 대해 카메라 내부 정보(K)를 이용하여 다시 영상 좌표계의 픽셀 좌표(p'(x,y))로 변환한다. 즉 3차원 픽셀 좌표를 2차원 픽셀 좌표로 재변환하여 변환 픽셀 좌표(p'(x,y))를 획득한다. 이때 카메라 내부 정보(K)는 4×3 행렬인 카메라 외부 정보(R|t)에 대응하여 도 3에 도시된 바와 같이, 4×3 행렬 형식으로 수정되어 변환을 수행할 수 있다.

[0034] 따라서 리지드 플로우 생성 모듈(25)은 다수의 프레임의 각 픽셀 좌표(p(x, y))를 카메라의 내부 정보와 객체의 위치에 따른 깊이 정보 및 카메라의 움직임을 반영하여 3차원 좌표 형식으로 변환한 후, 이를 다시 카메라 내부 정보(K)에 따른 2차원 픽셀 좌표(p'(x,y))로 변환함으로써, 다수의 프레임의 픽셀 좌표가 3차원 위치 정보 및 카메라 움직임이 고려된 2차원 공간에 다시 투영되도록 한다.

[0035] 그리고 리지드 플로우 생성 모듈(25)은 다수의 프레임 각각에서 다수의 픽셀 각각의 픽셀 좌표(p(x, y))가 변환 픽셀 좌표(p'(x,y))로 변환되면, 각 픽셀 좌표(p(x, y))가 변환 픽셀 좌표(p'(x,y)) 사이의 위치 변화에 따른 벡터를 추출하여 리지드 플로우(RF)를 획득한다. 즉 리지드 플로우(RF)는 다수의 프레임 사이에서 카메라 움직

입에 의한 각 픽셀의 위치 변화를 나타낸다. 따라서 해당 프레임에 동적 객체가 포함되어 있지 않다면, 변환 픽셀 좌표($p'(x,y)$)는 다음 프레임에서 대응하는 픽셀의 픽셀 좌표와 동일해야 한다.

- [0036] 정적 영역 검출 모듈(26)은 옵티컬 플로우 획득 모듈(21)에서 획득된 옵티컬 플로우(OF)와 3차원 플로우 획득 모듈(22)에서 획득된 리지드 플로우(RF)를 기반으로 동적 객체가 포함된 동적 영역과 동적 객체가 포함되지 않은 정적 영역을 구분하고, 구분된 동적 영역이 제거된 정적 플로우를 획득한다.
- [0037] 리지드 플로우(RF)가 카메라의 움직임에 따른 픽셀의 위치 변화를 나타내므로, 영상에 동적 객체가 포함되지 않은 경우, 각 픽셀에 대한 옵티컬 플로우(OF)와 리지드 플로우(RF)의 벡터는 서로 유사하게 나타나게 된다. 따라서 옵티컬 플로우(OF)와 리지드 플로우(RF)에서 서로 상이한 벡터값을 갖는 픽셀은 동적 객체가 포함된 영역으로 볼 수 있다. 이에 정적 영역 검출 모듈(26)은 옵티컬 플로우(OF)와 리지드 플로우(RF)에서 서로 대응하는 픽셀에 대한 벡터 사이의 차를 계산하고, 계산된 차가 기준값 이상인 픽셀의 영역을 동적 영역으로 검출한다.
- [0038] 그리고 정적 영역 검출 모듈(26)은 획득된 옵티컬 플로우에서 검출된 동적 영역을 제거하여 정적 영역에만 옵티컬 플로우가 적용된 정적 플로우(SF)를 획득할 수 있다. 일 예로 정적 영역 검출 모듈(26)은 검출된 동적 영역에 따라 동적 영역 마스크를 생성하고, 생성된 동적영역 마스크를 옵티컬 플로우에 적용함으로써, 정적 플로우(SF)를 획득할 수도 있다.
- [0039] 영상 보정 모듈(30)은 영상에 포함된 다수의 프레임과 정적 플로우 획득 모듈(20)에서 획득된 정적 플로우를 인가받아 신경망 연산하여, 다수의 프레임 각각에서의 흐림을 보상한다.
- [0040] 상기한 바와 같이, 기존에 영상을 안정화하는 장치 또한 옵티컬 플로우(OF)를 이용하여 보상을 수행하여 영상을 안정화하였다. 그러나 옵티컬 플로우가 2D에서의 픽셀의 위치 변화만을 고려하므로, 카메라의 움직임뿐만 아니라 동적 객체의 움직임이 함께 포함된 영상에서의 픽셀 위치 변화를 정확하게 검출하지 못하는 한계가 있었다. 이러한 옵티컬 플로우(OF)의 한계로 인해, 기존에는 잘못된 옵티컬 플로우(OF) 정보를 이용하여 영상의 흐림을 보상하게 되어, 영상을 정상적으로 안정화시키지 못하는 문제가 있었다. 그러나 본 개시의 영상을 안정화하는 장치는 옵티컬 플로우(OF)와 함께 3차원 정보로서 깊이 맵, 카메라 내부 정보 및 외부 정보를 획득하고, 획득된 3차원 정보를 기반으로 카메라의 움직임에 따른 픽셀의 위치 변화를 나타내는 리지드 플로우(RF)를 생성하며, 생성된 리지드 플로우(RF)와 옵티컬 플로우(OF) 사이의 차이에 따라 동적 영역을 검출하여 옵티컬 플로우(OF)에서 제거한 정적 플로우(SF)를 이용하여 흐림을 보상하도록 한다. 따라서 영상 보정 모듈(30)은 상대적으로 정확한 픽셀의 위치 변화를 나타내는 정적 플로우(SF)를 기반으로 다수의 프레임에서 각 픽셀을 보상함으로써, 영상의 흐림을 보정할 수 있다.
- [0041] 상기에서는 옵티컬 플로우 획득 모듈(21)과 깊이 맵 획득 모듈(23) 및 카메라 정보 추출 모듈(24)이 독립적인 신경망 모델로 구현되는 것으로 도시하였으나, 옵티컬 플로우 획득 모듈(21)과 깊이 맵 획득 모듈(23) 및 카메라 정보 추출 모듈(24)을 구현하는 신경망 모델은 공통 레이어를 갖도록 구성될 수 있다. 예로서, 각 신경망 모델에서 특징을 추출하는 인코딩 레이어는 공통으로 구성되고, 추출된 특징으로부터 옵티컬 플로우와 깊이 맵 및 카메라 정보를 각각 추출하는 디코딩 레이어는 서로 상이하게 구성될 수도 있다.
- [0042] 또한 도시하지 않았으나, 영상을 안정화하는 장치는 영상 획득 모듈(10)에서 획득된 영상을 인가받아 노이즈를 제거하여 정적 플로우 획득 모듈(20)로 전달하는 노이즈 제거 모듈(미도시)이 더 포함될 수 있다. 여기서 노이즈 제거 모듈 또한 영상의 다수의 프레임을 인가받아 신경망 연산하여 노이즈를 제거하도록 학습된 신경망 모델로 구현될 수 있다.
- [0043] 도 4은 도 1의 영상을 안정화하는 장치의 학습 방식을 설명하기 위한 도면이다.
- [0044] 도 1에 도시된 본 개시의 영상을 안정화하는 장치는 다수의 신경망 모델을 포함하여 구성됨에 따라 학습이 수행되어야 한다. 다만 여기서는 재귀적 반복을 통해 별도의 학습 데이터를 필요로 하지 않는 자기 지도 학습 방식으로 학습될 수 있다.
- [0045] 도 1의 영상을 안정화하는 장치의 동작은 도 4에 도시된 바와 같이, 크게 정적 플로우 획득 모듈(20)에서의 동작과 영상 보정 모듈(30)에서의 동작으로 구분된다. 정적 플로우 획득 모듈(20)은 영상 획득 모듈(10)에서 획득된 영상을 인가받아, 2차원 정보인 옵티컬 플로우(OF)와 3차원 정보를 추출한 후 다시 2차원 로 투영한 리지드 플로우(RF) 사이의 차이를 기반으로 동적 영역을 검출하고, 검출된 동적 영역을 옵티컬 플로우(OF)에서 제거하여 정적 플로우(SF)를 획득한다.
- [0046] 그리고 영상 보정 모듈(30)은 정적 플로우 획득 모듈(20)에서 획득된 정적 플로우(SF)를 기반으로 영상의 흐림

을 보정하여 영상 안정화를 수행한다.

[0047] 따라서 재귀적 학습에 의해 안정화된 영상이 영상을 안정화하는 장치의 입력으로 재귀적으로 반복 입력되면, 이전 획득된 옵티컬 플로우(OF)와 현재 반복 입력되어 획득된 옵티컬 플로우(OF') 사이의 차이가 점차 줄어들게 된다. 따라서 이전 옵티컬 플로우(OF)와 현재 옵티컬 플로우(OF') 사이의 차이가 기준값 이하가 되도록 옵티컬 플로우 획득 모듈(21)과 영상 보정 모듈(30)을 학습시킬 수 있다. 다만 여기서는 옵티컬 플로우 획득 모듈(21)과 영상 보정 모듈(30)이 이미 학습된 신경망 모델을 이용하는 것으로 가정하므로, 사전에 프리 트레이닝된 것으로 가정한다.

[0048] 한편 리지드 플로우(RF)의 경우, 다수의 프레임 사이에서 특정 픽셀의 3차원 공간 상의 일치성(Consistency)이 유지되어야 한다. 즉 리지드 플로우(RF)에 따라 두 프레임에서 서로 대응하는 픽셀이 3차원 투영된 3차원 좌표(x, x')는 서로 동일해야 한다.

[0049] 이에 3차원 플로우 획득 모듈(22)은 수학식 1의 3차원 손실(L_{3D})에 따라 학습될 수 있다.

수학식 1

[0050]
$$L_{3D} = \|x' - x\|_1 = \left\| \begin{bmatrix} D(p')K^{-1}p' \\ 1 \end{bmatrix} - \text{Rlt}_{matrix} \begin{bmatrix} D(p)K^{-1}p \\ 1 \end{bmatrix} \right\|$$

[0051] 또한 재귀적 학습 과정 동안 획득되는 옵티컬 플로우(OF)와 리지드 플로우(RF) 사이의 차에 따른 플로우 손실을 추가적으로 계산하여 함께 학습에 반영함으로써 옵티컬 플로우 획득 모듈(21)과 3차원 플로우 획득 모듈(22) 및 영상 보정 모듈(30)의 신경망 모듈에 대한 파인 튜닝 학습을 수행할 수 있다.

[0052] 도시된 실시예에서, 각 구성들은 이하에 기술된 것 이외에 상이한 기능 및 능력을 가질 수 있고, 이하에 기술되지 것 이외에도 추가적인 구성을 포함할 수 있다. 또한, 일 실시예에서, 각 구성은 물리적으로 구분된 하나 이상의 장치를 이용하여 구현되거나, 하나 이상의 프로세서 또는 하나 이상의 프로세서 및 소프트웨어의 결합에 의해 구현될 수 있으며, 도시된 예와 달리 구체적 동작에 있어 명확히 구분되지 않을 수 있다.

[0053] 그리고 도 1에 도시된 영상을 안정화하는 장치는 하드웨어, 펌웨어, 소프트웨어 또는 이들의 조합에 의해 로직 회로 내에서 구현될 수 있고, 범용 또는 특정 목적 컴퓨터를 이용하여 구현될 수도 있다. 장치는 고정배선형(Hardwired) 기기, 필드 프로그램 가능한 게이트 어레이(Field Programmable Gate Array, FPGA), 주문형 반도체(Application Specific Integrated Circuit, ASIC) 등을 이용하여 구현될 수 있다. 또한, 장치는 하나 이상의 프로세서 및 컨트롤러를 포함한 시스템온칩(System on Chip, SoC)으로 구현될 수 있다.

[0054] 뿐만 아니라 영상을 안정화하는 장치는 하드웨어적 요소가 마련된 컴퓨팅 장치 또는 서버에 소프트웨어, 하드웨어, 또는 이들의 조합하는 형태로 탑재될 수 있다. 컴퓨팅 장치 또는 서버는 각종 기기 또는 유무선 통신망과 통신을 수행하기 위한 통신 모듈 등의 통신장치, 프로그램을 실행하기 위한 데이터를 저장하는 메모리, 프로그램을 실행하여 연산 및 명령하기 위한 마이크로프로세서 등을 전부 또는 일부 포함한 다양한 장치를 의미할 수 있다.

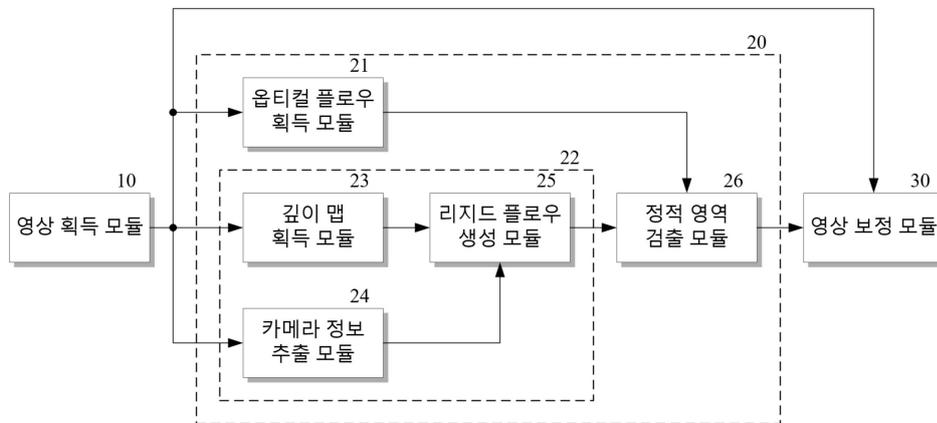
[0055] 도 5는 일 실시예 따른 영상을 안정화하는 방법을 나타낸다.

[0056] 도 1 내지 도 4를 참조하여, 본 개시에 따른 영상을 안정화하는 방법을 설명하면, 우선 다수의 프레임을 갖는 영상을 획득한다(51). 그리고 획득된 영상의 다수의 프레임에 대해 신경망 연산을 수행하여 옵티컬 플로우(OF)를 획득한다(52). 이와 함께 다수의 프레임에 대해 신경망 연산을 수행하여 깊이 맵(D)을 획득한다(53). 또한 카메라 내부 정보(K)와 외부 정보(Rlt)를 다수의 프레임에 대한 신경망 연산으로 획득한다(54). 여기서는 옵티컬 플로우(OF)와 깊이 맵(D) 및 카메라 정보(K, (Rlt))를 순차적으로 획득하는 것으로 도시하였으나, 옵티컬 플로우(OF)와 깊이 맵(D) 및 카메라 정보(K, (Rlt))는 병렬적으로 함께 획득될 수 있다.

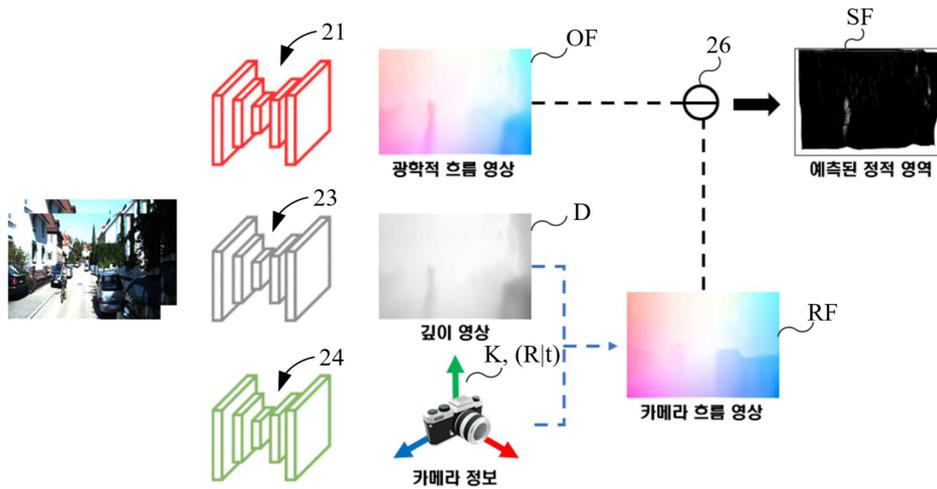
[0057] 깊이 맵(D)과 카메라 정보(K, (Rlt))가 획득되면, 프레임의 각 픽셀의 픽셀 좌표(p(x,y))를 획득된 깊이 맵(D)과 카메라 정보(K, (Rlt))에 따라 이동 변환시켜 변환 픽셀 좌표(p'(x,y))를 획득하여 리지드 플로우를 획득한다(55). 이를 위해 프레임의 각 픽셀을 카메라 내부 정보(K)의 역행렬(K⁻¹)을 이용하여 카메라 좌표로 변환하고, 깊이 맵(D)의 깊이 정보를 부가하여 3차원 공간에 투영하고, 투영된 3차원 픽셀 좌표를 카메라 외부 정보(Rlt)를 이용하여 프레임간 카메라의 움직임에 따른 3차원 공간으로 이동시킨 후, 다시 카메라 내부 정보

도면

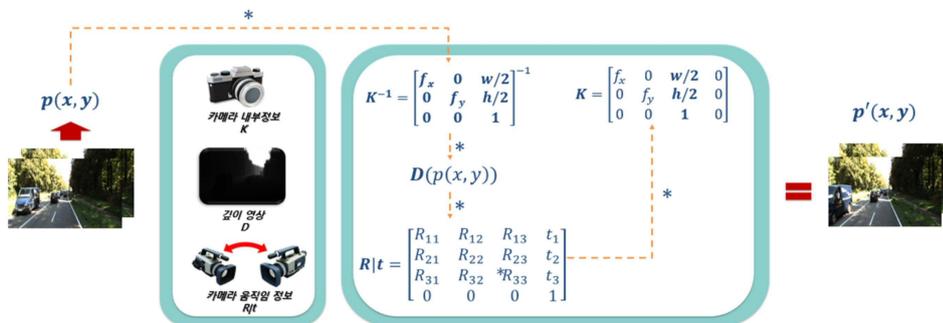
도면1



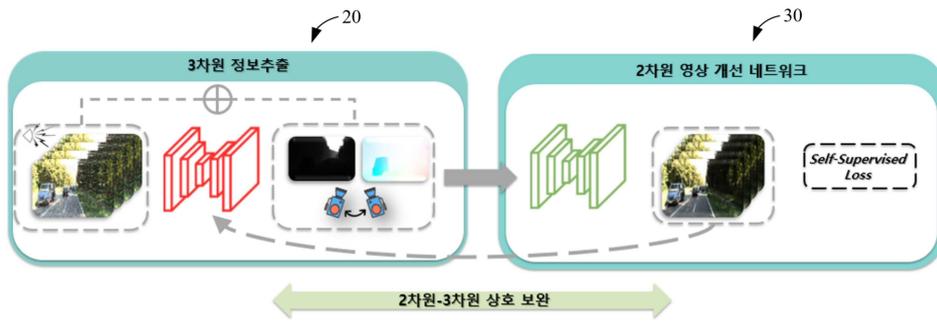
도면2



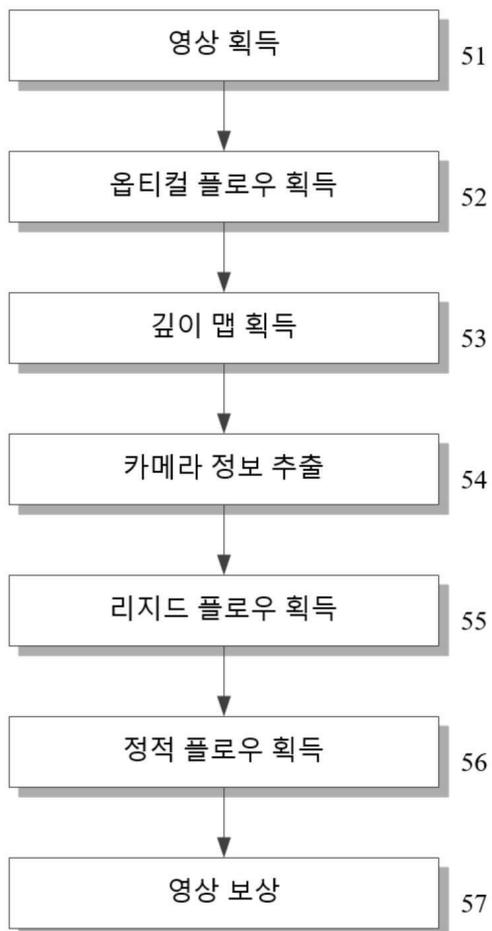
도면3



도면4



도면5



도면6

