



(19) 대한민국특허청(KR)  
(12) 공개특허공보(A)

(11) 공개번호 10-2025-0027477  
(43) 공개일자 2025년02월26일

(51) 국제특허분류(Int. Cl.)  
G06T 17/20 (2006.01) G06T 15/04 (2011.01)  
G06T 15/50 (2006.01)  
(52) CPC특허분류  
G06T 17/20 (2013.01)  
G06T 15/04 (2013.01)  
(21) 출원번호 10-2023-0196891  
(22) 출원일자 2023년12월29일  
심사청구일자 2023년12월29일  
(30) 우선권주장  
1020230108139 2023년08월18일 대한민국(KR)

(71) 출원인  
고려대학교 산학협력단  
서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)  
(72) 발명자  
김상필  
서울특별시 성북구 종암로24길 35, 204동 902호 (종암동, 종암2차 SK VIEW)  
남길주  
100 7th St. #501 Pittsburgh, PA 15222 United States  
류원정  
서울특별시 성북구 안암로9가길 53 (안암동5가) 402호  
(74) 대리인  
특허법인위솔

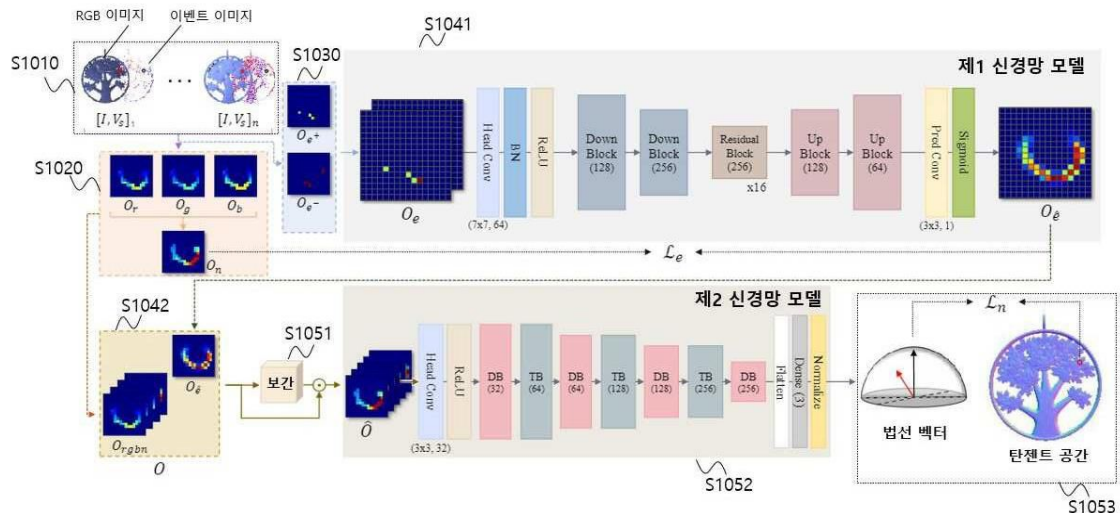
전체 청구항 수 : 총 10 항

(54) 발명의 명칭 RGB 카메라 및 이벤트 카메라를 이용한 3차원 법선 벡터 예측 방법 및 장치

(57) 요약

일 실시예에 따른 법선 벡터 예측 장치는 n 차수(n은 2 이상의 자연수)로 광원의 방향을 달리하여 객체를 촬영한 RGB 이미지 및 이벤트 이미지를 획득하는 동작; 차수 별 RGB 이미지의 제1 픽셀에서 픽셀 강도에 대한 정보를 포함하는 RGB 관찰맵을 추출하는 동작; 차수 별 이벤트 이미지의 제1 픽셀에서 이벤트의 변화에 대한 정보를 포함하는 이벤트 관찰맵을 추출하는 동작; 객체의 관찰맵과 신경망 모델의 출력맵의 차이를 손실 함수로 설정하여 학습시킨 제1 신경망 모델에 상기 이벤트 관찰맵을 입력하여 출력맵을 생성하는 동작; 및 객체의 법선 벡터와 신경망 모델의 출력 벡터의 차이를 손실 함수로 설정하여 학습시킨 제2 신경망 모델에 상기 RGB 관찰맵 및 상기 출력맵을 입력하여 상기 제1 픽셀의 법선 벡터를 특징하는 출력 벡터를 생성하는 동작을 수행할 수 있다.

대표도



(52) CPC특허분류

G06T 15/50 (2013.01)

G06T 2207/20084 (2013.01)

이 발명을 지원한 국가연구개발사업

|             |                                             |
|-------------|---------------------------------------------|
| 과제고유번호      | 1375027485                                  |
| 과제번호        | R2022020068-0001                            |
| 부처명         | 문화체육관광부                                     |
| 과제관리(전문)기관명 | 한국콘텐츠진흥원                                    |
| 연구사업명       | 차세대 실감콘텐츠 저작권핵심기술개발                         |
| 연구과제명       | 인공지능을 활용한 4D 실감 콘텐츠 창작물 생성 및 저작권 핵심 기술 연구개발 |
| 기 여 율       | 1/4                                         |
| 과제수행기관명     | 고려대학교 산학협력단                                 |
| 연구기간        | 2022.04.01 ~ 2022.12.31                     |

이 발명을 지원한 국가연구개발사업

|             |                         |
|-------------|-------------------------|
| 과제고유번호      | 1711193973              |
| 과제번호        | 2019-0-00079-005        |
| 부처명         | 과학기술정보통신부               |
| 과제관리(전문)기관명 | 정보통신기획평가원               |
| 연구사업명       | 정보통신방송혁신인재양성            |
| 연구과제명       | 인공지능대학원지원(고려대학교)        |
| 기 여 율       | 1/4                     |
| 과제수행기관명     | 고려대학교 산학협력단             |
| 연구기간        | 2023.01.01 ~ 2023.12.31 |

이 발명을 지원한 국가연구개발사업

|             |                                 |
|-------------|---------------------------------|
| 과제고유번호      | 1711140458                      |
| 과제번호        | 2021R1G1A1093855                |
| 부처명         | 과학기술정보통신부                       |
| 과제관리(전문)기관명 | 한국연구재단                          |
| 연구사업명       | 개인기초연구(과기정통부)(R&D)              |
| 연구과제명       | 동적 비전 센서를 이용한 초고속 3D 모델 생성 인공지능 |
| 기 여 율       | 1/1                             |
| 과제수행기관명     | 고려대학                            |
| 연구기간        | 2021.09.01 ~ 2022.08.31         |

이 발명을 지원한 국가연구개발사업

|             |                            |
|-------------|----------------------------|
| 과제고유번호      | 1711171488                 |
| 과제번호        | 2022R1F1A1074334           |
| 부처명         | 과학기술정보통신부                  |
| 과제관리(전문)기관명 | 한국연구재단                     |
| 연구사업명       | 개인기초연구(과기정통부)              |
| 연구과제명       | 메타버스 3차원 콘텐츠 생성 인공지능 기술 개발 |
| 기 여 율       | 1/4                        |
| 과제수행기관명     | 고려대학교                      |
| 연구기간        | 2022.06.01 ~ 2023.05.31    |

## 명세서

### 청구범위

#### 청구항 1

프로세서에 의해 동작하는 법선 벡터 예측 장치가 수행하는 동작의 방법에 있어서,

$n$  차수( $n$ 은 2 이상의 자연수)로 광원의 방향을 달리하여 객체를 촬영한 RGB 이미지 및 이벤트 이미지를 획득하는 동작;

차수 별 RGB 이미지의 제1 픽셀에서 픽셀 강도에 대한 정보를 포함하는 RGB 관찰맵을 추출하는 동작;

차수 별 이벤트 이미지의 제1 픽셀에서 이벤트의 변화에 대한 정보를 포함하는 이벤트 관찰맵을 추출하는 동작;

객체의 관찰맵과 신경망 모델의 출력맵의 차이를 손실 함수로 설정하여 학습시킨 제1 신경망 모델에 상기 이벤트 관찰맵을 입력하여 출력맵을 생성하는 동작; 및

객체의 법선 벡터와 신경망 모델의 출력 벡터의 차이를 손실 함수로 설정하여 학습시킨 제2 신경망 모델에 상기 RGB 관찰맵 및 상기 출력맵을 입력하여 상기 제1 픽셀의 법선 벡터를 특징하는 출력 벡터를 생성하는 동작을 포함하는,

방법.

#### 청구항 2

제1항에 있어서,

상기 RGB 관찰맵을 추출하는 동작은

차수 별 RGB 이미지의 제1 픽셀에서 R채널 픽셀 강도를 추출한 R채널 관찰맵, G채널 픽셀 강도를 추출한 G채널 관찰맵, B채널 픽셀 강도를 추출한 B채널 관찰맵을 추출하고, 상기 R채널 관찰맵, G채널 관찰맵 및 B채널 관찰맵을 정규화한 정규화 관찰맵을 추출하는 동작을 포함하는

방법.

#### 청구항 3

제2항에 있어서,

상기 이벤트 관찰맵을 추출하는 동작은

차수 별 이벤트 이미지의 제1 픽셀에서 양의 이벤트에 해당하는 차수를 특징한 양이벤트 관찰맵, 음의 이벤트에 해당하는 차수를 특징한 음이벤트 관찰맵을 추출하는 동작을 포함하는,

방법.

#### 청구항 4

제3항에 있어서,

상기 양이벤트 관찰맵 및 상기 음이벤트 관찰맵을 추출하는 동작은

상기 이벤트 이미지를 복셀 그리드로 변환하는 동작; 및

상기 이벤트 이미지 각각의 제1 픽셀을 기준으로, 특정 차수의 픽셀 강도가 이전 차수의 픽셀 강도보다 기 설정된 임계값 이상으로 높아진 경우의 차수를 특징한 양이벤트 관찰맵을 추출하고, 특정 차수의 픽셀 강도가 이전 차수의 픽셀 강도보다 기 설정된 임계값 이상으로 낮아진 경우의 차수를 특징한 음이벤트 관찰맵을 추출하는 동작을 포함하는,

방법.

## 청구항 5

제3항에 있어서,

상기 출력맵을 생성하는 동작은

객체의 정규화 관찰맵과 신경망 모델의 출력맵의 차이를 손실 함수로 설정하여 학습시킨 제1 신경망 모델에 상기 양이벤트 관찰맵 및 상기 음이벤트 관찰맵을 입력하여 출력맵을 생성하는 동작을 포함하는

방법.

## 청구항 6

제5항에 있어서,

상기 제1 신경망 모델은

객체의 양이벤트 관찰맵과 음이벤트 관찰맵이 입력되도록 설정된 입력 레이어 및 객체의 정규화 관찰맵이 입력되도록 설정된 출력 레이어를 포함하고, 지도 학습을 기초로 양이벤트 관찰맵 및 음이벤트 관찰맵과 정규화 관찰맵 간의 상관 관계를 갖는 모델 파라미터를 포함하는,

방법.

## 청구항 7

제5항에 있어서,

상기 제2 신경망 모델은

객체의 R채널 관찰맵, G채널 관찰맵, B채널 관찰맵, 정규화 관찰맵 및 출력맵이 입력되도록 설정된 입력 레이어 및 객체의 법선 벡터가 입력되도록 설정된 출력 레이어를 포함하고, 지도 학습을 기초로 R채널 관찰맵, G채널 관찰맵, B채널 관찰맵, 정규화 관찰맵 및 출력맵과 법선 벡터 간의 상관 관계를 갖는 모델 파라미터를 포함하는,

방법.

## 청구항 8

제5항에 있어서,

상기 출력 벡터를 생성하는 동작은

상기 R채널 관찰맵, 상기 G채널 관찰맵, 상기 B채널 관찰맵, 상기 정규화 관찰맵, 상기 출력맵 각각의 픽셀값을 컨볼루션하여 보간된 R채널 관찰맵, 보간된 G채널 관찰맵, 보간된 B채널 관찰맵, 보간된 정규화 관찰맵, 보간된 출력맵을 생성하는 동작; 및

상기 제2 신경망 모델에 보간된 R채널 관찰맵, 보간된 G채널 관찰맵, 보간된 B채널 관찰맵, 보간된 정규화 관찰맵, 보간된 출력맵을 입력하여 상기 제1 픽셀의 법선 벡터를 예측하는 출력 벡터를 생성하는 동작을 포함하는,

방법.

## 청구항 9

제1항에 있어서,

상기 출력 벡터를 생성하는 동작은

상기 RGB 이미지 및 상기 이벤트 이미지의 모든 픽셀에 대한 출력 벡터를 생성하는 동작; 및

상기 모든 픽셀에 대한 출력 벡터를 탄젠트 공간(tangent space)에 표현하는 동작을 포함하는,

방법.

## 청구항 10

명령어를 포함하는 메모리; 및

상기 명령어를 기초로 소정의 동작을 수행하는 프로세서를 포함하고,

상기 프로세서의 동작은,

$n$  차수( $n$ 은 2 이상의 자연수)로 광원의 방향을 달리하여 객체를 촬영한 RGB 이미지 및 이벤트 이미지를 획득하는 동작;

차수 별 RGB 이미지의 제1 픽셀에서 픽셀 강도에 대한 정보를 포함하는 RGB 관찰맵을 추출하는 동작;

차수 별 이벤트 이미지의 제1 픽셀에서 이벤트의 변화에 대한 정보를 포함하는 이벤트 관찰맵을 추출하는 동작;

객체의 관찰맵과 신경망 모델의 출력맵의 차이를 손실 함수로 설정하여 학습시킨 제1 신경망 모델에 상기 이벤트 관찰맵을 입력하여 출력맵을 생성하는 동작; 및

객체의 법선 벡터와 신경망 모델의 출력 벡터의 차이를 손실 함수로 설정하여 학습시킨 제2 신경망 모델에 상기 RGB 관찰맵 및 상기 출력맵을 입력하여 상기 제1 픽셀의 법선 벡터를 특정하는 출력 벡터를 생성하는 동작을 포함하는,

법선 벡터 예측 장치.

## 발명의 설명

### 기술 분야

[0001] 본 발명은 RGB 카메라 및 이벤트 카메라를 활용하여 객체를 촬영한 데이터를 함께 이용하여 객체 표면의 3차원 법선 벡터를 보다 정교하게 예측하는 기술에 관한 것이다.

### 배경 기술

[0002] 포토메트릭 스테레오는 최소 3개의 광원과 각 광원의 프레임을 사용하여 객체의 로컬 형상의 표면 법선을 추정하는 기술이다. 기존의 포토메트릭 스테레오 기법 중 확산된 양방향 반사 분포 함수를 사용하는 기법은 객체가 Lambertian 반사 표면이어야 한다는 제한 사항이 존재하였다. 다만, 실제 환경은 이러한 제약을 넘어서기 때문에 확산된 양방향 반사 분포 함수를 사용하는 기법은 실제 환경에 적용하기 어렵다는 단점이 존재한다.

[0003] 이에, 최근의 포토메트릭 스테레오 기법은 딥러닝을 이용하여 향상된 성능을 얻어내고 있는데, 딥러닝을 활용한 기법은 이미지와 빛의 방향을 하나의 쌍으로 만든 후, 공유 가중치 네트워크를 이용하여 모든 쌍들에 대한 특징맵을 추출하고 특징맵들을 융합하여 객체의 법선 벡터를 예측하는 방식을 사용한다. 하지만, 이러한 기법은 일반 카메라로 촬영된 이미지를 사용하여 법선 벡터를 예측하고 있는데, 이 경우 낮은 다이내믹 레인지 때문에 빛이 어둡거나 밝은 환경에서 촬영된 데이터에 대해서는 픽셀 강도가 제대로 표현되지 않는 문제가 있다.

[0004] 즉, 기존의 딥러닝을 활용한 포토메트릭 스테레오 기법은 단일 카메라로 획득된 데이터만을 사용하고 있어, 광대한 범위의 빛의 영향에 의해 발생하는 노이즈에 취약할 수 있으며, 이러한 문제는 법선 벡터 예측의 성능 저하와 직접적인 연관이 있으므로, 높은 다이내믹 레인지를 표현하면서 글로벌한 빛의 영향에 강인한 기술이 요구되고 있다.

## 선행기술문헌

### 특허문헌

[0005] (특허문헌 0001) 대한민국 공개특허공보 제10-2023-0004423호

## 발명의 내용

### 해결하려는 과제

[0006] 본 발명이 해결하고자 하는 과제는 RGB 카메라와 더불어, 광대한 범위의 빛의 영향에 의해 발생하는 노이즈에 강인한 이벤트 카메라의 데이터를 함께 사용하여, 넓은 다이내믹 레인지에서 객체의 3차원 법선 벡터를 예측할

수 있는 기술을 제공하는 것이다.

[0007] 한편, 본 발명의 기술적 과제들은 이상에서 언급한 기술적 과제들로 제한되지 않으며, 언급되지 않은 또 다른 기술적 과제들은 아래의 기재로부터 통상의 기술자에게 명확하게 이해될 수 있을 것이다.

### 과제의 해결 수단

[0008] 일 실시예에 따른 프로세서에 의해 동작하는 법선 벡터 예측 장치가 수행하는 동작의 방법에 있어서,  $n$  차수( $n$ 은 2 이상의 자연수)로 광원의 방향을 달리하여 객체를 촬영한 RGB 이미지 및 이벤트 이미지를 획득하는 동작; 차수 별 RGB 이미지의 제1 픽셀에서 픽셀 강도에 대한 정보를 포함하는 RGB 관찰맵을 추출하는 동작; 차수 별 이벤트 이미지의 제1 픽셀에서 이벤트의 변화에 대한 정보를 포함하는 이벤트 관찰맵을 추출하는 동작; 객체의 관찰맵과 신경망 모델의 출력맵의 차이를 손실 함수로 설정하여 학습시킨 제1 신경망 모델에 상기 이벤트 관찰맵을 입력하여 출력맵을 생성하는 동작; 및 객체의 법선 벡터와 신경망 모델의 출력 벡터의 차이를 손실 함수로 설정하여 학습시킨 제2 신경망 모델에 상기 RGB 관찰맵 및 상기 출력맵을 입력하여 상기 제1 픽셀의 법선 벡터를 특징하는 출력 벡터를 생성하는 동작을 포함할 수 있다.

[0009] 또한, 상기 RGB 관찰맵을 추출하는 동작은 차수 별 RGB 이미지의 제1 픽셀에서 R채널 픽셀 강도를 추출한 R채널 관찰맵, G채널 픽셀 강도를 추출한 G채널 관찰맵, B채널 픽셀 강도를 추출한 B채널 관찰맵을 추출하고, 상기 R채널 관찰맵, G채널 관찰맵 및 B채널 관찰맵을 정규화한 정규화 관찰맵을 추출하는 동작을 포함할 수 있다.

[0010] 또한, 상기 이벤트 관찰맵을 추출하는 동작은 차수 별 이벤트 이미지의 제1 픽셀에서 양의 이벤트에 해당하는 차수를 특징한 양이벤트 관찰맵, 음의 이벤트에 해당하는 차수를 특징한 음이벤트 관찰맵을 추출하는 동작을 포함할 수 있다.

[0011] 또한, 상기 양이벤트 관찰맵 및 상기 음이벤트 관찰맵을 추출하는 동작은 상기 이벤트 이미지를 복셀 그리드로 변환하는 동작; 및 상기 이벤트 이미지 각각의 제1 픽셀을 기준으로, 특정 차수의 픽셀 강도가 이전 차수의 픽셀 강도보다 기 설정된 임계값 이상으로 높아진 경우의 차수를 특징한 양이벤트 관찰맵을 추출하고, 특정 차수의 픽셀 강도가 이전 차수의 픽셀 강도보다 기 설정된 임계값 이상으로 낮아진 경우의 차수를 특징한 음이벤트 관찰맵을 추출하는 동작을 포함할 수 있다.

[0012] 또한, 상기 출력맵을 생성하는 동작은 객체의 정규화 관찰맵과 신경망 모델의 출력맵의 차이를 손실 함수로 설정하여 학습시킨 제1 신경망 모델에 상기 양이벤트 관찰맵 및 상기 음이벤트 관찰맵을 입력하여 출력맵을 생성하는 동작을 포함할 수 있다.

[0013] 또한, 상기 제1 신경망 모델은 객체의 양이벤트 관찰맵과 음이벤트 관찰맵이 입력되도록 설정된 입력 레이어 및 객체의 정규화 관찰맵이 입력되도록 설정된 출력 레이어를 포함하고, 지도 학습을 기초로 양이벤트 관찰맵 및 음이벤트 관찰맵과 정규화 관찰맵 간의 상관 관계를 갖는 모델 파라미터를 포함할 수 있다.

[0014] 또한, 상기 제2 신경망 모델은 객체의 R채널 관찰맵, G채널 관찰맵, B채널 관찰맵, 정규화 관찰맵 및 출력맵이 입력되도록 설정된 입력 레이어 및 객체의 법선 벡터가 입력되도록 설정된 출력 레이어를 포함하고, 지도 학습을 기초로 R채널 관찰맵, G채널 관찰맵, B채널 관찰맵, 정규화 관찰맵 및 출력맵과 법선 벡터 간의 상관 관계를 갖는 모델 파라미터를 포함할 수 있다.

[0015] 또한, 상기 출력 벡터를 생성하는 동작은 상기 R채널 관찰맵, 상기 G채널 관찰맵, 상기 B채널 관찰맵, 상기 정규화 관찰맵, 상기 출력맵 각각의 픽셀값을 컨볼루션하여 보간된 R채널 관찰맵, 보간된 G채널 관찰맵, 보간된 B채널 관찰맵, 보간된 정규화 관찰맵, 보간된 출력맵을 생성하는 동작; 및 상기 제2 신경망 모델에 보간된 R채널 관찰맵, 보간된 G채널 관찰맵, 보간된 B채널 관찰맵, 보간된 정규화 관찰맵, 보간된 출력맵을 입력하여 상기 제1 픽셀의 법선 벡터를 예측하는 출력 벡터를 생성하는 동작을 포함할 수 있다.

[0016] 또한, 상기 출력 벡터를 생성하는 동작은 상기 RGB 이미지 및 상기 이벤트 이미지의 모든 픽셀에 대한 출력 벡터를 생성하는 동작; 및 상기 모든 픽셀에 대한 출력 벡터를 탄젠트 공간(tangent space)에 표현하는 동작을 포함할 수 있다.

[0017] 일 실시예에 따른 법선 벡터 예측 장치는 명령어를 포함하는 메모리; 및 상기 명령어를 기초로 소정의 동작을 수행하는 프로세서를 포함하고, 상기 프로세서의 동작은  $n$  차수( $n$ 은 2 이상의 자연수)로 광원의 방향을 달리하여 객체를 촬영한 RGB 이미지 및 이벤트 이미지를 획득하는 동작; 차수 별 RGB 이미지의 제1 픽셀에서 픽셀 강도에 대한 정보를 포함하는 RGB 관찰맵을 추출하는 동작; 차수 별 이벤트 이미지의 제1 픽셀에서 이벤트의 변화

에 대한 정보를 포함하는 이벤트 관찰맵을 추출하는 동작; 객체의 관찰맵과 신경망 모델의 출력맵의 차이를 손실 함수로 설정하여 학습시킨 제1 신경망 모델에 상기 이벤트 관찰맵을 입력하여 출력맵을 생성하는 동작; 및 객체의 법선 벡터와 신경망 모델의 출력 벡터의 차이를 손실 함수로 설정하여 학습시킨 제2 신경망 모델에 상기 RGB 관찰맵 및 상기 출력맵을 입력하여 상기 제1 픽셀의 법선 벡터를 특정하는 출력 벡터를 생성하는 동작을 포함할 수 있다.

### 발명의 효과

- [0018] 본 발명은 광원의 방향을 달리하여 객체를 촬영한 RGB 이미지와 이벤트 이미지를 함께 이용하면서, RGB 이미지로부터 생성한 RGB 관찰맵을 통해 학습시킨 제1 신경망 모델을 이용하여 스파스(sparse)한 특성을 가진 이벤트 관찰맵이 조밀한 특성을 갖도록 보장할 수 있으며, 이로부터 생성한 RGB 관찰맵과 조밀한 이벤트 관찰맵을 함께 이용하여 객체의 법선 벡터를 예측할 수 있다.
- [0019] 이에 따라, 본 발명은 RGB 카메라로부터 획득한 데이터와 더불어, 광대한 범위의 빛의 영향에 의해 발생하는 노이즈 등에 강인한 이벤트 카메라의 데이터를 함께 사용함에 따라, 보다 넓은 다이내믹 레인지에서 객체의 3차원 법선 벡터를 정확하게 예측할 수 있다.
- [0020] 한편, 본 발명에 의한 효과는 이상에서 언급한 것들로 제한되지 않으며, 언급되지 않은 또 다른 기술적 효과들은 아래의 기재로부터 통상의 기술자에게 명확하게 이해될 수 있을 것이다.

### 도면의 간단한 설명

- [0021] 도 1은 일 실시예에 따른 법선 벡터 예측 장치의 구성도이다.
- 도 2는 일 실시예에 따른 법선 벡터 예측 장치가 수행하는 동작의 단계를 나타낸 흐름도이다.
- 도 3는 일 실시예에 따른 법선 벡터 예측 장치가 수행하는 동작을 기능 블록과 함께 나타낸 흐름도이다.
- 도 4는 일 실시예에 따른 제1 신경망 모델에서 은닉 레이어의 구성에 대한 예시도이다.
- 도 5는 일 실시예에 따라 각 맵이 갖는 픽셀 강도의 값이 서로 유사한 범위 내에 위치하도록 각 맵의 픽셀 강도 값을 보간하는 동작의 예시도이다.
- 도 6은 일 실시예에 따른 제2 신경망 모델에서 은닉 레이어의 구성에 대한 예시도이다.
- 도 7은 특정 객체에 대한 법선 벡터를 예측하는 기존 알고리즘의 성능과 본 발명의 실시예를 적용한 경우를 비교한 성능 평가 결과이다.
- 도 8은 기존 알고리즘을 이용하여 특정 객체에 대한 법선 벡터를 예측한 후 탄젠트 공간에 표현한 결과와, 본 발명의 실시예를 이용하여 특정 객체에 대한 법선 벡터를 예측한 후 탄젠트 공간에 표현한 결과이다.

### 발명을 실시하기 위한 구체적인 내용

- [0022] 본 발명의 목적과 기술적 구성 및 그에 따른 작용 효과에 관한 자세한 사항은 본 발명의 명세서에 첨부된 도면에 의거한 이하의 상세한 설명에 의해 보다 명확하게 이해될 것이다. 첨부된 도면을 참조하여 본 발명에 따른 실시예를 상세하게 설명한다.
- [0023] 본 명세서에서 개시되는 실시예들은 본 발명의 범위를 한정하는 것으로 해석되거나 이용되지 않아야 할 것이다. 이 분야의 통상의 기술자에게 본 명세서의 실시예를 포함한 설명은 다양한 응용을 갖는다는 것이 당연하다. 따라서, 본 발명의 상세한 설명에 기재된 임의의 실시예들은 본 발명을 보다 잘 설명하기 위한 예시적인 것이며 본 발명의 범위가 실시예들로 한정되는 것을 의도하지 않는다.
- [0024] 도면에 표시되고 아래에 설명되는 기능 블록들은 가능한 구현의 예들일 뿐이다. 다른 구현들에서는 상세한 설명의 사상 및 범위를 벗어나지 않는 범위에서 다른 기능 블록들이 사용될 수 있다. 또한, 본 발명의 하나 이상의 기능 블록이 개별 블록들로 표시되지만, 본 발명의 기능 블록들 중 하나 이상은 동일 기능을 실행하는 다양한 하드웨어 및 소프트웨어 구성들의 조합일 수 있다.
- [0025] 또한, 어떤 구성요소들을 포함한다는 표현은 “개방형”의 표현으로서 해당 구성요소들이 존재하는 것을 단순히 지칭할 뿐이며, 추가적인 구성요소들을 배제하는 것으로 이해되어서는 안 된다.
- [0026] 나아가 어떤 구성요소가 다른 구성요소에 “연결되어” 있다거나 “접속되어” 있다고 언급될 때에는, 그 다른

구성요소에 직접적으로 연결 또는 접속되어 있을 수도 있지만, 중간에 다른 구성요소가 존재할 수도 있다고 이해되어야 한다.

- [0027] 이하, 본 발명의 다양한 실시예가 첨부된 도면을 참조하여 기재된다. 그러나, 이는 본 발명을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 발명의 실시예의 다양한 변경(modification), 균등물(equivalent), 및/또는 대체물(alternative)을 포함하는 것으로 이해되어야 한다.
- [0028] 도 1은 일 실시예에 따른 법선 벡터 예측 장치(100)(이하, '장치(100)'로 지칭)의 구성도이다.
- [0029] 도 1을 참조하면, 일 실시예에 따른 장치(100)는 각각 메모리(110), 프로세서(120), 입출력 인터페이스(130) 및 통신 인터페이스(140)를 포함할 수 있다.
- [0030] 메모리(110)는 외부 장치로부터 획득한 데이터 또는 스스로 생성한 데이터를 저장할 수 있다. 메모리(110)는 프로세서(120)의 동작을 수행시킬 수 있는 명령어들을 저장할 수 있다. 또한, 메모리(110)는 후술할 객체 정보, 객체의 RGB 이미지, 객체의 이벤트 이미지, 제1 신경망 모델 및 제2 신경망 모델에 대한 정보를 저장할 수 있다.
- [0031] 프로세서(120)는 전반적인 동작을 제어하는 연산 장치이다. 프로세서(120)는 메모리(110)에 저장된 명령어들을 실행할 수 있다. 본 문서의 실시예에 따라 후술할 장치(100)의 동작은 프로세서(120)에 의해 수행되는 동작으로 이해될 수 있다.
- [0032] 입출력 인터페이스(130)는 정보를 입력하거나 출력하는 하드웨어 인터페이스 또는 소프트웨어 인터페이스를 포함할 수 있다.
- [0033] 통신 인터페이스(140)는 통신망을 통해 정보를 송수신 할 수 있게 한다. 이를 위해, 통신 인터페이스(140)는 무선 통신모듈 또는 유선 통신모듈을 포함할 수 있다.
- [0034] 장치(100)는 프로세서(120)를 통해 연산을 수행하고 네트워크를 통해 정보를 송수신할 수 있는 다양한 형태의 장치로 구현될 수 있다. 예를 들면, 서버, 컴퓨터 장치, 휴대용 통신 장치, 스마트 폰, 휴대용 멀티미디어 장치, 노트북, 태블릿 PC 등의 형태로 구현될 수 있으나, 이러한 예시에 한정되는 것은 아니다.
- [0035] 본 문서의 장치(100)는 특정 객체를 촬영한 각도에서 해당 객체의 표면에 대한 법선 벡터를 구하는 기술을 제공한다. 이를 위해, 장치(100)는 다음 도 2 및 도 3의 단계를 수행할 수 있다.
- [0036] 도 2는 일 실시예에 따른 법선 벡터 예측 장치가 수행하는 동작의 단계를 나타낸 흐름도이고, 도 3는 일 실시예에 따른 법선 벡터 예측 장치가 수행하는 동작을 기능 블록과 함께 나타낸 흐름도이다.
- [0037] 도 2 및 도 3의 실시예에 따른 장치(100)의 동작은 프로세서(120)에 의해 수행되는 동작으로 이해될 수 있다.
- [0038] 도 2 및 도 3에 개시된 각 단계는 본 발명의 목적을 달성함에 있어서 바람직한 실시예일 뿐이며, 필요에 따라 일부 단계가 추가 또는 삭제될 수 있음은 물론이고, 어느 한 단계가 다른 단계에 포함되어 수행될 수도 있다. 도 2 및 도 3에 개시된 각 동작의 순서는 이해의 편의를 위해 배치된 순서일 뿐, 이러한 순서가 시계열적인 순서로 한정되는 것이 아니며, 설계자의 선택에 따라 순서가 다르게 변경되어 동작될 수 있다.
- [0039] S1010 단계에서, 장치(100)는 RGB 카메라로 객체를 촬영한 RGB 이미지 및 이벤트 카메라로 객체를 촬영한 이벤트 이미지를 획득할 수 있다. 장치(100)는 RGB 카메라, 이벤트 카메라 및 객체의 위치를 유지한 채,  $n$  차수( $n$ 은 2 이상의 자연수) 만큼 광원의 방향을 달리하여 객체를 촬영함으로써,  $n$ 개의 RGB 이미지( $I_1 \sim I_n$ )와  $n$ 개의 이벤트 이미지( $V_{S1} \sim V_{Sn}$ )를 획득할 수 있다.
- [0040] S1020 단계에서, 장치(100)는 차수 별 RGB 이미지의 특정 픽셀(이하, '제1 픽셀'로 지칭)에서 픽셀 강도에 대한 정보를 포함하는 RGB 관찰맵을 추출할 수 있다. 예를 들어, 장치(100)는 1 차수 내지  $n$  차수까지 촬영된 RGB 이미지의 제1 픽셀에서 RED 채널의 픽셀 강도를 추출하고, 1차수부터  $n$ 차수까지 순서대로 각 픽셀 강도를 복셀 그리드의 형태로 표현한 R채널 관찰맵( $O_r$ )을 생성할 수 있다. 예를 들어, 장치(100)는 1 차수 내지  $n$  차수 까지 촬영된 RGB 이미지의 제1 픽셀에서 GREEN 채널의 픽셀 강도를 추출하고, 1차수부터  $n$ 차수까지 순서대로 각 픽셀 강도를 복셀 그리드의 형태로 표현한 G채널 관찰맵( $O_g$ )을 생성할 수 있다. 예를 들어, 장치(100)는 1 차수 내지  $n$  차수까지 촬영된 RGB 이미지의 제1 픽셀에서 BLUE 채널의 픽셀 강도를 추출하고, 1차수부터  $n$ 차수까지 순서대로 각 픽셀 강도를 복셀 그리드의 형태로 표현한 B채널 관찰맵( $O_b$ )을 생성할 수 있다.

- [0041] 또한, 장치(100)는 R채널 관찰맵, G채널 관찰맵 및 B채널 관찰맵을 정규화한 정규화 관찰맵( $O_n$ )을 생성할 수 있다. 예를 들어, 장치(100)는 동일 차수에 해당하는 R채널 관찰맵의 픽셀 강도, G채널 관찰맵의 픽셀 강도, B채널 관찰맵의 픽셀 강도를 평균한 값을 계산할 수 있고, 이 과정을 1 차수부터 n차수까지 모든 차수에 대해 수행하여 정규화 관찰맵을 생성할 수 있다.
- [0042] S1030 단계에서, 장치(100)는 차수 별 이벤트 이미지의 제1 픽셀에서 이벤트의 변화에 대한 정보를 포함하는 이벤트 관찰맵을 추출할 수 있다. 여기서, 이벤트 카메라는 빛의 변화에 따른 이벤트를 기록하는 카메라이다. 즉, 1 차수 내지 n 차수의 이벤트 이미지가 있는 경우, 특정 차수의 이벤트 이미지를 바로 이전 차수의 이벤트 이미지와 비교하였을 때, 빛의 세기(ex. 픽셀 강도)가 이전 차수보다 기 설정된 임계값 이상으로 강해진 픽셀에는 양의 이벤트가 매핑되고, 빛의 세기가 이전 차수 보다 기 설정된 임계값 이상으로 약해진 픽셀에는 음의 이벤트가 매핑되며, 빛의 세기가 기 설정된 임계값 이상으로 변하지 않은 픽셀에는 아무 정보가 매핑되지 않는다. 이에 따라, 장치(100)는 1 차수 내지 n차수 까지, 차수 별 이벤트 이미지의 제1 픽셀에서 양의 이벤트에 해당하는 차수를 특정한 양이벤트 관찰맵( $O_{e+}$ )과, 음의 이벤트에 해당하는 차수를 특정한 음이벤트 관찰맵( $O_{e-}$ )을 추출할 수 있다. 이때 1 차수의 이벤트 이미지는 비교가 될 대상이 없으므로, 1차수 이벤트 이미지는 아무런 정보가 매핑되지 않을 수 있으며, 1차수 이벤트 이미지는 다음 2 차수 이벤트 이미지의 픽셀 강도와 비교될 수 있다.
- [0043] 이때 이벤트 관찰맵은 이벤트 변화가 발생한 픽셀에만 정보가 기록되므로, 이벤트 관찰맵이 갖는 정보는 스파스(sparse)한 특성을 갖는다.
- [0044] S1040 단계에서, 장치(100)는 스파스한 정보를 갖는 이벤트 관찰맵으로부터 조밀한 정보를 갖는 이벤트 관찰맵이 출력되도록 학습시킨 제1 신경망 모델을 이용하여, S1030 단계에서 추출한 이벤트 관찰맵으로부터 조밀한 이벤트 출력맵을 생성할 수 있다.
- [0045] 일 예로, 제1 신경망 모델은 도 3에 도시된 입력 레이어, 은닉 레이어 및 출력 레이어를 포함할 수 있다. 제1 신경망 모델은 학습 단계에서, 입력 데이터로서 객체의 양이벤트 관찰맵과 음이벤트 관찰맵이 입력 레이어에 입력되도록 설정될 수 있고, 출력 데이터로서 객체의 정규화 관찰맵의 조밀성에 가까운 값을 갖는 출력맵이 출력되도록 설정할 수 있다. 이를 위해, 제1 신경망 모델은 지도 학습을 기초로 입력 레이어에 입력되는 양이벤트 관찰맵 및 음이벤트 관찰맵과, 출력 레이어에 입력되는 정규화 관찰맵 간의 상관 관계를 갖도록 은닉 레이어의 모델 파라미터가 학습될 수 있다. 또한, 제1 신경망 모델의 손실 함수는 학습에 이용되는 객체의 정규화 관찰맵과, 신경망 모델이 예측으로 출력하는 출력맵의 차이로 설정될 수 있다.
- [0046] 도 4는 일 실시예에 따른 제1 신경망 모델에서 은닉 레이어의 구성에 대한 예시도이다.
- [0047] 도 3 및 도 4를 함께 참조하면, 제1 신경망 모델의 은닉 레이어는 도 3과 같이, Head 부분, 2개의 Down 블록, 16개의 Residual 블록, 2개의 Up 블록 및 pred 부분으로 구성될 수 있다. Head 부분은 헤드 컨볼루션 레이어, 배치 정규화 레이어, ReLU 활성화 함수로 구성되며, 이 부분은 스파스한 이벤트 관찰맵을 여러 채널로 증폭할 수 있다. Down 블록은 도 4와 같이, 5x5 커널 컨볼루션 계층, 배치 정규화 계층 및 Leaky ReLU 활성화 함수로 구성되며, 이벤트 관찰맵의 크기를 절반으로 줄이면서 인코딩할 수 있고, 마지막 Down 블록에서 인코딩된 특징맵을 얻은 후 16개의 잔여 블록은 최소 강도 특징을 보간하기 위해 출력 특징맵을 처리할 수 있다. Residual 블록은 3x3 커널 컨볼루션 계층, 배치 정규화 계층 및 ReLU 활성화 함수로 구성되며, 최소 입력 기능맵에서 더 유창한 출력 기능맵을 도출할 수 있다. Up 블록은 도 4와 같이, 크기 조정된 입력을 디코딩하기 위해 입력 해상도와 일치하도록 잔여 블록 Mout의 마지막 레이어 출력을 업스케일링 진행할 수 있다. Pred 부분은 Pred convolution 레이어로 구성되며, 스파스한 이벤트 관찰맵을 조밀한 이벤트 관찰맵으로 변환하고 시그모이드 활성화 함수를 사용하여 0과 1 사이의 값을 조정하여 입력 값의 범위를 일치시킬 수 있다.
- [0048] 이에 따라, S1041 단계에서, 장치(100)는 객체의 정규화 관찰맵과 신경망 모델의 출력맵의 차이를 손실 함수로 설정하여 학습시킨 제1 신경망 모델에 양이벤트 관찰맵 및 음이벤트 관찰맵을 입력하여 출력맵을 생성할 수 있고, S1042 단계에서, 장치(100)는 객체의 R채널 관찰맵( $O_r$ ), G채널 관찰맵( $O_g$ ), B채널 관찰맵( $O_b$ ), 정규화 관찰맵( $O_n$ ) 및 제1 신경망 모델의 출력맵( $O^d$ )을 획득할 수 있다.
- [0049] S1050 단계에서, 장치(100)는 RGB 관찰맵 및 제1 신경망 모델의 출력맵으로부터 객체에서 제1 픽셀에 해당하는 위치의 법선 벡터를 출력하도록 학습시킨 제2 신경망 모델을 이용하여, 제1 픽셀에 해당하는 위치의 법선 벡터를 특징하는 출력 벡터를 생성할 수 있다.

- [0050] 일 예로, S1051 단계에서, 장치(100)는 제2 신경망 모델에 값을 입력하기 전에, 객체의 R채널 관찰맵, G채널 관찰맵, B채널 관찰맵, 정규화 관찰맵 및 제1 신경망 모델의 출력맵이 갖는 픽셀 강도의 값이 서로 유사한 범위 내에 위치하도록 각 맵의 픽셀 강도 값을 보간할 수 있다.
- [0051] 도 5는 일 실시예에 따라 각 맵이 갖는 픽셀 강도의 값이 서로 유사한 범위 내에 위치하도록 각 맵의 픽셀 강도 값을 보간하는 동작의 예시도이다.
- [0052] 도 5를 참조하면, S1051 단계에서, 장치(100)는 R채널 관찰맵, G채널 관찰맵, B채널 관찰맵, 정규화 관찰맵, 출력맵 각각의 픽셀값을 컨볼루션하고, 시그모이드 활성화 함수로 각 픽셀 강도의 값을 0에서 1 사이의 범위로 설정한 후, 각 채널에 대해 요소별 곱셈을 수행할 수 있다. 이에 따라, 장치(100)는 보간된 R채널 관찰맵, 보간된 G채널 관찰맵, 보간된 B채널 관찰맵, 보간된 정규화 관찰맵, 보간된 출력맵을 생성할 수 있다.
- [0053] 일 예로, S1052 단계에서, 장치(100)는 보간된 R채널 관찰맵, 보간된 G채널 관찰맵, 보간된 B채널 관찰맵, 보간된 정규화 관찰맵, 보간된 출력맵을 제2 신경망 모델에 입력하여 제1 픽셀의 법선 벡터를 예측하는 출력 벡터를 생성할 수 있다.
- [0054] 일 예로, 제2 신경망 모델은 도 3에 도시된 입력 레이어, 은닉 레이어 및 출력 레이어를 포함할 수 있다. 제2 신경망 모델은 학습 단계에서, 입력 데이터로서 보간된 R채널 관찰맵, 보간된 G채널 관찰맵, 보간된 B채널 관찰맵, 보간된 정규화 관찰맵, 보간된 출력맵이 입력 레이어에 입력되도록 설정될 수 있고, 출력 데이터로서 객체의 실제 법선 벡터에 가까운 값을 갖는 출력 벡터가 출력되도록 설정할 수 있다. 이를 위해, 제2 신경망 모델은 지도 학습을 기초로 입력 레이어에 입력되는 보간된 R채널 관찰맵, 보간된 G채널 관찰맵, 보간된 B채널 관찰맵, 보간된 정규화 관찰맵, 보간된 출력맵과, 출력 레이어에 입력되는 객체로부터 측정된 실제 법선 벡터값 간의 상관 관계를 갖도록 은닉 레이어의 모델 파라미터가 학습될 수 있다. 또한, 제2 신경망 모델의 손실 함수는 학습에 이용되는 객체의 실제 법선 벡터와, 신경망 모델이 예측으로 출력하는 출력 벡터의 차이로 설정될 수 있다.
- [0055] 도 6은 일 실시예에 따른 제2 신경망 모델에서 은닉 레이어의 구성에 대한 예시도이다.
- [0056] 도 3 및 도 6을 함께 참조하면, 제2 신경망 모델의 은닉 레이어는 전환 블록(TB)의 컨볼루션 레이어와 ReLU 활성화 함수 사이에 배치 정규화 레이어(BN)를 연결함으로써 객체의 법선 벡터를 예측하는 출력 벡터를 생성할 수 있다.
- [0057] 일 예로, S1053 단계에서, 장치(100)는 제1 픽셀을 기준으로 수행한 동작을 모든 픽셀에 대해 수행하여, RGB 이미지 및 이벤트 이미지의 모든 픽셀에 대한 출력 벡터를 생성할 수 있다. 장치(100)는 모든 픽셀에 대한 출력 벡터를 탄젠트 공간(tangent space)에 표현할 수 있다.
- [0058] 도 7은 특정 객체에 대한 법선 벡터를 예측하는 기존 알고리즘의 성능과 본 발명의 실시예를 적용한 경우를 비교한 성능 평가 결과이고, 도 8은 기존 알고리즘을 이용하여 특정 객체에 대한 법선 벡터를 예측한 후 탄젠트 공간에 표현한 결과와, 본 발명의 실시예를 이용하여 특정 객체에 대한 법선 벡터를 예측한 후 탄젠트 공간에 표현한 결과이다.
- [0059] 도 7 및 도 8의 법선 벡터 측정의 성능 평가 결과는 MAE (Mean Angular Error) score로 측정하였다. MAE는 객체의 각 픽셀 좌표에서 예측된 법선과 실제 법선 사이의 각도 차이에 대한 평균 오차를 나타낸다.
- [0060] 도 7 및 도 8을 참조하면, 본 발명의 실시예를 적용한 경우 기존의 알고리즘에 비해, 표면 법선 변화가 있는 경우에도 높은 충실도로 표면 법선 벡터를 가장 정확하게 예측하였으며, MAE 오류 점수가 대폭 감소하였음을 확인할 수 있다. 또한, 커널 크기, 컨볼루션 레이어의 채널 수, 잔여 블록수를 경량화한 lightweights 버전을 적용한 경우에도 기존 알고리즘의 성능을 넘어서는 것을 확인할 수 있다.
- [0061] 상술한 실시예에 따르면, 본 발명은 광원의 방향을 달리하여 객체를 촬영한 RGB 이미지와 이벤트 이미지를 함께 이용하면서, RGB 이미지로부터 생성한 RGB 관찰맵을 통해 학습시킨 제1 신경망 모델을 이용하여 스파스(sparse)한 특성을 가진 이벤트 관찰맵이 조밀한 특성을 갖도록 보강할 수 있으며, 이로부터 생성한 RGB 관찰맵과 조밀한 이벤트 관찰맵을 함께 이용하여 객체의 법선 벡터를 예측할 수 있다. 이에 따라, 본 발명은 RGB 카메라로부터 획득한 데이터와 더불어, 광대한 범위의 빛의 영향에 의해 발생하는 노이즈 등에 강인한 이벤트 카메라의 데이터를 함께 사용함에 따라, 보다 넓은 다이내믹 레인지에서 객체의 3차원 법선 벡터를 정확하게 예측할 수 있다.
- [0062] 본 문서의 다양한 실시예들 및 이에 사용된 용어들은 본 문서에 기재된 기술적 특징들을 특정한 실시예들로 한정하려는 것이 아니며, 해당 실시예의 다양한 변경, 균등물, 또는 대체물을 포함하는 것으로 이해되어야 한다.

도면의 설명과 관련하여, 유사한 또는 관련된 구성요소에 대해서는 유사한 참조 부호가 사용될 수 있다. 아이템에 대응하는 명사의 단수 형은 관련된 문맥상 명백하게 다르게 지시하지 않는 한, 아이템 한 개 또는 복수 개를 포함할 수 있다.

[0063] 본 문서에서, "A 또는 B", "A 및 B 중 적어도 하나", "A 또는 B 중 적어도 하나", "A, B 또는 C", "A, B 및 C 중 적어도 하나", "및 A, B, 또는 C 중 적어도 하나"와 같은 문구들 각각은 그 문구들 중 해당하는 문구에 함께 나열된 항목들의 모든 가능한 조합을 포함할 수 있다. "1", "제2", 또는 "첫째" 또는 "둘째"와 같은 용어들은 단순히 해당 구성요소를 다른 해당 구성요소와 구분하기 위해 사용될 수 있으며, 해당 구성요소들을 다른 측면(예: 중요성 또는 순서)에서 한정하지 않는다. 어떤(예: 제1) 구성요소가 다른(예: 제2) 구성요소에, "기능적으로" 또는 "통신적으로"라는 용어와 함께 또는 이런 용어 없이, "커플드" 또는 "커넥티드"라고 언급된 경우, 그것은 어떤 구성요소가 다른 구성요소에 직접적으로(예: 유선으로), 무선으로, 또는 제3 구성요소를 통하여 연결될 수 있다는 것을 의미한다.

[0064] 본 문서에서 사용된 용어 "모듈"은 하드웨어, 소프트웨어 또는 펌웨어로 구현된 유닛을 포함할 수 있으며, 예를 들면, 로직, 논리 블록, 부품, 또는 회로 등의 용어와 상호 호환적으로 사용될 수 있다. 모듈은, 일체로 구성된 부품 또는 하나 또는 그 이상의 기능을 수행하는, 부품의 최소 단위 또는 그 일부가 될 수 있다. 예를 들면, 일 실시예에 따르면, 모듈은 ASIC(application-specific integrated circuit)의 형태로 구현될 수 있다.

[0065] 본 문서의 다양한 실시예들은 기기(예: 전자 장치)에 의해 읽을 수 있는 저장 매체(예: 메모리)에 저장된 하나 이상의 명령어들을 포함하는 소프트웨어(예: 프로그램)로서 구현될 수 있다. 저장 매체는 RAM(random access memory), 메모리 버퍼, 하드 드라이브, 데이터베이스, EPROM(erasable programmable read-only memory), EEPROM(electrically erasable read-only memory), ROM(read-only memory) 및/또는 등등을 포함할 수 있다.

[0066] 또한, 본 문서의 실시예들의 프로세서는, 저장 매체로부터 저장된 하나 이상의 명령어들 중 적어도 하나의 명령어를 호출하고, 그것을 실행할 수 있다. 이것은 기기가 호출된 적어도 하나의 명령어에 따라 적어도 하나의 기능을 수행하도록 운영되는 것을 가능하게 한다. 이러한 하나 이상의 명령어들은 컴파일러에 의해 생성된 코드 또는 인터프리터에 의해 실행될 수 있는 코드를 포함할 수 있다. 프로세서는 범용 프로세서, FPGA(Field Programmable Gate Array), ASIC(Application Specific Integrated Circuit), DSP(Digital Signal Processor) 및/또는 등등 일 수 있다.

[0067] 기기로 읽을 수 있는 저장매체는, 비일시적(non-transitory) 저장매체의 형태로 제공될 수 있다. 여기서, '비일시적'은 저장매체가 실재(tangible)하는 장치이고, 신호(예: 전자기파)를 포함하지 않는다는 것을 의미할 뿐이며, 이 용어는 데이터가 저장매체에 반영구적으로 저장되는 경우와 임시적으로 저장되는 경우를 구분하지 않는다.

[0068] 본 문서에 개시된 다양한 실시예들에 따른 방법은 컴퓨터 프로그램 제품(computer program product)에 포함되어 제공될 수 있다. 컴퓨터 프로그램 제품은 상품으로서 판매자 및 구매자 간에 거래될 수 있다. 컴퓨터 프로그램 제품은 기기로 읽을 수 있는 저장 매체(예: compact disc read only memory (CD-ROM))의 형태로 배포되거나, 또는 어플리케이션 스토어(예: 플레이 스토어)를 통해 또는 두 개의 사용자 장치들(예: 스마트폰들) 간에 직접, 온라인으로 배포(예: 다운로드 또는 업로드)될 수 있다. 온라인 배포의 경우에, 컴퓨터 프로그램 제품의 적어도 일부는 제조사의 서버, 어플리케이션 스토어의 서버, 또는 서버의 메모리와 같은 기기로 읽을 수 있는 저장 매체에 적어도 일시 저장되거나, 임시적으로 생성될 수 있다.

[0069] 다양한 실시예들에 따르면, 기술한 구성요소들의 각각의 구성요소(예: 모듈 또는 프로그램)는 단수 또는 복수의 개체를 포함할 수 있다. 다양한 실시예들에 따르면, 기술한 해당 구성요소들 중 하나 이상의 구성요소들 또는 동작들이 생략되거나, 또는 하나 이상의 다른 구성요소들 또는 동작들이 추가될 수 있다. 대체적으로 또는 추가적으로, 복수의 구성요소들(예: 모듈 또는 프로그램)은 하나의 구성요소로 통합될 수 있다. 이런 경우, 통합된 구성요소는 복수의 구성요소들 각각의 구성요소의 하나 이상의 기능들을 통합 이전에 복수의 구성요소들 중 해당 구성요소에 의해 수행되는 것과 동일 또는 유사하게 수행할 수 있다. 다양한 실시예들에 따르면, 모듈, 프로그램 또는 다른 구성요소에 의해 수행되는 동작들은 순차적으로, 병렬적으로, 반복적으로, 또는 휴리스틱하게 실행되거나, 동작들 중 하나 이상이 다른 순서로 실행되거나, 생략되거나, 또는 하나 이상의 다른 동작들이 추가될 수 있다.

## 부호의 설명

[0070]

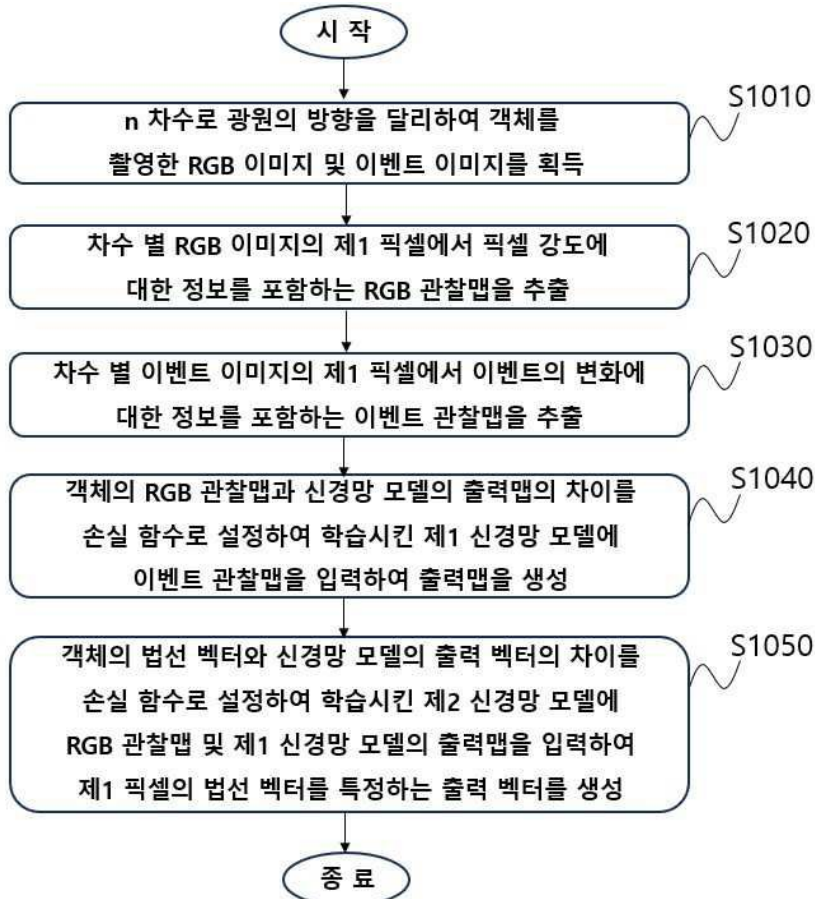
- 100: 장치
- 110: 메모리
- 120: 프로세서
- 130: 입출력 인터페이스
- 140: 통신 인터페이스

도면

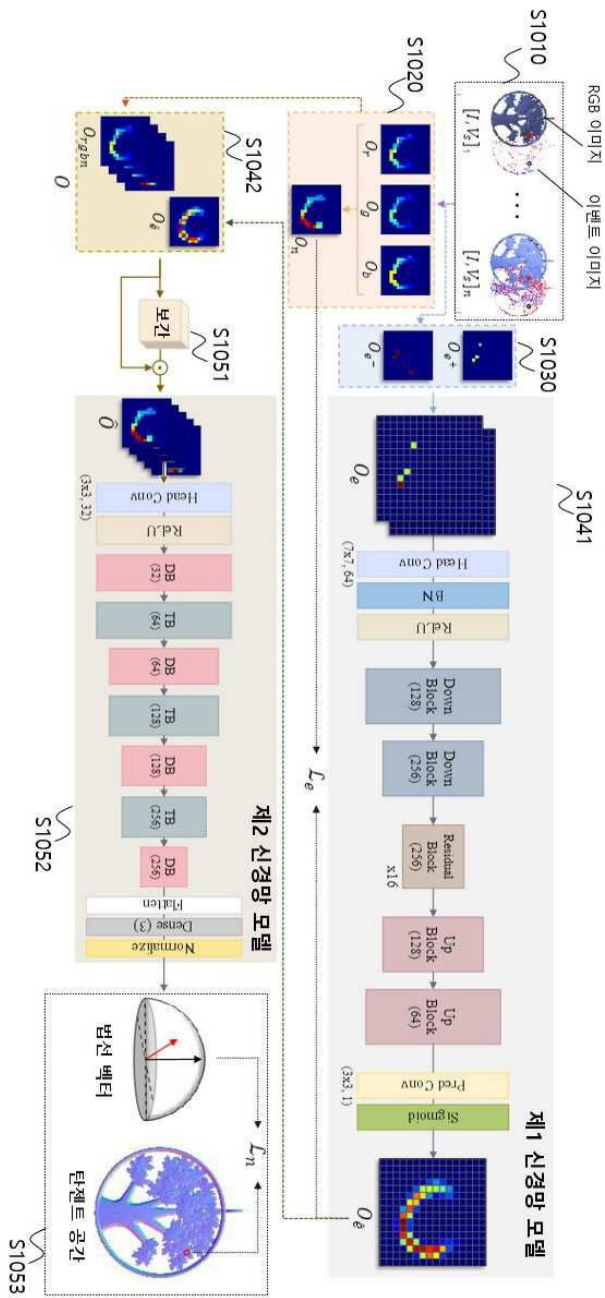
도면1



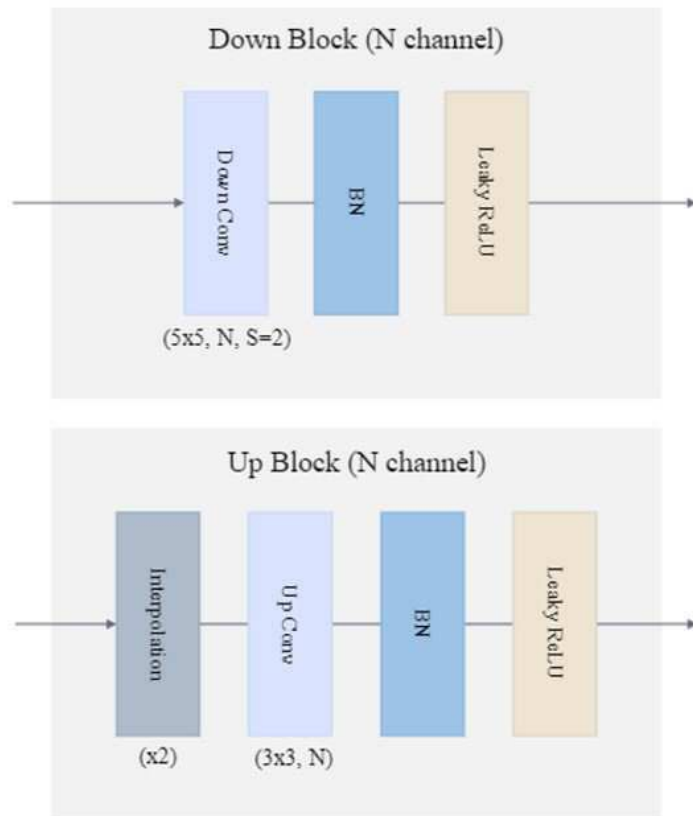
도면2



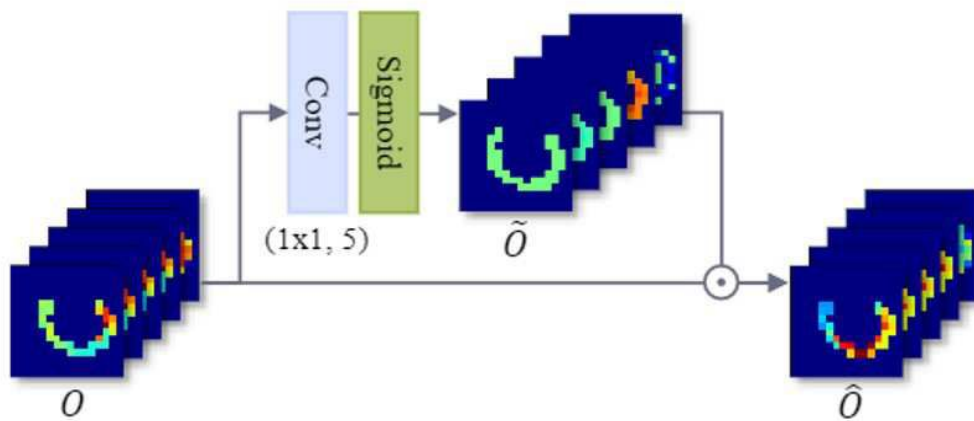
도면3



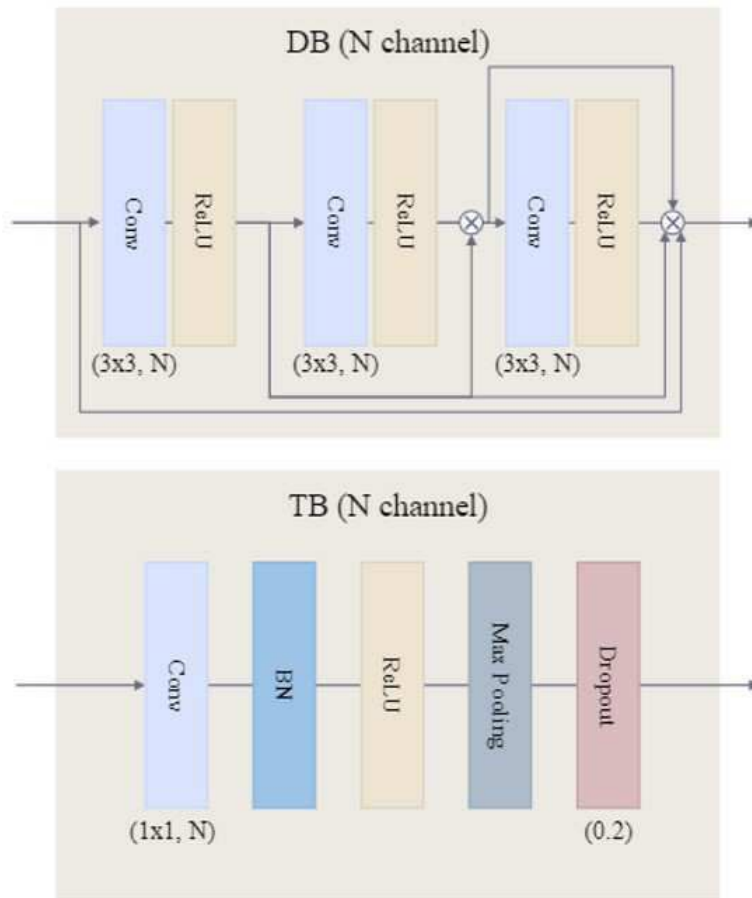
도면4



도면5



도면6



도면7

| Method                                   | Ball         | Bear         | Buddha       | Cat          |
|------------------------------------------|--------------|--------------|--------------|--------------|
| N-LS (Woodham, 1980)                     | 30.20        | 33.33        | 37.88        | 29.93        |
| N-L1 (Ikehata et al., 2012)              | 30.31        | 31.88        | 37.18        | 29.63        |
| N-SBL (Ikehata et al., 2012)             | 30.99        | 31.61        | 37.57        | 29.83        |
| PS-LS (Ikehata et al., 2014)             | 32.84        | 30.51        | 39.44        | 30.82        |
| L-PLSBL (Ikehata et al., 2014)           | 36.25        | 31.40        | 39.72        | 31.80        |
| NR-PS (Li & Li, 2022)                    | 29.33        | 30.15        | 34.26        | 28.30        |
| PS-FCN (Chen et al., 2018)               | 34.54        | 36.26        | 38.05        | 35.01        |
| NormAttention-PSN (Ju et al., 2022)      | 27.01        | 33.73        | 25.62        | 25.42        |
| CNN-PS (Ikehata, 2018), K = 10           | 12.38        | 20.09        | 29.18        | 23.18        |
| PX-Net (Logothetis et al., 2021), K = 10 | 13.40        | 17.96        | 25.68        | 21.11        |
| Ours (lightweights), K = 10              | <u>12.13</u> | <u>14.99</u> | <b>25.11</b> | <u>19.48</u> |
| Ours, K = 10                             | <b>10.46</b> | <b>14.89</b> | <u>25.26</u> | <b>17.43</b> |

도면8

