(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2024년11월19일
(11) 등록번호 10-2732100
(24) 등록일자 2024년11월14일

(51) 국제특허분류(Int. Cl.)
G06T 5/00 (2024.01) *G06N 20/00* (2019.01)
G06T 3/00 (2024.01) *G06T 3/40* (2024.01)
G06T 5/20 (2024.01) *G06T 7/11* (2017.01)
(52) CPC특허분류
G06T 5/80 (2024.01)
G06N 20/00 (2021.08)
(21) 출원번호 10-2023-0072171
(22) 출원일자 2023년06월05일
심사청구일자 2023년06월05일
(56) 선행기술조사문헌
KR1020210069388 A
KR1020190110304 A
KR1020210013834 A

(73) 특허권자
울산과학기술원
울산광역시 울주군 언양읍 유니스트길 50
(72) 발명자
심재영
울산광역시 울주군 언양읍 유니스트길 50
박종혁
울산광역시 울주군 언양읍 유니스트길 50
(뒷면에 계속)
(74) 대리인
특허법인 무한

전체 청구항 수 : 총 18 항 심사관 : 김광식

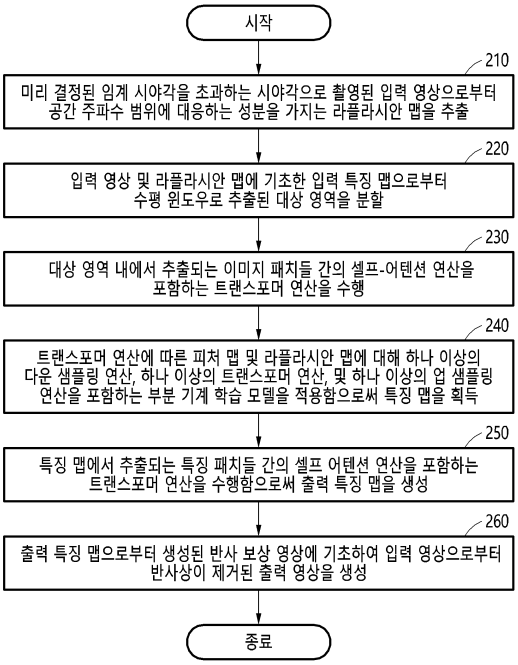
(54) 발명의 명칭 **영상 내 포함된 반사상을 수평 윈도우 트랜스포머 기반의 기계 학습 모델을 이용하여 제거하는 방법 및 장치**

(57) 요 약

반사상 제거 기술은 촬영된 영상의 유리 부분(이하, '유리영역')으로부터 유리영역 내부의 투과 영상을 재구성하기 위한 기술이다. 본 발명은 입력 영상에서 유리영역을 먼저 감지한 후 반사상에 대응되는 오브젝트 부분(이하, '참조영역')의 정보를 이용하여 반사상을 제거하여, 유리영역 영상을 복원하는 기술이다. 본 발명은

(뒷면에 계속)

대 표 도 - 도2



우선 유리영역과 참조영역을 자동으로 감지하고, 참조영역의 정보를 동시에 사용하여 유리영역에서 반사상을 제거하기 위한 end-to-end 프레임워크를 제안한다. 본 발명은, 유리영역과 참조영역 사이의 다양한 기하학적 관계(geometric relation)와 광도 변화(photometric variation)에 기초하여, 입력 영상에 대한 사실적인 반사상(realistic reflection artifact)을 합성함으로써 학습 데이터 세트를 구성한다. 또한, 본 발명은 수평 윈도우(horizontal window) 방식을 사용하는 트랜스포머(transformer) 기반의 U-net 구조를 제안하고, 셀프 어텐션 메커니즘(self-attention mechanism)을 통해 입력 영상에서 유리영역과 참조영역 간의 장거리 종속성(long-range dependency)을 획득하며, 참조영역을 사용하여 반사상을 제거한다.

(52) CPC특허분류
G06T 3/18 (2024.01)
G06T 3/40 (2024.01)
G06T 5/20 (2024.01)
G06T 7/11 (2017.01)
G06T 2207/20081 (2013.01)

(72) 발명자
김현아
울산광역시 울주군 언양읍 유니스트길 50

박은필
울산광역시 울주군 언양읍 유니스트길 50

이 발명을 지원한 국가연구개발사업
과제고유번호 1711193983
과제번호 2020-0-01336-004
부처명 과학기술정보통신부
과제관리(전문)기관명 정보통신기획평가원
연구사업명 정보통신방송혁신인재양성
연구과제명 인공지능대학원지원(울산과학기술원)
기 여 율 20/100
과제수행기관명 울산과학기술원
연구기간 2023.01.01 ~ 2023.12.31
이 발명을 지원한 국가연구개발사업
과제고유번호 1711193622
과제번호 2021-0-02068-003
부처명 과학기술정보통신부
과제관리(전문)기관명 정보통신기획평가원
연구사업명 정보통신방송혁신인재양성
연구과제명 인공지능 혁신 허브 연구 개발
기 여 율 10/100
과제수행기관명 고려대학교산학협력단
연구기간 2023.01.01 ~ 2023.12.31
이 발명을 지원한 국가연구개발사업
과제고유번호 1711166373
과제번호 2020R1A2B5B01002725
부처명 과학기술정보통신부
과제관리(전문)기관명 한국연구재단
연구사업명 개인기초연구(과기정통부)
연구과제명 딥러닝 기반 유리 왜곡 모델링 및 영상 복원 기술 개발
기 여 율 70/100
과제수행기관명 울산과학기술원
연구기간 2022.03.01 ~ 2023.02.28

명 세 서

청구범위

청구항 1

프로세서에 의해 수행되는 반사상 제거 방법에 있어서,

미리 결정된 임계 시야각을 초과하는 시야각으로 촬영된 입력 영상으로부터 공간 주파수 범위에 대응하는 성분을 가지는 라플라시안 맵(laplacian map)을 추출하는 단계;

상기 입력 영상(input image) 및 상기 라플라시안 맵을 연결(concatenate)한 입력 특징 맵으로부터 수평 윈도우(horizontal window)로 추출된 대상 영역을 분할(segment)하는 단계;

상기 대상 영역 내에서 추출되는 이미지 패치들 간의 셀프-어텐션(self-attention) 연산을 포함하는 트랜스포머(transformer) 연산을 수행하는 단계;

상기 트랜스포머 연산에 따른 피처 맵(feature map) 및 상기 라플라시안 맵에 대해 하나 이상의 다운 샘플링(down sampling) 연산, 하나 이상의 트랜스포머 연산, 및 하나 이상의 업 샘플링(up sampling) 연산을 포함하는 부분 기계 학습 모델을 적용함으로써 특징 맵을 획득하는 단계;

상기 특징 맵에서 추출되는 특징 패치들 간의 셀프 어텐션 연산을 포함하는 트랜스포머 연산을 수행함으로써 출력 특징 맵을 생성하는 단계; 및

상기 출력 특징 맵으로부터 생성된 반사 보상 영상(reflection compensation image)에 기초하여 상기 입력 영상으로부터 상기 반사상이 제거된 출력 영상을 생성하는 단계

를 포함하는 반사상 제거 방법.

청구항 2

제1항에 있어서,

상기 부분 기계 학습 모델은,

동일한 수의 다운 샘플링 연산 및 업 샘플링 연산을 포함하고,

트랜스포머 연산을 상기 다운 샘플링 연산의 수에 두 배에서 1회 차감한 수만큼 포함하는

반사상 제거 방법.

청구항 3

제1항에 있어서,

상기 부분 기계 학습 모델을 적용함으로써 특징 맵을 획득하는 단계는,

상기 하나 이상의 다운 샘플링 연산 중 제1 다운 샘플링 연산의 결과로부터 상기 하나 이상의 트랜스포머 연산을 거쳐 생성된 중간 피처 맵 및 상기 제1 다운 샘플링 연산을 위해 사용된 피처 맵에 대해, 상기 하나 이상의 업 샘플링 연산 중 제1 업 샘플링 연산을 수행하는 단계를 포함하는,

반사상 제거 방법.

청구항 4

제1항에 있어서,

상기 부분 기계 학습 모델을 적용함으로써 특징 맵을 획득하는 단계는,

상기 피처 맵 및 상기 라플라시안 맵에 대해 상기 하나 이상의 다운 샘플링 연산 및 상기 하나 이상의 트랜스포머 연산을 교대로 수행함으로써 중간 피처 맵을 생성하는 단계;

상기 생성된 중간 피처 맵에 상기 하나 이상의 업 샘플링 연산 및 상기 하나 이상의 트랜스포머 연산을 교대로 수행함으로써 상기 특징 맵을 획득하는 단계를 포함하는,

반사상 제거 방법.

청구항 5

제4항에 있어서,

상기 라플라시안 맵을 추출하는 단계는,

복수의 공간 주파수 범위들에 대응하는 복수의 라플라시안 맵들을 추출하는 단계를 포함하고,

상기 중간 피처 맵을 생성하는 단계는,

상기 복수의 라플라시안 맵들 중에서 상기 하나 이상의 다운 샘플링 연산 각각에 대응하는 라플라시안 맵에 기초하여 해당 다운 샘플링 연산을 수행하는 단계를 포함하는,

반사상 제거 방법.

청구항 6

제4항에 있어서,

상기 특징 맵을 획득하는 단계는,

상기 하나 이상의 다운 샘플링 연산 각각에 대응하는 상기 하나 이상의 업 샘플링 연산은 각 다운 샘플링 연산을 위해 사용된 피처 맵을 해당 업 샘플링 연산에 사용하는 단계

를 포함하는 반사상 제거 방법.

청구항 7

제1항에 있어서,

상기 입력 영상에 유리가 포함되었는지 여부, 상기 입력 영상에 반사상이 존재하는지 여부 및 상기 입력 영상에 상기 반사상에 대응하는 객체가 상기 입력 영상 내 특정 범위 내에서 나타나는지 여부에 기초하여 상기 라플라시안 맵을 추출하는 단계, 상기 대상 영역을 분할하는 단계, 상기 트랜스포머 연산을 수행하는 단계, 상기 특징 맵을 획득하는 단계, 상기 출력 특징 맵을 생성하는 단계, 및 상기 입력 영상으로부터 상기 반사상이 제거된 출력 영상을 생성하는 단계의 수행 여부를 결정하는 단계

를 더 포함하는 반사상 제거 방법.

청구항 8

제1항에 있어서,

카메라를 원점으로 상기 입력 영상을 상기 원점이 중심인 단위 구(unit sphere) 표면 레이어에 와핑(warping)하는 단계;

상기 와핑된 입력 영상을 미리 설정된 거리의 평면 레이어에 프로젝션(projection)하는 단계;

상기 프로젝션 된 입력 영상의 가장자리 픽셀 값을 랜덤하게 잘라낸 영역을 추출하는 단계; 및

상기 추출된 영역에 픽셀에 대한 노출 시간 파라미터, 상기 픽셀의 감쇠 파라미터 및 카메라에 의한 영상 보정 파라미터 중 적어도 하나에 기초하여 상기 반사상을 합성하는 단계

를 더 포함하는 반사상 제거 방법.

청구항 9

제8항에 있어서,

상기 감쇠 파라미터의 값이 1미만인

반사상 제거 방법.

청구항 10

제8항에 있어서,

상기 반사상을 합성하는 단계는,

상기 입력 영상으로부터 추출된 영역 이외의 영역에 포함된 픽셀의 광도를 상기 노출 시간 파라미터 및 상기 영상 보정 파라미터 중 적어도 하나에 기초하여 조절하는 단계

를 포함하는 반사상 제거 방법.

청구항 11

하드웨어와 결합되어 제1항 내지 제10항 중 어느 하나의 항의 방법을 실행시키기 위하여 컴퓨터 판독 가능한 기록매체에 저장된 컴퓨터 프로그램.

청구항 12

반사상 제거 장치에 있어서,

미리 결정된 임계 시야각을 초과하는 시야각으로 촬영된 입력 영상으로부터 공간 주파수 범위에 대응하는 성분을 가지는 라플라시안 맵(Laplacian map)을 추출하고,

상기 입력 영상(input image) 및 상기 라플라시안 맵을 연결(concatenate)한 입력 특징 맵으로부터 수평 윈도우(horizontal window)로 추출된 대상 영역을 분할(segment)하며,

상기 대상 영역내에서 추출되는 이미지 패치들 간의 셀프-어텐션(self-attention) 연산을 포함하는 트랜스포머(transformer) 연산을 수행하고,

상기 트랜스포머 연산에 따른 피처 맵(feature map) 및 상기 라플라시안 맵에 대해 하나 이상의 다운 샘플링(down sampling) 연산, 하나 이상의 트랜스포머 연산, 및 하나 이상의 업 샘플링(up sampling) 연산을 포함하는 부분 기계 학습 모델을 적용함으로써 특징 맵을 획득하며,

상기 특징 맵에서 추출되는 특징 패치들 간의 셀프 어텐션 연산을 포함하는 트랜스포머 연산을 수행함으로써 출력 특징 맵을 생성하고,

상기 출력 특징 맵으로부터 생성된 반사 보상 영상(reflection compensation image)에 기초하여 상기 입력 영상으로부터 상기 반사상이 제거된 출력 영상을 생성하는 프로세서

를 포함하는 반사상 제거 장치.

청구항 13

제12항에 있어서,

상기 부분 기계 학습 모델은,

동일한 수의 다운 샘플링 연산 및 업 샘플링 연산을 포함하고,

트랜스포머 연산을 상기 다운 샘플링 연산의 수에 두 배에서 1회 차감한 수만큼 포함하는

반사상 제거 장치.

청구항 14

제12항에 있어서,

상기 프로세서는,

상기 하나 이상의 다운 샘플링 연산 중 제1 다운 샘플링 연산의 결과로부터 상기 하나 이상의 트랜스포머 연산을 거쳐 생성된 중간 피처 맵 및 상기 제1 다운 샘플링 연산을 위해 사용된 피처 맵에 대해, 상기 하나 이상의 업 샘플링 연산 중 제1 업 샘플링 연산을 수행하는

반사상 제거 장치.

청구항 15

제12항에 있어서,

상기 프로세서는

상기 피처 맵 및 상기 라플라시안 맵에 대해 상기 하나 이상의 다운 샘플링 연산 및 상기 하나 이상의 트랜스포머 연산을 교대로 수행함으로써 중간 피처 맵을 생성하고,

상기 생성된 중간 피처 맵에 상기 하나 이상의 업 샘플링 연산 및 상기 하나 이상의 트랜스포머 연산을 교대로 수행함으로써 상기 특징 맵을 획득하는

반사상 제거 장치.

청구항 16

제15항에 있어서,

상기 프로세서는,

복수의 공간 주파수 범위들에 대응하는 복수의 라플라시안 맵들을 포함하는 상기 라플라시안 맵을 추출하고,

상기 복수의 라플라시안 맵들 중에서 상기 하나 이상의 다운 샘플링 연산 각각에 대응하는 라플라시안 맵에 기초하여 해당 다운 샘플링 연산을 수행함으로써 상기 중간 피처 맵을 생성하는

반사상 제거 장치.

청구항 17

제15항에 있어서,

상기 프로세서는,

상기 하나 이상의 다운 샘플링 연산 각각에 대응하는 상기 하나 이상의 업 샘플링 연산은 각 다운 샘플링 연산을 위해 사용된 피처 맵을 해당 업 샘플링 연산에 사용하여 상기 특징 맵을 획득하는

반사상 제거 장치.

청구항 18

제12항에 있어서,

상기 프로세서는

상기 입력 영상에 유리가 포함되었는지 여부, 상기 입력 영상에 반사상이 존재하는지 여부 및 상기 입력 영상에 상기 반사상에 대응하는 객체가 상기 입력 영상 내 특정 범위내에서 나타나는지 여부에 기초하여 상기 라플라시안 맵을 추출하고, 상기 대상 영역을 분할하며, 상기 트랜스포머 연산을 수행하고, 상기 특징 맵을 획득하며, 상기 출력 특징 맵을 생성하고, 및 상기 입력 영상으로부터 상기 반사상이 제거된 출력 영상을 생성할지 여부를 결정하는

반사상 제거 장치.

발명의 설명

기 술 분 야

[0001] 이하, 영상 내 포함된 반사상을 제거하는 이미지 처리 기술이 개시된다.

배 경 기 술

[0002] 유리가 촬영된 영상에서 일반적으로 반사상은 유리내부의 투과 영상과 다른 성질을 가진다. 예를 들어, 반사상

은 투과 영상과 비교하여 더욱 흐릿할(blur) 수 있고, 반사상은 유리의 두께로 인하여 고스트 컨투어(ghost contour) 현상이 발생할 수 있다. 단일 영상 반사상 제거(single image reflection removal) 기술은 상기 반사상의 특징들을 사용하여 제안되었다. 이러한 반사상을 제거하기 위해 딥 러닝 기반의 기술이 사용될 수 있다.

발명의 내용

과제의 해결 수단

[0003] 일 실시예에 따른 반사상 제거 방법은, 미리 결정된 임계 시야각을 초과하는 시야각으로 촬영된 입력 영상으로부터 공간 주파수 범위에 대응하는 성분을 가지는 라플라시안 맵(laplaican map)을 추출하는 단계, 상기 입력 영상(input image) 및 상기 라플라시안 맵에 기초한 입력 특징 맵으로부터 수평 윈도우(horizontal window)로 추출된 대상 영역을 분할(segment)하는 단계, 상기 대상 영역 내에서 추출되는 이미지 패치들 간의 셀프-어텐션(self-attention) 연산을 포함하는 트랜스포머(transformer) 연산을 수행하는 단계, 상기 트랜스포머 연산에 따른 피처 맵(feature map) 및 상기 라플라시안 맵에 대해 하나 이상의 다운 샘플링(down sampling) 연산, 하나 이상의 트랜스포머 연산, 및 하나 이상의 업 샘플링(up sampling) 연산을 포함하는 부분 기계 학습 모델을 적용함으로써 특징 맵을 획득하는 단계, 상기 특징 맵에서 추출되는 특징 패치들 간의 셀프 어텐션 연산을 포함하는 트랜스포머 연산을 수행함으로써 출력 특징 맵을 생성하는 단계, 및 상기 출력 특징 맵으로부터 생성된 반사 보상 영상(reflection compensation image)에 기초하여 상기 입력 영상으로부터 상기 반사상이 제거된 출력 영상을 생성하는 단계를 포함할 수 있다.

[0004] 상기 부분 기계 학습 모델은 동일한 수의 다운 샘플링 연산 및 업 샘플링 연산을 포함하고, 트랜스포머 연산을 상기 다운 샘플링 연산의 수에 두 배에서 1회 차감한 수만큼 포함할 수 있다.

[0005] 상기 부분 기계 학습 모델을 적용함으로써 특징 맵을 획득하는 단계는 상기 하나 이상의 다운 샘플링 연산 중 제1 다운 샘플링 연산의 결과로부터 상기 하나 이상의 트랜스포머 연산을 거쳐 생성된 중간 피처 맵 및 상기 제1 다운 샘플링 연산을 위해 사용된 피처 맵에 대해, 상기 하나 이상의 업 샘플링 연산 중 제1 업 샘플링 연산을 수행하는 단계를 포함할 수 있다.

[0006] 상기 부분 기계 학습 모델을 적용함으로써 특징 맵을 획득하는 단계는 상기 피처 맵 및 상기 라플라시안 맵에 대해 상기 하나 이상의 다운 샘플링 연산 및 상기 하나 이상의 트랜스포머 연산을 교대로 수행함으로써 중간 피처 맵을 생성하는 단계, 상기 생성된 중간 피처 맵에 상기 하나 이상의 업 샘플링 연산 및 상기 하나 이상의 트랜스포머 연산을 교대로 수행함으로써 상기 특징 맵을 획득하는 단계를 포함할 수 있다.

[0007] 상기 라플라시안 맵을 추출하는 단계는 상기 복수의 공간 주파수 범위들에 대응하는 복수의 라플라시안 맵들을 추출하는 단계를 포함하고, 상기 중간 피처 맵을 생성하는 단계는, 상기 복수의 라플라시안 맵들 중에서 상기 하나 이상의 다운 샘플링 연산 각각에 대응하는 라플라시안 맵에 기초하여 해당 다운 샘플링 연산을 수행하는 단계를 포함할 수 있다.

[0008] 상기 특징 맵을 획득하는 단계는 상기 하나 이상의 다운 샘플링 연산 각각에 대응하는 상기 하나 이상의 업 샘플링 연산은 각 다운 샘플링 연산을 위해 사용된 피처 맵을 해당 업 샘플링 연산에 사용하는 단계를 포함할 수 있다.

[0009] 일 실시예에 따른 반사상 제거 방법은 상기 입력 영상에 유리가 포함되었는지 여부, 상기 입력 영상에 반사상이 존재하는지 여부 및 상기 입력 영상에 상기 반사상에 대응하는 객체가 상기 입력 영상 내 특정 범위 내에서 나타나는지 여부에 기초하여 상기 반사상의 제거 동작 수행 여부를 결정하는 단계를 더 포함할 수 있다.

[0010] 일 실시예에 따른 반사상 제거 방법은 픽셀에 대한 노출 시간 파라미터, 상기 픽셀의 감쇠 파라미터 및 카메라에 의한 영상 보정 파라미터 중 적어도 하나에 기초하여 상기 반사상을, 상기 입력 영상 중 카메라를 원점으로 상기 입력 영상을 상기 원점이 중심인 단위 구(unit sphere) 표면 레이어에 와핑(warping)하고, 와핑된 영상을 미리 설정된 거리의 평면 레이어에 프로젝션(projection)하며, 프로젝션 된 입력 영상의 가장자리 픽셀 값을 랜덤하게 잘라내어 추출된 영역에 합성하는 단계를 더 포함할 수 있다.

[0011] 상기 감쇠 파라미터의 값이 1미만일 수 있다.

[0012] 상기 반사상을 합성하는 단계는 상기 입력 영상으로부터 상기 추출된 영역 이외의 영역에 포함된 픽셀의 광도를 상기 노출 시간 파라미터 및 상기 영상 보정 파라미터 중 적어도 하나에 기초하여 조절하는 단계를 포함할 수

있다.

[0014] 일 실시예에 따른 반사상 제거 장치는, 미리 결정된 임계 시야각을 초과하는 시야각으로 촬영된 입력 영상으로부터 공간 주파수 범위에 대응하는 성분을 가지는 라플라시안 맵(Laplacian map)을 추출하고, 상기 입력 영상(input image) 및 상기 라플라시안 맵에 기초한 입력 특징 맵으로부터 수평 윈도우(horizontal window)로 추출된 대상 영역을 분할(segment)하며, 상기 대상 영역내에서 추출되는 이미지 패치들 간의 셀프-어텐션(self-attention) 연산을 포함하는 트랜스포머(transformer) 연산을 수행하고, 상기 트랜스포머 연산에 따른 피처 맵(feature map) 및 상기 라플라시안 맵에 대해 하나 이상의 다운 샘플링(down sampling) 연산, 하나 이상의 트랜스포머 연산, 및 하나 이상의 업 샘플링(up sampling) 연산을 포함하는 부분 기계 학습 모델을 적용함으로써 특징 맵을 획득하며, 상기 특징 맵에서 추출되는 특징 패치들 간의 셀프 어텐션 연산을 포함하는 트랜스포머 연산을 수행함으로써 출력 특징 맵을 생성하고, 상기 출력 특징 맵으로부터 생성된 반사 보상 영상(reflection compensation image)에 기초하여 상기 입력 영상으로부터 상기 반사상이 제거된 출력 영상을 생성하는 프로세서를 포함할 수 있다.

[0015] 상기 부분 기계 학습 모델은, 동일한 수의 다운 샘플링 연산 및 업 샘플링 연산을 포함하고, 트랜스포머 연산을 상기 다운 샘플링 연산의 수에 두 배에서 1회 차감한 수만큼 포함할 수 있다.

[0016] 상기 프로세서는, 상기 하나 이상의 다운 샘플링 연산 중 제1 다운 샘플링 연산의 결과로부터 상기 하나 이상의 트랜스포머 연산을 거쳐 생성된 중간 피처 맵 및 상기 제1 다운 샘플링 연산을 위해 사용된 피처 맵에 대해, 상기 하나 이상의 업 샘플링 연산 중 제1 업 샘플링 연산을 수행할 수 있다.

[0017] 상기 프로세서는 상기 피처 맵 및 상기 라플라시안 맵에 대해 상기 하나 이상의 다운 샘플링 연산 및 상기 하나 이상의 트랜스포머 연산을 교대로 수행함으로써 중간 피처 맵을 생성하고, 상기 생성된 중간 피처 맵에 상기 하나 이상의 업 샘플링 연산 및 상기 하나 이상의 트랜스포머 연산을 교대로 수행함으로써 상기 특징 맵을 획득할 수 있다.

[0018] 상기 프로세서는, 상기 복수의 공간 주파수 범위들에 대응하는 복수의 라플라시안 맵들을 포함하는 상기 라플라시안 맵을 추출하고, 상기 복수의 라플라시안 맵들 중에서 상기 하나 이상의 다운 샘플링 연산 각각에 대응하는 라플라시안 맵에 기초하여 해당 다운 샘플링 연산을 수행함으로써 상기 중간 피처 맵을 생성할 수 있다.

[0019] 상기 프로세서는, 상기 하나 이상의 다운 샘플링 연산 각각에 대응하는 상기 하나 이상의 업 샘플링 연산은 각 다운 샘플링 연산을 위해 사용된 피처 맵을 해당 업 샘플링 연산에 사용하여 상기 특징 맵을 획득할 수 있다.

[0020] 상기 프로세서는, 상기 입력 영상에 유리가 포함되었는지 여부, 상기 입력 영상에 반사상이 존재하는지 여부 및 상기 입력 영상에 상기 반사상에 대응하는 객체가 상기 입력 영상 내 특정 범위내에서 나타나는지 여부에 기초하여 상기 반사상의 제거 동작 수행 여부를 결정할 수 있다.

도면의 간단한 설명

[0021] 도 1은 일 실시예에 따른 반사상 제거 장치의 구성을 도시한 블록도이다.

도 2는 일 실시예에 따른 반사상 제거 방법을 설명하기 위한 흐름도다.

도 3은 수평 윈도우 트랜스포머 연산 블록도 및 셀프 어텐션 연산을 위하여 대상 영역을 분할하고 대상 영역에서 이미지 패치를 추출하는 것을 나타낸다.

도 4는 셀프 어텐션 연산을 위한 대상 영역 추출 과정에서 활용되는 정사각형 윈도우(430)와 수평 윈도우(320)를 비교한 도면이다.

도 5는 도 4의 이미지(420)에서 추출된 대상 영역에 포함되는 일부 이미지 패치들 간의 트랜스포머의 셀프 어텐션 연산 메커니즘을 도시한다.

도 6은 부분 기계 학습 모델의 일 실시예에 따른 연산 블록도다.

도 7은 반사상을 제거하기 위한 전체 기계 학습 모델을 트레이닝 하는 단계에 관한 블록도다.

도 8은 반사상이 합성된 합성 영상과 반사상이 제거된 정답 영상 쌍으로 구성된 트레이닝 데이터 세트를 도시한다.

도 9는 전체 기계 학습 모델을 트레이닝하기 위한 단계(710)에서 미리 설정된 크기의 유리영역을 지정하는 방법

에 관해 도시한다.

발명을 실시하기 위한 구체적인 내용

[0022] 실시예들에 대한 특정한 구조적 또는 기능적 설명들은 단지 예시를 위한 목적으로 개시된 것으로서, 다양한 형태로 변경되어 구현될 수 있다. 따라서, 실제 구현되는 형태는 개시된 특정 실시예로만 한정되는 것이 아니며, 본 명세서의 범위는 실시예들로 설명한 기술적 사상에 포함되는 변경, 균등물, 또는 대체물을 포함한다.

[0023] 제1 또는 제2 등의 용어를 다양한 구성요소들을 설명하는데 사용될 수 있지만, 이런 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 해석되어야 한다. 예를 들어, 제1 구성요소는 제2 구성요소로 명명될 수 있고, 유사하게 제2 구성요소는 제1 구성요소로도 명명될 수 있다.

[0024] 어떤 구성요소가 다른 구성요소에 "연결되어" 있다고 언급된 때에는, 그 다른 구성요소에 직접적으로 연결되어 있거나 또는 접속되어 있을 수도 있지만, 중간에 다른 구성요소가 존재할 수도 있다고 이해되어야 할 것이다.

[0025] 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 명세서에서, "포함하다" 또는 "가지다" 등의 용어는 설명된 특징, 숫자, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것이 존재함으로 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.

[0026] 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 해당 기술분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가진다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥상 가지는 의미와 일치하는 의미를 갖는 것으로 해석되어야 하며, 본 명세서에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.

[0027] 이하, 실시예들을 첨부된 도면들을 참조하여 상세하게 설명한다. 첨부 도면을 참조하여 설명함에 있어, 도면 부호에 관계없이 동일한 구성 요소는 동일한 참조 부호를 부여하고, 이에 대한 중복되는 설명은 생략하기로 한다.

[0028] 단일 영상에서는 영상 전체 픽셀에서 반사가 발생하는 반면, 360도 영상에서는 일부 픽셀에서만 반사가 발생한다. 그러므로, 360도 영상 내에서 반사가 발생하는 유리영역을 찾을 필요가 있다. 그러나, 유리영역을 찾는 것은 쉬운 작업이 아니며, 기존의 방법은 유리영역이 사용자 상호 작용에 의해 제공된다고 가정한다. 또한, 휴리스틱 매칭 기법(heuristic matching scheme)을 이용하여 유리영역에 포함된 반사상을 제거하기 위하여, 반사상과 대응되는 오브젝트가 포함된 영역으로서, 반사상 제거 시 참조하는 영역(이하, '참조영역')의 정보를 활용한다. 일 실시예에 따르면, 반사상 제거 장치는 사용자와의 상호 작용 없이 유리영역과 참조영역을 찾고, 입력 영상에 대한 완전 자동 end-to-end 반사상 제거 네트워크를 포함한다.

[0029] 카메라로 유리부분을 촬영하면, 촬영된 영상에서 유리 내부 영상에 해당하는 투과 영상 T와 유리에 의한 반사상 R이 함께 촬영된 유리영역 M이 획득될 수 있다.

[0030] 유리영역 M은 하기의 수학식 1에 의해 계산될 수 있다.

수학식 1

[0031] $$\mathbf{M}(\mathbf{x}) = \Omega(\mathbf{x})\mathbf{T}(\mathbf{x}) + \Phi(\mathbf{x})\mathbf{R}(\mathbf{x})$$

[0032] 여기서, M(x), T(x), 및 R(x)는 입력 영상에서 유리영역, 투과 영상 및 반사상에서의 픽셀 x의 강도(intensity)를 나타낸다. Ω 및 Φ는 각각 유리의 특성과 관련된 굴절 진폭 계수 맵(refractive amplitude coefficient map)과 반사 진폭 계수 맵(reflective amplitude coefficient map)을 나타낸다. 반사상 제거 기술은 유리영역 M으로부터 투과 영상 T를 재구성하는 것이 목적이다. 유리영역 M에 투과 영상 T와 반사상 R이 함께 캡처되므로, T와 R을 구분하는 것은 모호할 수 있다. 그러한 모호성을 극복하기 위해, 반사상 제거 장치는 R과 T를 구분하기 위해, 부드러움 우선(smoothness prior) 및 고스팅 클루(ghosting clue)와 같은 R에 대한 통계치를 사용한다.

[0033] 반사상 제거 기술은, 반사상을 포함하는 유리영역 $\mathcal{M}$과 해당 반사상에 대응하는 오브젝트의 전면부를 포함하는 참조영역 $\mathcal{R}$을 모두 포함하는 영상에 대하여 반사상을 제거하기 위한 기술로서 제안된다. 반사상 제거 장치는 참조영역 $\mathcal{R}$의 정보를 이용하여 효과적으로 유리영역 $\mathcal{M}$에 포함된 반사상을 제거할 수 있다.

[0034] 그러나, 일반적인 영상에서는 반사가 영상 전체 픽셀에 대해 유효하게 적용된다고 가정하나, 360도 영상에서 반사상은 영상의 일부 픽셀에서 발생한다. 반사상 제거 장치는, 첫째, 360도 영상 내에서 유리영역 $\mathcal{M}$에 포함된 반사상에 대응하는 참조영역 $\mathcal{R}$뿐만 아니라, 반사상 제거 기술에 의해 복원될 유리영역 $\mathcal{M}$을 먼저 찾아야 한다. 둘째, 참조영역 $\mathcal{R}$의 정보를 이용하여, 유리영역 $\mathcal{M}$에서 투과 영상 T를 재구성해야 한다. 이 때, 반사상 제거 장치가 유리영역 및 참조영역을 탐지하는 것과 투과 영상을 재구성하는 것은 순환적인 관계로 인해 어떤 것을 먼저 해결해야 할지 결정하기 어려운 문제이다. 반사상 제거 장치가 참조영역 $\mathcal{R}$을 감지할수록, 투과 영상 T의 더 나은 재구성 결과가 산출될 수 있다. 마찬가지로, 반사상 제거 장치가 투과 영상 T에 대하여 더 나은 재구성 결과를 산출할수록, 참조영역 $\mathcal{R}$을 더욱 잘 감지할 수 있다.

[0035] 일 실시예에 따른, 반사상 제거 장치는, 완전 자동으로 유리영역 $\mathcal{M}$과 참조영역 $\mathcal{R}$을 감지하고 투과 영상을 재구성하는 360도 영상에 대한 end-to-end 반사상 제거 네트워크를 포함한다. 우선, 반사상 제거 장치는 360도 영상에 대한 사실적인 반사상을 합성하여 네트워크 훈련을 위한 합성영상 및 반사상이 제거된 정답영상 쌍 데이터 세트를 생성한다. 반사상 제거 장치는 다양한 기하학적 관계(geometric relation)를 시뮬레이션하고 유리영역 $\mathcal{M}$과 참조영역 $\mathcal{R}$ 사이의 광도 변화(photometric variation)를 조정하여 보다 사실적으로 반사상을 합성한다. 또한, 반사상 제거 장치는 360도 영상에서 유리영역 $\mathcal{M}$과 참조영역 $\mathcal{R}$ 사이의 관계를 캡처하기 위해 수평 윈도우 방식(horizontal windowing scheme)의 트랜스포머 연산 블록을 U-net 구조로 포함하는 기계 학습 모델을 포함한다. 또한, 반사상 제거 장치는 다중 스케일의 고주파 성분에 해당하는 특징을 활용하기 위해 라플라시안 피라미드를 사용한다.

[0036] 도 1은 일 실시예에 따른 반사상 제거 장치의 구성을 도시한 블록도이다.

[0037] 반사상 제거 장치(111)는 프로세서(112), 메모리(113), 영상획득부(114), 및 디스플레이(115)를 포함할 수 있다.

[0038] 프로세서(112)는, 촬영된 입력 영상(110)으로 부터, 수평 윈도우(horizontal window) 기반의 트랜스포머(transformer)에 기초하는 전체 기계 학습 모델(100)을 포함하고, 전체 기계 학습 모델을 통해 입력 영상(110) 내에서 반사상을 제거하여 출력 영상(120)을 생성할 수 있다.

[0039] 메모리(113)는, 촬영되거나 수신된 입력 영상(110)을 저장할 수 있다. 메모리(113)는 라플라시안 피라미드 생성기(130) 및 부분 기계 학습 모델(160)에 관한 정보를 저장할 수 있다. 메모리(113)는 라플라시안 피라미드 생성기(130)에 의하여 생성된 라플라시안 맵(140)에 관한 정보를 저장할 수 있다. 메모리(113)는 부분 기계 학습 모델(160)에 기초하여 생성된 피처 맵 또는 특징 맵 정보를 저장할 수 있다. 또한, 메모리(113)는 입력 영상(110)에서 반사상이 제거된 출력 영상(120)을 획득하기 위해 요구되는 데이터 또는 정보를 임시적으로 또는 영구적으로 저장할 수 있다.

[0040] 영상획득부(114)는 입력 영상(110)을 촬영하거나, 미리 촬영된 입력 영상(110)을 수신할 수도 있다. 예를 들어, 영상 획득부(114)는 카메라 및/또는 통신부를 포함할 수 있다.

[0041] 디스플레이(115)는 프로세서에 의해 생성된 출력 영상(120)을 시각화 할 수 있다.

[0042] 수평 윈도우 기반의 트랜스포머에 기초하는 전체 기계 학습 모델(100)은 라플라시안 피라미드 생성기(laplacian pyramid generator)(130)와 부분 기계 학습 모델(160)을 포함할 수 있다.

[0043] 전체 기계 학습 모델(100)은 라플라시안 피라미드 생성기(130)에 의해 생성된 라플라시안 맵(140)과 입력 영상(110)에 기초한 입력 특징 맵으로부터 반복적으로 트랜스포머 연산을 수행하여 출력 특징 맵을 생성하는 연산을 포함할 수 있다. 전체 기계 학습 모델(100)은 생성된 출력 특징 맵에 기초하여, 반사 보상 영상을 생성할 수

있다. 전체 기계 학습 모델(100)은 반사 보상 영상과 입력 영상에 기초하여 입력 영상에서 반사상이 제거된 출력 영상(120)을 생성할 수 있다.

[0044] 부분 기계 학습 모델(160)은 입력 영상(110)에 대하여 복수의 다운 샘플링 연산 및 트랜스포머 연산을 수행하여 중간 피처 맵을 생성할수 있다. 또한, 부분 기계 학습 모델(160)은 중간 피처 맵에 대하여 복수의 업 샘플링 연산 및 트랜스포머 연산을 수행하여 특징 맵을 생성할 수 있다. 부분 기계 학습 모델(160)에 대하여 도 6에서 구체적으로 후술한다.

[0045] 라플라시안 피라미드 생성기(130)는 복수의 공간 주파수 범위들에 대응하는 복수의 라플라시안 맵들을 생성할 수 있다. 예를 들어, 라플라시안 피라미드 생성기(130)는 복수의 공간 주파수 범위들의 각각에 대응하는 라플라시안 맵(Laplacian map)을 생성할 수 있다. 임의의 이미지에 대한 공간 주파수는, 예를 들어, 해당 이미지에서 일축(예: 세로축 또는 가로축)을 따른 픽셀 값의 변화율일 수 있다. 픽셀 값의 변화율이 낮은 주파수 범위는 저주파 범위, 픽셀값의 변화율이 높은 주파수 범위는 고주파 범위일 수 있다. 임의의 공간 주파수 범위에 대응하는 라플라시안 맵은 해당하는 공간 주파수 성분의 값들을 가질 수 있다. 라플라시안 피라미드 생성기(130)는 가장 낮은 주파수 범위에 대응하는 라플라시안 맵부터 가장 높은 주파수 범위에 대응하는 라플라시안 맵을 생성할 수 있다. 각 주파수 범위는 다른 주파수 범위와 중첩되지 않을 수 있으나, 이로 한정하는 것은 아니고, 일부 범위에서 중첩될 수도 있다. 라플라시안 피라미드 생성기(130)는 각 주파수 범위에 대응하는 라플라시안 맵을 추출하도록 설계 및 트레이닝된 기계 학습 모델(예: 뉴럴 네트워크)를 포함할 수 있다. 라플라시안 피라미드 생성기(130)는 입력 영상(110)으로부터 전술된 기계 학습 모델에 기초하여 라플라시안 맵을 생성할 수 있다. 기계 학습 모델의 연산에 기초하여 생성된 라플라시안 맵은, 영상에 대해 라플라시안 필터를 적용한 결과로부터 생성되는 맵과 동일 또는 유사할 수 있다.

[0046] 다만, 이로 한정하는 것은 아니고, 라플라시안 피라미드 생성기(130)는 입력 영상에 대해 라플라시안 필터를 적용한 결과로부터 라플라시안 맵을 획득할 수도 있다. 예를 들어, 라플라시안 피라미드 생성기(130)는 임의의 공간 주파수 범위의 상한 주파수로 라플라시안 필터를 영상에 적용한 결과로부터 해당 공간 주파수 범위의 하한 주파수로 라플라시안 필터를 영상에 적용한 결과를 차감함으로써, 해당 공간 주파수 범위의 라플라시안 맵을 생성할 수 있다.

[0047] 도 2는 일 실시예에 따른 반사상 제거 방법을 설명하기 위한 흐름도다.

[0048] 단계(210)에서, 반사상 제거 장치는, 미리 결정된 임계 시야각을 초과하는 시야각으로 촬영된 입력 영상으로부터 공간 주파수 범위에 대응하는 성분을 가지는 라플라시안 맵을 추출할 수 있다.

[0049] 일 실시예에 따르면 반사상 제거 장치는 미리 결정된 임계 시야각을 초과하는 시야각으로 촬영된 입력 영상을 획득할 수 있다. 반사상 제거 장치는 카메라를 이용하여 입력 영상을 직접 촬영하거나, 외부 카메라 모듈로부터 촬영된 입력 영상을 수신할 수 있다. 미리 결정된 임계 시야각은 180도를 초과하는 광시야각일 수 있다. 입력 영상이 광시야각으로 촬영될 경우, 입력 영상이 유리에 나타난 반사상과 반사상에 대응하는 오브젝트의 전면부 모두를 포함할 가능성이 높다. 미리 결정된 임계 시야각을 초과하는 시야각으로 촬영된 입력 영상이 예를 들어, 360도의 시야각을 갖는 영상(이하, 360도 영상)일 경우, 360도 영상은 장애물에 의해 가려지지 않는 한 반사상과 반사상에 대응하는 오브젝트의 전면부를 확정적으로 포함할 수 있다. 참고로, 오브젝트의 전면부의 적어도 일부가 확보되기 위해서는, 오브젝트가 촬영 카메라의 광축 대비 90도를 초과하는 각도를 형성하는 위치에 배치된다고 가정될 수 있다.

[0050] 일 실시예에 따르면, 반사상 제거 장치는, 복수의 라플라시안 맵들을 추출할 수 있다. 복수의 라플라시안 맵들은 복수의 공간 주파수 범위들 각각에 대응할 수 있다. 도 6에서 후술하겠으나, 복수의 라플라시안 맵들은 부분 기계 학습 모델에서 복수의 다운 샘플링 연산이 반복 또는 깊어짐에 따라, 개별적으로 다운 샘플링 연산될 수 있다.

[0051] 단계(220)에서, 반사상 제거 장치는, 입력 영상 및 라플라시안 맵에 기초한 입력 특징 맵으로부터 수평 윈도우로 추출된 대상 영역을 분할할 수 있다.

[0052] 일 실시예에 따르면, 반사상 제거 장치는, 입력 영상과 라플라시안 맵을 연결(concatenation)한 데이터를 생성할 수 있다. 반사상 제거 장치는 뉴럴 네트워크 연산에 기초하여, 입력 영상과 라플라시안 맵을 연결한 데이터로부터 입력 특징 맵을 생성할 수 있다. 예를 들어, 뉴럴 네트워크 연산은 컨볼루션 신경망(convolutional neural network, CNN)을 포함할 수 있으나, 이에 한정되지 않는다.

[0053] 수평 윈도우는 트랜스포머의 셀프 어텐션(self-attention) 연산 레이어에서 활용되며, 셀프 어텐션 커널

(kernel) 이라 지칭될 수 있다. 또한, 수평 윈도우의 크기는 가로×세로의 직사각형 모양일 수 있다. 가로 길이는 입력 영상의 가로 길이일 수 있으며, 세로 길이는 단위 길이일 수 있다. 예를 들어, n×W 크기의 수평 윈도우가 셀프 어텐션 연산에 활용될 수 있다. 이때, W는 입력 영상의 너비일 수 있고, n은 하이퍼 파라미터 일 수 있다.

[0054] 단계(230)에서, 반사상 제거 장치는, 대상 영역 내에서 추출되는 이미지 패치들 간의 셀프-어텐션(self-attention) 연산을 포함하는 트랜스포머(transformer) 연산을 수행할 수 있다.

[0055] 수평 윈도우를 통해 추출된 대상 영역에 영상 내 반사상을 포함하는 유리영역과 유리영역에 포함된 반사상을 제거하기 위하여, 반사상과 대응되는 오브젝트가 포함된 영역 으로서, 반사상 제거 시 참조하는 영역(이하, '참조영역')이 나타날 수 있다. 수평 윈도우를 통해 추출된 대상 영역이 유리영역과 참조영역을 포함하므로, 반사상 제거 장치는 대상 영역 내에서 유리영역에 포함되는 이미지 패치와 참조영역에 포함되는 이미지 패치 간의 어텐션 연산을 강화할 수 있다.

[0056] 단계(240)에서, 반사상 제거 장치는, 트랜스포머 연산에 따른 피처 맵(feature map) 및 라플라시안 맵에 대해 하나 이상의 다운 샘플링(down sampling) 연산, 하나 이상의 트랜스포머 연산, 및 하나 이상의 업 샘플링(up sampling) 연산을 포함하는 부분 기계 학습 모델을 적용함으로써 특징 맵을 획득할 수 있다.

[0057] 입력 영상 및 라플라시안 맵에 기초하여, 다운 샘플링 연산, 트랜스포머 연산 및 업 샘플링 연산을 거쳐 특징 맵을 생성하는 부분 기계 학습 모델은 트랜스포머 기반의 U-net 구조로 모델링 될 수 있다.

[0058] 반사상 제거 장치는, 다운 샘플링 연산을 통해 트랜스포머 연산에 따라 추출된 입력 영상의 특징을 보존하면서 정보를 압축할 수 있다. 다운 샘플링 연산은 트랜스포머 연산에 따른 피처 맵 및 라플라시안 맵을 연결(concatenation)한 정보에 대하여 수행될 수 있고, 트랜스포머 연산에 따른 피처 맵은 다운 샘플링에 대응하는 해당 업 샘플링 연산 처리될 수 있다. 부분 기계 학습 모델의 다운 샘플링 및 업 샘플링 과정은 도 6에서 구체적으로 후술한다.

[0059] 단계(250)에서, 반사상 제거 장치는, 특징 맵에서 추출되는 특징 패치들 간의 셀프 어텐션 연산을 포함하는 트랜스포머 연산을 수행함으로써 출력 특징 맵을 생성할 수 있다.

[0060] 특징 패치들은 부분 기계 학습 모델에 의해 추출된 특징 맵에 대하여, 반사상 제거 장치가 수평 윈도우 트랜스포머에서 셀프 어텐션 연산을 할 때, 수평 윈도우에 의해 추출된 대상 영역에 포함되는 이미지 패치일 수 있다. 예를 들어, 이미지 패치의 크기는 1 픽셀 × 1 픽셀일 수 있다.

[0061] 단계(260)에서, 반사상 제거 장치는, 출력 특징 맵으로부터 생성된 반사 보상 영상에 기초하여 입력 영상으로부터 반사상이 제거된 출력 영상을 생성할 수 있다.

[0062] 반사 보상 영상은 뉴럴 네트워크 연산에 기초하여, 출력 특징 맵으로부터 획득될 수 있다. 예를 들어, 뉴럴 네트워크 연산은 컨볼루션 신경망(CNN)을 포함할 수 있으나, 이에 한정되지 않는다.

[0063] 반사상 제거 장치는, 반사 보상 영상을 입력 영상과 행렬 성분 합(element wise addition)하여 입력 영상에서 반사상이 제거된 출력 영상을 생성할 수 있다.

[0064] 도 3은 수평 윈도우 트랜스포머 연산 블록도 및 셀프 어텐션 연산을 위하여 대상 영역을 분할하고 대상 영역에서 이미지 패치를 추출하는 것을 나타낸다.

[0065] 반사상 제거 장치는, 수평 윈도우 기반 셀프 어텐션 연산 블록(310)을 포함하는 트랜스포머 연산을 수행할 수 있다. 예를 들어, 반사상 제거 장치는, 입력 영상의 가로 길이와 같은 가로 길이를 갖는 수평 윈도우(320)를 이용하여 입력 영상에서 대상 영역(330)을 분할할 수 있다. 반사상 제거 장치는 분할된 각 대상 영역(330)내에서 가로 및 세로의 길이를 단위 픽셀 값으로 하여 추출할 수 있다. 반사상 제거 장치는, 추출된 이미지 패치(340)들 간의 셀프 어텐션 연산(350)을 수행할 수 있다.

[0066] 도 4는 셀프 어텐션 연산을 위한 대상 영역 추출 과정에서 활용되는 정사각형 윈도우(430)와 수평 윈도우(320)를 비교한 도면이다.

[0067] 일 실시예에 따르면, 이미지(410)는 정사각형 윈도우(430)에 의해 대상 영역으로 분할될 수 있다. 정사각형 윈도우(430)에 의해 추출된 대상 영역은 수평 길이가 제한되므로 유리영역(440a)과 참조영역(450a)을 함께 포함하기 어렵다.

[0068] 이미지(420)는 수평 윈도우(320)에 의해 대상 영역을 추출할 수 있다. 수평 윈도우(320)에 의해 추출된 대상 영역은 유리영역(440b)에 대한 정보 및 유리영역 내 반사상에 대응하는 오브젝트가 위치될 참조영역(450b)에 대한 정보를 포함할 수 있다.

[0069] 따라서, 수평 윈도우(320)에 의해 추출된 대상 영역 내에서, 유리 영역(440b)에 해당하는 이미지 패치와 참조영역(450b)에 해당하는 이미지 패치를 셀프 어텐션 연산 시 함께 활용함으로써, 이차적 계산 복잡성을 줄일 수 있다.

[0070] 도 5는 도 4의 이미지(420)에서 추출된 대상 영역에 포함되는 일부 이미지 패치들 간의 트랜스포머의 셀프 어텐션 연산 메커니즘을 도시한다.

[0071] 수평 윈도우를 사용함으로써, 유리영역과 참조영역에 대한 영역 추정에 기초하여 셀프 어텐션 연산하는 것을 도시한다. 프로세서는, 대상 영역 중 유리영역에 포함되는 이미지 패치(510a)와 대상영역 중 유리영역 및 참조영역 어디에도 포함되지 않는 이미지 패치(520)간의 내적 값을 산출할 수 있다. 산출된 내적 값은 대상 영역 중 유리영역에 포함되는 이미지 패치(510a)와 대상영역 중 유리영역 및 참조영역 어디에도 포함되지 않는 이미지 패치(520) 간의 어텐션 맵(또는 어텐션 가중치)을 포함할 수 있다. 예를 들어, 어텐션 맵은 1 채널의 흑백 이미지를 포함할 수 있다. 어텐션 맵에 포함된 어텐션 맵 값은 픽셀의 밝기 값에 대응될 수 있다. 픽셀의 밝기 값이 높아 흰색에 가까울수록 높은 어텐션 맵 값으로, 픽셀의 밝기 값이 낮아 검은색에 가까울수록 낮은 어텐션 맵 값으로 지칭될 수 있다.

[0072] 유리영역은 단일 영상 레이어 영역과 다른 특성(예를 들어, 광도, 선명도 등)을 나타낼 수 있다. 다른 특성을 가지는 유리영역과 유리영역 및 참조 영역 어디에도 포함되지 않는 영역 간에는 상대적으로 낮은 어텐션 맵 값을 가질 수 있다. 높은 어텐션 맵 값을 가지는 경우 유리영역으로 추정할 수 있다.

[0073] 참조영역은 반사상에 대응하는 오브젝트에 대한 영상을 포함하므로, 참조영역의 특징을 유리영역에서도 관측할 수 있다. 프로세서는 대상영역 중 참조영역에 포함되는 이미지 패치(530)와 대상영역 중 유리영역에 포함되는 이미지 패치(510b) 간의 내적 값을 산출할 수 있다. 산출된 내적 값은 대상영역 중 참조영역에 포함되는 이미지 패치(530)와 대상영역 중 유리영역에 포함되는 이미지 패치(510b) 간의 어텐션 맵을 포함할 수 있다. 유사한 특성이 관측되는 유리영역과 참조영역 간에는 상대적으로 높은 어텐션 맵 값을 가질수 있다. 프로세서는 높은 어텐션 맵 값을 통해 참조영역을 추정할 수 있다.

[0074] 도 6은 부분 기계 학습 모델의 일 실시예에 따른 연산 블록도다.

[0075] 일 실시예에 따른 반사상 제거 장치는, 동일한 수의 다운 샘플링 연산(630) 및 업 샘플링 연산(640)을 포함하는 부분 기계 학습 모델(160)을 포함할 수 있다. 부분 기계 학습 모델(160)은 다운 샘플링 연산(630)의 수에 대해 두 배에서 1회 차감한 개수만큼의 트랜스포머 연산(150)을 포함할 수 있다.

[0076] 반사상 제거 장치는 부분 기계 학습 모델(160)을 통해 트랜스포머 연산에 따라 생성된 피처 맵(610) 및 라플라시안 맵(140)에 대해 하나 이상의 다운 샘플링 연산(630) 및 하나 이상의 트랜스포머 연산(150)을 교대로 수행함으로써 중간 피처 맵(650)을 생성할 수 있다. 반사상 제거 장치는 부분 기계 학습 모델(160)에 포함된 다운 샘플링 연산(630)을 통하여, 다운 샘플링 연산 직전의 트랜스포머 연산(150)에 의해 생성된 특징 맵의 정보를 압축할 수 있다. 다운 샘플링 연산(630) 및 트랜스포머 연산(150)이 반복되며 깊어짐에 따라, 트랜스포머 연산(150)의 블록(block)의 수와 헤드(head)의 수가 증가할 수 있다. 트랜스포머 연산(150)의 블록의 수는 해당 트랜스포머 연산의 수행 횟수를 의미할 수 있다. 반사상 제거 장치는 부분 기계 학습 모델(160)에 포함된 트랜스포머 연산(150)을 할 때 입력 채널의 수를 헤드 수만큼 스플릿(split) 할 수 있다. 예를 들어, 반사상 제거 장치는 트랜스포머 연산에 대한 256개의 입력 채널이 있는 경우, 헤드가 4인 때 채널의 수를 64개씩 4개의 묶음으로 스플릿 할 수 있다. 반사상 제거 장치는 부분 기계 학습 모델(160)을 통하여 스플릿 된 채널들의 묶음에 해당하는 헤드에 수평 윈도우를 적용하여 셀프 어텐션 연산을 수행할 수 있다. 트랜스포머 연산(150)이 반복되며 깊어짐에 따라 트랜스포머 연산(150)의 블록수와 헤드수가 증가할 수 있고, 이로 인해 트랜스포머 연산량이 증폭될 수 있다. 반사상 제거 장치는 부분 기계 학습 모델(160)에 포함된 다운 샘플링 연산(630)을 통해 정보를 압축할 수 있다. 반사상 제거 장치는 다운 샘플링 연산(630)을 통해 블록수 및 헤드수가 증가함에 따라 발생하는 트랜스포머 연산량의 증폭을 상쇄할 수 있다. 다운 샘플링 연산(630)과 트랜스포머 연산(150)을 교대로 수행하여 중간 피처 맵(650)을 생성하는 연산 과정은 일반적인 U-net 구조에서의 인코더(encoder), 인코딩 절차(encoding process) 또는 인코딩 페이즈(encoding phase) 등으로 표현될 수 있다.

[0077] 일 실시예에 따른 반사상 제거 장치는 부분 기계 학습 모델(160)을 통해 생성된 중간 피처 맵(650)에 하나 이상

의 업 샘플링 연산(640) 및 하나 이상의 트랜스포머 연산(150)을 교대로 수행함으로써 특징 맵(620)을 생성할 수 있다. 반사상 제거 장치는 부분 기계 학습 모델(160)을 통해 중간 피처 맵(650)에 대하여 업 샘플링 연산(640)을 반복 수행함으로써 영상 데이터 정보를 복원할 수 있다. 업 샘플링 연산(640) 및 트랜스포머 연산(150)을 반복하여 획득되는 특징 맵의 정보량이 입력 영상의 정보량에 근사함에 따라, 트랜스포머 연산(150)의 블록수와 헤드수가 감소할 수 있다. 업 샘플링 연산(640)과 트랜스포머 연산(150)을 교대로 수행하는 연산 과정은 일반적인 U-net 구조에서의 디코더(decoder), 디코딩 절차(decoding process) 또는 디코딩 페이즈(decoding phase) 등으로 표현될 수 있다.

[0078] 부분 기계 학습 모델(160)에서 다운 샘플링 연산(630)과정과 업 샘플링 연산(640)과정은 대칭적인 구조를 형성할 수 있다. 하나 이상의 다운 샘플링 연산(630) 각각과 업 샘플링 연산(640) 각각은 U-net 구조의 동일한 레벨에서 연결될 수 있다. 일 실시예에 따라면, 부분 기계 학습 모델에서 마지막 다운 샘플링 연산에 해당하는 제1 다운 샘플링 연산(660)과 첫 업 샘플링 연산에 해당하는 제1 업 샘플링 연산(670)을 동일한 레벨이라 할 수 있다. 제1 다운 샘플링 연산(660) 직전의 다운 샘플링 연산과 제1 업 샘플링(670)연산 직후의 업 샘플링 연산을 동일한 레벨이라 할 수 있다. 즉, 다운 샘플링 연산(630)과 업 샘플링 연산(640)이 각각 N회 반복되는 경우, k번째 다운 샘플링 연산과 (N-k+1)번째 업 샘플링 연산을 동일한 레벨이라 할 수 있다. 반사상 제거 장치는 부분 기계 학습 모델(160)에 포함된 다운 샘플링 연산(630)에서 업 샘플링 연산(640)으로 전환하는 때, 제1 다운 샘플링 연산에 사용된 피처 맵을 대응되는 레벨의 제1 업 샘플링 연산에 다시 활용하는 스킵 커넥션(skip connection)을 수행할 수 있다. 또한, 반사상 제거 장치는 부분 기계 학습 모델(160)에 포함된 k 번째 다운 샘플링 연산에 사용된 피처 맵을 동일한 레벨의 (N-k+1)번째 업 샘플링 연산에 스킵 커넥션하여 연결할 수 있다. 다운 샘플링 연산 및 그에 대응되는 업 샘플링 연산 간의 스킵 커넥션은 다운 샘플링 연산에서 업 샘플링 연산으로의 전환 과정 뿐만 아니라, 모든 대응하는 레벨 간에서 활용될 수 있다. 즉, 반사상 제거 장치는, 하나 이상의 다운 샘플링 연산 각각과 동일한 레벨의 하나 이상의 업 샘플링 연산에 각 다운 샘플링 연산을 위해 사용된 피처 맵을 사용하여 특징 맵을 획득한다.

[0079] 일 실시예에 따르면, 반사상 제거 장치는, 복수의 공간 주파수 범위들에 대응하는 복수의 라플라시안 맵(140)을 추출하고, 복수의 라플라시안 맵(140) 중에서 하나 이상의 다운 샘플링 연산 각각에 대응하는 라플라시안 맵에 기초하여 해당 다운 샘플링 연산을 수행함으로써 중간 피처 맵을 생성한다. 예를 들어, 초기 다운 샘플링 연산은 입력 영상에서 높은 공간 주파수 범위에 해당하는 라플라시안 맵에 기초하여 수행될 수 있다. 높은 공간 주파수 범위에 해당하는 라플라시안 맵은 입력 영상에서 일 축(예를 들어, 가로축 또는 세로축)을 기준으로 변화율이 큰 부분에 대응하는 특징 맵일 수 있다. 예를 들어, 입력 영상의 엣지(edge) 성분, 선명한 경계, 또는 급격한 색상 변화와 같은 이미지 정보를 강조하는 라플라시안 맵일 수 있다. 더 깊은 단계의 다운 샘플링 연산에서는 낮은 공간 주파수 범위에 해당하는 라플라시안 맵을 이용할 수 있다. 낮은 공간 주파수 범위에 해당하는 라플라시안 맵은 입력 영상에서 일 축을 기준으로 변화율이 작은 부분에 대응하는 특징 맵일 수 있다. 예를 들어, 색감 정보, 부드러운 그라데이션 및 배경과 같은 이미지 정보를 강조하는 라플라시안 맵일 수 있다.

[0080] 도 7은 반사상을 제거하기 위한 전체 기계 학습 모델을 트레이닝 하는 단계에 관한 블록도다.

[0081] 단계(710)에서 반사상 제거 장치는 입력 영상을 수신하여, 입력 영상 내 미리 설정된 크기의 영역에 반사상을 합성한다. 반사상 제거 장치는, 반사상을 합성하여 획득된 합성 영상과 반사상이 제거된 정답 영상의 쌍을 트레이닝 데이터 세트로 구성하여, 전체 기계 학습 모델의 파라미터를 업데이트 함으로써 트레이닝 할 수 있다. 이때, 트레이닝 방법은 지도 학습(supervised learning) 방식을 포함할 수 있다. 예를 들어, 지도 학습 방식으로 트레이닝 하는 경우, 반사상 제거 장치는 합성 영상과 정답 영상으로 구성된 트레이닝 데이터 세트를 입력으로하여 임시 출력 영상을 획득하고, 임시 출력 영상과 정답 영상 간의 오차를 목적 함수 값으로 산출하여, 목적 함수의 값에 기초하여 기계 학습 모델의 파라미터를 반복적으로 업데이트 할 수 있다. 이때 목적 함수의 값은 최소화, 최대화 또는 특정 값에 수렴하는 등의 경우 중 하나 일 수 있다. 단계(710)에서 미리 설정된 크기의 영역을 랜덤하게 설정하는 방법에 대하여는 추후 도 9에서 상세 설명한다.

[0082] 도 8은 반사상이 합성된 합성 영상과 반사상이 제거된 정답 영상 쌍으로 구성된 트레이닝 데이터 세트를 도시한다.

[0083] 실제 반사상을 포함하는 360도 영상(801)의 중앙부에 유리 부분(810)이 나타날 수 있다. 유리 부분(810)에서 반사상이 나타날 수 있다. 유리 부분(810) 내부와 외부의 밝기 차이로 인해, 광포화 부분(820)이 360도 영상(801)에서 나타날 수 있다. 광포화 부분(820)은 카메라의 ISP 보정에 의해 광도가 포화된 부분을 의미할 수 있다.

[0084] 반사상이 포함되어 있지 않은 360도 영상(802)은 트레이닝 데이터 세트 제작을 위해 필요한 합성 영상을 생성하기 위한 일 실시 예이다. 반사상 제거 장치는 반사상이 포함되어 있지 않은 360도 영상(802)에 미리 설정된 크기의 영역을 랜덤하게 추출하여 유리영역(830a)으로 설정할 수 있다. 참조영역(840a)은 360도 영상(802)에서 유리영역(830a)의 정반대에 위치할 수 있다.

[0085] 기존 이미지 모델링을 이용한 반사상 합성 영상(803)은 360도 영상(802)에서 상기 수학식 1에 기초하여 반사상을 합성한 이미지를 나타낸다. 실제 반사상을 포함하는 360도 영상(801)의 유리 부분(810)과 광포화 부분(820)을 비교하면, 참조영역이 유리영역에 비하여 더 밝은 픽셀 값 강도를 가질 수 있다. 또한, 촬영 장치 자체 이미지 신호 처리(image signal processing, ISP)에 의해 입력 영상의 밝기가 보정될 수 있고, 투과 영상과 반사상 간의 광도 차이에 영향을 줄 수 있다. 수학식 1에 따라서, 반사상 제거 장치가 유리영역(830b)에 참조영역(840b)을 합성하는 경우, 합성전에 비하여 유리영역(830b)의 밝기가 영상 내 유리영역(830b) 주변에 비하여 밝아지거나, 광포화될 수 있다.

[0086] 일 실시 예에 따른 반사상 제거 장치의 합성 영상(804)은 수학식 2에 기초하여, 유리영역에 반사상을 합성한 이미지를 나타낸다.

수학식 2

[0087] $$\mathbf{M}(\mathbf{x}) = \kappa(\Omega(\mathbf{x}) \cdot \alpha\hat{\mathbf{T}}(\mathbf{x}) + \Phi(\mathbf{x}) \cdot \hat{\mathbf{R}}(\mathbf{x}))^{1/\gamma}$$

[0088] κ는 영상 촬영 시 픽셀에 대한 노출시간 파라미터를 나타내고, α는 감쇠 파라미터(attenuation parameter)를 나타내며, γ는 카메라에 의한 영상 보정 파라미터를 나타낸다. 반사상 제거 장치는, 픽셀에 대한 노출시간 파라미터(κ), 픽셀의 감쇠 파라미터(α) 및 카메라에 의한 영상 보정 파라미터(γ) 중 적어도 하나에 기초하여, 반사상을 입력 영상 중 랜덤하게 추출된 영역에 합성할 수 있다. 예를 들어, 픽셀의 감쇠 파라미터(α)의 값은 1 미만일 수 있다. 반사상 제거 장치는 픽셀에 대한 노출 시간 파라미터(κ) 및 카메라에 의한 영상 보정 파라미터(γ)의 값을 조절하여, 아래의 수학식 3에 따라 반사상이 합성된 360도 영상 I를 획득할 수 있다. 예를 들어, 반사상 제거 장치는, 상계 값과 하계 값을 각각 1.2와 1.0으로 지정하고 해당 범위 사이의 값을 소수점 넷째 자리까지 무작위로 추출하는 것에 기초하여 카메라에 의한 영상 보정 파라미터(γ)를 조절할 수 있다. 다른 예를 들어, 반사상 제거 장치는, 상계 값과 하계 값을 각각 1.1과 1로 지정하고 해당 범위 사이의 값을 소수점 넷째 자리까지 무작위로 추출하는 것에 기초하여 픽셀에 대한 노출 시간 파라미터(κ)를 조절할 수 있다.

수학식 3

[0089] $$\mathbf{I}(\mathbf{x}) = \begin{cases} \mathbf{M}(\mathbf{x}), & \text{for } \mathbf{x} \in \mathcal{M}, \\ \kappa(\hat{\mathbf{J}}(\mathbf{x}))^{1/\gamma}, & \text{for } \mathbf{x} \notin \mathcal{M}. \end{cases}$$

[0090] 수학식 3에서, ISP를 유리영역 이외의 다른 픽셀에만 적용하고, 합성 이미지 M으로 반사상이 포함되지 않은 이미지 $\hat{\mathbf{J}}$의 유리영역 부분을 대체하여 반사상이 합성된 360도 이미지 I를 획득할 수 있다. 정답 값 영상 J는 하기의 수학식 4와 같이 주어진다.

수학식 4

[0091] $$\mathbf{J}(\mathbf{x}) = \kappa(\hat{\mathbf{J}}(\mathbf{x}))^{1/\gamma}$$

[0092] 반사상 제거 장치의 기계 학습 모델은 합성된 트레이닝 데이터 세트를 이용하여, 합성된 360도 이미지 I로부터 정답 값 영상 J를 매핑(mapping)하도록 트레이닝 된다.

[0093] 입력 영상 중 반사상을 합성하기 위하여 랜덤하게 추출된 영역에 대한 설정 방법은 이하 도9에서 자세히 설명한다.

[0094] 반사상이 제거된 정답 값(ground truth) 360도 영상(805)은 일 실시예에 따른 반사상 제거 장치의 반사상 합성 영상(804)과 함께 반사상 제거 장치의 전체 기계 학습 모델 트레이닝을 위한 데이터 트레이닝 세트로 포함될 수 있다.

[0095] 도 9는 전체 기계 학습 모델을 트레이닝하기 위한 단계(710)에서 미리 설정된 크기의 유리영역을 지정하는 방법에 관해 도시한다.

[0096] 일 실시예에 따르면, 반사상 제거 장치는, 360도 영상을 촬영하는 카메라를 원점(910)으로하고, 원점(910)이 중심인 단위 구(units sphere)(920)를 가정할 수 있다. 반사상 제거 장치는, 입력 영상을 원점(910)이 중심인 단위 구(920)의 표면에 와핑(warping)할 수 있다. 반사상 제거 장치는, 와핑된 영상(930)을 미리 설정된 거리의 평면 레이어(940)에 프로젝션(projection)할 수 있다. 예를 들어, 평면 레이어(940)는 단위 구(920)와 동일한 원점(910)을 포함하면서, 단위 구(920)에 접하는 정육면체의 면에 대응될 수 있다. 단위 구(920)에 와핑된 영상(930)은 원점(910)을 꼭지점으로 하고 평면 레이어(940)를 포함하는 평면을 밑면으로 하는 원뿔을 가정할 때, 원뿔의 꼭지점을 지나고 평면 레이어(940)에 수직인 평면상에 위치하는 원뿔 모선 간의 각도가 90도인 경우에 응답하여 프로젝션 될 수 있다. 반사상 제거 장치는, 프로젝션 된 입력 영상(950)의 가장자리 픽셀 값을 랜덤하게 잘라냄으로써, 입력 영상에서 최종 유리영역을 지정할 수 있다. 다만, 이에 한정하는 것은 아니며, 원점(910)이 중심인 단위 구의 반지름은 최대 2배까지 설정될 수 있다.

[0097] 이상에서 설명된 실시예들은 하드웨어 구성요소, 소프트웨어 구성요소, 및/또는 하드웨어 구성요소 및 소프트웨어 구성요소의 조합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치, 방법 및 구성요소는, 예를 들어, 프로세서, 콘트롤러, ALU(arithmetic logic unit), 디지털 신호 프로세서(digital signal processor), 마이크로컴퓨터, FPGA(field programmable gate array), PLU(programmable logic unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(OS) 및 상기 운영 체제 상에서 수행되는 소프트웨어 어플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리 요소(processing element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알 수 있다. 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 컨트롤러를 포함할 수 있다. 또한, 병렬 프로세서(parallel processor)와 같은, 다른 처리 구성(processing configuration)도 가능하다.

[0098] 소프트웨어는 컴퓨터 프로그램(computer program), 코드(code), 명령(instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성요소(component), 물리적 장치, 가상 장치(virtual equipment), 또는 컴퓨터 저장 매체 또는 장치에 영구적으로, 또는 일시적으로 구체화(embody)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.

[0099] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있으며 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록

매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다.

[0100] 위에서 설명한 하드웨어 장치는 실시예의 동작을 수행하기 위해 하나 또는 복수의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.

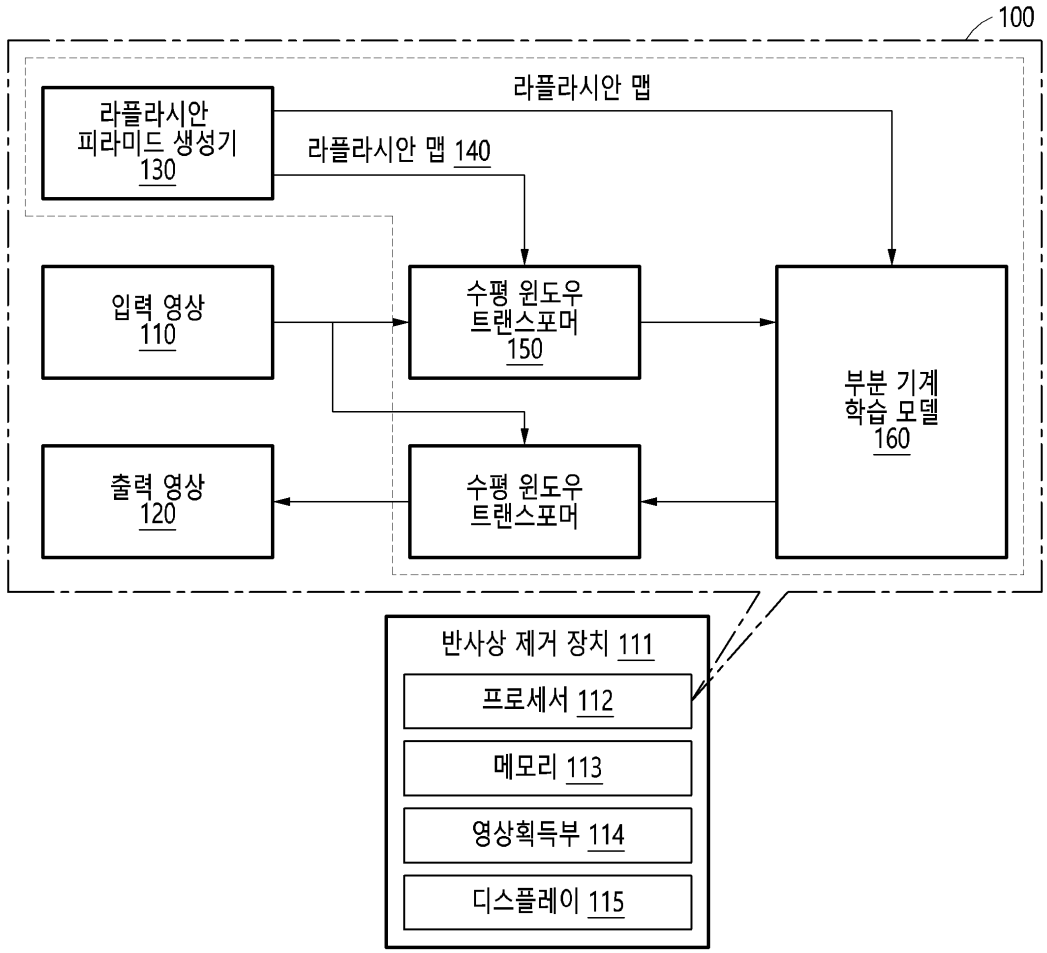
[0101] 본 문서에서, "A 또는 B", "A 및 B 중 적어도 하나", "A 또는 B 중 적어도 하나", "A, B 또는 C", "A, B 및 C 중 적어도 하나", 및 "A, B, 또는 C 중 적어도 하나"와 같은 문구들 각각은 그 문구들 중 해당하는 문구에 함께 나열된 항목들 중 어느 하나, 또는 그들의 모든 가능한 조합을 포함할 수 있다.

[0102] 이상과 같이 실시예들이 비록 한정된 도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 이를 기초로 다양한 기술적 수정 및 변형을 적용할 수 있다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다.

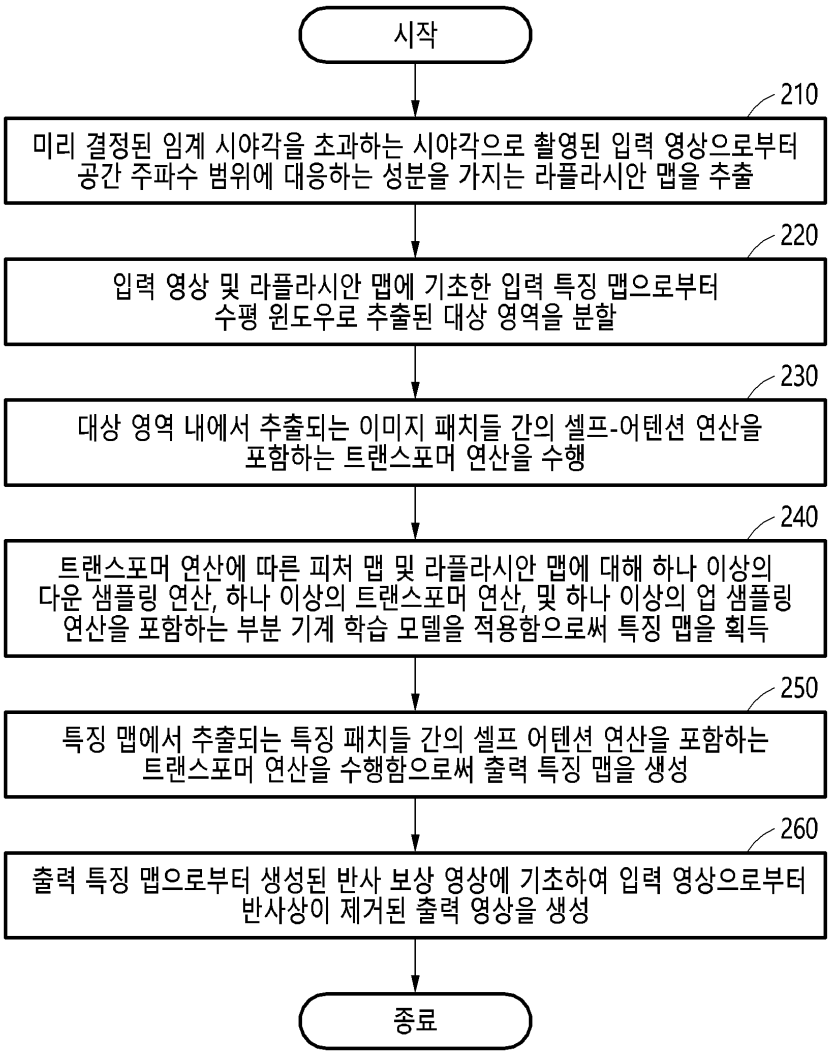
[0103] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

도면

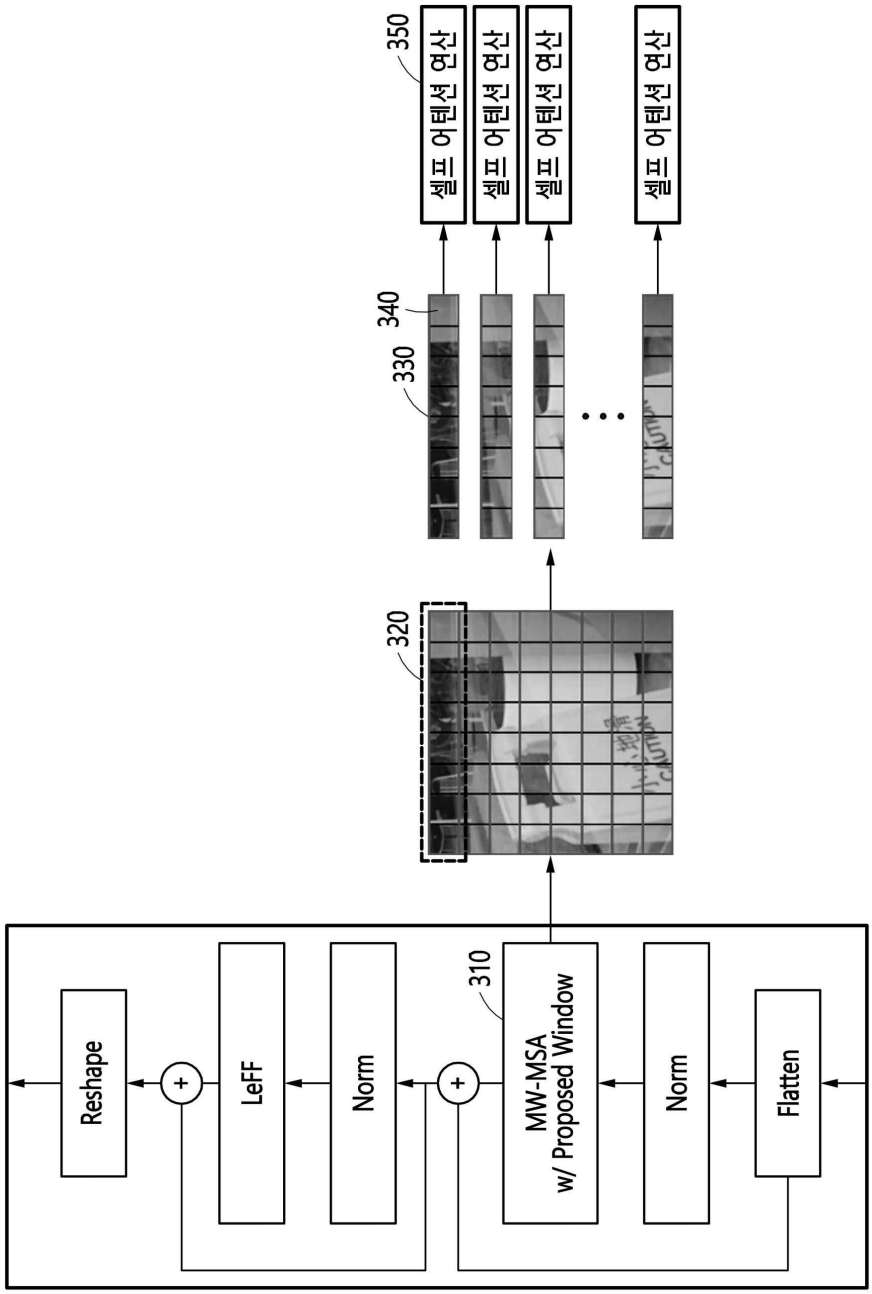
도면1



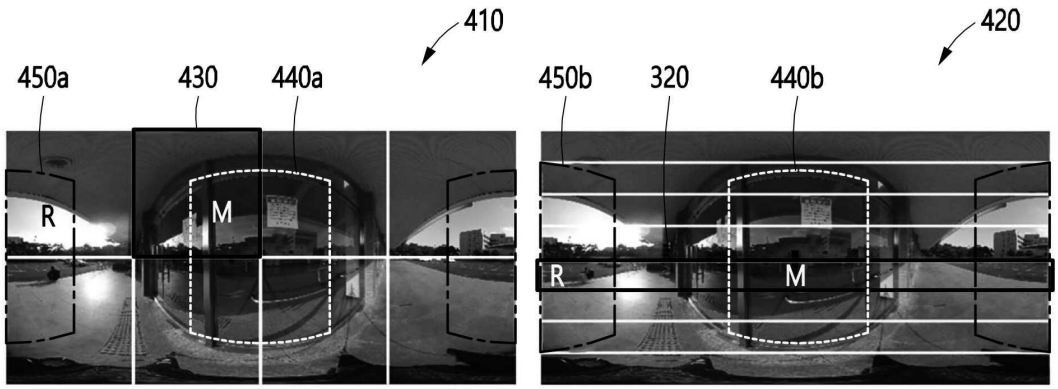
도면2


시작
210
미리 결정된 임계 시야각을 초과하는 시야각으로 촬영된 입력 영상으로부터 공간 주파수 범위에 대응하는 성분을 가지는 라플라시안 맵을 추출
220
입력 영상 및 라플라시안 맵에 기초한 입력 특징 맵으로부터 수평 윈도우로 추출된 대상 영역을 분할
230
대상 영역 내에서 추출되는 이미지 패치들 간의 셀프-어텐션 연산을 포함하는 트랜스포머 연산을 수행
240
트랜스포머 연산에 따른 피처 맵 및 라플라시안 맵에 대해 하나 이상의 다운 샘플링 연산, 하나 이상의 트랜스포머 연산, 및 하나 이상의 업 샘플링 연산을 포함하는 부분 기계 학습 모델을 적용함으로써 특징 맵을 획득
250
특징 맵에서 추출되는 특징 패치들 간의 셀프 어텐션 연산을 포함하는 트랜스포머 연산을 수행함으로써 출력 특징 맵을 생성
260
출력 특징 맵으로부터 생성된 반사 보상 영상에 기초하여 입력 영상으로부터 반사상이 제거된 출력 영상을 생성
종료

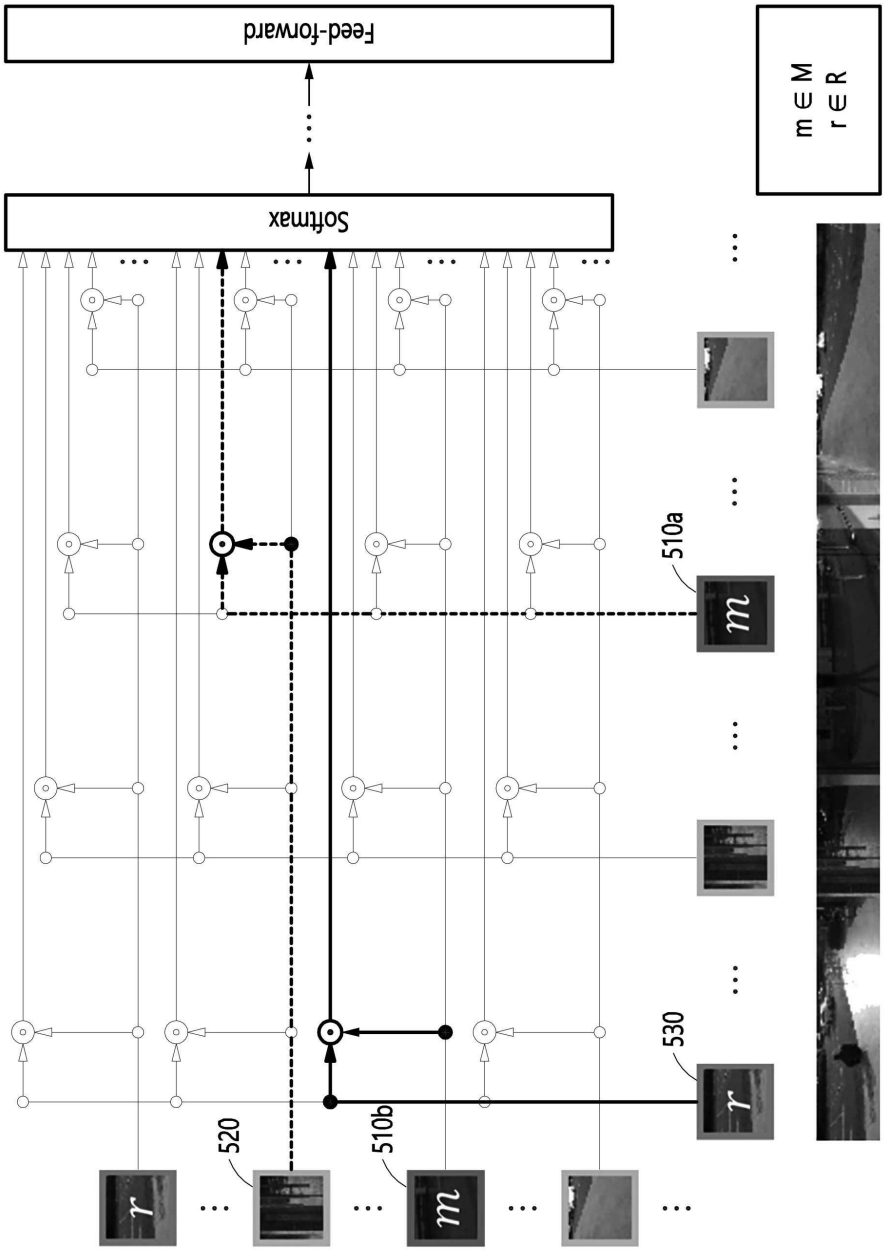
도면3


350
셀프 어텐션 연산
셀프 어텐션 연산
셀프 어텐션 연산
셀프 어텐션 연산
340
330
320
310
MW-MSA
w/ Proposed Window
Norm
Norm
LeFF
Reshape
Flatten

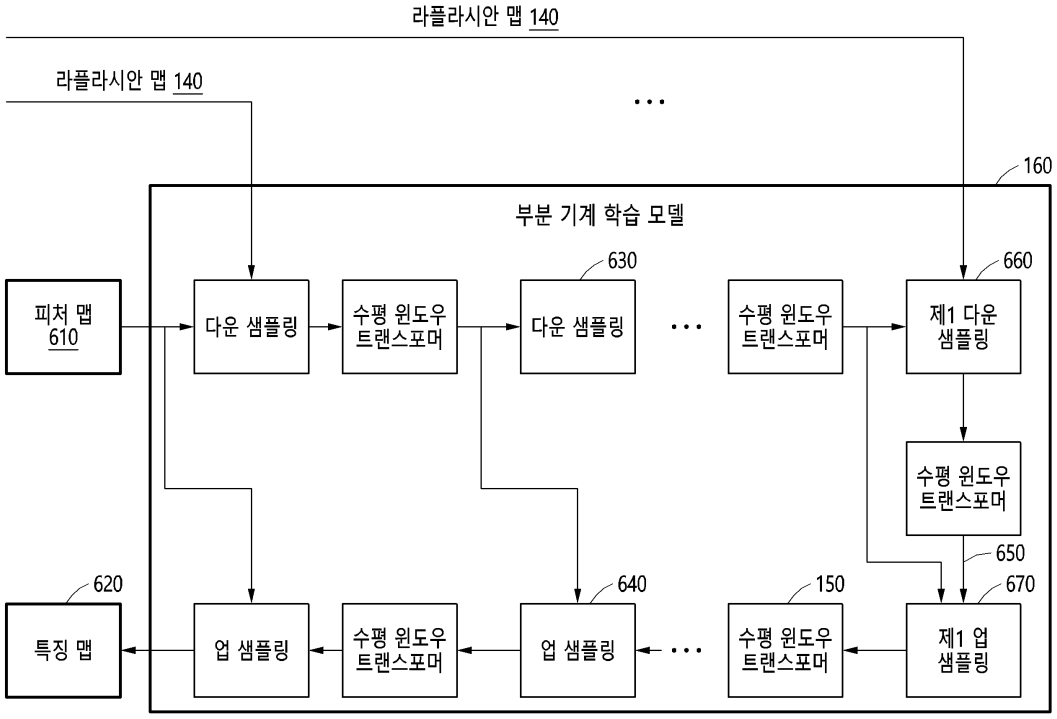
도면4


410
420
450a
430
440a
450b
320
440b
R
M
R
M

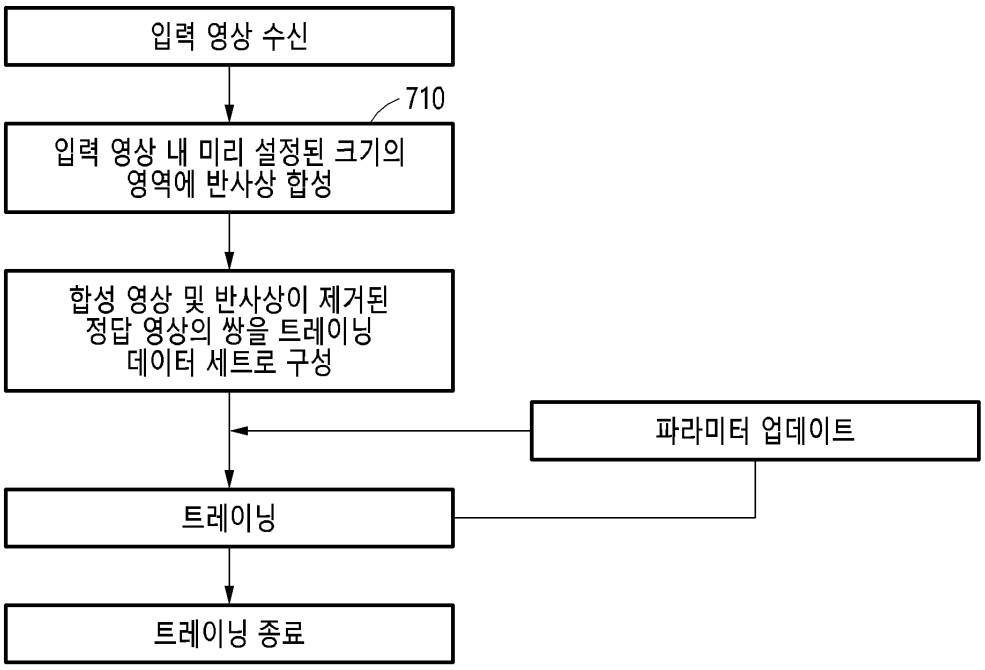
도면5



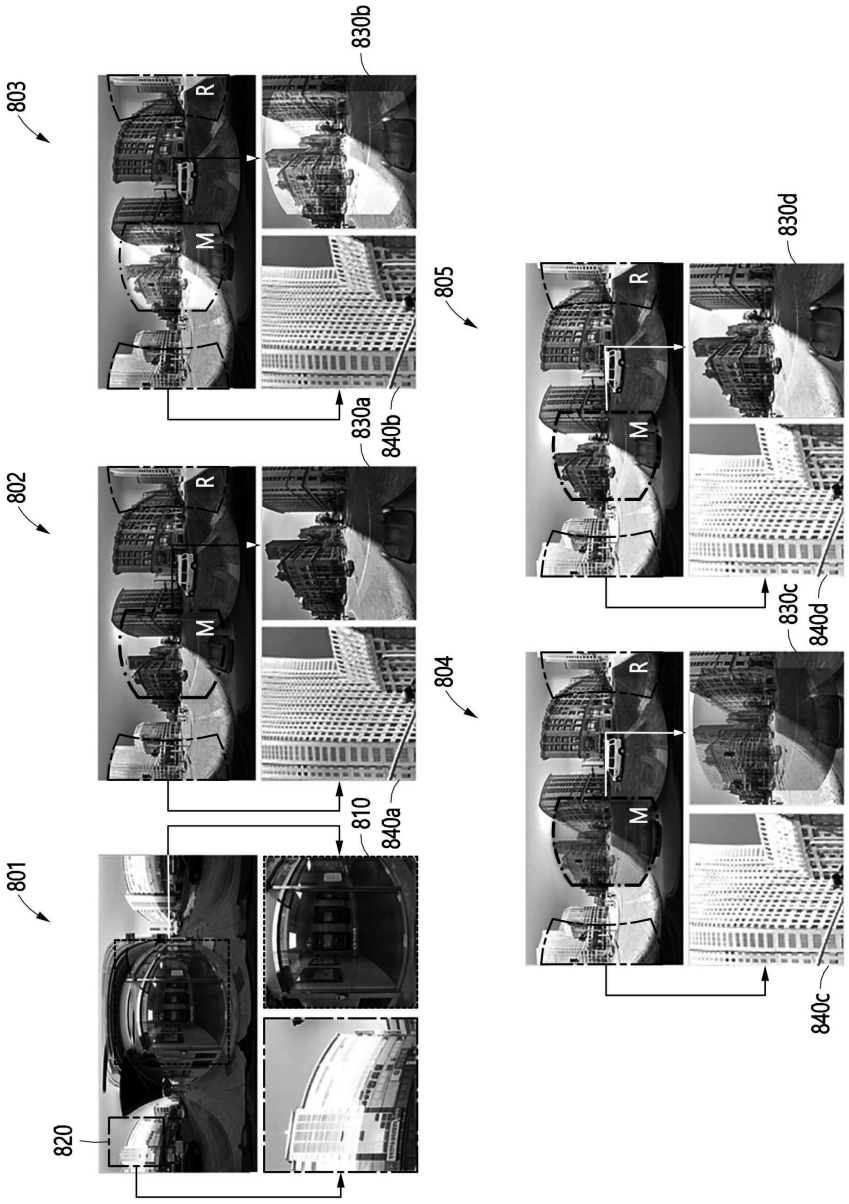
도면6


라플라시안 맵 140
라플라시안 맵 140
160
부분 기계 학습 모델
피처 맵 610
다운 샘플링
수평 윈도우 트랜스포머
630
다운 샘플링
수평 윈도우 트랜스포머
660
제1 다운 샘플링
수평 윈도우 트랜스포머
650
620
640
150
670
특징 맵
업 샘플링
수평 윈도우 트랜스포머
업 샘플링
수평 윈도우 트랜스포머
제1 업 샘플링

도면7


입력 영상 수신
710
입력 영상 내 미리 설정된 크기의 영역에 반사상 합성
합성 영상 및 반사상이 제거된 정답 영상의 쌍을 트레이닝 데이터 세트로 구성
파라미터 업데이트
트레이닝
트레이닝 종료

도면8


801
802
803
804
805
810
820
830a
830b
830c
830d
840a
840b
840c
840d
R
M

도면9

