



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2025-0036544
(43) 공개일자 2025년03월14일

(51) 국제특허분류(Int. Cl.)
G06T 15/50 (2006.01) G06N 3/08 (2023.01)
G06T 5/50 (2024.01) G06V 10/143 (2022.01)
(52) CPC특허분류
G06T 15/50 (2013.01)
G06N 3/08 (2023.01)
(21) 출원번호 10-2023-0119149
(22) 출원일자 2023년09월07일
심사청구일자 없음

(71) 출원인
삼성전자주식회사
경기도 수원시 영통구 삼성로 129 (매탄동)
고려대학교 산학협력단
서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)
(72) 발명자
김중욱
서울특별시 노원구 중계로 233, 108동 1304호(중계동, 청구3차아파트)
한동근
서울특별시 동대문구 무학로45길 32, 402호(용두동)
(74) 대리인
특허법인 무한

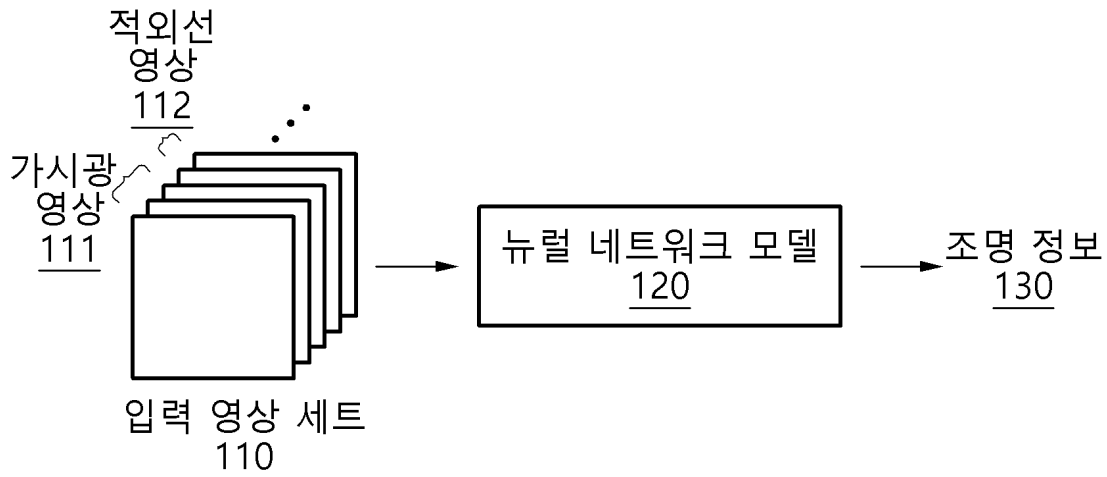
전체 청구항 수 : 총 20 항

(54) 발명의 명칭 뉴럴 네트워크 기반의 영상 처리 방법 및 장치

(57) 요약

뉴럴 네트워크 기반의 영상 처리 방법 및 장치가 제공된다. 그 방법은 입력 영상에 대응하는 가시광 파장 대역의 적어도 하나의 가시광 영상 및 입력 영상에 대응하는 적외선 파장 대역의 적어도 하나의 적외선 영상을 포함하는 입력 영상 세트를 구성하고, 입력 영상 세트의 제1 입력 영상 세트에 기초하여 입력 영상의 조명 구성을 나타내는 조명 맵을 추정하고, 입력 영상 세트의 제2 입력 영상 세트의 스펙트럼 정보에 기초하여 입력 영상의 확산 점수 맵을 추정하고, 조명 맵 및 확산 점수 맵을 융합하여 입력 영상의 조명 정보를 결정하는 단계들을 포함할 수 있다.

대표도 - 도1



(52) CPC특허분류

G06T 5/50 (2024.01)

G06V 10/143 (2023.08)

G06T 2207/20081 (2013.01)

G06T 2207/20081 (2013.01)

명세서

청구범위

청구항 1

입력 영상에 대응하는 가시광 파장 대역의 적어도 하나의 가시광 영상 및 상기 입력 영상에 대응하는 적외선 파장 대역의 적어도 하나의 적외선 영상을 포함하는 입력 영상 세트를 구성하는 단계;

상기 입력 영상 세트의 제1 입력 영상 세트에 기초하여 상기 입력 영상의 조명 구성을 나타내는 조명 맵을 추정하는 단계;

상기 입력 영상 세트의 제2 입력 영상 세트의 스펙트럼 정보에 기초하여 상기 입력 영상의 확산 점수 맵을 추정하는 단계; 및

상기 조명 맵 및 상기 확산 점수 맵을 융합하여 상기 입력 영상의 조명 정보를 결정하는 단계를 포함하는 영상 처리 방법.

청구항 2

제1항에 있어서,

상기 제1 입력 영상 세트는 상기 적어도 하나의 가시광 영상을 포함하고,

상기 제2 입력 영상 세트는 상기 적어도 하나의 적외선 영상을 포함하는,

영상 처리 방법.

청구항 3

제1항에 있어서,

상기 제1 입력 영상 세트 및 상기 제2 입력 영상 세트는 서로 수직인 제1 방향, 제2 방향, 및 제3 방향에 기초하여 정의되고,

상기 조명 맵은 상기 제1 방향 및 상기 제2 방향의 공간 정보에 기초하여 추정되고,

상기 확산 점수 맵은 상기 제3 방향의 상기 스펙트럼 정보에 기초하여 추정되는,

영상 처리 방법.

청구항 4

제1항에 있어서,

상기 제2 입력 영상 세트는 서로 다른 파장 대역들의 멀티밴드 영상들을 포함하고,

상기 스펙트럼 정보는 상기 멀티밴드 영상들의 대응 지점들의 픽셀 값들을 벡터 값들로 갖는 스펙트럼 벡터에 기초하여 결정되는,

영상 처리 방법.

청구항 5

제1항에 있어서,

상기 조명 맵을 추정하는 단계는

제1 뉴럴 네트워크 모델을 이용하여 상기 조명 맵을 추정하는 단계를 포함하고,

상기 확산 점수 맵을 추정하는 단계는

제2 뉴럴 네트워크 모델을 이용하여 상기 확산 점수 맵을 추정하는 단계

를 포함하는, 영상 처리 방법.

청구항 6

제5항에 있어서,

상기 제1 뉴럴 네트워크 모델은

상기 제1 입력 영상 세트로부터 공간 특징을 추출하는 공간 특징 추출 모델, 및

상기 공간 특징에 기초하여 상기 조명 맵을 추정하는 조명 추정 모델

을 포함하는, 영상 처리 방법.

청구항 7

제5항에 있어서,

상기 제2 뉴럴 네트워크 모델은

상기 제2 입력 영상 세트로부터 스펙트럼 특징을 추출하는 스펙트럼 특징 추출 모델, 및

상기 스펙트럼 특징에 기초하여 상기 확산 점수 맵을 추정하는 확산 추정 모델

을 포함하는, 영상 처리 방법.

청구항 8

제7항에 있어서,

상기 스펙트럼 특징 추출 모델은

상기 스펙트럼 정보의 인접 대역의 컨텍스트를 분석하여 상기 제2 입력 영상 세트의 상기 스펙트럼 특징을 결정하는,

영상 처리 방법.

청구항 9

제7항에 있어서,

상기 제2 입력 영상 세트는 서로 다른 파장 대역들의 멀티밴드 영상들을 포함하고,

상기 스펙트럼 특징 추출 모델은

회귀 모델에 기초하여 상기 멀티밴드 영상들의 대응 지점들의 픽셀 값들을 벡터 값들로 갖는 스펙트럼 벡터의 컨텍스트를 분석하고, 분석 결과를 상기 제2 입력 영상 세트에 반영하여 상기 스펙트럼 특징을 결정하는,

영상 처리 방법.

청구항 10

제1항에 있어서,

상기 조명 정보를 결정하는 단계는

상기 조명 맵 및 상기 확산 점수 맵 간의 가중 합을 구하는 단계; 및

상기 가중 합에 따라 상기 조명 정보를 결정하는 단계를 포함하는,

영상 처리 방법.

청구항 11

제10항에 있어서,

상기 가중 합을 구하는 단계는

상기 조명 맵에 따른 로컬 영역 별 벡터들을 상기 확산 점수 맵에 따른 가중치를 이용하여 합산하여 상기 가중합을 구하는 단계를 포함하는,

영상 처리 방법.

청구항 12

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

청구항 13

프로세서; 및

상기 프로세서에서 실행가능한 명령어들을 포함하는 메모리

를 포함하고,

상기 명령어들이 상기 프로세서에서 실행되면, 상기 프로세서는

입력 영상에 대응하는 가시광 파장 대역의 적어도 하나의 가시광 영상 및 상기 입력 영상에 대응하는 적외선 파장 대역의 적어도 하나의 적외선 영상을 포함하는 입력 영상 세트를 구성하고,

상기 입력 영상 세트의 제1 입력 영상 세트에 기초하여 상기 입력 영상의 조명 구성을 나타내는 조명 맵을 추정하고,

상기 입력 영상 세트의 제2 입력 영상 세트의 스펙트럼 정보에 기초하여 상기 입력 영상의 확산 점수 맵을 추정하고,

상기 조명 맵 및 상기 확산 점수 맵을 융합하여 상기 입력 영상의 조명 정보를 결정하는,

영상 처리 장치.

청구항 14

제13항에 있어서,

상기 제1 입력 영상 세트 및 상기 제2 입력 영상 세트는 서로 수직인 제1 방향, 제2 방향, 및 제3 방향에 기초하여 정의되고,

상기 조명 맵은 상기 제1 방향 및 상기 제2 방향의 공간 정보에 기초하여 추정되고,

상기 확산 점수 맵은 상기 제3 방향의 상기 스펙트럼 정보에 기초하여 추정되는,

영상 처리 장치.

청구항 15

제13항에 있어서,

상기 제2 입력 영상 세트는 서로 다른 파장 대역들의 멀티밴드 영상들을 포함하고,

상기 스펙트럼 정보는 상기 멀티밴드 영상들의 대응 지점들의 픽셀 값들을 벡터 값들로 갖는 스펙트럼 벡터에 기초하여 결정되는,

영상 처리 장치.

청구항 16

제13항에 있어서,

상기 프로세서는

제1 뉴럴 네트워크 모델을 이용하여 상기 조명 맵을 추정하고,

제2 뉴럴 네트워크 모델을 이용하여 상기 확산 점수 맵을 추정하는,

영상 처리 장치.

청구항 17

제16항에 있어서,

상기 제2 뉴럴 네트워크 모델은

상기 제2 입력 영상 세트로부터 스펙트럼 특징을 추출하는 스펙트럼 특징 추출 모델, 및

상기 스펙트럼 특징에 기초하여 상기 확산 점수 맵을 추정하는 확산 추정 모델

을 포함하는, 영상 처리 장치.

청구항 18

제17항에 있어서,

상기 스펙트럼 특징 추출 모델은

상기 스펙트럼 정보의 인접 대역의 컨텍스트를 분석하여 상기 제2 입력 영상 세트의 상기 스펙트럼 특징을 결정하는,

영상 처리 장치.

청구항 19

제17항에 있어서,

상기 제2 입력 영상 세트는 서로 다른 파장 대역들의 멀티밴드 영상들을 포함하고,

상기 스펙트럼 특징 추출 모델은

회귀 모델에 기초하여 상기 멀티밴드 영상들의 대응 지점들의 픽셀 값들을 벡터 값들로 갖는 스펙트럼 벡터의 컨텍스트를 분석하고, 분석 결과를 상기 제2 입력 영상 세트에 반영하여 상기 스펙트럼 특징을 결정하는,

영상 처리 장치.

청구항 20

입력 영상을 생성하는 카메라; 및

상기 입력 영상에 대응하는 가시광 파장 대역의 적어도 하나의 가시광 영상 및 상기 입력 영상에 대응하는 적외선 파장 대역의 적어도 하나의 적외선 영상을 포함하는 입력 영상 세트를 구성하고,

상기 입력 영상 세트의 제1 입력 영상 세트에 기초하여 상기 입력 영상의 조명 구성을 나타내는 조명 맵을 추정하고,

상기 입력 영상 세트의 제2 입력 영상 세트의 스펙트럼 정보에 기초하여 상기 입력 영상의 확산 점수 맵을 추정하고,

상기 조명 맵 및 상기 확산 점수 맵을 융합하여 상기 입력 영상의 조명 정보를 결정하는, 프로세서

를 포함하는, 전자 장치.

발명의 설명

기술 분야

[0001] 아래 실시예들은 뉴럴 네트워크 기반의 영상 처리 방법 및 장치에 관한 것이다.

배경 기술

[0002] 영상 처리에는 딥 러닝 기반의 뉴럴 네트워크가 이용될 수 있다. 뉴럴 네트워크는 딥 러닝에 기반하여 트레이닝된 후, 비선형적 관계에 있는 입력 데이터 및 출력 데이터를 서로 매핑함으로써 목적에 맞는 추론(inferenc

e)을 수행해낼 수 있다. 이러한 맵핑을 생성하는 트레이닝된 능력은 신경망의 학습 능력이라 할 수 있다. 더구나, 영상 복원과 같은 특화된 목적을 위해 트레이닝된 신경망은, 예를 들어 트레이닝하지 않은 입력 패턴에 대하여 비교적 정확한 출력을 발생시키는 일반화 능력을 가질 수 있다.

발명의 내용

과제의 해결 수단

[0003] 일 실시예에 따르면, 영상 처리 방법은 입력 영상에 대응하는 가시광 파장 대역의 적어도 하나의 가시광 영상 및 입력 영상에 대응하는 적외선 파장 대역의 적어도 하나의 적외선 영상을 포함하는 입력 영상 세트를 구성하는 단계, 입력 영상 세트의 제1 입력 영상 세트에 기초하여 입력 영상의 조명 구성을 나타내는 조명 맵을 추정하는 단계, 입력 영상 세트의 제2 입력 영상 세트의 스펙트럼 정보에 기초하여 입력 영상의 확산 점수 맵을 추정하는 단계, 및 조명 맵 및 확산 점수 맵을 융합하여 입력 영상의 조명 정보를 결정하는 단계를 포함한다.

[0004] 일 실시예에 따르면, 영상 처리 장치는 프로세서, 및 프로세서에서 실행가능한 명령어들을 포함하는 메모리를 포함하고, 명령어들이 프로세서에서 실행되면, 프로세서는 입력 영상에 대응하는 가시광 파장 대역의 적어도 하나의 가시광 영상 및 입력 영상에 대응하는 적외선 파장 대역의 적어도 하나의 적외선 영상을 포함하는 입력 영상 세트를 구성하고, 입력 영상 세트의 제1 입력 영상 세트에 기초하여 입력 영상의 조명 구성을 나타내는 조명 맵을 추정하고, 입력 영상 세트의 제2 입력 영상 세트의 스펙트럼 정보에 기초하여 입력 영상의 확산 점수 맵을 추정하고, 조명 맵 및 확산 점수 맵을 융합하여 입력 영상의 조명 정보를 결정한다.

[0005] 일 실시예에 따르면, 전자 장치는 입력 영상을 생성하는 카메라, 및 입력 영상에 대응하는 가시광 파장 대역의 적어도 하나의 가시광 영상 및 입력 영상에 대응하는 적외선 파장 대역의 적어도 하나의 적외선 영상을 포함하는 입력 영상 세트를 구성하고, 입력 영상 세트의 제1 입력 영상 세트에 기초하여 입력 영상의 조명 구성을 나타내는 조명 맵을 추정하고, 입력 영상 세트의 제2 입력 영상 세트의 스펙트럼 정보에 기초하여 입력 영상의 확산 점수 맵을 추정하고, 조명 맵 및 확산 점수 맵을 융합하여 입력 영상의 조명 정보를 결정하는, 프로세서를 포함한다.

도면의 간단한 설명

[0006] 도 1은 일 실시예에 따른 뉴럴 네트워크 모델을 이용한 영상 처리 과정을 개략적으로 나타내는 도면이다.
 도 2는 일 실시예에 따른 입력 영상, 가시광 영상들, 및 적외선 영상들을 나타내는 도면이다.
 도 3은 일 실시예에 따른 뉴럴 네트워크 모델의 예시적인 구조를 나타내는 도면이다.
 도 4는 일 실시예에 따른 제1 뉴럴 네트워크 모델의 예시적인 구조를 나타내는 도면이다.
 도 5는 일 실시예에 따른 컨볼루션 모델을 이용한 공간 특징의 생성 과정을 예시적으로 나타내는 도면이다.
 도 6은 일 실시예에 따른 제2 뉴럴 네트워크 모델의 예시적인 구조를 나타내는 도면이다.
 도 7은 일 실시예에 따른 회귀 모델을 이용한 스펙트럼 특징의 생성 과정을 예시적으로 나타내는 도면이다.
 도 8은 일 실시예에 따른 조명 벡터를 이용한 화이트밸런스 동작을 예시적으로 나타내는 도면이다.
 도 9는 일 실시예에 따른 확산 점수 맵 및 조명 맵의 예시적인 구조를 나타내는 도면이다.
 도 10은 일 실시예에 따른 변환 모델들을 이용한 가시광 영상들 및 적외선 영상들의 생성 과정을 예시적으로 나타내는 도면이다.
 도 11은 일 실시예에 따른 영상 처리 방법을 예시적으로 나타내는 플로우 차트이다.
 도 12는 일 실시예에 따른 영상 처리 장치의 구성을 예시적으로 나타내는 블록도이다.
 도 13은 일 실시예에 따른 전자 장치의 구성을 예시적으로 나타내는 블록도이다.

발명을 실시하기 위한 구체적인 내용

[0007] 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.

- [0008] 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.
- [0009] 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.
- [0010] 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.
- [0011] 본 명세서에서, "A 또는 B 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.
- [0012] 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 해당 기술 분야에서 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.
- [0013] 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.
- [0014] 도 1은 일 실시예에 따른 뉴럴 네트워크 모델을 이용한 영상 처리 과정을 개략적으로 나타내는 도면이다. 도 1을 참조하면, 영상 처리 장치는 뉴럴 네트워크 모델(120)을 이용하여 입력 영상 세트(110)의 조명 정보(illuminant information)(130)를 생성할 수 있다. 뉴럴 네트워크 모델(120)은 복수의 레이어들을 포함하는 딥 뉴럴 네트워크(deep neural network, DNN)를 포함할 수 있다. 복수의 레이어들은 입력 레이어(input layer), 적어도 하나의 히든 레이어(hidden layer), 및 출력 레이어(output layer)를 포함할 수 있다.
- [0015] 딥 뉴럴 네트워크는 완전 연결 네트워크(fully connected network, FCN), 컨볼루션 뉴럴 네트워크(convolutional neural network, CNN), 및 리커런트 뉴럴 네트워크(recurrent neural network, RNN) 중 적어도 하나를 포함할 수 있다. 예를 들어, 뉴럴 네트워크 내 복수의 레이어들 중 적어도 일부는 CNN에 해당할 수 있고, 다른 일부는 FCN에 해당할 수 있다. CNN은 컨볼루션 레이어로 지칭될 수 있고, FCN은 완전 연결 레이어로 지칭될 수 있다.
- [0016] CNN의 경우, 각 레이어에 입력되는 데이터는 입력 특징 맵(input feature map)으로 지칭될 수 있고, 각 레이어에서 출력되는 데이터는 출력 특징 맵(output feature map)으로 지칭될 수 있다. 입력 특징 맵 및 출력 특징 맵은 액티베이션 데이터(activation data)로 지칭될 수도 있다. 컨볼루션 레이어가 입력 레이어에 해당하는 경우, 입력 레이어의 입력 특징 맵은 영상일 수 있다.
- [0017] 뉴럴 네트워크는 딥 러닝에 기반하여 트레이닝된 후, 비선형적 관계에 있는 입력 데이터 및 출력 데이터를 서로 매핑함으로써 트레이닝 목적에 맞는 추론(inference)을 수행해낼 수 있다. 딥 러닝은 빅 데이터 세트로부터 영상 또는 음성 인식과 같은 문제를 해결하기 위한 기계 학습 기법이다. 딥 러닝은 준비된 트레이닝 데이터를 이용하여 뉴럴 네트워크를 트레이닝하면서 에너지가 최소화되는 지점을 찾아가는 최적화 문제 풀이 과정으로 이해될 수 있다.
- [0018] 딥 러닝의 지도식(supervised) 또는 비지도식(unsupervised) 학습을 통해 뉴럴 네트워크의 구조, 혹은 모델에 대응하는 웨이트(weight)가 구해질 수 있고, 이러한 웨이트를 통해 입력 데이터 및 출력 데이터가 서로 매핑될 수 있다. 뉴럴 네트워크의 폭과 깊이가 충분히 크면 임의의 함수를 구현할 수 있을 만큼의 용량(capacity)을 가질 수 있다. 뉴럴 네트워크가 적절한 트레이닝 과정을 통해 충분히 많은 양의 트레이닝 데이터를 학습하면 최적의 성능을 달성할 수 있다.
- [0019] 아래에서 뉴럴 네트워크가 '미리' 트레이닝된 것으로 표현될 수 있는데, 여기서 '미리'는 뉴럴 네트워크가 '시작'되기 전에 나타낼 수 있다. 뉴럴 네트워크가 '시작'되었다는 것은 뉴럴 네트워크가 추론을 위한 준비가 된

것을 의미할 수 있다. 예를 들어, 뉴럴 네트워크가 '시작'된 것은 뉴럴 네트워크가 메모리에 로드된 것, 혹은 뉴럴 네트워크가 메모리에 로드된 이후 뉴럴 네트워크에 추론을 위한 입력 데이터가 입력된 것을 포함할 수 있다.

- [0020] 입력 영상 세트(110)는 가시광 영상(visible light image)(111) 및 적외선 영상(infrared image)(112)을 포함할 수 있다. 가시광 영상(111)은 가시광 파장 대역(wavelength band)에 기초하여 생성될 수 있다. 적외선 영상(112)은 적외선 파장 대역에 기초하여 생성될 수 있다. 예를 들어, 적외선 파장 대역은 근적외선(near infrared, NIR) 파장 대역을 포함할 수 있다.
- [0021] 영상 처리 장치는 조명 정보(130)를 추정하는데 가시광 영상(111)뿐만 아니라 적외선 영상(112)도 사용할 수 있다. NIR은 가시광 대역을 벗어나므로 사람의 눈으로는 관측될 수 없으나 가시광 대역과 인접하여 가시광 대역과 일부 반사 특성을 공유할 수 있다. 물체의 표면 재질에 따라 NIR에 따른 반사 특성이 달라질 수 있다. 또한, 적외선 영상(112)은 가시광 영상(111)에 비해 촬영 환경에 강인한 특성을 가질 수 있다.
- [0022] 조명 정보(130)는 입력 영상에 대한 조명의 영향을 나타낼 수 있다. 예를 들어, 입력 영상은 RGB 영상과 같은 가시광 컬러 영상에 해당할 수 있다. 이하, RGB 영상이 입력 영상인 예시를 설명하겠지만, 이에 한정되는 것은 아니다. 조명 정보(130)는 조명 벡터(illuminant vector)를 포함할 수 있다. 입력 영상이 RGB 영상인 경우, 조명 벡터는 R 채널 조명 값, G 채널 조명 값, 및 B 채널 조명 값을 포함할 수 있다.
- [0023] 조명 정보(130)는 조명의 색상을 추정하는 색 항상성(color constancy) 기술에 적용될 수 있다. 예를 들어, 조명 정보(130)를 통해 입력 영상(예: 가시광 영상(111))에 대한 화이트밸런싱(white balancing)이 수행될 수 있다. 조명의 개입으로 인해 카메라로 생성된 영상 내 사물의 색상과 실제 사물의 색상은 서로 다를 수 있다. 색 항상성 기술은 이러한 조명의 영향을 제거할 수 있다. 색 항상성 기술은 영상 전처리(preprocessing) 과정에 속할 수 있다.
- [0024] 도 2는 일 실시예에 따른 입력 영상, 가시광 영상들, 및 적외선 영상들을 나타내는 도면이다. 도 2를 참조하면, 입력 영상(211)은 가시광 파장 대역의 전체 범위(221)에 기초하여 생성될 수 있다. 가시광 파장 대역은 약 400~700nm의 파장 대역을 포함할 수 있다. 가시광 영상들(231, 232, 233)은 가시광 파장 대역의 부분 범위들(241, 242, 243)에 기초하여 생성될 수 있고, 적외선 영상들(234, 235, 236)은 적외선 파장 대역의 부분 범위들(244, 245, 246)에 기초하여 생성될 수 있다. 적외선 파장 대역은 약 700~1000nm의 파장 대역을 포함할 수 있다. 가시광 영상들(231, 232, 233)은 멀티밴드 가시광 영상들, 적외선 영상들(234, 235, 236)은 멀티밴드 적외선 영상들, 부분 범위들(241 내지 246)은 멀티밴드로 부를 수 있다. 가시광 영상들(231, 232, 233) 및 적외선 영상들(234, 235, 236)은 멀티밴드 영상들로 통칭하여 부를 수 있다.
- [0025] 가시광 영상들(231, 232, 233)의 부분 범위들(241, 242, 243)의 폭은 서로 동일할 수 있다. 가시광 영상들(231, 232, 233)의 개수는 M, 가시광 파장 대역의 폭은 B1으로 나타낼 수 있다. 이 경우, 부분 범위들(241, 242, 243)의 폭은 각각 B1/M일 수 있다. 적외선 영상들(234, 235, 236)의 부분 범위들(244, 245, 246)의 폭은 서로 동일할 수 있다. 적외선 영상들(234, 235, 236)의 개수는 N, 적외선 파장 대역의 폭은 B2로 나타낼 수 있다. 이 경우, 부분 범위들(244, 245, 246)의 폭은 각각 B2/N일 수 있다.
- [0026] 입력 영상(211), 가시광 영상들(231, 232, 233), 및 적외선 영상들(234, 235, 236)을 동일한 장면(scene)을 나타낼 수 있다. 일 실시예에 따르면, 가시광 파장 대역의 전체 범위(221)의 영상 정보를 센싱하는 단일 센서에 기초하여 입력 영상(211)이 생성될 수 있고, 입력 영상(211)에 기초하여 가시광 영상들(231, 232, 233) 및 적외선 영상들(234, 235, 236)이 생성될 수 있다. 예를 들어, 가시광 영상들(231, 232, 233) 및 적외선 영상들(234, 235, 236)은 뉴럴 네트워크를 이용하여 생성될 수 있다. 일 실시예에 따르면, 가시광 파장 대역의 영상 정보를 센싱하는 센서에 기초하여 입력 영상(211)이 생성되고, 가시광 파장 대역의 부분 범위들(241, 242, 243)의 영상 정보를 센싱하는 적어도 하나의 센서에 기초하여 가시광 영상들(231, 232, 233)이 생성되고, 적외선 파장 대역의 부분 범위들(244, 245, 246)의 영상 정보를 센싱하는 적어도 하나의 센서에 기초하여 적외선 영상들(234, 235, 236)이 생성될 수 있다.
- [0027] 도 3은 일 실시예에 따른 뉴럴 네트워크 모델의 예시적인 구조를 나타내는 도면이다. 도 3을 참조하면, 영상 처리 장치는 입력 영상 세트(300)에 기초하여 뉴럴 네트워크 모델(310)을 실행하여 입력 영상의 조명 구성을 나타내는 조명 맵(321) 및 입력 영상의 확산 점수 맵(322)을 추정할 수 있다. 영상 처리 장치는 조명 맵(321) 및 확산 점수 맵(322)을 융합하여 입력 영상의 조명 정보를 결정할 수 있다.
- [0028] 일 실시예에 따르면, 영상 처리 장치는 입력 영상에 대응하는 가시광 파장 대역의 적어도 하나의 가시광 영상

및 상기 입력 영상에 대응하는 적외선 파장 대역의 적어도 하나의 적외선 영상을 포함하는 입력 영상 세트(300)를 구성할 수 있다. 영상 처리 장치는 뉴럴 네트워크 기반의 변환 모델을 이용하여 입력 영상에 대응하는 적어도 하나의 가시광 영상 및 입력 영상에 대응하는 적어도 하나의 적외선 영상을 생성할 수 있다.

[0029] 영상 처리 장치는 입력 영상 세트(300)의 제1 입력 영상 세트(301)에 기초하여 입력 영상의 조명 맵(321)을 추정하고, 입력 영상 세트(300)의 제2 입력 영상 세트(302)의 스펙트럼 정보에 기초하여 입력 영상의 확산 점수 맵(322)을 추정할 수 있다. 제1 입력 영상 세트(301)는 적어도 하나의 가시광 영상을 포함하고, 제2 입력 영상 세트(302)는 상기 적어도 하나의 적외선 영상을 포함할 수 있다. 예를 들어, 제1 입력 영상 세트(301)는 멀티밴드 가시광 영상들로 구성될 수 있고, 제2 입력 영상 세트(302)는 멀티밴드 가시광 영상들 및 멀티밴드 적외선 영상들로 구성될 수 있다. 이와 달리, 제1 입력 영상 세트(301)는 가시광 대역의 입력 영상으로 구성될 수 있고, 제2 입력 영상 세트(302)는 멀티밴드 가시광 영상들 및 멀티밴드 적외선 영상들로 구성될 수 있다. 이와 달리, 제1 입력 영상 세트(301)는 가시광 대역의 입력 영상으로 구성될 수 있고, 제2 입력 영상 세트(302)는 멀티밴드 가시광 영상들로 구성될 수 있다.

[0030] 제1 입력 영상 세트(301) 및 제2 입력 영상 세트(302)는 서로 수직인 제1 방향, 제2 방향, 및 제3 방향에 기초하여 정의될 수 있다. 예를 들어, 제1 입력 영상 세트(301) 및 제2 입력 영상 세트(302)는 $H \times W \times C$ 의 디멘전(dimension)을 가질 수 있다. H는 높이(height) 방향의 디멘전, W는 폭(width) 방향의 디멘전, C는 채널(channel) 방향의 디멘전을 나타낼 수 있다. 예를 들어, 높이 방향 및 폭 방향은 제1 방향 및 제2 방향에 해당할 수 있고, 채널 방향은 제3 방향에 해당할 수 있다.

[0031] 조명 맵(321)은 제1 방향 및 제2 방향의 공간 정보(spatial information)에 기초하여 추정될 수 있다. 영상 정보는 높이 방향 및 폭 방향에 기초한 공간 정보에 해당하므로, 제1 방향 및 제2 방향에 기초하여 공간 정보가 분석될 수 있다. 확산 점수 맵(322)은 제3 방향의 스펙트럼 정보(spectral information)에 기초하여 추정될 수 있다.

[0032] 제2 입력 영상 세트(302)는 $H \times W \times M$ 의 멀티밴드 가시광 영상들 및 $H \times W \times N$ 의 멀티밴드 적외선 영상들을 포함할 수 있다. 이 경우, 제2 입력 영상 세트(302)는 $H \times W \times (M+N)$ 으로 나타낼 수 있다. M은 멀티밴드 가시광 영상들의 멀티밴드의 수, N은 멀티밴드 적외선 영상들의 멀티밴드의 수일 수 있다. 제2 입력 영상 세트(302)는 $H \times W$ 의 픽셀들을 포함할 수 있고, 제2 입력 영상 세트(302)의 각 픽셀은 동일 장면에 대한 가시광 파장 대역 및 적외선 파장 대역의 스펙트럼 정보를 포함할 수 있다. 스펙트럼 정보는 $1 \times 1 \times (M+N)$ 의 스펙트럼 벡터로 나타낼 수 있다. 스펙트럼 정보는 채널 방향에 기초하므로, 제3 방향에 기초하여 스펙트럼 정보가 분석될 수 있다. 제2 입력 영상 세트(302)는 서로 다른 파장 대역들의 멀티밴드 영상들을 포함할 수 있다. 스펙트럼 정보는 멀티밴드 영상들의 대응 지점들의 픽셀 값들을 벡터 값들로 갖는 스펙트럼 벡터에 기초하여 결정될 수 있다.

[0033] 조명 맵(321)은 로컬 영역 별 조명 구성을 나타낼 수 있다. 일 실시예에 따르면, 입력 영상 세트(300)는 로컬 영역들을 포함할 수 있다. 예를 들어, 제1 입력 영상 세트(301)는 $H \times W \times C$ 의 디멘전을 가질 수 있고, 조명 맵(321)은 $H/k \times W/k \times C$ 의 디멘전을 가질 수 있다. k는 제1 입력 영상 세트(301)의 2차원 디멘전과 조명 맵(321)의 2차원 디멘전 간의 비율을 나타낼 수 있다. k는 뉴럴 네트워크 모델(311)에 따른 영상 축소(downscale) 비율을 의미할 수 있다. 제1 입력 영상 세트(301)를 채널 방향으로 투사하면, 2차원 영상은 $H \times W$ 의 디멘전을 갖고 $H/k \times W/k$ 의 로컬 영역들을 포함할 수 있다. 조명 맵(321)의 채널 방향의 각 채널 벡터는 2차원 영상의 대응 로컬 영역의 조명 구성을 나타낼 수 있다. 예를 들어, 조명 구성은 조명 벡터를 포함할 수 있다. 조명 벡터는 R 채널 조명 값, G 채널 조명 값, 및 B 채널 조명 값을 포함할 수 있다.

[0034] 확산 점수 맵(322)은 로컬 영역 별로 조명 벡터의 확산 점수를 나타낼 수 있다. 제2 입력 영상 세트(302)의 멀티밴드 영상들은 각각 2차원 영상으로 표현될 수 있고, $H/k \times W/k$ 의 로컬 영역들을 포함할 수 있다. 제1 입력 영상 세트(301)와 제2 입력 영상 세트(302)는 대응 로컬 영역들을 포함할 수 있다. 대응 로컬 영역들은 대응 쌍으로 부를 수 있다. 대응 쌍들은 $H/k \times W/k$ 개 존재할 수 있다. 확산 점수 맵(322)은 $H/k \times W/k$ 의 디멘전을 가질 수 있다. 확산 점수 맵(322)은 각 대응 쌍의 조명 벡터의 각 확산 점수를 포함할 수 있다.

[0035] 뉴럴 네트워크 모델(310)은 제1 뉴럴 네트워크 모델(311) 및 제2 뉴럴 네트워크 모델(312)을 포함할 수 있다. 영상 처리 장치는 제1 뉴럴 네트워크 모델(311)을 이용하여 조명 맵(321)을 추정할 수 있고, 제2 뉴럴 네트워크 모델(312)을 이용하여 확산 점수 맵(322)을 추정할 수 있다. 영상 처리 장치는 조명 맵(321) 및 확산 점수 맵(322)을 융합하여 입력 영상의 조명 정보를 결정할 수 있다.

[0036] 도 4는 일 실시예에 따른 제1 뉴럴 네트워크 모델의 예시적인 구조를 나타내는 도면이다. 도 4를 참조하면, 제

1 뉴럴 네트워크 모델(400)은 제1 입력 영상 세트(401)에 기초하여 조명 맵(421)을 추정할 수 있다. 제1 뉴럴 네트워크 모델(400)은 제1 입력 영상 세트(401)로부터 공간 특징을 추출하는 공간 특징 추출 모델(410) 및 공간 특징에 기초하여 조명 맵(421)을 추정하는 조명 추정 모델(420)을 포함할 수 있다. 공간 특징 추출 모델(410) 및 조명 추정 모델(420)은 뉴럴 네트워크에 기반할 수 있다.

[0037] 도 5는 일 실시예에 따른 컨볼루션 모델을 이용한 공간 특징의 생성 과정을 예시적으로 나타내는 도면이다. 도 5를 참조하면, 공간 특징 추출 모델은 제1 입력 영상 세트(501)로부터 제1 중간 특징(511)을 추출하고, 제1 중간 특징(511)으로부터 제2 중간 특징(512)을 추출하고, 제2 중간 특징(512)으로부터 제3 중간 특징(513)을 추출할 수 있다. 공간 특징 추출 모델은 복수의 컨볼루션 레이어들을 포함할 수 있고, 제1 중간 특징(511), 제2 중간 특징(512), 및 제3 중간 특징(513)의 추출에 복수의 컨볼루션 레이어들이 이용될 수 있다. 제1 입력 영상 세트(501), 제1 중간 특징(511), 제2 중간 특징(512), 및 제3 중간 특징(513)은 동일한 채널 개수를 가질 수 있다. 도 5는 3개의 컨볼루션 레이어들이 이용되는 예시를 나타내지만, 이에 제한되는 것은 아니다.

[0038] 제1 중간 특징(511), 제2 중간 특징(512), 및 제3 중간 특징(513)은 제4 중간 특징(514)으로 병합될 수 있다. 예를 들어, 제1 중간 특징(511)의 평균화를 통해 제4 중간 특징(514)의 하나의 채널 데이터가 형성될 수 있고, 제2 중간 특징(512)의 평균화를 통해 제4 중간 특징(514)의 다른 하나의 채널 데이터가 형성될 수 있고, 제3 중간 특징(513)의 평균화를 통해 제4 중간 특징(514)의 또 다른 하나의 채널 데이터가 형성될 수 있다. 공간 특징 추출 모델은 제4 중간 특징(514)에 관한 컨볼루션을 수행하여 제1 입력 영상 세트(501)의 공간 특징(521)을 결정할 수 있다.

[0039] 도 6은 일 실시예에 따른 제2 뉴럴 네트워크 모델의 예시적인 구조를 나타내는 도면이다. 도 6을 참조하면, 제2 뉴럴 네트워크 모델(600)은 제2 입력 영상 세트(601)에 기초하여 확산 점수 맵(621)을 추정할 수 있다. 제2 뉴럴 네트워크 모델(600)은 제2 입력 영상 세트(601)로부터 스펙트럼 특징을 추출하는 스펙트럼 특징 추출 모델(610) 및 스펙트럼 특징에 기초하여 확산 점수 맵(621)을 추정하는 확산 추정 모델(620)을 포함할 수 있다. 일 실시예에 따르면, 공간 특징 추출 모델에 의해 제1 입력 영상 세트의 공간 특징이 추출되면, 공간 특징 및 스펙트럼 특징은 연쇄화(concatenation)될 수 있고, 확산 추정 모델(620)은 연쇄화 결과물에 기초하여 확산 점수 맵(621)을 추정할 수 있다. 스펙트럼 특징 추출 모델(610) 및 확산 추정 모델(620)은 뉴럴 네트워크에 기반할 수 있다.

[0040] 도 7은 일 실시예에 따른 회귀 모델을 이용한 스펙트럼 특징의 생성 과정을 예시적으로 나타내는 도면이다. 도 7을 참조하면, 제2 입력 영상 세트(710)는 $H \times W \times (M+N)$ 의 디멘전을 가질 수 있다. 제2 입력 영상 세트(710)로부터 스펙트럼 벡터들(711, 712, 713)이 추출될 수 있다. 스펙트럼 벡터들(711, 712, 713)은 $1 \times 1 \times (M+N)$ 의 디멘전을 가질 수 있다. 스펙트럼 특징 추출 모델은 회귀 모델(recurrent model)을 이용하여 스펙트럼 벡터들(711, 712, 713)을 분석할 수 있다. 스펙트럼 특징 추출 모델은 회귀 모델에 기초하여 제2 입력 영상 세트(710)의 멀티밴드 영상들의 대응 지점들의 픽셀 값들을 벡터 값들로 갖는 스펙트럼 벡터들(711, 712, 713)의 컨텍스트를 분석하고, 분석 결과를 제2 입력 영상 세트(710)에 반영하여 스펙트럼 특징(730)을 결정할 수 있다. 분석 결과는 스펙트럼 벡터들(711, 712, 713)과 동일 디멘전을 갖는 분석 결과 벡터에 해당할 수 있다. 분석 결과 벡터들과 스펙트럼 벡터들(711, 712, 713) 간의 곱셈 결과가 스펙트럼 벡터들(711, 712, 713)을 대체하거나, 분석 결과 벡터들이 스펙트럼 벡터들(711, 712, 713)을 대체하여, 스펙트럼 특징(730)이 결정될 수 있다. 스펙트럼 벡터들(711, 712, 713)은 제2 입력 영상 세트(710)의 임의의 지점들로부터 추출될 수 있다. 도 7에는 3개의 스펙트럼 벡터들(711, 712, 713)이 이용되는 예시가 도시되어 있으나, 이에 제한되지 않는다.

[0041] 스펙트럼 특징 추출 모델은 스펙트럼 정보의 인접 대역의 컨텍스트를 분석하여 제2 입력 영상 세트(710)의 스펙트럼 특징을 결정할 수 있다. 제2 입력 영상 세트(710)의 멀티밴드 영상들의 인접 대역 영상들은 피사체의 특징(예: 재질, 반사 특성 등)에 따른 컨텍스트를 가질 수 있다. 회귀 모델은 이러한 컨텍스트를 분석하는데 적합할 수 있다. 회귀 모델을 이용하여 스펙트럼 정보의 컨텍스트가 분석됨에 따라 현실에 가까운 조명 정보가 도출될 수 있다. 회귀 모델은 뉴럴 네트워크에 기반할 수 있다. 예를 들어, 회귀 모델은 LSTM(long short term memory)을 포함할 수 있다. 회귀 모델은 도 7의 예시와 같이 복수의 모델들로 구현되거나, 도 7의 예시와 달리 하나의 모델로 구현될 수 있다.

[0042] 도 8은 일 실시예에 따른 조명 벡터를 이용한 화이트밸런스 동작을 예시적으로 나타내는 도면이다. 도 8을 참조하면, 영상 처리 장치는 확산 점수 맵(801)과 조명 맵(802) 간의 가중 합을 구하고, 가중 합에 따라 조명 정보를 결정할 수 있다. 예를 들어, 조명 정보는 조명 벡터(803)를 포함할 수 있다. 조명 벡터는 R 채널 조명 값, G 채널 조명 값, 및 B 채널 조명 값을 포함할 수 있다. 손실 함수(loss function)로 아래 수학적 식 1이 이용

될 수 있다.

수학식 1

$$L = \cos^{-1} \left(\frac{\Gamma_{gt} \cdot \Gamma_{est}}{\|\Gamma_{gt}\| \|\Gamma_{est}\|} \right)$$

[0043]

[0044]

수학식 1에서 L 은 손실을, Γ_{est} 는 조명 벡터를, Γ_{gt} 는 GT(ground truth)를 나타낼 수 있다. 손실(L)은 조명 벡터(Γ_{est})와 GT(Γ_{gt}) 간의 각도 오차(angular error)를 나타낼 수 있다. 조명 벡터(803)는 가시광 영상(804)에 대한 화이트밸런스 처리에 이용될 수 있다. 처리 결과(805)에 기초하여 출력 영상이 결정될 수 있다.

[0045]

도 9는 일 실시예에 따른 확신 점수 맵 및 조명 맵의 예시적인 구조를 나타내는 도면이다. 도 9를 참조하면, 확신 점수 맵(910)의 디멘전은 $H/k*W/k$ 일 수 있다. 확신 점수 맵(910)은 로컬 영역(901)과 같은 로컬 영역들을 포함할 수 있다. 로컬 영역들의 수는 $H/k*W/k$ 개일 수 있다. 확신 점수 맵(910)은 확신 점수 값(911)과 같은 확신 점수 값들을 포함할 수 있다. 확신 점수 값들의 수는 $H/k*W/k$ 개일 수 있다. 조명 맵(920)의 디멘전은 $H/k*W/k*c1$ 일 수 있다. 조명 맵(920)은 로컬 영역(902)과 같은 로컬 영역들을 포함할 수 있다. 로컬 영역들의 수는 $H/k*W/k$ 개일 수 있다. 대응 위치의 로컬 영역들(901, 902)은 연산 쌍을 구성할 수 있다. 조명 맵(920)은 조명 벡터(921)와 같은 조명 벡터들을 포함할 수 있다. 조명 벡터는 R 채널 조명 값, G 채널 조명 값, 및 B 채널 조명 값을 포함할 수 있다. 조명 벡터들의 수는 $H/k*W/k$ 개일 수 있다.

[0046]

영상 처리 장치는 조명 맵(920)에 따른 로컬 영역 별 조명 벡터들을 확신 점수 맵(910)에 따른 가중치를 이용하여 합산하여 가중 합을 구할 수 있다. 가시광 영상의 로컬 영역들과 적외선 영상의 로컬 영역들은 대응 쌍들을 구성할 수 있다. 확신 점수 맵(910)의 확신 점수 값들은 대응 쌍들에 대한 가중치들을 결정할 수 있다. 스펙트럼 분석 결과에 따라 대응 쌍들의 확신 점수 및 가중치가 결정될 수 있다. 조명 맵(920)과 확신 점수 맵(910) 간의 가중 합을 통해 $1*1*c1$ 의 조명 벡터가 결정될 수 있다. 영상의 로컬 영역에 따라 조명 추정치의 정확도가 다르기 때문에 뉴럴 네트워크 모델은 각 영역별 조명의 신뢰도를 함께 학습할 수 있다. 이러한 각 영역별 조명의 신뢰도에 따른 가중 합을 통해 하나의 조명 벡터가 결정될 수 있다. 조명 벡터는 입력 영상(예: 가시광 영상)의 화이트 밸런싱에 이용될 수 있다.

[0047]

도 10은 일 실시예에 따른 변환 모델들을 이용한 가시광 영상들 및 적외선 영상들의 생성 과정을 예시적으로 나타내는 도면이다. 도 10을 참조하면, 제1 변환 모델(1020)은 입력 영상(1010)에 따른 멀티밴드 가시광 영상들(1021)을 추정할 수 있다. 제2 변환 모델(1030)은 멀티밴드 가시광 영상들(1021)에 따른 멀티밴드 적외선 영상들(1031)을 추정할 수 있다. 제1 변환 모델(1020) 및 제2 변환 모델(1030)은 뉴럴 네트워크에 기반할 수 있다.

[0048]

도 11은 일 실시예에 따른 영상 처리 방법을 예시적으로 나타내는 플로우 차트이다. 도 11을 참조하면, 영상 처리 장치는 단계(1110)에서 입력 영상에 대응하는 가시광 파장 대역의 적어도 하나의 가시광 영상 및 입력 영상에 대응하는 적외선 파장 대역의 적어도 하나의 적외선 영상을 포함하는 입력 영상 세트를 구성하고, 단계(1120)에서 입력 영상 세트의 제1 입력 영상 세트에 기초하여 입력 영상의 조명 구성을 나타내는 조명 맵을 추정하고, 단계(1130)에서 입력 영상 세트의 제2 입력 영상 세트의 스펙트럼 정보에 기초하여 입력 영상의 확신 점수 맵을 추정하고, 단계(1140)에서 조명 맵 및 확신 점수 맵을 융합하여 입력 영상의 조명 정보를 결정할 수 있다.

[0049]

제1 입력 영상 세트는 적어도 하나의 가시광 영상을 포함하고, 제2 입력 영상 세트는 적어도 하나의 적외선 영상을 포함할 수 있다.

[0050]

제1 입력 영상 세트 및 제2 입력 영상 세트는 서로 수직인 제1 방향, 제2 방향, 및 제3 방향에 기초하여 정의되고, 조명 맵은 제1 방향 및 제2 방향의 공간 정보에 기초하여 추정되고, 확신 점수 맵은 제3 방향의 스펙트럼 정보에 기초하여 추정될 수 있다.

[0051]

제2 입력 영상 세트는 서로 다른 파장 대역들의 멀티밴드 영상들을 포함하고, 스펙트럼 정보는 멀티밴드 영상들의 대응 지점들의 픽셀 값들을 벡터 값들로 갖는 스펙트럼 벡터에 기초하여 결정될 수 있다.

[0052]

단계(1120)는 제1 뉴럴 네트워크 모델을 이용하여 조명 맵을 추정하는 단계를 포함하고, 단계(1130)는 제2 뉴럴

네트워크 모델을 이용하여 확신 점수 맵을 추정하는 단계를 포함할 수 있다.

- [0053] 제1 뉴럴 네트워크 모델은 제1 입력 영상 세트로부터 공간 특징을 추출하는 공간 특징 추출 모델, 및 공간 특징에 기초하여 조명 맵을 추정하는 조명 추정 모델을 포함할 수 있다.
- [0054] 제2 뉴럴 네트워크 모델은 제2 입력 영상 세트로부터 스펙트럼 특징을 추출하는 스펙트럼 특징 추출 모델, 및 스펙트럼 특징에 기초하여 확신 점수 맵을 추정하는 확신 추정 모델을 포함할 수 있다.
- [0055] 스펙트럼 특징 추출 모델은 스펙트럼 정보의 인접 대역의 컨텍스트를 분석하여 제2 입력 영상 세트의 스펙트럼 특징을 결정할 수 있다.
- [0056] 제2 입력 영상 세트는 서로 다른 파장 대역들의 멀티밴드 영상들을 포함하고, 스펙트럼 특징 추출 모델은 회귀 모델에 기초하여 멀티밴드 영상들의 대응 지점들의 픽셀 값들을 벡터 값들로 갖는 스펙트럼 벡터의 컨텍스트를 분석하고, 분석 결과를 제2 입력 영상 세트에 반영하여 스펙트럼 특징을 결정할 수 있다.
- [0057] 단계(1140)는 조명 맵 및 확신 점수 맵 간의 가중 합을 구하는 단계, 및 가중 합에 따라 조명 정보를 결정하는 단계를 포함할 수 있다.
- [0058] 가중 합을 구하는 단계는 조명 맵에 따른 로컬 영역 별 벡터들을 확신 점수 맵에 따른 가중치를 이용하여 합산하여 가중 합을 구하는 단계를 포함할 수 있다.
- [0059] 그 밖에, 도 11의 영상 처리 방법에 도 1 내지 도 10, 도 12, 및 도 13의 설명이 적용될 수 있다.
- [0060] 도 12는 일 실시예에 따른 영상 처리 장치의 구성을 예시적으로 나타내는 블록도이다. 도 12를 참조하면, 영상 처리 장치(1200)는 프로세서(1210) 및 메모리(1220)를 포함한다. 메모리(1220)는 프로세서(1210)에 연결되고, 프로세서(1210)에 의해 실행가능한 명령어들, 프로세서(1210)가 연산할 데이터 또는 프로세서(1210)에 의해 처리된 데이터를 저장할 수 있다. 메모리(1220)는 비일시적인 컴퓨터 판독가능 매체, 예컨대 고속 랜덤 액세스 메모리 및/또는 비휘발성 컴퓨터 판독가능 저장 매체(예컨대, 하나 이상의 디스크 저장 장치, 플래쉬 메모리 장치, 또는 기타 비휘발성 솔리드 스테이트 메모리 장치)를 포함할 수 있다.
- [0061] 프로세서(1210)는 도 1 내지 도 11, 및 도 13의 동작을 수행하기 위한 명령어들을 실행할 수 있다. 예를 들어, 프로세서(1210)는 입력 영상에 대응하는 가시광 파장 대역의 적어도 하나의 가시광 영상 및 입력 영상에 대응하는 적외선 파장 대역의 적어도 하나의 적외선 영상을 포함하는 입력 영상 세트를 구성하고, 입력 영상 세트의 제1 입력 영상 세트에 기초하여 입력 영상의 조명 구성을 나타내는 조명 맵을 추정하고, 입력 영상 세트의 제2 입력 영상 세트의 스펙트럼 정보에 기초하여 입력 영상의 확신 점수 맵을 추정하고, 조명 맵 및 확신 점수 맵을 융합하여 입력 영상의 조명 정보를 결정할 수 있다. 그 밖에, 영상 처리 장치(1200)에는 도 1 내지 도 11, 및 도 13의 설명이 적용될 수 있다.
- [0062] 도 13은 일 실시예에 따른 전자 장치의 구성을 예시적으로 나타내는 블록도이다. 도 13을 참조하면, 전자 장치(1300)는 프로세서(1310), 메모리(1320), 카메라(1330), 저장 장치(1340), 입력 장치(1350), 출력 장치(1360) 및 네트워크 인터페이스(1370)를 포함할 수 있으며, 이들은 통신 버스(1380)를 통해 서로 통신할 수 있다. 예를 들어, 전자 장치(1300)는 이동 전화, 스마트 폰, PDA, 넷북, 태블릿 컴퓨터, 랩톱 컴퓨터 등과 같은 모바일 장치, 스마트 워치, 스마트 밴드, 스마트 안경 등과 같은 웨어러블 디바이스, 데스크탑, 서버 등과 같은 컴퓨팅 장치, 텔레비전, 스마트 텔레비전, 냉장고 등과 같은 가전 제품, 도어 락 등과 같은 보안 장치, 자율주행 차량, 스마트 차량 등과 같은 차량의 적어도 일부로 구현될 수 있다. 전자 장치(1300)는 도 12의 영상 처리 장치(1200)를 구조적 및/또는 기능적으로 포함할 수 있다.
- [0063] 프로세서(1310)는 전자 장치(1300) 내에서 실행하기 위한 기능 및 명령어들을 실행한다. 예를 들어, 프로세서(1310)는 메모리(1320) 또는 저장 장치(1340)에 저장된 명령어들을 처리할 수 있다. 프로세서(1310)는 도 1 내지 도 12를 통하여 설명된 하나 이상의 동작을 수행할 수 있다. 메모리(1320)는 컴퓨터 판독가능한 저장 매체 또는 컴퓨터 판독가능한 저장 장치를 포함할 수 있다. 메모리(1320)는 프로세서(1310)에 의해 실행하기 위한 명령어들을 저장할 수 있고, 전자 장치(1300)에 의해 소프트웨어 및/또는 애플리케이션이 실행되는 동안 관련 정보를 저장할 수 있다.
- [0064] 카메라(1330)는 입력 영상 및/또는 입력 영상 세트를 생성할 수 있다. 입력 영상은 사진 및/또는 비디오를 포함할 수 있다. 카메라(1330)는 가시광 영상을 생성하는 가시광 카메라, 및 적외선 영상을 생성하는 적외선 카메라를 포함할 수 있다. 가시광 영상 및 적외선 영상은 입력 영상 세트를 구성할 수 있다. 저장 장치(1340)는 컴퓨터 판독가능한 저장 매체 또는 컴퓨터 판독가능한 저장 장치를 포함한다. 저장 장치(1340)는 메모리(132

0)보다 더 많은 양의 정보를 저장하고, 정보를 장기간 저장할 수 있다. 예를 들어, 저장 장치(1340)는 자기 하드 디스크, 광 디스크, 플래쉬 메모리, 플로피 디스크 또는 이 기술 분야에서 알려진 다른 형태의 비휘발성 메모리를 포함할 수 있다.

[0065] 입력 장치(1350)는 키보드 및 마우스를 통한 전통적인 입력 방식, 및 터치 입력, 음성 입력, 및 이미지 입력과 같은 새로운 입력 방식을 통해 사용자로부터 입력을 수신할 수 있다. 예를 들어, 입력 장치(1350)는 키보드, 마우스, 터치 스크린, 마이크로폰, 또는 사용자로부터 입력을 검출하고, 검출된 입력을 전자 장치(1300)에 전달할 수 있는 임의의 다른 장치를 포함할 수 있다. 출력 장치(1360)는 시각적, 청각적 또는 촉각적인 채널을 통해 사용자에게 전자 장치(1300)의 출력을 제공할 수 있다. 출력 장치(1360)는 예를 들어, 디스플레이, 터치 스크린, 스피커, 진동 발생 장치 또는 사용자에게 출력을 제공할 수 있는 임의의 다른 장치를 포함할 수 있다. 네트워크 인터페이스(1370)는 유선 또는 무선 네트워크를 통해 외부 장치와 통신할 수 있다.

[0066] 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 애플리케이션을 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.

[0067] 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), 컴퓨터 저장 매체 또는 장치, 또는 전송되는 신호 파(signal wave)에 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.

[0068] 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 저장할 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.

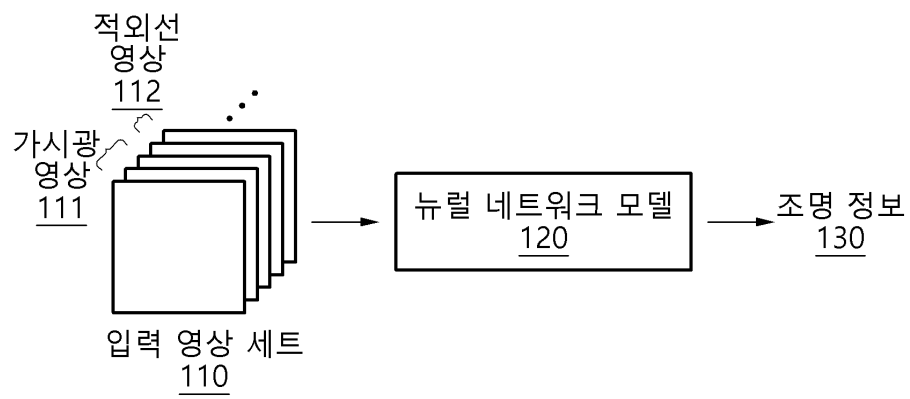
[0069] 위에서 설명한 하드웨어 장치는 실시예의 동작을 수행하기 위해 하나 또는 복수의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.

[0070] 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.

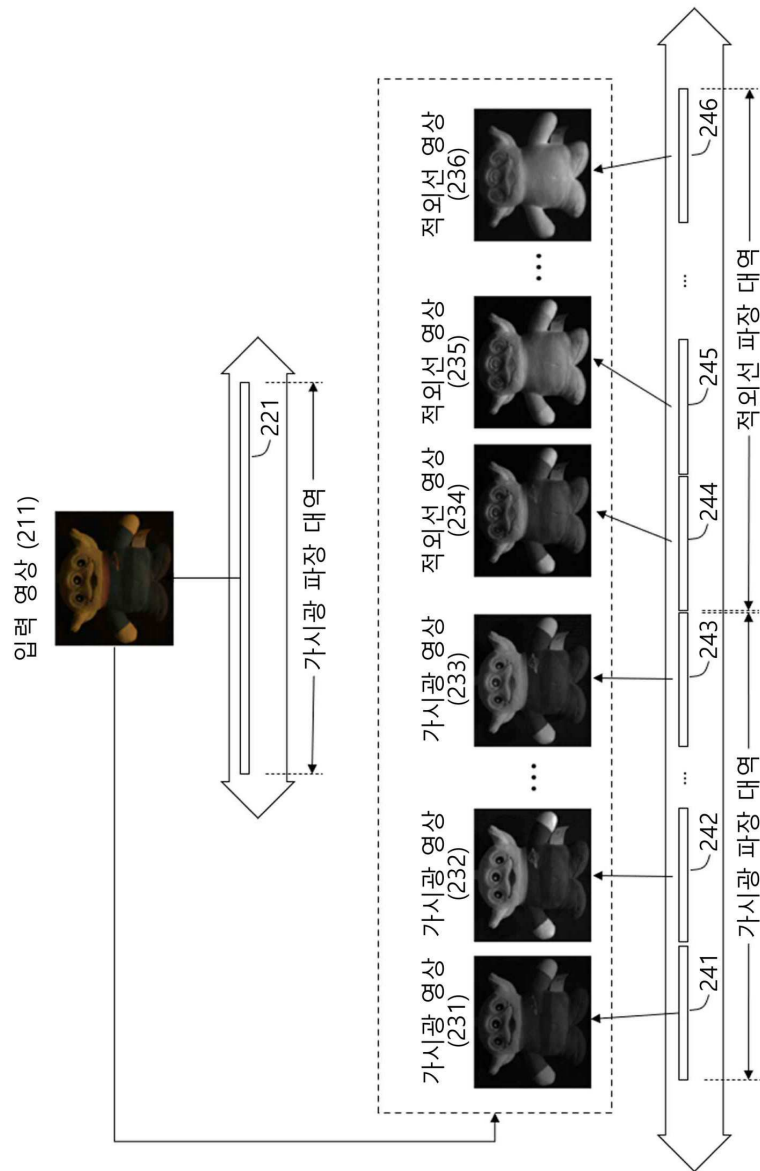
[0071] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

도면

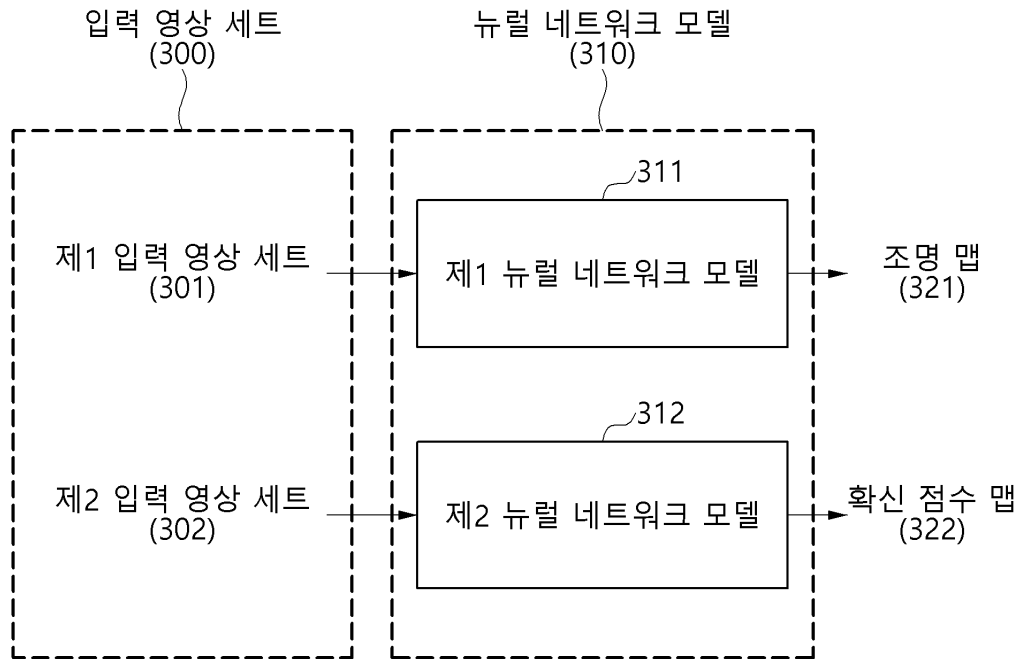
도면1



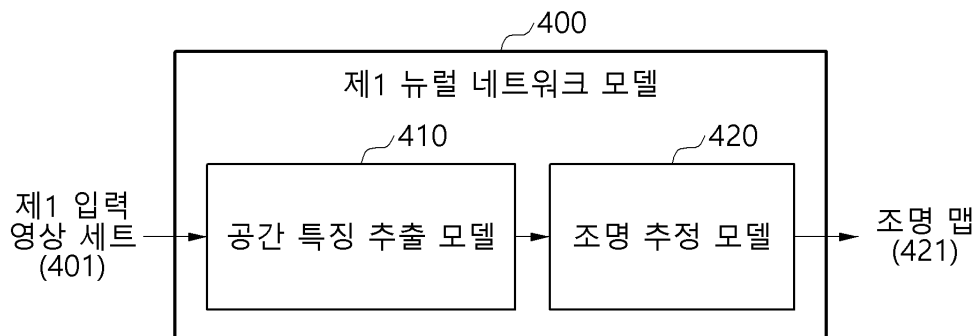
도면2



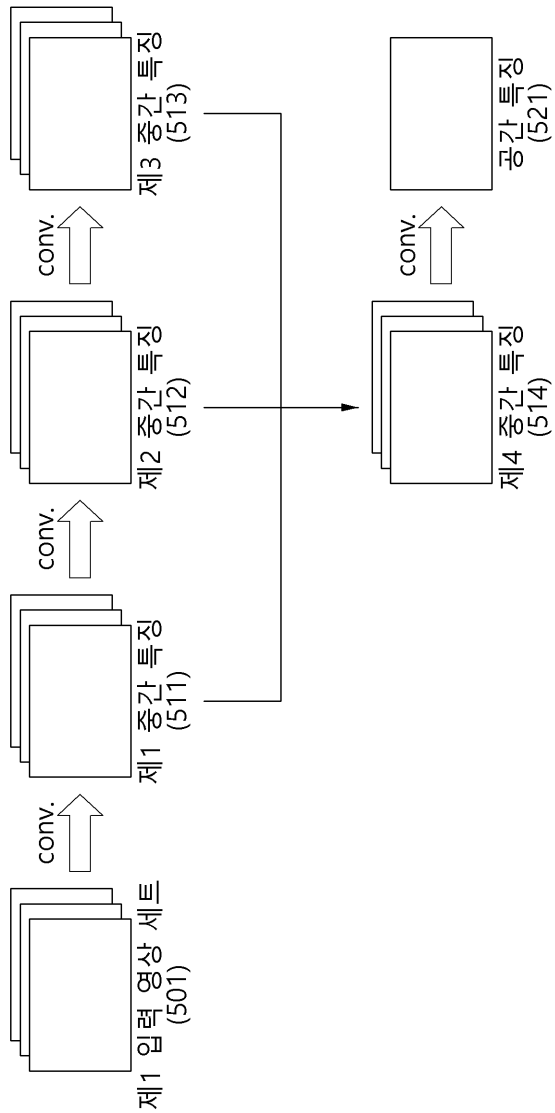
도면3



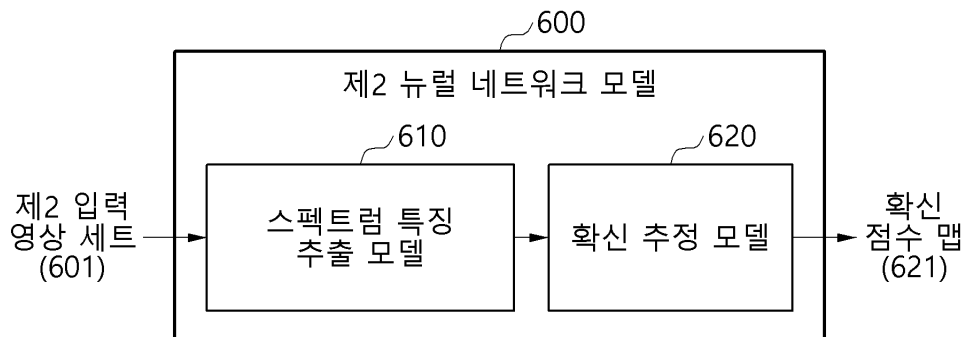
도면4



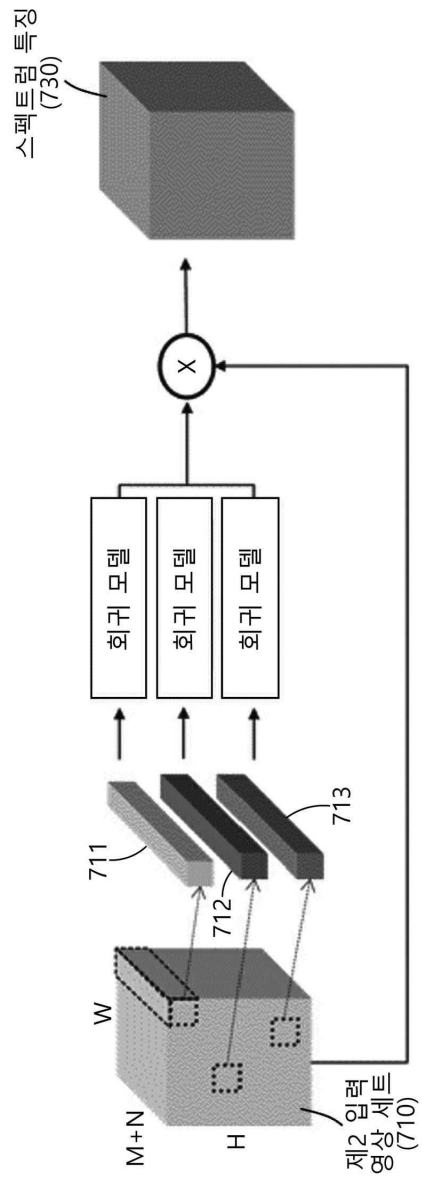
도면5



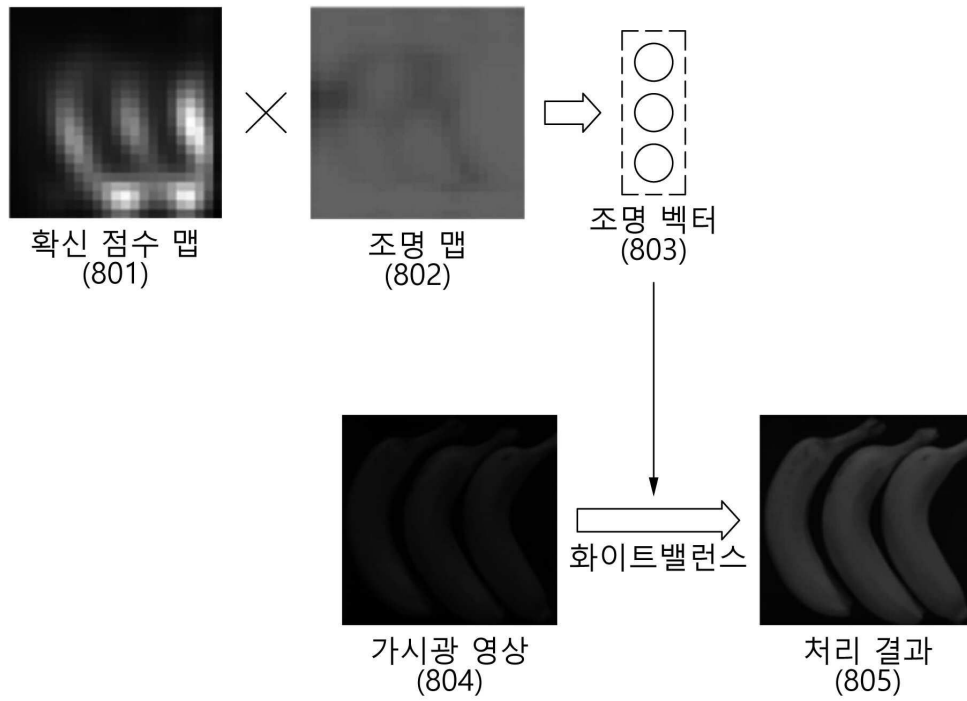
도면6



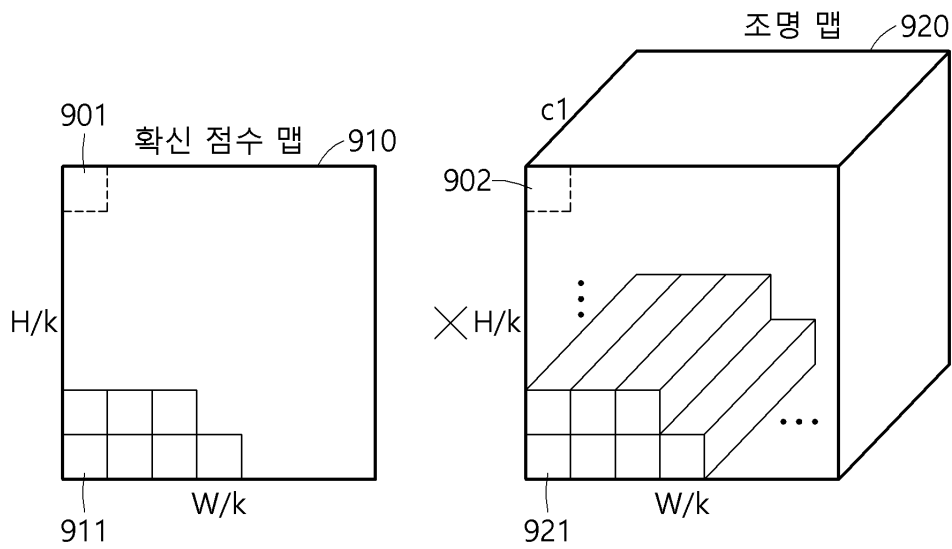
도면7



도면8



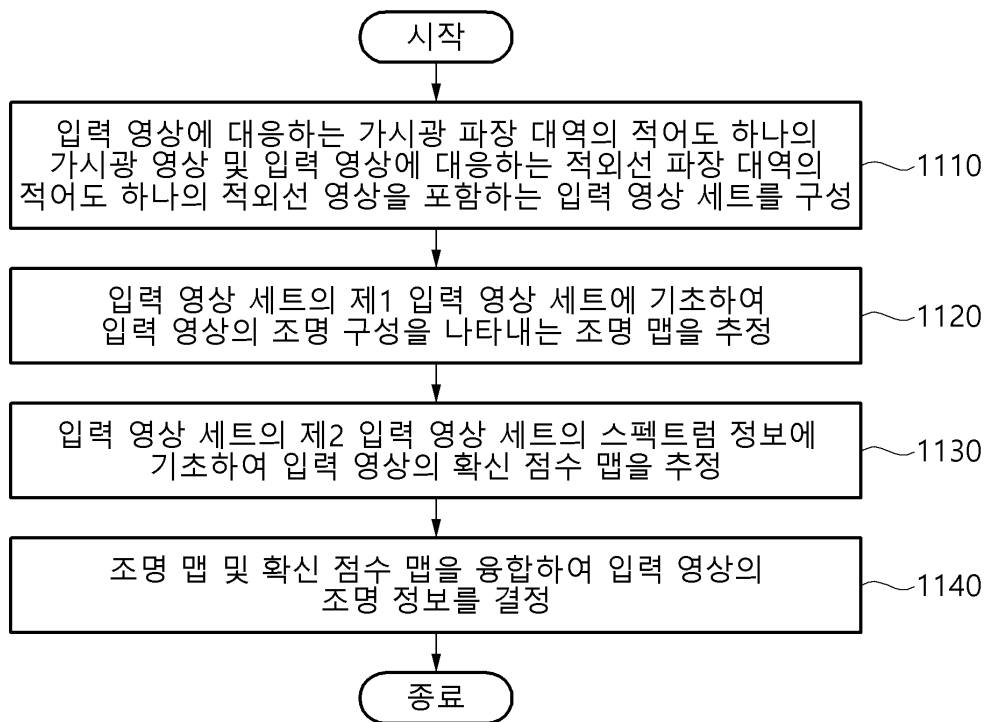
도면9



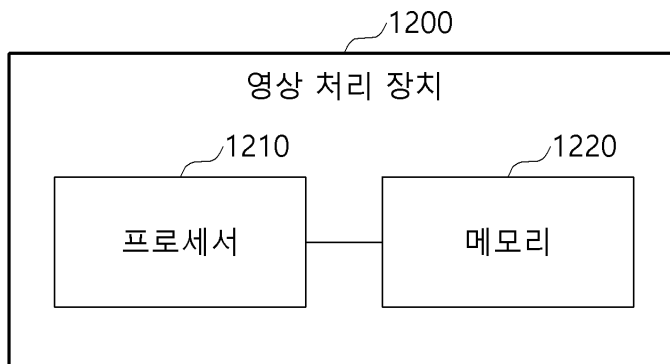
도면10



도면11



도면12



도면13

