

(12) 특허협력조약에 의하여 공개된 국제출원

(19) 세계지식재산권기구
국제사무국

(43) 국제공개일
2022년 5월 5일 (05.05.2022)

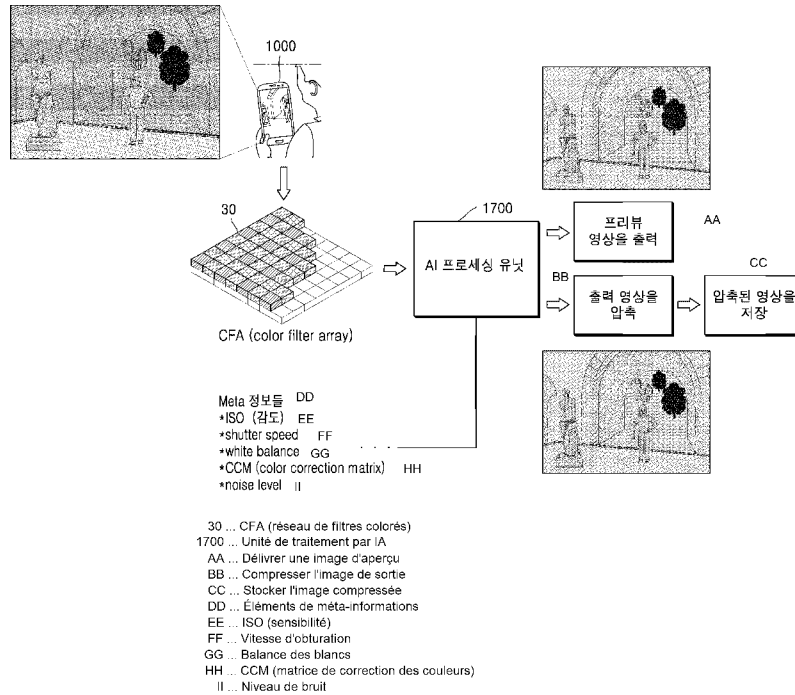


(10) 국제공개번호
WO 2022/092742 A1

- (51) 국제특허분류: G06T 5/00 (2006.01) G06N 3/08 (2006.01)
G06T 3/40 (2006.01) G06N 3/04 (2006.01)
- (21) 국제출원번호: PCT/KR2021/015056
- (22) 국제출원일: 2021년 10월 26일 (26.10.2021)
- (25) 출원언어: 한국어
- (26) 공개언어: 한국어
- (30) 우선권정보: 10-2020-0140682 2020년 10월 27일 (27.10.2020)KR
10-2020-0149890 2020년 11월 11일 (11.11.2020)KR
- (71) 출원인: 삼성전자 주식회사 (SAMSUNG ELECTRONICS CO., LTD.) [KR/KR]; 16677 경기도 수원시 영통구 삼성로 129, Gyeonggi-do (KR).
- (72) 발명자: 딘꾸오칸 (DINH, Quockhanh); 16677 경기도 수원시 영통구 삼성로 129, Gyeonggi-do (KR). 진경환 (JIN, Kyonghwan); 16677 경기도 수원시 영통구 삼성로 129, Gyeonggi-do (KR). 박영오 (PARK, Youngo); 16677 경기도 수원시 영통구 삼성로 129, Gyeonggi-do (KR). 최광표 (CHOI, Kwangpyo); 16677 경기도 수원시 영통구 삼성로 129, Gyeonggi-do (KR).
- (74) 대리인: 리앤목 특허법인 (Y.P.LEE, MOCK & PARTNERS); 06292 서울시 강남구 언주로30길 13 대림아크로텔 12층, Seoul (KR).
- (81) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 국내 권리의 보호를 위하여): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KW,

(54) Title: DEVICE AND METHOD FOR GENERATING IMAGE IN WHICH SUBJECT HAS BEEN CAPTURED

(54) 발명의 명칭: 피사체가 촬영된 영상을 생성하는 디바이스 및 방법



(57) Abstract: A device and a method for generating an image in which a subject has been captured are provided. A method by which a device generates, from a raw image, an output image in which a subject has been captured comprises the operations of: using a first processor for control of the device to control a camera sensor in the device, and thus acquire a raw image indicating the subject; using a second processor for image processing of the raw image through an artificial intelligence model, so as to input, into a first artificial intelligence model trained for adjusting the brightness of the image, the raw image indicating the subject; acquiring tone map data outputted from the first artificial intelligence model by using the second processor; and storing the output image generated on the basis of the tone map data.



WO 2022/092742 A1

KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

- (84) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 역내 권리의 보호를 위하여): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 유라시아 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 유럽 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

공개:

— 국제조사보고서와 함께 (조약 제21조(3))

(57) 요약서: 피사체가 촬영된 영상을 생성하는 디바이스 및 방법이 제공된다. 디바이스가 원시 영상으로부터 피사체가 촬영된 출력 영상을 생성하는 방법은, 상기 디바이스의 제어를 위한 제1 프로세서를 이용하여, 상기 디바이스 내의 카메라 센서를 제어함으로써 피사체를 나타내는 원시 영상을 획득하는 동작; 인공지능 모델을 이용하는 상기 원시 영상에 대한 영상 처리를 위한 제2 프로세서를 이용하여, 상기 피사체를 나타내는 상기 원시 영상을, 영상의 밝기를 조절하기 위해 훈련된 제1 인공지능 모델에 입력하는 동작; 상기 제2 프로세서를 이용하여, 상기 제1 인공지능 모델로부터 출력되는 톤 맵 데이터를 획득하는 동작; 및 상기 톤 맵 데이터에 기초하여 생성되는 출력 영상을 저장하는 동작;을 포함한다.

명세서

발명의 명칭: 피사체가 촬영된 영상을 생성하는 디바이스 및 방법 기술분야

- [1] 본 개시는 인공지능 모델을 이용하여 피사체가 촬영된 영상을 생성하는 디바이스 및 방법에 관한 것이다.

배경기술

- [2] 휴대 단말기가 널리 사용되고, 통신 네트워크 기술이 발전함에 따라 휴대 단말기와 관련된 소비자들의 요구가 다양화되고 있으며, 이에 따라 다양한 종류의 부가 장치들이 휴대 단말기에 장착되고 있다. 또한, 휴대 단말기가 소형화되고 피사체를 촬영하기 위한 카메라의 기능이 휴대 단말기에 의해 지원되고 있다.
- [3] 하지만, 휴대 단말기의 소형화로 인하여, 고성능의 카메라 센서를 휴대 단말에 장착하기 힘든 어려움이 있었으며, 카메라 센서를 통해 생성되는 로우 데이터를 프로세싱하는데 휴대 단말의 많은 리소스가 필요한 문제가 있었다.
- [4] 이에 따라, 휴대 단말의 카메라 센서를 이용하여 다양한 영상 특징을 가지는 영상을 생성하는 기술의 필요성이 높아지고 있으며, 나아가, 높은 품질의 HDR (High Dynamic Range) 이미지를 생성할 수 있는 인공지능 기술이 요구되고 있다.

발명의 상세한 설명

기술적 과제

- [5] 본 개시의 일 실시예는, 인공지능 모델을 이용하여 피사체가 촬영된 영상을 생성할 수 있는 디바이스 및 방법을 제공할 수 있다.
- [6] 또한, 본 개시의 일 실시예는, 카메라 센서로부터 생성되는 원시 영상을 인공지능 모델에 입력하여 피사체가 촬영된 출력 영상을 획득할 수 있는 디바이스 및 방법을 제공할 수 있다.
- [7] 또한, 본 개시의 일 실시예는, 톤 맵을 생성하기 위한 인공지능 모델을 이용하여 피사체가 촬영된 출력 영상을 획득할 수 있는 디바이스 및 방법을 제공할 수 있다.
- [8] 또한, 본 개시의 일 실시예는, 복수의 인공지능 모델을 순차적으로 이용하여 원시 영상으로부터 피사체가 촬영된 출력 영상을 획득할 수 있는 디바이스 및 방법을 제공할 수 있다.
- [9] 또한, 본 개시의 일 실시예는, 함께 훈련된 복수의 인공지능 모델을 적어도 일부를 이용하여, 피사체를 촬영하기 위한 라이브 뷰 영상 및 피사체가 촬영된 출력 영상을 획득할 수 있는 디바이스 및 방법을 제공할 수 있다.
- [10] 또한, 본 개시의 일 실시예는, 피사체가 촬영되는 상황 또는 사용자의 선호 중 적어도 하나에 관련하여 훈련된 인공지능 모델을 이용하여, 피사체를 촬영하기 위한 라이브 뷰 영상 및 피사체가 촬영된 출력 영상을 획득할 수 있는 디바이스

및 방법을 제공할 수 있다.

과제 해결 수단

- [11] 상술한 기술적 과제를 달성하기 위한 기술적 수단으로서, 본 개시의 제1 측면은, 디바이스의 제어를 위한 제1 프로세서를 이용하여, 상기 디바이스 내의 카메라 센서를 제어함으로써 상기 카메라 센서의 전방의 피사체를 나타내는 원시 영상을 획득하는 동작; 인공지능 모델을 이용하는 상기 원시 영상에 대한 영상 처리를 위한 제2 프로세서를 이용하여, 상기 카메라 센서의 전방의 상기 피사체를 나타내는 상기 원시 영상을, 영상의 밝기를 조절하기 위해 훈련된 제1 인공지능 모델에 입력하는 동작; 상기 제2 프로세서를 이용하여, 상기 제1 인공지능 모델로부터 출력되는 톤 맵 데이터를 획득하는 동작; 및 상기 톤 맵 데이터에 기초하여 생성되는 출력 영상을 저장하는 동작;을 포함하는, 디바이스가 원시 영상으로부터 피사체가 촬영된 출력 영상을 생성하는 방법을 제공할 수 있다.
- [12] 또한, 본 개시의 제2 측면은, 카메라 센서; 디스플레이; 상기 디바이스의 제어를 위한 제1 명령어들을 저장하는 제1 메모리; 상기 제1 메모리에 저장된 제1 명령어들을 실행하는 제1 프로세서; 상기 원시 영상에 대한 영상 처리를 위한 적어도 하나의 인공지능 모델 및 상기 인공지능 모델의 실행에 관련된 제2 명령어들을 저장하는 제2 메모리; 및 상기 제2 메모리에 저장된 상기 적어도 하나의 인공지능 모델 및 상기 제2 명령어들을 실행하는 제2 프로세서;를 포함하며, 상기 제1 프로세서가 상기 카메라 센서를 이용하여 상기 카메라 센서의 전방의 피사체를 나타내는 원시 영상을 획득하고, 상기 제2 프로세서가 상기 카메라 센서의 전방의 상기 피사체를 나타내는 상기 원시 영상을, 영상의 밝기를 조절하기 위해 훈련된 제1 인공지능 모델에 입력하고, 상기 제2 프로세서가 상기 제1 인공지능 모델로부터 출력되는 톤 맵 데이터를 획득하고, 상기 제1 프로세서가 상기 톤 맵 데이터에 기초하여 생성되는 출력 영상을 상기 제1 메모리에 저장하는 것인, 원시 영상으로부터 피사체가 촬영된 출력 영상을 생성하는 디바이스를 제공할 수 있다.
- [13] 또한, 본 개시의 제3 측면은, 제1 측면의 방법을 컴퓨터에서 실행시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체를 제공할 수 있다.

도면의 간단한 설명

- [14] 도 1은 본 개시의 일 실시예에 따른, 디바이스(1000)가 피사체를 촬영하여 출력 영상을 생성하는 예시를 나타내는 도면이다.
- [15] 도 2는 본 개시의 일 실시예에 따른 디바이스(1000)의 블록도이다.
- [16] 도 3은 본 개시의 일 실시예에 따른 원시 영상으로부터 피사체가 촬영된 출력 영상이 생성되는 과정을 설명하는 도면이다.
- [17] 도 4a는 본 개시의 일 실시예에 따른 디바이스(1000)가 원시 영상(30)으로부터 톤 맵을 생성하는 과정을 설명하는 도면이다.

- [18] 도 4b는 본 개시의 일 실시예에 따른 디바이스(1000)가 원시 영상(30)으로부터 특징들을 추출하는 과정을 설명하는 도면이다.
- [19] 도 4c는 본 개시의 일 실시예에 따른 디바이스(1000)가 특징 영상들을 보정하는 과정을 설명하는 도면이다.
- [20] 도 4d는 본 개시의 일 실시예에 따른 디바이스(1000)가 보정된 특징 영상들로부터 출력 영상을 생성하는 과정을 설명하는 도면이다.
- [21] 도 5는 본 개시의 일 실시예에 따른 특징 추출 모델(1732)의 구조의 일례를 나타내는 도면이다.
- [22] 도 6은 본 개시의 일 실시예에 따른 디바이스(1000)가 피사체를 촬영하여 출력 영상을 생성하는 방법의 흐름도이다.
- [23] 도 7은 본 개시의 일 실시예에 따른 인공지능 모델(1620)이 훈련되는 예시를 나타내는 도면이다.
- [24] 도 8은 본 개시의 일 실시예에 따른 인공지능 모델(1620)을 훈련하는 방법의 흐름도이다.
- [25] 도 9는 본 개시의 일 실시예에 따른 디바이스(1000)가 라이브 뷰 영상을 출력하는 방법의 흐름도이다.
- [26] 도 10a는 본 개시의 일 실시예에 따른 라이브 뷰 영상을 생성하기 위하여 인공지능 모델(1730) 내의 톤 맵 생성 모델(1731)을 비활성화하는 예시를 나타내는 도면이다.
- [27] 도 10b는 본 개시의 일 실시예에 따른 라이브 뷰 영상을 생성하기 위하여 인공지능 모델(1730) 내의 특징 추출 모델(1732) 및 영상 회복 모델(1734)을 비활성화하는 예시를 나타내는 도면이다.
- [28] 도 10c는 본 개시의 일 실시예에 따른 라이브 뷰 영상을 생성하기 위하여 인공지능 모델(1730) 내의 톤 맵 생성 모델(1731), 특징 추출 모델(1732) 및 영상 회복 모델(1734)을 비활성화하는 예시를 나타내는 도면이다.
- [29] 도 11은 본 개시의 일 실시예에 따른 디바이스(1000)가 리트레이닝된 인공지능 모델(1620)을 서버(미도시)로부터 수신하여 인공지능 모델(1620)을 업데이트하는 방법의 흐름도이다.
- [30] 도 12는 본 개시의 일 실시예에 따른 디바이스(1000)가 인공지능 모델(1620)을 리트레이닝하여 업데이트하는 방법의 흐름도이다.
- [31] 도 13a는 본 개시의 일 실시예에 따른 피사체를 촬영하는 디바이스(1000)의 촬영 설정을 위한 GUI의 예시를 나타내는 도면이다.
- [32] 도 13b는 본 개시의 일 실시예에 따른 피사체를 촬영하는 디바이스(1000)의 촬영 설정을 위한 GUI의 예시를 나타내는 도면이다.
- 발명의 실시를 위한 형태**
- [33] 아래에서는 첨부한 도면을 참조하여 본 개시가 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 본 개시의 실시예를 상세히

설명한다. 그러나 본 개시는 여러 가지 상이한 형태로 구현될 수 있으며 여기에서 설명하는 실시예에 한정되지 않는다. 그리고 도면에서 본 개시를 명확하게 설명하기 위해서 설명과 관계없는 부분은 생략하였으며, 명세서 전체를 통하여 유사한 부분에 대해서는 유사한 도면 부호를 붙였다.

- [34] 명세서 전체에서, 어떤 부분이 다른 부분과 "연결"되어 있다고 할 때, 이는 "직접적으로 연결"되어 있는 경우뿐 아니라, 그 중간에 다른 소자를 사이에 두고 "전기적으로 연결"되어 있는 경우도 포함한다. 또한 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있는 것을 의미한다.
- [35] 또한, 본 명세서에서, AI 프로세싱 유닛은 인공지능 기술을 이용하여 이미지를 처리하기 위한 프로세싱 유닛일 수 있다. 예를 들어, AI 프로세싱 유닛은 인공지능 모델을 이용한 이미지 처리를 위하여 설계된 프로세싱 유닛으로서, 이미지 처리를 위한 전용의 프로세싱 유닛일 수 있다. 또는, 예를 들어, AI 프로세싱 유닛은 NPU(Neural Processing Unit)에 인공지능 모델을 이용한 이미지 처리를 위한 설정을 함으로써 구현될 수 있다.
- [36] 또한, 본 명세서에서, 인공지능 모델은 원시 영상(raw image)으로부터 피사체가 촬영된 결과물인 출력 영상을 생성하기 위하여 훈련된 모델로서, 복수의 하위 인공지능 모델을 포함할 수 있다. 인공지능 모델에 포함된 복수의 하위 인공지능 모델은, 톤 맵 생성 모델, 특징 추출 모델, 영상 보정 모델 및 영상 회복 모델을 포함할 수 있다. 톤 맵 생성 모델은 원시 영상으로부터 톤 맵을 생성하기 위하여 훈련된 인공지능 모델이고, 특징 추출 모델은 특징 추출 모델에 입력된 원시 영상 내의 특징들을 추출하기 위하여 훈련된 인공지능 모델이며, 영상 보정 모델은 특징 추출 모델로부터 출력되는 특징 영상들을 보정하기 위하여 훈련된 인공지능 모델이며, 영상 회복 모델은 보정된 특징 영상들로부터 피사체가 촬영된 촬영 영상을 생성하기 위하여 훈련된 인공지능 모델일 수 있다.
- [37] 또한, 본 명세서에서, 톤 맵은 원시 영상 내의 픽셀들의 밝기를 스케일링하기 위한 정보를 포함하는 맵 데이터일 수 있다. 톤 맵은, 원시 영상 내의 부분 별로 픽셀의 밝기를 스케일링하는 로컬톤 매핑 또는 원시 영상의 전체에 대한 밝기를 스케일링하는 글로벌톤 매핑 중 적어도 하나를 위한 맵 데이터일 수 있다.
- [38] 또한, 본 명세서에서, 라이브 뷰 영상은 피사체를 촬영하는 사용자가 촬영되는 피사체를 확인할 수 있도록 디바이스의 화면 상에 출력되는 영상일 수 있다.
- [39] 이하 첨부된 도면을 참고하여 본 개시를 상세히 설명하기로 한다.
- [40] 도 1은 본 개시의 일 실시예에 따른, 디바이스(1000)가 피사체를 촬영하여 출력 영상을 생성하는 예시를 나타내는 도면이다.
- [41] 도 1을 참조하면, 디바이스(1000)는 피사체가 촬영된 영상을 생성하기 위한 인공지능 모델을 이용하여 피사체를 촬영할 수 있다. 디바이스(1000)는 피사체가 촬영된 영상을 생성하기 위한 AI 프로세싱 유닛(1700)을 포함할 수 있으며,

디바이스(1000)의 AI 프로세싱 유닛(1700)은 디바이스(1000) 내의 카메라 센서를 이용하여 생성되는 원시 영상 및 영상 보정을 위한 기설정된 설정 값을 인공지능 모델에 입력하고, 인공지능 모델로부터 출력되는 출력 영상을 획득할 수 있다. AI 프로세싱 유닛(1700)에 의해 이용되는 인공지능 모델은, 원시 영상피사체가 촬영된 결과물인 출력 영상을 원시 영상으로부터 생성하기 위하여 훈련된 모델로서, 복수의 하위 인공지능 모델을 포함할 수 있다. 인공지능 모델에 포함된 복수의 하위 인공지능 모델은, 톤 맵을 생성하기 위한 인공지능 모델, 영상의 특징을 추출하기 위한 인공지능 모델, 추출된 특징을 나타내는 특징 영상을 보정하기 위한 인공지능 모델 및 출력 영상을 생성하기 위한 인공지능 모델을 포함할 수 있다.

- [42] 또한, 디바이스(1000)의 AI 프로세싱 유닛(1700)은 복수의 하위 인공지능 모델 중 적어도 하나를 이용하여 피사체 촬영을 위한 라이브뷰 영상을 생성할 수 있다.
- [43] 디바이스(1000)는, 스마트폰, 태블릿 PC, PC, 스마트 TV, 휴대폰, PDA(personal digital assistant), 랩톱, 미디어 플레이어, GPS(global positioning system) 장치, 전자책 단말기, 디지털방송용 단말기, 네비게이션, 키오스크, 디지털 카메라, 가전기기 및 기타 모바일 또는 비모바일 컴퓨팅 장치일 수 있으나, 이에 제한되지 않는다. 또한, 디바이스(1000)는 통신 기능 및 데이터 프로세싱 기능을 구비한 시계, 안경, 헤어 밴드 및 반지 등의 웨어러블 디바이스일 수 있다. 그러나, 이에 제한되지 않으며, 디바이스(1000)는 피사체를 촬영할 수 있는 모든 종류의 기기를 포함할 수 있다.
- [44] 디바이스(1000)는 피사체가 촬영된 영상을 획득하기 위하여 서버(미도시)와 네트워크를 통하여 통신할 수 있다. 네트워크는 근거리 통신망(Local Area Network; LAN), 광역 통신망(Wide Area Network; WAN) 또는 부가가치 통신망(Value Added Network; VAN) 등과 같은 유선 네트워크나 이동 통신망(mobile radio communication network) 또는 위성 통신망 등과 같은 모든 종류의 무선 네트워크로 구현될 수 있다. 또한, 네트워크는 근거리 통신망(Local Area Network; LAN), 광역 통신망(Wide Area Network; WAN), 부가가치 통신망(Value Added Network; VAN), 이동 통신망(mobile radio communication network) 또는 위성 통신망 중 적어도 둘 이상의 상호 조합을 포함할 수 있으며, 각 네트워크 구성 주체가 서로 원활하게 통신을 할 수 있도록 하는 포괄적인 의미의 데이터 통신망이며, 유선 인터넷, 무선 인터넷 및 모바일 무선 통신망을 포함할 수 있다. 무선 통신은 예를 들어, 무선 랜(Wi-Fi), 블루투스, 블루투스 저 에너지(Bluetooth low energy), 지그비, WFD(Wi-Fi Direct), UWB(ultra wideband), 적외선 통신(IrDA, infrared Data Association), NFC(Near Field Communication) 등이 있을 수 있으나, 이에 한정되는 것은 아니다.
- [45] 도 2는 본 개시의 일 실시예에 따른 디바이스(1000)의 블록도이다.
- [46] 도 2를 참조하면, 본 개시의 일 실시예에 따른 디바이스(1000)는 사용자

- 입력부(1100), 디스플레이부(1200), 통신 인터페이스(1300), 카메라 센서부(1400), 제1 프로세서(1500), 제1 메모리(1600) 및 AI 프로세싱 유닛(1700)를 포함할 수 있다. 또한, AI 프로세싱 유닛(1700)은 인공지능 기술을 이용하여 이미지를 처리하기 위한 제2 프로세서(1710) 및 제2 메모리(1720)을 포함할 수 있다. 예를 들어, AI 프로세싱 유닛(1700)은 인공지능 모델을 이용한 이미지 처리를 위하여 설계된 프로세싱 유닛으로서, 이미지 처리를 위한 전용의 프로세싱 유닛일 수 있다. 또는, 예를 들어, AI 프로세싱 유닛(1700)은 NPU(Neural Processing Unit)에 인공지능 모델을 이용한 이미지 처리를 위한 설정을 함으로써 구현될 수 있다.
- [47] 사용자 입력부(1100)는, 사용자가 디바이스(1000)를 제어하기 위한 데이터를 입력하는 수단을 의미한다. 예를 들어, 사용자 입력부(1100)는 키 패드(key pad), 돔 스위치 (dome switch), 터치 패드(접촉식 정전 용량 방식, 압력식 저항막 방식, 적외선 감지 방식, 표면 초음파 전도 방식, 적분식 장력 측정 방식, 피에조 효과 방식 등), 조그 휠 또는 조그 스위치 중 적어도 하나를 포함할 수 있으나 이에 한정되는 것은 아니다. 사용자 입력부(1100)는 디바이스(1000)를 이용하는 사용자가 사진을 촬영하기 위한 사용자 입력을 수신할 수 있다.
- [48] 디스플레이부(1200)는 디바이스(1000)에서 처리되는 정보를 표시 출력한다. 예를 들어, 디스플레이부(1200)는, 사진 촬영을 위한 GUI, 라이브 뷰 영상, 사진 촬영의 결과물로 출력되는 출력 영상을 디스플레이할 수 있다.
- [49] 한편, 디스플레이부(1200)와 터치패드가 레이어 구조를 이루어 터치 스크린으로 구성되는 경우, 디스플레이부(1200)는 출력 장치 이외에 입력 장치로도 사용될 수 있다. 디스플레이부(1200)는 액정 디스플레이(liquid crystal display), 박막 트랜지스터 액정 디스플레이(thin film transistor-liquid crystal display), 유기 발광 다이오드(organic light-emitting diode), 플렉시블 디스플레이(flexible display), 3차원 디스플레이(3D display), 전기영동 디스플레이(electrophoretic display) 중에서 적어도 하나를 포함할 수 있다. 그리고 디바이스(1000)의 구현 형태에 따라 디바이스(1000)는 디스플레이부(1200)를 2개 이상 포함할 수도 있다.
- [50] 통신 인터페이스(1300)는, 다른 디바이스(미도시) 및 서버(미도시) 간의 통신을 하게 하는 하나 이상의 구성요소를 포함할 수 있다. 예를 들어, 통신 인터페이스(1300)는, 근거리 통신부, 이동 통신부, 방송 수신부를 포함할 수 있다.
- [51] 근거리 통신부(short-range wireless communication unit)는, 블루투스 통신부, BLE(Bluetooth Low Energy) 통신부, 근거리 무선 통신부(Near Field Communication unit), WLAN 통신부, 지그비(Zigbee) 통신부, 적외선(IrDA, infrared Data Association) 통신부, WFD(Wi-Fi Direct) 통신부, UWB(ultra wideband) 통신부, Ant+ 통신부 등을 포함할 수 있으나, 이에 한정되는 것은 아니다. 이동 통신부는, 이동 통신망 상에서 기지국, 외부의 단말, 서버 중 적어도 하나와 무선 신호를 송수신한다. 여기에서, 무선 신호는, 음성 호 신호, 화상 통화 호 신호 또는 문자/멀티미디어 메시지 송수신에 따른 다양한 형태의 데이터를

포함할 수 있다. 방송 수신부는, 방송 채널을 통하여 외부로부터 방송 신호 및/또는 방송 관련된 정보를 수신한다. 방송 채널은 위성 채널, 지상파 채널을 포함할 수 있다. 구현 예에 따라서 디바이스(1000)가 방송 수신부를 포함하지 않을 수도 있다.

[52] 또한, 통신 인터페이스(1300)는, 사진 촬영을 위한 인공지능 모델(1620)을 관리하기 위하여 필요한 정보를, 다른 디바이스(미도시) 및 서버(미도시)와 송수신할 수 있다.

[53] 카메라 센서부(1400)는 컬러 필터(1410) 및 이미지 센서(1420)를 포함하며, 후술할 제1 프로세서(1500) 또는 제2 프로세서(1710)에 의해 제어됨으로써, 피사체에 대한 원시 영상(raw image)을 생성할 수 있다. 카메라 센서부(1400)는 제1 프로세서(1500)의 제어 신호에 기초한 제2 프로세서(1710)의 제어 또는 제2 프로세서(1710)의 독립적인 제어에 의해, 피사체에 대한 원시 영상을 생성할 수 있다. 예를 들어, 제2 프로세서(1710)가 카메라 센서부(1400) 내에 포함되는 경우에, 제1 프로세서(1500)가 카메라 센서부(1400)에 대해서 촬영을 위한 제어 요청 신호를 보내면, 카메라 센서부(1400) 내의 제2 프로세서(1710)가 제어 요청 신호를 수신하여 카메라 센서부(1400)를 제어하여 원시 영상을 생성할 수 있다. 또는, 예를 들어, 제2 프로세서(1710)가 카메라 센서부(1400)의 외부에 있는 경우에, 제1 프로세서(1500)가 촬영을 위한 제어 요청 신호를 제2 프로세서(1710)에게 보내면, 제2 프로세서(1710)가 제어 요청 신호를 수신하여 카메라 센서부(1400)를 제어하여 원시 영상을 생성할 수 있다. 또는, 예를 들어, 제1 프로세서(1500)가 카메라 센서부(1400)를 제어하여 원시 영상을 획득하고 획득된 원시 영상을 제2 프로세서(1710)에게 제공할 수도 있다.

[54] CMOS(complementary metal-oxide semiconductor) 또는 CCD(charge coupled device) 등의 이미지 센서(1420)에 의해 획득되는 영상은 단색(monochrome)의 영상이므로, 이미지 센서(1420)의 전단부에 가시광선 영역에서 특정 주파수 대역만을 통과시키는 컬러 필터(1410)가 배치될 수 있으며, 컬러 필터(1410)를 통과하여 이미지 센서(1420)로 전달되는 광으로부터 컬러 영상이 획득될 수 있다. 컬러 필터(1410)는 피사체로부터 전달되는 광 중에서 특정 주파수 대역을 갖는 광을 통과시킨다. 컬러 필터(1410)는 복수의 색상에 대응되는 복수의 영역으로 구분되어 있으며, 복수의 영역 각각은, 예를 들어, 적색, 녹색, 청색의 3개의 색상 중 하나의 색상이 갖는 주파수 대역과 동일한 주파수 대역을 갖는 광만을 통과시킬 수 있다. 컬러 필터(1410)를 통과한 광은 CMOS 또는 CCD 등의 이미지 센서(1420)로 전달되며, 이미지 센서(1420)는 전달받은 광을 전기적 신호로 변환할 수 있다. 예를 들어, 이미지 센서(1420)에 의해 변환된 전기적 신호는 적색 값, 녹색 값 및 청색 값으로 구성될 수 있으며, 적색 값, 녹색 값 및 청색 값의 배열로 구성된 원시 영상(raw image)이 생성될 수 있다. 원시 영상은 컬러 필터(1410)의 어레이 패턴에 따른 패턴을 가질 수 있으며, 예를 들어, 바이어(bayer) 패턴, RGBE 패턴, RYYB 패턴, CYYM 패턴, CYGM 패턴, RGBW

- 베이어 패턴, 및 X-trans 패턴 중 하나의 패턴을 가질 수 있다.
- [55] 제1 메모리(1600)는, 후술할 제1 프로세서(1500) 또는 제2 프로세서(1700) 중 적어도 하나의 처리 및 제어를 위한 프로그램을 저장할 수 있고, 디바이스(1000)로 입력되거나 디바이스(1000)로부터 출력되는 데이터를 저장할 수 있다.
- [56] 제1 메모리(1600)는 플래시 메모리 타입(flash memory type), 하드디스크 타입(hard disk type), 멀티미디어 카드 마이크로 타입(multimedia card micro type), 카드 타입의 메모리(예를 들어 SD 또는 XD 메모리 등), 램(RAM, Random Access Memory) SRAM(Static Random Access Memory), 롬(ROM, Read-Only Memory), EEPROM(Electrically Erasable Programmable Read-Only Memory), PROM(Programmable Read-Only Memory), 자기 메모리, 자기 디스크, 광디스크 중 적어도 하나의 타입의 저장매체를 포함할 수 있다.
- [57] 제1 메모리(1600)에 저장된 프로그램들은 그 기능에 따라 복수 개의 모듈들로 분류할 수 있는데, 예를 들어, 영상 처리 모듈(1610), 적어도 하나의 인공지능 모델(1620) 및 모델 관리 모듈(1630)을 포함할 수 있다. 또한, 영상 처리 모듈(1610)은 전처리 모듈(1611), 출력 영상 생성 모듈(1612) 및 라이브 뷰 생성 모듈(1613)을 포함하고, 인공지능 모델(1620)은 톤 맵 생성 모델(tone map generation model), 특징 추출 모델(feature extraction model), 영상 보정 모델(image modifying model) 및 영상 회복 모델(image regression model)와 같은 복수의 하위 인공지능 모델들을 포함하며, 모델 관리 모듈(1630)은 모델 선택 모듈(1634), 다운로드 모듈(1631), 업데이트 모듈(1632) 및 리트레이닝 모듈(1633)을 포함할 수 있다. 제1 프로세서(1500)는, 통상적으로 디바이스(1000)의 전반적인 동작을 제어한다. 예를 들어, 제어부(1300)는, 메모리(1700)에 저장된 프로그램들을 실행함으로써, 사용자 입력부(1100), 디스플레이부(1200), 통신 인터페이스(1300), 카메라 센서부(1400) 및 제1 메모리(1600)를 제어할 수 있다. 제1 프로세서(1500)는 하나 또는 복수의 프로세서로 구성될 수 있다. 이때, 하나 또는 복수의 프로세서는 CPU, AP, DSP(Digital Signal Processor) 등과 같은 범용 프로세서 또는 GPU, VPU(Vision Processing Unit)와 같은 그래픽 전용 프로세서 중 적어도 하나를 포함할 수 있다.
- [58] AI 프로세싱 유닛(1700)은 제2 프로세서(1720) 및 제2 메모리(1720)를 포함할 수 있다. 제2 프로세서(1720) 및 제2 메모리(1720)는 인공지능 모델을 이용한 이미지 처리를 위하여 설계될 수 있으나, 이에 제한되지 않는다.
- [59] 제2 메모리(1720)는 제2 프로세서(1720)의 처리 및 제어를 위한 프로그램을 저장할 수 있다. 예를 들어, 제2 메모리(1720)는 제1 메모리(1600)에 저장된 인공지능 모델(1620) 중에서 선택된 인공지능 모델(1720)이 저장될 수 있으며, 제2 메모리(1720)에 저장된 인공지능 모델(1720)은 제2 프로세서(1720)에 의해 실행될 수 있다. 제2 메모리(1720)에 저장된 인공지능 모델(1720)은 톤 맵 생성 모델(tone map generation model)(1731), 특징 추출 모델(feature extraction

model)(1732), 영상 보정 모델(image modifying model)(1733) 및 영상 회복 모델(image regression model)(1734)와 같은 복수의 하위 인공지능 모델들을 포함할 수 있다.

- [60] 제2 메모리(1720)는 플래시 메모리 타입(flash memory type), 하드디스크 타입(hard disk type), 멀티미디어 카드 마이크로 타입(multimedia card micro type), 카드 타입의 메모리(예를 들어 SD 또는 XD 메모리 등), 램(RAM, Random Access Memory) SRAM(Static Random Access Memory), 롬(ROM, Read-Only Memory), EEPROM(Electrically Erasable Programmable Read-Only Memory), PROM(Programmable Read-Only Memory), 자기 메모리, 자기 디스크, 광디스크 중 적어도 하나의 타입의 저장매체를 포함할 수 있다.
- [61] 제1 메모리(1600)에 저장된 영상 처리 모듈(1610)은 제1 프로세서(1500) 또는 제2 프로세서(1720) 중 적어도 하나에 의해 실행됨으로써 피사체를 촬영하기 위한 라이브 뷰 영상 및 피사체가 촬영된 출력 영상을 생성할 수 있다. 라이브 뷰 영상은 피사체를 촬영하는 사용자가 촬영되는 피사체를 확인할 수 있도록 디스플레이부(1200) 상에 출력되는 영상이며, 피사체가 촬영된 출력 영상은 피사체를 촬영한 결과물로서 디바이스(1000)에 저장되는 피사체가 촬영된 영상일 수 있다.
- [62] 제1 프로세서(1500)는 제1 메모리(1600)에 저장된 전처리 모듈(1611)을 실행함으로써 피사체에 대한 원시 영상을 생성할 수 있다. 제1 프로세서(1500)는, 피사체를 촬영하는 사용자 입력이 수신됨에 따라, 사용자 입력이 수신된 때에 카메라 센서부(1400)의 컬러 필터(1410)를 통과하여 이미지 센서(1420)로 전달되는 광에 기초한 원시 영상을 생성할 수 있다. 제1 프로세서(1500)는, 예를 들어, 적색 값, 녹색 값 및 청색 값의 배열로 구성된 원시 영상(raw image)을 생성할 수 있다.
- [63] 제1 프로세서(1500)는 출력 영상 생성 모듈(1612)을 실행함으로써 피사체가 촬영된 출력 영상을 생성할 수 있다. 제1 프로세서(1500)는 원시 영상을 이용하여 출력 영상을 생성할 것을 제2 프로세서(1720)에게 요청할 수 있다. 이에 따라, 제2 프로세서(1720)는 톤 맵 생성 모델(1731), 특징 추출 모델(1732), 영상 보정 모델(1733) 및 영상 회복 모델(1734)을 포함하는 인공지능 모델(1720)에 원시 영상을 입력하고 인공지능 모델(1720)로부터 출력되는 출력 영상을 획득할 수 있다. 제2 프로세서(1710)는, 원시 영상을 톤 맵 생성 모델(1731)에 입력함으로써 톤 맵을 획득하고, 생성된 톤 맵 및 원시 영상을 특징 추출 모델(1732)에 입력함으로써 피사체의 복수의 특징들에 관한 특징 영상들을 획득하고, 특징 영상들을 영상 보정 모델(1733)에 입력함으로써 복수의 특징 영상들을 보정하고, 보정된 복수의 특징 영상들을 영상 회복 모델(1734)에 입력함으로써 디바이스(1000)에 저장될 출력 영상을 획득할 수 있다. 또한, 인공지능 모델(1720)로부터 출력되는 출력 영상은 제1 프로세서(1500)에 의해 디스플레이부(1200)에 표시되고 제1 메모리(1600)에 저장될 수 있다.

- [64] 또한, 사용자가 피사체를 촬영하는 상황에 관련된 인공지능 모델(1720)이 제1 메모리(1600)에 저장된 인공지능 모델들(1620)로부터 선택되어 제2 메모리(1720)에 로딩될 수 있으며, 이에 대하여는 후술하기로 한다.
- [65] 제2 프로세서(1710)는 원시 영상 내의 적색 값, 녹색 값 및 청색 값의 밝기(brightness)를 스케일링하기 위하여 이용되는 톤 맵을 생성하기 위하여, 원시 영상을 톤 맵 생성 모델(1731)에 입력할 수 있다. 예를 들어, 제2 프로세서(1710)가 카메라 센서부(1400) 내에 포함되는 경우에, 제1 프로세서(1500)가 카메라 센서부(1400)에 대해서 촬영을 위한 제어 요청 신호를 보내면, 카메라 센서부(1400) 내의 제2 프로세서(1710)가 제어 요청 신호를 수신하여 카메라 센서부(1400)를 제어하여 원시 영상을 생성하고, 제2 프로세서(1710)가 원시 영상을 톤 맵 생성 모델(1731)에 입력할 수 있다. 또는, 예를 들어, 제2 프로세서(1710)가 카메라 센서부(1400)의 외부에 있는 경우에, 제1 프로세서(1500)가 촬영을 위한 제어 요청 신호를 제2 프로세서(1710)에게 보내면, 제2 프로세서(1710)가 제어 요청 신호를 수신하여 카메라 센서부(1400)를 제어하여 원시 영상을 생성하고, 제2 프로세서(1710)가 원시 영상을 톤 맵 생성 모델(1731)에 입력할 수 있다. 예를 들어, 제1 프로세서(1500)가 카메라 센서부(1400)를 제어하여 원시 영상을 획득하고 획득된 원시 영상을 제2 프로세서(1710)에게 제공하며, 제2 프로세서(1710)가 원시 영상을 톤 맵 생성 모델(1731)에 입력할 수 있다. 톤 맵은 원시 영상 내의 픽셀들의 밝기를 스케일링하기 위한 정보를 포함하는 맵 데이터일 수 있다. 톤 맵은, 원시 영상 내의 부분 별로 픽셀의 밝기를 스케일링하는 로컬톤 매핑 또는 원시 영상의 전체에 대한 밝기를 스케일링하는 글로벌톤 매핑 중 적어도 하나를 위한 맵 데이터일 수 있다. 또한, 예를 들어, 톤 맵은, 원시 영상 내에서 어두운 영역의 밝기를 밝은 영역의 밝기보다 크게 스케일링하도록 생성될 수 있다. 톤 맵 생성 모델(1731)은 원시 영상으로부터 톤 맵을 생성하기 위하여 훈련된 인공지능 모델일 수 있다. 톤 맵 생성 모델(1731)은 복수의 신경망 레이어들로 구성될 수 있으며, 복수의 신경망 레이어들 각각은 복수의 가중치들(weight values)을 갖고 있으며, 이전(previous) 레이어의 연산 결과와 복수의 가중치들 간의 연산을 통해 신경망 연산을 수행할 수 있다. 복수의 신경망 레이어들이 갖고 있는 복수의 가중치들은 톤 맵 생성 모델(1731)의 훈련 결과에 의해 최적화될 수 있으며, 예를 들어, 톤 맵 생성 모델(1731)은 특징 추출 모델(1732), 영상 보정 모델(1733) 및 영상 회복 모델(1734)과 함께 훈련될 수 있다. 또한, 예를 들어, 톤 맵 생성 모델(1731)은 CNN (Convolutional Neural Network)를 포함할 수 있으나, 이에 제한되지 않는다.
- [66] 제2 프로세서(1710)는 원시 영상 및 톤 맵을 조합하여 특징 추출 모델(1732)에 입력하고 특징 추출 모델(1732)로부터 출력되는 복수의 특징 영상들을 획득할 수 있다. 예를 들어, 제2 프로세서(1710)는 톤 맵을 이용하여 원시 영상 내의 픽셀들의 밝기를 스케일링하고, 밝기가 스케일링된 원시 영상을 특징 추출

모델(1732)에 입력할 수 있다. 특징 추출 모델(1732)에 입력되는 원시 영상은, 어두운 부분의 픽셀이 밝은 값을 가지도록 스케일링될 수 있으며 이에 따라, 원시 영상 내의 어두운 부분에 위치한 피사체가 보다 효과적으로 식별될 수 있게 된다.

- [67] 특징 추출 모델(1732)은 특징 추출 모델(1732)에 입력된 원시 영상 내의 특징들을 추출할 수 있다. 원시 영상 내의 복수의 특징들을 각각 나타내는 복수의 특징 영상들이 특징 추출 모델(1732)로부터 출력될 수 있다. 예를 들어, 복수의 특징 영상들은, 원시 영상 내의 엣지에 관한 특징을 나타내는 특징 영상, 원시 영상 내의 라인에 관한 특징을 나타내는 특징 영상, 원시 영상 내의 공간에 관련된 특징을 나타내는 특징 영상, 원시 영상 내의 객체의 형상 및 깊이와 관련된 특징을 나타내는 특징 영상, 원시 영상 내의 인물에 관련된 특징을 나타내는 특징 영상 및 원시 영상 내의 사물에 관련된 특징을 나타내는 특징 영상 등을 포함할 수 있으나, 이에 제한되지 않는다.
- [68] 특징 추출 모델(1732)은 특징 추출 모델(1732)에 입력된 원시 영상 내의 특징들을 추출하기 위하여 훈련된 인공지능 모델일 수 있다. 특징 추출 모델(1732)은 복수의 신경망 레이어들로 구성될 수 있으며, 복수의 신경망 레이어들 각각은 복수의 가중치들(weight values)을 갖고 있으며, 이전(previous) 레이어의 연산 결과와 복수의 가중치들 간의 연산을 통해 신경망 연산을 수행할 수 있다. 복수의 신경망 레이어들이 갖고 있는 복수의 가중치들은 특징 추출 모델(1732)의 훈련 결과에 의해 최적화될 수 있으며, 예를 들어, 특징 추출 모델(1732)은 톤 맵 생성 모델(1731), 영상 보정 모델(1733) 및 영상 회복 모델(1734)과 함께 훈련될 수 있다. 또한, 예를 들어, 특징 추출 모델(1732)은 후술할 도 5와 같이 End-to-End 방식의 Fully-Convolutional Network 기반 구조를 가지는 U-NET에 의해 구현될 수 있으나, 이에 제한되지 않는다.
- [69] 제2 프로세서(1710)는 특징 추출 모델(1732)로부터 출력된 특징 영상들을 보정할 수 있다. 제2 프로세서(1710)는 기설정된 영상 속성들에 관련한 설정에 기초하여 특징 영상들을 보정할 수 있다. 예를 들어, 제2 프로세서(1710)는 화이트 밸런스의 보정 및 색상 보정을 위한 기설정된 기준에 따라, 영상 보정 모델(1733)을 이용하여 특징 영상들을 보정할 수 있다. 제2 프로세서(1720)는 특징 추출 모델(1732)로부터 출력된 특징 영상들 및 영상 속성에 관련된 기설정된 속성 값들을 영상 보정 모델(1733)에 입력하고, 영상 보정 모델(1733)로부터 출력되는 보정된 특징 영상들을 획득할 수 있다.
- [70] 이 경우, 제2 프로세서(1720)는 특징 영상들을 보정하기 위한 영상 속성을 결정할 수 있다. 예를 들어, 특징 영상의 보정을 위한 영상 속성은, 디바이스(1000)의 촬영 환경에 따라 결정될 수 있다. 제2 프로세서(1720)는 카메라 센서부(1400)를 통하여 피사체를 촬영하는 때의 디바이스(1000) 주변의 환경을 나타내는 센싱 데이터를 획득하며, 센싱 데이터에 따라 미리 설정된 기준에 기초하여 특징 영상의 보정을 위한 화이트 밸런스 매트릭스 값 및 색상

- 보정 매트릭스 값을 영상 보정 모델(1733)에 특징 영상들과 함께 입력할 수 있다.
- [71] 또는, 예를 들어, 제1 프로세서(1500)는 피사체 촬영의 설정을 위하여 GUI를 디바이스(1000)의 화면 상에 디스플레이할 수 있으며, GUI를 통한 사용자 입력에 기초하여 화이트 밸런스의 보정 및 색상 보정 등을 위한 설정을 미리 수행할 수 있다.
- [72] 예를 들어, 디바이스(1000)는 도 13a 및 도 13b에서와 같이 피사체를 촬영하는 디바이스(1000)의 촬영 설정을 위한 GUI를 디스플레이부(1200) 상에 디스플레이할 수 있다. 이에 따라, 사용자는 디스플레이부(1200) 상에 디스플레이된 GUI를 통하여, ISO, 셔터 스피드, 화이트 밸런스, 색온도, 틸트, 대비, 채도, 하이라이트 효과, 쉐도우 효과 등에 관련된 설정 값을 디바이스(1000)에 입력할 수 있다.
- [73] 이 경우, 제1 프로세서(1500)는 사용자에게 의해 미리 설정된 기준에 따라, 화이트 밸런스 매트릭스 값 및 색상 보정 매트릭스 값을 제1 메모리(1600)로부터 추출하고, 추출된 화이트 밸런스 매트릭스 값 및 색상 보정 매트릭스 값을 영상 보정 모델(1733)에 특징 영상들과 함께 입력할 수 있다.
- [74] 영상 보정 모델(1733)은 영상 보정 모델(1733)에 입력된 특징 영상들의 영상 속성을 보정하기 위하여 훈련된 인공지능 모델일 수 있다. 영상 보정 모델(1733)은 복수의 신경망 레이어들로 구성될 수 있으며, 복수의 신경망 레이어들 각각은 복수의 가중치들(weight values)을 갖고 있으며, 이전(previous) 레이어의 연산 결과와 복수의 가중치들 간의 연산을 통해 신경망 연산을 수행할 수 있다. 복수의 신경망 레이어들이 갖고 있는 복수의 가중치들은 영상 보정 모델(1733)의 훈련 결과에 의해 최적화될 수 있으며, 예를 들어, 영상 보정 모델(1733)은 톤 맵 생성 모델(1731), 특징 추출 모델(1732) 및 영상 회복 모델(1734)과 함께 훈련될 수 있다.
- [75] 한편, 상기에서는, 화이트 밸런스 매트릭스 값 및 색상 보정 매트릭스 값 값이 영상 보정 모델(1733)에 입력되는 것으로 설명되었지만 이에 제한되지 않는다. 복수의 인공지능 모델(1620)이 화이트 밸런스 및 색상 보정에 관한 설정 별로 각각 훈련될 수 있다. 이 경우에는, 출력 영상을 생성하기 위해 소정의 화이트 밸런스 및 색상 보정에 관한 설정에 대응되는 인공지능 모델(1730)이 제2 메모리(1720)에 로딩되어 제2 프로세서(1710)에 의해 이용될 수 있다. 제2 메모리(1720)에 로딩된 인공지능 모델(1730)에는 화이트 밸런스 매트릭스 값 및 색상 보정 매트릭스 값이 입력되지 않더라도, 제2 메모리(1720)에 로딩된 인공지능 모델(1730)로부터 소정의 화이트 밸런스 및 색상 보정에 관한 설정을 고려한 출력 영상이 출력될 수 있다.
- [76] 제2 프로세서(1710)는 보정된 특징 영상들을 영상 회복 모델(1734)에 입력하고, 영상 회복 모델(1734)로부터 출력되는 출력 영상을 획득할 수 있다. 영상 회복 모델(1734)로부터 출력되는 출력 영상은 피사체를 촬영한 결과물로서 디바이스(1000)에 저장될 영상일 수 있다.

- [77] 또한, 피사체를 촬영하기 위한 사용자 입력이 수신되면 인공지능 모델(1730)을 통해 출력된 영상이 소정 기준에 따라 압축되고, 압축된 영상이 제1 메모리(1600)에 저장될 수 있으나 이에 제한되지 않는다.
- [78] 영상 회복 모델(1734)은 특징 영상들로부터 피사체가 촬영된 촬영 영상을 생성하기 위하여 훈련된 인공지능 모델일 수 있다. 영상 회복 모델(1734)은 복수의 신경망 레이어들로 구성될 수 있으며, 복수의 신경망 레이어들 각각은 복수의 가중치들(weight values)을 갖고 있으며, 이전(previous) 레이어의 연산 결과와 복수의 가중치들 간의 연산을 통해 신경망 연산을 수행할 수 있다. 복수의 신경망 레이어들이 갖고 있는 복수의 가중치들은 영상 회복 모델(1734)의 훈련 결과에 의해 최적화될 수 있으며, 예를 들어, 영상 회복 모델(1734)은, 톤 맵 생성 모델(1731), 특징 추출 모델(1732) 및 영상 보정 모델(1733)과 함께 훈련될 수 있다. 또한, 예를 들어, 영상 회복 모델(1734)은 CNN (Convolutional Neural Network)를 포함할 수 있으나, 이에 제한되지 않는다.
- [79] 제1 프로세서(1500)는 라이브 뷰 생성 모듈(1613)을 실행함으로써, 라이브 뷰 영상을 생성할 수 있다. 피사체를 촬영하기 위한 촬영 기능을 활성화하는 사용자 입력이 수신되면, 제1 프로세서(1500)는 라이브 뷰 영상을 생성하고 생성된 라이브뷰 영상을 디스플레이부(1200) 상에 디스플레이할 수 있다. 예를 들어, 디바이스(1000)에 설치된 카메라 애플리케이션을 실행하는 사용자 입력이 수신되면, 제1 프로세서(1500)는 카메라 애플리케이션을 실행하고 사용자가 촬영될 피사체를 확인할 수 있도록 라이브 뷰 영상을 생성하여 디스플레이부(1200) 상에 디스플레이할 수 있다.
- [80] 제1 프로세서(1500)는, 디바이스(1000)의 촬영 기능이 활성화됨에 따라, 카메라 센서부(1400)를 통하여 입력되는 광에 기초하여 라이브 뷰 영상을 생성하는데 이용되는 원시 영상을 생성할 수 있다. 또한, 제2 프로세서(1710)는 라이브 뷰 영상을 생성하는데 소요되는 시간을 단축시키기 위하여, 인공지능 모델(1730) 내의 모델들 중에서 적어도 하나를 이용하지 않을 수 있다. 제1 프로세서(1500)는 톤 맵 생성 모델(1731), 특징 추출 모델(1732), 영상 보정 모델(1733) 또는 영상 회복 모델(1734) 중 적어도 하나를 이용하지 않을 수 있다. 이 경우, 라이브 뷰 영상의 생성을 위하여 톤 맵 생성 모델(1731), 특징 추출 모델(1732) 및 영상 회복 모델(1734) 중 어느 모델을 이용하지 않을 지는 미리 설정될 수 있다.
- [81] 제2 프로세서(1710)는 적어도 하나의 모델이 비활성화된 인공지능 모델(1730)에 원시 영상을 입력할 수 있으며, 제2 프로세서(1710)는 인공지능 모델(1730)로부터 출력되는 라이브 뷰 영상을 제1 프로세서(1500)에게 제공하며, 제1 프로세서(1500)는 라이브 뷰 영상을 디스플레이부(1200) 상에 디스플레이할 수 있다. 이 경우, 인공지능 모델(1730)은 인공지능 모델(1730) 내의 모델들 중에서 적어도 하나가 비활성화된 상태에서 양호한 라이브뷰 영상이 출력될 수 있도록 미리 훈련될 수 있다.
- [82] 이에 따라, 디바이스(1000)의 카메라 애플리케이션이 실행되어 촬영 기능이

활성화되면, 디바이스(1000)의 적은 리소스를 사용하여 생성되는 라이브 뷰 영상이 디바이스(1000)의 화면 상에 디스플레이될 수 있으며, 피사체를 촬영하는 사용자의 캡처 입력이 수신되면 캡처 입력이 수신될 때 생성된 원시 영상으로부터 인공지능 모델(1730)을 통해 생성되는 출력 영상이 제1 메모리(1600)에 저장될 수 있다.

- [83] 제1 프로세서(1500)는 모델 관리 모듈(1630)을 실행함으로써, 제2 프로세서(1720)에 의해 이용될 인공지능 모델(1730)을 선택하고, 제1 메모리(1600)에 저장된 인공지능 모델(1620)을 업데이트할 수 있다.
- [84] 제1 프로세서(1500)는 모델 선택 모듈(1634)을 실행함으로써, 제1 메모리(1600)에 저장된 적어도 하나의 인공지능 모델(1620) 중에서 제2 프로세서(1720)에 의해 이용될 인공지능 모델(1730)을 선택할 수 있다. 또한, 제1 프로세서(1500)에 의해 선택된 인공지능 모듈(1730)은 AI 프로세싱 유닛(1700)의 제2 메모리(1720)에 로딩될 수 있다.
- [85] 제1 메모리(1600)에는 복수의 인공지능 모델(1620)이 저장될 수 있으며, 복수의 인공지능 모델(1620)은 복수의 상황에 따라 각각 훈련된 인공지능 모델(1620)일 수 있다. 예를 들어, 카메라 필터, 카메라 렌즈, 카메라의 제조사, 디바이스 모델, 복수의 연속 촬영된 이미지, 촬영 환경, 피사체의 종류 또는 촬영 영상의 속성 중 적어도 하나에 관련된 상황을 기준으로 훈련된 인공지능 모델들일 수 있다.
- [86] 예를 들어, 카메라 필터 별로 훈련된 인공지능 모델들은, 비선명 마스크를 이용하여 촬영된 이미지들을 바탕으로 훈련된 인공지능 모델, Contrast 조정 마스크를 이용하여 촬영된 이미지들을 바탕으로 훈련된 인공지능 모델 및 컬러 필터 마스크를 이용하여 촬영된 이미지들을 바탕으로 훈련된 인공지능 모델을 포함할 수 있으나, 이에 제한되지 않는다. 또한, 예를 들어, 카메라 렌즈 별로 훈련된 인공지능 모델들은, 망원 렌즈를 이용하여 촬영된 이미지들을 바탕으로 훈련된 인공지능 모델, 광각 렌즈를 이용하여 촬영된 이미지들을 바탕으로 훈련된 인공지능 모델, 어안 렌즈를 이용하여 촬영된 이미지들을 바탕으로 훈련된 인공지능 모델을 포함할 수 있으나, 이에 제한되지 않는다. 또한, 예를 들어, 카메라의 제조사 별로 훈련된 인공지능 모델들은, 제조사 A의 카메라를 이용하여 촬영된 이미지들을 바탕으로 훈련된 인공지능 모델, B 제조사 B의 카메라를 이용하여 촬영된 이미지들을 바탕으로 훈련된 인공지능 모델, 제조사 C의 카메라를 이용하여 촬영된 이미지들을 바탕으로 훈련된 인공지능 모델을 포함할 수 있으나, 이에 제한되지 않는다. 또한, 예를 들어, 디바이스 별로 훈련된 인공지능 모델들은 갤럭시 S10으로 촬영된 이미지들을 바탕으로 훈련된 인공지능 모델, 갤럭시 S20으로 촬영된 이미지들을 바탕으로 훈련된 인공지능 모델, 갤럭시 노트 20으로 촬영된 이미지들을 바탕으로 훈련된 인공지능 모델을 포함할 수 있으나, 이에 제한되지 않는다.
- [87] 또한, 예를 들어, 촬영 환경 별로 훈련된 인공지능 모델은, 실내에서 촬영된 이미지들을 바탕으로 훈련된 인공지능 모델, 야외에서 촬영된 이미지들을

바탕으로 훈련된 인공지능 모델, 특정 조도 범위에서 촬영된 이미지들을 바탕으로 훈련된 인공지능 모델을 포함할 수 있으나, 이에 제한되지 않는다. 또한, 예를 들어, 피사체의 종류 별로 훈련된 인공지능 모델은, 인물이 촬영된 이미지들을 바탕으로 훈련된 인공지능 모델, 음식이 촬영된 이미지들을 바탕으로 훈련된 인공지능 모델, 건물이 촬영된 이미지들을 바탕으로 훈련된 인공지능 모델을 포함할 수 있으나 이에 제한되지 않는다. 또한, 예를 들어, 촬영 영상의 속성 별로 훈련된 인공지능 모델은, 특정 화이트 밸런스 값을 적용하여 촬영된 이미지들을 바탕으로 훈련된 인공지능 모델, 특정 ISO를 적용하여 촬영된 이미지들을 바탕으로 훈련된 인공지능 모델, 특정 셔터 스피드에서 촬영된 이미지들을 바탕으로 훈련된 인공지능 모델을 포함할 수 있으나, 이에 제한되지 않는다.

- [88] 디바이스(1000)의 카메라 애플리케이션이 실행되고 카메라 기능이 활성화되면 제1 프로세서(1500)는 모델 선택 모듈(1634)을 실행하여 카메라 필터, 카메라 렌즈, 카메라의 제조사, 디바이스 모델, 복수의 연속 촬영된 이미지, 촬영 환경, 피사체의 종류 또는 촬영 영상의 속성 중 적어도 하나에 관련된 상황을 식별할 수 있다. 예를 들어, 제1 프로세서(1500)는 디바이스(1000)의 주변 상황을 센싱한 센싱 값 및 카메라 애플리케이션의 설정 값에 기초하여, 적어도 하나의 상황을 식별할 수 있다.
- [89] 또한, 제1 프로세서(1500)는 사진 촬영의 설정을 위한 소정의 GUI를 디스플레이부(1200) 상에 디스플레이하고, GUI에 대한 사용자 입력에 기초하여 설정된 값에 기초하여, 적어도 하나의 상황을 식별할 수도 있다. 예를 들어, 제1 프로세서(1500)는 도 13a 및 도 13b에서와 같이 피사체를 촬영하는 디바이스(1000)의 촬영 설정을 위한 GUI를 디스플레이부(1200) 상에 디스플레이할 수 있다. 이에 따라, 사용자는 디스플레이부(1200) 상에 디스플레이된 GUI를 통하여, ISO, 셔터 스피드, 화이트 밸런스, 색온도, 틸트, 대비, 채도, 하이라이트 효과, 웨도우 효과 등에 관련된 설정 값을 디바이스(1000)에 입력할 수 있다.
- [90] 제1 프로세서(1500)는, 제1 메모리(1600)에 저장된 인공지능 모델(1620) 중에서, 식별된 적어도 하나의 상황에 대응되는 인공지능 모델(1730)을 추출하여 제2 메모리(1720)에 로딩할 수 있다. 예를 들어, 사용자는 디바이스(1000)의 디스플레이부(1200) 상에 디스플레이된 GUI를 통하여, ISO, 화이트 밸런스, 색온도, 틸트, 채도 및 대비 등에 관련된 설정을 할 수 있으며, 디바이스(1000)는 사용자에게 의해 설정된 값들에 기초하여, 제1 메모리(1600)에 저장된 복수의 인공지능 모델(1620) 중에서 설정 값에 대응되는 인공지능 모델(1730)을 선택하고, 선택된 인공지능 모델(1730)을 제2 메모리(1720)에 로딩할 수 있다.
- [91] 또한, 제1 프로세서(1500)는, 예를 들어, 사용자가 소정 기준 이상 사용한 카메라 필터, 사용자가 소정 기준 이상 사용한 카메라 렌즈, 사용자가 소정 기준 이상 사용한 카메라의 제조사, 사용자가 소정 기준 이상 사용한 디바이스 모델,

소정 횟수 이상의 촬영 환경, 사용자가 소정 기준 이상 촬영된 피사체의 종류 또는 사용자가 소정 기준 이상 사용한 촬영 영상의 속성 중 적어도 하나에 관한 정보에 기초하여, 사용자의 선호에 대응되는 인공지능 모델(1730)을 제1 메모리(1600)로부터 추출하여 제2 메모리(1720)에 로딩할 수 있다.

[92] 또한, 제1 프로세서(1500)는, 예를 들어, 기설정된 이미지들을 디스플레이하고, 디스플레이된 이미지들 중에서 사용자에게 의해 선택된 적어도 하나의 이미지에 기초하여, 사용자에게 의해 선택된 이미지와 유사한 특징의 이미지를 출력하도록 훈련된 인공지능 모델(1730)을 제2 메모리(1720)에 로딩할 수 있다. 예를 들어, 사용자가 옛지가 강조된 이미지를 많이 선택하면 옛지가 강조된 이미지를 출력하도록 훈련된 인공지능 모델(1730)이 제2 메모리(1720)에 로딩될 수 있다.

[93] 만약, 식별된 적어도 하나의 상황에 대응되는 인공지능 모델(1730)이 제1 메모리(1600)에 저장되어 있지 않으면, 제1 프로세서(1500)는 서버(미도시)에게 인공지능 모델(1730)을 요청하고, 서버(미도시)로부터 인공지능 모델(1730)을 수신하여 제2 메모리(1720)에 저장할 수 있다.

[94] 제1 프로세서(1500)는 다운로드 모듈(1631)을 실행함으로써, 리트레이닝된 인공지능 모델(1620), 또는 인공지능 모델(1620)을 리트레이닝하기 위한 데이터를 서버(미도시)로부터 수신할 수 있다. 제1 프로세서(1500)는 리트레이닝된 인공지능 모델(1620)을 서버(미도시)에게 요청할 수 있다. 인공지능 모델(1620)은 서버(미도시)에 의해 리트레이닝될 수 있으며, 서버(미도시)는 리트레이닝된 인공지능 모델(1620)이 존재함을 알리는 알림 정보를 디바이스(1000)에게 제공할 수 있다. 또한, 제1 프로세서(1500)는 서버(미도시)로부터 수신된 알림 정보를 화면 상에 디스플레이하고, 인공지능 모델(1620)을 업데이트하기 위한 사용자 입력을 수신할 수 있다. 제1 프로세서(1500)는 서버(미도시)에게 사용자가 선호하는 사진 속성에 관한 정보를 제공하면서 서버(미도시)에게 리트레이닝된 인공지능 모델(1620)을 요청할 수도 있다. 사용자가 선호하는 사진 속성에 관한 정보는, 예를 들어, 사용자가 소정 기준 이상 사용한 카메라 필터, 사용자가 소정 기준 이상 사용한 카메라 렌즈, 사용자가 소정 기준 이상 사용한 카메라의 제조사, 사용자가 소정 기준 이상 사용한 디바이스 모델, 소정 횟수 이상의 촬영 환경, 사용자가 소정 기준 이상 촬영된 피사체의 종류 또는 사용자가 소정 기준 이상 사용한 촬영 영상의 속성 중 적어도 하나에 관한 정보를 포함할 수 있다. 제1 프로세서(1500)가 사용자가 선호하는 사진 속성에 관한 정보를 서버(미도시)에게 제공한 경우에는, 제1 프로세서(1500)는 사용자가 선호하는 사진 속성과 관련하여 리트레이닝된 인공지능 모델(1620)을 서버(미도시)로부터 다운로드할 수 있다. 또한, 제1 프로세서(1500)는 업데이트 모듈(1632)을 실행함으로써, 디바이스(1000) 내의 인공지능 모델(1620)을 서버(미도시)로부터 수신된 리트레이닝된 인공지능 모델(1620)로 대체함으로써, 디바이스(1000) 내의 인공지능 모델(1620)을 업데이트할 수 있다.

- [95] 제1 프로세서(1500)는 다운로드 모듈(1631)을 실행함으로써, 리트레이닝을 위한 기준 원시 영상 및 기준 원시 영상에 대응되는 기준 영상을 서버(미도시)로부터 다운로드할 수 있다. 기준 영상은 리트레이닝을 위한 기준 원시 영상으로부터 생성된 영상일 수 있다. 제1 프로세서(1500)는 서버(미도시)에게 인공지능 모델(1620)의 리트레이닝을 위한 기준 원시 영상 및 기준 영상을 요청하고, 서버(미도시)로부터 리트레이닝을 위한 기준 원시 영상 및 기준 영상을 수신할 수 있다. 이 경우, 제1 프로세서(1500)는 사용자가 선호하는 사진 속성에 관한 정보를 서버(미도시)에게 제공할 수 있다. 사용자가 선호하는 사진 속성에 관한 정보는, 예를 들어, 사용자가 소정 기준 이상 사용한 카메라 필터, 사용자가 소정 기준 이상 사용한 카메라 렌즈, 사용자가 소정 기준 이상 사용한 카메라의 제조사, 사용자가 소정 기준 이상 사용한 디바이스 모델, 소정 횟수 이상의 촬영 환경, 사용자가 소정 기준 이상 촬영된 피사체의 종류 또는 사용자가 소정 기준 이상 사용한 촬영 영상의 속성 중 적어도 하나에 관한 정보를 포함할 수 있다. 이 경우, 서버(미도시)는 사용자가 선호하는 사진 속성과 관련하여 생성된 기준 영상 및 기준 원시 영상을 디바이스(1000)에게 제공할 수 있다.
- [96] 제1 프로세서(1500)는 리트레이닝 모듈(1633)을 실행함으로써 리트레이닝을 위한 기준 원시 영상 및 기준 원시 영상에 대응되는 기준 영상을 획득하여 제2 프로세서(1720)가 인공지능 모델(1620)을 리트레이닝할 수 있다. 제2 프로세서(1710)는 서버(미도시)로부터 수신된 기준 영상을 정답 영상으로 이용할 수 있으며, 제2 프로세서(1710)는 서버(미도시)로부터 수신된 기준 원시 영상을 인공지능 모델(1620)에 입력하고 인공지능 모델(1620)로부터 출력되는 출력 영상을 기준 영상과 비교함으로써 인공지능 모델(1620)을 리트레이닝할 수 있다.
- [97] 한편, 상기에서는 디바이스(1000)가 디바이스(1000) 내의 인공지능 모델(1620)을 이용하여 피사체가 촬영된 출력 영상을 생성하는 것으로 설명되었지만, 이에 제한되지 않는다. 디바이스(1000)는 서버(미도시)와 함께 피사체가 촬영된 출력 영상을 생성할 수도 있다. 예를 들어, 디바이스(1000)는 원시 영상을 생성하고 생성된 원시 영상을 서버(미도시)로 전송하면서 서버(미도시)에게 출력 영상을 요청할 수 있다. 디바이스(1000)는 원시 영상 및 영상 보정을 위한 설정 정보를 서버(미도시)에게 함께 제공할 수도 있다. 이 경우, 인공지능 모델(1620)은 서버(미도시)에 포함될 수 있으며, 서버(미도시)는 서버(미도시) 내의 인공지능 모델(1620)을 이용하여 원시 영상으로부터 출력 영상을 생성할 수 있다. 또한, 서버(미도시)는 생성된 출력 영상을 디바이스(1000)에게 제공할 수 있다.
- [98] 또는, 예를 들어, 디바이스(1000)는 톤 맵 생성 모델(1731)로부터 출력된 톤 맵 및 원시 영상을 서버(미도시)에게 제공하면서 서버(미도시)에게 출력 영상을 요청할 수 있다. 디바이스(1000)는 원시 영상, 톤 맵 및 영상 보정을 위한 설정

정보를 서버(미도시)에게 함께 제공할 수도 있다. 이 경우, 인공지능 모델(1620) 내의 특징 추출 모델(1732), 영상 보정 모델(1733) 및 영상 회복 모델(1734)은 서버(미도시)에 포함될 수 있으며, 서버(미도시)는 서버(미도시) 내의 특징 추출 모델(1732), 영상 보정 모델(1733) 및 영상 회복 모델(1734)을 이용하여 출력 영상을 생성할 수 있다. 또한, 서버(미도시)는 생성된 출력 영상을 디바이스(1000)에게 제공할 수 있다.

- [99] 또는, 예를 들어, 디바이스(1000)는 특징 추출 모델(1732)로부터 출력된 특징 영상들을 서버(미도시)에게 제공하면서 서버(미도시)에게 출력 영상을 요청할 수 있다. 디바이스(1000)는 특징 영상들 및 영상 보정을 위한 설정 정보를 서버(미도시)에게 함께 제공할 수도 있다. 이 경우, 인공지능 모델(1620) 내의 영상 보정 모델(1733) 및 영상 회복 모델(1734)은 서버(미도시)에 포함될 수 있으며, 서버(미도시)는 서버(미도시) 내의 영상 보정 모델(1733) 및 영상 회복 모델(1734)을 이용하여 출력 영상을 생성할 수 있다. 또한, 서버(미도시)는 생성된 출력 영상을 디바이스(1000)에게 제공할 수 있다.
- [100] 또한, 디바이스(1000)는 디바이스(1000)의 상황에 따라, 디바이스(1000)가 혼자 출력 영상을 생성할지, 디바이스(1000)가 서버(미도시)와 함께 출력 영상을 생성할지를 결정할 수 있다. 예를 들어, 디바이스(1000)는 배터리 잔량, 리소스 사용량 및 통신 상태 등을 고려하여 디바이스(1000)가 혼자서 출력 영상을 생성할지, 디바이스(1000)가 서버(미도시)와 함께 출력 영상을 생성할지를 결정할 수 있다. 예를 들어, 디바이스(1000)의 배터리 잔량이 임계치보다 작은 경우에, 디바이스(1000)는 서버(미도시)에게 출력 영상을 생성할 것을 요청할 수 있다. 또한, 예를 들어, 디바이스(1000)의 리소스 사용량이 임계치보다 큰 경우에, 디바이스(1000)는 서버(미도시)에게 출력 영상을 생성할 것을 요청할 수 있다. 또한, 예를 들어, 디바이스(1000)는 디바이스(1000)의 통신 상태가 양호한 경우에, 서버(미도시)에게 출력 영상을 생성할 것을 요청할 수 있다. 이 경우, 디바이스(1000)가 서버(미도시)에게 출력 영상을 요청하면서, 원시 영상을 제공할지, 원시 영상 및 톤 맵을 제공할지, 특징 영상들을 제공할지는, 디바이스(1000)의 상황에 따른 다양한 기준에 따라 설정될 수 있다. 예를 들어, 촬영 기능이 활성화되면, 디바이스(1000)는 배터리 잔량, 리소스 사용량 및 통신 상태 등을 식별할 수 있으며, 서버(미도시)에게 원시 영상을 제공할지, 원시 영상 및 톤 맵을 제공할지, 특징 영상들을 제공할지를 결정할 수 있다. 또한, 디바이스(1000)는 결정에 기초하여 원시 영상, 톤 맵, 특징 영상들 중 적어도 하나를 서버(미도시)에게 제공하면서 출력 영상을 제공해 줄 것을 요청할 수 있다.
- [101] 도 3은 본 개시의 일 실시예에 따른 원시 영상으로부터 피사체가 촬영된 출력 영상이 생성되는 과정을 설명하는 도면이다.
- [102] 도 3을 참조하면, 디바이스(1000)는, AI 프로세싱 유닛(1700)을 이용하여, 피사체에 대한 원시 영상(30)을 톤 맵 생성 모델(1731)에 입력하고 톤 맵 생성

모델(1731)로부터 출력되는 톤 맵을 획득할 수 있다. 또한, 디바이스(1000)는 원시 영상(30) 및 톤 맵을 조합하여 특징 추출 모델(1732)에 입력하고 특징 추출 모델(1732)로부터 출력되는 복수의 특징 영상들을 획득할 수 있다. 이후, 디바이스(1000)는 복수의 특징 영상들을 영상 보정 모델(1733)에 입력하고, 영상 보정 모델(1733)로부터 출력되는 보정된 특징 영상들을 획득할 수 있다. 이후, 디바이스(1000)는 보정된 특징 영상들을 영상 회복 모델(1734)에 입력함으로써 영상 회복 모델(1734)로부터 출력되는 출력 영상(38)을 획득할 수 있다.

- [103] 톤 맵 생성 모델(1731), 특징 추출 모델(1732), 영상 보정 모델(1733) 및 영상 회복 모델(1734) 각각은, 복수의 신경망 레이어들로 구성될 수 있다. 복수의 신경망 레이어들 각각은 복수의 가중치들(weight values)을 갖고 있으며, 이전(previous) 레이어의 연산 결과와 복수의 가중치들 간의 연산을 통해 신경망 연산을 수행한다. 복수의 신경망 레이어들이 갖고 있는 복수의 가중치들은 인공지능 모델의 학습 결과에 의해 최적화될 수 있다. 예를 들어, 학습 과정 동안 인공지능 모델에서 획득한 로스(loss) 값 또는 코스트(cost) 값이 감소 또는 최소화되도록 복수의 가중치들이 갱신될 수 있다. 인공 신경망은 심층 신경망(DNN:Deep Neural Network)를 포함할 수 있으며, 예를 들어, CNN (Convolutional Neural Network), DNN (Deep Neural Network), RNN (Recurrent Neural Network), RBM (Restricted Boltzmann Machine), DBN (Deep Belief Network), BRDNN(Bidirectional Recurrent Deep Neural Network) 또는 심층 Q-네트워크 (Deep Q-Networks) 등이 있으나, 전술한 예에 한정되지 않는다. 톤 맵 생성 모델(1731), 특징 추출 모델(1732), 영상 보정 모델(1733) 및 영상 회복 모델(1734)은, 후술할 도 7에서와 같이, 기준 원시 영상 및 기준 출력 영상을 이용하여 함께 훈련될 수 있다.
- [104] 도 4a는 본 개시의 일 실시예에 따른 디바이스(1000)가 원시 영상(30)으로부터 톤 맵을 생성하는 과정을 설명하는 도면이다.
- [105] 도 4a를 참조하면, 디바이스(1000)의 카메라 애플리케이션이 실행되고 촬영 기능이 활성화됨에 따라, 디바이스(1000)는 원시 영상(30)을 생성할 수 있다. 디바이스(1000)는 피사체를 촬영하기 위한 기능이 활성화됨에 따라 카메라 센서부(1400)를 이용하여 피사체로부터 제공되는 광에 기초한 원시 영상들을 생성할 수 있으며, 피사체를 촬영하는 사용자 입력이 수신됨에 따라 사용자 입력이 수신된 때에 카메라 센서부(1400)에 의해 생성된 원시 영상(30)을 획득할 수 있다. 원시 영상(30)은, 예를 들어, 적색 값, 녹색 값 및 청색 값의 배열로 구성될 수 있다.
- [106] 또한, 디바이스(1000)는 촬영 기능이 활성화됨에 따라, 피사체 촬영을 위한 적어도 하나의 상황을 식별할 수 있으며, 식별된 적어도 하나의 상황에 대응되는 인공지능 모델(1720)을 제1 메모리(1600)로부터 추출하여 AI 프로세싱 유닛(1700) 내의 제2 메모리(1720)에 로딩할 수 있다. 만약, 식별된 적어도 하나의 상황에 대응되는 인공지능 모델(1720)이 제1 메모리(1600)에 저장되어 있지 않은

경우에는, 디바이스(1000)는 식별된 적어도 하나의 상황에 대응되는 인공지능 모델(1720)을 서버(미도시)에게 요청하고, 서버(미도시)로부터 인공지능 모델(1720)을 수신하여 제2 메모리(1720)에 로딩할 수 있다.

- [107] 이후, 디바이스(1000)는 원시 영상(30) 내의 적색 값, 녹색 값 및 청색 값의 밝기(brightness)를 스케일링하기 위하여 이용되는 톤 맵(40)을 생성하기 위하여, 카메라 애플리케이션이 실행되고 피사체를 촬영하는 사용자 입력을 수신하고 수신된 사용자 입력에 응답하여, 원시 영상(30)을 톤 맵 생성 모델(1731)에 입력할 수 있다. 톤 맵(40)은 원시 영상(30) 내의 픽셀들의 밝기를 스케일링하기 위한 정보를 포함할 수 있다. 또한, 예를 들어, 톤 맵(40)은, 원시 영상(30) 내에서 어두운 영역의 밝기를 밝은 영역의 밝기보다 크게 스케일링하도록 생성될 수 있다.
- [108] 톤 맵 생성 모델(1731)은 복수의 신경망 레이어들로 구성될 수 있으며, 복수의 신경망 레이어들 각각은 복수의 가중치들(weight values)을 갖고 있으며, 이전(previous) 레이어의 연산 결과와 복수의 가중치들 간의 연산을 통해 신경망 연산을 수행할 수 있다. 복수의 신경망 레이어들이 갖고 있는 복수의 가중치들은 톤 맵 생성 모델(1731)의 훈련 결과에 의해 최적화될 수 있으며, 예를 들어, 톤 맵 생성 모델(1731)은 특징 추출 모델(1732), 영상 보정 모델(1733) 및 영상 회복 모델(1734)과 함께 훈련될 수 있다. 또한, 예를 들어, 톤 맵 생성 모델(1731)은 CNN (Convolutional Neural Network)를 포함할 수 있으나, 이에 제한되지 않는다.
- [109] 도 4b는 본 개시의 일 실시예에 따른 디바이스(1000)가 원시 영상(30)으로부터 특징들을 추출하는 과정을 설명하는 도면이다.
- [110] 도 4b를 참조하면, 디바이스(1000)는 원시 영상(30) 및 톤 맵(40)을 조합하여 특징 추출 모델(1732)에 입력하고 특징 추출 모델(1732)로부터 출력되는 복수의 특징 영상들(42)을 획득할 수 있다. 예를 들어, 디바이스(1000)는 톤 맵(40)을 이용하여 원시 영상(30) 내의 픽셀들의 밝기를 스케일링하고, 밝기가 스케일링된 원시 영상(30)을 특징 추출 모델(1732)에 입력할 수 있다. 특징 추출 모델(1732)에 입력되는 원시 영상(30)은, 어두운 부분의 픽셀이 밝은 값을 가지도록 스케일링될 수 있으며 이에 따라, 원시 영상(30) 내의 어두운 부분에 위치한 피사체가 보다 효과적으로 식별될 수 있게 된다.
- [111] 특징 추출 모델(1732)은 특징 추출 모델(1732)에 입력된 원시 영상(30) 내의 특징들을 추출할 수 있다. 원시 영상(30) 내의 복수의 특징들을 각각 나타내는 복수의 특징 영상들(42)이 특징 추출 모델(1732)로부터 출력될 수 있다. 예를 들어, 복수의 특징 영상들(42)은, 원시 영상(30) 내의 엣지에 관한 특징을 나타내는 특징 영상, 원시 영상(30) 내의 라인에 관한 특징을 나타내는 특징 영상, 원시 영상(30) 내의 공간에 관련된 특징을 나타내는 특징 영상, 원시 영상(30) 내의 객체의 형상 및 깊이에 관련된 특징을 나타내는 특징 영상, 원시 영상(30) 내의 인물에 관련된 특징을 나타내는 특징 영상 및 원시 영상(30) 내의 사물에 관련된 특징을 나타내는 특징 영상 등을 포함할 수 있으나, 이에 제한되지

않는다.

- [112] 특징 추출 모델(1732)은 복수의 신경망 레이어들로 구성될 수 있으며, 복수의 신경망 레이어들 각각은 복수의 가중치들(weight values)을 갖고 있으며, 이전(previous) 레이어의 연산 결과와 복수의 가중치들 간의 연산을 통해 신경망 연산을 수행할 수 있다. 복수의 신경망 레이어들이 갖고 있는 복수의 가중치들은 특징 추출 모델(1732)의 훈련 결과에 의해 최적화될 수 있으며, 예를 들어, 특징 추출 모델(1732)은 톤 맵 생성 모델(1731), 영상 보정 모델(1733) 및 영상 회복 모델(1734)과 함께 훈련될 수 있다. 또한, 예를 들어, 특징 추출 모델(1732)은 도 5와 같이 End-to-End 방식의 Fully-Convolutional Network 기반 구조를 가지는 U-NET에 의해 구현될 수 있으나, 이에 제한되지 않는다.
- [113] 도 4c는 본 개시의 일 실시예에 따른 디바이스(1000)가 특징 영상들을 보정하는 과정을 설명하는 도면이다.
- [114] 도 4c를 참조하면, 디바이스(1000)는 특징 추출 모델(1732)로부터 출력된 특징 영상들(42)을 보정할 수 있다. 디바이스(1000)는 특징 추출 모델(1732)로부터 출력된 특징 영상들(42) 및 영상 속성들에 관련한 설정 값들을 영상 보정 모델(1733)에 입력하여 특징 영상들(42)을 보정할 수 있다. 예를 들어, 디바이스(1000)는 화이트 밸런스의 보정 및 색상 보정을 위한 기설정된 기준에 따라, 특징 영상들(42)을 보정할 수 있다.
- [115] 예를 들어, 디바이스(1000)의 제1 프로세서(1500)는 디바이스(1000) 내의 센서를 이용하여 피사체를 촬영하는 때의 디바이스(1000) 주변의 환경을 나타내는 센싱 데이터를 생성하며, 생성된 센싱 데이터에 따라 디바이스(1000)의 주변 환경을 자동으로 식별하고, 미리 설정된 기준에 기초하여 특징 영상의 보정을 위한 화이트 밸런스 매트릭스(44) 및 색상 보정 매트릭스(45)를 제1 메모리(1600)로부터 추출할 수 있다. 또한, 제1 프로세서(1500)는 추출된 화이트 밸런스 매트릭스(44) 및 색상 보정 매트릭스(45)를 제2 프로세서(1720)에게 제공할 수 있다. 또한, 제2 프로세서(1720)는 제1 프로세서(1500)로부터 제공받은 화이트 밸런스 매트릭스(44) 및 색상 보정 매트릭스(45)를 영상 보정 모델(1733)에 특징 영상들(42)과 함께 입력할 수 있다.
- [116] 또는, 예를 들어, 디바이스(1000)의 제1 프로세서(1500)는 촬영 기능이 활성화됨에 따라 특징 영상들(42)의 화이트 밸런스, 색상을 보정하기 위한 설정을 위하여 GUI를 디바이스(1000)의 디스플레이부(1200) 상에 디스플레이할 수 있으며, GUI를 통한 사용자 입력에 기초하여 화이트 밸런스의 보정 및 색상 보정을 위한 설정을 미리 수행할 수 있다. 또한, 제1 프로세서(1500)는 사용자 입력에 따른 설정에 기초하여 제1 메모리(1600)로부터 화이트 밸런스 매트릭스(44) 및 색상 보정 매트릭스(45)를 추출하고, 화이트 밸런스 매트릭스(44) 및 색상 보정 매트릭스(45)를 제2 프로세서(1720)에게 제공할 수 있다. 또한, 제2 프로세서(1720)는 제1 프로세서(1500)로부터 제공받은 화이트 밸런스 매트릭스(44) 및 색상 보정 매트릭스(45)를 영상 보정 모델(1733)에 특징

영상들(42)과 함께 입력할 수 있다.

- [117] 상기에서는 디바이스(1000)가 영상 보정 모델(1733)을 이용하여 특징 영상들(42)을 보정하는 것으로 설명되었지만, 이에 제한되지 않는다. 디바이스(1000)는 영상 보정 모델(1733)을 이용하지 않고 특징 영상들(42)을 보정할 수 있다. 이 경우, 디바이스(1000)는 영상 속성을 보정하기 위한 매트릭스를 이용하여 특징 영상들을 보정할 수 있다. 예를 들어, 디바이스(1000)는 화이트 밸런스를 보정하기 위한 매트릭스(44)를 특징 영상들(42)에 각각 곱함(multiply)으로써 특징 영상들(42)의 화이트 밸런스를 보정할 수 있다. 또한, 디바이스(1000)는 색상을 보정하기 위한 매트릭스(45)를 특징 영상들(42)에 각각 곱함으로써 특징 영상들(42)의 색상을 보정할 수 있다.
- [118] 한편, 상기에서는, 화이트 밸런스 매트릭스 값 및 색상 보정 매트릭스 값 값이 영상 보정 모델(1733)에 입력되는 것으로 설명되었지만 이에 제한되지 않는다. 복수의 인공지능 모델(1620)이 화이트 밸런스 및 색상 보정에 관한 설정 별로 각각 훈련될 수 있다. 이 경우에는, 출력 영상을 생성하기 위해 소정의 화이트 밸런스 및 색상 보정에 관한 설정에 대응되는 인공지능 모델(1730)이 제2 메모리(1720)에 로딩되어 제2 프로세서(1710)에 의해 이용될 수 있다. 또한, 제2 메모리(1720)에 로딩된 인공지능 모델(1730)에는 화이트 밸런스 매트릭스 값 및 색상 보정 매트릭스 값이 입력되지 않더라도, 제2 메모리(1720)에 로딩된 인공지능 모델(1730)로부터 소정의 화이트 밸런스 및 색상 보정에 관한 설정을 고려한 출력 영상이 출력될 수 있게 된다.
- [119] 도 4d는 본 개시의 일 실시예에 따른 디바이스(1000)가 보정된 특징 영상들로부터 출력 영상을 생성하는 과정을 설명하는 도면이다.
- [120] 디바이스(1000)는 보정된 특징 영상들(46)을 영상 회복 모델(1734)에 입력하고, 영상 회복 모델(1734)로부터 출력되는 출력 영상(38)을 획득할 수 있다. 영상 회복 모델(1734)로부터 출력되는 출력 영상(38)은 피사체를 촬영한 결과물로서 디바이스(1000)에 저장될 영상일 수 있다. 영상 회복 모델(1734)은 복수의 신경망 레이어들로 구성될 수 있으며, 복수의 신경망 레이어들 각각은 복수의 가중치들(weight values)을 갖고 있으며, 이전(previous) 레이어의 연산 결과와 복수의 가중치들 간의 연산을 통해 신경망 연산을 수행할 수 있다. 복수의 신경망 레이어들이 갖고 있는 복수의 가중치들은 영상 회복 모델(1734)의 훈련 결과에 의해 최적화될 수 있으며, 예를 들어, 영상 회복 모델(1734)은, 톤 맵 생성 모델(1731) 및 특징 추출 모델(1732)과 함께 훈련될 수 있다. 또한, 예를 들어, 영상 회복 모델(1734)은 CNN (Convolutional Neural Network)를 포함할 수 있으나, 이에 제한되지 않는다.
- [121] 도 5는 본 개시의 일 실시예에 따른 특징 추출 모델(1732)의 구조의 일례를 나타내는 도면이다.
- [122] 도 5를 참조하면, 특징 추출 모델(1732)은 End-to-End 방식의 Fully-Convolutional Network 기반 구조를 가지는 U-NET에 의해 구현될 수 있다. 특징 추출

모델(1732)은 입력된 영상의 컨텍스트를 포착하기 위한 목적으로 구성된 수축 단계(contraction path)의 레이어들 및 수축 단계의 특징 영상으로부터 높은 해상도의 결과를 얻기 위한 업 샘플링(Up-Sampling)을 수행하는 팽창 단계(expanding path)의 레이어들을 포함할 수 있다. 또한, 수축 단계(contraction path)의 레이어들 및 팽창 단계(expanding path)의 레이어들은 서로 대칭되는 구조를 가질 수 있다.

- [123] 도 6은 본 개시의 일 실시예에 따른 디바이스(1000)가 피사체를 촬영하여 출력 영상을 생성하는 방법의 흐름도이다.
- [124] 동작 S600에서 디바이스(1000)는 디바이스(1000) 주변의 피사체를 촬영하는 사용자 입력을 수신하고, 동작 S605에서 디바이스(1000)는 카메라 센서부(1400)를 통하여 입력되는 광에 기초하여, 원시 영상을 생성할 수 있다. 디바이스(1000)는 피사체를 촬영하기 위한 기능이 활성화됨에 따라 카메라 센서부(1400)를 이용하여 피사체로부터 제공되는 광에 기초한 원시 영상들을 생성할 수 있으며, 피사체를 촬영하는 사용자 입력이 수신됨에 따라 사용자 입력이 수신된 때에 카메라 센서부(1400)에 의해 생성된 원시 영상을 획득할 수 있다.
- [125] 동작 S610에서 디바이스(1000)는 원시 영상을 톤 맵 생성 모델(1731)에 입력할 수 있다. 디바이스(1000)는 원시 영상을 톤 맵 생성 모델(1731)에 입력하고, 톤 맵 생성 모델(1731)로부터 출력되는 톤 맵을 획득할 수 있다.
- [126] 동작 S615에서 디바이스(1000)는 생성된 톤 맵 및 원시 영상을 특징 추출 모델(1732)에 입력할 수 있다. 디바이스(1000)는 톤 맵을 이용하여 원시 영상 내의 픽셀들의 밝기를 스케일링하고, 밝기가 스케일링된 원시 영상을 특징 추출 모델에 입력할 수 있다. 또한, 디바이스(1000)는 특징 추출 모델(1732)로부터 출력되는 특징 영상들을 획득할 수 있다.
- [127] 동작 S620에서 디바이스(1000)는 특징 추출 모델(1732)에 의해 생성된 특징 영상들을 기설정된 기준에 따라 보정할 수 있다. 디바이스(1000)는 기설정된 영상 속성들에 관련한 설정에 기초하여 특징 영상들을 보정할 수 있다. 예를 들어, 디바이스(1000)는 화이트 밸런스의 보정 및 색상 보정을 위한 기설정된 기준에 따라, 특징 영상들을 보정할 수 있으나, 특징 영상의 보정을 위한 기설정된 기준은 이에 제한되지 않는다.
- [128] 동작 S625에서 디바이스(1000)는 보정된 특징 영상들을 영상 회복 모델(1734)에 입력하고, 동작 S630에서 디바이스(1000)는 영상 회복 모델(1734)로부터 출력되는 출력 영상을 저장할 수 있다.
- [129] 도 7은 본 개시의 일 실시예에 따른 인공지능 모델(1620)이 훈련되는 예시를 나타내는 도면이다.
- [130] 도 7을 참조하면, 기준 원시 영상에 대한 ISP 프로세싱을 통해 생성되는 기준 영상(72)이 인공지능 모델(1620)의 훈련을 위한 정답 영상(GT image: Ground Truth image)으로 이용될 수 있다. 예를 들어, 피사체가 짧은 시간 구간동안 연속

촬영됨으로써 생성되는 복수의 원시 영상들(70)이 조합됨으로써 기준 원시 영상이 생성될 수 있으며, 기준 원시 영상에 대한 기존의 ISP 프로세싱을 통해 기준 영상(72)이 출력될 수 있다. 기존의 ISP 프로세싱은 인공지능 모델을 이용하지 않고 원시 영상을 영상 처리하는 기존의 프로세싱일 수 있으며, 예를 들어, 원시 영상에 대한 전처리, 화이트 밸런스 조정, 디모자이킹, 감마 보정, 색 변환 등의 처리를 인공지능 모델을 이용하지 않고 수행하는 이미지 프로세싱일 수 있다. 출력된 기준 영상(72)은 인공지능 모델의 정답 영상(GT image: Ground Truth image)으로 이용될 수 있다. 또는, 피사체가 짧은 시간 구간동안 연속 촬영됨으로써 생성되는 복수의 원시 영상들(70) 중 하나가 선택되고, 선택된 원시 영상으로부터 기존의 ISP 프로세싱을 통해 출력된 영상이 정답 영상으로 이용될 수 있다.

[131] 인공지능 모델(1620)의 훈련을 위하여, 연속 촬영된 복수의 원시 영상들(70) 중 하나 또는 기준 원시 영상이 인공지능 모델(1620)에 입력될 수 있다. 예를 들어, 버스트 샷에 의해 8장의 원시 영상들이 생성되는 경우에, 8장의 원시 영상을 조합하여 기준 원시 영상이 생성될 수 있으며, 기준 원시 영상 및 기준 원시 영상으로부터 기존의 ISP 프로세싱을 통해 생성되는 기준 영상이 인공지능 모델(1620)의 훈련에 이용될 수 있다. 또는, 예를 들어, 버스트 샷에 의해 8장의 원시 영상들이 생성되는 경우에, 8장의 원시 영상들 각각으로부터 기존의 ISP 프로세싱을 통해 8장의 출력 영상들이 생성될 수 있다. 이 경우, 8장의 원시 영상들 중 하나가 선택되고, 선택된 원시 영상 및 선택된 원시 영상에 대응되는 출력 영상이 인공지능 모델(1620)의 훈련을 위하여 이용될 수 있다.

[132] 또한, 특징 영상들의 영상 속성을 보정하기 위한 설정 정보가 인공지능 모델(1620)에 입력될 수 있다. 영상 속성을 보정하기 위한 설정 정보는, 예를 들어, 기설정된 화이트 밸런스 보정을 위한 매트릭스 및 기설정된 색상 보정을 위한 매트릭스를 포함할 수 있다. 다양한 영상 속성에 대하여 인공지능 모델(1620)이 훈련될 수 있도록 인공지능 모델(1620)에 다양한 설정에 따른 설정 정보가 입력될 수 있다.

[133] 또한, 소정의 노이즈가 포함된 기준 원시 영상이 인공지능 모델(1620)의 훈련을 위하여 인공지능 모델(1620)에 입력될 수 있다. 이 경우, 점차적으로 많은 노이즈를 포함하는 기준 원시 영상들을 인공지능 모델(1620)에 입력하여 인공지능 모델(1620)이 훈련될 수 있다. 예를 들어, 서버(미도시) 또는 디바이스(1000)는 기준 원시 영상에 제1~n 레벨의 노이즈를 각각 포함시켜, n 개의 입력 영상을 생성하고 n 개의 입력 영상을 인공지능 모델(1620)에 각각 입력하여 인공지능 모델(1620)을 훈련시킬 수 있다. 이에 따라, 훈련된 인공지능 모델(1620)은 노이즈가 포함된 원시 영상으로부터 디노이징된 출력 영상을 출력하도록 훈련될 수 있다.

[134] 또한, 기준 영상(72) 및 인공지능 모델(1620)로부터 출력되는 출력 영상(74)을 비교함으로써 기준 영상(72)와 출력 영상(74) 간의 로스(loss)를 감소시킬 수

있도록 인공지능 모델(1620) 내의 신경망 레이어들의 가중치들이 튜닝될 수 있다. 이 경우, 인공지능 모델(1620)은 톤 맵 생성 모델(1731), 특징 추출 모델(1732) 및 영상 회복 모델(1734)을 포함할 수 있으며, 이에 따라, 톤 맵 생성 모델(1731) 내의 신경망 레이어들의 가중치들, 특징 추출 모델(1732) 내의 신경망 레이어들의 가중치들 및 영상 회복 모델(1734) 내의 신경망 레이어들의 가중치들이 함께 튜닝될 수 있다.

- [135] 도 8은 본 개시의 일 실시예에 따른 인공지능 모델(1620)을 훈련하는 방법의 흐름도이다.
- [136] 동작 S800에서 서버(미도시)는 피사체를 연속 촬영함으로써 생성되는 복수의 원시 영상들을 획득할 수 있다. 서버(미도시)는 짧은 시간 구간동안 피사체를 연속 촬영함으로써 생성되는 복수의 원시 영상들을 획득할 수 있다. 복수의 원시 영상들은 연속 촬영을 통하여 생성되므로, 복수의 원시 영상들의 영상 정보는 서로 유사할 수 있다.
- [137] 동작 S805에서 서버(미도시)는 복수의 원시 영상들을 조합하여 기준 원시 영상을 생성할 수 있다. 서버(미도시)는 영상 융합(fusion) 기법을 이용하여 복수의 원시 영상들을 조합함으로써 하나의 기준 원시 영상을 생성할 수 있다.
- [138] 동작 S810에서 서버(미도시)는 기준 원시 영상으로부터 ISP 프로세싱을 통해 생성되는 기준 영상을 획득할 수 있다. ISP 프로세싱은 인공지능 모델을 이용하지 않고 원시 영상을 영상 처리하는 기존의 프로세싱일 수 있으며, 예를 들어, 원시 영상에 대한 전처리, 화이트 밸런스 조정, 디모자이킹, 감마 보정, 색 변환 등의 처리를 인공지능 모델을 이용하지 않고 수행하는 이미지 프로세싱일 수 있다. ISP 프로세싱을 통해 생성된 기준 영상은 인공지능 모델(1620)의 정답 영상(GT image: Ground Truth image)으로 이용될 수 있다.
- [139] 동작 S815에서 서버(미도시)는 인공지능 모델(1620)로부터 출력되는 제1 출력 영상을 획득할 수 있다. 서버(미도시)는 복수의 원시 영상들 중 하나 또는 기준 원시 영상을 인공지능 모델(1620)에 입력하고, 인공지능 모델(1620)로부터 출력되는 제1 출력 영상을 획득할 수 있다. 서버(미도시)에 의해 훈련될 인공지능 모델(1620)은 도 3에 기재된 인공지능 모델(1620)의 구성들을 포함할 수 있다.
- [140] 동작 S820에서 서버(미도시)는 기준 영상 및 제1 출력 영상 간의 로스(loss)를 분석할 수 있다. 서버(미도시)는 기준 영상을 정답 영상(Ground Truth image)로 이용할 수 있으며, 기준 영상 및 제1 출력 영상 간의 차이를 비교할 수 있다.
- [141] 동작 S825에서 서버(미도시)는 분석된 로스에 기초하여, 인공지능 모델(1620)의 가중치를 변경할 수 있다. 서버(미도시)는 기준 영상 및 제1 출력 영상 간의 로스(loss)를 감소시킬 수 있도록 인공지능 모델(1620) 내의 신경망 레이어들의 가중치들을 조정할 수 있다. 이 경우, 인공지능 모델(1620)은 톤 맵 생성 모델(1731), 특징 추출 모델(1732) 및 영상 회복 모델(1734)을 포함할 수 있으며, 이에 따라, 톤 맵 생성 모델(1731) 내의 신경망 레이어들의 가중치들, 특징 추출 모델(1732) 내의 신경망 레이어들의 가중치들 및 영상 회복 모델(1734)

내의 신경망 레이어들의 가중치들이 함께 조정될 수 있다.

- [142] 동작 S830에서 서버(미도시)는 원시 영상들 중 적어도 하나 또는 기준 원시 영상을 가중치가 변경된 인공지능 모델(1620)에 입력하고, 동작 S835에서 서버(미도시)는 가중치가 변경된 인공지능 모델(1620)로부터 출력되는 제2 출력 영상을 획득하고, 동작 S840에서 서버(미도시)는 기준 영상 및 제2 출력 영상 간의 로스를 분석할 수 있다.
- [143] 또한, 동작 S845에서 서버(미도시)는 인공지능 모델(1620)의 훈련을 종료할 지를 결정할 수 있다. 동작 S840에서 분석된 로스가 기설정된 임계치보다 작을 경우에 서버(미도시)는 인공지능 모델(1620)의 훈련을 종료할 것을 결정할 수 있으며, 동작 S840에서 분석된 로스가 기설정된 임계치보다 큰 경우에 서버(미도시)는 인공지능 모델(1620)의 가중치를 변경하기 위한 동작들을 반복하여 수행할 수 있다.
- [144] 한편, 복수의 상황에 따라 복수의 인공지능 모델들(1620)이 훈련될 수 있다. 예를 들어, 카메라 필터, 카메라 렌즈, 카메라의 제조사, 디바이스 모델, 복수의 연속 촬영된 이미지, 촬영 환경, 피사체의 종류 또는 촬영 영상의 속성 중 적어도 하나에 관련된 상황을 기준으로 복수의 인공지능 모델들(1620)이 훈련될 수 있다.
- [145] 예를 들어, 비선명 마스크를 이용하여 촬영된 기준 원시 영상들 및 촬영 영상들을 이용하여 인공지능 모델(1620)이 훈련되고, Contrast 조정 마스크를 이용하여 촬영된 기준 원시 영상들 및 촬영 영상들을 이용하여 인공지능 모델(1620)이 훈련되고, 컬러 필터 마스크를 이용하여 촬영된 기준 원시 영상들 및 촬영 영상들을 이용하여 인공지능 모델(1620)이 훈련될 수 있으나, 이에 제한되지 않는다. 또한, 예를 들어, 카메라 렌즈 별로 촬영된 기준 원시 영상들 및 촬영 영상들을 이용하여 인공지능 모델(1620)이 훈련될 수 있으며, 카메라 렌즈들은, 예를 들어, 망원 렌즈, 광각 렌즈, 어안 렌즈 등을 포함할 수 있으나, 이에 제한되지 않는다. 또한, 예를 들어, 카메라의 제조사 별로 카메라에 의해 촬영된 기준 원시 영상들 및 촬영 영상들을 이용하여 인공지능 모델(1620)이 훈련될 수 있으며, 디바이스의 모델 별로 디바이스에 의해 촬영된 기준 원시 영상들 및 촬영 영상들을 이용하여 인공지능 모델(1620)이 훈련될 수 있으나, 이에 제한되지 않는다. 또한, 예를 들어, 실내에서 촬영된 기준 원시 영상들 및 촬영 영상들을 이용하여 인공지능 모델(1620)이 훈련되고, 야외에서 촬영된 기준 원시 영상들 및 촬영 영상들을 이용하여 인공지능 모델(1620)이 훈련되고, 특정 조도 범위에서 촬영된 기준 원시 영상들 및 촬영 영상들을 이용하여 인공지능 모델(1620)이 훈련될 수 있으나, 이에 제한되지 않는다. 또한, 예를 들어, 인물이 촬영된 기준 원시 영상들 및 촬영 영상들을 이용하여 인공지능 모델(1620)이 훈련되고, 음식이 촬영된 기준 원시 영상들 및 촬영 영상들을 이용하여 인공지능 모델(1620)이 훈련되고, 건물이 촬영된 기준 원시 영상들 및 촬영 영상들을 이용하여 인공지능 모델(1620)이 훈련될 수 있으나 이에 제한되지 않는다. 또한,

예를 들어, 촬영 영상의 속성 별로 촬영된 기준 원시 영상들 및 촬영 영상들을 이용하여 인공지능 모델(1620)이 훈련될 수 있으며, 예를 들어, 촬영 영상의 속성은, 화이트 밸런스, ISO, 셔터 스피드 등을 포함할 수 있으나, 이에 제한되지 않는다.

[146] 한편, 도 8에서는 인공지능 모델(1620)의 훈련이 서버(미도시)에 의해 수행되는 것으로 설명되었지만, 이에 제한되지 않는다. 인공지능 모델(1620)의 훈련은 디바이스(1000)에 의해 수행될 수도 있다. 이 경우, 디바이스(1000)는 인공지능 모델(1620)의 훈련을 위하여 이용될 기준 원시 영상 및 기준 영상을 직접 생성하거나 서버(미도시)에게 요청하여 수신할 수 있다.

[147] 예를 들어, 디바이스(1000)는 디바이스(1000)에서 촬영된 영상들을 서버(미도시)에게 제공하면서 서버(미도시)에게 인공지능 모델(1620)의 훈련을 위한 기준 원시 영상 및 기준 영상을 요청할 수 있다. 서버(미도시)는 디바이스(1000)로부터 수신된 영상들을 분석하여 디바이스(1000)의 사용자가 선호하는 상황에 관련된 기준 원시 영상 및 기준 영상을 디바이스(1000)에게 제공할 수 있다. 예를 들어, 서버(미도시)는 디바이스(1000)로부터 수신된 영상들을 분석함으로써, 사용자가 이용한 디바이스(1000)의 종류, 디바이스(1000)에서 촬영된 영상 내의 피사체의 종류, 사용자가 선호하는 영상 스타일, 사용자가 주로 촬영하는 장소의 환경(예를 들어, 실내, 실외, 조도, 날씨 등) 등과 같은 사용자의 선호 상황을 식별할 수 있다. 또는, 예를 들어, 디바이스(1000)는 디바이스(1000)의 카메라의 종류, 렌즈 종류에 관한 정보를 서버(미도시)에게 제공하면서 서버(미도시)에게 인공지능 모델(1620)의 훈련을 위한 기준 원시 영상 및 기준 영상을 요청할 수 있다. 서버(미도시)는 디바이스(1000)로부터 수신된 카메라의 종류 및 렌즈 종류에 관련된 기준 원시 영상 및 기준 영상을 디바이스(1000)에게 제공할 수 있다.

[148] 또는, 디바이스(1000)는 사용자의 선호하는 영상 스타일, 촬영 환경 및 피사체 등에 관한 사용자 선호 정보를 서버(미도시)에게 제공하면서 서버(미도시)에게 인공지능 모델(1620)의 훈련을 위한 기준 원시 영상 및 기준 영상을 요청할 수 있다. 서버(미도시)는 디바이스(1000)로부터 수신된 사용자 선호 정보에 관련된 기준 원시 영상 및 기준 영상을 디바이스(1000)에게 제공할 수 있다.

[149] 도 9는 본 개시의 일 실시예에 따른 디바이스(1000)가 라이브 뷰 영상을 출력하는 방법의 흐름도이다.

[150] 동작 S900에서 디바이스(1000)는 피사체를 촬영하기 위한 촬영 기능을 활성화하는 사용자 입력을 수신할 수 있다. 예를 들어, 디바이스(1000)는 디바이스(1000)에 설치된 카메라 애플리케이션을 실행하는 사용자 입력을 수신할 수 있다.

[151] 동작 S905에서 디바이스(1000)는 카메라 센서부(1400)를 통하여 입력되는 광에 기초하여 원시 영상을 생성할 수 있다. 디바이스(1000)는 촬영 기능이 활성화됨에 따라 라이브 뷰 영상을 생성하는데 이용되는 원시 영상을 생성할 수

있다.

- [152] 동작 S910에서 디바이스(1000)는 인공지능 모델(1730) 내의 모델들 중에서 적어도 하나를 비활성화할 수 있다. 디바이스(1000)는 톤 맵 생성 모델(1731), 특징 추출 모델(1732), 영상 보정 모델(1733) 또는 영상 회복 모델(1734) 중 적어도 하나를 비활성화할 수 있다. 라이브 뷰 영상의 생성을 위하여 톤 맵 생성 모델(1731), 특징 추출 모델(1732), 영상 보정 모델(1733) 또는 영상 회복 모델(1734) 중 어느 모델을 비활성화할 지는 미리 설정될 수 있다.
- [153] 동작 S915에서 디바이스(1000)는 적어도 하나의 모델이 비활성화된 인공지능 모델(1730)에 원시 영상을 입력할 수 있으며, 동작 S920에서 디바이스(1000)는 인공지능 모델(1730)로부터 출력되는 라이브 뷰 영상을 화면 상에 디스플레이할 수 있다. 이 경우, 인공지능 모델(1620)은 인공지능 모델(1620) 내의 모델들 중에서 적어도 하나가 비활성화된 상태에서 양호한 라이브 뷰 영상이 출력될 수 있도록 미리 훈련될 수 있다.
- [154] 한편, 상기에서는 톤 맵 생성 모델(1731), 특징 추출 모델(1732), 영상 보정 모델(1733) 또는 영상 회복 모델(1734) 중 적어도 하나가 비활성된 인공지능 모델(1720)을 이용하여 라이브 뷰 영상을 생성하는 것으로 설명되었지만 이에 제한되지 않는다. 톤 맵 생성 모델(1731), 특징 추출 모델(1732), 영상 보정 모델(1733) 및 영상 회복 모델(1734)이 모두 활성화된 인공지능 모델(1730)을 이용하여 라이브 뷰 영상이 생성될 수도 있다.
- [155] 한편, 톤 맵 생성 모델(1731), 특징 추출 모델(1732), 영상 보정 모델(1733) 또는 영상 회복 모델(1734) 중 적어도 하나가 비활성된 인공지능 모델(1720)을 이용하여, 피사체가 촬영된 영상이 생성되고 저장될 수도 있다.
- [156] 도 10a는 본 개시의 일 실시예에 따른 라이브 뷰 영상을 생성하기 위하여 인공지능 모델(1730) 내의 톤 맵 생성 모델(1731)을 비활성화하는 예시를 나타내는 도면이다.
- [157] 도 10a를 참조하면, 인공지능 모델(1730) 내의 톤 맵 생성 모델(1731)이 비활성화된 상태에서, 인공지능 모델(1730)에 원시 영상이 입력될 수 있다. 원시 영상은 특징 추출 모델(1732)에 입력되고, 특징 추출 모델(1732)로부터 출력된 특징 영상들이 보정되고, 보정된 특징 영상들이 영상 회복 모델(1734)에 입력되며, 영상 회복 모델(1734)로부터 라이브 뷰 영상이 출력될 수 있다. 이 경우, 인공지능 모델(1730)은 톤 맵 생성 모델(1731)이 비활성화된 상태에서 도 7에서와 같은 방법으로 훈련된 모델일 수 있다.
- [158] 도 10b는 본 개시의 일 실시예에 따른 라이브 뷰 영상을 생성하기 위하여 인공지능 모델(1730) 내의 특징 추출 모델(1732) 및 영상 회복 모델(1734)을 비활성화하는 예시를 나타내는 도면이다.
- [159] 도 10b를 참조하면, 인공지능 모델(1730) 내의 특징 추출 모델(1732) 및 영상 회복 모델(1734)이 비활성화된 상태에서, 인공지능 모델(1730)에 원시 영상이 입력될 수 있다. 원시 영상은 톤 맵 생성 모델(1731)에 입력되고, 톤 맵 생성

모델(1731)로부터 출력된 톤 맵에 기초하여 원시 영상의 밝기가 조정되고, 밝기가 조정된 원시 영상의 보정됨으로써 라이브뷰 영상이 생성될 수 있다. 이 경우, 인공지능 모델(1730)은 특징 추출 모델(1732) 및 영상 회복 모델(1734)이 비활성화된 상태에서 도 7에서와 같은 방법으로 훈련된 모델일 수 있다.

- [160] 도 10c는 본 개시의 일 실시예에 따른 라이브뷰 영상을 생성하기 위하여 인공지능 모델(1730) 내의 톤 맵 생성 모델(1731), 특징 추출 모델(1732) 및 영상 회복 모델(1734)을 비활성화하는 예시를 나타내는 도면이다.
- [161] 도 10c를 참조하면, 인공지능 모델(1730) 내의 톤 맵 생성 모델(1731), 특징 추출 모델(1732) 및 영상 회복 모델(1734)이 비활성화된 상태에서, 원시 영상이 보정됨으로써 라이브뷰 영상이 생성될 수도 있다.
- [162] 한편, 도 10a 내지 도 10c에서는, 화이트 밸런스 매트릭스 값 및 색상 보정 매트릭스 값 값이 영상 보정 모델(1733)에 입력되는 것으로 설명되었지만 이에 제한되지 않는다. 복수의 인공지능 모델(1620)이 화이트 밸런스 및 색상 보정에 관한 설정 별로 각각 훈련될 수 있다. 이 경우에는, 라이브뷰 영상을 생성하기 위해 소정의 화이트 밸런스 및 색상 보정에 관한 설정에 대응되는 인공지능 모델(1730)이 제2 메모리(1720)에 로딩되어 제2 프로세서(1710)에 의해 이용될 수 있다. 또한, 제2 메모리(1720)에 로딩된 인공지능 모델(1730)에는 화이트 밸런스 매트릭스 값 및 색상 보정 매트릭스 값이 입력되지 않더라도, 제2 메모리(1720)에 로딩된 인공지능 모델(1730)로부터 소정의 화이트 밸런스 및 색상 보정에 관한 설정을 고려한 라이브뷰 영상이 출력될 수 있게 된다.
- [163] 도 11은 디바이스(1000)가 리트레이닝된 인공지능 모델(1620)을 서버(미도시)로부터 수신하여 인공지능 모델(1620)을 업데이트하는 방법의 흐름도이다.
- [164] 동작 S1100에서 디바이스(1000)는 리트레이닝된 인공지능 모델(1620)을 서버(미도시)에게 요청할 수 있다. 인공지능 모델(1620)은 서버(미도시)에 의해 리트레이닝될 수 있으며, 서버(미도시)는 리트레이닝된 인공지능 모델(1620)이 존재함을 알리는 알림 정보를 디바이스(1000)에게 제공할 수 있다. 또한, 디바이스(1000)는 서버(미도시)로부터 수신된 알림 정보를 화면 상에 디스플레이하고, 인공지능 모델(1620)을 업데이트하기 위한 사용자 입력을 수신할 수 있다. 디바이스(1000)는 서버(미도시)에게 사용자가 선호하는 사진 속성에 관한 정보를 제공하면서 서버(미도시)에게 리트레이닝된 인공지능 모델(1620)을 요청할 수도 있다.
- [165] 디바이스(1000)는 디바이스(1000)에서 촬영된 영상들을 서버(미도시)에게 제공하면서 서버(미도시)에게 리트레이닝된 인공지능 모델(1620)을 요청할 수 있다. 서버(미도시)는 디바이스(1000)로부터 수신된 영상들을 분석하여 디바이스(1000)의 사용자가 선호하는 상황에 관련된 기준 원시 영상 및 기준 영상을 이용하여 인공지능 모델(1620)을 리트레이닝할 수 있다. 예를 들어, 서버(미도시)는 디바이스(1000)로부터 수신된 영상들을 분석함으로써, 사용자가

이용한 디바이스(1000)의 종류, 디바이스(1000)에서 촬영된 영상 내의 피사체의 종류, 사용자가 선호하는 영상 스타일, 사용자가 주로 촬영하는 장소의 환경(예를 들어, 실내, 실외, 조도, 날씨 등) 등과 같은 사용자의 선호 상황을 식별할 수 있다. 또한, 서버(미도시)는 사용자의 선호 상황에 관련된 기준 원시 영상 및 기준 영상을 이용하여 인공지능 모델(1620)을 리트레이닝할 수 있다.

- [166] 또는, 예를 들어, 디바이스(1000)는 디바이스(1000)의 카메라의 종류, 렌즈 종류에 관한 정보를 서버(미도시)에게 제공하면서 서버(미도시)에게 리트레이닝된 인공지능 모델(1620)을 요청할 수 있다. 서버(미도시)는 디바이스(1000)로부터 수신된 카메라의 종류 및 렌즈 종류에 관련된 기준 원시 영상 및 기준 영상을 이용하여 인공지능 모델(1620)을 리트레이닝할 수 있다.
- [167] 또는, 디바이스(1000)는 사용자의 선호하는 영상 스타일, 촬영 환경 및 피사체 등에 관한 사용자 선호 정보를 서버(미도시)에게 제공하면서 서버(미도시)에게 리트레이닝된 인공지능 모델(1620)을 요청할 수 있다. 서버(미도시)는 디바이스(1000)로부터 수신된 사용자 선호 정보에 관련된 기준 원시 영상 및 기준 영상을 이용하여 인공지능 모델(1620)을 리트레이닝할 수 있다.
- [168] 동작 S1110에서 디바이스(1000)는 리트레이닝된 인공지능 모델(1620)을 서버(미도시)로부터 수신할 수 있다. 디바이스(1000)가 사용자가 선호하는 사진 속성에 관한 정보를 서버(미도시)에게 제공한 경우에는, 서버(미도시)는 사용자가 선호하는 사진 속성과 관련하여 리트레이닝된 인공지능 모델(1620)을 디바이스(1000)에게 제공할 수 있다.
- [169] 동작 S1120에서 디바이스(1000)는 리트레이닝된 인공지능 모델(1620)에 기초하여, 디바이스(1000) 내의 인공지능 모델(1620)을 업데이트할 수 있다. 디바이스(1000)는, 예를 들어, 디바이스(1000) 내의 인공지능 모델(1620)을 리트레이닝된 인공지능 모델(1620)로 대체함으로써, 디바이스(1000) 내의 인공지능 모델(1620)을 업데이트할 수 있다.
- [170] 도 12는 디바이스(1000)가 인공지능 모델(1620)을 리트레이닝하여 업데이트하는 방법의 흐름도이다.
- [171] 동작 S1200에서 디바이스(1000)는 리트레이닝을 위한 기준 원시 영상 및 기준 원시 영상에 대응되는 기준 영상을 획득할 수 있다. 기준 영상은 리트레이닝을 위한 기준 원시 영상으로부터 생성되는 영상일 수 있다. 디바이스(1000)는 서버(미도시)에게 인공지능 모델(1620)의 리트레이닝을 위한 기준 원시 영상 및 기준 영상을 요청하고, 서버(미도시)로부터 리트레이닝을 위한 기준 원시 영상 및 기준 영상을 수신할 수 있다. 이 경우, 디바이스(1000)는 디바이스(1000)에서 촬영된 사진 또는 사용자가 선호하는 사진 속성에 관한 정보를 서버(미도시)에게 제공할 수 있으며, 서버(미도시)는 사용자가 선호하는 상황과 관련하여 생성된 기준 영상 및 기준 원시 영상을 디바이스(1000)에게 제공할 수 있다.
- [172] 예를 들어, 디바이스(1000)는 디바이스(1000)에서 촬영된 영상들을 서버(미도시)에게 제공하면서 서버(미도시)에게 인공지능 모델(1620)의

리트레이닝을 위한 기준 원시 영상 및 기준 영상을 요청할 수 있다. 서버(미도시)는 디바이스(1000)로부터 수신된 영상들을 분석하여 디바이스(1000)의 사용자가 선호하는 상황에 관련된 기준 원시 영상 및 기준 영상을 디바이스(1000)에게 제공할 수 있다. 예를 들어, 서버(미도시)는 디바이스(1000)로부터 수신된 영상들을 분석함으로써, 사용자가 이용한 디바이스(1000)의 종류, 디바이스(1000)에서 촬영된 영상 내의 피사체의 종류, 사용자가 선호하는 영상 스타일, 사용자가 주로 촬영하는 장소의 환경(예를 들어, 실내, 실외, 조도, 날씨 등) 등과 같은 사용자의 선호 상황을 식별할 수 있다. 또한, 서버(미도시)는 사용자의 선호 상황에 관련된 기준 원시 영상 및 기준 영상을 디바이스(1000)에게 제공할 수 있다.

[173] 또는, 예를 들어, 디바이스(1000)는 디바이스(1000)의 카메라의 종류, 렌즈 종류에 관한 정보를 서버(미도시)에게 제공하면서 서버(미도시)에게 인공지능 모델(1620)의 리트레이닝을 위한 기준 원시 영상 및 기준 영상을 요청할 수 있다. 서버(미도시)는 디바이스(1000)로부터 수신된 카메라의 종류 및 렌즈 종류에 관련된 기준 원시 영상 및 기준 영상을 디바이스(1000)에게 제공할 수 있다.

[174] 또는, 디바이스(1000)는 사용자의 선호하는 영상 스타일, 촬영 환경 및 피사체 등에 관한 사용자 선호 정보를 서버(미도시)에게 제공하면서 서버(미도시)에게 인공지능 모델(1620)의 리트레이닝을 위한 기준 원시 영상 및 기준 영상을 요청할 수 있다. 서버(미도시)는 디바이스(1000)로부터 수신된 사용자 선호 정보에 관련된 기준 원시 영상 및 기준 영상을 이용하여 디바이스(1000)에게 제공할 수 있다.

[175] 동작 S1210에서 디바이스(1000)는 리트레이닝을 위한 기준 원시 영상 및 기준 원시 영상에 대응되는 기준 영상을 이용하여, 디바이스(1000) 내의 인공지능 모델(1620)을 업데이트할 수 있다. 디바이스(1000)는 서버(미도시)로부터 수신된 기준 영상을 정답 영상으로 이용할 수 있으며, 디바이스(1000)는 서버(미도시)로부터 수신된 기준 원시 영상을 인공지능 모델(1620)에 입력하고 인공지능 모델(1620)로부터 출력되는 출력 영상을 기준 영상과 비교함으로써 인공지능 모델(1620)을 리트레이닝할 수 있다.

[176] 본 개시에 따른 인공지능과 관련된 기능은 프로세서와 메모리를 통해 동작된다. 프로세서는 하나 또는 복수의 프로세서로 구성될 수 있다. 이때, 하나 또는 복수의 프로세서는 CPU, AP, DSP(Digital Signal Processor) 등과 같은 범용 프로세서, GPU, VPU(Vision Processing Unit)와 같은 그래픽 전용 프로세서 또는 NPU와 같은 인공지능 전용 프로세서일 수 있다. 하나 또는 복수의 프로세서는, 메모리에 저장된 기 정의된 동작 규칙 또는 인공지능 모델에 따라, 입력 데이터를 처리하도록 제어한다. 또는, 하나 또는 복수의 프로세서가 인공지능 전용 프로세서인 경우, 인공지능 전용 프로세서는, 특정 인공지능 모델의 처리에 특화된 하드웨어 구조로 설계될 수 있다.

[177] 기 정의된 동작 규칙 또는 인공지능 모델은 학습을 통해 만들어진 것을

특징으로 한다. 여기서, 학습을 통해 만들어진다는 것은, 기본 인공지능 모델이 학습 알고리즘에 의하여 다수의 학습 데이터들을 이용하여 학습됨으로써, 원하는 특성(또는, 목적)을 수행하도록 설정된 기 정의된 동작 규칙 또는 인공지능 모델이 만들어짐을 의미한다. 이러한 학습은 본 개시에 따른 인공지능이 수행되는 기기 자체에서 이루어질 수도 있고, 별도의 서버 및/또는 시스템을 통해 이루어 질 수도 있다. 학습 알고리즘의 예로는, 지도형 학습(supervised learning), 비지도형 학습(unsupervised learning), 준지도형 학습(semi-supervised learning) 또는 강화 학습(reinforcement learning)이 있으나, 전술한 예에 한정되지 않는다.

[178] 인공지능 모델은, 복수의 신경망 레이어들로 구성될 수 있다. 복수의 신경망 레이어들 각각은 복수의 가중치들(weight values)을 갖고 있으며, 이전(previous) 레이어의 연산 결과와 복수의 가중치들 간의 연산을 통해 신경망 연산을 수행한다. 복수의 신경망 레이어들이 갖고 있는 복수의 가중치들은 인공지능 모델의 학습 결과에 의해 최적화될 수 있다. 예를 들어, 학습 과정 동안 인공지능 모델에서 획득한 로스(loss) 값 또는 코스트(cost) 값이 감소 또는 최소화되도록 복수의 가중치들이 갱신될 수 있다. 인공 신경망은 심층 신경망(DNN:Deep Neural Network)를 포함할 수 있으며, 예를 들어, CNN (Convolutional Neural Network), DNN (Deep Neural Network), RNN (Recurrent Neural Network), RBM (Restricted Boltzmann Machine), DBN (Deep Belief Network), BRDNN(Bidirectional Recurrent Deep Neural Network) 또는 심층 Q-네트워크 (Deep Q-Networks) 등이 있으나, 전술한 예에 한정되지 않는다.

[179] 본 개시의 일 실시예에서 원시 영상이 인공지능 모델의 입력 데이터로 이용되어 인공지능 모델로부터 출력 영상 데이터가 출력될 수 있다. 인공지능 모델은 학습을 통해 만들어 질 수 있다. 여기서, 학습을 통해 만들어진다는 것은, 기본 인공지능 모델이 학습 알고리즘에 의하여 다수의 학습 데이터들을 이용하여 학습됨으로써, 원하는 특성(또는, 목적)을 수행하도록 설정된 기 정의된 동작 규칙 또는 인공지능 모델이 만들어짐을 의미한다. 인공지능 모델은, 복수의 신경망 레이어들로 구성될 수 있다. 복수의 신경망 레이어들 각각은 복수의 가중치들(weight values)을 갖고 있으며, 이전(previous) 레이어의 연산 결과와 복수의 가중치들 간의 연산을 통해 신경망 연산을 수행한다. 인공지능 모델은, 객체 인식(Object Recognition), 객체 추적(Object Tracking), 영상 검색(Image Retrieval), 사람 인식(Human Recognition), 장면 이해(Scene Recognition), 공간 이해(3D Reconstruction/Localization), 영상 개선(Image Enhancement) 등에 이용될 수 있다.

[180] 본 개시의 일 실시예는 컴퓨터에 의해 실행되는 프로그램 모듈과 같은 컴퓨터에 의해 실행가능한 명령어를 포함하는 기록 매체의 형태로도 구현될 수 있다. 컴퓨터 판독 가능 매체는 컴퓨터에 의해 액세스될 수 있는 임의의 가용 매체일 수 있고, 휘발성 및 비휘발성 매체, 분리형 및 비분리형 매체를 모두

포함한다. 또한, 컴퓨터 판독가능 매체는 컴퓨터 저장 매체 및 통신 매체를 포함할 수 있다. 컴퓨터 저장 매체는 컴퓨터 판독가능 명령어, 데이터 구조, 프로그램 모듈 또는 기타 데이터와 같은 정보의 저장을 위한 임의의 방법 또는 기술로 구현된 휘발성 및 비휘발성, 분리형 및 비분리형 매체를 모두 포함한다. 통신 매체는 전형적으로 컴퓨터 판독가능 명령어, 데이터 구조, 또는 프로그램 모듈과 같은 변조된 데이터 신호의 기타 데이터를 포함할 수 있다.

[181] 또한, 컴퓨터에 의해 읽을 수 있는 저장매체는, 비일시적(non-transitory)

저장매체의 형태로 제공될 수 있다. 여기서, ‘비일시적 저장매체’는 실재(tangible)하는 장치이고, 신호(signal)(예: 전자기파)를 포함하지 않는다는 것을 의미할 뿐이며, 이 용어는 데이터가 저장매체에 반영구적으로 저장되는 경우와 임시적으로 저장되는 경우를 구분하지 않는다. 예로, ‘비일시적 저장매체’는 데이터가 임시적으로 저장되는 버퍼를 포함할 수 있다.

[182] 일 실시예에 따르면, 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 두개의 사용자 장치들(예: 스마트폰들) 간에 직접, 온라인으로 배포(예: 다운로드 또는 업로드)될 수 있다. 온라인 배포의 경우에, 컴퓨터 프로그램 제품(예: 다운로드할 앱(downloadable app))의 적어도 일부는 제조사의 서버, 어플리케이션 스토어의 서버, 또는 중계 서버의 메모리와 같은 기기로 읽을 수 있는 저장 매체에 적어도 일시 저장되거나, 임시적으로 생성될 수 있다.

[183] 또한, 본 명세서에서, “부”는 프로세서 또는 회로와 같은 하드웨어 구성(hardware component), 및/또는 프로세서와 같은 하드웨어 구성에 의해 실행되는 소프트웨어 구성(software component)일 수 있다.

[184] 또한, 본 명세서에서, “a, b 또는 c 중 적어도 하나를 포함한다”는 “a만 포함하거나, b만 포함하거나, c만 포함하거나, a 및 b를 포함하거나, b 및 c를 포함하거나, a 및 c를 포함하거나, a, b 및 c를 모두 포함하는 것을 의미할 수 있다.

[185] 전술한 본 개시의 설명은 예시를 위한 것이며, 본 개시가 속하는 기술분야의 통상의 지식을 가진 자는 본 개시의 기술적 사상이나 필수적인 특징을 변경하지 않고서 다른 구체적인 형태로 쉽게 변형이 가능하다는 것을 이해할 수 있을 것이다. 그러므로 이상에서 기술한 실시예들은 모든 면에서 예시적인 것이며 한정적이 아닌 것으로 이해해야만 한다. 예를 들어, 단일형으로 설명되어 있는 각 구성 요소는 분산되어 실시될 수도 있으며, 마찬가지로 분산된 것으로 설명되어 있는 구성 요소들도 결합된 형태로 실시될 수 있다.

[186] 본 개시의 범위는 상기 상세한 설명보다는 후술하는 특허청구범위에 의하여 나타내어지며, 특허청구범위의 의미 및 범위 그리고 그 균등 개념으로부터

도출되는 모든 변경 또는 변형된 형태가 본 개시의 범위에 포함되는 것으로 해석되어야 한다.

청구범위

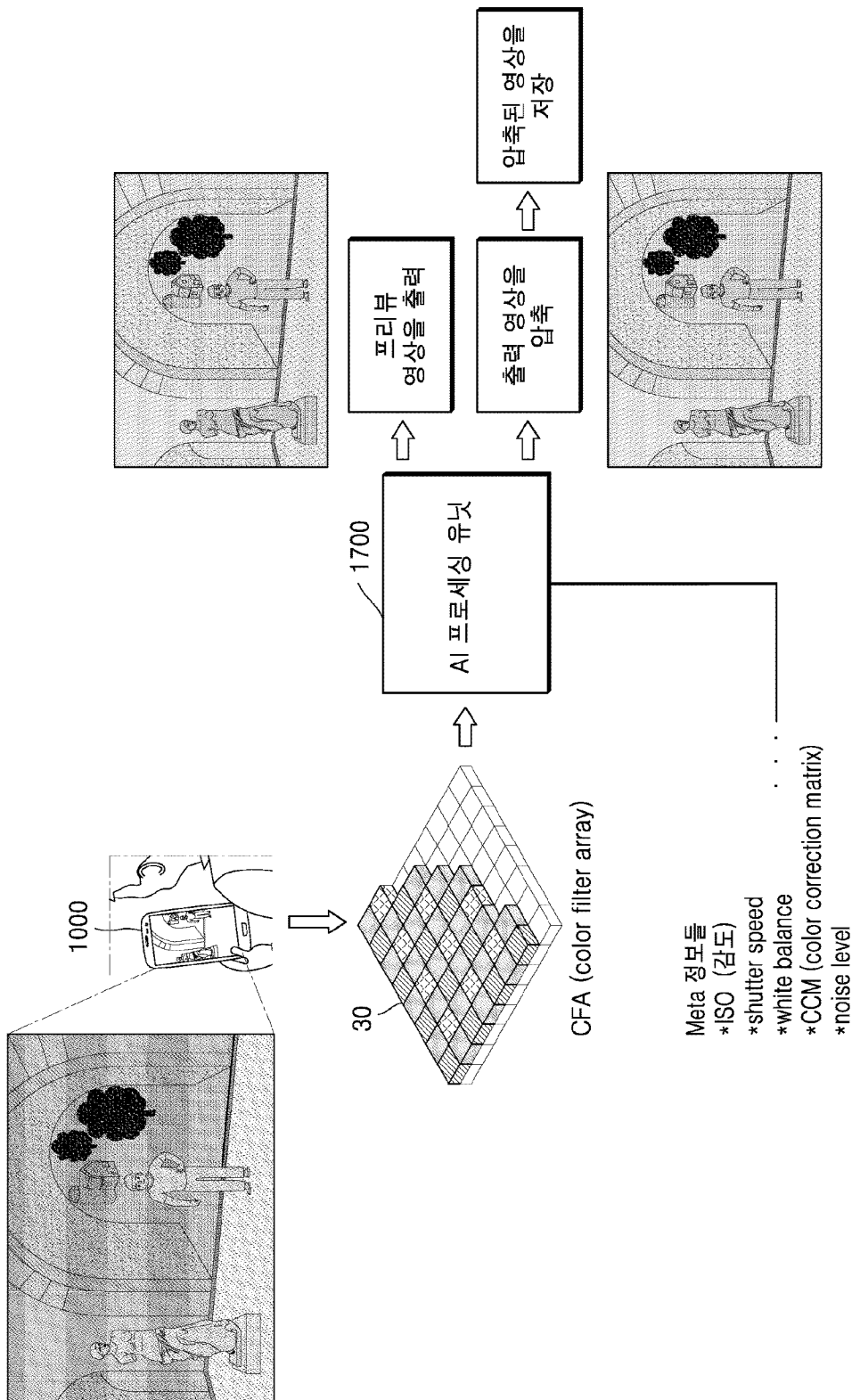
- [청구항 1] 디바이스가 원시 영상으로부터 피사체가 촬영된 출력 영상을 생성하는 방법에 있어서,
 상기 디바이스의 제어를 위한 제1 프로세서를 이용하여, 상기 디바이스 내의 카메라 센서를 제어함으로써 상기 카메라 센서의 전방의 피사체를 나타내는 원시 영상을 획득하는 동작;
 인공지능 모델을 이용하는 상기 원시 영상에 대한 영상 처리를 위한 제2 프로세서를 이용하여, 상기 카메라 센서의 전방의 상기 피사체를 나타내는 상기 원시 영상을, 영상의 밝기를 조절하기 위해 훈련된 제1 인공지능 모델에 입력하는 동작;
 상기 제2 프로세서를 이용하여, 상기 제1 인공지능 모델로부터 출력되는 톤 맵 데이터를 획득하는 동작; 및
 상기 톤 맵 데이터에 기초하여 생성되는 출력 영상을 저장하는 동작; 을 포함하는, 방법.
- [청구항 2] 제1 항에 있어서,
 상기 제2 프로세서를 이용하여:
 상기 카메라 센서의 전방의 피사체를 나타내는 상기 원시 영상 및 상기 제1 인공지능 모델로부터 출력되는 상기 톤 맵 데이터를, 영상의 특징들을 분석하기 위해 훈련된 제2 인공지능 모델에 입력하는 동작;
 상기 제2 인공지능 모델로부터 출력되는 복수의 특징 영상들을 획득하는 동작;
 영상의 특징을 보정하기 위한 적어도 하나의 설정에 기초하여, 상기 제2 인공지능 모델로부터 출력된 복수의 특징 영상들을 보정하는 동작; 및
 상기 적어도 하나의 설정에 기초하여 보정된 복수의 특징 영상들에 기초하여 상기 피사체를 포함하는 출력 영상을 생성하는 동작; 을 더 포함하는, 방법.
- [청구항 3] 제2 항에 있어서,
 상기 출력 영상을 생성하는 동작은, 상기 제2 프로세서를 이용하여, 상기 보정된 복수의 특징 영상들을, 출력 영상을 생성하기 위한 제3 인공지능 모델에 입력하는 동작을 포함하는 것인, 방법.
- [청구항 4] 제3 항에 있어서,
 상기 제1 인공지능 모델은 상기 원시 영상의 픽셀 별로 밝기를 조절하기 위한 톤 맵을 생성하도록 미리 훈련된 모델이고,
 상기 제2 인공지능 모델은, 영상 내의 기설정된 복수의 특징들을 분석하기 위하여 미리 훈련된 모델이며,
 상기 제3 인공지능 모델은, 복수의 특징 영상들로부터 출력 영상을 회복하도록 미리 훈련된 모델인 것인, 방법.

- [청구항 5] 제3 항에 있어서,
상기 제1 인공지능 모델, 상기 제2 인공지능 모델 및 상기 제3 인공지능 모델은, 기준 원시 영상 및 상기 기준 원시 영상으로부터 기설정된 ISP 프로세싱 (Image Signal Processing)을 통해 생성되는 기준 출력 영상을 바탕으로 함께 훈련된 것인, 방법.
- [청구항 6] 제4 항에 있어서,
연속 촬영으로 인하여 생성되는 복수의 원시 영상들을 조합함으로써 생성되는 상기 기준 원시 영상으로부터 기설정된 ISP 프로세싱을 통해 생성되는 상기 기준 출력 영상, 및 상기 기준 원시 영상으로부터 상기 제1 인공지능 모델, 상기 제2 인공지능 모델 및 상기 제3 인공지능 모델을 통해 출력되는 출력 영상 간의 로스(loss) 기초하여, 상기 제1 인공지능 모델, 상기 제2 인공지능 모델 및 상기 제3 인공지능 모델이 함께 훈련된 것인, 방법.
- [청구항 7] 제1 항에 있어서,
상기 원시 영상은 상기 카메라 센서 내의 이미지 센서 및 컬러 필터를 통하여 생성되는 영상이며, 베이어(bayer) 패턴, RGBE 패턴, RYYB 패턴, CYYM 패턴, CYGM 패턴, RGBW 베이어 패턴, 및 X-trans 패턴 중 어느 하나의 패턴을 가지는 것인, 방법.
- [청구항 8] 제2 항에 있어서,
상기 적어도 하나의 설정은, 화이트 밸런스를 보정하기 위한 설정 및 색상을 보정하기 위한 설정 중 적어도 하나를 포함하는 것인, 방법.
- [청구항 9] 제3 항에 있어서,
상기 제1 인공지능 모델, 상기 제2 인공지능 모델 및 상기 제3 인공지능 모델은, 상기 디바이스의 사용자의 선호도를 반영하여 선택된 것인, 방법.
- [청구항 10] 제3 항에 있어서,
상기 제1 인공지능 모델, 상기 제2 인공지능 모델 및 상기 제3 인공지능 모델 중 적어도 하나를 이용하지 않고, 상기 원시 영상으로부터 라이브뷰 영상을 생성하는 동작; 및
상기 생성된 라이브뷰 영상을 디스플레이하는 동작;
을 더 포함하는, 방법.
- [청구항 11] 제3 항에 있어서,
상기 제1 인공지능 모델, 상기 제2 인공지능 모델 및 상기 제3 인공지능 모델을 리트레이닝하는 동작;
을 더 포함하는, 방법.
- [청구항 12] 제11 항에 있어서,
상기 리트레이닝하는 동작은,
상기 리트레이닝을 위한 기준 영상 및 상기 기준 영상에 대응되는 기준 원시 영상을 획득하는 동작; 및

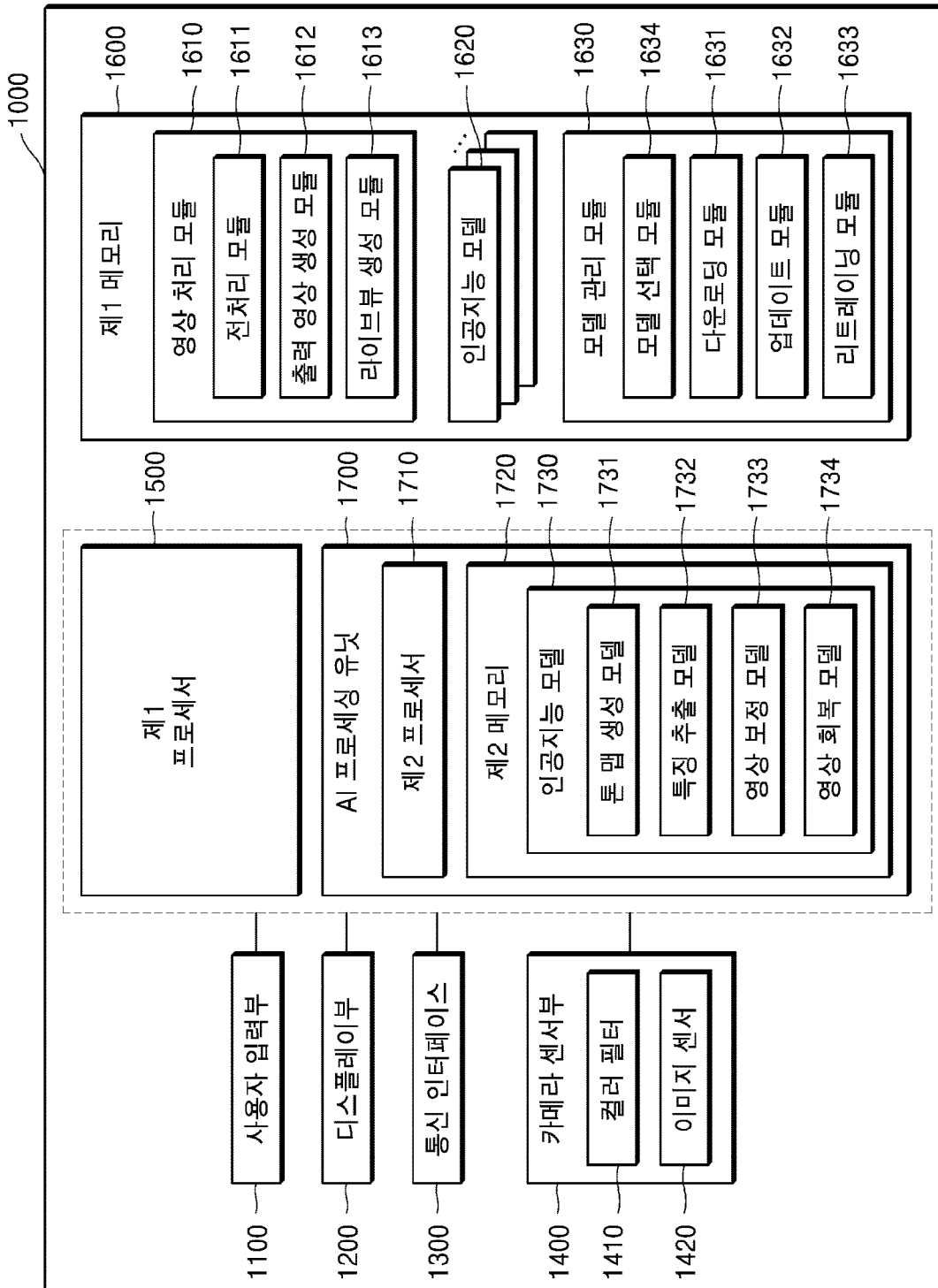
상기 획득된 기준 영상 및 상기 기준 원시 영상을 이용하여, 상기 제1 인공지능 모델, 상기 제2 인공지능 모델 및 상기 제3 인공지능 모델을 리트레이닝하는 동작;
을 포함하는 것인, 방법.

- [청구항 13] 원시 영상으로부터 피사체가 촬영된 출력 영상을 생성하는 디바이스에 있어서,
카메라 센서;
디스플레이;
상기 디바이스의 제어를 위한 제1 명령어들을 저장하는 제1 메모리;
상기 제1 메모리에 저장된 제1 명령어들을 실행하는 제1 프로세서;
상기 원시 영상에 대한 영상 처리를 위한 적어도 하나의 인공지능 모델 및
상기 인공지능 모델의 실행에 관련된 제2 명령어들을 저장하는 제2
메모리; 및
상기 제2 메모리에 저장된 상기 적어도 하나의 인공지능 모델 및 상기 제2
명령어들을 실행하는 제2 프로세서;
를 포함하며,
상기 제1 프로세서가 상기 카메라 센서를 이용하여 상기 카메라 센서의
전방의 피사체를 나타내는 원시 영상을 획득하고,
상기 제2 프로세서가 상기 카메라 센서의 전방의 상기 피사체를 나타내는
상기 원시 영상을, 영상의 밝기를 조절하기 위해 훈련된 제1 인공지능
모델에 입력하고,
상기 제2 프로세서가 상기 제1 인공지능 모델로부터 출력되는 톤 맵
데이터를 획득하고,
상기 제1 프로세서가 상기 톤 맵 데이터에 기초하여 생성되는 출력
영상을 상기 제1 메모리에 저장하는 것인, 디바이스.
- [청구항 14] 제13 항에 있어서,
상기 제2 프로세서가, 상기 제2 명령어들을 실행함으로써:
상기 카메라 센서의 전방의 피사체를 나타내는 상기 원시 영상 및 상기
제1 인공지능 모델로부터 출력되는 톤 맵 데이터를, 영상의 특징들을
분석하기 위해 훈련된 제2 인공지능 모델에 입력하고,
상기 제2 인공지능 모델로부터 출력되는 복수의 특징 영상들을
획득하고,
영상의 특징을 보정하기 위한 적어도 하나의 설정에 기초하여, 상기 제2
인공지능 모델로부터 출력된 복수의 특징 영상들을 보정하고,
상기 적어도 하나의 설정에 기초하여 보정된 복수의 특징 영상들에
기초하여 상기 피사체를 포함하는 출력 영상을 생성하는 것인, 디바이스.
- [청구항 15] 제1 항의 방법을 컴퓨터에서 실행시키기 위한 프로그램을 기록한
컴퓨터로 읽을 수 있는 기록매체.

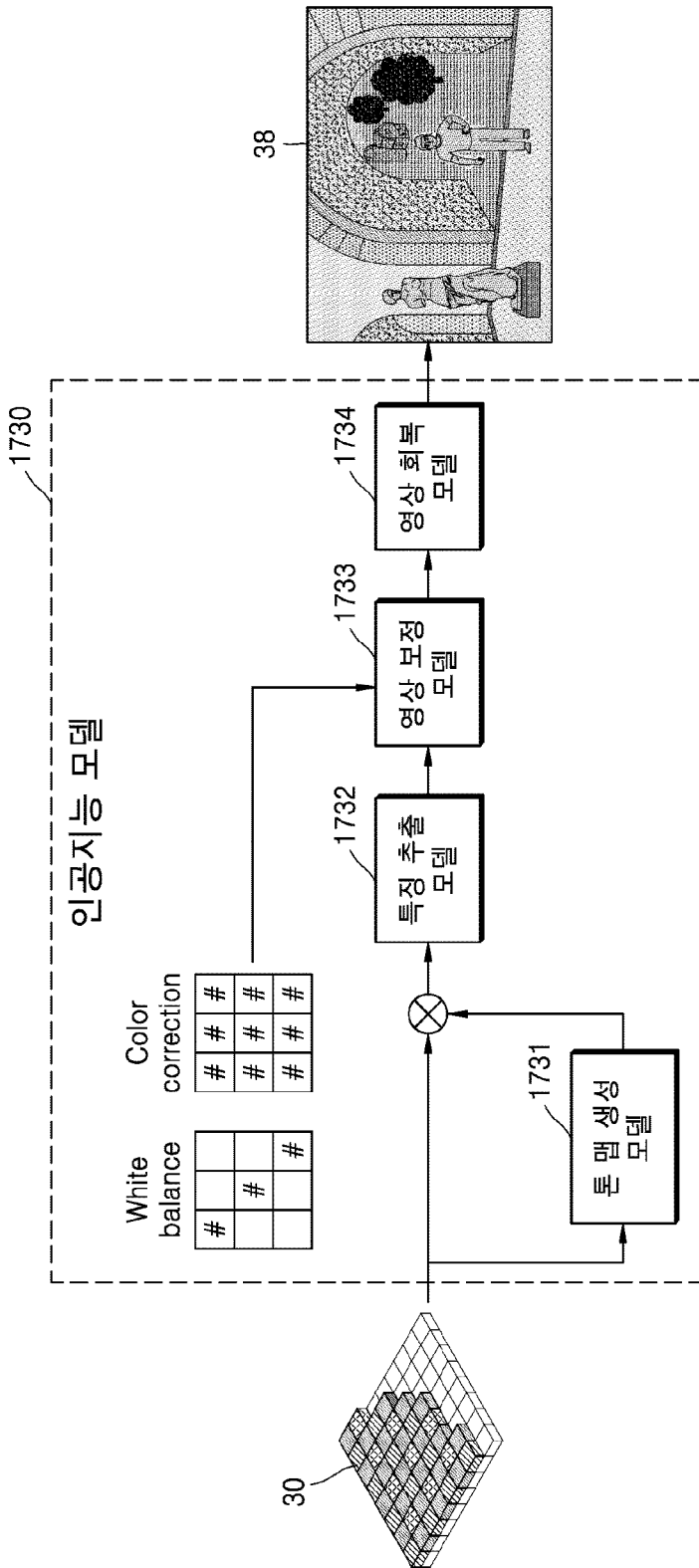
[도 1]



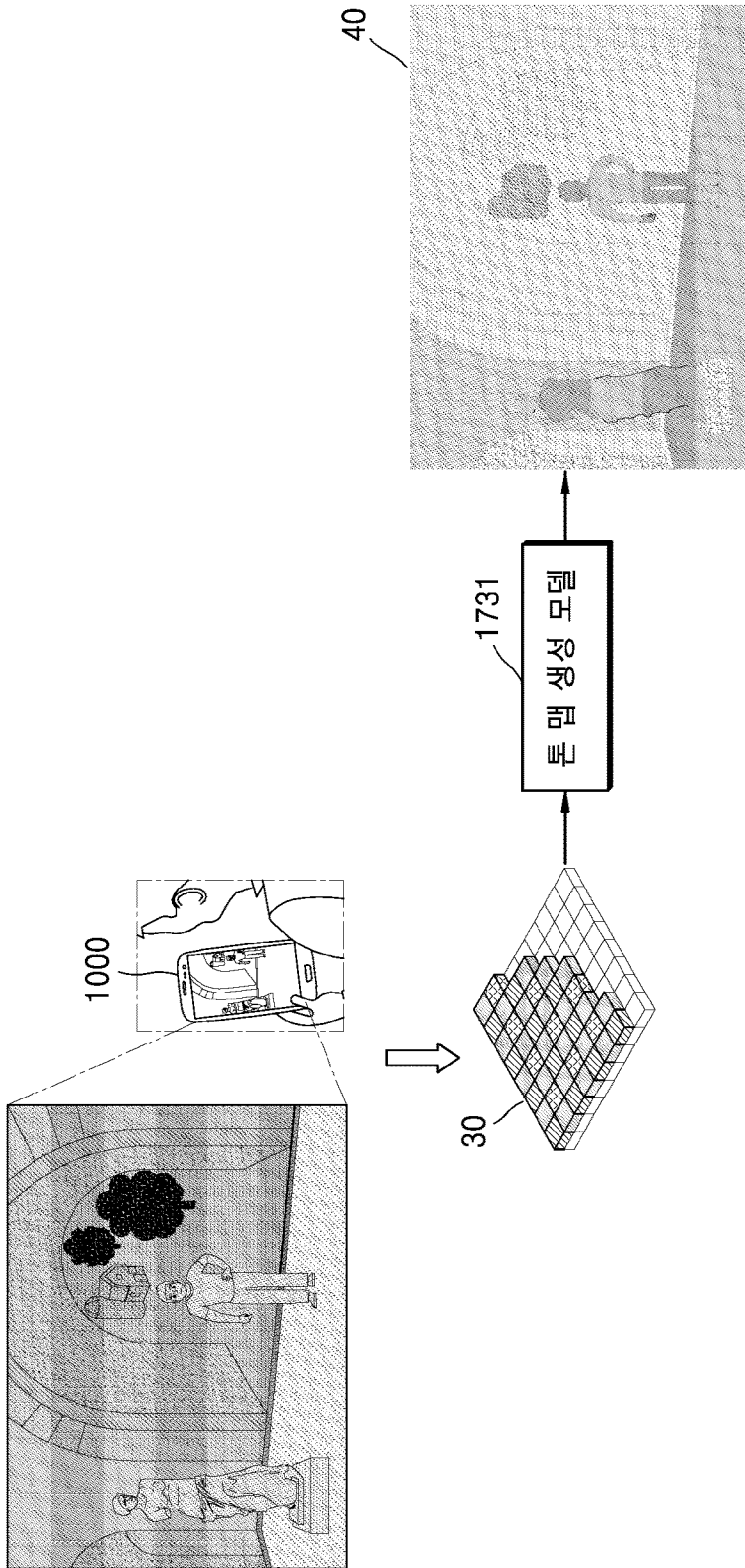
[도2]



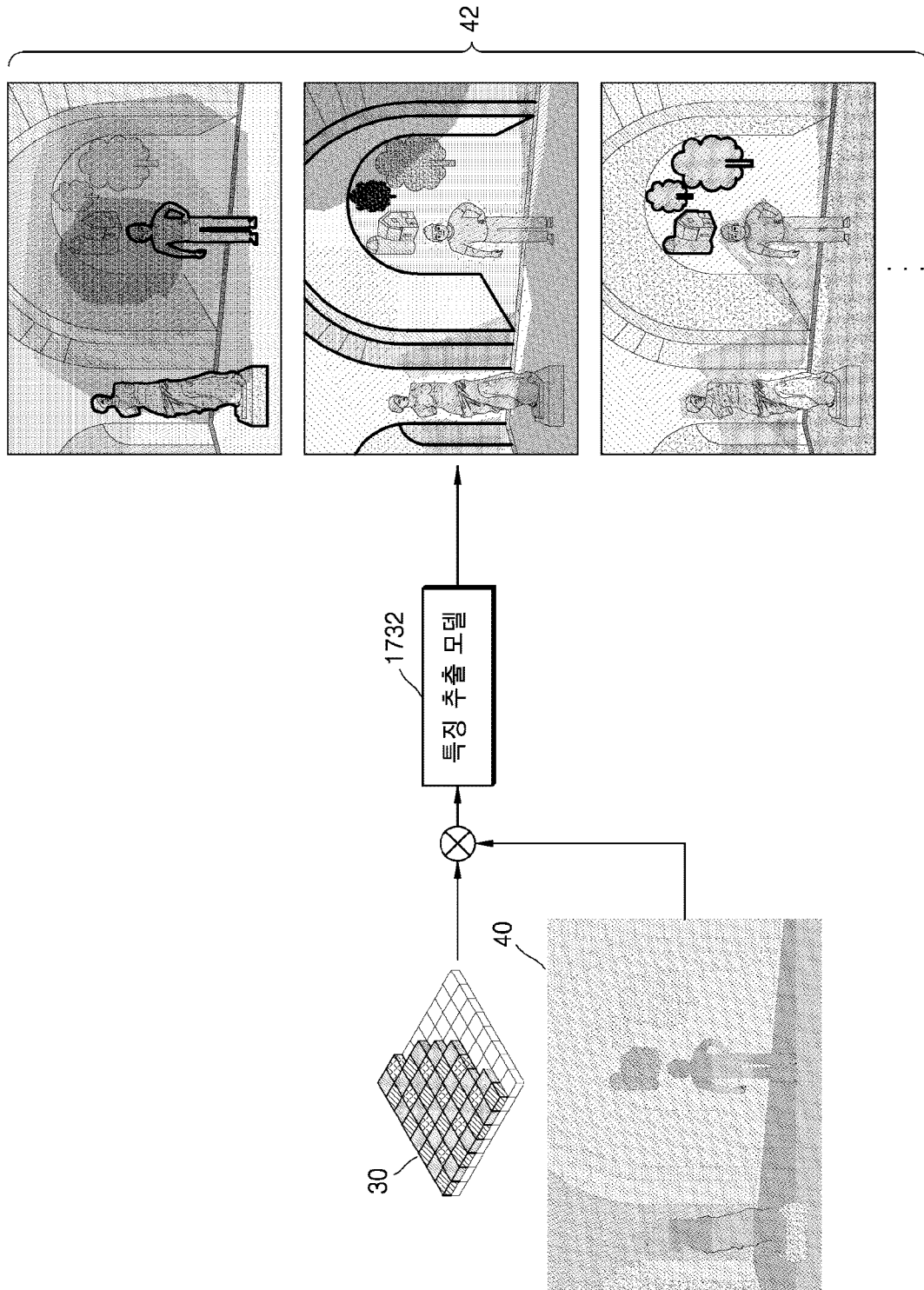
[도3]



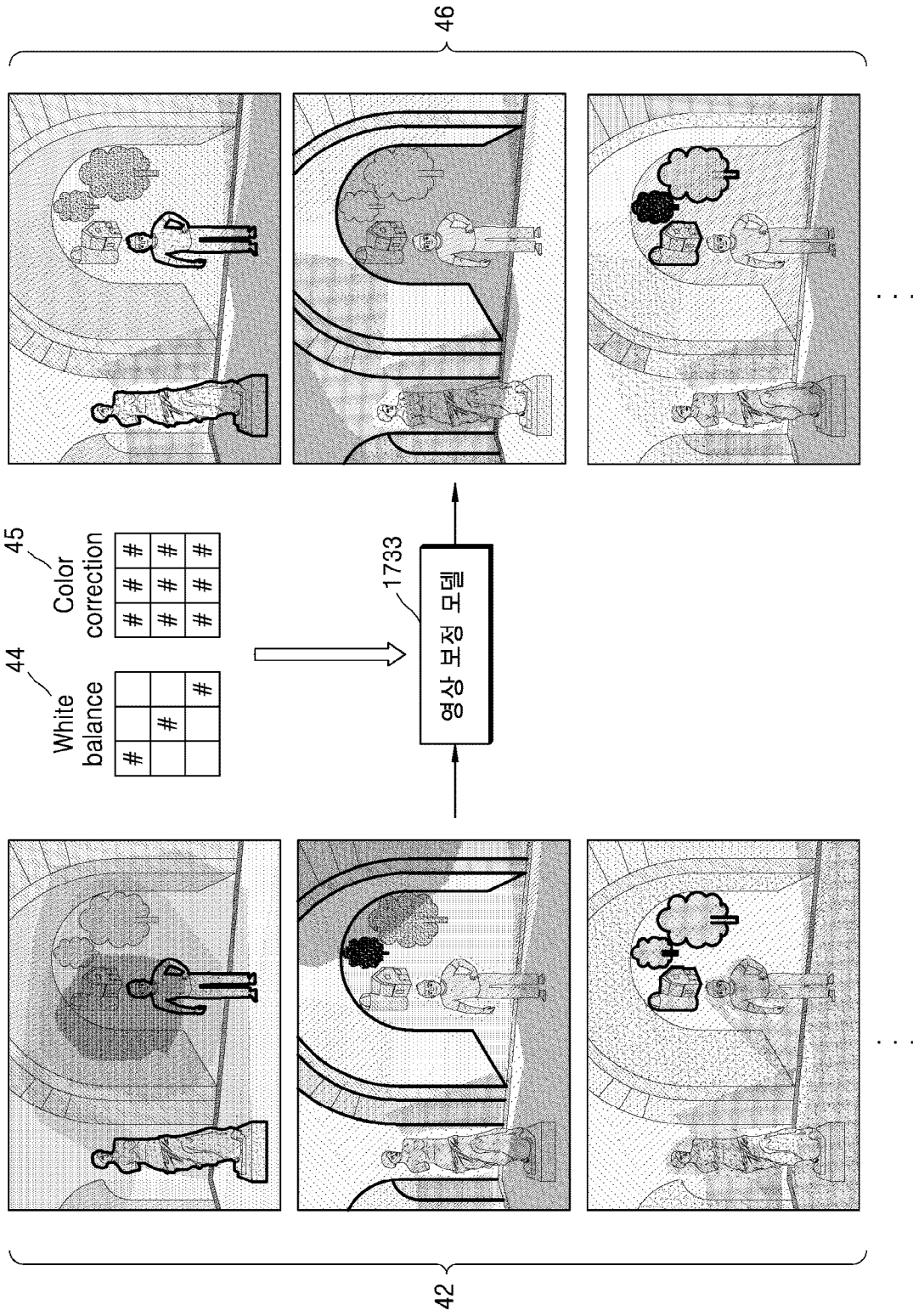
[도4a]



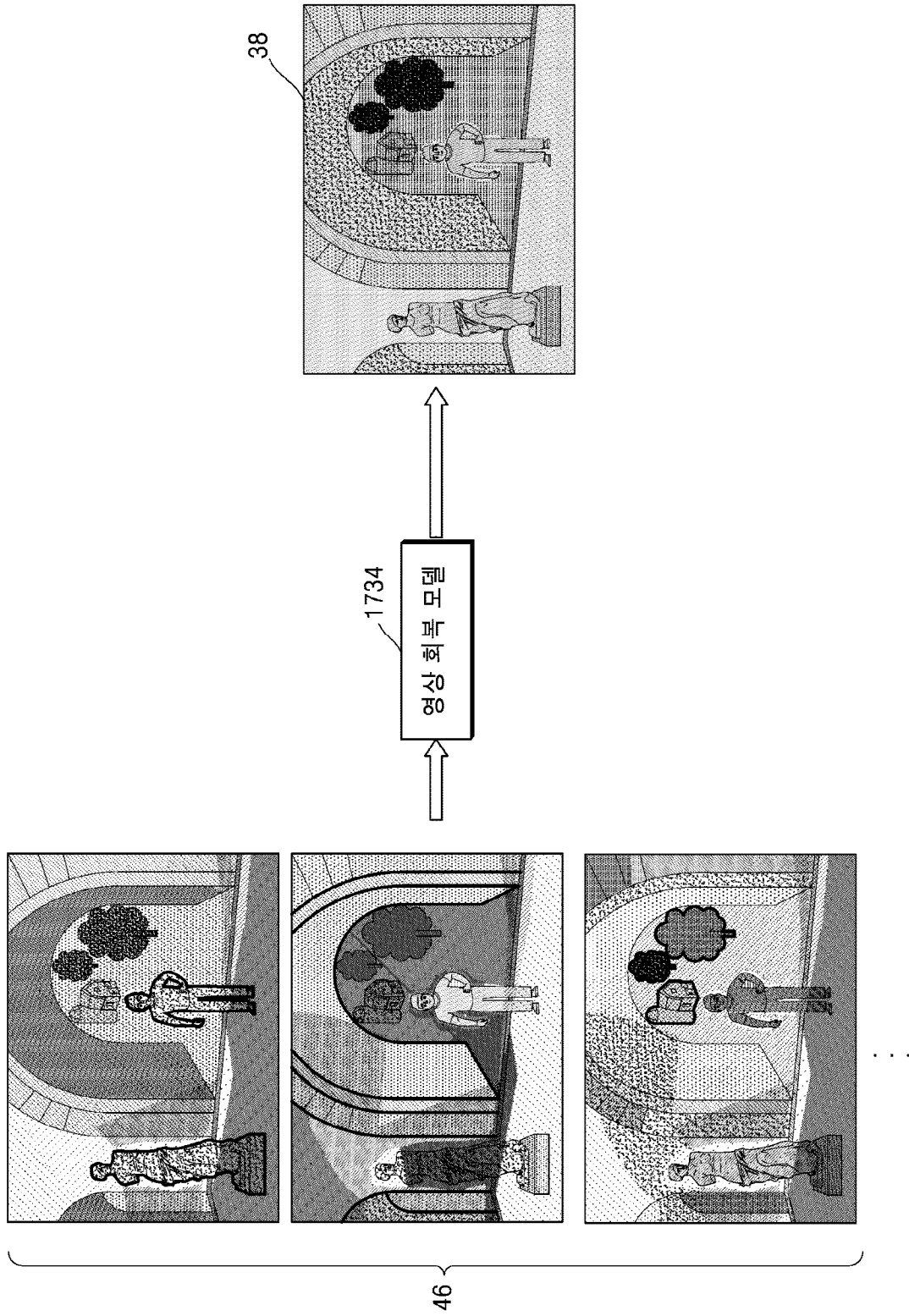
[도4b]



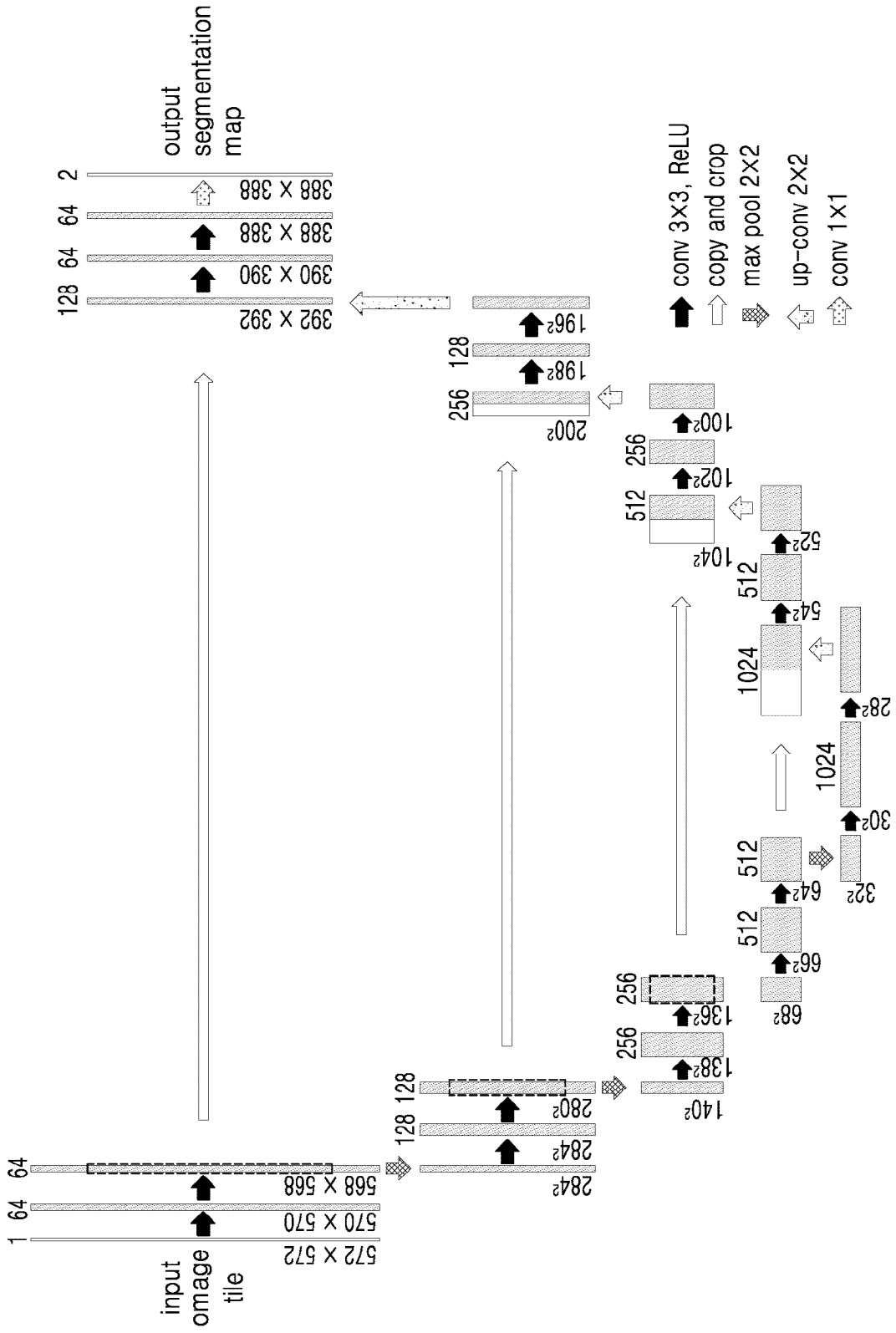
[도4c]



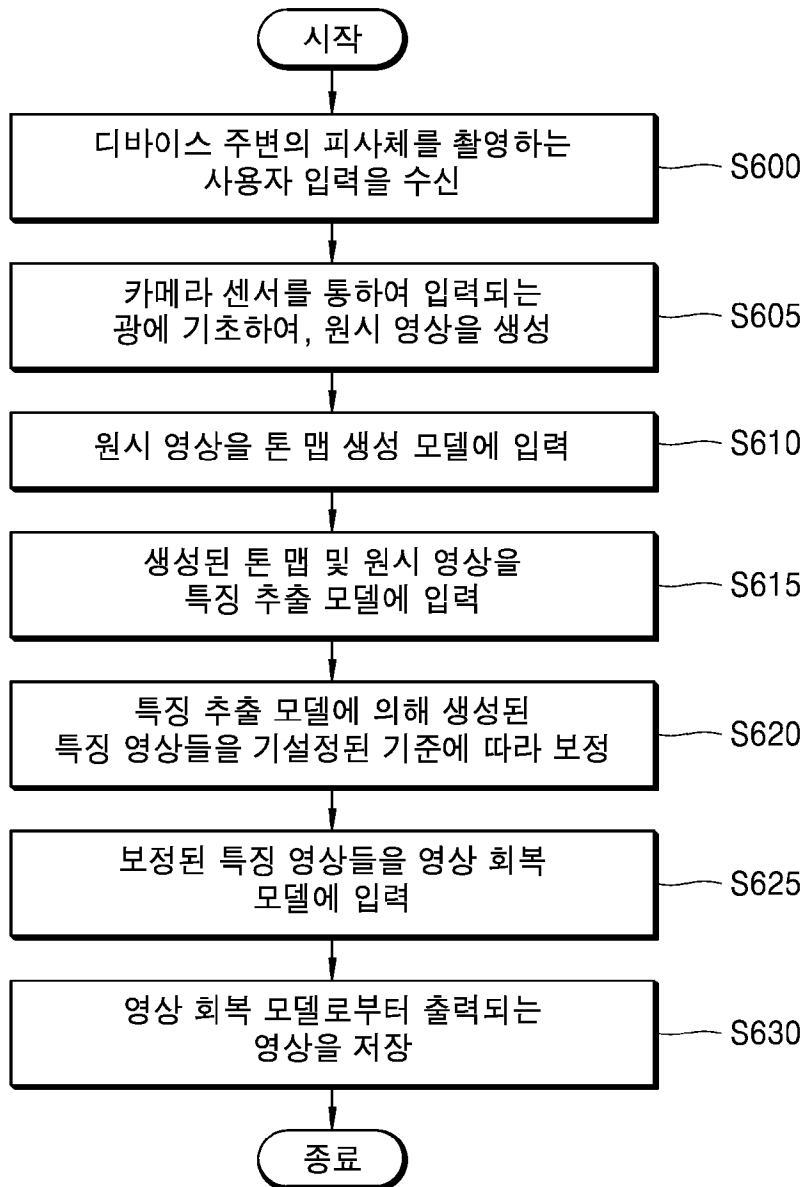
[도4d]



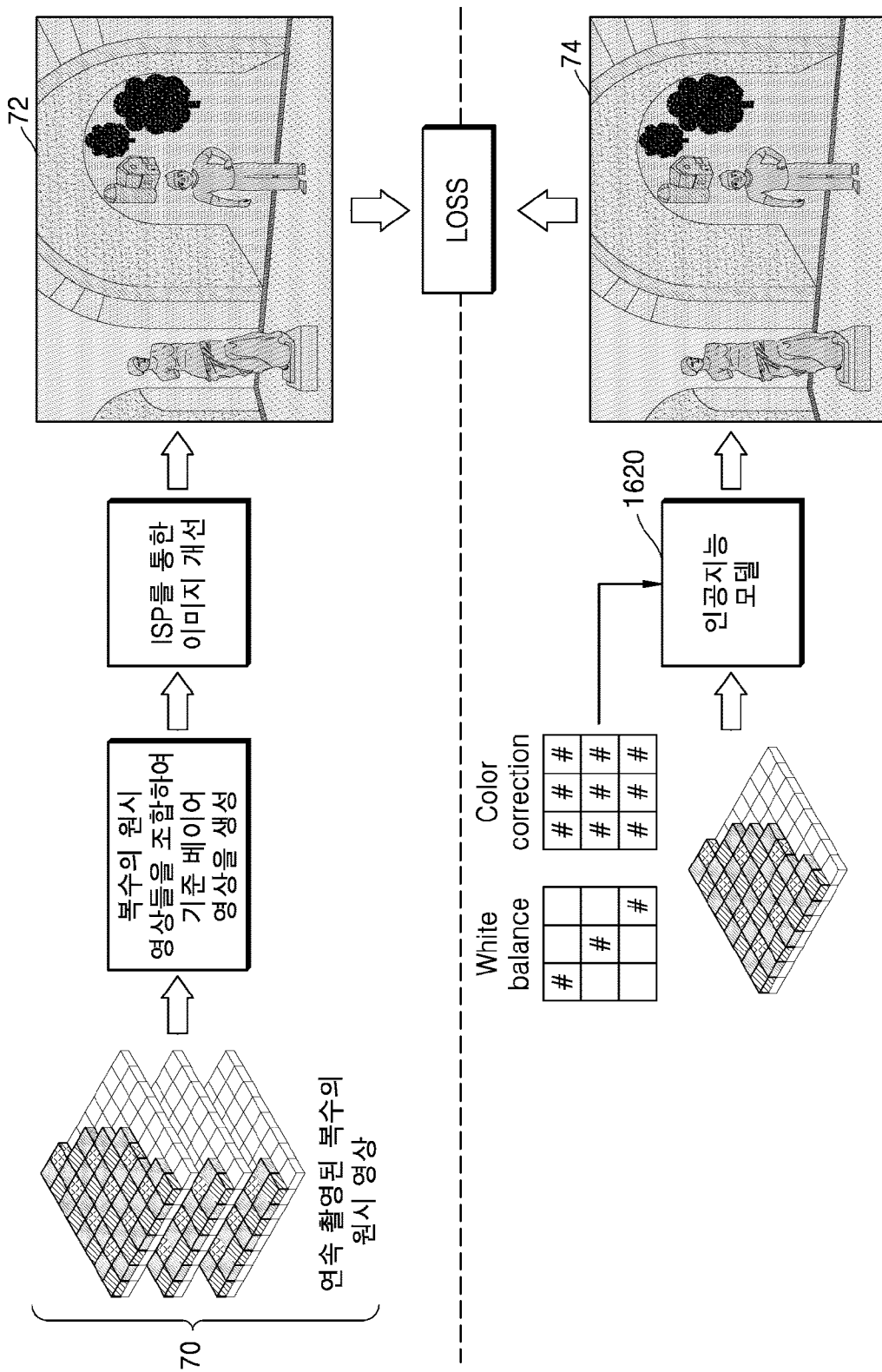
[도5]



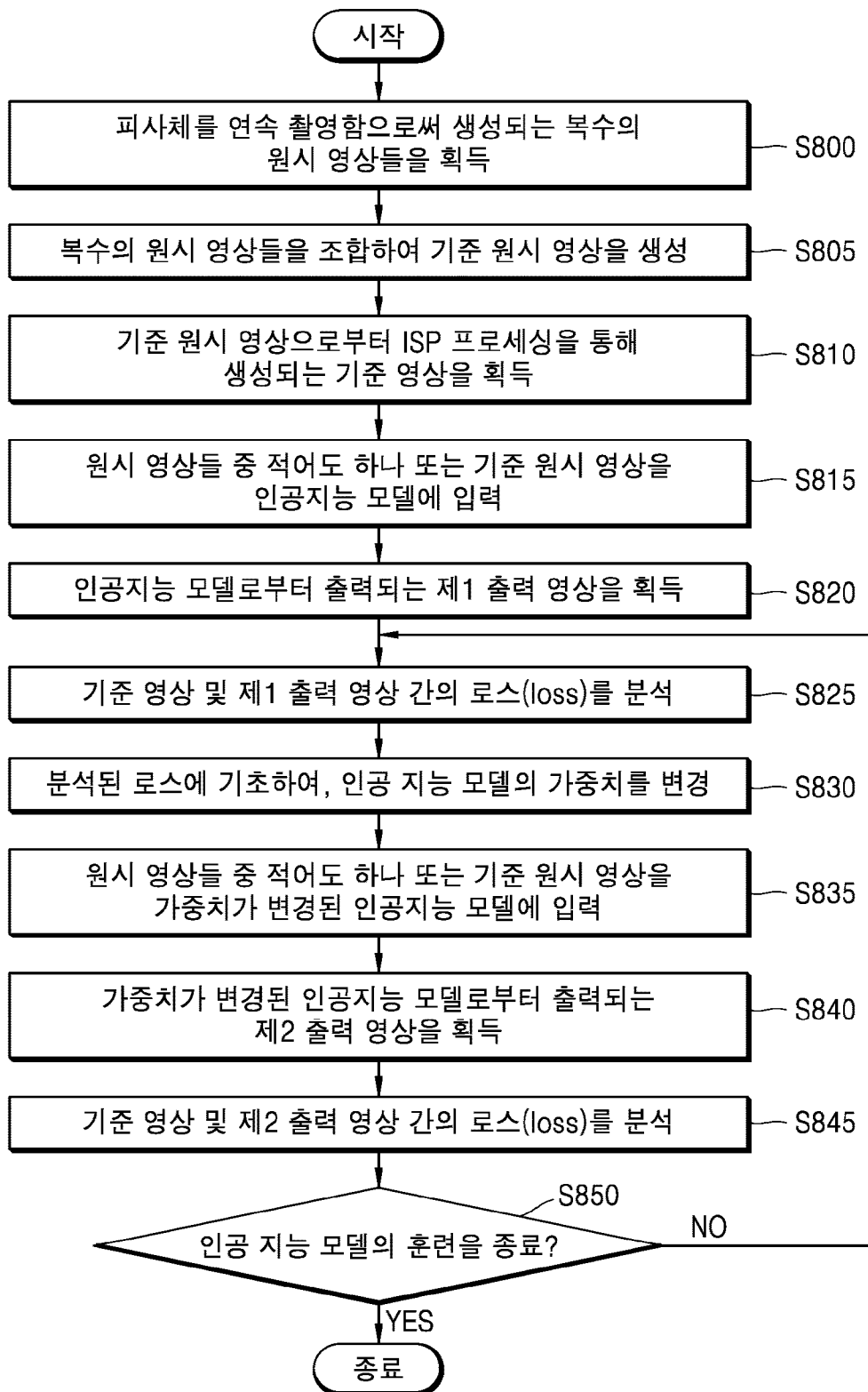
[도6]



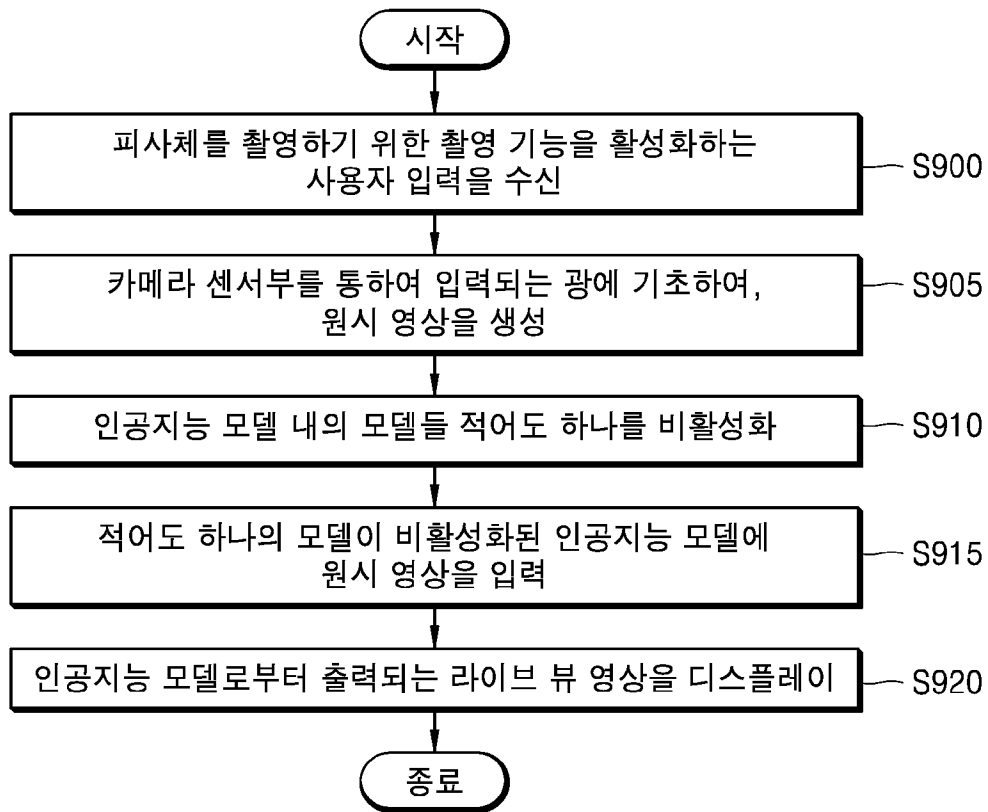
[도7]



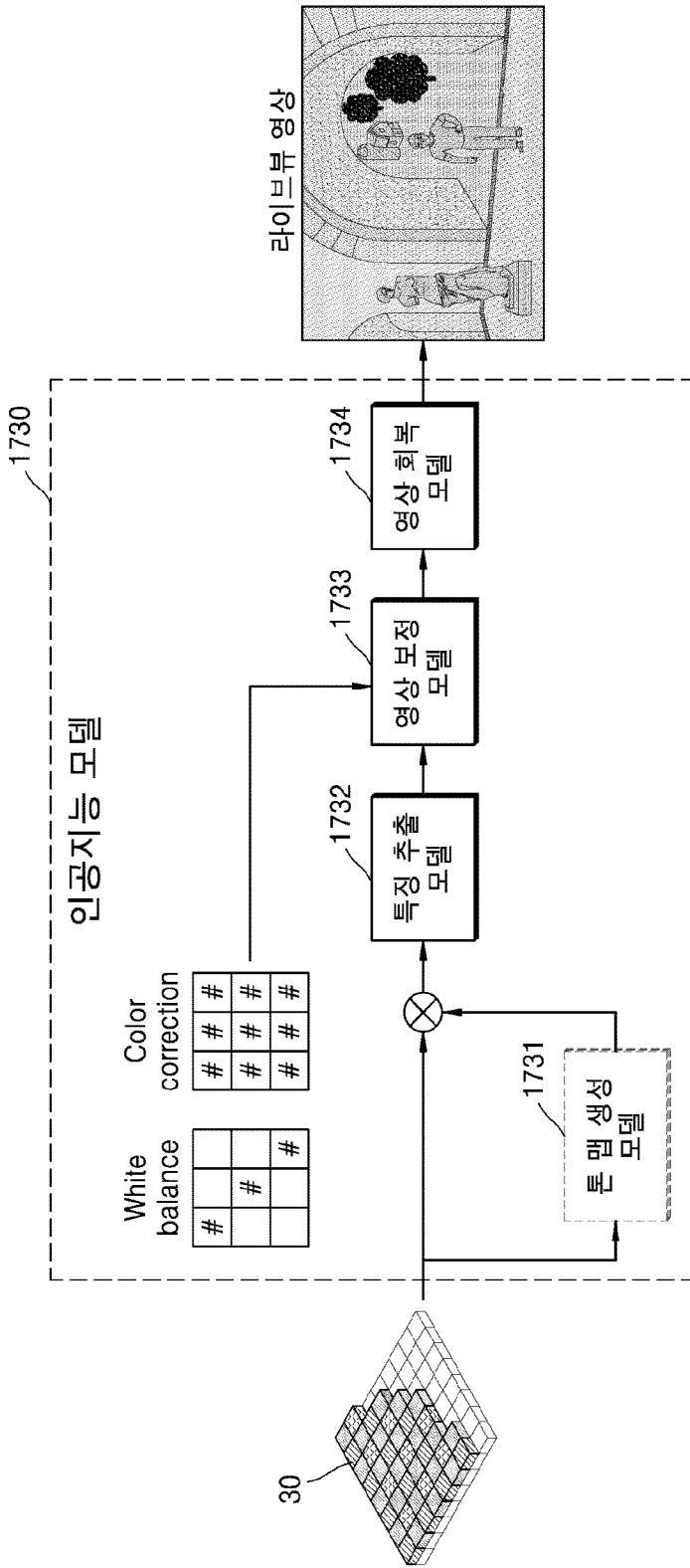
[도8]



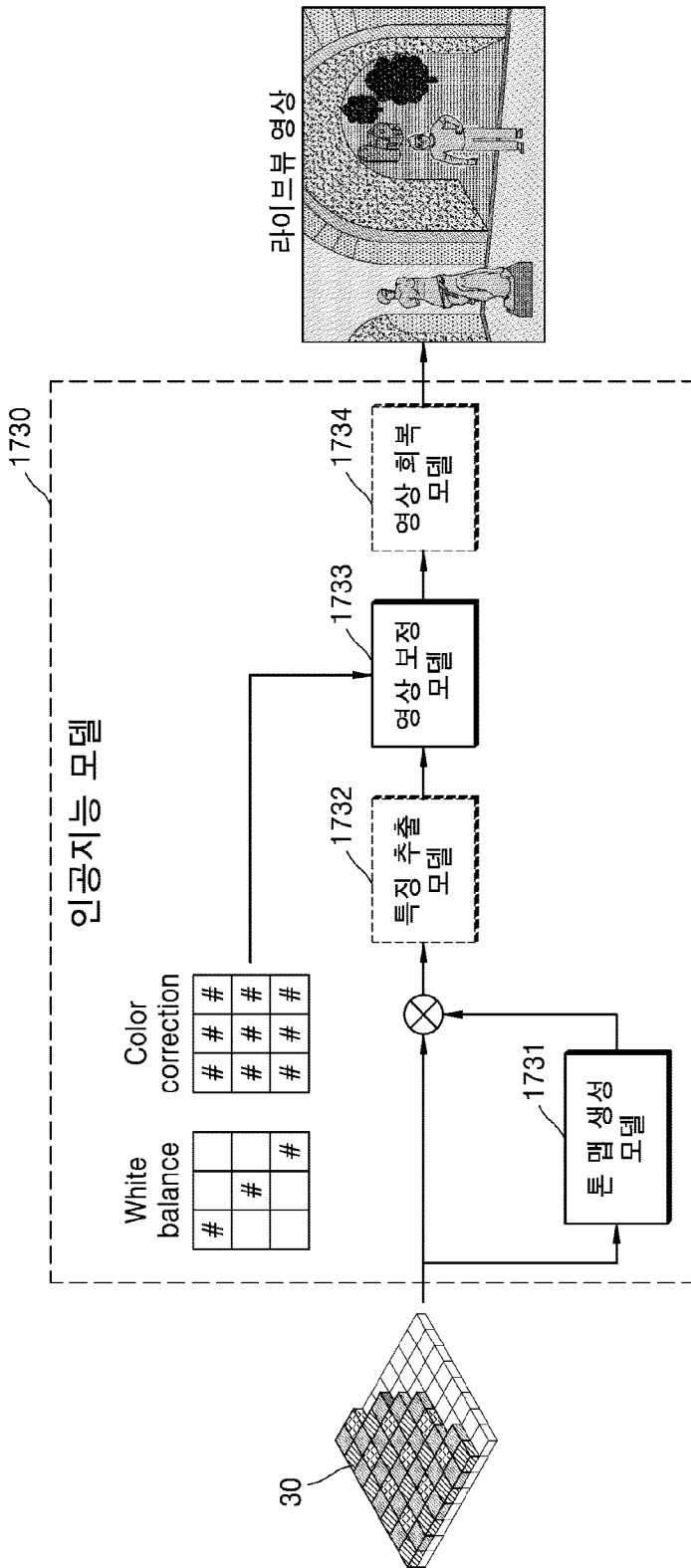
[도9]



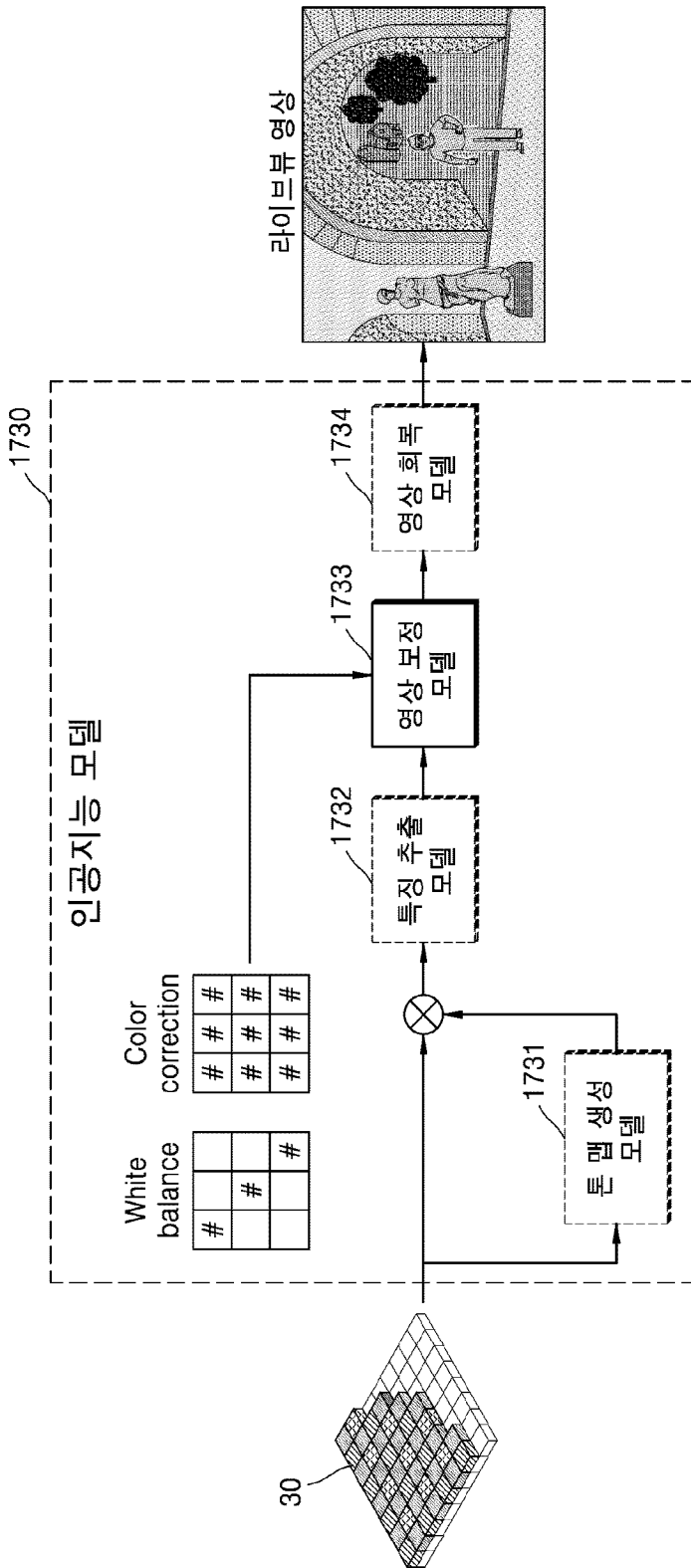
[도 10a]



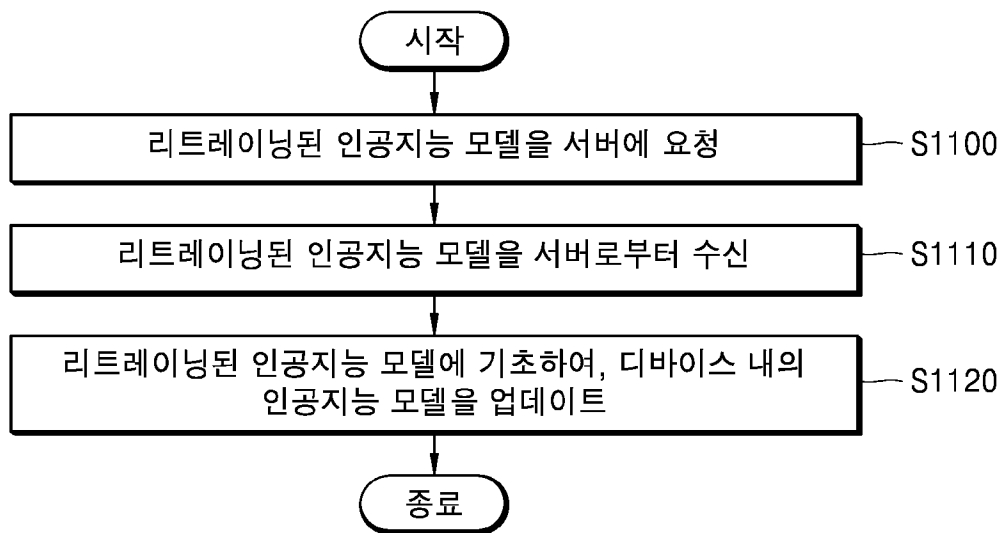
[도 10b]



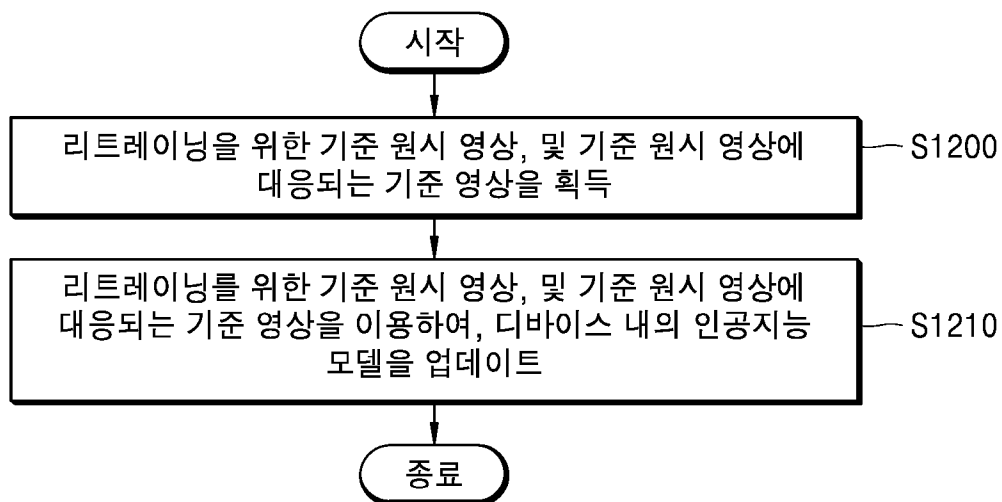
[도 10c]



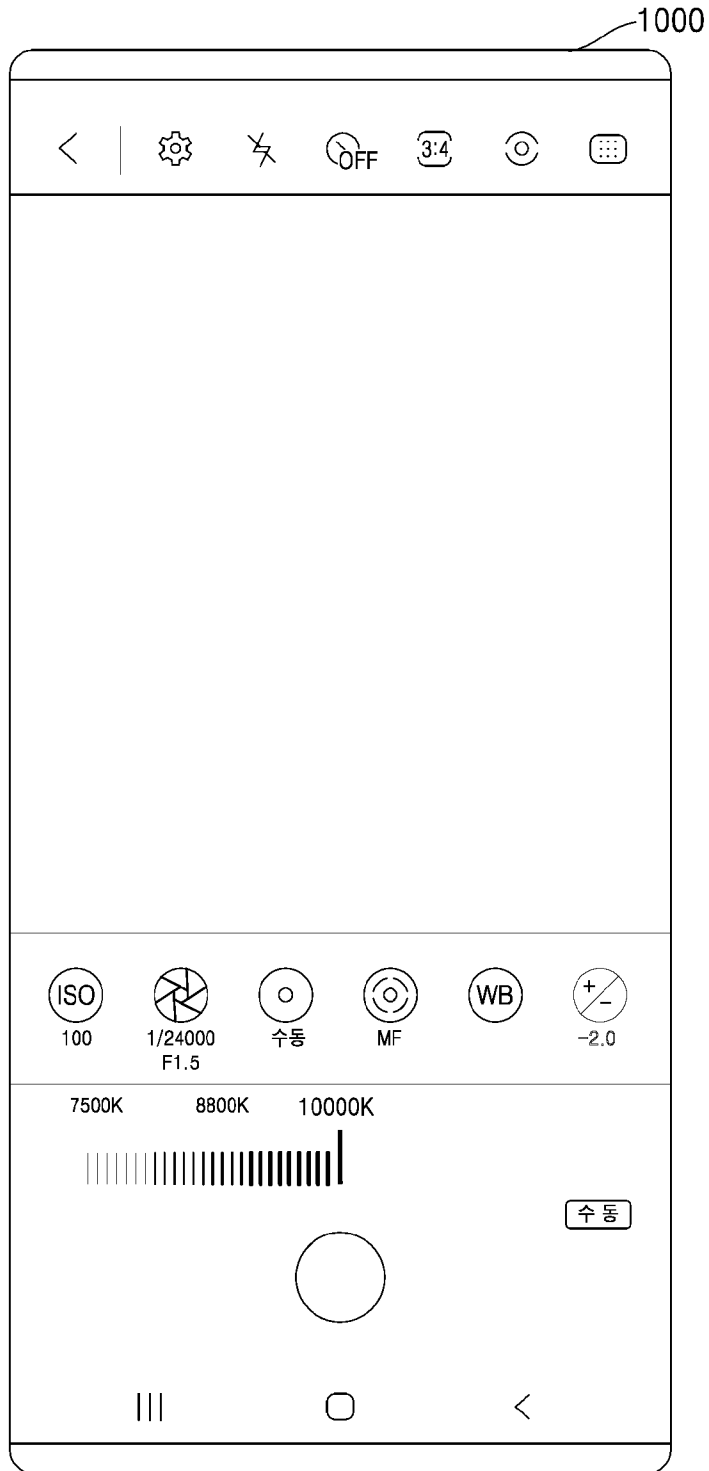
[도11]



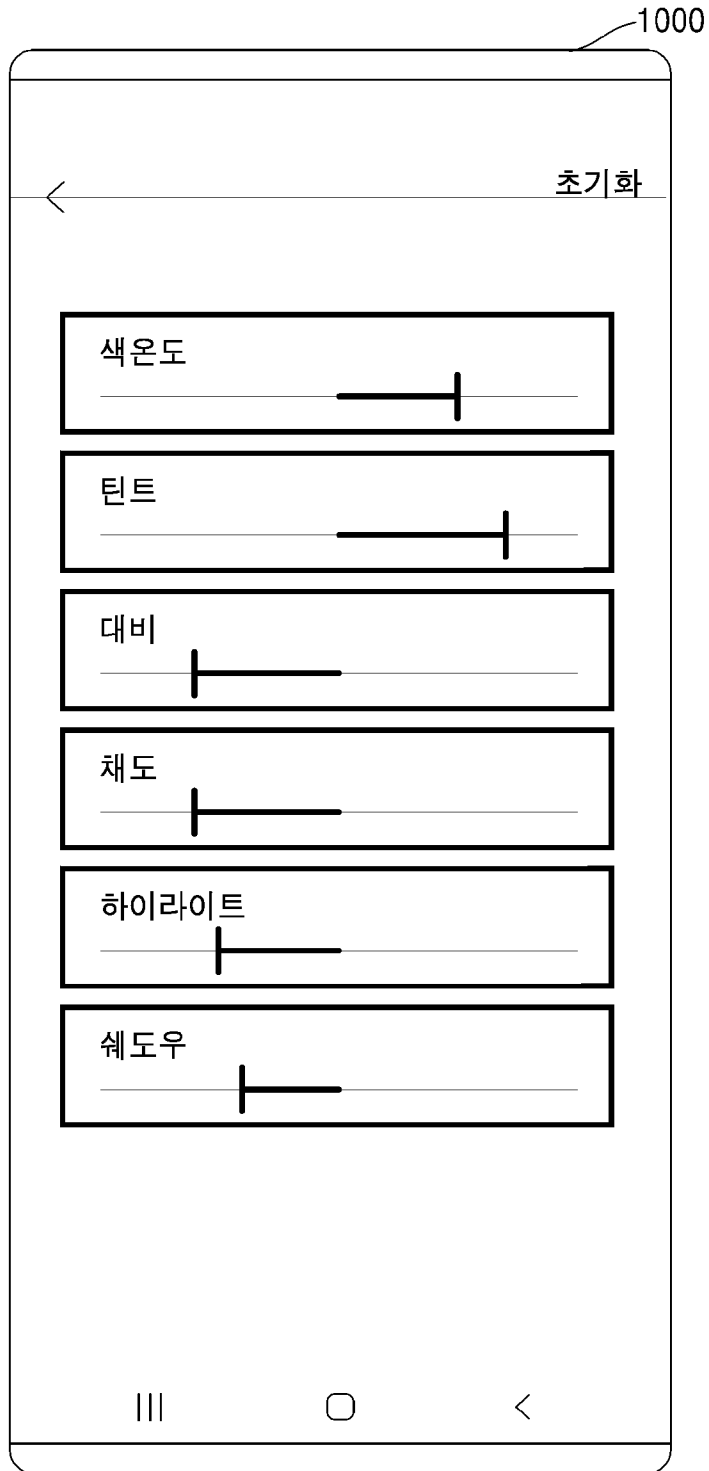
[도12]



[도 13a]



[도13b]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2021/015056

A. CLASSIFICATION OF SUBJECT MATTER		
G06T 5/00(2006.01)i; G06T 3/40(2006.01)i; G06N 3/08(2006.01)i; G06N 3/04(2006.01)i		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED		
Minimum documentation searched (classification system followed by classification symbols) G06T 5/00(2006.01); G06K 9/00(2006.01); G06K 9/62(2006.01); G06N 3/04(2006.01); G06N 3/08(2006.01); G06T 3/40(2006.01); H04N 5/225(2006.01); H04N 5/232(2006.01)		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Korean utility models and applications for utility models: IPC as above Japanese utility models and applications for utility models: IPC as above		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) eKOMPASS (KIPO internal) & keywords: 인공 지능(artificial intelligence), 원시 영상(raw image), 톤 매핑(tone mapping), 색상 보정(color correction), 선호도(preference), 라이브 뷰(live view), 리트레이닝(retraining)		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	KR 10-2017-0078516 A (SAMSUNG ELECTRONICS CO., LTD.) 07 July 2017 (2017-07-07) See paragraphs [0074]-[0075], [0080] and [0087]; and claims 6, 8-9, 12 and 14.	1,7,13,15
Y		2-6,8-12,14
Y	KR 10-2018-0055573 A (SAMSUNG ELECTRONICS CO., LTD.) 25 May 2018 (2018-05-25) See paragraphs [0017]-[0018], [0059]-[0060], [0068], [0092] and [0105].	2-6,8-12,14
Y	KR 10-2020-0009922 A (SAMSUNG ELECTRONICS CO., LTD.) 30 January 2020 (2020-01-30) See paragraphs [0272] and [0308].	10
A	KR 10-2029852 B1 (INDUSTRY ACADEMY COOPERATION FOUNDATION OF SEJONG UNIVERSITY) 08 October 2019 (2019-10-08) See claim 1.	1-15
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "D" document cited by the applicant in the international application "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search 07 February 2022		Date of mailing of the international search report 09 February 2022
Name and mailing address of the ISA/KR Korean Intellectual Property Office Government Complex-Daejeon Building 4, 189 Cheongsaro, Seo-gu, Daejeon 35208 Facsimile No. +82-42-481-8578		Authorized officer Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2021/015056

C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2020-0089994 A1 (CANON KABUSHIKI KAISHA) 19 March 2020 (2020-03-19) See paragraphs [0057]-[0058].	1-15
<hr/>		

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/KR2021/015056

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
KR	10-2017-0078516	A	07 July 2017	CN	106934426	A	07 July 2017
				US	10460231	B2	29 October 2019
				US	2017-0185871	A1	29 June 2017
KR	10-2018-0055573	A	25 May 2018	US	10474933	B2	12 November 2019
				US	2018-0137395	A1	17 May 2018
KR	10-2020-0009922	A	30 January 2020	US	2021-0272249	A1	02 September 2021
				WO	2020-017936	A1	23 January 2020
KR	10-2029852	B1	08 October 2019	None			
US	2020-0089994	A1	19 March 2020	JP	2020-046928	A	26 March 2020
				US	11188788	B2	30 November 2021

A. 발명이 속하는 기술분류(국제특허분류(IPC)) G06T 5/00(2006.01)i; G06T 3/40(2006.01)i; G06N 3/08(2006.01)i; G06N 3/04(2006.01)i		
B. 조사된 분야 조사된 최소문헌(국제특허분류를 기재) G06T 5/00(2006.01); G06K 9/00(2006.01); G06K 9/62(2006.01); G06N 3/04(2006.01); G06N 3/08(2006.01); G06T 3/40(2006.01); H04N 5/225(2006.01); H04N 5/232(2006.01) 조사된 기술분야에 속하는 최소문헌 이외의 문헌 한국등록실용신안공보 및 한국공개실용신안공보: 조사된 최소문헌란에 기재된 IPC 일본등록실용신안공보 및 일본공개실용신안공보: 조사된 최소문헌란에 기재된 IPC 국제조사에 이용된 전산 데이터베이스(데이터베이스의 명칭 및 검색어(해당하는 경우)) eKOMPASS(특허청 내부 검색시스템) & 키워드: 인공 지능(artificial intelligence), 원시 영상(raw image), 톤 매핑(ton e mapping), 색상 보정(color correction), 선호도(preference), 라이브 뷰(live view), 리트레이닝(retraining)		
C. 관련 문헌		
카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
X	KR 10-2017-0078516 A (삼성전자주식회사) 2017.07.07 단락 [0074]-[0075], [0080], [0087]; 및 청구항 6, 8-9, 12, 14	1,7,13,15
Y		2-6,8-12,14
Y	KR 10-2018-0055573 A (삼성전자주식회사) 2018.05.25 단락 [0017]-[0018], [0059]-[0060], [0068], [0092], [0105]	2-6,8-12,14
Y	KR 10-2020-0009922 A (삼성전자주식회사) 2020.01.30 단락 [0272], [0308]	10
A	KR 10-2029852 B1 (세종대학교 산학협력단) 2019.10.08 청구항 1	1-15
A	US 2020-0089994 A1 (CANON KABUSHIKI KAISHA) 2020.03.19 단락 [0057]-[0058]	1-15
<input type="checkbox"/> 추가 문헌이 C(계속)에 기재되어 있습니다. <input checked="" type="checkbox"/> 대응특허에 관한 별지를 참조하십시오.		
* 인용된 문헌의 특별 카테고리: “A” 특별히 관련이 없는 것으로 보이는 일반적인 기술수준을 정의한 문헌 “D” 본 국제출원에서 출원인이 인용한 문헌 “E” 국제출원일보다 빠른 출원일 또는 우선일을 가지나 국제출원일 이후에 공개된 선출원 또는 특허 문헌 “L” 우선권 주장에 의문을 제기하는 문헌 또는 다른 인용문헌의 공개일 또는 다른 특별한 이유(이유를 명시)를 밝히기 위하여 인용된 문헌 “O” 구두 개시, 사용, 전시 또는 기타 수단을 언급하고 있는 문헌 “P” 우선일 이후에 공개되었으나 국제출원일 이전에 공개된 문헌 “T” 국제출원일 또는 우선일 후에 공개된 문헌으로, 출원과 상충하지 않으며 발명의 기초가 되는 원리나 이론을 이해하기 위해 인용된 문헌 “X” 특별한 관련이 있는 문헌. 해당 문헌 하나만으로 청구된 발명의 신규성 또는 진보성이 없는 것으로 본다. “Y” 특별한 관련이 있는 문헌. 해당 문헌이 하나 이상의 다른 문헌과 조합하는 경우로 그 조합이 당업자에게 자명한 경우 청구된 발명은 진보성이 없는 것으로 본다. “&” 동일한 대응특허문헌에 속하는 문헌		
국제조사의 실제 완료일	국제조사보고서 발송일	
2022년02월07일 (07.02.2022)	2022년02월09일 (09.02.2022)	
ISA/KR의 명칭 및 우편주소	심사관	
대한민국 특허청 (35208) 대전광역시 서구 청사로 189, 4동 (둔산동, 정부대전청사)	양정록	
팩스 번호 +82-42-481-8578	전화번호 +82-42-481-5709	

국제조사보고서에서 인용된 특허문헌	공개일	대응특허문헌	공개일
KR 10-2017-0078516 A	2017/07/07	CN 106934426 A	2017/07/07
		US 10460231 B2	2019/10/29
		US 2017-0185871 A1	2017/06/29
KR 10-2018-0055573 A	2018/05/25	US 10474933 B2	2019/11/12
		US 2018-0137395 A1	2018/05/17
KR 10-2020-0009922 A	2020/01/30	US 2021-0272249 A1	2021/09/02
		WO 2020-017936 A1	2020/01/23
KR 10-2029852 B1	2019/10/08	없음	
US 2020-0089994 A1	2020/03/19	JP 2020-046928 A	2020/03/26
		US 11188788 B2	2021/11/30