(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0161425
(43) 공개일자 2024년11월12일

(51) 국제특허분류(Int. Cl.)
G06V 40/10 (2022.01) *G06N 3/02* (2023.01)
G06V 10/141 (2022.01) *G06V 10/147* (2022.01)
G06V 10/24 (2022.01) *G06V 10/72* (2022.01)
G06V 10/77 (2022.01) *G06V 10/82* (2022.01)
(52) CPC특허분류
G06V 40/103 (2022.01)
G06N 3/02 (2023.01)
(21) 출원번호 10-2023-0058556
(22) 출원일자 2023년05월04일
심사청구일자 2023년05월04일

(71) 출원인
포항공과대학교 산학협력단
경상북도 포항시 남구 청암로 77 (지곡동)
(72) 발명자
곽수하
경상북도 포항시 남구 청암로 77
임재성
경상북도 포항시 남구 청암로 77
(뒷면에 계속)
(74) 대리인
리앤목특허법인

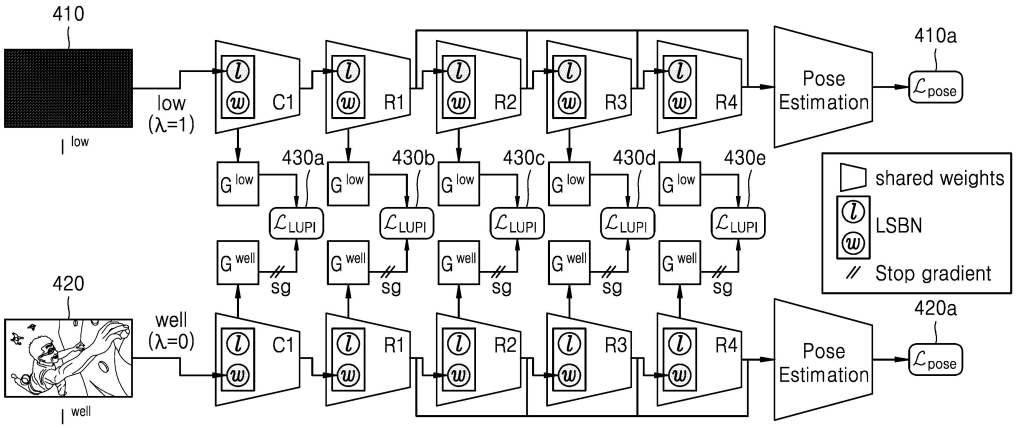
전체 청구항 수 : 총 15 항

(54) 발명의 명칭 **저조도 환경에서 인간자세추정을 학습하는 방법 및 장치**

(57) 요 약

본 발명의 바람직한 일 실시예에서 인간자세추정을 학습하는 장치는 밝은영상과 저조도영상을 동시에 획득하고, 밝은 영상에서 주석 처리를 수행하여 저조도 영상에 전달함으로써 인간 자세 추정을 위한 데이터셋을 구성한다. 그리고 데이터셋에 포함된 밝은영상을 교사모델의 입력으로, 저조도영상을 학생모델의 입력으로 이용하여, 학생모델이 교사모델의 특권정보를 이용하여 저조도 환경에서 높은 정확도로 인간자세추정을 학습하는 것을 특징으로 한다.

대 표 도



(52) CPC특허분류
G06V 10/141 (2023.08)
G06V 10/147 (2023.08)
G06V 10/24 (2023.08)
G06V 10/72 (2023.08)
G06V 10/77 (2022.01)
G06V 10/82 (2022.01)
G06T 2207/10152 (2013.01)
G06T 2207/20081 (2013.01)
G06T 2207/30196 (2013.01)
(72) 발명자
우병주
경상북도 포항시 남구 청암로 77
김건우
경상북도 포항시 남구 청암로 77
조성현
경상북도 포항시 남구 청암로 77
이소현
경상북도 포항시 남구 청암로 77
이해찬
경상북도 포항시 남구 청암로 77

명 세 서

청구범위

청구항 1

데이터셋구축부에서 밝은영상(I^{well})과 저조도영상(I^{low})을 하나의 쌍(I^{well},I^{low})의 형태로 이용하여 데이터셋을 구성하는 단계;

학습부에서 학생모델은 상기 데이터셋의 저조도영상을 기설정된 픽셀값으로 조정하여 입력받고, 교사모델은 상기 데이터셋의 밝은영상을 입력받으며, 상기 학생모델과 상기 교사모델은 공통된 인간자세추정손실로 학습하고, 상기 학생모델은 또한 상기 교사모델의 특권정보(Learning Under Privileged Information, LUPI)를 모방하면서 학습하는 단계;를 포함하고,

상기 교사모델과 상기 학생모델은 신경망에서 각각 별개의 배치정규화 파라미터를 이용하고, 그 외의 다른 파라미터를 공유하는 것을 특징으로 하는 저조도 환경에서 인간자세추정을 학습하는 방법.

청구항 2

제 1 항에 있어서, 특권 정보 사용 학습 손실($\mathcal{L}_{LUPI}$)을 통해

상기 학생모델의 그람행렬이 상기 교사모델의 그람행렬을 따르도록 학습하여 상기 밝은영상의 특징맵과 상기 저조도영상의 특징맵의 스타일 격차를 줄이는 것을 특징으로 하는 저조도 환경에서 인간자세추정을 학습하는 방법.

청구항 3

제 1 항에 있어서,

상기 밝은영상과 상기 저조도영상은 하나의 쌍(I^{well},I^{low})의 형태로 신경망에 입력되어, 서로 다른 배치정규화 레이어에서 서로 다른 배치정규화 방식으로 처리되는 것을 특징으로 하는 저조도 환경에서 인간자세추정을 학습하는 방법.

청구항 4

제 1 항에 있어서, 상기 밝은영상(I^{well})과 상기 저조도영상(I^{low})은 동시에 촬영되는 것을 특징으로 하는 저조도 환경에서 인간자세추정을 학습하는 방법.

청구항 5

제 1 항에 있어서,

동일한 장면을 동시에 촬영하여 상기 밝은영상(I^{well})과 상기 저조도영상(I^{low})을 획득하고, 호모그래피 행렬을 이용하여 상기 밝은영상(I^{well})과 상기 저조도영상(I^{low})의 위치 오차를 조정하는 것을 특징으로 하는 저조도 환경에서 인간자세추정을 학습하는 방법.

청구항 6

제 4 항에 있어서,

하나의 영상촬영장치에서 상기 밝은영상(I^{well})과 상기 저조도영상(I^{low})을 동시에 촬영하고, 상기 영상촬영장치는

밝은영상을 촬영하는 제 1 카메라모듈;

ND필터를 통해 저조도영상을 촬영하는 제 2 카메라모듈;및

렌즈로 들어온 빛을 균등하게 상기 제 1 카메라모듈과 상기 제 2 카메라모듈에 분배하는 빔스플리터;를 포함하

는 것을 특징으로 하는 저조도 환경에서 인간자세추정을 학습하는 방법.

청구항 7

제 4 항에 있어서,

영상촬영장치는 동기화된 셔터를 이용하여 상기 밝은영상과 상기 저조도영상을 동시에 촬영하여 하나의 쌍(I^{well},I^{low})의 형태로 획득하는 것을 특징으로 하는 저조도 환경에서 인간자세추정을 학습하는 방법.

청구항 8

제 1 항에 있어서,

상기 저조도영상에서 상기 인간자세추정손실의 예측값이 상기 밝은영상에서 주석처리한 결과와 유사해지도록 학습을 수행하는 것을 특징으로 하는 저조도 환경에서 인간자세추정을 학습하는 방법.

청구항 9

밝은영상(I^{well})과 저조도영상(I^{low})을 하나의 쌍(I^{well},I^{low})의 형태로 이용하여 데이터셋을 구성하는 데이터셋구축부;

학생모델은 상기 데이터셋의 저조도영상을 기설정된 픽셀값으로 조정하여 입력받고, 교사모델은 상기 데이터셋의 밝은영상을 입력받으며, 상기 학생모델과 상기 교사모델은 공통된 인간자세추정손실로 학습하고, 상기 학생모델은 또한 상기 교사모델의 특권정보(LUPI)를 모방하면서 학습하는 학습부;를 포함하고,

상기 교사모델과 상기 학생모델은 신경망에서 각각 별개의 배치정규화 레이어를 이용하고, 그 외의 다른 파라미터값을 공유하는 것을 특징으로 하는 저조도 환경에서 인간자세추정을 학습하는 장치.

청구항 10

제 9 항에 있어서, 상기 학습부는

특권 정보 사용 학습 손실($\mathcal{L}_{LUPI}$)을 통해 상기 학생모델의 그람행렬이 상기 교사모델의 그람행렬을 따르도록 학습하여 상기 밝은영상의 특징맵과 상기 저조도영상의 특징맵의 스타일 격차를 줄이는 것을 특징으로 하는 저조도 환경에서 인간자세추정을 학습하는 장치.

청구항 11

제 9 항에 있어서, 상기 저조도영상에서 상기 인간자세추정손실의 예측값이

동시에 촬영된 밝은영상에서 주석처리한 결과와 유사해지도록 학습을 수행하는 것을 특징으로 하는 저조도 환경에서 인간자세추정을 학습하는 것을 특징으로 하는 저조도 환경에서 인간자세추정을 학습하는 장치.

청구항 12

제 9 항에 있어서,

상기 밝은영상과 상기 저조도영상은 하나의 쌍(I^{well},I^{low})의 형태로 신경망에 입력되어, 서로 다른 배치정규화 레이어에서 서로 다른 배치정규화 방식으로 처리되는 것을 특징으로 하는 저조도 환경에서 인간자세추정을 학습하는 장치.

청구항 13

제 9 항에 있어서,

하나의 영상촬영장치에서 상기 밝은영상(I^{well})과 상기 저조도영상(I^{low})을 동시에 촬영하고, 상기 영상촬영장치는

밝은영상을 촬영하는 제 1 카메라모듈;

ND필터를 통해 저조도영상을 촬영하는 제 2 카메라모듈;및

렌즈로 들어온 빛을 균등하게 상기 제 1 카메라모듈과 상기 제 2 카메라모듈에 분배하는 빔스플리터;를 포함하는 것을 특징으로 하는 저조도 환경에서 인간자세추정을 학습하는 장치.

청구항 14

제 9 항에 있어서,

상기 저조도영상에서 상기 인간자세추정손실의 예측값이 상기 밝은영상에서 주석처리한 결과와 유사해지도록 학습을 수행하는 것을 특징으로 하는 저조도 환경에서 인간자세추정을 학습하는 장치.

청구항 15

제 9 항에 있어서,

상기 저조도영상에서 상기 인간자세추정손실의 예측값이 상기 밝은영상에서 주석처리한 결과와 유사해지도록 학습을 수행하는 것을 특징으로 하는 저조도 환경에서 인간자세추정을 학습하는 장치.

발명의 설명

기 술 분 야

[0001] 본 발명은 인간자세를 추정하는 방법에 관한 것이다. 보다 상세히 저조도 환경에서 인간자세추정을 학습하기 위한 방법에 관한 것이다. 본 연구는 삼성미래기술육성사업의 지원을 받아 수행되었음(과제번호: SRFC-IT1801-05).

배 경 기 술

[0002] 인간자세추정은 영상에서 인간의 관절의 위치를 추정하는 방법으로, 자율주행자동차, 감시카메라 시스템 등에서 이용되고 있다. 그러나, 야간과 같이 빛이 부족한 환경에서 촬영된 저조도 영상들은 주로 저 노출, 잡음, 블러 등의 열화에 의해 크게 훼손되어 인간자세추정이 쉽지 않은 문제가 있다.

[0003] 기존에는 저조도 영상을 밝은 영상으로 복원하거나, 쌍을 이루지 않는 저조도 영상과 밝은 영상을 이용하여 융합하는 방법을 이용하였다. 그러나, 저조도 영상을 밝은 영상으로 개선하는 기술은 계산량이 과도한 문제가 있다. 또한 NOD 데이터셋이나 ARID 데이터셋을 이용하여 인간자세를 추정하는 방법은 트레이닝을 위해 저조도 영상만을 제공하여 정확도가 높지 않으며, 극한의 저조도 상황에 대해서는 학습하지 않는 한계가 있다.

선행기술문헌

특허문헌

[0004] (특허문헌 0001) KR 10-2021-0036840 A

발명의 내용

해결하려는 과제

[0005] 본 발명의 바람직한 일 실시예에서는 극한의 저조도 환경에서도 인간 자세를 추정할 수 있는 데이터셋을 구축하고자 한다.

[0006] 본 발명의 또 다른 바람직한 일 실시예에서는 저조도 환경에서 인간 자세 추정을 학습할 수 있는 방법을 제공하고자 한다.

과제의 해결 수단

[0007] 본 발명의 바람직한 일 실시예로서, 저조도 환경에서 인간자세추정을 학습하는 방법은 데이터셋구축부에서 밝은영상(I^{well})과 저조도영상(I^{low})을 하나의 쌍(I^{well},I^{low})의 형태로 이용하여 데이터셋을 구성하는 단계; 학습부에서

학생모델은 상기 데이터셋의 저조도영상을 기설정된 픽셀값으로 조정하여 입력받고, 교사모델은 상기 데이터셋의 밝은영상을 입력받으며, 상기 학생모델과 상기 교사모델은 공통된 인간자세추정손실로 학습하고, 상기 학생모델은 또한 상기 교사모델의 특권정보(Learning Under Privileged Information, LUPI)를 모방하면서 학습하는 단계;를 포함하고, 상기 교사모델과 상기 학생모델은 신경망에서 각각 별개의 배치정규화 파라미터를 이용하고, 그 외의 다른 파라미터를 공유하는 것을 특징으로 한다.

[0008] 본 발명의 바람직한 일 실시예로서, 저조도 환경에서 인간자세추정을 학습하는 방법은 특권 정보 사용 학습 손실($\mathcal{L}_{LUPI}$)을 통해 학생모델의 그람행렬이 교사모델의 그람행렬을 따르도록 학습하여 상기 밝은영상의 특징맵과 상기 저조도영상의 특징맵의 스타일 격차를 줄이는 것을 특징으로 한다.

[0009] 본 발명의 바람직한 일 실시예로서, 저조도 환경에서 인간자세추정을 학습하는 방법은 하나의 영상촬영장치에서 상기 밝은영상(I^{well})과 상기 저조도영상(I^{low})을 동시에 촬영하고, 상기 영상촬영장치는 밝은영상을 촬영하는 제 1 카메라모듈; ND필터를 통해 저조도영상을 촬영하는 제 2 카메라모듈;및 렌즈로 들어온 빛을 균등하게 상기 제 1 카메라모듈과 상기 제 2 카메라모듈에 분배하는 빔스플리터;를 포함하는 것을 특징으로 한다.

[0010] 상기 영상촬영장치는 동기화된 셔터를 이용하여 상기 밝은영상과 상기 저조도영상을 동시에 촬영하여 하나의 쌍(I^{well},I^{low})의 형태로 획득하는 것을 특징으로 한다. 동시에 촬영된 상기 밝은영상(I^{well})과 상기 저조도영상(I^{low})은 호모그래피 행렬을 이용하여 위치오차를 조정하는 것을 특징으로 한다.

[0011] 본 발명의 바람직한 일 실시예로서, 저조도 환경에서 인간자세추정을 학습하는 방법에서 상기 밝은영상과 상기 저조도영상은 하나의 쌍(I^{well},I^{low})의 형태로 신경망에 입력되어, 서로 다른 배치정규화 레이어에서 서로 다른 배치정규화 방식으로 처리되는 것을 특징으로 한다.

[0012] 본 발명의 또 다른 바람직한 일 실시예로서, 저조도 환경에서 인간자세추정을 학습하는 장치는 밝은영상(I^{well})과 저조도영상(I^{low})을 하나의 쌍(I^{well},I^{low})의 형태로 이용하여 데이터셋을 구성하는 데이터셋구축부; 학생모델은 상기 데이터셋의 저조도영상을 기설정된 픽셀값으로 조정하여 입력받고, 교사모델은 상기 데이터셋의 밝은영상을 입력받으며, 상기 학생모델과 상기 교사모델은 공통된 인간자세추정손실로 학습하고, 상기 학생모델은 또한 상기 교사모델의 특권정보(LUPI)를 모방하면서 학습하는 학습부;를 포함하고, 상기 교사모델과 상기 학생모델은 신경망에서 각각 별개의 배치정규화 레이어를 이용하고, 그 외의 다른 파라미터값을 공유하는 것을 특징으로 한다.

발명의 효과

[0013] 본 발명의 바람직한 일 실시예에서는 밝은영상과 저조도영상을 동시에 획득하고, 밝은 영상에서 주석 처리를 수행하여 저조도 영상에 전달함으로써 인간 자세 추정을 위한 데이터셋을 구성할 수 있다.

[0014] 본 발명의 바람직한 일 실시예에서는 학생 모델은 저조도영상을, 교사 모델은 밝은영상을 입력받도록 교사-학습 모델을 구성하고, 학생모델과 교사모델은 공통된 인간자세추정손실로 학습하며, 학생모델은 또한 교사모델의 특권정보를 전달받도록 신경망을 학습하여 저조도 환경에서 높은 정확도로 인간자세추정을 학습할 수 있는 효과가 있다.

도면의 간단한 설명

[0015] 도 1 은 본 발명의 바람직한 일 실시예로서, 저조도 환경에서 인간자세 추정을 위한 데이터셋을 구축하기 위해 이용하는 영상촬영장치의 내부 구성도를 도시한다.

도 2 는 본 발명의 바람직한 일 실시예로서, 영상촬영장치에서 동시에 획득한 밝은영상과 저조도 영상의 쌍을 도시한다.

도 3 은 본 발명의 바람직한 일 실시예로서, 인간자세추정을 학습하는 장치의 내부 구성도를 도시한다.

도 4 는 본 발명의 바람직한 일 실시예로서, 인간자세추정을 학습하는 장치에서 신경망을 통해 인간자세추정을 학습하는 과정을 도시한다.

발명을 실시하기 위한 구체적인 내용

[0016] 이하에서 도면을 참고하여 설명한다.

[0017] 도 1 은 본 발명의 바람직한 일 실시예로서, 저조도 환경에서 인간자세 추정을 위한 데이터셋을 구축하기 위해 이용하는 영상촬영장치의 내부 구성도를 도시한다.

[0018] 본 발명의 바람직한 일 실시예로서, 영상촬영장치(100)는 렌즈, 두 개의 카메라 모듈(110, 120), 빔스플리터(130) 및 ND-필터(140)를 포함한다.

[0019] 본 발명의 영상촬영장치(100)는 같은 장면에서 밝은 영상 및 저조도 영상을 동시에 획득한다. 상세히, 빔스플리터(130)는 렌즈로 들어온 빛을 균등하게 두 개의 카메라 모듈(110, 120)에 균등하게 분배한다. 두 개의 카메라 모듈(110, 120) 중 제 1 카메라모듈(110)은 밝은 영상을 촬영하고, 제 2 카메라모듈(120)은 저조도 영상을 촬영한다. 도 1 에서 제 1 카메라모듈(110)과 제 2 카메라모듈(120)는 위치가 서로 변경될 수 있다.

[0020] 제 2 카메라모듈(120)은 ND필터(140)를 이용하여 빔스플리터(130)로부터 들어온 빛의 양을 예를 들어 100배 줄여, 제 1 카메라모듈(110)에서 촬영된 영상보다 100배 어두운 저조도 영상을 촬영할 수 있다.

[0021] 영상촬영장치(100)는 동기화된 셔터를 이용하여 두 개의 카메라 모듈(110, 120)로부터 밝은영상(I^{well})과 저조도 영상(I^{low})을 동시에 촬영하여 하나의 쌍(I^{well},I^{low})의 형태로 획득하여 메모리에 저장한다.

[0022] 도 2 는 본 발명의 바람직한 일 실시예로서, 영상촬영장치(100)에서 카메라 모듈(110, 120)의 노출시간과 게인값을 조절하여 획득한 밝은영상(I^{well})(210)과 저조도 영상(I^{low})(220)의 쌍(I^{well},I^{low})을 도시한다. 도 2 에서 저조도 영상(220)은 시각화를 위해 30배 스케일링된 결과이다. 밝은영상(I^{well})(210)과 저조도 영상(I^{low})(220)은 호모그래피 행렬 등을 이용하여 위치적인 오차를 조정하여 정렬된다.

[0023] 본 발명의 바람직한 일 실시예로서, 인간자세추정을 학습하는 장치(300)는 밝은영상(I^{well})(210)에 일반적인 인간자세추정 기술을 적용하여 주석을 획득한다. 적용획득된 주석에 오류가 있는 경우 추가적인 수정이 가능하다. 인간자세추정을 학습하는 장치(300)는 저조도 영상(I^{low})(220)에서 밝은영상(I^{well})(210)에서 획득하고 수정한 주석을 그대로 이용한다.

[0024] 도 3 은 본 발명의 바람직한 일 실시예로서, 인간자세추정을 학습하는 장치(300)의 내부 구성도를 도시한다.

[0025] 인간자세추정학습장치(300)는 도 1에 도시된 영상촬영장치를 통해 동시에 획득한 밝은영상 및 저조도영상을 이용하여 데이터셋을 생성한다. 그리고, 도 4를 참고하면, 인간자세추정학습장치(300)는 밝은영상을 이용하는 교사모델과 저조도영상을 이용하는 학생모델 각각에서 동일한 인간자세추정손실 $\mathcal{L}_{pose}$(410a, 420a)로 학습하고, 학생모델은 교사모델의 특권정보 LUPI(430a, 430b, 430c, 430d, 430e)를 모방하면서 인간자세를 예측하도록 학습한다.

[0026] 인간자세추정학습장치(300)의 내부 구성을 참고하여 설명한다.

[0027] 인간자세추정학습장치(300)는 데이터셋구축부(310) 및 학습부(320)를 포함한다.

[0028] 데이터셋구축부(310)는 영상촬영장치(100)에서 동시에 획득한 밝은영상(I^{well})과 저조도 영상(I^{low})의 쌍(I^{well},I^{low})을 인간자세추정 학습을 위한 데이터셋으로 이용한다. 이 경우, 저조도 영상(I^{low})은 그에 대응하는 밝은영상(I^{well})에서 획득한 정확도가 높은 주석 정보를 이용한다.

[0029] 인간자세추정학습장치(300)는 인간자세추정을 학습하기 위해 저조도영상을 이용하므로 영상의 품질이 낮아 학습이 어렵다. 이러한 문제점을 해결하기 위하여 학습부(320)는 지식 증류(Knowledge Distillation) 기반으로 교사-학생모델을 이용한다. 학습부(320)에서 교사모델은 데이터셋의 밝은영상을 입력받는다. 학생모델은 데이터셋의 저조도영상을 기설정된 픽셀값으로 조정하여 입력받는다. 기설정된 픽셀값은 예를 들어 0.4이다.

[0030] 학습부(320)에서 학생모델과 교사모델은 수학식 1과 같은 공통된 인간자세추정손실로 학습한다. 저조도 영상과 밝은영상은 동일한 장면을 촬영한 영상이므로, 저조도영상은 밝은영상에 대해 주석처리된 값을 인간자세에 대한 정답으로 공유한다. 또한, 학생모델은 교사모델의 특권정보(Learning Under Privileged Information, LUPI)를 모방하면서 인간자세를 예측하도록 학습을 수행한다.

수학식 1

[0031]
$$\mathcal{L}_{\text{pose}}(P,Y)=\frac{1}{K}\sum_{i=1}^{K}||P_i-Y_i||_2^2$$

[0032] 수학식 1에서 Pi와 Yi는 각각 i번째 관절에 대해 예측된 히트맵과 예측값을 나타내고, K는 관절 수를 나타낸다.

[0033] 도 4를 참고하면, 인간자세추정학습장치(400)에서 학생모델과 교사모델은 공통된 자세에 대한 정답을 추정하도록 동일한 인간자세추정손실 $\mathcal{L}_{\text{pose}}$(410a, 420a)로 학습된다. 그리고, 학생모델은 교사모델의 특권정보 LUPI(430a, 430b, 430c, 430d, 430e)를 모방하도록 추가적인 감독을 받는다.

[0034] 도 4 는 본 발명의 바람직한 일 실시예에서, 학생모델과 교사모델을 통합한 인간자세추정을 학습하는 장치의 내부 구성도를 도시한다.

[0035] 인간자세추정학습장치(400)에서 학생모델과 교사모델은 신경망 내에서 각각 별개의 배치 정규화 파라미터를 이용하고, 상기 신경망 내의 다른 파라미터를 공유한다. 도 4 에서 신경망 내의 다른 파라미터의 예로 가중치가 표시되어 있다.

[0036] 학생모델과 교사모델을 통합한 인간자세추정학습장치(400)는 영상의 조명조건에 따라 수학식 2와 같이 배치 정규화 레이어를 분리한다.

수학식 2

[0037]
$$\begin{aligned}\text{LSBN}(\mathbf{x},\lambda)=&\lambda\left(\gamma^{\text{low}}\cdot\frac{\mathbf{x}-\mu}{\sqrt{\sigma^2+\epsilon}}+\beta^{\text{low}}\right)\\&+(1-\lambda)\left(\gamma^{\text{well}}\cdot\frac{\mathbf{x}-\mu}{\sqrt{\sigma^2+\epsilon}}+\beta^{\text{well}}\right)\end{aligned}$$

[0038] 수학식 2에서 LSBN(lighting-condition specific batch normalization)은 조명 조건 별 배치 정규화를 나타낸다. λ는 조명조건을 나타낸다. 예를 들어, λ는 x 의 조명 조건이 저조도(low)이면 1이고 밝은 조건(well)이면 0의 값을 가진다. x는 저조도영상 또는 밝은영상으로부터 계산된 채널을 나타낸다. μ과 σ^2은 x에서의 활성화의 평균과 분산을 의미하고, ε은 수치 안정성을 위한 아주 작은 값을 나타낸다.

[0039] 인간자세추정학습장치(400)의 학습부는 밝은 영상(I^{well})과 이에 대응하는 저조도 영상(I^{low}) 을 함께 입력받은 후, 신경망 내에서 밝은 영상에 이용하는 정규화레이어와 저조도 영상에 이용하는 정규화레이어를 각각 분리하고, 그 외의 모든 파라미터를 공유한다.

[0040] 밝은 영상(I^{well})은 제 1 정규화레이어에서 아핀 변환 파라미터 (γ^{well},β^{well})를 이용하고, 저조도 영상(I^{low})은 제 2 정규화레이어에서 아핀 변환 파라미터(γ^{low},β^{low})를 이용한다.

[0041] 학습부는 영상이 저조도인지 밝은 조도인지와 같은 조명 조건에 따라 각각 별개의 배치정규화 파라미터를 이용하여 배치 정규화를 수행한다. 그리고, 학생모델이 수학식 3과 같은 특권정보사용 학습손실($\mathcal{L}_{\text{LUPI}}$)을 통해 교사모델로부터 추출된 특권 정보를 따르도록 하는 추가적인 감독을 받음으로써 저조도와 밝은 조도 간의 특징맵의 스타일 차이를 줄일 수 있다.

수학식 3

[0042] $$\mathcal{L}_{\text{LUPI}} = \sum_{l} \frac{1}{4C_l^2 N_l^2} \sum_{i=1}^{C_l} \sum_{j=1}^{C_l} \left(\mathbf{G}_{i,j}^{\text{low},l} - \mathbf{G}_{i,j}^{\text{well},l} \right)^2$$

[0043] 수학식 3에서 G는 그람행렬, C_l와 N_l는 채널 수와 l번째 레이어로부터 특징 맵의 공간적 크기를 의미한다. $G^{low,l}$은 학생 모델의 l번째 레이어의 특징 맵으로부터 계산된 저조도 영상의 그람 행렬, $G^{well,l}$은 교사 모델의 l번째 레이어의 특징 맵으로부터 계산된 밝은 영상의 그람 행렬을 나타낸다. 본 발명의 바람직한 일 실시예에서는 학생 모델이 수학식 3의 특권정보사용 학습손실($\mathcal{L}_{LUPI}$)이 최소화하도록 교사 모델의 스타일을 모방한다.

[0044] 도 4 를 참고하면, 저조도 영상(410)의 특징맵 스타일은 그람 행렬 G^{low}, 밝은 영상(420)의 특징맵 스타일은 그람 행렬 G^{well}로 표시되었다. 도 4에서 특권정보사용 학습손실($\mathcal{L}_{LUPI}$)은 도 4 에 도시된 신경망인 ResNet의 제 1 컨볼루션레이어(C1)과 다른 4개의 잔여블록 R1, R2, R3 및 R4의 특징맵에 모두 적용된다(430a, 430b, 430c, 430d, 430e).

[0045] 본 발명의 바람직한 일 실시예로서, 학습부는 밝은 영상(I^{well})과 이에 대응하는 저조도 영상(I^{low})간의 특징맵 스타일 차이를 줄이기 위해, 스탑그레디언트 sg(stop gradient)를 통해 G^{low}가 G^{well}을 따르도록 학습한다.

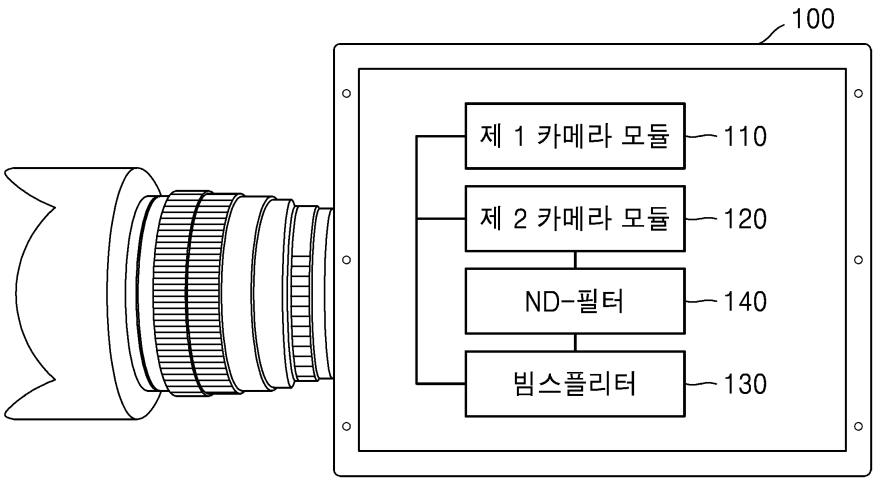
[0046] 이상에서 설명된 장치는 하드웨어 구성요소, 소프트웨어 구성요소, 및/또는 하드웨어 구성요소 및 소프트웨어 구성요소의 조합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치 및 구성요소는, 예를 들어, 프로세서, 콘트롤러, ALU(arithmetic logic unit), 디지털 신호 프로세서(digital signal processor), 마이크로컴퓨터, FPGA(field programmable gate array), PLU(programmable logic unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 하나 이상의 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(OS) 및 상기 운영 체제 상에서 수행되는 하나 이상의 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 콘트롤러를 포함할 수 있다. 또한, 병렬 프로세서(parallel processor)와 같은, 다른 처리 구성(processing configuration)도 가능하다.

[0047] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 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.

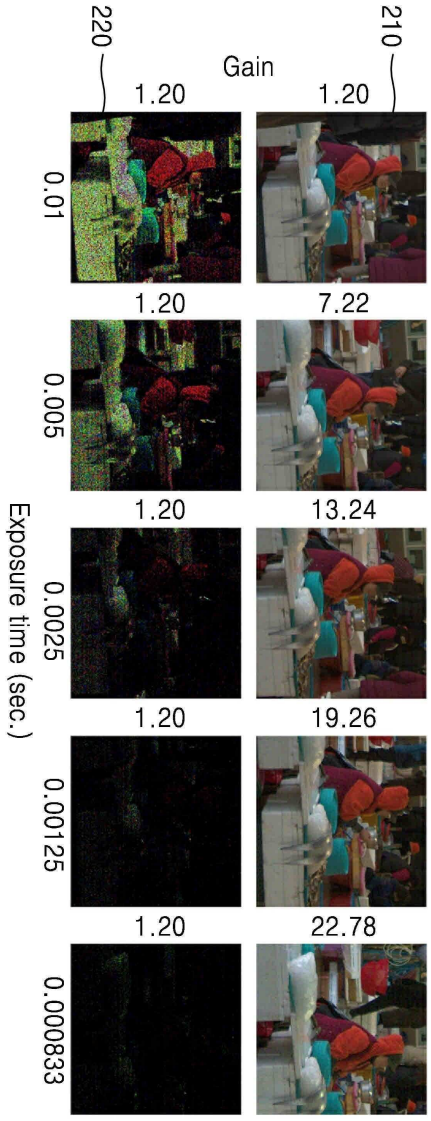
[0048] 이상과 같이 실시예들이 비록 한정된 실시예와 도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 상기의 기재로부터 다양한 수정 및 변형이 가능하다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다. 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

도면

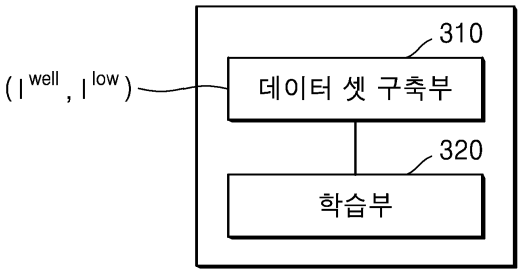
도면1


100
제 1 카메라 모듈
110
제 2 카메라 모듈
120
ND-필터
140
빔스플리터
130

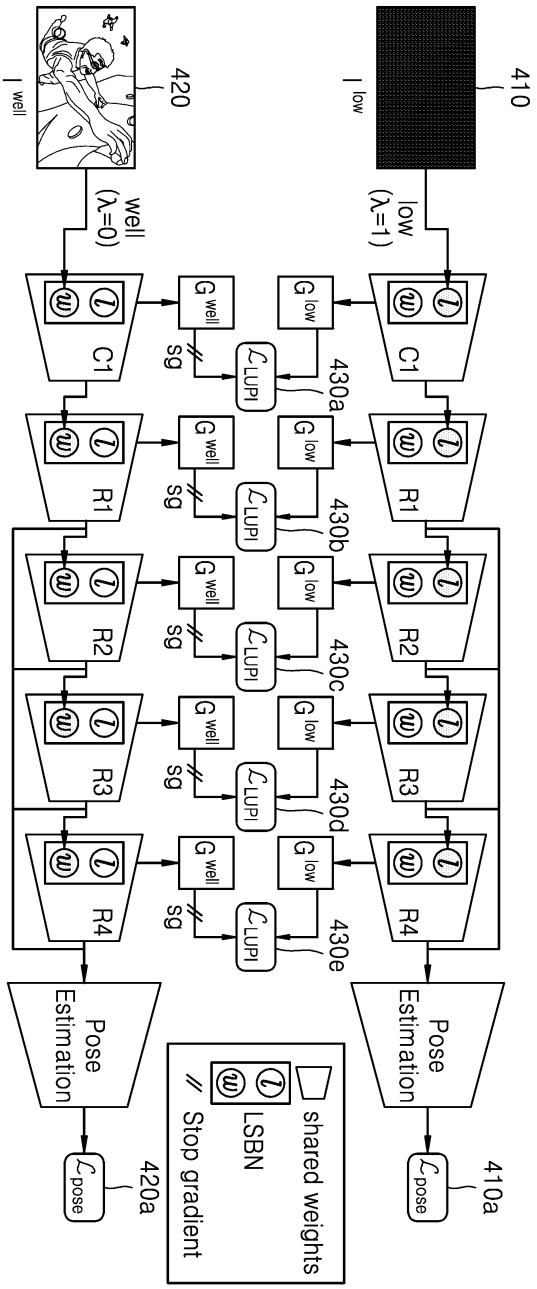
도면2


220
210
Gain
1.20
1.20
1.20
7.22
1.20
13.24
1.20
19.26
1.20
22.78
Exposure time (sec.)
0.01
0.005
0.0025
0.00125
0.000833

도면3


(I^{well}, I^{low})
310
데이터 셋 구축부
320
학습부

도면4


I_{well}
420
I_{low}
410
well (λ=0)
low (λ=1)
C1
R1
R2
R3
R4
G well
G low
sg
$\mathcal{L}_{LUPI}$
430a
430b
430c
430d
430e
Pose Estimation
$\mathcal{L}_{pose}$
420a
410a
shared weights
LSBN
// Stop gradient