



(19) 대한민국특허청(KR)  
(12) 등록특허공보(B1)

(45) 공고일자 2022년02월15일  
(11) 등록번호 10-2363049  
(24) 등록일자 2022년02월10일

(51) 국제특허분류(Int. Cl.)  
G06T 5/00 (2019.01) G06T 7/00 (2017.01)  
(52) CPC특허분류  
G06T 5/002 (2013.01)  
G06T 7/00 (2013.01)  
(21) 출원번호 10-2020-0176867  
(22) 출원일자 2020년12월16일  
심사청구일자 2020년12월16일  
(65) 공개번호 10-2021-0086493  
(43) 공개일자 2021년07월08일  
(30) 우선권주장  
1020190178478 2019년12월30일 대한민국(KR)  
(56) 선행기술조사문헌  
JP2019212275 A\*  
Amir Atapour-Abarghouei, Toby P. Breckon,  
"Real-Time Monocular Depth Estimation using  
Synthetic Data with Domain Adaptation via  
Image Style Transfer", Proceedings of the  
IEEE Conference on Computer Vision and  
Pattern Recognition (CVPR), 2018, pp.  
2800-2810(2018.06.18.) 1부.\*  
Marcela Carvalho, Bertrand Le Saux, Pauline  
Trouve-Peloux, Andres Almansa, Frederic  
Champagnat, "Deep Depth from Defocus: how can  
defocus blur improve 3D estimation using  
dense neural networks", Proceedings of the  
European Conference on Computer Vision (ECCV)  
Workshops, 2018(2018.09.08.) 1부.\*  
\*는 심사관에 의하여 인용된 문헌

(73) 특허권자  
포항공과대학교 산학협력단  
경상북도 포항시 남구 청암로 77 (지곡동)  
(72) 발명자  
이승용  
경상북도 포항시 남구 지곡로 155, 8동 803호  
조성현  
경상북도 포항시 남구 지곡로 278, 207동 102호  
이준용  
경상북도 포항시 북구 양덕로50번길 33, 101동  
1602호  
(74) 대리인  
특허법인이상

전체 청구항 수 : 총 14 항

심사관 : 김광식

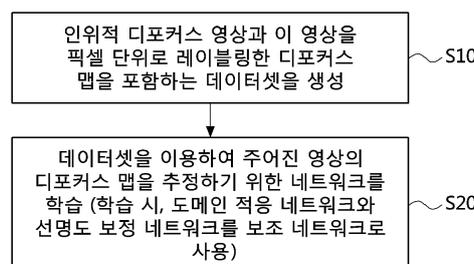
(54) 발명의 명칭 디포커스 맵 추정을 위한 기계 학습 방법 및 장치

(57) 요약

블러 추정 네트워크의 보조 네트워크로서 도메인 적응 네트워크와 선명도 보정 네트워크를 사용하여 자연적 디포커스 영상에 대해 강인하게 작동하는 디포커스 맵 추정을 위한 기계 학습 방법 및 장치가 개시된다. 디포커스 맵 추정을 위한 기계 학습 방법은 인위적 디포커스 영상과 인위적 디포커스 영상의 디포커스 정도를 픽셀 단위로 레

(뒷면에 계속)

대표도 - 도1



이블링한 디포커스 맵을 포함한 데이터셋을 생성하는 단계와, 데이터셋을 이용하여 주어진 영상의 디포커스 맵을 추정하기 위한 네트워크를 학습하는 단계를 포함하고, 학습하는 단계에서 블러 추정 네트워크가 주어진 입력 영상의 디포커스 맵을 출력할 때 도메인 적응 네트워크 및 선명도 보정 네트워크를 보조 네트워크로 사용하여 자연적 디포커스 영상에 대한 강인성을 증대시킨다.

(52) CPC특허분류

G06T 2207/20081 (2013.01)

G06T 2207/20084 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711081119
과제번호	2015-0-00174-005
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기술진흥센터
연구사업명	SW컴퓨팅산업원천기술개발
연구과제명	(SW 스타랩) 빅 비주얼 데이터 기반의 고품질 사진 메이크업 SW 개발
기 여 율	7/10
과제수행기관명	포항공과대학교 산학협력단
연구기간	2019.01.01 ~ 2019.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호	1711093801
과제번호	GK19P0300
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	범부처GigaKOREA사업
연구과제명	초실감 서비스를 위한 동적 객체의 실시간 4D 복원 기술 개발
기 여 율	3/10
과제수행기관명	광주과학기술원
연구기간	2019.01.01 ~ 2019.12.31

공지예외적용 : 있음

## 명세서

### 청구범위

#### 청구항 1

인위적 디포커스 영상과 상기 인위적 디포커스 영상의 디포커스 정도를 픽셀 단위로 레이블링한 디포커스 맵을 포함한 데이터셋을 생성하는 단계; 및

상기 데이터셋을 이용하여 주어진 영상의 디포커스 맵을 추정하기 위한 네트워크를 학습하는 단계;를 포함하고, 상기 학습하는 단계는 블러 추정 네트워크가 주어진 입력 영상의 디포커스 맵을 출력할 때 도메인 적응 네트워크 및 선명도 보정 네트워크를 보조 네트워크로 사용하여 자연적 디포커스 영상에 대한 강인성을 증대시키며, 상기 학습하는 단계는 모양 유지 네트워크를 사용하여 상기 블러 추정 네트워크에서 출력되는 디포커스 맵과 교차 영상의 디포커스 맵과의 L2 손실 함수에 의해 발생하는 번짐을 방지하는, 디포커스 맵 추정을 위한 기계 학습 방법.

#### 청구항 2

청구항 1에 있어서,

상기 학습하는 단계는 상기 도메인 적응 네트워크를 사용하여 상기 블러 추정 네트워크가 인위적 디포커스 영상과 자연적 디포커스 영상의 도메인 차이를 줄이도록 하는, 디포커스 맵 추정을 위한 기계 학습 방법.

#### 청구항 3

청구항 2에 있어서,

상기 도메인 적응 네트워크는 상기 블러 추정 네트워크의 중간 피쳐맵을 입력으로 사용하며, 상기 중간 피쳐맵은 인위적 영상에 대한 피쳐맵과 자연적 영상에 대한 피쳐맵을 포함하고, 상기 도메인 적응 네트워크는 인위적 영상에 대한 피쳐맵과 자연적 영상에 대한 피쳐맵을 한번씩 입력으로 받아 각 피쳐맵을 판별하도록 동작하며, 여기서 상기 블러 추정 네트워크는 상기 도메인 적응 네트워크의 인위적 판별에 의해 인위적 영상 도메인과 자연적 영상 도메인에 대해 무더지는 도메인 적응 분위기에서 학습하는, 디포커스 맵 추정을 위한 기계 학습 방법.

#### 청구항 4

청구항 2에 있어서,

상기 학습하는 단계는 상기 선명도 보정 네트워크를 사용하여 상기 블러 추정 네트워크가 상기 입력 영상의 블러 정도를 이진 학습 대비 더 정확히 학습하도록 동작하는, 디포커스 맵 추정을 위한 기계 학습 방법.

#### 청구항 5

청구항 4에 있어서,

상기 선명도 보정 네트워크는 1X1 컨볼루션 레이어로 이루어지는, 디포커스 맵 추정을 위한 기계 학습 방법.

#### 청구항 6

삭제

#### 청구항 7

청구항 1에 있어서,

상기 모양 유지 네트워크는 VGG 네트워크(VGGNet)를 포함하는, 디포커스 맵 추정을 위한 기계 학습 방법.

#### 청구항 8

청구항 1에 있어서,

상기 학습하는 단계는 상기 블러 추정 네트워크가 주어진 입력 영상의 디포커스 맵을 출력하는 과정에서 상기 블러 추정 네트워크의 디코더의 각 레벨의 마지막 레이어에 적용되는 보조 네트워크 또는 보조 모듈의 보조 손실함수를 더 사용하며, 상기 보조 손실함수는 상기 블러 추정 네트워크에서 출력되는 디포커스 맵과 교사 영상의 디포커스 맵과의 L2 손실 함수인, 디포커스 맵 추정을 위한 기계 학습 방법.

**청구항 9**

프로그램을 저장하는 메모리; 및

상기 메모리에 연결되어 상기 프로그램을 실행하는 프로세서;를 포함하되,

상기 프로세서는 상기 프로그램에 의해, 인위적 디포커스 영상과 상기 인위적 디포커스 영상의 디포커스 정도를 픽셀 단위로 레이블링한 디포커스 맵을 포함하는 데이터셋을 생성하고, 상기 데이터셋을 이용하여 주어진 영상의 디포커스 맵을 추정하기 위한 네트워크를 학습하며, 상기 학습하는 단계에서 블러 추정 네트워크가 주어진 입력 영상의 디포커스 맵을 출력할 때 도메인 적응 네트워크 및 선명도 보정 네트워크를 상기 블러 추정 네트워크의 보조 네트워크로 사용하여 자연적 디포커스 영상에 대한 강인성을 증대시키며,

상기 프로세서는 상기 블러 추정 네트워크의 출력단에 연결되는 모양 유지 네트워크를 통해 상기 블러 추정 네트워크가 입력 영상 대비 더 정확한 모양의 디포커스 맵을 출력하도록 학습하는, 디포커스 맵 추정을 위한 기계 학습 장치.

**청구항 10**

삭제

**청구항 11**

청구항 9 있어서,

상기 프로세서는 상기 블러 추정 네트워크가 주어진 입력 영상의 디포커스 맵을 출력하는 과정에서 상기 블러 추정 네트워크의 디코더의 각 레벨의 마지막 레이어에 적용되는 보조 네트워크 또는 보조 모듈의 보조 손실함수를 더 사용하며, 상기 보조 손실함수는 상기 블러 추정 네트워크의 디포커스 맵 출력과 교사 영상과의 L2 손실 함수인, 디포커스 맵 추정을 위한 기계 학습 장치.

**청구항 12**

인위적 디포커스 영상과 상기 인위적 디포커스 영상의 디포커스 정도를 픽셀 단위로 레이블링한 디포커스 맵을 포함한 데이터셋을 저장하는 메모리; 및

상기 메모리에 연결되고 상기 데이터셋을 이용하여 주어진 영상의 디포커스 맵을 추정하기 위한 네트워크를 학습하는 프로세서;를 포함하되,

상기 프로세서에는 상기 주어진 영상의 디포커스 맵을 추정하는 블러 추정 네트워크와, 상기 블러 추정 네트워크가 인위적 디포커스 영상과 자연적 디포커스 영상의 도메인 차이를 줄이도록 보조하는 도메인 적응 네트워크와, 상기 블러 추정 네트워크가 입력 영상의 블러 정도를 이전 학습 대비 더 정확히 학습하도록 보조하는 선명도 보정 네트워크가 탑재되고,

상기 도메인 적응 네트워크 및 상기 선명도 보정 네트워크는 상기 블러 추정 네트워크의 자연적 디포커스 영상에 대한 강인성을 증대시키고,

상기 프로세서에는 상기 블러 추정 네트워크에서 출력되는 디포커스 맵과 교사 영상의 디포커스 맵과의 L2 손실 함수에 의해 발생하는 번짐을 방지하는 모양 유지 네트워크가 더 탑재되는, 디포커스 맵 추정을 위한 기계 학습 장치.

**청구항 13**

청구항 12에 있어서,

상기 도메인 적응 네트워크는 상기 블러 추정 네트워크의 중간 피쳐맵을 입력으로 사용하며, 상기 중간 피쳐맵

은 인위적 영상에 대한 피쳐맵과 자연적 영상에 대한 피쳐맵을 포함하고, 상기 도메인 적응 네트워크는 인위적 영상에 대한 피쳐맵과 자연적 영상에 대한 피쳐맵을 한번씩 입력으로 받아 각 피쳐맵을 판별하도록 동작하며, 여기서 상기 블러 추정 네트워크는 상기 도메인 적응 네트워크의 인위적 판별에 의해 인위적 영상 도메인과 자연적 영상 도메인에 대해 무더지는 도메인 적응 분위기에서 학습하는, 디포커스 맵 추정을 위한 기계 학습 장치.

**청구항 14**

청구항 12에 있어서,  
상기 선명도 보정 네트워크는 1X1 컨볼루션 레이어로 이루어지는, 디포커스 맵 추정을 위한 기계 학습 장치.

**청구항 15**

삭제

**청구항 16**

청구항 12에 있어서,  
상기 모양 유지 네트워크는 VGG 네트워크(VGGNet)를 포함하는, 디포커스 맵 추정을 위한 기계 학습 장치.

**청구항 17**

청구항 12에 있어서,  
상기 프로세서는 상기 블러 추정 네트워크의 디코더의 각 레벨의 마지막 레이어에 보조 손실함수를 더 적용하며, 상기 보조 손실함수는 상기 블러 추정 네트워크에서 출력되는 디포커스 맵과 교사 영상의 디포커스 맵과의 L2 손실 함수인, 디포커스 맵 추정을 위한 기계 학습 장치.

**발명의 설명**

**기술 분야**

[0001] 본 발명은 디포커스 블러를 포함한 단일 영상의 디포커스 맵을 추정하는 기술에 관한 것으로, 보다 구체적으로는, 블러 추정 네트워크의 보조 네트워크로서 도메인 적응 네트워크와 선명도 보정 네트워크를 사용하여 자연적 디포커스 영상에 대해 강인하게 작동하는 디포커스 맵 추정을 위한 기계 학습 방법 및 장치에 관한 것이다.

**배경 기술**

[0002] 기존의 디포커스 맵(Defocus map) 추정 방법은 먼저 주어진 영상에서 엣지(edge) 즉, 경계 영역을 추정하고 추정된 엣지의 디포커스 블러 정도를 측정한다. 그리고 측정된 값을 보간하여 엣지 사이 영역 즉, 균일 단면의 디포커스 블러 정도를 측정한다.

[0003] 하지만, 전술한 기존 방법으로 생성된 디포커스 맵은 몇가지 문제를 가진다. 첫째, 디포커스 블러를 포함하는 단일 영상에서 엣지를 추정하는 것이 힘들어 최종 디포커스 맵 또한 정확하지 않을 수 있다. 둘째, 엣지가 잘 추정되었을지라도, 서로 다른 정도로 디포커스되어 있는 엣지의 블러 정도를 정확히 측정하기 어렵다. 셋째, 앞선 두 문제로 인해 오류를 가질 확률이 매우 높은 엣지의 디포커스 정도를 보간하므로 측정된 균일단면의 디포커스 정도는 정확하지 않을 확률이 매우 높다.

[0004] 이와 같이 영상의 디포커스 정도를 정확하게 측정할 수 있는 효율적인 방안이 요구되고 있다.

**발명의 내용**

**해결하려는 과제**

[0005] 본 발명은 전술한 종래 기술의 요구에 부응하기 위해 도출된 것으로, 본 발명의 목적은 엣지 기반의 문제를 해결하여 디포커스 정보를 정확하게 측정할 수 있는 디포커스 맵 추정을 위한 기계 학습 방법 및 장치를 제공하는 데 있다.

[0006] 본 발명의 다른 목적은 상기의 디포커스 맵 추정을 위한 기계 학습 방법을 컴퓨터에서 실행시키기 위한 프로그

램을 기록한 기록 매체를 제공하는데 있다.

**과제의 해결 수단**

- [0007] 상기 기술적 과제를 해결하기 위한 본 발명의 일 측면에 따른 디포커스 맵 추정을 위한 기계 학습 방법은 인위적 디포커스 영상과 인위적 디포커스 영상의 디포커스 정도를 픽셀 단위로 레이블링한 디포커스 맵을 포함한 데이터셋을 생성하는 단계; 및 데이터셋을 이용하여 주어진 영상의 디포커스 맵을 추정하기 위한 네트워크를 학습하는 단계;를 포함하고, 상기 학습하는 단계는 블러 추정 네트워크가 주어진 입력 영상의 디포커스 맵을 출력할 때 도메인 적응 네트워크 및 선명도 보정 네트워크를 보조 네트워크로 사용하여 자연적 디포커스 영상에 대한 강인성을 증대시킨다.
- [0008] 일실시예에서, 상기 학습하는 단계는 도메인 적응 네트워크를 사용하여 블러 추정 네트워크가 인위적 디포커스 영상과 자연적 디포커스 영상의 도메인 차이를 줄이도록 학습한다.
- [0009] 일실시예에서, 상기 도메인 적응 네트워크는 블러 추정 네트워크의 중간 피쳐맵을 입력으로 사용한다. 중간 피쳐맵은 인위적 영상에 대한 피쳐맵과 자연적 영상에 대한 피쳐맵을 포함한다. 도메인 적응 네트워크는 인위적 영상에 대한 피쳐맵과 자연적 영상에 대한 피쳐맵을 한번씩 입력으로 받아 각 피쳐맵을 판별하도록 동작한다. 여기서 블러 추정 네트워크는 도메인 적응 네트워크의 인위적 판별에 의해 인위적 영상 도메인과 자연적 영상 도메인에 대해 무너지는 도메인 적응 분위기에서 학습한다.
- [0010] 일실시예에서, 상기 학습하는 단계는 선명도 조정 네트워크를 사용하여 블러 추정 네트워크가 입력 영상의 블러 정도를 이전 학습 대비 더 정확히 학습하도록 동작한다.
- [0011] 일실시예에서, 상기 선명도 조정 네트워크는 1X1 컨볼루션 레이어 혹은 이와 동일하거나 유사한 기능을 수행하는 기능부나 이에 대응하는 구성부로 이루어질 수 있다.
- [0012] 일실시예에서, 상기 학습하는 단계는 모양 유지 네트워크를 사용하여 블러 추정 네트워크에서 출력되는 디포커스 맵과 교사 영상의 디포커스 맵과의 L2 손실 함수에 의해 발생하는 번짐을 방지한다.
- [0013] 일실시예에서, 상기 모양 유지 네트워크는 블러 추정 네트워크의 출력측에 연결되는 VGG 네트워크를 포함한다.
- [0014] 일실시예에서, 상기 학습하는 단계는 블러 추정 네트워크가 주어진 입력 영상의 디포커스 맵을 출력하는 과정에서 블러 추정 네트워크의 디코더의 각 레벨의 마지막 레이어에 적용되는 보조 네트워크 또는 보조 모듈의 보조 손실함수를 더 사용한다. 상기 보조 손실함수는 블러 추정 네트워크에서 출력되는 디포커스 맵과 교사 영상의 디포커스 맵과의 L2 손실 함수일 수 있다.
- [0015] 상기 기술적 과제를 해결하기 위한 본 발명의 다른 측면에 따른 디포커스 맵 추정을 위한 기계 학습 장치는, 프로그램을 저장하는 메모리; 및 메모리에 연결되어 상기 프로그램을 실행하는 프로세서;를 포함하되, 상기 프로세서는 프로그램에 의해, 인위적 디포커스 영상과 인위적 디포커스 영상의 디포커스 정도를 픽셀 단위로 레이블링한 디포커스 맵을 포함하는 데이터셋을 생성하고, 데이터셋을 이용하여 주어진 영상의 디포커스 맵을 추정하기 위한 네트워크를 학습한다. 그리고 프로세서는 학습하는 단계에서 블러 추정 네트워크가 주어진 입력 영상의 디포커스 맵을 출력할 때 도메인 적응 네트워크 및 선명도 보정 네트워크를 블러 추정 네트워크의 보조 네트워크로 사용하여 자연적 디포커스 영상에 대한 강인성을 증대시킨다.
- [0016] 일실시예에서, 상기 프로세서는 블러 추정 네트워크의 출력단에 연결되는 모양 유지 네트워크를 통해 블러 추정 네트워크가 입력 영상 대비 더 정확한 모양의 디포커스 맵을 출력하도록 학습한다.
- [0017] 일실시예에서, 상기 프로세서는 상기 블러 추정 네트워크가 주어진 입력 영상의 디포커스 맵을 출력하는 과정에서 상기 블러 추정 네트워크의 디코더의 각 레벨의 마지막 레이어에 적용되는 보조 네트워크 또는 보조 모듈의 보조 손실함수를 더 사용하며, 상기 보조 손실함수는 상기 블러 추정 네트워크의 디포커스 맵 출력과 교사 영상과의 L2 손실 함수일 수 있다.
- [0018] 상기 기술적 과제를 해결하기 위한 본 발명의 또 다른 측면에 따른 디포커스 맵 추정을 위한 기계 학습 장치는, 인위적 디포커스 영상과 상기 인위적 디포커스 영상의 디포커스 정도를 픽셀 단위로 레이블링한 디포커스 맵을 포함한 데이터셋을 저장하는 메모리; 및 메모리에 연결되고 데이터셋을 이용하여 주어진 영상의 디포커스 맵을 추정하기 위한 네트워크를 학습하는 프로세서;를 포함하되, 상기 프로세서에는 주어진 영상의 디포커스 맵을 추정하는 블러 추정 네트워크와, 블러 추정 네트워크가 인위적 디포커스 영상과 자연적 디포커스 영상의 도메인 차이를 줄이도록 보조하는 도메인 적응 네트워크와, 블러 추정 네트워크가 입력 영상의 블러 정도를 이전 학습

대비 더 정확히 학습하도록 보조하는 선명도 조정 네트워크가 탑재된다. 도메인 적응 네트워크 및 선명도 보정 네트워크는 블러 추정 네트워크의 자연적 디포커스 영상에 대한 강인성을 증대시킨다.

[0019] 일실시예에서, 상기 도메인 적응 네트워크는 블러 추정 네트워크의 중간 피쳐맵을 입력으로 사용한다. 중간 피쳐맵은 인위적 영상에 대한 피쳐맵과 자연적 영상에 대한 피쳐맵을 포함한다. 도메인 적응 네트워크는 인위적 영상에 대한 피쳐맵과 자연적 영상에 대한 피쳐맵을 한번씩 입력으로 받아 각 피쳐맵을 판별하도록 동작한다. 여기서 블러 추정 네트워크는 도메인 적응 네트워크의 인위적 판별에 의해 인위적 영상 도메인과 자연적 영상 도메인에 대해 무더지는 도메인 적응 분위기에서 학습한다. 상기 선명도 조정 네트워크는 1X1 컨볼루션 레이어로 이루어질 수 있다.

[0020] 일실시예에서, 상기 프로세서에는 블러 추정 네트워크에서 출력되는 디포커스 맵과 교사 영상의 디포커스 맵과의 L2 손실 함수에 의해 발생하는 번짐을 방지하는 모양 유지 네트워크가 더 탑재될 수 있다. 모양 유지 네트워크는 VGG 네트워크를 포함할 수 있다.

[0021] 일실시예에서, 상기 프로세서는 상기 블러 추정 네트워크의 디코더의 각 레벨의 마지막 레이어에 보조 손실함수를 더 적용할 수 있다. 보조 손실함수는 블러 추정 네트워크에서 출력되는 디포커스 맵과 교사 영상의 디포커스 맵과의 L2 손실 함수일 수 있다.

**발명의 효과**

[0022] 진술한 디포커스 맵 추정을 위한 기계 학습 방법 및 장치를 사용하는 경우에는 기존의 문제점 즉, 디포커스 블러를 포함하는 단일 영상에서 엷지를 추정하는 것이 힘들어 최종 디포커스 맵이 정확하지 않는 문제, 엷지가 잘 추정되었을지라도 서로 다른 정도로 디포커스되어 있는 엷지의 블러 정도를 정확히 측정하기 어려운 문제, 앞의 두 문제로 인해 오류를 가질 확률이 매우 높은 엷지의 디포커스 정도를 보간하므로 측정된 균일단면의 디포커스 정도가 정확하지 않을 확률이 매우 높은 문제를 해결하여 디포커스 정보가 정확한 디포커스 맵을 추정할 수 있다.

[0023] 또한, 본 발명에 의하면, 인위적인 디포커스 영상과 이 영상의 디포커스 정도를 픽셀 단위로 레이블링한 디포커스 맵을 포함하는 데이터셋을 이용하여 학습하고, 학습 시, 도메인 적응 네트워크와 선명도 보정 네트워크를 보조 네트워크로 사용함으로써 자연적 디포커스 영상에 대해 강인하게 작동하는 디포커스 맵 추적을 위한 기계 학습 방법 및 장치를 제공할 수 있다.

**도면의 간단한 설명**

- [0024] 도 1은 본 발명의 일실시예에 따른 디포커스 맵 추정을 위한 기계 학습 방법에 대한 흐름도이다.
- 도 2는 도 1의 디포커스 맵 추정 방법에 채용할 수 있는 데이터셋 생성의 주요 과정을 설명하기 위한 도면이다.
- 도 3은 도 2의 데이터셋 생성 과정을 통해 얻어지는 데이터셋 영상을 예시한 도면이다.
- 도 4는 도 1의 디포커스 맵 추정 방법에 채용할 수 있는 네트워크 학습 과정을 설명하기 위한 도면이다.
- 도 5는 도 4의 네트워크 학습 과정에서 보조 네트워크로 사용되는 선명도 보정 네트워크의 데이터셋을 예시하는 도면이다.
- 도 6은 도 4의 네트워크 학습 과정에서 모듈별 효과 결과를 예시한 도면이다.
- 도 7은 본 실시예에 따른 디포커스 맵 추정을 위한 기계 학습 방법의 디포커스 맵 추정 결과와 기존 방법들의 추정 결과를 비교하여 나타낸 예시도이다.
- 도 8은 본 발명의 다른 실시예에 따른 디포커스 맵 추정을 위한 기계 학습 장치에 대한 개략적인 블록도이다.

**발명을 실시하기 위한 구체적인 내용**

[0025] 본 발명은 다양한 변경을 가할 수 있고 여러 가지 실시예를 가질 수 있는 바, 특정 실시예들을 도면에 예시하고 상세하게 설명하고자 한다. 그러나, 이는 본 발명을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다.

[0026] 제1, 제2 등의 용어는 다양한 구성요소들을 설명하는데 사용될 수 있지만, 상기 구성요소들은 상기 용어들에 의해 한정되어서는 안 된다. 상기 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 사용된

다. 예를 들어, 본 발명의 권리 범위를 벗어나지 않으면서 제1 구성요소는 제2 구성요소로 명명될 수 있고, 유사하게 제2 구성요소도 제1 구성요소로 명명될 수 있다. 및/또는 이라는 용어는 복수의 관련된 기재된 항목들의 조합 또는 복수의 관련된 기재된 항목들 중의 어느 항목을 포함한다.

[0027] 어떤 구성요소가 다른 구성요소에 "연결되어" 있다거나 "접속되어" 있다고 언급된 때에는, 그 다른 구성요소에 직접적으로 연결되어 있거나 또는 접속되어 있을 수도 있지만, 중간에 다른 구성요소가 존재할 수도 있다고 이해되어야 할 것이다. 반면에, 어떤 구성요소가 다른 구성요소에 "직접 연결되어" 있다거나 "직접 접속되어" 있다고 언급된 때에는, 중간에 다른 구성요소가 존재하지 않는 것으로 이해되어야 할 것이다.

[0028] 본 출원에서 사용한 용어는 단지 특정한 실시예를 설명하기 위해 사용된 것으로, 본 발명을 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 출원에서, "포함하다" 또는 "가지다" 등의 용어는 명세서상에 기재된 특징, 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.

[0029] 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가지고 있다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥 상 가지는 의미와 일치하는 의미를 가진 것으로 해석되어야 하며, 본 출원에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.

[0030] 이하, 첨부한 도면들을 참조하여, 본 발명의 바람직한 실시예를 보다 상세하게 설명하고자 한다. 본 발명을 설명함에 있어 전체적인 이해를 용이하게 하기 위하여 도면상의 동일한 구성요소에 대해서는 동일한 참조부호를 사용하고 동일한 구성요소에 대해서 중복된 설명은 생략한다.

[0031] 도 1은 본 발명의 실시예에 따른 디포커스 맵 추정을 위한 기계 학습 방법에 대한 흐름도이다.

[0032] 도 1을 참조하면, 본 실시예에 따른 디포커스 맵 추정을 위한 기계 학습 방법(이하 간략히 '디포커스 맵 추정 방법'이라 한다)은 인위적 디포커스 영상과 이 영상을 픽셀 단위로 레이블링한 디포커스 맵(defocus map)을 포함하는 데이터셋을 생성하는 단계(S10)를 포함한다.

[0033] 또한, 디포커스 맵 추정 방법은 앞서 생성한 데이터셋을 이용하여 주어진 영상의 디포커스 맵을 추정하기 위한 네트워크를 학습하는 단계(S20)를 포함한다.

[0034] 상기의 학습하는 단계(S20)에서는 블러 추정 네트워크에서 디포커스 맵을 추정할 때 보조 네트워크로서 도메인 적응 네트워크와 선명도 보정 네트워크를 사용하며, 그에 의해 블러 추정 네트워크가 자연적 디포커스 영상에 대해 강인하게 동작하도록 이루어진다.

[0035] 도 2는 도 1의 디포커스 맵 추정 방법에 채용할 수 있는 데이터셋 생성의 주요 과정을 설명하기 위한 도면이다. 도 3은 도 2의 데이터셋 생성 과정을 통해 얻어지는 데이터셋 영상을 예시한 도면이다.

[0036] 본 실시예에 따른 디포커스 맵 추정 방법에서는 컨볼루션 신경망(Convolutional Neural Network, CNN)을 지도 학습하여 단일 영상의 디포커스 맵을 추정하도록 한다. 이때 네트워크의 입력은 인위적으로 생성한 디포커스 영상을 포함하고, 네트워크를 학습하기 위한 교사 영상 혹은 레이블(label) 영상은 인위적으로 생성한 디포커스 영상의 디포커스 맵을 사용한다.

[0037] 본 실시예에서의 데이터셋 생성 방법의 주요 과정을 예시하면 다음과 같다.

[0038] 인위적 디포커스 영상을 만들기 위해서는 입력 영상과 그 영상의 깊이 맵이 필요하다. 이 두 정보를 이용하여 디포커스 영상과 이에 해당하는 디포커스 맵을 구한다. 인위적 디포커스 영상을 최대한 자연스럽게 만들기 위해서 본 실시예에서는 광학 모델의 일종인 얇은 렌즈 모델(Thin Lens Model)을 사용한다.

[0039] [수학식 1]

$$c(x) = \alpha \frac{|x - S_1|}{x}, \text{ where } \alpha = \frac{f_1}{S_1} D$$

[0040]

[0041] 위의 수학식 1과 도 2에 도시한 바와 같이, S<sub>1</sub>은 초점이 맞는 깊이로서 객체-공간 초점 거리[mm], D는 조리개

값(aperture diameter), F는 초점 거리(focal length), 그리고  $f_1$ 은 카메라 렌즈 중심에서 센서까지의 거리 즉, 이미지-공간 초점 거리(image-space focal distance)를 나타낸다.

[0042] 즉, 상기의 변수들을 이용해 얇은 렌즈 모델을 사용하면 실제 삼차원(3D) 공간에서의 디포커스 정도인  $C(x)$ 와 이차원(2D) 이미지 공간에서의 디포커스 정도인  $c(x)$ 를 구할 수 있다.

[0043] 본 실시예에서는 주어진 깊이맵이 있는 영상에 대해  $S_1, D, F, f_1$ 을 랜덤으로 생성하여 영상의 각 픽셀에 대해  $c(x)$  크기를 구하고, 이 크기 만큼 가우시안 블러를 적용하여 인위적인 디포커스 영상에 대한 디포커스 맵을 생성하고 이것들을 포함하는 데이터셋을 생성한다.

[0044] 예를 들어, 본 실시예에 따른 기계 학습 방법의 적어도 일부의 단계들을 실행하는 컴퓨팅 장치는 얇은 렌즈 모델을 통해 디포커스 블러를 가진 영상을 생성하기 위해 먼저 영상의 깊이 맵으로부터 최소 깊이 경계(minimum depth bound,  $x_{near}$ )와 최대 깊이 경계(maximum depth bound,  $x_{far}$ )를 추출할 수 있다. 그런 다음,  $[x_{near}, x_{far}]$ 의 범위에서 객체-공간 초점 거리( $S_1$ )를 랜덤하게 샘플링할 수 있다. 그리고, 레이블링된 디포커스 블러의 양을 소정의 최대치까지 제한함으로써 일정 범위에서 랜덤하게 샘플링되는 디포커스 맵의 특정한 물리적 상수를 상한(upper bound)를 결정할 수 있다.

[0045] 데이터셋은 예를 들어 도 3에 도시한 바와 같이 입력 영상(a)에 대한 인위적 디포커스 영상(defocused image, b)과 디포커스 영상의 디포커스 정도를 픽셀 단위로 레이블링한 디포커스 맵(c)의 쌍을 포함하는 수천 세트 이상(예를 들어 총 8,231장)의 디포커스 영상과 디포커스 맵의 쌍을 포함할 수 있다.

[0046] 도 4는 도 1의 디포커스 맵 추정 방법에 채용할 수 있는 네트워크 학습 과정을 설명하기 위한 도면이다. 도 5는 도 4의 네트워크 학습 과정에서 보조 네트워크로 사용되는 선명도 보정 네트워크의 데이터셋을 예시하는 도면이다.

[0047] 도 4를 참조하면, 본 실시예에 따른 디포커스 맵 추정 방법은 블러 추정 네트워크(Blur Estimation Network, B), 도메인 적응 네트워크(Domain adaptation Network, D), 선명도 조정 네트워크(Sharpness Calibration Network, S) 및 모양 유지 네트워크(Contents Preservation Network, C)의 4가지 모듈로 구성되는 신경망에서 사용될 수 있다.

[0048] 블러 추정 네트워크(B)는 메인 네트워크로써 주어진 입력 영상의 디포커스 맵을 출력한다.

[0049] 도메인 적응 네트워크(D)는 블러 추정 네트워크(B)로 하여금 인위적 디포커스 영상과 자연적 디포커스 영상의 도메인 차이를 줄이게 한다.

[0050] 선명도 조정 네트워크(S)는 블러 추정 네트워크(B)로 하여금 입력 영상의 블러 정도를 조금 더 정확히 학습할 수 있도록 지원한다.

[0051] 모양 유지 네트워크(C)는 블러 추정 네트워크(B)로 하여금 입력 영상과 비교하여 조금 더 정확한 모양의 디포커스 맵을 출력하도록 한다.

[0052] 네트워크 학습 시에는 상기의 4개의 모듈 전체를 사용하며, 테스트 시에는 블러 추정 네트워크(B)만을 사용한다.

[0054] 각 네트워크의 학습 과정을 좀더 구체적으로 살펴보면 다음과 같다.

[0055] 블러 추정 네트워크(B) 학습

[0056] 블러 추정 네트워크(B)에 인위적 디포커스 영상을 입력으로 주고 네트워크가 출력한 디포커스 추정맵과, 디포커스 교사맵(ground-truth 디포커스 맵)의 차이를 0으로 만드는 손실 함수(도 4의 blur estimation loss,  $L_B$ )를 사용한다. 본 실시예에서는 L2(픽셀 값의 차이의 제곱의 평균, Mean Squared Error라고도 부름) 함수를 사용하여 손실 함수를 계산할 수 있다. 이러한 손실 함수( $L_B$ )를 식으로 나타내는 수학적 식 2와 같다.

[0057] [수학적 식 2]

$$L_B = \frac{1}{WH} \sum_{i=1}^W \sum_{j=1}^H (\mathbf{B}(I_S; \theta_B)_{i,j} - y_{i,j})^2 + \lambda_{aux} L_{aux}$$

[0058]

[0059] [수학식 2]에서  $W, H$  는 이미지의 길이와 높이를 각각 의미한다.

[0060] 즉, [수학식 2]에서 손실 함수( $L_B$ )는 인위적 디포커스 영상( $I_S$ )에 대해 가중치( $\theta_B$ )를 가지는 블러 추정 네트워크(B)의 디포커스 추정맵의  $(i, j)$ 번째 픽셀들과 디포커스 교사맵  $y$ 의  $(i, j)$ 번째 픽셀들의 차이의 평균을 구한다. 블러 추정 네트워크(B)는 상기의 차이를 최소화하도록 학습된다.  $\lambda_{aux}$ 는 ...를 의미한다.

[0061] 또한, [수학식 2]에서 블러 추정 네트워크(B)는 디코더의 각 레벨의 마지막 레이어에 보조(auxiliary) 네트워크를 적용하여 디포커스 맵을 출력한 뒤, 그 출력과 교사 영상과의 L2 손실 함수( $L_{aux}$ )를 적용한다(도 4의 auxiliary loss 참조). L2 손실 함수( $L_{aux}$ )를 식으로 나타내면 다음의 [수학식 3]과 같다.

[0062] [수학식 3]

$$L_{aux} = \sum_{\ell=1}^{L_B} \frac{1}{W_\ell H_\ell} \sum_{i=1}^{W_\ell} \sum_{j=1}^{H_\ell} (\mathbf{B}_\ell(I_S; \theta_B, \theta_{aux})_{i,j} - y_{\ell,i,j})^2,$$

[0063] [수학식 3]에서  $W_\ell, H_\ell$ 은 디코더 각 레벨의 디포커스 맵의 길이와 높이를 각각 의미하며,  $\theta_{aux}$ 는 보조 네트워크( $B_1$ )의 가중치를 의미한다.

[0066] 도메인 적응 네트워크(D) 학습

[0067] 블러 추정 네트워크(B)에 중간 피쳐맵을 도메인 적응 네트워크(D)의 입력으로 사용한다. 이때 피쳐맵은 2가지 종류로 나뉘어지는데, 이는 인위적 영상에 대한 피쳐맵(주황색 선 또는 점선)과 자연적 영상에 대한 피쳐맵(파란색 선 또는 실선)이다. 도메인 적응 네트워크(D)는 두 피쳐맵을 한번씩 입력으로 받아 각 피쳐맵을 잘 판별하도록 학습한다.

[0068] 다시 말해서, 블러 추정 네트워크(B)는 기본적으로 자연적 디포커스 영상을 입력을 받았을 때 생성한 피쳐맵을 도메인 적응 네트워크(D)가 인위적 피쳐맵이라 판별하도록 학습한다. 한편, 도메인 적응 네트워크(D)가 인위적 피쳐맵과 자연적 피쳐맵의 판별을 제대로 학습시키도록 하는 중에 블러 추정 네트워크(B)는 인위적 영상에 대한 피쳐맵을 도메인 적응 네트워크(D)가 인위적으로 판단하게 되고, 자연적 영상에 대한 피쳐맵도 도메인 적응 네트워크(D)가 인위적으로 판단하도록 학습된다. 따라서, 블러 추정 네트워크(B)는 인위적 자연적 영상 도메인에 대해 무더지게 된다. 본 실시예에서는 이러한 현상을 도메인 적응(Domain Apaptation)이라 한다.

[0069] 도메인 적응 네트워크(D)를 학습하기 위해 사용한 손실함수( $L_D$ )를 식으로 나타내면 다음의 [수학식 4]와 같다.

[0070] [수학식 4]

$$L_D = (z - 1) \cdot \log(1 - \mathbf{D}(\mathbf{B}_{last}(I; \theta_B), \theta_D)) - z \cdot \log(\mathbf{D}(\mathbf{B}_{last}(I; \theta_B), \theta_D)),$$

[0071] 여기서 블러 추정 네트워크(B)를 학습하기 위해 사용한 손실함수( $L_{adv}$ )를 수학식으로 나타내면 다음과 같다.

[0073] [수학식 5]

$$L_{adv} = -\log(\mathbf{D}(\mathbf{B}_{last}(I_R; \theta_B), \theta_D)),$$

[0075] 위의 [수학식 5]에서  $\theta_D$ 는 도메인 적응 네트워크(D)의 학습 가중치이다.

[0077] 선명도 조정 네트워크(S) 학습

[0078] 선명도 조정 네트워크(S)는 자연적 디포커스 영상에 대한 블러 추정 네트워크(B)의 출력을 입력으로 받는다. 선명도 조정 네트워크(S)는  $1 \times 1$  컨볼루션 레이어로 구성되어 있는데, 이는 학습 시 선명도 조정 네트워크(S)로부터 발생된 그래디언트가 블러 추정 네트워크(B)에서 출력되는 디포커스 맵의 번짐 현상을 방지하도록 하기 위함이다. 본 실시예의 선명도 조정 네트워크(S)에서는 다른 네트워크들과 다르게 CUHK(Chinese University of Hong Kong)의 CUHK 데이터셋에서 제공하는 블러맵을 교사 영상으로 사용할 수 있다.

[0079] CUHK 블러맵은 도 5의 (a) 및 (b)로 도시된 바와 같이 이진 영상으로 선명한 영역은 0, 선명하지 않은 영역은 1로 표기되어 있는데, 이를 교사 영상으로 사용하면, 블러 추정 네트워크(B)가 선명한 영역만큼은 확실히 배우게

되어 즉, 선명도 조정 효과가 보다 더 정확하게 되어 정확한 디포커스 맵을 추정할 수 있게 된다. 선명도 조정 네트워크(S)에서 사용한 판별 손실함수(classification loss)는 다음의 [수학식 6]과 같다.

[0080] [수학식 6]

$$L_S = \frac{1}{WH} \sum_{i=1}^W \sum_{j=1}^H \left\| \frac{1}{1 + \exp(-\mathbf{S}(\mathbf{B}(I_B; \theta_B); \theta_S)_{i,j})} - \mathbf{b}_{i,j} \right\|_2^2,$$

[0082] 위의 [수학식 6]은 각 픽셀에 대한 판별 손실함수이다. [수학식 6]에서  $\theta_S$ 는 선명도 조정 네트워크(S)의 학습 가중치이다.

[0084] 모양 유지 네트워크(C) 학습

[0085] 모양 유지 네트워크(C)는 블러 추정 네트워크를 학습할 때 사용한 L2 손실 함수에 의해 발생하는 번짐 현상을 방지하는 역할을 한다. 본 실시예에서 모양 유지 네트워크(C)는 ImageNet의 데이터셋에 대하여 미리 학습된 VGG19 네트워크를 사용하였으나, 이에 한정되지는 아니한다. 모양 유지 네트워크(C)에서 사용한 손실 함수를 나타내면 다음의 [수학식 7]과 같다.

[0086] [수학식 7]

$$L_C = \frac{1}{W_\ell H_\ell} \sum_{i=1}^{W_\ell} \sum_{j=1}^{H_\ell} (\phi_\ell(\mathbf{B}(I_S; \theta_B))_{i,j} - \phi_\ell(y)_{i,j})^2,$$

[0088] 위의 [수학식 7]에서  $\phi_\ell$ 는 VGG19 네트워크를 의미하며,  $\ell$ 은 사용한 VGG 네트워크의 레이어 번호를 의미한다.

[0089] 위에서 4개의 모듈(B, D, S, C)을 구체적으로 살핀 바와 같이, 본 실시예에 따른 디포커스 맵 추정을 위한 기계 학습 방법은 인위적 디포커스 영상과 이 영상의 디포커스 정도를 픽셀 단위로 레이블링한 디포커스 맵의 쌍을 포함하는 데이터셋을 사용하여 주어진 영상의 디포커스 맵을 추정하도록 학습하며, 학습 과정에서 블러 추정 네트워크(B)가 도메인 적응 네트워크(D)의 동작에 의해 인위적 영상에 대한 피쳐맵을 인위적으로 판단하고 자연적 영상에 대한 피쳐맵을 인위적으로 판단하도록 하여 도메인 적응 성능을 갖도록 하며, 선명도 조정 네트워크(S)의 동작에 의해 블러 추정 네트워크(B)에서 출력되는 디포커스 맵의 번짐 현상을 방지하고, 모양 유지 네트워크(C)의 동작을 통해 블러 추정 네트워크(B)의 디포커스 맵의 출력과 교사 영상과의 L2 손실 함수에 의해 발생하는 번짐 현상을 방지함으로써 정확한 디포커스 정도를 가진 디포커스 맵을 추정할 수 있다.

[0090] 도 6은 도 4의 네트워크 학습 과정에서 모듈별 효과 결과를 예시한 도면이다.

[0091] 도 6을 참조하면, 본 실시예에 따른 디포커스 맵 추정을 위한 기계 학습 방법에서 인위적 입력 결과(상단 행)와 자연적 입력 결과(하단 행)를 확인할 수 있다.

[0092] 도 6에서 왼쪽의 (a)는 입력영상 또는 입력(Input)으로, 입력에 대하여 (b)는 블러 추정 네트워크(B)만을 사용한 경우, (c)는 블러 추정 네트워크(B)와 도메인 적응 네트워크(D)를 사용한 경우, (d)는 블러 추정 네트워크(B)와 도메인 적응 네트워크(D)와 모양 유지 네트워크(C)를 사용한 경우, (e)는 블러 추정 네트워크(B)와 도메인 적응 네트워크(D)와 선명도 조정 네트워크(S)와 모양 유지 네트워크(C)를 사용하고 보조 손실함수를 사용하지 않은 경우, (f)는 블러 추정 네트워크(B)와 도메인 적응 네트워크(D)와 선명도 조정 네트워크(S)와 모양 유지 네트워크(C)를 사용하고 보조 손실함수를 사용한 경우의 결과를 각각 나타낸다.

[0093] 도 6에서 볼 수 있듯이, 블러 추정 네트워크(B)만 사용한 경우(b), 인위적 입력에 대한 디포커스 맵의 결과는 나쁘지 않지만 도메인 차이 때문에 자연적 입력에 대한 결과가 좋지 않다. 그리고 추가로 도메인 적응 네트워크(D)를 사용한 경우(c), 도메인이 적응돼 자연적 입력에 대해서도 네트워크가 작동하려 함을 볼 수 있다.

[0094] 여기에 모양 유지 네트워크(C)를 붙이면 즉, (d)의 경우 디포커스 맵이 이전 결과에 비해 노이즈가 줄어들고 한층 더 부드러워짐을 확인할 수 있다. 또한 추가로 선명도 조정 네트워크(S)를 붙이면 즉, (e)의 경우 디포커스 블러 추정 정확도가 더욱 높아짐을 확인할 수 있다. 아울러, (f)의 경우와 같이, 보조 손실함수(auxiliary loss)를 추가로 사용하는 경우, 보조 손실함수를 사용하지 않는 경우에 비해 디포커스 블러 추정 결과가 더욱 좋음을 알 수 있다.

[0095] 도 7은 본 실시예에 따른 디포커스 맵 추정을 위한 기계 학습 방법의 디포커스 맵 추정 결과와 기존 방법들의

추정 결과를 비교하여 나타낸 예시도이다.

- [0096] 도 7을 참조하면, (a) 4개의 입력(Input)과 각 입력에 대하여 비교예 1(S. Zhuo and T. Sim. Defocus map estimation from a single image. Pattern Recognition, 44(9): 1852-1858, 2011.)에 의해 추정된 블러맵(b), 비교예 2(Shi, L. Xu, and J. Jia. Just noticeable defocus blur detection and estimation. In Proc. CVPR, 2015.)에 의해 추정된 블러맵(c), 비교예 3(J. Park, Y. Tai, D. Cho, and I. S. Kweon. A unified approach of multi-scale deep and hand-crafted features for defocus estimation. In Proc. CVPR, 2017.)에 의해 추정된 블러맵(d), 비교예 4(A. Karaali and C. Jung. Edge-based defocus blur estimation with adaptive scale selection. IEEE Trans. Image Processing (TIP), 27(3):1126-1137, 2018.)에 의해 추정된 블러맵(e), 비교예 5(A. Karaali and C. Jung. Edge-based defocus blur estimation with adaptive scale selection. IEEE Trans. Image Processing(TIP), 27(3):1126-1137, 2018.)에 의해 추정된 블러맵(f)과 비교할 때, 본 실시예에 의한 블러맵(g)의 디포커스 정도가 더욱 명확한 것을 확인할 수 있다. 여기서, (g)는 CUHK 데이터셋의 블러맵이다. 용어 블러맵은 디포커스 맵에 포함될 수 있다.
- [0097] 이와 같이, 본 실시예에 의한 블러맵은 비교예들의 다른 블러맵과 비교하여 볼 때, 예를 들어 도 8의 (a) 내지 (f)의 맨 위쪽의 1행에서와 같이 손등위의 블러가 더 부드럽게 표현될 수 있고 하트 모양을 완벽히 잡을 수 있으며, 2행에서와 같이 사람 테두리가 다른 비교예의 결과와 비교해서 완벽히 보존되어 있으며, 3행에서와 같이 박스 위의 마크가 표시되지 않도록 박스의 경계를 정확하게 디포커스 할 수 있고, 4행에서와 같이 하늘을 잘 표시할 수 있는 등 다른 비교예와 대비할 때, 월등히 디포커스 정도의 정확도에서 매우 우수한 성능을 나타내고 있음을 알 수 있다.
- [0098] 도 8은 본 발명의 다른 실시예에 따른 디포커스 맵 추정을 위한 기계 학습 장치에 대한 개략적인 블록도이다.
- [0099] 도 4와 함께 도 8을 참조하면, 본 실시예에 따른 디포커스 맵 추정을 위한 기계 학습 장치(이하 간략히 '디포커스 맵 추정 장치'라고 한다)는 메모리 및 메모리에 연결되는 프로세서를 포함한다. 메모리는 인위적 디포커스 영상과 인위적 디포커스 영상의 디포커스 정도를 픽셀 단위로 레이블링한 디포커스 맵을 포함한 데이터셋을 저장한다. 프로세서는 데이터셋을 이용하여 주어진 영상의 디포커스 맵을 추정하기 위한 네트워크를 학습하거나 주어진 영상의 디포커스 맵을 추정하도록 트레이닝 위상(training phase)이나 테스트 위상(test phase)에서 동작한다.
- [0100] 프로세서에는 주어진 영상의 디포커스 맵을 추정하는 블러 추정 네트워크(B)와, 블러 추정 네트워크(B)가 인위적 디포커스 영상과 자연적 디포커스 영상의 도메인 차이를 줄이도록 보조하는 도메인 적응 네트워크(D)와, 블러 추정 네트워크(B)가 입력 영상의 블러(Blur) 정도를 이전 학습 대비 더 정확히 학습하도록 보조하는 선명도 조정 네트워크(S)가 탑재된다. 여기서 도메인 적응 네트워크(D) 및 선명도 보정 네트워크(S)는 블러 추정 네트워크(B)의 자연적 디포커스 영상에 대한 강인성을 증대시킨다.
- [0101] 또한, 프로세서에는 블러 추정 네트워크(B)에서 출력되는 디포커스 맵과 교사 영상의 디포커스 맵과의 L2 손실 함수에 의해 발생하는 번짐을 방지하는 모양 유지 네트워크(Content Preservation Network, C)가 더 탑재될 수 있다.
- [0102] 각 구성요소를 좀더 구체적으로 살펴보면, 블러 추정 네트워크(B)는 사전훈련된(Pretrained) VGG19 모델(B1), U-net 디코더(Decoder)(B2) 및 잔차 컨볼루션(Residual convolutions, B3)을 구비하며, 사전훈련된 VGG19 모델(B1)과 U-net 디코더(B2) 사이에 스킵 커넥션(Skip connections)을 구비한다.
- [0103] 여기서 프로세서는 블러 추정 네트워크(B)에서 출력되는 예측 임의적 디포커스 맵(Predicted synthetic defocus map,  $y_2$ )과 진짜 타겟(ground truth, GT) 임의적 디포커스 맵(synthetic defocus map,  $y_1$ )과의 차이에 기초하는 블러 추정 손실함수(blur estimation loss, LB)를 계산할 수 있다.
- [0104] 도메인 적응 네트워크(D)는 U-net 디코더(B2)와 잔차 컨볼루션(B3) 사이의 피쳐맵 즉, 블러 추정 네트워크의 중간 피쳐맵을 입력으로 사용한다. 중간 피쳐맵은 인위적 영상에 대한 피쳐맵(점선)과 자연적 영상에 대한 피쳐맵(실선)을 포함한다. 도메인 적응 네트워크(D)는 인위적 영상에 대한 피쳐맵과 자연적 영상에 대한 피쳐맵을 한 번씩 입력으로 받아 각 피쳐맵을 판별하도록 동작한다.
- [0105] 도메인 적응 네트워크(D)는 판별기(Discriminator, D1)를 통해 인위적(Synthetic) 피쳐맵인지 자연적(Real) 피쳐맵인지를 판별하고, 도메인 적응 손실함수(domain adapatation loss)를 생성한다. 도메인 적응 손실함수는 도메인 적응 네트워크(D)를 학습하기 위해 사용한 손실함수( $L_D$ )와 블러 추정 네트워크(B)를 학습하기 위해 사용한

손실함수( $L_{adv}$ )를 포함한다.

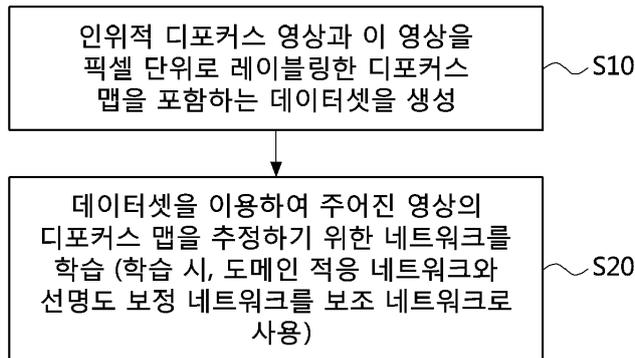
- [0106] 이러한 구성에 의하면, 블러 추정 네트워크(B)는 도메인 적응 네트워크(D)의 두 피쳐맵들에 대한 인위적 판별에 의해 인위적 영상 도메인과 자연적 영상 도메인에 대해 무더지는 도메인 적응(Domain Adaptation) 분위기에서 학습할 수 있다.
- [0107] 선명도 조정 네트워크(S)는 자연적 디포커스 영상에 대한 블러 추정 네트워크(B)의 출력 즉, 예측 자연적 디포커스 맵(Predicted real defocus map,  $y_3$ )을 입력으로 받는다. 선명도 조정 네트워크(S)는 1X1 컨볼루션 레이어(1X1 CNN, S1) 혹은 이와 동일하거나 유사한 기능을 수행하는 기능부나 이러한 기능부에 대응하는 구성부로 이루어질 수 있다.
- [0108] 선명도 조정 네트워크(S)는 1X1 컨볼루션 레이어(S1)를 통해 예측한 자연적 블러맵(Predicted real blur map,  $b_1$ )과 GT 자연적 블러맵(real blur map,  $b_2$ )과의 차이에 기초하여 판별 손실함수(classification loss)를 생성한다. 판별 손실함수는 선명도 조정 손실함수(sharpness calibration loss,  $L_s$ )로 지칭될 수 있다.
- [0109] 선명도 조정 네트워크(S)의 선명도 조정 효과를 위해 블러 추정 네트워크(B)의 교사 영상으로는 CUHK 블러맵이나 이와 유사한 구성의 블러맵을 사용할 수 있다. CUHK 블러맵은 이진영상으로 선명한 영역은 0으로 표기되고, 선명하지 않은 영역은 1로 표기되어 있으므로 블러 추정 네트워크(B)가 블러맵의 선명한 영역만큼은 확실하게 학습할 수 있어 블러 추정 네트워크(b)가 더욱 정확한 디포커스 맵을 추정하도록 사용될 수 있다.
- [0110] 모양 유지 네트워크(C)는 블러 추정 네트워크(B)를 학습시킬 때 사용한 L2 손실 함수에 의해 발생하는 번짐 현상을 방지하기 위해 블러 추정 네트워크(B)의 출력측에 연결된다. 이러한 모양 유지 네트워크(C)는 ImageNet 데이터셋에 대하여 사전학습된 일정 사이즈(fixed)의 VGG19(C1) 등의 VGG 네트워크(VGGNet)를 포함할 수 있다.
- [0111] 모양 유지 네트워크(C)는 블러 추정 네트워크(B)에서 출력되는 예측 임의적 디포커스 맵( $y_2$ )과 GT 임의적 디포커스 맵( $y_1$ )을 입력으로 받고, 예측 액티베이션(predicted activation,  $a_1$ )을 출력하고, 예측 액티베이션( $a_1$ )과 GT 액티베이션( $a_2$ )과의 디포커싱 정도의 차이에 기초하는 모양 유지 손실함수(content preservation loss,  $L_c$ )를 계산할 수 있다.
- [0112] 전술한 손실함수들 각각의 손실값이나 손실점수는 반복 학습을 위해 액티베이터 등의 최적화 모듈을 통해 블러 추정 네트워크의 입력측과 출력측 사이의 은닉층들 각각의 가중치를 업데이트시키는데 이용될 수 있다.
- [0113] 또한, 도 8에 도시한 바와 같이, 프로세서는 블러 추정 네트워크(B)의 디코더(B2)의 각 레벨의 마지막 레이어에 보조 손실함수(Auxiliary loss,  $L_{aux}$ )를 추가 적용하기 위해 또다른 보조 네트워크로서 보조 모듈(auxiliary module, Aux)을 더 탑재할 수 있다. 보조 손실함수( $L_{aux}$ )는 블러 추정 네트워크(B)에서 출력되는 디포커스 맵과 교사 영상의 디포커스 맵과의 L2 손실 함수일 수 있다.
- [0114] 위에서 살핀 바와 같이, 본 실시예에 따른 디포커스 맵 추정을 위한 기계 학습 장치는 인위적 디포커스 영상과 이 영상의 디포커스 정도를 픽셀 단위로 레이블링한 디포커스 맵의 쌍을 포함하는 데이터셋을 사용하여 주어진 영상의 디포커스 맵을 추정하도록 학습하며, 학습 과정에서 블러 추정 네트워크가 도메인 적응 네트워크의 동작에 의해 인위적 영상에 대한 피쳐맵을 인위적으로 판단하고 자연적 영상에 대한 피쳐맵을 인위적으로 판단하도록 하여 도메인 적응 분위기에서 학습하도록 동작하며, 선명도 조정 네트워크의 동작에 의해 블러 추정 네트워크에서 출력되는 디포커스 맵의 번짐 현상을 방지하고, 모양 유지 네트워크의 동작에 의해 블러 추정 네트워크의 디포커스 맵과 교사 영상과의 L2 손실 함수에 의해 발생하는 번짐 현상을 방지함으로써 정확한 디포커스 정도를 가진 디포커스 맵을 추정한다.
- [0115] 한편, 전술한 실시예들을 통해 설명한 디포커스 맵 추적을 위한 기계 학습 방법은 다양한 컴퓨터 수단을 통해 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 컴퓨터 판독 가능 매체에 기록되는 프로그램 명령은 본 발명을 위해 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용가능한 것일 수 있다.
- [0116] 컴퓨터 판독 가능 매체의 예에는 롬(rom), 램(ram), 플래시 메모리(flash memory) 등과 같이 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러(compiler)에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터(interpreter) 등을 사용해서 컴퓨터에 의해 실행될 수 있는 고급 언어 코드를 포함한다. 상술한 하드웨어 장치는 본 실시예에 따른 디포커스 맵 추적을 위한

기계 학습 방법의 일련의 동작을 수행하기 위해 적어도 하나의 소프트웨어 모듈로 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.

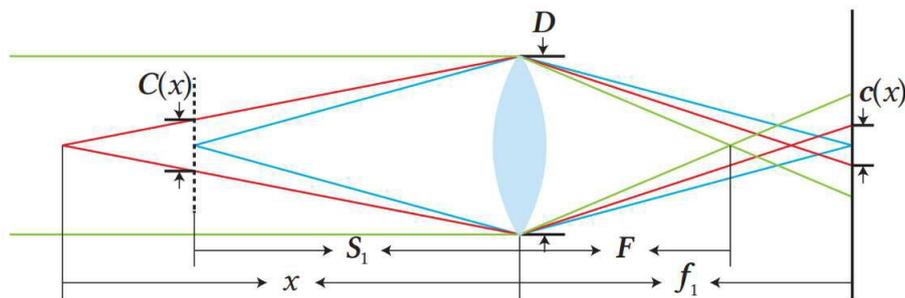
[0117] 이상 실시예를 참조하여 설명하였지만, 해당 기술 분야의 숙련된 당업자는 하기의 청구범위에 기재된 본 발명의 사상 및 영역으로부터 벗어나지 않는 범위 내에서 본 발명을 다양하게 수정 및 변경시킬 수 있음을 이해할 수 있을 것이다.

도면

도면1



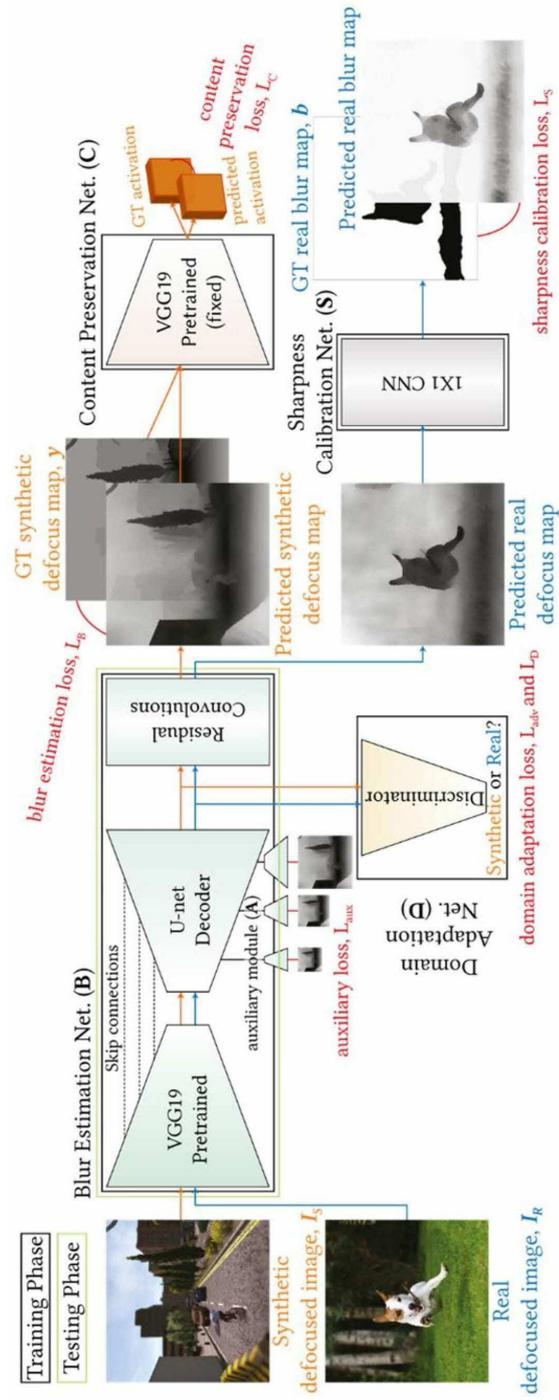
도면2



도면3



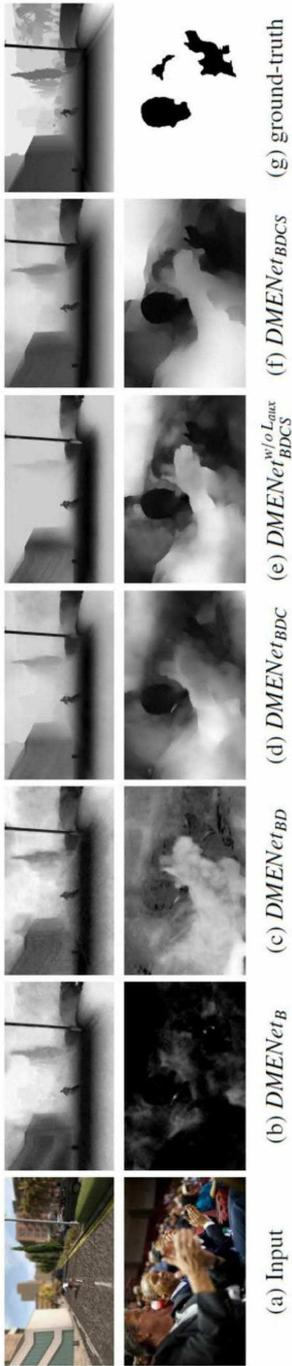
도면4



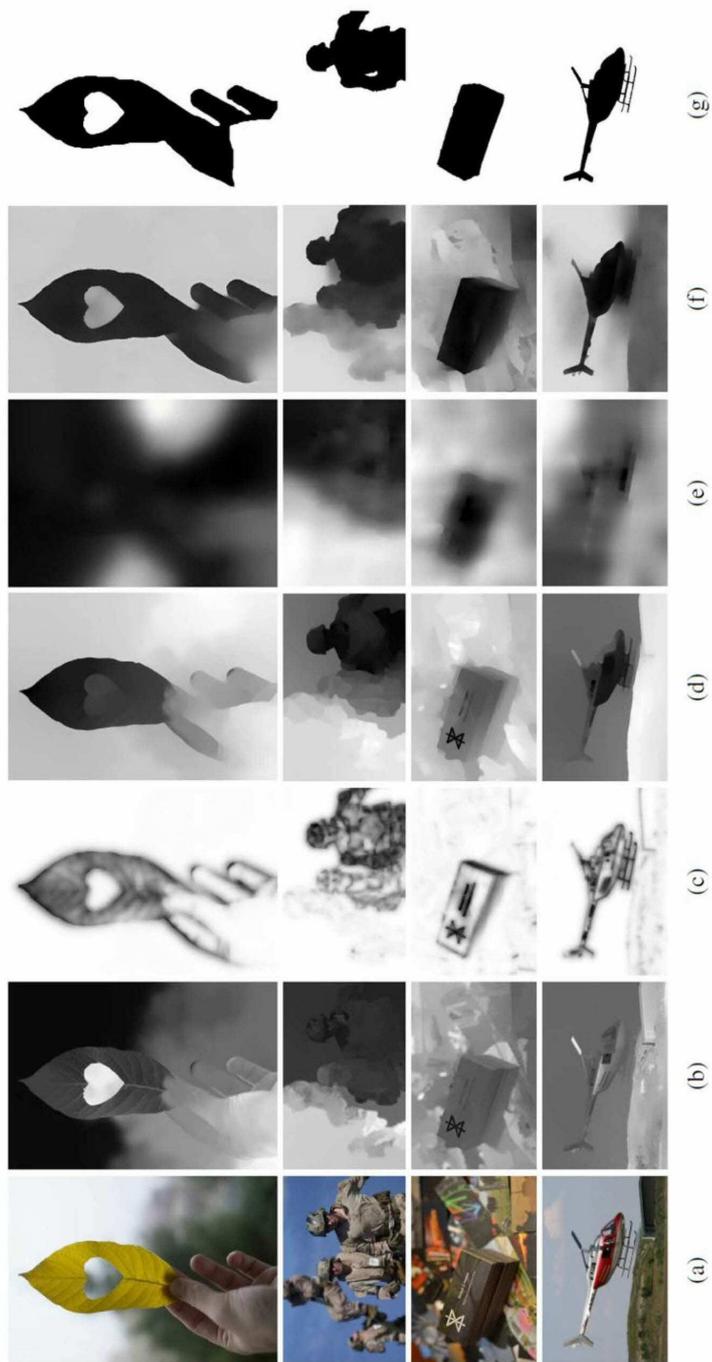
도면5



도면6



도면7



도면8

