(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2021년09월16일
(11) 등록번호 10-2303002
(24) 등록일자 2021년09월10일

(51) 국제특허분류(Int. Cl.)
G06T 5/00 (2019.01) *G06N 20/00* (2019.01)
(52) CPC특허분류
G06T 5/003 (2013.01)
G06N 20/00 (2021.08)
(21) 출원번호 10-2021-0041690
(22) 출원일자 2021년03월31일
심사청구일자 2021년03월31일
(56) 선행기술조사문헌
KR102197089 B1
KR1020190114340 A
CN110210432 A
KR1020200096026 A

(73) 특허권자
인하대학교 산학협력단
인천광역시 미추홀구 인하로 100(용현동, 인하대학교)
(72) 발명자
박인규
서울특별시 강남구 선릉로 221, 103동 1403호(도곡동, 도곡렉슬아파트)
조나단사무엘
인천광역시 미추홀구 소성로 40, 502호(학익동)
(74) 대리인
양성보

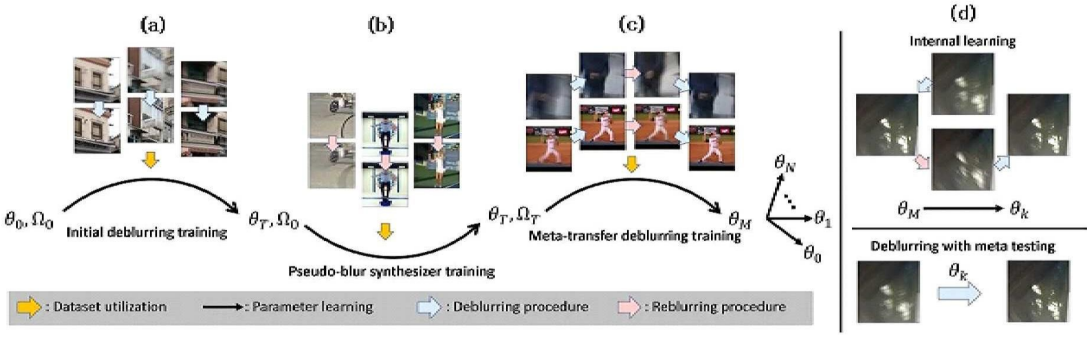
전체 청구항 수 : 총 8 항 심사관 : 김광식

(54) 발명의 명칭 **의사 블러 합성기를 이용한 사람이 포함된 영상의 디블러링 방법 및 장치**

(57) 요 약

의사 블러 합성기를 이용한 사람이 포함된 영상의 디블러링 방법 및 장치가 제시된다. 본 발명에서 제안하는 의사 블러 합성기를 이용한 사람이 포함된 영상의 디블러링 방법은 장면 블러 및 인간 블러를 포함하는 입력 영상에 대하여 초기 디블러링부를 통해 디블러링 네트워크가 수렴될 때까지 초기화 및 학습하여 학습된 디블러링 가중치를 생성하는 단계, 학습된 디블러링 가중치를 입력 받아 의사 블러 합성부를 통해 학습된 리블러 영상을 생성하는 단계, 학습된 리블러 영상을 의사 블러 합성부를 통해 학습한 후 메타 전송 디블러링 가중치를 생성하는 단계 및 최종 디블러링부를 통해 메타 전송 디블러링 가중치를 이용하여 메타 테스트를 수행하는 단계를 포함한다.

대 표 도



(52) CPC특허분류
G06T 2207/20081 (2013.01)

이 발명을 지원한 국가연구개발사업
과제고유번호 1711110995
과제번호 2019R1A2C1006706
부처명 과학기술정보통신부
과제관리(전문)기관명 한국연구재단
연구사업명 개인기초연구(과기정통부)(R&D)
연구과제명 제한없는 임의 시공간에서 취득한 영상으로부터 완전한 얼굴 복원
기 여 율 1/2
과제수행기관명 인하대학교
연구기간 2020.03.01 ~ 2021.02.28
이 발명을 지원한 국가연구개발사업
과제고유번호 1711116910
과제번호 2020-0-01389-001
부처명 과학기술정보통신부
과제관리(전문)기관명 정보통신기획평가원
연구사업명 인공지능융합연구센터지원사업(국고)
연구과제명 [Ezbaro][정부] 인공지능융합연구센터지원(1차년도)
기 여 율 1/2
과제수행기관명 인하대학교 산학협력단
연구기간 2020.04.01 ~ 2020.12.31
공지예외적용 : 있음

명 세 서

청구범위

청구항 1

장면 블러 및 인간 블러를 포함하는 입력 영상에 대하여 초기 디블러링부를 통해 디블러링 네트워크가 수렴될 때까지 초기화 및 학습하여 학습된 디블러링 가중치를 생성하는 단계;

의사 블러 합성부를 통해 상기 초기 디블러링부에서 학습된 디블러링 가중치를 입력 받아 학습된 리블러 영상을 생성하는 단계;

학습된 리블러 영상을 의사 블러 합성부를 통해 학습한 후 메타 전송 디블러링 가중치를 생성하는 단계; 및

최종 디블러링부를 통해 메타 전송 디블러링 가중치를 이용하여 메타 테스트를 수행하는 단계

를 포함하는 디블러링 방법.

청구항 2

제1항에 있어서,

장면 블러 및 인간 블러를 포함하는 입력 영상에 대하여 초기 디블러링부를 통해 디블러링 네트워크가 수렴될 때까지 초기화 및 학습하여 학습된 디블러링 가중치를 생성하는 단계는,

초기 디블러링부는 블러 RGB 영상을 입력 받아 복수의 컨볼루션 블록을 통해 전달하여 중간 특징을 생성하고,

각 컨볼루션 블록은 정상화 함수 및 활성화 함수에 이어 컨볼루션(CONV) 레이어를 포함하며, 인스턴스 정규화(Instance Normalization; IN) 및 ReLU를 적용하여 각 블록에서 시퀀스(CONV→IN→ReLU)를 출력한 후, 중간 특징은 일련의 복수의 잔여 블록(ResBlock)을 통해 전달되며,

복수의 잔여 블록은 각 블록에서 인스턴스 정규화 및 ReLU를 적용하여 각 블록에서 시퀀스(CONV→IN→ReLU→CONV→IN)로 출력하고, 시퀀스 출력은 최종 컨볼루션 레이어와 융합되고 입력된 블러 RGB 영상과 함께 디블러 영상을 출력하는

디블러링 방법.

청구항 3

제1항에 있어서,

의사 블러 합성부를 통해 상기 초기 디블러링부에서 학습된 디블러링 가중치를 입력 받아 학습된 리블러 영상을 생성하는 단계는,

의사 블러 합성부가 예비 훈련을 위해 초기 디블러링부의 상기 초기화 및 학습 동작을 정지하여 디블러링 네트워크의 성능을 유지하고, 의사 블러 합성부를 최적화 하기 위해 선명한 영상을 이용하여 먼저 학습한 후, 디블러링된 영상을 이용하여 다시 학습하여 리블러 영상을 생성하는

디블러링 방법.

청구항 4

제1항에 있어서,

학습된 리블러 영상을 의사 블러 합성부를 통해 학습한 후 메타 전송 디블러링 가중치를 생성하는 단계는,

블러, 디블러, 리블러, 디블러 시퀀스를 적용하기 위한 최적화된 디블러링 파라미터를 찾기 위해 메타 테스트에 전송될 초기 안정 가중치인 메타 전송 디블러링 가중치를 생성하도록 메타 전송 학습을 수행하고, 메타 전송 디블러링 가중치는 최적화된 테스트 손실의 평균을 사용하여 계산된 값으로 업데이트되는

디블러링 방법.

청구항 5

장면 블러 및 인간 블러를 포함하는 입력 영상에 대하여 디블러링 네트워크가 수렴될 때까지 초기화 및 학습하여 학습된 디블러링 가중치를 생성하는 초기 디블러링부;

학습된 디블러링 가중치를 입력 받아 학습된 리블러 영상을 생성하고, 학습된 리블러 영상을 학습한 후 메타 전송 디블러링 가중치를 생성하는 의사 블러 합성부; 및

메타 전송 디블러링 가중치를 이용하여 메타 테스트를 수행하는 최종 디블러링부

를 포함하는 디블러링 장치.

청구항 6

제5항에 있어서,

초기 디블러링부는,

초기 디블러링부는 블러 RGB 영상을 입력 받아 복수의 컨볼루션 블록을 통해 전달하여 중간 특징을 생성하고,

각 컨볼루션 블록은 정상화 함수 및 활성화 함수에 이어 컨볼루션(CONV) 레이어를 포함하며, 인스턴스 정규화(Instance Normalization; IN) 및 ReLU를 적용하여 각 블록에서 시퀀스(CONV→IN→ReLU)를 출력한 후, 중간 특징은 일련의 복수의 잔여 블록(ResBlock)을 통해 전달되며,

복수의 잔여 블록은 각 블록에서 인스턴스 정규화 및 ReLU를 적용하여 각 블록에서 시퀀스(CONV→IN→ReLU→CONV→IN)로 출력하고, 시퀀스 출력은 최종 컨볼루션 레이어와 융합되고 입력된 블러 RGB 영상과 함께 디블러 영상을 출력하는

디블러링 장치.

청구항 7

제5항에 있어서,

의사 블러 합성부는,

의사 블러 합성부가 예비 훈련을 위해 초기 디블러링부의 상기 초기화 및 학습 동작을 정지하여 디블러링 네트워크의 성능을 유지하고, 의사 블러 합성부를 최적화 하기 위해 선명한 영상을 이용하여 먼저 학습한 후, 디블러링된 영상을 이용하여 다시 학습하여 리블러 영상을 생성하는

디블러링 장치.

청구항 8

제5항에 있어서,

의사 블러 합성부는,

블러, 디블러, 리블러, 디블러 시퀀스를 적용하기 위한 최적화된 디블러링 파라미터를 찾기 위해 메타 테스트에 전송될 초기 안정 가중치인 메타 전송 디블러링 가중치를 생성하도록 메타 전송 학습을 수행하고, 메타 전송 디블러링 가중치는 최적화된 테스트 손실의 평균을 사용하여 계산된 값으로 업데이트되는

디블러링 장치.

발명의 설명

기 술 분 야

[0001] 본 발명은 의사 블러 합성기를 이용한 사람이 포함된 영상의 디블러링 방법 및 장치에 관한 것이다.

배 경 기 술

[0002] 블러 영상을 선명한 버전으로 복구하는 아이디어는 10여 년 전에 제시되었으며 컴퓨터 비전 분야에서 여전히 활

발한 연구 분야로 남아 있다. 픽셀의 확산은 일반적으로 캡처 중 모션 효과로 인해 영상이 흐릿해진다. 이 모션은 특정 포인트 스프레드 함수(Point Spread Function; PSF)에 의해 모델링되며 블러 커널로 표현될 수 있다. 초기 모션의 블러 영상은 소수의 추가 노이즈로 선명한 영상에서 직접 변환되는 블러 커널에 의해 모델링된다. PSF 기반 저하(PSF-based degradation)를 복원하는 작업을 모션 디블러링이라고 한다. 이전의 가정을 기반으로, 기존 방법은 예측된 블러 커널로 블러 영상을 디컨볼루션함으로써 모션 디블러링을 해결한다. 이 아이디어는 종래기술[1]~[7]에서 제시되었으며, 이전 작업들도 디블러링 과정을 돕기 위해 적용된다.

[0003] 딥 러닝의 발전과 함께, 블러 커널 프리 디블러링(blur-kernel free deblurring)에 대한 많은 접근법이 도입된다. Nah[8]의 연구는 적대적 생성망(Generative-Adversarial-Network; GAN)[10] 최적화에서 심층 ResNet 기반[9] 아키텍처로 복원된 영상의 직접 근사치를 소개한다. 이후, 이 작업은 최근 DebGAN-V1[11] 및 DebGAN-V2[12]에서 각각 Wasserstein[13] 근사치 및 경량 기능을 추가하여 개선되었다. 최근에는 로컬 영상 영역의 핸드 크래프트 정보를 활용하여 적대적 학습을 개선한다. 이 아이디어는 Iizuka[14]에 의해 판별기 모듈의 로컬 영역을 가리는 영상 인페인팅 방법에 의해 시작되었다. 이 접근법에 의해 네트워크 이전에 마스크 지도를 추가로 활용하는 것이 최근의 디블러링 알고리즘 [15]-[17]이다.

[0004] 디블러링 알고리즘이 광범위하게 이용됨에 따라, 현재 영상 복원에 비 정통적인 디블러링 접근 방식이 포함되는 것이 관찰된다. Chen[18]의 초기 작업은 디블러(deblur) 출력을 리블러링하여 디블러링을 해결한다. 그런 다음 블러 입력을 통해 리블러 출력을 감독하는 방식으로 학습한다. 간소화를 위해 Chen[18]은 테스트 체계에서 블러-디블러(B→D)를 유지하면서 훈련 과정에서 블러-디블러-리블러(B→D→R) 시퀀스를 적용한다. 최근에 Zhang[19]은 페어링되지 않은 전략을 위해 디블러링 모듈의 상단에 리블러 네트워크를 제공하여 비지도 접근법을 제공한다. 이 방법은 훈련 단계에서 B→D의 직접적인 시퀀스를 피하고, 디블러링 프레임워크에서 새로운 시퀀스를 설계한다. 이들의 접근 방식[19]은 훈련 과정에서 R→D로 표현된다. 이들의 리블러링(R) 기능[19]은 실제 선명하고 블러 영상의 페어링되지 않은 케이스를 통해 학습된다. Chen[18]과 Zhang[19]에 의해 도입된 이러한 고유한 시퀀스는 훈련 기간에만 수행되는 반면, B→D는 딥 러닝 기반 방법에 일반적인 테스트 기간 동안 수행되는 유일한 시퀀스이다.

발명의 내용

해결하려는 과제

[0005] 본 발명이 이루고자 하는 기술적 과제는 훈련 및 테스트 단계 중에 블러 데이터의 즉각적인 증강을 제공하고, 이러한 증강을 달성하기 위해, 공동 프레임워크에서 블러-디블러-리블러(reblur)-디블러 단계를 통해 디블러링하는 방식을 통합하는 방법 및 장치를 제공하는데 있다.

과제의 해결 수단

[0006] 일 측면에 있어서, 본 발명에서 제안하는 의사 블러 합성기를 이용한 사람이 포함된 영상의 디블러링 방법은 장면 블러 및 인간 블러를 포함하는 입력 영상에 대하여 초기 디블러링부를 통해 디블러링 네트워크가 수렴될 때까지 초기화 및 학습하여 학습된 디블러링 가중치를 생성하는 단계, 학습된 디블러링 가중치를 입력 받아 의사 블러 합성부를 통해 학습된 리블러 영상을 생성하는 단계, 학습된 리블러 영상을 의사 블러 합성부를 통해 학습한 후 메타 전송 디블러링 가중치를 생성하는 단계 및 최종 디블러링부를 통해 메타 전송 디블러링 가중치를 이용하여 메타 테스트를 수행하는 단계를 포함한다.

[0007] 장면 블러 및 인간 블러를 포함하는 입력 영상에 대하여 초기 디블러링부를 통해 디블러링 네트워크가 수렴될 때까지 초기화 및 학습하여 학습된 디블러링 가중치를 생성하는 단계는 초기 디블러링부는 블러 RGB 영상을 입력 받아 복수의 컨볼루션 블록을 통해 전달하여 중간 특징을 생성하고, 각 컨볼루션 블록은 정상화 함수 및 활성화 함수에 이어 컨볼루션(CONV) 레이어를 포함하며, 인스턴스 정규화(Instance Normalization; IN) 및 ReLU를 적용하여 각 블록에서 시퀀스(CONV→IN→ReLU)를 출력한 후, 중간 특징은 일련의 복수의 잔여 블록(ResBlock)을 통해 전달되며, 복수의 잔여 블록은 각 블록에서 인스턴스 정규화 및 ReLU를 적용하여 각 블록에서 시퀀스(CONV→IN→ReLU→CONV→IN)로 출력하고, 시퀀스 출력은 최종 컨볼루션 레이어와 융합되고 입력된 블러 RGB 영상과 함께 디블러 영상을 출력한다.

[0008] 학습된 디블러링 가중치를 입력 받아 의사 블러 합성부를 통해 학습된 리블러 영상을 생성하는 단계는 의사 블러 합성부가 예비 훈련을 위해 초기 디블러링부를 동결하여 성능을 유지하고, 의사 블러 합성부를 최적화 하기 위해 선명한 영상을 이용하여 먼저 학습한 후, 디블러링된 영상을 이용하여 다시 학습하여 리블러 영상을 생성

한다.

[0009] 학습된 리블러 영상을 의사 블러 합성부를 통해 학습한 후 메타 전송 디블러링 가중치를 생성하는 단계는 블러, 디블러, 리블러, 디블러 시퀀스를 적용하기 위한 최적화된 디블러링 파라미터를 찾기 위해 메타 테스트에 전송될 초기 안정 가중치인 메타 전송 디블러링 가중치를 생성하도록 메타 전송 학습을 수행하고, 메타 전송 디블러링 가중치는 최적화된 테스트 손실의 평균을 사용하여 계산된 값으로 업데이트된다.

[0010] 또 다른 일 측면에 있어서, 본 발명에서 제안하는 의사 블러 합성기를 이용한 사람이 포함된 영상의 디블러링 장치는 장면 블러 및 인간 블러를 포함하는 입력 영상에 대하여 디블러링 네트워크가 수렴될 때까지 초기화 및 학습하여 학습된 디블러링 가중치를 생성하는 초기 디블러링부, 학습된 디블러링 가중치를 입력 받아 학습된 리블러 영상을 생성하고, 학습된 리블러 영상을 학습한 후 메타 전송 디블러링 가중치를 생성하는 의사 블러 합성부 및 메타 전송 디블러링 가중치를 이용하여 메타 테스트를 수행하는 최종 디블러링부를 포함한다.

발명의 효과

[0011] 본 발명의 실시예들에 따르면 훈련 및 테스트 단계에 구현되는 블러-디블러-리블러-디블러 시퀀스에 의한 디블러링의 프레임워크를 통해 영상의 지역적인 인간과 비인간 영역을 사용하여 단일 입력에서 블러 영상을 합성할 수 있다. 또한, 의사 블러 합성기를 통해 핸드 크래프트된 휴먼 프라이어 및 적대적 학습의 조합을 학습 가능하여 디블러링 성능을 향상시킬 수 있다.

도면의 간단한 설명

[0012] 도 1은 본 발명의 일 실시예에 따른 블러 이미지 입력 및 디블러링 결과를 나타내는 도면이다.

도 2는 본 발명의 일 실시예에 따른 의사 블러 합성기를 이용한 사람이 포함된 영상의 디블러링 방법을 설명하기 위한 흐름도이다.

도 3은 본 발명의 일 실시예에 따른 의사 블러 합성기를 이용한 사람이 포함된 영상의 디블러링 장치의 구성을 나타내는 도면이다.

도 4는 본 발명의 일 실시예에 따른 디블러링 과정을 설명하기 위한 도면이다.

도 5는 본 발명의 일 실시예에 따른 본 발명의 일 실시예에 따른 디블러링 및 리블러링 과정을 설명하기 위한 도면이다.

도 6은 본 발명의 일 실시예에 따른 블러 입력 및 블러 결과를 나타내는 도면이다.

발명을 실시하기 위한 구체적인 내용

[0013] 현재 딥 러닝 기반 모션 디블러링(deblurring) 방법은 합성 블러와 날카로운 데이터 쌍을 사용하여 특정 프레임워크를 회귀시킨다. 이 작업은 흐릿한 영상 입력을 복원된 버전의 출력으로 직접 변환하도록 설계되었다. 앞서 언급된 접근법은 훈련 단계 이전에만 사용할 수 있는 합성 블러 데이터(synthetic blurry data)의 품질에 크게 의존한다.

[0014] 본 발명에서는 훈련 및 테스트 단계 중에 블러 데이터의 즉각적인 증강을 제공한다. 이러한 증강을 달성하기 위해, 공동 프레임워크에서 블러-디블러-리블러(reblur)-디블러 단계를 통해 디블러링하는 방식을 통합한다. 리블러링 모듈(다시 말해, 합성기(synthesizer))은 선명하거나 디블러된 대상의 리블러 버전(의사 블러(pseudo-blur))을 제공한다. 의사-블러 합성기는 또한 인체의 통계 모델을 사용하여 추출된 핸드 그래프트 프라이어(hand-crafted prior)가 장착되어 있다. 이러한 프라이어 정보는 적대적 학습 중 인간과 비인간 영역을 매핑하여 인간이 만든 장면 모션 블러 특성을 활용한다. 이러한 방식을 통해 본 발명의 실시예에 따른 전체 프레임워크는 최근의 딥 러닝 기반 디블러링 알고리즘에 비해 우수한 결과를 달성할 수 있다.

[0015] 본 발명에서는 인간 및 장면 모션 디블러링을 해결하기 위한 접근법으로 블러-디블러-리블러-디블러(B→D→R→D) 순서를 제안한다. 종래기술[18], [19]와는 반대로, 본 발명에서는 훈련 및 테스트 과정에서 수행된다. 이 접근 방식은 메타 학습 접근법[23]을 사용하여 수행된 종래기술[20]-[22]에 의해 동기 부여된다. 제안된 순서가 메타 학습 접근 방식[23]과 정렬되고, 먼저 B→D→R→D 시퀀스를 디블러링 네트워크를 훈련하기 위한 메타 전송 학습 과정에 매핑한다. 메타 전송 디블러링 가중치는 메타 테스트 작업 시 테스트 기간에 사용된다. 테스트 기간에는 자체 지도 전략이 적용되어 내부 학습을 가능하게 한다.

[0016] 본 발명의 실시예에 따른 시퀀스를 수용하기 위해 의사-블러 합성기 모듈을 리블러링 네트워크에 제공한다. 의사-블러 합성기 모듈은 인간 모델을 기반하여 핸드 크래프트로 사전 제작된 것으로 훈련된다 [24]. 훈련되면, 합성기는 추가적인 메타 전송 학습 및 메타 테스트 단계에 즉각적인 데이터 증강 장치로 사용된다. 이 방법을 적용하여 본 발명의 실시예에 따른 프레임워크가 원래의 블러와 합성 블러를 동시에 학습하는 동시에 정량적 및 정성적 결과에서 탁월한 향상을 얻을 수 있다.

[0018] 도 1은 본 발명의 일 실시예에 따른 블러 이미지 입력 및 디블러링 결과를 나타내는 도면이다.

[0019] 도 1(a) 및 도 1(b)는 이미지 입력 및 의사 블러 합성기를 적용하지 않은 디블러 결과를 나타내는 도면이다. 도 1(c) 및 도 1(d)는 초기 디블러링된 입력의 의사 블러링된 이미지와 제안하는 의사 블러를 통한 디블러링 결과를 나타내는 도면이다. 제안하는 의사 블러 합성기(다시 말해, 리블러러(reblurrer))는 단일 RGB 입력 이미지만 수신하고, 증강으로 그것의 블러 출력 버전을 생성한다. 본 발명의 실시예에 따르면, 리블러러는 훈련 및 테스트 단계에서 사용된다.

[0020] 본 발명의 실시예에 따르면, 훈련 데이터가 없는 블러 시나리오에서 우수한 성능을 보여주면서 훈련 및 테스트 절차에 구현되는 B→D→R→D 시퀀스에 의한 디블러링의 프레임워크를 제안한다. 또한, 영상의 지역적인 인간과 비인간 영역을 사용하여 단일 입력에서 블러 영상을 합성하는 새로운 방법을 제공한다.

[0021] 본 발명의 실시예에 따른 의사 블러 합성기에서 핸드 크래프트된 휴먼 프라이어(human prior)와 적대적 학습의 조합의 학습 가능하며, 이는 결과적으로 디블러링 성능을 향상시킬 수 있다.

[0022] 모션 디블러링 초기의 디블러링 알고리즘은 먼저 블러 커널을 추정하여 디블러링의 전형적인 방법을 활용한다. 추정된 커널은 선명한 입력을 얻기 위해 블러 입력을 디컨볼루션하는 데 사용된다. 이러한 접근법[1]-[7]을 개선하기 위해 다양한 정규화 프라이어(regularization priors)가 활용된다. 이러한 작업은 스테레오의 [25]와 라이트 필드영상의 [26], [27]의 방법에 나타낸 것처럼 다중 뷰 영상을 추가 대상으로 한다. [8]은 다양한 규모에서 견고성을 보여주는 다중 스케일 심층 프레임워크를 제안한다. 이러한 접근 방식은 학습 능력을 자동으로 증가시키는 프레임워크의 깊이 레벨에 도움이 된다. 최근 딥 러닝은 Isola[28]에 의해 대중화된 특정 도메인으로 영상을 변환하는 것으로 알려진 GAN[10]이 있고, 이 외에도 다양한 [8], [11], [12]가 있다. 디블러링의 최근 연구는 적대적 학습을 위해 핸드 크래프트된 프라이어를 포함한다. 이러한 프라이어 정보는 Shen[15]과 Ren[29]의 연구 및 Shen[16]과 Lumentut[17]의 인간 디블러링 작업에 활용된다. [15]는 각 얼굴 부위에 세그먼트 맵을 사용하여 판별기를 훈련시킨다. [29]는 생성기 앞에 렌더링된 표면을 연결한다. Shen[16]의 연구는 데이터셋에서 인간 영역에 배치된 직사각형 상자에 의해 안내되는 신경망을 활용한다. 하지만, 상자는 선명한 영상에서 주석을 달기 때문에 훈련 중 블러 영역을 나타내지 않는다. Lumentut[17]의 연구는 정교한 통계적 다인 선형 모델(SMPL)[24]을 사용하여 인간 및 비인간 영역의 실시간 매핑을 제공함으로써 이 접근 방식을 반대한다. 앞서 설명된 것처럼 블러 영상의 복원(B→D) 작업은 Chen[18]과 Zhang[19] 등의 작업에 적용된다. 이러한 접근 방식은 훈련 단계에서 리블러링 단계를 추가하면 디블러링 성능이 향상된다는 것을 보여준다.

[0023] 본 발명에서는 네트워크가 훈련 및 테스트 단계에서 내부 학습을 경험할 수 있도록 하는 새로운 B→D→R→D 방식을 제공한다.

[0024] 블러 영상을 생성하는 초기 작업은 [30]에 의해 시작된다. 여기서, 작업은 카메라를 정확한 위치로 움직이는 로봇 시스템으로 카메라 모션을 만들어낸다. 이 접근 방식은 부피가 커서 매일 사용하기 어렵다. 더 복잡한 블러 데이터셋이 [31]에 의해 도입되어 소비자 휴대전화의 관성 센서(inertial sensor)를 사용하여 기록된 모션 블러 집합이 수집된다. 이러한 모션 블러들은 합성 블러 영상을 생성하기 위해 선명한 영상으로 직접 합성된다. 최근의 접근 방식은 평균 다중 프레임이 [8]로 표현된 현실적인 장면 블러를 생성할 수 있을 만큼 충실하다는 것을 보여준다. Brooks와 Barron[32]은 두 개의 연속적인 선명한 프레임에서 흐릿한 영상을 생성하는 또 다른 접근 방식을 도입했다. 여기서, 아이디어는 부드러운 블러 결과를 생성하기 위해 두 입력 내에서 여러 개의 중간 프레임을 생성했기 때문에 프레임 보간 작업에 기초한다. [16]는 특정 이슈, 특히 인간 디블러링을 해결하기 위한 데이터셋으로 블러 인간 영상을 제공한다. 그러나 이것의 블러 결과는 인간이 만든 모션 제약 조건을 고려하지 않고 장면 모션에 의해서만 영향을 받는다. [16]과 달리, 본 발명에서는 지역적 인체 영역을 이용함으로써 달성된 합성 블러 생성에 대한 두 가지 제약 조건을 모두 고려한다. 본 발명과 가장 가까운 작업은 훈련(R→D) 중 선명한 영상을 흐리게 하는 리블러링 네트워크를 제공하는 Zhang[33]의 연구이다. 그러나 이 방법은 입력의 추가 노이즈에 전적으로 의존한다. 반대로, 본 발명의 실시예에 따른 의사 블러 합성기는 일관된 리블러 출력을 보장하는 단일RGB 입력만을 필요로 한다.

[0025] 메타 학습의 목표는 테스트 기간 동안 적응할 수 있는 네트워크의 업데이트된 버전을 제공하는 것이다. 일반적으로 메타 학습은 세 가지 그룹으로 분류된다. 첫 번째 그룹은 메트릭 기반 방법[34], [35]에 속한다. 이 접근 방식은 몇 가지 샘플에서 효율적인 학습을 제공하는 메트릭 공간을 추구하는 목표를 가지고 있다. 두 번째 그룹은 메모리 네트워크 기반 접근법[36]-[38]에 속하며, 여기서 목표는 보이지 않는 작업에 견고하도록 다양한 작업을 학습하는 네트워크를 훈련시키는 것이다. 마지막 그룹은 경사(gradient) 기반 학습이 사용되는 최적화 기반 접근 방식에 속한다. 주요 아이디어는 몇 가지 경사 업데이트[23], [39], [40] 내에서 네트워크가 적응하는 데 도움이 되는 초기 전송 가능 지점을 찾는 것이다. 경사 강하를 활용하는 최근의 (Model-Agnostic Meta-Learning; MAML)[23] 방법은 복원 작업에 상당한 영향을 미친다[20], [22]. 이러한 작업은 초고해상도 작업에서 최적화 목적으로 복원된 출력을 다시 저하하여 유사한 전략을 활용한다.

[0026] 본 발명에서는 메타 학습을 통해 디블러링 네트워크를 향상시키기 위해 의사-블러 합성기를 사용한다.

[0028] 도 2는 본 발명의 일 실시예에 따른 의사 블러 합성기를 이용한 사람이 포함된 영상의 디블러링 방법을 설명하기 위한 흐름도이다.

[0029] 제안하는 의사 블러 합성기를 이용한 사람이 포함된 영상의 디블러링 방법은 장면 블러 및 인간 블러를 포함하는 입력 영상에 대하여 초기 디블러링부를 통해 디블러링 네트워크가 수렴될 때까지 초기화 및 학습하여 학습된 디블러링 가중치를 생성하는 단계(210), 학습된 디블러링 가중치를 입력 받아 의사 블러 합성부를 통해 학습된 리블러 영상을 생성하는 단계(220), 학습된 리블러 영상을 의사 블러 합성부를 통해 학습한 후 메타 전송 디블러링 가중치를 생성하는 단계(230) 및 최종 디블러링부를 통해 메타 전송 디블러링 가중치를 이용하여 메타 테스트를 수행하는 단계(240)를 포함한다.

[0030] 단계(210)에서, 장면 블러 및 인간 블러를 포함하는 입력 영상에 대하여 초기 디블러링부를 통해 디블러링 네트워크가 수렴될 때까지 초기화 및 학습하여 학습된 디블러링 가중치를 생성한다.

[0031] 초기 디블러링부는 블러 RGB 영상을 입력 받아 복수의 컨볼루션 블록을 통해 전달하여 중간 특징을 생성한다. 각 컨볼루션 블록은 정상화 함수 및 활성화 함수에 이어 컨볼루션(CONV) 레이어를 포함하며, 인스턴스 정규화(Instance Normalization; IN) 및 ReLU를 적용하여 각 블록에서 시퀀스(CONV→IN→ReLU)를 출력한다. 이후, 중간 특징은 일련의 복수의 잔여 블록(ResBlock)을 통해 전달된다. 복수의 잔여 블록은 각 블록에서 인스턴스 정규화 및 ReLU를 적용하여 각 블록에서 시퀀스(CONV→IN→ReLU→CONV→IN)로 출력하고, 시퀀스 출력은 최종 컨볼루션 레이어와 융합되고 입력된 블러 RGB 영상과 함께 디블러 영상을 출력한다.

[0032] 단계(220)에서, 학습된 디블러링 가중치를 입력 받아 의사 블러 합성부를 통해 학습된 리블러 영상을 생성한다.

[0033] 의사 블러 합성부가 예비 훈련을 위해 초기 디블러링부를 동결하여 성능을 유지한다. 의사 블러 합성부를 최적화 하기 위해 선명한 영상을 이용하여 먼저 학습한 후, 디블러링된 영상을 이용하여 다시 학습하여 리블러 영상을 생성한다.

[0034] 단계(230)에서, 학습된 리블러 영상을 의사 블러 합성부를 통해 학습한 후 메타 전송 디블러링 가중치를 생성한다.

[0035] 블러, 디블러, 리블러, 디블러 시퀀스를 적용하기 위한 최적화된 디블러링 파라미터를 찾기 위해 메타 테스트에 전송될 초기 안정 가중치인 메타 전송 디블러링 가중치를 생성하도록 메타 전송 학습을 수행한다. 이때, 메타 전송 디블러링 가중치는 최적화된 테스트 손실의 평균을 사용하여 계산된 값으로 업데이트된다.

[0036] 단계(240)에서, 최종 디블러링부를 통해 메타 전송 디블러링 가중치를 이용하여 메타 테스트를 수행한다.

[0038] 도 3은 본 발명의 일 실시예에 따른 의사 블러 합성기를 이용한 사람이 포함된 영상의 디블러링 장치의 구성을 나타내는 도면이다.

[0039] 제안하는 의사 블러 합성기를 이용한 사람이 포함된 영상의 디블러링 장치(300)는 초기 디블러링부(310), 의사 블러 합성부(320) 및 최종 디블러링부(330)를 포함한다.

[0040] 초기 디블러링부(310), 의사 블러 합성부(320) 및 최종 디블러링부(330)는 도 2의 단계들(210~240)을 수행하기 위해 구성될 수 있다.

[0041] 초기 디블러링부(310)는 장면 블러 및 인간 블러를 포함하는 입력 영상에 대하여 디블러링 네트워크가 수렴될 때까지 초기화 및 학습하여 학습된 디블러링 가중치를 생성한다.

[0042] 초기 디블러링부(310)는 블러 RGB 영상을 입력 받아 복수의 컨볼루션 블록을 통해 전달하여 중간 특징을 생성한다. 각 컨볼루션 블록은 정상화 함수 및 활성화 함수에 이어 컨볼루션(CONV) 레이어를 포함하며, 인스턴스 정규화(Instance Normalization; IN) 및 ReLU를 적용하여 각 블록에서 시퀀스(CONV→IN→ReLU)를 출력한다. 이후, 중간 특징은 일련의 복수의 잔여 블록(ResBlock)을 통해 전달된다. 복수의 잔여 블록은 각 블록에서 인스턴스 정규화 및 ReLU를 적용하여 각 블록에서 시퀀스(CONV→IN→ReLU→CONV→IN)로 출력하고, 시퀀스 출력은 최종 컨볼루션 레이어와 융합되고 입력된 블러 RGB 영상과 함께 디블러 영상을 출력한다.

[0043] 의사 블러 합성부(320)는 학습된 디블러링 가중치를 입력 받아 학습된 리블러 영상을 생성하고, 학습된 리블러 영상을 학습한 후 메타 전송 디블러링 가중치를 생성한다.

[0044] 의사 블러 합성부(320)가 예비 훈련을 위해 초기 디블러링부(310)를 동결하여 성능을 유지한다. 의사 블러 합성부(320)를 최적화 하기 위해 선명한 영상을 이용하여 먼저 학습한 후, 디블러링된 영상을 이용하여 다시 학습하여 리블러 영상을 생성한다.

[0045] 의사 블러 합성부(320)는 학습된 리블러 영상을 학습한 후 메타 전송 디블러링 가중치를 생성한다.

[0046] 의사 블러 합성부(320)는 블러, 디블러, 리블러, 디블러 시퀀스를 적용하기 위한 최적화된 디블러링 파라미터를 찾기 위해 메타 테스트에 전송될 초기 안정 가중치인 메타 전송 디블러링 가중치를 생성하도록 메타 전송 학습을 수행한다. 이때, 메타 전송 디블러링 가중치는 최적화된 테스트 손실의 평균을 사용하여 계산된 값으로 업데이트된다.

[0047] 최종 디블러링부(330)는 최종 디블러링부를 통해 메타 전송 디블러링 가중치를 이용하여 메타 테스트를 수행한다. 도 4 내지 도 6을 참조하여 본 발명의 실시예에 따른 의사 블러 합성기를 이용한 사람이 포함된 영상의 디블러링 과정을 더욱 상세히 설명한다.

[0049] 도 4는 본 발명의 일 실시예에 따른 디블러링 과정을 설명하기 위한 도면이다.

[0050] 본 발명의 실시예에 따르면, 인간 관절 및 장면 모션의 제약을 동시에 사용하여 얻은 증강 블러에서 학습하는 디블러링 방식을 제안한다. 이러한 블러들은 특성에 따라 구별된다. 장면 블러는 카메라나 물체의 모션에 의해 생성되는 반면, 인간의 블러는 신체 관절에 의해 움직인다. 본 발명에서는 이러한 제약 조건을 해결하는 리블러 함수로서 블러 증강 장치(의사 블러 합성기(pseudo-blur synthesizer)를 구성한다. 본 발명의 실시예에서는 훈련과 테스트 단계 중에 적용할 수 있는 B→D→R→D의 방식을 제안한다. 이를 위해 도 4에 도시된 몇 가지 단계를 준비한다.

[0051] B→D→R→D를 수행하기 위해 디블러링 네트워크 θ는 여러 단계를 거친다. 리블러링 단계는 이전에 핸드 크래프트된 인간과 함께 훈련된 의사 블러 합성기 Ω에 의해 수행된다. 그런 다음 θ는 리블러링 절차를 포함하는 메타 학습 단계를 통과한다. 마지막으로, 테스트 단계에서는 먼저 리블러링 작업을 사용하여 내부 학습을 수행하고, 최종 업데이트된 θ는 최종 디블러링된 출력을 생성하는 데 사용된다.

[0052] 도 4(a)를 참조하면, 초기에는 디블러링 네트워크가 수렴될 때까지 초기화된다. 이 단계의 출력은 학습된 디블러링 가중치 θ_T이다. 그런 다음 도 4(b)를 참조하면, 디블러링 네트워크는 리블러링 네트워크(의사 블러 합성기)를 훈련한다. 이 단계의 출력은 리블러(Ω_T)의 학습된 버전이다. 그런 다음 도 4(c)를 참조하면, 이러한 출력은 메타 전송 디블러링 가중치 Ω_M을 생성하는 것을 목표로 하는 메타 전송 학습 절차를 통해 전달된다. 이 메타 학습 절차는 B→D→R→D 시퀀스를 완전히 구현한다. 이것 외에도, 나이브 방법(naive method)으로 제안하는 디블러링 네트워크를 훈련시킨다. 도 4(d)를 참조하면, 메타 전송 디블러링 가중치를 이용하여 메타 테스트를 수행한다. 나이브 방법은 여전히 훈련에서 B→D→R→D 시퀀스를 활용하지만 테스트 단계에서는 B→D만 실행한다.

[0053] 본 발명의 실시예에 따르면, 메타 학습 접근 방식을 Ours-M으로, 나이브 미세조정 접근 방식을 Ours-F로 나타낸다.

[0055] 도 5는 본 발명의 일 실시예에 따른 디블러링 및 리블러링 과정을 설명하기 위한 도면이다.

[0056] 초기 디블러링부를 통한 초기 디블러링 훈련 중에는 모듈(521, 522) 만 훈련된다. 이후, 의사 블러 합성부를 훈련하기 위해 모듈(511, 512, 513, 514, 515)이 역전파된다. 그런 다음 훈련된 리블러러(다시 말해, 의사 블러

합성부)가 메타 학습 목적으로 사용된다. 최종 디블러링부를 통한 디블러링 네트워크의 미세 조정 버전은 훈련된 버전의 모듈(521, 522) 및 모듈(511, 512, 513, 514, 515)을 모두 사용한다.

[0057] 본 발명의 실시예에 따르면, 디블러링 네트워크를 초기 훈련함으로써 전체 과정을 시작한다. 이 네트워크는 GoPro[8] 및 HIDE[16] 데이터셋에 관련하여 훈련된다. 도 5에 도시된 바와 같이, 디블러링 네트워크(521, 522)는 블러 RGB 영상 B의 입력을 수신한다. 초기에는 B가 여러 개의 컨볼루션 블록을 통해 전달되어 중간 특징을 생성한다. 각 블록은 정상화와 활성화 함수에 이어 컨볼루션 레이어로 설계된다. 본 발명에서는 인스턴스 정규화(IN) 및 ReLU 활성화를 채택하여 각 블록에서 이 시퀀스(CONV→IN→ReLU)를 출력한다. 제2 레이어와 제3 레이어의 스트라이드(strides)는 2로서 특징을 최대 4배까지 축소한다. 그런 다음 중간 특징은 일련의 잔여 블록(ResBlock)[9]을 통해 전달된다[8], [11]. 본 발명에서는 원래의 ResNet 버전과 비교하여 다르게 조정되는 9개의 ResBlock을 활용한다[9]. 특히, 각 블록에서 IN과 ReLU를 활용하고, 이는 다음 시퀀스(CONV→IN→ReLU→CONV→IN)로 출력된다. ResBlock의 출력 특징은 두 개의 전치 컨볼루션 레이어(transpose convolutional layers)를 통해 전달되며, 이 레이어는 특징을 최대 4배까지 확장한다. 마찬가지로, 레이어 시퀀스는 DEVCONV→IN→ReLU로 표시된다. 마지막으로, 출력은 최종 컨볼루션 계층과 융합되고 입력 B와 직접 추가되어 최종 디블러 출력 D(함수 f(B) = D로)를 생성한다. 본 발명의 실시예에 따른 디블러링 절차의 주요 모듈인 이 네트워크도 다음 방식을 통해 처리된다. 이 디블러링 생성기 모듈의 자세한 구성은 표 1에 나타내었다.

[0058] <표 1>

Layer	Detail	Output size	Stride
Input (B or $\hat{B}$)	-	$(H \times W \times 3)$	-
Conv 7×7	*IN+ReLU*	$(H \times W \times 64)$	1
Conv 3×3	*IN+ReLU*	$(H/2 \times W/2 \times 64)$	2
Conv 3×3	*IN+ReLU*	$(H/4 \times W/4 \times 128)$	2
Res_Blocks_1-9 3×3	*IN+ReLU*	$(H/4 \times W/4 \times 256)$	1
ConvTrans 3×3	*IN+ReLU*	$(H/2 \times W/2 \times 128)$	2
ConvTrans 3×3	*IN+ReLU*	$(H \times W \times 64)$	2
Conv 7×7	*Tanh*	$(H \times W \times 3)$	1

[0059]

[0060] 공간 출력 크기는 스트라이드 수에 따라 다운 샘플링 및 업 샘플링된다.

[0061] 정규화 및 활성화 함수는 각 표의 세부 열에 정의되어 있다. Conv 7X7의 최종 레이어는 원래 블러 입력(B 또는 $\hat{B}$)과 함께 추가되어 디블러링 출력(D)을 생성한다.

[0062] 초기 디블러링 네트워크를 최적화하기 위해 미니 배치 b, $(L_1(S,D) = \frac{1}{b}\sum_c^b \|S_c - D_c\|)$에서 디블러링된 출력 D와 선명한 실측자료 영상 S 사이의 단순 절대 오류(L_1) 계산을 활용한다. 여기서 c는 각 배치의 데이터를 나타낸다. 이 손실은 도 5에서와 같이 단일 전역(global) 디블러 판별기로 대응된다. 판별기는 [41]에 소개된 최근 최소 제곱 GAN(LSGAN)에 이어 실제 사례에 대해 선명한 S와 의사 케이스에 대해 모두 입력 받으면서 가중치를 공유한다. 디블러링 판별기의 적대적 실제 및 의사 손실은 다음과 같이 표현된다:

[0063]
$$L^d_{\mathbf{Real}} = \frac{1}{b}\sum_c^b \Pi_{\mathbf{Glo}}(S_c); L^d_{\mathbf{Fake}} = \frac{1}{b}\sum_c^b \Pi_{\mathbf{Glo}}(D_c), \quad (1)$$

[0064] 여기서 $\Pi(\cdot)$는 판별 함수로 작용한다. 두 함수는 디블러링의 생성기와 판별기 손실에 결합되며, 다음과 같이 기록된다:

[0065] 두 손실 모두 학습율이 0.0001인 ADAM 최적화 도구 집합을 사용하여 학습된다. 디블러링 판별기의 자세한 구성은 표 2에 반영되어 있다.

[0066] <표 2>

Layer	Detail	Output size	Stride
Input (S or D)	-	$(H \times W \times 3)$	-
Conv 4×4	*IN+LeakyReLU*	$(H/2 \times W/2 \times 64)$	2
Conv 4×4	*IN+LeakyReLU*	$(H/4 \times W/4 \times 128)$	2
Conv 4×4	*IN+LeakyReLU*	$(H/8 \times W/8 \times 256)$	2
Conv 4×4	*IN+LeakyReLU*	$(H/16 \times W/16 \times 512)$	2
Conv 4×4	*IN+LeakyReLU*	$(H/16 \times W/16 \times 512)$	1
Conv 4×4	*IN+Sigmoid*	$(H/16 \times W/16 \times 1)$	1

[0067]

[0068] 제안하는 판별기는 16으로 완전히 분할된 입력 영상을 필요로 하므로, 훈련 과정에서 128X128의 패치 크기를 활용한다.

[0069] 본 발명의 실시예에 따른 의사-블러 합성기(다시 말해, 리블러 네트워크)의 학습 프로세스이다. 관련 모듈은 도 5에 모듈(511, 512, 513, 514, 515)로 나타내었다. 제안하는 리블러링 프레임워크는 출력에 입력 데이터를 추가하는 것을 제외하고 디블러링 네트워크와 유사하게 구성된다. 이 모듈의 입력은 선명한 영상 S 또는 디블러된 영상 D일 수 있다. 첫 번째 단계에서, 처음 합성기를 훈련시키기 위해 S를 활용한다. 이렇게 하면 네트워크가 원래의 선명한 입력 영상에서 올바른 특징을 완전히 학습할 수 있다. 다음 단계에서는 디블러링 모듈 D의 출력을 사용하여 리블러도 학습된다. 리블러 네트워크의 훈련된 가중치는 Ω_T로 표시된다. 리블러 모듈의 자세한 구성은 표 3에 나타내고, 리블러 판별기는 표 4에 나타내었다.

[0070] <표 3>

Layer	Detail	Output size	Stride
Input (S or D)	-	$(H \times W \times 3)$	-
Conv 7×7	*IN+ReLU*	$(H \times W \times 64)$	1
Conv 3×3	*IN+ReLU*	$(H/2 \times W/2 \times 64)$	2
Conv 3×3	*IN+ReLU*	$(H/4 \times W/4 \times 128)$	2
Res_Blocks_1-9 3×3	*IN+ReLU*	$(H/4 \times W/4 \times 256)$	1
ConvTrans 3×3	*IN+ReLU*	$(H/2 \times W/2 \times 128)$	2
ConvTrans 3×3	*IN+ReLU*	$(H \times W \times 64)$	2
Conv 7×7	*Tanh*	$(H \times W \times 3)$	1

[0071]

[0072] <표 4>

Layer	Detail	Output size	Stride
Input (B or $\hat{B}$)	-	$(H \times W \times 3)$	-
Conv 4×4	*IN+LeakyReLU*	$(H/2 \times W/2 \times 64)$	2
Conv 4×4	*IN+LeakyReLU*	$(H/4 \times W/4 \times 128)$	2
Conv 4×4	*IN+LeakyReLU*	$(H/8 \times W/8 \times 256)$	2
Conv 4×4	*IN+LeakyReLU*	$(H/16 \times W/16 \times 512)$	2
Conv 4×4	*IN+LeakyReLU*	$(H/16 \times W/16 \times 512)$	1
Conv 4×4	*IN+Sigmoid*	$(H/16 \times W/16 \times 1)$	1

[0073]

[0074] 인체 및 장면 판별기는 인체영역 정보를 사용하여 추출한 생성된 지도 M으로 마스킹된 입력을 수신한다.

[0075] 리블러링 네트워크에서 훈련하기 위해 디블러링 네트워크를 동결하여 원래 θ 성능을 유지한다. 이 접근 방식과 관련된 데이터셋은 (i) HIDE[16] 및 (ii) 재구성된 블러 휴먼 데이터셋(blurry human datase)이다. 예비 훈련 리블러링 네트워크에서 훈련하기 위해 디블러링 네트워크를 동결하여 원래 성능을 유지한다. HIDE 데이터셋은 실측자료 값으로 하나의 선명한 영상과 출력으로 여러 개의 블러 영상의 특징을 제공하기 때문에 사용된다(일대 다 효과). HIDE 데이터셋[16]이 휴먼 존재(human presence)를 제공하는 것으로 알려져 있지만, 대부분의 블러는 장면 블러에서 생성된다. 본 발명에서는 이 단점을 해결하기 위해 휴먼 모션과 장면 모션을 포함하는 새로운 블러 휴먼 데이터셋을 생성한다. 이 데이터셋은 중간 영역의 선명한 인간 영상을 포함하는 Leeds Sport 데이터셋[42]의 선명한 영상에서 합성된다. 본 발명에서는 [17]에 따라 선명한 영상을 준비하고 다른 포즈로 인간 영상을 생성한다. 본 발명은 [43]의 포즈 합성기 모듈을 사용하여 약간 다른 인체 자세를 가진 7개의 연속적인 프레임을 경험적으로 생산한다. 이러한 영상은 블러 단일 출력을 생성하기 위해 함께 평균화된다. 생성된 데이터셋에는 2,000 쌍의 선명한 인간 관절 모션 블러 인간 영상이 포함되어 있다.

[0076] 리블러링 훈련의 첫 단계에서 제안하는 의사 블러 합성기는 디블러링 네트워크 D의 출력을 수신하고 이를

$(r(D)=\hat{B}^{D})$로 직접 변환하여 합성된 블러 영상을 예측한다. 그런 다음 $\hat{B}^{S}$를 예측하기 위해 선명한 입력 데이터 S를 사용하여 리블러 네트워크를 두 번째로 실행한다. 이 모듈은 적대적 방법을 통해 학습되므로 우선 $\hat{B}^{D}$및 $\hat{B}^{S}$의 차이를 데이터셋로부터의 실제 블러 영상 B와 식 4의 절대 손실 방식으로 결합한다. 그러나 일대다 효과는 B_L과 ($\hat{B}^{D}$, $\hat{B}^{S}$) 사이에 약간 상이한 색상을 생성한다. 따라서, 이러한 파라미터가 RGB에서 YUV 공간으로 변환된 후에 리블러링 손실(L_1^r)에서만 Y 채널을 활용한다. L_1^r의 전체 표현은 다음과 같다:

[0077]
$$L_1^r = \frac{1}{b}\sum_{c}^{b} \|y(B) - y(\hat{B}_c^D)\| + \|y(B) - y(\hat{B}_c^S)\|, \quad (4)$$

[0078] 여기서 $y(\cdot)$는 색상 변환 및 Y 채널 선택을 포함한다. 마지막으로, 진짜와 가짜 리블러링 데이터를 구별하기 위해 판별기 손실을 결정한다. 본 발명의 실시예에 따른 블러 휴먼 데이터와 HIDE[16] 데이터셋은 적대적 훈련에 완전히 사용된다. 진짜와 가짜 블러 영상의 입력을 마스킹하기 전에 핸드 크래프트된 인체를 활용한다. 본 발명은 Lumentut[17]와 유사한 방법으로 이진 지도 prior(M)의 추출을 따른다. 구체적으로, D와 $\hat{B}^{D}$사이의 Sobel 필터를 사용하여 먼저 에지 차이를 찾아 오차를 구한다. 이 차이 지도는 Kanazawa[44]의 정교한 인체 예측 모형을 통해 검출된 주요 포인트에서 추출한 가장 위쪽, 오른쪽, 아래쪽 및 왼쪽 좌표를 사용하여 홀을 피하도록 최대 풀링되고 잘린다. 이 인체 모델은 통계 인체 모델[24]에서 인체 관절과 체형 파라미터를 직접 나타내는 훈련된 가중치를 제공한다. 인체 관절에서 2D로 투영된 키포인트는 정점 정보보다 적은 수의 데이터를 제공하기 때문에 선택된다 [44]. 지도와 그 역 버전은 각각 마스크 바디 M_u와 장면 M_v 영역에 제공된다. 전체 블러 영상은 마스크되지 않은 (전역) 리블러 판별기를 통해 학습되는 반면, 마스크된 카운터파트(counterpart)는 인체 및 장면 판별기를 통해 처리된다. 이것은 인간의 관절 모션 블러에만 영향을 받는 인체와 그 주변 영역을 고려하기 때문에 수행된다. 배경의 나머지 영역은 장면 블러로 처리된다. 이들의 완전 적대적 손실은 다음과 같이 표현된다.

[0079]
$$L_{\mathbf{Real}}^r = \frac{1}{3b}\sum_{c}^{b} \Pi_{\mathbf{Glo}}(B_c) + \Pi_{\mathbf{u}}\mathcal{M}_{\mathbf{u}} \odot B_c) + \Pi_{\mathbf{v}}(\mathcal{M}_{\mathbf{v}} \odot B_c). \quad (5)$$

[0080] 유사하게, 본 발명에서는 리블러링의 의사 케이스에 대한 손실을 다음과 같이 모델링한다:

[0081]
$$L_{\mathbf{Fake}}^r = \frac{1}{6b}\sum_{c}^{b} \Pi_{\mathbf{Glo}}(\hat{B}_c^D) + \Pi_{\mathbf{u}}(\mathcal{M}_{\mathbf{u}} \odot \hat{B}_c^D) + \Pi_{\mathbf{Glo}}(\hat{B}_c^S)$$
$$\Pi_{\mathbf{v}}(\mathcal{M}_{\mathbf{v}} \odot \hat{B}_c^D) + \Pi_{\mathbf{u}}(\mathcal{M}_{\mathbf{u}} \odot \hat{B}_c^S) + \Pi_{\mathbf{v}}(\mathcal{M}_{\mathbf{v}} \odot \hat{S}_c^D). \quad (6)$$

[0082] 마지막으로, 다음과 같이 LSGAN[41] 접근 방식을 통해 의사 케이스에 대한 손실을 생성기 및 판별기 손실로 함께 구성한다:

[0083]
$$L_{\mathbf{Disc}}^r = 0.5 \times (\|L_{\mathbf{Real}}^r - 1\|^2 + \|L_{\mathbf{Fake}}^r\|^2); \quad (7)$$

[0084]
$$L_{\mathbf{Gen}}^r = L_1^r + 0.5 \times (\|L_{\mathbf{Fake}}^r - 1\|^2). \quad (8)$$

[0085] 본 발명의 실시예에 따른 리블러링 모듈도 학습율이 0.0001인 ADAM 방법을 사용하여 최적화되었다. 합성된 블러의 시각적 표현은 도 6에 나타내었다.

[0087] 도 6 본 발명의 일 실시예에 따른 블러 입력 및 블러 결과를 나타내는 도면이다.

[0088] 도 6(b)는 블러 입력을 나타내고, 도 6(c)는 다른 Ω_T를 사용하여 합성된 블러 결과(의사 블러)의 예시를 나타낸다. 도 6(a)는 선명한 이미지 입력을 나타낸다.

[0089] 장면 블러의 예시는 왼쪽의 처음 두 열에 표시된다. 영상 입력으로부터 전체 인체를 수신하는 관절 모션 블러는 오른쪽에서 마지막 두 열에 표시된다. 훈련 중에 데이터셋에 배치가 포함된 경우 전역, 장면 및 인간 리블러 판

별기가 활용된다. HID 데이터셋을 선택하면 전역 리블러 판별기만 사용된다. 데이터셋 선택은 각 훈련 반복 시 랜덤으로 이루어진다.

[0090] 이전 단계까지, 본 발명의 실시예에 따른 방법은 도 4와 같이 초기 디블러링 θ_T와 리블러링 가중치 Ω_T를 출력한다. 다음 목표는 B→D→R→D 절차에 적합한 최적화된 디블러링 파라미터를 찾는 것이다. 이를 달성하기 위해 메타 테스트 중에 전송될 초기 안정 가중치 θ_M을 찾는 메타 전송 학습 연산을 적용한다. 알고리즘 1은 본 발명의 실시예에 따른 접근 방식을 설명한다.

Algorithm 1: Meta-transfer Learning

Input : Pairs of blurry B and sharp images S from data distribution $\mathcal{D}$
Input : Learning rates α, β
Input : Reblurring model Ω_T
Output: Deblurring model θ_M

1 Initialize θ with θ_T
2 Initialize Ω with Ω_T
3 Generate task distribution $p(\mathcal{T})$ from $\mathcal{D}$
4 **while** *not done* **do**
5 Sample task batch $\mathcal{T}_i^{tr}$ and $\mathcal{T}_i^{te}$ from $p(\mathcal{T})$
6 **foreach** i **do**
7 **if** $\mathcal{T}_i^{tr}$ **then**
8 Deblur $f_\theta(B^{tr})$ of $\mathcal{T}_i^{tr}$: D^{tr}
9 Scale $\downarrow_{0.5\times}$ (B^{tr}, D^{tr}) : $B_\downarrow^{tr}, D_\downarrow^{tr}$
10 Deblur $f_\theta(B_\downarrow^{tr})$: D_{in}^{tr}
11 Reblur $r_\Omega(D_{in}^{tr})$: $R_\downarrow^{tr}$
12 Deblur $f_\theta(R_\downarrow^{tr})$: D_{out}^{tr}
13 Evaluate training loss: $L^{tr}(\theta) = |D_\downarrow^{tr} - D_{in}^{tr}| + |y(D_\downarrow^{tr}) - y(D_{out}^{tr})|$
14 Compute adapted parameters with gradient descent using training loss: $\theta_i \leftarrow \theta - \alpha \nabla_\theta L^{tr}(\theta)$
15 **else if** $\mathcal{T}_i^{te}$ **then**
16 Deblur $f_{\theta_i}(B^{te})$ of $\mathcal{T}_i^{te}$: D^{te}
17 Scale $\downarrow_{0.5\times}$ (B^{te}, D^{te}) : $B_\downarrow^{te}, D_\downarrow^{te}$
18 Deblur $f_{\theta_i}(B_\downarrow^{te})$: D_{in}^{te}
19 Reblur $r_\Omega(D_{in}^{te})$: $R_\downarrow^{te}$
20 Deblur $f_{\theta_i}(R_\downarrow^{te})$: D_{out}^{te}
21 Evaluate test loss: $L^{te}(\theta_i) = |D_\downarrow^{te} - D_{in}^{te}| + |y(D_\downarrow^{te}) - y(D_{out}^{te})|$
22
23 **end foreach**
24 Update θ with respect to average test loss : $\theta \leftarrow \theta - \beta \nabla_\theta \Sigma_i L^{te}(\theta_i)$
25
26 **end while**
27 **return** meta-transferred θ : θ_M

[0091]

[0092] 여기서 6-22 행은 내부 루프 구현을 보여준다. 훈련 손실($L^{tr}(\theta)$)을 사용하여 각 작업 $\mathcal{T}_i$ 내에서 α = 0.01로 가중치가 점차 업데이트된다. 메타 데이터는 β = 0.0001인 ADAM을 사용하여 최적화된 테스트 손실 ($\Sigma_i L^{te}(\theta_i)$) 의 평균을 사용함으로써 23행의 최종 가중치를 업데이트한다. 9행과 17행의 다운 스케일링은 확대 중 명확한 블러 차이를 시뮬레이션하는 것을 목표로 한다. B→D→R→D는 각각 훈련($\mathcal{T}_i^{tr}$)과 테스트($\mathcal{T}_i^{te}$) 작업을 위해 10-12행과 18-20행에서 식별된다.

[0093] 손실 함수는 디블러링 결과$(D_{in}^{tr}, D_{out}^{tr}, D_{in}^{te}, D_{out}^{te})$와 훈련 및 테스트 작업 모두의 다운 스케일링 초기 디블러링 출력$(D_\downarrow^{tr}, D_\downarrow^{te})$을 각각 비교한다. 이러한 함수는 알고리즘 1의 13과 21행에 반영된다.

[0094] 본 발명의 실시예에 따른 메타 테스트 방법은 초해상도 작업에서 제로샷 응용프로그램과 유사하게 블러 단일 영상 B만 입력으로 처리되는 내부 학습을 활용한다[22], [45]. 알고리즘 2는 테스트 절차 중 구현에 대해 설명한

다.

Algorithm 2: Meta-testing

Input : Blurry image B, meta-transfer trained model θ_M, number of gradient updates n, and learning rate α

Output: Deblurred image D

Initialize θ with θ_M
Initialize Ω with Ω_T
Deblur $f_\theta(B)$: D
Scale $\downarrow_{0.5\times}(B, D) : B_\downarrow, D_\downarrow$
Deblur $f_\theta(B_\downarrow) : D_{in}$
Reblur $r_\Omega(D_{in}) : R_\downarrow$
Deblur $f_\theta(R_\downarrow) : D_{out}$
foreach n **do**
    Evaluate test loss:
    $L(\theta) = |D_\downarrow - D_{in}| + |y(D_\downarrow) - y(D_{out})|$
    Update $\theta \leftarrow \theta - \alpha \nabla_\theta L(\theta)$

end foreach
Final deblur $f_\theta(B)$: D

[0095]

[0096] 메타 전달 학습 범위 내의 작업과 유사하게, 다운 스케일링은 원래의 블러 입력 B와 비교하여 명확한 다른 블러 스케일을 제공하기 위해 수행된다. 훈련된 리블러링 네트워크는 6행에 표시된 것처럼 의사 블러 버전을 합성하는 데 사용된다. 절대 손실 계산(9행)이 활용되며, 가중치 파라미터는 경사 강하 최적화 도구(10행)를 사용하여 점진적으로 업데이트된다. 최종 디블러링 작업(13행)은 반복 후 가장 정교한 가중치를 사용하여 수행한다. 본 발명의 실시예에 따른 구현은 TensorFlow를 사용하여 작성되고 Titan RTX GPU에서 실행된다. 모든 훈련은 128X128 공간 입력을 사용하여 공급되며 8의 미니 배치로 처리된다. 본 발명의 실시예에 따르면, Ω_0, θ_0에서 Ω_T, θ_M 까지 네트워크를 훈련하는 데 필요한 총 시간은 3일이다(도 4 참조).

[0098] 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, FPA(field programmable 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.

[0099] 소프트웨어는 컴퓨터 프로그램(computer program), 코드(code), 명령(instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성요소(component), 물리적 장치, 가상 장치(virtual equipment), 컴퓨터 저장 매체 또는 장치에 구체화(embody)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.

[0100] 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.

[0101] 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.

[0102] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

[0104] <참고자료>

[0105] [1] S. Cho and S. Lee, "Fast motion deblurring," ACM Trans. on Graphics, vol. 28, no. 5, pp. 145:1-145:8, 2009.

[0106] [2] O. Whyte, J. Sivic, A. Zisserman, and J. Ponce, "Non-uniform deblurringfor shaken images," in Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, 2010, pp. 491-498.

[0107] [3] L. Xu and J. Jia, "Two-phase kernel estimation for robust motion deblurring," in European Conference on Computer Vision, vol. 6311, 2010, pp. 157-170.

[0108] [4] D. Krishnan, T. Tay, and R. Fergus, "Blind deconvolution using a normalized sparsity measure," in Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, 2011, pp. 233-240.

[0109] [5] O. Whyte, J. Sivic, A. Zisserman, and J. Ponce, "Non-uniform deblurring for shaken images," International Journal of Computer Vision, vol. 98, no. 2, pp. 168-186, 2012.

[0110] [6] J. Pan, Z. Hu, Z. Su, and M.-H. Yang, "Deblurring text images via L0-regularized intensity and gradient prior," in Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 2901-2908.

[0111] [7] J. Pan, D. Sun, H. Pfister, and M.-H. Yang, "Blind image deblurring using dark channel prior," in Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 1628-1636.

[0112] [8] S. Nah, T. H. Kim, and K. M. Lee, "Deep multi-scale convolutional neural network for dynamic scene deblurring," in Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 257-265.

[0113] [9] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770-778.

[0114] [10] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in Advances in Neural Information Processing Systems, 2014, pp. 2672-2680.

[0115] [11] O. Kupyn, V. Budzan, M. Mykhailych, D. Mishkin, and J. Matas, "DeblurGAN: Blind motion deblurring using conditional adversarial networks," in Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 8183-8192.

[0116] [12] O. Kupyn, T. Martyniuk, J. Wu, and Z. Wang, "DeblurGAN-v2: Deblurring (orders-of-magnitude) faster and better," in Proc. of the IEEE International Conference on Computer Vision, 2019, pp. 8877-8886.

[0117] [13] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. C. Courville, "Improved training of wasserstein gans," in Advances in Neural Information Processing Systems, 2017, pp. 5767-5777.

[0118] [14] S. Iizuka, E. Simo-Serra, and H. Ishikawa, "Globally and Locally Consistent Image Completion,"

ACM Trans. on Graphics, vol. 36, no. 4, pp. 107:1-107:14, 2017.

[0119] [15] Z. Shen, W.-S. Lai, T. Xu, J. Kautz, and M.-H. Yang, "Deep semantic face deblurring," in Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 8260-8269.

[0120] [16] Z. Shen, W. Wang, X. Lu, J. Shen, H. Ling, T. Xu, and L. Shao, "Humanaware motion deblurring," in Proc. of the IEEE International Conference on Computer Vision, 2019, pp. 5571-5580.

[0121] [17] J. S. Lumentut, J. Santoso, and I. K. Park, "Human motion deblurring using localized body prior," in Asian Conference on Computer Vision, 2020.

[0122] [18] H. Chen, J. Gu, O. Gallo, M.-Y. Liu, A. Veeraraghavan, and J. Kautz, "Reblur2deblur: Deblurring videos via self-supervised learning," in IEEE International Conference on Computational Photography. IEEE, 2018, pp. 1-9.

[0123] [19] K. Zhang, W. Luo, Y. Zhong, L. Ma, B. Stenger, W. Liu, and H. Li, "Deblurring by realistic blurring," in Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, 2020, pp. 2737-2746.

[0124] [20] S. Park, J. Yoo, D. Cho, J. Kim, and T. H. Kim, "Fast adaptation to super-resolution networks via meta-learning," in Proc. of the European Conference on Computer Vision.

[0125] [21] M. Choi, J. Choi, S. Baik, T. H. Kim, and K. M. Lee, "Scene-adaptive video frame interpolation via meta-learning," in Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, 2020, pp. 9444-9453.

[0126] [22] J. W. Soh, S. Cho, and N. I. Cho, "Meta-transfer learning for zero-shot super-resolution," in Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, 2020, pp. 3516-3525.

[0127] [23] C. Finn, P. Abbeel, and S. Levine, "Model-agnostic meta-learning for fast adaptation of deep networks," in Proc. of the 34th International Conference on Machine Learning, pp. 1126-1135.

[0128] [24] M. Loper, N. Mahmood, J. Romero, G. Pons-Moll, and M. J. Black, "SMPL: a skinned multi-person linear model," ACM Trans. On Graphics, vol. 34, no. 6, pp. 248:1-248:16, 2015. [Online]. Available: https://doi.org/10.1145/2816795.2818013

[0129] [25] A. Sellent, C. Rother, and S. Roth, "Stereo video deblurring," in Proc. Of the European Conference on Computer Vision, vol. 9906. Springer, 2016, pp. 558-575.

[0130] [26] P. P. Srinivasan, R. Ng, and R. Ramamoorthi, "Light field blind motion deblurring," in Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 2354-2362.

[0131] [27] J. S. Lumentut, T. H. Kim, R. Ramamoorthi, and I. K. Park, "Deep recurrent network for fast and full-resolution light field deblurring," IEEE Signal Processing Letters, vol. 26, no. 12, pp. 1788-1792, 2019.

[0132] [28] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 5967-5976.

[0133] [29] W. Ren, J. Yang, S. Deng, D. Wipf, X. Cao, and X. Tong, "Face video deblurring using 3D facial priors," in Proc. of the IEEE International Conference on Computer Vision, 2019, pp. 9387-9396.

[0134] [30] R. Kohler, M. Hirsch, B. Mohler, B. Scholkopf, and S. Harmeling, "Recording and playback of camera shake: Benchmarking blind deconvolution with a real-world database," in European Conference on Computer Vision, vol. 7578, 2012, pp. 27-40.

[0135] [31] W.-S. Lai, J.-B. Huang, Z. Hu, N. Ahuja, and M.-H. Yang, "A comparative study for single image blind deblurring," in Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 1701-1709.

[0136] [32] T. Brooks and J. T. Barron, "Learning to synthesize motion blur," in Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, 2019, pp. 6840-6848.

[0137] [33] S. Zhang, Y. Lin, and H. Sheng, "Residual networks for light field image super-resolution," in Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, 2019, pp. 11 046-11 055.

[0138] [34] O. Vinyals, C. Blundell, T. Lillicrap, D. Wierstra et al., "Matching networks for one shot learning," in Advances in neural information processing systems, 2016, pp. 3630-3638.

[0139] [35] J. Snell, K. Swersky, and R. Zemel, "Prototypical networks for few-shot learning," in Advances in neural information processing systems, 2017, pp. 4077-4087.

[0140] [36] A. Santoro, S. Bartunov, M. Botvinick, D. Wierstra, and T. Lillicrap, "Meta-learning with memory-augmented neural networks," in International conference on machine learning, 2016, pp. 1842-1850.

[0141] [37] B. Oreshkin, P. Rodriguez Lopez, and A. Lacoste, "Tadam: Task dependent adaptive metric for improved few-shot learning," Advances in Neural Information Processing Systems, vol. 31, pp. 721-731, 2018.

[0142] [38] N. Mishra, M. Rohaninejad, X. Chen, and P. Abbeel, "A simple neural attentive meta-learner," in International conference on learning representation, 2018.

[0143] [39] E. Grant, C. Finn, S. Levine, T. Darrell, and T. Griffiths, "Recasting gradient-based meta-learning as hierarchical bayes," in International conference on learning representation, 2018.

[0144] [40] C. Finn and S. Levine, "Meta-learning and universality: Deep representations and gradient descent can approximate any learning algorithm," in International conference on learning representation, 2018.

[0145] [41] X. Mao, Q. Li, H. Xie, R. Y. Lau, Z. Wang, and S. Paul Smolley, "Least squares generative adversarial networks," in Proc. of the IEEE International Conference on Computer Vision, 2017, pp. 2813-2821.

[0146] [42] S. Johnson and M. Everingham, "Clustered pose and nonlinear appearance models for human pose estimation," in Proc. of the British Machine Vision Conference. BMVA Press, 2010, pp. 1-11.

[0147] [43] G. Balakrishnan, A. Zhao, A. V. Dalca, F. Durand, and J. Guttag, "Synthesizing images of humans in unseen poses," in Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 8340-8348.

[0148] [44] A. Kanazawa, M. J. Black, D. W. Jacobs, and J. Malik, "End-to-end recovery of human shape and pose," in Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 7122-7131.

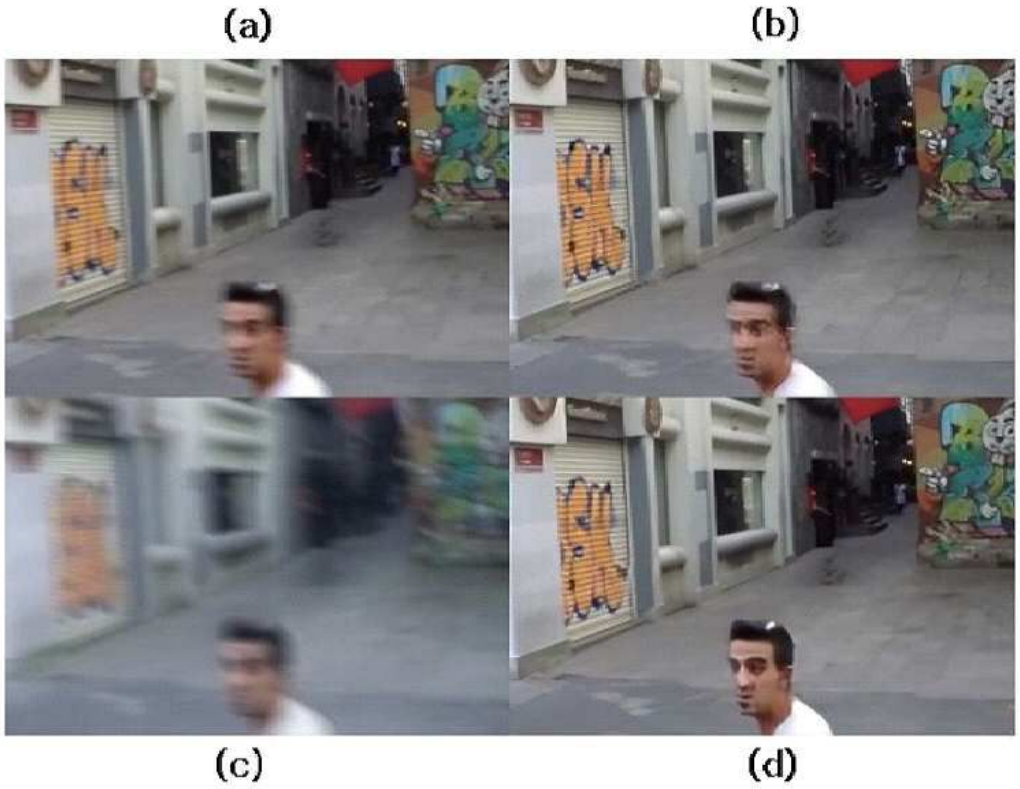
[0149] [45] A. Shocher, N. Cohen, and M. Irani, ""zero-shot" super-resolution using deep internal learning," in Proc. of the IEEE conference on computer vision and pattern recognition, 2018, pp. 3118-3126.

[0150] [46] H. Zhang, Y. Dai, H. Li, and P. Koniusz, "Deep stacked hierarchical multipatch network for image deblurring," in Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, 2019, pp. 5978-5986.

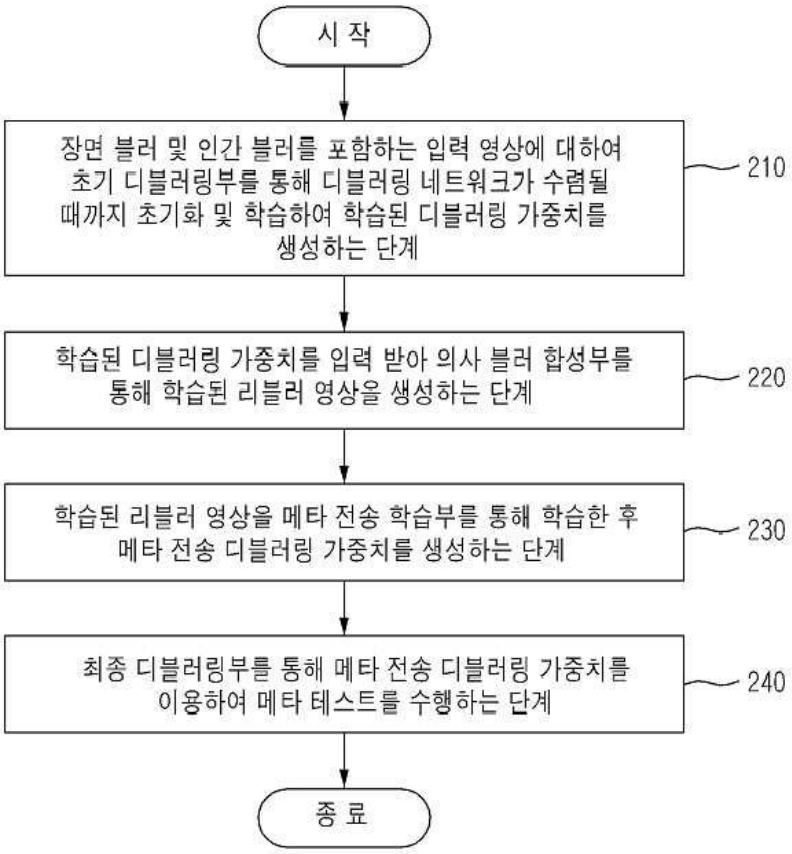
[0151] [47] M. Suin, K. Purohit, and A. Rajagopalan, "Spatially-attentive patchhierarchical network for adaptive motion deblurring," in Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, 2020, pp. 3606-3615.

도면

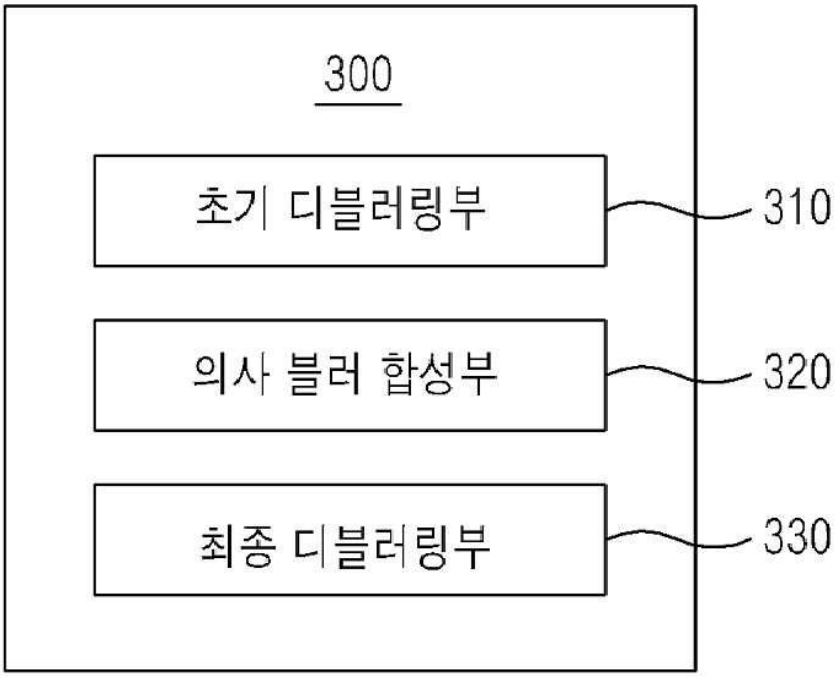
도면1



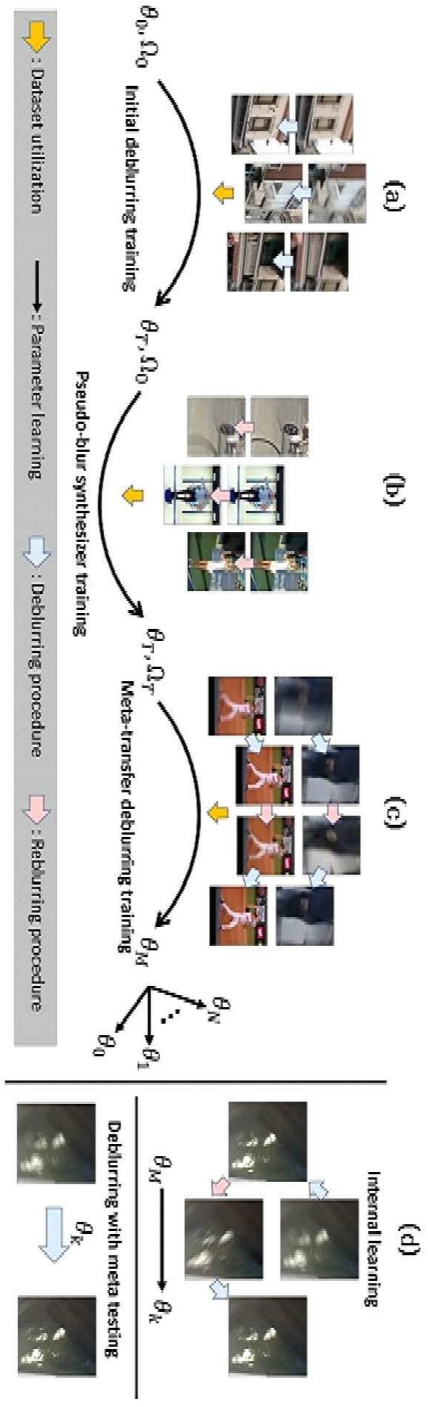
도면2


시 작
장면 블러 및 인간 블러를 포함하는 입력 영상에 대하여 초기 디블러링부를 통해 디블러링 네트워크가 수렴될 때까지 초기화 및 학습하여 학습된 디블러링 가중치를 생성하는 단계
210
학습된 디블러링 가중치를 입력 받아 의사 블러 합성부를 통해 학습된 리블러 영상을 생성하는 단계
220
학습된 리블러 영상을 메타 전송 학습부를 통해 학습한 후 메타 전송 디블러링 가중치를 생성하는 단계
230
최종 디블러링부를 통해 메타 전송 디블러링 가중치를 이용하여 메타 테스트를 수행하는 단계
240
종 료

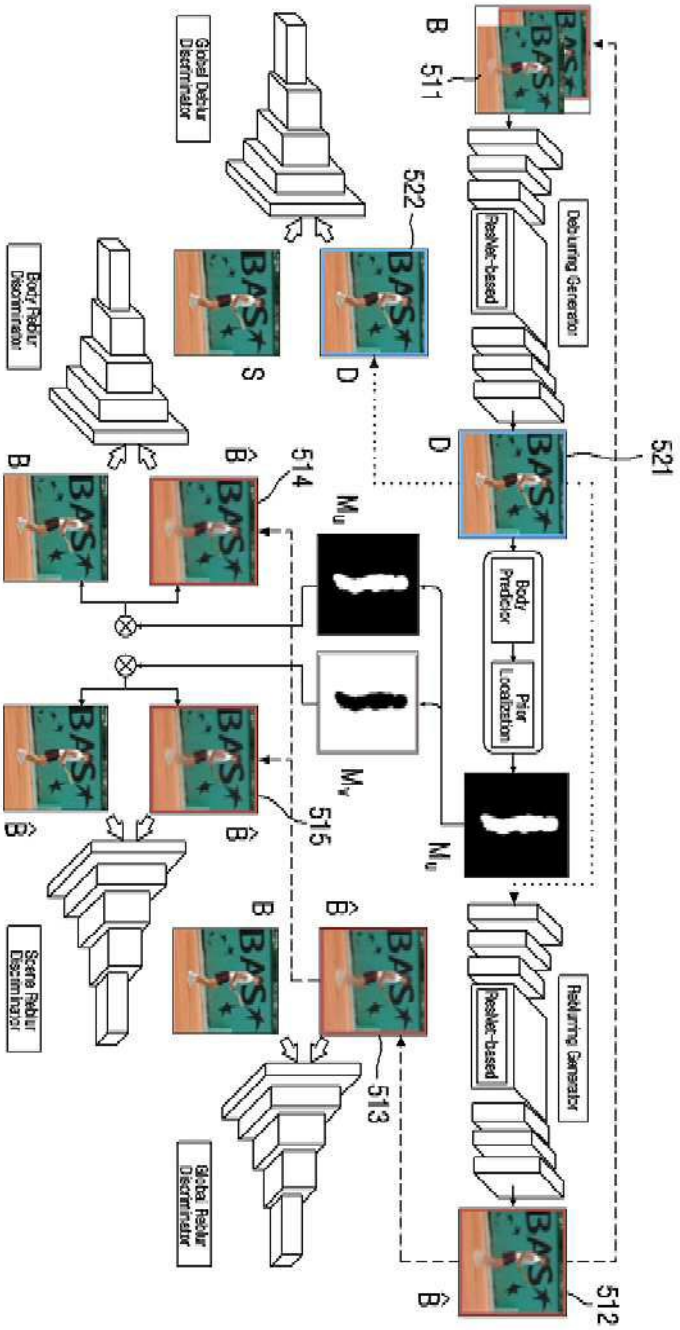
도면3


300
초기 디블러링부
310
의사 블러 합성부
320
최종 디블러링부
330

도면4


: Dataset utilization
: Parameter learning
: Deblurring procedure
: Reblurring procedure
Initial deblurring training
Pseudo-blur synthesizer training
Meta-transfer deblurring training
(a)
(b)
(c)
(d)
Internal learning
Deblurring with meta testing

도면5


Global Deblur Discriminator
Body Prediction
Prior Localization
Deblurring Generator
ResNet-based
Reblurring Generator
ResNet-based
Body Reblur Discriminator
Scene Reblur Discriminator
Global Reblur Discriminator
B
S
D
511
512
513
514
515
521
522

도면6