(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2024년05월10일
(11) 등록번호 10-2664276
(24) 등록일자 2024년05월02일

(51) 국제특허분류(Int. Cl.)
G06T 5/00 (2024.01) *G06N 3/0464* (2023.01)
G06T 5/20 (2024.01) *G06T 7/11* (2017.01)
G06T 9/00 (2019.01)
(52) CPC특허분류
G06T 5/73 (2024.01)
G06N 3/0464 (2023.01)
(21) 출원번호 10-2023-0011268
(22) 출원일자 2023년01월27일
심사청구일자 2023년01월27일
(30) 우선권주장
1020220167044 2022년12월02일 대한민국(KR)
(56) 선행기술조사문헌
Jiawei Zhang ET AL: "Learning Fully Convolutional Networks for Iterative Non-Blind Deconvolution", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3817-3825, 21-26 July 2017(2017.07.21.) 1부.*
Kiyeon Kim ET AL: "MSSNet: Multi-Scale-Stage Network for Single Image Deblurring", Computer Science_Computer Vision and Pattern Recognition, 5 April 2022(2022.04.05.) 1부.*
KR1020210028218 A
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
고려대학교 산학협력단
서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)
(72) 발명자
고성제
서울특별시 서초구 방배로 270, 나-304호
조성진
경기도 성남시 수정구 위례중앙로 216, 5313동 1102호
지서원
서울특별시 마포구 마포대로 115-8, 101동 1703호
(74) 대리인
윤귀상

전체 청구항 수 : 총 3 항 심사관 : 김광식

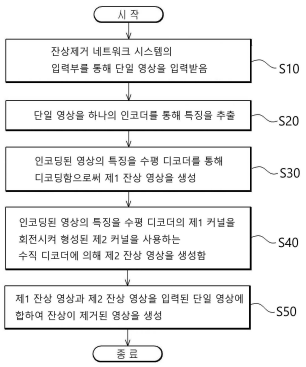
(54) 발명의 명칭 **수직-수평 분할을 통한 단일 영상 잔상 제거 방법 및 이를 위한 네트워크 시스템**

(57) 요 약

본 발명의 일실시예는 수직-수평 분할을 통한 단일 영상 잔상 제거 방법에 있어서, 잔상제거 네트워크 시스템의 입력부가 단일 영상을 입력받는 단계; 입력된 단일 영상을 하나의 인코더가 특징을 추출하는 단계; 상기 인코더로부터 제공된 특징에 대해 수평 디코더가 제1 컨볼루션 커널을 사용하여 디코딩함으로써 제1 잔상(residual) 영

(뒷면에 계속)

대 표 도 - 도4



상을 생성하는 단계; 상기 인코더로부터 제공된 특징에 대해 수직 디코더가, 상기 제1 컨볼루션 커널을 회전하여 만들어진 제2 컨볼루션 커널을 사용하여 디코딩함으로써 제2 잔상 영상을 생성하는 단계; 및 합성부가 상기 제1 잔상 영상과 제2 잔상 영상을 입력된 단일 영상에 합하여 잔상이 제거된 영상을 생성하는 단계;를 포함하는 것을 특징으로 하는, 수직-수평 분할을 통한 단일 영상 잔상 제거 방법을 제공한다.

(52) CPC특허분류
G06T 5/20 (2024.01)
G06T 7/11 (2017.01)
G06T 9/00 (2019.01)
G06T 2207/20084 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711159675
과제번호	2014-3-00077-009
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	혁신성장동력프로젝트(과기정통부)
연구과제명	(딥뷰-2세부) 대규모 실시간 비디오 분석에 의한 전역적 다중 관심객체 추적 및 상황 예측 기술 개발
기 여 율	1/1
과제수행기관명	광주과학기술원
연구기간	2014.04.01 ~ 2024.02.29

공지예외적용 : 있음

명 세 서

청구범위

청구항 1

수직-수평 분할을 통한 단일 영상 잔상 제거 방법에 있어서,

잔상제거 네트워크 시스템의 입력부가 단일 영상을 입력받는 단계;

입력된 단일 영상을 하나의 인코더가 특징을 추출하는 단계;

상기 인코더로부터 제공된 특징에 대해 수평 디코더가 제1 컨볼루션 커널을 사용하여 디코딩함으로써 제1 잔상(residual) 영상을 생성하는 단계;

상기 인코더로부터 제공된 특징에 대해 수직 디코더가, 상기 제1 컨볼루션 커널을 회전하여 만들어진 제2 컨볼루션 커널을 사용하여 디코딩함으로써 제2 잔상 영상을 생성하는 단계; 및

합성부가 상기 제1 잔상 영상과 제2 잔상 영상을 입력된 단일 영상에 합하여 잔상이 제거된 영상을 생성하는 단계;를 포함하고,

상기 제2 잔상 영상을 생성하는 단계에서, 상기 제1 컨볼루션 커널을 반시계 방향으로 회전시켜 상기 제2 컨볼루션 커널이 형성됨으로써 상기 수평 디코더가 상기 제1 잔상 영상 생성시 사용한 딥러닝 파라미터가 상기 수직 디코더와 쉐어링(sharing)되고,

상기 입력부, 상기 인코더 및 상기 수평 디코더에 의해 제1 U-shape network를 이루어 상기 제1 잔상 영상 생성을 위한 딥러닝을 통해 제1 파라미터들을 학습하고,

상기 입력부, 상기 인코더 및 상기 수직 디코더에 의해 제2 U-shape network를 이루어 상기 제1 파라미터들을 쉐어링하여 상기 제2 잔상 영상 생성을 하고,

상기 제1 잔상 영상 및 상기 제2 잔상 영상은 상기 수평 디코더 및 상기 수직 디코더에 의해 각각 추정된 잔상으로서, 상기 제1 잔상 영상 및 상기 제2 잔상 영상에는 수평 및 수직 움직임 성분이 포함되어 있고, 인코딩된 특징에 내재된 정보를 x축(수평축) 및 y축(수직축)을 따라 있는 정보로 분해할 수 있는 것을 특징으로 하는, 수직-수평 분할을 통한 단일 영상 잔상 제거 방법.

청구항 2

삭제

청구항 3

삭제

청구항 4

청구항 1에 있어서,

외부적인 감독(supervision) 없이 상기 수평 디코더는 가로 방향 잔상을 제거하도록 학습되어 상기 제1 잔상 영상을 출력하고, 상기 수직 디코더는 세로 방향 잔상을 제거하도록 학습되어 상기 제2 잔상 영상을 출력하는 것을 특징으로 하는, 수직-수평 분할을 통한 단일 영상 잔상 제거 방법.

청구항 5

삭제

청구항 6

단일 영상을 입력받는 입력부;

상기 입력된 단일 영상에서 특징을 추출하는 인코더;

상기 인코더로부터 제공된 특징에 대해 제1 컨볼루션 커널을 사용하여 디코딩함으로써 제1 잔상(residual) 영상을 생성하는 수평 디코더;

상기 인코더로부터 제공된 특징에 대해 상기 제1 컨볼루션 커널을 회전하여 만들어진 제2 컨볼루션 커널을 사용하여 디코딩함으로써 제2 잔상 영상을 생성하는 수직 디코더; 및

상기 제1 잔상 영상과 제2 잔상 영상을 입력된 단일 영상에 합하여 잔상이 제거된 영상을 생성하는 합성부;를 포함하고,

상기 수직 디코더는 제2 잔상 영상을 생성하기 위해 상기 제1 컨볼루션 커널을 반시계 방향으로 회전시켜 형성된 상기 제2 컨볼루션 커널을 사용하며, 상기 수평 디코더가 상기 제1 잔상 영상 생성시 사용한 딥러닝 파라미터를 상기 수평 디코더와 쉐어링(sharing)하고,

상기 입력부, 상기 인코더 및 상기 수평 디코더는 제1 U-shape network를 이루어 상기 제1 잔상 영상 생성을 위한 딥러닝을 통해 제1 파라미터들을 학습하고,

상기 입력부, 상기 인코더 및 상기 수직 디코더는 제2 U-shape network를 이루어 상기 제1 파라미터들을 쉐어링하여 상기 제2 잔상 영상을 생성하고,

상기 제1 잔상 영상 및 상기 제2 잔상 영상은 상기 수평 디코더 및 상기 수직 디코더에 의해 각각 추정된 잔상으로서, 상기 제1 잔상 영상 및 상기 제2 잔상 영상에는 수평 및 수직 움직임 성분이 포함되어 있고, 인코딩된 특징에 내재된 정보를 x축(수평축) 및 y축(수직축)을 따라 있는 정보로 분해할 수 있는 것을 특징으로 하는, 수직-수평 분할을 통한 단일 영상 잔상 제거 네트워크 시스템.

청구항 7

삭제

발명의 설명

기 술 분 야

[0001] 본 발명은 수직-수평 분할을 통한 단일 영상 잔상 제거 방법 및 이를 위한 네트워크 시스템에 관한 것으로, 더욱 상세하게는 딥러닝 알고리즘을 기반으로 한 영상 복원 알고리즘을 사용한 수직-수평 분할을 통한 단일 영상 잔상 제거 방법 및 이를 위한 네트워크 시스템에 관한 것이다.

배 경 기 술

[0003] 일반적으로 영상을 취득할 때, 다양한 원인으로 영상에 잔상 (blur)이 생길 수 있다. 도 1은 열화된 영상을 선명한 영상으로 복원하는 잔상제거(deblurring)의 일 예를 보여준다. 이미지 잔상제거라고 하는 것은 영상에서 잔상을 제거하는 기술로 도 1의 상측 이미지와 같은 어떤 블러한 영상을 입력으로 받았을 때 이 영상에서 잔상을 제거를 해서 하측 이미지와 같은 샤프한 영상을 만들어내는 것을 목표로 한다. 영상내 잔상은 손떨림, 객체의 움직임, out-focus 등 다양한 원인으로 발생한다. 영상 내에서 위치마다도 블러의 정도가 달라질 수 있기 때문에 잔상제거는 굉장히 어려운 문제로 여겨지고 있다. 영상에서 잔상을 제거하는 기술은 다양한 기술에 전처리로 활용되고 있으며, 심미적으로도 매우 중요할 뿐만 아니라, 비전 시스템 전체적으로도 매우 중요한 역할을 수행하고 있다. 예를 들어, 잔상으로 인한 영상의 열화는 영상 자체의 품질을 저하시켜, 미적으로 좋지 않을 뿐만 아니라 잔상이 있는 영상에서는 객체 인식 등 컴퓨터 비전 알고리즘의 성능을 하락시킬 수 있기 때문에 이를 해결하는 것은 매우 중요한 일이다.

[0004] 도 2는 U-shape network의 일 예를 나타내는 도면이다.

[0005] 일반적으로 영상의 블러를 제거하기 위한 잔상 제거 네트워크로는, 도 2에 도시된 바와 같은 U-shape의 네트워크를 사용하고 있다. U-Net은 일반적으로 영상을 입력으로 받아 몇몇 인코더로 특징맵(feature map)의 해상도를

줄여 정보를 압축하고 이를 디코더에서 다시 복원하는 구조를 가지고 있다. 이러한 UNet은 입력 영상에서 특징맵을 추출하는데, 특징맵의 해상도를 점차적으로 줄였다가 키우는 구조를 가진다. 일반적으로 UNet에서 특징맵을 줄이는 부분을 Encoder 키우는 부분을 Decoder라고 명명한다. 대부분의 네트워크는 전체적으로 하나의 인코더와 하나의 디코더로 구성이 되어있다.

[0006] 일반적으로 모션 블러는 영상 내에서 객체의 움직임, 카메라의 흔들림과 같은 복합적인 블러를 포함하고 있는데도 2과 같은 일반적인 구조는 영상내에 존재하는 블러를 효과적으로 다루기는 매우 어려울 수 있다. 특히, 전술된 복합적인 블러를 단일 디코더로는 해결하기 어렵다.

발명의 내용

해결하려는 과제

[0008] 본 발명이 이루고자 하는 기술적 과제는 하나의 디코더의 컨볼루션 커널을 다른 디코더에서 이를 회전시켜 재사용함으로 특징들을 보다 효과적으로 사용하는 수직-수평 분할을 통한 단일 영상 잔상 제거 방법 및 이를 위한 네트워크 시스템을 제공하는 것이다.

[0009] 본 발명이 이루고자 하는 기술적 과제는 이상에서 언급한 기술적 과제로 제한되지 않으며, 언급되지 않은 또 다른 기술적 과제들은 아래의 기재로부터 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에게 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

[0011] 상기 기술적 과제를 달성하기 위하여, 본 발명의 일실시예는 수직-수평 분할을 통한 단일 영상 잔상 제거 방법에 있어서, 잔상제거 네트워크 시스템의 입력부가 단일 영상을 입력받는 단계; 입력된 단일 영상을 하나의 인코더가 특징을 추출하는 단계; 상기 인코더로부터 제공된 특징에 대해 수평 디코더가 제1 컨볼루션 커널을 사용하여 디코딩함으로써 제1 잔상(residual) 영상을 생성하는 단계; 상기 인코더로부터 제공된 특징에 대해 수직 디코더가, 상기 제1 컨볼루션 커널을 회전하여 만들어진 제2 컨볼루션 커널을 사용하여 디코딩함으로써 제2 잔상 영상을 생성하는 단계; 및 합성부가 상기 제1 잔상 영상과 제2 잔상 영상을 입력된 단일 영상에 합하여 잔상이 제거된 영상을 생성하는 단계;를 포함하는 것을 특징으로 하는, 수직-수평 분할을 통한 단일 영상 잔상 제거 방법을 제공한다.

[0012] 본 발명의 실시예에 있어서, 상기 제2 잔상 영상을 생성하는 단계에서, 상기 제1 컨볼루션 커널을 반시계 방향으로 회전시켜 상기 제2 컨볼루션 커널이 형성됨으로써 상기 수평 디코더가 상기 제1 잔상 영상 생성시 사용한 딥러닝 파라미터가 상기 수직 디코더와 쉐어링(sharing)될 수 있다.

[0013] 본 발명의 실시예에 있어서, 상기 입력부, 상기 인코더 및 상기 수평 디코더에 의해 제1 U-shape network를 이루어 상기 제1 잔상 영상 생성을 위한 딥러닝을 통해 제1 파라미터들을 학습하고, 상기 입력부, 상기 인코더 및 상기 수직 디코더에 의해 제2 U-shape network를 이루어 상기 제1 파라미터들을 쉐어링하여 상기 제2 잔상 영상 생성을 할 수 있다.

[0014] 본 발명의 실시예에 있어서, 외부적인 감독(supervision) 없이 상기 수평 디코더는 가로 방향 잔상을 제거하도록 학습되어 상기 제1 잔상 영상을 출력하고, 상기 수직 디코더는 세로 방향 잔상을 제거하도록 학습되어 상기 제2 잔상 영상을 출력할 수 있다.

[0015] 본 발명의 실시예에 있어서, 상기 제1 잔상 영상 및 상기 제2 잔상 영상은 상기 수평 디코더 및 상기 수직 디코더에 의해 각각 추정된 잔상으로서, 상기 제1 잔상 영상 및 상기 제2 잔상 영상에는 수평 및 수직 움직임 성분이 포함되어 있고, 인코딩된 특징에 내재된 정보를 x축 및 y축을 따라 있는 정보로 분해할 수 있다.

[0016] 상기 기술적 과제를 달성하기 위하여, 본 발명의 다른 실시예는 단일 영상을 입력받는 입력부; 상기 입력된 단일 영상에서 특징을 추출하는 인코더; 상기 인코더로부터 제공된 특징에 대해 제1 컨볼루션 커널을 사용하여 디코딩함으로써 제1 잔상(residual) 영상을 생성하는 수평 디코더; 상기 인코더로부터 제공된 특징에 대해 상기 제1 컨볼루션 커널을 회전하여 만들어진 제2 컨볼루션 커널을 사용하여 디코딩함으로써 제2 잔상 영상을 생성하는 수직 디코더; 및 상기 제1 잔상 영상과 제2 잔상 영상을 입력된 단일 영상에 합하여 잔상이 제거된 영상을 생성하는 합성부;를 포함하는 것을 특징으로 하는, 수직-수평 분할을 통한 단일 영상 잔상 제거 네트워크 시스템을 제공한다.

[0017] 본 발명의 실시예에 있어서, 상기 수직 디코더는 제2 잔상 영상을 생성하기 위해 상기 제1 컨볼루션 커널을 반시계 방향으로 회전시켜 형성된 상기 제2 컨볼루션 커널을 사용하며, 상기 수평 디코더가 상기 제1 잔상 영상 생성시 사용한 딥러닝 파라미터를 상기 수평 디코더와 쉐어링(sharing)할 수 있다.

발명의 효과

[0019] 본 발명의 실시예에 의하면, 하나의 디코더의 컨볼루션 커널을 다른 디코더에서 이를 회전시켜 재사용함으로 특징들을 보다 효과적으로 사용하는 수직-수평 분할을 통한 단일 영상 잔상 제거 방법 및 이를 위한 네크워크 시스템을 제공할 수 있다.

[0020] 본 발명의 실시예에 의하면, 수직-수평 분할을 통한 단일 영상 잔상 제거 방법 및 이를 위한 네트워크 시스템은 UNet 등 여러 네트워크에 적용될 수 있으며, 해당 네크워크의 잔상 제거 성능을 향상시킬 수 있다.

[0021] 본 발명의 효과는 상기한 효과로 한정되는 것은 아니며, 본 발명의 상세한 설명 또는 특허청구범위에 기재된 발명의 구성으로부터 추론 가능한 모든 효과를 포함하는 것으로 이해되어야 한다.

도면의 간단한 설명

[0023] 도 1은 열화된 영상을 선명한 영상으로 복원하는 잔상제거(deblurring)의 일 예를 보여준다.

도 2는 U-shape network의 일 예를 나타내는 도면이다.

도 3은 본 발명의 일 실시예에 따른 수직-수평 분할을 통한 단일 영상 잔상 제거 방법을 구현하는 잔상제거 네트워크 시스템의 일 예를 나타낸다.

도 4는 본 발명의 일 실시예에 따른 수직-수평 분할을 통한 단일 영상 잔상 제거 방법을 나타내는 순서도이다.

도 5는 GoPro 데이터셋에서의 잔상제거 결과를 나타낸다.

도 6는 RealBlur 데이터셋에서의 잔상제거 결과를 나타낸다.

도 7은 GoPro 데이터셋 및 RealBlur 데이터셋에서 잔상제거 결과를 비교하는 표이다.

도 8 및 도 9는 GoPro 데이터셋에서의 결과영상을 나타낸다.

발명을 실시하기 위한 구체적인 내용

[0024] 본 발명은 다양한 변경을 가할 수 있고 여러 가지 형태를 가질 수 있는 바, 특정 실시예들을 도면에 예시하고 본문에 상세하게 설명하고자 한다. 그러나, 이는 본 발명을 특정한 개시 형태에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다. 각 도면을 설명하면서 유사한 참조부호를 유사한 구성요소에 대해 사용하였다.

[0025] 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 가지는 것으로 해석되어야 하며, 본 출원에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.

[0026] 이하, 첨부한 도면들을 참조하여, 본 발명의 바람직한 실시예를 보다 상세하게 설명하고자 한다.

[0027] 단일 이미지 디블러링(deblurring)을 위한 많은 컨볼루션 신경망(CNN)은 U-Net 구조를 사용하여 잠재적인 선명한 이미지를 추정한다. 오랫동안 이미지 복원 작업에서 효과적인 것으로 입증된 인코더-디코더 아키텍처의 단일 레인은 얽힌 동작으로 인해 복잡한 블러 커널에서 흐릿한 이미지가 생성되는 디블러링의 특성을 간과한다. 단일 이미지 디블러링을 위한 효과적인 네트워크 아키텍처를 위해 본 발명은 하나의 인코더-2-디코더 아키텍처로 보완적인 하위 솔루션 학습을 통해서 잔상을 제거하는 방법을 제시한다. 본 발명은 여러 디코더가 인코딩된 특징 정보를 방향성 구성 요소로 분해하는 방법을 성공적으로 학습하며, 어느 하나의 디코더에서 활용되는 컨볼루션 커널을 회전시켜서 생성된 커널을 다른 디코더가 공유함으로써, 디코더가 색상 이동과 같은 불필요한 구성 요소를 분리하는 것을 방지하며, 또한 네트워크 효율성과 디블러링 성능을 더욱 향상시킨다. 본 발명에 따른 수직-수평 분할을 통한 단일 영상 잔상 제거 방법 및 이를 위한 잔상제거 네트워크 시스템은 네트워크 파라미터를 유지하면서 UNet에 비해 우수한 결과를 보이며 다른 네트워크에 기반 네트워크로 채택하여 사용하면 기존의 디블

러링 네트워크의 성능을 향상시킬 수 있다.

[0028] 본 발명의 일 실시예에 따른 수직-수평 분할을 통한 단일 영상 잔상 제거 방법은 잔상 제거라는 큰 문제를 여러 하위 문제로 쪼개고 이를 해결하기 위해 다중 디코더 구조를 사용한다. 각 디코더의 결과는 최종적으로 입력영상과 합쳐지기 때문에, 잔상제거 네트워크 시스템은 잔상 내지 잔차(Residual)를 학습하게 된다. 본 발명의 실시예들은 실험적으로 인코더-디코더 네트워크를 학습시키는 과정에서 외부적인 감독(supervision) 없이 각 디코더가 가로 방향 블러 및 세로 방향의 블러(blur)를 제거하도록 학습되는 특징의 발견에 근거하고 있다.

[0029] 도 3은 본 발명의 일 실시예에 따른 수직-수평 분할을 통한 단일 영상 잔상 제거 방법을 구현하는 잔상제거네트워크 시스템의 일 예를 나타낸다.

[0030] 도 3을 참조하면, 본 실시예의 잔상제거 네트워크 시스템은 입력부(10), 인코더(20), 수평 디코더(30), 수직 디코더(40) 및 합성부(50)를 포함한다.

[0031] 입력부(10)는 단일 영상을 입력받을 수 있다. 인코더(20)는 입력된 단일 영상을 인코딩하여 특징을 추출한다. 여기서 인코딩된 이미지 특징은 다음 식(1)과 같이 얻을 수 있다.

[0032] $$\mathbf{z} = \mathcal{E}(\mathbf{x}; \Theta_e) \quad (1)$$

[0033] 여기서 $\mathcal{E}$는 매개변수가 Θe인 인코더를 나타낸다.

[0034] 일반적으로 인코더 구조는 입력부를 통해 입력된 단일 영상에 대해 해상도를 줄여가면서 순차적으로 배열된 채널들을 통해 정보를 추출한다. 그러면, 어떤 채널은 이전 채널에 비해 높이가 절반으로 감소하면 가로 반 세로 반이 줄어들어서 정보가 4배 감소될 수 있다. 그러면 이를 보존하기 위해서 보통 채널을 2배로 늘리게 된다. 즉, 인코더에서는 단계가 진행될수록 채널이 늘어나는 형태를 가지게 된다.

[0035] 수평 디코더(30)는 인코더(20)로부터 제공된 특징에 대해 제1 컨볼루션 커널을 사용하여 디코딩함으로써 제1 잔상(residual) 영상을 생성할 수 있다. 수직 디코더(40)는 인코더(20)로부터 제공된 특징에 대해 수직 디코더가, 상기 제1 컨볼루션 커널을 회전하여 만들어진 제2 컨볼루션 커널을 사용하여 디코딩함으로써 제2 잔상 영상을 생성할 수 있다.

[0036] 일반적으로 잔상 제거 네트워크는 U자 모양의 Unet(U-shape network)을 기반으로 개발되고 있다. 이러한 UNet은 입력 영상에서 특징맵을 추출하는데, 특징맵의 해상도를 점차적으로 줄였다가 키우는 구조를 가진다. UNet에서 특징맵을 줄이는 부분을 인코더(Encoder) 키우는 부분을 디코더(Decoder)라고 명명한다. 대부분의 네트워크는 전체적으로 하나의 인코더와 하나의 디코더로 구성이 되어있다. 반면, 도 3에 제시된 본 실시예의 잔상제거 네트워크 시스템은 하나의 인코더(20)와 두개의 디코더(30,40)로 된(one-encoder-two-decoder) 네트워크 시스템이다. 도 4에서 인코더 및 디코더는 박스들의 배열로 표현되어 있다. 인코더 및 두개의 디코더들에서 박스들의 세로 방향 길이, 즉 높이는 해상도를 나타낸다. 각 박스의 가로폭은 딥러닝 네트워크에서의 채널에 대응된다.

[0037] 이와 같은 one-encoder-two-decoder 네트워크를 사용하여 흐릿한 이미지(입력된 단일 영상)에서 잔상 r = y - x를 추정하도록 네트워크를 설계할 수 있다. 잔상 추정 $\hat{\mathbf{r}}$은 수식(2)과 같이 표현되는 학습 가능한 매개변수 Θ를 사용하여 도 3에 예시된 네트워크 $\mathcal{F}$를 통해 얻을 수 있다.

[0038] $$\hat{\mathbf{r}} = \mathcal{F}(\mathbf{x}; \Theta) \quad (2)$$

[0039] 본 실시예에서는 다음 식(3)과 같이 두 개의 별도 디코더 네트워크 Dhor 및 Dver를 사용하여 잔상 추정을 얻을 수 있다.

[0040] $$\hat{\mathbf{r}} = \mathcal{D}_{\text{hor}}(\mathbf{z}; \Theta_{\text{hor}}) + \mathcal{D}_{\text{ver}}(\mathbf{z}; \Theta_{\text{ver}}) \quad (3)$$

[0041] 여기서 Θhor와 Θver는 각각 Dhor와 Dver의 네트워크 매개변수이고 z는 인코딩된 이미지 특징을 나타낸다. 두 개의 디코더 Dhor(수평 디코더, 30)와 Dver(수직 디코더, 40)가 동일한 구조를 가질 수 있다. 인코더와 두 개의 디코더 중 하나만 사용하면 네트워크 구조가 공통 U-Net 구조로 축소된다.

[0042] 도 3의 제1 잔상 영상(r1) 및 제2 잔상 영상(r2)을 참조하면, 입력부(10), 인코더(20) 및 상기 수평 디코더(30)는 제1 U-shape network로서 상기 제1 잔상 영상 생성을 위한 딥러닝을 통해 제1 파라미터들을 학습할 수 있다. 또한, 입력부(10), 인코더(20) 및 수직 디코더(40)는 제2 U-shape network로서 상기 제1 파라미터들을 쉐어

링하여 상기 제2 잔상 영상 생성을 위한 학습을 할 수 있다. 도 3의 제1 잔상 영상 및 제2 잔상 영상은 두 디코더 Dhor(수평 디코더)와 Dver(수직 디코더)에서 추정된 잔상의 예를 보여준다. 본 실시예의 잔상제거 네트워크 시스템에서, 분리된 두 디코더를 사용하여 얻은 두 잔상에는 수평 및 수직 움직임 성분이 포함되어 있다. 즉, 수평 디코더가 출력하는 제1 잔상에는 수평 움직임 성분이 포함되어 있고, 수직 디코더가 출력하는 제2 잔상에는 수직 움직임 성분이 포함되어 있다. 이와 같은 제1 잔상 영상 및 제2 잔상 영상의 성질은 본 발명에 따른 잔상제거 네트워크 시스템을 통해 네트워크가 명시적 제약을 적용하지 않고 인코딩된 특징에 내재된 정보를 x축(수평축)및 y축(수직축)을 따라 있는 정보로 분해할 수 있음을 의미한다. 이는 보다 효율적인 디블러링을 위해 디코더의 매개변수를 공유하는 방식을 사용할 수 있도록 하는 관찰적 근거가 될 수 있다. 본 실시예의 두 디코더는 서로 직각인 축을 따라 고유 정보를 학습할 수 있으므로 다음 식(4)와 같이 두 디코더의 네트워크 매개변수를 공유할 수 있다.

[0043] $$\hat{\mathbf{r}} = \mathcal{D}_{\mathrm{hor}}(\mathbf{z};\Theta_{\mathrm{d}}) + \mathcal{D}_{\mathrm{ver}}(\mathbf{z};\Theta_{\mathrm{rd}}) \quad (4)$$

[0044] 여기서 Θrd는 Θd의 네트워크 매개변수를 시계 반대 방향으로 90도 회전시킨 매개변수이다. 기본 디코더로 Dhor(수평 디코더)를 선택할 수 있다. 평균적으로 Dhor의 결과 잔상은 Dver(수직 디코더)의 신호 전력(power)보다 더 큰 신호 전력을 가지며, 이는 수직축에 비해 수평축을 따라 더 큰 변화가 발생함을 의미한다. 이것은 카메라가 yaw 축을 중심으로 360도 회전할 수 있기 때문에 자연스러운 현상이지만 사진을 찍는 동안 피치 회전은 다소 제한적이다. 수평 디코더가 항상 기본 디코더가 되는 것은 아니고, 블러한 이미지에서 잔상 내지 잔차 학습 결과 신호 전력이 가장 크게 나오는 방향을 기본 방향으로 하여 기본 디코더를 정하고 이에 수직인 방향으로 보완 디코더를 설정하는 것도 가능할 것이다. 예를 들어, 낙하하는 물체에 대한 영상 또는 상하로 카메라가 움직이는 경우는 수직 방향으로 더 신호 전력이 출력될 수도 있다.

[0045] 합성부(50)는 제1 잔상 영상과 제2 잔상 영상을 입력된 단일 영상(Blurry Image)에 합하여 잔상이 제거된 영상(Restored Image)을 생성할 수 있다.

[0046] 도 4는 본 발명의 일 실시예에 따른 수직-수평 분할을 통한 단일 영상 잔상 제거 방법을 나타내는 순서도이다.

[0047] 도 4를 참조하면, 수직-수평 분할을 통한 단일 영상 잔상 제거 방법에서, 먼저, 잔상제거 네트워크 시스템의 입력부(10)가 단일 영상을 입력받는다(S10). 입력된 단일 영상을 하나의 인코더(20)가 특징을 추출한다(S20). 인코더로부터 제공된 특징에 대해 수평 디코더(30)가 제1 컨볼루션 커널을 사용하여 디코딩함으로써 제1 잔상(residual) 영상을 생성한다(S30). 인코더로부터 제공된 특징에 대해 수직 디코더(40)가, 상기 제1 컨볼루션 커널을 회전하여 만들어진 제2 컨볼루션 커널을 사용하여 디코딩함으로써 제2 잔상 영상을 생성한다(S40). 합성부(50)가 상기 제1 잔상 영상(Residual1)과 제2 잔상 영상(Residual2)을 입력된 단일 영상에 합하여 잔상이 제거된 영상을 생성한다(S50).

[0048] 이하, 더 상세히 설명한다.

[0049] 본 실시예의 수직-수평 분할을 통한 단일 영상 잔상 제거 방법에서는, 일반적인 네트워크와는 다르게, 인코더로부터 제공된 이미지 특징에 대해 수평 디코더를 통해 디코딩하여 제1 잔상 영상을 생성하고(S30), 인코더로부터 제공된 이미지 특징에 대해 수직 디코더를 통해, 상기 수평 디코더가 상기 제1 잔상 영상 생성을 위해 사용한 제1 컨볼루션 커널을 반시계 방향으로 90도 회전하여 만들어진 제2 컨볼루션 커널을 사용하여 디코딩하여 제2 잔상 영상을 생성한다(S40). 이와 같이, 본 실시예에서는 하나의 디코더 형태를 두 개의 디코더, 즉 수평 디코더 및 수직 디코더로 분할하고, 수평 디코더는 수평 방향 잔상을 제거하도록 학습되고, 수직 디코더는 수직 방향의 잔상을 제거하도록 학습될 수 있다. 본 실시예에서는 잔상제거 네트워크 시스템의 인코더 및 디코더로만 학습을 하는 것이 아니라 잔상 영상 자체를 학습에 제공할 수 있다.

[0050] 예를 들어, 수평 디코더는 주어진 2D 장면에서 기본 잔상(제1 잔상 영상)을 암시적으로 생성하고, 수직 디코더는 분리된 디코더의 보완 잔상(제2 잔상 영상)을 생성할 수 있다. 본 실시예의 방법은 명시적인 감독 없이 기본 잔상 및 보완 잔상이 각각 이미지 평면의 가로(수평) 및 세로(수직) 축 방향을 따라 흐릿한 가장자리(blurred edges)를 포함한다는 것을 관찰한 바에 근거하여 X축 방향(수평 방향) 및 Y축 방향(수직 방향)으로 잔상 제거를 분할(decomposition)하여 실행하는 방법이다. 이러한 관찰을 기반으로 수평 디코더에서 제1 콘볼루션 커널을 공간적으로 회전하고 수직 디코더와 매개변수를 공유하여 1-인코더-2-디코더 구조를 더욱 개선할 수 있다. 또한, 본 실시예의 방법은 네트워크 효율성을 위해 성능을 희생하는 CNN의 기존 매개변수 공유 접근 방식과 달리 회전에 의한 커널 및 매개변수의 공유 내지 재활용을 하는 분리된 디코더 구조는 매개변수 수를 줄일 뿐만 아니라 디블러링 작업과 관련이 없는 특징의 바람직하지 않은 분리를 제거하여 성능을 향상시킬 수 있다.

[0051] 수평 디코더(30) 및 수직 디코더(40)의 작업 이후, 합성부(50)가 제1 잔상 영상과 제2 잔상 영상을 입력된 단일 영상에 합하여 잔상이 제거된 영상을 생성한다(S50). 즉, 합성부는 두 개의 디코더로부터 출력되는 잔상 영상(제1 잔상 영상, 제2 잔상 영상)에 입력 영상을 더해서 복원 네트워크를 학습시킬 수 있다. 즉, 잔상 영상은 어떤 의미에서는 마이너스 블록 컴포넌트로서, 인코더-디코더 네트워크는 이러한 마이너스 블록 컴포넌트를 만들어내도록 학습될 수 있다.

[0052] 본 발명의 실시예에 따른 수직-수평 분할을 통한 단일 영상 잔상 제거 방법은 마이너스 잔차만 학습하는 것이 아니라 이 잔차를 가로 방향 성분 및 세로 방향 성분으로 쪼개서 두 개로 분할하여 각각 해당 블러를 제거하도록 하는 점이 주요 특징이다. 구체적으로, 전술한 바와 같이, 하나의 네트워크의 컨볼루션 커널을 재사용한다. 이때, 본 실시예에서는 수평 디코더 및 수직 디코더가 수평(가로), 수직(세로)으로 잔상(블러)을 제거하기 때문에, 컨볼루션 커널의 재활용 방법으로서 수평 디코더의 제1 컨볼루션 커널을 반시계 방향으로 회전시켜 이루어진 제2 컨볼루션 커널을 수직 디코더가 사용하여 제2 잔상 영상 생성을 할 수 있다.

[0053] 본 실시예와 같이 디코더를 수평 디코더 및 수직 디코더로 두 갈래로 분할하여 수직 및 수평 방향으로 잔상을 제거하는 기술을 적용을 하였을 때, 파라미터적으로 연산량에 있어서는 유의미한 변동은 없다고 볼 수 있다. 왜냐하면 기본적으로 디코더를 두 개로 분할하게 되면 딥러닝에 있어서 통상적으로 컨볼루션(convolution) 커널에 웨이트들이 증가하기 때문에 파라미터가 증가하기 마련인데, 본 실시예의 방법에서는 수직 디코더의 제2 컨볼루션 커널이 전술한 바와 같이 수평 디코더의 제1 컨볼루션 커널의 반시계 방향 회전에 의해 생성되므로, 수평 방향의 학습에서의 파라미터 또는 웨이트(가중치)를 수직 방향의 학습에도 사용(웨이트를 셰어)한다. 즉, 컨볼루션 커널을 재사용 내지 쉐어링을 하므로 파라미터의 증가가 없는 것과 마찬가지의 효과를 얻게 된다.

[0054] 본 실시예와 다르게, 두 개의 디코더를 구비하고 전술된 커널 셰어링을 통하지 않는 경우나, 단일 디코더 구조에서는 이러한 디코더 자체가 앞서 인코더에서 만들어 추출한 이 특징을 사용하여 한번에 잔차를 학습해야 한다. 본 실시예에서는 이 잔차를 가로 방향(수평 방향)과 세로 방향(수직 방향)으로 분할하여 각각 문제를 해결해 나가서 최종적으로 블러 제거라는 큰 문제를 해결하는 효과를 달성한다.

[0055] 이와 같은 본 실시예의 수직-수평 분할을 통한 단일 영상 잔상 제거 방법 및 이를 위한 네트워크 시스템을 기존의 잔상제거 네트워크에 적용하여 시험하였다. 기존에 잔상을 제거하는 U-shape network 네트워크 중 대표적으로 세가지 종류, 즉 PSS-NSC, DMPHN, MSCNN에 대해 적용하였다.

[0056] 도 5는 GoPro 데이터셋에서의 잔상제거 결과를 나타낸다. 도 6는 RealBlur 데이터셋에서의 잔상제거 결과를 나타낸다.

[0057] 도 5 및 도 6에서 가로축은 잔상제거 네트워크 시스템의 학습에 사용된 파라미터의 개수를 나타내고, 세로축은 신호대잡음비(PSNR)를 나타낸다. PSNR은 클수록 원본과 유사한 특징을 보인다. 각각의 그래프에서 원의 반지름은 GMAC을 의미한다. 본 실시예의 수직-수평 분할을 통한 단일 영상 잔상 제거 방법이 이 적용된 경우($PSS\text{-}NSC^{Ours}$, $DMPHN^{Ours}$, $MSCNN^{Ours}$)는 전술한 바와 같이, 파라미터를 재사용 하였기 때문에 파라미터 증가 없이, 효과적으로 성능을 향상 시키고 있다. 또한 연산량을 의미하는 GMAC 대비 성능 향상 폭 또한 매우 높다. 수평 디코더 및 수직 디코더 2개를 사용하는 경우 당연히 수직 방향의 특징(features) 및 수평 방향의 특징을 각각 추출해야 한다. 그러나, 본 실시예의 방법은 전술한 바와 같이, 수평 디코더 및 수직 디코더가 컨볼루션 커널을 회전시켜서 웨이트를 셰어하도록 사용을 하고 있다. 그렇기 때문에 파라미터적으로는 전혀 증가가 없다. 다만, 디코더를 하나로 푸는 것이 아니라 두 개로 잔상제거 문제를 풀게 되므로, 도 5 및 도 6에서 원의 지름으로 표현된 맥(GMAC)이라고 하는 계산양적인 측면에서는 매우 조금 증가하는 것을 볼 수 있다. PSNR은 얼마나 원본 영상에 가깝게 복원했는지를 나타내는 지표이다. 본 발명의 실시예에 의한 방법의 경우 계산량이 매우 조금 증가하지만 PSNR이 크게 증가하는 것을 볼 수 있다. 즉, 수직 디코더가 수평 디코더의 컨볼루션 커널을 회전시켜 자신의 컨볼루션 커널로 사용함으로써 파라미터 셰어링을 통해서 파라미터 증가는 없고, 디코더를 수평 및 수직으로 분할하기 때문에 약간의 계산량이 증가했지만 그에 비해 성능이 크게 향상되는 것을 확인할 수 있다.

[0058] 도 7은 GoPro 데이터셋 및 RealBlur 데이터셋에서 잔상제거 결과를 비교하는 표이다.

[0059] 도 7을 참조하면, 공평한 실험을 위해서 채널을 올린 것과 레이어를 올린 것과, 본 실시예의 방법의 실험 결과 데이터가 제시되어 있다. 매개변수(Params)의 수는 백만 단위로 측정되었다. GMAC는 720P(1280 X 720)의 입력 해상도로 추정된다. VRAM 사용량은 GB 단위로 측정되며 입력 크기는 256 X 256이다. 기준선 방법(Baseline)에 비해 증가한 PSNR과 SSIM 값은 각각 빨간색과 파란색으로 표시되었다. 표를 확인하면, 최종적으로 본 실시예의 방법, 즉 디코더를 수평 및 수직으로 분할하고, 커널 및 파라미터를 셰어(share)하는 알고리즘이 가장 성능이 좋은 것

을 확인할 수 있다.

[0060] 도 8 및 도 9는 GoPro 데이터셋에서의 결과영상을 나타낸다.

[0061] 도 8 및 도 9에는 GOPRO 데이터 세트로 본 실시예의 방법이 적용된 경우(W/Ours), 적용되지 않는 경우(W/o, Ours)에 대해 MSCNN, DMPHN, PSS-NSC에 대한 실험 결과를 보여준다. 도 8 및 도 9에서 빨간색 박스 부분이 확대된 영역이다. 상기 다양한 네트워크에서 본 실시예의 방법을 적용했을 때와 아닐 때를 위아래로 비교하면 본 실시예의 방법이 적용된 경우에 훨씬 더 나은 성능을 나타냄을 확인할 수 있다. r1과 r2은 각각 수직 방향과 수평 방향의 특징의 잔차 성분을 의미한다.

[0062] 본 발명의 실시예에 의하면, 한 디코더에서 다른 디코더로 회전된 콘볼루션 커널을 공유하므로 표준 U-Net 아키텍처와 동일한 수의 매개변수를 사용하면서도 디블러링 성능을 크게 향상시킨다.

[0063] 본 발명의 실시예와 다르게, 각각의 디코더를 독립적인 파라미터를 갖는 구조로 잔상 제거를 학습하게 되면 문제를 해결하는 과정에서 영상의 전체적인 색상이 변화하는 현상이 발생할 수 있다. 본 발명에서는 하나의 디코더의 커널을 회전시켜 생성된 커널을 다른 디코더가 사용함으로써, 디코더의 파라미터를 재사용하여 이러한 문제를 제거하고 또한 네트워크 전체적인 파라미터 증가를 줄인다.

[0064] 영상에서 잔상을 제거하는 것은 매우 고전적이지만 영상 취득을 위해 매우 핵심적인 기술로 판단된다. 기본적으로 스마트폰, 카메라의 ISP (ISP)에 탑재되어 그 취득되는 영상의 품질을 크게 향상시킬 수 있다. 따라서 본 발명의 실시예들은 기본적으로 카메라를 활용하는 모든 분야에 적용 가능하다.

[0065] 전술한 본 발명의 설명은 예시를 위한 것이며, 본 발명이 속하는 기술분야의 통상의 지식을 가진 자는 본 발명의 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.

[0066] 본 발명의 범위는 후술하는 특허청구범위에 의하여 나타내어지며, 특허청구범위의 의미 및 범위 그리고 그 균등 개념으로부터 도출되는 모든 변경 또는 변형된 형태가 본 발명의 범위에 포함되는 것으로 해석되어야 한다.

부호의 설명

[0068] 입력부 : 10

인코더 : 20

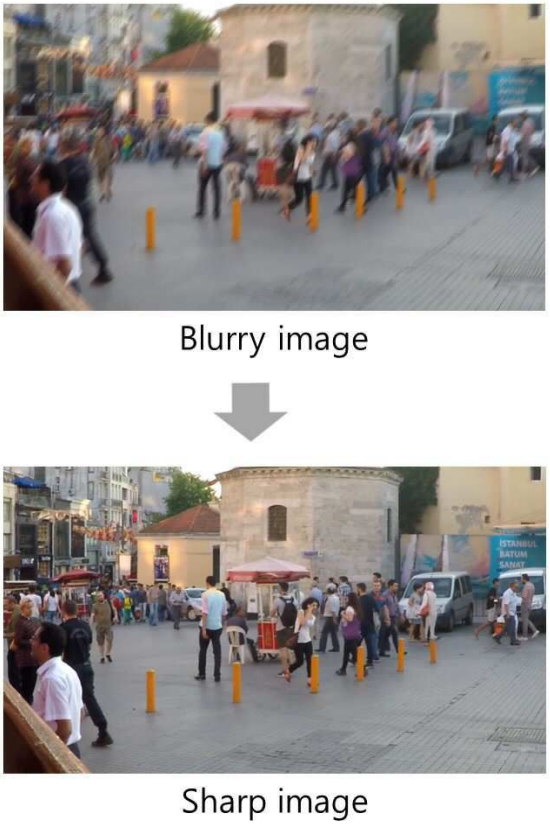
수평 디코더 : 30

수직 디코더 :40

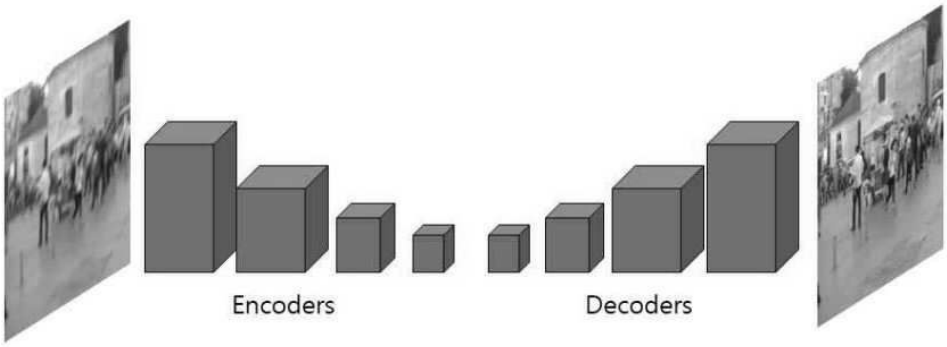
합성부 : 50

도면

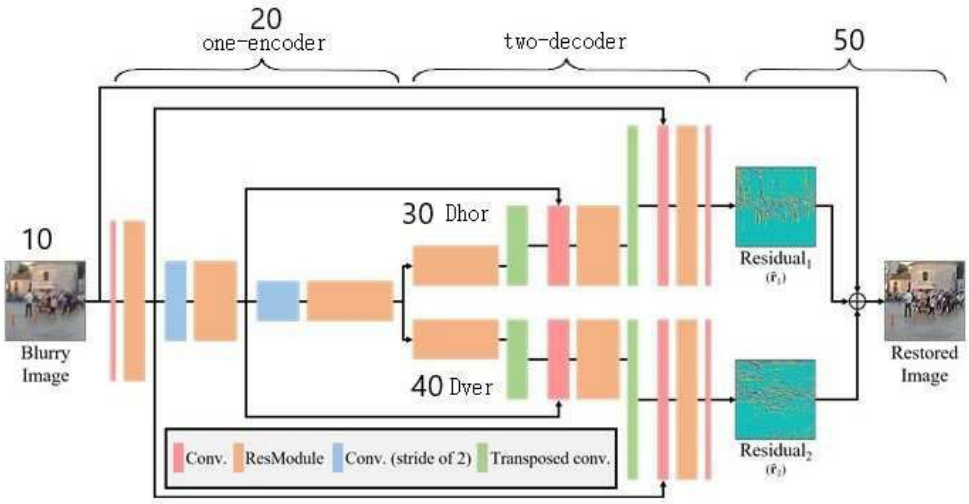
도면1



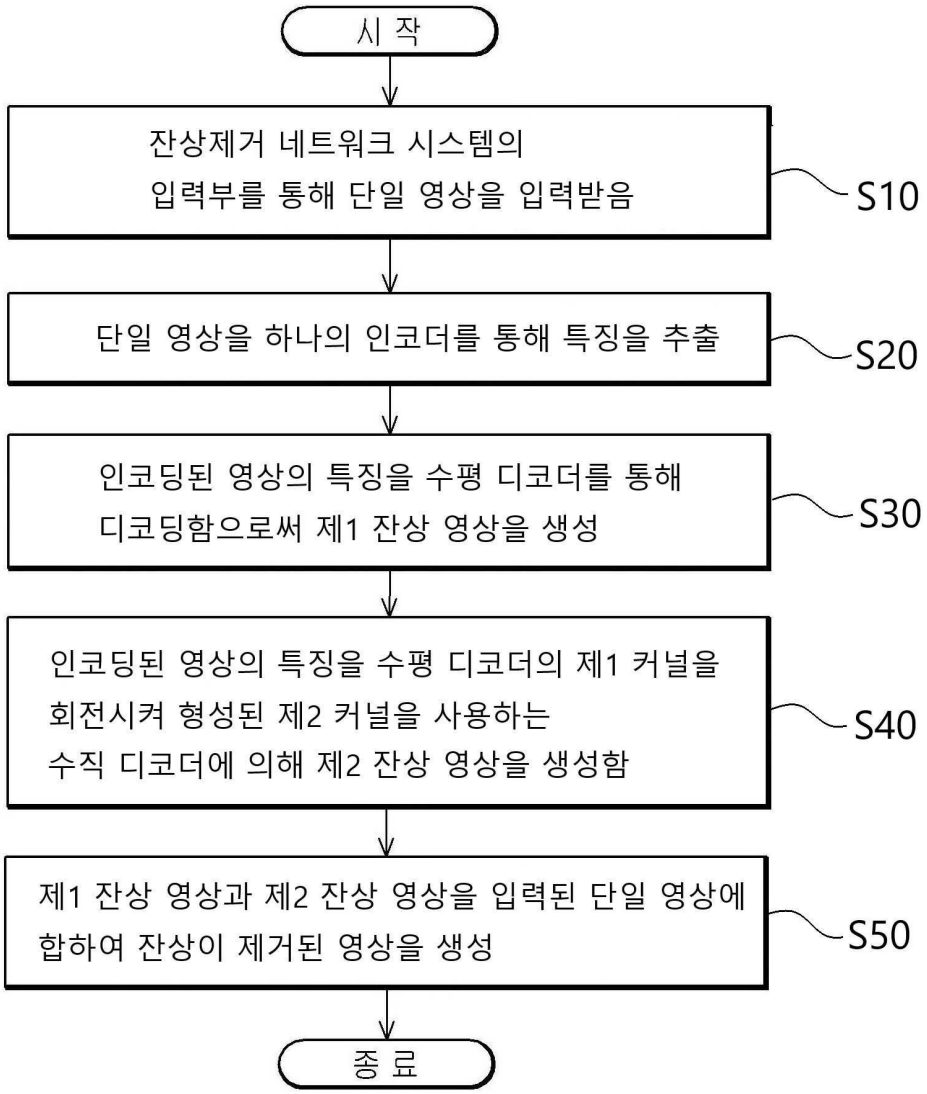
도면2



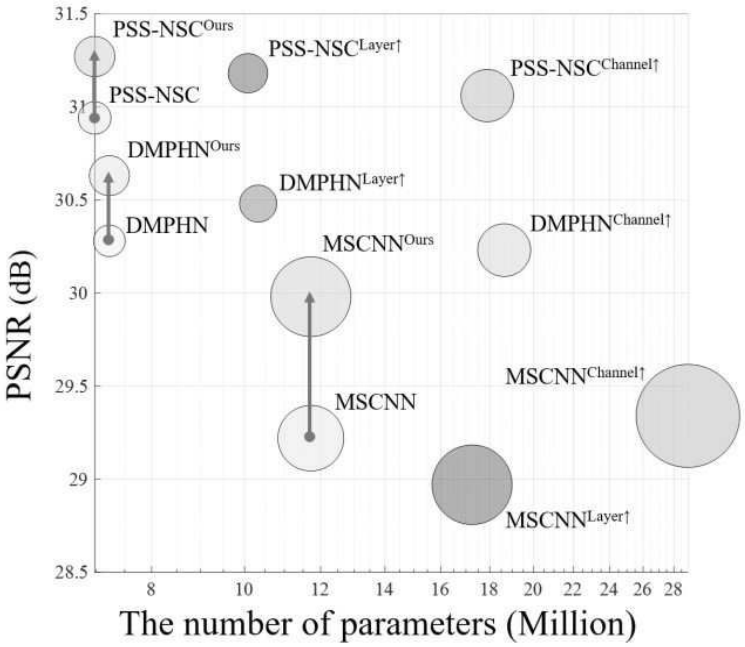
도면3


20
one-encoder
two-decoder
50
10
Blurry Image
30 Dhor
40 Dver
Residual1
Residual2
Restored Image
Conv.
ResModule
Conv. (stride of 2)
Transposed conv.

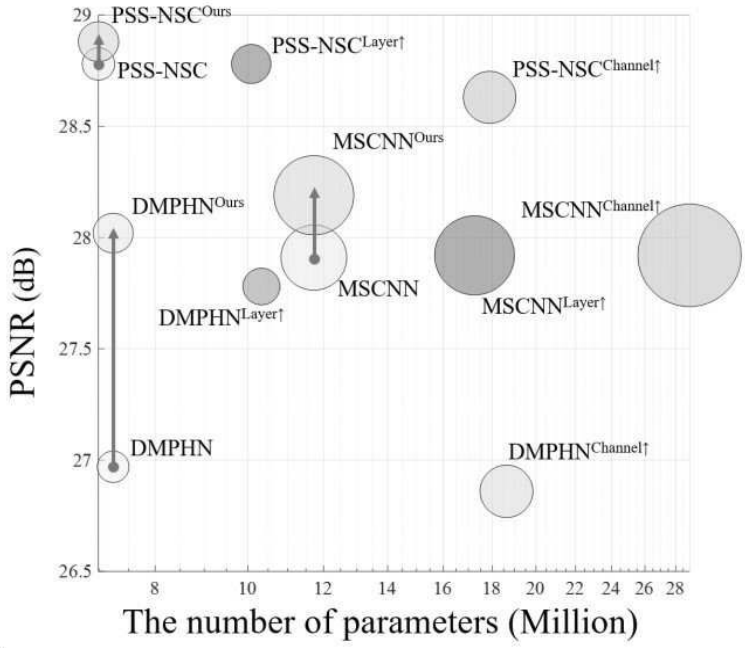
도면4


시 작
잔상제거 네트워크 시스템의
입력부를 통해 단일 영상을 입력받음
S10
단일 영상을 하나의 인코더를 통해 특징을 추출
S20
인코딩된 영상의 특징을 수평 디코더를 통해
디코딩함으로써 제1 잔상 영상을 생성
S30
인코딩된 영상의 특징을 수평 디코더의 제1 커널을
회전시켜 형성된 제2 커널을 사용하는
수직 디코더에 의해 제2 잔상 영상을 생성함
S40
제1 잔상 영상과 제2 잔상 영상을 입력된 단일 영상에
합하여 잔상이 제거된 영상을 생성
S50
종 료

도면5


PSS-NSC^Ours
PSS-NSC^Layer↑
PSS-NSC^Channel↑
PSS-NSC
DMPHN^Ours
DMPHN^Layer↑
DMPHN
DMPHN^Channel↑
MSCNN^Ours
MSCNN^Channel↑
MSCNN
MSCNN^Layer↑
PSNR (dB)
The number of parameters (Million)

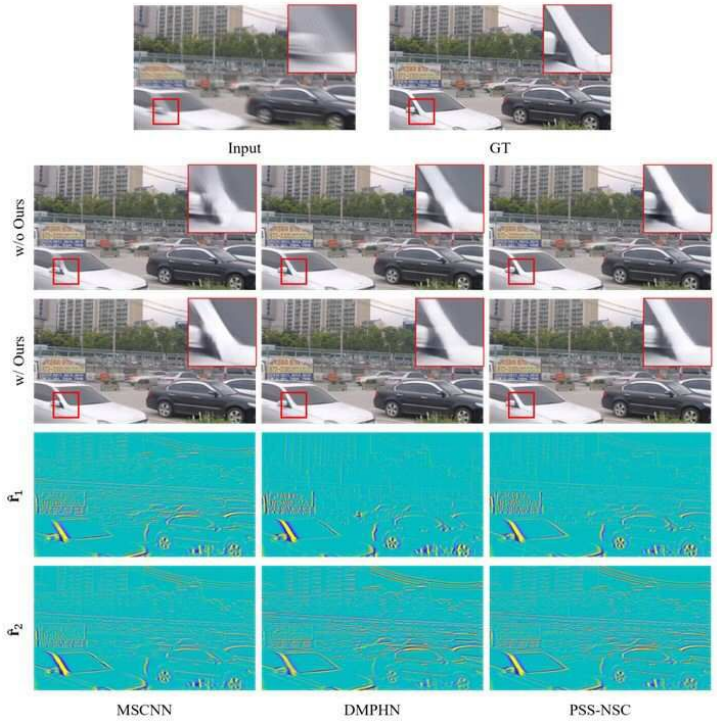
도면6


PSS-NSC^Ours
PSS-NSC^Layer↑
PSS-NSC^Channel↑
PSS-NSC
MSCNN^Ours
DMPHN^Ours
MSCNN^Channel↑
MSCNN
DMPHN^Layer↑
MSCNN^Layer↑
DMPHN
DMPHN^Channel↑
PSNR (dB)
The number of parameters (Million)

도면7

Method	Variations	GoPro		RealBlur		Complexity†		
		PSNR	SSIM	PSNR	SSIM	Params	GMACs	VRAM
PSS-NSC	Baseline*	$30.94_{(30.92)}$	0.9494	28.78	0.8716	6.98	1167.24	3.79
	Channel↑	31.06	0.9509	28.63	0.8667	17.89	3020.81	5.97
	Layer↑	31.18	0.9524	28.78	0.8721	10.08	1703.95	5.55
	Ours	31.27	0.9531	28.88	0.8765	6.98	1784.39	5.68
DMPHN	Baseline*	$30.28_{(30.25)}$	0.9408	26.97	0.8170	7.23	1100.18	1.62
	Channel↑	30.23	0.9414	26.86	0.8201	18.63	3060.28	2.52
	Layer↑	30.48	0.9443	27.78	0.8366	10.33	1509.11	2.30
	Ours	30.63	0.9458	28.02	0.8459	7.23	1754.07	2.41
MSCNN	Baseline*	$29.22_{(29.08)}$	0.9273	27.91	0.8442	11.72	4728.69	3.21
	Channel↑	29.34	0.9292	27.92	0.8432	28.94	11673.07	4.84
	Layer↑	28.97	0.9250	27.92	0.8449	17.25	6960.32	4.71
	Ours	29.98	0.9382	28.19	0.8550	11.72	6966.13	4.78

도면8



도면9

