



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2020년12월30일
(11) 등록번호 10-2197089
(24) 등록일자 2020년12월23일

(51) 국제특허분류(Int. Cl.)
G06T 5/00 (2019.01) G06N 20/00 (2019.01)
G06T 7/215 (2017.01)
(52) CPC특허분류
G06T 5/002 (2013.01)
G06N 20/00 (2019.01)
(21) 출원번호 10-2020-0121985
(22) 출원일자 2020년09월22일
심사청구일자 2020년09월22일
(56) 선행기술조사문헌
KR1020200096026 A*
KR100860967 B1
KR101152525 B1
KR101350460 B1
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
엘아이지넥스원 주식회사
경기도 용인시 기흥구 마북로 207 (마북동)
연세대학교 산학협력단
서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)
(72) 발명자
정형주
서울특별시 서대문구 연세로 50, 연세대학교 (신촌동)
손광훈
서울특별시 서대문구 연세로 50, 연세대학교 (신촌동)
(뒷면에 계속)
(74) 대리인
특허법인우인

전체 청구항 수 : 총 10 항

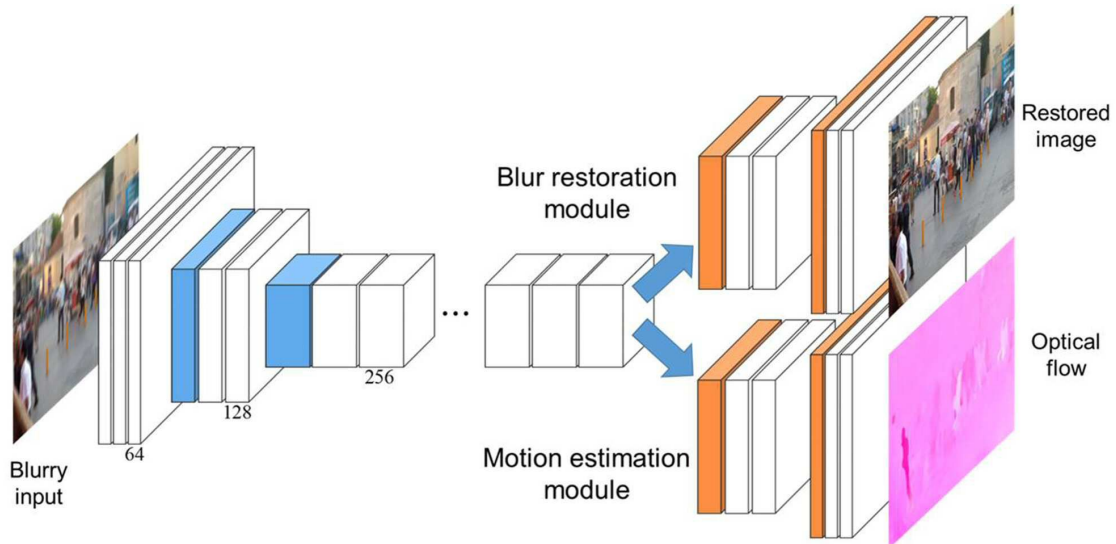
심사관 : 김광식

(54) 발명의 명칭 단일 영상 비균일 블러 제거를 위한 딥러닝 기반 다중 학습 구조

(57) 요약

본 실시예들은 인코더-디코더 구조로 설계된 블러 개선 네트워크 및 모션 예측 네트워크가 인코더를 공유하고 동시 학습을 거친 다중 학습 네트워크 모델을 통해 단일 영상에서 비균일 블러를 제거할 수 있는 블러 개선 방법 및 장치를 제공한다.

대표도 - 도5



(52) CPC특허분류

G06T 7/215 (2017.01)

G06T 2207/20081 (2013.01)

(72) 발명자

장현성

경기도 성남시 분당구 판교로 333 (삼평동)

하남구

경기도 성남시 분당구 판교로 333 (삼평동)

권구용

경기도 성남시 분당구 판교로 333 (삼평동)

이민석

경기도 성남시 분당구 판교로 333 (삼평동)

공지예외적용 : 있음

명세서

청구범위

청구항 1

블러 개선 장치에 의한 블러 개선 방법에 있어서,

단일 블러 이미지를 입력받는 단계; 및

상기 입력된 단일 블러 이미지로부터 다중 학습 네트워크 모델을 통해 비균일 블러를 제거하는 단계를 포함하며,

상기 다중 학습 네트워크 모델의 블러 개선 네트워크 및 모션 예측 네트워크는 인코더를 상호 공유하며,

상기 블러 개선 네트워크 및 상기 모션 예측 네트워크는 동시에 학습되는 것을 특징으로 하는 블러 개선 방법.

청구항 2

제1항에 있어서,

상기 인코더는 컨볼루션 레이어(Convolutional Layer)가 적용된 모션 기반의 특징 추출 모듈인 것을 특징으로 하는 블러 개선 방법.

청구항 3

제2항에 있어서,

상기 인코더는 상기 컨볼루션 레이어를 통해 표현자를 추출하고, 스킵 구조를 갖는 레지듀얼 블록을 통해 상기 표현자의 채널 개수를 증가시키는 것을 특징으로 하는 블러 개선 방법.

청구항 4

제1항에 있어서,

상기 블러 개선 네트워크는 상기 인코더와 제1 디코더를 포함하며, 상기 제1 디코더는 상기 인코더에서 추출한 특징으로부터 이미지를 복원하는 것을 특징으로 하는 블러 개선 방법.

청구항 5

제4항에 있어서,

상기 제1 디코더는 디컨볼루션 레이어(Deconvolutional Layer)를 통해 해상도를 증가시키고, 상기 디컨볼루션 레이어의 후단에 제1 컨볼루션 레이어를 추가하고, 상기 제1 컨볼루션 레이어에서 표현자의 채널 개수를 상기 인코더와 대칭이 되도록 감소시키고, 상기 제1 컨볼루션 레이어는 블러 개선된 이미지를 추정하며,

대칭이 되는 상기 인코더와 상기 제1 디코더 간에 스킵 연결을 통해 상기 표현자의 공간적 구조를 유지하는 것을 특징으로 하는 블러 개선 방법.

청구항 6

제1항에 있어서,

상기 모션 예측 네트워크는 상기 인코더와 제2 디코더를 포함하며, 상기 제2 디코더는 상기 인코더에서 추출한 특징으로부터 옵티컬 플로우를 추정하는 것을 특징으로 하는 블러 개선 방법.

청구항 7

제6항에 있어서,

상기 제2 디코더는 디컨볼루션 레이어를 통해 상기 옵티컬 플로우를 추정하고, 상기 디컨볼루션 레이어의 후단에 제2 컨볼루션 레이어를 추가하고, 상기 제2 컨볼루션 레이어에서 표현자의 채널 개수를 상기 인코더와 대칭

이 되도록 감소시키고, 상기 제2 컨볼루션 레이어는 전방 업티컬 플로우와 후방 업티컬 플로우를 추정하며, 대칭이 되는 상기 인코더와 상기 제2 디코더 간에 스킵 연결을 통해 상기 표현자의 공간적 구조를 유지하는 것을 특징으로 하는 블러 개선 방법.

청구항 8

제6항에 있어서,

상기 모션 예측 네트워크는 학습 과정에서 업티컬 플로우에 대한 슈도 정답을 적용하는 것을 특징으로 하는 블러 개선 방법.

청구항 9

제8항에 있어서,

상기 업티컬 플로우에 대한 슈도 정답은 공간 피라미드 배합과 심층 학습을 결합한 업티컬 플로우 추정 모델에 의해 산출되는 것을 특징으로 하는 블러 개선 방법.

청구항 10

제9항에 있어서,

상기 업티컬 플로우 추정 모델은 비교할 이미지 쌍을 스택한 후 컨볼루션 레이어를 거쳐 추정하거나 비교할 이미지 쌍을 각각의 컨볼루션 레이어를 거친 후 서로의 입력을 이용하는 상관 레이어(Correlation Layer)를 통해 특징을 비교하는 것을 특징으로 하는 블러 개선 방법.

청구항 11

삭제

발명의 설명

기술 분야

[0001] 본 발명이 속하는 기술 분야는 딥러닝 기반의 다중 학습 구조를 이용한 단일 영상 비균일 블러 개선 방법 및 장치에 관한 것이다.

배경 기술

[0002] 이 부분에 기술된 내용은 단순히 본 실시예에 대한 배경 정보를 제공할 뿐 종래기술을 구성하는 것은 아니다.

[0003] 카메라의 센서는 장면에 대한 정보를 얻기 위해 빛을 축적하며, 이때 충분한 노출 시간이 필요하다. 그 결과, 노출 시간 중의 카메라/물체의 움직임은 사진에서 모션 블러를 발생시킨다. 카메라/물체의 움직임에 의한 모션 블러는 사진의 가시성을 떨어뜨리며, 이는 다양한 물체 인식 분야의 성능에 영향을 준다.

[0004] 모션 블러에 의해 열화된 영상을 복원하는 것은 가시적으로 향상된 영상을 얻을 수 있으며, 다양한 객체 탐지 분야에서 매우 필수적인 전처리 과정이다.

선행기술문헌

특허문헌

- [0005] (특허문헌 0001) 한국등록특허공보 제10-0860967호 (2008.09.24.)
- (특허문헌 0002) 한국등록특허공보 제10-1152525호 (2012.05.25.)
- (특허문헌 0003) 한국등록특허공보 제10-1350460호 (2014.01.03.)
- (특허문헌 0004) 한국등록특허공보 제10-1844332호 (2018.03.27.)

발명의 내용

해결하려는 과제

- [0006] 본 발명의 실시예들은 인코더-디코더 구조로 설계된 블러 개선 네트워크 및 모션 예측 네트워크가 인코더를 공유하고 동시 학습을 거친 다중 학습 네트워크 모델을 통해 단일 영상에서 비균일 블러를 제거하는 데 주된 목적이 있다.
- [0007] 본 발명의 명시되지 않은 또 다른 목적들은 하기의 상세한 설명 및 그 효과로부터 용이하게 추론할 수 있는 범위 내에서 추가적으로 고려될 수 있다.

과제의 해결 수단

- [0008] 본 실시예의 일 측면에 의하면 컴퓨팅 디바이스에 의한 블러 개선 방법에 있어서, 단일 블러 이미지를 입력받는 단계, 및 상기 입력된 단일 블러 이미지로부터 다중 학습 네트워크 모델을 통해 비균일 블러를 제거하는 단계를 포함하며, 상기 다중 학습 네트워크 모델의 블러 개선 네트워크 및 모션 예측 네트워크는 인코더를 상호 공유하는 것을 특징으로 하는 블러 개선 방법을 제공한다.
- [0009] 상기 인코더는 컨볼루션 레이어(Convolutional Layer)가 적용된 모션 기반의 특징 추출 모듈일 수 있다.
- [0010] 상기 인코더는 컨볼루션 레이어(Convolutional Layer)를 통해 표현자를 추출하고, 스킵 구조를 갖는 레지듀얼 블록을 통해 상기 표현자의 채널 개수를 증가시킬 수 있다.
- [0011] 상기 블러 개선 네트워크는 상기 인코더와 제1 디코더를 포함하며, 상기 제1 디코더는 상기 인코더에서 추출한 특징으로부터 이미지를 복원할 수 있다.
- [0012] 상기 제1 디코더는 디컨볼루션 레이어(Deconvolutional Layer)를 통해 해상도를 증가시키고, 상기 디컨볼루션 레이어의 후단에 제1 컨볼루션 레이어를 추가하고, 상기 제1 컨볼루션 레이어에서 표현자의 채널 개수를 상기 인코더와 대칭이 되도록 감소시키고, 상기 제1 컨볼루션 레이어는 블러 개선된 이미지를 추정하며, 대칭이 되는 상기 인코더와 상기 제1 디코더 간에 스킵 연결을 통해 상기 표현자의 공간적 구조를 유지할 수 있다.
- [0013] 상기 모션 예측 네트워크는 상기 인코더와 제2 디코더를 포함하며, 상기 제2 디코더는 상기 인코더에서 추출한 특징으로부터 옵티컬 플로우를 추정할 수 있다.
- [0014] 상기 제2 디코더는 디컨볼루션 레이어를 통해 상기 옵티컬 플로우를 추정하고, 상기 디컨볼루션 레이어의 후단에 제2 컨볼루션 레이어를 추가하고, 상기 제2 컨볼루션 레이어에서 표현자의 채널 개수를 상기 인코더와 대칭이 되도록 감소시키고, 상기 제2 컨볼루션 레이어는 전방 옵티컬 플로우와 후방 옵티컬 플로우를 추정하며, 대칭이 되는 상기 인코더와 상기 제2 디코더 간에 스킵 연결을 통해 상기 표현자의 공간적 구조를 유지할 수 있다.
- [0015] 상기 모션 예측 네트워크는 학습 과정에서 옵티컬 플로우에 대한 슈도 정답을 적용할 수 있다.
- [0016] 상기 옵티컬 플로우에 대한 슈도 정답은 공간 피라미드 배합과 심층 학습을 결합한 옵티컬 플로우 추정 모델에 의해 산출될 수 있다.
- [0017] 상기 옵티컬 플로우 추정 모델은 비교할 이미지 쌍을 스택한 후 컨볼루션 레이어를 거쳐 추정하거나 비교할 이미지 쌍을 각각의 컨볼루션 레이어를 거친 후 서로의 입력을 이용하는 상관 레이어(Correlation Layer)를 통해 특징을 비교할 수 있다.
- [0018] 상기 블러 개선 네트워크 및 상기 모션 예측 네트워크는 동시에 학습될 수 있다.

발명의 효과

- [0019] 이상에서 설명한 바와 같이 본 발명의 실시예들에 의하면, 인코더-디코더 구조로 설계된 블러 개선 네트워크 및 모션 예측 네트워크가 인코더를 공유하고 동시 학습을 거친 다중 학습 네트워크 모델을 통해 단일 영상에서 비균일 블러를 제거할 수 있는 효과가 있다.
- [0020] 여기에서 명시적으로 언급되지 않은 효과라 하더라도, 본 발명의 기술적 특징에 의해 기대되는 이하의 명세서에서 기재된 효과 및 그 잠정적인 효과는 본 발명의 명세서에 기재된 것과 같이 취급된다.

도면의 간단한 설명

- [0021] 도 1 및 도 2는 기존의 블러 개선 모델을 예시한 도면이다.
- 도 3은 본 발명의 일 실시예에 따른 블러 개선 장치를 예시한 블록도이다.
- 도 4 및 도 5는 본 발명의 일 실시예에 따른 블러 개선 장치의 다중 학습 네트워크 모델을 예시한 도면이다.
- 도 6은 본 발명의 일 실시예에 따른 블러 개선 장치의 다중 학습 네트워크 모델이 처리하는 입력 이미지, 모션 정보, 및 복원 이미지를 예시한 도면이다.
- 도 7은 본 발명의 다른 실시예에 따른 블러 개선 방법을 예시한 흐름도이다.

발명을 실시하기 위한 구체적인 내용

- [0022] 이하, 본 발명을 설명함에 있어서 관련된 공지기능에 대하여 이 분야의 기술자에게 자명한 사항으로서 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명을 생략하고, 본 발명의 일부 실시예들을 예시적인 도면을 통해 상세하게 설명한다.
- [0023] 본 발명은 표적 탐지, 군 기반 시설 감시 체계 등에 적용할 수 있다.
- [0024] 도 1는 기존의 블러 커널 기반의 블러 개선 모델이고, 도 2는 기존의 블러 커널을 사용하지 않는 블러 개선 모델을 예시한 도면이다.
- [0025] 도 1에 도시된 기존의 방식은 블러 커널을 추정한 후 논-블라인드(Non-Blind) 디블러를 통해 블러 제거된 영상을 획득한다. 영상의 모든 픽셀에 대한 블러 커널을 찾는 것은 잘 정의되지 않은 문제(ill-posed problem)에 해당한다. ill-posed 문제를 해결하기 위하여 블러 커널에 다양한 가정을 설정하고, 블러 커널 추정 과정에서 카메라와 객체의 움직임 정보를 한정하기 위해 다양한 물리적 모델을 사용한다. 수작업을 사용하므로 정확한 커널 추정이 곤란한 문제가 있다.
- [0026] 도 2에 도시된 기존의 방식은 커널 추정없이 바로 블러 영상으로부터 개선 영상의 고차원의 비선형 매핑 관계를 딥러닝 기술을 통해 학습한다. 학습 네트워크는 움직임 정보에 대해서 블러 영상과 정답 영상 간의 고차원의 비선형 매핑 관계를 통해 학습한다. 이렇게 학습된 네트워크의 표현자(feature)는 의미론적으로 움직임 정보를 포함할 수 있게 하는 직접적인 학습 과정이 없다.
- [0027] 본 실시예에 따른 블러 개선 장치는 입력 정답 간의 매핑을 통한 움직임 정보를 학습할 뿐만 아니라, 움직임 추정 네트워크를 추가로 학습하는 다중 학습 구조를 통해 분명하게(explicit) 움직임 정보를 학습한다. 이는 동일한 네트워크를 사용하였을 때 기존의 딥러닝 방식보다 실행시간(runtime)의 손실없이 더욱 향상된 블러 제거 성능을 확보할 수 있다. 기존의 커널 기반의 방식과 달리 테스트 과정에서 정확한 블러 정보(즉 움직임 정보)를 추정할 필요가 없으며, 복잡한 반복 연산을 필요로 하지 않는다.
- [0028] 도 3은 본 발명의 일 실시예에 따른 블러 개선 장치를 예시한 블록도이다.
- [0029] 블러 개선 장치(110)는 적어도 하나의 프로세서(120), 컴퓨터 판독 가능한 저장매체(130) 및 통신 버스(170)를 포함한다.
- [0030] 프로세서(120)는 블러 개선 장치(110)로 동작하도록 제어할 수 있다. 예컨대, 프로세서(120)는 컴퓨터 판독 가능한 저장 매체(130)에 저장된 하나 이상의 프로그램들을 실행할 수 있다. 하나 이상의 프로그램들은 하나 이상의 컴퓨터 실행 가능 명령어를 포함할 수 있으며, 컴퓨터 실행 가능 명령어는 프로세서(120)에 의해 실행되는 경우 블러 개선 장치(110)로 하여금 예시적인 실시예에 따른 동작들을 수행하도록 구성될 수 있다.
- [0031] 컴퓨터 판독 가능한 저장 매체(130)는 컴퓨터 실행 가능 명령어 내지 프로그램 코드, 프로그램 데이터 및/또는 다른 적합한 형태의 정보를 저장하도록 구성된다. 컴퓨터 판독 가능한 저장 매체(130)에 저장된 프로그램(140)은 프로세서(120)에 의해 실행 가능한 명령어의 집합을 포함한다. 일 실시예에서, 컴퓨터 판독 가능한 저장 매체(130)는 메모리(랜덤 액세스 메모리와 같은 휘발성 메모리, 비휘발성 메모리, 또는 이들의 적절한 조합), 하나 이상의 자기 디스크 저장 디바이스들, 광학 디스크 저장 디바이스들, 플래시 메모리 디바이스들, 그 밖에 블러 개선 장치(110)에 의해 액세스되고 원하는 정보를 저장할 수 있는 다른 형태의 저장 매체, 또는 이들의 적합한 조합일 수 있다.
- [0032] 통신 버스(170)는 프로세서(120), 컴퓨터 판독 가능한 저장 매체(140)를 포함하여 블러 개선 장치(110)의 다른

다양한 컴포넌트들을 상호 연결한다.

- [0033] 블러 개선 장치(110)는 또한 하나 이상의 입출력 장치를 위한 인터페이스를 제공하는 하나 이상의 입출력 인터페이스(150) 및 하나 이상의 통신 인터페이스(160)를 포함할 수 있다. 입출력 인터페이스(150) 및 통신 인터페이스(160)는 통신 버스(170)에 연결된다. 입출력 장치(미도시)는 입출력 인터페이스(150)를 통해 블러 개선 장치(110)의 다른 컴포넌트들에 연결될 수 있다.
- [0034] 도 4 및 도 5는 본 발명의 일 실시예에 따른 블러 개선 장치의 다중 학습 네트워크 모델을 예시한 도면이다.
- [0035] 다중 학습 네트워크 모델은 복수의 레이어가 연결된 학습 네트워크 구조이며, 인코더-디코더 구조를 갖는다. 학습 네트워크 모델은 다수의 레이어가 네트워크로 연결되며 히든 레이어를 포함한다. 레이어는 파라미터를 포함할 수 있고, 레이어의 파라미터는 학습가능한 필터 집합을 포함한다. 파라미터는 노드 간의 가중치 및/또는 바이어스를 포함한다.
- [0036] 다중 학습 네트워크 모델은 블러 개선 네트워크 및 모션 예측 네트워크를 포함한다. 두 네트워크 모두 넓은 범위의 영역을 참조하기 위해 스트라이드 컨볼루션을 통한 인코더-디코더 구조로 설계되며 서로 일부 네트워크 파라미터를 공유한다.
- [0037] 블러 개선 네트워크는 먼저 단일 입력 영상으로부터 세 개의 컨볼루션 레이어를 통해 64 채널의 표현자를 추출한다. 그 후, 스트라이드 컨볼루션을 포함하는 두 개의 레지듀얼 블록을 통해 공간적 크기를 반으로 줄이면서 동시에 표현자 채널의 수를 2배씩 증가시킨다.
- [0038] 레지듀얼 블록은 네트워크 학습 시 그라디언트가 사라지는 문제를 해결할 수 있어 더욱 좋은 표현자를 학습할 수 있는 장점이 있다. 레지듀얼 블록은 세 개의 컨볼루션 레이어로 구성되며, 첫 번째 레이어는 스트라이드(stride)가 2이다. 또한, 레지듀얼 블록의 결과는 마지막 컨볼루션 레이어의 결과와 첫 번째 레이어의 결과를 합하여 구한다. 1/4 크기로 줄여진 표현자 (채널은 256)는 그 후 세 개의 레지듀얼 블록을 통과하며, 이때 모든 컨볼루션 레이어의 스트라이드는 1이다.
- [0039] 마지막으로 입력 영상과 동일한 해상도를 갖는 개선 영상을 추정하기 위해 두 개의 디컨볼루션 레이어를 사용하여 해상도를 2배씩 증가시킨다. 각 디컨볼루션 레이어 이후 한 개의 컨볼루션 레이어를 추가로 설계하였으며, 표현자의 채널 수는 인코더와 대칭이 되도록 반으로 감소시켰다. 마지막 컨볼루션 레이어는 세 채널의 개선 영상을 바로 추정한다. 한편 대칭이 되는 인코더와 디코더 표현자 간 스킵 연결을 통해 앞 단의 표현자의 공간적 구조를 유지시킨다.
- [0040] 네트워크의 비선형성을 높이기 위한 비선형 함수로 ReLU(Rectified linear unit) 레이어를 사용한다. 또한, 컨볼루션 표현자의 내부 코베리언스 변화(Internal Covariate Shift) 문제를 해결하기 위한 Batch normalization 레이어를 사용할 수 있다.
- [0041] 다중 학습 구조의 다른 네트워크인 모션 예측 네트워크(옵티컬 플로우 추정 네트워크)는 블러 개선 네트워크의 인코더 구조를 공유한다. 옵티컬 플로우 추정을 위한 추가적인 네트워크는 디코더 부분의 디컨볼루션 레이어와 컨볼루션 레이어이다. 이는 블러 개선 네트워크의 디코더 부분과 유사하게 설계할 수 있다.
- [0042] 모션 예측 네트워크의 마지막 컨볼루션은 네 채널의 옵티컬 플로우를 바로 추정한다. 각각의 두 채널짜리 플로우는 각각 전방 옵티컬 플로우와 후방 옵티컬 플로우를 나타낸다.
- [0043] 다중 학습 네트워크 모델은 학습 과정에서 제1 피드백 경로를 통해 블러 개선 네트워크의 제1 디코더의 출력을 모션 기반의 특징 추출 모듈인 인코더의 입력으로 변환하여 피드백할 수 있다. 변환 모듈은 출력 및 입력의 포맷에 맞게 데이터 값을 변환하며 제1 피드백 경로 상에서 제1 변환 모듈을 통해 데이터 값을 변환한다.
- [0044] 다중 학습 네트워크 모델은 학습 과정에서 제2 피드백 경로를 통해 모션 예측 네트워크의 제2 디코더의 출력을 모션 기반의 특징 추출 모듈인 인코더의 입력으로 변환하여 피드백할 수 있다. 변환 모듈은 출력 및 입력의 포맷에 맞게 데이터 값을 변환하며 제2 피드백 경로 상에서 제2 변환 모듈을 통해 데이터 값을 변환한다.
- [0045] 다중 학습 네트워크 모델은 학습 과정에서 제3 피드백 경로를 통해 블러 개선 네트워크의 제1 디코더의 출력을 모션 예측 네트워크의 제2 디코더의 입력으로 변환하여 피드백할 수 있다. 변환 모듈은 출력 및 입력의 포맷에 맞게 데이터 값을 변환하며 제3 피드백 경로 상에서 제3 변환 모듈을 통해 데이터 값을 변환한다.
- [0046] 다중 학습 네트워크 모델은 학습 과정에서 제4 피드백 경로를 통해 모션 예측 네트워크의 제2 디코더의 출력을 블러 개선 네트워크의 제1 디코더의 입력으로 변환하여 피드백할 수 있다. 변환 모듈은 출력 및 입력의 포맷에

맞게 데이터 값을 변환하며 제4 피드백 경로 상에서 제4 변환 모듈을 통해 데이터 값을 변환한다.

[0047] 다중 학습 네트워크 모델은 손실 함수를 최소화하는 파라미터를 도출하도록 학습된다.

[0048] 전체 손실 함수는 블러 개선 네트워크 및 모션 예측 네트워크를 학습하기 위한 L_{deblur} 와 L_{flow} 의 비용 함수의 합으로 설계된다. L_{deblur} 는 최신의 CNN 기반 초해상도 영상 복원 기술을 따라 수학식 1과 같이 Perceptual Loss와 Adversarial Loss의 합으로 설계할 수 있다. λ 는 0.01로 설정될 수 있으며, 요구되는 설계 사항에 따라 다른 수치 적용이 가능하다.

수학식 1

$$L_{deblur} = L_{VGG} + \lambda L_{adv}$$

[0050] L_{VGG} 는 ImageNet 데이터베이스를 이용한 영상 분류 분야에 이미 학습된 VGG 네트워크의 표현자를 통해 계산된다.

L_{VGG} 는 개선 영상 S_{deblur} 와 정답 $S_{\lfloor M/2 \rfloor}$ 를 수학식 2와 같이 설계된다.

수학식 2

$$L_{VGG} = \| \Phi(S_{deblur}) - \Phi(S_{\lfloor M/2 \rfloor}) \|^2$$

[0052] Φ 는 VGG 네트워크의 표현자 추정 네트워크를 나타낸다. 이러한 의미론적 특성을 갖는 표현자를 이용한 비용 함수는 가시적으로 향상된 개선 영상을 획득할 수 있다. L_{adv} 는 개선 영상과 정답 영상 간의 구분이 안 되도록 하는 비용 함수이다. 이러한 적대적(Adversarial) 학습을 위해서는 입력 영상이 실제 뚜렷한 영상인지 개선된 영상인지를 분류하기 위한 판별자(Discriminator) 네트워크가 필요하다. GAN(Generative Adversarial Network) 모델을 고려하여 표 1과 같이 판별자 네트워크를 설계할 수 있다.

표 1

Layer block	Weight dimension (stride)
Conv + LReLU	$64 \times 3 \times 3 \times 3$ (1)
Conv + LReLU + BN	$64 \times 64 \times 4 \times 4$ (2)
Conv + LReLU + BN	$128 \times 64 \times 4 \times 4$ (2)
Conv + LReLU + BN	$256 \times 128 \times 4 \times 4$ (2)
Conv + LReLU + BN	$512 \times 256 \times 4 \times 4$ (2)
Conv + LReLU + BN	$512 \times 512 \times 4 \times 4$ (2)
FC	$1 \times 512 \times 4 \times 4$ (1)

[0053]

[0054] 여기서, Conv, LReLU, BN, 및 FC는 각각 컨볼루션, Leaky ReLU, 배치 정규화(Batch Normalization), 풀리 커넥티드 레이어(Fully Connected Layer)를 의미한다. 표 1의 판별자는 블러 개선 네트워크의 결과 영상 또는 정답 영상을 입력으로 하며, 수학식 3의 비용 함수로 학습된다.

수학식 3

$$L_D = D(S_{\lfloor M/2 \rfloor}) - D(G(B))$$

D와 G는 각각 판별자 및 블러 개선 네트워크를 나타낸다. L_{deblur} 에서 L_{adv} 가 개선 영상과 정답 영상의 구분 안 되게 학습하기 위해서는 수학식 4의 비용 함수를 최소화해야 한다.

수학식 4

$$L_{adv} = -D(G(B))$$

단일 블러 영상으로부터 모션 예측 네트워크(옵티컬 플로우 추정 네트워크)를 학습하기 위해서는 먼저 플로우에 대한 정답을 알아야 한다. 그러나 일반적으로 이러한 정답은 존재하지 않는다.

본 실시예에 따른 블러 개선 장치는 S[0]와 S[M-1] 두 영상을 이용하여 옵티컬 플로우 추정 모델을 이용하여 플로우에 대한 슈도 정답($\hat{F}_{0 \rightarrow M-1}$ 및 $\hat{F}_{0 \leftarrow M-1}$)을 도출한다. 옵티컬 플로우 추정 모델은 비교할 이미지 쌍을 스택한 후 컨볼루션 레이어를 거쳐 추정하거나 비교할 이미지 쌍을 각각의 컨볼루션 레이어를 거친 후 서로의 입력을 이용하는 상관 레이어(Correlation Layer)를 통해 특징을 비교한다.

다중 학습 네트워크 구조를 학습하기 위한 입력 영상 및 블러 개선과 플로우 추정에 대한 정답 영상의 예시기도 6에 예시되어 있다.

L_{flow} 는 수학식 5와 같이 표현된다.

수학식 5

$$L_{flow} = \|F_{forward} - \hat{F}_{0 \rightarrow M-1}\|^2 + \|F_{backward} - \hat{F}_{0 \leftarrow M-1}\|^2$$

$F_{forward}$ 와 $F_{backward}$ 는 모션 예측 네트워크(옵티컬 플로우 추정 네트워크)로 추정된 전방 옵티컬 플로우와 후방 옵티컬 플로우를 나타낸다.

도 7은 본 발명의 다른 실시예에 따른 블러 개선 방법을 예시한 흐름도이다. 블러 개선 방법은 블러 개선 장치 또는 컴퓨팅 디바이스에 의하여 수행될 수 있다.

단계 S210에서 프로세서는 단일 블러 이미지를 입력받는다.

단계 S220에서 프로세서는 입력된 단일 블러 이미지로부터 다중 학습 네트워크 모델을 통해 비균일 블러를 제거한다.

다중 학습 네트워크 모델의 블러 개선 네트워크 및 모션 예측 네트워크는 인코더를 상호 공유한다. 블러 개선 네트워크 및 모션 예측 네트워크는 동시에 학습된다.

인코더는 컨볼루션 레이어(Convolutional Layer)가 적용된 모션 기반의 특징 추출 모듈이다. 인코더는 컨볼루션 레이어를 통해 표현자를 추출하고, 스킵 구조를 갖는 레지듀얼 블록을 통해 상기 표현자의 채널 개수를 증가시킨다.

블러 개선 네트워크는 인코더와 제1 디코더를 포함하며, 제1 디코더는 인코더에서 추출한 특징으로부터 이미지

를 복원한다. 제1 디코더는 디컨볼루션 레이어(Deconvolutional Layer)를 통해 해상도를 증가시키고, 디컨볼루션 레이어의 후단에 제1 컨볼루션 레이어를 추가하고, 제1 컨볼루션 레이어에서 표현자의 채널 개수를 인코더와 대칭이 되도록 감소시키고, 제1 컨볼루션 레이어는 블러 개선된 이미지를 추정한다. 대칭이 되는 인코더와 제1 디코더 간에 스킵 연결을 통해 표현자의 공간적 구조를 유지한다.

[0070] 모션 예측 네트워크는 인코더와 제2 디코더를 포함하며, 제2 디코더는 인코더에서 추출한 특징으로부터 옵티컬 플로우를 추정한다. 제2 디코더는 디컨볼루션 레이어를 통해 상기 옵티컬 플로우를 추정하고, 디컨볼루션 레이어의 후단에 제2 컨볼루션 레이어를 추가하고, 제2 컨볼루션 레이어에서 표현자의 채널 개수를 인코더와 대칭이 되도록 감소시키고, 제2 컨볼루션 레이어는 전방 옵티컬 플로우와 후방 옵티컬 플로우를 추정한다. 대칭이 되는 인코더와 제2 디코더 간에 스킵 연결을 통해 표현자의 공간적 구조를 유지한다.

[0071] 모션 예측 네트워크는 학습 과정에서 옵티컬 플로우에 대한 슈도 정답을 적용한다. 옵티컬 플로우에 대한 슈도 정답은 공간 피라미드 배합과 심층 학습을 결합한 옵티컬 플로우 추정 모델에 의해 산출된다. 옵티컬 플로우 추정 모델은 비교할 이미지 쌍을 스택한 후 컨볼루션 레이어를 거쳐 추정하거나 비교할 이미지 쌍을 각각의 컨볼루션 레이어를 거친 후 서로의 입력을 이용하는 상관 레이어(Correlation Layer)를 통해 특징을 비교한다.

[0072] 블러 개선 장치는 하드웨어, 펌웨어, 소프트웨어 또는 이들의 조합에 의해 로직회로 내에서 구현될 수 있고, 범용 또는 특정 목적 컴퓨터를 이용하여 구현될 수도 있다. 장치는 고정배선형(Hardwired) 기기, 필드 프로그램 가능한 게이트 어레이(Field Programmable Gate Array, FPGA), 주문형 반도체(Application Specific Integrated Circuit, ASIC) 등을 이용하여 구현될 수 있다. 또한, 장치는 하나 이상의 프로세서 및 컨트롤러를 포함한 시스템온칩(System on Chip, SoC)으로 구현될 수 있다.

[0073] 블러 개선 장치는 하드웨어적 요소가 마련된 컴퓨팅 디바이스 또는 서버에 소프트웨어, 하드웨어, 또는 이들의 조합하는 형태로 탑재될 수 있다. 컴퓨팅 디바이스 또는 서버는 각종 기기 또는 유무선 통신망과 통신을 수행하기 위한 통신 모듈 등의 통신장치, 프로그램을 실행하기 위한 데이터를 저장하는 메모리, 프로그램을 실행하여 연산 및 명령하기 위한 마이크로프로세서 등을 전부 또는 일부 포함한 다양한 장치를 의미할 수 있다.

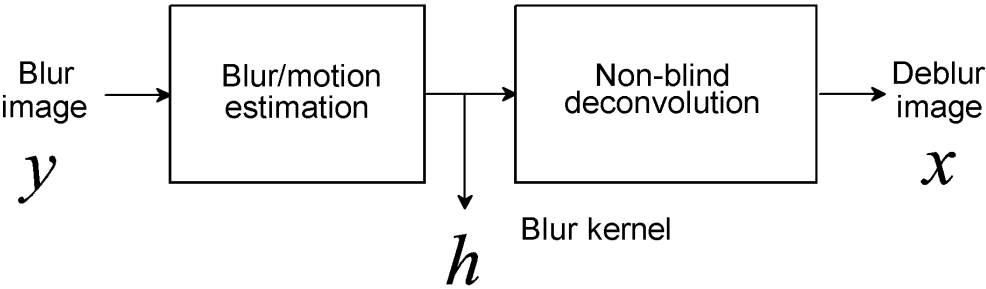
[0074] 도 7에서는 각각의 과정을 순차적으로 실행하는 것으로 기재하고 있으나 이는 예시적으로 설명한 것에 불과하고, 이 분야의 기술자라면 본 발명의 실시예의 본질적인 특성에서 벗어나지 않는 범위에서 도 7에 기재된 순서를 변경하여 실행하거나 또는 하나 이상의 과정을 병렬적으로 실행하거나 다른 과정을 추가하는 것으로 다양하게 수정 및 변형하여 적용 가능할 것이다.

[0075] 본 실시예들에 따른 동작은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능한 매체에 기록될 수 있다. 컴퓨터 판독 가능한 매체는 실행을 위해 프로세서에 명령어를 제공하는 데 참여한 임의의 매체를 나타낸다. 컴퓨터 판독 가능한 매체는 프로그램 명령, 데이터 파일, 데이터 구조 또는 이들의 조합을 포함할 수 있다. 예를 들면, 자기 매체, 광기록 매체, 메모리 등이 있을 수 있다. 컴퓨터 프로그램은 네트워크로 연결된 컴퓨터 시스템 상에 분산되어 분산 방식으로 컴퓨터가 읽을 수 있는 코드가 저장되고 실행될 수도 있다. 본 실시예를 구현하기 위한 기능적인(Functional) 프로그램, 코드, 및 코드 세그먼트들은 본 실시예가 속하는 기술분야의 프로그래머들에 의해 용이하게 추론될 수 있을 것이다.

[0076] 본 실시예들은 본 실시예의 기술 사상을 설명하기 위한 것이고, 이러한 실시예에 의하여 본 실시예의 기술 사상의 범위가 한정되는 것은 아니다. 본 실시예의 보호 범위는 아래의 청구범위에 의하여 해석되어야 하며, 그와 동등한 범위 내에 있는 모든 기술 사상은 본 실시예의 권리범위에 포함되는 것으로 해석되어야 할 것이다.

도면

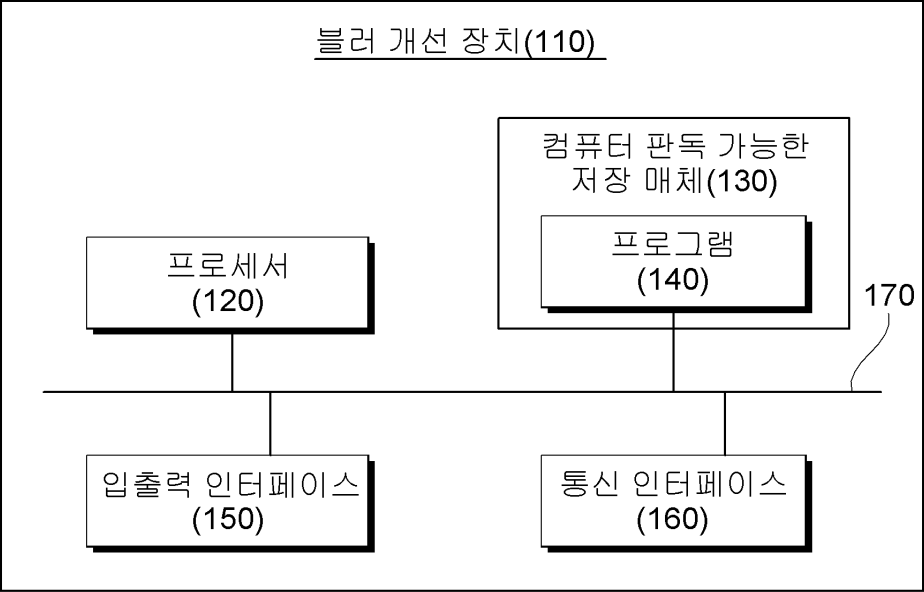
도면1



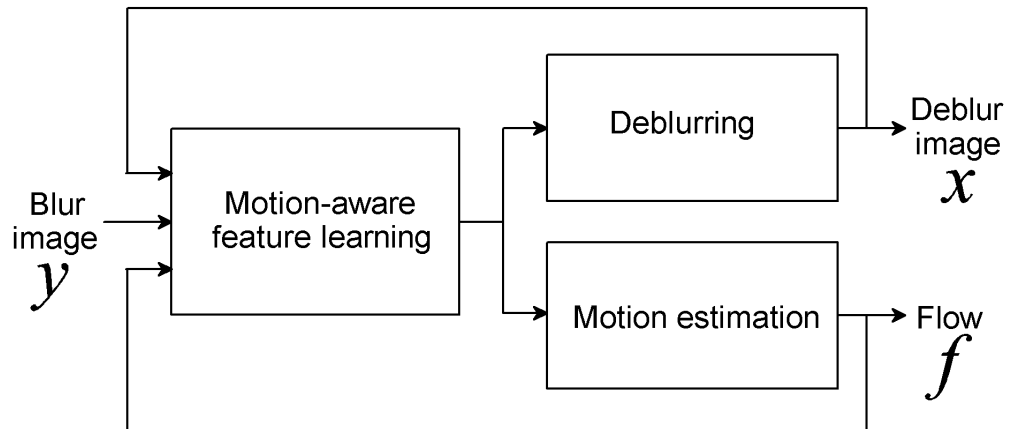
도면2



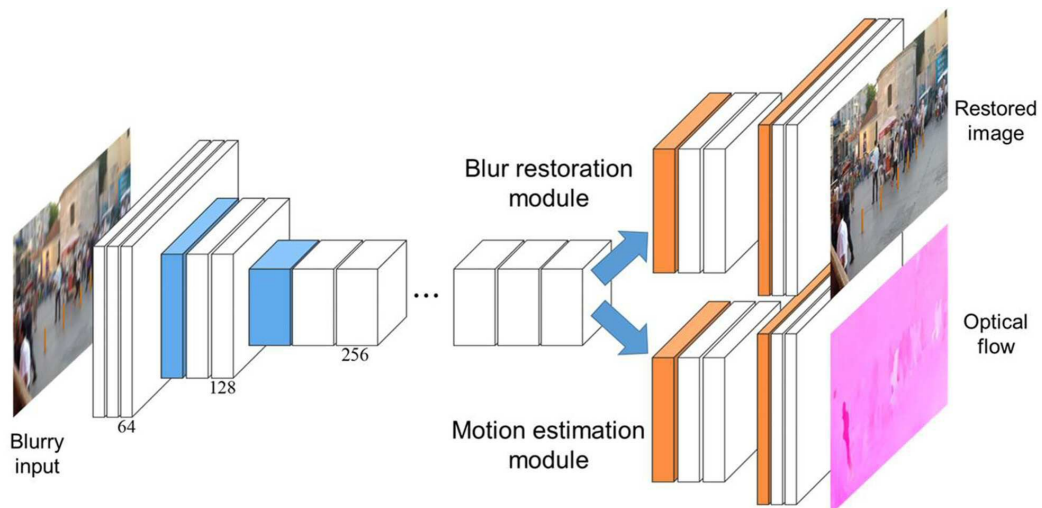
도면3



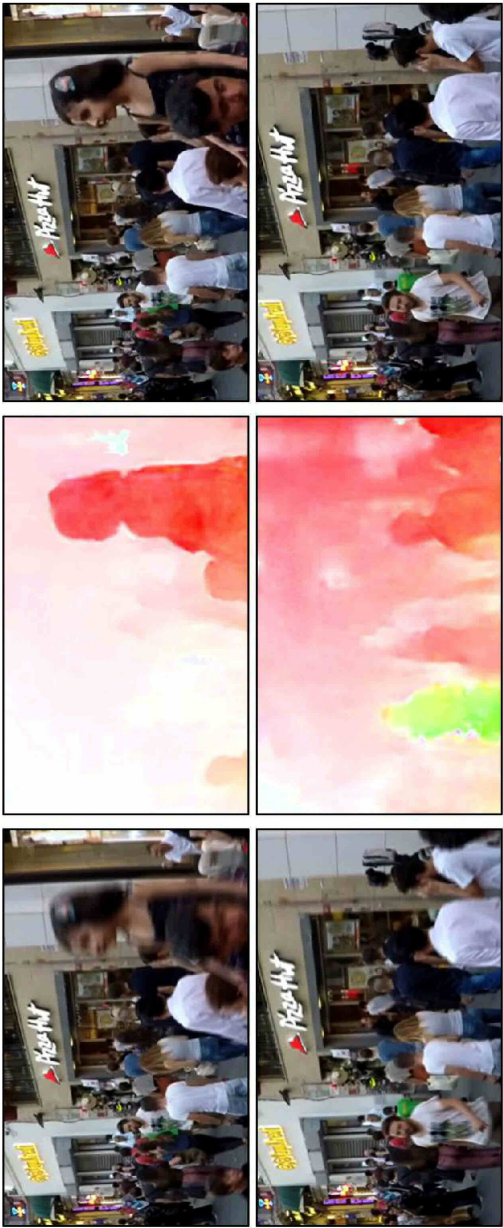
도면4



도면5



도면6



도면7

