



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2020년10월23일
(11) 등록번호 10-2169242
(24) 등록일자 2020년10월19일

(51) 국제특허분류(Int. Cl.)
G06T 5/00 (2019.01) G06N 3/08 (2006.01)
G06T 3/40 (2006.01) G06T 5/20 (2006.01)
(52) CPC특허분류
G06T 5/001 (2013.01)
G06N 3/08 (2013.01)
(21) 출원번호 10-2018-0169475
(22) 출원일자 2018년12월26일
심사청구일자 2018년12월26일
(65) 공개번호 10-2020-0084434
(43) 공개일자 2020년07월13일
(56) 선행기술조사문헌
US20180075581 A1*
Xintao Wang ET AL: "ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks", The European Conference on Computer Vision (ECCV) Workshops, September 8-14, 2018(2018.09.08.) 1부.*
JP6448839 B1
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
포항공과대학교 산학협력단
경상북도 포항시 남구 청암로 77 (지곡동)
(72) 발명자
이승용
경상북도 포항시 남구 청암로 77 포항공과대학교
정보통신연구소 333호
박성진
경기도 수원시 영통구 광고중앙로 145, C동 1122호 (이의동, 광고엘포트 아이파크)
손형석
경상북도 포항시 남구 청암로 77 포항공과대학교
정보통신연구소 233호
(74) 대리인
김합곤, 안광석, 박영일

전체 청구항 수 : 총 11 항

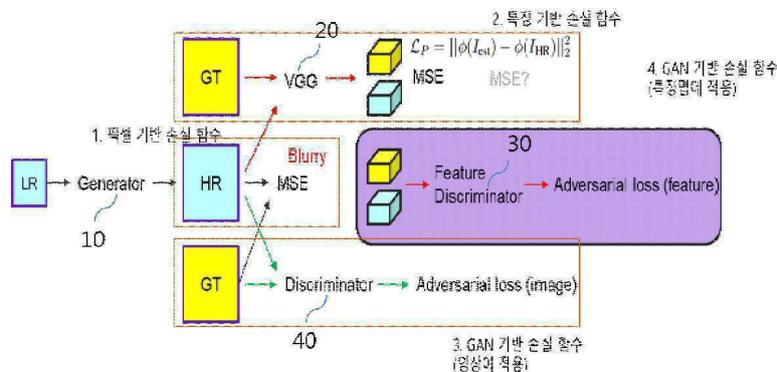
심사관 : 김광식

(54) 발명의 명칭 초해상도 영상 복원을 위한 기계 학습 방법

(57) 요약

본 발명은 교사 영상들의 특징맵들이 가지는 특성 분포를 사용하는 GAN 손실 함수를 이용하여 네트워크가 좀 더 손쉽게 자연스러운 디테일 생성을 학습할 수 있게 한 초해상도 영상 복원을 위한 기계 학습 방법에 관한 것으로, 초해상도 영상 복원 장치가 초해상도 교사 영상 및 상기 초해상도 교사 영상을 다운 샘플링시켜 생성한 저해상도 중간 영상을 이용하여 초해상도 영상 복원 모델을 기계 학습하는 방법은, 초해상도 영상 복원 모델을 이용하여 저해상도 중간 영상의 해상도를 향상시켜 보정 영상을 생성하는 단계; 특징 추출 모델을 이용하여 보정 영상과 초해상도 교사 영상으로부터 보정 특징맵과 교사 특징맵을 생성하는 단계; 및 제 1 선명도 판별 모델을 이용하여 측정된 보정 특징맵의 판별 에러가 최소가 되는 조건을 만족하도록 초해상도 영상 복원 모델을 학습하는 단계;를 포함하여 구성된다.

대표도



(52) CPC특허분류

G06T 3/4076 (2013.01)

G06T 5/20 (2013.01)

G06T 2207/20081 (2013.01)

G06T 2207/20084 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711065143
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기술진흥센터
연구사업명	SW컴퓨팅산업원천기술개발
연구과제명	(SW 스타랩) 빅 비주얼 데이터 기반의 고품질 사진 메이크업 SW 개발
기 여 율	70/100
과제수행기관명	포항공과대학교
연구기간	2018.01.01 ~ 2018.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호	1711074242
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	차세대정보·컴퓨팅기술개발
연구과제명	초실감 원격가상 인터랙션을 위한 teleportation 기술 개발
기 여 율	30/100
과제수행기관명	포항공과대학교
연구기간	2018.06.01 ~ 2019.05.31

공지예외적용 : 있음

명세서

청구범위

청구항 1

초해상도 영상 복원 장치가 초해상도 교사 영상 및 상기 초해상도 교사 영상을 다운 샘플링시켜 생성한 저해상도 중간 영상을 이용하여 초해상도 영상 복원 모델을 기계 학습하는 방법에 있어서,

상기 초해상도 영상 복원 모델을 이용하여 상기 저해상도 중간 영상의 해상도를 향상시켜 보정 영상을 생성하는 단계;

특징 추출 모델을 이용하여 상기 보정 영상과 상기 초해상도 교사 영상으로부터 보정 특징맵과 교사 특징맵을 생성하는 단계; 및

제 1 선명도 판별 모델을 이용하여 측정된 상기 보정 특징맵의 판별 에러가 최소가 되는 조건을 만족하도록 상기 초해상도 영상 복원 모델을 학습하는 단계;를 포함하여 구성되고,

상기 제 1 선명도 판별 모델은 컨볼루션 신경망으로 이루어지고, 상기 컨볼루션 신경망은 상기 보정 특징맵 또는 상기 교사 특징맵을 입력으로 하여, 상기 보정 특징맵을 false, 상기 교사 특징맵을 true로 판별하도록 학습하는 것을 특징으로 하는 초해상도 영상 복원을 위한 기계 학습 방법.

청구항 2

제 1 항에 있어서,

상기 초해상도 영상 복원 모델을 학습하는 단계는, 상기 보정 특징맵과 상기 교사 특징맵 사이의 제곱 에러와, 제 1 선명도 판별 모델을 이용하여 측정된 상기 보정 특징맵의 판별 에러와, 제 2 선명도 판별 모델을 이용하여 측정된 상기 보정 영상의 판별 에러의 합이 최소가 되는 조건을 만족하도록 상기 초해상도 영상 복원 모델을 학습하는 단계임을 특징으로 하는 초해상도 영상 복원을 위한 기계 학습 방법.

청구항 3

제 1 항에 있어서,

상기 초해상도 영상 복원 모델은 컨볼루션 신경망으로 이루어짐을 특징으로 하는 초해상도 영상 복원을 위한 기계 학습 방법.

청구항 4

제 3 항에 있어서,

상기 컨볼루션 신경망은

컨볼루션 필터를 이용하여 특징맵을 생성하는 컨볼루션 레이어로 구성되어 순차적으로 이어진 복수의 레지듀얼 블록들; 및

영상의 해상도를 향상시키는 적어도 하나의 업샘플링 레이어;를 포함하여 구성되는 것을 특징으로 하는 초해상도 영상 복원을 위한 기계 학습 방법.

청구항 5

제 4 항에 있어서,

상기 순차적으로 이어진 복수의 레지듀얼 블록들 중 마지막 레지듀얼 블록의 뒤에 위치하고 각 레지듀얼 블록의 가중치가 부여된 출력 결과를 합산하는 노드와, 각 레지듀얼 블록의 출력 결과에 가중치를 부여하는 복수의 1x1 컨볼루션 레이어들을 더 포함하는 것을 특징으로 하는 초해상도 영상 복원을 위한 기계 학습 방법.

청구항 6

삭제

청구항 7

제 1 항에 있어서,

상기 컨볼루션 신경망은,

컨볼루션 필터들을 이용하여 특징맵을 생성하는 컨볼루션 레이어로 구성되어 순차적으로 이어진 복수의 컨볼루션 블록들; 및

상기 복수의 컨볼루션 블록들의 특징맵으로부터 상기 컨볼루션 신경망의 입력 특징맵이 선명한 특성을 가지는지 아닌지 판별 결과를 생성하는 완전하게 연결된(fully connected) 적어도 하나의 레이어로 이루어진 판별기;를 포함하는 것을 특징으로 하는 초해상도 영상 복원을 위한 기계 학습 방법.

청구항 8

삭제

청구항 9

제 1 항에 있어서,

상기 초해상도 영상 복원 모델에 대한 학습과 상기 제 1 선명도 판별 모델에 대한 학습은 동시에 또는 번갈아가며 이루어지는 것을 특징으로 하는 초해상도 영상 복원을 위한 기계 학습 방법.

청구항 10

제 2 항에 있어서,

상기 제 2 선명도 판별 모델은 컨볼루션 신경망으로 이루어짐을 특징으로 하는 초해상도 영상 복원을 위한 기계 학습 방법.

청구항 11

제 10 항에 있어서,

상기 컨볼루션 신경망은,

컨볼루션 필터들을 이용하여 특징맵을 생성하는 컨볼루션 레이어로 구성되어 순차적으로 이어진 복수의 컨볼루션 블록들; 및

상기 복수의 컨볼루션 블록들의 특징맵으로부터 상기 컨볼루션 신경망의 입력 영상이 선명한 특성을 가지는지 아닌지 판별 결과를 생성하는 완전하게 연결된(fully connected) 적어도 하나의 레이어로 이루어진 판별기;를 포함하는 것을 특징으로 하는 초해상도 영상 복원을 위한 기계 학습 방법.

청구항 12

제 1 항 내지 제5항, 제7항, 제9항 내지 제 11 항 중 어느 한 항의 초해상도 영상 복원을 위한 기계 학습 방법을 컴퓨터에서 실행시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록 매체.

청구항 13

제 1 항 내지 제5항, 제7항, 제9항 내지 제 11 항 중 어느 한 항의 초해상도 영상 복원을 위한 기계 학습 방법을 컴퓨터에서 실행시키기 위하여 매체에 저장된 컴퓨터 프로그램.

발명의 설명

기술 분야

본 발명은 초해상도 영상 복원 기술에 관한 것으로, 보다 상세하게는 초해상도 영상 복원을 위한 기계 학습 방

[0001]

법에 관한 것이다.

배경 기술

- [0002] 딥러닝 및 CNN(convolution neural network)을 이용한 초해상도(super-resolution) 영상 복원 방법은 지도학습(supervised learning)을 통해 네트워크를 학습한다. 이 때, 보통 일반 선명한 영상들을 지도 학습의 교사 영상 혹은 레이블(label) 개념으로 사용하고, 다운샘플링(down sampling)을 통해 해상도를 낮춘 영상을 입력 저해상도 영상으로 사용한다.
- [0003] 기계 학습 과정은 다음과 같다.
- [0004] 매 반복(iteration)마다 입력 저해상도 영상이 컨볼루션 레이어와 업샘플링 레이어로 이루어진 네트워크를 통과하여 초해상도 영상으로 변환되고, 초해상도 영상을 교사 영상과 비교하여 차이를 측정하고, 차이가 줄어들도록 차이를 역전파(back-propagate)하여 컨볼루션 레이어의 웨이트(weights)를 변화시킨다. 학습이 진행될수록 학습 데이터에 대해서 초해상도 영상(네트워크의 결과)과 교사 영상 사이의 차이 값이 점점 줄어들고, 데이터가 다양할수록 학습 효과가 일반화되어 학습에 사용되지 않은 다른 영상에 대해서도 초해상도 영상 복원 성능이 향상되게 된다.
- [0005] 딥러닝 및 CNN을 이용한 초해상도 영상 복원 방법에서는, 초해상도 영상과 교사 영상 사이의 차이를 측정하여 네트워크에 역전파(back-propagate)하는 손실 함수(loss function)의 역할이 중요하다. 손실 함수로 초해상도 영상 교사 영상 사이의 픽셀 마다 차이를 측정하는 L1(픽셀 값의 차이의 절대값의 평균), L2(픽셀 값의 차이의 제곱의 평균, Mean Squared Error라고도 부름) 함수를 사용하는 방법이 있다.
- [0006] 2차원 그래프에서 커브 피팅을 생각해 보면 알 수 있듯이, 어떤 수학적 커브 모델도 모든 관찰점을 완벽히 지나지는 커브를 만들 수 없다. 결국 관찰점들의 사이를 적당히 지나는 커브가 나오는 경우를 많이 볼 수 있다. 초해상도 영상 복원에서조차 마찬가지로 아무리 잘 학습해도 모든 픽셀에서 초해상도 영상(네트워크의 결과)과 교사 영상(ground-truth 고해상도 영상)의 차이가 0이 될 수 없다. 이 때 픽셀 당 차이에 기반하는 L1, L2 손실 함수를 사용할 경우, 모든 픽셀에 대해 적당한 차이를 갖고 디테일이 뭉개진 초해상도 영상 결과를 얻게 된다. 다시 말하면, L1, L2 손실 함수는 영상을 선명하게 만드는 것이 아니고 픽셀들의 평균 차이를 줄이려고 하기 때문에 결과 영상에서 고주파수 디테일이 복원되지 않고 블러된(Blurred) 결과를 만든다.
- [0007] 이와 같은 문제점으로 인하여, GAN(Generative Adversarial Networks) 손실 함수를 사용하여 네트워크가 고주파수 디테일을 포함한 선명해 보이는 초해상도 복원 결과를 생성하도록 학습시키는 방법이 제안되었다. GAN 손실 함수를 사용한 학습 방법은 비지도 학습 기법을 통해 학습이 되며, 무수히 많은 교사 영상들이 가지는 선명한 영상의 특성 분포에 기반하여 네트워크의 결과가 선명한지 선명하지 않은지를 판별하여, 인지적으로 선명하게 보이는 결과를 생성하도록 네트워크가 학습된다.
- [0008] 그러나 GAN 손실 함수 방법의 경우 교사 영상들이 가지는 특성 분포가 매우 다양하기 때문에 자연스러운 고주파수 디테일 생성을 학습하기 어렵다는 문제가 있다.

선행기술문헌

비특허문헌

- [0009] (비특허문헌 0001) Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network, Ledig et al., CVPR 2017. (SRGAN)
- (비특허문헌 0002) EnhanceNet: Single Image Super-Resolution Through Automated Texture Synthesis, Sajjadi et al., ICCV 2017. (EnhanceNet)

발명의 내용

해결하려는 과제

- [0010] 본 발명은 교사 영상들의 특징맵들이 가지는 특성 분포를 사용하는 GAN 손실 함수를 이용하여 네트워크가 좀 더 손쉽게 자연스러운 디테일 생성을 학습할 수 있게 한 초해상도 영상 복원을 위한 기계 학습 방법을 제공함을 그

목적으로 한다.

과제의 해결 수단

- [0011] 상기의 목적을 달성하기 위하여, 본 발명에 의한 초해상도 영상 복원 장치가 초해상도 교사 영상 및 상기 초해상도 교사 영상을 다운 샘플링시켜 생성한 저해상도 중간 영상을 이용하여 초해상도 영상 복원 모델을 기계 학습하는 방법은, 상기 초해상도 영상 복원 모델을 이용하여 상기 저해상도 중간 영상의 해상도를 향상시켜 보정 영상을 생성하는 단계; 특징 추출 모델을 이용하여 상기 보정 영상과 상기 초해상도 교사 영상으로부터 보정 특징맵과 교사 특징맵을 생성하는 단계; 및 제 1 선명도 판별 모델을 이용하여 측정된 상기 보정 특징맵의 판별 에러가 최소가 되는 조건을 만족하도록 상기 초해상도 영상 복원 모델을 학습하는 단계;를 포함하여 구성된다.
- [0012] 상기 초해상도 영상 복원을 위한 기계 학습 방법에 있어서, 상기 초해상도 영상 복원 모델을 학습하는 단계는, 상기 보정 특징맵과 상기 교사 특징맵 사이의 제곱 에러와, 제 1 선명도 판별 모델을 이용하여 측정된 상기 보정 특징맵의 판별 에러와, 제 2 선명도 판별 모델을 이용하여 측정된 상기 보정 영상의 판별 에러의 합이 최소가 되는 조건을 만족하도록 상기 초해상도 영상 복원 모델을 학습하는 단계임을 특징으로 한다.
- [0013] 상기 초해상도 영상 복원을 위한 기계 학습 방법에 있어서, 상기 초해상도 영상 복원 모델은 컨볼루션 신경망으로 이루어짐을 특징으로 한다.
- [0014] 상기 초해상도 영상 복원을 위한 기계 학습 방법에 있어서, 상기 컨볼루션 신경망은 컨볼루션 필터를 이용하여 특징맵을 생성하는 컨볼루션 레이어로 구성되어 순차적으로 이어진 복수의 레지듀얼 블록들; 및 영상의 해상도를 향상시키는 적어도 하나의 업샘플링 레이어;를 포함하여 구성되는 것을 특징으로 한다.
- [0015] 상기 초해상도 영상 복원을 위한 기계 학습 방법에 있어서, 상기 순차적으로 이어진 복수의 레지듀얼 블록들 중 마지막 레지듀얼 블록의 뒤에 위치하고 각 레지듀얼 블록의 가중치가 부여된 출력 결과를 합산하는 노드와, 각 레지듀얼 블록의 출력 결과에 가중치를 부여하는 복수의 1x1 컨볼루션 레이어들을 더 포함하는 것을 특징으로 한다.
- [0016] 상기 초해상도 영상 복원을 위한 기계 학습 방법에 있어서, 상기 제 1 선명도 판별 모델은 컨볼루션 신경망으로 이루어짐을 특징으로 한다.
- [0017] 상기 초해상도 영상 복원을 위한 기계 학습 방법에 있어서, 상기 컨볼루션 신경망은, 컨볼루션 필터들을 이용하여 특징맵을 생성하는 컨볼루션 레이어로 구성되어 순차적으로 이어진 복수의 컨볼루션 블록들; 및 상기 복수의 컨볼루션 블록들의 특징맵으로부터 상기 컨볼루션 신경망의 입력 특징맵이 선명한 특성을 가지는지 아닌지 판별 결과를 생성하는 완전하게 연결된(fully connected) 적어도 하나의 레이어로 이루어진 판별기;를 포함하는 것을 특징으로 한다.
- [0018] 상기 초해상도 영상 복원을 위한 기계 학습 방법에 있어서, 상기 컨볼루션 신경망은, 상기 보정 특징맵 또는 상기 교사 특징맵을 입력으로 하여, 상기 보정 특징맵을 false, 상기 교사 특징맵을 true로 판별하도록 학습하는 것을 특징으로 한다.
- [0019] 상기 초해상도 영상 복원을 위한 기계 학습 방법에 있어서, 상기 초해상도 영상 복원 모델에 대한 학습과 상기 제 1 선명도 판별 모델에 대한 학습은 동시에 또는 번갈아 가며 이루어지는 것을 특징으로 한다.
- [0020] 상기 초해상도 영상 복원을 위한 기계 학습 방법에 있어서, 상기 제 2 선명도 판별 모델은 컨볼루션 신경망으로 이루어짐을 특징으로 한다.
- [0021] 상기 초해상도 영상 복원을 위한 기계 학습 방법에 있어서, 상기 컨볼루션 신경망은, 컨볼루션 필터들을 이용하여 특징맵을 생성하는 컨볼루션 레이어로 구성되어 순차적으로 이어진 복수의 컨볼루션 블록들; 및 상기 복수의 컨볼루션 블록들의 특징맵으로부터 상기 컨볼루션 신경망의 입력 영상이 선명한 특성을 가지는지 아닌지 판별 결과를 생성하는 완전하게 연결된(fully connected) 적어도 하나의 레이어로 이루어진 판별기;를 포함하는 것을 특징으로 한다.
- [0022] 상기의 목적을 달성하기 위하여, 본 발명에 의한 컴퓨터로 읽을 수 있는 기록 매체는 상기 초해상도 영상 복원을 위한 기계 학습 방법을 컴퓨터에서 실행시키기 위한 프로그램을 기록한다.
- [0023] 상기의 목적을 달성하기 위하여, 본 발명에 의한 컴퓨터 프로그램은 상기 초해상도 영상 복원을 위한 기계 학습 방법을 컴퓨터에서 실행시키기 위하여 매체에 저장된다.

발명의 효과

[0024] 본 발명에 의하면, 픽셀보다 인지적으로 더욱 함축적인 정보를 담고 있는 특징맵에 기반하여, 교사 영상들의 특징맵들이 가지는 특성 분포를 사용하는 GAN 손실 함수를 이용해 네트워크가 좀 더 손쉽게 자연스러운 디테일 생성을 학습할 수 있게 된다.

도면의 간단한 설명

[0025] 도 1은 본 발명에 의한 초해상도 영상 복원 장치가 초해상도 영상 복원 모델을 기계 학습하는 방법을 설명하는 도면이다.

도 2는 본 발명에 의한 Generator Network의 구조를 도시한 것이다.

도 3은 본 발명에 의한 Discriminator Network의 구조를 도시한 것이다.

도 4는 저해상도 입력 영상에서 초해상도 영상을 복원한 결과를 종래기술과 대비하여 보여준다.

발명을 실시하기 위한 구체적인 내용

[0026] 이하에서, 첨부된 도면을 참조하면서 본 발명의 바람직한 실시예에 대하여 상세히 설명하기로 한다.

[0027] 도 1을 참조하면, 본 발명에 의한 초해상도 영상 복원 장치는 제너레이터 네트워크(Generator Network)(10), 특징 추출 네트워크(20), 특징 판별기 네트워크(Feature Discriminator Network)(30), 판별기 네트워크(Discriminator Network)(40)를 포함하여 구성된다.

[0028] 제너레이터 네트워크(10), 특징 추출 네트워크(20), 특징 판별기 네트워크(30) 및 판별기 네트워크(40)는 정보 처리장치 상의 하드웨어 및 소프트웨어의 결합체로 이루어지고, 개념적으로 각각 초해상도 영상 복원 모델, 특징 추출 모델, 제 1 선명도 판별 모델 및 제 2 선명도 판별 모델을 구현한다.

[0029] 본 발명에 의한 초해상도 영상 복원 장치에서 제너레이터 네트워크(10)는 저해상도 영상을 초해상도 영상으로 복원하는 역할을 하고, 특징 판별기 네트워크(30)와 판별기 네트워크(40)는 제너레이터 네트워크(10)의 결과가 실제 HR 영상 같은지 판별하는 역할을 한다.

[0030] 도 2를 참조하면, 제너레이터 네트워크(Generator Network)(10)는 컨볼루션 신경망으로, 컨볼루션 필터를 이용하여 특징맵을 생성하는 컨볼루션 레이어로 구성되고 순차적으로 이어진 복수의 레지듀얼 블록들(12), 영상의 해상도를 향상시키는 적어도 하나의 업샘플링 레이어(14) 그리고 스킵 커넥션 구조(16)를 포함하여 구성된다. 스킵 커넥션 구조(16)에 대해서는 아래에서 보다 상세히 설명하기로 한다.

[0031] 특징 추출 네트워크(20)는 예를 들어, VGG network와 같은 것으로, ImageNet과 같이 초거대 데이터셋을 이용해 영상 분류(classification)를 미리 학습하여 영상으로부터 특징(feature)을 추출한다.

[0032] 도 3을 참조하면, 판별기 네트워크(Discriminator Network)(40)는 컨볼루션 신경망으로, 컨볼루션 필터들을 이용하여 특징맵을 생성하는 컨볼루션 레이어로 구성되어 순차적으로 이어진 복수의 컨볼루션 블록들(32)과 복수의 컨볼루션 블록들의 특징맵으로부터 판별기 네트워크(40)의 입력 영상이 선명한 특성을 가지는지 아닌지 판별 결과를 생성하는 완전하게 연결된(fully connected) 적어도 하나의 레이어로 이루어진 판별기(34)를 포함하여 구성된다. 여기서, Fully connected layer는 레이어의 입력 및 출력 특징(feature)들의 각 노드(특징맵의 해상도(H*W)*채널 수(C)개수만큼 존재 존재)들이 모두 서로 연결된 형태이다.

[0033] 한편, 특징 판별기 네트워크(Feature Discriminator Network)(30)는 도 3에 도시된 판별기 네트워크(Discriminator Network)(40)와 그 구조는 동일하지만, 영상 대신 특징 추출 네트워크(20)에 의해 생성된 특징맵을 입력으로 사용한다.

[0035] 이하에서, 도 1을 참조하여 초해상도 영상 복원 장치가 초해상도 교사 영상(GT) 및 (초해상도 교사 영상을 다운 샘플링시켜 생성한) 저해상도 중간 영상(LR)을 이용하여 초해상도 영상 복원 모델을 기계 학습하는 방법에 대하여 설명하기로 한다.

[0036] 먼저, 초해상도 영상 복원 모델을 이용하여 저해상도 중간 영상(LR)의 해상도를 향상시켜 보정 영상(HR)을 생성한다.

[0037] 이후, 특징 추출 모델을 이용하여 보정 영상(HR)과 초해상도 교사 영상(GT)으로부터 보정 특징맵과 교사 특징맵

을 생성한다.

[0038] 이후, 제 1 선명도 판별 모델을 이용하여 측정한 보정 특징맵의 판별 에러가 최소가 되는 조건을 만족하도록, 초해상도 영상 복원 모델을 학습한다.

[0039] 여기서, 판별 에러는 GAN 손실함수의 결과 값이다. 영상 픽셀 기반 GAN 손실함수의 식을 예로 들어보면 수학적 1과 같이 표현된다.

수학적 1

[0040]
$$L_a^i = -\log(d^i(I^g))$$

[0041] 제 1 선명도 판별 모델은 특징맵 기반이므로, 수학적 1에서 d()에 영상 대신 특징맵이 들어가게 된다.

[0042] 즉, L_a^i 는 GAN 손실함수 중 초해상도 영상 복원 모델을 학습하기 위한 손실함수로, 보정 특징맵의 선명도 판별 결과가 1(true)로 판별되도록 하는 손실함수이다. Feature Discriminator Network(30)의 d()의 값의 범위는 0~1이고, d()의 값이 1일 땐 판별 에러(L_a^i)가 0이 되게 되고, d()의 값이 1이 아닌 0~1 사이의 값일 땐 판별 에러(L_a^i)>0이 되게 된다. 따라서, L_a^i 손실함수는 Generator Network(10)가 Feature Discriminator Network(30)를 속일 수 있는 선명한 영상을 만들도록 학습하는 역할을 한다.

[0043] 한편, 네트워크의 학습은 back-propagation을 통해서 이루어지는데, back-propagation은 손실함수의 판별 에러 값이 발생한 부분에서 역방향으로 레이어들을 거치며 chain-rule에 의해 gradient를 전달한다. 이때 GAN 손실함수의 경우 Feature Discriminator Network(30)의 끝에서 에러가 발생하기 때문에 Generator Network(10)까지 이 에러를 전달하기 위해선 매우 많은 레이어들을 거쳐야 하므로, Generator Network(10)의 초반 부분까지 에러가 제대로 전달되지 않을 수 있다.

[0044] 도 2에 의하면, 본 발명에 의한 Generator Network(10)에서는 순차적으로 이어진 복수의 레지듀얼 블록들(12) 중 마지막 레지듀얼 블록의 뒤에 위치하고 각 레지듀얼 블록의 출력 결과를 합산하는 노드에 의한 스킵 커넥션 구조(16)를 포함한다. 따라서, Feature Discriminator Network(30)에서 전달된 gradient를 Generator Network(10)의 구석 구석으로 전달될 수 있게 된다.

[0045] 이 때 일반 스킵 커넥션(skip-connection)을 사용하게 되면 복수의 레지듀얼 블록들(12)의 끝 부분에서 다양한 특징(feature)들이 합쳐지면서 중요하지 않은 특징(feature)들도 똑 같은 가중치를 갖고 더해져서 오히려 학습을 방해할 수 있다. 따라서, 특징(feature)들에 가중치를 두어서 더하기 위해 본 발명에 의한 Generator Network(10)의 스킵 커넥션 구조(16)에는 1x1 컨볼루션 레이어(convolution layer)를 추가하여 가중치를 학습하도록 한다.

[0046] 이후, 보정 특징맵 또는 교사 특징맵을 입력으로 하여, 보정 특징맵을 false, 교사 특징맵을 true로 판별하도록 제 1 선명도 판별 모델을 학습한다.

[0047] 제 1 선명도 판별 모델을 학습하기 위한 영상 픽셀 기반 GAN 손실함수는 수학적 2와 같이 표현된다.

수학적 2

[0048]
$$L_d^i = -\log(d^i(I^h)) - \log(1 - d^i(I^g))$$

[0049] 제 1 선명도 판별 모델은 특징맵 기반이므로, 수학적 2에서 d()에 영상 대신 특징맵이 들어가게 된다.

[0050] 구체적으로 설명하면, 교사 특징맵은 무조건 true, 보정 특징맵은 무조건 false라고 레이블링 한다. 이렇게 되면 Feature Discriminator Network(30)가 교사 특징맵을 false라고 판별하게 되면, 레이블 true와 차이가 발생하여 Feature Discriminator Network(30)가 해당 교사 특징맵을 true로 판별하도록 네트워크의 웨이트

(weights)를 업데이트 하게 된다. 반대로 Feature Discriminator Network(30)가 보정 특징맵을 true로 판별하게 되면, 레이블 (false)와 차이가 발생하여 Feature Discriminator Network(30)가 보정 특징맵을 false로 판별하도록 네트워크의 웨이트(weights)를 업데이트 하게 된다.

[0051] 여기서, 초해상도 영상 복원 모델에 대한 학습과 제 1 선명도 판별 모델에 대한 학습은 동시에 또는 번갈아가며 이루어질 수 있다.

[0053] 본 발명의 다른 실시예에서는, 초해상도 영상 복원 모델을 학습할 때, 제 1 선명도 판별 모델을 이용하여 측정된 보정 특징맵의 판별 에러가 최소가 되는 조건이 아니라, 보정 특징맵과 교사 특징맵 사이의 제곱 에러와, 제 1 선명도 판별 모델을 이용하여 측정된 보정 특징맵의 판별 에러와, 제 2 선명도 판별 모델을 이용하여 측정된 보정 영상의 판별 에러의 합이 최소가 되는 조건을 만족하도록 초해상도 영상 복원 모델을 학습한다.

[0054] 여기서, 보정 특징맵과 교사 특징맵 사이의 제곱 에러는 특징맵의 사이의 L2 손실 함수(특징 기반 손실 함수)로 정의 된다. 특징맵의 각 특징점(일반 영상에서는 픽셀에 대응)들은 넓은 패치 정보가 인지적으로 함축되었기 때문에 특징맵의 차이를 최소화함으로써 더욱 인지적인 영상 복원이 가능하다. 즉, 전체 픽셀들의 차이를 최소화하는 것이 아니라 중요한 특징들의 차이를 최소화함으로써 덜 blurry한 결과를 얻을 수 있다.

[0055] 특징 기반 손실 함수는 픽셀 기반 손실 함수와 함께 입력 영상의 콘텐츠를 유지하는 역할을 한다. GT(ground-truth)가 있는 짝(pair) 데이터셋을 이용해 지도 학습을 수행한다. 픽셀 기반 손실은 픽셀당 차이를 강제로 최소화하려고 하기 때문에 GAN 손실 함수와 함께 사용하면 충돌이 많이 일어나지만, 특징 기반 손실 함수는 이미 정보가 many-to-one으로 함축된 특징들의 차이를 최소화하기 때문에 좀 더 자유도가 높아서 GAN 손실 함수와 함께 사용했을 때 조화로운 결과를 얻을 수 있다.

[0056] 보정 영상의 판별 에러는 GAN 손실함수의 결과 값이다. 영상 픽셀 기반 GAN 손실함수는 수학적 식 1과 같이 표현된다.

[0057] 즉, L_a^i 는 GAN 손실함수 중 초해상도 영상 복원 모델을 학습하기 위한 손실함수로, 보정 영상의 선명도 판별 결과가 1(true)로 판별되도록 하는 손실함수이다. Discriminator Network(40)의 $d()$ 의 값의 범위는 0~1이고, $d()$ 의 값이 1일 땐 판별 에러(L_a^i)가 0이 되게 되고, $d()$ 의 값이 1이 아닌 0~1 사이의 값일 땐 판별 에러(L_a^i)>0이 되게 된다. 따라서, L_a^i 손실함수는 Generator Network(10)가 Discriminator Network(40)를 속일 수 있는 선명한 영상을 만들도록 학습하는 역할을 한다.

[0058] 본 발명의 제2 실시예에서는, 제 1 선명도 판별 모델뿐만 아니라, 제 2 선명도 판별 모델도 학습한다. 제 1 선명도 판별 모델에 대한 학습 방법은 제1 실시예에서 설명한 방식과 동일하다. 한편, 제 2 선명도 판별 모델에 대해서는 보정 영상 또는 교사 영상을 입력으로 하여, 보정 영상을 false, 교사 영상을 true로 판별하도록 학습한다. 제 2 선명도 판별 모델을 학습하기 위한 GAN 손실함수는 수학적 식 2와 같이 표현된다.

[0059] 구체적으로 설명하면, 교사 영상은 무조건 true, 보정 영상은 무조건 false라고 레이블링 한다. 이렇게 되면 Discriminator Network(40)가 교사 영상을 false라고 판별하게 되면, 레이블 (true)와 차이가 발생하여 Discriminator Network(40)가 해당 교사 영상을 true로 판별하도록 네트워크의 웨이트(weights)를 업데이트 하게 된다. 반대로 Discriminator Network(40)가 보정 영상을 true로 판별하게 되면, 레이블 (false)와 차이가 발생하여 Discriminator Network(40)가 보정 영상을 false로 판별하도록 네트워크의 웨이트(weights)를 업데이트 하게 된다.

[0060] 여기서, 초해상도 영상 복원 모델에 대한 학습과 제 1,2 선명도 판별 모델에 대한 학습은 동시에 또는 번갈아가며 이루어질 수 있다.

[0061] 도 4에 의하면, 본 발명을 이용하여 저해상도 입력 영상을 이용하여 초해상도 영상을 복원한 결과를 종래의 기술과 대비하여 보여준다. 도 4에서 (a)는 저해상도 입력 영상이고, (b)는 픽셀 기반 손실 함수를 이용한 결과 영상이고, (c)는 GAN 기반 손실 함수를 이용한 결과 영상이고, (d)는 본 발명에 의한 결과 영상이다.

[0063] 한편, 상술한 본 발명의 실시예는 개인용 컴퓨터를 포함한 범용 컴퓨터에서 사용되는 매체에 기록될 수 있다. 상기 매체는 마그네틱 기록매체(예를 들면, 롬, 플로피 디스크, 하드 디스크 등), 광학적 판독매체(예를 들면, 씨디롬, 디브이디 등) 및 전기적 기록매체(예를 들면, 플래쉬 메모리, 메모리 스틱 등)와 같은 기록매체를 포함

도면4

