



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2020-0073267
(43) 공개일자 2020년06월23일

(51) 국제특허분류(Int. Cl.)
G06T 3/40 (2006.01) G06K 9/46 (2006.01)
G06K 9/62 (2006.01)
(52) CPC특허분류
G06T 3/4053 (2013.01)
G06K 9/46 (2013.01)
(21) 출원번호 10-2020-7014462
(22) 출원일자(국제) 2019년09월25일
심사청구일자 2020년05월20일
(85) 번역문제출일자 2020년05월20일
(86) 국제출원번호 PCT/CN2019/107761
(87) 국제공개번호 WO 2020/063648
국제공개일자 2020년04월02일
(30) 우선권주장
201811155930.9 2018년09월30일 중국(CN)
(뒷면에 계속)

(71) 출원인
보에 테크놀로지 그룹 컴퍼니 리미티드
중국 베이징 100016, 차오양 디스트릭트, 지우시
양치아오 로드 10호
(72) 발명자
류, 한원
중국 100176 베이징 비디에이 디저 로드 넘버9
주, 단
중국 100176 베이징 비디에이 디저 로드 넘버9
나바레테 미첼리니, 파블로
중국 100176 베이징 비디에이 디저 로드 넘버9
(74) 대리인
양영준, 김성운, 백만기

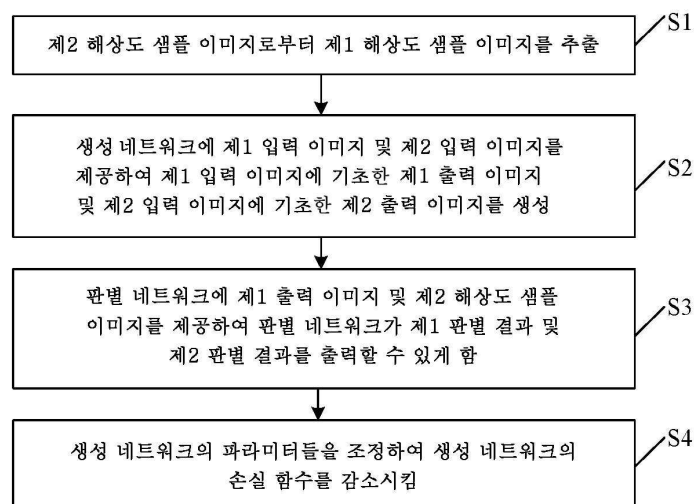
전체 청구항 수 : 총 14 항

(54) 발명의 명칭 생성 적대 네트워크를 위한 트레이닝 방법, 이미지 처리 방법, 디바이스 및 저장 매체

(57) 요약

생성 적대 네트워크를 위한 트레이닝 방법, 이미지 처리 방법, 컴퓨터 디바이스 및 컴퓨터 판독가능 저장 매체. 트레이닝 방법은 생성 네트워크 트레이닝 단계를 포함한다. 생성 네트워크 트레이닝 단계는: 제2 해상도 샘플 이미지로부터 제1 해상도 샘플 이미지를 추출하는 단계(S1); 생성 네트워크에 제1 입력 이미지 및 제2 입력 이미지를 개별적으로 제공하여 제1 입력 이미지-기반 제1 출력 이미지 및 제2 입력 이미지-기반 제2 출력 이미지를 각각 생성하는 단계(S2); 판별 네트워크에 제1 출력 이미지 및 제2 해상도 샘플 이미지를 개별적으로 제공하고, 판별 네트워크가 제1 출력 이미지-기반 제1 판별 결과 및 제2 해상도 샘플 이미지-기반 제2 판별 결과를 출력하는 단계(S3); 및 생성 네트워크의 파라미터들을 조정하여 생성 네트워크의 손실 함수를 감소시키는 단계(S4)를 포함하고, 그에 의해 요구된 이미지들을 획득한다.

대표도 - 도2



(52) CPC특허분류

G06K 9/62 (2013.01)
G06T 3/4046 (2013.01)
G06T 2207/20081 (2013.01)

(30) 우선권주장

201811155326.6	2018년09월30일	중국(CN)
201811155147.2	2018년09월30일	중국(CN)
201811155252.6	2018년09월30일	중국(CN)

명세서

청구범위

청구항 1

생성 적대 네트워크를 위한 트레이닝 방법으로서,

상기 생성 적대 네트워크는 생성 네트워크 및 판별 네트워크를 포함하고, 생성 네트워크는 제1 해상도 이미지를 제2 해상도 이미지로 변환하도록 구성되고, 상기 제2 해상도 이미지의 해상도는 상기 제1 해상도 이미지의 해상도보다 높고, 상기 트레이닝 방법은 생성 네트워크 트레이닝 절차를 포함하며, 상기 생성 네트워크 트레이닝 절차는:

제2 해상도 샘플 이미지로부터 제1 해상도 샘플 이미지를 추출하는 단계- 상기 제2 해상도 샘플 이미지의 해상도는 상기 제1 해상도 샘플 이미지의 해상도보다 높음 -;

상기 생성 네트워크에 제1 입력 이미지 및 제2 입력 이미지를 개별적으로 제공하여 상기 제1 입력 이미지에 기초한 제1 출력 이미지, 및 상기 제2 입력 이미지에 기초한 제2 출력 이미지를 각각 생성하는 단계; - 상기 제1 입력 이미지는 상기 제1 해상도 샘플 이미지 및 제1 진폭을 갖는 노이즈 샘플에 대응하는 제1 노이즈 이미지를 포함하고; 상기 제2 입력 이미지는 상기 제1 해상도 샘플 이미지 및 제2 진폭을 갖는 노이즈 샘플에 대응하는 제2 노이즈 이미지를 포함하고; 상기 제1 진폭은 0보다 크고, 상기 제2 진폭은 0과 동일함 -;

상기 판별 네트워크에 상기 제1 출력 이미지 및 상기 제2 해상도 샘플 이미지를 개별적으로 제공하여 상기 판별 네트워크가 상기 제1 출력 이미지에 기초한 제1 판별 결과, 및 상기 제2 해상도 샘플 이미지에 기초한 제2 판별 결과를 출력할 수 있게 하는 단계; 및

상기 생성 네트워크의 파라미터들을 조정하여 상기 생성 네트워크의 손실 함수를 감소시키는 단계- 상기 생성 네트워크의 상기 손실 함수는 제1 손실, 제2 손실 및 제3 손실을 포함하고, 상기 제1 손실은 상기 제2 출력 이미지와 상기 제2 해상도 샘플 이미지 사이의 재구성 오차에 기초하고; 상기 제2 손실은 상기 제1 출력 이미지와 상기 제2 해상도 샘플 이미지 사이의 지각 오차에 기초하고; 상기 제3 손실은 상기 제1 판별 결과 및 상기 제2 판별 결과에 기초함 -

를 포함하는 트레이닝 방법.

청구항 2

제1항에 있어서,

상기 제2 출력 이미지와 상기 제2 해상도 샘플 이미지 사이의 상기 재구성 오차는 상기 제2 출력 이미지와 상기 제2 해상도 샘플 이미지 사이의 차이 이미지 매트릭스(difference image matrix)의 L1 놈(norm), 상기 제2 출력 이미지와 상기 제2 해상도 샘플 이미지 사이의 평균 제곱 오차, 및 상기 제2 출력 이미지와 상기 제2 해상도 샘플 이미지 사이의 구조적 유사성 인덱스 중 임의의 하나에 따라 결정되는 트레이닝 방법.

청구항 3

제1항에 있어서,

상기 제1 출력 이미지 및 상기 제2 출력 이미지 둘 다는 해상도 향상 절차의 반복 프로세스를 통해 상기 생성 네트워크에 의해 생성되고, 상기 생성 네트워크의 상기 손실 함수의 상기 제1 손실은 $\lambda_1 L_{rec}(X, Y_{n=0})$ 이고,

$$L_{rec}(X, Y_{n=0}) = E \left[\sum_{l=1}^L \|Y_{n=0}^l - HR^l\|_1 \right] + E \left[\sum_{l=1}^L \|D_{bic}^l(Y_{n=0}^l) - LR\|_1 \right]$$

여기서 X는 상기 제2 해상도 샘플 이미지를 나타내고;

$Y_{n=0}$ 은 상기 제2 출력 이미지를 나타내고;

$L_{rec}(X, Y_{n=0})$ 은 상기 제2 출력 이미지와 상기 제2 해상도 샘플 이미지 사이의 재구성 오차를 나타내고;

L 은 상기 반복 프로세스에서의 상기 해상도 향상 절차의 총 횟수를 나타내고, $L \geq 1$ 이고;

$Y_{n=0}^l$ 은 상기 제2 입력 이미지에 기초하여 상기 생성 네트워크에 의해 수행되는 상기 반복 프로세스에서의 1번째 상기 해상도 향상 절차의 종료 시 생성된 이미지를 나타내고, $1 \leq L$ 이고;

LR 은 상기 제1 해상도 샘플 이미지를 나타내고;

$D_{bic}^l(Y_{n=0}^l)$ 은 $Y_{n=0}^l$ 을 다운샘플링함으로써 획득된 이미지를 나타내고, 상기 이미지의 해상도는 상기 제1 해상도 샘플 이미지의 해상도와 동일하고;

HR^l 은 상기 제2 해상도 샘플 이미지를 다운샘플링함으로써 획득된 이미지를 나타내고, 상기 이미지의 해상도는 $Y_{n=0}^l$ 의 해상도와 동일하고;

$E[]$ 는 매트릭스 에너지의 계산을 나타내고;

λ_1 은 미리 설정된 가중치인 트레이닝 방법.

청구항 4

제3항에 있어서,

상기 생성 네트워크의 상기 손실 함수의 상기 제2 손실은 $\lambda_2 L_{per}(X, Y_{n=1})$ 이고,

$$L_{per}(X, Y_{n=1}) = E \left[\sum_{l=1}^L L_{CX}(Y_{n=1}^l, HR^l) \right] + E \left[\sum_{l=1}^L L_{CX}(D_{bic}^l(Y_{n=1}^l), LR) \right]$$

여기서, $Y_{n=1}$ 은 상기 제1 출력 이미지를 나타내고;

$L_{per}(X, Y_{n=1})$ 은 상기 제1 출력 이미지와 상기 제2 해상도 샘플 이미지 사이의 상기 지각 오차를 나타내고;

$Y_{n=1}^l$ 은 상기 제1 입력 이미지에 기초하여 상기 생성 네트워크에 의해 수행되는 상기 반복 프로세스에서의 1번째 상기 해상도 향상 절차의 종료 시 생성된 이미지를 나타내고;

$D_{bic}^l(Y_{n=1}^l)$ 은 $Y_{n=1}^l$ 을 다운샘플링함으로써 획득된 이미지를 나타내고, 상기 이미지의 해상도는 상기 제1 해상도 샘플 이미지의 해상도와 동일하고;

$L_{CX}()$ 는 컨텍스트적 손실 계산 함수이고;

λ_2 는 미리 설정된 가중치인 트레이닝 방법.

청구항 5

제4항에 있어서,

상기 생성 네트워크의 상기 손실 함수의 상기 제3 손실은 $\lambda_3 L_{GAN}(Y_{n=1})$ 이고,

$$L_{GAN}(Y_{n=1}) = E \left[\log(1 - D(Y_{n=1}^{1, 2, \dots, L})) \right] + E \left[\log(D(HR^{1, 2, \dots, L})) \right]$$

여기서, $Y_{n=1}^{1, 2, \dots, L}$ 은 상기 제1 입력 이미지에 기초하여 상기 생성 네트워크에 의해 수행되는 상기 반복 프로세스에서 생성되는 이미지 그룹을 나타내고, 상기 이미지 그룹은 매회의 상기 해상도 향상 절차의 종료 시 각각 생성되는 이미지들을 포함하고;

$HR^{1, 2, \dots, L}$ 은 상기 제2 해상도 샘플 이미지를 다운샘플링함으로써 획득된 이미지들을 나타내고, 상기 이미지들

은 상기 $Y_{n=1}^{1, 2, \dots, L}$ 에서의 이미지들과 일대일 대응하고, 각각은 상기 대응하는 이미지의 해상도와 동일한 해상도를 갖고;

$D(Y_{n=1}^{1, 2, \dots, L})$ 는 상기 제1 판별 결과를 나타내고; $D(HR^{1, 2, \dots, L})$ 은 상기 제2 판별 결과를 나타내고;

λ_3 은 미리 설정된 가중치인 트레이닝 방법.

청구항 6

제5항에 있어서,

$\lambda_1 : \lambda_2 : \lambda_3 = 10 : 0.1 : 0.001$ 인 트레이닝 방법.

청구항 7

제1항에 있어서,

상기 노이즈 샘플은 랜덤 노이즈인 트레이닝 방법.

청구항 8

제1항에 있어서,

판별 네트워크 트레이닝 절차를 추가로 포함하고, 상기 판별 네트워크 트레이닝 절차는: 상기 판별 네트워크에 상기 제1 출력 이미지 및 상기 제2 해상도 샘플 이미지를 개별적으로 제공하여 상기 판별 네트워크가 상기 제1 출력 이미지에 기초한 판별 결과 및 상기 제2 해상도 샘플 이미지에 기초한 판별 결과를 각각 출력할 수 있게 하는 단계; 및 상기 판별 네트워크의 파라미터들을 조정하여 상기 판별 네트워크의 손실 함수를 감소시키는 단계

를 포함하고;

상기 판별 네트워크 트레이닝 절차 및 상기 생성 네트워크 트레이닝 절차는 미리 설정된 트레이닝 조건이 충족될 때까지 교대로 수행되는 트레이닝 방법.

청구항 9

제8항에 있어서,

상기 제1 출력 이미지 및 상기 제2 출력 이미지 둘 다는 해상도 향상 절차의 반복 프로세스를 통해 상기 생성 네트워크에 의해 생성되고, 상기 반복 프로세스에서의 상기 해상도 향상 절차의 총 횟수는 L이고; L이 1보다 클 때, 상기 제1 입력 이미지에 기초하여 상기 생성 네트워크에 의해 수행되는 상기 반복 프로세스에서의 이전 L-1회의 상기 해상도 향상 절차에서, 상기 생성 네트워크는 상기 해상도 향상 절차가 수행될 때마다 중간 이미지를 생성하고;

상기 판별 네트워크 트레이닝 절차에서, 상기 제1 출력 이미지가 상기 판별 네트워크에 제공되는 동안, 상기 제1 입력 이미지에 기초하여 상기 생성 네트워크에 의해 생성되는 각각의 중간 이미지가 상기 판별 네트워크에 제공되고; 상기 제2 해상도 샘플 이미지가 상기 판별 네트워크에 제공되는 동안, 상기 제2 해상도 샘플 이미지를 다운샘플링함으로써 획득된 제3 해상도 샘플 이미지들이 상기 판별 네트워크에 제공되고, 상기 제3 해상도 샘플 이미지들은 중간 이미지들과 일대일 대응하고, 각각은 상기 대응하는 중간 이미지의 해상도와 동일한 해상도를 갖는 트레이닝 방법.

청구항 10

제1항 내지 제9항 중 어느 한 항에 따른 상기 트레이닝 방법에 의해 획득된 상기 생성 적대 네트워크의 상기 생성 네트워크를 사용하는 이미지 처리 방법으로서,

상기 이미지 처리 방법은 이미지의 해상도를 증가시키기 위해 사용되고,

상기 생성 네트워크에 입력 이미지 및 기준 노이즈에 대응하는 노이즈 이미지를 제공하여 상기 생성 네트워크가

상기 입력 이미지에 기초하여 제2 해상도 이미지를 생성할 수 있게 하는 단계를 포함하는 이미지 처리 방법.

청구항 11

제10항에 있어서,

상기 기준 노이즈의 진폭은 0에서 제1 진폭까지의 범위인 이미지 처리 방법.

청구항 12

제10항에 있어서,

상기 기준 노이즈는 랜덤 노이즈인 이미지 처리 방법.

청구항 13

컴퓨터 프로그램들을 저장하는 메모리 및 프로세서를 포함하는 컴퓨터 디바이스로서,

상기 컴퓨터 프로그램들이 상기 프로세서에 의해 실행될 때 제1항 내지 제9항 중 어느 한 항에 따른 상기 트레이닝 방법이 수행되는 컴퓨터 디바이스.

청구항 14

컴퓨터 프로그램들을 저장하는 컴퓨터 판독가능 저장 매체로서,

상기 컴퓨터 프로그램들이 프로세서에 의해 실행될 때 제1항 내지 제9항 중 어느 한 항의 트레이닝 방법이 수행되는 컴퓨터 판독가능 저장 매체.

발명의 설명

기술 분야

[0001] 관련 출원에 대한 상호참조

[0002] 본 출원은 2018년 9월 30일자로 SIPO에 출원된, 발명의 명칭이 "TRAINING METHOD FOR GENERATIVE ADVERSARIAL NETWORK, IMAGE PROCESSING METHOD, DEVICE AND STORAGE MEDIUM"인 중국 특허 출원 제201811155930.9호의 우선권을 주장하고; 본 출원은 2018년 9월 30일자로 SIPO에 출원된, 발명의 명칭이 "IMAGE DISCRIMINATION METHOD, DISCRIMINATOR AND COMPUTER-READABLE STORAGE MEDIUM"인 중국 특허 출원 제201811155326.6호의 우선권을 주장하고; 본 출원은 2018년 9월 30일자로 SIPO에 출원된, 발명의 명칭이 "IMAGE PROCESSING METHOD AND SYSTEM, RESOLUTION ENHANCEMENT METHOD, AND READABLE STORAGE MEDIUM"인 중국 특허 출원 제201811155147.2호의 우선권을 주장하고; 본 출원은 2018년 9월 30일자로 SIPO에 출원된, 발명의 명칭이 "PREPROCESSING METHOD AND MODULE FOR TRAINING IMAGES, DISCRIMINATOR, AND READABLE STORAGE MEDIUM"인 중국 특허 출원 제201811155252.6호의 우선권을 주장하며, 그 개시내용은 본 명세서에 참조로 포함된다.

[0003] 기술 분야

[0004] 본 개시내용은 이미지 처리 분야에 관한 것이지만, 이에 제한되는 것은 아니며, 구체적으로는 생성 적대 네트워크(generative adversarial network)를 위한 트레이닝 방법, 트레이닝 방법에 의해 획득되는 생성 적대 네트워크를 사용하는 이미지 처리 방법, 컴퓨터 디바이스, 및 컴퓨터 판독가능 저장 매체에 관한 것이다.

배경 기술

[0005] 컨볼루션 신경 네트워크는 일반적인 딥 러닝 네트워크이고, 이미지 식별, 이미지 분류, 초해상도 이미지 재구성 등을 달성하기 위해 요즘 이미지 처리 분야에 널리 적용되어 왔다.

[0006] 현재의 초해상도 재구성 방법들에서, 제1 해상도 이미지에 기초하여 재구성된 제2 해상도 이미지(제2 해상도 이미지의 해상도는 제1 해상도 이미지의 해상도보다 높음)는 보통 상세 정보가 없고, 이는 제2 해상도 이미지가 비현실적으로 보이게 한다.

발명의 내용

[0007] 종래 기술에서의 기술적 문제점들 중 적어도 하나를 해결하기 위해, 본 개시내용은 생성 적대 네트워크를 위한

트레이닝 방법, 트레이닝 방법에 의해 획득된 생성 적대 네트워크를 사용하는 이미지 처리 방법, 컴퓨터 디바이스, 및 컴퓨터 판독가능 저장 매체를 제공한다.

- [0008] 위의 언급된 기술적 문제들 중 하나를 해결하기 위해, 본 개시내용은 생성 적대 네트워크를 위한 트레이닝 방법을 제공하고, 생성 적대 네트워크는 생성 네트워크 및 판별 네트워크를 포함하고, 생성 네트워크는 제1 해상도 이미지를 제2 해상도 이미지로 변환하도록 구성되고, 제2 해상도 이미지의 해상도는 제1 해상도 이미지의 해상도보다 높고, 트레이닝 방법은 생성 네트워크 트레이닝 절차를 포함하며, 이 생성 네트워크 트레이닝 절차는 다음을 포함한다:
- [0009] 제2 해상도 샘플 이미지로부터 제1 해상도 샘플 이미지를 추출하는 단계- 제2 해상도 샘플 이미지의 해상도는 제1 해상도 샘플 이미지의 해상도보다 높음 -;
- [0010] 생성 네트워크에 제1 입력 이미지 및 제2 입력 이미지를 개별적으로 제공하여 제1 입력 이미지에 기초한 제1 출력 이미지, 및 제2 입력 이미지에 기초한 제2 출력 이미지를 각각 생성하는 단계; - 제1 입력 이미지는 제1 해상도 샘플 이미지 및 제1 진폭을 갖는 노이즈 샘플에 대응하는 제1 노이즈 이미지를 포함하고, 제2 입력 이미지는 제1 해상도 샘플 이미지 및 제2 진폭을 갖는 노이즈 샘플에 대응하는 제2 노이즈 이미지를 포함하고; 제1 진폭은 0보다 크고, 제2 진폭은 0과 동일함 -;
- [0011] 판별 네트워크에 제1 출력 이미지 및 제2 해상도 샘플 이미지를 개별적으로 제공하여 판별 네트워크가 제1 출력 이미지에 기초한 제1 판별 결과, 및 제2 해상도 샘플 이미지에 기초한 제2 판별 결과를 출력할 수 있게 하는 단계; 및
- [0012] 생성 네트워크의 파라미터들을 조정하여 생성 네트워크의 손실 함수를 감소시키는 단계; - 생성 네트워크의 손실 함수는 제1 손실, 제2 손실 및 제3 손실을 포함하고, 손실 함수의 제1 손실은 제2 출력 이미지와 제2 해상도 샘플 이미지 사이의 재구성 오차에 기초하고; 손실 함수의 제2 손실은 제1 출력 이미지와 제2 해상도 샘플 이미지 사이의 지각 오차에 기초하고; 손실 함수의 제3 손실은 제1 판별 결과 및 제2 판별 결과에 기초함 -.
- [0013] 일부 구현들에서, 제2 출력 이미지와 제2 해상도 샘플 이미지 사이의 재구성 오차는 제2 출력 이미지와 제2 해상도 샘플 이미지 사이의 차이 이미지 매트릭스(difference image matrix)의 L1 놈(norm), 제2 출력 이미지와 제2 해상도 샘플 이미지 사이의 평균 제곱 오차, 및 제2 출력 이미지와 제2 해상도 샘플 이미지 사이의 구조적 유사성 인덱스 중 임의의 하나에 따라 결정된다.
- [0014] 일부 구현들에서, 제1 출력 이미지 및 제2 출력 이미지 둘 다는 해상도 향상 절차의 반복 프로세스를 통해 생성 네트워크에 의해 생성되고; 생성 네트워크의 손실 함수의 제1 손실은 $\lambda_1 L_{rec}(X, Y_{n=0})$ 이고,
- $$L_{rec}(X, Y_{n=0}) = E \left[\sum_{l=1}^L \|Y_{n=0}^l - HR^l\|_1 \right] + E \left[\sum_{l=1}^L \|D_{bic}^l(Y_{n=0}^l) - LR\|_1 \right]$$
- [0015]
- [0016] 여기서, X는 제2 해상도 샘플 이미지를 나타내고;
- [0017] $Y_{n=0}$ 은 제2 출력 이미지를 나타내고;
- [0018] $L_{rec}(X, Y_{n=0})$ 은 제2 출력 이미지와 제2 해상도 샘플 이미지 사이의 재구성 오차를 나타내고;
- [0019] L은 반복 프로세스에서의 해상도 향상 절차의 총 횟수를 나타내고; $L \geq 1$ 이고,
- [0020] $Y_{n=0}^l$ 은 제2 입력 이미지에 기초하여 생성 네트워크에 의해 수행되는 반복 프로세스에서의 l번째 해상도 향상 절차의 종료 시 생성된 이미지를 나타내고; $1 \leq L$ 이고;
- [0021] LR은 제1 해상도 샘플 이미지를 나타내고;
- [0022] $D_{bic}^l(Y_{n=0}^l)$ 는 $Y_{n=0}^l$ 를 다운샘플링함으로써 획득된 이미지를 나타내고, 이미지의 해상도는 제1 해상도 샘플 이미지의 해상도와 동일하고;
- [0023] HR^l 은 제2 해상도 샘플 이미지를 다운샘플링함으로써 획득된 이미지를 나타내고, 이미지의 해상도는 $Y_{n=0}^l$ 의 해상도와 동일하고;

- [0024] $E[\]$ 는 매트릭스 에너지의 계산을 나타내고;
- [0025] λ_1 은 미리 설정된 가중치이다.
- [0026] 일부 구현들에서, 생성 네트워크의 손실 함수의 제2 손실은 $\lambda_2 L_{per}(X, Y_{n=1})$ 이고,
- [0027]
$$L_{per}(X, Y_{n=1}) = E \left[\sum_{l=1}^L L_{CX}(Y_{n=1}^l, HR^l) \right] + E \left[\sum_{l=1}^L L_{CX}(D_{bic}^l(Y_{n=1}^l), LR) \right]$$
- [0028] 여기서, $Y_{n=1}$ 은 제1 출력 이미지를 나타내고;
- [0029] $L_{per}(X, Y_{n=1})$ 은 제1 출력 이미지와 제2 해상도 샘플 이미지 사이의 지각 오차를 나타내고;
- [0030] $Y_{n=1}^l$ 은 제1 입력 이미지에 기초하여 생성 네트워크에 의해 수행되는 반복 프로세스에서의 1번째 해상도 향상 절차의 종료 시 생성된 이미지를 나타내고;
- [0031] $D_{bic}^l(Y_{n=1}^l)$ 는 $Y_{n=1}^l$ 를 다운샘플링함으로써 획득된 이미지를 나타내고, 이미지의 해상도는 제1 해상도 샘플 이미지의 해상도와 동일하고;
- [0032] $L_{CX}(\)$ 는 컨텍스트적 손실 계산 함수이고;
- [0033] λ_2 는 미리 설정된 가중치이다.
- [0034] 일부 구현들에서, 생성 네트워크의 손실 함수의 제3 손실은 $\lambda_3 L_{GAN}(Y_{n=1})$ 이고,
- [0035]
$$L_{GAN}(Y_{n=1}) = E \left[\log(1 - D(Y_{n=1}^{1, 2, \dots, L})) \right] + E \left[\log(D(HR^{1, 2, \dots, L})) \right]$$
- [0036] 여기서, $Y_{n=1}^{1, 2, \dots, L}$ 은 제1 입력 이미지에 기초하여 생성 네트워크에 의해 수행되는 반복 프로세스에서 생성되는 이미지 그룹을 나타내고, 이미지 그룹은 매회의 해상도 향상 절차의 종료 시 각각 생성되는 이미지들을 포함하고;
- [0037] $HR^{1, 2, \dots, L}$ 은 제2 해상도 샘플 이미지를 다운샘플링함으로써 획득된 이미지들을 나타내고, 이미지들은 $Y_{n=1}^{1, 2, \dots, L}$ 에서의 이미지들과 일대일 대응하고, 각각은 대응하는 이미지의 해상도와 동일한 해상도를 갖고;
- [0038] $D(Y_{n=1}^{1, 2, \dots, L})$ 는 제1 판별 결과를 나타내고;
- [0039] $D(HR^{1, 2, \dots, L})$ 는 제2 판별 결과를 나타내고;
- [0040] λ_3 은 미리 설정된 가중치이다.
- [0041] 일부 구현들에서, $\lambda_1 : \lambda_2 : \lambda_3 = 10 : 0.1 : 0.001$ 이다.
- [0042] 일부 구현들에서, 노이즈 샘플은 랜덤 노이즈이다. 일부 구현들에서, 트레이닝 방법은 판별 네트워크 트레이닝 절차를 추가로 포함하고, 이 판별 네트워크 트레이닝 절차는: 판별 네트워크에 제1 출력 이미지 및 제2 해상도 샘플 이미지를 개별적으로 제공하여 판별 네트워크가 제1 출력 이미지에 기초한 판별 결과 및 제2 해상도 샘플 이미지에 기초한 판별 결과를 각각 출력할 수 있게 하는 단계; 및 판별 네트워크의 파라미터들을 조정하여 판별 네트워크의 손실 함수를 감소시키는 단계를 포함하고;
- [0043] 판별 네트워크 트레이닝 절차 및 생성 네트워크 트레이닝 절차는 미리 설정된 트레이닝 조건이 충족될 때까지 교대로 수행된다.
- [0044] 일부 구현들에서, 제1 출력 이미지 및 제2 출력 이미지 둘 다는 해상도 향상 절차의 반복 프로세스를 통해 생성 네트워크에 의해 생성되고, 반복 프로세스에서의 해상도 향상 절차의 총 횟수는 L 이고; L 이 1보다 클 때, 제1 입력 이미지에 기초하여 생성 네트워크에 의해 수행되는 반복 프로세스에서의 이전 $L-1$ 회의 해상도 향상 절차에

서, 생성 네트워크는 해상도 향상 절차가 수행될 때마다 중간 이미지를 생성하고;

[0045] 판별 네트워크 트레이닝 절차에서, 제1 출력 이미지가 판별 네트워크에 제공되는 동안, 제1 입력 이미지에 기초하여 생성 네트워크에 의해 생성되는 각각의 중간 이미지가 판별 네트워크에 제공되고; 제2 해상도 샘플 이미지가 판별 네트워크에 제공되는 동안, 그리고 제2 해상도 샘플 이미지를 다운샘플링함으로써 획득된 제3 해상도 샘플 이미지들이 판별 네트워크에 제공되고, 제3 해상도 샘플 이미지들은 중간 이미지들과 일대일 대응하고, 각각은 대응하는 중간 이미지의 해상도와 동일한 해상도를 갖는다.

[0046] 따라서, 본 개시내용은 트레이닝 방법에 의해 획득되는 생성 적대 네트워크의 생성 네트워크를 사용하는 이미지 처리 방법을 추가로 제공하고, 이미지 처리 방법은 이미지의 해상도를 증가시키기 위해 사용되고, 다음을 포함한다:

[0047] 생성 네트워크에 입력 이미지 및 기준 노이즈에 대응하는 노이즈 이미지를 제공하여 생성 네트워크가 입력 이미지에 기초하여 제2 해상도 이미지를 생성할 수 있게 하는 단계.

[0048] 일부 구현들에서, 기준 노이즈의 진폭은 0에서 제1 진폭까지의 범위이다.

[0049] 일부 구현들에서, 기준 노이즈는 랜덤 노이즈이다.

[0050] 따라서, 본 개시내용은 컴퓨터 프로그램들을 저장하는 메모리 및 프로세서를 포함하는 컴퓨터 디바이스를 추가로 제공하고, 컴퓨터 프로그램들이 프로세서에 의해 실행될 때 위의 트레이닝 방법이 수행된다.

[0051] 따라서, 본 개시내용은 컴퓨터 프로그램들을 저장하는 컴퓨터 판독가능 저장 매체를 추가로 제공하고, 컴퓨터 프로그램들이 프로세서에 의해 실행될 때 위의 트레이닝 방법이 수행된다.

도면의 간단한 설명

[0052] 첨부 도면들은 본 개시내용의 추가의 이해를 제공하도록 의도되고, 본 명세서 내에 포함되고 그 일부를 구성한다. 도면들은, 이하의 특정 실시예들과 함께, 본 개시내용을 설명하도록 의도되지만, 본 개시내용에 임의의 제한을 행하지 않는다. 도면들에서:

도 1은 재구성 왜곡과 지각 왜곡 사이의 관계를 나타내는 개략도이다;

도 2는 본 개시내용의 실시예들에 따른 생성 네트워크 트레이닝 절차를 예시하는 흐름도이다; 및

도 3은 본 개시내용의 실시예들에 따른 생성 네트워크의 개략 구조도이다.

발명을 실시하기 위한 구체적인 내용

[0053] 본 개시내용의 특정 실시예들은 첨부 도면들을 참조하여 이하에서 상세하게 설명될 것이다. 본 명세서에 설명된 특정 실시예들은 단지 본 개시내용을 예시하고 설명하기 위한 것이며, 본 개시내용에 임의의 제한을 행하지 않는다는 것을 이해해야 한다.

[0054] 초해상도 이미지 재구성은 더 높은 해상도를 갖는 이미지를 획득하기 위해 초기 이미지의 해상도를 증가시키기 위한 기술이다. 초해상도 이미지 재구성에서, 재구성 왜곡 및 지각 왜곡은 초해상도 재구성 효과를 평가하기 위해 사용된다. 재구성 왜곡은 재구성 이미지와 기준 이미지 사이의 차이를 측정하기 위해 사용되고, 특정 평가 기준들은 평균 제곱 오차(MSE), 구조적 유사성 인덱스(SSIM), 및 피크 신호 대 노이즈비(PSNR)를 포함하고; 지각 왜곡은 주로 이미지가 더 자연 이미지처럼 보이게 하는 것에 초점을 맞춘다. 도 1은 재구성 왜곡과 지각 왜곡 사이의 관계를 나타내는 개략도이다. 도 1에 도시된 바와 같이, 재구성 왜곡이 비교적 작을 때, 지각 왜곡은 비교적 크고, 이 경우 재구성된 이미지는 더 평활하지만 상세들은 없다. 지각 왜곡이 비교적 작을 때, 재구성 왜곡은 비교적 크고, 이 경우 재구성된 이미지는 더 많은 상세들을 갖는다. 현재의 초해상도 이미지 재구성 방법들은 일반적으로 비교적 작은 재구성 왜곡을 목표로 하지만, 사람들은 일부 응용 시나리오들에서 풍부한 상세들을 갖는 재구성된 이미지들을 획득하는 것을 선호한다.

[0055] 본 개시내용은 생성 적대 네트워크를 위한 트레이닝 방법을 제공하고, 생성 적대 네트워크는 생성 네트워크 및 판별 네트워크를 포함하고, 생성 네트워크는 제1 해상도 이미지를 제2 해상도 이미지로 변환하여 타겟 해상도를 갖는 제2 해상도 이미지를 획득하도록 구성되고, 제2 해상도 이미지의 해상도는 제1 해상도 이미지의 해상도보다 높다. 생성 네트워크는 해상도 향상 절차를 한번 수행함으로써 또는 해상도 향상 절차를 복수 회 반복함으로써 제2 해상도 이미지를 획득할 수 있다. 예를 들어, 처리될 이미지(즉, 제1 해상도 이미지)는 128x128의 해

상도를 갖고, 타겟 해상도는 1024x1024이고, 생성 네트워크는 해상도를 8배 증가시키는 해상도 향상 절차를 한번 수행함으로써 1024x1024의 해상도를 갖는 제2 해상도 이미지를 획득할 수 있거나; 또는 생성 네트워크는 해상도 향상 절차를 반복함으로써 순차적으로 256x256 이미지, 512x512 이미지 및 1024x1024 이미지를 획득할 수 있으며, 이는 해상도를 2배, 3배 증가시킨다.

[0056] 생성 적대 네트워크를 위한 트레이닝 방법은 생성 네트워크 트레이닝 절차를 포함한다. 도 2는 본 개시내용의 실시예들에 따른 생성 네트워크 트레이닝 절차를 예시하는 흐름도이다. 도 2에 도시된 바와 같이, 생성 네트워크 트레이닝 절차는 다음의 S1 내지 S4를 포함한다.

[0057] S1, 제2 해상도 샘플 이미지로부터 제1 해상도 샘플 이미지를 추출하는 단계- 제2 해상도 샘플 이미지의 해상도는 제1 해상도 샘플 이미지의 해상도보다 높음 -. 구체적으로, 제1 해상도 샘플 이미지는 제2 해상도 샘플 이미지를 다운샘플링함으로써 획득될 수 있다.

[0058] S2, 생성 네트워크에 제1 입력 이미지 및 제2 입력 이미지를 개별적으로 제공하여 제1 입력 이미지에 기초한 제1 출력 이미지, 및 제2 입력 이미지에 기초한 제2 출력 이미지를 각각 생성하는 단계- 제1 입력 이미지는 제1 해상도 샘플 이미지 및 제1 진폭을 갖는 노이즈 샘플에 대응하는 제1 노이즈 이미지를 포함하고, 제2 입력 이미지는 제1 해상도 샘플 이미지 및 제2 진폭을 갖는 노이즈 샘플에 대응하는 제2 노이즈 이미지를 포함함 -. 제1 진폭은 0보다 크고, 제2 진폭은 0과 동일하다.

[0059] 노이즈 샘플의 진폭은 노이즈 샘플의 평균 요동 진폭이다. 예를 들어, 노이즈 샘플은 랜덤 노이즈이고, 노이즈 샘플에 대응하는 이미지의 평균은 μ 이고, 노이즈 샘플에 대응하는 이미지의 분산은 σ 인데, 즉, 노이즈 샘플에 대응하는 이미지의 대부분의 픽셀 값들은 $\mu - \sigma$ 로부터 $\mu + \sigma$ 까지 요동하고, 이 경우 노이즈 진폭은 μ 이다. 이미지 처리 프로세스에서 임의의 이미지가 매트릭스의 형태로 도시되고, 픽셀 값들이 이미지 매트릭스의 요소 값들을 나타낸다는 점을 이해해야 한다. 노이즈 샘플의 진폭이 0일 때, 이미지 매트릭스의 요소 값이 0보다 작지 않기 때문에, 이미지 매트릭스의 각각의 요소 값은 0인 것으로 간주될 수 있다.

[0060] 또한, 생성 적대 네트워크를 위한 트레이닝 방법은 복수의 생성 네트워크 트레이닝 절차를 포함하고; 단일 생성 네트워크 트레이닝 절차에서, 제1 해상도 샘플 이미지는 단일의 하나이고, 제1 입력 이미지 및 제2 입력 이미지를 수신할 때 생성 네트워크의 모델 파라미터들은 동일하다는 점에 유의해야 한다.

[0061] S3, 판별 네트워크에 제1 출력 이미지 및 제2 해상도 샘플 이미지를 개별적으로 제공하여 판별 네트워크가 제1 출력 이미지에 기초한 제1 판별 결과, 및 제2 해상도 샘플 이미지에 기초한 제2 판별 결과를 출력할 수 있게 하는 단계. 제1 판별 결과는 제1 출력 이미지와 제2 해상도 샘플 이미지 사이의 매칭 정도를 나타내기 위해 사용되는데, 예를 들어, 제1 판별 결과는 제1 출력 이미지가 제2 해상도 샘플 이미지와 동일한 것으로 판별 네트워크에 의해 결정되는 확률을 나타내기 위해 사용되고; 제2 판별 결과는 제2 해상도 샘플 이미지가 실제로 제2 해상도 샘플 이미지인 것으로 판별 네트워크에 의해 결정되는 확률을 나타내기 위해 사용된다.

[0062] 판별 네트워크는 스코어링 기능을 갖는 분류자로서 간주될 수 있다. 판별 네트워크는 수신된 판별될 이미지를 스코어링할 수 있고, 판별될 이미지(제1 출력 이미지)가 제2 해상도 샘플 이미지와 동일할 확률을 나타내는 스코어, 즉, 0에서 1까지의 범위일 수 있는, 위에서 언급된 매칭 정도를 나타내는 스코어를 출력한다. 판별 네트워크의 출력 스코어가 0이거나 0에 가까운 경우, 판별 네트워크가 수신된 판별될 이미지를 비-고해상도 샘플 이미지로서 분류하는 것으로 표시되고; 판별 네트워크의 출력 스코어가 1이거나 1에 가까운 경우, 수신된 판별될 이미지는 제2 해상도 샘플 이미지와 동일한 것으로 표시된다.

[0063] 판별 네트워크의 스코어링 기능은 미리 결정된 스코어들을 갖는 "참" 샘플 및 "거짓" 샘플의 사용에 의해 트레이닝될 수 있다. 예를 들어, "거짓" 샘플은 생성 네트워크에 의해 생성된 이미지이고, "참" 샘플은 제2 해상도 샘플 이미지이다. 판별 네트워크의 트레이닝 프로세스는 판별 네트워크가 "참" 샘플을 수신할 때 1에 가까운 스코어를 출력할 수 있게 하고, "거짓" 샘플을 수신할 때 0에 가까운 스코어를 출력할 수 있게 하기 위해 판별 네트워크의 파라미터들을 조정하는 프로세스이다.

[0064] S4, 생성 네트워크의 파라미터들을 조정하여 생성 네트워크의 손실 함수를 감소시키는 단계. "생성 네트워크의 손실 함수를 감소시키는 것"은 손실 함수의 값이 이전의 생성 네트워크 트레이닝 절차에서의 것과 비교하여 감소되거나, 복수의 생성 네트워크 트레이닝 절차에서의 손실 함수의 값들이 감소하는 추세를 나타낸다는 것을 의미한다. 생성 네트워크의 손실 함수는 제1 손실, 제2 손실, 및 제3 손실을 포함하고; 구체적으로, 손실 함수는 제1 손실, 제2 손실 및 제3 손실의 중첩이고, 제1 손실은 제2 출력 이미지와 제2 해상도 샘플 이미지 사이의 재구성 오차에 기초하고; 제2 손실은 제1 출력 이미지와 제2 해상도 샘플 이미지 사이의 지각 오차에 기초하고;

제3 손실은 제1 판별 결과 및 제2 판별 결과에 기초한다.

[0065] 초해상도 재구성에서, 재구성된 제2 해상도 이미지에서의 상세 특징들(예를 들어, 머리카락들, 라인들 등)은 일반적으로 노이즈와 관련된다. 생성 네트워크의 트레이닝에 노이즈가 추가되지 않을 때, 생성 네트워크에 의해 생성된 제2 해상도 이미지는 작은 재구성 왜곡 및 큰 지각 왜곡을 가져서, 제2 해상도 이미지는 육안으로 비현실적으로 보이게 되고; 생성 네트워크의 트레이닝에 노이즈가 추가될 때, 재구성된 제2 해상도 이미지는 명백한 상세 특징들을 갖지만, 그 재구성 왜곡은 비교적 크다. 본 개시내용의 생성 네트워크 트레이닝 절차에서, 0의 진폭을 갖는 노이즈 이미지를 포함하는 제2 입력 이미지 및 1의 진폭을 갖는 노이즈 이미지를 포함하는 제1 입력 이미지는 트레이닝을 위해 생성 네트워크에 개별적으로 제공되고, 손실 함수의 제1 손실은 생성 네트워크에 의해 생성된 결과의 재구성 왜곡을 반영하고, 제2 손실은 생성 네트워크에 의해 생성된 결과의 지각 왜곡을 반영하는데, 즉 손실 함수는 2개의 왜곡 평가 기준을 결합한다. 트레이닝된 생성 네트워크가 이미지에 대해 해상도 향상을 수행하기 위해 사용될 때, 입력 노이즈의 진폭은 실제 요구들(즉, 이미지의 상세들이 강조될 필요가 있는지 그리고 상세들이 어느 정도로 강조되는지)에 따라 조정될 수 있어서, 재구성된 이미지가 실제 요구들을 충족시킬 수 있게 된다. 예를 들어, 재구성 왜곡의 주어진 범위 내에서, 최소 지각 왜곡은 입력 노이즈의 진폭을 조정함으로써 달성되거나; 또는 지각 왜곡의 주어진 범위 내에서, 최소 재구성 왜곡은 입력 노이즈의 진폭을 조정함으로써 달성된다.

[0066] 실시예에서 1인, 제1 입력 이미지의 노이즈 이미지의 진폭은 노이즈 이미지의 진폭을 정규화함으로써 획득되는 진폭 값이라는 점에 유의해야 한다. 본 개시내용의 다른 실시예들에서, 제1 입력 이미지의 노이즈 이미지의 진폭이 1과 동일하지 않을 수 있도록, 노이즈 이미지의 진폭을 정규화하지 않는 것이 가능하다.

[0067] 일부 구현들에서, 노이즈 샘플은 랜덤 노이즈이고; 제1 노이즈 이미지의 평균은 1이다. 일부 구현에서, 제1 노이즈 이미지의 평균은 제1 노이즈 이미지의 정규화된 이미지의 평균이다. 예를 들어, 제1 노이즈 이미지가 그레이스케일 이미지인 경우, 제1 노이즈 이미지를 정규화함으로써 획득된 이미지에서의 모든 픽셀 값들의 평균은 제1 노이즈 이미지의 평균이고; 다른 예로서, 제1 노이즈 이미지가 컬러 이미지인 경우, 제1 노이즈 이미지의 모든 채널을 정규화함으로써 획득되는 이미지에서의 모든 픽셀 값들의 평균은 제1 노이즈 이미지의 평균이다. 본 개시내용의 실시예에서의 이미지의 채널은 처리를 위해 이미지를 분할함으로써 획득되는 하나 이상의 채널을 나타내는데, 예를 들어, RGB-모드 컬러 이미지는 3개의 채널, 즉, 적색 채널, 녹색 채널, 및 청색 채널로 분할될 수 있고; 이미지가 그레이스케일 이미지인 경우, 그것은 1-채널 이미지이고; 컬러 이미지가 HSV 컬러 시스템에 따라 분할되는 경우, 이미지는 3개의 채널, 즉, 색조(H) 채널, 채도(S) 채널, 및 값(V) 채널로 분할될 수 있다는 점에 유의해야 한다.

[0068] 일부 구현들에서, 생성 네트워크의 손실 함수는 다음과 같은 수학적식으로 표현되고:

$$[0069] \text{Loss} = \lambda_1 L_{rec}(X, Y_{n=0}) + \lambda_2 L_{per}(X, Y_{n=1}) + \lambda_3 L_{GAN}(Y_{n=1})$$

[0070] 여기서, 손실 함수 Loss의 제1 손실 $\lambda_1 L_{rec}(X, Y_{n=0})$ 에서, $L_{rec}(X, Y_{n=0})$ 은 제2 출력 이미지와 제2 해상도 샘플 이미지 사이의 재구성 오차를 나타내고; 손실 함수 Loss의 제2 손실 $\lambda_2 L_{per}(X, Y_{n=1})$ 에서, $L_{per}(X, Y_{n=1})$ 은 제2 출력 이미지와 제2 해상도 샘플 이미지 사이의 지각 오차를 나타내고; 손실 함수 Loss의 제3 손실 $\lambda_3 L_{GAN}(Y_{n=1})$ 에서, $L_{GAN}(Y_{n=1})$ 은 제1 판별 결과와 제2 판별 결과의 합을 나타내고; λ_1 , λ_2 , λ_3 은 모두 미리 설정된 가중치들이다. $\lambda_1 : \lambda_2 : \lambda_3$ 은 실제 요구들에 따라 조정될 수 있는데, 예를 들어, $\lambda_1 : \lambda_2 : \lambda_3 = 10 : 0.1 : 0.001$ 이거나, $\lambda_1 : \lambda_2 : \lambda_3 = 1 : 1 : 0.5$ 이다. 일부 실시예들에서, $\lambda_1 : \lambda_2 : \lambda_3$ 은 로컬 이미지들의 연속성에 따라 설정될 수 있다. 한편 일부 다른 실시예들에서, $\lambda_1 : \lambda_2 : \lambda_3$ 은 이미지의 타겟 픽셀들에 따라 설정될 수 있다.

[0071] 구체적으로, 제2 출력 이미지 $Y_{n=0}$ 과 제2 해상도 샘플 이미지 X 사이의 재구성 오차 $L_{rec}(X, Y_{n=0})$ 은 다음의 수학적식:

$$[0072] L_{rec}(X, Y_{n=0}) = E \left[\sum_{l=1}^L \|Y_{n=0}^l - HR^l\|_1 \right] + E \left[\sum_{l=1}^L \|D_{bic}^l(Y_{n=0}^l) - LR\|_1 \right]$$

[0073] 에 따라 계산되고, 여기서 제1 출력 이미지 및 제2 출력 이미지 둘 다는 해상도 향상 절차의 반복 프로세스를

통해 생성 네트워크에 의해 생성되고; 반복 프로세스에서의 해상도 향상 절차의 총 횟수는 L 이고, $L \geq 1$ 이다.

[0074] $Y_{n=0}^l$ 은 제2 입력 이미지에 기초하여 생성 네트워크에 의해 수행되는 반복 프로세스에서의 1번째 해상도 향상 절차의 종료 시 생성된 이미지를 나타내고; $1 \leq L$ 이다. 생성 네트워크는 $1=L$ 일 때 제2 출력 이미지 $Y_{n=0}$ 을 생성한다는 것을 이해해야 한다.

[0075] LR은 제1 해상도 샘플 이미지를 나타내고; $D_{bic}^l(Y_{n=0}^l)$ 은 $Y_{n=0}^l$ 을 다운샘플링함으로써 획득된 이미지를 나타내고, 이미지의 해상도는 제1 해상도 샘플 이미지의 해상도와 동일하다. 다운샘플링은 단계 S1에서 제2 해상도 샘플 이미지로부터 제1 해상도 샘플 이미지를 추출하기 위한 방식과 동일한 방식으로 수행될 수 있다.

[0076] HR^l 은 제2 해상도 샘플 이미지를 다운샘플링함으로써 획득된 이미지를 나타내고, 이미지의 해상도는 $Y_{n=0}^l$ 의 해상도와 동일하다. $1=L$ 일 때, $Y_{n=0}^l$ 이 제2 출력 이미지 $Y_{n=0}$ 이고, HR^l 이 제2 해상도 샘플 이미지 자체이거나, 또는 제2 해상도 샘플 이미지를 1회 다운샘플링함으로써 획득된 이미지로서 간주될 수 있다는 점에 유의해야 한다.

[0077] $E[\]$ 는 매트릭스 에너지의 계산을 나타낸다. 예를 들어, $E[\]$ 는 "[]"에서의 매트릭스 내의 요소들의 최대 또는 평균을 계산할 수 있다.

[0078] 생성 네트워크에 의해 해상도 향상 절차를 복수 회 반복하는 경우에, 재구성 오차를 계산할 때, 제2 출력 이미지 자체와 제2 해상도 샘플 이미지 사이의 차이 이미지 매트릭스의 $L1$ 노름이 계산되고, 생성 네트워크에 의해 생성된 제3 해상도 이미지(즉, $Y_{n=0}^1, Y_{n=0}^2, \dots, Y_{n=0}^{L-1}$)와 제3 해상도 샘플 이미지(즉, $HR^1, HR^2, \dots, HR^{L-1}$) 사이의 차이 이미지 매트릭스의 $L1$ 노름이 또한 계산되고, 제3 해상도 샘플 이미지의 해상도는 제3 해상도 이미지의 해상도와 동일하다. 동시에, 제2 출력 이미지를 다운샘플링함으로써 생성된 이미지인 제3 해상도 이미지와 제1 해상도 샘플 이미지 사이의 차이 이미지의 $L1$ 노름이 또한 계산된다. 이러한 방식으로, 생성 네트워크가 해상도 향상을 위해 사용되고 0의 진폭을 갖는 노이즈가 입력될 때, 생성 네트워크에 의해 최종적으로 출력되는 이미지는 최소 재구성 왜곡을 달성할 수 있다. 제3 해상도 이미지의 해상도는 제1 해상도 샘플 이미지의 해상도보다 높고, 제3 해상도 샘플 이미지의 해상도와 동일하다는 점에 유의해야 한다.

[0079] 위의 실시예에서, 제2 출력 이미지와 제2 해상도 샘플 이미지 사이의 재구성 오차 $L_{rec}(X, Y_{n=0})$ 은 제2 출력 이미지와 제2 해상도 샘플 이미지 사이의 차이 이미지 매트릭스의 $L1$ 노름에 기초하여 획득되거나, 재구성 오차는 제2 출력 이미지와 제2 해상도 샘플 이미지 사이의 평균 제곱 오차(MSE)에 기초하여 획득될 수 있거나, 또는 제2 출력 이미지와 제2 해상도 샘플 이미지 사이의 구조적 유사성 인덱스(SSIM)에 기초하여 획득될 수 있다.

[0080] 일부 구현들에서, 제1 출력 이미지 $Y_{n=1}$ 과 제2 해상도 샘플 이미지 X 사이의 지각 오차 $L_{per}(X, Y_{n=1})$ 은 다음의 수학적식:

$$L_{per}(X, Y_{n=1}) = E \left[\sum_{l=1}^L L_{CX}(Y_{n=1}^l, HR^l) \right] + E \left[\sum_{l=1}^L L_{CX}(D_{bic}^l(Y_{n=1}^l), LR) \right]$$

[0082] 에 따라 계산되고, $Y_{n=1}^l$ 은 제1 입력 이미지에 기초하여 생성 네트워크에 의해 수행되는 반복 프로세스에서의 1번째 해상도 향상 절차의 종료 시 생성된 이미지를 나타내고; $1 \leq L$ 이다. 생성 네트워크는 $1=L$ 일 때 제1 출력 이미지 $Y_{n=1}$ 을 생성한다는 것을 이해해야 한다.

[0083] $D_{bic}^l(Y_{n=1}^l)$ 은 $Y_{n=1}^l$ 을 다운샘플링함으로써 획득된 이미지를 나타내고, 이미지의 해상도는 제1 해상도 샘플 이미지 LR의 해상도와 동일하다. 다운샘플링은 단계 S1에서 제2 해상도 샘플 이미지로부터 제1 해상도 샘플 이미지를 추출하기 위한 방식과 동일한 방식으로 수행될 수 있다. HR^l 및 $E[\]$ 의 의미들에 대한 위의 설명에 대해 참조가 이루어질 수 있고, 이는 본 명세서에서 반복되지 않을 것이다.

[0084] $L_{CX}(\)$ 는 컨텍스트적 손실 계산 함수를 나타낸다.

[0085] 재구성 오차의 계산과 유사하게, 지각 오차의 계산은 컨텍스트적 손실 계산 함수의 사용에 의해 제1 출력 이미

지와 제2 해상도 샘플 이미지 사이의 차이의 계산을 수반할 뿐만 아니라, 제1 입력 이미지에 기초하여 생성 네트워크에 의해 생성된 제3 해상도 이미지(즉, $Y_{n=1}^1, Y_{n=1}^2, \dots, Y_{n=1}^{L-1}$)와 제3 해상도 샘플 이미지(즉, $HR^1, HR^2, \dots, HR^{L-1}$) 사이의 차이의 계산을 수반하고, 제3 해상도 샘플 이미지의 해상도는 제3 해상도 이미지의 해상도와 동일하고, 제2 출력 이미지를 다운샘플링함으로써 생성된 이미지인 제3 해상도 이미지와, 제1 해상도 샘플 이미지 사이의 차이의 계산을 추가로 수반한다. 이러한 방식으로, 생성 네트워크가 해상도 향상을 위해 사용되고 제1 진폭을 갖는 노이즈가 입력될 때, 생성 네트워크에 의해 최종적으로 출력되는 이미지는 최소 지각 왜곡을 달성할 수 있다.

[0086] 일부 구현들에서, 생성 네트워크의 손실 함수의 제3 손실에서의 $L_{GAN}(Y_{n=1})$ 은 다음의 수학적식:

$$L_{GAN}(Y_{n=1}) = E \left[\log(1 - D(Y_{n=1}^{1, 2, \dots, L})) \right] + E \left[\log(D(HR^{1, 2, \dots, L})) \right]$$

[0088] 에 따라 계산되고, 여기서, $Y_{n=1}^{1, 2, \dots, L}$ 은 제1 입력 이미지에 기초하여 생성 네트워크에 의해 수행되는 반복 프로세스에서 생성되는 이미지 그룹을 나타내고, 이미지 그룹은 매회의 해상도 향상 절차의 종료 시 각각 생성되는 이미지들을 포함한다. $L=1$ 일 때, 이미지 그룹은 제1 출력 이미지를 단독으로 포함하고; 그리고 $L>1$ 일 때, 이미지 그룹은 $Y_{n=1}^1$ 내지 $Y_{n=1}^{L-1}$, 및 제1 출력 이미지 $Y_{n=1}^L$ 을 포함한다.

[0089] $HR^{1, 2, \dots, L}$ 은 제2 해상도 샘플 이미지를 다운샘플링함으로써 획득된 이미지들을 나타내고, 이미지들은 $Y_{n=1}^{1, 2, \dots, L}$ 의 이미지들과 일대일 대응하고, 각각은 대응하는 이미지의 해상도와 동일한 해상도를 갖는다. HR^L 은 제2 해상도 샘플 이미지 자체이다.

[0090] $D(Y_{n=1}^{1, 2, \dots, L})$ 은 $Y_{n=1}^{1, 2, \dots, L}$ 에 기초하여 판별 네트워크에 의해 생성된 판별 결과, 즉 제1 판별 결과를 나타내고; $D(HR^{1, 2, \dots, L})$ 은 $HR^{1, 2, \dots, L}$ 에 기초하여 판별 네트워크에 의해 생성된 판별 결과, 즉 제2 판별 결과를 나타낸다.

[0091] 본 개시내용의 트레이닝 방법에서, 생성 네트워크 트레이닝 절차에 더하여, 트레이닝 방법은 판별 네트워크 트레이닝 절차를 추가로 포함하고, 이 판별 네트워크 트레이닝 절차는: 판별 네트워크에 제1 출력 이미지 및 제2 해상도 샘플 이미지를 개별적으로 제공하여 판별 네트워크가 제1 출력 이미지에 기초한 판별 결과 및 제2 해상도 샘플 이미지에 기초한 판별 결과를 각각 출력할 수 있게 하는 단계; 및 판별 네트워크의 파라미터들을 조정하여 판별 네트워크의 손실 함수를 감소시키는 단계를 포함한다.

[0092] 판별 네트워크 트레이닝 절차 및 생성 네트워크 트레이닝 절차는 미리 설정된 트레이닝 조건이 충족될 때까지 교대로 수행된다. 예를 들어, 미리 설정된 트레이닝 조건은 교대 횟수가 미리 결정된 값에 도달하는 것일 수 있다.

[0093] 초기화 프로세스에서, 생성 네트워크 및 판별 네트워크의 파라미터들은 미리 설정되거나 랜덤이다.

[0094] 전술한 바와 같이, 제1 출력 이미지 및 제2 출력 이미지 둘 다는 해상도 향상 절차의 반복 프로세스를 통해 생성 네트워크에 의해 생성되고, 반복 프로세스에서의 해상도 향상 절차의 총 횟수는 L 이다. $L=1$ 일 때, 이미지가 판별 네트워크에 공급될 때마다, 제1 출력 이미지 또는 제2 해상도 샘플 이미지 단독으로 판별 네트워크에 공급하는 것이 가능하다. $L>1$ 일 때, 제1 입력 이미지에 기초하여 생성 네트워크에 의해 수행되는 이전 $L-1$ 회의 해상도 향상 절차에서, 생성 네트워크는 해상도 향상 절차가 수행될 때마다 중간 이미지를 생성하고; 해상도 향상 절차가 L 번째 반복될 때, 생성 네트워크에 의해 생성된 이미지는 제1 출력 이미지이다. 이러한 경우에, 판별 네트워크는 복수의 이미지를 동시에 수신하기 위해 복수의 입력 단자를 구비하고, 최고 해상도를 갖는 수신된 복수의 이미지 중 하나와, 제2 해상도 샘플 이미지 사이의 매칭 정도를 결정한다. 판별 네트워크 트레이닝 절차에서, 제1 출력 이미지가 판별 네트워크에 제공되는 동안, 제1 입력 이미지에 기초하여 생성 네트워크에 의해 생성되는 각각의 중간 이미지가 판별 네트워크에 제공되고; 제2 해상도 샘플 이미지가 판별 네트워크에 제공되는 동안, 제2 해상도 샘플 이미지를 다운샘플링함으로써 획득된 제3 해상도 샘플 이미지들이 판별 네트워크에 제공되고, 제3 해상도 샘플 이미지들은 중간 이미지들과 일대일 대응하고, 각각은 대응하는 중간 이미지의 해상도와 동일한 해상도를 갖는다.

- [0095] 생성 네트워크의 트레이닝 프로세스에서, 생성 네트워크의 파라미터들은, 생성 네트워크의 출력 결과가 판별 네트워크에 입력된 후에, 판별 네트워크가 판별 결과로서 가능한 한 1에 가까운 매칭 정도를 출력할 수 있게 하도록, 즉, 판별 네트워크가 생성 네트워크의 출력 결과를 제2 해상도 샘플 이미지로서 간주할 수 있게 하도록 조정된다. 판별 네트워크의 트레이닝 프로세스에서, 판별 네트워크의 파라미터들은, 제2 해상도 샘플 이미지가 판별 네트워크에 입력된 후에, 판별 네트워크가 가능한 한 1에 가까운 매칭 정도를 출력할 수 있게 하고, 또한 생성 네트워크의 출력 결과가 판별 네트워크에 입력된 후에, 판별 네트워크가 가능한 한 0에 가까운 매칭 정도를 출력할 수 있게 하도록 조정되는데; 즉, 판별 네트워크는 수신된 이미지가 제2 해상도 샘플 이미지인지를 결정할 수 있도록 트레이닝될 수 있다. 생성 네트워크 및 판별 네트워크를 교대로 트레이닝함으로써, 판별 네트워크는 판별 능력을 개선하도록 지속적으로 최적화되고, 생성 네트워크는 가능한 한 제2 해상도 샘플 이미지에 가까운 결과를 출력하도록 지속적으로 최적화된다. 이 방법에 의해, 2개의 "적대하는" 모델은 서로 경쟁하고 각각은 각각의 트레이닝 프로세스에서 다른 하나로부터의 점점 더 나은 결과에 기초하여 개선되어, 획득된 생성 적대 네트워크 모델이 점점 더 나아진다.
- [0096] 본 개시내용은 위의 트레이닝 방법에 의해 획득되는 생성 적대 네트워크를 사용하는 이미지 처리 방법을 추가로 제공하고, 이미지 처리 방법은 생성 적대 네트워크의 생성 네트워크를 사용하여 이미지의 해상도를 증가시키기 위해 사용되고, 생성 네트워크에 입력 이미지 및 기준 노이즈에 대응하는 노이즈 이미지를 제공하여 생성 네트워크가 입력 이미지보다 더 높은 해상도를 갖는 이미지를 생성할 수 있게 하는 단계를 포함한다. 기준 노이즈의 진폭은 0에서 제1 진폭까지의 범위이다. 구체적으로, 기준 노이즈는 랜덤 노이즈이다.
- [0097] 본 개시내용에 따른 생성 적대 네트워크의 생성 네트워크의 트레이닝 프로세스에서, 0의 진폭을 갖는 노이즈 샘플 및 제1 진폭을 갖는 노이즈 샘플은 생성 네트워크에 개별적으로 제공되고, 생성 네트워크의 손실 함수는 재구성 왜곡 및 지각 왜곡을 평가하기 위한 2개의 왜곡 평가 기준을 결합하여, 생성 네트워크가 이미지의 해상도를 증가시키기 위해 사용될 때, 실제 요구들을 충족시키기 위해 기준 노이즈의 진폭이 실제 요구들에 따라 조정될 수 있다. 예를 들어, 재구성 왜곡의 주어진 범위 내에서, 최소 지각 왜곡은 기준 노이즈의 진폭을 조정함으로써 달성되거나; 또는 지각 왜곡의 주어진 범위 내에서, 최소 재구성 왜곡은 기준 노이즈의 진폭을 조정함으로써 달성된다.
- [0098] 도 3은 본 개시내용의 실시예들에 따른 생성 네트워크의 개략 구조도이다. 생성 네트워크는 도 3과 관련하여 아래에 설명된다. 해상도 향상 절차를 반복하기 위해 생성 네트워크가 사용되고, 해상도 향상 절차가 수행될 때마다 처리될 이미지 I_{l-1} 의 해상도가 증가되어, 증가된 해상도를 갖는 이미지 I_l 를 획득한다. 해상도 향상 절차의 총 반복 횟수가 1일 때, 처리될 이미지 I_{l-1} 은 초기 입력 이미지이고; 해상도 향상 절차의 총 반복 횟수가 L 이고 $L>1$ 일 때, 처리될 이미지 I_{l-1} 은 해상도 향상 절차를 $(l-1)$ 번째 반복한 후에 출력되는 이미지이다. 생성 네트워크는 다음과 같이 예로서 아래에 예시된다: 초기 입력 이미지는 128×128 의 해상도를 갖고, 각각의 해상도 향상 절차에서 해상도는 2배 증가되고, $l=2$ 이다. 이 예에서, 도 3에서의 처리될 이미지 I_{l-1} 은 해상도 향상 절차를 한번 수행한 후 획득된 256×256 이미지이다.
- [0099] 도 3에 도시된 바와 같이, 생성 네트워크는 제1 분석 모듈(11), 제2 분석 모듈(12), 제1 연결 모듈(21), 제2 연결 모듈(22), 보간 모듈(31), 제1 업샘플링 모듈(41), 제1 다운샘플링 모듈(51), 중첩 모듈(70), 및 반복을 위한 잔차 보정 시스템을 포함한다.
- [0100] 제1 분석 모듈(11)은 처리될 이미지 I_{l-1} 의 특징 이미지 R^u_{l-1} 을 생성하도록 구성되고, 특징 이미지 R^u_{l-1} 의 채널들의 수는 처리될 이미지 I_{l-1} 의 채널들보다 더 크다.
- [0101] 제1 연결 모듈(21)은 처리될 이미지의 특징 이미지 R^u_{l-1} 및 노이즈 이미지를 연결하여 제1 병합된 이미지 RC^u_{l-1} 을 획득하도록 구성되고; 제1 병합된 이미지 RC^u_{l-1} 의 채널들의 수는 특징 이미지 R^u_{l-1} 의 채널들의 수와 노이즈 이미지의 채널들의 수의 합이다.
- [0102] 노이즈 이미지의 해상도는 처리될 이미지 I_{l-1} 의 해상도와 동일하다는 점에 유의해야 한다. 따라서, 생성 네트워크에 의해 수행되는 해상도 향상 절차의 총 반복 횟수가 1보다 클 때, 생성 네트워크 트레이닝 절차에서, 생성 네트워크에 제공되는 제1 입력 이미지 및 제2 입력 이미지 각각은 제1 해상도 샘플 이미지 및 상이한 해상도들을 갖는 복수의 노이즈 샘플 이미지를 포함할 수 있거나; 또는 제1 입력 이미지 및 제2 입력 이미지 각각은

제1 해상도 샘플 이미지 및 하나의 노이즈 샘플 이미지를 포함할 수 있고, 해상도 향상 절차가 1번째 반복될 때, 생성 네트워크는 노이즈 샘플의 진폭에 따라 요구된 배율로 노이즈 샘플 이미지를 생성한다.

[0103] 보간 모듈(31)은 처리될 이미지 I_{1-1} 에 대해 보간을 수행하여 그에 기초하여 제4 해상도 이미지를 획득하도록 구성되고, 제4 해상도 이미지는 512x512의 해상도를 갖는다. 보간 모듈은 바이큐빅(bicubic) 보간과 같은 전통적인 보간 방법들을 사용하는 것에 의해 보간을 수행할 수 있다. 제4 해상도 이미지의 해상도는 처리될 이미지 I_{1-1} 의 해상도보다 높다.

[0104] 제2 분석 모듈(12)은 제4 해상도 이미지의 특징 이미지를 생성하도록 구성되고, 특징 이미지의 채널들의 수는 제4 해상도 이미지의 것보다 더 크다.

[0105] 제1 다운샘플링 모듈(51)은 제4 해상도 이미지의 특징 이미지를 다운샘플링하여 256x256의 해상도를 갖는 제1 다운샘플링된 특징 이미지를 획득하도록 구성된다.

[0106] 제2 연결 모듈(22)은 제1 병합된 이미지 RC^{μ}_{1-1} 와 제1 다운샘플링된 특징 이미지를 연결하여 제2 병합된 이미지를 획득하도록 구성된다.

[0107] 제1 업샘플링 모듈(41)은 제2 병합된 이미지를 업샘플링하여 제1 업샘플링된 특징 이미지 R_1^0 를 획득하도록 구성된다.

[0108] 반복을 위한 잔차 보정 시스템은 역-투영(back-projection)을 통해 제1 업샘플링된 특징 이미지에 대해 잔차 보정을 적어도 1회 수행하여, 잔차 보정된 특징 이미지를 획득하도록 구성된다.

[0109] 반복을 위한 잔차 보정 시스템은 제2 다운샘플링 모듈(52), 제2 업샘플링 모듈(42), 및 잔차 결정 모듈(60)을 포함한다. 제2 다운샘플링 모듈(52)은 수신된 이미지를 2배 다운샘플링하도록 구성되고, 제2 업샘플링 모듈(42)은 수신된 이미지를 2배 업샘플링하도록 구성되고; 잔차 결정 모듈(60)은 2개의 수신된 이미지 사이의 차이 이미지를 결정하도록 구성된다.

[0110] 제1 잔차 보정에서, 제1 업샘플링된 특징 이미지 R_1^0 이 첫 번째 것인 제2 다운샘플링 모듈(52)에 의해 2배 다운샘플링되어 특징 이미지 R_1^{01} 를 획득하고; 특징 이미지 R_1^{01} 이 두 번째 것인 제2 다운샘플링 모듈(52)에 의해 2배 다운샘플링되어 초기 입력 이미지와 동일한 해상도를 갖는 특징 이미지 R_1^{02} 를 획득하고; 그 후, 하나의 잔차 결정 모듈은 첫 번째의 해상도 향상 절차에서 획득된 제1 병합된 이미지 RC^{μ}_0 (즉, 초기 입력 이미지의 특징 이미지와 노이즈 이미지를 병합함으로써 획득된 제1 병합된 이미지 RC^{μ}_0)와 특징 이미지 R_1^{02} 사이의 차이 이미지를 획득하기 위해 사용되고; 그 후, 차이 이미지가 제2 업샘플링 모듈에 의해 업샘플링되어 특징 이미지를 획득하고, 획득된 업샘플링된 특징 이미지가 중첩 모듈(70)에 의해 특징 이미지 R_1^{01} 에 대해 중첩되어, 제1 병합된 이미지 R_{1-1}^1 과 동일한 해상도를 갖는 특징 이미지 R_1^{03} 를 획득하고; 그 후, 다른 잔차 결정 모듈이 특징 이미지 R_1^{03} 과 제1 병합된 이미지 RC^{μ}_{1-1} 사이의 차이 이미지를 획득하기 위해 사용되고; 그 후, 차이 이미지가 제2 업샘플링 모듈(42)에 의해 2배 업샘플링되어 업샘플링된 이미지를 획득하고, 업샘플링된 이미지가 제1 업샘플링된 특징 이미지 R_1^0 에 대해 중첩되어, 제1 잔차 보정을 겪은 특징 이미지 R_1^1 를 획득한다.

[0111] 그 후, 특징 이미지 R_1^1 이 동일한 방식으로 제2 잔차 보정을 겪어, 제2 잔차 보정을 겪은 특징 이미지 R_1^2 를 획득할 수 있고; 특징 이미지 R_1^2 는 동일한 방식으로 제3 잔차 보정을 겪을 수 있는 등이다. 도 3에서, μ 는 잔차 보정의 횟수를 나타낸다.

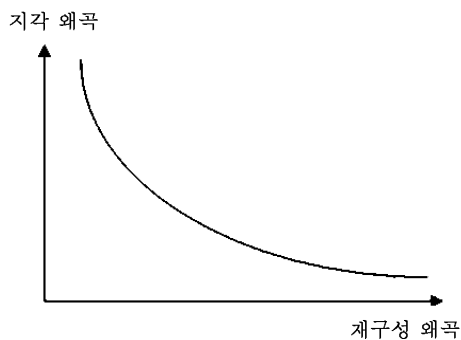
[0112] 생성 네트워크는, 복수 회의 잔차 보정 후에 획득된 특징 이미지 R_1^{μ} 를 합성하여 제5 해상도 이미지를 획득하도록 구성된 합성 모듈(80)을 추가로 포함하고, 제5 해상도 이미지의 채널들의 수는 제4 해상도 이미지의 것과 동일하고; 제5 해상도 이미지 및 제4 해상도 이미지는, 해상도 향상 절차가 1번째 수행된 후에 출력 이미지 I_1 를

획득하기 위해 중첩된다. 제5 해상도 이미지의 해상도는 제4 해상도 이미지의 해상도와 동일하다.

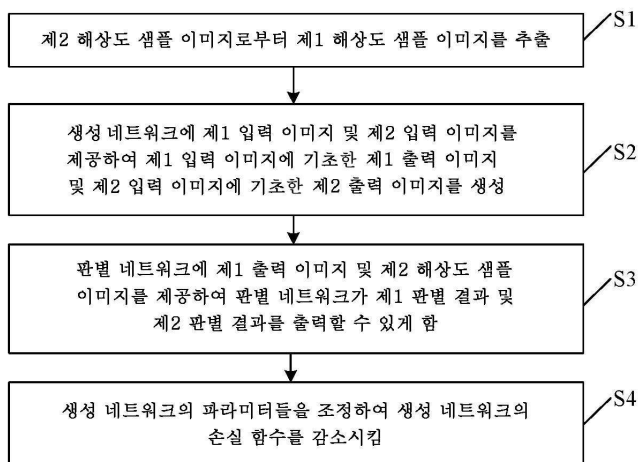
- [0113] 생성 네트워크에서, 제1 분석 모듈(11), 제2 분석 모듈(12), 제1 업샘플링 모듈(41), 제2 업샘플링 모듈(42), 제1 다운샘플링 모듈(51), 제2 다운샘플링 모듈(52), 및 합성 모듈(80)은 컨볼루션 레이어를 통해 대응하는 기능들을 수행할 수 있다.
- [0114] 반복 프로세스에서의 두 번째의 해상도 향상 절차는 $1=2$ 의 예로서 위에 예시되고; 다른 회의 해상도 향상 절차는 두 번째와 유사하므로, 본 명세서에서 상세히 설명되지 않을 것이다.
- [0115] 본 개시내용은 컴퓨터 프로그램들을 저장하는 메모리 및 프로세서를 포함하는 컴퓨터 디바이스를 추가로 제공하고, 컴퓨터 프로그램들이 프로세서에 의해 실행될 때 전술한 생성 적대 네트워크를 위한 트레이닝 방법이 수행된다.
- [0116] 본 개시내용은 컴퓨터 프로그램들을 저장하는 컴퓨터 판독가능 저장 매체를 추가로 제공하고, 컴퓨터 프로그램들이 프로세서에 의해 실행될 때 전술한 생성 적대 네트워크를 위한 트레이닝 방법이 수행된다.
- [0117] 전술한 메모리 및 컴퓨터 판독가능 저장 매체는 다음의 판독가능 매체: 랜덤 액세스 메모리들(RAM들), 판독 전용 메모리들(ROM들), 비휘발성 랜덤 액세스 메모리들(NVRAM들), 프로그램가능 판독 전용 메모리들(PROM들), 소거가능 프로그램가능 판독 전용 메모리들(EPROM들), 전기적 소거가능 프로그램가능 판독 전용 메모리들(EEPROM들), 플래시 메모리들, 자기 또는 광학 데이터 메모리들, 레지스터들, 자기 디스크들 또는 테이프들, 콤팩트 디스크들(CD들) 또는 디지털 다목적 디스크들(DVD들)과 같은 광학 저장 매체, 및 다른 비일시적 매체를 포함하지만, 이에 제한되지 않는다. 프로세서의 예들은 범용 프로세서, 중앙 처리 유닛(CPU), 마이크로프로세서, 디지털 신호 프로세서(DSP), 컨트롤러, 마이크로컨트롤러, 상태 머신 등을 포함하지만, 이에 제한되지 않는다.
- [0118] 전술한 실시예들은 본 개시내용의 원리들을 설명하기 위한 단지 예시적인 실시예들일 뿐이고, 본 개시내용은 그것에 제한되지 않는다는 점을 이해해야 한다. 본 개시내용의 사상 및 본질을 벗어나지 않고서, 통상의 기술자에 의해 다양한 변경들 및 수정들이 이루어질 수 있으며, 본 개시내용의 범위 내에 속하는 것으로 간주되어야 한다.

도면

도면1



도면2



도면3

