



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2023년05월18일
(11) 등록번호 10-2534657
(24) 등록일자 2023년05월16일

(51) 국제특허분류(Int. Cl.)
G06T 3/40 (2006.01) G06T 3/00 (2019.01)
(52) CPC특허분류
G06T 3/4046 (2013.01)
G06N 3/08 (2023.01)
(21) 출원번호 10-2021-0077174
(22) 출원일자 2021년06월15일
심사청구일자 2021년06월15일
(65) 공개번호 10-2022-0167901
(43) 공개일자 2022년12월22일
(56) 선행기술조사문헌
Cagri Ozcinar et al., "Super-resolution of Omnidirectional Images Using Adversarial Learning", 2019 IEEE 21st International Workshop on Multimedia Signal Processing (MMSP), (2019.11.18.)*
Jingwei GUAN et al., "SRDGAN: learning the noise prior for Super Resolution with Dual Generative Adversarial Networks", arXiv:1903.11821, (2019.03.28.)*
Max Jaderberg et al., "Spatial Transformer Networks", Advances in Neural Information Processing Systems 28(NIPS), (2015)*
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
연세대학교 산학협력단
서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)
(72) 발명자
김선주
서울특별시 서대문구 연세로 50 연세대학교
조영현
서울특별시 서대문구
(74) 대리인
권성현, 유광철, 백두진, 강일신, 김정연

전체 청구항 수 : 총 9 항

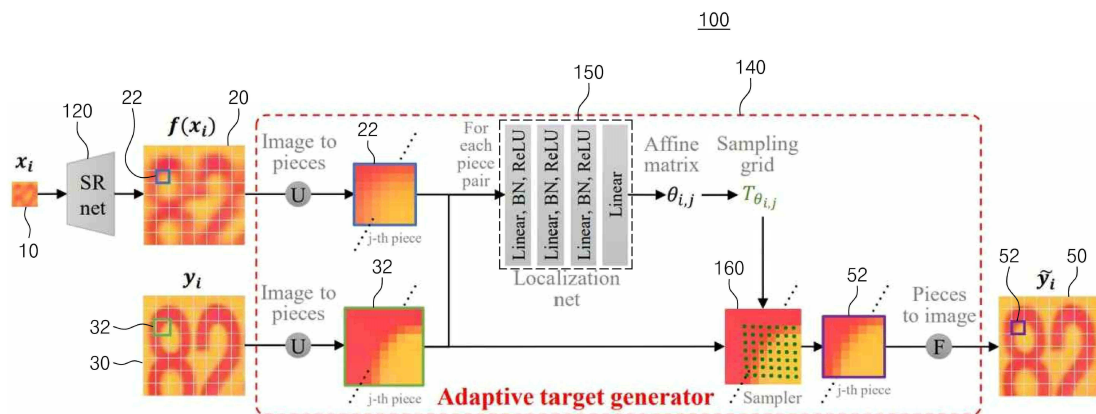
심사관 : 이정은

(54) 발명의 명칭 초해상도 신경망 학습 장치 및 그 방법

(57) 요약

초해상도 알고리즘의 학습 과정에서 일대다(One-to-many) 문제를 효과적으로 처리할 수 있는 초해상도 신경망 학습 장치 및 그 방법이 개시된다. 본 발명의 실시예에 따른 초해상도 신경망 학습 방법은: 초해상도 신경망에 의해, 저해상도 입력 영상을 초해상도 처리하여 고해상도 출력 영상을 획득하는 단계; 적응형 타겟 생성부에 의해, 상기 고해상도 출력 영상과 고해상도 정답 영상을 비교하여 상기 고해상도 정답 영상으로부터 적응형 타겟 영상을 생성하는 단계; 상기 고해상도 정답 영상으로부터 변환된 상기 적응형 타겟 영상을 이용하여 상기 초해상도 신경망을 학습하는 단계;를 포함한다.

대표도



(52) CPC특허분류

G06T 3/0006 (2013.01)

G06T 3/4053 (2013.01)

G06T 2207/20084 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711103312
과제번호	2014-3-00123-007
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	혁신성장동력프로젝트(R&D)(과기정통부)
연구과제명	(답류-1세부) 실시간 대규모 영상 데이터 이해·예측을 위한 고성능 비주얼 디스커
버리 플랫폼 개발	
기 여 율	1/2
과제수행기관명	한국전자통신연구원
연구기간	2021.01.01 ~ 2021.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호	1711126082
과제번호	2020-0-01361-002
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	정보통신방송혁신인재양성(R&D)
연구과제명	인공지능대학원지원(연세대학교)
기 여 율	1/2
과제수행기관명	연세대학교 산학협력단
연구기간	2021.01.01 ~ 2021.12.31

명세서

청구범위

청구항 1

초해상도 신경망에 의해, 저해상도 입력 영상을 초해상도 처리하여 고해상도 출력 영상을 획득하는 단계;

적응형 타겟 생성부에 의해, 상기 고해상도 출력 영상과 고해상도 정답 영상을 비교하여 상기 고해상도 정답 영상으로부터 상기 초해상도 신경망의 훈련을 위한 적응형 타겟 영상을 생성하는 단계;

상기 고해상도 정답 영상으로부터 변환된 상기 적응형 타겟 영상을 이용하여 상기 초해상도 신경망을 학습하는 단계;를 포함하고,

상기 적응형 타겟 영상을 생성하는 단계는:

상기 고해상도 출력 영상과 고해상도 정답 영상 간의 대응하는 패치 영역별로 설정된 왜곡 탐색 범위 내에서 아핀 변환(Affine transformation) 정보를 산출하는 단계; 및

상기 패치 영역 별로 산출된 상기 아핀 변환 정보를 기반으로 상기 고해상도 정답 영상을 상기 패치 영역별로 변환하여 상기 적응형 타겟 영상을 생성하는 단계;를 포함하고,

상기 적응형 타겟 영상을 생성하는 단계는:

상기 패치 영역 별로 상기 고해상도 정답 영상을 상기 아핀 변환 정보에 따라 보정하여 상기 적응형 타겟 영상을 생성하는 단계;를 포함하는 초해상도 신경망 학습 방법.

청구항 2

삭제

청구항 3

삭제

청구항 4

제1항에 있어서,

상기 적응형 타겟 영상을 생성하는 단계에서, 상기 적응형 타겟 생성부는 하기 수식 1의 손실 함수가 최소화되는 아핀 변환 행렬을 추정하도록 선행 학습되고, 선행 학습된 상기 적응형 타겟 생성부에 의해 상기 적응형 타겟 영상을 생성하고,

[수식 1]

$$\sum_i \left[\sum_j \ell(\theta_{i,j}, \theta_{i,j}^{Syn}) + \lambda \ell(\tilde{y}_i, \tilde{y}_i^{Syn}) \right]$$

상기 수식 1에서, i, j는 각 패치 영역의 위치(좌표), ℓ 은 제곱 평균 오차, $\theta_{i,j}$ 는 상기 각 패치 영역의 아핀 변환 행렬, $\theta_{i,j}^{Syn}$ 는 상기 고해상도 정답 영상의 상기 각 패치 영역으로부터 설정된 왜곡 탐색 범위 내에서 임의로 모사되어 생성되는 모의 아핀 변환 행렬, \tilde{y}_i^{Syn} 는 상기 고해상도 정답 영상으로부터 상기 각 패치 영역 별로 상기 모의 아핀 변환 행렬에 의해 임의로 왜곡되어 생성되는 모의 타겟 영상, \tilde{y}_i 는 상기 각 패치 영역 별로 상기 고해상도 정답 영상으로부터 상기 아핀 변환 행렬에 의해 변환되어 생성되는 상기 적응형 타겟 영상, λ 는 설정된 스케일 파라미터인 초해상도 신경망 학습 방법.

청구항 5

제1항에 있어서,

상기 초해상도 신경망을 학습하는 단계는 하기 수식 2의 손실 함수가 최소화되도록 상기 초해상도 신경망을 학습하고,

[수식 2]

$$\sum_i \ell(\tilde{y}_i, f(x_i))$$

상기 수식 2에서, ℓ 은 제곱 평균 오차, \tilde{y}_i 는 상기 패치 영역 별로 상기 고해상도 정답 영상으로부터 상기 아핀 변환 정보에 의해 변환되어 생성되는 상기 적응형 타겟 영상, $f(x_i)$ 는 상기 고해상도 출력 영상인 초해상도 신경망 학습 방법.

청구항 6

제1항에 있어서,

상기 초해상도 신경망을 학습하는 단계는:

상기 초해상도 신경망에 의해 제1 저해상도 입력 영상을 초해상도 처리하여 획득되는 제1 고해상도 출력 영상과, 상기 제1 저해상도 입력 영상과 관련된 제1 고해상도 정답 영상을 이용하여 상기 초해상도 신경망을 1차 학습하는 단계; 및

상기 초해상도 신경망에 의해 제2 저해상도 입력 영상을 초해상도 처리하여 획득되는 제2 고해상도 출력 영상과, 상기 제2 저해상도 입력 영상과 관련된 제2 고해상도 정답 영상으로부터 상기 적응형 타겟 생성부에 의해 생성되는 적응형 타겟 영상을 이용하여 상기 초해상도 신경망을 2차 학습하는 단계;를 포함하는 초해상도 신경망 학습 방법.

청구항 7

제1항, 제4항 내지 제6항 중 어느 한 항의 초해상도 신경망 학습 방법을 실행시키도록 컴퓨터로 판독 가능한 기록 매체에 기록된 컴퓨터 프로그램.

청구항 8

저해상도 입력 영상을 초해상도 처리하여 고해상도 출력 영상을 획득하도록 구성되는 초해상도 신경망;

상기 고해상도 출력 영상과 고해상도 정답 영상을 비교하여 상기 고해상도 정답 영상으로부터 상기 초해상도 신경망의 훈련을 위한 적응형 타겟 영상을 생성하도록 구성되는 적응형 타겟 생성부; 및

상기 고해상도 정답 영상으로부터 변환된 상기 적응형 타겟 영상을 이용하여 상기 초해상도 신경망을 학습하도록 구성되는 초해상도 신경망 학습부;를 포함하고,

상기 적응형 타겟 생성부는:

상기 고해상도 출력 영상과 고해상도 정답 영상 간의 대응하는 패치 영역별로 설정된 왜곡 탐색 범위 내에서 아핀 변환(Affine transformation) 정보를 산출하고;

상기 패치 영역 별로 산출된 상기 아핀 변환 정보를 기반으로 상기 고해상도 정답 영상을 상기 패치 영역별로 변환하여 상기 적응형 타겟 영상을 생성하도록 구성되고,

상기 적응형 타겟 생성부는:

상기 패치 영역 별로 상기 고해상도 정답 영상을 상기 아핀 변환 정보에 따라 보정하여 상기 적응형 타겟 영상을 생성하도록 구성되는 초해상도 신경망 학습 장치.

청구항 9

삭제

청구항 10

삭제

청구항 11

제8항에 있어서,

상기 적응형 타겟 생성부는 하기 수식 1의 손실 함수가 최소화되는 아핀 변환 행렬을 추정하도록 선행 학습되고, 선행 학습된 상기 적응형 타겟 생성부는 상기 적응형 타겟 영상을 생성하고,

[수식 1]

$$\sum_i \left[\sum_j \ell(\theta_{i,j}, \theta_{i,j}^{Syn}) + \lambda \ell(\tilde{y}_i, \tilde{y}_i^{Syn}) \right]$$

상기 수식 1에서, i, j는 각 패치 영역의 위치(좌표), ℓ 은 제곱 평균 오차, $\theta_{i,j}$ 는 상기 각 패치 영역의 아핀 변환 행렬, $\theta_{i,j}^{Syn}$ 는 상기 고해상도 정답 영상의 상기 각 패치 영역으로부터 설정된 왜곡 탐색 범위 내에서 임의로 모사되어 생성되는 모의 아핀 변환 행렬, \tilde{y}_i^{Syn} 는 상기 고해상도 정답 영상으로부터 상기 각 패치 영역 별로 상기 모의 아핀 변환 행렬에 의해 임의로 왜곡되어 생성되는 모의 타겟 영상, \tilde{y}_i 는 상기 각 패치 영역 별로 상기 고해상도 정답 영상으로부터 상기 아핀 변환 행렬에 의해 변환되어 생성되는 상기 적응형 타겟 영상, λ 는 설정된 스케일 파라미터인 초해상도 신경망 학습 장치.

청구항 12

제8항에 있어서,

상기 초해상도 신경망 학습부는 하기 수식 2의 손실 함수가 최소화되도록 상기 초해상도 신경망을 학습하고,

[수식 2]

$$\sum_i \ell(\tilde{y}_i, f(x_i))$$

상기 수식 2에서, ℓ 은 제곱 평균 오차, \tilde{y}_i 는 상기 패치 영역 별로 상기 고해상도 정답 영상으로부터 상기 아핀 변환 정보에 의해 변환되어 생성되는 상기 적응형 타겟 영상, $f(x_i)$ 는 상기 고해상도 출력 영상인 초해상도 신경망 학습 장치.

청구항 13

제8항에 있어서,

상기 초해상도 신경망 학습부는:

상기 초해상도 신경망에 의해 제1 저해상도 입력 영상을 초해상도 처리하여 획득되는 제1 고해상도 출력 영상과, 상기 제1 저해상도 입력 영상과 관련된 제1 고해상도 정답 영상을 이용하여 상기 초해상도 신경망을 1차 학습하고;

상기 초해상도 신경망에 의해 제2 저해상도 입력 영상을 초해상도 처리하여 획득되는 제2 고해상도 출력 영상과, 상기 제2 저해상도 입력 영상과 관련된 제2 고해상도 정답 영상으로부터 상기 적응형 타겟 생성부에 의해 생성되는 적응형 타겟 영상을 이용하여 상기 초해상도 신경망을 2차 학습하도록 구성되는 초해상도 신경망 학습 장치.

발명의 설명

기술 분야

[0001] 본 발명은 초해상도 알고리즘의 일대다(One-to-many) 문제를 처리할 수 있는 초해상도 신경망 학습 장치 및 그 방법에 관한 것이다.

배경 기술

[0002] 초해상도(super-resolution) 처리는 입력 영상의 영상 크기(해상도)를 증가시키기 위한 기술이다. 초해상도 처리는 학습 과정에서 초해상도 문제의 일대다(One-to-many) 특성으로 인해 하나의 저해상도 영상이 많은 고해상도 영상에 매핑될 수 있는 문제가 있다. 그러나 학습 기반의 초해상도 알고리즘은 저해상도 입력 영상을 훈련 데이터에 주어진 하나의 정답 고해상도 영상으로만 매핑되도록 훈련된다.

[0003] 초해상도 알고리즘이 초해상도 문제에서 수학적으로 유효한 결과를 생성한 경우에도 이것이 훈련 데이터에 존재하는 정답과 일치하지 않으면 손실 함수(loss function) 값이 증가하여 알고리즘에 불이익을 줄 수 있다. 이는 블라인드 초해상도(blind super-resolution)에서 더 큰 문제가 되는데, 그 이유는 다양한 알 수 없는 블러 커널(blur kernel)로 인해 유효한 해의 범위가 더 넓어져 초해상도 일대다 문제가 더욱 악화될 수 있기 때문이다.

발명의 내용

해결하려는 과제

[0004] 본 발명은 초해상도 알고리즘의 학습 과정에서 일대다(One-to-many) 문제를 효과적으로 처리할 수 있는 초해상도 신경망 학습 장치 및 그 방법을 제공하기 위한 것이다.

과제의 해결 수단

[0005] 본 발명의 실시예에 따른 초해상도 신경망 학습 방법은: 초해상도 신경망에 의해, 저해상도 입력 영상을 초해상도 처리하여 고해상도 출력 영상을 획득하는 단계; 적응형 타겟 생성부에 의해, 상기 고해상도 출력 영상과 고해상도 정답 영상을 비교하여 상기 고해상도 정답 영상으로부터 적응형 타겟 영상을 생성하는 단계; 및 상기 고해상도 정답 영상으로부터 변환된 상기 적응형 타겟 영상을 이용하여 상기 초해상도 신경망을 학습하는 단계;를 포함한다.

[0006] 상기 적응형 타겟 영상을 생성하는 단계는: 상기 고해상도 출력 영상과 고해상도 정답 영상 간의 대응하는 패치 영역별로 설정된 왜곡 범위 내에서 아핀 변환(Affine transformation) 정보를 산출하는 단계; 및 상기 패치 영역 별로 산출된 상기 아핀 변환 정보를 기반으로 상기 고해상도 정답 영상을 상기 패치 영역별로 변환하여 상기 적응형 타겟 영상을 생성하는 단계;를 포함할 수 있다.

[0007] 상기 적응형 타겟 영상을 생성하는 단계는: 상기 패치 영역 별로 상기 고해상도 정답 영상을 상기 아핀 변환 정보에 따라 보정하여 상기 적응형 타겟 영상을 생성하는 단계;를 포함할 수 있다.

[0008] 상기 적응형 타겟 영상을 생성하는 단계에서, 상기 적응형 타겟 생성부는 하기 수식 1의 손실 함수가 최소화되는 아핀 변환 행렬을 추정하도록 선행 학습되고, 선행 학습된 상기 적응형 타겟 생성부에 의해 상기 적응형 타겟 영상을 생성할 수 있다.

[0009] [수식 1]

$$\sum_i \left[\sum_j \ell(\theta_{i,j}, \theta_{i,j}^{Syn}) + \lambda \ell(\tilde{y}_i, \tilde{y}_i^{Syn}) \right]$$

[0010] 상기 수식 1에서, i, j 는 각 패치 영역의 위치(좌표), ℓ 은 제곱 평균 오차, $\theta_{i,j}$ 는 상기 각 패치 영역의 아핀 변환 행렬, $\theta_{i,j}^{Syn}$ 는 상기 고해상도 정답 영역의 상기 각 패치 영역으로부터 설정된 왜곡 범위 내에서 임의로 모사되어 생성되는 모의 아핀 변환 행렬, \tilde{y}_i^{Syn} 는 상기 고해상도 정답 영상으로부터 상기 각 패치 영역 별로 상기 모의 아핀 변환 행렬에 의해 임의로 왜곡되어 생성되는 모의 타겟 영상, \tilde{y}_i 는 상기 각 패치 영역 별로 상기 고해상도 정답 영역으로부터 상기 아핀 변환 행렬에 의해 변환되어 생성되는 상기 적응형 타겟 영상, λ 는 설정된

스케일 파라미터이다.

[0012] 상기 초해상도 신경망을 학습하는 단계는 하기 수식 2의 손실 함수가 최소화되도록 상기 초해상도 신경망을 학습할 수 있다.

[0013] [수식 2]

$$\sum_i \ell(\tilde{y}_i, f(x_i))$$

[0015] 상기 수식 2에서, ℓ 은 제곱 평균 오차, \tilde{y}_i 는 상기 패치 영역 별로 상기 고해상도 정답 영역으로부터 상기 아핀 변환 정보에 의해 변환되어 생성되는 상기 적응형 타겟 영상, $f(x_i)$ 는 상기 고해상도 출력 영상이다.

[0016] 상기 초해상도 신경망을 학습하는 단계는: 상기 초해상도 신경망에 의해 제1 저해상도 입력 영상을 초해상도 처리하여 획득되는 제1 고해상도 출력 영상과, 상기 제1 저해상도 입력 영상과 관련된 제1 고해상도 정답 영상을 이용하여 상기 초해상도 신경망을 1차 학습하는 단계; 및 상기 초해상도 신경망에 의해 제2 저해상도 입력 영상을 초해상도 처리하여 획득되는 제2 고해상도 출력 영상과, 상기 제2 저해상도 입력 영상과 관련된 제2 고해상도 정답 영상으로부터 상기 적응형 타겟 생성부에 의해 생성되는 적응형 타겟 영상을 이용하여 상기 초해상도 신경망을 2차 학습하는 단계;를 포함할 수 있다.

[0017] 본 발명의 실시예에 따르면, 상기 초해상도 신경망 학습 방법을 실행시키도록 컴퓨터로 판독 가능한 기록 매체에 기록된 컴퓨터 프로그램이 제공된다.

[0018] 본 발명의 실시예에 따른 초해상도 신경망 학습 장치는: 저해상도 입력 영상을 초해상도 처리하여 고해상도 출력 영상을 획득하도록 구성되는 초해상도 신경망; 상기 고해상도 출력 영상과 고해상도 정답 영상을 비교하여 상기 고해상도 정답 영상으로부터 적응형 타겟 영상을 생성하도록 구성되는 적응형 타겟 생성부; 및 상기 고해상도 정답 영상으로부터 변환된 상기 적응형 타겟 영상을 이용하여 상기 초해상도 신경망을 학습하도록 구성되는 초해상도 신경망 학습부;를 포함한다.

[0019] 상기 적응형 타겟 생성부는: 상기 고해상도 출력 영상과 고해상도 정답 영상 간의 대응하는 패치 영역별로 설정된 왜곡 범위 내에서 아핀 변환(Affine transformation) 정보를 산출하고; 상기 패치 영역 별로 산출된 상기 아핀 변환 정보를 기반으로 상기 고해상도 정답 영상을 상기 패치 영역별로 변환하여 상기 적응형 타겟 영상을 생성하도록 구성될 수 있다.

[0020] 상기 적응형 타겟 생성부는: 상기 패치 영역 별로 상기 고해상도 정답 영상을 상기 아핀 변환 정보에 따라 보정하여 상기 적응형 타겟 영상을 생성하도록 구성될 수 있다.

[0021] 상기 적응형 타겟 생성부는 하기 수식 1의 손실 함수가 최소화되는 아핀 변환 행렬을 추정하도록 선행 학습되고, 선행 학습된 상기 적응형 타겟 생성부는 상기 적응형 타겟 영상을 생성할 수 있다.

[0022] [수식 1]

$$\sum_i \left[\sum_j \ell(\theta_{i,j}, \theta_{i,j}^{Syn}) + \lambda \ell(\tilde{y}_i, \tilde{y}_i^{Syn}) \right]$$

[0024] 상기 수식 1에서, i, j 는 각 패치 영역의 위치(좌표), ℓ 은 제곱 평균 오차, $\theta_{i,j}$ 는 상기 각 패치 영역의 아핀 변환 행렬, $\theta_{i,j}^{Syn}$ 는 상기 고해상도 정답 영역의 상기 각 패치 영역으로부터 설정된 왜곡 범위 내에서 임의로 모사되어 생성되는 모의 아핀 변환 행렬, \tilde{y}_i^{Syn} 는 상기 고해상도 정답 영상으로부터 상기 각 패치 영역 별로 상기 모의 아핀 변환 행렬에 의해 임의로 왜곡되어 생성되는 모의 타겟 영상, \tilde{y}_i 는 상기 각 패치 영역 별로 상기 고해상도 정답 영역으로부터 상기 아핀 변환 행렬에 의해 변환되어 생성되는 상기 적응형 타겟 영상, λ 는 설정된 스케일 파라미터이다.

[0025] 상기 초해상도 신경망 학습부는 하기 수식 2의 손실 함수가 최소화되도록 상기 초해상도 신경망을 학습할 수 있다.

[0026] [수식 2]

$$\sum_i \ell(\tilde{y}_i, f(x_i))$$

[0027]

[0028] 상기 수식 2에서, ℓ 은 제곱 평균 오차, \tilde{y}_i 는 상기 패치 영역 별로 상기 고해상도 정답 영역으로부터 상기 아핀 변환 정보에 의해 변환되어 생성되는 상기 적응형 타겟 영상, $f(x_i)$ 는 상기 고해상도 출력 영상이다.

[0029] 상기 초해상도 신경망 학습부는: 상기 초해상도 신경망에 의해 제1 저해상도 입력 영상을 초해상도 처리하여 획득되는 제1 고해상도 출력 영상과, 상기 제1 저해상도 입력 영상과 관련된 제1 고해상도 정답 영상을 이용하여 상기 초해상도 신경망을 1차 학습하고; 상기 초해상도 신경망에 의해 제2 저해상도 입력 영상을 초해상도 처리하여 획득되는 제2 고해상도 출력 영상과, 상기 제2 저해상도 입력 영상과 관련된 제2 고해상도 정답 영상으로부터 상기 적응형 타겟 생성부에 의해 생성되는 적응형 타겟 영상을 이용하여 상기 초해상도 신경망을 2차 학습하도록 구성될 수 있다.

발명의 효과

[0030] 본 발명의 실시예에 의하면, 초해상도 알고리즘의 학습 과정에서 일대다(One-to-many) 문제를 효과적으로 처리할 수 있는 초해상도 신경망 학습 장치 및 그 방법이 제공된다.

도면의 간단한 설명

[0031] 도 1은 본 발명의 실시예에 따른 초해상도 신경망 학습 장치의 구성도이다.

도 2는 본 발명의 실시예에 따른 초해상도 신경망 학습 장치를 구성하는 적응형 타겟 생성부의 기능을 설명하기 위한 개념도이다.

도 3은 본 발명의 실시예에 따른 초해상도 신경망 학습 장치를 구성하는 적응형 타겟 생성부를 구체적으로 나타낸 개념도이다.

도 4는 본 발명의 실시예에 따른 초해상도 신경망 학습 장치의 구성도이다.

도 5는 본 발명의 실시예에 따른 초해상도 신경망 학습 방법의 효과를 설명하기 위한 개념도이다.

도 6은 본 발명의 실시예에 따른 초해상도 신경망 학습 방법에 의해 학습된 초해상도 신경망을 이용하여 생성되는 고해상도 출력 영상의 영상 품질을 보여주는 예시도이다.

도 7은 본 발명의 실시예에 따른 초해상도 신경망 학습 방법의 성능을 종래 기술과 비교하여 나타낸 예시도이다.

도 8 및 도 9는 다양한 테스트셋 영상을 이용하여 본 발명의 실시예에 따른 초해상도 신경망 학습 방법의 영상 품질을 종래 기술과 비교한 결과이다.

도 10은 본 발명의 실시예에 따른 초해상도 신경망 학습 방법의 평가 조사 결과이다.

발명을 실시하기 위한 구체적인 내용

[0032] 본 발명의 이점 및 특징, 그리고 그것들을 달성하는 방법은 첨부되는 도면과 함께 상세하게 후술되어 있는 실시예들을 참조하면 명확해질 것이다. 그러나 본 발명은 이하에서 개시되는 실시예들에 한정되는 것이 아니라 서로 다른 다양한 형태로 구현될 수 있으며, 단지 본 실시예들은 본 발명의 개시가 완전하도록 하고, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 발명의 범주를 완전하게 알려주기 위해 제공되는 것이며, 본 발명은 청구항의 범주에 의해 정의될 뿐이다. 명세서 전체에 걸쳐 동일 참조 부호는 동일 구성 요소를 지칭한다.

[0033] 본 명세서에서, 어떤 부분이 어떤 구성요소를 '포함'한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있는 것을 의미한다. 본 명세서에서 사용되는 '~부'는 적어도 하나의 기능이나 동작을 처리하는 단위로서, 예를 들어 소프트웨어, FPGA 또는 하드웨어 구성요소를 의미할 수 있다. '~부'에서 제공하는 기능은 복수의 구성요소에 의해 분리되어 수행되거나, 다른 추가적인 구성요소와 통합될 수도 있다. 본 명세서의 '~부'는 반드시 소프트웨어 또는 하드웨어에 한정되지 않으며, 어드레싱할 수 있는 저장 매체에 있도록 구성될 수도 있고, 하나 또는 그 이상의 프로세서들을 재생시키도록 구성될

수도 있다. 이하에서는 도면을 참조하여 본 발명의 실시예에 대해서 구체적으로 설명하기로 한다.

[0034] 도 1은 본 발명의 실시예에 따른 초해상도 신경망 학습 장치의 구성도이다. 도 1을 참조하면, 본 발명의 실시예에 따른 초해상도 신경망 학습 장치(100)는 초해상도 알고리즘의 학습 과정에서 일대다(One-to-many) 문제를 효과적으로 처리할 수 있도록 하기 위한 것으로, 이를 위해 초해상도 신경망(120), 및 적응형 타겟 생성부(140)를 포함할 수 있다.

[0035] 초해상도 신경망(super-resolution network)(120)은 저해상도 입력 영상(low resolution input image)을 초해상도(super-resolution) 처리하여, 저해상도 입력 영상 보다 높은 해상도를 가지는 고해상도 출력 영상(high resolution output image)을 출력하도록 구성될 수 있다.

[0036] 적응형 타겟 생성부(140)는 고해상도 출력 영상과 고해상도 정답 영상(high resolution ground truth image or high resolution original target image)을 비교하여 고해상도 정답 영상으로부터 초해상도 신경망(120)의 학습(훈련)을 위한 적응형 타겟 영상(adaptive target image)을 생성하도록 구성될 수 있다.

[0037] 도 2는 본 발명의 실시예에 따른 초해상도 신경망 학습 장치를 구성하는 적응형 타겟 생성부의 기능을 설명하기 위한 개념도이다. 도 1 및 도 2를 참조하면, 적응형 타겟 생성부(140)는 고해상도 출력 영상 $f(x_i)$ (20)과, 고해상도 정답 영상 y_i (30) 간의 대응하는 패치 영역 (22, 32) (24, 34) 별로, 설정된 왜곡 탐색 범위(search space)(36) 내에서 고해상도 패치 영역 (22, 24)에 대해 최소 오차를 갖는 영역(best match)(40)을 도출 후 재배열함으로써 적응형 타겟 영상(50)을 생성할 수 있다. 새롭게 생성된 적응형 타겟 영상(50)은 기존의 원본 데이터(original target)(38)와는 다른 영역들을 포함할 수 있다.

[0038] 도 3은 본 발명의 실시예에 따른 초해상도 신경망 학습 장치를 구성하는 적응형 타겟 생성부를 구체적으로 나타낸 개념도이다. 도 1 내지 도 3을 참조하면, 적응형 타겟 생성부(140)는 국소 신경망(150)과, 샘플러(sampler)(160)를 포함할 수 있다.

[0039] 국소 신경망(150)은 고해상도 출력 영상(20)과 고해상도 정답 영상(30)에 대해 각각 선형(Linear) 처리, 배치 정규화(batch normalization) 처리, ReLU 함수 처리 등을 반복적으로 수행하는 국소 신경망(localization) 처리에 의해, 아핀 변환 행렬(Affine matrix)을 포함하는 아핀 변환 정보를 출력하도록 구성될 수 있다.

[0040] 샘플러(160)는 국소 신경망(150)에 의해 출력되는 아핀 변환 정보(아핀 변환 행렬)에 따라 고해상도 정답 영상(30)을 샘플링하여 각 패치 영역(32)별로 변환 영상(52)을 출력하도록 구성될 수 있다. 패치 영역(32)별로 고해상도 정답 영상(30)으로부터 변환된 변환 영상들(52)은 적응형 타겟 영상(50) 생성을 위해 재정렬될 수 있다.

[0041] 적응형 타겟 생성부(140)는 패치 영역 (22, 32) (24, 34) 별로 고해상도 정답 영상(30)을 아핀 변환 정보에 따라 보정하여 재정렬함으로써 적응형 타겟 영상(50)을 생성하도록 구성될 수 있다. 실시예에서, 적응형 타겟 생성부(140)는 하기 수학적 식 1의 손실 함수가 최소화되는 아핀 변환 행렬을 추정하도록 선행 학습된다. 선행 학습된 적응형 타겟 생성부(140)는 적응형 타겟 영상(50)을 생성할 수 있다.

[0042] [수학적 식 1]

$$\sum_i \left[\sum_j \ell(\theta_{i,j}, \theta_{i,j}^{Syn}) + \lambda \ell(\tilde{y}_i, \tilde{y}_i^{Syn}) \right]$$

[0043] 수학적 식 1에서, i, j 는 고해상도 정답 영상(30)의 각 패치 영역(32, 34)의 위치(좌표), ℓ 은 제곱 평균 오차(mean square error), $\theta_{i,j}$ 는 각 패치 영역(32, 34)의 아핀 변환 행렬, $\theta_{i,j}^{Syn}$ 는 고해상도 정답 영상(30)의 각 패치 영역(32, 34)으로부터 설정된 왜곡 탐색 범위(36) 내에서 임의로 모사되어 생성되는 모의 아핀 변환 행렬, \tilde{y}_i^{Syn} 는 고해상도 정답 영상(30)으로부터 각 패치 영역(32, 34) 별로 모의 아핀 변환 행렬에 의해 임의로 왜곡되어 생성되는 모의 타겟 영상, \tilde{y}_i 는 각 패치 영역(32, 34) 별로 고해상도 정답 영상(30)으로부터 아핀 변환 행렬에 의해 변환되어 생성되는 적응형 타겟 영상(50, 52), λ 는 설정된 스케일 파라미터이다.

[0045] 고해상도 출력 영상(20)과 고해상도 정답 영상(30)은 도 2의 예에서는 각각 5×5 개의 패치 영역들(22, 24) (32, 34)로 분할되어 있고, 도 3의 예에서는 각각 8×8 개의 패치 영역들(22, 24) (32, 34)로 분할되어 있으나, 패치 영역들의 분할 개수는 다양하게 변형될 수 있다.

[0046] 고해상도 출력 영상(20)의 패치 영역들(22, 24)은 서로 중첩되지 않도록 고해상도 출력 영상(20)으로부터 분할될 수 있다. 고해상도 출력 영상(20)과 달리, 고해상도 정답 영상(30)의 패치 영역들(32, 34)은 서로 부분적으로 중첩되도록 고해상도 정답 영상(30)으로부터 분할될 수 있다.

[0047] 따라서, 고해상도 정답 영상(30)의 패치 영역(32, 34)의 크기는 고해상도 출력 영상(20)의 패치 영역(22, 24)보다 크게 설정될 수 있다. 고해상도 정답 영상(30)의 패치 영역(32, 34)의 크기는 왜곡 탐색 범위(36)에 따라 설정될 수 있다. 즉, 고해상도 정답 영상(30)의 패치 영역(32, 34)은 고해상도 출력 영상(20)의 각 패치 영역(22, 24)을 둘러싸는 왜곡 탐색 범위(36)를 커버하도록 설정될 수 있다.

[0048] 왜곡 탐색 범위(36)는 설정된 화소 거리 범위와, 설정된 회전 각도 범위 내로 설정될 수 있다. 실시예에서, 왜곡 탐색 범위(36)의 화소 거리 범위는 $[-1, 1]$ 화소 범위로 설정되고, 회전 각도 범위는 $[-10^\circ, 10^\circ]$ 각도 범위로 설정될 수 있으나, 이에 한정되는 것은 아니다. 적응형 타겟 생성부(140)는 하기 수학적 식 2를 만족하는 적응형 타겟 영상 \hat{y}_i^* 을 산출하도록 구성될 수 있다.

[0049] [수학적 식 2]

$$\hat{y}_i^* = \underset{\hat{y}_i}{\operatorname{argmin}} \ell(\hat{y}_i, f(x_i)), \text{ s.t. } x_i = (\hat{y}_i * k_i^m) \downarrow_s$$

[0050]

[0051] 수학적 식 2에서, \hat{y}_i 은 초해상도 신경망의 학습에 활용될 수 있는 가능한 고해상도 영상들 중의 하나, $f(x_i)$ 는 초해상도 신경망의 출력, ℓ 은 제곱 평균 오차(mean square error), k_i^m 는 블러 커널, x_i 는 저해상도 입력 영상이다. 이때, 저해상도 입력 영상 x_i 는 고해상도 영상 \hat{y}_i 를 블러 커널 k_i^m 을 이용하여 다운 샘플링(down sampling)한 결과와 같아지게 되는 조건을 만족하도록 설정되며, 적응형 타겟 영상 \hat{y}_i^* 은 초해상도 신경망의 출력과 제곱 평균 오차가 최소화되는 고해상도 영상으로 결정될 수 있다.

[0052] 예를 들어, 고해상도 출력 영상(20)의 각 패치 영역(22, 24)이 7×7 화소 크기를 가지고, 왜곡 탐색 범위(36)가 $[-1, 1]$ 화소 범위로 설정되는 경우, 고해상도 정답 영상(30)의 각 패치 영역(32, 34)은 9×9 화소 크기를 가지도록 설정될 수 있다.

[0053] 실시예에서, 초해상도 신경망 학습부는 초해상도 신경망(120)에 의해 제1 저해상도 입력 영상을 초해상도 처리하여 획득되는 제1 고해상도 출력 영상과, 제1 저해상도 입력 영상과 관련된 제1 고해상도 정답 영상을 이용하여 초해상도 신경망(120)을 1차 학습할 수 있다.

[0054] 초해상도 신경망 학습부는 제1 고해상도 정답 영상을 이용한 1차 학습 후, 초해상도 신경망(120)에 의해 제2 저해상도 입력 영상을 초해상도 처리하여 획득되는 제2 고해상도 출력 영상과, 제2 저해상도 입력 영상과 관련된 제2 고해상도 정답 영상으로부터 적응형 타겟 생성부(140)에 의해 생성되는 적응형 타겟 영상을 이용하여 초해상도 신경망(120)을 2차 학습할 수 있다.

[0055] 적응형 타겟 생성부(140)에 의해 생성된 고해상도 적응형 타겟 영상(50)은 초해상도 신경망 학습부(도시 생략됨)에 의해 초해상도 신경망(120)을 학습하는데 활용될 수 있다. 초해상도 신경망 학습부는 고해상도 정답 영상(30)으로부터 변환된 적응형 타겟 영상(50)을 이용하여 초해상도 신경망을 학습하도록 구성될 수 있다. 실시예에서, 초해상도 신경망 학습부는 하기 수학적 식 3의 손실 함수가 최소화되도록 초해상도 신경망(120)을 학습할 수 있다.

[0056] [수학적 식 3]

$$\sum_i \ell(\tilde{y}_i, f(x_i))$$

[0057]

[0058] 수학적 식 3에서, \tilde{y}_i 는 패치 영역 별로 고해상도 정답 영상(30)으로부터 아핀 변환 정보에 의해 변환되어 생성되는 적응형 타겟 영상(50), $f(x_i)$ 는 초해상도 신경망(120)에서 출력되는 고해상도 출력 영상, ℓ 은 고해상도 출력 영상과, 적응형 타겟 영상(50) 간의 제곱 평균 오차(mean square error)이다.

- [0059] 상술한 바와 같이, 본 발명의 실시예에 따른 초해상도 신경망 학습 장치는 주어진 훈련 데이터의 고해상도 정답 영상(30)을 초해상도 알고리즘의 현재 출력에 맞게 아핀 변환(Affine transformation)하여 적응형 타겟 영상(50)을 생성하고, 고해상도 정답 영상(30) 자체가 아닌 이로부터 초해상도 신경망의 출력에 맞게 변형된 적응형 타겟 영상을 이용하여 초해상도 신경망을 학습할 수 있다.
- [0060] 상술한 바와 같은 본 발명의 실시예에 따른 초해상도 신경망 학습 장치에 의하면, 다양한 유효한 초해상도 처리의 해답을 수용할 수 있는 유연성을 제공할 수 있으며, 초해상도 알고리즘이 초해상도 문제의 일대다 특성을 효과적으로 처리할 수 있어 초해상도 처리에 의해 생성되는 고해상도 출력 영상의 영상 품질을 개선할 수 있다.
- [0061] 도 4는 본 발명의 실시예에 따른 초해상도 신경망 학습 방법의 순서도이다. 도 1 내지 도 4를 참조하면, 먼저 초해상도 신경망(120)에 의해, 저해상도 입력 영상(low resolution input image)을 초해상도(super-resolution) 처리하여 저해상도 입력 영상 보다 높은 해상도를 가지는 고해상도 출력 영상을 획득할 수 있다(S10).
- [0062] 적응형 타겟 생성부(140)는 일대다 문제의 효과적 처리를 위해 고해상도 출력 영상(20)과 고해상도 정답 영상(30)을 비교하고, 비교 결과에 따라 고해상도 정답 영상(30)을 변환하여 초해상도 신경망(120)의 학습(훈련)을 위한 적응형 타겟 영상(50)을 생성할 수 있다(S20).
- [0063] 적응형 타겟 생성부(140)는 고해상도 출력 영상(20)과, 고해상도 정답 영상(30) 간의 대응하는 패치 영역 별로, 설정된 왜곡 탐색 범위(36) 내에서 고해상도 패치 영역 (22, 24)에 대해 최소 오차를 갖는 영역(best match)(40)을 도출 후 재배열함으로써 적응형 타겟 영상(50)을 생성할 수 있다. 새롭게 생성된 적응형 타겟 영상(50)은 기존의 원본 데이터(38)와는 다른 영역들을 포함할 수 있다.
- [0064] 이때, 적응형 타겟 생성부(140)는 설정된 왜곡 탐색 범위(36) 내에서 아핀 변환 정보를 산출한 후, 각 패치 영역 별로 산출된 아핀 변환 정보를 기반으로 고해상도 정답 영상(30)의 패치 영역 별로 최소 오차 영역으로 회전 및 이동 변환하여 재배열함으로써 적응형 타겟 영상(50)을 생성할 수 있다.
- [0065] 적응형 타겟 생성부(140)에 의해 생성된 고해상도 적응형 타겟 영상(50)은 초해상도 신경망 학습부(도시 생략됨)에 의해 초해상도 신경망(120)을 학습하는데 활용될 수 있다. 초해상도 신경망 학습부는 고해상도 정답 영상(30)으로부터 변환된 적응형 타겟 영상(50)을 이용하여 초해상도 신경망을 학습할 수 있다(S30).
- [0066] 실시예에서, 초해상도 신경망 학습부는 초해상도 신경망(120)에 의해 제1 저해상도 입력 영상을 초해상도 처리하여 획득되는 제1 고해상도 출력 영상과, 제1 저해상도 입력 영상과 관련된 제1 고해상도 정답 영상을 이용하여 초해상도 신경망(120)을 1차 학습할 수 있다.
- [0067] 초해상도 신경망 학습부는 제1 고해상도 정답 영상을 이용한 1차 학습 후, 초해상도 신경망(120)에 의해 제2 저해상도 입력 영상을 초해상도 처리하여 획득되는 제2 고해상도 출력 영상과, 제2 저해상도 입력 영상과 관련된 제2 고해상도 정답 영상으로부터 적응형 타겟 생성부(140)에 의해 생성되는 적응형 타겟 영상을 이용하여 초해상도 신경망(120)을 2차 학습할 수 있다.
- [0068] 도 5는 본 발명의 실시예에 따른 초해상도 신경망 학습 방법의 효과를 설명하기 위한 개념도이다. 도 1 내지 도 5를 참조하면, 본 발명의 실시예에 따른 초해상도 신경망 학습 방법은 주어진 훈련 데이터의 고해상도 정답 영상 $y_i(30)$ 을 초해상도 알고리즘의 현재 출력에 맞게 아핀 변환하여 적응형 타겟 영상 $\tilde{y}_i(50)$ 을 생성하고, 고해상도 정답 영상 $y_i(30)$ 가 아닌, 이로부터 초해상도 신경망의 출력에 맞게 변형된 적응형 타겟 영상 $\tilde{y}_i(50)$ 을 이용하여 초해상도 신경망을 학습할 수 있다.
- [0069] 도 6은 본 발명의 실시예에 따른 초해상도 신경망 학습 방법에 의해 학습된 초해상도 신경망을 이용하여 생성되는 고해상도 출력 영상의 영상 품질을 보여주는 예시도이다. 업스케일 인자(upscale factor)를 4로 설정하고, 패치 사이즈를 35×35 로 설정하여 본 발명의 실시예에 따른 초해상도 신경망 학습 방법을 구현하였다.
- [0070] 본 발명의 실시예에 따라 적응형 타겟 영상 \tilde{y}_i 을 이용하여 초해상도 신경망을 학습한 경우, 적응형 타겟 영상 \tilde{y}_i 을 이용하여 초해상도 신경망을 학습하지 않은 경우와 비교하여 초해상도 신경망에 의해 출력되는 고해상도 출력 영상 $f(x_i)$ 의 영상 품질이 개선되는 것을 시각적으로 확인할 수 있다.

표 1

Method	Set5			Set14			BSDS100			Urban100			Manga109		
	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS
Bicubic	25.85	0.7289	0.4460	24.20	0.6222	0.5468	24.59	0.5940	0.6445	21.68	0.5745	0.5904	22.79	0.7097	0.4427
KG [2] + ZSSR [34]	-	-	-	-	-	-	-	-	-	22.06	0.6223	0.3788	24.75	0.7786	0.2454
IKC [8]	<u>31.62</u>	<u>0.8808</u>	<u>0.2016</u>	<u>28.18</u>	<u>0.7608</u>	<u>0.3207</u>	<u>27.33</u>	<u>0.7161</u>	<u>0.4174</u>	<u>25.31</u>	<u>0.7487</u>	<u>0.2895</u>	<u>28.82</u>	<u>0.8756</u>	<u>0.1664</u>
Ours	<u>31.58</u>	0.8814	0.1932	<u>28.14</u>	0.7626	0.3122	27.43	0.7216	0.4030	25.72	0.7683	0.2518	29.97	0.8955	0.1286

표 1은 본 발명의 실시예(Ours)에 따라 블라인드(blind) 방식으로 학습된 초해상도 신경망에 의해 출력되는 고해상도 출력 영상의 품질(PSNR(peak signal-to-noise ratio), SSIM(structural similarity image metric), LPIPS(learned perceptual similarity metrics))을 종래 기술과 비교한 결과이다.

표 1의 Set5, Set14, BSDS100, Urban100, Manga109는 성능 검증에 활용된 테스트셋으로, "David Martin, Charless Fowlkes, Doron Tal, and Jitendra Malik. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In ICCV, volume 2, pages 416-423. IEEE, 2001.", "Jia-Bin Huang, Abhishek Singh, and Narendra Ahuja. Single image super-resolution from transformed self-exemplars. In CVPR, pages 5197-5206, 2015.", "Yusuke Matsui, Kota Ito, Yuji Aramaki, Azuma Fujimoto, Toru Ogawa, Toshihiko Yamasaki, and Kiyoharu Aizawa. Sketch-based manga retrieval using manga109 dataset. Multimedia Tools and Applications, 76(20):21811-21838, 2017." 등에 개시된 테스트셋이 성능 검증에 사용되었다.

표 1의 'KG+ZSSR'은 "Sefi Bell-Kligler, Assaf Shocher, and Michal Irani. Blind super-resolution kernel estimation using an internal-gan. In NeurIPS, pages 284-293, 2019." 및 "Assaf Shocher, Nadav Cohen, and Michal Irani. zero-shot super-resolution using deep internal learning. In CVPR, pages 3118-3126, 2018."에 개시된 기술, 'IKC'는 "Jinjin Gu, Hannan Lu, Wangmeng Zuo, and Chao Dong. Blind super-resolution with iterative kernel correction. In CVPR, pages 1604-1613, 2019."에 개시된 기술이다.

도 7은 본 발명의 실시예에 따른 초해상도 신경망 학습 방법의 성능을 종래 기술과 비교하여 나타낸 예시도이다. 'KG+ZSSR' 방식의 경우, 1920×1080 크기의 출력 영상을 생성하는데 458초가 소요되었으며, 'IKC' 방식의 경우, 2.72초가 소요되었다. 이에 반해 본 발명의 실시예(Ours)의 경우, 1920×1080 크기의 출력 영상을 생성하는데 소요되는 시간이 0.44초로 단축되었으며, 영상 품질 또한 대체로 종래 기술보다 우수한 것을 알 수 있다.

도 7에 도시된 'USRNet'은 "Kai Zhang, Luc Van Gool, and Radu Timofte. Deep unfolding network for image super-resolution. In CVPR, pages 3217-3226, 2020."에 개시된 비블라인드(non-blind) 기술이고, 'GT'는 정답 영상(ground truth image)이다. 본 발명의 실시예에 의하면 블라인드 방식으로 학습한 기술임에도 비블라인드 기술인 USRNet과 동등 수준 혹은 그 이상의 영상 품질을 나타내는 것을 알 수 있다.

도 8 및 도 9는 다양한 테스트셋 영상을 이용하여 본 발명의 실시예에 따른 초해상도 신경망 학습 방법의 영상 품질을 종래 기술과 비교한 결과이다. 테스트를 위해 앞서 설명한 'KG' (Bell-Kligler, Shocher, and Irani. NeurIPS 19) 내의 'DIV2K' 테스트셋이 활용되었다. 표 2는 본 발명의 실시예와 종래 기술의 영상 품질을 정량적으로 평가한 결과이다.

표 2

Method	DIV2K		
	PSNR	SSIM	LPIPS
Bicubic	25.37	0.6812	0.5635
KG [2] + ZSSR [34]	26.80	0.7319	0.4389
IKC [8]	<u>27.69</u>	<u>0.7658</u>	<u>0.3856</u>
Ours	28.42	0.7854	0.3394
SRMD [44]	27.63	0.7623	0.3816
MZSR [36]	24.99	0.6919	0.4096
USRNet [43]	29.80	0.8148	0.3088

[0079] PSNR, SSIM은 각각 신호 잡음비, 구조적 영상 유사도 지수로, 그 값이 높을수록 영상 품질이 우수한 것을 의미하며, LPIPS는 지각적 영상 유사도 지수로, 그 값이 낮을수록 지각적 영상 품질이 우수한 것을 의미한다. 도 9에 도시된 'GT'는 정답 영상(ground truth image)이다. 도 8, 도 9 및 표 2로부터, 본 발명의 실시예(Ours)의 경우, 종래 기술(Bicubic, KG+ZSSR, IKC) 보다 영상 품질이 개선되는 것을 알 수 있다. 도 10은 본 발명의 실시예에 따른 초해상도 신경망 학습 방법의 평가 조사 결과이다. 도 10의 (a), (b), (c)는 각각 가우시안 커널(Isotropic Gaussian kernels), 랜덤 커널(Random kernels), 바이큐빅 다운샘플링(Bicubic downsampling)에 의해 초해상도 신경망을 학습한 것이다.

[0080] 도 10에 도시된 'IKC', 'KG+ZSSR'은 앞서 설명한 바와 같고, 'SRMD'는 "Kai Zhang, Wangmeng Zuo, and Lei Zhang. Learning a single convolutional super-resolution network for multiple degradations. In CVPR, pages 3262-3271, 2018."에 개시된 기술, 'ESRGAN'은 "Xintao Wang, Ke Yu, Shixiang Wu, Jinjin Gu, Yihao Liu, Chao Dong, Yu Qiao, and Chen Change Loy. Esrgan: Enhanced super-resolution generative adversarial networks. In ECCV Workshops, September 2018."에 개시된 기술, 'NatSR'은 "Jae Woong Soh, Gu Yong Park, Junho Jo, and Nam Ik Cho. Natural and realistic single image super-resolution with explicit natural manifold discrimination. In CVPR, June 2019."에 개시된 기술, 'RankSRGAN'은 "Wenlong Zhang, Yihao Liu, Chao Dong, and Yu Qiao. Ranksgan: Generative adversarial networks with ranker for image super-resolution. In ICCV, October 2019."에 개시된 기술이다.

[0081] Amazon Mechanical Turk 조사 기관을 통해, 본 발명의 실시예(Ours)에 따라 초해상도 처리에 의해 출력되는 고해상도 영상과, 종래 기술(IKC, SRMD, KG+ZSSR, ESRGAN, NatSR, RankSRGAN)에 따라 출력되는 고해상도 영상을 다수의 평가자들에게 제시하여 1순위(Rank1) 부터 4순위(Rank4)까지 순위를 평가하도록 하였으며, 그 결과 평가자들은 모든 경우에 대해 본 발명의 실시예에 따라 생성된 고해상도 영상의 품질이 가장 우수한 것으로 평가한 것으로 조사되었다.

표 3

Learned transformations	DIV2KRRK		
	PSNR	SSIM	LPIPS
Integer step Tr (NN search)	27.75	0.7772	0.3195
Tr	28.42	0.7832	0.3469
Tr + Rot (Ours)	28.42	<u>0.7854</u>	<u>0.3394</u>
Tr + Rot ($p = 7$ and $s = 11$)	28.27	0.7778	0.3679
Tr + Rot + Shear	<u>28.39</u>	0.7851	0.3402
Tr + Rot + Shear + Scale	28.26	0.7855	0.3382

[0082]

[0083] 표 3은 학습 변환의 유형에 따른 영상 품질 변화를 나타낸 것이다. 표 3에서 'Tr'은 이동(translation) 변환을 나타내고, 'Rot'는 회전(rotation) 변환을 나타낸다. 이동 변환과 회전 변환을 모두 사용하여 초해상도 신경망을 학습한 경우, 영상 품질이 우수한 것을 알 수 있다. 상대적으로 좁은 탐색 영역을 설정한 'Tr+Rot(Ours)'($p=7$ and $s=9$)의 경우, 보다 넓은 왜곡 탐색 영역을 가지는 'Tr+Rot($p=7$ and $s=11$)'의 경우에 비해 PSNR, SSIM 및 LPIPS 성능이 우수한 것을 알 수 있다. 이동/회전 변환과 함께 시어(shear), 스케일(scale) 변환을 함께 적용한 경우에도 우수한 PSNR, SSIM, LPIPS 성능을 얻을 수 있음을 알 수 있다.

[0084] 상술한 바와 같이, 본 발명의 실시예에 따른 초해상도 신경망 학습 방법에 의하면, 다양한 유효한 초해상도 처리의 해답을 수용할 수 있는 유연성을 제공할 수 있으며, 초해상도 알고리즘이 초해상도 문제의 일대다 특성을 효과적으로 처리할 수 있어 초해상도 처리에 의해 생성되는 고해상도 출력 영상의 영상 품질을 개선할 수 있다.

[0085] 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

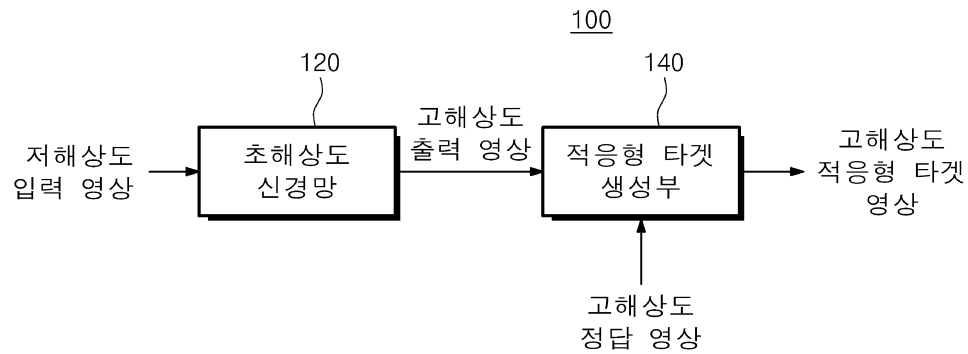
수 있다. 그러므로, 다른 구현들, 다른 실시예들 및 청구범위와 균등한 것들도 후술하는 청구범위의 범위에 속한다.

부호의 설명

[0086]	10: 저해상도 입력 영상	20: 고해상도 출력 영상
	22, 24: 패치 영역	30: 고해상도 정답 영상
	32, 34: 패치 영역	36: 왜곡 탐색 범위
	38: 원본 데이터	40: 최소 오차 영역
	50: 적응형 타겟 영상	100: 초해상도 신경망 학습 장치
	120: 초해상도 신경망	140: 적응형 타겟 생성부
	150: 신경망	160: 샘플러

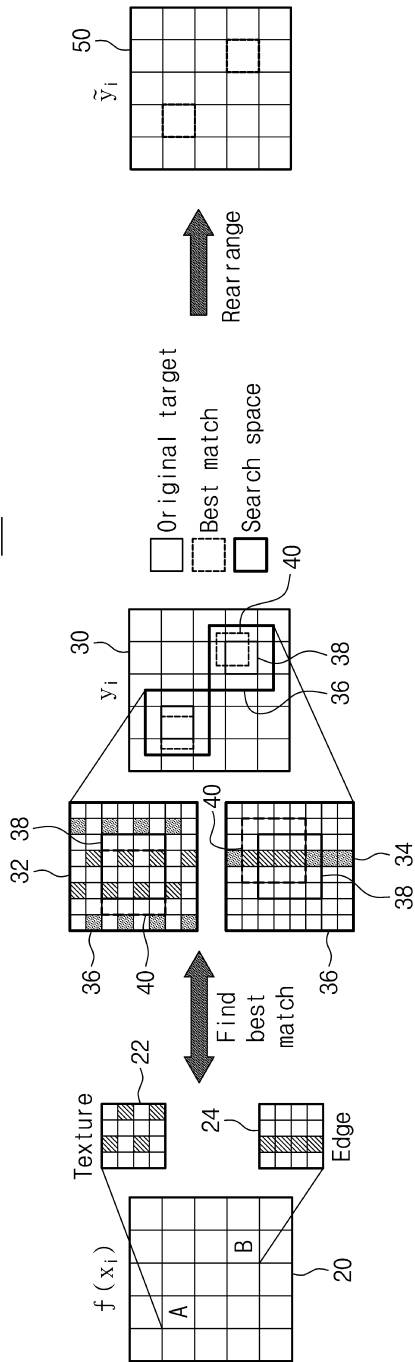
도면

도면1



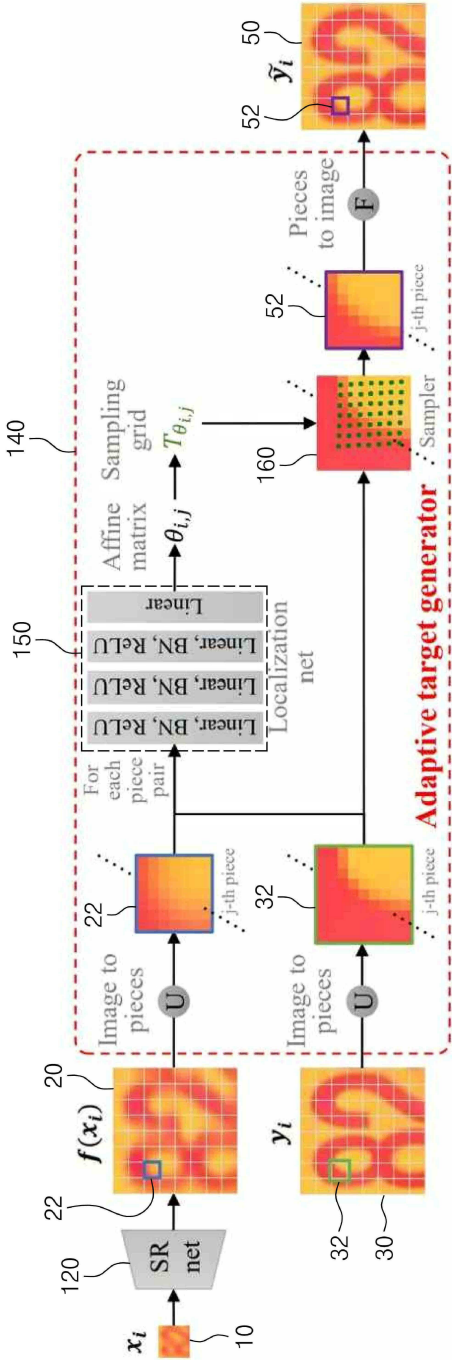
도면2

140

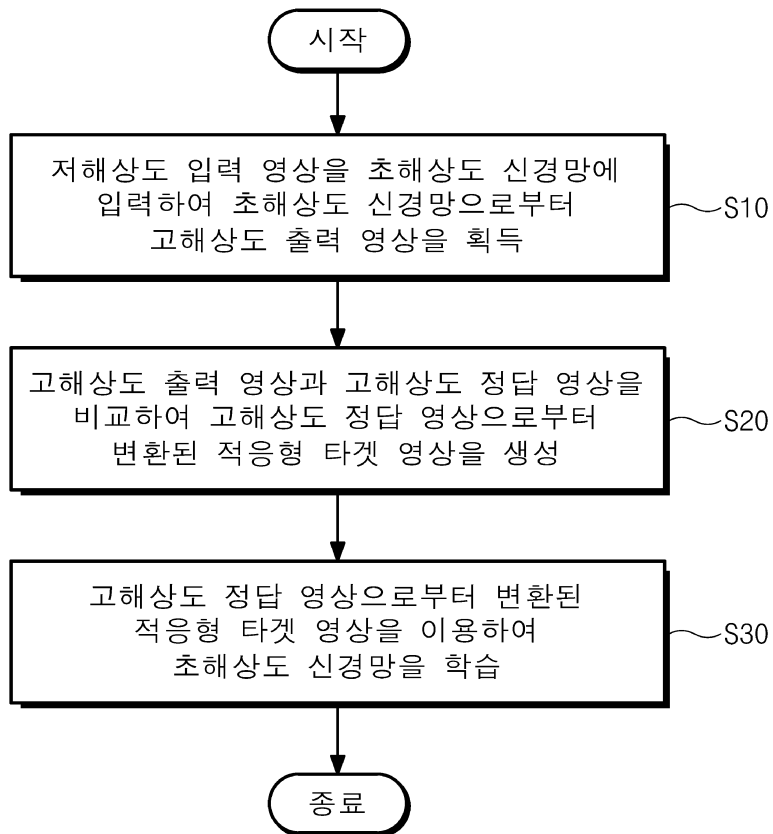


도면3

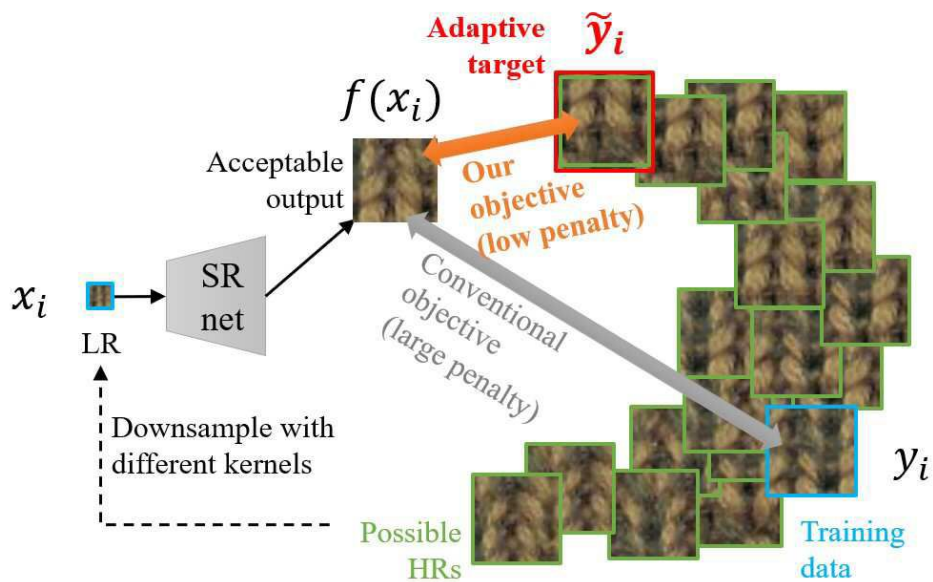
100









도면4



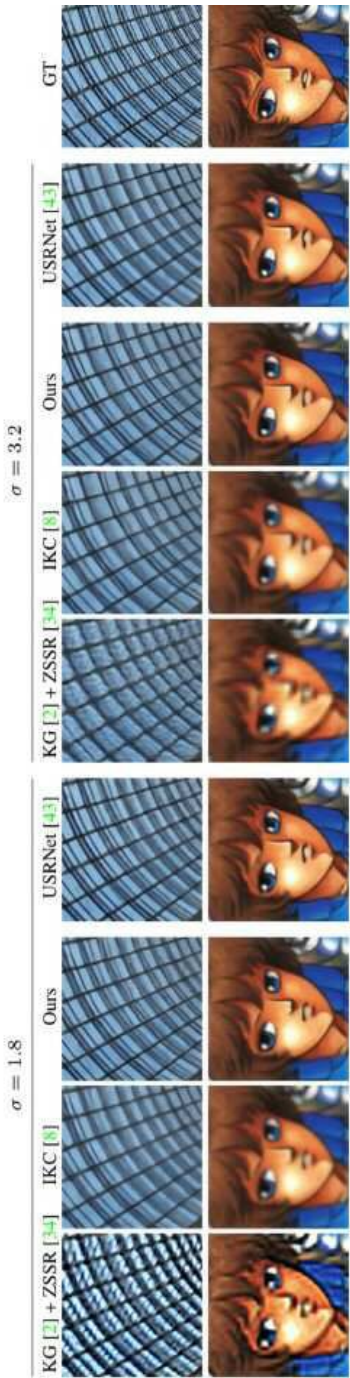
도면5



도면6

Before training with \tilde{y}_i				After training with \tilde{y}_i			
x_i	$f(x_i)$	\tilde{y}_i	y_i	$f(x_i)$	\tilde{y}_i	y_i	
		1.29e-2	1.50e-2		2.62e-3	7.46e-3	
		1.80e-2	1.92e-2		5.09e-3	8.29e-3	

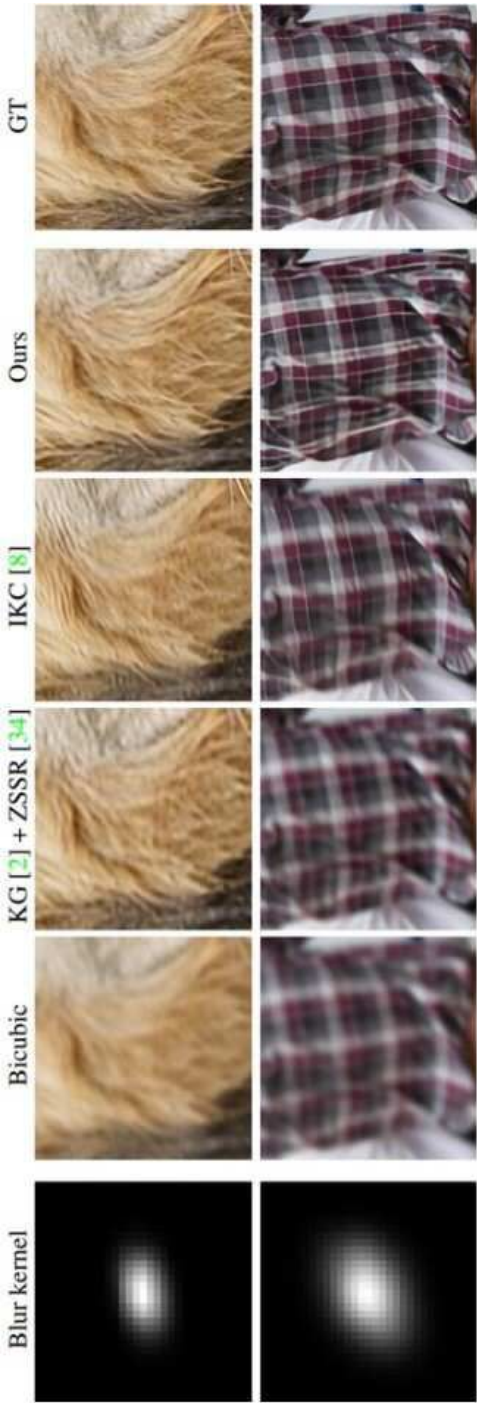
도면7



도면8



도면9



도면10

