



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0055841
(43) 공개일자 2022년05월04일

(51) 국제특허분류(Int. Cl.)
G06T 3/40 (2006.01) G06N 20/00 (2019.01)
(52) CPC특허분류
G06T 3/4053 (2013.01)
G06N 20/00 (2021.08)
(21) 출원번호 10-2020-0140333
(22) 출원일자 2020년10월27일
심사청구일자 2020년10월27일

(71) 출원인
연세대학교 산학협력단
서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)
(72) 발명자
이중석
인천광역시 남동구 소래역로 93, 906동 104호(논현동, 넷마을신영지웰아파트)
전근우
경상남도 창원시 의창구 의창대로211번길 6-5(서상동)
(뒷면에 계속)
(74) 대리인
특허법인우인

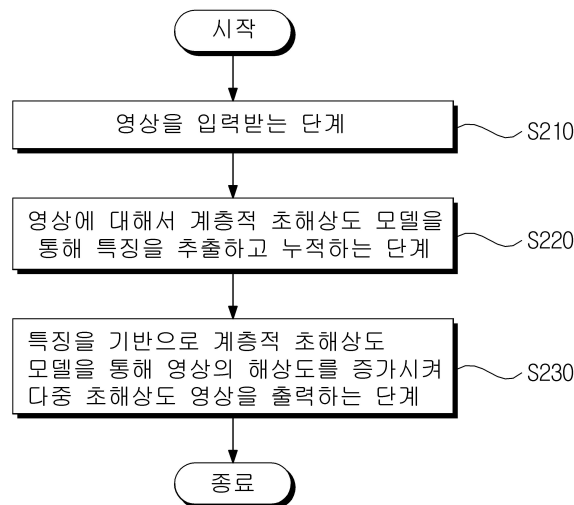
전체 청구항 수 : 총 5 항

(54) 발명의 명칭 다중 출력 아키텍처를 통한 점진적 초해상도 영상 변환 장치 및 방법

(57) 요약

본 실시예들은 계층적 초해상도 모델을 통해 특징을 점진적으로 향상시켜 기본 초해상도 영상, 중간 초해상도 영상, 및 최종 초해상도 영상에 해당하는 다중 초해상도 영상을 출력할 수 있는 초해상도 영상 변환 방법 및 장치를 제공한다.

대 표 도 - 도5



(52) CPC특허분류

G06T 2207/20016 (2013.01)

G06T 2207/20081 (2013.01)

(72) 발명자

최준호

인천광역시 남동구 논현로26번길 12, 605A호 (논현동, 부티크 646)

김준혁

인천광역시 연수구 송도과학로27번길 55, 201동 1102호(송도동, 롯데캐슬 캠퍼스타운)

이 발명을 지원한 국가연구개발사업

과제고유번호 1711116255

과제번호 2016-0-00564-005

부처명 과학기술정보통신부

과제관리(전문)기관명 정보통신기획평가원(한국연구재단부설)

연구사업명 정보통신방송연구개발사업

연구과제명 사용자의 의도와 맥락을 이해하는 지능형 인터랙션 기술연구개발

기 여 율 1/1

과제수행기관명 경북대학교산학협력단

연구기간 2020.03.01 ~ 2020.12.31

명세서

청구범위

청구항 1

하나 이상의 프로세서 및 상기 하나 이상의 프로세서에 의해 실행되는 하나 이상의 프로그램을 저장하는 메모리를 포함하는 초해상도 영상 변환 장치에 있어서,

상기 프로세서는,

영상을 입력받고,

상기 영상에 대해서 계층적 초해상도 모델을 통해 특징을 추출하고 누적하고,

상기 특징을 기반으로 상기 계층적 초해상도 모델을 통해 상기 영상의 해상도를 증가시켜 다중 초해상도 영상을 출력하는 것을 특징으로 하는 초해상도 영상 변환 장치.

청구항 2

제1항에 있어서,

상기 계층적 초해상도 모델은 레이어의 입력을 다음 레이어의 출력에 바로 연결하는 스킵 구조를 갖는 레지듀얼 블록을 계층마다 연결한 구조를 포함하고,

상기 프로세서가 상기 레지듀얼 블록을 이용하여 상기 특징을 추출하고 누적하는 것을 특징으로 하는 초해상도 영상 변환 장치.

청구항 3

제1항에 있어서,

상기 프로세서가 다중 초해상도 영상을 출력하는 것은,

기본 초해상도 영상을 출력하고, 중간 초해상도 영상을 출력하고, 최종 초해상도 영상을 출력하는 것을 특징으로 하는 초해상도 영상 변환 장치.

청구항 4

제3항에 있어서,

상기 계층적 초해상도 모델은 컨볼루션 레이어 쌍 사이에 활성화 함수를 갖는 구조에 픽셀을 재정렬하는 픽셀 셔플 블록이 연결된 구조를 계층마다 포함하고,

상기 프로세서가 상기 픽셀 셔플 블록을 이용하여 상기 중간 초해상도 영상을 출력하고,

상기 프로세서가 N (상기 N 은 자연수) 번째 계층에 해당하는 중간 초해상도 영상에 $N+1$ 번째 계층에 해당하는 중간 초해상도를 추가하는 것을 특징으로 하는 초해상도 영상 변환 장치.

청구항 5

제3항에 있어서,

상기 프로세서는 계층 전체의 특징을 모두 합하고 상기 계층 전체의 특징으로부터 상기 최종 초해상도 영상을 출력하고,

상기 계층적 초해상도 모델의 손실 함수는 (i) 상기 중간 초해상도 영상과 검증 영상 간의 차이 및 (ii) 상기 최종 초해상도 영상과 상기 검증 영상 간의 차이를 최소화하도록 학습되는 것을 특징으로 하는 초해상도 영상 변환 장치.

발명의 설명

기술 분야

[0001] 본 발명이 속하는 기술 분야는 다중 출력이 가능한 초해상도 영상 변환 장치 및 방법에 관한 것이다.

배경 기술

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

[0003] 저해상도의 영상을 세세한 부분까지 복원하여 고해상도로 확대하는 기술을 초해상도라고 한다.

[0004] 딥러닝 기반 초해상도 방식은 학습을 통해 픽셀값을 추정한다. 딥러닝 기반 초해상도 방식은 심층 신경 네트워크를 학습하는 방식이며 추론을 기반으로 스케일을 증가시킨다.

선행기술문헌

특허문헌

[0005] (특허문헌 0001) 한국등록특허공보 제10-1834512호 (2018.02.26.)

발명의 내용

해결하려는 과제

[0006] 본 발명의 실시예들은 계층적 초해상도 모델을 통해 특징을 점진적으로 향상시켜 기본 초해상도 영상, 중간 초해상도 영상, 및 최종 초해상도 영상에 해당하는 다중 초해상도 영상을 출력하는 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, 하나 이상의 프로세서 및 상기 하나 이상의 프로세서에 의해 실행되는 하나 이상의 프로그램을 저장하는 메모리를 포함하는 초해상도 영상 변환 장치에 있어서, 상기 프로세서는, 영상을 입력받고, 상기 영상에 대해서 계층적 초해상도 모델을 통해 특징을 추출하고 누적하고, 상기 특징을 기반으로 상기 계층적 초해상도 모델을 통해 상기 영상의 해상도를 증가시켜 다중 초해상도 영상을 출력하는 것을 특징으로 하는 초해상도 영상 변환 장치를 제공한다.

[0009] 상기 계층적 초해상도 모델은 레이어의 입력을 다음 레이어의 출력에 바로 연결하는 스킵 구조를 갖는 레지듀얼 블록을 계층마다 연결한 구조를 포함하고, 상기 프로세서가 상기 레지듀얼 블록을 이용하여 상기 특징을 추출하고 누적할 수 있다.

[0010] 상기 프로세서가 다중 초해상도 영상을 출력하는 것은 기본 초해상도 영상을 출력하고, 중간 초해상도 영상을 출력하고, 최종 초해상도 영상을 출력할 수 있다.

[0011] 상기 계층적 초해상도 모델은 컨볼루션 레이어 쌍 사이에 활성화 함수를 갖는 구조에 픽셀을 재정렬하는 픽셀 셔플 블록이 연결된 구조를 계층마다 포함할 수 있다.

[0012] 상기 프로세서가 상기 픽셀 셔플 블록을 이용하여 상기 중간 초해상도 영상을 출력하고, 상기 프로세서가 N(상기 N은 자연수) 번째 계층에 해당하는 중간 초해상도 영상에 N+1 번째 계층에 해당하는 중간 초해상도를 추가할 수 있다.

[0013] 상기 프로세서는 계층 전체의 특징을 모두 합하고 상기 계층 전체의 특징으로부터 상기 최종 초해상도 영상을 출력할 수 있다.

[0014] 상기 계층적 초해상도 모델의 손실 함수는 (i) 상기 중간 초해상도 영상과 검증 영상 간의 차이 및 (ii) 상기 최종 초해상도 영상과 상기 검증 영상 간의 차이를 최소화하도록 학습될 수 있다.

[0015] 본 실시예의 다른 측면에 의하면, 컴퓨팅 디바이스에 의한 초해상도 영상 변환 방법에 있어서, 영상을 입력받는 단계, 상기 영상에 대해서 계층적 초해상도 모델을 통해 특징을 추출하고 누적하는 단계, 및 상기 특징을 기반으로 상기 계층적 초해상도 모델을 통해 상기 영상의 해상도를 증가시켜 다중 초해상도 영상을 출력하는 단계를 포함하는 초해상도 영상 변환 방법을 제공한다.

발명의 효과

[0016] 이상에서 설명한 바와 같이 본 발명의 실시예들에 의하면, 계층적 초해상도 모델을 통해 특징을 점진적으로 향상시켜 기본 초해상도 영상, 중간 초해상도 영상, 및 최종 초해상도 영상에 해당하는 다중 초해상도 영상을 출력할 수 있는 효과가 있다.

[0017] 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

[0018] 도 1은 초해상도 기법을 예시한 도면이다.

도 2는 본 발명의 일 실시예에 따른 초해상도 영상 변환 장치를 예시한 블록도이다.

도 3 및 도 4는 본 발명의 일 실시예에 따른 초해상도 영상 변환 장치의 계층적 초해상도 모델을 예시한 도면이다.

도 5는 본 발명의 다른 실시예에 따른 초해상도 영상 변환 방법을 예시한 흐름도이다.

발명을 실시하기 위한 구체적인 내용

[0019] 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.

[0020] 도 1은 초해상도 기법을 예시한 도면이다.

[0021] 딥러닝 기반 초해상도 기법은 선명한 영상(Ground Truth)과 생성된 출력에 대해서 모델 학습을 통해 픽셀값을 추정한다. 딥러닝 기반 초해상도 기법은 심층 신경 네트워크를 학습하는 방식이며 추론을 기반으로 스케일을 증가시킨다.

[0022] 기존의 초해상도 모델은 네트워크의 마지막 레이어에서 하나의 출력을 생성한다.

[0023] 본 발명의 일 실시예에 따른 초해상도 영상 변환 장치는 계층적 네트워크 모델을 통해 네트워크의 중간에서도 출력을 생성하여, 다중 초해상도 영상을 출력할 수 있다.

[0024] 도 2는 본 발명의 일 실시예에 따른 초해상도 영상 변환 장치를 예시한 블록도이다.

[0025] 초해상도 영상 변환 장치(110)는 적어도 하나의 프로세서(120), 컴퓨터 판독 가능한 저장매체(130) 및 통신 버스(170)를 포함한다.

[0026] 프로세서(120)는 초해상도 영상 변환 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 초해상도 영상 변환 장치(110)로 하여금 예시적인 실시예에 따른 동작들을 수행하도록 구성될 수 있다.

[0027] 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 초해상도 영상 변환 장치(110)에 의해 액세스되고 원하는 정보를 저장할 수 있는 다른 형태의 저장 매체, 또는 이들의 적합한 조합일 수 있다.

- [0028] 통신 버스(170)는 프로세서(120), 컴퓨터 관독 가능한 저장 매체(140)를 포함하여 초해상도 영상 변환 장치(110)의 다른 다양한 컴포넌트들을 상호 연결한다.
- [0029] 초해상도 영상 변환 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 초해상도 영상 변환 장치(110)의 다른 컴포넌트들에 연결될 수 있다.
- [0030] 초해상도 영상 변환 장치(110)는 차례로 연결된 네트워크 블록에서 이전 블록에서 미리 만든 출력에 계속해서 정보를 추가하는 방식으로 동작한다. 이를 계층적 초해상도 모델이라고 칭한다. 계층적 초해상도 모델은 최종 출력의 정확도를 높일 수 있다. 하드웨어에 따른 시간 제약 등이 있을 때는 이전 블록에서 미리 생성한 출력을 활용할 수 있다. 느리고 성능이 좋은 네트워크 또는 빠르고 성능이 낮은 네트워크 사이에서 트레이드 오프하여 상황에 맞게 출력 품질 및 출력 시간을 조절할 수 있다.
- [0031] 도 3 및 도 4는 본 발명의 일 실시예에 따른 초해상도 영상 변환 장치의 계층적 초해상도 모델을 예시한 도면이다.
- [0032] 계층적 초해상도 모델은 복수의 레이어가 연결된 학습 네트워크 구조이다. 초해상도 모델은 특징 추출 모델과 복수의 업스케일 모델을 포함한다. 복수의 업스케일 모델은 계층적으로 연결된다. 특징 추출 모델과 업스케일 모델은 다수의 레이어가 네트워크로 연결되며 컨볼루션(Convolution) 레이어를 포함할 수 있다. 레이어는 파라미터를 포함할 수 있고, 레이어의 파라미터는 학습가능한 필터 집합을 포함한다. 파라미터는 노드 간의 가중치 및/또는 바이어스를 포함한다. 네트워크 모델은 손실 함수를 최소화하는 방향으로 네트워크 가중치를 갱신한다.
- [0033] 계층적 초해상도 모델은 레지듀얼(Residual) 블록을 포함할 수 있다. 레지듀얼 블록은 네트워크 구조의 출력에 다시 입력을 더해 다음 레이어로 넘기며, 레이어의 입력을 레이어의 출력에 바로 연결하는 스킵 구조를 가질 수 있다.
- [0034] 계층적 초해상도 모델은 업스케일 모델의 후단에 픽셀 셔플 모듈을 적용한다. 채널 확장 컨볼루션을 적용하지 않고 픽셀 셔플 모듈을 적용하여 파라미터의 개수와 실행 시간을 감소시킬 수 있다.
- [0035] 예컨대 계층적 초해상도 모델은 48 채널의 저해상도 영상 크기를 갖는 특징 맵으로부터 3 채널의 고해상도 영상 크기를 갖는 특징 맵으로 4배 업스케일을 수행할 수 있다. 픽셀 셔플 모듈에 의해 초해상도 특징 맵의 크기를 유지하면서 채널을 변경할 수 있다.
- [0036] 계층적 초해상도 모델은 헤드 모듈, 복수의 바디 모듈, 테일 모듈을 포함한다. 각각의 바디 모듈의 후단에 레그 모듈을 포함한다. 초해상도 모델(LarvaNet)의 수학적 정의는 수학식 1 내지 수학식8과 같이 표현되며, X 는 입력된 저해상도 영상이고, Y 는 검증 자료인 고해상도 영상이고, \hat{Y} 는 출력된 초해상도 영상을 의미한다.

수학식 1

$$\hat{Y} = \text{LarvaNet}(X)$$

- [0038] 계층적 초해상도 모델은 쌍삼차보간(bicubic interpolation)을 이용하여 기본 초해상도 영상을 산출할 수 있다. 쌍삼차보간은 인접한 4개의 픽셀값을 사용하는 삼차보간(cubic interpolation)을 2차원(예컨대, x축과 y축)으로 확장시킨 보간으로, 인접한 16개의 픽셀값과 거리에 따른 가중치의 곱을 사용하여 산출한다.

수학식 2

$$\hat{Y}_{base} = \text{bicubic}(X)$$

- [0040] 헤드 모듈은 입력된 저해상도 영상으로부터 초기 특징 맵 F_0 을 추출한다.

수학식 3

$$F_0 = head(X)$$

i번째 바디 모듈은 스킵 연결을 갖는 레지듀얼(Residual) 블록 집합을 사용하여 특징을 처리한다.

수학식 4

$$\begin{aligned} F_i &= body_i(F_{i-1}) \\ &= F_{i-1} + res_i(F_{i-1}) \quad (i \in \{1, 2, \dots, n\}) \end{aligned}$$

수학식 5

$$F_i = F_0 + res_1(F_0) + \dots + res_i(F_{i-1})$$

계층적 초해상도 모델은 수학식 4 및 수학식 5와 같이 특징을 누적할 수 있다.

각 바디 모듈의 레그 모듈은 기본 초해상도 영상에 추가하는 방식을 초해상도 영상을 생성한다. 레그 모듈은 컨볼루션 레이어 쌍 사이에 활성화 함수를 갖는 블록과 픽셀을 재정렬하는 픽셀 셔플 블록을 결합한 모듈이다.

수학식 6

$$\hat{Y}_i = \hat{Y}_{base} + leg_i(F_i)$$

테일 모듈은 전체 바디 모듈의 출력을 통합한다.

수학식 7

$$\hat{Y} = \hat{Y}_{base} + tail(F_1, F_2, \dots, F_n)$$

계층적 초해상도 모델의 손실 함수는 (i) 중간 초해상도 영상과 검증 영상 간의 차이 및 (ii) 최종 초해상도 영상과 검증 영상 간의 차이를 최소화하도록 학습된다.

수학식 8

$$Loss = \frac{1}{n+1} \left[\sum_{i=1}^n \|Y - \hat{Y}_i\|_1 + \|Y - \hat{Y}\|_1 \right]$$

손실 함수는 영상 간의 차이를 놈(Norm)으로 산출한다.

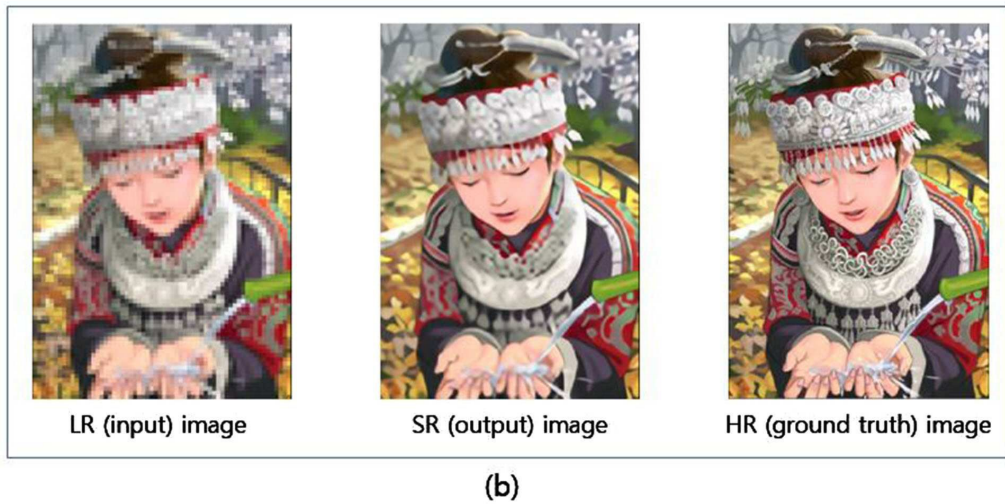
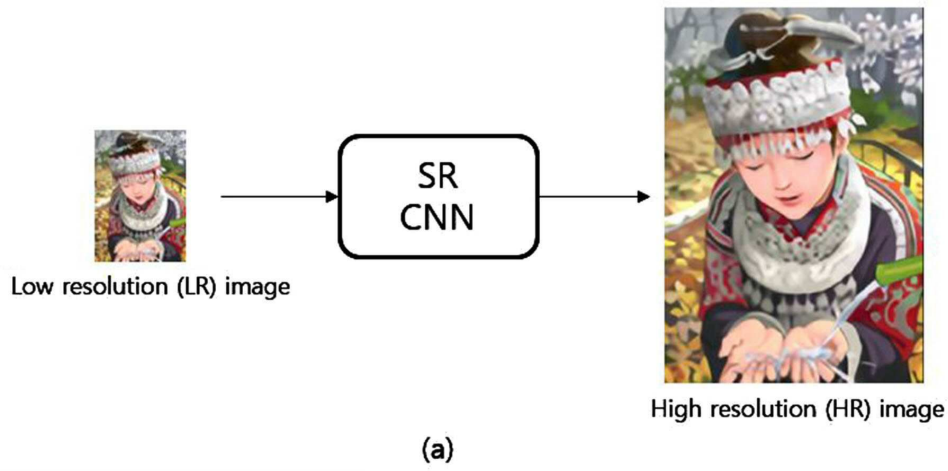
계층적으로 연결된 각 바디 모듈은 자신의 손실 함수에 따른 얇은 모델의 일부와 다음 모델들의 손실 함수들에 따른 깊은 모델의 일부를 모두 동시에 학습한다. 각 바디 모듈은 복수의 초해상도 영상을 복원하는데 효과적인

특징을 생성하도록 훈련된다.

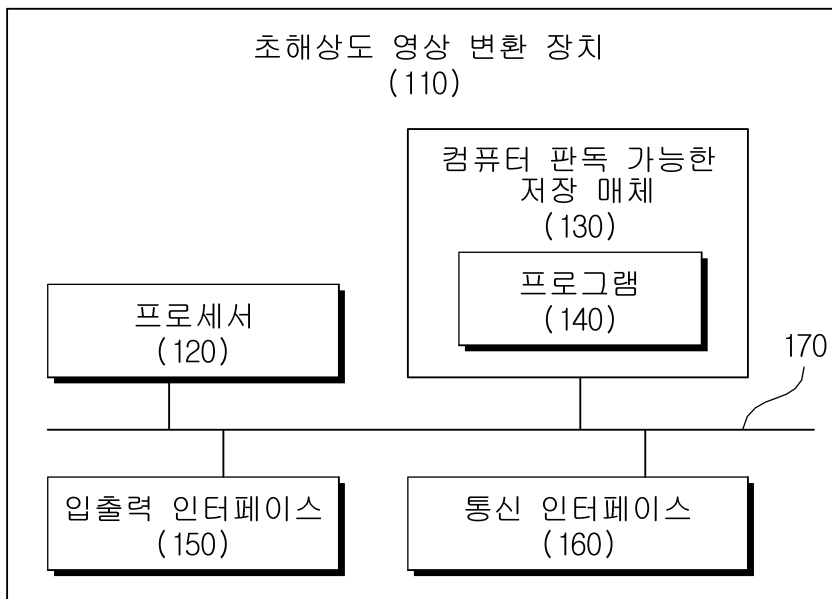
- [0054] 도 5는 본 발명의 다른 실시예에 따른 초해상도 영상 변환 방법을 예시한 흐름도이다. 초해상도 영상 변환 방법은 컴퓨팅 디바이스에 의하여 수행될 수 있으며, 초해상도 영상 변환 장치와 동일한 방식으로 동작한다.
- [0055] 단계 S210에서 프로세서는 영상을 입력받는다.
- [0056] 단계 S220에서 프로세서는 영상에 대해서 계층적 초해상도 모델을 통해 특징을 추출하고 누적한다.
- [0057] 단계 S230에서 프로세서는 특징을 기반으로 계층적 초해상도 모델을 통해 영상의 해상도를 증가시켜 다중 초해상도 영상을 출력한다.
- [0058] 프로세서는 기본 초해상도 영상을 출력하고, 중간 초해상도 영상을 출력하고, 최종 초해상도 영상을 출력한다. 프로세서는 N (N 은 자연수) 번째 계층에 해당하는 중간 초해상도 영상에 $N+1$ 번째 계층에 해당하는 중간 초해상도를 추가한다. 프로세서는 계층 전체의 특징을 모두 합하고 계층 전체의 특징으로부터 최종 초해상도 영상을 출력한다.
- [0059] 초해상도 영상 변환 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.
- [0060] 초해상도 영상 변환 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.
- [0061] 도 5에서는 각각의 과정을 순차적으로 실행하는 것으로 기재하고 있으나 이는 예시적으로 설명한 것에 불과하고, 이 분야의 기술자라면 본 발명의 실시예의 본질적인 특성에서 벗어나지 않는 범위에서 도 5에 기재된 순서를 변경하여 실행하거나 또는 하나 이상의 과정을 병렬적으로 실행하거나 다른 과정을 추가하는 것으로 다양하게 수정 및 변형하여 적용 가능할 것이다.
- [0062] 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.
- [0063] 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

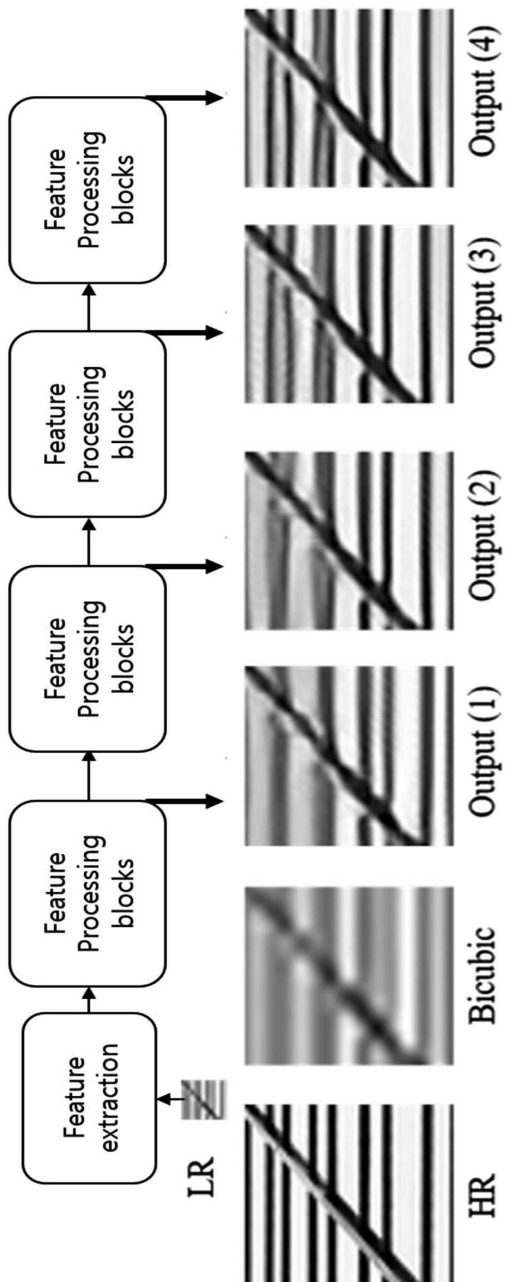
도면1



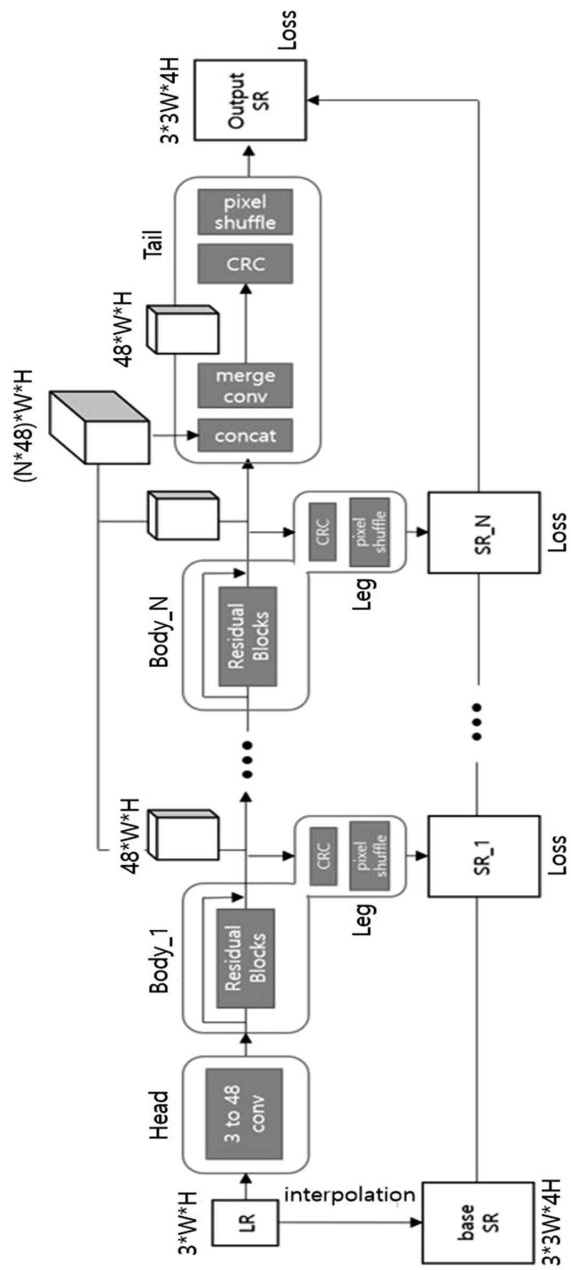
도면2



도면3



도면4



도면5

