(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2023년06월29일
(11) 등록번호 10-2549552
(24) 등록일자 2023년06월26일

(51) 국제특허분류(Int. Cl.)
G06T 3/40 (2006.01) *G06N 3/08* (2023.01)
H04N 19/105 (2014.01) *H04N 19/132* (2014.01)
(52) CPC특허분류
G06T 3/4053 (2013.01)
G06N 3/08 (2023.01)
(21) 출원번호 10-2021-0156430
(22) 출원일자 2021년11월15일
심사청구일자 2021년11월15일
(65) 공개번호 10-2023-0046174
(43) 공개일자 2023년04월05일
(30) 우선권주장
1020210128746 2021년09월29일 대한민국(KR)

(56) 선행기술조사문헌
Chunwei Tian et al., "Asymmetric CNN for image super-resolution", IEEE Transactions on Systems, Man, and Cybernetics: Systems, (2021.05.03.)*
Kwok-Wai Hung et al., "Real-Time Image Super-Resolution Using Recursive Depthwise Separable Convolution Network", IEEE Access, (2019.07.16.)*
Tianliang Fu et al., "An Enhanced Reference Structure For Reference icture Resampling (RPR) In VVC", 2021 IEEE International Conference on Image Processing (ICIP), (2021.09.22.)*
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
이화여자대학교 산학협력단
서울특별시 서대문구 이화여대길 52 (대현동, 이화여자대학교)
(72) 발명자
강제원
서울특별시 마포구 독막로 145, 104동 904호(창전동,서강쌍용예가아파트)
여은
인천광역시 연수구 학나래로6번길 42, 304호(선학동)
(74) 대리인
특허법인(유한)아이시스

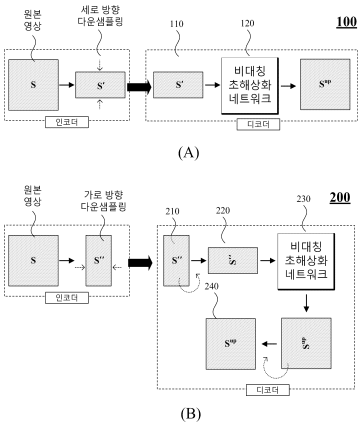
전체 청구항 수 : 총 8 항 심사관 : 이정은

(54) 발명의 명칭 **비대칭 초해상화 네트워크를 이용한 영상 업샘플링 방법 및 디코더**

(57) 요 약

비대칭 초해상화 네트워크를 이용한 영상 업샘플링 방법은 영상처리장치가 다운샘플링된 영상을 입력받는 단계 및 상기 영상처리장치가 상기 다운샘플링된 영상을 비대칭 초해상화 네트워크에 입력하여 업샘플링을 수행하는 단계를 포함한다. 상기 비대칭 초해상화 네트워크는 영상 프레임의 상기 제1 방향으로만 영상을 업샘플링한다.

대 표 도 - 도4



(52) CPC특허분류
G06T 3/4046 (2013.01)
H04N 19/105 (2015.01)
H04N 19/132 (2015.01)

이 발명을 지원한 국가연구개발사업
과제고유번호 1711140190
과제번호 2019R1C1C1010249
부처명 과학기술정보통신부
과제관리(전문)기관명 한국연구재단
연구사업명 개인기초연구(과기정통부)(R&D)
연구과제명 고화질 비디오의 장기간 예측과 해석 가능한 딥러닝 모델 기반의 비디오 신호처리 알고리즘 개발
기 여 율 1/2
과제수행기관명 이화여자대학교
연구기간 2021.03.01 ~ 2022.02.28
이 발명을 지원한 국가연구개발사업
과제고유번호 1711126409
과제번호 2020-0-01460-002
부처명 과학기술정보통신부
과제관리(전문)기관명 정보통신기획평가원
연구사업명 정보통신방송혁신인재양성(R&D)
연구과제명 Simulated Reality 콘텐츠 구현을 위한 핵심기술 연구 및 인력양성
기 여 율 1/2
과제수행기관명 이화여자대학교산학협력단
연구기간 2021.01.01 ~ 2021.12.31

명 세 서

청구범위

청구항 1

영상처리장치가 다운샘플링된 영상을 입력받는 단계; 및

상기 영상처리장치가 상기 다운샘플링된 영상을 비대칭 초해상화 네트워크에 입력하여 업샘플링을 수행하는 단계를 포함하되,

상기 다운샘플링된 영상은 VVC(Versatile Video Coding)의 RPR (Reference Picture Resampling) 매커니즘에 따라 다운샘플링된 것이고,

영상 프레임은 평면에서 제1 방향과 제2 방향의 축을 갖고,

상기 비대칭 초해상화 네트워크는 상기 영상 프레임의 상기 제1 방향으로만 영상을 업샘플링하되,

상기 다운샘플링된 영상이 상기 제1 방향으로만 다운샘플링된 영상인 경우, 상기 영상처리장치는 상기 다운샘플링된 영상을 비대칭 초해상화 네트워크에 입력하여 상기 제1 방향으로만 업샘플링하고,

상기 다운샘플링된 영상이 상기 제2 방향으로만 다운샘플링된 영상인 경우, 상기 영상처리장치는 상기 다운샘플링된 영상을 90도 회전시킨 후 상기 비대칭 초해상화 네트워크에 입력하여 영상을 업샘플링하는 비대칭 초해상화 네트워크를 이용한 영상 업샘플링 방법.

청구항 2

삭제

청구항 3

삭제

청구항 4

제1항에 있어서,

상기 다운샘플링된 영상이 상기 제1 방향 및 상기 제2 방향으로 다운샘플링된 경우,

상기 영상처리장치는 상기 다운샘플링된 영상을 상기 비대칭 초해상화 네트워크에 입력하여 영상을 업샘플링하고,

상기 업샘플링된 영상을 90도 회전시킨 후 상기 비대칭 초해상화 네트워크에 다시 입력하여 영상을 업샘플링하는 비대칭 초해상화 네트워크를 이용한 영상 업샘플링 방법.

청구항 5

제1항에 있어서,

상기 비대칭 초해상화 네트워크는 영상의 이미지 해상도를 업샘플링하는 디컨볼루션 계층을 포함하고, 상기 디컨볼루션 계층은 커널의 크기가 비대칭인 비대칭 초해상화 네트워크를 이용한 영상 업샘플링 방법.

청구항 6

제1항에 있어서,

상기 비대칭 초해상화 네트워크는 컨볼루션 계층을 포함하고, 상기 컨볼루션 계층은 분리 컨볼루션(separable convolution)을 사용하여 원본 영상의 크기를 갖는 상기 제2 방향으로는 초해상화를 수행하지 않는 비대칭 초해상화 네트워크를 이용한 영상 업샘플링 방법.

청구항 7

다운샘플링된 영상을 입력받는 입력장치;

영상 프레임이 제1 방향과 제2 방향의 축을 갖고, 상기 영상 프레임의 상기 제1 방향으로만 영상을 업샘플링하는 비대칭 초해상화 네트워크를 저장하는 저장장치; 및

상기 다운샘플링된 영상을 비대칭 초해상화 네트워크에 입력하여 업샘플링을 수행하는 연산장치를 포함하되,

상기 다운샘플링된 영상이 상기 제1 방향으로만 다운샘플링된 영상인 경우, 상기 연산장치는 상기 다운샘플링된 영상을 비대칭 초해상화 네트워크에 입력하여 상기 제1 방향으로만 업샘플링하고,

상기 다운샘플링된 영상이 상기 제2 방향으로만 다운샘플링된 영상인 경우, 상기 연산장치는 상기 다운샘플링된 영상을 90도 회전시킨 후 상기 비대칭 초해상화 네트워크에 입력하여 영상을 업샘플링하는 비대칭 초해상화 네트워크를 이용하여 영상을 업샘플링하는 디코더.

청구항 8

삭제

청구항 9

삭제

청구항 10

제7항에 있어서,

상기 다운샘플링된 영상이 상기 제1 방향 및 상기 제2 방향으로 다운샘플링된 경우,

상기 연산장치는 상기 다운샘플링된 영상을 상기 비대칭 초해상화 네트워크에 입력하여 영상을 업샘플링하고,

상기 업샘플링된 영상을 90도 회전시킨 후 상기 비대칭 초해상화 네트워크에 다시 입력하여 영상을 업샘플링하는

비대칭 초해상화 네트워크를 이용하여 영상을 업샘플링하는 디코더.

청구항 11

제7항에 있어서,

상기 비대칭 초해상화 네트워크는 영상의 이미지 해상도를 업샘플링하는 디컨볼루션 계층을 포함하고, 상기 디컨볼루션 계층은 커널의 크기가 비대칭인 비대칭 초해상화 네트워크를 이용하여 영상을 업샘플링하는 디코더.

청구항 12

제7항에 있어서,

상기 비대칭 초해상화 네트워크는 컨볼루션 계층을 포함하고, 상기 컨볼루션 계층은 분리 컨볼루션(separable convolution)을 사용하여 원본 영상의 크기를 갖는 상기 제2 방향으로는 초해상화를 수행하지 않는 비대칭 초해상화 네트워크를 이용하여 영상을 업샘플링하는 디코더.

발명의 설명

기 술 분 야

[0001] 이하 설명하는 기술은 비대칭적으로 초해상화를 수행하는 신경망 모델을 이용하여 영상을 업샘플링하는 기법에 관한 것이다.

배 경 기 술

[0002] 최근 표준화가 완료된 VVC(Versatile Video Coding)에 대한 표준화 과정에서 실시간 비디오 원격 통신 및 회의 상황과 같은 다양한 비디오 서비스 환경에서도 빠른 스트리밍을 위한 적응적 해상도 변경에 대한 논의가 있었다. 이에 VVC는 현재 영상과 다른 해상도를 가지는 영상을 참조할 수 있는 RPR (Reference Picture Resampling) 기술을 포함하게 되었다. RPR은 참조 프레임의 해상도가 중간에 변경되어 전체 비디오 비트 스트림

의 크기를 줄이면서 줄어든 해상도에 의하여 적은 복잡도로 부호화를 수행할 수 있는 장점이 있다. 그에 반하여 프레임을 복호하는 경우에는 줄어든 해상도를 다시 복원하기 위한 업샘플링 기술이 필요하다.

[0003] RPR 기술은 현재 픽쳐 또는 블록의 업샘플링 또는 다운샘플링 기법을 사용한다. RPR는 가로 및/또는 세로 방향으로 서로 다른 비대칭적인 스케일링이 가능하다. 즉, RPR은 원본 영상에서 세로 방향, 가로 방향 또는 가로 및 세로 방향으로 다운샘플링을 할 수 있다.

선행기술문헌

특허문헌

[0004] (특허문헌 0001) 한국등록특허 제10-1690139호

발명의 내용

해결하려는 과제

[0005] 이하 설명하는 기술은 종래 VVC의 RPR에서 비대칭적인 스케일링을 하는 경우에도 적용 가능한 영상 업샘플링 기법을 제공하고자 한다. 나아가, 이하 설명하는 기술은 가능한 모든 방향으로 스케일링된 경우에도 적응적으로 업샘플링이 가능한 기법을 제공하고자 한다.

과제의 해결 수단

[0006] 비대칭 초해상화 네트워크를 이용한 영상 업샘플링 방법은 영상처리장치가 다운샘플링된 영상을 입력받는 단계 및 상기 영상처리장치가 상기 다운샘플링된 영상을 비대칭 초해상화 네트워크에 입력하여 업샘플링을 수행하는 단계를 포함한다. 영상 프레임은 평면에서 제1 방향과 제2 방향의 축을 갖고, 상기 비대칭 초해상화 네트워크는 상기 영상 프레임의 상기 제1 방향으로만 영상을 업샘플링한다.

[0007] 대칭 초해상화 네트워크를 이용하여 영상을 업샘플링하는 디코더는 다운샘플링된 영상을 입력받는 입력장치, 영상 프레임이 제1 방향과 제2 방향의 축을 갖고, 상기 영상 프레임의 상기 제1 방향으로만 영상을 업샘플링하는 비대칭 초해상화 네트워크를 저장하는 저장장치 및 상기 다운샘플링된 영상을 비대칭 초해상화 네트워크에 입력하여 업샘플링을 수행하는 연산장치를 포함한다.

발명의 효과

[0008] 이하 설명하는 기술은 비대칭 부호화를 통해 원본 영상 대비 전송되는 비트율을 줄일 수 있고, 대칭 부호화에 비해 정보 손실을 줄여 초해상화 과정을 통해 더 높은 화질을 제공할 수 있다. 또한, 이하 설명하는 기술은 종전 대칭 초해상화 대비 연상량과 복잡도가 낮다. 따라서, 이하 설명하는 기술은 4K와 같은 대용량 영상 데이터 처리에 적합하다.

도면의 간단한 설명

[0009] 도 1은 Sharp 사가 제안한 신경망 기반의 초해상화 기법이다.

도 2는 Qualcomm 사가 제안한 신경망 기반의 초해상화 기법이다.

도 3은 Bytedance 사가 제안한 신경망 기반의 초해상화 기법이다.

도 4는 비대칭 RPR 적응적인 업샘플링 과정에 대한 예이다.

도 5는 RPR 적응적인 업샘플링 과정에 대한 다른 예이다.

도 6은 초해상화 네트워크에 대한 예이다.

도 7은 RPR 적응적 업샘플링을 하는 디코더에 대한 예이다.

발명을 실시하기 위한 구체적인 내용

[0010] 이하 설명하는 기술은 다양한 변경을 가할 수 있고 여러 가지 실시예를 가질 수 있는 바, 특정 실시예들을 도면

에 예시하고 상세하게 설명하고자 한다. 그러나, 이는 이하 설명하는 기술을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 이하 설명하는 기술의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다.

[0011] 제1, 제2, A, B 등의 용어는 다양한 구성요소들을 설명하는데 사용될 수 있지만, 해당 구성요소들은 상기 용어들에 의해 한정되지는 않으며, 단지 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 사용된다. 예를 들어, 이하 설명하는 기술의 권리 범위를 벗어나지 않으면서 제1 구성요소는 제2 구성요소로 명명될 수 있고, 유사하게 제2 구성요소도 제1 구성요소로 명명될 수 있다. 및/또는 이라는 용어는 복수의 관련된 기재된 항목들의 조합 또는 복수의 관련된 기재된 항목들 중의 어느 항목을 포함한다.

[0012] 본 명세서에서 사용되는 용어에서 단수의 표현은 문맥상 명백하게 다르게 해석되지 않는 한 복수의 표현을 포함하는 것으로 이해되어야 하고, "포함한다" 등의 용어는 설명된 특징, 개수, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것이 존재함을 의미하는 것이지, 하나 또는 그 이상의 다른 특징들이나 개수, 단계 동작 구성요소, 부분품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 배제하지 않는 것으로 이해되어야 한다.

[0013] 도면에 대한 상세한 설명을 하기에 앞서, 본 명세서에서의 구성부들에 대한 구분은 각 구성부가 담당하는 주기능 별로 구분한 것에 불과함을 명확히 하고자 한다. 즉, 이하에서 설명할 2개 이상의 구성부가 하나의 구성부로 합쳐지거나 또는 하나의 구성부가 보다 세분화된 기능별로 2개 이상으로 분화되어 구비될 수도 있다. 그리고 이하에서 설명할 구성부 각각은 자신이 담당하는 주기능 이외에도 다른 구성부가 담당하는 기능 중 일부 또는 전부의 기능을 추가적으로 수행할 수도 있으며, 구성부 각각이 담당하는 주기능 중 일부 기능이 다른 구성부에 의해 전담되어 수행될 수도 있음은 물론이다.

[0014] 또, 방법 또는 동작 방법을 수행함에 있어서, 상기 방법을 이루는 각 과정들은 문맥상 명백하게 특정 순서를 기재하지 않은 이상 명기된 순서와 다르게 일어날 수 있다. 즉, 각 과정들은 명기된 순서와 동일하게 일어날 수도 있고 실질적으로 동시에 수행될 수도 있으며 반대의 순서대로 수행될 수도 있다.

[0016] 초해상화 (Super Resolution) 기법은 저해상도 영상을 고해상도로 변환하기 위해 사용하는 방법이다. 적은 복잡도와 연산량을 갖는 바이큐빅 또는 선형 보간법과 같은 방법을 이용하여 저해상도 영상의 해상도를 생성할 수 있다. 최근 신경망 네트워크의 발전으로 인하여 CNN (Convolutional Neural Network) 기반의 초해상화 네트워크가 연구되었다.

[0017] VVC에서는 RPR 매커니즘이 도입되어 입력된 비디오의 해상도를 특정 배율에 맞추어 변형할 수 있도록 하였다. 또한, 입력된 비디오의 해상도를 RPR을 이용하여 기존 해상도보다 작은 크기로 부호화 및 복호화를 진행한 뒤 포스트 필터로서 초해상화 기법을 적용하여 원본 영상과 동일한 해상도로 복원하는 방법이 제안되었다.

[0018] VVC 표준화 회의에서도 신경망 기반의 초해상화 기법에 대한 논의가 있었다. 주요하게 논의되었던 초해상화 기법에 대하여 간략하게 설명한다.

[0019] 도 1은 Sharp 사가 제안한 신경망 기반의 초해상화 기법이다. 도 1(A)는 JVET 표준화 회의에서 Sharp 사가 제안한 RPR을 이용한 신경망 기반의 초해상화 프레임워크이다. 도 1(B)는 Sharp사에서 제안한 초해상화 네트워크 구조이다. VTM (VVC Test Model) 10.0 소프트웨어에서 RPR을 적용하여 여러 4K 영상의 가로, 세로 해상도를 각각 1.5배 혹은 2배로 다운 샘플링하여 부호화 및 복호화를 한다. 복호화 된 영상을 도 1(B)와 같은 신경망 기반의 네트워크를 통과시켜 원본의 해상도와 같은 해상도로 초해상화를 수행한다. 모든 입력 영상에 고정된 신경망 파라미터를 사용하게 된다면 파라미터를 따로 전송해줄 필요는 없고, 각 영상과 양자화 파라미터에 맞게 가변적인 파라미터를 사용하려면 디코더로 해당 파라미터를 전송해주어야 한다. 이는 SEI(supplemental enhancement information) 메시지를 통해 전송되며 디코더의 복잡도를 줄이기 위해 MPEG의 표준 중 하나인 NNR을 적용하여 파라미터 압축을 시행한다.

[0020] 도 2는 JVET 표준화 회의에서 Qualcomm 사가 제안한 신경망 기반의 초해상화 기법이다. 도 2(A)는 JVET 표준화 회의에서 Qualcomm 사가 제안한 RPR을 이용한 신경망 기반의 초해상화 프레임워크이다. 도 2(B)는 Qualcomm 사에서 제안한 초해상화 네트워크 구조이다. Sharp 사와 마찬가지로 VTM 11.0 소프트웨어에 내장된 RPR 매커니즘을 사용하여 입력된 원본 영상을 가로, 세로 각각 2배씩 다운 샘플링을 수행한 뒤 부호화와 복호화 과정을 거친다. 복호화된 영상은 복잡도를 줄인 단일 스케일의 EDSR 네트워크(도2(B))를 사용하여 원본 영상과 같은 해상도로 초해상화을 실행한다.

[0021] 도 3은 JVET 표준화 회의에서 Bytedance 사가 제안한 신경망 기반의 초해상화 기법이다. 앞에서 소개한 두 회사에서 제안한 내용과 마찬가지로 VTM 11.0 소프트웨어의 RPR을 이용하여 입력된 원본 영상의 가로, 세로를 각각

다운샘플링 하여 부호화 및 복호화를 진행한다. 도 3(A)는 Bytedance 사에서 제안한 복호화된 루마 영상을 초해상화 하는 네트워크이다. 도 3(B)는 Bytedance 사에서 제안한 복호화된 크로마 영상을 초해상화 하는 네트워크이다. 출력된 복호화된 영상은 루마와 크로마 채널로 각각 분리하여 신경망 기반의 네트워크의 입력으로 사용한다. 루마 업샘플링 네트워크의 경우 다운샘플링된 루마채널의 입력과 양자화 파라미터, 루마 채널의 예측 신호를 입력으로 사용하여 도 3(A)의 네트워크를 통과시켜 초해상화를 진행한다. 크로마 업샘플링 네트워크는 도 3(B)와 같은 구조로 구성되어있다. 이는 루마 업샘플링 네트워크와 유사한 구조로 도 3(A)에서 출력된 루마 채널의 영상을 다시 다운샘플링한 입력과 양자화 파라미터, 다운샘플링된 크로마 채널의 예측 신호를 입력으로 취해 도 3(B)의 네트워크를 통과시켜 초해상화를 수행한다.

[0022] 정리하면 종래 신경망 기반 RPR 방법들은 가로 W, 세로 H 크기의 영상이 입력되면. 가로 W/2 및 세로 H/2로 다운샘플링하고, EDSR 등의 초해상화 네트워크를 이용해서 가로 W 및 세로 H 크기의 영상으로 복원하는 방식을 사용한다.

[0024] 이하 설명하는 기술은 VVC의 RPR 과정에서 수행되는 업샘플링 기술이다. 특히 이하 설명하는 기술은 선택적인 비대칭 RPR에 대해서도 업샘플링이 가능한 기술이다. 비대칭 RPR은 세로(수직) 방향 또는 가로(수평) 방향으로만 영상이 다운샘플링된 경우를 의미한다.

[0025] 이하 설명하는 기술은 비대칭 초해상화 네트워크를 사용한다. 비대칭 초해상화 네트워크는 2차 프레임에서 하나의 방향으로만 초해상화를 수행하는 신경망 네트워크이다. 이하 비대칭 초해상화 네트워크는 세로(수직) 방향으로만 초해상화를 수행한다고 가정한다.

[0026] 이하 설명하는 업샘플링 과정은 영상의 디코더에서 수행될 수 있다. 나아가, 디코더를 포함하는 영상처리장치에서 비대칭 초해상화 네트워크를 이용한 업샘플링을 수행할 수도 있다.

[0027] 도 4는 비대칭 RPR 적응적인 업샘플링 과정에 대한 예이다.

[0028] 도 4(A)는 비대칭 RPR 매커니즘으로 세로 방향으로 다운샘플링되어 부호화된 경우이다. 인코더는 원본 영상 S를 세로 방향으로만 다운샘플링하여 S'를 생성하고, 일정한 부호화 과정을 통해 S'를 부호화한다. 도 4(A)는 세로 방향 비대칭 RPR에 대한 업샘플링 과정(100)을 도시한다. 디코더는 S'를 포함하는 스트림을 입력받는다. 디코더는 일정한 복호 과정을 통해 영상 S'를 복호한다(110). 디코더는 복호한 영상 S'를 비대칭 초해상화 네트워크에 입력하여 영상을 업샘플링한다(120). 비대칭 초해상화 네트워크는 세로 방향으로만 업샘플링을 수행한다. 업샘플링된 영상 S^{up}은 원본 영상 S와 동일한 해상도를 갖는 영상이다.

[0029] 도 4(B)는 비대칭 RPR 매커니즘으로 가로 방향으로 다운샘플링되어 부호화된 경우이다. 인코더는 원본 영상 S를 가로 방향으로만 다운샘플링하여 S''를 생성하고, 일정한 부호화 과정을 통해 S''를 부호화한다. 도 4(B)는 가로 방향 비대칭 RPR에 대한 업샘플링 과정(200)을 도시한다. 디코더는 S''를 포함하는 스트림을 입력받는다. 디코더는 일정한 복호 과정을 통해 영상 S''를 복호한다(210). 디코더는 복호한 영상 S''를 90° 회전시킨다(220). 디코더는 회전된 영상 S''를 비대칭 초해상화 네트워크에 입력하여 영상을 업샘플링한다(230). 비대칭 초해상화 네트워크는 세로 방향으로만 업샘플링을 수행한다. 업샘플링된 영상 S^{up}은 원본 영상 S와 동일한 해상도를 갖는 영상이다. 디코더는 업샘플링된 영상 S^{up}을 220과정과 반대 방향으로 회전시킨다(240). 도 4에서 디코더는 220 과정에서 반시계방향으로 프레임을 회전하고, 240에서 시계방향으로 프레임을 회전하는 예이다. 물론, 디코더는 220 과정에서 시계방향으로 프레임을 회전하고, 240 과정에서 반시계방향으로 프레임을 회전할 수도 있다.

[0031] 나아가, 디코더는 전술한 비대칭 초해상화 네트워크를 이용하여 가로 및 세로 방향으로 모두 다운샘플링이된 영상도 업샘플링할 수 있다. 도 5는 RPR 적응적인 업샘플링 과정(300)에 대한 다른 예이다. 도 5는 가로 및 세로 방향으로 RPR 매커니즘이 적용된 경우에 대한 업샘플링의 예이다. 인코더는 원본 영상 S를 가로 및 세로 방향으로 다운샘플링하여 S'''를 생성하고, 일정한 부호화 과정을 통해 S'''를 부호화한다.

[0032] 디코더는 S'''를 포함하는 스트림을 입력받는다. 디코더는 일정한 복호 과정을 통해 영상 S'''를 복호한다(310). 디코더는 복호한 영상 S'''를 초해상화 네트워크에 입력하여 영상을 업샘플링한다(320). 비대칭 초해상화 네트워크는 세로 방향으로만 업샘플링을 수행한다. 디코더는 업샘플링한 영상 S^{up}를 90° 회전시킨다(330). 디코더는 회전된 영상 S^{up}를 비대칭 초해상화 네트워크에 입력하여 영상을 다시 업샘플링한다(340). 비대칭 초해상화 네트워크는 세로 방향으로만 업샘플링을 수행한다. 결과적으로 업샘플링된 영상 S'^{up}은 원본 영상 S와 동일한

해상도를 갖는 영상이다. 디코더는 업샘플링된 영상 S'^{up}을 330과정과 반대 방향으로 회전시킨다(350). 도 5에서 디코더는 330 과정에서 반시계방향으로 프레임을 회전하고, 350에서 시계방향으로 프레임을 회전하는 예이다. 물론, 디코더는 330 과정에서 시계방향으로 프레임을 회전하고, 350 과정에서 반시계방향으로 프레임을 회전할 수도 있다.

[0034] 신경망 기반의 초해상화 네트워크의 종류 및 구조는 다양할 수 있다. 설명의 편의를 위하여 FSR-CNN의 구조를 기준으로 설명한다. 다만, 비대칭 초해상화 네트워크는 하나의 방향(가로 또는 세로)으로만 업샘플링을 하면 된다. 따라서, 이하 설명하는 구성은 다른 종류의 네트워크에도 적용 가능하다. 나아가, 도 4 및 도 5에서 비대칭 비율을 절반으로 설명하였지만, 비대칭 비율은 1.25, 1.75 등 다양한 비율이 사용될 수도 있다. 비대칭 초해상화 네트워크는 해당 비대칭 비율에 맞게 일방향으로 영상을 업샘플링하도록 구성되면 된다.

[0036] 도 6은 초해상화 네트워크에 대한 예이다.

[0037] 도 6(A)는 FSR-CNN 네트워크에 대한 예이다. FSR-CNN 네트워크는 입력되는 저해상도 이미지(LR image)를 가로 및 세로 방향으로 업샘플링하여 고해상도 이미지(SR image)를 생성한다. FSR-CNN은 각 계층을 통하여 이미지 특징 추출(Feature extraction), 특징 차원 축소(Shrinking), 매핑(Mapping), 특징 차원 증가(Expanding) 및 디컨볼루션(Deconvolution)을 수행한다. 기본적인 각 계층의 필터 크기(Filter size), 필터 개수(# of Filters) 및 입력 채널의 개수(# of Ch.)는 도 6(A)에 도시한 바와 같다.

[0038] 특징 추출 계층은 SR-CNN과 비슷하지만 FSR-CNN의 입력은 보간되지 않고, 바로 첫 번째 CNN(Convolutional. Neural Network)에 입력된다. 특징 차원 축소 계층은 매핑과정으로 통과되기 전에 특징의 차원을 줄여 연산 복잡도를 줄여준다. 주로 특징의 차원을 증가시키거나 축소할 때 CNN의 커널크기를 1×1로 고정하여 연산 복잡도를 줄인다. 매핑 계층은 초해상화 품질에 중요한 부분이다. 특히 CNN 층에 필터 개수(width)와 CNN 층의 수(depth)가 초해상화 성능에 가장 영향을 미치는 부분이라 할 수 있다. FSR-CNN은 커널의 크기를 3×3으로 축소하고 여러 개의 층으로 겹겹이 쌓아 SR-CNN보다 좋은 성능을 보인다. 디컨볼루션 계층은 FSR-CNN의 마지막 계층으로, 실질적으로 이미지 해상도를 업샘플링하여 고해상도 영상을 생성하는 구성이다. 디컨볼루션 계층은 스트라이드(stride) 부분을 조정하여 정수 단위로 해상도의 크기를 조절할 수 있다.

[0039] 도 6(B)는 일방향으로만 업샘플링을 수행하는 변형된 FSR-CNN 네트워크의 예이다. 도 6(B)의 네트워크는 도 6(A)와 유사하지만, 디컨볼루션 계층의 구성에 차이가 있다. 일반적인 FSR-CNN의 디컨볼루션 계층의 커널 크기가 9×9이다. 이에 반하여 도 6(B)의 디컨볼루션 계층의 커널 크기는 9×3이다. 따라서, 도 6(B)의 네트워크는 비대칭적인 업샘플링을 수행하게 된다. 또한, 도 6(B)의 네트워크는 커널의 크기를 줄여서 종전 FSR-CNN보다 연산량과 복잡도가 줄어든다. 도 6(B)의 네트워크는 입력되는 저해상도 이미지(LR image)를 세로 방향으로만 업샘플링하여 고해상도 이미지(SR image)를 생성한다.

[0041] 한편, 도 6과는 다른 접근으로 일 방향 업샘플링을 수행할 수도 있다. 예컨대, 종전의 초해상화를 위한 신경망 네트워크를 이용하되 분리 컨볼루션(Separable convolution)을 적용하여 일 방향으로만 업샘플링 되도록 할 수 있다. 예컨대, 입력 영상이 가로 또는 세로 방향으로 RPR을 수행하여 복호화된 영상의 해상도는 설정한 배율에 맞게 다운샘플링 된다. 이때 종전의 초해상화 네트워크에서 컨볼루션을 수행할 때 분리 컨볼루션을 적용하여 원본 영상과 동일한 크기를 갖는 방향에 대해서는 초해상화가 진행되지 않게 할 수 있다. 예컨대, 인코더에서 세로 방향으로만 다운샘플링 하였다면, 디코더는 가로 방향으로는 초해상화를 진행하지 않도록 분할 컨볼루션이 적용된 초해상화 네트워크를 이용하여 영상을 업샘플링할 수 있다. 종전에 연구되었던 다양한 초해상화 네트워크(도 1 내지 도 3에서 설명한 네트워크 포함)에 분리 컨볼루션을 적용할 수 있다.

[0043] 도 7은 RPR 적응적 업샘플링을 하는 디코더(400)에 대한 예이다. 디코더(400)는 물리적으로 다양한 형태일 수 있다. 예컨대, 디코더(400)는 PC와 같은 컴퓨터 장치, 네트워크의 서버, 데이터 처리 전용 칩셋, VR 장치 등의 형태를 가질 수 있다.

[0044] 디코더(400)는 저장장치(410), 메모리(420), 연산장치(430), 인터페이스 장치(440), 통신장치(450) 및 출력장치(460)를 포함할 수 있다.

[0045] 저장장치(410)는 인코더가 전달하는 영상 스트림을 저장할 수 있다.

[0046] 저장장치(410)는 영상 스트림을 복호하기 위한 코덱 내지 프로그램을 저장할 수 있다.

[0047] 저장장치(410)는 전술한 비대칭 초해상화 네트워크를 저장할 수 있다.

[0048] 저장장치(410)는 업샘플링한 영상을 저장할 수 있다.

[0049] 메모리(420)는 디코더(400)가 영상을 업샘플링하는 과정 등에서 생성되는 데이터 및 정보 등을 저장할 수 있다.

[0050] 인터페이스 장치(440)는 외부로부터 일정한 명령 및 데이터를 입력받는 장치이다. 인터페이스 장치(440)는 물리적으로 연결된 입력 장치 또는 외부 저장장치로부터 다운샘플링된 영상을 입력을 입력받을 수 있다. 인터페이스 장치(440)는 업샘플링된 영상을 다른 객체에 전달할 수도 있다.

[0051] 통신장치(450)는 유선 또는 무선 네트워크를 통해 일정한 정보를 수신하고 전송하는 구성을 의미한다. 통신장치(450)는 외부 객체로부터 다운샘플링된 영상을 수신할 수 있다. 또는 통신장치(450)는 업샘플링된 영상을 사용자 단말과 같은 외부 객체에 송신할 수도 있다.

[0052] 인터페이스 장치(440) 및 통신장치(450)는 사용자 또는 다른 물리적 객체로부터 일정한 데이터를 주고받는 구성이므로, 포괄적으로 입출력장치라고도 명명할 수 있다. 다운샘플링된 영상을 입력받는 기능에 한정하면 인터페이스 장치(440) 및 통신장치(450)는 입력장치라고 할 수도 있다.

[0053] 출력장치(460)는 일정한 정보를 출력하는 장치이다. 출력장치(460)는 영상 처리 과정에 필요한 인터페이스를 출력할 수 있다. 또한, 출력장치(460)는 업샘플링된 영상을 출력할 수 있다.

[0054] 연산 장치(430)는 저장장치(410)에 저장된 프로그램 및 비대칭 초해상화 네트워크를 이용하여 다운샘플링된 영상을 업샘플링할 수 있다.

[0055] 연산 장치(430)는 일정한 복호 방식에 따라 입력된 영상을 복호할 수 있다. 복호된 영상은 전술한 바와 같이 다운샘플링된 영상이다. 복호된 영상은 가로 또는 세로 방향으로만 비대칭적으로 다운샘플링된 영상일 수 있다. 또는 복호된 영상을 가로 및 세로 방향으로 다운샘플링된 영상일 수도 있다.

[0056] 연산 장치(430)는 영상 스트림을 복호하면서 특정 필드에 있는 정보를 기준으로 현재 다운샘플링된 영상의 압축 방향을 알 수 있다.

[0057] 연산 장치(430)는 세로 방향으로만 다운샘플링된 영상이라면, 도 4(A)에서 설명한 바와 같이 비대칭 초해상화 네트워크에 복호한 영상을 입력하여 업샘플링을 할 수 있다.

[0058] 연산 장치(430)는 가로 방향으로만 다운샘플링된 영상이라면, 도 4(B)에서 설명한 바와 같이 영상 회전과 비대칭 초해상화 네트워크를 이용한 초해상화를 통해 영상을 업샘플링할 수 있다.

[0059] 연산 장치(430)는 세로 방향 및 가로 방향으로 다운샘플링된 영상이라면, 도 5에서 설명한 바와 같이 영상 회전과 비대칭 초해상화 네트워크를 이용한 초해상화를 통해 영상을 업샘플링할 수 있다.

[0060] 연산 장치(430)는 데이터를 처리하고, 일정한 연산을 처리하는 프로세서, AP, 프로그램이 임베디드된 칩과 같은 장치일 수 있다.

[0062] 또한, 상술한 바와 같은 VVC의 RPR 비대칭 다운샘플링 영상에 대한 업샘플링 방법은 컴퓨터에서 실행될 수 있는 실행가능한 알고리즘을 포함하는 프로그램(또는 어플리케이션)으로 구현될 수 있다. 상기 프로그램은 일시적 또는 비일시적 판독 가능 매체(non-transitory computer readable medium)에 저장되어 제공될 수 있다.

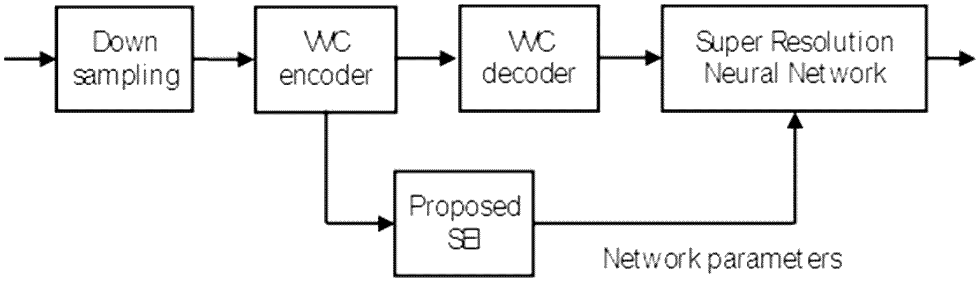
[0063] 비일시적 판독 가능 매체란 레지스터, 캐쉬, 메모리 등과 같이 짧은 순간 동안 데이터를 저장하는 매체가 아니라 반영구적으로 데이터를 저장하며, 기기에 의해 판독(reading)이 가능한 매체를 의미한다. 구체적으로는, 상술한 다양한 어플리케이션 또는 프로그램들은 CD, DVD, 하드 디스크, 블루레이 디스크, USB, 메모리카드, ROM(read-only memory), PROM (programmable read only memory), EPROM(Erasable PROM, EPROM) 또는 EEPROM(Electrically EPROM) 또는 플래시 메모리 등과 같은 비일시적 판독 가능 매체에 저장되어 제공될 수 있다.

[0064] 일시적 판독 가능 매체는 스태틱 램(Static RAM, SRAM), 다이내믹 램(Dynamic RAM, DRAM), 싱크로너스 디램(Synchronous DRAM, SDRAM), 2배속 SDRAM(Double Data Rate SDRAM, DDR SDRAM), 증강형 SDRAM(Enhanced SDRAM, ESDRAM), 동기화 DRAM(Synclink DRAM, SLDRAM) 및 직접 램버스 램(Direct Rambus RAM, DRRAM) 과 같은 다양한 RAM을 의미한다.

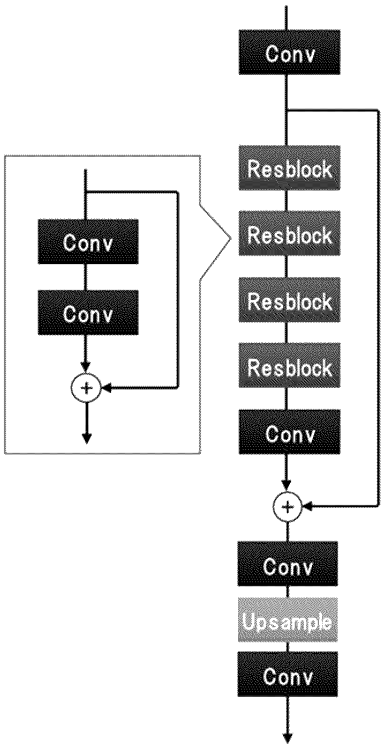
[0067] 본 실시예 및 본 명세서에 첨부된 도면은 전술한 기술에 포함되는 기술적 사상의 일부를 명확하게 나타내고 있는 것에 불과하며, 전술한 기술의 명세서 및 도면에 포함된 기술적 사상의 범위 내에서 당업자가 용이하게 유추할 수 있는 변형 예와 구체적인 실시례는 모두 전술한 기술의 권리범위에 포함되는 것이 자명하다고 할 것이다.

도면

도면1


Down sampling
WC encoder
WC decoder
Super Resolution Neural Network
Proposed SE
Network parameters

(A)

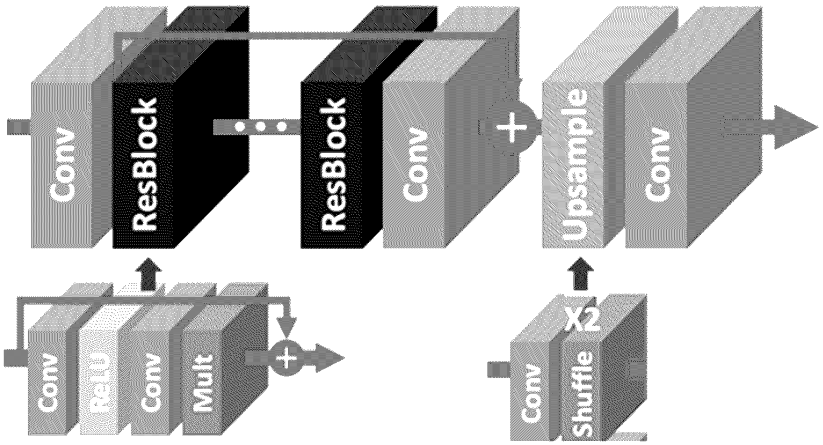

Conv
Resblock
Resblock
Resblock
Resblock
Conv
Conv
Conv
Conv
Upsample
Conv

(B)

도면2

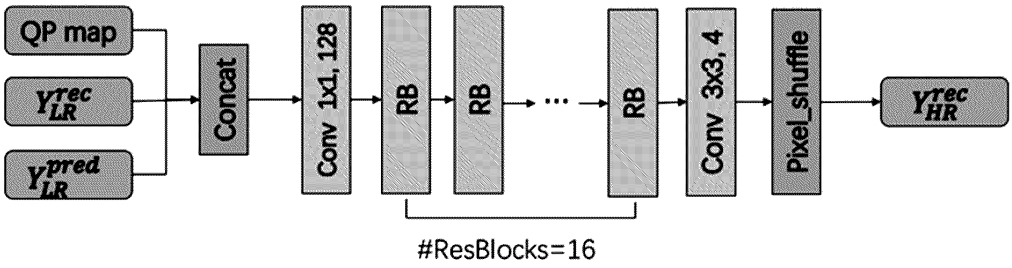

2x
Encode Frame
bitstream
Decode Frame
2x
Original sequence
VTM-11.0 Encoder
VTM-11.0 Decoder
Output pictures

(A)

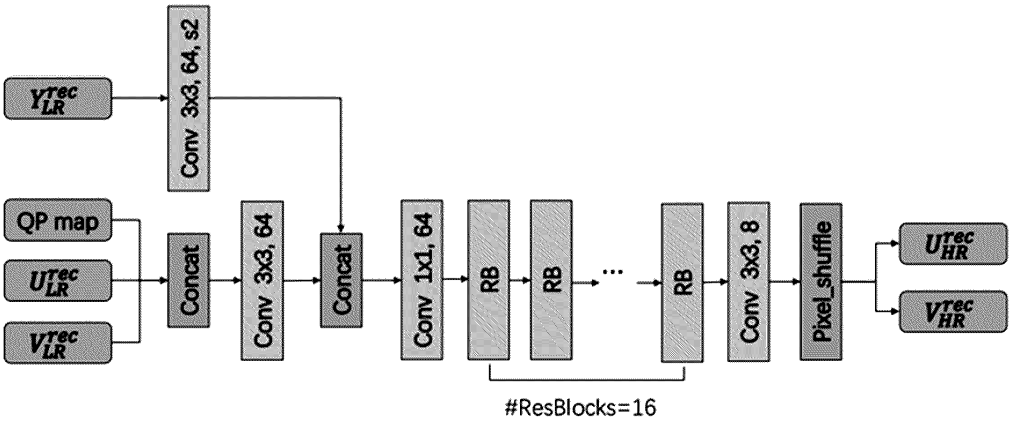

Conv
ResBlock
ResBlock
Conv
Upsample
Conv
Conv
ReLU
Conv
Mult
Conv
X2
Shuffle

(B)

도면3

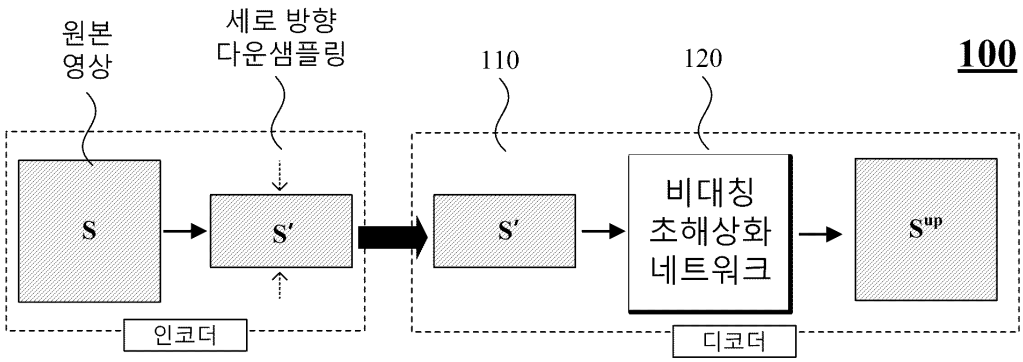

QP map
Y_{LR}^{rec}
Y_{LR}^{pred}
Concat
Conv 1x1, 128
RB
RB
RB
Conv 3x3, 4
Pixel_shuffle
Y_{HR}^{rec}
#ResBlocks=16

(A)

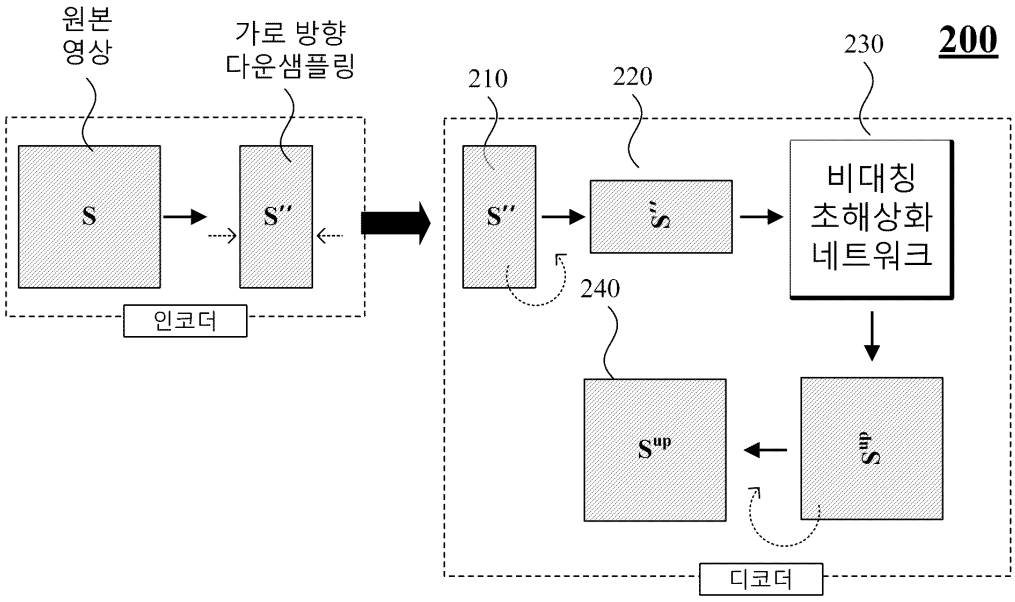

Y_{LR}^{rec}
Conv 3x3, 64, s2
QP map
U_{LR}^{rec}
V_{LR}^{rec}
Concat
Conv 3x3, 64
Concat
Conv 1x1, 64
RB
RB
RB
Conv 3x3, 8
Pixel_shuffle
U_{HR}^{rec}
V_{HR}^{rec}
#ResBlocks=16

(B)

도면4

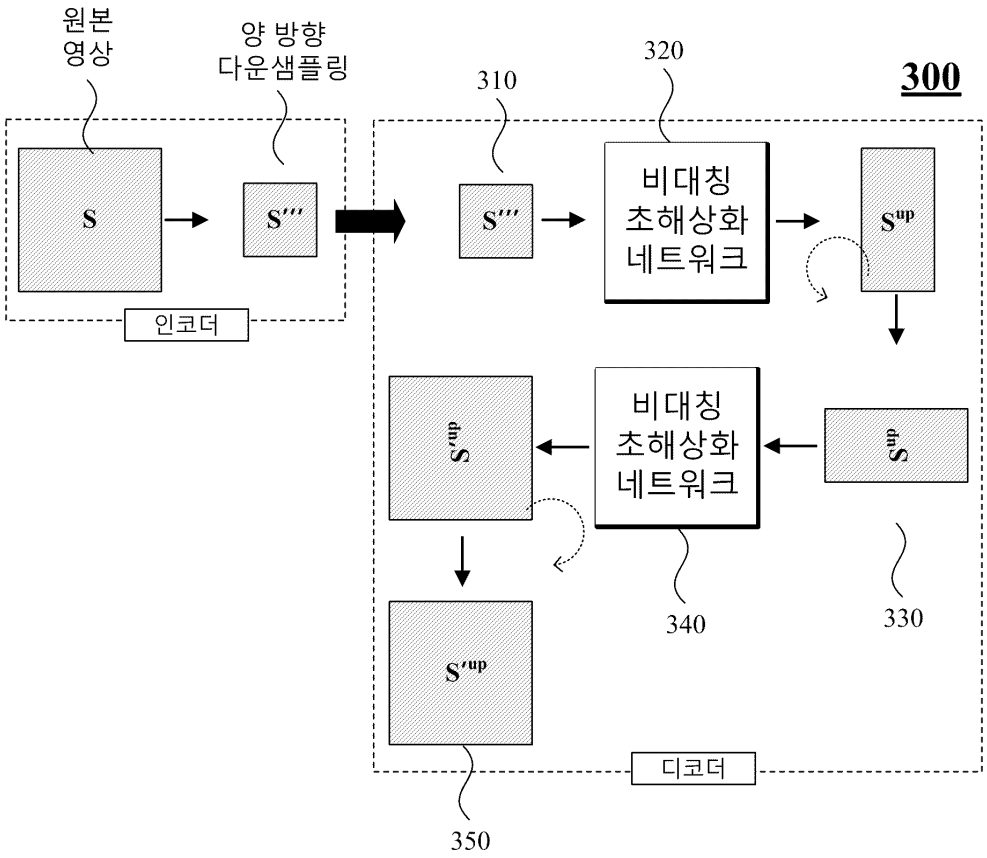

원본
영상
세로 방향
다운샘플링
110
120
100
S
S′
S′
비대칭
초해상화
네트워크
S^up
인코더
디코더

(A)

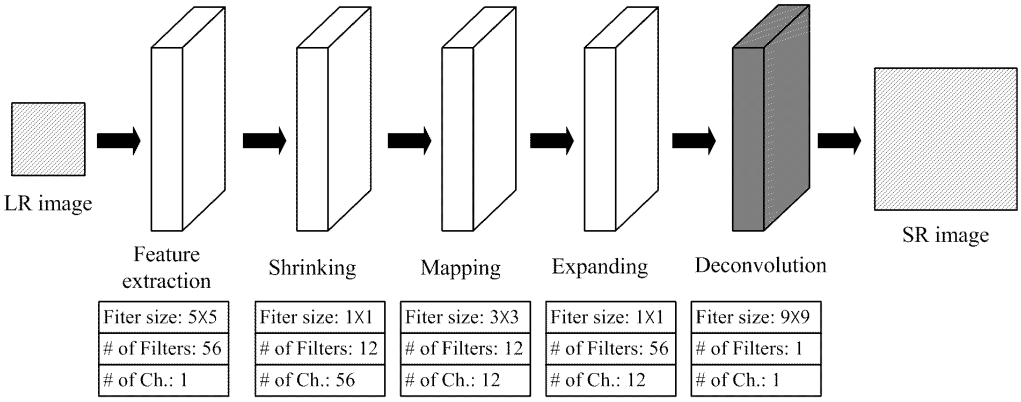

원본
영상
가로 방향
다운샘플링
210
220
230
200
S
S″
S″
비대칭
초해상화
네트워크
240
S^up
S^up
인코더
디코더

(B)

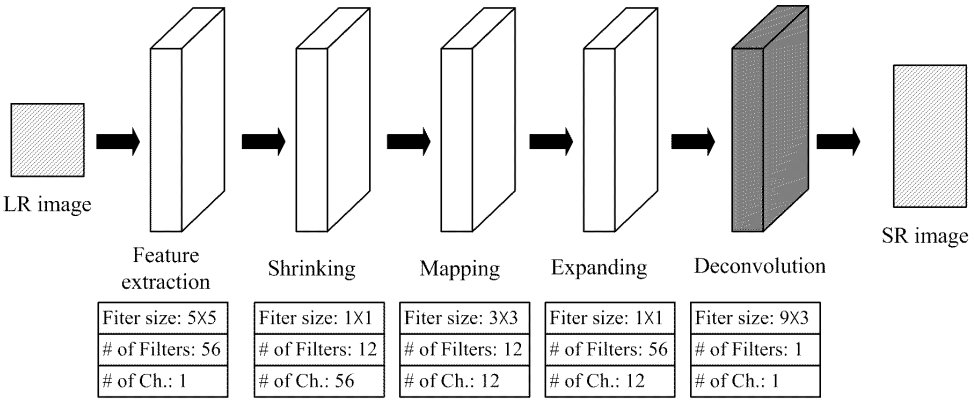
도면5


원본
영상
양 방향
다운샘플링
S
S‴
인코더
310
320
300
S‴
비대칭
초해상화
네트워크
S^up
S^up
비대칭
초해상화
네트워크
S^up
S′^up
340
330
S′^up
350
디코더

도면6

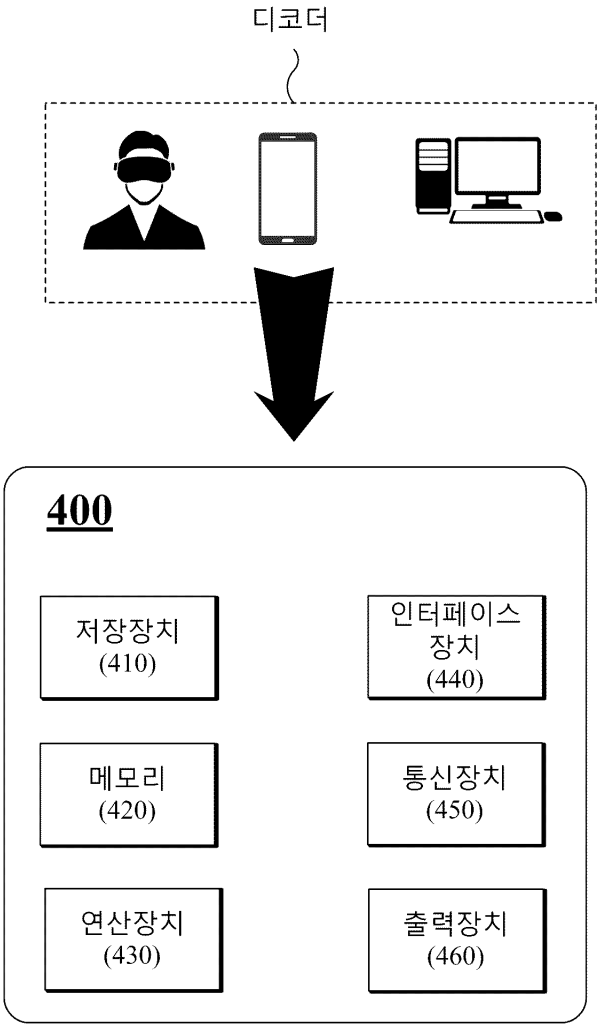

LR image
Feature extraction
Shrinking
Mapping
Expanding
Deconvolution
SR image
Fiter size: 5X5
# of Filters: 56
# of Ch.: 1
Fiter size: 1X1
# of Filters: 12
# of Ch.: 56
Fiter size: 3X3
# of Filters: 12
# of Ch.: 12
Fiter size: 1X1
# of Filters: 56
# of Ch.: 12
Fiter size: 9X9
# of Filters: 1
# of Ch.: 1

(A)


LR image
Feature extraction
Shrinking
Mapping
Expanding
Deconvolution
SR image
Fiter size: 5X5
# of Filters: 56
# of Ch.: 1
Fiter size: 1X1
# of Filters: 12
# of Ch.: 56
Fiter size: 3X3
# of Filters: 12
# of Ch.: 12
Fiter size: 1X1
# of Filters: 56
# of Ch.: 12
Fiter size: 9X3
# of Filters: 1
# of Ch.: 1

(B)

도면7



【심사관 직권보정사항】

【직권보정 1】

【보정항목】 청구범위

【보정세부항목】 청구항 1

【변경전】

영상처리장치가 다운샘플링된 영상을 입력받는 단계; 및

상기 영상처리장치가 상기 다운샘플링된 영상을 비대칭 초해상화 네트워크에 입력하여 업샘플링을 수행하는 단계를 포함하되,

상기 다운샘플링된 영상은 VVC(Versatile Video Coding)의 RPR (Reference Picture Resampling) 매커니즘에 따라 다운샘플링된 것이고,

영상 프레임은 평면에서 제1 방향과 제2 방향의 축을 갖고,

상기 비대칭 초해상화 네트워크는 상기 영상 프레임의 상기 제1 방향으로만 영상을 업샘플링하되,

상기 다운샘플링 영상이 상기 제1 방향으로만 다운샘플링된 영상인 경우, 상기 영상처리장치는 상기 다운샘플링된 영상을 비대칭 초해상화 네트워크에 입력하여 상기 제1 방향으로만 업샘플링하고,

상기 다운샘플링 영상이 상기 제2 방향으로만 다운샘플링된 영상인 경우, 상기 영상처리장치는 상기 다운샘플링된 영상을 90도 회전시킨 후 상기 비대칭 초해상화 네트워크에 입력하여 영상을 업샘플링하는 비대칭 초해상화 네트워크를 이용한 영상 업샘플링 방법.

【변경후】

영상처리장치가 다운샘플링된 영상을 입력받는 단계; 및

상기 영상처리장치가 상기 다운샘플링된 영상을 비대칭 초해상화 네트워크에 입력하여 업샘플링을 수행하는 단계를 포함하되,

상기 다운샘플링된 영상은 VVC(Versatile Video Coding)의 RPR (Reference Picture Resampling) 매커니즘에 따라 다운샘플링된 것이고,

영상 프레임은 평면에서 제1 방향과 제2 방향의 축을 갖고,

상기 비대칭 초해상화 네트워크는 상기 영상 프레임의 상기 제1 방향으로만 영상을 업샘플링하되,

상기 다운샘플링된 영상이 상기 제1 방향으로만 다운샘플링된 영상인 경우, 상기 영상처리장치는 상기 다운샘플링된 영상을 비대칭 초해상화 네트워크에 입력하여 상기 제1 방향으로만 업샘플링하고,

상기 다운샘플링된 영상이 상기 제2 방향으로만 다운샘플링된 영상인 경우, 상기 영상처리장치는 상기 다운샘플링된 영상을 90도 회전시킨 후 상기 비대칭 초해상화 네트워크에 입력하여 영상을 업샘플링하는 비대칭 초해상화 네트워크를 이용한 영상 업샘플링 방법.

【직권보정 2】

【보정항목】 청구범위

【보정세부항목】 청구항 4

【변경전】

제1항에 있어서,

상기 다운샘플링 영상이 상기 제1 방향 및 상기 제2 방향으로 다운샘플링된 경우,

상기 영상처리장치는 상기 다운샘플링된 영상을 상기 비대칭 초해상화 네트워크에 입력하여 영상을 업샘플링하고,

상기 업샘플링된 영상을 90도 회전시킨 후 상기 비대칭 초해상화 네트워크에 다시 입력하여 영상을 업샘플링하는 비대칭 초해상화 네트워크를 이용한 영상 업샘플링 방법.

【변경후】

제1항에 있어서,

상기 다운샘플링된 영상이 상기 제1 방향 및 상기 제2 방향으로 다운샘플링된 경우,

상기 영상처리장치는 상기 다운샘플링된 영상을 상기 비대칭 초해상화 네트워크에 입력하여 영상을 업샘플링하고,

상기 업샘플링된 영상을 90도 회전시킨 후 상기 비대칭 초해상화 네트워크에 다시 입력하여 영상을 업샘플링하는 비대칭 초해상화 네트워크를 이용한 영상 업샘플링 방법.

【직권보정 3】

【보정항목】 청구범위

【보정세부항목】 청구항 7

【변경전】

다운샘플링된 영상을 입력받는 입력장치;

영상 프레임이 제1 방향과 제2 방향의 축을 갖고, 상기 영상 프레임의 상기 제1 방향으로만 영상을 업샘플링하는 비대칭 초해상화 네트워크를 저장하는 저장장치; 및

상기 다운샘플링된 영상을 비대칭 초해상화 네트워크에 입력하여 업샘플링을 수행하는 연산장치를 포함하되,

상기 다운샘플링 영상이 상기 제1 방향으로만 다운샘플링된 영상인 경우, 상기 연산장치는 상기 다운샘플링된 영상을 비대칭 초해상화 네트워크에 입력하여 상기 제1 방향으로만 업샘플링하고,

상기 다운샘플링 영상이 상기 제2 방향으로만 다운샘플링된 영상인 경우, 상기 연산장치는 상기 다운샘플링된 영상을 90도 회전시킨 후 상기 비대칭 초해상화 네트워크에 입력하여 영상을 업샘플링하는 비대칭 초

해상화 네트워크를 이용하여 영상을 업샘플링하는 디코더.

【변경후】

다운샘플링된 영상을 입력받는 입력장치;

영상 프레임이 제1 방향과 제2 방향의 축을 갖고, 상기 영상 프레임의 상기 제1 방향으로만 영상을 업샘플링하는 비대칭 초해상화 네트워크를 저장하는 저장장치; 및

상기 다운샘플링된 영상을 비대칭 초해상화 네트워크에 입력하여 업샘플링을 수행하는 연산장치를 포함하되,

상기 다운샘플링된 영상이 상기 제1 방향으로만 다운샘플링된 영상인 경우, 상기 연산장치는 상기 다운샘플링된 영상을 비대칭 초해상화 네트워크에 입력하여 상기 제1 방향으로만 업샘플링하고,

상기 다운샘플링된 영상이 상기 제2 방향으로만 다운샘플링된 영상인 경우, 상기 연산장치는 상기 다운샘플링된 영상을 90도 회전시킨 후 상기 비대칭 초해상화 네트워크에 입력하여 영상을 업샘플링하는 비대칭 초해상화 네트워크를 이용하여 영상을 업샘플링하는 디코더.

【직권보정 4】

【보정항목】 청구범위

【보정세부항목】 청구항 10

【변경전】

제7항에 있어서,

상기 다운샘플링 영상이 상기 제1 방향 및 상기 제2 방향으로 다운샘플링된 경우,

상기 연산장치는 상기 다운샘플링된 영상을 상기 비대칭 초해상화 네트워크에 입력하여 영상을 업샘플링하고,

상기 업샘플링된 영상을 90도 회전시킨 후 상기 비대칭 초해상화 네트워크에 다시 입력하여 영상을 업샘플링하는

비대칭 초해상화 네트워크를 이용하여 영상을 업샘플링하는 디코더.

【변경후】

제7항에 있어서,

상기 다운샘플링된 영상이 상기 제1 방향 및 상기 제2 방향으로 다운샘플링된 경우,

상기 연산장치는 상기 다운샘플링된 영상을 상기 비대칭 초해상화 네트워크에 입력하여 영상을 업샘플링하고,

상기 업샘플링된 영상을 90도 회전시킨 후 상기 비대칭 초해상화 네트워크에 다시 입력하여 영상을 업샘플링하는

비대칭 초해상화 네트워크를 이용하여 영상을 업샘플링하는 디코더.