(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0018447
(43) 공개일자 2022년02월15일

(51) 국제특허분류(Int. Cl.)
G06T 9/00 (2019.01) *G06N 3/08* (2006.01)
G06T 3/40 (2006.01) *G06T 5/00* (2019.01)
H04N 19/513 (2014.01)
(52) CPC특허분류
G06T 9/002 (2013.01)
G06N 3/08 (2013.01)
(21) 출원번호 10-2021-0103569
(22) 출원일자 2021년08월06일
심사청구일자 없음
(30) 우선권주장
1020200098414 2020년08월06일 대한민국(KR)

(71) 출원인
현대자동차주식회사
서울특별시 서초구 헌릉로 12 (양재동)
기아 주식회사
서울특별시 서초구 헌릉로 12 (양재동)
이화여자대학교 산학협력단
서울특별시 서대문구 이화여대길 52 (대현동, 이화여자대학교)
(72) 발명자
강제원
서울특별시 마포구 서강로 53 서강쌍용예가아파트 104-904호
박승욱
경기도 용인시 수지구 태봉로 17, 403동 302호
임화평
경기도 화성시 남양읍 역골로 49번길 29 202호
(74) 대리인
이철희

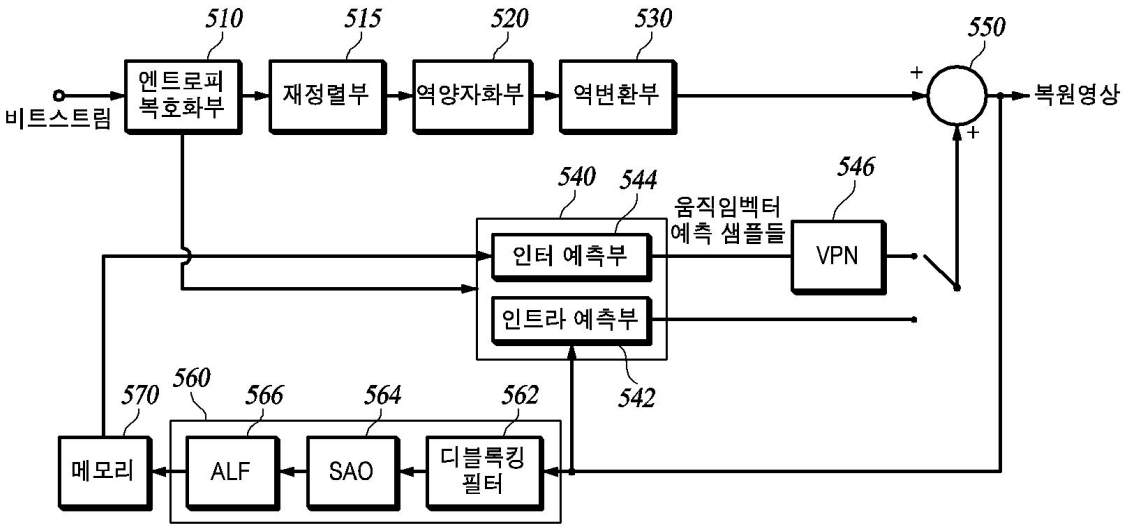
전체 청구항 수 : 총 16 항

(54) 발명의 명칭 **딥러닝 기반 인터 예측을 이용하는 영상 부호화 및 복호화**

(57) 요 약

본 실시예는, 현재블록에 대해 기존의 인터 예측을 수행하여 움직임벡터 및 예측 샘플들을 생성하고, 움직임벡터, 참조 샘플들, 예측 샘플들 등을 기반으로 딥러닝 기반 VPN(Video Prediction Network)을 이용하여 현재블록에 대한 개선된 예측 샘플들을 생성함으로써, 부호화 효율을 향상시키는 영상 부호화/복호화 방법을 제공한다.

대 표 도



(52) CPC특허분류
G06T 3/4046 (2013.01)
G06T 3/4053 (2013.01)
G06T 5/002 (2013.01)
H04N 19/513 (2015.01)

명 세 서

청구범위

청구항 1

비트스트림으로부터 현재블록에 대한 움직임벡터, 및 잔차 값들을 복호화하는 엔트로피 복호화부;

상기 움직임벡터에 의해 지시되는 참조픽처 내의 참조 샘플들을 이용하여 상기 현재블록에 대한 제1 예측 샘플들을 생성하는 인터 예측부;

상기 움직임벡터, 상기 참조 샘플들, 및 상기 제1 예측 샘플들의 전부 또는 일부를 이용하여 제2 예측 샘플들을 생성하는 VPN(Video Prediction Network); 및

상기 제2 예측 샘플들에 상기 잔차 값들을 가산하여 상기 현재블록에 대한 복원블록을 생성하는 가산기

를 포함하는 것을 특징으로 하는, 영상 복호화 장치.

청구항 2

제1항에 있어서,

상기 VPN은,

심층신경망 기반의 제1 영상복원 모델(image restoration model)을 포함하고, 상기 제1 영상복원 모델은 상기 제1 예측 샘플들을 입력으로 받아들여 상기 제2 예측 샘플들을 생성하되, 상기 제2 예측 샘플들이 상기 현재블록의 원본 영상에 근접하도록 트레이닝되는 것을 특징으로 하는, 영상 복호화 장치.

청구항 3

제1항에 있어서,

상기 VPN은,

심층신경망 기반의 제2 영상복원 모델을 포함하고, 상기 제2 영상복원 모델은 상기 제1 예측 샘플들, 과거 참조 샘플들, 미래 참조 샘플들, 및 상기 움직임벡터를 입력으로 받아들여 상기 제2 예측 샘플들을 생성하되, 상기 제2 예측 샘플들이 상기 현재블록의 원본 영상에 근접하도록 트레이닝되는 것을 특징으로 하는, 영상 복호화 장치.

청구항 4

제1항에 있어서,

상기 움직임벡터는,

AMVP(Advanced Motion Vector Prediction) 모드 또는 머지(merge) 모드를 적용하여 획득되되, 영상 부호화 장치로부터 수신한 플래그를 기반으로 상기 VPN의 적용 여부가 결정되는 것을 특징으로 하는, 영상 복호화 장치.

청구항 5

제1항에 있어서,

상기 현재블록의 인터 예측 모드가 AMVP 모드 또는 머지 모드이고, 상기 현재블록의 주변 블록이 상기 VPN을 이용하여 예측된 경우, 움직임벡터 후보 리스트에서 상기 주변 블록에 대응하는 움직임벡터의 우선 순위를 높게 설정되는 것을 특징으로 하는, 영상 복호화 장치.

청구항 6

제1항에 있어서,

머지 모드에 따른 인터 예측이 수행되는 경우, 상기 인터 예측부는, 움직임벡터 후보 리스트에서 주변 블록을 선택하고, 상기 선택된 주변 블록에 해당하는 움직임벡터가 지시하는 상기 참조픽처 내의 참조 샘플들을 이용하

여 상기 현재블록에 대한 제1 예측 샘플들을 생성하고, 상기 VPN은, 상기 움직임벡터, 상기 참조 샘플들, 및 상기 제1 예측 샘플들의 전부 또는 일부를 이용하여 가상 블록을 생성하고, 상기 가상블록을 상기 현재블록에 대한 제2 예측 샘플들로 이용하는 것을 특징으로 하는, 영상 복호화 장치.

청구항 7

제6항에 있어서,

상기 움직임벡터 후보 리스트에 포함된 기존의 주변 블록이 아닌, 상기 가상 블록이 선택되어 상기 제2 예측 샘플들로 이용되는 경우, 상기 가상 블록을 지시하기 위해, 상기 움직임벡터 후보 리스트에 추가된 신규 인덱스를 이용하되, 상기 가상 블록에 대한 움직임벡터는 0 벡터로 설정되는 것을 특징으로 하는, 영상 복호화 장치.

청구항 8

비트스트림으로부터 현재블록에 대한 움직임벡터, 및 잔차 값들을 복호화하는 엔트로피 복호화부;

상기 움직임벡터에 의해 지시되는 참조픽처 내의 참조 샘플들을 이용하여 상기 현재블록에 대한 제1 예측 샘플들을 생성하는 인터 예측부;

상기 움직임벡터, 또는 상기 현재블록의 주변 블록들의 움직임벡터들을 이용하여 복수의 제1 움직임벡터를 생성하고, 상기 복수의 제1 움직임벡터를 기반으로 상기 참조 샘플들 또는 상기 제1 예측 샘플들로부터 제2 예측 샘플들을 생성하는 VPN(Video Prediction Network); 및

상기 제2 예측 샘플들에 상기 잔차 값들을 가산하여 상기 현재블록에 대한 복원블록을 생성하는 가산기

를 포함하는 것을 특징으로 하는, 영상 복호화 장치.

청구항 9

제8항에 있어서,

상기 VPN은,

적어도 하나의 콘볼루션 레이어를 포함하는 SR(Super-resolution) 모델을 포함하고, 상기 복수의 제1 움직임벡터를 포함하는 2차원 배열을 상기 SR 모델에 입력하여, 복수의 제2 움직임벡터를 포함하는 2차원 배열을 생성하는, 영상 복호화 장치.

청구항 10

제9항에 있어서,

상기 복수의 제2 움직임벡터 각각은 상기 현재 블록의 픽셀 위치들 각각에 대한 옵티컬 플로우를 나타내고,

상기 VPN은,

상기 복수의 제2 움직임벡터 각각을 이용하여 상기 제1 예측 샘플들을 보정하기 위한 샘플 오프셋을 생성하고, 상기 제1 예측 샘플들과 상기 샘플 오프셋을 픽셀 단위로 가산하여 상기 제2 예측 샘플들을 생성하는 것을 특징으로 하는, 영상 복호화 장치.

청구항 11

제10항에 있어서,

상기 SR 모델은,

상기 복수의 제1 움직임벡터에 해당하는 학습용 데이터, 및 상기 옵티컬 플로우에 해당하는 레이블을 기반으로 상기 복수의 제1 움직임벡터를 포함하는 2차원 배열의 해상도를 증가시키도록 트레이닝되는 것을 특징으로 하는, 영상 복호화 장치.

청구항 12

제9항에 있어서,

상기 복수의 제2 움직임벡터 각각은 상기 현재 블록에 대한 참조 샘플과 제2 예측 샘플 간의 관계를 나타내는

픽셀 단위의 움직임벡터를 나타내고,

상기 VPN은,

상기 복수의 제2 움직임벡터 각각을 이용하여 상기 참조 샘플들로부터 상기 제2 예측 샘플들을 픽셀 단위로 생성하는 것을 특징으로 하는, 영상 복호화 장치.

청구항 13

제12항에 있어서,

상기 SR 모델은,

상기 복수의 제1 움직임벡터에 해당하는 학습용 데이터, 및 상기 픽셀 단위의 움직임벡터에 해당하는 레이블을 기반으로 상기 복수의 제1 움직임벡터를 포함하는 2차원 배열의 해상도를 증가시키도록 트레이닝되는 것을 특징으로 하는, 영상 복호화 장치.

청구항 14

영상 복호화 장치가 수행하는, 현재블록에 대한 인터 예측을 수행하기 위한 영상 복호화 방법에 있어서,

비트스트림으로부터 현재블록에 대한 움직임벡터, 및 잔차 값들을 복호화하는 단계;

상기 움직임벡터에 의해 지시되는 참조픽처 내의 참조 샘플들을 이용하여 상기 현재블록에 대한 제1 예측 샘플들을 생성하는 단계;

상기 움직임벡터, 상기 참조 샘플들, 및 상기 제1 예측 샘플들의 전부 또는 일부를 기반으로 딥러닝 기반 VPN(Video Prediction Network)을 이용하여 제2 예측 샘플들을 생성하는 단계; 및

상기 제2 예측 샘플들에 상기 잔차 값들을 가산하여 상기 현재블록에 대한 복원블록을 생성하는 단계

를 포함하는 것을 특징으로 하는, 영상 복호화 방법.

청구항 15

제14항에 있어서,

상기 VPN은 심층신경망 기반의 제1 영상복원 모델(image restoration model)을 포함하고,

상기 제2 예측 샘플들을 생성하는 단계는, 상기 제1 영상복원 모델을 이용하여 상기 제1 예측 샘플들로부터 상기 제2 예측 샘플들을 생성하되, 상기 제1 영상복원 모델은 상기 제2 예측 샘플들이 상기 현재블록의 원본 영상에 근접하도록 트레이닝되는 것을 특징으로 하는, 영상 복호화 방법.

청구항 16

제14항에 있어서,

상기 VPN은 심층신경망 기반의 제2 영상복원 모델을 포함하고,

상기 제2 예측 샘플들을 생성하는 단계는, 상기 제2 영상복원 모델을 이용하여 상기 제1 예측 샘플들, 과거 참조 샘플들, 미래 참조 샘플들, 및 움직임벡터로부터 상기 제2 예측 샘플들을 생성하되, 상기 제2 영상복원 모델은 상기 제2 예측 샘플들이 상기 현재블록의 원본 영상에 근접하도록 트레이닝되는 것을 특징으로 하는, 영상 복호화 방법.

발명의 설명

기 술 분 야

[0001] 본 개시는 영상(비디오)의 부호화 및 복호화에 관한 것이다. 더욱 상세하게는, 현재블록에 대해 기존의 인터 예측을 수행하여 움직임벡터 및 예측 샘플들을 생성하고, 움직임벡터, 참조 샘플들, 예측 샘플들 등을 기반으로 딥러닝 기반 VPN(Video Prediction Network)을 이용하여 현재블록에 대한 개선된 예측 샘플들을 생성함으로써,

부호화 효율을 향상시키는 영상 부호화/복호화 방법에 대한 것이다.

배 경 기 술

[0002] 이하에 기술되는 내용은 단순히 본 발명과 관련되는 배경 정보만을 제공할 뿐 종래기술을 구성하는 것이 아니다.

[0003] 비디오 데이터는 음성 데이터나 정지 영상 데이터 등에 비하여 많은 데이터량을 가지기 때문에, 압축을 위한 처리 없이 그 자체를 저장하거나 전송하기 위해서는 메모리를 포함하여 많은 하드웨어 자원을 필요로 한다.

[0004] 따라서, 통상적으로 비디오 데이터를 저장하거나 전송할 때에는 부호화기를 사용하여 비디오 데이터를 압축하여 저장하거나 전송하며, 복호화기에서는 압축된 비디오 데이터를 수신하여 압축을 해제하고 재생한다. 이러한 비디오 압축 기술로는 H.264/AVC, HEVC(High Efficiency Video Coding) 등을 비롯하여, HEVC에 비해 약 30% 이상의 부호화 효율을 향상시킨 VVC(Versatile Video Coding)가 존재한다.

[0005] 그러나, 영상의 크기 및 해상도, 프레임률이 점차 증가하고 있고, 이에 따라 부호화해야 하는 데이터량도 증가하고 있으므로 기존의 압축 기술보다 더 부호화 효율이 좋고 화질 개선 효과도 높은 새로운 압축 기술이 요구된다.

[0006] 최근, 딥러닝 기반 영상처리 기술이 기존의 부호화 요소 기술에 적용되고 있다. 기존 부호화 기술 중 인터 예측, 인트라 예측, 인루프 필터, 변환 등과 같은 압축 기술에 딥러닝 기반 영상처리 기술을 적용함으로써, 부호화 효율을 향상시킬 수 있다. 대표적인 응용 예로는, 딥러닝 모델 기반으로 생성된 가상 참조 프레임 기반 인터 예측, 잡음 제거 모델 기반의 인루프 필터 등이 있다. 따라서, 영상 부호화/복호화에 있어서, 부호화 효율을 향상시키기 위해, 딥러닝 기반 영상처리 기술의 지속적인 적용이 고려될 필요가 있다.

선행기술문헌

비특허문헌

[0007] (비특허문헌 0001) 비특허문헌 1: Chao Dong, Yubin Deng, Chen Change Loy, and Xiaoou Tang, Compression Artifacts Reduction by a Deep Convolutional Network, ICCV, 2015.

(비특허문헌 0002) 비특허문헌 2: Ren Yang, Mai Xu, Zulin Wang and Tianyi Li, Multi-Frame Quality Enhancement for Compressed Video, Arxiv:1803.04680.

발명의 내용

해결하려는 과제

[0008] 본 개시는, 현재블록에 대해 기존의 인터 예측을 수행하여 움직임벡터 및 예측 샘플들을 생성하고, 움직임벡터, 참조 샘플들, 예측 샘플들 등을 기반으로 딥러닝 기반 VPN(Video Prediction Network)을 이용하여 현재블록에 대한 개선된 예측 샘플들을 생성함으로써, 부호화 효율을 향상시키는 영상 부호화/복호화 방법을 제공하는 데 목적이 있다.

과제의 해결 수단

[0009] 본 개시의 실시예에 따르면, 비트스트림으로부터 현재블록에 대한 움직임벡터, 및 잔차 값들을 복호화하는 엔트로피 복호화부; 상기 움직임벡터에 의해 지시되는 참조픽처 내의 참조 샘플들을 이용하여 상기 현재블록에 대한 제1 예측 샘플들을 생성하는 인터 예측부; 상기 움직임벡터, 상기 참조 샘플들, 및 상기 제1 예측 샘플들의 전부 또는 일부를 이용하여 제2 예측 샘플들을 생성하는 VPN(Video Prediction Network); 및 상기 제2 예측 샘플들에 상기 잔차 값들을 가산하여 상기 현재블록에 대한 복원블록을 생성하는 가산기를 포함하는 것을 특징으로 하는, 영상 복호화 장치를 제공한다.

[0010] 본 개시의 다른 실시예에 따르면, 비트스트림으로부터 현재블록에 대한 움직임벡터, 및 잔차 값들을 복호화하는 엔트로피 복호화부; 상기 움직임벡터에 의해 지시되는 참조픽처 내의 참조 샘플들을 이용하여 상기 현재블록에 대한 제1 예측 샘플들을 생성하는 인터 예측부; 상기 움직임벡터, 또는 상기 현재블록의 주변 블록들의 움직임

벡터들을 이용하여 복수의 제1 움직임벡터를 생성하고, 상기 복수의 제1 움직임벡터를 기반으로 상기 참조 샘플들 또는 상기 제1 예측 샘플들로부터 제2 예측 샘플들을 생성하는 VPN(Video Prediction Network); 및 상기 제2 예측 샘플들에 상기 잔차 값들을 가산하여 상기 현재블록에 대한 복원블록을 생성하는 가산기를 포함하는 것을 특징으로 하는, 영상 복호화 장치를 제공한다.

[0011] 본 개시의 다른 실시예에 따르면, 영상 복호화 장치가 수행하는, 현재블록에 대한 인터 예측을 수행하기 위한 영상 복호화 방법에 있어서, 비트스트림으로부터 현재블록에 대한 움직임벡터, 및 잔차 값들을 복호화하는 단계; 상기 움직임벡터에 의해 지시되는 참조픽처 내의 참조 샘플들을 이용하여 상기 현재블록에 대한 제1 예측 샘플들을 생성하는 단계; 상기 움직임벡터, 상기 참조 샘플들, 및 상기 제1 예측 샘플들의 전부 또는 일부를 기반으로 딥러닝 기반 VPN(Video Prediction Network)을 이용하여 제2 예측 샘플들을 생성하는 단계; 및 상기 제2 예측 샘플들에 상기 잔차 값들을 가산하여 상기 현재블록에 대한 복원블록을 생성하는 단계를 포함하는 것을 특징으로 하는, 영상 복호화 방법을 제공한다.

발명의 효과

[0012] 이상에서 설명한 바와 같이 본 실시예에 따르면, 현재블록에 대해 기존의 인터 예측을 수행하여 움직임벡터 및 예측 샘플들을 생성하고, 움직임벡터, 참조 샘플들, 예측 샘플들 등을 기반으로 딥러닝 기반 VPN(Video Prediction Network)을 이용하여 현재블록에 대한 개선된 예측 샘플들을 생성하는 영상 부호화/복호화 방법을 제공함으로써, 인터 예측 성능의 개선에 따라 부호화 효율을 향상시키는 것이 가능해지는 효과가 있다.

도면의 간단한 설명

[0013] 도 1은 본 개시의 기술들을 구현할 수 있는 영상 부호화 장치에 대한 예시적인 블록도이다.

도 2는 QTBTTT 구조를 이용하여 블록을 분할하는 방법을 설명하기 위한 도면이다.

도 3a 및 도 3b는 광각 인트라 예측모드들을 포함한 복수의 인트라 예측모드들을 나타낸 도면이다.

도 4는 현재블록의 주변블록에 대한 예시도이다.

도 5는 본 개시의 기술들을 구현할 수 있는 영상 복호화 장치의 예시적인 블록도이다.

도 6은 본 개시의 일 실시예에 따른, 딥러닝 기반 인터 예측을 이용하는 영상 부호화 장치에 대한 예시적인 블록도이다.

도 7은 본 개시의 일 실시예에 따른, 딥러닝 기반 인터 예측을 이용하는 영상 복호화 장치에 대한 예시적인 블록도이다.

도 8은 본 개시의 일 실시예에 따른, ARCNN에 기반하는 VPN에 대한 개략적인 예시도이다.

도 9는 본 개시의 다른 실시예에 따른, RefAR에 기반하는 VPN에 대한 개략적인 예시도이다.

도 10은 본 개시의 다른 실시예에 따른, SRCNN에 기반하는 VPN에 대한 개략적인 예시도이다.

도 11은 본 개시의 일 실시예에 따른 머지 모드에서 VPN을 적용하는 방법에 대한 예시도이다.

도 12는 본 개시의 일 실시예에 따른, 딥러닝 기반 인터 예측을 이용하는 영상 부호화 방법의 흐름도이다.

도 13은 본 개시의 일 실시예에 따른, 딥러닝 기반 인터 예측을 이용하는 영상 복호화 방법의 흐름도이다.

도 14는 BDOF의 개념을 설명하기 위한 예시도이다.

발명을 실시하기 위한 구체적인 내용

[0014] 이하, 본 발명의 실시예들을 예시적인 도면을 참조하여 상세하게 설명한다. 각 도면의 구성요소들에 참조부호를 부가함에 있어서, 동일한 구성요소들에 대해서는 비록 다른 도면상에 표시되더라도 가능한 한 동일한 부호를 가지도록 하고 있음에 유의해야 한다. 또한, 본 실시예들을 설명함에 있어, 관련된 공지 구성 또는 기능에 대한 구체적인 설명이 본 실시예들의 요지를 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명은 생략한다.

[0015] 도 1은 본 개시의 기술들을 구현할 수 있는 영상 부호화 장치에 대한 예시적인 블록도이다. 이하에서는 도 1의 도시를 참조하여 영상 부호화 장치와 이 장치의 하위 구성들에 대하여 설명하도록 한다.

[0016] 영상 부호화 장치는 픽처 분할부(110), 예측부(120), 감산기(130), 변환부(140), 양자화부(145), 재정렬부(150), 엔트로피 부호화부(155), 역양자화부(160), 역변환부(165), 가산기(170), 루프 필터부(180) 및 메모리(190)를 포함하여 구성될 수 있다.

[0017] 영상 부호화 장치의 각 구성요소는 하드웨어 또는 소프트웨어로 구현되거나, 하드웨어 및 소프트웨어의 결합으로 구현될 수 있다. 또한, 각 구성요소의 기능이 소프트웨어로 구현되고 마이크로프로세서가 각 구성요소에 대응하는 소프트웨어의 기능을 실행하도록 구현될 수도 있다.

[0018] 하나의 영상(비디오)은 복수의 픽처들을 포함하는 하나 이상의 시퀀스로 구성된다. 각 픽처들은 복수의 영역으로 분할되고 각 영역마다 부호화가 수행된다. 예를 들어, 하나의 픽처는 하나 이상의 타일(Tile) 또는/및 슬라이스(Slice)로 분할된다. 여기서, 하나 이상의 타일을 타일 그룹(Tile Group)으로 정의할 수 있다. 각 타일 또는/슬라이스는 하나 이상의 CTU(Coding Tree Unit)로 분할된다. 그리고 각 CTU는 트리 구조에 의해 하나 이상의 CU(Coding Unit)들로 분할된다. 각 CU에 적용되는 정보들은 CU의 신택스로서 부호화되고, 하나의 CTU에 포함된 CU들에 공통적으로 적용되는 정보는 CTU의 신택스로서 부호화된다. 또한, 하나의 슬라이스 내의 모든 블록들에 공통적으로 적용되는 정보는 슬라이스 헤더의 신택스로서 부호화되며, 하나 이상의 픽처들을 구성하는 모든 블록들에 적용되는 정보는 픽처 파라미터 셋(PPS, Picture Parameter Set) 혹은 픽처 헤더에 부호화된다. 나아가, 복수의 픽처가 공통으로 참조하는 정보들은 시퀀스 파라미터 셋(SPS, Sequence Parameter Set)에 부호화된다. 그리고, 하나 이상의 SPS가 공통으로 참조하는 정보들은 비디오 파라미터 셋(VPS, Video Parameter Set)에 부호화된다. 또한, 하나의 타일 또는 타일 그룹에 공통으로 적용되는 정보는 타일 또는 타일 그룹 헤더의 신택스로서 부호화될 수도 있다. SPS, PPS, 슬라이스 헤더, 타일 또는 타일 그룹 헤더에 포함되는 신택스들은 상위수준(high level) 신택스로 칭할 수 있다.

[0019] 픽처 분할부(110)는 CTU(Coding Tree Unit)의 크기를 결정한다. CTU의 크기에 대한 정보(CTU size)는 SPS 또는 PPS의 신택스로서 부호화되어 영상 복호화 장치로 전달된다.

[0020] 픽처 분할부(110)는 영상을 구성하는 각 픽처(picture)를 미리 결정된 크기를 가지는 복수의 CTU(Coding Tree Unit)들로 분할한 이후에, 트리 구조(tree structure)를 이용하여 CTU를 반복적으로(recursively) 분할한다. 트리 구조에서의 리프 노드(leaf node)가 부호화의 기본 단위인 CU(coding unit)가 된다.

[0021] 트리 구조로는 상위 노드(혹은 부모 노드)가 동일한 크기의 네 개의 하위 노드(혹은 자식 노드)로 분할되는 쿼드트리(QuadTree, QT), 또는 상위 노드가 두 개의 하위 노드로 분할되는 바이너리트리(BinaryTree, BT), 또는 상위 노드가 1:2:1 비율로 세 개의 하위 노드로 분할되는 터너리트리(TernaryTree, TT), 또는 이러한 QT 구조, BT 구조 및 TT 구조 중 둘 이상을 혼용한 구조일 수 있다. 예컨대, QTBT(QuadTree plus BinaryTree) 구조가 사용될 수 있고, 또는 QTBTTT(QuadTree plus BinaryTree TernaryTree) 구조가 사용될 수 있다. 여기서, BTTT를 합쳐서 MTT(Multiple-Type Tree)라 지칭될 수 있다.

[0022] 도 2는 QTBTTT 구조를 이용하여 블록을 분할하는 방법을 설명하기 위한 도면이다.

[0023] 도 2에 도시된 바와 같이, CTU는 먼저 QT 구조로 분할될 수 있다. 쿼드트리 분할은 분할 블록(splitting block)의 크기가 QT에서 허용되는 리프 노드의 최소 블록 크기(MinQTSize)에 도달할 때까지 반복될 수 있다. QT 구조의 각 노드가 하위 레이어의 4개의 노드들로 분할되는지 여부를 지시하는 제1 플래그(QT_split_flag)는 엔트로피 부호화부(155)에 의해 부호화되어 영상 복호화 장치로 시그널링된다. QT의 리프 노드가 BT에서 허용되는 루트 노드의 최대 블록 크기(MaxBTSize)보다 크지 않은 경우, BT 구조 또는 TT 구조 중 어느 하나 이상으로 더 분할될 수 있다. BT 구조 및/또는 TT 구조에서는 복수의 분할 방향이 존재할 수 있다. 예컨대, 해당 노드의 블록이 가로로 분할되는 방향과 세로로 분할되는 방향 두 가지가 존재할 수 있다. 도 2의 도시와 같이, MTT 분할이 시작되면, 노드들이 분할되었는지 여부를 지시하는 제2 플래그(mtt_split_flag)와, 분할이 되었다면 추가적으로 분할 방향(vertical 혹은 horizontal)을 나타내는 플래그 및/또는 분할 타입(Binary 혹은 Ternary)을 나타내는 플래그가 엔트로피 부호화부(155)에 의해 부호화되어 영상 복호화 장치로 시그널링된다.

[0024] 대안적으로, 각 노드가 하위 레이어의 4개의 노드들로 분할되는지 여부를 지시하는 제1 플래그(QT_split_flag)를 부호화하기에 앞서, 그 노드가 분할되는지 여부를 지시하는 CU 분할 플래그(split_cu_flag)가 부호화될 수도 있다. CU 분할 플래그(split_cu_flag) 값이 분할되지 않았음을 지시하는 경우, 해당 노드의 블록이 분할 트리 구조에서의 리프 노드(leaf node)가 되어 부호화의 기본 단위인 CU(coding unit)가 된다. CU 분할 플래그(split_cu_flag) 값이 분할됨을 지시하는 경우, 영상 부호화 장치는 전술한 방식으로 제1 플래그부터 부호화를 시작한다.

[0025] 트리 구조의 다른 예시로서 QTBT가 사용되는 경우, 해당 노드의 블록을 동일 크기의 두 개 블록으로 가로로 분할하는 타입(즉, symmetric horizontal splitting)과 세로로 분할하는 타입(즉, symmetric vertical splitting) 두 가지가 존재할 수 있다. BT 구조의 각 노드가 하위 레이어의 블록으로 분할되는지 여부를 지시하는 분할 플래그(split_flag) 및 분할되는 타입을 지시하는 분할 타입 정보가 엔트로피 부호화부(155)에 의해 부호화되어 영상 복호화 장치로 전달된다. 한편, 해당 노드의 블록을 서로 비대칭 형태의 두 개의 블록으로 분할하는 타입이 추가로 더 존재할 수도 있다. 비대칭 형태에는 해당 노드의 블록을 1:3의 크기 비율을 가지는 두 개의 직사각형 블록으로 분할하는 형태가 포함될 수 있고, 혹은 해당 노드의 블록을 대각선 방향으로 분할하는 형태가 포함될 수도 있다.

[0026] CU는 CTU로부터의 QTBT 또는 QTBTTT 분할에 따라 다양한 크기를 가질 수 있다. 이하에서는, 부호화 또는 복호화하고자 하는 CU(즉, QTBTTT의 리프 노드)에 해당하는 블록을 '현재블록'이라 칭한다. QTBTTT 분할의 채용에 따라, 현재블록의 모양은 정사각형뿐만 아니라 직사각형일 수도 있다.

[0027] 예측부(120)는 현재블록을 예측하여 예측블록을 생성한다. 예측부(120)는 인트라 예측부(122)와 인터 예측부(124)를 포함한다.

[0028] 일반적으로, 픽처 내 현재블록들은 각각 예측적으로 코딩될 수 있다. 일반적으로 현재블록의 예측은 (현재블록을 포함하는 픽처로부터의 데이터를 사용하는) 인트라 예측 기술 또는 (현재블록을 포함하는 픽처 이전에 코딩된 픽처로부터의 데이터를 사용하는) 인터 예측 기술을 사용하여 수행될 수 있다. 인터 예측은 단방향 예측과 양방향 예측 모두를 포함한다.

[0029] 인트라 예측부(122)는 현재블록이 포함된 현재 픽처 내에서 현재블록의 주변에 위치한 픽셀(참조 픽셀)들을 이용하여 현재블록 내의 픽셀들을 예측한다. 예측 방향에 따라 복수의 인트라 예측모드가 존재한다. 예컨대, 도 3a에서 보는 바와 같이, 복수의 인트라 예측모드는 planar 모드와 DC 모드를 포함하는 2개의 비방향성 모드와 65개의 방향성 모드를 포함할 수 있다. 각 예측모드에 따라 사용할 주변 픽셀과 연산식이 다르게 정의된다.

[0030] 직사각형 모양의 현재블록에 대한 효율적인 방향성 예측을 위해, 도 3b에 점선 화살표로 도시된 방향성 모드들(67 ~ 80번, -1 ~ -14 번 인트라 예측모드들)이 추가로 사용될 수 있다. 이들은 "광각 인트라 예측모드들(wide angle intra-prediction modes)"로 지칭될 수 있다. 도 3b에서 화살표들은 예측에 사용되는 대응하는 참조샘플들을 가리키는 것이며, 예측 방향을 나타내는 것이 아니다. 예측 방향은 화살표가 가리키는 방향과 반대이다. 광각 인트라 예측모드들은 현재블록이 직사각형일 때 추가적인 비트 전송 없이 특정 방향성 모드를 반대방향으로 예측을 수행하는 모드이다. 이때 광각 인트라 예측모드들 중에서, 직사각형의 현재블록의 너비와 높이의 비율에 의해, 현재블록에 이용 가능한 일부 광각 인트라 예측모드들이 결정될 수 있다. 예컨대, 45도보다 작은 각도를 갖는 광각 인트라 예측모드들(67 ~ 80번 인트라 예측모드들)은 현재블록이 높이가 너비보다 작은 직사각형 형태일 때 이용 가능하고, -135도보다 큰 각도를 갖는 광각 인트라 예측모드들(-1 ~ -14 번 인트라 예측모드들)은 현재블록이 너비가 높이보다 큰 직사각형 형태일 때 이용 가능하다.

[0031] 인트라 예측부(122)는 현재블록을 부호화하는데 사용할 인트라 예측모드를 결정할 수 있다. 일부 예들에서, 인트라 예측부(122)는 여러 인트라 예측모드들을 사용하여 현재블록을 인코딩하고, 테스트된 모드들로부터 사용할 적절한 인트라 예측모드를 선택할 수도 있다. 예를 들어, 인트라 예측부(122)는 여러 테스트된 인트라 예측모드들에 대한 비트율 왜곡(rate-distortion) 분석을 사용하여 비트율 왜곡 값들을 계산하고, 테스트된 모드들 중 최선의 비트율 왜곡 특징들을 갖는 인트라 예측모드를 선택할 수도 있다.

[0032] 인트라 예측부(122)는 복수의 인트라 예측모드 중에서 하나의 인트라 예측모드를 선택하고, 선택된 인트라 예측모드에 따라 결정되는 주변 픽셀(참조 픽셀)과 연산식을 사용하여 현재블록을 예측한다. 선택된 인트라 예측모드에 대한 정보는 엔트로피 부호화부(155)에 의해 부호화되어 영상 복호화 장치로 전달된다.

[0033] 인터 예측부(124)는 움직임 보상 과정을 이용하여 현재블록에 대한 예측블록을 생성한다. 인터 예측부(124)는 현재 픽처보다 먼저 부호화 및 복호화된 참조픽처 내에서 현재블록과 가장 유사한 블록을 탐색하고, 그 탐색된 블록을 이용하여 현재블록에 대한 예측블록을 생성한다. 그리고, 현재 픽처 내의 현재블록과 참조픽처 내의 예측블록 간의 변위(displacement)에 해당하는 움직임벡터(Motion Vector: MV)를 생성한다. 일반적으로, 움직임 추정은 루마(luma) 성분에 대해 수행되고, 루마 성분에 기초하여 계산된 움직임벡터는 루마 성분 및 크로마 성분 모두에 대해 사용된다. 현재블록을 예측하기 위해 사용된 참조픽처에 대한 정보 및 움직임벡터에 대한 정보를 포함하는 움직임 정보는 엔트로피 부호화부(155)에 의해 부호화되어 영상 복호화 장치로 전달된다.

[0034] 인터 예측부(124)는, 예측의 정확성을 높이기 위해, 참조 픽처 또는 참조 블록에 대한 보간을 수행할 수도

있다. 즉, 연속한 두 정수 샘플 사이의 서브 샘플들은 그 두 정수 샘플을 포함한 연속된 복수의 정수 샘플들에 필터 계수들을 적용하여 보간된다. 보간된 참조 픽처에 대해서 현재블록과 가장 유사한 블록을 탐색하는 과정을 수행하면, 움직임벡터는 정수 샘플 단위의 정밀도(precision)가 아닌 소수 단위의 정밀도까지 표현될 수 있다. 움직임벡터의 정밀도 또는 해상도(resolution)는 부호화하고자 하는 대상 영역, 예컨대, 슬라이스, 타일, CTU, CU 등의 단위마다 다르게 설정될 수 있다. 이와 같은 적응적 움직임벡터 해상도(Adaptive Motion Vector Resolution: AMVR)가 적용되는 경우 각 대상 영역에 적용할 움직임벡터 해상도에 대한 정보는 대상 영역마다 시그널링되어야 한다. 예컨대, 대상 영역이 CU인 경우, 각 CU마다 적용된 움직임벡터 해상도에 대한 정보가 시그널링된다. 움직임벡터 해상도에 대한 정보는 후술할 차분 움직임벡터의 정밀도를 나타내는 정보일 수 있다.

[0035] 한편, 인터 예측부(124)는 양방향 예측(bi-prediction)을 이용하여 인터 예측을 수행할 수 있다. 양방향 예측의 경우, 두 개의 참조 픽처와 각 참조 픽처 내에서 현재블록과 가장 유사한 블록 위치를 나타내는 두 개의 움직임벡터가 이용된다. 인터 예측부(124)는 참조픽처 리스트 0(RefPicList0) 및 참조픽처 리스트 1(RefPicList1)로부터 각각 제1 참조픽처 및 제2 참조픽처를 선택하고, 각 참조픽처 내에서 현재블록과 유사한 블록을 탐색하여 제1 참조블록과 제2 참조블록을 생성한다. 그리고, 제1 참조블록과 제2 참조블록을 평균 또는 가중 평균하여 현재블록에 대한 예측블록을 생성한다. 그리고 현재블록을 예측하기 위해 사용한 두 개의 참조픽처에 대한 정보 및 두 개의 움직임벡터에 대한 정보를 포함하는 움직임 정보를 부호화부(150)로 전달한다. 여기서, 참조픽처 리스트 0은 기복원된 픽처들 중 디스플레이 순서에서 현재 픽처 이전의 픽처들로 구성되고, 참조픽처 리스트 1은 기복원된 픽처들 중 디스플레이 순서에서 현재 픽처 이후의 픽처들로 구성될 수 있다. 그러나 반드시 이에 한정되는 것은 아니며, 디스플레이 순서 상으로 현재 픽처 이후의 기복원 픽처들이 참조픽처 리스트 0에 추가로 더 포함될 수 있고, 역으로 현재 픽처 이전의 기복원 픽처들이 참조픽처 리스트 1에 추가로 더 포함될 수도 있다.

[0036] 움직임 정보를 부호화하는 데에 소요되는 비트량을 최소화하기 위해 다양한 방법이 사용될 수 있다.

[0037] 예컨대, 현재블록의 참조픽처와 움직임벡터가 주변블록의 참조픽처 및 움직임벡터와 동일한 경우에는 그 주변블록을 식별할 수 있는 정보를 부호화함으로써, 현재블록의 움직임 정보를 영상 복호화 장치로 전달할 수 있다. 이러한 방법을 '머지 모드(merge mode)'라 한다.

[0038] 머지 모드에서, 인터 예측부(124)는 현재블록의 주변블록들로부터 기 결정된 개수의 머지 후보블록(이하, '머지 후보'라 함)들을 선택한다.

[0039] 머지 후보를 유도하기 위한 주변블록으로는, 도 4에 도시된 바와 같이, 현재 픽처 내에서 현재블록에 인접한 좌측블록(L), 상단블록(A), 우상단블록(AR), 좌하단블록(BL), 좌상단블록(AL) 중에서 전부 또는 일부가 사용될 수 있다. 또한, 현재블록이 위치한 현재 픽처가 아닌 참조픽처(현재블록을 예측하기 위해 사용된 참조픽처와 동일할 수도 있고 다를 수도 있음) 내에 위치한 블록이 머지 후보로서 사용될 수도 있다. 예컨대, 참조픽처 내에서 현재블록과 동일 위치에 있는 블록(co-located block) 또는 그 동일 위치의 블록에 인접한 블록들이 머지 후보로서 추가로 더 사용될 수 있다. 이상에서 기술된 방법에 의해 선정된 머지 후보의 개수가 기설정된 개수보다 작으면, 0 벡터를 머지 후보에 추가한다.

[0040] 인터 예측부(124)는 이러한 주변블록들을 이용하여 기 결정된 개수의 머지 후보를 포함하는 머지 리스트를 구성한다. 머지 리스트에 포함된 머지 후보들 중에서 현재블록의 움직임정보로서 사용할 머지 후보를 선택하고 선택된 후보를 식별하기 위한 머지 인덱스 정보를 생성한다. 생성된 머지 인덱스 정보는 부호화부(150)에 의해 부호화되어 영상 복호화 장치로 전달된다.

[0041] 머지 스킵(merge skip) 모드는 머지 모드의 특별한 경우로서, 양자화를 수행한 후, 엔트로피 부호화를 위한 변환 계수가 모두 영(zero)에 가까울 때, 잔차신호의 전송 없이 주변블록 선택 정보만을 전송한다. 머지 스킵 모드를 이용함으로써, 움직임이 적은 영상, 정지 영상, 스크린 콘텐츠 영상 등에서 상대적으로 높은 부호화 효율을 달성할 수 있다.

[0042] 이하, 머지 모드와 머지 스킵 모드를 통칭하여, 머지/스킵 모드로 나타낸다.

[0043] 움직임 정보를 부호화하기 위한 또 다른 방법은 AMVP(Advanced Motion Vector Prediction) 모드이다.

[0044] AMVP 모드에서, 인터 예측부(124)는 현재블록의 주변블록들을 이용하여 현재블록의 움직임벡터에 대한 예측 움직임벡터 후보들을 유도한다. 예측 움직임벡터 후보들을 유도하기 위해 사용되는 주변블록으로는, 도 4에 도시된 현재 픽처 내에서 현재블록에 인접한 좌측블록(L), 상단블록(A), 우상단블록(AR), 좌하단블록(BL), 좌상단블록(AL) 중에서 전부 또는 일부가 사용될 수 있다. 또한, 현재블록이 위치한 현재 픽처가 아닌 참조픽처(현재블록을 예측하기 위해 사용된 참조픽처와 동일할 수도 있고 다를 수도 있음) 내에 위치한 블록이 예측 움직임벡터

후보들을 유도하기 위해 사용되는 주변블록으로서 사용될 수도 있다. 예컨대, 참조픽처 내에서 현재블록과 동일 위치에 있는 블록(collocated block) 또는 그 동일 위치의 블록에 인접한 블록들이 사용될 수 있다. 이상에서 기술된 방법에 의해 움직임벡터 후보의 개수가 기설정된 개수보다 작으면, 0 벡터를 움직임벡터 후보에 추가한다.

[0045] 인터 예측부(124)는 이 주변블록들의 움직임벡터를 이용하여 예측 움직임벡터 후보들을 유도하고, 예측 움직임벡터 후보들을 이용하여 현재블록의 움직임벡터에 대한 예측 움직임벡터를 결정한다. 그리고, 현재블록의 움직임벡터로부터 예측 움직임벡터를 감산하여 차분 움직임벡터를 산출한다.

[0046] 예측 움직임벡터는 예측 움직임벡터 후보들에 기 정의된 함수(예컨대, 중앙값, 평균값 연산 등)를 적용하여 구할 수 있다. 이 경우, 영상 복호화 장치도 기 정의된 함수를 알고 있다. 또한, 예측 움직임벡터 후보를 유도하기 위해 사용하는 주변블록은 이미 부호화 및 복호화가 완료된 블록이므로 영상 복호화 장치도 그 주변블록의 움직임벡터도 이미 알고 있다. 그러므로 영상 부호화 장치는 예측 움직임벡터 후보를 식별하기 위한 정보를 부호화할 필요가 없다. 따라서, 이 경우에는 차분 움직임벡터에 대한 정보와 현재블록을 예측하기 위해 사용한 참조픽처에 대한 정보가 부호화된다.

[0047] 한편, 예측 움직임벡터는 예측 움직임벡터 후보들 중 어느 하나를 선택하는 방식으로 결정될 수도 있다. 이 경우에는 차분 움직임벡터에 대한 정보 및 현재블록을 예측하기 위해 사용한 참조픽처에 대한 정보와 함께, 선택된 예측 움직임벡터 후보를 식별하기 위한 정보가 추가로 부호화된다.

[0048] 감산기(130)는 현재블록으로부터 인트라 예측부(122) 또는 인터 예측부(124)에 의해 생성된 예측블록을 감산하여 잔차블록을 생성한다.

[0049] 변환부(140)는 공간 영역의 픽셀 값들을 가지는 잔차블록 내의 잔차신호를 주파수 도메인의 변환 계수로 변환한다. 변환부(140)는 잔차블록의 전체 크기를 변환 단위로 사용하여 잔차블록 내의 잔차신호들을 변환할 수 있으며, 또는 잔차블록을 복수 개의 서브블록으로 분할하고 그 서브블록을 변환 단위로 사용하여 변환을 할 수도 있다. 또는, 변환 영역 및 비변환 영역인 두 개의 서브블록으로 구분하여, 변환 영역 서브블록만 변환 단위로 사용하여 잔차신호들을 변환할 수 있다. 여기서, 변환 영역 서브블록은 가로축 (혹은 세로축) 기준 1:1의 크기 비율을 가지는 두 개의 직사각형 블록 중 하나일 수 있다. 이런 경우, 서브블록 만을 변환하였음을 지시하는 플래그(cu_sbt_flag), 방향성(vertical/horizontal) 정보(cu_sbt_horizontal_flag) 및/또는 위치 정보(cu_sbt_pos_flag)가 엔트로피 부호화부(155)에 의해 부호화되어 영상 복호화 장치로 시그널링된다. 또한, 변환 영역 서브블록의 크기는 가로축 (혹은 세로축) 기준 1:3의 크기 비율을 가질 수 있으며, 이런 경우 해당 분할을 구분하는 플래그(cu_sbt_quad_flag)가 추가적으로 엔트로피 부호화부(155)에 의해 부호화되어 영상 복호화 장치로 시그널링된다.

[0050] 한편, 변환부(140)는 잔차블록에 대해 가로 방향과 세로 방향으로 개별적으로 변환을 수행할 수 있다. 변환을 위해, 다양한 타입의 변환 함수 또는 변환 행렬이 사용될 수 있다. 예컨대, 가로 방향 변환과 세로 방향 변환을 위한 변환 함수의 쌍을 MTS(Multiple Transform Set)로 정의할 수 있다. 변환부(140)는 MTS 중 변환 효율이 가장 좋은 하나의 변환 함수 쌍을 선택하고 가로 및 세로 방향으로 각각 잔차블록을 변환할 수 있다. MTS 중에서 선택된 변환 함수 쌍에 대한 정보(mts_idx)는 엔트로피 부호화부(155)에 의해 부호화되어 영상 복호화 장치로 시그널링된다.

[0051] 양자화부(145)는 변환부(140)로부터 출력되는 변환 계수들을 양자화 파라미터를 이용하여 양자화하고, 양자화된 변환 계수들을 엔트로피 부호화부(155)로 출력한다. 양자화부(145)는, 어떤 블록 혹은 프레임에 대해, 변환 없이, 관련된 잔차 블록을 곧바로 양자화할 수도 있다. 양자화부(145)는 변환블록 내의 변환 계수들의 위치에 따라 서로 다른 양자화 계수(스케일링 값)을 적용할 수도 있다. 2차원으로 배열된 양자화된 변환 계수들에 적용되는 양자화 행렬은 부호화되어 영상 복호화 장치로 시그널링될 수 있다.

[0052] 재정렬부(150)는 양자화된 잔차값에 대해 계수값의 재정렬을 수행할 수 있다.

[0053] 재정렬부(150)는 계수 스캐닝(coefficient scanning)을 이용하여 2차원의 계수 어레이를 1차원의 계수 시퀀스로 변경할 수 있다. 예를 들어, 재정렬부(150)에서는 지그-재그 스캔(zig-zag scan) 또는 대각선 스캔(diagonal scan)을 이용하여 DC 계수부터 고주파수 영역의 계수까지 스캔하여 1차원의 계수 시퀀스를 출력할 수 있다. 변환 단위의 크기 및 인트라 예측모드에 따라 지그-재그 스캔 대신 2차원의 계수 어레이를 열 방향으로 스캔하는 수직 스캔, 2차원의 블록 형태 계수를 행 방향으로 스캔하는 수평 스캔이 사용될 수도 있다. 즉, 변환 단위의 크기 및 인트라 예측모드에 따라 지그-재그 스캔, 대각선 스캔, 수직 방향 스캔 및 수평 방향 스캔 중에서 사용

될 스캔 방법이 결정될 수도 있다.

[0054] 엔트로피 부호화부(155)는, CABAC(Context-based Adaptive Binary Arithmetic Code), 지수 골롬(Exponential Golomb) 등의 다양한 부호화 방식을 사용하여, 재정렬부(150)로부터 출력된 1차원의 양자화된 변환 계수들의 시퀀스를 부호화함으로써 비트스트림을 생성한다.

[0055] 또한, 엔트로피 부호화부(155)는 블록 분할과 관련된 CTU size, CU 분할 플래그, QT 분할 플래그, MTT 분할 타입, MTT 분할 방향 등의 정보를 부호화하여, 영상 복호화 장치가 영상 부호화 장치와 동일하게 블록을 분할할 수 있도록 한다. 또한, 엔트로피 부호화부(155)는 현재블록이 인트라 예측에 의해 부호화되었는지 아니면 인터 예측에 의해 부호화되었는지 여부를 지시하는 예측 타입에 대한 정보를 부호화하고, 예측 타입에 따라 인트라 예측정보(즉, 인트라 예측모드에 대한 정보) 또는 인터 예측정보(머지 모드의 경우 머지 인덱스, AMVP 모드의 경우 참조픽처 인덱스 및 차분 움직임벡터에 대한 정보)를 부호화한다. 또한, 엔트로피 부호화부(155)는 양자화와 관련된 정보, 즉, 양자화 파라미터에 대한 정보 및 양자화 행렬에 대한 정보를 부호화한다.

[0056] 역양자화부(160)는 양자화부(145)로부터 출력되는 양자화된 변환 계수들을 역양자화하여 변환 계수들을 생성한다. 역변환부(165)는 역양자화부(160)로부터 출력되는 변환 계수들을 주파수 도메인으로부터 공간 도메인으로 변환하여 잔차블록을 복원한다.

[0057] 가산부(170)는 복원된 잔차블록과 예측부(120)에 의해 생성된 예측블록을 가산하여 현재블록을 복원한다. 복원된 현재블록 내의 픽셀들은 다음 순서의 블록을 인트라 예측할 때 참조 픽셀로서 사용된다.

[0058] 루프(loop) 필터부(180)는 블록 기반의 예측 및 변환/양자화로 인해 발생하는 블록킹 아티팩트(blocking artifacts), 링잉 아티팩트(ringing artifacts), 블러링 아티팩트(blurring artifacts) 등을 줄이기 위해 복원된 픽셀들에 대한 필터링을 수행한다. 필터부(180)는 인루프(in-loop) 필터로서 디블록킹 필터(182), SAO(Sample Adaptive Offset) 필터(184) 및 ALF(Adaptive Loop Filter, 186)의 전부 또는 일부를 포함할 수 있다.

[0059] 디블록킹 필터(182)는 블록 단위의 부호화/복호화로 인해 발생하는 블록킹 현상(blocking artifact)을 제거하기 위해 복원된 블록 간의 경계를 필터링하고, SAO 필터(184) 및 alf(186)는 디블록킹 필터링된 영상에 대해 추가적인 필터링을 수행한다. SAO 필터(184) 및 alf(186)는 손실 부호화(lossy coding)로 인해 발생하는 복원된 픽셀과 원본 픽셀 간의 차이를 보상하기 위해 사용되는 필터이다. SAO 필터(184)는 CTU 단위로 오프셋을 적용함으로써 주관적 화질뿐만 아니라 부호화 효율도 향상시킨다. 이에 비하여 ALF(186)는 블록 단위의 필터링을 수행하는데, 해당 블록의 에지 및 변화량의 정도를 구분하여 상이한 필터를 적용하여 왜곡을 보상한다. ALF에 사용될 필터 계수들에 대한 정보는 부호화되어 영상 복호화 장치로 시그널링될 수 있다.

[0060] 디블록킹 필터(182), SAO 필터(184) 및 ALF(186)를 통해 필터링된 복원블록은 메모리(190)에 저장된다. 한 픽처 내의 모든 블록들이 복원되면, 복원된 픽처는 이후에 부호화하고자 하는 픽처 내의 블록을 인터 예측하기 위한 참조 픽처로 사용될 수 있다.

[0061] 도 5는 본 개시의 기술들을 구현할 수 있는 영상 복호화 장치의 예시적인 기능 블록도이다. 이하에서는 도 5를 참조하여 영상 복호화 장치와 이 장치의 하위 구성들에 대하여 설명하도록 한다.

[0062] 영상 복호화 장치는 엔트로피 복호화부(510), 재정렬부(515), 역양자화부(520), 역변환부(530), 예측부(540), 가산기(550), 루프 필터부(560) 및 메모리(570)를 포함하여 구성될 수 있다.

[0063] 도 1의 영상 부호화 장치와 마찬가지로, 영상 복호화 장치의 각 구성요소는 하드웨어 또는 소프트웨어로 구현되거나, 하드웨어 및 소프트웨어의 결합으로 구현될 수 있다. 또한, 각 구성요소의 기능이 소프트웨어로 구현되고 마이크로프로세서가 각 구성요소에 대응하는 소프트웨어의 기능을 실행하도록 구현될 수도 있다.

[0064] 엔트로피 복호화부(510)는 영상 부호화 장치에 의해 생성된 비트스트림을 복호화하여 블록 분할과 관련된 정보를 추출함으로써 복호화하고자 하는 현재블록을 결정하고, 현재블록을 복원하기 위해 필요한 예측정보와 잔차신호에 대한 정보 등을 추출한다.

[0065] 엔트로피 복호화부(510)는 SPS(Sequence Parameter Set) 또는 PPS(Picture Parameter Set)로부터 CTU size에 대한 정보를 추출하여 CTU의 크기를 결정하고, 픽처를 결정된 크기의 CTU로 분할한다. 그리고, CTU를 트리 구조의 최상위 레이어, 즉, 루트 노드로 결정하고, CTU에 대한 분할정보를 추출함으로써 트리 구조를 이용하여 CTU를 분할한다.

[0066] 예컨대, QTBTTT 구조를 사용하여 CTU를 분할하는 경우, 먼저 QT의 분할과 관련된 제1 플래그(QT_split_flag)를 추출하여 각 노드를 하위 레이어의 네 개의 노드로 분할한다. 그리고, QT의 리프 노드에 해당하는 노드에 대해서는 MTT의 분할과 관련된 제2 플래그(MTT_split_flag) 및 분할 방향(vertical / horizontal) 및/또는 분할 타입(binary / ternary) 정보를 추출하여 해당 리프 노드를 MTT 구조로 분할한다. 이에 따라 QT의 리프 노드 이하의 각 노드들을 BT 또는 TT 구조로 반복적으로(recursively) 분할한다.

[0067] 또 다른 예로서, QTBTTT 구조를 사용하여 CTU를 분할하는 경우, 먼저 CU의 분할 여부를 지시하는 CU 분할 플래그(split_cu_flag)를 추출하고, 해당 블록이 분할된 경우, 제1 플래그(QT_split_flag)를 추출할 수도 있다. 분할 과정에서 각 노드는 0번 이상의 반복적인 QT 분할 후에 0번 이상의 반복적인 MTT 분할이 발생할 수 있다. 예컨대, CTU는 바로 MTT 분할이 발생하거나, 반대로 다수 번의 QT 분할만 발생할 수도 있다.

[0068] 다른 예로서, QTBT 구조를 사용하여 CTU를 분할하는 경우, QT의 분할과 관련된 제1 플래그(QT_split_flag)를 추출하여 각 노드를 하위 레이어의 네 개의 노드로 분할한다. 그리고, QT의 리프 노드에 해당하는 노드에 대해서는 BT로 더 분할되는지 여부를 지시하는 분할 플래그(split_flag) 및 분할 방향 정보를 추출한다.

[0069] 한편, 엔트로피 복호화부(510)는 트리 구조의 분할을 이용하여 복호화하고자 하는 현재블록을 결정하게 되면, 현재블록이 인트라 예측되었는지 아니면 인터 예측되었는지를 지시하는 예측 타입에 대한 정보를 추출한다. 예측 타입 정보가 인트라 예측을 지시하는 경우, 엔트로피 복호화부(510)는 현재블록의 인트라 예측정보(인트라 예측모드)에 대한 신택스 요소를 추출한다. 예측 타입 정보가 인터 예측을 지시하는 경우, 엔트로피 복호화부(510)는 인터 예측정보에 대한 신택스 요소, 즉, 움직임벡터 및 그 움직임벡터가 참조하는 참조픽처를 나타내는 정보를 추출한다.

[0070] 또한, 엔트로피 복호화부(510)는 양자화 관련된 정보, 및 잔차신호에 대한 정보로서 현재블록의 양자화된 변환계수들에 대한 정보를 추출한다.

[0071] 재정렬부(515)는, 영상 부호화 장치에 의해 수행된 계수 스캐닝 순서의 역순으로, 엔트로피 복호화부(510)에서 엔트로피 복호화된 1차원의 양자화된 변환계수들의 시퀀스를 다시 2차원의 계수 어레이(즉, 블록)로 변경할 수 있다.

[0072] 역양자화부(520)는 양자화된 변환계수들을 역양자화하고, 양자화 파라미터를 이용하여 양자화된 변환계수들을 역양자화한다. 역양자화부(520)는 2차원으로 배열된 양자화된 변환계수들에 대해 서로 다른 양자화 계수(스케일링 값)을 적용할 수도 있다. 역양자화부(520)는 영상 부호화 장치로부터 양자화 계수(스케일링 값)들의 행렬을 양자화된 변환계수들의 2차원 어레이에 적용하여 역양자화를 수행할 수 있다.

[0073] 역변환부(530)는 역양자화된 변환계수들을 주파수 도메인으로부터 공간 도메인으로 역변환하여 잔차신호들을 복원함으로써 현재블록에 대한 잔차블록을 생성한다.

[0074] 또한, 역변환부(530)는 변환블록의 일부 영역(서브블록)만 역변환하는 경우, 변환블록의 서브블록만을 변환하였음을 지시하는 플래그(cu_sbt_flag), 서브블록의 방향성(vertical/horizontal) 정보(cu_sbt_horizontal_flag) 및/또는 서브블록의 위치 정보(cu_sbt_pos_flag)를 추출하여, 해당 서브블록의 변환계수들을 주파수 도메인으로부터 공간 도메인으로 역변환함으로써 잔차신호들을 복원하고, 역변환되지 않은 영역에 대해서는 잔차신호로 “0”값을 채움으로써 현재블록에 대한 최종 잔차블록을 생성한다.

[0075] 또한, MTS가 적용된 경우, 역변환부(530)는 영상 부호화 장치로부터 시그널링된 MTS 정보(mts_idx)를 이용하여 가로 및 세로 방향으로 각각 적용할 변환 함수 또는 변환 행렬을 결정하고, 결정된 변환 함수를 이용하여 가로 및 세로 방향으로 변환블록 내의 변환계수들에 대해 역변환을 수행한다.

[0076] 예측부(540)는 인트라 예측부(542) 및 인터 예측부(544)를 포함할 수 있다. 인트라 예측부(542)는 현재블록의 예측 타입이 인트라 예측일 때 활성화되고, 인터 예측부(544)는 현재블록의 예측 타입이 인터 예측일 때 활성화된다.

[0077] 인트라 예측부(542)는 엔트로피 복호화부(510)로부터 추출된 인트라 예측모드에 대한 신택스 요소로부터 복수의 인트라 예측모드 중 현재블록의 인트라 예측모드를 결정하고, 인트라 예측모드에 따라 현재블록 주변의 참조 픽셀들을 이용하여 현재블록을 예측한다.

[0078] 인터 예측부(544)는 엔트로피 복호화부(510)로부터 추출된 인터 예측모드에 대한 신택스 요소를 이용하여 현재블록의 움직임벡터와 그 움직임벡터가 참조하는 참조픽처를 결정하고, 움직임벡터와 참조픽처를 이용하여 현재

블록을 예측한다.

[0079] 가산기(550)는 역변환부로부터 출력되는 잔차블록과 인터 예측부 또는 인트라 예측부로부터 출력되는 예측블록을 가산하여 현재블록을 복원한다. 복원된 현재블록 내의 픽셀들은 이후에 복호화할 블록을 인트라 예측할 때의 참조픽셀로서 활용된다.

[0080] 루프 필터부(560)는 인루프 필터로서 디블록킹 필터(562), SAO 필터(564) 및 ALF(566)를 포함할 수 있다. 디블록킹 필터(562)는 블록 단위의 복호화로 인해 발생하는 블록킹 현상(blocking artifact)을 제거하기 위해, 복원된 블록 간의 경계를 디블록킹 필터링한다. SAO 필터(564) 및 ALF(566)는 손실 부호화(lossy coding)으로 인해 발생하는 복원된 픽셀과 원본 픽셀 간의 차이를 보상하기 위해, 디블록킹 필터링 이후의 복원된 블록에 대해 추가적인 필터링을 수행한다. ALF의 필터 계수는 비스트림으로부터 복호한 필터 계수에 대한 정보를 이용하여 결정된다.

[0081] 디블록킹 필터(562), SAO 필터(564) 및 ALF(566)를 통해 필터링된 복원블록은 메모리(570)에 저장된다. 한 픽처 내의 모든 블록들이 복원되면, 복원된 픽처는 이후에 부호화하고자 하는 픽처 내의 블록을 인터 예측하기 위한 참조 픽처로 사용된다.

[0082] 본 실시예는 이상에서 설명한 바와 같은 영상(비디오)의 부호화 및 복호화에 관한 것이다. 더욱 상세하게는, 현재블록에 대해 기존의 인터 예측을 수행하여 움직임벡터 및 예측 샘플들을 생성하고, 움직임벡터, 참조 샘플들, 예측 샘플들 등을 기반으로 딥러닝 기반 VPN(Video Prediction Network)을 이용하여 현재블록에 대한 개선된 예측 샘플들을 생성함으로써, 부호화 효율을 향상시키는 영상 부호화/복호화 방법을 제공한다.

[0083] 이하의 설명에서, 영상 부호화 장치 및 방법은 부호화 장치 및 방법과 병행하여 사용되고, 영상 복호화 장치 및 방법은 복호화 장치 및 방법과 병행하여 사용된다.

[0084] '대상블록(target block)'이라는 용어는 전술한 바와 같은 현재블록 또는 코딩유닛(CU)과 동일한 의미로 사용될 수 있고, 또는 코딩유닛의 일부 영역을 의미할 수도 있다.

[0085] 양방향 옵티컬 플로우(Bidirectional OF: BDOF)는, 영상을 구성하는 샘플 또는 객체가 일정한 속도로 이동하고 샘플 값의 변화가 거의 없다는 가정을 기반으로, 양방향 움직임 예측을 이용하여 예측된 샘플들의 움직임을 옵티컬 플로우(Optical Flow: OF)를 기반으로 추가로 보상하는 기술이다.

[0086] 대상블록에 대한 (통상의) 양방향 움직임 예측에 의하여, 도 14에 예시된 바와 같이, 참조픽처들(Ref_0과 Ref_1)에서 현재 픽처의 부호화되는 대상블록과 가장 유사한 대응 영역들(즉, 참조 블록들)을 가리키는 양방향의 움직임벡터들(MV_0, MV_1)이 결정되었다고 가정하자. 이때, 두 개의 움직임벡터들은 대상블록 전체의 움직임을 나타내는 값이다. 도 14의 예시에서, 대상블록 내 샘플 P와 대응되는, 움직임벡터(MV_0)가 가리키는 참조픽처(Ref_0)내 샘플은 P_0이며, 대상블록 내 샘플 P와 대응되는, 움직임벡터(MV_1)가 가리키는 참조픽처(Ref_1)내 샘플은 P_1이다.

[0087] 이하, 도 14의 예시에서, 샘플 P에 대한 움직임은 대상블록의 전체적인 움직임과는 조금 상이하다라고 가정하자. 예컨대, Ref_0 내 샘플 A에 위치하는 물체가 현재 픽처의 대상블록 내 샘플 P를 지나 Ref1 내 샘플 B로 움직였다면, 샘플 A 및 샘플 B는 상당히 유사한 값을 가져야 한다. 따라서, 대상블록 내 샘플 P와 가장 유사한 Ref_0 내 지점은 움직임벡터(MV_0)가 가리키는 P0이 아니라, P0를 소정의 변위 벡터(v_xt_0,v_yt_0)만큼 이동시킨 A이다. 또한, 대상블록 내 샘플 P와 가장 유사한 Ref_1 내 지점은 움직임벡터(MV_1)가 가리키는 P1이 아니라, P1을 소정의 변위 벡터($-v_xt_1$,$-v_yt_1$)만큼 이동시킨 B이다. 여기서, t_0와 t_1은 현재 픽처를 기준으로 각각 Ref_0와 Ref_1에 대한 시간축 거리를 의미하며, POC(Picture Order Count)를 기반으로 계산된다. 이때, (v_x,v_y)를 OF라고 지칭한다. 따라서, OF는 픽처와 픽처 간에 발생하는 픽셀의 움직임을 표현하는, 픽셀 단위의 움직임벡터를 나타낸다.

[0088] BDOF에서는, 현재 픽처 내 대상블록의 샘플(P)값을 예측함에 있어서, 두 참조 샘플(A,B)의 값을 이용하여, 양방향 움직임벡터(MV_0,MV_1)가 가리키는 참조 샘플(P0,P1), 즉 예측 샘플을 이용하는 것보다 좀 더 정확한 예측을 수행한다.

[0089] 대상블록 내 (루마) 픽셀 (x,y)와 대응되는, 움직임벡터(MV_0)가 가리키는 참조픽처(Ref_0)내 예측샘플의 값을 $I^{(0)}$(x,y)라고 정의하고, 대상블록 내 샘플 (x,y)와 대응되는, 움직임벡터(MV_1)가 가리키는 참조픽처(Ref_1)내 예

측샘플의 값을 $I^{(1)}(x,y)$라고 정의한다.

[0090] BDOF에서는, 픽셀에 대한 예측샘플들의 값, OF 및 예측샘플에서의 그래디언트(gradient)를 이용하여, 수학식 1에 나타낸 바와 같이, 대상블록의 픽셀값이 예측될 수 있다.

수학식 1

[0091] $$pred(x,y) = \left(I^{(0)}(x,y) + I^{(1)}(x,y) + b(x,y) + o_{offset}\right) \gg shift$$

[0092] 여기서, 픽셀에 대한 예측샘플들의 값은, 인터 예측의 양방향 예측에 사용된 움직임 정보(두 개의 움직임벡터 및 두 개의 참조픽처)를 기반으로 예측된다. b(x,y)는 OF 및 예측샘플에서의 그래디언트를 이용하여 산정된 샘플 오프셋이다. shift는 스케일링 인자이고, o_{offset}은 반올림 연산을 위한 값으로서 shift의 절반이다.

[0093] 한편, 인터 예측은 병진 운동(translation motion) 모델을 반영하는 움직임 예측이다. 즉, 수평방향(x축 방향) 및 수직방향(y축 방향)으로의 움직임을 예측하는 기법이다. 그러나, 실제로는 병진 운동 이외에 회전(rotation), 줌-인(zoom-in) 또는 줌-아웃(zoom-out)과 같은 다양한 형태의 움직임이 존재할 수 있다. 아핀 움직임 예측(affine motion prediction)에서는, 이러한 다양한 형태의 움직임을 반영할 수 있는 움직임벡터를 생성한 후, 이 움직임벡터를 이용하여 대상블록의 서브블록 단위로 샘플들의 값이 예측된다.

[0094] PROF(Prediction refinement with OF)는, 서브블록 단위 아핀 움직임 예측에 따른 예측샘플들의 샘플값을, 옵티컬 플로우에 기반하여 샘플 단위로 조정하는 것을 나타낸다. 즉, 아핀 움직임 예측이 수행된 단위인 서브블록 내에서, 각 샘플의 위치에 따른 움직임이 추가로 보상된다.

[0095] 도 6은 본 개시의 일 실시예에 따른, 딥러닝 기반 인터 예측을 이용하는 영상 부호화 장치의 일부에 대한 예시적인 블록도이다.

[0096] 본 실시예에 따른 영상 부호화 장치는, 기존 구성요소인 인터 예측부(124)의 후단에 딥러닝 기반 VPN(126)을 추가적으로 포함한다. 여기서, 본 실시예에 따른 영상 부호화 장치에 포함되는 구성요소가 반드시 이에 한정되는 것은 아니다. 예컨대, 영상 부호화 장치는 VPN(126)에 포함된 심층신경망(이하, '딥러닝 모델'과 호환하여 사용)의 트레이닝을 위한 트레이닝부(미도시)를 추가로 구비하거나, 외부의 트레이닝부와 연동되는 형태로 구현될 수 있다.

[0097] 인터 예측부(124)는, 부호화하고자 하는 현재블록에 대해, 참조픽처를 이용하여 현재블록에 대한 참조블록을 탐색하고, 탐색된 참조블록의 참조 샘플들로부터 현재블록에 대한 예측 샘플들을 포함하는 예측블록을 생성한다. 또한, 인터 예측부(124)는, 참조블록과 현재블록 간의 변위에 해당하는 움직임벡터를 생성한다. 한편, 양방향 예측인 경우, 전술한 바와 같이, 참조블록과 움직임벡터는 각각 2 개일 수 있다.

[0098] 참조 샘플들, 움직임벡터 및 예측 샘플들의 전부 또는 일부를 이용하여 VPN(126)은 현재블록에 대한 개선된 예측 샘플들(이하, '개선 예측 샘플들')을 생성한다. 여기서, 개선 예측 샘플들은 최종 예측블록에 포함된다. VPN(126)은, 개선 예측 샘플들의 생성 과정에서 심층신경망을 이용할 수 있다.

[0099] 부호화 장치는 현재블록으로부터 최종 예측블록을 감산하여 잔차블록을 생성하고, 잔차블록의 잔차 값들을 변환부(140)에 입력하여, 전술한 바와 같은 부호화 과정을 수행할 수 있다.

[0100] 한편, 사전에 트레이닝된 심층신경망의 파라미터는 부호화 장치와 복호화 장치 간에 공유될 수 있다.

[0101] 도 7은 본 개시의 일 실시예에 따른, 딥러닝 기반 인터 예측을 이용하는 영상 복호화 장치에 대한 예시적인 블록도이다.

[0102] 본 실시예에 따른 영상 복호화 장치는, 기존 구성요소인 인터 예측부(544)의 후단에 딥러닝 기반 VPN(546)을 추가적으로 포함한다.

[0103] 엔트로피 복호화부(510)는 비트스트림으로부터 복호화하고자 하는 현재블록에 대한 움직임벡터, 및 잔차 값들을 포함하는 잔차블록을 복호한다. 이때, 움직임벡터는, AMVP 모드의 경우, 비트스트림으로부터 명시적으로 복호화될 수 있으나, 머지/스킵 모드의 경우, 주변의 움직임벡터 정보 또는 움직임 예측모드로부터 추정될 수 있다. 한편, 양방향 예측인 경우, 전술한 바와 같이, 움직임벡터는 2 개일 수 있다.

[0104] 인터 예측부(544)는 움직임벡터를 이용하여 참조 샘플들로부터 현재블록에 대한 예측 샘플들을 포함하는 예측블록을 생성한다. 여기서, 참조 샘플들은 움직임벡터가 지시하는 참조픽처로부터 생성될 수 있다.

[0105] 예측 샘플들은 움직임벡터 및 움직임 예측 모드에 따라 적응적으로 생성될 수 있다. 예를 들어, 움직임벡터의 크기가 영인 경우, 참조픽처 저장 리스트로부터 활용하는 샘플들과 동일하게 예측 샘플들이 생성될 수 있다.

[0106] 참조 샘플들, 움직임벡터 및 예측 샘플들의 전부 또는 일부를 이용하여 VPN(546)은 개선 예측 샘플들을 생성한다. 여기서, 개선 예측 샘플들은 최종 예측블록에 포함된다. 복호화 장치는 개선 예측 샘플들에 잔차 값들을 가산하여 현재블록에 대한 복원블록을 생성한다.

[0107] 이하, 도 8 내지 도 10을 이용하여 부호화 장치 내 VPN(546)의 동작에 대해 설명한다. 이하의 설명은 부호화 장치 내 VPN(126)에도 동일하게 적용될 수 있다.

[0108] 도 8은 본 개시의 일 실시예에 따른, ARCNN에 기반하는 VPN에 대한 개략적인 예시도이다.

[0109] 본 개시의 일 예에 있어서, VPN(546)은 영상잡음(artifact) 제거에 적합한 딥러닝 모델인 ARCNN(Artifacts Reduction Convolutional Neural Network, 비특허문헌 1 참조)을 포함할 수 있다. ARCNN은 영상복원(image restoration) 모델의 하나로서, 잡음(예컨대, 양자화 잡음(quantization noise))이 섞인 입력이 주어진 경우, 원본 영상에 근접한 출력 영상을 복원할 수 있도록 학습된다.

[0110] ARCNN은, 도 8에 예시된 바와 같이, 4 개의 블록을 포함하고, 각 블록은 콘볼루션 레이어(convolution layer)로 구성된다. 잡음이 포함된 입력 영상에 대하여, ARCNN의 4 개의 블록 각각은, 특성 추출(feature extraction), 특성 개선(enhancement), 특성 매핑(mapping), 및 영상 복원(restoration)을 차례대로 수행한다.

[0111] 본 실시예에 있어서, ARCNN은 예측 샘플들 X_{pred}를 입력으로 받아들여, 개선 예측 샘플들 $X_{f,pred}$를 생성한다.

[0112] ARCNN은 (예측 샘플들에 해당하는) 학습용 데이터, 및 원본 영상에 해당하는 레이블 X_t를 기반으로 트레이닝될 수 있다. 본 실시예에 따른 트레이닝부는, ARCNN이 학습용 데이터로부터 생성한 개선 예측 샘플들 $X_{f,pred}$와 레이블 X_t 간의 거리(distance)에 기반하여, 수학식 2에 나타낸 바와 같이, 손실 함수(loss function) L을 정의한다.

수학식 2

[0113] $$L = \left|X_{f,pred} - X_t\right|^2$$

[0114] 트레이닝부는 손실 함수를 감소시키는 방향으로 ARCNN의 파라미터를 업데이트함으로써, ARCNN에 대한 트레이닝을 진행한다. 한편, 수학식 2에서, 거리로서 L2 메트릭(metric)이 이용되었으나, L1 메트릭과 같이, 두 비교 대상 간의 거리 차이를 표현할 수 있는 어느 것이든 이용될 수 있다.

[0115] 도 9는 본 개시의 다른 실시예에 따른, RefAR에 기반하는 VPN에 대한 개략적인 예시도이다.

[0116] 본 개시의 다른 실시예에 있어서, VPN(546)은 예측 샘플들 X_{pred} 외에 과거 참조 샘플들 및 미래 참조 샘플들을 이용하는 딥러닝 모델 RefAR(Reference-based Artifacts Reduction, 비특허문헌 2 참조)을 포함할 수 있다. 여기서 참조 샘플들은, 인터 예측에서 이용되는 참조픽처에 있어서, 움직임벡터에 의해 참조되는 블록의 샘플들을 나타낸다. 양방향 예측인 경우, 과거 참조 샘플들과 미래 참조 샘플들이 모두 이용되나, 단방향 예측인 경우, 과거 참조 샘플들과 미래 참조 샘플들 중 하나가 이용될 수 있다.

[0117] 한편, RefAR은 영상복원 모델의 하나로서, 잡음이 섞인 입력이 주어진 경우, 원본 영상에 근접한 출력 영상을 복원하도록 학습된다.

[0118] RefAR은, 도 9에 예시된 바와 같이, 콘볼루션 레이어, 와프(warp), 및 융합 레이어(fusion layer)를 포함하는 세 단계로 구현될 수 있다. 콘볼루션 레이어는 예측 샘플들, 과거 참조 샘플들, 및 미래 참조 샘플들 각각의 특성을 추출한다. 과거/미래 참조 참조 샘플들과 예측 샘플들 간은 상호 시간이 일치하지 않으므로, 와프 단계는 OF(또는 움직임벡터)를 과거/미래 참조 샘플들에 적용하여, 과거/미래 참조 샘플들과 예측 샘플들 간의 시간을 일치시킨다. 융합 레이어는 예측 샘플들, 및 시간이 조정된 과거/미래 참조 샘플들을 융합하여 개선 예측 샘플

들을 생성한다.

[0119] 본 실시예에 있어서, RefAR은 예측 샘플들 X_{pred}, 과거 참조 샘플들 X_{ref0}, 및 미래 참조 샘플들 X_{ref1}를 입력으로 받아들여, 개선 예측 샘플들 $X_{f,pred}$를 생성한다.

[0120] RefAR은 (예측 샘플들, 미래 참조 샘플들, 및 과거 참조 샘플들에 해당하는) 학습용 데이터, 및 원본 영상에 해당하는 레이블 X_t를 기반으로 트레이닝될 수 있다. 본 실시예에 따른 트레이닝부는, RefAR이 학습용 데이터로부터 생성한 개선 예측 샘플들 $X_{f,pred}$와 레이블 X_t 간의 거리(distance)에 기반하여, 수학식 2에 나타낸 바와 같이, 손실 함수(loss function) L을 정의한다. 트레이닝부는 손실 함수를 감소시키는 방향으로 RefAR의 파라미터를 업데이트함으로써, RefAR에 대한 트레이닝을 진행한다.

[0121] 이상, 도 8 및 도 9에 예시된 방식은, VPN(546)에 포함된 딥러닝 모델을 기반으로 현재블록의 예측 샘플들에 대한 개선 예측 샘플들을 생성한다. 본 개시의 다른 실시예에 있어서, 복호화 장치는, 하나의 픽처 또는 슬라이스에 대한 예측 샘플들을 모두 생성한 후, 딥러닝 모델을 기반으로 하나의 픽처 또는 슬라이스에 대한 개선 예측 샘플들을 생성할 수 있다.

[0122] 도 8 및 도 9에 예시된 방식은, VPN(546)에 포함된 딥러닝 모델을 기반으로 예측 샘플들에 대한 개선 예측 샘플들을 바로 생성한다. 본 개시에 따른 다른 실시예에 있어서, 움직임벡터의 정밀도를 향상시킨 후, 정밀도가 향상된 움직임벡터를 이용하여 예측 샘플들로부터 개선 예측 샘플들을 생성할 수 있는데, 이때, 딥러닝 모델이, 움직임벡터의 정밀도를 향상시킨 개선 움직임벡터를 생성하기 위해 이용될 수 있다.

[0123] 도 10은 본 개시의 다른 실시예에 따른, SRCNN에 기반하는 VPN에 대한 개략적인 예시도이다.

[0124] 먼저, 참조 샘플들로부터 예측 샘플들을 생성하기 위한 움직임벡터를 m_t로 나타낸다. 움직임벡터는 블록 기반으로 취득하므로, 양방향 예측인 경우, 현재블록에 대해 두 개가 결정된다. 한편, 단방향 예측인 경우, 하나의 움직임벡터가 결정될 수 있다.

[0125] 따라서, 복수의 블록에 대한 m_t는, 도 10에 예시된 바와 같이, 2차원 배열을 형성할 수 있다. 이때, 이러한 이차원 배열에서, 각 움직임벡터 m_t는 (x, y)와 같이 두 개의 좌표값으로 표현될 수 있다. 현재블록에 대한 복수의 블록들로는, 예컨대, 머지 예측 모드인 경우, 후보 블록들(AR, A, AL, L 및 BL)이 이용될 수 있다. 즉, 후보 블록들의 움직임벡터들을 결합하여, 움직임벡터의 2차원 배열이 생성될 수 있다.

[0126] 본 개시에 따른 다른 실시예에 있어서, 현재블록을 서브샘플링(sub-sampling)한 픽셀 수만큼 현재블록의 움직임벡터 m_t를 반복하여 배열함으로써, 움직임벡터의 2차원 배열이 생성될 수 있다. 이때, 움직임벡터 m_t는 (x, y)와 같이 두 개의 좌표값으로 표현될 수 있다.

[0127] 다음, 예측 샘플들로부터 개선 예측 샘플들을 생성할 수 있는, 픽셀 단위의 개선 움직임벡터를 M_t로 나타낸다. 예컨대, M_t는 OF과 같이 픽셀 단위의 움직임벡터일 수 있다. 현재블록에 대해, 개선 움직임벡터 M_t도 (x, y) 두 개의 좌표값을 이용하여, 도 10에 예시된 바와 같이, 2차원 배열을 형성할 수 있다.

[0128] 이때, 개선 움직임벡터 M_t의 2차원 배열에 대한 해상도가 움직임벡터 m_t의 2차원 배열에 대한 해상도보다 크므로, 본 실시예에서, VPN(546)은, 개선 움직임벡터 M_t를 생성하기 위해 SR(Super-resolution) 모델인 SRCNN을 포함할 수 있다.

[0129] SRCNN으로는, 적어도 하나의 콘볼루션 레이어를 포함하여, 저해상도 영상을 입력으로 받아들여 고해상도 영상을 출력으로 생성하는, 어느 딥러닝 모델이든 이용될 수 있다.

[0130] 본 실시예에 있어서, SRCNN은 움직임벡터 m_t에 대한 2차원 배열을 입력으로 받아들여, 개선 움직임벡터 M_t에 대한 2차원 배열을 생성한다.

[0131] 이때, 각 m_t는 참조 샘플들 X_{ref0} 및 X_{ref1}으로부터 예측 샘플들 X_{pred}를 생성하기 위해 사용된다. 한편, 각 개선 움직임벡터 M_t는, 예컨대, 픽셀에 대한 OF인 경우, 전술한 바와 같은 BDOF에서와 같이, 예측 샘플들 X_{pred}를 보정하기 위한 샘플 오프셋을 생성하기 위해 이용될 수 있다. VPN(546)은, 수학식 1에 나타낸 바와 같이, 예측 샘플들과 샘플 오프셋을 픽셀 단위로 가산하여 개선 예측 샘플들 $X_{f,pred}$를 생성할 수 있다.

[0132] SRCNN은 움직임벡터에 해당하는 학습용 데이터, 및 픽셀 단위의 OF에 해당하는 레이블 M_{OF}를 기반으로 트레이닝 될 수 있다. 본 실시예에 따른 트레이닝부는, SRCNN이 학습용 데이터로부터 생성한 개선 움직임벡터 M_t(즉, M_t에 대한 2차원 배열)와 레이블 M_{OF}(즉, M_{OF}에 대한 2차원 배열) 간의 거리(distance)에 기반하여, 수학식 3에 나타낸 바와 같이, 손실 함수(loss function) L을 정의한다.

수학식 3

[0133] $$L = |M_t - M_{OF}|^2$$

[0134] 트레이닝부는 손실 함수를 감소시키는 방향으로 SRCNN의 파라미터를 업데이트함으로써, SRCNN에 대한 트레이닝을 진행한다. 한편, 수학식 3에서, 거리로서 L2 메트릭이 이용되었으나, L1 메트릭과 같이, 두 비교 대상 간의 거리 차이를 표현할 수 있는 어느 것이든 이용될 수 있다.

[0135] 한편, 본 개시에 따른 다른 실시예에 있어서, 개선 움직임벡터 M_t는 참조 샘플들로부터 개선 예측 샘플을 생성할 수 있는 픽셀 단위의 움직임벡터일 수 있다. 수학식 1은, 움직임벡터, OF 및 그래디언트를 이용하여, 하나의 픽셀에 대한 참조 샘플로부터 개선 예측 샘플이 산정되는 과정을 나타낸다. 따라서, 수학식 1에 나타낸 참조 샘플과 개선 예측 샘플 간의 픽셀 단위의 관계가 개선 움직임벡터 M_t로 할당될 수 있다. 이때, VPN(546)은 개선 움직임벡터 M_t를 이용하여 참조 샘플들로부터 개선 예측 샘플들을 픽셀 단위로 생성할 수 있다.

[0136] SRCNN은 움직임벡터에 해당하는 학습용 데이터, 및 픽셀 단위의 움직임벡터에 해당하는 레이블 M_{pix}를 기반으로 트레이닝될 수 있다. 본 실시예에 따른 트레이닝부는, SRCNN이 학습용 데이터로부터 생성한 개선 움직임벡터 M_t(즉, M_t에 대한 2차원 배열)와 레이블 M_{pix}(즉, M_{pix}에 대한 2차원 배열) 간의 거리(distance)에 기반하여, 수학식 4에 나타낸 바와 같이, 손실 함수(loss function) L을 정의한다.

수학식 4

[0137] $$L = |M_t - M_{pix}|^2$$

[0138] 트레이닝부는 손실 함수를 감소시키는 방향으로 SRCNN의 파라미터를 업데이트함으로써, SRCNN에 대한 트레이닝을 진행할 수 있다.

[0139] 전술한 바와 같은 ARCNN, RefAR 및 SRCNN은 모두 콘볼루션 레이어를 포함하는 CNN 기반의 심층신경망이나, 본 실시예에 따른 VPN이 포함하는 심층신경망이 ㅂ반드시 CNN 기반의 심층신경망에 한정되는 것은 아니다. 심층신경망은 구성요소로서, 콘볼루션 레이어, 풀링 레이어(pooling layer), 비선형 활성 함수(nonlinear activation funtion), 매트릭스 곱(matrix mutilplication) 레이어, 전연결 레이어(fully connected layer) 등을 포함할 수 있다. 따라서, 본 실시예에 따른 VPN은 전술한 바와 같은 구성요소의 전부 또는 일부를 포함하는 심층신경망을 포함하고, 개선 예측 샘플들을 생성하는 과정에서 이러한 심층신경망을 이용할 수 있다.

[0140] 이하, 기존의 인터 예측 기법들과 본 실시예에 따른 VPN을 결합하는 방안을 설명한다. 이하의 설명은 복호화 장치의 VPN(546)을 이용하나, 부호화 장치의 VPN(126)에도 동일하게 적용될 수 있다.

[0141] AMVP 모드 또는 머지 모드를 적용하여 획득한 움직임벡터 및 예측 샘플들에 대해, VPN(546)이 결합될 수 있다. 이때, 부호화 장치는 VPN(546)의 적용 여부를 지시하는 플래그를 전송하고, 복호화 장치에서는 수신한 플래그를 기반으로 VPN(546)의 적용 여부를 결정할 수 있다.

[0142] VPN(546)이 생성한 개선 예측 샘플들을 포함하는 최종 예측블록과 잔차 값들을 가산하여 생성된 복원블록을 포함하는 픽처는, 참조픽처 리스트에 포함되어 있는 기존 참조픽처를 대체하거나, 신규 참조픽처로 추가될 수 있다. 이후, 현재블록의 예측 과정에서, 참조픽처 리스트로부터 개선 예측 샘플들에 기반하는 참조픽처의 인덱스를 선택하여 참조 샘플들로 이용할 수 있다.

[0143] AMVP 모드 또는 머지 모드에 따른 인터 예측을 수행한 후, PROF 또는 BDOF를 적용하는 경우에 대해, 전술한 바

와 같이, VPN(546)은 SRCNN을 이용할 수 있다. 이때, BDOF 또는 PROF의 적용 조건과 동일하게 SRCNN의 적용 여부를 결정한다. 또한, 개선 움직임벡터 M_t의 정밀도는 BDOF 또는 PROF에서 사용하는 OF와 동일한 정밀도를 갖는다.

[0144] SRCNN에 기반하는 개선 움직임벡터 M_t를 이용하는 경우, 개선 움직임벡터의 최댓값은, 기존의 인터 예측에서 제공하는 움직임벡터의 최댓값을 초과하지 않도록 설정된다.

[0145] AMVP 모드 또는 머지 모드에 따른 인터 예측을 수행하는 경우, 전술한 바와 같이, 블록 위치에 따라 움직임벡터 후보 리스트에서의 움직임벡터 후보 순서가 결정된다. 그러나 VPN(546)을 이용하여 생성된 최종 예측블록이 현재블록 주변에 존재하는 경우, 최종 예측블록에 대한 움직임벡터(즉, 최종 예측블록의 생성에 이용된 예측 샘플들의 생성에 이용된 움직임벡터)의 경우에는 보다 높은 우선 순위가 부여될 수 있도록 한다. 예를 들어, 움직임벡터 후보 리스트에서 가장 높은 순서에 위치시킬 수 있다.

[0146] 이하, 본 실시예에 따른 VPN을 인터 예측 중 머지 모드에 적용하는 방법에 대해 설명한다.

[0147] 도 11은 본 개시의 일 실시예에 따른 머지 모드에서 VPN을 적용하는 방법에 대한 예시도이다.

[0148] VPN을 이용하는 경우에도, 부/복호화 장치는 과거에 부호화를 완료한 블록에 해당되는 움직임벡터를 기반으로 현재블록에 대한 머지 모드 예측을 수행할 수 있다.

[0149] 먼저, 부호화 장치는, 머지 모드 예측을 위해, 현재블록을 예측하기 위한 후보 블록 위치를 선택한다. 예를 들어, AR 블록을 선택하고 해당되는 움직임벡터 mv_1을 추출한다. 부호화 장치는 움직임벡터 mv_1을 이용하여 현재블록에 대한 예측 샘플들을 참조픽처 내의 참조 샘플들로부터 유도한다. 부호화 장치는 VPN(126)에 포함된 딥러닝 모델을 이용하여 움직임벡터, 참조 샘플들, 및 예측 샘플들의 전부 또는 일부로부터 가상 블록 V'을 생성한다. 부호화 장치는, AR 블록과 mv_1 이외의 후보 블록들(A, AL, L 및 BL), 및 mv_2~mv_5에 대해, 이러한 과정을 순서대로 적용하여, 비트율 왜곡 관점에서 최적의 가상블록 V'을 찾은 후, 해당되는 후보 블록의 위치(예컨대, AR)를 추출한다. 부호화 장치는, 현재블록이 최종 예측블록으로 가상블록 V'을 선택하였음을 표시하기 위해, 기존의 머지 후보 리스트에서 AR 인덱스를 대체하여 사용할 수 있다. 이때, 부호화 장치는 가상 블록 V'의 움직임벡터를 mv_1으로 설정한다.

[0150] 이상의 설명에서, V'을 생성하기 위한 VPN(126)에 포함되는 딥러닝 모델로는, 전술한 바와 같은, ARCNN 또는 RefAR이 이용될 수 있다.

[0151] 한편, 복호화 장치는, 비트스트림으로부터 머지 모드 예측을 위한 인덱스를 복호화한다. 예를 들어, 주변 블록 AR에 대한 인덱스를 복호한다. 복호화 장치는 AR에 해당하는 움직임벡터 mv_1을 이용하여 현재블록에 대한 예측 샘플들을 참조픽처 내의 참조 샘플들로부터 유도한다. 복호화 장치는 VPN(546)에 포함된 딥러닝 모델을 이용하여 움직임벡터, 참조 샘플들, 및 예측 샘플들의 전부 또는 일부로부터 가상 블록 V'을 생성하고, 이를 현재블록에 대한 최종 예측블록으로 이용할 수 있다.

[0152] 본 개시에 따른 다른 실시예에 있어서, 부/복호화 장치는 모든 후보 블록들 (AR, A, AL, L 및 BL)의 mv_1~mv_5을 기반으로, 현재블록에 대한 예측블록들을 생성한 후, 하나 또는 그 이상의 예측블록의 결합(예컨대, 가중합)을 기반으로 VPN을 이용하여 가상 블록 V'을 생성할 수 있다.

[0153] 본 개시에 따른 다른 실시예에 있어서, 부/복호화 장치는 과거에 부호화를 완료한 블록에 해당되는 움직임벡터를 개선하여 현재블록에 대한 머지 모드 예측을 수행할 수 있다.

[0154] 먼저, 부호화 장치는, 머지 모드 예측을 위해, 현재블록을 예측하기 위한 후보 블록 위치를 선택한다. 예를 들어, AR 블록을 선택하고 해당되는 움직임벡터 mv_1을 추출한다. 부호화 장치는 움직임벡터 mv_1을 이용하여 현재블록에 대한 예측 샘플들을 참조픽처 내의 참조 샘플들로부터 유도한다. 부호화 장치는 VPN(126)에 포함된 딥러닝 모델을 이용하여 움직임벡터로부터 개선 움직임벡터를 생성하고, 예측 샘플들 및 개선 움직임벡터를 이용하여 가상 블록 V'을 생성한다. 부호화 장치는, AR 블록과 mv_1 이외의 후보 블록(A, AL, L 및 BL), 및 mv_2~mv_5에 대해, 이러한 과정을 순서대로 적용하여, 비트율 왜곡 관점에서 최적의 가상블록 V'을 찾은 후, 해당되는 후보 블록의 위치(예컨대, AR)를 추출한다. 부호화 장치는, 현재블록이 최종 예측블록으로 가상블록 V'을 선택하였음을 표시하기 위해, 기존의 머지 후보 리스트에서 AR 인덱스를 대체하여 사용할 수 있다. 이때, 부호화 장치는

가상 블록 V'의 움직임벡터를 mv_1으로 설정한다.

[0155] 이상의 설명에서, V'을 생성하기 위한 VPN(126)에 포함되는 딥러닝 모델로는, 전술한 바와 같은, SRCNN이 이용될 수 있다. 현재블록을 서브샘플링한 픽셀 수만큼 현재블록의 움직임벡터를 반복하여 배열함으로써, SRCNN의 입력인 움직임벡터의 2차원 배열이 생성될 수 있다. 또한, SRCNN의 출력인 개선 움직임벡터의 2차원 배열에서, 각 개선 움직임벡터는 픽셀 단위의 OF일 수 있다.

[0156] 한편, 복호화 장치는, 비트스트림으로부터 머지 모드 예측을 위한 인덱스를 복호화한다. 예를 들어, 주변 블록 AR에 대한 인덱스를 복호한다. 복호화 장치는 AR에 해당하는 움직임벡터 mv_1을 이용하여 현재블록에 대한 예측 샘플들을 참조픽처 내의 참조 샘플들로부터 유도한다. 복호화 장치는 VPN(546)에 포함된 딥러닝 모델을 이용하여 움직임벡터로부터 개선 움직임벡터를 생성하고, 예측 샘플들 및 개선 움직임벡터를 이용하여 가상 블록 V'을 생성하여, 이를 현재블록에 대한 최종 예측블록으로 이용할 수 있다.

[0157] 본 개시에 따른 다른 실시예에 있어서, 부/복호화 장치는 모든 후보 블록 (AR, A, AL, L 및 BL)의 mv_1~mv_5의 기반으로, 현재블록에 대한 예측블록들을 생성하고, mv_1~mv_5 중 하나 또는 그 이상의 결합하여 형성된 2차원 배열을 기반으로 SRCNN을 이용하여 픽셀 단위의 개선 움직임벡터의 2차원 배열을 생성할 수 있다. 부/복호화 장치는, 하나 또는 그 이상의 예측블록의 결합(예컨대, 가중합) 및 픽셀 단위의 개선 움직임벡터를 이용하여 가상 블록 V'을 생성할 수 있다.

[0158] 본 개시에 따른 다른 실시예에 있어서, 부/복호화 장치는, 현재블록에 대한 최종 예측블록으로 기존의 후보 블록이 아닌 V'을 선택하였음을 표시하기 위해, 기존의 머지 후보 리스트에 별도의 인덱스를 추가할 수 있다. 이때, 부/복호화 장치는, V'의 해당 움직임벡터를 0 벡터로 설정하거나, 또는 기존의 움직임벡터 mv_1(즉, V'을 생성하기 위해 이용되었던 움직임벡터)으로 설정할 수 있다.

[0159] 도 12는 본 개시의 일 실시예에 따른, 딥러닝 기반 인터 예측을 이용하는 영상 부호화 방법의 흐름도이다.

[0160] 부호화 장치는, 참조픽처 내에서 부호화하고자 하는 현재블록에 대한 참조블록을 탐색하고, 탐색된 참조블록의 참조 샘플들로부터 현재블록에 대한 예측 샘플들을 생성한다(S1200).

[0161] 부호화 장치는 참조블록과 현재블록 간의 변위에 해당하는 움직임벡터를 생성한다(S1202). 여기서, 양방향 예측인 경우, 전술한 바와 같이, 참조블록과 움직임벡터는 각각 2 개일 수 있다.

[0162] 부호화 장치는 움직임벡터, 참조 샘플들, 및 예측 샘플들의 전부 또는 일부를 기반으로 딥러닝 기반 VPN을 이용하여 개선 예측 샘플들을 생성한다(S1204), 여기서, 개선 예측 샘플들은 최종 예측블록에 포함된다.

[0163] 또한, VPN(126)은, 잡음 제거에 적합한 딥러닝 모델인 ARCNN을 포함하고, ARCNN은 예측 샘플들을 입력으로 받아들여 개선 예측 샘플들을 생성하도록 트레이닝된다.

[0164] 본 개시의 다른 실시예에 있어서, VPN(126)은, 예측 샘플들 및 참조 샘플들을 이용하는 딥러닝 모델인 RefAR을 포함하고, RefAR은 예측 샘플들, 과거 참조 샘플들, 미래 참조 샘플들, 및 움직임벡터를 입력으로 받아들여, 개선 예측 샘플들을 생성하도록 트레이닝된다.

[0165] 본 개시의 다른 실시예에 있어서, VPN(126)은, 저해상도 영상을 입력으로 받아들여 고해상도 영상을 출력으로 생성하는 딥러닝 모델인 SRCNN을 포함하고, SRCNN은 복수의 움직임벡터를 포함하는 2차원 배열을 입력으로 받아들여, 복수의 개선 움직임벡터를 포함하는 2차원 배열을 생성하도록 트레이닝된다.

[0166] 복수의 개선 움직임벡터는, 예컨대, 픽셀에 대한 OF인 경우, 예측 샘플들을 보정하기 위한 샘플 오프셋을 생성하기 위해 이용될 수 있다. VPN(126)은 예측 샘플들과 샘플 오프셋을 픽셀 단위로 가산하여 개선 예측 샘플들을 생성할 수 있다.

[0167] 본 개시에 따른 다른 실시예에 있어서, 복수의 개선 움직임벡터는, 참조 샘플들로부터 개선 예측 샘플을 생성할 수 있는 픽셀 단위의 움직임벡터일 수 있다. VPN(126)은 개선 움직임벡터를 이용하여 참조 샘플들로부터 개선 예측 샘플들을 픽셀 단위로 생성할 수 있다.

[0168] 부호화 장치는 최종 예측블록을 현재블록으로부터 감산하여 잔차블록을 생성한다(S1206), 부호화 장치는 잔차블록의 잔차 값들을 변환부(140)에 입력하여, 전술한 바와 같은 부호화 과정을 수행할 수 있다.

[0169] 한편, 사전에 트레이닝된 딥러닝 모델의 파라미터는 부호화 장치와 복호화 장치 간에 공유될 수 있다.

[0170] 도 13은 본 개시의 일 실시예에 따른, 딥러닝 기반 인터 예측을 이용하는 영상 복호화 방법의 흐름도이다.

[0171] 복호화 장치는, 복호화하고자 하는 현재블록에 대해, 비트스트림으로부터 현재블록에 대한 움직임벡터, 및 현재블록의 잔차 값들을 포함하는 잔차블록을 획득한다(S1300). 양방향 예측인 경우, 전술한 바와 같이, 움직임벡터는 2 개일 수 있다.

[0172] 복호화 장치는 움직임벡터를 이용하여 참조 샘플들로부터 현재블록에 대한 예측 샘플들을 생성한다(S1302). 여기서, 참조 샘플들은 움직임벡터가 지시하는 참조픽처로부터 생성될 수 있다.

[0173] 복호화 장치는 움직임벡터, 참조 샘플들, 및 예측 샘플들의 전부 또는 일부를 기반으로 딥러닝 기반 VPN을 이용하여 개선 예측 샘플들을 생성한다(S1304). 여기서, 개선 예측 샘플들은 최종 예측블록에 포함된다.

[0174] 복호화 장치는 개선 예측 샘플들을 잔차 값들에 가산하여 현재블록에 대한 복원블록을 생성한다(S1306).

[0175] 본 실시예에 따른 각 순서도에서는 각각의 과정을 순차적으로 실행하는 것으로 기재하고 있으나, 반드시 이에 한정되는 것은 아니다. 다시 말해, 순서도에 기재된 과정을 변경하여 실행하거나 하나 이상의 과정을 병렬적으로 실행하는 것이 적용 가능할 것이므로, 순서도는 시계열적인 순서로 한정되는 것은 아니다.

[0176] 한편, 본 실시예에서 설명된 다양한 기능들 혹은 방법들은 하나 이상의 프로세서에 의해 판독되고 실행될 수 있는 비일시적 기록매체에 저장된 명령어들로 구현될 수도 있다. 비일시적 기록매체는, 예를 들어, 컴퓨터 시스템에 의하여 판독가능한 형태로 데이터가 저장되는 모든 종류의 기록장치를 포함한다. 예를 들어, 비일시적 기록매체는 EPROM(erasable programmable read only memory), 플래시 드라이브, 광학 드라이브, 자기 하드 드라이브, 솔리드 스테이트 드라이브(SSD)와 같은 저장매체를 포함한다.

[0177] 이상의 설명은 본 실시예의 기술 사상을 예시적으로 설명한 것에 불과한 것으로서, 본 실시예가 속하는 기술 분야에서 통상의 지식을 가진 자라면 본 실시예의 본질적인 특성에서 벗어나지 않는 범위에서 다양한 수정 및 변형이 가능할 것이다. 따라서, 본 실시예들은 본 실시예의 기술 사상을 한정하기 위한 것이 아니라 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.

부호의 설명

[0178] 124: 인터 예측부 126: VPN

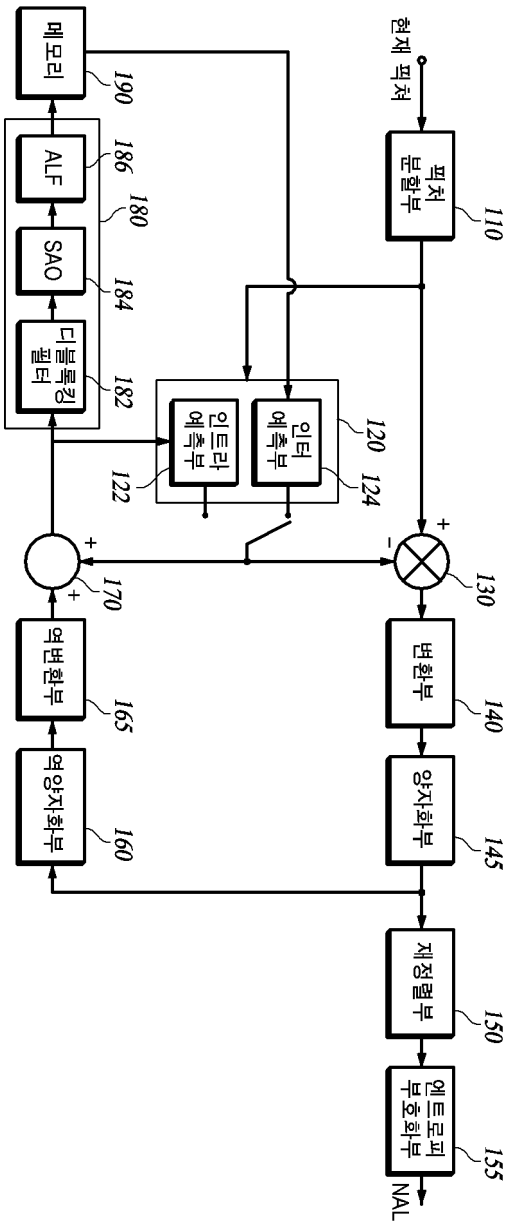
130: 감산기 140: 변환부

510: 엔트로피 복호화부 544: 인터 예측부

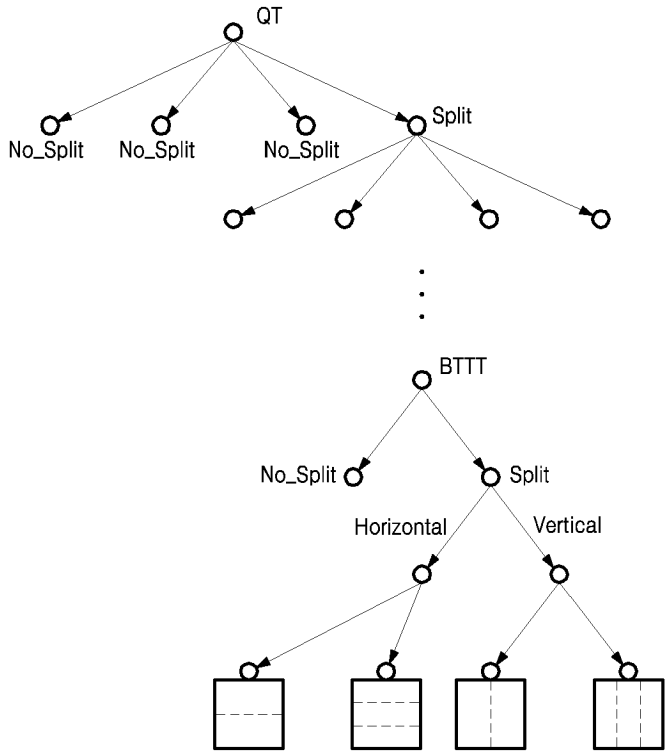
546: VPN 550: 가산기

도면

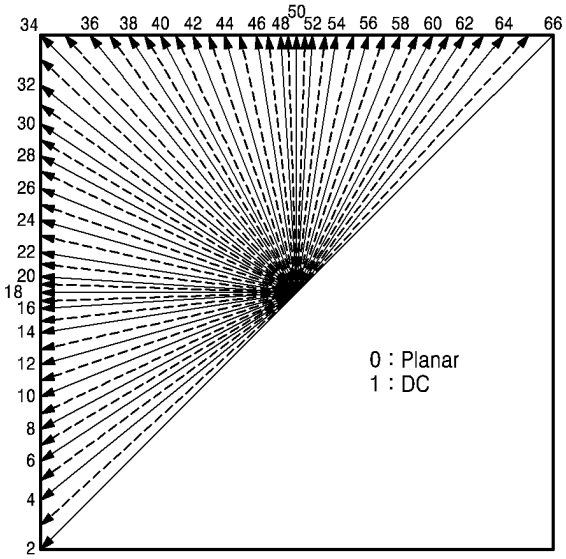
도면1


현재 픽쳐
픽처 분할부
110
변환부
140
양자화부
145
제정렬부
150
엔트로피 부호화부
155
NAL
역양자화부
160
역변환부
165
인터 예측부
120
인트라 예측부
122
124
디블록킹 필터
182
SAO
184
ALF
186
180
메모리
190
130
170

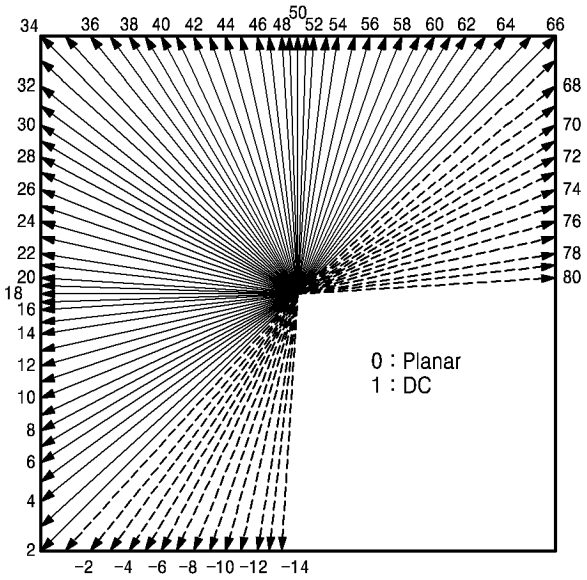
도면2


QT
No_Split
No_Split
No_Split
Split
BTTT
No_Split
Split
Horizontal
Vertical

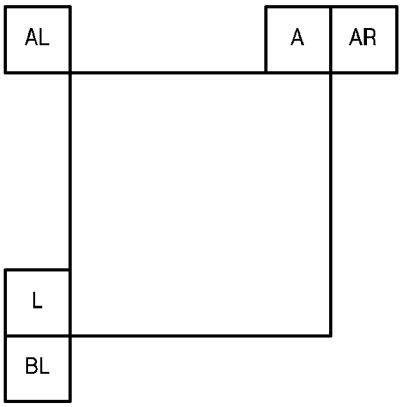
도면3a


34
36
38
40
42
44
46
48
50
52
54
56
58
60
62
64
66
32
30
28
26
24
22
20
18
16
14
12
10
8
6
4
2
0 : Planar
1 : DC

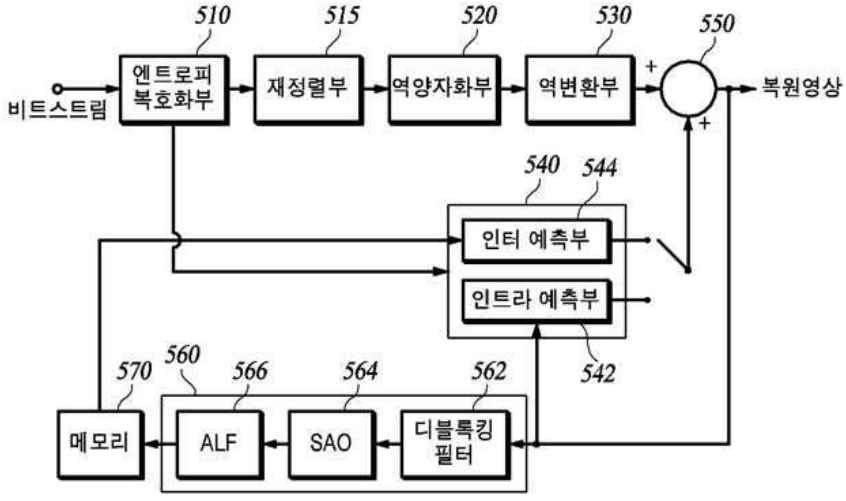
도면3b


50
34 36 38 40 42 44 46 48 52 54 56 58 60 62 64 66
32 30 28 26 24 22 20 18 16 14 12 10 8 6 4 2
68 70 72 74 76 78 80
-2 -4 -6 -8 -10 -12 -14
0 : Planar
1 : DC

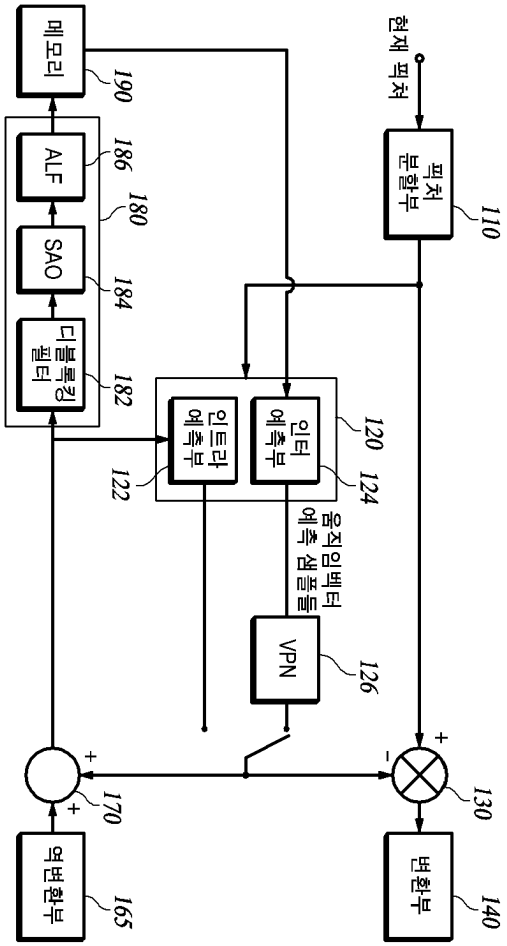
도면4


AL
A
AR
L
BL

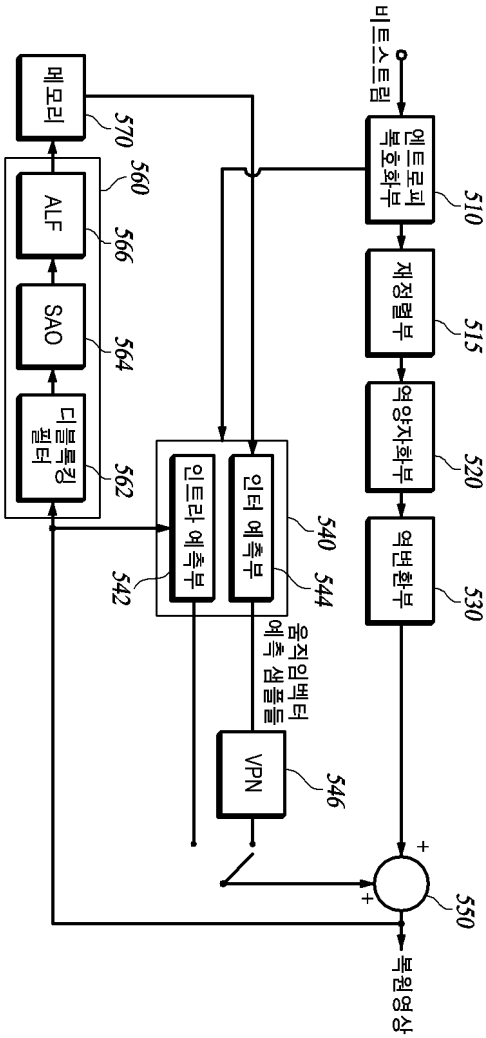
도면5


510
515
520
530
550
비트스트림
엔트로피 복호화부
재정렬부
역양자화부
역변환부
복원영상
540
544
인터 예측부
인트라 예측부
542
570
560
566
564
562
메모리
ALF
SAO
디블록킹 필터

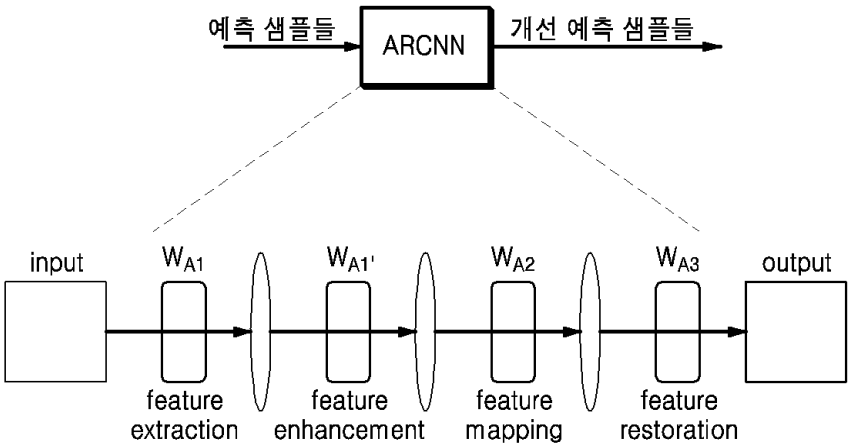
도면6


메모리
190
ALF
186
SAO
184
디블록킹 필터
182
180
현재 픽처
픽처 분할부
110
인트라 예측부
122
인터 예측부
124
120
움직임벡터 예측 샘플들
VPN
126
+
+
170
역변환부
165
+
-
130
변환부
140

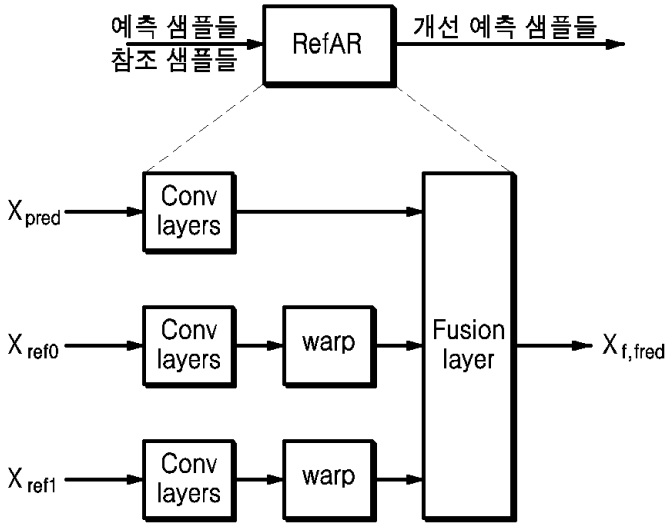
도면7


비트스트림
엔트로피 복호화부
510
재정렬부
515
역양자화부
520
역변환부
530
+
+
550
복원영상
메모리
570
ALF
560
566
SAO
564
디블록킹 필터
562
인트라 예측부
542
인터 예측부
540
544
움직임벡터 예측 샘플들
VPN
546

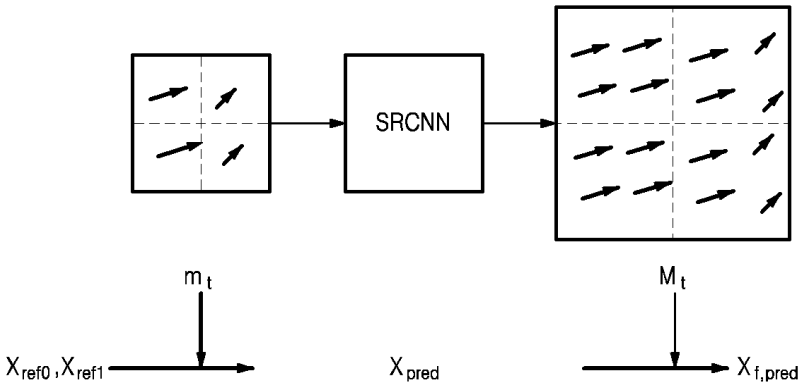
도면8


예측 샘플들
ARCNN
개선 예측 샘플들
input
W_{A1}
$W_{A1'}$
W_{A2}
W_{A3}
output
feature extraction
feature enhancement
feature mapping
feature restoration

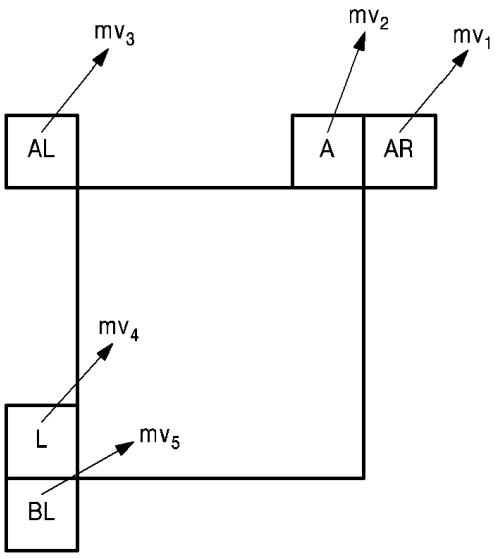
도면9



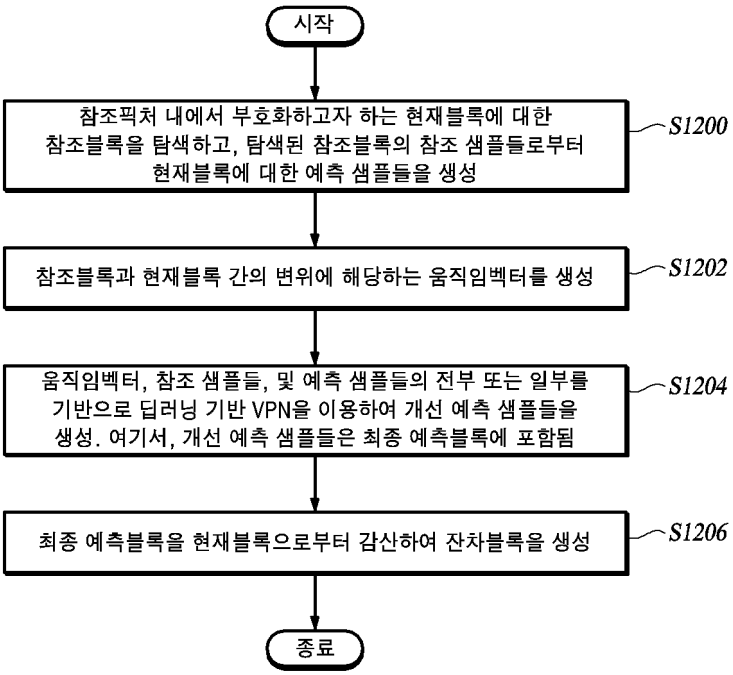
도면10



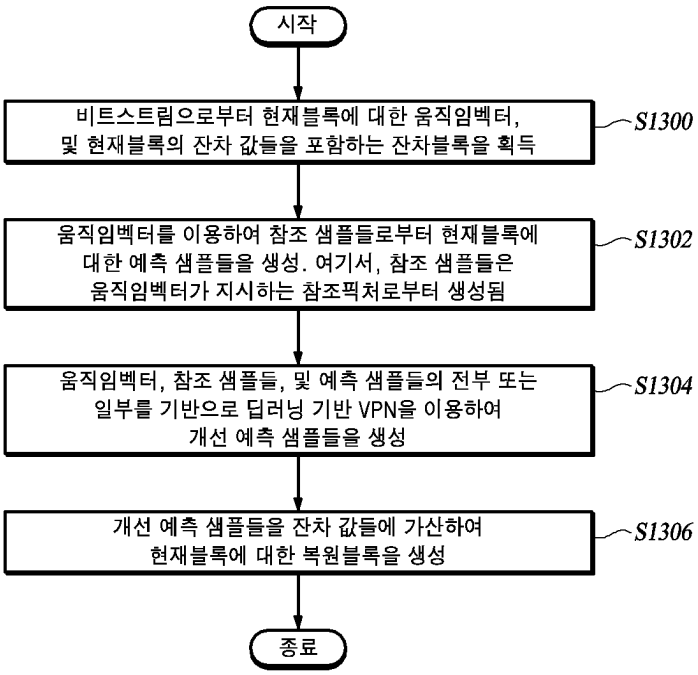
도면11



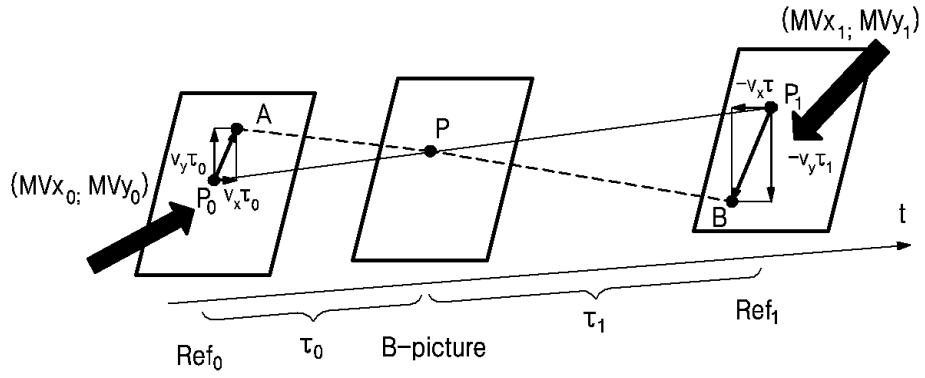
도면12


시작
참조픽처 내에서 부호화하고자 하는 현재블록에 대한 참조블록을 탐색하고, 탐색된 참조블록의 참조 샘플들로부터 현재블록에 대한 예측 샘플들을 생성
S1200
참조블록과 현재블록 간의 변위에 해당하는 움직임벡터를 생성
S1202
움직임벡터, 참조 샘플들, 및 예측 샘플들의 전부 또는 일부를 기반으로 딥러닝 기반 VPN을 이용하여 개선 예측 샘플들을 생성. 여기서, 개선 예측 샘플들은 최종 예측블록에 포함됨
S1204
최종 예측블록을 현재블록으로부터 감산하여 잔차블록을 생성
S1206
종료

도면13


시작
비트스트림으로부터 현재블록에 대한 움직임벡터, 및 현재블록의 잔차 값들을 포함하는 잔차블록을 획득
S1300
움직임벡터를 이용하여 참조 샘플들로부터 현재블록에 대한 예측 샘플들을 생성. 여기서, 참조 샘플들은 움직임벡터가 지시하는 참조픽처로부터 생성됨
S1302
움직임벡터, 참조 샘플들, 및 예측 샘플들의 전부 또는 일부를 기반으로 딥러닝 기반 VPN을 이용하여 개선 예측 샘플들을 생성
S1304
개선 예측 샘플들을 잔차 값들에 가산하여 현재블록에 대한 복원블록을 생성
S1306
종료

도면14


(MVx1; MVy1)
A
P
P1
−vxτ
vyτ0
−vyτ1
(MVx0; MVy0)
P0
vxτ0
B
t
Ref0
τ0
B-picture
τ1
Ref1