

(12) 특허협력조약에 의하여 공개된 국제출원

(19) 세계지식재산권기구
국제사무국

(43) 국제공개일
2022년 3월 31일 (31.03.2022) WIPO | PCT



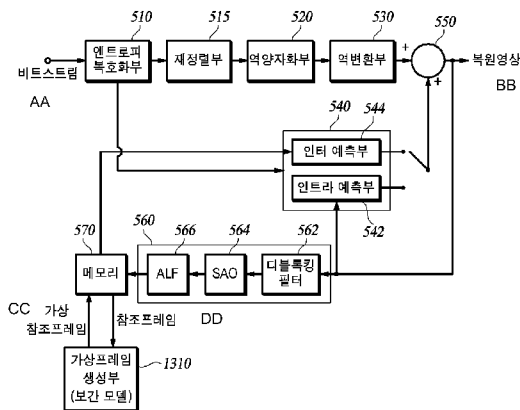
(10) 국제공개번호
WO 2022/065977 A1

- (51) 국제특허분류:
G06T 9/00 (2006.01) H04N 19/70 (2014.01)
G06T 3/40 (2006.01) H04N 19/109 (2014.01)
G06T 5/00 (2006.01) H04N 19/503 (2014.01)
H04N 19/82 (2014.01) G06N 3/08 (2006.01)
H04N 19/86 (2014.01) G06N 3/04 (2006.01)
- (21) 국제출원번호: PCT/KR2021/013218
- (22) 국제출원일: 2021년 9월 28일 (28.09.2021)
- (25) 출원언어: 한국어
- (26) 공개언어: 한국어
- (30) 우선권정보:
10-2020-0126111 2020년 9월 28일 (28.09.2020) KR
10-2021-0127795 2021년 9월 28일 (28.09.2021) KR
- (71) 출원인: 현대자동차주식회사 (HYUNDAI MOTOR COMPANY) [KR/KR]; 06797 서울시 서초구 헌릉로 12, Seoul (KR). 기아 주식회사 (KIA CORPORATION) [KR/KR]; 06797 서울시 서초구 헌릉로 12, Seoul (KR). 이화여자대학교 산학협력단 (EWhA UNIVERSITY - INDUSTRY COLLABORATION FOUNDATION) [KR/KR]; 03760 서울시 서대문구 이화여대길 52, Seoul (KR).

- (72) 발명자: 강제원 (KANG, Je Won); 04069 서울시 마포구 독막로 145, 104-904, Seoul (KR). 김나영 (KIM, Na Young); 02740 서울시 성북구 오패산로3길 17, 104동 101호, Seoul (KR). 이정경 (LEE, Jung Kyung); 07987 서울시 양천구 목동동로 350, 506동 1201호, Seoul (KR). 박승욱 (PARK, Seung Wook); 16803 경기도 용인시 수지구 태봉로 17, 403동 302호, Gyeonggi-do (KR). 임화평 (LIM, Wha Pyeong); 18271 경기도 화성시 남양읍 역골로 49번길 29 202호, Gyeonggi-do (KR).
- (74) 대리인: 이철희 (LEE, Chulhee); 06229 서울시 강남구 도곡로33길 26, 베리타스빌딩 2-4층, Seoul (KR).
- (81) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 국내 권리의 보호를 위하여): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(54) Title: INTER PREDICTION METHOD BASED ON VARIABLE COEFFICIENT DEEP LEARNING

(54) 발명의 명칭: 가변계수 딥러닝 기반 인터 예측 방법



510 ... Entropy-decoding unit
515 ... Realignment unit
520 ... Dequantizing unit
530 ... Inverse-transforming unit
542 ... Intra prediction unit
544 ... Inter prediction unit
562 ... Deblocking filter
570 ... Memory
1310 ... Virtual frame generation unit (interpolation model)
AA ... Bitstream
BB ... Reconstructed image
CC ... Virtual reference frame
DD ... Reference frame

(57) Abstract: The present embodiment provides an inter prediction method which: allows a variable coefficient deep learning model to adaptively learn characteristics of a video; transmits a variable coefficient deep learning model parameter generated from the learning from an image encoding device to an image decoding device; and refers to a virtual reference frame generated by the variable coefficient deep learning model.

(57) 요약서: 본 실시예는, 가변계수 딥러닝 모델로 하여금 비디오의 특성을 적응적으로 학습하도록 하고, 학습으로부터 생성된 가변계수 딥러닝 모델의 파라미터를 영상 부호화 장치로부터 영상 복호화 장치로 전송하며, 가변계수 딥러닝 모델에 의해 생성된 가상의 참조프레임을 참조하는 인터 예측 방법을 제공한다.



(84) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 역내 권리의 보호를 위하여): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 유라시아 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 유럽 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

공개:

— 국제조사보고서와 함께 (조약 제21조(3))

명세서

발명의 명칭: 가변계수 딥러닝 기반 인터 예측 방법

기술분야

- [1] 본 개시는 영상 부호화 또는 복호화 장치에 포함되는, 가변계수 딥러닝에 기반하는 인터 예측 방법에 관한 것이다.

배경기술

- [2] 이하에 기술되는 내용은 단순히 본 발명과 관련되는 배경 정보만을 제공할 뿐 종래기술을 구성하는 것이 아니다.
- [3] 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.
- [4] 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.
- [5] 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.
- [6] 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, 영상복원 모델 기반의 인루프 필터(비특허문헌 1 참조) 등이 있다. 따라서, 영상 부호화/복호화에 있어서, 부호화 효율을 향상시키기 위해, 딥러닝 기반 영상처리 기술의 지속적인 적용이 고려될 필요가 있다.
- [7] (선행기술문헌)
- [8] (비특허문헌)
- [9] (비특허문헌 1) (비특허문헌 1) Lee et al., Deep Video Prediction Network Based Inter-Frame Coding in HEVC, IEEE Access 2020.

발명의 상세한 설명

기술적 과제

- [10] 본 개시는, 가변계수 딥러닝 모델로 하여금 비디오의 특성을 적응적으로 학습하도록 하고, 학습으로부터 생성된 가변계수 딥러닝 모델의 파라미터를 영상 부호화 장치로부터 영상 복호화 장치로 전송하며, 가변계수 딥러닝 모델에

의해 생성된 가상의 참조 프레임을 참조하는 인터 예측 방법을 제공하는 데 목적이 있다.

과제 해결 수단

- [11] 본 개시의 실시예에 따르면, 영상 복호화 장치가 수행하는 영상 복호화 방법에 있어서, 비트스트림으로부터 가변계수 값들, 아핀(affine) 예측 플래그, 및 부호화 모드를 복호화하는 단계, 여기서, 상기 아핀 예측 플래그는 현재블록에 대한 아핀 움직임 예측의 적용 여부를 나타내고, 상기 부호화 모드는 상기 현재블록의 움직임 정보의 부호화 모드로서 머지 모드(merge mode) 또는 AMVP(Advanced Motion Vector Prediction) 모드를 나타냄; 보간 모델(interpolation model)을 이용하여 참조 프레임들에 기초하여 가상 참조프레임을 생성하는 단계, 여기서, 상기 가변계수 값들을 이용하여 상기 보간 모델 상의 가변계수 네트워크가 설정됨; 및 상기 부호화 모드가 상기 머지 모드인 경우, 상기 가상 참조프레임 및 상기 참조 프레임들에 기초하여 상기 현재블록의 머지 후보를 생성하는 단계를 포함하는 것을 특징으로 하는, 영상 복호화 방법을 제공한다.
- [12] 본 개시의 다른 실시예에 따르면, 영상 부호화 장치가 수행하는 영상 부호화 방법에 있어서, 사전에 생성된 가변계수 값들, 및 기설정된 부호화 모드를 획득하는 단계, 여기서, 상기 부호화 모드는 현재블록의 움직임 정보의 부호화 모드로서 머지 모드(merge mode) 또는 AMVP(Advanced Motion Vector Prediction) 모드를 나타냄; 보간 모델(interpolation model)을 이용하여 참조 프레임들로부터 가상 참조프레임을 생성하는 단계, 여기서, 상기 가변계수 값들을 이용하여 상기 보간 모델 상의 가변계수 네트워크가 설정됨; 및 상기 부호화 모드가 상기 머지 모드인 경우, 상기 가상 참조프레임 및 상기 참조 프레임들에 기초하여 상기 현재블록의 머지 후보를 생성하는 단계를 포함하는 것을 특징으로 하는, 영상 부호화 방법을 제공한다.
- [13] 본 개시의 다른 실시예에 따르면, 비트스트림으로부터 가변계수 값들, 및 부호화 모드를 복호화하는 엔트로피 복호화부, 여기서, 상기 부호화 모드는 현재블록의 움직임 정보의 부호화 모드로서 머지 모드(merge mode) 또는 AMVP(Advanced Motion Vector Prediction) 모드를 나타냄; 보간 모델(interpolation model)을 이용하여 참조 프레임들로부터 가상 참조프레임을 생성하는 가상프레임 생성부, 여기서, 상기 가변계수 값들을 이용하여 상기 보간 모델 상의 가변계수 네트워크가 설정됨; 및 상기 부호화 모드가 상기 머지 모드인 경우, 상기 가상 참조프레임 및 상기 참조 프레임들에 기초하여 상기 현재블록의 머지 후보를 생성하는 인터 예측부를 포함하는 것을 특징으로 하는, 영상 복호화 장치를 제공한다.

발명의 효과

- [14] 이상에서 설명한 바와 같이 본 실시예에 따르면, 가변계수 딥러닝 모델에 의해 생성된 가상의 참조 프레임을 참조하는 인터 예측 방법을 제공함으로써, 인터

예측의 효과를 증대시키고, 부호화 효율을 향상시키는 것이 가능해지는 효과가 있다.

도면의 간단한 설명

- [15] 도 1은 본 개시의 기술들을 구현할 수 있는 영상 부호화 장치에 대한 예시적인 블록도이다.
- [16] 도 2는 QTBT 구조를 이용하여 블록을 분할하는 방법을 설명하기 위한 도면이다.
- [17] 도 3a 및 도 3b는 광각 인트라 예측모드들을 포함한 복수의 인트라 예측모드들을 나타낸 도면이다.
- [18] 도 4는 현재블록의 주변블록에 대한 예시도이다.
- [19] 도 5는 본 개시의 기술들을 구현할 수 있는 영상 복호화 장치의 예시적인 블록도이다.
- [20] 도 6은 본 개시의 일 실시예에 따른 BDOF의 개념을 설명하기 위한 예시도이다.
- [21] 도 7은 본 개시의 일 실시예에 따른, 머지/스킵 모드에서 움직임벡터의 후보를 탐색하는 과정을 나타내는 순서도이다.
- [22] 도 8은 본 개시의 일 실시예에 따른 아핀 움직임 예측을 설명하기 위한 예시도이다.
- [23] 도 9는 본 개시의 일 실시예에 따른 서브블록 단위의 아핀 움직임 예측을 설명하기 위한 예시도이다.
- [24] 도 10은 주변블록의 병진 움직임벡터로부터 아핀 움직임을 위한 머지 후보들을 유도하는 방법을 설명하기 위한 예시도이다.
- [25] 도 11은 본 개시의 일 실시예에 따른, 아핀 AMVP 모드에서 아핀 AMVP 후보를 탐색하는 과정을 나타내는 순서도이다.
- [26] 도 12는 본 개시의 일 실시예에 따른 보간 모델을 포함하는 영상 부호화 장치를 나타내는 블록도이다.
- [27] 도 13은 본 개시의 일 실시예에 따른 보간 모델을 포함하는 영상 복호화 장치를 나타내는 블록도이다.
- [28] 도 14는 본 개시의 일 실시예에 따른, 영상 부호화 장치 및 영상 복호화 장치에서 이용하는 보간 모델의 동작을 나타내는 예시도이다.
- [29] 도 15는 본 개시의 실시예에 따른 가변계수 보간 모델의 예시들이다.
- [30] 도 16은 본 개시의 일 실시예에 따른 고정계수 네트워크를 더 포함하는 보간 모델을 나타내는 예시도이다.
- [31] 도 17은 본 개시의 일 실시예에 따른, 가변계수를 전송하는 방법을 나타내는 예시도이다.
- [32] 도 18은 본 개시의 일 실시예에 따른, 인터 예측의 머지/스킵 모드에서 가상 참조프레임을 이용하는 과정에 대한 순서도를 나타낸다.
- [33] 도 19는 본 개시의 일 실시예에 따른, 인터 예측의 아핀 AMVP 모드에서 가상

참조프레임을 이용하는 과정에 대한 순서도를 나타낸다.

발명의 실시를 위한 형태

- [34] 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.
- [35] 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.
- [36] 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.
- [37] 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.
- [38] 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.

- [39] 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.
- [40] 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.
- [41] 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 QTBT(TT(QuadTree plus BinaryTree TernaryTree) 구조가 사용될 수 있다. 여기서, BT(TT)를 합쳐서 MTT(Multiple-Type Tree)라 지칭될 수 있다.
- [42] 도 2는 QTBT(TT) 구조를 이용하여 블록을 분할하는 방법을 설명하기 위한 도면이다.
- [43] 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.
- [44] 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.
- [45] 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.
- [46] CU는 CTU로부터의 QTBT 또는 QTBT_TT 분할에 따라 다양한 크기를 가질 수 있다. 이하에서는, 부호화 또는 복호화하고자 하는 CU(즉, QTBT_TT의 리프 노드)에 해당하는 블록을 '현재블록'이라 칭한다. QTBT_TT 분할의 채용에 따라, 현재블록의 모양은 정사각형뿐만 아니라 직사각형일 수도 있다.
- [47] 예측부(120)는 현재블록을 예측하여 예측블록을 생성한다. 예측부(120)는 인트라 예측부(122)와 인터 예측부(124)를 포함한다.
- [48] 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.
- [49] 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.
- [50] 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.

- [51] 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.
- [52] 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.
- [53] 인트라 예측부(124)는 움직임 보상 과정을 이용하여 현재블록에 대한 예측블록을 생성한다. 인트라 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.
- [54] 인트라 예측부(124)는, 예측의 정확성을 높이기 위해, 참조픽처 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 참조픽처에 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.

- [55] 한편, 인터 예측부(124)는 양방향 예측(bi-prediction)을 이용하여 인터 예측을 수행할 수 있다. 양방향 예측의 경우, 두 개의 참조픽처와 각 참조픽처 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.
- [56] 움직임 정보를 부호화하는 데에 소요되는 비트량을 최소화하기 위해 다양한 방법이 사용될 수 있다.
- [57] 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.
- [58] 머지 모드에서, 인터 예측부(124)는 현재블록의 주변블록들로부터 기 결정된 개수의 머지 후보블록(이하, '머지 후보'라 함)들을 선택한다.
- [59] 머지 후보를 유도하기 위한 주변블록으로는, 도 4에 도시된 바와 같이, 현재 픽처 내에서 현재블록에 인접한 좌측블록(A0), 좌하단블록(A1), 상단블록(B0), 우상단블록(B1), 및 좌상단블록(A2) 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.
- [60] 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.

- [61] 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.
- [62] 이하, 머지 모드와 머지 스킵 모드를 통칭하여, 머지/스킵 모드로 나타낸다.
- [63] 움직임 정보를 부호화하기 위한 또 다른 방법은 AMVP(Advanced Motion Vector Prediction) 모드이다.
- [64] 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(A0), 좌하단블록(A1), 상단블록(B0), 우상단블록(B1), 및 좌상단블록(A2) 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.
- [65] 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.
- [66] 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.
- [67] 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.

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

[69] 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.

[70] 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.

[71] 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)를 적용할 수도 있다. 2차원으로 배열된 양자화된 변환 계수들에 적용되는 양자화 행렬은 부호화되어 영상 복호화 장치로 시그널링될 수 있다.

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

[73] 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 사용될 스캔 방법이 결정될 수도 있다.

- [74] 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.
- [75] 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(움직임 정보의 부호화 모드(머지 모드 또는 AMVP 모드), 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.
- [76] 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.
- [77] 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.
- [78] 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.
- [79] 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.

- [80] 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 참조픽처로 사용될 수 있다.
- [81] 도 5는 본 개시의 기술들을 구현할 수 있는 영상 복호화 장치의 예시적인 블록도이다. 이하에서는 도 5를 참조하여 영상 복호화 장치와 이 장치의 하위 구성들에 대하여 설명하도록 한다.
- [82] 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.
- [83] 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.
- [84] 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.
- [85] 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.
- [86] 예컨대, QTBT 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.
- [87] 또 다른 예로서, QTBT 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.

- [88] 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.
- [89] 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 선택스 요소를 추출한다. 예측 타입 정보가 인터 예측을 지시하는 경우, 엔트로피 복호화부(510)는 인터 예측정보에 대한 선택스 요소, 즉, 움직임벡터 및 그 움직임벡터가 참조하는 참조픽처를 나타내는 정보를 추출한다.
- [90] 또한, 엔트로피 복호화부(510)는 양자화 관련된 정보, 및 잔차신호에 대한 정보로서 현재블록의 양자화된 변환계수들에 대한 정보를 추출한다.
- [91] 제정렬부(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.
- [92] 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)를 적용할 수도 있다. 역양자화부(520)는 영상 부호화 장치로부터 양자화 계수(스케일링 값)들의 행렬을 양자화된 변환계수들의 2차원 어레이에 적용하여 역양자화를 수행할 수 있다.
- [93] 역변환부(530)는 역양자화된 변환계수들을 주파수 도메인으로부터 공간 도메인으로 역변환하여 잔차신호들을 복원함으로써 현재블록에 대한 잔차블록을 생성한다.
- [94] 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.

- [95] 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.
- [96] 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.
- [97] 인트라 예측부(542)는 엔트로피 복호화부(510)로부터 추출된 인트라 예측모드에 대한 선택스 요소로부터 복수의 인트라 예측모드 중 현재블록의 인트라 예측모드를 결정하고, 인트라 예측모드에 따라 현재블록 주변의 참조픽셀들을 이용하여 현재블록을 예측한다.
- [98] 인터 예측부(544)는 엔트로피 복호화부(510)로부터 추출된 인터 예측모드에 대한 선택스 요소를 이용하여 현재블록의 움직임벡터와 그 움직임벡터가 참조하는 참조픽처를 결정하고, 움직임벡터와 참조픽처를 이용하여 현재블록을 예측한다.
- [99] 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.
- [100] 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 블로킹 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.
- [101] 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 참조픽처로 사용된다.
- [102]
- [103] 본 실시예는 이상에서 설명한 바와 같은 영상(비디오)의 부호화 및 복호화에 관한 것이다. 보다 자세하게는, 가변계수 딥러닝 모델로 하여금 비디오의 특성을 적응적으로 학습하도록 하고, 학습으로부터 생성된 가변계수 딥러닝 모델의 파라미터를 영상 부호화 장치로부터 영상 복호화 장치로 전송하며, 가변계수 딥러닝 모델에 의해 생성된 가상의 참조 프레임을 참조하는 인터 예측 방법을 제공한다.

- [104] 이하의 실시예는 영상 부호화 장치의 인터 예측부(124) 및 영상 복호화 장치의 인터 예측부(544)에서 수행될 수 있다. 이하, '대상블록(target block)'이라는 용어는 전술한 바와 같은 현재블록 또는 코딩 유닛(CU)과 동일한 의미로 사용될 수 있고, 또는 코딩 유닛의 일부 영역을 의미할 수도 있다.
- [105] **I. 양방향 옵티컬 플로우(Bi-directional Optical Flow: BDOF)**
- [106] 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.
- [107] 대상블록에 대한 (통상의) 양방향 움직임 예측에 의하여, 도 6에 예시된 바와 같이, 참조픽처들(Ref₀과 Ref₁)에서 현재 픽처의 부호화되는 대상블록과 가장 유사한 대응 영역들(즉, 참조 블록들)을 가리키는 양방향의 움직임벡터들(MV₀, MV₁)이 결정되었다고 가정하자. 이때, 두 개의 움직임벡터들은 대상블록 전체의 움직임을 나타내는 값이다. 도 6의 예시에서, 대상블록 내 샘플 P와 대응되는, 움직임벡터(MV₀)가 가리키는 참조픽처(Ref₀)내 샘플은 P₀이며, 대상블록 내 샘플 P와 대응되는, 움직임벡터(MV₁)가 가리키는 참조픽처(Ref₁)내 샘플은 P₁이다.
- [108] 이하, 도 6의 예시에서, 샘플 P에 대한 움직임은 대상블록의 전체적인 움직임과는 조금 상이하다라고 가정하자. 예컨대, Ref₀ 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 Ref₀ 내 지점은 움직임벡터(MV₀)가 가리키는 P₀이 아니라, P₀를 소정의 변위 벡터($v_x t_0, v_y t_0$)만큼 이동시킨 A이다. 또한, 대상블록 내 샘플 P와 가장 유사한 Ref₁ 내 지점은 움직임벡터(MV₁)가 가리키는 P₁이 아니라, P₁을 소정의 변위 벡터($-v_x t_1, -v_y t_1$)만큼 이동시킨 B이다. 여기서, t_0 와 t_1 은 현재 픽처를 기준으로 각각 Ref₀와 Ref₁에 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.
- [109] BDOF에서는, 현재 픽처 내 대상블록의 샘플(P)값을 예측함에 있어서, 두 참조 샘플(A,B)의 값을 이용하여, 양방향 움직임벡터(MV₀, MV₁)가 가리키는 참조 샘플(P₀, P₁), 즉 예측 샘플을 이용하는 것보다 좀 더 정확한 예측을 수행한다.
- [110] 대상블록 내 (루마) 픽셀 (x,y)와 대응되는, 움직임벡터(MV₀)가 가리키는 참조픽처(Ref₀)내 예측샘플의 값을 $I^{(0)}(x,y)$ 라고 정의하고, 대상블록 내 샘플 (x,y)와 대응되는, 움직임벡터(MV₁)가 가리키는 참조픽처(Ref₁)내 예측샘플의 값을 $I^{(1)}(x,y)$ 라고 정의한다.
- [111] BDOF 움직임벡터 (v_x, v_y)가 대상블록 내 샘플에 대응되는 것으로 가리키는 참조픽처 Ref₀ 내의 샘플 A의 값은 $I^{(0)}(i + v_x t_0, j + v_y t_0)$ 로 정의될 수 있으며, 참조픽처 Ref₁ 내의 샘플 B의 값은 $I^{(1)}(i - v_x t_1, j - v_y t_1)$ 로 정의될 수 있다. 또한, 시간 t_0 와 t_1 이 동일하다고 가정하고, ' t_0+t_1 '를 이용하여 정규화하면, 샘플 A의 값은 $I^{(0)}(i$

$+ \frac{1}{2}v_{x,j} + \frac{1}{2}v_{y,j}$)로 정의될 수 있으며, 샘플 B의 값은 $I^{(0)}(i - \frac{1}{2}v_{x,j} - \frac{1}{2}v_{y,j})$ 로 정의될 수 있다

- [112] BDOF에서는, 픽셀에 대한 예측샘플들의 값, OF 및 예측샘플에서의 그라디언트(gradient)를 이용하여, 수학적 식 1에 나타난 바와 같이, 대상블록의 픽셀값이 예측될 수 있다.

- [113] [수식1]

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

- [114] 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 그라디언트를 이용하여 산정된 샘플 오프셋이다. $shift$ 는 스케일링 인자이고, o_{offset} 은 반올림 연산을 위한 값으로서 $shift$ 의 절반이다.

- [115] **II. 인터 예측의 머지/스킵 모드**

- [116] 이하, 도 7의 예시를 이용하여, 인터 예측의 머지/스킵 모드에서 움직임벡터의 머지 후보 리스트를 구성하는 방법을 설명한다. 머지 모드를 지원하기 위해, 인터 예측부(124)는 기설정된 개수(예컨대, 6 개)의 머지 후보를 선정하여 머지 후보 리스트를 구성할 수 있다.

- [117] 도 7은 본 개시의 일 실시예에 따른, 머지/스킵 모드에서 움직임벡터의 후보를 탐색하는 과정을 나타내는 순서도이다.

- [118] 인터 예측부(124)는 공간적 머지 후보를 탐색한다(S700). 인터 예측부(124)는, 도 4의 예시된 바와 같은 주변블록들로부터 공간적 머지 후보를 탐색한다. 공간적 머지 후보는 최대 4 개까지 선정될 수 있다.

- [119] 인터 예측부(124)는 시간적 머지 후보를 탐색한다(S702). 인터 예측부(124)는, 대상블록이 위치한 현재 픽처가 아닌 참조픽처(현재블록을 예측하기 위해 사용된 참조픽처와 동일할 수도 있고 다를 수도 있음) 내의 현재블록과 동일 위치에 있는 블록(co-located block)을 시간적 머지 후보로 추가할 수 있다. 시간적 머지 후보는 1 개가 선정될 수 있다.

- [120] 인터 예측부(124)는 HMVP(History-based Motion Vector Predictor) 후보를 탐색한다(S704). 인터 예측부(124)는 이전 n (여기서, n 은 자연수) 개의 CU의 움직임벡터를 테이블에 저장한 후, 머지 후보로 이용할 수 있다. 테이블의 크기는 6이며, FiFO(First-in First Out) 방식에 따라 이전 CU의 움직임벡터를 저장한다. 이는 HMVP 후보가 최대 6 개까지 테이블에 저장됨을 나타낸다. 인터 예측부(124)는 테이블에 저장된 HMVP 후보 중, 최근의 움직임벡터들을 머지 후보로 설정할 수 있다.

- [121] 인터 예측부(124)는 PAMVP(Pairwise Average MVP) 후보를 탐색한다(S706). 인터 예측부(124)는 머지 후보 리스트에서 첫 번째 후보와 두 번째 후보의 움직임벡터 평균을 머지 후보로 설정할 수 있다.

- [122] 위의 과정(S700 내지 S706)을 모두 수행해도 머지 후보 리스트를 채울 수 없는 경우(즉, 기설정된 개수를 충족하지 못하는 경우), 인터 예측부(124)는 제로(zero) 움직임벡터를 머지 후보로 추가한다(S708).
- [123] **III. 아핀 움직임 예측(affine motion prediction)**
- [124] 전술한 바와 같은 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. 본 실시예에 따른 일 측면은 이러한 다양한 형태의 움직임을 반영할 수 있는 아핀 움직임 예측(affine motion prediction)을 제공한다.
- [125] 도 8은 본 개시의 일 실시예에 따른 아핀 움직임 예측을 설명하기 위한 예시도이다.
- [126] 아핀 움직임 예측을 위한 두 가지 타입의 모델이 존재할 수 있다. 하나는, 도 8의 (a)에서 보는 바와 같이, 현재 부호화하고자 하는 대상블록의 좌상귀(top-left corner)와 우상귀(top-right corner) 두 개의 제어점 움직임벡터들(Control Point Motion Vectors: CPMVs), 즉, 4 개의 파라미터를 이용하는 모델이다. 다른 하나는, 도 8의 (b)에서 보는 바와 같이, 대상블록의 좌상귀, 우상귀, 좌하귀(bottom-left corner)의 세 개의 제어점의 움직임벡터들, 즉, 6 개의 파라미터를 이용하는 모델이다.
- [127] 4 파라미터 아핀 모델은 수학식 2에 나타난 바와 같이 표현된다. 대상블록 내의 샘플 위치 (x,y)에서의 움직임은 수학식 2에 나타난 바에 따라 연산될 수 있다. 여기서, 대상블록의 좌상단 샘플의 위치는 (0,0)으로 가정된다.
- [128] [수식2]
- $$\begin{cases} mv_x = \frac{mv_{1x} - mv_{0x}}{W}x - \frac{mv_{1y} - mv_{0y}}{W}y + mv_{0x} \\ mv_y = \frac{mv_{1y} - mv_{0y}}{W}x + \frac{mv_{1x} - mv_{0x}}{W}y + mv_{0y} \end{cases}$$
- [129] 또한, 6 파라미터 아핀 모델은 수학식 3에 나타난 바와 같이 표현된다. 대상블록 내의 샘플 위치 (x,y)에서의 움직임은 수학식 3에 나타난 바에 따라 연산될 수 있다.
- [130] [수식3]
- $$\begin{cases} mv_x = \frac{mv_{1x} - mv_{0x}}{W}x + \frac{mv_{2x} - mv_{0x}}{H}y + mv_{0x} \\ mv_y = \frac{mv_{1y} - mv_{0y}}{W}x + \frac{mv_{2y} - mv_{0y}}{H}y + mv_{0y} \end{cases}$$
- [131] 여기서, (mv_{0x},mv_{0y})는 좌상귀 제어점의 움직임벡터, (mv_{1x},mv_{1y})는 우상귀 제어점의 움직임벡터, (mv_{2x},mv_{2y})는 좌하귀 제어점의 움직임벡터이다. W는 대상블록의 가로 길이이고, H는 대상블록의 세로 길이이다.

- [132] 아핀 움직임 예측은, 대상블록 내의 각 샘플마다 수학식 2 또는 수학식 3에 따라 연산된 움직임벡터를 이용하여 수행될 수 있다.
- [133] 대안적으로, 연산의 복잡도를 줄이기 위해, 도 9에 예시된 바와 같이, 대상블록으로부터 분할된 서브블록 단위로 수행될 수도 있다. 예를 들어, 서브블록의 크기는 4×4 일 수 있고, 2×2 또는 8×8 일 수도 있다. 이하에서는, 대상블록에 대해 4×4 서브블록 단위로 아핀 움직임 예측을 수행하는 것을 예로 들어 설명한다. 이 예시는 설명의 편의를 위한 것일 뿐, 본 발명의 범위를 제한하는 것은 아니다.
- [134] 서브블록 단위 아핀 움직임 예측에서, 각 서브블록의 움직임벡터(아핀 움직임벡터)는 수학식 2 또는 3의 (x, y) 에 각 서브블록의 중심 위치를 대입함으로써 연산된다. 여기서, 중심 위치는 서브블록의 실제 중심점(center point)일 수도 있고 그 중심점의 우하단 샘플 위치일 수도 있다. 예컨대, 좌하단 샘플의 좌표가 $(0, 0)$ 인 4×4 서브블록의 경우, 서브블록의 중심 위치는 $(1.5, 1.5)$ 일 수 있고 또는 $(2, 2)$ 일 수 있다. 각 서브블록에 대한 예측블록은 해당 서브블록의 아핀 움직임벡터 (mv_x, mv_y) 를 이용하여 생성된다.
- [135] 움직임벡터 (mv_x, mv_y) 는 1/16 샘플 정확도(precision)를 갖도록 설정될 수 있다. 이 경우, 수학식 2 또는 3에 따라 산출된 움직임벡터 (mv_x, mv_y) 는 1/16 샘플 단위로 반올림될 수 있다.
- [136] 아핀 움직임 예측은 루마(luma) 성분뿐만 아니라 크로마(chroma) 성분에도 대해서도 수행될 수 있다. 4:2:0 비디오 포맷의 경우, 루마 성분에 대해 4×4 서브블록 단위의 아핀 움직임 예측이 수행되었다면, 크로마 성분에 대한 아핀 움직임 예측은 2×2 서브블록 단위로 수행될 수 있다. 크로마 성분의 각 서브블록의 움직임벡터 (mv_x, mv_y) 는 대응하는 루마 성분의 움직임벡터로부터 유도될 수 있다. 대안적으로, 크로마 성분의 아핀 움직임 예측을 위한 서브블록의 크기는 루마 성분의 경우와 동일할 수도 있다. 루마 성분에 대해 4×4 서브블록 단위로 아핀 움직임 예측이 수행되었다면, 크로마 성분에 대해서도 4×4 서브블록 크기로 아핀 움직임 예측이 수행된다. 이 경우, 크로마 성분에 대한 4×4 크기의 서브블록은 루마 성분에 대한 네 개의 4×4 서브블록에 대응하므로, 크로마 성분의 서브블록에 대한 움직임벡터 (mv_x, mv_y) 는 대응하는 루마 성분의 네 개의 서브블록의 움직임벡터들의 평균을 연산하여 산출될 수 있다.
- [137] 영상 부호화 장치는 인트라 예측, 인터 예측(병진 움직임 예측), 아핀 움직임 예측 등을 수행하고, RD(rate-distortion) 코스트(cost)를 연산하여 최적의 예측 방법을 선택한다. 아핀 움직임 예측을 수행하기 위해, 영상 부호화 장치의 인터 예측부(124)는 두 가지 타입의 모델 중 어느 타입을 사용할지 여부를 결정하고, 결정된 타입에 따라 두 개 또는 세 개의 제어점들을 결정한다. 인터 예측부(124)는 제어점들의 움직임벡터들을 이용하여 대상블록 내의 4×4 서브블록들 각각에 대한 움직임벡터 (mv_x, mv_y) 를 연산한다. 그리고, 각 서브블록의 움직임벡터 (mv_x, mv_y) 를 이용하여 서브블록 단위로 참조픽처 내에서

움직임 보상을 수행함으로써, 대상블록 내의 각 서브블록에 대한 예측블록을 생성한다.

- [138] 영상 부호화 장치의 엔트로피 부호화부(155)는 대상블록에 아핀 움직임 예측이 적용되었는지 여부를 나타내는 플래그, 아핀 모델의 타입을 나타내는 타입 정보, 및 각 제어점의 움직임벡터를 나타내는 움직임 정보 등을 포함하는 아핀 관련 선택스 요소들을 부호화하여 영상 복호화 장치로 전달한다. 타입 정보와 제어점들의 움직임 정보는 아핀 움직임 예측이 수행되는 경우에 시그널링될 수 있고, 제어점들의 움직임벡터들은 타입 정보에 따라 결정된 개수만큼 시그널링될 수 있다.
- [139] 영상 복호화 장치는 시그널링된 선택스들을 이용하여 아핀 모델의 타입과 제어점 움직임벡터(control point motion vector)들을 결정하고, 수학식 2 또는 3를 이용하여 대상블록 내의 각 4×4 서브블록에 대한 움직임벡터 (mv_x, mv_y)를 연산한다. 만약, 대상블록의 아핀 움직임벡터에 대한 움직임벡터 해상도 정보가 시그널링된 경우, 움직임벡터 (mv_x, mv_y)는 반올림 등의 연산을 이용하여 움직임벡터 해상도 정보에 의해 식별되는 정밀도로 수정된다.
- [140] 영상 복호화 장치는 각 서브블록에 대한 움직임벡터 (mv_x, mv_y)를 이용하여 참조픽처 내에서 움직임 보상을 수행함으로써, 각 서브블록에 대한 예측블록을 생성한다.
- [141] 제어점들의 움직임벡터들을 부호화하기 위해 소요되는 비트량을 줄이기 위해, 전술한 바와 같은 일반적인 인트라 예측(병진 움직임 예측) 방식이 적용될 수 있다.
- [142] 하나의 예시로서, 머지 모드의 경우, 영상 부호화 장치의 인터 예측부(124)는 기정의된 개수(예컨대, 5 개)의 아핀 머지 후보 리스트를 구성한다. 먼저, 영상 부호화 장치의 인터 예측부(124)는 대상블록의 주변블록들로부터 상속(inherited) 아핀 머지 후보를 유도한다. 예컨대, 도 4에 도시된 대상블록의 주변샘플들(A0, A1, B0, B1, B2)로부터 기정의된 개수의 상속 아핀 머지 후보들을 유도함으로써, 머지 후보 리스트를 생성한다. 후보 리스트에 포함되는 상속 아핀 머지 후보들 각각은 두 개 또는 세 개의 CPMVs의 쌍에 대응한다.
- [143] 인터 예측부(124)는 대상블록의 주변블록들 중 아핀 모드로 예측된 주변블록의 제어점 움직임벡터들로부터 상속 아핀 머지 후보를 유도한다. 일부 실시예에서, 아핀 모드로 예측된 주변블록들로부터 유도되는 머지 후보의 개수는 제한될 수 있다. 예컨대, 인터 예측부(124)는 A0와 A1 중에서 하나, 그리고 B0, B1, 및 B2 중에서 하나, 총 두 개의 상속 아핀 머지 후보를 아핀 모드로 예측된 주변블록으로부터 유도할 수 있다. 우선순위는 A0, A1의 순서, 그리고 B0, B1, B2의 순서일 수 있다.
- [144] 한편, 머지 후보의 총 개수가 3 개 이상인 경우, 인터 예측부(124)는 주변블록의 병진 움직임벡터(translational motion vector)로부터 부족한 개수만큼의 조합(constructed) 아핀 머지 후보들을 유도할 수 있다.

- [145] 도 10은 주변블록의 병진 움직임벡터로부터 아핀 움직임 예측을 위한 조합 아핀 머지 후보들을 유도하는 방법을 설명하기 위한 예시도이다.
- [146] 인터 예측부(124)는 주변블록 그룹 {B2, B3, A2}, 주변블록 그룹 {B1, B0}, 주변블록 그룹{A1, A0}로부터 각각 하나씩의 제어점 움직임벡터 CPMV1, CPMV2, CPMV3을 유도한다. 하나의 예시로서, 각 주변블록 그룹 내에서의 우선 순위는, B2, B3, A2의 순서, B1, B0의 순서, 그리고 A1, A0의 순서일 수 있다. 또한, 참조픽처 내의 동일 위치 블록(collocated block) T로부터 또 하나의 제어점 움직임벡터 CPMV4를 유도한다. 인터 예측부(124)는 네 개의 제어점 움직임벡터 중에서 두 개 또는 세 개의 제어점 움직임벡터들을 조합하여 부족한 개수만큼의 조합 아핀 머지 후보를 생성한다. 조합의 우선순위는 아래와 같다. 각 그룹 내의 요소들은 좌상귀, 우상귀, 좌하귀 제어점 움직임벡터의 순서로 나열되어 있다.
- [147] {CPMV1, CPMV2, CPMV3}, {CPMV1, CPMV2, CPMV4}, {CPMV1, CPMV3, CPMV4},
- [148] {CPMV2, CPMV3, CPMV4}, {CPMV1, CPMV2}, {CPMV1, CPMV3}
- [149] 상속 아핀 머지 후보 및 조합 아핀 머지 후보를 이용하여 머지 후보 리스트를 채울 수 없는 경우, 인터 예측부(124)는 제로 움직임벡터를 후보로 추가할 수 있다.
- [150] 인터 예측부(124)는 머지 후보 리스트로부터 머지 후보를 선택하여 대상블록에 대한 아핀 움직임 예측을 수행한다. 선택된 후보가 두 개의 제어점 움직임벡터로 이루어진 경우 4 파라미터 모델을 이용하여 아핀 움직임 예측을 수행한다. 반면, 선택된 후보가 세 개의 제어점 움직임벡터로 구성된 경우 6 파라미터 모델을 이용하여 아핀 움직임 예측을 수행한다. 영상 부호화 장치의 엔트로피 부호화부(155)는 머지 후보 리스트 내 머지 후보들 중 선택된 머지 후보를 나타내는 인덱스 정보를 부호화하여 영상 복호화 장치로 시그널링한다.
- [151] 영상 복호화 장치의 엔트로피 복호화부(510)은 영상 부호화 장치로부터 시그널링된 인덱스 정보를 복호화한다. 영상 복호화 장치의 인터 예측부(544)는 영상 부호화 장치와 동일한 방식으로 머지 후보 리스트를 구성하고, 인덱스 정보에 의해 지시되는 머지 후보에 대응하는 제어점 움직임벡터들을 이용하여 아핀 움직임 예측을 수행한다.
- [152] 다른 예시로서, AMVP 모드인 경우, 영상 부호화 장치의 인터 예측부(124)는 대상블록에 대한 아핀 모델의 타입과 제어점 움직임벡터들을 결정한다. 그리고, 대상블록의 실제 제어점 움직임벡터들과 각 제어점의 예측 움직임벡터들 간의 차이인 차분 움직임벡터(motion vector difference)를 연산하고 각 제어점에 대한 차분 움직임벡터를 부호화하여 전송한다. 이를 위해, 영상 부호화 장치의 인터 예측부(124)는 기정의된 개수(예컨대, 2 개)의 아핀 AMVP 후보 리스트를 구성한다. 대상블록이 4 파라미터 타입인 경우, 리스트에 포함된 후보들은 각각 두 개의 제어점 움직임벡터의 쌍으로 이루어진다. 반면, 대상블록이 6 파라미터 타입인 경우, 리스트에 포함된 후보들은 각각 세 개의 제어점 움직임벡터들의

쌍으로 이루어진다.

- [153] 이하, 도 11의 예시를 이용하여, 인터 예측의 아핀 AMVP 모드에서 후보 리스트를 구성하는 방법을 설명한다. 아핀 AMVP 후보 리스트는 전술한 아핀 머지 후보 리스트를 구성하는 방법과 유사하게 유도될 수 있다.
- [154] 도 11은 본 개시의 일 실시예에 따른, 아핀 AMVP 모드에서 아핀 AMVP 후보를 탐색하는 과정을 나타내는 순서도이다.
- [155] 인터 예측부(124)는 상속 아핀 AMVP 후보의 참조픽처와 현재블록의 참조픽처가 동일한지를 확인한다(S1100). 여기서, 상속 아핀 AMVP 후보는, 전술한 아핀 머지 모드에서와 같이, 도 4에 도시된 대상블록의 주변블록들(A0, A1, B0, B1, B2) 중 아핀 모드로 예측된 블록일 수 있다.
- [156] 상속 아핀 AMVP 후보의 참조픽처와 현재블록의 참조픽처가 동일한 경우, 인터 예측부(124)는 해당되는 상속 아핀 AMVP 후보를 추가한다(S1102).
- [157] 상속 아핀 AMVP 후보의 참조픽처와 현재블록의 참조픽처가 동일하지 않은 경우, 인터 예측부(124)는 조합 아핀 AMVP 후보의 모든 CPMV의 참조픽처와 현재블록의 참조픽처가 동일한지를 확인한다(S1104). 여기서, 조합 아핀 AMVP 후보의 모든 CPMV는, 전술한 아핀 머지 모드에서와 같이, 도 10에 도시된 주변샘플들의 움직임벡터로부터 유도될 수 있다.
- [158] 조합 아핀 AMVP 후보의 모든 CPMV의 참조픽처와 현재블록의 참조픽처가 동일한 경우, 인터 예측부(124)는 해당되는 조합 아핀 AMVP 후보를 추가한다(S1106).
- [159] 이때, 대상블록의 아핀 모델 타입이 고려되어야 한다. 영상 부호화 장치는 대상블록의 아핀 모델 타입이 4 파라미터 타입인 경우, 주변블록의 아핀 모델을 이용하여 두 개의 제어점 움직임벡터들(대상블록의 좌상귀 및 우상귀 제어점 움직임벡터들)을 유도한다. 대상블록의 아핀 모델 타입이 6 파라미터 타입인 경우, 주변블록의 아핀 모델을 이용하여 세 개의 제어점 움직임벡터들(대상블록의 좌상귀, 우상귀 및 좌하귀 제어점 움직임벡터들)을 유도한다.
- [160] 주변블록이 4 파라미터 타입인 경우, 주변블록의 두 개의 제어점 움직임벡터를 이용하여, 대상블록의 아핀 모델 타입에 따라 두 개 또는 세 개의 제어점 움직임벡터들이 예측된다. 예컨대, 수학식 2로 표현된 주변블록의 아핀 모델이 이용될 수 있다. 수학식 2에서 (mv_{0x}, mv_{0y}) 와 (mv_{1x}, mv_{1y}) 는 각각 주변블록의 좌상귀(top-left corner) 및 우상귀 제어점 움직임벡터로 대체된다. 대상블록의 각 제어점에 대한 예측 움직임벡터는, (x, y) 에 대상블록의 해당 제어점의 위치와 주변블록의 좌상귀 위치 간의 차이를 입력함으로써, 유도될 수 있다.
- [161] 주변블록이 6 파라미터 타입인 경우, 주변블록의 세 개의 제어점 움직임벡터를 이용하여, 대상블록의 아핀 모델 타입에 따라 두 개 또는 세 개의 제어점 움직임벡터들이 예측된다. 예컨대, 수학식 3으로 표현된 주변블록의 아핀 모델이 이용될 수 있다. 수학식 3에서 (mv_{0x}, mv_{0y}) , (mv_{1x}, mv_{1y}) , 및 (mv_{2x}, mv_{2y}) 는

각각 주변블록의 좌상귀, 우상귀 및 좌하귀의 제어점 움직임벡터로 대체된다. 대상블록의 각 제어점에 대한 예측 움직임벡터는, (x,y) 에 대상블록의 해당 제어점의 위치와 주변블록의 좌상귀 위치 간의 차이를 입력함으로써, 유도될 수 있다.

- [162] 모든 CPMV의 참조픽처와 현재블록의 참조픽처가 동일하지 않은 경우, 인터 예측부(124)는 아핀 AMVP 후보로서 병진 움직임벡터를 추가한다(S1108). 병진 움직임벡터는, mv_0, mv_1, mv_2 의 순서대로, 현재블록의 CPMV를 예측하기 위해 이용될 수 있다.
- [163] 위의 과정(S1100 내지 S1108)을 모두 수행해도 후보 리스트를 채울 수 없는 경우(즉, 기설정된 개수를 충족하지 못하는 경우), 인터 예측부(124)는 제로 움직임벡터를 아핀 AMVP 후보로 추가한다(S1110).
- [164] 영상 부호화 장치의 인터 예측부(124)는 아핀 AMVP 리스트로부터 하나의 후보를 선택하고, 실제 각 제어점의 움직임벡터들과 선택된 후보의 대응하는 제어점의 예측움직임벡터 간의 차분 움직임벡터를 생성한다. 영상 부호화 장치의 엔트로피 부호화부(155)는 대상블록의 아핀 모델 타입을 나타내는 타입 정보, 아핀 AMVP 리스트 내의 후보들 중에서 선택된 후보를 지시하는 인덱스 정보, 및 각 제어점에 대응하는 차분 움직임벡터를 부호화하여 영상 복호화 장치로 전달한다.
- [165] 영상 복호화 장치의 인터 예측부(544)는 영상 부호화 장치로부터 시그널링된 정보를 이용하여 아핀 모델 타입을 결정하고, 각 제어점의 차분 움직임벡터를 생성한다. 그리고, 영상 부호화 장치와 동일한 방식으로 아핀 AMVP 리스트를 생성하고, 아핀 AMVP 리스트 내에서 시그널링된 인덱스 정보에 의해 지시되는 후보를 선택한다. 영상 복호화 장치의 인터 예측부(544)는 선택된 후보의 각 제어점의 예측 움직임벡터와 대응하는 차분 움직임벡터를 가산하여 각 제어점의 움직임벡터를 산출한다.
- [166] **IV. 가상 참조프레임 생성 및 인터 예측 모드에서의 이용**
- [167] 가변계수 딥러닝에 기초하는 비디오 예측은 이중 루프 부호화(double-loop coding)로서, 다음과 같은 두 단계의 과정으로 구성된다. 두 단계는, 가변계수 딥러닝 모델을 기반으로 주어진 비디오를 분석하여 딥러닝 모델의 파라미터를 취득하는 단계, 및 취득한 딥러닝 모델의 파라미터를 비디오 예측 과정에 응용하되, 부호화된 비트스트림과 딥러닝 모델의 파라미터를 전송하여 영상 부호화 및 복호화에 사용하는 단계를 포함한다.
- [168] 첫 번째 단계에서는, 비디오의 움직임 예측에 관련된 딥러닝 모델의 파라미터가 개별 비디오에 적응되도록, 해당 개별 비디오를 학습용 데이터로 활용하여 영상 부호화 장치가 딥러닝 모델을 트레이닝한다. 예를 들어, 비디오의 움직임 예측 신호의 보정, 움직임벡터의 정밀화, 움직임벡터의 해상도 변경, 및 참조 프레임의 개선 중 하나 또는 이상의 목적을 달성하기 위해 딥러닝 모델이 사용될 수 있다.

- [169] 딥러닝 모델은, 움직임 예측, 움직임 보간(interpolation), 움직임 보외(extrapolation) 등의 기능을 제공하는 고정계수(fixed coefficient)를 포함하는 레이어들(이하, '고정계수 네트워크')과 가변계수를 포함하는 하나 이상의 콘볼루션 레이어(convolution layer)의 결합일 수 있다. 또는, 가변계수만을 포함하는 콘볼루션 레이어들(이하, '가변계수 네트워크')이 딥러닝 모델로 이용될 수 있다. 어느 형태의 딥러닝 모델이든, 전술한 바와 같이, 개별 비디오에 적응되도록 트레이닝된다.
- [170] 두 번째 과정에서는, 트레이닝된 딥러닝 모델을 이용하여, 영상 부호화 장치 및 영상 복호화 장치는 움직임 예측 신호의 개선, 움직임벡터의 정밀화와 초해상화, 참조프레임의 PSNR(Peak Signal to Noise Ratio) 개선 등을 수행한다. 이러한 기능을 수행하는 딥러닝 모델의 활용은 블록 단위 또는 비디오 시퀀스 단위에서 제어될 수 있다. 예를 들어, 비디오의 시퀀스 파라미터 세트(SPS), 비디오 파라미터 세트(VPS), 픽처 파라미터 세트(PPS), 또는 그 이하의 서브픽처나 타일에서 딥러닝 모델을 사용 가능하도록 하는 플래그가 이용될 수 있다. 또는 비디오의 블록 단위, 즉, CTU 또는 CU 단위에서 딥러닝 모델을 사용 가능하도록 하는 플래그가 이용될 수 있다. 영상 부호화 장치는 부호화된 비트스트림과 딥러닝 모델의 가변계수 값들을 영상 복호화 장치에게 전송한다.
- [171] 첫 번째 단계에서 딥러닝 모델의 가변계수 네트워크를 트레이닝하는데 시간이 소요될 수 있으므로, 가변계수 딥러닝은 저지연 부호화(low delay coding)가 요구되는 분야, 예컨대 실시간 화상 회의, 실시간 스트리밍 서비스 등에는 적용이 어려울 수 있다. 하지만, 저장된 비디오를 이용하여 첫 번째 단계를 사전에 수행하는 것이 가능하므로, 임의 접근 부호화(random access coding)를 이용하는 분야에는 가변계수 딥러닝이 적용이 용이할 수 있다.
- [172] 이하, 딥러닝 모델로서 가변계수 네트워크를 포함하는 비디오 보간 모델(이하, '가변계수 보간 모델' 또는 '보간 모델')을 이용하는 실시예를 설명한다. 전술한 바와 같이, 보간 모델은 고정계수 네트워크도 포함할 수 있다. 이러한 보간 모델을 이용하여 참조 프레임들로부터 가상(virtual) 참조프레임을 생성하고, 가상 참조프레임을 이용하여 비디오를 부호화하는 과정, 및 부호화된 비트스트림과 가변계수 딥러닝 모델을 전송하여 비디오를 복호화하는 과정을 설명한다. 하지만 전술한 사례에 한정하지 않고, 움직임 예측 신호의 개선 및 움직임벡터의 정밀화에도 가변계수 딥러닝 모델을 이용함으로써, 본 발명의 적용 분야가 확장될 수 있다.
- [173] 영상 부호화 및 복호화 장치는, 보간 모델이 생성한 가상 참조프레임을 참조픽처 리스트에 포함시킨 후, 머지 모드 및 AMVP 모드에서의 움직임 예측에 사용함으로써, 부호화 이득을 얻을 수 있다. 이때, 가상 참조프레임이 원본 프레임과 유사할수록 부호화 효율이 증가하는 것으로 알려져 있다(비특허문헌 1 참조). 가상 참조프레임의 생성 시, 입력 프레임에 과적합(overfitting)되도록 적응적으로 학습된 가변계수 보간 모델을 이용함으로써, 영상 부호화 및 복호화

장치는 더욱 용이하게 원본 프레임과 유사한 프레임을 생성할 수 있다. 따라서, 본 개시에서는 도 12 내지 도 14에 예시된 바와 같은, 가변계수 보간 모델을 이용하여 가상 참조프레임을 생성하고, 가상 참조프레임을 이용하여 부호화 효율을 증대시키는 방법을 설명한다.

[174] 도 12는 본 개시의 일 실시예에 따른 보간 모델을 포함하는 영상 부호화 장치를 나타내는 블록도이다.

[175] 도 12에 예시된 바와 같이, 영상 부호화 장치는, 메모리(190)에 저장된 이미 복호화된 주변 참조프레임으로부터 가상 참조프레임을 생성하는 가상프레임 생성부(1210)를 포함할 수 있다. 여기서, 가상프레임 생성부(1210)는 가변계수 보간 모델을 포함한다. 보간 모델의 동작은 추후 설명하기로 한다.

[176] 도 13은 본 개시의 일 실시예에 따른 보간 모델을 포함하는 영상 복호화 장치를 나타내는 블록도이다.

[177] 도 13에 예시된 바와 같이, 영상 복호화 장치는, 메모리(570)에 저장된 이미 복호화된 주변 참조프레임으로부터 가상 참조프레임을 생성하는 가상프레임 생성부(1310)를 포함할 수 있다. 여기서, 가상프레임 생성부(1310)는 가변계수 보간 모델을 포함한다. 보간 모델의 동작은 추후 설명하기로 한다.

[178] 도 14는 본 개시의 일 실시예에 따른, 영상 부호화 장치 및 영상 복호화 장치에서 이용하는 보간 모델의 동작을 나타내는 예시도이다.

[179] 영상 부호화 장치의 가상프레임 생성부(1210)는, 보간 모델을 이용하여 복호화된 주변 참조 프레임들 X_{t-1} 과 X_{t+1} 로부터 원본 프레임 X_t 와 매우 유사하도록 추론된 프레임 $X_{t,pred}$ 을 생성한다. 추론된 프레임 $X_{t,pred}$ 은 가상 참조프레임으로 사용된다. 원본 프레임과 매우 유사한 프레임 $X_{t,pred}$ 을 생성할 수 있도록 적응적으로 트레이닝된 보간 모델을 이용하면, 영상 부호화 장치는 높은 유사도의 가상 참조프레임을 생성할 수 있다. 따라서, 영상 부호화 장치는 부호화하고자 하는 입력 프레임을 적응적으로 학습한 가변계수 보간 모델을 이용할 수 있다. 이때, 영상 부호화 장치는 트레이닝된 보간 모델의 가변계수 값들을 영상 복호화 장치로 전송한다.

[180] 보간 모델이 포함하는 가변계수의 개수가 많을수록, 입력 프레임을 타겟 프레임에 과적합시키는 것이 용이하여지도록 가변계수 보간 모델을 트레이닝시킬 수 있다. 그러나 보간 모델을 이용하여 원본 프레임과 유사한 프레임을 생성하기 위해 가변계수의 개수가 많아진다면, 전송할 비트 수의 증가를 초래하여 부호화 효율을 감소시킬 수 있다. 따라서, 가변계수 보간 모델로는 가장 적은 수의 가변계수를 사용하면서도 부호화 효율이 좋은 모델이 고려되어야 한다.

[181] 한편, 영상 복호화 장치의 가상프레임 생성부(1310)는, 전송한 바와 같이, 영상 부호화 장치로부터 전송된 가변계수 값들을 이용하여 보간 모델의 가변계수를 설정한다. 보간 모델은 이미 복호화된 주변 참조 프레임들 X_{t-1} 과 X_{t+1} 로부터 원본 프레임 X_t 와 매우 유사한 프레임 $X_{t,pred}$ 을 추론한다. 추론된 프레임 $X_{t,pred}$ 은 가상

참조프레임으로 사용된다.

[182] 이하, 영상 부호화 장치에서 수행하는 보간 모델의 학습을 설명한다. 전술한 바와 같이, 보간 모델의 학습은 가변계수 딥러닝에 기초하는 비디오 예측의 첫 번째 단계에 해당한다.

[183] 보간 모델이 가상 참조프레임을 생성하는 것을 학습하도록 하기 위하여, 영상 부호화 장치는 트레이닝부(미도시)를 추가로 구비하거나, 외부의 트레이닝부와 연동되는 형태로 구현될 수 있다.

[184] 전술한 바와 같은, 이미 복호화된 주변 참조 프레임들 X_{t-1} 과 X_{t+1} 로부터 추론 프레임 $X_{t,pred}$ 을 생성하는 보간 모델 f 의 동작은 수학식 4와 같이 나타낼 수 있다.

[185] [수식4]

$$X_{t,pred} = f(X_{t-1}, X_{t+1}, \theta_t)$$

[186] 여기서, θ_t 는 가변계수에 해당하는 파라미터이다.

[187] 트레이닝부는 부호화하고자 하는 프레임에 과적합되도록 파라미터 θ_t 를 업데이트할 수 있다. 이때, 보간 모델을 트레이닝하기 위한 손실함수(loss function)는 수학식 5에 나타난 바와 같이 정의된다.

[188] [수식5]

$$L = \frac{1}{H \times W} \sum_i^{H \times W} |X_{t,i} - X_{t,i,pred}|^2$$

[189] 여기서, H 와 W 각각은 프레임의 세로 및 가로 픽셀 수를 나타낸다. $X_{t,i,pred}$ 는 추론 프레임 $X_{t,pred}$ 의 픽셀 단위의 값을 나타내고, $X_{t,i}$ 는 원본 프레임의 픽셀 단위의 값을 나타낸다. 전술한 바와 같이, 영상 부호화 장치는 프레임별로 트레이닝이 완료된 가변계수 θ_t 의 값들을 영상 복호화 장치로 전송한다. 영상 복호화 장치는, 전송 받은 가변계수 θ_t 의 값들을 이용하여 영상 부호화 장치에 포함된 것과 동일한 보간 모델의 가변계수 θ_t 를 설정한 후, 수학식 4에 나타난 바와 동일한 방법으로 가상 참조프레임을 생성할 수 있다.

[190] 도 15는 본 개시의 실시예에 따른 가변계수 보간 모델의 예시들이다.

[191] 도 15에 예시된 보간 모델들은, 영상 처리에 적합한 CNN(Convolutional Neural Network) 기반 딥러닝 모델일 수 있다.

[192] 본 개시에 따른 보간 모델은, 도 15의 (a)에 나타난 바와 같이, 주변 참조 프레임들로부터 추론 프레임을 바로 생성하는 딥러닝 모델일 수 있다.

[193] 본 개시에 따른 다른 실시예에 있어서, 보간 모델은, 도 15의 (b)에 나타난 바와 같이, 마스크를 이용하는 딥러닝 모델일 수 있다(비특허문헌 1 참조). 이때, 보간 모델은 참조 프레임들로부터 마스크를 생성한 후, 마스크를 이용하여 이미 복호화된 주변 참조 프레임들 X_{t-1} 과 X_{t+1} 을 가중합함으로써 추론 프레임 $X_{t,pred}$ 를 생성할 수 있다.

[194] 본 개시에 따른 또다른 실시예에 있어서, 보간 모델은, 도 15의 (c)에 나타난

바와 같이, 옵티컬 플로우를 이용하는 딥러닝 모델일 수 있다. BDOF와 관련하여 기술한 바와 같이, OF를 이용하는 딥러닝 모델은 시간 정규화를 가정하여, 추론 프레임 $X_{t,pred}$ 를 생성할 수 있다. 보간 모델은 참조 프레임들로부터 OF를 생성한 후, OF를 기반으로 참조 프레임들을 와핑(warping)함으로써 추론 프레임 $X_{t,pred}$ 를 생성할 수 있다.

- [195] 한편, 도 15에 예시된 어느 보간 모델이든 수학식 5에 나타낸 손실함수를 기반으로 가변계수가 업데이트될 수 있다.
- [196] 가변계수 보간 모델로는 도 15에 예시된 어느 딥러닝 모델이든 사용될 수 있다. 하지만, 가변계수 값들이 전송되어야 하므로, 가장 간략한 모델을 사용해야만 부호화 효율 향상시킬 수 있다. 반면, 가변계수의 개수를 적게 사용할수록 과적합이 어렵다는 문제가 있다.
- [197] 전술한 바와 같이, 보간 모델에서, 가변계수를 포함하는 부분을 가변계수 네트워크로 나타낸다.
- [198] 도 16은 본 개시의 일 실시예에 따른 고정계수 네트워크를 더 포함하는 보간 모델을 나타내는 예시도이다.
- [199] 따라서, 도 16의 예시와 같이, 고정계수 네트워크를 더 포함하는 보간 모델은, 입력 프레임들로부터 부호화하고자 하는 프레임과 유사한 프레임인 $X_{t-1 \rightarrow t}$ 와 $X_{t+1 \rightarrow t}$ 를 중간 출력으로 생성한다. 이때, 고정계수 네트워크는 전체 데이터에 기초하여 영상 부호화 장치에서 사전에 트레이닝될 수 있다. 또한, 영상 부호화 장치와 영상 복호화 장치에 동일한 고정계수 값들에 저장한 채로, 영상 부호화 장치와 영상 복호화 장치는 사전 약속에 따라 고정계수 네트워크를 설정할 수 있다.
- [200] 중간 출력 $X_{t-1 \rightarrow t}$ 와 $X_{t+1 \rightarrow t}$ 를 이용함으로써 보간 모델은 더 적은 개수의 가변계수를 이용하여 좋은 성능을 제공할 수 있다. 또한, 영상 부호화 장치는 가변계수 네트워크에 적용되는 가변계수 값들을 영상 복호화 장치로 전송한다.
- [201] 도 16의 예시와 같이, 고정계수 네트워크의 뒤에 가변계수 네트워크를 결합하여 부호화하고자 하는 프레임에 과적합하는 것이 가능하나, 다른 실시예로서, 고정계수 네트워크가 없이 가변 계수 네트워크만이 이용될 수 있다. 이때, 가변계수 네트워크만을 포함하는 보간 모델은, student-teacher network에 기반하는 학습 방법을 이용하여, 기존의 고정계수 네트워크를 포함하는 보간 모델의 축약 형태로 구현될 수 있다.
- [202] 이하, 도 17의 예시를 이용하여, 가변계수의 데이터량을 감소시키는 방안을 설명한다.
- [203] 도 17은 본 개시의 일 실시예에 따른, 가변계수를 전송하는 방법을 나타내는 예시도이다.
- [204] 영상 부호화 장치는 프레임 단위 또는 GOP(Group of Picture) 단위로 가변계수를 전송할 수 있다. 도 17의 (a)에 예시된 바와 같이, 모든 t 단위, 즉 매 프레임마다 가변계수 θ_t 를 전송하는 것이 가능하다. 하지만, 영상 부호화 장치는

GOP 단위 또는 복수의 GOP 단위로 가변계수를 전송함으로써 가변계수의 데이터량을 감소시킬 수 있다.

- [205] 예컨대, 도 17의 (b)에 예시된 바와 같이 GOP의 크기가 8인 경우, 영상 부호화 장치는 GOP 단위로 가변계수를 전송함으로써 4 장의 프레임에 대응되는 가변계수를 전송할 수 있다. 또한, 영상 부호화 장치는 2 개의 GOP, 3 개의 GOP 단위 등으로 전송 주기를 확장할 수 있다. 이때, 입력 프레임을 배치(batch)로 누적한 채로 보간 모델을 한번에 학습할 수 있는데, 트레이닝부는 수학적 식 6과 같이 시간 t 를 고려하는 손실함수를 사용할 수 있다.

- [206] [수식6]

$$L = \frac{1}{H \times W \times T} \sum_t \sum_i^{H \times W} |X_{t,i} - X_{t,i,pred}|^2$$

- [207] 여기서, T 는 학습용 배치에 포함되는 프레임의 개수를 나타낸다.

- [208] 한편, 트레이닝 과정에서 기본적으로 부동소수점(floating point) 32 비트 형식의 가변계수 θ_i 가 이용되나, 영상 부호화 장치는 16 비트, 8 비트, 또는 더 적은 비트수를 갖는 소수점 형식으로 가변계수 값들을 변경한 후, 영상 복호화 장치로 전송할 수 있다.

- [209] 본 개시에 따른 다른 실시예에 있어서, 영상 부호화 장치는 가변계수 θ_i 를 양자화한 후, 영상 복호화 장치로 전송할 수 있다. 예컨대, 영상 부호화 장치는 가변계수 θ_i 의 범위를 균일한 구간으로 분할하여 양자화할 수 있다.

- [210] 본 개시에 따른 다른 실시예에 있어서, 프레임 또는 GOP별로 가변계수 값들을 직접 전송하는 대신, 영상 부호화 장치는 기설정된 개수의 가변계수 세트를 전송할 수 있다. 또는, 영상 부호화 장치와 영상 복호화 장치에 동일한 가변계수 세트를 저장한 채로, 영상 부호화 장치는, 가변계수 세트를 지정하기 위해 기정의된 인덱스만을 전송할 수 있다. 여기서, 가변계수 세트는, 비디오 시퀀스의 신호 특성, 비디오 콘텐츠의 종류, 양자화 파라미터 등에 따라, 보간 모델의 학습 과정에서 사전에 결정될 수 있다.

- [211] 이하, 도 18 및 도 19의 순서도를 이용하여, 보간 모델에 의해 생성된 가상 참조프레임을 인터 예측에 이용하는 과정을 설명한다.

- [212] 도 18은 본 개시의 일 실시예에 따른, 인터 예측의 머지/스킵 모드에서 가상 참조프레임을 이용하는 과정에 대한 순서도를 나타낸다.

- [213] 인터 예측부(124)는 공간적 머지 후보를 탐색한다(S1800). 인터 예측부(124)는, 도 4의 예시된 바와 같은 주변블록들로부터 공간적 머지 후보를 탐색한다. 이때, 주변블록의 참조픽처는 가상 참조프레임 또는 기존의 참조 프레임일 수 있다.

- [214] 인터 예측부(124)는 가상 참조프레임을 고려하여 시간적 머지 후보를 탐색한다(S1802).

- [215] 인터 예측부(124)는 현재블록과 동일 위치에 있는 블록이 존재하는 픽처와 시간적 머지 후보의 참조 프레임이 가상 참조프레임인지 또는 기존의 참조

프레임인지 확인한다.

- [216] 현재블록과 동일 위치에 있는 블록이 존재하는 픽처와 시간적 머지 후보의 참조 프레임이 모두 가상 참조프레임인 경우, 인터 예측부(124)는 제로 움직임벡터를 후보로 설정하고, 현재 픽처의 가상 참조프레임을 스킵 또는 머지 모드에서 사용한다.
- [217] 현재블록과 동일 위치에 있는 블록이 존재하는 픽처가 가상 참조프레임이고 시간적 머지 후보의 참조 프레임은 기존의 참조 프레임인 경우, 인터 예측부(124)는 제로 움직임벡터를 후보로 설정하고, 현재 픽처의 가상 참조프레임을 스킵 또는 머지 모드에서 사용한다. 또는, 인터 예측부(124)는 해당 후보를 시간적 머지 후보로 설정하지 않을 수 있다.
- [218] 반면, 현재블록과 동일 위치에 있는 블록이 존재하는 픽처가 기존의 참조 프레임이고 시간적 머지 후보의 참조 프레임은 가상 참조프레임인 경우, 인터 예측부(124)는 제로 움직임벡터를 후보로 설정하고, 현재 픽처의 가상 참조프레임을 스킵 또는 머지 모드에서 사용한다. 또는, 인터 예측부(124)는 해당 후보를 시간적 머지 후보로 설정하지 않을 수 있다.
- [219] 인터 예측부(124)는 HMVP 후보를 탐색한다(S1804). 인터 예측부(124)는, 이전 n 개의 CU의 움직임벡터를 테이블에 저장 시, 가상 참조프레임을 참조하여 부호화한 CU를 포함시킬 수 있다. 전술한 바와 같이, 인터 예측부(124)는 테이블에 저장된 HMVP 후보 중, 최근의 움직임벡터들을 머지 후보로 설정할 수 있다.
- [220] 인터 예측부(124)는 PAMVP 후보를 탐색한다(S1806).
- [221] 우선, PAMVP의 참조픽처는 첫 번째 머지 후보의 참조픽처와 동일하게 설정된다. 첫 번째 머지 후보가 가상 참조프레임을 이용한다면, 인터 예측부(124)는 제로 움직임벡터와 두 번째 후보의 움직임벡터 간의 평균에 해당하는 움직임벡터를 머지 후보로 설정할 수 있다.
- [222] 위의 과정(S1800 내지 S1806)를 모두 수행해도 머지 후보 리스트를 채울 수 없는 경우(즉, 기설정된 개수를 충족하지 못하는 경우), 인터 예측부(124)는 제로 움직임벡터를 머지 후보로 추가한다(S1808).
- [223] 도 19는 본 개시의 일 실시예에 따른, 인터 예측의 아핀 AMVP 모드에서 가상 참조프레임을 이용하는 과정에 대한 순서도를 나타낸다.
- [224] 인터 예측부(124)는 상속 아핀 AMVP 후보의 참조 픽처와 현재블록의 참조 픽처가 동일한지를 확인한다(S1900).
- [225] 상속 아핀 AMVP 후보의 참조 픽처와 현재블록의 참조 픽처가 동일한 경우, 인터 예측부(124)는 동일한 참조 픽처가 가상 참조프레임인지 기존의 참조 프레임인지를 확인한다(S1902).
- [226] 동일한 참조 픽처가 가상 참조프레임인 경우, 인터 예측부(124)는 제로 움직임벡터를 아핀 AMVP 후보로 추가한다(S1904).
- [227] 동일한 참조 픽처가 기존의 참조 프레임인 경우, 인터 예측부(124)는 해당되는

상속 아핀 AMVP 후보를 추가한다(S1906). 또는, 인터 예측부(124)는 해당 후보를 상속 아핀 AMVP 후보로 추가하지 않을 수 있다.

- [228] 상속 아핀 AMVP 후보의 참조 픽처와 현재블록의 참조 픽처가 동일하지 않은 경우, 인터 예측부(124)는 조합 아핀 AMVP 후보의 모든 CPMV의 참조 픽처와 현재블록의 참조 픽처가 동일한지를 확인한다(S1910).
- [229] 조합 아핀 AMVP 후보의 모든 CPMV의 참조 픽처와 현재블록의 참조 픽처가 동일한 경우, 동일한 참조 픽처가 가상 참조프레임인지 기존의 참조 프레임인지를 확인한다(S1912).
- [230] 동일한 참조 픽처가 가상 참조프레임인 경우, 인터 예측부(124)는 제로 움직임벡터를 아핀 AMVP 후보로 추가한다(S1914).
- [231] 동일한 참조 픽처가 기존의 참조 프레임인 경우, 인터 예측부(124)는 해당되는 조합 아핀 AMVP 후보를 추가한다(S1916). 또는, 인터 예측부(124)는 해당 후보를 조합 아핀 AMVP 후보로 추가하지 않을 수 있다.
- [232] 조합 아핀 AMVP 후보의 모든 CPMV의 참조 픽처와 현재블록의 참조 픽처가 동일하지 않은 경우, 인터 예측부(124)는 아핀 AMVP 후보로서 병진 움직임벡터를 추가한다(S1920).
- [233] 위의 과정(S1900 내지 S1920)을 모두 수행해도 후보 리스트를 채울 수 없는 경우(즉, 기설정된 개수를 충족하지 못하는 경우), 인터 예측부(124)는 제로 움직임벡터를 아핀 AMVP 후보로 추가한다(S1922).
- [234] 한편, 이상의 설명에서, 병진 움직임벡터를 후보로 추가하는 단계를 제외한 나머지 단계들(S1900 내지 S1916, 및 S1922)은, 인터 예측부(124)에 의한 아핀 움직임 후보의 탐색 과정에도 유사하게 적용될 수 있다.
- [235] 이상에서 설명한 바와 같이 본 실시예에 따르면, 가변계수 딥러닝 기반 보간 모델에 의해 생성된 가상 참조프레임을 참조하는 인터 예측 방법을 제공함으로써, 인터 예측의 효과를 증대시키고, 부호화 효율을 향상시키는 것이 가능해지는 효과가 있다.
- [236] 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.
- [237] 이상의 설명에서 예시적인 실시예들은 많은 다른 방식으로 구현될 수 있다는 것을 이해해야 한다. 하나 이상의 예시들에서 설명된 기능들 혹은 방법들은 하드웨어, 소프트웨어, 펌웨어 또는 이들의 임의의 조합으로 구현될 수 있다. 본 명세서에서 설명된 기능적 컴포넌트들은 그들의 구현 독립성을 특히 더 강조하기 위해 "...부(unit)" 로 라벨링되었음을 이해해야 한다.
- [238] 한편, 본 실시예에서 설명된 다양한 기능들 혹은 방법들은 하나 이상의 프로세서에 의해 관독되고 실행될 수 있는 비일시적 기록매체에 저장된 명령어들로 구현될 수도 있다. 비일시적 기록매체는, 예를 들어, 컴퓨터

시스템에 의하여 관독가능한 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.

- [239] 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.

[240] (부호의 설명)

[241] 190, 570: 메모리 510: 엔트로피 복호화부

[242] 124, 544: 인터 예측부

[243] 1210, 1310: 가상프레임 생성부

[244]

[245] CROSS-REFERENCE TO RELATED APPLICATION

[246] 본 특허출원은 2020년 9월 28일 한국에 출원한 특허출원번호 제10-2020-0126111 호, 2021년 9월 28일 한국에 출원한 특허출원번호 제10-2021-0127795 호에 대해 우선권을 주장하며, 그 모든 내용은 참고문헌으로 본 특허출원에 병합된다.

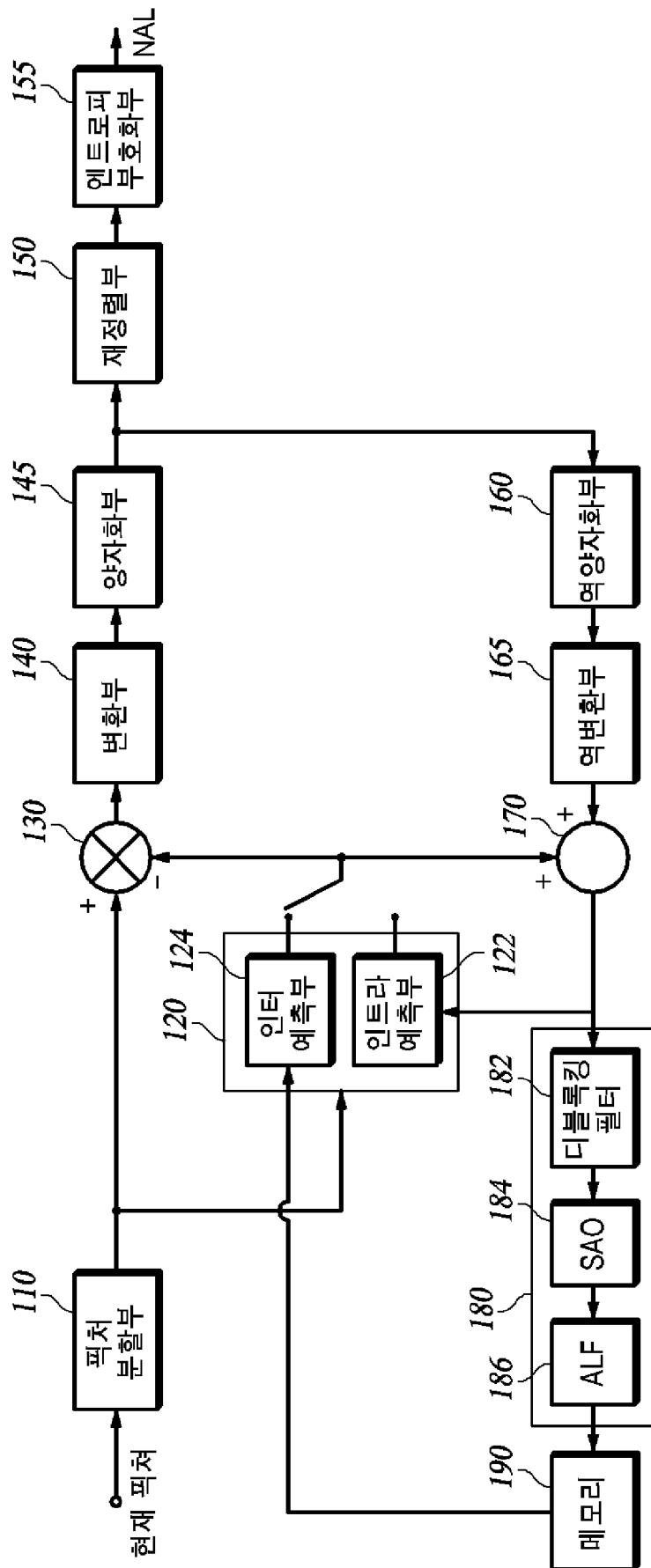
청구범위

- [청구항 1] 영상 복호화 장치가 수행하는 영상 복호화 방법에 있어서,
비트스트림으로부터 가변계수 값들, 아핀(affine) 예측 플래그, 및 부호화
모드를 복호화하는 단계, 여기서, 상기 아핀 예측 플래그는 현재블록에
대한 아핀 움직임 예측의 적용 여부를 나타내고, 상기 부호화 모드는 상기
현재블록의 움직임 정보의 부호화 모드로서 머지 모드(merge mode) 또는
AMVP(Advanced Motion Vector Prediction) 모드를 나타냄;
보간 모델(interpolation model)을 이용하여 참조 프레임들에 기초하여
가상 참조프레임을 생성하는 단계, 여기서, 상기 가변계수 값들을
이용하여 상기 보간 모델 상의 가변계수 네트워크가 설정됨; 및
상기 부호화 모드가 상기 머지 모드인 경우, 상기 가상 참조프레임 및
상기 참조 프레임들에 기초하여 상기 현재블록의 머지 후보를 생성하는
단계
를 포함하는 것을 특징으로 하는, 영상 복호화 방법.
- [청구항 2] 제1항에 있어서,
상기 보간 모델은,
상기 참조 프레임들로부터 상기 가상 참조프레임을 바로 생성하는
딥러닝 모델, 상기 참조 프레임들로부터 마스크를 생성한 후, 상기
마스크를 기반으로 상기 참조 프레임들을 가중합하여 상기 가상
참조프레임을 생성하는 딥러닝 모델, 또는, 상기 참조 프레임들로부터
옵티컬 플로우를 생성한 후, 상기 옵티컬 플로우를 기반으로 상기 참조
프레임들을 와핑(warping)하여 상기 가상 참조프레임을 생성하는 딥러닝
모델 중의 하나인 것을 특징으로 하는, 영상 복호화 방법.
- [청구항 3] 제1항에 있어서,
상기 보간 모델이 생성하는 추론 프레임과 원본 프레임 간의 차이에
기반하는 손실함수를 이용하여, 영상 부호화 장치가 상기 보간 모델의
가변계수 네트워크를 업데이트하여 상기 가변계수 값들을 생성하는 것을
특징으로 하는, 영상 복호화 방법.
- [청구항 4] 제1항에 있어서,
상기 보간 모델은,
고정계수 네트워크를 추가로 포함하되, 상기 고정계수 네트워크의
고정계수 값들은 원본 참조 프레임 전체에 기초하는 사전학습에서
생성되고, 사전 약속에 따라 설정되는 것을 특징으로 하는, 영상 복호화
방법.
- [청구항 5] 제4항에 있어서,
상기 보간 모델의 가변계수 네트워크는,
상기 고정계수 네트워크의 후단에 연결되는 것을 특징으로 하는, 영상

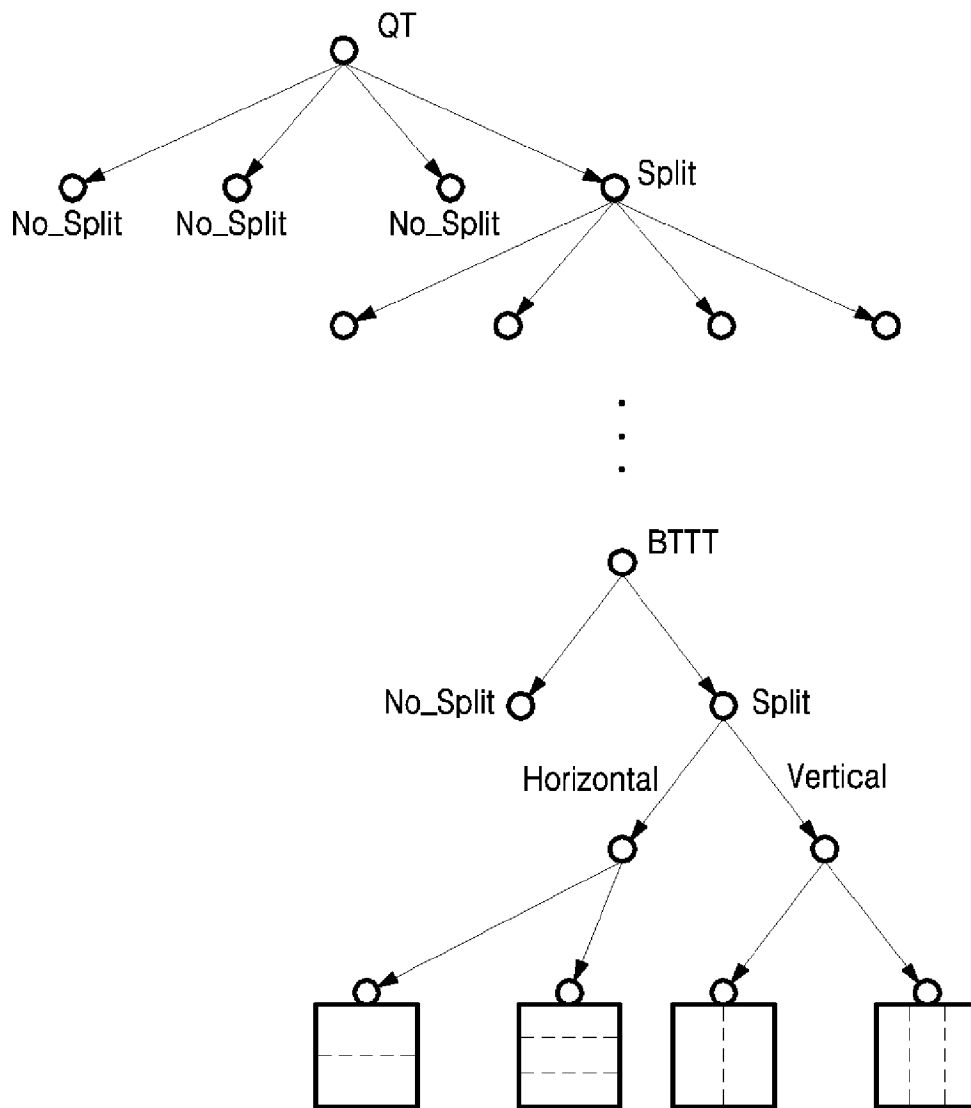
- 복호화 방법.
- [청구항 6] 제1항에 있어서,
상기 가변계수 값들은,
프레임 단위, 또는 적어도 하나의 GOP(Group of Pictures) 단위로
복호화되는 것을 특징으로 하는, 영상 복호화 방법.
- [청구항 7] 제1항에 있어서,
상기 머지 후보를 생성하는 단계는,
공간적 머지 후보를 탐색하는 단계; 및
상기 가상 참조프레임을 고려하여 시간적 머지 후보를 탐색하는 단계
를 포함하는 것을 특징으로 하는, 영상 복호화 방법.
- [청구항 8] 제7항에 있어서,
상기 시간적 머지 후보를 탐색하는 단계는,
상기 현재블록과 동일 위치에 있는(co-located) 블록이 존재하는 픽처와
상기 시간적 머지 후보의 참조 프레임이 모두 상기 가상 참조프레임인
경우, 제로 움직임벡터를 상기 머지 후보로 설정하는 것을 특징으로 하는,
영상 복호화 방법.
- [청구항 9] 제1항에 있어서,
상기 아핀 예측 플래그가 참이고, 상기 부호화 모드가 상기 AMVP 모드인
경우, 아핀 AMVP 후보를 생성하는 단계를 더 포함하되,
상기 아핀 AMVP 후보를 생성하는 단계는,
상기 가상 참조프레임을 고려하여 상속(inherited) 아핀 AMVP 후보를
추가하는 단계; 및
상기 가상 참조프레임을 고려하여 조합(constructed) 아핀 AMVP 후보를
추가하는 단계;
를 포함하는 것을 특징으로 하는, 영상 복호화 방법.
- [청구항 10] 제9항에 있어서,
상기 상속 아핀 AMVP 후보를 추가하는 단계는,
상속 아핀 AMVP 후보의 참조 픽처와 상기 현재블록의 참조 픽처가 동일
참조픽처이고, 상기 동일 참조픽처가 상기 가상 참조프레임인 경우, 제로
움직임벡터를 상기 아핀 AMVP 후보로 추가하는 것을 특징으로 하는,
영상 복호화 방법.
- [청구항 11] 제9항에 있어서,
상기 조합 아핀 AMVP 후보를 추가하는 단계는,
상기 조합 아핀 AMVP 후보의 모든 제어점 움직임벡터들(control point
motion vectors)의 참조 픽처와 상기 현재블록의 참조 픽처가 동일
참조픽처이고, 상기 동일 참조픽처가 상기 가상 참조프레임인 경우, 제로
움직임벡터를 상기 아핀 AMVP 후보로 추가하는 것을 특징으로 하는,
영상 복호화 방법.

- [청구항 12] 영상 부호화 장치가 수행하는 영상 부호화 방법에 있어서,
 사전에 생성된 가변계수 값들, 및 기설정된 부호화 모드를 획득하는 단계,
 여기서, 상기 부호화 모드는 현재블록의 움직임 정보의 부호화 모드로서
 머지 모드(merge mode) 또는 AMVP(Advanced Motion Vector Prediction)
 모드를 나타냄;
 보간 모델(interpolation model)을 이용하여 참조 프레임들로부터 가상
 참조프레임을 생성하는 단계, 여기서, 상기 가변계수 값들을 이용하여
 상기 보간 모델 상의 가변계수 네트워크가 설정됨; 및
 상기 부호화 모드가 상기 머지 모드인 경우, 상기 가상 참조프레임 및
 상기 참조 프레임들에 기초하여 상기 현재블록의 머지 후보를 생성하는
 단계
 를 포함하는 것을 특징으로 하는, 영상 부호화 방법.
- [청구항 13] 제12항에 있어서,
 상기 보간 모델이 생성하는 추론 프레임과 원본 프레임 간의 차이에
 기반하는 손실함수를 이용하여, 상기 보간 모델의 가변계수 네트워크를
 업데이트하여 상기 가변계수 값들을 생성하는 것을 특징으로 하는, 영상
 부호화 방법.
- [청구항 14] 제12항에 있어서,
 상기 보간 모델은,
 고정계수 부분을 추가로 포함하되, 상기 고정계수 부분의 고정계수
 값들은 원본 참조 프레임 전체에 기초하는 사전학습에서 생성되고, 사전
 약속에 따라 설정되는 것을 특징으로 하는, 영상 부호화 방법.
- [청구항 15] 제12항에 있어서,
 상기 가변계수 값들은,
 프레임 단위, 또는 적어도 하나의 GOP(Group of Pictures) 단위로
 부호화되는 것을 특징으로 하는, 영상 부호화 방법.
- [청구항 16] 비트스트림으로부터 가변계수 값들, 및 부호화 모드를 복호화하는
 엔트로피 복호화부, 여기서, 상기 부호화 모드는 현재블록의 움직임
 정보의 부호화 모드로서 머지 모드(merge mode) 또는 AMVP(Advanced
 Motion Vector Prediction) 모드를 나타냄;
 보간 모델(interpolation model)을 이용하여 참조 프레임들로부터 가상
 참조프레임을 생성하는 가상프레임 생성부, 여기서, 상기 가변계수
 값들을 이용하여 상기 보간 모델 상의 가변계수 네트워크가 설정됨; 및
 상기 부호화 모드가 상기 머지 모드인 경우, 상기 가상 참조프레임 및
 상기 참조 프레임들에 기초하여 상기 현재블록의 머지 후보를 생성하는
 인터 예측부
 를 포함하는 것을 특징으로 하는, 영상 복호화 장치.

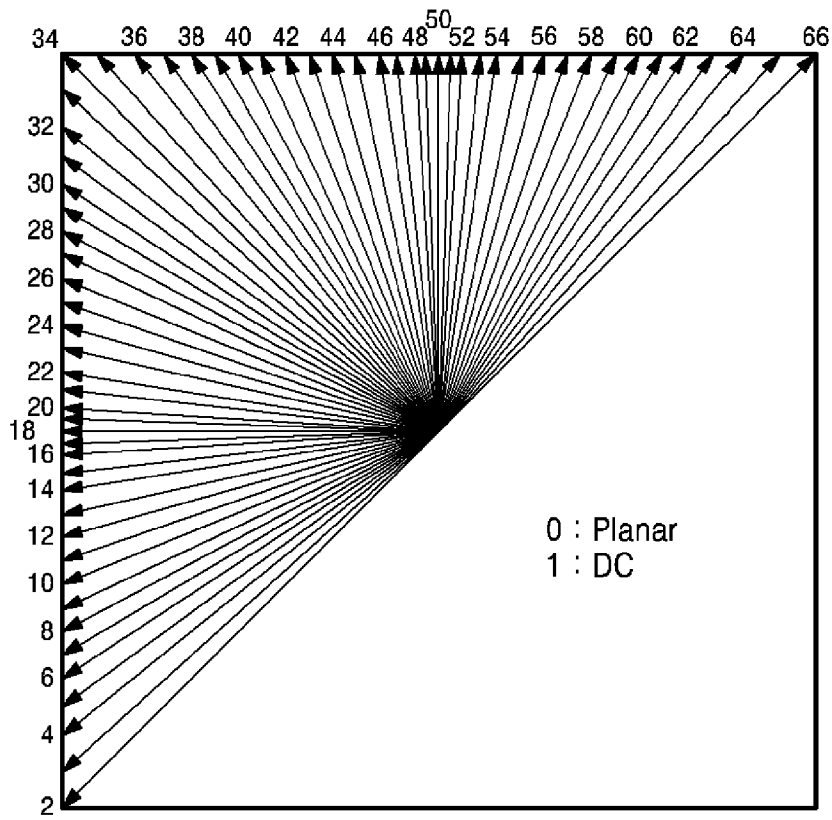
[도 1]



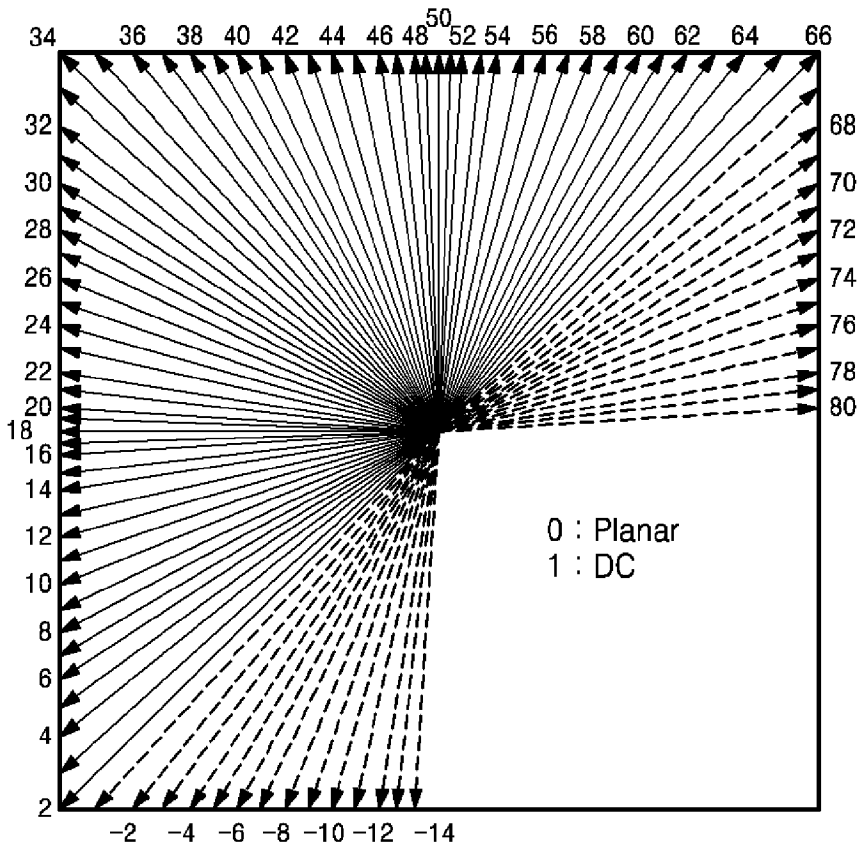
[도2]



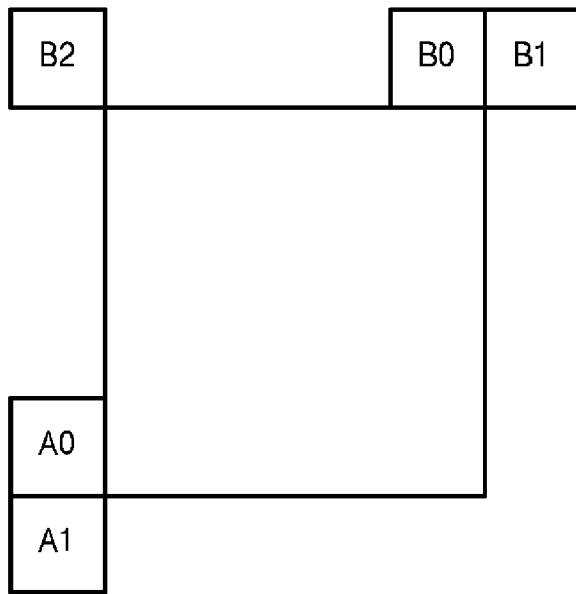
[도3a]



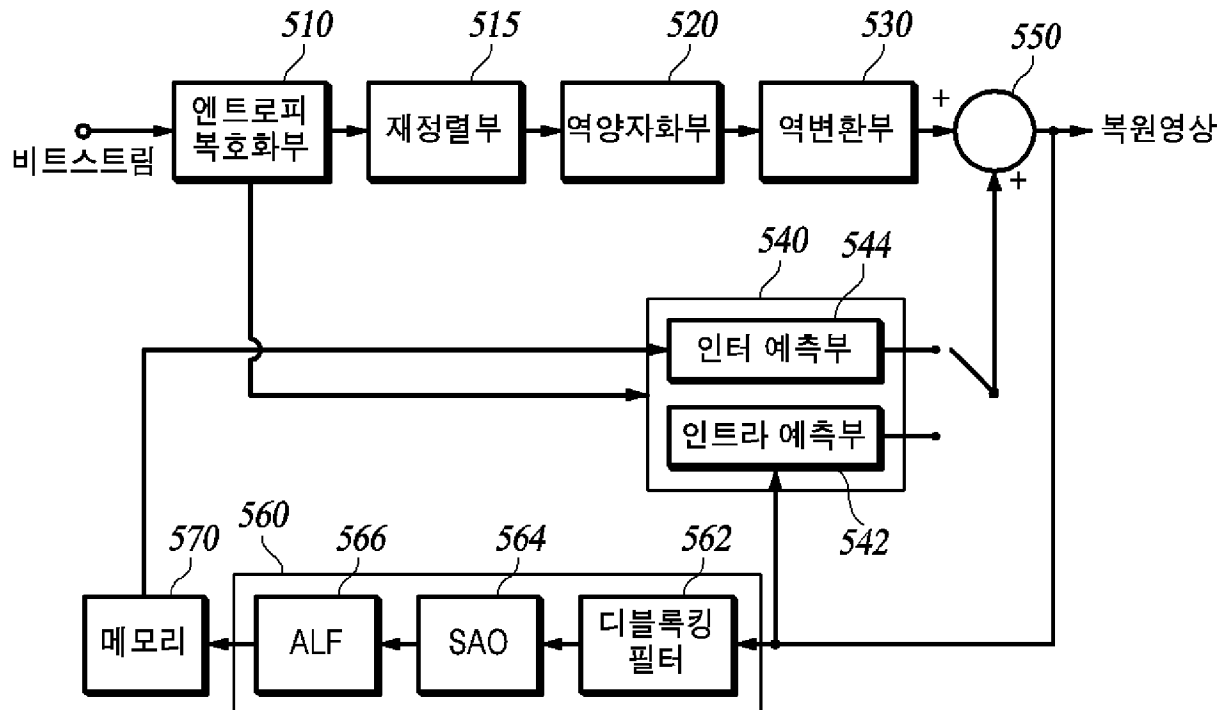
[도3b]



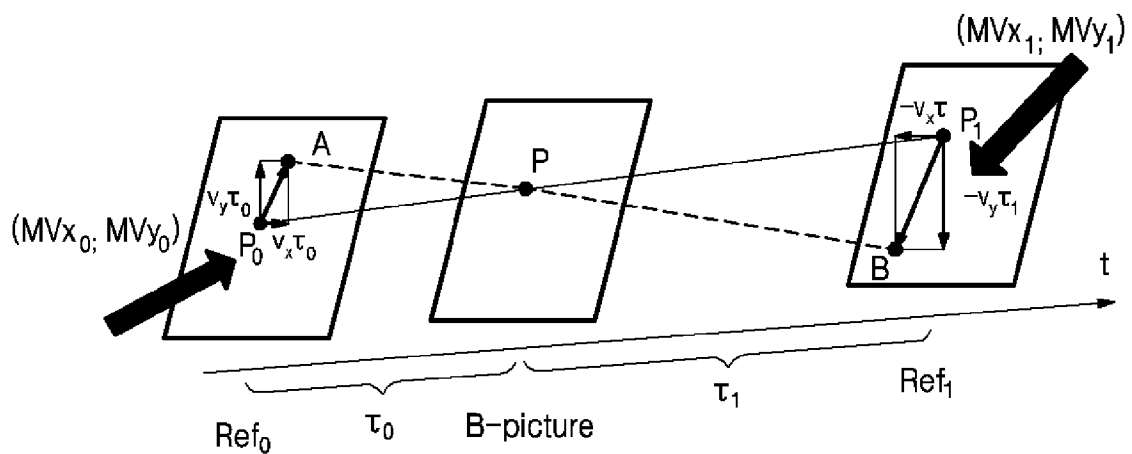
[도4]



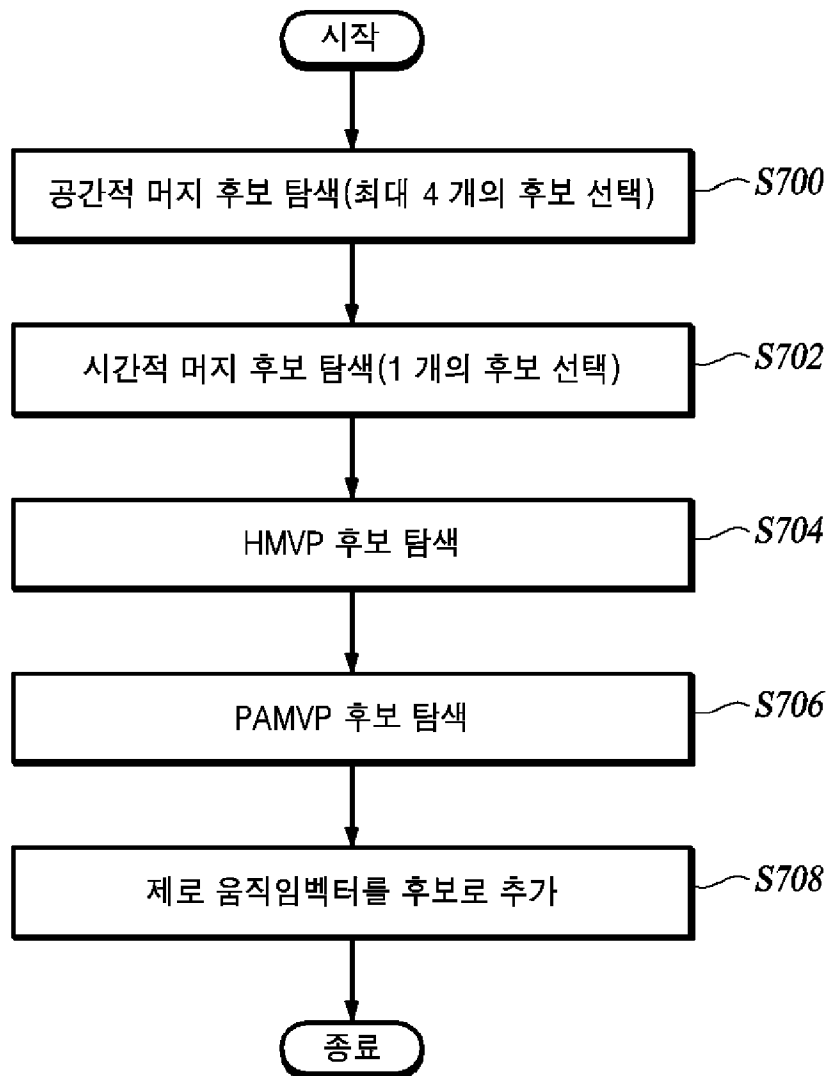
[도5]



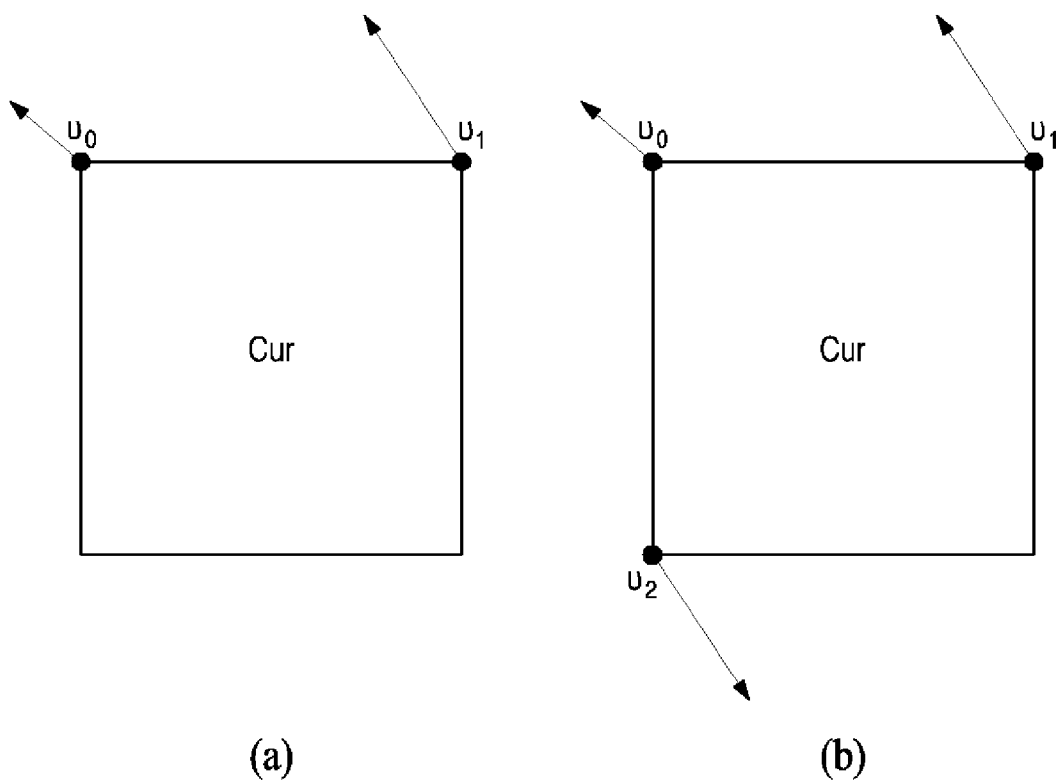
[도6]



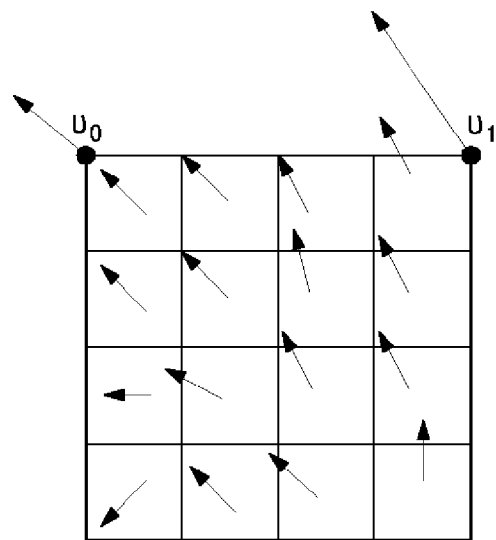
[도7]



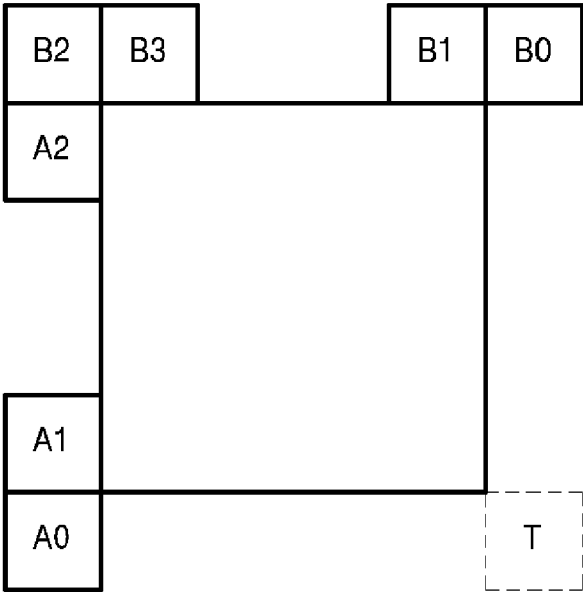
[도8]



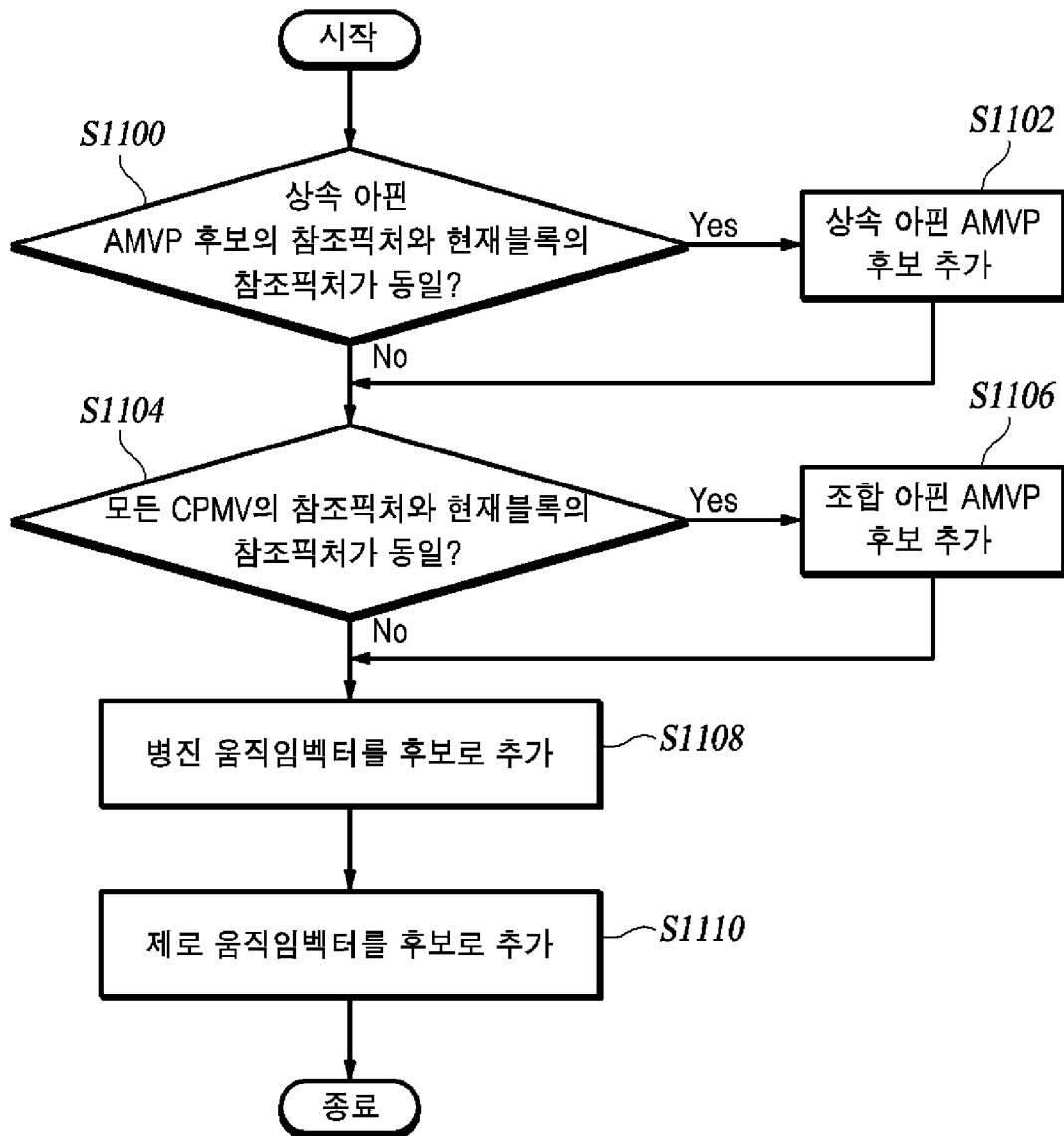
[도9]



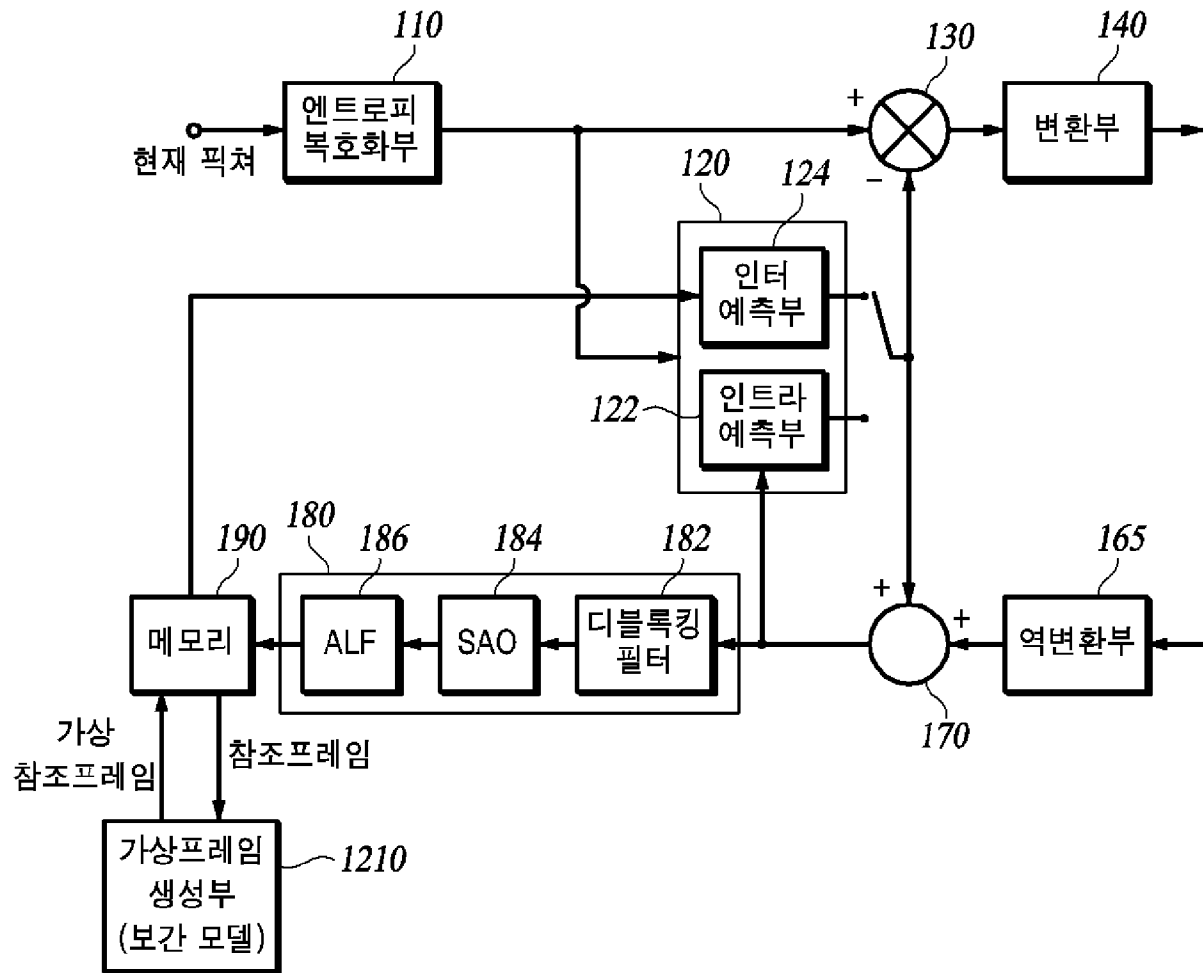
[도10]



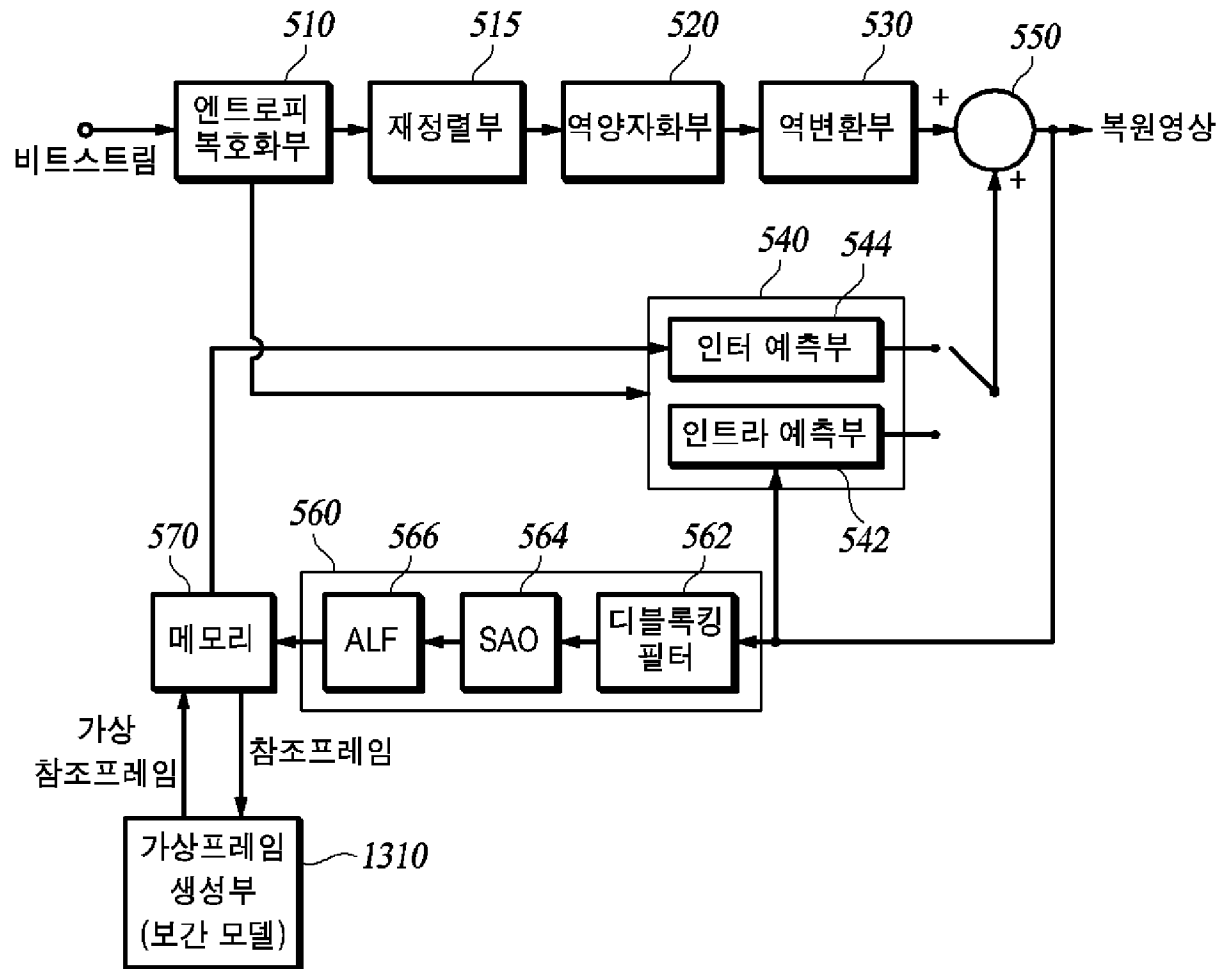
[도11]



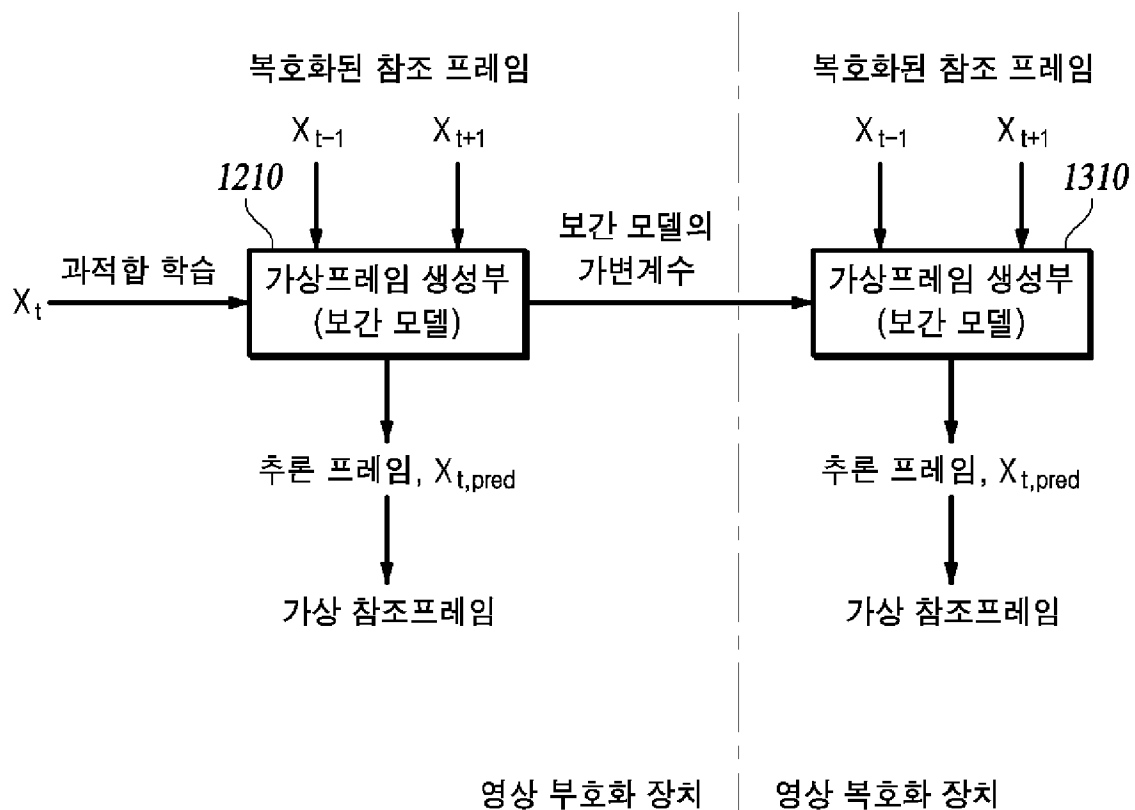
[도12]



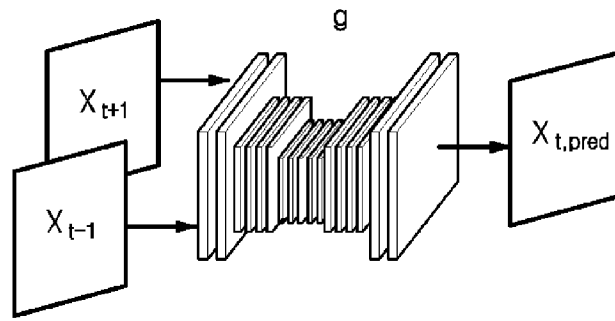
[도13]



[도14]



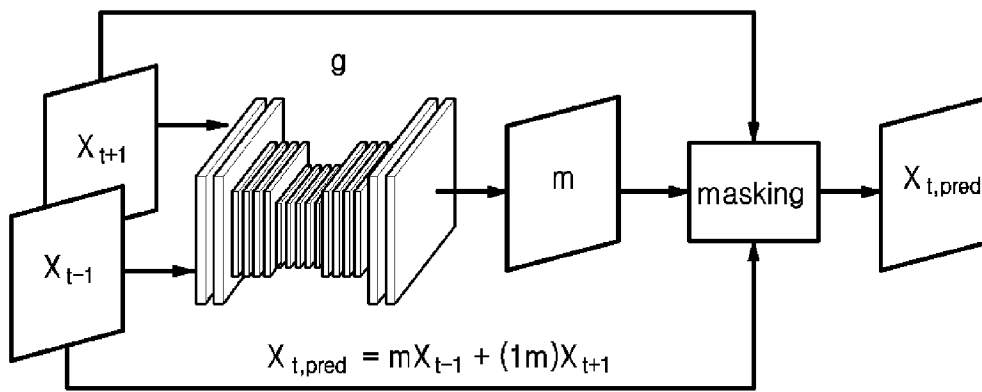
[도15]



추론된 프레임을 바로 생성하는 보간 모델

$$X_{t,pred} = g([X_{t-1}, X_{t+1}])$$

(a)

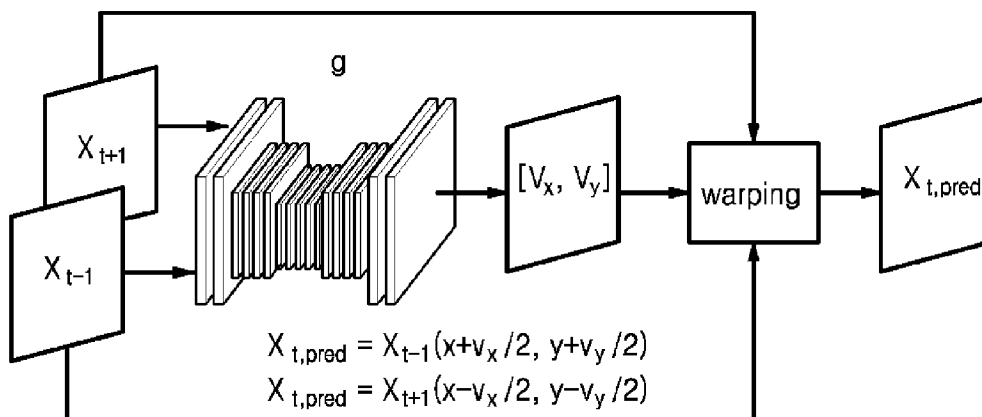


$$X_{t,pred} = mX_{t-1} + (1-m)X_{t+1}$$

마스크를 이용하는 보간 모델

$$m = g([X_{t-1}, X_{t+1}])$$

(b)



$$X_{t,pred} = X_{t-1}(x+v_x/2, y+v_y/2)$$

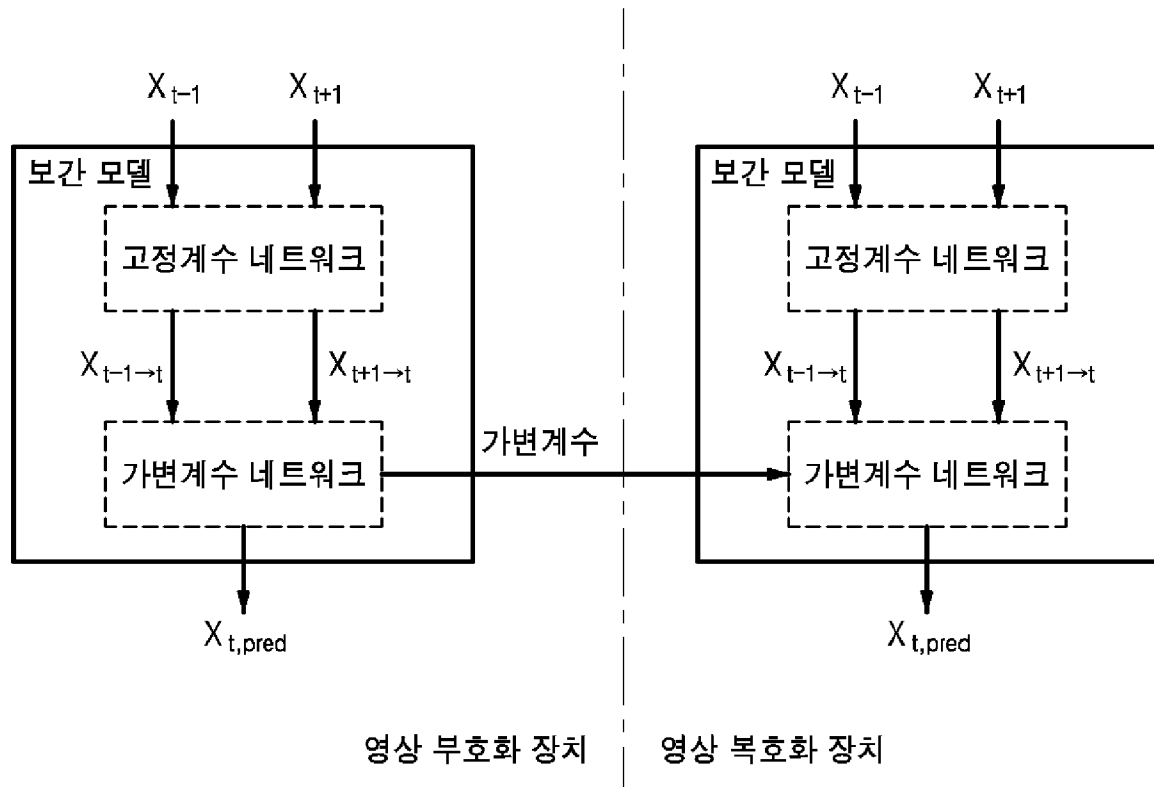
$$X_{t,pred} = X_{t+1}(x-v_x/2, y-v_y/2)$$

OF를 이용하는 보간 모델

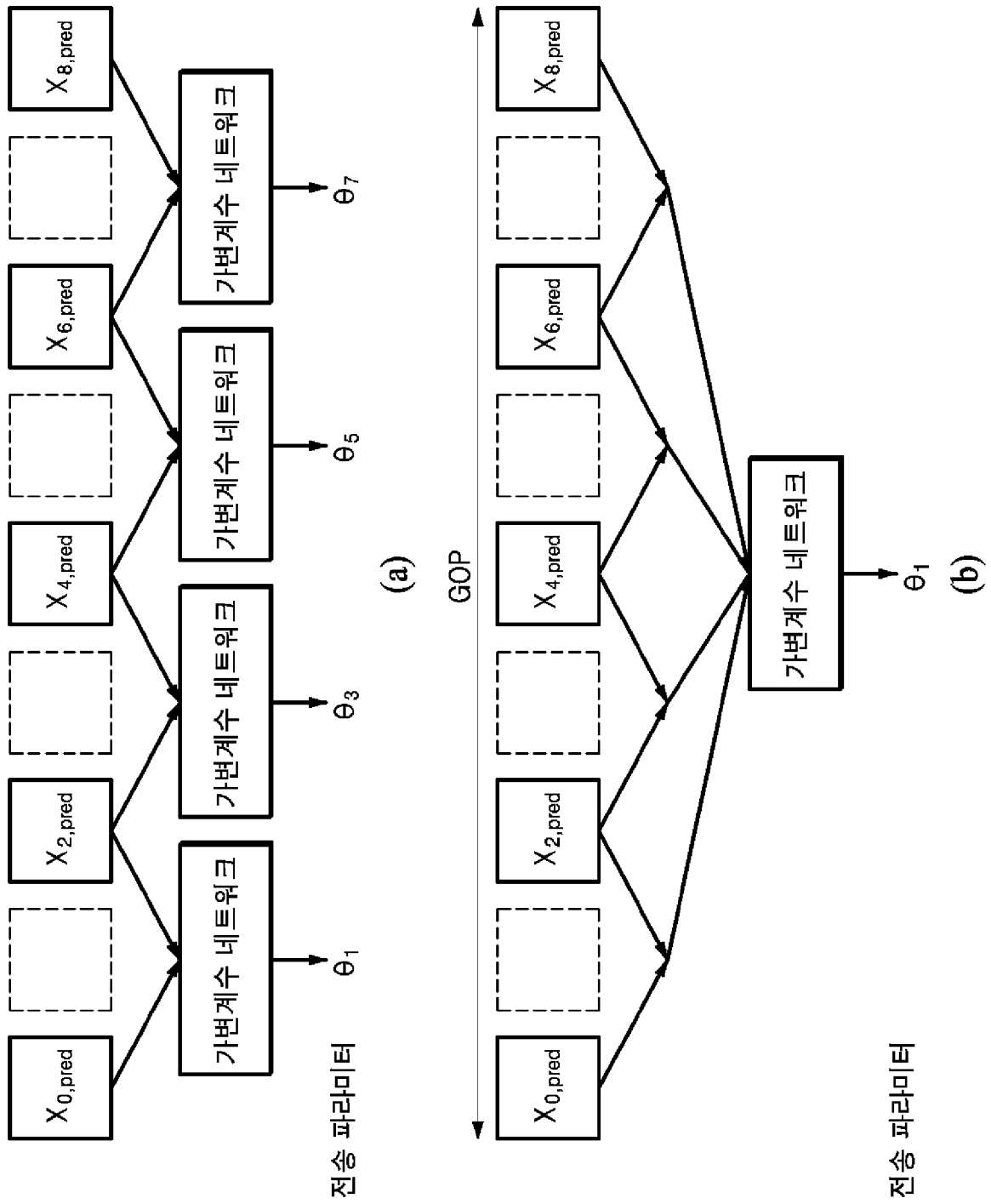
$$[V_x, V_y] = g([X_{t-1}, X_{t+1}])$$

(c)

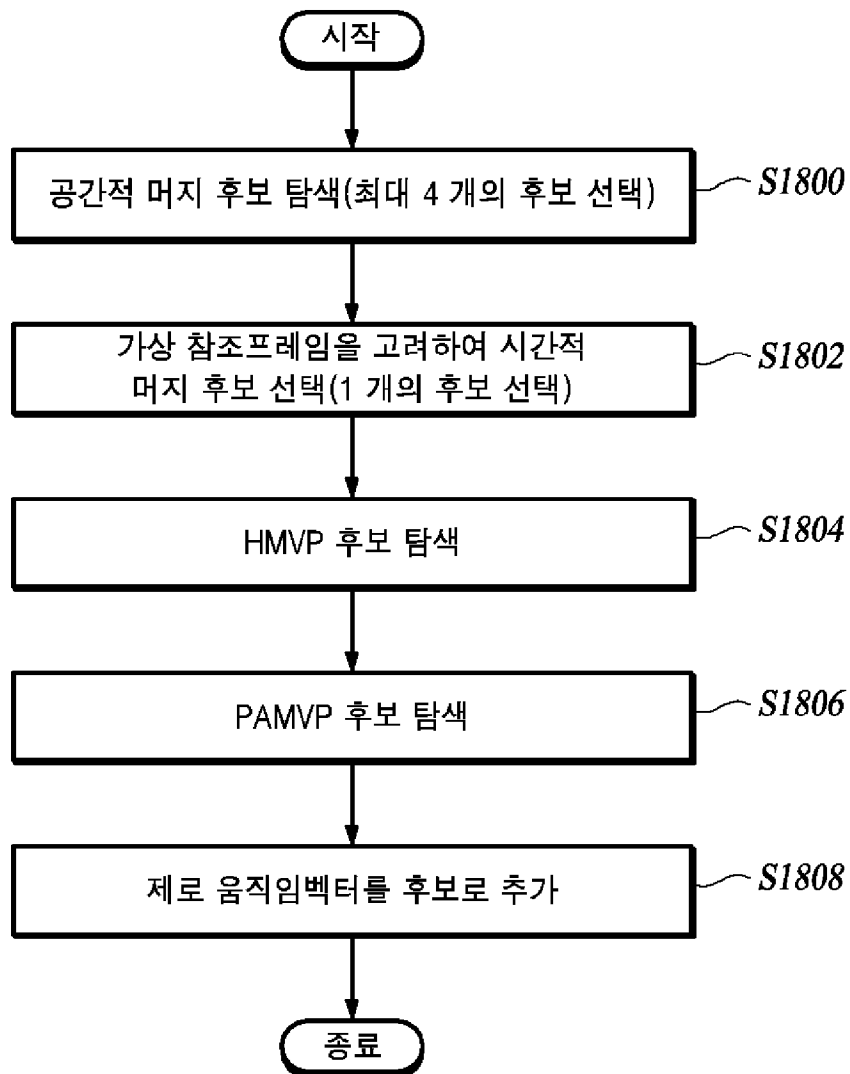
[도16]



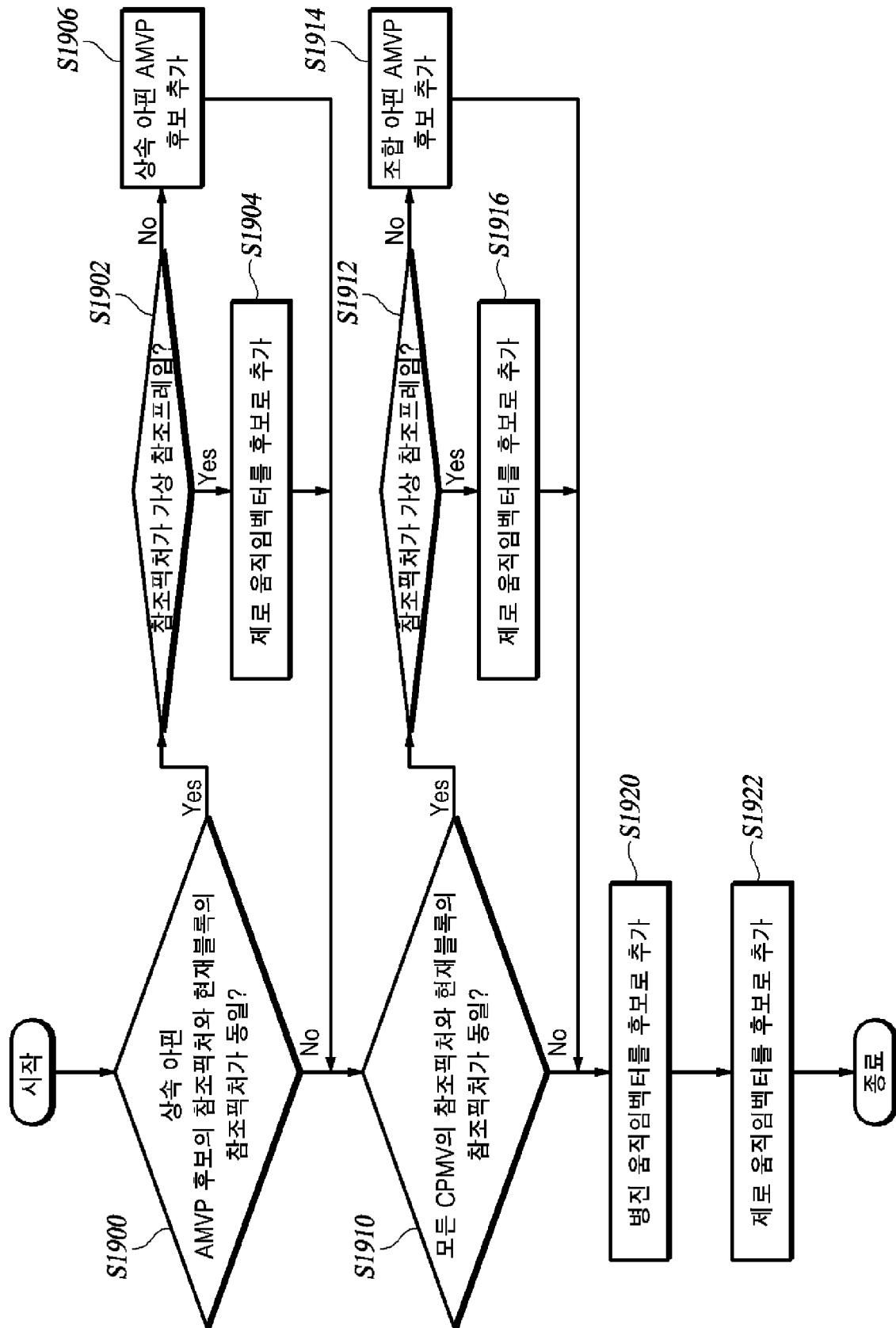
[도17]



[도18]



[도 19]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2021/013218

A. CLASSIFICATION OF SUBJECT MATTER

G06T 9/00(2006.01)i; G06T 3/40(2006.01)i; G06T 5/00(2006.01)i; H04N 19/82(2014.01)i; H04N 19/86(2014.01)i;
H04N 19/70(2014.01)i; H04N 19/109(2014.01)i; H04N 19/503(2014.01)i; G06N 3/08(2006.01)i; G06N 3/04(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06T 9/00(2006.01); H04N 19/105(2014.01); H04N 19/124(2014.01); H04N 19/132(2014.01); H04N 19/136(2014.01);
H04N 19/172(2014.01); H04N 19/184(2014.01); H04N 19/503(2014.01); H04N 21/44(2011.01); H04N 21/4402(2011.01)

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Korean utility models and applications for utility models: IPC as above
Japanese utility models and applications for utility models: IPC as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

eKOMPASS (KIPO internal) & keywords: 인터 예측(inter-prediction), 참조(reference), 보간(interpolation), 딥러닝(deep learning), 아핀(affine)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	KR 10-2019-0116067 A (ELECTRONICS AND TELECOMMUNICATIONS RESEARCH INSTITUTE et al.) 14 October 2019 (2019-10-14) See paragraphs [0693], [0982]-[1038] and [1061]-[1066]; claims 1-8; and figures 26, 41 and 44.	1-8,12-16
A		9-11
Y	KR 10-2020-0044661 A (SAMSUNG ELECTRONICS CO., LTD.) 29 April 2020 (2020-04-29) See paragraphs [0092]-[0105]; claim 1; and figure 5.	1-8,12-16
A	KR 10-2019-0096281 A (ELECTRONICS AND TELECOMMUNICATIONS RESEARCH INSTITUTE et al.) 19 August 2019 (2019-08-19) See claims 1-16.	1-16
A	KR 10-2018-0119753 A (KANG, Hyun-Inn) 05 November 2018 (2018-11-05) See claims 1-6.	1-16



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance
“D” document cited by the applicant in the international application
“E” earlier application or patent but published on or after the international filing date
“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
“O” document referring to an oral disclosure, use, exhibition or other means
“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

03 January 2022

Date of mailing of the international search report

03 January 2022

Name and mailing address of the ISA/KR

Korean Intellectual Property Office
Government Complex-Daejeon Building 4, 189 Cheongsaro, Seo-gu, Daejeon 35208

Facsimile No. +82-42-481-8578

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2021/013218

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	KR 10-2020-0054720 A (SAMSUNG ELECTRONICS CO., LTD.) 20 May 2020 (2020-05-20) See claims 1-10.	1-16

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/KR2021/013218

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
KR	10-2019-0116067	A	14 October 2019	US	11019355	B2	25 May 2021
				US	2019-0306526	A1	03 October 2019
KR	10-2020-0044661	A	29 April 2020	CN	112889282	A	01 June 2021
				EP	3844962	A1	07 July 2021
				KR	10-2021-0113131	A	15 September 2021
KR	10-2019-0096281	A	19 August 2019	US	10841577	B2	17 November 2020
				US	2019-0246102	A1	08 August 2019
KR	10-2018-0119753	A	05 November 2018	KR	10-2053242	B1	06 December 2019
				WO	2018-199459	A1	01 November 2018
KR	10-2020-0054720	A	20 May 2020	CN	111179163	A	19 May 2020
				EP	3651462	A1	13 May 2020
				US	2020-0154123	A1	14 May 2020
				WO	2020-101257	A1	22 May 2020

A. 발명이 속하는 기술분류(국제특허분류(IPC)) G06T 9/00(2006.01)i; G06T 3/40(2006.01)i; G06T 5/00(2006.01)i; H04N 19/82(2014.01)i; H04N 19/86(2014.01)i; H04N 19/70(2014.01)i; H04N 19/109(2014.01)i; H04N 19/503(2014.01)i; G06N 3/08(2006.01)i; G06N 3/04(2006.01)i		
B. 조사된 분야 조사된 최소문헌(국제특허분류를 기재) G06T 9/00(2006.01); H04N 19/105(2014.01); H04N 19/124(2014.01); H04N 19/132(2014.01); H04N 19/136(2014.01); H04N 19/172(2014.01); H04N 19/184(2014.01); H04N 19/503(2014.01); H04N 21/44(2011.01); H04N 21/4402(2011.01) 조사된 기술분야에 속하는 최소문헌 이외의 문헌 한국등록실용신안공보 및 한국공개실용신안공보: 조사된 최소문헌란에 기재된 IPC 일본등록실용신안공보 및 일본공개실용신안공보: 조사된 최소문헌란에 기재된 IPC 국제조사에 이용된 전산 데이터베이스(데이터베이스의 명칭 및 검색어(해당하는 경우)) eKOMPASS(특허청 내부 검색시스템) & 키워드: 인터 예측(inter-prediction), 참조(reference), 보간(interpolation), 딥러닝(deep learning), 아핀(affine)		
C. 관련 문헌		
카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
Y	KR 10-2019-0116067 A (한국전자통신연구원 등) 2019.10.14 단락 [0693], [0982]-[1038], [1061]-[1066]; 청구항 1-8; 및 도면 26, 41, 44	1-8,12-16
A		9-11
Y	KR 10-2020-0044661 A (삼성전자주식회사) 2020.04.29 단락 [0092]-[0105]; 청구항 1; 및 도면 5	1-8,12-16
A	KR 10-2019-0096281 A (한국전자통신연구원 등) 2019.08.19 청구항 1-16	1-16
A	KR 10-2018-0119753 A (강현인) 2018.11.05 청구항 1-6	1-16
A	KR 10-2020-0054720 A (삼성전자주식회사) 2020.05.20 청구항 1-10	1-16
☐ 추가 문헌이 C(계속)에 기재되어 있습니다. ☒ 대응특허에 관한 별지를 참조하십시오.		
* 인용된 문헌의 특별 카테고리: “A” 특별히 관련이 없는 것으로 보이는 일반적인 기술수준을 정의한 문헌 “D” 본 국제출원에서 출원인이 인용한 문헌 “E” 국제출원일보다 빠른 출원일 또는 우선일을 가지나 국제출원일 이후에 공개된 선출원 또는 특허 문헌 “L” 우선권 주장에 의문을 제기하는 문헌 또는 다른 인용문헌의 공개일 또는 다른 특별한 이유(이유를 명시)를 밝히기 위하여 인용된 문헌 “O” 구두 개시, 사용, 전시 또는 기타 수단을 언급하고 있는 문헌 “P” 우선일 이후에 공개되었으나 국제출원일 이전에 공개된 문헌 “T” 국제출원일 또는 우선일 후에 공개된 문헌으로, 출원과 상충하지 않으며 발명의 기초가 되는 원리나 이론을 이해하기 위해 인용된 문헌 “X” 특별한 관련이 있는 문헌. 해당 문헌 하나만으로 청구된 발명의 신규성 또는 진보성이 없는 것으로 본다. “Y” 특별한 관련이 있는 문헌. 해당 문헌이 하나 이상의 다른 문헌과 조합하는 경우로 그 조합이 당업자에게 자명한 경우 청구된 발명은 진보성이 없는 것으로 본다. “&” 동일한 대응특허문헌에 속하는 문헌		
국제조사의 실제 완료일 2022년01월03일(03.01.2022)		국제조사보고서 발송일 2022년01월03일(03.01.2022)
ISA/KR의 명칭 및 우편주소 대한민국 특허청 (35208) 대전광역시 서구 청사로 189, 4동 (둔산동, 정부대전청사) 팩스 번호 +82-42-481-8578		심사관 양정록 전화번호 +82-42-481-5709

국제조사보고서에서 인용된 특허문헌	공개일	대응특허문헌	공개일
KR 10-2019-0116067 A	2019/10/14	US 11019355 B2	2021/05/25
		US 2019-0306526 A1	2019/10/03
KR 10-2020-0044661 A	2020/04/29	CN 112889282 A	2021/06/01
		EP 3844962 A1	2021/07/07
		KR 10-2021-0113131 A	2021/09/15
KR 10-2019-0096281 A	2019/08/19	US 10841577 B2	2020/11/17
		US 2019-0246102 A1	2019/08/08
KR 10-2018-0119753 A	2018/11/05	KR 10-2053242 B1	2019/12/06
		WO 2018-199459 A1	2018/11/01
KR 10-2020-0054720 A	2020/05/20	CN 111179163 A	2020/05/19
		EP 3651462 A1	2020/05/13
		US 2020-0154123 A1	2020/05/14
		WO 2020-101257 A1	2020/05/22