

(12) 특허협력조약에 의하여 공개된 국제출원

(19) 세계지식재산권기구  
국제사무국

(43) 국제공개일  
2020년 4월 23일 (23.04.2020)



(10) 국제공개번호  
**WO 2020/080751 A1**

(51) 국제특허분류:

H04N 19/132 (2014.01) H04N 19/184 (2014.01)  
H04N 19/85 (2014.01) G06N 3/08 (2006.01)  
H04N 19/50 (2014.01)

(21) 국제출원번호: PCT/KR2019/013344

(22) 국제출원일: 2019년 10월 11일 (11.10.2019)

(25) 출원언어: 한국어

(26) 공개언어: 한국어

(30) 우선권정보:

10-2018-0125406 2018년 10월 19일 (19.10.2018) KR  
10-2019-0041111 2019년 4월 8일 (08.04.2019) KR  
10-2019-0076569 2019년 6월 26일 (26.06.2019) KR

(71) 출원인: 삼성전자 주식회사 (SAMSUNG ELECTRONICS CO., LTD.) [KR/KR]; 16677 경기도 수원시 영통구 삼성로 129, Gyeonggi-do (KR).

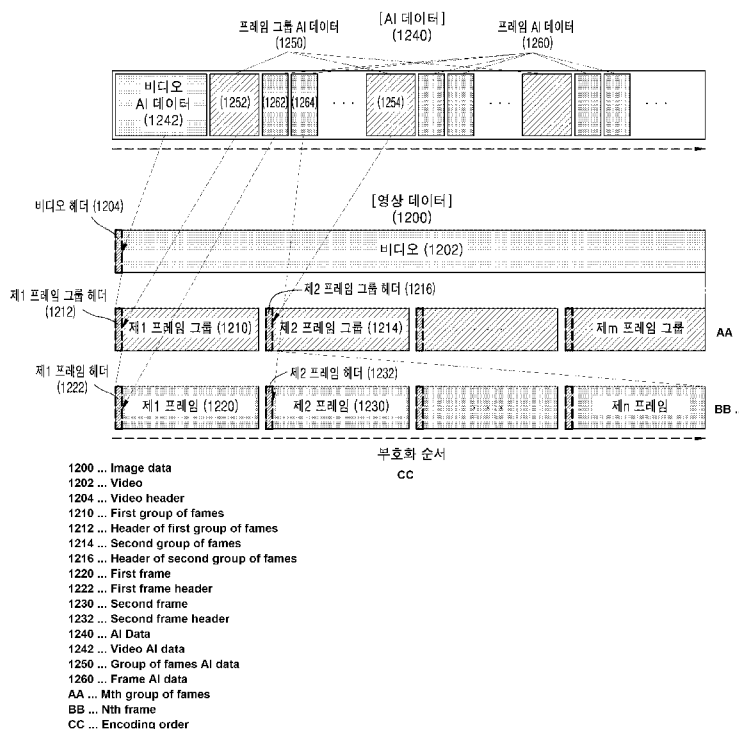
(72) 발명자: 박필규 (PARK, Pilkyu); 16677 경기도 수원시 영통구 삼성로 129, Gyeonggi-do (KR). 박영오 (PARK, Youngo); 16677 경기도 수원시 영통구 삼성로 129, Gyeonggi-do (KR). 이종석 (LEE, Jongseok); 16677 경기도 수원시 영통구 삼성로 129, Gyeonggi-do (KR). 손유미 (SOHN, Yumi); 16677 경기도 수원시 영통구 삼성로 129, Gyeonggi-do (KR). 엄명진 (EOM, Myungjin); 16677 경기도 수원시 영통구 삼성로 129, Gyeonggi-do (KR).

(74) 대리인: 리앤목 특허법인 (Y.P.LEE,MOCK & PARTNERS); 06292 서울시 강남구 언주로30길 13 대림아크로텔 12층, 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, JO, JP, KE, KG, KH, KN, KP, KW, KZ,

(54) Title: ENCODING METHOD AND APPARATUS THEREOF, AND DECODING METHOD AND APPARATUS THEREOF

(54) 발명의 명칭: 부호화 방법 및 그 장치, 복호화 방법 및 그 장치



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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

- (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))

(57) 요약서: AI 부호화 데이터를 포함하는 비디오 파일이 저장된 컴퓨터로 기록가능한 기록 매체에 있어서, AI 부호화 데이터는, 고해상도 영상을 AI 다운스케일하여 생성된 저해상도 영상의 부호화 정보를 포함하는 영상 데이터, 및 영상 데이터에 따라 복원된 저해상도 영상의 AI 업스케일에 대한 AI 데이터 포함하고, AI 데이터는, 하나 이상의 프레임에 AI 업스케일이 적용되는지 여부를 나타내는 AI 타겟 데이터, 및 하나 이상의 프레임에 AI 업스케일이 적용될 경우, 미리 설정된 복수의 디폴트 DNN 설정 정보 중에서, 하나 이상의 프레임의 AI 업스케일에 사용되는 업스케일 DNN 정보에 관한 AI 보조 데이터를 포함하는 것을 특징으로 하는 컴퓨터로 기록가능한 기록 매체가 제공된다.

## 명세서

### 발명의 명칭: 부호화 방법 및 그 장치, 복호화 방법 및 그 장치 기술분야

- [1] 본 개시는 영상 처리 분야에 관한 것이다. 보다 구체적으로, 본 개시는 딥 뉴럴 네트워크(Deep Neural Network)를 이용하여 영상을 부호화 및 복호화하는 장치 및 방법에 관한 것이다.

#### 배경기술

- [2] 영상은 소정의 데이터 압축 표준, 예를 들어 MPEG (Moving Picture Expert Group) 표준 등을 따르는 코덱(codec)에 의해 부호화된 후 비트스트림의 형태로 기록매체에 저장되거나 통신 채널을 통해 전송된다.
- [3] 고해상도/고화질의 영상을 재생, 저장할 수 있는 하드웨어의 개발 및 보급에 따라, 고해상도/고화질의 영상을 효과적으로 부호화 및 복호화할 수 있는 코덱의 필요성이 증대하고 있다.

#### 발명의 상세한 설명

##### 기술적 과제

- [4] 일 실시예에 따른 영상의 부호화 및 복호화 방법, 및 장치는 DNN(Deep Neural Network)를 이용하여 낮은 비트레이트로 영상을 부호화 및 복호화하는 것을 기술적 과제로 한다.
- [5] 더불어 본 개시의 일 실시 예에 따른 영상 부호화 방법 및 영상 복호화 방법을 컴퓨터에서 실행시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체가 개시된다.

##### 과제 해결 수단

- [6] 본 개시에 있어서, AI 부호화 데이터를 포함하는 비디오 파일이 저장된 컴퓨터로 기록가능한 기록 매체에 있어서, 상기 AI 부호화 데이터는, 고해상도 영상을 AI 다운스케일하여 생성된 저해상도 영상의 부호화 정보를 포함하는 영상 데이터, 및 상기 영상 데이터에 따라 복원된 저해상도 영상의 AI 업스케일에 대한 AI 데이터 포함하고, 상기 AI 데이터는, 하나 이상의 프레임에 AI 업스케일이 적용되는지 여부를 나타내는 AI 타겟 데이터, 및 상기 하나 이상의 프레임에 AI 업스케일이 적용될 경우, 미리 설정된 복수의 디폴트 DNN 설정 정보 중에서, 상기 하나 이상의 프레임의 AI 업스케일에 사용되는 업스케일 DNN 정보에 관한 AI 보조 데이터를 포함하는 것을 특징으로 하는 컴퓨터로 기록가능한 기록 매체가 제공된다.
- [7] 본 개시에 있어서, AI 업스케일을 이용한 비디오 복호화 방법에 있어서, 영상 데이터와 상기 영상 데이터의 AI 업스케일에 관한 AI 데이터가 포함된 AI 부호화 데이터를 포함하는 비디오 파일을 수신하는 단계, 상기 비디오 파일의 메타데이터 박스로부터 상기 AI 부호화 데이터의 AI 데이터를 획득하고, 상기

비디오 파일의 미디어데이터 박스로부터, 상기 AI 부호화 데이터의 영상 데이터를 획득하는 단계, 상기 영상 데이터를 복호화하여, 현재 프레임의 저해상도 영상을 복원하는 단계, 상기 AI 데이터로부터, 상기 현재 프레임의 업스케일 DNN 정보를 획득하는 단계, 및 상기 현재 프레임의 업스케일 DNN 정보에 따라, 상기 저해상도 영상을 AI 업스케일함으로써, 상기 저해상도 영상에 대응되는 고해상도 영상을 생성하는 단계를 포함하는 비디오 복호화 방법이 제공된다.

- [8] 본 개시에 있어서, AI 다운스케일을 이용한 비디오 부호화 방법에 있어서, 현재 프레임의 고해상도 영상을 저해상도 영상으로 AI 다운스케일하기 위한, 다운스케일 DNN 정보를 결정하는 단계, 상기 다운스케일 DNN 정보에 따라, 상기 현재 프레임의 고해상도 영상을 AI 다운스케일함으로써, 상기 현재 프레임의 저해상도 영상을 생성하는 단계, 상기 현재 프레임의 저해상도 영상의 AI 업스케일에 사용되는, 상기 다운스케일 DNN 정보에 대응되는 업스케일 DNN 정보에 대한, AI 데이터를 생성하는 단계, 상기 현재 프레임의 저해상도 영상을 부호화함으로써, 영상 데이터를 획득하는 단계, 상기 영상 데이터와 상기 AI 데이터를 포함하는 AI 부호화 데이터를 생성하는 단계, 및 상기 AI 부호화 데이터의 영상 데이터가 삽입된 미디어데이터 박스와 상기 AI 부호화 데이터의 AI 데이터가 삽입된 메타데이터 박스를 포함하는 비디오 파일을 출력하는 단계를 포함하는 비디오 부호화 방법이 제공된다.
- [9] 본 개시에 있어서, AI 업스케일을 이용한 비디오 복호화 방법을 수행하는 비디오 복호화 장치에 있어서, 영상 데이터와 상기 영상 데이터의 AI 업스케일에 관한 AI 데이터가 포함된 AI 부호화 데이터를 포함하는 비디오 파일을 수신하는 통신부, 상기 비디오 파일의 메타데이터 박스로부터 상기 AI 부호화 데이터의 AI 데이터를 획득하고, 상기 비디오 파일의 미디어데이터 박스로부터, 상기 AI 부호화 데이터의 영상 데이터를 획득하는 파싱부, 상기 영상 데이터를 복호화하여, 현재 프레임의 저해상도 영상을 복원하는 제1 복호화부, 및 상기 AI 데이터로부터, 상기 현재 프레임의 업스케일 DNN 정보를 획득하고, 상기 현재 프레임의 업스케일 DNN 정보에 따라, 상기 저해상도 영상을 AI 업스케일함으로써, 상기 저해상도 영상에 대응되는 고해상도 영상을 생성하는 AI 업스케일부를 포함하는 비디오 복호화 장치가 제공된다.
- [10] 본 개시에 있어서, AI 다운스케일을 이용한 비디오 부호화 방법을 수행하는 비디오 부호화 장치에 있어서, 현재 프레임의 고해상도 영상을 저해상도 영상으로 AI 다운스케일하기 위한, 다운스케일 DNN 정보를 결정하고, 상기 다운스케일 DNN 정보에 따라, 상기 현재 프레임의 고해상도 영상을 AI 다운스케일함으로써, 상기 현재 프레임의 저해상도 영상을 생성하고, 상기 현재 프레임의 저해상도 영상의 AI 업스케일에 사용되는, AI 데이터를 생성하는 AI 다운스케일부, 상기 현재 프레임의 저해상도 영상을 부호화함으로써, 영상 데이터를 결정하는 제1 부호화부, 상기 영상 데이터와 상기 AI 데이터를

포함하는 AI 부호화 데이터를 결정하는 데이터 처리부, 및 상기 AI 부호화 데이터의 영상 데이터가 삽입된 미디어데이터 박스와 상기 AI 부호화 데이터의 AI 데이터가 삽입된 메타데이터 박스를 포함하는 비디오 파일을 출력하는 통신부를 포함하는 비디오 부호화 장치가 제공된다.

- [11] 본 실시 예가 이루고자 하는 기술적 과제는 상기된 바와 같은 기술적 과제들로 한정되지 않으며, 이하의 실시 예들로부터 또 다른 기술적 과제들이 유추될 수 있다.

### 발명의 효과

- [12] 본 실시 예에서, 저해상도 영상의 AI 업스케일에 필요한 AI 데이터의 데이터 구조와 함께, 상기 데이터 구조에 따른 AI 데이터의 생성 방법 및 파싱 방법이 제공된다. 최적화된 데이터 구조에 따른 AI 데이터에 따라 AI 업스케일의 효율이 향상될 수 있다.

### 도면의 간단한 설명

- [13] 본 명세서에서 인용되는 도면을 보다 충분히 이해하기 위하여 각 도면의 간단한 설명이 제공된다.
- [14] 도 1은 일 실시예에 따른 AI 부호화 과정 및 AI 복호화 과정을 설명하기 위한 도면이다.
- [15] 도 2는 일 실시예에 따른 AI 복호화 장치의 구성을 나타내는 블록도이다.
- [16] 도 3은 제 2 영상의 AI 업스케일을 위한 제 2 DNN을 나타내는 예시적인 도면이다.
- [17] 도 4는 컨볼루션 레이어에 의한 컨볼루션 연산을 설명하기 위한 도면이다.
- [18] 도 5는 여러 영상 관련 정보들과 여러 DNN 설정 정보들 사이의 매핑 관계를 나타내는 예시적인 도면이다.
- [19] 도 6은 복수의 프레임으로 구성된 제 2 영상을 도시하는 도면이다.
- [20] 도 7은 일 실시예에 따른 AI 부호화 장치의 구성을 나타내는 블록도이다.
- [21] 도 8은 원본 영상의 AI 다운스케일을 위한 제 1 DNN을 나타내는 예시적인 도면이다.
- [22] 도 9는 제 1 DNN 및 제 2 DNN을 훈련시키는 방법을 설명하기 위한 도면이다.
- [23] 도 10는 훈련 장치에 의한 제 1 DNN 및 제 2 DNN의 훈련 과정을 설명하기 위한 도면이다.
- [24] 도 11은 원본 영상의 AI 다운스케일을 위한 장치 및 제 2 영상의 AI 업스케일을 위한 장치를 도시하는 예시적인 도면이다.
- [25] 도 12는 영상 데이터와 AI 데이터의 구조와 함께, 영상 데이터와 AI 데이터의 상응 관계를 설명한다.
- [26] 도 13A는, AI 데이터가 영상 데이터에 삽입된 경우, AI 복호화 장치에서 데이터의 흐름을 설명한다.
- [27] 도 13B는, AI 데이터와 영상 데이터가 분리되어 2개의 파일에 포함된 경우, AI

복호화 장치에서 데이터의 흐름을 설명한다.

- [28] 도 14는 단일 파일에서 AI 데이터와 영상 데이터가 분리되어 있는 경우의 AI 부호화 데이터의 일 실시 예를 도시한다.
- [29] 도 15A는 단일 파일에서 AI 데이터가 영상 데이터에 삽입된 경우의 AI 부호화 데이터의 일 실시 예를 도시한다.
- [30] 도 15B는 단일 파일에서 AI 데이터가 영상 데이터에 삽입된 경우의 AI 부호화 데이터의 일 실시 예를 도시한다.
- [31] 도 15C는 단일 파일에서 일부 AI 데이터가 영상 데이터에 삽입되고 나머지 AI 데이터는 영상 데이터와 분리된 경우의 AI 부호화 데이터의 일 실시 예를 도시한다.
- [32] 도 16은 도 14와 같이 AI 데이터와 영상 데이터가 분리된 경우, 비디오 세그먼트 단위로 분할된 AI 부호화 데이터의 일 실시 예를 도시한다.
- [33] 도 17은 2개의 파일로 분리되어 전송되는 AI 데이터와 영상 데이터의 일 실시 예를 도시한다.
- [34] 도 18A는 도 14 내지 도 17에서 설명된 비디오 AI 데이터에 적용될 수 있는 데이터 구조의 일 실시 예를 나타낸다.
- [35] 도 18B는 도 16에서 설명된 비디오 세그먼트 AI 데이터 또는 도 14, 도 15A 내지 15C, 도 17의 프레임 그룹 AI 데이터에 적용될 수 있는 데이터 구조의 일 실시 예를 나타낸다.
- [36] 도 19는 도 18A의 데이터 구조가 구현된 선택스 테이블을 나타낸다.
- [37] 도 20은 도 14 내지 도 17에서 설명된 프레임 그룹 AI 데이터 또는 프레임 AI 데이터에 적용될 수 있는 데이터 구조의 일 실시 예를 나타낸다.
- [38] 도 21은 도 20의 데이터 구조가 구현된 선택스 테이블을 나타낸다.
- [39] 도 22는 AI 복호화기에 따른 영상 복호화 방법의 일 실시예의 순서도이다.
- [40] 도 23은 AI 부호화기에 따른 영상 부호화 방법의 일 실시예의 순서도이다.
- [41] 도 24는 일 실시예에 따른 영상 복호화 장치의 구성을 나타내는 블록도이다.
- [42] 도 25는 일 실시예에 따른 영상 부호화 장치의 구성을 나타내는 블록도이다.

### 발명의 실시를 위한 최선의 형태

- [43] 본 개시에 있어서, AI 부호화 데이터를 포함하는 비디오 파일이 저장된 컴퓨터로 기록가능한 기록 매체에 있어서, 상기 AI 부호화 데이터는, 고해상도 영상을 AI 다운스케일하여 생성된 저해상도 영상의 부호화 정보를 포함하는 영상 데이터, 및 상기 영상 데이터에 따라 복원된 저해상도 영상의 AI 업스케일에 대한 AI 데이터 포함하고, 상기 AI 데이터는, 하나 이상의 프레임에 AI 업스케일이 적용되는지 여부를 나타내는 AI 타겟 데이터, 및 상기 하나 이상의 프레임에 AI 업스케일이 적용될 경우, 미리 설정된 복수의 디폴트 DNN 설정 정보 중에서, 상기 하나 이상의 프레임의 AI 업스케일에 사용되는 업스케일 DNN 정보에 관한 AI 보조 데이터를 포함하는 것을 특징으로 하는 컴퓨터로

기록가능한 기록 매체가 제공된다.

### 발명의 실시를 위한 형태

- [44] 본 개시는 다양한 변경을 가할 수 있고 여러 가지 실시예를 가질 수 있는 바, 특정 실시예들을 도면에 예시하고, 이를 상세한 설명을 통해 상세히 설명하고자 한다. 그러나, 이는 본 개시의 실시 형태에 대해 한정하려는 것이 아니며, 본 개시는 여러 실시예들의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다.
- [45] 실시예를 설명함에 있어서, 관련된 공지 기술에 대한 구체적인 설명이 본 개시의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우 그 상세한 설명을 생략한다. 또한, 명세서의 설명 과정에서 이용되는 숫자(예를 들어, 제 1, 제 2 등)는 하나의 구성요소를 다른 구성요소와 구분하기 위한 식별기호에 불과하다.
- [46] 또한, 본 명세서에서, 일 구성요소가 다른 구성요소와 "연결된다" 거나 "접속된다" 등으로 언급된 때에는, 상기 일 구성요소가 상기 다른 구성요소와 직접 연결되거나 또는 직접 접속될 수도 있지만, 특별히 반대되는 기재가 존재하지 않는 이상, 중간에 또 다른 구성요소를 매개하여 연결되거나 또는 접속될 수도 있다고 이해되어야 할 것이다.
- [47] 또한, 본 명세서에서 '~부(유닛)', '모듈' 등으로 표현되는 구성요소는 2개 이상의 구성요소가 하나의 구성요소로 합쳐지거나 또는 하나의 구성요소가 보다 세분화된 기능별로 2개 이상으로 분화될 수도 있다. 또한, 이하에서 설명할 구성요소 각각은 자신이 담당하는 주기능 이외에도 다른 구성요소가 담당하는 기능 중 일부 또는 전부의 기능을 추가적으로 수행할 수도 있으며, 구성요소 각각이 담당하는 주기능 중 일부 기능이 다른 구성요소에 의해 전담되어 수행될 수도 있음은 물론이다.
- [48] 또한, 본 명세서에서, '영상(image)' 또는 '픽처'는 정지영상, 복수의 연속된 정지영상(또는 프레임)으로 구성된 동영상, 또는 비디오를 나타낼 수 있다.
- [49] 또한, 본 명세서에서 'DNN(deep neural network)'은 뇌 신경을 모사한 인공신경망 모델의 대표적인 예시으로써, 특정 알고리즘을 사용한 인공신경망 모델로 한정되지 않는다.
- [50] 또한, 본 명세서에서 '파라미터'는 뉴럴 네트워크를 이루는 각 레이어의 연산 과정에서 이용되는 값으로서 예를 들어, 입력 값을 소정 연산식에 적용할 때 이용되는 가중치를 포함할 수 있다. 또한, 파라미터는 매트릭스 형태로 표현될 수 있다. 파라미터는 훈련의 결과로 설정되는 값으로서, 필요에 따라 별도의 훈련 데이터(training data)를 통해 갱신될 수 있다.
- [51] 또한, 본 명세서에서 '제 1 DNN'은 영상의 AI 다운스케일을 위해 이용되는 DNN을 의미하고, '제 2 DNN'은 영상의 AI 업스케일을 위해 이용되는 DNN을 의미한다.
- [52] 또한, 본 명세서에서 'DNN 설정 정보'는 DNN을 구성하는 요소와 관련된

정보로서 전술한 파라미터를 포함한다. DNN 설정 정보를 이용하여 제 1 DNN 또는 제 2 DNN이 설정될 수 있다.

[53] 또한, 본 명세서에서 '원본 영상'은 AI 부호화의 대상이 되는 영상을 의미하고, '제 1 영상'은 AI 부호화 과정에서 원본 영상의 AI 다운스케일 결과 획득된 영상을 의미한다. 또한, '제 2 영상'은 AI 복호화 과정에서 제 1 복호화를 통해 획득된 영상을 의미하고, '제 3 영상'은 AI 복호화 과정에서 제 2 영상을 AI 업스케일하여 획득된 영상을 의미한다.

[54] 또한, 본 명세서에서 'AI 다운스케일'은 AI 기반으로 영상의 해상도를 감소시키는 처리를 의미하고, '제 1 부호화'는 주파수 변환 기반의 영상 압축 방법에 의한 부호화 처리를 의미한다. 또한, '제 1 복호화'는 주파수 변환 기반의 영상 복원 방법에 의한 복호화 처리를 의미하고, 'AI 업스케일'은 AI 기반으로 영상의 해상도를 증가시키는 처리를 의미한다.

[55] 도 1은 일 실시예에 따른 AI(artificial intelligence) 부호화 과정 및 AI 복호화 과정을 설명하기 위한 도면이다.

[56] 전술한 바와 같이 영상의 해상도가 급격히 커짐에 따라 부호화/복호화를 위한 정보 처리량이 많아지게 되고, 이에 따라 영상의 부호화 및 복호화 효율을 향상시키기 위한 방안이 필요하다.

[57] 도 1에 도시된 바와 같이, 본 개시의 일 실시예에 따르면, 해상도가 큰 원본 영상(105)을 AI 다운스케일(110)하여 제 1 영상(115)을 획득한다. 그리고, 상대적으로 작은 해상도의 제 1 영상(115)을 대상으로 하여 제 1 부호화(120) 및 제 1 복호화(130)를 수행하므로, 원본 영상(105)을 대상으로 하여 제 1 부호화(120) 및 제 1 복호화(130)를 수행하는 경우에 비해 처리되는 비트레이트를 크게 감소시킬 수 있다.

[58] 도 1을 참조하여 상세히 설명하면, 일 실시예는 AI 부호화 과정에서, 원본 영상(105)을 AI 다운스케일(110)하여 제 1 영상(115)을 획득하고, 제 1 영상(115)을 제 1 부호화(120)한다. AI 복호화 과정에서는, AI 부호화 결과 획득된 AI 데이터와 영상 데이터를 포함하는 AI 부호화 데이터를 수신하고, 제 1 복호화(130)를 통해 제 2 영상(135)을 획득하고, 제 2 영상(135)을 AI 업스케일(140)하여 제 3 영상(145)을 획득한다.

[59] AI 부호화 과정을 좀 더 상세히 살펴보면, 원본 영상(105)을 입력받으면, 소정 해상도 또는 소정 화질의 제 1 영상(115)을 획득하기 위해 원본 영상(105)을 AI 다운스케일(110)한다. 이때, AI 다운스케일(110)은 AI 기반으로 수행되는데, AI 다운스케일(110)을 위한 AI는 제 2 영상(135)의 AI 업스케일(140)을 위한 AI와 연계되어 훈련되어야(joint trained) 한다. 왜냐하면, AI 다운스케일(110)을 위한 AI와 AI 업스케일(140)을 위한 AI가 분리되어 훈련되는 경우, AI 부호화 대상인 원본 영상(105)과 AI 복호화를 통해 복원된 제 3 영상(145) 사이의 차이가 커지게 되기 때문이다.

[60] 본 개시의 실시예에서는, AI 부호화 과정과 AI 복호화 과정에서 이러한 연계

관계를 유지하기 위해, AI 데이터를 이용할 수 있다. 따라서, AI 부호화 과정을 통해 획득된 AI 데이터는 업스케일 타겟을 나타내는 정보를 포함하여야 하고, AI 복호화 과정에서는 AI 데이터에 기초하여 확인되는 업스케일 타겟에 따라 제 2 영상(135)을 AI 업스케일(140)하여야 한다.

[61] AI 다운스케일(110)을 위한 AI 및 AI 업스케일(140)을 위한 AI는 DNN(deep neural network)으로 구현될 수 있다. 도 9를 참조하여 후술하는 바와 같이, 제 1 DNN과 제 2 DNN은 소정 타겟 하에 손실 정보의 공유를 통해 연계 훈련되므로, AI 부호화 장치는 제 1 DNN과 2 DNN이 연계 훈련할 때 이용된 타겟 정보를 AI 복호화 장치로 제공하고, AI 복호화 장치는 제공받은 타겟 정보에 기초하여 제 2 영상(135)을 타겟하는 해상도로 AI 업스케일(140)할 수 있다.

[62] 도 1에 도시된 제 1 부호화(120) 및 제 1 복호화(130)에 대해 상세히 설명하면, 원본 영상(105)으로부터 AI 다운스케일(110)된 제 1 영상(115)은 제 1 부호화(120)를 통해 정보량이 감축될 수 있다. 제 1 부호화(120)는, 제 1 영상(115)을 예측하여 예측 데이터를 생성하는 과정, 제 1 영상(115)과 예측 데이터 사이의 차이에 해당하는 잔차 데이터를 생성하는 과정, 공간 영역 성분인 잔차 데이터를 주파수 영역 성분으로 변환(transformation)하는 과정, 주파수 영역 성분으로 변환된 잔차 데이터를 양자화(quantization)하는 과정 및 양자화된 잔차 데이터를 엔트로피 부호화하는 과정 등을 포함할 수 있다. 이와 같은 제 1 부호화 과정(120)은 MPEG-2, H.264 AVC(Advanced Video Coding), MPEG-4, HEVC(High Efficiency Video Coding), VC-1, VP8, VP9 및 AV1(AOMedia Video 1) 등 주파수 변환을 이용한 영상 압축 방법 중의 하나를 통해 구현될 수 있다.

[63] 제 1 영상(115)에 대응하는 제 2 영상(135)은 영상 데이터의 제 1 복호화(130)를 통해 복원될 수 있다. 제 1 복호화(130)는, 영상 데이터를 엔트로피 복호화하여 양자화된 잔차 데이터를 생성하는 과정, 양자화된 잔차 데이터를 역양자화하는 과정, 주파수 영역 성분의 잔차 데이터를 공간 영역 성분으로 변환하는 과정, 예측 데이터를 생성하는 과정 및 예측 데이터와 잔차 데이터를 이용하여 제 2 영상(135)을 복원하는 과정 등을 포함할 수 있다. 이와 같은 제 1 복호화(130) 과정은 제 1 부호화(120) 과정에서 사용된 MPEG-2, H.264, MPEG-4, HEVC, VC-1, VP8, VP9 및 AV1 등 주파수 변환을 이용한 영상 압축 방법 중의 하나에 대응되는 영상 복원 방법을 통해 구현될 수 있다.

[64] AI 부호화 과정을 통해 획득된 AI 부호화 데이터는, 제 1 영상(115)의 제 1 부호화(120) 결과 획득된 영상 데이터 및 원본 영상(105)의 AI 다운스케일(110)과 관련된 AI 데이터를 포함할 수 있다. 영상 데이터는 제 1 복호화(130) 과정에서 이용될 수 있으며, AI 데이터는 AI 업스케일(140) 과정에서 이용될 수 있다.

[65] 영상 데이터는 비트스트림 형태로 전송될 수 있다. 영상 데이터는 제 1 영상(115) 내 픽셀 값들에 기초하여 획득되는 데이터, 예를 들어, 제 1 영상(115)과 제 1 영상(115)의 예측 데이터 사이의 차이인 잔차 데이터를 포함할 수 있다. 또한, 영상 데이터는 제 1 영상(115)의 제 1 부호화(120) 과정에서 이용된

정보들을 포함한다. 예를 들어, 영상 데이터는 제 1 영상(115)을 제 1 부호화(120)하는데 이용된 예측 모드(mode) 정보, 움직임 정보, 및 제 1 부호화(120)에서 이용된 양자화 파라미터 관련 정보 등을 포함할 수 있다. 영상 데이터는 MPEG-2, H.264 AVC, MPEG-4, HEVC, VC-1, VP8, VP9 및 AV1 등 주파수 변환을 이용하는 영상 압축 방법 중 제 1 부호화(120) 과정에서 이용된 영상 압축 방법의 규칙, 예를 들어, 신택스(syntax)에 따라 생성될 수 있다.

[66] AI 데이터는 제 2 DNN에 기반한 AI 업스케일(140)에 이용된다. 전술한 바와 같이, 제 1 DNN과 제 2 DNN은 연계 훈련되기 때문에, AI 데이터는 제 2 DNN을 통한 제 2 영상(135)의 정확한 AI 업스케일(140)이 수행될 수 있게 하는 정보를 포함한다. AI 복호화 과정에서는 AI 데이터에 기반하여 제 2 영상(135)을 타겟하는 해상도 및/또는 화질로 AI 업스케일(140)할 수 있다.

[67] AI 데이터는 비트스트림의 형태로 영상 데이터와 함께 전송될 수 있다. 또는, 구현예에 따라, AI 데이터는 프레임이나 패킷 형태로 영상 데이터와 구분되어 전송될 수도 있다. AI 부호화 결과 획득된 영상 데이터와 AI 데이터는 동일한 네트워크 또는 서로 상이한 네트워크를 통해 전송될 수 있다.

[68] 도 2는 일 실시예에 따른 AI 복호화 장치(200)의 구성을 도시하는 블록도이다.

[69] 도 2를 참조하면, 일 실시예에 따른 AI 복호화 장치(200)는 수신부(210) 및 AI 복호화부(230)를 포함할 수 있다. 수신부(210)는 통신부(212), 파싱부(214) 및 출력부(216)를 포함할 수 있다. AI 복호화부(230)는 제 1 복호화부(232) 및 AI 업스케일부(234)를 포함할 수 있다.

[70] 수신부(210)는 AI 부호화 결과 획득된 AI 부호화 데이터를 수신 및 파싱하고, 영상 데이터와 AI 데이터를 구분하여 AI 복호화부(230)로 출력한다.

[71] 구체적으로, 통신부(212)는 네트워크를 통해 AI 부호화 결과 획득된 AI 부호화 데이터를 수신한다. AI 부호화 결과 획득된 AI 부호화 데이터는 영상 데이터와 AI 데이터를 포함한다. 영상 데이터와 AI 데이터는 동종 네트워크 또는 이종 네트워크를 통해 수신될 수 있다.

[72] 파싱부(214)는 통신부(212)를 통해 수신된 AI 부호화 데이터를 전달받아 파싱(parsing)하여 영상 데이터와 AI 데이터로 구분한다. 예를 들어, 통신부(212)로부터 획득된 데이터의 헤더를 읽어, 해당 데이터가 영상 데이터인지 또는 AI 데이터인지를 구분할 수 있다. 일 예에서, 파싱부(214)는 통신부(212)를 통해 수신된 데이터의 헤더를 통해 영상 데이터와 AI 데이터를 구분하여 출력부(216)로 전달하고, 출력부(216)는 각각의 구분된 데이터를 제 1 복호화부(232) 및 AI 업스케일부(234)로 전달한다. 이 때, AI 부호화 데이터에 포함된 영상 데이터가 소정의 코덱(예를 들어, MPEG-2, H.264, MPEG-4, HEVC, VC-1, VP8, VP9 또는 AV1)을 통해 획득된 영상 데이터인 것으로 확인할 수도 있다. 이 경우, 영상 데이터가 상기 확인된 코덱으로 처리될 수 있도록, 출력부(216)를 통해 해당 정보를 제 1 복호화부(232)로 전달할 수 있다.

[73] 일 실시예에서, 파싱부(214)가 파싱하는 AI 부호화 데이터는, 하드 디스크,

플로피 디스크 및 자기 테이프와 같은 자기 매체, CD-ROM 및 DVD와 같은 광기록 매체, 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical medium) 등을 포함하는 데이터 저장 매체로부터 획득된 것일 수도 있다.

[74] 제 1 복호화부(232)는 영상 데이터에 기초하여 제 1 영상(115)에 대응하는 제 2 영상(135)을 복원한다. 제 1 복호화부(232)에 의해 획득된 제 2 영상(135)은 AI 업스케일부(234)로 제공된다. 구현예에 따라서는, 영상 데이터에 포함된 예측 모드 정보, 움직임 정보, 양자화 파라미터 정보 등의 제 1 복호화 관련 정보가 AI 업스케일부(234)로 더 제공될 수 있다.

[75] AI 데이터를 수신한 AI 업스케일부(234)는 AI 데이터에 기초하여 제 2 영상(135)을 AI 업스케일한다. 구현예에 따라서는, 영상 데이터에 포함된 예측 모드 정보, 양자화 파라미터 정보 등의 제 1 복호화 관련 정보를 더 이용하여 AI 업스케일 할 수 있다.

[76] 일 실시예에 따른 수신부(210) 및 AI 복호화부(230)는 개별적인 장치로 설명되었으나, 하나의 프로세서를 통해 구현될 수 있다. 이 경우, 수신부(210) 및 AI 복호화부(230)는 전용 프로세서로 구현될 수도 있고, AP 또는 CPU, GPU와 같은 범용 프로세서와 S/W의 조합을 통해 구현될 수도 있다. 또한, 전용 프로세서의 경우, 본 개시의 실시예를 구현하기 위한 메모리를 포함하여 구현되거나, 외부 메모리를 이용하기 위한 메모리 처리부를 포함하여 구현될 수 있다.

[77] 또한, 수신부(210) 및 AI 복호화부(230)는 복수의 프로세서로 구성될 수도 있다. 이 경우, 전용 프로세서들의 조합으로 구현될 수도 있고, AP 또는 CPU, GPU와 같은 다수의 범용 프로세서들과 S/W의 조합을 통해 구현될 수도 있다. 마찬가지로 AI 업스케일부(234)와 제 1 복호화부(232)도 각각 서로 다른 프로세서로 구현될 수 있다.

[78] AI 업스케일부(234)로 제공되는 AI 데이터는, 제 2 영상(135)을 AI 업스케일할 수 있게 하는 정보들을 포함한다. 이때, 업스케일 타겟은 제 1 DNN의 다운스케일에 대응하여야 한다. 따라서, AI 데이터는 제 1 DNN의 다운스케일 타겟을 확인할 수 있는 정보가 포함되어야 한다.

[79] AI 데이터에 포함된 정보를 구체적으로 예시하면, 원본 영상(105)의 해상도와 제 1 영상(115)의 해상도의 차이 정보, 제 1 영상(115) 관련 정보가 있다.

[80] 차이 정보는, 원본 영상(105) 대비 제 1 영상(115)의 해상도 변환 정도에 대한 정보(예를 들어, 해상도 변환률 정보)로 표현될 수 있다. 그리고, 복원된 제 2 영상(135)의 해상도를 통해 제 1 영상(115)의 해상도를 알게 되고 이를 통해 해상도 변환 정도를 확인할 수 있기 때문에, 차이 정보는 원본 영상(105)의 해상도 정보만으로 표현될 수도 있다. 여기서 해상도 정보는 가로/세로의 화면 사이즈로 표현될 수도 있고, 비율(16:9, 4:3 등)과 한 축의 사이즈로 표현될 수도 있다. 또한, 기 설정된 해상도 정보가 있는 경우는 인덱스 또는 플래그의 형태로 표현될 수도 있다.

- [81] 그리고, 제 1 영상(115) 관련 정보는, 제 1 영상(115)의 제 1 부호화 결과 획득된 영상 데이터의 비트레이트 및 제 1 영상(115)의 제 1 부호화시 이용된 코덱 타입 중 적어도 하나에 대한 정보를 포함할 수 있다.
- [82] AI 업스케일부(234)는 AI 데이터에 포함된 차이 정보 및 제 1 영상(115) 관련 정보 중 적어도 하나에 기초하여 제 2 영상(135)의 업스케일 타겟을 결정할 수 있다. 업스케일 타겟은 예를 들어, 제 2 영상(135)을 어느 정도의 해상도로 업스케일하여야 하는지를 나타낼 수 있다. AI 업스케일부(234)는 업스케일 타겟이 결정되면, 업스케일 타겟에 대응하는 제 3 영상(145)을 획득하기 위해 제 2 DNN을 통해 제 2 영상(135)을 AI 업스케일한다.
- [83] AI 업스케일부(234)가 업스케일 타겟에 맞춰 제 2 영상(135)을 AI 업스케일하는 방법을 설명하기에 앞서, 제 2 DNN을 통한 AI 업스케일 과정에 대해 도 3 및 도 4를 참조하여 설명한다.
- [84] 도 3은 제 2 영상(135)의 AI 업스케일을 위한 제 2 DNN(300)을 나타내는 예시적인 도면이고, 도 4는 도 3에 도시된 제 1 컨볼루션 레이어(310)에서의 컨볼루션 연산을 도시하고 있다.
- [85] 도 3에 도시된 바와 같이, 제 2 영상(135)은 제 1 컨볼루션 레이어(310)로 입력된다. 도 3에 도시된 제 1 컨볼루션 레이어(310)에 표시된 3X3X4는 3 x 3의 크기의 4개의 필터 커널을 이용하여 1개의 입력 영상에 대해 컨볼루션 처리를 하는 것을 예시한다. 컨볼루션 처리 결과 4개의 필터 커널에 의해 4개의 특징 맵이 생성된다. 각 특징 맵은 제 2 영상(135)의 고유한 특성들을 나타낸다. 예를 들어, 각 특징 맵은 제 2 영상(135)의 수직 방향 특성, 수평 방향 특성 또는 에지 특성 등을 나타낼 수 있다.
- [86] 도 4를 참조하여, 제 1 컨볼루션 레이어(310)에서의 컨볼루션 연산에 대해 상세히 설명한다.
- [87] 제 1 컨볼루션 레이어(310)에서 이용되는 3 X 3의 크기를 갖는 필터 커널(430)의 파라미터들과 그에 대응하는 제 2 영상(135) 내 픽셀 값들 사이의 곱 연산 및 덧셈 연산을 통해 하나의 특징 맵(450)이 생성될 수 있다. 제 1 컨볼루션 레이어(310)에서는 4개의 필터 커널이 이용되므로, 4개의 필터 커널을 이용한 컨볼루션 연산 과정을 통해 4개의 특징 맵이 생성될 수 있다.
- [88] 도 4에서 제 2 영상(135)에 표시된 I1 내지 I49는 제 2 영상(135)의 픽셀들을 나타내고, 필터 커널(430)에 표시된 F1 내지 F9는 필터 커널(430)의 파라미터들을 나타낸다. 또한, 특징 맵(450)에 표시된 M1 내지 M9는 특징 맵(450)의 샘플들을 나타낸다.
- [89] 도 4는 제 2 영상(135)이 49개의 픽셀을 포함하는 것으로 예시하고 있으나, 이는 하나의 예시일 뿐이며, 제 2 영상(135)이 4K의 해상도를 갖는 경우, 예를 들어, 3840 X 2160개의 픽셀을 포함할 수 있다.
- [90] 컨볼루션 연산 과정에서, 제 2 영상(135)의 I1, I2, I3, I8, I9, I10, I15, I16, I17의 픽셀 값들 각각과 필터 커널(430)의 F1, F2, F3, F4, F5, F6, F7, F8 및 F9 각각의 곱

- 연산이 수행되고, 곱 연산의 결과 값들을 조합(예를 들어, 덧셈 연산)한 값이 특징 맵(450)의 M1의 값으로 할당될 수 있다. 컨볼루션 연산의 스트라이드(stride)가 2라면, 제 2 영상(135)의 I3, I4, I5, I10, I11, I12, I17, I18, I19의 픽셀 값들 각각과 필터 커널(430)의 F1, F2, F3, F4, F5, F6, F7, F8 및 F9 각각의 곱 연산이 수행되고, 곱 연산의 결과 값들을 조합한 값이 특징 맵(450)의 M2의 값으로 할당될 수 있다.
- [91] 필터 커널(430)이 제 2 영상(135)의 마지막 픽셀에 도달할 때까지 스트라이드에 따라 이동하는 동안 제 2 영상(135) 내 픽셀 값들과 필터 커널(430)의 파라미터들 사이의 컨볼루션 연산이 수행됨으로써, 소정 크기를 갖는 특징 맵(450)이 획득될 수 있다.
- [92] 본 개시에 따르면, 제 1 DNN과 제 2 DNN의 연계 훈련을 통해 제 2 DNN의 파라미터들, 예를 들어, 제 2 DNN의 컨볼루션 레이어들에서 이용되는 필터 커널의 파라미터들(예를 들어, 필터 커널(430)의 F1, F2, F3, F4, F5, F6, F7, F8 및 F9)의 값이 최적화될 수 있다. 전술한 바와 같이, AI 업스케일부(234)는 AI 데이터에 기초하여, 제 1 DNN의 다운스케일 타겟에 대응하는 업스케일 타겟을 결정하고, 결정된 업스케일 타겟에 대응하는 파라미터들을 제 2 DNN의 컨볼루션 레이어들에서 이용되는 필터 커널의 파라미터들로 결정할 수 있다.
- [93] 제 1 DNN 및 제 2 DNN에 포함된 컨볼루션 레이어들은 도 4와 관련하여 설명한 컨볼루션 연산 과정에 따른 처리를 할 수 있으나, 도 4에서 설명한 컨볼루션 연산 과정은 하나의 예시일 뿐이며, 이에 한정되는 것은 아니다.
- [94] 다시 도 3을 참조하면, 제 1 컨볼루션 레이어(310)에서 출력된 특징 맵들은 제 1 활성화 레이어(320)로 입력된다.
- [95] 제 1 활성화 레이어(320)는 각각의 특징 맵에 대해 비선형(Non-linear) 특성을 부여할 수 있다. 제 1 활성화 레이어(320)는 시그모이드 함수(sigmoid function), Tanh 함수, ReLU(Rectified Linear Unit) 함수 등을 포함할 수 있으나, 이에 한정되는 것은 아니다.
- [96] 제 1 활성화 레이어(320)에서 비선형 특성을 부여하는 것은, 제 1 컨볼루션 레이어(310)의 출력인, 특징 맵의 일부 샘플 값을 변경하여 출력하는 것을 의미한다. 이때, 변경은 비선형 특성을 적용하여 수행된다.
- [97] 제 1 활성화 레이어(320)는 제 1 컨볼루션 레이어(310)로부터 출력되는 특징 맵들의 샘플 값들을 제 2 컨볼루션 레이어(330)로 전달할지 여부를 결정한다. 예를 들어, 특징 맵들의 샘플 값들 중 어떤 샘플 값들은 제 1 활성화 레이어(320)에 의해 활성화되어 제 2 컨볼루션 레이어(330)로 전달되고, 어떤 샘플 값들은 제 1 활성화 레이어(320)에 의해 비활성화되어 제 2 컨볼루션 레이어(330)로 전달되지 않는다. 특징 맵들이 나타내는 제 2 영상(135)의 고유 특성이 제 1 활성화 레이어(320)에 의해 강조된다.
- [98] 제 1 활성화 레이어(320)에서 출력된 특징 맵들(325)은 제 2 컨볼루션 레이어(330)로 입력된다. 도 3에 도시된 특징 맵들(325) 중 어느 하나는 도 4와 관련하여 설명한 특징 맵(450)이 제 1 활성화 레이어(320)에서 처리된 결과이다.

- [99] 제 2 컨볼루션 레이어(330)에 표시된 3X3X4는 3 x 3의 크기의 4개의 필터 커널을 이용하여 입력된 특징 맵들(325)에 대해 컨볼루션 처리하는 것을 예시한다. 제 2 컨볼루션 레이어(330)의 출력은 제 2 활성화 레이어(340)로 입력된다. 제 2 활성화 레이어(340)는 입력 데이터에 대해 비선형 특성을 부여할 수 있다.
- [100] 제 2 활성화 레이어(340)에서 출력된 특징 맵들(345)은 제 3 컨볼루션 레이어(350)로 입력된다. 도 3에 도시된 제 3 컨볼루션 레이어(350)에 표시된 3X3X1는 3 x 3의 크기의 1개의 필터 커널을 이용하여 1개의 출력 영상을 만들기 위해 컨볼루션 처리를 하는 것을 예시한다. 제 3 컨볼루션 레이어(350)는 최종 영상을 출력하기 위한 레이어로서 1개의 필터 커널을 이용하여 1개의 출력을 생성한다. 본 개시의 예시에 따르면, 제 3 컨볼루션 레이어(350)는 컨볼루션 연산 결과를 통해 제 3 영상(145)을 출력할 수 있다.
- [101] 제 2 DNN(300)의 제 1 컨볼루션 레이어(310), 제 2 컨볼루션 레이어(330) 및 제 3 컨볼루션 레이어(350)의 필터 커널의 개수, 필터 커널의 파라미터 등을 나타내는 DNN 설정 정보는 후술하는 바와 같이 복수 개일 수 있는데, 복수의 DNN 설정 정보는 제 1 DNN의 복수의 DNN 설정 정보와 연계되어야 한다. 제 2 DNN의 복수의 DNN 설정 정보와 제 1 DNN의 복수의 DNN 설정 정보 사이의 연계는, 제 1 DNN 및 제 2 DNN의 연계 학습을 통해 구현될 수 있다.
- [102] 도 3은 제 2 DNN(300)이 세 개의 컨볼루션 레이어(310, 330, 350)와 두 개의 활성화 레이어(320, 340)를 포함하고 있는 것으로 도시하고 있으나, 이는 하나의 예시일 뿐이며, 구현예에 따라서, 컨볼루션 레이어 및 활성화 레이어의 개수는 다양하게 변경될 수 있다. 또한, 구현예에 따라서, 제 2 DNN(300)은 RNN(recurrent neural network)을 통해 구현될 수도 있다. 이 경우는 본 개시의 예시에 따른 제 2 DNN(300)의 CNN 구조를 RNN 구조로 변경하는 것을 의미한다.
- [103] 일 실시예에서, AI 업스케일부(234)는 전술한 컨볼루션 연산 및 활성화 레이어의 연산을 위한 적어도 하나의 ALU(Arithmetic logic unit)를 포함할 수 있다. ALU는 프로세서로 구현될 수 있다. 컨볼루션 연산을 위해, ALU는 제 2 영상(135) 또는 이전 레이어에서 출력된 특징 맵의 샘플 값들과 필터 커널의 샘플 값들 사이의 곱 연산을 수행하는 곱셈기 및 곱셈의 결과 값들을 더하는 가산기를 포함할 수 있다. 또한, 활성화 레이어의 연산을 위해, ALU는 미리 결정된 시그모이드 함수, Tanh 함수 또는 ReLU 함수 등에서 이용되는 가중치를 입력된 샘플 값에 곱하는 곱셈기, 및 곱한 결과와 소정 값을 비교하여 입력된 샘플 값을 다음 레이어로 전달할지를 판단하는 비교기를 포함할 수 있다.
- [104] 이하에서, AI 업스케일부(234)가 업스케일 타겟에 맞춰 제 2 영상(135)을 AI 업스케일하는 방법을 설명한다.
- [105] 일 실시예에서, AI 업스케일부(234)는 제 2 DNN에 세팅 가능한 복수의 DNN 설정 정보를 저장할 수 있다.

- [106] 여기서, DNN 설정 정보는 제 2 DNN에 포함되는 컨볼루션 레이어의 수, 컨볼루션 레이어별 필터 커널의 개수 및 각 필터 커널의 파라미터 중 적어도 하나에 대한 정보를 포함할 수 있다. 복수의 DNN 설정 정보는 다양한 업스케일 타겟에 각각 대응될 수 있으며, 특정 업스케일 타겟에 대응되는 DNN 설정 정보에 기반하여 제 2 DNN이 동작할 수 있다. DNN 설정 정보에 따라 제 2 DNN이 서로 다른 구조를 가질 수 있다. 예를 들어, 어느 DNN 설정 정보에 따라 제 2 DNN이 3개의 컨볼루션 레이어를 포함할 수 있고, 다른 DNN 설정 정보에 따라 제 2 DNN이 4개의 컨볼루션 레이어를 포함할 수 있다.
- [107] 일 실시예에서, DNN 설정 정보는 제 2 DNN에서 사용되는 필터 커널의 파라미터만을 포함할 수도 있다. 이 경우, 제 2 DNN의 구조는 변경되지 않는 대신, DNN 설정 정보에 따라 내부의 필터 커널의 파라미터만이 달라질 수 있다.
- [108] AI 업스케일부(234)는 복수의 DNN 설정 정보 중 제 2 영상(135)의 AI 업스케일을 위한 DNN 설정 정보를 획득할 수 있다. 여기서 사용되는 복수의 DNN 설정 정보 각각은 미리 결정된 해상도 및/또는 미리 결정된 화질의 제 3 영상(145)을 획득하기 위한 정보로, 제 1 DNN과 연계하여 훈련된 것이다.
- [109] 예를 들어, 복수의 DNN 설정 정보 중 어느 하나의 DNN 설정 정보는 제 2 영상(135)의 해상도보다 2배 큰 해상도의 제 3 영상(145), 예를 들어, 2K (2048\*1080)의 제 2 영상(135)보다 2배 큰 4K(4096\*2160)의 제 3 영상(145)을 획득하기 위한 정보들을 포함할 수 있고, 다른 하나의 DNN 설정 정보는 제 2 영상(135)의 해상도보다 4배 큰 해상도의 제 3 영상(145), 예를 들어, 2K (2048\*1080)의 제 2 영상(135)보다 4배 큰 8K(8192\*4320)의 제 3 영상(145)을 획득하기 위한 정보들을 포함할 수 있다.
- [110] 복수의 DNN 설정 정보 각각은 AI 부호화 장치(600)의 제 1 DNN의 DNN 설정 정보와 연계되어 만들어진 것이며, AI 업스케일부(234)는 제 1 DNN의 DNN 설정 정보의 축소 비율에 대응되는 확대 비율에 따라 복수의 DNN 설정 정보 중 하나의 DNN 설정 정보를 획득한다. 이를 위해, AI 업스케일부(234)는 제 1 DNN의 정보를 확인하여야 한다. 제 1 DNN의 정보를 AI 업스케일부(234)가 확인하기 위해, 일 실시예에 따른 AI 복호화 장치(200)는 AI 부호화 장치(600)로부터 제 1 DNN의 정보를 포함하는 AI 데이터를 수신한다.
- [111] 다시 말하면, AI 업스케일부(234)는 AI 부호화 장치(600)로부터 수신되는 정보들을 이용하여, 제 1 영상(115)을 획득하기 위해 이용된 제 1 DNN의 DNN 설정 정보가 타겟하는 정보를 확인하고, 그와 연계 훈련된 제 2 DNN의 DNN 설정 정보를 획득할 수 있는 것이다.
- [112] 복수의 DNN 설정 정보 중 제 2 영상(135)의 AI 업스케일을 위한 DNN 설정 정보가 획득되면, 획득된 DNN 설정 정보에 따라 동작하는 제 2 DNN에 기초하여 입력 데이터가 처리될 수 있다.
- [113] 예를 들어, 어느 하나의 DNN 설정 정보가 획득되면, 도 3에 도시된 제 2 DNN(300)의 제 1 컨볼루션 레이어(310), 제 2 컨볼루션 레이어(330) 및 제 3

컨볼루션 레이어(350) 각각에 대해서, 각 레이어에 포함되는 필터 커널의 개수와 필터 커널의 파라미터들을, 상기 획득된 DNN 설정 정보에 포함된 값으로 설정한다.

- [114] 구체적으로, 도 4에 도시된 제 2 DNN의 어느 하나의 컨볼루션 레이어에서 이용되는 3 X 3의 필터 커널의 파라미터들이 {1, 1, 1, 1, 1, 1, 1, 1, 1}로 세팅되도록 하고, 이후 DNN 설정 정보의 변경이 있는 경우, 이를 변경된 DNN 설정 정보에 포함된 파라미터들인 {2, 2, 2, 2, 2, 2, 2, 2, 2}로 교체할 수 있는 것이다.
- [115] AI 업스케일부(234)는 AI 데이터에 포함된 정보에 기초하여 복수의 DNN 설정 정보 중 제 2 영상(135)을 업스케일하기 위한 DNN 설정 정보를 획득할 수 있는데, DNN 설정 정보를 획득하는데 이용되는 AI 데이터에 대해 구체적으로 설명한다.
- [116] 일 실시예에서, AI 업스케일부(234)는 AI 데이터에 포함된 차이 정보에 기초하여, 복수의 DNN 설정 정보 중 제 2 영상(135)을 업스케일하기 위한 DNN 설정 정보를 획득할 수 있다. 예를 들어, 차이 정보에 기초하여 원본 영상(105)의 해상도(예를 들어, 4K(4096\*2160))가 제 1 영상(115)의 해상도(예를 들어, 2K(2048\*1080))보다 2배 큰 것으로 확인된 경우, AI 업스케일부(234)는 제 2 영상(135)의 해상도를 2배 증가시킬 수 있는 DNN 설정 정보를 획득할 수 있다.
- [117] 다른 실시예에서, AI 업스케일부(234)는 AI 데이터에 포함된 제 1 영상(115) 관련 정보에 기초하여, 복수의 DNN 설정 정보 중 제 2 영상(135)을 AI 업스케일하기 위한 DNN 설정 정보를 획득할 수 있다. AI 업스케일부(234)는 영상 관련 정보들과 DNN 설정 정보들 사이의 매핑 관계를 미리 결정하고, 제 1 영상(115) 관련 정보에 매핑된 DNN 설정 정보를 획득할 수 있다.
- [118] 도 5는 여러 영상 관련 정보들과 여러 DNN 설정 정보들 사이의 매핑 관계를 나타내는 예시적인 도면이다.
- [119] 도 5에 따른 실시예를 통해, 본 개시의 일 실시예의 AI 부호화/AI 복호화 과정은 해상도의 변화만을 고려하는 것이 아님을 알 수 있다. 도 5에 도시된 바와 같이, SD, HD, Full HD와 같은 해상도, 10Mbps, 15Mbps, 20Mbps와 같은 비트레이트, 그리고 AV1, H.264, HEVC와 같은 코덱 정보를 개별적으로 또는 모두 고려하여 DNN 설정 정보의 선택이 이루어질 수 있다. 이러한 고려를 위해서는 AI 훈련 과정에서 각각의 요소들을 고려한 훈련이 부호화, 복호화 과정에 연계되어 이루어져야 한다(도 9 참조).
- [120] 따라서, 훈련 내용에 따라서 도 5에 도시된 바와 같이, 코덱 타입, 영상의 해상도 등을 포함하는 영상 관련 정보에 기반하여 복수의 DNN 설정 정보가 구비된 경우, AI 복호화 과정에서 전달받는 제 1 영상(115) 관련 정보에 기반하여 제 2 영상(135)의 AI 업스케일을 위한 DNN 설정 정보를 획득할 수 있다.
- [121] 즉, AI 업스케일부(234)는 도 5에 도시된 표의 좌측에 도시된 영상 관련 정보와 표 우측의 DNN 설정 정보를 매칭하고 있음으로써, 영상 관련 정보에 따른 DNN

설정 정보를 사용할 수 있는 것이다.

- [122] 도 5에 예시된 바와 같이, 제 1 영상(115) 관련 정보로부터 제 1 영상(115)의 해상도가 SD이고, 제 1 영상(115)의 제 1 부호화 결과 획득된 영상 데이터의 비트레이트가 10Mbps이고, 제 1 영상(115)이 AV1 코덱으로 제 1 부호화된 것으로 확인되면, AI 업스케일부(234)는 복수의 DNN 설정 정보 중 A DNN 설정 정보를 사용할 수 있다.
- [123] 또한, 제 1 영상(115) 관련 정보로부터 제 1 영상(115)의 해상도가 HD이고, 제 1 부호화 결과 획득된 영상 데이터의 비트레이트가 15Mbps이고, 제 1 영상(115)이 H.264 코덱으로 제 1 부호화된 것으로 확인되면, AI 업스케일부(234)는 복수의 DNN 설정 정보 중 B DNN 설정 정보를 사용할 수 있다.
- [124] 또한, 제 1 영상(115) 관련 정보로부터 제 1 영상(115)의 해상도가 Full HD이고, 제 1 영상(115)의 제 1 부호화 결과 획득된 영상 데이터의 비트레이트가 20Mbps이고, 제 1 영상(115)이 HEVC 코덱으로 제 1 부호화된 것으로 확인되면, AI 업스케일부(234)는 복수의 DNN 설정 정보 중 C DNN 설정 정보를 사용하고, 제 1 영상(115)의 해상도가 Full HD이고, 제 1 영상(115)의 제 1 부호화 결과 획득된 영상 데이터의 비트레이트가 15Mbps이고, 제 1 영상(115)이 HEVC 코덱으로 제 1 부호화된 것으로 확인되면, AI 업스케일부(234)는 복수의 DNN 설정 정보 중 D DNN 설정 정보를 사용할 수 있다. 제 1 영상(115)의 제 1 부호화 결과 획득된 영상 데이터의 비트레이트가 20Mbps인지 또는 15Mbps인지에 따라 C DNN 설정 정보와 D DNN 설정 정보 중에서 어느 하나가 선택된다. 동일 해상도의 제 1 영상(115)을 동일 코덱으로 제 1 부호화하였을 때, 영상 데이터의 비트레이트가 서로 상이하다는 것은, 복원되는 영상의 화질이 서로 상이하다는 것을 의미한다. 따라서, 제 1 DNN과 제 2 DNN은 소정 화질에 기반하여 연계 훈련될 수 있으며, 이에 따라 AI 업스케일부(234)는 제 2 영상(135)의 화질을 나타내는, 영상 데이터의 비트레이트에 따라 DNN 설정 정보를 획득할 수 있다.
- [125] 또 다른 실시예에서, AI 업스케일부(234)는 제 1 복호화부(232)로부터 제공되는 정보(예측 모드 정보, 움직임 정보, 양자화 파라미터 정보 등)와 AI 데이터에 포함된 제 1 영상(115) 관련 정보를 모두 고려하여 복수의 DNN 설정 정보 중 제 2 영상(135)을 AI 업스케일하기 위한 DNN 설정 정보를 획득할 수도 있다. 예를 들어, AI 업스케일부(234)는 제 1 복호화부(232)로부터 제 1 영상(115)의 제 1 부호화 과정에서 이용된 양자화 파라미터 정보를 전달받고, AI 데이터로부터 제 1 영상(115)의 부호화 결과 획득된 영상 데이터의 비트레이트를 확인하고, 양자화 파라미터 및 비트레이트에 대응하는 DNN 설정 정보를 획득할 수도 있다. 동일한 비트레이트라도 영상의 복잡도에 따라 복원 영상의 화질 정도에 차이가 있을 수 있으며, 비트레이트는 제 1 부호화되는 제 1 영상(115) 전체를 대표하는 값으로서 제 1 영상(115) 내에서도 각 프레임의 화질이 서로 다를 수 있다. 따라서, 제 1 복호화부(232)로부터 각 프레임 별로 획득할 수 있는 예측 모드 정보, 움직임 정보 및/또는 양자화 파라미터를 함께 고려하면, AI 데이터만

이용하는 것에 비해 제 2 영상(135)에 보다 적합한 DNN 설정 정보를 획득할 수 있다.

- [126] 또한, 구현예에 따라, AI 데이터는 상호 약속된 DNN 설정 정보의 식별자를 포함할 수 있다. DNN 설정 정보의 식별자는, 제 1 DNN의 다운스케일 타겟에 대응하는 업스케일 타겟으로, 제 2 영상(135)을 AI 업스케일할 수 있도록, 제 1 DNN과 제 2 DNN 간 연계 훈련된 DNN 설정 정보의 쌍을 구분하기 위한 정보이다. AI 업스케일부(234)는 AI 데이터에 포함된 DNN 설정 정보의 식별자를 획득한 후, DNN 설정 정보의 식별자에 대응하는 DNN 설정 정보를 사용하여 제 2 영상(135)을 AI 업스케일할 수 있다. 예를 들어, 제 1 DNN에 설정 가능한 복수의 DNN 설정 정보 각각을 가리키는 식별자와 제 2 DNN에 설정 가능한 복수의 DNN 설정 정보 각각을 가리키는 식별자가 미리 지정되어 있을 수 있다. 이 경우, 제 1 DNN 및 제 2 DNN 각각에 설정 가능한 DNN 설정 정보의 쌍에 대해 동일한 식별자가 지정될 수 있다. AI 데이터는 원본 영상(105)의 AI 다운스케일을 위해 제 1 DNN에 설정된 DNN 설정 정보의 식별자를 포함할 수 있다. AI 데이터를 수신한 AI 업스케일부(234)는 복수의 DNN 설정 정보 중 AI 데이터에 포함된 식별자가 가리키는 DNN 설정 정보를 이용하여 제 2 영상(135)을 AI 업스케일할 수 있다.
- [127] 또한, 구현예에 따라, AI 데이터는 DNN 설정 정보를 포함할 수도 있다. AI 업스케일부(234)는 AI 데이터에 포함된 DNN 설정 정보를 획득한 후, 해당 DNN 설정 정보를 사용하여 제 2 영상(135)을 AI 업스케일할 수도 있다.
- [128] 구현예에 따라, DNN 설정 정보를 구성하는 정보들(예를 들어, 컨볼루션 레이어의 수, 컨볼루션 레이어별 필터 커널의 개수, 각 필터 커널의 파라미터 등)들이 룩업 테이블 형태로 저장되어 있는 경우, AI 업스케일부(234)는 AI 데이터에 포함된 정보에 기초하여 룩업 테이블 값들 중에서 선택된 일부를 조합하여 DNN 설정 정보를 획득하고, 획득된 DNN 설정 정보를 사용하여 제 2 영상(135)을 AI 업스케일할 수도 있다.
- [129] 구현예에 따라, AI 업스케일부(234)는 업스케일 타겟에 대응되는 DNN의 구조가 결정되면, 결정된 DNN의 구조에 대응하는 DNN 설정 정보, 예를 들어, 필터 커널의 파라미터들을 획득할 수도 있다.
- [130] AI 업스케일부(234)는 제 1 DNN과 관련된 정보를 포함하는 AI 데이터를 통해 제 2 DNN의 DNN 설정 정보를 획득하고, 획득된 DNN 설정 정보로 세팅된 제 2 DNN을 통해 제 2 영상(135)을 AI 업스케일하는 데, 이는, 제 2 영상(135)의 특징을 직접 분석하여 업스케일하는 것과 비교하여 메모리 사용량과 연산량이 감소될 수 있다.
- [131] 일 실시예에서, 제 2 영상(135)이 복수의 프레임으로 구성된 경우, AI 업스케일부(234)는 소정 개수의 프레임 별로 DNN 설정 정보를 독립적으로 획득할 수 있고, 또는, 전체 프레임에 대해 공통된 DNN 설정 정보를 획득할 수도 있다.

- [132] 도 6은 복수의 프레임으로 구성된 제 2 영상(135)을 도시하는 도면이다.
- [133] 도 6에 도시된 바와 같이, 제 2 영상(135)은  $t_0$  내지  $t_n$ 에 대응하는 프레임들로 이루어질 수 있다.
- [134] 일 예시에서, AI 업스케일부(234)는 AI 데이터를 통해 제 2 DNN의 DNN 설정 정보를 획득하고, 획득한 DNN 설정 정보에 기초하여  $t_0$  내지  $t_n$ 에 대응하는 프레임들을 AI 업스케일할 수 있다. 즉,  $t_0$  내지  $t_n$ 에 대응하는 프레임들이 공통된 DNN 설정 정보에 기초하여 AI 업스케일될 수 있다.
- [135] 다른 예시에서, AI 업스케일부(234)는  $t_0$  내지  $t_n$ 에 대응하는 프레임들 중 일부의 프레임들, 예를 들어,  $t_0$  내지  $t_a$ 에 대응하는 프레임들을 AI 데이터로부터 획득한 'A' DNN 설정 정보로 AI 업스케일하고,  $t_{a+1}$  내지  $t_b$ 에 대응하는 프레임들을 AI 데이터로부터 획득한 'B' DNN 설정 정보로 AI 업스케일할 수 있다. 또한, AI 업스케일부(234)는  $t_{b+1}$  내지  $t_n$ 에 대응하는 프레임들을 AI 데이터로부터 획득한 'C' DNN 설정 정보로 AI 업스케일할 수 있다. 다시 말하면, AI 업스케일부(234)는 복수의 프레임들 중 소정 개수의 프레임을 포함하는 그룹마다 DNN 설정 정보를 독립적으로 획득하고, 그룹 각각에 포함된 프레임들을 독립적으로 획득한 DNN 설정 정보로 AI 업스케일할 수 있다.
- [136] 또 다른 예시에서, AI 업스케일부(234)는 제 2 영상(135)을 구성하는 프레임별로 DNN 설정 정보를 독립적으로 획득할 수도 있다. 즉, 제 2 영상(135)이 3개의 프레임으로 구성되어 있는 경우, AI 업스케일부(234)는 첫 번째 프레임과 관련하여 획득한 DNN 설정 정보로 첫 번째 프레임을 AI 업스케일하고, 두 번째 프레임과 관련하여 획득한 DNN 설정 정보로 두 번째 프레임을 AI 업스케일하고, 세 번째 프레임과 관련하여 획득한 DNN 설정 정보로 세 번째 프레임을 AI 업스케일할 수 있다. 전술한, 제 1 복호화부(232)로부터 제공되는 정보(예측 모드 정보, 움직임 정보, 양자화 파라미터 정보 등)와 AI 데이터에 포함된 제 1 영상(115) 관련 정보에 기초하여 DNN 설정 정보가 획득되는 방식에 따라, 제 2 영상(135)을 구성하는 프레임별로 DNN 설정 정보가 독립적으로 획득될 수 있다. 왜냐하면, 모드 정보, 양자화 파라미터 정보 등은 제 2 영상(135)을 구성하는 프레임마다 독립적으로 결정될 수 있기 때문이다.
- [137] 또 다른 예시에서, AI 데이터는 AI 데이터에 기초하여 획득되는 DNN 설정 정보가 어느 프레임까지 유효한 것인지를 나타내는 정보를 포함할 수 있다. 예를 들어, AI 데이터 내에 DNN 설정 정보가  $t_a$  프레임까지 유효하다는 정보가 포함되어 있는 경우, AI 업스케일부(234)는 AI 데이터에 기초하여 획득한 DNN 설정 정보로  $t_0$  내지  $t_a$  프레임들을 AI 업스케일한다. 그리고, 다른 AI 데이터 내에 DNN의 설정 정보가  $t_n$  프레임까지 유효하다는 정보가 포함되어 있는 경우, AI 업스케일부(234)는 상기 다른 AI 데이터에 기초하여 획득한 DNN 설정 정보로  $t_{a+1}$  내지  $t_n$  프레임들을 AI 업스케일할 수 있다.
- [138] 이하에서는, 도 7을 참조하여, 원본 영상(105)의 AI 부호화를 위한 AI 부호화

장치(600)에 대해 설명한다.

[139] 도 7은 일 실시예에 따른 AI 부호화 장치(600)의 구성을 도시하는 블록도이다.

[140] 도 7을 참조하면, AI 부호화 장치(600)는 AI 부호화부(610) 및 전송부(630)를 포함할 수 있다. AI 부호화부(610)는 AI 다운스케일부(612) 및 제 1 부호화부(614)를 포함할 수 있다. 전송부(630)는 데이터 처리부(632) 및 통신부(634)를 포함할 수 있다.

[141] 도 7은 AI 부호화부(610) 및 전송부(630)를 개별적인 장치로 도시하고 있으나, AI 부호화부(610) 및 전송부(630)는 하나의 프로세서를 통해 구현될 수 있다. 이 경우, 전용 프로세서로 구현될 수도 있고, AP 또는 CPU, GPU와 같은 범용 프로세서와 S/W의 조합을 통해 구현될 수도 있다. 또한, 전용 프로세서의 경우, 본 개시의 실시예를 구현하기 위한 메모리를 포함하여 구현되거나, 외부 메모리를 이용하기 위한 메모리 처리부를 포함하여 구현될 수 있다.

[142] 또한, AI 부호화부(610) 및 전송부(630)는 복수의 프로세서로 구성될 수도 있다. 이 경우, 전용 프로세서들의 조합으로 구현될 수도 있고, AP 또는 CPU, GPU와 같은 다수의 범용 프로세서들과 S/W의 조합을 통해 구현될 수도 있다. AI 다운스케일부(612)와 제 1 부호화부(614)도 각각 서로 다른 프로세서로 구현될 수 있다.

[143] AI 부호화부(610)는 원본 영상(105)의 AI 다운스케일 및 제 1 영상(115)의 제 1 부호화를 수행하고, AI 데이터 및 영상 데이터를 전송부(630)로 전달한다. 전송부(630)는 AI 데이터 및 영상 데이터를 AI 복호화 장치(200)로 전송한다.

[144] 영상 데이터는 제 1 영상(115)의 제 1 부호화 결과 획득된 데이터를 포함한다. 영상 데이터는 제 1 영상(115) 내 픽셀 값들에 기초하여 획득되는 데이터, 예를 들어, 제 1 영상(115)과 제 1 영상(115)의 예측 데이터 사이의 차이인 잔차 데이터를 포함할 수 있다. 또한, 영상 데이터는 제 1 영상(115)의 제 1 부호화 과정에서 이용된 정보들을 포함한다. 예를 들어, 영상 데이터는 제 1 영상(115)을 제 1 부호화하는데 이용된 예측 모드 정보, 움직임 정보 및 제 1 영상(115)을 제 1 부호화하는데 이용된 양자화 파라미터 관련 정보 등을 포함할 수 있다.

[145] AI 데이터는, AI 업스케일부(234)가 제 1 DNN의 다운스케일 타겟에 대응하는 업스케일 타겟으로 제 2 영상(135)을 AI 업스케일할 수 있게 하는 정보들을 포함한다. 일 예에서, AI 데이터는 원본 영상(105)과 제 1 영상(115) 사이의 차이 정보를 포함할 수 있다. 또한, AI 데이터는 제 1 영상(115) 관련 정보를 포함할 수도 있다. 제 1 영상(115) 관련 정보는, 제 1 영상(115)의 해상도, 제 1 영상(115)의 제 1 부호화 결과 획득된 영상 데이터의 비트레이트 및 제 1 영상(115)의 제 1 부호화시 이용된 코덱 타입 중 적어도 하나에 대한 정보를 포함할 수 있다.

[146] 일 실시예에서, AI 데이터는, 제 1 DNN의 다운스케일 타겟에 대응하는 업스케일 타겟으로 제 2 영상(135)이 AI 업스케일될 수 있도록, 상호 약속된 DNN 설정 정보의 식별자를 포함할 수 있다.

- [147] 또한, 일 실시예에서, AI 데이터는 제 2 DNN에 세팅 가능한 DNN 설정 정보를 포함할 수도 있다.
- [148] AI 다운스케일부(612)는 제 1 DNN을 통해 원본 영상(105)으로부터 AI 다운스케일된 제 1 영상(115)을 획득할 수 있다. AI 다운스케일부(612)는 미리 결정된 기준에 기초하여 원본 영상(105)의 다운스케일 타겟을 결정할 수 있다.
- [149] 다운스케일 타겟에 부합하는 제 1 영상(115)의 획득을 위해, AI 다운스케일부(612)는 제 1 DNN에 세팅 가능한 복수의 DNN 설정 정보를 저장할 수 있다. AI 다운스케일부(612)는 복수의 DNN 설정 정보 중 다운스케일 타겟에 대응하는 DNN 설정 정보를 획득하고, 획득된 DNN 설정 정보로 세팅된 제 1 DNN을 통해 원본 영상(105)을 AI 다운스케일한다.
- [150] 상기 복수의 DNN 설정 정보 각각은 미리 결정된 해상도 및/또는 미리 결정된 화질의 제 1 영상(115)을 획득하기 위해 훈련된 것일 수 있다. 예를 들어, 복수의 DNN 설정 정보 중 어느 하나의 DNN 설정 정보는 원본 영상(105)의 해상도보다 1/2배만큼 작은 해상도의 제 1 영상(115), 예를 들어, 4K(4096\*2160)의 원본 영상(105)보다 1/2배 작은 2K (2048\*1080)의 제 1 영상(115)을 획득하기 위한 정보들을 포함할 수 있고, 다른 하나의 DNN 설정 정보는 원본 영상(105)의 해상도보다 1/4배만큼 작은 해상도의 제 1 영상(115), 예를 들어, 8K(8192\*4320)의 원본 영상(105)보다 1/4배 작은 2K (2048\*1080)의 제 1 영상(115)을 획득하기 위한 정보들을 포함할 수 있다.
- [151] 구현예에 따라, DNN 설정 정보를 구성하는 정보들(예를 들어, 컨볼루션 레이어의 수, 컨볼루션 레이어별 필터 커널의 개수, 각 필터 커널의 파라미터 등)들이 룩업 테이블 형태로 저장되어 있는 경우, AI 다운스케일부(612)는 다운스케일 타겟에 따라 룩업 테이블 값들 중에서 선택된 일부를 조합하여 DNN 설정 정보를 획득하고, 획득된 DNN 설정 정보를 사용하여 원본 영상(105)을 AI 다운스케일할 수도 있다.
- [152] 구현예에 따라, AI 다운스케일부(612)는 다운스케일 타겟에 대응되는 DNN의 구조를 결정하고, 결정된 DNN의 구조에 대응하는 DNN 설정 정보, 예를 들어, 필터 커널의 파라미터들을 획득할 수도 있다.
- [153] 원본 영상(105)의 AI 다운스케일을 위한 복수의 DNN 설정 정보는, 제 1 DNN과 제 2 DNN이 연계 훈련됨으로써, 최적화된 값을 가질 수 있다. 여기서, 각 DNN 설정 정보는 제 1 DNN에 포함되는 컨볼루션 레이어의 수, 컨볼루션 레이어별 필터 커널의 개수 및 각 필터 커널의 파라미터 중 적어도 하나를 포함한다.
- [154] AI 다운스케일부(612)는 원본 영상(105)의 AI 다운스케일을 위해 결정된 DNN 설정 정보로 제 1 DNN을 세팅하여, 소정 해상도 및/또는 소정 화질의 제 1 영상(115)을 제 1 DNN을 통해 획득할 수 있다. 복수의 DNN 설정 정보 중 원본 영상(105)의 AI 다운스케일을 위한 DNN 설정 정보가 획득되면, 제 1 DNN 내 각 레이어는 DNN 설정 정보에 포함된 정보들에 기초하여 입력된 데이터를 처리할 수 있다.

- [155] 이하에서는, AI 다운스케일부(612)가 다운스케일 타겟을 결정하는 방법에 대해 설명한다. 상기 다운스케일 타겟은 예를 들어, 원본 영상(105)으로부터 얼마나 해상도가 감소한 제 1 영상(115)을 획득해야 하는지를 나타낼 수 있다.
- [156] 일 실시예에서, AI 다운스케일부(612)는 압축률(예를 들어, 원본 영상(105)과 제 1 영상(115) 사이의 해상도 차이, 타겟 비트레이트), 압축 품질(예를 들어, 비트레이트 타입), 압축 히스토리 정보 및 원본 영상(105)의 타입 중 적어도 하나에 기초하여 다운스케일 타겟을 결정할 수 있다.
- [157] 일 예에서, AI 다운스케일부(612)는 미리 설정되거나, 사용자로부터 입력받은 압축률 또는 압축 품질 등에 기반하여 다운스케일 타겟을 결정할 수 있다.
- [158] 다른 예로, AI 다운스케일부(612)는 AI 부호화 장치(600)에 저장된 압축 히스토리 정보를 이용하여 다운스케일 타겟을 결정할 수도 있다. 예를 들어, AI 부호화 장치(600)가 이용할 수 있는 압축 히스토리 정보에 따르면, 사용자가 선호하는 부호화 품질 또는 압축률 등이 결정될 수 있으며, 압축 히스토리 정보에 기초하여 결정된 부호화 품질 등에 따라 다운스케일 타겟이 결정될 수 있다. 예를 들면, 압축 히스토리 정보에 따라 가장 많이 이용된 적이 있는 부호화 품질에 따라 제 1 영상(115)의 해상도, 화질 등이 결정될 수 있다.
- [159] 또 다른 예로, AI 다운스케일부(612)는 압축 히스토리 정보에 따라 소정의 임계값보다 많이 이용된 적이 있는 부호화 품질(예를 들면, 소정의 임계값보다 많이 이용된 적이 있는 부호화 품질의 평균 품질)에 기초하여 다운스케일 타겟을 결정할 수도 있다.
- [160] 또 다른 예로, AI 다운스케일부(612)는 원본 영상(105)의 해상도, 타입(예를 들어, 파일의 형식)등에 기초하여 다운스케일 타겟을 결정할 수도 있다.
- [161] 일 실시예에서, 원본 영상(105)이 복수의 프레임으로 구성된 경우, AI 다운스케일부(612)는 소정 개수의 프레임 별로 다운스케일 타겟을 독립적으로 결정할 수 있고, 또는, 전체 프레임에 대해 공통된 다운스케일 타겟을 결정할 수도 있다.
- [162] 일 예시에서, AI 다운스케일부(612)는 원본 영상(105)을 구성하는 프레임들을 소정 개수의 그룹으로 구분하고, 각 그룹별로 다운스케일 타겟을 독립적으로 결정할 수 있다. 각 그룹에 대해 서로 동일하거나 서로 상이한 다운스케일 타겟이 결정될 수 있다. 그룹들에 포함된 프레임들의 개수는 그룹별로 동일하거나 상이할 수 있다.
- [163] 다른 예시에서, AI 다운스케일부(612)는 원본 영상(105)을 구성하는 프레임별로 다운스케일 타겟을 독립적으로 결정할 수 있다. 각각의 프레임에 대해 서로 동일하거나 서로 상이한 다운스케일 타겟이 결정될 수 있다.
- [164] 이하에서는, AI 다운스케일의 기반이 되는 제 1 DNN(700)의 예시적인 구조에 대해 설명한다.
- [165] 도 8은 원본 영상(105)의 AI 다운스케일을 위한 제 1 DNN(700)을 나타내는 예시적인 도면이다.

- [166] 도 8에 도시된 바와 같이, 원본 영상(105)은 제 1 컨볼루션 레이어(710)로 입력된다. 제 1 컨볼루션 레이어(710)는  $5 \times 5$ 의 크기의 32개의 필터 커널을 이용하여 원본 영상(105)에 대해 컨볼루션 처리를 한다. 컨볼루션 처리 결과 생성된 32개의 특징 맵은 제 1 활성화 레이어(720)로 입력된다. 제 1 활성화 레이어(720)는 32개의 특징 맵에 대해 비선형(Non-linear) 특성을 부여할 수 있다.
- [167] 제 1 활성화 레이어(720)는 제 1 컨볼루션 레이어(710)로부터 출력되는 특징 맵들의 샘플 값들을 제 2 컨볼루션 레이어(730)로 전달할지 여부를 결정한다. 예를 들어, 특징 맵들의 샘플 값들 중 어떤 샘플 값들은 제 1 활성화 레이어(720)에 의해 활성화되어 제 2 컨볼루션 레이어(730)로 전달되고, 어떤 샘플 값들은 제 1 활성화 레이어(720)에 의해 비활성화되어 제 2 컨볼루션 레이어(730)로 전달되지 않는다. 제 1 컨볼루션 레이어(710)로부터 출력되는 특징 맵들이 나타내는 정보가 제 1 활성화 레이어(720)에 의해 강조된다.
- [168] 제 1 활성화 레이어(720)의 출력(725)이 제 2 컨볼루션 레이어(730)로 입력된다. 제 2 컨볼루션 레이어(730)는  $5 \times 5$ 의 크기의 32개의 필터 커널을 이용하여 입력 데이터에 대해 컨볼루션 처리를 한다. 컨볼루션 처리 결과 출력된 32개의 특징 맵은 제 2 활성화 레이어(740)로 입력되고, 제 2 활성화 레이어(740)는 32개의 특징 맵에 대해 비선형 특성을 부여할 수 있다.
- [169] 제 2 활성화 레이어(740)의 출력(745)은 제 3 컨볼루션 레이어(750)로 입력된다. 제 3 컨볼루션 레이어(750)는  $5 \times 5$ 의 크기의 1개의 필터 커널을 이용하여 입력된 데이터에 대해 컨볼루션 처리를 한다. 컨볼루션 처리 결과 제 3 컨볼루션 레이어(750)로부터 1개의 영상이 출력될 수 있다. 제 3 컨볼루션 레이어(750)는 최종 영상을 출력하기 위한 레이어로서 1개의 필터 커널을 이용하여 1개의 출력을 획득한다. 본 개시의 예시에 따르면, 제 3 컨볼루션 레이어(750)는 컨볼루션 연산 결과를 통해 제 1 영상(115)을 출력할 수 있다.
- [170] 제 1 DNN(700)의 제 1 컨볼루션 레이어(710), 제 2 컨볼루션 레이어(730) 및 제 3 컨볼루션 레이어(750)의 필터 커널의 개수, 필터 커널의 파라미터 등을 나타내는 DNN 설정 정보는 복수 개일 수 있는데, 복수의 DNN 설정 정보는 제 2 DNN의 복수의 DNN 설정 정보와 연계되어야 한다. 제 1 DNN의 복수의 DNN 설정 정보와 제 2 DNN의 복수의 DNN 설정 정보 사이의 연계는, 제 1 DNN 및 제 2 DNN의 연계 학습을 통해 구현될 수 있다.
- [171] 도 8은 제 1 DNN(700)이 세 개의 컨볼루션 레이어(710, 730, 750)와 두 개의 활성화 레이어(720, 740)를 포함하고 있는 것으로 도시하고 있으나, 이는 하나의 예시일 뿐이며, 구현예에 따라서, 컨볼루션 레이어 및 활성화 레이어의 개수는 다양하게 변경될 수 있다. 또한, 구현예에 따라서, 제 1 DNN(700)은 RNN(recurrent neural network)을 통해 구현될 수도 있다. 이 경우는 본 개시의 예시에 따른 제 1 DNN(700)의 CNN 구조를 RNN 구조로 변경하는 것을 의미한다.
- [172] 일 실시예에서, AI 다운스케일부(612)는 컨볼루션 연산 및 활성화 레이어의

연산을 위한 적어도 하나의 ALU를 포함할 수 있다. ALU는 프로세서로 구현될 수 있다. 컨볼루션 연산을 위해, ALU는 원본 영상(105) 또는 이전 레이어에서 출력된 특징 맵의 샘플 값들과 필터 커널의 샘플 값들 사이의 곱 연산을 수행하는 곱셈기 및 곱셈의 결과 값들을 더하는 가산기를 포함할 수 있다. 또한, 활성화 레이어의 연산을 위해, ALU는 미리 결정된 시그모이드 함수, Tanh 함수 또는 ReLU 함수 등에서 이용되는 가중치를 입력된 샘플 값에 곱하는 곱셈기, 및 곱한 결과와 소정 값을 비교하여 입력된 샘플 값을 다음 레이어로 전달할지를 판단하는 비교기를 포함할 수 있다.

[173] 다시 도 7을 참조하면, AI 다운스케일부(612)로부터 제 1 영상(115)을 전달받은 제 1 부호화부(614)는 제 1 영상(115)을 제 1 부호화하여 제 1 영상(115)이 가지는 정보량을 감축시킬 수 있다. 제 1 부호화부(614)에 의한 제 1 부호화 결과 제 1 영상(115)에 대응하는 영상 데이터가 획득될 수 있다.

[174] 데이터 처리부(632)는 AI 데이터 및 영상 데이터 중 적어도 하나가 소정의 형태로 전송될 수 있게 처리한다. 예를 들어, AI 데이터 및 영상 데이터를 비트스트림 형태로 전송하여야 하는 경우, 데이터 처리부(632)는 AI 데이터가 비트스트림 형태로 표현되도록 AI 데이터를 처리하고, 통신부(634)를 통해 하나의 비트스트림 형태의 AI 데이터 및 영상 데이터를 전송한다. 다른 예로, 데이터 처리부(632)는 AI 데이터가 비트스트림 형태로 표현되도록 AI 데이터를 처리하고, 통신부(634)를 통해 AI 데이터에 해당하는 비트스트림 및 영상 데이터에 해당하는 비트스트림 각각을 통신부(634)를 통해 전송한다. 또 다른 예로, 데이터 처리부(632)는 AI 데이터가 프레임 또는 패킷으로 표현되도록 AI 데이터를 처리하고, 비트스트림 형태의 영상 데이터와, 프레임 또는 패킷 형태의 AI 데이터를 통신부(634)를 통해 전송한다.

[175] 통신부(634)는 네트워크를 통해 AI 부호화 결과 획득된 AI 부호화 데이터를 전송한다. AI 부호화 결과 획득된 AI 부호화 데이터는 영상 데이터와 AI 데이터를 포함한다. 영상 데이터와 AI 데이터는 동종 네트워크 또는 이종 네트워크를 통해 전송될 수 있다.

[176] 일 실시예에서, 데이터 처리부(632)의 처리 결과 획득된 AI 부호화 데이터는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체, CD-ROM 및 DVD와 같은 광기록 매체, 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical medium) 등을 포함하는 데이터 저장 매체에 저장될 수도 있다.

[177] 이하에서는, 도 9를 참조하여, 제 1 DNN(700)과 제 2 DNN(300)을 연계 훈련시키는 방법에 대해 설명한다.

[178] 도 9는 제 1 DNN(700) 및 제 2 DNN(300)을 훈련시키는 방법을 설명하기 위한 도면이다.

[179] 일 실시예에서 AI 부호화 과정을 통해 AI 부호화된 원본 영상(105)이 AI 복호화 과정을 통해 제 3 영상(145)으로 복원되는데, AI 복호화 결과 획득된 제 3

영상(145)과 원본 영상(105)과의 유사성을 유지하기 위해서는 AI 부호화 과정 및 AI 복호화 과정에 연관성이 필요하다. 즉, AI 부호화 과정에서 손실된 정보를 AI 복호화 과정에서 복원할 수 있어야 하는데, 이를 위해 제 1 DNN(700)과 제 2 DNN(300)의 연계 훈련이 요구된다.

[180] 정확한 AI 복호화를 위해서는 궁극적으로 도 9에 도시된 제 3 훈련 영상(804)과 원본 훈련 영상(801) 사이의 비교 결과에 대응하는 퀄리티 손실 정보(830)를 감소시킬 필요가 있다. 따라서, 퀄리티 손실 정보(830)는 제 1 DNN(700) 및 제 2 DNN(300)의 훈련 모두에 이용된다.

[181] 먼저, 도 9에 도시된 훈련 과정에 대해 설명한다.

[182] 도 9에서, 원본 훈련 영상(original training image)(801)은 AI 다운스케일의 대상이 되는 영상이고, 제 1 훈련 영상(first training image)(802)은 원본 훈련 영상(801)로부터 AI 다운스케일된 영상이다. 또한, 제 3 훈련 영상(third training image)(804)은 제 1 훈련 영상(802)으로부터 AI 업스케일된 영상이다.

[183] 원본 훈련 영상(801)은 정지 영상 또는 복수의 프레임으로 이루어진 동영상상을 포함한다. 일 실시예에서, 원본 훈련 영상(801)은 정지 영상 또는 복수의 프레임으로 이루어진 동영상으로부터 추출된 휘도 영상을 포함할 수도 있다. 또한, 일 실시예에서, 원본 훈련 영상(801)은 정지 영상 또는 복수의 프레임으로 이루어진 동영상에서 추출된 패치 영상을 포함할 수도 있다. 원본 훈련 영상(801)이 복수의 프레임으로 이루어진 경우, 제 1 훈련 영상(802), 제 2 훈련 영상 및 제 3 훈련 영상(804) 역시 복수의 프레임으로 구성된다. 원본 훈련 영상(801)의 복수의 프레임이 순차적으로 제 1 DNN(700)에 입력되면, 제 1 DNN(700) 및 제 2 DNN(300)을 통해 제 1 훈련 영상(802), 제 2 훈련 영상 및 제 3 훈련 영상(804)의 복수의 프레임이 순차적으로 획득될 수 있다.

[184] 제 1 DNN(700) 및 제 2 DNN(300)의 연계 훈련을 위해, 원본 훈련 영상(801)이 제 1 DNN(700)으로 입력된다. 제 1 DNN(700)으로 입력된 원본 훈련 영상(801)은 AI 다운스케일되어 제 1 훈련 영상(802)으로 출력되고, 제 1 훈련 영상(802)이 제 2 DNN(300)에 입력된다. 제 1 훈련 영상(802)에 대한 AI 업스케일 결과 제 3 훈련 영상(804)이 출력된다.

[185] 도 9를 참조하면, 제 2 DNN(300)으로 제 1 훈련 영상(802)이 입력되고 있는데, 구현예에 따라, 제 1 훈련 영상(802)의 제 1 부호화 및 제 1 복호화 과정을 거쳐 획득된 제 2 훈련 영상(second training image)이 제 2 DNN(300)으로 입력될 수도 있다. 제 2 훈련 영상을 제 2 DNN으로 입력시키기 위해 MPEG-2, H.264, MPEG-4, HEVC, VC-1, VP8, VP9 및 AV1 중 어느 하나의 코덱이 이용될 수 있다. 구체적으로, 제 1 훈련 영상(802)의 제 1 부호화 및 제 1 훈련 영상(802)에 대응하는 영상 데이터의 제 1 복호화에, MPEG-2, H.264, MPEG-4, HEVC, VC-1, VP8, VP9 및 AV1 중 어느 하나의 코덱이 이용될 수 있다.

[186] 도 9를 참조하면, 제 1 DNN(700)을 통해 제 1 훈련 영상(802)이 출력되는 것과 별개로, 원본 훈련 영상(801)으로부터 레거시 다운스케일된 축소 훈련

영상(803)이 획득된다. 여기서, 레거시 다운스케일은 바이리니어(bilinear) 스케일, 바이큐빅(bicubic) 스케일, 란조스(lanczos) 스케일 및 스테어 스텝(stair step) 스케일 중 적어도 하나를 포함할 수 있다.

- [187] 원본 영상(105)의 구조적 특징을 기준으로 제 1 영상(115)의 구조적 특징이 크게 벗어나는 것을 방지하기 위해, 원본 훈련 영상(801)의 구조적 특징을 보존하는 축소 훈련 영상(803)을 획득하는 것이다.
- [188] 훈련의 진행 전 제 1 DNN(700) 및 제 2 DNN(300)은 미리 결정된 DNN 설정 정보로 세팅될 수 있다. 훈련이 진행됨에 따라 구조적 손실 정보(810), 복잡성 손실 정보(820) 및 퀄리티 손실 정보(830)가 결정될 수 있다.
- [189] 구조적 손실 정보(810)는 축소 훈련 영상(803)과 제 1 훈련 영상(802)의 비교 결과에 기초하여 결정될 수 있다. 일 예에서, 구조적 손실 정보(810)는 축소 훈련 영상(803)의 구조적 정보와 제 1 훈련 영상(802)의 구조적 정보 사이의 차이에 해당할 수 있다. 구조적 정보는, 영상의 휘도, 대비, 히스토그램 등 영상으로부터 추출 가능한 다양한 특징을 포함할 수 있다. 구조적 손실 정보(810)는 원본 훈련 영상(801)의 구조적 정보가 제 1 훈련 영상(802)에서 어느 정도로 유지되고 있는지를 나타낸다. 구조적 손실 정보(810)가 작을수록 제 1 훈련 영상(802)의 구조적 정보가 원본 훈련 영상(801)의 구조적 정보와 유사해진다.
- [190] 복잡성 손실 정보(820)는 제 1 훈련 영상(802)의 공간적 복잡도에 기초하여 결정될 수 있다. 일 예에서, 공간적 복잡도로서, 제 1 훈련 영상(802)의 총 분산(total variance)값이 이용될 수 있다. 복잡성 손실 정보(820)는 제 1 훈련 영상(802)을 제 1 부호화하여 획득한 영상 데이터의 비트레이트와 관련된다. 복잡성 손실 정보(820)가 작을수록 영상 데이터의 비트레이트가 작은 것으로 정의된다.
- [191] 퀄리티 손실 정보(830)는 원본 훈련 영상(801)과 제 3 훈련 영상(804)의 비교 결과에 기초하여 결정될 수 있다. 퀄리티 손실 정보(830)는 원본 훈련 영상(801)과 제 3 훈련 영상(804)의 차이에 대한 L1-norm 값, L2-norm 값, SSIM(Structural Similarity) 값, PSNR-HVS(Peak Signal-To-Noise Ratio-Human Vision System) 값, MS-SSIM(Multiscale SSIM) 값, VIF(Variance Inflation Factor) 값 및 VMAF(Video Multimethod Assessment Fusion) 값 중 적어도 하나를 포함할 수 있다. 퀄리티 손실 정보(830)는 제 3 훈련 영상(804)이 원본 훈련 영상(801)과 어느 정도로 유사한지를 나타낸다. 퀄리티 손실 정보(830)가 작을수록 제 3 훈련 영상(804)이 원본 훈련 영상(801)에 더 유사해진다.
- [192] 도 9를 참조하면, 구조적 손실 정보(810), 복잡성 손실 정보(820) 및 퀄리티 손실 정보(830)가 제 1 DNN(700)의 훈련에 이용되고, 퀄리티 손실 정보(830)는 제 2 DNN(300)의 훈련에 이용된다. 즉, 퀄리티 손실 정보(830)는 제 1 DNN(700) 및 제 2 DNN(300)의 훈련에 모두 이용된다.
- [193] 제 1 DNN(700)은 구조적 손실 정보(810), 복잡성 손실 정보(820) 및 퀄리티 손실 정보(830)에 기초하여 결정된 최종 손실 정보가 감소 또는 최소화되도록

파라미터를 갱신할 수 있다. 또한, 제 2 DNN(300)은 퀄리티 손실 정보(830)가 감소 또는 최소화되도록 파라미터를 갱신할 수 있다.

[194] 제 1 DNN(700) 및 제 2 DNN(300)의 훈련을 위한 최종 손실 정보는 아래의 수학적 식 1과 같이 결정될 수 있다.

[195] [수학적 식 1]

[196]  $\text{LossDS} = a * \text{구조적 손실 정보} + b * \text{복잡성 손실 정보} + c * \text{퀄리티 손실 정보}$

[197]  $\text{LossDS} = d * \text{퀄리티 손실 정보}$

[198] 상기 수학적 식 1에서, LossDS는 제 1 DNN(700)의 훈련을 위해 감소 또는 최소화되어야 할 최종 손실 정보를 나타내고, LossUS는 제 2 DNN(300)의 훈련을 위해 감소 또는 최소화되어야 할 최종 손실 정보를 나타낸다. 또한, a, b, c, d는 미리 결정된 소정의 가중치에 해당할 수 있다.

[199] 즉, 제 1 DNN(700)은 수학적 식 1의 LossDS가 감소되는 방향으로 파라미터들을 갱신하고, 제 2 DNN(300)은 LossUS가 감소되는 방향으로 파라미터들을 갱신하게 된다. 훈련 과정에서 도출된 LossDS에 따라 제 1 DNN(700)의 파라미터들이 갱신되면, 갱신된 파라미터에 기초하여 획득되는 제 1 훈련 영상(802)이 이전 훈련 과정에서의 제 1 훈련 영상(802)과 달라지게 되고, 그에 따라 제 3 훈련 영상(804) 역시 이전 훈련 과정에서의 제 3 훈련 영상(804)과 달라지게 된다. 제 3 훈련 영상(804)이 이전 훈련 과정에서의 제 3 훈련 영상(804)과 달라지게 되면, 퀄리티 손실 정보(830) 역시 새롭게 결정되며, 그에 따라 제 2 DNN(300)은 파라미터들을 갱신한다. 퀄리티 손실 정보(830)가 새롭게 결정되면, LossDS 역시 새롭게 결정되므로, 제 1 DNN(700)은 새롭게 결정된 LossDS에 따라 파라미터들을 갱신한다. 즉, 제 1 DNN(700)의 파라미터 갱신은, 제 2 DNN(300)의 파라미터 갱신을 야기하고, 제 2 DNN(300)의 파라미터 갱신은 제 1 DNN(700)의 파라미터 갱신을 야기하는 것이다. 다시 말하면, 제 1 DNN(700) 및 제 2 DNN(300)은 퀄리티 손실 정보(830)의 공유를 통해 연계 훈련되므로, 제 1 DNN(700)의 파라미터들과 제 2 DNN(300)의 파라미터들이 서로 연관성을 가지고 최적화될 수 있는 것이다.

[200] 수학적 식 1을 참조하면, LossUS가 퀄리티 손실 정보(830)에 따라 결정되는 것을 알 수 있으나, 이는 하나의 예시이며, LossUS는 구조적 손실 정보(810) 및 복잡성 손실 정보(820) 중 적어도 하나와, 퀄리티 손실 정보(830)에 기초하여 결정될 수도 있다.

[201] 앞서, AI 복호화 장치(200)의 AI 업스케일부(234) 및 AI 부호화 장치(600)의 AI 다운스케일부(612)는 복수의 DNN 설정 정보를 저장하는 것으로 설명하였는데, AI 업스케일부(234) 및 AI 다운스케일부(612)에 저장되는 복수의 DNN 설정 정보 각각을 훈련시키는 방법에 대해 설명한다.

[202] 수학적 식 1과 관련하여 설명한 바와 같이, 제 1 DNN(700)의 경우, 제 1 훈련 영상(802)의 구조적 정보와 원본 훈련 영상(801)의 구조적 정보 사이의 유사 정도(구조적 손실 정보(810)), 제 1 훈련 영상(802)의 제 1 부호화 결과 획득되는

영상 데이터의 비트레이트(복잡성 손실 정보(820)) 및 제 3 훈련 영상(804)과 원본 훈련 영상(801) 사이의 차이(퀄리티 손실 정보(830))를 고려하여 파라미터를 갱신하게 된다.

- [203] 자세히 설명하면, 원본 훈련 영상(801)의 구조적 정보와 유사하면서, 제 1 부호화를 하였을 때 획득되는 영상 데이터의 비트레이트가 작은 제 1 훈련 영상(802)이 획득 가능하도록 하는 동시에, 제 1 훈련 영상(802)을 AI 업스케일하는 제 2 DNN(300)이 원본 훈련 영상(801)에 유사한 제 3 훈련 영상(804)을 획득할 수 있도록, 제 1 DNN(700)의 파라미터가 갱신될 수 있다.
- [204] 수학적 식 1의 a, b, c의 가중치가 조절됨으로써, 제 1 DNN(700)의 파라미터들이 최적화되는 방향이 상이해지게 된다. 예를 들어, b의 가중치를 높게 결정하는 경우, 제 3 훈련 영상(804)의 퀄리티보다 비트레이트가 낮아지는 것에 더 중요도를 두고 제 1 DNN(700)의 파라미터가 갱신될 수 있다. 또한, c의 가중치를 높게 결정하는 경우, 비트레이트가 높아지는 것이나, 원본 훈련 영상(801)의 구조적 정보가 유지되는 것보다 제 3 훈련 영상(804)의 퀄리티가 증가하도록 하는 것에 더 중요도를 두고 제 1 DNN(700)의 파라미터가 갱신될 수 있다.
- [205] 또한, 제 1 훈련 영상(802)을 제 1 부호화하는데 이용되는 코덱의 타입에 따라 제 1 DNN(700)의 파라미터들이 최적화되는 방향이 상이해질 수 있다. 왜냐하면, 코덱의 종류에 따라, 제 2 DNN(300)으로 입력될 제 2 훈련 영상이 달라질 수 있기 때문이다.
- [206] 즉, 가중치 a, 가중치 b, 가중치 c 및 제 1 훈련 영상(802)의 제 1 부호화를 위한 코덱의 종류에 기반하여 제 1 DNN(700)의 파라미터들과 제 2 DNN(300)의 파라미터들이 연계하여 갱신될 수 있는 것이다. 따라서, 가중치 a, 가중치 b, 가중치 c 각각을 소정의 값으로 결정하고, 코덱의 종류를 소정의 종류로 결정한 후, 제 1 DNN(700)과 제 2 DNN(300)을 훈련시키면, 서로 연계되어 최적화된 제 1 DNN(700)의 파라미터들과 제 2 DNN(300)의 파라미터들이 결정될 수 있다.
- [207] 그리고, 가중치 a, 가중치 b, 가중치 c 및 코덱의 종류를 변경한 후, 제 1 DNN(700)과 제 2 DNN(300)을 훈련시키면, 서로 연계되어 최적화된 제 1 DNN(700)의 파라미터들과 제 2 DNN(300)의 파라미터들이 결정될 수 있다. 다시 말하면, 가중치 a, 가중치 b, 가중치 c 및 코덱의 종류 각각의 값을 변경하면서 제 1 DNN(700)과 제 2 DNN(300)을 훈련시키면 서로 연계되어 훈련된 복수의 DNN 설정 정보가 제 1 DNN(700) 및 제 2 DNN(300)에서 결정될 수 있는 것이다.
- [208] 앞서, 도 5와 관련하여 설명한 바와 같이, 제 1 DNN(700) 및 제 2 DNN(300)의 복수의 DNN 설정 정보들은 제 1 영상 관련 정보들에 매핑되어 있을 수 있다. 이러한 매핑 관계의 설정을 위해, 제 1 DNN(700)에서 출력되는 제 1 훈련 영상(802)을 특정 비트레이트에 따라 특정 코덱으로 제 1 부호화하고, 제 1 부호화 결과 획득된 비트스트림을 제 1 복호화하여 획득한 제 2 훈련 영상을 제 2 DNN(300)으로 입력할 수 있다. 즉, 특정 해상도의 제 1 훈련 영상(802)이 특정 코덱에 의해 특정 비트레이트로 제 1 부호화되도록 환경을 설정한 후, 제 1

DNN(700) 및 제 2 DNN(300)을 훈련시킴으로써, 제 1 훈련 영상(802)의 해상도, 제 1 훈련 영상(802)의 제 1 부호화에 이용된 코덱의 종류 및 제 1 훈련 영상(802)의 제 1 부호화 결과 획득된 비트스트림의 비트레이트에 매핑된 DNN 설정 정보 쌍이 결정될 수 있는 것이다. 제 1 훈련 영상(802)의 해상도, 제 1 훈련 영상(802)의 제 1 부호화에 이용된 코덱의 종류 및 제 1 훈련 영상(802)의 제 1 부호화에 따라 획득되는 비트스트림의 비트레이트를 다양하게 변경시킴으로써, 제 1 DNN(700) 및 제 2 DNN(300)의 복수의 DNN 설정 정보들과 제 1 영상 관련 정보들 사이의 매핑 관계가 결정될 수 있다.

- [209] 도 10는 훈련 장치(1000)에 의한 제 1 DNN(700) 및 제 2 DNN(300)의 훈련 과정을 설명하기 위한 도면이다.
- [210] 도 9와 관련하여 설명한 제 1 DNN(700) 및 제 2 DNN(300)의 훈련은 훈련 장치(1000)에 의해 수행될 수 있다. 훈련 장치(1000)는 제 1 DNN(700) 및 제 2 DNN(300)을 포함한다. 훈련 장치(1000)는 예를 들어, AI 부호화 장치(600) 또는 별도의 서버일 수 있다. 훈련 결과 획득된 제 2 DNN(300)의 DNN 설정 정보들은 AI 복호화 장치(200)에 저장된다.
- [211] 도 10를 참조하면, 훈련 장치(1000)는 제 1 DNN(700) 및 제 2 DNN(300)의 DNN 설정 정보를 초기 세팅한다(S840, S845). 이에 의해, 제 1 DNN(700) 및 제 2 DNN(300)은 미리 결정된 DNN 설정 정보에 따라 동작할 수 있다. DNN 설정 정보는 제 1 DNN(700) 및 제 2 DNN(300)에 포함되는 컨볼루션 레이어의 수, 컨볼루션 레이어별 필터 커널의 개수, 컨볼루션 레이어별 필터 커널의 크기 및 각 필터 커널의 파라미터 중 적어도 하나에 대한 정보를 포함할 수 있다.
- [212] 훈련 장치(1000)는 원본 훈련 영상(801)을 제 1 DNN(700)으로 입력한다(S850). 원본 훈련 영상(801)은 정지 영상 또는 동영상을 구성하는 적어도 하나의 프레임을 포함할 수 있다.
- [213] 제 1 DNN(700)은 초기 세팅된 DNN 설정 정보에 따라 원본 훈련 영상(801)을 처리하고, 원본 훈련 영상(801)으로부터 AI 다운스케일된 제 1 훈련 영상(802)을 출력한다(S855). 도 10은 제 1 DNN(700)으로부터 출력된 제 1 훈련 영상(802)이 제 2 DNN(300)으로 바로 입력되는 것으로 도시되어 있으나, 제 1 DNN(700)으로부터 출력된 제 1 훈련 영상(802)이 훈련 장치(1000)에 의해 제 2 DNN(300)으로 입력될 수 있다. 또한, 훈련 장치(1000)는 제 1 훈련 영상(802)을 소정의 코덱으로 제 1 부호화 및 제 1 복호화한 후, 제 2 훈련 영상을 제 2 DNN(300)으로 입력할 수 있다.
- [214] 제 2 DNN(300)은 제 1 훈련 영상(802) 또는 제 2 훈련 영상을 초기 세팅된 DNN 설정 정보에 따라 처리하고, 제 1 훈련 영상(802) 또는 제 2 훈련 영상으로부터 AI 업스케일된 제 3 훈련 영상(804)을 출력한다(S860).
- [215] 훈련 장치(1000)는 제 1 훈련 영상(802)에 기초하여 복잡성 손실 정보(820)를 산출한다(S865).
- [216] 훈련 장치(1000)는 축소 훈련 영상(803)과 제 1 훈련 영상(802)을 비교하여

구조적 손실 정보(810)를 산출한다(S870).

[217] 훈련 장치(1000)는 원본 훈련 영상(801)과 제 3 훈련 영상(804)을 비교하여  
퀄리티 손실 정보(830)를 산출한다(S875).

[218] 제 1 DNN(700)은 최종 손실 정보에 기초한 역전사(back propagation) 과정을  
통해 초기 세팅된 DNN 설정 정보를 갱신한다(S880). 훈련 장치(1000)는 복잡성  
손실 정보(820), 구조적 손실 정보(810) 및 퀄리티 손실 정보(830)에 기초하여 제  
1 DNN(700)의 훈련을 위한 최종 손실 정보를 산출할 수 있다.

[219] 제 2 DNN(300)은 퀄리티 손실 정보 또는 최종 손실 정보에 기초한 역전사  
과정을 통해 초기 세팅된 DNN 설정 정보를 갱신한다(S885). 훈련 장치(1000)는  
퀄리티 손실 정보(830)에 기초하여 제 2 DNN(300)의 훈련을 위한 최종 손실  
정보를 산출할 수 있다.

[220] 이후, 훈련 장치(1000), 제 1 DNN(700) 및 제 2 DNN(300)은 최종 손실 정보들이  
최소화될 때까지 S850 내지 S885 과정을 반복하면서 DNN 설정 정보를  
갱신한다. 이 때, 각 반복 과정 동안 제 1 DNN(700) 및 제 2 DNN은 이전 과정에서  
갱신된 DNN 설정 정보에 따라 동작한다.

[221] 아래의 표 1은 본 개시의 일 실시예에 따라 원본 영상(105)을 AI 부호화 및 AI  
복호화한 경우와, HEVC로 원본 영상(105)을 부호화 및 복호화한 경우의 효과를  
나타낸다.

[222] [표 1]

| Content    | Resolution        | Frame<br>Number | Information<br>Amount (Bitrate)<br>(Mbps) |                             | Subjective Image<br>Quality Score<br>(VMAF) |                             |
|------------|-------------------|-----------------|-------------------------------------------|-----------------------------|---------------------------------------------|-----------------------------|
|            |                   |                 | HEVC                                      | AI Encoding/<br>AI Decoding | HEVC                                        | AI Encoding/<br>AI Decoding |
| Content_01 | 8K<br>(7680×4320) | 300 frames      | 46.3                                      | 21.4                        | 94.80                                       | 93.54                       |
| Content_02 |                   |                 | 46.3                                      | 21.6                        | 98.05                                       | 98.98                       |
| Content_03 |                   |                 | 46.3                                      | 22.7                        | 96.08                                       | 96.00                       |
| Content_04 |                   |                 | 46.1                                      | 22.1                        | 86.26                                       | 92.00                       |
| Content_05 |                   |                 | 45.4                                      | 22.7                        | 93.42                                       | 92.98                       |
| Content_06 |                   |                 | 46.3                                      | 23.0                        | 95.99                                       | 95.61                       |
|            |                   | Average         | 46.11                                     | 22.25                       | 94.10                                       | 94.85                       |

[224] 표 1에서 알 수 있듯이, 본 개시의 일 실시예에 따라 8K 해상도의 300개의  
프레임들로 이루어진 콘텐츠들을 AI 부호화 및 AI 복호화한 경우의 주관적  
화질이 HEVC로 부호화 및 복호화한 경우의 주관적 화질보다 높음에도,  
비트레이트가 50% 이상 감소한 것을 알 수 있다.

[225] 도 11은 원본 영상(105)의 AI 다운스케일을 위한 장치(20) 및 제 2 영상(135)의  
AI 업스케일을 위한 장치(40)를 도시하는 예시적인 도면이다.

[226] 장치(20)는 원본 영상(105)을 수신하고, AI 다운스케일부(1124) 및 변환 기반  
부호화부(1126)를 이용하여 영상 데이터(25) 및 AI 데이터(30)를 장치(40)로

제공한다. 일 실시예에서, 영상 데이터(25)는 도 1의 영상 데이터에 대응하고, AI 데이터(30)는 도 1의 AI 데이터에 대응한다. 또한, 일 실시예에서, 변환 기반 부호화부(1126)는 도 7의 제 1 부호화부(614)에 대응하고, AI 다운스케일부(1124)는 도 7의 AI 다운스케일부(612)에 대응한다.

- [227] 장치(40)는 AI 데이터(30) 및 영상 데이터(25)를 수신하고, 변환 기반 복호화부(1146) 및 AI 업스케일부(1144)를 이용하여 제 3 영상(145)을 획득한다. 일 실시예에서, 변환 기반 복호화부(1146)는 도 2의 제 1 복호화부(232)에 대응하고, AI 업스케일부(1144)는 도 2의 AI 업스케일부(234)에 대응한다.
- [228] 일 실시예에서, 장치(20)는 CPU, 메모리 및 인스트럭션들을 포함하는 컴퓨터 프로그램을 포함한다. 컴퓨터 프로그램은 메모리에 저장된다. 일 실시예에서, CPU에 의한 컴퓨터 프로그램의 실행에 따라, 장치(20)는 도 11과 관련하여 설명할 기능을 수행한다. 일 실시예에서, 전용 하드웨어 칩 및/또는 CPU에 의해 도 11과 관련하여 설명할 기능이 수행된다.
- [229] 일 실시예에서, 장치(40)는 CPU, 메모리 및 인스트럭션들을 포함하는 컴퓨터 프로그램을 포함한다. 컴퓨터 프로그램은 메모리에 저장된다. 일 실시예에서, CPU에 의한 컴퓨터 프로그램의 실행에 따라, 장치(40)는 도 11과 관련하여 설명할 기능을 수행한다. 일 실시예에서, 전용 하드웨어 칩 및/또는 CPU에 의해 도 11과 관련하여 설명할 기능이 수행된다.
- [230] 도 11에서, 구성 제어부(1122)는 하나 이상의 입력 값(10)을 수신한다. 일 실시예에서, 하나 이상의 입력 값(10)은 AI 다운스케일부(1124) 및 AI 업스케일부(1144)를 위한 타겟 해상도 차이, 영상 데이터(25)의 비트레이트, 영상 데이터(25)의 비트레이트 타입(예를 들어, variable bitrate 타입, constant bitrate 타입 또는 average bitrate 타입 등) 및 변환 기반 부호화부(1126)를 위한 코덱 타입 중 적어도 하나를 포함할 수 있다. 하나 이상의 입력 값(10)은 장치(20)에 미리 저장되거나, 사용자로부터 입력받은 값을 포함할 수 있다.
- [231] 구성 제어부(1122)는 수신된 입력 값(10)에 기초하여 AI 다운스케일부(1124) 및 변환 기반 부호화부(1126)의 동작을 제어한다. 일 실시예에서, 구성 제어부(1122)는 수신된 입력 값(10)에 따라 AI 다운스케일부(1124)를 위한 DNN 설정 정보를 획득하고, 획득한 DNN 설정 정보로 AI 다운스케일부(1124)를 세팅한다. 일 실시예에서, 구성 제어부(1122)는 수신된 입력 값(10)을 AI 다운스케일부(1124)로 전달하고, AI 다운스케일부(1124)는 수신된 입력 값(10)에 기초하여 원본 영상(105)을 AI 다운스케일하기 위한 DNN 설정 정보를 획득할 수 있다. 일 실시예에서, 구성 제어부(1122)는 입력 값(10)과 함께 추가 정보, 예를 들어, AI 다운스케일이 적용되는 컬러 포맷(휘도 성분, 색차 성분, 레드 성분, 그린 성분 또는 블루 성분 등) 정보, HDR(high dynamic range)의 톤 매핑(tone mapping) 정보 등을 AI 다운스케일부(1124)로 제공하고, AI 다운스케일부(1124)는 입력 값(10) 및 추가 정보를 고려하여 DNN 설정 정보를 획득할 수도 있다. 일 실시예에서, 구성 제어부(1122)는 수신된 입력 값(10)의

적어도 일부를 변환 기반 부호화부(1126)로 전달하여 변환 기반 부호화부(1126)가 특정 값의 비트레이트, 특정 타입의 비트레이트 및 특정 코덱으로 제 1 영상(115)을 제 1 부호화하게 한다.

- [232] AI 다운스케일부(1124)는 원본 영상(105)을 수신하고, 제 1 영상(115)을 획득하기 위해 도 1, 도 7, 도 8, 도 9 및 도 10 중 적어도 하나와 관련하여 설명한 동작을 수행한다.
- [233] 일 실시예에서, AI 데이터(30)는 장치(40)로 제공된다. AI 데이터(30)는 원본 영상(105)과 제 1 영상(115) 사이의 해상도 차이 정보 및 제 1 영상(115) 관련 정보 중 적어도 하나를 포함할 수 있다. 해상도 차이 정보는 입력 값(10)의 타겟 해상도 차이에 기반하여 결정될 수 있고, 제 1 영상(115) 관련 정보는 타겟 비트레이트, 비트레이트 타입 및 코덱 타입 중 적어도 하나에 기반하여 결정될 수 있다. 일 실시예에서, AI 데이터(30)는 AI 업스케일 과정에 이용되는 파라미터들을 포함할 수도 있다. AI 데이터는 AI 다운스케일부(1124)로부터 장치(40)로 제공될 수도 있다.
- [234] 제 1 영상(105)은 변환 기반 부호화부(1126)에 의해 처리되어 영상 데이터(25)가 획득되고, 영상 데이터(25)는 장치(40)로 전달된다. 변환 기반 부호화부(1126)는 MPEG-2, H.264 AVC, MPEG-4, HEVC, VC-1, VP8, VP9 또는 AV1에 따라 제 1 영상(115)을 처리할 수 있다.
- [235] 구성 제어부(1142)는 AI 데이터(30)에 기초하여 AI 업스케일부(1144)의 동작을 제어한다. 일 실시예에서, 구성 제어부(1142)는 수신된 AI 데이터(30)에 따라 AI 업스케일부(1144)를 위한 DNN 설정 정보를 획득하고, 획득한 DNN 설정 정보로 AI 업스케일부(1144)를 세팅한다. 일 실시예에서, 구성 제어부(1142)는 수신된 AI 데이터(30)를 AI 업스케일부(1144)로 전달하고, AI 업스케일부(1144)는 AI 데이터(30)에 기초하여 제 2 영상(135)을 AI 업스케일하기 위한 DNN 설정 정보를 획득할 수 있다. 일 실시예에서, 구성 제어부(1142)는 AI 데이터(30)와 함께 추가 정보, 예를 들어, AI 업스케일이 적용되는 컬러 포맷(휘도 성분, 색차 성분, 레드 성분, 그린 성분 또는 블루 성분 등) 정보, HDR(high dynamic range)의 톤 매핑 정보 등을 AI 업스케일부(1144)로 제공하고, AI 업스케일부(1144)는 AI 데이터(30) 및 추가 정보를 고려하여 DNN 설정 정보를 획득할 수도 있다. 일 실시예에서, AI 업스케일부(1144)는 구성 제어부(1142)로부터 AI 데이터(30)를 수신하고, 변환 기반 복호화부(1146)로부터 예측 모드 정보, 움직임 정보 및 양자화 파라미터 정보 중 적어도 하나를 수신하고, 예측 모드 정보, 움직임 정보 및 양자화 파라미터 정보 중 적어도 하나와 AI 데이터(30)에 기초하여 DNN 설정 정보를 획득할 수도 있다.
- [236] 변환 기반 복호화부(1146)는 영상 데이터(25)를 처리하여 제 2 영상(135)을 복원한다. 변환 기반 복호화부(1146)는 MPEG-2, H.264 AVC, MPEG-4, HEVC, VC-1, VP8, VP9 또는 AV1에 따라 영상 데이터(25)를 처리할 수 있다.
- [237] AI 업스케일부(1144)는 세팅된 DNN 설정 정보를 기초로 변환 기반

복호화부(1146)로부터 제공되는 제 2 영상(135)을 AI 업스케일하여 제 3 영상(145)을 획득한다.

- [238] AI 다운스케일부(1124)는 제 1 DNN을 포함하고, AI 업스케일부(1144)는 제 2 DNN을 포함할 수 있는데, 일 실시예에서, 제 1 DNN 및 제 2 DNN을 위한 DNN 설정 정보들은 도 9 및 도 10과 관련하여 설명한 훈련 방법에 따라 훈련된다.
- [239] 도 12 내지 25에서, AI 데이터의 데이터 구조와 함께, AI 부호화 데이터에서 AI 데이터와 영상 데이터의 관계가 설명된다. 이하, 업스케일 DNN은 도 3의 제2 DNN(300)과 마찬가지로 제2 영상(135)과 같은 저해상도 영상을 제3 영상(145)과 같은 고해상도 영상으로 AI 업스케일하는데 사용되는 DNN이다. 업스케일 DNN 정보는 AI 데이터에 따라 특정된 DNN 설정 정보를 나타내며, 업스케일 DNN 정보에 기반하여 업스케일 DNN이 설정될 수 있다. 저해상도 영상은 제1 영상(115) 및 제2 영상(135)와 같이 해상도가 작은 영상을 나타낸다. 고해상도 영상은 원본 영상(105) 및 제3 영상(145)과 같이 해상도가 높은 영상을 나타낸다.
- [240] 도 12는 영상 데이터(1200)와 AI 데이터(1240)의 구조와 함께, 영상 데이터(1200)와 AI 데이터(1240)의 상응 관계를 설명한다.
- [241] 도 12에서 영상 데이터(1200)의 비디오-프레임 그룹-프레임 계층 구조가 설명된다. 도 12의 비디오(1202)는 영상 데이터(1200)의 연속된 모든 프레임을 포함하는 데이터 단위이다. 비디오(1202)에 포함된 모든 프레임에는 비디오 파라미터 세트의 파라미터 정보가 적용될 수 있다. 비디오 파라미터 세트는 비디오 헤더(1204)에 포함된다.
- [242] 비디오(1202)는 복수의 프레임 그룹을 포함할 수 있다. 프레임 그룹은 프레임 그룹 파라미터 세트의 파라미터 정보를 공유하는 하나 이상의 연속된 프레임으로 구성된 데이터 단위이다. 프레임 그룹은 GOP(Group of Pictures) 또는 CVS(Coded Video Sequence)일 수 있다. 프레임 그룹 파라미터 세트는 프레임 그룹 헤더에 포함될 수 있다. 예를 들어, 제1 프레임 그룹(1210)의 프레임 그룹 파라미터 세트는 제1 프레임 그룹 헤더(1212)에 포함될 수 있다. 그리고 제2 프레임 그룹(1214)의 프레임 그룹 파라미터 세트는 제2 프레임 그룹 헤더(1216)에 포함될 수 있다.
- [243] 하나의 프레임 그룹은 다른 프레임 그룹을 참조하지 않고 독립적으로 부호화된다. 프레임 그룹은 다른 프레임을 참조하지 않고 부호화된 IDR(Instantaneous decoding refresh) 프레임 또는 IRAP(Intra Random Access Pictures) 프레임을 포함한다. 그리고 프레임 그룹의 나머지 프레임은 IDR 프레임 (또는 IRAP 프레임)을 참조하여 부호화된다. 따라서, 제1 프레임 그룹(1210)는 비디오(1202)의 다른 프레임 그룹을 참조하지 않고 독립적으로 부호화된다. 그리고 제1 프레임 그룹(1210)의 첫 번째로 부호화된 프레임인 제1 프레임(1220)은 IDR 프레임 (또는 IRAP 프레임)이다. 제2 프레임(1230)을 비롯한 제1 프레임 그룹(1210)의 나머지 프레임들은 제1 프레임(1220)을 참조하여 부호화된다.

- [244] 프레임은 비디오에 포함된 하나의 정지 영상을 나타낸다. 프레임 헤더에는 프레임에 적용되는 파라미터 정보를 포함하는 프레임 파라미터 세트가 포함될 수 있다. 예를 들어, 제1 프레임(1220)의 제1 프레임 헤더(1222)는 제1 프레임(1220)에 적용되는 프레임 파라미터 세트가 포함될 수 있다. 마찬가지로, 제2 프레임(1230)의 제2 프레임 헤더(1232)는 제2 프레임(1230)에 적용되는 프레임 파라미터 세트가 포함될 수 있다.
- [245] 도 12에서 상기 설명된 비디오-프레임 그룹-프레임 계층 구조에 따른 AI 데이터(1240)의 분류 방법이 설명된다. AI 데이터(1240)는 적용 범위에 따라 비디오 AI 데이터(1242), 프레임 그룹 AI 데이터(1250), 및 프레임 AI 데이터(1260)로 분류될 수 있다. 비디오 AI 데이터(1242)는 비디오에 포함된 모든 프레임 그룹들에 공통적으로 적용되는 AI 데이터를 의미한다. 그리고 프레임 그룹 AI 데이터(1250)는 현재 프레임 그룹에 포함된 프레임에 공통적으로 적용되는 AI 데이터를 의미한다. 또한 프레임 AI 데이터(1260)는 현재 프레임에 적용되는 AI 데이터를 의미한다.
- [246] 비디오 AI 데이터(1242)는 비디오 헤더(1204)에 대응된다. 따라서 비디오 AI 데이터(1242)는 비디오 헤더(1204)와 병렬적으로 복호화될 수 있다. 또는 비디오 AI 데이터(1242)는 비디오 헤더(1204)의 복호화 바로 이전에 복호화될 수 있다. 또는 비디오 AI 데이터(1242)는 비디오 헤더(1204)의 복호화 바로 이후에 복호화될 수 있다.
- [247] 프레임 그룹 AI 데이터(1250)는 프레임 그룹 헤더에 대응된다. 도 12에 따르면, 제1 프레임 그룹 AI 데이터(1252)는 제1 프레임 그룹 헤더(1212)에 대응된다. 그리고 제2 프레임 그룹 AI 데이터(1254)는 제2 프레임 그룹 헤더(1216)에 대응된다. 제1 프레임 그룹 AI 데이터(1252) 및 제2 프레임 그룹 AI 데이터(1254)는 각각 제1 프레임 그룹 헤더(1212) 및 제2 프레임 그룹 헤더(1216)와 병렬적으로 복호화될 수 있다. 또는 제1 프레임 그룹 AI 데이터(1252) 및 제2 프레임 그룹 AI 데이터(1254)는 각각 제1 프레임 그룹 헤더(1212) 및 제2 프레임 그룹 헤더(1216)의 복호화 바로 이전에 복호화될 수 있다. 또는 제1 프레임 그룹 AI 데이터(1252) 및 제2 프레임 그룹 AI 데이터(1254)는 각각 제1 프레임 그룹 헤더(1212) 및 제2 프레임 그룹 헤더(1216)의 복호화 바로 이후에 복호화될 수 있다.
- [248] 프레임 AI 데이터(1260)는 프레임 헤더에 대응된다. 도 12에 따르면, 제1 프레임 AI 데이터(1262)는 제1 프레임 헤더(1222)에 대응된다. 그리고 제2 프레임 AI 데이터(1264)는 제2 프레임 헤더(1232)에 대응된다. 제1 프레임 AI 데이터(1262) 및 제2 프레임 AI 데이터(1264)는 각각 제1 프레임 헤더(1222) 및 제2 프레임 헤더(1232)와 병렬적으로 복호화될 수 있다. 또는 제1 프레임 AI 데이터(1262) 및 제2 프레임 AI 데이터(1264)는 각각 제1 프레임 헤더(1222) 및 제2 프레임 헤더(1232)의 복호화 바로 이전에 복호화될 수 있다. 또는 제1 프레임 AI 데이터(1262) 및 제2 프레임 AI 데이터(1264)는 각각 제1 프레임 헤더(1222) 및

제2 프레임 헤더(1232)의 복호화 바로 이후에 복호화될 수 있다.

- [249] 일 실시 예에 따르면, 도 7의 데이터 처리부(632)는 영상 데이터(1200)와 AI 데이터(1240)를 모두 포함한 단일 파일 형태의 AI 부호화 데이터를 생성할 수 있다. 통신부(634)는 단일 파일 형태의 AI 부호화 데이터를 도 2의 통신부(212)로 전송한다. 본 명세서에서 파일은 메모리에 저장된 데이터의 모임을 의미한다. 그리고 비디오 파일은 메모리에 저장된 영상 데이터의 모임을 의미하며, 영상 데이터는 비트스트림의 형태로 구현될 수 있다.
- [250] 일 실시 예에 따르면, AI 데이터(1240)는 영상 데이터(1200)에 삽입되지 않고, 단일 파일 내에서 영상 데이터(1200)와 분리되어 구성될 수 있다. 따라서 AI 부호화 데이터가 단일 파일로 구성되었음에도, AI 데이터(1240)와 영상 데이터(1200)가 분리되었으므로, AI 데이터(1240) 및/또는 영상 데이터(1200)는 AI 데이터(1240)와 영상 데이터(1200)의 동기화에 관한 정보를 포함할 수 있다.
- [251] 통신부(212)는 AI 부호화 데이터를 수신할 수 있다. 그리고 파싱부(214)는 AI 부호화 데이터로부터 AI 데이터와 영상 데이터를 추출할 수 있다. 그리고 출력부(216)는 영상 데이터를 제1 복호화부(232)로, AI 데이터를 AI 업스케일부(234)에 전송한다. 그리고 제1 복호화부(232)는 영상 데이터를 복호화하여, 저해상도 영상을 생성한다. 그리고 AI 업스케일부(234)는 AI 데이터에 기초하여 저해상도 영상의 업스케일에 적합한 업스케일 DNN 정보를 획득하고, 상기 업스케일 DNN 정보에 따라 세팅된 업스케일 DNN을 이용하여 저해상도 영상을 AI 업스케일한다.
- [252] 일 실시 예에 따라, AI 데이터와 영상 데이터의 동기화를 위한 동기화 데이터가 AI 부호화 데이터에 포함될 수 있다. 동기화 데이터는 AI 데이터와 영상 데이터와 독립적으로 AI 부호화 데이터에 포함될 수 있다. 또는 동기화 데이터는 AI 데이터 또는 영상 데이터에 포함될 수 있다. 파싱부(214)는 동기화 데이터에 따라, 영상 데이터와 AI 데이터를 동기화할 수 있다. 또는 AI 업스케일부(234)가 동기화 데이터에 따라, 영상 데이터와 AI 데이터를 동기화할 수 있다. 동기화 데이터에 따라, 저해상도 영상을 AI 업스케일하기 위한 적절한 업스케일 DNN 정보가 선택될 수 있다.
- [253] 일 실시 예에 따르면, AI 데이터(1240)는 영상 데이터(1200)에 삽입(embed)될 수 있다. 예를 들어, 비디오 AI 데이터(1242)는 비디오 헤더(1204)에 삽입될 수 있다. 따라서 비디오 헤더(1204)는 비디오 파라미터 세트와 함께, 비디오 AI 데이터(1242)를 포함할 수 있다. 그러므로 비디오 AI 데이터(1242)는 비디오 파라미터 세트와 함께 복호화될 수 있다. 또는 비디오 AI 데이터(1242)는, 비디오 헤더(1204)로부터 독립적으로, 비디오 헤더(1204) 이전 또는 이후에 위치하도록 단일 파일에 삽입될 수 있다. 그러므로 비디오 AI 데이터(1242)의 복호화는 비디오 헤더(1204)의 복호화 바로 이전, 또는 바로 이후에 진행될 수 있다.
- [254] 마찬가지로 제1 프레임 그룹 헤더(1212)는 프레임 그룹 파라미터 세트와 함께, 제1 프레임 그룹 AI 데이터(1252)를 포함할 수 있다. 그러므로 제1 프레임 그룹

AI 데이터(1252)는 프레임 그룹 파라미터 세트와 함께 복호화될 수 있다. 또는 제1 프레임 그룹 AI 데이터(1252)는, 제1 프레임 그룹 헤더(1212)로부터 독립적으로, 제1 프레임 그룹 헤더(1212) 이전 또는 이후에 위치하도록 단일 파일에 삽입될 수 있다. 그러므로 제1 프레임 그룹 AI 데이터(1252)의 복호화는 제1 프레임 그룹 헤더(1212)의 복호화 바로 이전, 또는 바로 이후에 진행될 수 있다.

- [255] 마찬가지로 제1 프레임 헤더(1222)는 프레임 파라미터 세트와 함께, 제1 프레임 AI 데이터(1262)를 포함할 수 있다. 그러므로 제1 프레임 AI 데이터(1262)는 프레임 파라미터 세트와 함께 복호화될 수 있다. 또는 제1 프레임 AI 데이터(1262)는, 제1 프레임 헤더(1222)로부터 독립적으로, 제1 프레임 헤더(1222) 이전 또는 이후에 위치하도록 단일 파일에 삽입될 수 있다. 그러므로 제1 프레임 AI 데이터(1262)의 복호화는 제1 프레임 헤더(1222)의 복호화 바로 이전, 또는 바로 이후에 진행될 수 있다.
- [256] 도 13A에 기초하여, AI 데이터가 영상 데이터에 삽입된 경우, AI 복호화 장치(200)에서 데이터의 흐름이 설명된다.
- [257] AI 부호화 데이터에서, AI 데이터가 영상 데이터에 삽입된 경우, AI 데이터는 영상 데이터의 복호화 없이 독립적으로 복호화될 수 없다. 그러므로 제1 복호화부(232)가 영상 데이터를 복호화하는 중에, 영상 데이터에 삽입된 AI 데이터가 영상 데이터로부터 추출된다. 그리고 영상 데이터로부터 추출된 AI 데이터는 제1 복호화부(232)로부터 AI 업스케일부(234)로 전송된다.
- [258] 일 실시 예에 따르면, AI 데이터의 일부는 영상 데이터에 삽입되고, AI 데이터의 나머지 일부는 영상 데이터와 독립적으로 AI 부호화 데이터에 포함될 수 있다. 예를 들어, 비디오 AI 데이터는 영상 데이터와 독립적으로 AI 부호화 데이터에 포함되고, 프레임 그룹 AI 데이터 및 프레임 AI 데이터는 영상 데이터에 삽입될 수 있다.
- [259] 영상 데이터와 독립적으로 존재하는 제1 AI 데이터는 파싱부(214)에서 영상 데이터로부터 분리될 수 있다. 그리고 영상 데이터로부터 분리된 제1 AI 데이터는 출력부(216)로부터 AI 업스케일부(234)로 전송될 수 있다. 상기 제1 AI 데이터는 비디오 AI 데이터 및/또는 비디오 세그먼트 AI 데이터일 수 있다.
- [260] 그리고 영상 데이터에 삽입된 제2 AI 데이터는 제1 복호화부(232)에 의하여 영상 데이터로부터 추출된다. 상기 추출된 제2 AI 데이터는 제1 복호화부(232)로부터 AI 업스케일부(234)로 전송된다. 그리고 AI 업스케일부(234)는 제1 AI 데이터와 제2 AI 데이터에 따라, 저해상도 영상의 AI 업스케일에 필요한 업스케일 DNN 정보를 획득한다. 상기 제2 AI 데이터는 프레임 그룹 AI 데이터 및/또는 프레임 AI 데이터일 수 있다.
- [261] 일 실시 예에 따르면, 데이터 처리부(632)는 영상 데이터(1200)에 해당하는 파일과 AI 데이터(1240)에 해당하는 파일을 별도로 생성할 수 있다. 따라서 통신부(634)는 AI 부호화 데이터를 2개의 파일 형태로 통신부(212)로 전송한다.

따라서 통신부(634)는 영상 데이터(1200)에 해당하는 파일과 AI 데이터(1240)에 해당하는 파일을 다른 통신 채널로 전송할 수 있다. 또한 통신부(634)는 영상 데이터(1200)에 해당하는 파일과 AI 데이터(1240)에 해당하는 파일을 시간적 차이를 두고 순차적으로 전송할 수 있다.

- [262] 상기 실시 예에서, 도 12에서 설명된 영상 데이터(1200)와 AI 데이터(1240)의 상응 관계에 따라, AI 데이터(1240)에 해당하는 파일은 영상 데이터(1200)에 해당하는 파일의 복호화 프로세스에 종속적으로 복호화될 수 있다. 또한 영상 데이터(1200)에 해당하는 파일과 AI 데이터(1240)에 해당하는 파일이 분리되어 있으므로, 영상 데이터(1200)에 해당하는 파일과 AI 데이터(1240)에 해당하는 파일은 상기 2개의 파일의 동기화에 관한 정보를 포함할 수 있다.
- [263] 도 13B는 AI 데이터와 영상 데이터가 분리되어 2개의 파일로 구성되는 경우, AI 복호화 장치(200)에서의 데이터의 흐름을 설명한다.
- [264] 통신부(212)는 AI 부호화 데이터를 포함한 단일 파일 대신 영상 데이터를 포함한 파일과 AI 데이터를 포함하는 파일을 별도로 수신할 수 있다. 또한 통신부(212)는 영상 데이터와 AI 데이터의 동기화에 필요한 동기화 데이터를 영상 데이터를 포함한 파일 또는 AI 데이터를 포함하는 파일로부터 획득할 수 있다. 실시 예에 따라, 동기화 데이터는 별도의 파일로부터 전송될 수 있다. 도 13B에서는 동기화 데이터가 AI 데이터와 독립적인 데이터로 표현되었으나, 실시 예에 따라, AI 데이터 또는 영상 데이터가 동기화 데이터를 포함할 수 있다.
- [265] 일 실시 예에 따라, 파싱부(214)는 영상 데이터와 AI 데이터를 동기화 데이터에 따라 동기화할 수 있다. 그리고 출력부(216)는 동기화된 영상 데이터를 제1 복호화부(232)에, 동기화된 AI 데이터를 AI 업스케일부(234)에 전송할 수 있다.
- [266] 일 실시 예에 따라, 출력부(216)는 영상 데이터를 제1 복호화부(232)에 전송할 수 있다. 그리고 AI 데이터와 동기화 데이터를 AI 업스케일부(234)에 전송할 수 있다. 그리고 AI 업스케일부(234)는 제1 복호화부(232)로부터 출력된 저해상도 영상을 동기화 데이터 및 AI 데이터에 따라 획득된 업스케일 DNN 정보를 이용하여 AI 업스케일한다.
- [267] 도 14는 단일 파일 내에서 AI 데이터와 영상 데이터가 분리되어 있는 경우의 AI 부호화 데이터의 일 실시 예를 도시한다. 도 14에서 AI 부호화 데이터는 소정 컨테이너 포맷의 비디오 파일(1400)에 포함된다. 상기 소정 컨테이너 포맷은 MP4, AVI, MKV, FLV 등일 수 있다.
- [268] 비디오 파일(1400)은 메타데이터 박스(1410)와 미디어데이터 박스(1430)로 구성된다. 메타데이터 박스(1410)에는 미디어데이터 박스(1430)에 포함된 미디어 데이터에 관한 정보가 포함된다. 예를 들어, 메타데이터 박스(1410)에는 미디어 데이터의 종류, 미디어 데이터의 부호화에 사용된 코덱의 종류 및 미디어의 재생 시간 등에 관한 정보가 포함될 수 있다.
- [269] 메타데이터 박스(1410)에는 동기화 데이터(1415) 및 AI 데이터(1420)가 포함될 수 있다. 동기화 데이터(1415) 및 AI 데이터(1420)는 소정 컨테이너 포맷에서

- 제공하는 부호화 방식에 따라 부호화되어, 메타데이터 박스(1410)에 저장된다.
- [270] 파싱부(214)는 메타데이터 박스(1410)로부터 동기화 데이터(1415) 및 AI 데이터(1420)를 추출할 수 있다. 그리고 파싱부(214)는 미디어데이터 박스(1430)로부터 영상 데이터(1431)를 추출할 수 있다. 출력부(216)는 동기화 데이터(1415)에 따라, 영상 데이터(1431)를 제1 복호화부(232)에 전송하고, AI 데이터(1420)를 AI 업스케일부(234)에 전송할 수 있다. 또는 출력부(216)는 동기화 데이터(1415)를 AI 업스케일부(234)에 전송하고, AI 업스케일부(234)가, 동기화 데이터(1415)에 따라, 영상 데이터(1431)와 AI 데이터(1420)의 동기화를 할 수 있다.
- [271] AI 데이터(1420)는 비디오 AI 데이터(1422), 프레임 그룹 AI 데이터(1424) 및 프레임 AI 데이터(1426)을 포함한다. 비디오 AI 데이터(1422)는 비디오 헤더(1432)와, 프레임 그룹 AI 데이터(1424)는 프레임 그룹 헤더(1436)와, 프레임 AI 데이터(1426)는 프레임 헤더(1440)와 상응하도록 설정된다.
- [272] 실시 예에 따라, 비디오의 모든 프레임들에 동일한 업스케일 DNN 정보가 이용될 경우, AI 데이터(1420)에서 프레임 그룹 AI 데이터(1424)와 프레임 AI 데이터(1426)는 생략될 수 있다. 또는 업스케일 DNN 정보가 프레임 그룹 단위마다 적응적으로 획득되고, 프레임 그룹의 모든 프레임들에 동일한 업스케일 DNN 정보가 사용될 경우, AI 데이터(1420)에서 프레임 AI 데이터(1426)는 생략될 수 있다.
- [273] 동기화 데이터(1415)는 비디오 AI 데이터(1422), 프레임 그룹 AI 데이터(1424) 및 프레임 AI 데이터(1426)와 비디오 헤더(1432), 프레임 그룹 헤더(1436), 프레임 헤더(1440)의 동기화에 관한 정보를 포함한다. 예를 들어, 동기화 데이터(1415)는 미디어데이터 박스(1430)의 영상 데이터(1431)의 재생 순서 정보 또는 복호화 순서 정보를 포함할 수 있다. 따라서 AI 업스케일부(234)는, 동기화 데이터(1415)의 재생 순서 정보 또는 복호화 순서 정보에 따라 결정된 AI 데이터로부터, 저해상도 영상의 AI 업스케일에 필요한 업스케일 DNN 정보를 획득할 수 있다.
- [274] 예를 들어, 프레임(1438)의 저해상도 영상을 AI 업스케일하기 위하여, 파싱부(214) 또는 AI 업스케일부(234)는 동기화 데이터(1415)에 기초하여 프레임 그룹(1434)에 대응되는 프레임 그룹 AI 데이터(1424)와 프레임(1438)에 대응되는 프레임 AI 데이터(1426)를 결정할 수 있다. 그리고 프레임 그룹 AI 데이터(1424) 및 프레임 AI 데이터(1426)와 함께 모든 프레임에 적용되는 비디오 AI 데이터(1422)를 고려하여, AI 업스케일부(234)는 프레임(1438)의 저해상도 영상의 AI 업스케일을 위한 업스케일 DNN 정보를 획득할 수 있다.
- [275] 메타데이터 박스(1410)의 AI 데이터(1420)는 미디어 데이터 박스(1430)의 영상 데이터(1431)보다 먼저 복호화될 수 있다. 따라서, AI 업스케일부(234)는, AI 데이터(1420)에 따라, 업스케일 DNN 정보를 영상 데이터(1431)의 복호화 이전에 획득할 수 있다. 상기 업스케일 DNN 정보는 비디오 전체에 대하여 적용될 수

있다. 또는 프레임 그룹 단위의 적응적 AI 업스케일을 위하여, 프레임 그룹마다 다른 업스케일 DNN 정보가 획득될 수 있다. 또는, 프레임 단위의 적응적 AI 업스케일을 위하여, 업스케일 DNN 정보는 프레임마다 다르게 미리 획득될 수 있다.

- [276] AI 업스케일부(234)는 메타데이터 박스(1410)의 AI 데이터(1420)를, 미디어 데이터 박스(1430)의 영상 데이터(1431)의 복호화 순서에 따라, 복호화할 수 있다. 비디오 AI 데이터(1422)의 복호화는 비디오 헤더(1432)의 복호화 바로 이전 또는 이후에 수행될 수 있다. 또는 비디오 AI 데이터(1422)는 비디오 헤더(1432)와 병렬적으로 복호화될 수 있다. 비디오 AI 데이터(1422)를 비디오 헤더(1432)의 복호화 순서에 따라 복호화하기 위하여, 동기화 데이터(1415)가 참조될 수 있다.
- [277] 마찬가지로, AI 업스케일부(234)는 프레임 그룹 AI 데이터(1424)의 복호화를, 제1 복호화부(232)에 따른 프레임 그룹 헤더(1436)의 복호화 바로 이전 또는 이후에, 수행할 수 있다. 또는 AI 업스케일부(234)는 프레임 그룹 AI 데이터(1424)를, 제1 복호화부(232)에 따른 프레임 그룹 헤더(1436)의 복호화와 병렬적으로, 복호화할 수 있다. 프레임 그룹 AI 데이터(1424)를 프레임 그룹 헤더(1436)의 복호화 순서에 따라 복호화하기 위하여, 동기화 데이터(1415)가 참조될 수 있다.
- [278] 그리고 AI 업스케일부(234)는 프레임 AI 데이터(1426)의 복호화를, 제1 복호화부(232)에 따른 프레임 헤더(1440)의 복호화 바로 이전 또는 이후에, 수행할 수 있다. 또는 AI 업스케일부(234)는 프레임 AI 데이터(1426)는, 제1 복호화부(232)에 따른 프레임 헤더(1440)의 복호화와 병렬적으로, 복호화될 수 있다. 프레임 AI 데이터(1426)를 프레임 헤더(1440)의 복호화 순서에 따라 복호화하기 위하여, 동기화 데이터(1415)가 참조될 수 있다.
- [279] 도 14에서는 설명의 편의를 위하여, 비디오 파일(1400)에 하나의 메타데이터 박스(1410)와 하나의 미디어데이터 박스(1430)가 포함된 것으로 설명되었다. 그러나 비디오 파일(1400)은 2개 이상의 메타데이터 박스와 2개 이상의 미디어 데이터 박스를 포함할 수 있다. 따라서, 영상 데이터가 소정의 시간 단위로 분할된 2개 이상의 영상 데이터 세그먼트들이, 2개 이상의 미디어 데이터 박스에 저장될 수 있다. 그리고 상기 2개 이상의 미디어 데이터 박스에 저장된 영상 데이터 세그먼트들에 대한 정보가 2개 이상의 메타 데이터 박스에 포함될 수 있다. 또한 2개 이상의 메타데이터 박스는 각각 AI 데이터를 포함할 수 있다.
- [280] 도 15A는 AI 데이터가 영상 데이터에 삽입된 경우의 AI 부호화 데이터의 일 실시 예를 도시한다.
- [281] 비디오 파일(1500)은, 도 14의 비디오 파일(1400)과 마찬가지로, 메타데이터 박스(1502)와 미디어데이터 박스(1504)을 포함한다. AI 데이터가 영상 데이터에 삽입된 바, 메타데이터 박스(1502)는 AI 데이터를 포함하지 않는다. 대신 AI 데이터가 삽입된 영상 데이터가 미디어데이터 박스(1504)에 포함된다. AI

데이터는 영상 데이터의 비디오 코덱에 따라 부호화될 수 있다. 예를 들어, 상기 영상 데이터의 비디오 코덱은 H.264, HEVC, AVS2.0, Xvid 등일 수 있다.

[282] 따라서 파싱부(214)는 메타데이터 박스(1502)로부터 AI 데이터를 추출하지 않는다. 대신 제1 복호화부(232)는 영상 데이터(1505)로부터 AI 데이터를 추출하고, 추출된 AI 데이터를 AI 업스케일부(234)에 전달할 수 있다. 그리고 상기 AI 데이터에 의해 획득된 업스케일 DNN 정보에 기초하여, AI 업스케일부(234)는 제1 복호화부(232)에 의하여 복원된 저해상도 영상을 AI 업스케일 할 수 있다.

[283] 도 14의 비디오 파일(1400)과 다르게, AI 데이터가 영상 데이터(1505) 안에서 영상 데이터(1505)의 복호화 순서에 따라 삽입되어 있으므로, 비디오 파일(1500)은 동기화 데이터를 포함하지 않을 수 있다. 따라서, AI 데이터는 영상 데이터(1505)가 복호화됨에 따라, 순차적으로 복호화된다.

[284] 비디오 AI 데이터(1508)는 비디오의 파라미터 정보를 포함하는 비디오 헤더(1506)의 바로 이후에 위치할 수 있다. 따라서 비디오 AI 데이터(1508)는 비디오 헤더(1506)에 포함된 비디오 파라미터들이 복호화된 후에 복호화될 수 있다. 실시 예에 따라, 도 15A와 다르게, 비디오 AI 데이터(1508)가 비디오 헤더(1506)의 바로 이전에 위치할 수 있다.

[285] 프레임 그룹 AI 데이터(1514)는 프레임 그룹(1510)의 파라미터 정보를 포함하는 프레임 그룹 헤더(1512)의 바로 이후에 위치할 수 있다. 따라서 프레임 그룹 AI 데이터(1514)는 프레임 그룹 헤더(1512)에 포함된 프레임 그룹 파라미터들이 복호화된 후에 복호화될 수 있다. 실시 예에 따라, 도 15A와 다르게, 프레임 그룹 AI 데이터(1514)가 프레임 그룹 헤더(1512)의 바로 이전에 위치할 수 있다. 프레임 그룹(1510) 이후에 복호화되는 나머지 프레임 그룹들의 프레임 그룹 헤더와 프레임 그룹 AI 데이터의 복호화 순서도 프레임 그룹 헤더(1512)와 프레임 그룹 AI 데이터(1514)의 복호화 순서와 동일하게 결정될 수 있다.

[286] 프레임 AI 데이터(1520)는 프레임(1516)의 파라미터 정보를 포함하는 프레임 헤더(1518)의 바로 이후에 위치할 수 있다. 따라서 프레임 AI 데이터(1520)는 프레임 헤더(1518)에 포함된 프레임 그룹 파라미터들이 복호화된 후에 복호화될 수 있다. 실시 예에 따라, 도 15A와 다르게, 프레임 AI 데이터(1520)가 프레임 헤더(1518)의 바로 이전에 위치할 수 있다. 프레임 (1516) 이후에 복호화되는 나머지 프레임들의 프레임 헤더와 프레임 AI 데이터의 복호화 순서도 프레임 헤더(1518)와 프레임 AI 데이터(1520)의 복호화 순서와 동일하게 결정될 수 있다.

[287] 도 15B는 AI 데이터가 영상 데이터에 삽입된 경우의 AI 부호화 데이터의 다른 실시 예를 도시한다.

[288] 도 15A의 비디오 파일(1500)과 마찬가지로, 비디오 파일(1520)은 메타데이터 박스(1522)와 미디어데이터 박스(1524)를 포함한다. 그리고 메타데이터 박스(1522)는 AI 데이터를 포함하지 않고, 대신 AI 데이터가 삽입된 영상

데이터(1525)가 미디어데이터 박스(1524)에 포함된다. 다만 도 15A의 비디오 파일(1500)과 다르게, 비디오 파일(1520)에는 AI 데이터가 대응되는 데이터 헤더에 삽입된다.

- [289] 비디오 AI 데이터(1528)는 비디오의 파라미터 정보를 포함하는 비디오 헤더(1526)에 포함할 수 있다. 따라서 비디오 AI 데이터(1528)는 비디오 헤더(1526)에 포함된 비디오 파라미터들과 함께 복호화될 수 있다.
- [290] 프레임 그룹 AI 데이터(1534)는 프레임 그룹(1530)의 파라미터 정보를 포함하는 프레임 그룹 헤더(1532)에 포함할 수 있다. 따라서 프레임 그룹 AI 데이터(1534)는 프레임 그룹 헤더(1532)에 포함된 프레임 그룹 파라미터들과 함께 복호화될 수 있다. 프레임 그룹(1530) 이후에 복호화되는 나머지 프레임 그룹들의 프레임 그룹 AI 데이터도 프레임 그룹 헤더에 포함될 수 있다.
- [291] 프레임 AI 데이터(1540)는 프레임(1536)의 파라미터 정보를 포함하는 프레임 헤더(1538)에 포함할 수 있다. 따라서 프레임 AI 데이터(1540)는 프레임 헤더(1538)에 포함된 프레임 그룹 파라미터들과 함께 복호화될 수 있다. 프레임(1536) 이후에 복호화되는 나머지 프레임들의 프레임 AI 데이터도 프레임 헤더에 포함될 수 있다.
- [292] 도 15C는 일부 AI 데이터가 영상 데이터에 삽입되고 나머지 AI 데이터는 영상 데이터와 분리된 경우의 AI 부호화 데이터의 일 실시 예를 도시한다.
- [293] 도 15A의 비디오 파일(1500)과 마찬가지로, 비디오 파일(1550)은 메타데이터 박스(1552)와 미디어데이터 박스(1556)를 포함한다. 메타데이터 박스(1552)는 비디오의 모든 프레임에 적용되는 비디오 AI 데이터(1554)를 포함한다. 그리고 프레임 그룹 AI 데이터와 프레임 AI 데이터는 미디어데이터 박스(1556)의 영상 데이터에 포함된다.
- [294] 도 15A의 비디오 AI 데이터(1508) 및 도 15B의 비디오 AI 데이터(1528)와 달리 메타데이터 박스(1552)에 포함된 비디오 AI 데이터(1554)는 영상 데이터의 복호화 이전에 복호화될 수 있다. 그리고 프레임 그룹 AI 데이터와 프레임 AI 데이터는 영상 데이터가 복호화됨에 따라, 순차적으로 복호화된다.
- [295] 따라서 파싱부(214)는 메타데이터 박스(1552)로부터 비디오 AI 데이터(1554)를 추출할 수 있다. 그리고 통신부(216)는 비디오 AI 데이터(1554)를 AI 업스케일부(234)에 전송할 수 있다. 또한 통신부(216)는 영상 데이터(1557)를 제1 복호화부(232)에 전송할 수 있다. 그리고 제1 복호화부(232)는, 영상 데이터(1557)를 복호화함으로써, 저해상도 영상을 복원하고, 프레임 그룹 AI 데이터(1564)와 프레임 AI 데이터(1570)를 추출할 수 있다. 또한 제1 복호화부(232)는 프레임 그룹 AI 데이터(1564)와 프레임 AI 데이터(1570)를 AI 업스케일부(234)에 전송할 수 있다. 그리고 AI 업스케일부(234)는 비디오 AI 데이터(1554), 프레임 그룹 AI 데이터(1564), 및 프레임 AI 데이터(1570)에 따라 저해상도 영상의 AI 업스케일을 위한 업스케일 DNN 정보를 획득할 수 있다.
- [296] 프레임 그룹 AI 데이터(1564)는 프레임 그룹(1560)의 파라미터 정보를

포함하는 프레임 그룹 헤더(1562)의 바로 이후에 위치할 수 있다. 그러나 실시 예에 따라, 프레임 그룹 AI 데이터(1564)는 프레임 그룹 헤더(1562)의 바로 이전에 위치할 수 있다. 또한 프레임 그룹 AI 데이터(1564)는 프레임 그룹 헤더(1562)에 포함될 수 있다.

[297] 프레임 AI 데이터(1570)는 프레임(1566)의 파라미터 정보를 포함하는 프레임 헤더(1568)의 바로 이후에 위치할 수 있다. 그러나 실시 예에 따라, 프레임 AI 데이터(1570)가 프레임 헤더(1568)의 바로 이전에 위치할 수 있다. 또한 프레임 AI 데이터(1570)는 프레임 헤더(1568)에 포함될 수 있다.

[298] 도 15C에서는, 비디오 AI 데이터(1554)만 메타데이터 박스(1552)에 포함된다고 표현되었으나, 프레임 그룹 AI 데이터가 추가적으로 메타데이터 박스(1552)에 포함될 수 있다. 또는 프레임 그룹 AI 데이터의 일부가 메타데이터 박스(1552)에 포함될 수 있다. 또는 프레임 AI 데이터의 일부가 메타데이터 박스(1552)에 포함될 수 있다.

[299] 도 15A 및 15C에서 미디어데이터 박스에 삽입된 프레임 그룹 AI 데이터(1514, 1564) 및 프레임 AI 데이터(1520, 1570)는 SEI (Supplementary Enhancement Information) 메시지의 형태로 삽입될 수 있다. SEI 메시지는 영상의 복호화에 반드시 필요하지 않은 영상에 관한 부가 정보가 포함되는 데이터 유닛이다. SEI 메시지는 프레임 그룹 단위 또는 프레임 단위로 전송될 수 있다. 그리고 SEI 메시지는 도 13A에서 기술한 제2 AI 데이터와 마찬가지로 제 1 복호화부(232)에서 추출되어 AI 업스케일부(234)로 전달될 수 있다.

[300] 도 16은 도 14와 같이 AI 데이터와 영상 데이터가 분리된 경우, 비디오 세그먼트 단위로 분할된 AI 부호화 데이터의 일 실시 예를 도시한다.

[301] AI 데이터를 비롯한 메타 데이터의 크기가 크거나, 영상 데이터가 통신 채널을 통해 라이브 스트리밍되는 경우, AI 데이터를 비롯한 메타 데이터의 처리에 필요한 오버헤드를 줄이기 위하여, AI 부호화 데이터가 비디오 세그먼트 단위로 분할되어 비디오 파일(1600)에 저장될 수 있다. 비디오 세그먼트는 비디오의 일부분으로, 소정의 시간동안의 프레임을 포함한다. 예를 들어, 비디오 세그먼트는 1개 이상의 프레임 그룹만 포함할 수 있다. 만약, 비디오 세그먼트가 1개의 프레임 그룹을 포함할 경우, 비디오 파일(1600)은 영상 데이터의 프레임 그룹의 개수만큼의 비디오 세그먼트들을 포함할 수 있다.

[302] 비디오 세그먼트는 각각 메타데이터 박스와 미디어데이터 박스를 포함할 수 있다. AI 데이터를 비롯한 메타데이터가 복수의 비디오 세그먼트에 분할되어 저장된다. 따라서, AI 데이터가 영상 데이터 전체에 대한 메타데이터 박스(1610)뿐만 아니라, 비디오 세그먼트들에 대한 메타데이터 박스들에 분할되어 저장된다. 그러므로, 영상 데이터의 특정 부분을 저해상도 영상을 AI 업스케일할 때, 영상 데이터 전체에 대한 메타데이터 박스(1610)에 모든 AI 데이터가 저장된 경우에 비하여, AI 데이터가 상기 메타데이터 박스(1610)를 비롯하여 비디오 세그먼트들에 대한 메타데이터 박스들에 분산되어 저장된

경우, 더 빠르게 상기 특정 부분에 적합한 업스케일 DNN 정보를 획득하기 위한 AI 데이터가 획득될 수 있다.

- [303] 예를 들어, 현재 비디오 세그먼트부터 재생이 될 때, 현재 비디오 세그먼트의 저해상도 영상의 AI 업스케일을 위하여, 메타데이터 박스(1610)의 비디오 AI 데이터(1612)와 현재 비디오 세그먼트 데이터 박스(1620)의 세그먼트 메타데이터 박스(1630)의 AI 데이터만 참조된다. 그리고 다른 비디오 세그먼트들의 메타데이터 박스들의 AI 데이터는 참조되지 않는다. 따라서, AI 데이터의 복호화에 따른 오버헤드가 감소할 수 있다.
- [304] 또한, 영상 데이터가 라이브 스트리밍되는 경우, AI 데이터가 최초 재생시에 한 번에 전송되지 않고, 비디오 세그먼트 단위로 분산되어 전송될 수 있다. 따라서, AI 데이터가 분산되어 순차적으로 전송됨으로써, AI 데이터의 복호화에 따른 오버헤드가 감소할 수 있다. 그러므로 비디오 세그먼트 단위로 분할하여 전송하는 것이 유리할 수 있다.
- [305] 도 16에서, 전체 영상 데이터에 대한 메타데이터 박스(1610)는 비디오 AI 데이터(1612)를 포함한다. 비디오 AI 데이터(1612)는 비디오에 포함된 모든 비디오 세그먼트들에 적용된다. 메타데이터 박스(1610)는 현재 비디오 세그먼트 데이터 박스(1620)보다 먼저 복호화될 수 있다.
- [306] 현재 비디오 세그먼트 데이터 박스(1620)는 세그먼트 메타데이터 박스(1630)와 세그먼트 미디어데이터 박스(1640)를 포함한다. 세그먼트 메타데이터 박스(1630)는 동기화 데이터(1631)와 AI 데이터(1632)를 포함할 수 있다. 그리고 세그먼트 미디어데이터 박스(1640)는 비디오 세그먼트 데이터(1641)을 포함한다.
- [307] 현재 비디오 세그먼트 데이터 박스(1620)의 AI 데이터(1632)는 비디오 세그먼트 AI 데이터(1634), 프레임 그룹 AI 데이터(1636) 및 프레임 AI 데이터(1638)를 포함할 수 있다. 비디오 세그먼트 AI 데이터(1634)는 현재 비디오 세그먼트에 포함된 모든 프레임 그룹들에 적용된다. 그리고 프레임 그룹 AI 데이터(1636) 현재 프레임 그룹에 포함된 모든 프레임들에 적용된다. 프레임 AI 데이터(1638)는 현재 프레임에 적용된다.
- [308] 실시 예에 따라, 비디오 세그먼트의 모든 프레임들에 동일한 업스케일 DNN 정보가 이용될 경우, 현재 비디오 세그먼트 데이터 박스(1620)의 AI 데이터(1632)에서 프레임 그룹 AI 데이터(1636)와 프레임 AI 데이터(1638)는 생략될 수 있다. 또는 업스케일 DNN 정보가 프레임 그룹 단위마다 적응적으로 결정되고, 프레임 그룹의 모든 프레임들에 동일한 업스케일 DNN 정보가 사용될 경우, 현재 비디오 세그먼트 데이터 박스(1620)의 AI 데이터(1632)에서 프레임 AI 데이터(1638)는 생략될 수 있다.
- [309] 실시 예에 따라, 비디오 세그먼트가 하나의 프레임 그룹만 포함할 경우, 현재 비디오 세그먼트 데이터 박스(1620)의 AI 데이터(1632)에서 비디오 세그먼트 AI 데이터(1634)는 생략될 수 있다. 그리고 프레임 그룹 AI 데이터(1636)가 비디오

세그먼트 AI 데이터(1634)의 역할을 수행할 수 있다. 또는 비디오 세그먼트가 하나의 프레임 그룹만 포함하고, 프레임 그룹의 모든 프레임들에 동일한 업스케일 DNN 정보가 사용될 경우, 현재 비디오 세그먼트 데이터 박스(1620)의 AI 데이터(1632)에서 비디오 세그먼트 AI 데이터(1634)와 프레임 AI 데이터(1638)는 생략될 수 있다. 그리고 프레임 그룹 AI 데이터(1636)에 따른 업스케일 DNN 정보에 기초하여, 비디오 세그먼트의 모든 프레임에 AI 업스케일이 적용될 수 있다. 동기화 데이터(1631)는 비디오 세그먼트 AI 데이터(1634), 프레임 그룹 AI 데이터(1636) 및 프레임 AI 데이터(1638)와 비디오 세그먼트 헤더(1642), 프레임 그룹 헤더(1646), 프레임 헤더(1648)의 동기화에 관한 정보를 포함한다. 비디오 세그먼트 헤더(1642)는 비디오 세그먼트에 포함된 프레임들에 공통적으로 적용되는 비디오 세그먼트 파라미터를 포함한다. 예를 들어, 동기화 데이터(1631)는 세그먼트 미디어데이터 박스(1640)의 비디오 세그먼트 데이터(1641)의 재생 순서 정보 또는 복호화 순서 정보를 포함할 수 있다. 따라서 AI 업스케일부(234)는, 동기화 데이터(1631)에 따라 결정된 AI 데이터로부터, 저해상도 영상의 AI 업스케일에 필요한 업스케일 DNN 정보를 획득할 수 있다.

[310] 예를 들어, 프레임(1649)의 저해상도 영상을 AI 업스케일하기 위하여, 파싱부(214) 또는 AI 업스케일부(234)는 동기화 데이터(1631)에 기초하여 프레임 그룹(1644)에 대응되는 프레임 그룹 AI 데이터(1636)와 프레임(1649)에 대응되는 프레임 AI 데이터(1636)를 획득할 수 있다. 그리고 프레임 그룹 AI 데이터(1636) 및 프레임 AI 데이터(1638)와 함께 모든 프레임에 적용되는 비디오 세그먼트 AI 데이터(1634)를 고려하여, AI 업스케일부(234)는 프레임(1649)의 저해상도 영상의 AI 업스케일을 위한 업스케일 DNN 정보를 획득할 수 있다.

[311] 일 실시 예에 따라, 세그먼트 메타데이터 박스(1630)의 AI 데이터(1632)는 세그먼트 미디어데이터 박스(1640)의 비디오 세그먼트 데이터(1641)보다 먼저 복호화될 수 있다. 따라서, AI 업스케일부(234)는, AI 데이터(1632)에 따라, 업스케일 DNN 정보를 비디오 세그먼트 데이터(1641)의 복호화 이전에 획득할 수 있다. 상기 획득된 업스케일 DNN 정보는 비디오 세그먼트 전체에 대하여 적용될 수 있다. 또는 프레임 그룹 단위의 적응적 AI 업스케일을 위하여, 업스케일 DNN 정보는 프레임 그룹마다 다르게 미리 획득될 수 있다. 또는, 프레임 단위의 적응적 AI 업스케일을 위하여, 업스케일 DNN 정보는 프레임마다 다르게 미리 획득될 수 있다.

[312] 일 실시 예에 따라, AI 업스케일부(234)는 세그먼트 메타데이터 박스(1630)의 AI 데이터(1632)를, 세그먼트 미디어데이터 박스(1640)의 비디오 세그먼트 데이터(1641)의 복호화 순서에 따라, 복호화할 수 있다. 예를 들어, AI 업스케일부(234)는 프레임 그룹 AI 데이터(1636)의 복호화를, 제1 복호화부(232)에 따른 프레임 그룹 헤더(1646)의 복호화 순서에 따라, 수행할 수 있다. 그리고 AI 업스케일부(234)는 프레임 AI 데이터(1638)의 복호화를, 제1

복호화부(232)에 따른 프레임 헤더(1648)의 복호화 순서에 따라, 수행할 수 있다. 동기화 데이터(1631)를 참조하여, 비디오 세그먼트 AI 데이터(1634), 프레임 그룹 AI 데이터(1636) 및 프레임 AI 데이터(1638)의 복호화 순서와 비디오 세그먼트 헤더(1642), 프레임 그룹 헤더(1646), 프레임 헤더(1648)의 복호화 순서가 동기화될 수 있다.

- [313] 현재 비디오 세그먼트 데이터 박스(1620) 이후의 나머지 비디오 세그먼트 데이터 박스들도 순차적으로 현재 비디오 세그먼트 데이터 박스(1620)와 동일한 방법으로 복호화될 수 있다.
- [314] 도 17은 2개의 별도 파일로 전송되는 AI 데이터(1740)와 영상 데이터(1700)의 일 실시 예를 도시한다.
- [315] 통신부(212)가 AI 데이터(1740)를 수신하지 않은 경우, 영상 데이터(1700)로부터 획득된 저해상도 영상은 AI 업스케일되지 않는다. 만약 통신부(212)가 AI 데이터(1740)를 수신한 경우, 출력부(216)를 거쳐 AI 업스케일부(234)에 전송된 AI 데이터(1740)에 따라 저해상도 영상의 AI 업스케일에 필요한 업스케일 DNN 정보가 획득된다.
- [316] 영상 데이터(1700)는 비디오 헤더(1710), 프레임 그룹(1720)의 프레임 그룹 헤더(1722) 및 프레임(1730)의 프레임 헤더(1732)를 포함할 수 있다. 그리고 AI 데이터(1740)는 비디오 AI 데이터(1742), 프레임 그룹 AI 데이터(1750), 프레임 AI 데이터(1760)을 포함할 수 있다. 영상 데이터(1700)와 AI 데이터(1740)는 별개의 파일로 전송되므로, 영상 데이터(1700) 및/또는 AI 데이터(1740)는 영상 데이터(1700)와 AI 데이터(1740)의 동기화에 필요한 동기화 데이터를 포함할 수 있다. 동기화 데이터는 영상 데이터(1700)의 복호화 순서 또는 재생 순서를 나타낼 수 있다.
- [317] 파싱부(214) 또는 AI 업스케일부(234)는 동기화 데이터에 따라 비디오 AI 데이터(1742), 프레임 그룹 AI 데이터(1750), 프레임 AI 데이터(1760)의 재생 순서 또는 복호화 순서가 비디오 헤더(1710), 프레임 그룹 헤더(1722) 및 프레임 헤더(1732)의 재생 순서 또는 복호화 순서와 일치하도록 설정할 수 있다. 도 17의 비디오 AI 데이터(1742)와 비디오 헤더(1710) 간의 점선, 프레임 그룹 AI 데이터(1750)와 프레임 그룹 헤더(1722) 간의 점선, 및 프레임 AI 데이터(1760)와 프레임 헤더(1732) 간의 점선은 AI 데이터와 데이터 헤더 간의 동기화를 나타낸다.
- [318] 영상 데이터(1700)와 AI 데이터(1740)에 두 데이터의 매칭을 위한 식별번호가 포함될 수 있다. 예를 들어, AI 데이터(1740)는 AI 데이터(1740)가 적용되는 영상 데이터(1700)의 식별번호를 포함할 수 있다. 반대로, 영상 데이터(1700)가 영상 데이터(1700)에 적용되는 AI 데이터(1740)의 식별 번호를 포함할 수 있다. 또는 영상 데이터(1700)와 AI 데이터(1740)에 모두 식별번호가 포함될 수 있다. 따라서, 영상 데이터(1700)와 AI 데이터(1740)가 동시에 전송되지 않은 경우에도, 영상 데이터(1700)와 AI 데이터(1740)는 식별 번호에 따라 서로 매칭될 수 있다.

- [319] 실시 예에 따라, 프레임 그룹 단위로 업스케일 DNN 정보가 획득되는 경우, AI 데이터(1740)에서 프레임 AI 데이터(1760)는 생략될 수 있다. 그리고 비디오의 모든 프레임에 동일한 업테이트 DNN 정보가 사용되는 경우, AI 데이터(1740)에서 프레임 그룹 AI 데이터(1750) 및 프레임 AI 데이터(1760)는 생략될 수 있다.
- [320] 도 18A는 도 14 내지 도 17에서 설명된 비디오 AI 데이터에 적용될 수 있는 데이터 구조(1800)의 일 실시 예를 나타낸다.
- [321] 비디오 AI 데이터의 데이터 구조(1800)에는 AI 업스케일에 사용되는 업스케일 DNN 정보에 관련된 엘리먼트들이 포함된다. 상기 엘리먼트들은 ai\_codec\_info (1802), ai\_codec\_applied\_channel\_info (1804), target\_bitrate\_info (1806), res\_info (1808), ai\_codec\_DNN\_info (1814), ai\_codec\_supplementary\_info (1816) 등을 포함할 수 있다. 도 18A에 나타난 엘리먼트의 배열 순서는 예시일 뿐, 통상의 기술자는 상기 엘리먼트의 배열 순서를 변경하여 실시할 수 있다.
- [322] 일 실시 예에서, ai\_codec\_info (1802)는 저해상도 영상에 AI 업스케일이 적용되는지 여부를 나타내는 AI 타겟 데이터를 의미한다. ai\_codec\_info (1802)가 영상 데이터에 따라 복원된 저해상도 영상에 AI 업스케일이 적용됨을 나타낼 경우, 데이터 구조(1800)는 AI 업스케일에 사용되는 업스케일 DNN 정보를 획득하기 위한 AI 보조 데이터를 포함한다. 반대로 ai\_codec\_info (1802)가 영상 데이터에 따라 복원된 저해상도 영상에 AI 업스케일이 적용되지 않음을 나타낼 경우, 데이터 구조(1800)는 AI 업스케일에 관한 AI 보조 데이터를 포함하지 않는다. 상기 AI 보조 데이터는 ai\_codec\_applied\_channel\_info (1804), target\_bitrate\_info (1806), res\_info (1808), ai\_codec\_DNN\_info (1814), ai\_codec\_supplementary\_info (1816) 등이 포함된다.
- [323] ai\_codec\_applied\_channel\_info (1804)는 AI 업스케일이 적용되는 색 채널을 나타내는 채널 정보이다. 영상은 RGB 포맷, YUV 포맷, YCbCr 포맷 등으로 표현될 수 있다. 만약 영상 데이터로부터 복원된 저해상도 영상이 YCbCr 포맷일 경우, 저해상도 영상은 휘도(luminance)에 관한 Y 채널의 저해상도 영상과 색차(chrominance)에 관한 Cb 채널의 저해상도 영상, 및 Cr 채널의 저해상도 영상을 포함한다. ai\_codec\_applied\_channel\_info (1804)는 3개의 채널 중 AI 업스케일이 적용되는 색 채널을 나타낼 수 있다. 그리고 AI 업스케일부(234)는 ai\_codec\_applied\_channel\_info (1804)가 나타내는 색 채널의 저해상도 영상을 AI 업스케일할 수 있다. AI 업스케일부(234)는 색 채널마다 서로 다른 업스케일 DNN 정보를 획득할 수 있다.
- [324] 실시 예에 따라, AI 업스케일에 따른 연산량을 감소시키기 위해, 휘도에 관한 Y 채널의 저해상도 영상에 대하여만 AI 업스케일이 적용될 수 있다. 사람의 눈은 색차보다는 휘도의 화질에 민감하게 반응하므로, 모든 색 채널의 저해상도 영상이 AI 업스케일되는 경우와 Y 채널의 저해상도 영상만 AI 업스케일되는 경우 간에, 사람이 느끼는 주관적 화질은 차이가 미미할 수 있다.

- [325] 그러므로 `ai_codec_applied_channel_info` (1804)는 Cb 채널의 저해상도 영상, 및 Cr 채널의 저해상도 영상이 AI 업스케일되는지 여부를 나타낼 수 있다. `ai_codec_applied_channel_info` (1804)가 Cb 채널의 저해상도 영상, 및 Cr 채널의 저해상도 영상이 AI 업스케일되지 않는다고 나타낼 경우, AI 업스케일부(234)는 Y 채널의 저해상도 영상에 대하여만 AI 업스케일을 적용할 수 있다. 반대로 `ai_codec_applied_channel_info` (1804)가 Cb 채널의 저해상도 영상, 및 Cr 채널의 저해상도 영상이 AI 업스케일된다고 나타낼 경우, AI 업스케일부(234)는 모든 채널의 저해상도 영상들에 대하여 AI 업스케일을 적용할 수 있다.
- [326] `target_bitrate_info` (1806)는 제 1 부호화(614) 결과 획득된 영상 데이터의 비트레이트를 나타내는 정보이다. AI 업스케일부(234)는 `target_bitrate_info` (1806)에 따라, 저해상도 영상의 화질에 적합한 업스케일 DNN 정보를 획득할 수 있다.
- [327] `res_info` (1808)는 제3 영상(145)과 같은 AI 업스케일된 고해상도 영상의 해상도에 관련된 해상도 정보를 나타낸다. `res_info` (1808)는 `pic_width_org_luma`(1810), `pic_height_org_luma`(1812)을 포함할 수 있다. `pic_width_org_luma`(1810), `pic_height_org_luma`(1812)는 각각 고해상도 영상의 너비와 높이를 나타내는 고해상도 영상 너비 정보와 고해상도 영상 높이 정보이다. AI 업스케일부(234)는 `pic_width_org_luma`(1810), `pic_height_org_luma`(1812)에 따라 결정된 고해상도 영상의 해상도와 제1 복호화부(232)에서 복원된 저해상도 영상의 해상도에 따라, AI 업스케일 비율을 결정할 수 있다.
- [328] 일 실시 예에서, `res_info` (1808)는 `pic_width_org_luma`(1810), `pic_height_org_luma`(1812)을 포함하지 않고, 대신 저해상도 영상과 고해상도 영상의 해상도 비율을 나타내는 해상도 비율 정보를 포함할 수 있다. 따라서, AI 업스케일부(234)는, 제1 복호화부(232)에서 복원된 저해상도 영상의 해상도와 해상도 비율 정보에 따른 해상도 비율에 따라, 고해상도 영상의 해상도를 결정할 수 있다. 상기 해상도 비율 정보는 수직 해상도 비율 정보와 수평 해상도 비율 정보를 포함할 수 있다.
- [329] AI 업스케일부(234)는 `res_info` (1808)에 따라 결정된 AI 업스케일 비율에 따라, 저해상도 영상의 AI 업스케일에 적합한 업스케일 DNN 정보를 획득할 수 있다. 또는 AI 업스케일부(234)는 `res_info` (1808)에 따라 결정된 저해상도 영상의 해상도와 고해상도 영상의 해상도에 따라, 저해상도 영상의 AI 업스케일에 적합한 업스케일 DNN 정보를 획득할 수 있다.
- [330] `ai_codec_DNN_info` (1814)는 저해상도 영상의 AI 업스케일에 사용되는 상호 약속된 업스케일 DNN 정보를 가리키는 정보이다. AI 업스케일부(234)는 `ai_codec_applied_channel_info` (1804), `target_bitrate_info` (1806), `res_info` (1808) 등에 따라, 미리 저장된 복수의 디폴트 DNN 설정 정보들 중에 하나를 업스케일 DNN 정보로 결정할 수 있다. 또한 AI 업스케일부(234)는 고해상도 영상의 기타

특성(영상의 장르, 최대 휘도, 색역 등)을 추가로 고려하여, 미리 저장된 복수의 디폴트 DNN 설정 정보 중에 하나를 업스케일 DNN 정보로 결정할 수 있다.

- [331] 만약, 특정 조건에 대응되는 디폴트 DNN 설정 정보가 2개 이상일 때, ai\_codec\_DNN\_info (1814)는 2개 이상의 디폴트 DNN 설정 정보 중 하나를 나타낼 수 있다. 그리고 ai\_codec\_DNN\_info (1814)가 나타내는 업스케일 DNN 정보를 이용하여, AI 업스케일부(234)는 저해상도 영상을 AI 업스케일할 수 있다.
- [332] 일 실시 예에 따르면, 특정 조건에 대응되는 디폴트 DNN 설정 정보가 2개 이상일 때, ai\_codec\_DNN\_info (1814)는 2개 이상의 디폴트 DNN 설정 정보 중 현재 비디오 파일에 적용될 수 있는 2개 이상의 업스케일 DNN 정보를 나타낼 수 있다. 그리고 프레임 그룹 또는 프레임 단위로, ai\_codec\_DNN\_info (1814)가 나타내는 2개 이상의 업스케일 DNN 정보 중 하나가 적응적으로 선택될 수 있다.
- [333] ai\_codec\_supplementary\_info (1816)는 AI 업스케일에 관한 부가 정보를 나타낸다. ai\_codec\_supplementary\_info (1816)는 비디오에 적용되는 업스케일 DNN 정보의 결정에 필요한 정보를 포함할 수 있다. ai\_codec\_supplementary\_info (1816)는 장르, HDR 최대 조도, HDR 색역, HDR PQ 정보, 레이트 콘트롤 타입 등에 관한 정보, 제1 부호화에 사용되는 codec에 관한 정보 등을 포함할 수 있다.
- [334] 도 18A의 ai\_codec\_applied\_channel\_info (1804), target\_bitrate\_info (1806), res\_info (1808), ai\_codec\_DNN\_info (1814), ai\_codec\_supplementary\_info (1816)는 실시 예에 따라 일부가 생략될 수 있다. 상기 엘리먼트들이 포함된 데이터 구조(1800)를 가지는 비디오 AI 데이터는 모든 프레임 그룹들에 적용될 수 있다.
- [335] 도 18B는 도 16에서 설명된 비디오 세그먼트 AI 데이터 또는 도 14, 도 15A 내지 15C, 도 17의 프레임 그룹 AI 데이터에 적용될 수 있는 데이터 구조(1820)의 일 실시 예를 나타낸다.
- [336] 데이터 구조(1820)는 도 18A의 데이터 구조(1800)와 유사한 구조를 가진다. 도 18B의 ai\_codec\_info (1822), ai\_codec\_applied\_channel\_info (1826), target\_bitrate\_info (1828), res\_info (1830), pic\_width\_org\_luma (1832), pic\_height\_org\_luma (1834), ai\_codec\_DNN\_info (1836), ai\_codec\_supplementary\_info (1838)은 도 18A의 ai\_codec\_applied\_channel\_info (1802), ai\_codec\_applied\_channel\_info (1804), target\_bitrate\_info (1806), res\_info (1808), pic\_width\_org\_luma (1810), pic\_height\_org\_luma (1812), ai\_codec\_DNN\_info (1814), ai\_codec\_supplementary\_info (1816)와 동일한 기능을 수행할 수 있다.
- [337] 다만 데이터 구조(1820)는 연속된 이전 데이터 단위와 현재 데이터 단위 간에 AI 보조 데이터가 동일한지 여부를 나타내는 AI 보조 데이터 종속 정보(dependent\_ai\_condition\_info (1824))를 추가적으로 포함할 수 있다. 데이터 구조(1820)는, ai\_codec\_info (1822)가 저해상도 영상에 AI 업스케일이 적용됨을 나타낼 경우, dependent\_ai\_condition\_info (1824)를 포함할 수 있다. 만약,

ai\_codec\_info (1822)가 저해상도 영상에 AI 업스케일이 적용되지 않음을 나타낼 경우, 비디오 dependent\_ai\_condition\_info (1824)는 데이터 구조(1820)에서 생략될 수 있다.

- [338] dependent\_ai\_condition\_info (1824)가 연속된 이전 데이터 단위와 현재 데이터 단위 간에 AI 보조 데이터가 동일하다고 나타낼 경우, ai\_codec\_applied\_channel\_info (1826), target\_bitrate\_info (1828), res\_info (1830), pic\_width\_org\_luma (1832), pic\_height\_org\_luma (1834), ai\_codec\_DNN\_info (1836), ai\_codec\_supplementary\_info (1838)은 데이터 구조(1820)에서 생략될 수 있다. 그리고 채널 정보, 목표 비트레이트 정보, 해상도 정보, DNN 정보, 및 부가 정보는 이전 데이터 단위의 ai\_codec\_applied\_channel\_info (1826), target\_bitrate\_info (1828), res\_info (1830), pic\_width\_org\_luma (1832), pic\_height\_org\_luma (1834), ai\_codec\_DNN\_info (1836), ai\_codec\_supplementary\_info (1838)로부터 결정된다. 따라서, 복수의 데이터 단위에 대하여 동일한 AI 보조 데이터가 적용되는 경우, dependent\_ai\_condition\_info (1824)에 따라 AI 데이터의 크기가 감소할 수 있다.
- [339] dependent\_ai\_condition\_info (1824)가 연속된 이전 데이터 단위와 현재 데이터 단위 간에 AI 보조 데이터가 동일하지 않다고 나타낼 경우, 데이터 구조는 ai\_codec\_applied\_channel\_info (1826), target\_bitrate\_info (1828), res\_info (1830), pic\_width\_org\_luma (1832), pic\_height\_org\_luma (1834), ai\_codec\_DNN\_info (1836), ai\_codec\_supplementary\_info (1838)를 포함한다. 그리고, 현재 데이터 단위의 업스케일 DNN 정보는 이전 데이터 단위의 업스케일 DNN 정보와 상관없이 독립적으로 획득될 수 있다.
- [340] 일 실시 예에서, 최초로 전송된 데이터 단위의 AI 보조 데이터 종속 정보는 생략될 수 있다. 따라서 두 번째로 전송된 비디오 세그먼트부터, AI 보조 데이터 종속 정보가 비디오 세그먼트의 헤더에 포함된다.
- [341] 일 실시 예에서, AI 보조 데이터 종속 정보는 특정 AI 보조 데이터의 종속 여부만을 나타낼 수 있다. 예를 들어, AI 보조 데이터 종속 정보는 현재 데이터 단위가 이전 데이터 단위의 해상도 정보를 승계하는지 여부를 나타낼 수 있다. AI 보조 데이터 종속 정보에 따라 승계되는 AI 데이터는 ai\_codec\_info(1100), ai\_codec\_applied\_channel\_info (1804), target\_bitrate\_info (1806), res\_info (1808), ai\_codec\_DNN\_info (1814), ai\_codec\_supplementary\_info (1816)중 적어도 하나를 포함할 수 있다.
- [342] 도 18B에서, 상기 데이터 단위는 비디오 세그먼트 또는 프레임 그룹일 수 있다. 상기 데이터 단위가 비디오 세그먼트일 때, dependent\_ai\_condition\_info (1824)는 연속된 이전 비디오 세그먼트와 현재 비디오 세그먼트 간의 AI 보조 데이터의 동일 여부를 나타낼 수 있다. 상기 데이터 단위가 프레임 그룹일 때, dependent\_ai\_condition\_info (1824)는 연속된 이전 프레임 그룹과 현재 프레임 그룹 간의 AI 보조 데이터의 동일 여부를 나타낼 수 있다.

- [343] 도 19는 도 18A의 데이터 구조(1800)가 구현된 선택스 테이블(1900)을 나타낸다.
- [344] 선택스 테이블 (1900)에 따르면, ai\_codec\_info, ai\_codec\_applied\_channel\_info, target\_bitrate, pic\_width\_org\_luma, pic\_height\_org\_luma, ai\_codec\_DNN\_info, ai\_codec\_supplementary\_info\_flag 와 같은 선택스 요소(syntax element)가 파싱된다.
- [345] ai\_codec\_info 는 도 18A의 ai\_codec\_info (1802)에 대응되는 AI 타겟 데이터이다. ai\_codec\_info 가 AI 업스케일이 허용됨을 나타낼 경우(if(ai\_codec\_info)), AI 보조 데이터에 대응되는 선택스 요소들이 파싱된다. ai\_codec\_info 가 AI 업스케일이 허용되지 않음을 나타낼 경우, AI 보조 데이터에 대응되는 선택스 요소들은 파싱되지 않는다.
- [346] AI 보조 데이터에 대응되는 선택스 요소는 ai\_codec\_applied\_channel\_info, target\_bitrate, pic\_width\_org\_luma, pic\_height\_org\_luma, ai\_codec\_DNN\_info, ai\_codec\_supplementary\_info\_flag를 포함한다.
- [347] ai\_codec\_applied\_channel\_info 는 도 18A의 ai\_codec\_applied\_channel\_info (1804)에 대응되는 채널 정보이다. target\_bitrate는 도 18A의 target\_bitrate\_info (1806)에 대응되는 목표 비트레이트 정보이다. pic\_width\_org\_luma, pic\_height\_org\_luma 는 각각 도 18A의 pic\_width\_org\_luma(1810), pic\_height\_org\_luma(1812)에 대응되는 고해상도 영상 너비 정보와 고해상도 영상 높이 정보이다. ai\_codec\_DNN\_info는 도 18A의 ai\_codec\_DNN\_info(1814)에 대응되는 DNN 정보이다.
- [348] ai\_codec\_supplementary\_info\_flag는 도 18A의 ai\_codec\_supplementary\_info (1816)가 선택스 테이블(1900)에 포함되어 있는지 여부를 나타내는 부가 정보 플래그이다. 만약 ai\_codec\_supplementary\_info\_flag가 AI 업스케일에 이용되는 보조 정보가 파싱되지 않는다고 나타내는 경우, 추가적인 보조 정보는 획득되지 않는다. 그러나 ai\_codec\_supplementary\_info\_flag가 AI 업스케일에 이용되는 보조 정보가 파싱된다고 나타내는 경우(if(ai\_codec\_supplementary\_info\_flag)), 추가적인 보조 정보는 획득된다.
- [349] 획득되는 추가적인 보조 정보는 genre\_info, hdr\_max\_luminance, hdr\_color\_gamut, hdr\_pq\_type, 및 rate\_control\_type, 및 codec\_type 등이 포함될 수 있다. genre\_info는 영상 데이터의 콘텐츠의 장르, hdr\_max\_luminance는 고해상도 영상에 적용되는 HDR (High Dynamic Range) 최대 조도, hdr\_color\_gamut는 고해상도 영상에 적용되는 HDR 색역, hdr\_pq\_type는 고해상도 영상에 적용되는 HDR PQ (Perceptual Quantizer) 정보, rate\_control\_type는 제 1 부호화 결과 획득된 영상 데이터에 적용된 레이트 컨트롤 타입, codec\_type은 제1 부호화에 사용되는 codec을 나타낸다. 실시 예에 따라, 보조 정보에 해당되는 상기 선택스 요소들 중 특정 선택스 요소가 파싱될 수 있다.
- [350] 도 19의 선택스 테이블 (1900)은 예시에 불과하며, 도 18A의 데이터

구조(1800)의 엘리먼트들 중 일부가 선택스 테이블 (1900)에 포함될 수 있다. 또한, 데이터 구조(1800)에 포함되지 않은 엘리먼트가 선택스 테이블 (1900)에 포함될 수 있다.

- [351] 실시 예에 따라, 도 19의 선택스 테이블 (1900)은, 도 18B의 데이터 구조(1820)와 마찬가지로, `dependent_ai_condition_info`와 같은 AI 보조 데이터 종속 정보를 포함할 수 있다. 따라서 AI 보조 데이터 종속 정보가 추가된 선택스 테이블 (1900)은 비디오 세그먼트 또는 프레임 그룹에 적용될 수 있다.
- [352] 도 20은 도 14 내지 도 17에서 설명된 (프레임 그룹 AI 데이터 또는) 프레임 AI 데이터에 적용될 수 있는 데이터 구조(2000)의 일 실시 예를 나타낸다.
- [353] 상기 데이터 구조(2000)에는 프레임 단위로 적응적으로 업스케일 DNN 정보를 결정하기 위한 엘리먼트들이 포함된다. 상기 엘리먼트들은 `ai_codec_frame_info` (2002), `dependent_ai_condition_frame_info` (2004), `ai_codec_frame_DNN_info` (2006), `ai_codec_enhancement_flag` (2008), `ai_codec_artifact_removal_flag` (2014) 등을 포함할 수 있다. 도 20에 나타난 엘리먼트의 배열 순서는 예시일 뿐, 통상의 기술자는 상기 엘리먼트의 배열 순서를 변경하여 실시할 수 있다.
- [354] `ai_codec_frame_info` (2002)는 현재 프레임에 AI 업스케일이 허용되는지 여부를 나타내는 프레임 AI 타겟 데이터이다. `ai_codec_frame_info` (2002)에 따라, 현재 프레임에 AI 업스케일이 허용될 경우, 데이터 구조(2000)는 현재 프레임의 AI 업스케일에 관련된 프레임 AI 보조 데이터를 포함한다. 그렇지 않을 경우, 현재 프레임에는 AI 업스케일이 적용되지 않으며, 데이터 구조(2000)는 프레임 AI 보조 데이터를 포함하지 않는다. 프레임 AI 보조 데이터는 프레임에 적용되는 AI 보조 데이터를 의미한다.
- [355] 일 실시 예에 따라, `ai_codec_frame_info` (2002)에 따라 현재 프레임에 AI 업스케일이 허용될 경우, 데이터 구조(2000)는 `dependent_ai_condition_frame_info` (2004)를 포함할 수 있다. `dependent_ai_condition_frame_info` (2004)는 연속된 이전 프레임과 현재 프레임 간에 프레임 AI 보조 데이터가 동일한지 여부를 나타내는 프레임 AI 보조 데이터 종속 정보이다. `dependent_ai_condition_frame_info` (2004)가 연속된 이전 프레임과 현재 프레임 간에 프레임 AI 보조 데이터가 동일하다고 나타낼 경우, 데이터 구조(2000)는 현재 프레임에 대한 추가적인 프레임 AI 보조 데이터를 포함하지 않고, 현재 프레임의 업스케일 DNN 정보는 이전 프레임의 업스케일 DNN 정보와 동일하게 결정된다. `dependent_ai_condition_frame_info` (2004)가 연속된 이전 프레임과 현재 프레임 간에 프레임 AI 보조 데이터가 동일하지 않다고 나타내는 경우, 데이터 구조(2000)는 현재 프레임에 대한 추가적인 프레임 AI 보조 데이터를 포함하고, 현재 프레임의 업스케일 DNN 정보는 이전 프레임의 업스케일 DNN 정보와는 독립적으로 획득된다. 상기 추가적인 프레임 AI 보조 데이터는 `ai_codec_frame_DNN_info` (2006), `ai_codec_enhancement_flag` (2008), `ai_codec_artifact_removal_flag` (2014)을 포함할 수 있다.

- [356] ai\_codec\_frame\_DNN\_info (2006)는 현재 프레임의 상위 데이터 유닛에 대한 복수의 업스케일 DNN 정보 중 현재 프레임의 업스케일 DNN 정보를 나타내는 프레임 DNN 정보이다. 도 18A의 ai\_codec\_DNN\_info (1814)가 비디오에 대한 2개 이상의 업스케일 DNN 정보를 나타낼 경우, ai\_codec\_frame\_DNN\_info (2006)는 상기 2개 이상의 업스케일 DNN 정보로부터 현재 프레임의 업스케일 DNN 정보를 결정할 수 있다. 또는 도 18B의 ai\_codec\_DNN\_info (1834)가 비디오 세그먼트 또는 프레임 그룹에 대한 2개 이상의 업스케일 DNN 정보를 나타낼 경우, ai\_codec\_frame\_DNN\_info (2006)는 상기 2개 이상의 업스케일 DNN 정보로부터 현재 프레임의 업스케일 DNN 정보를 결정할 수 있다. 만약, 도 18A의 ai\_codec\_DNN\_info (1814) 또는 도 18B의 ai\_codec\_DNN\_info (1834)가 하나의 업스케일 DNN 정보만 나타낼 경우, ai\_codec\_frame\_DNN\_info (2006)는 생략될 수 있다.
- [357] ai\_codec\_enhancement\_flag (2008)는 AI 업스케일 정확도의 향상 프로세스가 활성화되는지 여부를 나타내는 AI 향상 정보이다. ai\_codec\_enhancement\_flag (2008)가 AI 업스케일 정확도의 향상 프로세스가 활성화되었음을 나타내는 경우, AI 업스케일된 고해상도 영상의 샘플들 중 일부가 부호화 파라미터 정보에 따라 조정된다. ai\_codec\_enhancement\_flag (2008)가 AI 업스케일 정확도의 향상 프로세스가 활성화되지 않음을 나타내는 경우, 상기 향상 프로세스는 생략된다.
- [358] 부호화 파라미터는 부호화단에서 원본 영상(105) 또는 제1 영상(115)를 부호화할 때 발생된다. 부호화 파라미터는 데이터 단위(최대 부호화 단위, 부호화 단위, 예측 단위, 변환 단위 또는 픽셀 단위)의 예측, 변환 및 인루프 필터링 프로세스에 따라 생성될 수 있다. 부호화 파라미터는 움직임 벡터, 예측 움직임 벡터, 인트라 모드, 잔차 신호 관련 정보, SAO 파라미터 등의 정보를 포함한다.
- [359] 부호화 파라미터 정보는 상기 부호화 파라미터에 따른 향상 프로세스에 필요한 정보이다. 부호화 파라미터 정보는 향상 프로세스를 위하여 참조되는 부호화 파라미터의 타입을 나타내는 부호화 파라미터 타입 정보와 현재 프레임에서 향상 프로세스의 적용 영역을 나타내는 부호화 파라미터 맵 정보를 포함할 수 있다.
- [360] ai\_codec\_enhancement\_flag (2008)가 AI 업스케일 정확도의 향상 프로세스가 활성화되었음을 나타내는 경우, 데이터 구조(2000)는 부호화 파라미터 타입 정보인 encod\_param\_type (2010)과 부호화 파라미터 맵 정보인 encod\_param\_map (2012)를 포함할 수 있다. 예를 들어, encod\_param\_type (2010)이 움직임 벡터를 나타낼 수 있다. 그리고 encod\_param\_map (2012)은 AI 업스케일된 고해상도 영상에서 움직임 벡터에 따른 향상 프로세스의 적용 영역을 나타낼 수 있다. 그리고 움직임 벡터에 따라 상기 적용 영역의 픽셀이 수정될 수 있다.
- [361] 실시 예에 따라, encod\_param\_type (2010)은 2개 이상의 부호화 파라미터들을 나타낼 수 있다. 그리고 encod\_param\_map (2012)은 각 부호화 파라미터에 대한

향상 프로세스의 적용 영역을 나타낼 수 있다.

- [362] 실시 예에 따라, 데이터 구조(2000)에서 `encod_param_type` (2010) 및 `encod_param_map` (2012)은 생략될 수 있다. 따라서, 상기 향상 프로세스에 참조되는 하나 이상의 부호화 파라미터와 각 부호화 파라미터의 적용 영역이 미리 결정될 수 있다.
- [363] `ai_codec_artifact_removal_flag` (2014)는 AI 업스케일된 고해상도 영상의 아티팩트(artifact) 제거가 수행되는지 여부를 나타내는 아티팩트 제거 정보이다. 고해상도 영상의 아티팩트 제거를 위하여, 상기 아티팩트 제거 정보에 따라, AI 업스케일 전 저해상도 영상의 픽셀들을 보정하거나, 제2 DNN에 따른 AI 업스케일 후 고해상도 영상의 픽셀들을 보정함으로써, 아티팩트가 제거될 수 있다.
- [364] 아티팩트 제거 정보는 아티팩트의 타입을 나타내는 아티팩트 타입 정보와 아티팩트가 위치한 아티팩트 영역을 나타내는 아티팩트 맵 정보를 포함할 수 있다. 또한 아티팩트 제거 정보는 영상의 아티팩트 타입의 수를 나타내는 아티팩트 타입 개수 정보를 포함할 수 있다. 따라서, 데이터 구조(2000)는 아티팩트 타입 개수 정보에 따른 개수만큼 아티팩트 타입 정보와 아티팩트 맵 정보를 포함할 수 있다.
- [365] 아티팩트의 타입은 컨투어(contour) 아티팩트, 링잉(ringing) 아티팩트 및 앨리어싱(aliasing) 아티팩트 등을 포함한다. 그리고 아티팩트의 타입마다 한 개 이상의 아티팩트 영역이 결정된다. 예를 들어, 컨투어 아티팩트에 대하여 한 개 이상의 아티팩트 영역이 결정되고, 링잉 아티팩트에 대하여 한 개 이상의 아티팩트 영역이 결정될 수 있다.
- [366] `ai_codec_artifact_removal_flag` (2014)가 AI 업스케일된 고해상도 영상의 아티팩트 제거가 수행됨을 나타내는 경우, 데이터 구조(2000)는 아티팩트 타입 개수 정보인 `num_artifact_type` (2016) 등을 포함할 수 있다. 만약 `ai_codec_artifact_removal_flag` (2014)가 AI 업스케일된 고해상도 영상의 아티팩트 제거가 수행되지 않음을 나타내는 경우, 데이터 구조(2000)에는 `num_artifact_type` (2016) 등이 생략된다.
- [367] 데이터 구조(2000)는 `num_artifact_type` (2016)가 나타내는 개수만큼 아티팩트 타입 정보인 `artifact_type` (2018)를 포함할 수 있다. 또한 데이터 구조(2000)는 각 `artifact_type` (2018)에 대한 아티팩트 맵 정보를 포함할 수 있다. 데이터 구조(2000)의 아티팩트 맵 정보는 아티팩트 영역의 개수를 나타내는 `num_artifact_map` (2020)를 포함할 수 있다. 그리고 데이터 구조(2000)는 각 아티팩트 영역의 위치와 크기를 나타내는 `map_x_pos` (2022), `map_y_pos` (2024), `map_width` (2026), `map_height` (2028)을 포함할 수 있다.
- [368] 실시 예에 따라, 데이터 구조(2000)에서, 도 20의 `dependent_ai_condition_frame_info` (2004), `ai_codec_frame_DNN_info` (2006), `ai_codec_enhancement_flag` (2008), `ai_codec_artifact_removal_flag` (2014)의 일부가

생략될 수 있다.

- [369] 실시 예에 따라, 프레임 대신 프레임 그룹에 도 20의 데이터 구조(2000)가 적용될 수 있다. 프레임 그룹에 데이터 구조(2000)가 적용될 경우, 프레임 그룹에 포함된 모든 프레임에 동일한 업스케일 DNN 정보, AI 향상 정보 및 아티팩트 제거 정보가 적용될 수 있다. 또한 프레임 그룹에 데이터 구조(2000)가 적용될 경우, 데이터 구조(2000)는 ai\_codec\_frame\_info (2002), dependent\_ai\_condition\_frame\_info (2004), 및 ai\_codec\_frame\_DNN\_info (2006) 대신 프레임 그룹에 적용되는 ai\_codec\_frame\_group\_info, dependent\_ai\_condition\_frame\_group\_info, 및 ai\_codec\_frame\_group\_DNN\_info를 포함할 수 있다. 그리고 ai\_codec\_enhancement\_flag (2008), encod\_param\_type (2010), encod\_param\_map (2012), ai\_codec\_artifact\_removal\_flag (2014), num\_artifact\_type (2016), artifact\_type (2018), num\_artifact\_map (2020), map\_x\_pos (2022), map\_y\_pos (2024), map\_width (2026), 및 map\_height (2028) 은 프레임 그룹의 모든 프레임에 적용될 수 있다.
- [370] 일 실시 예에 따르면, 도 18A의 데이터 구조(1800)는 비디오 AI 데이터에 적용될 수 있다. 그리고 도 20의 데이터 구조(2000)은 프레임 AI 데이터에 적용될 수 있다. 데이터 구조(1800)에 따른 비디오 AI 데이터는 모든 프레임에 공통적으로 적용되는 AI 데이터를 포함하고, 데이터 구조(2000)에 따른 프레임 AI 데이터는 프레임에 적응적으로 적용되는 AI 데이터를 포함한다.
- [371] 일 실시 예에 따르면, 도 18A의 데이터 구조(1800)는 비디오 AI 데이터에 적용될 수 있다. 그리고 도 20의 데이터 구조(2000)은 프레임 그룹 AI 데이터에 적용될 수 있다. 데이터 구조(1800)에 따른 비디오 AI 데이터는 모든 프레임에 공통적으로 적용되는 AI 데이터를 포함하고, 데이터 구조(2000)에 따른 프레임 그룹 AI 데이터는 프레임 그룹에 적응적으로 적용되는 AI 데이터를 포함한다. 그리고 프레임 그룹에 포함된 모든 프레임에 동일한 AI 데이터가 적용된다.
- [372] 일 실시 예에 따르면, 도 18B의 데이터 구조(1820)는 비디오 세그먼트 AI 데이터 또는 프레임 그룹 AI 데이터에 적용될 수 있다. 그리고 도 20의 데이터 구조(2000)은 프레임 AI 데이터에 적용될 수 있다. 데이터 구조(1820)에 따른 비디오 세그먼트 AI 데이터 또는 프레임 그룹 AI 데이터는 비디오 세그먼트 또는 프레임 그룹의 모든 프레임에 공통적으로 적용되는 AI 데이터를 포함하고, 데이터 구조(2000)에 따른 프레임 AI 데이터는 프레임에 적응적으로 적용되는 AI 데이터를 포함한다.
- [373] 일 실시 예에 따르면, 도 18B의 데이터 구조(1820)는 비디오 세그먼트 AI 데이터에 적용될 수 있다. 그리고 도 20의 데이터 구조(2000)은 프레임 그룹 AI 데이터에 적용될 수 있다. 데이터 구조(1800)에 따른 비디오 세그먼트 AI 데이터는 비디오 세그먼트의 모든 프레임에 공통적으로 적용되는 AI 데이터를 포함하고, 데이터 구조(2000)에 따른 프레임 그룹 AI 데이터는 프레임 그룹에 적응적으로 적용되는 AI 데이터를 포함한다. 그리고 프레임 그룹에 포함된 모든

프레임에 동일한 AI 데이터가 적용된다.

[374] 도 21은 도 20의 데이터 구조(2000)가 구현된 선택스 테이블(2100)을 나타낸다.

[375] 선택스 테이블 (2100)에 따르면, ai\_codec\_frame\_info, dependent\_ai\_condition\_frame\_info, ai\_codec\_frame\_DNN\_info, ai\_codec\_enhancement\_flag, 및 ai\_codec\_artifact\_removal\_flag 와 같은 선택스 요소가 파싱된다.

[376] ai\_codec\_frame\_info 는 도 20의 ai\_codec\_frame\_info (2002)에 대응되는 프레임 AI 타겟 데이터이다. ai\_codec\_frame\_info가 AI 업스케일이 허용됨을 나타낼 경우(if(ai\_codec\_frame\_info)), 프레임 AI 보조 데이터에 대응되는 선택스 요소들이 파싱된다. ai\_codec\_frame\_info가 AI 업스케일이 허용되지 않음을 나타낼 경우, 프레임 AI 보조 데이터에 대응되는 선택스 요소들은 파싱되지 않는다.

[377] 프레임 AI 보조 데이터에 대응되는 선택스 요소는 dependent\_ai\_condition\_frame\_info, ai\_codec\_frame\_DNN\_info, ai\_codec\_enhancement\_flag, ai\_codec\_artifact\_removal\_flag 를 포함할 수 있다. 도 21의 dependent\_ai\_condition\_frame\_info, ai\_codec\_frame\_DNN\_info, ai\_codec\_enhancement\_flag, ai\_codec\_artifact\_removal\_flag는 도 20의 dependent\_ai\_condition\_frame\_info (2004), ai\_codec\_frame\_DNN\_info (2006), ai\_codec\_enhancement\_flag (2008), ai\_codec\_artifact\_removal\_flag (2014)에 대응된다.

[378] ai\_codec\_frame\_info가 AI 업스케일이 허용됨을 나타낼 경우, dependent\_ai\_condition\_frame\_info가 획득된다. dependent\_ai\_condition\_frame\_info가 연속된 이전 프레임과 현재 프레임 간에 프레임 AI 보조 데이터가 동일하다고 나타낼 경우, 현재 프레임의 업스케일 DNN 정보는 이전 프레임의 업스케일 DNN 정보와 동일하게 결정된다. 만약 dependent\_ai\_condition\_frame\_info가 연속된 이전 프레임과 현재 프레임 간에 프레임 AI 보조 데이터가 동일하지 않다고 나타낼 경우, ai\_codec\_frame\_DNN\_info, ai\_codec\_enhancement\_flag, ai\_codec\_artifact\_removal\_flag에 따라, 현재 프레임의 업스케일 DNN 정보, AI 향상 정보, 아티팩트 제거 정보가 결정된다.

[379] ai\_codec\_frame\_DNN\_info는 현재 프레임의 상위 데이터 유닛에 대한 복수의 업스케일 DNN 정보중 현재 프레임의 업스케일 DNN 정보를 나타낸다.

[380] ai\_codec\_enhancement\_flag는 AI 업스케일 정확도의 향상 프로세스가 활성화되는지 여부를 나타낸다. ai\_codec\_enhancement\_flag가 향상 프로세스가 활성화됨을 나타낼 경우(if(ai\_codec\_enhancement\_flag)), 부호화 파라미터 타입을 나타내는 encod\_param\_type과 부호화 파라미터 영역을 나타내는 encod\_param\_map이 획득된다.

[381] ai\_codec\_artifact\_removal\_flag는 AI 업스케일된 고해상도 영상의 아티팩트

제거가 수행되는지 여부를 나타낸다. ai\_codec\_artifact\_removal\_flag가 고해상도 영상의 아티팩트 제거가 수행됨을 나타낼 때(if(ai\_codec\_artifact\_removal\_flag)), 아티팩트 타입의 개수를 나타내는 num\_artifact\_type이 획득된다.

num\_artifact\_type이 나타내는 개수만큼(for(i=0; i<num\_artifact\_type; i++)), 아티팩트의 타입을 나타내는 artifact\_type과 아티팩트 영역의 개수를 나타내는 num\_artifact\_map이 획득된다. 그리고 num\_artifact\_map의 개수만큼(for(i=0; i<num\_artifact\_map; i++)), 아티팩트 영역의 위치 및 크기를 나타내는 map\_x\_pos, map\_y\_pos, map\_width, map\_height가 획득된다.

- [382] 도 21의 선택스 테이블 (2100)은 예시에 불과하며, 도 20의 데이터 구조(2000)의 엘리먼트들 중 일부가 선택스 테이블 (2100)에 포함될 수 있다. 또한, 데이터 구조(2000)에 포함되지 않은 엘리먼트가 선택스 테이블 (2100)에 포함될 수 있다.
- [383] 도 22는 AI 복호화기에 따른 영상 복호화 방법의 일 실시예의 순서도이다.
- [384] 단계 2210에서, 영상 데이터와 상기 영상 데이터의 AI 업스케일에 관한 AI 데이터가 포함된 AI 부호화 데이터를 포함하는 비디오 파일이 수신된다. 단계 2210에 따른 AI 부호화 데이터 수신은 통신부(212)에 의하여 수행될 수 있다.
- [385] 단계 2220에서, 상기 비디오 파일의 메타데이터 박스로부터 상기 AI 부호화 데이터의 AI 데이터를 획득하고, 상기 비디오 파일의 미디어데이터 박스로부터, 상기 AI 부호화 데이터의 영상 데이터가 획득된다. 영상 데이터와 AI 데이터가 분리된 경우, 단계 2220에 따른 영상 데이터와 AI 데이터의 획득은 파싱부(212)에 의하여 수행될 수 있다. 또는 AI 데이터가 영상 데이터에 포함된 경우, 제1 복호화부(232)에 의하여, 영상 데이터로부터 AI 데이터가 획득될 수 있다.
- [386] 일 실시 예에 따르면, AI 부호화 데이터는 영상 데이터와 AI 데이터의 동기화에 관한 동기화 데이터를 포함할 수 있다. 동기화 데이터는 영상 데이터의 복호화 순서 또는 재생 순서에 따라, 영상 데이터와 AI 데이터의 관계를 나타낼 수 있다. 동기화 데이터는 비디오 헤더와 비디오 AI 데이터의 동기화, 프레임 그룹 헤더와 프레임 그룹 AI 데이터의 동기화, 및 프레임 헤더와 프레임 AI 데이터의 동기화에 대한 정보를 포함할 수 있다.
- [387] 단계 2230에서, 영상 데이터를 복호화하여, 현재 프레임의 저해상도 영상이 복원된다. 단계 2230에 따른 저해상도 영상의 복원은 제1 복호화부(232)에 의하여 수행될 수 있다.
- [388] 단계 2240에서, AI 데이터로부터, 현재 프레임의 업스케일 DNN 정보가 획득된다. 단계 2240에 따른 현재 프레임의 업스케일 DNN 정보의 획득은 AI 업스케일부(234)에 의하여 수행될 수 있다.
- [389] 일 실시 예에 따르면, AI 데이터는 도 12의 AI 데이터의 계층 구조에 따라 비디오 AI 데이터, 프레임 그룹 AI 데이터 및 프레임 AI 데이터로 구성될 수 있다. 추가적으로 AI 데이터는 도 16의 비디오 세그먼트 AI 데이터를 포함할 수 있다. 상기 계층 구조에 따른 AI 데이터는, 대응되는 계층에 적용되는 업스케일 DNN

정보를 나타낼 수 있다.

- [390] 일 실시 예에 따르면, AI 데이터는 비디오 AI 데이터를 포함할 수 있다. AI 데이터에 비디오 AI 데이터만 포함된 경우, 비디오 AI 데이터에 따라 복수의 디폴트 DNN 설정 정보로부터 영상 데이터의 모든 프레임들에 적용되는 업스케일 DNN 정보가 획득될 수 있다.
- [391] 일 실시 예에 따르면, AI 데이터는 프레임 그룹 AI 데이터를 포함할 수 있다. AI 데이터에 프레임 그룹 AI 데이터만 포함된 경우, 프레임 그룹 AI 데이터에 따라 복수의 디폴트 DNN 설정 정보로부터 프레임 그룹의 모든 프레임에 적용되는 업스케일 DNN 정보가 프레임 그룹에 적응적으로 획득될 수 있다.
- [392] 일 실시 예에 따르면, AI 데이터는 프레임 AI 데이터를 포함할 수 있다. AI 데이터에 프레임 AI 데이터만 포함된 경우, 프레임 AI 데이터에 따라 복수의 디폴트 DNN 설정 정보로부터 프레임에 적용되는 업스케일 DNN 정보가 프레임에 적응적으로 될 수 있다.
- [393] 일 실시 예에 따르면, AI 데이터는 비디오 AI 데이터와 함께 프레임 그룹 AI 데이터를 포함할 수 있다. AI 데이터에 비디오 AI 데이터 및 프레임 그룹 AI 데이터만 포함된 경우, 비디오 AI 데이터에 따라 복수의 디폴트 DNN 설정 정보로부터 하나 이상의 업스케일 DNN 정보가 획득될 수 있다. 그리고 프레임 그룹 AI 데이터에 따라, 상기 하나 이상의 업스케일 DNN 정보로부터, 프레임 그룹의 프레임들에 적용되는 업스케일 DNN 정보가 선택될 수 있다.
- [394] 일 실시 예에 따르면, AI 데이터는 비디오 AI 데이터와 함께 프레임 AI 데이터를 포함할 수 있다. AI 데이터에 비디오 AI 데이터 및 프레임 AI 데이터만 포함된 경우, 비디오 AI 데이터에 따라 복수의 디폴트 DNN 설정 정보로부터 하나 이상의 업스케일 DNN 정보가 획득될 수 있다. 그리고 프레임 AI 데이터에 따라, 상기 하나 이상의 업스케일 DNN 정보로부터, 프레임에 적용되는 업스케일 DNN 정보가 선택될 수 있다.
- [395] 일 실시 예에 따르면, AI 데이터는 프레임 그룹 AI 데이터와 함께 프레임 AI 데이터를 포함할 수 있다. AI 데이터에 프레임 그룹 AI 데이터 및 프레임 AI 데이터만 포함된 경우, 프레임 그룹 AI 데이터에 따라 복수의 디폴트 DNN 설정 정보로부터 하나 이상의 업스케일 DNN 정보가 획득될 수 있다. 그리고 프레임 AI 데이터에 따라, 상기 하나 이상의 업스케일 DNN 정보로부터, 프레임에 적용되는 업스케일 DNN 정보가 선택될 수 있다.
- [396] 일 실시 예에 따르면, AI 데이터는 비디오 세그먼트 AI 데이터를 포함할 수 있다. AI 데이터에 비디오 세그먼트 AI 데이터만 포함된 경우, 비디오 세그먼트 AI 데이터에 따라, 복수의 디폴트 DNN 설정 정보로부터 비디오 세그먼트의 모든 프레임들에 적용되는 업스케일 DNN 정보가 획득될 수 있다.
- [397] 일 실시 예에 따르면, AI 데이터는 비디오 세그먼트 AI 데이터와 함께 비디오 AI 데이터를 포함할 수 있다. AI 데이터에 비디오 AI 데이터 및 비디오 세그먼트 AI 데이터만 포함된 경우, 비디오 AI 데이터에 따라 복수의 디폴트 DNN 설정

정보로부터 하나 이상의 업스케일 DNN 정보가 획득될 수 있다. 그리고 비디오 세그먼트 AI 데이터에 따라, 상기 하나 이상의 업스케일 DNN 정보로부터, 비디오 세그먼트의 모든 프레임들에 적용되는 업스케일 DNN 정보가 선택될 수 있다.

[398] 일 실시 예에 따르면, AI 데이터는 비디오 세그먼트 AI 데이터와 함께 프레임 그룹 AI 데이터 또는 프레임 AI 데이터를 포함할 수 있다. AI 데이터에 비디오 세그먼트 AI 데이터와 함께 프레임 그룹 AI 데이터 또는 프레임 AI 데이터가 포함된 경우, 비디오 세그먼트 AI 데이터에 따라 복수의 디폴트 DNN 설정 정보로부터 하나 이상의 업스케일 DNN 정보가 획득될 수 있다. 그리고 프레임 그룹 AI 데이터 또는 프레임 AI 데이터에 따라, 상기 하나 이상의 업스케일 DNN 정보로부터, 프레임 그룹 또는 프레임에 적용되는 업스케일 DNN 정보가 선택될 수 있다.

[399] 일 실시 예에 따르면, AI 데이터는 비디오 AI 데이터, 비디오 세그먼트 AI 데이터, 프레임 그룹 AI 데이터 및 프레임 AI 데이터를 모두 포함할 수 있다. AI 데이터가 비디오 AI 데이터, 비디오 세그먼트 AI 데이터, 프레임 그룹 AI 데이터 및 프레임 AI 데이터를 모두 포함할 때, 비디오 AI 데이터에 따라, 복수의 디폴트 DNN 설정 정보로부터, 비디오에 적용될 수 있는 하나 이상의 업스케일 DNN 정보가 한정될 수 있다. 그리고 비디오 세그먼트 AI 데이터에 따라, 상기 비디오에 적용될 수 있는 하나 이상의 업스케일 DNN 정보로부터, 비디오 세그먼트에 적용될 수 있는 하나 이상의 업스케일 DNN 정보가 선택될 수 있다. 그리고 프레임 그룹 AI 데이터에 따라, 상기 비디오 세그먼트에 적용될 수 있는 하나 이상의 업스케일 DNN 정보로부터, 프레임 그룹에 적용될 수 있는 하나 이상의 업스케일 DNN 정보가 선택될 수 있다. 그리고 프레임 AI 데이터에 따라, 상기 프레임 그룹에 적용될 수 있는 하나 이상의 업스케일 DNN 정보로부터, 프레임에 적용되는 업스케일 DNN 정보가 선택될 수 있다. 상기 실시 예에서, 비디오 AI 데이터, 비디오 세그먼트 AI 데이터, 프레임 그룹 AI 데이터 및 프레임 AI 데이터 중 일부가 상기 설명된 계층적 AI 데이터 구조에서 제외될 수 있다.

[400] 일 실시 예에 따르면, AI 데이터는, AI 업스케일이 적용되는지 여부를 나타내는 AI 타겟 데이터, 및 AI 업스케일이 적용될 경우, AI 업스케일에 사용되는 업스케일 DNN 정보에 관한 AI 보조 데이터를 포함할 수 있다.

[401] 비디오 AI 데이터는 영상 데이터에 AI 업스케일이 적용되는지 여부를 나타내는 비디오 AI 타겟 데이터 및 영상 데이터에 포함된 프레임들의 AI 업스케일에 적용될 수 있는 하나 이상의 업스케일 DNN 정보에 관한 비디오 AI 보조 데이터를 포함할 수 있다. 비디오 AI 타겟 데이터가 영상 데이터에 포함된 프레임들에 AI 업스케일이 적용됨을 나타낼 때, 비디오 AI 보조 데이터에 따라 하나 이상의 업스케일 DNN 정보가 획득된다.

[402] 비디오 세그먼트 AI 데이터는 비디오 세그먼트에 AI 업스케일이 적용되는지 여부를 나타내는 비디오 세그먼트 AI 타겟 데이터 및 비디오 세그먼트에 포함된

프레임들의 AI 업스케일에 적용될 수 있는 하나 이상의 업스케일 DNN 정보에 관한 비디오 세그먼트 AI 보조 데이터를 포함할 수 있다. 비디오 세그먼트 AI 타겟 데이터가 비디오 세그먼트에 포함된 프레임들에 AI 업스케일이 적용됨을 나타낼 때, 비디오 세그먼트 AI 보조 데이터에 따라 비디오 세그먼트의 하나 이상의 업스케일 DNN 정보가 획득된다.

- [403] 프레임 그룹 AI 데이터는 프레임 그룹에 AI 업스케일이 적용되는지 여부를 나타내는 프레임 그룹 AI 타겟 데이터 및 프레임 그룹에 포함된 프레임들의 AI 업스케일에 적용될 수 있는 하나 이상의 업스케일 DNN 정보에 관한 프레임 그룹 AI 보조 데이터를 포함할 수 있다. 프레임 그룹 AI 타겟 데이터가 프레임 그룹에 포함된 프레임들에 AI 업스케일이 적용됨을 나타낼 때, 프레임 그룹 AI 보조 데이터에 따라 프레임 그룹의 하나 이상의 업스케일 DNN 정보가 획득된다.
- [404] 프레임 AI 데이터는 프레임에 AI 업스케일이 적용되는지 여부를 나타내는 프레임 AI 타겟 데이터 및 프레임의 AI 업스케일에 사용되는 업스케일 DNN 정보에 관한 프레임 AI 보조 데이터를 포함할 수 있다. 프레임 AI 타겟 현재 프레임에 AI 업스케일이 적용됨을 나타낼 때, 프레임 AI 보조 데이터에 따라 현재 프레임의 업스케일 DNN 정보가 획득된다.
- [405] 일 실시 예에 따르면, 비디오 세그먼트 AI 데이터는 연속된 이전 비디오 세그먼트와 현재 비디오 세그먼트 간에 비디오 세그먼트 AI 보조 데이터가 동일한지 여부를 나타내는 비디오 세그먼트 AI 보조 데이터 종속 정보를 포함할 수 있다. 비디오 세그먼트 AI 보조 데이터 종속 정보가 연속된 이전 비디오 세그먼트와 현재 비디오 세그먼트 간에 비디오 세그먼트 AI 보조 데이터가 동일함을 나타낼 때, 현재 비디오 세그먼트의 비디오 세그먼트 AI 보조 데이터는 이전 비디오 세그먼트의 비디오 세그먼트 AI 보조 데이터와 동일하게 결정된다.
- [406] 또한 프레임 그룹 AI 데이터는 연속된 이전 프레임 그룹과 현재 프레임 그룹 간에 프레임 그룹 AI 보조 데이터가 동일한지 여부를 나타내는 프레임 그룹 AI 보조 데이터 종속 정보를 포함할 수 있다. 그리고 프레임 그룹 AI 보조 데이터 종속 정보가 연속된 이전 프레임 그룹과 현재 프레임 그룹 간에 프레임 그룹 AI 보조 데이터가 동일함을 나타낼 때, 현재 프레임 그룹의 프레임 그룹 AI 보조 데이터는 이전 프레임 그룹의 프레임 그룹 AI 보조 데이터와 동일하게 결정된다.
- [407] 또한 프레임 AI 데이터는 연속된 이전 프레임과 현재 프레임 간에 프레임 AI 보조 데이터가 동일한지 여부를 나타내는 프레임 AI 보조 데이터 종속 정보를 포함할 수 있다. 그리고 프레임 AI 보조 데이터 종속 정보가 연속된 이전 프레임과 현재 프레임 간에 프레임 AI 보조 데이터가 동일함을 나타낼 때, 현재 프레임의 프레임 AI 보조 데이터는 이전 프레임의 프레임 AI 보조 데이터와 동일하게 결정된다.
- [408] 일 실시 예에 따르면, 동기화 데이터에 따라, 영상 데이터 전체에 적용되는 비디오 AI 데이터, 비디오 세그먼트에 대응되는 비디오 세그먼트 AI 데이터, 프레임 그룹에 적용되는 프레임 그룹 AI 데이터, 및 프레임에 적용되는 프레임

AI 데이터가 결정될 수 있다.

- [409] 일 실시 예에 따르면, AI 데이터는 AI 업스케일이 적용되는 색 채널을 나타내는 채널 정보를 포함할 수 있다. 그리고 채널 정보가 나타내는 색 채널에 대하여 업스케일 DNN 정보가 획득될 수 있다.
- [410] 일 실시 예에 따르면, AI 데이터는 영상 데이터에 따른 저해상도 영상의 비트레이트를 나타내는 목표 비트레이트 정보 및 AI 업스케일된 고해상도 영상의 해상도에 관련된 해상도 정보 중 적어도 하나를 포함할 수 있다. 그리고 목표 비트레이트 정보 및 해상도 정보 중 적어도 하나에 따라, 비디오, 비디오 세그먼트, 또는 프레임 그룹에 대한 2개 이상의 업스케일 DNN 정보가 결정될 수 있다. 또한 프레임 AI 데이터에 따라, 상기 2개 이상의 업스케일 DNN 정보로부터 현재 프레임의 업스케일 DNN 정보가 결정될 수 있다.
- [411] 단계 2250에서, 현재 프레임의 업스케일 DNN 정보에 따라, 저해상도 영상을 AI 업스케일함으로써, 저해상도 영상에 대응되는 고해상도 영상이 생성된다. 단계 2260에 따른 고해상도 영상 생성은 AI 업스케일부(234)에 의하여 수행될 수 있다.
- [412] 일 실시 예에 따르면, 생성된 고해상도 영상의 정확도는 부호화 파라미터에 따른 AI 업스케일 향상 프로세스에 따라 향상될 수 있다. 또는 생성된 고해상도 영상의 아티팩트는 아티팩트 제거 프로세스에 따라 제거될 수 있다.
- [413] 도 12 내지 도 21에서 설명된 AI 업스케일에 대한 기능이 도 22의 비디오 복호화 방법에 적용될 수 있다.
- [414] 도 23은 AI 부호화기에 따른 영상 부호화 방법의 일 실시예의 순서도이다.
- [415] 단계 2310에서, 현재 프레임의 고해상도 영상을 저해상도 영상으로 AI 다운스케일하기 위한, 다운스케일 DNN 정보가 결정된다. 단계 2310에 따른 AI 데이터의 결정은 AI 다운스케일부(612)에 의하여 수행될 수 있다.
- [416] 목표 비트레이트 정보 및 해상도 정보 중 적어도 하나에 따라, 비디오, 비디오 세그먼트, 또는 프레임 그룹에 대하여 2개 이상의 다운스케일 DNN 정보가 결정될 수 있다. 그리고 프레임 AI 데이터에 따라, 상기 2개 이상의 다운스케일 DNN 정보로부터 현재 프레임의 다운스케일 DNN 정보가 결정될 수 있다.
- [417] 단계 2320에서, 다운스케일 DNN 정보를 이용하여, 현재 프레임의 고해상도 영상을 AI 다운스케일함으로써, 현재 프레임의 저해상도 영상이 생성된다. 단계 2330에 따른 저해상도 영상 생성은 AI 다운스케일부(612)에 의하여 수행될 수 있다.
- [418] 단계 2330에서, 현재 프레임의 저해상도 영상의 AI 업스케일에 사용되는, AI 데이터가 생성된다. 단계 2330에 따른 저해상도 영상 생성은 AI 다운스케일부(612)에 의하여 수행될 수 있다. AI 업스케일에 사용되는 AI 데이터는 단계 2310에서 생성된 다운스케일 DNN 정보를 참조하여 생성될 수 있다.
- [419] 일 실시 예에 따르면, AI 데이터는 도 12의 AI 데이터의 계층 구조에 따라

비디오 AI 데이터, 프레임 그룹 AI 데이터 및 프레임 AI 데이터로 구성될 수 있다. 추가적으로 AI 데이터는 도 16의 비디오 세그먼트 AI 데이터를 포함할 수 있다. 상기 계층 구조에 따른 AI 데이터는, 대응되는 계층에 적용되는 업스케일 DNN 정보를 나타낼 수 있다.

[420] 일 실시 예에 따르면, AI 데이터는 비디오 AI 데이터를 포함할 수 있다. AI 데이터에 비디오 AI 데이터만 포함된 경우, 비디오 AI 데이터는 복수의 디폴트 DNN 설정 정보로부터 영상 데이터의 모든 프레임들에 적용되는 업스케일 DNN 정보를 나타낼 수 있다.

[421] 일 실시 예에 따르면, AI 데이터는 프레임 그룹 AI 데이터를 포함할 수 있다. AI 데이터에 프레임 그룹 AI 데이터만 포함된 경우, 프레임 그룹 AI 데이터는 복수의 디폴트 DNN 설정 정보로부터 프레임 그룹의 모든 프레임에 적용되는 업스케일 DNN 정보를 나타낼 수 있다.

[422] 일 실시 예에 따르면, AI 데이터는 프레임 AI 데이터를 포함할 수 있다. AI 데이터에 프레임 AI 데이터만 포함된 경우, 프레임 AI 데이터는 복수의 디폴트 DNN 설정 정보로부터 프레임에 적용되는 업스케일 DNN 정보를 나타낼 수 있다.

[423] 일 실시 예에 따르면, AI 데이터는 비디오 AI 데이터와 함께 프레임 그룹 AI 데이터를 포함할 수 있다. AI 데이터에 비디오 AI 데이터 및 프레임 그룹 AI 데이터만 포함된 경우, 비디오 AI 데이터는 복수의 디폴트 DNN 설정 정보로부터 하나 이상의 업스케일 DNN 정보를 나타낼 수 있다. 그리고 프레임 그룹 AI 데이터에 따라, 상기 하나 이상의 업스케일 DNN 정보로부터, 프레임 그룹의 프레임들에 적용되는 업스케일 DNN 정보를 나타낼 수 있다.

[424] 일 실시 예에 따르면, AI 데이터는 비디오 AI 데이터와 함께 프레임 AI 데이터를 포함할 수 있다. AI 데이터에 비디오 AI 데이터 및 프레임 AI 데이터만 포함된 경우, 비디오 AI 데이터는 복수의 디폴트 DNN 설정 정보로부터 하나 이상의 업스케일 DNN 정보를 나타낼 수 있다. 그리고 프레임 AI 데이터는, 상기 하나 이상의 업스케일 DNN 정보로부터, 프레임에 적용되는 업스케일 DNN 정보를 나타낼 수 있다.

[425] 일 실시 예에 따르면, AI 데이터는 프레임 그룹 AI 데이터와 함께 프레임 AI 데이터를 포함할 수 있다. AI 데이터에 프레임 그룹 AI 데이터 및 프레임 AI 데이터만 포함된 경우, 프레임 그룹 AI 데이터는 복수의 디폴트 DNN 설정 정보로부터 하나 이상의 업스케일 DNN 정보를 나타낼 수 있다. 그리고 프레임 AI 데이터는, 상기 하나 이상의 업스케일 DNN 정보로부터, 프레임에 적용되는 업스케일 DNN 정보를 나타낼 수 있다.

[426] 일 실시 예에 따르면, AI 데이터는 비디오 세그먼트 AI 데이터를 포함할 수 있다. AI 데이터에 비디오 세그먼트 AI 데이터만 포함된 경우, 비디오 세그먼트 AI 데이터는, 복수의 디폴트 DNN 설정 정보로부터 비디오 세그먼트의 모든 프레임들에 적용되는 업스케일 DNN 정보를 나타낼 수 있다.

- [427] 일 실시 예에 따르면, AI 데이터는 비디오 세그먼트 AI 데이터와 함께 비디오 AI 데이터를 포함할 수 있다. AI 데이터에 비디오 AI 데이터 및 비디오 세그먼트 AI 데이터만 포함된 경우, 비디오 AI 데이터는, 복수의 디폴트 DNN 설정 정보로부터 하나 이상의 업스케일 DNN 정보를 나타낼 수 있다. 그리고 비디오 세그먼트 AI 데이터는, 상기 하나 이상의 업스케일 DNN 정보로부터, 비디오 세그먼트의 모든 프레임들에 적용되는 업스케일 DNN 정보를 나타낼 수 있다.
- [428] 일 실시 예에 따르면, AI 데이터는 비디오 세그먼트 AI 데이터와 함께 프레임 그룹 AI 데이터 또는 프레임 AI 데이터를 포함할 수 있다. AI 데이터에 비디오 세그먼트 AI 데이터와 함께 프레임 그룹 AI 데이터 또는 프레임 AI 데이터가 포함된 경우, 비디오 세그먼트 AI 데이터는, 복수의 디폴트 DNN 설정 정보로부터 하나 이상의 업스케일 DNN 정보를 나타낼 수 있다. 그리고 프레임 그룹 AI 데이터 또는 프레임 AI 데이터는, 상기 하나 이상의 업스케일 DNN 정보로부터, 프레임 그룹 또는 프레임에 적용되는 업스케일 DNN 정보를 나타낼 수 있다.
- [429] 일 실시 예에 따르면, AI 데이터는 비디오 AI 데이터, 비디오 세그먼트 AI 데이터, 프레임 그룹 AI 데이터 및 프레임 AI 데이터를 모두 포함할 수 있다. AI 데이터가 비디오 AI 데이터, 비디오 세그먼트 AI 데이터, 프레임 그룹 AI 데이터 및 프레임 AI 데이터를 모두 포함할 때, 비디오 AI 데이터는, 복수의 디폴트 DNN 설정 정보로부터, 비디오에 적용될 수 있는 하나 이상의 업스케일 DNN 정보를 나타낼 수 있다. 그리고 비디오 세그먼트 AI 데이터는, 상기 비디오에 적용될 수 있는 하나 이상의 업스케일 DNN 정보로부터, 비디오 세그먼트에 적용될 수 있는 하나 이상의 업스케일 DNN 정보를 나타낼 수 있다. 그리고 프레임 그룹 AI 데이터는, 상기 비디오 세그먼트에 적용될 수 있는 하나 이상의 업스케일 DNN 정보로부터, 프레임 그룹에 적용될 수 있는 하나 이상의 업스케일 DNN 정보를 나타낼 수 있다. 그리고 프레임 AI 데이터는, 상기 프레임 그룹에 적용될 수 있는 하나 이상의 업스케일 DNN 정보로부터, 프레임에 적용되는 업스케일 DNN 정보를 나타낼 수 있다. 상기 실시 예에서, 비디오 AI 데이터, 비디오 세그먼트 AI 데이터, 프레임 그룹 AI 데이터 및 프레임 AI 데이터 중 하나가 상기 설명된 계층적 AI 데이터 구조에서 제외될 수 있다.
- [430] 일 실시 예에 따르면, AI 데이터는 AI 업스케일이 적용되는지 여부를 나타내는 AI 타겟 데이터를 포함할 수 있다. 또한 AI 데이터는, AI 업스케일이 적용되는 경우, AI 다운스케일에 사용된 다운스케일 DNN 정보에 대응되는 업스케일 DNN 정보에 관한 AI 보조 데이터를 포함할 수 있다.
- [431] 비디오 AI 데이터는 영상 데이터에 AI 업스케일이 적용되는지 여부를 나타내는 비디오 AI 타겟 데이터 및 영상 데이터에 포함된 프레임들의 AI 업스케일에 적용될 수 있는 하나 이상의 업스케일 DNN 정보에 관한 비디오 AI 보조 데이터를 포함할 수 있다. 영상 데이터에 AI 업스케일이 적용되는지 여부에 따라 비디오 AI 타겟 데이터가 결정될 수 있다. 그리고 영상 데이터에 포함된

프레임들의 하나 이상의 업스케일 DNN 정보에 따라 비디오 AI 보조 데이터가 결정될 수 있다.

[432] 비디오 세그먼트 AI 데이터는 비디오 세그먼트에 AI 업스케일이 적용되는지 여부를 나타내는 비디오 세그먼트 AI 타겟 데이터 및 비디오 세그먼트에 포함된 프레임들의 AI 업스케일에 적용될 수 있는 하나 이상의 업스케일 DNN 정보에 관한 비디오 세그먼트 AI 보조 데이터를 포함할 수 있다. 비디오 세그먼트 AI 데이터는 비디오 세그먼트에 AI 업스케일이 적용되는지 여부에 따라 비디오 세그먼트 AI 타겟 데이터가 결정될 수 있다. 비디오 세그먼트에 포함된 프레임들의 하나 이상의 업스케일 DNN 정보에 따라 비디오 세그먼트 AI 보조 데이터를 포함할 수 있다.

[433] 프레임 그룹 AI 데이터는 프레임 그룹에 AI 업스케일이 적용되는지 여부를 나타내는 프레임 그룹 AI 타겟 데이터 및 프레임 그룹에 포함된 프레임들의 AI 업스케일에 적용될 수 있는 하나 이상의 업스케일 DNN 정보에 관한 프레임 그룹 AI 보조 데이터를 포함할 수 있다. 프레임 그룹에 AI 업스케일이 적용되는지 여부에 따라 프레임 그룹 AI 타겟 데이터가 결정될 수 있다. 그리고 프레임 그룹에 포함된 프레임들의 AI 업스케일에 적용될 수 있는 하나 이상의 업스케일 DNN 정보에 따라 프레임 그룹 AI 보조 데이터가 결정될 수 있다.

[434] 프레임 AI 데이터는 프레임에 AI 업스케일이 적용되는지 여부를 나타내는 프레임 AI 타겟 데이터 및 프레임의 AI 업스케일에 사용되는 업스케일 DNN 정보에 관한 프레임 AI 보조 데이터를 포함할 수 있다.

[435] 현재 프레임에 AI 업스케일이 적용되는지 여부에 따라 프레임 AI 타겟 데이터가 결정될 수 있다. 또한 현재 프레임의 AI 업스케일에 사용되는 업스케일 DNN 정보에 따라, 프레임 AI 보조 데이터가 결정될 수 있다.

[436] 일 실시 예에 따르면, 비디오 세그먼트 AI 데이터는 연속된 이전 비디오 세그먼트와 현재 비디오 세그먼트 간에 비디오 세그먼트 AI 보조 데이터가 동일한지 여부를 나타내는 비디오 세그먼트 AI 보조 데이터 종속 정보를 포함할 수 있다. 연속된 이전 비디오 세그먼트와 현재 비디오 세그먼트 간에 비디오 세그먼트 AI 보조 데이터가 동일한지 여부에 따라, 비디오 세그먼트 AI 보조 데이터 종속 정보가 결정된다.

[437] 그리고 프레임 그룹 AI 데이터는 연속된 이전 프레임 그룹과 현재 프레임 그룹 간에 프레임 그룹 AI 보조 데이터가 동일한지 여부를 나타내는 프레임 그룹 AI 보조 데이터 종속 정보를 포함할 수 있다. 연속된 이전 프레임 그룹과 현재 프레임 그룹 간에 프레임 그룹 AI 보조 데이터가 동일한지 여부에 따라, 프레임 그룹 AI 보조 데이터 종속 정보가 결정된다.

[438] 그리고 프레임 AI 데이터는 연속된 이전 프레임과 현재 프레임 간에 프레임 AI 보조 데이터가 동일한지 여부를 나타내는 프레임 AI 보조 데이터 종속 정보를 포함할 수 있다. 연속된 이전 프레임과 현재 프레임 간에 프레임 AI 보조 데이터가 동일한지 여부에 따라, 프레임 AI 보조 데이터 종속 정보가 결정된다.

- [439] 일 실시 예에 따르면, AI 부호화 데이터는 AI 데이터와 영상 데이터의 동기화에 대한 동기화 데이터를 포함할 수 있다. 동기화 데이터는 영상 데이터와 비디오 AI 데이터의 동기화, 비디오 세그먼트와 비디오 세그먼트 AI 데이터의 동기화, 프레임 그룹과 프레임 그룹 AI 데이터의 동기화, 현재 프레임과 프레임 AI 데이터의 동기화에 관한 데이터를 포함할 수 있다.
- [440] 일 실시 예에 따르면, AI 데이터는 현재 프레임의 AI 업스케일에 적용되는 색 채널을 나타내는 채널 정보를 포함할 수 있다. 그리고 채널 정보가 나타내는 색 채널에 대한 업스케일 DNN 정보가 결정될 수 있다.
- [441] 일 실시 예에 따르면, AI 데이터는 저해상도 영상의 비트레이트를 나타내는 목표 비트레이트 정보 및 고해상도 영상의 해상도를 나타내는 해상도 정보 중 적어도 하나를 포함할 수 있다.
- [442] 단계 2340에서, 현재 프레임의 저해상도 영상을 부호화함으로써, 영상 데이터가 획득된다. 단계 2340에 따른 영상 데이터 획득은 제1 부호화부(614)에 의하여 수행될 수 있다.
- [443] 단계 2350에서, 영상 데이터와 AI 데이터를 포함하는 AI 부호화 데이터가 생성된다. 단계 2350에 따른 AI 부호화 데이터 생성은 데이터 처리부(632)에 의하여 수행될 수 있다. 실시 예에 따라, 영상 데이터와 AI 데이터는 하나의 단일 파일에 포함되지 않고, 별개의 파일로 구성될 수 있다.
- [444] 단계 2360에서 상기 생성된 AI 부호화 데이터의 영상 데이터가 삽입된 미디어데이터 박스와 상기 AI 부호화 데이터의 AI 데이터가 삽입된 메타데이터 박스를 포함하는 비디오 파일이 출력된다. 단계 2360에 따른 AI 부호화 데이터 출력은 통신부(634)에 의하여 수행될 수 있다.
- [445] 도 12 내지 도 21에서 설명된 AI 업스케일에 대응되는 AI 다운스케일의 기능이 도 23의 비디오 부호화 방법에 적용될 수 있다.
- [446] 도 24는 일 실시 예에 따른 영상 복호화 장치의 구성을 나타내는 블록도이다.
- [447] 도 24를 참조하면, 일 실시 예에 따른 영상 복호화 장치(2400)는 통신부(2410), 프로세서(2420) 및 메모리(2430)를 포함할 수 있다.
- [448] 통신부(2410)는 AI 부호화 데이터를 수신할 수 있다. 또는 통신부(2410)는 프로세서(2420)의 제어에 의해 외부 장치(예를 들어, 서버)로부터 AI 데이터와 영상 데이터를 수신할 수 있다.
- [449] 프로세서(2420)는 영상 복호화 장치(2400)를 전반적으로 제어할 수 있다. 일 실시 예에 따른 프로세서(2420)는 메모리(2430)에 저장되는 하나 이상의 프로그램들을 실행할 수 있다. 프로세서(2420)는 제1 복호화부(232) 및 AI 업스케일부(234)의 기능을 수행할 수 있다. 프로세서(2420)는 하나 이상의 범용 프로세서로 구성될 수 있다.
- [450] 일 실시 예에 따르면, 프로세서(2420)는 그래픽 프로세서(2422)와 AI 전용 프로세서(2424)를 포함할 수 있다. 또한 프로세서는 그래픽 프로세서(2422) 및 AI 전용 프로세서(2424) 중 적어도 하나를 통합한 SoC(System On Chip) 형태로

구현될 수 있다. 프로세서(2422)는 영상 복호화 장치(2400)의 전반적인 동작 및 영상 복호화 장치(2400)의 내부 구성 요소들 사이의 신호 흐름을 제어하고, 데이터를 처리하는 기능을 수행한다.

[451] 그래픽 프로세서(2422)는 영상의 복호화 및 후처리에 특화하도록 설계된 프로세서이다. 따라서 그래픽 프로세서(2422)는 영상 복호화 장치(2400)가 수신한 영상 데이터에 대한 처리를 수행할 수 있으며, 제1 복호화부(232)의 저해상도 영상 복원 기능을 효율적으로 수행할 수 있다. AI 전용 프로세서(2424)는 AI 연산에 특화하도록 설계된 프로세서이다. 따라서 AI 전용 프로세서(2424)는 AI 업스케일부(234)의 저해상도 영상의 AI 업스케일 기능을 효율적으로 수행할 수 있다.

[452] 통신부(2410)를 통해 영상 복호화 장치(2400)에 입력된 영상 데이터와 AI 데이터는 프로세서(2420)에 의하여 처리된다. 영상 데이터의 복호화에 관한 연산이 그래픽 프로세서(2422)에 의하여 수행됨으로써, 저해상도 영상이 생성될 수 있다. 그리고 저해상도 영상의 AI 업스케일에 관한 연산은 AI 전용 프로세서(2424)에 의하여 수행될 수 있다. 따라서 AI 전용 프로세서(2424)에 의하여, 저해상도 영상이 AI 업스케일된 고해상도 영상이 생성될 수 있다.

[453] 프로세서(2420)는 하나의 그래픽 프로세서(2422)를 포함한 것으로 묘사되었지만, 실시 예에 따라, 2개 이상의 그래픽 프로세서(2422)를 포함할 수도 있다. 또한 프로세서(2420)는 하나의 AI 전용 프로세서(2424)를 포함한 것으로 묘사되었지만, 실시 예에 따라, 2개 이상의 AI 전용 프로세서(2424)를 포함할 수도 있다. 또한 프로세서(2420)는 하나 이상의 범용 프로세서를 포함할 수 있다. 그리고 하나 이상의 범용 프로세서에 따라, AI 업스케일에 필요한 추가적인 프로세스가 수행될 수 있다. AI 전용 프로세서(2424)는 하드웨어 FPGA로 구현될 수 있다.

[454] 일 실시예에 따른 메모리(2430)는 영상 복호화 장치(2400)를 구동하고 제어하기 위한 다양한 데이터, 프로그램 또는 어플리케이션을 저장할 수 있다. 메모리(2430)에 저장되는 프로그램은 하나 이상의 인스트럭션들을 포함할 수 있다. 그리고 메모리(2430)에 저장된 프로그램(하나 이상의 인스트럭션들) 또는 어플리케이션은 프로세서(2420)에 의해 실행될 수 있다. 메모리(2430)는 통신부(2410)와 프로세서(2420)로부터 비롯된 데이터를 저장할 수 있다. 또한 메모리(2430)는 프로세서(2420)가 요구하는 데이터를 프로세서(2420)에 전송할 수 있다.

[455] 일 실시예에 따른 영상 복호화 장치(2400)는 도 2에서 설명한 영상 복호화 장치의 각 기능들 및 도 22에서 설명한 영상 복호화 방법의 각 단계들 중 적어도 하나를 수행할 수 있다.

[456] 도 25는 일 실시예에 따른 영상 부호화 장치의 구성을 나타내는 블록도이다.

[457] 도 25를 참조하면, 일 실시예에 따른 영상 부호화 장치(2500)는 통신부(2510), 프로세서(2520) 및 메모리(2530)를 포함할 수 있다.

- [458] 일 실시예에 따른 프로세서(2520)는 영상 부호화 장치(2500)를 전반적으로 제어할 수 있다. 일 실시예에 따른 프로세서(2520)는 메모리(2530)에 저장되는 하나 이상의 프로그램들을 실행할 수 있다. 프로세서(2520)는 AI 다운스케일부(612) 및 제1 부호화부(614)의 기능을 수행할 수 있다. 프로세서(2520)는 하나 이상의 범용 프로세서로 구성될 수 있다.
- [459] 일 실시예에 따르면, 프로세서(2520)는 그래픽 프로세서(2522)와 AI 전용 프로세서(2524)를 포함할 수 있다. 또한 프로세서는 그래픽 프로세서(2522) 및 AI 전용 프로세서(2524) 중 적어도 하나를 통합한 SoC(System On Chip) 형태로 구현될 수 있다. 프로세서(2522)는 영상 부호화 장치(2500)의 전반적인 동작 및 영상 부호화 장치(2500)의 내부 구성 요소들 사이의 신호 흐름을 제어하고, 데이터를 처리하는 기능을 수행한다.
- [460] 그래픽 프로세서(2522)는 영상의 부호화 및 후처리에 특화하도록 설계된 프로세서이다. 따라서 그래픽 프로세서(2522)는 제1 부호화부(614)의 저해상도 영상 부호화 기능을 효율적으로 수행할 수 있다. AI 전용 프로세서(2524)는 AI 연산에 특화하도록 설계된 프로세서이다. 따라서 AI 전용 프로세서(2524)는 AI 다운스케일부(612)의 고해상도 영상의 AI 다운스케일 기능을 효율적으로 수행할 수 있다. AI 전용 프로세서(2524)는 하드웨어 FPGA로 구현될 수 있다.
- [461] 프로세서(2520)에 의하여, 고해상도 영상의 AI 다운스케일과 저해상도 영상의 부호화가 수행된다. 고해상도 영상의 AI 다운스케일을 위한 연산이 AI 전용 프로세서(2524)에 의하여 수행됨으로써, 저해상도 영상이 생성될 수 있다, 그리고 AI 전용 프로세서(2524)에 의하여, 저해상도 영상의 AI 업스케일에 필요한, AI 데이터가 생성될 수 있다. 저해상도 영상의 부호화를 위한 연산은 그래픽 프로세서(2522)에 의하여 수행됨으로써, 영상 데이터가 생성될 수 있다.
- [462] 통신부(2510)는, 프로세서(2520)의 제어에 따라, AI 데이터와 영상 데이터를 포함하는 AI 부호화 데이터의 단일 파일을 생성할 수 있다. 그리고 통신부(2510)는, 프로세서(2520)의 제어에 따라, AI 부호화 데이터의 단일 파일을 영상 부호화 장치(2500) 외부로 출력할 수 있다. 또는 통신부(2510)는, 프로세서(2520)의 제어에 따라, AI 데이터를 포함하는 파일과 영상 데이터를 포함하는 파일을 각각 생성할 수 있다. 그리고 통신부(2510)는, 프로세서(2520)의 제어에 따라, AI 데이터를 포함하는 파일과 영상 데이터를 포함하는 파일을 각각 영상 부호화 장치(2500) 외부로 출력할 수 있다.
- [463] 프로세서(2520)는 하나의 그래픽 프로세서(2522)를 포함한 것으로 묘사되었지만, 실시예에 따라, 2개 이상의 그래픽 프로세서(2522)를 포함할 수도 있다. 또한 프로세서(2520)는 하나의 AI 전용 프로세서(2524)를 포함한 것으로 묘사되었지만, 실시예에 따라, 2개 이상의 AI 전용 프로세서(2524)를 포함할 수도 있다. 또한 프로세서(2520)는 하나 이상의 범용 프로세서를 포함할 수 있다. 그리고 하나 이상의 범용 프로세서에 따라, AI 업스케일에 필요한 추가적인 프로세스가 수행될 수 있다.

- [464] 일 실시예에 따른 메모리(2530)는 영상 부호화 장치(2500)를 구동하고 제어하기 위한 다양한 데이터, 프로그램 또는 어플리케이션을 저장할 수 있다. 메모리(2530)에 저장되는 프로그램은 하나 이상의 인스트럭션들을 포함할 수 있다. 메모리(2530)에 저장된 프로그램(하나 이상의 인스트럭션들) 또는 어플리케이션은 프로세서(2520)에 의해 실행될 수 있다. 메모리(2530)는 원본 영상(105)과 같은 고해상도 영상을 저장할 수 있다. 메모리(2530)는 통신부(2510)와 프로세서(2520)로부터 비롯된 데이터를 저장할 수 있다. 또한 메모리(2530)는 프로세서(2520)가 요구하는 데이터를 프로세서(2520)에 전송할 수 있다.
- [465] 일 실시예에 따른 영상 부호화 장치(2500)는 도 7에서 설명한 영상 복호화 장치의 각 기능들 및 도 23에서 설명한 영상 부호화 방법의 각 단계들 중 적어도 하나를 수행할 수 있다.
- [466] 한편, 상술한 본 개시의 실시예들은 컴퓨터에서 실행될 수 있는 프로그램으로 작성가능하고, 작성된 프로그램은 매체에 저장될 수 있다.
- [467] 매체는 컴퓨터로 실행 가능한 프로그램을 계속 저장하거나, 실행 또는 다운로드를 위해 임시 저장하는 것일 수도 있다. 또한, 매체는 단일 또는 수개 하드웨어가 결합된 형태의 다양한 기록수단 또는 저장수단일 수 있는데, 어떤 컴퓨터 시스템에 직접 접속되는 매체에 한정되지 않고, 네트워크 상에 분산 존재하는 것일 수도 있다. 매체의 예시로는, 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체, CD-ROM 및 DVD와 같은 광기록 매체, 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical medium), 및 ROM, RAM, 플래시 메모리 등을 포함하여 프로그램 명령어가 저장되도록 구성된 것이 있을 수 있다. 또한, 다른 매체의 예시로, 어플리케이션을 유통하는 앱 스토어나 기타 다양한 소프트웨어를 공급 내지 유통하는 사이트, 서버 등에서 관리하는 기록매체 내지 저장매체가 있다.
- [468] 한편, 상술한 DNN과 관련된 모델은, 소프트웨어 모듈로 구현될 수 있다. 소프트웨어 모듈(예를 들어, 명령어(instruction)를 포함하는 프로그램 모듈)로 구현되는 경우, DNN 모델은 컴퓨터로 읽을 수 있는 판독 가능한 기록매체에 저장될 수 있다.
- [469] 또한, DNN 모델은 하드웨어 칩 형태로 집적되어 전술한 영상 복호화 장치(200) 또는 영상 부호화 장치(600)의 일부가 될 수도 있다. 예를 들어, DNN 모델은 인공 지능을 위한 전용 하드웨어 칩 형태로 제작될 수도 있고, 또는 기존의 범용 프로세서(예를 들어, CPU 또는 어플리케이션 프로세서) 또는 그래픽 전용 프로세서(예를 들어, GPU)의 일부로 제작될 수도 있다.
- [470] 또한, DNN 모델은 다운로드 가능한 소프트웨어 형태로 제공될 수도 있다. 컴퓨터 프로그램 제품은 제조사 또는 전자 마켓을 통해 전자적으로 배포되는 소프트웨어 프로그램 형태의 상품(예를 들어, 다운로드 가능한 어플리케이션)을 포함할 수 있다. 전자적 배포를 위하여, 소프트웨어 프로그램의 적어도 일부는

저장 매체에 저장되거나, 임시적으로 생성될 수 있다. 이 경우, 저장 매체는 제조사 또는 전자 마켓의 서버, 또는 중계 서버의 저장매체가 될 수 있다.

- [471] 이상, 본 개시의 기술적 사상을 바람직한 실시 예를 들어 상세하게 설명하였으나, 본 개시의 기술적 사상은 상기 실시 예들에 한정되지 않고, 본 개시의 기술적 사상의 범위 내에서 당 분야에서 통상의 지식을 가진 자에 의하여 여러 가지 변형 및 변경이 가능하다.

## 청구범위

- [청구항 1] AI 부호화 데이터를 포함하는 비디오 파일이 저장된 컴퓨터로 기록가능한 기록 매체에 있어서,  
상기 AI 부호화 데이터는,  
고해상도 영상을 AI 다운스케일하여 생성된 저해상도 영상의 부호화 정보를 포함하는 영상 데이터; 및  
상기 영상 데이터에 따라 복원된 저해상도 영상의 AI 업스케일에 대한 AI 데이터 포함하고,  
상기 AI 데이터는,  
하나 이상의 프레임에 AI 업스케일이 적용되는지 여부를 나타내는 AI 타겟 데이터; 및  
상기 하나 이상의 프레임에 AI 업스케일이 적용될 경우, 미리 설정된 복수의 디폴트 DNN 설정 정보 중에서, 상기 하나 이상의 프레임의 AI 업스케일에 사용되는 업스케일 DNN 정보에 관한 AI 보조 데이터를 포함하는 것을 특징으로 하는 컴퓨터로 기록가능한 기록 매체.
- [청구항 2] 제1항에 있어서,  
상기 AI 타겟 데이터는,  
영상 데이터에 포함된 프레임들에 AI 업스케일이 적용되는지 여부를 나타내는 비디오 AI 타겟 데이터;  
비디오 세그먼트에 포함된 프레임들에 AI 업스케일이 적용되는지 여부를 나타내는 비디오 세그먼트 AI 타겟 데이터;  
프레임 그룹에 포함된 프레임들에 AI 업스케일이 적용되는지 여부를 나타내는 프레임 그룹 AI 타겟 데이터; 및  
상기 현재 프레임에 AI 업스케일이 적용되는지 여부를 나타내는 프레임 AI 타겟 데이터 중 적어도 하나를 포함하는 것을 특징으로 하는 컴퓨터로 기록가능한 기록 매체.
- [청구항 3] 제1항에 있어서,  
상기 AI 보조 데이터는,  
영상 데이터에 포함된 프레임들의 AI 업스케일에 공통적으로 적용되는 하나 이상의 업스케일 DNN 정보에 대한 비디오 AI 보조 데이터;  
비디오 세그먼트에 포함된 프레임들의 AI 업스케일에 공통적으로 적용되는 하나 이상의 업스케일 DNN 정보에 대한 비디오 세그먼트 AI 보조 데이터  
프레임 그룹에 포함된 프레임들의 AI 업스케일에 공통적으로 적용되는 하나 이상의 업스케일 DNN 정보에 대한 프레임 그룹 AI 보조 데이터; 및  
상기 현재 프레임의 AI 업스케일에 적용되는 업스케일 DNN 정보에 대한 프레임 AI 보조 데이터 중 적어도 하나를 포함하는 것을 특징으로 하는

- 컴퓨터로 기록가능한 기록 매체.
- [청구항 4] 제1항에 있어서,  
 상기 비디오 파일은,  
 상기 영상 데이터를 포함하는 미디어데이터 박스; 및  
 상기 영상 데이터에 대한 메타 데이터를 포함하는 메타데이터 박스를 포함하고,  
 상기 AI 데이터는 메타데이터 박스에 포함된 것을 특징으로 하는  
 컴퓨터로 기록가능한 기록 매체.
- [청구항 5] 제4항에 있어서,  
 상기 메타데이터 박스는,  
 상기 영상 데이터와 상기 AI 데이터의 동기화에 관한 동기화 데이터를 포함하고,  
 상기 AI 데이터는,  
 상기 동기화 데이터에 기초하여, 상기 영상 데이터의 프레임들의 재생  
 순서 또는 복호화 순서에 따라, 복호화되도록 설정된 것을 특징으로 하는  
 컴퓨터로 기록가능한 기록 매체.
- [청구항 6] 제1항에 있어서,  
 상기 AI 데이터는 상기 영상 데이터에 삽입된 것을 특징으로 하고,  
 상기 영상 데이터의 AI 업스케일에 대한 비디오 AI 데이터는 상기 영상  
 데이터의 프레임들에 공통적으로 적용되는 비디오 부호화 파라미터와  
 함께 복호화되도록 상기 영상 데이터 안에 위치하고,  
 상기 현재 프레임이 포함된 현재 비디오 세그먼트의 AI 업스케일에 대한  
 비디오 세그먼트 AI 데이터는 상기 비디오 세그먼트의 프레임들에  
 공통적으로 적용되는 비디오 세그먼트 부호화 파라미터와 함께  
 복호화되도록 상기 영상 데이터 안에 위치하고,  
 상기 현재 프레임이 포함된 현재 프레임 그룹의 AI 업스케일에 대한  
 프레임 그룹 AI 데이터는 상기 프레임 그룹의 프레임들에 공통적으로  
 적용되는 프레임 그룹 파라미터와 함께 복호화되도록 상기 영상 데이터  
 안에 위치하고,  
 상기 현재 프레임의 AI 업스케일에 대한 프레임 AI 데이터는 상기 현재  
 프레임의 부호화 정보와 함께 복호화되도록 상기 영상 데이터 안에  
 위치한 것을 특징으로 하는 컴퓨터로 기록가능한 기록 매체.
- [청구항 7] 제6항에 있어서,  
 상기 영상 데이터는,  
 상기 영상 데이터의 비디오 부호화 파라미터를 포함하는 비디오 헤더,  
 상기 현재 비디오 세그먼트의 비디오 세그먼트 부호화 파라미터를  
 포함하는 비디오 세그먼트 헤더,  
 상기 현재 프레임 그룹의 프레임 그룹 부호화 파라미터를 포함하는

프레임 그룹 헤더, 및

상기 현재 프레임의 부호화 정보를 포함하는 프레임 헤더를 포함하고,  
상기 비디오 AI 데이터는 상기 비디오 헤더에 포함되거나 상기 비디오  
헤더의 바로 이전 또는 이후에 위치하는 것을 특징으로 하고,

상기 비디오 세그먼트 AI 데이터는 상기 비디오 세그먼트 헤더에  
포함되거나 상기 비디오 세그먼트 헤더의 바로 이전 또는 이후에  
위치하는 것을 특징으로 하고,

상기 프레임 그룹 AI 데이터는 상기 프레임 그룹 헤더에 포함되거나 상기  
프레임 그룹 헤더의 바로 이전 또는 이후에 위치하는 것을 특징으로 하고,  
상기 프레임 AI 데이터는 상기 프레임 헤더에 포함되거나, 상기 프레임  
헤더의 바로 이전 또는 이후에 위치하는 것을 특징으로 하는 컴퓨터로  
기록가능한 기록 매체.

[청구항 8]

제1항에 있어서,

상기 비디오 파일은,

상기 영상 데이터에 대한 메타 데이터를 포함하는 메타데이터 박스; 및  
상기 AI 부호화 데이터를 소정의 시간을 분할한 상기 복수의 비디오  
세그먼트 데이터 박스를 포함하고,

상기 복수의 비디오 세그먼트는 각각,

비디오 세그먼트 데이터를 포함하는 세그먼트 미디어데이터 박스; 및  
비디오 세그먼트 데이터에 대한 메타 데이터를 포함하는 세그먼트  
메타데이터 박스를 포함하고,

상기 AI 데이터는,

상기 영상 데이터의 모든 프레임의 AI 업스케일에 적용되는 비디오 AI  
데이터 및 상기 비디오 세그먼트의 모든 프레임의 AI 업스케일에  
적용되는 비디오 세그먼트 AI 데이터를 포함하고,

상기 비디오 AI 데이터는 상기 메타데이터 박스에 포함되고, 상기 비디오  
세그먼트 AI 데이터는 상기 세그먼트 메타데이터 박스에 포함된 것을  
특징으로 하는 컴퓨터로 기록가능한 기록 매체.

[청구항 9]

제3항에 있어서,

상기 AI 데이터는,

연속된 이전 비디오 세그먼트와 현재 비디오 세그먼트 간에 비디오  
세그먼트 AI 보조 데이터가 동일한지 여부를 나타내는 비디오 세그먼트  
AI 보조 데이터 종속 정보,

연속된 이전 프레임 그룹과 현재 프레임 그룹 간에 프레임 그룹 AI 보조  
데이터가 동일한지 여부를 나타내는 프레임 그룹 AI 보조 데이터 종속  
정보, 및

연속된 이전 프레임과 현재 프레임 간에 프레임 AI 보조 데이터가  
동일한지 여부를 나타내는 프레임 AI 보조 데이터 종속 정보 중 적어도

하나를 포함하고,  
 상기 비디오 세그먼트 AI 보조 데이터 종속 정보가 연속된 이전 비디오 세그먼트와 현재 비디오 세그먼트 간에 비디오 세그먼트 AI 보조 데이터가 동일함을 나타낼 때, 상기 현재 비디오 세그먼트에 대한 상기 비디오 세그먼트 AI 보조 데이터는 상기 AI 데이터로부터 생략되고,  
 상기 프레임 그룹 AI 보조 데이터 종속 정보가 연속된 이전 프레임 그룹과 현재 프레임 그룹 간에 프레임 그룹 AI 보조 데이터가 동일함을 나타낼 때, 상기 현재 프레임 그룹에 대한 상기 프레임 그룹 AI 보조 데이터는 상기 AI 데이터로부터 생략되고,  
 프레임 AI 보조 데이터 종속 정보가 연속된 이전 프레임과 현재 프레임 간에 프레임 AI 보조 데이터가 동일함을 나타낼 때, 상기 현재 프레임에 대한 상기 프레임 AI 보조 데이터는 상기 AI 데이터로부터 생략되는 것을 특징으로 하는 컴퓨터로 기록가능한 기록 매체.

- [청구항 10] 제1항에 있어서,  
 상기 AI 보조 데이터는,  
 AI 업스케일이 적용되는 색 채널을 나타내는 채널 정보;  
 상기 저해상도 영상의 비트레이트를 나타내는 목표 비트레이트 정보; 및  
 상기 AI 업스케일된 고해상도 영상의 해상도에 관련된 해상도 정보 중 적어도 하나를 포함하는 컴퓨터로 기록가능한 기록 매체.
- [청구항 11] AI 업스케일을 이용한 비디오 복호화 방법에 있어서,  
 영상 데이터와 상기 영상 데이터의 AI 업스케일에 관한 AI 데이터가 포함된 AI 부호화 데이터를 포함하는 비디오 파일을 수신하는 단계;  
 상기 비디오 파일의 메타데이터 박스로부터 상기 AI 부호화 데이터의 AI 데이터를 획득하고, 상기 비디오 파일의 미디어데이터 박스로부터, 상기 AI 부호화 데이터의 영상 데이터를 획득하는 단계;  
 상기 영상 데이터를 복호화하여, 현재 프레임의 저해상도 영상을 복원하는 단계;  
 상기 AI 데이터로부터, 상기 현재 프레임의 업스케일 DNN 정보를 획득하는 단계; 및  
 상기 현재 프레임의 업스케일 DNN 정보에 따라, 상기 저해상도 영상을 AI 업스케일함으로써, 상기 저해상도 영상에 대응되는 고해상도 영상을 생성하는 단계를 포함하는 비디오 복호화 방법.

- [청구항 12] 제11항에 있어서,  
 상기 AI 데이터는,  
 상기 현재 프레임의 AI 업스케일에 대한 프레임 AI 데이터, 상기 현재 프레임이 포함된 프레임 그룹의 AI 업스케일에 대한 프레임 그룹 AI 데이터, 상기 현재 프레임이 포함된 비디오 세그먼트의 AI 업스케일에 대한 비디오 세그먼트 AI 데이터, 또는 현재 프레임이 포함된 비디오의 AI

업스케일에 대한 비디오 AI 데이터 중 적어도 하나를 포함하는 것을 특징으로 하는 비디오 복호화 방법.

[청구항 13] 제11항에 있어서,  
 상기 AI 데이터는,  
 하나 이상의 프레임들에 AI 업스케일이 적용되는지 여부를 나타내는 AI 타겟 데이터; 및  
 상기 하나 이상의 프레임들의 AI 업스케일에 사용되는 업스케일 DNN 정보에 관한 AI 보조 데이터를 포함하고,  
 상기 업스케일 DNN 정보를 결정하는 단계는,  
 상기 AI 타겟 데이터에 따라, 상기 하나 이상의 프레임들에 AI 업스케일이 적용되는지 여부를 결정하는 단계; 및  
 상기 하나 이상의 프레임들에 AI 업스케일이 적용됨이 결정된 경우, 상기 AI 보조 데이터에 따라, 상기 하나 이상의 프레임들에 적용되는 하나 이상의 업스케일 DNN 정보를 결정하는 단계를 포함하는 비디오 복호화 방법.

[청구항 14] 제11항에 있어서,  
 상기 AI 데이터는,  
 상기 저해상도 영상의 비트레이트를 나타내는 목표 비트레이트 정보 및  
 상기 AI 업스케일된 고해상도 영상의 해상도에 관련된 해상도 정보 중 적어도 하나를 포함하고,  
 상기 현재 프레임의 업스케일 DNN 정보를 획득하는 단계는,  
 상기 목표 비트레이트 정보 및 상기 해상도 정보 중 적어도 하나에 따라, 복수의 디폴트 DNN 설정 정보로부터 상기 하나 이상의 프레임의 2개 이상의 업스케일 DNN 정보를 획득하는 것을 특징으로 하는 비디오 복호화 방법.

[청구항 15] 제11항에 있어서,  
 상기 비디오 파일의 메타데이터 박스는,  
 상기 영상 데이터와 상기 AI 데이터의 동기화에 대한 동기화 데이터를 포함하고,  
 상기 AI 데이터를 획득하는 단계는,  
 상기 동기화 데이터가 나타내는 상기 영상 데이터와 상기 AI 데이터의 부호화 순서 또는 재생 순서에 따라, 상기 현재 프레임에 대응되는 상기 AI 데이터를 획득하는 것을 특징으로 하는 비디오 복호화 방법.

[청구항 16] AI 다운스케일을 이용한 비디오 부호화 방법에 있어서,  
 현재 프레임의 고해상도 영상을 저해상도 영상으로 AI 다운스케일하기 위한, 다운스케일 DNN 정보를 결정하는 단계;  
 상기 다운스케일 DNN 정보에 따라, 상기 현재 프레임의 고해상도 영상을 AI 다운스케일함으로써, 상기 현재 프레임의 저해상도 영상을 생성하는

단계;

상기 현재 프레임의 저해상도 영상의 AI 업스케일에 사용되는, 상기 다운스케일 DNN 정보에 대응되는 업스케일 DNN 정보에 대한, AI 데이터를 생성하는 단계;

상기 현재 프레임의 저해상도 영상을 부호화함으로써, 영상 데이터를 획득하는 단계;

상기 영상 데이터와 상기 AI 데이터를 포함하는 AI 부호화 데이터를 생성하는 단계; 및

상기 AI 부호화 데이터의 영상 데이터가 삽입된 미디어데이터 박스와 상기 AI 부호화 데이터의 AI 데이터가 삽입된 메타데이터 박스를 포함하는 비디오 파일을 출력하는 단계를 포함하는 비디오 부호화 방법.

[청구항 17]

제16항에 있어서,

상기 AI 데이터는,

상기 현재 프레임의 AI 업스케일에 대한 프레임 AI 데이터, 상기 현재 프레임이 포함된 프레임 그룹의 AI 업스케일에 대한 프레임 그룹 AI 데이터, 상기 현재 프레임이 포함된 비디오 세그먼트의 AI 업스케일에 대한 비디오 세그먼트 AI 데이터, 또는 현재 프레임이 포함된 비디오의 AI 업스케일에 대한 비디오 AI 데이터인 것을 특징으로 하는 비디오 부호화 방법.

[청구항 18]

제16항에 있어서,

상기 AI 데이터는,

하나 이상의 프레임들에 AI 업스케일이 적용되는지 여부를 나타내는 AI 타겟 데이터; 및

상기 하나 이상의 프레임들에 AI 업스케일이 적용되는 경우, 상기 하나 이상의 프레임들의 AI 업스케일에 사용되는 업스케일 DNN 정보에 관한 AI 보조 데이터를 포함하고,

상기 AI 데이터를 결정하는 단계는,

상기 하나 이상의 프레임들에 AI 업스케일이 적용되는지 여부에 따라,

상기 AI 타겟 데이터를 결정하는 단계; 및

상기 AI 업스케일이 적용될 상기 하나 이상의 프레임들의 업스케일 DNN 정보에 따라, 상기 AI 보조 데이터를 결정하는 단계를 포함하는 비디오 부호화 방법.

[청구항 19]

제16항에 있어서,

상기 AI 데이터는,

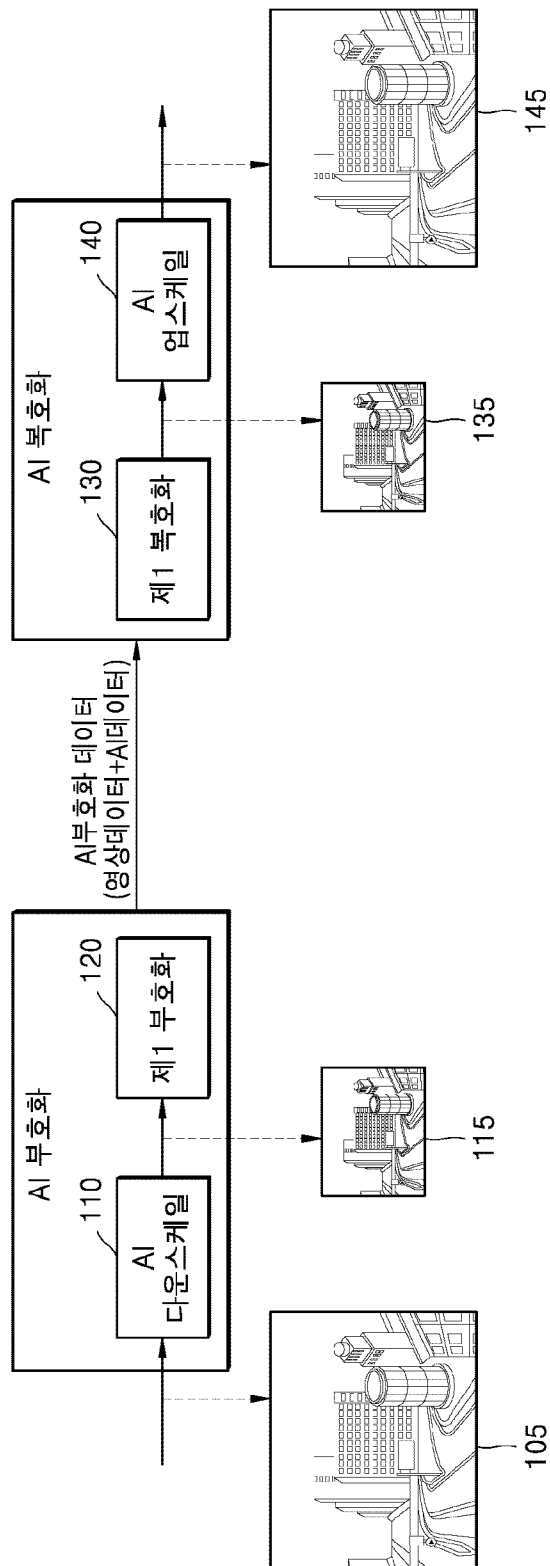
상기 저해상도 영상의 비트레이트를 나타내는 목표 비트레이트 정보 및 상기 고해상도 영상의 해상도를 나타내는 해상도 정보 중 적어도 하나를 포함하고,

상기 AI 데이터를 생성하는 단계는,

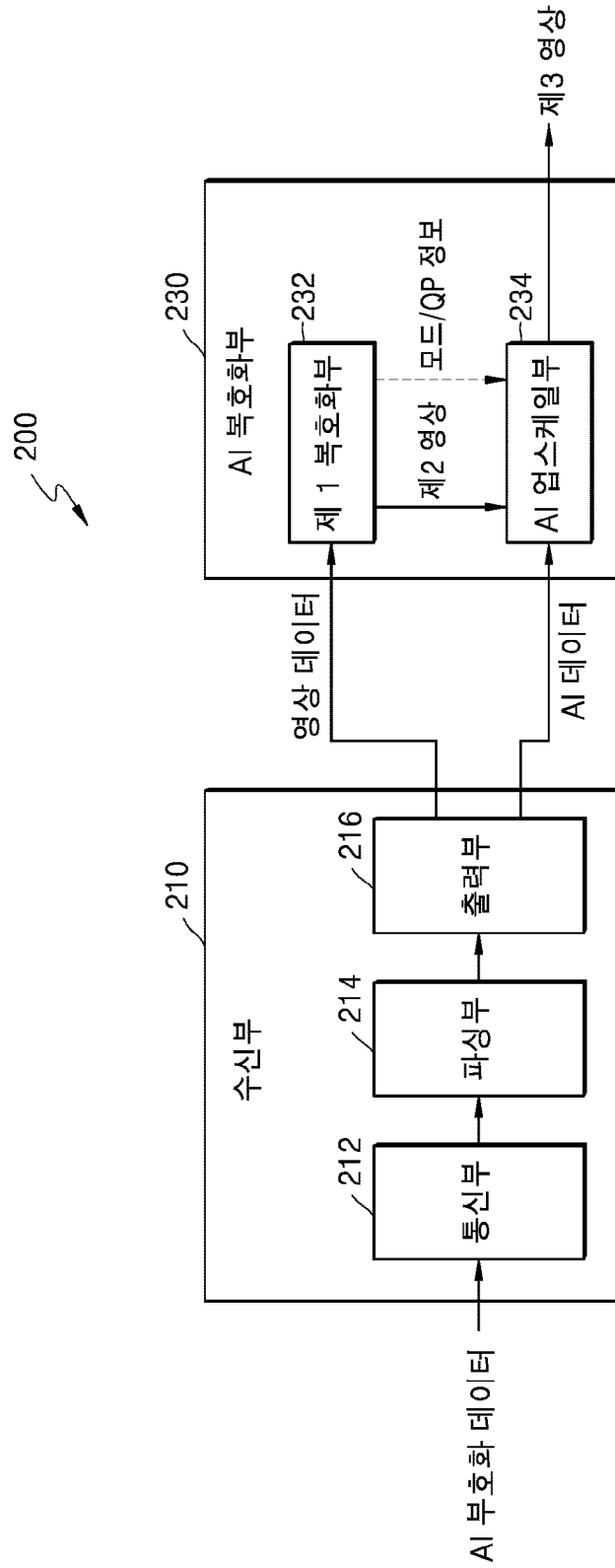
- 상기 목표 비트레이트 정보 및 상기 해상도 정보 중 적어도 하나에 따라, 복수의 디폴트 DNN 설정 정보로부터 하나 이상의 프레임의 2개 이상의 업스케일 DNN 정보를 결정하는 것을 특징으로 하는 비디오 부호화 방법.
- [청구항 20] 제16항에 있어서,  
상기 비디오 파일의 메타데이터 박스는,  
상기 AI 데이터와 상기 영상 데이터의 동기화에 대한 동기화 데이터를 포함하고,  
상기 AI 데이터를 결정하는 단계는,  
상기 영상 데이터와 상기 AI 데이터의 부호화 순서 또는 재생 순서에 따라, 동기화 데이터를 결정하는 단계를 포함하는 것을 특징으로 하는 비디오 부호화 방법.
- [청구항 21] AI 업스케일을 이용한 비디오 복호화 방법을 수행하는 비디오 복호화 장치에 있어서,  
영상 데이터와 상기 영상 데이터의 AI 업스케일에 관한 AI 데이터가 포함된 AI 부호화 데이터를 포함하는 비디오 파일을 수신하는 통신부;  
상기 비디오 파일의 메타데이터 박스로부터 상기 AI 부호화 데이터의 AI 데이터를 획득하고, 상기 비디오 파일의 미디어데이터 박스로부터, 상기 AI 부호화 데이터의 영상 데이터를 획득하는 파싱부;  
상기 영상 데이터를 복호화하여, 현재 프레임의 저해상도 영상을 복원하는 제1 복호화부; 및  
상기 AI 데이터로부터, 상기 현재 프레임의 업스케일 DNN 정보를 획득하고, 상기 현재 프레임의 업스케일 DNN 정보에 따라, 상기 저해상도 영상을 AI 업스케일함으로써, 상기 저해상도 영상에 대응되는 고해상도 영상을 생성하는 AI 업스케일부를 포함하는 비디오 복호화 장치.
- [청구항 22] AI 다운스케일을 이용한 비디오 부호화 방법을 수행하는 비디오 부호화 장치에 있어서,  
현재 프레임의 고해상도 영상을 저해상도 영상으로 AI 다운스케일하기 위한, 다운스케일 DNN 정보를 결정하고, 상기 다운스케일 DNN 정보에 따라, 상기 현재 프레임의 고해상도 영상을 AI 다운스케일함으로써, 상기 현재 프레임의 저해상도 영상을 생성하고, 상기 현재 프레임의 저해상도 영상의 AI 업스케일에 사용되는, AI 데이터를 생성하는 AI 다운스케일부;  
상기 현재 프레임의 저해상도 영상을 부호화함으로써, 영상 데이터를 결정하는 제1 부호화부;  
상기 영상 데이터와 상기 AI 데이터를 포함하는 AI 부호화 데이터를 결정하는 데이터 처리부; 및  
상기 AI 부호화 데이터의 영상 데이터가 삽입된 미디어데이터 박스와  
상기 AI 부호화 데이터의 AI 데이터가 삽입된 메타데이터 박스를

포함하는 비디오 파일을 출력하는 통신부를 포함하는 비디오 부호화 장치.

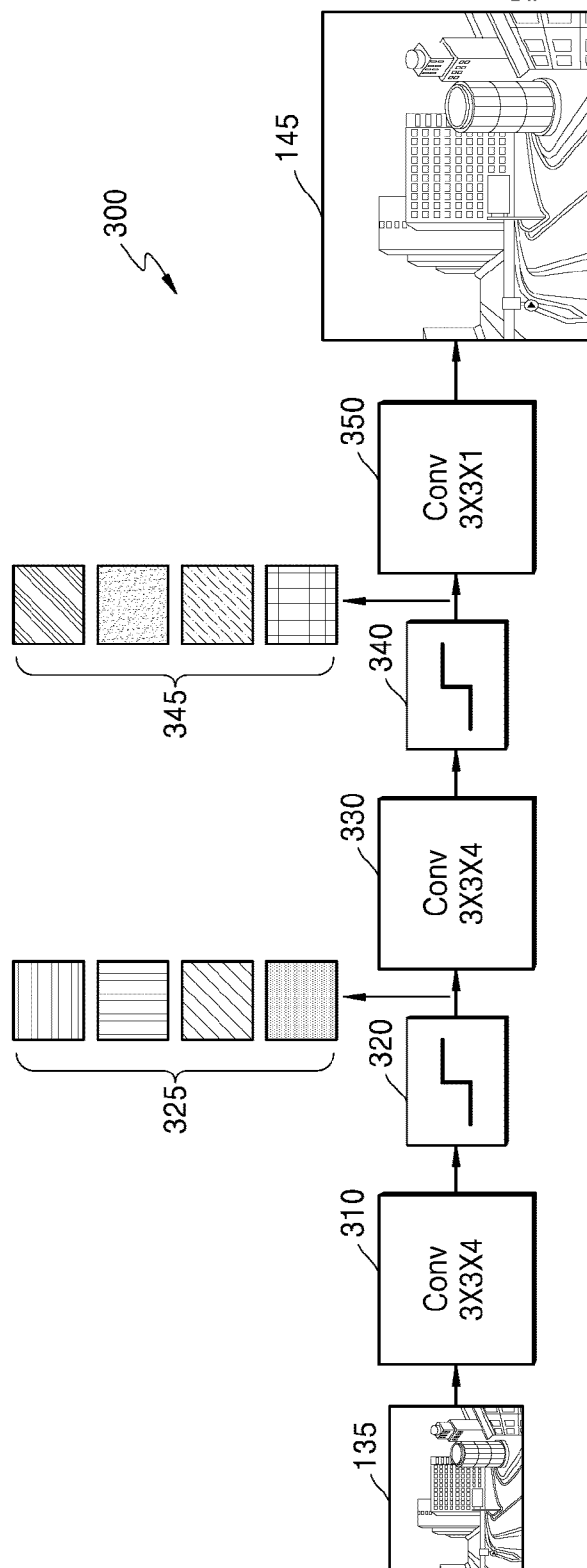
[도 1]



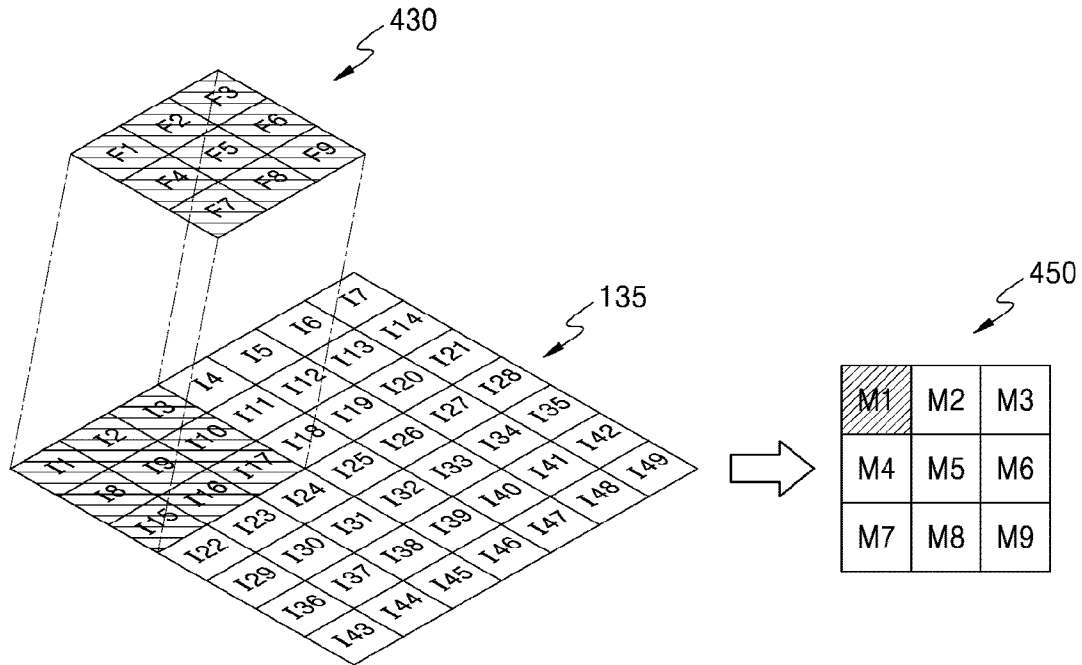
[도2]



[도3]



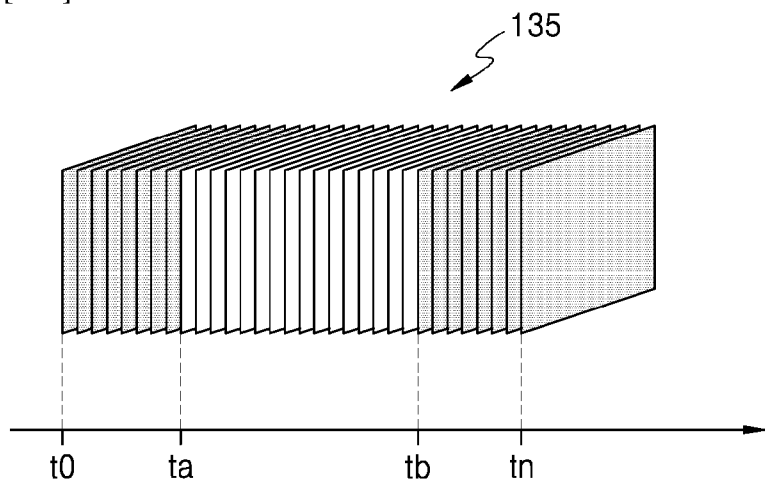
[도4]



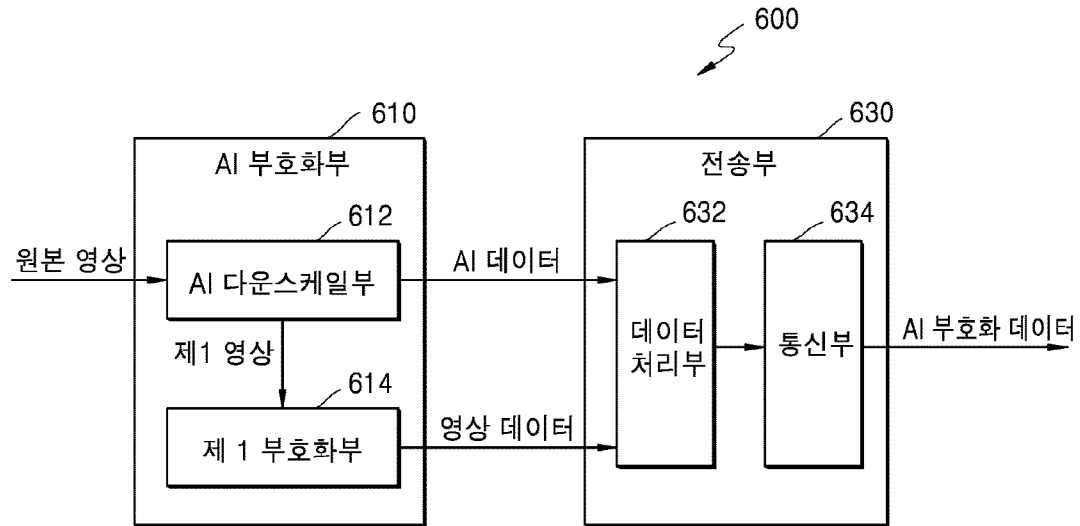
[도5]

| 영상 관련 정보              | DNN 설정 정보   |
|-----------------------|-------------|
| SD, 10Mbps, AV1       | A DNN 설정 정보 |
| HD, 15Mbps, H.264     | B DNN 설정 정보 |
| Full HD, 20Mbps, HEVC | C DNN 설정 정보 |
| Full HD, 15Mbps, HEVC | D DNN 설정 정보 |

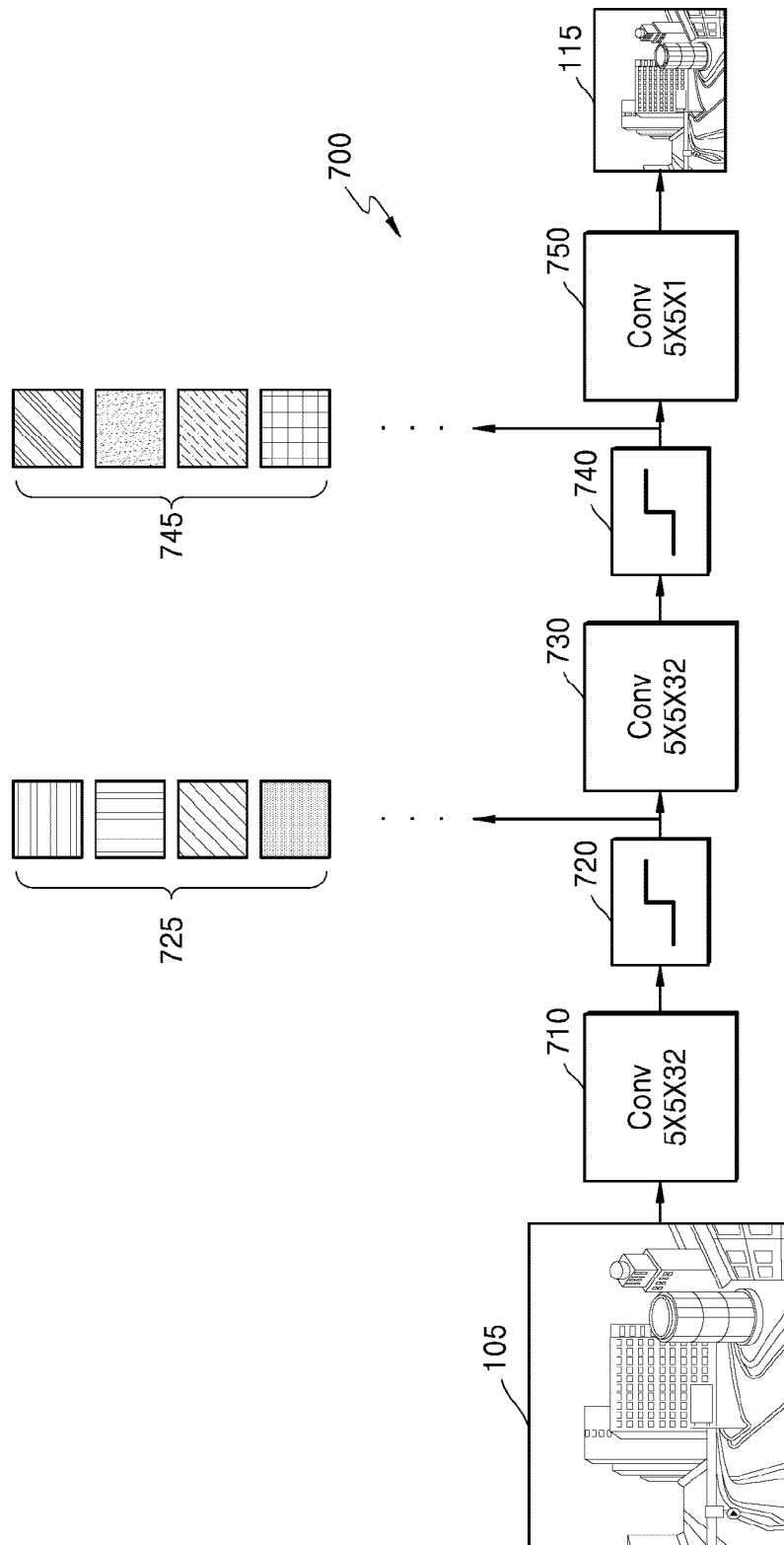
[도6]



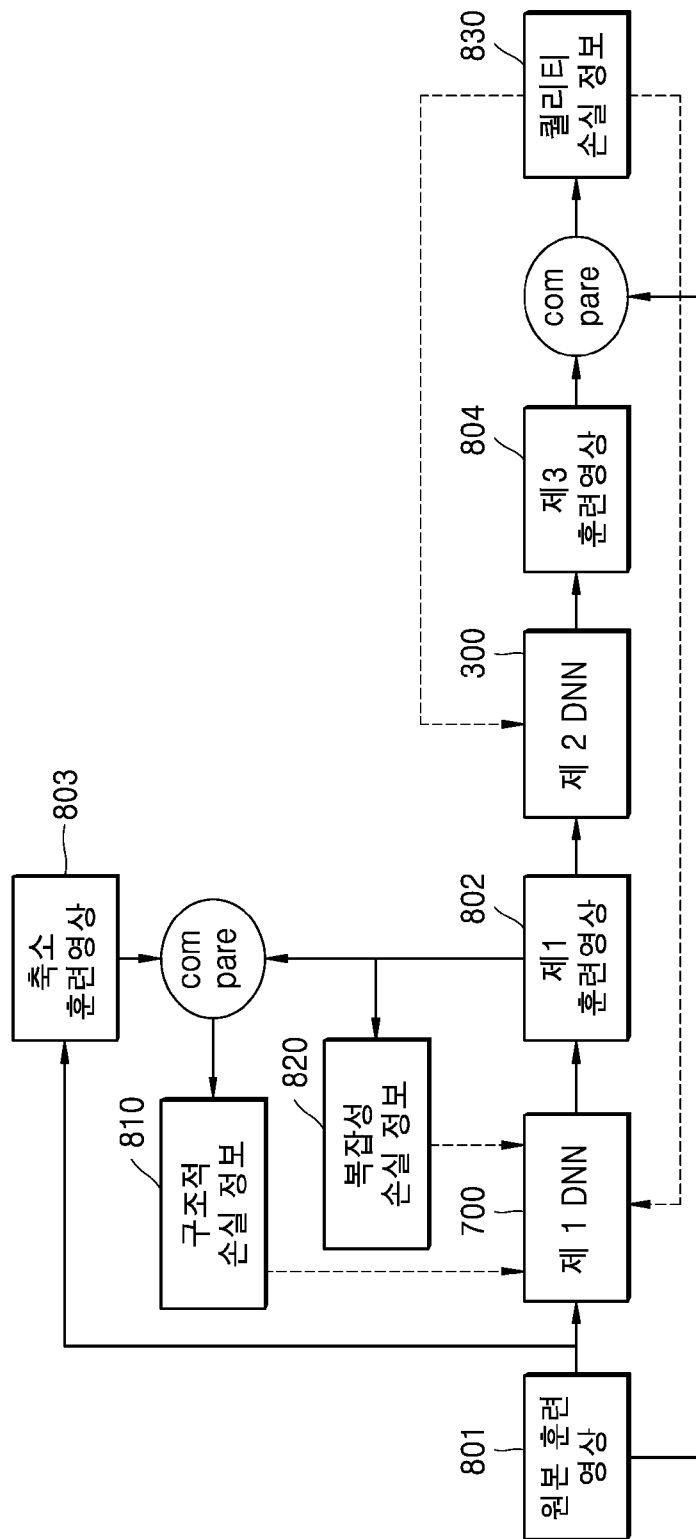
[도7]



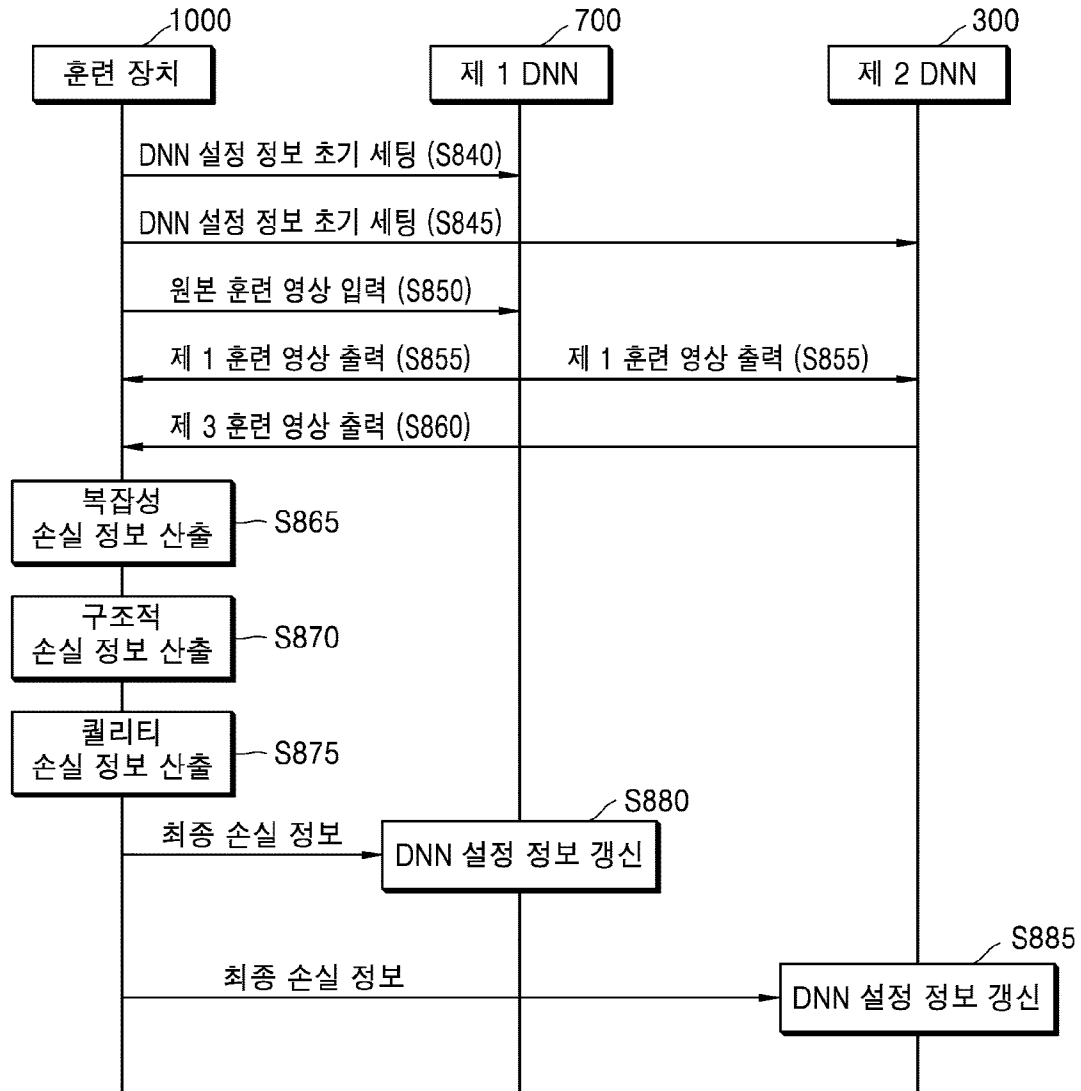
[도8]



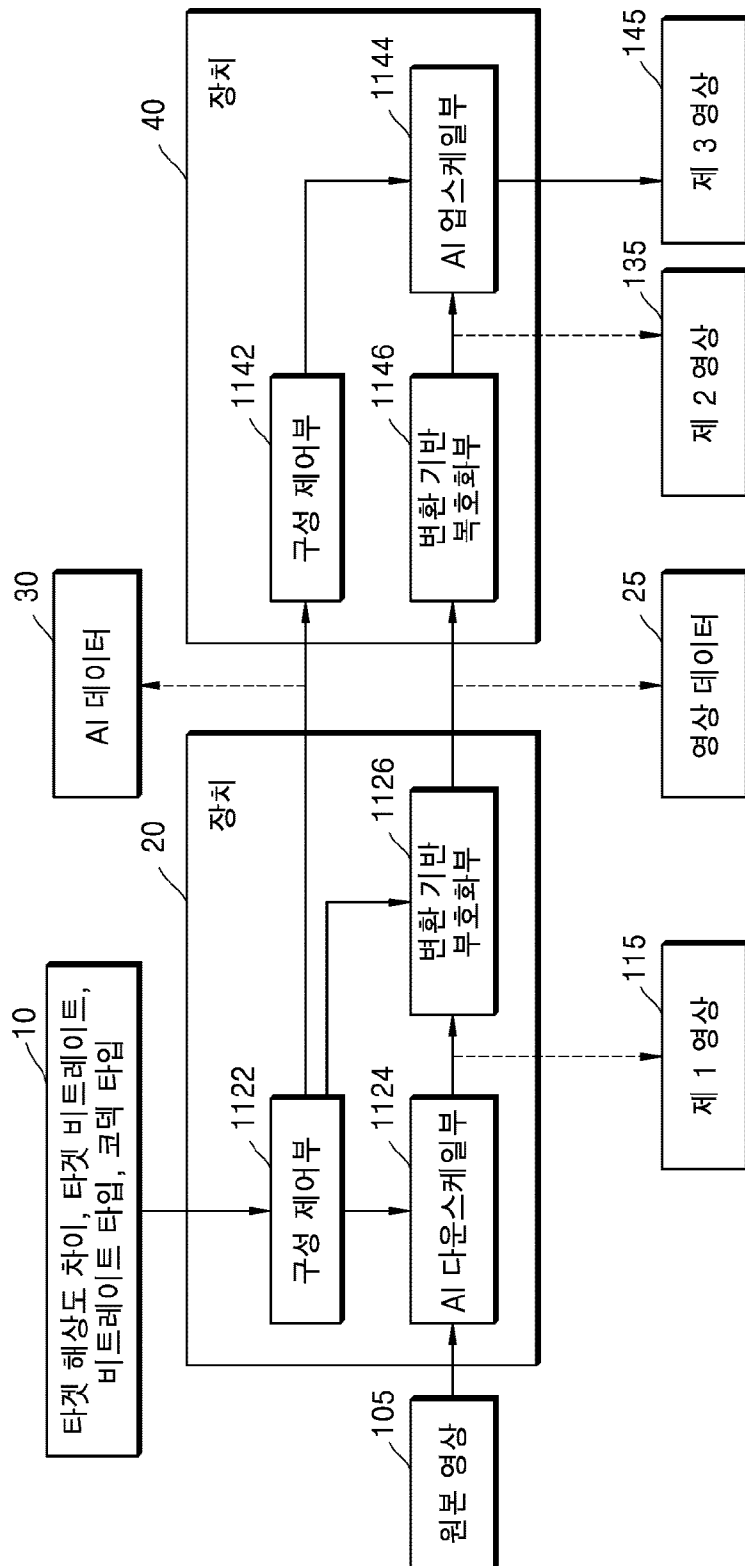
[도9]



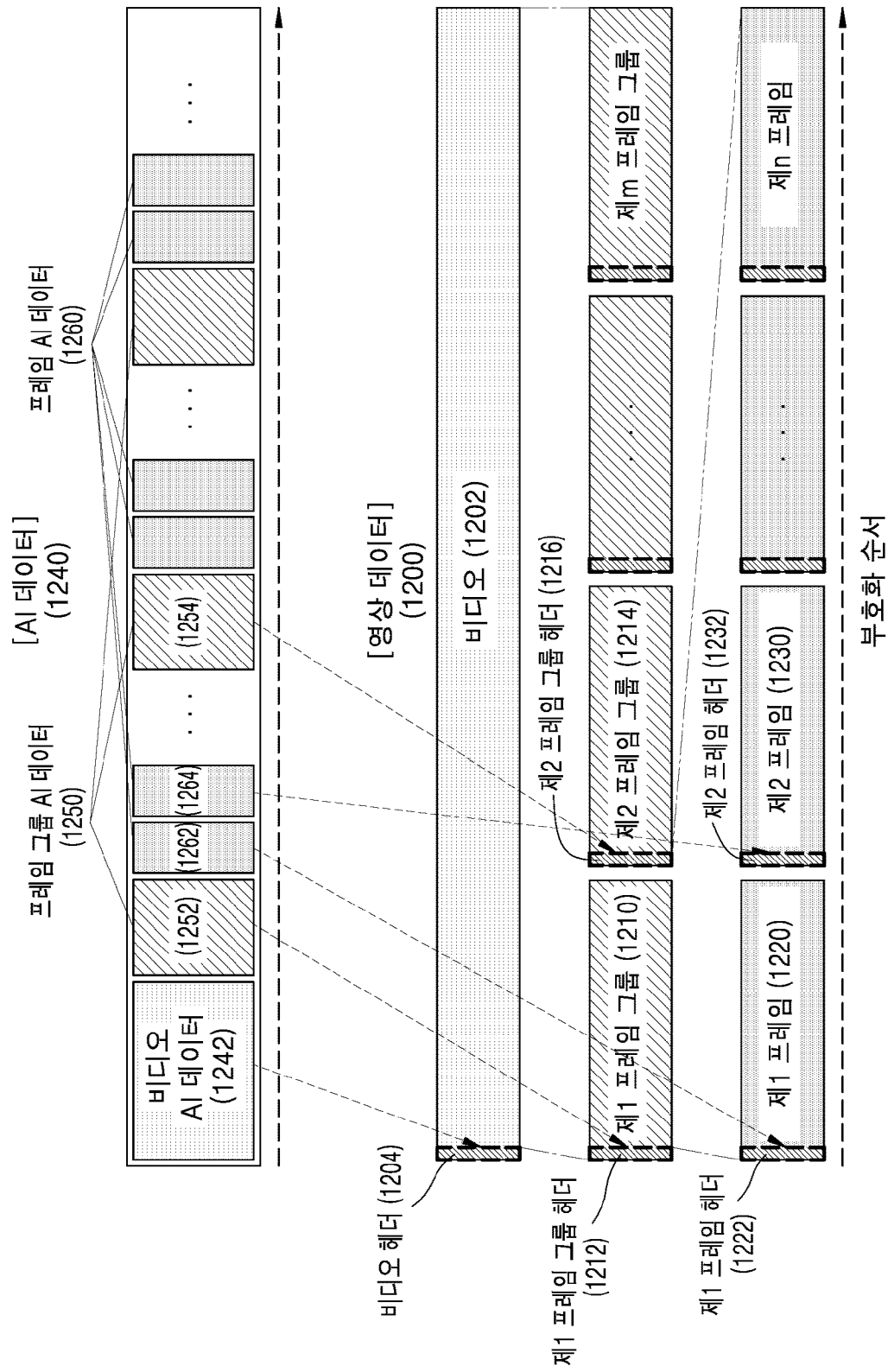
[도10]



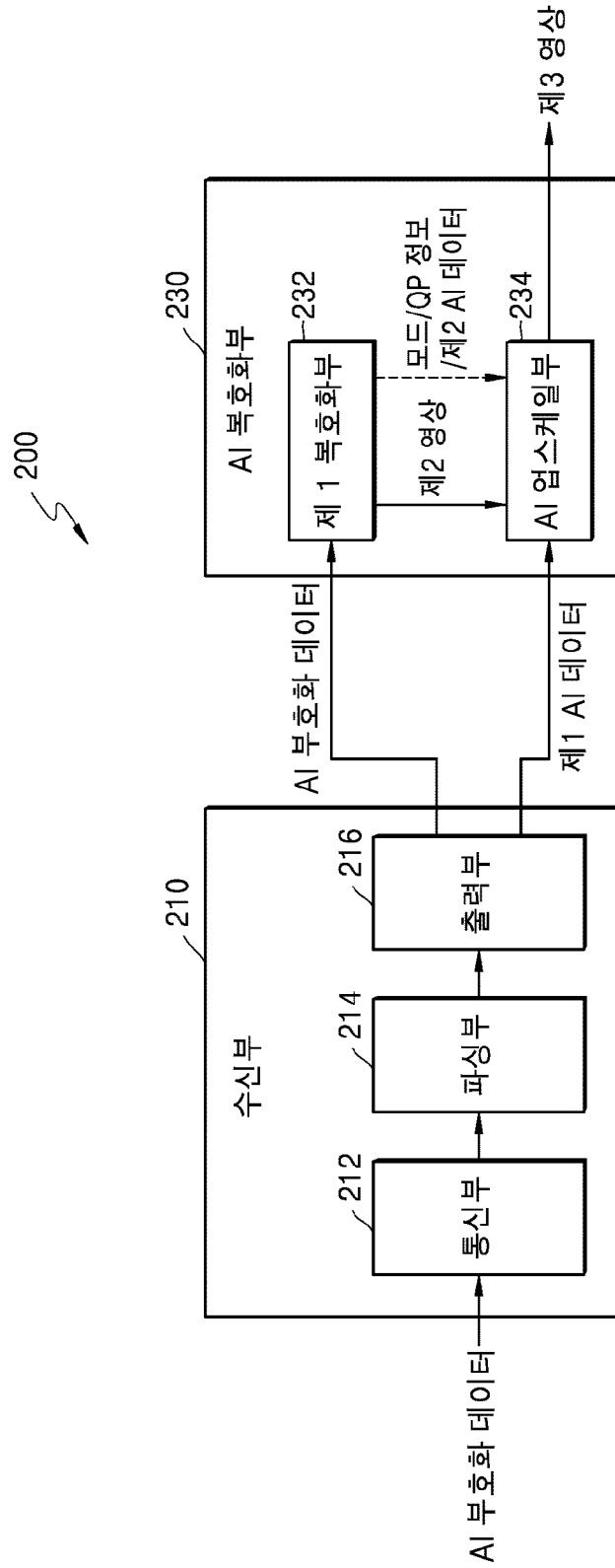
[도 11]



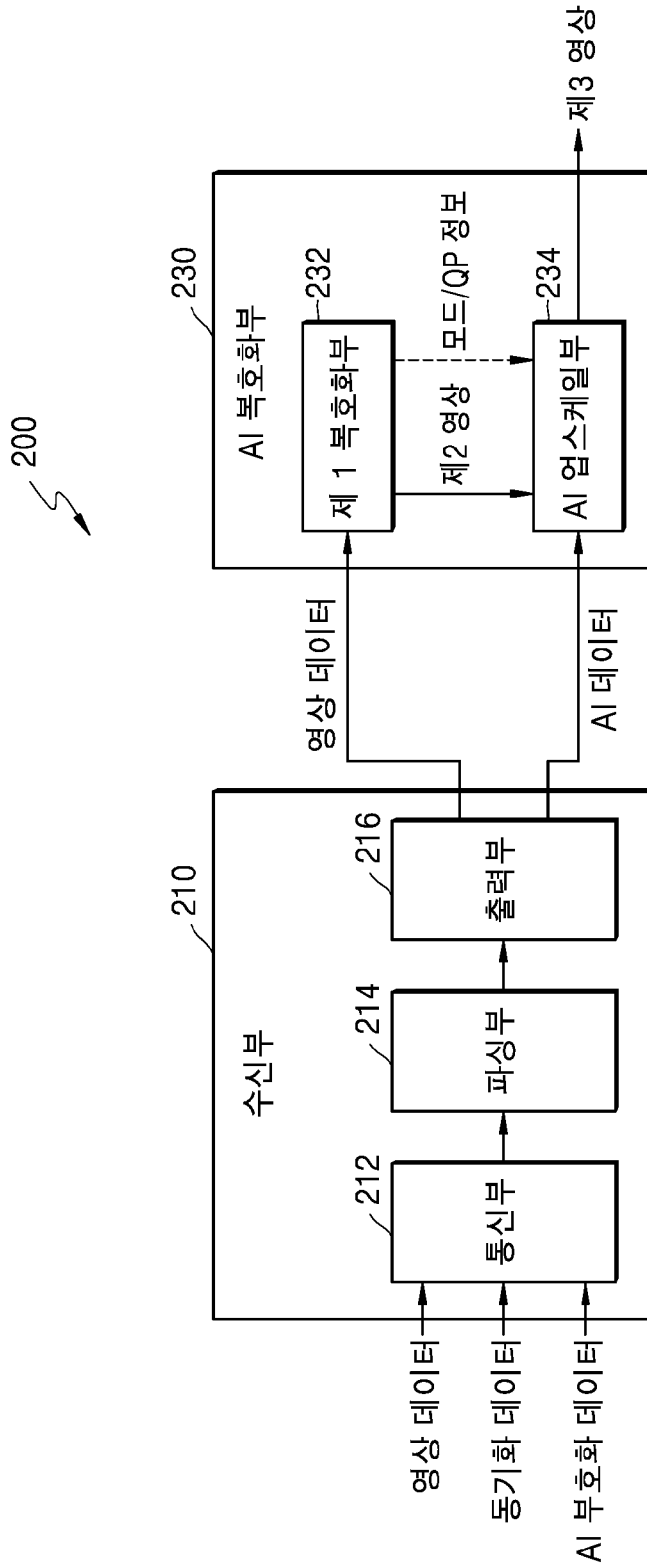
[도12]



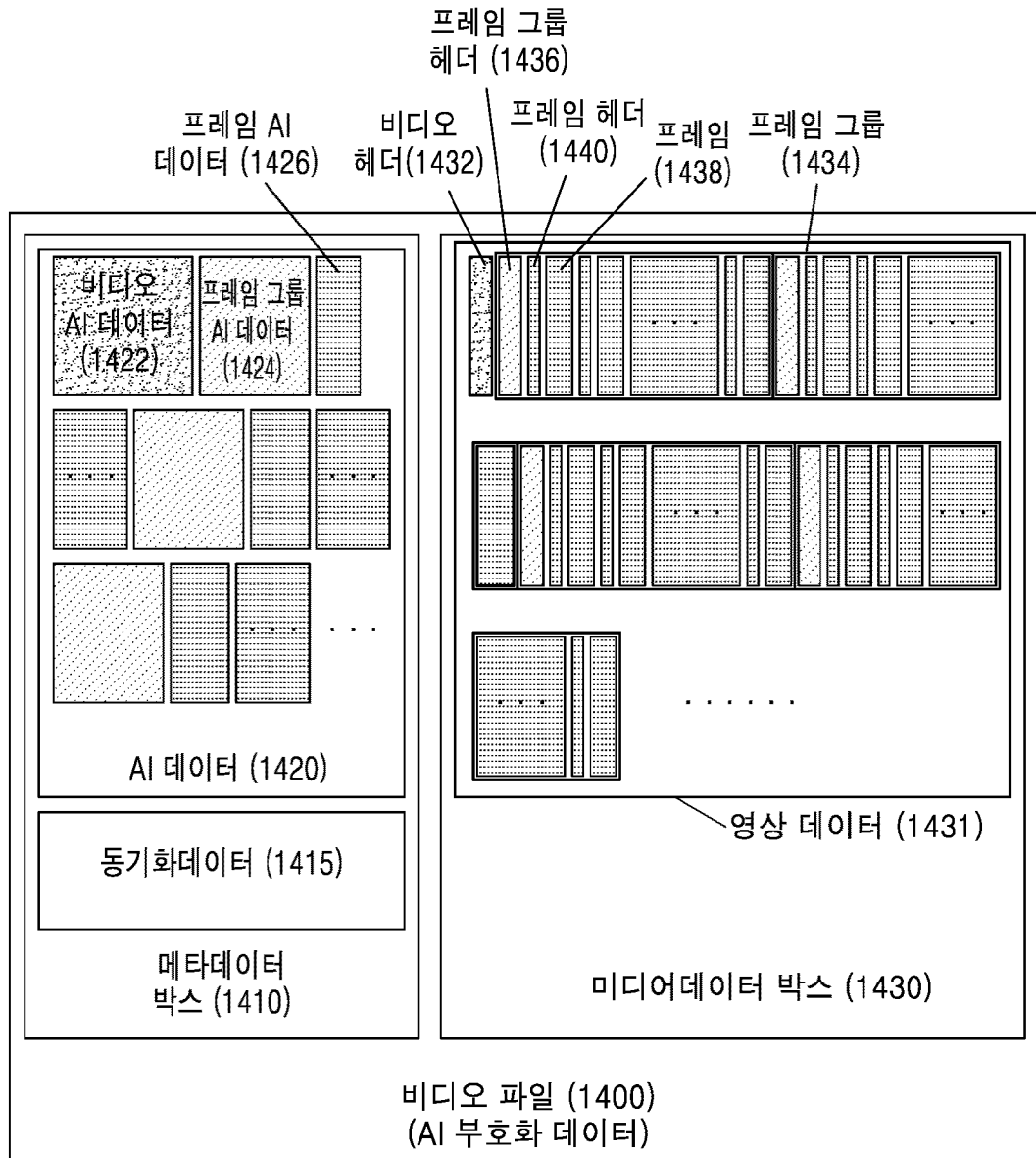
[도 13a]



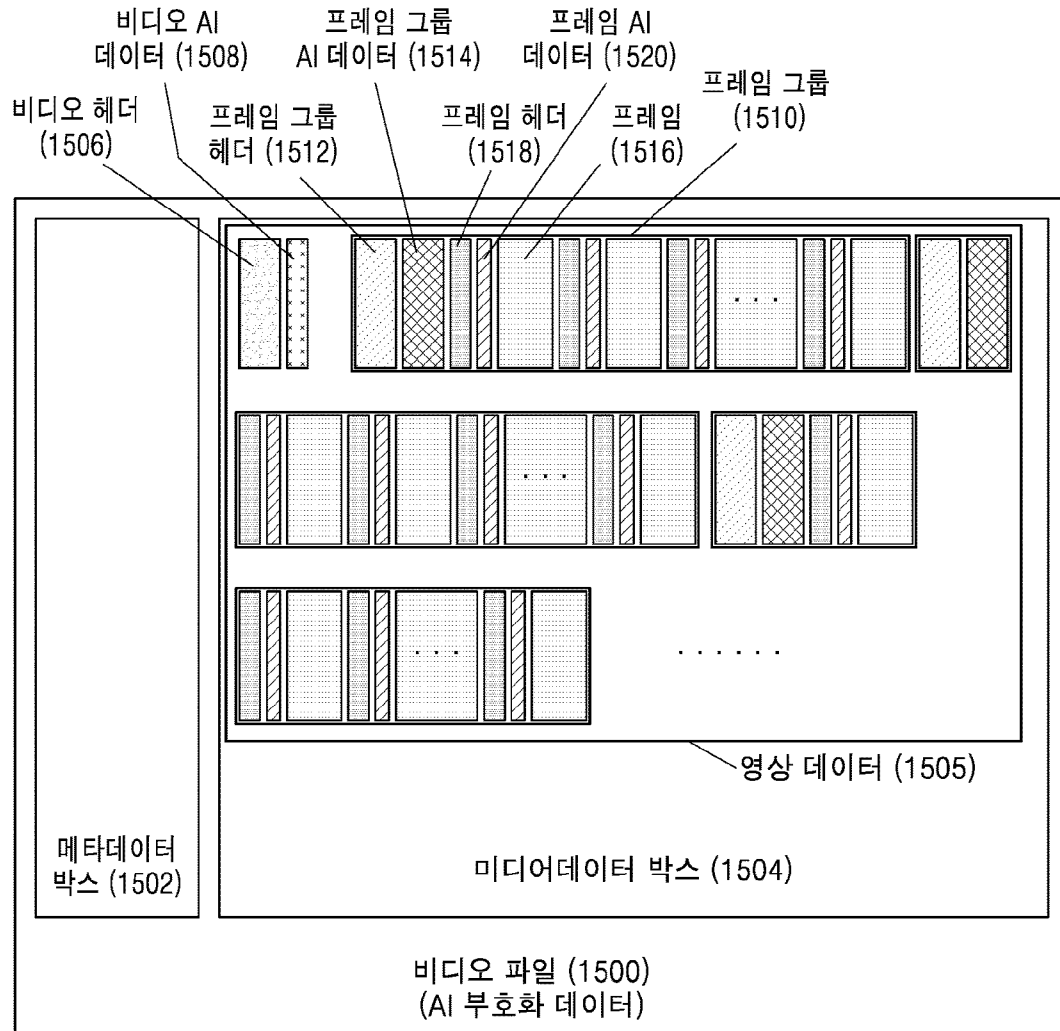
[도13b]



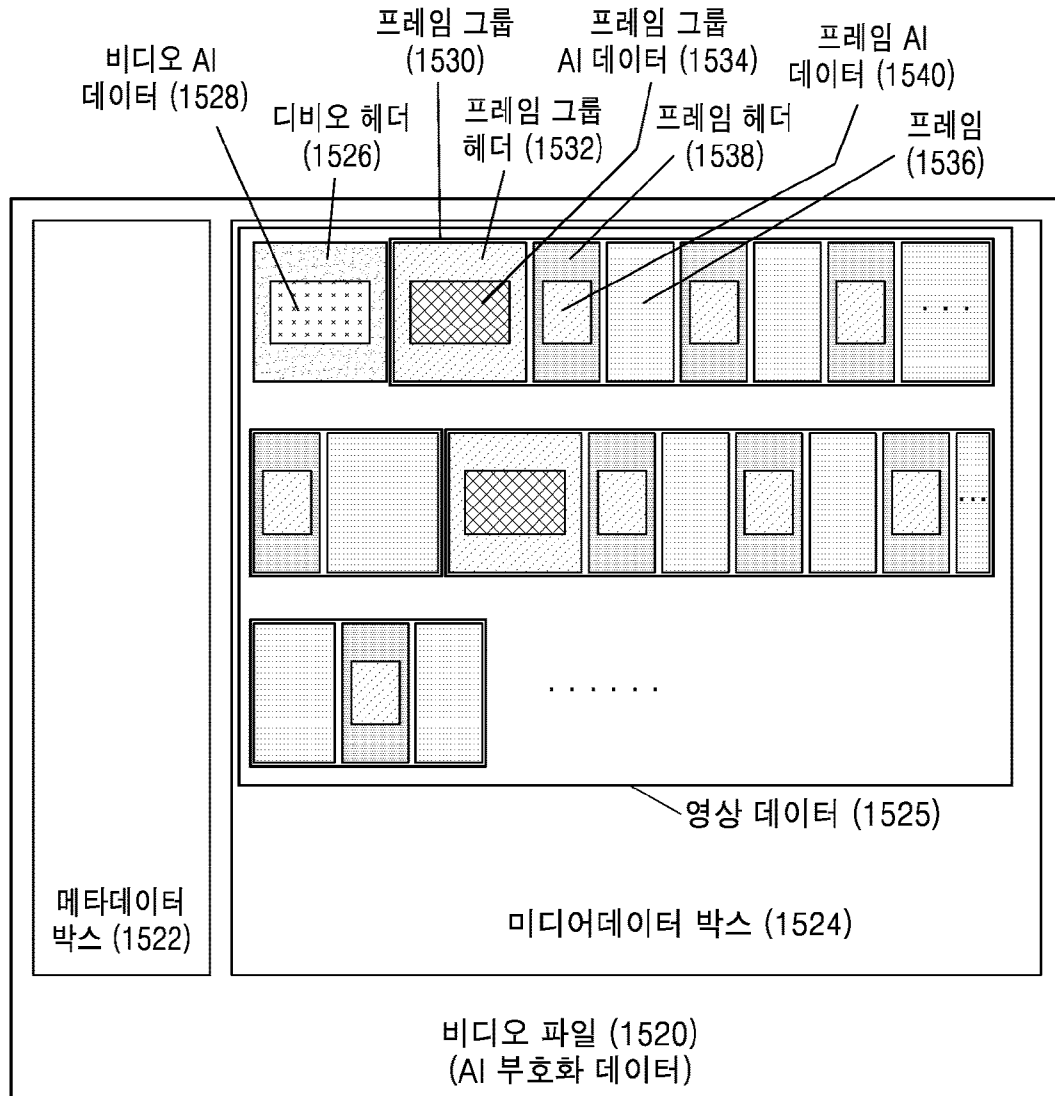
[도14]



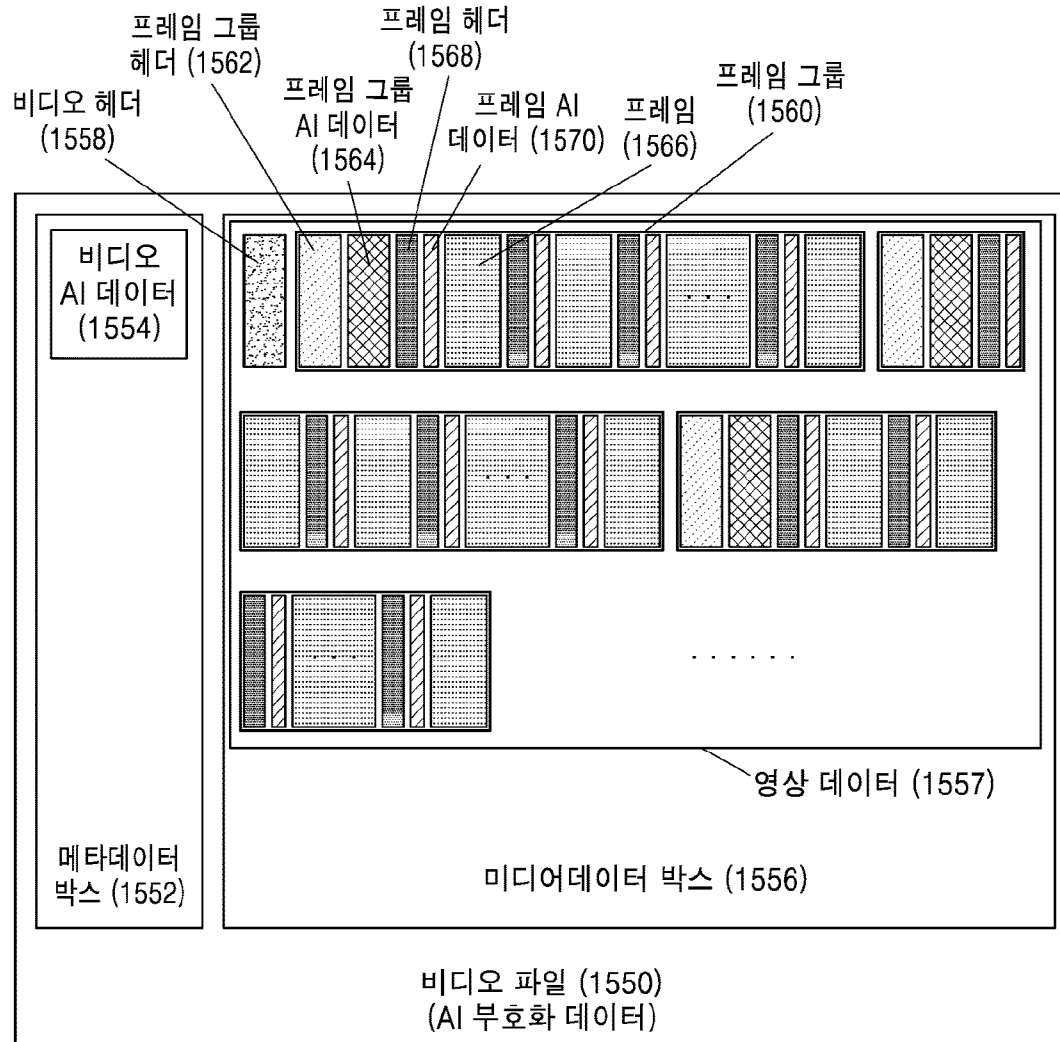
[도 15a]



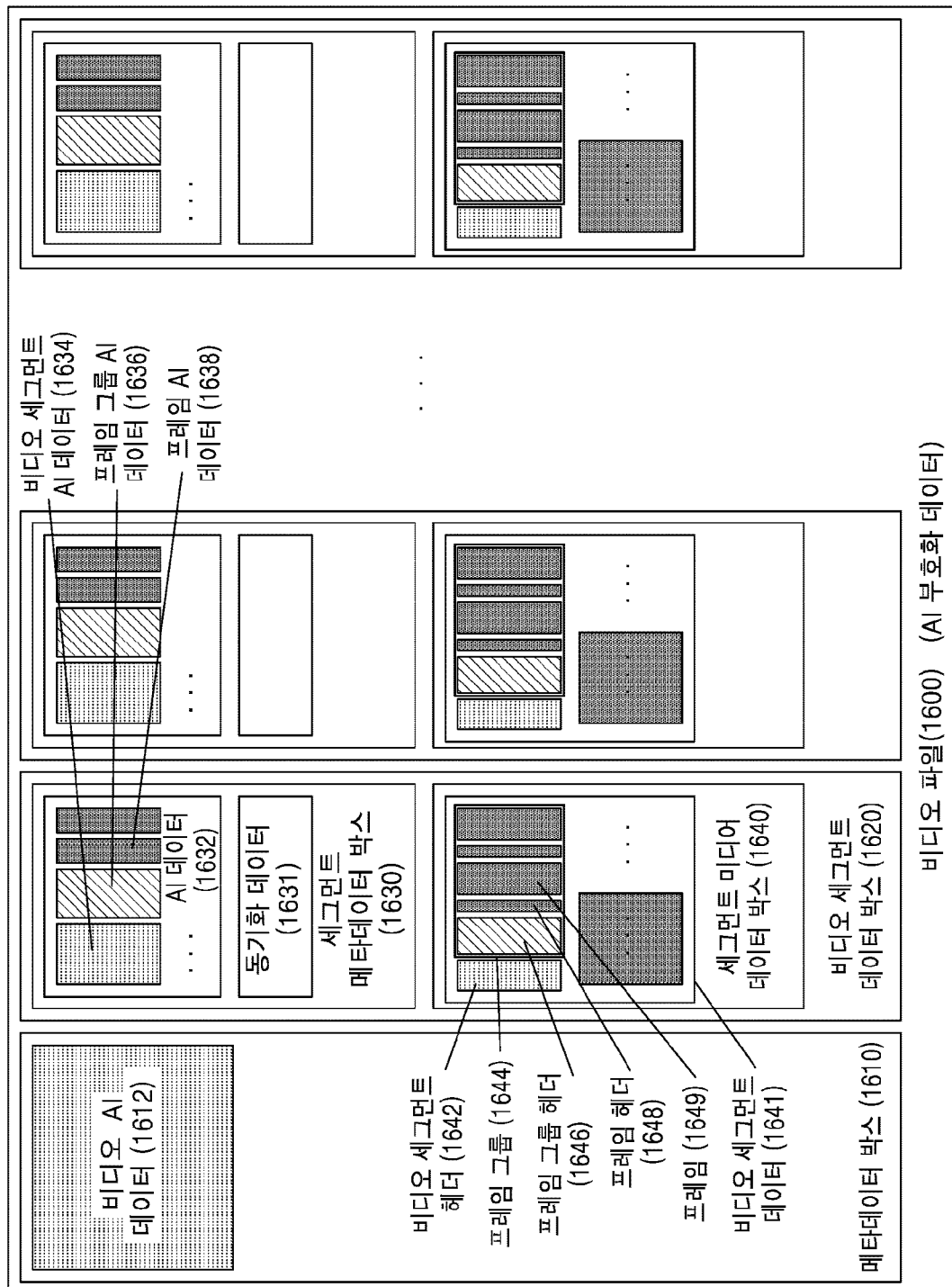
[도 15b]



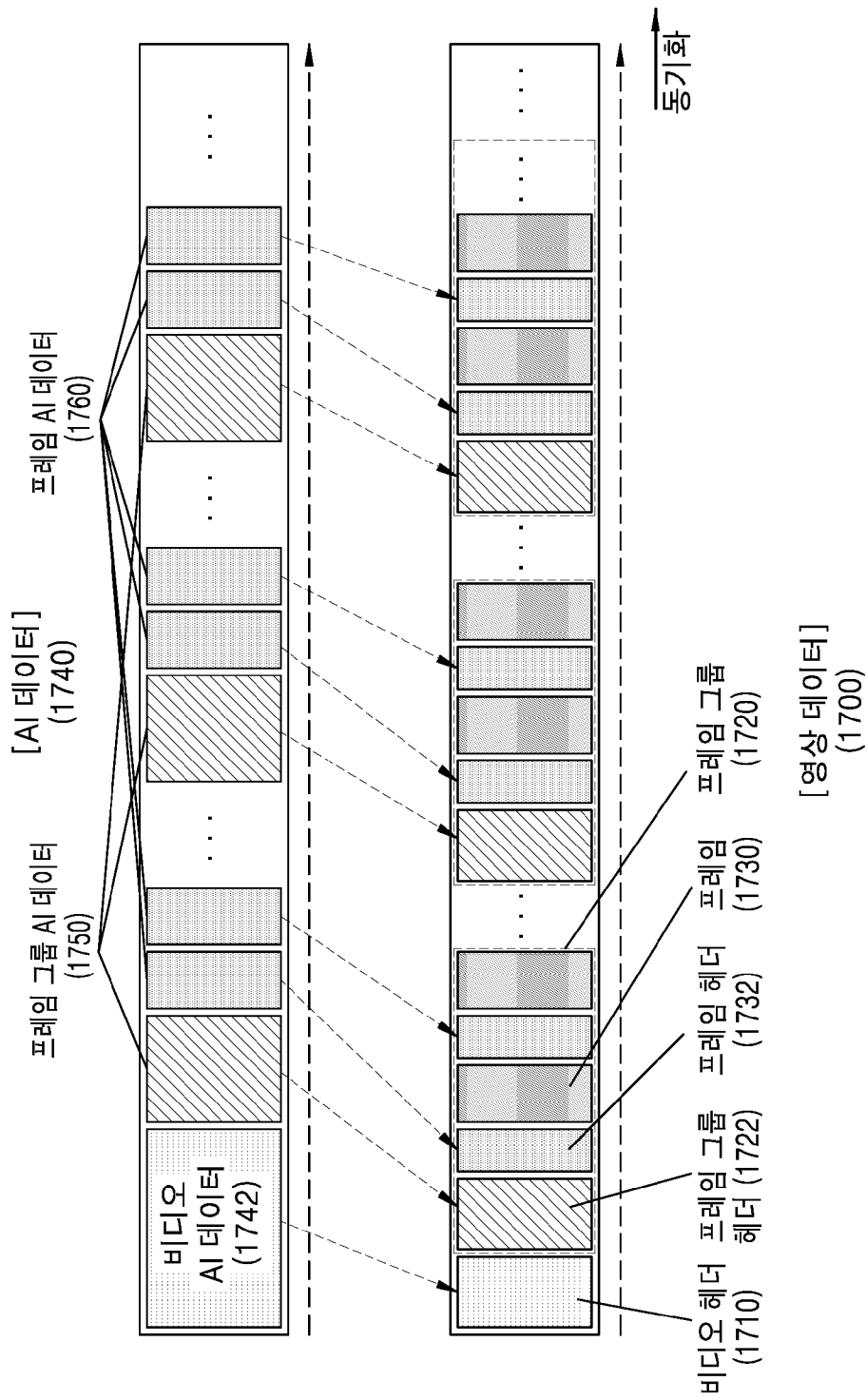
[도 15c]



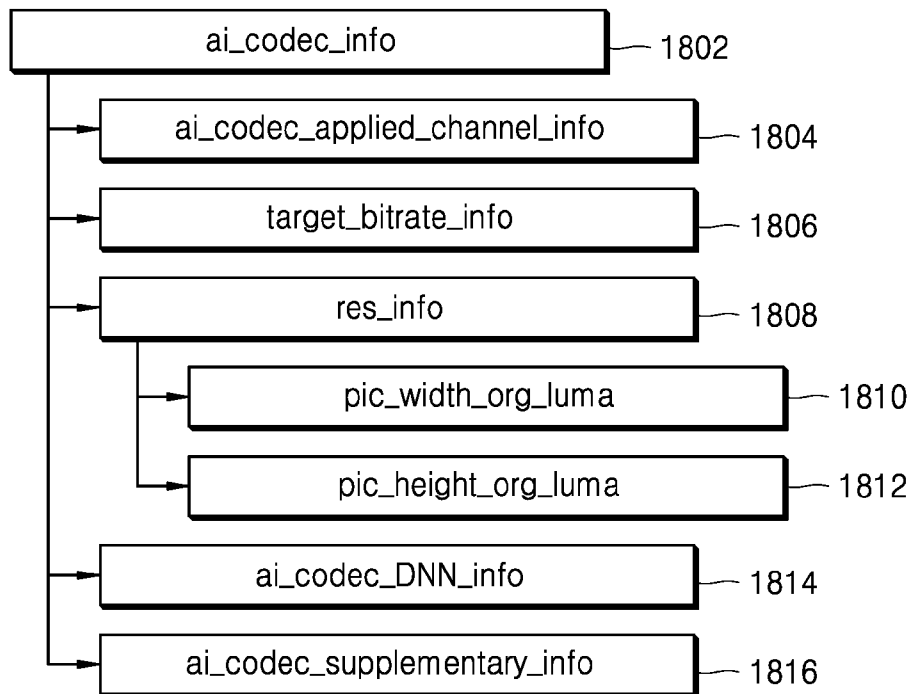
[도16]



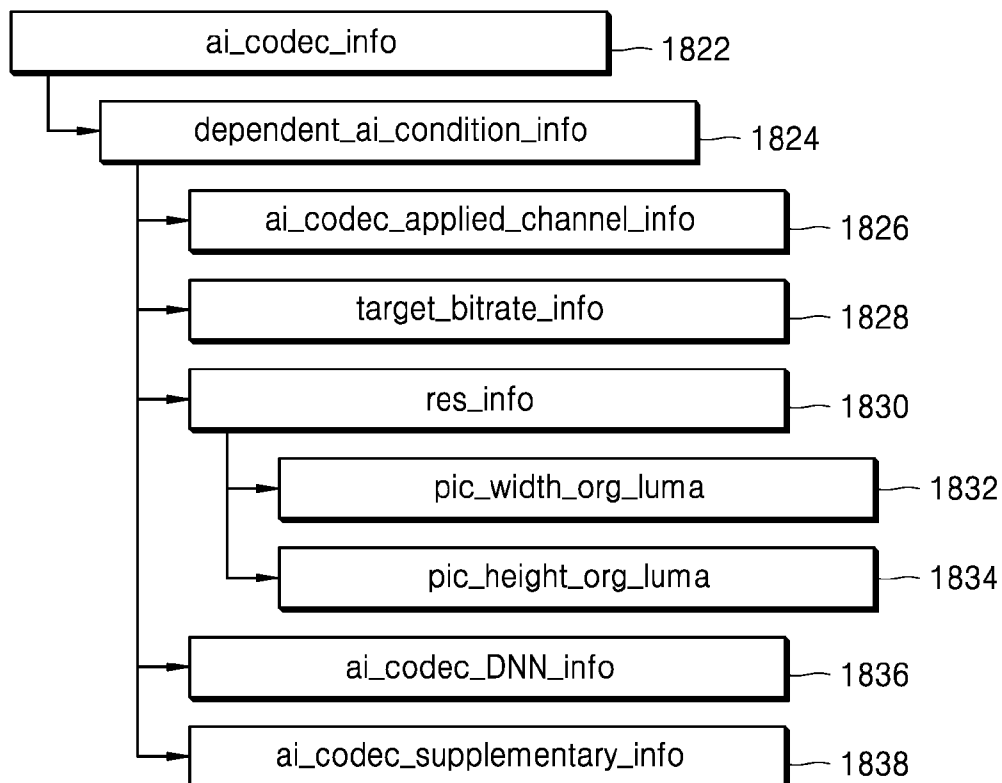
[도17]



[도 18a]

1800

[도 18b]

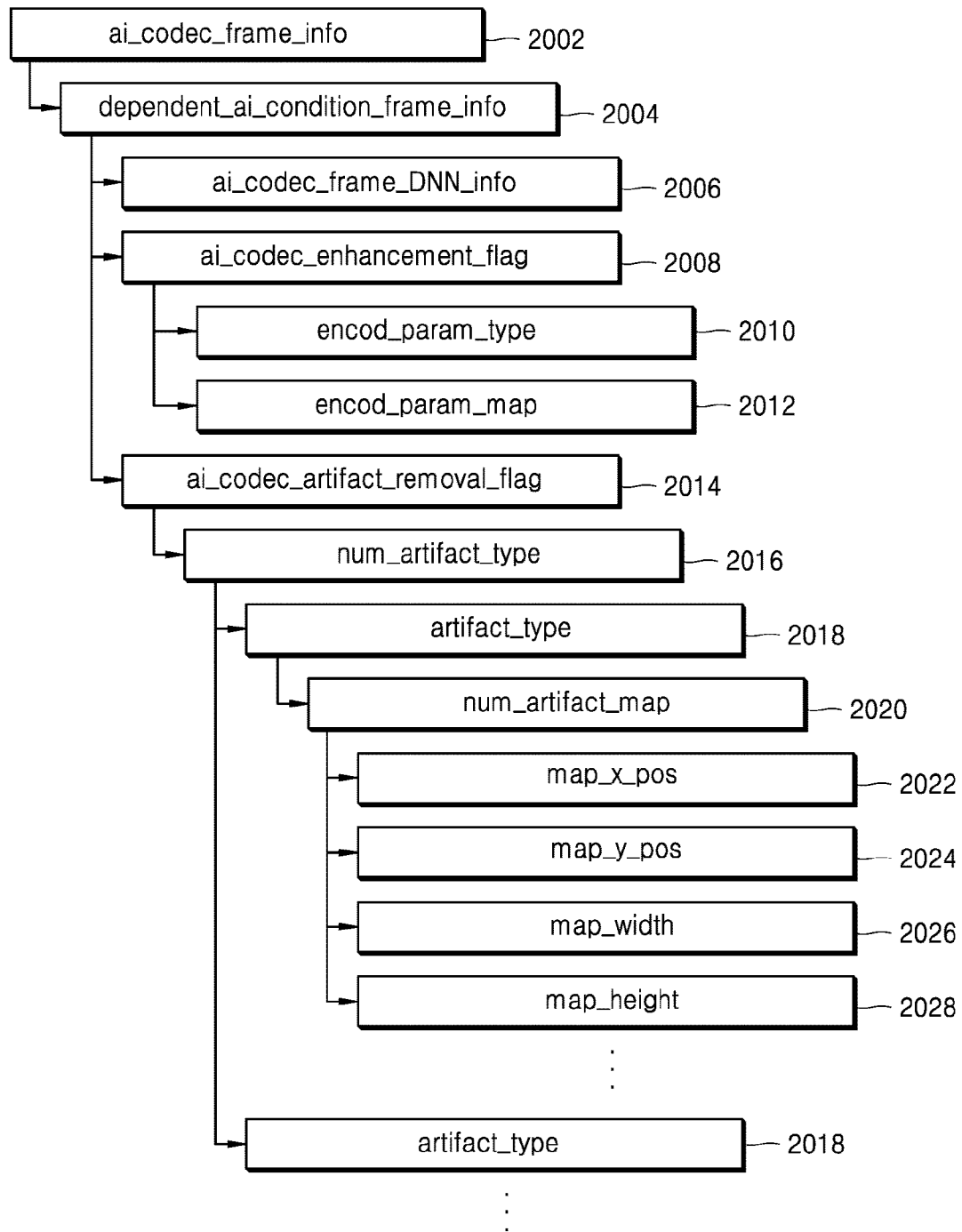
1820

[도 19]

1900

|                                        |            |
|----------------------------------------|------------|
| ai_codec_usage_main(payloadSize) {     | Descriptor |
| ai_codec_info {                        | ue(v)      |
| if(ai_codec_info) {                    |            |
| ai_codec_applied_channel_info          | ue(v)      |
| target_bitrate                         | ue(v)      |
| pic_width_in_org_luma_samples          | ue(v)      |
| pic_height_in_org_luma_samples         | ue(v)      |
| ai_codec_DNN_info                      | ue(v)      |
| ai_codec_supplementary_info_flag       | u(1)       |
| if(ai_codec_supplementary_info_flag) { |            |
| genre_info                             | ue(v)      |
| hdr_max_luminance                      | ue(v)      |
| hdr_color_gamut                        | ue(v)      |
| hdr_pq_type                            | ue(v)      |
| rate_control_type                      | ue(v)      |
| codec_type                             | ue(v)      |
| }                                      |            |
| }                                      |            |
| }                                      |            |
| }                                      |            |

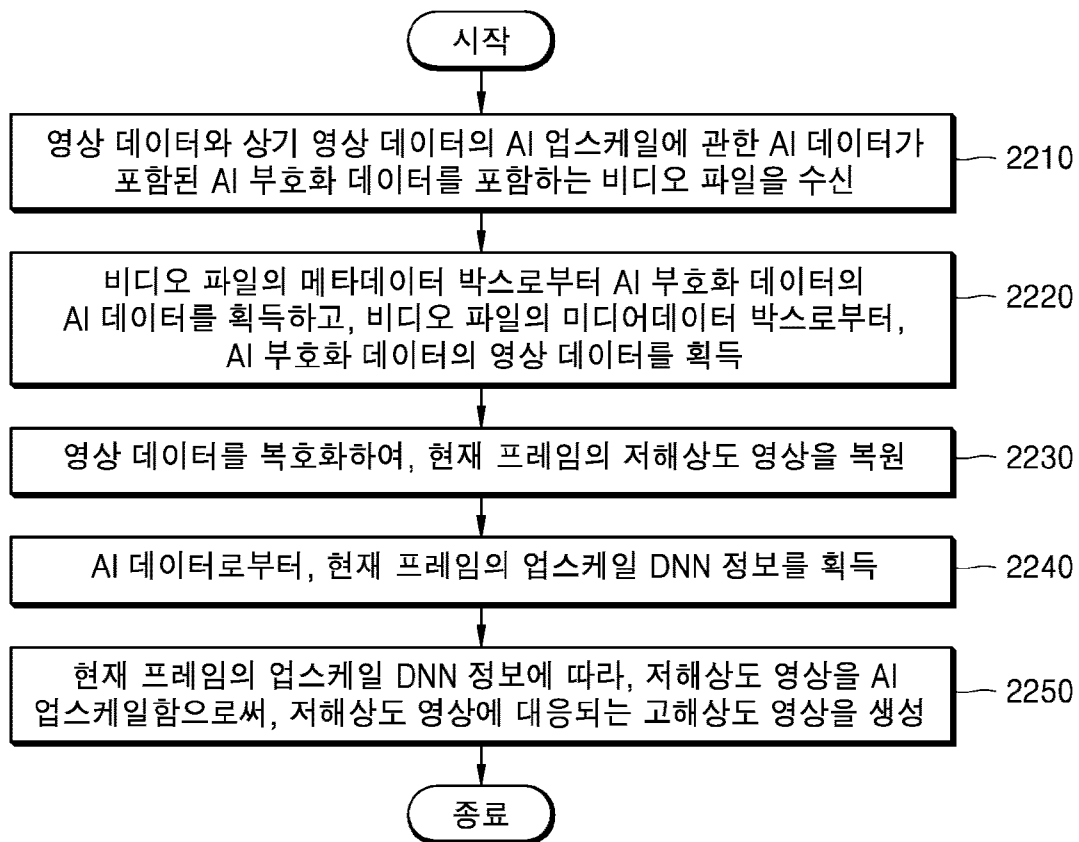
[도20]

2000

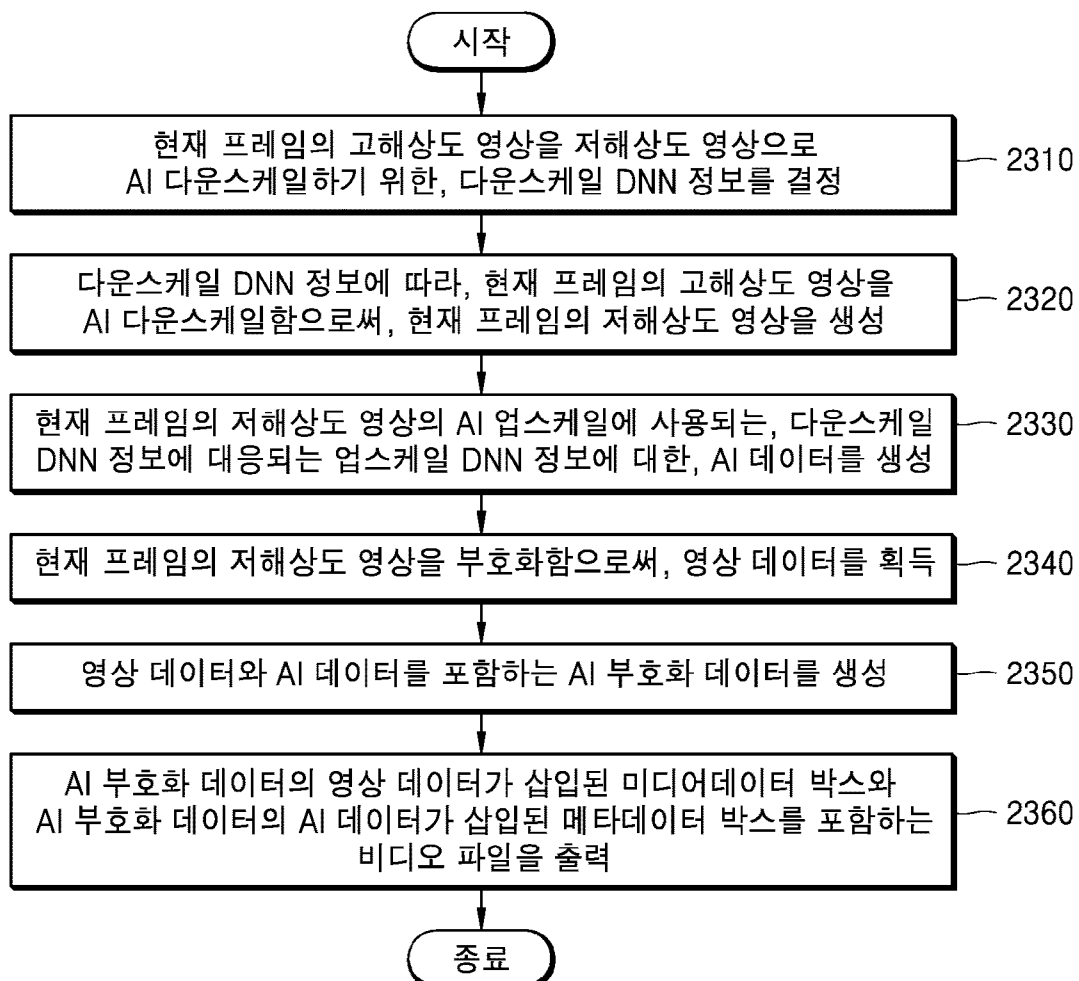
[도21]

|                                      |            |
|--------------------------------------|------------|
| ai_codec_usage_main(payloadSize) {   | Descriptor |
| ai_codec_frame_info                  | ue(v)      |
| if (ai_codec_frame_info) {           |            |
| ai_codec_frame_DNN_info              | ue(v)      |
| ai_codec_enhancement_flag            | ue(1)      |
| if(ai_codec_enhancement_flag) {      |            |
| encod_param_type                     | ue(v)      |
| encod_param_map                      | ue(v)      |
| }                                    |            |
| ai_codec_artifact_removal_flag{      | u(1)       |
| if(ai_codec_artifact_removal_flag) { |            |
| num_artifact_type                    | ue(v)      |
| for(i=0; i<num_artifact_type; i++) { |            |
| artifact_type                        | ue(v)      |
| num_artifact_map                     | ue(v)      |
| for(i=0; i<num_artifact_map; i++) {  |            |
| map_x_pos                            | ue(v)      |
| map_y_pos                            | ue(v)      |
| map_width                            | ue(v)      |
| map_height                           | ue(v)      |
| }                                    |            |
| }                                    |            |
| }                                    |            |
| }                                    |            |

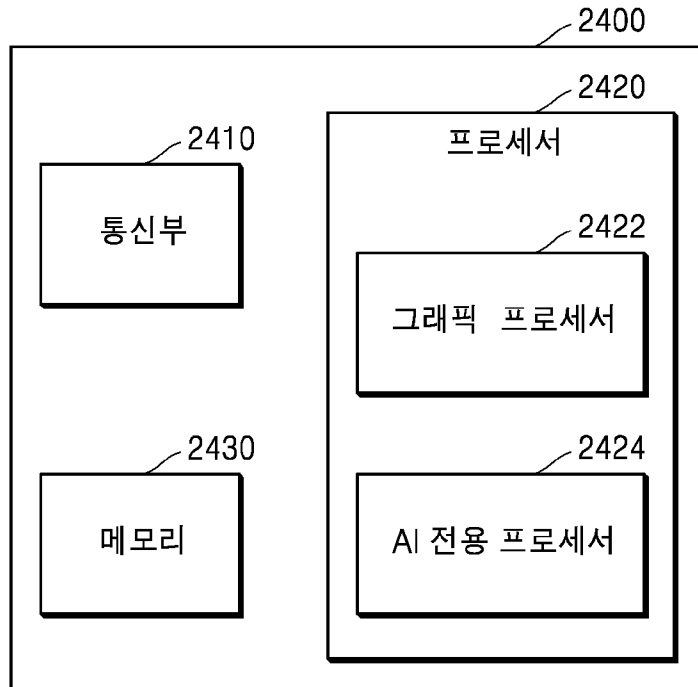
[도22]



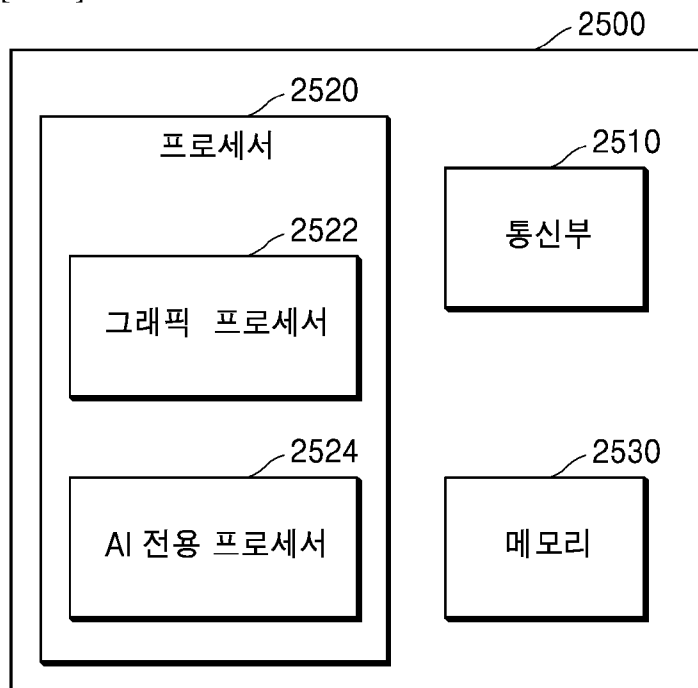
[도23]



[도24]



[도25]



## INTERNATIONAL SEARCH REPORT

International application No.

**PCT/KR2019/013344****A. CLASSIFICATION OF SUBJECT MATTER*****H04N 19/132(2014.01)i, H04N 19/85(2014.01)i, H04N 19/50(2014.01)i, H04N 19/184(2014.01)i, G06N 3/08(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)

H04N 19/132; G06N 3/04; G06T 3/40; H04N 19/124; H04N 19/176; H04N 19/85; H04N 19/90; H04N 21/2343; H04N 7/01; H04N 19/50; H04N 19/184; G06N 3/08

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) &amp; Keywords: encoding, down-scaling, decoding, up-scaling, neural network, metadata

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

| Category* | Citation of document, with indication, where appropriate, of the relevant passages                                                                   | Relevant to claim No. |
|-----------|------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| Y         | US 2014-0177706 A1 (SAMSUNG ELECTRONICS CO., LTD.) 26 June 2014<br>See paragraphs [0005], [0008], [0081], [0088]-[0089], [0100]; and claims 6-8, 14. | 1-22                  |
| Y         | US 2017-0347061 A1 (MAGIC PONY TECHNOLOGY LIMITED) 30 November 2017<br>See paragraph [0531].                                                         | 1-22                  |
| A         | KR 10-1885855 B1 (INDUSTRY-ACADEMIC COOPERATION FOUNDATION, DANKOOK UNIVERSITY) 07 August 2018<br>See paragraphs [0005]-[0012].                      | 1-22                  |
| A         | KR 10-2018-0108288 A (NCSOFT CORPORATION et al.) 04 October 2018<br>See paragraphs [0005]-[0019].                                                    | 1-22                  |
| A         | KR 10-2018-0100976 A (ELECTRONICS AND TELECOMMUNICATIONS RESEARCH INSTITUTE) 12 September 2018<br>See paragraphs [0009]-[0018].                      | 1-22                  |



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

“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

“&amp;” document member of the same patent family

Date of the actual completion of the international search

**22 JANUARY 2020 (22.01.2020)**

Date of mailing of the international search report

**22 JANUARY 2020 (22.01.2020)**

Name and mailing address of the ISA/KR

Korean Intellectual Property Office  
Government Complex Daejeon Building 4, 189, Cheongsu-ro, Seo-gu,  
Daejeon, 35208, Republic of Korea

Facsimile No. +82-42-481-8578

Authorized officer

Telephone No.

**INTERNATIONAL SEARCH REPORT**  
Information on patent family members

International application No.

**PCT/KR2019/013344**

| Patent document<br>cited in search report | Publication<br>date | Patent family<br>member | Publication<br>date |
|-------------------------------------------|---------------------|-------------------------|---------------------|
| US 2014-0177706 A1                        | 26/06/2014          | None                    |                     |
| US 2017-0347061 A1                        | 30/11/2017          | EP 3259910 A1           | 27/12/2017          |
|                                           |                     | EP 3278559 A1           | 07/02/2018          |
|                                           |                     | EP 3298575 A1           | 28/03/2018          |
|                                           |                     | EP 3493149 A1           | 05/06/2019          |
|                                           |                     | US 2018-0122048 A1      | 03/05/2018          |
|                                           |                     | US 2018-0130180 A1      | 10/05/2018          |
|                                           |                     | US 2018-0139458 A1      | 17/05/2018          |
|                                           |                     | WO 2016-132154 A1       | 25/08/2016          |
|                                           |                     | WO 2016-156864 A1       | 06/10/2016          |
|                                           |                     | WO 2017-144881 A1       | 31/08/2017          |
|                                           |                     | WO 2017-158363 A1       | 21/09/2017          |
| KR 10-1885855 B1                          | 07/08/2018          | None                    |                     |
| KR 10-2018-0108288 A                      | 04/10/2018          | KR 10-1990092 B1        | 17/06/2019          |
| KR 10-2018-0100976 A                      | 12/09/2018          | None                    |                     |

**A. 발명이 속하는 기술분류(국제특허분류(IPC))**

H04N 19/132(2014.01)i, H04N 19/85(2014.01)i, H04N 19/50(2014.01)i, H04N 19/184(2014.01)i, G06N 3/08(2006.01)i

**B. 조사된 분야**

조사된 최소문헌(국제특허분류를 기재)

H04N 19/132; G06N 3/04; G06T 3/40; H04N 19/124; H04N 19/176; H04N 19/85; H04N 19/90; H04N 21/2343; H04N 7/01; H04N 19/50; H04N 19/184; G06N 3/08

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) &amp; 키워드: 부호화(encoding), 다운스케일링(down-scaling), 복호화(decoding), 업스케일링(up-scaling), 신경망(neural network), 메타데이터(metadata)

**C. 관련 문헌**

| 카테고리* | 인용문헌명 및 관련 구절(해당하는 경우)의 기재                                                                                                           | 관련 청구항 |
|-------|--------------------------------------------------------------------------------------------------------------------------------------|--------|
| Y     | US 2014-0177706 A1 (SAMSUNG ELECTRONICS CO., LTD.) 2014.06.26<br>단락 [0005], [0008], [0081], [0088]-[0089], [0100]; 및 청구항 6-8, 14 참조. | 1-22   |
| Y     | US 2017-0347061 A1 (MAGIC PONY TECHNOLOGY LIMITED) 2017.11.30<br>단락 [0531] 참조.                                                       | 1-22   |
| A     | KR 10-1885855 B1 (단국대학교 산학협력단) 2018.08.07<br>단락 [0005]-[0012] 참조.                                                                    | 1-22   |
| A     | KR 10-2018-0108288 A (주식회사 엔씨소프트 등) 2018.10.04<br>단락 [0005]-[0019] 참조.                                                               | 1-22   |
| A     | KR 10-2018-0100976 A (한국전자통신연구원) 2018.09.12<br>단락 [0009]-[0018] 참조.                                                                  | 1-22   |

☐ 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 이후 “X”에 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.

“&amp;” 동일한 대응특허문헌에 속하는 문헌

국제조사의 실제 완료일

2020년 01월 22일 (22.01.2020)

국제조사보고서 발송일

2020년 01월 22일 (22.01.2020)

ISA/KR의 명칭 및 우편주소

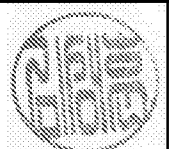
대한민국 특허청  
(35208) 대전광역시 서구 청사로 189,  
4동 (둔산동, 정부대전청사)

팩스 번호 +82-42-481-8578

심사관

김성훈

전화번호 +82-42-481-8710



| 국제조사보고서에서<br>인용된 특허문헌 | 공개일        | 대응특허문헌             | 공개일        |
|-----------------------|------------|--------------------|------------|
| US 2014-0177706 A1    | 2014/06/26 | 없음                 |            |
| US 2017-0347061 A1    | 2017/11/30 | EP 3259910 A1      | 2017/12/27 |
|                       |            | EP 3278559 A1      | 2018/02/07 |
|                       |            | EP 3298575 A1      | 2018/03/28 |
|                       |            | EP 3493149 A1      | 2019/06/05 |
|                       |            | US 2018-0122048 A1 | 2018/05/03 |
|                       |            | US 2018-0130180 A1 | 2018/05/10 |
|                       |            | US 2018-0139458 A1 | 2018/05/17 |
|                       |            | WO 2016-132154 A1  | 2016/08/25 |
|                       |            | WO 2016-156864 A1  | 2016/10/06 |
|                       |            | WO 2017-144881 A1  | 2017/08/31 |
|                       |            | WO 2017-158363 A1  | 2017/09/21 |
| KR 10-1885855 B1      | 2018/08/07 | 없음                 |            |
| KR 10-2018-0108288 A  | 2018/10/04 | KR 10-1990092 B1   | 2019/06/17 |
| KR 10-2018-0100976 A  | 2018/09/12 | 없음                 |            |