(12) 특허협력조약에 의하여 공개된 국제출원

(19) 세계지식재산권기구
국제사무국

(43) 국제공개일
2022 년 5 월 12 일 (12.05.2022)

WIPO | PCT

(10) 국제공개번호
WO 2022/098207 A1

(51) 국제특허분류:
H04N 19/86 (2014.01) *H04N 19/124* (2014.01)
H04N 19/82 (2014.01) *G06T 9/00* (2006.01)
H04N 19/423 (2014.01) *G06N 3/08* (2006.01)

(21) 국제출원번호: PCT/KR2021/016236

(22) 국제출원일: 2021 년 11 월 9 일 (09.11.2021)

(25) 출원언어: 한국어

(26) 공개언어: 한국어

(30) 우선권정보:
10-2020-0148914 2020 년 11 월 9 일 (09.11.2020) KR
10-2020-0177841 2020 년 12 월 17 일 (17.12.2020) KR

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

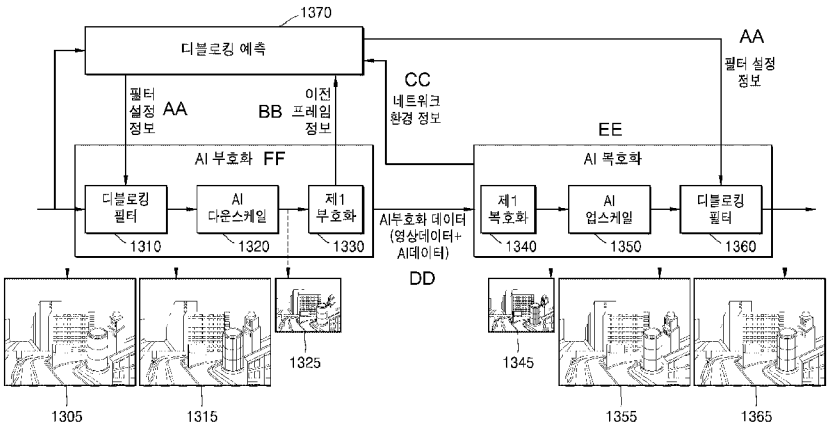
(72) 발명자: **양희철 (YANG, Heechul)**; 16677 경기도 수원시 영통구 삼성로 129, Gyeonggi-do (KR). **나인학 (NA, Inhak)**; 16677 경기도 수원시 영통구 삼성로 129, Gyeonggi-do (KR). **정현권 (CHUNG, Hyunkwon)**; 16677 경기도 수원시 영통구 삼성로 129, Gyeonggi-do (KR).

(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, IT, JO, JP, KE, KG, KH, KN, KP, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA,

(54) **Title:** METHOD AND DEVICE FOR PERFORMING ARTIFICIAL INTELLIGENCE ENCODING AND ARTIFICIAL INTELLIGENCE DECODING

(54) **발명의 명칭:** 인공지능 부호화 및 인공지능 복호화를 수행하기 위한 방법 및 장치



1310, 1360 ... Deblocking filter
1320 ... AI downscaling
1330 ... First encoding
1340 ... First decoding
1350 ... AI upscaling
1370 ... Deblocking prediction

AA ... Filter configuration information
BB ... Previous frame information
CC ... Network environment information
DD ... AI encoding data (image data + AI data)
EE ... AI decoding
FF ... AI encoding

(57) **Abstract:** Proposed is an AI encoding device comprising a memory and a processor, wherein the memory stores instructions executable by the processor, and the processor obtains an original image, previously encoded frame information, and network environment information, obtains deblocking filter configuration information on the basis of the original image, the previously encoded frame information, and the network environment information, applies deblocking filtering to the original image on the basis of the deblocking filter configuration information and obtains an AI-downscaled first image via a DNN for downscaling, performs first encoding of the first image to generate image data, and transmits the deblocking filter configuration information, AI data including information related to the AI downscaling, and the image data.

WO 2022/098207 A1

[다음 쪽 계속]

PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(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: 메모리; 및 프로세서를 포함하고, 상기 메모리는 상기 프로세서에 의해 실행 가능한 명령어들을 저장하고 상기 프로세서는, 원본 영상, 이전에 부호화된 프레임 정보, 및 네트워크 환경 정보를 획득하고, 상기 원본 영상, 상기 이전에 부호화된 프레임 정보, 및 상기 네트워크 환경 정보에 기초하여 디블로킹 필터 설정 정보를 획득하고, 상기 디블로킹 필터 설정 정보에 기초하여 상기 원본 영상에 디블로킹 필터링을 적용하고 다운스케일용 DNN을 통해 AI 다운스케일된 제1 영상을 획득하고, 상기 제1 영상을 제1 부호화하여 영상 데이터를 생성하고, 상기 디블로킹 필터 설정 정보, 상기 AI 다운스케일과 관련된 정보를 포함하는 AI 데이터, 및 상기 영상 데이터를 전송하는 것을 특징으로하는, AI 부호화 장치를 제안한다.

명세서

발명의 명칭: 인공지능 부호화 및 인공지능 복호화를 수행하기 위한 방법 및 장치

기술분야

[1] 본 개시는 영상을 처리하기 위한 방법 및 장치에 관한 것으로, 보다 구체적으로, 본 개시는 인공지능(AI) 부호화 및 인공지능 복호화를 수행하기 위한 방법 및 장치에 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 AI 부호화 및 AI 복호화 방법, 및 장치는 실시간 스트리밍 서비스에서 가변적이고 제한적인 네트워크 환경에서 블로킹 아티팩트의 발생을 예측하여 디블로킹 필터링을 추가로 수행하고 AI 기반으로 다운스케일 및 업스케일하여, 영상을 부호화 및 복호화하는 것을 기술적 과제로 한다.

과제 해결 수단

[5] 상기 기술적 과제를 해결하기 위해 본 개시에서 제안하는 AI 복호화 장치는, 하나 이상의 인스트럭션들을 저장하는 메모리; 및 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행하는 프로세서를 포함하고, 상기 프로세서는, 제 1 영상의 제 1 부호화 결과 생성된 영상 데이터, 원본 영상으로부터 상기 제 1 영상으로의 AI 다운스케일과 관련된 AI 데이터, 및 디블로킹 필터 설정 정보를 획득하고, 상기 영상 데이터를 제 1 복호화하여 상기 제 1 영상에 대응하는 제 2 영상을 획득하고, 복수의 DNN 설정 정보 중 상기 제 2 영상의 AI 업스케일을 위한 DNN 설정 정보를, 상기 AI 데이터에 기초하여 획득하고, 상기 획득된 DNN 설정 정보로 동작하는 업스케일용 DNN(Deep Neural Network)을 통해 상기 제 2 영상으로부터 AI 업스케일되고 상기 디블로킹 필터 설정 정보에 기초하여 디블로킹 필터링이 적용된 복원 영상을 생성할 수 있다.

[6] 상기 기술적 과제를 해결하기 위해 본 개시에서 제안하는 AI 복호화 방법은, 제 1 영상의 제 1 부호화 결과 생성된 영상 데이터, 원본 영상으로부터 상기 제 1 영상으로의 AI 다운스케일과 관련된 AI 데이터, 및 디블로킹 필터 설정 정보를

획득하는 단계; 상기 영상 데이터를 제 1 복호화하여 상기 제 1 영상에 대응하는 제 2 영상을 획득하는 단계; 복수의 DNN 설정 정보 중 상기 제 2 영상의 AI 업스케일을 위한 DNN 설정 정보를, 상기 AI 데이터에 기초하여 획득하는 단계; 및 상기 획득된 DNN 설정 정보로 동작하는 업스케일용 DNN(Deep Neural Network)을 통해 상기 제 2 영상으로부터 AI 업스케일되고 상기 디블로킹 필터 설정 정보에 기초하여 디블로킹 필터링이 적용된 복원 영상을 생성하는 단계를 포함할 수 있다.

[7] 상기 기술적 과제를 해결하기 위해 본 개시에서 제안하는 AI 부호화 장치는, 하나 이상의 인스트럭션들을 저장하는 메모리; 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행하는 프로세서를 포함하고, 상기 프로세서는, 원본 영상, 이전에 부호화된 프레임 정보, 및 네트워크 환경 정보를 획득하고, 상기 원본 영상, 상기 이전에 부호화된 프레임 정보, 및 상기 네트워크 환경 정보에 기초하여 디블로킹 필터 설정 정보를 획득하고, 상기 디블로킹 필터 설정 정보에 기초하여 상기 원본 영상에 디블로킹 필터링을 적용하고 다운스케일용 DNN을 통해 AI 다운스케일된 제1 영상을 획득하고, 상기 제1 영상을 제1 부호화하여 영상 데이터를 생성하고, 상기 디블로킹 필터 설정 정보, 상기 AI 다운스케일과 관련된 정보를 포함하는 AI 데이터, 및 상기 영상 데이터를 전송할 수 있다.

[8] 상기 기술적 과제를 해결하기 위해 본 개시에서 제안하는 AI 부호화 방법은, 원본 영상, 이전에 부호화된 프레임 정보, 및 네트워크 환경 정보를 획득하는 단계; 상기 원본 영상, 상기 이전에 부호화된 프레임 정보, 및 상기 네트워크 환경 정보에 기초하여 디블로킹 필터 설정 정보를 획득하는 단계; 상기 디블로킹 필터 설정 정보에 기초하여 상기 원본 영상에 디블로킹 필터링을 적용하고 다운스케일용 DNN을 통해 AI 다운스케일된 제1 영상을 획득하는 단계; 상기 제1 영상을 제1 부호화하여 영상 데이터를 생성하는 단계; 상기 디블로킹 필터 설정 정보, 상기 AI 다운스케일과 관련된 정보를 포함하는 AI 데이터, 및 상기 영상 데이터를 전송하는 단계를 포함할 수 있다.

발명의 효과

[9] 부호화시에 원본 영상, 이전 프레임 정보, 및 네트워크 환경 정보에 기초하여 디블로킹 필터 설정 정보를 획득하고, 디블로킹 필터 설정 정보에 기초하여 원본 영상에 디블로킹 필터링을 적용하고, 디블로킹 필터링된 원본 영상을 AI 다운스케일하여 다운스케일된 영상과 디블로킹 필터 설정 정보를 전송하고, 복호화시에 전송된 다운스케일된 영상을 AI 업스케일하여 업스케일된 영상을 획득하고, 전송된 디블로킹 필터 설정 정보에 기초하여 디블로킹 필터링을 업스케일된 영상에 적용하여 복원 영상을 획득함으로써, 네트워크 상황이 갑자기 변하는 경우에도 AI에 기반하여 고해상도의 영상이 효과적으로 부호화 및 복호화될 수 있다.

도면의 간단한 설명

[10] 본 명세서에서 인용되는 도면을 보다 충분히 이해하기 위하여 각 도면의 간단한 설명이 제공된다.

[11] 도 1은 일 실시예에 따른 AI 부호화 과정 및 AI 복호화 과정을 설명하기 위한 도면이다.

[12] 도 2는 일 실시예에 따른 AI 복호화 장치의 구성을 나타내는 블록도이다.

[13] 도 3은 제 2 영상의 AI 업스케일을 위한 제 2 DNN을 나타내는 예시적인 도면이다.

[14] 도 4는 컨볼루션 레이어에 의한 컨볼루션 연산을 설명하기 위한 도면이다.

[15] 도 5는 여러 영상 관련 정보들과 여러 DNN 설정 정보들 사이의 매핑 관계를 나타내는 예시적인 도면이다.

[16] 도 6은 복수의 프레임으로 구성된 제 2 영상을 도시하는 도면이다.

[17] 도 7은 일 실시예에 따른 AI 부호화 장치의 구성을 나타내는 블록도이다.

[18] 도 8은 원본 영상의 AI 다운스케일을 위한 제 1 DNN을 나타내는 예시적인 도면이다.

[19] 도 9는 일 실시예에 따른 AI 부호화 데이터의 구조를 나타내는 도면이다.

[20] 도 10은 다른 실시예에 따른 AI 부호화 데이터의 구조를 나타내는 도면이다.

[21] 도 11은 제 1 DNN 및 제 2 DNN을 훈련시키는 방법을 설명하기 위한 도면이다.

[22] 도 12는 훈련 장치에 의한 제 1 DNN 및 제 2 DNN의 훈련 과정을 설명하기 위한 도면이다.

[23] 도 13은 일 실시예에 따른 디블로킹 필터링을 추가하는 AI 부호화 과정 및 AI 복호화 과정을 설명하기 위한 도면이다.

[24] 도 14는 일 실시예에 따른 AI 복호화 장치의 구성을 나타내는 블록도이다.

[25] 도 15는 일 실시예에 따른 AI 부호화 장치의 구성을 나타내는 블록도이다.

[26] 도 16은 다른 실시예에 따른 디블로킹 필터링을 적응적으로 추가하는 AI 부호화 과정 및 AI 복호화 과정을 설명하기 위한 도면이다.

[27] 도 17은 다른 실시예에 따른 디블로킹 필터링을 적응적으로 추가하는 AI 부호화 과정 및 AI 복호화 과정을 설명하기 위한 도면이다.

[28] 도 18은 다른 실시예에 따른 디블로킹 필터링을 적응적으로 추가하는 AI 부호화 과정 및 AI 복호화 과정을 설명하기 위한 도면이다.

[29] 도 19는 AI에 기반한 디블로킹 필터를 위한 제 3 DNN 및 제 4 DNN을 나타내는 예시적인 도면이다.

[30] 도 20은 제 3 DNN(2010) 및 제 4 DNN(2020)을 훈련시키는 방법을 설명하기 위한 도면이다.

[31] 도 21는 훈련 장치(2100)에 의한 제 3 DNN(2010) 및 제 4 DNN(2020)의 훈련 과정을 설명하기 위한 도면이다.

[32] 도 22는 일 실시예에 따른 AI 부호화 방법을 설명하기 위한 순서도이다.

[33] 도 23은 일 실시예에 따른 AI 복호화 방법을 설명하기 위한 순서도이다.

발명의 실시를 위한 형태

[34] 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.

[35] 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.

[36] 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.

[37] 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.

[38] 또한, 본 명세서에서, '영상(image)' 또는 '픽처'는 정지영상, 복수의 연속된 정지영상(또는 프레임)으로 구성된 동영상, 또는 비디오를 나타낼 수 있다.

[39] 또한, 본 명세서에서 'DNN(deep neural network)'은 뇌 신경을 모사한 인공신경망 모델의 대표적인 예시로써, 특정 알고리즘을 사용한 인공신경망 모델로 한정되지 않는다.

[40] 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.

[41] 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.

[42] 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.

[43] 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.

[44] 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.

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

[46] 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.

[47] 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 비트레이트를 크게 감소시킬 수 있다.

[48] 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.

[49] AI 부호화 과정을 좀 더 상세히 살펴보면, 원본 영상(105)을 입력받으면, 소정 해상도 및/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.

[50] 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.

[51] AI 다운스케일(110)을 위한 AI 및 AI 업스케일(140)을 위한 AI는 DNN(deep neural network)으로 구현될 수 있다. 도 11을 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 화질 및/또는 해상도로 AI 업스케일(140)할 수 있다.

[52] 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 방법들 중의 하나를 통해 구현될 수 있다.

[53] 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 방법들 중의 하나에 대응되는 영상 복원 방법을 통해 구현될 수 있다.

[54] 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.

[55] 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.

[56] 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.

[57] AI 데이터는 비트스트림의 형태로 영상 데이터와 함께 전송될 수 있다. 구현예에 따라, AI 데이터는 프레임이나 패킷 형태로 영상 데이터와 구분되어 전송될 수도 있다.

[58] 또는 구현예에 따라, AI 데이터는 영상 데이터에 포함되어 전송될 수도 있다.

[59] 영상 데이터와 AI 데이터는 동일한 네트워크 또는 서로 상이한 네트워크를 통해 전송될 수 있다.

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

[61] 도 2를 참조하면, 일 실시예에 따른 AI 복호화 장치(200)는 수신부(210) 및 AI 복호화부(230)를 포함한다. AI 복호화부(230)는 파싱부(232), 제 1 복호화부(234), AI 업스케일부(236) 및 AI 설정부(238)를 포함할 수 있다.

[62] 도 2에는 수신부(210) 및 AI 복호화부(230)가 개별적인 장치로 도시되어 있으나, 수신부(210) 및 AI 복호화부(230)는 하나의 프로세서를 통해 구현될 수 있다. 이 경우, 수신부(210) 및 AI 복호화부(230)는 전용 프로세서로 구현될 수도 있고, AP(application processor) 또는 CPU(central processing unit), GPU(graphic processing unit)와 같은 범용 프로세서와 S/W의 조합을 통해 구현될 수도 있다. 또한, 전용 프로세서의 경우, 본 개시의 실시예를 구현하기 위한 메모리를 포함하거나, 외부 메모리를 이용하기 위한 메모리 처리부를 포함할 수 있다.

[63] 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. 일 실시예에서, 수신부(210)는 제 1 프로세서로 구현되고, 제 1 복호화부(234)는 제 1 프로세서와 상이한 제 2 프로세서로 구현되고, 파싱부(232), AI 업스케일부(236) 및 AI 설정부(238)는 제 1 프로세서 및 제 2 프로세서와 상이한 제 3 프로세서로 구현될 수 있다.

[64] 수신부(210)는 AI 부호화 결과 획득된 AI 부호화 데이터를 수신한다. 일 예로, AI 부호화 데이터는 mp4, mov 등의 파일 형식을 갖는 비디오 파일일 수 있다.

[65] 수신부(210)는 네트워크를 통해 전달되는 AI 부호화 데이터를 수신할 수 있다. 수신부(210)는 AI 부호화 데이터를 AI 복호화부(230)로 출력한다.

[66] 일 실시예에서, 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.

[67] 파싱부(232)는 AI 부호화 데이터를 파싱하여 제 1 영상(115)의 제 1 부호화 결과로 생성된 영상 데이터를 제 1 복호화부(234)로 전달하고, AI 데이터를 AI 설정부(238)로 전달한다.

[68] 일 실시예에서, 파싱부(232)는 AI 부호화 데이터 내에 서로 분리되어 포함된 영상 데이터와 AI 데이터를 파싱할 수 있다. 파싱부(232)는 AI 부호화 데이터 내의 헤더를 읽어 AI 부호화 데이터 내에 포함되어 있는 AI 데이터와 영상 데이터를 구분할 수 있다. 일 예에서, AI 데이터는 HDMI 스트림 내의 VSIF(Vendor Specific InfoFrame)에 포함될 수 있다.

[69] 서로 분리된 AI 데이터와 영상 데이터를 포함하는 AI 부호화 데이터의 구조는 도 9를 참조하여 후술한다.

[70] 다른 실시예에서, 파싱부(232)는 AI 부호화 데이터에서 영상 데이터를 파싱하고, 영상 데이터로부터 AI 데이터를 추출한 후, AI 데이터를 AI 설정부(238)로 전달하고, 나머지 영상 데이터를 제 1 복호화부(234)로 전달할 수 있다. 즉, AI 데이터는 영상 데이터에 포함될 수 있는데, 예를 들어, AI 데이터는 영상 데이터에 해당하는 비트스트림의 부가 정보 영역인 SEI(Supplemental enhancement information)에 포함될 수 있다. AI 데이터가 포함된 영상 데이터를 포함하는 AI 부호화 데이터의 구조는 도 10을 참조하여 후술한다.

[71] 다른 실시예에서, 파싱부(232)는 영상 데이터에 해당하는 비트스트림을 제 1 복호화부(234)에서 처리될 비트스트림과 AI 데이터에 해당하는 비트스트림으로 분할하고, 분할된 각각의 비트스트림을 제 1 복호화부(234)와 AI 설정부(238)로 출력할 수 있다.

[72] 파싱부(232)는 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, 해당 정보를 제 1 복호화부(234)로 전달할 수 있다.

[73] 제 1 복호화부(234)는 파싱부(232)로부터 수신된 영상 데이터에 기초하여 제 1 영상(115)에 대응하는 제 2 영상(135)을 복원한다. 제 1 복호화부(234)에 의해 획득된 제 2 영상(135)은 AI 업스케일부(236)로 제공된다.

[74] 구현예에 따라, 예측 모드 정보, 움직임 정보, 양자화 파라미터 정보 등의 제 1 복호화 관련 정보가 제 1 복호화부(234)로부터 AI 설정부(238)로 제공될 수 있다. 제 1 복호화 관련 정보는 DNN 설정 정보를 획득하는데 이용될 수 있다.

[75] AI 설정부(238)로 제공되는 AI 데이터는, 제 2 영상(135)을 AI 업스케일할 수 있게 하는 정보들을 포함한다. 이때, 제 2 영상(135)의 업스케일 타겟은 제 1 DNN의 다운스케일 타겟에 대응하여야 한다. 따라서, AI 데이터는 제 1 DNN의

다운스케일 타겟을 확인할 수 있는 정보가 포함되어야 한다.

[76] 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.

[77] 차이 정보는, 원본 영상(105) 대비 제 1 영상(115)의 해상도 변환 정도에 대한 정보(예를 들어, 해상도 변환율 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.

[78] 그리고, 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.

[79] AI 설정부(238)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(236)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.

[80] AI 설정부(238)가 AI 데이터에 기초하여 업스케일 타겟을 결정하는 방법을 설명하기에 앞서, 제 2 DNN을 통한 AI 업스케일 과정에 대해 도 3 및 도 4를 참조하여 설명한다.

[81] 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.

[82] 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.

[83] 도 4를 참조하여, 제 1 컨볼루션 레이어(310)에서의 컨볼루션 연산에 대해 상세히 설명한다.

[84] 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.

[85] 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.

[86] 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.

[87] 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.

[88] 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.

[89] 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. AI 설정부(238)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.

[90] 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.

[91] 다시 도 3을 참조하면, 제 1 컨볼루션 레이어(310)에서 출력된 특징 맵들은 제 1 활성화 레이어(320)로 입력된다.

[92] 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.

[93] 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.

[94] 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.

[95] 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.

[96] 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.

[97] 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 컨벌루션 연산을 통해 제 3 영상(145)을 출력할 수 있다.

[98] 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.

[99] 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.

[100] 일 실시예에서, AI 업스케일부(236)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.

[101] 이하에서, AI 설정부(238)가 업스케일 타겟을 결정하고, AI 업스케일부(236)가 업스케일 타겟에 맞춰 제 2 영상(135)을 AI 업스케일하는 방법을 설명한다.

[102] 일 실시예에서, AI 설정부(238)는 제 2 DNN에 세팅 가능한 복수의 DNN 설정 정보를 저장할 수 있다.

[103] 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.

[104] 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.

[105] AI 설정부(238)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.

[106] 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.

[107] 복수의 DNN 설정 정보 각각은 AI 부호화 장치(700)의 제 1 DNN의 DNN 설정 정보와 연계되어 만들어진 것이며, AI 설정부(238)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 설정부(238)는 제 1 DNN의 정보를 확인하여야 한다. 제 1 DNN의 정보를 AI 설정부(238)가 확인하기 위해,

일 실시예에 따른 AI 복호화 장치(200)는 AI 부호화 장치(700)로부터 제 1 DNN의 정보를 포함하는 AI 데이터를 수신한다.

[108] 다시 말하면, AI 설정부(238)는 AI 부호화 장치(700)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.

[109] 복수의 DNN 설정 정보 중 제 2 영상(135)의 AI 업스케일을 위한 DNN 설정 정보가 획득되면, DNN 설정 정보는 AI 업스케일부(236)로 전달되고, DNN 설정 정보에 따라 동작하는 제 2 DNN에 기초하여 입력 데이터가 처리될 수 있다.

[110] 예를 들어, AI 업스케일부(236)는 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.

[111] 구체적으로, AI 업스케일부(236)는 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}이고, DNN 설정 정보의 변경이 있는 경우, 필터 커널의 파라미터들을 변경된 DNN 설정 정보에 포함된 파라미터들인 {2, 2, 2, 2, 2, 2, 2, 2, 2}로 교체할 수 있는 것이다.

[112] AI 설정부(238)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.

[113] 일 실시예에서, AI 설정부(238)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 설정부(238)는 제 2 영상(135)의 해상도를 2배 증가시킬 수 있는 DNN 설정 정보를 획득할 수 있다.

[114] 다른 실시예에서, AI 설정부(238)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 설정부(238)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.

[115] 도 5는 여러 영상 관련 정보들과 여러 DNN 설정 정보들 사이의 매핑 관계를 나타내는 예시적인 도면이다.

[116] 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 11 참조).

[117] 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.

[118] 즉, AI 설정부(238)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.

[119] 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 설정부(238)는 복수의 DNN 설정 정보 중 A DNN 설정 정보를 획득할 수 있다.

[120] 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 설정부(238)는 복수의 DNN 설정 정보 중 B DNN 설정 정보를 획득할 수 있다.

[121] 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 설정부(238)는 복수의 DNN 설정 정보 중 C DNN 설정 정보를 획득하고, 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 설정부(238)는 복수의 DNN 설정 정보 중 D DNN 설정 정보를 획득할 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 설정부(238)는 제 2 영상(135)의 화질을 나타내는, 영상 데이터의 비트레이트에 따라 DNN 설정 정보를 획득할 수 있다.

[122] 또 다른 실시예에서, AI 설정부(238)는 제 1 복호화부(234)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 설정부(238)는 제 1 복호화부(234)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(234)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.

[123] 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 설정부(238)는 AI 데이터에 포함된 DNN 설정 정보의 식별자를 획득한 후, DNN 설정 정보의 식별자에 대응하는 DNN 설정 정보를 획득하고, AI 업스케일부(236)는 해당 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 설정부(238)는 복수의 DNN 설정 정보 중 AI 데이터에 포함된 식별자가 가리키는 DNN 설정 정보를 획득하고, AI 업스케일부(236)는 해당 DNN 설정 정보를 이용하여 제 2 영상(135)을 AI 업스케일할 수 있다.

[124] 또한, 구현예에 따라, AI 데이터는 DNN 설정 정보를 포함할 수도 있다. AI 설정부(238)는 AI 데이터에 포함된 DNN 설정 정보를 획득하고, AI 업스케일부(236)는 해당 DNN 설정 정보를 사용하여 제 2 영상(135)을 AI 업스케일할 수도 있다.

[125] 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 설정부(238)는 AI 데이터에 포함된 정보에 기초하여 룩업 테이블 값들 중에서 선택된 일부를 조합하여 DNN 설정 정보를 획득하고, AI 업스케일부(236)는 해당 DNN 설정 정보를 사용하여 제 2 영상(135)을 AI 업스케일할 수도 있다.

[126] 구현예에 따라, AI 설정부(238)는 업스케일 타겟에 대응되는 DNN의 구조가

결정되면, 결정된 DNN의 구조에 대응하는 DNN 설정 정보, 예를 들어, 필터 커널의 파라미터들을 획득할 수도 있다.

[127] 전술한 바와 같이, AI 설정부(238)는 제 1 DNN과 관련된 정보를 포함하는 AI 데이터를 통해 제 2 DNN의 DNN 설정 정보를 획득하고, AI 업스케일부(236)는 해당 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.

[128] 일 실시예에서, 제 2 영상(135)이 복수의 프레임으로 구성된 경우, AI 설정부(238)는 소정 개수의 프레임 별로 DNN 설정 정보를 독립적으로 획득할 수 있고, 또는, 전체 프레임에 대해 공통된 DNN 설정 정보를 획득할 수도 있다.

[129] 도 6은 복수의 프레임으로 구성된 제 2 영상(135)을 도시하는 도면이다.

[130] 도 6에 도시된 바와 같이, 제 2 영상(135)은 t0 내지 tn에 대응하는 프레임들로 이루어질 수 있다.

[131] 일 예시에서, AI 설정부(238)는 AI 데이터를 통해 제 2 DNN의 DNN 설정 정보를 획득하고, AI 업스케일부(236)는 해당 DNN 설정 정보에 기초하여 t0 내지 tn에 대응하는 프레임들을 AI 업스케일할 수 있다. 즉, t0 내지 tn에 대응하는 프레임들이 공통된 DNN 설정 정보에 기초하여 AI 업스케일될 수 있다.

[132] 다른 예시에서, AI 설정부(238)는 t0 내지 tn에 대응하는 프레임들 중 일부의 프레임들, 예를 들어, t0 내지 ta에 대응하는 프레임들에 대해 AI 데이터로부터 'A' DNN 설정 정보를 획득하고, ta+1 내지 tb에 대응하는 프레임들에 대해 AI 데이터로부터 'B' DNN 설정 정보획득할 수 있다. 또한, AI 설정부(238)는 tb+1 내지 tn에 대응하는 프레임들에 대해 AI 데이터로부터 'C' DNN 설정 정보획득할 수 있다. 다시 말하면, AI 설정부(238)는 복수의 프레임들 중 소정 개수의 프레임을 포함하는 그룹마다 DNN 설정 정보를 독립적으로 획득하고, AI 업스케일부(236)는그룹 각각에 포함된 프레임들을 독립적으로 획득한 DNN 설정 정보로 AI 업스케일할 수 있다.

[133] 또 다른 예시에서, AI 설정부(238)는 제 2 영상(135)을 구성하는 프레임별로 DNN 설정 정보를 독립적으로 획득할 수도 있다. 예를 들어, 제 2 영상(135)이 3개의 프레임으로 구성되어 있는 경우, AI 설정부(238)는 첫 번째 프레임과 관련하여 DNN 설정 정보를 획득하고, 두 번째 프레임과 관련하여 DNN 설정 정보를 획득하고, 세 번째 프레임과 관련하여 DNN 설정 정보를 획득할 수 있다. 즉, 첫 번째 프레임, 두 번째 프레임 및 세 번째 프레임 각각에 대해 DNN 설정 정보가 독립적으로 획득될 수 있다. 전술한, 제 1 복호화부(234)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.

[134] 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 ta 프레임까지 유효하다는 정보가 포함되어 있는 경우, AI 설정부(238)는 AI 데이터에 기초하여 DNN 설정 정보를 획득하고, AI 업스케일부(236)는 t0 내지 ta 프레임들을 해당 DNN 설정 정보로 AI 업스케일한다. 그리고, 다른 AI 데이터 내에 DNN 설정 정보가 tn 프레임까지 유효하다는 정보가 포함되어 있는 경우, AI 설정부(238)는 상기 다른 AI 데이터에 기초하여 DNN 설정 정보를 획득하고, AI 업스케일부(236)는 획득한 DNN 설정 정보로 ta+1 내지 tn 프레임들을 AI 업스케일할 수 있다.

[135] 이하에서는, 도 7을 참조하여, 원본 영상(105)의 AI 부호화를 위한 AI 부호화 장치(700)에 대해 설명한다.

[136] 도 7은 일 실시예에 따른 AI 부호화 장치(700)의 구성을 나타내는 블록도이다.

[137] 도 7을 참조하면, AI 부호화 장치(700)는 AI 부호화부(710) 및 전송부(730)를 포함할 수 있다. AI 부호화부(710)는 AI 다운스케일부(712), 제 1 부호화부(714), 데이터 처리부(716) 및 AI 설정부(718)를 포함할 수 있다.

[138] 도 7은 AI 부호화부(710) 및 전송부(730)를 개별적인 장치로 도시하고 있으나, AI 부호화부(710) 및 전송부(730)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 포함하거나, 외부 메모리를 이용하기 위한 메모리 처리부를 포함할 수 있다.

[139] 또한, AI 부호화부(710) 및 전송부(730)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.

[140] 일 실시예에서, 제 1 부호화부(714)는 제 1 프로세서로 구성되고, AI 다운스케일부(712), 데이터 처리부(716) 및 AI 설정부(718)는 제 1 프로세서와 상이한 제 2 프로세서로 구현되고, 전송부(730)는 제 1 프로세서 및 제 2 프로세서와 상이한 제 3 프로세서로 구현될 수 있다.AI 부호화부(710)는 원본 영상(105)의 AI 다운스케일 및 제 1 영상(115)의 제 1 부호화를 수행하고, AI 부호화 데이터를 전송부(730)로 전달한다. 전송부(730)는 AI 부호화 데이터를 AI 복호화 장치(200)로 전송한다.

[141] 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.

[142] AI 데이터는, AI 업스케일부(236)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. 일 예에서, 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.

[143] 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.

[144] 또한, 일 실시예에서, AI 데이터는 제 2 DNN에 세팅 가능한 DNN 설정 정보를 포함할 수도 있다.

[145] AI 다운스케일부(712)는 제 1 DNN을 통해 원본 영상(105)으로부터 AI 다운스케일된 제 1 영상(115)을 획득할 수 있다. AI 다운스케일부(712)는 AI 설정부(718)로부터 제공된 DNN 설정 정보를 이용하여 원본 영상(105)을 AI 다운스케일할 수 있다. AI 설정부(718)는 미리 결정된 기준에 기초하여 원본 영상(105)의 다운스케일 타겟을 결정할 수 있다.

[146] 다운스케일 타겟에 부합하는 제 1 영상(115)의 획득을 위해, AI 설정부(718)는 제 1 DNN에 세팅 가능한 복수의 DNN 설정 정보를 저장할 수 있다. AI 설정부(718)는 복수의 DNN 설정 정보 중 다운스케일 타겟에 대응하는 DNN 설정 정보를 획득하고, 획득된 DNN 설정 정보를 AI 다운스케일부(712)로 제공한다.

[147] 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.

[148] 구현예에 따라, DNN 설정 정보를 구성하는 정보들(예를 들어, 컨볼루션 레이어의 수, 컨볼루션 레이어별 필터 커널의 개수, 각 필터 커널의 파라미터 등)이 룩업 테이블 형태로 저장되어 있는 경우, AI 설정부(718)는 다운스케일 타겟에 따라 룩업 테이블 값들 중에서 선택된 일부를 조합하여 획득한 DNN 설정 정보를 AI 다운스케일부(712)로 제공할 수도 있다.

[149] 구현예에 따라, AI 설정부(718)는 다운스케일 타겟에 대응되는 DNN의 구조를

결정하고, 결정된 DNN의 구조에 대응하는 DNN 설정 정보, 예를 들어, 필터 커널의 파라미터들을 획득할 수도 있다.

[150] 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.

[151] AI 다운스케일부(712)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.

[152] 이하에서는, AI 설정부(718)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.

[153] AI 설정부(718)는 하나 이상의 입력 정보를 획득한다. 일 실시예에서, 입력 정보는 제 1 영상(115)의 타겟 해상도, 영상 데이터의 타겟 비트레이트, 영상 데이터의 비트레이트 타입(예를 들어, variable bitrate 타입, constant bitrate 타입 또는 average bitrate 타입 등), AI 다운스케일이 적용되는 컬러 포맷(휘도 성분, 색차 성분, 레드 성분, 그린 성분 또는 블루 성분 등), 제 1 영상(115)의 제 1 부호화를 위한 코덱 타입, 압축 히스토리 정보, 원본 영상(105)의 해상도 및 원본 영상(105)의 타입 중 적어도 하나를 포함할 수 있다.

[154] 하나 이상의 입력 정보는 AI 부호화 장치(700)에 미리 저장되거나, 사용자로부터 입력받은 정보를 포함할 수 있다.

[155] AI 설정부(718)는 입력 정보에 기초하여 AI 다운스케일부(712)의 동작을 제어한다. 일 실시예에서, AI 설정부(718)는 입력 정보에 따라 다운스케일 타겟을 결정하고, 결정된 다운스케일 타겟에 대응하는 DNN 설정 정보를 AI 다운스케일부(712)로 제공할 수 있다.

[156] 일 실시예에서, AI 설정부(718)는 입력 정보의 적어도 일부를 제 1 부호화부(714)로 전달하여 제 1 부호화부(714)가 특정 값의 비트레이트, 특정 타입의 비트레이트 및 특정 코덱으로 제 1 영상(115)을 제 1 부호화하게 할 수도 있다.

[157] 일 실시예에서, AI 설정부(718)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.

[158] 일 예에서, AI 설정부(718)는 미리 설정되거나, 사용자로부터 입력받은 압축률 또는 압축 품질 등에 기반하여 다운스케일 타겟을 결정할 수 있다.

[159] 다른 예로, AI 설정부(718)는 AI 부호화 장치(700)에 저장된 압축 히스토리 정보를 이용하여 다운스케일 타겟을 결정할 수도 있다. 예를 들어, AI 부호화 장치(700)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.

[160] 또 다른 예로, AI 설정부(718)는 압축 히스토리 정보에 따라 소정의 임계 값보다 많이 이용된 적이 있는 부호화 품질(예를 들면, 소정의 임계값보다 많이 이용된 적이 있는 부호화 품질들의 평균 품질)에 기초하여 다운스케일 타겟을 결정할 수도 있다.

[161] 또 다른 예로, AI 설정부(718)는 원본 영상(105)의 해상도, 타입(예를 들어, 파일의 형식)등에 기초하여 다운스케일 타겟을 결정할 수도 있다.

[162] 일 실시예에서, 원본 영상(105)이 복수의 프레임으로 구성된 경우, AI 설정부(718)는 소정 개수의 프레임 별로 DNN 설정 정보를 독립적으로 획득하고, 독립적으로 획득한 DNN 설정 정보를 AI 다운스케일부(712)로 제공할 수도 있다.

[163] 일 예시에서, AI 설정부(718)는 원본 영상(105)을 구성하는 프레임들을 소정 개수의 그룹으로 구분하고, 각 그룹별로 DNN 설정 정보를 독립적으로 획득할 수 있다. 각 그룹에 대해 서로 동일하거나 서로 상이한 DNN 설정 정보가 획득될 수 있다. 그룹들에 포함된 프레임들의 개수는 그룹별로 동일하거나 상이할 수 있다.

[164] 다른 예시에서, AI 설정부(718)는 원본 영상(105)을 구성하는 프레임별로 DNN 설정 정보를 독립적으로 결정할 수 있다. 각각의 프레임에 대해 서로 동일하거나 서로 상이한 DNN 설정 정보가 획득될 수 있다.

[165] 이하에서는, AI 다운스케일의 기반이 되는 제 1 DNN(800)의 예시적인 구조에 대해 설명한다.

[166] 도 8은 원본 영상(105)의 AI 다운스케일을 위한 제 1 DNN(800)을 나타내는 예시적인 도면이다.

[167] 도 8에 도시된 바와 같이, 원본 영상(105)은 제 1 컨볼루션 레이어(810)로 입력된다. 제 1 컨볼루션 레이어(810)는 5 x 5의 크기의 32개의 필터 커널을 이용하여 원본 영상(105)에 대해 컨볼루션 처리를 한다. 컨볼루션 처리 결과 생성된 32개의 특징 맵은 제 1 활성화 레이어(820)로 입력된다. 제 1 활성화 레이어(820)는 32개의 특징 맵에 대해 비선형(Non-linear) 특성을 부여할 수 있다.

[168] 제 1 활성화 레이어(820)는 제 1 컨볼루션 레이어(810)로부터 출력되는 특징 맵들의 샘플 값들을 제 2 컨볼루션 레이어(830)로 전달할지 여부를 결정한다. 예를 들어, 특징 맵들의 샘플 값들 중 어떤 샘플 값들은 제 1 활성화 레이어(820)에 의해 활성화되어 제 2 컨볼루션 레이어(830)로 전달되고, 어떤 샘플 값들은 제 1 활성화 레이어(820)에 의해 비활성화되어 제 2 컨볼루션 레이어(830)로 전달되지 않는다. 제 1 컨볼루션 레이어(810)로부터 출력되는

특징 맵들이 나타내는 정보가 제 1 활성화 레이어(820)에 의해 강조된다.

[169] 제 1 활성화 레이어(820)의 출력(825)이 제 2 컨볼루션 레이어(830)로 입력된다. 제 2 컨볼루션 레이어(830)는 5 x 5의 크기의 32개의 필터 커널을 이용하여 입력 데이터에 대해 컨볼루션 처리를 한다. 컨볼루션 처리 결과 출력된 32개의 특징 맵은 제 2 활성화 레이어(840)로 입력되고, 제 2 활성화 레이어(840)는 32개의 특징 맵에 대해 비선형 특성을 부여할 수 있다.

[170] 제 2 활성화 레이어(840)의 출력(845)은 제 3 컨볼루션 레이어(850)로 입력된다. 제 3 컨볼루션 레이어(850)는 5 x 5의 크기의 1개의 필터 커널을 이용하여 입력된 데이터에 대해 컨볼루션 처리를 한다. 컨볼루션 처리 결과 제 3 컨볼루션 레이어(850)로부터 1개의 영상이 출력될 수 있다. 제 3 컨볼루션 레이어(850)는 최종 영상을 출력하기 위한 레이어로서 1개의 필터 커널을 이용하여 1개의 출력을 획득한다. 본 개시의 예시에 따르면, 제 3 컨볼루션 레이어(850)는 컨벌루션 연산 결과를 통해 제 1 영상(115)을 출력할 수 있다.

[171] 제 1 DNN(800)의 제 1 컨볼루션 레이어(810), 제 2 컨볼루션 레이어(830) 및 제 3 컨볼루션 레이어(850)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.

[172] 도 8은 제 1 DNN(800)이 세 개의 컨볼루션 레이어(810, 830, 750)와 두 개의 활성화 레이어(820, 840)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(800)은 RNN(recurrent neural network)을 통해 구현될 수도 있다. 이 경우는 본 개시의 예시에 따른 제 1 DNN(800)의 CNN 구조를 RNN 구조로 변경하는 것을 의미한다.

[173] 일 실시예에서, AI 다운스케일부(712)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.

[174] 다시 도 7을 참조하면, AI 설정부(718)는 AI 데이터를 데이터 처리부(716)로 전달한다. AI 데이터는, AI 업스케일부(236)가 제 1 DNN의 다운스케일 타겟에 대응하는 업스케일 타겟으로 제 2 영상(135)을 AI 업스케일할 수 있게 하는

정보들을 포함한다. AI 다운스케일부(712)로부터 제 1 영상(115)을 전달받은 제 1 부호화부(714)는 주파수 변환 기반의 영상 압축 방법에 따라 제 1 영상(115)을 제 1 부호화하여 제 1 영상(115)이 가지는 정보량을 감축시킬 수 있다. 소정의 코덱(예를 들어, MPEG-2, H.264, MPEG-4, HEVC, VC-1, VP8, VP9 또는 AV1)을 통한 제 1 부호화 결과, 영상 데이터가 획득된다. 영상 데이터는 소정의 코덱의 규칙, 즉 신택스에 따라 획득된다. 예를 들어, 영상 데이터는 제 1 영상(115)과 제 1 영상(115)의 예측 데이터 사이의 차이인 잔차 데이터, 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. 제 1 부호화부(714)의 제 1 부호화 결과 획득된 영상 데이터는 데이터 처리부(716)로 제공된다.

[175] 데이터 처리부(716)는 제 1 부호화부(714)로부터 수신된 영상 데이터와 AI 설정부(718)로부터 수신된 AI 데이터를 포함하는 AI 부호화 데이터를 생성한다.

[176] 일 실시예에서, 데이터 처리부(716)는 영상 데이터와 AI 데이터를 각각 분리된 상태로 포함하는 AI 부호화 데이터를 생성할 수 있다. 일 예로, AI 데이터는 HDMI 스트림 내의 VSIF(Vendor Specific InfoFrame)에 포함될 수 있다.

[177] 다른 실시예에서, 데이터 처리부(716)는 제 1 부호화부(714)에 의한 제 1 부호화 결과로 획득되는 영상 데이터 내에 AI 데이터를 포함시키고, 해당 영상 데이터를 포함하는 AI 부호화 데이터를 생성할 수 있다. 일 예로, 데이터 처리부(716)는 영상 데이터에 해당하는 비트스트림과 AI 데이터에 해당하는 비트스트림을 결합하여 하나의 비트스트림 형태의 영상 데이터를 생성할 수도 있다. 이를 위해, 데이터 처리부(716)는 AI 데이터를 0 또는 1의 값을 갖는 비트들, 즉 비트스트림으로 표현할 수 있다. 일 실시예에서, 데이터 처리부(716)는 제 1 부호화 결과로 획득되는 비트스트림의 부가 정보 영역인 SEI(Supplemental enhancement information)에 AI 데이터에 해당하는 비트스트림을 포함시킬 수 있다.

[178] AI 부호화 데이터는 전송부(730)로 전송된다. 전송부(730)는 네트워크를 통해 AI 부호화 결과 획득된 AI 부호화 데이터를 전송한다.일 실시예에서, 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.

[179] 도 9는 일 실시예에 따른 AI 부호화 데이터(900)의 구조를 나타내는 도면이다.

[180] 전술한 바와 같이, AI 부호화 데이터(900) 내에 AI 데이터(912)와 영상 데이터(932)가 분리되어 포함될 수 있다. 여기서, AI 부호화 데이터(900)는 MP4, AVI, MKV, FLV 등의 컨테이너 포맷일 수 있다. AI 부호화 데이터(900)는 메타데이터 박스(910)와 미디어데이터 박스(930)로 구성될 수 있다.

[181] 메타데이터 박스(910)에는 미디어데이터 박스(930)에 포함된 영상

데이터(932)에 관한 정보가 포함된다. 예를 들어, 메타데이터 박스(910)에는 제 1 영상(115)의 종류, 제 1 영상(115)의 부호화에 사용된 코덱의 종류 및 제 1 영상(115)의 재생 시간 등에 관한 정보가 포함될 수 있다. 또한, 메타데이터 박스(910)에는 AI 데이터(912)가 포함될 수 있다. AI 데이터(912)는 소정 컨테이너 포맷에서 제공하는 부호화 방식에 따라 부호화되어, 메타데이터 박스(910)에 저장될 수 있다.

[182] 미디어데이터 박스(930)는 소정의 영상 압축 방식의 신택스에 따라 생성된 영상 데이터(932)를 포함할 수 있다.

[183] 도 10은 다른 실시예에 따른 AI 부호화 데이터(1000)의 구조를 나타내는 도면이다.

[184] 도 10을 참조하면, AI 데이터(1034)는 영상 데이터(1032)에 포함될 수도 있다. AI 부호화 데이터(1000)는 메타데이터 박스(1010)와 미디어데이터 박스(1030)를 포함할 수 있는데, AI 데이터(1034)가 영상 데이터(1032)에 포함된 경우, 메타데이터 박스(1010)에는 AI 데이터(1034)가 포함되지 않을 수 있다.

[185] 미디어데이터 박스(1030)에는 AI 데이터(1034)를 포함하는 영상 데이터(1032)가 포함된다. 일 예로, AI 데이터(1034)는 영상 데이터(1032)의 부가 정보 영역에 포함될 수 있다.

[186] 이하에서는, 도 11을 참조하여, 제 1 DNN(800)과 제 2 DNN(300)을 연계 훈련시키는 방법에 대해 설명한다.

[187] 도 11은 제 1 DNN(800) 및 제 2 DNN(300)을 훈련시키는 방법을 설명하기 위한 도면이다.

[188] 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(800)과 제 2 DNN(300)의 연계 훈련이 요구된다.

[189] 정확한 AI 복호화를 위해서는 궁극적으로 도 11에 도시된 제 3 훈련 영상(1104)과 원본 훈련 영상(1101) 사이의 비교 결과에 대응하는 퀄리티 손실 정보(1130)를 감소시킬 필요가 있다. 따라서, 퀄리티 손실 정보(1130)는 제 1 DNN(800) 및 제 2 DNN(300)의 훈련 모두에 이용된다.

[190] 먼저, 도 11에 도시된 훈련 과정에 대해 설명한다.

[191] 도 11에서, 원본 훈련 영상(original training image)(1101)은 AI 다운스케일의 대상이 되는 영상이고, 제 1 훈련 영상(first training image)(1102)은 원본 훈련 영상(1101)로부터 AI 다운스케일된 영상이다. 또한, 제 3 훈련 영상(third training image)(1104)은 제 1 훈련 영상(1102)으로부터 AI 업스케일된 영상이다.

[192] 원본 훈련 영상(1101)은 정지 영상 또는 복수의 프레임으로 이루어진 동영상을 포함한다. 일 실시예에서, 원본 훈련 영상(1101)은 정지 영상 또는 복수의

프레임으로 이루어진 동영상으로부터 추출된 휘도 영상을 포함할 수도 있다. 또한, 일 실시예에서, 원본 훈련 영상(1101)은 정지 영상 또는 복수의 프레임으로 이루어진 동영상에서 추출된 패치 영상을 포함할 수도 있다. 원본 훈련 영상(1101)이 복수의 프레임으로 이루어진 경우, 제 1 훈련 영상(1102), 제 2 훈련 영상 및 제 3 훈련 영상(1104) 역시 복수의 프레임으로 구성된다. 원본 훈련 영상(1101)의 복수의 프레임이 순차적으로 제 1 DNN(800)에 입력되면, 제 1 DNN(800) 및 제 2 DNN(300)을 통해 제 1 훈련 영상(1102), 제 2 훈련 영상 및 제 3 훈련 영상(1104)의 복수의 프레임이 순차적으로 획득될 수 있다.

[193] 제 1 DNN(800) 및 제 2 DNN(300)의 연계 훈련을 위해, 원본 훈련 영상(1101)이 제 1 DNN(800)으로 입력된다. 제 1 DNN(800)으로 입력된 원본 훈련 영상(1101)은 AI 다운스케일되어 제 1 훈련 영상(1102)으로 출력되고, 제 1 훈련 영상(1102)이 제 2 DNN(300)에 입력된다. 제 1 훈련 영상(1102)에 대한 AI 업스케일 결과 제 3 훈련 영상(1104)이 출력된다.

[194] 도 11을 참조하면, 제 2 DNN(300)으로 제 1 훈련 영상(1102)이 입력되고 있는데, 구현예에 따라, 제 1 훈련 영상(1102)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(1102)의 제 1 부호화 및 제 1 훈련 영상(1102)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.

[195] 도 11을 참조하면, 제 1 DNN(800)을 통해 제 1 훈련 영상(1102)이 출력되는 것과 별개로, 원본 훈련 영상(1101)으로부터 레거시 다운스케일된 축소 훈련 영상(1103)이 획득된다. 여기서, 레거시 다운스케일은 바이리니어(bilinear) 스케일, 바이큐빅(bicubic) 스케일, 란조스(lanczos) 스케일 및 스테어 스탭(stair step) 스케일 중 적어도 하나를 포함할 수 있다.

[196] 원본 영상(105)의 구조적 특징을 기준으로 제 1 영상(115)의 구조적 특징이 크게 벗어나는 것을 방지하기 위해, 원본 훈련 영상(1101)의 구조적 특징을 보존하는 축소 훈련 영상(1103)을 획득하는 것이다.

[197] 훈련의 진행 전 제 1 DNN(800) 및 제 2 DNN(300)은 미리 결정된 DNN 설정 정보로 세팅될 수 있다. 훈련이 진행됨에 따라 구조적 손실 정보(1110), 복잡성 손실 정보(1120) 및 퀄리티 손실 정보(1130)가 결정될 수 있다.

[198] 구조적 손실 정보(1110)는 축소 훈련 영상(1103)과 제 1 훈련 영상(1102)의 비교 결과에 기초하여 결정될 수 있다. 일 예에서, 구조적 손실 정보(1110)는 축소 훈련 영상(1103)의 구조적 정보와 제 1 훈련 영상(1102)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(1110)는 원본 훈련 영상(1101)의 구조적 정보가 제 1 훈련 영상(1102)에서 어느 정도로 유지되고

있는지를 나타낸다. 구조적 손실 정보(1110)가 작을수록 제 1 훈련 영상(1102)의 구조적 정보가 원본 훈련 영상(1101)의 구조적 정보와 유사해진다.

[199] 복잡성 손실 정보(1120)는 제 1 훈련 영상(1102)의 공간적 복잡도에 기반하여 결정될 수 있다. 일 예에서, 공간적 복잡도로서, 제 1 훈련 영상(1102)의 총 분산(total variance)값이 이용될 수 있다. 복잡성 손실 정보(1120)는 제 1 훈련 영상(1102)을 제 1 부호화하여 획득한 영상 데이터의 비트레이트와 관련된다. 복잡성 손실 정보(1120)가 작을수록 영상 데이터의 비트레이트가 작은 것으로 정의된다.

[200] 퀄리티 손실 정보(1130)는 원본 훈련 영상(1101)과 제 3 훈련 영상(1104)의 비교 결과에 기초하여 결정될 수 있다. 퀄리티 손실 정보(1130)는 원본 훈련 영상(1101)과 제 3 훈련 영상(1104)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(1130)는 제 3 훈련 영상(1104)이 원본 훈련 영상(1101)과 어느 정도로 유사한지를 나타낸다. 퀄리티 손실 정보(1130)가 작을수록 제 3 훈련 영상(1104)이 원본 훈련 영상(1101)에 더 유사해진다.

[201] 도 11을 참조하면, 구조적 손실 정보(1110), 복잡성 손실 정보(1120) 및 퀄리티 손실 정보(1130)가 제 1 DNN(800)의 훈련에 이용되고, 퀄리티 손실 정보(1130)는 제 2 DNN(300)의 훈련에 이용된다. 즉, 퀄리티 손실 정보(1130)는 제 1 DNN(800) 및 제 2 DNN(300)의 훈련에 모두 이용된다.

[202] 제 1 DNN(800)은 구조적 손실 정보(1110), 복잡성 손실 정보(1120) 및 퀄리티 손실 정보(1130)에 기초하여 결정된 최종 손실 정보가 감소 또는 최소화되도록 파라미터를 갱신할 수 있다. 또한, 제 2 DNN(300)은 퀄리티 손실 정보(1130)가 감소 또는 최소화되도록 파라미터를 갱신할 수 있다.

[203] 제 1 DNN(800) 및 제 2 DNN(300)의 훈련을 위한 최종 손실 정보는 아래의 수학식 1과 같이 결정될 수 있다.

[204] [수학식 1]

[205] LossDS = a*구조적 손실 정보 + b*복잡성 손실 정보 + c*퀄리티 손실 정보

[206] LossUS = d*퀄리티 손실 정보

[207] 상기 수학식 1에서, LossDS는 제 1 DNN(800)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.

[208] 즉, 제 1 DNN(800)은 수학식 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(800)의 파라미터들이 갱신되면, 갱신된 파라미터에 기초하여 획득되는 제 1 훈련

영상(1102)이 이전 훈련 과정에서의 제 1 훈련 영상(1102)과 달라지게 되고, 그에 따라 제 3 훈련 영상(1104) 역시 이전 훈련 과정에서의 제 3 훈련 영상(1104)과 달라지게 된다. 제 3 훈련 영상(1104)이 이전 훈련 과정에서의 제 3 훈련 영상(1104)과 달라지게 되면, 퀄리티 손실 정보(1130) 역시 새롭게 결정되며, 그에 따라 제 2 DNN(300)은 파라미터들을 갱신한다. 퀄리티 손실 정보(1130)가 새롭게 결정되면, LossDS 역시 새롭게 결정되므로, 제 1 DNN(800)은 새롭게 결정된 LossDS에 따라 파라미터들을 갱신한다. 즉, 제 1 DNN(800)의 파라미터 갱신은, 제 2 DNN(300)의 파라미터 갱신을 야기하고, 제 2 DNN(300)의 파라미터 갱신은 제 1 DNN(800)의 파라미터 갱신을 야기하는 것이다. 다시 말하면, 제 1 DNN(800) 및 제 2 DNN(300)은 퀄리티 손실 정보(1130)의 공유를 통해 연계 훈련되므로, 제 1 DNN(800)의 파라미터들과 제 2 DNN(300)의 파라미터들이 서로 연관성을 가지고 최적화될 수 있는 것이다.

[209] 수학식 1을 참조하면, LossUS가 퀄리티 손실 정보(1130)에 따라 결정되는 것을 알 수 있으나, 이는 하나의 예시이며, LossUS는 구조적 손실 정보(1110) 및 복잡성 손실 정보(1120) 중 적어도 하나와, 퀄리티 손실 정보(1130)에 기초하여 결정될 수도 있다.

[210] 앞서, AI 복호화 장치(200)의 AI 설정부(238) 및 AI 부호화 장치(700)의 AI 설정부(718)는 복수의 DNN 설정 정보를 저장하는 것으로 설명하였는데, AI 설정부(238) 및 AI 설정부(718)에 저장되는 복수의 DNN 설정 정보 각각을 훈련시키는 방법에 대해 설명한다.

[211] 수학식 1과 관련하여 설명한 바와 같이, 제 1 DNN(800)의 경우, 제 1 훈련 영상(1102)의 구조적 정보와 원본 훈련 영상(1101)의 구조적 정보 사이의 유사 정도(구조적 손실 정보(1110)), 제 1 훈련 영상(1102)의 제 1 부호화 결과 획득되는 영상 데이터의 비트레이트(복잡성 손실 정보(1120)) 및 제 3 훈련 영상(1104)과 원본 훈련 영상(1101) 사이의 차이(퀄리티 손실 정보(1130))를 고려하여 파라미터를 갱신하게 된다.

[212] 자세히 설명하면, 원본 훈련 영상(1101)의 구조적 정보와 유사하면서, 제 1 부호화를 하였을 때 획득되는 영상 데이터의 비트레이트가 작은 제 1 훈련 영상(1102)이 획득 가능하도록 하는 동시에, 제 1 훈련 영상(1102)을 AI 업스케일하는 제 2 DNN(300)이 원본 훈련 영상(1101)에 유사한 제 3 훈련 영상(1104)을 획득할 수 있도록, 제 1 DNN(800)의 파라미터가 갱신될 수 있다.

[213] 수학식 1의 a, b, c의 가중치가 조절됨으로써, 제 1 DNN(800)의 파라미터들이 최적화되는 방향이 상이해지게 된다. 예를 들어, b의 가중치를 높게 결정하는 경우, 제 3 훈련 영상(1104)의 퀄리티보다 비트레이트가 낮아지는 것에 더 중요도를 두고 제 1 DNN(800)의 파라미터가 갱신될 수 있다. 또한, c의 가중치를 높게 결정하는 경우, 비트레이트가 높아지는 것이나, 원본 훈련 영상(1101)의 구조적 정보가 유지되는 것보다 제 3 훈련 영상(1104)의 퀄리티가 증가하도록 하는 것에 더 중요도를 두고 제 1 DNN(800)의 파라미터가 갱신될 수 있다.

[214] 또한, 제 1 훈련 영상(1102)을 제 1 부호화하는데 이용되는 코덱의 타입에 따라 제 1 DNN(800)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.

[215] 즉, 가중치 a, 가중치 b, 가중치 c 및 제 1 훈련 영상(1102)의 제 1 부호화를 위한 코덱의 종류에 기반하여 제 1 DNN(800)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(800)과 제 2 DNN(300)을 훈련시키면, 서로 연계되어 최적화된 제 1 DNN(800)의 파라미터들과 제 2 DNN(300)의 파라미터들이 결정될 수 있다.

[216] 그리고, 가중치 a, 가중치 b, 가중치 c 및 코덱의 종류를 변경한 후, 제 1 DNN(800)과 제 2 DNN(300)을 훈련시키면, 서로 연계되어 최적화된 제 1 DNN(800)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(800)과 제 2 DNN(300)을 훈련시키면 서로 연계되어 훈련된 복수의 DNN 설정 정보가 제 1 DNN(800) 및 제 2 DNN(300)에서 결정될 수 있는 것이다.

[217] 앞서, 도 5와 관련하여 설명한 바와 같이, 제 1 DNN(800) 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(800)에서 출력되는 제 1 훈련 영상(1102)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(1102)이 특정 코덱에 의해 특정 비트레이트로 제 1 부호화되도록 환경을 설정한 후, 제 1 DNN(800) 및 제 2 DNN(300)을 훈련시킴으로써, 제 1 훈련 영상(1102)의 해상도, 제 1 훈련 영상(1102)의 제 1 부호화에 이용된 코덱의 종류 및 제 1 훈련 영상(1102)의 제 1 부호화 결과 획득된 비트스트림의 비트레이트에 매핑된 DNN 설정 정보 쌍이 결정될 수 있는 것이다. 제 1 훈련 영상(1102)의 해상도, 제 1 훈련 영상(1102)의 제 1 부호화에 이용된 코덱의 종류 및 제 1 훈련 영상(1102)의 제 1 부호화에 따라 획득되는 비트스트림의 비트레이트를 다양하게 변경시킴으로써, 제 1 DNN(800) 및 제 2 DNN(300)의 복수의 DNN 설정 정보들과 제 1 영상 관련 정보들 사이의 매핑 관계가 결정될 수 있다.

[218] 도 12는 훈련 장치(1200)에 의한 제 1 DNN(800) 및 제 2 DNN(300)의 훈련 과정을 설명하기 위한 도면이다.

[219] 도 11과 관련하여 설명한 제 1 DNN(800) 및 제 2 DNN(300)의 훈련은 훈련 장치(1200)에 의해 수행될 수 있다. 훈련 장치(1200)는 제 1 DNN(800) 및 제 2 DNN(300)을 포함한다. 훈련 장치(1200)는 예를 들어, AI 부호화 장치(700) 또는 별도의 서버일 수 있다. 훈련 결과 획득된 제 2 DNN(300)의 DNN 설정 정보들은 AI 복호화 장치(200)에 저장된다.

[220] 도 12를 참조하면, 훈련 장치(1200)는 제 1 DNN(800) 및 제 2 DNN(300)의 DNN 설정 정보를 초기 세팅한다(S1240, S1245). 이에 의해, 제 1 DNN(800) 및 제 2 DNN(300)은 미리 결정된 DNN 설정 정보에 따라 동작할 수 있다. DNN 설정 정보는 제 1 DNN(800) 및 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.

[221] 훈련 장치(1200)는 원본 훈련 영상(1101)을 제 1 DNN(800)으로 입력한다(S1250). 원본 훈련 영상(1101)은 정지 영상 또는 동영상을 구성하는 적어도 하나의 프레임을 포함할 수 있다.

[222] 제 1 DNN(800)은 초기 세팅된 DNN 설정 정보에 따라 원본 훈련 영상(1101)을 처리하고, 원본 훈련 영상(1101)으로부터 AI 다운스케일된 제 1 훈련 영상(1102)을 출력한다(S1255). 도 12는 제 1 DNN(800)으로부터 출력된 제 1 훈련 영상(1102)이 제 2 DNN(300)으로 바로 입력되는 것으로 도시되어 있으나, 제 1 DNN(800)으로부터 출력된 제 1 훈련 영상(1102)이 훈련 장치(1200)에 의해 제 2 DNN(300)으로 입력될 수 있다. 또한, 훈련 장치(1200)는 제 1 훈련 영상(1102)을 소정의 코덱으로 제 1 부호화 및 제 1 복호화한 후, 제 2 훈련 영상을 제 2 DNN(300)으로 입력할 수 있다.

[223] 제 2 DNN(300)은 제 1 훈련 영상(1102) 또는 제 2 훈련 영상을 초기 세팅된 DNN 설정 정보에 따라 처리하고, 제 1 훈련 영상(1102) 또는 제 2 훈련 영상으로부터 AI 업스케일된 제 3 훈련 영상(1104)을 출력한다(S1260).

[224] 훈련 장치(1200)는 제 1 훈련 영상(1102)에 기초하여 복잡성 손실 정보(1120)를 산출한다(S1265).

[225] 훈련 장치(1200)는 축소 훈련 영상(1103)과 제 1 훈련 영상(1102)을 비교하여 구조적 손실 정보(1110)를 산출한다(S1270).

[226] 훈련 장치(1200)는 원본 훈련 영상(1101)과 제 3 훈련 영상(1104)을 비교하여 퀄리티 손실 정보(1130)를 산출한다(S1275).

[227] 제 1 DNN(800)은 최종 손실 정보에 기초한 역전사(back propagation) 과정을 통해 초기 세팅된 DNN 설정 정보를 갱신한다(S1280). 훈련 장치(1200)는 복잡성 손실 정보(1120), 구조적 손실 정보(1110) 및 퀄리티 손실 정보(1130)에 기초하여 제 1 DNN(800)의 훈련을 위한 최종 손실 정보를 산출할 수 있다.

[228] 제 2 DNN(300)은 퀄리티 손실 정보 또는 최종 손실 정보에 기초한 역전사 과정을 통해 초기 세팅된 DNN 설정 정보를 갱신한다(S1285). 훈련 장치(1200)는 퀄리티 손실 정보(1130)에 기초하여 제 2 DNN(300)의 훈련을 위한 최종 손실 정보를 산출할 수 있다.

[229] 이후, 훈련 장치(1200), 제 1 DNN(800) 및 제 2 DNN(300)은 최종 손실 정보들이 최소화될 때까지 S1250 내지 S1285 과정을 반복하면서 DNN 설정 정보를 갱신한다. 이 때, 각 반복 과정 동안 제 1 DNN(800) 및 제 2 DNN은 이전 과정에서 갱신된 DNN 설정 정보에 따라 동작한다.

[230] 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.

[231] [표 1]

[232]

컨텐츠	해상도	프레임수	정보량 (Bitrate) (Mbps)		주관적 화질 스코어 (VMAF)	
			HEVC	AI 부호화/ AI 복호화	HEVC	AI 부호화/ AI 복호화
컨텐츠_01	8K (7680×4320)	300 frames	46.3	21.4	94.80	93.54
컨텐츠_02			46.3	21.6	98.05	98.98
컨텐츠_03			46.3	22.7	96.08	96.00
컨텐츠_04			46.1	22.1	86.26	92.00
컨텐츠_05			45.4	22.7	93.42	92.98
컨텐츠_06			46.3	23.0	95.99	95.61
		평균	46.11	22.25	94.10	94.85

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

[234] 이하에서는, 도 1 내지 도 12에서 전술된 AI 기반의 업스케일 및 다운스케일에 따라, 영상 전송시의 해상도와 비트레이트에 따라 결정된 AI 업스케일 모델 및 AI 다운스케일 모델을 이용하여 영상을 실시간으로 스트리밍할 때, 스트리밍 채널의 환경 변화로 인해 대역폭이 줄어들어 발생하는 블로킹 현상을 해결하기 위해 디블로킹 필터를 추가하여 영상을 전송하는 방법이 후술된다.

[235] 도 13은 일 실시예에 따른 디블로킹 필터링을 추가하는 AI 부호화 과정 및 AI 복호화 과정을 설명하기 위한 도면이다.

[236] 도 1에서 전술된 것과 같이, 스트리밍 채널 환경이 정상적인 경우, 스트리밍 채널 환경에 따라 적합한 AI 다운스케일 모델 및 AI 업스케일 모델에 기초하여 원본 영상(105)을 AI 다운스케일(110)하여 제 1 영상(115)을 획득하고, 제 1 영상(115)을 제 1 부호화(120)하여 전송하고, 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.

[237] 반면에, 스트리밍 채널의 환경 변화가 발생하는 경우, 예를 들어, 갑작스럽게 많은 사용자들이 몰리거나 실외에서 단말의 통신 환경이 안좋아지는 경우에는 대역폭이 줄어들기 때문에 비트레이트를 낮추기 위해 양자화 파라미터를 증가시켜 영상의 해상도 및 화면 특성에 따라 블로킹 현상이 발생될 수 있다. 블로킹 현상, 즉, 블로킹 아티팩트가 발생하면 시청자는 영상의 화질이 나쁘다고 생각한다. 이러한 블로킹 현상을 방지하기 위해, 디블로킹 필터를 추가하여 AI

다운스케일 및 AI 업스케일함으로써 영상의 화질을 개선할 수 있다.

[238] 도 13을 참고하면, 디블로킹 예측(1370)에서, 현재 스트리밍 채널의 네트워크 환경 정보, 원본 영상(1305), 이전에 부호화된 프레임의 정보를 획득하여, 입력되는 원본 영상에 블로킹 현상이 발생할 지를 예측하고, 블록이 생기는 영역에 대한 영역 맵과 같은 필터 설정 정보를 AI 부호화 및 AI 복호화가 수행될 때 전송한다. AI 부호화 과정에서, 이러한 디블로킹 필터를 위한 필터 설정 정보에 기초하여 원본 영상(1305)에 디블로킹 필터(1310)을 적용하여 디블로킹 필터링된 원본 영상(1315)을 획득하고, 디블로킹 필터링된 영상(1315)을 AI 다운스케일(1320)하여 제 1 영상(1325)을 획득하고, 제 1 영상(1325)을 제 1 부호화(1330)하여 전송한다, AI 복호화 과정에서는, AI 부호화 결과 획득된 AI 데이터와 영상 데이터를 포함하는 AI 부호화 데이터를 수신하고, 제 1 복호화(1340)를 통해 제 2 영상(1345)을 획득하고, 제 2 영상(1345)을 AI 업스케일(1350)하여 AI 업스케일된 영상(1355)을 획득하고, 디블로킹 예측(1370)에서 전송된 필터 설정 정보에 따른 디블로킹 필터(1360)을 AI 업스케일된 영상(1355)에 적용하여 복원 영상(1365)을 획득한다.

[239] 네트워크 환경 정보는 이동 단말 간의 통신, 서버와 클라이언트 간의 통신과 같은 네트워크 통신에서 이용되는 실시간 프로토콜(예를 들어, 실시간 전송 프로토콜(Real-time Transport Protocol, RTP), 실시간 전송 제어 프로토콜(Real-time Transport Control Protocol, RTCP))에 따라 전송되는 네트워크 환경의 변화를 나타내는 정보이다.

[240] 미리정해진 해상도와 비트레이트로 스트리밍이 수행되는 도중에 네트워크 환경의 변화로 대역폭이 줄어들면 블로킹 현상이 발생될 수 있다. 예를 들어, 해상도가 4K이고 비트레이트가 20Mbps인 경우, 해상도가 4K이고 비트레이트가 15Mbps인 경우, 해상도가 FHD이고 비트레이트가 10Mbps인 경우, 해상도가 FHD이고 비트레이트가 7.5Mbps인 경우, 해상도가 HD이고 비트레이트가 5Mbps인 경우 등의 각각의 경우에 적합한 AI 다운스케일 및 AI 업스케일 모델이 선택되어 AI 부호화 및 AI 복호화 과정이 수행되는 도중에 네트워크 환경의 변화(사용자의 급격한 증가, 실외에서의 통신 환경이 안좋아지는 경우)에 따라, 디블로킹 예측 과정을 통해 AI 부호화 및 AI 복호화에 네트워크 환경의 변화, 원본 영상의 특성, 이전 프레임의 블로킹 아티팩트 검출 여부에 따라 디블로킹 필터링을 수행함으로써, 가변적이고 제한적인 네트워크 대역폭에서 블로킹 아티팩트의 발생을 예측하고 인지하여 블로킹 아티팩트를 제거하여 영상의 화질이 개선될 수 있다. 일반적으로 영상의 부호화 및 복호화에서 인루프 필터로 블록 단위로 블로킹 아티팩트를 제거하는 디블로킹 필터가 있지만, 인루프 필터이기 때문에 디블로킹 필터링을 강하게 걸면, 실제 데이터가 손상될 수 있기 때문에, 추가적인 디블로킹 필터를 통해 디블로킹 필터링이 수행되어 영상의 화질이 개선될 수 있다.

[241] 디블로킹 예측(1370)은 원본 영상(1305)으로부터 블로킹이 생길 가능성이 있는

영역이 존재하는지 여부를 획득하고, 이전 프레임 정보로부터 현재 프레임의 이전 프레임의 블로킹 현상 검출 여부를 획득하고, 네트워크 환경 정보로부터 대역폭, 비트레이트 등의 변화에 대한 정보를 획득하여, 블록이 생길 수 있는 영역에 대한 블로킹 영역 맵 등을 포함하는 필터 설정 정보를 AI 복호화 및 AI 부호화가 수행될 때 전송한다. 전송된 필터 설정 정보에 기초하여 AI 복호화 및 AI 부호화 과정 각각이 수행될 때 디블로킹 필터링이 추가로 수행된다.

[242] AI 부호화 과정에서의 디블로킹 필터(1310)는 전처리-필터(pre-filter)로서, 할당된 비트보다 전체 복잡도가 더 복잡한 경우, 복잡한 영역과 복잡하지 않은 영역의 화질을 재분배한다. 예를 들어, 영상에서 가만히 서있는 나무와 같은 단순한 영역과 분수에서 물줄기가 나오는 것과 같은 복잡하고 움직임이 빠른 영역에 있는 경우, 복잡한 영역을 단순하게 블러 처리하여 복잡도를 줄여 비트레이트를 절감하고, 줄어든 복잡도에 따른 비트레이트는 부호화시 단순한 영역에 분배된다. 블로킹 현상은 화질 분배가 제대로 되지 않을 때 나타나는 현상이므로, 화질을 재분배함으로써 블로킹 현상의 발생 확률을 줄일 수 있다. 복잡하고 움직임이 빠른 영역은 블러 처리하여 블로킹 아티팩트가 눈에 덜 띄도록 하고, 복잡하고 움직임이 빠른 영역에 대한 블러 처리를 통해 절감된 비트레이트를 단순한 영역에 할당하여 이전보다 더 많은 비트레이트로 단순한 영역의 블로킹 현상의 발생을 방지한다.

[243] AI 복호화 과정에서의 디블로킹 필터(1360)는 후처리-필터(post-filter)로서, 필터 설정 정보에 포함된 블로킹 영역 맵에 따라 디블로킹 필터링을 적용하여, 블러 처리된 부분에 대하여 경계면이 선명하게 보이도록 하는 샤픈(sharpen) 효과를 내고, 블로킹 아티팩트를 제거하여 복원한다.

[244] 디블로킹 예측(1370)에서, 현재 프레임의 블로킹 아티팩트 발생 확률과 영역 별 발생 확률에 대한 블로킹 영역 맵을 포함하는 필터 설정 정보가 생성될 수 있다. 원본 영상에서 움직임이 빠르고 복잡한 영역과 평평한 영역은 상대적으로 블록이 더 잘 보인다. 이러한 점에 기초하여 블로킹 아티팩트의 발생 확률을 나타내는 블로킹 영역 맵이 생성될 수 있다.

[245] 또한, 여러가지 기준에 따라, 스코어 맵을 생성하여 가중치 평균에 따라 블로킹 영역 맵이 생성될 수 있다.

[246] 예를 들어, A 스코어 맵은 원본 영상으로부터 블록 또는 픽셀 별로 복잡도를 계산하여 복잡도가 높은 영역을 찾고, 복잡도가 높은 영역에 높은 스코어를 주어 생성할 수 있고, 복잡도는 총 분산 값 등에 기초하여 계산될 수 있다. B 스코어 맵은 N-1 번째 프레임과 옵티컬 플로(optical flow) 등으로부터 움직임을 확인하여 움직임의 변화가 많은 영역에 높은 스코어를 주어 생성할 수 있다.

[247] 또한, 이전 영상에 블로킹 아티팩트가 많으면 현재 영상에도 블로킹 아티팩트가 많을 수 있다. 이에 따라, C 스코어 맵은 이전 영역에서 블로킹 아티팩트가 많은 영역에 높은 스코어를 주어 생성할 수 있고, D 스코어 맵은 블록 경계에서 픽셀의 경계값 차이가 큰 영역에 높은 스코어를 주어 생성할 수

있다.

[248] 마지막으로, 부호화시에 현재 프레임의 복잡도, 타겟 비트레이트, 양자화 파라미터(QP)에 따라 블로킹 아티팩트가 발생될 수 있다. 예를 들어, E 스코어 맵은 현재 프레임의 블록의 픽셀들의 평균 QP와 각각의 픽셀들의 QP의 차이를 나타내는 DQP(delta quantization parameter) 맵에 기초하여 생성할 수 있다. F 스코어 맵은 부호화에 따른 복잡도, 즉 코딩 스코어에 기초하여 생성하고, 코딩 스코어는 인트라 코딩 스코어와 인터 코딩 스코어의 합으로 결정되고, 인트라 코딩 스코어는 현재 이미지를 입력으로 하여 원본 이미지와 인트라 코딩된 이미지의 차이와 이미지를 코딩하는데 이용되는 예측 비트에 기초하여 결정되고, 인터 코딩 스코어는 현재 이미지와 적어도 하나의 이전 이미지를 입력으로하여 원본 이미지와 인터 코딩된 이미지의 차이와 이미지를 코딩하는데 이용되는 예측 비트에 기초하여 결정될 수 있다. G 스코어 맵은 F 스코어 맵의 코딩 스코어를 타겟 비트레이트(채널 정보, 대역폭)로 나눈 값에 기초하여 생성할 수 있다. 이에 따라, G 스코어 맵은 인코더가 네트워크 상 이용가능한 대역폭에 기반하여 부호화된 비트스트림의 비트레이트를 추정하는 레이트 컨트롤(rate control)에 기초하여 생성할 수 있다. 코딩 스코어가 크더라도 타겟 비트레이트가 상대적으로 크면, G 스코어 맵에서는 스코어가 크지 않을 수 있다.

[249] 위와 같이, A 내지 G 스코어 맵에 대한 가중치 평균이 블로킹 영역 맵이 될 수 있다. 다만, A 내지 G 스코어 맵은 하나의 예시이고 이에 한정되지 않는다.

[250] 디블로킹 예측 과정에서 이러한 블로킹 영역 맵을 포함하는 필터 설정 정보에 따라 디블로킹 필터링이 수행될 수 있다.

[251] 일 실시예에 따라, 디블로킹 필터는 레거시 디블로킹 필터일 수 있다.

[252] 다른 실시예에 따라, 디블로킹 필터는 훈련된 DNN일 수 있다. 이에 따라, 디블로킹 예측에서 디블로킹 필터를 위해서 전송하는 필터 설정 정보에는 블로킹 영역 맵과 DNN 설정 정보가 포함될 수 있다. 도 20 및 도 21에서는 디블로킹 필터가 DNN인 경우에 AI 다운스케일 전에 수행되는 전처리 필터용 디블로킹 필터링을 위한 제 3 DNN과 AI 업스케일 후에 수행되는 후처리 필터용 디블로킹 필터링을 위한 제 4 DNN을 훈련시키는 방법과 훈련 과정이 후술된다.

[253] 이에 따라, 통신 환경의 변화가 발생하더라도, 초기에 이미 선택된 AI 다운스케일 모델 및 AI 스케일 모델을 변경하지 않고, 네트워크 환경, 원본 영상, 이전 프레임 정보에 따라 디블로킹 필터링을 수행함으로써, AI 다운스케일 모델 및 AI 스케일 모델을 유지한 채 영상의 품질을 조절할 수 있다.

[254] 일 실시예에 따라, 디블로킹 예측(1370)에서 획득된 원본 영상(1305), 이전 프레임 정보, 및 네트워크 환경 정보에 기초하여 DQP 맵 또는 레이트 컨트롤(rate control)에 대한 정보가 추가적으로 제 1 부호화(1330)에 전송되어 제 1 부호화를 수행하는데 이용될 수도 있다.

[255] 도 14는 일 실시예에 따른 AI 복호화 장치(1400)의 구성을 나타내는

블록도이다.

[256] 도 14를 참고하면, 일 실시예에 따른 AI 복호화 장치(1400)는 수신부(1410), AI 복호화부(1420), 및 디블로킹 예측부(1430)를 포함한다. AI 복호화부(230)는 파싱부(1421), 제 1 복호화부(1422), AI 업스케일부(1423), 디블로킹 필터부(1424) 및 AI 설정부(1425)를 포함할 수 있다.

[257] 도 14에는 수신부(1410), AI 복호화부(1420), 및 디블로킹 예측부(1430)가 개별적인 장치로 도시되어 있으나, 수신부(1410), AI 복호화부(1420), 및 디블로킹 예측부(1430)는 하나의 프로세서를 통해 구현될 수 있다. 이 경우, 수신부(1410), AI 복호화부(1420), 및 디블로킹 예측부(1430)는 전용 프로세서로 구현될 수도 있고, AP(application processor) 또는 CPU(central processing unit), GPU(graphic processing unit)와 같은 범용 프로세서와 S/W의 조합을 통해 구현될 수도 있다. 또한, 전용 프로세서의 경우, 본 개시의 실시예를 구현하기 위한 메모리를 포함하거나, 외부 메모리를 이용하기 위한 메모리 처리부를 포함할 수 있다.

[258] 수신부(1410), AI 복호화부(1420), 및 디블로킹 예측부(1430)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. 일 실시예에서, 수신부(1410)는 제 1 프로세서로 구현되고, 제 1 복호화부(1422)는 제 1 프로세서와 상이한 제 2 프로세서로 구현되고, 파싱부(1421), AI 업스케일부(1423), 디블로킹 필터부(1424) 및 AI 설정부(1425)는 제 1 프로세서 및 제 2 프로세서와 상이한 제 3 프로세서로 구현되고, 디블로킹 예측부(1430)는 제 1 프로세서, 제 2 프로세서, 및 제 3 프로세서와 상이한 제 3 프로세서로 구현될 수 있다.

[259] 수신부(1410)는 AI 부호화 결과 획득된 AI 부호화 데이터 및 AI 부호화 장치의 디블로킹 예측부에서 전송된 필터 설정 정보를 수신한다. 일 예로, AI 부호화 데이터는 mp4, mov 등의 파일 형식을 갖는 비디오 파일일 수 있다.

[260] 수신부(1410)는 네트워크를 통해 전달되는 AI 부호화 데이터 및 필터 설정 정보를 수신할 수 있다. 수신부(1410)는 AI 부호화 데이터 및 필터 설정 정보를 AI 복호화부(1420)로 출력한다.

[261] 다른 실시예에서, 수신부(1410)는 필터 설정 정보를 포함하는 AI 부호화 데이터를 수신할 수 있다.

[262] 일 실시예에서, 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.

[263] 파싱부(1421)는 AI 부호화 데이터를 파싱하여 제 1 영상(1325)의 제 1 부호화 결과로 생성된 영상 데이터를 제 1 복호화부(1422)로 전달하고, AI 데이터를 AI 설정부(1425)로 전달하고 필터 설정 정보를 디블로킹 예측부(1430)로 전달한다.

[264] 일 실시예에서, 디블로킹 예측부(1430)는 필터 설정 정보를 디블로킹 필터부(1424)에 전달한다. 필터 설정 정보는 블록 아티팩트가 발생되는 영역을 나타내는 블로킹 영역 맵을 포함할 수 있다.

[265] 다른 실시예에서, 필터 설정 정보는 블록 아티팩트가 발생되는 영역을 나타내는 블로킹 영역 맵 및 디블로킹 필터링을 위한 DNN의 DNN 설정 정보를 포함할 수 있다.

[266] 일 실시예에서, 파싱부(1421)는 AI 부호화 데이터 내에 서로 분리되어 포함된 영상 데이터와 AI 데이터를 파싱할 수 있다. 파싱부(1421)는 AI 부호화 데이터 내의 헤더를 읽어 AI 부호화 데이터 내에 포함되어 있는 AI 데이터와 영상 데이터를 구분할 수 있다. 일 예에서, AI 데이터는 HDMI 스트림 내의 VSIF(Vendor Specific InfoFrame)에 포함될 수 있다. 서로 분리된 AI 데이터와 영상 데이터를 포함하는 AI 부호화 데이터의 구조는 도 9를 참조하여 전술된 바 생략한다.

[267] 다른 실시예에서, 파싱부(1421)는 AI 부호화 데이터에서 영상 데이터를 파싱하고, 영상 데이터로부터 AI 데이터를 추출한 후, AI 데이터를 AI 설정부(1425)로 전달하고, 나머지 영상 데이터를 제 1 복호화부(1422)로 전달할 수 있다. 즉, AI 데이터는 영상 데이터에 포함될 수 있는데, 예를 들어, AI 데이터는 영상 데이터에 해당하는 비트스트림의 부가 정보 영역인 SEI(Supplemental enhancement information)에 포함될 수 있다. AI 데이터가 포함된 영상 데이터를 포함하는 AI 부호화 데이터의 구조는 도 10을 참조하여 전술된 바 생략한다.

[268] 다른 실시예에서, 파싱부(1421)는 AI 부호화 데이터 내에 서로 분리되어 포함된 영상 데이터, AI 데이터, 및 필터 설정 정보를 파싱하고, 제 1 영상(1325)의 제 1 부호화 결과로 생성된 영상 데이터를 제 1 복호화부(1422)로 출력하고, AI 데이터를 AI 설정부(1425)로 출력하고 필터 설정 정보를 디블로킹 예측부(1430)로 출력할 수 있다.

[269] 다른 실시예에서, 파싱부(1421)는 영상 데이터에 해당하는 비트스트림을 제 1 복호화부(1422)에서 처리될 비트스트림과 AI 데이터에 해당하는 비트스트림으로 분할하고, 분할된 각각의 비트스트림을 제 1 복호화부(1422)와 AI 설정부(1425)로 출력할 수 있다.

[270] 파싱부(1421)는 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, 해당 정보를 제 1 복호화부(1422)로 전달할 수 있다.

[271] 제 1 복호화부(1422)는 파싱부(1421)로부터 수신된 영상 데이터에 기초하여 제 1 영상(1325)에 대응하는 제 2 영상(1345)을 복원한다. 제 1 복호화부(1422)에 의해 획득된 제 2 영상(1345)은 AI 업스케일부(1423)로 제공된다.

[272] 구현예에 따라, 예측 모드 정보, 움직임 정보, 양자화 파라미터 정보 등의 제 1

복호화 관련 정보가 제 1 복호화부(1422)로부터 AI 설정부(1425)로 제공될 수 있다. 제 1 복호화 관련 정보는 DNN 설정 정보를 획득하는데 이용될 수 있다.

[273] AI 설정부(1425)로 제공되는 AI 데이터는, 제 2 영상(1345)을 AI 업스케일할 수 있게 하는 정보들을 포함한다. 이때, 제 2 영상(1345)의 업스케일 타겟은 제 1 DNN의 다운스케일 타겟에 대응하여야 한다. 따라서, AI 데이터는 제 1 DNN의 다운스케일 타겟을 확인할 수 있는 정보가 포함되어야 한다.

[274] AI 데이터에 포함된 정보를 구체적으로 예시하면, 원본 영상(1305)의 해상도와 제 1 영상(1325)의 해상도의 차이 정보, 제 1 영상(132) 관련 정보가 있다.

[275] 차이 정보는, 원본 영상(1305) 대비 제 1 영상(1325)의 해상도 변환 정도에 대한 정보(예를 들어, 해상도 변환율 정보)로 표현될 수 있다. 그리고, 복원된 제 2 영상(1345)의 해상도를 통해 제 1 영상(1325)의 해상도를 알게 되고 이를 통해 해상도 변환 정도를 확인할 수 있기 때문에, 차이 정보는 원본 영상(1305)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.

[276] 제 1 영상(1325) 관련 정보는, 제 1 영상(1325)의 해상도, 제 1 영상(1325)의 제 1 부호화 결과 획득된 영상 데이터의 비트레이트 및 제 1 영상(1325)의 제 1 부호화시 이용된 코덱 타입 중 적어도 하나에 대한 정보를 포함할 수 있다.

[277] AI 설정부(1425)는 AI 데이터에 포함된 차이 정보 및 제 1 영상(1325) 관련 정보 중 적어도 하나에 기초하여 제 2 영상(1345)의 업스케일 타겟을 결정할 수 있다. 업스케일 타겟은 예를 들어, 제 2 영상(1345)을 어느 정도의 해상도로 업스케일하여야 하는지를 나타낼 수 있다.

[278] AI 업스케일부(1423)는 업스케일 타겟이 결정되면, 업스케일 타겟에 대응하는 업스케일된 영상(1355)을 획득하기 위해 제 2 DNN을 통해 제 2 영상(1345)을 AI 업스케일한다.

[279] 디블로킹 필터부(1424)는 디블로킹 예측부(1430)로부터 획득된 필터 설정 정보에 기초하여 업스케일된 제 2 영상(1355)에 디블로킹 필터링을 적용하여 제 3 영상(1465)을 획득한다.

[280] 제 2 DNN을 통한 AI 업스케일 과정은 도 3 및 도 4에서 전술된 바 생략한다.

[281] 이하에서, AI 설정부(1425)가 AI 데이터에 기초하여 업스케일 타겟을 결정하고, AI 업스케일부(1423)가 업스케일 타겟에 맞춰 제 2 영상(1345)을 AI 업스케일하는 방법을 설명한다.

[282] 일 실시예에서, AI 설정부(1425)는 제 2 DNN에 세팅 가능한 복수의 DNN 설정 정보를 저장할 수 있다.

[283] 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.

[284] 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.

[285] 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.

[286] AI 설정부(1425)는 복수의 DNN 설정 정보 중 제 2 영상(1345)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 AI 업스케일된 제 2 영상(1355)을 획득하기 위한 정보로, 제 1 DNN과 연계하여 훈련된 것이다.

[287] 예를 들어, 복수의 DNN 설정 정보 중 어느 하나의 DNN 설정 정보는 제 2 영상(1345)의 해상도보다 2배 큰 해상도의 제 3 영상(1365), 예를 들어, 2K (2048*1080)의 제 2 영상(1345)보다 2배 큰 4K(4096*2160)의 제 3 영상(1365)을 획득하기 위한 정보들을 포함할 수 있고, 다른 하나의 DNN 설정 정보는 제 2 영상(1345)의 해상도보다 4배 큰 해상도의 제 3 영상(1365), 예를 들어, 2K (2048*1080)의 제 2 영상(1345)보다 4배 큰 8K(8192*4320)의 제 3 영상(1365)을 획득하기 위한 정보들을 포함할 수 있다.

[288] 복수의 DNN 설정 정보 각각은 AI 부호화 장치(1500)의 제 1 DNN의 DNN 설정 정보와 연계되어 만들어진 것이며, AI 설정부(1425)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 설정부(1425)는 제 1 DNN의 정보를 확인하여야 한다. 제 1 DNN의 정보를 AI 설정부(1425)가 확인하기 위해, 일 실시예에 따른 AI 복호화 장치(1400)는 AI 부호화 장치(1500)로부터 제 1 DNN의 정보를 포함하는 AI 데이터를 수신한다.

[289] 다시 말하면, AI 설정부(1425)는 AI 부호화 장치(1500)로부터 수신되는 정보들을 이용하여, 제 1 영상(1325)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.

[290] 복수의 DNN 설정 정보 중 제 2 영상(1345)의 AI 업스케일을 위한 DNN 설정 정보가 획득되면, DNN 설정 정보는 AI 업스케일부(1423)로 전달되고, DNN 설정 정보에 따라 동작하는 제 2 DNN에 기초하여 입력 데이터가 처리될 수 있다.

[291] 예를 들어, AI 업스케일부(1423)는 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.

[292] 구체적으로, AI 업스케일부(1423)는 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}이고, DNN 설정 정보의 변경이 있는 경우, 필터 커널의 파라미터들을 변경된 DNN 설정 정보에 포함된 파라미터들인 {2, 2, 2, 2, 2, 2, 2, 2, 2}로 교체할 수 있는 것이다.

[293] AI 설정부(1425)는 AI 데이터에 포함된 정보에 기초하여 복수의 DNN 설정 정보 중 제 2 영상(1345)을 업스케일하기 위한 DNN 설정 정보를 획득할 수 있는데, DNN 설정 정보를 획득하는데 이용되는 AI 데이터에 대해 구체적으로 설명한다.

[294] 일 실시예에서, AI 설정부(1425)는 AI 데이터에 포함된 차이 정보에 기초하여, 복수의 DNN 설정 정보 중 제 2 영상(1345)을 업스케일하기 위한 DNN 설정 정보를 획득할 수 있다. 예를 들어, 차이 정보에 기초하여 원본 영상(1305)의 해상도(예를 들어, 4K(4096*2160))가 제 1 영상(1325)의 해상도(예를 들어, 2K (2048*1080))보다 2배 큰 것으로 확인된 경우, AI 설정부(238)는 제 2 영상(1345)의 해상도를 2배 증가시킬 수 있는 DNN 설정 정보를 획득할 수 있다.

[295] 다른 실시예에서, AI 설정부(1425)는 AI 데이터에 포함된 제 1 영상(1325) 관련 정보에 기초하여, 복수의 DNN 설정 정보 중 제 2 영상(1345)을 AI 업스케일하기 위한 DNN 설정 정보를 획득할 수 있다. AI 설정부(1425)는 영상 관련 정보들과 DNN 설정 정보들 사이의 매핑 관계를 미리 결정하고, 제 1 영상(1325) 관련 정보에 매핑된 DNN 설정 정보를 획득할 수 있다.

[296] 도 15는 일 실시예에 따른 AI 부호화 장치(1500)의 구성을 나타내는 블록도이다.

[297] 도 15를 참조하면, AI 부호화 장치(1500)는 AI 부호화부(1510), 전송부(1520), 및 디블로킹 예측부(1530)를 포함할 수 있다. AI 부호화부(1510)는 디블로킹 필터부(1511), AI 다운스케일부(1512), 제 1 부호화부(1513), 데이터 처리부(1514) 및 AI 설정부(1515)를 포함할 수 있다.

[298] 도 15는 AI 부호화부(1510), 전송부(1520), 및 디블로킹 예측부(1530)를 개별적인 장치로 도시하고 있으나, AI 부호화부(1510), 전송부(1520), 및 디블로킹 예측부(1530)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 포함하거나, 외부 메모리를 이용하기 위한 메모리 처리부를 포함할 수 있다.

[299] 또한, AI 부호화부(1510), 전송부(1520), 및 디블로킹 예측부(1530)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. 일 실시예에서, 제 1 부호화부(1513)는 제 1 프로세서로

구성되고, 디블로킹 필터부(1511), AI 다운스케일부(1512), 데이터 처리부(1514) 및 AI 설정부(1515)는 제 1 프로세서와 상이한 제 2 프로세서로 구현되고, 전송부(1520)는 제 1 프로세서 및 제 2 프로세서와 상이한 제 3 프로세서로 구현되고, 디블로킹 예측부(1530)는 제 1 프로세서, 제 2 프로세서, 및 제 3 프로세서와 상이한 제 4 프로세서로 구현될 수 있다.

[300] 디블로킹 예측부(1530)는 원본 영상(1305), 이전 프레임 정보, 및 네트워크 환경 정보를 획득하고, 원본 영상(1305), 이전 프레임 정보, 및 네트워크 환경 정보에 기초하여 필터 설정 정보를 디블로킹 필터부(1511)에 전송한다.

[301] 디블로킹 예측부(1530)는 제 1 부호화부(1513)로부터 이전 프레임 정보를 획득하고, AI 복호화 장치(1400)로부터 실시간 프로토콜에 따라 네트워크 환경 정보를 획득한다.

[302] 일 실시예에서, 디블로킹 예측부(1530)는 필터 설정 정보를 디블로킹 필터부(1511)에 전달한다. 필터 설정 정보는 블록 아티팩트가 발생되는 영역을 나타내는 블로킹 영역 맵을 포함할 수 있다.

[303] 다른 실시예에서, 필터 설정 정보는 블록 아티팩트가 발생되는 영역을 나타내는 블로킹 영역 맵 및 디블로킹 필터링을 위한 DNN의 DNN 설정 정보를 포함할 수 있다.

[304] 디블로킹 필터부(1511)는 디블로킹 예측부(1530)로부터 획득된 필터 설정 정보에 기초하여 원본 영상(1305)에 디블로킹 필터링을 적용하여 디블로킹 필터링된 원본 영상(1315)을 획득한다.

[305] AI 부호화부(1513)는 디블로킹 필터링된 원본 영상(1315)의 AI 다운스케일 및 제 1 영상(1325)의 제 1 부호화를 수행하고, AI 부호화 데이터를 전송부(1520)로 전달한다. 전송부(1520)는 AI 부호화 데이터 및 필터 설정 정보를 AI 복호화 장치(1400)로 전송한다.

[306] 다른 실시예에서, 전송부(1520)는 필터 설정 정보, AI 데이터, 및 영상데이터를 포함하는 AI 부호화 데이터를 AI 복호화 장치(1400)로 전송한다.

[307] 영상 데이터는 제 1 영상(1325)의 제 1 부호화 결과 획득된 데이터를 포함한다. 영상 데이터는 제 1 영상(1325) 내 픽셀 값들에 기초하여 획득되는 데이터, 예를 들어, 제 1 영상(1325)과 제 1 영상(1325)의 예측 데이터 사이의 차이인 잔차 데이터를 포함할 수 있다. 또한, 영상 데이터는 제 1 영상(1325)의 제 1 부호화 과정에서 이용된 정보들을 포함한다. 예를 들어, 영상 데이터는 제 1 영상(1325)을 제 1 부호화하는데 이용된 예측 모드 정보, 움직임 정보 및 제 1 영상(1325)을 제 1 부호화하는데 이용된 양자화 파라미터 관련 정보 등을 포함할 수 있다.

[308] AI 데이터는, AI 업스케일부(1513)가 제 1 DNN의 다운스케일 타겟에 대응하는 업스케일 타겟으로 제 2 영상(1345)을 AI 업스케일할 수 있게 하는 정보들을 포함한다.

[309] 일 예에서, AI 데이터는 원본 영상(1305)과 제 1 영상(1325) 사이의 차이 정보를

포함할 수 있다.

[310] 일 예에서, AI 데이터는 제 1 영상(1325) 관련 정보를 포함할 수도 있다. 제 1 영상(1325) 관련 정보는, 제 1 영상(1325)의 해상도, 제 1 영상(1325)의 제 1 부호화 결과 획득된 영상 데이터의 비트레이트 및 제 1 영상(1325)의 제 1 부호화시 이용된 코덱 타입 중 적어도 하나에 대한 정보를 포함할 수 있다.

[311] 일 실시예에서, AI 데이터는, 제 1 DNN의 다운스케일 타겟에 대응하는 업스케일 타겟으로 제 2 영상(1345)이 AI 업스케일될 수 있도록, 상호 약속된 DNN 설정 정보의 식별자를 포함할 수 있다.

[312] 또한, 일 실시예에서, AI 데이터는 제 2 DNN에 세팅 가능한 DNN 설정 정보를 포함할 수도 있다.

[313] AI 다운스케일부(1512)는 디블로킹 필터링된 원본 영상(1315)으로부터 제 1 DNN을 통해 AI 다운스케일된 영상(1325)을 획득할 수 있다. AI 다운스케일부(1512)는 AI 설정부(1515)로부터 제공된 DNN 설정 정보를 이용하여 필터링된 원본 영상(1315)을 AI 다운스케일할 수 있다.

[314] AI 설정부(1515)는 미리 결정된 기준에 기초하여 원본 영상(105)의 다운스케일 타겟을 결정할 수 있다.

[315] 다운스케일 타겟에 부합하는 제 1 영상(1325)의 획득을 위해, AI 설정부(1515)는 제 1 DNN에 세팅 가능한 복수의 DNN 설정 정보를 저장할 수 있다. AI 설정부(1515)는 복수의 DNN 설정 정보 중 다운스케일 타겟에 대응하는 DNN 설정 정보를 획득하고, 획득된 DNN 설정 정보를 AI 다운스케일부(1512)로 제공한다.

[316] 상기 복수의 DNN 설정 정보 각각은 미리 결정된 해상도 및/또는 미리 결정된 화질의 제 1 영상(1325)을 획득하기 위해 훈련된 것일 수 있다. 예를 들어, 복수의 DNN 설정 정보 중 어느 하나의 DNN 설정 정보는 원본 영상(1305)의 해상도보다 1/2배만큼 작은 해상도의 제 1 영상(1325), 예를 들어, 4K(4096*2160)의 원본 영상(1305)보다 1/2배 작은 2K (2048*1080)의 제 1 영상(1325)을 획득하기 위한 정보들을 포함할 수 있고, 다른 하나의 DNN 설정 정보는 원본 영상(1305)의 해상도보다 1/4배만큼 작은 해상도의 제 1 영상(1325), 예를 들어, 8K(8192*4320)의 원본 영상(1305)보다 1/4배 작은 2K (2048*1080)의 제 1 영상(1325)을 획득하기 위한 정보들을 포함할 수 있다.

[317] 구현예에 따라, DNN 설정 정보를 구성하는 정보들(예를 들어, 컨볼루션 레이어의 수, 컨볼루션 레이어별 필터 커널의 개수, 각 필터 커널의 파라미터 등)이 룩업 테이블 형태로 저장되어 있는 경우, AI 설정부(1515)는 다운스케일 타겟에 따라 룩업 테이블 값들 중에서 선택된 일부를 조합하여 DNN 설정 정보를 획득하고, 획득한 DNN 설정 정보를 AI 다운스케일부(1512)로 제공할 수도 있다.

[318] 구현예에 따라, AI 설정부(1515)는 다운스케일 타겟에 대응되는 DNN의 구조를 결정하고, 결정된 DNN의 구조에 대응하는 DNN 설정 정보, 예를 들어, 필터 커널의 파라미터들을 획득할 수도 있다.

[319] 필터링된 원본 영상(1315)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.

[320] AI 다운스케일부(1512)는 필터링된 원본 영상(1315)의 AI 다운스케일을 위해 결정된 DNN 설정 정보로 제 1 DNN을 세팅하여, 소정 해상도 및/또는 소정 화질의 제 1 영상(1325)을 제 1 DNN을 통해 획득할 수 있다. 복수의 DNN 설정 정보 중 원본 영상(1305)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.

[321] AI 설정부(1515)가 다운스케일 타겟을 결정하는 방법에 대해 설명한다. 상기 다운스케일 타겟은 예를 들어, 디블로킹 필터링된 원본 영상(1315)으로부터 얼마나 해상도가 감소한 제 1 영상(1325)을 획득해야 하는지를 나타낼 수 있다.

[322] AI 설정부(1515)는 하나 이상의 입력 정보를 획득한다. 일 실시예에서, 입력 정보는 제 1 영상(1325)의 타겟 해상도, 영상 데이터의 타겟 비트레이트, 영상 데이터의 비트레이트 타입(예를 들어, variable bitrate 타입, constant bitrate 타입 또는 average bitrate 타입 등), AI 다운스케일이 적용되는 컬러 포맷(휘도 성분, 색차 성분, 레드 성분, 그린 성분 또는 블루 성분 등), 제 1 영상(1325)의 제 1 부호화를 위한 코덱 타입, 압축 히스토리 정보, 원본 영상(1305)의 해상도 및 원본 영상(1305)의 타입 중 적어도 하나를 포함할 수 있다.

[323] 하나 이상의 입력 정보는 AI 부호화 장치(1500)에 미리 저장되거나, 사용자로부터 입력받은 정보를 포함할 수 있다.

[324] AI 설정부(1515)는 입력 정보에 기초하여 AI 다운스케일부(1512)의 동작을 제어한다. 일 실시예에서, AI 설정부(1515)는 입력 정보에 따라 다운스케일 타겟을 결정하고, 결정된 다운스케일 타겟에 대응하는 DNN 설정 정보를 AI 다운스케일부(1512)로 제공할 수 있다.

[325] 일 실시예에서, AI 설정부(1515)는 입력 정보의 적어도 일부를 제 1 부호화부(1513)로 전달하여 제 1 부호화부(1513)가 특정 값의 비트레이트, 특정 타입의 비트레이트 및 특정 코덱으로 제 1 영상(1325)을 제 1 부호화하게 할 수도 있다.

[326] 다른 실시예에서, 디블로킹 예측부(1530)는 원본 영상(1305), 이전 프레임 정보, 및 네트워크 환경 정보에 기초하여 DQP 맵 또는 레이트 컨트롤(rate control)에 대한 정보를 추가로 획득하고, DQP 맵 또는 레이트 컨트롤(rate control)에 대한 정보를 제 1 부호화부(1513)에 전송할 수 있다. 또한, 제 1 부호화부(1513)는 제 1 영상(1325)을 제 1 부호화하는데 DQP 맵 또는 레이트 컨트롤(rate control)에 대한 정보를 이용할 수 있다.

[327] 일 실시예에서, AI 설정부(1515)는 압축률(예를 들어, 원본 영상(1305)과 제 1

영상(1325) 사이의 해상도 차이, 타겟 비트레이트), 압축 품질(예를 들어, 비트레이트 타입), 압축 히스토리 정보 및 원본 영상(1305)의 타입 중 적어도 하나에 기초하여 다운스케일 타겟을 결정할 수 있다.

[328] 일 예에서, AI 설정부(1515)는 미리 설정되거나, 사용자로부터 입력받은 압축률 또는 압축 품질 등에 기반하여 다운스케일 타겟을 결정할 수 있다.

[329] 다른 예로, AI 설정부(1515)는 AI 부호화 장치(1500)에 저장된 압축 히스토리 정보를 이용하여 다운스케일 타겟을 결정할 수도 있다. 예를 들어, AI 부호화 장치(1500)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.

[330] 또 다른 예로, AI 설정부(1515)는 압축 히스토리 정보에 따라 소정의 임계값보다 많이 이용된 적이 있는 부호화 품질(예를 들면, 소정의 임계값보다 많이 이용된 적이 있는 부호화 품질들의 평균 품질)에 기초하여 다운스케일 타겟을 결정할 수도 있다.

[331] 또 다른 예로, AI 설정부(1515)는 원본 영상(1305)의 해상도, 타입(예를 들어, 파일의 형식)등에 기초하여 다운스케일 타겟을 결정할 수도 있다.

[332] 일 실시예에서, 원본 영상(1305)이 복수의 프레임으로 구성된 경우, AI 설정부(1515)는 소정 개수의 프레임 별로 DNN 설정 정보를 독립적으로 획득하고, 독립적으로 획득한 DNN 설정 정보를 AI 다운스케일부(1512)로 제공할 수도 있다.

[333] 일 예시에서, AI 설정부(1515)는 원본 영상(1305)을 구성하는 프레임들을 소정 개수의 그룹으로 구분하고, 각 그룹별로 DNN 설정 정보를 독립적으로 획득할 수 있다. 각 그룹에 대해 서로 동일하거나 서로 상이한 DNN 설정 정보가 획득될 수 있다. 그룹들에 포함된 프레임들의 개수는 그룹별로 동일하거나 상이할 수 있다.

[334] 다른 예시에서, AI 설정부(1515)는 원본 영상(1305)을 구성하는 프레임별로 DNN 설정 정보를 독립적으로 결정할 수 있다. 각각의 프레임에 대해 서로 동일하거나 서로 상이한 DNN 설정 정보가 획득될 수 있다.

[335] 이하, 도 16 내지 도 18에서는 도 13과 비교하여 디블로킹 필터의 위치가 다른 실시예들이 후술된다.

[336] 도 16은 다른 실시예에 따른 디블로킹 필터링을 적응적으로 추가하는 AI 부호화 과정 및 AI 복호화 과정을 설명하기 위한 도면이다.

[337] 도 16은 도 13의 디블로킹 필터링을 적응적으로 추가하는 AI 부호화 과정 및 AI 복호화 과정과 비교하여 AI 부호화 과정에서의 디블로킹 필터의 위치가 달라진 실시예를 나타낸다.

[338] 도 16을 참고하면, 디블로킹 예측 과정(1670)에서, 현재 스트리밍 채널의 네트워크 환경 정보, 원본 영상(1605), 이전에 부호화된 프레임의 정보를

획득하여, 입력되는 원본 영상에 블로킹 현상이 발생할 지를 예측하고, 블록이 생기는 영역에 대한 블로킹 영역 맵과 같은 필터 설정 정보를 AI 부호화 및 AI 복호화가 수행될 때 전송한다. AI 부호화 과정에서, 원본 영상(1605)을 AI 다운스케일(1610)하여 제 1 영상(1615)을 획득하고, 디블로킹 예측 과정(1670)에서 획득된 필터 설정 정보에 기초하여 제 1 영상(1615)에 디블로킹 필터(1620)을 적용하여 디블로킹 필터링된 제 1 영상(1625)을 획득하고, 디블로킹 필터링된 제 1 영상(1625)을 제 1 부호화(1630)하여 전송한다, AI 복호화 과정에서는, AI 부호화 결과 획득된 AI 데이터와 영상 데이터를 포함하는 AI 부호화 데이터를 수신하고, 제 1 복호화(1640)를 통해 제 2 영상(1645)을 획득하고, 제 2 영상(1645)을 AI 업스케일(1650)하여 AI 업스케일된 영상(1355)을 획득하고, 디블로킹 예측(1370)에서 전송된 필터 설정 정보에 따른 디블로킹 필터(1660)을 적용하여 제 3 영상(1665)을 획득한다.

[339] 도 17은 다른 실시예에 따른 디블로킹 필터링을 적응적으로 추가하는 AI 부호화 과정 및 AI 복호화 과정을 설명하기 위한 도면이다.

[340] 도 17은 도 13의 디블로킹 필터링을 적응적으로 추가하는 AI 부호화 과정 및 AI 복호화 과정과 비교하여 AI 부호화 과정에서의 디블로킹 필터 및 AI 복호화 과정에서의 디블로킹 필터의 위치가 모두 달라진 실시예를 나타낸다.

[341] 도 17을 참고하면, 디블로킹 예측 과정(1770)에서, 현재 스트리밍 채널의 네트워크 환경 정보, 원본 영상(1705), 이전에 부호화된 프레임의 정보를 획득하여, 입력되는 원본 영상에 블로킹 현상이 발생할 지를 예측하고, 블록이 생기는 영역에 대한 블로킹 영역 맵과 같은 필터 설정 정보를 AI 부호화 및 AI 복호화가 수행될 때 전송한다. AI 부호화 과정에서, 원본 영상(1705)을 AI 다운스케일(1710)하여 제 1 영상(1715)을 획득하고, 디블로킹 예측 과정(1770)에서 획득된 필터 설정 정보에 기초하여 제 1 영상(1715)에 디블로킹 필터(1720)을 적용하여 디블로킹 필터링된 제 1 영상(1725)을 획득하고, 디블로킹 필터링된 제 1 영상(1725)을 제 1 부호화(1730)하여 전송한다, AI 복호화 과정에서는, AI 부호화 결과 획득된 AI 데이터와 영상 데이터를 포함하는 AI 부호화 데이터를 수신하고, 제 1 복호화(1740)를 통해 제 2 영상(1745)을 획득하고, 제 2 영상(1745)을 디블로킹 예측(1370)에서 전송된 필터 설정 정보에 따른 디블로킹 필터(1750)을 적용하여 디블로킹 필터링된 제 2 영상(1755)을 획득하고, 디블로킹 필터링된 제 2 영상(1755)을 AI 업스케일(1760)하여 제 3 영상(1765)을 획득한다.

[342] 도 18은 다른 실시예에 따른 디블로킹 필터링을 적응적으로 추가하는 AI 부호화 과정 및 AI 복호화 과정을 설명하기 위한 도면이다.

[343] 도 18은 도 13의 디블로킹 필터링을 적응적으로 추가하는 AI 부호화 과정 및 AI 복호화 과정과 비교하여 AI 복호화 과정에서의 디블로킹 필터의 위치가 달라진 실시예를 나타낸다.

[344] 도 18을 참고하면, 디블로킹 예측 과정(1870)에서, 현재 스트리밍 채널의

네트워크 환경 정보, 원본 영상(1805), 이전에 부호화된 프레임의 정보를 획득하여, 입력되는 원본 영상에 블로킹 현상이 발생할 지를 예측하고, 블록이 생기는 영역에 대한 블로킹 영역 맵과 같은 필터 설정 정보를 AI 부호화 및 AI 복호화가 수행될 때 전송한다. AI 부호화 과정에서, 디블로킹 예측 과정(1870)에서 획득된 필터 설정 정보에 기초하여 원본 영상(1805)에 디블로킹 필터(1810)을 적용하여 디블로킹 필터링된 원본 영상(1815)을 AI 다운스케일(1820)하여 제 1 영상(1825)을 획득하고, 제 1 영상(1825)을 제 1 부호화(1830)하여 전송한다, AI 복호화 과정에서는, AI 부호화 결과 획득된 AI 데이터와 영상 데이터를 포함하는 AI 부호화 데이터를 수신하고, 제 1 복호화(1840)를 통해 제 2 영상(1845)을 획득하고, 제 2 영상(1845)을 디블로킹 예측 과정(1870)에서 전송된 필터 설정 정보에 따른 디블로킹 필터(1850)을 적용하여 디블로킹 필터링된 제 2 영상(1855)을 획득하고, 디블로킹 필터링된 제 2 영상(1855)을 AI 업스케일(1860)하여 제 3 영상(1865)을 획득한다.

[345] 이하에서는, 도 13의 디블로킹 필터링이 레거시 디블로킹 필터가 아닌 DNN으로 훈련된 AI 디블로킹 필터인 경우에 디블로킹 필터링을 위한 DNN이 후술된다.

[346] 도 19는 AI에 기반한 디블로킹 필터를 위한 제 3 DNN 및 제 4 DNN을 나타내는 예시적인 도면이다.

[347] 도 19에 도시된 바와 같이, 원본 영상(1305) 또는 업스케일된 제 2 영상(1355)는 제 1 컨볼루션 레이어(1910)로 입력된다. 제 1 컨볼루션 레이어(1910)는 3 x 3의 크기의 4개의 필터 커널을 이용하여 원본 영상(1305) 또는 업스케일된 제 2 영상(1355)에 대해 컨볼루션 처리를 한다. 컨볼루션 처리 결과 4개의 특징 맵이 생성된다. 각 특징 맵은 원본 영상(1305) 또는 업스케일된 제 2 영상(1355)의 고유한 특성들을 나타낸다. 예를 들어, 각 특징 맵은 원본 영상(1305) 또는 업스케일된 제 2 영상(1355)의 수직 방향 특성, 수평 방향 특성, 에지 특성, 또는 블로킹 아티팩트의 발생 가능성이 있는지 등을 나타낼 수 있다. 생성된 4개의 특징 맵은 제 1 활성화 레이어(1920)로 입력된다.

[348] 제 1 활성화 레이어(1920)는 4개의 특징 맵에 대해 비선형(Non-linear) 특성을 부여할 수 있다.

[349] 제 1 활성화 레이어(1920)는 제 1 컨볼루션 레이어(1910)로부터 출력되는 특징 맵들의 샘플 값들을 제 2 컨볼루션 레이어(1930)로 전달할지 여부를 결정한다. 예를 들어, 특징 맵들의 샘플 값들 중 어떤 샘플 값들은 제 1 활성화 레이어(1920)에 의해 활성화되어 제 2 컨볼루션 레이어(1930)로 전달되고, 어떤 샘플 값들은 제 1 활성화 레이어(1920)에 의해 비활성화되어 제 2 컨볼루션 레이어(1930)로 전달되지 않는다. 제 1 컨볼루션 레이어(1910)로부터 출력되는 특징 맵들이 나타내는 정보가 제 1 활성화 레이어(1920)에 의해 강조된다.

[350] 제 1 활성화 레이어(1920)의 출력(1925)이 제 2 컨볼루션 레이어(1930)로 입력된다. 제 2 컨볼루션 레이어(1930)는 3 x 3의 크기의 4개의 필터 커널을

이용하여 입력 데이터에 대해 컨볼루션 처리를 한다. 컨볼루션 처리 결과 출력된 4개의 특징 맵은 제 2 활성화 레이어(1940)로 입력되고, 제 2 활성화 레이어(1940)는 4개의 특징 맵에 대해 비선형 특성을 부여할 수 있다.

[351] 제 2 활성화 레이어(1940)의 출력(1945)은 제 3 컨볼루션 레이어(1950)로 입력된다. 제 3 컨볼루션 레이어(1950)는 3 x 3의 크기의 1개의 필터 커널을 이용하여 입력된 데이터에 대해 컨볼루션 처리를 한다. 컨볼루션 처리 결과 제 3 컨볼루션 레이어(1950)로부터 1개의 영상이 출력될 수 있다. 제 3 컨볼루션 레이어(1950)는 최종 영상을 출력하기 위한 레이어로서 1개의 필터 커널을 이용하여 1개의 출력을 획득한다. 본 개시의 예시에 따르면, 제 3 컨볼루션 레이어(1950)는 컨벌루션 연산 결과를 통해 디블로킹 필터링된 원본 영상(1315) 또는 제 3 영상(1365)을 출력할 수 있다.

[352] 제 3 DNN 또는 제 4 DNN의 제 1 컨볼루션 레이어(1910), 제 2 컨볼루션 레이어(1930) 및 제 3 컨볼루션 레이어(1950)의 필터 커널의 개수, 필터 커널의 파라미터 등을 나타내는 DNN 설정 정보는 복수 개일 수 있다.

[353] 제 3 DNN과 제 4 DNN의 복수의 DNN 설정 정보는 서로 연계되어야 한다. 제 3 DNN의 복수의 DNN 설정 정보와 제 4 DNN의 복수의 DNN 설정 정보 사이의 연계는, 제 3 DNN 및 제 4 DNN의 연계 학습을 통해 구현될 수 있다.

[354] 도 19는 제 3 DNN 또는 제 4 DNN이 세 개의 컨볼루션 레이어(1910, 1930, 1950)와 두 개의 활성화 레이어(1920, 1940)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 3 DNN 또는 제 4 DNN은 RNN(recurrent neural network)을 통해 구현될 수도 있다. 이 경우는 본 개시의 예시에 따른 제 3 DNN 또는 제 4 DNN의 CNN 구조를 RNN 구조로 변경하는 것을 의미한다.

[355] 일 실시예에서, 디블로킹 필터부(1424) 또는 디블로킹 필터부(1511)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(1305), 업스케일된 제 2 영상(1355) 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.

[356] 이하에서는, 도 20을 참조하여, 제 3 DNN(2010)과 제 4 DNN(2020)을 연계 훈련시키는 방법에 대해 설명한다. 제 3 DNN(2010)과 제 4 DNN(2020)의 연계 훈련에 이용되는 제 1 DNN(800) 및 제 2 DNN(300)은 도 11 및 도 12에서 전술된 방식으로 훈련된 DNN이다.

[357] 도 20은 제 3 DNN(2010) 및 제 4 DNN(2020)을 훈련시키는 방법을 설명하기 위한 도면이다.

[358] 일 실시예에서, 스트리밍 환경의 변화가 있는 경우(사용자가 많이 몰리거나 실외에서 통신 환경의 변화가 있는 경우) 디블로킹 예측 과정을 통해 원본 영상(1305), 이전 프레임 정보, 및 네트워크 환경 정보를 획득하고, 원본 영상(1305), 이전 프레임 정보, 및 네트워크 환경 정보에 기초하여 디블로킹 필터 설정 정보를 AI 부호화 및 AI 복호화를 위해 전송하고, AI 부호화 과정을 통해, 디블로킹 예측 과정을 통해 전송된 디블로킹 필터 설정 정보에 기초하여 원본 영상이 디블로킹 필터링되고 AI 다운스케일된 제 1 영상(1325)가 제 1 부호화되어 전송되고, 제 1 복호화된 제 2 영상(1345)이 AI 업스케일되고, 디블로킹 예측 과정을 통해 전송된 필터 설정 정보에 기초하여 디블로킹 필터링된 제 3 영상(1365)으로 복원된다.

[359] 도 1 내지 도 12에서 전술된 AI 부호화 및 AI 복호화 과정은 스트리밍 환경이 유지되는 경우에 제 1 DNN과 제 2 DNN의 DNN 설정 정보가 초기 네트워크 환경에 따른 해상도와 비트레이트에 적합하게 선택되어 추가적인 디블로킹 필터 없이 영상이 전송되지만, 스트리밍 환경에 문제가 생기는 경우에 제 1 DNN과 제 2 DNN의 DNN 설정 정보 변경 없이 추가적인 디블로킹 필터들을 통해 영상의 품질을 유지하여 영상을 전송할 수 있다.

[360] 따라서, 디블로킹 필터링을 위한 제 3 DNN 및 제 4 DNN은 특정 해상도와 특정 비트레이트에 대응하는 훈련된 제 1 DNN과 훈련된 제 2 DNN 각각의 DNN 설정 정보로 AI 다운스케일 및 AI 업스케일되고, 네트워크 환경의 변화를 가정해서 네트워크 채널의 대역폭 또는 비트레이트가 일부 변경된 조건 하에서 훈련된다.

[361] 이에 따라, 디블로킹 필터링이 추가된 AI 부호화 및 AI 복호화 결과 획득된 제 3 영상(1365)과 원본 영상(1305)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 3 DNN(2010)과 제 4 DNN(2020)의 연계 훈련이 요구된다.

[362] 정확한 디블로킹 필터링 및 AI 복호화를 위해서는 궁극적으로 도 20에 도시된 제 3 훈련 영상(2004)과 원본 훈련 영상(2001) 사이의 비교 결과에 대응하는 퀄리티 손실 정보(2050)를 감소시킬 필요가 있다. 따라서, 퀄리티 손실 정보(2050)는 제 3 DNN(2010) 및 제 4 DNN(2020)의 훈련 모두에 이용된다.

[363] 먼저, 도 20에 도시된 훈련 과정에 대해 설명한다.

[364] 도 20에서, 원본 훈련 영상(original training image)(2001)은 디블로킹 필터링 및 AI 다운스케일의 대상이 되는 영상이고, 제 1 훈련 영상(first training image)(2002)은 원본 훈련 영상(1101)로부터 디블로킹 필터링 및 AI 다운스케일된 영상이다. 또한, 제 3 훈련 영상(third training image)(2004)은 제 1 훈련 영상(2002)으로부터 AI 업스케일되고 디블로킹 필터링된 영상이다.

[365] 원본 훈련 영상(2001)은 정지 영상 또는 복수의 프레임으로 이루어진 동영상을

포함한다. 일 실시예에서, 원본 훈련 영상(2001)은 정지 영상 또는 복수의 프레임으로 이루어진 동영상으로부터 추출된 휘도 영상을 포함할 수도 있다. 또한, 일 실시예에서, 원본 훈련 영상(2001)은 정지 영상 또는 복수의 프레임으로 이루어진 동영상에서 추출된 패치 영상을 포함할 수도 있다. 원본 훈련 영상(2001)이 복수의 프레임으로 이루어진 경우, 제 1 훈련 영상(2002), 제 2 훈련 영상 및 제 3 훈련 영상(2004) 역시 복수의 프레임으로 구성된다. 원본 훈련 영상(2001)의 복수의 프레임이 순차적으로 제 3 DNN(2010) 및 훈련된 제 1 DNN(800)에 입력되면, 제 3 DNN(2010), 훈련된 제 1 DNN(800), 훈련된 제 2 DNN(300), 및 제 4 DNN(2020)을 통해 제 1 훈련 영상(2002), 제 2 훈련 영상 및 제 3 훈련 영상(2004)의 복수의 프레임이 순차적으로 획득될 수 있다.

[366] 제 3 DNN(2010) 및 제 4 DNN(2020)의 연계 훈련을 위해, 원본 훈련 영상(2001)이 제 3 DNN(2010) 및 훈련된 제 1 DNN(800)으로 입력된다. 제 3 DNN(2010) 및 훈련된 제 1 DNN(800)으로 입력된 원본 훈련 영상(2001)은 디블로킹 필터링되고 AI 다운스케일되어 제 1 훈련 영상(2002)으로 출력되고, 제 1 훈련 영상(2002)이 훈련된 제 2 DNN(300) 및 제 4 DNN(2020)에 입력된다. 제 1 훈련 영상(2002)에 대한 AI 업스케일 및 디블로킹 필터링 결과 제 3 훈련 영상(2004)이 출력된다.

[367] 도 20을 참조하면, 훈련된 제 2 DNN(300)으로 제 1 훈련 영상(2002)이 입력되고 있는데, 구현예에 따라, 제 1 훈련 영상(2002)의 제 1 부호화 및 제 1 복호화 과정을 거쳐 획득된 제 2 훈련 영상(second training image)이 훈련된 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(2002)의 제 1 부호화 및 제 1 훈련 영상(2002)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.

[368] 도 20을 참조하면, 제 3 DNN(2010) 및 훈련된 제 1 DNN(800)을 통해 제 1 훈련 영상(2002)이 출력되는 것과 별개로, 원본 훈련 영상(2001)으로부터 레거시 다운스케일된 축소 훈련 영상(2003)이 획득된다. 여기서, 레거시 다운스케일은 바이리니어(bilinear) 스케일, 바이큐빅(bicubic) 스케일, 란조스(lanczos) 스케일 및 스테어 스탭(stair step) 스케일 중 적어도 하나를 포함할 수 있다.

[369] 원본 영상(1305)의 구조적 특징을 기준으로 제 1 영상(1325)의 구조적 특징이 크게 벗어나는 것을 방지하기 위해, 원본 훈련 영상(2001)의 구조적 특징을 보존하는 축소 훈련 영상(2003)을 획득하는 것이다.

[370] 훈련의 진행 전 제 3 DNN(2010) 및 제 4 DNN(2020)은 미리 결정된 DNN 설정 정보로 세팅될 수 있다. 훈련이 진행됨에 따라 구조적 손실 정보(2030), 복잡성 손실 정보(2040) 및 퀄리티 손실 정보(2050)가 결정될 수 있다.

[371] 구조적 손실 정보(2030)는 축소 훈련 영상2003)과 제 1 훈련 영상(1102)의 비교 결과에 기초하여 결정될 수 있다. 일 예에서, 구조적 손실 정보(2030)는 축소 훈련

영상(2003)의 구조적 정보와 제 1 훈련 영상(2002)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(2030)는 원본 훈련 영상(2001)의 구조적 정보가 제 1 훈련 영상(2002)에서 어느 정도로 유지되고 있는지를 나타낸다. 구조적 손실 정보(2030)가 작을수록 제 1 훈련 영상(2002)의 구조적 정보가 원본 훈련 영상(2001)의 구조적 정보와 유사해진다.

[372] 복잡성 손실 정보(2040)는 제 1 훈련 영상(2002)의 공간적 복잡도에 기반하여 결정될 수 있다. 일 예에서, 공간적 복잡도로서, 제 1 훈련 영상(2002)의 총 분산(total variance)값이 이용될 수 있다. 복잡성 손실 정보(2040)는 제 1 훈련 영상(2002)을 제 1 부호화하여 획득한 영상 데이터의 비트레이트와 관련된다. 복잡성 손실 정보(2040)가 작을수록 영상 데이터의 비트레이트가 작은 것으로 정의된다.

[373] 퀄리티 손실 정보(2050)는 원본 훈련 영상(2001)과 제 3 훈련 영상(2004)의 비교 결과에 기초하여 결정될 수 있다. 퀄리티 손실 정보(2050)는 원본 훈련 영상(2001)과 제 3 훈련 영상(2004)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(2050)는 제 3 훈련 영상(2004)이 원본 훈련 영상(2001)과 어느 정도로 유사한지를 나타낸다. 퀄리티 손실 정보(2050)가 작을수록 제 3 훈련 영상(2004)이 원본 훈련 영상(2001)에 더 유사해진다.

[374] 도 20을 참조하면, 구조적 손실 정보(2030), 복잡성 손실 정보(2040) 및 퀄리티 손실 정보(2050)가 제 3 DNN(2010)의 훈련에 이용되고, 퀄리티 손실 정보(2050)는 제 4 DNN(2020)의 훈련에 이용된다. 즉, 퀄리티 손실 정보(2050)는 제 3 DNN(2010) 및 제 4 DNN(2020)의 훈련에 모두 이용된다.

[375] 제 3 DNN(2010)은 구조적 손실 정보(2030), 복잡성 손실 정보(2040) 및 퀄리티 손실 정보(2050)에 기초하여 결정된 최종 손실 정보가 감소 또는 최소화되도록 파라미터를 갱신할 수 있다. 또한, 제 4 DNN(2020)은 퀄리티 손실 정보(2050)가 감소 또는 최소화되도록 파라미터를 갱신할 수 있다.

[376] 제 3 DNN(2010) 및 제 4 DNN(2020)의 훈련을 위한 최종 손실 정보는 아래의 수학식 2와 같이 결정될 수 있다.

[377] [수학식 2]

[378] LossDDS = e*구조적 손실 정보+f*복잡성 손실 정보+g*퀄리티 손실 정보

[379] LossDUS = h*퀄리티 손실 정보

[380] 상기 수학식 2에서, LossDDS는 제 3 DNN(2010)의 훈련을 위해 감소 또는 최소화되어야 할 최종 손실 정보를 나타내고, LossDUS는 제 4 DNN(2020)의 훈련을 위해 감소 또는 최소화되어야 할 최종 손실 정보를 나타낸다. 또한, e, f, g, h는 미리 결정된 소정의 가중치에 해당할 수 있다.

[381] 즉, 제 3 DNN(2010)은 수학식 2의 LossDDS가 감소되는 방향으로 파라미터들을 갱신하고, 제 4 DNN(2020)은 LossDUS가 감소되는 방향으로 파라미터들을 갱신하게 된다. 훈련 과정에서 도출된 LossDDS에 따라 제 3 DNN(2010)의 파라미터들이 갱신되면, 갱신된 파라미터에 기초하여 획득되는 제 1 훈련 영상(2002)이 이전 훈련 과정에서의 제 1 훈련 영상(2002)과 달라지게 되고, 그에 따라 제 3 훈련 영상(2004) 역시 이전 훈련 과정에서의 제 3 훈련 영상(2004)과 달라지게 된다. 제 3 훈련 영상(2004)이 이전 훈련 과정에서의 제 3 훈련 영상(2004)과 달라지게 되면, 퀄리티 손실 정보(2050) 역시 새롭게 결정되며, 그에 따라 제 4 DNN(2020)은 파라미터들을 갱신한다. 퀄리티 손실 정보(2050)가 새롭게 결정되면, LossDDS 역시 새롭게 결정되므로, 제 3 DNN(2010)은 새롭게 결정된 LossDDS에 따라 파라미터들을 갱신한다. 즉, 제 3 DNN(2010)의 파라미터 갱신은, 제 4 DNN(2020)의 파라미터 갱신을 야기하고, 제 4 DNN(2020)의 파라미터 갱신은 제 3 DNN(2010)의 파라미터 갱신을 야기하는 것이다. 다시 말하면, 제 3 DNN(2010) 및 제 4 DNN(2020)은 퀄리티 손실 정보(2050)의 공유를 통해 연계 훈련되므로, 제 3 DNN(2010)의 파라미터들과 제 4 DNN(2020)의 파라미터들이 서로 연관성을 가지고 최적화될 수 있는 것이다.

[382] 수학식 2를 참조하면, LossDUS가 퀄리티 손실 정보(2050)에 따라 결정되는 것을 알 수 있으나, 이는 하나의 예시이며, LossDUS는 구조적 손실 정보(2030) 및 복잡성 손실 정보(2040) 중 적어도 하나와, 퀄리티 손실 정보(2050)에 기초하여 결정될 수도 있다.

[383] 앞서, AI 복호화 장치(1400)의 디블로킹 예측부(1430) 및 AI 부호화 장치(1500)의 디블로킹 예측부(1530)는 필터 설정 정보를 저장하는 것으로 설명하였는데, 필터 설정 정보는 복수의 DNN 설정 정보도 포함할 수 있다. 디블로킹 예측부(1430) 및 디블로킹 예측부(1530)에 저장되는 복수의 DNN 설정 정보 각각을 훈련시키는 방법에 대해 설명한다.

[384] 수학식 2와 관련하여 설명한 바와 같이, 제 3 DNN(2010)의 경우, 제 1 훈련 영상(2002)의 구조적 정보와 원본 훈련 영상(2001)의 구조적 정보 사이의 유사 정도(구조적 손실 정보(2030)), 제 1 훈련 영상(2002)의 제 1 부호화 결과 획득되는 영상 데이터의 비트레이트(복잡성 손실 정보(2040)) 및 제 3 훈련 영상(2004)과 원본 훈련 영상(2001) 사이의 차이(퀄리티 손실 정보(2050))를 고려하여 파라미터를 갱신하게 된다.

[385] 자세히 설명하면, 원본 훈련 영상(2001)의 구조적 정보와 유사하면서, 제 1 부호화를 하였을 때 획득되는 영상 데이터의 비트레이트가 작은 제 1 훈련 영상(2002)이 획득 가능하도록 하는 동시에, 제 1 훈련 영상(2002)을 AI 업스케일하는 제 4 DNN(2020)이 원본 훈련 영상(2001)에 유사한 제 3 훈련 영상(2004)을 획득할 수 있도록, 제 3 DNN(2010)의 파라미터가 갱신될 수 있다.

[386] 수학식 2의 e, f, g의 가중치가 조절됨으로써, 제 3 DNN(2010)의 파라미터들이 최적화되는 방향이 상이해지게 된다. 예를 들어, f의 가중치를 높게 결정하는

경우, 제 3 훈련 영상(2004)의 퀄리티보다 비트레이트가 낮아지는 것에 더 중요도를 두고 제 3 DNN(2010)의 파라미터가 갱신될 수 있다. 또한, g의 가중치를 높게 결정하는 경우, 비트레이트가 높아지는 것이나, 원본 훈련 영상(2001)의 구조적 정보가 유지되는 것보다 제 3 훈련 영상(2004)의 퀄리티가 증가하도록 하는 것에 더 중요도를 두고 제 3 DNN(2010)의 파라미터가 갱신될 수 있다.

[387] 또한, 제 1 훈련 영상(2002)을 제 1 부호화하는데 이용되는 코덱의 타입에 따라 제 3 DNN(2010)의 파라미터들이 최적화되는 방향이 상이해질 수 있다. 왜냐하면, 코덱의 종류에 따라, 제 4 DNN(2020)으로 입력될 제 2 훈련 영상이 달라질 수 있기 때문이다.

[388] 즉, 가중치 e, 가중치 f, 가중치 g 및 제 1 훈련 영상(2002)의 제 1 부호화를 위한 코덱의 종류에 기반하여 제 3 DNN(2010)의 파라미터들과 제 4 DNN(2020)의 파라미터들이 연계하여 갱신될 수 있는 것이다. 따라서, 가중치 e, 가중치 f, 가중치 g 각각을 소정의 값으로 결정하고, 코덱의 종류를 소정의 종류로 결정한 후, 제 3 DNN(2010)과 제 4 DNN(2020)을 훈련시키면, 서로 연계되어 최적화된 제 3 DNN(2010)의 파라미터들과 제 4 DNN(2020)의 파라미터들이 결정될 수 있다.

[389] 그리고, 가중치 e, 가중치 f, 가중치 g 및 코덱의 종류를 변경한 후, 제 3 DNN(2010)과 제 4 DNN(2020)을 훈련시키면, 서로 연계되어 최적화된 제 3 DNN(2010)의 파라미터들과 제 4 DNN(2020)의 파라미터들이 결정될 수 있다. 다시 말하면, 가중치 e, 가중치 f, 가중치 g 및 코덱의 종류 각각의 값을 변경하면서 제 3 DNN(2010)과 제 4 DNN(2020)을 훈련시키면 서로 연계되어 훈련된 복수의 DNN 설정 정보가 제 3 DNN(2010)과 제 4 DNN(2020)에서 결정될 수 있는 것이다.

[390] 도 21는 훈련 장치(2100)에 의한 제 3 DNN(2010) 및 제 4 DNN(2020)의 훈련 과정을 설명하기 위한 도면이다.

[391] 도 20과 관련하여 설명한 제 3 DNN(2010) 및 제 4 DNN(2020)의 훈련은 훈련 장치(2100)에 의해 수행될 수 있다. 훈련 장치(2100)는 제 3 DNN(2010), 제 4 DNN(2020), 훈련된 제 1 DNN(800) 및 훈련된 제 2 DNN(300)을 포함한다. 훈련 장치(2100)는 예를 들어, AI 부호화 장치(1500) 또는 별도의 서버일 수 있다. 훈련 결과 획득된 제 4 DNN(2020)의 DNN 설정 정보들은 AI 복호화 장치(1400)에 저장된다.

[392] 도 21을 참조하면, 훈련 장치(2100)는 제 3 DNN(2010) 및 제 4 DNN(2020)의 DNN 설정 정보를 초기 세팅한다(S2110, S2115). 이에 의해, 제 3 DNN(2010) 및 제 4 DNN(2020)은 미리 결정된 DNN 설정 정보에 따라 동작할 수 있다. DNN 설정 정보는 제 3 DNN(2010) 및 제 4 DNN(2020)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.

[393] 훈련 장치(2100)는 원본 훈련 영상(2001)을 제 3 DNN(2010)으로

입력한다(S2120). 원본 훈련 영상(1101)은 정지 영상 또는 동영상을 구성하는 적어도 하나의 프레임을 포함할 수 있다. 제 3 DNN(2010)에서 필터링 결과가 출력되어 훈련된 제 1 DNN에 입력된다(S2125).

[394] 제 3 DNN(2010)은 초기 세팅된 DNN 설정 정보에 따라 원본 훈련 영상(2001)을 처리하고, 원본 훈련 영상(2001)으로부터 디블로킹 필터링되고 훈련된 제 1 DNN(800)을 통해 AI 다운스케일된 제 1 훈련 영상(2002)을 출력한다(S2130). 도 21은 제 3 DNN(2010) 및 훈련된 제 1 DNN(800)으로부터 출력된 제 1 훈련 영상(2002)이 훈련된 제 2 DNN(300)으로 바로 입력되는 것으로 도시되어 있으나, 제 3 DNN(2010) 및 훈련된 제 1 DNN(800)으로부터 출력된 제 1 훈련 영상(2002)이 훈련 장치(2100)에 의해 훈련된 제 2 DNN(300)으로 입력될 수 있다. 또한, 훈련 장치(2100)는 제 1 훈련 영상(2002)을 소정의 코덱으로 제 1 부호화 및 제 1 복호화한 후, 제 2 훈련 영상을 훈련된 제 2 DNN(300)으로 입력할 수 있다.

[395] 훈련된 제 2 DNN(300)은 제 1 훈련 영상(2002) 또는 제 2 훈련 영상을 AI 업스케일(S2135)하고, 제 1 훈련 영상(2002) 또는 제 2 훈련 영상으로부터 디블로킹 필터링된 제 3 훈련 영상(2004)을 출력한다(S2140).

[396] 훈련 장치(2100)는 제 1 훈련 영상(2002)에 기초하여 복잡성 손실 정보(2040)를 산출한다(S2145).

[397] 훈련 장치(2100)는 축소 훈련 영상(2003)과 제 1 훈련 영상(2002)을 비교하여 구조적 손실 정보(2030)를 산출한다(S2150).

[398] 훈련 장치(2100)는 원본 훈련 영상(2001)과 제 3 훈련 영상(2004)을 비교하여 퀄리티 손실 정보(2050)를 산출한다(S1255).

[399] 제 3 DNN(2010)은 최종 손실 정보에 기초한 역전사(back propagation) 과정을 통해 초기 세팅된 DNN 설정 정보를 갱신한다(S2160). 훈련 장치(2100)는 복잡성 손실 정보(2040), 구조적 손실 정보(2030) 및 퀄리티 손실 정보(2050)에 기초하여 제 3 DNN(2010)의 훈련을 위한 최종 손실 정보를 산출할 수 있다.

[400] 제 4 DNN(2020)은 퀄리티 손실 정보 또는 최종 손실 정보에 기초한 역전사 과정을 통해 초기 세팅된 DNN 설정 정보를 갱신한다(S2165). 훈련 장치(2100)는 퀄리티 손실 정보(2050)에 기초하여 제 4 DNN(2020)의 훈련을 위한 최종 손실 정보를 산출할 수 있다.

[401] 이후, 훈련 장치(2100), 제 3 DNN(2010) 및 제 4 DNN(2020)은 최종 손실 정보들이 최소화될 때까지 S2110 내지 S2165 과정을 반복하면서 DNN 설정 정보를 갱신한다. 이 때, 각 반복 과정 동안 제 3 DNN(2010) 및 제 4 DNN(2020)은 이전 과정에서 갱신된 DNN 설정 정보에 따라 동작한다.

[402] 도 22는 일 실시예에 따른 AI 부호화 방법을 설명하기 위한 순서도이다.

[403] S2210 단계에서, AI 부호화 장치(1500)는 원본 영상, 이전에 부호화된 프레임 정보, 및 네트워크 환경 정보를 획득한다.

[404] 일 실시예에 따라, 이전 프레임 정보는 부호화된 이전 프레임의 블록 아티팩트 검출 여부를 나타낼 수 있다.

[405] 일 실시예에 따라, 네트워크 환경 정보는 네트워크 채널의 비트레이트 및 대역폭 중 적어도 하나의 변화를 나타낼 수 있다.

[406] S2220 단계에서, AI 부호화 장치(1500)는 상기 원본 영상, 상기 이전에 부호화된 프레임 정보, 및 전송 환경 네트워크 정보에 기초하여 디블로킹 필터 설정 정보를 획득한다.

[407] 일 실시예에 따라, 네트워크 환경 정보가 네트워크 환경이 변화가 없음을 나타내면, 상기 디블로킹 필터 설정 정보는 획득되지 않고, 디블로킹 필터링은 수행되지 않을 수 있다. 구체적으로, 여러 명이 동시에 접속하거나, 실외에서 네트워크 채널의 변화가 있는 경우가 아닌 정상적인 상황에서는, 화질이 열화되는 블로킹 현상이 발생하지 않으므로, 디블로킹 필터링이 추가적으로 수행될 필요가 없기 때문이다.

[408] 일 실시예에 따라, 디블로킹 필터 설정 정보는 블록 아티팩트가 발생되는 영역을 나타내는 블로킹 영역 맵을 포함할 수 있다. 구체적으로, 도 13에서 전술한 바와 같이, A 스코어 맵 내지 G 스코어 맵들의 가중치 평균을 통해 블로킹 영역 맵이 결정될 수 있다.

[409] S2230 단계에서, AI 부호화 장치(1500)는 상기 디블로킹 필터 설정 정보에 기초하여 상기 원본 영상에 디블로킹 필터링을 적용하고 다운스케일용 DNN을 통해 AI 다운스케일된 제1 영상을 획득한다.

[410] 일 실시예에 따라, AI 부호화에서의 디블로킹 필터링은 디블로킹 필터 설정 정보에 기초하여 비트의 할당량을 재분배하는 필터링일 수 있다. 구체적으로, 디블로킹 필터는 영상의 복잡하고 움직임이 빠른 영역은 블러 처리하여 비트량을 절감하고, 절감한 비트량을 단순한 영역에 분배하여 단순한 영역의 블로킹 현상이 눈에 덜 띄도록 비트의 할당량을 재분배할 수 있다.

[411] S2240 단계에서AI 부호화 장치(1500)는 상기 제1 영상을 제1 부호화하여 영상 데이터를 생성한다.

[412] S2250 단계에서, AI 부호화 장치(1500)는 상기 디블로킹 필터 설정 정보, 상기 AI 다운스케일과 관련된 정보를 포함하는 AI 데이터, 및 상기 영상 데이터를 전송한다.

[413] 도 23은 일 실시예에 따른 AI 복호화 방법을 설명하기 위한 순서도이다.

[414] S2310 단계에서, AI 복호화 장치(1400)는 제 1 영상의 제 1 부호화 결과 생성된 영상 데이터, 원본 영상으로부터 상기 제 1 영상으로의 AI 다운스케일과 관련된 AI 데이터, 및 디블로킹 필터 설정 정보를 획득한다.

[415] 일 실시예에 따라, 디블로킹 필터 설정 정보가 획득되지 않으면, 디블로킹 필터링은 수행되지 않을 수 있다. 구체적으로, 디블로킹 필터 설정 정보가 획득되지 않았다는 것은 디블로킹 필터가 적용되어야 하는 경우, 즉, 여러 명이 동시에 접속하거나, 실외에서 네트워크 채널의 변화가 있는 경우가 아닌 정상적인 상황임을 의미하므로, 디블로킹 필터링이 추가적으로 수행될 필요가 없기 때문이다.

[416] 일 실시예에 따라, 디블로킹 필터 설정 정보는 블록 아티팩트가 발생되는 영역을 나타내는 블로킹 영역 맵을 포함할 수 있다. 구체적으로, 도 13에서 전술한 바와 같이, A 스코어 맵 내지 G 스코어 맵들의 가중치 평균을 통해 블로킹 영역 맵이 결정될 수 있다.

[417] S2320 단계에서, AI 복호화 장치(1400)는 상기 영상 데이터를 제 1 복호화하여 상기 제 1 영상에 대응하는 제 2 영상을 획득한다.

[418] S2330 단계에서, AI 복호화 장치(1400)는 복수의 DNN 설정 정보 중 상기 제 2 영상의 AI 업스케일을 위한 DNN 설정 정보를, 상기 AI 데이터에 기초하여 획득한다.

[419] S2340 단계에서, AI 복호화 장치(1400)는 상기 획득된 DNN 설정 정보로 동작하는 업스케일용 DNN(Deep Neural Network)을 통해 상기 제 2 영상으로부터 AI 업스케일되고 상기 하나의 디블로킹 필터에 기초하여 디블로킹 필터링이 적용된 복원 영상을 생성한다.

[420] 일 실시예에 따라, AI 복호화에서의 디블로킹 필터링은 디블로킹 필터 설정 정보에 기초하여 영상의 블록 경계를 블러 처리하고 블로킹 아티팩트를 제거하는 필터링일 수 있다. 구체적으로, 디블로킹 필터는 영상의 블록 사이의 경계는 블러 처리하고, 블로킹 아티팩트는 제거하여 각각의 영역에 맞게 디블로킹 필터링할 수 있다.

[421] 한편, 상술한 본 개시의 실시예들은 컴퓨터에서 실행될 수 있는 프로그램 또는 인스트럭션으로 작성가능하고, 작성된 프로그램 또는 인스트럭션은 매체에 저장될 수 있다.

[422] 매체는 컴퓨터로 실행 가능한 프로그램 또는 인스트럭션을 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, 애플리케이션을 유통하는 앱 스토어나 기타 다양한 소프트웨어를 공급 내지 유통하는 사이트, 서버 등에서 관리하는 기록매체 내지 저장매체도 들 수 있다.

[423] 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.

[424] 또한, DNN 모델은 하드웨어 칩 형태로 집적되어 전술한 AI 복호화 장치(1400) 또는 AI 부호화 장치(1500)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 애플리케이션 프로세서) 또는 그래픽 전용 프로세서(예를 들어, GPU)의 일부로 제작될 수도 있다.

[425] 한, 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 애플리케이션)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.

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

청구범위

[청구항 1] 메모리; 및
프로세서를 포함하고,
상기 메모리는 상기 프로세서에 의해 실행 가능한 명령어들을 저장하고
상기 프로세서는,
원본 영상, 이전에 부호화된 프레임 정보, 및 네트워크 환경 정보를 획득하고,
상기 원본 영상, 상기 이전에 부호화된 프레임 정보, 및 상기 네트워크 환경 정보에 기초하여 디블로킹 필터 설정 정보를 획득하고,
상기 디블로킹 필터 설정 정보에 기초하여 상기 원본 영상에 디블로킹 필터링을 적용하고 다운스케일용 DNN을 통해 AI 다운스케일된 제1 영상을 획득하고,
상기 제1 영상을 제1 부호화하여 영상 데이터를 생성하고,
상기 디블로킹 필터 설정 정보, 상기 AI 다운스케일과 관련된 정보를 포함하는 AI 데이터, 및 상기 영상 데이터를 전송하는 것을 특징으로하는, AI 부호화 장치.

[청구항 2] 제1항에 있어서,
상기 디블로킹 필터 설정 정보는 블록 아티팩트가 발생되는 영역을 나타내는 블로킹 영역 맵을 포함하는, AI 부호화 장치.

[청구항 3] 제2항에 있어서,
상기 블로킹 영역 맵은 상기 원본 영상의 복잡도 및 움직임 변화도 중 적어도 하나에 기초하여 결정되는, AI 부호화 장치.

[청구항 4] 제1항에 있어서,
상기 디블로킹 필터링은 상기 디블로킹 필터 설정 정보에 기초하여 비트의 할당량을 재분배하는 필터링인, AI 부호화 장치.

[청구항 5] 제1항에 있어서,
상기 다운스케일용 DNN의 DNN 설정 정보는, 상기 다운스케일용 DNN과 상기 제 2 영상의 AI 업스케일을 위한 업스케일용 DNN의 연계 훈련을 통해 획득되는, AI 부호화 장치.

[청구항 6] 제1항에 있어서,
상기 네트워크 환경 정보가 네트워크 환경이 변화가 없음을 나타내면,
상기 디블로킹 필터 설정 정보는 획득되지 않고, 상기 디블로킹 필터링은 수행되지 않는, AI 부호화 장치.

[청구항 7] 제1항에 있어서,
상기 네트워크 환경 정보는 네트워크 채널의 비트레이트 및 대역폭 중 적어도 하나의 변화를 나타내는, AI 부호화 장치.

[청구항 8] 제1항에 있어서,

상기 이전 프레임 정보는 부호화된 이전 프레임의 블록 아티팩트 검출 여부를 나타내는, AI 부호화 장치.

[청구항 9] 제1항에 있어서,
상기 원본 영상, 상기 이전에 부호화된 프레임 정보, 및 상기 네트워크 환경 정보에 기초하여 양자화 파라미터 차분값 맵 및 레이트 컨트롤에 대한 정보를 추가로 획득하고,
상기 양자화 파라미터 차분값 맵 및 레이트 컨트롤에 대한 정보에 기초하여 상기 제 1 부호화를 수행하는, AI 부호화 장치.

[청구항 10] 메모리; 및
프로세서를 포함하고,
상기 메모리는 상기 프로세서에 의해 실행 가능한 명령어들을 저장하고
상기 프로세서는,
제 1 영상의 제 1 부호화 결과 생성된 영상 데이터, 원본 영상으로부터 상기 제 1 영상으로의 AI 다운스케일과 관련된 AI 데이터, 및 디블로킹 필터 설정 정보를 획득하고,
상기 영상 데이터를 제 1 복호화하여 상기 제 1 영상에 대응하는 제 2 영상을 획득하고,
복수의 DNN 설정 정보 중 상기 제 2 영상의 AI 업스케일을 위한 DNN 설정 정보를, 상기 AI 데이터에 기초하여 획득하고,
상기 획득된 DNN 설정 정보로 동작하는 업스케일용 DNN(Deep Neural Network)을 통해 상기 제 2 영상으로부터 AI 업스케일되고 상기 디블로킹 필터 설정 정보에 기초하여 디블로킹 필터링이 적용된 복원 영상을 생성하는 것을 특징으로 하는, AI 복호화 장치.

[청구항 11] 제10항에 있어서,
상기 디블로킹 필터 설정 정보는 블록 아티팩트가 발생되는 영역을 나타내는 블로킹 영역 맵을 포함하는, AI 복호화 장치.

[청구항 12] 제10항에 있어서,
상기 블로킹 영역 맵은 원본 영상의 복잡도 및 움직임 변화도 중 적어도 하나에 기초하여 결정되는, AI 복호화 장치.

[청구항 13] 제10항에 있어서,
상기 디블로킹 필터링은 상기 디블로킹 필터 설정 정보에 기초하여 영상의 블록 경계를 블러 처리하고 블로킹 아티팩트를 제거하는 필터링인, AI 복호화 장치.

[청구항 14] 제10항에 있어서,
상기 복수의 DNN 설정 정보는, 상기 업스케일용 DNN과 상기 원본 영상의 AI 다운스케일에 이용되는 다운스케일용 DNN의 연계 훈련을 통해 획득되는, AI 복호화 장치.

[청구항 15] AI 부호화 장치에 의한 AI 부호화 방법에 있어서,

원본 영상, 이전에 부호화된 프레임 정보, 및 네트워크 환경 정보를 획득하는 단계;

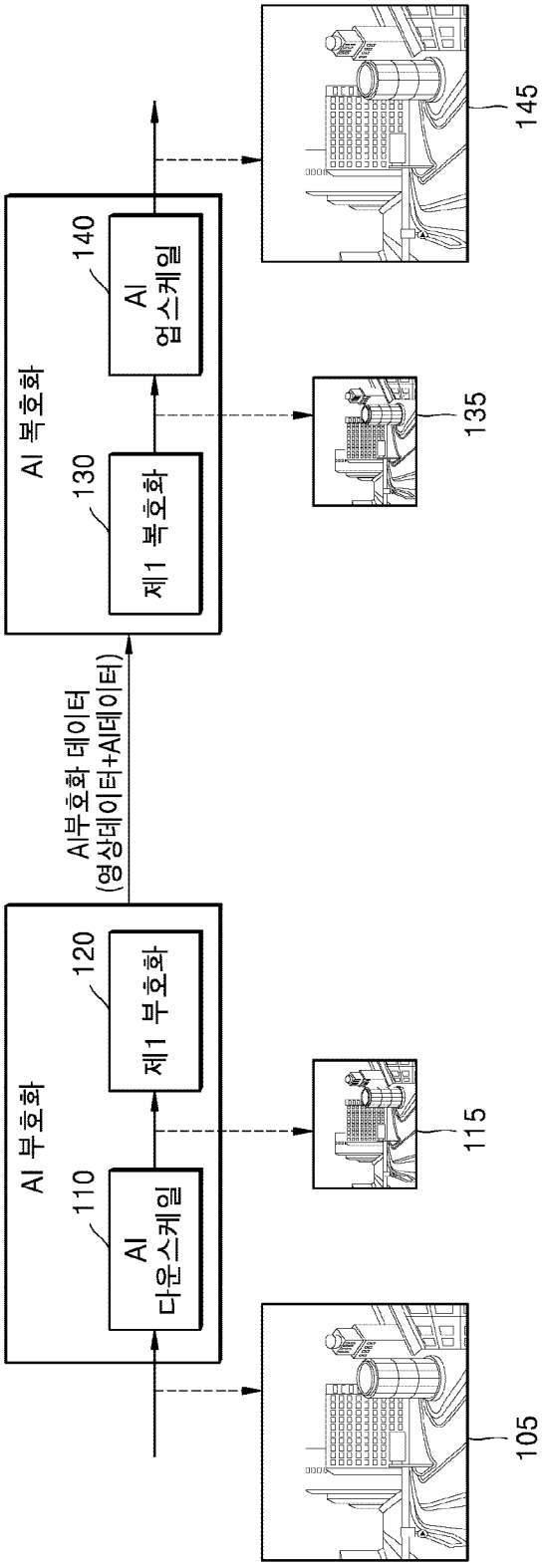
상기 원본 영상, 상기 이전에 부호화된 프레임 정보, 및 상기 네트워크 환경 정보에 기초하여 디블로킹 필터 설정 정보를 획득하는 단계;

상기 디블로킹 필터 설정 정보에 기초하여 상기 원본 영상에 디블로킹 필터링을 적용하고 다운스케일용 DNN을 통해 AI 다운스케일된 제1 영상을 획득하는 단계;

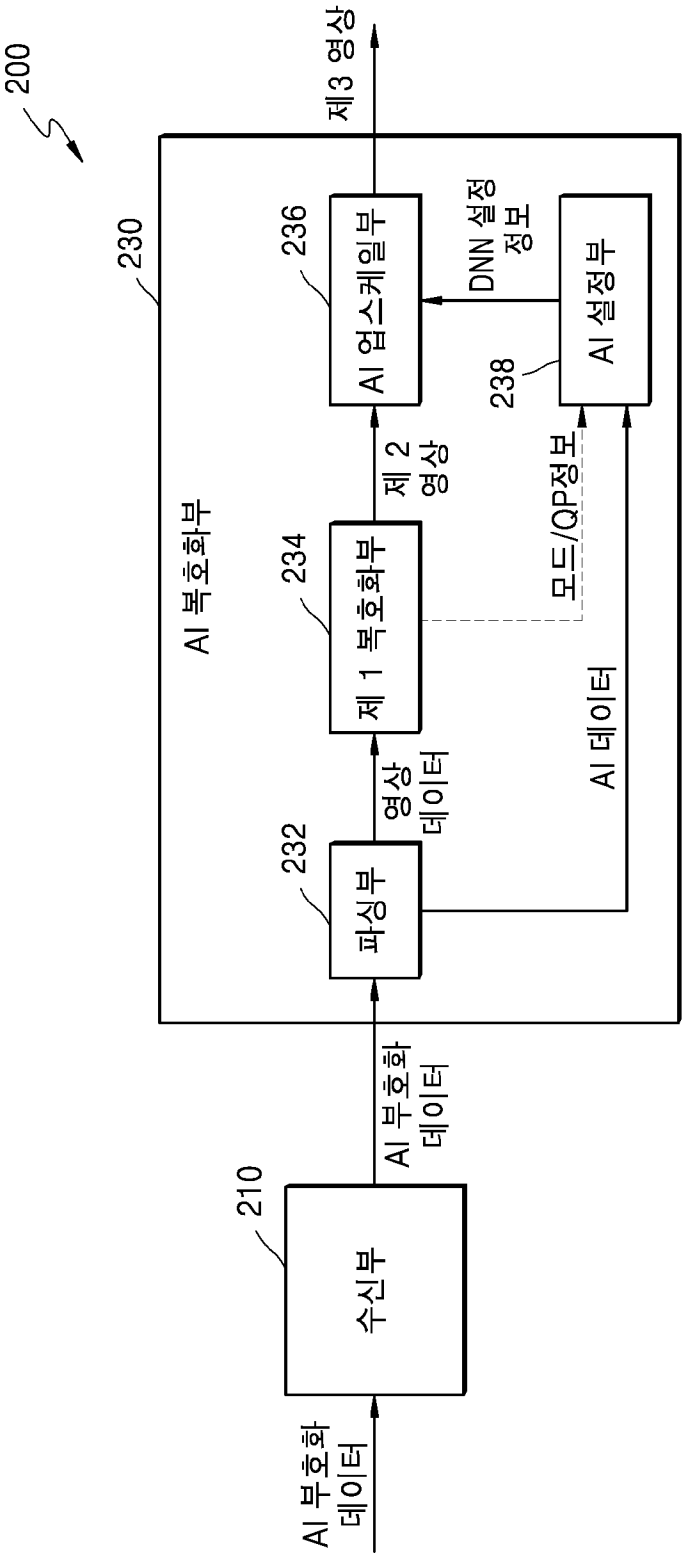
상기 제1 영상을 제1 부호화하여 영상 데이터를 생성하는 단계;

상기 디블로킹 필터 설정 정보, 상기 AI 다운스케일과 관련된 정보를 포함하는 AI 데이터, 및 상기 영상 데이터를 전송하는 단계를 포함하는 것을 특징으로하는 AI 부호화 방법.

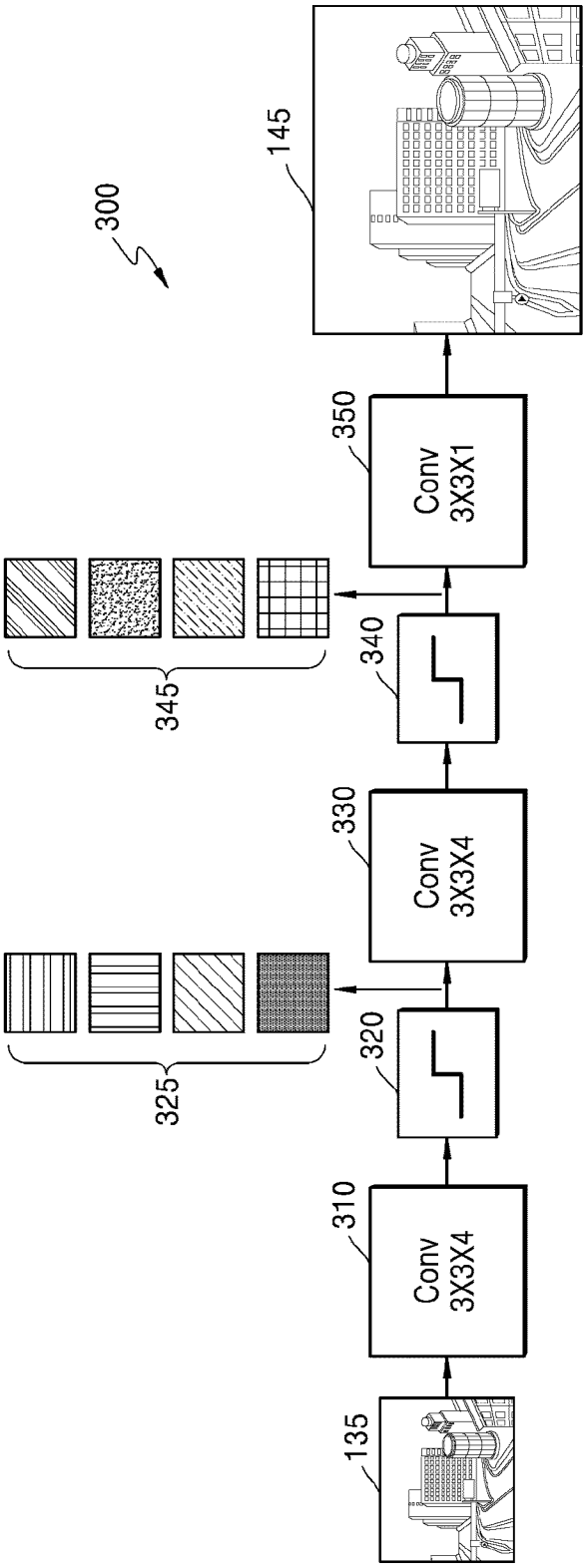
[도1]


AI 부호화
110
AI
다운스케일
120
제1 부호화
AI부호화 데이터
(영상데이터+AI데이터)
AI 복호화
130
제1 복호화
140
AI
업스케일
105
115
135
145

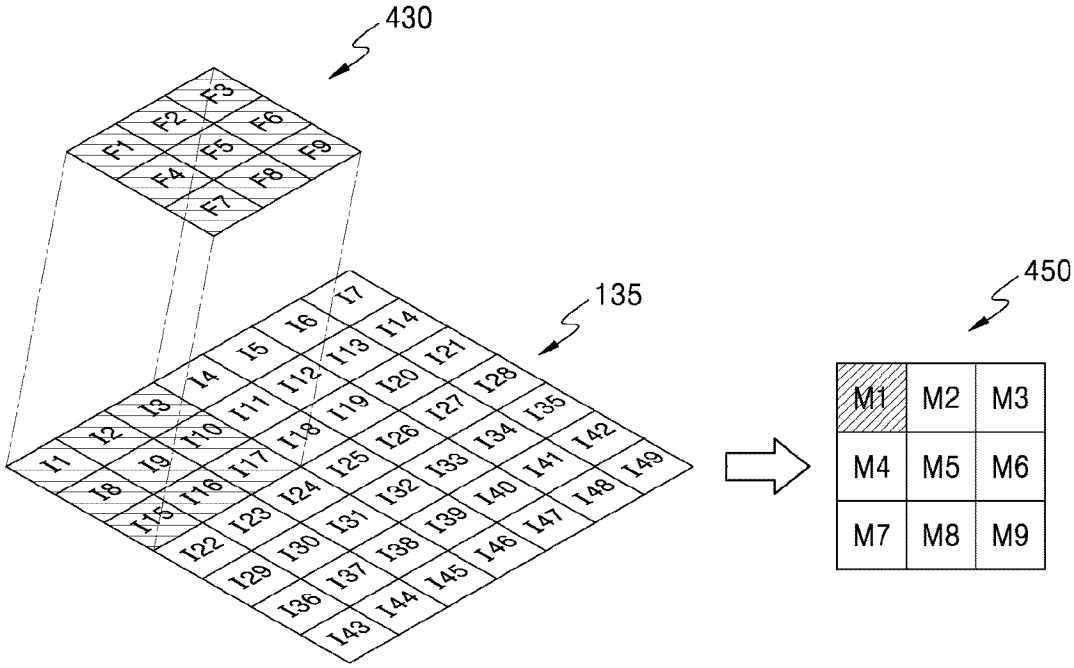
[도2]


200
제3 영상
230
236
AI 업스케일부
DNN 설정 정보
238
AI 설정부
AI 복호화부
234
제 1 복호화부
제 2 영상
모드/QP정보
232
파싱부
영상 데이터
AI 데이터
210
수신부
AI 부호화 데이터
AI 부호화 데이터

[도3]


300
145
350
Conv 3X3X1
340
345
330
Conv 3X3X4
320
325
310
Conv 3X3X4
135

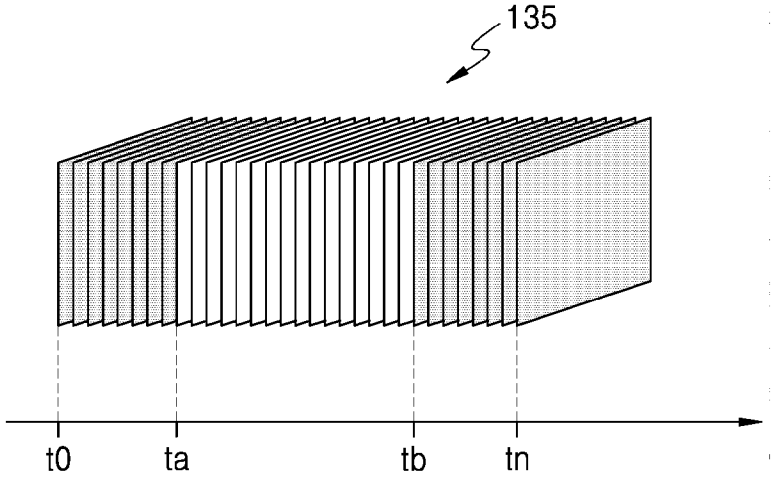
[도4]



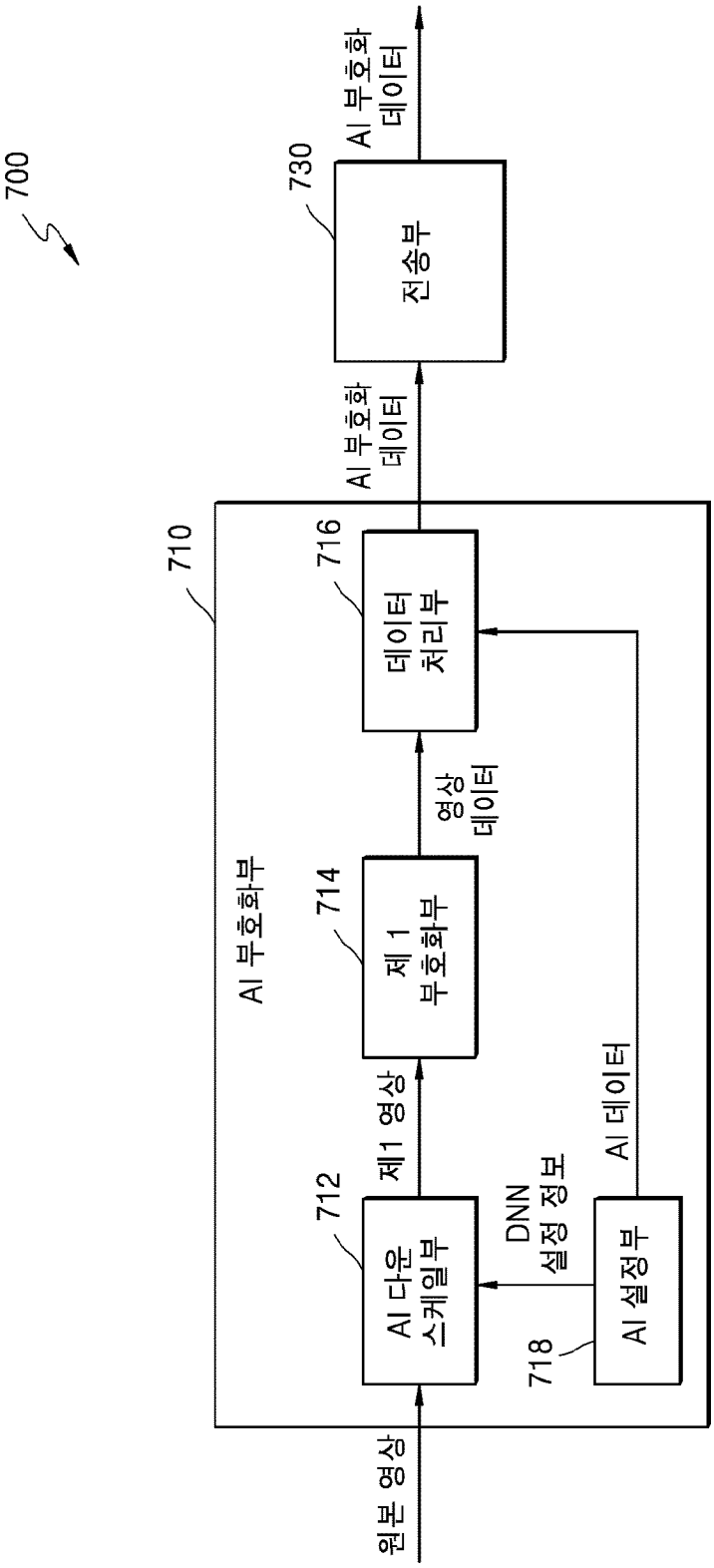
[도5]

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

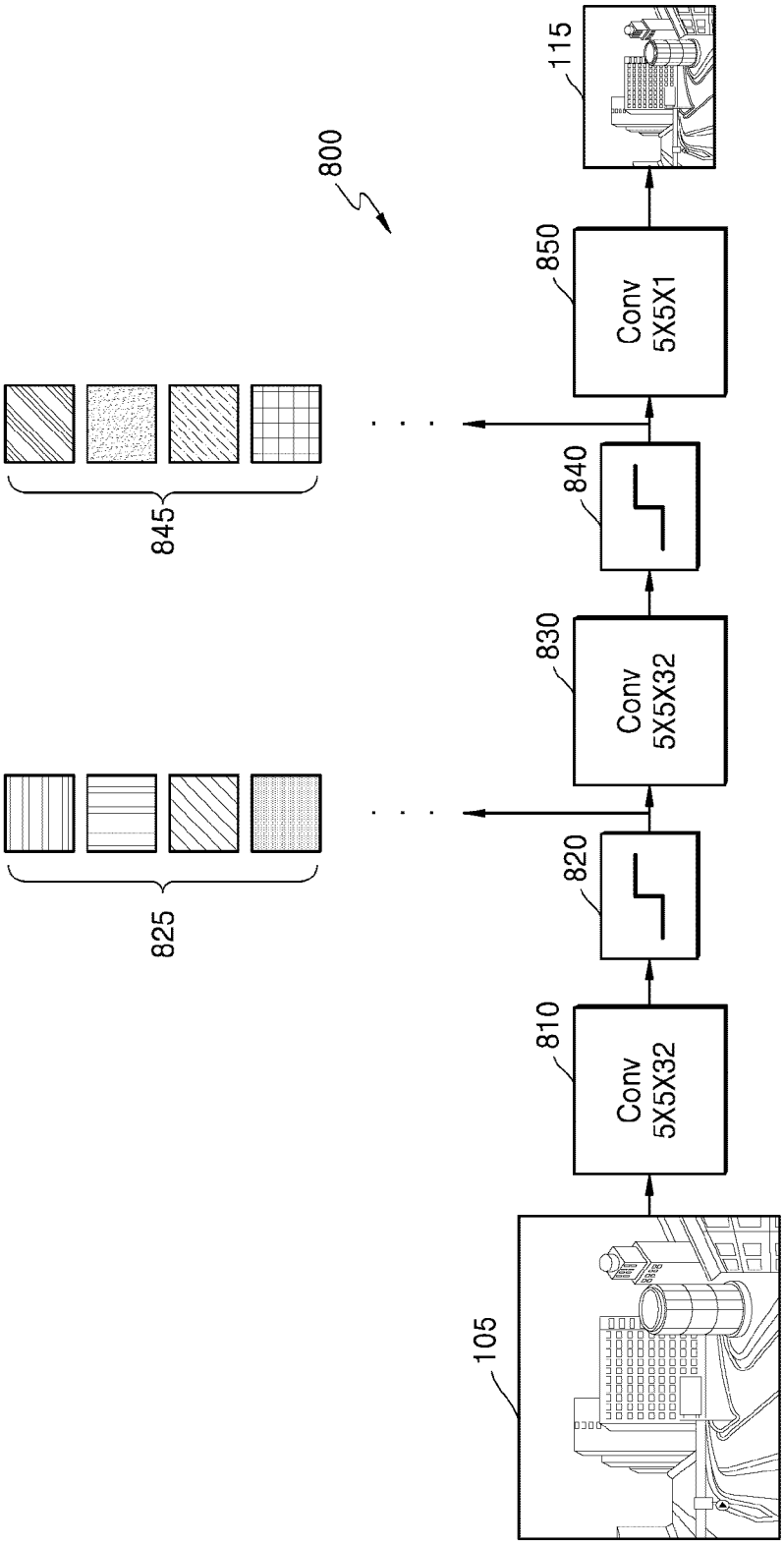
[도6]



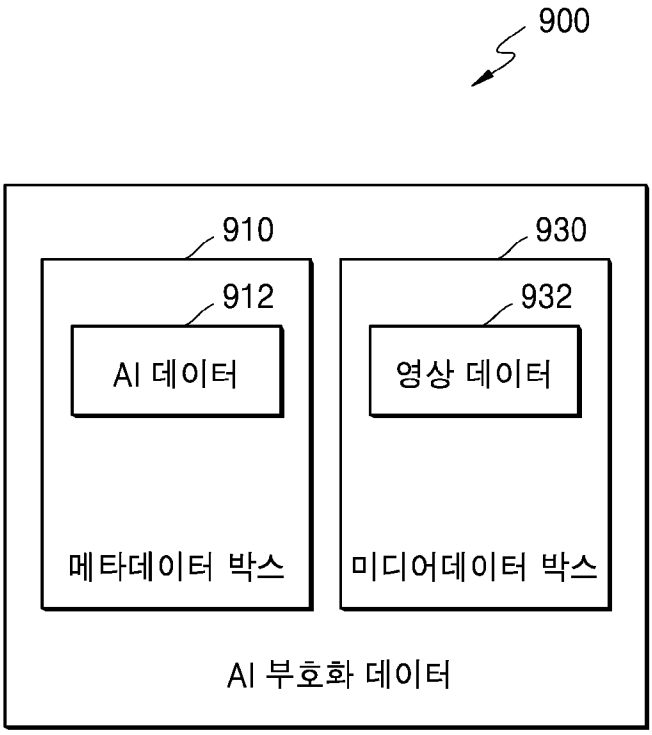
[도7]


700
730
전송부
AI 부호화 데이터
AI 부호화 데이터
710
716
데이터 처리부
AI 부호화부
714
제 1 부호화부
영상 데이터
712
AI 다운 스케일부
제1 영상
DNN 설정 정보
AI 데이터
718
AI 설정부
원본 영상

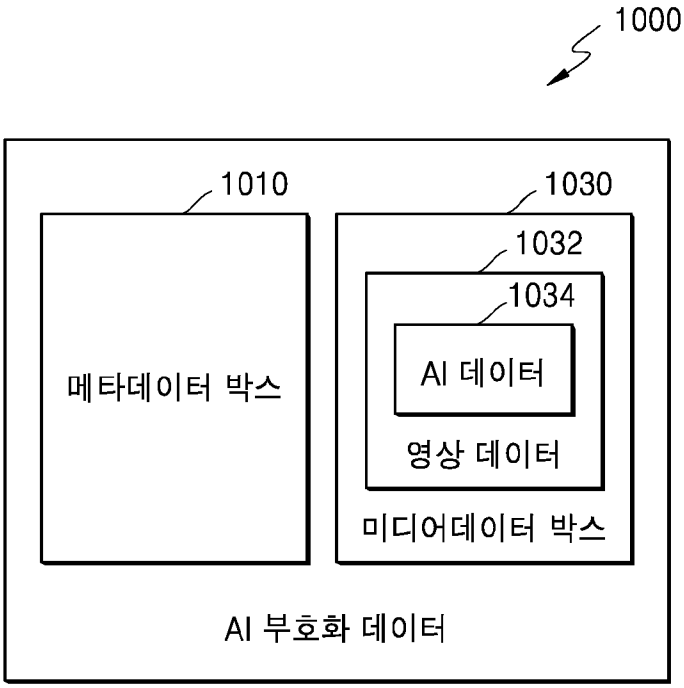
[도8]


105
Conv 5X5X32
810
820
Conv 5X5X32
830
840
Conv 5X5X1
850
115
800
825
845

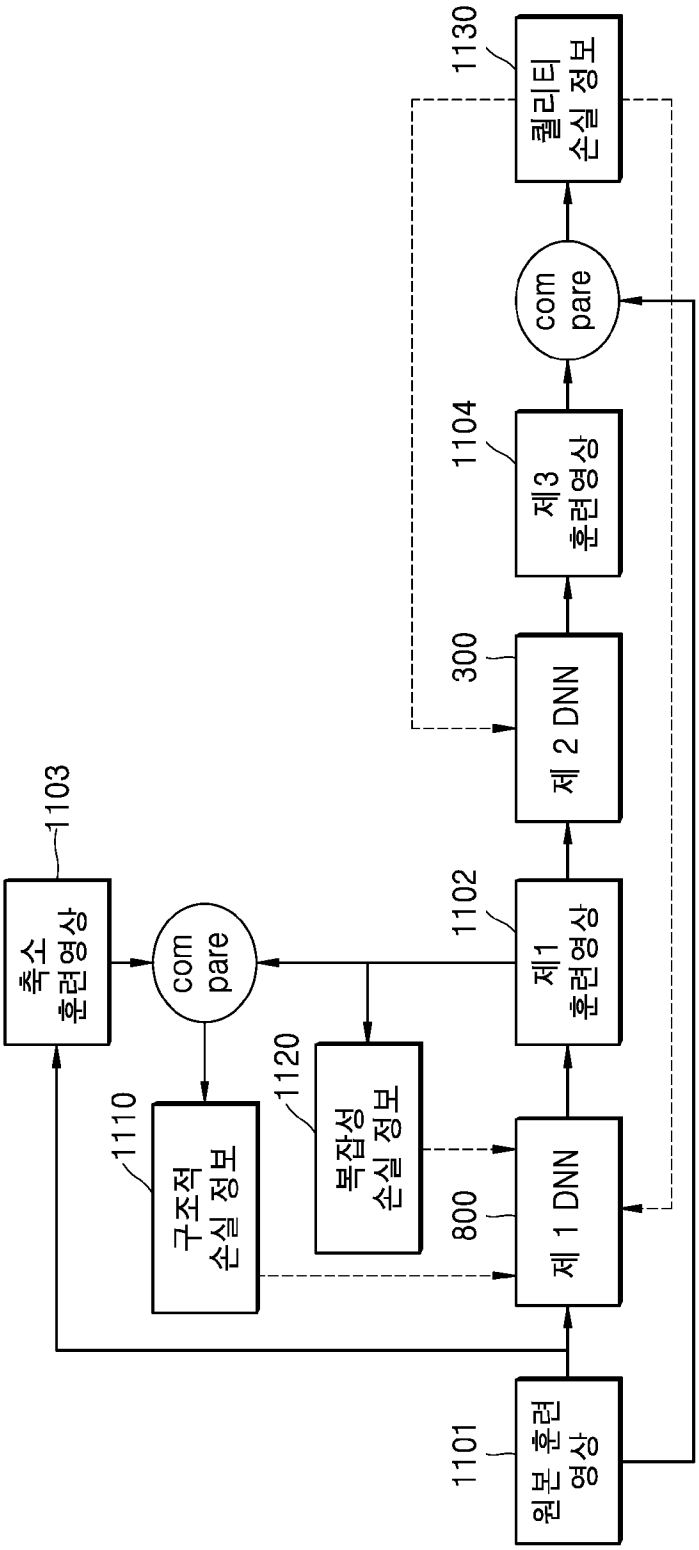
[도9]


900
910
912
AI 데이터
메타데이터 박스
930
932
영상 데이터
미디어데이터 박스
AI 부호화 데이터

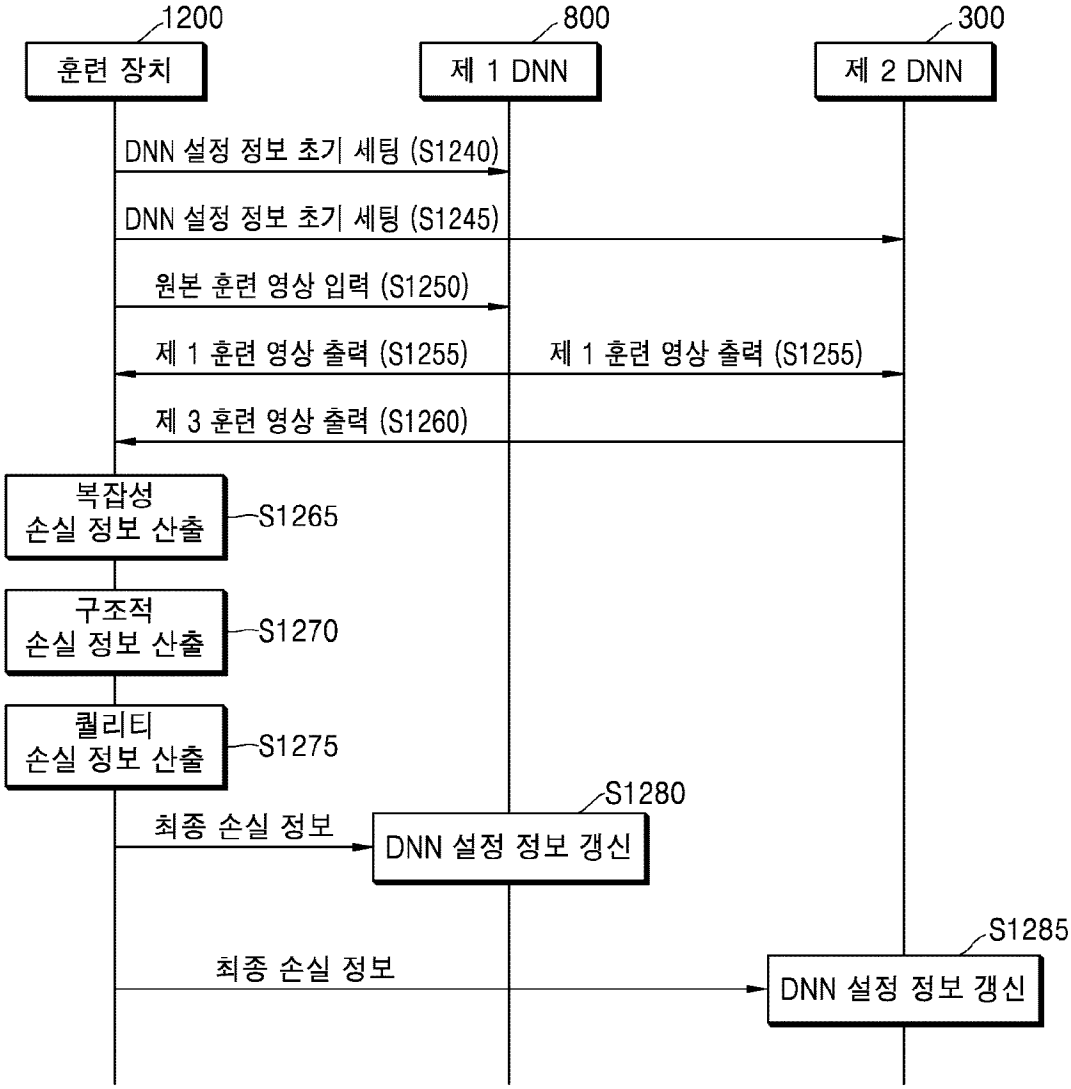
[도10]


1000
1010
메타데이터 박스
1030
1032
1034
AI 데이터
영상 데이터
미디어데이터 박스
AI 부호화 데이터

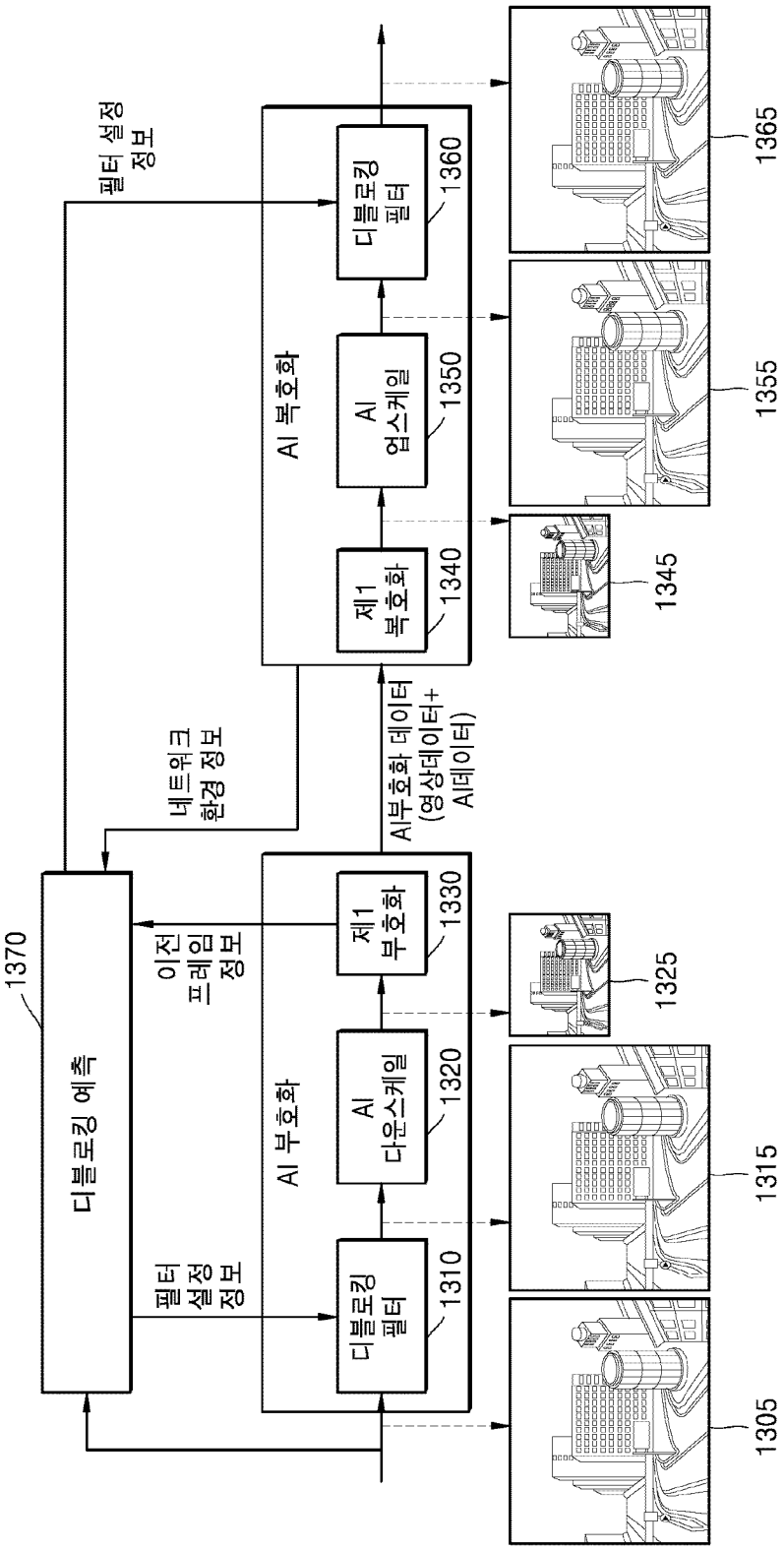
[도11]


1130
퀄리티 손실 정보
com pare
1104
제3 훈련영상
300
제 2 DNN
1102
제1 훈련영상
1103
축소 훈련영상
com pare
1120
복잡성 손실 정보
1110
구조적 손실 정보
800
제 1 DNN
1101
원본 훈련 영상

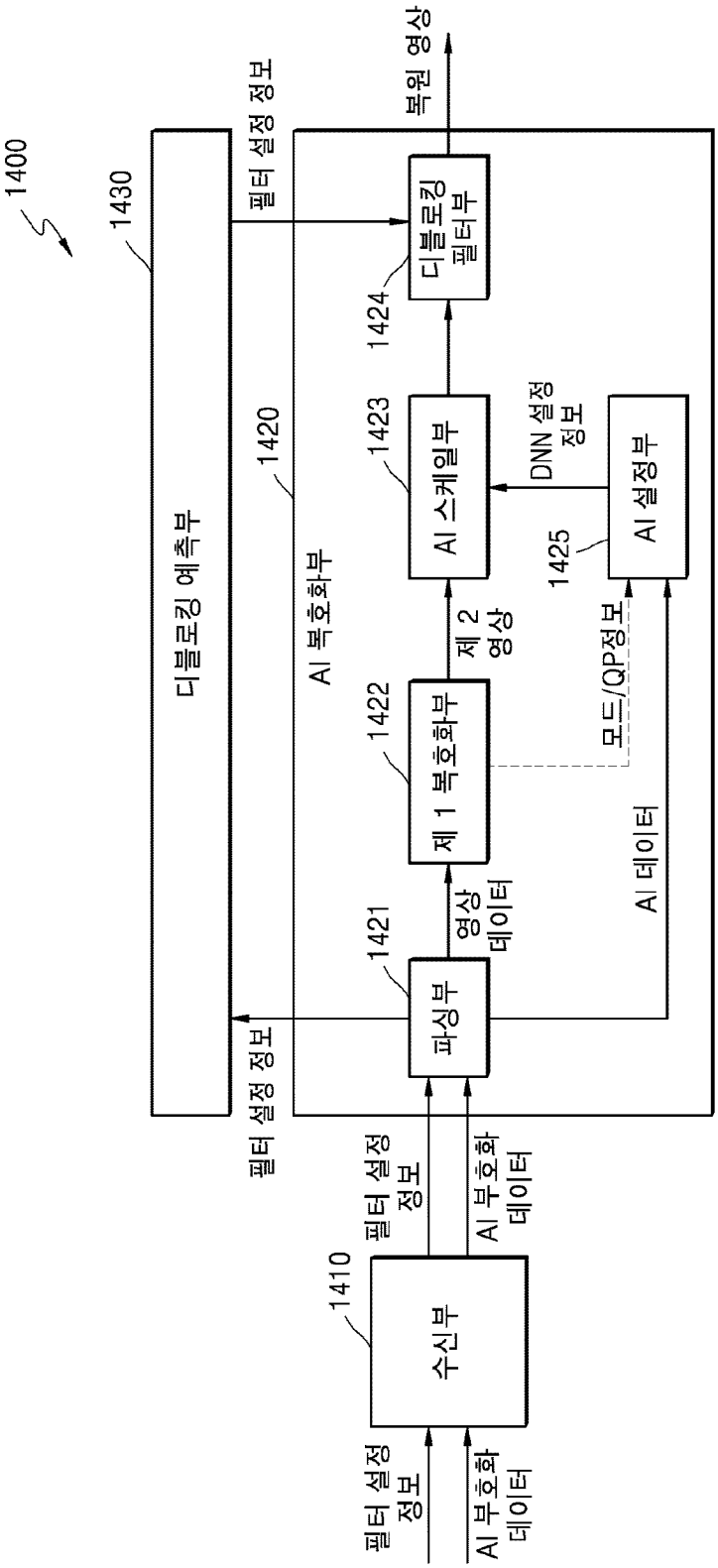
[도12]


1200
훈련 장치
800
제 1 DNN
300
제 2 DNN
DNN 설정 정보 초기 세팅 (S1240)
DNN 설정 정보 초기 세팅 (S1245)
원본 훈련 영상 입력 (S1250)
제 1 훈련 영상 출력 (S1255)
제 1 훈련 영상 출력 (S1255)
제 3 훈련 영상 출력 (S1260)
복잡성 손실 정보 산출
S1265
구조적 손실 정보 산출
S1270
퀄리티 손실 정보 산출
S1275
S1280
최종 손실 정보
DNN 설정 정보 갱신
S1285
최종 손실 정보
DNN 설정 정보 갱신

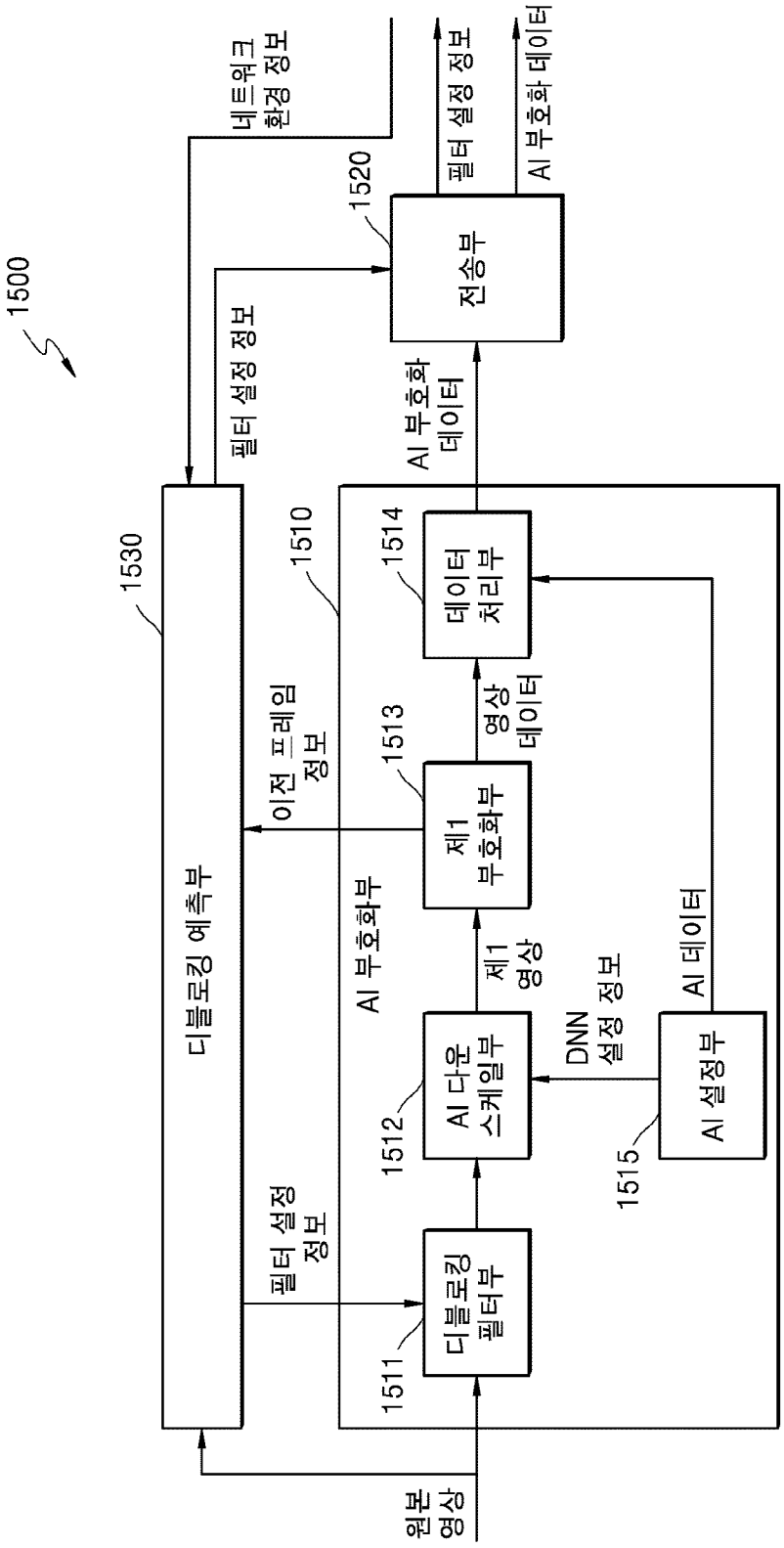
[도13]


1370
디블로킹 예측
필터 설정 정보
이전 프레임 정보
네트워크 환경 정보
필터 설정 정보
AI 부호화
디블로킹 필터
1310
AI 다운스케일
1320
제1 부호화
1330
AI부호화 데이터 (영상데이터+ AI데이터)
AI 복호화
제1 복호화
1340
AI 업스케일
1350
디블로킹 필터
1360
1305
1315
1325
1345
1355
1365

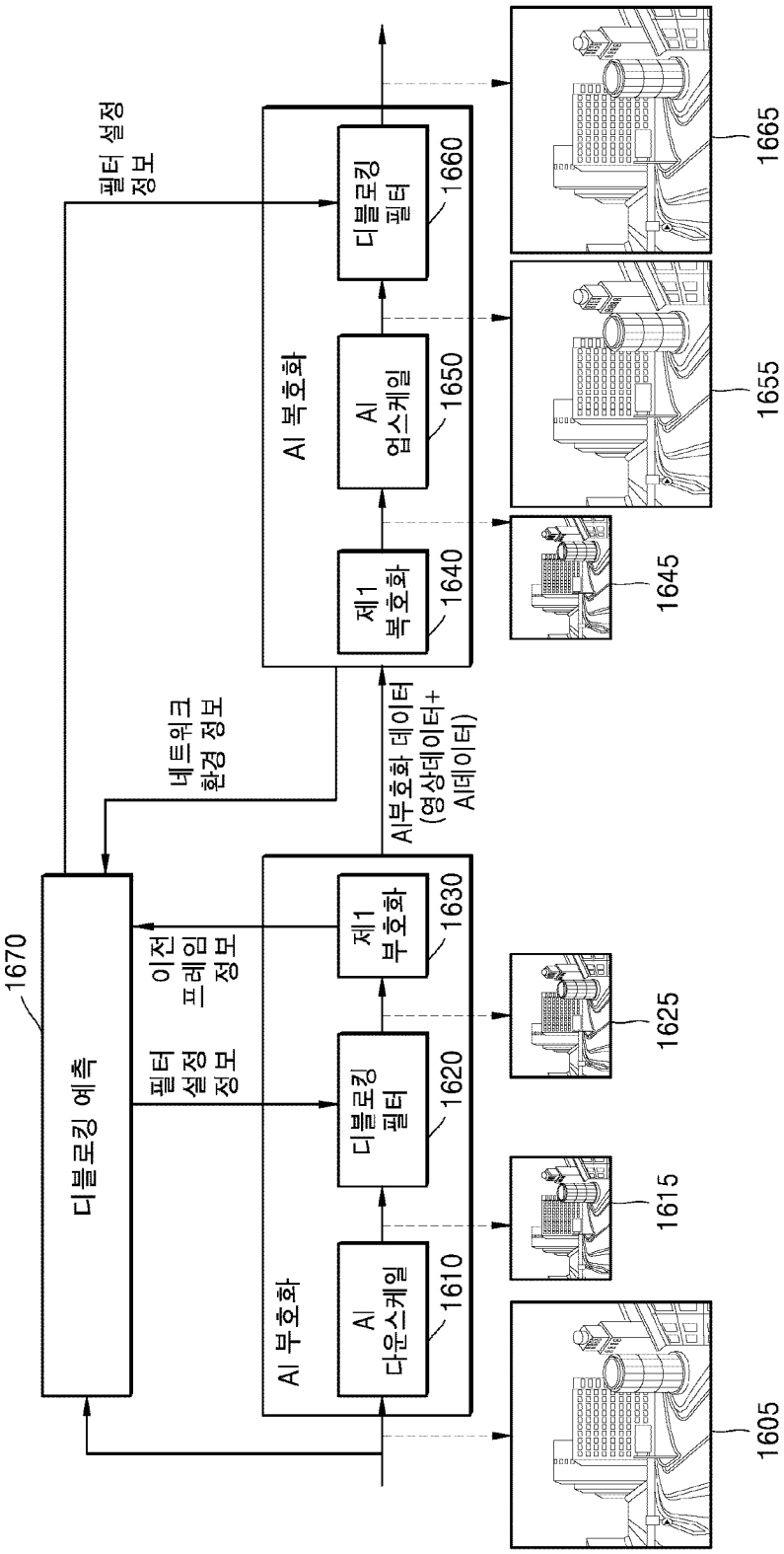
[도14]


1400
1430
디블로킹 예측부
1420
AI 복호화부
1421
파싱부
1422
제 1 복호화부
1423
AI 스케일부
1424
디블로킹 필터부
1425
AI 설정부
수신부
1410
필터 설정 정보
AI 부호화 데이터
영상 데이터
제 2 영상
DNN 설정 정보
모드/QP정보
AI 데이터
복원 영상

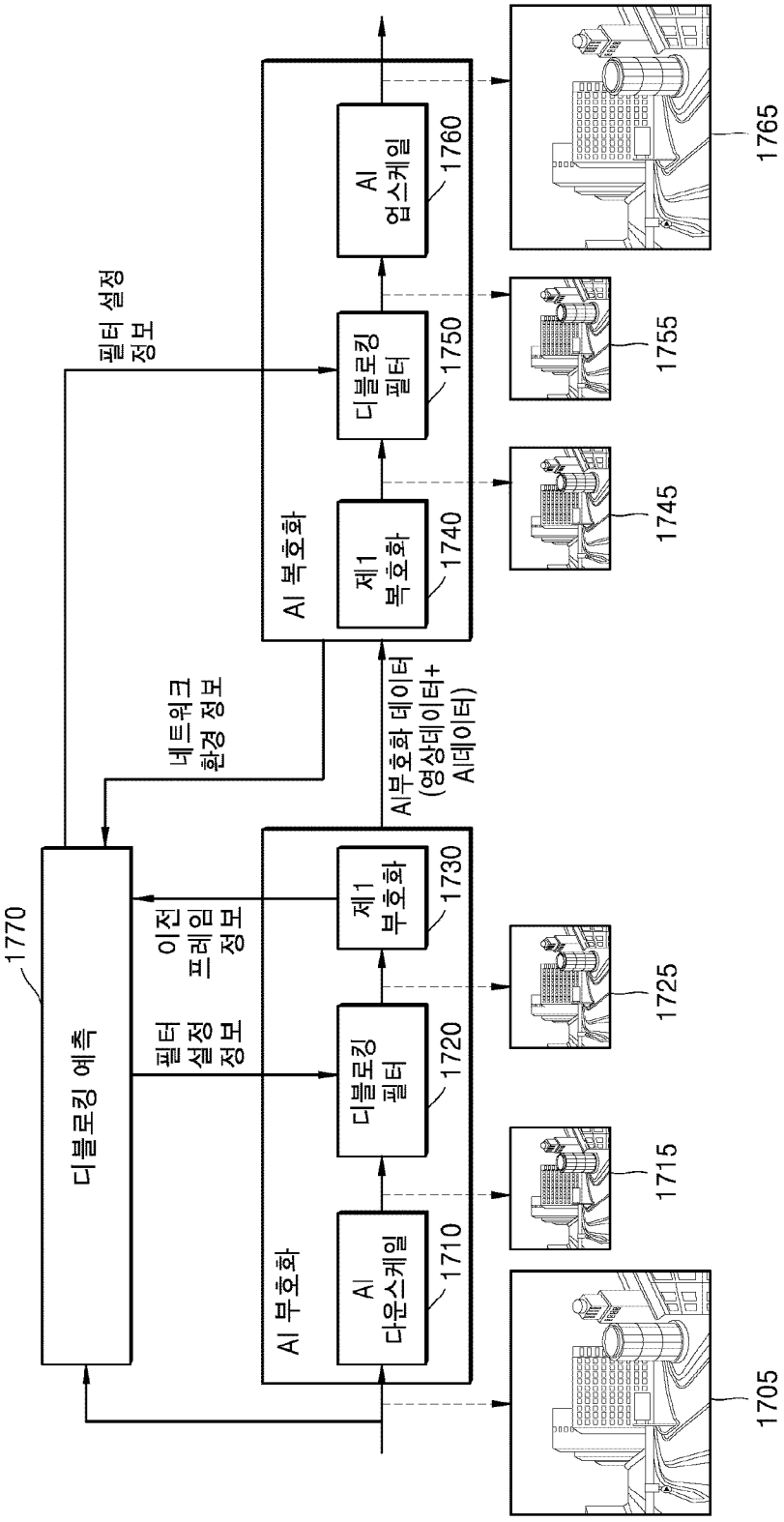
[도15]


1500
네트워크 환경 정보
필터 설정 정보
AI 부호화 데이터
1520
전송부
1530
디블로킹 예측부
1510
AI 부호화부
1514
데이터 처리부
1513
제1 부호화부
이전 프레임 정보
영상 데이터
1512
AI 다운 스케일부
제1 영상
DNN 설정 정보
1515
AI 설정부
AI 데이터
1511
디블로킹 필터부
필터 설정 정보
원본 영상

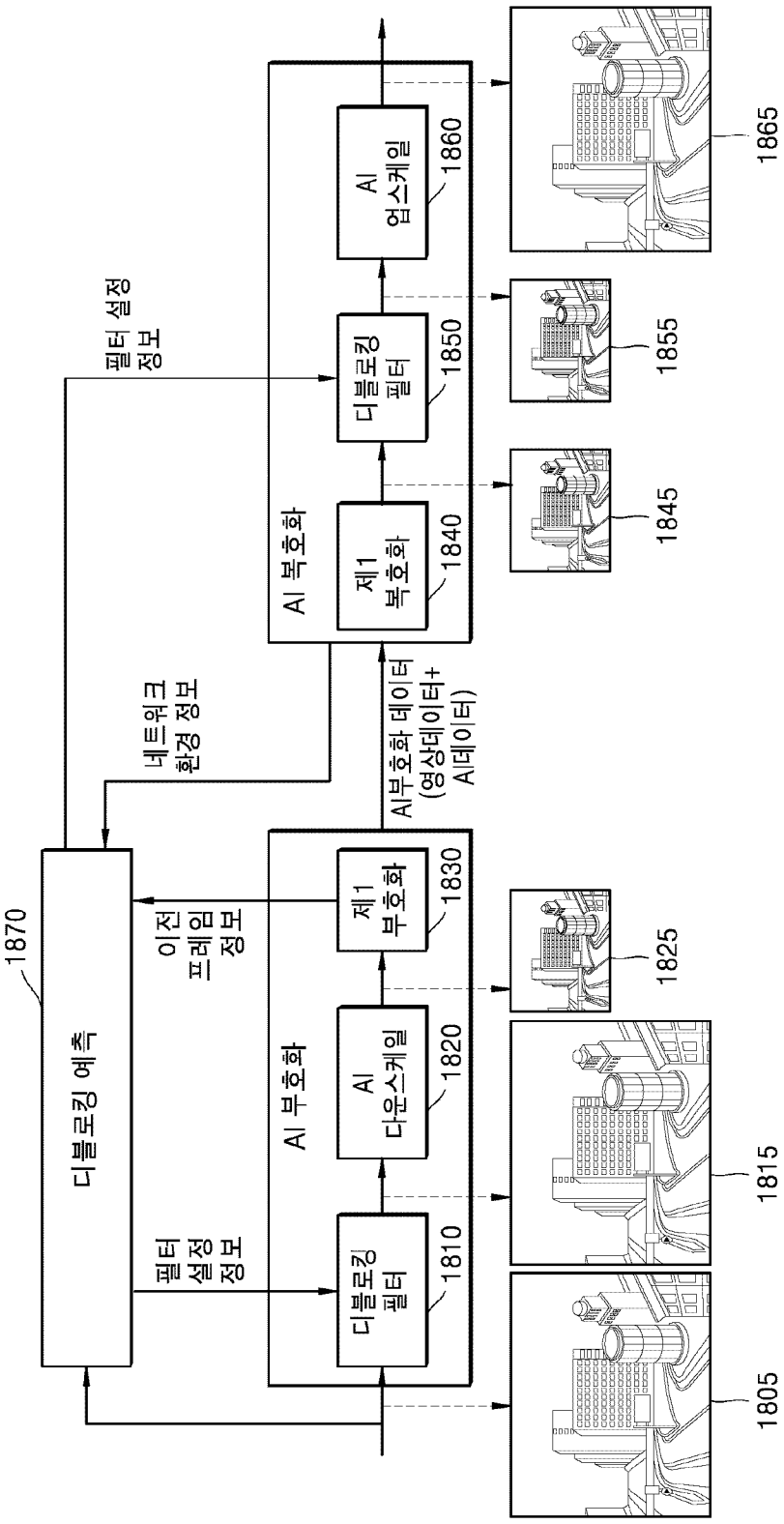
[도16]


1670
디블로킹 예측
필터 설정 정보
이전 프레임 정보
네트워크 환경 정보
AI 부호화
AI 다운스케일
1610
디블로킹 필터
1620
제1 부호화
1630
AI부호화 데이터 (영상데이터+ AI데이터)
AI 복호화
제1 복호화
1640
AI 업스케일
1650
디블로킹 필터
1660
필터 설정 정보
1605
1615
1625
1645
1655
1665

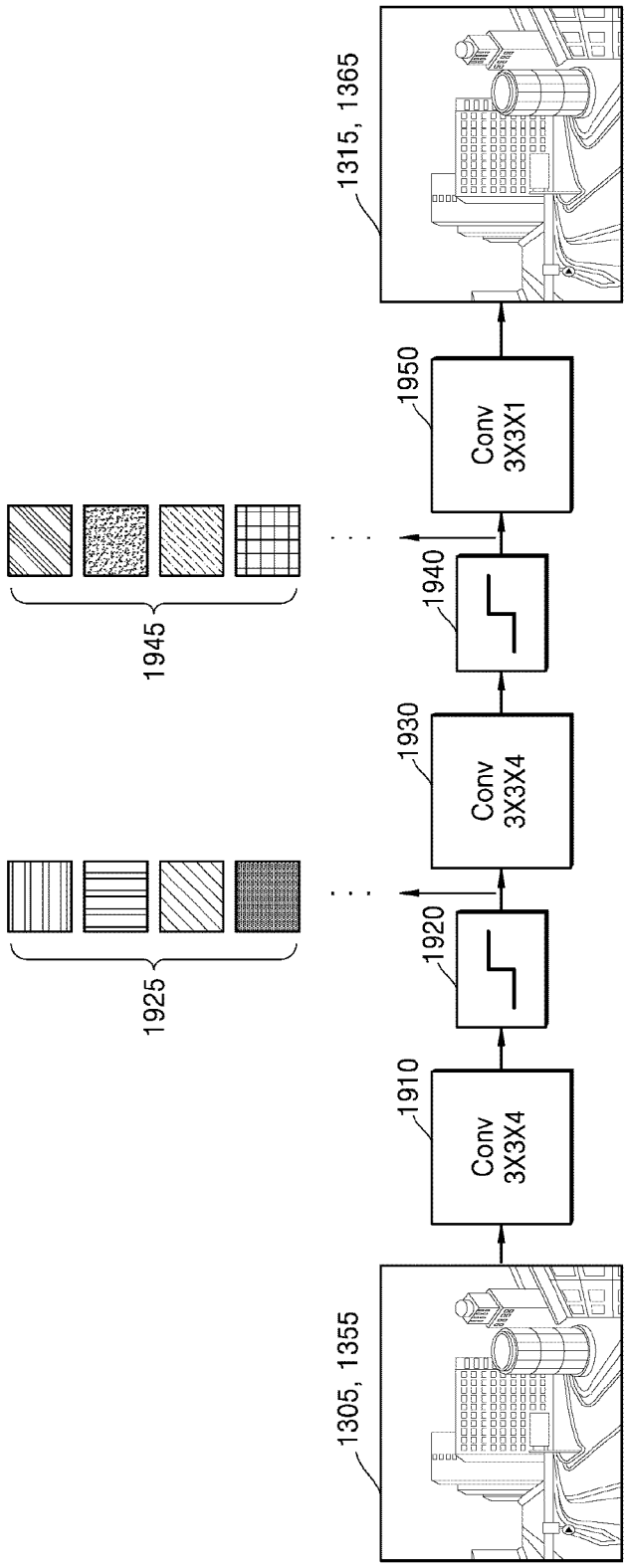
[도17]


1770
디블로킹 예측
이전 프레임 정보
필터 설정 정보
AI 부호화
AI 다운스케일
1710
디블로킹 필터
1720
제1 부호화
1730
AI부호화 데이터 (영상데이터+ AI데이터)
네트워크 환경 정보
AI 복호화
제1 복호화
1740
디블로킹 필터
1750
AI 업스케일
1760
필터 설정 정보
1705
1715
1725
1745
1755
1765

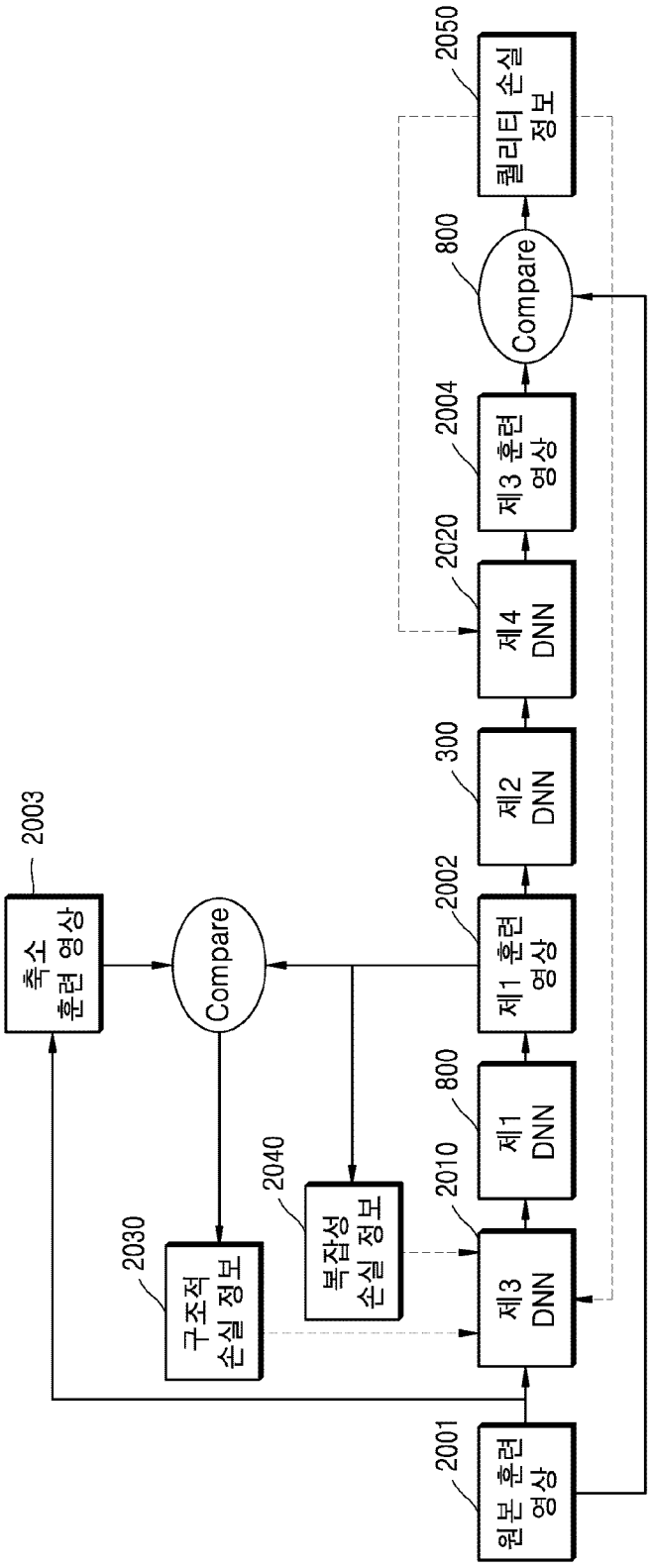
[도18]


1870
디블로킹 예측
필터 설정 정보
이전 프레임 정보
네트워크 환경 정보
AI 부호화
디블로킹 필터
1810
AI 다운스케일
1820
제1 부호화
1830
AI부호화 데이터 (영상데이터+ AI데이터)
AI 복호화
제1 복호화
1840
디블로킹 필터
1850
AI 업스케일
1860
필터 설정 정보
1805
1815
1825
1845
1855
1865

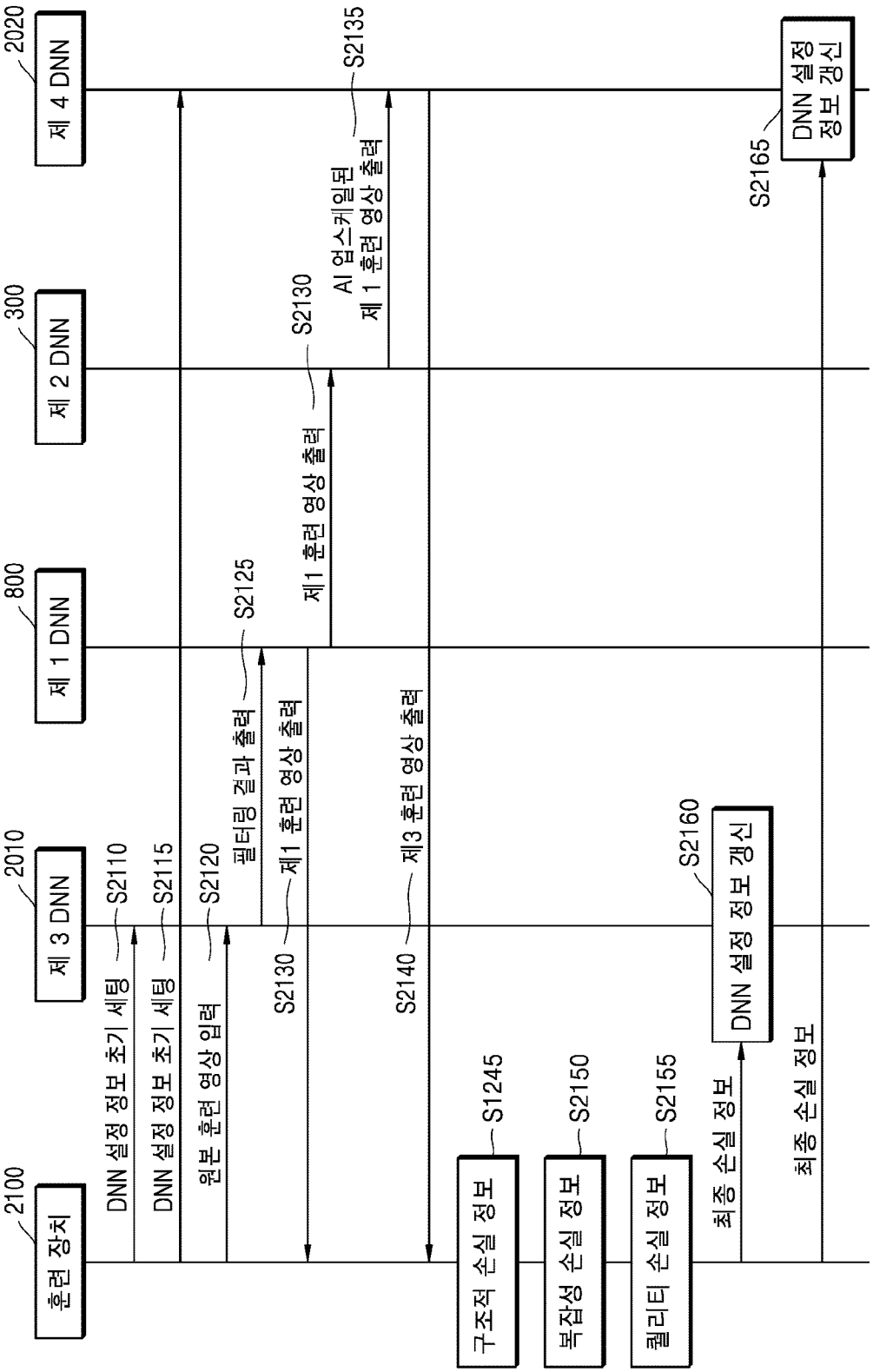
[도19]


1315, 1365
1950
Conv 3X3X1
1940
1930
Conv 3X3X4
1920
1910
Conv 3X3X4
1945
1925
1305, 1355

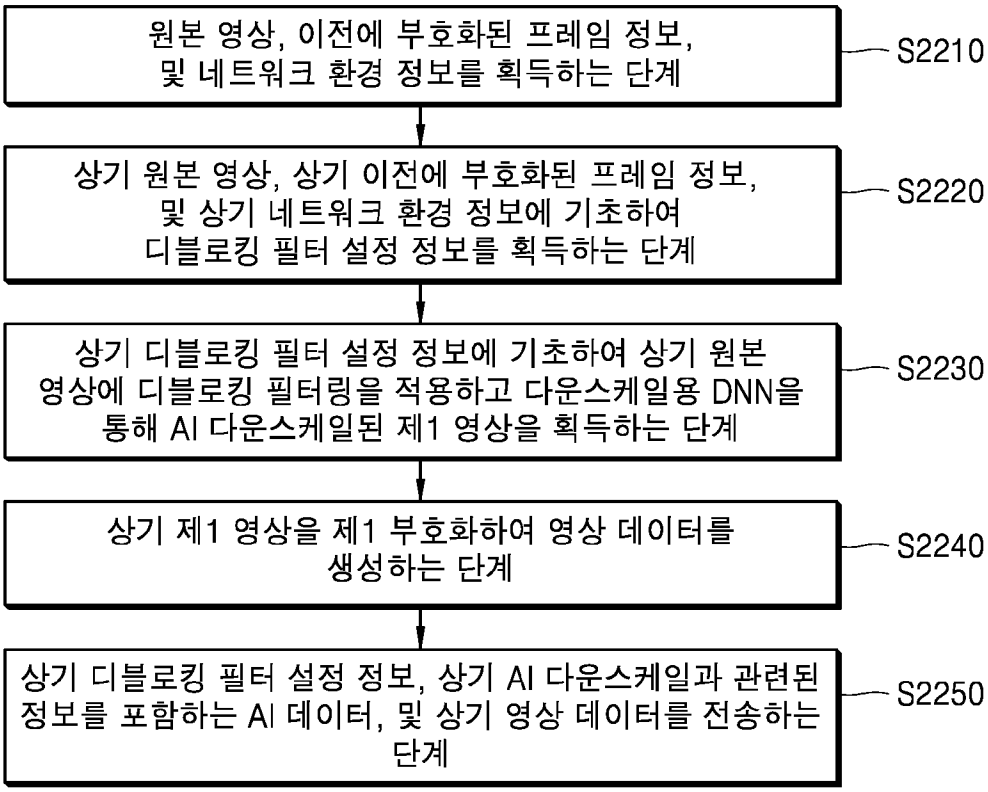
[도20]


2001
원본 훈련 영상
2003
축소 훈련 영상
Compare
2030
구조적 손실 정보
2040
복잡성 손실 정보
2010
제3 DNN
800
제1 DNN
2002
제1 훈련 영상
300
제2 DNN
2020
제4 DNN
2004
제3 훈련 영상
800
Compare
2050
퀄리티 손실 정보

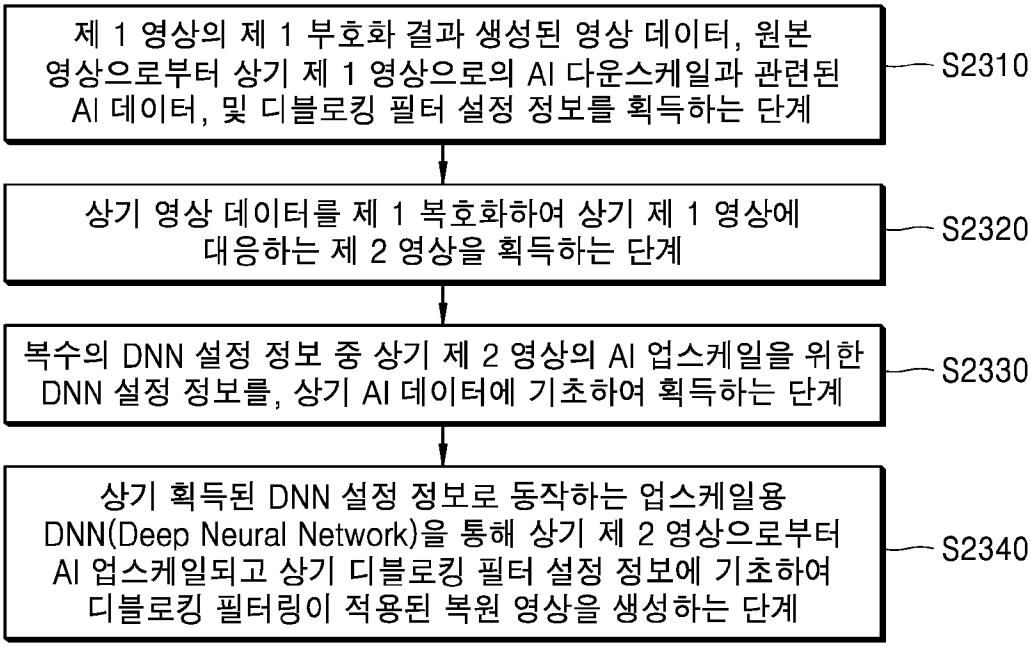
[도21]


2100
훈련 장치
2010
제 3 DNN
800
제 1 DNN
300
제 2 DNN
2020
제 4 DNN
DNN 설정 정보 초기 세팅
S2110
DNN 설정 정보 초기 세팅
S2115
원본 훈련 영상 입력
S2120
필터링 결과 출력
S2125
제1 훈련 영상 출력
S2130
제1 훈련 영상 출력
S2130
AI 업스케일된 제 1 훈련 영상 출력
S2135
제3 훈련 영상 출력
S2140
구조적 손실 정보
S1245
복잡성 손실 정보
S2150
퀄리티 손실 정보
S2155
최종 손실 정보
DNN 설정 정보 갱신
S2160
최종 손실 정보
DNN 설정 정보 갱신
S2165

[도22]


원본 영상, 이전에 부호화된 프레임 정보, 및 네트워크 환경 정보를 획득하는 단계
S2210
상기 원본 영상, 상기 이전에 부호화된 프레임 정보, 및 상기 네트워크 환경 정보에 기초하여 디블로킹 필터 설정 정보를 획득하는 단계
S2220
상기 디블로킹 필터 설정 정보에 기초하여 상기 원본 영상에 디블로킹 필터링을 적용하고 다운스케일용 DNN을 통해 AI 다운스케일된 제1 영상을 획득하는 단계
S2230
상기 제1 영상을 제1 부호화하여 영상 데이터를 생성하는 단계
S2240
상기 디블로킹 필터 설정 정보, 상기 AI 다운스케일과 관련된 정보를 포함하는 AI 데이터, 및 상기 영상 데이터를 전송하는 단계
S2250

[도23]


제 1 영상의 제 1 부호화 결과 생성된 영상 데이터, 원본 영상으로부터 상기 제 1 영상으로의 AI 다운스케일과 관련된 AI 데이터, 및 디블로킹 필터 설정 정보를 획득하는 단계
S2310
상기 영상 데이터를 제 1 복호화하여 상기 제 1 영상에 대응하는 제 2 영상을 획득하는 단계
S2320
복수의 DNN 설정 정보 중 상기 제 2 영상의 AI 업스케일을 위한 DNN 설정 정보를, 상기 AI 데이터에 기초하여 획득하는 단계
S2330
상기 획득된 DNN 설정 정보로 동작하는 업스케일용 DNN(Deep Neural Network)을 통해 상기 제 2 영상으로부터 AI 업스케일되고 상기 디블로킹 필터 설정 정보에 기초하여 디블로킹 필터링이 적용된 복원 영상을 생성하는 단계
S2340

INTERNATIONAL SEARCH REPORT

International application No.
PCT/KR2021/016236

A. CLASSIFICATION OF SUBJECT MATTER

H04N 19/86(2014.01)i; **H04N 19/82**(2014.01)i; **H04N 19/423**(2014.01)i; **H04N 19/124**(2014.01)i; **G06T 9/00**(2006.01)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 19/86(2014.01); G06N 3/04(2006.01); G06T 3/40(2006.01); H04N 19/00(2014.01); H04N 19/105(2014.01); H04N 19/124(2014.01); H04N 19/164(2014.01); H04N 19/172(2014.01); H04N 19/44(2014.01); H04N 19/51(2014.01)

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

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

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

eKOMPASS (KIPO internal) & keywords: 부호화(encoding), 복호화(decoding), 디블록킹(deblocking), 필터(filter), 환경(environment), DNN(Deep Neural Network)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2017-0180740 A1 (FASTVDO LLC) 22 June 2017 (2017-06-22) See paragraphs [0025], [0030] and [0049]-[0050]; and claims 1 and 6.	1-15
Y	US 10825140 B1 (SAMSUNG ELECTRONICS CO., LTD.) 03 November 2020 (2020-11-03) See column 31, line 5 - column 32, line 14; column 49, lines 45-61; claims 1-13; and figures 3, 9, 11 and 25.	1-15
A	KR 10-2016-0011580 A (RENESAS ELECTRONICS CORPORATION) 01 February 2016 (2016-02-01) See claims 1-13.	1-15
A	KR 10-2014-0116992 A (SAMSUNG ELECTRONICS CO., LTD. et al.) 07 October 2014 (2014-10-07) See claims 1-13; and figure 3.	1-15
A	KR 10-1611408 B1 (SAMSUNG ELECTRONICS CO., LTD.) 12 April 2016 (2016-04-12) See claims 2-12.	1-15

☐ Further documents are listed in the continuation of Box C. ☑ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"D" document cited by the applicant in the international application
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
10 February 2022	**10 February 2022**

Name and mailing address of the ISA/KR	Authorized officer
Korean Intellectual Property Office **Government Complex-Daejeon Building 4, 189 Cheongsa-ro, Seo-gu, Daejeon 35208**	
Facsimile No. **+82-42-481-8578**	Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/KR2021/016236

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
US 2017-0180740 A1	22 June 2017	US 10306238 B2	28 May 2019
		US 2014-0307785 A1	16 October 2014
		US 2020-0014936 A1	09 January 2020
		US 9609336 B2	28 March 2017
		WO 2014-172166 A2	23 October 2014
		WO 2014-172166 A3	28 May 2015
US 10825140 B1	03 November 2020	KR 10-2020-0140096 A	15 December 2020
		US 10817990 B1	27 October 2020
		WO 2020-246756 A1	10 December 2020
KR 10-2016-0011580 A	01 February 2016	CN 105306883 A	03 February 2016
		CN 105306883 B	03 March 2020
		EP 2977930 A1	27 January 2016
		EP 2977930 B1	28 October 2020
		JP 2016-025494 A	08 February 2016
		JP 6392572 B2	19 September 2018
		US 2016-0029021 A1	28 January 2016
		US 2017-0118468 A1	27 April 2017
		US 9554137 B2	24 January 2017
KR 10-2014-0116992 A	07 October 2014	US 2014-0286442 A1	25 September 2014
KR 10-1611408 B1	12 April 2016	KR 10-2010-0003167 A	07 January 2010
		US 2009-0324112 A1	31 December 2009
		US 8577165 B2	05 November 2013

국 제 조 사 보 고 서

국제출원번호
PCT/KR2021/016236

A. 발명이 속하는 기술분류(국제특허분류(IPC))

H04N 19/86(2014.01)i; **H04N 19/82**(2014.01)i; **H04N 19/423**(2014.01)i; **H04N 19/124**(2014.01)i; **G06T 9/00**(2006.01)i; **G06N 3/08**(2006.01)i

B. 조사된 분야

조사된 최소문헌(국제특허분류를 기재)

H04N 19/86(2014.01); G06N 3/04(2006.01); G06T 3/40(2006.01); H04N 19/00(2014.01); H04N 19/105(2014.01); H04N 19/124(2014.01); H04N 19/164(2014.01); H04N 19/172(2014.01); H04N 19/44(2014.01); H04N 19/51(2014.01)

조사된 기술분야에 속하는 최소문헌 이외의 문헌

한국등록실용신안공보 및 한국공개실용신안공보: 조사된 최소문헌란에 기재된 IPC
일본등록실용신안공보 및 일본공개실용신안공보: 조사된 최소문헌란에 기재된 IPC

국제조사에 이용된 전산 데이터베이스(데이터베이스의 명칭 및 검색어(해당하는 경우))

eKOMPASS(특허청 내부 검색시스템) & 키워드: 부호화(encoding), 복호화(decoding), 디블록킹(deblocking), 필터(filter), 환경(environment), DNN(Deep Neural Network)

C. 관련 문헌

카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
Y	US 2017-0180740 A1 (FASTVDO LLC) 2017.06.22 단락 [0025], [0030], [0049]-[0050]; 및 청구항 1, 6	1-15
Y	US 10825140 B1 (SAMSUNG ELECTRONICS CO., LTD.) 2020.11.03 컬럼 31, 라인 5 - 컬럼 32, 라인 14; 컬럼 49, 라인 45-61; 청구항 1-13; 및 도면 3, 9, 11, 25	1-15
A	KR 10-2016-0011580 A (르네사스 일렉트로닉스 가부시키가이샤) 2016.02.01 청구항 1-13	1-15
A	KR 10-2014-0116992 A (삼성전자주식회사 등) 2014.10.07 청구항 1-13; 및 도면 3	1-15
A	KR 10-1611408 B1 (삼성전자주식회사) 2016.04.12 청구항 2-12	1-15

☐ 추가 문헌이 C(계속)에 기재되어 있습니다.

☑ 대응특허에 관한 별지를 참조하십시오.

* 인용된 문헌의 특별 카테고리:

"A" 특별히 관련이 없는 것으로 보이는 일반적인 기술수준을 정의한 문헌
"D" 본 국제출원에서 출원인이 인용한 문헌
"E" 국제출원일보다 빠른 출원일 또는 우선일을 가지나 국제출원일 이후에 공개된 선출원 또는 특허 문헌
"L" 우선권 주장에 의문을 제기하는 문헌 또는 다른 인용문헌의 공개일 또는 다른 특별한 이유(이유를 명시)를 밝히기 위하여 인용된 문헌
"O" 구두 개시, 사용, 전시 또는 기타 수단을 언급하고 있는 문헌
"P" 우선일 이후에 공개되었으나 국제출원일 이전에 공개된 문헌
"T" 국제출원일 또는 우선일 후에 공개된 문헌으로, 출원과 상충하지 않으며 발명의 기초가 되는 원리나 이론을 이해하기 위해 인용된 문헌
"X" 특별한 관련이 있는 문헌. 해당 문헌 하나만으로 청구된 발명의 신규성 또는 진보성이 없는 것으로 본다.
"Y" 특별한 관련이 있는 문헌. 해당 문헌이 하나 이상의 다른 문헌과 조합하는 경우로 그 조합이 당업자에게 자명한 경우 청구된 발명은 진보성이 없는 것으로 본다.
"&" 동일한 대응특허문헌에 속하는 문헌

국제조사의 실제 완료일	국제조사보고서 발송일
2022년02월10일(10.02.2022)	**2022년02월10일(10.02.2022)**

ISA/KR의 명칭 및 우편주소

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

팩스 번호 **+82-42-481-8578**

심사관

양정록

전화번호 **+82-42-481-5709**

국 제 조 사 보 고 서
대응특허에 관한 정보

국제출원번호
PCT/KR2021/016236

국제조사보고서에서 인용된 특허문헌	공개일	대응특허문헌	공개일
US 2017-0180740 A1	2017/06/22	US 10306238 B2	2019/05/28
		US 2014-0307785 A1	2014/10/16
		US 2020-0014936 A1	2020/01/09
		US 9609336 B2	2017/03/28
		WO 2014-172166 A2	2014/10/23
		WO 2014-172166 A3	2015/05/28
US 10825140 B1	2020/11/03	KR 10-2020-0140096 A	2020/12/15
		US 10817990 B1	2020/10/27
		WO 2020-246756 A1	2020/12/10
KR 10-2016-0011580 A	2016/02/01	CN 105306883 A	2016/02/03
		CN 105306883 B	2020/03/03
		EP 2977930 A1	2016/01/27
		EP 2977930 B1	2020/10/28
		JP 2016-025494 A	2016/02/08
		JP 6392572 B2	2018/09/19
		US 2016-0029021 A1	2016/01/28
		US 2017-0118468 A1	2017/04/27
		US 9554137 B2	2017/01/24
KR 10-2014-0116992 A	2014/10/07	US 2014-0286442 A1	2014/09/25
KR 10-1611408 B1	2016/04/12	KR 10-2010-0003167 A	2010/01/07
		US 2009-0324112 A1	2009/12/31
		US 8577165 B2	2013/11/05