(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2021년07월02일
(11) 등록번호 10-2272660
(24) 등록일자 2021년06월29일

(51) 국제특허분류(Int. Cl.)
H04N 19/86 (2014.01) *G06N 3/08* (2006.01)
H04N 19/597 (2014.01)
(52) CPC특허분류
H04N 19/86 (2015.01)
G06N 3/08 (2013.01)
(21) 출원번호 10-2019-0178253
(22) 출원일자 2019년12월30일
심사청구일자 2019년12월30일
(56) 선행기술조사문헌
B.Salahieh et al., "Test Model 2 for Immersive Video," ISO/IEC JTC 1/SC 29/WG 11 N18577,(2019.08.02.)*
(뒷면에 계속)

(73) 특허권자
이화여자대학교 산학협력단
서울특별시 서대문구 이화여대길 52 (대현동, 이화여자대학교)
(72) 발명자
강제원
서울특별시 마포구 독막로 145, 104동 904호(창전동,서강쌍용예가아파트)
박상효
서울특별시 강서구 강서로5라길 77, 403호(화곡동, 자라하우스)
박현수
제주특별자치도 제주시 서광로29길 25, 103동 1001호(이도일동, 장원월드컵아파트)
(74) 대리인
특허법인(유한)아이시스

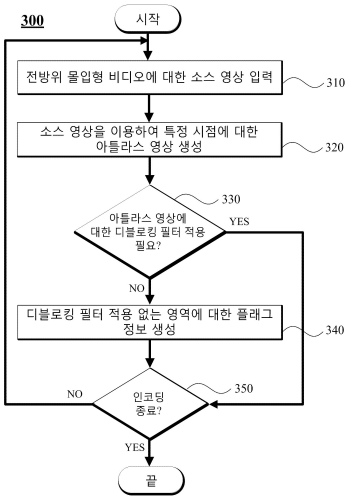
전체 청구항 수 : 총 6 항 심사관 : 조우연

(54) 발명의 명칭 **전방위 몰입형 비디오에 대한 적응적 디블로킹 필터 적용 방법 및 디코딩 장치**

(57) 요 약

전방위 몰입형 비디오에 대한 적응적 디블로킹 필터 적용 방법은 인코더가 전방위 몰입형 비디오에 대한 소스 영상들을 입력받는 단계, 상기 인코더가 상기 소스 영상들 사용하여 아틀라스(Atlas) 영상을 생성하는 단계 및 상기 인코더가 상기 아틀라스 영상에 대한 인코딩 단계를 포함한다. 상기 인코딩 단계에서 상기 인코더는 상기 아틀라스 영상의 종류 또는 특성에 따라 상기 아틀라스 영상 전체 또는 일부에 대한 디블로킹 필터(deblocking filter) 적용 여부를 결정한다.

대 표 도 - 도4



(52) CPC특허분류
H04N 19/597 (2015.01)

(56) 선행기술조사문헌
KR1020100046289 A*
이광순 외 3명, “이머시브미디어를 3DoF+ 비디오 부호화 표준 동향,” 전자통신동향분석, Vol. 34 No. 6, (2019.12.01.)
KR1020190061002 A
KR1020190072450 A
*는 심사관에 의하여 인용된 문헌

이 발명을 지원한 국가연구개발사업

과제고유번호	1711081006
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기술진흥센터
연구사업명	차세대(UHD)방송서비스활성화기술개발(R&D)
연구과제명	6DoF지원 초고화질 몰입형 비디오의 압축 및 전송 핵심 기술 개발
기 여 율	1/1
과제수행기관명	이화여자대학교 산학협력단
연구기간	2019.01.01 ~ 2019.12.31

공지예외적용 : 있음

명 세 서

청구범위

청구항 1

인코더가 전방위 몰입형 비디오에 대한 소스 영상들을 입력받는 단계;

상기 인코더가 상기 소스 영상들 사용하여 아틀라스(Atlas) 영상을 생성하는 단계; 및

상기 인코더가 상기 아틀라스 영상에 대한 인코딩 단계를 포함하되,

상기 인코딩 단계에서 상기 인코더는 상기 아틀라스 영상의 종류 또는 특성에 따라 상기 아틀라스 영상 전체 또는 일부에 대한 디블로킹 필터(deblocking filter) 적용 여부를 결정하고,

상기 영상의 종류는 기본 시점 영상, 추가 시점 영상 및 서로 다른 시점인 복수의 영상 중 어느 하나이고, 상기 인코더는 인코딩 과정에서 처리할 영상의 종류가 상기 기본 시점 영상인 경우 디블로킹 필터를 적용하고, 상기 추가 시점 영상인 경우 디블로킹 필터를 적용하지 않고, 상기 복수의 영상인 경우 디블로킹 필터를 적용하지 않고,

상기 특성은 딥러닝 모델을 이용하여 결정되되, 상기 인코더는 상기 아틀라스 영상의 특정 블록의 픽셀값을 딥러닝 모델에 입력하여 상기 특정 블록 또는 상기 특정 블록을 포함하는 슬라이스에 대한 디블로킹 필터 적용 여부를 결정하는 전방위 몰입형 비디오에 대한 적응적 디블로킹 필터 적용 방법.

청구항 2

삭제

청구항 3

삭제

청구항 4

삭제

청구항 5

제1항에 있어서,

상기 딥러닝 모델은 복수의 컨볼루션 계층 및 복수의 전연결 계층을 포함하고,

최초 CTU(Coding tree unit) 단위의 픽셀값을 입력받는 전방위 몰입형 비디오에 대한 적응적 디블로킹 필터 적용 방법.

청구항 6

제1항에 있어서,

상기 인코더는 프레임 단위, 픽쳐(picture) 헤더 단위, 슬라이스 단위 또는 블록 단위로 디블로킹 필터 제외 정보를 생성하는 전방위 몰입형 비디오에 대한 적응적 디블로킹 필터 적용 방법.

청구항 7

제1항에 있어서,

상기 인코더는 HEVC(High Efficiency Video Coding) 또는 VVC(Versatile Video Coding) 코덱에 따라 상기 아틀라스 영상을 인코딩하는 몰입형 비디오에 대한 적응적 디블로킹 필터 적용 방법.

청구항 8

디코더가 전방위 몰입형 비디오 영상을 입력받는 단계; 및

상기 디코더가 상기 영상에 포함된 아틀라스(Atlas) 영상에 대한 디코딩을 하는 단계를 포함하되,

상기 디코딩 단계에서 상기 디코더가 상기 아틀라스 영상 전체 또는 일부에 대한 디블로킹 필터(deblocking filter) 적용 여부를 나타내는 정보를 이용하여 디블로킹 필터링을 하고,

상기 정보는 아틀라스 영상의 종류 또는 특성을 포함하고,

상기 아틀라스 영상의 종류는 기본 시점 영상, 추가 시점 영상 및 서로 다른 시점인 복수의 영상 중 어느 하나이고, 상기 디코더는 디코딩할 영상의 종류가 상기 기본 시점 영상인 경우 디블로킹 필터를 적용하고, 상기 추가 시점 영상인 경우 디블로킹 필터를 적용하지 않고, 상기 복수의 영상인 경우 디블로킹 필터를 적용하지 않고,

상기 특성은 인코더에서 아틀라스 영상의 특정 블록의 픽셀값을 딥러닝 모델에 입력하여 결정되는 몰입형 비디오에 대한 적응적 디블로킹 필터 적용 방법.

청구항 9

삭제

청구항 10

삭제

청구항 11

삭제

청구항 12

아틀라스(Atlas) 영상을 포함하는 전방위 몰입형 비디오 영상에 대한 디코딩과정에서 적응적 디블로킹(deblocking)을 제어하는 프로그램을 저장하는 저장장치; 및

상기 프로그램을 사용하여 전방위 몰입형 비디오 영상에 대한 디코딩을 수행하는 프로세서를 포함하되,

상기 프로세서는 상기 아틀라스 영상 전체 또는 일부에 대한 디블로킹 필터(deblocking filter) 적용 여부를 나타내는 정보를 이용하여 디블로킹 필터링을 하고,

상기 정보는 아틀라스 영상의 종류 또는 특성을 포함하고,

상기 아틀라스 영상의 종류는 기본 시점 영상, 추가 시점 영상 및 서로 다른 시점인 복수의 영상 중 어느 하나이고, 상기 프로세서는 디코딩할 영상의 종류가 상기 기본 시점 영상인 경우 디블로킹 필터를 적용하고, 상기 추가 시점 영상인 경우 디블로킹 필터를 적용하지 않고, 상기 복수의 영상인 경우 디블로킹 필터를 적용하지 않고,

상기 특성은 인코더에서 아틀라스 영상의 특정 블록의 픽셀값을 딥러닝 모델에 입력하여 결정되는 몰입형 비디오에 대한 적응적 디블로킹을 수행하는 디코딩 장치.

청구항 13

삭제

청구항 14

삭제

발명의 설명

기 술 분 야

[0001] 이하 설명하는 기술은 전방위 몰입형 비디오에 대한 코딩 기법에 관한 것이다. 특히, 이하 설명하는 기술은 전

방위 몰입형 비디오에 대한 적응적 디블로킹 필터 적용 기법에 관한 것이다.

배 경 기 술

[0002] MPEG에서는 몰입형 미디어를 위한 차세대 프로젝트인 MPEG-I (Immersive) 표준화를 시작하였다. 특히 ISO/IEC 23090 Part 7 (Metadata for Immersive Media (Video))에서는 단계별 로드맵을 따라 표준화 과정이 진행 중이다. 현재 3DoF (Degree of Freedom)+ 표준이 진행중이며, 2022년까지 6DoF 표준 제정을 목표로 한다.

[0003] 표준화 기구는 몰입형 비디오 코덱 기술을 평가하기 위한 테스트 소프트웨어 TMIV(Test Model for Immersive Video)를 개발하였다. TMIV는 기준 시점과 추가 시점의 영상에서 공통된 신호를 제거한 잔차 패치를 생성하는 아틀라스(atlas) 기술을 포함한다. TMIV는 기본적으로 HEVC(High Efficiency Video Coding) 코딩에 기반하여 영상을 압축한다.

선행기술문헌

특허문헌

[0004] (특허문헌 0001) 한국공개특허 제10-2017-0110557호

발명의 내용

해결하려는 과제

[0005] HEVC 코딩은 코딩 블록의 양자화 적용에 따른 블록 경계면을 필터링하는 디블로킹 필터(deblocking filter)를 사용한다. 다만, 아틀라스 영상은 디블로킹 필터링이 영상 품질을 저하시킬 수 있는 문제점이 있다.

[0006] 이하 설명하는 기술은 몰입형 비디오 코덱을 위한 적응적 디블로킹 필터링 기법을 제공하고자 한다.

과제의 해결 수단

[0007] 전방위 몰입형 비디오에 대한 적응적 디블로킹 필터 적용 방법은 인코더가 전방위 몰입형 비디오에 대한 소스 영상들을 입력받는 단계, 상기 인코더가 상기 소스 영상들 사용하여 아틀라스(Atlas) 영상을 생성하는 단계 및 상기 인코더가 상기 아틀라스 영상에 대한 인코딩 단계를 포함한다. 상기 인코딩 단계에서 상기 인코더는 상기 아틀라스 영상의 종류 또는 특성에 따라 상기 아틀라스 영상 전체 또는 일부에 대한 디블로킹 필터(deblocking filter) 적용 여부를 결정한다.

[0008] 다른 측면에서 몰입형 비디오에 대한 적응적 디블로킹 필터 적용 방법은 디코더가 전방위 몰입형 비디오 영상을 입력받는 단계 및 상기 디코더가 상기 영상에 포함된 아틀라스(Atlas) 영상에 대한 디코딩을 하는 단계를 포함한다. 상기 디코딩 단계에서 상기 디코더가 상기 아틀라스 영상 전체 또는 일부에 대한 디블로킹 필터(deblocking filter) 적용 여부를 나타내는 정보를 이용하여 디블로킹 필터링을 한다.

발명의 효과

[0009] 이하 설명하는 기술은 아틀라스 영상에 대한 적응적 디블로킹 필터링을 하여 보다 품질이 좋은 몰입형 비디오 영상을 제공한다.

도면의 간단한 설명

[0010] 도 1은 TMIV 인코더 및 디코더 구성에 대한 예이다.

도 2는 아틀라스 영상에 대한 예이다.

도 3은 아틀라스 영상에 대한 디블로킹 필터 적용에 대한 예이다.

도 4는 적응적 디블로킹 필터 적용 과정에 대한 순서도의 예이다.

도 5는 적응적 디블로킹 필터를 판단하는 딥러닝 모델에 대한 예이다.

도 6은 적응적 디블로킹 필터를 사용하는 디코딩 장치의 구성에 대한 예이다.

발명을 실시하기 위한 구체적인 내용

[0011] 이하 설명하는 기술은 다양한 변경을 가할 수 있고 여러 가지 실시례를 가질 수 있는 바, 특정 실시례들을 도면에 예시하고 상세하게 설명하고자 한다. 그러나, 이는 이하 설명하는 기술을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 이하 설명하는 기술의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다.

[0012] 제1, 제2, A, B 등의 용어는 다양한 구성요소들을 설명하는데 사용될 수 있지만, 해당 구성요소들은 상기 용어들에 의해 한정되지는 않으며, 단지 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 사용된다. 예를 들어, 이하 설명하는 기술의 권리 범위를 벗어나지 않으면서 제1 구성요소는 제2 구성요소로 명명될 수 있고, 유사하게 제2 구성요소도 제1 구성요소로 명명될 수 있다. 및/또는 이라는 용어는 복수의 관련된 기재된 항목들의 조합 또는 복수의 관련된 기재된 항목들 중의 어느 항목을 포함한다.

[0013] 본 명세서에서 사용되는 용어에서 단수의 표현은 문맥상 명백하게 다르게 해석되지 않는 한 복수의 표현을 포함하는 것으로 이해되어야 하고, "포함한다" 등의 용어는 설명된 특징, 개수, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것이 존재함을 의미하는 것이지, 하나 또는 그 이상의 다른 특징들이나 개수, 단계 동작 구성요소, 부분품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 배제하지 않는 것으로 이해되어야 한다.

[0014] 도면에 대한 상세한 설명을 하기에 앞서, 본 명세서에서의 구성부들에 대한 구분은 각 구성부가 담당하는 주기능 별로 구분한 것에 불과함을 명확히 하고자 한다. 즉, 이하에서 설명할 2개 이상의 구성부가 하나의 구성부로 합쳐지거나 또는 하나의 구성부가 보다 세분화된 기능별로 2개 이상으로 분화되어 구비될 수도 있다. 그리고 이하에서 설명할 구성부 각각은 자신이 담당하는 주기능 이외에도 다른 구성부가 담당하는 기능 중 일부 또는 전부의 기능을 추가적으로 수행할 수도 있으며, 구성부 각각이 담당하는 주기능 중 일부 기능이 다른 구성부에 의해 전담되어 수행될 수도 있음은 물론이다.

[0015] 또, 방법 또는 동작 방법을 수행함에 있어서, 상기 방법을 이루는 각 과정들은 문맥상 명백하게 특정 순서를 기재하지 않은 이상 명기된 순서와 다르게 일어날 수 있다. 즉, 각 과정들은 명기된 순서와 동일하게 일어날 수도 있고 실질적으로 동시에 수행될 수도 있으며 반대의 순서대로 수행될 수도 있다.

[0017] 이하 설명하는 기술은 몰입형 비디오에 대한 코딩 기법에 해당한다. 몰입형 비디오는 시점 자유도가 있는 영상이다. 시점 자유도(DoF)는 몇 가지 유형으로 구분할 수 있다. 3DoF는 특정 지점을 기준으로 어느 방향으로든 자유롭게 시청할 수 있는 방식이다. 3DoF는 사용자 머리를 중심으로 세 가지 방향(yaw, pitch 및 roll)의 자유도를 제공한다. 3DoF+는 3DoF에 추가적으로 머리를 중심으로 움직임의 자유도를 제공한다. 3DoF+는 운동 시차(motion parallax)와 실감도에 강인한 요소에 대한 지각 효과를 가져온다. 6DoF는 시청자의 머리 회전 운동뿐만 아니라 3차원 공간에서의 이동에 따른 자유로운 시점 변환을 제공한다. 이하 몰입형 비디오 또는 전방위 몰입형 비디오는 3DoF+ 또는 6DoF와 같이 사용자의 움직임에 따른 영상을 제공하는 비디오를 의미한다고 정의한다.

[0019] 먼저, TMIV에서의 코딩 과정에 대하여 간략하게 설명한다. 몰입형 비디오 부호화를 위한 전/후처리의 핵심은 제한된 공간에서의 다양한 방향 및 각도로 획득한 다수의 뷰 비디오 간에 존재하는 많은 중복성을 제거하여 압축 전에 미리 압축 대상이 되는 데이터를 최소화하는 것이다.

[0021] 도 1은 TMIV 인코더 및 디코더 구성에 대한 예이다.

[0022] 도 1(A)는 인코더에 대한 예이다. 인코더는 복수의 카메라가 획득한 영상을 처리하기 위하여 영상 압축에 해당하는 인코딩 과정 전에 일정하게 영상을 처리한다. 다수의 소스 시점 비디오(텍스처 및 깊이)와 해당 시점의 카메라 파라미터가 입력된다. 시점 최적화기(View Optimizer, 110)는 카메라 파라미터정보를 기반으로 입력 비디오를 그대로 부호화할 기본 시점(Basic View) 영상과 기본 시점과 중복되는 영역이 제거될 추가 시점(Additional View) 영상으로 분류한다.

[0023] 아틀라스 생성기(Atlas Constructor, 120)는 분류된 기본 시점 영상과 추가 시점 영상을 입력받는다. Pruner는 마스크들을 사용하여 기본 시점 영상을 보존하고, 추가 시점 영상은 중복되는 부분을 제거한다. Aggregator는 중복성이 제거되고 남은 영역을 일정 시간만큼 누적시킨다. Patch packer는 각 영역들을 감싸는 사각형 영역으로 정의되는 각 패치(pathc)들을 하나의 프레임으로 패킹(packing)한다. 이 과정을 통하여 아틀라스 생성기(120)는 인코더(130)에 입력된 아틀라스 영상을 생성한다. 아틀라스 생성기(120)는 아트라스를 생성하고, 아틀라스 구성정보 리스트 및 카메라 파라미터 리스트를 출력한다.

[0024] 코어 인코더(130)는 입력된 아틀라스 영상을 일정하게 인코딩한다. 코어 인코더(130)는 입력된 영상을 특정 코

덱에 따라 압축하는 구성이다. 압축 방식은 다양한 방식을 사용할 수 있다. 예컨대, 코어 인코더(130)는 HEVC(High Efficiency Video Coding) 코덱에 따라 아틀라스 영상을 인코딩할 수 있다. 나아가, 코어 인코더(130)는 VVC(Versatile Video Coding) 코덱에 따라 아틀라스 영상을 인코딩할 수 있다.

[0025] 또한, 메타데이터 작성기(Metadata Composer, 140)는 아틀라스 구성 정보를 나타내는 아틀라스 파라미터 리스트와 카메라 파라미터 리스트를 받아 메타데이터를 생성한다. 메타데이터는 디코더에서 가상 시점을 복원할 때 사용된다.

[0027] 도 1(B)는 디코더(100)에 대한 예이다. 디코더(200)는 복수의 카메라가 획득한 영상을 처리하기 위하여 영상 복호에 해당하는 디코딩 과정 후에 일정하게 영상을 처리한다.

[0028] 코어 디코더(210)는 먼저 입력되는 영상을 일정한 방식으로 디코딩한다. 코어 디코더(210)는 인코딩 방식에 대응하는 코덱을 사용하여 영상을 디코딩한다. 예컨대, 코어 디코더(210)는 HEVC 코덱에 따라 아틀라스 영상을 디코딩할 수 있다. 나아가, 코어 디코더(210)는 VVC 코덱에 따라 아틀라스 영상을 디코딩할 수 있다. 코어 디코더(210)는 아틀라스 영상을 디코딩한다.

[0029] 메타데이터 파서(Metadata Parser, 220)는 입력 스트림에 포함된 메타데이터를 일정하게 처리한다. 메타데이터 파서(220)는 메타데이터에 포함된 아틀라스 파라미터 리스트 및 카메라 파라미터 리스트를 구분한다.

[0030] 패치 점유 지도 생성기(Atlas Patch Occupancy Map Generator, 230)는 사용자가 원하는 시점의 영상을 합성하기 위하여 입력된 아클라스 패킹 정보를 기준으로 점유 맵을 생성한다. 점유 맵은 아클라스에 패킹된 각 패치가 가지고 있는 유효 영역만을 시그널링하여, 가상시점 합성 시 필요한 시점의 유효 화소들을 이용하여 영상을 합성할 수 있게 한다.

[0031] 렌더러(Renderer, 240)는 아틀라스 영상, 점유 맵, 파라미터 리스트 및 카메라 파라미터 리스트를 이용하여 사용자가 원하는 시점의 영상을 합성한다. 렌더러(240)는 컨트롤러(Controller), 합성기(Synthesizer) 및 인페인터(Inpainter)로 구성되어 주로 중간 시점의 합성을 위하여 사용이 된다. 합성기는 당양한 알고리즘을 사용하여 영상을 합성할 수 있다. 합성기는 인접한 시점의 영상 외에도 멀리 있는 영상도 선택하여 합성할 수 있다. 인페인터는 누락된 픽셀을 보간하는 역할을 한다

[0033] 도 2는 아틀라스 영상에 대한 예이다. 도 2(A)는 기본 시점 영상에 대한 예이고, 도 2(B)는 추가 시점 영상에 대한 예이다. 도 2(A) 및 도 2(B)에서 각각 좌측 영상은 텍스쳐 영상이고, 우측 영상은 깊이 영상이다. 아틀라스 생성기는 시점 최적화 단계에서 선택된 기본 시점들을 하나의 완전한 단일 아틀라스로 구성한다. 아틀라스 생성기는 도 2(A)와 같이 기본 시점의 텍스쳐 및 깊이 정보를 원본과 같이 구성한다. 아틀라스 생성기는 추가 시점들은 기본 시점 기준으로 중복된 부분을 제거한 잔차 패치로 구성한다. 아틀라스 생성기는 도 2(B)와 같이 추가 시점에 대해서는 텍스쳐 및 깊이 정보 모두 잔차 정보를 블록 패치 형태로 구성한다. 한편, 디코더에서는 잔차 블록 패치를 재구성하여 추가 시점의 비디오 프레임을 복원한다.

[0035] HEVC 또는 VVC의 경우, 코딩 블록의 양자화 적용에 따른 블록 경계면을 필터링하는 디블로킹 필터가 주관적 화질 향상을 위해 필수이며, 중요한 코딩 툴 중 하나로 여겨진다. 기본 시점과 추가 시점의 잔차 신호를 패킹할 때, 잔차 신호들의 경계면 때문에 주파수 도메인에서 상당한 부분이 고주파수에 해당하게 된다. 이때 디블로킹 필터를 적용하게 되면 보존되어야 하는 잔차 신호의 경계 부분이 보존되기 어렵다. 아틀라스 영상의 경우, 오히려 이러한 경계면을 수정(디블로킹 필터링)이 원치 않는 코딩 에러를 유발할 수 있고, 추후 아틀라스로 추가 시점 영상들을 생성하려 할 때 그 에러가 치명적으로 영향을 줄 수 있다.

[0036] 도 3은 아틀라스 영상에 대한 디블로킹 필터 적용에 대한 예이다. 도 3(A)는 아틀라스 원본 영상이고, 도 3(B)는 도 3(A)의 영상에 디블로킹 필터링이 적용된 영상이고, 도 3(B)는 도 3(A)의 영상에 디블로킹 필터링을 적용하지 않은 영상이다. 도 3을 살펴보면, 디블로킹 필터를 적용하게 되면, 원치 않는 블러링 효과(blurring noise)와 픽셀 값이 번지는 에러(bleeding)가 발생하는 것을 확인할 수 있다.

[0037] 이와 같이 TMIV의 경우, 기존의 동영상 기반의 압축기술이 오히려 성능을 저해할 수 있다. 디블로킹 필터에 대한 시점별 VVC 압축 결과를 실험해 보았다. 동일한 VTM 버전에 동일한 TMIV의 MIV 모드 영상들에 대하여 기존 VTM 대비 디블로킹 필터를 사용하지 않게 한 VTM 인코딩의 실험을 진행하였다. 아래 표 1에서 보는 바와 같이, VVC로 TMIV를 압축할 경우, 오히려 디블로킹 필터를 적용하지 않는 경우가 전반적인 압축률 상승이 나타나는 것을 볼 수 있다. 아틀라스의 경우에 약 7.83%의 압축률 향상이 나타나는 것을 볼 수 있다. 디블로킹 필터와 같이 일반적인 동영상에서 압축 효율을 증진시키던 기술이 아틀라스의 경우 오히려 압축률을 떨어뜨리는 작용을 할

수 있음을 확인하였다.

표 1

Sequence Name / View	QP	TMIV+VVC				TMIV+VVC (without deblocking filter)			
		Total kbps	Y-PSNR	U-PSNR	V-PSNR	Total kbps	Y-PSNR	U-PSNR	V-PSNR
ClassroomVideo / Basic view	27	17271.93	35.49	49.03	50.95	17284.87	35.51	49.10	50.99
	32	4116.47	34.73	47.59	49.44	4117.05	34.74	47.63	49.50
	37	1169.73	33.99	46.04	47.91	1172.17	33.99	46.03	47.88
	42	508.67	33.14	44.79	46.44	509.46	33.13	44.70	46.51
BD-rate		-1.18%							
ClassroomVideo / Atlas	27	4614.59	41.47	52.06	53.93	4615.15	41.52	52.47	54.37
	32	1613.33	40.47	49.89	51.48	1611.21	40.56	50.26	51.92
	37	642.61	39.34	47.42	49.09	640.28	39.47	47.75	49.37
	42	355.80	37.95	45.75	47.17	353.24	38.15	45.86	47.28
BD-rate		-7.83%							

[0038]

[0040] 따라서, 몰입형 비디오의 코딩 과정에 이미 개발된 코덱(HEVC) 또는 현재 개발중인 영상 코덱(VVC)을 적용하는 경우, 디블로킹 필터링은 몰입형 비디오를 위한 적응적 기법이 필요하다. 다음과 같은 적응적 디블로킹 필터링 기법을 제안한다.

[0042] 도 4는 적응적 디블로킹 필터 적용 과정(300)에 대한 순서도의 예이다. 도 4는 인코더에서 수행하는 디블로킹 필터링 과정에 대한 예이다.

[0043] 인코더는 전방위 몰입형 비디오에 대한 소스 영상을 입력받는다(310). 인코더는 소스 영상을 이용하여 특정 시점에 대한 아틀라스 영상을 생성한다(320). 특정 시점은 특정한 기본 시점과 기본 시점과 함께 전달되는 추가 시점을 포함한다.

[0044] 인코더는 아틀라스 영상에 대한 디블로킹 필터 적용 여부를 결정한다(330). 예컨대, 도 1의 코어 인코더(130)는 입력되는 아틀라스 영상의 정보, 또는 아틀라스 영상을 별도로 분석한 결과를 기준으로 디블로킹 필터링을 여부를 결정할 수 있다. 이하 적응적 디블로킹 필터링을 위한 기준을 몇 가지 설명한다.

[0045] (1) 비디오 프레임을 전달하는 단위로 픽처 헤더 또는 슬라이스 헤더 등이 있다. 프레임을 전달하는 단위는 해당 단위별로 현재 비디오 프레임이 어떠한 정보를 포함하고 있는지를 표시한다. TMIV에서도 현재 비디오가 기본 시점인지 추가 시점인지를 표시하는 정보가 있다. 예컨대, 패치 아틀라스에 대해 Atlas parameters syntax 중 num_patches_minus1가 어떤 시점인지 정보를 나타낼 수 있다. 인코더는 이 값이 0과 같으면 기본 시점이므로 디블로킹 필터를 적용하고, 그렇지 않은 경우에는 추가 시점이므로 디블로킹 필터를 적용하지 않을 수 있다.

[0046] (2) 모든 시점이 단독으로 최종 영상을 구성하지 않고 모든 시점이 결합하여 최종 영상을 복원하는 경우가 있다. 인코더는 이와 같은 경우, 모든 비디오 프레임에 대해서 디블로킹 필터를 적용하지 않을 수 있다.

[0047] (3) 만약 아틀라스 시점인지 파악하기 어려운 상황이거나, 아틀라스 시점이라 하더라도 패치가 복잡하고 정교하여 디블로킹 필터를 적용하는 것이 유리한지 여부를 정밀하게 평가해야 할 수도 있다. 이 경우, 블록 단위로 디블로킹 필터의 적응적 적용 여부를 결정할 수 있다. 예를 들어, 심층 신경망학습 기반의 모델을 사용하여 적응적 디블로킹 필터링을 결정할 수 있다.

[0048] 도 5는 적응적 디블로킹 필터를 판단하는 딥러닝 모델(400)에 대한 예이다. 도 5는 VVC의 기본 CTU(Coding tree unit) 크기인 128×128을 가정하여 구성한 모델의 예이다. 도 5는 CNN(Convolutional Neural Network) 구조의 일종이다. 물론, 딥러닝 모델은 구체적은 구성은 도 5와는 다를 수도 있다. 도 5의 딥러닝 모델은 3개의 컨볼루션 블록(411, 412 및 413) 및 두 개의 전연결 계층(421 및 422)으로 구성된다. 각 컨볼루션 블록은 하나의 컨볼루션 계층(conv)와 풀링 계층(pool)로 구성된 예이다.

[0049] 컨볼루션 계층은 입력 영상에 대한 컨볼루션 연산을 통해 특징맵(feature map)을 출력한다. 이때 컨볼루션 연산을 수행하는 필터(filter)를 커널(kernel) 이라고도 부른다. 필터의 크기를 필터 크기 또는 커널 크기라고 한다. 도 5에서 필터 크기는 128일 수 있다. 이는 전술한 바와 같이 입력 블록의 크기에 따라 달라질 수 있다. 커널을 구성하는 연산 파라미터(parameter)를 커널 파라미터(kernel parameter), 필터 파라미터(filter parameter), 또는 가중치(weight)라고 한다.

[0050] 컨볼루션 계층은 컨볼루션 연산과 비선형 연산을 수행한다. 컨볼루션 연산은 일정한 크기의 윈도우에서 수행된다. 윈도우는 영상의 좌측 상단에서 우측 하단까지 한 칸씩 이동할 수 있고, 한 번에 이동하는 이동 크기를 조절할 수 있다. 이동 크기를 스트라이드(stride)라고 한다. 컨볼루션 계층은 입력이미지에서 윈도우를 이동하면서 입력이미지의 모든 영역에 대하여 컨볼루션 연산을 수행한다. 컨볼루션 계층은 영상의 가장 자리에 패딩(padding)을 하여 컨볼루션 연산 후 입력 영상의 차원을 유지할 수 있다. 비선형 연산 계층(nonlinear operation layer)은 뉴런(노드)에서 출력값을 결정하는 계층이다. 비선형 연산 계층은 전달 함수(transfer function)를 사용한다. 전달 함수는 Relu, sigmoid 함수 등이 있다.

[0051] 풀링 계층(pooling layer)은 컨볼루션 계층에서의 연산 결과로 얻은 특징맵을 서브 샘플링(sub sampling)한다. 풀링 연산은 최대 풀링(max pooling)과 평균 풀링(average pooling) 등이 있다. 최대 풀링은 윈도우 내에서 가장 큰 샘플 값을 선택한다. 평균 풀링은 윈도우에 포함된 값의 평균 값으로 샘플링한다.

[0052] 전연결 계층은 최종적으로 입력 영상을 분류한다. 전연결 계층은 이전 컨볼루션 계층에서 출력하는 값을 모두 입력받아 최종적인 분류를 한다. 예컨대, 전연결 계층은 소프트맥스(softmax) 함수를 사용하여 분류 결과를 출력할 수 있다. 도 5에서 전열결 계층(421)은 다수의 출력(예컨대, 128개)을 출력할 수 있고, 전열결 계층(422)는 두 개의 출력(이진 분류)을 할 수도 있다.

[0053] 인코더는 특정 블록(CUT) 내의 픽셀값들을 딥러닝 모델에 입력으로 전달한다. 이후, 사전에 학습된 딥러닝 모델은 현재 입력된 블록이 아틀라스 영역인지 아닌지 혹은 경계면들이 복잡하게 존재하는 영역인지 아닌지를 판별한다. 인코더는 딥러닝 모델의 출력 결과를 기준으로, 해당 블록에 대한 디블로킹 필터링 여부를 결정할 수 있다. 즉, 인코더는 HEVC 또는 VVC 코덱 기반의 코어 인코더에 해당 블록 또는 해당 슬라이스에 대한 디블로킹 필터링 여부를 전달할 수 있다.

[0054] 인코더는 디블로킹 필터링이 필요한지 여부를 결정한다. 디블로킹 필터링이 필요 없다면(330의 NO), 디블로킹 필터링 적용 없는 영역에 대한 플래그(flag) 정보를 생성할 수 있다(340). 플래그는 영상을 디코딩하는 경우, 해당 영역에 디블로킹 필터링을 적용하지 않았다는 정보를 전달하는데 사용된다.

[0055] VVC에서 아틀라스에 해당하는 영상에 대한 디블로킹 필터링를 제외하기 위하여, 예컨대 (1)프레임 별로 slice_deblocking_filter_disabled_flag를 True값으로 설정할 수 있다. 또는, (2) 픽쳐 헤더단위에서 이를 제어하기 위하여, pic_deblocking_filter_disabled_flag 를 True값으로 설정할 수 있다. (3)또는 Picture parameter set(PPS) 및 상위 레벨에서 디블로킹 필터링 적용을 하지 않게 하는 pps_deblocking_filter_disabled_flag를 True로 설정할 수 있다. 디블로킹 필터링을 수행하지 않으면서 동시에 관련 syntax 요소를 전송하지 않게 함으로써 압축 효율을 높일 수 있다. HEVC 경우에도 이와 유사한 방식으로 디블로킹 필터링 적용 여부에 대한 정보를 생성할 수 있다.

[0056] 인코더는 생성한 플래그 정보를 전술한 바와 같이 영상이 전달되는 단위별로 전달할 수 있다. 인코더는 프레임, 슬라이스, 픽쳐 또는 상위 레벨에서 디블로킹 필터링 비적용에 대한 플래그 정보를 저장하여 영상과 함께 인코딩할 수 있다.

[0057] 인코더는 인코딩이 종료(350에서 판단)할 때까지 이와 같은 동작을 반복한다.

[0059] 디코더는 인코더가 영상과 함께 인코딩한 상기 플래그 정보를 이용하여 아틀라스 영상을 디코딩할 수 있다. 디코더는 전방위 몰입형 비디오 영상을 입력받고, 아틀라스(Atlas) 영상에 대한 디코딩을 한다. 이때, 디코더는 아틀라스 영상 전체 또는 일부에 대한 디블로킹 필터 적용 여부를 나타내는 정보(플래그)를 이용하여 해당 영역에 대한 디블로킹 필터링을 적응적으로 한다.

[0060] 도 6은 적응적 디블로킹 필터를 사용하는 디코딩 장치(500)의 구성에 대한 예이다. 디코딩 장치(500)는 코어 디코더의 내부 컴포넌트일 수 있다. 디코딩 장치(500)는 아날로그 회로 또는 디지털 회로로 구현될 수 있다. 도 5는 적응적 디블로킹 필터링을 위한 코드가 임베드된 장치를 예로 도시한다.

[0061] 디코딩 장치(500)는 저장 장치(510), 메모리(520), 프로세서(520) 및 인터페이스 장치(530)를 포함한다. 나아가, 디코딩 장치는 통신장치(550) 또는 출력장치(560)를 포함할 수도 있다.

[0062] 저장 장치(510)는 아틀라스 영상을 포함하는 전방위 몰입형 비디오 영상에 대한 디코딩과정에서 적응적 디블로킹(deblocking)을 제어하는 프로그램을 저장한다. 저장 장치(510)는 입력되는 영상을 저장할 수 있다. 저장장치(510)로 하나의 블록을 도시하였지만, 저장 장치(510)는 임시 정보를 저장하는 메모리와 움직임 예측 코드를 저장하는 저장 매체로 구분될 수도 있다.

[0063] 메모리(520)는 디코딩 과정 및 적응적 디블로킹 필터링 과정에 생성되거나 필요한 정보를 임시로 저장할 수 있다.

[0064] 인터페이스 장치(540)는 디코더 내부에서 데이터 및 명령을 전달하는 구성을 의미한다. 인터페이스 장치(540)는 내부 통신을 위한 물리적 장치 및 통신 프로토콜을 포함할 수 있다. 인터페이스 장치(540)는 정보를 전달하는 회선, 정보 전달을 제어하는 구성(스위치 등), 정보를 가공하는 구성(합산기 등) 등으로 구성될 수 있다.

[0065] 통신장치(550)는 유선 또는 무선 통신을 통해 외부 객체로부터 일정한 정보를 수신할 수 있다. 예컨대, 통신장치(500)는 서비스 서버가 제공하는 몰입형 비디오 영상을 수신할 수 있다.

[0066] 프로세서(530)는 주어진 데이터 내지 정보를 처리하는 구성을 의미한다. 프로세서(530)는 일종의 연산 장치이다. 프로세서(530)는 저장 장치(510)에 저장된 코드 내지 프로그램을 이용하여 아틀라스 영상에 대한 적응적 필터링을 수행한다.

[0067] 프로세서(530)는 아틀라스 영상 전체 또는 일부에 대한 디블로킹 필터 적용 여부를 나타내는 정보(플래그)를 이용하여 해당 영상 또는 영상의 특정 영역(프레임, 슬라이스, 블록 단위 등)에 대한 디블로킹을 하지 않을 수 있다. 예컨대, 프로세서(530)는 디코딩할 영상이 기본 시점 영상인 경우 디블로킹 필터를 적용하고, 추가 시점 영상인 경우 디블로킹 필터를 적용하는 않을 수 있다. 프로세서(530)는 디코딩할 영상이 상기 복수의 영상을 이용하여 복원되는 영상인 경우, 복수의 영상에 대한 디블로킹 필터를 적용하지 않을 수 있다.

[0069] 또한, 상술한 바와 같은 적응적 디블로킹 적용 방법, 적응적 디블로킹 기반의 인코딩 방법, 적응적 디블로킹 기반의 디코딩 방법은 컴퓨터에서 실행될 수 있는 실행가능한 알고리즘을 포함하는 프로그램(또는 어플리케이션)으로 구현될 수 있다. 상기 프로그램은 일시적 또는 비일시적 판독 가능 매체(non-transitory computer readable medium)에 저장되어 제공될 수 있다.

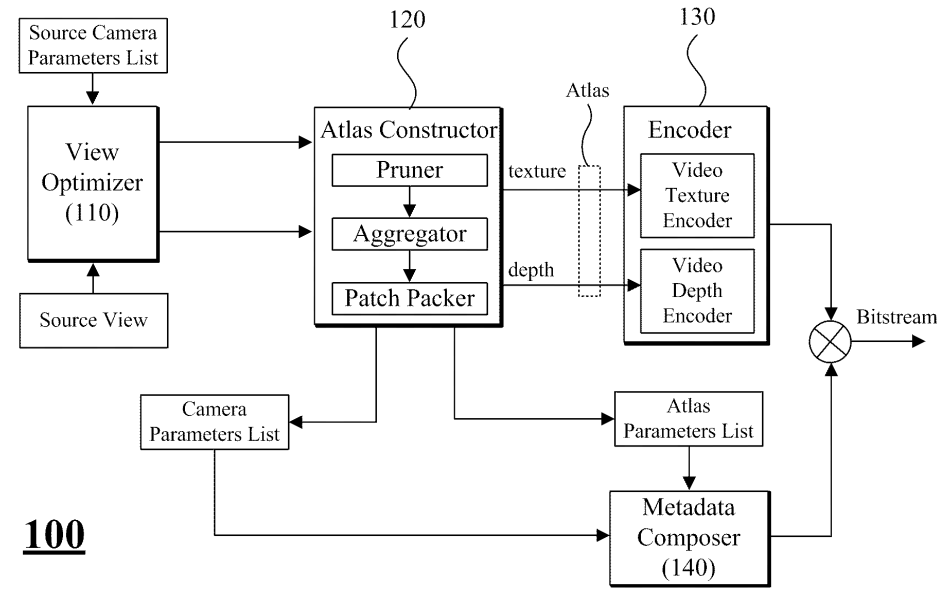
[0070] 비일시적 판독 가능 매체란 레지스터, 캐쉬, 메모리 등과 같이 짧은 순간 동안 데이터를 저장하는 매체가 아니라 반영구적으로 데이터를 저장하며, 기기에 의해 판독(reading)이 가능한 매체를 의미한다. 구체적으로는, 상술한 다양한 어플리케이션 또는 프로그램들은 CD, DVD, 하드 디스크, 블루레이 디스크, USB, 메모리카드, ROM (read-only memory), PROM (programmable read only memory), EPROM(Erasable PROM, EPROM) 또는 EEPROM(Electrically EPROM) 또는 플래시 메모리 등과 같은 비일시적 판독 가능 매체에 저장되어 제공될 수 있다.

[0071] 일시적 판독 가능 매체는 스태틱 램(Static RAM, SRAM), 다이내믹 램(Dynamic RAM, DRAM), 싱크로너스 디램(Synchronous DRAM, SDRAM), 2배속 SDRAM(Double Data Rate SDRAM, DDR SDRAM), 증강형 SDRAM(Enhanced SDRAM, ESDRAM), 동기화 DRAM(Synclink DRAM, SLDRAM) 및 직접 램버스 램(Direct Rambus RAM, DRRAM) 과 같은 다양한 RAM을 의미한다.

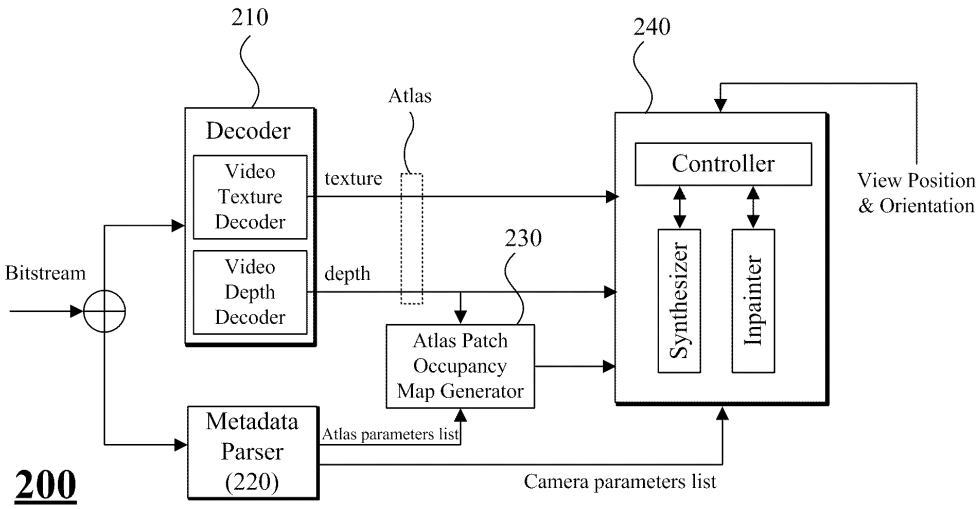
[0074] 본 실시례 및 본 명세서에 첨부된 도면은 전술한 기술에 포함되는 기술적 사상의 일부를 명확하게 나타내고 있는 것에 불과하며, 전술한 기술의 명세서 및 도면에 포함된 기술적 사상의 범위 내에서 당업자가 용이하게 유추할 수 있는 변형 예와 구체적인 실시례는 모두 전술한 기술의 권리범위에 포함되는 것이 자명하다고 할 것이다.

도면

도면1


Source Camera Parameters List
View Optimizer (110)
Source View
120
Atlas Constructor
Pruner
Aggregator
Patch Packer
Atlas
texture
depth
130
Encoder
Video Texture Encoder
Video Depth Encoder
Bitstream
Camera Parameters List
Atlas Parameters List
Metadata Composer (140)
100

(A)

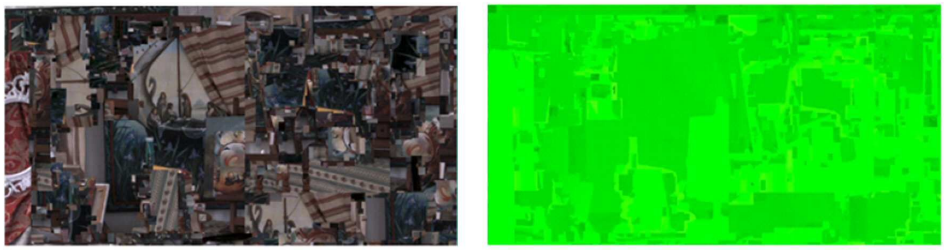

210
Decoder
Video Texture Decoder
Video Depth Decoder
Atlas
texture
depth
230
Atlas Patch Occupancy Map Generator
240
Controller
Synthesizer
Inpainter
View Position & Orientation
Bitstream
Metadata Parser (220)
Atlas parameters list
Camera parameters list
200

(B)

도면2

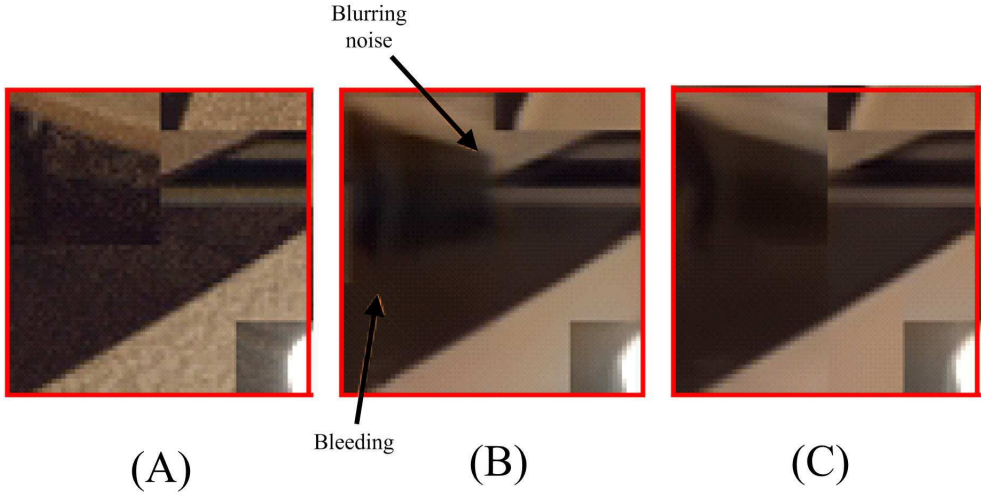


(A)

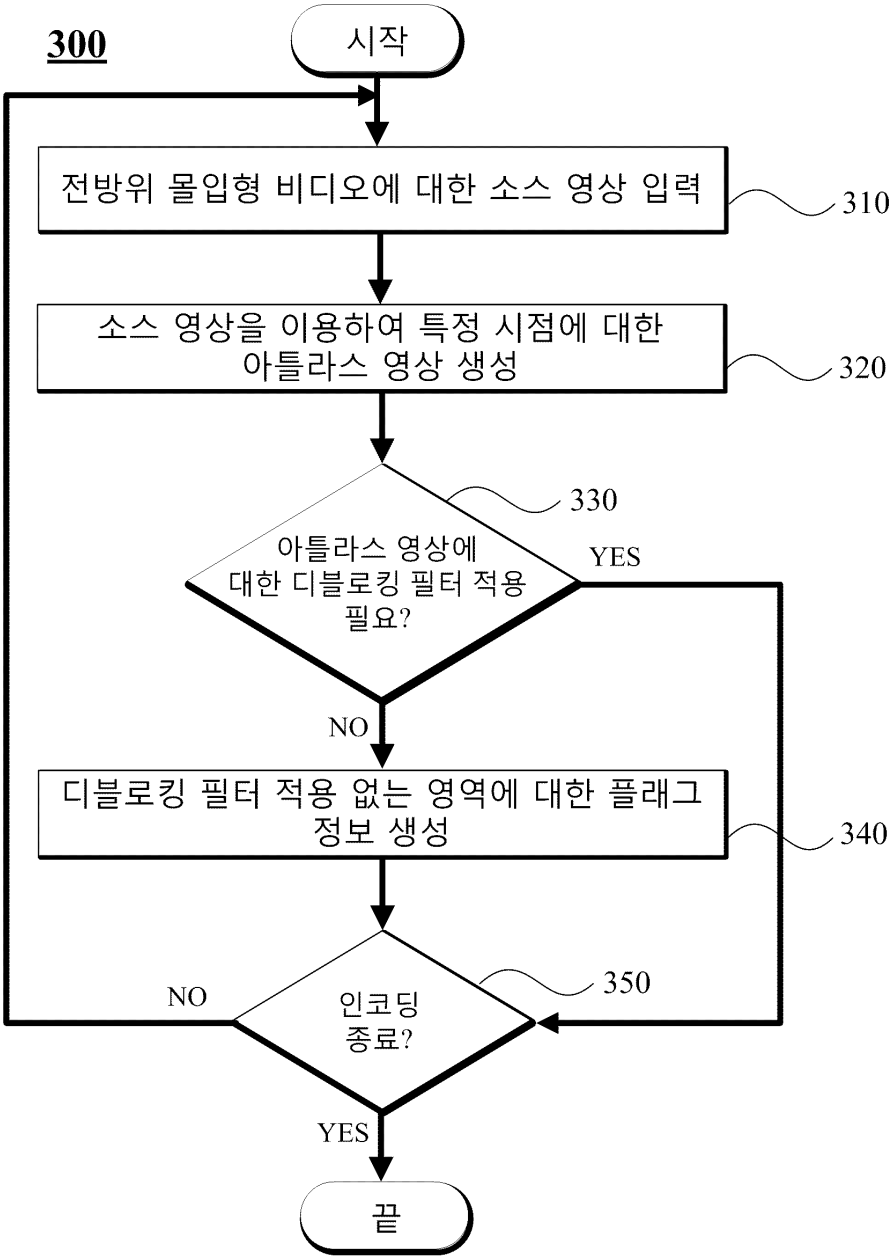


(B)

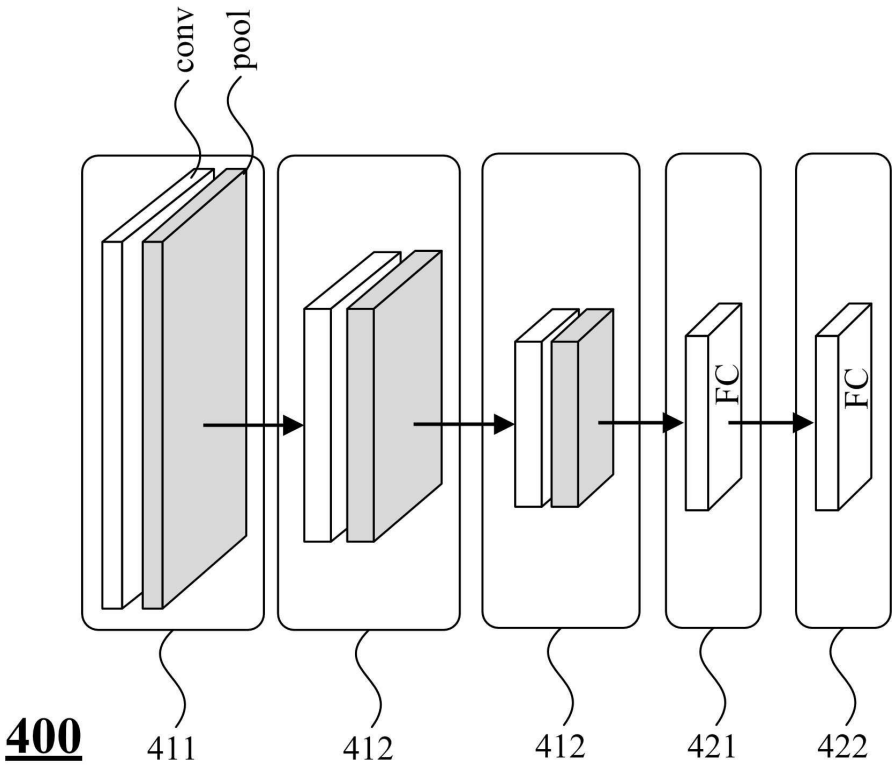
도면3



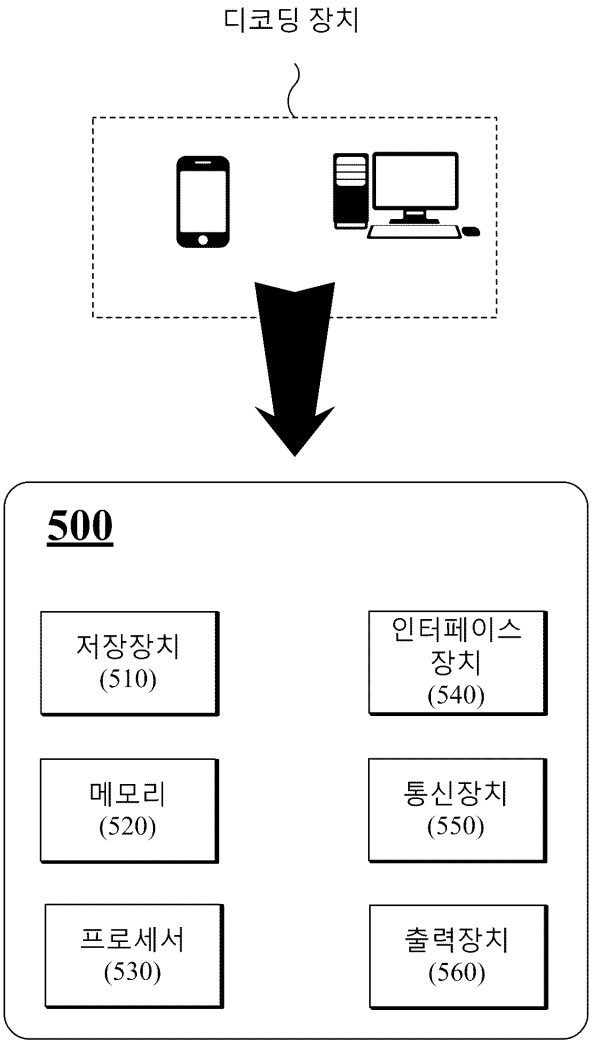
도면4


300
시작
전방위 몰입형 비디오에 대한 소스 영상 입력
310
소스 영상을 이용하여 특정 시점에 대한 아틀라스 영상 생성
320
330
아틀라스 영상에 대한 디블로킹 필터 적용 필요?
YES
NO
디블로킹 필터 적용 없는 영역에 대한 플래그 정보 생성
340
350
NO
인코딩 종료?
YES
끝

도면5



도면6


디코딩 장치
500
저장장치
(510)
인터페이스
장치
(540)
메모리
(520)
통신장치
(550)
프로세서
(530)
출력장치
(560)