



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2021년05월14일
(11) 등록번호 10-2252238
(24) 등록일자 2021년05월10일

- (51) 국제특허분류(Int. Cl.)
H04N 21/2343 (2011.01) H04L 29/06 (2006.01)
H04N 19/132 (2014.01) H04N 19/162 (2014.01)
H04N 19/167 (2014.01) H04N 19/176 (2014.01)
H04N 19/70 (2014.01) H04N 21/4728 (2011.01)
H04N 21/81 (2011.01) H04N 21/84 (2011.01)
H04N 21/845 (2011.01)
- (52) CPC특허분류
H04N 21/234345 (2013.01)
H04L 65/601 (2013.01)
- (21) 출원번호 10-2018-7032877
- (22) 출원일자(국제) 2017년05월09일
심사청구일자 2020년04월23일
- (85) 번역문제출일자 2018년11월13일
- (65) 공개번호 10-2019-0009290
- (43) 공개일자 2019년01월28일
- (86) 국제출원번호 PCT/US2017/031789
- (87) 국제공개번호 WO 2017/200804
국제공개일자 2017년11월23일
- (30) 우선권주장
62/339,009 2016년05월19일 미국(US)
15/589,782 2017년05월08일 미국(US)
- (56) 선행기술조사문헌
W02015197815 A1
(뒷면에 계속)

- (73) 특허권자
켈컴 인코포레이티드
미국 92121-1714 캘리포니아주 샌 디에고 모어하우스 드라이브 5775
- (72) 발명자
왕 예-쿠이
미국 92121-1714 캘리포니아주 샌디에고 모어하우스 드라이브 5775
헨드리 프누
미국 92121-1714 캘리포니아주 샌디에고 모어하우스 드라이브 5775
- (74) 대리인
특허법인코리어나

전체 청구항 수 : 총 44 항

심사관 : 김응권

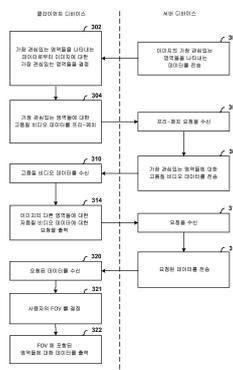
(54) 발명의 명칭 이미지에서의 가장 관심있는 영역

(57) 요약

비디오 데이터에 대한 정보를 결정하기 위한 디바이스는 비디오 데이터의 이미지의 복수의 영역들 중 하나 이상의 가장 관심있는 영역들을, 하나 이상의 가장 관심있는 영역들을 나타내는 데이터에 기초하여 결정하도록 구성되는 회로부로 구현된 하나 이상의 프로세서들을 포함한다. 하나 이상의 가장 관심있는 영역들을 나타내는

(뒷면에 계속)

대표도 - 도5



데이터는 비디오 데이터를 포함하는 비트스트림의 비디오 코딩 계층 (VCL) 데이터의 외부에 있다. 프로세서들은 이미지의 하나 이상의 가장 관심있는 영역들에 대한 상대적으로 고품질 비디오 데이터에 대한 요청을 출력하고, 상대적으로 고품질 비디오 데이터에 대한 요청을 출력한 후, 이미지의 하나 이상의 다른 영역들에 대한 상대적으로 저품질 비디오 데이터에 대한 요청을 출력한다.

(52) CPC특허분류

H04N 19/132 (2015.01)
H04N 19/162 (2015.01)
H04N 19/167 (2015.01)
H04N 19/176 (2015.01)
H04N 19/70 (2015.01)
H04N 21/4728 (2013.01)
H04N 21/8146 (2013.01)
H04N 21/84 (2013.01)
H04N 21/8456 (2013.01)

(56) 선행기술조사문헌

KR1020150070260 A
 W02014057131 A1
 KR1020120059214 A
 KR1020100026361 A
 KR1020150122781 A
 KR1020080069890 A
 KR1020140044664 A

명세서

청구범위

청구항 1

비디오 데이터에 대한 정보를 결정하는 방법으로서,

회로부로 구현된, 클라이언트 디바이스의 프로세서에 의해, 비디오 데이터의 이미지의 복수의 영역들 중 하나 이상의 가장 관심있는 영역들을, 상기 하나 이상의 가장 관심있는 영역들을 나타내는 데이터에 기초하여 결정하는 단계로서, 상기 하나 이상의 가장 관심있는 영역들을 나타내는 상기 데이터는 상기 비디오 데이터를 포함하는 비트스트림의 비디오 코딩 계층 (VCL) 데이터의 외부에 있는, 상기 하나 이상의 가장 관심있는 영역들을 결정하는 단계;

상기 클라이언트 디바이스의 상기 프로세서에 의해, 상기 이미지의 상기 하나 이상의 가장 관심있는 영역들에 대한 상대적으로 고품질 비디오 데이터에 대한 요청을 출력하는 단계;

상기 클라이언트 디바이스의 상기 프로세서에 의해, 상기 이미지의 상기 하나 이상의 가장 관심있는 영역들에 대한 상기 상대적으로 고품질 비디오 데이터를 수신하는 단계;

상기 클라이언트 디바이스의 상기 프로세서에 의해, 상기 상대적으로 고품질 비디오 데이터에 대한 요청을 출력한 후, 상기 이미지의 하나 이상의 다른 영역들에 대한 상대적으로 저품질 비디오 데이터에 대한 요청을 출력하는 단계로서, 상기 상대적으로 저품질 비디오 데이터에 대한 요청은 상기 이미지의 상기 하나 이상의 다른 영역들의 표시를 포함하는, 상기 이미지의 하나 이상의 다른 영역들에 대한 상대적으로 저품질 비디오 데이터에 대한 요청을 출력하는 단계; 및

상기 클라이언트 디바이스의 상기 프로세서에 의해, 상기 이미지의 상기 하나 이상의 다른 영역들에 대한 상기 상대적으로 저품질 비디오 데이터를 수신하는 단계를 포함하는, 비디오 데이터에 대한 정보를 결정하는 방법.

청구항 2

제 1 항에 있어서,

상기 클라이언트 디바이스의 상기 프로세서에 의해, 상기 이미지의 상기 하나 이상의 가장 관심있는 영역들에 대한 상기 상대적으로 고품질 비디오 데이터에 대한 요청을 출력한 후, 시야 (FOV) 에 기초하여 후속 이미지의 하나 이상의 영역들에 대한 상대적으로 고품질 비디오 데이터에 대한 요청을 출력하는 단계를 더 포함하는, 비디오 데이터에 대한 정보를 결정하는 방법.

청구항 3

제 1 항에 있어서,

상기 하나 이상의 가장 관심있는 영역들을 결정하는 단계는, 상기 하나 이상의 가장 관심있는 영역들을 나타내는 상기 데이터를 포함하는 보충 강화 정보 (SEI) 메시지에 기초하여 상기 하나 이상의 가장 관심있는 영역들을 결정하는 단계를 포함하는, 비디오 데이터에 대한 정보를 결정하는 방법.

청구항 4

제 1 항에 있어서,

상기 하나 이상의 가장 관심있는 영역들을 결정하는 단계는, 상기 하나 이상의 가장 관심있는 영역들을 나타내는 상기 데이터를 포함하는 샘플 그룹 박스에 기초하여 상기 하나 이상의 가장 관심있는 영역들을 결정하는 단계를 포함하고, 상기 샘플 그룹 박스는 상기 VCL 데이터를 포함하는 파일의 파일 포맷 헤더 정보에 포함되는, 비디오 데이터에 대한 정보를 결정하는 방법.

청구항 5

제 1 항에 있어서,

상기 이미지의 복수의 영역들 중 하나 이상의 가장 관심있는 영역들을 결정하는 단계는, 상기 하나 이상의 가장 관심있는 영역들을 나타내는 상기 데이터를 포함하는 타이밍된 메타데이터 트랙에 기초하여 상기 이미지의 상기 복수의 영역들 중 상기 하나 이상의 가장 관심있는 영역들을 결정하는 단계를 포함하는, 비디오 데이터에 대한 정보를 결정하는 방법.

청구항 6

제 1 항에 있어서,

상기 하나 이상의 가장 관심있는 영역들을 결정하는 단계는, 상기 하나 이상의 가장 관심있는 영역들을 나타내는 상기 데이터를 포함하는 DASH (dynamic adaptive streaming over HTTP) 미디어 프리젠테이션 디스크립션 (MPD) 에 기초하여 상기 하나 이상의 가장 관심있는 영역들을 결정하는 단계를 포함하는, 비디오 데이터에 대한 정보를 결정하는 방법.

청구항 7

제 1 항에 있어서,

상기 하나 이상의 가장 관심있는 영역들을 결정하는 단계는, 상기 하나 이상의 가장 관심있는 영역들을 나타내는 상기 데이터의 적어도 하나의 선택스 엘리먼트에 기초하여 상기 하나 이상의 가장 관심있는 영역들을 결정하는 단계를 포함하고, 상기 선택스 엘리먼트는 상기 하나 이상의 가장 관심있는 영역들의 수를 나타내는 값을 갖는, 비디오 데이터에 대한 정보를 결정하는 방법.

청구항 8

제 1 항에 있어서,

상기 하나 이상의 가장 관심있는 영역들을 결정하는 단계는, 상기 하나 이상의 가장 관심있는 영역들의 각각의 가장 관심있는 영역에 대해, 개별의 가장 관심있는 영역의 포지션 및 상기 개별의 가장 관심있는 영역의 크기를 나타내는 복수의 선택스 엘리먼트들에 대한 복수의 값들을 결정하는 단계를 포함하는, 비디오 데이터에 대한 정보를 결정하는 방법.

청구항 9

제 8 항에 있어서,

상기 하나 이상의 가장 관심있는 영역들의 각각의 가장 관심있는 영역에 대해, 개별의 가장 관심있는 영역의 포지션 및 상기 개별의 가장 관심있는 영역의 크기를 나타내는 복수의 선택스 엘리먼트들에 대한 복수의 값들을 결정하는 단계는, 상기 하나 이상의 가장 관심있는 영역들의 각각의 가장 관심있는 영역에 대해, 상기 개별의 가장 관심있는 영역의 상기 포지션 및 상기 개별의 가장 관심있는 영역의 상기 크기를 나타내는 상기 개별의 가장 관심있는 영역에 대한 루프를 결정하는 단계를 포함하는, 비디오 데이터에 대한 정보를 결정하는 방법.

청구항 10

제 8 항에 있어서,

상기 하나 이상의 가장 관심있는 영역들의 각각의 가장 관심있는 영역에 대해, 개별의 가장 관심있는 영역의 포지션 및 상기 개별의 가장 관심있는 영역의 크기를 나타내는 복수의 선택스 엘리먼트들에 대한 복수의 값들을 결정하는 단계는, 상기 하나 이상의 가장 관심있는 영역들의 각각의 가장 관심있는 영역에 대해, 상기 개별의 가장 관심있는 영역의 상부를 나타내는 제 1 선택스 엘리먼트, 상기 개별의 가장 관심있는 영역의 하부를 나타내는 제 2 선택스 엘리먼트, 상기 개별의 가장 관심있는 영역의 좌측을 나타내는 제 3 선택스 엘리먼트, 및 상기 개별의 가장 관심있는 영역의 우측을 나타내는 제 4 선택스 엘리먼트를 결정하는 단계를 포함하는, 비디오 데이터에 대한 정보를 결정하는 방법.

청구항 11

제 8 항에 있어서,

상기 하나 이상의 가장 관심있는 영역들의 각각의 가장 관심있는 영역에 대해, 개별의 가장 관심있는 영역의 포

지션 및 상기 개별의 가장 관심있는 영역의 사이즈를 나타내는 복수의 신택스 엘리먼트들에 대한 복수의 값들을 결정하는 단계는, 상기 하나 이상의 가장 관심있는 영역들의 각각의 가장 관심있는 영역에 대해, 상기 개별의 가장 관심있는 영역의 좌상측 (upper-left) 코너를 나타내는 제 1 신택스 엘리먼트 및 상기 개별의 가장 관심있는 영역의 우하측 (lower-right) 코너를 나타내는 제 2 신택스 엘리먼트를 결정하는 단계를 포함하는, 비디오 데이터에 대한 정보를 결정하는 방법.

청구항 12

제 8 항에 있어서,

상기 하나 이상의 가장 관심있는 영역들의 각각의 가장 관심있는 영역에 대해, 개별의 가장 관심있는 영역의 포지션 및 상기 개별의 가장 관심있는 영역의 사이즈를 나타내는 복수의 신택스 엘리먼트들에 대한 복수의 값들을 결정하는 단계는, 상기 하나 이상의 가장 관심있는 영역들의 각각의 가장 관심있는 영역에 대해, 상기 개별의 가장 관심있는 영역의 좌상측 코너를 나타내는 제 1 신택스 엘리먼트, 상기 개별의 가장 관심있는 영역의 폭을 나타내는 제 2 신택스 엘리먼트, 및 상기 개별의 가장 관심있는 영역의 높이를 나타내는 제 3 신택스 엘리먼트를 결정하는 단계를 포함하는, 비디오 데이터에 대한 정보를 결정하는 방법.

청구항 13

비디오 데이터에 대한 정보를 결정하기 위한 디바이스로서,

상기 디바이스는 회로부로 구현된 하나 이상의 프로세서들을 포함하고, 상기 하나 이상의 프로세서들은,

비디오 데이터의 이미지의 복수의 영역들 중 하나 이상의 가장 관심있는 영역들을, 상기 하나 이상의 가장 관심있는 영역들을 나타내는 데이터에 기초하여 결정하는 것으로서, 상기 하나 이상의 가장 관심있는 영역들을 나타내는 상기 데이터는 상기 비디오 데이터를 포함하는 비트스트림의 비디오 코딩 계층 (VCL) 데이터의 외부에 있는, 상기 하나 이상의 가장 관심있는 영역들을 결정하고;

상기 이미지의 상기 하나 이상의 가장 관심있는 영역들에 대한 상대적으로 고품질 비디오 데이터에 대한 요청을 출력하고;

상기 이미지의 상기 하나 이상의 가장 관심있는 영역들에 대한 상기 상대적으로 고품질 비디오 데이터를 수신하고;

상기 상대적으로 고품질 비디오 데이터에 대한 요청을 출력한 후, 상기 이미지의 하나 이상의 다른 영역들에 대한 상대적으로 저품질 비디오 데이터에 대한 요청을 출력하는 것으로서, 상기 상대적으로 저품질 비디오 데이터에 대한 요청은 상기 이미지의 상기 하나 이상의 다른 영역들의 표시를 포함하는, 상기 이미지의 하나 이상의 다른 영역들에 대한 상대적으로 저품질 비디오 데이터에 대한 요청을 출력하고; 그리고

상기 이미지의 제 2 부분을 형성하는 상기 이미지의 상기 하나 이상의 다른 영역들에 대한 상기 상대적으로 저품질 비디오 데이터를 수신하도록 구성되는, 비디오 데이터에 대한 정보를 결정하기 위한 디바이스.

청구항 14

제 13 항에 있어서,

상기 하나 이상의 가장 관심있는 영역들을 결정하기 위해, 상기 하나 이상의 프로세서들은,

상기 하나 이상의 가장 관심있는 영역들을 나타내는 상기 데이터를 포함하는 샘플 그룹 박스에 기초하여 상기 하나 이상의 가장 관심있는 영역들을 결정하도록 추가로 구성되고, 상기 샘플 그룹 박스는 상기 VCL 데이터를 포함하는 파일의 파일 포맷 헤더 정보에 포함되는, 비디오 데이터에 대한 정보를 결정하기 위한 디바이스.

청구항 15

제 13 항에 있어서,

상기 하나 이상의 가장 관심있는 영역들을 결정하기 위해, 상기 하나 이상의 프로세서들은,

상기 하나 이상의 가장 관심있는 영역들을 나타내는 상기 데이터를 포함하는 타이밍된 메타데이터 트랙에 기초하여 상기 하나 이상의 가장 관심있는 영역들을 결정하도록 추가로 구성되는, 비디오 데이터에 대한 정보를 결

정하기 위한 디바이스.

청구항 16

제 13 항에 있어서,

상기 하나 이상의 가장 관심있는 영역들을 결정하기 위해, 상기 하나 이상의 프로세서들은,

상기 하나 이상의 가장 관심있는 영역들을 나타내는 상기 데이터의 적어도 하나의 선택스 엘리먼트에 기초하여 상기 하나 이상의 가장 관심있는 영역들을 결정하도록 추가로 구성되고, 상기 선택스 엘리먼트는 상기 하나 이상의 가장 관심있는 영역들의 수를 나타내는 값을 갖는, 비디오 데이터에 대한 정보를 결정하기 위한 디바이스.

청구항 17

제 13 항에 있어서,

상기 하나 이상의 가장 관심있는 영역들을 결정하기 위해, 상기 하나 이상의 프로세서들은,

상기 하나 이상의 가장 관심있는 영역들의 각각의 가장 관심있는 영역에 대해, 개별의 가장 관심있는 영역의 포지션 및 상기 개별의 가장 관심있는 영역의 사이즈를 나타내는 복수의 선택스 엘리먼트들에 대한 복수의 값들을 결정하도록 추가로 구성되는, 비디오 데이터에 대한 정보를 결정하기 위한 디바이스.

청구항 18

제 17 항에 있어서,

상기 하나 이상의 가장 관심있는 영역들의 각각의 가장 관심있는 영역에 대해, 개별의 가장 관심있는 영역의 포지션 및 상기 개별의 가장 관심있는 영역의 사이즈를 나타내는 복수의 선택스 엘리먼트들에 대한 복수의 값들을 결정하기 위해, 상기 하나 이상의 프로세서들은,

상기 하나 이상의 가장 관심있는 영역들의 각각의 가장 관심있는 영역에 대해, 상기 개별의 가장 관심있는 영역의 상기 포지션 및 상기 개별의 가장 관심있는 영역의 상기 사이즈를 나타내는 상기 개별의 가장 관심있는 영역에 대한 루프를 결정하도록 추가로 구성되는, 비디오 데이터에 대한 정보를 결정하기 위한 디바이스.

청구항 19

제 17 항에 있어서,

상기 하나 이상의 가장 관심있는 영역들의 각각의 가장 관심있는 영역에 대해, 개별의 가장 관심있는 영역의 포지션 및 상기 개별의 가장 관심있는 영역의 사이즈를 나타내는 복수의 선택스 엘리먼트들에 대한 복수의 값들을 결정하기 위해, 상기 하나 이상의 프로세서들은,

상기 하나 이상의 가장 관심있는 영역들의 각각의 가장 관심있는 영역에 대해, 상기 개별의 가장 관심있는 영역의 상부를 나타내는 제 1 선택스 엘리먼트, 상기 개별의 가장 관심있는 영역의 하부를 나타내는 제 2 선택스 엘리먼트, 상기 개별의 가장 관심있는 영역의 좌측을 나타내는 제 3 선택스 엘리먼트, 및 상기 개별의 가장 관심있는 영역의 우측을 나타내는 제 4 선택스 엘리먼트를 결정하도록 추가로 구성되는, 비디오 데이터에 대한 정보를 결정하기 위한 디바이스.

청구항 20

제 17 항에 있어서,

상기 하나 이상의 가장 관심있는 영역들의 각각의 가장 관심있는 영역에 대해, 개별의 가장 관심있는 영역의 포지션 및 상기 개별의 가장 관심있는 영역의 사이즈를 나타내는 복수의 선택스 엘리먼트들에 대한 복수의 값들을 결정하기 위해, 상기 하나 이상의 프로세서들은,

상기 하나 이상의 가장 관심있는 영역들의 각각의 가장 관심있는 영역에 대해, 상기 개별의 가장 관심있는 영역의 좌상측 코너를 나타내는 제 1 선택스 엘리먼트, 상기 개별의 가장 관심있는 영역의 폭을 나타내는 제 2 선택스 엘리먼트, 및 상기 개별의 가장 관심있는 영역의 높이를 나타내는 제 3 선택스 엘리먼트를 결정하도록 추가로 구성되는, 비디오 데이터에 대한 정보를 결정하기 위한 디바이스.

청구항 21

비디오 데이터에 대한 정보를 결정하기 위한 디바이스로서,

비디오 데이터의 이미지의 복수의 영역들 중 하나 이상의 가장 관심있는 영역들을, 상기 하나 이상의 가장 관심있는 영역들을 나타내는 데이터에 기초하여 결정하기 위한 수단으로서, 상기 하나 이상의 가장 관심있는 영역들을 나타내는 상기 데이터는 상기 비디오 데이터를 포함하는 비트스트림의 비디오 코딩 계층 (VCL) 데이터의 외부에 있는, 상기 하나 이상의 가장 관심있는 영역들을 결정하기 위한 수단;

상기 이미지의 상기 하나 이상의 가장 관심있는 영역들에 대한 상대적으로 고품질 비디오 데이터에 대한 요청을 출력하기 위한 수단;

상기 이미지의 상기 하나 이상의 가장 관심있는 영역들에 대한 상기 상대적으로 고품질 비디오 데이터를 수신하기 위한 수단;

상기 상대적으로 고품질 비디오 데이터에 대한 요청을 출력한 후, 상기 이미지의 하나 이상의 다른 영역들에 대한 상대적으로 저품질 비디오 데이터에 대한 요청을 출력하기 위한 수단으로서, 상기 상대적으로 저품질 비디오 데이터에 대한 요청은 상기 이미지의 상기 하나 이상이 다른 영역들의 표시를 포함하는, 상기 이미지의 하나 이상의 다른 영역들에 대한 상대적으로 저품질 비디오 데이터에 대한 요청을 출력하기 위한 수단; 및

상기 이미지의 제 2 부분을 형성하는 상기 이미지의 상기 하나 이상의 다른 영역들에 대한 상기 상대적으로 저품질 비디오 데이터를 수신하기 위한 수단을 포함하는, 비디오 데이터에 대한 정보를 결정하기 위한 디바이스.

청구항 22

명령들을 저장한 컴퓨터 판독가능 저장 매체로서,

상기 명령들은, 실행될 때, 프로세서로 하여금,

비디오 데이터의 이미지의 복수의 영역들 중 하나 이상의 가장 관심있는 영역들을, 상기 하나 이상의 가장 관심있는 영역들을 나타내는 데이터에 기초하여 결정하게 하는 것으로서, 상기 하나 이상의 가장 관심있는 영역들을 나타내는 상기 데이터는 상기 비디오 데이터를 포함하는 비트스트림의 비디오 코딩 계층 (VCL) 데이터의 외부에 있는, 상기 하나 이상의 가장 관심있는 영역들을 결정하게 하고;

상기 이미지의 상기 하나 이상의 가장 관심있는 영역들에 대한 상대적으로 고품질 비디오 데이터에 대한 요청을 출력하게 하고;

상기 이미지의 상기 하나 이상의 가장 관심있는 영역들에 대한 상기 상대적으로 고품질 비디오 데이터를 수신하게 하고;

상기 상대적으로 고품질 비디오 데이터에 대한 요청을 출력한 후, 상기 이미지의 하나 이상의 다른 영역들에 대한 상대적으로 저품질 비디오 데이터에 대한 요청을 출력하게 하는 것으로서, 상기 상대적으로 저품질 비디오 데이터에 대한 요청은 상기 이미지의 상기 하나 이상의 다른 영역들의 표시를 포함하는, 상기 이미지의 하나 이상의 다른 영역들에 대한 상대적으로 저품질 비디오 데이터에 대한 요청을 출력하게 하고; 그리고

상기 이미지의 제 2 부분을 형성하는 상기 이미지의 상기 하나 이상의 다른 영역들에 대한 상기 상대적으로 저품질 비디오 데이터를 수신하게 하는, 컴퓨터 판독가능 저장 매체.

청구항 23

비디오 데이터에 대한 정보를 시그널링하는 방법으로서,

회로부로 구현된, 소스 디바이스의 프로세서에 의해, 비디오 데이터의 이미지의 복수의 영역들 중 하나 이상의 가장 관심있는 영역들을 결정하는 단계로서, 상기 하나 이상의 가장 관심있는 영역들은 클라이언트 디바이스에 의해 추출될 가능성이 가장 큰 하나 이상의 영역들을 포함하는, 상기 하나 이상의 가장 관심있는 영역들을 결정하는 단계;

상기 소스 디바이스의 상기 프로세서에 의해, 상기 하나 이상의 가장 관심있는 영역들을 나타내는 데이터를 생성하는 단계로서, 상기 하나 이상의 가장 관심있는 영역들을 나타내는 상기 데이터는 상기 비디오 데이터를 포함하는 비트스트림의 비디오 코딩 계층 (VCL) 데이터의 외부에 있는, 상기 하나 이상의 가장 관심있는 영역들을

나타내는 데이터를 생성하는 단계;

상기 소스 디바이스의 상기 프로세서에 의해, 상기 이미지의 상기 하나 이상의 가장 관심있는 영역들에 대한 상대적으로 고품질 비디오 데이터에 대한 요청에 응답하여, 상기 이미지의 상기 하나 이상의 가장 관심있는 영역들에 대한 상기 상대적으로 고품질 비디오 데이터를 출력하는 단계; 및

상기 소스 디바이스의 상기 프로세서에 의해, 상기 상대적으로 고품질 비디오 데이터에 대한 요청 후에 수신되는 상기 이미지의 하나 이상의 다른 영역들에 대한 상대적으로 저품질 비디오 데이터에 대한 요청에 응답하여, 상기 이미지의 상기 하나 이상의 다른 영역들에 대한 상기 상대적으로 저품질 비디오 데이터를 출력하는 단계로서, 상기 상대적으로 저품질 비디오 데이터에 대한 요청은 상기 이미지의 상기 하나 이상의 다른 영역들의 표시를 포함하는, 상기 이미지의 상기 하나 이상의 다른 영역들에 대한 상기 상대적으로 저품질 비디오 데이터를 출력하는 단계를 포함하는, 비디오 데이터에 대한 정보를 시그널링하는 방법.

청구항 24

제 23 항에 있어서,

상기 하나 이상의 가장 관심있는 영역들을 나타내는 데이터를 생성하는 단계는, 상기 하나 이상의 가장 관심있는 영역들을 나타내는 상기 데이터를 포함하는 보충 강화 정보 (SEI) 메시지를 생성하는 단계를 포함하는, 비디오 데이터에 대한 정보를 시그널링하는 방법.

청구항 25

제 23 항에 있어서,

상기 하나 이상의 가장 관심있는 영역들을 나타내는 데이터를 생성하는 단계는, 상기 하나 이상의 가장 관심있는 영역들을 나타내는 상기 데이터를 포함하는 샘플 그룹 박스를 생성하는 단계를 포함하고, 상기 샘플 그룹 박스는 상기 VCL 데이터를 포함하는 파일의 파일 포맷 헤더 정보에 포함되는, 비디오 데이터에 대한 정보를 시그널링하는 방법.

청구항 26

제 23 항에 있어서,

상기 하나 이상의 가장 관심있는 영역들을 나타내는 데이터를 생성하는 단계는, 상기 하나 이상의 가장 관심있는 영역들을 나타내는 상기 데이터를 포함하는 타이밍된 메타데이터 트랙을 생성하는 단계를 포함하는, 비디오 데이터에 대한 정보를 시그널링하는 방법.

청구항 27

제 23 항에 있어서,

상기 하나 이상의 가장 관심있는 영역들을 나타내는 데이터를 생성하는 단계는, 상기 하나 이상의 가장 관심있는 영역들을 나타내는 상기 데이터를 포함하는 DASH (dynamic adaptive streaming over HTTP) 미디어 프리젠테이션 디스크립션 (MPD) 을 생성하는 단계를 포함하는, 비디오 데이터에 대한 정보를 시그널링하는 방법.

청구항 28

제 23 항에 있어서,

상기 하나 이상의 가장 관심있는 영역들을 나타내는 데이터를 생성하는 단계는, 상기 하나 이상의 가장 관심있는 영역들을 나타내는 상기 데이터의 적어도 하나의 선택스 엘리먼트를 생성하는 단계를 포함하고, 상기 선택스 엘리먼트는 상기 하나 이상의 가장 관심있는 영역들의 수를 나타내는 값을 갖는, 비디오 데이터에 대한 정보를 시그널링하는 방법.

청구항 29

제 23 항에 있어서,

상기 하나 이상의 가장 관심있는 영역들을 나타내는 데이터를 생성하는 단계는, 상기 하나 이상의 가장 관심있는 영역들의 각각의 가장 관심있는 영역에 대해, 개별의 가장 관심있는 영역의 포지션 및 상기 개별의 가장 관

심있는 영역의 사이즈를 나타내는 복수의 선택스 엘리먼트들에 대한 복수의 값들을 생성하는 단계를 포함하는, 비디오 데이터에 대한 정보를 시그널링하는 방법.

청구항 30

제 29 항에 있어서,

상기 하나 이상의 가장 관심있는 영역들의 각각의 가장 관심있는 영역에 대해, 개별의 가장 관심있는 영역의 포지션 및 상기 개별의 가장 관심있는 영역의 사이즈를 나타내는 복수의 선택스 엘리먼트들에 대한 복수의 값들을 생성하는 단계는, 상기 하나 이상의 가장 관심있는 영역들의 각각의 가장 관심있는 영역에 대해, 상기 개별의 가장 관심있는 영역의 상기 포지션 및 상기 개별의 가장 관심있는 영역의 상기 사이즈를 나타내는 상기 개별의 가장 관심있는 영역에 대한 루프를 생성하는 단계를 포함하는, 비디오 데이터에 대한 정보를 시그널링하는 방법.

청구항 31

제 29 항에 있어서,

상기 하나 이상의 가장 관심있는 영역들의 각각의 가장 관심있는 영역에 대해, 개별의 가장 관심있는 영역의 포지션 및 상기 개별의 가장 관심있는 영역의 사이즈를 나타내는 복수의 선택스 엘리먼트들에 대한 복수의 값들을 생성하는 단계는, 상기 하나 이상의 가장 관심있는 영역들의 각각의 가장 관심있는 영역에 대해, 상기 개별의 가장 관심있는 영역의 상부를 나타내는 제 1 선택스 엘리먼트, 상기 개별의 가장 관심있는 영역의 하부를 나타내는 제 2 선택스 엘리먼트, 상기 개별의 가장 관심있는 영역의 좌측을 나타내는 제 3 선택스 엘리먼트, 및 상기 개별의 가장 관심있는 영역의 우측을 나타내는 제 4 선택스 엘리먼트를 생성하는 단계를 포함하는, 비디오 데이터에 대한 정보를 시그널링하는 방법.

청구항 32

제 29 항에 있어서,

상기 하나 이상의 가장 관심있는 영역들의 각각의 가장 관심있는 영역에 대해, 개별의 가장 관심있는 영역의 포지션 및 상기 개별의 가장 관심있는 영역의 사이즈를 나타내는 복수의 선택스 엘리먼트들에 대한 복수의 값들을 생성하는 단계는, 상기 하나 이상의 가장 관심있는 영역들의 각각의 가장 관심있는 영역에 대해, 상기 개별의 가장 관심있는 영역의 좌상측 코너를 나타내는 제 1 선택스 엘리먼트 및 상기 개별의 가장 관심있는 영역의 우하측 코너를 나타내는 제 2 선택스 엘리먼트를 생성하는 단계를 포함하는, 비디오 데이터에 대한 정보를 시그널링하는 방법.

청구항 33

제 29 항에 있어서,

상기 하나 이상의 가장 관심있는 영역들의 각각의 가장 관심있는 영역에 대해, 개별의 가장 관심있는 영역의 포지션 및 상기 개별의 가장 관심있는 영역의 사이즈를 나타내는 복수의 선택스 엘리먼트들에 대한 복수의 값들을 생성하는 단계는, 상기 하나 이상의 가장 관심있는 영역들의 각각의 가장 관심있는 영역에 대해, 상기 개별의 가장 관심있는 영역의 좌상측 코너를 나타내는 제 1 선택스 엘리먼트, 상기 개별의 가장 관심있는 영역의 폭을 나타내는 제 2 선택스 엘리먼트, 및 상기 개별의 가장 관심있는 영역의 높이를 나타내는 제 3 선택스 엘리먼트를 생성하는 단계를 포함하는, 비디오 데이터에 대한 정보를 시그널링하는 방법.

청구항 34

제 23 항에 있어서,

상기 하나 이상의 가장 관심있는 영역들을 결정하는 단계는 디렉터의 의도 또는 사용자 통계 중 하나 이상을 이용하여 상기 가장 관심있는 영역들을 결정하는 단계를 포함하는, 비디오 데이터에 대한 정보를 시그널링하는 방법.

청구항 35

비디오 데이터에 대한 정보를 시그널링하기 위한 디바이스로서,

상기 디바이스는 회로부로 구현된 하나 이상의 프로세서들을 포함하고, 상기 하나 이상의 프로세서들은,

비디오 데이터의 이미지의 복수의 영역들 중 하나 이상의 가장 관심있는 영역들을 결정하는 것으로서, 상기 하나 이상의 가장 관심있는 영역들은 클라이언트 디바이스에 의해 추출될 가능성이 가장 큰 하나 이상의 영역들을 포함하는, 상기 하나 이상의 가장 관심있는 영역들을 결정하고;

상기 하나 이상의 가장 관심있는 영역들을 나타내는 데이터를 생성하는 것으로서, 상기 하나 이상의 가장 관심있는 영역들을 나타내는 상기 데이터는 상기 비디오 데이터를 포함하는 비트스트림의 비디오 코딩 계층 (VCL) 데이터의 외부에 있는, 상기 하나 이상의 가장 관심있는 영역들을 나타내는 데이터를 생성하고;

상기 이미지의 상기 하나 이상의 가장 관심있는 영역들에 대한 상대적으로 고품질 비디오 데이터에 대한 요청에 응답하여, 상기 이미지의 상기 하나 이상의 가장 관심있는 영역들에 대한 상기 상대적으로 고품질 비디오 데이터를 출력하고; 그리고

상기 상대적으로 고품질 비디오 데이터에 대한 요청 후에 수신되는 상기 이미지의 하나 이상의 다른 영역들에 대한 상대적으로 저품질 비디오 데이터에 대한 요청에 응답하여, 상기 이미지의 상기 하나 이상의 다른 영역들에 대한 상기 상대적으로 저품질 비디오 데이터를 출력하는 것으로서, 상기 상대적으로 저품질 비디오 데이터에 대한 요청은 상기 이미지의 상기 하나 이상의 다른 영역들의 표시를 포함하는, 상기 이미지의 상기 하나 이상의 다른 영역들에 대한 상기 상대적으로 저품질 비디오 데이터를 출력하도록 구성되는, 비디오 데이터에 대한 정보를 시그널링하기 위한 디바이스.

청구항 36

제 35 항에 있어서,

상기 하나 이상의 가장 관심있는 영역들을 나타내는 데이터를 생성하기 위해, 상기 하나 이상의 프로세서들은, 상기 하나 이상의 가장 관심있는 영역들을 나타내는 상기 데이터를 포함하는 샘플 그룹 박스를 생성하도록 추가로 구성되고, 상기 샘플 그룹 박스는 상기 VCL 데이터를 포함하는 파일의 파일 포맷 헤더 정보에 포함되는, 비디오 데이터에 대한 정보를 시그널링하기 위한 디바이스.

청구항 37

제 35 항에 있어서,

상기 하나 이상의 가장 관심있는 영역들을 나타내는 데이터를 생성하기 위해, 상기 하나 이상의 프로세서들은, 상기 하나 이상의 가장 관심있는 영역들을 나타내는 상기 데이터를 포함하는 타이밍된 메타데이터 트랙을 생성하도록 추가로 구성되는, 비디오 데이터에 대한 정보를 시그널링하기 위한 디바이스.

청구항 38

제 35 항에 있어서,

상기 하나 이상의 가장 관심있는 영역들을 나타내는 데이터를 생성하기 위해, 상기 하나 이상의 프로세서들은, 상기 하나 이상의 가장 관심있는 영역들의 수를 나타내는 값을 생성하도록 추가로 구성되는, 비디오 데이터에 대한 정보를 시그널링하기 위한 디바이스.

청구항 39

제 35 항에 있어서,

상기 하나 이상의 가장 관심있는 영역들을 나타내는 데이터를 생성하기 위해, 상기 하나 이상의 프로세서들은, 상기 하나 이상의 가장 관심있는 영역들의 각각의 가장 관심있는 영역에 대해, 개별의 가장 관심있는 영역의 포지션 및 상기 개별의 가장 관심있는 영역의 사이즈를 나타내는 복수의 신택스 엘리먼트들에 대한 복수의 값들을 생성하도록 추가로 구성되는, 비디오 데이터에 대한 정보를 시그널링하기 위한 디바이스.

청구항 40

제 39 항에 있어서,

상기 하나 이상의 가장 관심있는 영역들의 각각의 가장 관심있는 영역에 대해, 개별의 가장 관심있는 영역의 포지션 및 상기 개별의 가장 관심있는 영역의 사이즈를 나타내는 복수의 신택스 엘리먼트들에 대한 복수의 값들을 생성하기 위해, 상기 하나 이상의 프로세서들은,

상기 하나 이상의 가장 관심있는 영역들의 각각의 가장 관심있는 영역에 대해, 상기 개별의 가장 관심있는 영역의 상기 포지션 및 상기 개별의 가장 관심있는 영역의 상기 사이즈를 나타내는 상기 개별의 가장 관심있는 영역에 대한 루프를 생성하도록 추가로 구성되는, 비디오 데이터에 대한 정보를 시그널링하기 위한 디바이스.

청구항 41

제 39 항에 있어서,

상기 하나 이상의 가장 관심있는 영역들의 각각의 가장 관심있는 영역에 대해, 개별의 가장 관심있는 영역의 포지션 및 상기 개별의 가장 관심있는 영역의 사이즈를 나타내는 복수의 신택스 엘리먼트들에 대한 복수의 값들을 생성하기 위해, 상기 하나 이상의 프로세서들은,

상기 하나 이상의 가장 관심있는 영역들의 각각의 가장 관심있는 영역에 대해, 상기 개별의 가장 관심있는 영역의 상부를 나타내는 제 1 신택스 엘리먼트, 상기 개별의 가장 관심있는 영역의 하부를 나타내는 제 2 신택스 엘리먼트, 상기 개별의 가장 관심있는 영역의 좌측을 나타내는 제 3 신택스 엘리먼트, 및 상기 개별의 가장 관심있는 영역의 우측을 나타내는 제 4 신택스 엘리먼트를 생성하도록 추가로 구성되는, 비디오 데이터에 대한 정보를 시그널링하기 위한 디바이스.

청구항 42

제 39 항에 있어서,

상기 하나 이상의 가장 관심있는 영역들의 각각의 가장 관심있는 영역에 대해, 개별의 가장 관심있는 영역의 포지션 및 상기 개별의 가장 관심있는 영역의 사이즈를 나타내는 복수의 신택스 엘리먼트들에 대한 복수의 값들을 생성하기 위해, 상기 하나 이상의 프로세서들은,

상기 하나 이상의 가장 관심있는 영역들의 각각의 가장 관심있는 영역에 대해, 상기 개별의 가장 관심있는 영역의 좌상측 코너를 나타내는 제 1 신택스 엘리먼트, 상기 개별의 가장 관심있는 영역의 폭을 나타내는 제 2 신택스 엘리먼트, 및 상기 개별의 가장 관심있는 영역의 높이를 나타내는 제 3 신택스 엘리먼트를 생성하도록 추가로 구성되는, 비디오 데이터에 대한 정보를 시그널링하기 위한 디바이스.

청구항 43

비디오 데이터에 대한 정보를 시그널링하기 위한 디바이스로서,

비디오 데이터의 이미지의 복수의 영역들 중 하나 이상의 가장 관심있는 영역들을 결정하기 위한 수단으로서, 상기 하나 이상의 가장 관심있는 영역들은 클라이언트 디바이스에 의해 추출될 가능성이 가장 큰 하나 이상의 영역들을 포함하는, 상기 하나 이상의 가장 관심있는 영역들을 결정하기 위한 수단; 및

상기 하나 이상의 가장 관심있는 영역들을 나타내는 데이터를 생성하기 위한 수단으로서, 상기 하나 이상의 가장 관심있는 영역들을 나타내는 상기 데이터는 상기 비디오 데이터를 포함하는 비트스트림의 비디오 코딩 계층(VCL) 데이터의 외부에 있는, 상기 하나 이상의 가장 관심있는 영역들을 나타내는 데이터를 생성하기 위한 수단;

상기 이미지의 상기 하나 이상의 가장 관심있는 영역들에 대한 상대적으로 고품질 비디오 데이터에 대한 요청에 응답하여, 상기 이미지의 상기 하나 이상의 가장 관심있는 영역들에 대한 상기 상대적으로 고품질 비디오 데이터를 출력하기 위한 수단; 및

상기 상대적으로 고품질 비디오 데이터에 대한 요청 후에 수신되는 상기 이미지의 하나 이상의 다른 영역들에 대한 상대적으로 저품질 비디오 데이터에 대한 요청에 응답하여, 상기 이미지의 상기 하나 이상의 다른 영역들에 대한 상기 상대적으로 저품질 비디오 데이터를 출력하기 위한 수단으로서, 상기 상대적으로 저품질 비디오 데이터에 대한 요청은 상기 이미지의 상기 하나 이상의 다른 영역들의 표시를 포함하는, 상기 이미지의 상기 하나 이상의 다른 영역들에 대한 상기 상대적으로 저품질 비디오 데이터를 출력하기 위한 수단을 포함하는, 비디오 데이터에 대한 정보를 시그널링하기 위한 디바이스.

청구항 44

명령들을 저장한 컴퓨터 판독가능 저장 매체로서,

상기 명령들은, 실행될 때, 프로세서로 하여금,

비디오 데이터의 이미지의 복수의 영역들 중 하나 이상의 가장 관심있는 영역들을 결정하게 하는 것으로서, 상기 하나 이상의 가장 관심있는 영역들은 클라이언트 디바이스에 의해 추출될 가능성이 가장 큰 하나 이상의 영역들을 포함하는, 상기 하나 이상의 가장 관심있는 영역들을 결정하게 하고;

상기 하나 이상의 가장 관심있는 영역들을 나타내는 데이터를 생성하게 하는 것으로서, 상기 하나 이상의 가장 관심있는 영역들을 나타내는 상기 데이터는 상기 비디오 데이터를 포함하는 비트스트림의 비디오 코딩 계층 (VCL) 데이터의 외부에 있는, 상기 하나 이상의 가장 관심있는 영역들을 나타내는 데이터를 생성하게 하고;

상기 이미지의 상기 하나 이상의 가장 관심있는 영역들에 대한 상대적으로 고품질 비디오 데이터에 대한 요청에 응답하여, 상기 이미지의 상기 하나 이상의 가장 관심있는 영역들에 대한 상기 상대적으로 고품질 비디오 데이터를 출력하게 하고; 그리고

상기 상대적으로 고품질 비디오 데이터에 대한 요청 후에 수신되는 상기 이미지의 하나 이상의 다른 영역들에 대한 상대적으로 저품질 비디오 데이터에 대한 요청에 응답하여, 상기 이미지의 상기 하나 이상의 다른 영역들에 대한 상기 상대적으로 저품질 비디오 데이터를 출력하게 하는 것으로서, 상기 상대적으로 저품질 비디오 데이터에 대한 요청은 상기 이미지의 상기 하나 이상의 다른 영역들의 표시를 포함하는, 상기 이미지의 상기 하나 이상의 다른 영역들에 대한 상기 상대적으로 저품질 비디오 데이터를 출력하게 하는, 컴퓨터 판독가능 저장 매체.

발명의 설명

기술 분야

[0001] 본 출원은 2016년 5월 19일자로 출원된 미국 가출원 제62/339,009호의 이익을 주장하고, 이는 이로써 전부 참조로 통합된다.

[0002] 본 개시는 인코딩된 비디오 데이터의 저장 및 전송에 관한 것이다.

배경 기술

[0003] 디지털 비디오 능력들은, 디지털 텔레비전들, 디지털 다이렉트 브로드캐스트 시스템들, 무선 브로드캐스트 시스템들, PDA (personal digital assistant) 들, 랩톱 또는 데스크톱 컴퓨터들, 디지털 카메라들, 디지털 레코딩 디바이스들, 디지털 미디어 플레이어들, 비디오 게이밍 디바이스들, 비디오 게임 콘솔들, 셀룰러 또는 위성 무선 전화기들, 비디오 텔레컨퍼런싱 디바이스들 등을 포함하는, 광범위한 디바이스들에 통합될 수 있다. 디지털 비디오 디바이스들은 MPEG-2, MPEG-4, ITU-T H.263 또는 ITU-T H.264/MPEG-4, Part 10, AVC (Advanced Video Coding), ITU-T H.265 (고효율 비디오 코딩 (High Efficiency Video Coding; HEVC) 으로도 또한 지칭됨)에 의해 정의된 표준들, 및 이러한 표준들의 확장들에서 설명된 것들과 같은 비디오 압축 기법들을 구현하여, 디지털 비디오 정보를 보다 효율적으로 송신 및 수신한다.

[0004] 비디오 압축 기법들은 비디오 시퀀스들에 내재하는 리던던시를 감소 또는 제거하기 위해 공간 예측 및/또는 시간 예측을 수행한다. 블록-기반 비디오 코딩에 대해, 비디오 프레임 또는 슬라이스는 매크로블록들 또는 코딩 트리 유닛들 (CTU들) 로 파티셔닝될 수도 있다. 이 문헌의 나머지는, 2 개의 용어들 매크로블록들 및 CTU들이 상호교환가능하게 사용된다. 각각의 매크로블록은 추가로 파티셔닝될 수 있다. 인트라-코딩된 (I) 프레임들 또는 슬라이스에서의 매크로블록들은 이웃하는 매크로블록들에 대한 공간 예측을 이용하여 인코딩된다. 인터-코딩된 (P 또는 B) 프레임 또는 슬라이스에서의 매크로블록들은 동일한 프레임 또는 슬라이스에서의 이웃하는 매크로블록들에 대한 공간 예측 또는 다른 참조 프레임들에 대한 시간 예측을 이용할 수도 있다.

[0005] 비디오 데이터가 인코딩된 후, 비디오 데이터는 송신 또는 저장을 위해 패키징될 수도 있다. 비디오 데이터는 ISO (International Organization for Standardization) 기반 미디어 파일 포맷 및 그 확장들, 이를 테면 AVC 파일 포맷과 같은 다양한 표준들 중 임의의 표준을 준수하는 비디오 파일로 어셈블링될 수도 있다.

[0006] 비디오 데이터가 송신 또는 저장을 위해 패키징된 후, 클라이언트 디바이스는 적합한 품질로 패키징된 비디오

데이터를 요청할 수도 있다. 예를 들어, 클라이언트 디바이스는, 패킷화된 비디오를 송신하는 디지털 비디오 디바이스와 클라이언트 디바이스 사이의 대역폭이 높을 때 상대적으로 고품질로 패킷화된 비디오 데이터를 요청하고 패킷화된 비디오를 송신하는 디지털 비디오 디바이스와 클라이언트 디바이스 사이의 대역폭이 낮을 때 상대적으로 저품질로 패킷화된 비디오 데이터를 요청할 수도 있다.

발명의 내용

해결하려는 과제

과제의 해결 수단

- [0007] 일반적으로, 본 개시는 비디오 데이터의 이미지의 하나 이상의 가장 관심있는 영역들을 나타내는 데이터를 포함하는 비트스트림을 생성하기 위한 기법들을 설명한다. 이미지의 하나 이상의 가장 관심있는 영역들은 이미지의 다른 영역들과 비교하여 이미지의 프리젠테이션 시간에 사용자에게 렌더링될 가능성이 더 크다. 디바이스는 하나 이상의 가장 관심있는 영역들을 이용하여 하나 이상의 가장 관심있는 영역들에 대한 비디오 데이터를 프리-페치할 수도 있다. 예를 들어, 디바이스는 이미지의 다른 영역들에 대한 비디오 데이터보다 저레이턴시로 및/또는 고품질로 하나 이상의 가장 관심있는 영역들에 대한 비디오 데이터를 페치할 수도 있고, 여기서 디바이스는 비디오 디코더를 포함하는 엔티티, 콘텐츠 전달 네트워크 (content delivery network; CDN) 노드와 같은 네트워크 엘리먼트, 또는 다른 엔티티인 것을 지칭할 수도 있다.
- [0008] 하나의 예에서, 비디오 데이터에 대한 정보를 결정하는 방법은, 회로부로 구현된, 클라이언트 디바이스의 프로세서에 의해, 비디오 데이터의 이미지의 복수의 영역들 중 하나 이상의 가장 관심있는 영역들을, 하나 이상의 가장 관심있는 영역들을 나타내는 데이터에 기초하여 결정하는 단계를 포함한다. 하나 이상의 가장 관심있는 영역들을 나타내는 데이터는 비디오 데이터를 포함하는 비트스트림의 비디오 코딩 계층 (video coding layer; VCL) 데이터의 외부에 있다. 방법은, 클라이언트 디바이스의 프로세서에 의해, 이미지의 하나 이상의 가장 관심있는 영역들에 대한 상대적으로 고품질 비디오 데이터에 대한 요청을 출력하는 단계 및 클라이언트 디바이스의 프로세서에 의해, 상대적으로 고품질 비디오 데이터에 대한 요청을 출력한 후, 이미지의 하나 이상의 다른 영역들에 대한 상대적으로 저품질 비디오 데이터에 대한 요청을 출력하는 단계를 더 포함한다.
- [0009] 다른 예에서, 비디오 데이터에 대한 정보를 결정하기 위한 디바이스는, 비디오 데이터의 이미지의 복수의 영역들 중 하나 이상의 가장 관심있는 영역들을, 하나 이상의 가장 관심있는 영역들을 나타내는 데이터에 기초하여 결정하도록 구성되는 회로부로 구현된 하나 이상의 프로세서들을 포함한다. 하나 이상의 가장 관심있는 영역들을 나타내는 데이터는 비디오 데이터를 포함하는 비트스트림의 VCL 데이터의 외부에 있다. 하나 이상의 프로세서들은 이미지의 하나 이상의 가장 관심있는 영역들에 대한 상대적으로 고품질 비디오 데이터에 대한 요청을 출력하고, 상대적으로 고품질 비디오 데이터에 대한 요청을 출력한 후, 이미지의 하나 이상의 다른 영역들에 대한 상대적으로 저품질 비디오 데이터에 대한 요청을 출력하도록 추가로 구성된다.
- [0010] 다른 예에서, 비디오 데이터에 대한 정보를 결정하기 위한 디바이스는, 비디오 데이터의 이미지의 복수의 영역들 중 하나 이상의 가장 관심있는 영역들을, 하나 이상의 가장 관심있는 영역들을 나타내는 데이터에 기초하여 결정하기 위한 수단을 포함한다. 하나 이상의 가장 관심있는 영역들을 나타내는 데이터는 비디오 데이터를 포함하는 비트스트림의 VCL 데이터의 외부에 있다. 디바이스는 이미지의 하나 이상의 가장 관심있는 영역들에 대한 상대적으로 고품질 비디오 데이터에 대한 요청을 출력하기 위한 수단 및 상대적으로 고품질 비디오 데이터에 대한 요청을 출력한 후, 이미지의 하나 이상의 다른 영역들에 대한 상대적으로 저품질 비디오 데이터에 대한 요청을 출력하기 위한 수단을 더 포함한다.
- [0011] 다른 예에서, 컴퓨터 프로그램 제품은, 실행될 때, 비디오 데이터에 대한 정보를 결정하기 위한 소스 디바이스의 프로세서로 하여금, 비디오 데이터의 이미지의 복수의 영역들 중 하나 이상의 가장 관심있는 영역들을, 하나 이상의 가장 관심있는 영역들을 나타내는 데이터에 기초하여 결정하게 하는 명령들을 포함하는 컴퓨터 판독가능 저장 매체를 포함한다. 하나 이상의 가장 관심있는 영역들을 나타내는 데이터는 비디오 데이터를 포함하는 비트스트림의 VCL 데이터의 외부에 있다. 명령들은 또한, 프로세서로 하여금, 이미지의 하나 이상의 가장 관심있는 영역들에 대한 상대적으로 고품질 비디오 데이터에 대한 요청을 출력하게 하고 상대적으로 고품질 비디오 데이터에 대한 요청을 출력한 후, 이미지의 하나 이상의 다른 영역들에 대한 상대적으로 저품질 비디오 데이터에 대한 요청을 출력하게 한다.

[0012] 다른 예에서, 비디오 데이터에 대한 정보를 시그널링하는 방법은, 회로부로 구현된, 소스 디바이스의 프로세서에 의해, 비디오 데이터의 이미지의 복수의 영역들 중 하나 이상의 가장 관심있는 영역들을 결정하는 단계를 포함한다. 하나 이상의 가장 관심있는 영역들은 클라이언트 디바이스에 의해 추출될 가능성이 가장 큰 하나 이상의 영역들을 포함한다. 방법은 소스 디바이스의 프로세서에 의해, 하나 이상의 가장 관심있는 영역들을 나타내는 데이터를 생성하는 단계를 더 포함한다. 하나 이상의 가장 관심있는 영역들을 나타내는 데이터는 비디오 데이터를 포함하는 비트스트림의 VCL 데이터의 외부에 있다.

[0013] 다른 예에서, 비디오 데이터에 대한 정보를 시그널링하기 위한 디바이스는 비디오 데이터의 이미지의 복수의 영역들 중 하나 이상의 가장 관심있는 영역들을 결정하도록 구성되는 회로부로 구현된 하나 이상의 프로세서들을 포함한다. 하나 이상의 가장 관심있는 영역들은 클라이언트 디바이스에 의해 추출될 가능성이 가장 큰 영역들을 포함한다. 하나 이상의 프로세서들은 하나 이상의 가장 관심있는 영역들을 나타내는 데이터를 생성하도록 추가로 구성된다. 하나 이상의 가장 관심있는 영역들을 나타내는 데이터는 비디오 데이터를 포함하는 비트스트림의 VCL 데이터의 외부에 있다.

[0014] 다른 예에서, 비디오 데이터에 대한 정보를 시그널링하기 위한 디바이스는 비디오 데이터의 이미지의 복수의 영역들 중 하나 이상의 가장 관심있는 영역들을 결정하기 위한 수단을 포함한다. 하나 이상의 가장 관심있는 영역들은 클라이언트 디바이스에 의해 추출될 가능성이 가장 큰 하나 이상의 영역들을 포함한다. 디바이스는 하나 이상의 가장 관심있는 영역들을 나타내는 데이터를 생성하기 위한 수단을 더 포함한다. 하나 이상의 가장 관심있는 영역들을 나타내는 데이터는 비디오 데이터를 포함하는 비트스트림의 VCL 데이터의 외부에 있다.

[0015] 다른 예에서, 컴퓨터 프로그램 제품은, 프로세서로 하여금, 비디오 데이터의 이미지의 복수의 영역들 중 하나 이상의 가장 관심있는 영역들을 결정하게 하는 명령들을 포함하는 컴퓨터 판독가능 저장 매체를 포함한다. 하나 이상의 가장 관심있는 영역들은 클라이언트 디바이스에 의해 추출될 가능성이 가장 큰 하나 이상의 영역들을 포함한다. 명령들은 또한, 프로세서로 하여금, 하나 이상의 가장 관심있는 영역들을 나타내는 데이터를 생성하게 한다. 하나 이상의 가장 관심있는 영역들을 나타내는 데이터는 비디오 데이터를 포함하는 비트스트림의 VCL 데이터의 외부에 있다.

[0016] 하나 이상의 예들의 상세들은 이하의 설명 및 첨부하는 도면들에 제시된다. 다른 피쳐들, 목적들, 및 이점들은 설명 및 도면들로부터, 그리고 청구항들로부터 명백할 것이다.

도면의 간단한 설명

[0017] 도 1 은 네트워크를 통해 미디어 데이터를 스트리밍하기 위한 기법들을 구현하는 예의 시스템을 예시하는 블록 다이어그램이다.

도 2 는 예의 멀티미디어 콘텐츠의 엘리먼트들을 예시하는 개념적 다이어그램이다.

도 3 은 도 2 의 표현에 대응할 수도 있는 예의 비디오 파일의 엘리먼트들을 예시하는 블록 다이어그램이다.

도 4 는 입방형 프로젝션에서의 예의 타일 코딩을 예시하는 개념적 다이어그램이다.

도 5 는 본 개시에서 설명된 하나 이상의 기법들에 따른 비디오 데이터에 대한 정보를 결정하기 위한 예의 접근법을 예시하는 플로우차트이다.

도 6 은 본 개시에서 설명된 하나 이상의 기법들에 따른 비디오 데이터에 대한 정보를 시그널링하기 위한 예의 접근법을 예시하는 플로우차트이다.

도 7 은 본 개시에서 설명된 하나 이상의 기법들에 따른 비디오 데이터를 프리-페치하기 위한 예의 접근법을 예시하는 플로우차트이다.

발명을 실시하기 위한 구체적인 내용

[0018] 본 발명은 가상 현실 (virtual reality; VR) 에 관한 것이다. 보다 구체적으로, 본 발명은 VR 비디오에서의 하나 이상의 가장 관심있는 영역들의 세트의 생성, 가장 관심있는 영역들의 정보의 시그널링, 및 VR 비디오 콘텐츠의 인코딩, 스트리밍, 플레이백, 및 관리에 있어서의 정보의 이용을 위한 일부 방법들을 개시한다.

[0019] 일반적으로, 본 개시는 비디오 데이터의 이미지의 하나 이상의 가장 관심있는 영역들을 나타내는 데이터를 포함하는 비디오 콘텐츠를 생성하기 위한 기법들을 설명한다. 디렉터 또는 프로듀서가 하나 이상의 가장 관심있

는 영역들을 선택할 수도 있거나 또는 컴퓨팅 디바이스가 서비스 또는 콘텐츠 제공자에 의한 사용자 통계 (user statistics) 로부터, 예를 들어, 전방향 비디오 콘텐츠가 스트리밍 서비스를 통하여 제공되었을 때 사용자들이 어느 영역들을 가장 많이 요청하였는지/보았는지의 통계를 통하여 하나 이상의 가장 관심있는 영역들을 도출할 수도 있다. 어느 경우나, 클라이언트 디바이스는 하나 이상의 가장 관심있는 영역들을 이용하여 하나 이상의 가장 관심있는 영역들에 대한 비디오 데이터를 프리-페치하여, 비디오, 이를 테면 360 비디오로 또한 지칭되는 가상 현실 비디오의 플레이백을 향상시킬 수도 있다.

[0020] 구체적으로, 클라이언트 디바이스는, 예를 들어, 비트스트림의 비디오 코딩 계층 데이터를 송신하기 위한 메시지 또는 헤더에서 수신된 정보를 이용하여, 하나 이상의 가장 관심있는 영역들을 나타내는 데이터를 결정할 수도 있다. 하나 이상의 가장 관심있는 영역들을 나타내는 데이터로부터 하나 이상의 가장 관심있는 영역들을 결정 시에, 클라이언트 디바이스는 상대적으로 고품질로 하나 이상의 가장 관심있는 영역들을 요청할 수도 있다. 상대적으로 고품질로 하나 이상의 가장 관심있는 영역들을 요청한 후, 클라이언트 디바이스는 사용자의 시야 (field of view; FOV) 에 기초하여 고품질보다 상대적으로 저품질로 이미지의 하나 이상의 다른 영역들을 요청할 수도 있다. 이 방식으로, 하나 이상의 가장 관심있는 영역들은 고품질로 하나 이상의 가장 관심있는 영역들이 수신되는 것을 허락하기 위해 프리-페치 (예를 들어, 사용자의 FOV 를 수신하기 이전에 페치) 되어, 비디오, 이를 테면, 가상 현실 비디오의 플레이백을 향상시킬 수도 있다. 즉, 디바이스는 하나 이상의 가장 관심있는 영역들을 이용하여 하나 이상의 가장 관심있는 영역들에 대한 비디오 데이터를 프리-페치할 수도 있다. 예를 들어, 디바이스는 이미지의 다른 영역들에 대한 비디오 데이터보다 저레이턴시로 및/또는 고품질로 하나 이상의 가장 관심있는 영역들에 대한 비디오 데이터를 페치할 수도 있고, 여기서 디바이스는 비디오 디코더를 포함하는 엔티티, 콘텐츠 전달 네트워크 (CDN) 노드와 같은 네트워크 엘리먼트, 또는 다른 엔티티인 것을 지칭할 수도 있다.

[0021] 본 개시의 기법들은 ISO 기반 미디어 파일 포맷, 스케일러블 비디오 코딩 (Scalable Video Coding; SVC) 파일 포맷, 어드밴스드 비디오 코딩 (Advanced Video Coding; AVC) 파일 포맷, 제 3 세대 파트너십 프로젝트 (3GPP) 파일 포맷, 및/또는 멀티뷰 비디오 코딩 (Multiview Video Coding; MVC) 파일 포맷, 또는 다른 유사한 비디오 파일 포맷들 중 임의의 것에 따라 캡슐화된 비디오 데이터를 준수하는 비디오 파일들에 적용될 수도 있다.

[0022] HTTP 스트리밍에서, 자주 사용되는 연산들은 HEAD, GET, 및 부분 (partial) GET 을 포함한다. HEAD 연산은 주어진 유니폼 리소스 로케이터 (uniform resource locator; URL) 또는 유니폼 리소스 이름 (uniform resource name; URN) 과 연관된 파일의 헤더를, URL 또는 URN 과 연관된 페이로드를 추출하지 않고 추출한다. GET 연산은 주어진 URL 또는 URN 과 연관된 전체 파일을 추출한다. 부분 GET 연산은 입력 파라미터로서 바이트 범위를 수신하고 파일의 연속 바이트들의 수를 추출하고, 여기서 바이트들의 수는 수신된 바이트 범위에 대응한다. 따라서, 무비 프래그먼트들은, 부분 GET 연산이 하나 이상의 개개의 무비 프래그먼트들을 얻을 수 있기 때문에 HTTP 스트리밍에 대해 제공될 수도 있다. 무비 프래그먼트에서, 상이한 트랙들의 여러 트랙 프래그먼트들이 존재할 수 있다. HTTP 스트리밍에서, 미디어 프리젠테이션은 클라이언트에 액세스가능한 데이터의 구조화된 컬렉션일 수도 있다. 클라이언트는 사용자에게 스트리밍 서비스를 제시하기 위해 미디어 데이터 정보를 요청 및 다운로드할 수도 있다.

[0023] HTTP 스트리밍을 이용하여 3GPP 데이터를 스트리밍하는 예에서, 멀티미디어 콘텐츠의 비디오 및/또는 오디오 데이터에 대한 다중 표현들이 존재할 수도 있다. 이하에 설명한 바와 같이, 상이한 표현들은 상이한 코딩 특성들 (예를 들어, 비디오 코딩 표준의 상이한 프로파일들 또는 레벨들), 상이한 코딩 표준들 또는 코딩 표준들의 확장들 (이를 테면 멀티뷰 및/또는 스케일러블 확장들), 또는 상이한 비트레이트들에 대응할 수도 있다. 이러한 표현들의 매니페스트 (manifest) 는 미디어 프리젠테이션 디스크립션 (Media Presentation Description; MPD) 데이터 구조에서 정의될 수도 있다. 미디어 프리젠테이션은 HTTP 스트리밍 클라이언트 디바이스에 액세스가능한 데이터의 구조화된 컬렉션에 대응할 수도 있다. HTTP 스트리밍 클라이언트 디바이스는 클라이언트 디바이스의 사용자에게 스트리밍 서비스를 제시하기 위해 미디어 데이터 정보를 요청 및 다운로드할 수도 있다. 미디어 프리젠테이션은 MPD 의 업데이트들을 포함할 수도 있는 MPD 데이터 구조에서 설명될 수도 있다.

[0024] 미디어 프리젠테이션은 하나 이상의 주기들의 시퀀스를 포함할 수도 있다. 각각의 주기는 다음 주기의 시작까지, 또는 마지막 주기의 경우에, 미디어 프리젠테이션의 종료까지 확장할 수도 있다. 각각의 주기는 동일한 미디어 콘텐츠에 대한 하나 이상의 표현들을 포함할 수도 있다. 표현은 오디오, 비디오, 타이밍된 텍스트, 또는 다른 이러한 데이터의 다수의 대안의 인코딩된 버전들 중 하나일 수도 있다. 표현들은 인코딩 타입들에 의해, 예를 들어, 비디오 데이터에 대한 비트레이트, 레졸루션, 및/또는 코덱 및 오디오 데이터에 대한

비트레이트, 언어, 및/또는 코덱에 의해 상이할 수도 있다. 용어 표현은 멀티미디어 콘텐츠의 특정한 주기에 대응하고 특정한 방식으로 인코딩된 인코딩된 오디오 또는 비디오 데이터의 섹션을 지칭하는데 사용될 수도 있다.

[0025] 특정한 주기의 표현들은 표현들이 속하는 적응 세트를 나타내는 MPD에서의 속성에 의해 표시된 그룹에 배정될 수도 있다. 동일한 적응 세트에서의 표현들은 일반적으로, 예를 들어, 대역폭 적응을 수행하기 위해, 클라이언트 디바이스가 이들 표현들 사이에서 동적으로 그리고 심리스로 스위칭할 수 있다는 점에서, 서로에 대한 대안물들로 간주된다. 예를 들어, 특정한 주기에 대한 비디오 데이터의 각각의 표현은, 대응하는 주기에 대한 멀티미디어 콘텐츠의, 비디오 데이터 또는 오디오 데이터와 같은 미디어 데이터를 제시하기 위해 표현들 중 임의의 것이 디코딩을 위해 선택될 수도 있도록, 동일한 적응 세트에 배정될 수도 있다. 하나의 주기 내의 미디어 콘텐츠는, 존재하면, 그룹 0 으로부터의 하나의 표현, 또는 일부 예들에서, 각각의 비-제로 그룹으로부터의 많아야 하나의 표현의 조합 중 어느 하나에 의해 표현될 수도 있다. 주기의 각각의 표현에 대한 타이밍 데이터는 주기의 시작 시간에 대하여 표현될 수도 있다.

[0026] 표현은 하나 이상의 세그먼트들을 포함할 수도 있다. 각각의 표현은 초기화 세그먼트를 포함할 수도 있거나, 또는 표현의 각각의 세그먼트는 자체-초기화중일 수도 있다. 존재 시, 초기화 세그먼트는 표현에 액세스하기 위한 초기화 정보를 포함할 수도 있다. 일반적으로, 초기화 세그먼트는 미디어 데이터를 포함하지 않는다. 세그먼트는 식별자, 이를 테면, 유니폼 리소스 로케이터 (URL), 유니폼 리소스 이름 (URN), 또는 유니폼 리소스 식별자 (uniform resource identifier; URI) 에 의해 고유하게 언급될 수도 있다. MPD는 각각의 세그먼트에 대한 식별자들을 제공할 수도 있다. 일부 예들에서, MPD는 또한, URL, URN, 또는 URI 에 의해 액세스가능한 파일 내의 세그먼트에 대한 데이터에 대응할 수도 있는, 범위 속성의 형태로 바이트 범위들을 제공할 수도 있다.

[0027] 상이한 표현들은 상이한 타입들의 미디어 데이터에 대한 실질적으로 동시 취출을 위해 선택될 수도 있다. 예를 들어, 클라이언트 디바이스는 세그먼트들을 취출할 오디오 표현, 비디오 표현, 및 타이밍된 텍스트 표현을 선택할 수도 있다. 일부 예들에서, 클라이언트 디바이스는 대역폭 적응을 수행하기 위해 특정한 적응 세트들을 선택할 수도 있다. 즉, 클라이언트 디바이스는 비디오 표현들을 포함하는 적응 세트, 오디오 표현들을 포함하는 적응 세트, 및/또는 타이밍된 텍스트를 포함하는 적응 세트를 선택할 수도 있다. 대안으로, 클라이언트 디바이스는 소정의 타입들의 미디어 (예를 들어, 비디오) 에 대한 적응 세트들을 선택하고, 다른 타입들의 미디어 (예를 들어, 오디오 및/또는 타이밍된 텍스트) 에 대한 표현들을 직접 선택할 수도 있다.

[0028] 도 1 은 네트워크를 통해 미디어 데이터를 스트리밍하기 위한 기법들을 구현하는 예의 시스템 (10) 을 예시하는 블록 다이어그램이다. 이 예에서, 시스템 (10) 은 콘텐츠 준비 디바이스 (20), 서버 디바이스 (60), 및 클라이언트 디바이스 (40) 를 포함한다. 클라이언트 디바이스 (40) 및 서버 디바이스 (60) 는 인터넷을 포함할 수도 있는 네트워크 (74) 에 의해 통신가능하게 커플링된다. 일부 예들에서, 콘텐츠 준비 디바이스 (20) 및 서버 디바이스 (60) 는 또한, 네트워크 (74) 또는 다른 네트워크에 의해 커플링될 수도 있거나, 또는 직접 통신가능하게 커플링될 수도 있다. 일부 예들에서, 콘텐츠 준비 디바이스 (20) 및 서버 디바이스 (60) 는 동일한 디바이스를 포함할 수도 있다.

[0029] 콘텐츠 준비 디바이스 (20) 는, 도 1 의 예에서, 오디오 소스 (22) 및 비디오 소스 (24) 를 포함한다. 오디오 소스 (22) 는, 예를 들어, 오디오 인코더 (26) 에 의해 인코딩될 캡처된 오디오 데이터를 나타내는 전기 신호들을 생성하는 마이크로폰을 포함할 수도 있다. 대안으로, 오디오 소스 (22) 는 이전에 레코딩된 오디오 데이터를 저장하는 저장 매체, 컴퓨터화된 합성기와 같은 오디오 데이터 생성기, 또는 오디오 데이터의 임의의 다른 소스를 포함할 수도 있다. 비디오 소스 (24) 는 비디오 인코더 (28) 에 의해 인코딩될 비디오 데이터를 생성하는 비디오 카메라, 이전에 레코딩된 비디오 데이터로 인코딩된 저장 매체, 컴퓨터 그래픽스 소스와 같은 비디오 데이터 생성 유닛, 또는 비디오 데이터의 임의의 다른 소스를 포함할 수도 있다. 콘텐츠 준비 디바이스 (20) 가 반드시 모든 예들에서 서버 디바이스 (60) 에 통신가능하게 커플링되는 것은 아니며, 서버 디바이스 (60) 에 의해 관독되는 별도의 매체에 멀티미디어 콘텐츠를 저장할 수도 있다.

[0030] 원시 (raw) 오디오 및 비디오 데이터는 아날로그 또는 디지털 데이터를 포함할 수도 있다. 아날로그 데이터는 오디오 인코더 (26) 및/또는 비디오 인코더 (28) 에 의해 인코딩되기 전에 디지털화될 수도 있다. 오디오 소스 (22) 는 말하는 참가자가 말하고 있는 동안 말하는 참가자로부터 오디오 데이터를 획득할 수도 있고, 비디오 소스 (24) 는 동시에 말하는 참가자의 비디오 데이터를 획득할 수도 있다. 다른 예들에서, 오디오 소스 (22) 는 저장된 오디오 데이터를 포함하는 컴퓨터 관독가능 저장 매체를 포함할 수도 있고, 비디오 소스

(24) 는 저장된 비디오 데이터를 포함하는 컴퓨터 판독가능 저장 매체를 포함할 수도 있다. 이 방식으로, 본 개시에서 설명된 기법들은 라이브, 스트리밍, 실시간 오디오 및 비디오 데이터에 또는 아카이브된, 프리-레코딩된 오디오 및 비디오 데이터에 적용될 수도 있다.

[0031] 비디오 프레임들에 대응하는 오디오 프레임들은 일반적으로 오디오 소스 (22) 에 의해 캡처된 (또는 생성된) 오디오 데이터를 비디오 프레임들 내에 포함되는 비디오 소스 (24) 에 의해 캡처된 (또는 생성된) 비디오 데이터와 동시에 포함하는 오디오 프레임들이다. 예를 들어, 말하는 참가자가 일반적으로 말하는 것으로 오디오 데이터를 생성하는 동안에, 오디오 소스 (22) 는 오디오 데이터를 캡처하고, 비디오 소스 (24) 는 동시에, 즉 오디오 소스 (22) 가 오디오 데이터를 캡처하고 있는 동안에 말하는 참가자의 비디오 데이터를 캡처한다. 이런 이유로, 오디오 프레임은 하나 이상의 특정한 비디오 프레임들에 일시적으로 대응할 수도 있다. 이에 따라, 비디오 프레임에 대응하는 오디오 프레임은 일반적으로, 오디오 데이터 및 비디오 데이터가 동시에 캡처되었고 오디오 프레임 및 비디오 프레임이 동시에 캡처된 오디오 데이터 및 비디오 데이터를 각각 포함하는 상황에 대응한다.

[0032] 일부 예들에서, 오디오 인코더 (26) 는 인코딩된 오디오 프레임에 대한 오디오 데이터가 레코딩된 시간을 나타내는 각각의 인코딩된 오디오 프레임에서의 타임스탬프를 인코딩할 수도 있고, 유사하게, 비디오 인코더 (28) 는 인코딩된 비디오 프레임에 대한 비디오 데이터가 레코딩된 시간을 나타내는 각각의 인코딩된 비디오 프레임에서의 타임스탬프를 인코딩할 수도 있다. 이러한 예들에서, 비디오 프레임에 대응하는 오디오 프레임은 타임스탬프를 포함하는 오디오 프레임 및 동일한 타임스탬프를 포함하는 비디오 프레임을 포함할 수도 있다. 콘텐츠 준비 디바이스 (20) 는, 오디오 인코더 (26) 및/또는 비디오 인코더 (28) 가 타임스탬프들을 생성할 수도 있거나, 또는 오디오 소스 (22) 및 비디오 소스 (24) 가 오디오 및 비디오 데이터를, 각각, 타임스탬프와 연관시키기 위해 이용할 수도 있는 내부 클럭을 포함할 수도 있다.

[0033] 일부 예들에서, 오디오 소스 (22) 는 오디오 데이터가 레코딩된 시간에 대응하는 데이터를 오디오 인코더 (26) 로 전송할 수도 있고, 비디오 소스 (24) 는 비디오 데이터가 레코딩된 시간에 대응하는 데이터를 비디오 인코더 (28) 로 전송할 수도 있다. 일부 예들에서, 오디오 인코더 (26) 는 인코딩된 오디오 데이터의 상대 시간 순서를 표시하기 위해 (그러나 오디오 데이터가 레코딩된 절대 시간을 반드시 표시하는 것은 아님) 인코딩된 오디오 데이터에서의 시퀀스 식별자를 인코딩할 수도 있고, 유사하게, 비디오 인코더 (28) 는 또한 인코딩된 비디오 데이터의 상대 시간 순서를 표시하기 위해 시퀀스 식별자들을 이용할 수도 있다. 유사하게, 일부 예들에서, 시퀀스 식별자는 타임스탬프와 맵핑 또는 다르게는 상관될 수도 있다.

[0034] 오디오 인코더 (26) 는 일반적으로 인코딩된 오디오 데이터의 스트림을 생성하는 한편, 비디오 인코더 (28) 는 인코딩된 비디오 데이터의 스트림을 생성한다. (오디오이든 비디오이든) 데이터의 각 개개의 스트림은 기본 스트림 (elementary stream) 으로 지칭될 수도 있다. 기본 스트림은 표현의 단일, 디지털로 코딩된 (어쩌면 압축된) 컴포넌트이다. 예를 들어, 표현의 코딩된 비디오 또는 오디오 파트는 기본 스트림일 수 있다. 기본 스트림은 비디오 파일 내에 캡슐화되기 전에 패킷화된 기본 스트림 (packetized elementary stream; PES) 으로 컨버팅될 수도 있다. 동일한 표현 내에서, 스트림 ID 는 하나의 기본 스트림에 속하는 PES-패킷들을 다른 것과 구별하는데 이용될 수도 있다. 기본 스트림의 데이터의 베이직 유닛은 패킷화된 기본 스트림 (PES) 패킷이다. 따라서, 코딩된 비디오 데이터는 일반적으로 기본 비디오 스트림들에 대응한다. 유사하게, 오디오 데이터는 하나 이상의 개별의 기본 스트림들에 대응한다.

[0035] 많은 비디오 코딩 표준들, 이를 테면 ITU-T H.264/AVC 및 업커밍 고효율 비디오 코딩 (HEVC) 표준은, 임의의 것이 소정의 프로파일 또는 레벨을 준수하는, 에러-프리 비트스트림들에 대한 신택스, 시맨틱스, 및 디코딩 프로세스를 정의한다. 비디오 코딩 표준들은 통상적으로 인코더를 특정하지 않지만, 인코더는 생성된 비트스트림들이 디코더에 대해 표준-순응 (standard-compliant) 함을 보장하는 임무를 맡는다. 비디오 코딩 표준들의 문맥에서, "프로파일" 은 알고리즘들, 피처들, 또는 툴들 및 그들에 적용하는 제약들의 서브세트에 대응한다. H.264 표준에 의해 정의한 바와 같이, 예를 들어, "프로파일" 은 H.264 표준에 의해 특정되는 전체 비트스트림 신택스의 서브세트이다. "레벨" 은, 픽처들의 레졸루션, 비트 레이트, 및 블록 프로세싱 레이트에 관련되는, 예를 들어 디코더 메모리 및 컴퓨테이션과 같은 디코더 리소스 소비의 한계들에 대응한다. 프로파일은 profile-idc (프로파일 표시자) 값으로 시그널링될 수도 있는 한편, 레벨은 level-idc (레벨 표시자) 값으로 시그널링될 수도 있다.

[0036] H.264 표준은, 예를 들어, 주어진 프로파일의 신택스에 의해 부과된 한도 내에, 디코딩된 픽처들의 특정된 사이즈와 같은 비트스트림에서의 신택스 엘리먼트들에 의해 취해진 값들에 의존하여 인코더들 및 디코더들의 성능의

큰 변동을 요구하는 것이 여전히 가능하다는 것을 인정한다. H.264 표준은 또한, 많은 애플리케이션들에서, 특정한 프로파일 내의 선택스의 모든 가정적 사용 (hypothetical use) 들을 다루는 것이 가능한 디코더를 구현하는 것이 현실적이지도 경제적이지도 않다는 것을 인정한다. 이에 따라, H.264 표준은 비트스트림에서의 선택스 엘리먼트들의 값들에 부과된 제약들의 특정된 세트로서 "레벨" 을 정의한다. 이들 제약들은 값들에 대한 단순 제한들일 수도 있다. 대안으로, 이들 제약들은 값들의 산술 조합들 (예를 들어, 픽처 폭 곱하기 픽처 높이 곱하기 초당 디코딩된 픽처들의 수) 에 대한 제약들의 형태를 취할 수도 있다. H.264 표준은 또한, 개개의 구현들이 각각의 지원된 프로파일에 대해 상이한 레벨을 지원할 수도 있다는 것을 규정한다.

[0037] 프로파일을 준수하는 디코더는 보통 프로파일에서 정의된 모든 픽처들을 지원한다. 예를 들어, 코딩 픽처로서, B-픽처 코딩은 H.264/AVC 의 베이스라인 프로파일에서 지원되지 않고 H.264/AVC 의 다른 프로파일들에서 지원된다. 레벨을 준수하는 디코더는 레벨에서 정의된 한계들을 넘어서 리소스들을 요구하지 않는 임의의 비트스트림을 디코딩하는 것이 가능해야 한다. 프로파일들 및 레벨들의 정의들은 해석능력 (interpretability) 에 도움이 될 수도 있다. 예를 들어, 비디오 송신 동안, 한 쌍의 프로파일 및 레벨 정의들은 전체 송신 세션 동안 협상되고 합의될 수도 있다. 보다 구체적으로, H.264/AVC 에서, 레벨은 디코딩된 픽처 버퍼 (DPB) 사이즈, 코딩된 픽처 버퍼 (CPB) 사이즈, 수직 모션 벡터 범위, 2 개의 연속적인 MB들 당 모션 벡터들의 최대 수, B-블록이 8x8 픽셀들 미만의 서브-매크로블록 파티션들을 가질 수 있는 여부, 및 프로 세싱되어야 하는 매크로블록들의 수에 대한 한계들을 정의할 수도 있다. 이 방식으로, 디코더는 디코더가 비트스트림을 적절히 디코딩하는 것이 가능한지 여부를 결정할 수도 있다.

[0038] 도 1 의 예에서, 콘텐츠 준비 디바이스 (20) 의 캡슐화 유닛 (30) 은 비디오 인코더 (28) 로부터 코딩된 비디오 데이터를 포함하는 기본 스트림들을, 그리고 오디오 인코더 (26) 로부터 코딩된 오디오 데이터를 포함하는 기본 스트림들을 수신한다. 일부 예들에서, 비디오 인코더 (28) 및 오디오 인코더 (26) 는 인코딩된 데이터로부터 PES 패킷들을 형성하기 위해 패킷화기들을 각각 포함할 수도 있다. 다른 예들에서, 비디오 인코더 (28) 및 오디오 인코더 (26) 는 인코딩된 데이터로부터 PES 패킷들을 형성하기 위해 개별의 패킷화기들과 각각 인터 페이스할 수도 있다. 여전히 다른 예들에서, 캡슐화 유닛 (30) 은 인코딩된 오디오 및 비디오 데이터로부터 PES 패킷들을 생성하기 위해 패킷화기들을 포함할 수도 있다.

[0039] 비디오 인코더 (28) 는, 다양한 비트레이트들로 그리고 다양한 특성들, 이를 테면 픽셀 레졸루션들, 프레임 레이트들, 다양한 코딩 표준들에 대한 적합성 (conformance), 다양한 코딩 표준들에 대한 다양한 프로파일들 및/ 또는 프로파일들의 레벨들에 대한 적합성, (예를 들어, 2 차원 또는 3 차원 플레이백을 위한) 하나 또는 다중 뷰들을 갖는 표현들, 또는 다른 이러한 특성들을 가진 멀티미디어 콘텐츠의 상이한 표현들을 생성하기 위해, 다양한 방식들로 멀티미디어 콘텐츠의 비디오 데이터를 인코딩할 수도 있다. 표현은, 본 개시에서 사용한 바와 같이, 오디오 데이터, 비디오 데이터, 텍스트 데이터 (예를 들어, 클로즈드 캡션들을 위한), 또는 다른 이러한 데이터 중 하나를 포함할 수도 있다. 표현은 기본 스트림, 이를 테면 오디오 기본 스트림 또는 비디오 기본 스트림을 포함할 수도 있다. 각각의 PES 패킷은 PES 패킷이 속하는 기본 스트림을 식별하는 stream_id 를 포함할 수도 있다. 캡슐화 유닛 (30) 은 기본 스트림들을 다양한 표현들의 비디오 파일들 (예를 들어, 세그먼트들) 로 어셈블링하는 것을 담당한다.

[0040] 캡슐화 유닛 (30) 은 오디오 인코더 (26) 및 비디오 인코더 (28) 로부터 표현의 기본 스트림들에 대한 PES 패킷들을 수신하고 PES 패킷들로부터 대응하는 네트워크 추상화 계층 (network abstraction layer; NAL) 유닛들을 형성한다. 코딩된 비디오 세그먼트들은, 비디오 텔레포니, 저장, 브로드캐스트, 또는 스트리밍과 같은 애플리케이션들을 어드레싱하는 "네트워크-친화적 (network-friendly)" 비디오 표현을 제공하는, NAL 유닛들로 조직 화될 수도 있다. NAL 유닛들은 비디오 코딩 계층 (Video Coding Layer; VCL) NAL 유닛들 (예를 들어, VCL 데이터) 및 비-VCL NAL 유닛들로 카테고리화될 수 있다. VCL 유닛들 (VCL 데이터로도 또한 지칭됨) 은 코어 압축 엔진을 포함할 수도 있고 블록, 매크로블록, 및/ 또는 슬라이스 레벨 데이터를 포함할 수도 있다. 다른 NAL 유닛들은 비-VCL NAL 유닛들일 수도 있다. 일부 예들에서, 보통 프라이어리 코딩된 픽처로서 제시되는, 하나의 시간 인스턴스에서의 코딩된 픽처는, 하나 이상의 NAL 유닛들을 포함할 수도 있는 액세스 유닛에 포함될 수도 있다.

[0041] 비-VCL NAL 유닛들은 다른 것들 중에서도, 파라미터 세트 NAL 유닛들 및 SEI NAL 유닛들을 포함할 수도 있다. 파라미터 세트들은 (시퀀스 파라미터 세트들 (SPS) 에서의) 시퀀스-레벨 헤더 정보 및 (픽처 파라미터 세트들 (PPS) 에서의) 드물게 변화하는 픽처-레벨 헤더 정보를 포함할 수도 있다. 파라미터 세트들 (예를 들어, PPS 및 SPS) 로, 드물게 변화하는 정보는 각각의 시퀀스 또는 픽처에 대해 반복될 필요가 없고, 이런 이유로 코딩 효율이 향상될 수도 있다. 더욱이, 파라미터 세트들의 이용은 중요한 헤더 정보의 대역외 (out-of-band)

송신을 가능하게 하여, 에러 내성을 위한 리던던트 송신들에 대한 필요성을 회피할 수도 있다. 대역의 송신 예들에서, 파라미터 세트 NAL 유닛들은 SEI NAL 유닛들과 같은 다른 NAL 유닛들과는 상이한 채널 상에서 송신될 수도 있다.

[0042] 보충 강화 정보 (Supplemental Enhancement Information; SEI) 는, VCL NAL 유닛들로부터 코딩된 픽처들 샘플들을 디코딩하는데 필요하지 않지만, 디코딩, 디스플레이, 에러 내성, 및 다른 목적들에 관련된 프로세스들을 도울 수도 있는 정보를 포함할 수도 있다. SEI 메시지들은 비-VCL NAL 유닛들에 포함될 수도 있다. SEI 메시지들은 일부 표준 규격들의 규범적 (normative) 파트이고, 따라서 표준 순응 디코더 구현에 대해 항상 의무적인 것은 아니다. SEI 메시지들은 시퀀스 레벨 SEI 메시지들 또는 픽처 레벨 SEI 메시지들일 수도 있다.

일부 시퀀스 레벨 정보는 SEI 메시지들, 이를 테면 SVC 의 예에서의 스케일러빌리티 정보 SEI 메시지들 및 MVC 에서의 뷰 스케일러빌리티 정보 SEI 메시지들에 포함될 수도 있다. 이들 예의 SEI 메시지들은, 예를 들어, 동작 포인트들의 추출 및 동작 포인트들의 특성들에 대한 정보를 전달할 수도 있다. 추가로, 캡슐화 유닛 (30) 은 매니페스트 파일, 이를 테면 표현들의 특성들을 설명하는 미디어 프리젠테이션 디스크립터 (media presentation descriptor; MPD) 를 형성할 수도 있다. 캡슐화 유닛 (30) 은 확장가능 마크업 언어 (extensible markup language; XML) 에 따라 MPD 를 포맷팅할 수도 있다.

[0043] 캡슐화 유닛 (30) 은 출력 인터페이스 (32) 에 매니페스트 파일 (예를 들어, MPD) 과 함께, 멀티미디어 콘텐츠의 하나 이상의 표현들에 대한 데이터를 제공할 수도 있다. 출력 인터페이스 (32) 는 네트워크 인터페이스 또는 저장 매체에 기입하기 위한 인터페이스, 이를 테면 범용 시리얼 버스 (USB) 인터페이스, CD 또는 DVD 라이터 또는 버너, 자기 또는 플래시 저장 매체들에 대한 인터페이스, 또는 미디어 데이터를 저장 또는 송신하기 위한 다른 인터페이스들을 포함할 수도 있다. 캡슐화 유닛 (30) 은 출력 인터페이스 (32) 에 멀티미디어 콘텐츠의 표현들의 각각의 데이터를 제공할 수도 있고, 그 출력 인터페이스는 네트워크 송신 또는 저장 매체들을 통해 데이터를 서버 디바이스 (60) 로 전송할 수도 있다. 도 1 의 예에서, 서버 디바이스 (60) 는 개별의 매니페스트 파일 (66) 및 하나 이상의 표현들 (68A 내지 68N) (표현들 (68)) 을 각각 포함하는 다양한 멀티미디어 콘텐츠들 (64) 을 저장하는 저장 매체 (62) 를 포함한다. 일부 예들에서, 출력 인터페이스 (32) 는 또한 데이터를 바로 네트워크 (74) 로 전송할 수도 있다.

[0044] 일부 예들에서, 표현들 (68) 은 적응 세트들로 분리될 수도 있다. 즉, 표현들 (68) 의 다양한 서브세트들은 개별의 공통 세트들의 특성들, 이를 테면 코덱, 프로파일 및 레벨, 레졸루션, 뷰들의 수, 세그먼트들에 대한 파일 포맷, 예를 들어, 화자들에게 의해 디코딩 및 제시될 오디오 데이터 및/또는 표현과 함께 디스플레이될 텍스트의 언어 또는 다른 특성들을 식별할 수도 있는 텍스트 타입 정보, 적응 세트에서의 표현들에 대한 장면의 카메라 각도 또는 실세계 카메라 관점을 설명할 수도 있는 카메라 각도 정보, 특정한 관중들에 대한 콘텐츠 적합성을 설명하는 레이팅 (rating) 정보 등을 포함할 수도 있다.

[0045] 매니페스트 파일 (66) 은, 적응 세트들에 대한 공통 특성들은 물론, 특정한 적응 세트들에 대응하는 표현들 (68) 의 서브세트들을 나타내는 데이터를 포함할 수도 있다. 매니페스트 파일 (66) 은 또한, 적응 세트들의 개개의 표현들에 대해, 비트레이트들과 같은 개개의 특성들을 나타내는 데이터를 포함할 수도 있다. 이 방식으로, 적응 세트는 단순화된 네트워크 대역폭 적응을 위해 제공될 수도 있다. 적응 세트에서의 표현들은 매니페스트 파일 (66) 의 적응 세트 엘리먼트의 자식 엘리먼트들을 이용하여 표시될 수도 있다.

[0046] 서버 디바이스 (60) 는 요청 프로세싱 유닛 (70) 및 네트워크 인터페이스 (72) 를 포함한다. 일부 예들에서, 서버 디바이스 (60) 는 복수의 네트워크 인터페이스들을 포함할 수도 있다. 더욱이, 서버 디바이스 (60) 의 피처들 중 임의의 것 또는 전부는 라우터들, 브릿지들, 프록시 디바이스들, 스위치들, 또는 다른 디바이스들과 같은, 콘텐츠 전달 네트워크의 다른 디바이스들 상에서 구현될 수도 있다. 일부 예들에서, 콘텐츠 전달 네트워크의 중간 디바이스들은 멀티미디어 콘텐츠 (64) 의 데이터를 캐시하고, 실질적으로 서버 디바이스 (60) 의 컴포넌트들을 준수하는 컴포넌트들을 포함할 수도 있다. 일반적으로, 네트워크 인터페이스 (72) 는 네트워크 (74) 를 통해 데이터를 전송 및 수신하도록 구성된다.

[0047] 요청 프로세싱 유닛 (70) 은 저장 매체 (62) 의 데이터를 위해, 클라이언트 디바이스 (40) 와 같은 클라이언트 디바이스들로부터 네트워크 요청들을 수신하도록 구성된다. 예를 들어, 요청 프로세싱 유닛 (70) 은 RFC 2616, "Hypertext Transfer Protocol - HTTP/1.1", by R. Fielding et al, Network Working Group, IETF, June 1999 에서 설명한 바와 같이, 하이퍼텍스트 전송 프로토콜 (HTTP) 버전 1.1 을 구현할 수도 있다. 즉, 요청 프로세싱 유닛 (70) 은 HTTP GET 또는 부분 GET 요청들을 수신하고 요청들에 응답하여 멀티미디어 콘텐츠 (64) 의 데이터를 제공하도록 구성될 수도 있다. 요청들은 예를 들어 세그먼트의 URL 을 이용하여, 표현들

(68) 중 하나의 세그먼트를 특정할 수도 있다. 일부 예들에서, 요청들은 또한 세그먼트의 하나 이상의 바이트 범위들을 특정하여, 부분 GET 요청들을 포함할 수도 있다. 요청 프로세싱 유닛 (70) 은 표현들 (68) 중 하나의 세그먼트의 헤더 데이터를 제공하기 위해 HTTP HEAD 요청들을 서비스하도록 추가로 구성될 수도 있다. 어느 경우나, 요청 프로세싱 유닛 (70) 은 요청들을 프로세싱하여 클라이언트 디바이스 (40) 와 같은 요청 디바이스에 요청된 데이터를 제공하도록 구성될 수도 있다.

[0048] 추가로 또는 대안으로, 요청 프로세싱 유닛 (70) 은 eMBMS 와 같은 브로드캐스트 또는 멀티캐스트 프로토콜을 통해 미디어 데이터를 전달하도록 구성될 수도 있다. 콘텐츠 준비 디바이스 (20) 는 설명한 바와 실질적으로 동일한 방식으로 DASH 세그먼트들 및/또는 서브-세그먼트들을 생성할 수도 있지만, 서버 디바이스 (60) 는 eMBMS 또는 다른 브로드캐스트 또는 멀티캐스트 네트워크 전송 프로토콜을 이용하여 이들 세그먼트들 또는 서브-세그먼트들을 전달할 수도 있다. 예를 들어, 요청 프로세싱 유닛 (70) 은 클라이언트 디바이스 (40) 로부터 멀티캐스트 그룹 가입 요청을 수신하도록 구성될 수도 있다. 즉, 서버 디바이스 (60) 는 특정한 미디어 콘텐츠 (예를 들어, 라이브 이벤트의 브로드캐스트) 와 연관된, 클라이언트 디바이스 (40) 를 포함하는, 클라이언트 디바이스들에 멀티캐스트 그룹과 연관된 인터넷 프로토콜 (IP) 어드레스를 광고할 수도 있다. 클라이언트 디바이스 (40) 는 차례로, 멀티캐스트 그룹에 가입하기 위한 요청을 제출할 수도 있다. 이 요청은 네트워크 (74), 예를 들어, 네트워크 (74) 를 구성하는 라우터들 전반에 걸쳐 전파될 수도 있어, 라우터들은 멀티캐스트 그룹과 연관된 IP 어드레스를 목적지로 한 트래픽을 클라이언트 디바이스 (40) 와 같은 가입 클라이언트 디바이스들로 향하게 한다.

[0049] 요청 프로세싱 유닛 (70) 은 요청들로부터 데이터를 어셈블링하도록 구성될 수도 있다. 예를 들어, 요청 프로세싱 유닛 (70) 은, 클라이언트 디바이스들로부터의 각각의 요청으로부터, 요청된 이미지의 하나 이상의 영역들의 표시를 추출할 수도 있다. 예를 들어, 요청 프로세싱 유닛 (70) 은 클라이언트 디바이스 (40) 로부터의 요청으로부터 특정한 이미지에 대한 사용자의 시야를 추출할 수도 있다. 추출한 후, 요청 프로세싱 유닛 (70) 은, 각각의 이미지에 대해, 클라이언트 디바이스 (예를 들어, 클라이언트 디바이스 (40)) 에 의해 요청된 사용자의 시야의 하나 이상의 영역들을 표시하는 테이블 또는 데이터베이스에서의 엔트리를 생성할 수도 있다.

[0050] 요청 프로세싱 유닛 (70) 은 어셈블링된 데이터로부터 하나 이상의 가장 관심있는 영역들을 도출할 수도 있다. 보다 구체적으로, 예를 들어, 이미지의 하나 이상의 영역들이 이미지의 다른 영역들과 비교하여 이미지의 가장 자주 요청된 영역들이라고 결정하는 것에 응답하여, 요청 프로세싱 유닛 (70) 은 하나 이상의 영역들을 이미지의 하나 이상의 가장 관심있는 영역들로서 선택할 수도 있다. 일부 예들에서, 요청 프로세싱 유닛 (70) 은 하나 이상의 영역들이 임계치를 충족하는 빈도로 사용자의 시야에 포함되었을 때 하나 이상의 영역들을 이미지의 하나 이상의 가장 관심있는 영역들로서 선택할 수도 있다. 예를 들어, 요청 프로세싱 유닛 (70) 은 사용자의 시야들의 적어도 절반이 하나 이상의 영역들을 포함한다는 것을 어셈블링된 데이터가 표시할 때 하나 이상의 영역들을 이미지의 하나 이상의 가장 관심있는 영역들로서 선택할 수도 있다.

[0051] 도 1 의 예에 예시한 바와 같이, 멀티미디어 콘텐츠 (64) 는 미디어 프리젠테이션 디스크립션 (MPD) 에 대응할 수도 있는 매니페스트 파일 (66) 을 포함한다. 매니페스트 파일 (66) 은 상이한 대안의 표현들 (68) (예를 들어, 상이한 품질들을 가진 비디오 서비스들) 의 디스크립션들을 포함할 수도 있고 디스크립션은, 예를 들어, 코덱 정보, 프로파일 값, 레벨 값, 비트레이트, 및 표현들 (68) 의 다른 설명적 특성들을 포함할 수도 있다. 클라이언트 디바이스 (40) 는 표현들 (68) 의 세그먼트들에 액세스하는 방법을 결정하기 위해 미디어 프리젠테이션의 MPD 를 추출할 수도 있다.

[0052] 특히, 추출 유닛 (52) 은 비디오 디코더 (48) 의 디코딩 능력들 및 비디오 출력 (44) 의 렌더링 능력들을 결정하기 위해 클라이언트 디바이스 (40) 의 구성 데이터 (미도시) 를 추출할 수도 있다. 구성 데이터는 또한, 클라이언트 디바이스 (40) 의 사용자에게 의해 선택된 언어 선호도, 클라이언트 디바이스 (40) 의 사용자에게 의해 설정된 심도 선호도들에 대응하는 하나 이상의 카메라 관점들, 및/또는 클라이언트 디바이스 (40) 의 사용자에게 의해 선택된 레이팅 선호도 중 임의의 것 또는 전부를 포함할 수도 있다. 추출 유닛 (52) 은, 예를 들어, HTTP GET 및 부분 GET 요청들을 제출하도록 구성된 웹 브라우저 또는 미디어 클라이언트를 포함할 수도 있다. 추출 유닛 (52) 은 클라이언트 디바이스 (40) 의 하나 이상의 프로세서들 또는 프로세싱 유닛들 (미도시) 에 의해 실행된 소프트웨어 명령들에 대응할 수도 있다. 일부 예들에서, 추출 유닛 (52) 에 대하여 설명된 기능성의 전부 또는 부분들은 하드웨어로, 또는 하드웨어, 소프트웨어, 및/또는 펌웨어의 조합으로 구현될 수도 있고, 여기서 필요한 하드웨어가 소프트웨어 또는 펌웨어에 대한 명령들을 실행하기 위해 제공될 수도 있다.

[0053] 추출 유닛 (52) 은 클라이언트 디바이스 (40) 의 디코딩 및 렌더링 능력들을 매니페스트 파일 (66) 의 정보에

의해 표시된 표현들 (68) 의 특성들과 비교할 수도 있다. 추출 유닛 (52) 은 표현들 (68) 의 특성들을 결정하기 위해 매니페스트 파일 (66) 의 적어도 부분을 초기에 추출할 수도 있다. 예를 들어, 추출 유닛 (52) 은 하나 이상의 적응 세트들의 특성들을 설명하는 매니페스트 파일 (66) 의 부분을 요청할 수도 있다. 추출 유닛 (52) 은 클라이언트 디바이스 (40) 의 코딩 및 렌더링 능력들에 의해 충족될 수 있는 특성들을 갖는 표현들 (68) 의 서브세트 (예를 들어, 적응 세트) 를 선택할 수도 있다. 추출 유닛 (52) 은 적응 세트에서의 표현들에 대한 비트레이트들을 결정하고, 네트워크 대역폭의 현재 이용가능한 양을 결정하고, 네트워크 대역폭에 의해 충족될 수 있는 비트레이트를 갖는 표현들 중 하나로부터 세그먼트들을 추출할 수도 있다.

[0054] 일반적으로, 보다 높은 비트레이트 표현들은 보다 고품질 비디오 플레이백을 산출할 수도 있는 한편, 보다 낮은 비트레이트 표현들은 이용가능한 네트워크 대역폭이 감소할 때 충분한 품질의 비디오 플레이백을 제공할 수도 있다. 이에 따라, 이용가능한 네트워크 대역폭이 상대적으로 높을 때, 추출 유닛 (52) 은 상대적으로 높은 비트레이트 표현들로부터 데이터를 추출할 수도 있는 반면, 이용가능한 네트워크 대역폭이 낮을 때, 추출 유닛 (52) 은 상대적으로 낮은 비트레이트 표현들로부터 데이터를 추출할 수도 있다. 이 방식으로, 클라이언트 디바이스 (40) 는 네트워크 (74) 의 네트워크 대역폭 이용가능성을 변화시키는 것에 또한 적응하면서 네트워크 (74) 를 통해 멀티미디어 데이터를 스트리밍할 수도 있다.

[0055] 추가로 또는 대안으로, 추출 유닛 (52) 은 eMBMS 또는 IP 멀티캐스트와 같은 브로드캐스트 또는 멀티캐스트 네트워크 프로토콜에 따라 데이터를 수신하도록 구성될 수도 있다. 이러한 예들에서, 추출 유닛 (52) 은 특정한 미디어 콘텐츠와 연관된 멀티캐스트 네트워크 그룹에 가입하기 위한 요청을 제출할 수도 있다. 멀티캐스트 그룹에 가입한 후, 추출 유닛 (52) 은 서버 디바이스 (60) 또는 콘텐츠 준비 디바이스 (20) 에 이슈된 추가 요청들 없이 멀티캐스트 그룹의 데이터를 수신할 수도 있다. 추출 유닛 (52) 은, 예를 들어, 플레이백을 중단하거나 또는 채널들을 상이한 멀티캐스트 그룹으로 변화시키기 위해, 멀티캐스트 그룹의 데이터가 더 이상 필요하지 않을 때 멀티캐스트 그룹을 떠나기 위한 요청을 제출할 수도 있다.

[0056] 네트워크 인터페이스 (54) 는 선택된 표현의 세그먼트들의 데이터를 추출 유닛 (52) 에 수신 및 제공할 수도 있고, 그 추출 유닛은 차례로 세그먼트들을 디캡슐화 유닛 (50) 에 제공할 수도 있다. 디캡슐화 유닛 (50) 은, 비디오 파일의 엘리먼트들을 구성 PES 스트림들로 디캡슐화하고, PES 스트림들을 디패킷화하여 인코딩된 데이터를 추출하고, 그리고 예를 들어, 스트림의 PES 패킷 헤더들에 의해 표시한 바와 같이, 인코딩된 데이터가 오디오 또는 비디오 스트림의 파트인지 여부에 의존하여, 인코딩된 데이터를 오디오 디코더 (46) 또는 비디오 디코더 (48) 중 어느 하나로 전송할 수도 있다. 오디오 디코더 (46) 는 인코딩된 오디오 데이터를 디코딩하고 디코딩된 오디오 데이터를 오디오 출력 (42) 으로 전송하는 한편, 비디오 디코더 (48) 는 인코딩된 비디오 데이터를 디코딩하고 스트림의 복수의 뷰들을 포함할 수도 있는 디코딩된 비디오 데이터를 비디오 출력 (44) 으로 전송한다.

[0057] 비디오 인코더 (28), 비디오 디코더 (48), 오디오 인코더 (26), 오디오 디코더 (46), 캡슐화 유닛 (30), 추출 유닛 (52), 및 디캡슐화 유닛 (50) 각각은 하나 이상의 마이크로프로세서들, 디지털 신호 프로세서들 (DSP들), 주문형 집적 회로들 (ASIC들), 필드 프로그래밍가능 게이트 어레이들 (FPGA들), 이산 로직 회로부, 소프트웨어, 하드웨어, 펌웨어 또는 그 임의의 조합들과 같이, 적용가능한 대로, 다양한 적합한 프로세싱 회로부 중 임의의 것으로서 구현될 수도 있다. 비디오 인코더 (28) 및 비디오 디코더 (48) 의 각각은, 어느 하나가 결합된 비디오 인코더/디코더 (CODEC) 의 파트로서 통합될 수도 있는, 하나 이상의 인코더들 또는 디코더들에 포함될 수도 있다. 마찬가지로, 오디오 인코더 (26) 및 오디오 디코더 (46) 의 각각은, 어느 하나가 결합된 CODEC 의 파트로서 통합될 수도 있는, 하나 이상의 인코더들 또는 디코더들에 포함될 수도 있다. 비디오 인코더 (28), 비디오 디코더 (48), 오디오 인코더 (26), 오디오 디코더 (46), 캡슐화 유닛 (30), 추출 유닛 (52), 및/또는 디캡슐화 유닛 (50) 을 포함하는 장치는 집적 회로, 마이크로프로세서, 및/또는 무선 통신 디바이스, 이를테면 셀룰러 전화기를 포함할 수도 있다.

[0058] 비디오 코딩 표준들은 ITU-T H.261, ISO/IEC MPEG-1 Visual, ITU-T H.262 또는 ISO/IEC MPEG-2 Visual, ITU-T H.263, ISO/IEC MPEG-4 Visual, ITU-T H.264 또는 ISO/IEC MPEG-4 AVC (그 스케일러블 비디오 코딩 (SVC) 및 멀티뷰 비디오 코딩 (MVC) 확장들을 포함함), 및 ITU-T H.265 및 ISO/IEC 23008-2 로도 또한 알려진, 고효율 비디오 코딩 (HEVC) (그 스케일러블 코딩 확장 (즉, 스케일러블 고효율 비디오 코딩, SHVC), 멀티뷰 확장 (즉, 멀티뷰 고효율 비디오 코딩, MV-HEVC), 및 3D 확장 (즉, 3D 고효율 비디오 코딩, 3D-HEVC) 을 포함함) 을 포함한다. MVC 의 하나의 공동 초안은 "Advanced video coding for generic audiovisual services", ITU-T Recommendation H.264, March, 2010 에서 설명된다.

- [0059] 캡슐화 유닛 (30) 은 이미지의 하나 이상의 가장 관심있는 영역들을 결정할 수도 있다. 본 명세서에서 사용한 바와 같이, 가장 관심있는 영역은 사용자에게 의해 뷰잉될 가능성이 높은 이미지의 영역을 지칭할 수도 있다. 예를 들어, 가장 관심있는 영역은 디렉터 또는 프로듀서에 의해 선택되는 이미지의 영역을 지칭할 수도 있다. 일부 예들에서, 가장 관심있는 영역은, 서비스 또는 콘텐츠 제공자에 의한 사용자 통계로부터, 예를 들어, 전방향 비디오 콘텐츠가 스트리밍 서비스를 통하여 제공되었을 때 사용자들이 어느 영역들을 가장 많이 요청하였는지/보았는지의 통계를 통하여 컴퓨팅 디바이스에 의해 도출되는 이미지의 영역을 지칭할 수도 있다.
- [0060] 캡슐화 유닛 (30) 은 하나 이상의 가장 관심있는 영역들을 나타내는 데이터를 캡슐화할 수도 있다. 일부 예들에서, 캡슐화 유닛 (30) 은 비디오 코딩 엔진에 의해 이용되는 실제 비디오 데이터를 포함할 수도 있는 VCL 데이터의 외부에 있는 하나 이상의 가장 관심있는 영역들을 나타내는 데이터를 캡슐화할 수도 있다. 일부 예들에서, 하나 이상의 가장 관심있는 영역들을 나타내는 데이터는 비트스트림에 포함될 수도 있다. 예를 들어, 캡슐화 유닛 (30) 은 VCL 데이터의 외부에 있는 하나 이상의 가장 관심있는 영역들을 나타내는 데이터를 포함하는 SEI 메시지를 캡슐화할 수도 있다. 일부 예들에서, 캡슐화 유닛 (30) 은 비디오 비트스트림의 외부에 있는 (그리고 결과적으로 또한 비트스트림의 VCL 데이터의 외부에 있는) 하나 이상의 가장 관심있는 영역들을 나타내는 데이터를 포함하는 파일 포맷 데이터를 캡슐화할 수도 있다. 예를 들어, 캡슐화 유닛 (30) 은 VCL 데이터의 외부에 있는 하나 이상의 가장 관심있는 영역들을 나타내는 데이터를 포함하는 박스 (예를 들어, 샘플-투-그룹) 또는 타이밍된 메타데이터 트랙을 캡슐화할 수도 있다.
- [0061] VR 비디오 애플리케이션들에 대하여, 가장 관심있는 영역들에 대한 정보를 활용하는 방식들을 고려할 때, 다음의 문제들 중 적어도 일부를 식별한다:
- [0062] 1. 사용자 제어 없이 가장 관심있는 영역들을 디스플레이하기 위해 가장 관심있는 영역들에 대한 정보를 이용하는 것은, 사용자가 실제로는 VR 비디오에서의 다른 영역들에 관심이 있다면 VR 비디오 경험을 낮출 수도 있다. 대신에, 가장 관심있는 영역들에 대한 정보는 다른 방식으로 향상된 VR 비디오 경험을 위해 이용될 수도 있다.
- [0063] 2. VR 비디오의 픽처들의 특정한 세트에서의 가장 관심있는 영역들은 디렉터의 컷과는 다른 방식들에 의해 결정 또는 생성될 수도 있다.
- [0064] 3. 가장 관심있는 영역들에 대한 정보를 시그널링하는 구체적인 방식이 없다.
- [0065] 본 개시는 VR 비디오에서의 하나 이상의 가장 관심있는 영역의 시그널링 및 재구성 및 VR 비디오 인코딩, 디코딩, 스트리밍, 및 플레이백에서의 대응하는 프로세싱을 위한 여러 기법들을 설명한다. 이들 기법들 중 하나 이상은 독립적으로, 또는 다른 것들과 조합하여 적용될 수도 있다.
- [0066] 제 1 예에서, 가장 관심있는 영역들은 디렉터 또는 프로듀서의 의도에 의해, 또는 서비스 또는 콘텐츠 제공자에 의한 사용자 통계에 의해, 예를 들어, VR 비디오 콘텐츠가 스트리밍 서비스를 통하여 제공되었을 때 사용자들이 어느 영역들을 가장 많이 요청하였는지/보았는지를 통하여 결정될 수도 있다.
- [0067] 제 2 예에서, 비디오 인코딩 또는 트랜스코딩 동안, 가장 관심있는 영역들은 (예를 들어, 보다 낮은 양자화 파라미터들을 이용하는 것에 의해) 보다 고품질, (비디오 픽처들의 상이한 영역들이 비디오 스티칭 또는 비디오 인코딩 동안 상이한 서브-샘플링 비율들로 서브-샘플링되면) 보다 높은 레졸루션, 또는 다른 영역들보다 더 높은 빈도의 랜덤 액세스가능성으로 의도적으로 인코딩될 수도 있다.
- [0068] 제 3 예에서, 본 개시는 비디오 비트스트림에서의 SEI 메시지들, 미디어 파일에서의 파일 포맷 메타데이터, 또는 DASH (dynamic adaptive streaming over HTTP) 미디어 프리젠테이션 디스크립션 (MPD) 엘리먼트들 또는 속성을 이용하는 가장 관심있는 영역들에 대한 정보의 시그널링을 제안한다.
- [0069] VR 비디오 비트스트림의 픽처들의 특정한 세트에 대한 가장 관심있는 영역들에 대한 정보는 SEI 메시지를 이용하여 시그널링될 수도 있다. 이 SEI 메시지의 인스턴스는 신택스, 시맨틱스 또는 양자 모두에 의해 (하나 이상의) 픽처들의 세트에 적용가능한 것으로 특정된다. SEI 메시지는 적어도 다음의 신택스 엘리먼트들을 포함할 수도 있다:
- [0070] i. 가장 관심있는 영역들의 수를 표시하는 신택스 엘리먼트.
- [0071] ii. 각각의 세트가 가장 관심있는 영역의 포지션 및 사이즈를 설명하는, 신택스 엘리먼트들의 세트의 루프. 예를 들어, 각각의 세트는 가장 관심있는 영역인 직사각형의 4 개의 코너들의, 풀 픽처에 대한 루마 샘플들의 수에서의, 상부 (top), 하부 (bottom), 왼쪽 (left) 및 오른쪽 (right) 오프셋들을 표시하는 4 개의 신택스 엘

리먼트를 포함할 수도 있다. 다른 예에서, 각각의 세트는 가장 관심있는 영역인 직사각형의 좌상측 (upper-left) 코너의, 풀 픽처에 대한 루마 샘플들의 수에서의, 좌표를 표시하는 2 개의 선택스 엘리먼트, 및 직사각형의 폭 및 높이를 표시하는 2 개의 선택스 엘리먼트들을 포함할 수도 있다.

- [0072] iii. 가장 관심있는 영역의 각각에 대해, 우선순위가 선택스 엘리먼트에 의해 표시될 수도 있고, 여기서 높은 우선순위는 영역의 보다 높은 관심을 표시한다.
- [0073] iv. 디코딩 순서에서 이 SEI 메시지의 모든 선행 인스턴스들의 효과를 캔슬할지 여부를 표시하는 선택스 엘리먼트.
- [0074] v. 이 SEI 메시지의 현재 인스턴스들에서의 정보가 출력 순서에서의 다음 픽처, 또는 새로운 코딩된 비디오 시퀀스의 시작, 또는 비트스트림의 종료까지 지속되는지 여부를 표시하는 선택스 엘리먼트.
- [0075] SEI 메시지는 H.264/AVC, H.265/HEVC, 또는 임의의 다른 비디오 코덱 규격의 파트로서 포함될 수 있다. SEI 메시지를 포함하는 NAL 유닛은 1) 파일 포맷에서의 정보의 시그널링을 위한 파일 포맷 박스 또는 구조, 2) SDP 에서의 정보의 시그널링을 위한 SDP 필드, 3) DASH MPD 에서의 정보의 시그널링을 위한 DASH MPD 속성 또는 엘리먼트, 및 4) MPEG-2 TS 에서의 정보의 시그널링을 위한 MPEG-2 전송 스트림 (TS) 디스크립터에 직접 포함될 수 있다.
- [0076] 상기 설명한 바와 동일한 정보는, 예를 들어, 새로운 샘플 그룹을 이용하는 것에 의해, 또는 타이밍된 메타데이터 트랙에의 포함에 의해, 파일 포맷 데이터 구조로 직접 반송될 수도 있다.
- [0077] 상기 설명한 바와 동일한 정보는 DASH MPD 에서의 일부 새로운 엘리먼트들, 속성들, 및/또는 디스크립터를 이용하여 직접 반송될 수도 있다.
- [0078] 제 4 예에서, DASH 와 같은 클라이언트-제어된 비디오 스트리밍 동안, 콘텐츠 전달 네트워크 (CDN) 가 사용중일 때, 예지 서버 (예를 들어, 서버 디바이스 (60)) 또는 캐시는 가장 관심있는 영역들에 대한 정보를 이용하여 보다 고품질 표현의 비디오 데이터를 획득하고 클라이언트 디바이스 (40) 로부터 요청을 수신하기 전에 약간 업커밍 이미지의 가장 관심있는 영역들을 커버하여, 비디오 데이터가 곧 요청될 가능성이 가장 클 것이라고 예측할 수도 있다. 업커밍 이미지는 하나 이상의 영역들이 클라이언트 디바이스 (40) 로 최근에 전송되었고 클라이언트 디바이스 (40) 에 의해 제시되고 있거나 또는 곧 제시될 이미지에 바로 후속하는 이미지일 수도 있다.
- [0079] 제 5 예에서, DASH 와 같은 클라이언트-제어된 비디오 스트리밍 동안, 클라이언트 디바이스 (40) 는 가장 관심있는 영역들에 대한 정보를 이용하여 보다 고품질 표현의 비디오 데이터를 요청하고 사용자가 그/그녀의 머리를 돌리고 요청들을 전송하기 전에 약간 업커밍 이미지의 가장 관심있는 영역들을 커버하여, 비디오 데이터가 곧 요청될 가능성이 가장 클 것이라고 예측할 수도 있다. 다시, 업커밍 이미지는 하나 이상의 영역들이 클라이언트 디바이스 (40) 로 최근에 전송되었고 클라이언트 디바이스 (40) 에 의해 제시되고 있거나 또는 곧 제시될 이미지에 바로 후속하는 이미지일 수도 있다.
- [0080] 제 6 예에서, 서버-제어된 비디오 스트리밍 시스템 (예를 들어, 서버 디바이스 (60)), 이를 태면 3GPP TS 26.234 에서의 3GPP 패킷-스위칭된 스트리밍 (PSS) 은, 서버가 전송될 가장 관심있는 영역들을 커버하는 적절한 비디오 데이터를 선정하도록 현재 FOV 의 방향에 대한 피드백을 스트리밍 클라이언트들 (예를 들어, 클라이언트 디바이스 (40)) 가 전송하는 것을 가능하게 하는 피쳐들로 구성될 수도 있다. 이러한 구성된 서버-제어된 비디오 스트리밍 시스템에서, 서버 디바이스 (60) 는 가장 관심있는 영역들에 대한 정보를 이용하여 보다 고품질 표현의 비디오 데이터를 전송하고 현재 FOV 의 방향에 대한 스트리밍 클라이언트의 피드백을 수신하기 전에 약간 업커밍 이미지의 가장 관심있는 영역들을 커버하여, 비디오 데이터가 곧 전송될 가능성이 가장 클 것이라고 예측할 수도 있다. 예를 들어, 특정한 스트리밍 클라이언트가 특정한 이미지를 출력하고 있다고 결정하는 것에 응답하여, 서버 디바이스 (60) 는 특정한 스트리밍 클라이언트에, 보다 고품질 표현의 비디오 데이터를 출력하고 특정한 이미지에 바로 후속하는 업커밍 이미지의 가장 관심있는 영역들을 커버할 수도 있다.
- [0081] 제 7 예에서, (예를 들어, 서버 디바이스 (60) 에서) 캐시를 수반하는 콘텐츠 (예를 들어, VR 콘텐츠) 의 저장 관리를 위해, 서버 디바이스 (60) 는 특정한 이미지에서의 하나 이상의 가장 관심있는 영역들을 나타내는 데이터를 이용하여 캐시에 저장할 특정한 이미지의 영역들을 결정할 수도 있다. 개발자는 저장 선택 알고리즘을 설계할 수도 있고 서버 디바이스 (60) 는 알고리즘을 실행하여 영역들에서의 선택된 관심 정도에 기초하여 (예를 들어, 디렉터의 의견에 기초하여) 및/또는 영역에서의 도출된 관심 정도에 기초하여 (예를 들어, 콘텐츠가 한동안 이용가능한 후 거동을 보는 사용자의 통계에 기초하여) 이미지의 영역들에 랭크를 부여할 수도 있다. 서버 디바이스 (60) 는, 캐시에, 임계치를 충족하는 (예를 들어, 임계치보다 큰) 랭크 값을 가진 특정한 이미

지의 영역들에 대한 비디오 데이터를 저장할 수도 있다. 예를 들어, 단지 특정한 이미지의 가장 관심있는 영역들에 대한 랭크 값들이 임계치를 충족할 때, 서버 디바이스 (60) 는 캐시에 특정한 이미지의 가장 관심있는 영역들에 대한 비디오 데이터를 단지 저장할 수도 있다. 이 예에서, 클라이언트 디바이스 (40) 는 캐시로부터 특정한 이미지의 가장 관심있는 영역들에 대한 비디오 데이터를 추출할 수도 있고 캐시보다 사용자들로부터 더 멀리 포지셔닝될 수도 있는 메인 서버 스토리지로부터 특정한 이미지의 다른 영역들에 대한 비디오 데이터를 추출할 수도 있다. 추가로, 또는 대안으로, 서버 디바이스 (60) 는 캐시에서의 이용가능한 공간의 양에 기초하여 임계치를 동적으로 결정할 수도 있다.

[0082] 제 8 예에서, VR 콘텐츠의 장기 저장을 위해, 서버 디바이스 (60) 는 (예를 들어, VR 비디오의) 특정한 이미지에서의 하나 이상의 가장 관심있는 영역들을 나타내는 데이터를 이용하여 저장할 이미지의 영역들을 결정할 수도 있다. 개발자는 저장 선택 알고리즘을 설계할 수도 있고 서버 디바이스 (60) 는 알고리즘을 실행하여 영역들에서의 선택된 관심 정도에 기초하여 (예를 들어, 디렉터의 의견에 기초하여) 및/또는 영역에서의 도출된 관심 정도에 기초하여 (예를 들어, 콘텐츠가 한동안 이용가능한 후 거동을 보는 사용자의 통계에 기초하여) 영역들에 랭크를 부여할 수도 있다. 서버 디바이스 (60) 는 (예를 들어, 수주 (weeks), 수월 (months), 수년 (years) 등 동안 비디오 데이터를 저장하도록 구성된) 장기 스토리지에, 임계치를 충족하는 (예를 들어, 임계치보다 큰) 랭크 값을 가진 특정한 이미지의 영역들에 대한 비디오 데이터를 저장할 수도 있다. 예를 들어, 단지 특정한 이미지의 가장 관심있는 영역들에 대한 랭크 값들이 임계치를 충족할 때, 서버 디바이스 (60) 는 장기 스토리지에 특정한 이미지의 가장 관심있는 영역들에 대한 비디오 데이터를 단지 저장하고 장기 스토리지로부터 특정한 이미지의 다른 영역들에 대한 비디오 데이터를 제거할 수도 있다.

[0083] 클라이언트 디바이스 (40), 서버 디바이스 (60), 및/또는 콘텐츠 준비 디바이스 (20) 는 본 개시의 기법들에 따라 동작하도록 구성될 수도 있다. 예의 목적들을 위해, 본 개시는 클라이언트 디바이스 (40) 및 서버 디바이스 (60) 에 대하여 이들 기법들을 설명한다. 그러나, 콘텐츠 준비 디바이스 (20) 는 서버 디바이스 (60) 대신에 (또는 그에 더하여) 이들 기법들을 수행하도록 구성될 수도 있는 것으로 이해되어야 한다.

[0084] 캡슐화 유닛 (30) 은 NAL 유닛이 속하는 프로그램을 식별하는 헤더, 뿐만 아니라 페이로드, 예를 들어, 오디오 데이터, 비디오 데이터, 또는 NAL 유닛이 대응하는 전송 또는 프로그램 스트림을 설명하는 데이터를 포함하는 NAL 유닛들을 형성할 수도 있다. 예를 들어, H.264/AVC 에서, NAL 유닛은 1-바이트 헤더 및 가변 사이즈의 페이로드를 포함한다. 그 페이로드에 비디오 데이터를 포함하는 NAL 유닛은 비디오 데이터의 다양한 입도 레벨들을 포함할 수도 있다. 예를 들어, NAL 유닛은 비디오 데이터의 블록, 복수의 블록들, 비디오 데이터의 슬라이스, 또는 비디오 데이터의 전체 픽처를 포함할 수도 있다. 캡슐화 유닛 (30) 은 기본 스트림들의 PES 패킷들의 형태로 비디오 인코더 (28) 로부터 인코딩된 비디오 데이터를 수신할 수도 있다. 캡슐화 유닛 (30) 은 각각의 기본 스트림을 대응하는 프로그램과 연관시킬 수도 있다.

[0085] 캡슐화 유닛 (30) 은 또한, 복수의 NAL 유닛들로부터 액세스 유닛들을 어셈블링할 수도 있다. 일반적으로, 액세스 유닛은 비디오 데이터의 프레임, 뿐만 아니라 오디오 데이터가 이용가능할 때 그 프레임에 대응하는 이러한 오디오 데이터를 표현하기 위한 하나 이상의 NAL 유닛들을 포함할 수도 있다. 액세스 유닛은 일반적으로 하나의 출력 시간 인스턴스에 대해 모든 NAL 유닛들, 예를 들어 하나의 시간 인스턴스에 대해 모든 오디오 및 비디오 데이터를 포함한다. 예를 들어, 각각의 뷰가 20 fps (frames per second) 의 프레임 레이트를 가지면, 각각의 시간 인스턴스는 0.05 초의 시간 간격에 대응할 수도 있다. 이 시간 간격 동안, 동일한 액세스 유닛 (동일한 시간 인스턴스) 의 모든 뷰들에 대한 특정 프레임들이 동시에 렌더링될 수도 있다. 하나의 예에서, 액세스 유닛은 프라이머리 코딩된 픽처로서 제시될 수도 있는 하나의 시간 인스턴스에서 코딩된 픽처를 포함할 수도 있다.

[0086] 이에 따라, 액세스 유닛은 공통 시간 인스턴스의 모든 오디오 및 비디오 프레임들, 예를 들어, 시간 X 에 대응하는 모든 뷰들을 포함할 수도 있다. 본 개시는 또한, "뷰 컴포넌트" 로서 특정한 뷰의 인코딩된 픽처와 관련된다. 즉, 뷰 컴포넌트는 특정한 시간에 특정한 뷰에 대한 인코딩된 픽처 (또는 프레임) 를 포함할 수도 있다. 이에 따라, 액세스 유닛은 공통 시간 인스턴스의 모든 뷰 컴포넌트들을 포함하는 것으로서 정의될 수도 있다. 액세스 유닛들의 디코딩 순서는 출력 또는 디스플레이 순서와 반드시 동일할 필요는 없다.

[0087] 미디어 프리젠테이션은 상이한 대안의 표현들 (예를 들어, 상이한 품질들을 가진 비디오 서비스들) 의 디스크립션들을 포함할 수도 있는 미디어 프리젠테이션 디스크립션 (MPD) 을 포함할 수도 있고 디스크립션은, 예를 들어, 코덱 정보, 프로파일 값, 및 레벨 값을 포함할 수도 있다. MPD 는 매니페스트 파일, 이를 테면 매니페스트 파일 (66) 의 하나의 예이다. 클라이언트 디바이스 (40) 는 미디어 프리젠테이션의 MPD 를 추출하여

다양한 프리젠테이션들의 무비 프래그먼트들에 액세스하는 방법을 결정할 수도 있다. 무비 프래그먼트들은 비디오 파일들의 무비 프래그먼트 박스들 (moof 박스들) 에 로케이트될 수도 있다.

[0088] 매니페스트 파일 (66) (이는 예를 들어, MPD 를 포함할 수도 있음) 은 표현들 (68) 의 세그먼트들의 이용가능성을 광고할 수도 있다. 즉, MPD 는 표현들 (68) 중 하나의 제 1 프래그먼트가 이용가능해지는 월-클록 (wall-clock) 시간을 표시하는 정보, 뿐만 아니라 표현들 (68) 내의 세그먼트들의 지속기간들을 표시하는 정보를 포함할 수도 있다. 이 방식으로, 클라이언트 디바이스 (40) 의 취출 유닛 (52) 은, 시작 시간 뿐만 아니라 특정한 세그먼트에 선행하는 세그먼트들의 지속기간들에 기초하여, 각각의 세그먼트가 이용가능한 때를 결정할 수도 있다. 일부 예들에서, 매니페스트 파일 (66) 은 본 개시의 기법들에 대하여, 특정한 이미지에서의 하나 이상의 가장 관심있는 영역들을 나타내는 데이터를 포함할 수도 있다.

[0089] 캡슐화 유닛 (30) 이 수신된 데이터에 기초하여 비디오 파일로 NAL 유닛들 및/또는 액세스 유닛들을 어셈블링한 후, 캡슐화 유닛 (30) 은 비디오 파일을 출력을 위해 출력 인터페이스 (32) 로 전달한다. 일부 예들에서, 캡슐화 유닛 (30) 은 비디오 파일을 바로 클라이언트 디바이스 (40) 로 전송하기 보다는, 비디오 파일을 로컬로 저장하거나 또는 비디오 파일을 출력 인터페이스 (32) 를 통해 원격 서버로 전송할 수도 있다. 출력 인터페이스 (32) 는, 예를 들어, 송신기, 트랜시버, 컴퓨터 판독가능 매체에 데이터를 기입하기 위한 디바이스, 이를 테면, 예를 들어, 광학 드라이브, 자기 미디어 드라이브 (예를 들어, 플로피 드라이브), 범용 직렬 버스 (USB) 포트, 네트워크 인터페이스, 또는 다른 출력 인터페이스를 포함할 수도 있다. 출력 인터페이스 (32) 는 비디오 파일을 컴퓨터 판독가능 매체, 이를 테면, 예를 들어, 송신 신호, 자기 매체, 광학 매체, 메모리, 플래시 드라이브, 또는 다른 컴퓨터 판독가능 매체로 출력한다.

[0090] 네트워크 인터페이스 (54) 는 NAL 유닛 또는 액세스 유닛을 네트워크 (74) 를 통해 수신하고 NAL 유닛 또는 액세스 유닛을 취출 유닛 (52) 을 통해 디캡슐화 유닛 (50) 에 제공할 수도 있다. 디캡슐화 유닛 (50) 은, 비디오 파일의 엘리먼트들을 구성 PES 스트림들로 디캡슐화하고, PES 스트림들을 디패킷화하여 인코딩된 데이터를 취출하고, 그리고 예를 들어, 스트림의 PES 패킷 헤더들에 의해 표시한 바와 같이, 인코딩된 데이터가 오디오 또는 비디오 스트림의 파트인지 여부에 의존하여, 인코딩된 데이터를 오디오 디코더 (46) 또는 비디오 디코더 (48) 중 어느 하나로 전송할 수도 있다. 오디오 디코더 (46) 는 인코딩된 오디오 데이터를 디코딩하고 디코딩된 오디오 데이터를 오디오 출력 (42) 으로 전송하는 한편, 비디오 디코더 (48) 는 인코딩된 비디오 데이터를 디코딩하고 스트림의 복수의 뷰들을 포함할 수도 있는 디코딩된 비디오 데이터를 비디오 출력 (44) 으로 전송한다.

[0091] 이 방식으로, 콘텐츠 준비 디바이스 (20) 는 비디오 데이터에 대한 정보를 시그널링하기 위한 디바이스의 예를 표현하고, 그 디바이스는 비디오 데이터의 이미지의 복수의 영역들 중 하나 이상의 가장 관심있는 영역들을 결정하도록 구성된 프로세서를 포함한다. 가장 관심있는 영역들은 클라이언트 디바이스에 의해 취출될 가능성이 가장 큰 하나 이상의 영역들을 포함한다. 프로세서는 가장 관심있는 영역들을 나타내는 데이터를 생성하도록 추가로 구성될 수도 있다. 가장 관심있는 영역들을 나타내는 데이터는 비디오 데이터를 포함하는 비트스트림의 VCL 데이터의 외부에 있다.

[0092] 추가적으로, 클라이언트 디바이스 (40) 는 비디오 데이터에 대한 정보를 결정하기 위한 디바이스의 예를 표현하고, 그 디바이스는 비디오 데이터의 이미지의 복수의 영역들 중 하나 이상의 가장 관심있는 영역들을, 가장 관심있는 영역들을 나타내는 데이터에 기초하여 결정하도록 구성된 프로세서를 포함한다. 가장 관심있는 영역들을 나타내는 데이터는 비디오 데이터를 포함하는 비트스트림의 VCL 데이터의 외부에 있다. 프로세서는 이미지의 하나 이상의 가장 관심있는 영역들에 대한 고품질 비디오 데이터에 대한 요청을 출력하고, 이미지의 하나 이상의 가장 관심있는 영역들에 대한 고품질 비디오 데이터에 대한 요청을 출력한 후, 이미지의 하나 이상의 다른 영역들에 대한 저품질 비디오 데이터에 대한 요청을 출력하도록 추가로 구성될 수도 있다.

[0093] 도 2 는 예의 멀티미디어 콘텐츠 (120) 의 엘리먼트들을 예시하는 개념적 다이어그램이다. 멀티미디어 콘텐츠 (120) 는 멀티미디어 콘텐츠 (64) (도 1), 또는 저장 매체 (62) 에 저장된 다른 멀티미디어 콘텐츠에 대응할 수도 있다. 도 2 의 예에서, 멀티미디어 콘텐츠 (120) 는 미디어 프리젠테이션 디스크립션 (MPD) (122) 및 복수의 표현들 (124A 내지 124N) (표현들 (124)) 을 포함한다. 표현 (124A) 은 옵션의 헤더 데이터 (126) 및 세그먼트들 (128A 내지 128N) (세그먼트들 (128)) 을 포함하는 한편, 표현 (124N) 은 옵션의 헤더 데이터 (130) 및 세그먼트들 (132A 내지 132N) (세그먼트들 (132)) 을 포함한다. 문자 N 은 편의상 표현들 (124) 의 각각에서 마지막 무비 프래그먼트를 지정하는데 사용된다. 일부 예들에서, 표현들 (124) 간에는 상이한 수들의 무비 프래그먼트들이 존재할 수도 있다.

- [0094] 도 2 는 시스템 (10) (도 1) 이 DASH 를 위해 구성되는 예를 설명하지만, 일부 예들에서, 시스템 (10) 은 다른 미디어 플레이백 제어 프로토콜들, 예를 들어, 실시간 스트리밍 프로토콜 (RTSP), 실시간 전송 프로토콜 (RTP), RTP 제어 프로토콜 (RTCP), 세션 디스크립션 프로토콜 (SDP), 다른 미디어 플레이백 제어 프로토콜, 또는 그 조합 (그러나 이들에 제한되지는 않음) 을 위해 구성될 수도 있다.
- [0095] MPD (122) 는 표현들 (124) 과는 별개의 데이터 구조를 포함할 수도 있다. MPD (122) 는 도 1 의 매니페스트 파일 (66) 에 대응할 수도 있다. 마찬가지로, 표현들 (124) 은 도 1 의 표현들 (68) 에 대응할 수도 있다. 일반적으로, MPD (122) 는 표현들 (124) 의 특성들, 이를 테면 코딩 및 렌더링 특성들, 적응 세트들, MPD (122) 가 대응하는 프로파일, 텍스트 타입 정보, 카메라 각도 정보, 레이팅 정보, 트릭 모드 정보 (예를 들어, 시간 서브-시퀀스들을 포함하는 표현들을 나타내는 정보), 및/또는 (예를 들어, 플레이백 동안 미디어 콘텐츠로의 타겟팅된 광고 삽입을 위한) 원격 주기들을 추출하기 위한 정보를 일반적으로 설명하는 데이터를 포함할 수도 있다.
- [0096] 헤더 데이터 (126) 는, 존재 시, 세그먼트들 (128) 의 특성들, 예를 들어, 세그먼트들 (128) 이 랜덤 액세스 포인트들을 포함하는, 랜덤 액세스 포인트들 (RAP들, 스트림 액세스 포인트들 (SAP들) 로도 또한 지칭됨) 의 시간 로케이션들, 세그먼트들 (128) 내의 랜덤 액세스 포인트들에 대한 바이트 오프셋들, 세그먼트들 (128) 의 유니폼 리소스 로케이터들 (URL들), 또는 세그먼트들 (128) 의 다른 양태들을 설명할 수도 있다. 헤더 데이터 (130) 는, 존재 시, 세그먼트들 (132) 에 대한 유사한 특성들을 설명할 수도 있다. 추가로 또는 대안으로, 이러한 특성들은 MPD (122) 내에 완전히 포함될 수도 있다.
- [0097] 세그먼트들 (128, 132) 은, 각각이 비디오 데이터의 프레임들 또는 슬라이스들을 포함할 수도 있는, 하나 이상의 코딩된 비디오 샘플들을 포함한다. 세그먼트들 (128) 의 코딩된 비디오 샘플들의 각각은 유사한 특성들, 예를 들어, 높이, 폭, 및 대역폭 요건들을 가질 수도 있다. 이러한 특성들은 MPD (122) 의 데이터에 의해 설명될 수도 있지만, 이러한 데이터는 도 2 의 예에 예시되지 않는다. MPD (122) 는, 본 개시에서 설명된 시그널링된 정보 중 임의의 것 또는 전부의 추가와 함께, 3GPP 규격에 의해 설명한 바와 같은 특성들을 포함할 수도 있다.
- [0098] 세그먼트들 (128, 132) 의 각각은 고유한 유니폼 리소스 로케이터 (URL) 와 연관될 수도 있다. 따라서, 세그먼트들 (128, 132) 의 각각은 DASH 와 같은 스트리밍 네트워크 프로토콜을 이용하여 독립적으로 추출가능할 수도 있지만, 다른 스트리밍 네트워크 프로토콜들이 이용될 수도 있다. 이 방식으로, 목적지 디바이스, 이를 테면 클라이언트 디바이스 (40) 는 HTTP GET 요청을 이용하여 세그먼트들 (128 또는 132) 을 추출할 수도 있다. 일부 예들에서, 클라이언트 디바이스 (40) 는 HTTP 부분 GET 요청들을 이용하여 특정 바이트 범위들의 세그먼트들 (128 또는 132) 을 추출할 수도 있다. 예를 들어, 클라이언트 디바이스 (40) 는 제 1 HTTP 부분 GET 요청을 이용하여 세그먼트들 (128, 132) 중 하나로부터 가장 관심있는 영역들을 나타내는 데이터를 초기에 추출하고, 가장 관심있는 영역들을 결정하고, 그 후 제 2 HTTP 부분 GET 요청을 이용하여 세그먼트들 (128, 132) 중 하나로부터 가장 관심있는 영역들에 대한 비디오 데이터의 고품질 버전들을 프리-페치하여, 예를 들어, 파일 포맷 정보에서, 가장 관심있는 영역들 데이터가 세그먼트들 (128, 132) 내에 포함됨을 가정할 수도 있다.
- [0099] 추가로 또는 대안으로, 본 개시의 기법들에 따르면, MPD (122) 는 하나 이상의 가장 관심있는 영역들을 표시하는 엘리먼트들 또는 속성들을 포함할 수도 있다. 예를 들어, MPD (122) 는, 세그먼트들 (128, 132) 의 각각에 대해, 가장 관심있는 영역의 개별의 포지션 및 가장 관심있는 영역의 개별의 사이즈를 나타내는 엘리먼트들 또는 속성들을 포함할 수도 있다.
- [0100] 보다 구체적으로는, 예를 들어, MPD (122) 는 세그먼트들 (128, 132) 의 각각 (또는 세그먼트들 (128, 132) 내의 하나 이상의 픽처들) 에 대해, 가장 관심있는 영역에 대해 루프하는 엘리먼트들 또는 속성들을 포함할 수도 있고, 그 루프의 각각의 반복은 가장 관심있는 영역들 중 하나를 나타내고 가장 관심있는 영역의 포지션 및 가장 관심있는 영역의 사이즈를 나타내는 선택스 엘리먼트들을 포함한다. 일부 예들에서, MPD (122) 는, 가장 관심있는 영역들의 각각에 대해, 가장 관심있는 영역의 상부를 나타내는 제 1 선택스 엘리먼트, 가장 관심있는 영역의 하부를 나타내는 제 2 선택스 엘리먼트, 가장 관심있는 영역의 좌측, 및 가장 관심있는 영역의 우측을 포함하는 엘리먼트들 또는 속성들을 세그먼트들 (128, 132) 의 각각에 대해 포함할 수도 있다. 일부 예들에서, MPD (122) 는, 가장 관심있는 영역들의 각각에 대해, 가장 관심있는 영역의 좌상측 코너를 나타내는 제 1 선택스 엘리먼트 및 가장 관심있는 영역의 우하측 (lower-right) 코너를 나타내는 제 2 선택스 엘리먼트를 포함하는 엘리먼트들 또는 속성들을 세그먼트들 (128, 132) 의 각각에 대해 포함할 수도 있다. 일부 예들에서, MPD (122) 는, 가장 관심있는 영역들의 각각에 대해, 가장 관심있는 영역의 좌상측 코너를 나타내는 제 1 선택

스 엘리먼트, 가장 관심있는 영역의 폭을 나타내는 제 2 선택스 엘리먼트, 및 가장 관심있는 영역의 높이를 나타내는 제 3 선택스 엘리먼트를 포함하는 엘리먼트들 또는 속성들을 세그먼트들 (128, 132) 의 각각에 대해 포함할 수도 있다. 일부 예들에서, MPD (122) 는 루마 샘플들에서 가장 관심있는 영역의 개별의 포지션 및 가장 관심있는 영역의 개별의 사이즈를 표현할 수도 있다.

[0101] 도 3 은 도 2 의 세그먼트들 (128, 132) 중 하나와 같은 표현의 세그먼트에 대응할 수도 있는 예의 비디오 파일 (150) 의 엘리먼트들을 예시하는 블록 다이어그램이다. 세그먼트들 (128, 132) 의 각각은 실질적으로 도 3 의 예에 예시된 데이터의 배열을 준수하는 데이터를 포함할 수도 있다. 비디오 파일 (150) 은 세그먼트를 캡슐화한다고 할 수도 있다. 상기 설명한 바와 같이, ISO 기반 미디어 파일 포맷 및 그 확장들에 따른 비디오 파일들은 "박스들" 로 지칭된 일련의 오브젝트들에 데이터를 저장한다. 도 3 의 예에서, 비디오 파일 (150) 은 파일 타입 (FTYP) 박스 (152), 무비 (MOOV) 박스 (154), 하나 이상의 쌍들의 무비 프래그먼트 (MOOF) 박스 (164) 및 미디어 데이터 (MDAT) 박스 (167), 및 무비 프래그먼트 랜덤 액세스 (MFRA) 박스 (166) 를 포함한다. 도 3 은 비디오 파일의 예를 표현하지만, 다른 미디어 파일들이 ISO 기반 미디어 파일 포맷 및 그 확장들에 따라 비디오 파일 (150) 의 데이터와 유사하게 구조화되는 다른 타입들의 미디어 데이터 (예를 들어, 오디오 데이터, 타이밍된 텍스트 데이터 등) 를 포함할 수도 있는 것으로 이해되어야 한다.

[0102] 파일 타입 (FTYP) 박스 (152) 는 일반적으로 비디오 파일 (150) 에 대한 파일 타입을 설명한다. 파일 타입 박스 (152) 는 비디오 파일 (150) 에 대한 최선의 사용을 설명하는 규격을 식별하는 데이터를 포함할 수도 있다. 파일 타입 박스 (152) 는 대안으로 MOOV 박스 (154), MOOF 박스 (164), 및/또는 MFRA 박스 (166) 랑에 배치될 수도 있다.

[0103] MOOV 박스 (154) 는, 도 3 의 예에서, 무비 헤더 (MVHD) 박스 (156) 및 트랙 (TRAK) 박스 (158) 를 포함한다. 일반적으로, MVHD 박스 (156) 는 비디오 파일 (150) 의 일반적인 특성들을 설명할 수도 있다. 예를 들어, MVHD 박스 (156) 는 비디오 파일 (150) 이 원래 생성되었을 때, 비디오 파일 (150) 이 가장 최근에 수정되었을 때, 비디오 파일 (150) 에 대한 타임스케일, 비디오 파일 (150) 에 대한 플레이어의 지속기간을 설명하는 데이터, 또는 비디오 파일 (150) 을 일반적으로 설명하는 다른 데이터를 포함할 수도 있다.

[0104] MOOV 박스 (154) 는 비디오 파일 (150) 에서의 트랙들의 수와 동일한 수의 TRAK 박스들을 포함할 수도 있다. TRAK 박스 (158) 는 비디오 파일 (150) 의 대응하는 트랙의 특성들을 설명할 수도 있다. 예를 들어, TRAK 박스 (158) 는 대응하는 트랙에 대한 시간 및/또는 공간 정보를 설명할 수도 있다. TRAK 박스 (158) 는 비디오 파일 (150) 의 트랙에 대한 데이터를 포함할 수도 있다. TRAK 박스 (158) 는 TRAK 박스 (158) 에 대응하는 트랙의 특성들을 설명하는 트랙 헤더 (TKHD) 박스를 포함할 수도 있다. 일부 예들에서, TRAK 박스 (158) 는 코딩된 비디오 픽처들을 포함할 수도 있는 한편, 다른 예들에서, 트랙의 코딩된 비디오 픽처들은 TRAK 박스 (158) 의 데이터에 의해 참조될 수도 있는 MDAT 박스 (167) 에 포함될 수도 있다.

[0105] 비디오 파일 (150) 은 타이밍된 메타데이터 트랙을 포함할 수도 있다. 예를 들어, 비디오 파일 (150) 은 하나 이상의 타이밍된 메타데이터 트랙들을 포함하는 표현에 대한 것일 수도 있다. 예를 들어, 비디오 파일 (150) 은 하나 이상의 타이밍된 메타데이터 트랙들을 포함하고 연관된 미디어 트랙(들)을 포함하지 않는 표현 (124N) 에 대한 것일 수도 있다. 일부 예들에서, 비디오 파일 (150) 은 연관된 미디어 트랙과 함께 멀티플렉싱된 하나 이상의 타이밍된 메타데이터 트랙들을 포함하는 표현에 대한 것일 수도 있다. 예를 들어, 비디오 파일 (150) 은 연관된 미디어 트랙(들)과 함께 멀티플렉싱된 하나 이상의 타이밍된 메타데이터 트랙들을 포함하는 표현 (124A) 에 대한 것일 수도 있다. 어느 경우나, 타이밍된 메타데이터 트랙은 비디오 파일 (150) 의 트랙에 대한 설명적 또는 주석적 메타데이터를 포함할 수도 있다.

[0106] 비디오 파일 (150) 은 비디오 파일 (150) 에 포함된 각 개별의 트랙에 대해 개별의 쌍의 MOOF 박스 (164) 및 MDAT 박스 (167) 를 포함할 수도 있다. 도시한 바와 같이, 각각의 MOOF 박스 (164) 는 트랙 프래그먼트 (TRAF) 박스 (165) 를 포함할 수도 있다. TRAF 박스 (165) 는, 도 3 의 예에서, 샘플-투-그룹 (sample-to-group; SBGP) (162) 을 포함한다. SBGP 박스 (162) 는 샘플 또는 트랙이 속하는 그룹 및 그 샘플 그룹의 연관된 디스크립션을 설명할 수도 있다. TRAF 박스 (165) 는 캡슐화 유닛 (30) (도 1) 이 비디오 파일 (150) 과 같은 비디오 파일에 파라미터 세트 트랙을 포함할 때, 파라미터 세트 트랙의 특성들을 설명할 수도 있다. 캡슐화 유닛 (30) 은 파라미터 세트 트랙을 설명하는 TRAK 박스 (158) 내의 파라미터 세트 트랙에서 시퀀스 레벨 SEI 메시지들의 존재를 시그널링할 수도 있다.

[0107] MDAT 박스 (167) 는 개별의 트랙에 대해 SEI 메시지 (172) 및 VCL NAL 유닛 (170) 을 포함할 수도 있다. VCL NAL 유닛 (170) 은 하나 이상의 코딩된 비디오 픽처들을 포함할 수도 있다. 일부 예들에서, VCL NAL 유

닛 (170) 은, 각각이 다수의 코딩된 비디오 픽처들, 예를 들어, 프레임들 또는 픽처들을 포함할 수도 있는, 하나 이상의 픽처들의 그룹들 (GOP들) 을 포함할 수도 있다.

[0108] 상기 언급한 바와 같이, 캡슐화 유닛 (30) 은 실제 코딩된 비디오 데이터를 포함하지 않는 비디오 샘플에 시퀀스 데이터 세트를 저장할 수도 있다. 비디오 샘플은 일반적으로, 특정 시간 인스턴스에서 코딩된 픽처의 표현인 액세스 유닛에 대응할 수도 있다. AVC 의 문맥에서, 액세스 유닛은 액세스 유닛의 모든 픽셀들을 구성하기 위해 정보를 포함하는 적어도 하나의 VCL NAL 유닛 (170) 및 다른 연관된 비-VCL NAL 유닛들, 이를 테면 SEI 메시지 (172) 를 포함한다. 이에 따라, 캡슐화 유닛 (30) 은 SEI 메시지 (172) 및 VCL NAL 유닛 (170) 을 포함할 수도 있는 시퀀스 데이터 세트를 포함할 수도 있다. 캡슐화 유닛 (30) 은 시퀀스 데이터 세트 및/또는 SEI 메시지 (172) 의 존재를 MDAT 박스 (167) 내에 존재하는 것으로서 추가로 시그널링할 수도 있다.

[0109] 비디오 파일 (150) 은 본 개시의 기법들에 따른 하나 이상의 가장 관심있는 영역들을 나타내는 데이터를 포함할 수도 있다. 보다 구체적으로, 비디오 파일 (150) 은 비디오 파일 (150) 의 파일 포맷 헤더 정보에 하나 이상의 가장 관심있는 영역들을 나타내는 데이터를 포함할 수도 있다. 예를 들어, 비디오 파일 (150) 은 가장 관심있는 영역들 정보를 시그널링하는 SEI 메시지 및/또는 파일 포맷 정보를 포함할 수도 있다. 예를 들어, 타이밍된 메타 데이터 트랙은 하나 이상의 가장 관심있는 영역들을 나타내는 데이터를 포함할 수도 있다. 일부 인스턴스들에서, SBGP 박스 (162) 는 하나 이상의 가장 관심있는 영역들을 나타내는 데이터를 포함할 수도 있다. 일부 인스턴스들에서, SEI 메시지 (172) 는 하나 이상의 가장 관심있는 영역들을 나타내는 데이터를 포함할 수도 있다.

[0110] SEI 메시지 (172) 는 H.264/AVC, H.265/HEVC, 또는 다른 비디오 코덱 규격들의 파트로서 포함될 수 있다. 도 3 의 예에서, SEI 메시지는 예를 들어, VCL NAL 유닛 (170) 의 대응하는 픽처에 대한 가장 관심있는 영역들을 나타내는 데이터를 포함한다. 다른 예들에서, 이러한 가장 관심있는 영역들 데이터는 1) 파일 포맷에서의 정보의 시그널링을 위한 파일 포맷 박스 또는 구조, 2) SDP 에서의 정보의 시그널링을 위한 SDP 필드, 3) DASH MPD (예를 들어, 도 2 의 MPD (122)) 에서의 정보의 시그널링을 위한 DASH MPD 속성 또는 엘리먼트, 4) MPEG-2 TS 에서의 정보의 시그널링을 위한 MPEG-2 전송 스트림 (TS) 디스크립터에, 다른 박스 또는 구조에, 또는 그 조합에 포함될 수도 있다.

[0111] MFRA 박스 (166) 는 비디오 파일 (150) 의 무비 프래그먼트들 내의 랜덤 액세스 포인트들을 설명할 수도 있다. 이것은 트릭 모드들을 수행하는 것, 이를 테면 비디오 파일 (150) 에 의해 캡슐화된 세그먼트 내에서 특정한 시간 로케이션들 (즉, 플레이백 시간들) 에 대한 탐색들을 수행하는 것을 도울 수도 있다. MFRA 박스 (166) 는 일반적으로 옵션이고 일부 예들에서, 비디오 파일들에 포함될 필요가 없다. 마찬가지로, 클라이언트 디바이스, 이를 테면 클라이언트 디바이스 (40) 는 비디오 파일 (150) 의 비디오 데이터를 정확히 디코딩 및 디스플레이 하기 위해 반드시 MFRA 박스 (166) 를 참조할 필요는 없다. MFRA 박스 (166) 는 비디오 파일 (150) 의 트랙들의 수와 동일한, 또는 일부 예들에서, 비디오 파일 (150) 의 미디어 트랙들 (예를 들어, 비-힌트 트랙들) 의 수와 동일한 수의 트랙 프래그먼트 랜덤 액세스 (TFRA) 박스들 (미도시) 을 포함할 수도 있다.

[0112] 일부 예들에서, 비디오 파일 (150) 의 무비 프래그먼트들은 IDR 픽처들과 같은 하나 이상의 스트림 액세스 포인트들 (SAP들) 을 포함할 수도 있다. 마찬가지로, MFRA 박스 (166) 는 SAP들의 비디오 파일 (150) 내의 로케이션들의 표시들을 제공할 수도 있다. 이에 따라, 비디오 파일 (150) 의 시간 서브-시퀀스는 비디오 파일 (150) 의 SAP들로부터 형성될 수도 있다. 시간 서브-시퀀스는 또한, 다른 픽처들, 이를 테면 SAP들에 의존하는 P-프레임들 및/또는 B-프레임들을 포함할 수도 있다. 시간 서브-시퀀스의 프레임들 및/또는 슬라이스들은 서브-시퀀스의 다른 프레임들/슬라이스들에 의존하는 시간 서브-시퀀스의 프레임들/슬라이스들이 적절히 디코딩될 수 있도록 세그먼트들 내에 배열될 수도 있다. 예를 들어, 데이터의 계층적 배열에서, 다른 데이터에 대한 예측을 위해 이용되는 데이터는 또한 시간 서브-시퀀스에 포함될 수도 있다.

[0113] 비디오 파일 (150) 은 가장 관심있는 영역들의 수를 표시하는 선택스 엘리먼트를 포함할 수도 있다. 예를 들어, 타이밍된 메타데이터 트랙은 가장 관심있는 영역들의 수를 표시하는 선택스 엘리먼트를 포함할 수도 있다. 일부 예들에서, SBGP 박스 (162) 는 가장 관심있는 영역들의 수를 표시하는 선택스 엘리먼트를 포함할 수도 있다. 일부 예들에서, SEI 메시지 (172) 는 가장 관심있는 영역들의 수를 표시하는 선택스 엘리먼트를 포함할 수도 있다.

[0114] 비디오 파일 (150) 은 현재 인스턴스들에서의 정보가 출력 순서에서의 다음 이미지, 또는 새로운 코딩된 비디오 시퀀스의 시작, 또는 비트스트림의 종료까지 지속되는지 여부를 표시하는 선택스 엘리먼트를 포함할 수도 있다. 예를 들어, 타이밍된 메타데이터 트랙은 타이밍된 메타데이터 트랙의 현재 인스턴스들에서의 정보가 출력 순

서에서의 다음 이미지, 또는 새로운 코딩된 비디오 시퀀스의 시작, 또는 비트스트림의 종료까지 지속되는지 여부를 표시하는 선택스 엘리먼트를 포함할 수도 있다. 일부 예들에서, SBGP 박스 (162) 는 SBGP 박스 (162) 의 현재 인스턴스들에서의 정보가 출력 순서에서의 다음 이미지, 또는 새로운 코딩된 비디오 시퀀스의 시작, 또는 비트스트림의 종료까지 지속되는지 여부를 표시하는 선택스 엘리먼트를 포함할 수도 있다. 일부 예들에서, SEI 메시지 (172) 는 SEI 메시지 (172) 의 현재 인스턴스들에서의 정보가 출력 순서에서의 다음 이미지, 또는 새로운 코딩된 비디오 시퀀스의 시작, 또는 비트스트림의 종료까지 지속되는지 여부를 표시하는 선택스 엘리먼트를 포함할 수도 있다.

[0115] 비디오 파일 (150) 은 디코딩 순서에서 모든 선행 인스턴스들의 효과를 캔슬할지 여부를 표시하는 선택스 엘리먼트를 포함할 수도 있다. 달리 말하면, 하나 이상의 가장 관심있는 영역들을 나타내는 데이터가 하나 이상의 가장 관심있는 영역들이 출력 순서에서의 다음 이미지, 또는 새로운 코딩된 비디오 시퀀스의 시작, 또는 비트스트림의 종료까지 지속된다는 것을 표시하는 인스턴스들에서, 효과를 캔슬할지 여부를 표시하는 선택스 엘리먼트는 하나 이상의 가장 관심있는 영역들의 지속성을 제거할 수도 있다. 예를 들어, 타이밍된 메타데이터 트랙은 디코딩 순서에서 타이밍된 메타데이터 트랙의 모든 선행 인스턴스들의 효과를 캔슬할지 여부를 표시하는 선택스 엘리먼트를 포함할 수도 있다. 일부 예들에서, SBGP 박스 (162) 는 디코딩 순서에서 SBGP 박스 (162) 의 모든 선행 인스턴스들의 효과를 캔슬할지 여부를 표시하는 선택스 엘리먼트를 포함할 수도 있다. 일부 예들에서, SEI 메시지 (172) 는 디코딩 순서에서 SEI 메시지 (172) 의 모든 선행 인스턴스들의 효과를 캔슬할지 여부를 표시하는 선택스 엘리먼트를 포함할 수도 있다.

[0116] 비디오 파일 (150) 은, 가장 관심있는 영역들의 각각에 대해, 선택스 엘리먼트에 의해 우선순위를 표시할 수도 있으며, 여기서 높은 우선순위는 영역의 보다 높은 관심을 표시한다. 예를 들어, 타이밍된 메타데이터 트랙은 가장 관심있는 영역의 각각에 대해, 선택스 엘리먼트에 의해 우선순위를 표시할 수도 있고, 여기서 높은 우선순위는 영역의 보다 높은 관심을 표시한다. 일부 예들에서, SBGP 박스 (162) 는 가장 관심있는 영역의 각각에 대해, 선택스 엘리먼트에 의해 우선순위를 표시할 수도 있고, 여기서 높은 우선순위는 영역의 보다 높은 관심을 표시한다. 일부 예들에서, SEI 메시지 (172) 는 가장 관심있는 영역의 각각에 대해, 선택스 엘리먼트에 의해 우선순위를 표시할 수도 있고, 여기서 높은 우선순위는 영역의 보다 높은 관심을 표시한다.

[0117] 본 개시의 기법들에 따르면, VCL NAL 유닛 (170) 에 포함된 비디오 데이터를 포함하는 비디오 파일 (150) 의 파일 포맷 정보에서의 데이터 구조는 가장 관심있는 영역의 개별의 포지션 및 가장 관심있는 영역의 개별의 사이즈의 표시를 포함할 수도 있다.

[0118] 예를 들어, 타이밍된 메타데이터 트랙은 가장 관심있는 영역들을 나타내는 데이터를 포함할 수도 있다. 예를 들어, 타이밍된 메타데이터 트랙은, 각각의 무비 프래그먼트에 대해, 가장 관심있는 영역의 개별의 포지션 및 가장 관심있는 영역의 개별의 사이즈의 표시를 포함할 수도 있다.

[0119] 보다 구체적으로, 일부 예들에서, 타이밍된 메타데이터 트랙은, 이미지를 형성하는 하나 이상의 무비 프래그먼트들의 각각의 하나에 대해, 가장 관심있는 영역에 대한 루프의 표시를 포함할 수도 있고, 루프의 각각의 반복은 가장 관심있는 영역들 중 하나를 나타내고 이미지에 대한 루마 샘플들의 수에서의 포지션 및 이미지에 대한 루마 샘플들의 수에서의 가장 관심있는 영역의 사이즈를 나타내는 선택스 엘리먼트들을 포함한다.

[0120] 일부 예들에서, 타이밍된 메타데이터 트랙은, 하나 이상의 무비 프래그먼트들의 각각의 하나에 대해, 이미지에 대한 루마 샘플들의 수에서의 가장 관심있는 영역의 상부를 나타내는 제 1 선택스 엘리먼트, 이미지에 대한 루마 샘플들의 수에서의 가장 관심있는 영역의 하부를 나타내는 제 2 선택스 엘리먼트, 이미지에 대한 루마 샘플들의 수에서의 가장 관심있는 영역의 좌측, 및 이미지에 대한 루마 샘플들의 수에서의 가장 관심있는 영역의 우측의 표시를 포함할 수도 있다.

[0121] 일부 예들에서, 타이밍된 메타데이터 트랙은, 가장 관심있는 영역들의 각각에 대해, 이미지에 대한 루마 샘플들의 수에서의 가장 관심있는 영역의 좌상측 코너를 나타내는 제 1 선택스 엘리먼트 및 이미지에 대한 루마 샘플들의 수에서의 가장 관심있는 영역의 우하측 코너를 나타내는 제 2 선택스 엘리먼트를 포함하는 엘리먼트들 또는 속성들을 각각의 무비 프래그먼트에 대해 포함할 수도 있다.

[0122] 일부 예들에서, 타이밍된 메타데이터 트랙은, 가장 관심있는 영역들의 각각에 대해, 이미지에 대한 루마 샘플들의 수에서의 가장 관심있는 영역의 좌상측 코너를 나타내는 제 1 선택스 엘리먼트, 이미지에 대한 루마 샘플들의 수에서의 가장 관심있는 영역의 폭을 나타내는 제 2 선택스 엘리먼트, 및 이미지에 대한 루마 샘플들의 수에서의 가장 관심있는 영역의 높이를 나타내는 제 3 선택스 엘리먼트를 포함하는 엘리먼트들 또는 속성들을 각각

의 무비 프래그먼트에 대해 포함할 수도 있다.

- [0123] 다른 예에서, SBGP 박스 (162) 는 가장 관심있는 영역들을 나타내는 데이터를 포함할 수도 있다. 예를 들어, SBGP 박스 (162) 는 각각의 무비 프래그먼트에 대해, 가장 관심있는 영역의 개별의 포지션 및 가장 관심있는 영역의 개별의 사이즈의 표시를 포함할 수도 있다.
- [0124] 보다 구체적으로, 일부 예들에서, SBGP 박스 (162) 는, 이미지를 형성하는 각각의 무비 프래그먼트에 대해, 가장 관심있는 영역에 대한 그 루프의 표시를 포함할 수도 있고, 그 루프의 각각의 반복은 가장 관심있는 영역들 중 하나를 나타내고 이미지에 대한 루마 샘플들의 수에서의 가장 관심있는 영역의 포지션 및 이미지에 대한 루마 샘플들의 수에서의 가장 관심있는 영역의 사이즈를 나타내는 신택스 엘리먼트들을 포함한다.
- [0125] 일부 예들에서, SBGP 박스 (162) 는 이미지를 형성하는 각각의 무비 프래그먼트에 대해, 이미지에 대한 루마 샘플들의 수에서의 가장 관심있는 영역의 상부를 나타내는 제 1 신택스 엘리먼트, 이미지에 대한 루마 샘플들의 수에서의 가장 관심있는 영역의 하부를 나타내는 제 2 신택스 엘리먼트, 이미지에 대한 루마 샘플들의 수에서의 가장 관심있는 영역의 좌측, 및 이미지에 대한 루마 샘플들의 수에서의 가장 관심있는 영역의 우측의 표시를 포함할 수도 있다.
- [0126] 일부 예들에서, SBGP 박스 (162) 는, 가장 관심있는 영역들의 각각에 대해, 이미지에 대한 루마 샘플들의 수에서의 가장 관심있는 영역의 좌상측 코너를 나타내는 제 1 신택스 엘리먼트 및 이미지에 대한 루마 샘플들의 수에서의 가장 관심있는 영역의 우하측 코너를 나타내는 제 2 신택스 엘리먼트를 포함하는 엘리먼트들 또는 속성들을 이미지를 형성하는 각각의 무비 프래그먼트에 대해 포함할 수도 있다.
- [0127] 일부 예들에서, SBGP 박스 (162) 는, 가장 관심있는 영역들의 각각에 대해, 이미지에 대한 루마 샘플들의 수에서의 가장 관심있는 영역의 좌상측 코너를 나타내는 제 1 신택스 엘리먼트, 이미지에 대한 루마 샘플들의 수에서의 가장 관심있는 영역의 폭을 나타내는 제 2 신택스 엘리먼트, 및 이미지에 대한 루마 샘플들의 수에서의 가장 관심있는 영역의 높이를 나타내는 제 3 신택스 엘리먼트를 포함하는 엘리먼트들 또는 속성들을 이미지를 형성하는 각각의 무비 프래그먼트에 대해 포함할 수도 있다.
- [0128] 다른 예에서, SEI 메시지 (172) 는 가장 관심있는 영역들을 나타내는 데이터를 포함할 수도 있다. 예를 들어, SEI 메시지 (172) 는, 이미지를 형성하는 각각의 무비 프래그먼트에 대해, 가장 관심있는 영역의 개별의 포지션 및 가장 관심있는 영역의 개별의 사이즈의 표시를 포함할 수도 있다.
- [0129] 보다 구체적으로, 예를 들어, SEI 메시지 (172) 는, 이미지를 형성하는 각각의 무비 프래그먼트에 대해, 가장 관심있는 영역에 대한 그 루프의 표시를 포함할 수도 있고, 그 루프의 각각의 반복은 가장 관심있는 영역들 중 하나를 나타내고 이미지에 대한 루마 샘플들의 수에서의 가장 관심있는 영역의 포지션 및 이미지에 대한 루마 샘플들의 수에서의 가장 관심있는 영역의 사이즈를 나타내는 신택스 엘리먼트들을 포함한다.
- [0130] 일부 예들에서, SEI 메시지 (172) 는, 이미지를 형성하는 각각의 무비 프래그먼트에 대해, 이미지에 대한 루마 샘플들의 수에서의 가장 관심있는 영역의 상부를 나타내는 제 1 신택스 엘리먼트, 이미지에 대한 루마 샘플들의 수에서의 가장 관심있는 영역의 하부를 나타내는 제 2 신택스 엘리먼트, 이미지에 대한 루마 샘플들의 수에서의 가장 관심있는 영역의 좌측, 및 이미지에 대한 루마 샘플들의 수에서의 가장 관심있는 영역의 우측의 표시를 포함할 수도 있다.
- [0131] 일부 예들에서, SEI 메시지 (172) 는, 가장 관심있는 영역들의 각각에 대해, 이미지에 대한 루마 샘플들의 수에서의 가장 관심있는 영역의 좌상측 코너를 나타내는 제 1 신택스 엘리먼트 및 이미지에 대한 루마 샘플들의 수에서의 가장 관심있는 영역의 우하측 코너를 나타내는 제 2 신택스 엘리먼트를 포함하는 엘리먼트들 또는 속성들을 이미지를 형성하는 각각의 무비 프래그먼트에 대해 포함할 수도 있다.
- [0132] 일부 예들에서, SEI 메시지 (172) 는, 가장 관심있는 영역들의 각각에 대해, 이미지에 대한 루마 샘플들의 수에서의 가장 관심있는 영역의 좌상측 코너를 나타내는 제 1 신택스 엘리먼트, 이미지에 대한 루마 샘플들의 수에서의 가장 관심있는 영역의 폭을 나타내는 제 2 신택스 엘리먼트, 및 이미지에 대한 루마 샘플들의 수에서의 가장 관심있는 영역들의 높이를 나타내는 제 3 신택스 엘리먼트를 포함하는 엘리먼트들 또는 속성들을 이미지를 형성하는 각각의 무비 프래그먼트에 대해 포함할 수도 있다.
- [0133] 도 4 는 입방형 프로젝션 (230) 에서의 예의 타일 코딩을 예시하는 개념적 다이어그램이다. 입방형 프로젝션 (230) 은 도시한 바와 같이 입방체 면들 (232A, 232B, 232C) 을 포함하고, 가려져 있기 때문에 미도시되는 추가적인 입방체 면들을 포함한다. 이 예에서, 입방형 프로젝션 (230) 의 입방체 면들의 각각은 4 개의 타

일들로 분할된다: 입방체 면 (232A) 은 4 개의 타일들 (234A 내지 234D) 로 분할되고, 입방체 면 (232B) 은 4 개의 타일들 (236A 내지 236D) 로 분할되고, 그리고 입방체 면 (232C) 은 4 개의 타일들 (238A 내지 238D) 로 분할된다. 다른, 미도시된 입방체 면들은 또한, 이 예에서, 총 24 개의 작은 타일들을 위해, 4 개의 개별의 타일들로 분할될 것이다. 일반적으로, 콘텐츠 준비 디바이스 (20) 는 타일들 (예를 들어, 타일들 (234A 내지 234D, 236A 내지 236D, 238A 내지 238-D), 및 미도시된 입방체 면들의 타일들) 의 각각에 대한 비디오 데이터를 제공할 수도 있다.

[0134] 추가로, 콘텐츠 준비 디바이스 (20) 는 입방체 면들의 각각을 커버할 수도 있는 큰 타일들에 대한 비디오 데이터를 제공할 수도 있다. 예를 들어, 콘텐츠 준비 디바이스 (20) 는 입방체 면 (232A) 을 커버하는 타일에 대한 비디오 데이터, 입방체 면 (232B) 을 커버하는 타일에 대한 비디오 데이터, 입방체 면 (232C) 을 커버하는 타일에 대한 비디오 데이터, 그리고 미도시된 입방체 면들을 커버하는 타일들에 대한 비디오 데이터를 제공할 수도 있다. 일부 예들에서, 비디오 인코더 (28) 는 뷰잉될 가능성이 더 큰 그 뷰포인트들에 대해서만 보다 큰 타일 사이즈들을 인코딩할 수도 있다. 따라서, 사용자가 위로 또는 아래로 볼 가능성이 낮으면, 비디오 인코더 (28) 는 단지 입방체 면들 (232B 및 232C), 및 입방체 면들 (232B 및 232C) 과 대향하는 입방체 면들을 커버하는 큰 타일들에 대한 비디오 데이터를 코딩할 수도 있지만, 예를 들어, 입방체 면 (232A) 은 아니다. 옵션으로, 타일들의 각각은 스테레오 비디오 데이터에 대한 좌측 및 우측 뷰들을 가질 수도 있다.

[0135] VR 은 몰입된 사용자의 움직임들에 의해 상관된 사운드와 자연 및/또는 합성 이미지들의 렌더링에 의해 생성된 비-물리적 세계에 가상으로 존재하는 능력이어서, 사용자가 그 세계와 상호작용하는 것을 허용한다. 렌더링 디바이스들, 이를 테면 헤드 장착 디스플레이들 (head mounted displays; HMD들), 및 VR 비디오 생성 (종종 360 도 비디오로도 지칭됨) 에서 행해진 최근의 진보로, 상당한 품질의 경험이 제공될 수 있다. VR 애플리케이션들은 게이밍, 트레이닝, 교육, 스포츠, 비디오, 온라인 쇼핑, 성인 오락, 등을 포함한다.

[0136] 통상의 VR 시스템은 다음의 컴포넌트들 및 기법들을 포함할 수도 있다:

[0137] 1) 카메라 세트, 이는 상이한 방향들을 가리키고 이상적으로는 카메라 세트 주위의 모든 뷰포인트들을 집합적으로 커버하는 다중 개개의 카메라들로 통상 이루어진다.

[0138] 2) 이미지 스티칭, 여기서 다중 개개의 카메라들에 의해 촬영된 비디오 픽처들은 시간 도메인에서 동기화되고 공간 도메인에서 스티칭되어, 구면 비디오가 되지만, 직사각형 포맷, 이를 테면 (월드 맵 (world map) 같은) 등-직사각형 (equi-rectangular) 또는 입방체 맵에 맵핑된다.

[0139] 3) 맵핑된 직사각형 포맷의 비디오는 비디오 코덱, 예를 들어, H.265/HEVC 또는 H.264/AVC 를 이용하여 인코딩/압축된다.

[0140] 4) 압축된 비디오 비트스트림(들)은 미디어 포맷으로 저장 및/또는 캡슐화되고 네트워크를 통하여 수신기에 (어쩌면 사용자에게 의해 보이는 에어리어만을 커버하는 서브세트만) 송신될 수도 있다.

[0141] 5) 수신기는, 어쩌면 포맷으로 캡슐화된, 비디오 비트스트림(들) 또는 그 파트를 수신하고, 디코딩된 비디오 신호 또는 그 파트를 렌더링 디바이스로 전송한다.

[0142] 6) 렌더링 디바이스는 예를 들어, HMD 일 수 있고, 이는 머리 움직임 및 심지어 안구 이동 순간을 추적하여 실감나는 경험이 사용자에게 전달되도록 비디오의 대응하는 파트를 렌더링할 수 있다.

[0143] 보통의 비디오와 비교한 VR 비디오의 피쳐는, VR 에서는 통상적으로 단지 뷰포트 (viewport) 로도 또한 지칭되는, 현재 FOV 에 대응하는, 비디오 픽처들에 의해 표현된 전체 비디오 영역의 서브세트, 예를 들어, 사용자가 현재 보고 있는 에어리어만이 디스플레이되는 한편, 보통의 비디오 애플리케이션들에서는 통상적으로 전체 비디오 영역이 디스플레이된다는 것이다. 이 피쳐는 예를 들어, 뷰포트 의존 프로젝션 맵핑 또는 뷰포트 의존 비디오 코딩을 이용하는 것에 의해, VR 비디오 시스템들의 성능을 향상시키기 위해 활용될 수도 있다. 성능 향상은 사용자에게 제시된 비디오 파트의 동일한 레졸루션/품질 하에서 종래의 VR 비디오 시스템들과 비교하여 보다 낮은 송신 대역폭 및 보다 낮은 디코딩 복잡성 중 어느 하나 또는 양자 모두일 수 있다.

[0144] 뷰포트 의존 프로젝션 맵핑은 또한 비대칭 프로젝션 맵핑으로 지칭될 수도 있다. 하나의 예는 서브-샘플링된 입방체-맵, 예를 들어, 입방형 프로젝션 (230) 이다. 통상의 cub-map 은 6 개의 동일한 사이즈의 면들을 포함할 수도 있다. 서브-샘플링된 cub-map 의 하나의 예에서, 면들 중 하나는 변화하지 않을 수 있는 한편, 대향 측의 면은 원래의 면 형상의 중심 에어리어에 로케이트된 보다 작은 사이즈로 서브-샘플링되거나 또는 다운-스케일링될 수 있고, 다른 면들은 이에 따라 기하학적으로 스케일링된다 (여전히 6 개의 면들을 유지함).

극단 (extreme) 은 대향 측의 면을 단일 포인트인 것으로 다운-스케일링하는 것일 수도 있으며, 따라서 입방체는 피라미드가 된다. 서버-샘플링된 cub-map 의 다른 예에서, 일부 면들은 예를 들어, 2x2 비율 (즉, 면의 에지에 평행한 각각의 방향에서 2:1) 만큼 비례하여 다운-사이징된다.

- [0145] 이러한 다운-사이징된 스케일링은 또한, 등-직사각형과 같은 다른 프로젝션 맵핑들에 대한 상이한 영역들에 적용될 수 있다. 하나의 예는 상부 및 하부 영역들 (즉, 극 (pole) 들) 양자 모두를 다운-사이징하는 것이다.
- [0146] 뷰포트 의존 비디오 코딩은 또한, 현재 FOV 또는 뷰포트의 디스플레이에 대해 충분한 정보를 제공하기 위해 전체 인코딩된 비디오 영역을 단지 부분적으로 디코딩하는 것을 가능하게 하는, 뷰포트 기반 부분 비디오 디코딩으로 지칭될 수도 있다.
- [0147] 뷰포트 의존 비디오 코딩의 하나의 예에서, VR 비디오는 FOV 또는 뷰포트를 커버하는 각각의 잠재적인 영역이 다른 영역들과는 독립적으로 디코딩될 수 있도록 타일들의 모션-제한된 세트들을 이용하여 코딩된다. 특정한 현재 FOV 또는 뷰포트에 대해, 현재 FOV 또는 뷰포트를 커버하는 독립적으로 디코딩가능한 타일들의 최소 세트의 코딩된 비디오 데이터는 클라이언트로 전송되고, 디코딩되고, 디스플레이된다. 이 접근법의 결점은, 사용자가 그 또는 그녀의 머리를 전송된 비디오 데이터에 의해 커버되지 않는 새로운 FOV 로 빠르게 돌릴 때, 새로운 FOV 를 커버하는 데이터가 도달하기 전에는 비디오 데이터의 타일들에 의해 커버되지 않는 에어리어에서 어떤 것도 볼 수 없다는 것이다. 이것은 예를 들어 10 밀리초의 매그니튜드로 매우 낮지 않은 한, 네트워크의 라운드 트립 시간이 쉽게 일어날 수 있으며, 이는 실현가능하지 않거나 또는 적어도 오늘날 또는 가까운 미래의 큰 도전과제이다.
- [0148] 뷰포트 의존 비디오 코딩의 다른 예는, IMRC (independent multi-resolution coding) 로 명명된, 그 전체 내용이 본 명세서에 참조로 통합되는, 2016년 6월 28일자로 출원된 미국 특허출원 제15/195,439호에서 제안되었으며, 여기서 구형/파노라마 비디오는 서로 독립적으로 다중 상이한 레졸루션들에서 인코딩되고, 각각의 표현은 타일들의 모션-제한된 세트를 이용하여 코딩된다. 수신기는 상이한 레졸루션들에서 구형/파노라마 비디오의 상이한 부분들을 디코딩하도록 선정한다. 통상적으로, 사용자에게 의해 현재 관측되고 있는 구형 비디오의 부분, 즉 현재 FOV 또는 현재 뷰포트는 가장 높은 레졸루션에서 코딩된 비디오 비트스트림의 파트이다. 현재 FOV 를 둘러싸는 영역은 약간 더 낮은 레졸루션을 이용하여 코딩된 비디오 비트스트림의 파트 등이다. 관측자의 머리 바로 뒤의 파노라마의 부분은 가장 낮은 레졸루션에서 코딩된 비디오 비트스트림의 파트이다. 사용자에게 의한 머리 움직임의 경우에, 사용자 경험은 대개의 경우에 작은 양만큼만 열화될 것이고, 품질 열화는 매우 큰 갑작스런 머리 움직임의 경우에만 가장 심각하며, 이는 드물다고 주장된다.
- [0149] 뷰포트 의존 비디오 코딩의 또 다른 예에서, 다중-레졸루션들의 비디오 비트스트림들은 SHVC 와 같은 스케일러블 비디오 코덱을 이용하여 스케일러블-코딩된다. 추가로, 가장 낮은 레졸루션 비디오 (또는 베이스 계층) 의 전체 비트스트림은 항상 전송된다. 가장 낮은 레졸루션 비디오는 타일들 또는 타일들의 모션-제한된 세트들을 이용하여 코딩될 필요가 없고, 하지만 그것이 타일들 또는 타일들의 모션-제한된 세트들을 이용하여 코딩되면 그것은 또한 작동할 것이다. 다른 양태들에 대해, 상기 설명한 바와 동일한 전략이 적용된다. 이 접근법은 타일들 또는 타일들의 타일 모션-제한된 세트들의 코딩이 코딩 효율을 감소시키기 때문에 가장 낮은 레졸루션 비디오의 보다 효율적인 코딩, 그리고 또한, 층간 예측의 사용으로 인한 보다 높은 레졸루션들의 스트림들의 보다 효율적인 코딩을 허용한다. 더욱이, 이 접근법은 또한, 일단 FOV 스위칭이 시작되면, 서버 또는 에지 서버가 (베이스 계층보다) 더 높은 레졸루션의 비디오 스트림에서 오고 새로운 FOV 의 어떤 영역도 커버하고 있지 않은 비디오 데이터를 전송하는 것을 즉시 중단할 수 있기 때문에, 보다 효율적인 FOV 스위칭을 허용한다. 현재 FOV 를 커버하는 가장 낮은 레졸루션의 비디오 데이터가 전송되지 않으면 (즉, 현재 FOV 를 커버하는 가장 높은 레졸루션의 비디오 데이터가 단지 전송됨), FOV 스위칭 동안, 서버는 때때로 가장 높은 레졸루션의 비디오 스트림에서 오고 단지 이전의 (old) FOV 또는 그 파트를 커버하고 있는 비디오 데이터를 계속 전송하여, 사용자가 그/그녀가 다시 이전의 FOV 또는 그 파트로 돌리는 경우에 가장 낮은 품질의 비디오를 적어도 볼 수 있도록 준비되게 해야 할 것이다.
- [0150] MPEG 기여 m37819 에서, 사용 케이스는, 심지어 사용자가 그/그녀의 머리를 돌리거나 또는 다른 UI 를 통하여 뷰포트를 변화시키지 않아도 디렉터가 관중이 집중하길 원하는 동적으로 변화하는 뷰포트를 VR 플레이백이 디스플레이할 수도 있도록, 디렉터의 컷에 대한 정보의 시그널링 및 이용에 대해 논의되었다. 이러한 뷰포트들에는 장면 단위로 전방향 비디오가 제공될 수도 있다는 것이 언급되었다.
- [0151] 타일들 (예를 들어, 타일들 (234A 내지 234D, 236A 내지 236D, 238A 내지 238D), 및 미도시된 입방체 면들의 타일들) 을 요청하기 위해 품질을 선택하도록 단지 사용자의 시야에 반드시 의존하기 보다는, 본 명세서에서 설

명된 하나 이상의 기법들은 클라이언트 디바이스 (40) 및/또는 서버 디바이스 (60) 가 하나 이상의 가장 관심있는 영역들을 포함하는 하나 이상의 타일들에 대한 고품질 데이터를 프리-페치하는 것을 허락한다. 예를 들어, 클라이언트 디바이스 (40) 는 하나 이상의 가장 관심있는 영역들을 나타내는 데이터가 타일들 (234A 내지 234D) 이 하나 이상의 가장 관심있는 영역들을 포함한다는 것을 표시할 때 사용자의 FOV 를 수신하기 전에 타일들 (234A 내지 234D) 을 요청할 수도 있다. FOV 가 하나 이상의 가장 관심있는 영역들에 포함되지 않는 이미지의 영역들을 포함하는 예에서, 사용자의 FOV 를 수신 시에, 클라이언트 디바이스 (40) 는 저품질로 사용자의 현재 FOV 를 충족하기 위해 추가 타일들 (예를 들어, 236A 내지 236D, 238A 내지 238D) 을 요청할 수도 있다. 이 방식으로, 하나 이상의 가장 관심있는 영역들은 하나 이상의 가장 관심있는 영역들이 고품질로 수신되는 것을 허락하기 위해 프리-페치 (예를 들어, 사용자의 FOV 를 수신하기 전에 페치) 되어, 비디오, 이를 테면 가상 현실 비디오의 플레이백을 향상시킬 수도 있다.

[0152] 도 5 는 본 개시에서 설명된 하나 이상의 기법들에 따른 비디오 데이터에 대한 정보를 결정하기 위한 예의 접근법을 예시하는 플로우차트이다. 도 5 의 방법들은 도 1 의 클라이언트 디바이스 (40) 및 서버 디바이스 (60) 에 대하여 설명된다. 그러나, 다른 디바이스들이 이들 기법들을 수행하도록 구성될 수도 있는 것으로 이해되어야 한다.

[0153] 초기에, 서버 디바이스 (60) 는 이미지의 가장 관심있는 영역들을 나타내는 데이터를 전송할 수도 있다 (300). 가장 관심있는 영역들을 나타내는 데이터는 비디오 데이터를 포함하는 비트스트림의 VCL 데이터의 외부에 있을 수도 있다. 예를 들어, 서버 디바이스 (60) 는 가장 관심있는 영역들을 나타내는 데이터를 포함하는 보충 강화 정보 (SEI) 메시지를 전송한다. 일부 예들에서, 서버 디바이스 (60) 는 가장 관심있는 영역들을 나타내는 데이터를 포함하는 샘플 그룹 박스 (예를 들어, 도 3 의 SBGP 박스 (162)) 를 전송한다. 일부 예들에서, 서버 디바이스 (60) 는 가장 관심있는 영역들을 나타내는 데이터를 포함하는 타이밍된 메타데이터 트랙을 전송한다.

[0154] 어느 경우나, 클라이언트 디바이스 (40) 는 가장 관심있는 영역들을 나타내는 데이터에 기초하여 이미지의 가장 관심있는 영역들을 결정할 수도 있다 (302). 예를 들어, 클라이언트 디바이스 (40) 는 하나 이상의 가장 관심있는 영역들의 각각의 하나에 대해, 가장 관심있는 영역들을 나타내는 데이터를 이용하여 개별의 가장 관심있는 영역의 포지션 및 개별의 가장 관심있는 영역의 사이즈를 나타내는 신택스 엘리먼트들에 대한 값들을 결정할 수도 있다.

[0155] 클라이언트 디바이스 (40) 는 하나 이상의 가장 관심있는 영역들에 대한 고품질 비디오 데이터를 프리-페치할 수도 있다 (304). 예를 들어, 클라이언트 디바이스 (40) 는 이미지의 다른 영역들과 비교하여 상대적으로 높은 비트레이트를 이용하여 하나 이상의 가장 관심있는 영역들에 대한 비디오 데이터를 요청할 수도 있다. 서버 디바이스 (60) 는 프리-페치 요청을 수신할 수도 있고 (306) 하나 이상의 가장 관심있는 영역들에 대한 고품질 비디오 데이터를 전송할 수도 있다 (308). 클라이언트 디바이스 (40) 는 고품질 비디오 데이터를 수신할 수도 있다 (310). 예를 들어, 클라이언트 디바이스 (40) 는 이미지의 다른 영역들에 대한 비디오 데이터보다 저레이턴시로 및/또는 고품질로 하나 이상의 가장 관심있는 영역들에 대한 비디오 데이터를 페치할 수도 있다.

[0156] 클라이언트 디바이스 (40) 는 이미지의 하나 이상의 다른 영역들에 대한 저품질 비디오 데이터에 대한 요청을 출력할 수도 있다 (314). 예를 들어, 클라이언트 디바이스 (40) 는 하나 이상의 가장 관심있는 영역들이 도 4 의 이미지들 (238A-D) 을 포함할 때 그리고 사용자의 FOV 가 도 4 의 이미지들 (238A-D 및 234C-D) 을 포함할 때 하나 이상의 가장 관심있는 영역들에 대한 비디오 품질과 비교하여 도 4 의 이미지들 (234C-D) 에 대한 상대적으로 저품질 비디오 데이터를 요청할 수도 있다. FOV 밖의 영역들 (예를 들어, 도 4 의 이미지들 (234A-B)) 은 요청되지 않을 수도 있거나 또는 FOV 를 이용하여 결정된 이미지들 (예를 들어, 도 4 의 이미지들 (238A-D 및 234C-D)) 보다 상대적으로 저품질로 요청될 수도 있는 것으로 이해되어야 한다.

[0157] 서버 디바이스 (60) 는 요청을 수신할 수도 있고 (316) 이미지의 하나 이상의 다른 영역들에 대해 요청된 데이터를 전송할 수도 있다 (318). 클라이언트 디바이스 (40) 는 요청된 데이터를 수신할 수도 있다 (320). 클라이언트 디바이스 (40) 는 사용자의 FOV 를 결정하고 (321) FOV 에 포함된 영역들에 대한 데이터를 사용자에게 출력할 수도 있다 (322). 예를 들어, 클라이언트 디바이스 (40) 는 사용자의 FOV 가 도 4 의 이미지들 (238A-D 및 234C-D) 을 포함한다고 결정하고 고품질 및 저품질 비디오 데이터를 이용한 디스플레이를 위해 도 4 의 이미지들 (238A-D 및 234C-D) 을 출력할 수도 있다 (322).

[0158] 이 방식으로, 도 5 의 방법은, 회로부로 구현된, 클라이언트 디바이스의 프로세서에 의해, 비디오 데이터의 이

미지의 복수의 영역들 중 하나 이상의 가장 관심있는 영역들을, 하나 이상의 가장 관심있는 영역들을 나타내는 데이터에 기초하여 결정하는 단계로서, 하나 이상의 가장 관심있는 영역들을 나타내는 데이터는 비디오 데이터를 포함하는 비트스트림의 비디오 코딩 계층 (VCL) 데이터의 외부에 있는, 상기 하나 이상의 가장 관심있는 영역들을 결정하는 단계; 클라이언트 디바이스의 프로세서에 의해, 이미지의 하나 이상의 가장 관심있는 영역들에 대한 상대적으로 고품질 비디오 데이터에 대한 요청을 출력하는 단계; 및 클라이언트 디바이스의 프로세서에 의해, 상대적으로 고품질 비디오 데이터에 대한 요청을 출력한 후, 이미지의 하나 이상의 다른 영역들에 대한 상대적으로 저품질 비디오 데이터에 대한 요청을 출력하는 단계를 포함하는 방법의 예를 표현한다.

[0159] 마찬가지로, 도 5의 방법은 또한, 회로부로 구현된, 소스 디바이스의 프로세서에 의해, 비디오 데이터의 이미지의 복수의 영역들 중 하나 이상의 가장 관심있는 영역들을 결정하는 단계로서, 하나 이상의 가장 관심있는 영역들은 클라이언트 디바이스에 의해 추출될 가능성이 가장 큰 하나 이상의 영역들을 포함하는, 상기 하나 이상의 가장 관심있는 영역들을 결정하는 단계; 및 소스 디바이스의 프로세서에 의해, 하나 이상의 가장 관심있는 영역들을 나타내는 데이터를 생성하는 단계로서, 하나 이상의 가장 관심있는 영역들을 나타내는 데이터는 비디오 데이터를 포함하는 비트스트림의 비디오 코딩 계층 (VCL) 데이터의 외부에 있는, 상기 하나 이상의 가장 관심있는 영역들을 나타내는 데이터를 생성하는 단계를 포함하는 방법의 예를 표현한다.

[0160] 도 6은 본 개시에서 설명된 하나 이상의 기법들에 따른 비디오 데이터에 대한 정보를 시그널링하기 위한 예의 접근법을 예시하는 플로우차트이다. 도 6의 방법은 도 1의 클라이언트 디바이스 (40) 및 서버 디바이스 (60)에 대하여 설명된다. 그러나, 다른 디바이스들이 이들 기법들을 수행하도록 구성될 수도 있는 것으로 이해되어야 한다.

[0161] 초기에, 서버 디바이스 (60)는 이미지의 하나 이상의 가장 관심있는 영역들을 결정할 수도 있다 (330). 예를 들어, 서버 디바이스 (60)는 하나 이상의 가장 관심있는 영역들의 디렉터 또는 프로듀서에 의한 선택의 표시를 수신할 수도 있거나 또는 서버 디바이스 (60) (또는 다른 디바이스)는 사용자 통계로부터 하나 이상의 가장 관심있는 영역들을 도출할 수도 있다. 어느 경우나, 서버 디바이스 (60)는 가장 관심있는 영역들을 나타내는 데이터를 생성할 수도 있다 (332). 가장 관심있는 영역들을 나타내는 데이터는 비디오 데이터를 포함하는 비트스트림의 VCL 데이터의 외부에 있을 수도 있다. 예를 들어, 서버 디바이스 (60)는 가장 관심있는 영역들을 나타내는 데이터를 포함하는 보충 강화 정보 (SEI) 메시지를 생성할 수도 있다. 일부 예들에서, 서버 디바이스 (60)는 가장 관심있는 영역들을 나타내는 데이터를 포함하는, 샘플 그룹 박스 (예를 들어, 도 3의 SBGP 박스 (162))와 같은 파일-포맷 데이터를 생성할 수도 있다. 일부 예들에서, 서버 디바이스 (60)는 가장 관심있는 영역들을 나타내는 데이터를 포함하는 타이밍된 메타데이터 트랙을 생성할 수도 있다.

[0162] 어느 경우나, 서버 디바이스 (60)는 가장 관심있는 영역들을 나타내는 데이터를 전송할 수도 있고 (334) 클라이언트 디바이스 (40)는 가장 관심있는 영역들을 나타내는 데이터를 수신할 수도 있다 (336). 가장 관심있는 영역들을 나타내는 데이터를 수신하는 것에 응답하여, 클라이언트 디바이스 (40)는 하나 이상의 가장 관심있는 영역들에 대한 고품질 비디오 데이터를 프리-페치하기 위한 요청을 출력할 수도 있다 (338). 서버 디바이스 (60)는 하나 이상의 가장 관심있는 영역들에 대한 고품질 비디오 데이터를 프리-페치하기 위한 요청을 수신하고 (340) 하나 이상의 가장 관심있는 영역들에 대한 고품질 비디오 데이터를 전송할 수도 있다 (342). 클라이언트 디바이스 (40)는 고품질 비디오 데이터를 수신할 수도 있다 (344).

[0163] 클라이언트 디바이스 (40)는 이미지의 하나 이상의 다른 영역들에 대한 저품질 비디오 데이터에 대한 요청을 출력할 수도 있다 (346). 서버 디바이스 (60)는 이미지의 하나 이상의 다른 영역들에 대한 저품질 비디오 데이터에 대한 요청을 수신할 수도 있다 (348). 이미지의 하나 이상의 다른 영역들에 대한 저품질 비디오 데이터에 대한 요청을 수신하는 것에 응답하여, 서버 디바이스 (60)는 이미지의 하나 이상의 다른 영역들에 대한 저품질 비디오 데이터를 전송할 수도 있다 (350). 클라이언트 디바이스 (40)는 저품질 비디오 데이터를 수신할 수도 있다 (352). 도 6의 예에 도시하지 않았지만, 클라이언트 디바이스 (40)는 그 후, 예를 들어, 사용자의 시야에 기초하여, 예를 들어, 도 5에 대하여 상기 논의한 바와 같이, 사용자에게 수신된 비디오 데이터의 적어도 파트를 제시하도록 진행할 수도 있다.

[0164] 이 방식으로, 도 6의 방법은, 회로부로 구현된, 클라이언트 디바이스의 프로세서에 의해, 비디오 데이터의 이미지의 복수의 영역들 중 하나 이상의 가장 관심있는 영역들을, 하나 이상의 가장 관심있는 영역들을 나타내는 데이터에 기초하여 결정하는 단계로서, 하나 이상의 가장 관심있는 영역들을 나타내는 데이터는 비디오 데이터를 포함하는 비트스트림의 비디오 코딩 계층 (VCL) 데이터의 외부에 있는, 상기 하나 이상의 가장 관심있는 영

역들을 결정하는 단계; 클라이언트 디바이스의 프로세서에 의해, 이미지의 하나 이상의 가장 관심있는 영역들에 대한 상대적으로 고품질 비디오 데이터에 대한 요청을 출력하는 단계; 및 클라이언트 디바이스의 프로세서에 의해, 상대적으로 고품질 비디오 데이터에 대한 요청을 출력한 후, 이미지의 하나 이상의 다른 영역들에 대한 상대적으로 저품질 비디오 데이터에 대한 요청을 출력하는 단계를 포함하는 방법의 예를 표현한다.

[0165] 마찬가지로, 도 6의 방법은 또한, 회로부로 구현된, 소스 디바이스의 프로세서에 의해, 비디오 데이터의 이미지의 복수의 영역들 중 하나 이상의 가장 관심있는 영역들을 결정하는 단계로서, 하나 이상의 가장 관심있는 영역들은 클라이언트 디바이스에 의해 추출될 가능성이 가장 큰 하나 이상의 영역들을 포함하는, 상기 하나 이상의 가장 관심있는 영역들을 결정하는 단계; 및 소스 디바이스의 프로세서에 의해, 하나 이상의 가장 관심있는 영역들을 나타내는 데이터를 생성하는 단계로서, 하나 이상의 가장 관심있는 영역들을 나타내는 데이터는 비디오 데이터를 포함하는 비트스트림의 비디오 코딩 계층 (VCL) 데이터의 외부에 있는, 상기 하나 이상의 가장 관심있는 영역들을 나타내는 데이터를 생성하는 단계를 포함하는 방법의 예를 표현한다.

[0166] 도 7은 본 개시에서 설명된 하나 이상의 기법들에 따른 비디오 데이터를 프리-페치하기 위한 예의 접근법을 예시하는 플로우차트이다. 도 7의 방법은 도 1의 클라이언트 디바이스 (40) 및 서버 디바이스 (60)에 대하여 설명된다. 그러나, 다른 디바이스들이 이들 기법들을 수행하도록 구성될 수도 있는 것으로 이해되어야 한다. 도 7의 예에서, 서버 디바이스 (60)는 근원 서버로부터 데이터를 초기에 추출하고 그 데이터를 일시적으로 저장하는 CDN 서버로서 구성될 수도 있다. 도 7의 예는 클라이언트 디바이스 (40) 및 서버 디바이스 (60) 중 어느 하나 또는 양자 모두가 가장 관심있는 영역들을 나타내는 데이터를 이용할 수도 있다는 것을 입증하는 것으로 이해되어야 한다.

[0167] 초기에, 서버 디바이스 (60)는 이미지의 가장 관심있는 영역들을 나타내는 데이터를 수신할 수도 있다 (360). 가장 관심있는 영역들을 나타내는 데이터는 비디오 데이터를 포함하는 비트스트림의 VCL 데이터의 외부에 있을 수도 있다. 예를 들어, 서버 디바이스 (60)는 가장 관심있는 영역들을 나타내는 데이터를 포함하는 보충 강화 정보 (SEI) 메시지를 수신한다. 일부 예들에서, 서버 디바이스 (60)는 가장 관심있는 영역들을 나타내는 데이터를 포함하는 샘플 그룹 박스 (예를 들어, 도 3의 SGBP 박스 (162))를 수신한다. 일부 예들에서, 서버 디바이스 (60)는 가장 관심있는 영역들을 나타내는 데이터를 포함하는 타이밍된 메타데이터 트랙을 수신한다. 어느 경우나, 서버 디바이스 (60)는 하나 이상의 관심있는 영역들에 대한 고품질 비디오 데이터를 프리-페치할 수도 있다 (362). 예를 들어, 서버 디바이스 (60)는 하나 이상의 가장 관심있는 영역들을 이용하여 하나 이상의 가장 관심있는 영역들에 대한 비디오 데이터를 프리-페치할 수도 있다. 예를 들어, 서버 디바이스 (60)는 이미지의 다른 영역들에 대한 비디오 데이터보다 저레이턴시로 및/또는 고품질로 하나 이상의 가장 관심있는 영역들에 대한 비디오 데이터를 페치할 수도 있다.

[0168] 하나 이상의 관심있는 영역들에 대한 고품질 비디오 데이터를 프리-페치 시에, 서버 디바이스 (60)는 하나 이상의 가장 관심있는 영역들에 대한 프리-페치된, 고품질 비디오 데이터를 저장할 수도 있다 (364). 서버 디바이스 (60)가 비디오 데이터를 저장 (예를 들어, 캐시)하도록 구성된 예지 서버인 예에서, 서버 디바이스 (60)는 고품질 비디오 데이터를 저장할 수도 있다. 일부 예들에서, 서버 디바이스 (60)는 이미지의 가장 관심있는 영역들을 나타내는 데이터에 기초하여 하나 이상의 가장 관심있는 영역들에 대한 고품질 비디오 데이터에 대한 저장의 지속시간을 결정할 수도 있다. 예를 들어, 서버 디바이스 (60)는 고품질 비디오 데이터가 가장 관심있는 영역들에 대한 것이라고 서버 디바이스 (60)가 결정할 때 고품질 비디오 데이터를 보다 높은 것으로 랭크할 수도 있어, 가장 관심있는 영역들에 대한 고품질 비디오 데이터가 장기 저장을 위해 유지되는 한편, 가장 관심있는 영역들에 대한 고품질 비디오 데이터와 유사한 시간에 추출된 다른 비디오 데이터 (예를 들어, 다른 영역들에 대한 것임)는 제거된다.

[0169] 클라이언트 디바이스 (40)는 하나 이상의 가장 관심있는 영역들에 대한 고품질 비디오 데이터를 프리-페치하기 위한 요청을 출력할 수도 있다 (366). 예를 들어, 클라이언트 디바이스 (40)는 이미지의 다른 영역들과 비교하여 상대적으로 높은 비트레이트를 이용하여 하나 이상의 가장 관심있는 영역들에 대한 비디오 데이터를 요청할 수도 있다. 서버 디바이스 (60)는 프리-페치 요청을 수신할 수도 있고 (368) 하나 이상의 가장 관심있는 영역들에 대한 고품질 비디오 데이터를 전송할 수도 있다 (370). 클라이언트 디바이스 (40)는 고품질 비디오 데이터를 수신할 수도 있다 (372).

[0170] 클라이언트 디바이스 (40)는 이미지의 하나 이상의 다른 영역들에 대한 저품질 비디오 데이터에 대한 요청을 출력할 수도 있다 (374). 서버 디바이스 (60)는 이미지의 하나 이상의 다른 영역들에 대한 저품질 비디오 데이터에 대한 요청을 수신할 수도 있다 (376). 이미지의 하나 이상의 다른 영역들에 대한 저품질 비디오

데이터에 대한 요청을 수신하는 것에 응답하여, 서버 디바이스 (60) 는 이미지의 하나 이상의 다른 영역들에 대한 저품질 비디오 데이터를 폐치하고 (378) 이미지의 하나 이상의 다른 영역들에 대한 저품질 비디오 데이터를 전송할 수도 있다 (380). 클라이언트 디바이스 (40) 는 저품질 비디오 데이터를 수신할 수도 있다 (382).

도 7 의 예에 도시하지 않았지만, 클라이언트 디바이스 (40) 는 그 후, 예를 들어, 사용자의 시야에 기초하여, 예를 들어, 도 5 에 대하여 상기 논의한 바와 같이, 사용자에게 수신된 비디오 데이터의 적어도 파트를 제시하도록 진행할 수도 있다.

[0171] 이 방식으로, 도 7 의 방법은, 회로부로 구현된, 클라이언트 디바이스의 프로세서에 의해, 비디오 데이터의 이미지의 복수의 영역들 중 하나 이상의 가장 관심있는 영역들을, 하나 이상의 가장 관심있는 영역들을 나타내는 데이터에 기초하여 결정하는 단계로서, 하나 이상의 가장 관심있는 영역들을 나타내는 데이터는 비디오 데이터를 포함하는 비트스트림의 비디오 코딩 계층 (VCL) 데이터의 외부에 있는, 상기 하나 이상의 가장 관심있는 영역들을 결정하는 단계; 클라이언트 디바이스의 프로세서에 의해, 이미지의 하나 이상의 가장 관심있는 영역들에 대한 상대적으로 고품질 비디오 데이터에 대한 요청을 출력하는 단계; 및 클라이언트 디바이스의 프로세서에 의해, 상대적으로 고품질 비디오 데이터에 대한 요청을 출력한 후, 이미지의 하나 이상의 다른 영역들에 대한 상대적으로 저품질 비디오 데이터에 대한 요청을 출력하는 단계를 포함하는 방법의 예를 표현한다.

[0172] 마찬가지로, 도 7 의 방법은 또한, 회로부로 구현된 소스 디바이스의 프로세서에 의해, 비디오 데이터의 이미지의 복수의 영역들 중 하나 이상의 가장 관심있는 영역들을 결정하는 단계로서, 하나 이상의 가장 관심있는 영역들은 클라이언트 디바이스에 의해 추출될 가능성이 가장 큰 하나 이상의 영역들을 포함하는, 상기 하나 이상의 가장 관심있는 영역들을 결정하는 단계; 및 소스 디바이스의 프로세서에 의해, 하나 이상의 가장 관심있는 영역들을 나타내는 데이터를 생성하는 단계로서, 하나 이상의 가장 관심있는 영역들을 나타내는 데이터는 비디오 데이터를 포함하는 비트스트림의 비디오 코딩 계층 (VCL) 데이터의 외부에 있는, 상기 하나 이상의 가장 관심있는 영역들을 나타내는 데이터를 생성하는 단계를 포함하는 방법의 예를 표현한다.

[0173] 하나 이상의 예들에서, 설명된 기능들은 하드웨어, 소프트웨어, 펌웨어, 또는 그 임의의 조합으로 구현될 수도 있다. 소프트웨어로 구현되면, 기능들은 컴퓨터 판독가능 매체 상에 하나 이상의 명령들 또는 코드로서 저장 또는 이를 통해 송신되고 하드웨어 기반 프로세싱 유닛에 의해 실행될 수도 있다. 컴퓨터 판독가능 매체들은 데이터 저장 매체들과 같은 유형의 매체에 대응하는 컴퓨터 판독가능 저장 매체들, 또는 예를 들어, 통신 프로토콜에 따라 일 장소로부터 다른 장소로의 컴퓨터 프로그램의 전송을 용이하게 하는 임의의 매체를 포함하는 통신 매체들을 포함할 수도 있다. 이 방식으로, 컴퓨터 판독가능 매체들은 일반적으로 (1) 비일시적인 (non-transitory) 유형의 컴퓨터 판독가능 저장 매체들 또는 (2) 신호 또는 캐리어파와 같은 통신 매체에 대응할 수도 있다. 데이터 저장 매체들은 본 개시에서 설명된 기법들의 구현을 위해 명령들, 코드, 및/또는 데이터 구조들을 추출하기 위해 하나 이상의 컴퓨터들 또는 하나 이상의 프로세서들에 의해 액세스될 수 있는 임의의 이용가능한 매체들일 수도 있다. 컴퓨터 프로그램 제품은 컴퓨터 판독가능 매체를 포함할 수도 있다.

[0174] 제한이 아닌 일 예로, 이러한 컴퓨터 판독가능 저장 매체들은 RAM, ROM, EEPROM, CD-ROM 또는 다른 광학 디스크 스토리지, 자기 디스크 스토리지, 또는 다른 자기 저장 디바이스들, 플래시 메모리, 또는 명령들 또는 데이터 구조들의 형태로 원하는 프로그램 코드를 저장하는데 이용될 수 있고 컴퓨터에 의해 액세스될 수 있는 임의의 다른 매체를 포함할 수 있다. 또한, 임의의 커넥션이 적절히 컴퓨터 판독가능 매체라 불린다. 예를 들어, 명령들이 웹사이트, 서버, 또는 다른 원격 소스로부터 동축 케이블, 광섬유 케이블, 트위스티드 페어, 디지털 가입자 회선 (DSL), 또는 적외선, 무선, 및 마이크로파와 같은 무선 기술들을 이용하여 송신되면, 매체의 정의에는 동축 케이블, 광섬유 케이블, 트위스티드 페어, DSL, 또는 적외선, 무선, 및 마이크로파와 같은 무선 기술들이 포함된다. 그러나, 컴퓨터 판독가능 저장 매체들 및 데이터 저장 매체들은 커넥션들, 캐리어파들, 신호들, 또는 다른 일시적 매체들을 포함하지 않고, 대신에 비일시적, 유형의 저장 매체들과 관련되는 것으로 이해되어야 한다. 디스크 (disk) 및 디스크 (disc) 는, 본 명세서에서 사용한 바와 같이, 콤팩트 디스크 (CD), 레이저 디스크, 광학 디스크, 디지털 다기능 디스크 (DVD), 플로피 디스크 및 Blu-ray 디스크를 포함하고, 여기서 디스크 (disk) 들은 보통 데이터를 자기적으로 재생하는 한편, 디스크 (disc) 들은 레이저들로 데이터를 광학적으로 재생한다. 상기의 조합들이 또한 컴퓨터 판독가능 매체들의 범위 내에 포함되어야 한다.

[0175] 명령들은 하나 이상의 프로세서들, 이를 테면 하나 이상의 디지털 신호 프로세서들 (DSP들), 범용 마이크로프로세서들, 주문형 집적 회로들 (ASIC들), 필드 프로그래밍가능 로직 어레이들 (FPGA들), 또는 다른 등가의 통합된 또는 별개의 로직 회로부에 의해 실행될 수도 있다. 이에 따라, 용어 "프로세서" 는, 본 명세서에서 사용한 바와 같이, 본 명세서에서 설명된 기법들의 구현에 적합한 전술한 구조 또는 임의의 다른 구조 중 임의의 것을

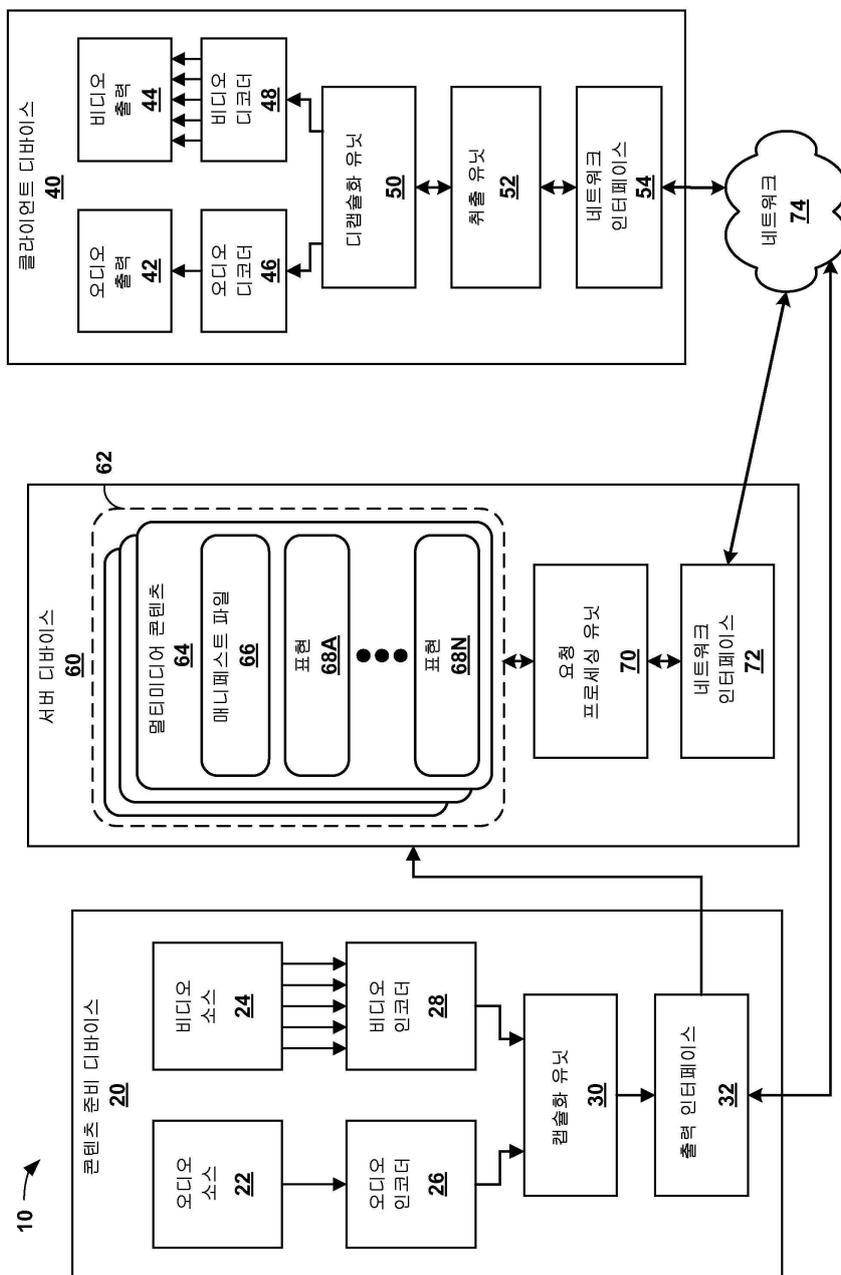
지칭할 수도 있다. 추가로, 일부 양태들에서, 본 명세서에서 설명된 기능성은 인코딩 및 디코딩을 위해 구성된 전용 하드웨어 및/또는 소프트웨어 모듈들 내에서 제공되거나, 또는 결합된 코덱에 통합될 수도 있다. 또한, 기법들은 하나 이상의 회로들 또는 로직 엘리먼트들에서 완전히 구현될 수 있다.

[0176] 본 개시의 기법들은 무선 핸드셋, 집적 회로 (IC) 또는 IC들의 세트 (예를 들어, 칩 세트) 를 포함하는 광범위한 디바이스들 또는 장치들로 구현될 수도 있다. 다양한 컴포넌트들, 모듈들, 또는 유닛들은 본 개시에서, 개시된 기법들을 수행하도록 구성된 디바이스들의 기능적 양태들을 강조하도록 설명되지만, 상이한 하드웨어 유닛들에 의한 실현을 반드시 요구하는 것은 아니다. 오히려, 상기 설명한 바와 같이, 다양한 유닛들은 코덱 하드웨어 유닛에서 결합되거나 또는 적합한 소프트웨어 및/또는 펌웨어와 함께, 상기 설명한 바와 같은 하나의 이상의 프로세서들을 포함하는, 상호운용적 하드웨어 유닛들의 콜렉션에 의해 제공될 수도 있다.

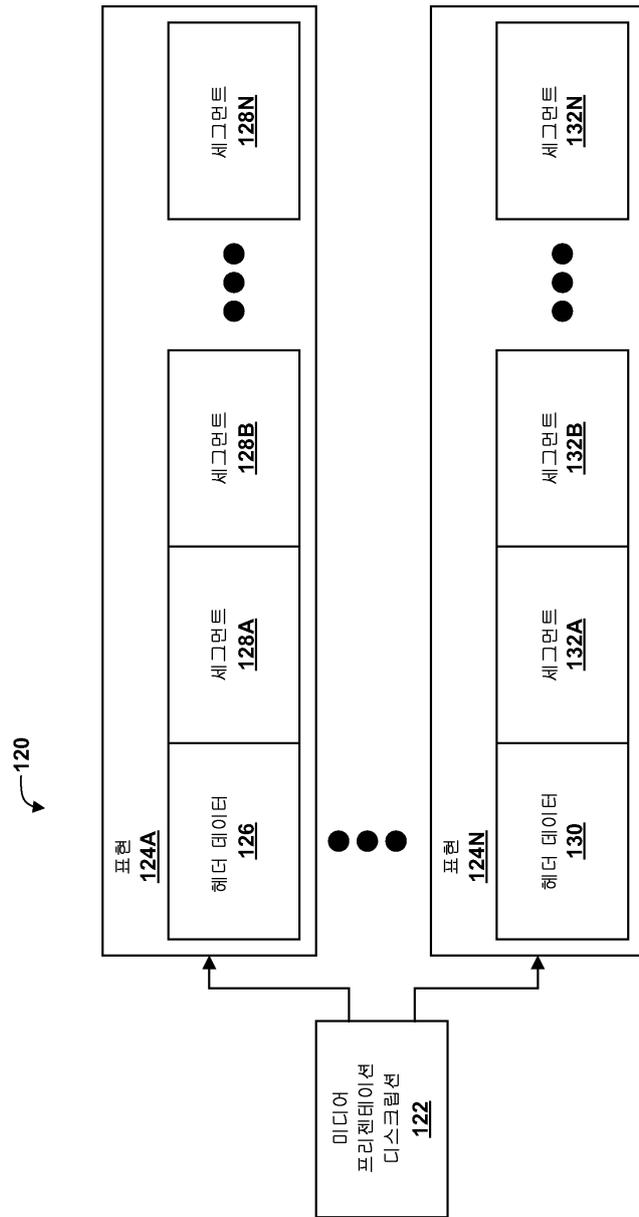
[0177] 다양한 예들이 설명되었다. 이들 및 다른 예들은 다음의 청구항들의 범위 내에 있다.

도면

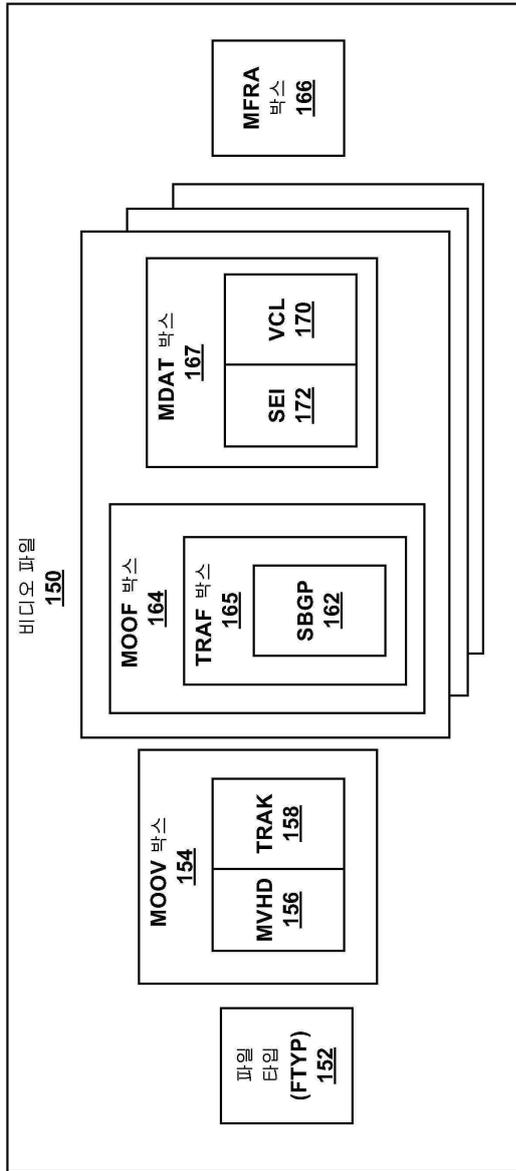
도면1



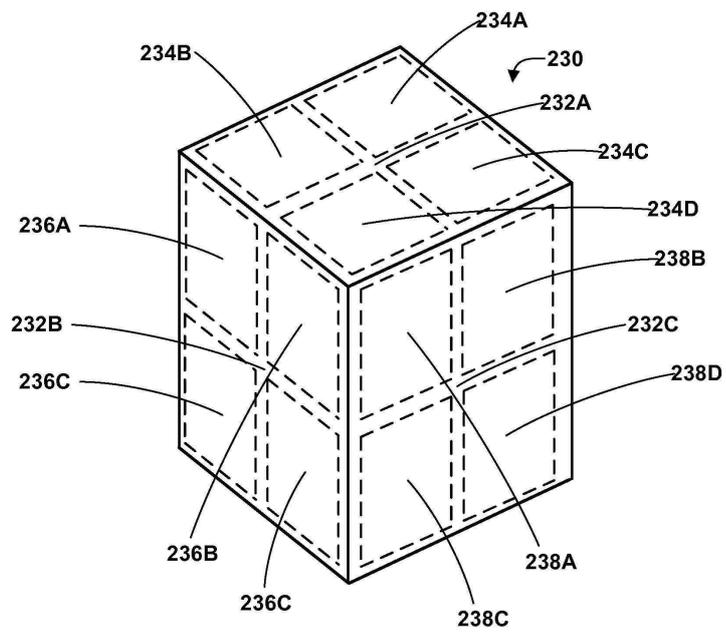
도면2



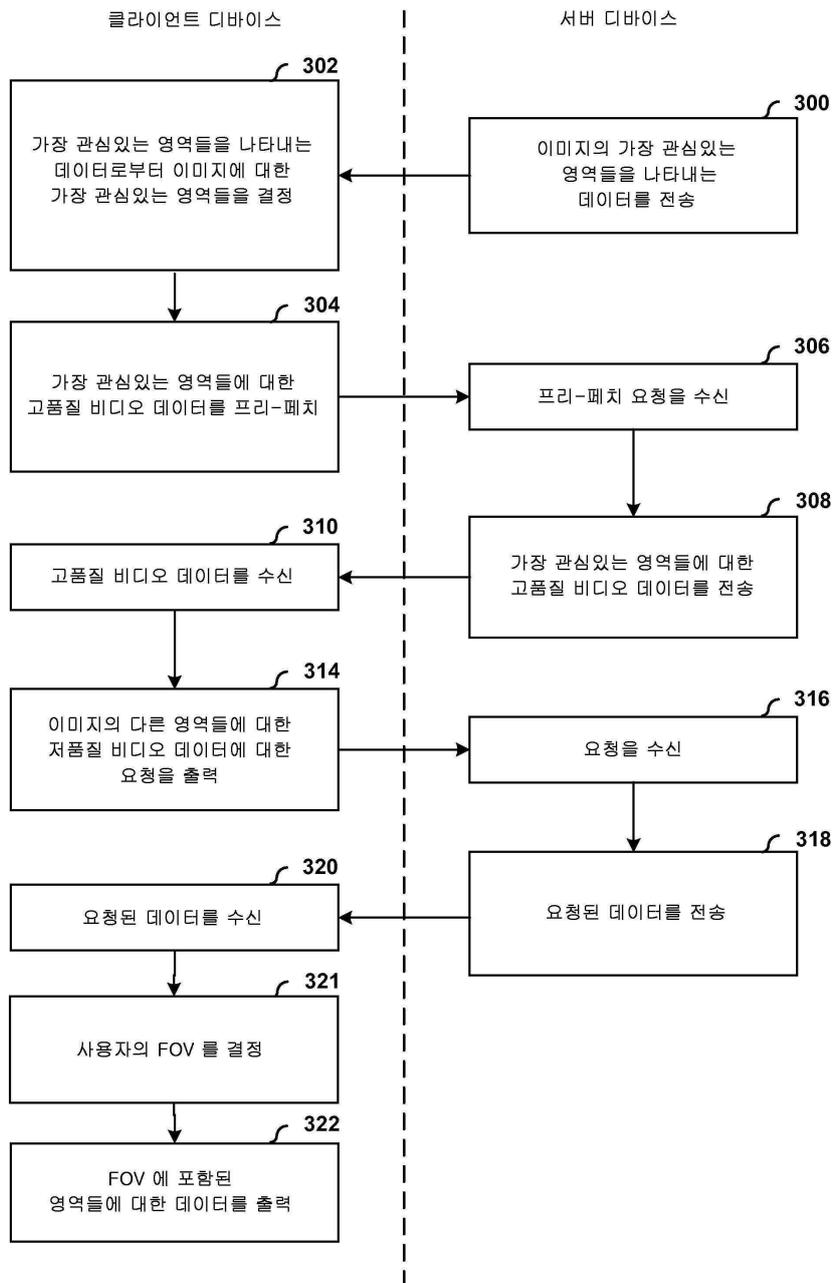
도면3



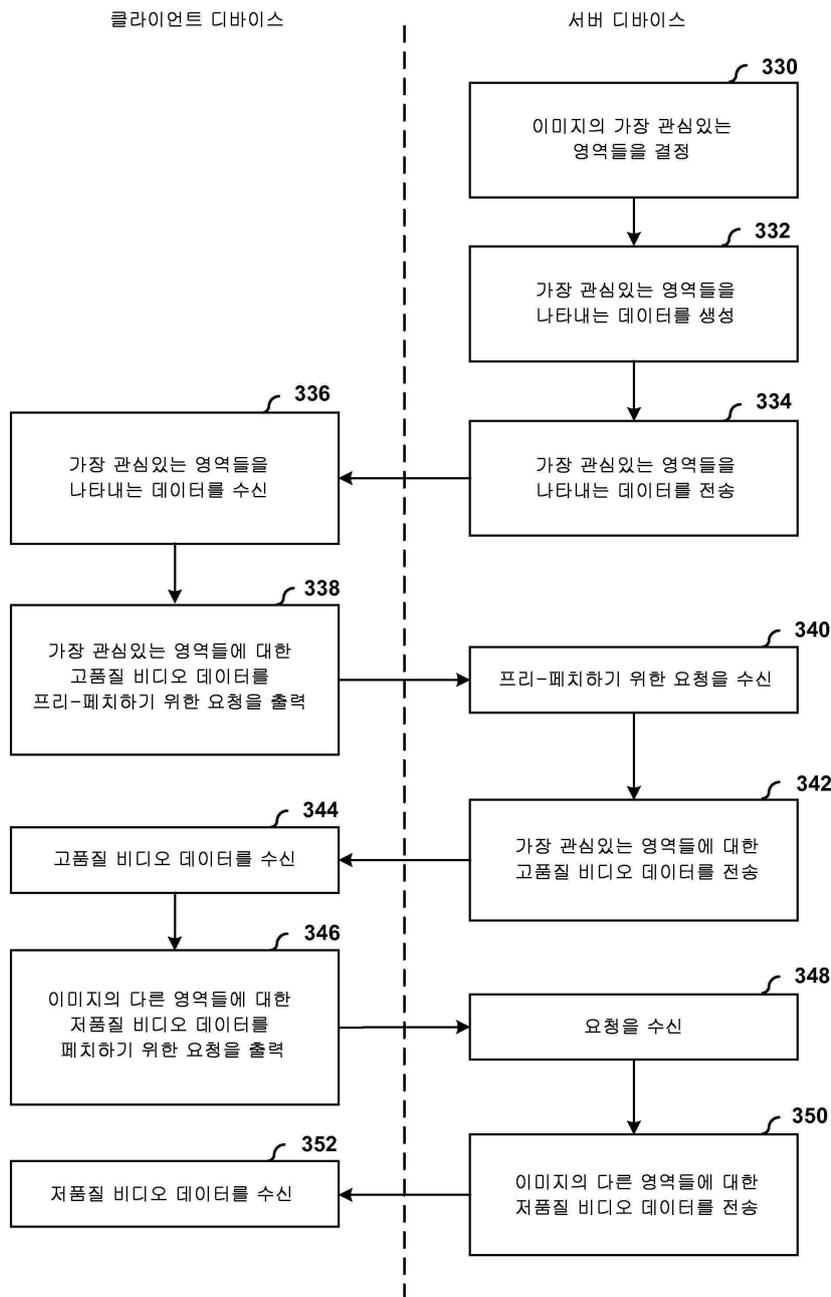
도면4



도면5



도면6



도면7

