



(19) 대한민국특허청(KR)  
(12) 등록특허공보(B1)

(45) 공고일자 2014년03월24일  
(11) 등록번호 10-1377928  
(24) 등록일자 2014년03월18일

- (51) 국제특허분류(Int. Cl.)  
H04N 19/70 (2014.01) H04N 13/00 (2006.01)
- (21) 출원번호 10-2012-7000403
- (22) 출원일자(국제) 2010년06월05일  
심사청구일자 2012년01월05일
- (85) 번역문제출일자 2012년01월05일
- (65) 공개번호 10-2012-0024947
- (43) 공개일자 2012년03월14일
- (86) 국제출원번호 PCT/US2010/037554
- (87) 국제공개번호 WO 2010/141927  
국제공개일자 2010년12월09일
- (30) 우선권주장  
12/643,632 2009년12월21일 미국(US)  
61/184,649 2009년06월05일 미국(US)
- (56) 선행기술조사문헌

Guofeng Zhang et al., "Stereoscopic Video Synthesis from a Monocular Video", IEEE Transactions on Visualization and Computer Graphics, Vol.13, Issue 4, 2007, pp.686-696.

Kimata, H. et al., "Real-time MVC viewer for free viewpoint navigation", IEEE International Conference on Multimedia and Expo, 2008, pp.1437-1440.

- (73) 특허권자  
퀄컴 인코포레이티드  
미국 92121-1714 캘리포니아주 샌 디에고 모어하우스 드라이브 5775
- (72) 발명자  
천 잉  
미국 92121-1714 캘리포니아주 샌디에고 모어하우스 드라이브 5775
- 카르체비츠 마르타  
미국 92121-1714 캘리포니아주 샌디에고 모어하우스 드라이브 5775
- (74) 대리인  
특허법인코리어나

전체 청구항 수 : 총 50 항

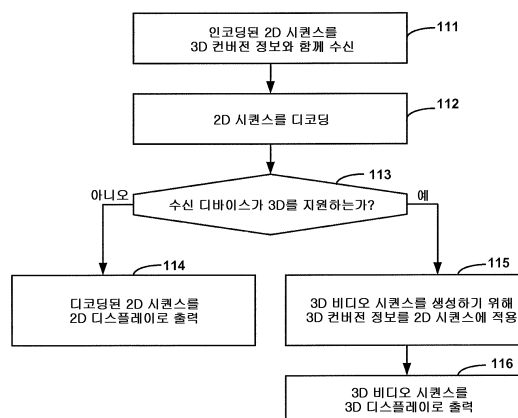
심사관 : 권오성

(54) 발명의 명칭 2-차원 비디오 시퀀스를 이용한 3-차원 컨버전 정보의 인코딩

(57) 요약

본 개시물은 비디오 프레임들의 2-차원 (2D) 비디오 시퀀스를 3-차원 (3D) 컨버전 정보와 함께 인코딩하기 위한 기술들을 설명하며, 3D 컨버전 정보는 2D 시퀀스의 비디오 프레임들 각각에 적용되어 3D 비디오 데이터를 생성할 수 있는 파라미터 세트를 포함한다. 파라미터 세트는 2D 시퀀스의 오리지널 비디오 프레임들 각각에 적용되어 오리지널 비디오 프레임들의 각각에 대한 2차 뷰들을 생성할 수 있는 비교적 소량의 데이터를 포함할 수도 있다. 오리지널 비디오 프레임들 및 2차 뷰 비디오 프레임들은 집합적으로 입체적인 3D 비디오 시퀀스를 정의할 수도 있다. 2D 시퀀스 및 파라미터 세트는 3D 시퀀스를 통신하기 위해 요구되는 것보다 상당히 적은 데이터를 포함할 수도 있다. 본 개시물은 또한 효과적이고 효율적인 방식으로 파라미터 세트를 인코딩하기 위해 이용될 수도 있는 일부 예시적인 신택스 (syntax)를 설명한다.

대표도 - 도8



## 특허청구의 범위

### 청구항 1

비디오 인코더에서 비디오 프레임들의 제 1 2-차원 (2D) 시퀀스를 인코딩하는 단계;

상기 비디오 인코더를 통해 3-차원 (3D) 컨버전 정보를 인코딩하는 단계로서, 상기 3D 컨버전 정보는 상기 제 1 2D 시퀀스로부터 유도되어, 상기 3D 컨버전 정보가, 상기 제 1 2D 시퀀스의 상기 비디오 프레임들의 각각에 적용되어, 상기 3D 컨버전 정보와 상기 제 1 2D 시퀀스로부터 3D 비디오 데이터의 입체적인 비디오 시퀀스가 유도 가능하도록 함으로써 상기 제 1 2D 시퀀스와 함께 상기 3D 비디오 데이터의 입체적인 비디오 시퀀스를 정의하는 제 2 2D 시퀀스를 생성할 수 있는 파라미터 세트를 포함하도록 하는, 상기 3D 컨버전 정보를 인코딩하는 단계; 및

상기 인코딩된 제 1 2D 시퀀스를 상기 3D 컨버전 정보와 함께 통신하는 단계를 포함하는, 방법.

### 청구항 2

제 1 항에 있어서,

상기 제 1 2D 시퀀스를 인코딩하는 단계는, ITU H.264 비디오 인코딩 표준에 따라 상기 제 1 2D 시퀀스를 인코딩하는 단계를 포함하고, 그리고

상기 3D 컨버전 정보를 인코딩하는 단계는 상기 3D 컨버전 정보를 상기 ITU H.264 비디오 인코딩 표준에 의해 지원되는 하나 이상의 SEI (Supplemental Enhancement Information) 메시지들로 인코딩하는 단계를 포함하는, 방법.

### 청구항 3

삭제

### 청구항 4

제 1 항에 있어서,

상기 3D 컨버전 정보는, 상기 제 1 2D 시퀀스에 적용되어 상기 3D 비디오 데이터를 생성하기 위한 3D 컨버전 프로세스를 식별하는 정보를 포함하는, 방법.

### 청구항 5

제 1 항에 있어서,

상기 3D 컨버전 정보는 상기 제 1 2D 시퀀스를 캡처하는 것과 연관된 카메라 파라미터들 및 값들을 포함하는, 방법.

### 청구항 6

제 1 항에 있어서,

상기 3D 컨버전 정보는,

상기 제 1 2D 시퀀스를 캡처한 카메라와 연관된 초점 길이를 표시하는 초점 길이 값,

상기 3D 비디오 데이터에서의 최소 깊이를 명시하는 근거리-깊이 (near-depth) 값,

상기 3D 비디오 데이터에서의 최대 깊이를 명시하는 원거리-깊이 (far-depth) 값, 및

상기 3D 비디오 데이터와 연관된 2 개의 카메라들 사이에서 추정된 거리를 정량화하는 평행 이동 (translate) 값을 포함하는, 방법.

### 청구항 7

제 1 항에 있어서,  
 상기 3D 컨버전 정보는,  
 상기 제 1 2D 시퀀스를 캡처한 카메라와 연관된 초점 길이를 표시하는 초점 길이 값,  
 상기 3D 비디오 데이터에서의 최소 깊이를 명시하는 근거리-깊이 값,  
 상기 3D 비디오 데이터에서의 최대 깊이를 명시하는 원거리-깊이 값,  
 상기 3D 비디오 데이터와 연관된 2 개의 카메라들 사이에서 추정된 거리를 정량화하는 평행 이동 값, 및  
 수렴 이미지 평면의 상기 2 개의 카메라들로의 추정된 거리를 정량화하는 수렴-깊이값을 포함하는, 방법.

#### 청구항 8

제 1 항에 있어서,  
 상기 3D 컨버전 정보는 명시적인 3D 파라미터 세트가 상기 3D 컨버전 정보 내에 포함되는지 여부 또는 디폴트 3D 파라미터 세트가 이용될 것인지 여부를 나타내는 플래그를 포함하는, 방법.

#### 청구항 9

제 1 항에 있어서,  
 상기 3D 컨버전 정보는 상기 제 1 2D 시퀀스의 제 2 뷰가 상기 제 1 2D 시퀀스의 좌측에 생성될 것인지 또는 우측에 생성될 것인지 여부를 나타내는 플래그를 포함하는, 방법.

#### 청구항 10

제 1 항에 있어서,  
 상기 3D 컨버전 정보는 상기 3D 비디오 데이터로부터 제거될 크롭 (crop) 영역을 식별하는 플래그를 포함하는, 방법.

#### 청구항 11

제 1 항에 있어서,  
 상기 3D 컨버전 정보는,  
 명시적인 3D 파라미터 세트가 상기 3D 컨버전 정보 내에 포함되는지 여부 또는 디폴트 3D 파라미터 세트가 이용될 것인지 여부를 나타내는 제 1 플래그로서, 상기 명시적인 3D 파라미터 세트는 상기 제 1 플래그가 설정되는 경우에 상기 3D 컨버전 정보 내에 포함되는, 상기 제 1 플래그,  
 상기 제 1 2D 시퀀스의 제 2 뷰가 상기 제 1 2D 시퀀스의 좌측에 생성될 것인지 또는 우측에 생성될 것인지 여부를 나타내는 제 2 플래그, 및  
 상기 3D 비디오 데이터로부터 제거될 크롭 영역을 식별하는 제 3 플래그를 포함하며,  
 상기 크롭 영역을 정의하는 정보는 상기 제 3 플래그가 설정되는 경우에 상기 3D 컨버전 정보 내에 포함되는, 방법.

#### 청구항 12

비디오 디코더에서 비디오 프레임들의 제 1 2-차원 (2D) 시퀀스를 수신하는 단계;  
 상기 비디오 디코더에서 상기 제 1 2D 시퀀스와 함께 3-차원 (3D) 컨버전 정보를 수신하는 단계로서, 상기 3D 컨버전 정보는 상기 제 1 2D 시퀀스로부터 유도되어, 상기 3D 컨버전 정보가, 상기 제 1 2D 시퀀스의 상기 비디오 프레임들의 각각에 적용되어 상기 제 1 2D 시퀀스와 함께 3D 비디오 데이터의 입체적인 비디오 시퀀스를 정의하는 제 2 2D 시퀀스를 생성할 수 있는 파라미터 세트를 포함하도록 하는, 상기 3D 컨버전 정보를 수신하는 단계;  
 상기 비디오 디코더를 통해 상기 제 1 2D 시퀀스를 디코딩하는 단계; 및

상기 제 1 2D 시퀀스 및 상기 3D 컨버전 정보에 기초하여 상기 비디오 디코더를 통해 상기 3D 비디오 데이터를 생성하는 단계를 포함하는, 방법.

#### 청구항 13

제 12 항에 있어서,

상기 제 1 2D 시퀀스는 ITU H.264 비디오 인코딩 표준에 따라 인코딩되고, 그리고

상기 3D 컨버전 정보는 상기 ITU H.264 비디오 인코딩 표준에 의해 지원되는 하나 이상의 SEI (Supplemental Enhancement Information) 메시지들에서 수신되는, 방법.

#### 청구항 14

삭제

#### 청구항 15

제 12 항에 있어서,

상기 3D 컨버전 정보는, 상기 제 1 2D 시퀀스에 적용되어 상기 3D 비디오 데이터를 생성하기 위한 3D 컨버전 프로세스를 식별하는 정보를 포함하는, 방법.

#### 청구항 16

제 12 항에 있어서,

상기 3D 컨버전 정보는 상기 제 1 2D 시퀀스를 캡처하는 것과 연관된 카메라 파라미터들 및 값들을 포함하는, 방법.

#### 청구항 17

제 12 항에 있어서,

상기 3D 컨버전 정보는,

상기 제 1 2D 시퀀스를 캡처한 카메라와 연관된 초점 길이를 표시하는 초점 길이 값,

상기 3D 비디오 데이터에서의 최소 깊이를 명시하는 근거리-깊이 값,

상기 3D 비디오 데이터에서의 최대 깊이를 명시하는 원거리-깊이 값, 및

상기 3D 비디오 데이터와 연관된 2 개의 카메라들 사이에서 추정된 거리를 정량화하는 평행 이동 값을 포함하는, 방법.

#### 청구항 18

제 12 항에 있어서,

상기 3D 컨버전 정보는,

상기 제 1 2D 시퀀스를 캡처한 카메라와 연관된 초점 길이를 표시하는 초점 길이 값,

상기 3D 비디오 데이터에서의 최소 깊이를 명시하는 근거리-깊이 값,

상기 3D 비디오 데이터에서의 최대 깊이를 명시하는 원거리-깊이 값,

상기 3D 비디오 데이터와 연관된 2 개의 카메라들 사이에서 추정된 거리를 정량화하는 평행 이동 값, 및

수렴 이미지 평면의 상기 2 개의 카메라들로의 추정된 거리를 정량화하는 수렴-깊이 값을 포함하는, 방법.

#### 청구항 19

제 12 항에 있어서,

상기 3D 컨버전 정보는 명시적인 3D 파라미터 세트가 상기 3D 컨버전 정보 내에 포함되는지 여부 또는 디폴트

3D 파라미터 세트가 이용될 것인지 여부를 나타내는 플래그를 포함하는, 방법.

#### 청구항 20

제 12 항에 있어서,

상기 3D 컨버전 정보는 상기 제 1 2D 시퀀스의 제 2 뷰가 상기 제 1 2D 시퀀스의 좌측에 생성될 것인지 또는 우측에 생성될 것인지 여부를 나타내는 플래그를 포함하는, 방법.

#### 청구항 21

제 12 항에 있어서,

상기 3D 컨버전 정보는 상기 3D 비디오 데이터로부터 제거될 크롭 영역을 식별하는 플래그를 포함하는, 방법.

#### 청구항 22

제 12 항에 있어서,

상기 3D 컨버전 정보는,

명시적인 3D 파라미터 세트가 상기 3D 컨버전 정보 내에 포함되는지 여부 또는 디폴트 3D 파라미터 세트가 이용될 것인지 여부를 나타내는 제 1 플래그로서, 상기 명시적인 3D 파라미터 세트는 상기 제 1 플래그가 설정되는 경우에 상기 3D 컨버전 정보 내에 포함되는, 상기 제 1 플래그,

상기 제 1 2D 시퀀스의 제 2 뷰가 상기 제 1 2D 시퀀스의 좌측에 생성될 것인지 또는 우측에 생성될 것인지 여부를 나타내는 제 2 플래그, 및

상기 3D 비디오 데이터로부터 제거될 크롭 영역을 식별하는 제 3 플래그를 포함하며,

상기 크롭 영역을 정의하는 정보는 상기 제 3 플래그가 설정되는 경우에 상기 3D 컨버전 정보 내에 포함되는, 방법.

#### 청구항 23

제 12 항에 있어서,

수신 디바이스가 상기 3D 비디오 데이터를 생성하고 렌더링할 수 있는지 여부를 결정하는 단계;

상기 수신 디바이스가 상기 3D 비디오 데이터를 생성하고 렌더링할 수 있는 경우에, 상기 제 1 2D 시퀀스 및 상기 3D 컨버전 정보에 기초하여 상기 3D 비디오 데이터를 생성하고 렌더링하는 단계; 및

상기 수신 디바이스가 상기 3D 비디오 데이터를 생성하거나 렌더링할 수 없는 경우에, 상기 제 1 2D 시퀀스를 렌더링하는 단계를 더 포함하는, 방법.

#### 청구항 24

비디오 프레임들의 제 1 2-차원 (2D) 시퀀스를 인코딩하고, 상기 제 1 2D 시퀀스와 함께 3-차원 (3D) 컨버전 정보를 인코딩하는 비디오 인코더를 포함하며,

상기 3D 컨버전 정보는 상기 제 1 2D 시퀀스로부터 유도되어, 상기 3D 컨버전 정보가, 상기 제 1 2D 시퀀스의 상기 비디오 프레임들의 각각에 적용되어, 상기 3D 컨버전 정보와 상기 제 1 2D 시퀀스로부터 3D 비디오 데이터의 입체적인 비디오 시퀀스가 유도가능하도록 함으로써 상기 제 1 2D 시퀀스와 함께 상기 3D 비디오 데이터의 입체적인 비디오 시퀀스를 정의하는 제 2 2D 시퀀스를 생성할 수 있는 파라미터 세트를 포함하도록 하는, 장치.

#### 청구항 25

제 24 항에 있어서,

상기 인코딩된 제 1 2D 시퀀스를 상기 3D 컨버전 정보와 함께 다른 디바이스로 통신하는 송신기를 더 포함하는, 장치.

#### 청구항 26

제 24 항에 있어서,

상기 비디오 인코더는 ITU H.264 비디오 인코딩 표준에 따라 상기 제 1 2D 시퀀스를 인코딩하고, 상기 3D 컨버전 정보를 상기 ITU H.264 비디오 인코딩 표준에 의해 지원되는 하나 이상의 SEI (Supplemental Enhancement Information) 메시지로 인코딩하는, 장치.

#### 청구항 27

삭제

#### 청구항 28

제 24 항에 있어서,

상기 3D 컨버전 정보는, 상기 제 1 2D 시퀀스에 적용되어 상기 3D 비디오 데이터를 생성하기 위한 3D 컨버전 프로세스를 식별하는 정보를 포함하는, 장치.

#### 청구항 29

제 24 항에 있어서,

상기 3D 컨버전 정보는 상기 제 1 2D 시퀀스를 캡처하는 것과 연관된 카메라 파라미터들 및 값들을 포함하는, 장치.

#### 청구항 30

제 24 항에 있어서,

상기 3D 컨버전 정보는,

상기 제 1 2D 시퀀스를 캡처한 카메라와 연관된 초점 길이를 표시하는 초점 길이 값,

상기 3D 비디오 데이터에서의 최소 깊이를 명시하는 근거리-깊이 값,

상기 3D 비디오 데이터에서의 최대 깊이를 명시하는 원거리-깊이 값, 및

상기 3D 비디오 데이터와 연관된 2 개의 카메라들 사이에서 추정된 거리를 정량화하는 평행 이동 값을 포함하는, 장치.

#### 청구항 31

제 24 항에 있어서,

상기 3D 컨버전 정보는,

상기 제 1 2D 시퀀스를 캡처한 카메라와 연관된 초점 길이를 표시하는 초점 길이 값,

상기 3D 비디오 데이터에서의 최소 깊이를 명시하는 근거리-깊이 값,

상기 3D 비디오 데이터에서의 최대 깊이를 명시하는 원거리-깊이 값,

상기 3D 비디오 데이터와 연관된 2 개의 카메라들 사이에서 추정된 거리를 정량화하는 평행 이동 값, 및

수렴 이미지 평면의 상기 2 개의 카메라들로의 추정된 거리를 정량화하는 수렴-깊이 값을 포함하는, 장치.

#### 청구항 32

제 24 항에 있어서,

상기 3D 컨버전 정보는 명시적인 3D 파라미터 세트가 상기 3D 컨버전 정보 내에 포함되는지 여부 또는 디폴트 3D 파라미터 세트가 이용될 것인지 여부를 나타내는 플래그를 포함하는, 장치.

#### 청구항 33

제 24 항에 있어서,

상기 3D 컨버전 정보는 상기 제 1 2D 시퀀스의 제 2 뷰가 상기 제 1 2D 시퀀스의 좌측에 생성될 것인지 또는 우측에 생성될 것인지 여부를 나타내는 플래그를 포함하는, 장치.

#### 청구항 34

제 24 항에 있어서,

상기 3D 컨버전 정보는 상기 3D 비디오 데이터로부터 제거될 크롭 영역을 식별하는 플래그를 포함하는, 장치.

#### 청구항 35

제 24 항에 있어서,

상기 3D 컨버전 정보는,

명시적인 3D 파라미터 세트가 상기 3D 컨버전 정보 내에 포함되는지 여부 또는 디폴트 3D 파라미터 세트가 이용될 것인지 여부를 나타내는 제 1 플래그로서, 상기 명시적인 3D 파라미터 세트는 상기 제 1 플래그가 설정되는 경우에 상기 3D 컨버전 정보 내에 포함되는, 상기 제 1 플래그,

상기 제 1 2D 시퀀스의 제 2 뷰가 상기 제 1 2D 시퀀스의 좌측에 생성될 것인지 또는 우측에 생성될 것인지 여부를 나타내는 제 2 플래그, 및

상기 3D 비디오 데이터로부터 제거될 크롭 영역을 식별하는 제 3 플래그를 포함하며,

상기 크롭 영역을 정의하는 정보는 상기 제 3 플래그가 설정되는 경우에 상기 3D 컨버전 정보 내에 포함되는, 장치.

#### 청구항 36

제 24 항에 있어서,

상기 장치는,

집적 회로;

마이크로프로세서; 및

상기 비디오 인코더를 포함하는 무선 통신 디바이스

중 적어도 하나를 포함하는, 장치.

#### 청구항 37

비디오 프레임들의 제 1 2-차원 (2D) 시퀀스를 수신하고,

상기 제 1 2D 시퀀스와 함께 3-차원 (3D) 컨버전 정보를 수신하는 것으로서, 상기 3D 컨버전 정보는 상기 제 1 2D 시퀀스로부터 유도되어, 상기 3D 컨버전 정보가, 상기 제 1 2D 시퀀스의 상기 비디오 프레임들의 각각에 적용되어 상기 제 1 2D 시퀀스와 함께 3D 비디오 데이터의 입체적인 비디오 시퀀스를 정의하는 제 2 2D 시퀀스를 생성할 수 있는 파라미터 세트를 포함하도록 하는, 상기 3D 컨버전 정보를 수신하고,

상기 제 1 2D 시퀀스를 디코딩하며, 그리고

상기 제 1 2D 시퀀스 및 상기 3D 컨버전 정보에 기초하여 상기 3D 비디오 데이터를 생성하는

비디오 디코더를 포함하는, 장치.

#### 청구항 38

제 37 항에 있어서,

상기 제 1 2D 시퀀스는 ITU H.264 비디오 인코딩 표준에 따라 인코딩되고, 그리고

상기 3D 컨버전 정보는 상기 ITU H.264 비디오 인코딩 표준에 의해 지원되는 하나 이상의 SEI (Supplemental Enhancement Information) 메시지들에서 수신되는, 장치.

#### 청구항 39

삭제

#### 청구항 40

제 37 항에 있어서,

상기 3D 컨버전 정보는, 상기 제 1 2D 시퀀스에 적용되어 상기 3D 비디오 데이터를 생성하기 위한 3D 컨버전 프로세스를 식별하는 정보를 포함하는, 장치.

#### 청구항 41

제 37 항에 있어서,

상기 3D 컨버전 정보는 상기 제 1 2D 시퀀스를 캡처하는 것과 연관된 카메라 파라미터들 및 값들을 포함하는, 장치.

#### 청구항 42

제 37 항에 있어서,

상기 3D 컨버전 정보는,

상기 제 1 2D 시퀀스를 캡처한 카메라와 연관된 초점 길이를 표시하는 초점 길이 값,

상기 3D 비디오 데이터에서의 최소 깊이를 명시하는 근거리-깊이 값,

상기 3D 비디오 데이터에서의 최대 깊이를 명시하는 원거리-깊이 값, 및

상기 3D 비디오 데이터와 연관된 2 개의 카메라들 사이에서 추정된 거리를 정량화하는 평행 이동 값을 포함하는, 장치.

#### 청구항 43

제 37 항에 있어서,

상기 3D 컨버전 정보는,

상기 제 1 2D 시퀀스를 캡처한 카메라와 연관된 초점 길이를 표시하는 초점 길이 값,

상기 3D 비디오 데이터에서의 최소 깊이를 명시하는 근거리-깊이 값,

상기 3D 비디오 데이터에서의 최대 깊이를 명시하는 원거리-깊이 값,

상기 3D 비디오 데이터와 연관된 2 개의 카메라들 사이에서 추정된 거리를 정량화하는 평행 이동 값, 및

수렴 이미지 평면의 상기 2 개의 카메라들로의 추정된 거리를 정량화하는 수렴-깊이 값을 포함하는, 장치.

#### 청구항 44

제 37 항에 있어서,

상기 3D 컨버전 정보는 명시적인 3D 파라미터 세트가 상기 3D 컨버전 정보 내에 포함되는지 여부 또는 디폴트 3D 파라미터 세트가 이용될 것인지 여부를 나타내는 플래그를 포함하는, 장치.

#### 청구항 45

제 37 항에 있어서,

상기 3D 컨버전 정보는 상기 제 1 2D 시퀀스의 제 2 뷰가 상기 제 1 2D 시퀀스의 좌측에 생성될 것인지 또는 우측에 생성될 것인지 여부를 나타내는 플래그를 포함하는, 장치.

#### 청구항 46



제 37 항에 있어서,

상기 3D 컨버전 정보는 상기 3D 비디오 데이터로부터 제거될 크롭 영역을 식별하는 플래그를 포함하는, 장치.

#### 청구항 47

제 37 항에 있어서,

상기 3D 컨버전 정보는,

명시적인 3D 파라미터 세트가 상기 3D 컨버전 정보 내에 포함되는지 여부 또는 디폴트 3D 파라미터 세트가 이용될 것인지 여부를 나타내는 제 1 플래그로서, 상기 명시적인 3D 파라미터 세트는 상기 제 1 플래그가 설정되는 경우에 상기 3D 컨버전 정보 내에 포함되는, 상기 제 1 플래그,

상기 제 1 2D 시퀀스의 제 2 뷰가 상기 제 1 2D 시퀀스의 좌측에 생성될 것인지 또는 우측에 생성될 것인지 여부를 나타내는 제 2 플래그, 및

상기 3D 비디오 데이터로부터 제거될 크롭 영역을 식별하는 제 3 플래그를 포함하며,

상기 크롭 영역을 정의하는 정보는 상기 제 3 플래그가 설정되는 경우에 상기 3D 컨버전 정보 내에 포함되는, 장치.

#### 청구항 48

제 37 항에 있어서,

상기 장치는 디스플레이를 포함하고,

상기 장치는,

상기 장치가 상기 3D 비디오 데이터를 생성하고 렌더링할 수 있는지 여부를 결정하고;

상기 장치가 상기 3D 비디오 데이터를 생성하고 렌더링할 수 있는 경우에, 상기 제 1 2D 시퀀스 및 상기 3D 컨버전 정보에 기초하여 상기 디스플레이 상에 상기 3D 비디오 데이터를 생성하고 렌더링하며; 및

상기 장치가 상기 3D 비디오 데이터를 생성하거나 렌더링할 수 없는 경우에, 상기 디스플레이 상에 상기 제 1 2D 시퀀스를 렌더링하는, 장치.

#### 청구항 49

제 37 항에 있어서,

상기 장치는,

집적 회로;

마이크로프로세서; 및

상기 비디오 디코더를 포함하는 무선 통신 디바이스

중 적어도 하나를 포함하는, 장치.

#### 청구항 50

비디오 인코더에서 비디오 프레임들의 제 1 2-차원 (2D) 시퀀스를 인코딩하는 수단;

상기 비디오 인코더를 통해 3-차원 (3D) 컨버전 정보를 인코딩하는 수단으로서, 상기 3D 컨버전 정보는 상기 제 1 2D 시퀀스로부터 유도되어, 상기 3D 컨버전 정보가, 상기 제 1 2D 시퀀스의 상기 비디오 프레임들의 각각에 적용되어, 상기 3D 컨버전 정보와 상기 제 1 2D 시퀀스로부터 3D 비디오 데이터의 입체적인 비디오 시퀀스가 유도가능하도록 함으로써 상기 제 1 2D 시퀀스와 함께 상기 3D 비디오 데이터의 입체적인 비디오 시퀀스를 정의하는 제 2 2D 시퀀스를 생성할 수 있는 파라미터 세트를 포함하도록 하는, 상기 3D 컨버전 정보를 인코딩하는 수단; 및

상기 인코딩된 제 1 2D 시퀀스를 인코딩된 파라미터들과 함께 통신하는 수단을 포함하는, 디바이스.

## 청구항 51

비디오 디코더에서 비디오 프레임들의 제 1 2-차원 (2D) 시퀀스를 수신하는 수단;

상기 비디오 디코더에서 상기 제 1 2D 시퀀스와 함께 3-차원 (3D) 컨버전 정보를 수신하는 수단으로서, 상기 3D 컨버전 정보는 상기 제 1 2D 시퀀스로부터 유도되어, 상기 3D 컨버전 정보가, 상기 제 1 2D 시퀀스의 상기 비디오 프레임들의 각각에 적용되어 상기 제 1 2D 시퀀스와 함께 3D 비디오 데이터의 입체적인 비디오 시퀀스를 정의하는 제 2 2D 시퀀스를 생성할 수 있는 파라미터 세트를 포함하도록 하는, 상기 3D 컨버전 정보를 수신하는 수단;

상기 제 1 2D 시퀀스를 디코딩하는 수단; 및

상기 제 1 2D 시퀀스 및 상기 3D 컨버전 정보에 기초하여 상기 3D 비디오 데이터를 생성하는 수단을 포함하는, 디바이스.

## 청구항 52

명령들을 포함하는 컴퓨터 판독가능 저장 매체로서,

상기 명령들은, 프로세서에 의해 실행될 시, 상기 프로세서로 하여금,

비디오 프레임들의 제 1 2-차원 (2D) 시퀀스를 인코딩하게 하고; 그리고

3-차원 (3D) 컨버전 정보를 인코딩하게 하도록 하며,

상기 3D 컨버전 정보는 상기 제 1 2D 시퀀스로부터 유도되어, 상기 3D 컨버전 정보가, 상기 제 1 2D 시퀀스의 상기 비디오 프레임들의 각각에 적용되어, 상기 3D 컨버전 정보와 상기 제 1 2D 시퀀스로부터 3D 비디오 데이터의 입체적인 비디오 시퀀스가 유도가능하도록 함으로써 상기 제 1 2D 시퀀스와 함께 상기 3D 비디오 데이터의 입체적인 비디오 시퀀스를 정의하는 제 2 2D 시퀀스를 생성할 수 있는 파라미터 세트를 포함하도록 하는, 컴퓨터 판독가능 저장 매체.

## 청구항 53

명령들을 포함하는 컴퓨터 판독가능 저장 매체로서,

상기 명령들은, 프로세서에 의해 실행될 시, 상기 프로세서로 하여금,

비디오 프레임들의 제 1 2-차원 (2D) 시퀀스를 수신할 시에, 그리고 상기 제 1 2D 시퀀스와 함께 3-차원 (3D) 컨버전 정보를 수신할 시에,

상기 제 1 2D 시퀀스를 디코딩하게 하고; 그리고

상기 제 1 2D 시퀀스 및 상기 3D 컨버전 정보에 기초하여 상기 3D 비디오 데이터를 생성하게 하며,

상기 3D 컨버전 정보는 상기 제 1 2D 시퀀스로부터 유도되어, 상기 3D 컨버전 정보가, 상기 제 1 2D 시퀀스의 상기 비디오 프레임들의 각각에 적용되어 상기 제 1 2D 시퀀스와 함께 3D 비디오 데이터의 입체적인 비디오 시퀀스를 정의하는 제 2 2D 시퀀스를 생성할 수 있는 파라미터 세트를 포함하도록 하는, 컴퓨터 판독가능 저장 매체.

## 청구항 54

비디오 프레임들의 제 1 2-차원 (2D) 시퀀스를 수신하는 단계;

상기 제 1 2D 시퀀스와 함께 3-차원 (3D) 컨버전 정보를 수신하는 단계;

상기 제 1 2D 시퀀스를 디코딩하는 단계; 및

상기 3D 컨버전 정보를 상기 비디오 프레임들의 상기 제 1 2D 시퀀스에 적용하여 3D 비디오 데이터를 생성하는 단계를 포함하며,

상기 3D 컨버전 정보는 상기 제 1 2D 시퀀스로부터 유도되어, 상기 3D 컨버전 정보가, 상기 제 1 2D 시퀀스의 상기 비디오 프레임들의 각각에 적용되어, 상기 3D 컨버전 정보와 상기 제 1 2D 시퀀스로부터 3D 비디오 데이터의 입체적인 비디오 시퀀스가 유도가능하도록 함으로써 상기 디코딩된 제 1 2D 시퀀스와 함께 상기 3D 비디오

데이터의 입체적인 비디오 시퀀스를 정의하는 제 2 2D 시퀀스를 생성할 수 있는 파라미터 세트를 포함하도록 하는, 방법.

## 명세서

### 기술분야

[0001] 본 출원은, 2009년 6월 5일에 출원된 미국 가출원 제61/184,649호의 이익을 주장하며, 그 전체 내용은 본 명세서에서 참조로서 통합된다.

[0002] 기술분야

[0003] 본 개시물은 비디오 코딩, 및 2-차원 (2D) 비디오 데이터의 3-차원 (3D) 비디오 데이터로의 컨버전에 관한 것이다.

### 배경기술

[0004] 디지털 멀티미디어 능력들은 디지털 텔레비전들, 디지털 다이렉트 브로드캐스트 시스템들, 무선 통신 디바이스들, 무선 브로드캐스트 시스템들, 개인 휴대 정보 단말기들 (PDAs), 랩톱 또는 데스크톱 컴퓨터들, 디지털 카메라들, 디지털 레코딩 디바이스들, 비디오 게임 디바이스들, 비디오 게임 콘솔들, 셀룰러 또는 위성 무선 전화기들, 디지털 미디어 플레이어들, 등등을 포함하는 광범위한 디바이스들 내에 통합될 수 있다. 디지털 멀티미디어 디바이스들은 디지털 비디오 데이터를 더 효율적으로 전송 및 수신하거나 저장 및 검색하기 위해 MPEG-2, ITU-H.263, MPEG-4, 또는 ITU-H.264/MPEG-4 Part 10, AVC (Advanced Video Coding) 와 같은 비디오 코딩 기술들을 구현할 수도 있다. 비디오 인코딩 기술들은 비디오 시퀀스들 내에서 고유한 리던던시를 감소시키거나 제거하기 위해 공간적인 및 시간적인 예측을 통해 비디오 압축을 수행할 수도 있다.

[0005] 대부분의 종래의 비디오 시퀀스들은 2-차원 (2D) 뷰잉 포맷으로 제공된다. 그러나, 3-차원 (3D) 시퀀스들 또한 가능하며, 이 경우에 비디오 시퀀스들은 각각의 비디오 프레임과 연관된 2 이상의 뷰들을 갖는다. 이 경우에, 2 이상의 뷰들은 3D 디스플레이에서 결합되어 3D 비디오를 렌더링할 수도 있다. 3D 비디오 시퀀스들의 통신은 2D 비디오 시퀀스들과 비교하여 상당한 양의 추가 데이터를 요구할 수 있다. 예를 들면, 3D 비디오 시퀀스를 통신하기 위해, 2 개의 개별 비디오 프레임들은 각각의 2D 비디오 프레임의 2 개의 서로 다른 뷰들을 제공하도록 요구되며, 따라서 대략 2 배의 데이터 양이 통신된다.

### 발명의 내용

#### 과제의 해결 수단

[0006] 본 개시물은 비디오 프레임들의 2-차원 (2D) 비디오 시퀀스를 3-차원 (3D) 컨버전 정보와 함께 인코딩하기 위한 기술들을 설명하며, 3D 컨버전 정보는 2D 시퀀스의 비디오 프레임들 각각에 적용되어 3D 비디오 데이터를 생성할 수 있는 파라미터 세트를 포함한다. 본 개시물은 또한 2D 비디오 시퀀스와 3D 컨버전 정보의 통신과 디코딩을 설명한다. 파라미터 세트는 2D 시퀀스의 오리지널 비디오 프레임들 각각에 적용되어 오리지널 비디오 프레임들의 각각에 대한 2차 뷰들을 생성할 수 있는 비교적 소량의 데이터를 포함할 수도 있다. 오리지널 비디오 프레임들 및 2차 뷰 비디오 프레임들은 집합적으로 입체적인 3D 비디오 시퀀스를 정의할 수도 있다. 2D 시퀀스 및 파라미터 세트는 3D 시퀀스를 통신하기 위해 요구되는 것 보다 상당히 적은 데이터를 포함할 수도 있다. 2D 시퀀스 및 파라미터 세트는 오리지널 2D 시퀀스를 통신하기 위해 요구되는 데이터 대하여 무시할만한 증가를 포함할 수 있다. 본 개시물은 또한 효과적이고 효율적인 방식으로 파라미터 세트를 인코딩하기 위해 이용될 수도 있는 일부 예시적인 신택스 (syntax) 를 설명한다.

[0007] 수신 디바이스는 수신 디바이스가 3D 디코딩 또는 3D 렌더링을 지원하지 않는 경우에도 2D 시퀀스를 디코딩하고 렌더링할 수도 있다. 그러나, 수신 디바이스가 본 개시물과 부합하는 3D 디코딩 및 3D 렌더링을 지원하면, 수신 디바이스는 2D 시퀀스 및 파라미터 세트에 기초하여 3D 시퀀스를 생성하고 렌더링할 수도 있다. 이러한 방식으로, 본 개시물의 기술들은 2D 비디오 출력 또는 3D 비디오 출력을 렌더링하기 위해 동일한 비트스트림이 이용될 수 있는, 역방향 호환성의 2D - 3D 비디오 코딩 및 컨버전을 지원할 수도 있다. 더욱이, 언급된 것과 같이, 설명되는 기술들은 3D 비디오 시퀀스의 통신을 위해 요구되는 데이터의 양을 감소시킬 수도 있다.

[0008] 일 실시예에서, 본 개시물은 비디오 인코더 내에서 비디오 프레임들의 2-차원 (2D) 시퀀스를 인코딩하는 단계, 비디오 인코더를 통해 3-차원 (3D) 컨버전 정보를 인코딩하는 단계로서, 3D 컨버전 정보는 2D 시퀀스의 비디오

프레임들의 각각에 적용되어 3D 비디오 데이터를 생성할 수 있는 파라미터 세트를 포함하는, 상기 3D 컨버전 정보를 인코딩하는 단계, 및 인코딩된 2D 시퀀스를 3D 컨버전 정보와 함께 통신하는 단계를 포함하는 방법을 설명한다.

- [0009] 또 다른 실시예에서, 본 개시물은 비디오 디코더에서 비디오 프레임들의 2-차원 (2D) 시퀀스를 수신하는 단계, 비디오 디코더에서 2D 시퀀스와 함께 3-차원 (3D) 컨버전 정보를 수신하는 단계로서, 3D 컨버전 정보는 2D 시퀀스의 비디오 프레임들의 각각에 적용되어 3D 비디오 데이터를 생성할 수 있는 파라미터 세트를 포함하는, 상기 3D 컨버전 정보를 수신하는 단계, 비디오 디코더를 통해 2D 시퀀스를 디코딩하는 단계, 및 2D 시퀀스 및 3D 컨버전 정보에 기초하여 비디오 디코더를 통해 3D 비디오 데이터를 생성하는 단계를 포함하는 방법을 설명한다.
- [0010] 또 다른 실시예에서, 본 개시물은 비디오 프레임들의 2-차원 (2D) 시퀀스를 인코딩하고, 2D 시퀀스로 3-차원 (3D) 컨버전 정보를 인코딩하는 비디오 인코더를 포함하는 장치를 설명하며, 상기 3D 컨버전 정보는 2D 시퀀스의 비디오 프레임들의 각각에 적용되어 3D 비디오 데이터를 생성할 수 있는 파라미터 세트를 포함한다.
- [0011] 또 다른 실시예에서, 본 개시물은 비디오 프레임들의 2-차원 (2D) 시퀀스를 수신하고, 2D 시퀀스와 함께 3-차원 (3D) 컨버전 정보를 수신하는 것으로서, 3D 컨버전 정보는 2D 시퀀스의 비디오 프레임들의 각각에 적용되어 3D 비디오 데이터를 생성할 수 있는 파라미터 세트를 포함하는, 상기 3D 컨버전 정보를 수신하며, 2D 시퀀스를 디코딩하며, 2D 시퀀스 및 3D 컨버전 정보에 기초하여 3D 비디오 데이터를 생성하는 디코더를 포함하는 장치를 설명한다.
- [0012] 또 다른 실시예에서, 본 개시물은 비디오 인코더 내에서 비디오 프레임들의 2-차원 (2D) 시퀀스를 인코딩하는 수단, 비디오 인코더를 통해 3-차원 (3D) 컨버전 정보를 인코딩하는 수단으로서, 상기 3D 컨버전 정보는 2D 시퀀스의 비디오 프레임들의 각각에 적용되어 3D 비디오 데이터를 생성할 수 있는 파라미터 세트를 포함하는, 상기 3D 컨버전 정보를 인코딩하는 수단, 및 인코딩된 2D 시퀀스를 인코딩된 파라미터들과 함께 통신하는 수단을 포함하는 디바이스를 설명한다.
- [0013] 또 다른 실시예에서, 본 발명은 비디오 디코더에서 비디오 프레임들의 2-차원 (2D) 시퀀스를 수신하는 수단, 비디오 디코더에서 2D 시퀀스와 함께 3-차원 (3D) 컨버전 정보를 수신하는 수단으로서, 3D 컨버전 정보는 2D 시퀀스의 비디오 프레임들의 각각에 적용되어 3D 비디오 데이터를 생성할 수 있는 파라미터 세트를 포함하는, 상기 3D 컨버전 정보를 수신하는 수단, 2D 시퀀스를 디코딩하는 수단, 및 2D 시퀀스 및 3D 컨버전 정보에 기초하여 3D 비디오 데이터를 생성하는 수단을 포함하는 디바이스를 설명한다.
- [0014] 또 다른 실시예에서, 본 발명은 3D 컨버전 정보를 2D 시퀀스에 적용하여 3D 비디오 데이터를 생성하는 방법들, 장치들 또는 디바이스들을 설명하며, 상기 3D 컨버전 정보는 2D 시퀀스의 비디오 프레임들의 각각에 적용되어 3D 비디오 데이터를 생성할 수 있는 파라미터 세트를 포함한다.
- [0015] 본 개시물에서 설명되는 기술들은 하드웨어, 소프트웨어, 펌웨어 또는 이들의 임의의 조합으로 구현될 수도 있다. 소프트웨어로 구현되는 경우에, 소프트웨어는 범용 마이크로프로세서, 주문형 집적 회로 (ASIC), 필드 프로그램가능 게이트 어레이 (FPGA), 또는 디지털 신호 프로세서 (DSP) 와 같은 하나 이상의 프로세서들에서 실행될 수도 있다. 기술들을 실행하는 소프트웨어는 초기에 컴퓨터 판독가능 매체 내에 저장되고, 프로세서 내에 로딩되어 실행될 수도 있다.
- [0016] 따라서, 본 개시물은 또한 프로세서에 의해 실행될 때 프로세서로 하여금, 비디오 프레임들의 2-차원 (2D) 시퀀스를 인코딩하고, 3-차원 (3D) 컨버전 정보를 인코딩하도록 하는 명령들을 포함하는 컴퓨터 판독가능 저장 매체를 고려하며, 3D 컨버전 정보는 2D 시퀀스의 비디오 프레임들의 각각에 적용되어 3D 비디오 데이터를 생성할 수 있는 파라미터 세트를 포함한다.
- [0017] 추가로, 본 개시물은 프로세서에 의해 실행될 때 프로세서로 하여금, 비디오 프레임들의 2-차원 (2D) 시퀀스를 수신하고, 2D 시퀀스와 함께 3-차원 (3D) 컨버전 정보를 수신하는 것으로서, 3D 컨버전 정보는 2D 시퀀스의 비디오 프레임들의 각각에 적용되어 3D 비디오 데이터를 생성할 수 있는 파라미터 세트를 포함하는, 상기 3D 컨버전 정보를 수신하며, 2D 시퀀스를 디코딩하고, 2D 시퀀스 및 3D 컨버전 정보에 기초하여 3D 비디오 데이터를 생성하도록 하는 명령들을 포함하는 컴퓨터 판독가능 저장 매체를 설명한다.
- [0018] 본 기술들의 하나 이상의 양태들의 세부사항을 첨부한 도면 및 하기의 설명에서 설명한다. 본 개시물에서 기술되는 기술들의 다른 특징들, 목적들 및 이점들은 상세한 설명 및 도면과 청구범위로부터 명백해질 것이다.

## 도면의 간단한 설명

- [0019] 도 1 은 본 개시물의 기술들을 구현할 수도 있는 예시적인 비디오 인코딩 및 디코딩 시스템을 예시한 블록도이다.
- 도 2 는 본 개시물과 부합하는 2-차원 (2D) 및 3-차원 (3D) 비디오 인코딩을 수행할 수도 있는 예시적인 비디오 인코더를 예시한 블록도이다.
- 도 3 은 본 개시물과 부합하는 2D 및 3D 비디오 디코딩을 수행할 수도 있는 예시적인 비디오 디코더를 예시한 블록도이다.
- 도 4 내지 도 6 은 3D 컨버전 파라미터들에 기초하여 적용될 수도 있는 2D - 3D 컨버전의 양태들을 예시하는 개념도들이다.
- 도 7 은 본 개시물과 부합하는 비디오 인코딩 디바이스에 의해 수행되는 예시적인 프로세스를 예시하는 흐름도이다.
- 도 8 은 본 개시물과 부합하는 비디오 디코딩 디바이스에 의해 수행되는 예시적인 프로세스를 예시하는 흐름도이다.

### 발명을 실시하기 위한 구체적인 내용

- [0020] 본 개시물은 비디오 프레임들의 2-차원 (2D) 비디오 시퀀스를 3-차원 (3D) 컨버전 정보와 함께 인코딩하기 위한 기술들을 설명하며, 3D 컨버전 정보는 2D 시퀀스의 비디오 프레임들 각각에 적용되어 3D 비디오 데이터를 생성할 수 있는 파라미터 세트를 포함한다. 3D 컨버전 정보는 비디오 시퀀스의 서로 다른 프레임들에 대하여 달라지는 것이 아니라, 2D 시퀀스의 오리지널 비디오 프레임들 각각에 적용되어 오리지널 비디오 프레임들의 각각에 대한 2차 뷰 비디오 프레임들을 생성할 수 있는 비교적 소량의 데이터를 형성하는 공통 파라미터 세트를 포함한다. 오리지널 비디오 프레임들 및 2차 뷰 비디오 프레임들은 집합적으로 3D 디스플레이에 렌더링될 수 있는 입체적인 3D 비디오 시퀀스를 정의할 수도 있다. 본 개시물에 따라, 2D 시퀀스 및 파라미터 세트는 3D 시퀀스를 통신하기 위해 요구되는 것 보다 상당히 적은 데이터를 포함할 수도 있다.
- [0021] 일 실시예에서, 3D 컨버전 정보는 2D 시퀀스의 오리지널 비디오 프레임들 각각에 적용되어 오리지널 비디오 프레임들의 각각의 2차 뷰들을 생성할 수도 있는 20 바이트 미만의 데이터를 포함할 수도 있다. 본 개시물의 기술들은 MPEG-2, MPEG-4, ITU H.263, ITU H.264, 독점적인 코딩 표준들 또는 향후 코딩 표준들과 같은 다수의 코딩 세팅들에서 유용할 수도 있다. ITU H.264 프레임워크에 따라, 본 개시물은 비디오 표준-순응 2D 비디오 시퀀스와 3D 컨버전 정보를 통신하기 위한 메커니즘으로서 SEI (Supplemental Enhancement Information) 메세지들을 이용할 수 있다.
- [0022] 수신 디바이스는 수신 디바이스가 3D 디코딩 또는 3D 렌더링을 지원하지 않는 경우에도 2D 시퀀스를 디코딩하고 렌더링할 수도 있다. 그러나, 수신 디바이스가 본 개시물과 부합하는 3D 디코딩 및 3D 렌더링을 지원하면, 수신 디바이스는 2D 시퀀스 및 파라미터 세트에 기초하여 3D 시퀀스를 생성하고 렌더링할 수도 있다. 이러한 방식으로, 본 개시물의 기술들은 2D 비디오 출력 또는 3D 비디오 출력을 렌더링하기 위해 동일한 비트스트림이 이용될 수 있는, 스케일러블 2D - 3D 비디오 코딩을 지원할 수도 있다. 더욱이, 언급된 것과 같이, 설명되는 기술들은 3D 비디오 시퀀스의 통신을 위해 요구되는 데이터의 양을 감소시킬 수도 있다.
- [0023] 본 개시물은 또한 효과적이고 효율적인 방식으로 파라미터 세트를 인코딩하기 위해 이용될 수도 있는 일부 예시적인 신택스 (syntax) 를 설명한다. 예를 들어, 일부 구현들에서, 신택스 엘리먼트들은 3D 컨버전 정보의 통신을 위해 ITU H.264 의 SEI 메세지들에서 이용될 수도 있다. 일 실시예에서, 하기에서 더 상세히 논의되는 것과 같이, 3D 컨버전 정보는 3D 컨버전 정보 내에 명시적인 3D 파라미터 세트가 포함되는지 여부 또는 디폴트 3D 파라미터 세트가 이용될 것인지의 여부를 나타내는 제 1 플래그를 포함할 수도 있고, 상기 명시적인 3D 파라미터 세트는 제 1 플래그가 설정될 때 3D 컨버전 정보 내에 포함된다. 이 경우에, 제 1 플래그가 설정되지 않으면, 디코더는 여전히 디폴트 3D 파라미터들을 적용할 수도 있다.
- [0024] 3D 컨버전 정보는 또한 2D 시퀀스의 2차 뷰가 2D 시퀀스의 좌측 또는 우측에 생성될 것인지 여부를 나타내는 제 2 플래그를 포함할 수도 있다. 이 경우에, 제 2 플래그는 디코더에서 (예컨대, 오리지널 비디오 프레임들의 좌측 또는 우측에) 생성될 2차 뷰들의 방향을 제공함으로써 3D 렌더링을 보조할 수 있다. 또한, 3D 컨버전 정보는 크롭 (crop) 영역이 3D 비디오 데이터로부터 제거될 것인지 여부를 확인하는 제 3 플래그를 포함할 수도 있고, 상기 크롭 영역을 정의하는 정보는 제 3 플래그가 설정될 때 3D 컨버전 정보 내에 포함된다. 제 3 플래그가 설정되지 않으면, 3D 비디오 데이터의 생성 및 3D 렌더링에서 크롭핑이 회피될 수도 있다. 일부 경



우에, 제 1 플래그가 설정되지 않으면, 제 2 및 제 3 플래그들이 비트스트림으로부터 제외될 수도 있다. 플래그들은 단일-비트 또는 멀티-비트 플래그들을 포함할 수도 있다.

[0025] 도 1 은 본 개시물의 기술들을 구현할 수 있는 예시적인 비디오 인코딩 및 디코딩 시스템 (10) 을 예시한 블록도이다. 도 1 에 도시된 바와 같이, 시스템 (10) 은 인코딩된 비디오를 통신 채널 (15) 을 통해 목적지 디바이스 (16) 로 송신하는 소스 디바이스 (12) 를 포함한다. 소스 디바이스 (12) 및 목적지 디바이스 (16) 는 모바일 디바이스들 또는 일반적으로 고정된 디바이스들을 포함하는 광범위한 디바이스들 중 임의의 것을 포함할 수도 있다. 일부 경우에, 소스 디바이스 (12) 및 목적지 디바이스 (16) 는 무선 핸드셋들, 소위 셀룰러 또는 위성 무선 전화기들, 개인 휴대 정보 단말기들 (PDAs), 모바일 미디어 플레이어들, 또는 무선일 수도 있거나 무선이 아닐 수도 있는 통신 채널 (15) 을 통해 비디오 정보를 통신할 수 있는 임의의 디바이스와 같은 무선 통신 디바이스들을 포함한다. 그러나, 2D 비디오 시퀀스와 함께 3D 컨버전 정보의 생성, 통신 및 이용에 관한 본 개시물의 기술들은 다수의 서로 다른 시스템들 및 세팅들에서 사용될 수도 있다. 도 1 은 그러한 시스템의 하나의 실시예일 뿐이다.

[0026] 도 1 의 실시예에서, 소스 디바이스 (12) 는 비디오 소스 (20), 비디오 인코더 (22), 변조기/복조기(모뎀) (23) 및 송신기 (24) 를 포함할 수도 있다. 목적지 디바이스 (16) 는 수신기 (26), 모뎀 (27), 비디오 디코더 (28) 및 디스플레이 디바이스 (30) 를 포함할 수도 있다. 본 개시물에 따라, 소스 디바이스 (12) 의 비디오 인코더 (22) 는 비디오 프레임들의 2D 시퀀스를 인코딩하고, 3D 컨버전 정보를 인코딩하도록 구성될 수도 있고, 상기 3D 컨버전 정보는 2D 시퀀스의 비디오 프레임들의 각각에 적용되어 3D 비디오 데이터를 생성할 수 있는 파라미터 세트를 포함한다. 모뎀 (23) 및 송신기 (24) 는 무선 신호들을 변조하여 목적지 디바이스로 전송할 수도 있다. 이러한 방식으로, 소스 디바이스 (12) 는 인코딩된 2D 시퀀스를 3D 컨버전 정보와 함께 목적지 디바이스 (16) 로 통신한다.

[0027] 수신기 (26) 및 모뎀 (27) 은 소스 디바이스 (12) 로부터 수신된 무선 신호들을 수신하고 복조한다. 따라서, 비디오 디코더 (28) 는 2D 시퀀스 및 2D 시퀀스를 디코딩하는 3D 컨버전 정보를 수신할 수도 있다. 본 개시물에 따라, 비디오 디코더 (28) 는 2D 시퀀스 및 3D 컨버전 정보에 기초하여 3D 비디오 데이터를 생성할 수도 있다. 다시 말해서, 3D 컨버전 정보는 2D 시퀀스의 비디오 프레임들 각각에 적용되어 3D 비디오 데이터를 생성할 수 있는 파라미터 세트를 포함할 수도 있고, 상기 3D 비디오 데이터는 3D 시퀀스를 통신하는데 요구되는 것보다 상당히 적은 데이터를 포함할 수도 있다.

[0028] 언급되는 바와 같이, 도 1 의 예시된 시스템 (10) 은 단지 예시적이다. 본 개시물의 기술들은 1차 블록-기반 비디오 코딩을 지원하는 임의의 코딩 디바이스 또는 기술로 확대될 수도 있다. 소스 디바이스 (12) 및 목적지 디바이스 (16) 는 소스 디바이스 (12) 가 목적지 디바이스 (16) 로의 전송을 위해 코딩된 비디오 데이터를 생성하는, 코딩 디바이스들의 실시예일 뿐이다. 일부 경우에, 디바이스들 (12, 16) 은 디바이스들 (12, 16) 각각이 비디오 인코딩 및 디코딩 컴포넌트들을 포함하도록 실질적으로 대칭적인 방식으로 동작할 수도 있다. 따라서, 시스템 (10) 은 예컨대, 비디오 스트리밍, 비디오 재생, 비디오 브로드캐스팅 또는 비디오 전화를 위해 비디오 디바이스들 (12, 16) 사이에 단방향 또는 양방향 비디오 전송을 지원할 수도 있다.

[0029] 소스 디바이스 (12) 의 비디오 소스 (20) 는 비디오 카메라와 같은 비디오 캡처 디바이스, 이전에 캡처된 비디오를 저장하는 비디오 아카이브 (archive), 또는 비디오 콘텐츠 제공자로부터의 비디오 피드를 포함할 수도 있다. 추가의 대안으로서, 비디오 소스 (20) 는 소스 비디오로서 컴퓨터 그래픽-기반 데이터, 또는 라이브 비디오, 보관된 비디오, 및 컴퓨터-생성된 비디오의 조합을 생성할 수도 있다. 일부 경우에, 비디오 소스 (20) 가 비디오 카메라이면, 소스 디바이스 (12) 및 목적지 디바이스 (16) 는 소위 카메라폰 또는 비디오폰을 형성한다. 각각의 경우에, 캡처되거나, 미리 캡처되거나, 컴퓨터-생성된 비디오는 비디오 인코더 (22) 에 의해 인코딩될 수도 있다. 인코딩된 비디오 정보는 예컨대, 코드 분할 다중 액세스 (CDMA) 또는 다른 통신 표준과 같은 통신 표준에 따라 모뎀 (23) 에 의해 변조되고, 송신기 (24) 를 통해 목적지 디바이스 (16) 로 전송될 수도 있다. 모뎀 (23) 은 신호 변조를 위해 설계된 다양한 믹서들, 필터들, 증폭기들 또는 다른 컴포넌트들을 포함할 수도 있다. 송신기 (24) 는 증폭기들, 필터들, 및 하나 이상의 안테나들을 포함하는, 데이터를 전송하기 위해 설계된 회로들을 포함할 수도 있다.

[0030] 목적지 디바이스 (16) 의 수신기 (26) 는 채널 (15) 을 통해 정보를 수신하고, 모뎀 (27) 은 정보를 복조한다. 다시 말해서, 비디오 인코딩 프로세스는 2D 시퀀스의 비디오 프레임들 각각에 적용되어 3D 비디오 데이터를 생성할 수 있는 파라미터 세트를 결정하기 위해 본 명세서 내에서 설명된 기술들 중 하나 이상을 구현할 수도 있다. 채널 (15) 을 통해 통신된 정보는 본 개시물과 부합하는 비디오 디코더 (28) 에 의해 이용될 수도 있다.

는 비디오 인코더 (22) 에 의해 정의된 정보를 포함할 수도 있다. 디스플레이 디바이스 (30) 는 디코딩된 비디오 데이터를 사용자에게 디스플레이하고, 음극선관, 액정 디스플레이 (LCD), 플라즈마 디스플레이, 유기 발광 다이오드 (OLED) 디스플레이, 또는 다른 타입의 디스플레이 디바이스와 같은 다양한 디스플레이 디바이스들 중 임의의 것을 포함할 수도 있다.

[0031] 도 1 의 예에서, 통신 채널 (15) 은 무선 주파수 (RF) 스펙트럼 또는 하나 이상의 물리적인 전송 라인들과 같은 임의의 무선 또는 유선 통신 매체, 또는 무선 및 유선 매체의 임의의 조합을 포함할 수도 있다. 따라서, 모뎀 (23) 및 송신기 (24) 는 다수의 가능한 무선 프로토콜들, 유선 프로토콜들, 또는 유선 및 무선 프로토콜들을 지원할 수도 있다. 통신 채널 (15) 은 근거리 네트워크 (LAN) 와 같은 패킷-기반 네트워크, 광역 네트워크 (WAN), 또는 하나 이상의 네트워크들의 상호 접속을 포함하는, 인터넷과 같은 글로벌 네트워크의 일부를 형성할 수도 있다. 통신 채널 (15) 은 일반적으로 소스 디바이스 (12) 로부터 목적지 디바이스 (16) 로 비디오 데이터를 전송하기 위한, 임의의 적절한 통신 매체 또는 서로 다른 통신 매체의 집합을 나타낸다. 통신 채널 (15) 은 라우터들, 스위치들, 기지국들, 또는 소스 디바이스 (12) 로부터 목적지 디바이스 (16) 로의 통신을 용이하게 하는데 유용할 수도 있는 임의의 다른 장비를 포함할 수도 있다. 본 개시물의 기술들은 반드시 하나의 디바이스로부터 다른 디바이스로 인코딩된 데이터의 통신을 요구하는 것이 아니며, 상호 디코딩 없이 인코딩 시나리오들에 적용할 수도 있다. 또한, 본 개시물의 양태들은 상호 인코딩 없이 디코딩 시나리오들에 적용할 수도 있다.

[0032] 비디오 인코더 (22) 및 비디오 디코더 (28) 는 대안적으로 MPEG-4, Part 10, AVC (Advanced Video Coding) 로서 설명되는 ITU-T H.264 표준과 같은 비디오 압축 표준에 따라 동작할 수도 있다. 그러나, 본 개시물의 기술들은 임의의 특정 코딩 표준 또는 그 확장에 제한되지 않는다. 도 1 에 도시되지는 않았지만, 일부 양태에서, 비디오 인코더 (22) 및 비디오 디코더 (28) 가 오디오 인코더 및 디코더와 각각 통합될 수도 있고, 적절한 MUX-DEMUX 유닛들, 또는 다른 하드웨어 및 소프트웨어를 포함하여, 공통 데이터 스트림 또는 개별 데이터 스트림들에서 오디오 및 비디오 양쪽 모두의 인코딩을 핸들링할 수도 있다. 적용가능하다면, MUX-DEMUX 유닛들은 ITU H.223 멀티플렉서 프로토콜, 또는 사용자 데이터그램 프로토콜 (UDP) 과 같은 다른 프로토콜들에 따를 수도 있다.

[0033] ITU-T H.264/MPEG-4 (AVC) 표준은 JVT (Joint Video Team) 로 알려진 단체 파트너십의 제품으로서 ISO/IEC MPEG (Moving Picture Experts Group) 과 함께 ITU-T VCEG (Video Coding Experts Group) 에 의해 만들어졌다. H.264 표준은 2005 년 3 월에 ITU-T Study Group 에 의한, 일반적인 시청각 서비스들을 위한 ITU-T 권고 H.264, AVC (Advanced Video Coding) 에서 설명되며, 이는 본 명세서에서 H.264 표준 또는 H.264 규격, 또는 H.264/AVC 표준 또는 규격으로 지칭될 수도 있다. JVT (Joint Video Team) 은 H.264/MPEG-4 AVC 로의 확장을 위해 노력하는 것을 계속하고 있다.

[0034] H.264/MPEG-4 AVC 표준을 진전하기 위한 작업은 KTA (Key Technologies Area) 포럼과 같은 ITU-T 의 다양한 포럼들에서 발생한다. KTA 포럼은 H.264/AVC 표준에 의해 보여지는 것보다 더 높은 코딩 효율을 보이는 코딩 기술들을 개발하려 한다. 본 개시물에서 설명되는 기술들은 특히 3D 비디오를 위한 H.264/AVC 표준과 관련하여 코딩 개선들을 제공할 수도 있다. 일부 양태들에서, 본 개시물은 본 명세서에 설명되는 3D 컨버전 정보를 인코딩하고 통신하기 위한 메커니즘으로서 ITU-T H.264 프레임워크 내에서 SEI (supplemental enhancement information) 메세지들의 이용을 고려한다.

[0035] 비디오 인코더 (22) 및 비디오 디코더 (28) 각각은 하나 이상의 마이크로프로세서들, 디지털 신호 프로세서 (DSP) 들, 주문형 집적회로 (ASIC) 들, 필드 프로그램가능 게이트 어레이 (FPGA) 들, 개별 로직 회로, 마이크로프로세서 또는 다른 플랫폼에 실행하는 소프트웨어, 하드웨어, 펌웨어 또는 이들의 임의의 조합들로서 구현될 수도 있다. 비디오 인코더 (22) 및 비디오 디코더 (28) 각각은 하나 이상의 인코더들 또는 디코더들에 포함될 수 있고, 그 중 어느 하나는 각각의 모바일 디바이스, 가입자 디바이스, 브로드캐스트 디바이스, 서버 등등 내에 결합된 인코더/디코더 (CODEC) 의 일부로서 통합될 수 있다.

[0036] 비디오 시퀀스는 통상적으로 일련의 비디오 프레임들을 포함한다. 비디오 인코더 (22) 및 비디오 디코더 (28) 는 비디오 데이터를 인코딩하고 디코딩하기 위해 개개의 비디오 프레임들 내의 픽셀들의 블록들 상에서 동작할 수도 있다. 비디오 블록들은 고정 또는 가변 사이즈들을 가질 수도 있고, 특정된 코딩 표준에 따라 사이즈가 달라질 수도 있다. 각각의 비디오 프레임은 일련의 슬라이스들 또는 다른 독립적으로 디코딩가능한 유닛들을 포함할 수도 있다. 각각의 슬라이스는 일련의 매크로블록들을 포함할 수도 있고, 그 매크로블록들은 서브-블록들로 배열될 수도 있다. 일 예로서, ITU-T H.264 표준은, 루마 성분들에 대하여 16×16, 8×8,

4×4 및 크로마 성분들에 대하여 8×8 과 같은 다양한 블록 사이즈들에서의 인트라 예측 (intra prediction) 뿐만 아니라, 루마 성분들에 대하여 16×16, 16×8, 8×16, 8×8, 8×4, 4×8 및 4×4 및 크로마 성분들에 대하여 대응하는 스케일링된 사이즈들과 같은 다양한 블록 사이즈들에서의 인터 예측 (inter prediction) 을 지원한다.

비디오 블록들은 예컨대, 이산 코사인 변환과 같은 변환 프로세스 또는 개념적으로 유사한 변환 프로세스들 이후에 변환 계수들의 블록들 또는 픽셀 데이터의 블록들을 포함할 수도 있다.

[0037] 비디오 블록들이 작을수록 더 양호한 분해능을 제공할 수 있고, 이는 높은 세부 레벨을 포함하는 비디오 프레임의 위치들을 위해 이용될 수도 있다. 일반적으로, 매크로블록들 및 다양한 서브-블록들 또는 분할들과 같은 일련의 비디오 블록들인 것으로 고려될 수도 있다. 일반적으로, 매크로블록은 픽셀들의 16×16 영역을 정의하는 크로미넌스 및 루미넌스 값들의 세트를 지칭할 수도 있다. 루미넌스 블록은 값들의 16×16 세트를 포함할 수도 있지만, 8×8 블록들, 4×4 블록들, 8×4 블록들, 4×8 블록들 또는 다른 사이즈들과 같은 더 작은 비디오 블록들로 추가로 분할될 수도 있다. 2 개의 서로 다른 크로미넌스 블록들은 매크로 블록에 대한 컬러를 정의할 수도 있고, 각각 픽셀들의 16×16 영역과 연관된 컬러 값들의 8×8 서브-샘플링된 블록들을 포함한다. 매크로블록들은 매크로블록들에 적용된 코딩 모드들 및/또는 코딩 기술들을 정의하기 위한 신택스 (syntax) 정보를 포함할 수도 있다.

[0038] 매크로블록들 또는 다른 비디오 블록들은 슬라이스들, 프레임들 또는 다른 독립적인 유닛들과 같은 디코딩 가능한 유닛들로 그룹화될 수도 있다. 각각의 슬라이스는 비디오 프레임의 독립적으로 디코딩 가능한 유닛일 수도 있다. 대안적으로, 프레임들은 자체적으로 디코딩 가능한 유닛들이 될 수도 있거나, 프레임의 다른 부분들이 디코딩 가능한 유닛들로 정의될 수도 있다. 본 개시물에서, 용어 "코딩된 유닛" 은 전체 프레임, 프레임의 슬라이스, 픽처 그룹 (GOPs), 또는 이용되는 코딩 기술들에 따라 정의되는 다른 독립적으로 디코딩 가능한 유닛과 같은 비디오 프레임의 임의의 독립적으로 디코딩 가능한 유닛을 지칭한다.

[0039] 인트라-기반 또는 인터-기반 예측 인코딩 이후 및 (H.264/AVC 에서 이용되는 4×4 또는 8×8 정수 변환 또는 이산 코사인 변환 또는 DCT 와 같은) 임의의 변환들 이후에, 양자화가 수행될 수도 있다. 양자화는 일반적으로 계수들을 나타내기 위해 이용되는 데이터의 양을 감소시킬 수 있도록 계수들이 양자화되는 프로세스를 지칭한다. 양자화 프로세스는 계수들 중 일부 또는 전부와 연관된 비트 깊이를 감소시킬 수도 있다. 예를 들면, 16-비트 값은 양자화 동안 15-비트 값으로 절단 (round down) 될 수도 있다. 양자화 이후, 예컨대, 콘텐츠 적응형 가변 길이 코딩 (CAVLC), 컨텍스트 적응형 이진 산술 코딩 (CABAC), 또는 다른 엔트로피 코딩 방법에 따라 엔트로피 코딩이 수행될 수도 있다.

[0040] 3D 비디오는 각각의 최초 인코딩된 프레임과 연관된 하나 이상의 추가의 비디오 프레임들 (예컨대, 추가의 뷰들) 을 요청할 수도 있다. 예를 들어, 2 개의 서로 다른 뷰들은 비디오 프레임의 입체적인 3D 렌디션 (rendition) 을 정의하기 위해 이용될 수도 있다. 3 개 이상의 뷰들을 포함할 수도 있는 다중 뷰들은 멀티-뷰 3D 렌디션들을 지원할 수도 있다. 3D 비디오의 서로 다른 뷰들은 2 이상의 뷰들이 비디오 시퀀스의 동일한 시간 인스턴스에 대응하도록 유사한 타이밍을 가질 수도 있다. 이러한 방식으로, 2 이상의 뷰들은 일반적으로 3D 시퀀스를 함께 형성하는 2 이상의 2D 시퀀스들을 정의할 수도 있고, 이러한 3D 시퀀스는 3D 비디오를 제공하도록 집합적으로 렌더링될 수도 있다.

[0041] 3D 비디오의 효율적인 인코딩, 통신 및 디코딩을 지원하기 위해, 본 개시물은 2D 시퀀스의 비디오 프레임들 각각에 적용되어 3D 비디오 데이터를 생성할 수도 있는 파라미터 세트를 포함하는 3D 컨버전 정보를 이용한다. 그러한 3D 컨버전 정보는 2D 시퀀스와 함께 통신될 수도 있다. 따라서, 수신 디바이스는 2D 시퀀스를 생성하고 디스플레이할 수 있거나, 수신 디바이스가 3D 비디오를 지원하면, 수신 디바이스는 3D 시퀀스를 생성하고 디스플레이할 수 있다. 일부 실시예들에서, 본 개시물의 3D 컨버전 정보는 100 바이트 미만의 데이터 및 더욱 상세하게는 20 바이트 미만의 데이터를 포함할 수도 있으며, 이들 데이터는 2D 시퀀스의 2D 프레임들 전부 또는 일부에 적용되어 3D 입체 비디오를 위한 2차 뷰를 생성할 수 있다. 이러한 방식으로, 적어도 일부 프레임들에 대하여 2 개의 뷰들의 전송을 제어함으로써, 본 개시물의 기술들은 3D 비디오를 통신하는 효율적인 방식을 제공한다.

[0042] 도 2 는 본 개시물과 부합하는 기술들을 수행할 수도 있는 비디오 인코더 (50) 의 일 실시예를 예시한 블록도이다. 비디오 인코더 (50) 는 소스 디바이스 (12) 의 비디오 인코더 (22) 또는 다른 디바이스의 비디오 인코더에 대응할 수도 있다. 비디오 인코더 (50) 는 비디오 프레임들 내의 블록들의 인트라-코딩 및 인터-코딩을 수행할 수도 있다. 인트라-코딩은 공간적 예측에 의존하여, 소정의 비디오 프레임 내에서 비디오의 공간



적 리턴던시를 감소 또는 제거한다. 인터-코딩은 시간적 예측에 의존하여, 비디오 시퀀스의 인접 프레임들 내에서 비디오의 시간적 리턴던시를 감소 또는 제거한다. 인트라-모드 (I-모드) 는 공간 기반 압축 모드를 지칭할 수도 있고, 예측 모드 (P-모드) 또는 양방향 모드 (B-모드) 와 같은 인터-모드들은 시간 기반 압축 모드를 지칭할 수도 있다.

[0043] 도 2 에 도시된 바와 같이, 비디오 인코더 (50) 는 인코딩될 비디오 프레임 또는 슬라이스 내의 현재 비디오 블록을 수신한다. 도 2 의 예에서, 비디오 인코더 (50) 는 예측 유닛 (35), 메모리 (34), 합산기 (48), 변환 유닛 (38), 양자화 유닛 (40), 및 엔트로피 코딩 유닛 (46) 을 포함한다. 비디오 블록 복원을 위해, 비디오 인코더 (50) 는 또한 역 양자화 유닛 (42), 역변환 유닛 (44) 및 합산기 (51) 를 포함한다. 또한, 본 개시물에 따라, 비디오 인코더 (50) 는 본 명세서에서 설명되는 3D 컨버전 정보를 생성하는, 2D - 3D 컨버전 유닛 (36) 을 포함할 수도 있다. 비디오 인코더 (50) 는 또한 복원된 비디오로부터 블록화 아티팩트들 (blockiness artifacts) 을 제거하기 위해 블록 경계들을 필터링하는 디블로킹 (deblocking) 필터 (비도시) 와 같은 다른 컴포넌트들을 포함할 수도 있다. 요구되는 경우, 디블로킹 필터는 통상적으로 합산기 (51) 의 출력을 필터링할 것이다.

[0044] 인코딩 프로세스 동안, 비디오 인코더 (50) 는 코딩될 비디오 블록을 수신하고, 예측 유닛 (35) 은 인트라-예측 또는 인터-예측 코딩을 수행한다. 예를 들어, 인코더 (50) 의 예측 유닛 (35) 은 코딩된 유닛 (예컨대, 프레임 또는 슬라이스) 의 각각의 비디오 블록 또는 비디오 블록 분할에 대한 모션 추정 및 모션 보상을 수행할 수 있다. 예측 유닛 (35) 은 특정 블록을 인코딩하는 것과 연관된 각각의 적용가능한 모드에 대한 레이트-왜곡 비용 (rdcost) 을 계산할 수도 있고, 최저 비용을 산출하는 코딩 모드를 선택할 수도 있다. rdcost 는 이용되는 비트들의 수 및 오리지널 비디오 데이터에 대하여 코딩된 데이터에서의 왜곡의 레벨과 관련하여 비용을 정량화할 수도 있다.

[0045] 레이트-왜곡 (RD) 분석은 비디오 코딩에서 상당히 통상적이며, 일반적으로 코딩 비용을 나타내는 비용 메트릭의 계산을 수반한다. 비용 메트릭은 코딩을 위해 요구되는 비트들의 수 (레이트) 와 코딩과 연관된 품질 레벨 (왜곡) 사이에서 균형을 이룰 수도 있다. 통상적인 레이트-왜곡 비용 계산은 일반적으로 하기의 포맷에 대응할 수도 있다:

[0046] 
$$J(\lambda) = \lambda R + D,$$

[0047] 상기  $J(\lambda)$  는 비용이고, R 은 비트 레이트이고, D 는 왜곡이며,  $\lambda$  는 라그랑제 (Lagrange) 승수이다. 예측 유닛 (35) 은 비디오 블록 인코딩을 수행하기 위해 이용될 수 있는 다양한 인트라-코딩 및 인터-코딩 모드들 (및 적용가능한 분할 사이즈들) 을 비교하기 위해 상기와 같은 타입의 비용 함수를 적용할 수도 있다.

[0048] 요구되는 예측 데이터가 예측 유닛 (35) 에 의해 식별되면, 비디오 인코더 (50) 는 잔여 블록을 생성하기 위해 코딩되는 오리지널 비디오 블록으로부터 예측 데이터를 감산함으로써 잔여 비디오 블록을 형성한다. 합산기 (48) 는 이러한 감산 연산들을 수행하는 컴포넌트 또는 컴포넌트들을 표시한다. 변환 유닛 (38) 은 잔여 변환 블록 계수들을 포함하는 비디오 블록을 생성할 때, 이산 코사인 변환 (DCT) 또는 개념적으로 유사한 변환과 같은 변환을 잔여 블록에 적용한다. 변환 유닛 (38) 은 DCT 와 개념적으로 유사한 H.264 표준에 의해 정의되는 것과 같은 변환들을 수행할 수도 있다. 웨이블릿 (wavelet) 변환들, 정수 변환들, 서브-대역 변환들 또는 다른 타입의 변환들이 이용될 수 있다. 임의의 경우에, 변환 유닛 (38) 은 잔여 변환 계수들의 블록을 생성할 때 변환을 잔여 블록에 적용한다. 변환은 잔여 정보를 픽셀 영역으로부터 주파수 영역으로 컨버팅할 수도 있다.

[0049] 양자화 유닛 (40) 은 잔여 변환 계수들을 양자화하여, 비트 레이트를 더 감소시킨다. 양자화 프로세스는 이러한 계수들의 전부 또는 일부와 연관된 비트 깊이를 감소시킬 수도 있다. 예를 들면, 9-비트 값은 양자화 동안 8-비트 값으로 절단될 수도 있다. 추가로, 양자화 유닛 (40) 은 오프셋이 이용되는 경우들에 대하여 서로 오프셋들을 양자화할 수도 있다.

[0050] 양자화 이후에, 엔트로피 코딩 유닛 (46) 은 양자화된 변환 계수들을 엔트로피 코딩한다. 예를 들면, 엔트로피 코딩 유닛 (46) 은 콘텐츠 적응형 가변 길이 코딩 (CAVLC), 콘텍스트 적응형 이진 산술 코딩 (CABAC), 또는 다른 엔트로피 코딩 방법을 수행할 수도 있다. 엔트로피 코딩 유닛 (46) 에 의한 엔트로피 코딩 이후에, 인코딩된 비디오는 다른 디바이스로 전송되거나, 이후 전송 또는 검색을 위해 보관될 수도 있다. 코딩된 비트스트림은 엔트로피 코딩된 잔여 블록들, 이러한 블록들에 대한 모션 벡터들 및 2D - 3D 비디오를 지원하기 위

해 본 명세서에 설명된 선택스와 같은 다른 선택스를 포함할 수도 있다.

- [0051] 역 양자화 유닛 (42) 및 역 변환 유닛 (44) 은 각각 역 양자화 및 역 변환을 적용하여, 픽셀 영역 내의 잔여 블록을, 예컨대 향후 이용을 위해 전송된 방식의 참조 데이터로서 복원한다. 합산기 (51) 는 복원된 잔여 블록을 모션 보상 유닛 (35) 에 의해 생성된 1차 및/또는 2차 예측 블록들에 합산하여 메모리 (34) 내의 저장을 위한 복원된 비디오 블록을 생성한다. 복원된 비디오 블록 및 잔여 데이터는 모션 보상 유닛 (35) 에 의해 후속 비디오 프레임 또는 다른 코딩된 유닛 내의 블록을 인터-인코딩하기 위한 참조 블록으로서 이용될 수도 있다.
- [0052] 3D 비디오를 지원하기 위해, 비디오 인코더 (50) 는 메모리 (34) 내에 저장된 복원된 2D 비디오 시퀀스와 관련하여 동작하는 2D - 3D 컨버전 유닛 (36) 을 추가로 포함할 수도 있다. 이러한 방식으로, 2D - 3D 컨버전 유닛 (36) 은 디코딩 프로세스 이후에 디코더에서 이용가능할 동일한 복원 데이터 상에 동작한다. 본 개시물에 따라, 2D - 3D 컨버전 유닛 (36) 은 2D 시퀀스의 비디오 프레임들 각각에 적용되어 3D 비디오 데이터를 생성할 수 있는 파라미터 세트를 포함하는 3D 컨버전 정보를 식별하거나, 결정하거나, 생성한다. 3D 컨버전 정보는 소정의 2D 시퀀스에 대하여 한 번 생성될 수도 있다.
- [0053] 3D 컨버전 정보가 적용하는 2D 시퀀스는 전체 비디오 시퀀스, 장면 (scene), 또는 가능하게는 디코딩가능한 세트를 형성하는 픽처 그룹을 포함할 수도 있다. 픽처 그룹들의 예들은 IBPBP 시나리오에서 5 개 프레임 세트 또는 IBBPBP 시나리오에서 7 개 프레임 세트를 포함하며, 이 경우 I 는 인트라-코딩을 나타내고, P 는 예측 인터-코딩 또는 단방향 인터-코딩을 나타내고, B 는 양방향-예측 (bi-predictive) 인터-코딩 또는 양방향 인터-코딩을 나타낸다. 이 경우에, 픽처 그룹 내의 프레임들은 상호의존적이고, 집합적으로 디코딩 가능하다. 일부 경우에, 3D 컨버전 정보는 픽처 그룹마다 한 번 전송될 수도 있지만, 3D 컨버전 정보는 장면마다 한 번 또는 전체 비디오 시퀀스마다 한 번 전송될 수 있다. 그러나, 중요한 것은, 3D 컨버전 정보는 복수의 프레임들에 대하여 적용되고, 따라서 복수의 프레임들의 각각의 개별 프레임에 대하여 서로 다른 3D 컨버전 정보가 요구되지 않는다.
- [0054] 비디오 인코더 (50) 는 ITU H.264 비디오 인코딩 표준에 따라 2D 시퀀스를 인코딩할 수도 있고, 2D - 3D 컨버전 유닛은 ITU H.264 비디오 인코딩 표준에 의해 지원되는 3D 컨버전 정보 SEI 메세지들을 인코딩할 수도 있다. 파라미터 세트는 제 1 의 2D 시퀀스의 비디오 프레임들 각각에 적용되어 비디오 프레임들의 제 2 의 2D 시퀀스를 생성할 수 있으며, 상기 제 1 및 제 2 의 2D 시퀀스들은 집합적으로 3D 입체 비디오 시퀀스를 정의한다. 3D 컨버전 정보는 2D 시퀀스에 적용되어 3D 비디오 데이터를 생성하기 위한 3D 컨버전 프로세스를 식별하는 정보를 포함할 수도 있다. 일부 경우에, 3D 컨버전 정보는 2D 시퀀스를 캡처하는 것과 연관된 카메라 파라미터들 및 값들을 포함할 수도 있다. 예를 들어, 하기에서 더 상세히 설명되는 것과 같이, 3D 컨버전 정보는 2D 시퀀스를 캡처한 카메라와 연관된 초점 길이를 나타내는 초점 길이 값, 3D 비디오 데이터에서의 최소 깊이를 명시하는 근거리-깊이 값, 3D 비디오 데이터에서의 최대 깊이를 명시하는 원거리-깊이값, 및 3D 비디오 데이터와 연관된 2 개의 카메라들 사이에서 추정된 거리를 정량화하는 평행 이동 값을 포함할 수도 있다.
- [0055] 3D 컨버전 정보를 인코딩하기 위해, 2D - 3D 컨버전 유닛 (36) 은 특정 시나리오들을 나타내도록 설정될 수도 있는 비트들인, 플래그들을 이용할 수도 있다. 예를 들면, 3D 컨버전 정보는 명시적인 3D 파라미터 세트가 3D 컨버전 정보 내에 포함되는지의 여부 또는 디폴트 3D 파라미터 세트가 이용될 것인지 여부를 나타내는 플래그를 포함할 수도 있다. 이 경우에, 명시적인 3D 파라미터 세트는 플래그가 설정된 경우에 3D 컨버전 정보 내에 포함된다. 또한, 3D 컨버전 정보는 2D 시퀀스의 2차 뷰가 2D 시퀀스의 좌측에 생성될 것인지 또는 우측에 생성될 것인지 여부를 나타내는 플래그를 포함할 수도 있다. 추가로, 3D 컨버전 정보는 3D 비디오 데이터로부터 제거될 크롭 (crop) 영역을 식별하는 플래그를 포함할 수도 있다. 이 경우에, 크롭 영역을 정의하는 정보는 플래그가 설정될 때 3D 컨버전 정보 내에 포함된다. 이러한 플래그들은 각각 3D 컨버전 정보를 집합적으로 인코딩하기 위해 이용될 수도 있고, 다른 플래그들은 본 개시물과 일치하여 이용되거나 정의될 수도 있다.
- [0056] 도 3 은 전송된 인코딩 기술들에 대한 상호 디코딩 기술들을 수행할 수 있는 예시적인 비디오 디코더 (70) 를 예시한 블록도이다. 비디오 디코더 (70) 는 엔트로피 디코딩 유닛 (72), 예측 유닛 (75), 역 양자화 유닛 (76), 역 변환 유닛 (78), 메모리 (74), 및 합산기 (79) 를 포함할 수도 있다. 예측 유닛 (75) 은 모션 보상 유닛뿐만 아니라 공간 예측 컴포넌트들을 포함할 수도 있다.
- [0057] 비디오 디코더 (70) 는 본 명세서에서 설명되는 방식으로 인코딩된 2D 시퀀스 및 비디오 블록들의 적절한 디코딩을 용이하게 하기 위해 디코더 (70) 에 의해 이용될 수도 있는 다양한 선택스 엘리먼트들을 포함하는 인코딩

된 비디오 비트스트림을 수신할 수도 있다. 더욱 상세하게는, 비디오 비트스트림은 비디오 프레임들의 2D 시퀀스에 기초하여 3D 비디오 데이터의 생성을 용이하게 하기 위해 본 명세서에서 설명된 3D 컨버전 정보를 포함할 수도 있다. 3D 컨버전 정보는 2D 시퀀스의 비디오 프레임들 각각에 적용되어 3D 비디오 데이터를 생성할 수 있는 파라미터 세트를 포함할 수도 있다. 다시 말해서, 3D 컨버전 정보는 비디오 시퀀스의 상이한 프레임들에 대하여 달라지는 것이 아니라, 2D 시퀀스의 오리지널 비디오 프레임들 각각에 적용되어 오리지널 비디오 프레임들 각각의 2차 뷰를 생성할 수 있는, 비교적 소량의 데이터를 형성하는 공통 파라미터 세트를 포함한다. 어구 (phrase) 2D 시퀀스는 전체 비디오 파일, 비디오 클립, 더 큰 비디오 파일 내의 비디오 장면 또는 가능하면 더 큰 비디오 시퀀스 내에 디코딩 가능한 프레임 세트를 형성하는 픽처 그룹을 포함할 수도 있는 복수의 비디오 프레임들을 지칭한다.

[0058] 엔트로피 디코딩 유닛 (72) 은 2D 시퀀스의 잔여 비디오 블록들에 대하여 양자화된 계수들을 생성하기 위해 비트스트림의 엔트로피 디코딩을 수행한다. 엔트로피 디코딩 유닛 (72) 은 비트스트림으로부터 선택스 엘리먼트들을 분석하여 이러한 선택스 엘리먼트를 비디오 디코더 (70) 의 다양한 유닛들로 포워딩할 수도 있다. 모션 정보 (예컨대, 모션 벡터들) 및 다른 선택스 정보는 예컨대, 예측 유닛 (75) 으로 포워딩될 수도 있다. 또한, 3D 컨버전 정보는 엔트로피 디코딩 유닛 (72) 에 의해 비트스트림으로부터 분석된 후, 2D - 3D 컨버전 유닛 (79) 로 포워딩될 수도 있다. 2D - 3D 컨버전 유닛 (79) 은 이 유닛 (79) 이 선택적이기 때문에 파선으로 도시된다. 2D - 3D 컨버전 유닛 (79) 을 가지지 않는 비디오 디코더는 단지 2D 비디오 시퀀스를 디코딩하고, 3D 컨버전 정보를 폐기한다. 이러한 방식으로, 3D 컨버전 정보는 일부 디바이스들이 2D 비디오를 디코딩할 수 있지만, 다른 디바이스들은 2D 비디오를 디코딩하고 또한 3D 컨버전 정보에 적용하여 3D 비디오를 생성할 수 있는 스케일러블 2D - 3D 비디오를 가능하게 한다.

[0059] 2D 비디오 시퀀스를 디코딩하기 위해, 예측 유닛 (75) 은 인코딩에서 이용되었던 예측 블록을 식별하기 위해 모션 정보 또는 다른 선택스 엘리먼트들을 이용한다. 인터-기반 디코딩을 위해, 모션 벡터는 모션 보상 프로세스에 적용되어, 하나 이상의 예측 데이터 리스트들로부터 예측 블록을 생성할 수도 있다. 인터-기반 디코딩을 위해, 선택스는 인트라 모드를 포함할 수도 있으며, 인트라 모드는 디코딩되고 있는 비디오 블록과 동일한 프레임으로부터의 데이터에 기초하여 예측 블록이 생성되는 방법을 정의할 수도 있다.

[0060] 역 양자화 유닛 (76) 은 엔트로피 디코딩 유닛 (72) 으로부터 데이터를 역 양자화하고, 역 변환 유닛은 역 변환을 수행하여 픽셀 영역 내에 잔여 블록을 생성한다. 그 후에, 합산기 (79) 는 예측 유닛 (75) 에 의해 생성된 예측 블록과 잔여 블록을 결합하여 오리지널 비디오 블록의 복원을 생성하며, 이는 메모리 (74) 내에 저장되고 및/또는 디코딩된 2D 비디오 데이터로서 디스플레이로 출력될 수도 있다. 다수의 비디오 블록들은 비디오 프레임들을 복원하고, 궁극적으로는 메모리 (74) 내의 비디오 프레임들의 전체 2D 시퀀스를 복원하기 위해 상기와 같은 방식으로 디코딩될 수도 있다. 이러한 방식으로, 비디오 디코더 (70) 는 전송된 비디오 인코더 (50) 에 의해 수행된 인코딩에 대하여 상호 디코딩을 수행한다.

[0061] 본 개시물에 따라, 2D - 3D 컨버전 유닛 (79) 은 비트스트림 내에서 전송된 3D 컨버전 정보를 메모리 (74) 내에 저장된 디코딩된 2D 비디오 시퀀스에 적용할 수도 있다. 예를 들어, 3D 컨버전 정보를 메모리 (74) 내에 저장된 디코딩된 2D 비디오 시퀀스에 적용할 때, 2D - 3D 컨버전 유닛 (79) 은 2D 시퀀스 내의 비디오 프레임들 전부 또는 일부와 연관된 2차 뷰들을 생성할 수도 있다. 오리지널 뷰들 및 2차 뷰들은 2D - 3D 컨버전 유닛 (79) 으로부터 3D 비디오 데이터로서 출력될 수도 있다.

[0062] 멀티-뷰 비디오 코딩 (MVC) 은 H.264/AVC 로의 확장을 형성하지만, MVC 는 다른 비디오 코딩 표준들에 대하여 적용될 수 있다. MVC 의 하나의 공통안 (joint draft) 은 JVT-AB204, "Joint Draft 8.0 on Multiview Video Coding," 28<sup>th</sup> JVT meeting, Hannover, Germany, July 2008 에서 설명된다. H.264/AVC 와 부합하여, 코딩된 비디오 비트들은 비디오 전화, 저장, 브로드캐스트 또는 스트리밍과 같은 애플리케이션들을 어드레싱하는 "네트워크-프렌들리" 비디오 표현을 제공하는 NAL (Network Abstraction Layer) 유닛들로 구성될 수도 있다. NAL 유닛들은 VCL (Video Coding Layer) NAL 유닛들 및 비-VCL NAL 유닛들로 분류될 수 있다. VCL 유닛들은 코어 압축 엔진을 포함하고, 블록, MB 및 슬라이스 레벨들을 포함할 수도 있다. 다른 NAL 유닛들은 비-VCL NAL 유닛들이다.

[0063] H.264/AVC 에 따라, SEI (Supplemental Enhancement Information) 메세지들은 VCL NAL 유닛들로부터 코딩된 픽처들의 샘플들을 디코딩하는데 필요하지 않은 정보를 포함할 수도 있다. SEI 메세지들은 또한 비-VCL NAL 유닛들 내에 포함된다. SEI 메세지들은 H.264/AVC 에 대한 표준 규격의 규범적 (normative) 부분이다. 표준-순응 디코더 구현에 있어 필수적인 것은 아니지만, SEI 메세지들은 디코딩, 디스플레이, 에러 내성 및 다

른 목적들과 관련된 프로세스들을 보조할 수도 있다. H.264/AVC 의 필수적인 부분은 완결되었지만, H.264/AVC 규격은 여전히 SEI 메세지들에 대하여 열려있다. 본 개시물은, 일부 양태들에서, 본 명세서에 설명된 3D 컨버전 정보를 인코딩하고, 통신하고, 디코딩하기 위한 메커니즘으로서 SEI 메세지들 또는 다른 유사한 메세지들의 이용을 제안한다.

[0064] 3D 비디오 포맷을 지원하기 위해, JVT 및 MPEG 는 새로운 표준들 및 특징들을 도입할 수도 있다. MPEG-2 멀티뷰 프로파일과 일치하여, 예컨대, 좌측 뷰와 같은 하나의 뷰가 감소된 프레임 레이트 (예컨대, 초당 15 프레임) 로 코딩될 수도 있고, 다른 뷰는 더 높은 프레임 레이트 (예컨대, 초당 30 프레임) 를 요구하는 시간 향상 계층으로서 코딩될 수도 있다. 그러나, 이는 비트스트림 내에서 양자의 뷰들의 통신을 요구하며, 이는 종래의 2D 시퀀스의 통신과 비교하여 비트스트림 내의 데이터의 양을 상당히 증가시킬 수 있다.

[0065] H.264/AVC 는 또한 입체 비디오 콘텐츠 접근방식들을 적용할 수도 있다. H.264/AVC 에서, 예컨대, 스테레오 비디오 정보 SEI 메세지는 하나의 비트스트림 내에 2 개의 뷰들이 배열되는 방식을 나타내기 위해 사용될 수 있다. 이러한 경우에, 2 개의 뷰들은 교번하는 프레임들 또는 상보적 필드 쌍들일 수도 있다. 2 개의 뷰들이 교번하는 프레임들인 경우, 2 개의 뷰들은 시간 인터리빙 모드로 순서화되며, 상보적 필드 쌍들에 대한 2 개의 뷰들인 경우에, 2 개의 뷰들로부터의 이미지 쌍은 실제로 하나의 픽처 내에 인터리빙된 로우 (row) 이다. 그러나, 다시 말해서, 이는 비트스트림 내에서 양자의 뷰들의 통신을 요구하며, 이는 종래의 2D 시퀀스의 통신과 비교하여 비트스트림 내의 데이터의 양을 상당히 증가시킬 수 있다.

[0066] 픽처들의 공간 인터리빙을 채택하는 것이 가능할 수도 있고, SEI 내의 공간 인터리빙의 존재를 시그널링 할 수도 있으며, 이는 시간 인터리빙 및 로우 인터리빙으로부터의 2 개의 뷰들의 지원을 더 유연한 공간 인터리빙 모드들로 확대한다. 대안적으로, SEI 메세지는 사이드-바이-사이드 (side-by-side) 인터리빙, 상/하 (above/below) 인터리빙, 컬럼 인터리빙, 또는 체커보드 (checkerboard) 인터리빙으로서 이미지 쌍의 결합을 지원할 수도 있다. 이러한 접근 방식들의 각각에서, 다른 바람직하지 않은 접근 방식들과 같이, 상이한 뷰들은 실제로 일부 방식으로 비트스트림 내에서 전송되며, 이는 종래의 2D 시퀀스와 비교하여 3D 비디오에 대하여 요구되는 정보의 양을 상당히 증가시킨다.

[0067] 본 개시물의 기술들은 2차 뷰들의 실제 전송을 방지함으로써 3D 비디오를 전달하기 위해 비트스트림 내에서 요구되는 정보의 양을 상당히 감소시킬 수도 있다. 이러한 경우에, 2차 뷰들을 실제로 인코딩하고 전송하는 것보다, 본 개시물은 디코더에서 오리지널 2D 시퀀스에 적용되어 디코더에서 2차 뷰들을 생성할 수 있는 파라미터 세트를 포함하는 3D 컨버전 정보를 전송한다. 이러한 방식으로, 실제로 2차 뷰들을 전달해야 하는 필요성이 방지되고, 대신에 파라미터 세트는 디코더에 적용되어 비트스트림내에서 2차 뷰들을 전달할 필요없이 2차 뷰들을 생성할 수 있다.

[0068] H.264/AVC 에서 SEI 메세지들은 3D 입력이 2D 비디오 시퀀스였던 것처럼 3D 입력을 수용하고, 2 개의 인터리빙된 뷰들을 구별하는 방법을 디코더에 알리기 위해 SEI 메세지를 활용함으로써 하나의 비디오 시퀀스가 2 개의 뷰들로 분할될 수 있도록 하는 케이스를 지원할 수도 있다. 재차 강조되지만, 그러한 케이스로의 입력은 2 개의 뷰들에 의해 표시되는 3D 입력이다. 대조적으로, 본 개시물의 기술들은 비트스트림 내에서 2 개의 뷰들을 전송해야 하는 필요성을 방지하며, 대신에 비트스트림 내의 3D 컨버전 정보에 기초하여 임의의 2차 뷰들을 생성하기 위해 디코더에 의존한다.

[0069] 일부 MVC 디코딩 순서들은 시간-우선 (time-first) 코딩으로 지칭될 수도 있다. 이 경우에, 각각의 액세스 유닛은 하나의 출력 시간 인스턴스에 대한 모든 뷰들의 코딩된 픽처들을 포함하는 것으로 정의된다. 그러나, 액세스 유닛들의 디코딩 순서는 출력 또는 디스플레이 순서와 동일하지 않을 수도 있다.

[0070] MVC 예측은 각각의 뷰 내의 인터-픽처 예측과 인터-뷰 예측 양자를 포함할 수도 있다. MVC 는 H.264/AVC 디코더들에 의해 디코딩될 수 있는, 이른바 기본 뷰를 포함할 수도 있으며, 2 개의 뷰들은 또한 MVC 에 의해 지원될 수 있다. 이 경우에, MVC 의 장점은 3 이상의 뷰들을 3D 비디오 입력으로서 취득하고 다수의 뷰들에 의해 표시되는 이 3D 비디오를 디코딩하는 케이스를 지원할 수 있다는 것이다. MVC 디코더에 의한 디코딩은 다수의 뷰들을 가지는 3D 비디오 콘텐츠를 예상할 수도 있다.

[0071] MPEG 표준들은 또한 MPEG-C 파트 3 에서 정규 비디오 스트림에 대하여 깊이 맵 (depth map) 을 부가하기 위한 포맷을 명시한다. 이 규격은 하기에 포함된다:



“Text of ISO/IEC FDIS 23002-3 Representation of Auxiliary Video and Supplemental Information,” *ISO/IEC JTC 1/SC 29/WG 11, MPEG Doc, N8768*, Marrakech, Morocco, January 2007.

[0072]

[0073]

MPEG 파트 3 에서, 이른바 보조 비디오는 깊이 맵 또는 시차 (parallax) 맵 일 수 있다. 깊이 맵을 표시하는 것은 깊이 맵의 각각의 깊이 값 및 해상도를 표시하기 위해 사용되는 비트 수와 관련하여 융통성을 제공할 수도 있다. 예를 들어, 깊이 맵은 소정의 이미지의 폭의 1/4 및 높이의 1/2 일 수도 있다. 유감스럽게도, 깊이 맵들은 통상적으로 시퀀스의 모든 프레임에 대하여 요구된다. 즉, 동일한 깊이 맵은 비디오 시퀀스의 모든 프레임에 적용하지 않는다. 따라서, 깊이 맵들의 통신은 몇몇 깊이 맵들이 전체 비디오 시퀀스에 걸쳐 요구되기 때문에 그 합계가 매우 상당한 양의 데이터에 달할 수 있다.

[0074]

MPEG 비디오 서브그룹은 3D 시나리오들을 연구하기 위해 3D 비디오 코딩에서 탐구 실험들을 정의하였다. MPEG 비디오 서브그룹은 각각의 뷰에 대하여 깊이 맵들을 가지는 것이 잠재적으로는 뷰 합성에 도움이 되지만, MPEG 에서 이 행위는 표준화의 일부가 아닐 수도 있다고 나타내었다. MPEG 3D 비디오와 일치하여, 2 개의 중요한 개념들은 깊이 추정 및 뷰 합성을 포함한다. 대부분의 비디오 콘텐츠는 멀티-카메라 시스템에 의해 캡처되고, 깊이 맵들은 코딩 이전에 생성되어 이 깊이 맵들이 텍스처 비디오 시퀀스들과 함께 전송될 수 있어야 한다고 가정될 수 있다. 그러나, 본 개시물에 따른 뷰 합성은 비트스트림 내에서 전송되지 않는 더 많은 뷰들을 생성하기 위해 비디오 렌더링 시기에 적용될 수도 있는 수단이다. 따라서, 뷰 합성의 개념은 본 명세서에서 설명되는 3D 컨버전 정보를 용이하게 함으로써 본 개시물의 기술들의 일부를 형성할 수도 있다.

[0075]

3D 비디오 통신 시스템들에서, 원시 비디오 데이터는 캡처되고, 인코딩 이전에 사전-처리된다. 깊이 맵들을 가질 수도 있는 원시 데이터는 인코딩될 수도 있고, 인코딩된 비디오 콘텐츠는 저장되거나 전송될 수도 있다. 목적지 디바이스는 3D 비디오를 디코딩하고 디스플레이할 수도 있다. 그러나, 전송된 것과 같이, 통신 및 대역폭의 관점에서, 비디오 시퀀스의 몇몇 이미지들에 대하여 추가의 뷰들 또는 깊이 맵들을 통신하는 것은 바람직하지 않을 수도 있다. 본 개시물에 따른 더 양호한 접근 방식은, 예컨대 디코더에 의해 2D 시퀀스의 모든 비디오 프레임에 적용되어, 실제로 비트스트림 내에서 2차 뷰들을 통신할 필요 없이 2차 뷰들을 생성할 수 있는 3D 컨버전 정보를 통신하는 것일 수도 있다.

[0076]

3D 콘텐츠의 캡처는 하나의 카메라 또는 카메라 어레이에 의해 수행될 수도 있거나, 심지어 깊이 맵들을 생성할 수 있는 디바이스들과 연관될 수도 있다. 일부 예들에서와 같이, 콘텐츠 캡처는 하기의 카테고리들 중 적어도 하나로 분류될 수도 있다:

[0077]

- 일반적으로 3D 콘텐츠를 제공하지 않는 2D 비디오 캡처.

[0078]

- 입체 비디오를 캡처 및/또는 제공할 수 있는 2-카메라 시스템들.

[0079]

- 다수의 뷰들을 캡처하는 카메라 어레이들.

[0080]

- 하나의 뷰 캡처와 깊이. 예컨대, 일부 디바이스들은 캡처된 이미지들과 연관된 깊이들을 캡처할 수 있다.

[0081]

- 깊이 정보를 캡처하고 및/또는 3D 모델을 생성할 수 있는 다른 기술들.

[0082]

3D 사전-프로세싱 및 인코딩은 또한 인코더에서 수행될 수도 있다. 여기에서, 3D 사전-프로세싱은 잡음 제거 또는 장면 검출과 관련된 통상적인 프로세싱을 지칭하지 않는다. 3D 사전-프로세싱은 3D 비디오 콘텐츠의 일부로서 코딩될 깊이 맵들을 생성할 수도 있다. 이러한 프로세스는 몇몇 전송된 뷰들에 대하여 깊이 맵들 또는 모든 캡처된 뷰에 대하여 하나의 깊이 맵들을 생성할 수 있다. 그러나, 깊이 맵들의 통신은 대역폭의 관점에서는 바람직하지 않을 수도 있다.

[0083]

비디오 콘텐츠가 디코더에 의해 수신되는 경우에, 하나 이상의 뷰들 뿐만 아니라 임의의 경우에 복원된 깊이 맵들을 포함할 수도 있는 전송된 데이터를 획득하도록 디코딩될 수 있다. 깊이 맵들이 디코더에서 사용가능한 경우에, 뷰 합성 알고리즘은 전송되지 않은 다른 뷰들에 대한 텍스처를 생성하도록 채택될 수 있다. 일반적인 3D 디스플레이들은 2 이상의 뷰들을 렌더링할 수도 있다. 고 프레임 레이트 비디오들을 디스플레이할 수 있는 일부 2D 디스플레이들은 입체 안경 (shutter glasses) 의 도움으로 3D 디스플레이로서 이용될 수 있다. 편광화 (polarization) 는 2 개의 뷰들을 출력으로서 제공하는 3D 디스플레이 기술이다. 일부 디스플레이들 또는 3D 텔레비전이 입력의 일부로서 깊이를 취득하고 있지만, 2 이상의 뷰들을 출력으로 생성해야 하는 빌트인 "뷰 합성" 모듈이 항상 존재해야 한다.

[0084] 3D 래핑 (wrapping) 은 본 개시물의 기술들에 유용할 수도 있는 뷰 합성의 하나의 형태이다. 도 4 내지 도 7 은 3D 래핑 및 다른 뷰 합성 개념들을 설명하기 위해 이용될 개념도들이다. 샘플링 이론에 기초한 뷰 합성은 임의의 뷰 각도에서 임의의 뷰를 완벽하게 생성하기 위해 고밀도로 (densely) 샘플링된 뷰들을 요구하는 샘플링 문제가 될 수도 있다. 그러나, 실질적인 접근방식들에서, 조밀하게 샘플링된 뷰들에 의해 요구되는 저장 또는 전송 대역폭은 통상적으로 실현되기에 너무 크다. 따라서, 일부 연구는 저밀도로 (sparsely) 샘플링된 뷰들 및 깊이 맵들에 기초하는 뷰 합성에 초점을 맞춰왔다.

[0085] 저밀도로 샘플링된 뷰들에 기초한 뷰 합성 알고리즘들은 3D 래핑의 개념에 의존할 수도 있다. 도 4 는 3D 래핑의 개념을 예시한다. 도 4 에 도시된 것과 같이, 3D 래핑에서, 깊이 및 카메라 모델이 주어질 때, 참조 뷰의  $\overline{u_1}$  내의 픽셀은 2D 카메라 좌표로부터 월드-스페이스 좌표 시스템 내의 포인트  $P$  로 투영될 수 있다.

그 후에, 포인트  $P$  는  $\overrightarrow{PO_2}$  의 방향을 따라 (생성될 가상 뷰인) 목적지 뷰로 투영될 수도 있다. 이 경우,  $\overrightarrow{PO_2}$  의 방향은 목적지 뷰의 뷰 각도에 대응한다. 투영된 좌표가  $\overline{u_2}$  라고 가정함으로써, 참조 뷰 내에서  $\overline{u_1}$  의 (상이한 컬러 성분들 내에서의) 픽셀 값들은 가상 뷰 내에서  $\overline{u_2}$  에 대한 픽셀 값들로 간주될 수 있다.

[0086] 때때로, 2 이상의 뷰는 참조 뷰들로 간주될 수 있다. 즉, 전술된  $\overline{u_1}$  으로부터  $\overline{u_2}$  로의 투영은 반드시 일대일 투영은 아니다. 그러나, 2 이상의 픽셀이 목적지 픽셀  $\overline{u_2}$  로 투영되는 경우에, 가시성 문제가 발생할 수도 있다. 다른 한편으로, 하나의 픽셀이 목적지 픽셀  $\overline{u_2}$  로 투영되는 경우에, 가상 뷰의 픽처 내에 홀 (hole) 이 나오거나 존재할 수도 있다. 이른바 가시성 문제는  $\overline{u_2}$  에 대한 픽셀 값들을 구성하기 위해 어떤 픽셀들을 이용할 것인지에 대하여 결정들을 요구할 수도 있다. 연속하는 영역의 픽처 내에 홀들이 존재하는 경우에, 이 현상은 오클루전 (occlusion) 이라 불린다. 그렇지 않으면, 화상 내에서 홀들이 저밀도로 분산되는 경우에, 홀들은 핀홀 (pinhole) 들이라 불릴 수도 있다. 오클루전은 하나의 참조 뷰를 서로 다른 방향으로 제공함으로써 해결될 수 있다. (예컨대, 핀홀들이 픽셀 값들을 충전하기 위한) 핀홀 필링은 일반적으로 이웃 픽셀들을 홀을 위한 후보들로서 취득한다. 핀홀 필링에 대한 기술들은 오클루전 문제들을 해결하기 위해 이용될 수 있다.

[0087] 2 이상의 픽셀이  $u_2$  의 픽셀 값들로 간주되는 경우에, 가중 평균 방법들이 채택될 수도 있다. 이러한 프로세스들은 통상적으로 뷰 합성에서 복원으로 지칭된다. 가시성, 오클루전, 핀홀 필링 및 복원은 뷰 합성에 기초한 3D 래핑의 구현에서 주요한 문제 및 장애물을 나타낸다. 카메라 모델은 이러한 문제들을 어드레싱하는 것을 도울 수도 있다.

[0088] 내부 (intrinsic) 파라미터 및 외부 (extrinsic) 파라미터를 포함하는 카메라 모델은 예컨대, 월드 좌표 (world coordinate) 시스템으로부터 카메라 평면으로의 변환 또는 그 반대의 변환을 설명하기 위해 이용될 수도 있다. 간단함을 위해, 본 개시물에서 설명되는 모든 언급되는 좌표 시스템들은 직교 좌표 시스템들이지만, 본 개시물의 기술들은 이 점에 관해서 반드시 제한되지 않는다.

[0089] 외부 파라미터들은 하기의 변환에 기초하여 월드 좌표에서 카메라 중심의 위치 및 카메라의 헤딩을 정의할 수 있다:

$$\begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = A \begin{pmatrix} x_w \\ y_w \\ z_w \\ 1 \end{pmatrix} \quad (1)$$

[0090]

[0091] 상기  $(x \ y \ z)^T$  는 3D 카메라 좌표 시스템에서의 좌표이고,  $(x_w \ y_w \ z_w)^T$  는 월드 좌표 시스템에서의 좌표이다. 행렬  $A$  는  $4 \times 4$  행렬을 포함할 수도 있고, 하기와 같이 예시될 수 있는 직교 정규

(orthonormal) 변환일 수도 있다:

$$\mathbf{A} = \begin{pmatrix} \mathbf{R} & \mathbf{T} \\ \mathbf{0} & 1 \end{pmatrix} \quad (2)$$

상기  $\mathbf{R}$  은  $3 \times 3$  회전 행렬이고,  $\mathbf{T}$  는 평행 이동 (translation) 이다. 이 경우에,  $\mathbf{T}$  는 카메라의 위치가 아니다.

3D 카메라 좌표 시스템에서,  $z$  축은 주광축 (principal optical axis) 이라 불릴 수 있고,  $x$  및  $y$  축들은 이미지 평면을 정의할 수 있다. 예를 들면, 도 6 에 도시된 것과 같이,  $\overrightarrow{O_1P}$  는 주광축을 정의할 수도 있다.  $u_1$  을 포함하는, 주광축에 수직인 평면은 이미지 평면을 정의할 수도 있다.

월드 좌표 시스템은 카메라의 3D 카메라 좌표 시스템과 동일한 것으로 정의될 수 있다. 이 경우에,  $\mathbf{A}=\mathbf{I}$  이다. 3D 카메라 좌표 시스템이 월드 좌표 시스템으로부터 평행 이동되면:

$$\mathbf{A} = \begin{pmatrix} \mathbf{I} & \mathbf{T} \\ \mathbf{0} & 1 \end{pmatrix}, \quad (3)$$

이고

$$\begin{pmatrix} x & y & z \end{pmatrix}^T = \begin{pmatrix} x_w & y_w & z_w \end{pmatrix}^T + \mathbf{T}^T \text{ 이다.}$$

내부 파라미터는 3D 카메라 좌표 시스템으로부터 2D 이미지 평면으로의 변환을 명시한다. 이러한 컨버전을 위한 모델은 도 5에 개념적으로 예시된 핀홀 카메라 모델로 지칭될 수도 있다. 이 경우에,  $\mathbf{O}$  는 3D 카메라 좌표 시스템의 원점이고, 카메라 평면 (또는 센서 평면) 의 중심을 정의할 수도 있다. 이러한 모델에서,  $\frac{u}{x} = \frac{v}{y} = \frac{-f}{z}$  이고, 상기  $f$  는 초점 길이를 지칭하고,  $(u,v)^T$  는 이미지 평면 내의 좌표들을 지칭한다.

핀홀 카메라 모델은 초점 길이  $f$  가 음의 값이면 곤란할 수도 있다. 이러한 문제를 처리하기 위해, 핀홀 카메라 모델은 또한 도 6 에 도시된 것과 같이 정면 (frontal) 핀홀 카메라 모델로 표시될 수도 있다. 정면

핀홀 카메라 모델에서, 관계식은  $\frac{u}{x} = \frac{v}{y} = \frac{f}{z}$  이 된다. 이러한 변환은 하기와 같이 표시될 수 있다:

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} fx \\ fy \\ z \end{pmatrix} = \mathbf{Q} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad (4)$$

상기  $(u,v)$  는 이미지 평면에서의 좌표들이고,  $\mathbf{Q}$  는 내부 파라미터들의 가장 간단한 표현이다. 하나의 파라미터가  $\bar{\mathbf{u}} = \begin{pmatrix} u & v & 1 \end{pmatrix}^T$  를 표시하면, 월드 좌표 시스템으로부터 이미지 평면으로의 전체 변환은 하기 식에 의해 제공된다:

$$\bar{\mathbf{u}} = \mathbf{Q} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = \mathbf{A} \begin{pmatrix} x_w \\ y_w \\ z_w \\ 1 \end{pmatrix} \quad (5)$$

일부 구현들에서, 내부 카메라 파라미터는 전술된 것보다 더 복잡할 수 있다.  $\mathbf{Q}$  로서 전술된 변형은 하기 식에 의해 제공될 수도 있다:

$$\mathbf{Q} = \begin{pmatrix} f_x & Skew & principal_x \\ 0 & f_y & principal_y \\ 0 & 0 & 1 \end{pmatrix} \quad (6)$$

이 경우에,  $Skew$  는 카메라의 스큐 (skew) 인자를 지칭하고,  $(principal_x, principal_y)^T$  는 이미지 평

면 내의 주점 (principal point) 의 좌표이다. 주점은 주광축이 이미지 평면을 교차하는 점이다.  $f_x$

및  $f_y$  의 값들은 x 및 y 축들에서 초점 길이 값들이다.

[0106] 또한, 일부 구현들에서, 외부 카메라 파라미터는 전술된 것보다 더 복잡할 수도 있다. 더 실현 가능한 경우에, 예컨대,  $R$  은 오직 xz 면에서의 회전을 정의하고, 하기와 같이 설명될 수 있다:

$$R = \begin{pmatrix} \cos \theta & 0 & -\sin \theta \\ 0 & 1 & 0 \\ \sin \theta & 0 & \cos \theta \end{pmatrix} \quad (7)$$

[0107]

[0108] 입체 비디오에서, 카메라들은 동일한 내부 파라미터들을 가질 수도 있다. 예컨대, 2 개의 카메라들 사이에 평행 이동만이 존재하고, 카메라들 중 하나는 월드 좌표 시스템에 따라 정렬되는 경우에, 이는 사실일 수도 있다.

이 경우에,  $R=I$ ,  $T=0$ ,  $u_1 = fx_w/z_w$ ,  $v_1 = fy_w/z_w$  이다. 제 2 카메라가 제 1 카메라와 평행하여 배치되면,  $R=I$ ,  $T=(d \ 0 \ 0)^T$  이다. 이 경우에, 하기의 식이 유도될 수도 있다:

$$u_2 = f(x_w + d)/z_w = u_1 + fd/z_w \quad (8).$$

[0109]

$fd/z_w$  는 또한 디스패리티 (disparity) 라 불린다. 이 경우에, 3D 랩핑은 단지 디스패리티의 계산을 요구할 수도 있지만, 전술된 문제점들은 여전히 존재할 수도 있다.

[0110]

[0111] 각각의 컬러 성분 내에서 각각의 픽셀의 값들이 8-비트로 양자화되고 표현되어 저장되기 때문에, 깊이 값들은 한정된 동적 범위로 제시되어야만 할 수도 있다. 깊이 값들은 예컨대, 8-비트 동적 범위 구현에서 독점적으로 0 과 255 사이일 수도 있다. 깊이 값들은 큰 범위로부터 변화할 수 있다. 그러나, 통상적으로 최단 깊이 및 최장 깊이는 각각 0 및 255 로 맵핑되며, 임의의 다른 깊이 값은 0 내지 255의 범위 밖의 값으로 맵핑될 것이다.

[0112]

일부 통상적인 깊이 값 양자화 방법들이 하기와 같이 열거된다.

$$\left(1 - \frac{v}{255}\right)(z_{far} - z_{near}) = z - z_{near} \quad (9)$$

[0113]

또는

$$\frac{1}{z} - \frac{1}{z_{far}} = \frac{v}{255} \left( \frac{1}{z_{near}} - \frac{1}{z_{far}} \right) \quad (10)$$

[0115]

[0116] 앞의 2 개 식들에서,  $v$  는  $[0, 255]$  로 양자화된 값이고,  $z$  는 깊이 값이며, 또한 1 바이트 내에 저장될 수도 있다. 값  $z$  은  $[0, 255]$  로 정규화될 수도 있다. 일반적으로, 카메라에 더 인접하는 픽셀은 더 큰 값을 정의하고, 더 큰 깊이를 가지는 픽셀은  $[0, 255]$  의 더 작은 값으로 컨버팅된다. 따라서, 깊이 값을  $[z_{near}, z_{far}]$  로부터  $[0, 255]$  로 선형 컨버팅하는 것이 바람직하며,  $z_{near}$  은 255 로 맵핑되고,  $z_{far}$  은 0 으로 맵핑된다. 이는 앞서 도시된 식 (9) 의 개념이다. 컨버전을 달성하는 또 다른 방식은 깊이의 역의 값을  $[1/z_{far}, 1/z_{near}]$  로부터  $[0, 255]$  로 선형 컨버팅하는 것이며,  $1/z_{near}$  는 255 로 맵핑되고,  $1/z_{far}$  는 0 으로 맵핑된다. 이러한 다른 방식은 앞서 도시된 식 (10) 의 개념이다.

[0117]

H.264/AVC 에 기초한 2D 비디오 통신 시스템들은 임의의 3D 지원을 고려하지 않고 광범위하게 전개되었다. 그러한 시스템에서 전달되는 2D 콘텐츠의 대부분에 대하여, 3D 비디오가 요구되는 경우에 몇가지 문제점들이 발생할 수도 있다. 특히, 비디오 콘텐츠는 통상적으로 다중 카메라 시스템에 의해 캡처되거나 심지어 3D 모델로부터 변환되는 3D 비디오 소스를 가지지 않을 수도 있다. 3D 비디오 소스로부터 존재하지 않는 비디오 콘텐츠에 대하여, 그러한 비디오 콘텐츠가 디코더에서의 일부 프로세싱 이후에 3D 디스플레이를 위해 사용될 수 있는지 여부에 대한 시그널링이 부족할 수도 있다.

[0118]

3D 디스플레이가 H.264/AVC 에서 코딩된 2D 콘텐츠에 대하여 가능하면, 기존의 뷰로부터 추가 뷰를 생성할 때, 카메라 파라미터들, 장면의 깊이 범위, 또는 다른 파라미터들과 같은 몇몇 사이드 정보는 유용할 수 있다. 그러나, 현재의 H.264/AVC 비트스트림에서 그러한 정보를 통신하기 위한 메커니즘이 존재해야 할 수도 있다.



이를 위해, 본 개시물의 기술들은 H.264/AVC 의 SEI 메시지들 또는 유사한 타입의 메시지들을 이용할 수도 있다.

- [0119] 또 다른 쟁점 또는 문제는 통상적으로, 추가의 뷰가 생성되는 경우에 전송된 뷰와 생성된 뷰에 대하여 2 개의 추정되는 카메라들의 수평적인 배치로 인해 생성된 뷰 내에 볼 수 없는 영역이 존재한다는 것이다. 이러한 바람직하지 않은 영역이 뚜렷한 아티팩트를 도입하는 경우에, 생성된 뷰 내에 이러한 영역을 제시하는 것은 바람직하지 않을 수도 있다. 이러한 문제를 처리하기 위해, 본 개시물의 기술들은 크롭 영역을 정의하는 능력을 제공하고, 이를 비트스트림 내에 시그널링한다.
- [0120] 일부 양태들에서, 본 개시물의 기술들은 비디오 프레임들의 2D 시퀀스를 포함하는 코딩된 비디오 스트림, 예컨대 H.264/AVC 비트스트림이 예컨대, 입체 디스플레이에 따라 3D 방식으로 컨버팅되고 디스플레이되도록 한다. 이 경우에, 디코더는 오리지널 디코딩된 비디오 시퀀스와 함께 3D 디스플레이를 가능하게 하는 또 다른 비디오 시퀀스 (예컨대, 제 2 뷰) 를 생성할 수도 있다.
- [0121] 본 개시물의 기술들에서, 2D - 3D 컨버전 정보가 비트스트림 내에 제공된다. 비트스트림 내의 인코딩된 비디오 정보는 일반적으로 추가의 뷰들 또는 깊이 맵들을 포함하지 않는, 장면의 2D 표현이다. 따라서, 코딩된 2D 콘텐츠의 대역폭은 2D - 3D 컨버전과 관련된 임의의 SEI 메시지들을 포함하지 않는, 대응하는 2D 시퀀스의 대역폭과 매우 유사하다. 일부 경우에, 2D 비디오 시퀀스는 인코딩 디바이스에 미리 저장되며, 인코딩 디바이스에서 캡처되거나 인코딩되지 않는다. 이러한 경우에, 2D - 3D 컨버전 정보는 디폴트 설정들에 기초하여 입력되고 정의되는 카메라 파라미터들을 포함할 수도 있다. 다른 경우들에서, 2D 비디오 콘텐츠는 캡처되지만 인코딩되지 않는다. 이 경우에, 인코더는 입력으로서 임의의 3D 콘텐츠 없이 2D 콘텐츠를 인코딩할 수도 있다. 그러나, 2D - 3D 컨버전을 인식하는 인코더는 3D 컨버전에 필수적인 정보를 생성할 수도 있고, 이러한 정보를 H.264/AVC 비트스트림 내에 포함시킬 수도 있다. 또 다른 경우들에서, 인코딩 디바이스는 2D 비디오 콘텐츠를 캡처하고 인코딩할 수도 있다. 이 경우에, 인코딩 디바이스는 (예컨대, 디코딩 프로세스 동안) 2D 비디오 비트스트림을 분석함으로써 2D - 3D 컨버전 정보를 합산할 수도 있다.
- [0122] 비디오 콘텐츠가 2D - 3D 컨버전을 지원하지 않고 H.264/AVC 디코더에 의해 디코딩되는 경우에, 2D 비디오가 복원되어 디스플레이될 수도 있다. 그러나, 디코더가 2D - 3D 컨버전 정보를 지원하고, 따라서 2D - 3D 컨버전 기능을 가지면, 디코더는 디코딩된 2D 비디오 시퀀스 및 연관된 2D - 3D 컨버전 정보에 기초하여 2 차 뷰와 연관된 프레임들을 생성할 수도 있다. 2 개의 뷰들 (디코딩된 뷰 및 생성된 뷰) 는 그 후에 3D 디스플레이 내에 도시될 수 있다.
- [0123] 본 개시물에 따라, 3D 컨버전 정보는 2D 시퀀스의 비디오 프레임들의 각각에 적용되어 3D 비디오 데이터를 생성할 수 있는 파라미터 세트를 포함한다. 언급되는 바와 같이, H.264/AVC 의 SEI 메시지들은 비트스트림 내의 3D 컨버전 정보를 통신하기 위한 하나의 메커니즘일 수도 있지만, 다른 메시지들 또는 메커니즘들이 특히 다른 표준들과 함께 이용될 수도 있다. 3D 컨버전 정보는 하기의 사항 중 전부 또는 일부를 포함할 수도 있다:
- [0124] 1 - 연관된 코딩된 비디오 시퀀스가 3D 로 컨버팅될 수 있다는 표시.
- [0125] 2 - 카메라들의 초점 길이 및/또는 가상의 뷰에 대응하는 추정된 카메라의 평행 이동과 같은 중요한 카메라 파라미터들.
- [0126] 3 - 현재 장면의 깊이 범위.
- [0127] 4 - 특별한 처리 또는 크롭핑 (cropping) 을 요구할 수도 있는, 디코딩된 비디오 (오리지널 뷰) 와 가상 뷰 내의 영역들.
- [0128] 하기의 표 1 은 다양한 SEI 메시지 선택스 엘리먼트들을 포함하는 SEI 메시지 구분의 형식으로 3D 컨버전 정보 (2D - 3D 컨버전 정보) 의 일 예를 제공한다.

표 1

2D_to_3D_conversion_info( payloadSize ) {	C	Descriptor
<b>camera_scene_para_present_flag</b>	5	u(1)
<b>left_view_original_flag</b>	5	u(1)
<b>dummy_region_flag</b>	5	u(1)
if ( camera_scene_para_present_flag ) {		
<b>focal_length</b>	5	ue(v)
<b>near_depth</b>	5	ue(v)
<b>far_depth</b>	5	ue(v)
<b>translate_x</b>	5	ue(v)
}		
if ( dummy_region_flag ) {		
dum_reg_width	5	ue(v)
}		
}		

변수들 u(1) 및 ue(v) 은 H.264/AVC 규격 내에서 정의될 수도 있으며, 상기 u(1) 는 비트스트림 내의 플래그를 코딩하는 카테고리이고, ue(v) 는 익스포넨셜 골롬 (Exponential Golomb) (가변 길이) 코딩을 이용하여 비트스트림 내의 부호 미표시 정수를 코딩하는 카테고리이다.

표 1 의 예시적인 3D 컨버전 정보는 하기에서 설명되는 것과 같은 시맨틱스 (semantics) 를 가질 수도 있다. 존재하는 경우에, 이러한 SEI 메시지는 다음의 2D - 3D 컨버전 SEI 메시지가 발생할 때까지 또는 코딩 비디오 시퀀스의 종료가 발생할 때까지 디코딩 순서로 현재의 액세스 유닛 및 모든 후속 액세스 유닛들을 포함하는 타겟 액세스 유닛들에 적용될 수도 있다. 제 1 의 2D - 3D 컨버전 SEI 메시지는, 존재하는 경우에, 코딩된 비디오 시퀀스의 제 1 액세스 유닛에서 나타날 수 있다. 이러한 SEI 메시지는 존재하는 경우에 디코딩된 비디오에 기초하여 또 다른 뷰가 생성될 수 있는 것을 나타낼 수도 있다. 디코딩된 비디오는 오리지널 뷰로 정의될 수 있고, 디코더에 의해 생성된 2차 뷰는 생성된 뷰로 표시될 수 있다. 2 개의 카메라들은 오리지널 뷰 및 생성된 뷰에 대하여 가정될 수 있다.

표 1 의 변수 camera\_scene\_para\_present\_flag 는 1 의 값으로 설정되는 경우에, 2 개의 카메라들의 초점 길이, 깊이 범위 값들 및 평행 이동이 SEI 메시지 내에 명시되는 것을 나타낼 수도 있다. camera\_scene\_para\_present\_flag 가 0 과 동일하면, 이는 focal\_length, near\_depth, far\_depth, 및 translate\_x 이 일부 디폴트 값들에 의해 유추될 수 있음을 나타낼 수도 있다. 일 예로서, focal\_length, near\_depth, far\_depth, 및 translate\_x 의 디폴트 값은 각각 1000, 2000, 4000 및 PicWidthInSamples<sub>L</sub>/16 으로 정의될 수도 있다. camera\_scene\_para\_present\_flag 가 1과 동일하면, 초점 길이, 깊이 범위 값들 및 평행 이동 값은 명시적으로 비트스트림 내에 포함되지만, camera\_scene\_para\_present\_flag 가 0 과 동일하면, 이 값들은 명시적으로 비트스트림 내에 포함되지 않는다.

표 1 의 변수 left\_view\_original\_flag 는 생성된 뷰가 오리지널 뷰의 좌측에 있는 것을 나타내도록 1 과 동일할 수도 있으며, 이는 생성된 뷰에 대한 추정된 카메라가 오리지널 뷰의 오리지널 카메라의 좌측에 있음을 의미한다. 0 과 동일한 변수 left\_view\_original\_flag 는 생성된 뷰가 오리지널 뷰의 우측에 있는 것을 나타낸다. 물론, 이러한 우측 및 좌측 표시들은 반전될 수 있다.

표 1 의 변수 dummy\_region\_flag 는 2 개의 뷰들에 대하여 더미 (dummy) 영역이 존재하며, 이 영역은 디스플레이 이전에 크롭 (crop) 되는 것을 나타내도록 1 과 동일할 수도 있다. 표 1 의 변수 focal\_length 는 카메라들이 초점 길이를 명시할 수도 있다. 일 예로서, focal\_length 의 값은 포괄적으로 1 내지  $2^{32} - 1$  의 범위 내에 있을 수도 있다. 표 1 의 변수 near\_depth 는 오리지널 뷰와 생성된 뷰 내의 픽셀들의 최소 깊이 값을 명시한다. 변수 near\_depth 는 포괄적으로 1 내지 far\_depth-1 의 범위 내에 있을 수도 있다. 표 1 의 변수 far\_depth 는 오리지널 뷰 및 생성된 뷰 내의 픽셀들의 최대 깊이 값을 명시한다. far\_depth 의 값은 포괄적으로 near\_depth+1 내지  $2^{32} - 1$  의 범위 내에 있을 수도 있다.

표 1 의 변수 translate\_x 는 오리지널 뷰와 생성된 뷰의 추정된 카메라들 사이의 거리를 명시한다. 변수 dum\_reg\_width 는 출력을 위해 이용되지 않고 오리지널 뷰와 생성된 뷰에 대하여 크롭되는 영역의 폭을 루마 샘플들의 단위로 명시한다. dum\_reg\_width 는 dummy\_region\_flag 가 1 과 동일한 경우에만 포함될 수도

있다.

[0137] left\_view\_original\_flag 가 1 과 동일한 경우에, 출력 영역은 수평 방향에서 포괄적으로 0 부터 PicWidthInSamples<sub>L</sub> - dum\_reg\_width-1 까지의 영역으로 크롭된다. left\_view\_original\_flag 가 0 과 동일한 경우에, 출력 영역은 포괄적으로 dum\_reg\_width 부터 PicWidthInSamples<sub>L</sub>-1 까지의 영역으로 크롭된다. dum\_reg\_width 의 값은 PicWidthInSamples<sub>L</sub>/16 로서 유추될 수도 있다.

[0138] 예컨대, SEI 메시지 내에서 3D 컨버전 정보는 추가의 뷰를 생성하기 위해 디코더에서 뷰 합성을 위해 적용할 알고리즘을 반드시 명시하는 것은 아니다. 그러나, 3D 래핑 기반 알고리즘에서, 전술된 식 (8) 은 오리지널 뷰 내의 하나의 픽셀을 가상의 뷰로 실질적으로 맵핑하는 3D 래핑을 위해 이용될 수 있다. 이 경우에,  $z_w$  는 오리지널 뷰 내의 픽셀의 깊이이고, 이는 [near\_depth, far\_depth] 의 간격 내에 있으며,  $f$  는 focal\_length 이고,  $d$  의 절대값은 translate\_x 이다.

[0139] 대안적으로, focal\_length, near\_depth, far\_depth, 및 translate\_x 는 IEC 60559 규격 또는 다른 통신 프로토콜에 따라 부호가 있는 이중 값들로 시그널링 될 수 있다. IEC 60559 에서, 값  $X$  는 3개의 부분들, 부호  $s$  (+/-),  $N$  (가수 부분),  $E$  (지수 부분) 으로 시그널링된다. 값  $X$  를 계산하는 한가지 가능한 구현은 하기의 의사 코드에서 설명된다.

[0140] 만약  $0 < E < 63$  이면,  $X = (-1)^s \cdot 2^{E-31} \cdot (1.M)$  이다.

[0141] 만약  $E = 0$  이고,  $M$  이 비-제로이면,  $X = (-1)^s \cdot 2^{-30} \cdot (0.M)$  이다.

[0142] 만약  $E = 0$  이고  $M = 0$  이면,  $X = (-1)^s \cdot 0$  이다.

[0143] 여기서  $X$  는 계산될 변수이고,  $s$ ,  $N$  및  $E$  는 계산될 각각의 변수에 대하여 연관된 부호, 지수 및 가수 선택스 엘리먼트들에 대응하며,  $0 \leq M < 1$  일 때  $M = \text{bin2float}(N)$  이다. 각각의 카메라 파라미터 변수와 대응하는 선택스 엘리먼트들 사이의 관계는 광범위한 구현들의 대상이 된다.

[0144] 또한, 전술된 파라미터들에 부가하여, 추가의 파라미터들이 이용될 수도 있다. 예를 들면, 수렴 이미지 평면의 2 개의 상이한 카메라들로의 추정 거리를 정량화하는 수렴-깊이 값은 파라미터 세트 내에 포함될 수도 있다. 표 1 의 파라미터들에 따라, 수렴-깊이 값은 무한한 것으로 가정될 수 있지만, 다른 경우들에서, 수렴-깊이 값은 특히 파라미터 세트 내에 포함될 수도 있다.

[0145] 하기의 의사-코드는 분수의 이진 표현을 대응하는 소수로 변환하는 bin2float( ) 의 일 실시예를 제공한다. 함수  $M = \text{bin2float}(N)$  의 의사-코드는 하기와 같이 제공될 수도 있다:

```
float M = 0;
float factor = 2^(-v); /* v is the length of the mantissa */
for ( i = 0; i < v; i++) {
    M = M + factor * (N>> i) & 0x01;
    factor = factor * 2;
}
```

[0146] 대안적으로, 더 많은 카메라 파라미터들이 시그널링될 수 있다. 예를 들어, 식 (6) 에 도시된 것과 같이,  $x$  및  $y$  축에 대한 상이한 초점 길이들  $f_x$  및  $f_y$ , 스큐 인자, 또는 주점을 가지는 내부 파라미터들 또한 시그널링될 수 있다. 추가로, 생성된 뷰에 대한 추정된 카메라는  $xz$  평면 내에서 회전을 가질 수도 있으며, 따라서 이러한 값, 식 (7) 에 도시된 것과 같은  $\theta$  또한 시그널링될 수 있다.

[0149] 도 7 은 본 개시물과 부합하는 비디오 인코더 (50) 를 포함하는 디바이스에 의해 수행될 수도 있는 인코딩 프로세스를 예시하는 흐름도이다. 도 7 에 도시된 것과 같이, 비디오 인코더 (50) 는 비디오 프레임들의 2D 시퀀스를 인코딩하고 (도면부호 101), 2D - 3D 컨버전 유닛 (36) 은 비디오 인코더를 통해 3D 컨버전 정보를 인코

딩하며, 상기 3D 컨버전 정보는 2D 시퀀스의 비디오 프레임들의 각각에 적용되어 3D 비디오 데이터를 생성할 수 있는 파라미터 세트를 포함한다 (도면부호 102). 비디오 인코더 (50) 를 포함하는 디바이스는 인코딩된 2D 시퀀스를 3D 컨버전 정보와 함께 통신할 수도 있다 (도면부호 103). 도 1 을 참조하여, 예를 들어, 비디오 인코더 (22) 는 비디오 인코더 (50) 에 대응할 수도 있다. 이 경우에, 소스 디바이스 (12) 는 코드 분할 다중 액세스 (CDMA) 와 같은 무선 프로토콜에 따라 정보를 변조하고 전송하는 모뎀 (23) 및 송신기 (24) 를 통해, 인코딩된 2D 시퀀스를 3D 컨버전 정보와 함께 통신할 수도 있다 (도면부호 103).

[0150] 2D 시퀀스를 인코딩할 때, 비디오 인코더 (50) 는 ITU H.264 비디오 인코딩 표준에 따라 2D 시퀀스를 인코딩할 수도 있다. 이 경우에, 3D 컨버전 정보를 인코딩하는 것은 3D 컨버전 정보를 전송된 것과 같은 ITU H.264 비디오 인코딩 표준에 의해 지원되는 하나 이상의 SEI 메시지들로 인코딩하는 것을 포함할 수도 있다. 2D 시퀀스는 제 1 의 2D 시퀀스 (예컨대, 오리지널 시퀀스) 로 지칭될 수도 있다. 파라미터 세트는 제 1 의 2D 시퀀스 (오리지널 시퀀스) 의 비디오 프레임들의 각각에 적용되어 비디오 프레임들의 제 2 의 2D 시퀀스 (예컨대, 2차 뷰) 를 생성할 수도 있고, 상기 제 1 및 제 2 의 2D 시퀀스들은 집합적으로 3D 입체 비디오 시퀀스를 정의한다.

[0151] 3D 컨버전 정보는 2D 시퀀스에 적용되어 3D 비디오 데이터를 생성하는 3D 컨버전 프로세스를 식별하는 정보를 포함할 수도 있다. 전송된 것과 같이, 3D 컨버전 정보는 2D 시퀀스를 캡처하는 것과 연관된 카메라 파라미터들 및 값들을 포함할 수도 있다. 예를 들면, 3D 컨버전 정보는 2D 시퀀스를 캡처한 카메라와 연관된 초점 길이를 나타내는 초점 길이 값, 3D 비디오 데이터에서의 최소 깊이를 명시하는 근거리-깊이 값, 3D 비디오 데이터에서의 최대 깊이를 명시하는 원거리-깊이 값, 및 3D 비디오 데이터와 연관된 2 개의 카메라들 사이의 추정된 거리를 정량화하는 평행 이동 값을 포함할 수도 있다.

[0152] 3D 컨버전 정보는 명시적인 3D 파라미터 세트가 3D 컨버전 정보 내에 포함되는지 여부 또는 디폴트 3D 파라미터 세트가 이용될 것인지 여부를 나타내는 플래그를 포함할 수도 있다. 추가로, 3D 컨버전 정보는 2D 시퀀스의 제 2 뷰가 2D 시퀀스의 좌측에 생성될 것인지 또는 우측에 생성될 것인지 여부를 나타내는 플래그를 포함할 수도 있다. 또한, 3D 컨버전 정보는 3D 비디오 데이터로부터 제거될 크롭 영역을 식별하는 플래그를 포함할 수도 있다.

[0153] 일 실시예에서, 제 1 플래그는 명시적인 3D 파라미터 세트가 3D 컨버전 정보 내에 포함되는지 여부 또는 디폴트 3D 파라미터 세트가 이용될 것인지 여부를 나타낼 수도 있고, 상기 명시적인 3D 파라미터 세트는 제 1 플래그가 설정되는 경우에 3D 컨버전 정보 내에 포함된다. 제 2 플래그는 2D 시퀀스의 제 2 뷰가 2D 시퀀스의 좌측에 생성될 것인지 또는 우측에 생성될 것인지 여부를 나타낼 수도 있고, 제 3 플래그는 3D 비디오 데이터로부터 제거될 크롭 영역을 식별할 수도 있으며, 상기 크롭 영역을 정의하는 정보는 제 3 플래그가 설정되는 경우에 3D 컨버전 정보 내에 포함된다.

[0154] 도 8 은 본 개시물과 부합하는 디코딩 프로세스를 예시하는 흐름도이다. 비디오 디코더 (70) 는 인코딩된 2D 시퀀스와 함께 3D 컨버전 정보를 수신하고 (도면부호 111), 2D 시퀀스를 디코딩한다 (도면부호 112). 예를 들어, 엔트로피 디코딩 유닛 (72) 은 2D 시퀀스의 잔여 비디오 블록들에 대한 양자화된 계수들을 생성하기 위해 비트스트림의 엔트로피 디코딩을 수행할 수도 있다. 엔트로피 코딩 유닛 (72) 은 비트스트림으로부터 선택스 엘리먼트들을 분석하여, 그러한 선택스 엘리먼트들을 비디오 디코더 (70) 의 다양한 유닛들로 포워딩할 수도 있다. 엔트로피 코딩 유닛 (72) 은 SEI 메시지들을 분석하여 그 메시지들 내의 임의의 선택스를 식별할 수도 있다. 모션 정보 (예컨대, 모션 벡터들) 및 다른 선택스 정보는 예컨대, 예측 유닛 (75) 으로 포워딩될 수도 있다. 예측 유닛 (75) 은 인코딩에서 이용된 예측 블록을 식별하기 위해 모션 정보 또는 다른 선택스 엘리먼트들을 이용한다. 인터-기반 디코딩을 위해, 모션 벡터는 모션 보상 프로세스에 적용되어 예측 데이터의 하나 이상의 리스트들로부터 예측 블록을 생성할 수도 있다. 인터-기반 디코딩을 위해, 선택스는 예측 블록이 디코딩되고 있는 비디오 블록과 동일한 프레임으로부터의 데이터에 기초하여 생성되는 방법을 정의할 수도 있는 인트라 모드를 포함할 수도 있다.

[0155] 역 양자화 유닛 (76) 은 엔트로피 디코딩 유닛 (72) 으로부터 데이터를 역 양자화하고, 역 변환 유닛은 역 변환을 수행하여 픽셀 영역 내에 잔여 블록을 생성한다. 합산기 (79) 는 잔여 블록을 예측 유닛 (75) 에 의해 생성된 예측 블록과 결합하여 오리지널 비디오 블록의 복원을 생성하며, 이는 메모리 (74) 내에 저장될 수도 있고 및/또는 디스플레이에 디코딩된 2D 비디오 데이터로서 출력될 수도 있다. 다수의 비디오 블록들은 비디오 프레임들을 복원하고, 궁극적으로는 메모리 (74) 내의 비디오 프레임들의 전체 2D 시퀀스를 복원하기 위해 상기와 같은 방식으로 디코딩될 수도 있다.



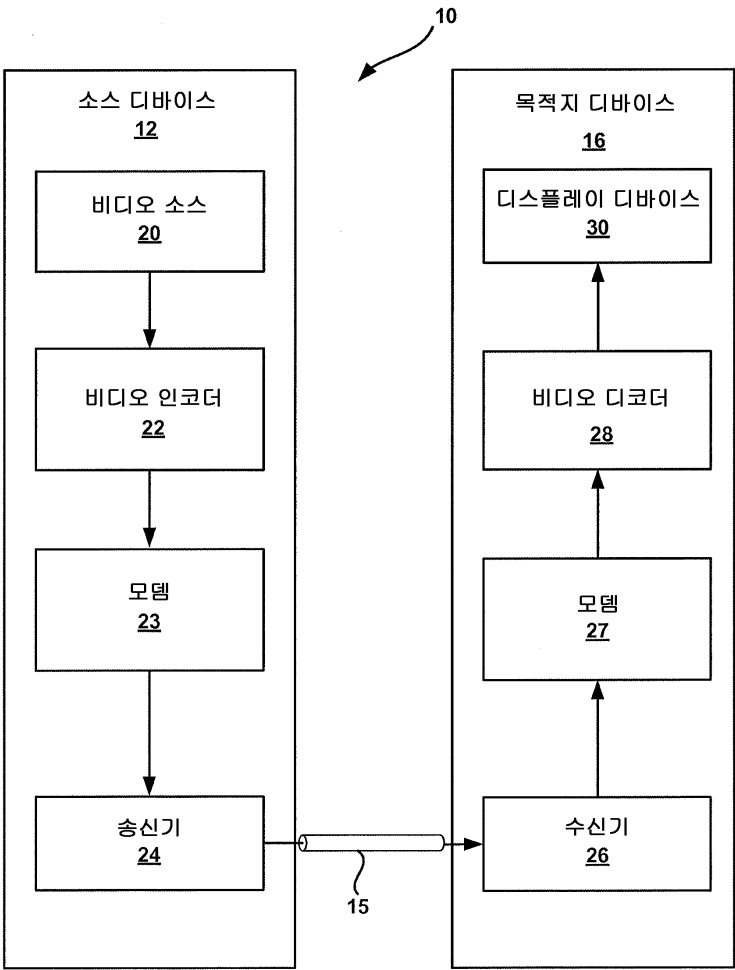
- [0156] 비디오 디코더 (70) 는 3D 비디오를 지원할지 여부를 결정할 수도 있다 (도면부호 113). 다시 말해서, 비디오 디코더 (70) 는 2D - 3D 컨버전 유닛 (79) 을 포함하는지 여부를 결정할 수도 있다. 이는 명시적으로 결정될 수도 있거나, 대안적으로 비디오 디코더 (70) 는 비디오 디코더 (70) 가 2D - 3D 컨버전 유닛 (79) 을 포함하지 않으면 2D 디코딩에 따라 고유하게 동작할 수도 있다. 만약 비디오 디코더 (70) 가 2D - 3D 컨버전 유닛 (79) 을 포함하면, 3D 비디오가 생성될 수도 있다.
- [0157] 따라서, 만약 비디오 디코더 (70) 가 3D 비디오를 지원하지 않으면 (도면부호 113 의 아니오 브랜치), 비디오 디코더는 디코딩된 2D 시퀀스를 2D 디스플레이로 출력할 수도 있다 (도면부호 114). 한편, 비디오 디코더 (70) 가 3D 비디오를 지원하면 (도면부호 113 의 예 브랜치), 2D - 3D 컨버전 유닛 (79) 은 디코딩된 2D 시퀀스에 3D 컨버전 정보를 적용하여 3D 비디오 시퀀스를 생성하며 (도면부호 115), 상기 시퀀스는 그 후에 비디오 디코더 (70) 로부터 3D 디스플레이로 출력될 수 있다 (도면부호 116). 이러한 방식으로, 3D 컨버전 정보와 통신되는 2D 시퀀스는 2D - 3D 컨버전을 지원하는 디코딩 디바이스들을 위한 3D 비디오를 지원할 수 있지만, 오직 2D 비디오만을 지원하는 레거시 (legacy) 디바이스들을 위해 종래의 2D 비디오를 지원할 수도 있다.
- [0158] 다른 실시예들에서, 3D 컨버전 정보를 3D 시퀀스와 함께 통신하는 것은 불필요할 수도 있다. 예를 들어, 인코더 또는 디코더는 3D 비디오 데이터를 생성하기 위해 본 명세서에서 설명된 3D 컨버전 정보를 저장되거나 수신된 2D 시퀀스에 적용할 수도 있다. 따라서, 본 개시물은 3D 컨버전 정보를 2D 시퀀스에 적용하여 3D 비디오 데이터를 생성하는 방법들, 컴퓨터 판독가능 저장 매체, 장치들 및 디바이스들을 고려하며, 상기 3D 컨버전 정보는 2D 시퀀스의 비디오 프레임의 각각에 적용되어 3D 비디오 데이터를 생성할 수 있는 파라미터 세트를 포함한다.
- [0159] 본 개시물의 기술들은 모바일 폰과 같은 무선 통신 디바이스 핸드셋, 집적 회로 (IC) 또는 IC들의 세트 (즉, 칩셋) 을 포함하는 폭넓게 다양한 디바이스들 또는 장치들로 구현될 수도 있다. 임의의 컴포넌트들, 모듈들, 또는 유닛들은 기능적 양태들을 강조하기 위해 제공되는 것으로 설명되었으며, 상이한 하드웨어 유닛들에 의한 실현을 반드시 요구하는 것은 아니다. 또한, 여기에 기술된 기술들은 하드웨어, 소프트웨어, 펌웨어, 또는 이들의 임의의 조합으로 구현될 수도 있다. 모듈들, 유닛들 또는 컴포넌트들로서 기술된 임의의 특징들은 통합된 로직 디바이스에서 함께 구현될 수도 있거나, 개별적이지만 상호 동작가능한 로직 디바이스들로서 별도로 구현될 수도 있다. 일부 경우, 다양한 특징들이 집적 회로 칩 또는 칩셋과 같은 집적 회로 디바이스로서 구현될 수도 있다.
- [0160] 소프트웨어로 구현되는 경우, 그 기술들은 프로세서에서 실행될 때에 상술된 방법들 중 하나 이상을 수행하는 명령들을 포함하는 컴퓨터 판독가능 매체에 의해 적어도 부분적으로 실행될 수도 있다. 컴퓨터 판독가능 매체는 물리적 구조물인 컴퓨터 판독가능 저장 매체를 포함할 수도 있고, 패키징 재료들을 포함할 수도 있는 컴퓨터 프로그램 제품의 일부를 형성할 수도 있다. 컴퓨터 판독가능 저장 매체는 동기식 동적 랜덤 액세스 메모리 (SDRAM : synchronous dynamic random access memory) 와 같은 랜덤 액세스 메모리 (RAM : random access memory), 판독 전용 메모리 (ROM : read-only memory), 비휘발성 랜덤 액세스 메모리 (NVRAM : non-volatile random access memory), 전기적 소거가능 프로그램가능 판독전용 메모리 (EEPROM : electrically erasable programmable read-only memory), 플래시 메모리 (FLASH memory), 자기 또는 광학 데이터 저장 매체 등을 포함할 수도 있다. 기술들은 추가로 또는 대안적으로, 명령들 또는 데이터 구조들의 형식으로 코드를 운반하거나 통신하고, 컴퓨터에 의해 액세스되고, 판독되고, 및/또는 실행될 수도 있는 컴퓨터 판독가능 통신 매체에 의해 적어도 부분적으로 실행될 수도 있다.
- [0161] 코드 또는 명령들은 하나 이상의 디지털 신호 프로세서 (DSP) 들, 범용 마이크로프로세서들, 주문형 집적 회로 (ASIC), 필드 프로그램가능 게이트 어레이 (FPGA) 들, 또는 다른 동등한 통합된 또는 개별적 로직 회로와 같은 하나 이상의 프로세서들에 의해 실행될 수도 있다. 따라서, 여기에 사용되는 "프로세서" 라는 용어는 여기에 기술된 기술들의 구현에 적합한 임의의 상기 구조 또는 임의의 다른 구조를 지칭할 수도 있다. 또한, 일부 양태들에서는, 여기에 기술된 기능성이 인코딩 및 디코딩을 위해 구성된 전용 소프트웨어 모듈들 또는 하드웨어 모듈들 내에 제공될 수도 있거나, 조합된 비디오 코덱 내에 포함될 수도 있다. 또한, 그 기술들은 하나 이상의 회로들 또는 로직 엘리먼트들에서 완전히 구현될 수 있다.
- [0162] 또한, 본 개시물은, 본 개시물에 기술된 기술들 중 하나 이상을 구현하기 위한 회로를 포함하는 다양한 집적 회로 디바이스들 중 임의의 집적 회로 디바이스를 고려한다. 이러한 회로가 소위 칩셋 내의 단일의 집적 회로 칩 또는 다수의 상호동작가능한 집적 회로 칩들로 제공될 수도 있다. 이러한 집적 회로 디바이스들은 그 일부가 포함될 수도 있는 다양한 애플리케이션들에서 이용될 수도 있고, 모바일 전화기 핸드셋들과 같은 무선 통

신 디바이스들에서 이용할 수도 있다.

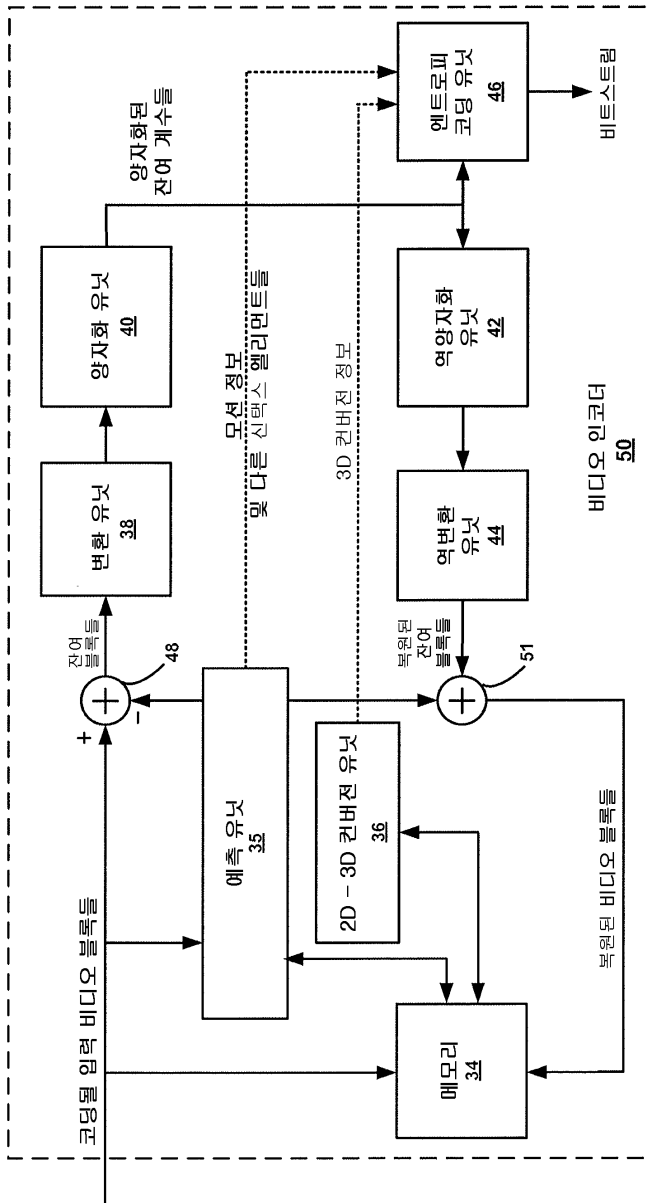
[0163] 기술들의 다양한 양태들이 기술되었다. 이들 양태들 및 다른 양태들은 다음의 청구항들의 범위 내에 있다.

도면

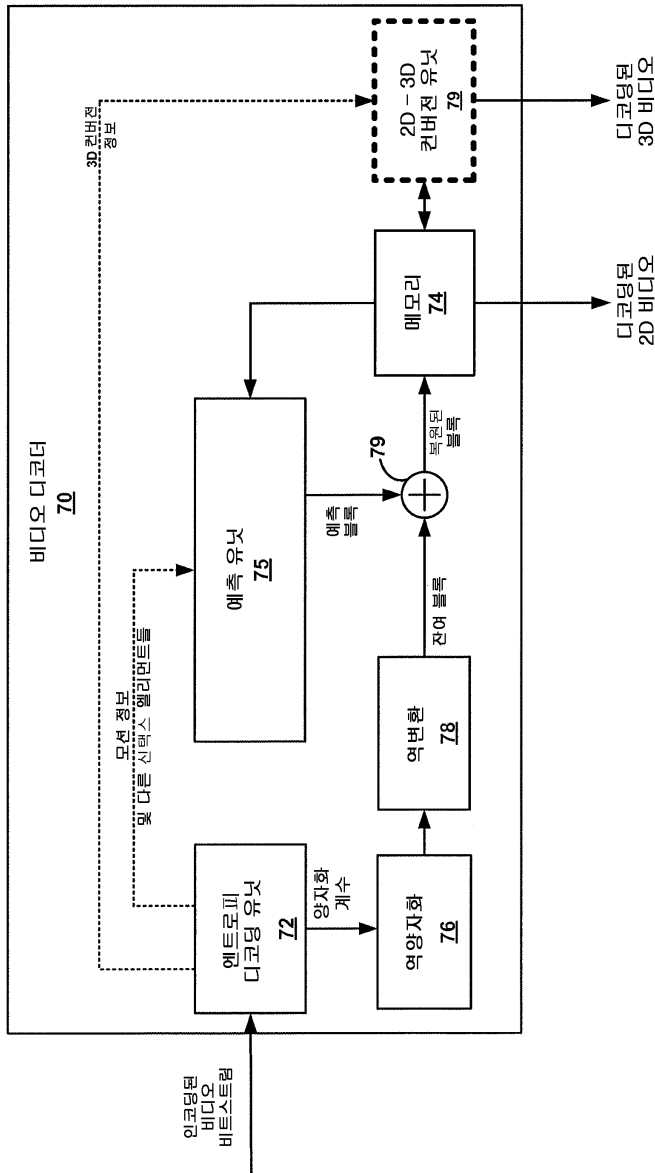
도면1



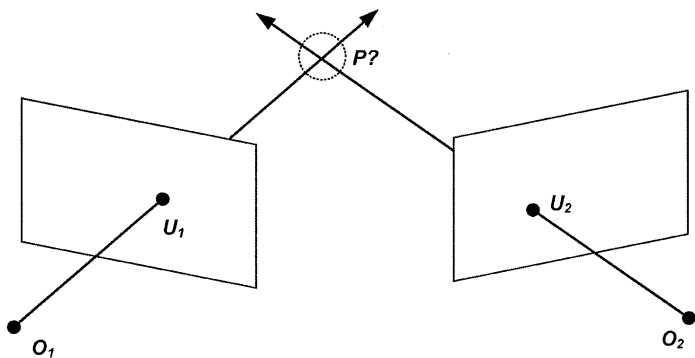
도면2



도면3

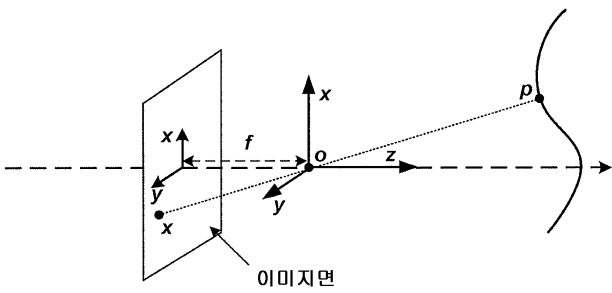


도면4

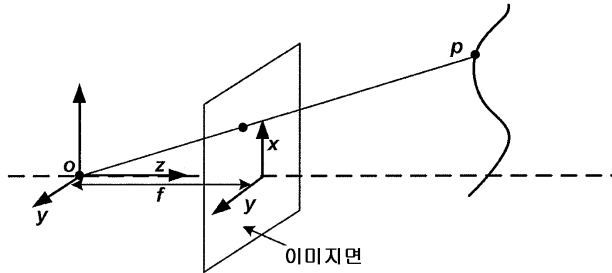




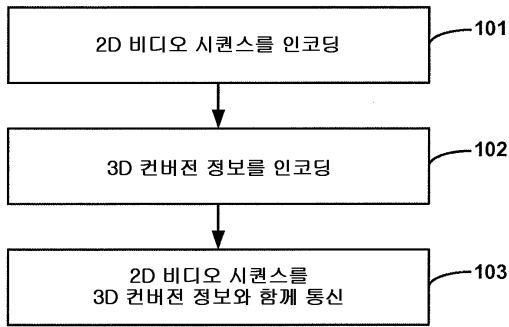
도면5



도면6



도면7



도면8

