



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2022년12월05일
(11) 등록번호 10-2474088
(24) 등록일자 2022년11월30일

(51) 국제특허분류(Int. Cl.)
G06T 19/00 (2011.01) G02B 27/01 (2006.01)
G06T 15/50 (2006.01) H04N 7/00 (2014.01)
(52) CPC특허분류
G06T 19/006 (2013.01)
G02B 27/017 (2013.01)
(21) 출원번호 10-2021-7037554(분할)
(22) 출원일자(국제) 2017년04월10일
심사청구일자 2021년11월17일
(85) 번역문제출일자 2021년11월17일
(65) 공개번호 10-2021-0144917
(43) 공개일자 2021년11월30일
(62) 원출원 특허 10-2018-7030032
원출원일자(국제) 2017년04월10일
심사청구일자 2020년03월26일
(86) 국제출원번호 PCT/EP2017/058571
(87) 국제공개번호 WO 2017/182315
국제공개일자 2017년10월26일
(30) 우선권주장
16305473.7 2016년04월22일
유럽특허청(EPO)(EP)
16306258.1 2016년09월29일
유럽특허청(EPO)(EP)
(56) 선행기술조사문헌
KR1020150116871 A*
JP2015231106 A*
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
인터디지탈 씨이 페이튼트 홀딩스, 에스에이에스
프랑스 75017 빠리 뒤 뒤 콜로넬 몰 3
(72) 발명자
띠에보, 실뱅
프랑스 35576 쉐송 셰비네 셰텍스 아브뉴 데 상
블랑 975 페르니콜로르 에르 에 데 프랑스
플뢰로, 줄리앙
프랑스 35576 쉐송 셰비네 셰텍스 아브뉴 데 상
블랑 975 페르니콜로르 에르 에 데 프랑스
제라르, 프랑스와
프랑스 35576 쉐송 셰비네 셰텍스 아브뉴 데 상
블랑 975 페르니콜로르 에르 에 데 프랑스
(74) 대리인
양영준, 이민호, 백만기

전체 청구항 수 : 총 20 항

심사관 : 옥윤철

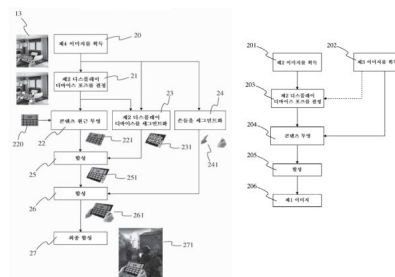
(54) 발명의 명칭 이미지를 합성하기 위한 방법 및 디바이스

(57) 요약

제1 이미지를 합성하고 및/또는 제1 디스플레이 디바이스에 송신하는 방법 및 디바이스로서, 방법은 장면을 표현하는 제2 이미지를 수신하는 단계 - 장면은 제3 이미지를 디스플레이하는 제2 디스플레이 디바이스를 포함함 - ; 제3 이미지를 수신하는 단계; 장면에 대한 제2 디스플레이 디바이스의 포즈를 표현하는 제1 정보를 획득하는 단

(뒷면에 계속)

대표도



계; 제1 정보에 따라 제3 이미지를 왜곡시키는 단계; 획득된 제1 정보를 사용하여 제2 이미지 및 왜곡된 제3 이미지를 결합함으로써 제1 이미지를 발생시키는 단계; 및 제1 이미지를 표현하는 데이터를 송신하는 단계를 포함한다.

(52) CPC특허분류

G06T 15/503 (2013.01)

H04N 7/00 (2013.01)

H04N 2201/3245 (2013.01)

명세서

청구범위

청구항 1

방법으로서,

실제 환경의 캡처된 이미지를 획득하는 단계 - 상기 실제 환경은 스크린을 갖는 디바이스를 포함하고, 상기 캡처된 이미지는 상기 스크린을 갖는 상기 디바이스를 포함함 -;

상기 캡처된 이미지에 기초하여 상기 스크린의 포즈를 결정하는 단계;

상기 캡처된 이미지 이외의 소스로부터, 가상 장면 내의 상기 스크린의 표현 상에 디스플레이될 2D 콘텐츠를 획득하는 단계;

투영된 2D 콘텐츠를 생성하기 위해 상기 2D 콘텐츠를 투영하는 단계 - 상기 투영된 2D 콘텐츠는 상기 스크린의 상기 포즈에 정렬됨 -; 및

가상 콘텐츠 아이템 및 상기 투영된 2D 콘텐츠의 조합으로서 상기 가상 장면을 생성하는 단계를 포함하는 방법.

청구항 2

제1항에 있어서, 사용자의 헤드 장착 디스플레이(HMD)에 상기 가상 장면을 디스플레이하는 단계를 추가로 포함하는 방법.

청구항 3

제2항에 있어서, 상기 캡처된 이미지는 상기 HMD의 적어도 하나의 카메라를 사용하여 획득되는 방법.

청구항 4

제1항에 있어서, 상기 캡처된 이미지 내의 상기 스크린은 적어도 하나의 마커를 디스플레이하고, 상기 스크린의 상기 포즈는 상기 캡처된 이미지 내의 상기 적어도 하나의 마커를 검출하는 것에 기초하여 결정되는 방법.

청구항 5

제1항에 있어서, 상기 2D 콘텐츠는 2D 이미지인 방법.

청구항 6

제1항에 있어서, 상기 2D 콘텐츠는 텍스트, 그래픽 및 그래픽 객체들 중 적어도 하나를 포함하는 방법.

청구항 7

제1항에 있어서, 상기 디바이스는 휴대용 디바이스인 방법.

청구항 8

제7항에 있어서, 상기 휴대용 디바이스는 그래픽 사용자 인터페이스(GUI)를 포함하고, 상기 방법은 상기 휴대용 디바이스 상의 상기 GUI의 그래픽 객체에 대한 누름 액션을 수신하는 단계를 추가로 포함하고, 상기 누름 액션은 상기 2D 콘텐츠의 하나 이상의 파라미터를 제어하는 방법.

청구항 9

제1항에 있어서, 상기 캡처된 이미지로부터 상기 디바이스의 이미지를 세그먼트화하는 단계를 추가로 포함하고, 상기 가상 장면은 상기 디바이스의 상기 세그먼트화된 이미지로부터 추가로 생성되는 방법.

청구항 10

제1항에 있어서, 상기 가상 장면은 상기 디바이스를 잡거나 상기 디바이스와 상호작용하는 사용자의 손들의 표

현으로부터 추가로 생성되는 방법.

청구항 11

장치로서,

프로세서 실행가능 프로그램 명령어들을 저장하는 메모리; 및

적어도 하나의 하드웨어 프로세서

를 포함하고,

상기 적어도 하나의 하드웨어 프로세서는 상기 프로그램 명령어들을 실행하여:

실제 환경의 캡처된 이미지를 획득하고 - 상기 실제 환경은 스크린을 갖는 디바이스를 포함하고, 상기 캡처된 이미지는 상기 스크린을 갖는 상기 디바이스를 포함함 -;

상기 캡처된 이미지에 기초하여 상기 스크린의 포즈를 결정하고;

상기 캡처된 이미지 이외의 소스로부터, 가상 장면 내의 상기 스크린의 표현 상에 디스플레이될 2D 콘텐츠를 획득하고;

투영된 2D 콘텐츠를 생성하기 위해 상기 2D 콘텐츠를 투영하고 - 상기 투영된 2D 콘텐츠는 상기 스크린의 상기 포즈에 정렬됨 -;

가상 콘텐츠 아이템과 상기 투영된 2D 콘텐츠의 조합으로서 상기 가상 장면을 생성하도록 구성되는 장치.

청구항 12

제11항에 있어서, 상기 적어도 하나의 하드웨어 프로세서는 사용자의 헤드 장착 디스플레이(HMD)에 상기 가상 장면을 디스플레이하기 위해 상기 프로그램 명령어들을 실행하도록 추가로 구성되는 장치.

청구항 13

제12항에 있어서, 상기 캡처된 이미지는 상기 HMD의 적어도 하나의 카메라를 사용하여 획득되는 장치.

청구항 14

제11항에 있어서, 상기 캡처된 이미지 내의 상기 스크린은 적어도 하나의 마커를 디스플레이하고, 상기 스크린의 상기 포즈는 상기 캡처된 이미지 내의 상기 적어도 하나의 마커를 검출하는 것에 기초하여 결정되는 장치.

청구항 15

제11항에 있어서, 상기 2D 콘텐츠는 2D 이미지인 장치.

청구항 16

제11항에 있어서, 상기 2D 콘텐츠는 텍스트, 그래픽, 및 그래픽 객체들 중 적어도 하나를 포함하는 장치.

청구항 17

제11항에 있어서, 상기 디바이스는 휴대용 디바이스인 장치.

청구항 18

제17항에 있어서, 상기 휴대용 디바이스는 그래픽 사용자 인터페이스(GUI)를 포함하고, 상기 적어도 하나의 하드웨어 프로세서는 상기 휴대용 디바이스 상의 상기 GUI의 그래픽 객체에 대한 누름 액션을 수신하도록 추가로 구성되고, 상기 누름 액션은 상기 2D 콘텐츠의 하나 이상의 파라미터를 제어하는 장치.

청구항 19

제11항에 있어서, 상기 적어도 하나의 하드웨어 프로세서는 상기 캡처된 이미지로부터 상기 디바이스의 이미지를 세그먼트화하도록 추가로 구성되고, 상기 가상 장면은 상기 디바이스의 상기 세그먼트화된 이미지로부터 추가로 생성되는 장치.

청구항 20

제11항에 있어서, 상기 가상 장면은 상기 디바이스를 잡거나 상기 디바이스와 상호작용하는 사용자의 손들의 표현으로부터 추가로 생성되는 장치.

발명의 설명

기술 분야

[0001] 본 개시내용은 예를 들어, 예컨대 헤드 장착 디스플레이(HMD: head-mounted display), 태블릿, 또는 스마트폰과 같은 모바일 디바이스 상에서 소비되는 몰입형 비디오 콘텐츠의 맥락에서의 이미지 처리의 영역에 관한 것이다.

배경 기술

[0002] 헤드 장착 디스플레이(HMD) 또는 태블릿과 같은 디스플레이 디바이스들은, 사용자가 몰입형 비디오 콘텐츠 내로 브라우징하는 것을 허용한다. 예를 들어, 헤드 장착 디스플레이들(HMD)은 한쪽 눈의 전방에(단안식 HMD) 또는 각각의 눈의 전방에(쌍안식 HMD) 작은 디스플레이 광학부를 갖는, 머리에 착용되거나 헬멧의 일부인 디스플레이 디바이스들이다. 그것들은 특히 사용자가 3D 가상 장면을 보고, 그/그녀의 머리를 회전하여 그 장면 내로 브라우징하는 것을 허용한다. 장면의 뷰는 사용자 헤드 포즈(user head pose)에 따라 업데이트된다. 그러한 디바이스들은 수십 년 전에 처음 제안되었지만, 최근에는 특히 OLED 및 LCD 스크린들의 진보로 인해 그것들의 가격 및 그것들의 성능이 극적으로 변화하여, 일부 HMD들이 현재 소비자 제품들로 되었을 정도에 이르렀다. 그러한 HMD의 가격은 합리적이고, 헤드 포즈 추정에 관하여 매우 낮은 레이턴시(예를 들어, 20ms)를 달성하면서도 110° 시야를 제공할 수 있다.

[0003] 다양한 디스패리티들이 전체적인 사용자 경험에 영향을 줄 수 있다. 예를 들어, 시각적 파라미터들이 정렬되지 않는 경우, 사용자는 메스꺼움을 느낄 수 있다. 그러므로, 몰입 경험을 생성하는 디바이스 및 능력은 사용자가 실세계와 상호 작용할 수 있는 능력을 향상시키거나, 아니면 그에 관련한 쟁점들을 발생시킬 수 있다.

발명의 내용

[0004] 본 명세서에서 "일 실시예", "실시예", "예시적인 실시예", "특정한 실시예"에 대한 언급들은 설명된 실시예가 특정한 특징, 구조, 또는 특성을 포함할 수 있음을 나타내지만, 모든 실시예가 반드시 그 특정한 특징, 구조, 또는 특성을 포함하지는 않을 수 있다. 더욱이, 그러한 문구들이 반드시 동일한 실시예를 지칭하는 것은 아니다. 또한, 특정한 특징, 구조, 또는 특성이 실시예와 관련하여 설명될 때, 그것은 명시적으로 설명되어 있는지 여부에 무관하게 다른 실시예들과 관련하여 그러한 특징, 구조, 또는 특성에 영향을 미치도록 본 기술분야의 통상의 기술자의 지식 범위 내에 있다는 것이 제시된다.

[0005] 본 개시내용은 제1 이미지를 제1 디스플레이 디바이스에 송신하는 방법에 관한 것으로서, 방법은:

[0006] 장면을 표현하는 제2 이미지를 수신하는 단계 - 장면은 제3 이미지를 디스플레이하는 제2 디스플레이 디바이스를 포함함 - ;

[0007] 제3 이미지를 수신하는 단계;

[0008] 장면에 대한 제2 디스플레이 디바이스의 포즈를 표현하는 제1 정보를 획득하는 단계;

[0009] 제1 정보에 따라 제3 이미지를 왜곡시키는 단계;

[0010] 획득된 제1 정보를 사용하여 제2 이미지와 왜곡된 제3 이미지를 결합함으로써 제1 이미지를 발생시키는 단계; 및

[0011] 제1 이미지를 표현하는 데이터를 송신하는 단계

[0012] 를 포함한다.

[0013] 특징에 따르면, 왜곡된 제3 이미지는 제2 디스플레이 디바이스의 스크린의 표현(representation)을 포함하는 제2 이미지의 부분에 오버레이된다.

[0014] 구체적인 특징에 따르면, 제2 이미지는 제1 디스플레이 디바이스에 연관된 카메라로부터 수신된다.

- [0015] 다른 특징에 따르면, 제3 이미지는 제2 디스플레이 디바이스 상에 디스플레이하기 위해 제3 이미지를 제2 디스플레이 디바이스에 송신하는 소스 디바이스로부터 수신된다.
- [0016] 특정한 특징에 따르면, 제1 디스플레이 디바이스는 헤드 장착 디스플레이이고, 제2 디스플레이 디바이스는 디스플레이 스크린을 포함하는 모바일 디바이스이다.
- [0017] 다른 특징에 따르면, 제1 정보는 수신된 제3 이미지를, 제3 이미지를 디스플레이하는 제2 디스플레이 디바이스를 포함하는 제2 이미지의 콘텐츠와 비교함으로써 획득된다.
- [0018] 구체적인 특징에 따르면, 획득하는 단계는 제2 디스플레이 디바이스로부터 제1 정보를 수신하는 단계를 포함한다.
- [0019] 다른 특징에 따르면, 획득하는 단계는:
- [0020] 제2 디스플레이 디바이스의 제4 이미지를 취득하는 단계 - 적어도 하나의 기준 마커(fiducial marker)가 제2 디스플레이 디바이스 상에 디스플레이됨 - ;
- [0021] 제4 이미지 상에서 적어도 하나의 기준 마커를 추적하는 단계;
- [0022] 추적된 적어도 하나의 기준 마커에 기초하여 제1 정보를 결정하는 단계
- [0023] 를 포함한다.
- [0024] 특정한 특징에 따르면, 제1 디스플레이 디바이스는 제4 이미지를 취득하도록 구성된 카메라를 포함한다.
- [0025] 구체적인 특징에 따르면, 제4 이미지는 제1 디스플레이 디바이스 외부의 카메라로 취득되고, 방법은 제1 디스플레이 디바이스의 위치를 표현하는 제2 정보를 획득하는 단계를 더 포함하고, 제1 이미지는 제2 정보를 사용하여 합성된다.
- [0026] 특정한 특징에 따르면, 제3 이미지는:
- [0027] 실제 환경을 표현하는 비디오 시퀀스의 이미지;
- [0028] 가상 환경을 표현하는 비디오 시퀀스의 이미지;
- [0029] 실제 및 가상 객체들을 포함하는 장면을 표현하는 비디오 시퀀스의 이미지;
- [0030] 그래픽 사용자 인터페이스의 적어도 하나의 그래픽 객체를 포함하는 이미지; 및
- [0031] 적어도 하나의 텍스트 요소를 포함하는 이미지
- [0032] 를 포함하는 이미지들의 세트에 속한다.
- [0033] 구체적인 특징에 따르면, 합성하는 단계는 제2 이미지를 결합하고, 상기 제2 디스플레이 디바이스를 들고 있는 손들을 표현하는 제5 이미지로 제3 이미지를 왜곡시키는 단계를 더 포함하며, 상기 왜곡된 제3 이미지는 상기 제2 디스플레이 디바이스에 오버레이된다.
- [0034] 본 개시내용은 또한 제1 이미지를 제1 디스플레이 디바이스에 송신하도록 구성되는 디바이스에 관한 것으로, 디바이스는:
- [0035] 장면을 표현하는 제2 이미지를 수신하기 위한 수신 수단 - 장면은 제3 이미지를 디스플레이하는 제2 디스플레이 디바이스를 포함함 - ;
- [0036] 제3 이미지를 수신하기 위한 수단;
- [0037] 장면에 대한 제2 디스플레이 디바이스의 포즈를 표현하는 제1 정보를 획득하기 위한 수단;
- [0038] 제1 정보에 따라 제3 이미지를 왜곡시키기 위한 수단;
- [0039] 획득된 제1 정보를 이용하여 제2 이미지와 왜곡된 제3 이미지를 결합함으로써 제1 이미지를 발생시키기 위한 수단; 및
- [0040] 제1 이미지를 표현하는 데이터를 송신하기 위한 수단
- [0041] 을 포함한다.

- [0042] 본 개시내용은 또한 제1 이미지를 제1 디스플레이 디바이스에 송신하도록 구성된 디바이스를 위한 디바이스에 관한 것으로, 디바이스는:
- [0043] 장면을 표현하는 제2 이미지를 수신하고 - 장면은 제3 이미지를 디스플레이하기 위한 제2 디스플레이 디바이스를 포함함 -, 제3 이미지를 수신하도록 구성된 수신기;
- [0044] 프로세서 - 프로세서는:
- 장면에 대한 제2 디스플레이 디바이스의 포즈를 표현하는 제1 정보를 획득하고;
- [0046] 제1 정보에 따라 제3 이미지를 왜곡시키고;
- [0047] 제1 정보를 이용하여 제2 이미지와 왜곡된 제3 이미지를 결합함으로써 제1 이미지를 발생시키도록 구성됨 - ; 및
- [0049] 제1 이미지를 표현하는 데이터를 송신하도록 구성되는 송신기
- [0050] 를 포함한다.
- [0051] 디바이스는 위에서 언급된 방법의 단계들을 수행하도록 구성된다.
- [0052] 본 개시내용은 또한 프로그램이 컴퓨터 상에서 실행될 때, 적어도 하나의 프로세서에 의해, 제1 이미지를 송신하는 위에서 언급된 방법을 실행하기 위한 프로그램 코드의 명령어들을 포함하는 컴퓨터 프로그램 제품에 관한 것이다.
- [0053] 본 개시내용은 또한 프로세서로 하여금 적어도, 제1 이미지를 송신하는 위에서 언급된 방법을 수행하게 하기 위한 명령어들이 저장되어 있는 (비-일시적) 프로세서 판독가능한 매체에 관한 것이다.

도면의 간단한 설명

- [0054] 이하의 설명을 읽으면, 본 개시내용이 더 잘 이해될 것이고, 다른 구체적인 특징들 및 이점들이 드러날 것이고, 이하의 설명은 첨부된 도면들을 참조한다.
- 도 1a 및 도 1b는 본 원리들의 예에 따라, 도 1c 및 도 1d에 각각 도시된 제2 이미지 및 제3 이미지로부터 획득된 제1 이미지의 예를 각각 도시한다.
- 도 2a는 본 원리들의 예시적인 실시예에 따라, 도 1a 또는 도 1b의 제1 이미지를 획득하는 제1 프로세스를 도시한다.
- 도 2b는 본 원리들의 예시적인 실시예에 따라, 도 6의 제1 이미지를 획득하는 제2 프로세스를 도시한다.
- 도 3은 본 원리들의 예시적인 실시예에 따라, 도 1a 또는 도 1b의 제1 이미지를 디스플레이 디바이스에 송신하는 방법을 도시한다.
- 도 4는 본 원리들의 예에 따라, 도 1a 및 도 1b의 하나 이상의 제1 이미지를 디스플레이하도록 구성된 제1 디스플레이 디바이스(4)의 구조를 개략적으로 도시한다.
- 도 5는 본 원리들의 예에 따라, 도 3 및/또는 도 7의 제1 이미지를 송신하는 방법을 구현하도록 적응된 장치의 구조를 개략적으로 도시한다.
- 도 6은 본 원리들의 예에 따라, 제2 이미지 및 제3 이미지로부터 획득된 제1 이미지의 다른 예를 도시한다.

발명을 실시하기 위한 구체적인 내용

- [0055] 이제, 발명의 주제는 도면들을 참조하여 설명되고, 여기서 유사한 참조번호들은 전체에서 유사한 구성요소들을 지칭하기 위해 사용된다. 이하의 설명에서, 설명의 목적으로, 발명의 주제에 대한 완전한 이해를 제공하기 위해 수많은 구체적인 상세들이 제시된다. 그러나, 발명의 주제의 실시예들은 이러한 구체적인 상세들 없이도 실시될 수 있음이 명백할 수 있다.
- [0056] 본 원리들은 제1 이미지를 제1 디스플레이 디바이스(예를 들어, HMD 또는 태블릿)에 송신하는 방법, 및 그 방법을 구현하도록 구성되는 디바이스의 특정한 실시예를 참조하여 설명될 것이다. 송신될 제1 이미지는 제2 이미

지(예를 들어, 사용자가 몰입되어 있는 비디오 콘텐츠로부터 추출됨)와 제3 이미지(예를 들어, 몰입형 콘텐츠를 보고 있는 사용자의 실제 환경을 표현하는 이미지, 또는 사용자가 자신의 손에 들고 있는 사용자의 스마트폰 또는 태블릿 상에서 사용자에게 의해 수신된 메시지의 이미지)를 결합함으로써 합성된다. 제3 이미지는, 예를 들어 사용자가 현재 들고 있는 제2 디스플레이 디바이스의 이미지에 오버레이된다. 사용자의 손들의 이미지는 제3 이미지를 디스플레이하는 제2 디스플레이 디바이스의 이미지 상의 제1 이미지에 임의로(optionally) 추가될 수 있다.

[0057] 제1 이미지를 형성하기 위해 제2 이미지에 시각적 콘텐츠(즉, 제3 이미지)를 추가하면, 예를 들어 몰입형 콘텐츠(즉, 제2 이미지)에 몰입해 있는 사용자는 자신의 실제 환경과 상호작용할 수 있게 된다. 예를 들어, 몰입형 콘텐츠에 몰입한 사용자(예를 들어, 몰입형 콘텐츠를 디스플레이하는 HMD를 착용한 때)가 자신의 스마트폰에서 통화 또는 SMS를 수신할 때, 그 사용자는 몰입형 콘텐츠에 여전히 몰입해 있으면서, 수신된 통화 또는 SMS에 대한 직접적인 뷰를 갖는 능력을 갖는다. 사용자가 자신의 손에 들고 있는 사용자의 스마트폰을 몰입형 콘텐츠 내에서 스마트폰의 이미지 상에 중첩된 사용자의 손들의 이미지와 함께 디스플레이하면, 사용자는 스마트폰을 자신의 실제 환경에서 보고 있는 것처럼 제어할 수 있게 된다. 통화 또는 SMS의 이미지는 예를 들어 HMD의 완전한 디스플레이 해상도로부터 이익을 얻도록 (예를 들어, 제어 장치를 통해) 스마트폰으로부터 수신된다.

[0058] 다른 실시예에 따르면, 제2 이미지는 제3 이미지, 예를 들어 비디오 콘텐츠를 디스플레이하는 제2 디스플레이 디바이스를 포함하는 장면의 이미지에 대응할 수 있다. 제2 이미지는, 제1 디스플레이 디바이스에 연관된 이미지 취득 디바이스에 내장된 이미지 취득 디바이스(예를 들어, 카메라, 웹캠), 예를 들어 제1 디스플레이 디바이스에 내장된 이미지 취득 디바이스(예를 들어, 태블릿 또는 스마트폰의 후방 카메라, HMD를 착용하고 있는 사용자가 그 안으로 이동하는 장면의 이미지들을 취득하기 위해 HMD의 정면에 통합된 카메라)로 취득된다. 제1 디스플레이 디바이스 상에 디스플레이되는 장면의 이미지의 최적 해상도로부터 이익을 얻기 위해, 제3 이미지는 제3 이미지를 제2 디스플레이 디바이스에 송신하는 소스 디바이스(예를 들어, 셋톱 박스, 제2 디스플레이 디바이스에 연관된 디코딩 디바이스)로부터 이미지 처리 디바이스에 의해 직접 수신될 수 있다. 이미지 처리 디바이스는 수신된 제3 이미지(또는 동일 이미지의 처리된 버전)와 제2 이미지를 결합함으로써 제1 이미지를 합성할 수 있다.

[0059] 도 1a 및 도 1b는 본 원리들의 특정한 비-제한적인 실시예에 따라, 제1 디스플레이 디바이스, 예를 들어 HMD 상에 디스플레이될 제1 이미지(10, 11)의 두 가지 예를 도시한다. 제1 이미지(10, 11)는 예를 들어 정지 영상, 또는 제1 이미지들의 시퀀스로 구성된 제1 비디오 시퀀스의 이미지이다. 제1 이미지(10, 11)는 예를 들어 2개 또는 3개의 상이한 소스로부터 획득된다. 제1 이미지(10, 11)는 도 1c에 도시된 제2 이미지(12)를, 예를 들어 도 1d에 도시된 제4 이미지(13)의 일부분, 예를 들어 제2 디스플레이 디바이스(130)를 표현하는 제4 이미지(13)의 부분과 결합함으로써 획득될 수 있고, 여기서 제2 디스플레이 디바이스는 예를 들어 인물, 예를 들어 제1 디스플레이 디바이스를 착용하거나 들고 있는 인물이 손에 들고 있는 태블릿 또는 스마트폰과 같은 모바일 디바이스이다.

[0060] 도 1a는 제2 이미지(12)를, 제2 디스플레이 디바이스(130)를 포함하는 이미지(13)의 부분과 결합함으로써 획득되는 제1 이미지(10)의 제1 예를 도시하며, 제3 이미지(100)는 제4 이미지(13)로부터 추출된 제2 디스플레이 디바이스(130)의 이미지에 오버레이된다. 이러한 비-제한적인 예에서, 제3 이미지(100)는 제4 이미지(13)에 도시된 제2 디스플레이 디바이스(130)의 카메라에 의해 취득된 실제 환경의 이미지에 대응한다. 도 1d의 예에서 하나 이상의 기준 마커가 제2 디스플레이 디바이스(130) 상에 디스플레이되고, 도 2와 관련하여 더 상세하게 설명되는 바와 같이, 기준 마커들은 제2 디스플레이 디바이스(130)를 위치 지정 및/또는 세그먼트화하기 위해 사용된다. 기준 마커는, 예를 들어, 정사각형들 또는 직사각형들과 같은 기하학적으로 단순한 형상들로 구성된 2차원 흑백 패턴을 포함한다. 제4 이미지(13)는 예를 들어 제1 디스플레이 디바이스를 착용한 사용자가 서 있는 실제 환경의 이미지에 대응하며, 사용자는 제2 디스플레이 디바이스(130)를 들고 있다. 제4 이미지(13)는 제1 디스플레이 디바이스를 착용한 사용자의 실제 환경의 이미지들을 취득하는 제1 디스플레이 디바이스의 카메라로 취득될 수 있고, 취득의 주된 방향은 예를 들어 제1 디스플레이 디바이스를 착용한 사용자의 주된 시각화 방향(main visualization direction)에 대응하며, 취득 방향은 제1 디스플레이 디바이스를 착용한 사용자의 머리의 이동에 따라 달라진다. 다른 예에 따르면, 제4 이미지(13)는 제1 디스플레이 디바이스 외부의 카메라, 예를 들어 비디오 감시 시스템의 카메라 또는 웹캠으로 취득될 수 있다. 제2 이미지(12)는 예를 들어 컴퓨터 생성 이미지(CGI), 즉 가상 환경을 표현하는 이미지, 실세계 뷰에 연관된 이미지 상에 중첩된 CGI(증강 현실 또는 혼합 현실이라고 지칭됨), 또는 실세계 뷰에 연관된 이미지에 대응한다. 제3 이미지의 오버레이에 대한 서포트(support)로서 제2 디스플레이 서포트의 이미지를 사용하는 것에 의한, 제2 이미지(12) 내의 사용자의 실제 환

경의 이미지들의 인크루스트(encrusting)는, 사용자가 제2 이미지의 (실제 또는 가상) 환경에 몰입되어 머무르면서 자신의 실제 환경의 뷰를 갖는 것을 가능하게 한다. 제3 이미지가 제2 디스플레이 디바이스의 카메라에 의해 제공되는 경우, 사용자는 (예를 들어, 인물이 사용자의 환경에 들어가서 사용자에게 말을 걸기 시작할 때) 제2 디스플레이 디바이스를 자신이 보고 싶은 실제 환경의 영역으로 향하게 함으로써, 자신이 보고 싶은 것을 선택할 수 있다. 제2 디스플레이 디바이스의 카메라에 의해 취득되는 제3 이미지는 왜곡된 제3 이미지를 제1 디스플레이 디바이스에 송신하는 제어 디바이스[예를 들어, 도 5와 관련하여 설명되는 장치(5)]에 송신된다. 왜곡된 제3 이미지는 제2 디스플레이 디바이스의 카메라의 해상도, 및 제1 디스플레이 디바이스의 스크린(들)의 해상도로부터 이익을 얻는다. 변형예에 따르면, 제3 이미지(100)는 제3 이미지에 대한 서포트로서 제2 디스플레이 디바이스(130)의 이미지를 사용하지 않고서 제2 이미지(12)에 직접 오버레이된다.

[0061] 도 1b는 제2 이미지(12)를, 제2 디스플레이 디바이스(130)를 포함하는 이미지(13)의 부분과 결합함으로써 획득되는 제1 이미지(11)의 제2 예를 도시하며, 제3 이미지(110)는 제4 이미지(13)로부터 추출된 제2 디스플레이 디바이스(130)의 이미지에 오버레이된다. 이러한 비-제한적인 예에서, 제3 이미지(110)는 하나 이상의 그래픽 객체(예를 들어, 메뉴, 버튼)를 포함하는 그래픽 사용자 인터페이스(GUI)의 이미지에 대응한다. 제1 이미지(11)에 몰입한 사용자는 제2 이미지(12)의 환경에 몰입해 있는 느낌을 여전히 가지면서도, 자신이 현재 자신의 손에 들고 있는 제2 디스플레이 디바이스 상에 디스플레이된 GUI를 물리적으로 보고 있는 것처럼 임의의 그래픽 객체를 누르는 능력을 갖는다. GUI의 그래픽 객체에 대한 임의의 작용은 제어 디바이스[예를 들어, 도 5와 관련하여 설명된 장치(5)]에 송신되고, 제어 디바이스는 눌러진 그래픽 객체에 연관된 제어 파라미터를 의도된 장치에, 예를 들어 GUI가 제1 이미지에 연관된 하나 이상의 파라미터(예를 들어, 밝기, 제1 이미지에 연관된 오디오 트랙의 강도, 대비, 재생/일시정지)를 제어하는 것을 목표로 할 때는 제1 디스플레이 디바이스에, 또는 GUI가 사용자의 실제 환경과 상호작용하거나 그것을 제어하는 것(예를 들어, 조명 요소들, 셔터들, 오디오 플레이어들을 제어하는 것)을 목표로 할 때는 실제 환경의 다른 구성요소들에 재지향시킬 수 있다. 변형예에 따르면, 제3 이미지(110)는 제3 이미지를 위한 서포트로서 제2 디스플레이 디바이스(130)의 이미지를 이용하지 않고서 제2 이미지(12)에 직접 오버레이된다.

[0062] 제2 이미지(12)는 예를 들어 제2 이미지들의 제2 비디오 시퀀스로부터 획득된 이미지에 대응한다. 제2 비디오 시퀀스는 예를 들어 몰입형 비디오 콘텐츠에 대응한다. 몰입형 비디오 콘텐츠는 그것을 보고 있는 사용자가 비디오 내로 브라우징할 수 있게 하는 비디오에 대응하며, 즉 결정된 시간에 몰입형 콘텐츠의 일부분만이 제1 디스플레이 디바이스의 스크린(들) 상에 디스플레이되며, 그 디스플레이된 부분은 제1 디스플레이 디바이스를 착용한 사용자의 뷰잉 방향에 의존한다. 몰입형 비디오 콘텐츠는 몰입형 콘텐츠를 시각화하기 위해 사용되는 제1 디스플레이 디바이스의 시야보다 큰 시야를 각각 갖는 하나 이상의 이미지의 시퀀스에 대응한다. 몰입형 비디오 오는 이미지 또는 이미지들의 시퀀스를 포함하며, 각각의 이미지는 예를 들어 컴퓨터 생성 이미지(CGI), 실제 세계 뷰 상에 중첩된 CGI(증강 현실 또는 혼합 현실이라고 지칭됨), 또는 예를 들어 실제 세계의 180°, 220°, 또는 360° 이미지[4 π 스테라디안 비디오(4 π steradians video)라고도 지칭됨]를 획득하기 위해 예를 들어 다수의 카메라로 취득된 실제 세계 뷰에 대응한다. 이러한 비디오 콘텐츠는 그것을 보고 있는 사용자가 비디오 콘텐츠에 몰입감을 느끼는 것을 가능하게 하므로 몰입형 비디오 콘텐츠라고 지칭되며, 사용자는 비디오 콘텐츠 내로 브라우징할 수 있다. 몰입형 비디오 콘텐츠 내로의 브라우징은 HMD와 같은 디스플레이 디바이스에 의해 가능해진다.

[0063] 도 6은 본 원리들의 추가의 특정한 비-제한적인 예에 따라, 제1 디스플레이 디바이스(60), 예를 들어 태블릿 상에 디스플레이된 제1 이미지(600)의 다른 예를 도시한다. 제1 이미지는 정지 이미지, 또는 제1 이미지들의 시퀀스를 포함하는 제1 비디오 시퀀스의 이미지일 수 있다. 제1 이미지(600)는 예를 들어 상이한 소스들로부터 획득된 복수의 이미지, 예를 들어 2개, 3개 또는 그 이상의 이미지를 결합함으로써 획득된다. 제1 이미지(600)는 예를 들어 제2 이미지와 제3 이미지의 결합에 대응한다. 제1 이미지가 제1 비디오 시퀀스의 이미지일 때, 제1 비디오 시퀀스는 제2 비디오 이미지들의 시퀀스를 포함하는 제2 비디오 시퀀스와 제3 비디오 이미지들의 시퀀스를 포함하는 제3 비디오 시퀀스를 결합함으로써 획득될 수 있고, 각각의 제1 이미지는 제2 및 제3 비디오 시퀀스 내에서 동일한 연관된 타임스탬프를 각각 갖는 제3 비디오 시퀀스의 하나의 제3 이미지와 제2 비디오 시퀀스의 하나의 제2 이미지를 결합한 것에 대응한다.

[0064] 예를 들어, 제2 이미지는 예를 들어 제1 디스플레이 디바이스(60)의 후방 카메라(601)로 장면(6)을 취득함으로써 획득된다. 제2 이미지에 포함된 장면(6)의 뷰는 예를 들어 점선들로 표현된 카메라(601)의 시야(또는 초점)에 의존한다. 제2 이미지는 예를 들어 큰 시야 이미지(최대 360°, 예를 들어, 110°, 135°, 180°, 240°, 또는 360°), 또는 표준 시야 이미지(예를 들어, 65°, 70°, 또는 90°)이다. 장면(6)은 하나 이상의

제2 디스플레이 디바이스(61)(예를 들어, 텔레비전 세트, 및/또는 프로젝터 및/또는 스마트폰 및/또는 태블릿에 연관된 디스플레이 스크린), 및 임의로(optionally), 장면의 성격에 의존할 수 있는 하나 이상의 다른 객체(예를 들어, 장면이 거실에 대응할 때는 소파, 장면이 침실에 대응할 때는 침대, 장면이 식당에 대응할 때는 테이블, 장면이 영화관에 대응할 때는 좌석)(도 6에 도시되지 않음)을 포함한다. 변형예에 따르면, 장면(6)의 제2 이미지는 제1 디스플레이 디바이스(60)에 내장되지 않은 카메라, 예를 들어 웹캠 또는 보안 카메라로 취득된다.

[0065] 제1 이미지(600)는 장면의 제2 디스플레이 디바이스(61)의 표현(601)을 제1 이미지의 일부분으로서 포함한다.

[0066] 제3 이미지(611)는 제2 이미지를 취득하는 동안 제2 디스플레이 디바이스(61) 상에 디스플레이된 콘텐츠(610)에 대응한다. 제1 이미지(600)를 획득하기 위해, 제1 이미지는 제3 이미지(611)를 사용하여 제2 이미지를 증강함으로써 생성된다. 증강은 예를 들어 제2 이미지 상에, 예를 들어 제2 디스플레이 디바이스(61)의 표현에 대응하는 제2 이미지의 영역 상에 제3 이미지(611)를 오버레이함으로써 획득될 수 있다. 예를 들어 객체(도시되지 않음)가 제2 이미지 내에서 디스플레이 디바이스(61)의 부분들을 가릴(occlude) 때, 가려지지 않은 영역들에서 제3 이미지에 의해 제2 이미지를 부분적으로 오버레이하는 것과 같은 다른 증강들이 가능하다. 다른 증강 모드들은 수퍼 해상도 이미지들을 생성하는 것, 제3 이미지의 줌인된 부분들을 오버레이하는 것, 증강 전에 제3 이미지에 컬러 관리 - 장면 조명, 및 제1 및 제2 디스플레이의 속성들에 연계됨 - 을 추가하는 것을 포함할 수 있다. 제2 이미지를 획득하기 위해, 제3 이미지의 콘텐츠(610)는 예를 들어 제2 이미지 및 수신된 제3 이미지(610)를 처리하도록 구성된 이미지 처리 장치에 의해 수신된다. 제3 이미지(610)는 제2 디스플레이 디바이스(61)에 제3 이미지(610)를 송신하는 소스 디바이스, 예를 들어 저장 디바이스, 디코딩 디바이스, 셋톱 박스, DVD 또는 블루레이 판독기, 랩톱, 개인용 컴퓨팅 디바이스, 태블릿 또는 스마트폰과 같은 모바일 디바이스로부터 수신된다. 소스로부터 직접 수신된 제3 이미지를 사용하면 제3 이미지의 디스플레이 품질을 향상시킬 수 있다. 디스플레이 품질은 해상도, 블러링, 노이즈 및 색상 양태들을 포함한다. 소스 디바이스로부터 획득된 제3 이미지는 카메라(601)로 취득되고 제2 이미지 내에서 입수가 가능한 제3 이미지의 취득된 버전보다 우수한 품질을 갖는다. 실제로, 제2 이미지는 카메라의 노이즈, 카메라 센서의 제한된 공간 해상도, 카메라 전자장치들의 제한된 시간 해상도, 카메라 센서 메타메리즘(metamerism)은 물론, 카메라(601) 및/또는 디스플레이(60)의 다른 제한된 색상 용량들로 인해 어려움을 겪는다.

[0067] 도 4는 비디오 콘텐츠를 디스플레이하도록 구성된 제1 디스플레이 디바이스(4)의 하드웨어 실시예를 개략적으로 도시하며, 그것의 콘텐츠는 적어도 부분적으로 제1 디스플레이 디바이스(4)의 위치에 따라 변화한다. 제1 디스플레이 디바이스(4)는 예를 들어 HMD에 대응한다. 제1 디스플레이 디바이스(4)는 적어도 하나의 관성 센서(41), 적어도 하나의 디스플레이 유닛[예를 들어, LCD(Liquid Crystal Display), OLED(Organic Light-Emitting Diode) 또는 LCOS(Liquid Crystal On Silicon)](42), 및 송신 유닛(43)을 포함한다. 관성 센서(41)는 예를 들어 자이로스코프에 대응하고, 제1 디스플레이 디바이스(4)는 예를 들어 3개의 자이로스코프를 포함하는데, 하나는 피치(pitch)를 위한 것이고, 하나는 롤(roll)을 위한 것이고, 하나는 요우(yaw)를 위한 것이다. 다른 예에 따르면, 관성 센서(41)는 IMU(Inertial Measurement Unit)에 대응한다. 제어기는 관성 센서에 연관될 수 있으며, 제어기는 관성 센서에 의해 취득된 디바이스의 위치 변화를 표현하는 데이터를 처리하고, 이 데이터를 제1 버스를 통해 송신 유닛(43)에 송신하도록 구성된다. 송신 유닛(43)은 예를 들어 무선 인터페이스이고, 관성 센서(41)로부터 수신된 데이터를 (예를 들어, Wi-Fi 또는 블루투스에 따라) 무선 링크를 통해 외부 디바이스(도시되지 않음)에 송신하도록 구성된다. 외부 디바이스는 예를 들어 셋톱 박스(STB), 컴퓨터, 또는 비디오 콘텐츠(들)을 저장 및/또는 합성 및/또는 디코딩하도록 적응된 임의의 디바이스이다. 송신 유닛(43)은 예를 들어 외부 디바이스로부터 무선 링크를 통해, 디스플레이 유닛(들)(42) 상에 디스플레이될 비디오 콘텐츠의 제1 이미지(들)(10)를 표현하는 데이터를 수신하도록 더 구성될 수 있다. 다음으로, 디스플레이될 제1 이미지(들)(10)를 표현하는 데이터는 제2 버스를 통해 디스플레이 유닛(들)(42)에 송신된다. 변형예에 따르면, 송신 유닛은 예를 들어 USB 또는 HDMI 표준에 따라 유선 링크를 통해 데이터를 송신/수신하도록 구성된 인터페이스이다. 외부 디바이스는 HMD의 측정된 위치에 대응하는 비디오 콘텐츠의 부분을 HMD에 송신할 수 있다. HMD의 측정된 위치에 대응하는 비디오 콘텐츠의 부분은, 예를 들어 실세계에서의 HMD에 연관된 관점(point of view)과 비디오 콘텐츠에 연관된 가상 카메라의 관점 사이의 관계를 확립하는 특정 함수로 결정된다. 제1 디스플레이 디바이스의 측정된 위치에 따라 제1 디스플레이 디바이스의 디스플레이 스크린(들) 상에 디스플레이될 비디오 콘텐츠의 부분들을 제어하는 것은, 특히 비디오 콘텐츠가 제1 디스플레이 디바이스의 디스플레이 스크린(들)에 연관된 시야보다 클 때, 제1 디스플레이 디바이스를 착용한 사용자가 비디오 콘텐츠 내로 브라우징할 수 있게 한다. 예를 들어, 제1 디스플레이 디바이스, 즉 HMD에 의해 제공되는 시야가 (예를 들어, 요우 축에 대해) 110° 와 동일한 경우, 및 비디오 콘텐츠가 180° 의 콘텐츠를 제공하는 경우, HMD를 착용한 사용자는 HMD에 의해 제공되는 시야 밖

의 비디오 콘텐츠의 부분들을 보기 위해, 자신의 머리를 오른쪽 또는 왼쪽으로 회전시킬 수 있다.

- [0068] 다른 예에 따르면, 제1 디스플레이 디바이스(4)는 태블릿[예를 들어, 도 6의 태블릿(60)] 또는 스마트폰과 같은 모바일 디바이스에 대응한다. 제1 이미지는 모바일 디바이스에 내장된 메모리(예를 들어, RAM 또는 SSD)에 연관된 처리 디바이스/유닛(예를 들어, CPU 및/또는 GPU)에 의해 획득될 수 있으며, 처리 디바이스/유닛은 제1 이미지를 획득하기 위해, 제2 이미지와 제3 이미지를 결합하도록 구성된다. 이 예에 따르면, 송신 유닛(43)은 제2 이미지 및/또는 제3 이미지를 수신하도록 구성된다. 모바일 디바이스는 또한 제2 이미지를 취득하도록 구성되는 취득 디바이스, 예를 들어 카메라를 포함할 수 있다.
- [0069] 도 2a는 본 원리들의 제1의 특정한 비-제한적인 실시예에 따라, 제1 이미지(10 또는 11)를 발생시키도록 적응된 프로세스를 도시한다. 프로세스는 예를 들면 제1 이미지를 디스플레이하는 제1 디스플레이 디바이스에 내장되거나 내장되지 않은 이미지 처리 장치에 의해 구현된다. 그러한 이미지 처리 장치의 예는 도 5에 도시되어 있다.
- [0070] 프로세스의 제1 동작(20)에서, 제4 이미지(13)가 획득된다. 제4 이미지(13)는 예를 들어 카메라와 같은 취득 디바이스로부터 수신된다. 카메라는, 예를 들어 도 2와 관련하여 설명된 프로세스로부터 야기된 제1 이미지(10, 11)를 디스플레이하기 위해 사용되는 제1 디스플레이 디바이스, 예를 들어 HMD에 연관된다(예를 들어, 통합된다). 카메라는 HMD를 착용한 사용자가 서 있는 실제 환경에 연관된 카메라일 수 있다. 다른 예에 따르면, 제4 이미지는 저장 디바이스, 예를 들어 RAM(Random Access Memory) 또는 SSD(Solid-State Drive) 메모리와 같은 메모리로부터 수신된다. 제4 이미지(13)는 예를 들어 제1 디스플레이 디바이스를 착용한 사용자가 서 있는 실제 환경의 결정된 뷰를 표현하며, 뷰는 제4 이미지를 취득하기 위해 사용되는 카메라의 취득 방향에 의해 결정된다. 도 2 및 도 1d의 예에서, 제4 이미지는 제1 디스플레이 디바이스를 착용한 사용자의 손에 들려 있는 제2 디스플레이 디바이스, 예를 들어 태블릿 또는 이동 전화와 같은 모바일 디바이스의 표현을 포함한다.
- [0071] 프로세스의 제2 동작(21)에서, 실제 환경에서의 제2 디스플레이 디바이스의 위치가 결정된다. 예를 들어, 위치는 제4 이미지 내에서 제2 디스플레이 디바이스 상에, 예를 들어 제2 디스플레이 디바이스의 하나 이상의 코너에 배열된 가시적인 랜드마크들을 검출함으로써 결정되며, 가시적인 랜드마크들을 사용하는 위치 결정 방법은 예를 들어 2011년의 *컴퓨터 비전 및 패턴 인식(Computer Vision and Pattern Recognition)*의 "무작위 회귀 포레스트들을 이용한 실시간 헤드 포즈 추정(Real time head pose estimation with random regression forests)"에서 Fanelli 등에 의해 설명되어 있다. 다른 예에 따르면, 위치는 1998년 2월 18일의 "*비디오 이미지들로부터의 신속한 전역적 수렴성의 포즈 추정(Fast and Globally Convergent Pose Estimation From Video Images)*"에서 Chien-Ping Lu 등에 의해 설명된 바와 같이 객체-공간 공선성 오차(object-space collinearity error)를 최소화하는 방법을 이용하거나 비선형 최적화 알고리즘을 이용하여 제4 이미지를 분석함으로써 결정될 수 있다. 다른 예에 따르면, 예를 들어 ACM SIGGRAPH VRCAI, 2004, pp. 2226-2229의 "실시간 증강 현실 시스템을 위한 2D 기준점 추적을 이용한 견고한 카메라 포즈 추정(Robust Camera Pose Estimation Using 2D Fiducials Tracking for Real-Time Augmented Reality Systems)"에서 Fakhr-eddine Ababsa 등에 의해 설명된 바와 같이, 하나 이상의 기준 마커가 제2 디스플레이 상에 디스플레이되고, 제2 디스플레이 디바이스의 포즈를 결정하기 위해 제4 이미지를 분석할 때 사용된다. 예를 들어, 제2 디스플레이 디바이스 상에 디스플레이된 기준 마커들의 보드와 결합된 OpenCV의 ArUco 소프트웨어가 사용되어, 마커 보드의 배향 및 제4 이미지 내에서의 각각의 마커의 위치를 표현하는 3개의 벡터를 제공할 수 있다.
- [0072] 변형예에 따르면, 위치 정보는 제2 디스플레이 디바이스에 의해 송신되며, 위치를 표현하는 파라미터들은 예를 들어 제2 디스플레이 디바이스의 센서들(예를 들어, 자이로스코프들)에 의해 결정된다.
- [0073] 다른 변형예에 따르면, 배향은 예를 들어 위치 정보가 충분하지 않을 때, 위치 정보와 함께 사용될 수 있다. 예를 들어, 제2 디스플레이 디바이스가 직사각형 또는 정사각형 형상을 갖는 경우, 4개의 코너의 좌표들을 결정하는 것은 후속 동작들을 수행하기 위해 제2 디스플레이 디바이스의 위치에 관한 충분한 정보를 제공한다. 반면에, 제2 디스플레이 디바이스의 단일 지점의 좌표들이 결정될 때, 위치 정보는 3D 공간에서 제2 디스플레이 디바이스의 배향으로 완성될 수 있다. 배향 정보를 갖는 위치는 포즈 정보를 형성한다.
- [0074] 제3 동작(22)에서, 제3 이미지(220)는 동작(21)에서 결정된 제2 디스플레이 디바이스의 위치 또는 포즈에 맞게 투영된다(projected). 제3 이미지(220)는 원근 투영(perspective projection)을 적용함으로써 제2 디스플레이 디바이스의 지각되는 기하형상에 맞게 왜곡된다. 그 목표에 도달하기 위해, 제4 이미지 내의 제2 디스플레이 디바이스의 표현에 오버레이되도록 제4 이미지를 왜곡시키기 위해 적용될 호모그래픽 행렬(homographic matrix)이 결정된다. 호모그래픽 행렬은 예를 들어 동작(21)에서 결정된 위치 또는 포즈 정보를 제4 이미지의 코너

들에 대응하는 제3 이미지의 포인트들과 함께 사용하여 결정되며, 제4 이미지의 코너들에 대응하는 제3 이미지의 포인트들과 제2 디스플레이 디바이스의 표현의 대응하는 포인트들 사이에 매핑이 확립된다. 호모그래픽 행렬을 입력된 제3 이미지(220)에 적용하면, 제2 디스플레이 디바이스의 표현의 형상의 대응하는 형상을 갖는 왜곡된 제3 이미지(221)가 제공된다.

[0075] 제4 이미지가 제1 디스플레이 디바이스 외부의 카메라로 취득될 때, 제3 이미지를 왜곡시키기 위해, 제1 디스플레이 디바이스의 포즈를 표현하는 제2 정보(제2 디스플레이 디바이스의 포즈 정보와 동일한 기준 프레임으로 표현됨)가 필요하다. 이러한 제2 정보는 제2 디스플레이 디바이스의 제1 포즈 정보와 동일한 방식으로, 즉 예를 들어 제1 디스플레이 디바이스 및 제2 디스플레이 디바이스를 포함하는 이미지를 분석함으로써 획득될 수 있거나, 제1 디스플레이 디바이스에 연관된 센서들로부터 직접 획득될 수 있다.

[0076] 순차적으로 또는 동시에 수행될 수 있는 제4 동작(23) 및 제5 동작(24)에서, 제4 이미지(13)에서의 제2 디스플레이 디바이스(231) 및 사용자의 손들(241)의 표현은 세그먼트화된다. 이미지 세그먼트화는 이미지를 복수의 세그먼트로 분할하는 프로세스이고, 각각의 세그먼트는 픽셀들의 세트에 대응한다. 예를 들어 위치 정보가 제2 디스플레이 디바이스의 코너들의 좌표들에 의해 주어질 때, 제2 디스플레이 디바이스의 세그먼트화는 위치를 표현하는 정보에 기초하여 수행된다. 변형예에 따르면, 예를 들어 제2 디스플레이 디바이스의 단 하나의 코너의 좌표들만이 알려질 때, 또는 제2 디스플레이 디바이스의 배향 정보만이 알려질 때, 세그먼트화는 치수들(예를 들어, 길이 및 폭)을 표현하는 정보에 더 기초한다. 상기 치수 정보는 예를 들어 제2 디스플레이 디바이스에 대해 수신되거나, 제2 디스플레이 디바이스들의 상이한 참조들의 목록에 대한 정보를 제공하는 룩업 테이블로부터 검색된다. 사용자의 손들의 세그먼트화는 예를 들어 "HSV 색 공간을 이용한 피부 검출(*Skin Detection using HSV color space*)"에서 V.A. Oliveira 등에 의해 설명된 바와 같이, HSV 색 공간에서 피부 색조를 검출함으로써 수행된다.

[0077] 제6, 제7 및 제8 동작(25, 26, 27)에서, 최종 제1 이미지(221)는 순차적으로 또는 동시에 수행될 수 있는 합성 동작들로 획득된다. 합성은 상이한 소스들로부터의 시각적 요소들을 단일 이미지(들)로 결합하는 것에 대응한다. 왜곡된 제3 이미지(271)는 제2 디스플레이 디바이스의 세그먼트화된 표현(231)에 오버레이되고, 손들(241)의 세그먼트화된 표현은 왜곡된 제3 이미지를 디스플레이하는 세그먼트화된 표현의 전방에 보여지도록 제1 이미지 내에 삽입된다. 최종 합성(27)은 최종 제1 이미지(271)를 발생시키기 위해 왜곡된 제3 이미지를 갖는 손들의 표현을 제2 이미지에 인크루스트하거나 오버레이하는 것에 있다. 이것은 사용자가 자신이 현재 제1 디스플레이 디바이스 상에서 보고 있는 콘텐츠 내에 보조 콘텐츠를 디스플레이하는 윈도우를 가질 수 있게 한다. 제3 이미지의 콘텐츠에 의존하여, 그것은 사용자가 예를 들어 자신이 서 있는 실제 환경에 대한 뷰를 가질 수 있게 하고, 및/또는 자신이 현재 서 있는 실제 환경의 하나 이상의 객체를 제어하고, 및/또는 자신이 현재 보고 있는 콘텐츠를 제어할 수 있게 한다.

[0078] 도 2b는 본 원리들의 제2의 특정한 비-제한적인 실시예에 따라, 제1 이미지(600)를 발생시키도록 적응된 추가 프로세스를 도시한다. 프로세스는 예를 들면 제1 이미지를 디스플레이하는 제1 디스플레이 디바이스에 내장되거나 내장되지 않는 이미지 처리 장치에 의해 구현된다. 그러한 이미지 처리 장치의 예는 도 5에 도시되어 있다.

[0079] 프로세스의 제1 동작(201)에서, 장면(6)을 표현하는 하나 이상의 제2 이미지가 획득된다. 제2 이미지는 예를 들어 카메라와 같은 취득 디바이스로부터 수신된다. 카메라는, 예를 들어 도 7과 관련하여 설명된 프로세스로부터 야기된 제1 이미지(600)를 디스플레이하기 위해 사용되는 제1 디스플레이 디바이스, 예를 들어 HMD 또는 태블릿에 연관된다(예를 들어, 통합된다). 카메라는 제1 디스플레이 디바이스를 착용한 사용자가 서 있는 실제 환경에 연관된 카메라일 수 있다. 다른 예에 따르면, 제2 이미지는 저장 디바이스, 예를 들어 RAM(Random Access Memory) 또는 SSD(Solid-State Drive) 메모리와 같은 메모리로부터 수신되거나 획득된다. 제2 이미지는 예를 들어 제1 디스플레이 디바이스를 착용한 사용자가 서 있는 실제 환경의 결정된 뷰를 표현하며, 뷰는 제2 이미지를 취득하기 위해 사용되는 카메라의 취득 방향에 의해 결정된다. 도 7 및 도 6의 예에서, 제2 이미지는 제2 디스플레이 디바이스(61), 예를 들어 텔레비전 세트의 표현을 포함한다.

[0080] 프로세스의 제2 동작(202)에서, 하나 이상의 제3 이미지(610)가 획득되는데, 예를 들어 소스 디바이스로부터 수신된다. 제3 이미지들은 예를 들어 장면(6)의 제2 디스플레이 디바이스 상에 디스플레이된 이미지들에 대응한다. 소스 디바이스는 예를 들어 제2 디스플레이 디바이스 상에 디스플레이하기 위해 제3 이미지들을 송신 및/또는 디코딩하는 소스 디바이스에 대응한다. 하나 이상의 제3 이미지는 이미지 처리 장치의 메모리(예를 들어, RAM 또는 SSD) 내에, 또는 추가 처리를 위해 캐시 내에 저장될 수 있다. 하나 이상의 제3 이미지는 무선(예를

들어, WiFi 또는 블루투스) 또는 유선(예를 들어, 인터넷 또는 USB 또는 HDMI) 접속을 통해 수신된다.

- [0081] 프로세스의 제3 동작(203)에서, 제2 이미지의 관점 및 뷰잉 방향(예를 들어, 제2 이미지의 취득 관점)에 관한 장면(6)에서의 제2 디스플레이 디바이스(61)의 위치, 배향, 스케일 및 형상이 결정된다. 그 목표에 도달하기 위해, 제2 이미지 내에서의 제2 디스플레이 디바이스의 표현이 예를 들어 제2 이미지 내에서 검출된다.
- [0082] 예를 들어, 위치는 제2 이미지 내에서 제2 디스플레이 디바이스 상에, 예를 들어 제2 디스플레이 디바이스의 하나 이상의 코너에 배열된 가시적인 랜드마크들을 검출함으로써 결정되며, 가시적인 랜드마크들을 사용하는 위치 결정 방법은 예를 들어 2011년의 *컴퓨터 비전 및 패턴 인식(Computer Vision and Pattern Recognition)*의 "무작위 회귀 포레스트들을 이용한 실시간 헤드 포즈 추정(Real time head pose estimation with random regression forests)"에서 Fanelli 등에 의해 설명되어 있다. 랜드마크들의 스케일 및 배향은 제2 디스플레이의 스케일 및 배향을 제공한다. 제2 디스플레이의 형상은 예를 들어 검출된 랜드마크들에 의해 점유되는 제2 이미지 내의 영역의 형상을 분석함으로써 획득될 수 있다.
- [0083] 다른 예에 따르면, 위치, 배향, 및 스케일은 1998년 2월 18일의 "*비디오 이미지들로부터의 신속한 전역적 수렴성의 포즈 추정(Fast and Globally Convergent Pose Estimation From Video Images)*"에서 Chien-Ping Lu 등에 의해 설명된 바와 같이 객체-공간 공선성 오차를 최소화하는 방법을 이용하거나 비선형 최적화 알고리즘을 이용하여 제2 이미지를 분석함으로써 결정될 수 있다.
- [0084] 다른 예에 따르면, 제2 디스플레이 디바이스의 위치는 제2 이미지의 콘텐츠를, 장면(6) 내에서 제2 디스플레이 디바이스 상에 디스플레이된 비디오 콘텐츠(즉, 제3 이미지)와 비교함으로써 결정되며, 비디오 콘텐츠(즉, 제3 이미지)는 후자의 비디오 콘텐츠를 제2 디스플레이 디바이스에 송신하는 소스로부터 직접 수신된다. 제2 디스플레이 디바이스(6) 상에 디스플레이되는 비디오의 콘텐츠(즉, 제3 이미지)를 안다면, 비교 프로세스의 목표는 제2 이미지 내에서의 그러한 비디오 콘텐츠의 위치를 지정하는 것이고, 이는 제2 이미지 내에서의 제2 디스플레이 디바이스의 표현의 위치, 배향, 스케일, 및 형상을 검출하는 것을 가능하게 한다.
- [0085] 또 다른 예에 따르면, 제2 디스플레이 디바이스의 위치는 상이한(시간적으로 말해서) 연속적인 제2 이미지들의 콘텐츠를 서로 비교함으로써 결정된다. 연속적인 제2 이미지들의 그룹(예를 들어, 제2 이미지들의 시퀀스에 대응하는 비디오 시퀀스의 10, 20 또는 20 프레임)에 연관된 관점이 변하지 않는다고 고려하면, 시간에 따라 변하는 제2 이미지의 콘텐츠의 유일한 부분은 제2 디스플레이 디바이스(61) 상에 디스플레이된 콘텐츠에 대응한다. 제2 이미지들은 그들 중 하나에 대하여 위치맞춤될(registered) 수 있고, 제2 디스플레이 디바이스의 표현은 제2 디스플레이 디바이스 상에 제3 이미지들을 디스플레이한 것으로 인한 제2 이미지들의 시퀀스 내의 변화들을 강조하는 이미지 구별(image differentiation)에 의해 검출될 수 있고, 제3 이미지들의 콘텐츠는 시간에 따라 변한다.
- [0086] 다른 예에 따르면, 예를 들어 ACM SIGGRAPH VRCAI, 2004, pp. 2226-2229의 "실시간 증강 현실 시스템을 위한 2D 기준점 추적을 이용한 견고한 카메라 포즈 추정(Robust Camera Pose Estimation Using 2D Fiducials Tracking for Real-Time Augmented Reality Systems)"에서 Fakhr-eddine Ababsa 등에 의해 설명된 바와 같이, 하나 이상의 기준 마커가 제2 디스플레이 디바이스 상에 디스플레이되고, 제2 디스플레이 디바이스의 포즈를 결정하기 위해 제4 이미지를 분석할 때 사용된다. 예를 들어, 제2 디스플레이 디바이스 상에 디스플레이된 기준 마커들의 보드와 결합된 OpenCV의 ArUco 소프트웨어가 사용되어, 마커 보드의 배향 및 제4 이미지 내에서의 각각의 마커의 위치를 표현하는 3개의 벡터를 제공할 수 있다.
- [0087] 변형예에 따르면, 배향은 예를 들어 위치 정보가 충분하지 않을 때, 위치 정보와 함께 사용될 수 있다. 예를 들어, 제2 디스플레이 디바이스가 직사각형 또는 정사각형 형상을 갖는 경우, 4개의 코너의 좌표들을 결정하는 것은 후속 동작들을 수행하기 위해 제2 디스플레이 디바이스의 위치에 관한 충분한 정보를 제공한다. 반면에, 제2 디스플레이 디바이스의 단일 지점의 좌표들이 결정될 때, 위치 정보는 3D 공간에서 제2 디스플레이 디바이스의 배향으로 완성될 수 있다. 배향 정보를 갖는 위치는 포즈 정보를 형성한다.
- [0088] 제4 동작(204)에서, 하나 이상의 제3 이미지(610)는 동작(203)에서 결정된 제2 디스플레이 디바이스의 위치 또는 포즈에 맞게 투영된다. 제3 이미지(610)는 투영을 적용함으로써 제2 디스플레이 디바이스의 지각되는 기하 형상에 맞게 왜곡된다. 투영은 투시(perspective), 약한 투시(weak perspective), 정사영(orthographic), 또는 예를 들어 충분한 기하학적 정밀도 또는 충분히 가벼운 구현 풋프린트(sufficiently light implementation footprint)를 갖는 임의의 다른 적절한 유형일 수 있다. 그 목표에 도달하기 위해, 비-제한적인 예에 따라, 제2 이미지 내의 제2 디스플레이 디바이스의 표현에 오버레이되도록 수신된 제3 이미지(610)를 왜곡시키기 위해

적용될 호모그래픽 행렬이 결정된다. 호모그래픽 행렬은 예를 들어 동작(203)에서 결정된 위치 또는 포즈 정보를 수신된 직사각형 제3 이미지(610)의 코너들에 대응하는 제3 이미지(610)의 포인트들과 함께 사용하여 결정되며, 수신된 제3 이미지의 코너들과 제2 디스플레이 디바이스의 표현의 대응하는 포인트들(예를 들어, 제2 이미지 내에서 검출된 제2 디스플레이 디바이스의 프레임의 코너들) 사이에 매핑이 확립된다. 호모그래픽 행렬을 수신된 제3 이미지(610)에 적용하면, 제2 디스플레이 디바이스의 표현의 형상의 대응하는 형상을 갖는 왜곡된 제3 이미지(611)가 제공된다.

[0089] 제2 이미지가 제1 디스플레이 디바이스 외부의 카메라로 취득될 때, 수신된 제3 이미지(610)를 왜곡시키기 위해, (제2 디스플레이 디바이스의 포즈 정보와 동일한 기준 프레임으로 표현되는) 제1 디스플레이 디바이스의 포즈를 표현하는 제2 정보가 필요하다. 이러한 제2 정보는 제2 디스플레이 디바이스의 제1 포즈 정보와 동일한 방식으로, 즉 예를 들어 제1 디스플레이 디바이스 및 제2 디스플레이 디바이스를 포함하는 이미지를 분석함으로써 획득될 수 있거나, 제1 디스플레이 디바이스에 연관된 센서들로부터 직접 획득될 수 있다.

[0090] 제5 동작(205)에서, 순차적으로 또는 동시에 수행될 수 있는 합성 동작들로 최종 제1 이미지(600)가 획득된다. 합성은 서로 다른 소스들로부터의 시각 요소들을 단일 이미지(들)로 결합 또는 증강하는 것에 대응한다. 왜곡된 제3 이미지(611)는 예를 들어 제2 디스플레이 디바이스의 표현(601)에 오버레이된다. 이는 사용자가 제1 디스플레이 디바이스의 스크린의 해상도와 함께, 수신된 제3 이미지(610)의 해상도로부터 야기되는 최적의 해상도로부터 이익을 얻을 수 있게 한다.

[0091] 변형예에 따르면, 왜곡된 제3 이미지(611)는 제2 이미지들 내에서 시간에 따라 고정된 위치에 오버레이된다. 예를 들어, 연속적인 제2 이미지들의 시퀀스에서, 최초의 제3 이미지는 제2 디스플레이 디바이스의 표현에 대응하는 시퀀스의 최초의 제2 이미지의 영역에 오버레이될 수 있고, 다음으로, 제2 이미지들의 시퀀스 내에서 제2 디스플레이 디바이스의 표현의 위치가 시간에 따라 변하더라도, (제3 이미지들의 시퀀스의) 다른 제3 이미지들은 시퀀스의 다른 제2 이미지들 내의 동일 위치에 오버레이될 수 있다. 제1 디스플레이 디바이스를 들고서 제2 이미지들을 취득하고 있는 사용자가 장면 내에서 이동할 때, 제2 디스플레이 디바이스의 표현의 위치가 달라질 수 있다. 제3 이미지(들)의 오버레이를 고정된 위치에 고정하면, 예를 들어 제2 이미지들을 취득하는 사용자의 손떨림에 기인하는 제3 이미지들의 작은 흔들림을 방지할 수 있다.

[0092] 임의적 변형예에 따르면, 합성 프로세스 동안, 예를 들어 증강 현실 장면 또는 증강 가상 장면을 표현하는 제1 이미지를 발생시킬 때, 가상 객체들이 제1 이미지에 추가될 수 있다.

[0093] 도 3은 본 원리들의 특정한 비-제한적인 실시예에 따라, 제1 이미지를 제1 디스플레이 디바이스에 송신하는 방법을 도시한다.

[0094] 초기화 단계(30)에서, 장치(5)의 상이한 파라미터들이 갱신된다. 구체적으로, 포즈 정보, 원근 투영의 파라미터들은 어떤 식으로든 초기화된다.

[0095] 단계(31)에서, 제2 이미지가 수신된다. 제2 이미지는 예를 들어 컴퓨터 발생 이미지(예를 들어, 게임 애플리케이션을 위한 것)와 같이 가상 환경을 표현한다. 다른 예들에 따르면, 제2 이미지는 실제 환경, 예를 들어 하나 이상의 카메라로 취득된 이미지, 또는 실제 및 가상 요소들을 포함하는 이미지를 표현하는 이미지이다. 제2 이미지는 연속적인 제2 이미지들의 시퀀스(즉, 비디오 시퀀스)에 속할 수 있다. 특정 실시예에 따르면, 제2 이미지는 몰입형 콘텐츠의 일부이고, 제2 이미지의 치수들은 제1 디스플레이 디바이스가 디스플레이할 수 있는 콘텐츠의 크기에 대응한다. 제2 이미지는 예를 들어 인터넷을 통해, 또는 LAN(Local Area Network)의 무선 또는 유선 접속을 통해 원격 저장 디바이스(예를 들어, 서버)로부터 수신된다. 다른 예에 따르면, 제2 이미지는 제2 이미지를 처리하도록 구성되는 처리 유닛(예를 들어, 프로세서, 예를 들어 CPU 및/또는 하나 이상의 GPU)으로부터 로컬 메모리를 접속하는 버스를 통해 로컬 메모리(예를 들어, RAM 또는 SSD)로부터 수신된다. 다른 예에 따르면, 하나 이상의 제2 이미지는 이미지 취득 디바이스, 예를 들어 제1 디스플레이 디바이스에 연관된 카메라로부터 수신된다.

[0096] 단계(32)에서, 제2 디스플레이 디바이스의 위치를 표현하는 제1 정보가 획득, 예를 들면 수신 또는 결정된다. 제1 정보는 예를 들어 유선 또는 무선 접속을 통해 제2 디스플레이 디바이스로부터 수신될 수 있다. 다른 예에 따르면, 도 2a의 동작(21) 또는 도 2b의 동작(203)에서 더 상세하게 설명되는 바와 같이, 제2 디스플레이 디바이스의 표현을 포함하는 제4 이미지에 포즈 추정 알고리즘을 적용함으로써 제1 정보가 결정된다. 제4 이미지(220)(도 6 및 도 2b의 예에서 제2 이미지라고도 지칭됨)는 제1 디스플레이 디바이스에 연관된 카메라, 또는 제1 디스플레이 디바이스 외부에 있는, 예를 들어 제1 디스플레이 디바이스를 착용하거나 들고 있는 사용자가 서

있는 실제 환경에 속하는 카메라로 취득될 수 있다. 임의적 변형예에 따르면, 제2 디스플레이 디바이스의 배향 정보가 획득될 수 있고, 함께 결합된 위치 및 배향 정보는 제2 디스플레이 디바이스의 포즈를 제공한다.

- [0097] 단계(33)에서, 제3 이미지가 수신된다. 제3 이미지는 예를 들어 인터넷을 통해, 또는 LAN(Local Area Network)에서의 무선 또는 유선 접속을 통해 원격 저장 디바이스(예를 들어, 서버)로부터, 또는 원격 디바이스(예를 들어, 사용자가 서 있는 실제 환경에 속하는 카메라 또는 제2 디스플레이 디바이스)로부터 수신된다. 제3 이미지는 다음을 포함하는 이미지들의 그룹에 속한다:
- [0098] - 실제 환경, 예를 들어 제1 디스플레이 디바이스를 착용하거나 들고 있는 사용자가 서 있는 실제 환경을 표현하는 영화 또는 비디오 시퀀스와 같은 비디오 시퀀스의 이미지. 상기 이미지는 예를 들어 (예를 들어, 사용자가 들고 있는) 제2 디스플레이 디바이스에 연관된 카메라, 또는 상기 실제 환경 내에 위치한 카메라로 취득된다. 변형예에 따르면, 이미지는 사용자의 환경으로부터 떨어져 있는 실제 환경을 표현하며, 예를 들면 사용자가 서 있는 룸과는 다른 룸을 표현한다. 다른 변형예에 따르면, 이미지는 인물, 예를 들어 인물의 얼굴을 표현하며, 이미지는 예를 들어 상기 인물의 스마트폰의 카메라 또는 태블릿의 카메라로 취득되고 제2 디스플레이 디바이스 상에서 수신된다;
- [0099] - 가상 환경을 표현하는 비디오 시퀀스의 이미지;
- [0100] - 실제 및 가상 요소들 또는 객체들을 결합하는 장면을 표현하는 비디오 시퀀스의 이미지;
- [0101] - 예를 들어 사용자의 환경의 요소들[예를 들어, 알람, 서터들, 조명의 스위치 온/오프]을 제어하기 위해 제1 및/또는 제2 이미지의 파라미터들을 제어하도록 적응된 그래픽 사용자 인터페이스(GUI)의 하나 이상의 그래픽 객체를 포함하는 이미지; 및
- [0102] - 텍스트 요소들, 예를 들어 제2 디스플레이 디바이스 상에 수신된 SMS(Short Message Service)의 텍스트 요소들을 포함하는 이미지.
- [0103] 제3 이미지는 위에서 언급된 이미지들 중 수 개 또는 전부의 조합의 위에서 언급된 이미지들 중의 임의의 것일 수 있고, 제3 이미지들의 시퀀스는 예를 들어 실제 환경의 하나 이상의 이미지, 다음으로 제2 디스플레이 디바이스 상의 사용자를 부르는 인물의 얼굴의 하나 이상의 이미지, 다음으로 하나 이상의 텍스트 요소를 포함할 수 있다. 변형예에 따르면, 제3 이미지는 위에서 언급된 이미지들, 예를 들어 그래픽 및/또는 텍스트 요소를 내장한 사용자의 실제 환경의 이미지의 수 개의 합성일 수 있다.
- [0104] 다음으로, 제3 이미지는 도 2의 동작(22)과 관련하여 더 상세히 설명된 바와 같이, 단계(32)에서 획득된 제1 위치 정보에 따라(또는 제1 포즈 정보에 따라) 왜곡된다.
- [0105] 단계(34)에서, 왜곡된 제3 이미지와 수신된 제2 이미지를 결합함으로써 제1 이미지가 발생(예를 들어, 합성)된다. 왜곡된 제3 이미지는 예를 들어 수신된 제2 이미지의 결정된 영역 상에, 예를 들어 제2 이미지의 좌측 하단 부분 또는 제2 이미지의 우측 하단 부분 또는 사용자에게 의해 정의된 부분 내에서 중첩된다. 변형예에 따르면, 합성은 제1 디스플레이 디바이스를 착용하거나 들고 있는 사용자에게 의해 들려 있는 제2 디스플레이 디바이스의 표현을 포함하는 제4 이미지의 세그먼트화된 부분들의 결합을 더 포함한다. 이러한 변형예에 따르면, 제3 이미지가 디스플레이되는 제2 디스플레이 디바이스를 들고 있는 사용자의 손들과 함께 제2 디스플레이 디바이스를 표현하는 세그먼트화된 부분에 제3 이미지가 오버레이된다. 다음으로, 제1 디스플레이 디바이스를 착용한 사용자는 자신이 몰입해 있는 환경을 떠나지 않고서, 보조 콘텐츠, 예를 들어 자신의 환경에 대한 뷰 또는 제2 디스플레이 디바이스(예를 들어, 스마트폰 또는 태블릿)에 착신되는 메시지에 대한 뷰에 접근할 수 있다. 소스로부터 직접 수신된 제3 이미지의 오버레이는 상기 제3 이미지를 디스플레이 할 때 제1 디스플레이 디바이스의 최대 해상도로부터 이익을 얻을 수 있게 한다.
- [0106] 단계(35)에서, 단계(34)에서 획득된 제1 이미지를 표현하는 데이터는 예를 들어 제1 디스플레이 디바이스의 스크린(들) 상에 디스플레이되기 위해, 제1 디스플레이 디바이스에 송신된다. 데이터는 예를 들어 제1 이미지의 픽셀들에 연관된 텍스처 정보, 예를 들어 각각의 픽셀에 연관된 RGB 정보를 포함한다.
- [0107] 단계들은 제1 이미지들의 시퀀스를 발생시키기 위해, 수 개의 제2 이미지(즉, 제2 이미지들의 시퀀스) 및/또는 수 개의 제3 이미지(즉, 제3 이미지들의 시퀀스), 및/또는 수 개의 제4 이미지(즉, 제4 이미지들의 시퀀스)에 대해 반복될 수 있다.
- [0108] 도 5는 제1 이미지를 합성 및/또는 송신하도록 구성된 장치(5)의 하드웨어 실시예를 개략적으로 도시한다. 디바이스(5)는 또한 1개 또는 수 개의 제1 이미지의 디스플레이 신호들의 생성을 위해 구성된다. 디바이스(5)는

예를 들어 태블릿, 스마트폰, 게임 콘솔, 컴퓨터, 랩톱, 또는 셋톱 박스에 대응한다.

- [0109] 장치(5)는 또한 클록 신호를 전송하는 어드레스들 및 데이터의 버스(55)에 의해 서로 접속된 다음의 구성요소들을 포함한다:
- [0110] - 마이크로프로세서(51)(또는 CPU),
- [0111] - 그래픽 카드(52) - 그래픽 카드는
- [0112] • 수 개의 그래픽 프로세서 유닛(또는 GPU)(520),
- [0113] • 그래픽 랜덤 액세스 메모리(GRAM)(521)
- [0114] 를 포함함 -,
- [0115] - ROM(Read Only Memory) 유형의 비휘발성 메모리(56),
- [0116] - 랜덤 액세스 메모리 또는 RAM(57),
- [0117] - 제1 이미지들을 표현하는 데이터를 예를 들어 제1 디스플레이 신호에 송신하도록 구성된 송신기(58),
- [0118] - 제2 디스플레이 디바이스로부터의, 카메라(들)로부터의, 제1 디스플레이 디바이스로부터의 데이터(예를 들어, 뷰잉 방향을 표현하는 정보), 제2 및/또는 제3 및/또는 제4 이미지들을 표현하는 데이터, 제2 디스플레이 디바이스로부터의 제어 파라미터들을 수신하도록 구성되는 수신기(59),
- [0119] - 예를 들어, 촉각 인터페이스, 마우스, 웹캠 등과 같은 1개 또는 수 개의 I/O(입력/출력) 디바이스(54), 및
- [0120] - 전원(59).
- [0121] 장치(5)는 그래픽 카드에서 계산된 이미지들, 예를 들어 라이브를 디스플레이하기 위해 그래픽 카드(52)에 직접 접속된 디스플레이 스크린 유형의 하나 이상의 디스플레이 디바이스(53)를 또한 포함할 수 있다. 디스플레이 디바이스(53)를 그래픽 카드(52)에 접속하기 위한 전용 버스의 사용은 훨씬 더 큰 데이터 송신 비트레이트들을 갖고, 따라서 그래픽 카드에 의해 합성된 이미지들의 디스플레이에 대한 레이턴시 시간을 감소시키는 이점을 제공한다. 변형예에 따르면, 디스플레이 디바이스는 장치(5)의 외부에 있으며, 디스플레이 신호들을 송신하기 위해 케이블에 의해 또는 무선으로 장치(5)에 접속된다. 장치(5), 예를 들어 그래픽 카드(52)는, 예를 들어 제1 디스플레이 디바이스(예를 들어, HMD), LCD 또는 플라즈마 스크린 또는 비디오 프로젝터와 같은 외부 디스플레이 수단에 디스플레이 신호를 송신하도록 적응된 송신 또는 접속을 위한 인터페이스(도 5에 도시되지 않음)를 포함한다.
- [0122] 메모리들(521, 56, 및 57)의 설명에서 사용된 "레지스터"라는 단어는 언급된 메모리들 각각에서 저장량의 메모리 구역(약간의 바이너리 데이터)은 물론, 대용량의 메모리 구역(전체 프로그램이 저장될 수 있게 하거나, 계산된 또는 디스플레이될 데이터를 표현하는 데이터의 데이터 전체 또는 일부) 둘 다를 지정한다는 점에 유의해야 한다.
- [0123] 스위치 온될 때, 마이크로프로세서(51)는 RAM(57)에 포함된 프로그램의 명령어들을 로딩하고 실행한다.
- [0124] 랜덤 액세스 메모리(57)는 특히 다음을 포함한다:
- [0125] - 레지스터(570) 내에서, 장치(5)의 스위치 온을 담당하는 마이크로프로세서(51)의 운영 프로그램,
- [0126] - 제1, 제2, 제3, 및/또는 제4 이미지(들)를 표현하는 데이터(571)(예를 들어, RGB 데이터),
- [0127] - 제2 디스플레이 디바이스 및/또는 제1 디스플레이 디바이스의 위치(및/또는 포즈)를 표현하는 정보(572).
- [0128] 본 개시내용에 특정한 방법(들)(예를 들어, 제1 이미지를 송신하는 방법 및/또는 제1 이미지를 합성하는 방법)의 단계들을 구현하는 알고리즘들은 이러한 단계들을 구현하는 장치(5)에 연관된 그래픽 카드(52)의 메모리(GRAM)(521) 내에 저장된다. 스위치 온되고, 데이터(571) 및 정보(572)가 RAM(57)에 로딩되고 나면, 그래픽 카드(52)의 그래픽 프로세서들(520)은 이러한 파라미터들을 GRAM(521) 내에 로딩하고, 이러한 알고리즘들의 명령어들을 예를 들어 HLSL(High Level Shader Language) 언어 또는 GLSL(OpenGL Shading Language)을 사용하여 "셰이더(shader)" 유형의 마이크로프로그램들의 형태로 실행한다.
- [0129] 랜덤 액세스 메모리(GRAM)(521)는 특히 다음을 포함한다:

- [0130] - 레지스터 내에서, 제2 이미지들을 표현하는 데이터;
- [0131] - 레지스터 내에서, 제3 및/또는 제4 이미지들을 표현하는 데이터;
- [0132] - 레지스터 내에서, 제1 이미지를 표현하는 데이터;
- [0133] - 레지스터 내에서, 위치 또는 포즈 정보를 표현하는 데이터.
- [0134] 다른 변형예에 따르면, GRAM(521) 내에서 이용가능한 메모리 저장 공간이 불충분한 경우, RAM(57)의 일부가 식별자들 및 거리들의 저장을 위해 CPU(51)에 의해 할당된다. 그러나, 이러한 변형예는 GPU들 내에 포함된 마이크로프로그램들로부터 합성된 환경의 표현을 포함하는 이미지의 합성에서 더 큰 레이턴시 시간을 야기하는데, 왜냐하면 데이터는 GPU들로부터 GRAM으로의, 및 그 반대로의 데이터 송신을 위해 그래픽 카드 내에서 이용가능한 것들에 비해 일반적으로 열등한 송신 용량들을 갖는 버스(55)를 지나 그래픽 카드로부터 랜덤 액세스 메모리(57)에 송신되어야 하기 때문이다.
- [0135] 다른 변형예에 따르면, 전력 공급부(58)는 장치(5)의 외부에 있다.
- [0136] 대안적인 실시예에서, 장치(5)는 어떠한 ROM도 포함하지 않고 RAM만을 포함하며, 도 2 또는 도 3과 관련하여 설명된 본 개시내용에 특정한 방법의 단계들을 구현하는 알고리즘들은 RAM 내에 저장된다. 다른 변형예에 따르면, 장치(5)는 ROM 및/또는 RAM을 대신하여 SSD(Solid-State Drive) 메모리를 포함한다.
- [0137] 당연히, 본 개시내용은 앞에서 설명된 실시예들에 제한되지 않는다.
- [0138] 구체적으로, 본 개시내용은 제1 이미지를 송신하는 방법에 제한되는 것이 아니라, 제1 이미지를 합성/발생시키는 방법(및 그것을 위해 구성된 디바이스), 및/또는 제1 이미지를 디스플레이하는 방법(및 그것을 위해 구성된 디바이스)으로 확장된다.
- [0139] 본 명세서에 설명된 구현예들은 예를 들어 방법 또는 프로세스, 장치, 소프트웨어 프로그램, 데이터 스트림, 또는 신호로 구현될 수 있다. 단일 형태의 구현예의 맥락에서만 논의(예를 들어, 방법 또는 디바이스로서만 논의)되더라도, 논의된 특징들의 구현은 다른 형태들(예를 들어, 프로그램)로도 구현될 수 있다. 장치는 예를 들어 적절한 하드웨어, 소프트웨어 및 펌웨어로 구현될 수 있다. 예를 들어, 방법들은, 예를 들어 컴퓨터, 마이크로프로세서, 집적 회로, 또는 프로그램가능한 로직 디바이스를 포함하는 처리 디바이스들을 일반적으로 지칭하는 프로세서와 같은 장치에서 구현될 수 있다. 프로세서들은 또한 예를 들어, 스마트폰, 태블릿, 컴퓨터, 이동 전화, 휴대용/개인용 정보 단말("PDA"), 및 최종 사용자들 간의 정보의 통신을 용이하게 하는 다른 디바이스들과 같은 통신 디바이스들을 포함한다.
- [0140] 본 명세서에 설명된 다양한 프로세스들 및 특징들의 구현예들은 다양한 상이한 장비 또는 애플리케이션들, 특히 예를 들어 데이터 인코딩, 데이터 디코딩, 뷰 발생, 텍스처 처리, 및 이미지 및 관련 텍스처 정보 및/또는 깊이 정보의 다른 처리에 연관된 장비 또는 애플리케이션들에 구현될 수 있다. 그러한 장비의 예들은 인코더, 디코더, 디코더로부터의 출력을 처리하는 포스트-프로세서, 인코더로의 입력을 제공하는 프리-프로세서, 비디오 코더, 비디오 디코더, 비디오 코덱, 웹 서버, 셋톱 박스, 랩톱, 개인용 컴퓨터, 셀폰, PDA 및 다른 통신 디바이스들을 포함한다. 명백한 바와 같이, 장비는 이동식일 수 있고, 심지어는 이동 차량에 설치될 수 있다.
- [0141] 추가로, 방법들은 프로세서에 의해 수행되는 명령어들에 의해 구현될 수 있으며, 그러한 명령어들(및/또는 구현에 의해 생성된 데이터 값들)은 예를 들어, 집적 회로, 소프트웨어 캐리어, 또는 예컨대 하드 디스크, 콤팩트 디스켓("CD"), (예를 들어, 종종 디지털 다용도 디스크 또는 디지털 비디오 디스크라고 지칭되는 DVD와 같은) 광 디스크, 랜덤 액세스 메모리("RAM") 또는 읽기 전용 메모리("ROM")와 같은 다른 저장 디바이스와 같은 프로세서 판독가능한 매체 상에 저장될 수 있다. 명령어들은 프로세서 판독가능한 매체 상에 유형으로(tangibly) 구현된 애플리케이션 프로그램을 형성할 수 있다. 명령어들은 예를 들어 하드웨어, 펌웨어, 소프트웨어 또는 조합일 수 있다. 명령어들은 예를 들어 운영 체제, 별도의 애플리케이션, 또는 이 둘의 조합에서 발견될 수 있다. 그러므로, 프로세서는 프로세스를 수행하도록 구성되는 디바이스, 및 프로세스를 수행하기 위한 명령어들을 갖는 프로세서 판독가능한 매체(예컨대, 저장 디바이스)를 포함하는 디바이스 둘 다로서 특징지어질 수 있다. 또한, 프로세서 판독가능한 매체는 명령어들에 추가하여 또는 명령어들을 대신하여, 구현에 의해 생성된 데이터 값들을 저장할 수 있다.
- [0142] 본 기술분야의 통상의 기술자에게 명백한 바와 같이, 구현예들은 예를 들어 저장되거나 송신될 수 있는 정보를 운반하도록 포맷된 다양한 신호들을 생성할 수 있다. 정보는 예를 들어 방법을 수행하기 위한 명령어들, 또는 설명된 구현예들 중 하나에 의해 생성된 데이터를 포함할 수 있다. 예를 들어, 신호는 설명된 실시예의 선택스

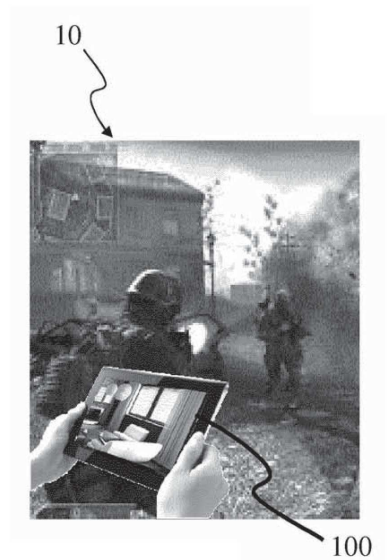
를 기입 또는 판독하기 위한 규칙들을 데이터로서 운반하거나, 설명된 실시예에 의해 기입된 실제 신호스 값들을 데이터로서 운반하도록 포맷될 수 있다. 이러한 신호는 예를 들어 전자기파(예를 들어, 스펙트럼의 무선 주파수 부분을 사용함) 또는 기저대역 신호로서 포맷될 수 있다. 포맷은 예를 들어 데이터 스트림을 인코딩하고 인코딩된 데이터 스트림으로 캐리어를 변조하는 것을 포함할 수 있다. 신호가 운반하는 정보는 예를 들어 아날로그 또는 디지털 정보일 수 있다. 신호는 알려진 바와 같이 다양한 다른 유선 또는 무선 링크들을 통해 송신될 수 있다. 신호는 프로세서 판독가능한 매체 상에 저장될 수 있다.

[0143]

다수의 구현예가 설명되었다. 그럼에도 불구하고, 다양한 수정들이 이루어질 수 있음이 이해될 것이다. 예를 들어, 상이한 구현예들의 구성요소들은 결합, 보충, 수정 또는 제거되어 다른 구현예들을 생성할 수 있다. 추가로, 본 기술분야의 통상의 기술자는 다른 구조들 및 프로세스들이 개시된 것들을 대체할 수 있고, 결과적인 구현예들은 개시된 구현예들과 적어도 실질적으로 동일한 결과(들)를 달성하기 위해 적어도 실질적으로 동일한 방식(들)으로 적어도 실질적으로 동일한 기능(들)을 수행할 것임을 이해할 것이다. 따라서, 이들 및 다른 구현예들이 본 출원에 의해 고려된다.

도면

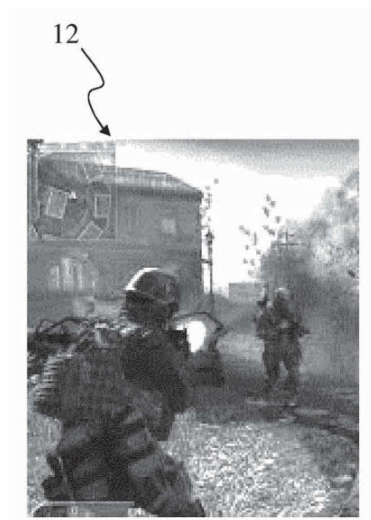
도면1a



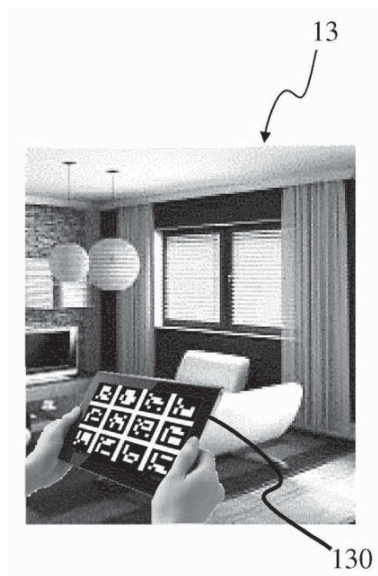
도면1b



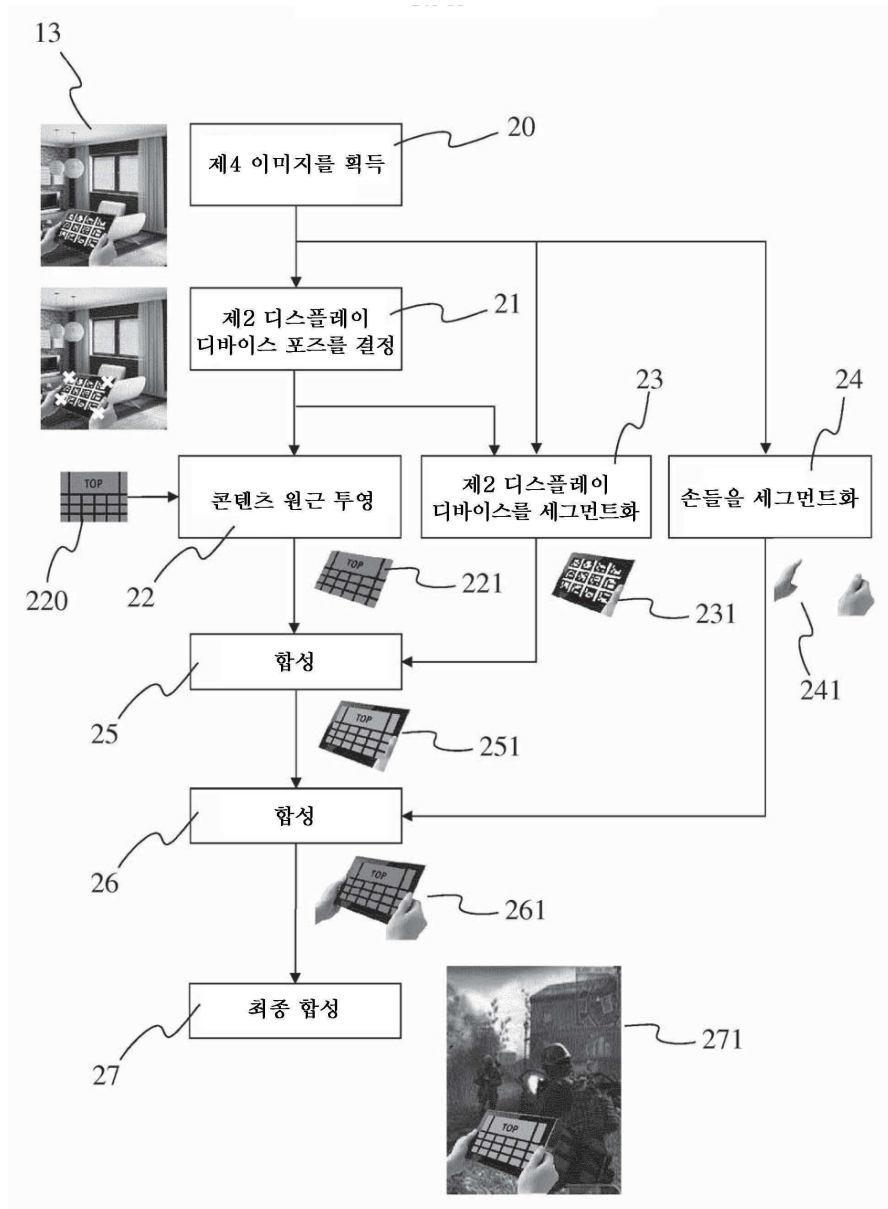
도면1c



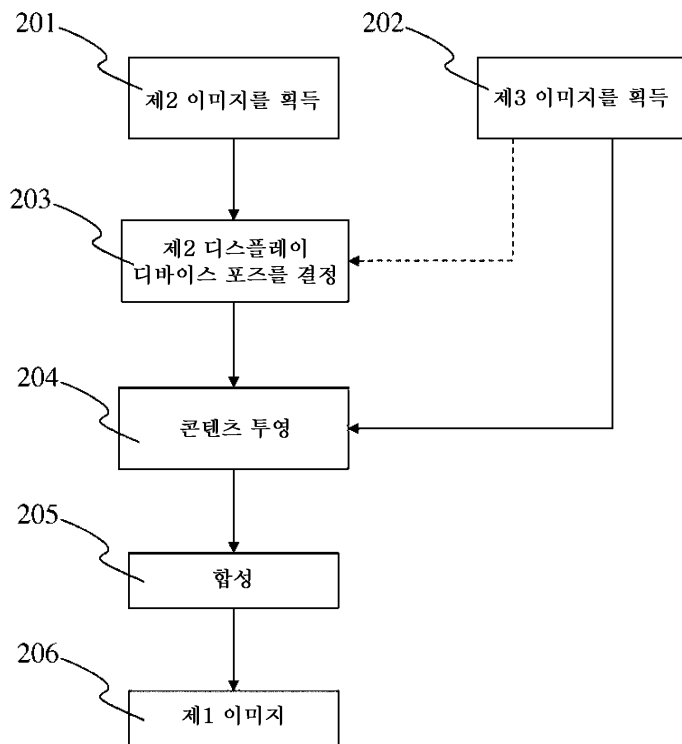
도면1d



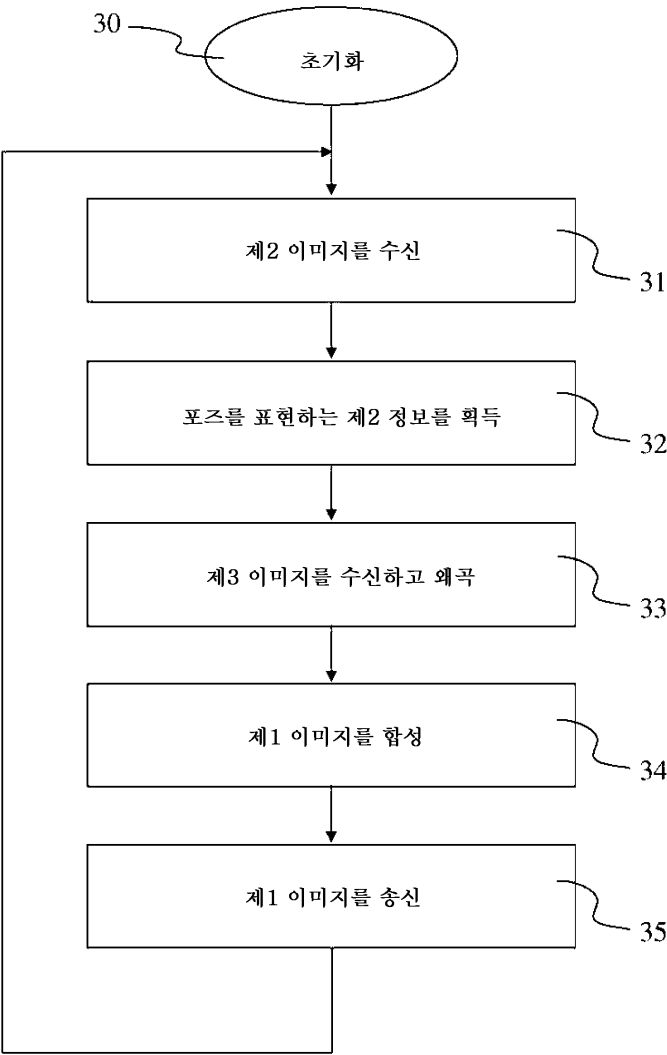
도면2a



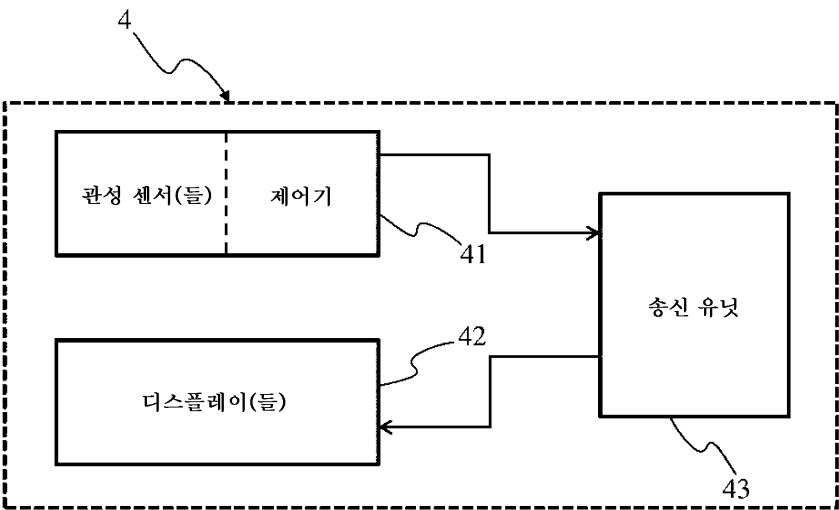
도면2b



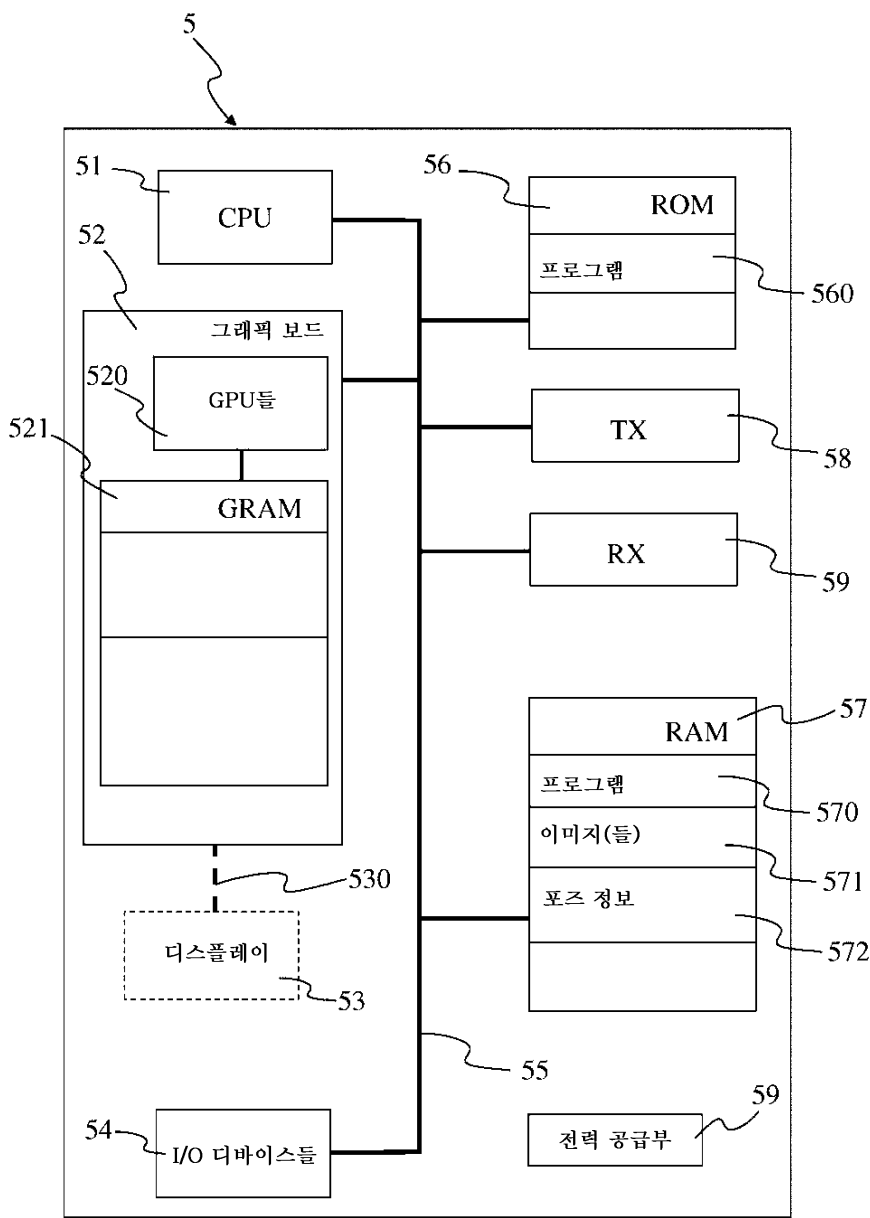
도면3



도면4



도면5



도면6

