

	(19) 대한민국특허청(KR) (12) 공개특허공보(A)	(11) 공개번호 10-2013-0124188 (43) 공개일자 2013년11월13일
(51) 국제특허분류(Int. Cl.) G06T 3/00 (2006.01) H04N 7/15 (2006.01)	(71) 출원인 커먼웰쓰 사이언티픽 앤드 인더스트리얼 리서치 오가니제이션	
(21) 출원번호 10-2013-0038522	오스트레일리아 오스트레일리안 캐피탈 테리토리 2612 캄벨 라임스톤 애비뉴	
(22) 출원일자 2013년04월09일 심사청구일자 2013년04월09일	(72) 발명자 루시, 사이먼	
(30) 우선권주장 2012901830 2012년05월04일 오스트레일리아(AU)	오스트레일리아 오스트레일리안 캐피탈 테리토리 2612 캄벨 라임스톤 애비뉴	
	(74) 대리인 특허법인 수	

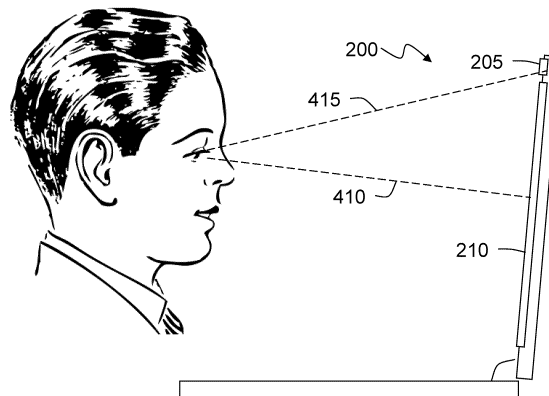
전체 청구항 수 : 총 26 항

(54) 발명의 명칭 비디오에서의 시선 정렬 시스템 및 방법

(57) 요약

본 발명의 이미지 조작 시스템은 화상 회의 방식을 개선할 수 있다. 이 시스템은 카메라, 상기 카메라에 인접한 디스플레이 화면, 상기 카메라 및 상기 디스플레이 화면에 연결된 프로세서, 및 상기 프로세서에 연결된 메모리를 포함한다. 프로세서에 의해 실행가능한 명령은 상기 카메라로부터 소스 이미지를 수신하고, 상기 소스 이미지에 기초하여 합성 이미지를 생성하는 것을 가능하게 한다. 상기 합성 이미지는 상기 디스플레이 화면에 배치된 가상 카메라의 뷰에 해당한다.

대표도 - 도4



특허청구의 범위

청구항 1

이미지 조작 시스템으로서,

카메라,

상기 카메라에 인접한 디스플레이 화면,

상기 카메라 및 상기 디스플레이 화면에 연결된 프로세서, 및

상기 프로세서에 연결되고, 상기 카메라로부터 소스 이미지를 수신하고, 상기 소스 이미지에 기초하여 상기 디스플레이 화면에 배치된 가상 카메라의 뷰(view)에 해당하는 합성 이미지를 생성하도록, 상기 프로세서에 의해 실행가능한 명령을 포함하는 메모리

를 포함하는 시스템.

청구항 2

제1항에 있어서,

상기 카메라 및 디스플레이 화면은 단일 구조로 통합된 시스템.

청구항 3

제1항에 있어서,

상기 카메라 및 디스플레이 화면은 독립적으로 이동가능하며, 상기 메모리는 상기 카메라에 대한 상기 디스플레이 화면의 위치를 추정하는 명령을 더 포함하는 시스템.

청구항 4

제1항 내지 제3항 중 어느 한 항에 있어서,

상기 합성 이미지를 생성하는 것은,

상기 소스 이미지에서 얼굴을 검출하고,

상기 가상 카메라의 위치에 따라 상기 얼굴을 변환하는 것을 더 포함하는 시스템.

청구항 5

제4항에 있어서,

상기 얼굴을 변환하는 것은 상기 얼굴에 능동적 외양 모델(Active Appearance Model: AAM)을 적용하고, 상기 AAM의 파라미터를 수정하는 것을 포함하는 시스템.

청구항 6

제4항 또는 제5항에 있어서,

상기 합성 이미지를 생성하는 것은,

상기 얼굴의 3D 모델을 생성하고,

상기 3D 모델을 상기 카메라와 상기 가상 카메라 간의 변위에 따라 회전시키는 것을 더 포함하는 시스템.

청구항 7

제1항 내지 제6항 중 어느 한 항에 있어서,

상기 가상 카메라의 위치는 상기 디스플레이 화면에 대해 중심에 고정되는 시스템.

청구항 8

제1항 내지 제6항 중 어느 한 항에 있어서,
상기 가상 카메라의 위치는 상기 디스플레이 화면 상에 제시된 콘텐츠에 기초하여 결정되는 시스템.

청구항 9

제1항 내지 제8항 중 어느 한 항에 있어서,
상기 이미지는 비디오 시퀀스의 이미지를 포함하는 시스템.

청구항 10

제9항에 있어서,
상기 비디오 시퀀스는 2차원 또는 3차원 비디오 시퀀스 중 하나인 시스템.

청구항 11

제1항 내지 제10항 중 어느 한 항에 있어서,
상기 메모리는,
상기 카메라로부터 제2 소스 이미지를 수신하고,
상기 소스 이미지 및 상기 제2 소스 이미지에 기초하여 상기 가상 카메라의 뷰에 해당하는 제2 합성 이미지를 생성하는 명령을 더 포함하는 시스템.

청구항 12

제1항 내지 제11항 중 어느 한 항에 있어서,
상기 합성 이미지를 생성하는 것은 상기 소스 이미지에 관한 메타데이터(metadata)를 수정하는 것을 포함하는 시스템.

청구항 13

제13항에 있어서,
상기 메타데이터는 카메라 파라미터를 포함하는 시스템.

청구항 14

화상 회의(video conferencing)의 컴퓨터 구현 방법으로서,
데이터 인터페이스 상에서, 카메라로부터 소스 이미지를 수신하는 단계, 및
프로세서에 의해, 상기 소스 이미지에 기초하여 상기 카메라에 인접한 디스플레이 화면에 위치하는 가상 카메라의 뷰(view)에 해당하는 합성 이미지를 생성하는 단계를 포함하는 방법.

청구항 15

제14항에 있어서,
상기 카메라에 대한 상기 디스플레이 화면의 위치를 추정하는 단계를 더 포함하는 방법.

청구항 16

제14항 또는 제15항에 있어서,
상기 합성 이미지를 생성하는 단계는,
상기 소스 이미지에서 얼굴을 검출하는 단계, 및

상기 가상 카메라의 위치에 따라 상기 얼굴을 변환하는 단계를 더 포함하는 방법.

청구항 17

제16항에 있어서,

상기 얼굴을 변환하는 단계는 상기 얼굴에 능동적 외양 모델(Active Appearance Model: AAM)을 적용하고, 상기 AAM의 파라미터를 수정하는 단계를 포함하는 방법.

청구항 18

제16항 또는 제17항에 있어서,

상기 합성 이미지를 생성하는 단계는,

상기 얼굴의 3D 모델을 생성하는 단계, 및

상기 3D 모델을 상기 카메라와 상기 가상 카메라 간의 변위에 따라 회전시키는 단계를 더 포함하는 방법.

청구항 19

제16항에 있어서,

상기 얼굴을 변환하는 단계는 상기 얼굴로부터 아바타 이미지로 얼굴 표현을 전이하는 단계를 포함하는 방법.

청구항 20

제14항 내지 제19항 중 어느 한 항에 있어서,

상기 가상 카메라의 위치는 상기 프로세서에 의해 상기 디스플레이 화면 상에 제시된 콘텐츠에 기초하여 결정되는 방법.

청구항 21

제14항 내지 제19항 중 어느 한 항에 있어서,

상기 가상 카메라의 위치는 미리 설정되는 방법.

청구항 22

제14항 내지 제21항 중 어느 한 항에 있어서,

데이터 인터페이스 상에서, 제1 원격 장치로부터 제1 원격 이미지를 수신하는 단계,

상기 제1 원격 이미지를 상기 디스플레이 화면 상에 디스플레이하는 단계, 및

상기 데이터 인터페이스 상에서, 상기 합성 이미지를 상기 제1 원격 장치로 송신하는 단계를 더 포함하는 방법.

청구항 23

제22항에 있어서,

상기 방법은,

상기 데이터 인터페이스 상에서, 제2 원격 장치로부터 제2 원격 이미지를 수신하는 단계, 및

상기 제1 원격 이미지에 인접한 상기 제2 원격 이미지를 상기 디스플레이 화면 상에 디스플레이하는 단계를 더 포함하는 방법.

청구항 24

제23항에 있어서,

상기 데이터 인터페이스 상에서, 상기 합성 이미지를 상기 제2 원격 장치로 송신하는 단계를 더 포함하는 방법.

청구항 25

제23항에 있어서,

상기 프로세서에 의해, 상기 소스 이미지에 기초하여 상기 디스플레이 화면에 배치된 제2 가상 카메라의 뷰에 해당하는 제2 합성 이미지를 생성하는 단계, 및

상기 데이터 인터페이스 상에서, 상기 제2 합성 이미지를 상기 제2 원격 장치로 송신하는 단계를 더 포함하는 방법.

청구항 26

제14항 내지 제25항 중 어느 한 항에 있어서,

상기 카메라로부터 추가 소스 이미지를 수신하는 단계, 및

상기 소스 이미지 및 상기 추가 소스 이미지에 기초하여 상기 가상 카메라의 뷰에 해당하는 추가 합성 이미지를 생성하는 단계를 더 포함하는 방법.

명세서

기술 분야

[0001] 본 발명은 이미지 조작(image manipulation)에 관한 것이다. 특히, 본 발명은, 비록 한정되지는 않지만 비디오에서의 시선 정렬(eye alignment)에 관한 것이다.

배경 기술

[0002] 비언어적인 사회적 단서(cues)는 의사 소통에서 중요한 역할을 한다. 눈은 특히 의사 소통에 중요하고 주의를 끈 사람들에게 암시(indication)를 제공한다. 어떤 문화에서는, 지배적인 사람의 눈을 바라보지 않는 것이 존경을 표하는 것이고, 서양 문화에서 사람들이 서로 눈을 마주치지 않으면 나쁘게 판단할 수 있다.

[0003] 그러나, 비언어적인 단서는 종종 종래 기술의 화상 회의(video conferencing) 시스템에서 왜곡되어, 오해 또는 의사 불통으로 이어진다.

[0004] 가장 단순한 형태로서의 화상 회의 시스템은 한 쌍의 카메라 및 한 쌍의 화면을 포함한다. 한 쌍의 카메라 및 한 쌍의 디스플레이 화면 각각은 사용자들 중 한 사용자를 향한다. 이러한 카메라는 화면을 막지 않도록 하기 위해 디스플레이 화면 옆이나 그 위에 배치되고, 화상 회의 장비를 갖춘 최신 랩탑은 종종 디스플레이 화면 위에 고정적으로 위치한 내장형 카메라를 포함한다.

[0005] 이러한 시스템의 단점은 사용자가 사용자의 위치와 다른 위치에서 기록되면서 디스플레이 화면 상에서 다른 사용자의 이미지를 보기 때문에 눈동자 단서(eye cue)가 왜곡된다는 것이다. 따라서, 사용자들은 직접 눈을 마주치는 것으로 보이지 않는다. 그 대신, 사용자는 실제로 그렇게 하지 않을 경우 눈을 마주치는 것으로 보일 수 있다.

[0006] 도 1의 화상 회의 시스템(100)에 예시된 바와 같이, 이러한 문제점을 줄인 특수화된 하드웨어 시스템이 개발되었다. 시스템(100)은 디스플레이 화면(105) 및 카메라(110)를 포함한다. 사용자(120)와 카메라(110) 사이에 빔 스플리터(115)가 배치되고 사용자(120)로부터의 빛이 카메라에 수신되게 하면서 디스플레이 화면(105)으로부터 이미지를 반사하도록 구성된다.

[0007] 시스템(100)과 같은 시스템의 단점은 가격이 비싸고, 부서지기 쉬우며, 부피가 커서 얇은 랩탑 화면에 통합될 수 없고, 사용자(120)의 초점이 디스플레이 화면(105)의 중심이 아닌 부분에 맞춰질 때에는 잘 작동하지 못한다는 것이다.

[0008] 다른 유사한 형태의 화상 회의 시스템은 반투명성 화면을 포함하며 카메라가 반투명성 화면 뒤에 배치되지만, 이 또한 마찬가지로 단점이 존재한다.

발명의 내용

해결하려는 과제

[0009] 본 발명의 실시예의 목적은 전술한 소비자에게 종래 기술에 비해 개선된 점 및 이점을 제공하고, 및/또는 전술

한 종래 기술의 단점 중 하나 이상을 극복 및 경감하며, 및/또는 유용한 상업적 선택을 제공하는데 있다.

과제의 해결 수단

- [0010] 본 발명의 일 양태에 따르면, 이미지 조작 시스템으로서, 카메라, 상기 카메라에 인접한 디스플레이 화면, 상기 카메라 및 상기 디스플레이 화면에 연결된 프로세서, 및 상기 프로세서에 연결되고, 상기 카메라로부터 소스 이미지를 수신하고, 상기 소스 이미지에 기초하여 상기 디스플레이 화면에 배치된 가상 카메라의 뷰(view)에 해당하는 합성 이미지를 생성하도록, 상기 프로세서에 실행가능한 명령을 포함하는 메모리를 포함한다.
- [0011] 바람직하게는, 상기 카메라 및 디스플레이 화면은 단일 구조로 통합된다. 또한, 상기 카메라 및 디스플레이 화면은 독립적으로 이동가능하며, 상기 메모리는 상기 카메라에 대한 상기 디스플레이 화면의 위치를 추정하는 명령을 더 포함한다.
- [0012] 바람직하게는, 상기 합성 이미지를 생성하는 것은, 상기 소스 이미지에서 얼굴을 검출하고, 상기 가상 카메라의 위치에 따라 상기 얼굴을 변환하는 것을 더 포함한다.
- [0013] 바람직하게는, 상기 얼굴을 변환하는 것은 상기 얼굴에 능동적 외양 모델(Active Appearance Model: AAM)을 적용하고, 상기 AAM의 파라미터를 수정하는 것을 포함한다.
- [0014] 바람직하게는, 상기 합성 이미지를 생성하는 것은, 상기 얼굴의 3D 모델을 생성하고, 상기 3D 모델을 상기 카메라와 상기 가상 카메라 간의 변위에 따라 회전시키는 것을 포함한다.
- [0015] 바람직하게는, 상기 가상 카메라의 위치는 상기 디스플레이 화면에 대해 중심에 고정된다. 또한, 상기 가상 카메라의 위치는 상기 디스플레이 화면 상에 제시된 콘텐츠에 기초하여 결정된다.
- [0016] 바람직하게는, 상기 이미지는 비디오 시퀀스의 이미지를 포함한다. 상기 비디오 시퀀스는 2차원 또는 3차원 비디오 시퀀스일 수 있다.
- [0017] 바람직하게는, 상기 메모리는, 상기 카메라로부터 제2 소스 이미지를 수신하고, 상기 소스 이미지 및 상기 제2 소스 이미지에 기초하여 상기 가상 카메라의 뷰에 해당하는 제2 합성 이미지를 생성하는 명령을 더 포함한다.
- [0018] 소정의 실시예에 따르면, 상기 합성 이미지를 생성하는 것은 상기 소스 이미지에 관한 메타데이터(metadata)를 수정하는 것을 포함한다. 바람직하게는, 상기 메타데이터는 카메라 파라미터를 포함한다.
- [0019] 본 발명의 제2 양태에 따르면, 화상 회의(video conferencing)의 컴퓨터 구현 방법으로서, 데이터 인터페이스 상에서, 카메라로부터 소스 이미지를 수신하는 단계, 및 프로세서에 의해, 상기 소스 이미지에 기초하여 상기 카메라에 인접한 디스플레이 화면에 위치하는 가상 카메라의 뷰(view)에 해당하는 합성 이미지를 생성하는 단계를 포함한다.
- [0020] 소정의 실시예에 따르면, 상기 방법은 상기 프로세서에 의해 상기 카메라에 대한 상기 디스플레이 화면의 위치를 추정하는 단계를 더 포함한다.
- [0021] 바람직하게는, 상기 합성 이미지를 생성하는 단계는, 상기 소스 이미지에서 얼굴을 검출하는 단계, 및 상기 가상 카메라의 위치에 따라 상기 얼굴을 변환하는 단계를 더 포함한다.
- [0022] 바람직하게는, 상기 얼굴을 변환하는 단계는 상기 얼굴에 능동적 외양 모델(Active Appearance Model: AAM)을 적용하고, 상기 AAM의 파라미터를 수정하는 단계를 포함한다.
- [0023] 또한, 상기 얼굴을 변환하는 단계는 상기 얼굴로부터 아바타 이미지로 얼굴 표현을 전이하는 단계를 포함한다.
- [0024] 바람직하게는, 상기 합성 이미지를 생성하는 단계는, 상기 얼굴의 3D 모델을 생성하는 단계, 및 상기 3D 모델을 상기 카메라와 상기 가상 카메라 간의 변위에 따라 회전시키는 단계를 더 포함한다.
- [0025] 소정의 실시예에 따르면, 상기 가상 카메라의 위치는 상기 프로세서에 의해 상기 디스플레이 화면 상에 제시된 콘텐츠에 기초하여 결정된다. 또한, 상기 가상 카메라의 위치는 미리 설정된다.
- [0026] 바람직하게는, 상기 이미지는 비디오 시퀀스의 이미지를 포함한다. 상기 비디오 시퀀스는 2차원 또는 3차원 비디오 시퀀스일 수 있다.
- [0027] 바람직하게는, 상기 방법은, 데이터 인터페이스 상에서, 제1 원격 장치로부터 제1 원격 이미지를 수신하는

단계, 상기 제1 원격 이미지를 상기 디스플레이 화면 상에 디스플레이하는 단계, 및 상기 데이터 인터페이스 상에서, 상기 합성 이미지를 상기 제1 원격 장치로 송신하는 단계를 더 포함한다.

[0028] 바람직하게는, 상기 방법은, 상기 데이터 인터페이스 상에서, 제2 원격 장치로부터 제2 원격 이미지를 수신하는 단계, 및 상기 제1 원격 이미지에 인접한 상기 제2 원격 이미지를 상기 디스플레이 화면 상에 디스플레이하는 단계를 더 포함한다.

[0029] 바람직하게는, 상기 방법은, 상기 데이터 인터페이스에서, 상기 합성 이미지를 상기 제2 원격 장치로 송신하는 단계를 더 포함한다. 또한, 상기 방법은, 상기 프로세서에 의해, 상기 소스 이미지에 기초하여 상기 디스플레이 화면에 배치된 제2 가상 카메라의 뷰에 해당하는 제2 합성 이미지를 생성하는 단계, 및 상기 데이터 인터페이스 상에서, 상기 제2 합성 이미지를 상기 제2 원격 장치로 송신하는 단계를 더 포함한다.

[0030] 바람직하게는, 상기 방법은, 상기 카메라로부터 추가 소스 이미지를 수신하는 단계, 및 상기 소스 이미지 및 상기 추가 소스 이미지에 기초하여 상기 가상 카메라의 뷰에 해당하는 추가 합성 이미지를 생성하는 단계를 더 포함한다.

[0031] 소정의 실시예에 따르면, 상기 합성 이미지를 생성하는 단계는 상기 소스 이미지에 관한 메타데이터를 수정하는 단계를 포함한다. 바람직하게는, 상기 메타데이터는 카메라 파라미터를 포함한다.

발명의 효과

[0032] 본 발명은 전술한 소비자에게 종래 기술에 비해 개선된 점 및 이점을 제공하고, 및/또는 전술한 종래 기술의 단점 중 하나 이상을 극복 및 경감하며, 및/또는 유용한 상업적 선택을 제공할 수 있다.

도면의 간단한 설명

[0033] 본 발명의 이해를 돕고 당업자가 본 발명을 실시가능하게 하기 위해, 첨부된 도면을 참조하여 본 발명의 바람직한 실시예가 단지 예시로서 이하에서 설명된다.

도 1은 종래 기술의 화상 회의 시스템을 예시하는 도면이다.

도 2는 본 발명의 일 실시예에 따른 비디오 조작 시스템을 예시하는 도면이다.

도 3은 도 2의 시스템의 블록도를 예시하는 도면이다.

도 4는 도 2의 시스템을 예시하는 도면이다.

도 5는 본 발명의 대안의 실시예에 따른 비디오 조작 시스템을 예시하는 도면이다.

도 6은 본 발명의 일 실시예에 따른 화상 회의 시스템을 예시하는 도면이다.

도 7은 본 발명의 일 실시예에 따른 비디오 조작 방법을 예시하는 도면이다.

도 8은 본 발명의 일 실시예에 따른 컴퓨팅 장치를 예시하는 도면이다.

당업자는 도면에 예시된 바와 같은 구성 요소의 배치를 약간 변경하는 것만으로 본 발명의 개시된 실시예의 고유한 기능이 손상되지 않는 것을 인식할 것이다.

발명을 실시하기 위한 구체적인 내용

[0034] 본 발명의 실시예는 시선 정렬 시스템 및 방법을 포함한다. 본 발명의 구성 요소는 본 발명의 실시예를 이해하는데 필요한 정도의 특정 상세 내용만을 보여주지만, 본 설명에 비추어 당업자에게 자명한 사항까지 지나치게 상세히 설명하여 본 내용이 불명확하게 되지 않도록 하기 위해, 도면에는 간결한 개요 형태로 예시된다.

[0035] 본 명세서에서, 제1 및 제2, 좌측 및 우측, 전방 및 후방, 상부 및 하부 등과 같은 용어는 그 용어에 의해 기술된 특정 상대 위치 또는 순서를 반드시 필요로 하는 것은 아니고, 단지 다른 구성 요소 또는 방법 단계로부터 하나의 구성 요소 또는 방법 단계를 정의하는데 사용된다. "포함한다" 또는 "구비한다"와 같은 단어는 구성 요소 또는 방법 단계의 배타적인 세트(set)를 정의하는데 사용되지 않는다. 그보다, 상기 단어는 단지 본 발명의 특정 실시예에 포함된 최소한의 구성 요소 또는 방법 단계의 세트만을 정의한다.

[0036] 본 명세서에서의 종래 기술이라는 언급은 그러한 종래 기술이 통상의 일반적인 지식의 일부를 구성한다는 것을 인정하거나 또는 암시하는 어떤 형태가 아니고, 또 그러한 것으로 간주되지 않아야 한다.

- [0037] 일 태양에 따르면, 본 발명은 이미지 조작 시스템에 관한 것으로서, 상기 시스템은 카메라, 상기 카메라에 인접한 디스플레이 화면, 상기 카메라 및 상기 디스플레이 화면에 연결된 프로세서, 및 상기 프로세서에 연결되고, 상기 프로세서에 의해 상기 카메라로부터 소스 이미지를 수신하고, 및 상기 소스 이미지에 기초하여 합성 이미지를 생성하도록 실행가능한 명령을 포함하는 메모리를 포함하고, 상기 합성 이미지는 상기 디스플레이 화면에 배치된 가상 카메라의 뷰(view)에 해당한다.
- [0038] 본 발명의 특정 실시예의 이점은 화상 회의 방식이 개선되고, 비디오 또는 이미지 기반 통신이 개선되며, 비디오 캡처 또는 회의 시스템이 더 간단하고 저렴해진다는 것이다. 또한, 본 발명은 레거시(legacy) 화상 회의 시스템에 적용될 수도 있다.
- [0039] 도 2는 본 발명의 일 실시예에 따른 비디오 조작 시스템(200)을 예시하는 도면이고, 도 3은 시스템(200)의 블록도를 예시하는 도면이다.
- [0040] 시스템(200)은 카메라(205) 및 디스플레이 화면(210)을 포함한다. 시스템(200)은 이미지를 디스플레이하고, 카메라(205)는 비디오를 캡처한다. 시스템(200)은, 예를 들면, 화상 회의에 사용될 수 있으며, 여기서 제1 사용자는 카메라(205)를 이용하여 레코딩되고 제2 사용자는 디스플레이 화면(210)을 이용하여 디스플레이된다. 그 대신, 시스템(200)은 뉴스 데스크 스타일의 레코딩에 사용될 수 있으며, 이 경우 사용자는 카메라(205)를 이용하여 레코딩되면서 디스플레이 화면(210) 상에 디스플레이된 스크립트(script)를 판독한다.
- [0041] 시스템(200)은 카메라(205) 및 디스플레이 화면(210)과 연결된 프로세서(305)를 더 포함한다. 프로세서(305)는 카메라(205)에 의해 레코딩된 비디오를 조작하는 명령 코드(315)를 포함하는 메모리(310)와도 연결된다.
- [0042] 명령 코드(315)는 카메라(205)로부터 소스 이미지를 수신하는 명령, 및 소스 이미지에 기초하여 합성 이미지를 생성하는 명령을 포함한다. 소스 이미지 및 합성 이미지는 비디오 시퀀스를 구성한다.
- [0043] 도 4에 도시된 바와 같이, 카메라(205)는 디스플레이 화면(210) 상부에 배치된다. 광 경로(410)와 같이, 사용자가 디스플레이 화면(210)을 똑바로 쳐다볼 때, 광 경로(415)를 통해 보여지는 카메라(205)에 의해 캡처된 소스 이미지는 카메라(205)와 디스플레이 화면(210) 간의 위치 차로 인해 사용자가 내려다보는 모습을 보여준다.
- [0044] 합성 이미지는 디스플레이 화면(210)에 위치하는 가상 카메라에 해당한다. 가상 카메라 위치는 디스플레이 화면(210)의 중심에 있을 수 있거나, 또는 애플리케이션 윈도우 위치와 같은 디스플레이 화면(210)의 특정 부분, 또는 다른 관심 영역에 있을 수 있다.
- [0045] 도 5는 가상 카메라 위치가 디스플레이 화면에 대해 중심에 오지 않는 시스템(500)을 예시하는 도면이다. 시스템(500)은 화상 회의 윈도우(505) 및 프리젠테이션(presentation) 공유 윈도우(510)를 포함한다. 시스템(500)은 발표자의 비디오와 함께 디지털 프리젠테이션을 볼 수 있도록 한다.
- [0046] 이 경우, 카메라(205) 및 가상 카메라는 가상 카메라 오프셋(515)으로 예시된 바와 같이 수직 및 수평 방향의 두가지 방향으로 오프셋(offset)된다.
- [0047] 소정의 실시예에 따르면, 카메라(205)와 디스플레이 화면(210) 간의 오프셋, 및 가상 카메라와 카메라(205) 간의 오프셋은 알려져 있다. 이는 카메라 및 디스플레이 화면이 단일 셸(shell) 또는 프레임으로 만들어지는 장치 또는 랩탑을 형성하는 목적에 관한 경우일 수 있다.
- [0048] 이러한 관계에 따라 가상 카메라 위치가 설정될 수 있다.
- [0049] 또 다른 실시예(미도시)에 따르면, 카메라(205)와 디스플레이 화면(210) 간의 오프셋은 가변적이다. 이는, 예를 들면, 카메라(205)가 플렉시블 케이블을 통해 PC에 연결된 웹 카메라를 포함하는 개인용 컴퓨터(PC) 설정시의 경우이다. 이 경우, 명령 코드(315)는 카메라에 대한 디스플레이 화면의 위치를 추정하는 명령을 포함할 수 있다. 또한, 명령 코드(315)는 사용자 또는 다른 소스로부터 카메라 위치를 수신하는 명령을 포함할 수도 있다.
- [0050] 다음으로 카메라(205)의 가상 카메라 오프셋에 따라 합성 이미지가 생성된다. 합성 이미지는 카메라(205) 및 가상 카메라의 위치를 이용하여, 폐색(occlusion) 처리 알고리즘과 같은 관련 알고리즘과 함께, 기하학적 뷰 합성과 같은 종래 기술의 뷰 합성 방법에 따라 생성될 수 있다.
- [0051] 그 대신, 뷰 합성은 인간의 얼굴에 특정될 수 있다. 이 경우, 모델이 적용되는 얼굴은 소스 이미지에서 검출된

다. 다음에, 얼굴은 카메라(205)의 가상 카메라 오프셋에 따라 변환된다.

- [0052] 얼굴의 3D 모델은 카메라(205)의 가상 카메라 오프셋에 따라 생성되고 회전될 수 있다. 모델은 능동적 외양 모델(Active Appearance Model: AAM)일 수 있으며, 여기서 AAM의 파라미터는 카메라(205)의 가상 카메라 오프셋에 따라 수정된다. 그리고, 이미지 데이터는 변환된 AAM로 렌더링된다.
- [0053] 또 다른 실시예에 따르면, 뷰 합성은 아바타에 기반할 수 있다. 사용자의 얼굴 표현은 아바타 이미지로 전이될 수 있으며, 아바타 이미지는 카메라(205)의 가상 카메라 오프셋에 따라 회전된다.
- [0054] 아바타는 사용자의 얼굴 이미지에 해당하는 아바타, 또는 어떤 다른 형태의 아바타일 수 있다. 사용자 및/또는 아바타의 모델은 PCT/AU2012/000295에 기술된 바와 같이 사용자의 이미지에 따라 생성될 수 있다.
- [0055] 소정의 실시예에 따르면, 이러한 변환은 사용자가 카메라(205)에서 기 설정된 거리보다 적게 떨어져 있는 경우에만 적용된다. 사용자가 카메라(205)로부터 멀리 떨어진 때, 카메라(205)와 가상 카메라 간의 각도는 줄어든다.
- [0056] 전술한 바와 같이, 가장 간단한 형태로서의 본 방법은 단일 이미지에 적용된다. 그러나, 소정의 실시예에 따르면, 본 방법은 아래에서 더 설명되는 바와 같은 복수의 이미지에도 동시에 적용될 수 있다.
- [0057] 이러한 이미지는 비디오 시퀀스를 포함할 수 있으며, 비디오 시퀀스는 2차원 또는 3차원 비디오 시퀀스일 수 있다. 3차원 비디오의 경우, 단일 시간 인스턴스 - 즉 뷰 - 로부터의 모든 이미지가 동시에 처리된다. 따라서, 가상 카메라를 향해 이미지를 이동시키고 3차원 데이터를 유지하는 변환이 선택될 수 있다. 멀티 뷰 3차원 비디오의 경우, 중앙 뷰보다 디스플레이 화면(210)에 더 밀접하게 대응하는 현재의 뷰(existing view)를 선택하는 것이 가능할 수 있다.
- [0058] 3차원 비디오의 경우, 합성 이미지를 생성하는 것은 소스 이미지의 카메라 파라미터를 수정하는 것을 포함할 수 있다.
- [0059] 2차원 비디오의 경우, 이전 이미지가 소스 이미지와 함께 이용되어 합성 이미지를 생성할 수 있다. 이전 이미지는 합성 이미지의 폐색(occluded) 영역에 관한 데이터를 제공하거나, 또는 합성 이미지의 품질을 향상시키는 데 사용될 수 있다.
- [0060] 도 6은 본 발명의 일 실시예에 따른 화상 회의 시스템(600)을 예시하는 도면이다.
- [0061] 화상 회의 시스템(600)은 제1 컴퓨팅 장치(615a)에 연결되어 제1 위치에서 화상 회의 시스템의 제1 부분을 함께 구성하는 제1 카메라(605a) 및 제1 디스플레이 화면(610a)을 포함한다. 화상 회의 시스템(600)은 제2 컴퓨팅 장치(615b)에 연결되어 제2 위치에서 화상 회의 시스템(600)의 제2 부분을 함께 구성하는 제2 카메라(605b) 및 제2 디스플레이 화면(610b)을 더 포함한다.
- [0062] 제1 및 제2 컴퓨팅 장치는 인터넷과 같은 통신 네트워크(620)를 통해 연결된다.
- [0063] 일련의 이미지가 각각 제1 및 제2 카메라(605a, 605b)에 의해 캡처되고, 제1 및 제2 컴퓨팅 장치(615a, 615b)에 의해 제1 카메라(605a) 및 제1 디스플레이 화면(610a)과, 제2 카메라(605b) 및 제2 디스플레이 화면(610b) 간의 거리에 따라 각각 변환된다. 그리고, 변환된 이미지는 디스플레이 화면(610b, 610a) 상에 디스플레이하기 위해 컴퓨팅 장치(615a, 615b)에 의해 전송된다. 이러한 이미지는, 당업자가 쉽게 이해하는 바와 같이, 전송 전에 압축될 수 있다.
- [0064] 두 카메라(605a, 605b)에 의해 캡처된 이미지를 전송함으로써, 서로 다른 위치에 있는 사용자들이 이들 각각의 디스플레이 화면(610a, 610b)을 통해 서로 눈으로 볼 수 있으며, 반면에 카메라(605a, 605b)는 디스플레이 화면(610a, 610b)에서 멀리 떨어져 배치된다.
- [0065] 소정의 실시예(미도시)에 따르면, 화상 회의 시스템(600)은 적어도 제3 위치에서 제3 컴퓨팅 장치를 포함하는 다자간 화상 회의인 경우를 더 포함한다. 각 디스플레이 화면(610a, 610b)은 서로 인접한 화상 회의 시스템(600)의 다른 사용자들의 이미지를 디스플레이한다.
- [0066] 제1 컴퓨팅 장치(615a)는 변환된 이미지를 제2 컴퓨팅 장치(615b) 및 제3 컴퓨팅 장치 양쪽으로 전송할 수 있다. 그 대신, 제1 컴퓨팅 장치는 변환된 이미지를 제2 컴퓨팅 장치(615b)로 송신하고, 제2 그룹의 변환된 이미지를 생성하여 제3 컴퓨팅 장치로 송신할 수 있다. 제2 그룹의 변환된 이미지는 디스플레이 화면에 배치된 제2 가상 카메라의 뷰에 해당할 수 있다.

- [0067] 이러한 점에서, 본 발명은 다수의 참가자 간의 화상 회의를 가능하게 하며 각 참가자는 자신들을 위해 특별하게 구성되는 비디오를 수신한다.
- [0068] 도 7은 본 발명의 일 실시예에 따른 비디오 조작 방법을 예시하는 도면이다.
- [0069] 단계 705에서, 데이터 인터페이스에서 소스 이미지를 카메라로부터 수신한다. 소스 이미지는 비디오 시퀀스의 이미지를 포함할 수 있다.
- [0070] 단계 710에서, 카메라에 대한 디스플레이 화면의 위치를 추정한다. 당업자가 용이하게 이해할 수 있는 바와 같이, 단일 장치 또는 프레임에 화면 및 디스플레이 카메라가 고정적으로 장착될 수 있고, 카메라와 디스플레이 화면 간의 상대적인 위치는 이미 인식되므로, 본 단계는 선택적이다. 디스플레이 화면 상에서 가상 카메라 위치가 결정된다.
- [0071] 단계 715에서, 소스 이미지에서 얼굴을 검출한다.
- [0072] 단계 720에서, 가상 카메라 위치에 따라 얼굴을 변환한다. 이러한 변환은 3차원 모델을 얼굴에 적용하고, 3차원 모델을 회전시키는 것을 포함할 수 있다.
- [0073] 단계 725에서, 소스 이미지 및 카메라와 가상 카메라 간의 변위에 따라 합성 이미지를 생성한다.
- [0074] 당업자가 쉽게 이해하는 바와 같이, 단계 715 및 단계 720는 선택적이다. 소스 이미지 전체는 얼굴이 이미지의 다른 부분과 같은 방식으로 변환되는 기하학적 뷰 합성을 이용하여 변환될 수 있다.
- [0075] 도 8은 본 발명의 일 실시예에 따른 컴퓨팅 장치(800)를 도시적으로 설명하는 도면이다. 도 2의 시스템(200) 및 도 6의 컴퓨팅 장치(615a, 615b)는 컴퓨팅 장치(800)와 동일하거나 유사할 수 있다. 마찬가지로, 도 7의 방법(700)은 컴퓨팅 장치(800)를 이용하여 구현될 수 있다.
- [0076] 컴퓨팅 장치(800)는 중앙 프로세서(802), 시스템 메모리(804) 및 시스템 메모리(804)를 중앙 프로세서(802)에 연결하는 것을 포함하여 여러 시스템 구성 요소 간을 연결하는 시스템 버스(806)를 포함한다. 시스템 버스(806)는 각종 버스 구조 중 어떤 구조라도 이용할 수 있으며, 메모리 버스 또는 메모리 제어기, 주변 버스, 및 로컬 버스를 포함하는 여러 형태의 버스 구조 중 어떤 것이라도 될 수 있다. 시스템 메모리(804)의 구조는 당업자에게 공지되어 있으며, 판독 전용 메모리(ROM)에 저장된 기본 입/출력 시스템(BIOS), 및 랜덤 액세스 메모리(RAM)에 저장된 운영 체제, 응용 프로그램과 프로그램 데이터와 같은 하나 이상의 프로그램 모듈을 포함할 수 있다.
- [0077] 또한, 컴퓨팅 장치(800)는 데이터를 판독 및 레코딩하는 다양한 인터페이스 수단 및 드라이브를 포함할 수 있다. 전술한 바와 같이, 데이터는 예를 들어 카메라와 디스플레이 화면 간의 변위를 포함할 수 있다.
- [0078] 특히, 컴퓨팅 장치(800)는 각각이 하드 디스크 드라이브(812) 및 제거가능한(removable) 메모리 드라이브(814)를 시스템 버스(806)에 연결하는 하드 디스크 인터페이스(808) 및 제거가능한 메모리 인터페이스(810)를 포함한다. 제거가능한 메모리 드라이브(814)의 예는 자기 디스크 드라이브 및 광학 디스크 드라이브를 포함한다. 이러한 드라이브 및 디지털 다기능 디스크(DVD)(816)와 같이 드라이브와 연관된 컴퓨터 판독가능한 매체는 컴퓨터 시스템(800)에 필요한 컴퓨터 판독가능한 명령, 데이터 구조, 프로그램 모듈 및 다른 데이터의 비휘발성 저장을 제공한다. 하나의 하드 디스크 드라이브(812) 및 하나의 제거가능한 메모리 드라이브(814)는 단지 설명의 목적으로 도시되며, 컴퓨팅 장치(800)가 여러 유사 드라이브를 포함할 수 있다는 점에 대한 이해가 수반된다. 또한, 컴퓨팅 장치(800)는 여러 형태의 컴퓨터 판독가능한 매체에 접속하는 드라이브를 포함할 수 있다.
- [0079] 컴퓨팅 장치(800)는 장치들을 시스템 버스(806)에 연결하는 추가 인터페이스를 포함할 수 있다. 도 8은 하나의 장치를 시스템 버스(806)에 연결하는데 사용될 수 있는 범용 직렬 버스(USB) 인터페이스(818)를 도시한다. 예를 들어, 추가 장치를 컴퓨팅 장치(800)에 연결하는데 IEEE 1394 인터페이스(820)가 사용될 수 있다. 추가 장치의 예는 도 1의 훈련(training) 이미지와 같은 이미지 또는 비디오를 수신하는 카메라를 포함한다.
- [0080] 컴퓨팅 장치(800)는 하나 이상의 원격 컴퓨터 또는 서버, 라우터, 네트워크 개인용 컴퓨터, 피어(peer) 장치 또는 다른 공통 네트워크 노드, 무선 전화기 또는 무선 개인 휴대정보 단말기(PDA, personal digital assistant)와 같은 다른 장치와의 로직 연결을 이용하여 네트워크(networked) 환경에서 동작할 수 있다. 컴퓨팅 장치(800)는 시스템 버스(806)를 근거리 네트워크(LAN)(824)에 연결하는 네트워크 인터페이스(822)를 포함한다. 네트워킹 환경은 사무실, 전사적(enterprise-wide) 컴퓨터 네트워크 및 가정용 컴퓨터 시스템에서 일반적이다.
- [0081] 또한, 인터넷과 같은 광역 네트워크(WAN)는, 예를 들어, 직렬 포트 인터페이스(826)에 연결된 모뎀 유닛을 통해

또는 LAN(824)을 통해 컴퓨팅 장치에 의해 액세스될 수 있다.

[0082] 화상 회의는 LAN(824), WAN, 또는 이들의 조합을 이용하여 수행될 수 있다.

[0083] 여기에서 도시되고 설명되는 네트워크 연결은 예시에 불과하고, 컴퓨터 간에 통신 링크를 구축하는 다른 방식이 사용될 수 있음은 자명한 것이다. TCP/IP, 프레임 릴레이, 이더넷, FTP, 및 HTTP 등과 같은 다양한 공지의 프로토콜 중 어떤 것도 이용될 수 있으며, 컴퓨팅 장치는 사용자로 하여금, 예를 들어, 웹 기반 서버로부터 데이터를 검색하게 하는 클라이언트-서버 구성에서 동작할 수 있다.

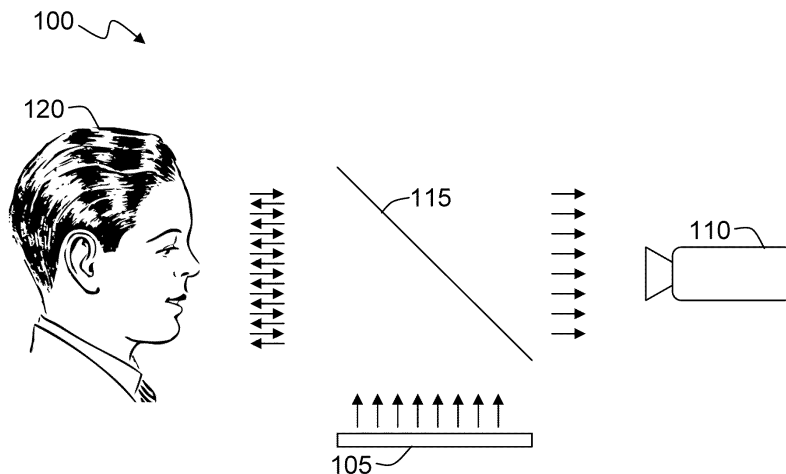
[0084] 컴퓨팅 장치의 동작은 각종 다른 프로그램 모듈에 의해 제어될 수 있다. 프로그램 모듈의 예는 특정 작업을 수행하거나 또는 특정한 추상적인 데이터 형태를 실행하는 루틴, 프로그램, 객체, 컴포넌트, 및 데이터 구조이다. 또한, 본 발명은 핸드헬드(hand-held) 장치, 다중프로세서 시스템, 마이크로프로세서 기반 또는 프로그램가능 고객 전자장치, 네트워크 PC, 미니컴퓨터, 메인프레임 컴퓨터, 및 개인 휴대정보 단말기 등을 포함하여 다른 컴퓨터 시스템 구성으로도 실시될 수 있다. 더욱이, 본 발명은 통신 네트워크를 통해 링크된 원격 처리 장치에 의해 작업을 수행하는 분산형 컴퓨팅 환경에서도 실시될 수 있다. 분산형 컴퓨팅 환경에서, 프로그램 모듈은 로컬 및 원격 메모리 저장 장치 양 쪽에 배치될 수 있다.

[0085] 요약하면, 본 발명의 일 실시예의 이점은 화상 회의 방식이 개선되고, 비디오 또는 이미지 기반 통신이 개선되며, 비디오 캡처 또는 회의 시스템이 더 간단하고 저렴하다는 것이다. 또한, 본 발명은 레거시(legacy) 화상 회의 시스템에 적용될 수 있다.

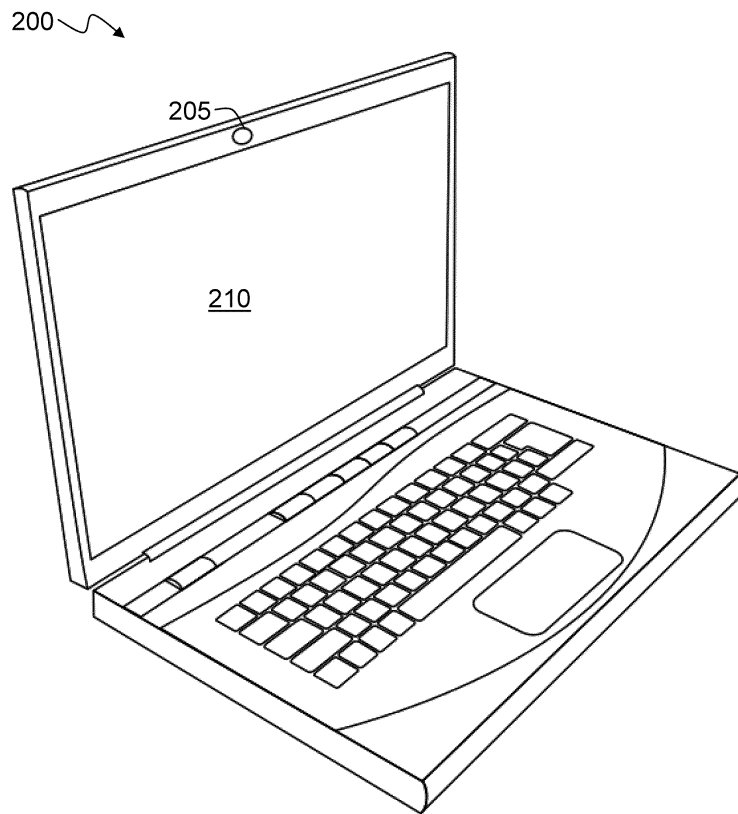
[0086] 전술한 본 발명의 여러 실시예는 당업자에게 설명하기 위한 목적으로 제공되는 것이다. 이러한 설명은 권리 범위를 무한정 확대하거나, 본 발명을 단일의 개시된 실시예로 한정하려는 것은 아니다. 전술한 바와 같이, 본 발명에 관한 많은 대체물 및 변경은 당업자에게 자명할 것이다. 따라서, 일부 대체 실시예가 특별히 기술되었지만, 당업자에 의해 다른 실시예도 자명하거나 비교적 쉽게 개발될 것이다. 따라서, 본 명세서는 본 명세서에 기술된 본 발명의 모든 대체물, 변형 및 변경, 및 전술한 본 발명의 정신 및 범주 내에 속하는 다른 실시예를 수용한다.

도면

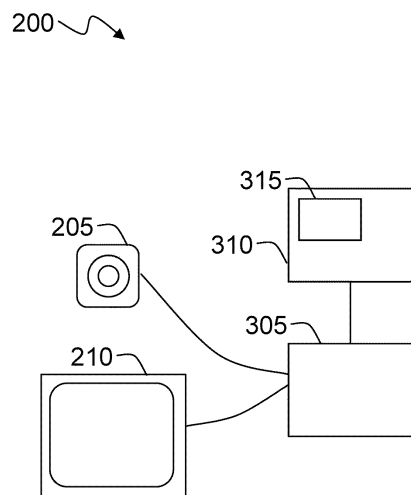
도면1



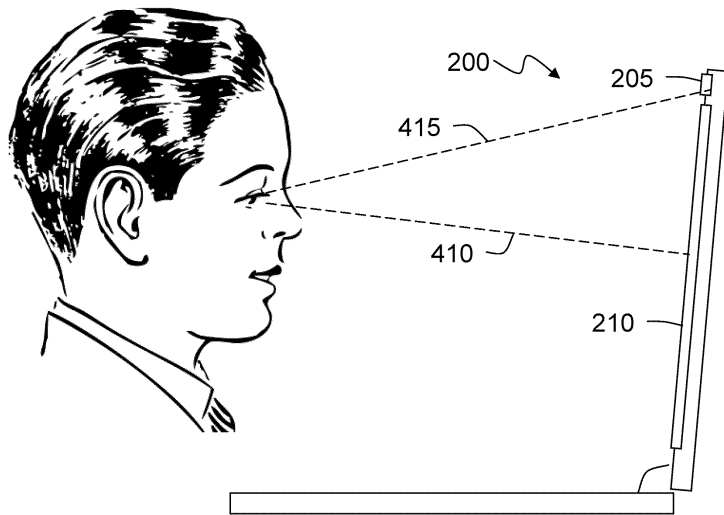
도면2



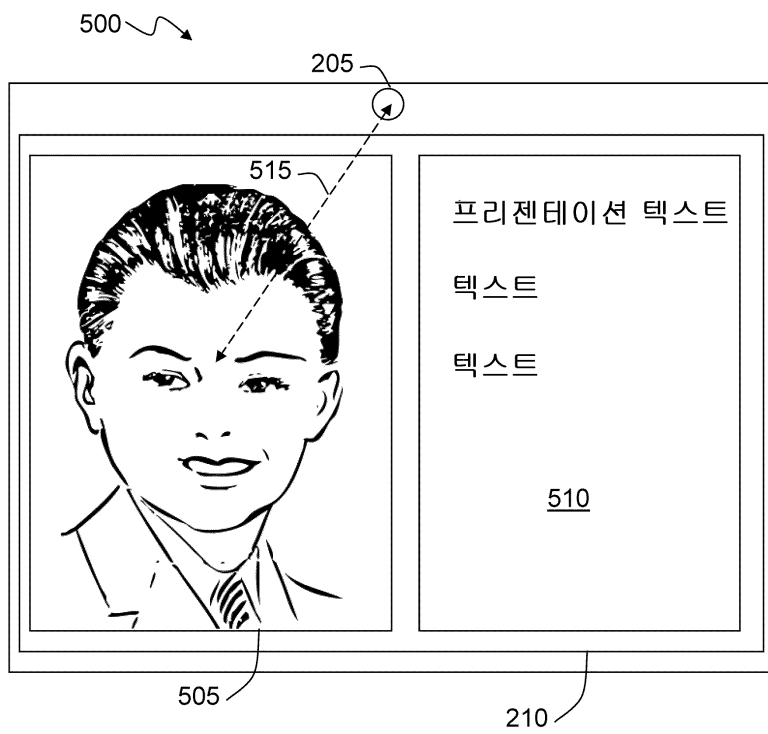
도면3



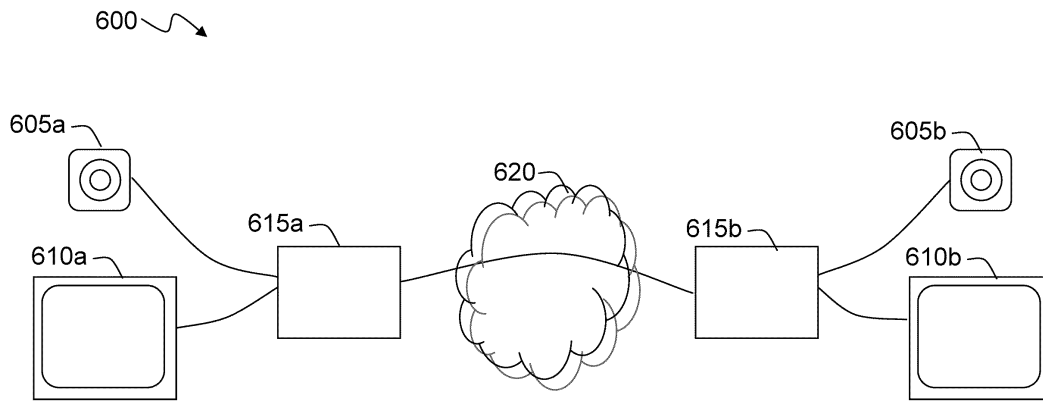
도면4



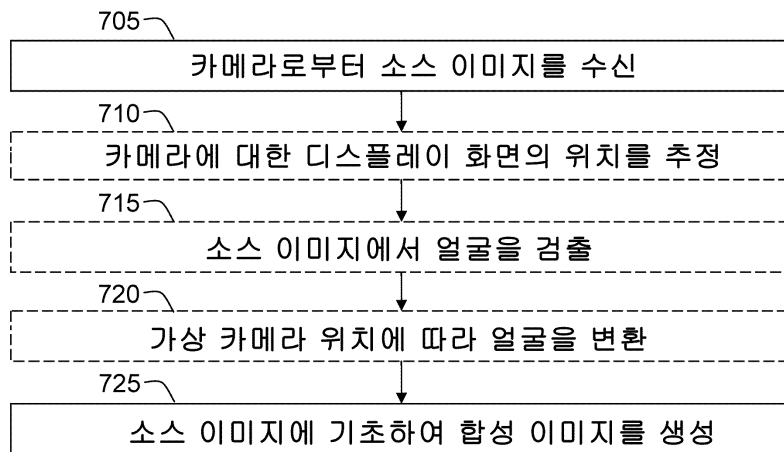
도면5



도면6



도면7



도면8

