



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2015-0082379
(43) 공개일자 2015년07월15일

<p>(51) 국제특허분류(Int. Cl.) G06T 7/00 (2006.01) G06K 9/00 (2006.01)</p> <p>(52) CPC특허분류 G06T 7/0042 (2013.01) G06K 9/00201 (2013.01)</p> <p>(21) 출원번호 10-2015-7014234</p> <p>(22) 출원일자(국제) 2013년10월18일 심사청구일자 없음</p> <p>(85) 번역문제출일자 2015년05월28일</p> <p>(86) 국제출원번호 PCT/US2013/065654</p> <p>(87) 국제공개번호 WO 2014/070483 국제공개일자 2014년05월08일</p> <p>(30) 우선권주장 61/722,091 2012년11월02일 미국(US) 13/831,405 2013년03월14일 미국(US)</p>	<p>(71) 출원인 켈컴 인코퍼레이티드 미국 92121-1714 캘리포니아주 샌 디에고 모어하우스 드라이브 5775</p> <p>(72) 발명자 라이트마이어 게르하르트 미국 92121-1714 캘리포니아주 샌디에고 모어하우스 드라이브 5775</p> <p>플로니 알렉산드로 미국 92121-1714 캘리포니아주 샌디에고 모어하우스 드라이브 5775</p> <p>(74) 대리인 특허법인코리아나</p>
---	--

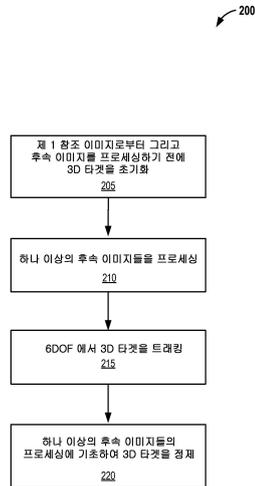
전체 청구항 수 : 총 40 항

(54) 발명의 명칭 단안 시각 SLAM 을 위한 고속 초기화

(57) 요약

고속 시각 동시적 위치인식 및 맵핑을 위한 장치 및 방법이 설명된다. 일 실시형태에서, 3차원 (3D) 타겟이 제 1 참조 이미지로부터 즉시 그리고 후속 이미지를 프로세싱하기 전에 초기화된다. 일 실시형태에서, 하나 이상의 후속 참조 이미지들이 프로세싱되고, 3D 타겟이 6자유도에서 트래킹된다. 일 실시형태에서, 3D 타겟은 프로세싱된 하나 이상의 후속 이미지들에 기초하여 정제된다.

대표도 - 도2



(52) CPC특허분류
G06T 2207/10016 (2013.01)

명세서

청구범위

청구항 1

시각 동시적 위치인식 및 지도작성 (visual simultaneous localization and mapping) 을 위한 프로세서 구현 방법으로서,

제 1 참조 이미지로부터 그리고 후속 이미지를 프로세싱하기 전에 3차원 (3D) 타겟을 초기화하는 단계;

하나 이상의 후속 이미지들을 프로세싱하는 단계;

6자유도에서 상기 3D 타겟을 트래킹하는 단계; 및

상기 하나 이상의 후속 이미지들의 프로세싱에 기초하여 상기 3D 타겟을 정제하는 단계

를 포함하는, 시각 동시적 위치인식 및 지도작성을 위한 프로세서 구현 방법.

청구항 2

제 1 항에 있어서,

상기 3D 타겟을 초기화할 시에 상기 3D 타겟의 올바르게 정렬, 배치 및 이격된 증강 현실 표현을 디스플레이하는 단계

를 더 포함하는, 시각 동시적 위치인식 및 지도작성을 위한 프로세서 구현 방법.

청구항 3

제 2 항에 있어서,

6자유도에서 상기 3D 타겟을 트래킹하면서 상기 3D 타겟의 상기 증강 현실 표현을 업데이트하는 단계

를 더 포함하는, 시각 동시적 위치인식 및 지도작성을 위한 프로세서 구현 방법.

청구항 4

제 1 항에 있어서,

상기 하나 이상의 후속 이미지들을 프로세싱하는 단계는

상기 하나 이상의 후속 이미지들로부터 2D 관심 포인트들을 추출하는 단계

를 더 포함하는, 시각 동시적 위치인식 및 지도작성을 위한 프로세서 구현 방법.

청구항 5

제 1 항에 있어서,

상기 3D 타겟을 초기화하는 단계는

상기 참조 이미지로부터 2D 관심 포인트들의 참조 세트를 추출하는 단계;

타겟 3D 포인트들을 결정하는 단계로서, 각각의 타겟 3D 포인트는 상기 2D 관심 포인트들의 참조 세트에 있는 2D 관심 포인트에 대응하는, 상기 타겟 3D 포인트들을 결정하는 단계; 및

미리결정된 초기 깊이 값을 각각의 타겟 3D 포인트에 할당하는 단계

를 더 포함하는, 시각 동시적 위치인식 및 지도작성을 위한 프로세서 구현 방법.

청구항 6

제 5 항에 있어서,

상기 타겟 3D 포인트들의 각각은 동일한 미리결정된 초기 깊이 값을 갖는 평면을 따라 초기화되는, 시각 동시적 위치인식 및 지도작성을 위한 프로세서 구현 방법.

청구항 7

제 5 항에 있어서,

상기 3D 타겟을 트래킹하는 단계는 상기 타겟 3D 포인트들의 적어도 하나의 타겟 3D 포인트의 참조 위치와 상기 하나 이상의 후속 이미지들로부터 추출된 업데이트된 2D 위치를 비교하는 단계를 더 포함하는, 시각 동시적 위치인식 및 지도작성을 위한 프로세서 구현 방법.

청구항 8

제 5 항에 있어서,

상기 3D 타겟을 정제하는 단계는:

상기 타겟 3D 포인트들의 하나 이상에 대한 업데이트된 깊이 값을 결정하는 단계; 및
할당된 상기 미리결정된 초기 깊이 값을 상기 업데이트된 깊이 값으로 대체하는 단계를 더 포함하는, 시각 동시적 위치인식 및 지도작성을 위한 프로세서 구현 방법.

청구항 9

제 8 항에 있어서,

업데이트된 깊이 값을 갖는 상기 타겟 3D 포인트들의 임계 개수가 충족되는 것을 결정하는 단계로서, 상기 임계 개수는 상기 하나 이상의 후속 이미지들의 각각의 후속 이미지로부터 업데이트된 깊이 값들에 기초하여 충족되는, 상기 타겟 3D 포인트들의 임계 개수가 충족되는 것을 결정하는 단계; 및

상기 각각의 후속 이미지를 제 2 참조 이미지로 할당하는 단계

를 더 포함하는, 시각 동시적 위치인식 및 지도작성을 위한 프로세서 구현 방법.

청구항 10

제 9 항에 있어서,

상기 3D 타겟의 상기 정제는 상기 제 2 참조 이미지로부터 추출되는 복수의 2D 관심 포인트들과 상기 타겟 3D 포인트들을 삼각측량하는 것에 의해 상기 3D 타겟의 추가 정제를 수행하는 단계를 더 포함하는, 시각 동시적 위치인식 및 지도작성을 위한 프로세서 구현 방법.

청구항 11

데이터 프로세싱 디바이스로 하여금 시각 동시적 위치인식 및 지도작성 (visual simultaneous localization and mapping) 을 위한 방법을 수행하게 하는 실행가능한 프로그램 명령들을 포함하는 컴퓨터 판독가능 비일시적 저장 매체로서, 상기 방법은

제 1 참조 이미지로부터 그리고 후속 이미지를 프로세싱하기 전에 3차원 (3D) 타겟을 초기화하는 단계;

하나 이상의 후속 이미지들을 프로세싱하는 단계;

6자유도에서 상기 3D 타겟을 트래킹하는 단계; 및

상기 하나 이상의 후속 이미지들의 프로세싱에 기초하여 상기 3D 타겟을 정제하는 단계

를 포함하는, 컴퓨터 판독가능 비일시적 저장 매체.

청구항 12

제 1 항에 있어서,

상기 3D 타겟을 초기화할 시에 상기 3D 타겟의 올바르게 정렬, 배치 및 이격된 증강 현실 표현을 디스플레이하

는 단계
를 더 포함하는, 컴퓨터 판독가능 비일시적 저장 매체.

청구항 13

제 12 항에 있어서,
6자유도에서 상기 3D 타겟을 트래킹하면서 상기 3D 타겟의 상기 증강 현실 표현을 업데이트하는 단계
를 더 포함하는, 컴퓨터 판독가능 비일시적 저장 매체.

청구항 14

제 11 항에 있어서,
상기 하나 이상의 후속 이미지들을 프로세싱하는 단계는
상기 하나 이상의 후속 이미지들로부터 2D 관심 포인트들을 추출하는 단계
를 더 포함하는, 컴퓨터 판독가능 비일시적 저장 매체.

청구항 15

제 11 항에 있어서,
상기 3D 타겟을 초기화하는 단계는
상기 참조 이미지로부터 2D 관심 포인트들의 참조 세트를 추출하는 단계;
타겟 3D 포인트들을 결정하는 단계로서, 각각의 타겟 3D 포인트는 상기 2D 관심 포인트들의 참조 세트에 있는
2D 관심 포인트에 대응하는, 상기 타겟 3D 포인트들을 결정하는 단계; 및
미리결정된 초기 깊이 값을 각각의 타겟 3D 포인트에 할당하는 단계
를 더 포함하는, 컴퓨터 판독가능 비일시적 저장 매체.

청구항 16

제 15 항에 있어서,
상기 타겟 3D 포인트들의 각각은 동일한 미리결정된 초기 깊이 값을 갖는 평면을 따라 초기화되는, 컴퓨터 판독
가능 비일시적 저장 매체.

청구항 17

제 15 항에 있어서,
상기 3D 타겟을 트래킹하는 단계는 상기 타겟 3D 포인트들의 적어도 하나의 타겟 3D 포인트의 참조 위치와 상기
하나 이상의 후속 이미지들로부터 추출된 업데이트된 2D 위치를 비교하는 단계를 더 포함하는, 컴퓨터 판독가능
비일시적 저장 매체.

청구항 18

제 15 항에 있어서,
상기 3D 타겟을 정제하는 단계는
상기 타겟 3D 포인트들의 하나 이상에 대한 업데이트된 깊이 값을 결정하는 단계; 및
할당된 상기 미리결정된 초기 깊이 값을 상기 업데이트된 깊이 값으로 대체하는 단계
를 더 포함하는, 컴퓨터 판독가능 비일시적 저장 매체.

청구항 19

제 18 항에 있어서,

업데이트된 깊이 값을 갖는 상기 타겟 3D 포인트들의 임계 개수가 충족되는 것을 결정하는 단계로서, 상기 임계 개수는 상기 하나 이상의 후속 이미지들의 각각의 후속 이미지로부터 업데이트된 깊이 값들에 기초하여 충족되는, 상기 타겟 3D 포인트들의 임계 개수가 충족되는 것을 결정하는 단계; 및

상기 각각의 후속 이미지를 제 2 참조 이미지로 할당하는 단계

를 더 포함하는, 컴퓨터 관독가능 비밀시적 저장 매체.

청구항 20

제 19 항에 있어서,

상기 3D 타겟의 상기 정제는 상기 제 2 참조 이미지로부터 추출되는 복수의 2D 관심 포인트들과 상기 타겟 3D 포인트들을 삼각측량하는 것에 의해 상기 3D 타겟의 추가 정제를 수행하는 단계를 더 포함하는, 컴퓨터 관독가능 비밀시적 저장 매체.

청구항 21

시각 동시적 위치인식 및 지도작성 (visual simultaneous localization and mapping) 을 위한 데이터 프로세싱 디바이스로서

프로세서; 및

상기 프로세서에 연결되고 명령들을 저장하기 위해 구성가능한 저장 디바이스를 포함하고,

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

제 1 참조 이미지로부터 그리고 후속 이미지를 프로세싱하기 전에 3차원 (3D) 타겟을 초기화하게 하고;

하나 이상의 후속 이미지들을 프로세싱하게 하고;

6자유도에서 상기 3D 타겟을 트래킹하게 하고; 및

상기 하나 이상의 후속 이미지들의 프로세싱에 기초하여 상기 3D 타겟을 정제하게 하는, 시각 동시적 위치인식 및 지도작성을 위한 데이터 프로세싱 디바이스.

청구항 22

제 21 항에 있어서,

상기 3D 타겟을 초기화할 시에 상기 3D 타겟의 올바르게 정렬, 배치 및 이격된 증강 현실 표현을 디스플레이하게 하는 명령들을 더 포함하는, 시각 동시적 위치인식 및 지도작성을 위한 데이터 프로세싱 디바이스.

청구항 23

제 22 항에 있어서,

6자유도에서 상기 3D 타겟을 트래킹하면서 상기 3D 타겟의 증강 현실 표현을 업데이트하게 하는 명령들을 더 포함하는, 시각 동시적 위치인식 및 지도작성을 위한 데이터 프로세싱 디바이스.

청구항 24

제 21 항에 있어서,

상기 하나 이상의 후속 이미지들을 프로세싱하는 것은

상기 하나 이상의 후속 이미지들로부터 2D 관심 포인트들을 추출하게 하는 명령들을 더 포함하는, 시각 동시적 위치인식 및 지도작성을 위한 데이터 프로세싱 디바이스.

청구항 25

제 21 항에 있어서,

상기 3D 타겟을 초기화하는 것은

상기 참조 이미지로부터 2D 관심 포인트들의 참조 세트를 추출하게 하고;

타겟 3D 포인트들을 결정하게 하는 것으로서, 각각의 타겟 3D 포인트는 상기 2D 관심 포인트들의 참조 세트에 있는 2D 관심 포인트에 대응하는, 상기 타겟 3D 포인트들을 결정하게 하고; 및

미리결정된 초기 깊이 값을 각각의 타겟 3D 포인트에 할당하게 하는 명령들을 더 포함하는, 시각 동시적 위치인식 및 지도작성을 위한 데이터 프로세싱 디바이스.

청구항 26

제 25 항에 있어서,

상기 타겟 3D 포인트들의 각각은 동일한 미리결정된 초기 깊이 값을 갖는 평면을 따라 초기화되는, 시각 동시적 위치인식 및 지도작성을 위한 데이터 프로세싱 디바이스.

청구항 27

제 25 항에 있어서,

상기 3D 타겟을 트래킹하는 것은 상기 타겟 3D 포인트들의 적어도 하나의 타겟 3D 포인트의 참조 위치와 상기 하나 이상의 후속 이미지들로부터 추출된 업데이트된 2D 위치를 비교하는 것을 더 포함하는, 시각 동시적 위치인식 및 지도작성을 위한 데이터 프로세싱 디바이스.

청구항 28

제 25 항에 있어서,

상기 3D 타겟을 정제하는 것은

상기 타겟 3D 포인트들의 하나 이상에 대한 업데이트된 깊이 값을 결정하게 하고; 및

할당된 상기 미리결정된 초기 깊이 값을 상기 업데이트된 깊이 값으로 대체하게 하는 명령들을 더 포함하는, 시각 동시적 위치인식 및 지도작성을 위한 데이터 프로세싱 디바이스.

청구항 29

제 28 항에 있어서,

업데이트된 깊이 값을 갖는 상기 타겟 3D 포인트들의 임계 개수가 충족되는 것을 결정하게 하는 것으로서, 상기 임계 개수는 상기 하나 이상의 후속 이미지들의 각각의 후속 이미지로부터 업데이트된 깊이 값들에 기초하여 충족되는, 상기 타겟 3D 포인트들의 임계 개수가 충족되는 것을 결정하게 하고; 및

상기 각각의 후속 이미지를 제 2 참조 이미지로 할당하게 하는 명령들을 더 포함하는, 시각 동시적 위치인식 및 지도작성을 위한 데이터 프로세싱 디바이스.

청구항 30

제 29 항에 있어서,

상기 3D 타겟의 상기 정제는 상기 제 2 참조 이미지로부터 추출되는 복수의 2D 관심 포인트들과 상기 타겟 3D 포인트들을 삼각측량하는 것에 의해 상기 3D 타겟의 추가 정제를 수행하게 하는 명령들을 더 포함하는, 시각 동시적 위치인식 및 지도작성을 위한 데이터 프로세싱 디바이스.

청구항 31

시각 동시적 위치인식 및 지도작성 (visual simultaneous localization and mapping) 을 위한 장치로서

제 1 참조 이미지로부터 그리고 후속 이미지를 프로세싱하기 전에 3차원 (3D) 타겟을 초기화하는 수단;

하나 이상의 후속 이미지들을 프로세싱하는 수단;

6자유도에서 상기 3D 타겟을 트래킹하는 수단; 및

상기 하나 이상의 후속 이미지들의 프로세싱에 기초하여 상기 3D 타겟을 정제하는 수단을 포함하는, 시각 동시적 위치인식 및 지도작성을 위한 장치.

청구항 32

제 31 항에 있어서,

상기 3D 타겟을 초기화할 시에 상기 3D 타겟의 올바르게 정렬, 배치 및 이격된 증강 현실 표현을 디스플레이하는 수단

을 더 포함하는, 시각 동시적 위치인식 및 지도작성을 위한 장치.

청구항 33

제 32 항에 있어서,

6자유도에서 상기 3D 타겟을 트래킹하면서 상기 타겟의 상기 증강 현실 표현을 업데이트하는 수단

을 더 포함하는, 시각 동시적 위치인식 및 지도작성을 위한 장치.

청구항 34

제 31 항에 있어서,

상기 하나 이상의 후속 이미지들을 프로세싱하는 것은

상기 하나 이상의 후속 이미지들로부터 2D 관심 포인트들을 추출하는 수단

을 더 포함하는, 시각 동시적 위치인식 및 지도작성을 위한 장치.

청구항 35

제 31 항에 있어서,

상기 3D 타겟을 초기화하는 것은

상기 참조 이미지로부터 2D 관심 포인트들의 참조 세트를 추출하는 수단;

타겟 3D 포인트들을 결정하는 수단으로서, 각각의 타겟 3D 포인트는 상기 2D 관심 포인트들의 참조 세트에 있는 2D 관심 포인트에 대응하는, 상기 타겟 3D 포인트들을 결정하는 수단; 및

상기 2D 관심 포인트들의 참조 세트에 있는 2D 관심 포인트에 대응하는 각각의 타겟 3D 포인트에 미리결정된 초기 깊이 값을 할당하는 수단

을 더 포함하는, 시각 동시적 위치인식 및 지도작성을 위한 장치.

청구항 36

제 35 항에 있어서,

상기 타겟 3D 포인트들의 각각은 동일한 미리결정된 초기 깊이 값을 갖는 평면을 따라 초기화되는, 시각 동시적 위치인식 및 지도작성을 위한 장치.

청구항 37

제 35 항에 있어서,

상기 3D 타겟을 트래킹하는 것은 상기 타겟 3D 포인트들의 적어도 하나의 타겟 3D 포인트의 참조 위치와 상기 하나 이상의 후속 이미지들로부터 추출된 업데이트된 2D 위치를 비교하는 수단을 더 포함하는, 시각 동시적 위치인식 및 지도작성을 위한 장치.

청구항 38

제 35 항에 있어서,

상기 3D 타겟을 정제하는 것은

상기 타겟 3D 포인트들의 하나 이상에 대한 업데이트된 깊이 값을 결정하는 수단; 및
 할당된 상기 미리결정된 초기 깊이 값을 상기 업데이트된 깊이 값으로 대체하는 수단을 더 포함하는, 시각 동시적 위치인식 및 지도작성을 위한 장치.

청구항 39

제 38 항에 있어서,

업데이트된 깊이 값을 갖는 상기 타겟 3D 포인트들의 임계 개수가 충족되는 것을 결정하는 수단으로서, 상기 임계 개수는 상기 하나 이상의 후속 이미지들의 각각의 후속 이미지로부터 업데이트된 깊이 값들에 기초하여 충족되는, 상기 타겟 3D 포인트들의 임계 개수가 충족되는 것을 결정하는 수단; 및

상기 각각의 후속 이미지를 제 2 참조 이미지로 할당하는 수단을

을 더 포함하는, 시각 동시적 위치인식 및 지도작성을 위한 장치.

청구항 40

제 39 항에 있어서,

상기 3D 타겟의 상기 정제는 상기 제 2 참조 이미지로부터 추출되는 복수의 2D 관심 포인트들과 상기 타겟 3D 포인트들을 삼각측량하는 것에 의해 상기 3D 타겟의 추가 정제를 수행하는 수단을 더 포함하는, 시각 동시적 위치인식 및 지도작성을 위한 장치.

발명의 설명

기술 분야

[0001] 관련 출원에 대한 상호 참조

[0002] 본원은 2013년 3월 14일자로 출원된 U.S. 출원 번호 13/831,405 에 대한 혜택 및 우선권을 주장하고, 이는 차례로 2012년 11월 2일자로 출원된 U.S. 가출원 번호 61/722,091 에 대한 혜택 및 우선권을 주장하고, 이들 양자 모두는 참조에 의해 본원에 전체적으로 인용된다.

[0003] 분야

[0004] 본원에 개시된 요지는 일반적으로 시각 동시적 위치인식 및 지도작성에 관한 것이다.

배경 기술

[0005] 동시적 위치인식 및 지도작성 (simultaneous localization and mapping; SLAM) 은 환경 또는 장면으로부터 타겟을 수립하기 위하여 증강 현실 시스템 및 로봇 방법에 사용된다. 시각 SLAM (VSLAM) 은 환경의 타겟 또는 모델을 수립하기 위한 입력으로서 카메라 또는 시각 센서 데이터 또는 이미지들을 사용한다. VSLAM 이 증강 현실 (AR) 시스템과 함께 사용될 때, 시각 오브젝트들이 현실 세계의 사용자의 뷰속에 삽입되고 디바이스 (예를 들어, 모바일 디바이스, 휴대 전화 또는 유사한 것) 상에 디스플레이될 수 있다.

[0006] VSLAM 이 카메라 위치 및 배향 (포즈) 을 트래킹 또는 결정하기 위한 하나의 공통 전제 조건은 기지의 참조를 사용하는 것이다. 예를 들어, 기지의 또는 이전에 획득된 참조는 현실 세계속에 삽입된 환경 또는 인공 마커의 3차원 (3D) 모델일 수 있다. 전통적인 VSLAM 는 또한, 제 1 참조 이미지가 초기화 및 트래킹 전에 환경에서 평면 표면의 정밀한 정면 뷰일 것을 요구할 수도 있다. 그와 달리, 기지의 참조 또는 정밀하게 캡처된 초기 이미지가 없으면, 오브젝트들은 잘못된 위치에서 나타날 수 있거나 또는 환경의 지도작성이 완전히 실패할 수도 있다.

[0007] 단일 카메라로 VSLAM 을 이용하는 트래킹 시스템은 또한, 단일 카메라에 의해 캡처된 2개의 분리된 참조 이미지들로부터 3D 타겟을 초기화하는 것에 의존할 수도 있다. 2개의 참조 이미지들에 기초한 전통적인 기법들을 이용하여 3D 타겟을 생성하는 것은, 2개 참조 이미지들간의 카메라 모션이 적절하고 또한 양쪽 모두의 이미지들에 있는 장면들간의 충분한 오버랩이 유지되는 경우에만, 가능하다. 참조 이미지들은, 2개의 특별히 정의된

참조 이미지들 사이의 충분한 최소 병진 (translation) 이 있을 때 적절한 것으로 결정될 수도 있다.

[0008] 전통적인 VSLAM 구현들은 또한, 3D 타겟이 초기화될 수 있기 전에 6자유도 (6 Degrees of Freedom; 6DoF) 카메라 모션을 기록하기 위하여 2개의 참조 이미지들을 선택하거나 또는 추가적인 시각 타겟을 제공하기 위한 직접적인 사용자 입력에 의존할 수도 있다.

[0009] 예를 들어, 일부 트래킹 방법들은, 3D 재구성 방법들이 환경에 있는 실제 평면을 찾고 이 평면으로부터 3D 타겟을 초기화하는데 이용될 수 있도록 시각 피드백 없이 사용자에게 특정 비직관적 모션 시퀀스를 수행할 것을 요구한다.

[0010] 전통적인 VSLAM 방법들의 위의 제한들의 결과로서, 현재 증강 현실 사용자의 경험은 종종 좌절스럽고 부자연스럽게 느껴질 수 있다. 게다가, 대부분의 사용자들은 전통적인 VSLAM 초기화에 필요한 카메라 모션들을 알거나 또는 이해하기 쉽지 않다. 통상적인 사용자들은 또한 종종, 증강 현실 시스템이 장면을 위한 트래킹 업데이트들을 디스플레이할 수 있기 전에는 그들이 특정 모션들을 왜 실시해야 하는지에 대해 혼란스러워 한다.

발명의 내용

해결하려는 과제

[0011] 따라서, 향상된 VSLAM 초기화 및 트래킹이 요망된다.

과제의 해결 수단

[0012] 요약

[0013] 본원에 개시된 실시형태들은 시각 동시적 위치인식 및 지도작성 (visual simultaneous localization and mapping) 방법에 관한 것일 수도 있다. 그 방법은 제 1 참조 이미지로부터 그리고 후속 이미지를 프로세싱하기 전에 3차원 타겟을 초기화하는 단계를 포함한다. 그 방법은 또한, 하나 이상의 후속 이미지들을 프로세싱하는 단계, 6DoF 에서 3D 타겟을 트래킹하는 단계, 및 하나 이상의 후속 이미지들의 프로세싱에 기초하여 3D 타겟을 정제 (refining) 하는 단계를 포함한다.

[0014] 본원에 개시된 실시형태들은 또한, 시각 동시적 위치인식 및 지도작성을 수행하기 위한 명령들을 갖는 컴퓨터 판독가능한 비일시적 저장 매체에 관한 것일 수도 있다. 그 매체는 제 1 참조 이미지로부터 그리고 후속 이미지를 프로세싱하기 전에 3차원 타겟을 초기화하기 위한 명령들을 포함한다. 그 매체는 또한, 하나 이상의 후속 이미지들을 프로세싱하는 것, 6DoF 에서 3D 타겟을 트래킹하는 것, 및 하나 이상의 후속 이미지들의 프로세싱에 기초하여 3D 타겟을 정제하는 것을 위한 명령들을 포함한다.

[0015] 본원에 개시된 실시형태들은 또한, 제 1 참조 이미지로부터 그리고 후속 이미지를 프로세싱하기 전에 3차원 타겟을 초기화하기 위한 수단을 포함하는 장치에 관한 것일 수도 있다. 그 장치는 또한, 하나 이상의 후속 이미지들을 프로세싱하는 것, 6DoF 에서 3D 타겟을 트래킹하는 것, 및 하나 이상의 후속 이미지들의 프로세싱에 기초하여 3D 타겟을 정제하는 것을 위한 명령들을 포함한다.

[0016] 본원에 개시된 실시형태들은 또한, 시각 동시적 위치인식 및 지도작성을 수행하기 위한 명령들을 저장하도록 구성가능한 저장 디바이스 및 프로세서를 포함하는 데이터 프로세싱 시스템에 관한 것일 수도 있다. 그 명령들은 프로세서로 하여금, 제 1 참조 이미지로부터 그리고 하나 이상의 후속 이미지들을 프로세싱하기 전에 3차원 타겟을 초기화하게 한다. 그 명령들은 또한 프로세서로 하여금, 후속 이미지를 프로세싱하게 하고, 6DoF 에서 3D 타겟을 트래킹하게 하고, 그리고 하나 이상의 후속 이미지들의 프로세싱에 기초하여 3D 타겟을 정제하게 한다.

[0017] 다른 특징 및 이점들은 첨부 도면들 및 상세한 설명으로부터 분명해질 것이다.

도면의 간단한 설명

- [0018] 도 1 은 개시된 방법들을 수행할 수 있는 시스템의 일 실시형태의 블록도이고;
- 도 2 는 고속 VSLAM 초기화의 일 실시형태의 흐름도를 도시하고;
- 도 3 은 고속 VSLAM 초기화의 또 다른 실시형태의 흐름도를 도시하고;

- 도 4 는 평면으로 초기화된 타겟의 이미지를 캡처하는 카메라의 예시적인 측면 개략도를 도시하고;
- 도 5 는 장면의 이미지를 캡처하는 카메라 및 그 카메라에 의한 작은 모션의 예시적인 측면 개략도를 도시하고;
- 도 6 은 장면의 이미지를 캡처하는 카메라 및 그 카메라에 의한 큰 모션의 예시적인 측면 개략도를 도시하고;
- 도 7 은 카메라에 의한 큰 모션 후의 장면의 표현에서 업데이트된 관심 포인트들의 세트의 예시적인 측면 개략도를 도시한다.

발명을 실시하기 위한 구체적인 내용

- [0019] "예시적" 또는 "예" 이라는 용어는 "예, 실례, 또는 예시의 역할을 하는 것" 을 의미하는 것으로 여기에서 사용된다. "예시적" 또는 "예" 로서 여기에 설명된 임의의 양태 또는 실시형태는 반드시 다른 양태 또는 실시형태보다 바람직하거나 또는 유리한 것으로 해석되어야 하는 것은 아니다.
- [0020] 도 1 은 개시된 방법들을 수행할 수 있는 시스템을 도시하는 블록도이다. 시스템은, 범용 프로세서 (161), 이미지 프로세서 (166), 포즈 프로세서 (168), 그래픽스 엔진 (167) 및 메모리 (164) 를 포함할 수도 있는 디바이스 (100) 를 포함할 수도 있다. 디바이스 (100) 는 또한, 프로세서들 (161, 166, 및 168) 중의 적어도 하나에 더 연결된 하나 이상의 버스들 (177) 또는 신호 라인들에 연결된 다수의 디바이스 센서들을 포함할 수도 있다. 디바이스 (100) 는 모바일 디바이스, 무선 디바이스, 휴대 전화, 개인 휴대 정보 단말기, 웨어러블 디바이스 (예를 들어, 안경, 시계, 헤드 웨어, 또는 유사한 신체 부착 디바이스), 모바일 컴퓨터, 태블릿, 개인용 컴퓨터, 랩톱 컴퓨터, 또는 프로세싱 능력을 가진 임의의 유형의 디바이스일 수도 있다.
- [0021] 일 실시형태에서, 디바이스 (100) 는 모바일/휴대용 플랫폼일 수도 있다. 디바이스 (100) 는, 카메라 (114) 및/또는 CMOS/시각 센서 (미도시) 와 같은 이미지를 캡처하기 위한 수단을 포함할 수 있고 선택적으로, 가속도계, 자이로스코프, 전자 나침반, 또는 다른 유사한 모션 센싱 요소와 같은 모션 센서들 (111) 을 포함할 수도 있다. 디바이스 (100) 는 또한, 전방 및/또는 후방을 향하는 카메라 (예를 들어, 카메라들 (114)) 상에서 이미지들을 캡처할 수도 있다. 디바이스 (100) 는, 디스플레이 (112) 와 같은 증강 현실 이미지를 디스플레이하기 위한 수단을 포함하는 사용자 인터페이스 (150) 를 더 포함할 수도 있다. 사용자 인터페이스 (150) 는 또한, 키보드, 키패드 (152), 또는 사용자가 정보를 디바이스 (100) 로 입력할 수 있는 다른 입력 디바이스를 포함할 수도 있다. 원한다면, 터치 스크린/센서를 갖는 디스플레이 (112) 속에 가상 키패드를 통합하는 것이 키보드 또는 키패드 (152) 를 제거할 수도 있다. 사용자 인터페이스 (150) 는 또한, 예를 들면 디바이스 (100) 가 휴대 전화와 같은 모바일 플랫폼인 경우, 마이크론 (154) 및 스피커 (156) 를 포함할 수도 있다. 디바이스 (100) 는, 위성 위치 시스템 수신기, 전력 디바이스 (예를 들어, 배터리) 와 같은 다양한 다른 요소들, 그리고 휴대용 및 비휴대용 전자 디바이스들과 통상적으로 연관된 다른 컴포넌트들을 포함할 수도 있다.
- [0022] 디바이스 (100) 는 모바일 또는 무선 디바이스로서 기능할 수도 있고, 임의의 적합한 무선 통신 기술에 기초하거나 또는 그렇지 않으면 지원하는 무선 네트워크를 통한 하나 이상의 무선 통신 링크들을 경유하여 통신할 수도 있다. 예를 들어, 일부 양태들에서, 디바이스 (100) 는 클라이언트 또는 서버일 수도 있고, 무선 네트워크와 어울릴 수도 있다. 몇몇 양태들에서 네트워크는 신체 영역 네트워크 (body area network) 또는 개인 영역 네트워크 (personal area network) (예를 들면, 초광대역) 을 포함할 수도 있다. 몇몇 양태들에서 네트워크는 로컬 영역 네트워크 (local area network) 또는 와이드 영역 네트워크 (wide area network) 를 포함할 수도 있다. 무선 디바이스는 예를 들면 CDMA, TDMA, OFDM, OFDMA, WiMAX, 및 Wi-Fi와 같은 다양한 무선 통신 기술들, 프로토콜들 또는 표준들 중 하나 이상을 지원하거나 또는 그렇지 않으면 사용할 수도 있다. 유사하게, 무선 디바이스는 다양한 대응하는 변조 또는 멀티플렉싱 스킴 (shceme) 들 중 하나 이상을 지원하거나 또는 그렇지 않으면 사용할 수도 있다. 모바일 무선 디바이스는 다른 모바일 디바이스들, 휴대폰, 다른 유선 및 무선 컴퓨터, 인터넷 웹 사이트 등과 무선으로 통신할 수도 있다.
- [0023] 상술된 바처럼, 디바이스 (100) 는 휴대용 전자 디바이스 (예를 들어, 스마트 폰, 전용 증강 현실 (AR) 디바이스, 게임 디바이스, 또는 AR 프로세싱 및 디스플레이 능력을 가진 다른 디바이스) 일 수 있다. 본원에 설명된 AR 시스템을 구현할 수도 있는 디바이스 (100) 는 다양한 환경들 (예를 들어, 쇼핑몰, 거리, 사무실, 가정 또는 사용자가 그의 디바이스를 사용할 수도 있는 곳 어디든지) 에서 사용될 수도 있다. 사용자들은 폭넓게 다양한 상황들에서 디바이스 (100) 의 다수의 특징부 (feature) 들과 인터페이스 접속가능할 수도 있다. AR 콘텍스트에서, 사용자는 디스플레이 (112) 를 통해 현실 세계의 표현을 보는데 디바이스 (100) 를 사용할 수도 있다. 사용자는 현실 세계 이미지들/비디오를 수신하는데 카메라 (114) 를 사용함으로써 AR 가능 디바이스 (100) 와 상호작용할 수도 있다. 다음으로, 디바이스 (100) 는 디스플레이된 현실 세계 이미지들/비디오 상

에 추가 또는 대안의 정보를 겹쳐 놓는 방식으로 이미지들을 프로세싱할 수도 있다. 일부 실시형태들에서, 디스플레이 (100) 상의 AR 구현에서, 현실 세계 오브젝트들 또는 장면들이 실시간으로 또는 근실시간으로, 또는 이미지 캡처의 단시간 윈도우 내에서, 대체되거나 또는 변경되고, 사용자에게 디스플레이 (112) 상에 디스플레이 될 수도 있다. 가상 오브젝트들 (예를 들어, 텍스트, 이미지, 비디오) 은 디바이스 디스플레이 상에 묘사된 장면의 표현속에 삽입될 수도 있다.

[0024] 일 실시형태에서, 본원에서 설명된 고속 VSLAM 초기화 모듈 (Fast VSLAM Initialization module; FVI) 은, 하나 이상의 관심 포인트들 (예를 들어, 타겟 오브젝트 또는 오브젝트들의 그룹/장면과 연관된 3D 관심 포인트들) 을 추출 또는 검출할 수도 있고, 포인트 대응성들의 세트로부터 6DoF 카메라 위치 및 배향 (포즈) 를 추정할 수 있다. 용어 모듈은, 언급된 기능들을 수행할 수 있는, 하드웨어, 펌웨어, 소프트웨어, 또는 이들의 일부 조합을 나타내는데 사용된다. 일부 실시형태들에서, FVI 는 실시간으로, 근실시간으로, 또는 이미지 캡처의 단시간 윈도우 내에서 동작할 수도 있다. 예를 들어, 일 실시형태에서, FVI 는, 증강들이 카메라 (114) 에 의해 캡처된 이미지들과 동시에 그리고 최소 사용자 인식가능 지연 또는 시간 지체로 나타나도록 입력 이미지들의 증강 표현을 디스플레이할 수도 있다.

[0025] FVI 는 이미지들 또는 비디오 입력을 수신한 후에, 입력 이미지들 또는 비디오의 증강 표현 (AR) 을 출력 (예를 들어, 디스플레이) 할 수도 있다. 본원에 사용된, 관심 포인트들의 검출 및 6DoF 카메라 포즈의 추정은 타겟을 "트래킹 (tracking)" 하는 것으로 지칭된다. FVI 는 사용자의 환경에 대한 사전 지식 없이 (예를 들어, 미리뱃붙여진 지도 (prepopulated map), CAD 모델, 장면에서의 마커 또는 유사한 사전정의된 타겟 디스크립터 없이) 그리고 시각 (예를 들어, 디스플레이된 타겟 조정) 피드백을 제공하기 전에 2개의 상이한 정밀 참조 이미지들의 사용 없이 초기화할 수도 있다.

[0026] FVI 는 단일 카메라 (114) 또는 단안 시각 입력 (monocular visual input) 으로부터 이미지들 또는 비디오를 수신하고 트래킹 및 카메라 포즈 결정을 제공할 수도 있다. 일 실시형태에서, FVI 는 실시간으로, 근실시간으로, 또는 단시간 윈도우 내에서, 또는 순시적으로 단일 참조 이미지의 수신시 장면의 증강 (예를 들어, 현실 세계 장면속으로의 가상 오브젝트의 추가 또는 변경) 을 디스플레이할 수도 있다. 일 실시형태에서, FVI 는 정밀 및 실시간, 근실시간, 또는 순시 트래킹을, 단일 참조 이미지 (예를 들어, 카메라 (114) 로부터 캡처된 이미지 프레임) 를 사용하여 그리고 가속도계, 레이저 거리계, 자이로스코프, GPS 또는 위치를 결정하는데 사용되는 다른 센서들로부터의 추가 센서 입력 없이, 제공할 수도 있다. 따라서, 비싸고 복잡한 다수의 카메라 어레이 시스템들이 회피될 수도 있고 FVI 는 일반적으로 이용가능한 카메라 센서들에 연결될 수도 있다. 예를 들어, 카메라는 스마트폰과 같은 모바일 디바이스 또는 AR 안경과 같은 웨어러블 디바이스에 연결될 수도 있고, AR 능력은 모바일 디바이스의 프로세서에 의해 또는 AR 안경에 있는 프로세서에 의해 실행되는 소프트웨어 애플리케이션 또는 프로그램으로서 구현될 수도 있다.

[0027] 도 2 는 VSLAM 초기화의 일 실시형태의 흐름도를 도시한다. 블록 (205) 에서, FVI 는 제 1 참조 이미지 (예를 들어, 단일 참조 이미지) 로부터 그리고 후속 이미지를 프로세싱하기 전에 3D 타겟을 초기화할 수도 있다. 예를 들어, 제 1 이미지 (예를 들어, 카메라 (114) 로부터 캡처된 단일 참조 이미지) 를 초기화하는 것은 아래에서 더 자세히 설명되는 바처럼 초기 추정 깊이를 갖는 3차원 공간에서 관심 포인트들의 세트를 결정하는 것을 포함할 수도 있다. 일부 실시형태들에서, 초기화는 즉시, 실시간으로, 근 실시간으로, 또는 참조 이미지 캡처의 단시간 윈도우 내에서, 일어날 수도 있다.

[0028] 일 실시형태에서, FVI 는 초기 캡처된 참조 이미지로부터 타겟을 생성할 수도 있다. 본원에 사용된 타겟은 또한, 3D 오브젝트 또는 장면의 모델, 또는 지도인 것으로 이해될 수도 있다. FVI 는 단일 카메라 센서 또는 단안 소스로부터의 단일 이미지로부터 타겟을 생성할 수도 있다. 타겟은 메모리 (예를 들어, 디바이스 (100) 의 메모리 (164)) 에 저장되고 이미지로부터 추출되거나 계산된 하나 이상의 3D 관심 포인트들에 의해 표현되거나 또는 이와 연관될 수도 있다. 일부 실시형태들에서, 3D 타겟은 초기 참조 이미지 또는 제 2 참조 이미지를 선택할 때 임의의 사용자 입력 없이 초기화될 수도 있다.

[0029] 블록 (210) 에서, FVI 는 하나 이상의 후속 이미지들을 프로세싱할 수도 있다. 프로세싱은 후속 이미지들의 각각 내에서 관심 포인트들의 세트를 결정하는 것을 포함할 수도 있다. 일부 실시형태들에서, 관심 포인트들의 세트는, 참조 이미지를 프로세싱한 후에 결정되는 동일한 관심 포인트들일 수도 있지만, 위치 및 깊이 값들은 참조 이미지와는 상이할 수도 있다.

[0030] 블록 (215) 에서, FVI 는 6자유도에서 3D 타겟을 트래킹할 수도 있다. FVI 는 수신된 제 1 이미지 프레임에 기초하여 타겟 (예를 들어, 장면, 오브젝트, 또는 지도) 을 트래킹하는 것을 시작할 수도 있다. 일부 실시

형태들에서, 트래킹은 즉시, 순시적으로, 실시간으로, 근실시간으로, 또는 제 1 단일 참조 이미지의 수신 다음에 단시간 기간 내에서 일어날 수도 있다. 일부 실시형태들에서, VSLAM 초기화는, 3D 타겟이 순간적으로, 실시간으로, 근실시간으로, 또는 프레임 캡처의 단시간 기간 내에 초기화될 수도 있기 때문에, 증강 현실 사용자에게 손쉽게 분명해질 수도 있다. 일부 실시형태들에서, 3D 타겟은 또한, 사용자가 카메라 (114) 를 임의의 방향으로 이동시키고 추가 이미지들이 수신되고 프로세싱될 때, 정제될 수도 있다. FVI 는 아래에서 더 상세하게 설명되는 바처럼, 트래킹하면서 타겟의 표현을 디스플레이할 수도 있다.

[0031] 블록 (220) 에서, FVI 는 하나 이상의 후속 이미지들의 프로세싱에 기초하여 3D 타겟을 정제할 수도 있다. FVI 는 관심 포인트들이 추정될 수도 있는지 여부를 결정하고 작은 이동들에 대해 타겟을 증가적으로 정제할 수도 있다. 일 실시형태에서, FVI 는 타겟을 가정 깊이 값을 갖는 평면으로 근사함으로써 초기 작은 모션들에 대한 카메라 모션 추정을 제공할 수도 있다. 또한, 타겟이 평면으로 근사되는 실시형태들에서, (충분한 병진을 갖는 것에 의존할 수도 있는) 등극선 기하 (epipolar geometry) 에 대한 명시적 풀이 단계가 생략될 수 있다. 따라서, 일부 실시형태들에서, FVI 는 자동으로 초기화하고 또한 제 1 참조 이미지로부터 계속 연속적인 카메라 (114) 트래킹을 제공할 수도 있다. FVI 는 또한, 카메라 (114) 에 의해 기록되는 추가적인 이미지들 (예를 들어, 제 1 초기 참조 이미지 후의 후속 이미지 캡처들) 로부터 제 2 참조 이미지를 자동 선택할 수도 있다. 제 2 참조 이미지를 선택하는 방법은 아래에서 더 상세하게 설명된다. 제 2 참조 이미지가 선택된 후 트래킹 및 카메라 (114) 포즈 결정은 제 1 및 제 2 참조 이미지로부터 트래킹 및 카메라 포즈 결정에 기초하여 계속될 수도 있다.

[0032] 일 실시형태에서, 초기 참조 이미지로부터 충분한 병진 (예를 들어, 보다 큰 이동) 을 갖는 이미지의 발견시, 발견된 이미지는 제 2 참조 이미지로서 자동으로 할당될 수도 있다. 제 2 참조 이미지로부터의 관심 포인트들은 타겟을 삼각측량 (triangulate) 하여 트래킹 정확성의 추가 증가를 낳는데 사용될 수도 있다.

[0033] 일부 실시형태들에서, 디바이스 (100) 는 임의의 직접 또는 수동 사용자 상호작용 또는 트리거 없이, FVI 를 (예를 들어, FVI 엔진, 소프트웨어 프로세스 또는 다른 구현으로서) (예를 들어, 프로세서 (161) 를 사용하여) 자동으로 시작, 실행 또는 실시 가능할 수도 있다. 대안적으로, 다른 실시형태들에서, FVI 는 디바이스 (100) 로 하여금 터치 스크린을 터치하는 것, 버튼을 누르는 것, 또는 유사한 입력에 의해 타겟의 초기화 및 트래킹을 시작하도록 사용자에게 프롬프트하게 할 수도 있다. 또 다른 구현들에서, FVI 는 애플리케이션 또는 프로그램속에 통합될 수도 있고 애플리케이션 또는 프로그램은 사용자에게 프롬프트하거나 또는 초기 참조 이미지를 자동으로 캡처하고 트래킹을 시작한다.

[0034] 제 1 이미지는 초기 또는 제 1 참조 이미지일 수도 있다. 디바이스 (100) 는 카메라 (114) 로 초기 이미지를 캡처하고 증강 현실 프로세싱을 위해 그 이미지를 FVI 에 전송할 수도 있다. 일 실시형태에서, 카메라 이동 또는 모션이 검출될 때 초기 이미지 프레임은 이 카메라 (114) 에 의해 자동으로 캡처될 수도 있다. 여기에 사용된 바처럼, 캡처된 이미지는 정지/사진 이미지 프레임 또는 비디오 프레임일 수도 있다는 것이 이해되어야 한다. 예를 들어, 카메라 (114) 는 비디오 그리고 정지 사진 이미지 캡처 능력을 가질 수도 있다.

[0035] 일 실시형태에서, 제 1 참조 이미지 (예를 들어, 카메라 (114) 에 의해 캡처된 제 1 이미지) 로부터 시작되는 고속 카메라 트래킹은, 단일 제 1 이미지 (예를 들어, 카메라 (114) 로부터 캡처된 초기화 이미지) 를 프로세싱하고 초기 추정 깊이를 이용하여 추출된 관심 포인트들의 세트를 계산하는 것에 의해 적어도 부분적으로, 가능하다. 일 실시형태에서, 증강 현실 프로그램 또는 애플리케이션은, 임의의 카메라 (114) 모션 (예를 들어, 카메라 병진) 이 수신되기 전에, 초기화로부터 계속 (예를 들어, 모바일 디바이스 또는 컴퓨터를 위한 디스플레이 상의) 장면에 등록된 콘텐츠 (그래픽스) 를 디스플레이할 수도 있다. FVI 는 또한, 회전 운동만이 수신될 때 장면에 등록된 콘텐츠를 디스플레이할 수도 있다.

[0036] 일 실시형태에서, FVI 는 트래킹을 초기화하기 위하여 사용자로부터의 특정 모션들을 요구함이 없이 타겟 트래킹을 위한 정확한 데이터셋을 생성할 수도 있다. 일 실시형태에서, FVI 는 장면 또는 오브젝트를 즉시 트래킹하고 실시간으로, 근실시간으로, 순시적으로, 또는 단시간 윈도우내에서 디스플레이에 AR 업데이트들 (예를 들어, 타겟의 증강) 을 제공할 수도 있다. 단일 초기 참조 이미지의 초기화로부터 타겟 트래킹을 가능하게 함으로써, 사용자들은 카메라를 이동/재배치하고 디바이스 카메라 (114) 로 타겟 또는 장면을 탐사하는 것을 계속하도록 권장된다. 상이한 각도 및 관점에서의 장면의 보다 큰 탐사는 타겟에 관한 보다 많은 정보를 드러낼 수도 있다. FVI 는 사용자가 타겟을 정제하기 위하여 카메라를 이동시키는 동안 학습되는 추가 정보를 사용할 수도 있다. 일부 실시형태들에서, FVI 는 사용자가 카메라를 이동시킴에 따라 정제된 타겟을 사용하여 실시간 디스플레이 피드백을 제공할 수도 있다. 일부 실시형태들에서, 타겟 및 카메라 포즈에 관한 FVI

의 수집된 정보를 더 정제하는데 사용될 수도 있는 제 2 참조 이미지가, 실시간 디스플레이 피드백을 제공하는 동안 발견 및 선택될 수도 있다.

[0037] 증강 현실 (AR) 구현들에서, FVI 는 선택적으로 타겟을 증강하기 위한 추가 정보를 위해 사용자에게 프롬프트할 수도 있다. 예를 들어, 사용자는 디바이스 (100) 디스플레이 상의 표현에 사용자 생성 또는 선택 콘텐츠를 추가 가능할 수도 있다. 사용자 콘텐츠는 타겟의 표현과 통합되거나, 오버레이되거나, 또는 이를 대체할 수도 있는, 이미지, 3D 오브젝트, 비디오, 텍스트 또는 다른 콘텐츠 타입일 수도 있다.

[0038] 일 실시형태에서, 제 1 참조 이미지 (예를 들어, 카메라 (114) 에 의해 캡처된 제 1 이미지) 로부터의 고속 타겟 트래킹은, 제 1 이미지 (예를 들어, 카메라 (114) 로부터 캡처된 초기화 이미지) 를 3D 타겟으로서 (예를 들어, 평면 또는 다른 기하학 형상의 3D 타겟으로서) 프로세싱하는 것에 의해 적어도 부분적으로, 가능해 질 수도 있다. 일 실시형태에서, FVI 는 관찰된 카메라 (114) 모션을 고려하여 업데이트될 수도 있는 장면 (예를 들어, 3D 지도) 의 부분들 (예를 들어, 관심 포인트들) 을 자동으로 선택한다. 따라서, 일부 실시형태들에서, FVI 는 자동으로 초기화하고 제 1 참조 이미지로부터 계속 후속하여 캡처된 이미지들로 연속적인 카메라 (114) 트래킹을 제공할 수도 있다. FVI 는 또한, 카메라 (114) 에 의해 기록되는 추가적인 이미지들 (예를 들어, 제 1 초기 참조 이미지 후의 후속 이미지 캡처들) 로부터 제 2 참조 이미지를 자동 선택할 수도 있다. 제 2 참조 이미지가 선택된 후 트래킹 및 카메라 (114) 포즈 결정은 제 1 및 제 2 참조 이미지로부터 트래킹 및 카메라 포즈 결정에 기초하여 더욱/완전히 정제될 수도 있다.

[0039] 일 실시형태에서, FVI 는 특정 타겟 형상, 마커 (현실 세계 또는 가상) 또는 태그에 관한 임의의 미리결정되거나 또는 사전 초기화 입력 없이 3D 타겟 트래킹을 위한 정확한 데이터셋을 생성할 수도 있다. 일부 실시형태들에서, FVI 는 3D 타겟의 형상의 사전 지식 없이 그리고 3D 타겟 내의 특정 피쳐들의 존재의 지식 없이 3D 타겟들을 초기화할 수도 있다. 예를 들어, 일 실시형태에서, 기지의 좌표들을 갖는 미리정의된 타겟 오브젝트 또는 타겟을 수신하는 대신에, FVI 는 3D 타겟을 초기화할 수도 있고 모든 발견되는 특징부에 동일한 깊이를 설정할 수도 있다. 그러므로, 일 실시형태들에서, 3D 타겟들을 위한 데이터셋은 제 1 단일 참조 이미지로부터 모든 관심 포인트들을 위한 깊이 및 위치로 완전히 덧붙여진다. 깊이에서의 임의의 에러들은 3D 타겟에 대해 카메라 포즈가 변화함에 따라 정정될 수도 있다.

[0040] 일 실시형태에서, AR 프로그램 또는 애플리케이션은, 임의의 카메라 (114) 또는 회전 모션이 수신되기 전에, 초기화로부터 계속 (예를 들어, 모바일 디바이스 또는 컴퓨터를 위한 디스플레이 상에) 장면에 등록된 콘텐츠 (그래픽스) 를 디스플레이할 수도 있다. 예를 들어, AR 프로그램은 모바일 폰 또는 핸드헬드 디바이스의 디스플레이 상에 타겟의 AR 표현을 제공할 수도 있다.

[0041] 본원에 사용된 관심 포인트 (interest point) 는 이미지의 흥미롭거나 또는 주목할만한 부분으로 정의될 수도 있다. 관심 포인트 검출은 관심 포인트가 특정 픽셀에 존재하는지 여부를 결정하기 위해 모든 픽셀을 조사하는 저레벨 이미지 프로세싱 동작일 수도 있다. 다르게는, 고레벨 알고리즘이 또한 관심 포인트 검출에 사용될 수도 있다. 관심 포인트 검출은 전체 이미지 프레임 또는, 다르게는 이미지의 서브섹션 (subsection) 을 프로세싱할 수도 있다.

[0042] 이미지로부터 추출된 관심 포인트들은 3차원 공간 (예를 들어, 축 X, Y, 및 Z 의 좌표들) 을 따라 구별되는 포인트들을 나타낼 수도 있다. 본원에 사용된 타겟은, 이미지 내의 단일 격리 오브젝트 또는 다수의 오브젝트들로부터 추출되거나 또는 이와 연관된 관심 포인트들을 포함할 수도 있다. 예를 들어, 이미지에서 캡처된 전체 장면은 다수의 오브젝트들을 포함할 수도 있고 각각의 오브젝트는 하나 이상의 추출된 관심 포인트들을 가질 수도 있다. 이미지 또는 장면 내의 오브젝트들의 그룹은 또한, 전체 그룹과 연관된 공동의 조합된 관심 포인트 세트 (collective combined interest point set) 를 가질 수도 있다.

[0043] 일 실시형태에서, 초기 추정 깊이를 갖는 관심 포인트들의 FVI 의 추출은 비평면 장면들 (예를 들어, 초기 참조 이미지 또는 이미지에 평행한 단일 평면이 아니라, 임의의 위치에 있는 평면, 또는 상이한 3D 표면에 있을 수도 있는 장면들, 또는 지배적인 평면 구조를 전혀 갖지 않는 장면들) 을 위한 트래킹 및 카메라 포즈 결정을 허용한다.

[0044] 각각의 캡처된 이미지 또는 비디오 프레임에 대해, 관심 포인트들이 검출되고 나면, 관심 포인트 주위 로컬 이미지 패치가 추출될 수도 있다. 관심 포인트들은, 관심 포인트들을 로컬화하고 그들의 디스크립션 (description) 을 생성하는, SIFT (Scale Invariant Feature Transform) 와 같은, 주지의 기법들을 사용하여 추출될 수도 있다. 원한다면, 다른 기법들, 이를테면 SURF (Speed Up Robust Features), GLOH (Gradient

Location-Orientation histogram), 또는 다른 비슷한 기법들이 사용될 수도 있다. 일부 실시형태들에서, 이미지에 대한 추출된 관심 포인트들의 수가 임계치 (예를 들어, 100개 관심 포인트들 또는 다른 수의 포인트들) 을 초과하는 것으로 결정될 때, 이미지는 제 1 참조 이미지로서 저장될 수도 있고 추출된 관심 포인트들은 참조 포인트들로서 정의될 수도 있다.

[0045] 일부 실시형태들에서, FVI 는 임의의 가속도계 데이터의 사용 없이 타겟을 트래킹할 수도 있다 (예를 들어, 타겟은 중력에 대해 임의의 배향에 있을 수도 있고, 수평 또는 수직 정렬된 오브젝트들이 FVI 트래킹을 위해 동일하게 실행가능하다). 게다가, FVI 는 임의의 형상의 타겟을 트래킹할 수도 있다 (예를 들어, FVI 는 타겟을 식별 및 트래킹하기 위하여 직사각형 또는 또 다른 정의된 기하학적 구조에 의존하지 않는다). 일부 실시형태들에서, FVI 는 타겟을 트래킹하기 위하여 완전한 타겟 형상/구조를 필요로 하거나 또는 취하지 않기 때문에, FVI 는 타겟이 부분적으로 가려져 있을 때에도 타겟을 트래킹할 수도 있다. 예를 들어, FVI 는, 타겟의 하나 이상의 부분들이 초기 참조 프레임으로부터 모호하거나 또는 빠진 경우에도 타겟을 트래킹할 수도 있다. FVI 는 카메라가 6DoF 에서 이동함에 따라 추가 발견된 관심 포인트들로 타겟을 업데이트할 수도 있다.

[0046] 일부 실시형태들에서, FVI 는 타겟을 트래킹하는 것을 시작하기 위해, 사용자 입력을 필요로 하지도, 다른 트래킹 초기화 방법들에 의존하지도 않을 수도 있다. 예를 들어, FVI 트래킹은 사용자가 특정 모션들에서 카메라를 이동시키는 것, 특정 위치에 서있는 것, 카메라를 수평으로 들고 있는 것, 또는 타겟을 초기화하기 전 다른 형태의 트래킹 초기화 방법들을 수행하는 것에 의존하지 않는다.

[0047] 상술된 바처럼, 디바이스 (100) 는 휴대용 전자 디바이스 (예를 들어, 스마트 폰, 전용 증강 현실 (AR) 디바이스, 게임 디바이스, 안경과 같은 웨어러블 디바이스, 또는 AR 프로세싱 및 디스플레이 능력을 갖는 다른 디바이스) 일 수도 있다. 본원에 설명된 AR 시스템을 구현하는 디바이스는 다양한 환경들, 이를테면, 쇼핑물, 거리, 방, 또는 사용자가 휴대용 디바이스를 가져갈 수도 있는 곳 어디든지에서 사용될 수도 있다. AR 콘텍스트에서, 사용자는 그들 디바이스의 디스플레이를 통해 타겟 및 현실 세계의 표현을 보는데 그 디바이스 (100) 를 사용할 수도 있다.

[0048] 사용자는 현실 세계 이미지들/비디오를 수신하고 디바이스 상의 디스플레이된 현실 세계 이미지들/비디오에 추가 또는 대안의 정보를 겹쳐놓거나 또는 오버레이하는데 디바이스의 카메라를 이용함으로써 AR 가능 디바이스와 상호작용할 수도 있다. 사용자가 그들의 디바이스 상에서 AR 구현을 볼 때, 현실 세계 오브젝트들 또는 장면들이 디바이스 디스플레이 상에서 실시간으로 대체 또는 변경될 수도 있다. 가상 오브젝트들 (예를 들어, 텍스트, 이미지, 비디오) 은 디바이스 디스플레이 상에 묘사된 장면의 표현속에 삽입될 수도 있다. 예를 들어, 맞춤형 가상 사진이 현실 세계 사진 (real world sign), 포스터 또는 사진 프레임 위에 삽입될 수도 있다. 또 다른 예는 카메라 및 AR 디바이스로 보여지는 사무실 또는 가정 환경 내와 같은 현실 세계 장면속에 3D 가상 캐릭터 (예를 들어, 비디오 게임 캐릭터) 를 배치한다.

[0049] 일 실시형태에서, AR 디바이스의 사용자의 경험은, 사용자의 환경의 사전 지식 없이 그리고 사용자가 디바이스를 이동시킬 때 디바이스 상에 디스플레이된 AR 을 자동으로 업데이트함으로써 크게 향상될 수도 있다. 예를 들어, 일부 실시형태들에서, FVI 는 지도, CAD 모델, 장면에서의 마커 또는 유사한 것 없이 동작할 수도 있다. FVI 는 또한, 다수의 상이하고 정밀한 이미지 캡처들 없이 시각 피드백 (예를 들어, 디스플레이 (112) 상에 표현되는 타겟에 대한 AR 업데이트) 를 제공함으로써 사용자의 경험을 향상시킬 수도 있다. 일부 실시형태들에서, AR 시스템에 대한 시각 업데이트들은 디스플레이 및 사용자에게 실시간으로, 근실시간으로, 거의 순시적으로, 또는 제 1 참조 이미지를 캡처하는 단시간 윈도우 내에서 제공될 수도 있다.

[0050] 일 실시형태에서, 애니웨어 증강 (anywhere augmentation) 을 사용함으로써, FVI 는, 사용자가 카메라를 포인트에 가리키자마자, 어떻게 증강들이 환경에서 그 선택된 포인트에 고정될 것인지에 대해 사용자에게 상호작용적인 피드백을 제공할 수도 있다. 예를 들어, 디바이스 상의 초기화시, FVI 는 사용자로 하여금 단지 카메라를 타겟에 가리킴으로써 그 타겟을 자동으로 선택할 수 있게 할 수도 있다.

[0051] 디바이스 (100) 및 카메라 (114) 의 이동은 디스플레이로 하여금, 트래킹되는 타겟 (예를 들어, 하나 이상의 오브젝트들 또는 장면들) 의 증강을, 실시간으로, 업데이트하게 할 수도 있다. 디바이스의 이동이 초기 참조 이미지 위치로부터 멀어짐에 따라, 디바이스는 대안의 뷰들로부터 추가 이미지들을 캡처할 수도 있다. 대안의 뷰들이 디스플레이될 때, FVI 가 추가 이미지들을 프로세싱하면서 장면 증강이 더 정확해질 수도 있다. FVI 는 카메라가 바라보는 환경의 3D 지식을 획득하기 위하여 타겟과 연관된 추출된 관심 포인트들의 3D 위치를 추정할 수도 있다. 관심 포인트들을 트래킹하는데 비전 기반 솔루션 (vision-based solution) 을 사용하여, 로컬 법선 벡터 (normal vector) 및 포인트들간의 상대 거리가 추론될 수도 있다.

- [0052] 일 실시형태에서, 오브젝트 또는 그래픽은 카메라 (114) 에 의해 캡처되고 디스플레이 (112) 상에 디스플레이된 비디오 스트림 (또는 이미지) 속으로 삽입되거나 또는 통합될 수도 있다. 디스플레이는, 일부 실시형태들에서, 원래 장면으로부터 심리스 트래킹 (seamless tracking) 에 의해 실시간으로 업데이트할 수도 있다. 예를 들어, 사인 상의 텍스트는 대안의 텍스트로 대체될 수도 있거나, 또는 3D 오브젝트는 장면에 전략적으로 배치되고 디바이스 (100) 상에 디스플레이될 수도 있다. 사용자가 카메라 (114) 의 위치 및 배향을 변화시킬 때, 그래픽 또는 오브젝트는 카메라 (114) 의 상대적인 이동을 매칭하기 위하여 조정 또는 증강될 수도 있다.
- [0053] 일 실시형태에서, FVI 는 레이 (ray) 를 카메라로부터 카메라 뷰 방향에서 추정된 지배적 평면 (예를 들어, 상술된 초기 깊이에서 초기화된 평면) 으로 보낼 수도 있다. FVI 는 가속도계를 이용하여 중력의 방향을 추정하고, 카메라가 어디를 가르키든지 간에, 카메라로부터 보내진 레이에 추가하여, 중력의 방향을 이용해, 디스플레이를 위해 타겟의 3D 증강을 올바르게 정렬, 배치 및 이격할 수도 있다. 예를 들어, 가상 오브젝트가 증강 현실 디스플레이 내에 삽입되면, 가상 오브젝트로부터 멀어지는 카메라 이동은 카메라 (114) 에 의해 이동되는 거리에 상대적으로 가상 오브젝트의 크기를 감소시킬 수도 있다. 예를 들어, 가상 오브젝트로부터 네 걸음 뒤로 물러서는 것은, 모든 다른 변수들이 동일하다면, 가상 오브젝트로부터 반 걸음 물러서는 것에 비해 가상 오브젝트의 크기에 있어서 더 큰 감소를 야기한다. 모션 그래픽스 또는 애니메이션이 FVI 에 의해 표현되는 장면내에서 애니메이션화될 수도 있다. 예를 들어, 애니메이션화된 오브젝트는 증강 현실 디스플레이에 묘사된 장면 내에서 "움직일" 수도 있다.
- [0054] 일 실시형태에서, 카메라 (114) 가 이동하고 카메라 포즈가 (예를 들어, 사용자 또는 로봇에 의해 개시된 이동에 의해) 변화함에 따라, FVI 는 제 2 참조 이미지로서 사용될 적절한 이미지를 선택할 수도 있다. 제 2 참조 이미지는 이미지 피드 또는 스트림으로부터 선택될 수도 있다. 관심 포인트들을 추출하고 제 2 참조 이미지로부터 삼각측량한 후에, 증강의 증가된 정확도가 달성될 수도 있다 (예를 들어, 오브젝트 주위 경계들이 더 정확하게 맞을 수도 있고, 장면에서 오브젝트의 표현이 더 현실적으로 보일 것이고, 타겟 배치가 카메라 (114) 포즈에 대해 보다 정확해질 수도 있다).
- [0055] 당업자는 본원에 기재된 실시형태들이 AR 외의 방식 (예를 들어, 로봇 위치결정 또는 FVI 를 이용한 다른 구현) 으로 구현될 수도 있다는 것을 인식할 것이다.
- [0056] 일 실시형태에서, FVI 는 이미지들 또는 비디오를 입력으로서 수신하기 위하여 프로세서에 의해 실행되는 엔진 또는 모듈로서 구현된다. FVI 는 단일 이미지 I_0 를 수신하는 것으로 시작될 수도 있다. 초기 이미지 I_0 와 연관된 카메라 포즈는 $C_0 = I^{4 \times 4}$, 항등식 (identity) 이다. 2D 관심 포인트들 $p_i = (u_i, v_i)$ 는 단일 이미지로부터 추출될 수도 있다. 각각의 관심 포인트는 초기 깊이 $z_i = 1$ 와 연관될 수도 있다.
- $w_i = \frac{1}{z_i}$
- 깊이는 역의 깊이 (inverse depth) 로서 저장될 수도 있다. 다음으로, 관심 포인트 p_i 에 대응하는 포인트 X_i 의 3D 위치는 동차 좌표 (homogeneous coordinate) 에서 표현되는 $X_i = (u_i, v_i, 1, w_i)^T$ 이다.
- [0057] 시간 t 에서 캡처된 임의의 후속 카메라 이미지 I_t 에 대해, 카메라 위치 C_t 는 3D 포인트들 X_i 의 2D 위치들의 측정으로부터 추정된다. 포인트들 X_i 은 마지막 기지의 카메라 포즈 C_{t-1} 를 사용하여 현재 이미지 I_t 속으로 재투영되고 이미지 I_t 에서의 2D 위치의 새로운 측정들이 이루어질 수도 있다. 일 실시형태에서, 이미지 I_t 에서 2D 위치의 새로운 측정들이 이미지 I_0 으로부터 취해진 이미지 패치들 사이의 정규화된 교차 상관, 또는 이미지 대응성을 획득하는 또 다른 방법을 이용하여 획득될 수도 있다. 그러한 방법을 사용하여, 2D 위치 m_i 는 각각의 포인트 X_i 에 대해 관측될 수 있다. 다음으로, 카메라 위치 C_t 그리고 각각의 포인트의 역의 깊이 w_i 가 가우스 뉴턴 비선형 정제 체계 (Gauss-Newton non-linear refinement scheme) 를 이용하여 최적화될 수 있다.

투영 함수는 하기와 같다:

$$p = \text{proj} \left(\begin{pmatrix} R & T \\ 0 & 1 \end{pmatrix}_t \begin{pmatrix} u \\ v \\ 1 \\ w \end{pmatrix} \right) = \text{proj} \left(R \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} + Tw \right)$$

R, T 및 w 에 관한 자코비안은 하기와 같다:

$$J_{R,T} = \frac{\partial \text{proj}(X)}{\partial X} \partial(R, T) \begin{pmatrix} u \\ v \\ 1 \\ w \end{pmatrix} \quad \text{및}$$

$$J_w = \frac{\partial \text{proj}(X)}{\partial X} T.$$

[0058]

[0059]

C_{t-1} 의 카메라 병진 T 이 너무 작을 때, 3D 포인트의 역의 깊이 w_i 에 대한 관찰 m_i 의 자코비안 J_w 이 사라지는데 (작아지고 거의 제로가 되는데), 이는 역의 깊이의 정확한 추정을 불가능하게 할 수 있다. 그러므로, 가우스 뉴턴 반복 동안, 각각의 포인트에 대해 역의 깊이 파라미터의 정보 매트릭스 $J_w^T J_w$ 가 테스트된다. 정보 매트릭스는 이 경우에 오직 음이 아닌 스칼라 값이다; 그것이 임계치, 이를테면 10^{-3} 보다 낮으면, 그의 역은 0 으로 설정된다. 그 역을 0 으로 설정하는 것은 깊이 좌표를 업데이트하는 것을 회피하는데, 왜냐하면 이것은 이 경우에 신뢰적으로 행해질 수 없기 때문이다.

[0060]

일 실시형태에서, FVI 는 항상 카메라 포즈를 추정할 수 있는데, 왜냐하면 포인트들 X_i 은 기지의 깊이 좌표들을 갖기 때문이다. 그러므로, 일 실시형태에서, 최적화의 카메라 파라미터 부분은 항상 잘 한정 (constrain) 된다.

[0061]

카메라 포즈 파라미터 및 모든 포인트의 역의 깊이 좌표 w_i 양자 모두를 최적화하고 업데이트하는 것은 자코비안이 사라지지 않은 모든 포인트에 대해 새로운 카메라 포즈 C_t 및 새로운 깊이 추정을 낳을 수 있다.

[0062]

도 3 는 고속 VSLAM 초기화의 일 실시형태의 흐름도를 도시한다.

[0063]

블록 (305) 에서, FVI 는 참조 이미지로서 사용할 이미지를 관독 또는 수신할 수도 있다. 참조 이미지는 카메라 센서/이미지 피드로부터의 초기 캡처된 이미지일 수도 있거나, 또는 사용자에게 의해 선택될 수도 있다. 일 실시형태에서, FVI 는 초기 이미지를 선택하기 위한 임의의 사용자 입력 없이 3D 타겟을 초기화할 수 있다. 일 실시형태에서, 3D 타겟을 수립하는 것은 카메라 (114) 가 임의의 방향으로 이동할 때 개시된다. 참조 이미지는, 카메라 (114) 가 6DoF 에서 이동하고 후속 이미지들을 캡처할 때 타겟의 본래의/정제된 형상을 연속적으로 추정하기 위한 출발 포인트일 수 있다. 일 실시형태에서, 사용자는 카메라 (114) 를 임의의 방향으로 이동시킬 수 있고 3D 타겟의 트래킹 및 초기화는, 원래 장면이 적어도 부분적으로 볼 수 있게 머물러 있을 때, 수행될 수도 있다. 일 실시형태에서, 포인트들에 가정 깊이 값을 할당함으로써, 카메라 트래킹은 순간적으로, 실시간으로, 근실시간으로, 또는 제 1 이미지 (참조 이미지) 로부터 단시간 윈도우 내에서 계속 수행될 수도 있다.

[0064]

블록 (310) 에서, FVI 는 상술된 방법들의 하나를 이용하여 참조 이미지로부터 3D 관심 포인트들의 세트를 결정할 수도 있다.

[0065]

블록 (315) 에서, FVI 는 타겟을 기하학적 형상으로 설정할 수도 있다. 일 실시형태에서, 기하학적 형상은, 평면의 형태를 취할 수도 있지만, 다른 형상들이 타겟을 설정하는데 사용될 수도 있다. 일 실시형태에서, FVI 는 3D 관심 포인트들의 각각을 동일/동등 고정 깊이 값에 할당할 수도 있다. 관심 포인트들은 초기에, 이미지 평면에 평행한 2D 평면을 따라 놓일 수도 있다. 일 실시형태에서, 모든 3D 관심 포인트들이 이미지 평면 (예를 들어, 센서 또는 카메라 평면) 에 평행한 평면 상에 있다는 가정은 3D 포인트 깊이들이 업데이트됨에 따라 증가하는 정제로 타겟의 실시간, 근실시간 또는 고속 트래킹을 가능하게 한다.

[0066]

도 4 는 평면으로 초기화된 타겟의 이미지를 캡처하는 카메라의 측면 개략도의 일 실시형태를 도시한다. 도 4 는 또한, 위치 405 에서 카메라 (114) 로 부터 봤을 때, 미리정의된 초기 거리에 있는 평면 (420) 으로 설정된 타겟의 관심 포인트들 (예를 들어, 관심 포인트들 (425, 430, 435, 440, 445, 및 450)). 관심 포인트들

은 위에 개시된 추출을 위한 주지의 기법들 중 하나를 사용하여 추출될 수도 있다. 예를 들어, 위에 개시된 바처럼, 2D 관심 포인트들 $p_i = (u_i, v_i)$ 이 이미지로부터 추출될 수 있다. 각각의 관심 포인트는 초기 깊이 (예를 들어, 위에 개시된 바처럼, $z_i = 1$ 또는 또 다른 미리결정된 초기화 값) 과 연관되어, 모든 관심 포인트들이 초기에 동일한 평면을 따라 있다 (예를 들어, 포인트들 (425, 430, 435, 440, 445, 및 450) 는 단일 평면 (420) 을 따라 있다).

[0067] 다른 실시형태들에서, 관심 포인트들은, 임의의 다른 기하학적 형상, 또는 임의의 배열의 초기화된 깊이 값들로 초기화될 수도 있다. 예를 들어, FVI 는 또 다른 기하학적 형상이 타겟을 위한 베이스라인 출발 포인트이어야 한다는 것을 결정하고, 더 많은 타겟 정보들이 발견됨에 따라 정제할 수도 있다. 게다가, 동일한 깊이 값을 갖는 모든 관심 포인트들 대신에, 초기화된 깊이 값들이 업계에 알려져 있는 다른 기법들에 의해 개별적으로 추정 또는 결정될 수도 있다. 예를 들어, 초기 깊이 값들은, 평균 관심 포인트 깊이가 미리결정된 임계 평균 깊이 내에 속하게 하는 그러한 값들의 범위일 수도 있다.

[0068] 도 4 는 또한 카메라 (114) 에 의해 관찰되는 3D 장면을 표현하는 오브젝트들 (예를 들어, 오브젝트 (410)) 을 예시한다. 라인 (415) 은 오브젝트 (410) 와 연관된 관심 포인트 (425) 가 보여지는 레이 (ray) 를 예시한다 (예를 들어, 오브젝트 (410) 의 에지와 연관된 관심 포인트 (425) 는 평면 (420) 을 따라 검출되는 다른 관심 포인트들처럼 초기 동일 깊이를 갖는다). 위치 405 의 카메라 (114) 로부터 라인 (415) 을 따라 관심 포인트 (425) 까지의 거리가 관심 포인트 (425) 의 가정 깊이이다.

[0069] 도 3 을 참조하면, 블록 (320) 에서, FVI 는 후속 이미지를 프로세싱하는 것을 시작할 수 있다. 일 실시형태에서, 디바이스 (100) 가 참조 이미지 내의 오브젝트들 (예를 들어, 오브젝트 (410)) 에 대하여 이동함에 따라, 카메라 (114) 는 상이한 관점들로부터 타겟의 후속 이미지들의 하나 이상을 계속 캡처하고 프로세싱할 수 있다. 일부 경우들에서, 후속 이미지(들) 은 참조 이미지를 바로 뒤따를 수도 있다. 대안적으로, 후속 이미지(들) 은 이미지들의 스트림으로부터 캡처되고 참조 이미지 후의 임의의 나중 시간에 캡처될 수도 있다 (예를 들어, 다른 이미지들은 참조 이미지와 후속 이미지 사이에 즉시 캡처될 수도 있다). 후속 이미지를 프로세싱하는 것은 후속 이미지 내의 3D 관심 포인트들의 세트를 추출 또는 결정하는 것을 포함할 수 있다. 후속 이미지(들) 에서 3D 관심 포인트들의 세트는, 후속 이미지(들) 에서의 3D 관심 포인트들이 참조 이미지에 비해 상이한 깊이 또는 위치에 있을 수도 있다는 것을 제외하고는, 참조 이미지로부터 결정된 3D 관심 포인트들과 동일 또는 동등할 수도 있다.

[0070] FVI 는 참조 이미지의 결정 직후에 후속 이미지들을 프로세싱하는 것을 시작할 수도 있거나, 또는 다르게는, (예를 들어, 카메라 피드 내에서) 2개 이상의 이미지들이 검출될 때까지 프로세싱을 지연할 수도 있다. 예를 들어, 카메라 (114) 는 초당 30 프레임으로 이미지들을 캡처할 수도 있지만, FVI 는 5개의 직후 프레임들 (예를 들어, 또는 초기 참조 이미지의 캡처 바로 다음의 일정기간내에 캡처된 프레임들) 은 측정가능한 카메라 병진을 낳기 쉽지 않다는 것을 결정할 수도 있다. 그러므로, 일부 실시형태들에서, FVI 는 하나 이상의 후속 프레임들을 건너뛴 수도 있거나 또는 초기 참조 프레임들에 후속한 일정기간이 건너뛰어질 수도 있다. 유사하게, FVI 는 전체 프로세싱 시간을 감소시키기 위하여 이미지 피드에서 모든 이미지를 항상 프로세싱하는 것은 아닐 수도 있다. 예를 들어, FVI 는 관심 포인트들을 프로세싱하기 위하여 이미지들의 피드의 샘플 (예를 들어, 2개의 기록된 이미지마다 하나, 또는 5개의 기록된 이미지들 중 하나) 을 사용할 수도 있다.

[0071] 블록 (325) 에서, FVI 는 후속 이미지의 관심 포인트들을 트래킹할 수 있다. 일부 실시형태들에서, 타겟을 트래킹하는 것은 참조 이미지로부터 각각의 3D 관심 포인트와 후속 이미지로부터의 대응하는 3D 관심 포인트의 위치들을 비교하는 것 및 양쪽 모두의 이미지들에서 발견되는 동일한 3D 관심 포인트들의 하나 이상의 위치의 변화를 결정하는 것을 포함한다. 각각의 후속 이미지 (예를 들어, 초기 참조 이미지 후에 캡처되고 프로세싱된 이미지) 에 대해 FVI 는 후속 이미지가 아래에서 더 상세하게 개시되는 바처럼 제 2 참조 이미지로서 할당을 위한 후보인지 여부를 결정할 수도 있다. 일부 모션에 대해, 평면 근사는 임의의 남아있는 에러들이 대체로 사용자에 의해 눈에 띄지 않도록 트래킹 에러들을 감소 또는 삭제한다. 예를 들어, 작은 이동들에 대해, 장면의 제 1 평면 근사는 트래킹 에러들이 눈에 띄지 않도록 충분히 정확할 수 있다. 이들 작은 이동들로부터 생기는 이미지들은 후속 참조 이미지로서 사용되지 않지만, 이것은 또한 제 2 참조 이미지가 임의의 에러들을 더 감소시키기 위하여 획득될 때까지 사용자에 의한 장면 탐색을 증진시킨다.

[0072] 평면으로의 작은 이동 트래킹의 일 실시형태가 카메라 위치 405 로부터 카메라 위치 505 로의 작은 이동 (510) 의 측면 개략도로서 도 5에 예시되어 있다. 도 5에 예시된 바처럼, 레이 라인 및 관심 포인트들이 카메라

위치 405 로부터의 작은 이동 (510) 으로 초기에 벗어날 수도 있다.

[0073] 도 6 은 장면의 이미지를 캡처하는 카메라 및 위치 405 로부터 새로운 카메라 위치 615 로 카메라에 의한 큰 모션 (610) 의 측면 개략도의 일 실시형태를 도시한다. 도 6은 관심 포인트 (425) 와 사전에 연관된 레이 415 가 레이 620 이 되고, 더 이상 관심 포인트 (425) 를 정확히 트래킹하지 않는 것을 도시한다.

[0074] 도 3을 참조하면, 블록 (330) 에서, FVI 는 각각의 관심 포인트 깊이, 즉 3D 포인트 위치가 추정될 수 있는지 여부를 결정할 수 있다. 3D 포인트 추정은, 카메라 모션과 모든 가능한 3D 포인트들을 공동으로 추정하기 때문에, 등극 제약 (epipolar constraint) 을 준수한다. 위에 논의된 바처럼, 각각의 관심 포인트에 대해

역의 깊이 파라미터의 정보 매트릭스 $J_w^T J_w$ 는 관심 포인트의 깊이가 추정될 수 있는지를 결정하기 위하여 테스트된다. 관심 포인트 깊이가 추정될 수 없으면, FVI 는 블록 (340) 으로 진행되고 카메라 제약을 추가할 수도 있다. 그와 달리, 관심 포인트가 추정될 수 있으면, 추정된 관심 포인트는 블록 (335) 에서 업데이트된 포인트들의 리스트에 추가된다. 대안적으로 추정된 관심 포인트는, 업데이트된 포인트들에 대해 별도의 리스트 또는 위치를 유지하는 대신에, 관심 포인트들에 대한 주된 저장 위치 (예를 들어, 데이터베이스, 플랫폼 파일 또는 다른 저장 유형) 을 업데이트할 수도 있다. 관심 포인트들을 업데이트하거나 또는 업데이트 없음을 결정하는 것 중 어느 하나 후에, FVI 는 관심 포인트에 대한 카메라 제약을 추가하기 위하여 블록 (340) 으로 진행할 수도 있다.

[0075] 블록 (340) 에서, 카메라 제약이 관심 포인트들에 추가된다. 본원에 사용된 제약들은 자코비안 및 관측된 2D 측정에 의해 주어질 선형 제약들이다. 투영 함수 및 R , T , 및 w 에 관한 자코비안의 상세한 논의에 대해서는 위를 참조한다.

[0076] 블록 (345) 에서, FVI 는 카메라 및 타겟을 최적화할 수 있다. 일 실시형태에서, FVI 는 카메라 (114) 및 타겟을 카메라 (114) 의 이동에 기초하여 최적화한다. 사용자가 카메라 (114) 를 이동시키기에 따라, 디스플레이는 실시간으로 사용자에게 업데이트를 제공할 수 있다. 실시간 피드백을 수신하는 것은 사용자로 하여금 카메라 (114) 를 계속 이동하도록 권장할 수 있다. 추가적인 카메라 이동은 타겟의 추가 정제 및 트래킹에 있어서의 더 큰 정확성을 제공할 수 있다. 예를 들어, AR 콘텍스트에서, 디바이스 (100) 는 타겟을 갖는 통합 그래픽 또는 오브젝트를 디스플레이할 수도 있고, 그래픽 또는 오브젝트는, 카메라가 (예를 들어, 6DoF 의 하나에서) 이동할 때 그래픽 또는 오브젝트가 다른 오브젝트들 또는 환경에 대한 그의 상대적인 (제 1 참조 이미지로부터 결정되는 그들의 위치 및 배향에 상대적인) 위치를 유지하도록 장면에 배치될 수 있다.

[0077] 블록 (350) 에서, FVI 는 충분한 포인트들이 추정되었는지 여부를 결정할 수 있다. 포인트들의 개수에 대한 임계치는 상이한 장면들에서 방법을 테스트하는 것을 통해 설정된다, 예를 들어 임계치는 설계 및 시스템 파라미터들에 기초하여 설정될 수도 있다. 일 실시형태에서, 임계치는 100개의 포인트들로, 또는 다르게는 150개의 포인트들, 또는 어느 다른 개수들로 설정될 수도 있다. 충분한 포인트들이 있다면, FVI 는 블록 (355) 으로 진행할 수도 있다. 일 실시형태에서, 포인트들의 임계 개수가 충족될 때, FVI 는 제 2 참조 이미지로서 포인트들과 연관된 이미지를 할당하고 블록 (355) 으로 진행할 수도 있다. 그렇지 않으면, FVI 는 블록 (320) 에서 후속 이미지를 계속 프로세싱한다. 일 실시형태에서, FVI 가, 각각의 이미지를 프로세싱하는 한편, 충분한 관심 포인트들이 추정되었는지 결정하는 것을 (예를 들어, 카메라 (114) 로부터 캡처된 일련의 이미지 프레임들로부터) 다수의 후속 이미지들을 통해 반복할 수 있다.

[0078] 블록 (355) 에서, FVI 는 완료된 것으로 결정될 수 있고 증가된 정확성 트래킹이 후속 이미지들에 제공된다. 예를 들어, 도 7에 예시된 바처럼, FVI 는 제 1 및 제 2 참조 이미지로부터 관심 포인트 위치들을 삼각측량할 수 있다. 2개 참조 이미지들은 상이한 보기 방향들로부터 동일한 타겟 (예를 들어, 오브젝트 또는 장면의 섹션) 을 묘사할 수도 있다. FVI 는 대응성 (즉, 제 1 및 제 2 이미지들 양자 모두에서의 관심 포인트 위치들) 을 찾고 이들 대응하는 관심 포인트들의 3D 구조를, 제 1 참조 이미지로부터 제 2 참조 이미지로 카메라를 이동시키는 모션에 따라, 계산할 수도 있다. 관심 포인트들의 삼각측량 (triangulation) 은, 상술된 초기 작은 모션 정제 후에 제 2 레벨의 정제로 고려될 수도 있다. 삼각측량은 2개 참조 이미지들이 결정된 후에 일어날 수 있다.

[0079] 도 7 은 카메라 (405) 의 측면 개략도로서, 큰 카메라 이동 후 그리고 제 2 참조 이미지가 카메라 (114) (615) 에 의해 관찰되는 포인트들 (725, 735, 745, 750, 740, 및 760) 의 세트의 각각에 대한 보다 정확한 (즉, 업데이트된) 관심 포인트 깊이들을 허용하는 후를 예시한다. 라인 (620) 은 업데이트된 관심 포인트 (760) 가 보여지는 레이를 예시한다.

- [0080] 여기서의 교시는 다양한 장치들 (예를 들면, 디바이스들) 에 포함 (예를 들면, 그들 내부에 구현되거나 또는 그들에 의해 수행) 될 수도 있다. 예를 들면, 여기에 교시된 하나 이상의 양태들은 폰 (예를 들면, 휴대폰), 휴대 정보 단말 ("PDA"), 태블릿, 모바일 컴퓨터, 랩톱 컴퓨터, 태블릿, 엔터테인먼트 디바이스 (예를 들면, 음악 또는 비디오 디바이스), 헤드셋 (예를 들면, 헤드폰, 이어폰 등), 의료 디바이스 (예를 들면, 생체인식 센서, 심박수 모니터, 계보기, EKG 디바이스 등), 사용자 I/O 디바이스, 컴퓨터, 서버, 포스 디바이스 (point-of-sale device), 엔터테인먼트 디바이스, 셋톱 박스, 또는 임의의 다른 적합한 디바이스에 포함될 수도 있다. 이들 디바이스들은 상이한 전력 및 데이터 요건들을 가질 수도 있고 각각의 관심 포인트 또는 관심 포인트들에 대해 생성된 상이한 전력 프로파일들을 낳을 수도 있다.
- [0081] 몇몇 양태들에서 무선 디바이스는 통신 시스템을 위한 액세스 디바이스 (예를 들면, Wi-Fi 액세스 포인트) 를 포함할 수도 있다. 그러한 액세스 디바이스는, 예를 들면, 유선 또는 무선 통신 링크를 경유하여 트랜시버 (140) 를 통해 다른 네트워크 (예를 들면, 인터넷과 같은 와이드 영역 네트워크 또는 셀룰러 네트워크) 에의 접속성 (connectivity) 을 제공할 수도 있다. 따라서, 액세스 디바이스는 다른 디바이스 (예를 들면, Wi-Fi 국) 이 다른 네트워크 또는 일부 다른 기능 (functionality) 에 액세스하는 것을 가능하게 할 수도 있다. 또한, 디바이스들 중 하나 또는 양자 모두는 휴대용이거나 또는 몇몇 경우들에서는 상대적으로 휴대용이 아닐 수도 있다는 것이 인식되어야 한다.
- [0082] 당업자는 정보 및 신호가 임의의 다양한 상이한 기술 및 기법을 이용하여 표현될 수도 있음을 이해할 것이다. 예를 들어, 위의 상세한 설명 전반에 걸쳐 참조될 수도 있는 데이터, 명령, 커맨드, 정보, 신호, 비트, 심볼, 및 칩은 전압, 전류, 전자기파, 자기장 또는 자기입자, 광학장 (optical field) 또는 광학 입자, 또는 이들의 임의의 조합에 의해 표현될 수도 있다.
- [0083] 당업자는 또한, 여기에 개시된 실시형태와 관련하여 설명된 다양한 예시적인 논리 블록, 모듈, 엔진, 회로, 및 알고리즘 단계가 전자 하드웨어, 컴퓨터 소프트웨어, 또는 이 양자의 조합으로 구현될 수도 있음을 인식할 것이다. 하드웨어와 소프트웨어의 이러한 상호교환가능성을 명확히 예시하기 위해, 다양한 예시적인 컴포넌트, 블록, 모듈, 엔진, 회로, 및 단계가 일반적으로 그들의 기능성의 관점에서 위에서 설명되었다. 그러한 기능성이 하드웨어 또는 소프트웨어로 구현될지 여부는, 전체 시스템에 부과된 설계 제약 및 특정 애플리케이션에 의존한다. 당업자는 설명된 기능성을 특정 애플리케이션 각각에 대한 다양한 방식으로 구현할 수도 있지만, 이러한 구현 결정이 본 발명의 범위를 벗어나게 하는 것으로 해석되지는 않아야 한다.
- [0084] 여기에 개시된 실시형태와 관련하여 설명된 다양한 예시적인 논리 블록, 모듈, 및 회로는 범용 프로세서, 디지털 신호 프로세서 (DSP), 주문형 집적 회로 (ASIC), 필드 프로그래밍가능 게이트 어레이 (FPGA) 또는 다른 프로그래밍가능 로직 디바이스, 이산 게이트 또는 트랜지스터 로직, 이산 하드웨어 컴포넌트 또는 여기에 설명된 기능을 수행하도록 설계된 이들의 임의의 조합으로 구현 또는 수행될 수도 있다. 범용 프로세서는 마이크로프로세서일 수도 있지만, 다르게는, 프로세서는 임의의 종래의 프로세서, 제어기, 마이크로제어기, 또는 상태 머신일 수도 있다. 또한, 프로세서는 컴퓨팅 디바이스들의 조합, 예를 들어, DSP 와 마이크로프로세서의 조합, 복수의 마이크로프로세서, DSP 코어와 결합한 하나 이상의 마이크로프로세서, 또는 임의의 다른 이러한 구성요소로서 구현될 수도 있다.
- [0085] 여기에 개시된 실시형태들과 관련하여 설명된 방법 또는 알고리즘의 단계들은, 직접 하드웨어로, 프로세서에 의해 실행되는 소프트웨어로, 또는 양자의 조합으로 구체화될 수도 있다. 소프트웨어는 RAM 메모리, 플래시 메모리, ROM 메모리, EPROM 메모리, EEPROM 메모리, 레지스터, 하드디스크, 리무버블 디스크, CD-ROM, 또는 업계에 공지된 임의의 다른 형태의 저장 매체에 상주할 수도 있다. 예시적 저장 매체는 프로세서가 저장 매체로부터 정보를 판독할 수 있고 저장 매체에 정보를 기입할 수 있도록 프로세서에 커플링된다. 다르게는, 저장 매체는 프로세서에 통합될 수도 있다. 프로세서 및 저장 매체는 ASIC 에 상주할 수도 있다. ASIC는 사용자 단말기에 상주할 수도 있다. 다르게는, 프로세서 및 저장 매체는 사용자 단말기 내에 이산 컴포넌트로서 상주할 수도 있다.
- [0086] 하나 이상의 예시적 실시형태들에서, 설명된 기능 또는 모듈은 하드웨어 (예를 들어, 하드웨어 (162)), 소프트웨어 (예를 들어, 소프트웨어 (165)), 펌웨어 (예를 들어, 펌웨어 (163)), 또는 이들의 임의의 조합으로 구현될 수도 있다. 컴퓨터 프로그램 제품으로서 소프트웨어로 구현되면, 그 기능 또는 모듈은 비일시적 컴퓨터 판독가능 매체 상의 하나 이상의 명령 또는 코드로서 저장되거나 또는 송신될 수도 있다. 컴퓨터 판독가능 매체는 일 장소로부터 다른 장소로의 컴퓨터 프로그램의 전송을 용이하게 하는 임의의 매체를 포함하는 통신 매체 및 컴퓨터 저장 매체 양자를 포함할 수 있다. 저장 매체는 컴퓨터에 의해 액세스될 수 있는 임의의 가용 매

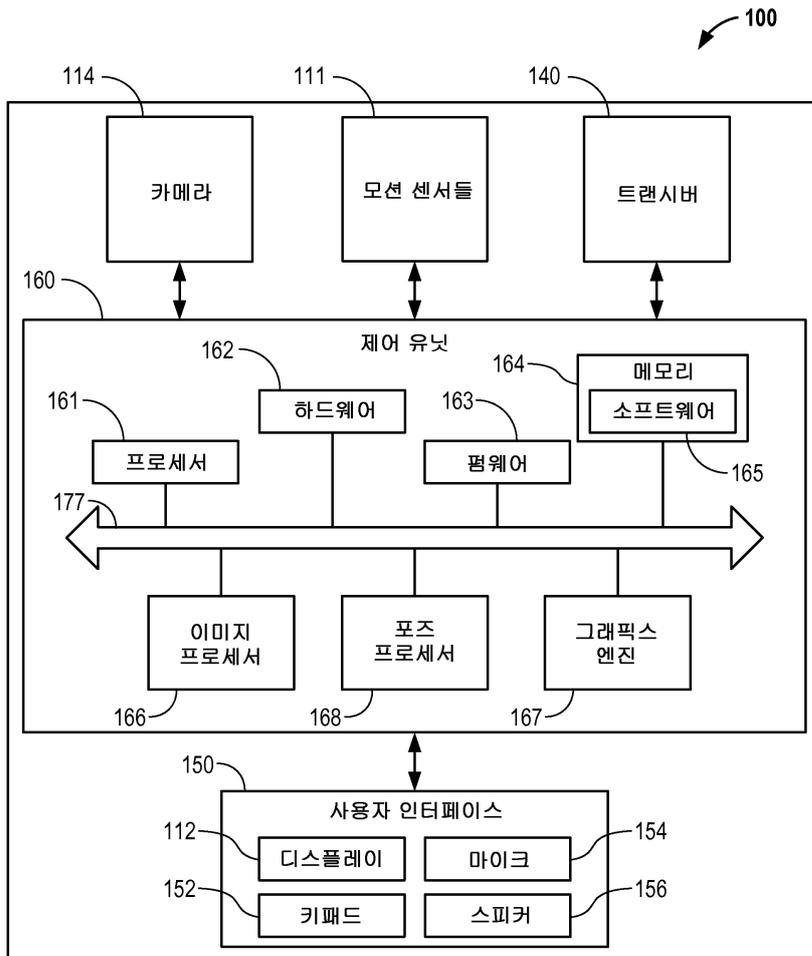
체일 수도 있다. 비한정적 예로서, 이러한 비밀시적 컴퓨터 판독가능 매체는 RAM, ROM, EEPROM, CD-ROM 또는 다른 광학 디스크 저장, 자기 디스크 저장 또는 다른 자기 저장 디바이스들, 또는 명령 또는 데이터 구조의 형태로 원하는 프로그램 코드를 전달 또는 저장하는데 사용될 수 있고 컴퓨터에 의해 액세스될 수 있는 임의의 다른 매체를 포함할 수 있다. 또한, 임의의 접속이 컴퓨터 판독가능 매체로 적절히 칭해진다. 예를 들어, 소프트웨어가 동축 케이블, 광섬유 케이블, 연선 (twisted pair), 디지털 가입자 라인 (DSL), 또는 적외선, 전화, 및 마이크로파와 같은 무선 기술을 사용하여 웹사이트, 서버, 또는 다른 원격 소스로부터 송신되면, 그 동축 케이블, 광섬유 케이블, 연선, DSL, 또는 적외선, 전화, 및 마이크로파와 같은 무선 기술은 매체의 정의 내에 포함된다. 여기에 사용된 바와 같이, 디스크 (disk) 및 디스크 (disc) 는 콤팩트 디스크 (CD), 레이저 디스크, 광 디스크, DVD (digital versatile disc), 플로피 디스크 및 블루레이 디스크를 포함하며, 여기서, 디스크 (disk) 는 보통 데이터를 자기적으로 재생하지만, 디스크 (disc) 는 레이저를 이용하여 광학적으로 데이터를 재생한다. 또한, 상기의 조합은 비밀시적 컴퓨터 판독가능 매체의 범위 내에 포함되어야 한다.

[0087]

개시된 실시형태의 이전의 설명은 당업자가 본 발명을 제조 또는 사용하는 것을 가능하게 하기 위하여 제공된다. 이들 실시형태에 대한 다양한 변형은 당업자에게는 용이하게 명백할 것이며, 여기에 정의된 일반적인 원리는 본 발명의 사상 또는 범위를 벗어남이 없이 다른 실시형태에 적용될 수도 있다. 따라서, 본 발명은 여기에 나타난 실시형태로 한정되도록 의도된 것이 아니라, 여기에 개시된 원리 및 신규한 특징들에 부합하는 최광의 범위가 허여되어야 한다.

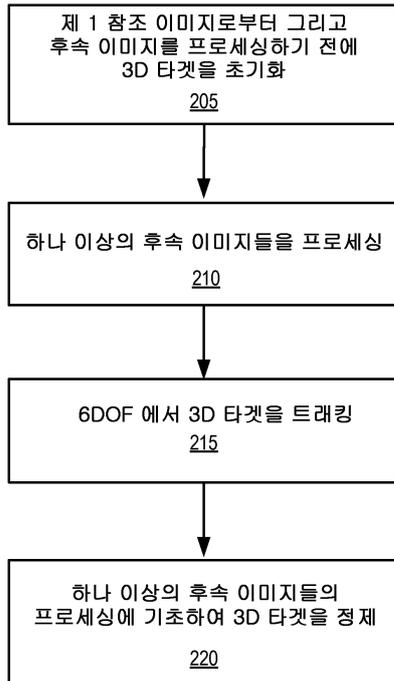
도면

도면1

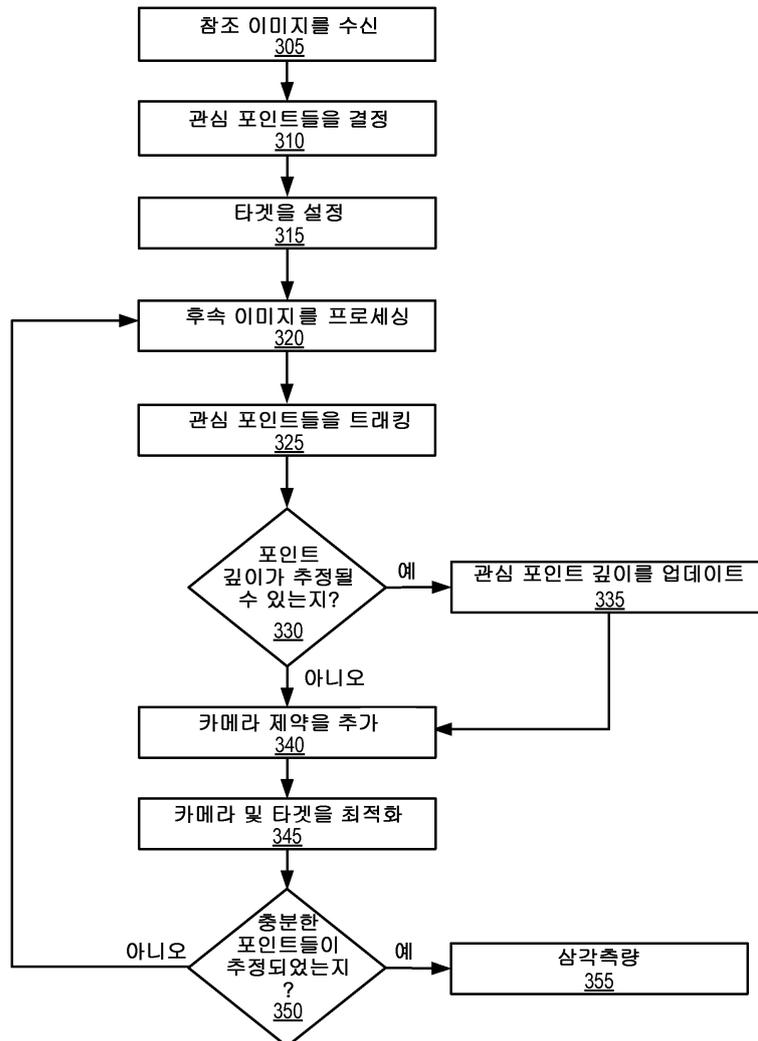


도면2

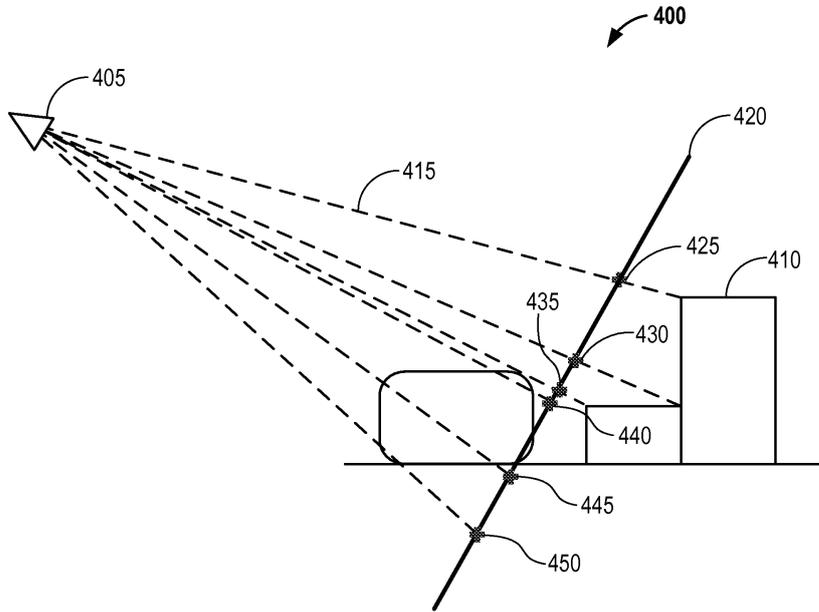
200



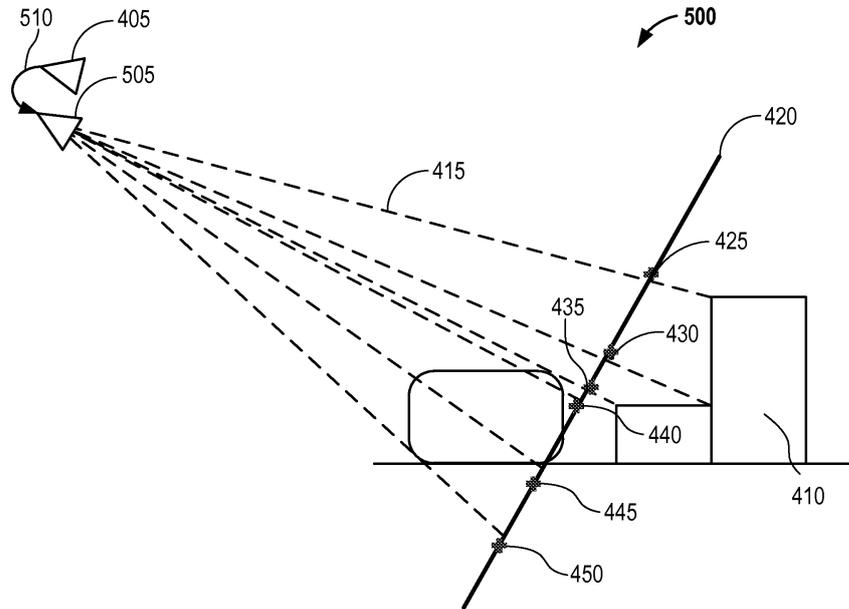
도면3



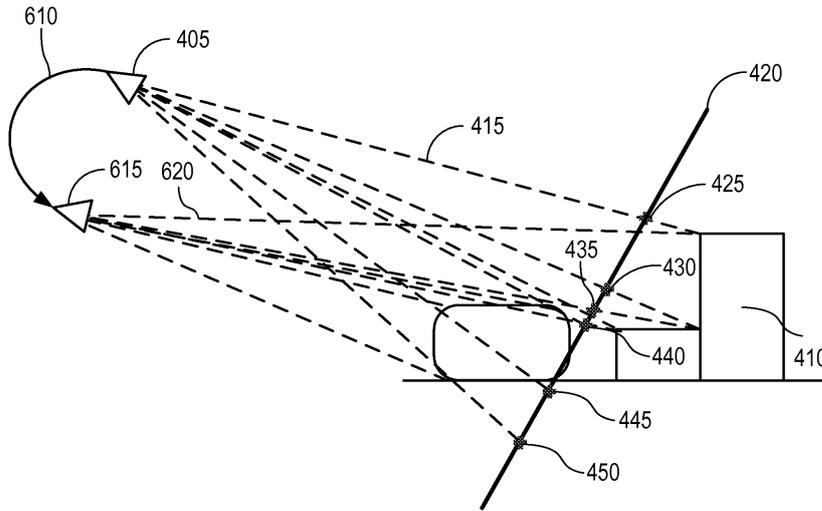
도면4



도면5



도면6



도면7

