



(19) 대한민국특허청(KR)  
(12) 공개특허공보(A)

(11) 공개번호 10-2015-0013709  
(43) 공개일자 2015년02월05일

- (51) 국제특허분류(Int. Cl.)  
*H04N 5/272* (2006.01) *G06T 7/00* (2006.01)
- (21) 출원번호 10-2014-7034271
- (22) 출원일자(국제) 2013년05월09일  
심사청구일자 없음
- (85) 번역문제출일자 2014년12월05일
- (86) 국제출원번호 PCT/GB2013/051205
- (87) 국제공개번호 WO 2013/167901  
국제공개일자 2013년11월14일
- (30) 우선권주장  
1208088.3 2012년05월09일 영국(GB)

- (71) 출원인  
엔씨에이엠 테크놀로지스 리미티드  
영국 더블유1디 3비피 런던 칼라일 스트리트 8-9
- (72) 발명자  
부아뱅, 사무엘  
프랑스 91690 사끌라 뤼 드 퐁 아 샤 6배  
미슈우, 브리스  
프랑스 60500 샹티이 마리 아멜리에 에비뉴 24
- (74) 대리인  
장훈

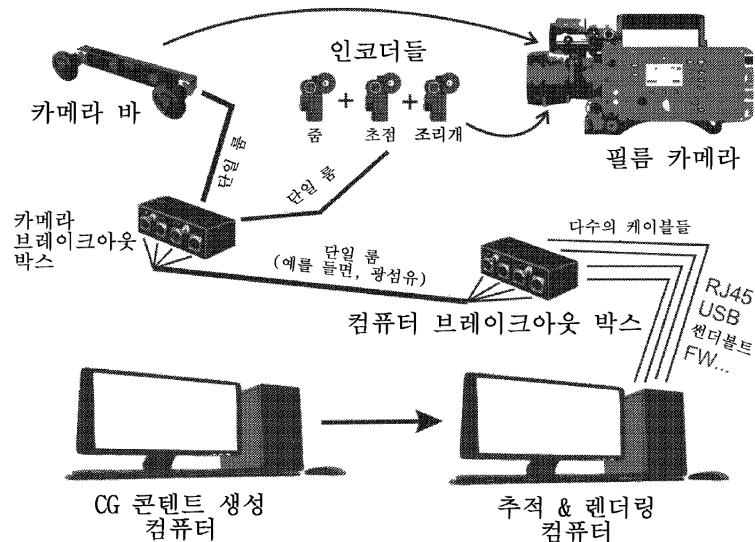
전체 청구항 수 : 총 60 항

(54) 발명의 명칭 컴퓨터 생성된 3D 객체들 및 필름 카메라로부터의 비디오 공급을 실시간으로 믹싱 또는 합성하기 위한 시스템

### (57) 요 약

컴퓨터 생성된 3D 객체들 및 필름 카메라로부터의 비디오 공급을 실시간으로 믹싱 또는 합성하는 방법이 개시되고, 필름 카메라의 몸체는 3D 내에서 움직일 수 있고, 카메라 내의 또는 그것에 부착된 센서들은 카메라의 3D 포지션 및 3D 방향을 정의하거나 3D 포지션이 계산되게 하는 실시간 포지셔닝 데이터를 제공한다.

### 대 표 도



## 특허청구의 범위

### 청구항 1

TV 방송, 영화, 또는 비디오 게임들을 위한 실시간 증강 현실 비디오를 생성하기 위해, 컴퓨터 생성된 3D 객체들 및 비디오 카메라와 같은, 필름 카메라로부터의 비디오 공급을 실시간으로 믹싱 또는 합성하기 위한 시스템에 있어서,

- (a) 상기 필름 카메라의 몸체는 3D 내에서 움직일 수 있고, 상기 필름 카메라 내의 또는 상기 필름 카메라에 직접 또는 간접적으로 부착된 센서들은 상기 필름 카메라의 3D 포지션(position) 및 3D 방향을 정의하거나, 상기 필름 카메라의 3D 포지션 및 3D 방향이 계산되게 할 수 있는 실시간 포지셔닝 데이터를 제공하고,
- (b) 상기 실시간 포지셔닝 데이터는 그 후, 컴퓨터 생성된 3D 객체들을 생성하거나, 다시 불러내거나, 렌더링(rendering)하거나, 수정하도록 상기 시스템에 의해 자동으로 사용되고,
- (c) 결과로 발생한 상기 컴퓨터 생성된 3D 객체들은 그 후, TV 방송, 영화 또는 비디오 게임들에 증강 현실 비디오를 제공하기 위해 상기 필름 카메라로부터의 상기 비디오공급에 믹싱 또는 합성되는, 시스템.

### 청구항 2

제 1 항에 있어서,

상기 컴퓨터 생성된 3D 객체들은 상기 필름 카메라로부터의 상기 실시간 비디오 공급과 함께 실시간으로 믹싱 또는 합성되는, 시스템.

### 청구항 3

제 1 항 또는 제 2 항에 있어서,

상기 필름 카메라의 실시간 줌(zoom), 초점, 및 조리개 설정들은 예를 들면, 종래의 인코더들을 사용하여 측정되고, 상기 실시간 포지셔닝 데이터와 함께 사용되어, 상기 3D 객체들이 3D 장면에서 원하는 위치 및 방향으로 올바르게 렌더링되게 하는, 시스템.

### 청구항 4

제 1 항 내지 제 3 항 중 어느 한 항에 있어서,

상기 센서들은 가속도계 및 자이로(gyro)(6DOF 센서)를 포함하는, 시스템.

### 청구항 5

제 1 항 내지 제 4 항 중 어느 한 항에 있어서,

상기 센서들은 3D 내에서의 병진 가속(translational acceleration)을 측정하는 3-축 가속도계, 3D 내에서의 각속도를 측정하는 3-축 자이로, 및 3D 내에서의 절대적인 방향지시(absolute heading)를 측정하는 자력계를 포함하고, 따라서 9DOF 센서를 구성하는, 시스템.

### 청구항 6

제 1 항 내지 제 5 항 중 어느 한 항에 있어서,

상기 센서들은 구조화된 광 또는 ToF(time-of-flight) 카메라와 같은, 3D 거리 센서를 포함하는, 시스템.

### 청구항 7

제 6 항에 있어서,

상기 3D 거리 센서는 상기 카메라로부터의 비디오 출력 내의 각각의 화소의 심도를 캡처하는, 시스템.

### 청구항 8

제 7 항에 있어서,

가장자리들의 심도는 3D 거리 센서 심도들을 상기 필름 카메라의 고해상도 비디오 공급 상으로 다시 투영함으로써 개선되는, 시스템.

#### 청구항 9

제 1 항 내지 제 8 항 중 어느 한 항에 있어서,

상기 센서들은 상기 필름 카메라에 안전하게 고정될 수 있는 유닛으로 형성되는, 시스템.

#### 청구항 10

제 9 항에 있어서,

상기 유닛은 1개 또는 2개의 위트니스 카메라들(witness cameras)을 포함하는, 시스템.

#### 청구항 11

제 9 항에 있어서,

상기 유닛은 비디오 출력 내의 각각의 화소의 심도를 캡처하는 3D 거리 센서를 포함하는, 시스템.

#### 청구항 12

제 9 항에 있어서,

상기 유닛은 복잡한 장면을 조사하고, 조사되는 상기 3D 장면을 정의하는 데이터를 무선으로 컴퓨터로 송신하기 위해 사용될 수 있는 조사 장치를 형성할 수 있고, 상기 컴퓨터는 후에 상기 장면을 추적하거나 상기 장면의 추적을 복원하는, 시스템.

#### 청구항 13

제 1 항 내지 제 12 항 중 어느 한 항에 있어서,

180도 어안 렌즈들일 수 있는 렌즈들이 구비된, 하나의 단일 위트니스 카메라(평면적인 경우) 또는 2개의 위트니스 카메라들(입체적인 경우)을 포함하는, 시스템.

#### 청구항 14

제 13 항에 있어서,

상기 위트니스 카메라(들)는 상기 필름 카메라로부터 오프-세트(off-set)되고, 상기 오프-세트는 제 1 및 제 2 그룹의 원들을 포함하는 교정 차트를 사용하여 획득되고, 각각의 원은 하나 이상의 다른 원들에 대하여 알려진 위치에 있고, 각각의 원은 블롭(blob) 이미지 알고리즘을 사용하여 인식되는, 시스템.

#### 청구항 15

제 1 항 내지 제 14 항 중 어느 한 항에 있어서,

상기 필름 카메라의 렌즈는 몇개의 원들을 포함하는 교정 차트를 사용하여 광학 왜곡에 대해 교정되고, 각각의 원은 하나 이상의 다른 원들에 대하여 알려진 위치에 있고, 각각의 원은 블롭 이미지 알고리즘을 사용하여 인식되는, 시스템.

#### 청구항 16

제 1 항 내지 제 15 항 중 어느 한 항에 있어서,

상기 필름 카메라의 3D 포지션 및 3D 방향은 상기 센서들로부터의 상기 실시간 3D 포지셔닝 데이터와, 광학 흐름을 사용함으로써 부분적으로 생성된 실세계의 3D 지도를 참조하여 결정되고, 상기 광학 흐름 내에서 위트니스 카메라(들)가 장면을 조사하고 처리기 상에서 실행되는 소프트웨어가 상기 장면에 수동으로 또는 인위적으로 추가된 것이 아닌 상기 장면 내의 자연스러운 마커들('특징 점들')을 검출하는, 시스템.

**청구항 17**

제 1 항 내지 제 16 항 중 어느 한 항에 있어서,

추적될 상기 장면을 오직 조사하는('인스턴트(instant) 조사'로 불리는) 별개의 단계 없이 상기 시스템이 완전히 개시되게 하기 위해 1개 또는 2개의 (적어도 100 fps와 같은) 고속 워트니스 카메라들을 사용하지만, 대신 조사는 상기 카메라가 비디오를 캡쳐하기 위해 사용되는 동안 계속 발생하는, 시스템.

**청구항 18**

제 17 항에 있어서,

상기 2개의 고속 워트니스 카메라들은 소프트웨어가 상기 이미지들을 처리할 수 있게 하고, 상기 카메라 시스템이 전혀 움직이지 않더라도, 인스턴트 3D 점 클라우드를 생성할 수 있게 하는 (예를 들면, 상기 2개의 카메라와 등극선 기하 간의 분리에 대한 지식을 사용하여 상기 장면 내의 다수의 점들을 3D 공간 내의 그들의 포지션에 연관시키는) 입체적인 시스템을 형성하는, 시스템.

**청구항 19**

제 17 항에 있어서,

상기 3D 점 클라우드 내의 각각의 화소의 심도는 각각의 입체적인 워트니스 카메라 및 등극선 탐색 알고리즘으로부터 획득된 대응하는 2D 질감 패치들(patches)을 사용하여 획득되는, 시스템.

**청구항 20**

제 1 항 내지 제 19 항 중 어느 한 항에 있어서,

상기 워트니스 카메라 시스템으로부터의 광학 흐름 데이터와 상기 하드웨어 센서들로부터의 상기 실시간 포지셔닝 데이터를 조합하는 결합 알고리즘을 실행시키는, 시스템.

**청구항 21**

제 20 항에 있어서,

상기 결합 알고리즘은 상기 카메라의 포지션 및 방향을 결정하기 위해 가속도계, 자이로스코프, 자력계, 3D 거리 센서를 포함할 수 있는 상기 센서들 모두로부터의 출력들을 통합하고 상기 센서들 모두를 재교정하기 위해 확장형 칼만(Kalman) 필터 예측/정정 기술에 기초하는, 시스템.

**청구항 22**

제 21 항에 있어서,

상기 확장형 칼만 필터 결합 알고리즘은 각각의 센서로부터의 상기 데이터를 어떻게 병합할지를 결정할 때, 각각의 센서로부터의 상기 출력에 연관된 신뢰 수준 데이터를 사용하는, 시스템.

**청구항 23**

제 1 항 내지 제 22 항 중 어느 한 항에 있어서,

상기 워트니스 카메라(들)에 의해 생성된 키프레임들은 상기 가시적 추적 처리의 부분이고 상기 워트니스 카메라 비디오 공급의 4개의 상이한 해상도 수준들에서 컴퓨팅된 실시간 이미지들인, 시스템.

**청구항 24**

제 1 항 내지 제 23 항 중 어느 한 항에 있어서,

(a) 가상 인물들, 객체들, 및 장소들의 3D 컴퓨터 생성된 애니메이션을 제공하는 콘텐트 생성 컴퓨터 및 (b) (상기 콘텐트 생성 컴퓨터로부터 분리될 수 있거나 분리되지 않을 수 있는) 렌더링 컴퓨터를 포함하고, 상기 필름 카메라의 3D 포지션을 정의하는 상기 실시간 포지셔닝 데이터는 상기 비디오 공급에서 보여지는 상기 장면의 자연스러운 부분을 형성하기 위해 상기 필름 카메라로부터의 상기 비디오 공급에 실시간으로 삽입되고 상기 비디오 공급에 띄워질 수 있는 실시간, 컴퓨터 생성된 3D 객체들로 하여금 생성되도록 하기 위해, 상기 콘텐트 생

성 컴퓨터 및 상기 렌더링 컴퓨터 중 하나 또는 그 둘 모두에 의해 사용되는, 시스템.

### 청구항 25

제 1 항 내지 제 24 항 중 어느 한 항에 있어서,

상기 컴퓨터 생성된 3D 객체들은 상기 콘텐트 생성 컴퓨터에 의해 결정된 방식으로 상기 장면 내의 어느 곳으로나 움직일 수 있고 그들의 모양 및 겉모습을 바꿀 수 있는 애니메이션들인, 시스템.

### 청구항 26

제 1 항 내지 제 25 항 중 어느 한 항에 있어서,

상기 컴퓨터 생성된 3D 객체들은 상기 장면에 믹싱될 때 현실적인 방식으로 움직이는(예를 들면, 달리기, 춤추기, 걷기, 싸우기, 날기, 점프하기...) 사람 또는 생명체들의 애니메이션된 인물들인, 시스템.

### 청구항 27

제 1 항 내지 제 26 항 중 어느 한 항에 있어서,

상기 카메라 포지셔닝 또는 추적 데이터는 또한 후반-제작 CGI를 용이하게 하기 위해 후반-제작에서 사용될 수 있게 만들어진, 시스템.

### 청구항 28

제 1 항 내지 제 27 항 중 어느 한 항에 있어서,

3D 거리 센서는 재구성된 3D 점과 연관된 심도 측정의 정확도를 개선시키거나, 상기 재구성된 3D 점을 거부하기 위해 사용되는, 시스템.

### 청구항 29

제 28 항에 있어서,

상기 3D 거리 센서는 동적인 페색을 가능하게 하고 녹색 스테이지의 궁극적인 사용을 억제하기 위해 실시간 심도 키잉(keying)을 위해 사용되는, 시스템.

### 청구항 30

제 1 항 내지 제 29 항 중 어느 한 항에 있어서,

상기 장면 내에 위치된, 알려진 크기의 및 알려진 패턴으로 커버된 계시판과 같은, 소형 카메라 등록 객체를 사용하여, 상기 검출된 패턴의 코너가 상기 3D 점 클라우드(및 따라서 상기 세계)에 대한 근원으로서 처리되게 하는, 시스템.

### 청구항 31

제 30 항에 있어서,

상기 카메라 등록 객체는 실제 수직으로 정렬되고 블롭 이미지 인식 알고리즘을 사용하여 인식되는 알려진 크기의 적어도 2개의 구들을 포함하는, 시스템.

### 청구항 32

제 1 항 내지 제 31 항 중 어느 한 항에 있어서,

자기 N극을 나타내기 위한 자력계, 중력 방향을 나타내기 위한 (그리고 그에 따라 실제 수직을 제공하는) 가속도계, 상기 시스템이 위로/아래로 기울어졌는지나 좌측 또는 우측으로 팬(pan)되었거나 광학 축에 대하여 회전되었는지를 나타내기 위한 자이로, 및 시작 포지션으로부터의 3D 내에서의 병진이 추론되게 하기 위한 3-축 가속도계를 포함하는, 시스템.

### 청구항 33

제 1 항 내지 제 32 항 중 어느 한 항에 있어서,

소프트웨어는 추적 손실들을 상당히 줄이고 추적 정확도를 높이기 위해 상기 3D 지도에 균일하게 분포된 점 클라우드를 생성하는 것을 시도하여, 시차(parallax)가 보다 많이 생성되어, 추정된 카메라 포지션이 보다 정확한, 시스템.

#### **청구항 34**

제 1 항 내지 제 33 항 중 어느 한 항에 있어서,

카메라 추적 시스템은 상기 필름 카메라에 무선으로 연결될 수 있으므로, 상기 3D 점 클라우드를 생성할 때 상기 세트 주변을 빠르게 움직일 수 있는, 시스템.

#### **청구항 35**

제 1 항 내지 제 34 항 중 어느 한 항에 있어서,

감독/촬영기사가 상기 필름 카메라에 부착된 상기 카메라 추적 시스템을 추적, 팬(pan), 기울일 때, 상기 필름 카메라를 추적하는 인스턴트 조사(입체적인 경우) 또는 3D 장면의 실시간 조사(평면적인 경우)를 조합하는 카메라 추적 시스템을 제공하는, 시스템.

#### **청구항 36**

제 1 항 내지 제 35 항 중 어느 한 항에 있어서,

회전, 병진, 초점, 조리개, 초점 거리를 포함하는 필름 카메라 추적에 대한 모든 양태들을 완벽히 자동화하고; 상기 비디오에 믹싱될 상기 3D 컴퓨터 생성된 콘텐트의 스케일링, 포지셔닝 및 방향설정을 자동화하는, 시스템.

#### **청구항 37**

제 1 항 내지 제 36 항 중 어느 한 항에 있어서,

장면의 실시간, 계속적인 조사가 상기 장면을 정의하는 보다 완벽한 점 클라우드를 생성하게 하는, 시스템.

#### **청구항 38**

제 1 항 내지 제 37 항 중 어느 한 항에 있어서,

추적의 복원을 용이하게 하기 위해, 예를 들면, ORB를 사용하는 회전 불변 서술자들(rotation invariant descriptors)을 상기 장면에서 검출된 특징 점들에 부착하는, 시스템.

#### **청구항 39**

제 1 항 내지 제 38 항 중 어느 한 항에 있어서,

이전에 올바르게 컴퓨팅되거나 확인된 포지션을 사용하여 상기 필름 카메라의 다음 포지션을 예측하기 위해 상기 센서들에 의해 제공된 정보에 연관된 고정 속도 모델을 사용하는, 시스템.

#### **청구항 40**

제 39 항에 있어서,

점 매칭 알고리즘이 상기 위트니스 카메라 시스템으로부터의 상기 실시간 비디오 공급 내의 식별된 점들과 새로운 3D 점 클라우드 내의 상기 투영된 점들을 매칭하게 하도록, 3D 점 클라우드를 현재의 프레임 상에 다시 투영하기 위해 상기 예측을 사용하는, 시스템.

#### **청구항 41**

제 40 항에 있어서,

상기 위트니스 카메라 시스템으로부터의 상기 실시간 비디오 공급 내의 식별된 상기 점들과 상기 새로운 3D 점 클라우드 내의 상기 투영된 점들 간의 오류를 최소화하기 위해 카메라 추적에 대해 레벤버그-마콰트(Levenberg-Marquardt) 최소화 방식을 사용하는, 시스템.

**청구항 42**

제 1 항 내지 제 41 항 중 어느 한 항에 있어서,

사용자는 3D 가비지 매팅 마스크들(3D Garbage Matting masks)과 같은, 3D 마스크들을 정의하기 위해 상기 카메라 추적 시스템에 의해 생성된 3D 점 클라우드를 사용할 수 있는, 시스템.

**청구항 43**

제 1 항 내지 제 42 항 중 어느 한 항에 있어서,

상기 3D 객체들은 정적인 객체들, 동적인 애니메이션들, 가상 세계들, 가상 사람, 가상 건물들, 가상 정치, 가상 영화 세트들, 및 애니메이션 데이터베이스 내의 임의의 데이터를 포함하는, 시스템.

**청구항 44**

제 1 항 내지 제 43 항 중 어느 한 항에 있어서,

상기 필름 카메라 및 위트니스 카메라는 플레싱 LED와 연관된 광 세기 커브들을 비교함으로써 같이, 변조된 광원을 사용하여 프레임 획득 지역에 대하여 교정되는, 시스템.

**청구항 45**

제 1 항 내지 제 44 항 중 어느 한 항에 있어서,

상기 필름 카메라는 크레인 카메라(crane camera), 스테디 캠(steady cam), 핸드-헬드-카메라(hand-held-camera), 돌리 장착 카메라(dolly mounted camera), 삼각대-장착 카메라(tripod-mounted camera), 스마트폰, 및 증강 현실 안경 중 임의의 것인, 시스템.

**청구항 46**

실시간, 컴퓨터 생성된 3D 객체들 및 필름 카메라로부터의 비디오 공급을 믹싱 또는 합성하는 방법에 있어서,

상기 필름 카메라의 몸체는 3D 내에서 움직일 수 있고 상기 카메라 내의 또는 상기 카메라에 부착된 센서들은 상기 카메라의 3D 포지션 및 3D 방향을 정의하거나 상기 3D 포지션이 계산될 수 있게 하는 실시간 포지셔닝 데이터를 제공하는, 3D 객체들 및 비디오 공급을 믹싱 또는 합성하는 방법.

**청구항 47**

TV 방송, 영화, 또는 비디오 게임들을 위한 증강 현실 비디오를 생성하기 위해, 실시간, 컴퓨터 생성된 3D 객체들 및 비디오 카메라와 같은, 필름 카메라로부터의 비디오 공급을 믹싱 또는 합성하기 위한 방법에 있어서,

(a) 상기 필름 카메라의 몸체는 3D 내에서 움직일 수 있고, 상기 필름 카메라 내의 또는 상기 필름 카메라에 직접 또는 간접적으로 부착된 센서들은 상기 필름 카메라의 3D 포지션 및 3D 방향을 정의하거나, 상기 필름 카메라의 상기 3D 포지션 및 3D 방향이 계산되게 할 수 있는 실시간 포지셔닝 데이터를 제공하고,

(b) 상기 실시간 포지셔닝 데이터는 그 후 컴퓨터 생성된 3D 객체들을 생성하거나, 다시 불러내거나, 웹더링하거나, 수정하기 위해 시스템에 의해 자동으로 사용되고,

(c) 결과로 발생한 상기 컴퓨터 생성된 3D 객체들은 그 후, TV 방송, 영화 또는 비디오 게임들에 증강 현실 비디오를 제공하기 위해 상기 필름 카메라로부터 상기 비디오 공급에 믹싱 또는 합성되는, 3D 객체들 및 비디오 공급을 믹싱 또는 합성하기 위한 방법.

**청구항 48**

감독(또는 사진 감독)이 카메라 내의 사전-제작 컴퓨터 생성된 3D 자산들, 전형적으로 사전 가시화 또는 시각 효과들의 자산들을 상기 카메라에 의해 캡처된 비디오 이미지들 또는 실제 액션 영화 플레이트(live action film plate) 상으로 실시간으로 조작(frame-up)하게 하기 위해 사용된 제 1 항 내지 제 47 항 중 어느 한 항에 따른 방법들 또는 시스템들에 있어서,

상기 감독이 상기 감독에 의해 보여지는 것과 같은 상기 비디오로 자동으로 믹싱되는 상기 컴퓨터 생성된 3D 자

산들과 함께 실시간으로 가능한 카메라 각도들 및 움직임들을 조사하게 하는, 방법들 또는 시스템들.

#### 청구항 49

제 1 항 내지 제 48 항 중 어느 한 항에 있어서,

후반-제작 처리들에 매칭-움직임(match-move) 데이터를 제공하기 위해 실시간 포지셔닝 데이터가 기록되고 시간-코드 스탬핑되는, 방법들 또는 시스템들.

#### 청구항 50

제 1 항 내지 제 49 항 중 어느 한 항에 있어서,

가상 객체들 또는 이미지들이 방송 비디오 스트림에 삽입될 수 있게 하기 위해 사용된, 방법들 또는 시스템들.

#### 청구항 51

제 1 항 내지 제 50 항 중 어느 한 항에 있어서,

- 스튜디오 카메라들에 대한 실시간 추적
- 스테디캠에 대한 실시간 추적
- 크레인-장착 카메라들에 대한 실시간 추적
- 돌리-장착 카메라들에 대한 실시간 추적
- 스테디캠에 대한 실시간 추적
- OB(Outside Broadcast)에 대한 실시간 추적
- 2D 후반-제작을 위한 실시간 데이터(예를 들면, 추적 데이터)의 사용
- 3D 입체 콘텐트에 대한 후반-변환을 위한 실시간 데이터(예를 들면, 추적 데이터)의 사용
- 원시(native) 3D 입체 콘텐트를 위한 실시간 데이터(예를 들면, 추적 데이터)의 사용
- 3D 그래픽들 삽입
- 스튜디오-안 또는 세트-상 간접 광고(on-set product placement)를 위한 3D 그래픽들 삽입
- OB를 위한 3D 그래픽들 삽입
- 다른 후원된 이미지들에 대한 3D 그래픽들 삽입
- 시청자-위치 특정인 3D 그래픽들 삽입
- 시청자-특정인 3D 그래픽들 삽입
- 시간-특정인 3D 그래픽들 삽입
- 군중 장면들을 채우기 위한 3D 그래픽들 삽입
- 초록색 화면 대체를 위한 3D 그래픽들 삽입
- 박물관들에서 및 문화, 역사, 또는 자연 현장들 내의 번역 센터들에서 학습을 돋기 위한 교육 콘텐트의 3D 그래픽들 삽입, 및
- 장면 내의 객체들의 절대적인 또는 상대적인 크기의 측정 중 하나 이상을 가능하게 하는, 방법 또는 시스템.

### 청구항 52

제 1 항 내지 제 51 항 중 어느 한 항에 있어서,

필름 카메라가:

- 표준 테더(tether)를 갖는 모든 카메라들
- 전략적인 광섬유 접속을 필요로 하는 카메라들, 및
- RF/무선 접속을 필요로 하는 카메라 중 하나인, 방법들 또는 시스템들.

### 청구항 53

제 1 항 내지 제 52 항 중 어느 한 항에 있어서,

- 영화/TV (라이브(live) 아님)
- 광고 (라이브 아님)
- 광고들 라이브
- 방송 (스포츠 아님)
- 방송 OB
- 스포츠 스튜디오 기반
- 스포츠 OB 기반
- TV 간접 광고 라이브
- 인터넷 사용 (라이브 아님)
- 인터넷 라이브
- 인터넷 라이브 영역 기반
- 인터넷 간접 광고 라이브
- 박물관/문화유산 콘텐트
- 박물관/문화유산 광고들
- 건축학, 및
- 게임들 중 하나에 사용되는, 방법 또는 시스템들.

### 청구항 54

앞서 정의된 바와 같고 증강 현실 이미지들이 스마트폰, 및 증강 현실 안경을 포함하는, 임의의 디스플레이 장치에 의해 디스플레이되게 하기 위해 사용된 제 1 항 내지 제 53 항 중 어느 한 항에 따른 방법들 또는 시스템들에 있어서,

상기 증강 현실 이미지들의 결모습은 실시간 포지셔닝 데이터에 따라 자동으로 바뀌는, 방법들 또는 시스템들.

### 청구항 55

실시간, 컴퓨터 생성된 3D 객체들이 카메라로부터의 비디오 공급과 막싱되는 필름, 영화, TV 쇼 또는 비디오 게

임에 있어서,

상기 카메라의 몸체는 3D 내에서 움직일 수 있고, 상기 카메라 내의 또는 상기 카메라에 부착된 센서들은 상기 카메라의 3D 포지션을 정의하거나 상기 3D 포지션이 계산되게 하는 실시간 포지셔닝 데이터를 제공하는, 필름, 영화, TV 쇼, 또는 비디오 게임.

### 청구항 56

종래의 카메라에 부착되도록 적응된 장면 조사 및 추적 장치에 있어서,

상기 카메라의 몸체는 3D 내에서 움직일 수 있고, 상기 장면 조사 및 추적 장치 내의 하드웨어 센서들은 상기 카메라의 3D 포지션을 정의하거나 상기 3D 포지션이 계산되게 하는 실시간 포지셔닝 데이터를 제공하는, 장면 조사 및 추적 장치.

### 청구항 57

카메라 내의 또는 상기 카메라에 부착된 센서들을 포함하는 핸드-헬드 또는 휴대용 카메라에 있어서,

3D 기준 프레임에 대한 상기 카메라의 3D 포지션 및 3D 방향을 정의하거나 광학 흐름 데이터와 같은, 다른 데이터를 분석하는 시스템의 전체 또는 부분으로서 상기 3D 포지션이 계산되게 하는 실시간 포지셔닝 데이터를 제공하는, 핸드-헬드 또는 휴대용 카메라.

### 청구항 58

입체적 위트니스 카메라 시스템을 포함하거나 상기 입체적 위트니스 카메라 시스템에 부착된 필름 카메라에 있어서,

상기 위트니스 카메라 시스템은 광각(예를 들면, 180도) 입체적 이미지를 생성하고, 상기 필름 카메라는 상기 카메라 시스템이 전혀 추적/움직여지지 않으면서, 인스턴트 3D 점 클라우드를 생성하도록 소프트웨어가 상기 이미지를 처리하게 하는, 필름 카메라.

### 청구항 59

필름 카메라에 부착되기 위한 카메라 추적 시스템에 있어서,

위트니스 카메라 시스템으로서, 입체적 이미지를 생성하고, 소프트웨어가 이미지를 처리하게 하고, 상기 카메라 시스템이 전혀 움직이지 않으면서, 인스턴트 3D 점 클라우드를 생성하게 하고 상기 필름 카메라의 실시간 추적(포지션, 방향, 줌, 초점, 및 조리개)을 제공하게 하는, 상기 위트니스 카메라 시스템을 포함하는, 카메라 추적 시스템.

### 청구항 60

증강 현실 안경 내에 또는 상기 증강 현실 안경에 부착된 센서들을 포함하는 상기 증강 현실 안경에 있어서,

3D 기준 프레임에 대한 상기 증강 현실 안경의 3D 포지션 및 3D 방향을 정의하거나, 광학 흐름 데이터와 같은, 다른 데이터를 분석하는 시스템의 전체 또는 부분으로서 상기 3D 포지션이 계산되게 하는 실시간 포지셔닝 데이터를 제공하는, 증강 현실 안경.

## 명세서

### 기술분야

[0001] 본 발명은 TV 방송, 영화, 또는 비디오 게임들을 위한 실시간 증강 현실 비디오를 생성하기 위해, 컴퓨터 생성된 3D 객체들과 비디오 카메라와 같은 필름 카메라로부터의 비디오 공급을 실시간으로 딱싱 또는 합성하기 위한 시스템에 관한 것이다.

### 배경기술

[0002] 지난 20년 동안, 본 분야에서 상당한 연구 및 상업적 활동이 있었고, 시청자들이 보는 가상의 제 1 다운 라인을 피치(pitch) 상에 중첩시켜 추가하는, 미국 풋볼과 같은 스포츠 경기들에 대한 텔레비전 시청 개선들을 개발해

온, 스포츠비전, 인크.와 같은 회사들로부터의 비디오 삽입 또는 증강 현실 시스템에 대한 기준이 만들어졌다. 이 영역에서 시스템들을 개발한 다른 회사들은 PVI, Inc.가 포함된다. 알려진 시스템들의 하나의 공통 특징은, 그 시스템들이 원칙적으로, 카메라가 실세계 장면 중 어느 부분을 가리키고 있는지를 결정하기 위해 카메라로부터의 비디오 공급 내의 특징들을 분석하는 것에 의존한다는 것인데; 장면에서 카메라가 가리키는 곳을 시스템이 알게 되면, 객체의 포지션(position) 및 방향이 그것을 장면의 자연스러운 부분으로 보이게 하는 방식으로 비디오 공급 상에 (상기 가상의 제 1 다운 라인과 같은) 컴퓨터 생성된 객체들을 추가 또는 합성할 수 있도록, 시스템은 미리 그 장면의 3D 지도를 구축할 것이다. 이 방식에서 오직 광학 흐름에만 의존하는 것에 대한 하나의 단점은 이러한 시스템들이 신뢰적이지 않을 수 있다는 것이다.

[0003] 다른 시스템들은 순수한 마커(marker)-기반 접근법들(예를 들면, 라이트크래프트 기술들(Lightcraft Technologies))에 의존한다. 그들은 시스템에 의해 검출되도록 세트 상에 실제 물리적 마커들(1m x 1m 큐)을 놓도록 오퍼레이터(operator)에게 요청한다. 그것은 매우 비효율적인데, 이는 준비하는데 몇 시간 또는 몇 일 걸리고, 이는 간혹 영화 제작을 위해 작용하지 않을 가능성성이 크기 때문이다. 그것은 또한, 물리적 마커들이 항상 그들의 시스템의 시계(field of view) 내에 있어야 하기 때문에 많은 제약사항들을 갖는다. 또한 **색인 1**에 인용된 것들을 과학적으로 참조할 수 있다.

### 발명의 내용

#### 해결하려는 과제

[0004] 이 기술은 실시간 컴퓨터 생성된 3D 객체들과, 영화 촬영 세트, TV 방송, 또는 비디오 게임 상의 실제 카메라(본원의 나머지 부분에서 필름 카메라로 불리는)로부터의 실제 비디오 공급을 믹싱하는 것에 대한 것이다. 이 시스템의 주요 목적들은:

[0005] • 영화 감독들에게 가상 인물들 및 환경의 실시간 프레임 도구를 제공하고,

[0006] • 실제 비디오에 바로 병합된 가상 객체들의 사진-현실적(photo-realistic) 결과를 실시간으로 디스플레이하고,

[0007] • 세트-상 촬영 및 영화 후반-제작에 필름 카메라의 포지션의 정확한 추정(카메라 추적으로 불리는)을 제공하고 (모든 필름 카메라들은 손잡이가 있음: 크레인 카메라들, 스테디 캠, 핸드-헬드 카메라, 삼각대 상의 카메라 등),

[0008] • 실제 장면의 3D 기하학 모델링의 근사치를 제공하여, 실제 객체들(및/또는 연기자들)과 가상 환경(3D 객체들, 인물들 등) 사이의 폐색을 처리하기 위한 것이다.

#### 과제의 해결 수단

[0009] 본 발명은 하나의 양태에서, TV 방송, 영화, 또는 비디오 게임들을 위한 증강 현실 비디오를 생성하기 위해, 실시간, 컴퓨터 생성된 3D 객체들 및 비디오 카메라와 같은, 필름 카메라로부터의 비디오 공급을 믹싱 또는 합성하기 위한 시스템에 있어서,

[0010] (a) 필름 카메라의 몸체는 3D 내에서 움직일 수 있고, 필름 카메라 내의 및 필름 카메라에 직접 또는 간접적으로 부착된 센서들은 필름 카메라의 3D 포지션 및 3D 방향을 정의하거나, 필름 카메라의 3D 포지션 및 3D 방향이 계산되게 할 수 있는 실시간 포지셔닝 데이터를 제공하고,

[0011] (b) 실시간 포지셔닝 데이터는 그 후 컴퓨터 생성된 3D 객체들을 생성하거나, 다시 불러내거나, 렌더링(rendering)하거나, 수정하기 위해 시스템에 의해 자동으로 사용되고,

[0012] (c) 결과로 발생한 컴퓨터 생성된 3D 객체들은 그 후, TV 방송, 영화 또는 비디오 게임들에 증강 현실 비디오를 제공하기 위해 필름 카메라로부터의 비디오 공급에 믹싱 또는 합성된다.

[0013] 선택적인 특징들(그들 중 몇몇 또는 그들 모두는 서로 조합될 수 있음)에는 다음이 포함된다:

[0014] • 컴퓨터 생성된 3D 객체들은 필름 카메라로부터의 실시간 비디오 공급에 실시간으로 믹싱 또는 합성됨.

[0015] • 필름 카메라의 실시간 줌(zoom), 초점, 조리개 설정들이 측정되고, 실시간 포지셔닝 데이터와 함께

사용되어, 3D 객체들이 3D 장면에서 원하는 위치 및 방향으로 올바르게 렌더링되게 함.

- 센서들은 가속도계 및 자이로(gyro)(6DOF 센서)를 포함함.

- 센서들은 3D 내에서의 병진 가속(translational acceleration)을 측정하는 3-축 가속도계, 3D 내에서의 각 속도를 측정하는 3-축 자이로, 및 3D 내에서의 절대적인 방향지시(absolute heading)를 측정하는 자력계를 포함하고, 따라서 9DOF 센서를 구성함.

- 센서들은 구조화된 광 또는 ToF(time-of-flight) 카메라와 같은, 3D 거리 센서를 포함함.

- 3D 거리 센서는 카메라로부터의 비디오 출력 내의 각각의 화소의 심도를 캡쳐함.

- 가장자리들의 심도는 3D 거리 센서 심도들을 필름 카메라로부터의 고해상도 비디오 공급 상으로 다시 투영함으로써 개선됨.

- 센서들은 필름 카메라에 안전하게 고정될 수 있는 유닛(unit)으로 형성됨.

- 유닛은 1개 또는 2개의 위트니스 카메라들(witness cameras)을 포함함.

- 유닛은 비디오 출력 내의 각각의 화소의 심도를 캡쳐하는 3D 거리 센서를 포함함.

- 유닛은 6DOF 또는 9DOF 센서를 포함함.

- 유닛은 그것이 상이한 크기들 및 디자인들의 필름 카메라들 및 로드들(rods)에 고정되게 허용하도록 하기 위해 스위칭 가능한 행어(hanger)들을 사용함.

- 유닛은 다른 형태들의 센서들을 포함하도록 확장가능함.

- 유닛은 복잡한 장면을 조사(survey)하고, 조사되는 3D 장면을 정의하는 데이터를 무선으로 컴퓨터로 송신하기 위해 사용될 수 있는 조사 장치를 형성할 수 있으며, 상기 컴퓨터는 후에 장면을 추적하거나 장면의 추적을 복원함.

- 필름 카메라는 카메라 줌, 초점, 및 조리개를 읽기 위해 종래의 인코더들을 포함함.

- 시스템은 180도 어안 렌즈들일 수 있는 렌즈들을 구비한 하나의 단일 위트니스 카메라(평면적인 경우) 또는 2개의 위트니스 카메라들(입체적인 경우)을 포함함.

- 위트니스 카메라(들)는 필름 카메라로부터 오프-세트(off-set)되고, 오프-세트는 제 1 및 제 2 그룹의 원들을 포함하는 교정 차트를 사용하여 획득되고, 각각의 원은 하나 이상의 다른 원들에 대하여 알려진 위치에 있고, 각각의 원은 블롭(blob) 이미지 알고리즘을 사용하여 인식됨.

- 필름 카메라의 렌즈는 몇 개의 원들을 포함하는 교정 차트를 사용하여 광학 왜곡에 대해 교정되고, 각각의 원은 하나 이상의 다른 원들에 대하여 알려진 위치에 있고, 각각의 원은 블롭 이미지 알고리즘을 사용하여 인식됨.

- 필름 카메라는 크레인 카메라(crane camera); 스테디 캠(steady cam); 핸드-핸드-카메라(hand-held-camera); 돌리 장착 카메라(dolly mounted camera), 삼각대-장착 카메라(tripod-mounted camera), 스마트폰, 증강 현실 안경 중 임의의 것임.

- 필름 카메라의 3D 포지션 및 방향은 센서들로부터의 실시간 3D 포지셔닝 데이터와, 위트니스 카메라(들)가 장면을 조사하고 처리기 상에서 실행되는 소프트웨어가 장면에 수동으로 또는 인위적으로 추가된 것이 아닌 자연스러운 마커들("특징 점들")을 그 장면에서 검출하는 광학 흐름을 부분적으로 사용함으로써 생성된 실세계의 3D 지도를 참조하여 결정됨.

- 시스템은 추적될 장면을 오직 조사하는('인스턴트(instant) 조사'로 불리는) 별개의 단계가 없지만, (대신, 카메라가 비디오를 캡처하기 위해 사용되는 동안 계속 발생되는 조사를 통해) 시스템이 완벽히 개시되게 하기 위해 1개 또는 2개의 고속(적어도 100 fps와 같은) 위트니스 카메라들을 사용함.

- [0035] • 입체적인 시스템은, 소프트웨어가 이미지들을 처리하게 하고, 카메라 시스템이 전혀 움직이지 않더라도 인스턴트 3D 점 클라우드(cloud)를 생성하게 함(예를 들면, 2개의 카메라와 등극선 기하 사이의 분리에 대한 지식을 사용하여 장면 내의 다수의 점들을 3D 공간 내의 그들의 포지션에 연관시키는).
- [0036] • 3D 점 클라우드 내의 각각의 화소의 심도는 각각의 입체적인 위트니스 카메라 및 등극선 탐색 알고리즘으로부터 획득된 대응하는 2D 질감 패치들(patches)을 사용하여 획득됨.
- [0037] • 시스템은 위트니스 카메라 시스템으로부터의 광학 흐름 데이터를 하드웨어 센서들로부터의 실시간 포지셔닝 데이터와 조합하는 결합 알고리즘을 실행함.
- [0038] • 결합 알고리즘은 카메라의 포지션 및 방향을 결정하기 위해, 가속도계, 자이로스코프(gyroscope), 자력계, 3D 거리 센서를 포함할 수 있는 모든 센서들로부터의 출력들을 통합하고 재교정하기 위해 확장형 칼만(Kalman) 필터 예측/정정 기술에 기초함.
- [0039] • EKF 결합 알고리즘은 각각의 센서로부터의 데이터를 어떻게 병합할지를 결정할 때, 각각의 센서로부터의 출력과 연관된, 신뢰 수준 데이터를 사용함.
- [0040] • 위트니스 카메라 시스템에 의해 생성된 키프레임들은 가시적 추적 처리의 부분이고 위트니스 카메라 비디오 공급의 4개의 상이한 해상도 수준들에서 컴퓨팅된 실시간 이미지들임.
- [0041] • 시스템은 (a) 가상 인물들, 객체들, 및 장소들의 3D 컴퓨터 생성된 애니메이션을 제공하는 콘텐트 생성 컴퓨터, 및 (b) (콘텐트 생성 컴퓨터로부터 분리될 수 있거나 분리되지 않을 수 있는) 렌더링 컴퓨터를 포함하고, 필름 카메라의 3D 포지션을 정의하는 실시간 포지셔닝 데이터는, 비디오 공급에서 보여지는 장면의 자연스러운 부분을 형성하기 위해 필름 카메라로부터의 그 비디오 공급에 실시간으로 삽입되고 그것에 맞설 수 있는 실시간 컴퓨터 생성된 3D 객체들로 하여금 생성되도록 하기 위해, 콘텐트 생성 컴퓨터 및 렌더링 컴퓨터 중 하나 또는 그 둘 모두에 의해 사용될 수 있음.
- [0042] • 컴퓨터 생성된 3D 객체들은 콘텐트 생성 컴퓨터에 의해 결정된 방식으로 장면 내의 어느 곳으로나 움직일 수 있고 그들의 모양 및 겉모습을 바꿀 수 있는 애니메이션들을 임.
- [0043] • 컴퓨터 생성된 3D 객체들은 장면에 맞설 때 현실적인 방식으로 움직이는 (예를 들면, 달리기, 춤추기, 걷기, 싸우기, 날기, 점프하기...) 사람 또는 생명체들의 애니메이션된 인물들임.
- [0044] • 카메라 포지셔닝 또는 추적 데이터는 또한 후반-제작 CGI를 용이하게 하기 위해 후반-제작에서 사용될 수 있게 만들어짐.
- [0045] • 3D 거리 센서는 재구성된 3D 점과 연관된 심도 측정의 정확도를 개선시키거나, 그 재구성된 3D 점을 거부하기 위해 사용됨.
- [0046] • 3D 거리 센서는 실시간 심도 키잉(keying)이 녹색 스테이지의 동적인 폐색을 가능하게 하고 궁극적인 사용을 금지하기 위해 사용됨.
- [0047] • 시스템은 장면 내에 위치된, 알려진 크기의 및 알려진 패턴으로 커버된 계시판과 같은 소형 카메라 등록 객체를 사용하여, 검출된 패턴의 코너가 3D 점 클라우드(및 그에 따라 상기 세계)에 대한 근원으로서 처리되게 함.
- [0048] • 카메라 등록 객체는 블롭 이미지 등록 알고리즘을 사용하여 인식되고 실제 수직으로 정렬된 알려진 크기의 적어도 2개의 구들을 포함함.
- [0049] • 시스템은 자기 N극을 나타내기 위한 자력계, 중력 방향을 나타내기 위한 (그리고 그에 따라 실제 수직을 제공하는) 가속도계, 시스템이 위로/아래로 기울어졌는지나 좌측 또는 우측으로 팬(pan)되었거나 광학 축에 대하여 회전되었는지를 나타내기 위한 자이로, 및 시작 포지션으로부터의 3D 내에서의 병진(translation)이 추론되게 하기 위한 3-축 가속도계를 포함함.

[0050] • 소프트웨어는 추적 손실들을 상당히 줄이고 추적 정확도를 높이기 위해 3D 지도에 균일하게 분포된 점 클라우드를 생성하고자함(시차(parallax)가 보다 많이 생성될수록, 추정된 카메라 포지션은 보다 정확함).

[0051] • 카메라 추적 시스템은 카메라 시스템에 무선으로 연결될 수 있으므로, 3D 점 클라우드를 생성할 때 세트 주변을 빠르게 움직일 수 있다. 이전의 위트니스 카메라 시스템들과는 달리, 세트에 걸친 케이블들을 뒤쫓을 필요가 없다.

[0052] • 카메라 추적 시스템은, 감독/촬영기사가 필름 카메라에 부착된 카메라 추적 시스템을 추적, 팬, 기울일 때, 필름 카메라를 추적하는 인스턴트 조사(입체적인 경우) 또는 3D 장면의 실시간 조사(평면적인 경우)를 조합함.

[0053] • 시스템은 회전, 병진, 초점, 조리개, 초점 거리를 포함하는 필름 카메라 추적에 대한 모든 양태들을 완전히 자동화하고; 비디오에 믹싱될 3D 컴퓨터 생성된 콘텐트의 스케일링, 포지셔닝 및 방향설정을 자동화함.

[0054] • 시스템은 장면의 실시간, 계속적인 조사가 장면을 정의하는 보다 완벽한 점 클라우드를 생성하게 함.

[0055] • 시스템은 추적의 복원을 용이하게 하기 위해, 예를 들면, ORB를 사용하는 회전 불변 서술자들(rotation invariant descriptors)을 장면에서 검출된 특정 점들에 부착함.

[0056] • 시스템은 이전에 올바르게 컴퓨팅되거나 확인된 포지션을 사용하여 필름 카메라의 다음 포지션을 예측하기 위해 센서들에 의해 제공된 정보에 연관된 고정 속도 모델을 사용함. 점 매칭 알고리즘이 위트니스 카메라 시스템으로부터의 실시간 비디오 공급에 식별된 점들과 새로운 3D 점 클라우드 내의 투영된 점들을 매칭시킬 수 있게 하기 위해 3D 점 클라우드를 현재의 프레임 상에 다시 투영하도록 상기 예측을 사용함.

[0057] • 시스템은 위트니스 카메라 시스템으로부터의 실시간 비디오 공급에 식별된 점들과 새로운 3D 점 클라우드 내의 투영된 점들 간의 오류를 최소화하기 위해 카메라 추적에 대해 레벤버그-마콰트(Levenberg-Marquardt) 최소화 방식을 사용함.

[0058] • 사용자는 3D 가비지 매팅 마스크들(3D Garbage Matting masks)과 같은 3D 마스크들을 정의하기 위해 카메라 추적 시스템에 의해 생성된 3D 점 클라우드를 사용할 수 있음.

[0059] • 3D 객체들은 정적인 객체들, 동적인 애니메이션들, 가상 세계들, 가상 사람, 가상 건물들, 가상 정치, 가상 영화 세트들, 및 애니메이션 데이터베이스 내의 임의의 데이터를 포함함.

[0060] • 필름 카메라 및 위트니스 카메라는 플래시 LED와 연관된 광 세기 커브들을 비교함으로써와 같이, 변조된 광원을 사용하여 프레임 획득 지역에 대하여 교정됨.

[0061] 다른 개념들 - 각각은 앞에서 정의된 인물들 중 임의의 것 또는 다음에 정의된 임의의 다른 개념들과 조합될 수 있다:

[0062] 실시간, 컴퓨터 생성된 3D 객체들 및 필름 카메라로부터의 비디오 공급을 믹싱 또는 합성하는 방법으로서, 필름 카메라의 몸체는 3D 내에서 움직일 수 있고, 카메라 내의 또는 그것에 부착된 센서들은 카메라의 3D 포지션 및 3D 방향을 정의하거나 3D 포지션이 계산될 수 있게 하는 실시간 포지셔닝 데이터를 제공한다.

[0063] TV 방송, 영화, 또는 비디오 게임들을 위한 증강 현실 비디오를 생성하기 위해, 실시간, 컴퓨터 생성된 3D 객체들 및 비디오 카메라와 같은 필름 카메라로부터의 비디오 공급을 믹싱 또는 합성하기 위한 방법에 있어서,

[0064] (a) 필름 카메라의 몸체는 3D 내에서 움직일 수 있으며, 필름 카메라 내의 또는 그것에 직접 또는 간접적으로 부착된 센서들은 필름 카메라의 3D 포지션 및 3D 방향을 정의하거나 필름 카메라의 3D 포지션 및 3D 방향이 계산되게 할 수 있는 실시간 포지셔닝 데이터를 제공하고,

[0065] (b) 실시간 포지셔닝 데이터는 그 후 컴퓨터 생성된 3D 객체들을 생성하거나, 다시 불러내거나, 렌더링하거나, 수정하기 위해 시스템에 의해 자동으로 사용되고,

[0066] (c) 결과로 발생한 컴퓨터 생성된 3D 객체들은 그 후, TV 방송, 영화, 또는 비디오 게임들에 증강 현실 비디오를 제공하기 위해 필름 카메라로부터의 비디오 공급에 믹싱 또는 합성된다.

[0067] 앞서 정의된 바와 같은 그리고 감독(또는 사진 감독)이 카메라 내의 사전-제작 컴퓨터 생성된 3D 자산들, 전형

적으로 사전 가시화 또는 시각 효과들의 자산들을 카메라에 의해 캡쳐된 비디오 이미지들 또는 실제 액션 영화 플레이트(live action film plate) 상에서 실시간으로 조작(frame-up)하게 하기 위해 사용되는 방법들 및 시스템들은, 감독이 감독에 의해 보여지는 것과 같은 비디오로 자동으로 생성되는 컴퓨터 생성된 3D 자산들과 함께 실시간으로 가능한 카메라 각도들 및 움직임들을 조사하게 한다.

[0068] 앞서 정의된 바와 같은 방법들 또는 시스템들에 있어서, 실시간 포지셔닝 데이터는 후반-제작 처리들에 매칭-움직임 데이터를 제공하기 위해 기록되고 시간-코드 스템핑된다.

[0069] 앞서 정의된 바와 같은 방법들 또는 시스템들은 가상 객체들 또는 이미지들을 방송 비디오 스트림에 삽입되도록 하기 위해 사용된다.

[0070] 앞서 정의된 바와 같은 방법 또는 시스템은 다음의 것들 중 하나 이상을 가능하게 한다:

- 스튜디오 카메라들에 대한 실시간 추적

- 스테디캠에 대한 실시간 추적

- 크레인-장착 카메라들에 대한 실시간 추적

- 돌리-장착 카메라들에 대한 실시간 추적

- 스테디캠에 대한 실시간 추적

- OB(Outside Broadcast)에 대한 실시간 추적

- 2D 후반-제작을 위한 실시간 데이터(예를 들면, 추적 데이터)의 사용

- 3D 입체 콘텐트에 대한 후반-변환을 위한 실시간 데이터(예를 들면, 추적 데이터)의 사용

- 원시 3D 입체 콘텐트를 위한 실시간 데이터(예를 들면, 추적 데이터)의 사용

- 3D 그래픽들 삽입

- 스튜디오-안 또는 세트-상 간접 광고를 위한 3D 그래픽들 삽입

- OB를 위한 3D 그래픽들 삽입

- 다른 후원된 이미지들에 대한 3D 그래픽들 삽입

- 시청자-위치 특정인 3D 그래픽들 삽입

- 시청자-특정인 3D 그래픽들 삽입

- 시간-특정인 3D 그래픽들 삽입

- 군중 장면들을 채우기 위한 3D 그래픽들 삽입

- 초록색 화면 대체를 위한 3D 그래픽들 삽입

- 박물관들에서 및 문화, 역사, 또는 자연 현장들 내의 번역 센터들에서 학습을 돋기 위한 교육 콘텐트의 3D 그래픽들 삽입

- 장면 내의 객체들의 절대적인 또는 상대적인 크기의 측정

[0091] 앞서 정의된 바와 같은 방법들 또는 시스템들에 있어서, 필름 카메라는 다음 중 하나이다:

- 표준 테더(tether)를 갖는 모든 카메라들

- 전략적인 광섬유 접속을 필요로 하는 카메라들

- [0094] - RF/무선 접속을 필요로 하는 카메라
- [0095] 앞서 정의된 바와 같은 방법 또는 시스템들에 있어서, 다음의 시장들 중 하나에 사용된다:
- 영화/TV (라이브(live) 아님)
  - 광고 (라이브 아님)
  - 광고들 라이브
  - 방송 (스포츠 아님)
  - 방송 OB
  - 스포츠 스튜디오 기반
  - 스포츠 OB 기반
  - TV 간접 광고 라이브
  - 인터넷 사용 (라이브 아님)
  - 인터넷 라이브
  - 인터넷 라이브 영역 기반
  - 인터넷 간접 광고 라이브
  - 박물관/문화유산 콘텐트
  - 박물관/문화유산 광고들
  - 건축학
  - 게임들.
- [0112] 증강 현실 이미지들이 스마트폰, 및 증강 현실 안경을 포함하는, 임의의 디스플레이 장치에 의해 디스플레이되게 하는 앞서 정의된 바와 같은 방법들 또는 시스템들에 있어서, 증강 현실 이미지들의 겉모습은 실시간 포지셔닝 데이터에 의존하여 자동으로 바뀐다.
- [0113] 실시간 컴퓨터 생성된 3D 객체들이 카메라로부터의 비디오 공급과 맵팅되는 필름, 영화, TV 쇼 또는 비디오 게임에 있어서, 카메라의 몸체는 3D 내에서 움직일 수 있고, 카메라 내의 또는 그것에 부착된 센서들은 카메라의 3D 포지션을 정의하거나 3D 포지션이 계산되게 하는 실시간 포지셔닝 데이터를 제공한다.
- [0114] 종래의 카메라에 부착되도록 적응된 장면 조사 및 추적 장치에 있어서, 카메라의 몸체는 3D 내에서 움직일 수 있고, 장면 조사 및 추적 장치 내의 하드웨어 센서들은 카메라의 3D 포지션을 정의하거나 3D 포지션이 계산되게 하는 실시간 포지셔닝 데이터를 제공한다.
- [0115] 카메라 내의 또는 카메라에 부착된 센서들을 포함하는 핸드-헬드 또는 휴대용 카메라는 3D 기준 프레임에 대한 카메라의 3D 포지션 및 3D 방향을 정의하거나 광학 흐름 데이터와 같은 다른 데이터를 분석하는 시스템의 전체 또는 부분으로서 3D 포지션이 계산되게 하는 실시간 포지셔닝 데이터를 제공한다.
- [0116] 입체적 위트니스 카메라 시스템을 포함하거나 그것에 부착된 필름 카메라에 있어서, 위트니스 카메라 시스템은 광각(예를 들면, 180도) 입체적 이미지들을 생성하고, 상기 필름 카메라는 소프트웨어가 이미지들을 처리하게 하고, 카메라 시스템이 전혀 추적/움직여지지 않으면서, 인스턴트 3D 점 클라우드를 생성하게 한다.
- [0117] 필름 카메라에 부착되기 위한 카메라 추적 시스템은 입체적 이미지들을 생성하고, 소프트웨어가 이미지들을 처리하게 하고, 카메라 시스템이 전혀 움직이지 않으면서, 인스턴트 3D 점 클라우드를 생성하고 필름 카메라의 실시간 추적(포지션, 방향, 줌, 초점, 및 조리개)을 제공하게 하는 위트니스 카메라 시스템을 갖는다.
- [0118] 증강 현실 안경 내에 또는 그것에 부착된 센서들을 포함하는 증강 현실 안경은 3D 기준 프레임에 대한 안경의 3D 포지션 및 3D 방향을 정의하거나, 광학 흐름 데이터와 같은, 다른 데이터를 분석하는 시스템의 전체 또는 부분으로서 3D 포지션이 계산되게 하는 실시간 포지셔닝 데이터를 제공한다.
- [0119] 본 시스템은 N캡이라 불리는 시스템 내에 구현된다. N캡의 다양한 응용들에는 다음이 포함된다:

[0120] 영화, TV, 및 광고들 내의 시각 효과들

[0121] 사전-제작 및 제작 내의 시각 효과들: N캡은 촬영 처리(제작) 동안 프리비즈(previz)(사전-제작)와 VFX(후반-제작) 간의 틈을 채우는 것을 도우려고 한다. N캡은 실시간으로 촬영하면서 카메라 내의 VFX 상에서 조작하고자 하는 것들에 대한 솔루션을 제공한다. 이전에 생성된 콘텐트를 사용하여, 종종 프리비즈 처리로부터, N캡은 그러한 프리비즈 자산들, 보통 3D 모델들, 및 애니메이션을 실시간으로 라이브 액션 영화 플레이트 상에 합성할 수 있다.

[0122] 신중하게 만들어지고 입증된 프리비즈를 세트 상에 놓으면 시간을 상당히 많이 절약할 수 있다. 영화 제작자들은, 초록색 화면 상에서 촬영하고 가상 배경을 합성하거나 전경에 생명체 또는 객체를 중첩시키든 아니든, 프리비즈 또는 VFX를 조작할 수 있다. 영화 제작자들은 그 후, 통상의 제작사항들 없이, 그때 그때 및 실시간으로, 가능한 카메라 각도들 및 움직임들을 연구하면서 프리비즈 및 VFX의 제어를 되돌려받을 수 있다.

[0123] 다음으로, 데이터가 기록 및 시간-코드 스템핑되어, 사설(editorial)에 최종 VFX의 '템프(temp)'를 제공하는 동시에, VFX 부분에 카메라 매칭-움직임 데이터(6 DOF 트랙)를 제공한다.

[0124] 인-카메라(in-camera)로 '템프' VFX를 촬영함으로써 및 처리 내에서 보다 일찍 VFX를 디자인함으로써, 후반-제작 처리로서 VFX 이외의 많은 추측 작업을 할 수 있다. 우리는 VFX 디자인 처리를 줄일 수 있으므로, 낭비를 줄일 수 있다.

[0125] 가상 광고 - 라이브 방송

[0126] N캡 기술은 방송 공간을 갖는 실시간 가상 광고에 잘 제공된다. 인-카메라인 실제의 물리적 제품과 반대로, 촬영 후 후반-제작 동안 제품이 삽입되는 디지털 간접 광고는 더욱 인기있어지고 있다. 그러나, 디지털 간접 광고 라이브(live)를 증대시기 위한 능력으로, 우리는 다양한 가능성들을 시작할 수 있다. 예들에는 경기장의 상표 로고들 또는 드라마 내의 캔 음료수들이 포함될 수 있다. 이것은 사진-실제 CG 이미지를 촬영 환경 조명에 기초하여 실시간으로 렌더링하는 능력이 필수적이되는 경우이다.

[0127] 홀로그램 게임하기

[0128] N캡의 기술을 홀로그램 디스플레이 및 (구글 안경과 같은) 아이-웨어(eye-wear)와 조합하면, 완전히 이머시브한 (immersive) 게임 경험이 제공될 수 있다. 실세계를 가상 세계 및 인물들과 조합한다. 그 가능성은 다소 끝이 없다.

[0129] 박물관 & 문화유산

[0130] N캡의 기술은 박물관 및 문화유산 분야들에 광대한 가능한 응용들을 제공한다.

[0131] 발굴 도구로서, 그것은 폐허들과 같은 유적지들의 재구성을 위해 사용되어, 그 유적지가 그것을 옛날 영광의 시대에 어떤지를 보여줄 수 있다.

[0132] 박물관에서, N캡은 아마도 공룡의 크기 및 움직임, 그것의 내부 장기들 또는 해골 구조를 보여주기 위한 교육 도구로서 사용될 수 있다. 또 다른 예는 분해된 애니메이션된 다이어그램을 효율적으로 보여주지만 완전히 이며시브한 3D 공간 내에서 내연 기관의 동작을 탐험하기 위한 것일 수 있다.

[0133] 가상 카메라 - VCS

[0134] N캡 기술은 이 응용에 완전히 제공된다. 가상 카메라는 반드시 완벽한 CG 3D 환경을 보여주는 LCD 화면이다. 오퍼레이터가 LCD 화면을 움직이면, 그것은 움직임의 면에서 단지 물리적 카메라처럼 동작한다. LCD 카메라는 XYZ 병진들 및 회전들 내에서 추적되고 CG 엔진으로부터의 완전한 GC 환경을 실시간으로 보여준다. 현재 다양한 VCS(가상 카메라 시스템) 솔루션들이 사용가능하지만, 모두가 기술적인 면에서 제한사항들을 갖는다. 그들은 셋업하기 위해 시간 소모적인 경향이 있고, 그들이 동작할 공간 및 환경에 제한되고, 고가이다. N캡 기반 VCS는 테블릿의 모양을 취할 수 있으므로, 처리는 LCD 화면과 국부적으로 조합되어 계산될 것이다. CG 컴퓨팅들은 원격 PC로부터 취해지거나 요구조건들에 따라 국부적으로 취해질 수 있다. 그 아이디어는 사용가능성 및 기준 소매 가격들을 통해 이전에 이러한 도구들에 대한 액세스를 가질 수 없었던 필름 내의 많은 부분들에 대한 오픈 액세스(open access)를 가능하게 하기 위한 것이다. 이것은 디자이너들, 건축가들, VFX, 게임 회사들, CG 및 애니메이션 팀들 등에게 매우 유익하다.

[0135] 이 VCS는 또한 잠재적인 박물관 및 문화유산 솔루션에 대한 백본(backbone)을 형성할 것이다.

### [0136] 가상 스카우트 시스템 - VSS

[0137] 당신이 영화를 촬영하고, 그것이 대부분 파란색 화면/녹색 화면 상에서라고 상상하자. 그러나, 영화 제작사 또는 감독으로서, 나는 어떤 각도들 또는 샷들(shots)이 최고일 것이란 것을 어떻게 알아낼까? 어쨌든, 나는 적은 양의 물리적인 세트 및 적은 연기자들을 가질 수 있지만, 나는 내 상상과 달리, 어떤 카메라 각도들이 최고인지, 팀 전체에게 우리가 어디에 카메라들을 설치해야하는지 설명하고 유효한 렌즈들을 선택하는 것에 대한 아이디어가 거의 없을 수 있다. 현재 우리는 필름 카메라 및 모든 그것의 장비들을 세트 상의 위치로 움직이고, N캡을 설치하고, 단지 그것이 다른 포지션에서 더 나은지를 인식하기 위해 관찰할 것이다. 우리가 이러한 결정들을 빠르고 효율적으로 하기 위해 가벼운 휴대용 시스템을 유일하게 갖는다면, 이것이 N캡 VSS이다.

[0138] VSS는 오늘날의 N캡이지만 소형 폼 팩터 카메라(form factor camera)로 통합된다 - DSLR을 생각해보자. 그것은 반드시 N캡의 이점들을 갖는 디지털 뷰파인더(viewfinder)이다. 우리는 오늘날의 N캡의 모든 요소들을 필름 세트 상에서, 그러나 휴대용 장치 내에 조합한다. 박물관/유적지 및 건축가들이 잠재적인 투자자들에게 그들의 새로운 생성물을 보여주길 원하기 때문에, TV 드라마를 위한 가상 제작은 또한 이 도구로부터 상당히 득을 볼 것이다.

### 도면의 간단한 설명

[0139] 도 1은 배치(configuration) 2.3에 대한 하드웨어 설명을 도시한 도면. 그려진 버전은 추적 워크스테이션(workstation) 및 렌더링 워크스테이션이 병합된 것임에 주의한다.

도 2는 필름 카메라의 포지션 및 회전을 실시간으로 획득하기 위한 글로벌 추적 알고리즘을 도시한 도면.

도 3은 우리의 시스템을 위해 생성된 3D 자동-순위화 등록 객체를 도시한 도면. 그 객체는 우리의 소프트웨어 내에서 실세계의 근원, 방향, 및 크기를 자동으로 컴퓨팅하기 위해 실제 장면에서 사용된다. 평면 버전이 도 5의 우측 도면에 보여진다.

도 4는 자연스러운 마커들의 검출을 위해 사용되는 이미지 피라미드를 도시한 도면. 이 도면은 또한 그 마커들에 연관된 몇몇의 2D 질감 패치들을 보여준다.

도 5는 N캡 내에서 현상 및 사용되는 교정 차트들을 도시한 도면. 좌측: 모든 렌즈들의 눈금을 매기기 위해 사용되는 교정 차트이다. 중간: 필름 카메라와 위트니스 카메라 사이에 컴퓨팅하는 라인업 또는 사용되는 교정 차트. 좌측 상의 교정 차트는 또한 필름 카메라 상의 초점 거리가 충분히 짧은 경우들에 대하여 사용될 수 있다 (세트 상에서 실험적으로 결정됨). 우측: 등록 객체의 평면 버전이다. 보다 다루기 용이하고 바닥 표면 상에 어디든 위치될 수 있다. 가상 객체들이 실제 장면에 추가될 때, 원들은 그 세계에 방향, 스케일, 및 근원(0,0,0)을 제공하는 정규직교 기저를 제공한다. 상부 좌측 원의 중심은 세계의 근원이다.

도 6은 (15mm 및 19mm 로드들에 대한) 상이한 행에 크기들을 갖고 필름 카메라에 장착되는 카메라 바(bar)를 도시한 도면.

도 7은 무선 조사 장치를 도시한 도면. 조사가 요청되면(즉, 예를 들면, 평면적인 경우(2.1 섹션을 참조) 또는 매우 복잡한 세트(2.1.3 섹션의 끝부분을 참조)), 그 후 그 조사 장치를 사용하여 전체 추적 처리가 달성된다. 조사가 완료되면, 모든 추적 데이터(점 클라우드, 키포인트들 등)는 추적/렌더링 워크스테이션으로 무선으로 송신되고, 추적/렌더링 워크스테이션은 즉시 추적을 복원하고 추적 데이터 및 최종 합성된 결과 모두를 스트리밍 출력하기 시작한다.

도 8은 카메라/컴퓨터 브레이크아웃 박스(Breakout Box)를 도시한 도면. 다수의 RJ45 케이블들을 사용하는 버전이 여기에 나타나 있다는 것에 주의한다. 모든 이를 브레이크아웃 박스들은 또한 그들의 케이블들이 단일의 전략적인 광섬유 케이블들로 대체되는 이형(variant)을 갖는다.

### 발명을 실시하기 위한 구체적인 내용

[0140] 여기에 설명된 기술은 우선 강력한 센서 결합 알고리즘이므로, 3개의 (입체적인 경우) 또는 2개의 (평면적인 경우) 상이한 광학 흐름들로부터 컴퓨팅된 자연스러운 마커들(즉, 물리적인 장면에 수동으로 추가되지 않음)과 4개의 다른 센서들(사이로스코프, 가속도계, 자력계, 3D 거리 센서)의 물리적인 측정을 병합한다. 그것은 또한 어떤 물리적인 마커들 또는 그것의 작업을 수행하기 위해 실제 장면에서 바로 보여지는 어떤 것도 수동으로 추가하도록 요청하지 않는 완전히 거슬리지 않는(non-intrusive) 시스템이다.

## [0141] 1. 하드웨어

[0142] 이 기술은 스트리밍 출력해야하는 데이터의 종류에 따라 2개의 상이한 가능한 하드웨어 배치들을 갖는다.

[0143] 모든 이들 2개의 배치들은 카메라 바로 불리는 장치 상에 장착되는 다음의 하드웨어(상세한 도시를 위해 도 6을 참조) 및 필수적으로 모든 상이한 케이블들을 하나 또는 몇 개의 독립적인 룸들(looms)로 병합 또는 분리하는 블레이크아웃 박스인 개별적인 접속 상자(상세한 도시를 위해 도 8을 참조)를 공통으로 갖는다:

- [0144] • 선택된 접근법에 따라 (평면적인 또는 입체적인) 워트니스 카메라들로 불리는 1개 또는 2개의 고속 카메라들 (적어도 100fps).

- [0145] • 각각의 워트니스 카메라에는 180도 어안 렌즈가 장착됨.

- [0146] • 사이로스코프, 자력계, 및 가속도계를 포함하는 9-DOF 센서(또는 자력계가 사용되지 않을 때는 6-DOF).

- [0147] • 화소의 심도를 캡쳐하는 3D 거리 센서.

[0148] 카메라 바들 이외에, 모든 배치들은 줌, 초점, 및 조리개 값을 읽기 위해 필름 카메라 상에 장착된 정규 렌즈 인코더들을 갖는다. 그들은 또한 장면을 조사하고 그것을 학습하기 위해 무선 조사 장치(상세한 도시를 위해 도 7을 참조)를 공통으로 갖는다. 그러나, 완전히 입체적인 접근법의 경우(2.5 섹션을 참조), 그리고 보다 정확하게 인스턴트 조사인 경우(2.5.1 섹션을 참조), 그 조사 장치는 불필요하다.

[0149] 이를 공유된 인물들 외에, 다음과 같은 2개의 배치들이 설명된다:

### [0150] 배치 1: 카메라 스트리밍

[0151] 1. 무선 내장형 컴퓨터:

[0152] (a) 내장형 추적 컴퓨터가 필름 카메라 상에 장착된다. 그것은 인코더들로부터의 줌, 초점, 및 조리개(ZFI) 뿐만 아니라, 필름 카메라 병진 및 회전(RT)을 외장형 렌더링 워크스테이션에 무선 스트리밍 출력한다.

[0153] (b) 외장형 콘텐트 생성 워크스테이션. 이 기계는 전형적으로 최종 합성에 3D 컴퓨터 생성된 애니메이션 및 관련 CG 콘텐트를 제공하기 위한 컴퓨터 그래픽 소프트웨어를 실행한다. 그것은 또한 애니메이션된 데이터를 외장형 렌더링 워크스테이션(1-1c)에 스트리밍 출력한다.

[0154] (c) 한편으로는 내장형 컴퓨터(1-1a)로부터 데이터를 수신하고 다른 한편으로는 또한 콘텐트 생성 워크스테이션(1-1b)으로부터의 애니메이션된 데이터를 처리하는 외장형 렌더링 워크스테이션. 이 렌더링 워크스테이션은 1-1a로부터의 필름 카메라 RT+ZFI를 사용하여 1-1b로부터의 3D 가상 환경을 디스플레이하고, 필름 카메라로부터의 실제 비디오 흐름 내에 결과를 미싱한다. 그것은 또한 필름 카메라의 접안 렌즈 또는 제어 모니터에 최종 합성된 결과를 무선으로 스트리밍한다.

[0155] 2. 유선 내장형 컴퓨터

[0156] (a) 내장형 추적 컴퓨터는 필름 카메라 상에 장착된다. 그것은 전략적인 광섬유 케이블들을 사용하여 필름 카메라 병진, 회전(RT), 및 인코더 데이터(ZFI)를 외장형 렌더링 워크스테이션에 스트리밍 출력한다.

[0157] (b) 배치 1-1b와 유사함.

[0158] (c) 최종 합성된 결과가 HD-SDI 케이블을 통해 카메라에 다시 송신되는 것만 제외하고, 배치 1-1c와 유사함.

### [0159] 배치 2: 전체 스트리밍

[0160] 1. 무선 내장형 컴퓨터:

[0161] (a) 내장형 추적 컴퓨터는 필름 카메라 상에 장착된다. 어떤 실제적인 처리도 그 유닛 상에서는 달성되지 않는다. 그것은 단지 모든 센서 데이터를 외장형 추적 워크스테이션에 무선 스트리밍 출력한다. 워트니스 카메라들의 2x 비디오 흐름들, 1x9DOF 데이터, 1x 거리 센서 데이터가 스트리밍 출력된다. 총 데이터 양은 142MB/초보다 큰 최소 대역폭을 요청한다(워트니스 카메라들은 63MB/초를 요청하고, 거리 센서는 74MB/초를 요청하고, 9DOF 센서는 4.4MB/초를 요청함).

[0162] (b) 내장형 추적 컴퓨터로부터 데이터를 수신하는 외장형 추적 워크스테이션. 이 유닛은 필름 카메라의 회전 및

이동 행렬들을 컴퓨팅하도록 모든 센서 데이터를 처리한다. 그것은 카메라 RT+ZFI를 외장형 렌더링 워크스테이션(2-1d)에 스트리밍 출력한다.

[0163] (c) 외장형 콘텐트 생성 워크스테이션. 이 기계는 전형적으로 최종 합성에 3D 컴퓨터 생성된 애니메이션 및 관련 CG 콘텐트를 제공하도록 컴퓨터 그래픽 소프트웨어를 실행한다. 그것은 또한 애니메이션된 데이터를 외장형 렌더링 워크스테이션(2-1d)에 스트리밍 출력한다.

[0164] (d) 한편으로는 내장형 컴퓨터(2-1b)로부터 데이터를 수신하고, 다른 한편으로는 또한 콘텐트 생성 워크스테이션(2-1c)으로부터의 애니메이션된 데이터를 처리하는 외장형 렌더링 워크스테이션. 이 렌더링 워크스테이션은 2-1c로부터의 3D 가상 환경을 디스플레이하도록 2-1b로부터의 필름 카메라 RT+ZFI를 사용하고, 필름 카메라로부터의 실제 비디오 흐름 내에 결과를 먹생한다. 그것은 또한 최종 합성된 결과를 무선으로 필름 카메라 접안 렌즈 또는 제어 모니터에 다시 스트리밍한다.

## 2. 유선 내장형 컴퓨터:

[0166] (a) 내장형 추적 컴퓨터는 필름 카메라 상에 장착된다. 그것은 전략적 광섬유 케이블들 또는 다수의 Cat-6 RJ45 케이블들을 사용하여 필름 카메라 병진 및 회전(RT)을 외장형 렌더링 워크스테이션(2-2c)에 스트리밍 출력한다. 인코더들은 또다시 줌, 초점, 및 조리개(ZFI)를 위해 사용된다.

[0167] (b) 배치(2-1c)와 유사함.

[0168] (c) 최종 합성된 결과가 HD-SDI 케이블을 통해 카메라에 다시 송신되는 것만 제외하고, 배치(2-1d)와 유사함.

[0169] 3. 내장형 컴퓨터 없음(그래픽 표현을 위해 도 1을 참조: 모든 다른 버전들이 그것으로부터 용이하게 도출될 수 있음):

[0170] (a) 단일 카메라 추적 장치(카메라 바)가 필름 카메라에 부착되고, 브레이크아웃 박스에 접속된다.

[0171] (b) 내장형 추적 장치로부터 데이터를 수신하는 외장형 추적 워크스테이션. 이 유닛은 필름 카메라의 회전 및 이동 행렬들을 컴퓨팅하기 위해 모든 센서 데이터를 처리한다. 인코더들은 또다시 줌, 초점, 및 조리개(ZFI)를 위해 사용된다. 그것은 렌더링 워크스테이션(2-3d)에 카메라 RT+ZFI를 스트리밍 출력한다.

[0172] (c) 배치(2-1c)와 유사함.

[0173] (d) 최종 합성된 결과가 HD-SDI 케이블을 통해 카메라에 다시 송신되는 것만 제외하고, 배치(2-1d)와 유사함.

[0174] 이전 배치들 이외에, 모든 버전은 또한 2개의 다른 이형들을 갖는다. 하나는 실제로 외장형 렌더링 워크스테이션과 병합된 콘텐트 생성 워크스테이션을 갖고, 두번째 이형은 추적 워크스테이션과 병합된 외장형 렌더링 워크스테이션을 갖고는다.

## 2. 소프트웨어

[0175] 정확한 추적 결과들 및 실시간 증강 현실을 달성하기 위해 몇몇의 새로운 과학적인 방법들 및 소프트웨어들이 생성되었다[19]. 이들 방법들은 다음 섹션들에서 설명된다.

### 2.1 비디오 공급으로부터의 평면적인 마커없는 추적

#### 2.1.1 처리 개요

[0176] 이 기술에서 사용되는 평면적인 기술은, 하나는 사용자 개입을 요청하는 반면 다른 것은 완전히 자동적이란 점만 제외하고, 실제로 기술적으로 거의 동일한 **2개의 개별적인 패스(pass)**들로 구축된다.

[0177] 제 1 패스는 조사로 불린다. 사용자는 실제 장면 및 등록 객체를 스캔하기 위해 1 섹션의 두번째 문단에서 설명된 하드웨어 상에 구축된 장치를 사용한다(또한 도 7 참조). 그 단계 동안, 사용자가 주변을 걸어다니면서 환경을 스캔하므로, 자연스러운 마커들이 시스템에 의해 3D 내에서 자동으로 학습 및 컴퓨팅된다. 사용자가 조사하기 시작하자마자, 그는 또한 자동으로 그 세계의 스케일, 근원, 및 방향을 추정하기 위해 **등록 객체**를 캡쳐해야 한다(2.1.3.2 섹션을 참조). 이것이 달성되면, 사용자는 자연스러운 마커들의 최고의 가능한 3D 점 클라우드를 구축하기 위해 장면의 나머지를 계속 스캔할 수 있다(자연스러운 마커는 사람 눈에 보이지 않는 특징 점이고 우리 알고리즘들에 의해 컴퓨팅됨). 이 조사가 보다 철저하게 달성되면, 최종 결과가 더 양호해질 것이다. 조사 부분은 모든 것들 중에서 가장 중요한 작업이고 그것은 매우 중요하다.

[0178] 제 2 패스는 필름 카메라 상에 놓인 추적 장치로부터 수행되는 필름 카메라 추적(2.1.4 섹션들 참조)이다. 제목

이 "하드웨어"인 이전 섹션의 설명을 뒤따르고 다양한 상황들에 따라 상이한 배치들이 적용된다. 이 단계는 완전히 자동이고 보통 어떠한 사람의 개입도 요청하지 않는다.

### [0182] 2.1.2 알고리즘 개요

[0183] 여기에 설명된 기술은 다수의 센서들 결합 기술(2.3 섹션을 참조)에 기초한 증강 현실 프레임워크이다.

[0184] 실제로, 그것은 카메라를 추적하고 가상 객체를 삽입하기 위해 통상의 단독 광학 흐름 데이터에 의존하지 않는다. 이러한 기술들은 많은 경우들에서 그들의 과학적이고 실질적인 제한사항들을 나타낸다. 예를 들면, 사용자가 추적을 위해 사용되는 카메라를 숨기면, 추적은 실패하고 길을 잃게 된다. 카메라가 조사되지 않은 영역을 가리킬 때, 그것은 정확히 동일한 문제이다(2.1.3 섹션을 참조). 게다가, 필름 카메라가 전혀 움직이지 않으면, 완전히 가시적 추적 알고리즘에 의해 컴퓨팅된 검출된 자연스러운 마커들의 정확도 때문에 몇몇의 불가피한 약간의 혼들림 생점들이 여전히 존재한다. 데이터를 필터링하는 것을 그 생점을 부분적으로 해결하지만, 당신은 특히 매우 긴 초점 렌즈를 사용할 때 완전히 안정적인 가상 객체를 갖지 못한다. 우리의 기술은 카메라가 움직이는지 아닌지를 알기 위해 다른 센서들을 사용하므로, 우리의 추적 알고리즘은 그 생점을 갖지 않는다.

[0185] SLAM[9], PTAM[8], RSLAM[13] 등과 같은 정규 광학 흐름 기술들은 카메라에 의해 캡쳐된 비디오에 포함된 정보를 사용하여 그것의 포지션 및 방향의 근사치를 낸다. 우리의 기술은, 한편으로는, 추적의 핵심이 그것이 길을 잃었을 때 단순한 키프레임 탐색이 아니고, 다른 한편으로는, 그것이 카메라의 정확한 포지션 및 방향을 컴퓨팅하기 위해 7개의 사용 가능한 센서들(1개의 필름 카메라, 2개의 위트니스 카메라들, 자이로스코프, 가속도계, 자력계, 및 3D 거리 센서)의 데이터 모두를 사용한다는 점만 제외하고, 그것의 가시적 추적 부분(가시적 추적이라 불리는)에 대해 유사한 접근법을 갖는다.

[0186] 우리는 다음의 센서들을 통합하기 위해 확장형 칼만 필터(EKF, 2.3.2 섹션을 참조)[7, 21]를 사용한다: 자이로스코프, 가속도계, 자력계 센서들 및 위트니스 카메라들. EKF 기술은 우리의 소프트웨어의 핵심이고 모든 예측/정정 알고리즘은 각각의 센서 중 최고를 사용하도록 허용하는 데이터 병합 방법에 기초한다. 이것은 모든 다른 기술들이 실패하는 단순하고 중요한 경우들에 비할대 없는 강력함을 제공한다. 실제로 센서가 추적을 잃게 되면 (실제로 그것의 측정들이 더이상 신뢰적이지 않다는 것을 의미함), EKF는 다른 신뢰적인 및 남아있는 센서들의 데이터를 병합함으로써 포지션/방향의 추정을 여전히 획득할 수 있다. 그것 이외에, 모든 센서가 신뢰의 수준을 가지므로, 이것은 시스템이 부정확한 정보들을 사용할 것을 방지하도록 병합 알고리즘에 영향을 준다. 예를 들면, 정규 자이로스코프는 그것이 움직이지 않더라도 그것의 방향을 수정하는 경향이 있는 드리프트(drift)라 불리는 자연스러운 흐름을 갖는다. 시간이 지날수록, 보다 많은 오류들이 자이로스코프에 의해 생성된다. 가시적 추적에 의해 자이로스코프의 드리프트가 정정되고, 그것의 오류가 너무 커지면, 가시적 추적은 자이로스코프에 의해 정정된다. 따라서 EKF는 이전에 올바르게 추정된 값들을 사용하여 필름 카메라의 움직임들을 예측한 후 모든 센서들에 의해 주어지는 새로운 측정들에 따라 그것의 예측을 정정함으로써 필름 카메라의 최고의 포지션 및 회전을 찾아낼 수 있다.

[0187] 게다가, 가장 완전한 광학-기반 추적 기술들은 그것이 길을 잃었을 때 카메라의 위치를 지정하기 위해 키프레임(즉, 스냅샷들(snapshots))을 사용한다. 당신이 카메라로 가리키고 있는 현재의 뷰의 키프레임을 갖고 있지 않으면, 복원은 실패한다. 예를 들면, 당신이 카메라를 단순히 수직으로 회전시킬 경우들을 해결하기 위해 SIFT[11] 또는 SURF[1] 기술들과 같은 보다 강력한 기술들이 물론 사용될 수 있다. 실제로 모든 이들 시스템들은 현재의 프레임을 가장 유사한 키프레임과 비교하므로, 보다 양호하고 빠른 매칭을 획득하기 위해 회전 및 스케일링 불변 서술자들을 사용하는 것에 매우 관심이 있다. 그러나, 당신이 카메라가 존재하는 포지션으로부터 어떠한 키프레임도 갖지 않으면, 당신의 시스템이 그것의 포지션을 복원할 수 있는 방법은 없다. 우리의 기술은 모든 기술 중 가장 최고를 취하고(우리는 핵심 서술자들로서 ORB[18]을 사용할 것임의 주의), 우리는 잊어버린 추적을 복원하기 위해 3개의 상이한 스레드들(threads)을 동시에 실행시킨다(보다 상세한 설명을 위해 2.1.4.4 섹션을 참조). 복원 처리는 확률론적인 및 확률적인 접근법에 기초한 글로벌 알고리즘(global algorithm)을 보여줄 수 있다. 추적이 길을 잃으면, 3D 공간은 타원 접근법을 사용하여 가장 나중에 알려진 포지션들 주변이 즉각적으로 샘플링되고, 그 타원 내부의 모든 키프레임들이 현재의 키프레임과 비교된다. 게다가, 나머지 센서들(자이로스코프, 가속도계, 자력계, 3D 거리 센서)로부터의 모든 다른 데이터는, 우리의 소프트웨어가 모든 타당한 양호한 후보자들을 명확하게 하는 탐색에 여전히 병합된다. 물론, 양호한 충분한 후보인 키프레임이 없으면, 시스템은 현재의 포지션의 근사치를 컴퓨팅하기 위해 가시적 추적 이외에 모든 다른 센서들을 사용할 것이다. 그 기술의 자명한 결과는 우리의 기술이 사용자로 하여금 추적을 잊어버리지 않으면서 조사하지 않은 장소로 카메라를 가리키게 하는 것을 허용하는 것이다. 그러나, 우리의 알고리즘들은 현재의 스냅샷과 매칭하는 키프레임

(즉, RT에 대한 보다 양호한 근사치)을 찾기 위해 3D 공간을 샘플링하는 것을 절대 멈추지 않는다. 가시적 추적이 복원되면, 모든 다른 센서 데이터는 사용되고 카메라 것을 이외에 그들의 정보들을 정정하기 위해 생성된다 (확장형 칼만 필터(2.3.2 섹션)를 참조).

[0188] 우리 기술의 최신 (즉, 정규) 버전은 입체적인데(2.5 섹션을 참조), 이는 그것이 21 센티미터 떨어져 있는 2개의 위트니스 카메라들을 사용하는 것을 의미한다(도 6을 참조). 이 기술은 비디오 공급에서 검출된 자연스러운 마커들의 심도 정보를 컴퓨팅하기 위해 단일의 위트니스 카메라의 어떠한 수동적인 병진 또는 호모그래피(homography)의 임의의 컴퓨팅을 요청하지 않는다(2.1.3 섹션을 참조). 이 방법은 대부분의 경우들에서 조사 단계를 더 이상 쓸모없게 만들어서 우리의 주요 알고리즘을 단순화시킨다(2.5 섹션을 참조). 평면적인 기술에서, 가시적 추적 자체는 시스템이 이미지 피라미드의 4개의 상이한 수준들(도 4를 참조)로의 카메라에서의 포지션을 추적하도록 허용하는 완벽한 하위-화소 접근법을 갖는 반면에, 모든 다른 광학-흐름 기반 기술들은 특정 점들을 탐색할 때 이미지들의 단지 2개의 상이한 해상도 상에서 화소 접근법을 잘못 사용한다.

### [0189] 2.1.3 제 1 단계: 조사

#### [0190] 2.1.3.1 개시화 단계

[0191] 제 1 단계는 사용자가 2개의 프레임들을 캡처하기 위해 장치를 물리적/수동적 수평 병진하게 하는 것이다. 알고리즘은 FASTER[17]라 불리는 코너 검출 방법을 사용하여 이미지를 내의 자연스러운 마커들(즉, 특정 점들)을 자동으로 검출한다. 한 쌍의 점들 사이의 대응관계를 찾기 위해 매칭이 적용된다. 하나의 이미지에 대하여, 모든 검출된 점들은 동일한 3D 평면 상에 남아 있어야 하는데 그 평면으로부터 사진으로의 올바른 호모그래피를 컴퓨팅할 수 있기 위해서이다. 이것은 3D 평면에 대한 카메라의 회전 및 병진을 제공한다. 두번째 사진은 동일한 원리를 따르고, 우리는 현재의 사진에 대한 카메라의 두번째 포지션 및 병진을 획득한다. 하나의 카메라에서 다른 카메라로의 변환을 컴퓨팅하고 첫번째 3D 점 클라우드를 획득하는 것이 이제 가능하다.

[0192] 카메라에 의해 캡처되는 2개의 뷰들(좌측 및 우측)을 고려하자. 각각의 뷰에 대하여, 위트니스 카메라는 평면 상에 있는 점  $M$ 을 목표로 삼았다.  $M^L$  및  $M^R$ 은 각각의 좌측 및 우측 뷰들 내의  $M$ 의 각각의 투영들이다.

[0193] 우리는 다음을 작성할 수 있는데:

$$M^L = H \cdot M^R \quad (1)$$

$$H = (D \cdot R + \vec{T} \cdot \vec{N}^t) \quad (2)$$

[0194] 여기서:

[0195] •  $R$ 은  $M^L$ 이  $M^R$ 에 대하여 회전된 회전 행렬임.

[0196] •  $\vec{N}(A, B, C)$ 은 점  $M$ 이 속한 평면에 대한 법선 벡터임.

[0197] •  $D$ 는  $Ax + By + Cz + D = 0$ 과 같은 평면까지의 거리임.

[0199] 수학식 1에서 H에 대해 풀면 다음이 산출된다:

$$H = \begin{pmatrix} x_{M^L} & 0 \\ y_{M^L} & 0 \\ 1 & 0 \\ 0 & x_{M^L} \\ 0 & y_{M^L} \\ 0 & 1 \\ -x_{M^L} * x_{M^R} & -x_{M^L} * y_{M^R} \\ -y_{M^L} * x_{M^R} & -y_{M^L} * y_{M^R} \\ -x_{M^R} & -y_{M^R} \end{pmatrix}^t$$

[0200]

[0201] 우리의 알고리즘은 최소한 4개의 공면 점들에 정확한 개시화를 수행하도록 요청한다. 우리가 실제로 그것보다 훨씬 많은 점들을 갖는 것이 매우 통상적이다. 그것은 우리가 모든 호모그래피에 대한 점들의 전체 세트 중에서 4개의 점들을 사용하여  $n$ 개의 가능한 호모그래피들을 컴퓨팅하기 때문이다. 따라서 우리는 RANSAC[4] 방법을 사용하여 최고의 가능한 호모그래피를 실제로 찾고 인라이어들(inliers) 및 아웃라이어들(outliers) 후보들의 세트를 구축한다. 점들을 2개의 뷔들 내로 다시 투영하는 오류들을 최소화하는 가우스-뉴튼 방법에 의해 인라이어들은 개선될 것이지만, 아웃라이어들은 거부될 것이다. 최고의 호모그래피가 필터링된 점들의 세트로부터 컴퓨팅되면, 이제 좌측 및 우측 뷔들 사이에 회전 및 이동 행렬들을 추정하는 것이 필수적이 된다. 이것은 2개의 수직 행렬들 U 및 V와 대각선 행렬  $\Upsilon$ 로 단일 값 분해(decomposition)를 사용함으로써 용이하게 달성된다. 행렬 H는 이제  $H = U\Upsilon V^t$ 로 작성될 수 있다.

[0202]  $\Upsilon = (D' \cdot R' + T' \cdot N'^t)$  (3) 이고, 여기서,

$$R = \det(U) \cdot \det(V) \cdot U \cdot R' \cdot V^t$$

$$T = U' \cdot T'$$

$$N = V \cdot N'$$

$$D = \det(U) \cdot \det(V) \cdot D'$$

[0203] 이다.

[0204] 우리는 미결정된 호모그래피 경우들은 처리하지 않으므로, 수학식 3은 2개의 가능한 솔루션들을 갖는다:

- [0205] • 솔루션 1:

$$T' = (\lambda_1 - \lambda_3) \begin{pmatrix} \varepsilon_1 \sqrt{\frac{\lambda_1^2 - \lambda_2^2}{\lambda_1^2 - \lambda_3^2}} \\ 0 \\ -\varepsilon_3 \sqrt{\frac{\lambda_2^2 - \lambda_3^2}{\lambda_1^2 - \lambda_3^2}} \end{pmatrix} \quad (4)$$

[0206]

$$R' = \begin{pmatrix} \cos \theta & 0 & -\sin \theta \\ 0 & 1 & 0 \\ \sin \theta & 0 & \cos \theta \end{pmatrix} \quad (5)$$

[0207]

$$\begin{cases} \cos \theta = \frac{(\lambda_2^2 + \lambda_1 \lambda_3)}{\Lambda} \\ \sin \theta = (\lambda_1 - \lambda_3) \cdot \varepsilon_1 \varepsilon_3 \frac{\sqrt{(\lambda_1^2 - \lambda_2^2)(\lambda_2^2 - \lambda_3^2)}}{\Lambda} \\ \Lambda = \lambda_2(\lambda_1 + \lambda_3) \end{cases}$$

여기서, 이다.

 $\lambda_1, \lambda_2, \text{ 및 } \lambda_3$ 은 수학식 2로부터의 행렬 H의 고유치들이고,  $\lambda_1 \lambda_2 \lambda_3$  및  $\lambda_1 \neq \lambda_2 \neq \lambda_3$ 과 같이 정렬된다.

[0210]

- 솔루션 2:

$$T' = (\lambda_1 + \lambda_3) \begin{pmatrix} \varepsilon_1 \sqrt{\frac{\lambda_2^2 - \lambda_3^2}{\lambda_1^2 - \lambda_3^2}} \\ 0 \\ \varepsilon_3 \sqrt{\frac{\lambda_2^2 - \lambda_3^2}{\lambda_1^2 - \lambda_3^2}} \end{pmatrix} \quad (6)$$

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

$$\begin{cases} \cos \varphi = \frac{(\lambda_1 \lambda_3 - \lambda_2^2)}{\Lambda} \\ \sin \varphi = \varepsilon_1 \varepsilon_3 \frac{\sqrt{(\lambda_1^2 - \lambda_2^2)(\lambda_2^2 - \lambda_3^2)}}{\Lambda} \\ \Lambda = \lambda_2(\lambda_1 - \lambda_3) \end{cases}$$

여기서, 이다.

 $\lambda_1, \lambda_2, \text{ 및 } \lambda_3$ 은 수학식 2로부터의 행렬 H의 고유치들이고,  $\lambda_1 \lambda_2 \lambda_3$  및  $\lambda_1 \neq \lambda_2 \neq \lambda_3$ 과 같이 정렬된다.

[0214]

특징들 점들의 세트 이외에, 알고리즘은 또한 위트니스 카메라들이 모든 시점에서 무엇을 보는지의 스냅샷들과 같은 2개의 키 포지션들을 캡처한다. 추가적인 점들의 학습은 항상 새롭게 검출된 점들, 키 포지션들, 및 패치들(2.1.3.3 섹션에 설명됨)의 3개의 구성요소들의 합이다.

[0215]

마지막으로, 입체적인 점근법의 경우(2.5 섹션을 참조), 조사 단계는 자동적이고 일시적이 된다는 것을 주의하자. 수동적인 조사는 매우 복잡한 경우들(세트에 대해 실험적으로 결정됨)에 대하여 여전히 사용될 수 있지만, 이것은 입증되지 않은 채로 남아있게 된다.

### 2.1.3.2 3D 세계의 자동의 스케일, 균원, 및 방향

[0217]

평면적인 조사 동안, 그 세계의 정확한 균원, 스케일, 및 방향을 컴퓨팅하는 것이 불가능한데, 이는 그것의 정확한 모양 및 차원들을 포함하는 실제 객체에 대한 이전 지식 또는 호모그래피를 컴퓨팅하기 위한 2개의 뷰들 사이에 사용되는 거리를 요청할 것이기 때문이다. 마지막 경우에, 이것은 그 세계의 균원 또는 방향을 제공하지 않을 것임을 주의하자. 대부분의 시스템들을 그 부분을 고려하지 않는다. 그들은 종종 비디오들 내에 가시적인 3D 객체의 치수들을 입력함으로써 수동으로 스케일을 하도록 사용자에게 요청한다. 다른 옵션들은 사용자가 활영 동안 특정 모델링/애니메이션 소프트웨어 내에서 3D 가상 인물들을 스케일링하게 하고 또한 그것을 실세계에 옮바르게 정렬하는 것이다. 그러나, 비록 작더라도 임의의 스케일 오류는 최종 합성의 품질에 상당한 영향을 끼치므로 그것은 그것이 우리의 기술에서 매우 신중하게 달성되는 이유이다(아래를 참조). 불량 스케일링의 가장

일반적인 관측되는 결과들 중 하나는 객체들이 바닥에 완전히 부착되어 남아있지 않고 바닥에서 미끄러진다고 사용자가 느끼는 것이다.

[0218] 우리는 조사를 수행할 때 장면 내에 올바른 **동록 객체**를 놓도록 요청하는 새로운 방법을 제안한다. 이 객체는 매우 잘 알려진 모양을 갖고 그것의 모든 치수들도 또한 알려져있기 때문에 자동으로 검출된다. 사용자 개입은 전혀 요청되지 않는다.

[0219] 우리의 동록 객체는 반드시 도 3에 나타낸 것과 같은 구들 및 다리들인 몇몇의 부분들로 구성된다. 구들은 우리의 시스템 내의 그 객체의 중요한 부분들인데, 이는 그들이 비디오에 걸친 패턴을 인식하기 위해 블롭 검출 알고리즘들을 사용하도록 허용하기 때문이다. 그 후 투영된 구들(즉, 원들)의 중력 중심들이 컴퓨팅되고, 우리가 모든 구들 사이의 거리 및 치수들을 정확히 알고 있으므로, 시스템은 그 세계에 대한 매우 정확한 스케일을 제공할 수 있다. 우리는 또한 평면 버전의 동록 객체를 갖는데(도 5의 우측 도면을 참조), 그것은 종종 실제 스테이지 상에서 하나의 장소에서 다른 장소로 이동하기에 보다 편하다.

[0220] 이것이 달성되면, 사용자는 예를 들면, 1.80m 높이의 가상 객체를 실제 비디오에 놓을 수 있고, 그것은 정말로 그 높이로 보이고 올바를 것임을 보증한다. 다른 접근법들은 작업의 수동적인 양태로 인해 이러한 품질을 획득하는 것에 실패한다.

[0221] 자동적인 스케일링은 실제로 조사 자체 동안 수행되는 것이지만, 그것은 조사의 시작시에 행해지므로, 그것을 개시화 단계로서 또한 고려하는 것이 중요하다. 컴퓨팅된 스케일은 실제로 3D 자연스러운 마커들의 정확한 세트를 구축하기 위해 다음의 조사 단계들에서 사용된다. 이제 조사 자체를 상세하게 살펴보자.

### 2.1.3.3 조사 처리 및 점 클라우드 구조

[0223] 조사는 사용자가 조사 장치를 사용하여 장면에 걸쳐 움직이도록 요청한다. 시스템은 항상 그것이 이전에 학습된 것들에서 가진 신뢰에 따라 새로운 점들을 컴퓨팅하므로, 조사는 항상 위트니스 카메라가 새로운 후보자들을 추가하기 위해 이전 세트로부터 충분한 점들을 보는 방식으로 구축된다.

[0224] 위트니스 카메라 움직임들 동안, 알고리즘은 **이미지 피라미드**의 4개의 상이한 수준들에서 학습된 3개의 세트들의 데이터로 구성된 **지도**를 생성한다(도 4를 참조):

- 키 포지션. 키 포지션은 위트니스 카메라가 보는 것의 4개의 상이한 해상도들로의 스냅샷을 포함한다. 피라미드의 보다 낮은 수준은 위트니스 카메라의 기초 해상도이다(우리 경우에 640 x 480). 그 키 포지션은 또한 추적이 잘못될 때 카메라 포지션을 복원하기 위해 알고리즘에 의해 사용된다.

[0226] • 키 포지션의 모든 이미지를 내의 FASTER[17]에 의해 추정된 한 세트의 2D 특징 점들. FASTER는 주요 특징이 코너 검출기 알고리즘인 매우 알 알려진 알고리즘이다. 모든 키포인트는 추적을 잃어버렸을 때 훨씬 빠른 복원을 보증하기 위해 부착된 ORB 서술자를 갖는다(2.1.4.4 섹션을 참조).

[0227] • 이전 단계에 의해 검출된 모든 특징 점 상에 중심이 맞춰진 한 세트의 패치들(16 x 16 2D 질감들). 새로운 점들의 검출 동안, 그들의 심도를 컴퓨팅하기 위해 단일 뷰로부터 시작할 방법이 없다. 그것은 패치들이 무엇에 유용한지이다. 등극선 탐색(도 2의, 점 클라우드 구축자 직사각형을 참조)은 2개의 가능한 가까운 키 포지션들로 2개의 패치들 간의 대응관계를 찾음으로써 키 포지션들의 4개의 이미지들에 걸쳐 적용될 수 있다. 한 점(즉, 한 패치)이 두 뷰들 모두에서 검출되면, 3D 특징 점을 컴퓨팅하는 것이 가능하다. 3D 특징 점의 세트는 지도로 불린다. 또한 키 포지션 A와 B 사이의 이 탐색이 동일한 수준의 피라미드 A 및 B를 여행함으로써 (그러나, B는 피라미드 하위-수준임) 달성된다는 것을 이해하는 것이 중요하다(도 4를 참조).

[0228] 지도를 구성하는 동안, 카메라는 조사 오퍼레이터의 변위에 따라 움직인다. 이 때, 우리의 소프트웨어는 이전에 컴퓨팅된(즉, 추적된) 카메라의 매개변수들만을 알고 있다. 카메라의 새로운 포지션을 컴퓨팅하기 위해, 우리는 자연스러운 마커들의 3D 점 클라우드 및 현재의 프레임이 필요하다. 한편으로는, FASTER가 이미지 피라미드의 몇몇의 수준들에서 (현재의 수준 및 현재의 수준+1) 2D 마커들의 세트를 컴퓨팅하고, 다른 한편으로는, 3D 점 클라우드가 현재의 프레임 상에 다시 투영된다. 이 마지막 단계는, 단지 점들이 그것의 관점(viewpoint)으로부터 다시 투영될 때 우리가 카메라의 포지션을 알게 되면 달성될 수 있다. 그러나 그것은 정확히 우리가 컴퓨팅하고자 하는 것이다. 그러므로, 우리의 소프트웨어는 이전에 정확하게 컴퓨팅된 포지션을 사용하여 카메라의 다음 포지션을 예측하기 위해 우리의 9-DOF 센서(2.3 섹션을 참조)에 의해 제공된 정보에 연관된 일정한 속도 모델을 사용한다. 그 예측을 하용하여, 3D 점 클라우드는 현재의 프레임 상으로 다시 투영될 수 있고, 점 매칭 알

고리즘이 FASTER에 의해 검출된 2D 점들과 3D 점 클라우드의 투영된 점들 사이의 대응관계를 찾기 위해 적용된다. 2개의 세트의 마커들 간의 오류는 레벤버그-마倨트[10, 12, 15] 알고리즘을 사용하여 최소화되는데, 이는 그것이 그 종류의 문제들에 대한 최고의 최적화 알고리즘이라고 알려져있기 때문이다. 매칭된 점들의 개수를 투영된 점들의 총 개수로 나눈 것이 주어진 임계보다 크면, 시스템은 카메라를 성공적으로 추적할 수 있고(추적이 양호함) 새로운 키 포지션들을 추가하도록 허용된다. 3D 점 클라우드 내에서 매칭을 찾지 못한 FASTER에 의해 검출된 2D 점들 및 그들의 관련된 2D 16 x 16 질감 패치들은 나중의 사용자를 위해 메모리에 저장된다. 그들은 새로운 3D 특징 점들을 생성하도록 요청될 것이다(다음 문단을 참조).

[0229] 3개의 조건들이 만족되면, 새로운 키 포지션들(및 새로운 특징 점들)이 추가된다. 첫째, 우리가 이전 문단에서 언급한 바와 같이, 추적은 양호해야 하는데, 이는 그것이 충분히 정확하고 길을 잃지 않았다는 것을 의미한다. 둘째, 너무 많은 데이터 세트가 구축되지 않기 위해 키 포지션은 매 30 프레임마다 추가된다(초당 1개). 셋째, 새로운 키 포지션은 가장 가까운 키 포지션으로부터 10cm의 최소의 거리에 있어야 한다. 이것은 시스템이 비유동적일 때 추가의 점들을 학습하는 것을 방지한다.

[0230] 모든 이들 테스트들을 성공적으로 통과하면, 소프트웨어는 새로운 스냅샷 및 새로운 3D 특징 점들을 추가할 수 있다. 첫번째 단계는 새로운 프레임(스냅샷) 내에서 새로운 관련된 2D 점들을 검출하기 위해 FASTER를 또다시 사용하는 것이다. 우리는 추적으로부터의 매칭되지 않은 한 세트의 2D 특징 점들을 가지므로, 우리는 이제 "오래된" 2D 특징 점 세트와 새롭게 컴퓨팅된 것을 매칭시키고자 한다. 이것은 새로운 점들의 세트의 2D 질감 패치들을 사용하여 전형적인 등극선 탐색에 의해 수행된다. 그 패치들은 "오래된" 세트의 점들의 2D 패치들의 세트를 매칭시키고자하는 등극선들(도 2를 참조)을 따라 움직인다. 등극선 제약 때문에, 2개의 2D 점들의 매칭으로부터의 3D 점의 심도를 컴퓨팅하는 것이 가능하다. 이것은 호모그래피 단계(2 섹션을 참조)의 매칭 단계와 실제로 유사하다. 새로운 점들이 지도에 추가되면, 로컬 번들(bundle) 조정이 적용된다. 이들 새로운 특징 점들을 사용하여 카메라의 포지션 및 회전을 개선한 후, 모든 특징 점들의 3D 포지션들을 개선하기 위해 레벤버그-마倨트 최소화 알고리즘이 사용된다. 이것은 k+1개의 프레임들 원도우 상에서 수행되고, 이는 카메라의 현재 포지션 및 회전을 올바르게 추정하기 위해 (로컬 번들 조정) 그 최소화가 현재의 스냅샷과 k개의 가장 가까운 것들을 고려한다는 것을 의미한다. k의 값은 그 최소화 알고리즘에 대해 주어진 오류 임계에 따라 경험적으로 결정되거나 적응적으로 컴퓨팅될 수 있다.

[0231] 로컬 번들 조정 이외에, 글로벌 번들 조정이 또한 모든 기준의 키 포지션들에 적용된다. 키 포지션들의 세트가 시간이 지남에 따라 점점 더 커지므로, 시스템은 글로벌 조정을 수행하기 위해 점점 더 많은 시간을 취한다. 새로운 키 포지션들이 추가되고 우리의 소프트웨어가 로컬 및/또는 글로벌 번들 조정들을 수행하기에 충분한 시간을 갖지 못하면, 새로운 키 포지션들의 추가에 우선순위를 주기 위해 조정들은 취소된다. 그들은, 새로운 키 포지션들이 추가되자마자, 또다시 적용될 것이다.

[0232] 기준의 점 클라우드가 존재하면, 카메라는 추적될 수 있다. 이것은 다음 섹션에 설명되어 있다.

#### 2.1.4 두번째 단계: 필름 카메라 추적

##### 2.1.4.1 기하학적인 카메라 눈금

[0233] 카메라 눈금은 몇 개의 목표들을 갖는다. 한편으로는, 그것은 이미지가 센서들 상에서 어떻게 생성되는지를 설명하기 위한 수학적인 관계를 우리에게 제공한다. 다른 한편으로는, 그것은 3D 점들과 카메라의 이미지 평면 상으로의 그들의 투영 사이의 이동 행렬을 컴퓨팅한다.

[0234] 조사 및 추적 처리의 이 중요한 전제 조건은 우리가 본질적인 및 외적인 매개변수들의 2개의 상이한 종류의 속성을 컴퓨팅해야 한다는 것을 의미한다. 본질적인 매개변수들은 카메라 및 렌즈 속성들에 관련되고, 그들이 물리적으로 구축되는 방식에 따라 움직인다. 그들은 이미지 포맷, 카메라의 광학적 중심과 이미지 평면 사이의 거리, 및 주점을 포함한다. 이미지 포맷은 화소들을 거리에 관련시키는 2개의 스케일 인자들로 구성된다. 주점은 이미지 평면(이상적으로 이미지의 중간에 있음) 상으로의 광학적 중심의 투영이다. 그것 이외에, 몇몇의 렌즈들은 그들의 광학적 왜곡이 컴퓨팅되고 추적을 고려할 것을 요청한다. 눈금 처리 동안, 몇개의 원 패턴들을 갖는 평면 표면이 카메라에 의해 연속적으로 캡쳐된다(도 5를 참조). 모든 포지션들에 대하여, 외적인 매개변수들(즉, 카메라의 포지션 및 회전) 및 본질적인 매개변수들(광학 왜곡을 포함함)은 모든 스냅샷들 간의 재투영의 오류를 최소화하는 레벤버그-마倨트 최소화 알고리즘을 통해 컴퓨팅된다.

[0235] 이 처리는 데버네이 및 포즈하의 워크들(works)[3]에 기초하는 매우 유명한 기술이다.

#### [0238] 2.1.4.2 필름 카메라 오프셋

[0239] 여기서 개발된 글로벌 시스템은 추적 장치를 목표된 필름 카메라에 직접 부착되게 놓도록 요청한다. 위트니스 카메라 포지션 및 방향은 이전 서술자 추적 기술을 사용하여 추정된다. 그러나, 필름 카메라와 위트니스 카메라 모두에 대하여 동일한 패턴을 사용함으로써 스케일의 문제가 존재한다. 실제로, 우리의 위트니스 카메라들은 180도 렌즈들을 사용하므로, 눈금 패턴 내의 원 모양들이 위트니스 카메라 뷰 내에서는 정말로 작게 보이는 경향이 있어서 그들을 검출하기 어려워진다.

[0240] 우리는 실제로 2개의 상이한 크기 및 대조적인 패턴들로 구성된 우리의 교정 차트를 만들었다. 그 차트의 한 부분은 완전히 흰색의 배경 상에 완전히 검은색의 원들을 갖는다(위트니스 카메라들 용). 그 차트의 두번째 부분은 완전히 검은색의 배경 상에 완전히 흰색의 원들로 구성된다(필름 카메라). 따라서 우리는 매우 간단한 패턴 검출기를 사용하여 위트니스 및 필름 카메라들에 대한 두 패턴들 모두를 찾는다. 우리가 두 패턴들 모두의 정확한 관련 포지셔닝을 알기 때문에, 필름 카메라의 포지션 및 회전은 위트니스 카메라의 계산된 포지션 및 회전을 "전파"함으로써 찾아질 수 있고, 그 반대일 수 있다. 실제로, 두 카메라들 모두가 서로 직접 부착되므로, 우리는 그들 간의 3D 오프셋(R, T)을 정확하게 컴퓨팅하기 위해 라인업(lineup) 알고리즘을 사용할 수 있다. 우리는 이전에 설명된 원 패턴들에 기초하여 라인업 알고리즘을 개발했다. 정확한 크기들, 모양들, 및 포지션들이 알려진 패턴들은 두 카메라들 모두에 의해 개별적으로 인식되며, 따라서, 위트니스 카메라와 필름 카메라 사이의 사소한 수학적인 관계(회전 및 병진)가 높아진다.

#### [0241] 2.1.4.3 추적 단계

[0242] 조사를 추적 처리로부터 분리시키는 것은 시스템이 세트 상에서 사용되는 방법을 설명하는 단지 실용적인 방식이다. 실제로, 추적 처리는 도 2에 설명된 추적과 상이하지 않으며 조사 동안 사용된다. 그러나, 그것은 자명한 이유들을 위해 터치스크린을 포함하지 않는 점을 제외하고 조사 장치와 상당이 유사한 보다 단순한 장치(도 6을 참조)를 사용한다.

[0243] 도 2의 가시적 추적 직사각형은 새로운 프레임 내의 FASTER에 의해 검출된 2D 자연스러운 마커들과 기존의 3D 점 클라우드 사이에 어떻게 매칭이 발견되는지를 나타낸다. 모든 자연스러운 마커와 연관되고 두 데이터 세트를 모두로부터의 2D 질감 패치들은 선형적으로 비교된다. 모든 특징 점들이 처리되면, 최고의 가능한 매치를 찾고 카메라의 포지션 및 방향을 컴퓨팅하기 위해 가우스 뉴튼 알고리즘이 적용된다. 이러한 최소화 알고리즘은 솔루션을 모으기 위해 대략 10 회 반복을 보통 요청한다. 최소화하기 위한 기준을 강력하게 하기 위해, 우리는 고정적인 터키(Tukey)-M 추정기(최대 가능성 유형 추정기)[6]를 사용한다. 이것은 또한 모으는 것이 아웃라이어들에 의해 방해되지 않음을 보증한다.

[0244] 모든 추적은 카메라의 최종 매개변수들을 획득하기 전에 피라미드의 2개의 상이한 수준들 상에서 2번 수행된다. 첫째, 알고리즘의 모으기는 이미지 피라미드의 거친(coarse) 수준을 통해 카메라의 포지션 및 방향의 근사치를 컴퓨팅하고 자연스러운 마카들의 하위세트를 사용하여 상당히 증가된다. 둘째, 동일한 컴퓨팅이 매우 보다 큰 세트(대략 20배 큼) 상에서 수행되고 최종의 정확한 결과를 획득하기 위해 이전의 근사치들과 병합된다.

#### [0245] 2.1.4.4 추적 실패들 및 복원

[0246] 실제 세트 상에서의 카메라 추적 동안, 가시적 추적이 길을 잊는 많은 상황이 발생할 수 있다. 이것은, 예를 들면, 사람이 위트니스 카메라 앞에 서있거나 바가 객체들에 의해 단순히 가려져있을 때 종종 발생한다. 우리가 가시적 추적을 잊어버리는 모든 경우들에 대하여, 3개의 상이한 복원 스크립트들이 동시에 시작되고 첫번째로 복원을 획득하는 것이 받아들여진다.

[0247] 확장형 칼만 필터의 예측 기능을 사용하여(2.3.2 섹션을 참조), EKF 예측된 매개변수들로부터 탐색을 시작하는 카메라 포지션 복원을 첫번째로 시도하는 것이 가능하다. 실제로, EKF가 예측-정정 알고리즘이므로, 그것은 항상 카메라가 다음 포지션으로 움직이기 전에 무엇이 카메라의 다음 포지션이 될지를 예측고자 한다. 예측이 올바르면, 인스턴트 복원을 획득하는 것이 가능하다. 실제로, 그 예측은 위트니스 카메라가 숨겨지고 시스템이 IMU를 사용하여 추적할 때 매우 잘 동작한다. 그러나, 그것은, 위트니스 카메라들을 숨기면서 필름 카메라가 상당히 병진되면, 종종 실패한다.

[0248] 제 2 수준의 복원은 매우 상이하다. 모든 특징 점은 그것에 연관된 회전 불변 서술자를 가지며, ORB(지향성 FAST 및 회전된 BRIEF) 기술[18]을 따른다. 그것은 우선 해리스(Harris) 점수[5]에 연관된 다중-스케일 특징들(즉, 이미지 피라미드의 상이한 수준들로의)을 컴퓨팅하기 위해 유명한 FAST[16] 기술에 의존한다. 서술자의 방

향은 세기 도심(intensity centroid)을 사용하여 컴퓨팅된다. 제 2 서술자는 키포인트의 방향을 따라 회전된 BRIEF 서술자[2]인 rBRIEF이다. 이를 2개의 서술자들을 사용하여, 우리의 기술은 정규 가시적 추적 기술들이 실패하는 포지션들로부터 복원할 수 있다. 예를 들면, 우리가 임의의 이유로 가시적 추적을 잊어버렸다고 하자. 여전히 잊어버린 동안, 우리가 이제 카메라를 Z 축 주변에서 병진시키고 90도 회전시켰다고 가정하자. 카메라에 의해 보여지는 현재의 이미지는 이전에 절대 학습되지 않았었지만, 그것은 우리가 이전에 학습된 키 포인트들 및 키 포지션들을 추가하는 방향을 여전히 가리키고 있다. 그러나, 정규 복원 처리는 회전 불변적이므로, 그들 시스템들은 추적을 복원하는데 실패한다. 우리의 경우, 모든 키 포인트에 연관된 회전 불변적 서술자들의 사용은, 필름 카메라가 이전에 그 포지션으로 어느 곳을 가리키자마자(회전은 관련 없음), (무작위(brute force) 매칭 기술을 사용하여) 빠른 복원을 허용한다.

[0249] 마지막으로, 제 3 복원 스레드는 보다 무작위하고 컴퓨팅적으로 집약적이다. 타원체(3D 타원)는 마지막으로 알려진 카메라 포지션 주변에 구축된다. 그 타원체는 그러나 확장형 칼만 필터에 의해 컴퓨팅된 매개변수 공분산으로 구성된 크기를 갖는다. 그 후 다수의 샘플들이 EKF의 오류 공분산을 따르는 타원체 내에서 생성된다. 탐색된 카메라 포지션 및 방향에 대한 최고의 가능한 매칭을 찾기 위해 RANSAC 알고리즘이 그 세트에 적용된다.

## 2.2 동적인 줌, 초점, 및 조리개

[0251] 글로벌 추적 처리 동안, 감독 또는 사진 감독은 예를 들면, 장면 내의 인물/객체를 줌하거나 그것에 초점을 맞추도록 결정할 수 있다. 우리의 기술은 알리(Arriflex)에 의해 제안된 것과 같은 몇몇의 정규 인코더들을 통합하지만, 또한 실시간으로 줌, 초점, 및 조리개 값을 캡처하기 위해 임의의 다른 인코더를 다룰 수도 있다. 이것은 세트 상의 동적인 주밍(zooming) 및 피사계 심도를 허용한다.

## 2.3 센서 결합 기술

[0253] 모든 완전한 광학 기반 마커없는 기술들은 많은 상황들에서 추적을 실패할 수 있다. 가장 일반적인 경우는 위트니스 카메라의 현재의 뷰에 충분한 자연스러운 마커들이 남아있지 않은 것이다. 이상적인 상황들에서, 이것은 거의 일어나지 않는데, 이는, 예를 들면, 사용자들은 그의 손으로 렌즈를 가리지 않도록 주의를 기울이기 때문이다. 실제 세트들에서, 그 상황은 많이 발생하는데, 이는 단순히 카메라 팀원이 종종 렌즈를 바꾸고, 카메라 리그(rig)를 변경하고, 렌즈를 청소하고, 다른 촬영 장소로 움직이는 등을 할 필요가 있기 때문이다. 정규 마커-기반 및 마커없는 (광학 흐름 기반) 기술 상에서, 이것은 추적 시스템이 카메라의 포지션 및 회전에 대한 양호한 추정을 찾지 못하게 한다.

[0254] 우리는 위트니스 카메라들, 자이로스코프, 가속도계, 및 자력계의 다수의 센서들을 사용하여 위트니스 카메라의 포지션 및 회전을 자동으로 정정하는 센서 결합 방법을 만들었다. 그 기술은 우리 기술의 핵심이다. 제 1 단계는 시스템의 눈금 매기기인데, 이는 카메라에 상대적으로 9-DOF IMU의 눈금을 매기는 것을 의미한다. 그것은 모든 센서들이 실제로 동일한 기반에서 동작한다는 것을 보증하기 위해 중요한 단계이다. 9-DOF 센서와 위트니스 카메라들 간의 관계는 손-눈 기술[20]와 비슷하지만, 우리는 그것을 우리의 추적 알고리즘에 맞게 하기 위해 몇몇의 관심있는 과학적인 의견들을 추가했다.

### 2.3.1 IMU-위트니스 카메라 눈금

[0256] 9-DOF 센서(IMU)로부터 2개의 위트니스 카메라들(C) 중 하나로의 회전을 컴퓨팅하는 변형  $R_{IMU \rightarrow C}$ 를 고려하자.  $R_{IMU}$ 는 그것을 국부적인 기반에 알려져있고,  $R_C$ 도 또한 그것의 국부적인 기반에서 알려져있다. 그 목적은

$$R_C = R_{IMU \rightarrow C} \cdot R_{IMU} \text{에 대하여 } R_{IMU \rightarrow C} \text{ 를 푸는 것이다.}$$

[0257] IMU와 카메라들 사이의 관계가 시간이 지나도 변하지 않기 때문에,  $R_{IMU \rightarrow C}$ 는 항상 고정적이라는 것이 보증된다(그들은 모두 고정적인 카메라 바에 부착됨).

[0258] 따라서 그것은  $t_i, i \in [0, n]$ 에 대해 다음과 같이 작성될 수 있다:

$$R_C^{t_0} = R_{IMU \rightarrow C} \cdot R_{IMU}^{t_0}$$

⋮

$$R_C^{t_{n-1}} = R_{IMU \rightarrow C} \cdot R_{IMU}^{t_{n-1}}$$

$$R_C^{t_n} = R_{IMU \rightarrow C} \cdot R_{IMU}^{t_n}$$

[0259] 우리는 이전 수학식들로부터

$$R_C^{t_i - t_{i-1}} = R_C^{t_i} \cdot (R_C^{t_{i-1}})^{-1}$$

$$R_{IMU}^{t_i - t_{i-1}} = R_{IMU}^{t_i} \cdot (R_{IMU}^{t_{i-1}})^{-1}$$

[0261] 의 관계들을 갖는

$$R_C^{t_1 - t_0} = R_{IMU \rightarrow C} \cdot R_{IMU}^{t_1 - t_0}$$

⋮

$$R_C^{t_{n-1} - t_{n-2}} = R_{IMU \rightarrow C} \cdot R_{IMU}^{t_{n-1} - t_{n-2}}$$

$$R_C^{t_n - t_{n-1}} = R_{IMU \rightarrow C} \cdot R_{IMU}^{t_n - t_{n-1}}$$

[0262] 를 추론할 수 있다.

[0263] 자이로스코프에 의해 제공된 회전 측정 및 가시적 추적은 같도록 제안된다. 따라서,  $R_{IMU}$  및  $R_C$ 가 동일한 회전 각

도를 (그러나 상이한 기반들 내에서) 기술한다고 고려하는 것이 논리적이다.  $\text{Log} : SO(3) \rightarrow so(3)$ 로 정의되는 리대수(Algebra Lie) 내에서 회전 행렬의 로그 지도를 사용하면, 행렬  $R_C$ 를 축-각도  $(\vec{r}_c, \alpha)$  표현으로 변환하는 것이 가능한데:

$$\alpha = \arccos \left( \frac{\text{tr}(R_C) - 1}{2} \right)$$

$$\vec{r}_c = \frac{1}{2 \sin \alpha} \begin{bmatrix} R_c(3, 2) & R_c(2, 3) \\ R_c(1, 3) & R_c(3, 1) \\ R_c(2, 1) & R_c(1, 2) \end{bmatrix}$$

[0264] 여기서,  $\text{tr}(R_C)$ 는  $\sum_{j=1}^3 R_C(j, j)$ 와 같은 행렬의 대

각합(trace)이다. 우리는 이제 다음의 중복결정된(overdetermined) 수학식 시스템을 작성할 수 있다:

$$S = \begin{cases} \vec{r}_c^{t_1 - t_0} = R_{IMU \rightarrow C} \cdot \vec{r}_{IMU}^{t_1 - t_0} \\ \vdots \\ \vec{r}_c^{t_{n-1} - t_{n-2}} = R_{IMU \rightarrow C} \cdot \vec{r}_{IMU}^{t_{n-1} - t_{n-2}} \\ \vec{r}_c^{t_n - t_{n-1}} = R_{IMU \rightarrow C} \cdot \vec{r}_{IMU}^{t_n - t_{n-1}} \end{cases} \quad (8)$$

[0265] [0266] 임의의 중복 결정된 시스템에 대하여,  $S = U \sigma V^t$ 의 단일 값 분해(Singular Value Decomposition(SVD))을 사용하는 이전의 직사각형 행렬을 풀 수 있는데, 여기서,  $U$ 는  $S \cdot S^t$ 의 고유 벡터들의  $m \times n$  행렬이고,  $V$ 는 고유 벡터들  $S^t \cdot S$ 의  $n \times n$  행렬이고,  $D$ 는  $S$ 의 순서화된 단일 값들의  $n \times n$  대각선 행렬이다.

[0267] 수학식의 이전 해(solution)은 측정들로부터의 잡음에 영향을 받았다. 따라서, 우리는 오류를 최소화하기 위해

레벨버그-마쿼트 최소화 방식을 적용한다:

$$E = \|(R_{IMU \rightarrow C} \times \overrightarrow{r_{IMU}}^{t_i - t_{i-1}}) \cdot \overrightarrow{r_C}^{t_i - t_{i-1}} - 1\| \quad (9)$$

[0268] 이제 우리는 IMU와 위트니스 카메라들 사이의 수학적인 관계를 가지므로, 모든 센서 중 최고를 추출하고 필름 카메라의 회전 및 병진을 컴퓨팅하기 위해 확장형 칼만 필터에 모든 측정을 주입할 수 있다.

### 2.3.2 확장형 칼만 필터

[0271] 확장형 칼만 필터(EKF)는 다수의 상이한 센서들로부터의 상이한 데이터를 병합하기 위한, 항공우주 및 로보틱 산업 영역에서 매우 잘 알려진 기술이다. 그것은, 추정 오류를 컴퓨팅하고 예측 및 정정의 2개의 단계들을 적용함으로써 잡음 측정들로부터 비선형 시스템의 매개변수들을 추정하기 위해 반드시 사용된다. 예측 단계에서, 칼만 필터는 새로운 매개변수들 뿐만 아니라, 그 매개변수들을 사용하는 한 시간 단계(timestep)로의 관련된 오류 및 이전 시간 단계에서 추정된 오류를 컴퓨팅한다. 정정 단계는 새로운 잡음 측정들을 사용하여 예측을 개선한다. 그들 새로운 측정들은 예측된 매개변수들 및 오류로부터 새로운 추정된 매개변수들 및 오류를 컴퓨팅하도록 허용한다. 오류가 주어진 임계보다 크면, 매개변수들은 정정된다.

[0272] 여기 우리의 경우에서, 우리가 3D 회전들에 관련된 다수의 측정들을 갖고(자이로스코프 + 위트니스 카메라들) 그들이 용이하게 보간되지 못하므로, EKF는 매우 특정한 방식으로 개발된다.

[0273] 우리의 기술에 대한 EKF의 아이디어는 모든 측정(위트니스 카메라, 자이로스코프, 자력계, 가속도계)에 몇몇 종류의 신뢰 값을 주고 최고의 신뢰 점수를 획득하는 센서를 사용하여 현재의 추정된 포지션 및 회전을 반복적으로 정정하는 것이다. 실제로, EKF는 그것보다 약간 복잡하고, 비선형 시스템들에 대하여 수학적으로 예측-정정 알고리즘으로 기술될 수 있다. 시간 단계( $t$ )에서의 상태 벡터( $\hat{X}(t)$ )를 고려하면, 그 알고리즘은  $\hat{X}(t)^{\Delta t}$ 가 되는 벡터( $\hat{X}(t)$ )의 새로운 상태를 예측하기 위해 우리의 움직임 수학식들의 현재의 물리적 모델을 적용한다. 정정 단계는 그 후 시간 단계( $t + \Delta t$ )에서 새로운 상태 벡터( $\hat{X}(t + \Delta t)$ )를 생성한다.

[0274] 우리의 경우에, EKF 모델은 다음과 같이 작성될 수 있는데:

$$\begin{cases} Y(t) = h(X(t), B) \\ \hat{X}(t)^{\Delta t} = f(\hat{X}(t), \Delta t) \end{cases} \quad (10)$$

[0275] 여기서,  $h$ 는 우리의 가측치(observables)에 대한 상태 방정식들이고,  $X_t$ 는 글로벌 기반 내의 병진, 회전, 병진 속도 및 각속도의 벡터이고,  $f$ 는 상태 벡터의 개선 함수이다.  $B$ 는 시스템 내의 글로벌 잡음이다.

#### $h$ 의 정의

$$Y(t) = h(X(t), B) \quad (11)$$

[0276] 한편으로는 우리가  $h_{witness}$ 와 같은 마커없는 가시적 추적으로부터의 측정들을 고려하고(즉, 여기서 위트니스 카메라들), 다른 한편으로는  $h_{IMU}$ 와 같은 IMU로부터의 측정들을 고려하면, 우리는  $h$ 는 두 종류의 가측치  $h_{witness}$  및  $h_{IMU}$ 로 구성된다고 말할 수 있다. 정의에 의해,  $h$ 는 사소한 함수이고 그것의 일반적인 형태는 식별:

[0277]  $Y_{witness}(t) = X(t)$  및  $Y_{IMU}(t) = X(t)$ 에 의해 주어진다. 측정들을 송신하는 센서들의 종류에 따라, 우리는 이제 2개의 상이한 버전의  $h$ 를 작성할 수 있고 따라서 가시적 추적 경우에 대하여 수학식 (11)을 다음과 같이 다시 작성할 수 있는데:

$$[0280] Y(t) = \begin{cases} Y_{witness}(t).T &= X(t).T \\ Y_{witness}(t).R &= X(t).R \\ Y_{witness}(t).\vec{V} &= X(t).\vec{V} \\ Y_{witness}(t).\vec{\Omega} &= X(t).\vec{\Omega} \end{cases} \quad (12)$$

[0281]  $Y_{witness}(t).\{\alpha \in T, R, V, \vec{\Omega}\}$  는 병진(T, 미터로), 회전(R, 반지름으로), 속도( $\vec{V}$ ,  $m \cdot s^{-1}$ 로), 및 각속도( $\vec{\Omega}$ ,  $rad \cdot s^{-1}$ 로)에 대한 시간 단계(t)에서의 가시적 추적의 가측치( $y (y \in \mathbb{R}^{12})$ )이다.

[0282] 센서가 자이로스코프인 경우, 수학식 (11)은 다음과 같이 되는데:

$$[0283] Y(t) = \begin{cases} Y_{IMU}(t).R &= X(t).R \\ Y_{IMU}(t).\vec{\Omega} &= X(t).\vec{\Omega} \end{cases} \quad (13)$$

[0284]  $Y_{IMU}(t).\{\alpha \in R, \vec{\Omega}\}$  는 회전(R, 반지름으로) 및 각속도( $\vec{\Omega}$ ,  $rad \cdot s^{-1}$ 로)에 대한 시간 단계(t)에서의 IMU의 가측치이다. 실제로 9-DOF 센서를 사용하는 컴퓨팅 가능한 병진 또는 속도는 없다.

### f의 정의

$$[0286] \hat{X}(t)^{\Delta t} = f(\hat{X}(t), \Delta t) \quad (14) \quad \text{이고},$$

[0287]  $\hat{X}(t)^{\Delta t}$  는 시간 단계( $t + \Delta t$ )에서 예측된 상태 벡터이고,  $\hat{X}(t)$ 는 시간 단계(t)에서 추정된 상태 벡터이다.

[0288] 함수 f는 상태 벡터의 갱신 함수이고 다음과 같이 정의된다:

$$[0289] \hat{X}(t)^{\Delta t} = \begin{cases} \hat{X}(t).T + \hat{X}(t).V \cdot \Delta_t \\ \hat{X}(t).R \times \hat{X}(t).\vec{\Omega} \cdot \Delta_t \\ \hat{X}(t).\vec{V} \\ \hat{X}(t).\vec{\Omega} \end{cases} \quad (15)$$

[0290] 회전 행렬들이 보간되지 못하고 선형 형태  $A \cdot X + B$ 로 작성되므로, 칼만 필터의 정규 갱신 함수는 우리의 경우에 사용불가능하다는 것을 주의하자.

### 예측 단계

[0292] 예측 단계는 칼만 필터의 정의를 사용하여 작성될 수 있는데:

$$[0293] \hat{X}(t)^{\Delta t} = f(\hat{X}(t), \Delta t) \quad (16)$$

$$P_t^{\Delta t} = A_t \cdot P_t \cdot A_t^T + Q \cdot \Delta t \quad (17)$$

[0294] 여기서,

[0295] •  $P_t^{\Delta t}$  는 시간 단계( $t + \Delta t$ )에서의 예측된 오류의 공분산 행렬이고,

[0296] •  $P_t$ 는 시간 단계(t)에서의 오류의 공분산 행렬이고,

- [0297] •  $A_t = \frac{\partial f}{\partial X_t} | \hat{X}_{t-1}$  와 같은 전이 행렬이고,  $A_t$ 는  $f$ 의 야코비안 행렬(jacobian matrix)이고 유한 차분법의 합  $\frac{f(x+\Delta)-f(x)}{\Delta}$  으로 컴퓨팅되고,

- [0298] •  $Q$ 는 상태 벡터의 잡음의 공분산 행렬  $12 \times 12$ 이다. 이동 상태 벡터의 잡음은 빌경적으로  $1e^{-2}$ 에 놓일 뿐만 아니라, 회전 상태 벡터의 잡음은  $1.e^{-3}$ 에 놓인다.

$$Q = \begin{bmatrix} 0.01 & 0.0 & 0.0 & 0.0 \\ T & 0.001 & 0.0 & 0.0 \\ 0.0 & R & 0.0 & 0.0 \\ 0.0 & 0.0 & V & 0.0 \\ 0.0 & 0.0 & 0.0 & \Omega \end{bmatrix}$$

- [0299] [0300] 정정 단계

- [0301] 칼만 정의에 따라, 우리의 경우의 정정은 다음과 같이 작성되는데:

$$K(t + \Delta t) = P_t^{\Delta t} \cdot H^T(t + \Delta t) \cdot (U(t + \Delta t) + H(t + \Delta t) \cdot P_t^{\Delta t} \cdot H^T(t + \Delta t))^{-1} \quad (18)$$

$$P(t + \Delta t) = (I - K(t + \Delta t) \cdot H(t + \Delta t)) \cdot P_t^{\Delta t} \quad (19)$$

$$\hat{X}(t + \Delta t) = \begin{cases} \hat{X}(t)^{\Delta t} \cdot T + K(t + \Delta t) \times J_{witness|IMU} \cdot T \\ \hat{X}(t)^{\Delta t} \cdot R + K(t + \Delta t) \times J_{witness|IMU} \cdot R \\ \hat{X}(t)^{\Delta t} \cdot V + K(t + \Delta t) \times J_{witness|IMU} \cdot V \\ \hat{X}(t)^{\Delta t} \cdot \Omega + K(t + \Delta t) \times J_{witness|IMU} \cdot \Omega \end{cases} \quad (20)$$

- [0302] [0303] 여기서,

- [0304] •  $I$ 는 항등 행렬이고,

- [0305] •  $K(t + \Delta t)$ 는 시간 단계( $t + \Delta t$ )에서의 칼만 이득(gain)이고,

- [0306] •  $H$ 는 함수  $h$ 의 야코비안 행렬이고, 그 자체는 현재의 선택된 센서에 따라  $h_{witness}$  또는  $h_{IMU}$ 이고

$$H(t + \Delta t) = \frac{\partial h}{\partial X} | \hat{X}_t^{\Delta t} \quad (21)$$

- [0307] [0308] 우리의 경우, 야코비안은 유한 차분법의 합으로 컴퓨팅되고

$$\frac{h_{witness|IMU}(x + \Delta) - h_{witness|IMU}(x)}{\Delta},$$

- [0309] [0310] •  $U(t + \Delta t)$ 는 현재 선택된 센서(위트니스 카메라 또는 IMU)의 잡음 측정들의 공분산 행렬이다. 그 값은

측정의 품질(신뢰 欲)에 따라 수정된다. 가시적 추적의 경우, 그것을 추적 오류  $\varepsilon$ 에 의존하고

$$U(t + \Delta t) = \begin{cases} 0.0001 & \text{if } \varepsilon < 1.0 \text{ pixel} \\ 10.0 & \text{if } \varepsilon \geq 1.0 \text{ pixel} \end{cases} \quad (22)$$

[0312] 잡음 오류들은 모든 센서들에 대한 제조업자들에 의해 직접 주어지기 때문에, IMU에 대한  $U(t + \Delta t)$ 를 결정하는 것이 매우 용이하고, 이전 수학식은 그들 오류들에 의해 바로 대체되고,

- $P(t + \Delta t)$ 는 시간 단계( $t + \Delta t$ )에서의 오류의 공분산 행렬이고,

- $V$  및  $\Omega$ 는 각각 속도 벡터( $\vec{V}$ ) 및 각속도( $\vec{\Omega}$ )의 행렬 표현이고,

- 수학식(20)의  $J_{witness|IMU}$ 는 또한 센서들의 종류를 따라 다음과 같이 작성되고:

$$J_{witness}.T = y(t + \Delta t).T - \hat{Y}_t^{\Delta t}.T$$

$$J_{witness}.R = y(t + \Delta t).R \times \hat{Y}_t^{\Delta t}.R^{-1}$$

$$J_{witness}.V = y(t + \Delta t).V - \hat{Y}_t^{\Delta t}.V$$

$$J_{witness}.\Omega = y(t + \Delta t).\Omega \times \hat{Y}_t^{\Delta t}.\Omega$$

$$J_{IMU}.R = y(t + \Delta t).R \times \hat{Y}_t^{\Delta t}.R^T$$

$$J_{IMU}.\Omega = y(t + \Delta t).\Omega \times \hat{Y}_t^{\Delta t}.\Omega^{-1}$$

- $y(t + \Delta t)$ 은 선택된 센서에 의해 주어진 현재의 측정이다.  $y(t + \Delta t).(\alpha \in \{R, T, V, \Omega\})$

[0316] 는 병진, 회전, 속도, 및 각속도 각각의 측정이고,  $J.(\alpha \in \{R, T, V, \Omega\})$ 는 칼만 이노베이션(Kalman Innovation)으로 알려져 있다.

#### 2.4 3D 재구성 및 심도 키잉(keying)

[0317] 실시간 3D 재구성은 우리의 카메라 바에 통합된 3D 거리 센서를 사용하여 달성된다. 그 거리 센서는 구조된 가벼운 장치(예를 들면, 마이크로소프트 키넥트(Kinect), 에이수스 액션(Asus Xtion)), 또는 플래시 라이더(Flash Lidar)로도 알려진 ToF 카메라(예를 들면, 개선된 과학적 개념의 타이거아이(TigerEye)와 같음)일 수 있다. 두 기술들 모두는 제한사항들을 가지며, 우리의 기술은 우선은 단지 단거리 센서 에이수스 액션을 사용한다. 보통, 눈금 매기는 단계는 위트니스 카메라들에 상대적으로 에이수스 액션의 포지션을 알기 위해 요청된다. 우리는 단순히 필름 라인업에 대하여 위트니스에 대해서와 정확히 동일한 기술을 사용한다(2.1.4.2 섹션을 참조). 거리 센서는 그레이스케일 카메라를 가지므로, 우리는 거리 센서의 카메라 및 우리의 위트니스 카메라들 중 하나의 두 카메라들 모두에서 검출된 눈금 패턴을 갖기 위해 우리의 통상의 눈금 패턴을 사용한다. 그 처리는 매우 간단하고 2.1.4.2와 유사하다.

[0318] 거리 센서는 우리의 기술에서 2가지 목적을 위해 사용된다. 한편으로는, 예를 들면, 그것은 객체들 뒤로 멀리 떨어져 생성된 불량 특징 점들을 삭제하는 것을 돋는다. 거리 센서의 심도 지도 상으로 특징 점들을 간단히 재투영하는 것은, 특징 점이 올바른지 아닌지를 알려줄 수 있다. 입체적인 추적 동안(2.5 섹션을 참조), 등극선 기하에 의해 생성된 모든 점은 여전히 등극선들에 의해 주어진 추정보다 가깝거나 멀 수 있다. 우리는 재구성된 3D 점을 거리 센서의 심도 지도로의 그것의 재투영의 현재의 값과 비교하여, 그것의 정확도를 향상시키고 간단히 그것을 거부한다.

[0321] 거리 센서의 두번째 사용은 심도 지도 자체에 의존한다. 그것이 12m의 최대 거리에서 심도를 제공하면, 우리는 가상 인물들/객체들에 의해 세트 내에서 그 거리 내에 있는 모든 사람(연기자)을 폐색할 수 있고, 그 반대일 수도 있다. 이것은 동적인 폐색이라고 불리고 영화 및 생방송에서 시각 효과들에 대한 또 다른 현실감 수준을 높인다. 우리의 기술은, 심도를 필름 카메라 상으로 다시 투영하고 결과를 개선시킴으로써 거리 센서의 해상도를 향상시키도록 (가장자리 향상) 주요 필름 카메라 공급을 사용한다(이것은 또한 추가적인 고해상도 위트니스 카메라를 추가하여 달성을 할 수 있음). 이러한 기술은 실시간 심도 키잉을 제공할 수 있으며, 따라서, 특히 보다 강력한 거리 센서가 사용될 때(플래시 라이더), 바인딩(binding) 녹색 화면을 제거한다.

## 2.5 입체적인 마커없는 기술

[0323] 평면적인 접근법을 사용하는 이전에 설명된 기술은 실제로 실제 상황들에 충분히 효율적이지 않다. 그것은 알려진 실제 객체가 그 세계의 스케일, 방향, 및 포지션을 컴퓨팅하기 위해 등록으로서 사용될 것을 요청하므로, 당신이 객체를 등록(검출)하고자 할 때 많은 사람들이 항상 당신의 경로 상에 있는 영화들 상에서 그것을 사용하기가 종종 어렵다. 그러나, 그 평면적인 접근법의 주요 제한사항은 그것이 장면을 조사하도록 요청한다는 사실이다. 모노그래피가 그 시스템의 개시화를 위해 사용되어 하므로, 그것은 기술자가 그것을 조사하기 시작하기 위해 세트로 와야한다는 것을 의미한다. 사실, 세트 상에서 걸어다니는 사람들은 가장 마지막 초들 이전에 그 장소를 절대로 떠나지 않으므로, 그 동작은 성공적으로 달성이 힘들다. 우리는 그 제한사항을 제거하기 위해 입체적인 접근법을 구축하므로, 인스턴트 조사 및 추적의 완전히 근본적인 개념을 가져온다.

### 2.5.1 인스턴트 조사

[0325] 카메라 추적 내에 내장된 입체적인 기술은 이전에 설명된 우리의 평면적인 접근법을 포함하는 임의의 다른 기술에 비교되는 중대한 진일보이다. 실제로, 2개의 사전-눈금 매겨진 위트니스 카메라들의 사용은 몇몇의 장점들을 가져온다. 첫번째 것은 인스턴트 조사이다. 평면적인 접근법에서 설명된 기술과 달리, 장면에 대해 사전 조사가 요청되지 않는다. 2개의 위트니스 카메라들 간의 정확한 거리를 완전히 알고 그들이 그들의 내부 젠록(genlock) 하드웨어 성능들을 사용하여 동기화되므로, 우리는 바를 움직이지 않고 즉시 점 클라우드를 획득할 수 있다. 점의 3D 재구성이 입체시각(Stereovision)의 이론을 통해 2개의 위트니스 카메라에 의해 불려와지므로, 문단에서 설명된 모로그래피는 더이상 사용되지 않는다. 두번째 주요점은 수동 스케일이 요구되지 않는다는 것이다. 2개의 비디오 스트림들의 캡쳐 동안, 우리는 등극선 기하에 의해 추정된 3D 점의 투영과 현재의 비디오 내에서 검출된 2D 특징 점 간의 화소 오류의 합을 최소화하기 위해 레벨버그-마쿼트 최소화 방식을 실행한다.

### 2.5.2 자동-등록

[0327] 그러나, 인스턴트 조사를 사용하면, 평면적인 접근법에서 검출되었던 것과 같은 등록 객체를 사용하지 못하는 새로운 쟁점이 발생한다. 그 제한사항의 주요 이유는, 카메라 바가 필름 카메라에 직접 부착되고 등록 객체를 학습하기 위해 카메라 팀원에게 그들의 카메라 주변에서 움직이도록 요청하지 못한다는 사실에 기인한다. 그것은, 우리가 임의의 물리적인 등록 객체가 장면에 추가되도록 요구하지 않는 제 2 자동 등록을 추가한 이유이다. 그러나, 그것은 사용자가 점 클라우드와의 2개의 상이한 상호작용들을 갖도록 요구한다: 무엇이 우리가 바닥 기준으로 사용하고자 하는 점인지를 시스템에게 알려주는 점 선택, 및 CG 콘텐트를 비교적으로 장면에 올바르게 향하게하기 위한 Y (수직) 축 주변에서의 수동 회전.

[0328] 단지 이들 2개의 매우 빠르고 매우 단순한 사용자 상호작용들만을 갖도록 하기 위해, 우리는 모든 다른 남아있는 자유도들을 컴퓨팅해야 하는데, 이는 올바른 스케일링 및 완전히 수직인 Y 축을 갖는 정규직교 기저를 컴퓨팅하는 것을 의미한다. 우리의 기술은 단순히 중력 방향, 즉, Y 축을 컴퓨팅하기 위해 바 내의 내장형 가속도계의 장점을 취한다. 자력계는 N극 방향을 제공하므로, 우리가 정규의 벡터의 외적들을 사용하여 정규직교 기저를 컴퓨팅하도록 허용하는 근사 벡터를 제공한다.

[0329] 스케일은 컴퓨팅하기 훨씬 용이하다. 그것은 완전히 일시적이며 임의의 사용자 상호작용을 요구하지 않는다. 위트니스 카메라들 사이의 거리가 정확하게 알려져 있으므로, 모든 3D 재구성된 점들은 올바른 스케일을 바로 갖게되고, 따라서, 정규의 완벽한 SfM(structure from motion) 알고리즘에서 검출되는 것과 같은 추가적인 객체가 없다.

### 2.5.3 인스턴트 추적

[0330] 추적 처리는 2.1.4.3 섹션의 평면적인 접근법에 대하여 설명된 것과 약간 상이하다. 모든 새로운 점은 3D 특징 점의 스케일 및 심도가 항상 올바르다는 것을 보증하는 등극선 기하를 사용하여 생성된다. 이것은, 우리의 시스

템이 시작되자마자, 그 광속의 정확하고 인스턴트 조사로 인해 이미 추적을 하고 있다는 것을 의미한다.

[0332] 우리는 추적 처리의 최대 안정성을 보장하기 위해 좌측 및 우측 카메라들로부터 보여지는 모든 재투영된 3D 점들의 결합(union)을 컴퓨팅한다. 점 클라우드는 단순히 특징 매칭을 위해 ORB 서술자와 조합된 등극선들의 3D 교차에 의해 생성된다(2.1.4.4 섹션을 참조). 특징 매칭 및 글로벌 최소화 방식은 평면적인 접근법에서 사용되는 것과 유사하다.

[0333] 인스턴트 조사에 의존하기 위해, 우리는 또한 새로운 점들이 점 클라우드 내에 구축되는 방식을 많이 수정한다. 점들의 입체적인 재구성 이외에, 특징 점들의 균일성 및 그리드 분포의 개념을 추가했다. 정규 추적 알고리즘 동안, 2D 측정 검출 알고리즘이 시작되고 점들은 2.1.3.3 및 2.1.4.4 섹션에서 설명된 바와 같이 3D 내에서 재구성된다. 그러나, 많은 이들 점들은 비신뢰적이고(불량한 심도, 가장자리 점들, 움직이는 점들, 어두운 점들 등) 알고리즘은 매우 대조적인 영역들 내에 많은 점을 학습할 경향이 있고, 낮은 대조적인 영역들 내에는 아무 것도 없다. 이것은 추적 데이터 내에 카메라가 흔들리는 느낌을 주는 많은 잡음을 생성한다. 우리가 생성한 그리드 분포 기술은 시스템에 의해 가장 균일한 방식으로 학습된 자연스러운 마커들을 분포시키고자 한다. 우리는 우선 매우 높은 임계를 사용함으로써 최대 특징 점들을 생성하기 시작한다. 그 후 우리는 해리스 점수를 사용하여 (우리는 예를 들면, 2개의 최상의 점들을 선택함) 사용자에 의해 주어진 임계에 따라 특징 점들을 수를 줄인다(예를 들면, 16 x 16 그리드, 셀당 2개의 샘플들). 따라서 모든 점은 그것이 어떻게 신뢰적인지를 알리기 위해 사용되는 품질 기준을 갖는다. 균일성이 충분히 양호하지 않으면, 시스템은 또한 낮은 대조적인 영역들 내에서도 점들을 캐치(catch)할 수 있다는 것을 보장하기 위해 대조적인 임계를 국부적으로 변경시킨다. 우리의 알고리즘이 특징 점들을 사진에 걸쳐 분포시킨다는 사실은 또한 시스템이 쉽게 추적을 잃어버리지 않게 한다. 실제로, 우리의 것과 같은 그리드 분포 기술이 없는 정규의 평면적인 접근법들은 예를 들면, 사진의 보다 아래쪽의 좌측 부분에 추적할 점들을 많이 가질 수 있다. 연기가 모든 이들 자연스러운 마커들을 가리기 시작하면, 추적이 즉시 잃어버려지는데, 이것은 우리는 전체 사진 상에 특징을 분포시켰기 때문에 우리의 경우에는 발생하지 않을 수 있는 것이다.

## 2.6 증강하는 실세계

### 2.6.1 카메라/위트니스 지연 문제

[0335] 필름 카메라 및 위트니스 카메라들은 상이한 속도들에서 동작하므로, 그들은 프레임 획득 처리에서 상이한 지연들을 생성할 수 있다. 필름 카메라는 실제로 실시계를 올바르게 증강시키기 위해 우리가 지연을 컴퓨팅하게 하는 훨씬 최고의 지연을 생성하는 것이다.

[0336] 우리는 2개의 카메라들 사이의 지연을 자동으로 컴퓨팅하기 위해 또다시 우리의 눈금매기기 장치를 사용한다. 1Hz에서 플래싱하는 소형 태양-밝기 led가 두 카메라들 모두 앞에서 사용된다. 세기 임계화 기술에 의해 개선된 블롭 검출기 알고리즘은 두 카메라들 모두에서 led를 검출하도록 개발됐다. 카메라들 내의 led 세기의 캡쳐로부터 생성된 관련 세기 커브들은 샘플들 간의 최고의 인스턴트 간섭성(coherency)을 찾기 위해 최소화 알고리즘을 사용하여 비교된다. 따라서, 몇몇의 고정의 매개변수들이 최소화되고 1000분의 1초 내의 지연이 대략 10초 동안 눈금으로부터 자동으로 추출된다.

### 2.6.2 렌더링 및 증강 현실

[0338] 우리가 카메라 추적 작업을 하면, 우리의 소프트웨어는, 예를 들면, 오토데스크 모션빌더(Autodesk MotionBuilder)와 같은 상이한 CG 콘텐트 생성기 소프트웨어들로부터 데이터를 수신할 수 있다. 그들 데이터는 실제 배경 또는 전경(녹색 스테이지의 경우)과 함께 합성되도록 우리의 소프트웨어 내에서 바로 통합 및 렌더링 된다(추적 및 렌더링 워크스테이션이 병합되는 경우, 도 1을 참조).

[0339] 우리의 기술에 고유한 또 다른 관심있는 특징은 3D 가비지 매팅이다. 가비지 매팅은 분명히 우리의 발명이 아니지만 그것은 혁신적인 방식으로 우리의 경우에서 수행된다. 실제로, 우리의 추적 알고리즘은 정확한 3D 점 클라우드를 생성하기 때문에(입체적인 접근법으로 인해), 모든 이들 점들은 필름 카메라 비디오 상으로 다시 투영될 때 3D 및 2D 모두에서의 표현을 갖는다. 우리는 사용자가 3D 점 클라우드 내에서 다수의 점들을 선택하도록 허용하는 간단한 사용자 인터페이스를 개발했다. 따라서, 우리는 이어 삼각분할(Ear Triangulation) 알고리즘[1 4]을 사용하여 모든 선택된 점들을 삼각분할함으로써 3D 객체를 생성한다. 그 새롭게 생성된 표면은 이제 가비지 매팅 객체로서 활성화될 수 있으며, 이는 필름 카메라 비디오 상으로의 그 객체의 2D 투영에 의해 커버된 모든 화소가 이제 3D 콘텐트 생성 워크스테이션에 의해 생성된 3D 가상 환경에 의해 대체된다(1 섹션을 참조). 긴 거리(즉, 적어도 50m) 심도 센서의 사용의 경우, 가비지 매팅 기술은 쓸모 없는데, 이는 우리의 심도 키잉 방법

이 그것을 용이하게 대체하기 때문이다(2.4 섹션을 참조).

### [0341] 부록 1

#### [0342] 참고 문헌들

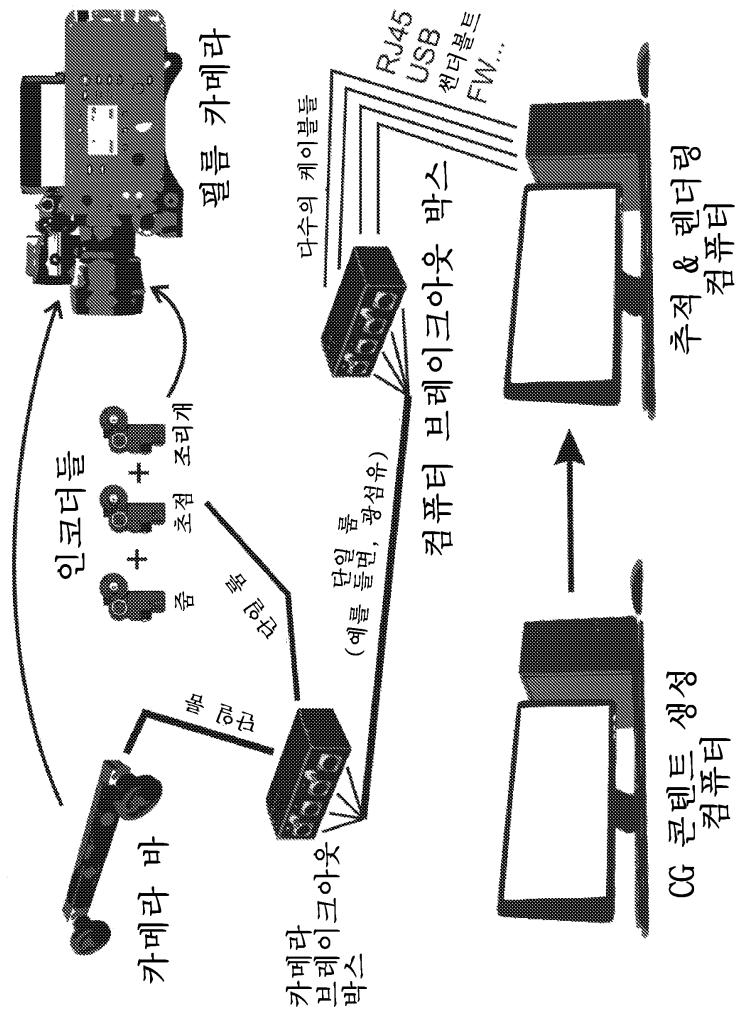
- [0343] [1] 헤르베르트 베이(Herbert Bay), 틴 투이텔라르스(Tinne Tuytelaars), 및 루크 반 쿨(Luc Van Gool). 속도가 높아진 강력한 특징들(Surf: speeded up robust features). 9차 유럽 컴퓨터 비전 협회의 의사록 내 - 볼륨 파트 I, ECCV'06, 404-417쪽, 베를린, 하이델베르그, 2006년. 스프링거 - 출판사.
- [0344] [2] 마이클 칼론더(Michael Calonder), 빈센트 레페티트(Vincent Lepetit), 크리스토프 스트레차(Christoph Strecha), 및 패스칼 푸아(Pascal Fua). 이진의 강력한 독립 기초 특징(Brief: binary robust independent elementary features). 11차 유럽 컴퓨터 비전 협회의 의사록 내 : 파트 IV, ECCV'10, 778-792쪽, 베를린, 하이델베르그, 2010년, 스프링거 - 출판사.
- [0345] [3] 프레데릭 데버네이(Frédéric Devernay) 및 올리비아 포츠하(Olivier Faugeras). 직선들은 곧아야 함(Straight lines have to be straight). SPIE 회, 볼륨 2567, 2001년.
- [0346] [4] 마틴 애이. 피슬러(Martin A. Fischler) 및 로베르트 씨. 볼레스(Robert C. Bolles). 무작위 샘플 합의들: 이미지 분석들 및 자동화된 지도학에 대한 응용들을 갖는 모델 피팅을 위한 패러다임(Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography). ACM의 커뮤니케이션스(Communications of the ACM), 24(6):381-395, 1981년.
- [0347] [5] 크리스 해리스(Chris Harris) 및 마이크 스텝스(Mike Stephens). 조합된 코너 및 가장자리 검출기(A combined corner and edge detector). 4차 엘비 비전 회의 내(In Proc. of Fourth Alvey Vision Conference), 147-151쪽, 1988년.
- [0348] [6] 디. 씨. 호그린(D.C. Hoaglin), 애프. 모스텔러(F. Mosteller), 및 제이.더블유. 터키(J.W. Tukey). 강력하고 탐구적인 데이터 분석 이해하기(Understanding Robust and Exploratory Data Analysis). 가능성 및 수학적 분석들 내의 와일리 시리즈들: 적용된 가능성 및 분석들(Wiley series in probability and mathematical statistics: applied probability and statistics). 존 와일리(John Wiley) 및 선스(Sons), 1983년.
- [0349] [7] 루돌프 에밀 칼만(Rudolph Emil Kalman). 선형 필터링 및 예측 문제들에 대한 새로운 접근법(A new approach to linear filtering and prediction problems). ASME의 트랜잭션스 - 기초 엔지니어링 저널(Transactions of the ASME - Journal of Basic Engineering), 82(시리즈 D):35-45, 1960년.
- [0350] [8] 게오르그 클레인(Georg Klein) 및 데이비드 머레이(David Murray). 소형 ar 작업공간들에 대한 평행 추적 및 매핑(Parallel tracking and mapping for small ar workspaces). 막싱 및 증강된 현실에 대한 2007년 6차 IEEE 및 ACM 국제 심포지움의 의사록 내(In Proceedings of the 2007 6th IEEE 및 ACM International Symposium on Mixed and Augmented Reality), ISMAR '07, 1-10쪽, 워싱턴, DC, 미국, 2007년. IEEE 컴퓨터 학회.
- [0351] [9] 제이. 제이. 레오나르드(J. J. Leonard) 및 두란트 에이치. 와이트(Durrant H. Whyte). 기하학 비콘들 추적에 의한 모바일 로보트 위치지정(Mobile robot localization by tracking geometric beacons). 로보틱스 및 자동화에 대한 IEEE 트랜잭션들(IEEE Transactions on Robotics and Automation), 7(3), 1991년.
- [0352] [10] 케이. 레벤버그(K. Levenberg). 최소 제곱 내의 특정 문제들의 솔루션에 대한 방법(A method for the solution of certain problems in least squares). Quart. Appl. Math., 2:164-168, 1944년.
- [0353] [11] 데이비드 지. 로우(David G. Lowe). 국부적 스케일-불변적인 특징들로부터의 객체 인식(Object recognition from local scale-invariant features). 컴퓨터 비전에 대한 국제 회의의 의사록 내 - 볼륨 2 - 볼륨 2, ICCV '99, 1150쪽부터, 워싱턴, DC, 미국, 1999년. IEEE 컴퓨터 학회.
- [0354] [12] 디. 마콰트(D. Marquardt). 선형 매개변수들의 최소-제곱들 추정에 대한 알고리즘(An algorithm for least-squares estimation of nonlinear parameters). SIAM J. Appl. Math., 11:431-441, 1963년.
- [0355] [13] 크리스토퍼 메이(Christopher Mei), 가베 시블리(Gabe Sibley), 마크 쿠민스(Mark Cummins), 폴 뉴먼(Paul Newman), 및 이안 레이드(Ian Reid). 스테레오를 사용하는 고정-시간 내의 큰-스케일 매핑을 위한 시스템(RsIam: A system for large-scale mapping in constant-time using stereo). Int. J. Comput. 비전,

94(2):198-214, 2011년 9월.

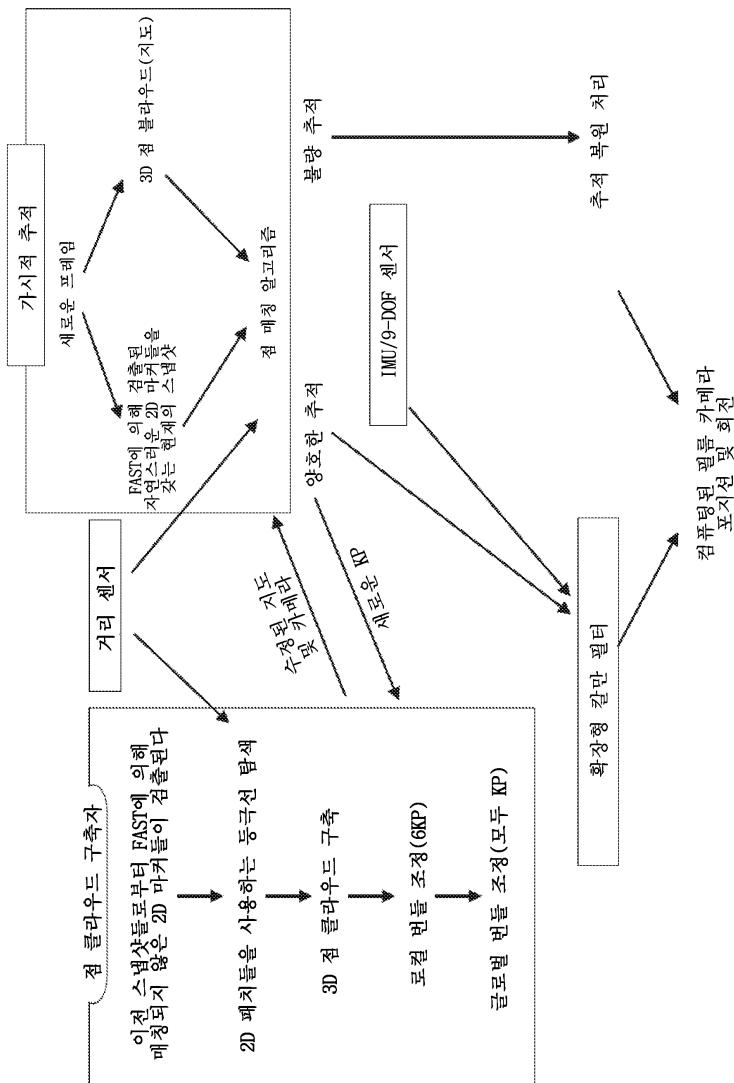
- [0356] [14] 게리 호시어 메이스터스(Gary Hosier Meisters). 다각형들은 귀들을 가짐(Polygons have ears). *Amer. Math. Monthly*, 82:648-651, 1975년.
- [0357] [15] 제이. 제이. 모어'(J. J. More'). 레벤버그-마콰트 알고리즘: 구현 및 이론(The Levenberg-Marquardt algorithm: implementation and theory), 105-116쪽. 수학 630의 강의록. 스프링거, 1977년.
- [0358] [16] 에드워드 로스턴(Edward Rosten) 및 톰 드루몬드(Tom Drummond). 고속 코너 검출에 대한 기계 학습(Machine learning for high-speed corner detection). 컴퓨터 비전에 대한 9차 유럽 회의의 의사록 내 - 볼륨 퍼트 I, ECCV'06, 430-443쪽, 베를린, 하이델베르그, 2006년. 스프링거-출판사.
- [0359] [17] 에드워드 로스턴, 레이드 포터(Reid Porter), 및 톰 드루몬드. 보다 빠르고 보다 양호한: 코너 검출에 대한 기계 학습 접근법(Faster and better: A machine learning approach to corner detection). IEEE 트랜잭션스. 패턴 분석 및 기계 지능, 32:105-119, 2010년.
- [0360] [18] 이든 루블리(Ethan Rublee). 빈센트 라바우드(Vincent Rabaud), 커트 코노리게(Kurt Konolige), 및 게리 브라드스키(Gary Bradski). Orb: sift 또는 surf에 대한 효율적인 대안(Orb: An efficient alternative to sift or surf). 컴퓨터 비전에 대한 2011년 국제 회의의 의사록 내, ICCV '11, 2564-2571쪽, 워싱턴, DC, 미국, 2011년. IEEE 컴퓨터 학회.
- [0361] [19] 이반 이. 서덜랜드(Ivan E. Sutherland). 머리-장착 3차원 디스플레이(A head-mounted three dimensional display). 1968년 12월 9-11일 추계연합컴퓨터회의 의사록 내(In Proceedings of the December 9-11, 1968, fall joint computer conference), 퍼트 I, AFIPS '68 (가을, 퍼트 I), 757-764쪽, 뉴욕, NY, 미국, 1968년. ACM.
- [0362] [20] 로저 와이. 차이(Roger Y. Tsai) 및 레이미르 케이. 렌즈(Reimer K. Lenz). 완벽히 자율적이고 효율적인 3d 로보틱스 손-눈 눈금 매기기에 대한 새로운 기술(A new technique for fully autonomous and efficient 3d robotics hand-eye calibration). 로보틱스 연구에 대한 4차 국제 심포지엄의 의사록 내, 287-297쪽, 캠브리지, MA, 미국, 1988년. MIT 프레스.
- [0363] [21] 제이.케이. 울만(J.K. Ulmann). 복수의 목표 추적에 대한 알고리즘들(Algorithms for multiple target tracking). 아메리칸 사이언티스트(American Scientist), 80(2):128-141, 1992년.

도면

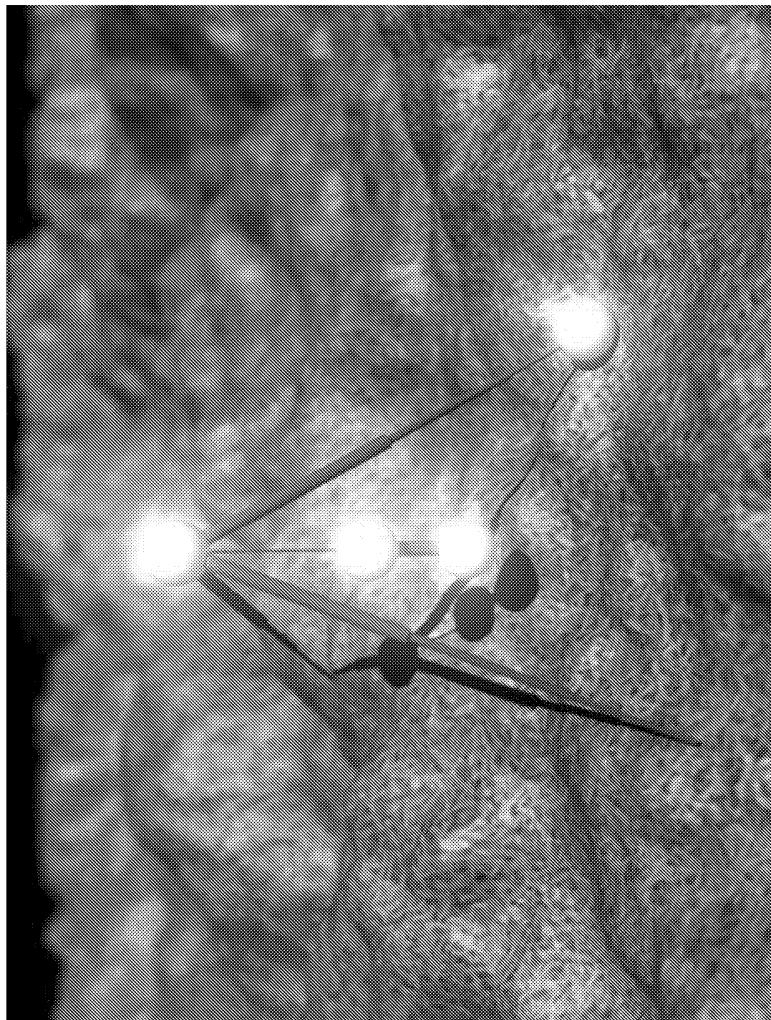
도면1



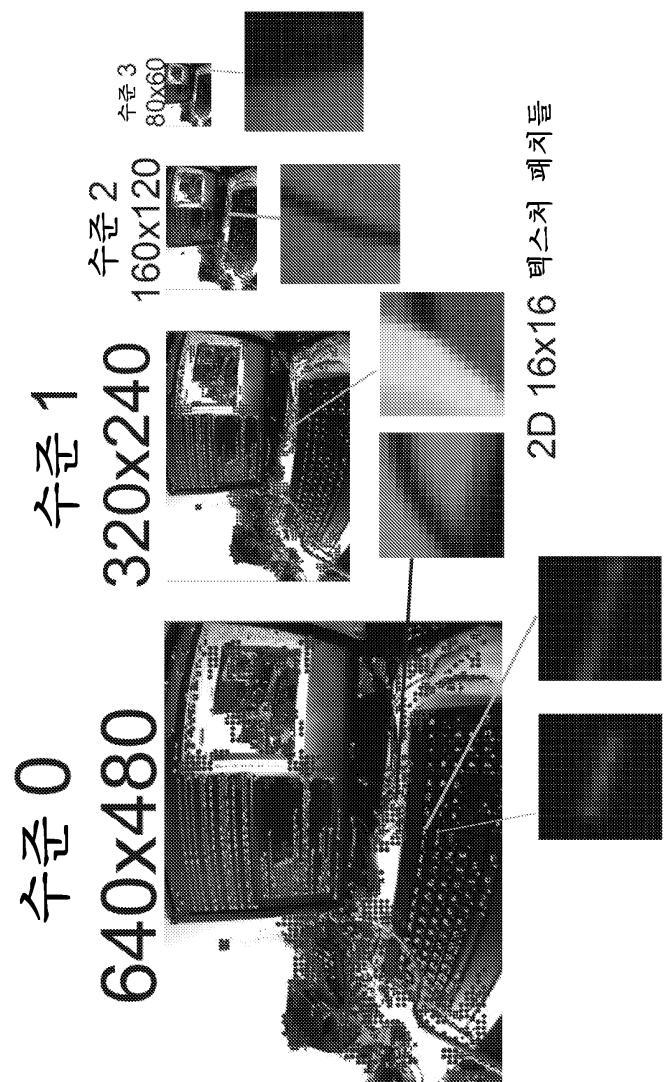
## 도면2



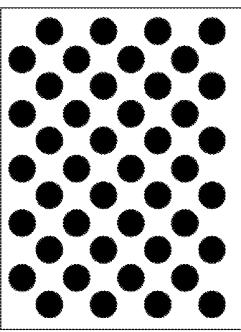
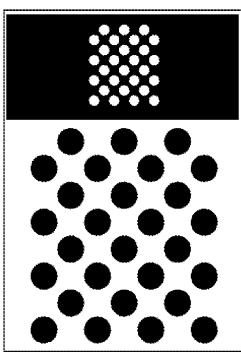
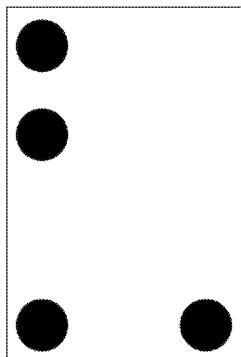
도면3



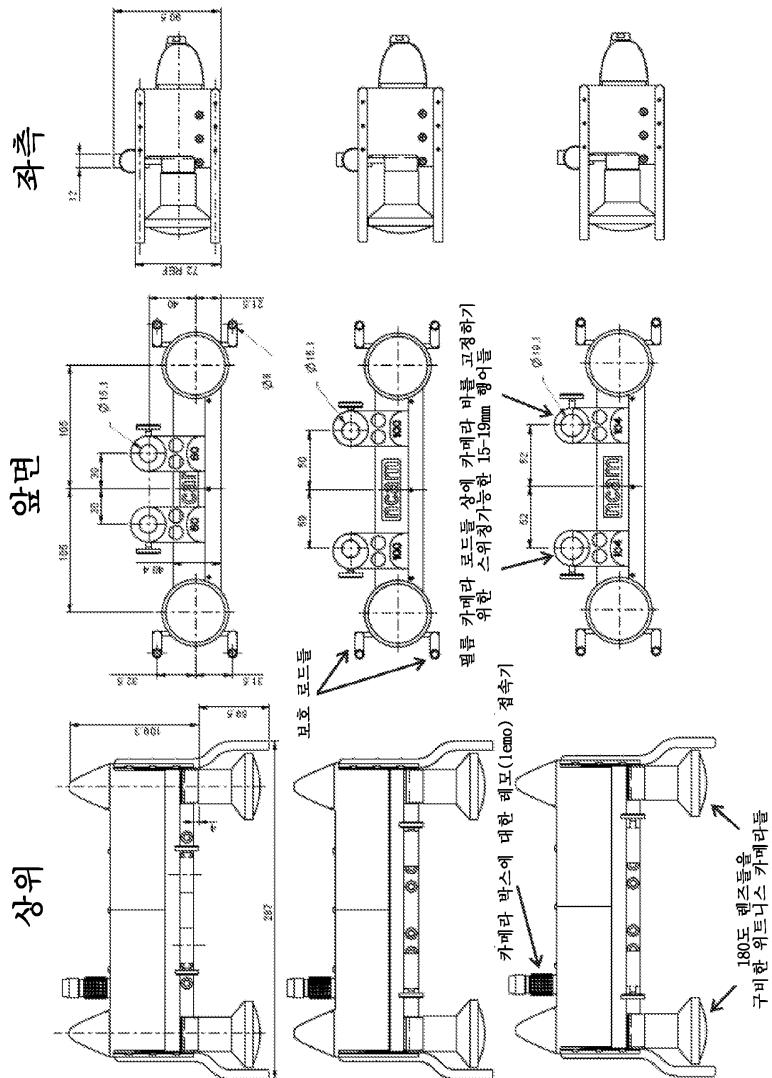
도면4



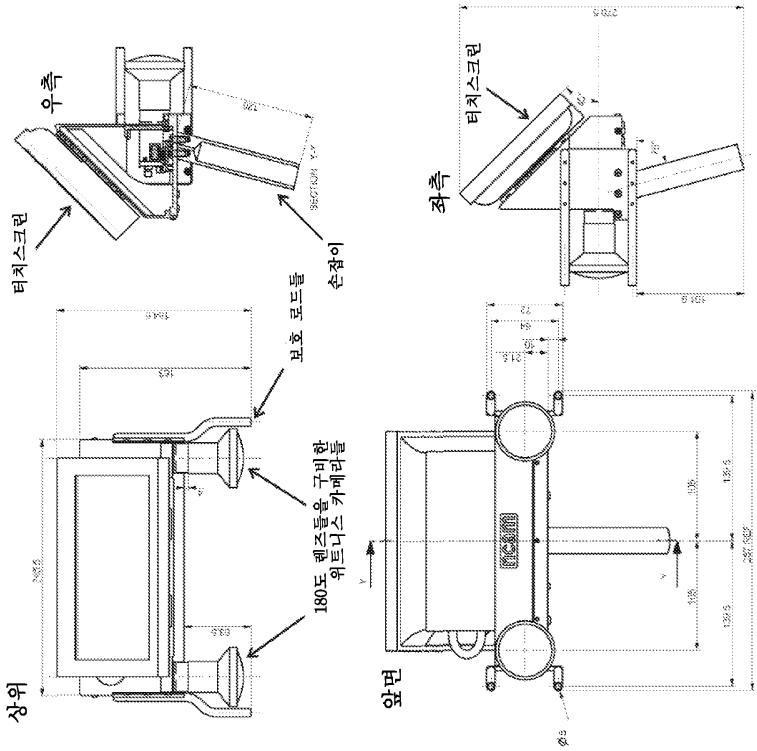
도면5



도면6



## 도면7



## 도면8

