



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2014년08월01일
 (11) 등록번호 10-1424942
 (24) 등록일자 2014년07월23일

(51) 국제특허분류(Int. Cl.)
 G06T 1/00 (2006.01)
 (21) 출원번호 10-2013-7014664 (분할)
 (22) 출원일자(국제) 2005년07월31일
 심사청구일자 2013년06월07일
 (85) 번역문제출일자 2013년06월07일
 (65) 공개번호 10-2013-0086061
 (43) 공개일자 2013년07월30일
 (62) 원출원 특허 10-2013-7000289
 원출원일자(국제) 2005년07월31일
 심사청구일자 2013년01월04일
 (86) 국제출원번호 PCT/IL2005/000813
 (87) 국제공개번호 WO 2006/011153
 국제공개일자 2006년02월02일
 (30) 우선권주장
 60/592,136 2004년07월30일 미국(US)
 (56) 선행기술조사문헌
 JP3960536 B2
 KR1020010095641 A
 KR1020040034606 A
 KR1020040055310 A
 전체 청구항 수 : 총 22 항

(73) 특허권자
 익스트림 리얼리티 엘티디.
 이스라엘, 허즐리야 피투아 46725, 16 아바 이벤
 블러마드
 (72) 발명자
 지본, 도르
 이스라엘, 리손-리지온 75241, 비니크 스트리트
 11/10
 (74) 대리인
 허용록

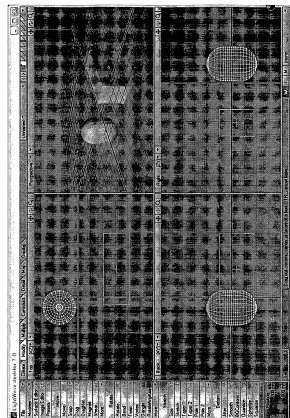
심사관 : 김평수

(54) 발명의 명칭 **이미지 프로세싱을 기반으로 한 3D 공간 차원용 시스템 및 방법**

(57) 요약

이미지 데이터의 3D 표현을 위한 장치로서, 이미지 데이터 내의 움직이는 구조를 식별하는 구조 식별기, 및 3차원 뼈대 모델 엘리먼트와 상기 식별된 구조를 결합시키는 뼈대 삽입 유닛을 포함한다. 상기 뼈대 모델 엘리먼트는 상기 구조와 함께 이동 가능하며 상기 입력 데이터에 프로젝트 될 수 있는 상기 이미지 데이터의 3차원 움직임 및 구조에 대한 이해를 제공한다. 개별 엘리먼트들 뿐만 아니라, 복합 몸체들도 다수의 엘리먼트를 갖는 복합 뼈대에 의해 모델링될 수 있다. 상기 뼈대 모델 엘리먼트들 자체는 복합 오브젝트들을 식별하는데 사용될 수 있다.

대표도 - 도1



특허청구의 범위

청구항 1

각각 이미지 데이터 세트로 구성된 2차원 이미지들의 세트를 포착 또는 수신하도록 구성된 이미지 포착 블록;

적어도 하나의 3차원 뼈대 모델을 저장하는 디지털 메모리로, 상기 3차원 뼈대 모델은 상대적 차원 및 인간 몸체의 구조를 나타내는 뼈대 모델 엘리먼트들의 세트 간의 공간적 제약을 정의하는, 디지털 메모리; 및

이미지 데이터 내에서 사람의 구조를 식별하고 식별된 구조들을 사용하여 3차원 뼈대 모델 엘리먼트들을 결합시키는 이미지 프로세싱 회로를 포함하고,

상기 이미지 프로세싱 회로는, 뼈대 모델 엘리먼트들 간의 공간적 제약이 정의된 모델 및 2차원 이미지 데이터의 세트 내에서 식별된 구조들 간의 공간적 관계가, 상기 2차원 이미지 데이터의 세트에 3차원 뼈대 모델을 맞추고, 신체의 구조들 중 적어도 하나의 3차원 위치의 근사치를 계산하기 위해 사용되는, 비디오 기반 인간 기계 인터페이스(Video Based Human Machine Interface) 시스템.

청구항 2

제1항에 있어서,

상기 이미지 프로세싱 회로는, 상기 이미지 데이터로부터 복수의 상호관련된 구조로 구성되는 복합 몸체를 인식하고, 상기 뼈대 모델 엘리먼트들 간의 3차원 움직임 관계를 정의하기 위해 상기 상호관련된 구조들에 3차원 뼈대 모델 엘리먼트를 삽입하도록 더 구성되는 것을 특징으로 하는 비디오 기반 인간 기계 인터페이스 시스템.

청구항 3

제1항에 있어서,

3차원 움직임 정보를 제공하기 위해, 상기 2차원 이미지의 시리즈를 통한 3차원 위치 변화의 대응하는 구조를 추정하도록 더 구성되는 것을 특징으로 하는 비디오 기반 인간 기계 인터페이스 시스템.

청구항 4

제1항에 있어서,

식별된 구조와 인식된 복합 몸체를 매칭하기 위하여, 기본 형상으로 구성되는 복수의 기결정된 뼈대 모델을 저장하도록 더 구성되는 것을 특징으로 하는 비디오 기반 인간 기계 인터페이스 시스템.

청구항 5

제1항에 있어서,

주어진 뼈대 모델 엘리먼트를 결합시키고, 상호관련된 기본 뼈대 형상을 사용하여 상기 주어진 뼈대 모델 엘리먼트를 상기 이미지로부터 식별된 구조로 변환하도록 더 구성되는 것을 특징으로 하는 비디오 기반 인간 기계 인터페이스 시스템.

청구항 6

제1항에 있어서,

주어진 뼈대 모델 엘리먼트를 결합시키고, 상기 이미지로부터 식별된 구조를 맞추기 위해 상기 주어진 뼈대 모델 엘리먼트를 변형하도록 더 구성되는 것을 특징으로 하는 비디오 기반 인간 기계 인터페이스 시스템.

청구항 7

제1항에 있어서,

주어진 뼈대 모델 엘리먼트를 결합시키고, 상기 주어진 뼈대 모델 엘리먼트에 텍스처를 적용하도록 더 구성되는 것을 특징으로 하는 비디오 기반 인간 기계 인터페이스 시스템.

청구항 8

제1항에 있어서,

제1 레벨에서 각각의 구조를 추적하고, 제2 레벨에서 대응하는 뼈대 모델 엘리먼트를 이동시키도록 더 구성되는 것을 특징으로 하는 비디오 기반 인간 기계 인터페이스 시스템.

청구항 9

제8항에 있어서,

상기 제2 레벨에서 각각의 뼈대 모델 엘리먼트를 추적하고, 상기 제1 레벨에서 구조의 편차를 계산하도록 더 구성되는 것을 특징으로 하는 비디오 기반 인간 기계 인터페이스 시스템.

청구항 10

제1항에 있어서,

3차원 동작 캡처를 제공하기 위해, 상기 뼈대 모델 엘리먼트 상에서 입력 데이터를 캡처하도록 더 구성되는 것을 특징으로 하는 비디오 기반 인간 기계 인터페이스 시스템.

청구항 11

제1항에 있어서,

3차원 이미지 데이터 압축 및 기록을 제공하기 위해, 상기 뼈대 모델 엘리먼트에 대한 데이터를 저장하도록 더 구성되는 것을 특징으로 하는 비디오 기반 인간 기계 인터페이스 시스템.

청구항 12

적어도 하나의 3차원 뼈대 모델을 저장하는 단계로, 상기 3차원 뼈대 모델은 상대적 차원 및 인간 몸체의 구조를 나타내는 뼈대 모델 엘리먼트들의 세트 간의 공간적 제약을 정의하는, 단계;

각각 이미지 데이터 세트로 구성된 2차원 이미지들의 세트를 포착 또는 수신하는 단계;

이미지 데이터 내에서 사람의 구조를 식별하는 단계; 및

식별된 구조들을 사용하여 3차원 뼈대 모델 엘리먼트들을 결합시키는 단계를 포함하고,

상기 결합시키는 단계는, 뼈대 모델 엘리먼트들 간의 공간적 제약이 정의된 모델 및 2차원 이미지 데이터의 세트 내에서 식별된 구조들 간의 공간적 관계가, 상기 2차원 이미지 데이터의 세트에 3차원 뼈대 모델을 맞추고, 신체의 구조들 중 적어도 하나의 3차원 위치의 근사치를 계산하기 위해 사용되는 것을 특징으로 하는 비디오 기반 인간 기계 인터페이스(Video Based Human Machine Interface) 방법.

청구항 13

제12항에 있어서,

상기 이미지 데이터로부터 복수의 상호관련된 구조로 구성되는 복합 몸체를 인식하고, 상기 뼈대 모델 엘리먼트들 간의 3차원 움직임 관계를 정의하기 위해 상기 상호관련된 구조들에 3차원 뼈대 모델 엘리먼트를 삽입하는 단계를 더 포함하는 것을 특징으로 하는 비디오 기반 인간 기계 인터페이스 방법.

청구항 14

제12항에 있어서,

3차원 움직임 정보를 제공하기 위해, 상기 2차원 이미지의 시리즈를 통한 3차원 위치 변화의 대응하는 구조를 추정하는 단계를 더 포함하는 것을 특징으로 하는 비디오 기반 인간 기계 인터페이스 방법.

청구항 15

제12항에 있어서,

식별된 구조와 인식된 복합 몸체를 매칭하기 위하여, 기본 형상으로 구성되는 복수의 기결정된 뼈대 모델을 저장하는 단계를 더 포함하는 것을 특징으로 하는 비디오 기반 인간 기계 인터페이스 방법.

청구항 16

제12항에 있어서,

주어진 뼈대 모델 엘리먼트를 결합하고, 상호관련된 기본 뼈대 형상을 사용하여 상기 주어진 뼈대 모델 엘리먼트를 상기 이미지로부터 식별된 구조로 변환하는 단계를 더 포함하는 것을 특징으로 하는 비디오 기반 인간 기계 인터페이스 방법.

청구항 17

제12항에 있어서,

주어진 뼈대 모델 엘리먼트를 결합하고, 상기 이미지로부터 식별된 구조를 맞추기 위해 상기 주어진 뼈대 모델 엘리먼트를 변형하는 단계를 더 포함하는 것을 특징으로 하는 비디오 기반 인간 기계 인터페이스 방법.

청구항 18

제12항에 있어서,

주어진 뼈대 모델 엘리먼트를 결합하고, 상기 주어진 뼈대 모델 엘리먼트에 텍스처를 적용하는 단계를 더 포함하는 것을 특징으로 하는 비디오 기반 인간 기계 인터페이스 방법.

청구항 19

제12항에 있어서,

제1 레벨에서 각각의 구조를 추적하고, 제2 레벨에서 대응하는 뼈대 모델 엘리먼트를 이동하는 단계를 더 포함하는 것을 특징으로 하는 비디오 기반 인간 기계 인터페이스 방법.

청구항 20

제19항에 있어서,

상기 제2 레벨에서 각각의 뼈대 모델 엘리먼트를 추적하고, 상기 제1 레벨에서 구조의 편차를 계산하는 단계를 더 포함하는 것을 특징으로 하는 비디오 기반 인간 기계 인터페이스 방법.

청구항 21

제12항에 있어서,

3차원 동작 캡처를 제공하기 위해, 상기 뼈대 모델 엘리먼트 상에서 입력 데이터를 캡처하는 단계를 더 포함하는 것을 특징으로 하는 비디오 기반 인간 기계 인터페이스 방법.

청구항 22

제12항에 있어서,

3차원 이미지 데이터 압축 및 기록을 제공하기 위해, 상기 뼈대 모델 엘리먼트에 대한 데이터를 저장하는 단계를 더 포함되는 것을 특징으로 하는 비디오 기반 인간 기계 인터페이스 방법.

명세서

기술분야

[0001] 본 발명은 사진, 이미지 프로세싱 및 애니메이션에 관한 것으로서, 보다 구체적으로는 3차원(3D) 사진, 3차원 이미지 프로세싱 및 3차원 애니메이션에 관한 것이나, 이에 한정되지는 않는다.

[0002] 3차원 사진에서의 본 기술은 시간 차원을 기반으로 한다.

[0003] 본 발명은 3D 형상 및 이미지 프로세싱 분야, 예컨대 입체 이미지, 구면 사진 시스템, 3D 컴퓨터 애니메이션,

3D 사진, 및 3D 이미지 프로세싱 알고리즘과 같은 분야에 해당하는 수개의 서로 다른 영역에 관한 것이다.

배경 기술

- [0004] 통상적인 3D 입체 사진 촬영은 평행한 광축 및 정렬된 렌즈들 사이의 고정된 거리를 갖는 트윈 카메라를 사용한다. 이러한 트윈 카메라들은 입체 디스플레이 및 시청을 위한 공지된 기술 중 어느 하나에 의해 디스플레이될 수 있는 한쌍의 이미지를 생성한다. 이러한 기술들은 일반적으로 오른쪽 렌즈에 의해 촬상된 이미지는 관찰자의 오른쪽 눈에 디스플레이되고, 왼쪽 렌즈에 의해 촬상된 이미지는 관찰자의 왼쪽 눈에 디스플레이된다는 원리를 기반으로 한다.
- [0005] 예를 들어, 미국등록특허 6,906,687 (출원인 텍사스 인스트루먼트, 발명의 명칭 "3D 디스플레이용 디지털 포맷터")는 오른쪽 눈 및 왼쪽 눈 디스플레이 모두에 대해 프로세싱된 비디오 데이터를 저장하고 판독하기 위해 4중 메모리 버퍼를 사용하는 3D 프로젝션 디스플레이를 개시하고 있다. 이 포맷터를 사용하여, 비디오 데이터는 48 frame/sec의 속도로 프로세싱되고, 2번 판독(반복)되어 96(최고 120) frame/sec의 플래시 레이트를 제공하며, 이것은 디스플레이 플리커 임계치를 넘는 수치이다. 그리고 나서, 상기 데이터는 헤드셋 또는 안경과 동기화되며, 오른쪽 눈 및 왼쪽 눈 프레임은 정밀하게 다른 위상을 갖게 되어 3D로 인식되는 이미지를 생성하게 된다.
- [0006] 구면 또는 파노라마 촬영은 통상적으로 "어안 렌즈(fish lens)"와 같은 매우 광각인 렌즈를 사용하거나, 오버랩되는 근접 이미지들을 스티치(stitch)하여 완전 구면의 시야각까지 광범위한 시야를 커버하도록 함으로써 수행된다. 이러한 기술들을 사용하여 촬상된 파노라마 또는 구형(spherical) 이미지들은 관찰자에게 깊이를 인식할 수 있도록 해주는 2차원 이미지 또는 입체 이미지가 될 수 있다.
- [0007] 예를 들어, 미국등록특허 6,833,843 (출원인 템피스트 마이크로시스템스)는 어안 렌즈 및 CCD와 같은 이미지 포착 장치를 사용하여 광각 이미지, 예컨대 반구형 시야각을 갖는 이미지를 촬상할 수 있는 이미지 포착 및 뷰잉 시스템을 개시하고 있다.
- [0008] 본 출원인의 또다른 출원인 미국특허 출원번호 10/416,533 (출원일 2001년 11월 28일) 또한, 본 명세서에 참조 문헌으로서 일체화된다. 상기 출원은 수평 및 수직으로 모두 360도로 관찰자를 둘러싸는 주변에 대한 완전 입체 구형 이미지들을 촬상할 수 있는 이미징 시스템에 대해 기술하고 있다. 입체 디스플레이에 적절한 수단에 의해 이미지를 디스플레이함으로써, 디스플레이되는 이미지의 입체 깊이를 인지할 수 있도록 하면서, 관찰자가 상하 방향뿐만 아니라 관찰자의 주변 전체를 볼 수 있도록 해준다. 상기 문헌은 카메라 어레이에 관해 기술하고 있으며 여기서 카메라의 렌즈들은 곡면에 배치되는데, 상기 곡면의 C개의 공통 중심으로부터 다른 방향을 향하도록 배치된다. 촬상된 이미지들은 배열 및 프로세싱되어 입체 이미지 쌍들의 집합을 생성하는데, 여기서 각 쌍의 이미지 중 하나는 관찰자의 오른쪽 눈에 지정되고, 나머지 하나는 왼쪽 눈으로 지정되어, 3차원 인식을 가능케 한다.
- [0009] 3D 컴퓨터 애니메이션은 "가상 현실" 분야에 관한 것인데, 이것은 최근에 관심을 모으고 있는 분야이다. 3D 가상 현실은 실제 이미지로부터 구성되며, 실제 이미지로부터 얻은 이미지들이 합성될 수 있다. 완전히 컴퓨터로 생성한 가상 현실도 존재한다. 3D 가상 현실은 촬상되는 이미지들의 3D 정보를 생성하기 위해 사진 촬영된 이미지의 3D 계산을 필요로 한다.
- [0010] 이것은 능동 방법들을 사용하여 실시간으로 수행될 수 있다.
- [0011] 예를 들어, 3DV 시스템즈 인코포레이티드 (<http://www.3dvsystems.com/>)은 컬러값 이외에 장면(scene)의 각 픽셀의 깊이 값을 실시간으로 캡처하는 ZCam™을 제공하며, 거리들에 대해 그레이 레벨을 스케일링함으로써 이미지의 모든 프레임에 대해 깊이 맵을 생성한다. ZCam™ 카메라는 적절한 폭을 갖는 라이트 월(light wall)을 사용하는 독특하게 설계된 카메라이다. 라이트 월(light wall)은 예컨대 정사각형 레이저 펄스로서 생성될 수 있다. 라이트 월이 촬상된 장면의 오브젝트들에 닿으면, ZCam™ 카메라 쪽으로 반사되어 오브젝트들의 임프린트(imprint)를 전달하게 된다. 임프린트들은 깊이 맵의 재구성에 필요한 모든 정보를 전달하게 된다.
- [0012] 사진촬영된 이미지들의 3D 계산은 수동 방법들을 사용하여 제공될 수도 있다.
- [0013] 깊이 구성을 위한 수동 방법들은 적어도 2개의 공지된 장면 지점을 이용하는 삼각 기술을 사용할 수 있다. 해당 피쳐들(features)이 식별되고, 광이 교차되어 각 피쳐의 3D 포지션을 찾아낸다. 공간-시간 스테레오는 일시 차원을 공간 매칭 기능에 사용되는 이웃들에 추가한다. 시간에 따라 다중 프레임을 사용하여 일시 스테레오를 추가함으로써, 제2 이미지에 대해 제1 이미지로부터의 단일 픽셀을 매칭시킬 수 있다. 이것은 관심 포인트들

(코너들)에 대조적으로, 통상적인 피쳐 기반의 이미지대 이미지 매칭에서 수행되는 바와 같이, 공간-시간 궤적을 매칭시킴으로써 수행될 수 있다. 이 시퀀스들은 대응되는 공간-시간 궤적을 따라 모든 포인트들의 일정한 매칭을 강화하고, 2개의 비디오 시퀀스들 사이의 서브 프레임 순간 일치(동기화)를 획득함으로써 공간 및 시스템 매칭된다.

- [0014] 3D 컴퓨터로 생성된 이미지들(CGI:3D computer generated images)은 가상 세계로서, 지정된 영역이며, 3D 생성 이미지 소프트웨어를 사용하여 생성된다. 가상 세계는 가상 세계의 모든 포인트가 컴퓨터로 생성된 포인트인 지정 영역에서 생성된다. 2D 또는 3D 실제 이미지들 또한 이 가상 세계에서 결합될 수 있다.
- [0015] 도1은 종래 기술에 따른 가상 세계를 나타낸다.
- [0016] 도1을 참조하면, 도1의 가상 세계에서 모든 포인트의 도3D 위치는 알려져 있다. 킬러, 휘도 등의 공간 디테일에서 임의의 포인트들을 부가함으로써 공간에서 형태를 생성한다(도1). 가상 카메라를 이 가상 세계에 유도함으로써 가상 세계에서의 시퀀스를 기반으로 한 시간, 스테레오 이미지 등을 생성할 수 있다.
- [0017] 공간 동기화 후에 시간 동기화를 사용하고, 공간 및 시간적으로 실제 세상 이미지를 가상 세계에 맞춤으로써, 사진 촬영된 이미지들과 컴퓨터로 생성된 세상을 동기화할 수 있다.
- [0018] 도2는 종래 기술에 따른 가상 스튜디오를 나타낸다.
- [0019] 도2a에 도시된 예에서, 블루/그린 스크린으로서 알려진 기술로 카메라가 사람의 형성과 그 배경을 분리할 수 있는 가상 스튜디오를 사용한다. 사람의 형상과 그 배경을 분리함으로써, 도2b에 도시된 바와 같이 컴퓨터에서 생성한 가상 세계에 상기 형상을 결합시킬 수 있다.
- [0020] 농구장과 같은 미리 설정된 공간에서 카메라들을 모니터링함으로써 이와 반대 작업도 수행될 수 있으며, 상기 미리 설정된 공간에서는 이미 알고 있는 고정 포인트들이 미리 설정되고 있고, 동기화된 고정 포인트들이 컴퓨터로 생성된 3D 세계에서 생성된다. 그러나 기술을 사용하여, CGI 형상을 분리하고, 그것을 농구장에 결합할 수 있다. 예를 들어, ORAD 인코포레이티드의 CyberSport™ 제품은 농구장, 축구장 등에서 발생하는 스포츠 이벤트들에 대한 경기장에 고정된 3D 그래픽의 실시간 삽입을 제공하여, 삽입된 그래픽 오브젝트들이 그 이벤트의 실제 부분인 것과 같은 장면을 생성한다.
- [0021] 상술한 바와 같이, 종래의 3D 이미징 및 입체 사진촬영용 방법 및 시스템은 특별한 카메라들, 특별한 렌즈들, 2개 이상의 카메라의 미리 설정된 포지셔닝 및 전용 알고리즘을 기반으로 한다.

발명의 내용

해결하려는 과제

- [0022] 따라서, 사진촬영 및 이미징용 시스템 및 방법이 요구되고, 그러한 시스템 방법을 사용하는 것은 많은 장점을 가져올 수 있을 것이다.

과제의 해결 수단

- [0023] 본 발명의 일 측면에 따라, 이미지 데이터를 3D로 표시하는 장치가 제공되며, 상기 장치는,
- [0024] 상기 이미지 데이터 내에 움직이는 구조들을 식별하는 구조 식별기; 및
- [0025] 상기 구조 식별기에 결합되어, 상기 뼈대 모델 엘리먼트가 상기 구조들과 함께 움직여서 상기 이미지 데이터의 3차원 움직임 및 구조 이해를 제공할 수 있도록 3차원 뼈대 모델 엘리먼트와 상기 구조들을 결합하는 뼈대 삽입 유닛을 포함한다.
- [0026] 본 발명의 제2 측면에 따라, 이미지 데이터를 3D로 표시하는 방법이 제공되며, 상기 방법은,
- [0027] 상기 이미지데이터 내의 구조들을 식별하는 단계; 및
- [0028] 상기 뼈대 모델 엘리먼트들이 상기 구조들과 함께 움직여서 상기 구조들의 3차원 이해를 제공할 수 있도록 3차원 뼈대 모델 엘리먼트와 상기 구조들을 결합하는 단계를 포함한다.
- [0029] 본 발명의 제3 측면에 따라, 깊이 정보와 함께 입력 데이터를 기록하는 기록 장치가 제공되며, 상기 기록 장치는,

- [0030] 상기 이미지 데이터 내의 움직이는 구조들을 식별하는 구조 식별기;
- [0031] 상기 구조 식별기와 결합되고, 상기 뼈대 모델 엘리먼트들이 상기 구조들과 함께 움직여서 상기 이미지 데이터의 3차원 움직임과 구조 이해를 제공할 수 있도록 상기 3차원 뼈대 모델 엘리먼트와 상기 구조들을 결합시키는 뼈대 삽입 유닛; 및
- [0032] 상기 뼈대 모델 엘리먼트들 및 배경 중 적어도 하나에 상대적으로 상기 입력 데이터를 기록하는 저장 유닛을 포함한다.
- [0033] 본 발명의 제4 측면에 따라, 깊이 정보와 함께 입력 데이터를 압축하는 압축 장치가 제공되며, 상기 장치는,
- [0034] 상기 이미지 데이터 내의 움직이는 구조들을 식별하는 구조 식별기;
- [0035] 상기 구조 식별기와 결합되고, 상기 뼈대 모델 엘리먼트들이 상기 구조들과 함께 움직여서 상기 이미지 데이터의 3차원 움직임 및 구조 이해를 제공할 수 있도록 상기 3차원 뼈대 모델 엘리먼트와 상기 구조들을 결합시키는 뼈대 삽입 유닛; 및
- [0036] 상기 입력 데이터의 압축을 제공하고, 깊이 정보를 제공할 수 있도록 상기 뼈대 모델 엘리먼트들 및 배경 적어도 하나에 상대적으로 상기 입력 데이터를 출력하는 압축 유닛을 포함한다.
- [0037] 본 발명의 제5 측면에 따라, 깊이 정보와 함께 입력 데이터를 기록하는 기록 방법이 제공되며, 상기 방법은,
- [0038] 상기 이미지 데이터 내의 움직이는 구조들을 식별하는 단계;
- [0039] 3차원 뼈대 모델 엘리먼트들을 상기 구조들과 결합하여, 상기 뼈대 모델 엘리먼트들이 상기 구조들과 함께 움직여서 상기 이미지 데이터의 3차원 움직임 및 구조 이해를 제공할 수 있도록 하는 단계; 및
- [0040] 상기 뼈대 모델 엘리먼트 및 배경 중 적어도 하나에 상대적으로 상기 입력 데이터를 기록하는 단계를 포함한다.
- [0041] 본 발명의 제6 측면에 따라, 깊이 정보와 함께 입력 데이터를 압축하는 압축 방법이 제공되며, 상기 방법은,
- [0042] 상기 이미지 데이터 내의 움직이는 구조들을 식별하는 단계;
- [0043] 3차원 뼈대 모델 엘리먼트와 상기 구조들을 결합하여, 상기 뼈대 모델 엘리먼트들이 상기 구조들과 함께 움직여서 상기 이미지 데이터의 3차원 움직임 및 구조 이해를 제공할 수 있도록 하는 단계; 및
- [0044] 상기 입력 데이터의 압축을 제공하고, 그 깊이 정보를 제공할 수 있도록 상기 뼈대 모델 엘리먼트 및 배경 중 적어도 하나에 상대적으로 상기 입력 데이터를 출력하는 단계를 포함한다.
- [0045] 달리 설명되지 않는 한, 본 명세서의 모든 기술 용어 및 과학 용어들은 본 발명이 속하는 기술 분야의 당업자들에게 공통적으로 이해되는 바와 동일한 의미를 갖는다. 본 명세서의 재료, 방법 및 실시예들은 설명을 위한 것이며, 본 발명의 범위를 제한하고자 하는 것이 아니다.
- [0046] 본 발명의 방법 및 시스템의 실시는 수동 또는 자동적으로 선택된 임의의 작업 또는 방법들 또는 그 결합을 실행 또는 완료하는 단계를 포함한다. 또한, 본 발명의 방법 및 시스템의 바람직한 실시예의 실제 도구 및 장비에 따르면, 수개의 선택된 단계들은 하드웨어 또는 임의의 펌웨어 또는 그것이 조합된 시스템 상에서 동작하는 소프트웨어에 의해 구현될 수 있을 것이다. 예를 들어, 하드웨어로서, 본 발명의 선택된 단계들은 칩 또는 회로로써 구현될 수 있을 것이다. 소프트웨어로서는, 본 발명의 선택된 단계들은 적절한 운영체제를 사용하는 컴퓨터에 의해 실행되는 복수의 소프트웨어 명령어로서 구현될 수 있을 것이다. 임의의 경우에, 본 발명의 방법 및 시스템의 선택된 단계들은 복수의 명령어를 실행하기 위한 컴퓨팅 플랫폼과 같은 데이터 프로세서에 의해 수행되는 것으로 설명될 수 있을 것이다.
- [0047] 이하, 첨부된 도면을 참조하여 본 발명의 실시예들을 설명한다. 이하의 실시예들에서, 특정 구성 요소들은 본 발명의 바람직한 실시예들에 대한 설명을 위한 것일 뿐이며, 본 발명의 원리와 개념에 대한 가장 유용하고 이해하기 쉽다고 생각되는 것을 제공하기 위해 제시된다. 이 점에 있어서, 본 발명의 가장 핵심적인 이해에 필요한 것 이상으로 상세한 구조적인 설명은 하지 않기로 하며, 당업자가 명확히 본 발명이 실제로 어떻게 구현되는가를 이해할 수 있도록 도면과 함께 설명한다.

도면의 간단한 설명

- [0048] 도 1은 종래 기술의 3D 컴퓨터가 생성한 가상 영상의 사진이다.

- 도 2a는 종래 기술 가상 스튜디오의 제1 사진이다.
- 도 2b는 종래 기술 가상 스튜디오의 제2 사진이다.
- 도 3은 본 발명의 바람직한 제1 실시예에 따른 간략화된 3D 이미지 분석용 블록도이다.
- 도 4는 본 발명의 바람직한 실시예에 따라 뼈대 모델 엘리먼트를 이미지 또는 일련의 이미지로부터 인지된 구조적 엘리먼트에 삽입하기 위한 과정을 설명하는 간략화된 흐름도이다.
- 도 5는 단일체로 인식되는 일련의 엘리먼트의 경우에 대한 도 4a의 과정의 변형을 나타내는 간략화된 흐름도이다.
- 도 6은 본 발명의 바람직한 실시예에 따라 2d 이미지 데이터의 3차원 이해를 제공하는데 있어서 뼈대 삽입 및 그 이후의 사용을 나타내는 간략화된 흐름도이다.
- 도 7은 2D 이미지의 3D 이해를 제공하기 위해 도 4a 내지 도 4c의 뼈대의 사용을 다층 포맷 구조로 나타내는 간략화된 다이어그램이다.
- 도 8은 본 발명의 바람직한 실시예에 따라 뼈대 모델 엘리먼트를 삽입할 2 차원 이미지로부터 오브젝트 인식을 획득하는 2가지 방법을 나타내는 흐름도이다.
- 도 9는 이미지의 2D 움직임 엘리먼트의 3D 이해를 제공하기 위해, 본 발명의 실시예에 따른 뼈대를 사용하는 과정을 나타내는 간략화된 흐름도이다.
- 도 10은 본 발명의 실시예에의 일련의 예시적인 응용예를 나타내는 풍선 차트이다.
- 도 11은 본 발명의 바람직한 실시예에 따른 가능한 이미징 프로세스를 나타내는 깊이맵이다.
- 도 12는 본 발명의 바람직한 실시예에 따른 가능한 이미징 프로세스를 나타내는, 깊이맵에 첨부된 뼈대이다.
- 도 13은 본 발명의 바람직한 실시예에 따른 가능한 이미징 프로세스를 나타내는 구조의 변형 과정을 설명하는 뼈대이다.
- 도 14는 일련의 이미지의 구조 엘리먼트가 최소 왜곡 위치가 존재하는 프레임으로부터 프로세싱되는 과정을 나타낸다.
- 도 15는 특정 방향으로부터 3D 정보를 제공하는 촬영된 이미지를 나타낸다.

발명을 실시하기 위한 구체적인 내용

- [0049] 본 발명은 촬영된 이미지들의 시간 기반 시퀀스들을 공간 기반 3차원 모델들로 변환하는 방법 및 장치를 포함하며, 3D 실제 이미지 애니메이션, 새로운 시간 기반 시퀀스, 이미지 프로세싱 조작, 2D/3D 모션 캡처 등과 같은 실시간 및 비 실시간 어플리케이션을 가능케 한다.
- [0050] 본 실시예들은 2차원 또는 부분적으로 3차원 데이터 내의 구조들을 식별하며, 그것들과 3차원 뼈대 또는 뼈대 모델 엘리먼트들을 결합한다. 뼈대 또는 뼈대 모델 엘리먼트들은 원 데이터로부터 분리된 레벨에서 적용될 수 있으며, 서로에게 프로젝트될 레벨로 하여금 이미지 데이터에 대한 정확한 깊이 정보를 제공할 수 있도록 한다.
- [0051] 본 발명에 따른 방법 및 장치의 원리 및 동작은 첨부 도면 및 이하의 설명을 참조하여 이해될 수 있을 것이다.
- [0052] 본 발명의 적어도 하나의 실시예를 상세히 설명하기 전에, 본 발명은 이하의 설명 또는 도면에 도시된 컴포넌트들의 구조 및 배열로 한정되지 않는다는 점이 이해되어야 한다. 본 발명은 다른 실시예들을 포함할 수 있으며, 다양한 방식으로 실시 또는 구현될 수 있다. 또한, 본 명세서에서 사용되는 어구 또는 용어는 설명을 위한 것이며 제한적인 것으로 해석되지 않는다.
- [0053] 도 3을 참조하면, 이미지 데이터에 대한 3차원 이해를 제공하는 장치를 나타내는 간략화된 블록도가 도시되어 있다. 이미지 데이터는 2차원 또는 부분적으로 3차원 정보일 수 있으며, 3차원 구조 및 3차원 움직임에 대해 통일적으로 이해될 것이다.
- [0054] 도 3의 장치는 이미지 데이터 내의 구조들을 식별하기 위한 구조 식별기(302)를 포함한다. 이하에서 설명할 바와 같이, 구조들은 인공 지능을 사용하여 자동적으로 식별되거나, 사용자 입력의 도움으로 또는 이것들의 조합에 의해 식별될 수 있다.

- [0055] 상기 장치는 고정(rigid) 구조 식별기와 결합된 뼈대 삽입 유닛(304)을 더 포함하며, 상기 뼈대 삽입 유닛(304)은 3차원 뼈대 모델 엘리먼트를 상기 이미지 데이터에서 식별된 구조들과 결합시킨다. 뼈대 모델 엘리먼트들은 블록, 튜브, 구형, 타원형, 또는 다른 엘리먼트 또는 보다 복잡한 3차원 기하학 구조일 수 있으며, 상기 엘리먼트들은 서로에게 결합될 수 있다. 상기 엘리먼트의 3차원 형태는 위와 같이 식별된 구조로 첨가되고, 상기 뼈대 모델 엘리먼트는 이동하거나, 상기 구조들과 같이 공존하여 구조들의 3차원 이해를 제공할 수 있다.
- [0056] 즉, 상기 뼈대 모델 엘리먼트는 공지된 3차원 구조를 가지며, 이것은 X,Y,Z 크기로 확장됨을 의미한다. 상기 구조의 이동은 X 및 Y 방향으로 관측될 수 있으며, Z방향으로의 구조의 이동은 뼈대 모델 엘리먼트와의 관계로부터 유추될 수 있다.
- [0057] 구조 식별기는 바람직하게는 개별 구조뿐만 아니라, 상호관련된 구조들, 그들 사이의 이동 관계를 정의한 상호관련된 의미로 구성된 복합 몸체(complex body)도 인식할 수 있다. 한 예로써, 팔뚝 및 어깨와 같은 구조로 구성되는 인체를 들 수 있다. 팔뚝은 정의된 방식으로 어깨의 끝단에서 피벗(pivot)하고, 이것은 본 실시예의 뼈대 모델 엘리먼트에 의해 모델링될 수 있다.
- [0058] 그러한 복합 몸체가 이미지 데이터로부터 인식되는 경우에, 상기 뼈대 삽입 유닛은 뼈대 모델 엘리먼트가 복합 몸체에 대한 것으로 정의되는 해당 복합 뼈대의 구성을 시도한다. 이하의 설명과 같이, 이것을 달성하는 한 방법은 복합 몸체, 예컨대 인간을 인식하고, 필요한 엘리먼트 그 내부에 미리 프로그램된 관계들로 뼈대들을 미리 설정하는 것이다.
- [0059] 그러한 미리 저장되거나 미리 설정된 뼈대를 사용하여, 구조 및 움직임 모두를 포함하는 복합 몸체의 3차원 측면을 이해할 수 있다. 즉, 이미지 내의 3차원 구조 및 움직임은 인식된 몸체에 대한 선형적인 지식으로부터 이해될 수 있다. 또한, 상기 뼈대를 기반으로 시스템 내에서 오브젝트에 대한 깊이 정보를 알 수 있는 경우에, 이미지에 대한 3차원 프로세싱의 프로세싱 로드는 상당히 감소될 수 있다.
- [0060] 상기 장치는 원 이미지 데이터 내의 상대적인 움직임을 분석하여 뼈대 삽입 유닛(304)에 움직임 관계 정의를 제공할 수 있는 움직임 분석기 유닛(306)을 더 포함할 수 있다. 상기 움직임 분석기는 공지된 추적 기술들(tracking techniques)을 사용하여 이미지를 구성하는 다수의 픽셀 내의 구조들을 인식하고, 픽셀 그룹들 내의 움직임을 식별할 수 있다.
- [0061] 뼈대 저장부(308)는 인식된 복합 몸체를 사용하여 뼈대를 미리 설정한다. 상기 저장부는 예컨대 인간에 대한 미리 설정된 뼈대를 저장할 수 있으며, 이것은 인간이 이미지 데이터에 인식될 때마다 사용된다.
- [0062] 구조가 미리 설정된 뼈대를 갖지 않는다고 가정해 보자. 그러면, 뼈대 삽입 유닛은 기하학적 엘리먼트를 삽입함으로써 스크래치로부터 뼈대의 형성을 시도한다. 그러나, 기하학적 엘리먼트는 이미지 데이터에 적합하게 될 때까지, 회전 및 왜곡시킬 필요가 있을 수 있다. 따라서, 이미지 데이터에 적합할 때까지 선택된 엘리먼트를 회전시키는 회전 유닛(310) 및 데이터에 맞추기 위해 다양한 방법으로 엘리먼트를 왜곡시키는 왜곡 유닛(312)이 제공된다. 상기 회전 및 왜곡 유닛은 사용자 입력을 통해 동작하거나, 또는 자동으로 동작할 수 있다.
- [0063] 뼈대를 맞춘 후에, 상기 이미지 내의 구조들은 3차원 모델들로 모델링된다. 추적 유닛(314)은 초기 이미지 데이터 내의 움직임을 추적하고, 이미지와 함께 뼈대를 이동시켜, 움직임에 대한 3차원 정보를 획득할 수 있게 된다. 뼈대 및 이미지 데이터 사이의 프로젝팅 과정이 수행될 수 있으며, 그에 따라 단일 카메라로부터 3차원 및 움직임 정보를 획득할 수 있게 된다.
- [0064] 애니메이션 유닛(316)은 움직임이 뼈대를 통해 적용될 수 있도록 하여, 일단 모델링된 형상 또는 다른 오브젝트가 애니메이션화될 수 있도록 한다.
- [0065] 어플리케이션에 따라, 상기 장치는 추적 유닛 및 애니메이션 유닛을 모두 가질 필요가 없음이 이해될 것이다. 애니메이션 어플리케이션은 통상적으로 애니메이션 유닛을 갖지만, 추적 유닛은 필요로 하지 않을 수 있는 반면에, 비디오 캡처 어플리케이션의 경우는 추적 유닛을 갖고, 애니메이션 유닛은 필요로 하지 않는다.
- [0066] 렌더링(rendering) 유닛(318)은 추적 유닛 및 애니메이션 유닛 중 하나 또는 둘 다에 연결되어, 요구된 방향으로부터 시각화되도록 모델링되는 화면을 표현한다. 즉, 3D 데이터를 갖는 경우의 장점은 이미지가 최초에 촬영된 각도뿐만 아니라, 모델링된 오브젝트들을 임의의 각도로부터 관측할 수 있다는 것이다. 렌더링 유닛은 3차원 모델을 요구되는 시각화 방향의 평면으로 프로젝션하고, 텍스처(texture)를 이하에 설명되는 바와 같이 적용만 하면, 주어진 방향으로부터 장면을 관측할 수 있다.
- [0067] 도 4를 참조하면, 본 발명의 바람직한 실시예에 따라, 움직임 데이터를 포함하는 3차원 모델을 획득하기 위한

프로세스를 설명하는 간략화된 다이어그램이 도시되어 있다.

- [0068] 단계(402)에서 이미지 데이터가 획득되고, 이 데이터는 2D 데이터 또는 부분적으로 또는 완전한 3D 데이터이다. 이 데이터 내의 엘리먼트들이 식별된다. 단계(406)에서 뼈대 모델 엘리먼트들이 식별된 구조 엘리먼트와의 결합을 위해 삽입된다. 그리고 나서, 단계(408)에서, 뼈대는 식별된 구조 엘리먼트에 적합하도록 회전되고, 변형 또는 스케일링된다. 변형은 왜곡을 포함한다. 그리고 나서, 단계(410)에서, 사용가능한 정보에 따라 뼈대 모델 엘리먼트들 사이의 움직임 관계들이 정의된다.
- [0069] *도 5를 참조하면, 사람과 같은 복합 몸체가 인식된 경우에 대한 도 4의 흐름도의 변형예가 도시되어 있다. 다시, 단계(502)에서 최초의 데이터가 획득된다. 최초 데이터로부터 복합 몸체가 식별된다. 단계(504)에서, 데이터 저장부로부터 적절한 뼈대가 검색되고, 단계(506)에서 복합 몸체와 결합되어 삽입된다. 그리고 나서, 단계(508)에서, 뼈대는 회전, 변형 또는 스케일링된다. 변형은 왜곡을 포함한다. 결과적으로 식별된 구조 및 뼈대 모델 엘리먼트 간의 맞춰진 상태(fit)를 생성할 수 있게 된다. 도 4와 같이 뼈대 모델 엘리먼트를 복합 몸체에 맞추려는 시도는 가령, 인간과 같은 복합 몸체로 인식되어, 결과적으로 적절한 복합 뼈대가 선택될 수 있음이 주목된다.
- [0070] 도 6을 참조하면, 도 4 및 5의 프로세스를 모델링되는 오브젝트의 움직임으로 확장시키는 것을 나타낸 도면이 도시되어 있다. 단계들(602 내지 608)은 앞서 설명한 바와 같다. 단계(612)는 3D 오브젝트를 모델링하는 뼈대를 포함하여, 오브젝트의 움직임이 뼈대에 프로젝트되고, 및/또는 뼈대의 움직임이 2D 이미지 데이터로 프로젝트된다. 그리고 나서, 단계(614)에서, 이미지 데이터는 임의의 바람직한 방향으로부터 표현될 수 있다.
- [0071] 도 7을 참조하면, 본 발명의 바람직한 실시예에 따라, 이미지 데이터가 계층화된 구조로 어떻게 관리되는가를 나타내는 간략화된 다이어그램이 도시되어 있다. 2차원 또는 부분적으로 또는 완전한 3차원 이미지 데이터가 제1층(702)에 저장된다. 3차원 뼈대는 하부층(706)에 저장된다. 3차원 뼈대의 2차원 프로젝트션은 그 사이의 가상층(704)에 존재한다. 외관상으로 2차원 이미지는 3차원 뼈대를 다른 방향으로부터 프로젝트션함으로써 층(702)의 원래 이미지의 방향과 다른 방향으로부터 관측될 수 있다. 프로젝트션은 가상층(704)에 저장된다.
- [0072] 도 8을 참조하면, 최초 데이터의 오브젝트들 및 구조들이 뼈대 모델 엘리먼트의 할당을 위해 어떻게 인식될 수 있는가를 나타내는 간략화된 다이어그램이 도시되어 있다.
- [0073] 2개의 경로가 도시되어 있으며, 하나는 수동 경로, 단계(802)이고, 사용자는 단순히 상기 장치에 대해 엘리먼트, 몸체 및 복합 몸체를 인식한다. 선택적으로, 구조들을 식별하기 위해 단계들(804, 806, 808)로 구성된 자동 경로가 제공될 수 있다.
- [0074] 그룹핑 및 계속 진행할 것인가에 대한 결정은 하나의 단계로 볼 수 있을 것이다. 포인트 및 픽셀들은 일련의 이미지들에 대해 추적되며, 함께 이동하는 포인트들은 함께 그룹핑된다. 그룹핑 과정은 안정된 식별에 도달할 때까지 반복된다.
- [0075] 2개의 프로세스들을 혼합하여 사용할 수도 있다. 예를 들어, 사용자는 시스템에 임의의 엘리먼트들 또는 복합 몸체를 지적할 수 있고, 시스템은 다른 엘리먼트들을 식별하거나, 상기 복합 몸체 내의 개별 엘리먼트들을 식별하게 된다.
- [0076] 도 9를 참조하면, 본 발명의 일 실시예를 사용하여 움직임을 상기 제3 차원에 외삽(extrapolation)하는 것의 반복적 속성을 설명하는 간략화된 다이어그램이 도시되어 있다. 단계(902)에서, 픽셀들은 최초 2D 또는 부분적으로 또는 완전한 3D 이미지에서 추적된다. 단계(904)에서, 그 밑의 뼈대는 결합된 구조로 추적된 포인트들의 움직임에 따라 이동하고, 단계(906)에서, 최초 이미지의 픽셀의 3차원 위치를 결정하기 위한 외삽이 수행된다. 단계(904, 906)는 동시발생적이고 서로간에 영향을 미치므로, 이들은 도면에서 2중 화살표로 표시되어 있다.
- [0077] 본 발명의 바람직한 실시예에서, 컴퓨터로 생성된 시간 기반의 사진 시퀀스는 3개의 3D 모델로 구성될 수 있다.
- [0078] 최초 데이터에 대한 입력은 예컨대 이미 기록되어 공지된 비디오 포맷으로 압축된 비디오 클립촬영된 이미지로부터, 또는 USB나 공지된 임의의 다른 디지털 또는 아날로그 연결을 사용하여 직접적으로 연결된 단일 또는 복수의 카메라들로부터, 비디오 클립의 수신 및 디지털 기록용 모듈에 의해 제공될 수 있다.
- [0079] 다시 도 4 및 5를 참조하면, 예컨대, 하나 이상의 시간 기반으로 촬영된 이미지 시퀀스로부터 최초 데이터를 획득할 수 있다. 이 기술 분야에 공지된 바와 같이, 비디오 또는 필름 시퀀스들은 관측자의 눈에 움직임의 시간

허상을 생성한다.

- [0080] 입력 데이터가 분석된다. 바람직하게는, 상기 분석은 입력 시퀀스의 깊이맵을 구성하는 것, 각 시간 기반 시퀀스들의 깊이맵을 생성하는 것, 및 이하의 알고리즘 섹션에 설명될 바와 같이, 깊이맵을 프로세싱하는 것을 포함한다.
- [0081] 바람직한 실시예에서, 본 방법은 궁극적으로 촬영된 이미지(530)의 시퀀스에 의해 캡처된 오브젝트들에 대한 3D 모델을 생성하게 된다.
- [0082] 이러한 모델들은 실제 이미지 또는 그래픽 클립 등으로부터 재구성되며, 여기서 시간 차원은 모든 피겨(figure) 및 정적 배경이 3차원 모델인 공간 차원으로 변환된다.
- [0083] 이러한 3D 모델들은 이전에는 컴퓨터로 생성된 3D 가상 세계에서만 가능하였던 많은 조작을 가능하게 할 수 있다.
- [0084] 본 발명의 바람직한 실시예에 따라, 개별 피겨들은 차례로 식별될 수 있다. 일단 식별되면, 피겨들은 독립된 3D 모델로 변환될 수 있다. 피겨의 움직임은 순서대로 사진의 타이밍과 비교되어 사진 전체를 통한 진행으로부터 피겨의 움직임을 매핑하는 기초를 제공할 수 있다. 그리고 나서, 각 피겨에 대해 타임 라인을 조정하여 서로 다른 일련의 이벤트를 제공하는 것이 가능하다. 이런 방식으로, 모델링되는 피겨들은 변형될 수 있다.
- [0085] 예를 들어, 이 시퀀스에서, A,B 두 사람이 거리를 건너고, A가 B보다 먼저 건너편에 도달한 경우에, 우리는 각 피겨를 별개로 모델링하였으므로, 각 피겨들의 타이밍을 변화시킬 수 있다. 따라서, 우리는 B가 A 보다 먼저 거리를 건너야 한다고 결정할 수 있고, 따라서 촬영된 시퀀스의 원래 타임 라인을 변경하여, 시퀀의 변형을 수행할 수 있다.
- [0086] 상기 장치는 사용자가 생성된 3D 공간 기반 모델을 기반으로 미디어로부터의 수개의 서로 다른 출력을 생성할 수 있도록 해준다. 사용자는 외부 이미지 프로세싱, 애니메이션, 브로드캐스트 등 공지된 기술 프로그램으로 생성된 3D 공간 기반 모델들을 사용하거나, 에디터와 같은 내부 툴을 사용할 수 있다. 그러한 에디터는 사용자가 2개의 메인 출력들, 선형 미디어 및 비선형 미디어를 생성 및 편집하는 것을 가능케 할 수 있다. 선형 미디어는 타임 라인 기반 미디어 즉, 특정 시간 간격으로 촬영된 일련의 이미지들이다. 사용자는 그가 표현하고자 하는 타임 라인 상의 이벤트들을 기반으로 한 클립을 생성할 수 있다. 그리고 나서, 사용자는 다양한 뷰잉(viewing) 포맷, 예컨대 실시간 라이브 비디오 이미지 프로세싱, 비디오 클립, 모션 캡처, 정지 영상, DVD, 구형 이미지, 2D 이미지, 입체영상 이미지 또는 공지된 임의의 포맷으로 그 결과를 익스포트(export)할 수 있다.
- [0087] 본 실시예의 장치는 또한 비 타임 라인 기반, 즉 비선형 미디어를 생성할 수 있다. 그러한 비 타임 라인 출력은 예컨대, 이미지, 애니메이션, 및 텍스트 세트를 포함하는 3D 서라운드를 포함할 수 있다.
- [0088] *본 발명의 장치는 이 출력을 예컨대, 사용자가 그가 선택한 임의의 경로를 통해 걸어가고, 임의의 포인트에 도달하고, 그 포인트에서 주위를 360도 둘러보고, 임의의 피겨와 상호작용을 하는 것이 가능한 3차원 가상 환경으로 표현할 수 있는 능력을 제공한다. 그러한 비선형 출력으로는 컴퓨터 게임, 의료 수술 시뮬레이터, 비행 시뮬레이터 등과 같이 많은 예가 있을 수 있다.
- [0089] 상기 장치는 도 3의 애니메이션 유닛(316)에 따른 애니메이션 에디터를 포함할 수 있다. 애니메이션 에디터(316)는 사용자가 선택하는 모든 오브젝트에 생명을 불어넣을 수 있는 툴이고, 애니메이션 에디터(316)는 또한 그 오브젝트에 바람에 흔들리는 나무 또는 걸어가는 사람과 같은 특정 움직임과 함께 걸을 때, 달릴때, 화가 났을때, 슬플때, 어떻게 움직이는지, 얼굴 표정, 입술의 움직임 등등 고유한 특징을 할당한다. 애니메이션 에디터는 또한 오브젝트에 컴퓨터 애니메이션으로부터 또는 외부 소스로부터의 모션 캡처를 첨가하거나 장치 모션 캡처 툴을 사용하여 미리 설정된 움직임 세트들을 첨가할 수도 있고, 또한, 작은 움직임, 예컨대 이마의 주름 등과 같은 오브젝트의 특징을 나타내는 특징들과 움직임들 세트를 정의할 수 있다. 이러한 움직임들은 피겨의 인격을 생성하는데 도움을 주는 특징들이다. 애니메이션 에디터는 또한 오브젝트에 대한 장치 모션 캡처 툴을 사용하여 음성 특징도 생성할 수 있으며, 이것은 피겨가 말을 하는 것도 가능케 해준다.
- [0090] 상기 소프트웨어는 바람직하게는 이하에 설명되는 방법 및 알고리즘을 사용한다.
- [0091] 3D 공간 기반 모델 생성에 대한 상세한 설명

- [0092] 본 발명의 바람직한 실시예에 대한 기본 플랫폼은 이하의 설명과 같이 컴퓨터로 생성된 3D 축 엔진에 배치되고, 상기 3D 축 엔진에 해당하는 3개의 벡터 및 공간-시간 벡터를 이용한다.
- [0093] 입력 이미지 시퀀스 S는 플랫폼에 입력되는 일련의 이미지들이다.
- [0094] 바람직한 실시예는 이하의 설명과 같이 알고리즘을 구현할 수 있다.
- [0095] 시퀀스 S는 $N_f(s)$ 개의 프레임, 예컨대 PAL 비디오 디스플레이 표준으로는 25fps으로 분할된다.
- [0096] 제1 프레임 S0 (제1 시퀀스)는 $s(0,0)$ 으로 표시된다.(제1 시퀀스의 제2 프레임은 $s(0,1)$ 이고, 마지막 프레임은 $s(0,n)$ 으로 표시된다. ($\Rightarrow N_f(S0)=n+1$ 개 프레임))
- [0097] 다수의 앵커 포인트들이 사용된다. 앵커 포인트들은 2개의 메이저 엘리먼트들을 가지며, 한 엘리먼트는 S_i ($0 \leq i \leq$ 시퀀스 수)내의 엘리먼트들 사이에 대응하고, 제2 엘리먼트는 S_i 와 F로 표시된 3D 축 엔진 사이의 대응관계이다.
- [0098] 입력, 깊이맵 및 앵커 포인트
- [0099] 본 발명의 바람직한 실시예에 따른 알고리즘은 입력으로 S0를 수신할 수 있으며, 시퀀스에 대한 깊이맵을 생성하기 위한 모든 시퀀스 프레임을 사용할 수 있다. 팩터 D는 $s(0,0)$ 의 깊이 벡터로 정의된다(이하에서 정의함). z 가 프레임 $s(0,0)$ 으로부터의 서로 다른 픽셀들에 대한 깊이 값들 세트라고 가정한다. $s(0,0)$ 에서, $d(0,0)$ 는 $z_0=\min(z)$ 라 할때, z_0 에 해당하는 프레임 마법(witch) 깊이 값으로부터의 포인트들 세트이다. $d(0,h)$ 는 $z_h=\max(z)$ 이고, $\{z_i | \min(z/z_i), i=0, \dots, h-1\}$ (여기서 z/z_i 는 z_0 부터 $z[i-1]$ 까지 엘리먼트가 없는 집합 z 를 의미함)일 때, z_h 에 해당하는 프레임 마법 깊이로부터의 포인트들 세트이다. 따라서, $\{z_0, z_1, z_2, \dots, z_h\}$ 는 프레임 $S(0,0) \Rightarrow$ 벡터 D0의 깊이 층의 세트이고, 이 세트는 아래에서 위로 정렬되어 있으며, 이 세트의 숫자들은 이하에 설명될 바와 같이 F의 해상도 팩터에 대한 벡터 D0의 층들이라는 것을 정의로부터 명확하다. 예를 들어, $D0=\{2,5,6,9,13,56,22,89\}$ 이면, $\{z_0=2, z_1=5, z_2=6, z_3=9, z_4=13, z_5=22, z_6=56, z_7=89\}$ 이다. D0은 $S(0,0)$ 에 대한 깊이 벡터이고, $D0=\{d(0,0), d(0,1), \dots, d(0,h)\}$ 이다.
- [0100] 프레임 $S(0,i)$ 의 D는 3D 행렬이다. 프레임 $S(0,i)$ 의 D는 2D 마스크 행렬의 벡터이다. 픽셀 $d(i,j)$ 의 깊이가 어떤 이유 때문에 정의되지 않은 경우에는, $d(i,j)=$ 무한대이고, 불리언 2D 행렬에서 0으로 정의될 것이다.
- [0101] Note : 프레임 $d(i,j)$ 로부터의 픽셀의 깊이값이 프레임 $S(0,i)$ 으로부터의 2D 이미지로부터의 이미지로 정의될 수 없는 경우에는, 알고리즘은 동일한 시퀀스 $S(S(0,i))$ 를 취한 마법으로부터의 시퀀스)로부터의 다중 프레임으로부터의 데이터를 이용하여 정의를 시도할 것이다. $s(0,0)$ 의 깊이맵 또는 그 일부가 불량 라이팅, 예컨대, SP가 $d(0,0)$ 를 처리한다거나, 정의된 부분이 일시적으로 " ∞ (무한대)"로 정의되고, $s(0,1) \dots s(0,i)$ $S_i(0 \leq i \leq$ 시퀀스의 수)를 사용하는 경우에는, $s(0,0)$ 의 계산을 시도한다.
- [0102] 프레임 $\{S(0,i) | i=0,1,2, \dots, n\}$ ($\in S$ (시퀀스 번호, 프레임 번호))의 깊이 맵이 계산되는 경우에, $D0=\{d(0,0), d(0,1), \dots, d(0,h)\}$ ($\in D$ (프레임 번호, 깊이 앵커 포인트))이다. 소프트웨어는 깊이 맵 행렬(깊이값 세트의 최대값)으로부터 값의 상한을 구한다.
- [0103] SP는 또한 깊이맵 행렬로부터 하한을 구한다. 시퀀스 S0 깊이맵의 모든 프레임이 성공적으로 프로세싱된 경우에는, SP는 시간상으로 모든 정의된 순간에 시퀀스 S0의 상한 및 하한 앵커 포인트를 구한다.
- [0104] $s(0,0)$ 에서, $d(0,0)$ 는 깊이가 가장 근접한 포인트이다. $s(0,0)$ 의 가장 깊은 포인트는 $d(0,h)$ 로 표시되며, $D0 \in \{d(0,0), d(0,1), \dots, d(0,h)\}$ 이다.
- [0105] 팩터 D는 알고리즘에서 깊이 벡터의 클래스이고, 여기서 이미지 깊이 구조맵들을 상호관련시키기 위한 작업 툴로서 데이터를 분석하기 위해 수개의 D 벡터들이 사용된다. SP 구조 맵은 F에 생성되고, 정적 및 이동하는 엘리먼트를 모델링하고, 엘리먼트들의 부분들(예컨대, 손, 다리 등등)을 표현하기 위해 F 내부에 오픈된 다수의 새로운 행렬들을 사용한다. D는 이 벡터를 따른 모든 포인트가 현재 깊이에서 해당 깊이 정보를 포함하도록 생성되고, 위상 맵(topological map)의 고도선과 같이 모든 포인트의 깊이 조각을 따라 깊이값을 표현한다.
- [0106] D는 3D 행렬이고, D를 따라 모든 Z 포인트에 대해 2D 불리언 이미지 행렬(x,y)로서 생성되며, 모든 2D 이미지 행렬(x,y)에서 해당 깊이 포인트(Z)의 이미지에 포함된 정보만 "1"로 표시한다.
- [0107] 비디오 시퀀스 S0의 프레임들 간의 참조 앵커 포인트를 구하기 위해, 시스템은 다음과 같은 과정을 수행한다.

- [0108] $\{x, y, z\} = \{\text{수평, 수직 깊이}\} = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ 는 F로부터의 공간 벡터이고, 우리는 그것을 "세계 좌표계"이라고 부를 것이다.
- [0109] 프레임 $S(0, t)$ (t 는 임의의 숫자)을 보면, 이 프레임은 그 로컬 좌표계 {칼럼 인덱스, 로우 인덱스, 깊이}을 갖는다.
- [0110] 이 프레임에서 k 번째 앵커 포인트를 가정해 보자. 프레임 로컬 좌표계는 $(u_{tk}=i, v_{tk}=j, \text{깊이}(i, j))$ 이다.
- [0111] 이 프레임 로컬 좌표계의 간격(베이스)은, 세계 좌표계에서는, $\{i_t, j_t, k_t\}$ 이고, $i_t, j_t, k_t = (i_t \times j_t)$ 는 세계 좌표계의 벡터들이다.
- [0112] 프레임 로컬 좌표계에서, i_t 는 $(1, 0, 0)$ 이고, j_t 는 $(0, 1, 0)$, $k_t(0, 0, 1)$ 이다.
- [0113] 비디오 시퀀스 S_0 의 프레임들 간의 참조 앵커 포인트를 구하기 위해, 시스템은 각 프레임 $(S(0, 0), S(0, 1) \dots S(0, n))$ 을 벡터 공간 S_0 의 서브 공간(즉 벡터 공간 자체)으로 취급하고, 이것은 필드 F 의 위에 존재한다. 시스템은 벡터 베이스 $W_0 \in S(0, 0)$ 를 계산하고, 여기서 $W_0(W_0 = \text{Sp}\{w_1, w_2, \dots, w_m\})$ 는 서브 공간 $S(0, 0)$ 을 넘어서 확장된다. 각 서브 공간에 대해 수개의 서로 다른 베이스가 존재할 수 있지만, 각 베이스의 간격은 W_0 를 넘어서 확장되고, 벡터 베이스 벡터의 역할을 수학에서의 역할과 유사하다. 이러한 벡터들은 서브공간 $W(0, 0) \in S(0, 0)$ 를 생성한다.

[0114] 깊이 정렬

- [0115] *이미지 배경과 같은 고정 오브젝트들의 깊이 정렬은 2단계로 수행된다. 제1 단계에서, 시스템은 $W(0, 0)$ 의 벡터 베이스를 구하며, 여기서 $\text{Sp}\{w_1, \dots, w_m\} = W(0, 0) \in S(0, 0)$ 이다. 빠른 정렬을 위해, 시스템은 벡터 베이스 $(0, 0)$ 로부터 4개의 참조 벡터, 수평, 수직, 깊이, 공간/시간 벡터를 생성한다.
- [0116] 제1 벡터 $Z \in S(0, 0)$ 는 $d_0 \in S(0, 0)$ 의 모든 포인트의 베이스 벡터의 수를 반영하고, 프레임의 베이스 벡터의 깊이 정보를 표현하는 Z 벡터를 생성한다. Z 의 중간점은 $d(\text{anc})$ 로도 표현되며, 이것은 프레임 자체의 중심점이다.

$$d(\text{anc}) = \frac{d(0, 0) + d(0, m)}{2}$$

- [0117]
- [0118] $d(\text{anc})$ 는 시스템이 0 포인트 축 $XYZ \in F$ 으로 표시한 포인트가 될 수 있다. 수평 및 수직 벡터들은 Z 벡터를 따라 모든 수평 및 수직 포인트의 벡터들을 표현한다. 제4 참조 벡터는 시간차원에서 공간차원로의 변형 벡터로 사용되는 공간/시간 벡터이다. 시스템은 D' 로 통일된 정렬에 대해 $S(0, 0)$ 및 $S(0, 1)$ 사이에 사용될 3개의 참조 벡터를 생성한다.

- [0119] 프레임들간의 차이점들은 라이팅(lighting), 프레임 내의 엘리먼트들의 이동 및 트랙 인/아웃, 트랙 좌측/우측, 크레인(crane) 업/다운, 틸트 업/다운, 팬 좌측/우측 및 줌과 같은 카메라 동작의 팩터가 될 수 있다(광학 또는 디지털 사진에 관해서는, 상기 차이점은 디지털 줌에서 더 낮은 인치당 픽셀의 양에서 발견할 수 있다). 프레임 간의 서로 다른 시프트는 대개 픽셀의 위치와 같은 형태로 발견되고, 프레임들 간에 몇가지 방식의 픽셀의 쉬프트가 존재할 수 있으며, 따라서 SP_1 은 공간/시간 벡터의 함수로서 S_0 프레임의 3개의 참조 벡터들을 계산한다. 이미지의 3D 정렬을 위해 3개의 해당 벡터들이 구성되고, 여기서 수직 및 수평 벡터는 공간 윈도우 (X, Y) 에 해당하고, Z 벡터는 깊이 벡터에 해당한다.

- [0120] 공간 (X, Y) 벡터의 각 팩터는 이미지의 모든 포인트에서 Z 벡터를 따라 공간 도메인의 이미지의 베이스 벡터를 반영한다.

- [0121] 매칭 함수는 0의 차이를 갖거나, 또는 가능한 한 많은 포인트에서 0의 차이를 갖도록 타게팅된, 이미지 1의 벡터들 $V(h_0/v_0, z_0)$ 과 이미지 2의 벡터들 $V(h_1/v_1, z_1)$ 간의 미리 설정된 포인트의 최소값을 갖는 것이 바람직하다. 벡터의 통일된 섹션의 정렬에 관해서는, 두 벡터 모두의 반대 양 에지 각각에는, 프레임 간의 차이점 포인트에 불일치가 존재할 수 있다. 이러한 포인트 차이는 새로운 프레임에 추가될 수 있는 서로 다른 정보를 가리키지만, 이전 프레임에는 나타나지 않는다.

$$V'0 \in V0 \in V0 \cap V1,$$

$$V'1 \in V1 \in V0 \cap V1,$$

[0122]

[0123] 3개의 벡터들은 이미지들의 3차원 위치 정보의 결과물이고, 시각적인 정보와는 관계가 없으며, 모든 포인트에서 이미지의 베이스 벡터들을 나타낸다.

[0124] 바람직하게는, 수평, 수직 및 깊이 벡터들은 각 벡터와 별개로 비교되어, 가능한 많은 포인트에서 최소 차이를 구하게 된다.

[0125] 프레임 간의 포인트 차이에 불일치가 존재할 수 있다. 이러한 불일치는 이전 프레임에 존재하지 않는 다른 정보가 현재의 프레임에 존재한다는 것을 표시할 수 있다.

[0126] 왜곡 및 카메라 이동 처리

[0127] 카메라 렌즈와 같은 광학 엘리먼트는 촬영된 이미지의 왜곡을 생성하고, 동일한 오브젝트에 기인한 깊이맵에서의 작은 차이를 생성할 수 있다.

[0128] 도 14를 참조하면, 2개의 이미지가 도시되어 있고, 카메라 이동 왜곡에 대한 포인트가 도시되어 있다.

[0129] 제1 프레임(1401)에서, 돌 칼럼(1403)은 프레임의 중심에 존재한다. 제2 프레임(1405)에서 동일한 칼럼(1403)은 프레임의 우측에 존재한다. 광학 프로세스의 결과로 약간의 왜곡이 존재할 수 있으며, 깊이 맵은 그러한 왜곡에 기인한 작은 차이를 가질 수 있다. 결과적으로, 정렬된 이미지들의 출력 결과에 약간의 불연속이 존재할 수 있으며, 이 불연속은 이들 왜곡의 결과물이다. 즉, 프레임의 중심에서 측면으로 이동할 때, 광학적인 이유로 블록의 구조가 변화되는 것으로 보인다.

[0130] 해결책은 주어진 오브젝트의 최상의 이미지를 하나로 식별하는 것인데, 이것은 상대적으로 프레임의 중심에 나타나게 된다. 픽셀들은 이 가장 정확한 측정으로부터 획득된 3D 위치를 수신한다.

[0131] Z 벡터를 정렬할 때, 우리는 또한, 이미지(광학 또는 디지털) 간의 줌 팩터 또는 카메라의 트랙 "인" 또는 "아웃"에 차이가 존재하는 문제점을 갖게 된다.

[0132] 프레임 간의 카메라 쉬프트가 작을 수록, 벡터간의 상관관계는 더 크며, 그 결과는 더 양호하다. 이미지 간의 차가 클 수록, 벡터들 간의 상관관계는 더 작다. SP는 일탈에 대한 임계값을 설정하고, 더 큰 차이점을 갖는 위치를 서로 다른 오브젝트에 속하는 것으로 간주한다.

[0133] 정렬 후에, 수평, 수직 및 깊이 벡터들의 통일된 섹션은 다음과 같다.

$$V'h/v(0) \in V h/v(0) \in V h/v(0) \cap V h/v(1),$$

$$V' h/v(1) \in V h/v(1) \in V h/v(0) \cap V h/v(1)$$

$$V' z0 \in V z0 \in V z0 \cap V z1$$

$$V' z1 \in V z1 \in V z0 \cap V z1$$

[0134]

[0135] 상기 통일된 섹션들은 이제 서브 공간으로 취급되고, 벡터들을 이 서브 공간의 참조로서 재산출한다. 줌 팩터는 고려대상이 되고, F 필드에 대한 "스칼라" 팩터, 벡터를 곱하거나 나누고 카메라의 줌/트랙 인 또는 아웃을 모방하는 스칼라로 벡터를 산출하며, 여기서 상기 프레임들의 엘리먼트들 간의 동일한 관계는 저장되지만, 해상도는 달라질 수 있다. 스칼라를 이용하여, 우리는 S(0,0)의 벡터를 S(0,1)과 정렬할 수 있다. 이 프로세스는 이미지를 정렬시킬 수 있고, 또한 다음 프레임의 정렬 방향을 지시할 수도 있다. 공간/시간 벡터는 시간 도메인에서 공간 도메인으로의 변환과 관련된 것이고, 새로운 정렬은 통일된 프레임 $\in F$ 로 간주되며, 다음 프레임은 이전의 통일된 프레임과 정렬된다. 이것은 특히 프레임이 이미 정렬된 영역을 반복하는 경우에 계산을 감소시킬 수도 있다. 공간/시간 벡터는 시간 차원에서 공간 차원으로의 변형에 대한 참조 벡터이다.

[0136] 일관되지 않은 깊이 정보의 처리

- [0137] 깊이 정보가 일관성이 없는 경우에, 장치는 바람직하게는 F1로 표시되는 새로운 벡터 평면을 오픈한다. 이 새로운 평면은 이 알고리즘의 과정이 처음부터 시작된 비어있는 XYZ 좌표계이다.
- [0138] 그러한 프로세스의 마지막에, 상기 좌표계의 사용자는 F0과 F1을 서로 다른 위치로 남겨놓기를 원하는지, 아니면 그들을 정렬하기를 원하는지를 질문받게 된다. 그 다음 사용자는 2개의 구조들을 수동으로 정렬하기 위해, 회전, 줌, 플립 등의 툴을 사용하여 F0 및 F1을 수동으로 정렬할 것인지를 질문받는다.(선택적으로, 사용자는 시스템에 자동으로 F0 및 F1을 정렬할 것을 명령할 수도 있다.
- [0139] 사용자가 F0과 F1을 수동으로 정렬한 후에, 사용자는 시스템에 이 정렬을 계산하도록 명령하고, 시스템은 정렬 알고리즘을 사용하여 필드의 정렬을 시도한다. 필드가 잘 정렬되면, 시스템은 그것을 공표하고, 그렇지 않으면 SP가 사용자에게 비정렬 팩터에 대한 외삽을 낮추어 줄 것을 요구한다(정확도가 낮은 정렬). 시스템은 공지된 이미지 프로세싱 툴을 사용하여 사용자에게 이미지 평면에서의 불연속을 극복하기 위한 툴 박스를 제공한다.
- [0140] SR 해상도
- [0141] 시스템 1은 "F0"(XYZ 축의 필드) 표시의 임시 해상도를 "R0"으로 정의한다. R은 인치당 참조 포인트의 수로 정의된다. 해상도는 시간 차원에서의 픽셀들에 의해 이미지 해상도의 조합 팩터 및 공간 차원에서의 깊이의 포인트 조합의 결과이다. 해상도 툴은 서로 다른 거리로부터 동일한 위치를 촬영하는 2개의 비디오 클립의 정렬에 있어서의 예로서 도움이 될 수 있다.
- [0142] 예를 들어, 테이블은 예컨대, 한쪽 다리에서 그 옆의 다리까지 테이블의 일부분들 간, 또는 더 가까운 위치에서 더 낮은 해상도를 갖는 제2 클립까지, 또는 적은 수의 참조 포인트가 존재하는 먼 거리로부터 디지털 줌을 사용함으로써 더 많은 참조 포인트가 존재하는 고 해상도 클립으로 촬영될 수 있다.
- [0143] 해상도 문제를 처리하는 참조 포인트는 현실을 참조하는 모든 픽셀의 3D 위치이다. 따라서, 공간 차원의 픽셀의 3D 위치는 시간 차원으로부터 변환된 후의 계산된 위치이다. 해상도는 $D0$ 이 $S0$ 와 일치되도록 한다. $D0=d0(anc)$ 의 중간점은 임시로 포인트, $(0,0,0) \in (X,Y,Z) \in F0$ 에서 축 필드의 중심에 배치될 수 있다.
- [0144] 참조 포인트의 시각 정보는 이하에 설명될 바와 같이, 정보의 시각 층으로서 F0에 배치될 수 있다.
- [0145] 움직이는 엘리먼트를 식별 및 재구성하기
- [0146] 이미지에 움직이는 엘리먼트가 존재하는 경우이다. 뼈대는 움직이는 그래픽 엘리먼트로 구성되고, 그 상대적인 위치 및 이동 패턴을 정의하여, 이동한 엘리먼트의 매우 정확한 3D 기하학적 모델을 구성하고, 그 모션 캡처를 저장하며, 촬영된 시각 정보를 3D에 모델에 첨가한다. 이 시스템은 식별 및 재구성 프로세스의 자동화를 가능케한다.
- [0147] 첫째, 시스템은 데이터 이미지 시퀀스에 이동하는 엘리먼트가 존재한다는 것을 알아야 한다. 다음 단계는 데이터 이미지의 상기 엘리먼트와 기정렬되거나 또는 사용자가 정의한 뼈대 모델 엘리먼트를 동일시하는 것이다. 마지막으로 상기 시스템은 미리 설정된 3D 구조 및 뼈대를 사용하여 엘리먼트의 3D 구조들의 재구성을 수행하거나, 상기 엘리먼트의 특성을 기반으로 단계적으로 구성된 새로운 독립 3D 구조를 생성한다.
- [0148] 카메라에 대해 그 배경과 다른 정보를 추가하는 움직이는 엘리먼트들은 시간에 따라 작은 정보를 추가하는 바람에 흔들리는 나무와 같은 준(semi) 정지 오브젝트 또는 프레임을 통과하고, 회전하고, 다른 편의 프레임으로부터 걸어 나오는 사람이다.
- [0149] 위에서 언급한 바와 같이, 시스템은 우선 시퀀스에 움직이는 오브젝트가 존재하는지를 알아낸다. 그 다음, 시스템은 미리 설정된 3D 엘리먼트 또는 뼈대 세트를 사용하여 이 오브젝트를 식별한다. 선택적으로, 사용자는 뼈대 또는 엘리먼트를 정의하여 상기 피겨에 첨가할 수 있다.
- [0150] 그리고 나서, 시스템은 미리 설정된 3D 엘리먼트 또는 뼈대 또는 새로운 사용자 정의 엘리먼트를 사용하여 피겨의 3D 구조를 구성한다.
- [0151] 프레임에 움직이는 엘리먼트가 존재하는지 식별하기 위해, 시스템은 공간 및 시간에 걸쳐 시퀀스의 깊이 픽셀의 불연속 점을 검색한다. 즉, $S0$ 에는 공간 차원의 카메라 및 배경에 대해 $S0$ 의 고정점과 일치하지 않고, 시간 차원에 따라 그 정보를 변화시키는 임의의 특별 3D 구조가 존재할 수 있다.

- [0152] 즉, 시간에 따라 공간의 비정렬이 존재한다. 예를 들어, 우리가 우측으로 이동하는 카메라를 사용하여 테이블을 촬영하는 경우, 테이블은 처음에는 프레임의 우측에 있다가 프레임의 좌측으로 이동하게 된다.
- [0153] 시간에 따라 정보가 변화하는 테이블의 앞에 3D 엘리먼트가 존재하는 경우에는, 상기 시스템은 프레임에 이동하는 오브젝트가 존재한다고 결론낼 수 있다. 시스템은, 상기 테이블은 정지 엘리먼트인데도, 이동하는 엘리먼트의 3D 모델을 재구성한다.
- [0154] 차원-기반 이미지 프로세싱에서 3D 엘리먼트를 생성하기 위해, 엘리먼트의 주변 영역으로부터 매칭 벡터가 구성될 수 있다.
- [0155] $N \times M \times Z$ 크기의 사각형 윈도우가 선택될 수 있으며, 3D 행렬이다. N 및 M 은 윈도우의 공간적 크기이고, Z 는 깊이 차원이다. 시간 차원에서 공간 차원에서의 엘리먼트 또는 오브젝트의 변환 차원을 정의하기 위해 제4 벡터가 제공되어 3D 엘리먼트 및 피겨를 구성할 수 있다. 시간 및 공간 차원 모두에서 엘리먼트를 매칭시키면 각 엘리먼트 또는 오브젝트에 대해 별개로 구성되는 3D 구조 맵에 해당하는 모든 포인트들을 일관되게 매칭시킬 수 있다.
- [0156] 본 실시예의 알고리즘은 그 배경 및 절대 3D 서라운드에서 상대적인 프레임의 움직이는 엘리먼트의 추적에 도움을 주기 위해 3D 정보 구조를 2D 이미지로 프로젝션한 것을 기반으로 한다.
- [0157] 엘리먼트의 식별
- [0158] 이미지 데이터로부터 현재 엘리먼트를 식별하는 과정은 기정렬된 3D 구조 집합의 도움으로 수행될 수 있다. 시스템은 전체 구조의 식별시까지 엘리먼트의 형태 또는 그 일부를 결정하기 위해 단계적으로 동작하고, 사용자가 새로운 구조를 형성하는 것을 보조한다.
- [0159] 상기 시스템에는 볼, 박스, 파이프 등의 단순한 3D 기하학적 3D 모델에서 부터, 고정 및 비고정(non-rigid) 몸체의 완전한 뼈대까지 뼈대 모델 엘리먼트에 대한 3D 구조의 집합을 포함하는 데이터베이스가 제공될 수 있다. 뼈대는 예컨대, 도12에 도시된 바와 같이, 자동차에서부터 동물 또는 사람의 뼈대까지를 식별하고 모델링하는 직사각형 영역이 될 수 있다.
- [0160] 뼈대는 다음 3개의 주요 엘리먼트를 포함하는 복합 3D 데이터 구조이다.
- [0161] 1. 뼈대의 물리적 결합, 즉 구성 뼈대 모델 엘리먼트들의 형태 및 상호관계
- [0162] 2. 입력 3D 정보에 따른 뼈대의 형태 정보
- [0163] 3. 몸체(뼈, 근육, 조인트 등)의 물리적 구조와 같은 내부 정보와 몸체의 물리적 행동의 결합
- [0164] 상기 3개의 관점들이 본 실시예의 알고리즘에 따른 식별 및 재구성에 요구된다.
- [0165] 1. 이러한 부분들, 뼈대 모델 엘리먼트들로부터, 시스템은 사용자의 요청에 따라 해당 몸체를 구성하거나, 새로운 몸체를 구성할 수 있다는 의미에서, 뼈대의 결합은 뼈대의 구조를 취하여 몸체 부분들의 최소 정의까지 그 부분들을 정의하는 것을 가리킨다.
- [0166] 예를 들어, 인간의 팔은 결합부를 통해 손을 상징하는 또 다른 실린더에 연결된 3D 실린더를 기반으로 할 수 있다. 또다른 예로, 머리는 3D 볼의 단순 형태로 시작하여, 목을 나타내는 결합부에 연결될 수 있다. 그 다음 목은 몸통을 나타내는 큰 실린더에 연결된다. 인간, 동물 등의 뼈대 부분들의 서로 다른 물리적 움직임 및 개별 엘리먼트들의 움직임은 본 기본적인 구성을 재구성하기 위해 통합되고, 시스템이 그 피겨를 식별하고 재구성하는 것을 보조하게 된다.
- [0167] 2. 이하의 설명과 같이, 일부분이 변형되는 뼈대를 나타내는 도 13을 참조하여, 입력된 3D 입력에 따라 뼈대를 형성하는 능력이 식별 과정 및 재구성 과정에서 사용된다.
- [0168] 3. 이하의 설명과 같이, 몸체(뼈, 근육, 조인트 등)에 물리적 구조와 같은 내부 정보 및 그 물리적 움직임이 식별 과정 및 재구성 과정에서 사용된다.
- [0169] 3D 구조 및 뼈대 집합을 사용하여, 시스템은 엘리먼트의 식별을 결정한다. 이 과정은 도 8에 도시된 바와 같이 자동 또는 사용자에 의해 수동으로 수행될 수 있으며, 시스템에 대한 해당 엘리먼트를 식별하고, 내부 뼈대를 피겨에 부착하건, 새로운 구조를 구성하는 과정을 포함한다.

- [0170] 움직이는 오브젝트가 일련의 프레임에 위치되어 있기 때문에, 시스템은 그것을 식별하고, 이전에 정의된 뼈대 모델 엘리먼트들 집합, 또는 사용자에게 의해 움직이는 오브젝트에 대해 정의된 특정 엘리먼트로부터 선택된 매칭되는 뼈대 모델 엘리먼트들의 부가를 시도하게 되며, 바람직하게는 시스템에서 제공하는 도구 집합을 사용하게 된다. 바람직하게는, 부가된 뼈대 모델 엘리먼트들은 크기, 모양 및 움직임 패턴에 의해 움직이는 오브젝트를 피트(fit)시키기 위해 움직이는 오브젝트의 크기, 모양 및 움직임 패턴에 자동적으로 맞춰진다. 시스템은 적절하게 중첩된 텍스처(texture)로 뼈대 모델 엘리먼트들 집합을 완성한다.
- [0171] 본 발명의 바람직한 실시예에서, 시스템은 움직이는 오브젝트를 임의의 바람직한 관점에 대한 2D 평면에 외삽(extrapolate)시키는 도구를 더 제공한다.
- [0172] 3D 구조 기반의 정렬의 속성을 이용하면, 서로 다른 비디오 시퀀스들 사이, 서로 다른 스케일(줌) 차이에서 매칭시키는 과정, 서로 다른 센싱 양상(modality)(IR 및 가시광 카메라) 등과 같은 다양한 상황에서 정보를 매칭시킬 수 있다.
- [0173] 움직이는 오브젝트로부터 엘리먼트의 3D 구조를 생성하는 것은 시스템이 서로 다른 시퀀스들 간의 표현의 차이점을 처리하는 것을 보조하는 기본 팩터이다.
- [0174] 엘리먼트는 튜브로 구성된 기본 뼈대, 팔, 다리 및 몸체에 결합된 조인트들, 및 머리에 결합된 볼이 부가될 수 있으며, 깊이 정렬은 3D 엘리먼트 구조의 생성과 3D 피겨의 상호 연관에 팔, 몸체, 다리, 머리에 결합된 볼의 크기에 대해, 기본 뼈대, 길이 및 튜브의 두께의 물리적 움직임과 같은 새로운 정보를 부가할 수 있다.
- [0175] 이러한 분리된 튜브 및 볼들로부터 완전 3D 피겨를 구성하는 것은 상호 간의 움직임, 즉 서로 어떻게 결합되어 있는가, 또는 어떻게 움직이는가를 보여줄 수 있다. 이 단계에서, 시스템은 어떤 엘리먼트가 마주보는가를 결정하거나, 또는 어떤 엘리먼트인지를 결정할 수 없다고 판단하고 사용자에게 하여금 해당 피겨를 결정하는 것을 도울 것을 요청할 수 있고, 또는 사용자가 미리 설정된 기본 피겨 집합에 존재하지 않는 새로운 구조를 형성할 수도 있다.
- [0176] 상술한 바와 같이, 소프트웨어는 깊이 정보를 사용하여 움직이는 오브젝트의 구조를 가능한 한 많이 식별하려고 한다. 이전에 정의된 3D 엘리먼트들 집합의 도움으로, 단계적으로, 소프트웨어는 오브젝트 부분들의 형태를 결정하고, 시각적인 정보 중 일부가 존재하지 않더라도 완전한 구조를 완성하게 된다.
- [0177] 제1 단계는 오브젝트를 식별하고 그 기본 형태를 결정하는 것이다. 그리고 나서, 시스템은 가능한 한 많이 그것을 완성시키려고 한다.
- [0178] 기본 형태를 사용하여, 사용자는 공간 도메인과 깊이 도메인에서 엘리먼트에 대한 정보를 얻게 된다. SP는 3D 뼈대 모델 엘리먼트들 집합(예컨대 볼, 박스, 파이프 등)을 사용하여 오브젝트의 상세한 부분들을 재구성한다.
- [0179] 단일 카메라를 사용하는 움직이는 엘리먼트들의 3D 구조 맵
- [0180] 시스템은 이미지의 완전 깊이 3D 맵을 수신할 수 있다. 움직이는 엘리먼트들을 포함하는 이미지의 깊이 맵을 구성하는 알고리즘이 이 기술분야에 알려져 있다. 예컨대, 공간 타임 스테레오 알고리즘을 사용하는 깊이 구조 맵은 적어도 2개의 카메라를 이용한다.
- [0181] 예컨대, 공간 시간 스테레오 알고리즘을 사용하는 정적 주변환경들에 대한 이미지의 깊이 맵을, 예컨대 하나의 카메라로 구성하는 기술 알고리즘이 알려져 있다.
- [0182] 깊이 맵을 외삽(extrapolate)하지 않고도, 하나의 카메라를 사용하여 비디오 시퀀스로부터 정적 모델들을 생성하는 기술 알고리즘이 알려져 있다.
- [0183] 본 알고리즘은 상술한 바와 같이, 하나의 카메라를 사용하여 움직이는 엘리먼트들의 깊이맵을 외삽하는데 사용될 수 있다.
- [0184] 움직이는 엘리먼트의 깊이 값의 구성을 위해, 시스템은 기술 알고리즘에 공지된 기술을 사용하여 정적이고 고정된 배경의 미리 획득한 깊이 값을 사용할 수 있고, 상기 움직이는 엘리먼트는 참조 포인트들을 사용하는, 그 배경과 관련된 독립 4D 행렬을 가리킨다.
- [0185] 3D 정보 구조(미리 설정된 3D 뼈대와 같은)를 2D 이미지 평면에 프로젝션하는 것은 깊이 축에 의해 각 프레임에서 움직이는 엘리먼트를 추적하는 것을 도울 수 있다. 2D 이미지 평면을 3D 공간으로 프로젝션하는 것과 함께,

엘리먼트의 깊이맵을 생성할 수 있는 능력이 제공된다. 뼈대와 기관(organ)을 부가함으로써 이미지를 깊이 맵에 맞추고, 임의의 오브젝트를 종합적으로 복사하고 그 움직임을 캡처한다. 후자는 재구성된 뼈대 상에 엘리먼트의 텍스처를 중첩시키는 과정을 더 포함하고, 이하의 설명과 같이 재구성 과정을 완료시킨다.

- [0186] 따라서, 본 발명은 프레임에서 움직이는 3D 맵의 생성을 강제한다.
- [0187] 깊이 외삽의 제1 단계는 각 프레임을 따라 모든 픽셀의 2D 위치를 추적하여, 각 픽셀의 궤도를 생성하는 것이다.
- [0188] 이 추적은 추적 알고리즘에 공지된 기술을 사용하여 행해진다.
- [0189] 앞서 설명된 바와 같이, 2개의 이미지 및 시간 차원을 따라 동일한 픽셀을 발견하는 수동적인 방법은 픽셀의 컬러, 음영, 밝기 및 모호성(ambiguities)를 사용하여, 2개의 프레임에서 시간축을 따라 동일한 픽셀의 위치를 찾는다.
- [0190] 시간상으로 2D 이미지 평면에서의 본 추적 알고리즘은 식별을 보조하기 위한 임의의 프로젝트된 패턴이 없으므로, 시간에 따른 에러를 수집하게 된다. 따라서, 2D에서의 추적은 복잡한 문제가 된다. 그러한 2D 추적로부터의 외삽 깊이 값은 실제 깊이 맵이 될 수 없다.
- [0191] 그러나, 2D 이미지 평면 상에 프로젝트된 데이터 구조로 3D 뼈대를 사용하는 경우에는, 각 프레임은 한편으로는 몸체의 식별된 기관(organ), 예컨대 손, 왼쪽 다리, 오른쪽 다리 등과 같은 3D 데이터 구조의 2D 프로젝트이고, 다른 한편으로는 2D 데이터 구조의 3D 프로젝트이다.
- [0192] 상기 결과로 추적 실패가 발생하지 않게 된다. 뼈대의 3D 자세가 시간 축을 따라 2D 평면상에 프로젝트되기 때문에, 모든 프레임에서 무한으로 새로운 추적 포인트들을 갖는 정확한 추적은 3D 프로젝트로부터 생성되고, 시스템은 3D 몸체의 숨겨진 부분들이 어디에 있는지 및 3D 공간에서 어디에 있는지를 알게 된다. 상기 시스템은 3D 부분들을 요구된 부분들이 현재는 소스 2D 이미지에서 볼 수 없는 경우라 하더라도, 임의의 요구된 2D 이미지 평면으로 프로젝트할 수 있다.
- [0193] 미리 결정된 3D 뼈대는 2D 이미지 평면으로 프로젝트된다. 상기 시스템은 사실상 도 7을 참조하여 설명된 바와 같이 정보의 추가 계층에서 이미지 같은 그림자를 생성한다. 정보의 추가 계층은 추적되어야 하는 이미지의 부분들을 정확히 조준하고, 에러를 즉시 제거하여, 에러의 성장을 방지한다. 그러한 단계는 고정 및 비고정 엘리먼트를 포함하는 걷는 사람과 같은 움직이는 엘리먼트의 깊이를 추적 및 외삽을 가능케한다. 그리고 나서, 3D 뼈대는 이미 추적된 움직이는 2D 엘리먼트의 3D 깊이 맵을 외삽하는데 사용될 수 있다. 2D 이미지에서 픽셀의 시간을 따라 궤도를 강제하기 위해 3D 뼈대 상에 위치한 무한 포인트들을 사용하는 경우에, 앞서 설명에서 2D 공간에 대한 3D 공간의 프로젝트 포인트로 설명된 4D 행렬이 엘리먼트를 둘러싸고, 그 반대도 가능하다는 점에서, 참조 포인트들을 사용하는 것이 가능하다. 따라서, 각 프레임의 움직이는 엘리먼트의 깊이 맵 정보와 함께 3D 뼈대 데이터 구조를 사용하여 요구되는 정확한 초 해상도 레벨의 생성을 강제함으로써 상기 시스템은 시간을 따라 삼각측량법 및 포인트들의 3D 추적을 사용할 수 있다. 일 예에서, 작업 흐름은 다음과 같다:
- [0194] M 을 $2D(x,y)$ 픽셀의 R,G,B 행렬이라 하면,
- [0195] n - 프레임 수,
- [0196] A - 배경에 대해 움직이는 엘리먼트,
- [0197] B - 층 같은 음영 - 그레이 스케일 "음영" 피겨의 2D 행렬,
- [0198] Q - 정의된 임계값을 갖는 각 프레임의 피쳐 포인트들,
- [0199] T - 궤도 (Q 의 2D 포인트 위치 벡터),
- [0200] δ - Q 에 대한 T 의 전이 함수,
- [0201] K - Q 가 갖는 프레임 수 (= T 의 길이),
- [0202] Z - 3D 외삽
- [0203] 따라서, 입력은 n 개의 프레임을 갖는 " M "이다. 시스템은 본 명세서의 다른 부분에서 설명된 바와 같이, 움직이는 엘리먼트를 식별하고, 3D 뼈대의 A 에의 2D 프로젝트 B 를 설정하는 과정 $F(b) \leftarrow G(A)$ 를 수행한다. 정렬은 초기에 정의된 임계값이며, 변경될 수 있다.

- [0204] 상기 과정은 Q를 검색하고 추적함으로써 계속되고, 따라서 각 Q에 대해 T를 생성한다. 여기서 δ 는 $\delta(f(a,b)i, \text{ and } q_{i+1})$ 는 궤도 벡터의 함수이다. 추적은 이미지 a 및 음영 b 상의 프레임 i에서 피쳐 Qj의 위치를 파악하여, 그 위치를 프레임 i+1에 부가하고, k개의 프레임에 대해 이 과정을 수행하게 된다. 추적에 의해 I+1의 새로운 포인트를 이미지 B에 부가하게 되고, 프레임 i+1 상의 새로운 정보는 이미지 A에 따라 이미지 B를 움직이게 할 수 있을 것이다. 따라서, B의 다리는 A의 다리를 따르게 될 것이다. 그리고 나서, 상기 과정은 모든 새로운 프레임에 정확하게 위치된 새로운 무한개의 Q를 수신한다. 각 T에 대해 우리는 Z를 외삽할 것이다, 출력은 움직이는 엘리먼트의 정확한 초해상도가 될 것이다.
- [0205] 그리고 나서, Z 차원은 t 및 t+1 등의 엘리먼트를 둘러싸는 4D 행렬로부터의 참조 포인트를 사용하여 외삽될 수 있다. 이것은 카메라의 배경에 대한 움직임 및 초점에 대해 수행될 수 있다. 참조 포인트로부터의 광으로부터 다음을 계산할 수 있다.
- [0206] 3D 데이터 구조에 대한 2D 변환,
- [0207] 픽셀의 2D 데이터 구조에 대한 3D 변환 또는 Z의 삼각측량을 사용하거나또는 픽셀의 3D 위치를 추적하기 위한 피쳐 포인트, 또는
- [0208] 단일 카메라의 비디오 클립을 형성하는, 움직이는 엘리먼트의 정확한 초해상도 깊이 맵을 생성하기 위한 피쳐 포인트
- [0209] 깊이 외삽 과정은 예컨대 다음과 같이 수행될 수 있다.
- [0210] {a,b,c,...}가 우리가 3D 좌표를 알고 있는 3D 참조 포인트라고 가정한다.
- [0211] t는 시간 't'에서의 (우리가 3D 포인트를 찾고자 하는) 엘리먼트 상의 3D 포인트(픽셀 또는 피쳐)이고,
- [0212] 't+1'은 시간 't+1'에서의 (마찬가지로, 우리가 3D 포인트를 찾고자 하는)동일한 3D 포인트이다.
- [0213] 참조 포인트에서 t 및 t+1까지의 프로젝팅 광은 삼각형 [t+1, t,a], [t+1, t,b], [t+1, t,c],...을 생성한다.
- [0214] 모든 삼각형에서, 우리는 그 참조 포인트의 3D 좌표를 알고 있다. 우리는 6개의 미지수(모든 삼각형은 미지수 {t,t+1}로 구성되고, t,t+1은 3D 포인트이다).
- [0215] 움직이는 엘리먼트의 3D 구조 맵의 추출은 시스템으로 하여금 엘리먼트의 3D 모델을 완전히 재구성할 수 있도록 하고, 서로 다른 시퀀스들 사이의 3D 구조를 재생하고, 서로 다른 시퀀스들 간의 영상의 차이점을 처리할 수 있도록 해준다.
- [0216] 3D 재구성
- [0217] 다음은 3D 뼈대의 3D 깊이 맵으로의 프로젝션 및 깊이 맵의 3D SP 공간으로의 프로젝션을 사용하여 엘리먼트의 모델을 생성하는 방법에 관한 것으로서, 맞는 뼈대 기관들을 깊이 맵 정보에 부가하고, 종합적으로 그것을 복사하고 그 움직임을 캡처하게 된다. 이하의 설명과 같이, 마지막 과정은 재구성된 뼈대 상의 엘리먼트의 텍스처와 중첩될 것이고, 재구성 과정을 완료시킬 것이다.
- [0218] 본 방법은 프레임에서 움직이는 엘리먼트의 3D 모델의 생성을 강제할 수 있도록 해준다.
- [0219] 본 알고리즘은 개념적으로 공간을 기반으로 한다. 한편으로는, 3D 정보를 2D 이미지 평면에 프로젝션하는 것도 3D 정보의 외삽을 가능케 하지만, 다른 한편으로는 공간 차원 기반의 알고리즘을 사용하면 3D 깊이 맵을 갖는 2D 세계를 공간 기반이 3D 세계로 프로젝트 하게 된다.
- [0220] 위에서 언급된 움직이는 엘리먼트 주위에서 선택된 NxMxZ 윈도우는, 사실상 (4D 행렬로 변화되는) 3D 행렬, 새로운 (XYZ) 축 필드 "f"이고, 이 필드에서 사용자는 미리 정의된 내부 뼈대 또는 3D 피겨(튜브, 조인트 등)의 일부를 부가할 수 있다.
- [0221] 깊이 외삽 과정은 또한 각 픽셀의 식별 및 프레임 간의 피쳐 움직임을 포함하여, 시간에 따라 픽셀 및 피쳐의 2D 움직임 플로우를 생성한다. 시스템은 2D 움직임 플로우를 3D 움직임 플로우로 변환한다.
- [0222] 3D 깊이 구조 맵 S(o), 또는 3D 궤도 집합, 또는 이하의 설명될 바와 같이 깊이 외삽의 과정 중에, 시스템은 각 프레임(및 통일된 프레임들)의 팩터 D를 시간에 따른 내부 부가된 뼈대(그 자신의 팩터 D')와 일치시키기 위한 재구성 알고리즘을 사용하고, 공간과 시간 사이의 변환인 전술한 제3 팩터를 사용하여 공간 및 시간 차원에 대

해 3D 행렬을 4D 행렬로 변환하는 내부 뼈대의 일부를 정의 및 구성한다.

- [0223] 깊이 외삽 및 재구성 과정은 직관적으로 말하면 3D 구조의 2D 반사 영상(reflection)이 2D 이미지 행렬 밑에 놓인 층상화된 장치이다. 3D 구조 자체는 프레임의 엘리먼트를 종합적으로 3D 재구성하는데 사용되는 3D 구조의 2D 반사 영상 밑에 놓이게 된다. 이 3층 구조는 도 7를 참조하여 설명된다. 공간 차원 하에서의 작업은 이하의 설명과 같이 3D 구조 및 텍스처를 재구성할 수 있고, 출력이 프레임의 엘리먼트의 움직임 캡처인 엘리먼트의 움직임을 3D 모델로서 보존할 수도 있다. 또는 그러한 이유 때문에 출력은 특정 2D 프로젝션일 수 있다.
- [0224] 움직이는 엘리먼트의 재구성 과정에 대해, 시스템은 완전한 3D 구조 깊이 맵을 생성하긴, 움직이는 엘리먼트의 완전한 3D 구조 깊이 맵을 수신할 수 있다. 정적 배경은 앞서 설명된 바와 같이 움직이는 엘리먼트로부터 분리되어 모델링된다.
- [0225] 본 실시예는 움직이는 사람과 같은 2D 몸체를, 원래 2D 또는 부분적으로 3D인 이미지에서, 3D 모델로 완전한 3D 초해상도 재구성을 가능케한다. 상기 과정은 또한 내부 3D 뼈대의 베이스 상에 구성된 3D 구조 텍스처 및 움직임을 캡처하는 과정도 포함한다. 상기 시스템은 3D로 재구성된 모델, 예컨대 애니메이션, 움직임 캡처, 실시간 모델링 등에 대한 무한대로 조작용 가능케 한다. 도 12는 별개의 뼈대 모델 엘리먼트들로부터 구성된 완전한 해부학적 모델을 나타낸다.
- [0226] 프로젝션은 이미지 및 몸체의 3D 깊이 맵을 사용하여 수행되며, 배경에 관해서는 엘리먼트 주위의 4D 행렬을 사용하고, 참조 포인트를 사용한다.
- [0227] 도 8을 참조하여 위에서 설명된 바와 같이, 엘리먼트들은 3D 뼈대 또는 그 부분의 자동 식별 또는 수동 식별하고, 그것을 3D 깊이 맵에 부가함으로써 식별된다. 시스템은 3D 데이터 구조를 프로젝션한다. 즉 3D 뼈대를 3D 깊이 맵으로 프로젝션한다.
- [0228] 프레임에 대한 엘리먼트의 3D 움직임의 추적은 픽셀들 및 궤도들의 DTM 광학 플로우를 기반으로 한다. 추적에 의해 엘리먼트의 3D 형성에 대한 가능한 많은 3D 정보를 얻을 수 있고, 뼈대의 3D 구조를 깊이 맵에 삽입함으로써 3D 구성을 획득할 수 있다. 뼈대의 3D 구조는 시간에 대해 획득한 3D 정보를 보존하여, 이미지의 엘리먼트를 가능한 한 정확하게 제공하기 위한 뼈대의 3D 구조를 설계한다.
- [0229] 예를 들어, 이미지의 구조가 사람이라는 것을 파악한 후에는, 시스템은 설정된 정책에 따라, 예컨대, 눈, 코 등에 대한 3D 및 시각 정보의 수신을 기대한다. SP가 예컨대, 머리의 특정 영역인 눈 및 눈썹을 기대하므로, SP가 3D 피겨에 대한 이 정보를 분석하는 것이 더 쉽고 빠르다.
- [0230] 움직이는 오브젝트의 기관들(organ) 사이의 위치 및 거리를 계산함으로써 SP가 오브젝트의 다른 기관들의 관계들을 정확히 평가하는 것을 보조하고, 3D 피겨의 이 정보를 조정하게 된다.
- [0231] 이미지 처리에서 공간 차원을 사용함으로써, 모든 서로 다른 프레임이 더 낮은 3D 및 시각 정보를 3D 엘리먼트 및 피겨에 제공한다는 점에 관해 시간 차원에 대해 엘리먼트의 움직임에 추가적인 가치가 존재한다.
- [0232] 재구성 과정과 병렬적으로, 본 실시예는 엘리먼트의 움직임을 포착하기 위한 움직임 캡처 도구로 사용될 수 있으며, 사용자가 2D 이미지로서 뿐만 아니라, 텍스처를 갖는 3D 모델로서도 프레임의 엘리먼트의 움직임을 캡처할 수 있도록 해준다.
- [0233] 재구성의 과정은 다음 방식으로 수행될 수 있다.
- [0234] 초미리 설정은 4D 행렬이고, 입력은 DTM일 것이다.
- [0235] DTM은 외부 알고리즘으로부터 제공될 수 있다. DTM은 또한 본 실시예의 깊이 외삽 알고리즘으로부터 제공될 수 있다. 처리 시간의 관점에서, 모델링 과정은 깊이 외삽 과정에 병렬적이고, 여기서 직관적으로 말하면, 행렬들은 서로의 밑에 위치하며, 제1 행렬은 이미지 2D 행렬이고, 그 아래에는 (3D 구조의) 2D 프로젝션 행렬이 배치되고, 그 아래에는 3D 데이터 구조가 배치된다. 입력은 또한 픽셀의 2D 추적, 특히 움직임이 추적될 수 있도록 프레임을 따라 세팅된 피쳐 포인트들을 기반으로 한 3D 궤도들도 포함한다. 피쳐 포인트는 추적하기 용이한 컬러 또는 다른 속성들을 기반으로 할 수 있다.
- [0236] 궤도들은 DTM으로 변환되고, 시스템은 그것들을 프레임의 픽셀 및 피쳐 포인트들의 3D 위치를 표시하는 3D 궤도들로 변환한다.
- [0237] 시스템은 3D 뼈대의 프로젝션과 입력 깊이 맵들 사이의 제한값을 설정하여, 엘리먼트의 식별된 기관들을 갖는

3D 프로젝션으로부터 생성된 모든 프레임에서 무한한 새로운 추적 포인트들로 시간 축을 따라 정확한 추적을 하게 된다. 이에 따라, 3D 물체의 숨겨진 부분들의 위치 및 3D 공간에서의 위치를 알 수 있게 된다. 시스템은 3D 뼈대 데이터 구조를 사용하여 포인트들의 3D 추적을 수행하여, 움직이는 엘리먼트의 정확한 초해상도 3D 모델의 생성을 강제한다.

- [0238] 작업 플로우는 다음과 같을 수 있다. 변수들은 다음과 같이 설정된다.
- [0239] E_{3d} - DTM의 시퀀스- 움직이는 엘리먼트의 (x,y,z) 의 3D 행렬,
- [0240] n - 프레임 수,
- [0241] S_{3d} - 3D 뼈대
- [0242] Q_{3d} - 모든 프레임에서의 3D 피쳐 포인트
- [0243] T_{3d} - 궤도들(Q_{3d} 의 3D 포인트 위치)
- [0244] δ_{3d} - Q_{3d} 상의 T_{3d} 의 전이 함수.
- [0245] K - 각 Q_{3d} 가 갖는 프레임 수(= T_{3d} 의 길이)
- [0246] 모델 - 재구성된 3D 모델
- [0247] 시스템은 3D 뼈대 S 가 DTM-E로 정렬되도록 $F(E) \leftarrow G(S)$ 를 정렬한다. 시스템은 E_t 상의 Q_{3d} 를 다음 DTM E_{t+1} 까지 3D 추적하기 위해 T_{3d} 를 사용한다. 여기서 δ_{3d} 는 $\delta_{3d}(f(s,e)_i, q_{3d} i+1)$, 궤도 벡터, 뼈대 S 및 DTM E 의 프레임 i 에서의 피쳐 Q_{3d} 의 위치의 함수이고, 프레임 $i+1$ 에서의 위치를 부가하고, 각 Q_{3d} 에 대한 (k) 프레임들에 대해서도 마찬가지로 수행한다. 각 프레임에 대해, $I+1$ 의 새로운 포인트를 S 에 부가하고, 프레임 $i+1$ 에 대한 새로운 정보는 E 의 새로운 위치(예컨대, S 의 다리는 E 의 다리와 정렬될 것이다)에 따라 S 의 정렬을 가능케한다.
- [0248] 결과적으로 모든 새로운 DTM에 정확하게 위치하는 새로운 무한개의 Q_{3d} 를 수신할 수 있게 된다. 팩터 D 및 D' 는 시스템이 E_{3d} 의 집합적인 정보의 형성에 따라, t 및 $t+1$ (기타 등등)의 엘리먼트 및 뼈대를 둘러싸는 4D 행렬에 S 의 형성을 변경시킬 수 있도록 해준다. 시스템은 모델로부터 다음 프레임의 어디에 사지(limb) 및 다른 엘리먼트들이 나타날지를 유추한다. D 는 3D 엘리먼트의 형태로 S_{3d} 를 종합적으로 복제하는 복잡한 수학적 구조에서의 키 팩터이다. 시스템은 3D 뼈대를 새로운 데이터 구조로 변환하고, 수집된 형태를 움직이는 엘리먼트의 정확한 초해상도 3D 모델 복제를 생성하기 위한 픽셀 또는 피쳐 포인트들의 3D 위치를 추적함으로써 3D 뼈대 데이터 구조의 DTM의 시퀀스로 수집 및 저장한다.
- [0249] $\{a,b,c,\dots\}$ 가 $E_{3d} i$ 상의 3D 좌표를 갖는 포인트들이라 가정한다. $S_{3d} i \{a',b',c'\}$ 상의 해당 포인트들을 $E_{3d} i$ 의 포인트들에 부가하기 위한 평가가 행해진다. 시스템은 4D 행렬의 서브 공간의 팩터 D' 를 $E_{3d} i$ 의 팩터 D 에 정렬시키고, $S_{3d} i$ 를 통합된 유닛으로 정렬시키고, 또한 S_{3d} 를 미리 정의된 미니어처 4D 행렬들로 분산시키며, 상기 4D 행렬들은 각각 독립 4D 서브 공간으로서 하나의 D' 팩터를 가지며, $S_{3d} i$ 의 형성을 재구성하여 $E_{3d} I$ 를 형성시키고, 그리고 나서, $i+1\dots$ 등에 대해서도 마찬가지로 작업을 수행한다. 출력은 엘리먼트의 정확한 초해상도 재구성(이하의 설명과 같은 형태 및 텍스처), 및 움직이는 엘리먼트의 3D 움직임 캡처가 된다.
- [0250] 이하의 추가로 설명될 바와 같이, 텍스처 중첩은 모델링 과정의 일부이다.
- [0251] 전술한 제한은 엘리먼트의 3D 모델을 완전하게 재구성할 수 있도록 해준다. 그것은 또한 서로 다른 시퀀스들 간의 3D 구조를 회복하고, 서로 다른 시퀀스들 간의 영상의 차이점을 처리할 수 있도록 해준다. 3D 구조 기반의 정렬의 속성을 이용함으로써, 우리는 다른 비디오 시퀀스들 사이와 같이 극도로 어려운 상황에서의 정보를 매칭시킬 수 있으며, 스케일(줌)에 차이점이 있는 상황, 서로 다른 센싱 양상(IR 및 가시광선 카메라)의 경우 등에서도 매칭을 수행할 수 있다.
- [0252] S_0 가 엘리먼트가 풍선에 바람을 불어넣는 경우와 같이 시간 및 공간에 따라 속도를 다르게 변화시키는 동안 촬영된 경우에, 3개의 선택사항이 독립 또는 통합된 해결책으로 존재한다. 시스템은 상기 풍선을 시간에 따라 속도를 변화시킴으로써 모델링할 수 있고, 프레임 또는 일련의 프레임, 또는 이상하게 움직이는 오브젝트를 문제가 있는 설정으로 표시하여, 그 문제를 시간 영역에 남겨둘 수 있다. 선택적으로, 사용자는 자동 시스템을 보조하여 프레임의 3D 피쳐를 정의할 수 있도록 하며, 따라서 어떤 3D 정보 및 시각 정보를 사용할 것인지를 알려준다.

- [0253] 과정의 마지막에, 엘리먼트의 움직임은 상기 모델은 그 3D 움직임 캡처와 관련시킬 수 있으며, 위 이미지 처리에서 공간 차원을 사용하여 오브젝트의 보다 정확한 3D 모델에 대한 재구성 과정에 대한 보다 많은 3D 및 시각 정보를 제공한다는 점에서 모든 프레임에 가치를 추가할 수 있다.
- [0254] 이러한 개별 튜브, 볼 및 다른 뼈대 모델 엘리먼트들로부터 시간 차원에 대해 완전한 3D 피거를 구성하게 되면, 이들이 상호 움직임에 대한 정보 - 서로 어떻게 결합되어 있는지, 어떻게 같이 움직이는지 및 피거의 추가적인 애니메이션에서 어떻게 보조되는지를 알 수 있게 된다.
- [0255] 일단 모델링이 완료되면, 원래 영상 및 배경으로부터 별개의 피거로써 보관될 수 있으며, 추후의 애니메이션에 사용될 수 있다. 그 원래 움직임은 움직임 캡처에 사용될 수 있다. 동일판 피거에 다른 시간 또는 위치로부터의 더 많은 시각정보를 부가할 수 있고, 사용자 행위에 따라 다각형 내부 뼈대 상에 컴퓨터로 생성된 이미지와 같은 새로운 3D 피거로 변경될 수도 있다.
- [0256] 또한, 피거는 배경에 대해 독립적이 되어, 추가적인 애니메이션에 사용될 수 있다. 오브젝트가 다른 시간 및 위치에서 촬영되는 경우에, SP는 다른 시간 또는 위치로부터 획득한 정보를 결합할 수 있다. 예를 들어, 시각 정보는 다른 시간 또는 위치로부터 획득될 수 있고, 계산된 정보는 피거의 주 3D 모델에 부가된다. 오브젝트의 3D 구조를 사용하여, 우리는 촬영된 배경에 영향을 받지 않고, 오브젝트 자체의 레벨에서, 애니메이션, 모방 얼굴을 생성하고, 음성을 부가하는 등의 작업을 수행할 수 있다.
- [0257] 시스템은 또한 움직이는 엘리먼트 내의 움직임을 캡처하는 능력을 사용하여, 얼굴의 모방과 같이 몸체 전체 애니메이션 또는 그 일부의 움직임 캡처를 사용하여 이미 존재하는 3D 모델을 애니메이션화할 수 있다.
- [0258] 상술한 예의 풍선에 바람을 불어넣는 경우와 같이, 시간 및 공간에 대해 상대적인 속도를 변화시키는 엘리먼트가 존재하는 경우에, 사용자의 보조가 요구될 수 있다. 사용자는 시스템에 프레임에 3D 피거를 정의하여, 어떤 3D 정보를 사용할 것인지 및 시간 기반 시퀀스로 애니메이션의 부가없이 이 피거를 남겨놓을 것인지를 표시하여 줄 것을 요청한다. 후자의 경우에, 이 엘리먼트는 정규 이미지 처리 툴을 사용하여 편집될 수 있다. 이번 예와 및 이전 예의 유일한 차이점은 이번 예의 경우에는 공간 기반 3D 오브젝트가 아니라 시간 기반 3D 오브젝트가 남는다는 점이다.
- [0259] 이미지 처리 툴은 사용자가 다른 시간 및 위치의 배경들을 함께 결합하고, 이미지의 왜곡을 정정하며, 엘리먼트를 제거하고 입력에 의해 생성된 정보를 기반으로 새로운 엘리먼트들을 생성하고, 또한 3D 컴퓨터로 생성된 피거를 생성할 수 있도록 하긴, 다른 3D 컴퓨터 애니메이션 프로그램으로부터 컴퓨터로 생성된 피거를 입력할 수 있도록 해준다.
- [0260] 시각 정보
- [0261] 공간 기반 3차원 모델에서 3차원 위치를 수신한 후에, 결정된 해상도에 따라, 각 포인트는 시각 정보 계층을 수신하며, 컬러 및 휘도 값은 상기 이미지의 디지털 정보에 기록된다.
- [0262] 고려할 수개의 시각 파라미터가 존재한다. 촬영된 이미지의 해상도, 와 비교된 모델의 해상도, 각 정보의 구형 정보, 및 다른 카메라 또는 다른 클립으로부터의 시각 정보의 다른 품질 등이 그것이다.
- [0263] 서로 다른 이미지 해상도에 대해서는 2가지 경우가 존재할 수 있다.
- [0264] 첫번째 경우에, 이미지 해상도는 결정된 해상도 F보다 높고, 3D 모델의 모든 픽셀에 대해 요구되는 정보의 양보다 더 많은 정보가 존재한다. 예를 들어, 촬영된 이미지가 인치당 픽셀 수의 관점에서 5배 더 클 경우, 시스템은 5개 픽셀마다 시각 정보를 합하여 평균을 내고, 새로 계산된 값으로 3D 모델의 새로운 픽셀을 생성한다.
- [0265] 두번째 경우는, 3D 모델의 해상도가 촬영된 이미지의 해상도 보다 큰 경우이다. 비디오 시퀀스를 사용하여, 모든 프레임은 프레임 내에 텍스처 픽셀을 생성하고, 카메라가 약간 움직이는 경우에는, 픽셀들은 이웃 3D 포인트들을 촬영하여, 통일된 모델에 대해 이미지의 총 픽셀 양보다 더 많은 시각 정보를 수집할 수 있게 된다. 그러한 경우는 예컨대, 먼 거리에서 이미지를 촬영하거나, 또는 디지털 줌을 사용하는 등의 경우에 발생할 수 있다. 이 경우에, 시스템은 시간 차원을 따라 이웃 픽셀들로부터 각 픽셀에 대한 정보를 추출한다. 여기서 각 픽셀에 대한 시각 정보의 다중 계층의 키 엘리먼트가 중요하며, 이하 이에 대해 논의하고자 한다.
- [0266] 새로운 픽셀들이 생성되어 시스템 해상도의 레벨에서 모델의 표면에 중첩된다. 각 새로운 픽셀은 이제 3D 공간

기반 모델에서 3차원 위치를 가지며, 실제 세계가 완전 360 각도로부터 관측될 수 있다.

- [0267] 일반적으로, 개별 픽셀들은 360도로부터 관측되지 않는다. 예를 들어, 벽의 한 점은 180도에서 관찰될 수 있고 (벽의 뒷면은 3D 위치에 따라 다른 픽셀에서 다른 정보를 갖는다), 모퉁이의 돌은 270도에서 관찰될 수 있다.
- [0268] 도 15를 참조하면, 특정 방향으로부터 시각 정보를 제공하는 촬영된 3D 이미지가 도시되어 있다. 각 촬영된 이미지는 특정 방향으로부터 시각 정보를 제공한다. SP가 특정 방향에서만 픽셀에 대한 시각 정보를 수신하는 경우에, 그것은 픽셀을 단조롭게 하고, 180도에서 그것을 관찰할 수 있도록 해준다. 이 경우는 측면 방향에서 이 픽셀을 관찰하는 경우에는 시각적 품질에서 약간의 왜곡을 발생시킨다.
- [0269] 본 발명의 바람직한 실시예는 공간 차원의 3D 위치에 관해 각 픽셀에 시각 정보의 다중 계층을 통일하는 반구 픽셀 구조를 제공한다. 무한개의 픽셀을 부가할 수 있고, 시각 품질의 관점에서 우리는 초해상도를 생성하고 있는 것이다.
- [0270] 초해상도 역시 깊이 포인트의 수와 관련되며, 초해상도 3D 포인트를 생성하는 통일된 모델을 수집하는 것이 가능하다. 깊이가 포인트들은 가장 정확한 방식으로 표면의 변형을 가능케 한다.
- [0271] 3차원 시각 정보의 관점에서, 시스템에 제공되는 시각 정보에 포함되는 각도가 많을수록, 소프트웨어가 모든 필요한 방향에서 픽셀을 관찰하는 능력을 이 더 좋고, 이 픽셀에 대한 더 좋은 구형 정보를 제공할 수 있으며, 모든 필요한 방향에서 픽셀을 관측하는 능력을 제공한다. 따라서 다중 이미지가 픽셀 주변의 다양한 각도에서 획득될 수 있다. 모든 픽셀은 각 클립을 따라 많은 프레임으로 촬영될 수 있다. 이 모든 정보가 필요한 것은 아니며, 이것은 이 모든 정보가 동일한 레벨의 품질을 갖지 않는다는 것을 의미한다. 그러한, 각 픽셀에 대한 다중 계층 시각 정보는 이미지 처리에 요구되는 계산을 낮추고, 이미지 품질을 개선하는데 도움이 될 수 있다.
- [0272] 시각 품질의 관점에서, 모든 픽셀은 각 클립을 따라 많은 프레임으로 촬영될 수 있다. 그럼에도 불구하고, 모든 정보가 동일한 레벨의 품질을 갖지는 아니하므로, 이 모든 정보가 요구되는 것은 아니다. 저품질 정보를 폐기하는 것이 이미지 처리에 요구되는 계산을 줄이는데 도움이 되는 포인트이지만, 모든 정보를 사용하여 불량 이미지 품질, 조명, 카메라 해상도 등에 기인하는 이미지 품질을 개선하는 것이 바람직하다.
- [0273] 시스템은 품질의 수준 Q를 생성하며, 모든 새로운 프레임으로부터 정보를 결정하는 정보의 각 새로운 계층은 그 시각 정보 및 해상도의 품질에 대해 검사된다. 시각 정보는 2개의 팩터에 의해 품질이 매겨지는데, 하나는 시간 차원에서의 이미지 품질이고, 다른 하나는 공간 차원에서의 이미지 품질이다.
- [0274] 예를 들어, SP는 빌딩 내부의 동일한 위치에서 촬영되며, 방의 내부 및 외부 정원을 촬영하기 위해 다른 개구를 사용하는 2개의 클립을 수신할 수 있다.
- [0275] 제1 클립에서, 카메라는 고 노출의 개구를 사용하고, 이것은 정원의 외부 부분들이 과노출되어 있고, 햇빛에 그 올리거나 지나치게 밝게 보이는 경우 카메라가 이미지의 내부 부분들에 대한 양호한 시각 정보를 수신할 수 있도록 해준다.
- [0276] 제2 클립에서, 카메라는 저 노출 개구를 사용하고, 이것은 이미지의 내부 부분들에 대한 매우 어두운 시각 정보를 생성하지만, 이미지의 외부 부분들은 매우 균형잡혀 있고, 잘 노출되게 된다.
- [0277] 이들 클립 각각은 독립 유닛으로서 밸런싱되지 않은 것은 당연하며, 이들 각각의 히스토그램은 밸런싱되지 않은 결과를 나타낼 것이다.
- [0278] 그러나, 모든 프레임이 공간 차원에서 별개로 검사되고, 배경들이 전체적으로 검사되면, 시스템이 시간 차원으로부터 변형되고, 시스템이 새로운 시각 정보를 수신하면서, 다음과 같이 2개의 팩터를 기반으로 클립을 체크한다.
- [0279] 제1 팩터는 시간 차원을 기반으로 하고, 별개 유닛으로서 모든 프레임에 대한 히스토그램을 결정하고, F에 대해 그 품질을 결정하며,
- [0280] 제2 팩터는 공간 차원을 기반으로 하고, 이 때 F에서 Q가 불량인 경우에도, 높은 품질을 달성하기 위해 이미 편집된 이미지들은 프레임의 임의의 영역을 가리킨다.
- [0281] 도 14에 관해 설명된 바와 같이, 시스템은 이웃하는 프레임 픽셀과 무관하고, 이웃 픽셀의 F의 불량 Q와 관련하여, SP에 요구되는 특정 부분들에 대해 더 나은 시각 품질을 위해 새로운 클립을 검색한다. 시스템은 동일한 이미지에서 동일한 시간에 다른 개구를 사용하여 촬영된 것처럼, 설명된 예에서, 최상의 품질로 내부 방 뿐만

아니라 외부 정원까지 나타내는 매우 노광이 잘된 이미지를 제공하는 양호하게 밸런싱된 이미지를 생성하게 된다.

- [0282] 시스템은 임의의 최소 레벨 Q 까지 이미지 정보를 고려하며, 이것은 이미지가 2개 팩터 모두에 대해 상기 최소 값보다 낮은 경우에는 이 정보에 사용할 것이 없거나, 픽셀 텍스처의 현재 값에 부가할 값이 없다는 것을 의미한다.
- [0283] 새로운 정보의 부가 과정은 $Q \in SP$ 를 기반으로 한다.
- [0284] Q가 높을 수록, 정보가 픽셀 값에서 갖는 참여값이 더 높으며, Q가 낮을 수록, 픽셀 값에서의 참여값이 더 낮다.
- [0285] 시스템은 2개의 클립 모두로부터의 정보를 통일하고, 방의 내부 및 외부 정원 모두에 대해 밸런싱되고 양호하게 노출된 뷰(view)를 생성한다.
- [0286] 시스템은 품질 Q에 대한 임계치를 설정할 수 있고, 그에 따라 시각 정보를 폐기할 수 있다.
- [0287] 이미지 처리는 또한 이미지를 밸런싱시키고, 인접한 이미지들 간의 균일성을 달성하기 위해 기준 카메라 제어 유닛(CCU)에 사용되는 것과 같은 처리 방법을 포함할 수 있다.
- [0288] 뷰잉(Viewing)
- [0289] 시각 정보를 포함하는 구성된 공간 기반 3D 모델은 SP에 제공된 모든 이미지의 캡처된 집합적인 결과이다.
- [0290] 집합적 필드에서의 임의의 포인트는 종래의 뷰잉 방법 중 임의의 것을 사용하여 임의의 뷰 포인트로부터 시각화될 수 있다. 다음 예를 살펴본다.
- [0291] 본 발명의 일 실시예에서, 가상 카메라들이 2개의 근접한 렌즈들의 시야각이 수평축에 대해 렌즈 측면에 놓인 2개의 근접한 렌즈들의 시야각에 의해 상당히 중첩되도록 정렬된다. 결과적으로, 입체적인(stereoscopic) 이미지가 생성될 수 있다.
- [0292] 바람직한 실시예는 전체 시간 기반 시퀀스, 라이브 시퀀스, 비선형 출력, 입체/3D 구형 이미징 등의 생성을 용이하게 한다.
- [0293] 입체 이미지를 제공하기 위한 바람직한 실시예에서, 가상 카메라들이 특정 설정으로 정렬되며, 여기서 임의의 렌즈의 시야각은 렌즈를 둘러싸는 모든 근접한 렌즈들의 시야각에 의해 원하는 만큼 중첩되며, 집합적 시야각은 완전히 환형의 이미지 집합을 포함하고, 여기서 각 시야각 내의 임의의 포인트는 입체 구형 이미징을 생성하기 위한 적어도 2개의 가상 렌즈에 의해 또는 2D 구형 이미징을 생성하기 위한 적어도 하나의 가상 렌즈로부터, 또는 임의의 뷰 포인트로부터 3D 구형 이미징을 생성하기 위한 적어도 2개의 가상 렌즈들로부터 캡처된다.
- [0294] 결과적으로, 단일 카메라에 의해 촬영된 장면을 뷰잉하기 위한 입체 데이터가 가용(available)하게 된다.
- [0295] SP에 의해 생성된 이미지는 다양한 포맷, 예컨대 정지 영상, 비디오, 입체 뷰잉, 가상 현실 등으로 시청자(viewer)에게 디스플레이될 수 있다. 형성된 이미지는 TV 또는 컴퓨터 스크린과 같은 평면 스크린 상에 디스플레이되거나, 또는 가상 현실 헤드셋과 같은 가상 현실용 디스플레이 장치에 의해 디스플레이될 수 있으며, 여기서 디스플레이되는 이미지의 일부는 사용자의 뷰포인트에 따라 변경된다. 시청자를 둘러싸므로써, 가상 현실 디스플레이용의 적절한 수단에 의해 수직 및 수평 모두 360도로 시청자에게 디스플레이된 이미지의 3D 깊이를 인지하면서, 위,아래를 포함하는 주위 모든 방향으로부터 볼 수 있는 능력을 제공한다.
- [0296] 종래의 가상 현실 분야에 공지된 수단을 사용하여, 가상 현실의 시각적 선형 및 비선형 정보가 사용자에게 제공된다. 그러한 수단은 시청자의 머리 위치를 검출하기 위한 센서를 갖는 헤드셋, 또는 손의 위치를 검출하기 위한 가상 장갑 또는 뷰잉 소프트웨어 분야에 공지된 임의의 장치일 수 있다.
- [0297] TV 또는 컴퓨터 스크린과 같은 평면 스크린 상에 디스플레이 하기 위해, 사용자의 뷰잉 파라미터들은 이 용도에 사용하기 위해 프로그래밍된 사용자가 갖고 있는 포인팅 장치(예 : 마우스 또는 조이스틱)으로부터 획득한다. 시스템은 예컨대 본 발명의 움직임 캡처 능력 또는 임의의 외부 장치에 의한 움직임 캡처를 사용하여 사용자 자신의 움직임을 수집할 수 있다.

- [0298] 사용자가 가상 현실 헤드셋을 쓴 상태에서 실제로 머리를 돌리거나 또는 컴퓨터 장치에 연결된 사용자가 잡고 있는 포인팅 장치에 의해, 특정 시각을 선택하면, 디스플레이 시스템에 의해 뷰잉 파라미터들이 검출 및 수신된다. 시청자의 뷰잉 파라미터들은 시청자의 뷰잉 방향 및 시청자가 놓인 수평면을 포함한다. 이들 파라미터들에 따라, 시청자의 시야각은 시청자를 둘러싸는 좌표에 의해 결정되며, 이미지는 뷰잉 수단으로 프로젝션된다.
- [0299] 카메라의 종류
- [0300] 본 발명은 이미지를 캡처하기 위해 사용되는 카메라의 종류나 SP에 입력되는 이미지의 시퀀스에 의해 제한을 받지 않는다. 카메라는 디지털 또는 아날로그 비디오 카메라 분야의 임의의 공지 기술로부터 선택될 수 있다. 카메라는 또한 비디지털일 수 있으며, 이 경우 공지된 기술을 사용하여 이미지를 디지털 포맷으로 변환할 수 있다.
- [0301] 바람직한 실시예에서, 디지털 이미지는 저장 및 공간 기반 3D 모델로의 변환에 앞서 품질을 개선하도록 조작될 수 있다.
- [0302] 어플리케이션
- [0303] 도 10을 참조하면, 본 발명의 다른 어플리케이션을 나타내는 풍선 차트가 도시되어 있다.
- [0304] 바람직한 실시예에 따라, 상술한 바와 같은 완전한 공간 기반 3D 모델이 가용하게 되면, 사용자는 가상 환경 내에서 뷰 포인트로부터 장면을 효과적으로 재촬영할 수 있도록, 원래 시퀀스에는 카메라가 배치되어 있지 않은 위치에 가상 카메라를 배치할 수 있다. 또한, 이것은 실시간으로 수행될 수 있다. 예를 들어, 농구 경기에서, 가상 카메라는 실제 카메라가 배치되어 있지 않은 뷰 포인트로부터 경기를 촬영할 수 있도록 배치될 수 있다. 필요한 것은 사전에 모델링된 경기장과 각 선수들이다. 사실, 선택적으로, 모델링은 방송 전에 미리 수행될 수 있다.
- [0305] 바람직한 실시예에 따라서, 위 모델링을 사용하여, 시퀀스로부터 캡처된 각 피겨는 사용자에게 의해 다시 애니메이션화될 수 있으며, 상기 사용자는 또한 예컨대 외부 소스로부터의 움직임 캡처 또는 SP의 움직임 캡처를 사용할 수 있으며, 원래 클립에서 피겨들의 움직임을 변경할 수 있다. 즉, 사용자는 원래 촬영된 이미지로부터 모델을 재구성할 수 있고, 실시간으로 피겨의 다른 움직임들을 출력할 수 있다.
- [0306] 바람직한 실시예에 따라서, 사용자는 이미지로부터 원래 피겨를 변형하거나, 또는 그 피겨를 완전히 새로이 조작된 피겨로 대체하는 것도 가능하다.
- [0307] 본 실시예에 따라, 앞서 설명한 기술들을 사용하여, 피겨를 3D 모델로 대체하고, 피겨로 새로운 영화 클립의 생성을 가능케함으로써 촬영된 클립 동안의 피겨의 원래 움직임에 영향을 받지 않고, 새로운 애니메이션이 각 피겨에 주어질 수 있다. 상기 피겨는 또한, 컴퓨터 게임, 콘솔 게임, TV 게임 등에서 사용자에게 의해 실시간으로 조작될 수 있다.
- [0308] 바람직한 실시예는 공지된 기술들을 사용하여, 새로운 조명을 3D 모델에 유도하여, 애니메이션의 장면 또는 비디오 클립의 제조 후에 조명을 추가할 수 있다.
- [0309] 바람직한 실시예는 각 엘리먼트 및 배경의 일부로서, 바람직한 정도로, 경기장에서의 깊이 외삽을 포함한다. 깊이 외삽은 시스템에 입력된 사진 피겨의 깊이 맵 분석을 포함하고, 이것은 다양한 방법에 의해 수행될 수 있으며, 이하 상세히 설명한다.
- [0310] *바람직한 실시예는 이미지 상의 움직임 블러(blur)와 같은 다양한 조작을 사용할 수 있도록 해준다.
- [0311] 여기 설명된 기술들을 사용하여, 장면을 촬영하는 동안 수행될 수 있는 모든 종류의 조작들, 예컨대 초점을 변경하거나, 줌 및 조명을 변경하는 등의 조작들이 3D 가상 경기장에서도 수행될 수 있다.
- [0312] 여기 설명된 기술들을 사용하여, 사용자는 피겨 및 배경으로부터 완전한 움직임 영상을 생성할 수 있다.
- [0313] 여기 설명된 기술들을 사용하여, 모든 피겨들이 실제 이미지 기반의 3D 모델인 3D 공간 기반 모델을 사용함으로써, 사용자는 완전한 컴퓨터 게임(콘솔 게임, TV 게임 등)을 생성할 수 있다

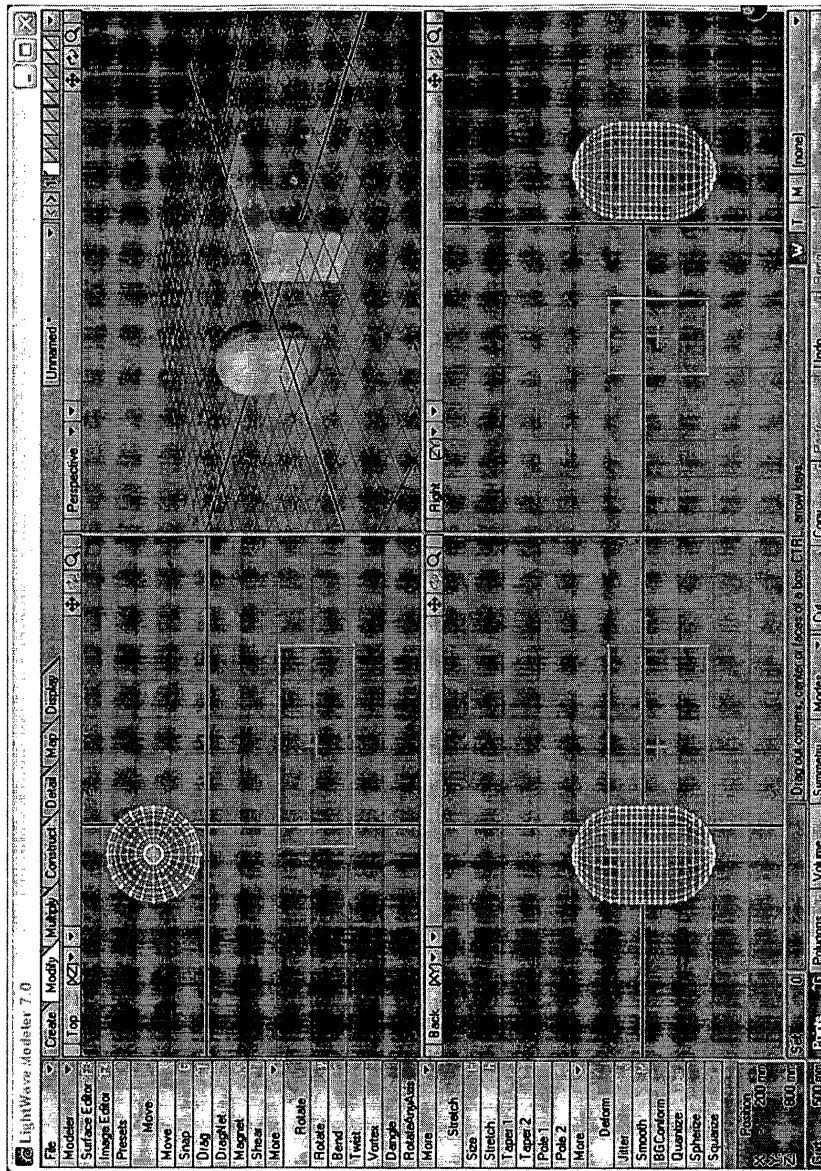
- [0314] 바람직하게는, 컴퓨터로 생성된 이미지들은 3차원 환경 및 3차원 모델들에 추가될 수 있다. 이 이미지들은 모델의 피부를 변경하거나, 추가적인 컴퓨터로 생성된 엘리먼트들을 모델 또는 배경 등에 추가하는 것과 같은 효과를 가질 수 있다.
- [0315] 본 실시예에 따라서, 사용자는 움직임 캡처 단계에서 피겨의 움직임을 재구성하기 위해 시퀀스 내의 각 피겨와 결합된 시간 라인 정보를 사용할 수 있다. 본 기술들은 단일 카메라로부터의 시퀀스 또는 2개 이상의 카메라로부터의 이미지들을 사용하여 수행된다.
- [0316] 여기 설명된 절차에서, 2차원 및 3차원 추적은 시간 기반 클립에서 그 움직임을 기반으로 식별된 임의의 피겨 및 배경에 적용될 수 있다. 추적은 실시간으로 또는 추후에 클립을 다시 애니메이션화하는 과정의 일부로서 수행될 수 있다.
- [0317] 본 발명의 바람직한 실시예에 따라서, 사용자는 또한 공간 기반 3D 환경에서 피겨 또는 배경에 움직이는 또는 정적인 엘리먼트들을 추가할 수 있다.
- [0318] 본 발명의 바람직한 실시예에 따라서, 사용자는 원래는 촬영되지 않은 새로운 경기장을 생성할 수 있다. 예컨대, 사용자는 수개의 서로 다른 환경들을 하나의 통일된 경기장으로 조합하거나, 또는 촬영된 경기장을 컴퓨터로 생성된 합성 경기장과 결합할 수 있다.
- [0319] 본 발명의 바람직한 실시예에 따라서, 사용자는 3D 모델에서의 본 실시예를 사용하여 재구성된 피겨를 사용하여, 배경으로부터 그것을 제거하거나, 다른 경기장에 재배치하거나, 임의의 컴퓨터로 생성된 프로그램으로 익스포트(export)할 수 있다.
- [0320] 본 발명의 바람직한 실시예에 따라서, 사용자는 재구성된 피겨들을 기반으로 새로운 피겨를 생성할 수 있다. 사용자는 나아가 그 텍스처, 기관 등을 추가하거나 변경할 수 있다.
- [0321] 본 발명의 바람직한 실시예에 따라서, 사용자는 기존의 촬영되어 있는 장면, 예컨대 옛 영화를 사용할 수 있고, 영화의 피겨 및 배경을 모델링하기 위해 영화의 데이터를 사용할 수 있다. 이것은 완전 3D 공간 기반 환경, 또는 피겨들의 경기장 및 거기서의 위치를 생성함으로써 수행될 수 있고, 생성한 3D 환경을 기반으로 원래 피겨 및 환경들로부터 만들어진 새로운 영화를 생성할 수 있다.
- [0322] 본 발명의 바람직한 실시예에 따라서, 가상 집합(gathering)이 사용자의 가상 3D 사본을 사용하여 수행될 수 있다. 그러한 가상 집합은 사용자의 움직임 캡처를 포함할 수 있다. 어플리케이션은 사용자를 가상 무술 수업에 참여시키고, 선생님이 사용자의 3D 피겨를 볼 수 있고 그의 움직임을 고쳐줄 수 있도록 하는 것이며, 각 학생은 다른 학생들의 3D 피겨를 볼 수도 있다. 움직임 캡처는 사용자 자신의 웹 카메라를 사용하여 수행될 수 있다.
- [0323] 그러한 어플리케이션은 또다른 교육 목적, 가상 체육 훈련, 가상 비디오 회의 등에 사용될 수 있다. 3D 모델 및 움직임 캡처는 또한, 가상 전시회, 멀티플레이어 게임 또는 가상 데이트에 까지도 사용될 수 있다.
- [0324] 본 발명의 바람직한 실시예에 따라서, 공간 기반 3D 모델은 군인들을 훈련하기 위한 전투장의 시뮬레이션, 비행 시뮬레이션 등에도 사용될 수 있다.
- [0325] 본 발명의 바람직한 실시예에 따라서, 3D 경기장은 의료용 장비에도 사용될 수 있다. 하나 이상의 센서로부터 포착된 이미지들을 조작하는데 사용될 수 있다. 이미지들은 실시간의 실제 외과 수술용 또는 시뮬레이션용의 신체 기관의 3D 모델을 생성하는데 사용될 수 있다.
- [0326] 본 발명의 특정 어플리케이션에서, 여기에 설명된 모델 및 환경들은 PAL 및 NTSC와 같은 서로 다른 비디오 표준 사이의 변환에도 사용될 수 있다.
- [0327] 여기 설명된 기술들의 한가지 어플리케이션은 비디오 압축이다. 어플리케이션에서, 촬영된 클립을 사용하는 공간 기반 3D 모델링은 모델의 전송을 가능케하고, 그리고 난 뒤에 필요한 것은 움직임 정보의 전송 정도이다. 그러한 기술은 비디오 프레임의 전송에 있어서 대역폭을 크게 절약할 수 있음을 의미한다. 이 어플리케이션은 영화에서부터 핸드폰으로 촬영한 비디오 클립까지 다양한 비디오 및 다양한 비디오 스펙에 적용될 수 있다.
- [0328] 또한, 본 실시예는 새로운 비디오 녹화 방법을 제공한다. 본 실시예의 3D 기반 모델로 직접 비디오 녹화가 되거나, 3D 기반 모델에 적용될 수 있다. 비디오 프레임 자체는 정보가 모델로 추출된 후에 재생성될 수 있다.
- [0329] 본 실시예의 3D 모델은 단일 소스로부터의 실시간으로 움직이는 엘리먼트를 캡처하고 모델링하고, 임의의 방향으로부터 그것을 관찰하는데 사용될 수 있다. 한 어플리케이션에서, 다른 스크린의 다수의 사용자들이 실시간

으로 임의의 방향 및 줌으로 이 피겨들을 시청할 수 있다.

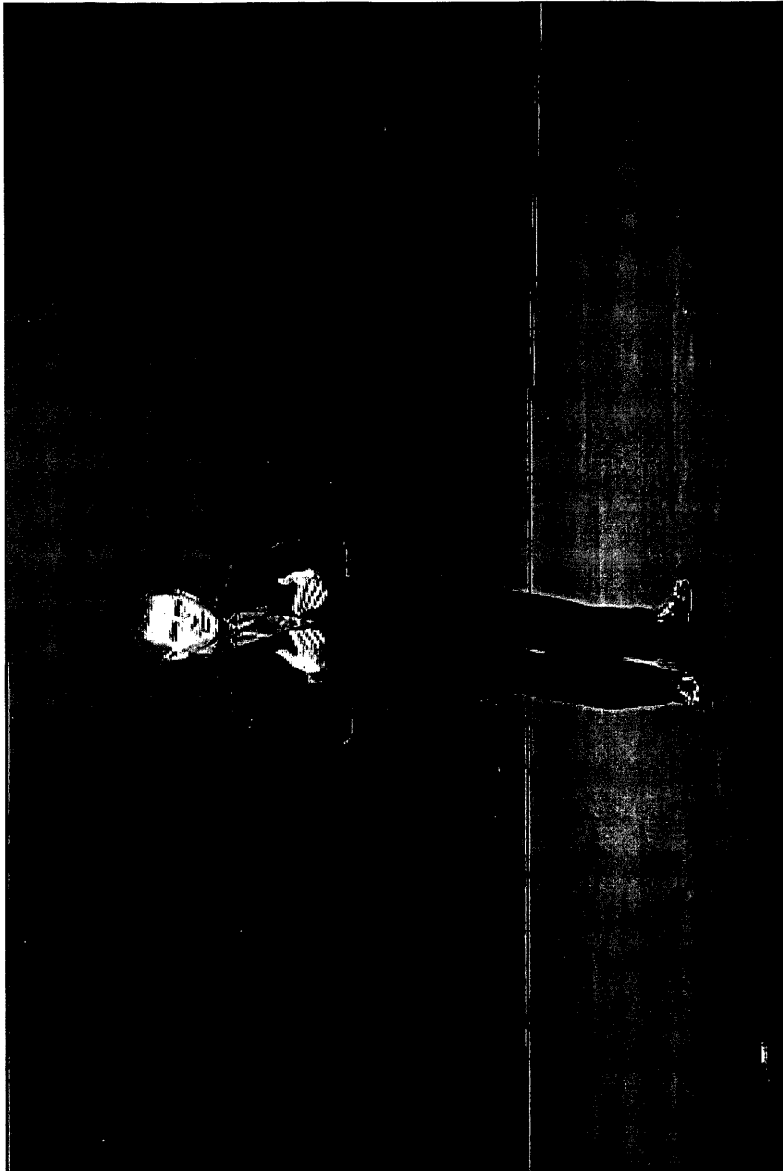
- [0330] 본 발명의 바람직한 실시예에 따른 장치는, 임의의 컴퓨터 프로그램에 대해, 사용자의 3D 움직임을 실시간으로 캡처하여, 손 또는 몸의 3D 움직임으로 그 컴퓨터를 완전히 동작시키는데 사용될 수 있다. 이러한 구현은 특정 카메라, 보통 비디오 카메라와 같은 보통 카메라, 정지 사진 카메라 또는 핸드폰용 카메라를 사용할 수 있다. 예를 들어, 사용자는 컴퓨터 게임에 포함되어 기존의 2D 또는 3D 캐릭터들이 사용자의 움직임에 따라 움직일 수 있다. 이것은 또한 핸드폰 또는 다른 휴대용 장치들의 사용자 인터페이스에서 수행될 수 있다.
- [0331] 본 발명의 바람직한 실시예에 따라서, 사용자들은 자신들을 전체 또는 부분적인 3D 모델로 모델링할 수 있고, 그 자신들을 컴퓨터 게임 또는 임의의 다른 관련된 컴퓨터 프로그램에 포함시킬 수 있다.
- [0332] 본 실시예의 어플리케이션들은 컴퓨터 게임, 시뮬레이터, 또는 다양한 플랫폼에서 완전한 실제 이미지 2D/3D 피겨들 및 배경을 생성할 수 있도록 해준다.
- [0333] 본 발명의 바람직한 실시예에 따라서, 3D 모델링은 적외선 등의 임의의 센서로 수집된 정보를 사용하여 수행될 수 있다.
- [0334] 본 발명의 바람직한 실시예에 따라서, 적절한 센서들로부터 수집된 데이터를 사용하여 초소형의 정보들도 새로운 3D 공간 기반 모델로 모델링될 수 있다.
- [0335] 본 발명의 바람직한 실시예에 따라서, 3D 모델 및 텍스처를 사용하여, 아원자(subatomic) 파티클, 멀리 떨어진 천체, 또는 센서가 캡처할 수 없는 영역(예컨대, 벽 뒤)과 같이 광학 정보가 없는 경우에도, 센서들로부터 수집된 데이터에 의해 새로운 사용자 정의 2D/3D 경기장을 생성할 수 있다.
- [0336] 본 발명의 바람직한 실시예에 따라서, 3D SP 방법은 기계 비전 장치에도 사용될 수 있다. 예를 들어, 로봇에 장면에 대한 3차원 공간 이해를 제공하는데 사용될 수 있다. 그에 따라, 로봇은 인간을 다중 프레임에서의 부분적인 이미지가 아닌 통일된 3차원 존재에 관련지을 수 있다. 결과적으로 로봇은 장애인을 돕는 등의 경우에도 사용될 수 있다.
- [0337] 어플리케이션에 의해 필요한 바와 같이, 3D SP 과정은 모델의 3D 정보를 구성하는 인치당 텍스처 영상의 수 및 깊이 포인트의 수에 의해 초해상도로 재구성된 3D 모델을 생성할 수 있다.
- [0338] 본 특허의 존속 기간 동안 많은 관련 사진 및 이미징 장치 및 시스템들이 개발될 것이고, 본 명세서에서 사용된 용어들, "3D 모델", "이미지 캡처", "깊이 맵", "클립", "가상 현실", "컴퓨터"의 범위는 그러한 개발된 모든 기술들을 포함하는 것으로 의도된다.
- [0339] 다음 예들을 검토하면, 본 발명의 추가적인 목적, 효과 및 신규한 특징이 당업자에게 명확해질 것이나, 본 발명의 범위가 이들 예로 제한되는 것은 아니다. 또한, 상세한 설명에 설명되고, 이하의 청구범위 섹션에 청구된 다양한 실시예 및 본 발명의 관점들 각각은 다음 예들에 의해 실험적으로 지지된다.
- [0340] 간명함을 위해 별개의 실시예들로 설명된 본 발명의 특징들은 하나의 실시예로 결합될 수 있다. 반대로, 간명함을 위해 하나의 실시예로 설명된 본 발명의 특징들이 별개로 또는 임의의 적절한 하위 조합으로 제공될 수 있다.
- [0341] 본 발명을 특정 실시예들과 함께 설명하였으나, 당업자에게는 그에 대한 다양한 변형이 가능할 것이다. 따라서, 첨부된 청구항의 범위에 그러한 변형예가 모두 포함되는 것으로 의도된다. 본 명세서에 언급된 모든 문서, 특허 및 특허 출원은, 각각이 참조에 의해 일체화되는 것으로 표시된 것과 마찬가지로, 모두 참조에 의해 본 명세서에 일체화된다. 또한, 본 출원에서 임의의 참조에 대한 인용 또는 식별은 그 참조가 본 발명에 대한 선행기술이라는 것을 인정하는 것으로 해석되지 않는다.

도면

도면1



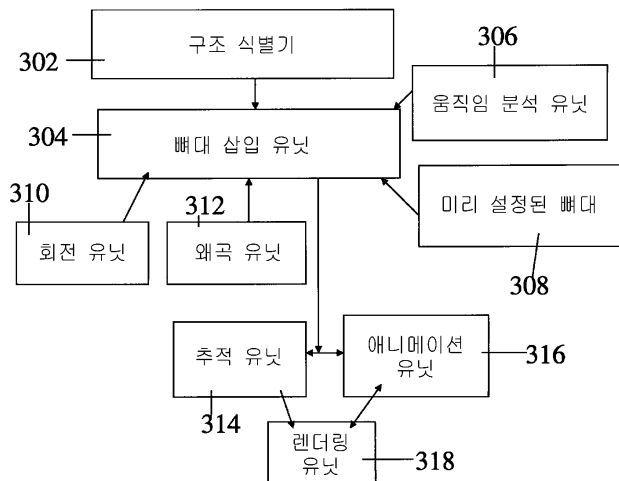
도면2a



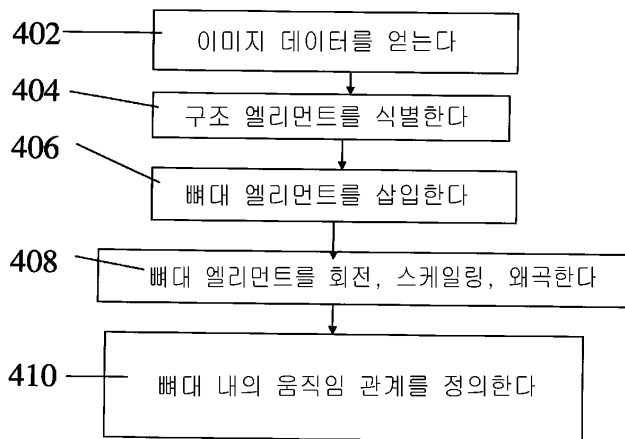
도면2b



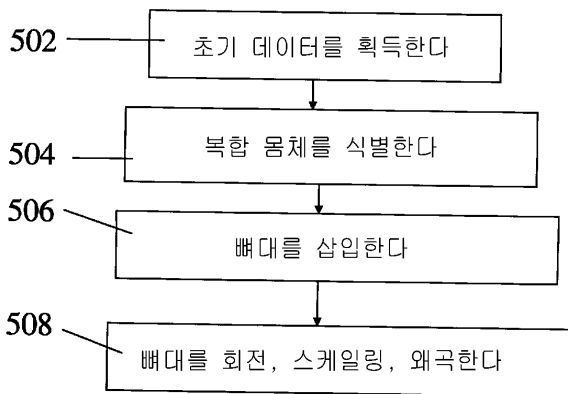
도면3



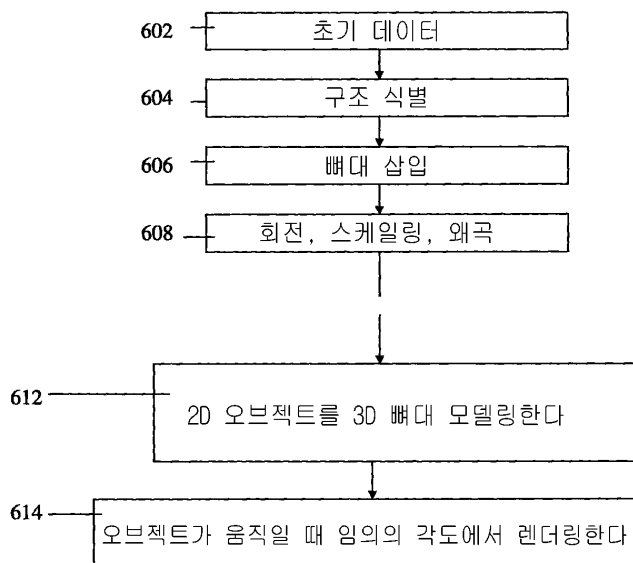
도면4



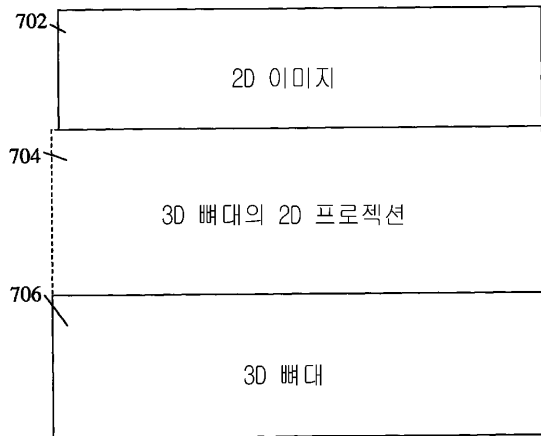
도면5



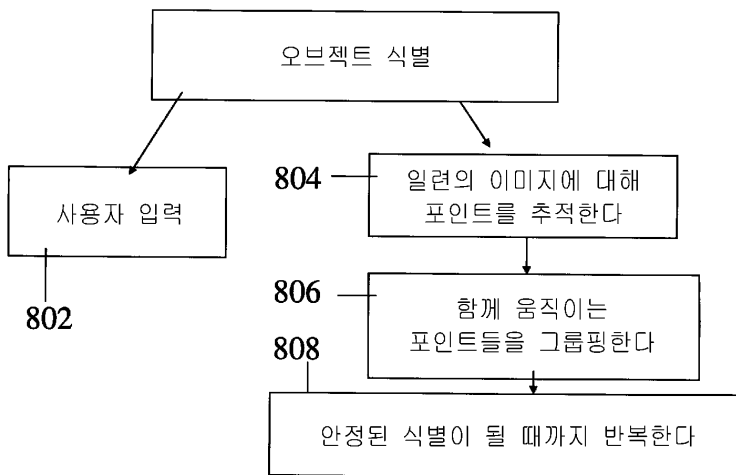
도면6



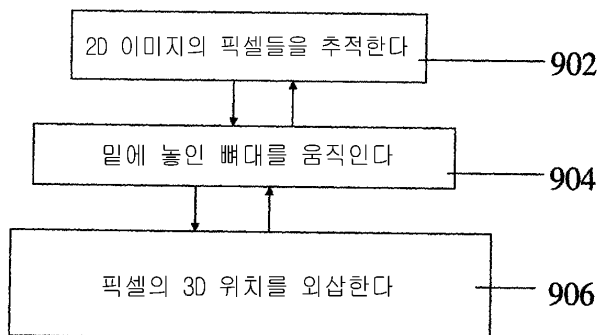
도면7



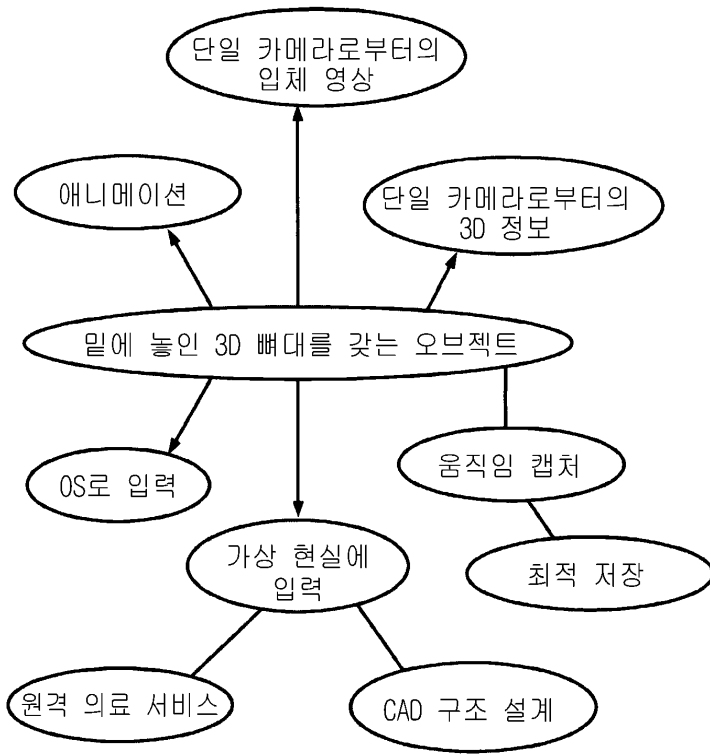
도면8



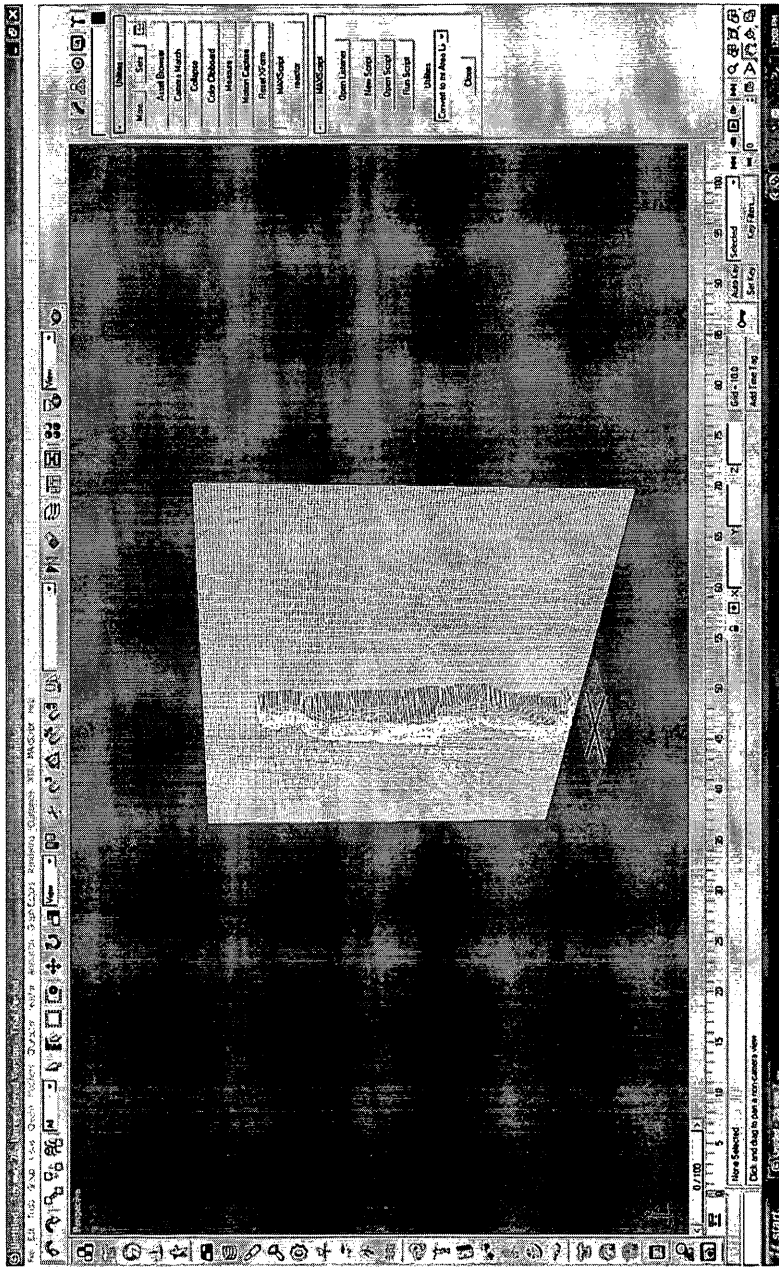
도면9



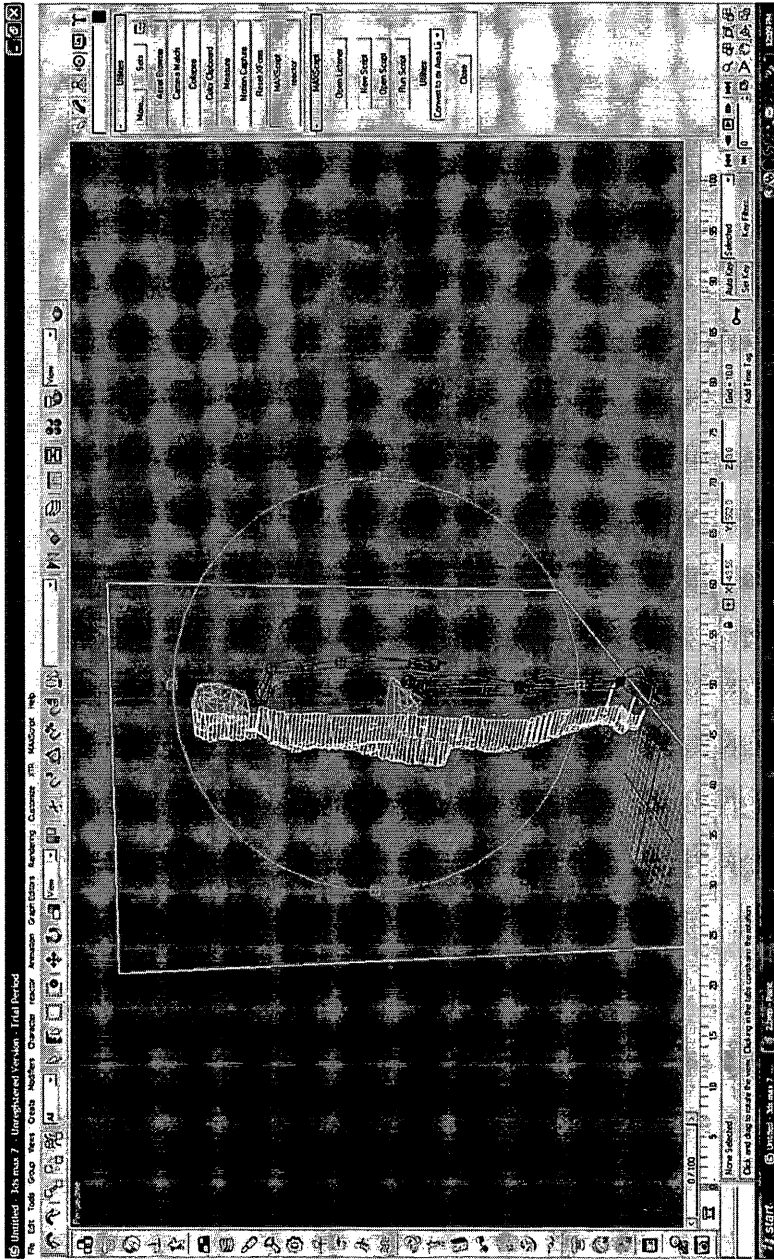
도면10



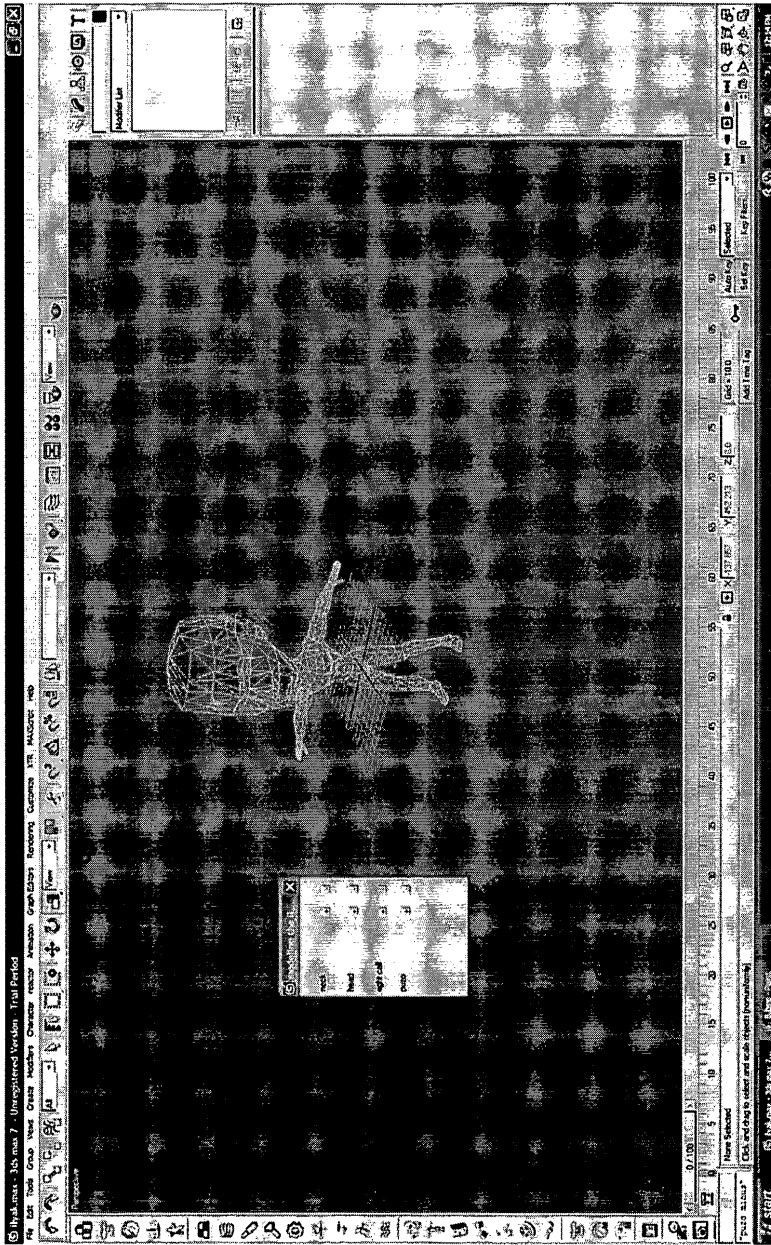
도면11



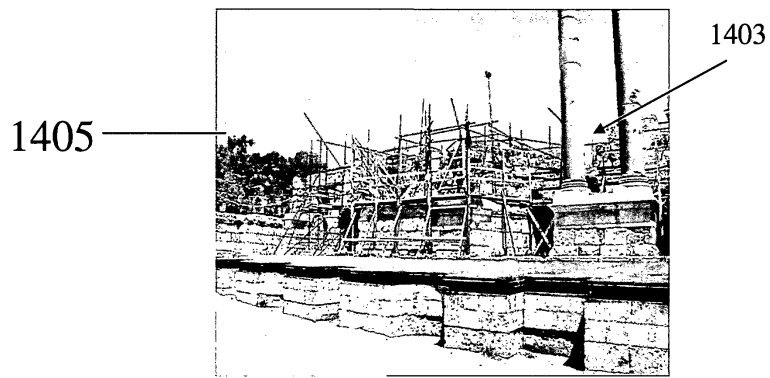
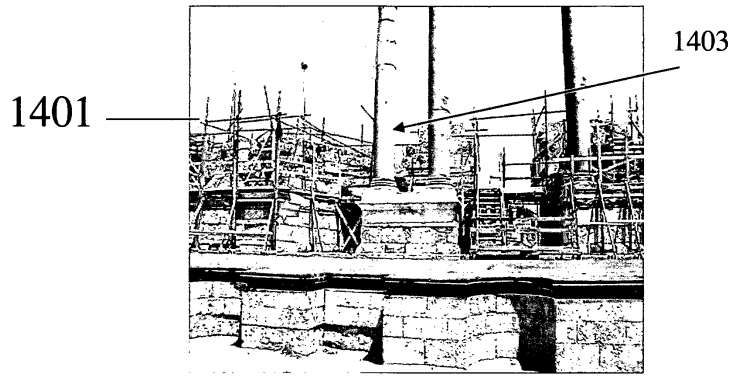
도면12



도면13



도면14



도면15

