



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2007-0119018
(43) 공개일자 2007년12월18일

- (51) Int. Cl.
G06T 17/20 (2006.01) *A63F 13/00* (2006.01)
G06T 15/00 (2006.01) *H04N 5/262* (2006.01)
- (21) 출원번호 10-2007-7021516
 (22) 출원일자 2007년09월19일
 심사청구일자 없음
 번역문제출일자 2007년09월19일
- (86) 국제출원번호 PCT/CA2006/000265
 국제출원일자 2006년02월23일
 (87) 국제공개번호 WO 2006/089417
 국제공개일자 2006년08월31일
- (30) 우선권주장
 60/655,514 2005년02월23일 캐나다(CA)

- (71) 출원인
 씨머스, 크레이그
 캐나다 비3제트 2알9 노바 스코티아 글렌 해븐
 테너 프레드릭 로드 68
- (72) 발명자
 씨머스, 크레이그
 캐나다 비3제트 2알9 노바 스코티아 글렌 해븐
 테너 프레드릭 로드 68
- (74) 대리인
 양영준, 백만기

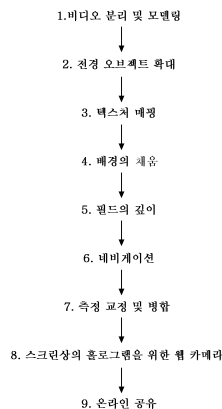
전체 청구항 수 : 총 68 항

(54) 3D 카메라 및 3D비디오를 위한 자동 썸 모델링

(57) 요약

단일 카메라 화상 처리 방법이 일반적인 이동성 비디오내에서의 3D 네비게이션에 대해 개시된다. 컬러 및 휘도에 따라, XYZ 좌표는 픽셀마다 정의될 수 있다. 이에 따른 기하학적 모델이, 레이저 범위 파인더와 같은 장비 및 현장 측량을 대체하여, 디지털 화상으로부터 측정치를 얻도록 사용될 수 있다. 모션 시차가 전경 오브젝트를 배경으로부터 분리시키는데 사용된다. 이는, 상이한 배경내의 비디오 요소의 배치, 프로덕트 배치, 및 컴퓨터 지원 설계(CAD) 모델 및 다른 소스로부터의 측정군과 비디오 요소의 병합을 위한 편리한 방법을 제공한다. 가정 내 사용자가 비디오 플라이 스루 또는 비디오로부터의 특정 3D 요소를 저장할 수 있으면, 이 방법은 사전적인 브랜드화된 매체 공유를 위한 기회를 제공한다. 이 화상 처리가 화상 회의 카메라에서 사용되면, 사용자의 움직임은, 통상적인 텔레비전 및 컴퓨터 스크린 상에 3D 홀로그램 효과를 생성하면서, 관점을 자동적으로 제어할 수 있다.

대표도 - 도3



특허청구의 범위

청구항 1

2차원의 디지털 화상의 시퀀스를 네비게이션가능한 3D 모델로 자동적으로 분할하는 방법에 있어서,

- a) 화상 시퀀스를 캡처하여, 비례적으로 더 큰 측방향 모션에 기초해서 더 근접한 매트(matte) 층 및/또는 깊이 맵(map)을 정의하는 단계;
- b) 단계 a)에서 캡처되어 처리된 미가공 비디오 데이터로부터 배경 및 전경 오브젝트를 위한 와이어프레임(wireframe) 표면을 생성하는 단계;
- c) 상이한 시각으로부터의 실루엣, 오브젝트의 폭 증가 및 감소에 비례해서 깊이식으로 돌출하는 중심 골격(spine), 또는 이용가능한 경우 모션 시차(motion parallax) 정보 중 하나를 사용하여 전경 오브젝트에 깊이를 부여하는 단계;
- d) 상기 와이어프레임상에 상기 미가공 비디오를 텍스처 매핑하는 단계;
- e) 블랭크 스팟(blank spot)의 중심으로 화상 에지를 안으로 확장시킴으로써, 상기 배경 및 보이지 않는 측부 모두 상에, 전경 오브젝트 뒤의 막힌 영역을 채우는 단계; 및
- f) 기존의 비디오 소프트웨어 개발 키트를 사용하거나, 또는 널리 공지된 컨볼루션 마스크(convolution mask)를 구현하는 화상 처리 코드를 기입함으로써, 더 근접한 오브젝트상에 표면 화상을 뚜렷하게 하고, 더 먼 화상을 희미하게 하여 깊이 지각을 더 생성함으로써, 화상 시퀀스를 3D 모델로 자동적으로 분할하는 단계를 포함하는 방법.

청구항 2

엔지니어링, 산업 및 다른 응용을 위해, 디지털 화상으로부터 생성된 3D 모델의 유닛 측정에 기초하여 씬의 오브젝트 및 특징들을 비접촉(non-contact) 측정하는 방법에 있어서,

- a) 점 또는 형상에 대해 X, Y, Z 좌표가 정의되면, 일상적인 수학적 계산이 거리 및 다른 측정치들을 계수 또는 계산하기 위해 사용될 수 있고,
- b) 측정되면, 특정 크기에서의 데이터 병합 또는 교정이 필요하고, 사용자는 소프트웨어 인터페이스에서 가시적 기준 오브젝트에 대해 일 길이만 지시할 수 있고, XYZ 좌표는 그 유닛으로 변환될 수 있고,
- c) 인터페이스는 사용자가 어디에 측정이 필요한 지를 지시하도록 할 수 있고, 그 결과의 거리, 부피, 또는 다른 측정치들을 보여지게 할 수 있는 방법.

청구항 3

웹 카메라를 사용하여 간단한 신체 움직임 및 몸짓으로 3D 비디오, 3D 컴퓨터 게임, 오브젝트 영화, 3D 오브젝트 및 파노라마 VR 씬에서의 관점 및 네비게이션을 제어하여 사용자의 전경 모션을 검출한 다음, 상기 전경 모션을 마우스 또는 키보드 입력과 같이 전송하여 상기 관점을 제어하거나 또는 네비게이션하는 방법.

청구항 4

제1항에 있어서,

타겟 또는 마스크 오브젝트를 배치하기 위한 수동적인 개입 없이, 전경 매트는 자동적으로 추출되고, 모션 시차를 사용해서 깊이가 정해지는 3D 모델의 생성 방법.

청구항 5

제1항에 있어서,

전체 3D 오브젝트는 단 3개의 화상으로부터 생성될 수 있고, 부분적 형상 및 깊이 모델은 2개의 일련의 또는 시각적 화상만으로부터 개발될 수 있는 3D 모델의 생성 방법.

청구항 6

제5항에 있어서,

상기 상이한 시각에서의 전경 오브젝트의 실루엣의 맵을 자동적으로 추출하기 위해, 모션 시차가, (청색-스크린 사진술 또는 수동적 배경 제거보다는) 상기 오브젝트가 일 시각으로부터 다른 시각으로 회전할 때의 비디오에서 사용될 수 있는, 2개 또는 3개의 화상으로부터 기하학적 형상을 생성하기 위한 절차.

청구항 7

제1항에 있어서,

상기 화상이 3D 점 및 깊이 맵 또는 와이어프레임을 생성하는데 사용된 다음, 또한 상기 깊이 맵 또는 와이어프레임 상에 텍스처 매핑되어 사실적인 3D 모델을 생성하는 3D 모델의 생성 방법.

청구항 8

제1항에 있어서, 실행되는 비디오로 변경될 수 있는 동적 와이어프레임 모델에 기초하는, 모션 시차를 사용하는 3D 모델의 생성 방법.

청구항 9

제1항에 있어서, 지정된 위치에 있을 필요가 없는 비디오 및/또는 스틸 카메라 모두로부터의 화상의 시퀀스를 사용하는 3D 모델의 생성 방법.

청구항 10

제1항에 있어서, 3D 모델은 자동적으로 생성되고, (입체적 영상(stereoscopy) 또는 다중-카메라 화상 캡처가 사용될 수 있지만) 단일의 촬상 장치만이 요구되는 3D 모델의 생성 방법.

청구항 11

제1항에 있어서, 프로덕트 배치를 위해 수행되는 것으로서, 3D 씬의 점들에 대한 상기 XYZ 좌표는 상기 씬에서 추가의 정적인 또는 이동하는 오브젝트의 배치를 가능하게 하도록 크기 조정될 수 있는 선형 비디오로부터 3D 씬을 자동적으로 생성하는 방법.

청구항 12

제1항에 있어서,

움직임의 상이한 속도를 식별하기 위한 프레임들 간 화상 비교는 절대적인 매치(absolute match) 보다는 "최적의" 특징 매치에 기초하는 3D 모델의 생성 방법.

청구항 13

제1항에 있어서,

3D 카메라의 화상 캡처 동안에, 또는 예를 들어, 셋탑 박스, 디지털 매체 허브(hub) 또는 컴퓨터에서의 시청 시에 처리가 수행될 수 있는 3D 모델의 생성 방법.

청구항 14

제2항에 있어서,

역사적인 영화 장면, 보다 광범위하게 임의의 선형 영화 장면으로부터 네비게이션가능한 3D 씬을 자동적으로 생성하기 위한 방법으로, 촬상 또는 시청시 중 하나에서 처리가 수행될 수 있는 방법.

청구항 15

제1항에 있어서,

상기 소프트웨어 인터페이스는, 전경 층과 배경 간의 팝 아웃(popout); 키프레임(keyframe) 주파수; 전경 오브젝트의 범위; 와이어프레임이 변하는 속도; 및 필드의 깊이에 대한 선택적 조정가능한 제어를 포함하는 3D 모델의 생성 방법.

청구항 16

제3항에 있어서,

상기 사용자가, 자신의 움직임이 마우스-위(mouse-over) 또는 마우스-아래(mouse-down) 제어에 영향을 미치는지에 관계없이, 자신의 움직임에 기초한 관점의 변화의 감도, 움직임의 방향의 역 방향, 및 키프레임 속도를 포함하는 변수를 조정할 수 있는, 화상 회의 카메라를 사용하여 통상적인 모니터상에 홀로그램 효과를 생성하는 방법.

청구항 17

제3항에 있어서,

상기 사용자의 신체 움직임은, 게임 엔진을 위한 소프트웨어 개발 키트(SDK;software development kit) 및/또는 다이나믹 링크 라이브러리(DLL;dynamic link library)를 통해 제공된 움직임 명령으로 화상 회의 카메라에 의해 검출되거나, 또는 마우스, 키보드, 조이스틱 및/또는 게임패드 드라이버 입력에 상기 사용자의 신체 움직임을 추가하는 운용 시스템 드라이버에 의해 검출되는 통상적인 모니터상에 홀로그램 효과를 생성하는 방법.

청구항 18

제1항에 있어서,

XYZ 관점은, 중심 또는 "노드(nodal)" 점을 벗어난 선내의 상기 관점이 이동할 때의 시차를 나타내는 전경 오브젝트 주변에서 이동할 수 있는 3D 모델의 생성 방법.

청구항 19

제1항에 있어서,

디스크상의 파일, 웹 카메라 출력, 온라인 비디오 스트리밍 및 케이블 방송을 포함하는 다양한 포맷의 디지털 비디오는, DirectX 또는 OpenGL과 같은 플랫폼의 소프트웨어 개발 키트(SDK)를 사용하여 3D에서 처리, 텍스처 매핑 및 재생될 수 있는 3D 모델의 생성 방법.

청구항 20

제1항에 있어서,

평면형, 원통형, 구형 또는 입방형 배경과 같은 좌표 시스템에 선형 비디오 또는 파노라마 비디오 중 하나를 사용하는 3D 모델의 생성 방법.

청구항 21

제1항에 있어서,

폐쇄된 영역에 대한 새로운 정보가 처리되거나 또는 초기에 폐쇄된 영역이 더이상 선에 존재하지 않을 때까지, 폐쇄된 영역의 화상 및 임의의 표면 구조를 저장함으로써, 배경이 더 보이게 되면 폐쇄부(occlusion)는 또한 채워질 수 있는 3D 모델의 생성 방법.

청구항 22

제3항에 있어서,

측 방향으로의 이동이 화상 회의 카메라에 의해 검출되어, 반대 방향의 마우스 드래그(drag) 명령으로 변환되어, 사용자가 통상적인 컴퓨터 데스크탑에서 전경 오브젝트 주변을 볼 수 있도록 하고, 스크린상의 윈도우 뒤를 볼 수 있는 능력을 제공하는, 화상 회의 카메라로 네비게이션 및 관점을 제어하는 방법.

청구항 23

제1항에 있어서,

분리된 선 요소들은, MPEG-4와 같은 비디오 압축 코드를 사용해서, 상이한 프레임 속도로 전송되어 보다 효과적

으로 대역폭을 사용할 수 있는 3D 모델의 생성 방법.

청구항 24

제1항에 있어서,

상기 모션 분석은, 단 한개의 개별 오브젝트가 아니라, 화상 시퀀스에서 가시적인 모든 씬 요소들에 대해 공간에서 XYZ 점을 자동적으로 생성하는 3D 모델의 생성 방법.

청구항 25

제1항에 있어서,

상이한 시각으로부터의 화상에 삼각법이 사용되어, 회전량을 알면서 상이한 각으로부터의 단면 폭을 XYZ 좌표로 변환할 수 있는 3D 모델의 생성 방법.

청구항 26

제25항에 있어서,

사진들 간의 회전 각은 사용자 인터페이스에서 주어질 수 있거나, 또는 사진은 3D 오브젝트 모델의 완전한 자동 렌더링을 위해 선정된 각에서 촬영될 수 있는, 상이한 각으로부터의 오브젝트 실루엣을 사용하여 오브젝트의 두께 및 형상을 정의하는 방법.

청구항 27

제1항 및 제25항에 있어서,

상기 오브젝트의 깊이는 상기 오브젝트상의 중심 리지(ridge) 아래의 일 에지에 의해 정의될 수 있거나, 또는 코너의 선명도(sharpness)를 사용자가 조정가능하게 선택하면서 더 곡선화된 폴리곤(polygon) 표면이 될 수 있는, 깊이를 정의하기 위해 중심 골격을 정의하는 방법.

청구항 28

제1항에 있어서,

와이어프레임 표면을 구성하도록 외부 오브젝트 데이터 점상에 삼각형이 생성되고, 데이터 포인트들의 쌍의 컬럼(또는 로우)을 사용해서 4개의 좌표 중 3개의 좌표 사이에 삼각형을 형성하는 컬럼을 만든 다음, 다음 컬럼으로 진행하기 전에, 동일한 컬럼 아래에 다른 삼각형으로 사각형을 채우는, 3D 모델의 생성 방법.

청구항 29

제28항에 있어서,

상기 사용자는, 단일의 용기된 표면 맵 또는 분할된 오브젝트를 생성하여, 오브젝트 에지로부터 배경으로 삼각형을 결합시키거나 결합시키지 않는 옵션을 갖는 3D 와이어프레임 모델의 생성 방법.

청구항 30

제1항에 있어서,

상기 소프트웨어는 픽처(경계선(seam))의 에지를 이동하여 상기 화상을 더 또는 덜 보이게 하는 변수를 포함하여, 상기 화상의 에지의 피트(fit)를 개선시킬 수 있는, 와이어프레임 모델상에 소스 화상을 표면-매핑하는 방법.

청구항 31

제1항에 있어서,

이동하는 오브젝트의 속도, 크기 또는 거리에 대한 불명료함은 더 근접한 측상의 오브젝트를 더 빨리 이동시키고, 상기 화상의 사실성이 상기 거리에 있어서의 정확도의 부족을 극복하도록 함으로써 간단하게 해결되는, 3D 모델의 생성 방법.

청구항 32

제1항에 있어서,

화상에 걸쳐서 픽셀마다 이동하는, 가변 크기, 형상 및 값들의 템플릿(template) 또는 "마스크"를 사용해서 일 프레임을 후속하는 프레임과 비교하여, 일 프레임으로부터 다음 프레임으로 각 픽셀의 명암 또는 컬러와 같은 변수에 대해 가장 근접한 매치를 추적하여, 상기 화상의 이동 영역을 결정하는, 3D 모델의 생성 방법.

청구항 33

제32항에 있어서,

마스크를 사용하여 전경 오브젝트를 정의하는 선택은, 프레임마다 변하는 영역을 정의하고, 이들 영역 각각의 중심점을 정의하고, 그 중심점을 추적하여 움직임의 위치, 속도 및 방향을 결정하는 것인, 움직임 및 시차를 검출하는 방법.

청구항 34

제1항에 있어서,

비교 프레임들 사이를 스킵하는 다수의 프레임, 마스크의 크기, 생성된 깊이 층의 수, 일 오브젝트내의 폴리곤의 수, 및 움직임의 이전 방향 및 속도에 기초한 검색 영역을 포함할 수 있는 조정가능한 파라미터로, 비디오 재생 및 깊이의 훌륭한 검지를 유지하면서 요구되는 기하학적 계산을 줄일 수 있는, 화상 시퀀스의 처리 방법.

청구항 35

제1항 및 제3항에 있어서,

소프트웨어의 기본 선전용 버전 및/또는 3D 모델 및 생성된 비디오 플라이 스루는 압축된 자기-실행(self-executing) 기록 파일로 ZIP 압축될 수 있고, MP3 3 음악과 같은 다른 매체 콘텐츠에 대해서도 사용되는 매체 공유 폴더로 디폴트로서 저장될 수 있는, 3D 모델을 생성 및 네비게이션하는 방법.

청구항 36

제1항에 있어서,

- a) 디폴트로서, 생성된 임의의 3D 모델 또는 비디오 플라이 스루는, 소프트웨어를 얻을 수 있는 웹사이트에 대한 링크를 포함할 수 있고, 링크의 XYZ 위치는 상기 사용자에게 의해 리셋될 수 있는 (1, 1, 1)과 같은 위치로 디폴팅되고,
- b) 상기 링크는, 다른 오브젝트들 및 컬러들이 사용될 수도 있지만, 반투과의 청색 구와 같은 단순한 형상 상에 배치될 수 있는 3D 모델의 생성 방법.

청구항 37

제1항에 있어서,

비디오의 연속적인 네비게이션이 사용될 수 있거나, 또는 선정된 경로에서 관점의 더 단순한 임시의 움직임을 일 버튼으로 제어할 수 있는, 화상으로부터 3D 모델을 생성하는 방법.

청구항 38

제1항에 있어서,

네비게이션가능한 3D 정보다는 거리 정보가, 상이한 시각을 제공하여 적-녹색의 편광된 또는 LCD 셔터 글래스와 같은, 단일 화상 세트로부터 각각의 눈에 상이한 시각을 제공하는 글래스 또는 입체경 뷰어로 시청하기 위해 스테레오(stereo) 화상의 디스패리티(disparity)를 정의하는데 사용되는, 화상으로부터 깊이 맵을 생성하는 방법.

청구항 39

2차원의 화상 시퀀스를 3D 모델로 자동적으로 분할하는 방법에 있어서,

- a) 디지털 환경에서 2차원의 좌표를 갖는 화상을 캡처하는데 사용되는 비디오 장치; 및
- b) 상기 비디오 캡처링 장치로부터 검출되어 캡처된 2차원의 화상을 수신, 변환 및 처리하도록 구성되는 프로세서 시스템 - 상기 시스템은 상기 2차원의 화상으로부터의 3D 좌표를 갖는 측정군(point cloud)을 생성하고, 상기 측정군으로부터의 에지를 정의하여 3D 좌표를 갖는 와이어프레임을 생성하고, 상기 와이어프레임에 와이어망(wiremesh)을 추가하고 후속하여 상기 와이어망상에 상기 비디오 캡처링 장치로부터의 화상을 텍스처 매핑하여, 상기 3D 모델을 스크린상에 디스플레이함 -
을 포함하는 방법.

청구항 40

제39항에 있어서,
상기 프로세서 시스템은 셋탑 박스, 디지털 매체 허브 또는 컴퓨터에 배치되는 방법.

청구항 41

제39항에 있어서,
상기 화상 장치는 비디오 캡처링 장치 또는 스틸 카메라인 방법.

청구항 42

제39항에 있어서,
상기 비디오 캡처링 장치는 화상 회의 카메라인 방법.

청구항 43

제39항 내지 제42항 중 어느 한 항에 있어서,
상기 프로세서는 또한 3D 화상 에지를 폐쇄된 영역의 중심으로 확장시킴으로써 상기 폐쇄된 영역을 채우는 방법.

청구항 44

제39항 내지 제43항 중 어느 한 항에 있어서,
상기 프로세서는 상기 전경내에 있는 화상을 뚜렷하게 하고, 상기 배경에서 더 멀리 떨어져 있는 화상을 약하게 또는 희미하게 하여, 깊이 지각을 더 생성하는 방법.

청구항 45

제39항에 있어서,
상기 프로세서는 조정가능한 제어를 포함하는 방법.

청구항 46

제45항에 있어서,
상기 조정가능한 제어는 상기 전경층과 상기 배경층 간의 거리를 조절하여 필드의 깊이를 조정하는 방법.

청구항 47

제39항에 있어서,
상기 2차원의 화상은, 디스크상의 파일, 웹 카메라 출력, 온라인 비디오 스트리밍 및 케이블 방송을 포함하는 다양한 포맷들 중 임의의 포맷인 방법.

청구항 48

제39항에 있어서,

평면, 원통형, 구형 또는 입방형 배경과 같은 좌표 시스템에 선형 비디오 또는 파노라마 비디오 중 하나를 사용하는 방법.

청구항 49

제39항에 있어서,

2차원의 화상 실루엣은 상이한 각에서 사용되어 3D 오브젝트의 두께 및 형상을 정의하는 방법.

청구항 50

제39항에 있어서,

상기 3D 관점은, 중심 또는 노드 점을 벗어난 썬내에서, 시차를 나타내는 전경 오브젝트 주변에서 이동할 수 있는 방법.

청구항 51

제3항에 있어서,

사용자의 움직임이 입체경 뷰잉 글래스에 대한 관점의 배향, 뷰잉각 및 거리를 제어하는데 사용되는, 화상 회의 카메라를 사용함으로써 3D 비디오, 3D 컴퓨터 게임, 오브젝트 영화, 3D 오브젝트 및 파노라마 VR 썬에서의 관점 및 네비게이션을 제어하기 위한 방법.

청구항 52

제51항에 있어서,

상기 입체경 뷰잉 글래스는 적-녹색의 입체사진(anaglyph) 글래스, 편광된 3D 글래스 또는 LCD 셔터 글래스인 방법.

청구항 53

제1항에 있어서,

상기 소프트웨어 인터페이스는 전경 오브젝트에 비해 상기 배경을 어둡게 하여 지각되는 깊이 및 팝 아웃을 향상시키는 최적의 조정가능한 제어를 포함하는 3D 모델의 생성 방법.

청구항 54

제4항에 있어서,

신뢰성 맵은 프레임 간 움직임의 보다 정확한 추적을 위해서 변화 맵 및 깊이 맵과 함께 추가될 수 있는 3D 모델의 생성 방법.

청구항 55

제4항에 있어서,

움기된 맵트가 상기 배경에 부착되어 남아 있는 것으로 보여질 수 있는, 전경 맵트의 깊이를 추론하기 위해 움직임을 분석하는 방법.

청구항 56

제4항에 있어서,

움기된 맵트가 상기 뷰어에 더 근접한 개별 오브젝트로서 보여질 수 있는, 전경 맵트의 깊이를 추론하기 위해 움직임을 분석하는 방법.

청구항 57

제1항에 있어서,

카메라 움직임은 움직임 해석을 위해 수동으로 설정될 수 있거나, 또는 썬 분석으로부터 계산될 수 있는, 3D 모

텔의 생성 방법.

청구항 58

제57항에 있어서,
상기 카메라는 정지 카메라인 방법.

청구항 59

제57항에 있어서,
카메라 움직임의 형태는 측방향일 수 있는 방법.

청구항 60

제1항에 있어서,
카메라 움직임의 형태는 제어되지 않는 방법.

청구항 61

제15항에 있어서,
상기 소프트웨어 인터페이스는, 상기 전경으로 팝 아웃하거나 또는 상기 배경으로 돌아가는 오브젝트를 계수하도록 검출 프레임에 따라서 조정되어, 안정적이고 정확한 깊이 모델링을 향상시킬 수 있는 3D 모델의 생성 방법.

청구항 62

제38항에 있어서,
좌안 및 우안 시각이 쌍안경에 디스플레이되어 깊이 팝 아웃을 생성하는 입체경 뷰를 생성하는 방법.

청구항 63

제14항에 있어서,
네비게이션을 위한 디폴트는 상기 관점의 흔들림을 조정가능한 양으로 제한하는 것인 네비게이션가능한 비디오를 렌더링하는 방법.

청구항 64

제63항에 있어서,
상기 디폴트 흔들림은 임의의 방향에서의 지정된 양인 방법.

청구항 65

제64항에 있어서,
상기 지정된 양은 임의의 방향에서 약 20도인 방법.

청구항 66

제14항에 있어서,
디폴트는 관점을 시작점으로 자동으로 되돌리는 것인 네비게이션가능한 비디오를 렌더링하는 방법.

청구항 67

제14항에 있어서,
움직임 제어는 키보드 키 및 마우스 움직임에 대해 설정되어, 사용자가 키보드를 사용해서 살펴보면서 마우스를 사용해서 씬 주변을 이동할 수 있도록 하는, 비디오로부터 네비게이션가능한 3D 씬을 렌더링하는 방법.

청구항 68

제14항에 있어서,

움직임 제어는 마우스 및 키보드 키 움직임에 대해 설정되어, 사용자가 마우스를 사용해서 살펴보면서 키보드 키를 사용해서 씬 주변을 이동할 수 있도록 하는, 비디오에 대해 네비게이션가능한 3D 씬을 렌더링하는 방법.

명세서

기술분야

<1> 본 발명은 화상 처리 기술에 관한 것으로, 구체적으로, 화상 시퀀스를 네비게이션가능한 3D 씬(scene)으로 자동적으로 분할하는 시스템 및 방법에 관한 것이다.

배경기술

<2> 현 시점에서 가상 여행은 3D 네비게이션에 대한 디지털 화상의 최대 애플리케이션이 되어 왔다. 2개의 어안(fisheye) 샷(shot)을 구면 화상(spherical image)으로 변환하는 선반 재고(off-the-shelf) 시스템, 및 360도 뷰(view)를 캡처(capture)하고 펼치는(unwrap) 포물면 거울(parabolic mirror) 시스템에 대해서, 사진을 스티칭(stitching)하여 파노라마로 만드는 다수의 포토-VR 방법이 존재한다. 불행히도, 이러한 접근법들은 단순한 동작에 대해 하나의 관점으로 제한되는 노드(nodal) 파노라마에 기초한다. 이러한 접근법들은 모두, 씬을 둘러보는 스크린상의 패닝(panning) 및 화상이 화소 처리될 때까지 확대 축소(zooming)하는 것을 가능하게 한다. 그러나, 3D 모델이 각 경우에 씬에 기초가 된다고 하더라도, 3D 모델에 있어서 이동할 수 없으며, 전경(foreground) 오브젝트(object)들을 결합시킬 수 없고, 전경 오브젝트가 배경에 대해서 움직일 때 시차(parallax)로부터 깊이를 지각할 수도 없다.

<3> 이러한 제한점은 360도 비디오에서 더욱 나쁘다. 가장 비싸고 높은 해상도를 갖는 카메라로 만들어졌다고 하더라도, 비디오에서의 해상도는 파노라마 씬에 대해 부적합하다. 하나의 장소에 고정된 관점을 갖는 것은 또한 모션(motion) 시차가 없다는 것을 의미한다. 실시간으로 움직일 때, 전경의 오브젝트는 배경의 오브젝트에 대해 움직인다. 이는 시각적 인지에 있어서 기본적인 깊이 단서이다.

<4> 대안적인 접근법으로서 3D 렌더링(rendering) 프로그램을 사용하여 3D 오브젝트 모델을 생성하는 방법이 있다. 그러나, 이 방법은 통상적으로 시간이 소모되는 접근법으로서, 값비싼 컴퓨터 하드웨어와 소프트웨어 및 값비싼 속력을 요구한다. 더욱이, 3D 렌더링 및 애니메이션을 위한 기술의 스테이트는 만화와 같은 오브젝트이다. 따라서, 사실적 3D 모델을 생성하고 나타낼 필요가 있다. 또한, 이 방법은 신속하고 비용이 저렴해야 한다.

<5> 현재의 포토모델링(photomodeling) 및 모션-추적 소프트웨어의 표준 사용은, 오브젝트 주변에 마커(marker)를 배치하거나, 사용자가 매 평면의 정점 및 형태를 표시하도록 하여, 상이한 시각으로부터 사진에 대응점이 표시되도록 하는 것이다. 그러나, 한번에 일 지점에 수작업으로 점들을 생성하는 것은 명백히 느린 것이다. 제조된 오브젝트에 대해서 실제 형태가 수작업으로 생성될 수 있다고 하더라도, 이는 유기적 오브젝트상의 미세한 변화 및 윤곽에 대해서는 잘 처리되지 못한다.

<6> Bracey, G.C., Goss, M.K. 및 Goss, Y.N.이 2001년에 국제 출원한 "물리적 특성을 이용한 3D 게임 아바타"라는 제목의 국제 출원 번호 제WO 01/635560호는 얼굴의 다양한 프로파일을 표시하여 3D 헤드(head) 모델을 구성하는 것에 관한 발명이다. 본 명세서에서 개시되는 발명은 상기 발명과 유사한 결과를 생성하도록 이용될 수 있지만, 이는 수동적 표시없이 자동적으로 생성되는 것이다. Bracey 등에 의해 정의된 헤드-모델링과 같은 사진측량법은 상이한 시각으로부터 화상에 개별적으로 특징점을 표시하는 것에 기초한다. Bracey 등은, 이것이 수동으로 또는 컴퓨터 프로그램으로 수행될 수 있다고 하였으나, 상이한 관점으로부터 상이한 형태를 갖는 소정의 것을 인식하는 것은 계산적으로 해결되지 않은 인위적인 인지의 기초적인 문제이다. Bracey 등은 이러한 오래 지속된 문제를 해결하기 위한 어떠한 방법도 특정하지 않았다. Bracey 등은, 컴퓨터 프로그램이 전면 및 측면으로부터 보여지는 경우 어떻게 눈썹을 동일한 오브젝트로 "인식"할 수 있는지에 대해서는 정의하지 않았다. Bracey 등이 설명한 방법은 사용자의 개입에 관련되어 다양한 대응 사진들에서 각 특징점을 수동으로 나타내도록 하는 것이다. Bracey 등에 의해 설명된 방법의 대상은, 실제 3D 모델링 보다는 선정된 일반적인 헤드 형태(와이어프레임(wireframe))상에 텍스처를 매핑하는 것으로 보여진다. 사람의 헤드의 형태 및 모양에 있어서 머리카락을 강조하면, 머리카락이 없는 기존의 마네킹 형태의 헤드를 사진 찍은 것은 명백한 결핍이 된다. 본 발

명의 방법은 임의의 형태에 대해 와이어프레임 오브젝트(및 텍스처 맵)를 정의할 것이다.

- <7> Bracey 등은 또한, 통상적으로 적어도 7개의 점을 표시하는 것 외에 대응하는 특징점이 사용하는 어떠한 제한도 특정하지 않은 것으로 보인다. 본 명세서에서 개시된 방법은 프레임마다 임의의 수의 픽셀을 매칭할 수 있고, 매우 확실한 방법으로 수행한다. 본 발명의 방법은 상이한 시각으로부터의 화상들 또는 와이어프레임 구조를 자동적으로 생성하는 모션 시차 중 하나를 사용한다. Bracey 등의 방법과 반대로, 본 발명의 방법은 컴퓨터 프로그램에 의해 자동적으로 수행되고, 수동으로는 거의 수행되지 않는 것을 의도한다. 본 발명의 방법은 머리만이 아니라 3D의 전체 씬을 렌더링할 것이다(이 방법은 또한 머리 및 얼굴의 근접 촬영을 포함한 사람들의 화상을 작업할 것이다). 본 발명의 방법은 Bracey 등의 방법에서와 같이 전면 및 측면 뷰를 필수적으로 사용할 필요는 없다. Bracey 등의 방법이 텍스처-매핑으로 한정되고 헤드 및 얼굴로만 한정되지만, Bracey 등의 수동적 특징 표시 방법은 포토모델링을 위한 기존의 상업화된 소프트웨어와 유사하다.
- <8> 또한, 실생활의 오브젝트로부터 3D 지오메트리(geometry)를 생성하기 위한 특성화된 하드웨어 시스템이 존재하지만, 이는 모두 노동 집약적이고 매우 비싼 장비를 요구한다.
- <9> · 스테레오(stereo) 비전: 2개의 렌즈 시스템이 소정의 거리만큼 떨어져서 조정된 특성화된 산업 카메라가 존재한다. 이것은 소비자의 사용을 위한 것이 아니고, 제조에 있어서 추가 비용이 든다. 뷰어는 통상 LCD 셔터 글래스 또는 적-녹색 3D 글래스와 같은 특수 장비를 요구한다.
- <10> · 레이저 범위 탐색: 선, 점, 또는 격자가 오브젝트상에 투사되어, 특정 광 지점이 식별될 때 광 이동 시간 또는 삼각측량(triangulation)을 이용해서 오브젝트의 거리 또는 형태가 정의된다. 이 접근법은 값비싼 장비를 요구하고, 대량의 데이터 세트에 기초하며, 느리고 실제적이지 않다.
- <11> 이러한 셋업은 실질적인 비용 및 특성화된 하드웨어의 불편함과 관련되고, 빌딩 또는 산 범위와 같은 오브젝트 보다는 작은 오브젝트에 적합한 것이다.
- <12> 이러한 상이한 영역들 모두에서 실시된 조사 및 제품 개발로부터, 화상 시퀀스의 XY 좌표로부터 자동적으로 XYZ 좌표를 생성하기 위한 수단이 거의 존재하지 않는 것으로 보여진다. 또한, XYZ 점으로부터 3D 표면 모델로 변환하기 위한 이용가능한 수단이 존재하지 않는다. 전문가 레벨 또는 소비자 레벨에서 모두, 사람들이 자신의 이동하는 비디오를 통해 네비게이션하도록 하는 시스템이 시장에 존재하지 않는다. 비디오로부터 지오메트리 모델을 자동적으로 생성하는 이용가능한 시스템이 존재하지 않는다. 사진 또는 비디오에서 작용하는 시스템도 존재하지 않고, 비교 픽처에서 타겟의 매칭을 수동으로 표시하지 않으면서 단지 몇몇의 화상으로부터 지오메트리 모델을 자동적으로 생성할 수 있는 시스템이 존재하지 않는다. 결국, 레이저 범위 탐색, 입체적 영상(stereoscopy), 다양한 형태의 3D 렌더링 및 사진측량(photogrammetry)과 같은 특성화된 접근법은 장비, 노동 및 숙련 비용과 깊이 관련되고, 소수의 주요 헐리우드 스튜디오 이외의 대부분의 영화 제작자 및 소비자의 범위에서 벗어난 기술에 기초한다.
- <13> 방송 및 영화 촬영술에 있어서, 매트(matte) 층을 추출하는 목적은 보통 교환가능한 전경 및 배경 층을 함께 합성하기 위한 것이다. 예를 들어, 밤의 날씨 방송을 위해 녹색 스크린 스튜디오를 사용해서, 사람이 얘기하는 뒤에 날씨의 맵이 디지털적으로 배치될 수 있다. 1940년대의 영화 촬영술에서도, 정교화된 씬 요소가 글래스상에 페인팅되고, 배우들은 이 "합성된" 윈도우를 통해 보면서 촬영된다. 디지털 특수 효과 이전의 시대에, 이 "매트 페인팅"은, 카메라의 시점에서 글래스상에 페인팅된 정교화된 방 가구를 사용해서 배우들이 통상적인 세트에서 촬영될 수 있도록 한다. 셀룰로이드(celluloid) 시트가 적층되어 전경 및 배경이 상이한 레이트로 다시 그려지는 셀 애니메이션에서 이와 유사한 기술이 전통적으로 사용되어 왔다. 또한, 역사적으로 디즈니의 다중면 카메라가, 합성된 글래스 윈도우상의 만화 요소를 통해 관점을 줌 인(zoom in)함으로써 깊이 시각을 생성하도록 개발되었다.
- <14> 디지털 화상 시퀀스에서 깊이를 나타내기 위한 모션 시차를 사용함으로써, 본 명세서에 개시된 방법은, 특성화된 카메라 하드웨어 또는 스튜디오 조명을 사용하지 않고 배경으로부터 전경 오브젝트들을 분리시킬 수 있다. X, Y 및 Z 좌표를 알아서 임의의 픽셀에 대한 3D 위치를 정의하면, 사람이 다른 관점으로부터 씬을 볼 수 있고 씬 요소를 통해 네비게이션할 수 있는 뷰잉이 가능하게 된다. 포토-기반의 오브젝트 영화 및 파노라마식 VR 씬과 달리, 이 움직임은 프레임에서 프레임으로 점프하지 않고 매끄러우며, 각각의 개별 뷰어에 대해 상이한 경로가 될 수 있다. 본 발명의 방법은, 씬에서 분할된 특정 오브젝트의 제거, 새로운 3D 전경 오브젝트의 추가, 또는 예를 들어, 벽상에 픽처를 교체하는 것과 같이 새로운 화상을 특정 표면상에 매핑하는 것을 가능하게 한다. 소비자가 점점 더 전통적인 텔레비전의 상업적 광고 모델을 통과시킬 수 있는 시대에, 이는 실시간 비디오에서

의 제품의 배치 방법이다. 따라서, 가정의 사용자가 비디오가 실행되는 것으로부터 비디오 플라이 스루(fly-through) 또는 특정 3D 요소를 줄이고자 하는 경우, 이 방법은 사전적인 브랜드화된(proactive branded) 매체 공유를 가능하게 할 수 있다.

- <15> 디지털 비디오 화상 회의의 카메라(또는 "웹 카메라")를 사용하는 경우, 사용자의 움직임을 쫓을 수 있고, 사용자가 시청하는 비디오의 관점을 변경할 수 있다. 이는 통상적인 텔레비전 및 컴퓨터 모니터상에서 3D 홀로그램의 효과를 제공한다. 액티브 제어를 요구하지 않는 양방향성 TV가 나와 있고, 여기에서 관점은 사용자가 이동하면 자동적으로 이동한다. 사용자는 TV를 수동적으로 시청할 수 있지만, 동작 및 신체 움직임을 사용해서, 3D 실행을 네비게이션하거나 및/또는 비디오 플레이어로서 살펴볼 수 있다.
- <16> 따라서, 2차원의 화상 시퀀스를 네비게이션가능한 3D 씬으로 자동적으로 분할하는 방법이 필요하다.

발명의 상세한 설명

- <17> 본 발명은, 2차원의 화상 시퀀스를 모션을 포함할 수 있는 네비게이션가능한 3D 씬으로 자동적으로 분할하는 방법 및 시스템에 관한 것이다.
- <18> 본 명세서에 개시된 방법은 "모션 시차"를 사용하여, 실행되는 비디오에서 전경 오브젝트를 자동적으로 분할하거나, 상이한 각으로부터의 오브젝트의 실루엣을 사용하여 오브젝트의 3D 형태를 자동적으로 생성한다. "모션 시차"는, 더 먼 배경 오브젝트의 광학적 플로(flow)보다 더 근접한 오브젝트가 상이한 속도 및 양으로 수평으로 이동하는 광학적 깊이의 단서이다. 모션 시차는 "매트", 즉, 층으로 구성될 수 있는 화상 세그먼트를 추출하는데 사용될 수 있다. 이는, 크로마킹(chromakeying)으로서 알려진 청색-스크린 매팅의 특수 조명, "로토스코핑(rotoscoping)" 영화 촬영 기법의 키프레임(keyframe)상의 수동 추적, 또는 대응점의 수동 표시를 요구하지 않는다. 모션 시차법은 또한 씬상에 어떤 종류의 그리드, 선 또는 패턴을 투사할 것을 요구하지도 않는다. 이는 3D 비디오를 위한 자동 씬 모델링을 위한 단일 카메라 방법이기 때문에, 이 기술은 "3D 카메라"내에서 동작할 수 있거나, 또는 기존의 또는 역사적인 영화 장면의 실행에 있어서 네비게이션가능한 3D 경험을 생성하는데 사용될 수 있다. 통상적인 비디오는 이 방법으로 3D내에서 연속적으로 보여질 수 있거나, 또는 3D 요소 및 플라이 스루가 절약되고 온라인상에서 공유될 수 있다.
- <19> 본 명세서에 개시된 화상 처리 기술이 도 1에서 도시된다. 이 기술은, 훌륭한 3D의 움직이는 시청각 환경으로 눈을 만족시키는 비디오의 3D 효과를 달성시켜 실행되는 밸런스(balance)를 구성한다. 모션 시차는 깊이(Z)를 프레임의 각 XY 좌표점에 부가하여, 3D 비디오를 위한 단일 카메라 자동 씬 모델링을 구성하는데 사용된다. 이 기술은 소비자가 사용하기 위해 저렴하고 비용 효율적이기 때문에 편리하게 설계됨과 동시에, 영화가 이동할 수 있으면서 시청자들도 이동할 수 있는, 전통적으로 영화로서 생각하는 것에 대한 완전히 새로운 인터페이스에 대해서도 열려 있다. 영화는 씬 내의 및 씬들 간의 이동을 고려하도록 만들어질 수 있다. 그러나, 제조 변화없이도, 셋탑 박스 및 컴퓨터용 소프트웨어는 임의의 비디오 신호가 이 시스템내에서 기하학적으로 렌더링될 수 있게 한다.
- <20> 편리성을 위해, 2D 좌표 시스템에서 수평축을 X, 수직축을 Y로 칭하는 관례에 따라서, Z는 깊이 디멘전(dimension)을 지칭하는 것으로 사용된다. 그러나, 이들 표시는 다소 임의적인 것으로서, 이와 다른 심볼이 3개의 디멘전을 지칭하기 위해 사용될 수 있다.
- <21> 일반적인 비디오로부터 3D 모델을 생성할 수 있는 기본 능력은 2개의 다른 능력들을 또한 끌어낸다. 비디오로부터 기하학적 구조를 생성할 수 있다면, 비디오의 프레임내의 특정 지점의 3D 좌표를 알아야만 한다. 이에 따라, 비디오의 오브젝트로부터 거리, 부피 및 다른 측정치들을 추출할 수 있고, 이는 이 화상 처리가 산업 애플리케이션에서 사용될 수 있도록 한다.
- <22> 그 다음 가능하게 되는 두번째 능력은 스크린상의 홀로그램 효과와 관련된다. 실행되는 비디오가 움직이는 3D 모델로 분할되면, 관점 파라미터는 시선(gaze)의 방향 및 XYZ 위치를 정의할 필요가 있을 것이다. 사람이 웹 카메라 또는 비디오 카메라를 이용하여 시청하는 경우, 시청하는 동안 이들의 움직임은 3D 비디오, VR 씬 또는 3D 게임에서의 관점 파라미터를 수정하는데 사용될 수 있다. 그 다음, 사람이 움직이면, 스크린상의 관점은 자동적으로 움직여서, 사람들이 전경 오브젝트를 볼 수 있게 한다. 이는 일반적인 텔레비전 또는 컴퓨터 모니터를 사용해서 3D 홀로그램과 유사한 효과를 발생시킨다.
- <23> 광대한 씬에서, 본 명세서에 개시된 방법의 목적은 "3D 카메라"에 매 픽셀이 저장될 수 있게 하는 것이고, 또한 XYZ 좌표에서의 위치를 정의할 수 있다. 이는 하나의 정지 관점으로부터의 비트맵을 벗어나서, 더 풍부한 3D

경험을 만들어내기 위해서 썬 지오메트리를 분석하기 위한 데이터 및 능력을 제공한다. 카메라의 화상 센서 또는 디스플레이 시에 화상 처리가 이루어질 수 있다. 본 명세서에 개시된 시스템은, 통상적인 비디오의 자동적 처리를 통해 통상적인 모니터 썬에서 강력한 시청 경험을 만들어 낼 수 있다. 특별한 카메라 하드웨어가 필요하지 않다. 이 시스템은, 수 백만개의 폴리곤(polygon)을 실제 썬으로 렌더링하려고 하는 표준 접근법보다는 화상으로부터 바로 썬을 생성하는 효율적인 방법을 사용한다.

- <24> 따라서, 본 발명의 목적은, 움직이는 비디오에서 상이한 광학 플로우에 기초하여 전경 오브젝트를 식별한 다음, 프레임의 각 XY 좌표점에 깊이(Z)를 추가하는 것이다.
- <25> 본 발명의 다른 목적은, 인구 통계, 또는 날씨 또는 위치와 같은 다른 변수들에 기초하는 동적 타겟팅의 경우에도 브랜드화된 프로덕트가 썬에 삽입될 때 간접 광고를 가능하게 하는 것이다.
- <26> 본 발명의 또 다른 목적은, 측정가능한 디멘전을 갖는 3D 모델을 얻도록 하는 화상 처리를 가능하게 하는 시스템을 구성하는 것이다.
- <27> 또한, 본 발명의 또 다른 목적은, 이용가능한 경우 웹 카메라로부터 사용자 움직임을 처리하여, 3D 스크린상에서 이동하는 경우 관점을 제어하도록 하는 것이다.
- <28> 통상적으로 3D 모델링에 있어서, 시각적 디테일(detail)은 와이어프레임 모델을 위해 최소화되어야만 한다는 것이 전제된다. 그러하더라도, 만화 영화를 위한 "플라이 스루"를 렌더링(즉, 3D 썬에 걸친 이동을 기록)하는 것은, 거대한 "렌더 팜(render farm)", 즉, 선형 비디오상에 기록된 단일 플라이 스루의 대량 다중-컴퓨터 렌더링에서 수 백만개의 폴리곤을 포함하는 와이어프레임의 처리를 요구한다. 반대로, 본 명세서에 개시된 방법 및 소프트웨어는, 3D 비디오가 어떻게 생성되어야만 하는지에 대한 전제에 있어서 매우 상이한 접근법을 취한다. 본 명세서에 정의된 방법은, 최소의 폴리곤 모델과 훌륭한 시청각 콘텐츠를 갖는 사실주의를 실현하기 위해서, 복잡하고 정교한 기하학적 모델에 대한 요구를 완화시키도록 설계된다. 이는 3D 경험을 가능하게 하여 어떤 사람이라도 가정의 컴퓨터에서 플라이 스루를 생성할 수 있게 된다. 수 백만개의 와이어프레임 표면을 렌더링하여 3D 플라이 스루를 생성하기 위해 수시간 또는 수일을 소비하는 산업 시스템이 아니라 통상적인 가정의 컴퓨터 또는 셋탑 박스면 충분하다.
- <29> 본 명세서에 개시된 방법은, 최소의 기하학적 모델을 생성하여 적당한 양의 처리로 비디오에 깊이를 더하고, 이 간단화된 기하학적 모델상에 매핑된 비디오를 간단하게 실행시키도록 설계된다. 렌더 팜이 요구되지 않는다. 제한된 수의 기하학적 오브젝트만을 생성하는 것은 렌더링을 계산적으로 더 간단하게 하고, 텍스처-매핑을 더 용이하게 한다. 이러한 방법으로 통상적인 일-카메라의 선형 비디오로부터의 이동하는 비디오에서의 3D 조정을 가능하게 함과 동시에, 모델의 단점은 비디오의 소리 및 모션에 의해 극복될 수 있다.
- <30> 이제, "사진을 찍는" 것을 의미하는 것의 본질을 바꾸는 기술적 능력을 가진다. 컬러 픽셀의 비트맵을 저장하는 대신에, "디지털 화상"은 썬 지오메트리를 또한 저장할 수 있다. 페인팅으로서 컬러의 점들을 저장할 수 있는 전통적인 능력을 사용하기 보다는, 디지털 화상은 컬러 점들 뿐만 아니라 3D 구조를 포함할 수 있다. 따라서, 소프트웨어는 사진 찍기 및 뷰잉 경험 둘 다의 기본적 특성을 변경할 수 있다.
- <31> 본 명세서에 개시된 방법을 사용해서, 전경 오브젝트는 비디오의 배경에서 분리되어 모델링, 처리 및 전송될 수 있다. 3D 비디오가 재생될 때 3D 비디오를 통한 네비게이션을 가정해 보자. 통상적인 비디오 카메라를 사용하면, 아마도 몇몇 사람들이 썬으로 걸어 들어올 것이다. 그 다음 비디오를 볼 때, 비디오를 통해 네비게이션하는 동안 그 사람들은 3D 썬에서 걸어다니는 것으로 보여질 수 있다. 인터페이스는 또한, 상공을 선회할 때 액션을 멈추게 하거나 빠르게 하거나 반대로 돌리는 것을 가능하게 한다. 이는 시간의 동결(frozen-in-time) 주변 회전 효과와 유사하지만, 이 경우에는 임의의 방향의 공간으로 이동할 수 있고, 또한 재생을 빠르게 하거나 일시 정지시키거나 반대로 돌릴 수 있다. 또한, 전경 및 배경을 분리시킬 수 있기 때문에, 걸어 감에 따라 상이한 3D 환경에 사람들을 놓을 수 있다.
- <32> 천문학자들은, 밤 동안 또는 연간 궤도에서 지구 자전시에 상이한 지점으로부터 찍은 사진에서의 거리를 추론함으로써, 모션 시차를 이용해서 행성과 별들과의 거리를 계산하는데 오래동안 관심을 가져 왔다. 또한, 본 명세서에 개시된 화상 처리는 지구의 궤도에 있어서 상이한 지점들에서 촬영된 일련의 화상으로부터 네비게이션가능한 3D 별 모델을 자동적으로 생성하는 새로운 방법을 제공한다.
- <33> 하나의 카메라로 촬영된 평평한 텔레비전 스크린 또는 모니터상의 선형 비디오로부터 가능한 뷰잉 경험의 특성에 있어서의 이러한 패러다임의 전환은, 어떻게 영화를 보는지 및 영화 생산의 특성을 본질적으로 바꿀 수 있다. 이러한 능력들을 지칭한 용어들이 "필름", "영화" 및 "모션 픽처"와 같은 용어로 제한되었지만, 비디오

가 재생될 때 네비게이션될 수 있는 비선형 비디오의 경험을 충분히 나타내는 용어는 없다. 씬에서 상호작용하는 경험이 매번 다를 수 있다는 의미에서 "재생"은 실제로 재생된다는 의미는 아니다.

<34> 양방향성 텔레비전의 생산자 및 사용자를 위한 새로운 가능성을 열 뿐만 아니라, 전경 오브젝트를 분리시킬 수 있도록 하는 것은, MPEG-4와 같은 압축 포맷의 정지 오브젝트보다 이동에 대해 더 높은 프레임 속도로 전송하여 비디오 대역폭을 낮출 수 있도록 한다.

실시예

<43> 컬러 및 명암과 같은 페인팅 및 화상의 전통적인 특성들과 함께 화상 처리에서 깊이 지오메트리가 분석되면 사진 및 비디오에서 더 좋은 뷰잉 경험이 얻어질 수 있다. 사진, 페인팅 또는 벽화에서와 같이 2차원의 화상상에 컬러 점을 표시하기 보다는, 본 명세서에 개시된 기술은 3D 씬 구조를 처리한다. 이 기술은 스틸 카메라든 비디오 카메라든 통상적인 디지털 촬상 장치에서 그와 같이 처리한다. 이 처리는 카메라에서 수행될 수 있지만, 통상적으로 시청자 측에서의 네비게이션으로 발생한다. 이 처리는 수동적인 개입없이 자동적으로 수행된다. 이는 역사적인 영화 장면에서도 처리될 수 있다.

<44> 일반적으로, 비디오에서 씬이 변화되고 카메라가 이동될 수 있으며, 이는 3D 구조에 영향을 미칠 것이다. 전체 광학 플로는 소정 형태의 카메라 움직임의 표시자로서 사용될 수 있으며, 예를 들어, 렌즈의 노드점 주변을 카메라가 선회하는 것은 시차를 없애고, 3D 모델을 평탄하게 할 수 있다. 카메라의 수평 움직임은 모션 시차 및 전경 오브젝트의 팝 아웃(pop-out)을 향상시킨다. 이동하는 오브젝트는 또한 전체 광학 플로우에 비해서 차동 모션에 기초해서 분할될 수 있다. 이는 뷰잉 경험에 대해서는 나쁘지 않을 수도 있지만, 감도 제어는 사용자가 팝 아웃의 양을 줄이게 할 수 있다. 비디오가 3D 좌표에서 재생되면, 디폴트(default)로서 이는 캡처된 초기 비디오로서 동일한 스크린 영역상에 설정된다.

<45> 현재 사용되는 모든 가상 여행과 달리, 이 시스템은 사용자가 사실적인 환경에서 이동할 수 있도록 하고, 카메라가 절대 있을 수 없는 장소에서도 어떤 시점에서든 사용자가 볼 수 있도록 한다. 거리 측정치는 기초가 되는 3D 모델때문에 씬으로부터 얻어질 수 있다.

<46> 본 발명의 일 실시예는 자동 매트 추출에 기초한다. 전경 오브젝트가 배경 광학 플로(즉, 모션 시차)와 다른 속도로 수평 움직임에 기초해서 분할된다. 그러나, 개시될 공통의 변수가 존재한다. 본래부터 일부 화상 시퀀스는 어떤 모션도 갖지 않으며, 이는 구체적으로, 사람의 또는 오브젝트의 얼굴 측면 뷰와 같은 직교 사진이다. 2개의 사진이 90도 또는 다른 특정 각도에서 찍어지면, 오브젝트 형태는 인간의 개입없이 여전히 자동적으로 렌더링될 수 있다. 배경의 움직임, 크로마킹, 또는 수동적 삭제 중 하나로 배경이 분리될 수 있는 방식으로 사진이 찍어지는 한, 상이한 시각에서의 2개의 실루엣은, 오브젝트를 정의하고, 확대시키고, 화상을 결과적인 와이어 프레임에 텍스처 매핑하기에 충분하다. 이 처리는, 배경이 키 아웃(key out)되고 사진이 선정된 회전 정도로 촬영된다면, 전체적으로 자동적이 될 수 있다. 사진이 선정된 회전량으로 촬영되지 않아도, 여전히 사용자 인터페이스로 상이한 시각의 회전 정도를 특정할 수 있다. 그 다음, 삼각법 공식이 점의 X, Y 및 Z 좌표를 계산하여 3차원에서의 와이어프레임의 외부 형태를 정의하는데 이용될 수 있다.

<47> 본 명세서에 개시된 화상 처리 시스템은 화상 캡처 장치의 형태에 관계없이 동작할 수 있고, 예를 들어, 디지털 비디오, 일련의 스틸 사진, 또는 입체경 카메라 입력에 적합할 수 있다. 이 시스템은 또한, 포물면 거울로부터 또는 외부에서 볼 수 있는 스틸 또는 비디오 카메라의 클러스터로부터 캡처될 때를 포함해서, 파노라마 화상으로 실행되도록 설계되었다. 파노라마 화상으로부터의 전경 오브젝트가 분리될 수 있거나, 파노라마는 다른 전경의 사람들 또는 오브젝트들이 배치될 수 있는 배경으로서 제공된다. 비디오로부터 3D 모델을 생성하는 대신, 본 명세서에서 개요된 방법을 사용할 수도 있으며, 이 방법은 2개의 상이한 관점을 생성하여 입체경 또는 적-녹색의 편광된 또는 LCD 서터 글래스로 깊이 지각을 만든다. 또한, 사용자의 움직임은 입체경 뷰잉 글래스에 대한 관점의 거리, 뷰잉 각도 및 배향을 제어하는데 사용될 수 있다.

<48> 이 시스템의 화상 처리는 3D 모델이 잘 정의된 디멘전을 갖도록 한다. 따라서, 생성된 씬으로부터 길이 측정치를 추출하는 것이 가능하다. 예를 들어, 기술자 및 부동산 중개업자에 있어서, 이 기술은, 현장에 가지 않고 물리적으로 측정 또는 조사하지 않고도 디지털 사진 및 비디오로부터 디멘전 및 측정치가 생성될 수 있도록 한다. 다수의 사용자로부터의 측정이 필요한 임의의 조직 또는 산업에 있어서, 관련된 값비싼 측정 하드웨어 및 요원을 방문하도록 계획할 필요없이, 다수의 사용자에 의해 처리된 또는 처리되기 위해 제공되는 화상으로, 데이터 수집이 분산화될 수 있다. 양호한 실시예는, 지시된 점 대 점 거리를 포함하여 인터페이스로부터 측정 디멘전을 얻고, 또한 렌더링된 오브젝트의 부피를 얻을 수 있는 능력과 관련된다.

- <49> 모션 시차를 이용해서 화상 시퀀스로부터 기하학적 구조를 얻는 것은 또한 네비게이션가능한 비디오 및 3D 오브젝트를 분리 또는 결합시키기 위한 일 방법이다. 이는, 새로운 MPEG-4 디지털 비디오 표준의, 고속 이동 씬 요소가 정지 요소보다 더 큰 프레임 레이트로 전송되는 압축 포맷의 객체에 적합하다. 개시된 본 발명은, 인구 통계, 또는 날씨 또는 위치와 같은 다른 변수들에 기초한 개인화된 타겟팅(targeting)의 경우에도, 브랜드화된 프로덕트가 씬으로 삽입되는 간접 광고를 가능하게 한다(페이지(phase) 7의 방법 설명 참조).
- <50> 소프트웨어는 또한, 3D 게임, 파노라마 VR 씬, 컴퓨터 데스크탑 제어 또는 3D 비디오의 네비게이션 제어 방법으로서, 텔레비전 회의 카메라(주로 "웹 카메라"로 불림)로 사용자 움직임을 검출하기 위해 사용될 수 있다. 웹 카메라는 텔레비전 회의를 위해 주로 컴퓨터 모니터 상에 장착되는 작은 디지털 비디오 카메라이다. 본 명세서에 개시된 발명으로서, 양호한 실시예는 전경에서 사용자의 움직임을 검출하여, 도 2에 도시된 바와 같이, 통상적인 텔레비전 또는 컴퓨터 모니터상의 3D 비디오 게임에서의 관점을 제어한다. 사용자의 움직임에 관한 정보는, 마우스, 키보드, 게임패드 및/또는 조이스틱으로부터 지시된 움직임 명령에 더하여, 컴퓨터로 보내져서 네비게이션 동안 관점을 제어한다. 양호한 실시예에서, 이는 운영 시스템에 설치된 드라이버를 통해 수행되고, 이 시스템은 웹 카메라로부터의 신체 움직임을 변환하여 예를 들어, 마우스 움직임의 형태로 컴퓨터에 보내지게 한다. 또한, 기능들을 3D 게임을 위한 그래픽 엔진에 부가하는 소프트웨어 개발 키트(SDK; software development kit) 및/또는 다이내믹 링크 라이브러리(DLL; dynamic link library)에서 웹 카메라 피드백을 실행시킬 수 있다. 당업자는 DLL 및 SDK의 사용이 컴퓨터 프로그래밍의 통상적인 절차라는 것을 인식할 것이다. 양호한 실시예에서는 저비용의 디지털 웹 카메라를 사용하였지만, 임의의 종류의 디지털화된 비디오 캡처 장치도 가능하다.
- <51> 웹 카메라로부터의 피드백은, 화상 처리 소프트웨어 내에서 또는 제어되는 3D 게임 또는 애플리케이션의 옵션으로, 상이한 형태의 네비게이션 및 움직임을 제어하도록 설정될 수 있다. 양호한 실시예에서, 사용자가 좌우 또는 앞뒤로 움직이는 경우, XYZ 관점 파라미터는 그에 따라서 움직인다. 그러나, 카 레이싱과 같은 일부 게임에서는, 게임에서 좌우로 움직이는 것이 관점을 변경시키고, 네비게이션을 제어한다. VRML과 같은 산업 표준에서와 같이, 공간을 통한 이동 또는 오브젝트의 회전이 선택되는 경우, 좌우 제어 움직임은 어떠한 형태의 씬 움직임이라도 사용자가 선택할 수 있도록 한다. 이는 보통 애플리케이션 또는 게임에서 정의되고, 웹 카메라 피드백의 부분으로서 설정될 필요는 없다.
- <52> 본 명세서에 개시된 방법은 또한, 영화, 스포츠 중계 방송 또는 다른 비디오 또는 화상 시퀀스를 시청할 때, 마우스로 네비게이션하는 것 보다 비디오 입력에 기초하여 관점을 제어하는데 사용될 수 있다. 영화가 소프트웨어 검출 시차에 의의해 분할되면, 웹 카메라에 소프트웨어를 사용해서 사용자 모션을 검출할 수 있다. 그 다음, 영화 재생 동안, 관점은 사용자 움직임 또는 마우스 제어를 통해 바뀔 수 있다.
- <53> 일 실시예에서, 웹 카메라가 사용되지 않으면, 움직임 제어는 키보드 키 및 마우스 움직임을 위해 설정되어, 사용자가 키보드를 사용해서 살펴보면서 마우스를 사용해서 또는 그 반대로 씬 주변을 이동할 수 있게 한다.
- <54> 소프트웨어를 사용하는 주요 기술적 절차는 도 3의 순서도에서 도시된다. 본 발명의 상기 및 다른 목적, 특성 및 장점들은 다음의 9개의 페이지에 대한 설명을 통해서 당업자에 의해 충분히 이해될 것이다.
- <55> 페이지 1: 비디오 분리 및 모델링
- <56> 광범위한 양상에서, 본 명세서에 개시된 발명은 상이한 움직임(모션 시차)의 영역들을 위한 미가공(raw) 비디오를 처리한다. 이 정보는 3D 비디오에 대한 깊이를 추론하는데 사용될 수 있거나, 또는 웹 카메라를 사용할 때, 사용자의 모션을 검출하여 3D 비디오, 포토-VR 씬 또는 3D 비디오 게임의 관점을 제어하는데 사용될 수 있다.
- <57> 프레임 간의 모션 검출의 일 실시예는 컬러 또는 명암과 같은 속성이 변화된 화상의 픽셀 및/또는 색선에 대한 검사에 기초한다. 변하는 영역들의 에지, 특징 또는 중심점을 추적하는 것은, 화상내의 움직임의 위치, 속도 및 방향을 결정하는데 사용될 수 있다. 본 발명은 본 발명의 사상 또는 본질적 특성에서 벗어나지 않으면서 임의의 이러한 특징들을 추적하는 것으로서 구현될 수 있다.
- <58> 에지 검출 및 광학 플로는 배경과 상이한 속도로 움직이는 전경 오브젝트(즉, 모션 시차)를 식별하는데 사용된다. 다수의 (또는 스테레오) 사진을 사용하던 비디오 프레임들 사용하던, 에지 검출은, 특징들의 절대적인 매치(absolute match)가 아니라, 프레임들 간 휘도, RGB값 또는 색조와 같은 특징들의 대응을 위한 최적의 매치에 기초한다. 다음 단계는 배경 및 전경 오브젝트를 위한 와이어프레임 표면을 생성하는 것이다. 배경은 입력의 디멘전에 기초한 직사각형의 비디오가 되거나, 또는 다수의 카메라, 광각 렌즈, 또는 포물면 거울과 같은 입력으로, 더 넓은 파노라마 필드의 뷰(예를 들어, 원통형, 구형 또는 입방형)가 될 수 있다. 비디오는 렌더링된

표면상에 텍스처 매핑된다. 그 다음, 재생될 때 씬내에서의 부드럽고 빠른 네비게이션 이동을 가능하게 하면서, 적합한 크로스 플랫폼(cross-platform)의 광범위하게 사용가능한 모델링 포맷(OpenGL, DirectX 또는 VRML을 포함하지만 이것으로 제한되지는 않음)으로 재생된다.

<59> 프레임들 간의 상대적인 픽셀 움직임을 측정하기 위해서, 저레벨 화상 처리의 일 실시예에서는 양 화상에서 동일한 점을 찾는다. 컴퓨터 영상 리서치에서, 이는 대응 문제(correspondence problem)로서 알려져 있다. 카메라 움직임의 인식 또는 다른 광학 플로우와 같은 정보는 이 조사를 좁힐 수 있다. 카메라가 어느 면(즉, 수평, 수직, 또는 소정의 다른 방향)으로 이동 또는 분리되는 지를 지정함으로써, 매칭 조사가 줄어든다. 프로그램은 3D 모델을 생성하기 위해 요구되는 처리 속도 및 해상도의 레벨에 따라서 컬럼을 건너 뛸 수 있다.

<60> 그 다음, 매칭 점에서의 픽셀 분리의 양은 깊이 점(즉, Z 좌표)으로 변환되고, XYZ 좌표의 (예를 들어, VRML 2.0 명세의) 3D 모델 데이터 파일로 기입된다. 또한, 더 낮은 해상도를 갖는 더 큰 형상들을 찾기 위해 처리 동안 화상의 크기를 줄여서 요구되는 처리 시간을 줄일 수 있다. 화상은, 대조점(2개의 또는 주어진 수의 픽셀에 대한 컬러 또는 휘도에서의 변화)의 식별을 간단하게 하기 위해서, 그레이 스케일(grayScale)로 줄여질 수 있다. 또한, 충분한 처리 정보만 얻는 것이 훌륭한 전략이다. 사용자는 소프트웨어 애플리케이션을 제어하여 거리 정보에서의 가장 큰 변화, 및 단지 이 정보를 찾을 것이다. 특정된 범위보다 작은 픽셀 시차에 대해서는, 단순히 화상의 이들 부분을 배경으로서 정의한다. 매칭이 이루어지면, 더이상의 추가 검색은 요구되지 않는다.

<61> 또한, 프레임 간 움직임의 보다 정확한 추적을 위해서 변화 맵 및 깊이 맵과 함께 신뢰성 맵이 부가될 수 있다. 용기된 맵이 배경에 부착되어 있는 것으로 보여질 수 있거나, 또는 시청자에게 더 가까운 개별 오브젝트로서 보여질 수 있다.

<62> 사용자 조정이 가능한 다수의 변수로서, 전경층과 배경 간의 팝 아웃의 정도를 위한 깊이 조정자, 키프레임 주파수를 위한 제어, 전경 오브젝트의 확대를 위한 감도 제어, 및 와이어프레임이 변하는 속도가 있다. 또한 필드의 깊이는 조정가능한 파라미터(페이지 5에서 구현됨)이다. 전경 오브젝트를 뚜렷하게 하여 초점을 제공하고, 이 오브젝트를 배경으로부터 더욱 잘 구별되게 하는(즉, 필드의 깊이를 짧게 하는) 것은 디폴트이다. 그 다음, 배경 비디오는 약화되고, 해상도가 낮아지며, 파노라마가 아니라면, 3D 배경에 부착되어 항상 고정되고 시청자가 그 뒤를 볼 수 없게 된다. VRML 2.0 명세에서와 같이, 디폴트 움직임은 항상 배경 앞의 XYZ 공간 내에 존재한다.

<63> 페이지 2: 전경 오브젝트 확대

<64> 오브젝트가 처음에 미가공 비디오에 기초해서 분할되면, 점들의 데이터 세트(중중 "측점군(point cloud)로서 지칭됨")가 생성된다. 이 점들은, 프로세서 자원에 기초한 디테일의 양이 특정된, 깊이가 변하는 표면으로 함께 결합될 수 있다. 함께 분할된 특징들의 그룹은 일반적으로 동일한 오브젝트의 부분으로서 정의된다. 사용자가 자신의 관점을 이동시키면, 전경 오브젝트가 두께를 갖는 경우 깊이의 착시가 더 강해질 것이다. 점들의 처리가 상세된 깊이 맵을 충분히 정의할 수 있다고 하더라도, 중심 골격(spine)을 생성하고 이를 폭에 비례해서 앞으로 당김으로써 전경 오브젝트에 깊이를 부여할 수 있다. 이것이 다소 구식이라고 하더라도, 이 알고리즘은 이동하는 비디오에서의 렌더링에 대해 고속이고, 비디오 스트림에서의 움직임 및 오디오가 임의의 인지된 결점을 극복할 수 있도록 한다.

<65> 개별 XYZ 데이터 점들의 측점군에서 와이어프레임 망으로 변환하기 위해서, 본 발명의 방법은 망의 요소들에 대해 삼각형을 사용하여 모든 폴리곤을 완전히 편평한 것을 보장한다. 삼각형은 임의의 형태를 생성하는데 사용될 수 있고, 2개의 삼각형은 사각형을 만들도록 함께 놓여질 수 있다. 삼각형의 와이어 망을 구성하기 위해서, 알고리즘은 오브젝트의 좌측 에지의 바닥부(도 6의 점 1)에서 시작한다. 가장 간단한 경우, 일 측에 그 형태를 정의하는 3 세트의 점, 즉, 도 6에 도시된 바와 같이 좌측 에지(점 1)에 대한 XYZ, 중심 두께에 대한 XYZ, 우측 에지에 대한 XYZ가 존재한다. 픽셀의 바닥부 로우(row)에서 시작하여, 좌측 에지와 중심 사이에 삼각형을 배치한다(1-2-4). 그 다음, 제1 삼각형(1-2-4)과 함께 사각형을 형성하는 제2 삼각형(5-4-2)으로 돌아간다. 이는, 먼저 하부 삼각형(1-2-4, 4-5-7, 7-8-10...)으로 그 다음 상부 삼각형(8-7-5, 5-4-2)으로 오브젝트의 최상부의 컬럼까지 반복된다. 오브젝트에 대해 3개(또는 임의의 특정 수)의 점들이 존재하는 것을 알게 되면, 그 다음, 삼각형의 정의와 화상을 삼각형에 표면 매핑하는 것 둘 다를 위해 삼각형의 코너 각각의 번호 지정(numbering)이 자동적으로 수행될 수 있다. 좌측 하부 좌표를 "1"로, 중간을 "2"로, 우측 에지를 "3"으로 정의한 다음, 각각의 더 높은 로우에서 번호 지정을 계속한다. 이것이 양호한 방법이지만, 당업자는 로우 또는 컬럼에 걸쳐서 다운 카운팅할 수도 있다는 것을 이해할 것이다.

- <66> 일 실시예에서, 도 4로부터 보여지는 바와 같이 하나 이상의 각으로부터의 측면뷰가 존재하는 경우, 오브젝트 두께의 보다 정확한 깊이 맵이 정의될 수 있더라도, 폭에 비례하여 깊이를 부여하도록 골격이 오브젝트상에 생성된다. 이 경우에, 소프트웨어는 각 픽처의 오브젝트의 실루엣을 사용하여 X 및 Y 좌표(각각, 수평 및 수직)를 정의하고, 상이한 각에서의 단면부를 사용하여 삼각법을 이용해서 Z 좌표(오브젝트의 깊이)를 정의할 수 있다. 도 5에 도시된 바와 같이, 표면 점들에 대한 X, Y, Z 좌표를 알게 되면, 와이어프레임 모델을 구성할 수 있고, 화상을 와이어프레임 표면에 텍스처 매핑할 수 있게 된다. 소프트웨어가 실루엣에 대한 명확한 에지를 검출하지 못하면, 그리기 도구가 포함되거나 또는 제3의 소프트웨어가 크로마킹 또는 마스킹(masking)을 위해 사용될 수 있다. 프레임이 충분히 가깝게 떨어져 있으면, 모션 시차는 충분할 것이다. 양 픽처를 교정하기 위해서, 프로그램은 해상도를 낮추고 픽처가 동일한 높이가 되도록 크기를 조절할 수 있다. 또한 사용자는 중앙의 형상 또는 오브젝트의 중심을 지시할 수 있어, Z 깊이가 양 픽처에서 동일한 기준을 갖게 된다. 각 사진에 대해 이 방법을 반복함으로써, 각 시각으로부터의 좌표의 세트가 생성되어 오브젝트를 정의한다. 이들 좌표는 동일한 크기로 하나의 대형 데이터 세트로 놓여짐으로써 융합될 수 있다. 이 알고리즘의 진정한 혁신적인 가치는, XYZ 좌표를 생성하기 위해 프로그램에 대해 카메라의 회전 및 크기 조절만이 요구된다는 것이다.
- <67> 제한된 수의 폴리곤이 사용될 때, 생성된 모델은 울퉁불퉁하거나 각이 져 있을 것이다. 이는 박스, 자동차 또는 빌딩과 같이 제조된 오브젝트에 대해 바람직할 수 있다. 하지만, 사람의 얼굴의 부드러움, 구름에서의 컬러의 변화 정도와 같은 유기적 오브젝트에 대해서는, 더 부드러운 곡선이 요구된다. 소프트웨어는 이를 위해 정점 및 코너에서의 에지의 부드러움 정도를 조정하는 인터페이스의 파라미터를 가질 것을 고려한다. 이는 VRML 2.0 명세에서의 유사한 파라미터와 상응한다.
- <68> 페이지 3: 텍스처 매핑
- <69> 축점군으로부터 와이어프레임 망으로 변환하고 나면, 3D 표면상에 화상을 얻을 필요가 있다. 각 프레임의 섹션에 대해 관련된 XY 좌표는, (Z를 떨어뜨리고 X 및 Y를 유지함으로써) 그 때에 존재하는 것으로서 XYZ 모델의 좌표로 매칭된다. 그 다음, OpenGL, DirectX 또는 VRML을 포함하지만 이것으로 제한되는 것은 아닌 산업 표준 모델링 포맷을 사용해서, 비디오는 모델의 표면상에서 재생된다. 이 방법은 또한 디지털 비디오를 위한 MPEG-4 표준에서 비디오 층을 분리시키는 것에 적합하다. (MPEG은 기술 표준을 정의하는 산업적 연합인 동영상 전문가 그룹을 지칭하는 약자이다.)
- <70> 와이어프레임 망으로 매핑하기 위한 본 명세서에 개시된 방법은 VRML 2.0 표준에 따른다. VRML 2.0에서 표면 맵에 대한 관례는 화상 맵 좌표가 수평축 및 수직축에서 0 내지 1의 크기가 되도록 하는 것이다. 따라서, XYZ로부터 좌표 변형이 수행될 필요가 있다. Z는 생략되고, Z 및 Y는 0과 1 사이의 소수로 변환된다. 이는 화상의 확장 및 배치를 정의하여 화상을 시각내에 놓이도록 한다. 상이한 화상이 중첩되어도, 이들은 반드시 시각내에 존재하고 함께 병합되어야 하기 때문에 문제되지 않는다.
- <71> 본 방법은 또한 다수의 중첩하는 화상들을 취하고, 화상을 함께 스티칭(stitching)하는 추가의 단계 없이 시각내의 화상들을 3D 표면에 적용할 수 있다는 점에서 혁신적이다. 인접한 사진들이 함께 스티칭되어 파노라마를 형성하면, 이 사진들은 보통 수동으로 정렬된 다음 2개의 화상이 혼합된다. 이는 시간을 요하고, 실제로, 경계진 인공물을 가져온다. 본 명세서에 정의된 접근법에 있어서의 중요한 혁신들 중 하나는 이 스티칭을 요구하지 않는다는 것이다. 화상은 모델을 정의하는 동일한 좌표로 매핑된다.
- <72> 페이지 4: 배경의 채움(filling)
- <73> 도 7에 도시된 바와 같이, 오브젝트가 전경으로 끌어 내어지면, 오브젝트는 배경에 빈 공간을 남기고, 이는 상이한 시각에서 보여질 때 보일 수 있다. 이상적으로, 관점이 이동하면, 전경 오브젝트 및 사람들 뒤를 볼 수 있지만 배경에 어떠한 홀(hole)도 인지하지 않는다. 본 명세서에 개시된 방법은 에지를 확대시킴으로써 배경을 채우는 것으로 시작하여, 홀의 중심에 주변 컬러를 끌어당겨 놓는다. 표면이 존재하기 때문에, 상이한 좌표는 간단히 화상을 빈 공간을 덮도록 확대시켜서 원 화상을 더 큰 영역상에 맞추도록 사용된다. 당업자는 본 발명의 사상 및 범위에서 벗어나지 않으면서 상기 설명의 측면에서 변형들이 있을 수 있다는 것을 이해할 것이다.
- <74> 후방부 또는 최상부 및 바닥부가 카메라에 보이지 않는 오브젝트에도 동일한 처리가 적용될 수 있다. 표면을 생성함으로써 홀의 에지를 연결시킬 수 있다. 그 다음, 주변의 화상 세그먼트들이 안으로 확대될 수 있다. 이 부분이 입력 화상에서 더욱 가시화될수록, 더 많은 표면이 부가될 수 있다.
- <75> 페이지 5: 필드의 깊이

<76> 깊이 지각을 향상시키기 위해 전경을 뚜렷하게 하고 배경을 약하게 또는 희미하게 한다. 이러한 종류의 화상 처리를 구현하는 기성 수단뿐만 아니라, 화상 처리에서 에지를 과장하거나 약하게 하기 위해서 콘볼루션 마스크(convolution mask)와 같은 표준 마스크 및 필터링 방법이 존재한다는 것을 당업자는 이해할 것이다. 이는 배경에서 홀을 숨기는 것을 돕고 배경에 대해 요구되는 해상도를 낮춘다. 이는 사용자에 대해 조정가능한 변수이다.

<77> 페이지 6: 네비게이션

<78> 최종 3D 모델이 생성되면, 이 모델이 보여지고 사용될 수 있는 다수의 방법이 존재한다. 네비게이션을 위해, 본 명세서에 개시된 절차는 VRML 2.0과 같은 표준에 따른다. VRML 2.0, OpenGL, 또는 DirectX와 같은 플랫폼에 대해 공공이 사용가능한 표준 요건들을 사용해서 3D 모델링 및 네비게이션을 위해 3D 데이터 및 최종의 비디오 파일을 어떻게 포맷시키는 지는 당업자에게 명백한 것이다.

<79> 또한, 본 명세서에 정의된 기술을 이용해서 3D 모델을 생성하고, 선형 비디오로서 플라이 스루로부터의 일련의 뷰를 저장할 수 있다. 상이한 플라이 스루 또는 재생을 저장함으로써, 예를 들어, 네비게이션 제어가 최소화될 수 있는, DVD 또는 스포츠 방송과 같은 인터페이스상의 소정의 상호적인 선택을 제공할 수 있다.

<80> 본 명세서에 정의된 화상 처리는 배경으로부터 전경 오브젝트를 분리시켜 모션 시차로부터 깊이 지각을 생성하는 것이기 때문에, MPEG-4 비디오의 모델의 사용에도 적합하다. 이러한 방법으로 생성된 데이터세트 및 3D 모델은 MPEG-4의 모델이 기반이 되는 VRML 2.0 표준에 따른다.

<81> 구체적으로 프로페셔널 스포츠에서, 경기장의 중심을 보면서 게임 동안 경기 화면을 앞 뒤 아래로 움직이게 하는 것은 매우 일반적이다. 네비게이션은, 움직임의 방향 및 속도 및 위치로부터 분리시켜 시선의 방향을 제어하도록 요구될 수 있다. 이는 3D 게임에서 선택적일 수 있지만, VRML과 같은 특정 모델링 플랫폼을 위해 시청자 측에서 설정될 수도 있다. 이러한 추가적 뷰잉 파라미터는 상이한 방향에서 플레이를 보면서 경기 화면을 위 아래로 이동시킬 수 있게 하고, 사용되는 카메라의 수 또는 관점에 관계없이 부드러운 움직임으로 수행되게 한다. 본 명세서에 개시된 방법으로서, 카메라의 위치를 알지 않고도 씬에서 네비게이션할 수 있다.

<82> 페이지 7: 측정 교정 및 병합

<83> 페이지 1, 2 및 3은, 깊이식 확대 전경 오브젝트를 합성한 다음 원 화상을 결과적 양각 표면에 텍스처 매핑하여, 모션 시차를 사용해서 비디오 매트릭스를 추출하기 위한 방법을 설명하였다. 임의의 픽셀이 XYZ 좌표 공간에 점으로서 정의되면, 임의의 다른 점으로부터 그 거리를 계산하는 것은 일상적인 수학적 계산의 문제이다. 양호한 실시예에서, 3D 비디오 소프트웨어의 버전은 사용자 인터페이스를 포함한다. 거리 또는 부피와 같은 측정치가 계산될 수 있는 점 또는 오브젝트를 나타내기 위한 수단들이 이 영역에서 사용가능하다.

<84> 또한, 다른 시스템들(예를 들어, 레이저 범위 파인더(finder))로부터의 이전의 측정군으로 병합시킬 수 있기를 원한다. 두 포맷은 데이터 점을 병합시키기 전에 크기 조정될 필요가 있다. 크기 조정을 위해서, 사용자 인터페이스는 또한 기준 오브젝트를 나타내기 위한 표시자, 및 실세계에서 오브젝트의 길이를 입력하기 위한 입력 박스를 포함할 필요가 있다. 길이가 공지된 기준 오브젝트는 목적상 원래의 사진 촬영에 포함될 수 있거나, 또는 씬에서 나타나는 오브젝트에 대해 길이 추정치가 생성될 수 있다. 길이가 씬내에서 크기 조절되면, 모든 데이터 점들은 새로운 유닛으로 변형될 수 있거나, 또는 요구에 따라 변환이 이루어질 수 있다.

<85> 다른 3D 모델과 병합시킬 수 있는 능력은 또한, 정확한 시각에서 눈에 보이는 간접 광고를 통상적인 비디오에 결합시킬 수 있게 한다. 이는 씬의 상업적 오브젝트를 배치시키는 것, 또는 정확한 시각에서 씬의 표면에 그 래픽을 매핑하는 것과 관련된다.

<86> 페이지 8. 스크린상의 홀로그램을 위한 웹 카메라

<87> 비디오에서 시차 움직임을 분석할 때, 웹 카메라, DV 카메라 또는 비디오 폰이 사용되는 경우, 사람의 시야에서 움직임을 추적하기 위해 동일한 알고리즘을 사용할 수 있다. 측면으로 이동하면, 스크린상에 3D 전경 오브젝트의 착시가 나타나면서 스크린상의 오브젝트 주변을 볼 수 있게 할 것이다. 도 2로부터 볼 수 있는 바와 같이, 관점 파라미터는 웹 카메라로 사용자의 움직임을 검출함으로써 변경될 수 있다. 사람이 움직이면, 3D 관점이 따라서 변경된다. 전경 오브젝트는 더욱 적응적으로 이동해야만 하고, 사용자는 자신의 측에서 더 잘 볼 수 있어야 한다. 3D 컴퓨터 게임에서, 사용자에게 의한 좌우 이동은, 제어되는 어떤 종류의 움직임에도 영향을 미치면서 화살 키, 마우스 또는 게임 패드로부터의 입력을 변경시킬 수 있다. 웹 카메라에 의한 모션 검출은 파노라마 포토-VR 씬과 같은 양방향성 멀티미디어의 네비게이션의 방향 및 속도를 제어하는데 사용될 수 있다.

<88> 본 명세서에 개시된 방법은 또한 3D 오브젝트 및 스크린상의 "오브젝트 영화"를 제어하기 위해 특별한 방법을 사용한다. 통상적으로, 예를 들어, 룸에서 네비게이션하는 경우에 좌측으로 이동하면, 스크린상의 움직임도 좌측으로 이동하는 것이 당연한 것이다. 그러나, 전경 오브젝트의 관점에 영향을 미치는 시차로서는, 관점이 좌측으로 이동하면, 오브젝트는 사실적으로 보이기 위해 실제로 우측으로 이동해야만 한다. 이들 제어 유형들 중 하나를 가능하게 하는 일 방법은 선택적 토글(toggle)을 제공하여, 필요한 경우 사용자가 움직임 방향을 반대로 할 수 있도록 하는 것이다.

<89> 페이지 9. 온라인 공유

<90> 본 명세서에 개시된 기술의 중요한 설계 부분은 소프트웨어 자체 및 3D 비디오 출력 모두를 공유하는 매체와 관련된다. 소프트웨어의 설계는 사용자 기반의 급격한 성장 및 고속 온라인 유포를 촉진하는 것을 의미한다. 비디오 플라이 스트루가 저장되면, 상업화된 소프트웨어 개발 키트가, 디폴트로서 폴더를 공유하는데 있어서 자체 플립 ZIP 압축(self-extracting zipped compression)으로 파일 또는 폴더를 저장하는데 사용된다. 이와 동시에, 3D 씬이 저장되면, 소프트웨어의 다운로드 사이트로의 연결이 또한 디폴트로서 씬에 지정될 수 있다. 디폴트는 초기화 동안 또는 소프트웨어 옵션에서 변경될 수 있다.

<91> 소프트웨어는 또한, 구매 후에 일련의 번호가 입력되면 시간 제한 또는 다른 제한을 삭제하는 "업그레이드" 능력을 갖도록 설계될 수 있다. 업그레이드의 구매는 다양한 상이한 유통 방법으로 이루어질 수 있지만, 양호한 실시예는 온라인 쇼핑 카트에서의 자동화된 비용 지불이다. 무료의 선전용 버전 및 업그레이드를 갖는 동일한 설치 시스템이 웹 카메라 소프트웨어에 사용될 수 있다.

<92> 본 명세서에 개시된 방법을 사용해서, 가정내 사용자들은 우선 (i) 비디오 플라이 스트루를 저장하고, (ii) 통상적인 비디오로부터 3D 요소를 추출할 수 있는 능력을 갖게 된다. 대부분의 디지털 매체를 사용해서, 이러한 능력들은 인스턴트 메시징(instant messaging), 이메일, 동등 계층 파일 공유 네트워크, 및 유사한 마찰없는 편리한 온라인 방법들을 통해 공유될 수 있다. 따라서 이 기술은 사전적인 브랜드화된 매체 공유를 가능하게 한다.

<93> 이 기술은 온라인 매체 공유에 대해 상당한 대중적 관심이 존재할 때 개발되었다. 디지털 비디오 기록기와 같은 장치를 사용해서, 가정내 소비자들은 또한 통상적인 끼어들기-기반의 텔레비전 광고들을 회피할 수 있는 능력을 점점 더 갖게 되었다. 이 기술은 또한, 방송 독점으로부터 "비제한된 채널 세계"로 이끌면서, 어떤 사람이라도 자신의 영화를 온라인상에 개봉하기 위해 이용할 수 있다. 따라서, 3D 비디오 소자를 분할, 크기 조정 및 합병할 수 있는 능력은 새롭고 중요한 브랜드화 및 간접 광고의 방법, 및 비디오 제조, 분배 및 웹 방송의 협찬에 대한 새로운 접근법을 제공한다. 또한, 상이한 데이터 스트림이 브랜드화 또는 간접 광고를 위해 사용될 수 있으며, 이는, 예를 들어, 개별화된 인구 통계, 장소 또는 시각에 기초한 가능성을 사용해서 상이한 요소가 동적으로 삽입될 수 있다는 것을 의미한다. 텔레비전, 방송, 비디오 및 웹 방송 협찬의 이러한 새로운 패러다임은 비디오를 3D 요소들로 나누는 기술적 능력을 통해서 가능하게 된다.

<94> 도면 및 명세서에서, 본 발명의 전형적인 양호한 실시예가 개시되고 특정 용어들이 사용되었지만, 이는 단지 일 반적이며 설명적인 의미로서 사용된 것으로서, 제한의 목적은 아니며, 본 발명의 범위는 후속하는 청구항들에 의해 주어지는 것이다.

도면의 간단한 설명

<35> 다음의 상세한 설명은 예시적인 것으로 제시되는 것이며 본 발명을 이것으로만 제한하려는 의도는 아니고, 첨부 도면을 참조해서 더욱 잘 이해될 것이다.

<36> 도 1은 전체 프로세스를 개략적으로 도시한 것으로서, 전경 오브젝트 매트는 배경으로부터 분리되고, 오브젝트가 (상이한 각도로부터 보여질 때) 있는 여백 부분이 생성되고, 와이어프레임이 부가되어 전경 매트에 두께를 제공한다.

<37> 도 2는, 관점이 스크린상에서 이동하도록 하여, 웹 카메라로부터의 피드백에서 사용자의 움직임을 검출하는 본 발명의 소프트웨어로 제어되는 스크린상의 홀로그램을 도시한다.

<38> 도 3은 본 발명의 처리 구성 요소들의 총괄적인 순서도를 도시한다.

<39> 도 4는 3D 모델이 렌더링되는 것으로부터 상이한 시각으로부터의 책상 램프의 2개의 사진을 도시한다.

<40> 도 5는 2개의 사진으로부터 생성된 책상 램프의 3D 모델을 도시한다. 평탄화된 와이어프레임 모델은 좌측에 도시된다. 표면상에 화상이 매핑된 최종 3D 오브젝트는 우측에 도시된다. 오브젝트의 후면 부분은 원 사진에서

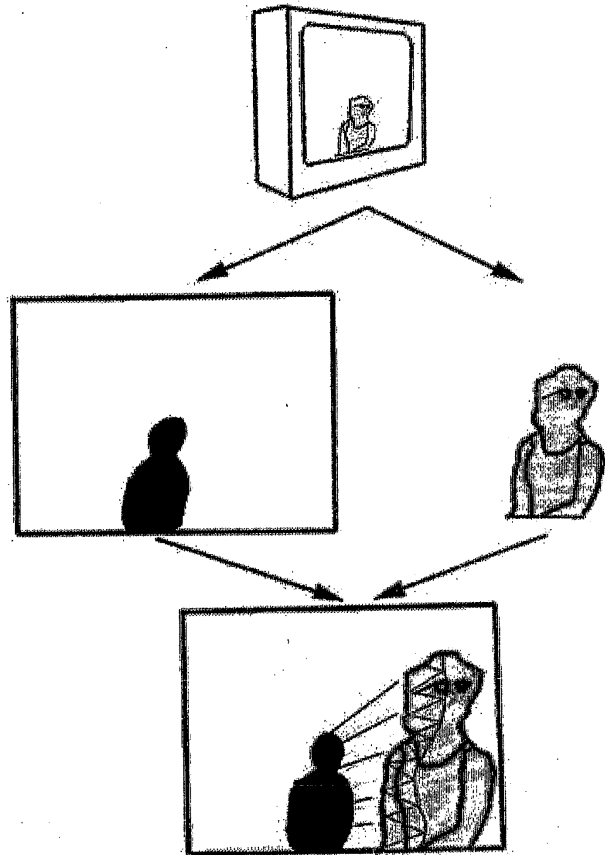
는 보이지 않는 공동(hollow)이지만, 그 표면은 폐쇄될 수 있다.

<41> 도 6은 와이어프레임 망(mesh)을 생성하기 위해, XYZ 좌표점상에 삼각형의 폴리곤을 정의하는 방법을 도시한다.

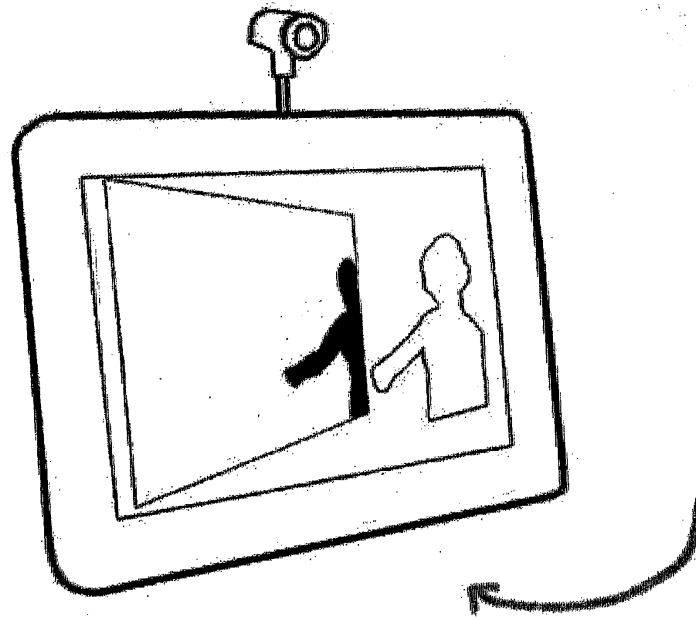
<42> 도 7은 배경상에 그림자를 나타내는 분리된 비디오의 기울어진 뷰를 도시한다.

도면

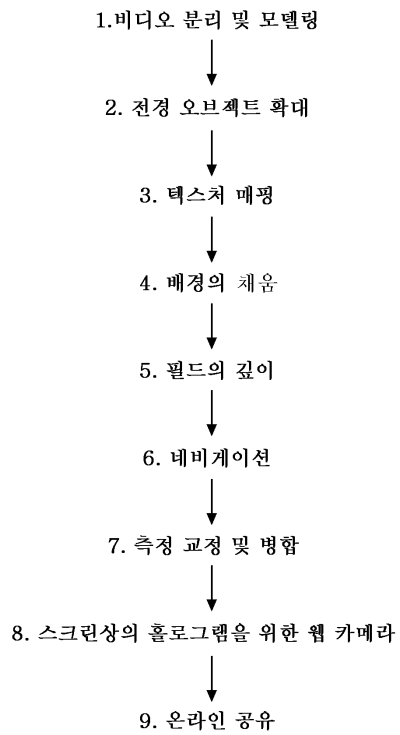
도면1



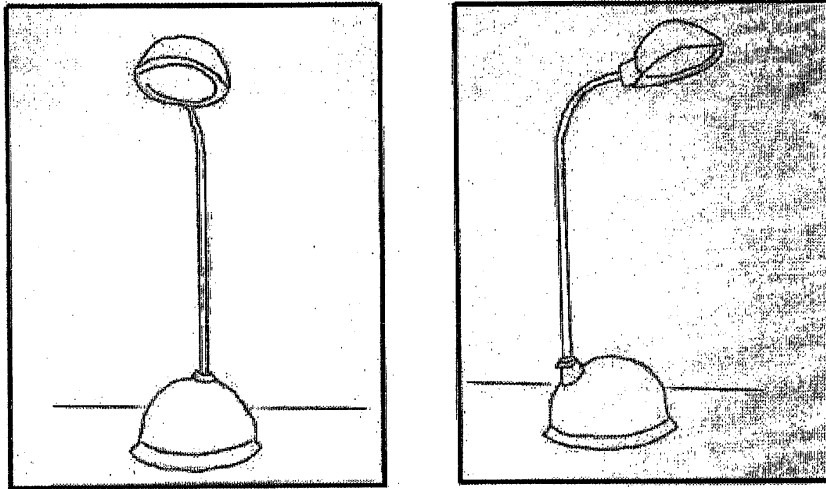
도면2



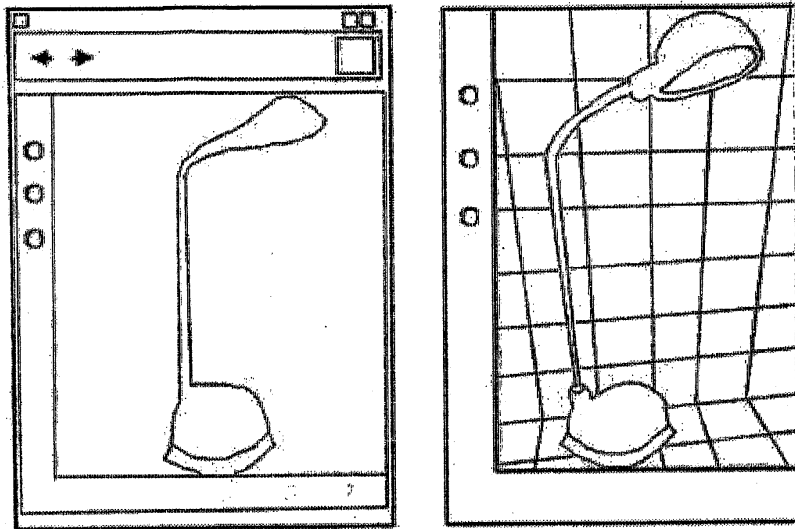
도면3



도면4

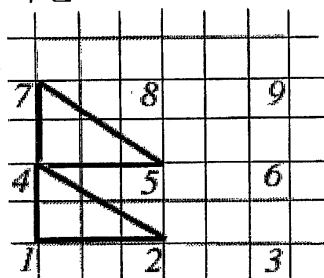


도면5



도면6

픽셀



도면7

