



(19) 대한민국특허청(KR)  
(12) 등록특허공보(B1)

(45) 공고일자 2017년07월24일  
(11) 등록번호 10-1761050  
(24) 등록일자 2017년07월18일

- (51) 국제특허분류(Int. Cl.)  
G06F 3/01 (2006.01) G06F 3/0481 (2013.01)  
G06F 3/0485 (2013.01)
- (52) CPC특허분류  
G06F 3/017 (2013.01)  
G06F 3/04815 (2013.01)
- (21) 출원번호 10-2015-7034957
- (22) 출원일자(국제) 2015년01월06일  
심사청구일자 2015년12월09일
- (85) 번역문제출일자 2015년12월09일
- (65) 공개번호 10-2016-0006751
- (43) 공개일자 2016년01월19일
- (86) 국제출원번호 PCT/EP2015/050091
- (87) 국제공개번호 WO 2015/104257  
국제공개일자 2015년07월16일
- (30) 우선권주장  
14150363.1 2014년01월07일  
유럽특허청(EPO)(EP)
- (56) 선행기술조사문헌  
US20110197263 A1\*  
US20130278504 A1\*  
\*는 심사관에 의하여 인용된 문헌

- (73) 특허권자  
소프트키네틱 소프트웨어  
벨기에 비-1050 브뤼셀 불바드 데 라 플레인 11
- (72) 발명자  
달 조트 데이비드  
벨기에 베-1360 오르베 뒤 뒤 코치계 15/003  
렌뤼 앙투와네  
프랑스 에프-59110 라 마텔레인 에비뉴 루이스 31  
사이먼스 케빈  
벨기에 베-1050 익셀레스 뒤 드 라 발레 29
- (74) 대리인  
양영준, 김현식

전체 청구항 수 : 총 13 항

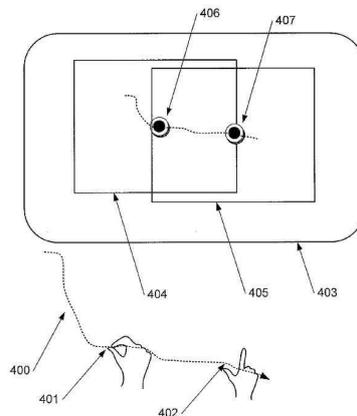
심사관 : 문영재

(54) 발명의 명칭 **인간-컴퓨터 자연스러운 3차원 손 제스처 기반 내비게이션 방법**

**(57) 요약**

인간-컴퓨터 3-차원 손 제스처-기반 자연스러운 상호 작용들을 가능하게 하기 위한 방법이 여기에서 설명된다. 거리 측정 이미징 시스템에 의해 제공된 깊이 이미지들에 대해, 방법은 제어(상호 작용) 제스처가 사용자가 그의 손들을 자유롭게 움직이는 환경에서 동작하는 기간의 한계를 정하기 위해 미리 결정된 유형의 자연스러운 제스처 (뒷면에 계속)

**대표도** - 도4



들의 시작(개시), 및 종료(정지)를 포함한 자연스러운 제스처들의 특정한 시퀀스의 효율적이며 강력한 검출을 가능하게 한다. 본 발명은 특히, 전적으로는 아니지만, 임의의 긍정 오류들도 지연도 없이, 종료하는 손가락 끝 또는 손 끝 기반 자연스러운 제스처가 검출될 때까지 그래픽 사용자 인터페이스 내에서 효율적이며 강력한 내비게이션, 주밍 및 스크롤링 상호 작용들을 제공하도록 개시하는 손가락 끝 또는 손 끝 기반 자연스러운 제스처 다음에 의도적으로 수행된 자연스러운 제스처의 검출에 의해 관심을 가진다.

(52) CPC특허분류

*G06F 3/0485* (2013.01)

---

## 명세서

### 청구범위

#### 청구항 1

3-차원 손 제스처 기반 인식 시스템을 사용하여 자연스러운 인간-컴퓨터 상호 작용(interaction)들을 제공하기 위한 방법에 있어서,

- a) 사용자의 손에서 적어도 두 개의 관심 포인트(point of interest)들을 결정하는 단계;
- b) 상기 적어도 두 개의 관심 포인트들에 의해 수행된 자연스러운 제스처(natural gesture)들을 인식하는 단계로서,

b1) 3-차원 공간내 두 개의 관심 포인트들이 미리 결정된 임계값 미만인 상기 두 개의 관심 포인트들 사이의 거리로 서로 가깝게 이동할 때를 결정함으로써 인간-컴퓨터(human-to-computer) 상호 작용의 시작(beginning)을 트리거하는 3-차원 손 포즈(hand pose)에 대응하는 제 1 자연스러운 제스처를 인식하는 단계;

b2) 상기 3-차원 공간내 적어도 하나의 관심 포인트의 움직임을 결정함으로써 연속적 인간-컴퓨터 상호 작용을 제어하는 상기 3-차원 손 포즈의 움직임들로부터 제 2 자연스러운 제스처를 인식하는 단계로서, 상기 제 1 3-차원 손 포즈가 지속되는 동안은 상기 인간-컴퓨터 상호 작용이 제어되는 단계; 및

b3) 상기 두개의 관심 포인트들이 제 2 미리 결정된 임계값보다 더 큰 상기 두 개의 관심 포인트들 사이의 거리로 서로에게서 멀리 이동할 때를 결정함으로써 상기 인간-컴퓨터 상호 작용의 종료(termination)를 트리거하는 제 3 자연스러운 제스처를 인식하는 단계를 포함하는 상기 자연스러운 제스처들을 인식하는 단계; 및

- c) 상기 관심 포인트들에 의해 수행된 상기 제 1, 제 2 및 제 3 자연스러운 제스처들의 시퀀스를 사용하여 그래픽 사용자 인터페이스(graphical user interface)와 상호 작용하는 단계를 포함하고,

상기 단계 b1)는 대응하는 3-차원 손 포즈가 미리 결정된 3-차원 형상 투시도(shape perspective)에서 검출된 홀(hole)을 포함할 때 상기 제 1 자연스러운 제스처를 확정하는 단계를 더 포함하는 것을 특징으로 하는 자연스러운 인간-컴퓨터 상호 작용들 제공 방법.

#### 청구항 2

청구항 1에 있어서,

단계 b2)는 단계 b1)에 이어 수행되고, 상기 적어도 하나의 관심 포인트의 위치의 연속적 변화를 검출하고 추적하는 단계를 포함하는 것을 특징으로 하는 자연스러운 인간-컴퓨터 상호 작용들 제공 방법.

#### 청구항 3

청구항 1에 있어서,

단계 b1) 및 단계 b3)는 단계 b2)를 위해 사용된 것들과 상이한 관심 포인트들을 사용하는 단계를 포함하는 것을 특징으로 하는 자연스러운 인간-컴퓨터 상호 작용들 제공 방법.

#### 청구항 4

삭제

#### 청구항 5

청구항 1에 있어서,

단계 b3)는 상기 홀이 더 이상 검출되지 않을 때 상기 제 3 자연스러운 제스처의 트리거링을 결정하는 단계를 더 포함하는 것을 특징으로 하는 자연스러운 인간-컴퓨터 상호 작용들 제공 방법.

#### 청구항 6

청구항 1 내지 청구항 3 중 어느 한 항에 있어서,

포인팅 제스처(pointing gesture)를 가능하게 하기 위해 상기 손의 관심 포인트와 적어도 하나의 포인터의 온-스크린 표현(on-screen representation)을 연관시키는 단계;

상기 제 1 자연스러운 제스처가 단계 b1)에서 인식될 때, 상기 포인터의 상기 온-스크린 표현을 그것이 가리키는 그래픽 사용자 인터페이스에서의 이동 가능한 요소, 및 그것이 가리키는 디스플레이된 그래픽 사용자 인터페이스에서의 위치 중 적어도 하나와 로크(lock)하는 단계; 및

상기 제 3 자연스러운 제스처가 단계 b3)에서 인식될 때 상기 온-스크린 표현을 언로크(unlock)하는 단계를 더 포함하는 것을 특징으로 하는 자연스러운 인간-컴퓨터 상호 작용들 제공 방법.

**청구항 7**

청구항 6에 있어서,

상기 적어도 하나의 포인터의 온-스크린 표현이 로크될 때, 단계 b2)는 상기 적어도 하나의 포인터가 상기 제 2 자연스러운 제스처에 대응하는 상기 적어도 하나의 관심 포인트의 위치의 변화에 대하여 로크되는 상기 그래픽 사용자 인터페이스에서의 상기 이동 가능한 요소의 위치의 변화를 야기하는 단계를 더 포함하는 것을 특징으로 하는 자연스러운 인간-컴퓨터 상호 작용들 제공 방법.

**청구항 8**

청구항 1 내지 청구항 3 중 어느 한 항에 있어서,

상기 제 1 및 제 3 자연스러운 제스처들은 각각: 개방 및 폐쇄 손 제스처; 손가락 끝 기반 "클릭킹(clicking)" 제스처; 손가락 끝 기반 "핀칭(pinching)" 제스처; 및 손 끝 기반 "잡기(grabbing)" 제스처 중 적어도 하나의 활성화 및 비활성화에 대응하는 것을 특징으로 하는 자연스러운 인간-컴퓨터 상호 작용들 제공 방법.

**청구항 9**

청구항 1 내지 청구항 3 중 어느 한 항에 있어서,

단계 b2)는 그래픽 사용자 인터페이스의 스크롤링(scrolling); 상기 그래픽 사용자 인터페이스에 대한 주밍 효과; 그래픽 사용자 인터페이스와 연관된 요소의 스크롤링; 및 그래픽 사용자 인터페이스의 변화 중 적어도 하나를 제어하기 위해 상기 제 2 자연스러운 제스처에 대한 적어도 하나의 관심 포인트의 위치의 변화를 이용하는 단계를 포함하는 것을 특징으로 하는 자연스러운 인간-컴퓨터 상호 작용들 제공 방법.

**청구항 10**

청구항 1 내지 청구항 3 중 어느 한 항에 있어서,

단계 b)는 사용자의 손에서 결정된 관심 포인트들을 사용하여 손 제스처들을 인식하는 단계를 포함하며, 상기 인식은: 3-차원 손 골격 모델링(hand skeleton modelling) 기술, 및 3-차원 손 끝들 및 파라미터들 검출 기술 중 적어도 하나를 사용하여 수행되는 것을 특징으로 하는 자연스러운 인간-컴퓨터 상호 작용들 제공 방법.

**청구항 11**

청구항 1 내지 청구항 3 중 어느 한 항에 있어서,

상기 제 1, 제 2 및 제 3 자연스러운 제스처들은 적어도 한쪽 손을 사용하여 수행되는 것을 특징으로 하는 자연스러운 인간-컴퓨터 상호 작용들 제공 방법.

**청구항 12**

청구항 1 내지 청구항 3 중 어느 한 항에 있어서,

단계 b1) 및 단계 b3)는 제 1 손을 사용하여 수행되며, 단계 b2)는 다른쪽 손을 사용하여 수행되는 것을 특징으로 하는 자연스러운 인간-컴퓨터 상호 작용들 제공 방법.

**청구항 13**

청구항 9에 있어서,

그래픽 사용자 인터페이스의 변화는 단계 b2)에서 수행된 상기 제 2 자연스러운 제스처의 방향에 의존하는 것을 특징으로 하는 자연스러운 인간-컴퓨터 상호 작용들 제공 방법.

**청구항 14**

컴퓨터 시스템과 자연스러운 인간-컴퓨터 상호 작용(interaction)들을 제공하기 위한 시스템에 있어서,

3-차원 장면을 캡처하고, 상기 캡처된 3-차원 장면을 나타내는 클러스터된 3-차원 포인트 클라우드를 생성하도록 구성된 3-차원 이미징 시스템; 및

3-차원 손 제스처 기반 인식 시스템을 포함하고,

상기 3-차원 손 제스처 기반 인식 시스템은:

상기 캡처된 3-차원 장면을 나타내는 상기 클러스터된 3-차원 포인트 클라우드를 수신하는 단계;

상기 클러스터된 3-차원 포인트 클라우드로부터 사용자의 손에서 적어도 두 개의 관심 포인트(point of interest)들을 결정하는 단계;

상기 적어도 두 개의 관심 포인트를 연속적으로 추적하는 단계;

상기 적어도 두 개의 관심 포인트에 의해 수행된 자연스러운 제스처(natural gesture)들을 인식하는 단계로서,

a) 3-차원 공간내 두 개의 관심 포인트들이 미리 결정된 임계값 미만인 상기 두 개의 관심 포인트들 사이의 거리로 서로 가깝게 이동할 때를 결정함으로써 인간-컴퓨터(human-to-computer) 상호 작용의 시작(beginning)을 트리거하는 3-차원 손 포즈(hand pose)에 대응하는 제 1 자연스러운 제스처를 인식하는 단계;

b) 상기 3-차원 공간내 적어도 하나의 관심 포인트의 움직임에 결정함으로써 연속적 인간-컴퓨터 상호 작용을 제어하는 상기 3-차원 손 포즈의 움직임들로부터 제 2 자연스러운 제스처를 인식하는 단계로서, 상기 제 1 3-차원 손 포즈가 지속되는 동안은 상기 인간-컴퓨터 상호 작용이 제어되는 단계; 및

c) 상기 두개의 관심 포인트들이 제 2 미리 결정된 임계값보다 더 큰 상기 두 개의 관심 포인트들 사이의 거리로 서로에게서 멀리 이동할 때를 결정함으로써 상기 인간-컴퓨터 상호 작용의 종료(termination)를 트리거하는 제 3 자연스러운 제스처를 인식하는 단계;

를 포함하는 상기 자연스러운 제스처들을 인식하는 단계; 및

상기 관심 포인트들에 의해 수행된 상기 제 1, 제 2 및 제 3 자연스러운 제스처들의 시퀀스를 사용하여 그래픽 사용자 인터페이스(graphical user interface)와 상호 작용하는 단계를 수행하도록 구성되고,

상기 단계 a)는 대응하는 3-차원 손 포즈가 미리 결정된 3-차원 형상 투시도에서 검출된 홀을 포함할 때 상기 제 1 자연스러운 제스처를 확정하는 단계를 더 포함하는 것을 특징으로 하는 컴퓨터 시스템과 자연스러운 인간-컴퓨터 상호 작용(interaction)들을 제공하기 위한 시스템.

**발명의 설명**

**기술 분야**

[0001] 본 발명은 3-차원 손 기반 제스처 인식을 사용한 인간-컴퓨터 상호 작용들에서의 또는 그것에 관한 개선들에 관한 것이다. 본 발명은 보다 특히, 전적으로는 아니지만, 임의의 긍정 오류(false positive)들 또는 지연 없이, 이벤트들을 트리거링하기 위한 어떤 디바이스 시스템 없이 자유롭게 움직이는 손의 연속적 모션에도 불구하고 종래의 그래픽 사용자 인터페이스 내에서 신뢰 가능한 내비게이션, 스크롤링 및 주밍 상호 작용들을 제공하도록 의도적으로 수행되는 자연스러운 제스처들의 특정한 시퀀스들의 실시간 검출과 관련이 있다.

**배경 기술**

[0002] 종래의 인간-컴퓨터 인터페이스들은 키보드들, 마우스들, 원격 제어 및 포인팅 디바이스들과 같은, 하드웨어 제어 시스템 인터페이스들을 포함한다. 이러한 인터페이스들을 갖고, 물리적 동작, 예를 들면, 터치하는 것, 움직이는 것, 잡는 것, 가리키는 것, 누르는 것, 움직이는 것, 클릭하는 것, 또는 심지어 함께, 순차적으로 또는

동시에 복수의 이들 동작들이, 트리거링된 이진 이벤트들 또는 연속적 값들과 같은 제어 명령어들이 인터페이스가 상호 작용하도록 의도되는 컴퓨터 시스템에 전송될 수 있도록 이들 디바이스 인터페이스들에 의해 가능해지는 방식으로, 하드웨어 디바이스 자체 상에서 수행될 필요가 있다.

[0003] 컴퓨터 시스템은 종종 윈도우들, 버튼 및 사용자에게 시각적 피드백을 제공하기 위해 디스플레이되는 다른 아이콘들 또는 요소들을 가진 그래픽 사용자 인터페이스(GUI)를 포함하며; 그것들은 종래의 인간-컴퓨터 하드웨어 인터페이스들의 유용성 및 인체 공학에 따라 설계된다. 예를 들면, 운영 시스템들은 종종, 맵, 이미지 또는 텍스트 박스와 같은 콘텐츠 내에서 내비게이션을 가능하게 하기 위한 스크롤 바들을 포함하는 GUI 윈도우들을 가지며, 그 크기는 디스플레이 스크린 크기 자체의 크기에 의해 한계가 정해진 면적 내에서 디스플레이된 것보다 크다. 스크롤 바들과의 상호 작용은 마우스 하드웨어 디바이스 상에서의 휠을 사용하기 위해, 또는 유지 클릭 동작과 마우스의 모션을 조합함으로써 최적화된다. 또한, 종래의 GUI들은 종종 사용자가 GUI의 콘텐츠의 줌 인 및 아웃을 위해 마우스 버튼들을 갖고 클릭하는 버튼들을 포함한다.

[0004] 게다가, 종래의 GUI들은 또한 종종 보통 마우스 움직임의 기능으로서 맵 스크롤을 만들거나 또는 하나의 맵 영역에서 다른 것으로 변경하기 위해 연속적인 마우스 움직임과 조합된 클릭을 요구하는 맵 내비게이션 GUI 상호 작용들을 포함한다.

[0005] 보다 최근에, 무접촉 상호 작용 시스템들이 이용 가능해져 왔다. 이들 시스템들은 시간에 대하여 장면의 이미지들을 순차적으로 캡처하기 위한 디바이스, 및 캡처된 장면 내에서 사용자에게 의해 수행된 제스처들을 결정하기 위한 방법에 기초한다. 이러한 무접촉 상호 작용 시스템들은 기존의 하드웨어 인터페이스들과 조합하여, 또는 선택적으로 종래의 하드웨어 인터페이스들과 동일하지만 장면의 순차적으로 캡처된 이미지들 내에서 인식된 제스처들로부터 온 제어 명령어들을 트리거함으로써 단독으로 사용되도록 의도된다.

[0006] US-A-2011/0115892에서, 종래의 2-차원(2D) 카메라는 가시 스펙트럼에서의 광을 캡처하기 위해 및 2D 이미지에서, 제스처가 인식될 수 있는 사용자의 손가락을 검출하기 위해 사용된다. 제스처는 컴퓨터화된 시스템과 상호 작용하기 위해 사용될 수 있다. 그러나, 종래의 2D 이미징 기술들의 내재적 한계들, 및 보다 특히 이미징 시스템의 내재적 한계들은 오브젝트들의 검출 및 인식이 변화하는 조명을 가진 환경에서 신뢰 가능하지 않게 만든다. 예를 들면, 펜들과 같은, 손가락-형 오브젝트들은 사용된 기술로 인해 손가락들로서 부정확하게 식별되거나 또는 검출될 수 있거나, 또는 순차적으로 캡처된 이미지들 내에서의 손가락의 추적은 장면 조명에 대하여 이미지 품질의 의존성으로 인해 손실될 수 있다.

[0007] 진보된 2D 이미징 프로세싱 기술들은 사용자의 손의 수용 가능한 검출을 하기 위해 및 몇몇 손 파라미터들 및 그것들의 모션의 검출 및 분석으로부터 매우 간단한 손 제스처 기반 상호 작용들을 제공하기 위해 종래의 2D 컬러 카메라를 사용하는 것을 가능하게 한다. 특히, 2D 배경 학습 기술들 및 후속하는 배경 제거 프로세싱과 조합된 최신식 자동 노출 기술들은 장면으로부터 관심 전경 오브젝트의 보다 효율적인 추출을 제공하며, 상기 장면은 오브젝트 인식이 수행될 수 있는 클러스터들로 분할된다. 이러한 방식으로, 사용자의 손은 피부 컬러 검출 및 형상 특징들로 인해 몇몇 상대적 강건성을 가진 2D 컬러 이미지로부터 추출될 수 있으며, 기본 제스처 인식이 수행될 수 있다. 그러나, 이미징 디바이스에 대한 사용자 부분들의 거리는 상당히 달라질 수 있으며 2D 캡처된 이미지에서의 그것들의 표현은 결과적으로 일관적이지 않게 될 수 있으므로 실제 3D 환경에서 효율적으로 및 정확하게 자유롭게 움직이는 사용자 부분들(예를 들면, 손)을 분석하는 것은 여전히 가능하지 않다.

[0008] US-A-2011/197263은 손 기준 포인트들을 매칭시키며 그것을 기준 손 구성들과 비교하는 것에 의존하는 공간-입력-기반 다중-사용자 공유 디스플레이 세션을 제공하기 위한 시스템 및 방법을 개시한다. 손의 검출된 부분들은 일치하는 손 부분들을 식별하기 위해, 손 부분들을 함께 연결하기 위해, 손 기준 포인트들의 벡터를 생성하기 위해, 및 일치하는 손 구성을 식별하기 위해 손 기준 데이터 저장소에 저장된 기준 손 오브젝트 모델들과 비교될 수 있다. 손 기준 포인트들의 벡터는 일치하는 손 구성을 식별하기 위해 손 기준 데이터 저장소들에 저장된 기준 손 구성들과 비교될 수 있다.

[0009] US-A-2013/278504는 경계 박스가 손 주위에서 효과적으로 제공되는 손 검출 시스템을 개시한다. 손과 연관된 형상 특징들, 예를 들면, 편심, 치밀성, 배향, 직사각형임, 폭 중심, 높이 중심, 최소 박스 각, 최소 박스 폭, 결합들의 수, 경계 박스의 좌측 및 우측 부분들 사이에서의 차이, 및/또는 경계 박스의 상부 및 하부 부분들 사이에서의 차이가 시스템 내에 저장된 데이터와의 비교를 위해 제스처 식별 모듈로 입력된다.

[0010] CN-A-102 236 409는 일단 손이 예지 검출을 사용하여 입력된 이미지에서 식별된다면, 손이 사전 설정된 개시 및 정지 제스처들이 수행되는지를 결정하기 위해 추적되는 손 검출 시스템을 개시한다. 제스처의 결정은 손에서

가시적인 손가락들의 수에 기초한다.

- [0011] US-A-2012/0069168에서, 컬러 정보는 손바닥 중심 및 손바닥의 맨 아래 부분뿐만 아니라 장면으로부터 추출된 손의 마스크를 사용하여 손바닥 중심에서 손의 윤곽 또는 말단들까지의 픽셀들의 거리들과 같은, 상이한 손 관련 데이터 또는 손 파라미터들을 찾기 위해 사용된다. 이들 거리들은 심지어 손이 쥐어졌는지 또는 펼쳐 있는지를 평가하기 위해 사용될 수 있으며, 상기 평가로부터, 손이 오브젝트 "선택" 제스처를 수행하고 있는지가 결정될 수 있다. 그러나, 이러한 제스처-기반 방법들은 한계들을 가지며 그러므로, 컴퓨터와의 상호 작용들과 같은 3D 포인팅을 해결하기 위한 로우-엔드 플랫폼들 상에서의 효율적인 실시간 프로세싱을 제공하는, 강력한 해법을 제공하지도 컬러들이 구별 가능하지 않을 수 있는 어두운 환경에서 이러한 방법들이 동작될 수도 없다. 게다가, 검출된 "선택" 제스처는 실제 객관적인 거리 및 크기 측정들이 없으므로 정확하지 않다. 이미징 디바이스로부터의 다양한 거리들에서 시각적 오브젝트를 정확하게 "포인팅하는 것" 및 "선택하는 것"은 그러므로 불확실하다. 더욱이, 이러한 해법은 3D 공간에서 손의 달라지는 배향의 지원과 같은, 고 신뢰성 레벨 요건도, 손가락-기반 3D 제스처들도 지원하기 위해 사용될 수 없다.
- [0012] 2D 이미지 캡처 해법들은 임의의 상황 또는 환경에서 강력하며 효율적인 손 제스처 인식을 가능하게 하기 위한 강력한 복구 가능하지 않은 한계들을 겪지만, 그것은 그것들이 신뢰 가능한 제어 명령어들 또는 신뢰 가능한 관련 인간-컴퓨터 상호 작용들을 제공하는 추가 도출된 신호 프로세싱 방법들을 가능하게 하기 위해 적절하지 않다는 것이 결과로서 인정되고 있다.
- [0013] 그러나, 3차원 공간에 관한 정보, 즉 깊이는 몇몇 새로운 범위 감지 카메라 시스템들; 3D 카메라들에 의해 최근에 이용 가능해진 데이터이다. 이들 3D 카메라들 또는 3D 이미징 시스템들의 하나의 이점은 특히, 예를 들면, 사용자의 손의, 및 손에 관련된 파라미터들의 개선된 검출 인식 및 추적을 가능하게 하는 절대 거리 메트릭 정보를 보다 정확하게 제공하기 위한 능력이다. 이러한 3차원은 예를 들면, EP-A-2613223에서와 같은, 정적 및/또는 동적 제스처들을 포함한, 공간에서의 3D 위치 및/또는 3D 자세들을 결정하는 것을 추가로 가능하게 하며, 여기에서 사용자의 부분, 즉 손의 말단에 대응하는 적어도 하나의 클러스터의, 및 손 푸시 제스처(push gesture); 손 클릭 제스처(click gesture); 손 풀 제스처(pull gesture); 손 스와이프 제스처(swipe gesture); 손 원 그리기 제스처(hand circle gesture); 또는 간단히 손 3D 위치-기반 포인팅 제스처 중 적어도 하나의 사용자를 포함한 클러스터링된 깊이 맵 내에서, 통상적으로 긴-범위 거리들(예를 들면, 약 3m)에서의 결정을 포함하는 방법이 개시된다. 제스처들은 모호하지 않은 일관성 있는 움직임을 수행하는 손 또는 말단이 검출되어 온 3D 공간의 국한된 면적에서 미리 결정되며 인식된다.
- [0014] 게다가, 이러한 거리 이미징 시스템들(3D 카메라들)은 그 자신의 조명을 제공할 수 있으며 따라서 가시 스펙트럼 대신에 적외선 스펙트럼에서 동작할 수 있다. 이러한 거리 감지 시스템들, 예를 들면, TOF(time-of-flight) 3D 카메라들은, 그건 그렇고, 장면 캡처가 장면의 자연 조명에 더 이상 의존적이지 않으므로 사용자 부분들(예를 들면, 손들)의 부분들의 보다 강력하고, 안정되고, 신뢰 가능하며 정확한 검출의 가능성을 여는 3D 정보를 제공한다. 또한, 오브젝트들 및 거리 측정들의 절대 크기의 제공은 이미지 감지 디바이스로부터 관심 오브젝트들(즉, 손들)의 거리에 대한 계산 전력 의존성(실시간 요건 양상)을 깬다.
- [0015] 손의 3D 차원들 및 3D 위치들의 실시간 측정을 다루는 몇몇 이미지 프로세싱 방법들, 및 연관된 손-기반 3D 제스처 인식 기술들이 이제 점진적으로 3D 카메라들에 의해 제공된 데이터를 사용하기 위해 적용되고 있다. 이러한 방법들은 적어도 깊이 정보, 및 선택적으로, 장면의 조명 레벨에 대응하는 컬러 이미지 및 그레이스케일 신뢰 이미지 중 적어도 하나를 사용한다.
- [0016] 3D 손-기반 제스처 인식의 도메인에서의 성장하는 종래 기술에도 불구하고, 거리 이미징 시스템들에 의해 제공된 데이터로부터 정확하고, 효율적이고(즉, 제한된 프로세싱 전력 요건들을 가진 지연이 없는), 강력하며 신뢰 가능한(즉, 긍정 오류 제스처 검출이 없는) 시스템을 획득하는 것은 여전히 도전이다. 손가락-기반 제스처 인식은 신규 도메인이다. 이것은 제스처 인식 기술들이 적용될 수 있는 사용자의 손의 및/또는 손가락들의 신규 검출 및 추적 기술들; 신규 제스처 인식 기술들 자체들; 및 신규 또는 적용된 인간-컴퓨터 상호 작용들의 조합을 요구하는 문제의 복잡성 때문이다.
- [0017] 특히, 제 1 이슈는 3D 손 제스처 인식 기반 상호 작용들이 현재 대부분 연속적 제어 값들과 조합된 이벤트 트리거들과 같은, 추가 명령어들을 요구하는 터치-기반 상호 작용들 또는 마우스 디바이스 기반 상호 작용들을 위해 설계된 GUI들에 대한 인간-컴퓨터 인터페이스들(HCI들)로서 사용된다. 손 움직임이 효과적으로 연속적인 현상임에 따라, 제 2 이슈는 의도적 상호 작용 및 비-의도적 상호 작용들 사이에서 구별하기 위한 능력과 함께 종래의 GUI들을 가진 3D 손 제스처 인식 시스템들을 사용하려고 노력할 때 발생한다. 훨씬 더 특히, 제 3 및 내재

적 이슈는 적어도 하나의 단일 손의 연속적 모션 내에서, 사용자가 자유롭게 움직이는 환경에서 미리 결정된 자연스러운 제스처의 시작(개시) 및 종료(정지)의 정확하며 신뢰 가능한 검출을 결정할 수 있다는 것이다.

[0018] 이것은 제스처들이 바람직하게는 수행하기에 자연스럽게 직관적이고 매우 용이할 필요가 있을 때, 및 손이 내재적으로 제어 제스처(또는 상호 작용)가 개시하며 정지할 수 있을 때를 표시한 이벤트를 트리거하기 위해 사용될 수 있는 임의의 버튼, 휠 또는 임의의 구체적으로 설계된 작동기를 포함하지 않기 때문에 훨씬 더 문제적이다.

[0019] 하나의 결과는 직관적으로, 지연 없이, 종래의 GUI의 하나의 버튼을 제어하는 것도, 잘해야 맵에서, 제어될 스크롤바들을 가진 애플리케이션의 GUI에서, 또는 명확하게 및 신뢰 가능하게 인식 가능하며 반복 가능한 간단한 제스처에 의해 액세스될 수 있는 복수의 윈도우들을 가진 GUI에서와 같은, GUI 내에서 자연스럽게, 정확하고, 빠르며 간단한 업-다운 및 좌-우 스크롤링을 위한 개방-폐쇄 상태를 포함하는 기존의 3D 손-기반 제스처 인식 시스템을 쉽게 사용하는 것도 거의 불가능하다는 것이다.

[0020] 여기에 설명된 본 발명은 이들 이슈들에 대한 해법을 제공한다.

**발명의 내용**

**과제의 해결 수단**

[0021] 본 발명의 일 양상에 따르면, 3-차원 손 제스처 기반 인식 시스템을 사용한 자연스러운 인간-컴퓨터 상호 작용(interaction)들을 제공하기 위한 방법이 제공되고 있으며, 상기 방법은:

[0022] a) 사용자의 손에서 적어도 두 개의 관심 포인트(point of interest)들을 결정하는 단계;

[0023] b) 상기 적어도 두 개의 관심 포인트들에 의해 수행된 자연스러운 제스처(natural gesture)들을 인식하는 단계로서,

[0024] b1) 3-차원 공간에서 두 개의 관심 포인트들이 서로 가깝게 이동할 때를 결정함으로써 인간-컴퓨터(human-to-computer) 상호 작용의 시작(beginning)을 트리거하는 3-차원 손 포즈(hand pose)에 대응하는 제 1 자연스러운 제스처를 인식하는 단계로서, 상기 두 개의 관심 포인트들 사이의 거리는 미리 결정된 임계값 미만인, 상기 제 1 자연스러운 제스처 인식 단계;

[0025] b2) 연속적 인간-컴퓨터 상호 작용을 제어하기 위해 상기 손의 적어도 하나의 관심 포인트의 움직임들로부터 제 2 자연스러운 제스처를 인식하는 단계로서, 상기 제 1 3-차원 손 포즈가 지속되는 동안은 상기 인간-컴퓨터 상호 작용이 제어되는, 상기 제2 자연스러운 제스처 인식 단계; 및

[0026] b3) 상기 인간-컴퓨터 상호 작용의 종료를 트리거하는 제 3 자연스러운 제스처를 인식하는 단계로서, 상기 제 3 자연스러운 제스처는 상기 제 1 3-차원 손 포즈가 해제되자마자 결정되는, 상기 제 3 자연스러운 제스처 인식 단계를 포함하는, 상기 자연스러운 제스처 인식 단계; 및

[0027] c) 상기 관심 포인트들에 의해 수행된 상기 제 1, 제 2, 및 제 3 자연스러운 제스처들의 시퀀스를 사용하여 그래픽 사용자 인터페이스(graphical user interface)와 상호 작용하는 단계를 포함한다.

[0028] 시간적으로 후속하는 특정 3-차원 손 포즈들의 발생을 검출함으로써, 자연스러운 손 움직임이 의도적인지 또는 비-의도적인지, 즉 제어 제스처가 의도적인지 또는 비-의도적인지, 및 그것이 개시할 때 및 그것이 정지할 때를 결정하는 것이 가능하다.

[0029] 적어도 두 개의 관심 포인트들을 사용하여 3-차원 손 포즈에 대응하는 제 1 3-차원 제스처를 검출하고 결정하며, 상기 제 1 3-차원 손 포즈가 지속되는 동안은 인간-컴퓨터 상호 작용을 제어함으로써, 연속적 상호 작용에 대응하는 제 2 후속 3-차원 제스처를 검출 및 결정하는 것이 가능하다. 상기 제 1 3-차원 손 포즈의 해제(즉, 개시 이벤트를 트리거한 3-차원 제스처)는 정지 이벤트에 대응하는 제 3 3-차원 제스처의 검출을 트리거링하며, 이것은 상기 제 2 후속 3-차원 연속 상호 작용의 정지를 의미한다.

[0030] 부가적으로, 3-차원 손 포즈에 대응하는 제 1 3-차원 제스처 내에서의 홀(hole)의 존재 및 상기 제 1 3-차원 손 포즈의 해체에 해당하는 제 3 3-차원 제스처에서의 홀의 후속 제거를 검출함으로써, 손 움직임이 의도적인지 또는 비-의도적인지를 보다 양호한 강건성을 갖고 결정하는 것이 가능하다. 이것은 인간-컴퓨터 상호 작용을 위해 사용된 자연스러운 3-차원 손 제스처들의 시퀀스의 3개의 주요 단계들, 즉 인간-컴퓨터 상호 작용의 개시, 인간-컴퓨터 상호 작용 자체 및 인간-컴퓨터 상호 작용으로의 종료를 정확하며 신뢰 가능한 검출을 제공한다.

- [0031] 본 발명은, 제 1 양상에서, 사용자의 손에서 파라미터들 및 관심 포인트들(POI들)을 결정하기 위한 방법에 관한 것이며, 이들 POI들은 이에 제한되지 않지만, 손가락 끝들을 포함한다. 파라미터들의 결정은 입력으로서 3D 거리 측정 이미징 시스템에 의해 캡처된 클러스터링 3D 포인트 클라우드(또는 클러스터링된 깊이 맵)를 사용한다. 검출은 적어도 한쪽 손의 3D 위치, 포즈, 모션 및 배향에 관계없이 가능해진다.
- [0032] 본 발명은, 보다 특히, 그것들이 손 끝, 손가락 끝들(즉, 손 및 손가락 말단들), 손바닥 중심 및 상기 말단들 및 상기 손바닥 중심 중에서 복수의 다른 POI들을 사용하여 결정된 POI 중에서 적어도 두 개의 POI들을 포함한다는 점에서 특성화된 손 POI들의 검출 및 추적에 관한 것이며, 이들 POI들 중 두 개는 신뢰 가능한 손가락 끝들 기반 "핀칭(pinching)" 3D 제스처, 신뢰 가능한 손가락 끝 기반 "클릭킹(clicking)" 3D 제스처 및 손 끝 "잡기(grabbing)" 3D 제스처 중 적어도 하나의 검출을 가능하게 하기 위해 선택되고 추적되며, 이들 POI들 중 하나는 손 기반 3D "포인팅(pointing)" 제스처(또는 상호 작용 제스처)를 가능하게 하기 위해 선택되며 추적된다.
- [0033] 본 발명은, 제 2 양상에서, 손가락 끝들 기반 "핀칭", 손가락 끝들 기반 "클릭킹" 및 손 기반 "잡기" 제스처들 중에서 미리 결정된 3D 제스처가 발생하고 해제될 때를 결정하기 위한 및 손 기반 "포인팅" 3D 제스처(또는 상호 작용 제스처)로부터 연속적 제어 명령어들을 결정하기 위한 방법에 관한 것이다.
- [0034] 본 발명은, 제 3 양상에서, 신뢰 가능한 손가락 끝 기반 "핀칭" 3D 제스처, 신뢰 가능한 손가락 끝 기반 "클릭킹" 3D 제스처 및 손 끝 "잡기" 3D 제스처 중 하나의 발생 및 해제의 검출로부터 결정되는 "시작" 및 "종료" 상호 작용들 이벤트들의 순차적 검출과 조합하여 손 기반 연속적 "포인팅" 제스처(또는 상호 작용 제스처)를 사용함으로써 그래픽 사용자 인터페이스의 스크롤링, 주밍 및 내비게이션 피쳐들을 제어하기 위한 3D 손 제스처 기반 인간 대 컴퓨터 상호 작용 프로세스에 관한 것이다.
- [0035] 하나의 유리한 실시예에서, 본 발명은 3-차원 손 및 손가락 3D 제스처들 기반 인식 시스템을 사용하여 자연스러운 인간-컴퓨터 상호 작용들을 제공하며, 상기 방법은 제 1 미리 결정된 3D 제스처(개시 이벤트)가 미리 결정된 인간-컴퓨터 상호 작용의 시작을 트리거하고, 제 2 후속(연속적 상호 작용) 3D 제스처가 인간 대 컴퓨터 상호 작용을 제어한다는 점에서 특성화되며, 상기 제 2 (상호 작용) 3D 제스처는 제 3 미리 결정된 3D 제스처(정지 이벤트)가 인식될 때 경과한다.
- [0036] 일 실시예에서, 단계 b2)는 단계 b1) 다음에 수행되며 적어도 하나의 관심 포인트의 위치에서의 연속적 변화를 검출하며 추적하는 것을 포함한다. 위치에서의 연속적 변화의 이러한 검출 및 추적은 손 내에서 적어도 하나의 POI의 3D 위치의 변화를 검출하며 추적하는 것을 포함할 수 있다.
- [0037] 일 바람직한 실시예에서, 단계들(b1) 및 b3))은 단계 b2)를 위해 사용된 것들과 상이한 관심 포인트들을 사용하는 것을 포함한다. 이 경우에, 제 1 및 제 3 (개시 및 정지 이벤트들)에 대응하는 미리 결정된 3D 제스처들은 제 2 후속(상호 작용) 3D 제스처를 결정하기 위해 사용된 것들과 상이한 손 POI들의 분석으로부터 인식된다.
- [0038] 보다 유리하게는, 일단 미리 결정된 3D(개시 이벤트) 제스처가 인식된다면, 적어도 하나의 POI의 온-스크린 포인터 표현은: 그것이 가리키는 그래픽 사용자 인터페이스에서의 포인팅된 이동 가능한 요소, 및 그것이 가리키는 디스플레이된 그래픽 사용자 인터페이스에서의 위치 중 적어도 하나와 로크되며, 일단 (정지 이벤트) 미리 결정된 3D 제스처가 인식된다면, 온-스크린 표현은 그것이 로크된 그래픽 사용자 인터페이스에서의 위치 또는 요소로부터 언로크된다.
- [0039] 여기에서 사용된 바와 같이 용어들("로크된(locked)" 또는 "로크하는(locking)")은 그래픽 사용자 인터페이스의 이동 가능한 요소와 포인터의 연관을 나타낸다. 대응하는 그래픽 표현은 그것을 언로크한 정지 이벤트의 검출이 전체 GUI 내에서 자유롭게 이동하는 것을 가능하게 할 때까지 그래픽 사용자 인터페이스의 상기 이동 가능한 요소의 공간에서의 위치에 결부시킨다.
- [0040] 여기에서 사용된 바와 같이 용어들("언로크된(unlocked)" 또는 "언로크하는(unlocking)")은 그래픽 사용자 인터페이스의 이동 가능한 요소와 포인트의 연관 해제를 나타낸다.
- [0041] 훨씬 더 유리하게는, 적어도 하나의 포인터의 온-스크린 표현이 로크될 때, 제 2 후속 (연속적 상호 작용) 3D 제스처는 포인터가 (연속적 상호 작용) 제스처를 제어하는 손에서의 관심 포인트의 3D 위치에 대하여 로크된 그래픽 사용자 인터페이스 요소의 위치의 변화를 야기한다.
- [0042] 바람직하게는, (개시 이벤트) 및 (정지 이벤트) 미리 결정된 3D 제스처들은 펴기/쥐기 손 3D 제스처, 손가락 끝

기반 "클릭킹" 3D 제스처, 손가락 끝 기반 "핀칭" 3D 제스처, 손 끝 기반 "잡기" 3D 제스처 중 하나의 발생 및 각각 해제이다.

- [0043] 보다 바람직하게는, 제 2 (연속적 상호 작용) 3D 제스처로서 사용된 POI의 위치의 변화는: 그래픽 사용자 인터페이스의 스크롤링, 그래픽 사용자 인터페이스에 대한 주밍 효과, 그래픽 사용자 인터페이스와 연관된 요소의 스크롤링, 및 그래픽 사용자 인터페이스의 변화 중 적어도 하나를 제어하기 위해 입력한다.
- [0044] 훨씬 더 바람직하게는, 손 제스처 인식들은 3-차원 손 골격 모델링 기술, 3-차원 손 끝들 및 파라미터들 검출 기술 중 적어도 하나에 의해 사용자의 손상에서 결정된 POI를 사용하여 수행된다.
- [0045] 유리하게는, 3개의 후속(개시, 상호 작용, 정지) 자연스러운 제스처들은 적어도 하나의 단일 손을 사용하여 수행된다.
- [0046] 하나의 또 다른 실시예에서, 손에서 결정된 적어도 하나의 POI의 위치의 변화를 검출하고 추적하는 것은 개시 및 정지 이벤트가 트리거링되는 것보다는 다른쪽 손으로 수행된다.
- [0047] 하나의 또 다른 실시예에서, 상호 작용은 그래픽 사용자 인터페이스 윈도우들이 후속 제 2 (연속적 상호 작용) 3D 제스처에 대하여 상대적으로 변화한다는 것이며; 나타나는 윈도우들은 상기 제 2 후속 자연스러운 제스처의 방향의 함수로서 및 상기 제 2 후속 자연스러운 제스처가 미리 결정된 임계치보다 높은 관심 포인트의 위치의 변화를 포함한다면 선택된다.
- [0048] 본 발명의 다른 상세한 특징들은 첨부된 청구항들에 의해 정의되며 이하의 설명에서 지원된다.

**도면의 간단한 설명**

- [0049] 본 발명의 보다 양호한 이해를 위해, 예로서, 첨부한 도면들에 대한 설명이 이제 이루어질 것이다.
  - 도 1은 본 발명에 따른 방법의 3개의 주요 프로세싱 단계들의 흐름도를 예시한다.
  - 도 2는 본 발명에 의해 결정된 바와 같이 손바닥 중심, 엄지 손가락 끝 및 손 끝의 위치들에 관한 POI들, 및 다른 POI들로부터 보간된 POI를 가진 손 기반 "잡기" 자세에서의 손의 측면도를 예시한다.
  - 도 3a 및 도 3b는 개방된 "핀칭" 포즈를 가정한 손의 깊이 맵 및 대응하는 응답 맵을 각각 예시한다.
  - 도 4는 자연스러운 손 및 손가락 기반 3D 제스처들을 가진 스크린상에서의 맵을 스크롤링하기 위해 수행된 연속적이며 자연스러운 손 움직임의 궤적을 예시한다.
  - 도 5는 자연스러운 손가락 끝 기반 개시 및 정지 상호 작용 3D 제스처들을 가진 스크린상에서 디스플레이된 윈도우들을 변경하고/슬라이딩하기 위해 수행된 손의 연속적 자연스러운 움직임의 궤적을 예시한다.
  - 도 6은 손가락 끝 기반 "핀칭" 3-차원 제스처를 수행할 때 두 개의 단계들의 2-차원 투시도를 예시한다.
  - 도 7은 손의 방향으로 손 바닥 중심으로부터 투사된 부근 라인들을 도시하며 직교 레이들이 홀 검출을 가능하게 하기 위해 연관되는 손가락 끝 기반 "핀칭" 3-차원 제스처의 2-차원 투시도를 예시한다.
  - 도 8은 손 바닥 중심을 근사하기 위해 사용된 거리 맵을 예시한다.

**발명을 실시하기 위한 구체적인 내용**

- [0050] 본 발명은 효율적이고, 강력한, 손 파라미터들 기반 3D 제스처들 인식 기술들을 사용하여 종래의 그래픽 사용자 인터페이스들(GUI들)을 제어하기 위한 손 및 손가락 기반 3-차원(3D) 자연스러운 제스처 상호 작용들에 관한 것이다. 그것은 특히, 3D 제스처-기반인 하나의 단일 방법에 기초하여 마우스들 및 터치 제스처 상호 작용들을 위해 원래 설계된 GUI들에서 내비게이션, 주밍 및 스크롤링 피쳐들을 다룬다. 그것은 게다가 자동차 인포테인먼트 (infotainment) (인포메이션(information)/엔터테인먼트(entertainment)) 시스템 제어들, 개인 컴퓨터 시스템 제어들, 태블릿들, 스마트폰들 및 다른 내장된 컴퓨터 제어들과 같은, 단거리 제스처 기반 상호 작용들에 완벽하게 적합하다.
- [0051] 여기에서 본 발명은 특정 실시예들에 대하여 및 특정한 도면들을 참조하여 설명될 것이지만, 이 기술분야의 사람은 본 발명이 이에 제한되지 않는다는 것을 인식할 것이다. 설명된 도면들은 단지 개략적이며 비-제한적이다. 도면들에서, 요소들 중 일부의 크기는 예시적인 목적들을 위해 과장될 수 있으며 일정한 비율로 그려지지 않는다. 또한, 이하에 설명되는 바와 같이 도면들의 특정한 배향들은 비 제한적이며 예로서 도시

된다.

- [0052] 여기에 사용된 바와 같이 용어들("자연스러운 상호 작용(natural interaction)" 또는 "자연스러운 3D 제스처-기반 상호 작용(natural 3D gesture-based interaction)")은 인간이 임의의 하드웨어 디바이스 또는 마커를 핸들링하고, 운반하거나 또는 사용할 필요가 없는 인간-컴퓨터 상호 작용들을 나타내며, 여기에서 상호 작용적 컴퓨터 시스템은 컴퓨터를 직관적으로 제어하기 위해 사용자 손 움직임들에 의해 제공된 정보를, 계산적 수단들을 사용하여 캡처하며 추출하기 위한 3D TOF(time-of-flight) 카메라와 같은 거리 측정 이미징 시스템과 같은, 이미징 수단과 연관되며, 상기 정보는 바람직하게는 미리 결정된 세트의 3D 자연스러운 제스처들의 형태에 있다.
- [0053] 여기에 사용된 바와 같이 용어들("시각적 표현(virtual representation)" 및 "표현(representation)")은 상호 작용 오브젝트의 GUI에서 또는 가상 세계에서서의 디지털 표현에 대응한다. 이러한 상호 작용 오브젝트는 가상 세계 자체로부터의 오브젝트, 예를 들면 사용자가 상호 작용할 수 있는 운영 시스템의 GUI로부터의 폴더일 수 있다. 이러한 상호 작용 오브젝트는 또한 실세계로부터 실제 상호 작용 오브젝트의 가상 세계에서서의 디지털 표현, 예를 들면, 그 위치가 운영 시스템의 GUI와의 포인팅 또는 마우스 포인팅 상호 작용들을 가능하게 하기 위해 시간에 걸쳐 추적되는 사용자의 손과 연관된 포인터 표현일 수 있다.
- [0054] 여기에서 사용된 바와 같이 용어("손 파라미터들")는 적어도 손에 관한 다음의 파라미터들: 손 및 손가락 끝들, 손바닥의 중심("손바닥 중심"으로 칭하여지는); 손바닥의 반경("손바닥 반경"으로 칭하여지는), 손바닥에 대한 법선("손바닥 법선(palm normal)"으로서 칭하여지는), 손바닥의 개방("손바닥 개방" 또는 "손바닥 개방-폐쇄 상태"로서 칭하여지는), 손바닥의 맨 아래 부분, 및 손바닥 중심에서 손 윤곽까지의 거리를 나타낸다.
- [0055] 여기에서 사용된 바와 같이, 용어들("단일 관심 포인트(singular point of interest)" 또는 "단일 관심 포인트들(singular points of interest)")은 사용자의 손상에서의 단일 포인트들, 즉 손의 말단들; 손가락 끝들; 손 끝들; 손의 질량 중심; 손바닥 중심; 또는 그 위치가 손으로부터 획득된 다른 신뢰 가능한 포인트들 중 적어도 하나의 위치를 사용하여 결정될 수 있으며 시간에 대하여 추적될 수 있는 임의의 다른 포인트를 나타낸다.
- [0056] 용어들("관심 포인트(point of interest)" 또는 "관심 포인트들(points of interest)")은 이후 일반적으로 각각 "POI" 및 "POI들"로서 불리운다.
- [0057] 여기에서 사용된 바와 같이 용어("손 끝(hand tip)")는 손의 주요 말단에 대응하는 단일 POI를 나타낸다. 그것은 보다 특히 손가락들의 말단들에 가까운 평균 위치에 위치될 수 있다.
- [0058] 여기에서 사용된 바와 같이 용어("손가락 끝들")는 각각 개개의 손가락 끝, 또는 손가락 말단을 표현한 공간에서의 여러 개의 위치들을 나타낸다. 각각의 개개의 손가락 끝은 통상적으로 단일 POI 또는 POI인 것으로 고려된다.
- [0059] 여기에서 사용된 바와 같이 용어들("포인터(pointer)" 또는 "포인터들")은 POI 또는 선택된 POI의 세트의 디스플레이 스크린상에서의 표현을 나타낸다.
- [0060] 여기에서 사용된 바와 같이 용어들("포즈(pose)" 또는 "자세(posture)")은 특정한 시간 인스턴스에서 물리적 손 자세를 나타낸다. 이것은 손의 특정 POI들의 세트의 위치들의 공간에서의 특정한 배열, 및 특히 서로에 대하여 및 손바닥 중심에 대하여 손가락 끝에 대응하는 POI들의 공간에서의 배열에 대응한다.
- [0061] 여기에서 사용된 바와 같이 용어("제스처" 또는 "3D 제스처")는 시간에 대하여 POI 또는 POI들의 세트의 공간에서의 배열에서의 변화를 나타낸다. 또한 또는 대안적으로, 용어("제스처")는 또한 시간에 대하여 공간(포즈)에서 POI들의 변경되지 않은 배열의 위치의 변화를 나타낸다. 포즈가 시간에 대하여 변하지 않는다면, 즉 공간 내에서의 POI의 배열이 변경되지 않은 채로 있다면, 그것은 정적 제스처로서 설명된다. POI의 배열이 변경되지 않은 채로 있으며 배열의 위치가 시간에 대하여 공간 내에서 변한다면, 이것은 정적 포즈를 가진 동적 제스처로서 설명된다. POI의 배열이 시간에 대하여 변한다면, 이것은 상이한 포즈들을 가진 동적 제스처로서 설명된다.
- [0062] 여기에서 사용된 바와 같이 용어("포인팅")는 "포인팅 제스처"를 나타낸다. 그것은 실세계 사용자 관련 기준 3D 위치, 즉 가상 공간에서, 포인터의 가상 표현이 위치되는 위치를 결정하기 위해 사용될 수 있는, 말단에 대응하는 공간에서의 위치로서, 예를 들면, 손가락 끝들, 또는 연관된 POI 중 적어도 하나를 사용하는 것에 대응한다. 포인터의 가상 표현의 위치의 결정은 예를 들면, 실세계 3D 공간에서 사용자 POI의 절대적 위치 결정을 사용하는 것, 결정된 기준 3D 위치에 대하여 실세계 3D 공간에서 사용자 POI의 상대적 위치 결정을 사용하는 것, 또는 두 개의 POI 사이에서의 벡터 세트의 주요 방향, 예를 들면, 손바닥 중심에서 집게 손가락 끝으로의

벡터에 의해 정의된 방향 또는 눈으로부터 손 말단, 예를 들면 집게 손가락-끝에 대응하는 POI들 중 하나로의 벡터에 의해 정의된 방향에 의해 결정된 상대적 위치 결정을 사용하는 것을 포함하는 여러 개의 방법들에 따라 달성될 수 있다. "포인팅"은 부가적으로, 예를 들면, 적어도 두 개의 POI들의 위치들의 통계적 모드를 사용하는 것을 포함하여, 여러 개의 포인터들을 사용할 수 있다는 것이 이해될 것이다. 통계적 모드는 예를 들면, 평균; 중간값; 최대; 최소 또는 그것들의 위치들의 가중 합 중 적어도 하나일 수 있다.

- [0063] 여기에서 사용된 바와 같이, 용어("상호 작용 3D 제스처")는 바람직하게는 컴퓨터 인터페이스에 의해 가능해진 상호 작용을 제어하기 위해, 예를 들면 포인터 또는 스크롤 바를 제어하기 위해 공간에서의 위치에 대하여 POI의 상대적 위치 결정을 사용한 "포인팅" 제스처와 같은, 연속적 제스처를 나타낸다.
- [0064] 여기에서 사용된 바와 같이 용어들("3D 제스처 개시"(개시 이벤트) 및 "3D 제스처 정지"(정지 이벤트))은 하나의 미리 결정된 자연스러운 3D 제스처가 수행되거나 또는 각각 해제된 것으로서 인식될 때 발생하는 이벤트들을 나타낸다.
- [0065] 여기에서 사용된 바와 같이, 용어들("잡기(grabbing)", "핀칭(pinching)" 및 "클릭킹(clicking)" 제스처들)은 바람직하게는 적어도 두 개의 POI들의 모션 및 포즈 분석으로부터 개시 또는 정지 제스처 기반 이벤트들을 트리거하는, 손 파라미터들을 사용하여 수행된 3D 자연스러운 제스처들을 나타낸다.
- [0066] 인체 공학 고려 사항들에 대하여, "핀칭" 제스처는 손가락 끝의 것들 중에서 선택된 적어도 두 개의 POI들이 서로 접촉하기 시작할 때까지 함께 더 가까워지도록 손의 폐쇄를 나타낸다. 두 개의 POI들 사이의 거리 또는 상대적 거리는 부가적으로 "핀칭" 제스처의 레벨 또는 양과 연관된 값을 결정하기 위해 사용될 수 있다. "핀칭 제스처"는 바람직하게는 엄지 손가락 끝 POI 및 집게 손가락 끝 POI의 사용을 포함한다. "핀칭" 제스처 검출은 하나의 선호된 투시도 내에서, 이용 가능하다면, 손의 형상 내에서, 홀의 이용 가능성을 부가적으로 검출함으로써 개선될 수 있다.
- [0067] "잡기 제스처(grabbing gesture)"는 제 1 POI가 엄지 손가락 끝이며 제 2 POI가 손 끝인 "핀칭 제스처(pinching gesture)"의 변형이다.
- [0068] "클릭킹 제스처(clicking gesture)"는 그 지속 기간이 시간 지연 임계치의 함수로서 결정될 수 있는, 매우 짧은 시간 기간 동안 유지되는 손 포즈의 검출과 같은, 산발적 이벤트에 대응한다는 점에서 특성화된 제스처이다. "클릭킹 제스처"는 보다 정확하게 고려된 두 개의 POI들 사이에서의 접촉이 매우 짧은 시간 기간들 동안 유지된다는 점에서 특성화된 이전 설명된 "핀칭" 및 "잡기" 제스처들에 대응하며, 지속 기간은 미리 결정된 시간 지연 임계치보다 낮지만, 이에 제한되지 않는다. "클릭킹 제스처"는 대안적으로 또한 하나의 단일 POI, 예를 들면 집게 손가락 끝 또는 손 끝을 갖고 수행될 수 있어서, 미리 결정된 거리 내에서 및 미리 결정된 시간 기간 내에서 손의 두 개의 주 축들(주 성분 분석(PCA)에 의해 결정된 바와 같이)에 대하여 위 아래로 한 번 빠르게 진동한다.
- [0069] 본 발명은 일반적으로 개선된 무접촉 자연스러운 인간-컴퓨터 상호 작용들을 제공하기 위한 방법, 및 보다 특히 직관적으로 및 용이하게 순차적으로 수행될 3D 제스처들을 사용하여 GUI들을 제어하기 위한 방법에 관한 것이다. 제스처들은 컴퓨터 시스템과 연관된 3D 카메라와 같은, 거리 이미징 시스템에 의해 시간에 대하여 순차적으로 캡처된 사용자의 손 움직임들에 의해 공간에서 실행된다. 여기에서 설명된 방법에 따라 동작하는 컴퓨터 시스템, 3D 카메라, 및 GUI는 무접촉 자연스러운 인간-컴퓨터 상호 작용 시스템을 형성한다. 거리 측정 카메라 디바이스에 대하여 사용자의 위치 결정은 상호 작용하는 손 및 상기 손의 POI가 거리 측정 카메라 디바이스의 이미징 센서로 이미징될 수 있는 한 중요하지 않다.
- [0070] 본 발명에 따르면, 손 제스처 정보는 거리 측정 카메라 디바이스에 의해 제공된 데이터로부터 추출되며, 상기 데이터는 적어도 픽셀 당 깊이 측정, 및 선택적으로 장면 조명 및/또는 컬러 조명을 포함하는 3D 포인트 클라우드(또한 꼭지점들의 세트(set of vertices)로서 불리우는)의 또는 깊이 맵의 형태에 있다.
- [0071] 본 발명의 실시예에 따르면, 도 1은 방법의 3개의 주요 프로세싱 단계들, 즉 상호 작용하는 손 파라미터들이 관련 있는 POI들을 정의하기 위해 시간에 걸쳐 검출되고, 식별되며, 추적되는 제 1 단계(단계 100), 3D 제스처들이 고려되는 관련 있는 POI들의 움직임 및 포즈로부터 인식되는 제 2 단계(단계 101), 및 단계(101)로부터 미리 결정된 인식된 3D 제스처들의 세트가 GUI와의 상호 작용을 제어하기 위해 사용되는 제 3 단계(단계 102)의 흐름도를 예시한다. GUI와의 상호 작용의 제어는 단계(100)에서 검출될 POI들의 식별 또는 수를 변경할 수 있는 피드백 루프(103), 또는 수를 변경할 수 있는 또 다른 피드백 루프(104)를 더 포함할 수 있다.
- [0072] 제 1 단계, 단계(100)는 그로부터 여러 개의 관련 있는 POI, 즉 특히 POI 또는 상기 POI가 3D 제스처들을 결정

하며 GUI와의 추가 상호 작용을 가능하게 하기 위해 선택될, 손 파라미터들의 세트를 검출하고, 식별하며 추적하는 것을 목표로 한다. 손 파라미터들은 바람직하게는, 전적으로는 아니지만, 적어도 하나의 안정된 및 신뢰 가능한 POI, 및 바람직하게는 두 개를 제공할 수 있도록, 손 끝들, 즉 손 끝 및 손가락 끝들을 포함할 수 있다. 부가적으로, 단계(100)는 또한 다음의 파라미터들 중 적어도 하나의 결정을 포함할 수 있다: 손바닥 반경; 손 개방; 및 손바닥 법선. 게다가, 이들 결정된 손 파라미터들로부터 적어도 하나의 신뢰 가능한 POI를 사용하여, 상기 방법은 GUI와 상호 작용하는 포인터와 연관되기 위해 적어도 하나의 POI를 결정하는 단계를 더 포함한다.

[0073] 본 발명의 제 1 단계는 사용자의 적어도 한쪽 손을 포함한 깊이 맵 또는 대응하는 3D 포인트 클라우드(즉, 꼭지점들의 세트)의 형태로 입력 정보를 사용한다. 깊이 맵은 이 기술분야에 알려진 방법에 따라 여러 개의 클러스터들로 분할되며, 클러스터들의 각각은 그에 따라 라벨링된다. 사용될 식별된 손(들)에 대응하는 라벨(들)이 또한 제공된다. 상호 작용하는 손은 임의로 또는 공간에서 그것의 위치 또는 그것의 움직임들로부터 자동으로(예를 들면, 모션을 보여주는 라벨링된 손이 선택될 수 있다) 결정될 수 있다(예를 들면, 손들 중 하나의 선택, 좌측 손). 특히, 단지 상호 작용하는 손에 대응하는 라벨링된 3D 포인트들만이 다음의 프로세싱 단계들에서 고려될 것이다.

[0074] 하나의 제 1 양상에서, 단계(100)는 먼저 그 출력이 사용될 식별된 손의 거리 맵인 "거리 변환"의 이전 계산을 요구하는 손의 손바닥 중심의 검출 및 결정을 포함할 수 있다. 손바닥 중심의 검출은 의무적이지 않지만, 그것은 그러나 포인팅 제스처 상호 작용들을 위한 기준으로서 사용될 비교적 안정된 POI를 제공하기 위한 유리한 모드이다. 손의 거리 맵은 손의 각각의 픽셀이 도 8에 예시된 바와 같이, 손의 가장 가까운 경계로의 실세계 메트릭 시스템에서의 거리와 같은 값을 갖는 이미지이며, 여기에서 사용될 라벨링된 손은 사용자의 임의의 다른 쓸모없는 라벨링된 부분들로부터 및 배경으로부터 분할된다. 예시에서, 손 표현에서의 각각의 픽셀의 콘텐츠가 어두울수록, 이 픽셀은 라벨링된 손의 중심으로부터 더 멀리 떨어진다. 이러한 거리 맵 이미지는 손의 이전 라벨 이미지에 대한 두 개의 패스들을 사용하여 계산될 수 있다. 여러 개의 방법들이 거리 변환을 수행하기 위해 사용될 수 있다. 그러나, 본 발명에서의 하나의 중요한 포인트는 라벨 이미지의 각각의 픽셀이 3D 포인트 클라우드에서의 위치에 대응하므로 및 거리 맵에서의 각각의 거리가 실세계에서의 거리 메트릭에 대응함에 따라 각각의 픽셀에 대해 획득된 거리 값이 실세계에서 메트릭 거리에 관련될 수 있다는 것이다. 예를 들면, 거리 변환은 제 1 패스가 이전 라벨 이미지의 상부-좌측 코너에서 하부-우측 코너로이며, 제 2 패스는 반대 방향인 2-패스 프로세스를 사용하여 수행될 수 있다. 각각의 패스에 대해, 거리들은 패스에 의해 이미 프로세싱된 직접 이웃하는 픽셀들의 모두의 최소 값들의 합에 대응하는 각각의 픽셀에 대해 누적된다. 예를 들면, 제 1 패스에서, 픽셀에 대한 직접 이웃은 좌측으로, 좌측 대각으로, 상부로 및 우측 대각으로의 픽셀일 수 있다. 현재 픽셀에 대해 결정된 거리는 4개의 이웃하는 픽셀들 더하기 1의 최소 값인 것으로서 결정될 수 있다. 더욱이, 픽셀들의 기본 기하학적 속성들을 고려하기 위해, 수평/수직 이웃 픽셀들의 값은 미리 결정된 양만큼, 예를 들면, 5의 값만큼 가중될 수 있는 반면, 대각 이웃 픽셀들의 것들은 또 다른 미리 결정된 가중 양만큼, 예를 들면 7의 값만큼 가중될 수 있다.  $7/5$  비는 통상적인 정방형 픽셀( $7/5 = 1.4 \approx \sqrt{2}$ )의 대각 및 측면 사이에서의 보존 비에 대응한다. 이러한 비의 사용은 제곱 근 계산들이 훨씬 더 많은 프로세싱 유닛 리소스들을 요구하므로 계산 최적화에서의 바람직한 실시예이다. 게다가, 개선된 프로세싱 효율을 획득하기 위해, 그 경계가 깊이 맵에서의 고려된 손에 대응하는 라벨링된 3D 포인트들의 최대 및 최소, 좌측, 우측, 상부 및 하부 위치들에 의해 정의되는 전체 깊이 맵의 제한된 부분이 프로세싱될 것이다. 손바닥 중심 위치를 근사하기 위해, 프로세스는 그 후 적어도 거리 정보 강도가 손의 모든 다른 픽셀들의 최대 값인 픽셀 위치를 결정하는 것을 포함하며, 바람직한 실시예에서, 이전 결정된 손바닥 중심으로의 거리를 최소화하는 것은 이전 결정된 손바닥 중심 위치의 위치에서의 평활한 변화들을 가능하게 하기 위해 고려될 수 있다. 이 기술분야로부터의 3D 안정화 기술들은 부가적으로 그렇게 근사된 손바닥 중심 위치를 개선하기 위해 사용될 수 있다.

[0075] 단계(100)는, 또 다른 양상에서, 또한 예를 들면 손의 3D 포인트 클라우드에 대해 수행된 PCA를 사용한 손 끝의 3D 위치의 검출 및 결정을 포함한다. 그 기원이 손바닥의 중심인 것으로 설정되며, 그 방향이 PCA에 의해 주어 진 주요 방향인 것으로 설정되는 3D 벡터를 사용하여, 히스토그램은 손의 주요 방향을 정의한 벡터의 각각의 섹션에 대해, 주요 방향에 직교하는 또 다른 벡터의 방향에 따라, 상기 주요 벡터 축으로부터 손의 경계로 픽셀들(또는 3D 포인트들)의 수를 카운팅하도록 계산된다.

[0076] 주요 방향을 표현한 벡터의 각각의 섹션은 예를 들면, 1mm 스텝의 거리 메트릭에 의해 표현될 수 있다. 주요 방향에 대한 직교 벡터는, 카메라 좌표 시스템의 두 개의 미리 결정된 축 및 손의 두 개의 주 방향들에 의해 형성된 두 개의 축에 의해 각각 형성된 평면들에 의해, 또는 생성된 히스토그램이 최고 에너지(즉, 적분 값이 최

고인 포인트)를 보여주는 정규 직교 방향을 선택함으로써 형성된 각도를 최대화하기 위해 카메라 위치의 기하학적 구조 대 손의 주요 방향에 대하여, PCA에 의해 획득된 손의 제 2 주 방향으로부터 임의로 선택될 수 있다. 그렇게 결정된 히스토그램을 사용하여, 손 끝은 그 후 손 픽셀들의 최소 수가 투사된 주요 벡터의 섹션에 대응하는 면적인 것으로서 결정될 수 있다. 일 바람직한 실시예에서, 덜 투사된 픽셀들을 가진 히스토그램의 섹션에 대응하며, 그에 따라 손 끝인 것으로 결정되는 공간에서의 손의 주요 방향을 따르는 손의 가장 두꺼운 섹션에 대응하는, 두 개의 미리 결정된 임계값들 사이에서의 범위, 예를 들면, 95% 내지 99% 범위가 정의될 수 있다. 이 기술분야에 알려진 3D 안정화 기술들은 부가적으로 근사된 손 끝 위치를 개선하기 위해 사용될 수 있다.

[0077] 단계(100)는 그 후 입력 깊이 맵(301)(도 3a에 대하여 이하에 설명될 바와 같이)에서의 손 포즈에 관계없이 손 끝들의 3D 위치들의 검출 및 결정을 포함한다. 손 끝들은 손의 말단들에 대응하며 손가락 끝들을 근사시키는 면적들을 포함한다. 이러한 검출을 완료하기 위해, "응답 이미지"(302)가 결정될 수 있다(도 3b를 참조하여 이하에서 설명될 바와 같이). 응답 이미지에서, 픽셀의 값은 3D 포인트들이 표현되는 기준 3D 좌표 시스템에 대하여 각각의 고려된 픽셀 및 그것의 8개의 이웃들 사이에서의 깊이의 차이를 나타낸다. 높은 값은 말단인 높은 확률을 표시한다. 각각의 픽셀(3D 포인트)에 대한 값은 공통의 8개의 방향들(즉, 위, 위-좌측, 우측, 아래-우측, 아래, 아래-좌측, 좌측, 위-좌측)을 따르는 8개의 이웃들로부터 계산된다. 각각의 방향에서, 현재 포인트의 2cm의 거리에 있는 픽셀(3D 포인트)이 고려된다. 어떤 픽셀이 각각의 방향에서 2cm의 거리의 투사에 대응하는지를 찾아내기 위해, 이미징 디바이스 또는 카메라의 규격이 현재 포인트의 깊이 값과 함께 사용되거나, 또는 기준 3D 좌표 시스템을 사용하여 간단히 결정된다. 이러한 동작은, 물론 단지 유효 픽셀들, 즉 미리 결정된 임계치보다 높은 신뢰도 값을 가진 픽셀들(3D 포인트들)에 대해 행해지며, 상기 신뢰도 값은 픽셀에서 수신된 적외선(IR) 조사의 양을 나타낸다. 이들 8개의 픽셀들에 대해, 현재 픽셀을 갖는 3D 위치에서의 차이가 결정되며, 현재 픽셀의 값은 제 3 최소 값인 것으로 취해진다. 예를 들면, 도 3a 및 도 3b에 대응하는 일 실시예에서 손이 카메라에 접하고 있을 때, 3D 위치에서의 차이는 대부분 깊이에서의 차이에 대응할 것이다. 제 3 최소 값은 임의의 비-말단이 일 측면(예를 들면, 손바닥의 측면) 상에서 연결되지 않음으로써 큰 값을 갖는 것이 용이하며 손의 이들 부분들이 손가락으로서 응답을 제공하지 않는 것이 요구되므로 선택된다. 그러나 손가락은 그것들이 손에 연결되는 몇몇 방향을 가질 수 있다. 이러한 이유로, 두 개의 최소 응답 값들을 가진 두 개의 방향들이 무시된다. 몇몇 손 포즈들은 손의 각각의 물리적 말단에 대한 적절한 응답을 제공하지 않을 수 있으며, 예를 들면, 닫힌 주먹이 손 끝들에 대응하는 하나의 단일 응답을 제공할 것임이 주의된다. 본 발명은 그러나 상호 작용들이, 두 개의 POI들(또는 말단들)이 이용 가능하며, "포인팅" 제스처가, 만약에 있다면, 손바닥 중심의 추정된 POI 위치로부터 수행될 수 있는 미리 결정된 제스처들의 세트에 의존하기 위해 설계되므로 이러한 확률에 의해 영향을 받지 않는다.

[0078] 더욱이, 응답 이미지 및 깊이 맵 이미지에 기초하여, 연결된 구성요소 이미지가 결정될 수 있다. 이것은 그것들 모두가 응답 이미지에서 긍정 응답 값을 가지는 경우에만, 즉, 그것들이 말단들이며, 그것들이 깊이에서 충분히 가까운(예를 들면, 1cm 내) 경우에만 두 개의 인접한 픽셀들(또는 3D 포인트들)에 대해 동일한 라벨을 설정하는 것을 가능하게 하는 라벨 이미지에 대응한다. 다시 말해서, 두 개의 이웃하는 픽셀들(또는 3D 포인트들)은 그것들이 미리 결정된 임계치, 예를 들면, 1cm보다 큰, 거리 값 차이, 예를 들면, 도 3에 따른 깊이 차이를 갖는다면, 또는 그것들 중 하나가 말단 응답 이미지에서의 긍정 응답 값을 갖지 않는다면 동일한 라벨을 갖지 않을 것이다. 이러한 이미지는 그 후 두 개의 목적들을 위해 사용될 수 있다: 첫 번째는 유일한 손 말단들이 손가락 끝들인 것처럼 손가락 끝 후보들을 형성하기 위해 말단 응답들을 함께 그룹핑하는 것을 허용하는 것이며; 두 번째로 이미지로부터 원치 않는 작은 응답들을 제거하는 것이다. 특히, 구성요소 크기가 미리 결정된 임계치, 예를 들면, 1cm보다 작다면, 그것은 거절될 수 있다.

[0079] 본 발명의 방법은 손바닥 중심, 손 끝, 손가락 끝, 또는 그 위치가 이들 파라미터들의 위치들의 조합으로부터 보간되는 포인트 중에서, 이전 결정된 손 파라미터들 중 적어도 하나를 선택함으로써 POI를 추가로 결정하는 단계를 포함한다. 바람직하게는, 두 개의 POI들은 본 발명의 특정 상호 작용들에 따라 결정되어야 한다. POI들의 수 또는 식별은 GUI에 의해 요구된 상호 작용에 의해 제어된 피드백 루프(103)에 의해 설정될 수 있다. 특히, POI들을 결정하는 것은 손 끝들에 대응하는 라벨링된 응답들을 사용하도록 요구한다. 이것은 하나에서 다른 것들로의 거리에 기초한 응답 이미지에서 각각의 라벨링된 손 끝 응답(즉, 구성요소)에 대한 가중 값을 계산함으로써 달성될 수 있다(구성요소가 더 많이 분리될수록, 가중치는 더 크다). 이러한 방식으로, 적어도 두 개의 POI들을 사전 선택하는 것은 두 개의 최대 가중치들을 가진 두 개의 구성요소들을 선택함으로써 간단히 수행될 수 있다.

- [0080] 본 발명의 일 바람직한 실시예에서, 응답 이미지에서의 모든 구성요소 중에서 중심 포인트가 결정될 수 있다. 손 구성요소들의 모든 3D 포인트들의 간단한 가중 합이 가중 평균으로서 취해진다면, 획득된 중심 포인트는 가시적인 보다 많은 손가락들이 있으며 따라서 보다 강한 응답들이 유일한 엄지 손가락 끝 응답이 위치되는 하부 영역에서보다 상부 영역에서 획득될 것이므로 손의 상부 부분을 향해 바이어싱될 것이다. 이러한 이슈를 회피하기 위해, 각각의 구성요소의 가중은 응답이 속하는 구성요소의 식별에 따라 개선된다. 구성요소가 모든 다른 구성요소들(예를 들면, 엄지 손가락 끝에 대응하는 구성요소)로부터 멀리 떨어져 있다면, 그것은 보다 큰 가중을 가질 것이며, 그것이 다른 구성요소들(예를 들면, 손가락 끝)에 가깝다면, 그것은 보다 작은 가중을 가질 것이다. 이것은 분리된 손가락(대부분의 경우들에서, 엄지 손가락)이 그것들이 가시적일 때 손가락들의 나머지보다 높은 가중을 얻는다는 것을 보장한다. 이것은 또한 적어도 손 끝에 비교하여 엄지 손가락 끝 또는 엄지 손가락의 끝, 및/또는 그것들의 위치 대 상기 엄지 손가락 및/대 손바닥 중심 위치 및 PCA로부터 획득된 손 주 방향들에 대하여 개별적으로 자체 식별될 수 있는 다른 손가락 끝들의 보다 강력한 식별을 보장한다.
- [0081] 이를 위해, 구성요소 사이에서의 거리는 예를 들면, 5의 미리 결정된 양으로 곱하여진 손 반경으로서 계산되는 손 너비(span)에 의해 정규화된다. 모든 말단들의 중심 포인트는 이전 정의된 구성요소 가중을 사용하여 가중 합으로서 계산된다. 일단 중심 포인트가 결정된다면, 그것은 3D 데카르트 공간에서의 구성요소를 구별하기 위해 사용된다. 그렇게 하기 위해, 결정된 중심 포인트 및 손 바닥 중심을 통과하는 평면이 사용된다. 목적은 바람직하게는 평면의 일 측면 상에서 손가락을 및 다른 측면 상에서 다른 손가락들을 갖는 것이므로, 중심 포인트 및 손 바닥 중심에 의해 형성된 축에 대한 평면의 배향은 평면에 대한 두 개의 주 구성요소들의 중심 포인트들의 거리를 동시에 최대화하는 것으로서 계속해서 결정될 수 있다. 그렇게 결정된 평면을 고려해볼 때, 손의 포인트들은 평면의 그것들의 측면에 따라 쉽게 분류될 수 있다. 게다가, 검출된 이용 가능한 구성요소들의 수, 그것들 각각의 크기 및 위치에 따라, 엄지 손가락 끝 구성요소, 손 끝 구성요소, 또는 다른 손가락 끝 구성요소 중 적어도 하나일 수 있는 이용 가능한 구성요소의 아이덴티티를 정확하게 결정하는 것이 가능할 것이다. 두 개의 "POI들"을 선택하는 것은 그 후 상호 작용을 위해 요구되며 단계(103)에서 피드백 루프에 의해 제공될 수 있는 구성요소들의 식별에 따라 실행될 수 있다.
- [0082] 또한, 본 발명이 자연스러운 3D 제스처들을 인식하기 위해 두 개의 POI들을 검출하며 추적하는 것을 요구한다면, 하나의 단일 POI는, 3D 제스처가 "포인팅" 상호 작용 제스처라면 충분할 수 있다. 이러한 단일 POI는 손 바닥 중심인 것으로 설정될 수 있다. 제 1 POI는 엄지 손가락 끝 POI이며 제 2 POI는 손 끝 POI 또는 집게 손가락 끝 POI와 같은 하나의 다른쪽 손가락 끝 POI인, 결정된 두 개의 "POI들"을 가진, 바람직한 실시예에서, 두 개의 POI 사이에서의 부가적인 안정된 가상 POI가 이 기술분야에 알려진 보간 기술로부터 계산될 수 있다. 가상 POI는 손바닥 중심의 움직임에 의해 조정되며 "포인팅 제스처들"을 가능하게 하기 위해 사용된 "POI들"의 중심일 수 있다.
- [0083] 도 2에서, "잡기" 자세에서의 손의 측면도가 손 끝 POI(201), 엄지 손가락 끝 POI(202), 손바닥 중심 POI(203)의 위치 결정, 및 부가적인 안정된 가상 POI(204)의 위치 결정과 오버레이되어 도시되며, 그 위치는 다른 앞서 언급된 POI를 사용하여 계산된다. 특히, 이러한 가상 POI의 위치는 단지 손바닥 중심이 미리 결정된 거리, 예를 들면 3mm 이상 이동하는 경우에만 업데이트될 수 있다. 이것은 본 발명의 하나의 유리한 양상을 보장하며, 즉 사용자는 그/그녀가 동시에 "핀칭 제스처들", "클릭킹 제스처들" 또는 "잡기 제스처들"을 수행해야 할지라도 그/그녀의 손가락들 끝들과의 상호 작용(즉, 포인팅)을 자연스럽게 및 편안하게 느낄 것이다. 더욱이, 손바닥 중심의 것들과 가상 POI의 움직임들을 연관시킴으로써, 이것은 가상 POI와 연관된 포인터가 손이 잡음, 지터 또는 작은 손가락 움직임들에 관계없이 이동하지 않을 때 이동하지 않을 것임을 보장한다. 또한, 가상 POI의 위치에서의 지터를 회피하기 위해, 이러한 프로세스는 이진, 즉 개시 또는 정지가 아닐 수 있지만, 대신에, 이전 프레임에서 가상 POI 위치 및 가상 POI의 새롭게 검출된 위치의 중심 사이에서의 간단한 선형 회귀에 의해 0 및 3mm 사이에서의 범위에서의 손 바닥 중심의 움직임들로부터 평탄화될 수 있다. 이 기술분야에 알려진 3D 안정화 기술은 바람직하게는 그렇게 근사된 POI 위치를 개선하기 위해 사용될 수 있다.
- [0084] 도 3a에서, 개방되거나 또는 해제된 "핀칭" 포즈를 가정한 손의 전면도 깊이 맵(301)이 도시된다. 여기에서, 집게 손가락 및 엄지 손가락의 끝들에 대응하는 POI가 도시된다. 도 3b는 도 3a에 도시된 깊이 맵에 대응하는 응답 맵(302)을 예시한다. 이해될 수 있는 바와 같이, 응답 맵(302)은 집게 손가락의 끝을 위한, 및 엄지 손가락 끝을 위한 모호하지 않은 응답 신호를 보여준다.
- [0085] 부가적으로, 하나의 추가 실시예에서, 손바닥 법선은 손바닥 자체의 배향을 결정하기 위해, 및 손가락 아이덴티티를 구별하기 위해 정의된 평면을 개선하기 위해 사용될 수 있다. 손바닥 법선은 손바닥 중심에서 손가락 끝들의 중심 포인트 위치까지의 벡터 및 이전 결정된 손바닥 법선 사이에서의 각도를 사용하여 조정될 수 있다.

이러한 각도는 두 개의 정보 세트들 사이에서의 평활한 혼합을 만들기 위해 몇몇 맞춤 파라미터화에 따라 가중될 수 있다. 손바닥 법선은 PCA를 사용하여 계산되며 가장 덜 중요한 방향인 것으로 설정된다.

- [0086] 부가적으로, 손바닥 반경의 정확한 추정들이 손바닥 중심 내지 손의 가장 가까운 경계 사이에서의 거리의 추정을 사용하며, 손바닥 중심의 현재 위치에서 깊이 값 및 카메라의 규격을 사용함으로써 거리 맵으로부터 획득될 수 있다.
- [0087] 부가적으로, 손이 퍼져 있는지 또는 쥐어져 있는지를 표현하는 통상적으로 부울 값일 수 있는 손 개방은 또한 이전 결정된 POI로부터 획득될 수 있다. 손은 검출된 어떤 손가락 끝들도 없다면(즉, 손가락 끝들에 대응하는 어떤 POI도 없다면) 쥐어진 것으로 고려된다. 두 개의 POI가 이용 가능하다면, 손 개방은 두 개의 선택된 POI들 사이의 거리 및 손바닥 반경 사이에서의 비를 사용하여 업데이트될 수 있다. 비가 미리 결정된 임계치보다 작다면 및 손이 이전에 퍼진 것으로 고려된다면, 손의 개방은 조정될 것이며 쥐어진 것으로 플래그될 것이다. 비가 또 다른 미리 결정된 임계치보다 커진다면, 손은 퍼진 것으로 고려될 것이다.
- [0088] 손에서 POI를 결정하기 위해 상기 설명된 방법이 바람직한 실시예이면, 본 발명의 단계(100)는 대안적으로 또는 보완적으로 3D 이미징 시스템의 시야 내에서 상호 작용하는 손의 모델을 사용할 수 있다. 손의 모델은 추적되며 손에 대응하는 3D 포인트 클라우드에 맞춰질 수 있다. 예를 들면, 손의 골격 모델이 사용될 수 있다. 알려진 맞춤 및 추적 기술들은 실시간으로, 손 파라미터들의 각각의 3D 위치의 결정을 가능하게 한다. 손 파라미터들 중 두 개는 그 후 3D 제스처들이 추가로 인식되는 두 개의 POI들이므로 선택될 수 있다. 선택은 각각의 손 파라미터가 모델에서 식별되며, 인터페이스가 제어될 제스처가 어떤 손 파라미터들이 선택되어야 하는지를 정의할 것이므로 쉽게 이루어진다. 예를 들면, 엄지 손가락 끝 및 집게 손가락 끝은 "핀칭" 3D 제스처를 가능하게 하기 위한 POI들로서 선택될 것이며, 손 끝을 가진 엄지 손가락 팁은 "잡기" 3D 제스처를 가능하게 하기 위한 두 개의 POI들로서 선택될 것이다.
- [0089] 본 방법에 관계 없이, POI의 검출 및 결정은 나중에 소개될 바와 같이 다중-손 상호 작용들을 가능하게 하기 위해 장면에서 복수의 손들로 수행될 수 있다는 것이 주의될 것이다.
- [0090] 본 발명의 제 2 단계(단계 101)는 GUI와의 상호 작용들을 제어하는 요구된 미리 결정된 3D 제스처들을 검출하기 위해 적어도 한쪽 손 결정 단계(100) 상에서 POI의 선택의 시간에 걸친 추적 및 분석을 포함한다. 단계(102)에서 예상되며 상세히 설명된 상호 작용에 따르면, 상이한 미리 결정된 3D 제스처들은 프레임에서 프레임으로 독립적으로, 순차적으로 및 바람직하게는 동시에 인식될 수 있다. 3D 제스처들의 인식을 수행하기 위해, 먼저 각각의 POI의 시간에 걸친 공간에서의 위치가 추적된다.
- [0091] 하나의 제 1 제스처 인식 양상에서, 단일 POI의 시간에 걸친 3D 위치는 종래의 "포인팅" 제스처로 입력을 제공하기 위해 사용될 수 있다. 선택된 POI의 공간에서의 상대적 위치는 스크린상에 디스플레이된 포인터가 그에 따라 이동할 수 있게 한다. 예를 들면, POI는 어떤 손가락 끝들도 이용 가능하지 않다면 손바닥 중심일 수 있고, 어떤 신뢰 가능한 손가락 끝도 이용 가능하지 않다면 손 끝이며, 바람직하게는 그것이 포인팅을 위해 자연스럽게 사용될 수 있기 때문에 이용 가능하다면 집게 손가락 팁이거나 또는 요구된다면 임의의 다른 손가락 끝일 수 있다. 하나의 다른 실시예에서, 다른 손가락 끝은, 그 위치가 3D 공간에서, 손바닥 중심, 손가락 끝들, 및 손 끝 중에서 미리 결정된 손 파라미터들 중 적어도 두 개로부터의 보간된 위치를 사용하여 컴퓨터화되는, 가상 POI일 수 있다. 더욱이, GUI에서의 포인터와 연관된 POI는 손 파라미터에 대응하는 임의의 다른 안정된 POI에 따라 개선된 위치를 가질 수 있다. 예를 들면, GUI에서의 포인터에 연관된 POI의 위치는 손바닥 중심 포인트가 사용자가 포인팅 제스처를 수행하고 있을 때 손의 실제 자발적 움직임을 입증하기 위해 사용될 수 있으므로 공간에서의 손바닥 중심 포인트의 위치에 따라 개선될 수 있다. 어떤 POI가 사용될지에 대한 결정은 유연성 없이 정적으로 설정될 수 있지만, 또한 파라메트릭이며 피드백 루프(103)를 사용하여 상호 작용에 의해 제어될 수 있다. 훨씬 더 유리한 실시예에서, 상호 작용의 연속성은 "포인팅 제스처"가 어떤 이용 가능한 손 파라미터들이 프레임 단위일지라도 실행될 수 있기 때문에 가능해질 수 있다. 보다 정확하게, 상태 머신을 포함한 자동화 프로세스가 어떤 손 파라미터가 이용 가능한지에 따라 GUI에서의 포인터와 연관된 POI를 설정하기 위해 사용될 수 있다. 예를 들면, 바람직하게는, 인체 공학 및 자연스러움에 대하여, "포인팅 제스처"는 먼저 두 개의 다른 것들로부터 보간된 가상 POI에, 두 번째로 집게 손가락 끝에 대응하는 POI로부터, 세 번째로 손 끝에 대응하는 POI에, 및 마지막으로 이것이 항상 이용 가능하므로 손바닥 중심 포인트에 연관될 것이다.
- [0092] 부가적으로, 상기 소개된 바와 같이, 포인터의 스크린상에서의 위치는 그것이 연관되는 POI의 공간에서의 위치로 상대적으로 설정될 수 있다. 하나의 다른 실시예에서, 포인터의 위치는 그것이 연관되는 손상에서 POI의 모션의 절대 메트릭 거리에 대하여 이동할 수 있다.

[0093] 제 2 제스처 인식 양상에서, 두 개의 POI들은 적어도 "클릭킹 제스처", "잡기 제스처", 또는 "핀칭 제스처" 중에서 적어도 하나의 동적 제스처의 발생 또는 해제의 검출에 대응하는 입력을 제공하기 위해 사용된다. 이들 제스처들의 인식, 그것들의 발생 또는 그것들의 해제는 개별적으로, 각각 "개시 이벤트" 및 "정지 이벤트"를 트리거하도록 의도된다. 바람직한 실시예에서, 두 개의 POI들은 잠재적으로 동시에, 수행한 것, 연속적 "포인팅 제스처"와 동일한 손상에서 선택될 수 있다. 그러나, 두 개의 POI들은 또한 거리 이미징 센서에 의해 이미징된 다른쪽 손상에서 미리 결정될 수 있다. 일 실시예에서, 본 발명은, 적어도 두 개의 POI들이 이용 가능할 때, 상태 머신의 사용으로 인해 자연스러운 제스처 기반 상호 작용들을 수행하기 위해 이용 가능한 가장 인체 공학적 POI를 자동으로 관리하며 결정하는 것을 포함할 수 있으며, "핀칭 제스처"로부터 "잡기 제스처"를 구별하는 것은 단계(100)에서 선택된 POI들에 대하여 자동으로 가능해진다. 선택된 두 개의 POI들이 손 끝 및 엄지 손가락 끝이면, 제스처는 "잡기 제스처"일 반면 선택된 두 개의 POI들은 집게 손가락 끝 및 엄지 손가락 끝이면, 제스처는 "핀칭 제스처"일 것이다. "클릭킹 제스처"는 또한 그것이 이들 두 개의 제스처들 중 임의의 하나를 수행할 때, 두 개의 POI들이 미리 결정된 시간 임계치(예를 들면, 0.5초)보다 낮은 짧은 시간 기간, 두 개의 POI들이 함께 병합하며 하나의 단일 POI(그 종류가 통상적으로 손 끝으로서 검출될 수 있는)를 형성하는 기간, 및 두 개의 초기 POI들이 무접촉 상태/자세로 되돌아가는 동안 다시 검출되는 기간 동안 서로 접근한다는 점에서 특성화된 "잡기 제스처" 또는 "핀칭 제스처" 중 임의의 것에 대응하므로 구별된다. "클릭킹 제스처"는 동적 손가락 기반 3D 제스처이다.

[0094] 하나의 단일 손 끝 또는 하나의 단일 손가락 끝이 이용 가능한 상태에 대응하는 하나의 특정한 실시예에서, "클릭킹 제스처"는 그것이 트리거될 수 있으므로 본 발명에 따라, 이러한 경우에, 상기 단일 손의 이용 가능한 말단과 연관된 POI의 특정 움직임에 의해 여전히 가능해진다. 보다 정확하게, "클릭킹 제스처" 이벤트는, POI가 그 모션이 미리 결정된 거리 임계치보다 낮으며(예를 들면 2cm 미만) 그 지속 기간이 시간 임계치 아래로 짧은(예를 들면, 0.5초 미만), 단일 진동의 형태에서의 움직임, 예를 들면, 위 아래 움직임을 수행하는 경우에만 트리거될 수 있다. 이러한 특정한 실시예는 상호 작용 방법이 POI 추적에 사용된 검출 및 추적 방법에 관계없이 손실될 수 있을 때 더 강력하게 한다.

[0095] 각각 "잡기 제스처" 및 "핀칭 제스처"의 발생 및 해제 이벤트들의 검출에 대하여, 이들 이벤트들은 각각 손 및 손가락 기반 동적 제스처들의 분석으로부터 결정된다. 보다 정확하게, 발생 및 해제 이벤트들의 검출은 두 개의 POI들이 하나의 단일 POI를 형성하기 위해 서로를 향해 이동할 때 또는 그것들 사이의 거리가 각각 미리 결정된 임계값(예를 들면, 0.2mm)보다 작게 될 때, 추적되는 두 개의 POI들(즉, 엄지 손가락 끝 및 손 손가락 끝, 또는 엄지 손가락 끝 및 집게 손가락 끝)의 아이덴티티의 함수로서 결정될 수 있고, "잡기 발생 이벤트" 또는 "핀칭 발생 이벤트"가 트리거될 수 있으며 "개시 제스처 이벤트"로서 설정될 수 있다. 두 개의 POI들이 하나의 단일 POI를 형성함으로써 "개시 제스처 이벤트"를 트리거할 때 또는 그것들 사이의 거리가 미리 결정된 임계값(예를 들면, 0.2mm) 아래로 떨어질 때, 및 단일 POI가 두 개의 초기 POI들로 재형성하고 및/또는 두 개의 POI들 사이의 거리가 각각 또 다른 미리 결정된 임계값(예를 들면, 0.3mm) 위로 증가한다면, "잡기 해제 이벤트" 또는 "핀칭 해제 이벤트"가 트리거될 수 있으며 "정지 제스처 이벤트"로서 설정될 수 있다. 바람직하게는, 히스테리시스 함수는 두 개의 POI들이 거리 임계값들에 가까울 때 불안정 및 쉼(jerky) 검출을 회피하도록 "개시 제스처 이벤트" 및 "정지 제스처 이벤트"를 검출하기 위한 거리 임계 기준들과 함께 사용될 수 있다. 예를 들면, 인체 공학 및 자연스러운 제스처들에 대하여 본 발명의 하나의 바람직한 실시예에서, 방법은 "핀칭" 동적 제스처 검출을 가능하게 하기 위해 손 끝들로부터 선택된 적어도 두 개의 POI들의 분석을 포함한다. 상기 설명된 바와 같이, "핀칭"은 서로를 향해 움직일 수 있는, 적어도 하나의 포인터로서 선택적으로 디스플레이된, 두 개의 대향하는 팁들에 대응할 수 있으며, 두 개의 손 끝들, 연관된 POI들 또는 포인터들 사이의 거리는 연속적 측정인 "핀칭"의 정도의 결정을 허용한다. 바람직하게는, "핀칭" 제스처는 도 6을 참조하여 이하에서 설명되는 바와 같이, 엄지 손가락 끝이 제 1 선택된 POI에 대응하며, 집게 손가락 끝이 제 2 선택된 POI에 대응할 때 수행된 제스처인 것으로서 미리 결정된다.

[0096] 또한, 본 발명의 제 2 단계(102)의 하나의 훨씬 더 바람직한 실시예에서, 예를 들면, 3D 카메라 디바이스가 자동차의 지붕에 위치되며 사용자가 자유롭게 동작하는 우측 손으로 앉을 때와 같은, 거리 이미징 카메라 디바이스가 상호 작용 손의 상부-부를 캡처할 때, 제 1 실시예의 방법은 그 목적이 손 형상의 하나의 미리 결정된 투시도에서 홀의 이용 가능성을 검출함으로써 "개시 제스처 이벤트" 및 "정지 제스처 이벤트"의 검출을 강화하는 몇몇 부가적인 프로세싱에 의해 개선될 수 있다. 우측 손의 검출은 좌측 핸들(예를 들면, 유럽, 미국, 캐나다 등에서)을 가진 자동차들을 위한 것이며, 좌측 손의 검출은 우측 핸들(예를 들면, 영국, 호주, 뉴질랜드 등에서)을 가진 자동차들에 적합하다는 것이 이해될 것이다.

- [0097] 인용된 예에 따르면, 도 6은 600에서 도시된 바와 같이, 손가락끝 기반 "핀칭" 3D 제스처의 2D 상부 뷰 또는 표현을 예시한다. 여기에서, 사전 선택된 상호 작용 손은 적어도 두 개의 명확하게 구별된 손가락 및 그것들의 손가락끝들, 및 손가락들이 "개시 3D 핀칭 제스처 이벤트" 또는 간단히 "개시 제스처 이벤트"를 트리거할 수 있는 "핀칭 발생" 포즈에 있을 때 (투시도에 따라) 홀(602)을 포함하는 610에서 도시된 2D 형상을 포함한다.
- [0098] 이러한 방식으로, "핀칭 발생" 포즈는 손으로부터, 또는 이에 제한되지 않지만, 하나의 미리 결정된 관심 포인트(예로서, 질량 중심)와 같은, 손으로부터의 적어도 하나의 특정 피처로부터의 임의의 후속 움직임이 어떤 3D 제스처가 수행되는지를 결정하기 위해 추적되며 분석될 "개시" 이벤트에 대응한다. 추적 및 분석은 손 포즈가 해제될 때까지, 즉 본 실시예에서, "핀칭 발생" 포즈가 정지할 때까지 유지될 것이다.
- [0099] 도 7은 도 6에서의 홀(602)을 가진 601로서 도시된 것과 동일한 손 형상 표현(700)을 도시한다. 손 형상 표현(700)은 PCA에 의해 결정된 바와 같이 손(705)의 주 방향을 따라 및 그것에 평행한 손 바닥 중심(703)으로부터 투사되며, 손의 물질(유한 평균 깊이 값)로부터 보이드 또는 강하게 상이한 깊이 값(즉, 미리 결정된 임계치 이상인 손 물질의 유한 평균 값으로부터의 차이를 갖는 깊이 값에 또는 배경에 대응하는 값)으로의 전이를 검출할 때, 적어도 두 개의 직교 선들(704)은 손 예지들 및 홀 후보 예지들(702)을 결정하기 위해 표현에 대응하는 깊이 맵의 경계들까지 투사되는, 복수의 평행 부근 라인들(706)의 예시를 가진 홀(701)을 포함한다. 보다 정확하게, "물질(matter)"에서 "물질 없음(no matter)"으로의 각각의 전이에 대해, 투사된 선은 이미지 내에서 손 윤곽들의 모두를 검출할 수 있을 것이다. 더욱이, 각각의 전이에 대해, 투사된 선이 전이 없이 이미지 경계에 도달한다면, 전이는 손 윤곽인 것으로 플래그되는 반면 캐스팅된 선들 상에서의 다른 검출된 예지들은 702에서 도시된 것과 유사한 홀 후보들인 것으로 플래그된다. 각각의 홀 예지 후보들에 대해, 경로는 그 후 루프를 정의한 것처럼 확인될 수 있다. 경로 파싱 알고리즘은 손 형상에서의 실제 홀로서 확증될 미리 결정된 임계치(예를 들면, 2cm)를 초과해야 하는 홀 후보 윤곽 길이를 측정하기 위해 사용될 수 있다. 복수의 부근 선들(706)은 손 내에서 완전하며 정확한 윤곽 검출을 가능하게 하기 위해 사용될 수 있다.
- [0100] 홀 검출 방법을 위한 다른 방법들이 또한 이용될 수 있다. 예를 들면, 추적되는 손의 투시도의 2D 표현 또는 3D 포인트 클라우드를 미리 정의된 수의 클러스터들로 클러스터링될 수 있으며, 이것은 손 파라미터들(예로서, 손가락 크기)에 따른 치수로 클러스터들의 평균 크기에 의해 제한된 이미징 디바이스 또는 카메라에 대한 손의 평균 거리의 함수로서 결정될 수 있다. 연결된 구성요소 분석(즉, 클러스터들의 연결)에 기초한 홀 검출은 그 후 이미징된 손에서의 홀들의 존재 또는 부재를 결정하기 위해 사용될 수 있다. 검출된 홀(들)의 수, 크기, 형상 또는 다른 파라미터들은 검출 단계를 강화하며 손에 의해 채택된 정적 포즈에서 유효한 홀의 존재를 확증하기 위해 추가로 사용될 수 있다.
- [0101] 각각 이전 설명된 손가락끝 또는 손 끝 기반 "핀칭 제스처" 발생 및 "잡기 제스처" 발생 및 해제 검출에 대하여, 이들 제스처들은 또 다른 상태 머신을 포함할 수 있는 앞서 언급한 홀 검출 프로세스를 사용함으로써 개선될 수 있다. 예를 들면, 각각 "핀칭 제스처" 발생 및 "잡기 제스처" 발생은 단지 홀 검출이 확증되는 경우에만 확증될 수 있다. 이를 위해, 상태 머신은 "대기" 상태를 갖고 초기화될 수 있다. 손 끝들로부터 두 개의 구별된 POI들이 검출되자마자, 상태 머신은 "개시" 상태로 설정된다. POI들의 수가 하나가 다른 것을 향해 두 개의 POI들의 움직임을 따라 2에서 1로 갈 때, 확증 단계가 시작되어 먼저 홀 후보의 이용 가능성, 및 두 번째로 이용 가능한 POI들의 수의 안정성을 확인한다. POI들 변화들의 수의 빈도가 미리 결정된 임계치보다 높거나 또는 어떤 홀도 검출되지 않는다면, "이벤트 제스처 개시"의 확증은 거절되고, 양쪽 모두가 만족된다면, "이벤트 제스처 개시"가 확증되며 상태 머신 상태가 그에 따라 설정된다. 히스테리시스 함수와 같은 최신 기술로부터의 추가 필터링은 상태에서의 저기 변화들을 추가로 회피하기 위해 상태 머신 상태에 적용될 수 있다.
- [0102] 본 발명의 제 3 및 마지막 단계(단계 102)는 입력 정보로서 인식된 3D 제스처들로부터의(단계(101)로부터의) 정보를 사용하여 유연하고, 강력하며 자연스러운 인간-컴퓨터 상호 작용들의 제공을 포함한다. 상호 작용 기법은 보다 특히 그것이 반드시 3개의 후속 제스처들이 수행되도록 요구하며, 이들 3개의 후속 제스처들이 자연스러운 상호 작용을 가능하게 한다는 점에서 특성화된다. 이들 제스처들 중에서, 적어도 하나는 연속적 동적 제스처(손 또는 손가락끝 기반)이며, 두 개의 다른 것들은 제스처 기반 이벤트들, 바람직하게는 손가락끝 제스처 기반 이벤트들이다. 보다 정확하게, 사용자의 손의 연속적 자유 움직임 내에서, 방법은 미리 결정된 인간-컴퓨터 상호 작용의 (개시) 시작을 트리거하는 제 1 미리 결정된 3D 제스처 이벤트, 인간-컴퓨터 상호 작용을 제어하는 제 2 후속(연속적 상호 작용) 3D 제스처를 검출하는 단계를 포함할 수 있으며, 상기 제 2 (연속적 상호 작용) 3D 제스처는 제 3 미리 결정된 (정지) 3D 제스처 이벤트가 인식될 때 종료한다.
- [0103] 훨씬 더 특히, 제 1 제스처 이벤트는 바람직하게는 "클릭킹 제스처", "핀칭 제스처", "잡기 제스처" 또는 개방/

폐쇄 손 제스처 중 하나를 갖고 수행된 3D 제스처이다. 제 2 제스처는 통상적으로 GUI에서의 포인터가 포인팅 제스처가 상대적인 POI의 위치와 연관될 때 GUI의 임의의 부분에 대한 액세스를 가능하게 하는 "포인팅 제스처"이며, 제 3 제스처는 "핀칭 제스처", "잡기 제스처", "폐쇄 손 이벤트 제스처", 또는 "클릭킹 이벤트 제스처"의 또는 "개방 손 이벤트 제스처"의 신규 발생을 고려하는 경우에, 상호 작용의 종료를 트리거하는 제 1 제스처의 해제이다.

[0104] 본 발명의 상호 작용 기법의 하나의 특정한 이점은 그것이 다양한 GUI와의 자연스러운 손 및 손가락 기반 상호 작용들을 가능하게 하는 3개의 후속 미리 결정된 제스처들을 필요하게 만든다는 것이며, 상기 상호 작용은 핸들링된 임의의 접촉 또는 접촉 요소의 요구 없이 공통 디바이스 기반 상호 작용적 시스템들(예를 들면, 마우스들 또는 터치 스크린)의 것들을 포함한다. 예를 들면, 본 발명의 일 실시예에 따르면, 자동차 GPS 시스템에서와 같은, 내비게이션 시스템으로의 손 3D 제스처 기반 상호 작용은 단일 손을 사용하여 안전하게 제어될 수 있다.

[0105] 도 4는 자연스러운 손 기반 3D 제스처들을 가진 스크린(403)상에서의 맵으로 내비게이션하기 위해 수행된 연속적 손 움직임(400)의 궤적의 후방 뷰를 예시한다. 내비게이션은 점선(400)에 의해 표시된 궤적을 가진 연속적 손 움직임 동안 수행된 도시된 개시 제스처 검출(401) 및 정지 제스처 검출 시간들(402) 사이에서 발생한 스크롤링 상호 작용을 포함한다. 온-스크린 맵 상대 변위/스크롤링은 상호 작용이 종료될 때 406에 중심을 둔 영역(404)이 407에 중심을 둔 영역(405)으로 변위되게 맵의 위치를 변경한다. 그/그녀가 수행하는 3D 제스처에 대하여 사용자로의 시각적 피드백을 가능하게 하기 위해, 사용자의 상호 작용하는 손에서의 하나의 POI에 연관된 포인터는 연속적으로 디스플레이될 수 있다.

[0106] 예를 들면, 본 발명의 하나의 또 다른 실시예에 따르면, 상이한 윈도우들을 포함한 GUI로의 손 3D 제스처 기반 내비게이션은 단일 손을 사용하여 안전하게 제어될 수 있다.

[0107] 도 5는 자연스러운 손 기반 3D 제스처들을 가진 디스플레이 스크린(503)으로 슬라이드들 또는 GUI(505)를 변경하기 위해 수행된 연속적 손 움직임(500)의 궤적의 후방 뷰를 예시한다. 사용자의 한쪽 손으로부터 POI와 연관된 포인터는 피드백 정보로서 디스플레이될 수 있다. 궤적(506)을 통한 그것의 모션은 예를 들면, 그것이 연관되는 POI의 점선(500)에 의해 표시된 바와 같이 절대 변위에 대하여 절대 변위이도록 이루어진다. 사용자의 상호 작용하는 손의 연속적 모션 동안, 제 1 "개시 이벤트 제스처"가 501에서 핀칭 제스처의 검출로부터 트리거될 때, 상호 작용하는 손의 자연스러운 연속적 모션의 거리 및 방향은 "핀칭 제스처"가 502에서 해제될 때 "정지 이벤트 제스처"가 검출될 때까지 측정된다. "개시 제스처" 이벤트(507) 및 "정지 제스처" 이벤트(508)가 도시되는 궤적(506)을 따라 포인터(원들로 도시된)의 변위에 대응하는 상기 움직임에 따르면, 움직임이 미리 결정된 거리 임계치(예를 들면, 5cm)보다 높다면, 궤적이 단일 방향으로 선형성 상관 값(예를 들면, 0.8의 선형 상관비)을 보여준다면, 및 움직임이 미리 결정된 시간 기간(예를 들면, 1초 미만) 내에서 수행된다면, 이것은 이전 디스플레이되지 않은 윈도우들 또는 GUI 또는 슬라이드(504)가 디스플레이되게 하며, 이전 디스플레이된 윈도우들 또는 GUI 또는 슬라이드(505)가 스크린으로부터 삭제되게 한다. 인터페이스에 따르면, 궤적은 위, 아래, 좌, 우 움직임, 또는 이들 모션들의 조합 중 적어도 하나일 수 있으며, 그것들은 자동차 인포테인먼트 시스템에서, 오디오 메뉴로부터 GPS 메뉴로 비디오 메뉴 등으로 전이를 가능하게 하는 것과 같은, 메뉴들을 제어하며 변경하도록 의도될 수 있다.

[0108] 예를 들면, 본 발명의 하나의 또 다른 실시예에 따르면, 상호 작용은 거리 및 시간 임계치들을 사용하지 않고, 그러나 포인팅 제스처를 사용함으로써 인터페이스들, 윈도우들 또는 GUI를 변경하는 것을 포함할 수 있다. 보다 정확하게, 일단 "클릭킹 제스처", "핀칭 제스처", "잡기 제스처" 중 하나로부터의 개시 제스처가 검출된다면, 포인터의 후속 위치는 디스플레이 스크린에 대한 그것의 위치의 함수로서 GUI에 대한 변화들을 결정할 것이다. 포인터가 스크린의 상부 경계에 도달한다면, GUI는 예를 들면, 오디오 시스템에서 사운드들의 플레이리스트의 슬라이드를 제작하기 위해 상부로 슬라이딩될 수 있다. 상호 작용은 임의의 방향에 따라 설정될 수 있으며 연속적이거나 또는 단계별 기반일 수 있다.

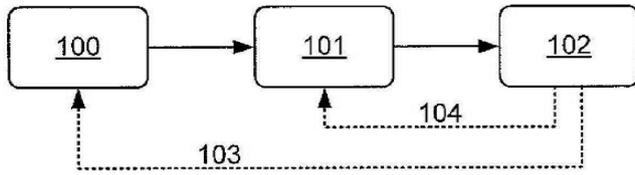
[0109] 예를 들면, 본 발명의 하나의 또 다른 실시예에 따르면, 상호 작용은 GUI를 줌 인하는 것을 포함할 수 있다. 줌 피쳐는 예를 들면, "잡기 제스처"와 연관될 수 있어서, "클릭킹" 및/또는 "핀칭 제스처"의 인식을 사용하여 다른 종류의 상호 작용을 가능하게 한다. 일단 "잡기 제스처" 발생으로서 식별된 바와 같이 "개시 이벤트"가 트리거된다면, 예를 들면, 디스플레이 스크린에 더 가까운 손의 연속적 이동은 디스플레이된 GUI를 줌 인할 것이며, 디스플레이 스크린으로부터 훨씬 더 멀리 이동하는 것은 GUI를 줌 아웃할 것인 반면, 잡기 제스처는 그것이 해제될 때까지 유지된다.

[0110] 본 발명의 또 다른 양상들이 특정 실시예들에 대하여 설명되었지만, 이들 양상들은, 한편으로 "개시 및 정지 제

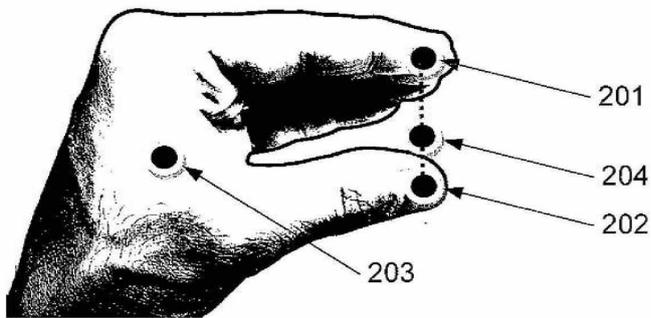
스처', 및 다른 한편으로 연속적 "상호 작용 제스처"를 트리거하는 두 개의 손들을 사용하여 조합하는 것을 포함한다, 다른 형상들로 구현될 수 있다는 것이 쉽게 이해될 것이다.

도면

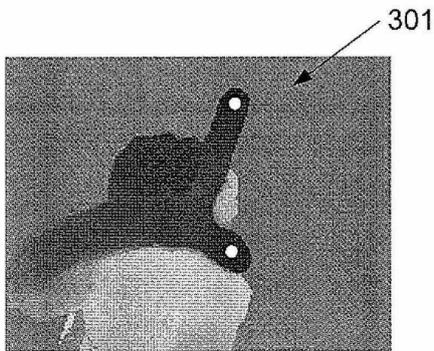
도면1



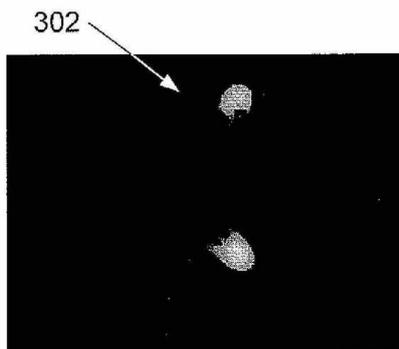
도면2



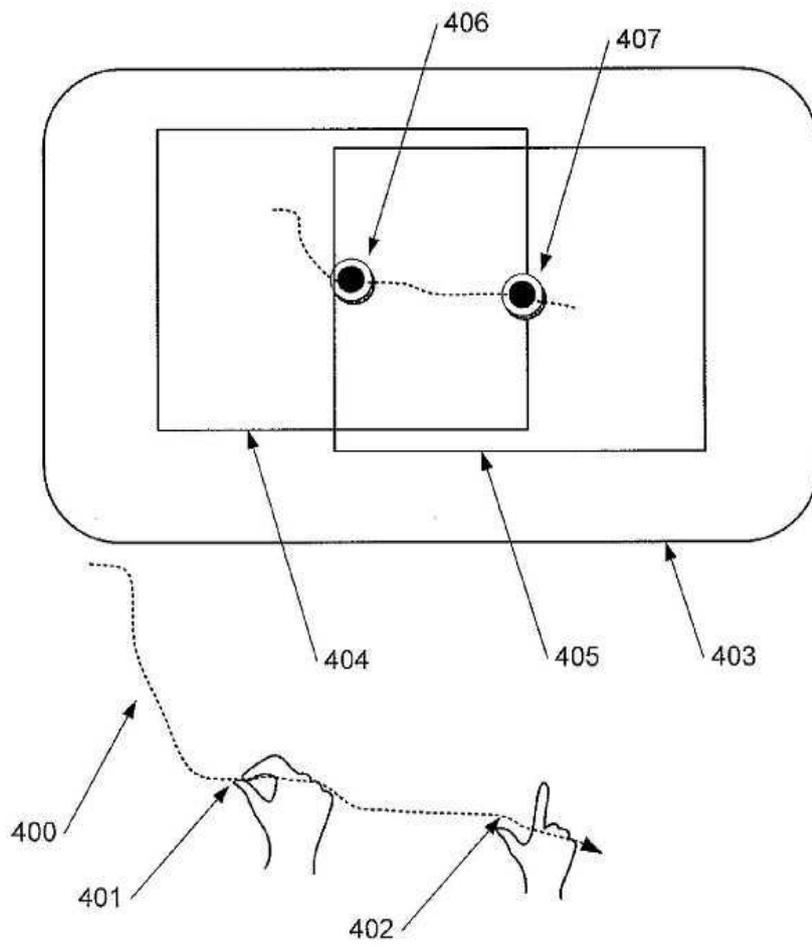
도면3a



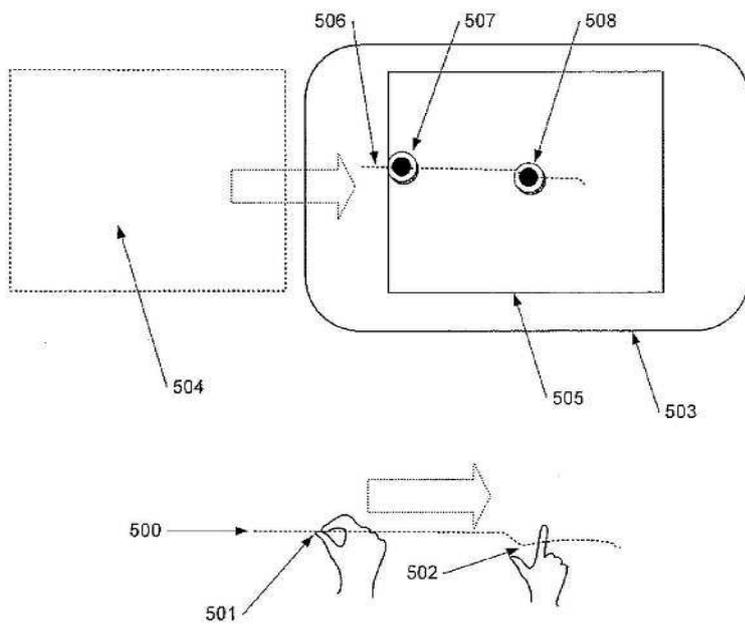
도면3b



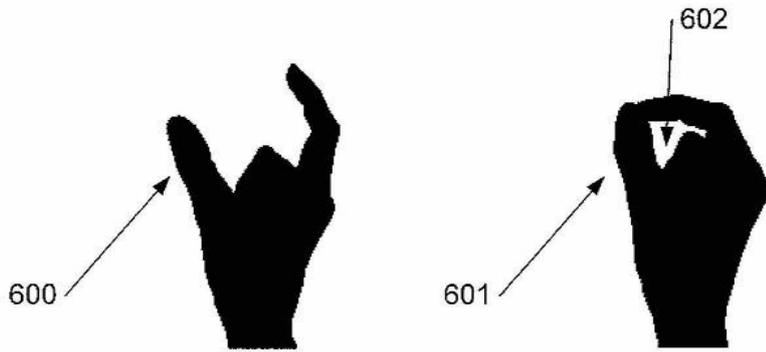
도면4



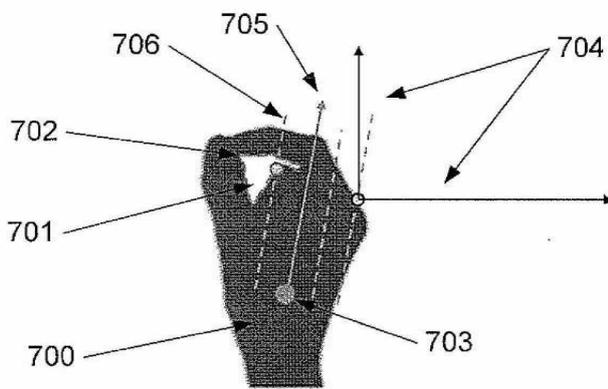
도면5



도면6



도면7



도면8

