



(12) 发明专利申请

(10) 申请公布号 CN 105378593 A

(43) 申请公布日 2016. 03. 02

(21) 申请号 201380036563. X

代理人 王琳 武晨燕

(22) 申请日 2013. 07. 12

(51) Int. Cl.

(30) 优先权数据

G06F 3/01(2006. 01)

12176458. 3 2012. 07. 13 EP

G06F 3/03(2006. 01)

G06K 9/00(2006. 01)

(85) PCT国际申请进入国家阶段日

2015. 01. 08

(86) PCT国际申请的申请数据

PCT/EP2013/064860 2013. 07. 12

(87) PCT国际申请的公布数据

W02014/009561 EN 2014. 01. 16

(71) 申请人 索夫特克尼特科软件公司

地址 比利时布鲁塞尔

(72) 发明人 吉勒·皮诺特 乌特库·萨利豪格路

卡洛斯·索扎 J·托洛

劳伦特·季格斯

(74) 专利代理机构 北京派特恩知识产权代理有

限公司 11270

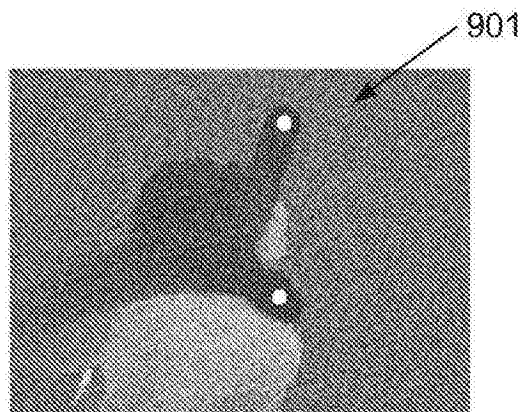
权利要求书3页 说明书15页 附图7页

(54) 发明名称

利用手上的奇异兴趣点基于手势进行人机同步交互的方法和系统

(57) 摘要

本文公开一种通过由测距成像系统提供的深度图像使能基于三维手势的人机自然交互的方法。该方法通过检测、跟踪和分析用户的单手上的奇异兴趣点使能对同步手势的识别,并向用户提供上下文反馈信息。该手的奇异兴趣点包括指尖、指尖、掌心和手的质心,并且用于定义指针的至少一个表示。兴趣点被随时间跟踪并被分析,以使能对由单手执行的“定点”和“激活”手势的顺序和/或同时确定。



1. 一种基于三维手势识别系统提供自然人机交互的方法,该方法包括以下步骤:
 - a) 对包括至少一个用户的至少一只手在内的场景进行成像;
 - b) 处理成像场景,以确定与所述至少一只手相关的至少一个兴趣点;
 - c) 跟踪所述至少一个兴趣点,以提供每个兴趣点随时间的跟踪运动;
 - d) 分析每个兴趣点的所述跟踪运动;
 - e) 从分析中确定至少一个手势的执行;以及
 - f) 使用所确定的所述至少一个手势的执行来进行人机交互。
2. 根据权利要求 1 所述的方法,其中,所述至少一个手势包括使用一个兴趣点的定点手势。
3. 根据权利要求 2 所述的方法,其中,所述一个兴趣点包括下列之一:指尖、手尖、掌心、手的质心以及指尖、手尖、掌心和手的质心中的至少两者的组合的衍生物。
4. 根据权利要求 1 所述的方法,其中,所述至少一个手势包括使用一个兴趣点的激活手势。
5. 根据权利要求 1 所述的方法,其中步骤 b) 确定至少两个兴趣点,并且步骤 e) 包括:使用对所述至少两个兴趣点的分析,来确定所述至少一个手势的执行。
6. 根据权利要求 5 所述的方法,其中,所述至少一个手势包括使用两个兴趣点的定点手势。
7. 根据权利要求 6 所述的方法,其中,所述定点手势使用从所述两个兴趣点内插的兴趣点。
8. 根据权利要求 5 至 7 中任一项所述的方法,其中,所述两个兴趣点包括下列任一者:两个指尖、一个指尖和一个手尖、一个指尖和掌心。
9. 根据权利要求 5 至 8 中任一项所述的方法,进一步包括以下步骤:通过所述两个兴趣点确定两个手势的同步执行。
10. 根据权利要求 9 所述的方法,其中,所述同步执行的两个手势包括定点手势和激活手势。
11. 根据权利要求 10 所述的方法,其中,两个兴趣点包括:提供定点手势的掌心和用于拍手势的指尖。
12. 根据权利要求 5 所述的方法,其中,步骤 b) 确定至少三个兴趣点,并且步骤 e) 包括:使用对所述至少三个兴趣点的分析,来确定所述至少一个手势的执行。
13. 根据权利要求 12 所述的方法,进一步包括以下步骤:通过所述三个兴趣点确定两个手势的同时执行。
14. 根据权利要求 13 所述的方法,其中,所述同步执行的两个手势包括定点手势和激活手势。
15. 根据权利要求 14 所述的方法,其中,所述定点手势由单一兴趣点执行,并且所述激活手势由两个兴趣点执行。
16. 根据权利要求 15 所述的方法,其中,所述单一兴趣点对应于掌心。
17. 根据权利要求 15 或 16 所述的方法,其中,所述激活手势由两个指尖执行,或由指尖和手尖执行。
18. 根据权利要求 15 所述的方法,其中,所述单一兴趣点包括两个指尖之间的内插兴

趣点,所述两个指尖提供所述激活手势。

19. 根据权利要求 15 和 18 中任一项所述的方法,其中,所述定点手势包括:使用两个指尖、掌心和内插兴趣点确定的、具有旋转不变性的定点手势。

20. 根据前述权利要求中任一项所述的方法,进一步包括以下步骤:响应于步骤 e),提供上下文反馈给用户。

21. 一种从由测距成像系统提供的对应于手的一组标记 3D 点中检测与手相关的参数的方法,确定与每个手参数相关的至少一个特征。

22. 根据权利要求 21 所述的方法,其中,所述与手相关的参数包括:指尖和手掌中心中的至少一个,并且其中,与指尖或手掌中心相关的特征至少为指尖或手掌中心在 3D 空间中的位置。

23. 根据权利要求 21 或 22 所述的方法,进一步包括:使用距离地图和 PCA 分析中的至少一种来确定掌心的位置。

24. 根据权利要求 21 或 22 所述的方法,进一步包括:通过以与手的所有正交像素被投影到的矢量的直方图中的预定范围对应的方式确定指尖在空间中的位置,来确定指尖的位置,该矢量的特征在于,该矢量的方向与在手 3D 点云上执行的 PCA 分析的主方向相关,该矢量的原点位于与手掌中心相关的位置。

25. 根据权利要求 21 至 24 中任一项所述的方法,进一步包括:通过计算响应地图来确定指尖或末端,每个指尖对应于一组像素,该组像素的值被确定为一组邻居像素中的第三最小值,每个邻居像素在一组预定方向上的预定距离处,并具有与被考虑的像素不同的、高于预定值的深度值。

26. 根据权利要求 21 至 25 中任一项所述的方法,进一步包括:根据手参数确定至少一个兴趣点在空间中的位置。

27. 根据权利要求 26 所述的方法,进一步包括:根据手参数中与至少两个兴趣点连接的部件相关的参数,确定所述至少两个兴趣点在空间中的位置。

28. 根据权利要求 26 或 27 所述的方法,进一步包括:自动确定至少两个相对的兴趣点在空间中的位置。

29. 根据权利要求 26 至 28 中任一项所述的方法,进一步包括:根据兴趣点和手参数,自动确定待表示在交互系统中的至少一个指针,该交互系统能够至少提供视觉反馈信息。

30. 根据权利要求 29 所述的方法,其中,能够提供反馈信息的所述系统是一图形用户界面,该图形用户界面包括能够渲染至少一个指针的虚拟元素,所述虚拟元素与所述至少一个指针进行交互。

31. 根据权利要求 26 至 30 中任一项所述的方法,进一步包括:跟踪和分析所述指针或相关兴趣点中的至少一个随时间的运动,以确定动态手势。

32. 根据权利要求 26 至 31 中任一项所述的方法,进一步包括:跟踪和分析多个指针或兴趣点随时间的运动,以确定动态手势。

33. 根据权利要求 26 至 32 中任一项所述的方法,进一步包括:分析多个兴趣点在空间中的几何布置,以确定静态手势。

34. 根据权利要求 31 至 33 中任一项所述的方法,其中,所述系统和由至少一个指针表示的用户的兴趣点之间的交互是包括对虚拟对象的操纵在内的自然交互,所述操纵包

括以下中的至少一种：触摸，挤压，捏，抓取，旋转，推，掉落，点击及其衍生的手势。

35. 根据权利要求 31 至 34 中任一项所述的方法，其中，交互包括使用由至少一个指针表示的至少两个兴趣点来操纵所述虚拟对象。

36. 根据权利要求 31 至 35 中任一项所述的方法，其中，交互包括使用由两个指针或抓取器表示的至少两个控制点来操纵虚拟对象。

37. 根据权利要求 31 至 36 中任一项所述的方法，其中，交互具体包括使用由两个指针或抓取器表示的至少两个控制点和另一个指针来操纵虚拟对象，所述另一个指针的位置被确定为在两个其他指针之间。

38. 根据权利要求 29 至 37 中任一项的方法，其中，所述反馈信息系统提供上下文信息，所述上下文信息控制手参数中的至少一个手参数的确定、待使用的控制点的确定以及指针的确定。

39. 根据权利要求 21 至 38 中任一项所述的方法，进一步包括：根据由用户执行的手势以及对象和指针各自在空间中的位置，来修改所述对象和指针在图形用户界面中的表示。

40. 一种用于执行根据权利要求 1 至 39 中任一项所述的方法的系统。

利用手上的奇异兴趣点基于手势进行人机同步交互的方法和系统

技术领域

[0001] 本发明涉及基于由用户的至少单手执行的手势的人机自然交互及其改进,并且更具体地说,涉及(非排他地)检测利用该单手同步执行的多个三维(3D)手势,诸如,定点(pointing)手势和激活手势。本发明包括:使用用于捕获包括用户的至少一只手在内的3D场景的深度传感成像装置,以及使能对可通过每个单手同时执行的多个3D手势的同步识别的计算装置。具体地,本发明包括:用于对手上的多个奇异兴趣点(POI)进行检测、空间跟踪和分析的装置,所述奇异兴趣点例如:手的末端,指尖,手尖,手的质心和掌心,这些奇异兴趣点被用作3D手势识别的输入。本发明进一步涉及一种图形用户界面系统,其使用由所识别的手和/或手指及其相关的兴趣点执行的三维手势,来使用户能够使用(特别是)同步定点和激活手势自然且直观地与计算机系统交互。

背景技术

[0002] 常规人机接口通常包括硬件控制系统接口,如,键盘、鼠标、遥控器和定点设备。在具有这些接口的情况下,还需要利用硬件设备本身以这些设备接口使能的方式,进行物理动作(例如,触摸、移动、保持、定点、按下、点击),或者甚至顺序或同时一起进行多个这些动作,从而使控制命令可以被发送到该接口欲与之交互的计算机系统。

[0003] 最近已经出现自然交互系统,例如,在US-A-2011/0115892中描述了传统的二维(2D)相机被用于捕获可见光谱中的光以及检测用户的手指。然而,由于该种技术的局限性,在被捕获的场景内的手指状的物体,例如,笔,可能被错误地识别或检测为手指,或者对手指的跟踪可能会由于场景的照明度而丢失。但是,先进的成像处理技术使得使用传统的照相机来检测手以及提供输入以允许分析手的末端成为可能。然而,使用这些技术仍然不能准确地(特别是对不同距离的强健壮性,或以对于背景照明度的强健壮性)分析存在于3D场景中的任何其他末端。

[0004] 在US-A-2012/0069168中,使用颜色信息,例如使用从场景中提取的手的掩模,来查找不同的与手有关的数据或手参数,如,掌心和掌基部,以及从掌心到手的轮廓或末端的距离。这些距离甚至可以用来评估手是闭合的还是伸开的,并且根据该评估可以确定手是否正在执行有关“选择”或“抓取”的手势。然而,这种基于手势的方法有其局限性,并且不能提供一种可靠的方法来解决与计算机的诸如3D定点之类的交互,也不能在无法区分颜色的黑暗环境下操作。此外,检测出的“抓取”手势不是非常精确,因为提供的距离仅仅是相对测量值,因此不能被用于准确“定点”和“抓取”在传感器或成像装置的不同距离处的虚拟对象。而且,这种方法不能提供以下信息,例如3D场景中两个独立兴趣点之间的准确目标距离,该准确目标距离可强制用于获取准确和可靠的“抓取”手势或还强制用于测量在场景中手的“捏”手势的水平或相对量。

[0005] 但是,与第三维,即,深度,相关的信息,作为一项重要补充,现在可以通过使用来自距离传感相机的输入来确定。此外,距离传感相机可以操作于,例如,红外光谱而非可见

光谱中。由于场景捕获独立于场景的自然照明度,并且无论对象以及兴趣点与图像传感装置的距离多远,对象的绝对尺寸和兴趣点之间的距离都能够确定,因此这样的距离传感相机提供的三维(3D)信息开启了获得更坚固、稳定、可靠和准确的用户手模型的可能性。

[0006] 到现在为止,尚未实现在三维空间中对手的健壮检测及其跟踪以及对手指和某些其它相关兴趣点的跟踪。此外,也尚未实现对由单手或与其相关的奇异兴趣点顺序或同步执行的不同种类的手势的健壮识别。特别是,目前还没有基于自然 3D 手势的交互系统,该系统能够确定在三维场景中至少单手上的多个奇异兴趣点,并且跟踪这些兴趣点,使定点和激活手势能够被识别而不会产生假阳性检测(即使这些手势与它们所关联的由单手同时执行)。

[0007] 在图形用户界面(GUI)技术领域,使用基于指针的界面是常见的,对触摸或多点触摸界面的使用一直在增加。使用基于指针的界面的表示性输入设备包括鼠标和触摸屏。这种基于至少一个指针的输入设备的优点在于,它们的操作是准确的,命令可以清楚地被区分,并传递到相关联的计算机系统的图形用户界面,例如,硬件鼠标装置,其使用点击按钮同时使能定向并激活一特征,向用户提供了关于他/她的交互的状态的清晰反馈。然而,使用硬件的缺点可能在于,反馈的一部分可能需要部分地通过与硬件设备本身接触来作出。

[0008] 在用于实现人机交互的图像处理的领域,围绕手指和手的检测、跟踪和识别,以及(以非常有限的比例)围绕手指和手在空间中的运动的识别,最近已开发出几种技术。此外,基于视觉的实时计算机人手指识别大多集中于用于身份验证应用的指纹识别和掌纹识别。此外,为了能够识别在复杂背景中的人手指,跟踪手指运动和用预定义的手势解释手指运动以往受到成像系统的能力和支撑成像系统的图像信号处理系统的能力的限制。一个后果是,在对基于手/手指 3D 手势的自然互动提供清晰无误的反馈方面,目前并没有真正付出任何努力。

[0009] 同时,已知的是,通过识别和跟踪用户身体一部分(例如,手或手上的手指)的 3D 维运动来控制指针的自然交互技术表现出相对低的识别率,因为仍然需要明确区分对应于控制的运动和与交互本身无关的运动。解决这个问题的一种常见技术,需要非直观的、很难用的特殊动作,比如,清楚顺序执行的“开始”和/或“停止”手势,这并不符合高效的单手同步“定点”和“激活”手势识别。此外,基于手或手指手势的自然交互技术也存在问题,因为仍难以使由用户在 3D 空间中的运动导致的指针位移对应于从用户的角度来看的鼠标位移。对于并未遵循基于自然的交互而开发使用的 GUI 或交互系统来说,特别是对于不能够向执行自然手势的用户提供指示该手势是否被识别出来的反馈的交互系统来说,尤其如此。这与其中激活按钮提供物理点击作为激活反馈的硬件鼠标是完全不同的。

发明内容

[0010] 在一个方面,本发明涉及在 3D 点云中检测用户的至少单手的参数和奇异兴趣点的具体实现,所述奇异兴趣点包括但不限于指尖和指尖,所述 3D 点云是从利用 3D 测距成像系统捕获的场景中获得的,并且其中,这些 3D 点的集群表示被考虑的手。不管手的姿态和取向如何,都可能被检测到。

[0011] 更具体地,本发明涉及检测在单手上的用于使能“定点”手势的至少第一兴趣点,

优选地为尖端或末端,并且在优选实施例中,检测至少两个奇异兴趣点,其用于使得可靠的“捏 (pinch)”、“抓取 (grab)”、“点击”或“拍 (snap)”手势能够在相关的顺序或同步交互中被识别出来。

[0012] 更优选地,本发明涉及检测一个单手上的两个 POI,根据该检测,例如,可以从所检测到的末端(如食指指尖)中确定至少一个第一 POI,而另一 POI 被确定为手的第二末端,如拇指指尖。通常,任何手指可以被指定为该交互所需的两个 POI 中的一个。

[0013] 在另一个实施例中,在 3D 点云内,第一 POI 可以被确定为例如指尖,如食指指尖,而至少第二 POI 可以被确定为掌心或手的质心。在本实施例中,第二 POI 可以被用于定点手势,而第一 POI 可以被用于由食指指尖的上下运动触发的激活手势,即“拍”手势,其可以与定点手势同时执行且其中任一种手势都不会被错误检测。

[0014] 在一个具体的形式中,该方法包括当一些手指粘在一起时允许健壮和有效地确定手尖的手段,例如,第一尖端可以被确定为所检测到的拇指指尖,其他尖端可以被确定为对应于手的末端,该末端具有其它各个手指粘在一起,该末端的位置对应于所有的其它指尖的近似平均位置。在其中两个尖端被检测到并且对应的兴趣点被使用的形式中,该方法可以进一步包括确定第三兴趣点,第三兴趣点的位置根据两个尖端来确定,以便生成“中间”兴趣点,该“中间”兴趣点在稍后的过程中用作指针。在一个具体的实施例中,该第三点被确定,以避免其在诸如“捏”手势的过程中移位,因为两个兴趣点(如尖端)的运动是不对称的,因此对于准确的自然交互可能是有问题的。

[0015] 应当理解的是,本发明的方法和系统不限于检测手的两个指尖,也可以用于同一只手和/或多只手的更多个指尖或 POI。特别是,该方法和系统包括从多个手尖中区分至少两个不同和相对的兴趣点的手段,以便自动确定和选择更适合基于“捏”或“抓取”的交互的兴趣点。

[0016] 在第二方面,本发明涉及分析一组确定的兴趣点在空间和时间中的布置的具体实现,该组确定的兴趣点至少包括手尖,并且优选地,为与这些手尖相关的一组兴趣点以及它们彼此之间的关系。该分析的目的是在第一子步骤中检测和识别兴趣点中的至少一个随时间的运动,以便使用本领域已知的方法确定动态手势,如“定点”手势或“扫 (swipe)”手势。该分析的目的是在第二子步骤中检测和识别兴趣点在一时间在空间中的布置,即,兴趣点在空间中各自的位置,以便识别具体布置和确定在给定时间点的手姿态。在给定时间点的手姿态分析可以使用本领域已知的任何其他方法,例如模式匹配识别方法。优选地,所述手姿态识别可以使用将兴趣点链接在一起的几何参数分析。该分析的目的是在第三子步骤中检测和识别兴趣点在空间中的布置随时间的改变,即,它们各自的位置改变(如果适用的话),以便确定具有手的非静态姿态的动态手势,并将其与具有静态姿态的动态手势区分开来。具体地说,该方法包括识别具有手的非静态姿态的动态手势,该动态手势可以至少为“捏”或“抓取”手势的形式。

[0017] 在第三方面,本发明涉及一种从兴趣点中确定用作用于与具有反馈装置(例如 GUI)的特定系统交互的指针的兴趣点的方法的具体实现,反馈装置的设计兼容至少一个指针的表示,并且优选地,兼容两个指针的表示。该系统的 GUI 还能够表示具有独立交互能力的不同虚拟对象,其可以与用户的手的兴趣点的表示,即指针,结合使用。更优选地,在 GUI 的表面的空间中表示的虚拟对象或一些部分可能能够根据由与兴趣点相关的指针或兴趣

点本身执行的手势或位置来提供视觉反馈给用户。甚至更优选地,当与用户的手的兴趣点相关的指针或兴趣点本身指向特定位置和 / 或正在执行手势时,在 GUI 的表面的空间中表示的虚拟对象或一些部分可能能够以失真、移动、闪烁、高亮或以其他方式向用户提供视觉反馈。另外,指针也可以根据其位置、根据所检测到的手势或根据该指针所交互的虚拟对象,来传送视觉反馈。在本发明的一个优选实施例中,交互可以进一步包括自动确定手上的兴趣点的数量,自动确定待显示的指针的数量,并且还自动确定与在 GUI 上显示的目标上下文相关的指针的表示以及其在 GUI 的区域内与所述目标的交互。

[0018] 本发明的其他特征由所附权利要求限定。

附图说明

[0019] 为了更好地理解本发明,现在将通过示例的方式参考附图进行描述,在附图中:

[0020] 图 1 示出根据本发明的方法的三个主要处理步骤的流程图;

[0021] 图 2 示出根据本发明的方法的第一主要步骤的四个子步骤的流程图;

[0022] 图 3a 至 3f 示出了当确定一些手参数时需要考虑的用户的手的不同的可用姿势或姿态;

[0023] 图 4a 和 4b 示出了使用两个手指抓取对象的原理,其中两个手指的指尖由根据本发明的方法来确定;

[0024] 图 5a 至 5d 示出了根据本发明的方法确定在手姿态中指尖、掌心和拳尖的定位;

[0025] 图 6 示出了呈现“抓取”姿势的手的侧视图,其叠加有指尖、拇指尖、掌心定位和一附加兴趣点的定位;

[0026] 图 7 示出了呈现“抓取”姿势的手的正视图,其叠加有指尖、拇指尖的定位以及可被设置为指针的一附加兴趣点的定位;

[0027] 图 8 示出了手的距离地图;

[0028] 图 9a 和 9b 分别示出了呈现“捏开”姿态的手的深度地图和相应的响应地图;

[0029] 图 10a 和 10b 分别示出了呈现“捏合”姿态的手的深度地图和相应的响应地图;

[0030] 图 11a 和 11b 中分别示出了深度地图和相应的响应地图,其中拇指产生清晰的响应信号,其它响应对应于指尖;

[0031] 图 12a 和 12b 分别示出了多个末端的深度地图和相应的响应图;

[0032] 图 13 示出了包括虚拟对象和与在手尖中确定的兴趣点相关的两个指针的表示的一些交互;以及

[0033] 图 14 示出了在使能同时定点和激活手势的姿势中手的侧视图。

具体实施方式

[0034] 将针对具体实施例并参照某些附图来描述本发明,但本发明不限于此。所描述的附图仅是示意性的和非限制性的。为了图示说明的目的,在附图中,一些元件的尺寸可能被夸大并且未按比例绘制。

[0035] 此外,如下所述的附图的特定取向不是限制性的,而是通过举例的方式示出。

[0036] 本文所用的术语“自然交互”或“基于三维 (3D) 手势的自然交互”是指其中人不必手持、携带或使用任何硬件设备或标记设备的人机交互,且其中交互式计算机系统与成像

装置相关,所述成像装置例如为诸如三维(3D)飞行时间(TOF)相机之类的测距成像系统,其用于使用一些计算装置捕获和提取用户所提供的信息(优选地,为可以由计算装置识别的手势的形式),从而控制计算机。

[0037] 本文所用的术语“姿态”或“姿势”是指在特定时刻的物理手势。这对应于手的一组特定点在空间中的位置的特定布置,尤其是,一组奇异兴趣点在空间中的布置。

[0038] 本文所用的术语“奇异兴趣点”或“兴趣点”是指用户的手上的奇异点,奇异点即为:手的末端;指尖;手尖;手的质心;掌心;或其位置可以利用从手获得的随时间可被跟踪的其它可靠点中的至少一个的位置来确定的任何其它点。

[0039] 下文中,术语“兴趣点”或“多个兴趣点”被统称为“POI”,从使用术语“POI”的上下文中可以明确地确定是否正在描述一个或多个兴趣点。

[0040] 本文所用的术语“手势”指的是一个兴趣点或一组兴趣点在空间中的布置随时间的变化。另外或替代地,术语“手势”也可以指POI的不变布置在空间中的位置随时间的变化。如果姿态不随时间改变,也就是,POI在空间内的布置保持不变,它被描述为静态手势。如果POI的布置保持不变,该布置在空间中的位置随时间变化,它可以被描述为具有静态姿态的动态手势。如果POI的布置随时间变化,则这被描述为具有不同姿态的动态手势。

[0041] 本文所用的术语“手尖”是指与手的主要末端对应的单一POI。它可以更具体地位于手指末端附近的平均位置。

[0042] 本文所用的术语“指尖”是指在空间中表示每个单独的指尖或手指末端的几个位置。每个单独的指尖通常被认为是一个POI。

[0043] 本文所用的术语“手尖”是指手尖,也指指尖。

[0044] 本文所用的术语“指针”或“多个指针”是指在一个兴趣点或一组选定的POI在虚拟世界中的表示。

[0045] 本文所用的术语“抓取器(grabber)”是指与手上的POI相关的预定数量的指针。它们基本上包括至少两个指针,用于当用户用他的至少一只手执行“捏手势”时识别手的有意义的末端。当使用拇指和食指时,食指和拇指的指尖分别包括上、下“抓取器”。如果使用多于两个手指时,下“抓取器”可以被定位(fixed)在拇指上,上“抓取器”可以是其他手指的平均位置。此外,使用的术语“下抓取器”和“上抓取器”是指已关联的POI相对于场景的取向。根据POI在场景内的取向,下抓取器”和“上抓取器”可能实际上被认为是左、右“抓取器”。

[0046] 本文所用的术语“定点”是指“定点手势”。它对应于,例如,使用至少一个手尖或相关的POI作为与真实世界用户相关的参考三维位置,即,在空间中与一末端相对应的位置,该位置可以用于确定在虚拟空间中指针的虚拟表示所在的位置。可根据多种方法来实现对指针的虚拟表示的位置的确定,包括,例如,使用用户POI在真实世界3D空间中的绝对定位,使用用户POI在真实3D空间中相对于一个确定的参考三维位置的相对定位,或使用由两个POI之间的矢量组的主方向(例如,由从掌心到食指指尖的矢量限定的方向)确定的相对定位。应该理解的是,“定点”可另外利用几个指针,包括,例如,使用至少两个POI的位置的统计模式。该统计模式可以是例如以下中的至少一个:均值;中值;最大值;最小值或它们的位置的加权和。

[0047] 本文所用的术语“激活手势”是指由系统识别的、被用作类似于鼠标点击的事件触

发器的手势。激活手势可能包括：“捏”手势，手指“拍”手势，和 / 或“抓取”的手势。

[0048] 就人体工程学方面的考虑，术语“捏”是指手的聚拢，使得在至少两个 POI 或其对应的从表示手的部分的指针中选取的指针变得彼此更近。“捏”可以是被识别为用于激活的手势。两个 POI 之间的距离或相对距离可另外被用于确定与“捏”手势的水平或量相关联的值。

[0049] 就人体工程学方面的考虑，术语“抓取”是指手的聚拢以使至少两个 POI 或其对应的从表示手的部分的指针中选取的指针变得彼此更近，直到它们在虚拟世界中的表示之间的距离对应于待“被抓取”的对象或虚拟对象的表示的尺寸。此时，对象可以被认为是“被抓取”，并可以根据用户执行的手势采用以下形式被操纵，例如，挤压，移动，取向，转换，拖曳和放下 (drop) 等。“抓取”可以是一个被识别的手势。

[0050] 就人体工程学方面的考虑，术语“拍手势”是指对应于指尖（优选地，食指尖或指尖或其相关联的指针）的 POI 的上下运动。“拍手势”通常为用作触发事件的激活手势的手势，例如，在人机交互系统中的点击。

[0051] 本文所用的术语“手参数”是指至少下列与手有关的参数：指尖、掌的中心（称为“掌心”）；掌的半径（称为“掌半径”）；手掌的法线（称为“掌法线”）时，掌的开放性（称为“掌开放性”），掌的基部，以及掌心与手轮廓的距离。

[0052] 本文所用的术语“虚拟表示”与“表示”对应于交互对象在虚拟世界中的数字图示。此交互对象可以是来自虚拟世界本身的对象，例如，来自与用户可以与之交互的操作系统的 GUI 的文件夹。此交互对象也可以是来自真实世界的真实交互对象在虚拟世界中的数字图示，例如，用户的手，其位置随时间被跟踪，从而实现在虚拟世界中与操作系统的 GUI 的定点或鼠标定点交互。

[0053] 本发明总体涉及一种用于提供无接触自然交互系统的设备和方法，并且，更具体地，涉及一种用于提供自然交互系统的装置和方法，该自然交互系统使用从手指和 / 或手本身的 3D 位置和 3D 手势和 / 或手的参数确定的控件以无接触的方式控制至少一个鼠标指针。特别是，利用根据定点和激活手势（其不仅可顺序执行，也可以同步执行）提供合适上下文反馈的系统，可以检测健壮和可靠的定点和激活手势，例如“定点”，“捏”，“抓取”或“拍”手势及其衍生物，以改进 3D 自然交互的人体工程学。

[0054] 本发明的方法并不限于经典鼠标类交互，因为它基于 3D 手势识别，3D 手势识别可以提供多点触摸屏交互，或者在有效检测定点和激活操作时，提供比传统鼠标更自然的交互。

[0055] 根据本发明，提供一种使能人与计算机化系统关于（从由测距相机装置提供的数据提取的）手势信息（即，深度地图或 3D 点云（也称为一组顶点或顶点））进行主观自然交互的方法和系统。交互利用预定的手参数，即手掌中心和指尖，以确定至少一个指针，并且，优选地至少两个指针，以便与相应设计的 GUI 进行上下文交互。

[0056] 该方法包括以下三个主要步骤：

[0057] 第一步骤包括：检测至少一个手参数及其奇异 POI，该 POI 至少包含掌心和指尖。该第一步骤的目的是要确定多个相关的 POI，即，奇异 POI 或 POI，其可以被用于单独地或组合地确定可用于进一步交互的指针。

[0058] 第二步骤包括：随时间跟踪和分析选择的至少一个 POI 或其相关指针，以检测手

势交互。根据所选择的 POI 的数量,至少一个定点手势可以被确定。如果根据本发明的一个优选实施例,使用至少两个 POI,可以与定点手势独立地或同步地确定“捏”,“抓取”和“手指拍”手势和其潜在的相关交互,如“点击”或“双击”。

[0059] 在此步骤中,有多种手势被识别,例如:

[0060] (i) “定点”手势,其使用手上的单一 POI,其中这种单一 POI 可以是指尖,手尖,手掌中心等。

[0061] (ii) “定点”手势,其使用手上的至少两个 POI,其中,在一个实施例中,两个 POI 位置之间的内插有效地提供了另一 POI 作为指针。用于内插的两个 POI 可以是:两个指尖;一个指尖与手尖,优选地,例如为拇指,因为从人体工程学角度上,拇指与指尖相对;一个指尖与手掌中心。在另一实施例中,使用两个 POI(例如指尖与掌心)之间的方向矢量来确定“定点”手势。将理解,任何两个合适的 POI 可以用作内插的基础或用于确定方向矢量。

[0062] (iii) “激活”或“活化”手势,其使用:至少一个单一 POI,例如,使用掌心的“推”手势;至少两个 POI 以执行“捏”和/或“抓取”;或者移动 POI 和固定 POI 以执行“拍”手势。

[0063] (iv) 同步“定点”和“激活”,其使用至少两个 POI,使用例如掌心进行定点,使用一个指尖在“拍”手势中进行激活,或使用两个指尖执行“捏”和/或“抓取”手势。

[0064] (v) 同步定点和激活,其使用至少三个 POI,使用:例如,掌心和两个指尖,掌心和两个指尖,两个指尖和内插 POI,或者两个指尖和掌心与内插 POI。在一个优选的实施例中,“定点”指针的位置是从两个 POI 内插的,并且当两个 POI 被同时用于执行除“定点”手势以外的手势时,该定点指针相对于这两个 POI 的旋转和单独运动具有不变性。

[0065] 第三步骤包括使用指针和手势,所述手势是结合具有 GUI 的系统检测到的,GUI 包括虚拟对象,并且提供上下文反馈,以使能符合人体工程学的、直观和自然的交互(例如,当与用户的手的 POI 的表示交互时,使目标高亮显示和挤压目标的外形)。

[0066] 根据本发明的一个实施例,图 1 示出了一个流程图,其示出该方法的三个主要处理步骤,即,检测手参数(步骤 100),确定 POI,并跟踪和分析 POI 的运动和手势(步骤 101),以及根据上下文确定兴趣点的表示与基于自然无接触手势的兼容性 GUI 之间的交互(步骤 102)。从步骤 102 中的根据上下文确定提供反馈控制回路,以在步骤 100 中检测手的参数,如箭头 103 所指示的。另外,从步骤 102 中的根据上下文确定到步骤 102 的确定 POI 并跟踪和分析其运动和姿势的另一反馈控制回路如箭头 104 所指示。

[0067] 本发明的第一步骤(步骤 100)的目的是检测手参数,尤其包括手尖,即手尖和指尖,以便能够提供至少一个稳定的 POI。优选地,该方法可以进一步包括检测上 POI 和下 POI,从中可以确定稳定的 POI。

[0068] 此外,如果尚未可用,该方法还可以包括确定以下参数中的至少一个:掌半径;手开放性;以及手掌法线。

[0069] 使用来自所确定的兴趣点中的至少一个兴趣点,该方法进一步包括:确定用于与系统进行交互的至少一个指针,这将在下文中结合步骤 102 更详细地进行说明。优选地,所述方法包括:确定至少两个指针,所述至少两个指针与对应于用于执行交互的手尖的上 POI 和下 POI 相关。

[0070] 本发明的第一步骤中使用的输入信息至少包括:深度地图或相应的 3D 点云(即,

一组顶点),该 3D 点云包括用户的至少一只手和标签图像。[3D 点云被分割为多个集群,其中每个集群被标记]。所识别的待使用的手的标签和预定义值用来确定待确定的控制点的数量(该预定义的控制点的数量可选地由用户界面通过箭头 103 所指示的反馈控制回路(如果使能)提供)。另外,也可以使用预定义值来确定待确定的指针数量(该预定义的指针的数量此可以由用户界面通过由箭头 103 所指示的反馈控制回路(如果使能)提供)。

[0071] 本发明的第一步骤,步骤 100,包括如图 2 所示的四个处理子步骤,即,检测和确定手掌中心和指尖(步骤 200),确定兴趣点(步骤 201),利用该 POI 确定指针(步骤 202),以及确定和/或精炼其它的附加手参数,如开放性和手掌法线(步骤 203)。

[0072] 步骤 200 包括:首先,检测和确定手掌中心,这需要事先计算一“距离变换”,其输出为所识别的待使用的手的距离地图。该手的距离地图是一个图像,其中手的各像素具有等于与手的最近边界的度量距离的值,如在图 8 中示出的。具体地说,在图 8 中,所标记的待使用的手与该用户的其它无用标记部分分离,并与背景分离。在该图中,在手的表示中每个像素的内容越暗,则该像素距所标记的手的边界越远。可以利用两次通过手的二进制标签图像来计算出这样的距离地图图像。

[0073] 几种方法可用于执行距离变换。然而,在本发明的重要的一点是,针对每个像素获得的距离值可以与真实世界中的一度量距离相关,这是因为标签图像的每个像素对应于该 3D 点云的一个位置,并且该距离地图中的每个距离对应于在真实世界中的距离度量。例如,可以使用两次通过方法来执行该距离变换,其中第一次通过是从二进制标签图像的左上角到右下角,并且第二次通过是相反的方向。对于每次通过,针对每个像素,距离被累加,其对应于由该次通过已处理的所有直接邻居像素的最小值的总和。例如,在第一次通过中,一个像素的直接邻居可以是左侧的像素、对角左侧的像素、上侧的像素和对角右侧的像素。针对当前像素确定的距离可以被确定为这四个邻居像素的最小值加一。此外,考虑到像素的基本几何特性,水平/垂直邻居像素的值可通过预定的因子被加权,例如,通过值 5,而对角邻居像素的值可以通过另一预定加权因子被加权,例如,通过值 7。7/5 的比率的值对应于典型方形像素($7/5 = 1.4 = \sqrt{2}$)的对角和侧面之间的保持比率。

[0074] 为了找到或逼近掌心位置,该方法包括:至少确定其中距离信息强度是手的所有其他像素的最大值的像素位置,并且,在优选的实施例中,可以考虑最小化与先前确定的掌心的距离,以用于实现先前确定的掌心位置的位置平滑变化。

[0075] 其次,步骤 200 包括:使用在手的 3D 点云上进行的主成分分析(PCA)来确定指尖。使用矢量(其中矢量的原点被设定为掌心,矢量的方向被设定为处于由 PCA 给定的主方向)来计算直方图,从而计数手的正交于该矢量且可以投影到该矢量的各个部分的像素的数目。所述矢量的部分可以例如通过 1mm 步长的距离度量表示。使用这样确定的直方图,则指尖可被确定为与矢量中手像素的最小值已被投影的部分相对应的区域。在一个优选的最健壮的实施例中,在两个预定的阈值之间的范围可以被限定,例如,95 至 99% 的范围,对应于直方图中具有较少的投影像素的部分,从而对应于手的沿着手在空间中的主方向(被认为是指尖)的最厚部分。

[0076] 第三,步骤 200 包括指尖的三维确定,即对应于接近指尖的手末端的区域。为此,“响应图像”被确定,如图 9b 所示。在响应图像中,像素的值表示被考虑的像素与其八个邻居之间的深度差。高值表示是末端的概率很高。每个像素的值从沿着共同的八个方向(即,

上,上左,右,下右,下,下左,左,上左)的八个邻居计算。在每个方向上,与当前点具有 2cm 的距离的像素被考虑。为了找出在每个方向上,哪个像素对应于距离为 2cm 的投影,成像设备或相机的说明书连同当前点的深度值被使用。当然,这种操作只对有效像素进行。对于这八个像素,其与当前像素的深度差被确定,并且当前像素的值被视为第三最小值。选择第三最小值是因为,任何非末端由于未连接到一侧上(例如,手掌的一侧)而具有大的值,并且期望手的那些部分不提供作为手指的响应。然而,即使是手指,也可以具有在它们连接到手的一定方向。由于这个原因,具有两个最小响应值的两个方向被忽略。

[0077] 不幸的是,某些情况下,将不提供对手的每个物理末端的适当的响应,例如,紧攥的拳头将提供对应于指尖的一个单一响应,如图 5d 所示;并且“停止”(伸开手掌,如图 3f 和图 5a 所示)将只提供针对拇指的响应,而不提供针对其它单个手指的响应,而是针对粘在一起的所有其他手指提供一个单一其他响应,例如,如在图 11a 和 11b 中的响应图像上的 1cm 半径的所示的。在图 10a 和 10b 中,1002 表示指尖的圆,在响应图像(图 10b)中,其中,从 1001(图 10a)的深度图像处可以清楚地看出,没有明确定义的末端。这说明了本发明方法的健壮性在于无论手姿态是什么,给定用于“定点”的至少一个 POI,它总是可以提供有效的解决方案。

[0078] 下面在图 11a 和 11b 所示的例子中,仅有针对拇指的一个单一响应和针对粘在一起的所有其他手指的一个单一反应,这提供两个可靠的 POI,其可以被选择为“抓取器”用于进一步的交互。

[0079] 此外,平滑滤波器可应用于响应图像以清理每个单独的响应,例如,可以使用已知的典型高斯滤波器。

[0080] 此外,根据平滑响应图像和 y 图像,可以确定连接的部件的图像。这对应于当且仅当两个相邻的像素都在其响应图像都具有正响应值,即它们是末端,并且深度足够接近(<1 厘米)时,标签图像给予这两个相邻的像素以相同的标签。换句话说,如果两个相邻的像素的深度值的差大于预定阈值,例如 1 厘米,或如果它们中的一个在末端响应图像中不具有正响应值,这两个相邻的像素将不会有相同的标签。此图像然后可用于两个目的:第一个是允许将末端响应组合到一起以形成指尖候选,并从图像中去除不想要的小的响应。如果部件比预定阈值小,例如 1 厘米,它可能会被丢弃。

[0081] 在一个实施例中,确定 POI 的步骤包括:通过从先前确定的手参数(至少包括:掌心,指尖,指尖,或这些参数的组合)中选择至少一个点,来确定至少一个单一 POI。这需要例如定义这些参数中的单个作为 POI。这可能需要例如定义这些参数中的两个作为 POI。这也可能需要定义手参数的一个集合或子集的平均位置作为 POI。

[0082] 在另一个实施例中,确定 POI 的步骤包括:通过选择手参数中的至少两个或一个手参数与其他几个手参数的组合来确定至少两个 POI。优选地,考虑到使用两个兴趣点的交互,例如,用于“抓取”时,第三 POI 可以被用于“定点”,并可以根据这两个 POI 来确定。

[0083] 具体地,根据对应于指尖的标记反应来确定 POI,可通过如下方式来实现:基于到其它部件的距离(部件离得越远,其权重越大),在响应图像中计算针对每个标记响应(即部件)的权重。响应图像需要包括至少一个标记部件。在其中预期具有两个指针的情况下,响应需要被分成至少两组。每个组将被用于计算指针或“抓取器”中的一个的位置。为此,有必要找到末端响应图像中具有正响应的手的点中的中心点。如果将使用在末端响应图像

中的响应的对手的所有点的简单加权和作为权重,获得的掌心将偏向手的上部,因为该区域有更多的手指可见,因此总体上具有更强的响应。为了避免这种情况,不仅响应的权重被使用,而且还使用与响应所属的部件相关的因子。

[0084] 如果一个部件远离所有其他部件,它将有较大的权重,并且如果它靠近其他部件,它会具有较小的权重。这确保了分离的手指(在大多数情况下,为拇指)获得比其余手指(如果它们是可见的)更高的权重。如果只有一个其它手指是可见的,则定义这两个手指将具有相同的权重。距离由手跨度标准化,手跨度被计算为手半径乘以例如因子 5。

[0085] 末端的中心被计算为使用其加权响应和与其部件相关联的权重的加权和。

[0086] 一旦中心被确定,它被用来分离在 3D 笛卡尔空间中的响应。为此,穿过所确定的中心点的平面被使用。由于目标是,优选地,拇指在该平面的一侧上和其余手指在另一侧上,从一个兴趣点(或“抓取器”)到另一兴趣点的方向被视为垂直于该平面,再加上中心点,提供平面的完整定义。

[0087] 然而,为了更好地随时间跟踪和提供可靠兴趣点(当将这些点视为“抓取器”时),即保证可靠地确定它们各自随时间的位置,可以优选地使用其前一帧 $t-1$ (其中当前帧为 t) 的位置。对于第一帧,向上轴被用作参考。这种选择没有影响,因为即使初始轴是错误的,该算法也会收敛到正确取向。选择向上轴是因为它是通常姿态中“抓取器”最可能的大体方向。

[0088] 考虑到在上一步骤中的平面,手的点根据其所在的平面的侧被排序,并且,针对每个组使用该重心,则可以获取针对两个“抓取器”的最佳候选,例如,那些更加接近各重心的点。

[0089] 然后,指针可被确定为以下之一:指尖;掌心;以及 POI。在一个优选的实施例中,在确定了“抓取器”之后,可以计算在“抓取器”之间的附加稳定定点坐标。它可以是“抓取器”的中心,由掌中的移动进行调整。然而,该点的值仅当掌心已经移动超过例如 3 毫米时被更新。这保证了,即使用户“捏”,如果他/她不移动他/她的手,则指针不动。为了避免抖动,这个过程不是二进制的,相反,对于 0 和 3mm 的范围内的手运动,通过新的“抓取器”的中心和前一帧的稳定化点之间的简单线性回归来变平滑。

[0090] 此外,手掌法线可用于确定手掌本身的方向。使用从掌心到指针位置的矢量与先前确定的手掌法线之间的角度来调节手掌法线。该角度是根据一些自定义参数化被加权,以在这两个信息集合之间进行平滑混合。手掌法线使用 PCA 来计算,最少重要方向是手掌法线。法线还混有诸如相机方向之类的信息。

[0091] 此外,可以从距离地图中获取手掌半径,即,距离地图的值除以预定因子,例如 5,以提供到手的最近边界的距离的估计,并且,使用相机的说明书和掌心的当前位置的深度值,可以以毫米精度估计手掌半径。

[0092] 另外,还可以获取作为一个布尔值的手开放性,其表达手是伸开还是闭合的。如果未检测到指尖,则手被认为是闭合的。可以使用在“抓取器”距离和手掌半径之间的比率来更新手开放性。如果该比例小于预定阈值,并且之前认为手是伸开的,则手的开放性将被调整,并标记为闭合。如果该比率变得大于另一预定阈值,手会被认为是伸开的。

[0093] 总之,本发明的第一步骤的输出包括:优选地,确定至少一个 POI,该 POI 的 3D 位置将被用于执行基于的 3D 手势的交互,优选地,确定至少两个 POI,其位置被用于基于 3D 手

势识别的、兼容单手同时执行定点和激活 3D 手势的交互。

[0094] 本发明的第二步骤（图 1 中的步骤 101）包括随时间跟踪和分析选择的 POI 或与其相关的指针，以检测的手势和交互。根据在至少一个单手上选择的 POI 的数量，可单独或同时地独立确定不同的手势。

[0095] 对于手势确定，首先将各个 POI 在空间中随时间的位置来用于跟踪和定点目的。特别是当使用由至少一个指针表示或优选地由两个指针表示的至少两个 POI 时，帧间跟踪确保一致的“定点。在一个实施例中，可以使用单一 POI。这 POI 可以被指定为用于使能定点特征的指针，并且可以被确定为上述前一步骤中检测到的指尖之一。例如，如果没有指尖可用，该指针可以是掌心；如果没有可靠的指尖可用，该指针可以是指尖；如果食指尖可用，则该指针可以是食指尖，因为它自然地用于定点，或者如果需要，该指针可以是任何其他指尖，或使用选自掌心、指尖和指尖的预定手参数的至少两个来计算机化其在 3D 空间中的位置的另一端。根据本发明的一个优选实施例，GUI 上的每个指针表示或至少一个兴趣点的位置的改变的分析还可以使用常规的手势识别方法来分析，所述常规的手势识别方法用于确定如下形式的动态手势，例如，“波”，“推”，“扫”，“圆”，“手指拍”或另一动态手势。确定这种动态手势用于触发将用作交互 GUI 的输入的事件。这可以与至少一个指针和由该指针指向的对象的位置相结合，但是这种结合不是必需的。

[0096] 第二，在手势确定中，确定的 POI 在给定的时间在空间中相对于彼此的位置用于确定手的姿势。其中手的位置不随时间改变的不变手姿势被确定为静态手势。静态手势是手参数的一个子集在空间中的布置。例如，拇指未连接的伸开的手可以对应于“停止”姿势或对应于其布置由拇指尖、掌心和指尖之间形成的直角定义的手势。可以使用在本领域中已知的多种方法来实现手势的检测，例如，使用形状匹配技术。然而，根据本发明的优选实施例中，确定手姿态包括检测 POI 的一个子集在空间中的布置的几何形状。每个确定的 POI（包括指尖和手掌中心中的至少一者）分别连接到彼此。POI 之间的链路的距离和方向以及它们的交点的特性（角度）被计算，以确定一组数据参数。根据 POI 的数量和它们所定义的数据参数，通过与包含待检测的主姿势的数据库进行比较来确定位置。对于每一帧，如果检测到了预期的手姿势之一，则可以触发事件以驱动交互系统。

[0097] 第三，在手势确定中，确定的 POI 在空间中的位置随时间相对于彼此的改变，也就是，在手的姿势随时间的改变，用于确定采用如下形式的手动态手势，例如，“捏”，“抓取”，“点击”，“拍”或其它手势。

[0098] 例如，在一个本发明的实施例中，该方法包括分析定义 POI 的至少两个指尖以及用于使能“捏”动态手势检测的相关指针。如上所述，“捏”可以对应于两个相对的尖端，这两个相对的尖端可选地显示为能够朝着彼此移动的指针，两个指尖或相关 POI 或指针之间的距离允许确定“捏”的程度。

[0099] 在本发明的另一个实施例中，如果“捏”为其中至少两个“抓取器”触摸彼此达一定时间段，然后返回到非接触状态 / 姿势，则可以确定为“点击”动态手势。

[0100] 当与在用户界面中表示的虚拟对象进行交互时，可上下文检测“抓取”手势。“抓取”包括检测至少两个 POI，这两个 POI 在 GUI 中可以以两个指针的形式表示。当两个指针触摸虚拟对象的两个相对侧时，可以检测到“抓取”。

[0101] 如果用户界面使能“推”手势的话，当至少一个指针的表示触摸虚拟对象表示时，

可以检测到“推”手势，并且指针的继续运动在 GUI 内推动虚拟对象表示。

[0102] 如果用户界面使能“手指拍”或更简单的“拍”手势的话，当在一只手上选择至少两个 POI 时，可以检测到“手指拍”或更简单的“拍”手势。第一 POI 优选对应于掌心或手的质心的近似，提供了可用于定点手势的 3D 位置，而第二个 POI 优选对应于一指尖，如食指尖，提供了用于“激活手势”识别的 3D 位置。该 POI 在确定速度下以确定幅度和持续时间的上下自然运动定义了可以用作可靠事件触发交互的“手指拍”（或“拍”）手势。该特定实施例的可靠性依赖于一个事实，即指尖和手掌中心尽管在同一只手上，但是两者的位置彼此无关。

[0103] 本发明的第三个步骤包括一种使能与 GUI 的人机无接触交互的方法和系统。该方法涉及用于控制界面的手段，该系统涉及用于显示界面和视觉反馈的手段。该方法和系统获益于包括距离测量值和相关信息在内的输入信息，例如从 POI 提取的定点和激活手势，其中相比于传统的 2D 无接触交互系统，该 POI 独立于手到测距成像系统或相机的距离。在本发明中，交互依赖于使用上述步骤确定的手 POI 和手势随时间的时空特性分析，特别是，对 POI 的虚拟表示的分析，即与虚拟对象上下文相关的指针，例如，“定点”交互，“抓取”交互，操纵交互（例如，“捏”，旋转，“拖放”，“推”），然后点击交互（例如，“手指拍”或“拍”）。

[0104] 下面的描述涉及的实施例例示本发明的不同方面，这些实施例应被视为非限制性的使用在上述步骤中提取的手参数来与屏幕上显示的虚拟对象进行交互。

[0105] 在描述“定点”交互的第一实施例中，该方法可以包括使用奇异 POI，即上面确定的掌心作为参考 POI，来确定指针表示在虚拟世界中的位置。手掌的位置随时间的改变触发其表示在改变其虚拟空间中的位置随时间的改变。在真实世界中的位置到虚拟世界中的位置的变换可以是使用 1:1 绝对比率的线性计算，或者使用 1:x 的相对比率的线性计算（其中 x 因子可以由用户预定），或者使用例如由用户预定义的多项式变换比率的非线性计算，或者可具有依赖于手到成像装置之间的距离的比率。

[0106] 在另一个实施例中，一个类似的方法可以包括使用指尖参考位置来确定指针表示在虚拟世界空间中随时间的位置。在这种特殊情况下，无论手指的配置如何，即，无论手指都结合在一起、部分地结合在一起还是彼此全部分离，指尖都被确定。手指的配置的示例示于图 3A 至 3F。

[0107] 在另一实施例中，一个类似的方法可以包括使用拳头参考位置（图 3c）来确定指针表示在虚拟世界空间中随时间的位置。这种特殊情况可以对应于无法检测到单个指尖的手，并且例如，手到测距成像装置的最接近的位置被确定为拳尖。

[0108] 在一个优选的实施例中，一个类似的方法可以包括使用至少一个单指尖参考位置来确定指针表示在虚拟世界空间中随时间的位置。这个在实际生活中特别常见和自然的使用单一指尖指向一事物的方式可以对应于例如使用“食”指尖随时间的位置，或在另一个实施例中，使用通过例如指尖的主方向确定的指尖方向。

[0109] 在另一个优选的实施例中，该方法可以包括使用多个手参数来确定单一指针表示在虚拟世界空间中随时间的位置。例如，该方法可以使用至少两个手参数的加权和来确定单一指针表示的位置。单一指针位置可以对应于所选择的手参数之间的位置。例如，它可以对应于指尖和分离指尖（例如，拇指尖）之间的位置，到所提到的尖端的相应距离由预定加权因子调制。具体地说，当考虑到分离指尖（例如，拇指尖）和另一指尖或指尖是，“捏”

运动的不对称性可能要求与分离指尖位置相关的加权因子比与另一尖端相关的加权因子更大,因为该另一尖端移动的幅度将大于拇指尖移动的幅度。这种特定的加权计算(或内插)目的在于获得可靠且稳定的用于定点手势的指针位置,当两个 POI(从中可得到指针的位置)正在执行另一手势,例如,激活手势,如“捏”和“抓取”手势,下指针位置被视为例如其中执行最小的运动,因此包括较大加权因子。当这些 POI 正在执行激活手势时,(从两个 POI 得到的)指针位置相对于单个 POI 的独立运动和旋转是不变的。

[0110] 再如,确定指针的合适加权因子是为了考虑对应于这些指针的每个手参数与掌心的距离;距离越近,其被视为移动得越少,因此具有最大的加权因子。另一示例特别适合于当期望加权相对于手取向不变且独立于指尖和指尖的标识的情况。

[0111] 在另一个优选的实施例中,该方法可以包括使用多个手参数来确定多个指针表示在虚拟世界空间中随时间的位置。例如,多个指针表示可以包括至少两个手参数的独立表示,例如,对应于指尖的“指针”和对应于拇指尖的“指针”。在衍生的实施例中,可以确定多于两个指针。另一个简单的示例对应于确定待在虚拟世界中表示的三个指针。这三个指针可以是指尖、拇指尖和对应于这两个手参数位置之间的加权位置的另一 POI。

[0112] 在又一实施例中,该方法可以包括根据手姿势来自动定义所表示的指针的数量,因为该姿势可以允许或不允许对手参数的可靠检测。

[0113] 在再一实施例中,该方法可以优选地包括:特别根据在所使用的 GUI 中表示的对象,自动定义待利用该 GUI 进行上下文表示的指针的数量。例如,该方法可以定义单一指针,该指针将对应于指尖,以利用指针表示解析屏幕区域。在屏幕区域包括例如文件夹表示(其例如由于可抓住而兼容至少两个指针交互)的特定位置,该界面将定义两个指针表示,这两个指针表示将对应于两个手参数,例如指尖和分离的指尖(例如,拇指尖)。这些指针的运动随后将允许由 GUI 认可的文件夹的“抓取”操纵。

[0114] 在第二实施例中,“抓取”交互,该方法包括使用如上所述被确定为定义参考位置的手参数的掌心来确定单一指针表示在虚拟世界空间中的位置。该方法包括进一步确定手上的两个 POI,例如,指尖和拇指尖。交互则包括指向使能“抓取”特征的虚拟对象。当两个 POI 移动到彼此之间的距离小于预定义阈值时,将发生虚拟对象的“抓取”。此时,虚拟对象的表示将与指针表示的位置相关,并且将相应地移动。当两个兴趣点之间的距离达到第二预定阈值时,虚拟对象将被释放(不再被“抓取”)。

[0115] 在另一实施例中,该方法包括从指尖中确定多个 POI,以及响应地确定两个指针,从而获取这两个指针的视觉反馈,这两个针对可对应于例如指尖和拇指尖。随后,交互将包括指向使能“抓取”特征的虚拟对象。当两个 POI 各自关于特定限制(这些限制涵盖某些预定阈值,例如以补偿相机或测距成像装置的潜在准确性问题)开始接触虚拟对象表示的边界时,将发生虚拟对象的“抓取”。此时,虚拟对象的表示将与指针表示的位置相关,并且将相应地移动。当至少一个指针和对象的边界之间的距离达到第二预定阈值时,虚拟对象将被释放(不再被“抓取”)。

[0116] 在另一个优选的实施例中,该方法包括使用从如上述的手参数选择指针,该指针可以对应于一特定 POI,该 POI 的位置被确定为在两个其它的 POI 之间的位置,这两个其它的 POI 可以对应于例如拇指尖和食指尖。当指针的表示到达一区域或对象的表示符合“抓取”特征时,指针然后可分裂成至少两个指针,所述至少两个指针与从中确定该特定 POI 的

至少两个 POI 相关。虚拟对象随后可以被“抓取”并如上所述被操纵。指针的表示保持可见,直到满足另一个条件,例如,所述指针彼此相隔大于预定阈值的距离。当“抓取”的条件不再相关时,重新建立先前的单指针状态。

[0117] 在第三个实施例中,操纵交互,该方法包括使用从如上述的手参数选择的至少一个指针,优选为两个指针,例如,所述指针可以对应于至少一个指尖。在屏幕上表示的指针相对于虚拟对象的位置可用于触发、控制和停止对虚拟对象的至少一个参数的操纵。例如,触发操纵的开始和结束可以链接到两个指针与虚拟对象的边界的距离;操纵可以包括根据指针的平均位置来确定虚拟对象的位置;虚拟对象的取向可以根据关于显示屏幕的水平和垂直轴确定的从一个指针到另一个指针的方向来确定。

[0118] 在另一个实施例中,该方法可包括“点击”交互,其使用多个手参数,如果满足一组特定的时空条件,则触发一个事件。例如,当指针位于屏幕上在预期将发生“点击”事件以触发诸如启动应用之类的系统事件的区域时,例如从一文件夹处,当两个 POI 之间的距离或两个显示的指针之间的距离变得小于预定义阈值达一最大预定义持续时间时,可以触发“点击”手势的检测和相关事件。

[0119] 在一个优选的实施例中,当指针比虚拟对象的尺寸更靠近时,GUI 的虚拟对象可以提供反馈,例如,根据指针间的距离,可以挤压对象。

[0120] 优选地,交互系统还可以提供视觉反馈信息,例如根据执行的交改变指针或对象的渲染。例如,当对象被两个指针“捏”时,可以生成虚拟对象挤压的视觉反馈。

[0121] 此外,通过上下文分析指针和虚拟对象各自的位置,可以优化交互,从而改进由虚拟对象使能的交互的效果。例如,使能“抓取”操纵的靠近虚拟对象的两个指针可以被吸引到虚拟对象的边界处,从而改善用户体验。

[0122] 此外,为了补偿测距成像装置的准确度,或为了改善用户体验从而允许在虚拟空间中并未进行准确接触的情况下与对象交互,可以引入容差。

[0123] 更优选地,虚拟对象交互能力可能在交互系统启动时预先设定,并根据交互和由用户执行的操纵而动态地改变。例如,对象可以只能旋转而成为旋钮控件,对象可以只能在一个方向上转换并成为滑动器,对象可以只能被按压并成为开关,对象仅在指针离得不是太远时出现,或者对象可以根据其被“抓取”的位置而呈现不同的行为。

[0124] 为了完整起见,图 4 示出了通过将至少两个指尖向对象的表面聚拢来“抓取”对象的一种方式。在本实例中,指尖包括拇指尖和食指尖,但也可以是拇指尖和中指、无名指或者小指的指尖。当抓取真实对象时,人们通常毫不困难地将两个或更多个手指向对象的相对侧聚拢,而当并非旨在真实对象而是在空气中执行抓取手势时,人们通常执行明显不对称的手势。通过表示虚拟对象和至少两个指针的视觉反馈(图 13),有可能重获以与真实对象类似的方式将指针向虚拟对象聚拢的能力。

[0125] 图 5a 至 5c 示出了如 501 指示的指尖的定位以及如 502 和 503 指示的掌心的定位,对于一些手的姿势,图 5d 示出了拳头尖和其相关的掌心 503 和指尖 503。

[0126] 图 6 示出了呈现“抓取”姿势的手的侧视图,其叠加有指尖(601)、拇指尖(602)、掌心(603)的定位和利用多个指尖计算出的一附加稳定 POI(604)的定位。

[0127] 图 7 示出了呈现“抓取”姿势的手的正视图,其叠加有指尖(701)、拇指尖(702)的定位以及利用多个指尖计算出的一附加稳定 POI(703)的定位。

[0128] 图 8 示出了手的距离地图,其中,手的外观或轮廓与手指和拇指一起是可见的。手掌与被视为掌心的部分一起也是可见的,其中掌心用手掌中心的高亮区域表示。

[0129] 在图 9a 中示出了呈现“捏开”姿态的手的深度地图 901。这里,对应于食指和拇指尖的 POI 被示出。图 9b 示出了对应于图 9a 中所示的深度地图的响应地图 902。如可以看到的,响应地图 902 具有的针对食指尖的清晰响应信号为上抓取器,针对拇指尖的清晰响应信号为下抓取器。

[0130] 在图 10a 中示出了呈现“捏合”姿态的手的深度地图 1001。只有对应于食指尖的 POI 是可见的。图 10b 示出了对应于图 10a 的深度地图的响应地图 1002,并且仅针对食指指尖获得清晰的响应信号。

[0131] 在图 11 中示出了张开的手的深度地图 1101,手上的 POI 对应于指尖和拇指尖。图 11b 示出了与图 11a 的深度地图对应的响应地图 1102。在本实例中,从拇指和指尖获得清晰的响应信号。

[0132] 在图 12a 中示出了手的深度地图。图 12b 示出了对应于图 12a 的深度地图的反应地图。在图 12b 中,针对手的多个末端获得清晰的响应,在本实例中,为拇指和手指。

[0133] 图 13 示出了一些交互,包括虚拟对象(由 1306 表示)以及与手的一些确定的兴趣点相关的两个指针的表示(由 1305 表示)的表示。正如 1301 所指示的,指针不接触虚拟对象。在 1302,两个指针都接触虚拟对象,虚拟对象以挤压的形状提供一视觉反馈,挤压的程度与所考虑的指针之间的距离相关。在这种情况下,该对象可被视为被抓取,然后可以根据该指针的进一步相应位置被操纵。在 1303,被挤压的虚拟对象可以被视为两个指针被抓住,并且然后可根据指针的相应位置(只要这些指针之间的距离保持等于或小于对象的原尺寸)被操纵。操纵可以是例如以下中的至少一种:例如与 1302 相比在 1303 处所示的旋转、转换、变换(例如缩放)或虚拟对象上下文允许的另一变换。

[0134] 在 1304,虚拟对象被完全挤压。在一个实施例中,这种状态可以被链接到触发对象表示的特定行为,例如,形式的变化,并且它还可以与触发与虚拟对象本身不相关的特定事件相关,例如,将应用的启动链接到用户界面。

[0135] 在图 14 中示出用户的手的侧视图,其与 POI 1401(表示食指指尖的尖)的定位叠加,可用于激活手势。还叠加有另一 POI 1402,其与手的质心相关,并且其可以用于定点手势。箭头 1403 展示“手指拍”激活手势的上下手指手势。

[0136] 尽管已经结合特定实施例对本发明的各方面进行了描述,但将容易理解,这些方面可以以其它的形式实现。

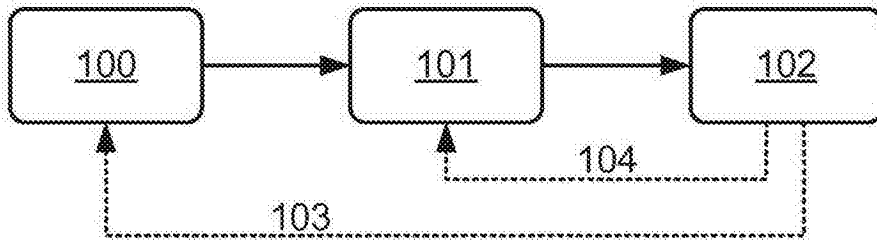


图 1

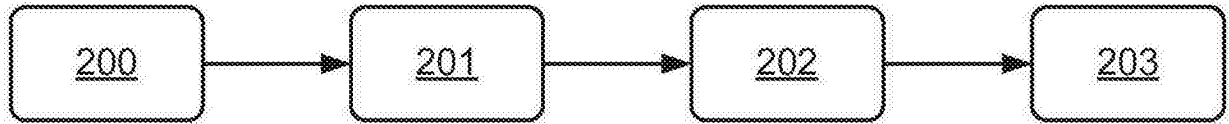
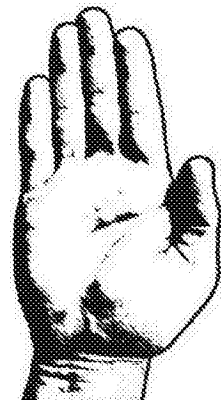
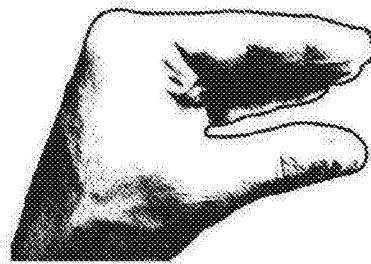
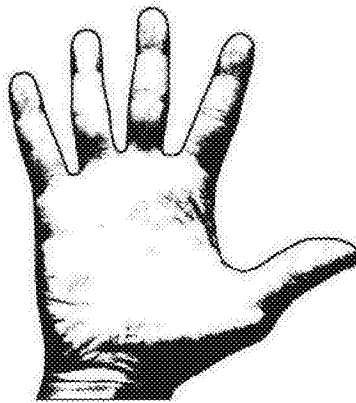
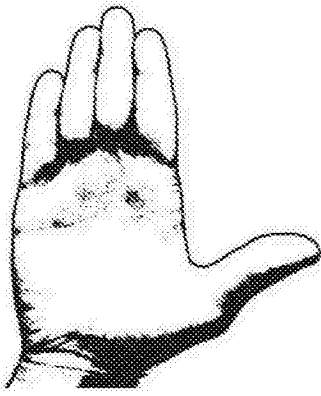


图 2



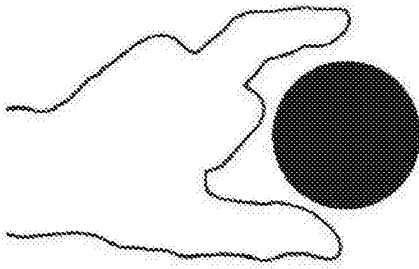


图 4a

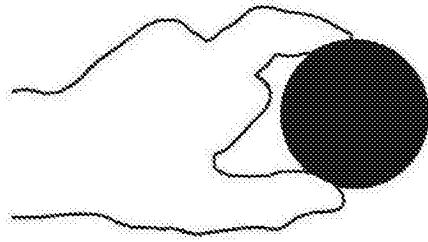


图 4b

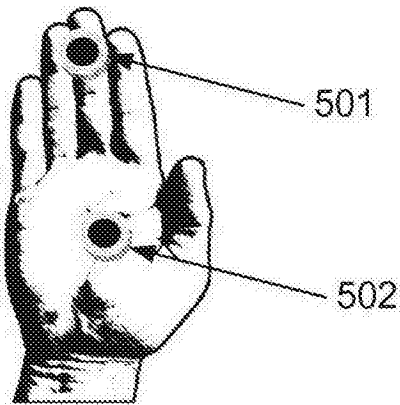


图 5a

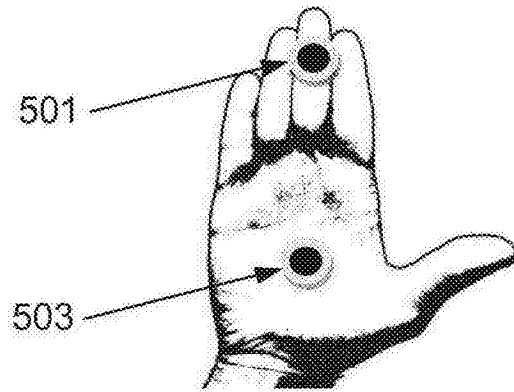


图 5b

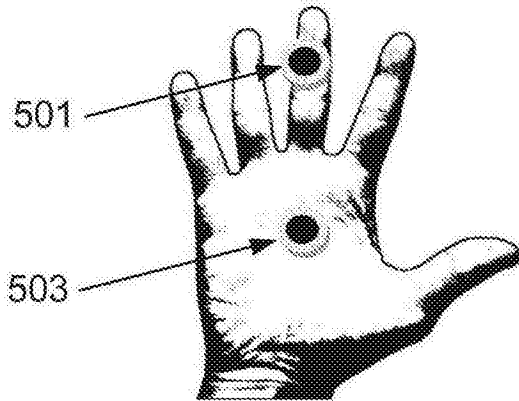


图 5c

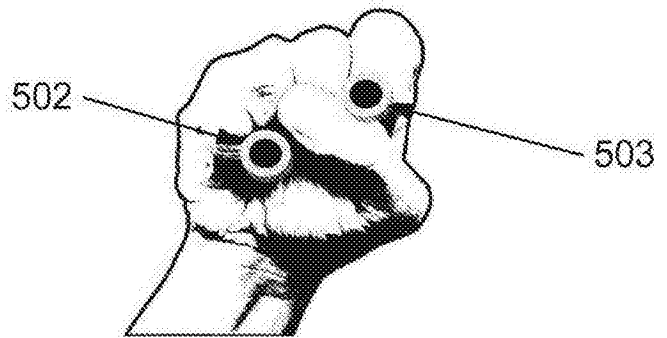


图 5d

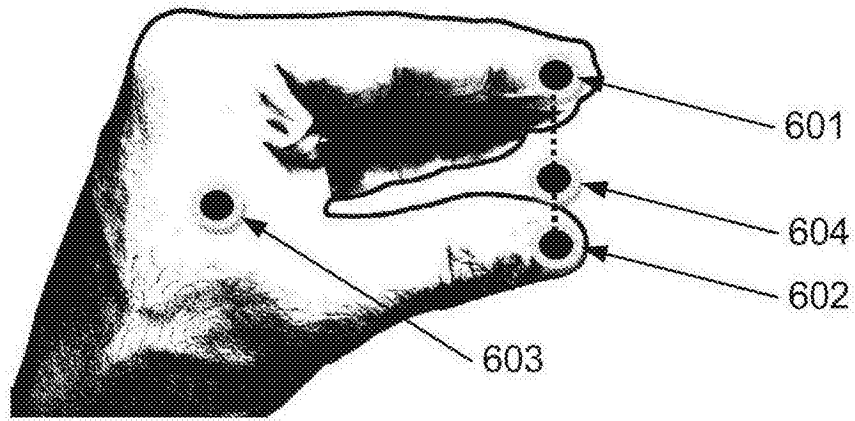


图 6

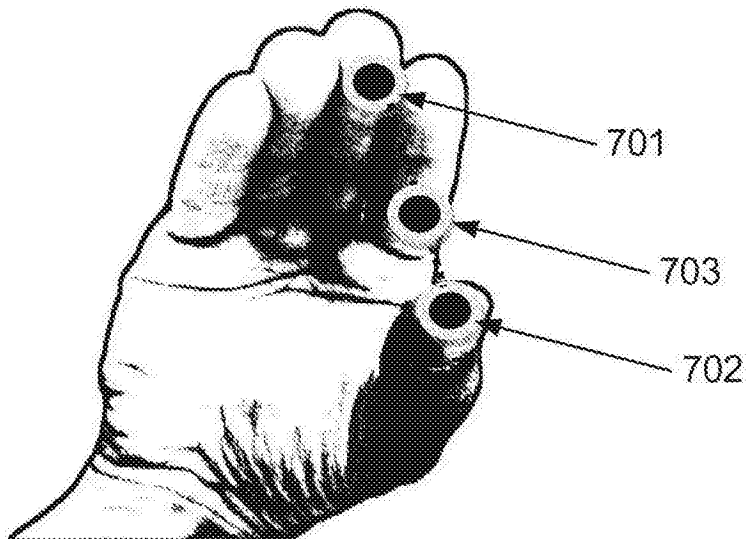


图 7

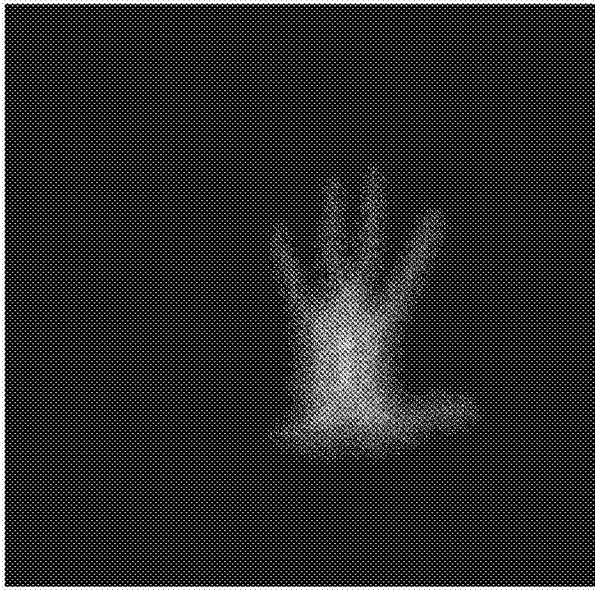


图 8

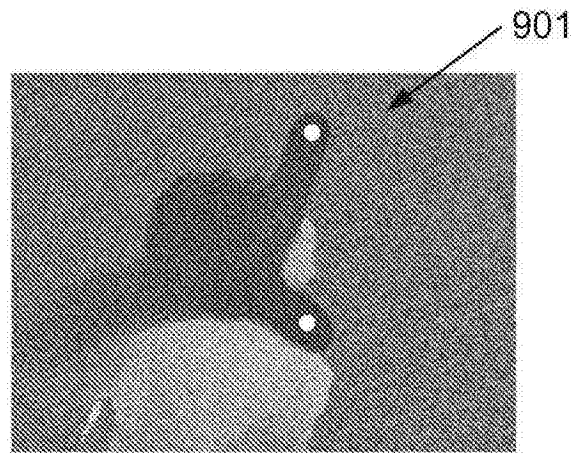


图 9a

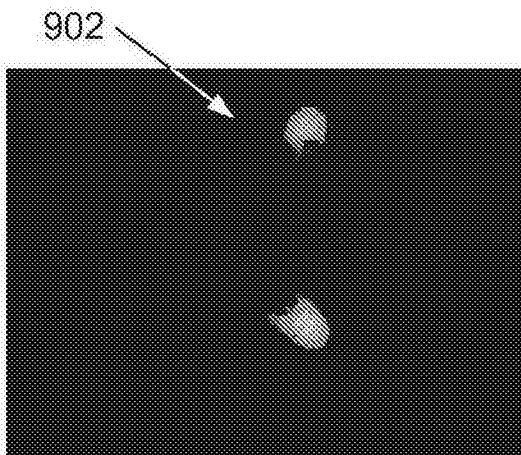


图 9b

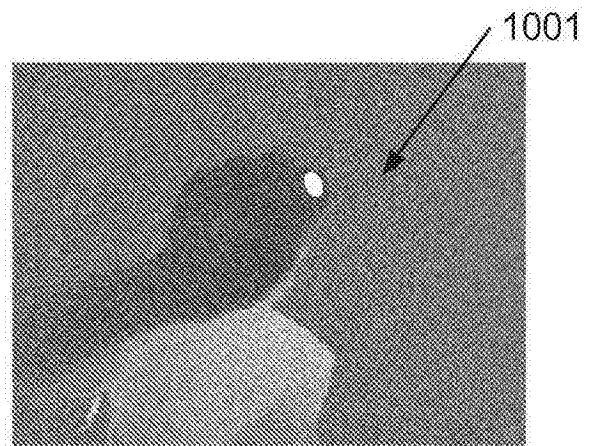


图 10a

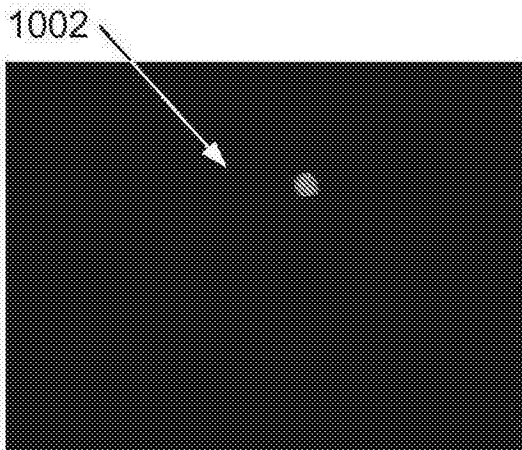


图 10b

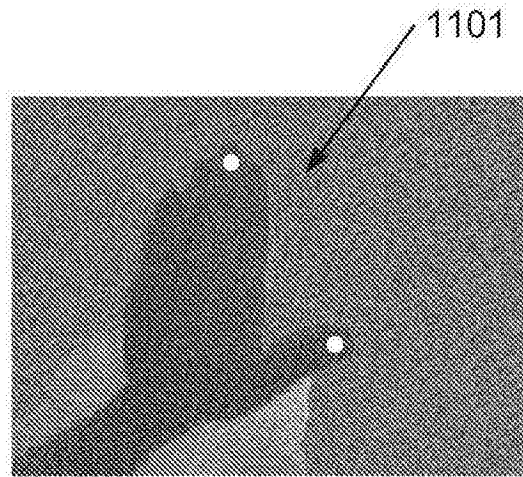


图 11a

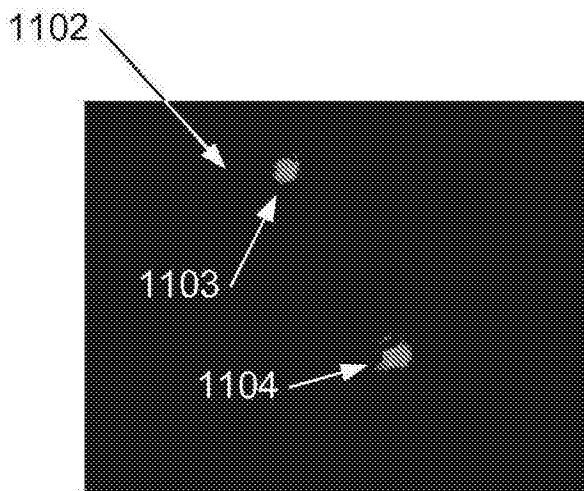


图 11b

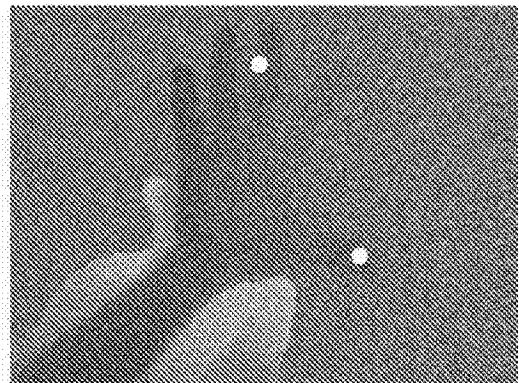


图 12a

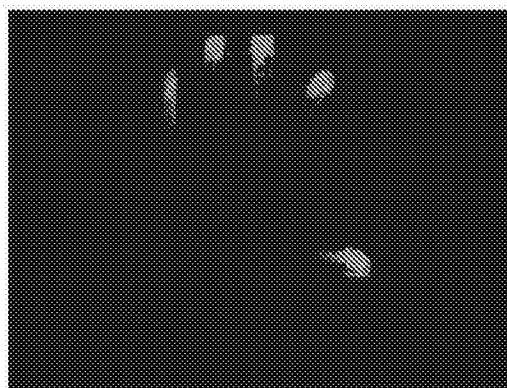


图 12b

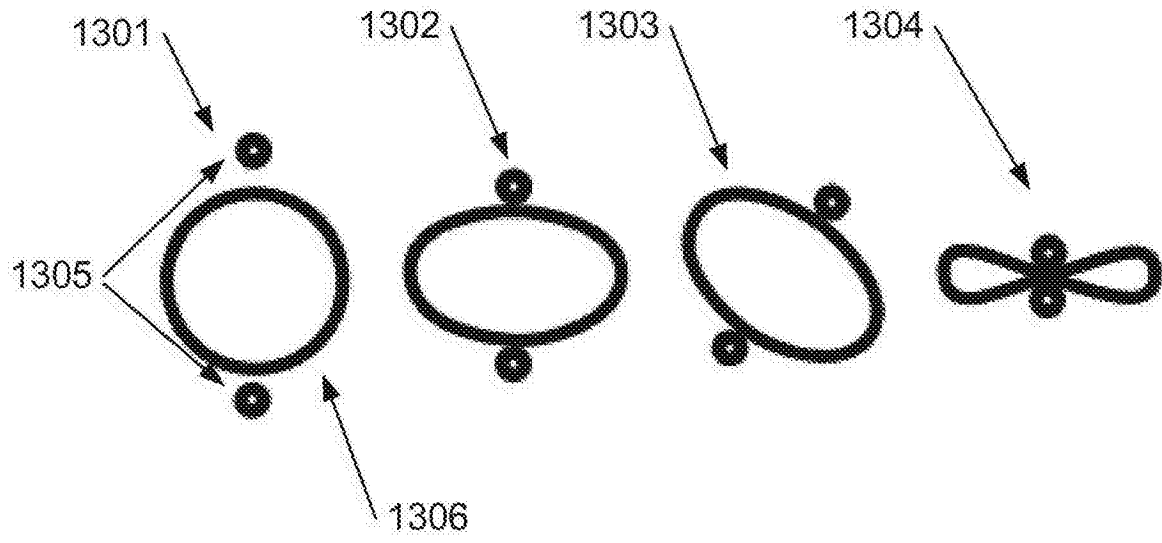


图 13

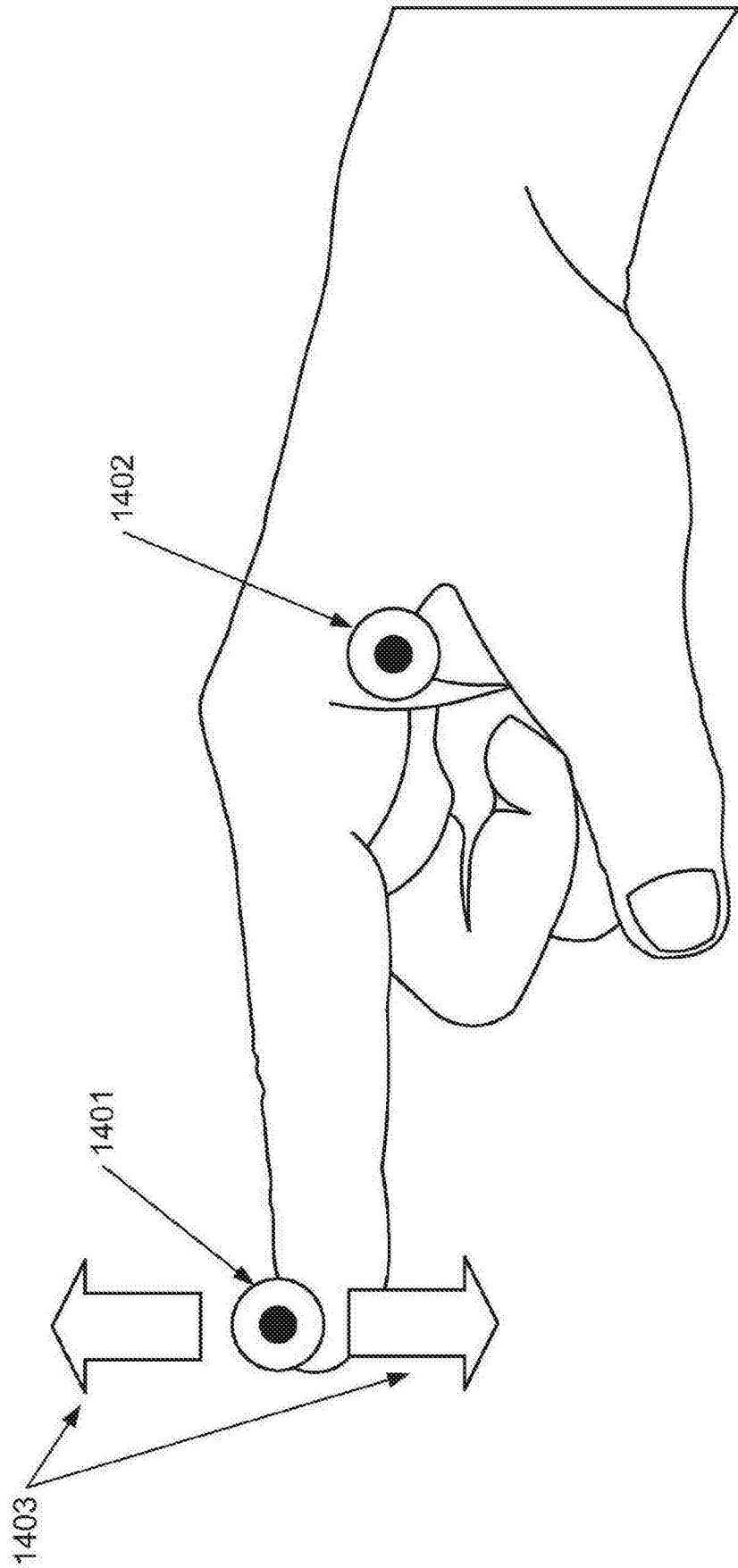


图 14