



(12) 发明专利

(10) 授权公告号 CN 101501614 B

(45) 授权公告日 2012. 04. 18

(21) 申请号 200780029522. 2

(22) 申请日 2007. 05. 29

(30) 优先权数据

11/463, 183 2006. 08. 08 US

(85) PCT申请进入国家阶段日

2009. 02. 09

(86) PCT申请的申请数据

PCT/US2007/012546 2007. 05. 29

(87) PCT申请的公布数据

W02008/018943 EN 2008. 02. 14

(73) 专利权人 微软公司

地址 美国华盛顿州

(72) 发明人 A·D·威尔森 M·J·辛格莱尔

(74) 专利代理机构 上海专利商标事务所有限公

司 31100

代理人 陈斌

(51) Int. Cl.

G06F 3/033 (2006. 01)

(56) 对比文件

US 2006/0136846 A1, 2006. 06. 22, 全文.

US 2002/0118880 A1, 2002. 08. 29, 全文.

WO 2006/020305 A2, 2006. 02. 23, 说明书第 24 页第 1 段至第 28 页第 1 段、附图 10A 至 15C.

US 2002/0140667 A1, 2002. 10. 03, 说明书第 [0124] 段, [0125] 段, [0127] 段, [0129] 段至 [0142] 段、附图 7, 8, 9b, 10a 至 10d, 11a 至 11c, 12a 至 12e, 13a 至 13b, 15a 至 15d.

审查员 刘长勇

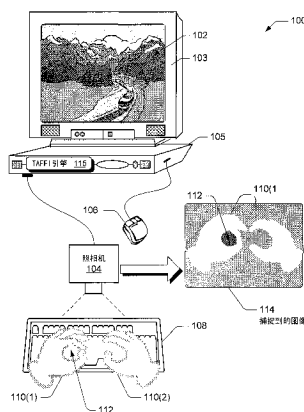
权利要求书 2 页 说明书 9 页 附图 5 页

(54) 发明名称

用于视觉显示器的虚拟控制器

(57) 摘要

公开了用于视觉显示的虚拟控制器。在一个实现中,照相机捕捉相对于背景的手的图像。该图像被分割成手区域和背景区域。各种手和手指姿势将背景的各部分隔离成各独立区域,然后将用于操纵视觉显示的控制参数分配给这些区域。多个控制参数可以与由两个手形成的、用于高级控制的多个独立区域的属性相关联,该高级控制包括点击、选择、执行、水平移动、垂直移动、卷屏、拖拽、旋转移动、缩放、最大化、最小化、执行文件功能、以及执行菜单选项的同时发生的功能。



1. 一种用于经由分配给手的移动的控制参数来改变用户界面上的显示的图像的方法，包括：

经由照相机检测相对于背景的一个或多个手的图像 (502)；

将所述图像分割成手区域和背景区域 (504)；

每隔一段时间对所述背景的多个独立区域进行计数 (506)，其中通过所述手区域中的一个的至少一部分而在视觉上独立于所述背景的其他部分的所述背景的每一部分都被定义为独立区域；

将用于操纵用户界面上的显示的图像的控制参数与每一计数的独立区域相关联 (508)；

经由所述控制参数与同所述控制参数相关联的所述独立区域的形状和 / 或位置的每一改变相关地改变所述显示的图像 (510)。

2. 如权利要求 1 所述的方法，其特征在于，还包括在一相关联的独立区域再次变为连接时将控制参数与该独立区域分离。

3. 如权利要求 1 所述的方法，其特征在于，所述控制参数包括当独立区域在阈值时间间隔内形成、展开、并重新形成时仿真鼠标按钮的动作。

4. 如权利要求 1 所述的方法，其特征在于，所述控制参数包括鼠标指针的仿真和移动。

5. 如权利要求 1 所述的方法，其特征在于，所述控制参数包括光标的移动。

6. 如权利要求 1 所述的方法，其特征在于，所述控制参数包括所述显示的图像相对于所述用户界面的移动。

7. 如权利要求 6 所述的方法，其特征在于，所述移动包括以下各项之一：垂直移动、水平移动、或者旋转移动。

8. 如权利要求 1 所述的方法，其特征在于，所述控制参数包括放大或缩小所述用户界面上的显示的图像的大小。

9. 如权利要求 1 所述的方法，其特征在于，还包括将所述控制参数与两个独立区域之间的关系相关联，每一个独立区域都由相应的手的图像形成。

10. 如权利要求 9 所述的方法，其特征在于，所述控制参数包括与所述两个独立区域之间的变化的距离相关地拉伸所述显示图像。

11. 如权利要求 9 所述的方法，其特征在于，与两个独立区域之间的关系相关联的控制参数控制以下动作之一：与两个独立区域之间的线的旋转相关地旋转所述显示图像；

当两个独立区域都以类似的弧方向移动时以顺时针或逆时针方向旋转所述显示的图像；

当两个独立区域以相同的方向同时移动时移动整个显示的图像；以及

当独立区域之间的距离改变时缩放所述显示的图像。

12. 如权利要求 1 所述的方法，其特征在于，所述手区域中的一个的拇指区域和食指区域定义拇指和食指接口 (TAFI)，所述拇指区域和所述食指区域在视觉上互相接触以便在视觉上将所述背景的一部分与所述背景的其他部分断开，以便形成独立区域。

13. 一种用于经由分配给手的移动的控制参数来改变用户界面上的显示的图像的系统，包括：

用于经由照相机检测相对于背景的一个或多个手的图像的装置；

用于将所述图像分割成手区域和背景区域的装置；

用于每隔一段时间对所述背景的多个独立区域进行计数的装置，其中通过所述手区域中的一个的至少一部分而在视觉上独立于所述背景的其他部分的所述背景的每一部分都被定义为独立区域；

用于将用于操纵用户界面上的显示的图像的控制参数与每一计数的独立区域相关联的装置；

用于经由所述控制参数与同所述控制参数相关联的所述独立区域的形状和 / 或位置的每一改变相关地改变所述显示的图像的装置。

14. 如权利要求 13 所述的系统，其特征在于，所述系统还包括用于在一相关联的独立区域再次变为连接时将控制参数与该独立区域分离的装置。

15. 如权利要求 13 所述的系统，其特征在于，所述控制参数包括当独立区域在阈值时间间隔内形成、展开、并重新形成时仿真鼠标按钮的动作。

16. 如权利要求 13 所述的系统，其特征在于，所述控制参数包括仿真视觉指示符的移动，所述视觉指示符选自鼠标指针图像、光标图像和所述显示图像的至少一部分组成的一组视觉指示符，并且其中所述移动包括选自垂直移动、水平移动、旋转移动、放大移动、以及缩小移动组成的一组移动的移动。

17. 如权利要求 13 所述的系统，其特征在于，所述系统还包括用于将所述控制参数与两个独立区域之间的关系相关联的装置，每一个独立区域都由相应的手的图像形成。

18. 如权利要求 17 所述的系统，其特征在于，所述控制参数包括与所述两个独立区域之间的变化的距离相关地拉伸所述显示的图像。

19. 如权利要求 17 所述的系统，其特征在于，所述控制参数包括与所述两个独立区域之间的线的旋转相关地旋转所述显示的图像。

## 用于视觉显示器的虚拟控制器

### [0001] 背景

[0002] 手移动和手信号是人类表达和交流的自然形式。该知识对人机交互的应用导致了允许人类手势作为计算机输入的基于视觉的计算机技术的开发。计算机视觉是能够以捕捉一个人的手或身体的不受阻碍的运动为目标来实现人类手势输入系统的技术。然而,当前正在开发的许多基于视觉的技术涉及需要非自然手势和附加设备的难以使用的运用。这些技术可以是复杂且庞大的,从而导致由于远离标准的计算机使用位置的重复的手移动而降低效率。

[0003] 当前计算机输入方法一般涉及使用键盘的文本输入和经由鼠标或指示笔的光标操纵两者。键盘和鼠标之间的反复切换随着时间的推移降低了用户的效率。计算机视觉技术已试图通过利用手移动作为输入来对人机输入任务的效率低下进行改进。在检测发生在诸如键盘等计算机使用期间的常见手位置的情况下,该利用将会是最有效的。许多现有的基于视觉的计算机技术将对定点或伸出的手指的使用用作输入手势。由于定点手势与打字期间的自然手定位类似,难以检测位于或在键盘位置附近的手势。

[0004] 大多数现有计算机视觉技术利用手势检测和跟踪范例来传感手势和移动。这些检测和跟踪范例是复杂的,其使用复杂的模式识别技术来恢复手的形状和位置。检测和跟踪受到若干因素的限制,包括难以实现合理的计算复杂性,由于人类手移动和手势中的歧义而导致的实际检测中的问题,以及缺乏对允许多于一个用户交互的技术的支持。

### [0005] 概述

[0006] 提供本概述以便介绍将在以下详细描述中进一步描述的用于视觉显示器的虚拟控制器的简化概念。本概述并不旨在标识所要求保护的主题的必要特征,也不旨在用于帮助确定所要求保护的主题的范围。

[0007] 在用于视觉显示器的虚拟控制器的一个实现中,照相机或其他传感器检测相对于背景的一个或多个手的图像。该图像被分割成手区域和背景区域,并且在各间隔处对在该图像中通过拇指和一手指做出封闭的环形成的不同的、独立的背景区域-“洞”进行计数(例如,每个手可创建一个洞)。拇指和食指在以此方式使用时被称作“拇指和食指接口”(TAFFI)。其他类型的手和手指接口也是有可能的。至少一个控制参数然后被分配到每一个经识别的洞,或捕捉到的图像中的独立背景区域,该控制参数通常允许用户的手操纵屏幕或监视器上所显示的图像的某一方面。例如,鼠标点击功能可以在手的拇指和食指互相接触以创建在视觉上独立的背景区域时被分配为控制参数。可以分配控制参数以使得显示的图像与同该控制参数相关联的独立区域的形状和/或位置的每一改变相关地,或与正在形成或展开(拇指和食指接触时的高状态以及拇指和食指打开时的低状态)的独立区域相关地改变。

### [0008] 附图简述

[0009] 在各附图中,使用相同的标号来指示相同的特征和组件:

[0010] 图 1 是其中可实现用于视觉显示器的示例性虚拟控制器的示例性基于计算机的系统的图示。

[0011] 图 2 是示例性虚拟控制器系统的框图。

[0012] 图 3 是图 2 的虚拟控制器系统的示例性分割器中所使用的图像分割的图示。

[0013] 图 4 是示例性拇指和食指接口控制的图示。

[0014] 图 5 是用手和手指姿势来控制视觉显示器的示例性方法的流程图。

[0015] 详细描述

[0016] 概览

[0017] 本发明描述了用于视觉显示器的虚拟控制器。在一个实现中,示例性系统通过利用应用于手和手指姿势的基于视觉的计算机技术来提供诸如计算机监视器所特有的视觉用户界面等显示器的导航。在一个实现中,用户在键盘上打字并且然后例如,通过暂停键盘键入并仅将一个手的拇指和食指碰在一起(犹如拿着一小指示笔)来调用“拇指和食指接口”即“TAFFI”。该示例性系统传感该事件并将控制参数分配给由该手指姿势形成的独立背景区域的各属性以便控制虚拟显示器上的图像。

[0018] “虚拟”或“虚拟控制器”指的是装置与用户的手缺少物理接触。因此,在一个实现中,虚拟控制器由定位在手和键盘上方的照相机以及从用户的手的视觉图像中导出一个或多个接口的相关联的逻辑组成。分割将手对象与背景(例如,包括键盘)分离开。如果用户将食指与拇指接触(上文中的 TAFFI),则该系统识别由该手势创建的独立背景区域并将其制表。即,该系统识别背景的一片段已通过拇指和食指接触以形成封入背景区域的椭圆形“环形洞”的完全封闭的“环”而在视觉上与主背景的剩余部分隔离开。通过除了计算机照相机之外的装置来检测视觉图像也是有可能的。例如,嵌入在键盘或桌子中的 2D 电极或天线阵列可使用静电或射频技术来对手势“成像”并且以类似于从照相机捕捉图像的方式来处理。

[0019] 在一个实现中,独立背景区域在其在视觉上由手区域或在一个变型中由图像和/或图像边界中的手区域来与背景的其它部分断开或隔离开时被认为是不同的视觉对象。当手和手指的图像是用于确定独立背景区域的边界的定界实体时,于是在拇指和食指“闭合”(互相接触)时创建的拇指和食指之间的椭圆区域就在大约拇指和食指接触的时刻被计作一新的独立背景区域。该新的独立背景区域可被认为是连接分量分析技术中的“连接分量”。这些连接分量或新的独立背景区域-“洞”-此处将被称为“独立背景区域”或仅为“独立区域”。应当理解,该术语指的是例如,连接分量分析技术中的、被认为是不同的视觉对象。

[0020] 当拇指和食指“打开”时,新形成的独立背景区域消失并再次变成较大的独立背景区域的一部分。

[0021] 就连接分量分析技术而言,连接分量是二值图像中具有类似属性的、由于属性相似性而分组在一起的一组像素。每一个连接分量通常对应于如由人类观察者观察到的不同的视觉对象。通过图像的手或手指区域的一部分而在视觉上独立于背景的其他部分的该背景的每一部分都可被定义为独立区域,或者在连接分量分析的语言中,被定义为与背景连接分量不同的新形成的连接分量。

[0022] 当然,其他实现可以使用手的其他手指的移动或接触来形成“洞”或“独立区域”。因此,“TAFFI”应被宽泛地解释为是指在视觉上将背景的一部分与一般背景的剩余部分隔离开的手指和手的配置。例如,人手的拇指和任何其他手指,或仅仅不是拇指的两个手指也

可形成“TAFFI”接口。然而,为了简化描述,各实现通常将按照“拇指和食指”来描述。

[0023] 一旦检测模块将新的独立背景区域与一般背景区域区分开来,系统就将该新识别的独立区域与使得用户能够操纵视觉用户界面上所显示的图像的一个或多个控制参数相关联。视觉用户界面上所显示的图像可以在跟踪独立背景区域的位置、形状、以及甚至存在时经由控制参数来改变。

[0024] 在一个实现中,一示例性系统能够检测多于一个独立区域,从而允许用户通过其中一个或两个手可参与的多个控制参数来控制显示的图像。多个控制参数与多个独立区域的关联使得对显示的图像的控制与每一检测到的独立区域的形状、位置、和存在的改变相关。因此,对显示的图像的操纵可包括对点击、选择、执行、水平移动、垂直移动、卷屏、拖拽、旋转移动、缩放、最大化和最小化、文件功能、菜单部署和使用等的控制。此外,控制参数还可被分配给多个识别出的独立区域之间的关系。即,例如,在两个独立区域相对于彼此移动时,各种控制参数可被附加到它们之间的距离。例如,当每个手的独立区域远离彼此移动时,图像可缩放或拉伸,或可以按其中独立区域之间的距离改变的维度或向量拉伸。

[0025] 虽然所描述的用于虚拟控制器的系统和方法的特征和概念可以在许多不同的环境中实现,但是虚拟控制器的实现是在以下示例性系统和环境的上下文中描述的。

[0026] 示例性环境

[0027] 图 1 示出了其中可实现诸如以上所介绍的拇指和食指接口,即 TAFFI 等虚拟控制器接口技术的示例性系统 100。示例性系统 100 包括视觉用户界面(监视器、屏幕或“显示器”103)上的“显示图像”102、与计算设备 105 耦合的照相机 104、鼠标 106、键盘 108、在上下文中示出的用户的手 110(当然不是该系统的硬件的一部分)、以及由正用作 TAFFI 的用户的手 110(1)形成的视觉上的独立区域 112。照相机获得将由示例性 TAFFI 引擎 115 使用的捕捉到的手的图像 114。(捕捉到的图像 114 仅出于描述的目的而示出,示例性系统 100 不必显示该照相机捕捉到了什么。)主存 TAFFI 引擎 115 的计算设备 105 可以是台式机、膝上型计算机、PDA、或能够成功地结合来自照相机 104 的输入以使得 TAFFI 引擎 115 能够检测特定手势并将其用作用户界面输入的其他计算设备 105。

[0028] 照相机 104 捕捉一个手 110(1) 构成 TAFFI 而另一个手 110(2) 仍然处于“常规”(非 TAFFI)打字位置的图像。捕捉到的图像 114 展示检测到对于形成 TAFFI 的手 110(1) 的独立区域 112,但是没有检测到对于仍然在打字或使用鼠标以获得附加输入条目的手 110(2) 的独立区域。由照相机 104 对独立区域 112 的检测被显示为捕捉到的图像 114 中的暗区(112)。该捕捉到的图像 114 展示了将在以下进一步描述的过程中的一个阶段,其中示例性系统 100 将手 110 和背景分隔成连续的、分割的区域,诸如大背景区域、手区域、以及构成由手 110(1) 的 TAFFI 形成的独立区域 112 的较小的背景区域。

[0029] 系统 100 可以是提供通过由照相机 104 或其他传感器检测到的手势输入来控制视觉用户界面的基于视觉的(“计算机视觉”)系统。换言之,示例性系统 100 可控制可以在包括基于 web 的显示器的计算设备上操作的许多不同类型的程序或应用的视觉用户界面显示输出。由此,示例性系统 100 可替换常规用户输入设备,诸如鼠标 106,且如果需要的话,键盘 108,包括它们的选择、移动、和改变视觉用户界面 102 中所显示的对象,或甚至输入文本的功能。

[0030] 虚拟控制器将特定手势和移动检测为用户输入。在所示实施例中,用于检测的照

相机 104 被放置在附连到显示器 103 的手和键盘上方的某处。放置在该位置的照相机 104 处理覆盖键盘 108 的至少大部分的视场并且大致聚焦于处于正常打字位置的用户的手 110 的平面。在一个实现中,可放置诸如红外或可见 LED 等灯以照亮手 110 和键盘 108,并且也可定位这些灯以缓解改变环境照明的影响。在某些情况下,环境光线可能是充足的,以使得照相机无需额外的灯来获取图像。在各变型中,照相机 104 和 / 或额外的灯可被放置在键盘 108 的各键之间,以使得照相机 104 面朝上并能够检测键盘 108 上方的手势和手的移动。

[0031] 可以在所示示例性系统 100 中使用的照相机 104 的示例是以 30Hz 的速率获取全分辨率灰度图像的罗技网络摄像头 104(加利福尼亚州弗里蒙特市)。照相机 104 可被附到或者键盘 108 或者显示器 103 或者任何合适的地方。

[0032] 在示例性系统 100 中,用户的手 110(1) 可形成 TAFFI,其在拇指和食指接触时创建独立于背景区域的剩余部分的视觉区域。在一个实现中,潜在的 TAFFI 以及一个或多个独立区域 112 的存在或缺失由实时图像处理例程来检测,该例程在计算设备 105 中执行以持续地监视并确定两个手 110 的状态,例如,手 110 是在打字还是在形成用于输入的姿势。该处理例程可首先确定用户的拇指和食指是否接触。如果手指接触而导致识别出 TAFFI 形成的独立区域 112,则可二维地跟踪该接触的位置。例如,拇指和食指接触的位置可以在计算机 105 中被注册为定点箭头的位置或光标位置。由此在一个实现中,可以使用该 TAFFI 形成位置及其相关联的独立区域 112 的识别来建立光标位置并且控制显示的图像。

[0033] 其中在一时间间隔内形成、展开并然后再次形成独立区域 112 的产生独立区域 112 的快速手移动可模拟或模仿鼠标的“点击”并允许用户选择正在显示的项目。独立区域 112 的快速形成、展开、并再次形成还可使得用户能够拖拽或滚动所显示项目的所选部分,以水平、垂直或对角方向移动对象,对显示的图像 102 进行旋转,缩放等。此外,在一个实现中,移动形成更靠近或更远离照相机 104 的独立区域 112 的 TAFFI 可产生对显示的图像的放大和缩小。

[0034] 经由多个 TAFFI 来控制显示的图像可涉及多于一个手 110。图 1 所示的示例性系统 100 是 TAFFI 控制的一实施例,其中图像操纵从一个手 110(1) 的 TAFFI 继续而另一个手 110(2) 打字并在键盘 108 处执行其他输入任务。但在 TAFFI 控制的另一实施例中,两个手 110 都可形成各自的 TAFFI,从而导致照相机 104 检测到至少两个独立区域 112。双手 TAFFI 控制可提供用于视觉用户界面的精细调整的导航的输入控制。该双手方法提供了除了放大、缩小和旋转移动之外的多方向图像操纵,其中该操纵因为多个 TAFFI 的独立区域 112 相对于彼此的交互而变得更复杂。

[0035] 示例性系统

[0036] 图 2 示出了示例性虚拟控制器系统 100 的各种组件。虚拟控制器系统 100 的所示配置仅仅是一个示例安排。所示组件或其他类似组件的许多安排在本主题的范围之内也是可能的。示例性虚拟控制器系统 100 具有一些组件,诸如可以用硬件、软件、或硬件、软件、固件等的组合来执行的 TAFFI 引擎 115 等。

[0037] 示例性系统 100 包括硬件 202,如照相机 104 或其他图像传感器、键盘 108 和显示器 103。TAFFI 引擎 115 包括其他组件,如图像分割器 204、独立区域跟踪器 206、包括链接模块 210 的控制参数引擎 208。

[0038] 在一个实现中,照相机 104 检测被解释为相对于背景的一个或多个手 110 的图像。

捕捉到的图像 114 的像素包括将用于将图像中的手 110 与该图像中的背景区分开来的对比度属性。对于与背景形成对比的手的适用的属性可包括亮度、灰度、色彩分量强度、色彩平面值、向量像素值、色彩图索引值等。在各变型中,照相机 104 可利用这些属性中的任何一个来将手像素与背景像素区分开来,例如,取决于是否使用红外照明来代替典型的可见光谱。有时,使用红外线来获得捕捉到的图像 114 导致不同肤色的大多数人的手以与背景类似的对比度出现,而不管由于种族、日晒等的不同而导致的在可见光谱中的皮肤颜色和色调的变化。因此,相对于图像中的背景的手的检测在不管可见的肤色的情况下在红外线中可以是容易实现的。

[0039] 分割器 204 由此例如通过根据上述对比度或亮度属性的二值图像分割来将捕捉到的图像 114 分隔成一个或多个手区域 110 和背景区域。该二值图像分割将背景区域像素与呈现在捕捉到的图像 114 中的任何其他(前景)对象或区域的像素区分开来。在一个实现中,分割器 204 通过首先确定对应于背景区域的像素来分隔图像。背景区域像素各自被分配到一个值,比如二进制值“一”(1)。捕捉到的图像 114 中的剩余像素各自被分配到一个不同的值,比如“零”(0)。

[0040] 图 3 示出了由分割器 204 执行的二值图像分割的示例 300。捕捉到的图像 114 包括背景对象 302 和前景中的手对象 304。存在用于产生分割的图像的各种技术,其中大多数技术在本领域内是公知的。在一个实现中,分割器 204 将背景区域像素与呈现在捕捉到的图像 114 或示例 300 中的任何其他对象或区域的像素辨别开来。区分二值图像中的像素是通过将对应于背景的每一像素认为是“开”或比如“一”的特定值来实现的。然后可将图像中的每一个其他像素与存储的背景图像的值进行比较。显著地亮于相应的背景像素的任何其他像素值都被认为是新的区域或图像对象的一部分,并且被标记为“关”,或被给予一不同的值,比如“零”。

[0041] 示例 300 还可将背景区域 302 与图像的其他区域的区别示为色差。背景区域 302 被示为等同于第一个值的较暗的颜色。被示为较亮颜色的手对象 304 等同于第二个值,从而将其与背景区域 302 区分开来。

[0042] 回到图 2,独立区域跟踪器 206 以固定的时间间隔确定背景的多个独立区域 112。通过至少一部分非背景手区域(或图像边界)在视觉上独立于背景的其他部分的该背景的每一部分都被定义为独立区域 112。对于传感的每一个独立区域 112,独立区域跟踪器 206 找到完全被“0”像素包围(即,不再连续地连接到构成主背景的“1”像素的剩余部分)的“1”像素区域。换言之,独立区域跟踪器 206 找到由 TAFFI 的接触的拇指和食指姿势来划定边界的隔离的背景区域。

[0043] 例如,当独立区域 112 完全位于由照相机 104 传感的捕捉到的图像 114 之内,即,在独立区域 112 未有任何一部分位于捕捉到的图像 114 的边界上时,可以确保准确地将该独立区域检测为指示用户选择显示器 103 上的对象的意图的单独的背景区域。

[0044] 然而,在一个实现中,独立区域跟踪器 206 的变型即使在独立区域 112 的一部分在“屏幕之外”-未作为捕捉到的图像 114 的一部分来包括时也可传感独立区域 112。这可通过将独立区域 112 定义为由手 110 的一部分或捕捉到的图像 114 的边界的一部分从主背景中剪切下来的背景区域。但这仅仅是如何划定独立背景区域的界限的变型。

[0045] 一旦确立了一个或多个独立区域的存在,链接模块 210 就将用于操纵用户界面上



的视觉图像显示 102 的控制参数与每一个计数的独立区域相关联。操纵可包括多个机制，包括视觉用户界面中的光标控制。视觉图像显示 102 的光标控制可以，且仅当检测到独立区域并将其与控制参数相关联时实现。如果对独立区域的检测停止，则控制参数关联停止，并且禁用光标控制和操纵。光标控制可包括多个操纵，包括模仿来自鼠标的输入的“点击”动作。该点击动作允许选择视觉图像显示 102 的所需部分，跟踪和拖拽，以及多方向移动和光标控制。

[0046] 链接模块 210 能够将特定控制参数与手或手指姿势或与手势改变相关联。一旦控制参数被分配或与手或手指姿势相关联，于是控制参数引擎 208 就可进一步细察 (nuance) 该手势与该控制参数是如何彼此相关的。例如，仅仅是拇指与食指的接触就可用作“开-关”、二值、高-低、或其他二状态接口或开关。然而可连续地改变的手势属性可被分配来提供对诸如显示图像 102 在连续区上的逐步移动等显示图像操纵的可变控制。

[0047] 当链接模块 210 将可变控制参数分配给例如与相应的独立区域的形状或位置的改变相关的显示图像 102 的控制时，该可变性方面可通过计算属于每一个独立区域的所有像素的平均位置并且然后跟踪在一个手形成 TAFFI 时创建的形状的位置的改变来实现。手的移动更改椭圆形的独立区域的方向并导致与所分配的控制参数相关联的显示属性的相应改变。

[0048] 对显示的图像的控制

[0049] 图 4 示出了在捕捉到的图像 114 的上下文中示出的示例 TAFFI 400。捕捉到的图像 114 的所示部分包括背景区域 302、手对象区域 110、独立区域 112 和图像边界 408。区域 302、110 和 406 中的每一个都可被描述为不同的连接区域或连接分量。TAFFI 引擎 115 将独立区域 112 与其他连接分量 302 和 110 区分开来。

[0050] TAFFI 引擎 115 由此可将图像的连接分量的计算用作实现用于视觉显示器的虚拟控制器的基础。更详细而言，连接分量是图像中的一像素子集或区域，其中每一像素都“连接”至该子集中的每一个其他像素。术语“连接”表示这样一个像素集：对于该像素集有可能通过遍历属于该集合的像素从任何其他像素到达每一个像素。当前存在用于计算图像中的一组连接分量的高效的技术。连接分量技术对于确定图像中的形状的特性而言可以是高效的途径，因为它们允许检查由整个图像的像素中的许多像素组成的小分量集合。

[0051] 计算连接分量的过程可引起对无关连接分量的检测。这些不需要的检测可能混淆由 TAFFI 或其他示例性接口形成的相关独立区域的判定，并且因此阻碍虚拟控制器的实现。在一个实现中，额外的连接分量的无关检测可通过丢弃具有的像素数量比预定可允许阈值少的连接分量来克服。

[0052] 在一个实现中，TAFFI 引擎 115 验证一经识别的独立区域 112 完全位于图像的边界内，即，完全在背景区域 302 的范围之内。有时，对大小足够并且不包括图像的边界 408 上的像素的独立区域 112 的该有限的检测加强了对所需独立区域 406 的可靠的标识。在这一个实现中，适当的检测通过避免假连接分量候选独立区域，或不是完全位于图像内以及包含该图像的边界 408 上的部分的独立区域来实现。

[0053] 但是，在另一实现中，TAFFI 引擎 115 通过检测独立区域 112 位于捕捉到的图像 114 中的一部分以及位于越过该图像的边界 408 的屏幕之外的一部分来检测独立区域 112。在该实现中，只要独立区域 112 连续到遇到和 / 或超过图像的边界 408 的点，连接分量分析

就继续。这可能在形成 TAFFI 的手和独立区域 112 只是部分地位于照相机的视场内,并因此只是部分地位于检测到的图像内时发生。

[0054] 在一个实现中,TAFFI 引擎 115 使用独立区域 112 的中心来建立显示的图像 102 中的光标位置和光标控制。TAFFI 引擎 115 可以对每一个识别出的独立区域 112 执行静态分析,其中独立区域跟踪器 206 计算属于每一个独立区域 112 的所有像素的“形心”或平均像素位置。该计算出的位置是对许多像素位置的概括,导致对于该实现的稳定性和精度。平均像素位置可以在与计算连接分量相同的阶段计算,从而产生以低处理成本提供快速结果的高效技术。

[0055] 关于作为控制视觉显示器的手段独立区域 406 的出现和消失,在一个实现中属于独立区域 112 的所有像素的平均像素位置仅在独立区域 112 是在重复的检测过程的一个间隔期间新检测到的时候才建立光标位置和控制。

[0056] 用对独立区域 406 的检测进行的光标控制可模仿鼠标输入设备 106。类似于鼠标 106,用于光标操纵的相对运动可以从由 TAFFI 400 形成的检测到的独立区域 112 的当前和过去的位置中计算出。拇指和食指联接在一起是允许如同鼠标输入设备一样的不费力的离合 (clutching) 行为的自然运动。将卡尔曼过滤器与 TAFFI 检测一起使用可平滑光标在视觉显示器 103 上的运动。

[0057] 示例性 TAFFI 引擎 115 支持通过在阈值时间间隔内快速形成、展开、并重新形成独立区域 112 来选择显示的图像 102 的对象。这些动作模仿用于“选择”或“执行”功能的鼠标按钮的“点击”,并且还支持从跟踪到拖拽所选对象的转换。例如,拖拽可通过紧接着独立区域 112 的最近一次形成模仿“鼠标按下”事件来实现。相应的“鼠标松开”事件在独立区域 112 通过打开拇指和食指消失时生成。例如,在独立区域形成的时刻,可以选择诸如视觉用户界面显示上的文档中的滚动条等对象。紧接着该选择,可以按与可移动鼠标 106 以便在文档中向下滚动相同的方式移动形成独立区域 112 的手的位置。

[0058] TAFFI 引擎 115 能够提供比只模仿常规的基于鼠标的功能更多的对视觉显示 102 的控制。独立区域 112 (连接分量) 的各像素位置的均值和协方差可通过计算像素位置的协方差矩阵的本征向量来与独立区域 112 的形状的定向椭圆模型相关。本征值大小的平方根给出其长轴和短轴大小的空间范围,而椭圆的方向被确定为多至 180 度歧义的本征向量中的一个的反正切。所产生的歧义可通过采用计算出的方向或 +180 度旋转的方向来最小化与前一帧的方向差别来解决。

[0059] TAFFI 引擎 115 可以从由示例性 TAFFI 400 创建的独立区域 112 的椭圆模型中计算出位置、方向和缩放的同时改变。在各种实现中,缩放的改变也可用于检测手朝照相机或远离照相机的移动。这假设形成独立区域 112 的用户的手一般保持在与照相机 104 的固定距离范围内以使得独立区域 112 的大小和形状仅在容差范围内变化,使得方向上的视觉改变对于背景区域 302 或键盘的平面在某种程度上是有限的。在一个实现中,一重要的考虑是,在整个交互过程中,用户必须在该用户将手相对于照相机或键盘移上移下时维持独立区域的大小 - 由 TAFFI 400 形成的椭圆形洞的大小 (即,在某些实现中,高度的改变与独立区域的形状的实际改变相混淆)。在其他实现中,TAFFI 引擎 115 使用计算机视觉逻辑来在手移上移下时补偿独立区域大小的改变。

[0060] 在一个示例性实现中,TAFFI 引擎 115 使用独立区域 112 的椭圆模型来单手导航诸

如由WINDOWS® LIVE VIRTUAL EARTH® web 服务或其他类似的因特网地图服务（华盛顿州雷蒙德市）提供的鸟瞰和卫星影像。通过在虚拟地图的整个视图中移动而进行的导航可由具有在诸如桌子或键盘等背景区域 302 中移动的独立区域 112 的 TAFFI 400 来实现。整张地图的旋转可通过在键盘的 2 维平面中旋转形成该独立区域 112 的手来实现，而放大和缩小功能可通过将手移近或远离照相机 104 来实现。

[0061] TAFFI 引擎 115 可实现两个或更多手对于光标控制和导航的使用。逐帧对应策略允许每一个独立区域 112 被连续地跟踪为被照相机检测到以供输入的第一、第二、第三等区域中的任一个。两个手为了被照相机检测而相对于背景区域 302 的放置，以及手相对于背景区域 302 的随后的移动更改独立区域 406 的椭圆模型的方向，并且导致与经由链接模块 210 所分配的控制参数的手部移动的定位和位置相关联的视觉用户界面显示的移动。

[0062] 对与多个手或手指姿势相对应的多个控制参数的同时跟踪允许各种双手交互。再次参考因特网虚拟地图示例，用于导航该虚拟地图的双手输入允许同时改变显示器 103 上的地图视图的旋转、转换和缩放。因为独立区域 406 的位置估算是从手的位置中导出的，所以该双手技术能够提供比单手技术所能提供的更稳定的运动估算。该双手技术因此允许：顺时针和逆时针旋转，其中两个手同时在旋转方向上移动；在垂直或水平方向上移动整个视觉用户界面显示视图，其中两个手在所需方向上移动；以及缩放功能，其中放大视觉用户界面显示在两个手开始靠在一起并且稍后伸展远离另一个手时实现，而缩小视觉用户界面显示通过将手从分开的起始位置拉拢在一起来执行。

[0063] 从独立区域 112 的椭圆模型中计算出的位置、方向和缩放的同时改变可以在除了标准计算设备环境之外的实现中使用。例如，TAFI 引擎 115 可控制包括桌子上的照相机和投影仪，但没有诸如鼠标、触摸屏或键盘等附加输入设备的交互式桌面系统。用户将手放在桌面上方，从而形成独立区域 406 以便提供操纵以及与该桌面和该表面上所显示的材料交互。类似的实现可包括将显示图像投影到墙上的系统，其中用户可通过担当 TAFI 400 的手和手指来控制该显示图像并与其交互。例如，TAFI 引擎 115 可允许该用户在投影仪演示期间更改幻灯片。

[0064] 示例性方法

[0065] 图 5 示出了经由手或手指姿势来控制视觉显示器的示例性方法 500。在该流程图中，操作被概括成单独的框。取决于实现，示例性方法 500 可以由硬件、软件、或硬件、软件、固件等的组合来执行，例如，由示例性虚拟控制器系统 100 的组件和 / 或示例性 TAFI 引擎 115 来执行。

[0066] 在框 502 处，经由照相机 104 捕捉相对于背景的一个或多个手 110 的图像。对比度、色彩或亮度可以是使得能够在手和周围背景区域之间进行区分的像素属性。手相对于对比背景可更容易地传感。传感手的一个场景是在键盘 108 处打字时。照相机 104 捕捉手 110 和传感为背景区域的一部分的键盘 108 的图像。红外 LED 照明也可用于该方法，其提供使得大多数手的肤色对于照相机 104 看似相似的受控光照。

[0067] 在框 504 处，通过二值分割将图像分割成手对象和背景区域。例如，从图像中的任何其他对象或区域的像素中标识并区分出背景区域像素。背景区域像素然后用一值来标记。随后标识图像中的其他对象或区域的像素并将其与存储的背景图像的像素值进行比较。显著地亮于相应的背景像素值的任何像素值都被标记为新的区域或图像的一部分，并

被给予不同于背景区域像素的值。图像的不同区域的这一区分和标记是该图像的二值分割。

[0068] 在框 506 处,以重复的检测间隔对背景的多个独立区域进行计数。独立区域 406 被定义为通过手对象 110 中的一个的至少一部分在视觉上独立于背景的其他部分的背景 302 的每一部分。例如,当一个手担当拇指和食指接口,即 TAFFI 时,这个手的拇指和食指创建独立于一般背景区域的剩余部分的封闭区域。该封闭区域形成可向其附加用于操纵视觉显示器的控制参数的新的独立区域 112。在一个实现中,该方法测试所检测到的独立区域是否真正独立,即,在一种情况下,独立区域是否具有图像边界上的像素。

[0069] 在框 508 处,将用于操纵显示器上的图像的控制参数与每一计数的独立区域或其属性相关联。例如,由用作 TAFFI 的手创建的独立区域 112 由照相机 104 来传感并且与使得用户能够选择用户界面显示上的对象的控制参数相关。随后,第二传感的独立区域 112 与使得用户能够将先前所选对象移至该用户界面显示上的不同位置的用户界面控制参数相关。该第一和第二独立区域 112 的快速连续传感可能由快速形成、展开和重新形成独立区域 406 产生,从而导致与传感的独立区域 112 相关联的类似鼠标的“点击”功能。

[0070] 在框 510 处,显示的图像经由控制参数与被分配给该控制参数的独立区域的属性的每一改变相关地改变。例如,独立区域 112 的位置可相对于传感照相机 104 向左或向右移动,并且显示的图像 102 可跟着移动。传感的独立区域 112 与控制参数的关联允许根据正用作 TAFFI 的手的移动、位置和关系来操纵显示的视觉图像 102。

[0071] 上述方法 500 以及其他相关方法可以在计算机可执行指令的一般上下文中实现。一般而言,计算机可执行指令可包括执行特定功能或实现特定的抽象数据类型的例程、程序、对象、组件、数据结构、过程、模块、功能等。这些方法还能在其中功能由通过通信网络链接的远程处理设备完成的分布式计算环境中实现。在分布式计算环境中,计算机可执行指令可以位于包括存储器存储设备在内的本地和远程计算机存储介质中。

[0072] 结论

[0073] 尽管已用对结构特征和 / 或方法动作专用的语言描述了示例性系统和方法,但可以理解,所附权利要求书中定义的主题不必限于所述具体特征或动作。相反,上述具体特征和动作是作为实现所要求保护的方法、设备、系统等的示例性形式而公开的。

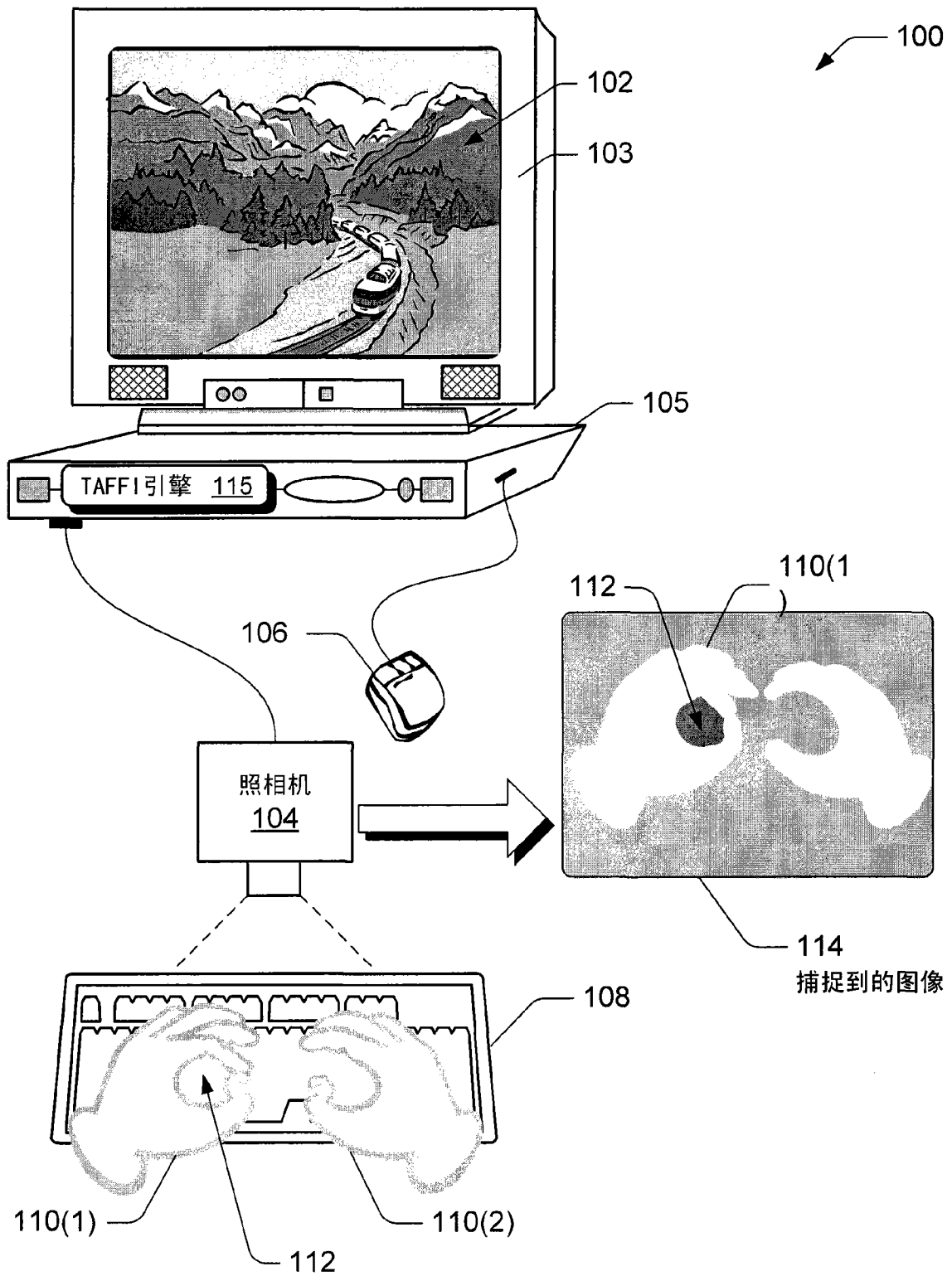


图 1

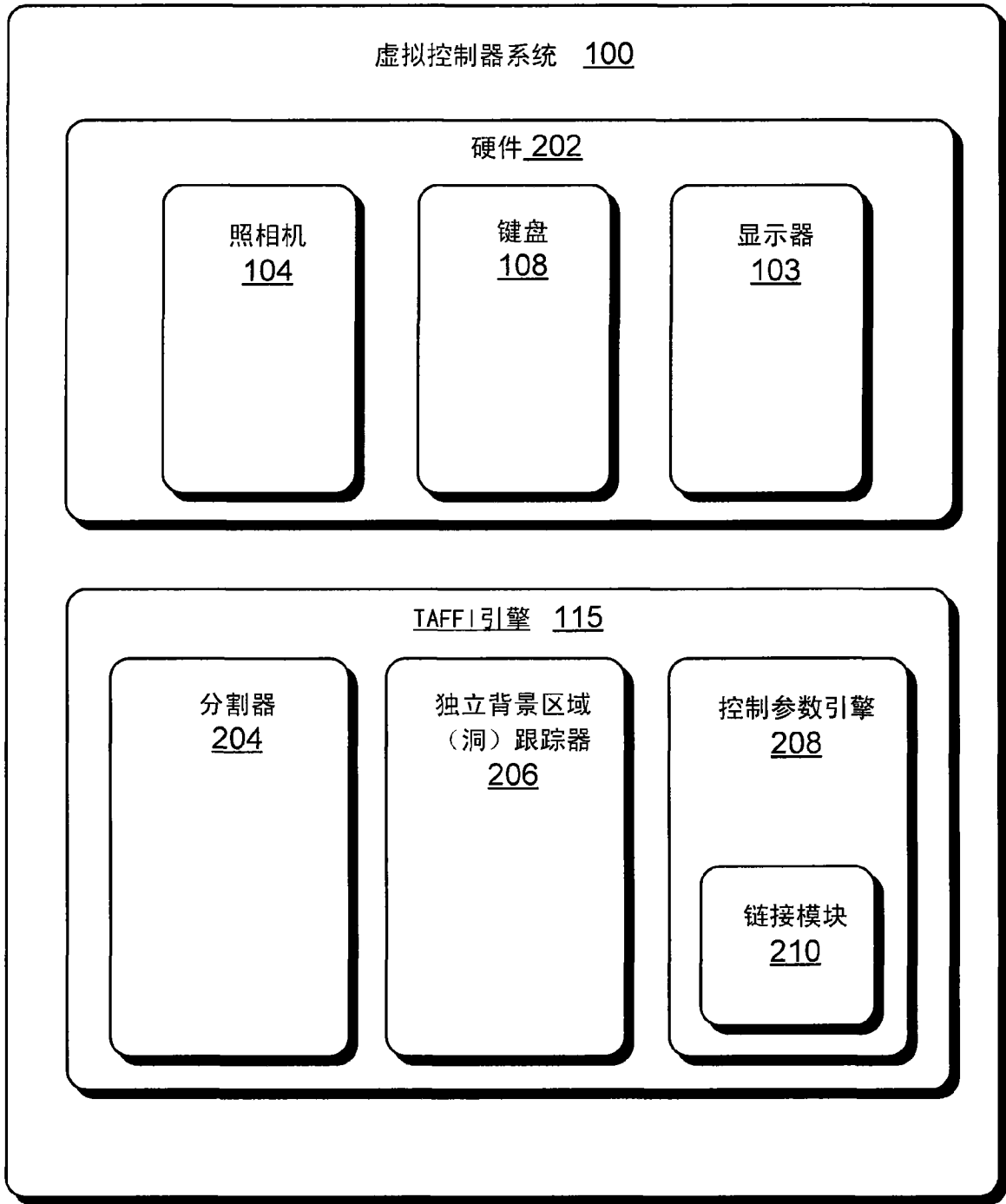


图 2

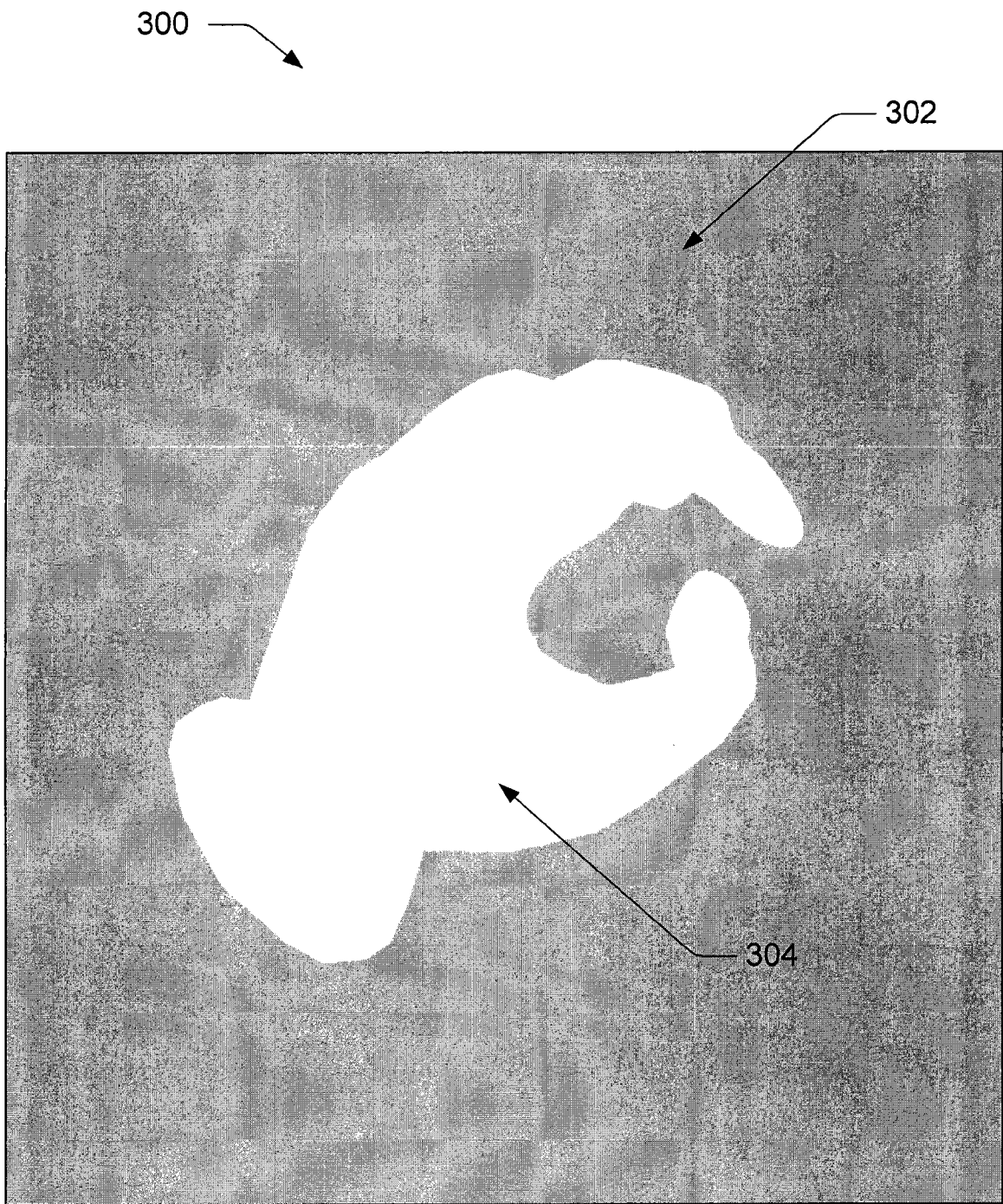


图 3

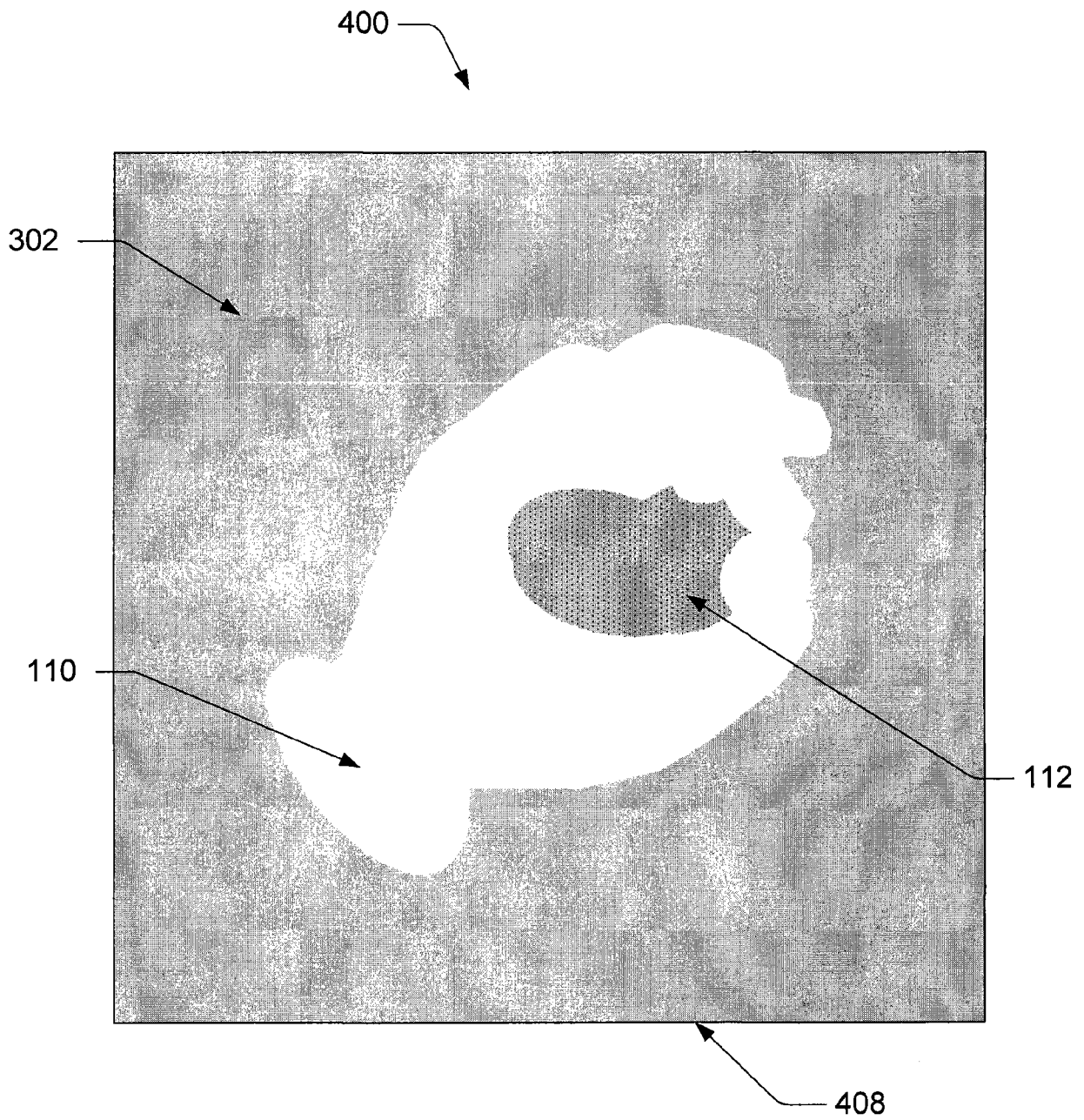


图 4



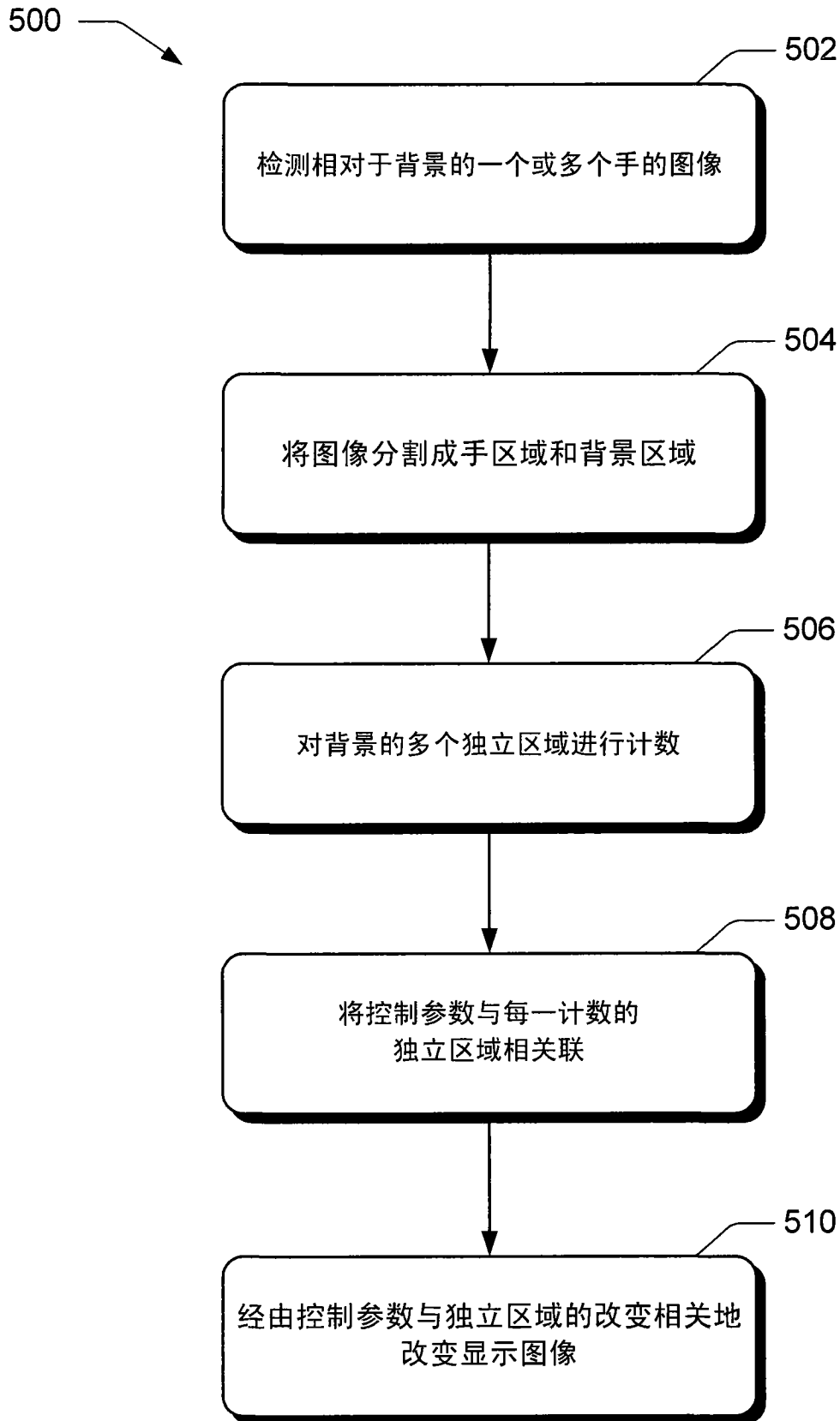


图 5