



(12) 发明专利

(10) 授权公告号 CN 101689244 B

(45) 授权公告日 2015. 07. 22

(21) 申请号 200880023087. 7

(22) 申请日 2008. 05. 02

(30) 优先权数据

60/916, 063 2007. 05. 04 US

(85) PCT国际申请进入国家阶段日

2009. 12. 31

(86) PCT国际申请的申请数据

PCT/US2008/062456 2008. 05. 02

(87) PCT国际申请的公布数据

W02008/137708 EN 2008. 11. 13

(73) 专利权人 高通股份有限公司

地址 美国加利福尼亚州

(72) 发明人 埃文·希尔德雷思

(74) 专利代理机构 北京律盟知识产权代理有限

责任公司 11287

代理人 宋献涛

(51) Int. Cl.

G06K 9/00(2006. 01)

(56) 对比文件

US 6204852 B1, 2001. 03. 20,

US 6204852 B1, 2001. 03. 20,

US 5827942 A, 1998. 10. 27,

CN 1694056 A, 2005. 11. 09,

审查员 李劲嫻

权利要求书4页 说明书18页 附图10页

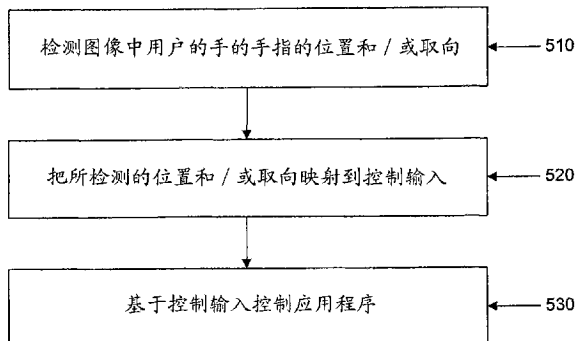
(54) 发明名称

用于紧凑设备的基于相机的用户输入

(57) 摘要

使用相机来检测诸如用户手指的对象的位置和 / 或取向作为用于提供用户输入的方法, 例如用于滚过数据、控制光标位置、以及基于用户手指的位置提供输入以控制视频游戏。可以把输入提供给手持设备, 包括例如, 蜂窝电话、视频游戏系统、便携式音乐 (MP3) 播放器、便携式视频播放器、个人数据助理 (PDA)、音频 / 视频设备遥控器和消费型数码相机, 或者其他类型的设备。

500



1. 一种用于从图像检测输入的方法,包括:

通过将用户的手的图像确定的一组经链接的段与手指的至少一个解剖模型相比较,来确定所述一组经链接的段是否是所述用户的手的手指或不是手指;和

在确定所述一组经链接的段是所述用户的所述手指的情况下:

使用手指的所述至少一个解剖模型在与电子设备进行交互的所述用户的手的所述图像中检测所述用户的手的所述手指的取向,其中所述解剖模型包括用于比较从所述用户的手的所述图像中所确定的经链接的段的信息;

把所检测的取向映射到控制输入;和

基于所述控制输入通过执行对应于所述控制输入的命令或操作来控制应用程序,

其中使用手指的所述至少一个解剖模型在与所述电子设备进行交互的所述用户的手的所述图像中检测所述用户的手的所述手指的取向包括:

识别所述图像的对应于手指皮肤中褶皱的区域,和

当在所述图像中检测手指时忽略所识别的所述图像的对应于所述手指皮肤中所述褶皱的区域。

2. 如权利要求 1 所述的方法,还包括:在确定所述一组经链接的段是所述用户的所述手指的情况下,在与所述电子设备进行交互的用户的手的图像中检测用户的手的手指的位置,其中,基于所检测的位置映射所述控制输入。

3. 如权利要求 2 所述的方法,其中:

所述控制输入定义了虚拟环境中对应于手指的所检测的位置和所检测的取向的位置和取向;并且

基于所述控制输入控制所述应用程序包括:

控制所述虚拟环境中的虚拟对象移动到所述虚拟环境中对应于手指的所检测的位置和所检测的取向的位置和取向;和

在所述虚拟环境中对应于手指的所检测的位置和所检测的取向的位置和取向表现所述虚拟对象的显示。

4. 如权利要求 1 所述的方法,其中,使用手指的解剖模型在与电子设备进行交互的用户的手的图像中检测用户的手的手指的取向包括:识别对应于指尖的手指的第一部分,识别对应于手指根部的手指的第二部分,计算手指的第一部分和手指的第二部分之间的向量,并基于所述向量确定手指的取向。

5. 如权利要求 1 所述的方法,其中,使用手指的解剖模型在与电子设备进行交互的用户的手的图像中检测用户的手的手指的取向包括:

获取手指的图像,和

对所述图像进行边缘处理。

6. 如权利要求 5 所述的方法,还包括:

沿着具有垂直于手指的取向的扫描线分割所述图像;

基于所述边缘处理来链接段;和

基于所链接的段的集合与所述手指的解剖模型的比较,选择所链接的段的集合作为所述手指。

7. 如权利要求 1 所述的方法,还包括检测手指周围的轮廓。

8. 如权利要求 7 所述的方法,其中,通过计算所述图像中图像密度的梯度,并识别所述图像密度的梯度中局部最小和局部最大中的至少一个来检测手指周围的轮廓。

9. 如权利要求 1 所述的方法,其中,使用手指的解剖模型在与电子设备进行交互的用户的手的图像中检测用户的手的手指的取向包括:基于对象的尺寸、对象的形状以及该对象是否和所述图像的边界相交来在所述图像中选择作为手指的对象。

10. 如权利要求 1 所述的方法,其中,使用解剖模型检测手指的取向包括:使用在平行于所述电子设备的显示器的平面中取向的手指的解剖模型来检测手指的取向,所述解剖模型是从手指的根部延伸到手指的指尖的手指模型。

11. 一种用于从图像检测输入的系统,包括:

用于通过将用户的手的图像确定的一组经链接的段与手指的至少一个解剖模型相比较,以确定所述一组经链接的段是否是所述用户的手的手指或不是手指的装置;

用于在确定所述一组经链接的段是所述用户的所述手指的情况下使用手指的所述至少一个解剖模型在与电子设备进行交互的所述用户的手的所述图像中检测所述用户的手的所述手指的取向的装置,其中所述解剖模型包括用于比较从所述用户的手的所述图像中所确定的经链接的段的信息;

用于把所检测的取向映射到控制输入的装置;和

用于基于所述控制输入通过执行对应于所述控制输入的命令或操作来控制应用程序的装置,

其中所述用于使用手指的所述至少一个解剖模型在与所述电子设备进行交互的所述用户的手的所述图像中检测所述用户的手的所述手指的取向的装置包括:

用于识别所述图像的对应于手指皮肤中褶皱的区域的装置,和

用于当在所述图像中检测手指时忽略所识别的所述图像的对应于所述手指皮肤中所述褶皱的区域的装置。

12. 一种用于从图像检测输入的方法,包括:

访问与电子设备进行交互的用户的手的图像;

检测所述图像中的一组经链接的段;

基于对象的尺寸是否是对应于手指的解剖模型的典型尺寸、对象的形状是否是对应于所述手指的解剖模型的拉长形状、所述一组经链接的段的比较是否与所述手指的解剖模型相匹配、以及该对象是否和所述图像的边界相交,在所述图像中识别作为用户的手的手指的取向,所述相交对应于进入所述图像的所述手指;

使用所述手指的解剖模型在与电子设备进行交互的所述用户的手的所述图像中检测所述用户的手的所述手指的取向;

确定在所述图像中被识别为用户的手的手指的取向的位置;

把所检测的位置映射到控制输入;和

基于所述控制输入通过执行对应于所述控制输入的命令或操作来控制应用程序,

其中使用所述手指的所述解剖模型在与所述电子设备进行交互的所述用户的手的所述图像中检测所述用户的手的所述手指的取向包括:

识别所述图像的对应于手指皮肤中褶皱的区域,和

当在所述图像中检测手指时忽略所识别的所述图像的对应于所述手指皮肤中所述褶

皱的区域。

13. 如权利要求 12 所述的方法, 其中, 在所述图像中识别作为用户的手的手指的对象包括: 基于在平行于所述电子设备的显示器的平面中取向的所述手指的解剖模型, 在所述图像中识别作为用户的手的手指的对象, 所述解剖模型是从手指的根部延伸到手指的指尖的手指模型。

14. 如权利要求 12 所述的方法, 其中:

把所检测的位置映射到控制输入包括: 把所检测的位置映射到显示区域中对应于手指的所检测的位置的位置; 和

基于所述控制输入控制应用程序包括: 把所显示的指示符移动到所述显示区域中对应于手指的所检测的位置的位置。

15. 如权利要求 12 所述的方法, 其中:

把所检测的位置映射到控制输入包括: 识别在所显示的用户界面中对应于手指的所检测的位置的位置处的输入控制, 并确定和所识别的输入控制相关联的操作; 和

基于所述控制输入控制应用程序包括: 执行和所识别的输入控制相关联的操作。

16. 如权利要求 12 所述的方法, 其中:

确定在所述图像中被识别为用户的手的手指的对象的位置包括:

确定所述图像中手指的宽度; 和

基于所确定的宽度, 估计手指相对于捕获所述图像的相机的深度位置; 和

基于所估计的深度位置映射所述控制输入。

17. 如权利要求 12 所述的方法, 其中, 确定在所述图像中被识别为用户的手的手指的对象的位置包括: 检测对应于指尖的手指部分, 并确定作为所述对应于指尖的手指部分的质心的位置。

18. 如权利要求 12 所述的方法, 还包括在所述图像中定义图像区域, 以使所述图像区域的边界与所述图像的边界分隔开至少所预期的指尖宽度的一半的距离, 其中, 在所定义的图像区域内检测手指的位置。

19. 如权利要求 18 所述的方法, 其中, 当在所述图像区域外检测到指尖时, 所述指尖的位置被映射到所述图像区域的边界上的位置。

20. 一种用于从图像检测输入的系统, 包括:

用于访问与电子设备进行交互的用户的手的图像的装置;

用于检测所述图像中的一组经链接的段的装置;

用于基于对象的尺寸是否是对应于手指的解剖模型的典型尺寸、对象的形状是否是对应于所述手指的解剖模型的拉长形状、所述一组经链接的段的比较是否与所述手指的解剖模型相匹配、以及该对象是否和所述图像的边界相交, 在所述图像中识别作为用户的手的手指的对象; 所述相交对应于进入所述图像的所述手指;

用于使用所述手指的解剖模型在与所述电子设备进行交互的所述用户的手的所述图像中检测所述用户的手的所述手指的取向的装置;

用于确定在所述图像中被识别为用户的手的手指的对象的位置的装置;

用于把所检测的位置映射到控制输入的装置; 和

用于基于所述控制输入通过执行对应于所述控制输入的命令或操作来控制应用程序

的装置，

其中所述用于使用所述手指的所述解剖模型在与所述设备进行交互的所述用户的手的所述图像中检测所述用户的手的所述手指的取向的装置包括：

用于识别所述图像的对应于手指皮肤中褶皱的区域的装置，和

用于当在所述图像中检测手指时忽略所识别的所述图像的对应于所述手指皮肤中所述褶皱的区域的装置。

用于紧凑设备的基于相机的用户输入

[0001] 相关申请的交叉引用

[0002] 本申请要求 2007 年 5 月 4 日递交的美国临时专利申请 No. 60/916,063 的权益,通过引用将该申请完全结合于此。

技术领域

[0003] 本公开一般涉及用户输入,并且至少一种特定配置涉及利用相机检测用户指尖的位置作为提供用户输入的机制。

背景技术

[0004] 包括蜂窝电话的手持设备可以包括一个或更多用于用户输入的机制,例如小键盘、控制杆、或者滚动轮。一些手持设备使用光学传感器来模拟 (emulate) 控制杆功能,并且一些手持设备包括触敏 (touch-sensitive) 显示屏。用于手持设备的输入机制典型地限于二维的控制。

发明内容

[0005] 本公开一般涉及用户输入,并且至少一个特定配置涉及利用相机检测用户指尖的位置作为提供用户输入的机制的新颖方法。

[0006] 根据一个一般的实施方式,使用相机来检测例如用户的手指的对象的位置和 / 或取向作为用于提供用户输入的方法,例如用于滚过 (scroll through) 数据、控制光标位置和基于用户的手指位置提供输入以控制视频游戏。把输入提供给手持设备,包括例如蜂窝电话、视频游戏系统、便携式音乐 (MP3) 播放器、便携式视频播放器、个人数据助理 (PDA)、音频 / 视频设备遥控器和消费型数码相机,或者其他类型的设备。手持设备可以包括相机,所以为了实施用于提供用户输入的系统,可以不要求额外的硬件。

[0007] 根据另一个一般实施方式,一种方法包括:使用手指的解剖模型在与电子设备进行交互的用户的手的图像中检测用户的手的手指的取向。所述方法还可以包括把所检测的取向映射到控制输入,并基于所述控制输入控制应用程序 (application)。

[0008] 实施方式可用包括一个或更多的下列特征。例如,所述方法还可以包括在与电子设备进行交互的用户的手的图像中检测用户的手的手指的位置,其中,可以基于所检测的位置映射控制输入。控制输入可以在虚拟环境中定义对应于手指的所检测的位置和所检测的取向的位置和取向。所述方法可以包括:控制虚拟环境中的虚拟对象移动到虚拟环境中对应于手指的所检测的位置和所检测的取向的位置和取向;以及,在虚拟环境中对应于手指的所检测的位置和所检测的取向的位置和取向表现虚拟对象的显示。

[0009] 使用手指的解剖模型在与电子设备进行交互的用户的手的图像中检测用户的手的手指的取向可以包括:识别对应于指尖的手指的第一部分,识别对应于手指根部的手指的第二部分,计算手指的第一部分和手指的第二部分之间的向量,以及,基于所述向量确定手指的取向。所述方法还可以包括获取手指的图像,并对所述图像进行边缘处理。此外,所

述方法可以包括沿着具有基本垂直于手指的取向的扫描线分割所述图像。

[0010] 在某些实施方式中,所述方法还可以包括检测手指周围的轮廓。可以通过计算所述图像内的图像密度的梯度,并识别所述图像密度的梯度中的局部最小和局部最大中的至少一个来检测手指周围的轮廓。所述方法可以包括识别所述图像对应于手指皮肤中的褶皱的区域,并且在所述图像内检测手指时忽略所述图像对应于手指皮肤中的所述褶皱的被识别区域。

[0011] 在某些例子中,所述方法可以包括在所述图像中基于对象的尺寸、对象的形状以及该对象是否和所述图像的边界相交选择作为手指的对象。所述方法还可以包括使用在平行于所述电子设备的显示器的平面中取向的手指的解剖模型检测手指的取向,所述解剖模型是从手指的根部延伸到手指的指尖的手指模型。

[0012] 根据另一个一般实施方式,一种设备包括:图像检测器,被配置成使用手指的解剖模型在与电子设备进行交互的用户的手的图像中检测所述用户的手的手指的取向。所述设备还可以包括处理器,该处理器被配置成把所检测的取向映射到控制输入,并基于所述控制输入控制应用程序。所述图像检测器可以被配置在移动电话、便携式视频游戏系统、便携式音乐播放器、个人数据助理、遥控器或者数码相机中。

[0013] 根据另一个一般实施方式,一种存储计算机程序产品的计算机可读介质。所述计算机程序产品包括指令,所述指令当被执行时工作以使计算机执行操作。所述操作包括使用手指的解剖模型在与电子设备进行交互的用户的手的图像中检测所述用户的手的手指的取向。所述操作还可以包括把所检测的取向映射到控制输入,并基于所述控制输入控制应用程序。

[0014] 根据另一个一般实施方式,一种方法包括:访问与电子设备进行交互的用户的手的图像,并且基于对象的尺寸、对象的形状以及该对象是否和所述图像的边界相交,在所述图像中识别作为所述用户的手的手指的对象。所述方法还可以包括确定在所述图像中被识别为所述用户的手的手指的所述对象的位置,把所检测的位置映射到控制输入,并基于所述控制输入控制应用程序。

[0015] 实施方式可以包括一个或更多的下列特征。例如,所述方法可以包括基于具有对应于手指的解剖模型的拉长形状的对象,在所述图像中识别作为所述用户的手的手指的对象。所述方法还可以包括基于在平行于所述电子设备的显示器的平面中取向的手指的解剖模型,在所述图像中识别作为所述用户的手的手指的对象。所述解剖模型可以从手指的根部延伸到手指的指尖的手指模型。

[0016] 在某些实施方式中,所述方法可以包括把所检测的位置映射到显示区域中对应于手指的所检测的位置的位置,并把所显示的指示符移动到所述显示区域中对应于手指的所检测的位置的位置。所述方法还可以包括识别在所显示的用户界面中对应于手指的所检测的位置的位置处的输入控制,并确定和被识别的输入控制相关联的操作。基于所述控制输入控制所述应用程序可以包括执行和被识别的输入控制相关联的操作。

[0017] 在某些例子中,所述方法可以包括确定所述图像中手指的宽度,并基于所确定的宽度,估计手指相对于捕获所述图像的相机的深度位置。在这些例子中,可以基于所估计的深度位置映射所述控制输入。所述方法还可以包括检测手指的对应于指尖的部分,并确定所述位置为手指的对应于指尖的部分的质心。

[0018] 所述方法可以包括在所述图像内定义图像区域以使所述图像区域的边界与所述图像的边界分隔开所预期的指尖宽度的至少一半的距离,其中,可以在所定义的图像区域内检测手指的位置。当在所述图像区域外部检测到指尖时,该指尖的位置可以被映射到所述图像区域的边界上的位置。

[0019] 根据另一个一般实施方式,一种设备包括相机,该相机被配置成捕获与电子设备进行交互的用户的手的图像。所述设备还可以包括处理器,该处理器被配置成访问与电子设备进行交互的用户的手的图像,并基于对象的尺寸、对象的形状以及该对象是否和所述图像的边界相交,在所述图像中识别作为所述用户的手的手指的对象。所述处理器还被配置成确定在所述图像中被识别为所述用户的手的手指的所述对象的位置,把所检测的位置映射到控制输入,并基于所述控制输入控制应用程序。

[0020] 根据另一个一般实施方式,一种存储计算机程序产品的计算机可读介质。所述计算机程序产品包含指令,所述指令当被执行时工作以使计算机执行操作。所述操作包括访问与电子设备进行交互的用户的手的图像,并基于对象的尺寸、对象的形状以及该对象是否和所述图像的边界相交,在所述图像中识别作为所述用户的手的手指的对象。所述操作还包括确定在所述图像中被识别为所述用户的手的手指的所述对象的位置,把所检测的位置映射到控制输入,并基于所述控制输入控制应用程序。

[0021] 在附图和下面的描述中给出了一个或更多实施方式的细节。从所述描述和附图,本公开的其他的潜在特征和优点将会清楚。

附图说明

[0022] 图 1 是展示检测用户指尖的位置作为用于提供用户输入的方法的情境图。

[0023] 图 2 示出了示范性手持设备。

[0024] 图 3A 示出了包括面朝着观察手持设备的显示器的用户的相机的另一个示范性手持设备。

[0025] 图 3B 示出了包括从观察手持设备的显示器的用户转开 (facing away) 的相机的另一个示范性手持设备。

[0026] 图 4 示出了用于向设备提供用户输入的用户手指的位置的例子。

[0027] 图 5 是示出把检测到的手指位置和 / 或取向映射到控制输入并基于所述控制输入控制应用程序的过程的流程图。

[0028] 图 6 是利用相机检测用户的指尖的位置作为用于提供用户输入的方法的过程的流程图。

[0029] 图 7A 示出了手指从底部图像边缘进入视场 (field of view) 的例子。

[0030] 图 7B 示出了手指在相机的视场内成角度出现的例子。

[0031] 图 8A 示出了沿着扫描线的像素密度的例子。

[0032] 图 8B 示出了通过边缘检测过程计算的梯度大小 (magnitude) 的例子。

[0033] 图 9A 示出了链接的段的例子。

[0034] 图 9B 示出了被识别为手指的链接的段的集合的例子。

[0035] 图 9C 示出了代表手指的部分的子集的例子。

[0036] 图 9D 示出了手指的宽度测量的例子。

- [0037] 图 10A 示出了在相机图像内定义的图像区域的例子。
- [0038] 图 10B 示出了被定位成使得指尖位置在图像区域外部的手指的例子。
- [0039] 图 11A 示出了把图像区域映射到显示区域的例子。
- [0040] 图 11B 示出了指示符位置直接和手指位置相关的例子。
- [0041] 图 12 示出了在显示屏幕上显示的应用程序的用户界面的例子。
- [0042] 图 13A 示出了在显示屏幕上显示的游戏的显示图像的例子。
- [0043] 图 13B 示出了包括主视角 (first-person view) 的手指控制射击游戏的例子。
- [0044] 图 13C 示出了包括第三人视角 (third-person view) 的手指控制射击游戏的例子。
- [0045] 相同的参考符号在各个附图中指示相同的元素。

具体实施方式

[0046] 根据一个实施方式,使用相机来检测例如用户指尖的对象的位置作为用于提供用户输入的方法。例如,检测用户指尖的位置可用于滚过数据、控制光标位置和基于用户的指尖位置提供输入以控制视频游戏。还可以检测额外的信息,例如手指的取向。输入被提供给手持(或其他)设备,包括例如,蜂窝电话、视频游戏系统、便携式音乐(MP3)播放器、便携式视频播放器、个人数据助理(PDA)、音频/视频设备遥控器,和消费型数码相机,或者其他类型的设备。手持设备可以包括相机(camera),所以为了实施用于提供用户输入的系统,可以不要求额外的硬件。

[0047] 图 1 是展示检测用户指尖的位置作为用于提供用户输入的方法的情景图。手持设备 101(下面更详细地描述)包括显示器 103,显示器 103 显示在手持设备 101 上运行的应用程序或者操作系统的输出。用户使用用户的手指 201 向手持设备 101 提供用户输入。如图所示,应用程序是格斗(melee combat)游戏,并且使用把手指位置和取向映射到虚拟对象 120(例如一把剑)的表示(representation)的用户输入方法来实现控制该游戏。

[0048] 例如,用户的手指 201 的移动 111 被映射到游戏内的虚拟对象 120(例如剑或其他武器)的移动 112。在这个例子中,游戏使用三维计算机图形技术显示虚拟对象 120 的表示,因而虚拟对象 120 具有在虚拟体(virtual volume)内的位置和取向。在图像区域或者三维图像区域内的指根位置被映射到虚拟体内的虚拟对象根部位置,并且手指取向被映射到虚拟体内的虚拟对象取向。至此,虚拟对象 120 的位置和取向对应于用户手指 201 的位置和取向,所以用户的手指 201 映射到了虚拟对象 120。

[0049] 使用这种映射,用户例如向右挥动其手指 201 向游戏提供用户输入以使剑 120 向右砍。用户也可以使其手指 201 在和代表显示器 103 的前表面的平面平行的平面中取向,使用户的指尖指向垂直于显示器 103 的最高边缘的方向。用户手指 201 的这种取向使得剑 120 在虚拟环境中出现在和代表显示器 103 的前表面的平面平行的平面中,剑 120 的尖端指向垂直于显示器 103 的最高边缘的方向。如果用户在保持着取向的同时向上移动其手指 201(即沿着垂直于显示器 103 的最高边缘的方向),则用户输入被提供给游戏,使剑 120 在显示器中以对应于用户手指 201 的方式向上刺。

[0050] 类似地,如果用户在保持着取向的同时向着显示器移动其手指 201,则剑 120 保持对应于用户手指 201 的取向,并在虚拟环境中沿着进入显示器中的方向行进。如果用户在

保持着取向的同时从显示器移动开其手指 201, 则剑 120 保持对应于用户手指 201 的取向, 并在虚拟环境中沿着离开显示器的方向行进。就这样, 用户手指 201 的位置的变化映射到虚拟环境中虚拟对象 (即剑) 120 的位置的变化, 所述虚拟对象 (即剑) 120 的位置的变化由显示器 103 上的虚拟对象 120 的显示中的变化代表。

[0051] 在某些例子中, 用户可以改变用户手指 201 的取向以改变剑 120 在虚拟环境中的取向。例如, 用户可以倾斜其手指 201 以使剑 120 在虚拟环境中倾斜。剑 120 可以在虚拟环境中倾斜特定的度数, 所述度数直接地或者成比例地对应于用户倾斜其手指 201 的度数。这样, 用户手指 201 的取向的变化映射到虚拟环境中虚拟对象 (即剑) 120 的取向的变化, 所述虚拟对象 (即剑) 120 的取向的变化由显示器 103 中虚拟对象 120 的显示中的变化代表。

[0052] 对用户手指 201 的位置和取向的变化都可以被映射到虚拟环境中剑 120 的位置和取向。例如, 用户可以向着显示器倾斜或者转动其手指 201 以便通过以对应于用户的手指 201 的移动的方式 (例如以对应于向着显示器倾斜或者转动用户的手指 201 的方式) 倾斜或者转动剑 120, 使剑 120 在虚拟环境中沿进入显示器的方向砍。用户也可以把其手指 201 指向显示器, 并在把其手指 201 指向显示器的同时向着显示器移动其手指 201。这种移动可以使剑 120 沿着进入显示器中的方向刺。在某些例子中, 深度感测相机 (例如立体相机或者飞行时间 (time-of-flight) 相机) 可被用来检测用户的手指 201 距相机的距离。

[0053] 应用程序 (即格斗游戏) 把虚拟环境 (例如虚拟体) 内的虚拟对象 120 的位置和取向与虚拟环境内的其他虚拟对象进行比较, 以控制游戏进行。基于所述比较, 应用程序 (即格斗游戏) 确定虚拟对象 120 是否与虚拟目标对象 (例如虚拟人物、另一个虚拟武器、目标等) 相交 (intersect)。虚拟目标对象也可以被表现在显示器上。在显示器上显示虚拟目标对象允许用户向着或向虚拟目标对象移动其手指 201, 以使虚拟对象 120 在虚拟环境中以对应的方式移动, 并因而与虚拟目标对象相交。把用户的手指 201 的位置和取向映射到虚拟对象 120 的位置和取向允许用户使用其手指 201 控制格斗游戏。

[0054] 参考图 2, 手持设备 101 包括相机 102 和为应用程序提供视觉显示的显示器 103 (例如 LCD 屏幕)。手持设备 101 还可以包括按钮 104 或者小键盘 105。内部的处理器被配置成从相机 102、按钮 104 和 / 或小键盘 105 接受输入, 并基于所述输入执行功能或者操作, 并在显示器 103 上表现图像。

[0055] 相交 102 是捕获图像的设备。例如, 相机 102 可以是数码相机、数码摄像机, 或者任何其他类型的捕获图像的设备。设备 101 可以包括一个或更多相机。相机 102 可以捕获与应用程序进行交互的物理对象的图像。例如, 相机 102 可以捕获在相机 102 的视场内与应用程序物理地进行交互的用户的手指或者指尖的图像。

[0056] 相机 102 还可以包括照明手指 201 的发射器 (emitter), 所以相机可以在低环境光中检测手指。发射器可以产生红外光, 并且相机 102 可以对红外光敏感。或者, 发射器可以产生可见光 (例如白光), 并且相机 102 可以对可见光敏感。代替发射器, 显示屏幕 103 可以发射光以便可以在低环境光中检测手指 201。

[0057] 相机 102 也可以是深度感测相机 (depth sensing camera)。例如, 相机 102 可以是包括多个传感器的立体相机, 以便从不同的视点获取手指 201 的图像。相机 102 能够通过比较每个视点的图像来感测深度, 或者, 该设备可以通过在每个视点中检测手指 201 并

比较手指位置来计算深度。在另一个例子中,相机 102 可以是飞行时间相机,其包括发射光脉冲的发射器和能够测量光脉冲被感测到的时间的传感器。飞行时间相机能够通过比较光脉冲被发射的时间和被感测的时间、从而测量光所占路径的长度来感测深度。

[0058] 显示器 103 向用户表现可以包括文本或者图像的视觉显示。视觉显示可以指示关于设备 101 的状态的信息,或者为诸如游戏的应用程序或者其他在设备 101 上执行的应用程序提供视觉界面 (visual interface)。显示器 103 也可以显示图形、图像或者文本,其包含用于所描述的实施方式中的软件应用程序以及操作设备 101 所需的操作系统程序的用户界面。

[0059] 按钮 104 是用户可以用来向设备 101 提供输入命令的输入控制按钮。按钮 104 可以对应于基于在显示器 103 上表现的界面改变或者不改变的特定功能。按钮 104 可用来激活使用户能够基于在由相机 102 拍摄的图像中捕获的用户手指的移动向设备 101 提供输入的输入控制程序。例如,用户手指的移动可以在按钮 104 被按下时被映射到用户输入,而在按钮 104 未被按下时可以被忽略。使用按钮 104 来控制用户输入程序的激活可以防止用户手指的偶然移动(例如当用户从其口袋或者包中取出设备 101 时)被映射到用户输入。按钮 104 的激活也可以控制相机 102 捕获图像,并控制处理器执行在把用户手指的移动映射到用户输入时使用的其他过程。

[0060] 设备 101 的用户使用小键盘 105 输入命令和数据来操作和控制操作系统程序以及应用程序。小键盘 105 包括用户可以激活以向设备 101 提供输入的一组一个或更多按钮或者压力激活的开关。小键盘 105 可以使用户能够把文本数据和用户命令输入设备 101 中。

[0061] 处理器(未示出)可以从按钮 104 和小键盘 105 接受输入,并可以分析由相机 102 捕获的图像以便把用户手指的移动映射到用户输入命令。处理器可以基于用户输入执行功能或者操作,并且可以在适当时响应于用户输入在显示器 103 上表现显示。处理器可以执行在手持设备 101 上运行的应用程序和操作系统程序。手持设备 101 可以包括多个处理器(或者其他控制电路),并且可以包括存储应用程序、操作系统、用户输入程序以及由应用程序、操作系统和用户输入程序使用的数据的存储器(或者其他计算机可读存储介质)。

[0062] 尽管在图 2 中设备 101 被示出为移动设备,但是在进一步的方面中,设备 101 可以包括个人计算机(PC)或者游戏系统、膝上型电脑、手持或平板计算机、个人数据助理(“PDA”)或者例如计算机键盘或者遥控器的其他类型的嵌入式系统的部件。设备 101 也可以是便携式音乐播放器、寻呼机或者其他通信设备,或者用于游戏、通信和/或数据组织的手持或便携式电子设备。

[0063] 参考图 3A,在一个一般实施方式中,手持设备 101 包括面朝着观察手持设备 101 的显示器的用户的相机 102。在这个实施方式中,用户将其手指 201 保持在手持设备 101 上方,在相机 102 前面较短的距离,以便向手持设备 101 提供用户输入。相机 102 可以被定位成使得相机的视场覆盖显示器前面的区域。用户也使用其手指 202 来操纵按钮 104。

[0064] 参考图 3B,在另一个一般实施方式中,手持设备 101 包括从观察手持设备 101 的显示器的用户转开的相机 102。在这个实施方式中,用户将其手指 201 保持在手持设备 101 下面,在相机 102 后面较短的距离,以便向手持设备 101 提供用户输入。如图所示,相机 102 的暴露部分(exposed portion)在手持设备 101 的后侧,并且不被看到。用户也使用其手指 202 来操纵按钮 104。尽管图 3A 和图 3B 示出了两个示范性相机结构,但是在其他的实施

方式中,相机也可以被设置在手持设备的任何表面或者侧面上。

[0065] 在图 3A 和图 3B 中所示的实施方式中,用户通过在相机 102 成像的区域内移动手指 201 来与手持设备 101 进行交互。此外,用户可以通过使用拇指 202 或者另一个手指来按按钮 104 或者小键盘 105 的按键来与手持设备 101 进行交互。

[0066] 可以以各种方式握持和 / 或操作设备。例如,用户可以用第一只手握持设备,同时使用第二只手的手指来指点。在另一个例子中,用户可以用用来握持设备的同一只手的手指来指点。在又一个例子中,设备可以是静态的(例如搁在桌子上),并且用户可以用任一手指指点。

[0067] 图 4 示出了向设备提供用户输入的用户手指 201 的位置的例子。例如,用户保持手指 201 以使指尖 401 处于相机的视场 402 内。用户的指尖 401 可以是距相机 102 二到十厘米的距离(尽管也可以使用更小或者更大的距离)。相机 102 可以是通常用于拍摄照片或者视频的类型或者结构,并且如果手指 201 是离焦的(out of focus),其仍可用于提供用户输入。具体来说,如下面更全面地描述的那样,处理器分析相机图像以便分割手指 201 的图像。处理器在分割相机图像之后计算指尖 401 的取向和位置。

[0068] 图 5 是示出把检测到的手指位置和 / 或取向映射到控制输入并基于控制输入控制应用程序的过程 500 的流程图。为了方便,执行所述过程时引用了针对图 2 到图 4 描述的具体部件。但是,类似的方法学可以应用在其他实施方式中,其中,使用不同的部件来限定系统的结构,或者,其中功能被在图 2 到图 4 所示的部件间不同地分布。

[0069] 设备 101 的处理器在与电子设备进行交互的用户的手的图像中检测用户的手的位置和 / 或取向(510)。例如,处理器 205 可以处理或者分析一个或更多访问的图像,以便在所访问的图像中检测用户的手指。在这个例子中,处理器可以通过把所访问的图像中的特征与图像中的手指的典型特征(例如手指的解剖模型)进行比较来检测手指,或者,可以通过检测所访问的图像中的运动来检测手指。所述处理器也可以基于所访问图像的皮肤颜色分析来检测手指,或者,可以通过把包括一个或更多手指的当前图像与背景模型进行比较以检测在当前图像和背景模型之间的差别来检测手指。处理器 205 可以使用任何类型的技术在单个图像内或者多个图像的集合内检测手指。

[0070] 在一个或更多图像中检测手指以后,处理器分析所述一个或更多图像的一部分以确定所检测的手指的位置和 / 或取向。处理器可以基于在检测到手指的相机的视场内的区域来确定手指的位置,并且处理器可以基于相机的视场内的手指的角度来确定手指的取向。处理器可以在二维或者三维中确定手指的位置。当在三维中确定手指的位置时,处理器可以基于手指的尺寸估计手指距相机的距离,或者可以处理从立体或者飞行时间相机捕获的包括深度信息的图像。类似地,处理器可以在二维空间或者三维空间中确定手指的取向。在检测手指的位置和取向时,处理器可以计算代表在三维空间中所检测的手指的位置和取向的向量。

[0071] 处理器把所检测的位置和 / 或取向映射到控制输入(520)。处理器可以基于在检测位置和 / 或取向时表现的显示,把所检测的手指的位置和 / 或取向映射到控制输入。例如,处理器可以把手指的位置和 / 或取向与在所显示的用户界面中表现的物件的位置和 / 或取向进行比较。基于所述比较,处理器可以把手指的所检测的位置和 / 或取向映射到和所显示的界面控制(例如所显示的对应手指的位置和 / 或取向的界面控制)相关联的特

定命令或操作。

[0072] 处理器也可以把所检测的位置和 / 或取向映射到在用户界面中显示的对象的位置和 / 或取向。例如,处理器可以把所检测的位置和 / 或取向映射到在用户界面中显示的光标(例如鼠标光标)的位置和 / 或取向。这样,用户的手指的移动可以被映射到光标在用户界面中的移动,使用户能够使用光标控制应用程序。

[0073] 处理器也可以把所检测的位置和 / 或取向映射到虚拟环境中的虚拟对象的位置和 / 或取向。例如,如上面针对图 1 所述,处理器可以把所检测的位置和 / 或取向映射到虚拟环境中的虚拟对象(例如格斗游戏中的武器),以使虚拟环境中的虚拟对象的位置和 / 或取向对应于所检测的位置和 / 或取向。

[0074] 处理器还可以基于由所检测的位置和 / 或取向代表的手势(gesture),把所检测的位置和 / 或取向映射到控制输入。处理器可以分析所检测的手指的移动(例如改变位置和 / 或取向),以确定用户是否做出了与控制输入相关的手势。例如,处理器可以把所检测的手指的移动与一组认识的(recognized)手势进行比较,基于所述比较检测所认识的手势的做出,并把移动(例如手指位置和 / 或取向)映射到和所认识的手势相关联的控制输入。在一个例子中,用户可以在相机前面按照字母“Z”的形状移动其手指来提供将设备 101 关闭或者使设备 101 进入休眠模式的输入。

[0075] 处理器基于控制输入控制应用程序(530)。例如,处理器可以执行对应于控制输入的应用程序中的操作或者命令。处理器可以基于控制输入选择所显示的界面控制,或者,可以基于控制输入控制所显示的对象(例如,光标或者虚拟对象)。控制输入可以和数据处理应用程序、游戏、操作系统、媒体应用程序、设备控制应用程序,或者任何其他类型的应用程序中的命令或者操作相关。

[0076] 图 6 是示出利用相机来检测用户的指尖的位置和 / 或取向作为用于提供用户输入的方法的过程 600 的流程图。如上针对标号 510 所描述的,过程 600 可以用于基于图像检测手指的位置和 / 或取向。为了方便,执行所述过程时引用了针对图 2 到图 4 描述的具体部件。但是,类似的方法学可以应用在其他实施方式中,其中,使用不同的部件来限定系统的结构,或者,其中功能被在图 2 到图 4 所示的部件间不同地分布。

[0077] 手持设备 101 的处理器从相机 102 获取相机图像(610),并对相机图像进行边缘处理(620),以检测手指的轮廓。使用该轮廓,处理器分割相机图像(630)以便识别手指。处理器基于所识别的手指计算指示手指的位置和 / 或取向的信息(640)。

[0078] 在某些实施方式中,设备(例如相机 102)捕获彩色格式的图像,在所述彩色格式中,彩色通道(color channels)被打包(packed)(例如,16 比特字包含 5 比特红色、6 比特绿色和 5 比特蓝色)。在一个示范性实施方式中,获取相机图像(610)包括把相机图像转换为对于后续处理更为有效和方便的彩色格式。例如,处理器可以把彩色格式转换为 8 比特灰阶(grayscale)或者 24 比特 RGB 彩色。

[0079] 在其他的实施方式中,由于手指接近相机,手指出现模糊,因此,图像中的高频成分可能不归因于手指。在这种情况下,获取相机图像(610)还包括向图像施加低通滤波器。施加低通滤波器可以去除比归因于手指的频率成分更高的频率成分,并且可以减小副边缘(ancillary edges)的大小和信号噪声的影响。

[0080] 在进一步的实施方式中,设备(例如相机 102)以比所描述的过程要求的更高的分

分辨率捕获图像。在这种情况下,获取相机图像(610)可以包括抽取(decimating)图像。抽取的方法可以包括合并(binning),藉此,被抽取的像素值被作为区域内多个像素值的平均来计算。合并过程可以减小信号噪声的影响,因为在多个像素的区域内平均噪声成分倾向于具有比该区域内单个像素的噪声成分更低的大小。可以使用子采样或者其他的抽取方法。

[0081] 边缘检测过程(620)检测手指周围的轮廓。边缘检测过程计算每一个像素处图像密度的梯度。处理器可以在边缘检测过程中使用索贝尔(Sobel)操作。索贝尔操作把一对内核与原始图像卷积来确定梯度的水平和垂直分量。下面在等式 1A 和 1B 中示出了这些内核计算:

$$[0082] \quad G_x = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} * A \quad (1A)$$

$$[0083] \quad G_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * A \quad (1B)$$

[0084] 在等式 1A 和 1B 中,A 是相机图像,并且 G_x 和 G_y 是水平和垂直梯度分量。水平和垂直分量被组合。组合结果计算为平方分量的总和的平方根,在下面等式 1C 中示出:

$$[0085] \quad G = \sqrt{G_x^2 + G_y^2} \quad (1C)$$

[0086] 可以使用更为有效的近似。使用近似可以使通常在手持设备上采用的处理器能够执行必要的计算。在近似中,组合结果计算为水平和垂直分量的绝对值的总和,在下面等式 1D 中示出:

$$[0087] \quad G = a \times |G_x| + b \times |G_y| \quad (1D)$$

[0088] 此外,水平和垂直分量可以和预先定义的标量 a 和 b 相乘。标量 a 可以大于或不大于标量 b。在某些实施方式中,标量 a 可以大于标量 b,所以系统可以对手指的侧面比对皮肤的褶皱更加敏感。可以使用其他的边缘检测方法替代索贝尔操作。

[0089] 作为手指接近相机的结果,手指在相机图像中可能显得模糊。形成手指的轮廓的梯度可能跨越几个像素的宽度。手指内的特征,例如皮肤褶皱、角质层和手指甲也可能形成边缘,但是模糊可以减小那些梯度的大小。

[0090] 在分割过程的实施方式中(630),相机图像可以按具有取向的扫描线处理。所述取向被选择成基本垂直于手指,以使每一个扫描线代表手指的一个横截面。取向可以预先确定或者不预先确定。例如,取向可以基于设备 101 内的相机 102 的取向和预期用户握持设备的通常取向。

[0091] 图 7A 示出了手指 201 从底部图像边缘 701 进入视场 402,所以手指取向 702 显得基本垂直的例子。在这个例子中,扫描线 703 被定义为像素的水平线。定义了间隔一个像素的多个平行的扫描线。

[0092] 图 7B 示出了手指 201 在相机的视场 402 内显得成角度的例子。在这个例子中,扫描线 705 被定义垂直于手指取向 704。定义了大致间隔一个像素的多个平行的扫描线。

[0093] 如图 7B 中所示,在扫描线取向成角度的例子中,处理器在边缘处理之前旋转相机图像。旋转处理旋转图像数据,以使扫描线呈现为被旋转的图像缓冲区 (image buffer) 内水平的像素行。在被旋转的图像缓冲区中,扫描线上的像素可以具有连续的存储器地址,所以不使用变换。

[0094] 在如图 7B 中所示,在扫描线取向成角度的其他实施方式中,分割过程在访问扫描线的像素的梯度结果时使用变换。在像素值被访问时,可以计算扫描线上的像素的存储器地址。在计算上,变换可能比或不比旋转图像更有效率。

[0095] 图 8A 示出了沿着扫描线 (例如图 7A 的扫描线 703) 的像素的强度的例子,并且图 8B 示出了由边缘检索过程计算的梯度大小的例子。在手指的任一侧的边缘被标记为 801 和 802。作为手指接近相机的结果,手指可以在相机图像中显得模糊,这使得边缘的斜率显得缓和,如图 8B 中所示。

[0096] 在分割相机图像时 (630),处理器识别局部最大,并在局部最大的点分割扫描线。局部最大被识别为作为 803 识别的、以所述局部最大为中心的区域内的最大值,并且其值比周围值至少大出被识别为 804 的阈值。扫描线段 (segments) 被链接形成图数据结构 (graph data structure)。如果段重叠,则段被链接到相邻扫描线的段。

[0097] 图 9A 示出了链接的段的例子,其中水平线代表段,而垂直线代表链接。在这个例子中,背景内的特征也可能形成边缘。在形成包括被链接的段的图数据结构以后,该图数据结构可以被修剪。在一个例子中,手指周围的边缘是不完整的,从而扫描线段包括手指的一部分和背景的一部分。通过修剪连接子集的那些段,被相对较少段 (相对于例如区域的尺寸) 链接的图的子集 (例如,区域) 可以被解除链接 (un-linked)。

[0098] 选择链接的段的集合作为手指。选择标准包括尺寸、形状,以及和预期手指进入图像 (例如如图 7A 中所示,在底部图像边缘 701) 的相机图像 (或被转动的图像) 的边界的相交。尺寸标准把像素计数和预先定义的范围进行比较。形状标准把宽度和高度的比率 and 预先定义的范围进行比较。边界相交标准确定链接的段的集合是否包括预先选择的扫描线的段。

[0099] 选择标准可以代表手指的解剖模型,并且手指的解剖模型可以和链接的段的集合进行比较,以确定链接的段的集合是否是手指。解剖模型可以包括定义相机图像中的手指的典型属性的信息,例如手指的典型尺寸和典型形状。由解剖模型定义的手指的典型形状可以反映从手指根部到手指尖的延伸的手指的拉长形状。解剖模型还可以包括皮肤纹理特征,例如指纹 (fingerprints) 预期的褶皱和预期的区域。手指的解剖模型可以代表整个手指,并且包括例如手指关节、手指和手的连接、手指甲、皮肤颜色等的特征。

[0100] 解剖模型还可以涉及以与设备进行交互的用户保持手指的方式取向的手指。例如,在用户通过将其手指保持在平行于设备的显示器的平面中与设备进行交互的实施方式中,解剖模型可以代表在平行于设备的显示器的平面中取向的手指的解剖模型。把解剖模型以与设备进行交互的用户保持手指的方式来取向可以有助于在与设备进行交互的图像中检测手指。

[0101] 在某些实施方式中,可以使用具有变化的特性的手指的多个解剖模型。例如,不同尺寸的解剖模型 (例如成年人和儿童模型) 或者不同性别模型 (例如男性和女性模型) 可被用来确定链接的段的集合是否是手指。具体来说,多个解剖模型中的每一个均可以和链

接的段的集合进行比较,直到找到匹配为止,或者所有的模型已经被使用,并且确定链接的段的集合不是手指为止。使用多个解剖模型可以实现更为大量用户的手指的检测。

[0102] 在与设备进行交互时,用户可以保持其手指处于相对于设备不同的取向。如果与设备进行交互的用户手指的取向和解剖模型相匹配,则可能不能正确地检测用户的手指。因此,具有变化的取向的手指的解剖模型可被用来帮助在较宽的取向范围上检测手指。把链接的段的集合与具有变化的取向的解剖模型进行比较也可以用于确定用户的手指的取向,所述取向可以被映射到控制输入。

[0103] 在某些结构中,可以使用手指登记过程把特定手指向特定设备进行登记(register)。例如,用户可以向设备提供识别信息并控制设备捕获用户手指的一个或更多图像。设备可以从所述一个或更多被捕获的图像产生对应于用户的手指的解剖模型,并把为用户的手指产生的解剖模型与识别所述用户的信息相关联地存储。当随后用户使用该设备并提供了识别信息(或者以其他方式被该设备识别)时,该设备使用为用户的手指产生的解剖模型检测与该设备进行交互的用户手指的位置和/或取向。使用专为特定用户的手指产生的解剖模型可以有助于检测该用户的手指。

[0104] 图 9B 示出了被识别为手指的链接的段的集合的例子。在手指轮廓周围的所有位置可能不都检测到边缘。例如,在 901 所指示的手指的轮廓的位置不包括检测到的边缘。结果,在 902 识别的段可以被修剪(例如,可以去除由虚线指示的部分)。在手指内检测到边缘。例如,在 903 识别的边缘是由指关节处的皮肤褶皱导致的,而在 904 识别的边缘是由手指指甲或者角质层导致的。被链接的段的路径绕过这些特征。

[0105] 处理器从所选择的链接的段的集合计算手指的位置和/或取向信息。选择靠近集合末端的段(例如在图 9B 中由 905 指示的段)作为代表指尖的子集。选择靠近集合末端的段(例如在图 9B 中由 906 指示的段)作为代表手指根部的子集的一部分。代表指尖或者手指根部的段的子集包括了在最远段的预先确定数量的扫描线内的段。

[0106] 图 9C 示出了代表手指的一部分的子集的例子。选择靠近集合末端的段的子集来代表指尖。例如,子集 907 代表指尖。靠近集合末端的段被选作代表手指根部的子集的一部分。例如,子集 908 代表手指的根部。

[0107] 处理器把指尖位置 909 计算为被识别的指尖部分(例如子集)907 的质心。可以使用浮点精度或者浮点的定点模拟来计算质心。处理器把根部位置 910 计算为被识别的手指根部部分(例如子集)908 的质心。可以使用浮点精度或者浮点的定点模拟来计算质心。浮点精度或者浮点的定点模拟允许以比相机图像更高的分辨率(或者,在图像被抽取的情况下,比被抽取的相机图像更高的分辨率)计算/估计质心。因此,使用浮点精度或者浮点的定点模拟可以通过允许边缘过程 620 和分割过程 630 处理较低分辨率图像来减小处理功耗要求。处理功率要求的减小可适于典型的手持设备的处理器能力。

[0108] 处理器把手指取向 911 计算为从根部位置 910 到指尖位置 909 形成的向量。在另一个实施方式中,处理器计算手指的矩(moments)(例如被修剪的所选择链接段的集合)以确定手指取向。作为选择,可以计算额外的特性,例如手指的平均宽度。

[0109] 参考图 9D,靠近指尖和手指根部,垂直于手指取向 911 测量手指的宽度。从指尖位置和根部位置之间的手指取向向量上的点测量指尖宽度 913。例如,在位于距指尖预先确定的距离的点 912 处测量指尖宽度 913。选择预先确定的距离以避免指尖的弯曲并避免

手指的指甲部分。避免手指的指甲部分可能是有益的,因为对手指的那个部分的分割可能不那么准确。从指尖位置和根部位置之间的手指取向向量的点测量手指根部宽度 915。例如,在位于距图像底部预先确定的距离的点 914 处(例如预先确定的扫描线)测量手指根部宽度 915。

[0110] 作为选择,可以使用手指宽度测量估计三维位置和取向。例如,指尖宽度 913 和手指根部宽度 915 可以和手指的解剖模型的对应宽度进行比较以估计手指的深度尺寸(例如,距相机的距离)。指尖宽度 913 依赖于指尖距相机的距离。测量的指尖宽度可以被反向映射到距离。例如,大量值的指尖宽度被映射到较短距离,并且小量值的指尖宽度被映射到较远距离。所计算的距离被包括作为指尖位置坐标的分量。同样地,手指根部宽度 915 被映射到手指根部距离,并被包括在手指根部位置坐标中。手指取向被计算为从手指根部位置到指尖位置形成的向量,包括被映射的距相机的指尖距离和被映射的距相机的手指根部距离坐标分量。所描述的分割过程可以被优化,以便在计算上有效率,并且可适于典型的手持设备的处理器能力。

[0111] 在分割(630)的另一个实施方式中,使用分水岭变换(Watershedtransformation)处理相机图像。边缘(例如梯度大小)被解析以识别局部最小。针对每一个局部最小定义一个区域。每一个区域被扩展,以包括具有小于或等于一个值的大小的相连接的像素。针对增加的值重复这个过程,直到区域集合被扩展成包括图像的所有像素。区域可以合并以形成更大的区域。基于和上面描述的类似的标准(例如尺寸、形状,以及在预期手指进入图像的情况下与相机图像的边界的相交)选择区域作为手指。和上面描述的类似,从所选择的区域计算手指的位置和取向信息。

[0112] 可以应用分水岭变换的各种实施方式。所选择的分水岭变换的实施方式可以依赖于设备的处理能力。例如,分水岭变换的计算较不密集的实施方式可用于具有有限的处理能力的设备。此外,获取图像可以包括抽取图像(如上所述)以便减小分水岭变换的处理要求。

[0113] 其他的方法可用于处理相机图像。例如,可以使用 Condensation 或“条件密度传播”(Conditional Density Propagation)过程。例如,可以利用已知的手指分割数据初始化 Condensation 过程。此后,当手指移动时,Condensation 过程预测连续的视频帧或者相机图像中的手指分割。Condensation 过程的每一次迭代均执行选择、预测和测量。选择使用参数化采样产生用于评估的潜在状态集合。预测使用动态模型估计目标的状态如何变化。测量把预测的状态和实际像素数据进行比较。Condensation 过程在追踪混乱背景中的目标时特别稳健,并且可能对手持设备有用,因为每视频帧迭代的数量和每一次迭代中采样的像素的数量可以根据手指设备的处理器能力调整。卡尔曼滤波(Kalman Filtering)是可用于处理相机图像的另一个过程。

[0114] 指尖位置坐标空间被变换成对应于显示器 103 以便在用户的手指移动和显示器 103 之间提供直观的映射。在相机朝着用户的实施方式中,如图 3A 所示,变换包括镜像变换。在这个实施方式中,相机从后侧观察手指,并且用户从前侧观察手指和显示器,所以把坐标系镜像允许相机对手指的观察对应于用户对手指和显示器的观察。在扫描线取向成角度的实施方式中,如图 7B 中所示,变换包括旋转与扫描线取向的角度相反的角度。在相机 102 的取向和显示器 103 的取向不同的实施方式中,变换校正了取向上的差别。

[0115] 参考 10A, 在相机图像 1002 内定义了图像区域 1001。处理器可以定义图像区域 1001, 并在基于所述图像检测手指的位置和 / 或取向时处理图像区域 1001, 如上面针对数字 510 所述。

[0116] 在指尖位置坐标空间变换包括旋转的实施方式中, 图像区域 1001 被相对于相机图像 1002 旋转。图像区域 1001 被定义为使得指尖位置 1007 可以移动到图像区域 1001 内的所有位置, 包括区域的边界。例如, 如距离 1003、1004、1005 和 1006 所指示, 距相机图像 1002 的边界的最小距离是预期的指尖宽度的一半。这个例子导致了指尖位置 1007 可以移动到图像区域 1001 内的所有位置的图像区域 1001, 因为指尖位置在指尖的中心, 并且指尖位置和指尖边缘之间的指尖的部分大约是指尖宽度的一半。此外, 在手指的根部, 选择从相机图像 1002 的边界到图像区域 1001 的距离 1006 以满足手指选择标准。

[0117] 图 10B 示出了手指定位使得指尖位置 1008 在图像区域 1001 外部的例子。在这个例子中, 指尖位置 1008 被映射到图像区域 1001 的边界上的最近位置 1009。

[0118] 图 11A 示出了把图像区域 1001 映射到显示区域 1101 的例子, 显示区域 1101 是显示器 103 的一部分。把图像区域 1001 映射到显示区域 1101 可用于把所检测的位置和 / 或取向映射到控制输入, 如上面针对数字 520 所述。显示区域 1101 可以包括整个显示器 103 或者显示器 103 的一部分。例如, 显示区域 1101 可以包括显示器 103 的活动区域, 状态条除外。相对于图像区域 1001 的指尖位置 1007 被映射到相对于显示区域 1101 的指示符位置 1102。

[0119] 图 11B 示出了指示符位置 1102 直接和手指位置相关的例子。对于在平行于显示器的平面上的手指位置, 手指位置和指示符位置之间存在一对一相关。手指 201 的移动 1103 被映射到指示符位置 1102 的类似移动 1104。例如, 如果用户把手指放在第一指尖位置, 则导致第一指示符位置。如果用户把手指移动到第二指尖位置, 则导致第二指示符位置。在这个例子中, 如果用户移动其手指所以其返回第一指尖位置, 则指示符也将返回第一指示符位置。用户可发现这比控制杆更直观。例如, 控制杆向指示符位置施加速度, 以使指示符位置以间接方式和控制杆的状态相关。

[0120] 在另一个实施方式中, 图像区域是三维区域, 并且指尖位置包括代表所计算的距相机的距离的分量。例如, 如上所述, 使用手指的宽度测量计算代表所计算的距相机的距离的分量。定义三维图像区域以使指尖位置能够移动到所定义空间内的所有位置。

[0121] 在一个例子中, 指尖靠近相机, 并且相机具有一个有限的范围, 在该范围中, 相机在平行于显示器的平面内对指尖移动成像 (例如小的移动可以导致指尖跑到相机的视场外)。在这个例子中, 区域边界被定义成包括距相机的一系列距离, 其中, 相机的视场不限制相机对阈值以下的指尖成像的区域。

[0122] 三维图像区域被映射到三维显示区域。三维显示区域是虚拟体的一部分, 并被使用三维计算机图形表现技术表现到显示器 103。

[0123] 在某些结构中, 相对于三维图像区域的指尖位置被映射到相对于三维虚拟体的指示符位置。指示符位置直接和手指的位置相关。在手指位置和指示符位置之间存在一对一相关。手指的移动被映射到虚拟体内指示符位置的类似移动。例如, 如果用户把手指放在第一指尖位置, 则导致虚拟体中的第一指示符位置。如果用户把手指移动到第二指尖位置, 则导致虚拟体中的第二指示符位置。在这个例子中, 如果用户移动手指所以其返回第一指

尖位置,则指示符也将返回虚拟体中的第一指示符位置。第二指示符位置在虚拟体中的高度、宽度和深度尺寸上可以和第一指示符位置不同。

[0124] 图 12 绘出了在显示屏 103 上显示的应用程序用户界面的例子。显示区域 1101 被示出,并且在对应于指尖位置的指示符位置 1102,显示光标图像 1202 被作为指示符。显示图像显示屏上对象,包括具有和个人计算机、手持设备或任何其他类型的电子设备上的那些类似的外观和功能的图标 1203、滚动条 1204、菜单 1205 和按钮(未示出)。基于位置指示符的控制功能和个人计算机上的鼠标光标类似。

[0125] 用户通过定位其指尖以使指示符位置在屏上对象的界限内,并点击按钮 104 或者小键盘 105 的按键来激活屏上对象。用户通过定位其指尖以使指示符位置在屏上对象的界限内,按下并保持按钮 104 或者小键盘 105 的按键,移动其指尖以使指示符位置移动,并释放按钮 104 或者小键盘 105 的按键来拖动屏上对象。

[0126] 在另一个实施方式中,用户通过定位其指尖以使指示符位置在屏上对象的界限内并做出手势来激活屏上对象。所述手势可以是移动指尖更靠近相机,其中,在小于预先定义的时限的持续时间上,距离减小了大于预先定义的阈值的量。如上所述,使用指尖宽度计算指尖的距离。使用开始所述手势时的指示符位置来确定输入,因为用户在做出所述手势的同时可能非有意地改变指示符位置。

[0127] 检测手势的过程的实施方式包括把所计算的指尖距离(或者指尖尺寸/宽度)存储在缓冲区中。数据被存储在环形数组(circular array)数据结构中,藉此新的数据样本代替了最老的数据样本。数据样本指尖距离(或者尺寸/宽度)被与当前指尖距离(或者尺寸/宽度)比较,以确定所述数据样本指尖距离(或者尺寸/宽度)是否比当前指尖距离(或者尺寸/宽度)更靠近了超过预先定义的阈值距离的距离。

[0128] 在另一个实施方式中,所描述的技术适于使用户能够滚过所显示的列表。例如,列表包括多个等级。在所述列表是联系人列表(例如地址和电话号码)的实施方式中,第一等级包括字母表字母的列表,第二等级包括归档在字母下的名字的列表,并且第三等级包括联系人的信息。在所述列表是歌曲列表(例如音乐)的实施方式中,第一等级包括艺术家列表,第二等级包括专辑列表(例如由艺术家创作的专辑),第三等级包括歌曲列表(例如专辑上包括的歌曲),并且第四等级包括歌曲。

[0129] 在列表是时间和日期的列表(例如日历)的实施方式中,第一等级包括月的列表,第二等级包括天(例如一月内的天)的列表,第三等级包括时间(例如一天内的时间),并且第四等级包括在一时间的信息(例如约会)。在另一个实施方式中,其中所述列表是时间和日期的列表(例如日历),第一等级包括星期的列表,第二等级包括天(例如一星期内的天)的列表,第三等级包括时间(例如一天内的时间),并且第四等级包括在一时间的信息(例如约会)。

[0130] 图像区域 1001 内的指尖位置 1007 的水平分量控制对列表的等级的选择。图像区域 1001 被划分为对应于等级的许多列。应用程序被配置成根据指尖位置位于的列显示信息。用户通过把其指尖向左或向右移动在等级之间切换。设备可以显示动画或者播放声音向用户指示等级的变化。

[0131] 在等级(level)内,项目的列表被显示,其中项目指示符按列排列。项目指示符可以包括文本和/或图标。图像区域 1001 内的指尖位置 1007 被映射到项目指示符排列的界

限,并与项目指示符排列的界限内的项目指示符位置比较。在对应于指尖位置的位置处的项目指示符被表现成指示其被选择。例如,可以用交替的颜色(突出显示)表现项目指示符,或者,当项目指示符对应于指尖位置时,可以用较大尺寸表现该项目指示符。其他的项目指示符可以移动以便为所述较大尺寸让出地方。设备也可以播放声音向用户指示选择的变化。

[0132] 在某些实施方式中,指尖距相机的距离控制对列表的等级的选择。预先定义的距离范围由许多个等级划分。应用程序根据对应于指尖距相机的距离的等级显示信息。用户通过移动其指尖更靠近或远离相机在等级之间切换。设备也可以显示动画或者播放声音向用户指示等级的变化。定义三维图像区域并且三维图像区域的距离分量被划分为层。

[0133] 在等级内,其中排列了项目指示符的项目列表被显示。项目指示符可以按列、行、网格,或者成行成列地,或者任意的布置排列,项目指示符可以包括文本或者图标。图像区域 1001 内的指尖位置 1007 被映射到项目指示符排列的界限,并与项目指示符排列的界限内的项目指示符位置比较。在对应于指尖位置的位置处的项目指示符被表现以指示其被选择。例如,可以用交替的颜色(突出显示)表现项目指示符,或者,当项目指示符对应于指尖位置时,可以用较大尺寸表现该项目指示符。其他的项目指示符可以移动以便为所述较大尺寸让出地方。设备也可以播放声音向用户指示选择的变化

[0134] 在某些排列中,用户可以通过改变用户手指的取向来改变等级。用户通过在平行于显示器 103 的平面内旋转用户手指来改变等级。用户可以通过把其手指定位在平行于显示器 103 的平面中并使其手指取向指着朝着九点钟方向来选择第一等级。然后,用户可以通过旋转其手指以便指着朝着十二点钟方向选择第二等级,然后通过旋转其手指以指着朝着三点钟方向来选择第三等级。

[0135] 此外,在某些例子中,用户通过沿着进入显示器 103 方向改变用户手指的取向或者角度。在这些例子中,用户通过朝着显示器 103 旋转用户的手指改变等级。用户通过保持用户手指的取向并保持着所述取向改变用户手指的位置在特定等级内选择项目。

[0136] 在图像区域被划分为行和列,或者三维图像区域被划分为行、列和层的实施方式中,可以动态地调整划分以使包括指尖的划分被放大。放大包括指尖的划分可以减小用户在试图切换行或列时非有意地切换层的可能性,或者在试图切换行时非有意地切换列的可能性。

[0137] 参考图 13A,在显示器 103 上显示了游戏。手指 201 的移动 1301 被映射到游戏内的移动 1302。在这个实施方式中,所描述的技术适于使用户能够控制射击游戏。例如,指尖位置被映射到瞄准功能。在手指控制的射击游戏的实施方式中,相对于图像区域的指尖位置被映射到相对于显示区域的靶位置。

[0138] 图 13B 描绘了包括主视角的手指控制的射击游戏的例子。显示区域 1101 被示出,并且瞄准图像 1303(例如十字瞄准线)在对应于指尖位置的指示符位置 1102 被显示为指示符。靶图像 1304 也被显示。用户可以移动其手指以便把瞄准图像 1303 定位在靶图像 1304 上。

[0139] 用户可以按按钮 104 或者小键盘 105 上的按键来激活射击功能。在另一个实施方式中,用户可以做出手势来激活射击功能。可以如上所述做出并检测所述手势。

[0140] 手指控制的射击游戏可以表现主视角,例如如图 13B 中所示。用户也能够改变虚

拟环境内的视角。如果用户移动其手指以使指示符位置 1102 在显示区域 1101 的左部分中的子区域 1305 内移动,则视角向左滚动(或者旋转)。如果用户移动其手指以使指示符位置 1102 在显示区域 1101 的右部分中的子区域 1306 内移动,则视角向右滚动(或者旋转)。类似地,如果用户移动其手指以使指示符位置 1102 在显示区域 1101 的上部分中的子区域(未示出)内移动,则游戏向上滚动(或者旋转),并且,如果用户移动其手指以使指示符位置 1102 在显示区域 1101 的下部分中的子区域(未示出)内移动,则游戏向上滚动(或者旋转)。如果用户移动其手指以使指尖距相机的距离比预先定义的阈值更近,则游戏向前移动(或者放大)。如果用户移动其手指以使指尖距相机的距离比预先定义的阈值更远,则游戏向后移动(或者缩小)。

[0141] 在另一个实施方式中,在具有主视角的手指控制的射击游戏中,用户通过按按钮 104、小键盘 105 上的按键、方向板或者控制杆改变虚拟环境内的视角。

[0142] 在又一个实施方式中,在手指控制的射击游戏中,手指位置控制虚拟环境内的视角。在这个实施方式中,瞄准图像可以是静态的,或者被自动功能控制。在一个例子中,速度被计算为图像区域(或者三维图像区域)内指尖位置相对于图像区域(或者三维图像区域)内的中心位置的差。在另一个例子中,速度被计算为图像区域(或者三维图像区域)内指尖位置相对于中心在图像区域(或者三维图像区域)内的子区域的边界上的最靠近部分的差,其中,子区域代表“死区”。速度的水平分量被施加到虚拟环境内的视点(viewpoint),以使虚拟环境的视角被水平地旋转(或滚动)。速度的垂直分量被施加到虚拟环境内的视点,以使虚拟环境的视角被垂直地旋转(或者滚动)。速度的距离分量被施加于虚拟环境内的视点,以使虚拟视点的位置在虚拟环境中沿着当前前视轨迹(forward looking trajectory)平移(或者缩放)。

[0143] 图 13C 示出了包括第三人视角的手指控制的射击游戏的例子。在这个例子中,在显示器 103 上表现玩家图像 1307。以使武器弹道(weapon trajectory) 1309 对准指示符位置 1102 的取向表现武器图像 1308。当用户移动其手指时,武器图像 1308 和武器弹道 1309 跟踪指示符位置 1102。在某些结构中,武器弹道 1309 对应于用户用其手指指着的方向,所以改变用户手指的取向改变了武器弹道 1309。

[0144] 在另一个实施方式中,在具有第三人视角的手指控制的射击游戏中,用户通过按按钮 104、小键盘 105 上的按键、方向板或者控制杆移动玩家图像 1307。手指控制的射击游戏可以采用各种形式。靶图像 1304 和玩家图像 1307 可以代表例如宇宙飞船、车辆、士兵、异形、怪物、卡通人物、猎人和猎物,等等。

[0145] 在另一个实施方式中,提供了用于在图像上滚动和缩放的用户输入技术。图像可以包括相片、地图,或者任何其他类型的图像。可以显示地图,其中,比例尺确定了被表现的细节的等级。例如,“缩小的”地图显示城市的位置,并且“放大的”地图显示各个街道。

[0146] 缩放图像(例如改变比例尺)的过程使用手指距相机的距离来控制比例尺或者图像的“缩放”。在手持被保持在设备上方的系统的实施方式中,靠近相机的距离对应于大比例尺(即被放大),而远离相机的距离对应于小比例尺(即被缩小)。

[0147] 在手指被保持在设备下方的系统的实施方式中,如图 3B 所示,靠近相机的距离可以对应于小比例尺(即被缩小),而远离相机的距离可以对应于大比例尺(即被放大)。在另一方面,靠近相机的距离可以对应于大比例尺(即被缩小),而远离相机的距离可以对应

于小比例尺（即被缩小）。用户可以选择优选的配置，并且所述优选的配置可以为用户应用。

[0148] 在把手指距相机的距离映射成控制图像的比例尺或者“缩放”的实施方式中，三维图像区域的距离分量被映射到图像的尺寸（或者“缩放”）范围。相对于三维图像区域的指尖位置的距离分量被映射到相对于图像的尺寸（或者“缩放”）范围的图像尺寸（或者“缩放”）。

[0149] 在另一个实施方式中，使用手指位置滚动（或者“摇动”）图像实现滚动图像的过程。图像区域 1001 内的指尖位置被映射到表现到显示器 103 的图像的显示区域位置。移动手指以使表现到显示器 103 的图像的显示区域移动，从而滚动或者“摇动”该图像。

[0150] 在又一个实施方式中，滚动和缩放图像的过程把图像区域 1001 映射到图像内的范围，其中，图像内的可滚动范围依赖于缩放等级。例如，用户在平行于显示器的平面中移动手指以便滚过整个图像，并且，一旦期望的图像部分被集中在显示器上，在用户沿着垂直于显示器的方向移动手指来“放大”。一旦“放大”，用户在平行于显示器的平面中移动手指以便滚过图像的一部分。对于小比例尺（即被缩小），可滚动范围被定义成包括整个图像，所以通过滚动图像，图像的所有部分都被显示。对于大比例尺（即被放大），可滚动范围被定义成包括一部分图像，所以图像被缩放的部分可以被精确地滚动。可滚动范围可以和比例尺成线性比例。

[0151] 在另一个实施方式中，提供了用于在虚拟环境内导航的用户输入技术。虚拟环境可以是游戏的一部分，包括但不限于上面所讨论的游戏类型。应用程序或者游戏使用计算机图形技术把虚拟环境表现到显示器。

[0152] 在虚拟环境内导航允许在虚拟三维空间内移动，并且系统提供控制在虚拟三维空间内导航的用户输入。例如，速度被计算为图像区域（或者三维图像区域）内指尖位置相对于图像区域（或者三维图像区域）内的中心位置的差。在另一个例子中，速度被计算为图像区域（或者三维图像区域）内指尖位置相对于中心在图像区域（或者三维图像区域）内的子区域的边界上的最靠近部分的差，其中，子区域代表“死区”。在计算过速度以后，该速度的水平分量被施加到虚拟环境内的视点，以使虚拟环境的视角被水平地旋转（或滚动）。速度的垂直分量被施加到虚拟环境内的视点，以使虚拟环境的视角被垂直地旋转（或者滚动）。速度的距离分量被施加于虚拟环境内的视点，以使虚拟视点的位置在虚拟环境中沿着当前前视轨迹平移（或者缩放）。

[0153] 在另一个实施方式中，提供了用于使用户能够定位和旋转虚拟对象的用户输入技术。在这个实施方式中，应用程序或者游戏使用计算机图形技术把虚拟环境表现到显示器。例如，用户通过在平行于显示器的平面中移动手指旋转虚拟对象。在这个例子中，图像区域 1001 被映射到旋转范围，并且图像区域内的指尖位置被映射到旋转范围内的角度。在另一个例子中，用户通过沿着垂直于显示器的方向移动手指“放大”和“缩小”。在这个例子中，指尖距相机的距离被映射到虚拟对象的比例尺。

[0154] 在除了上述那些以外的应用程序和游戏中，可以使用跟踪手指位置和取向的方法。此外，所描述的提供用户输入的技术可在包括相机或者其他成像设备的任何设备中使用。

[0155] 实施方式可以包括例如过程、设备或用于执行过程的设备。例如，实施方式可以包

括一个或更多被配置成执行一个或更多过程的设备。设备可以包括例如分立或者集成的硬件、固件和软件。设备可以包括例如计算设备或者其他的计算或处理设备,特别是假如被编程为执行一个或更多所描述的过程及其变体。这些计算或处理设备可以包括例如处理器、集成电路、可编程逻辑器件、个人计算机、个人数字助理、游戏设备、蜂窝电话、计算器和包含软件应用程序的设备。

[0156] 实施方式也可以具体实施在包括一个或更多计算机可读介质的设备中,所述计算机可读介质具有用于实现一个或更多过程的指令。所述计算机可读介质可以包括例如存储设备、存储器和编码或者传送指令的格式化电磁波。计算机可读介质还可以包括例如各种非易失或易失存储器结构,例如硬盘、快闪存储器、随机访问存储器、只读存储器和紧制盘。指令可以存储在例如硬件、固件、软件和电磁波中。

[0157] 因此,计算设备可以代表被编程为执行所描述的実施方式的计算设备的实施方式,并且存储设备可以代表存储用于实现所描述的實施方式的指令的计算机可读介质。

[0158] 已经描述了许多实施方式。尽管如此,将会理解,可以做出各种修改。因此,其他的实施方式在下列权利要求的范围内。

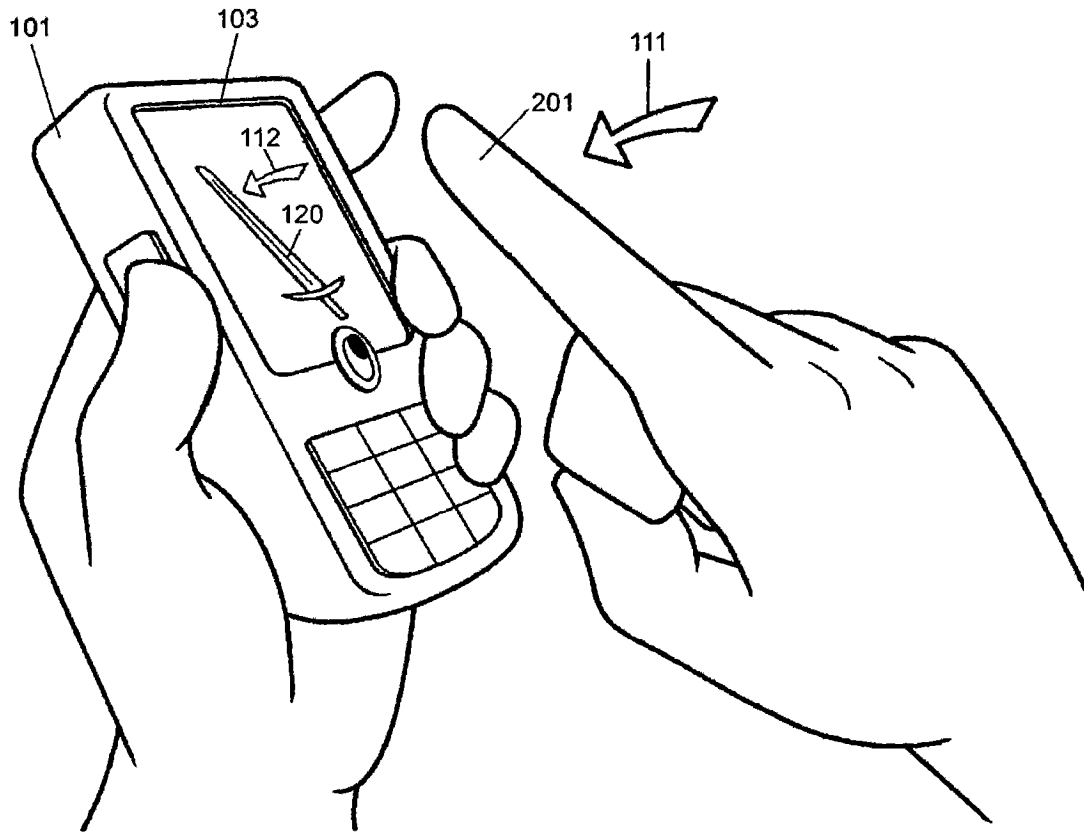


图 1

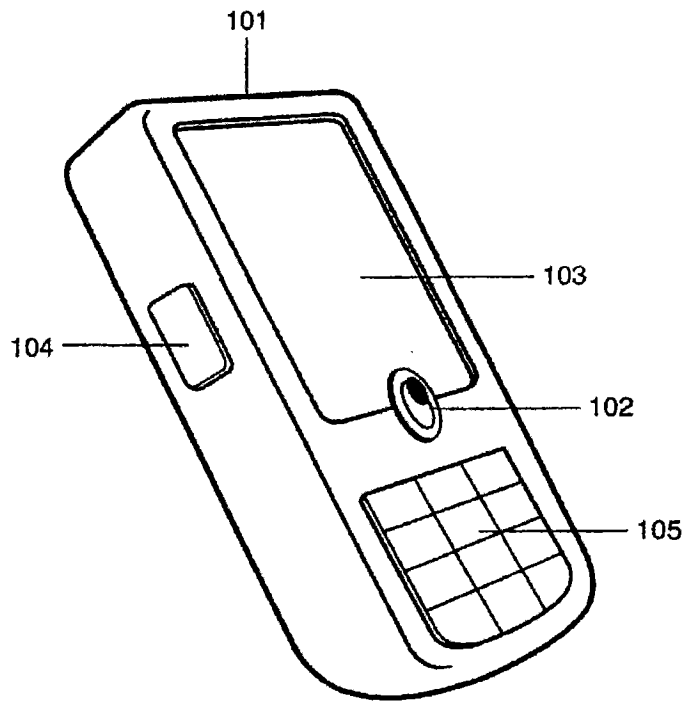


图 2

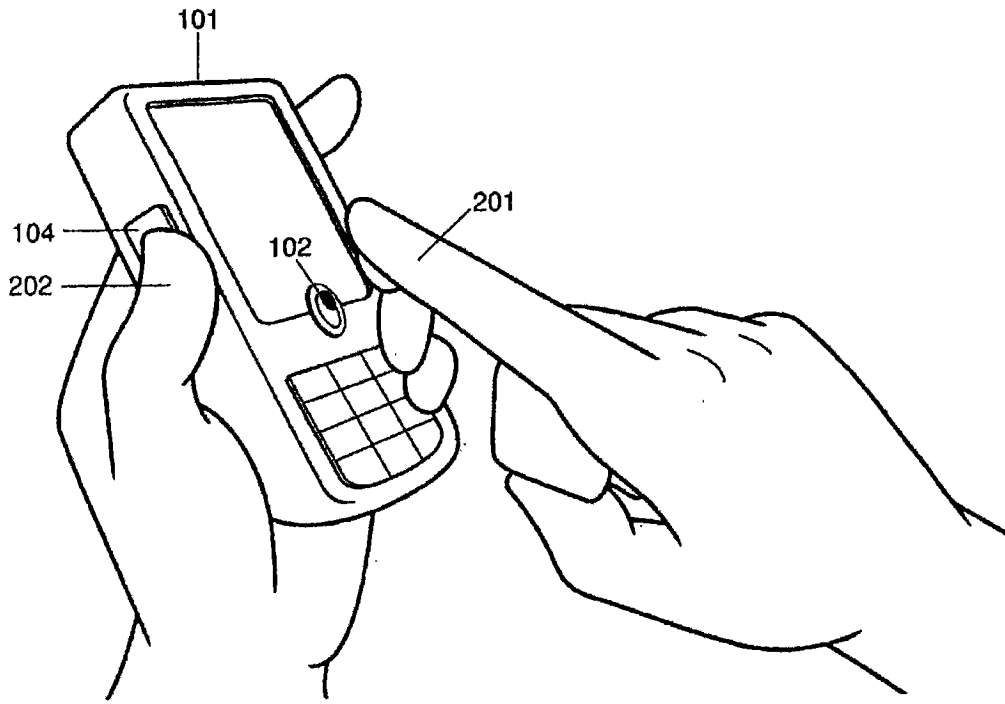


图 3A

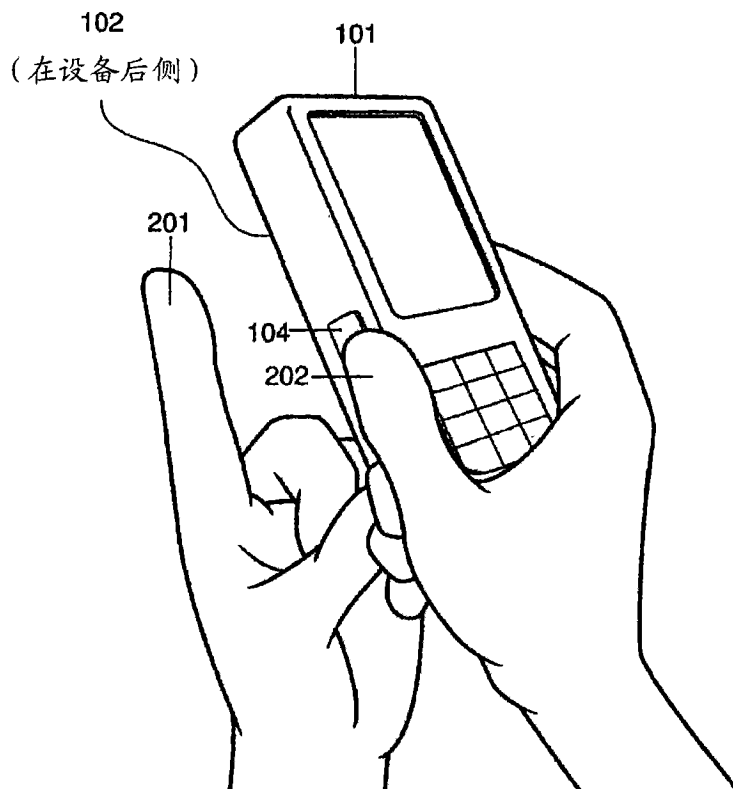


图 3B

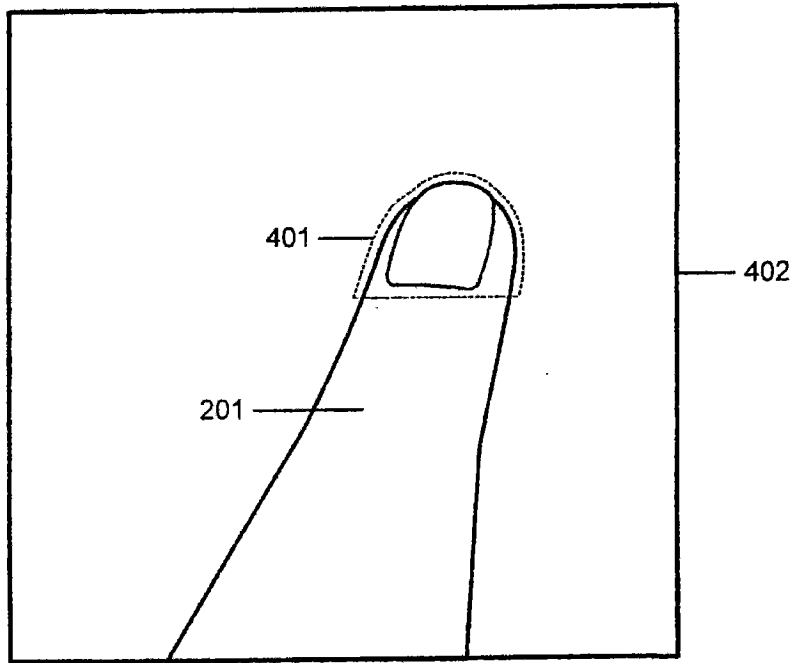


图 4

500

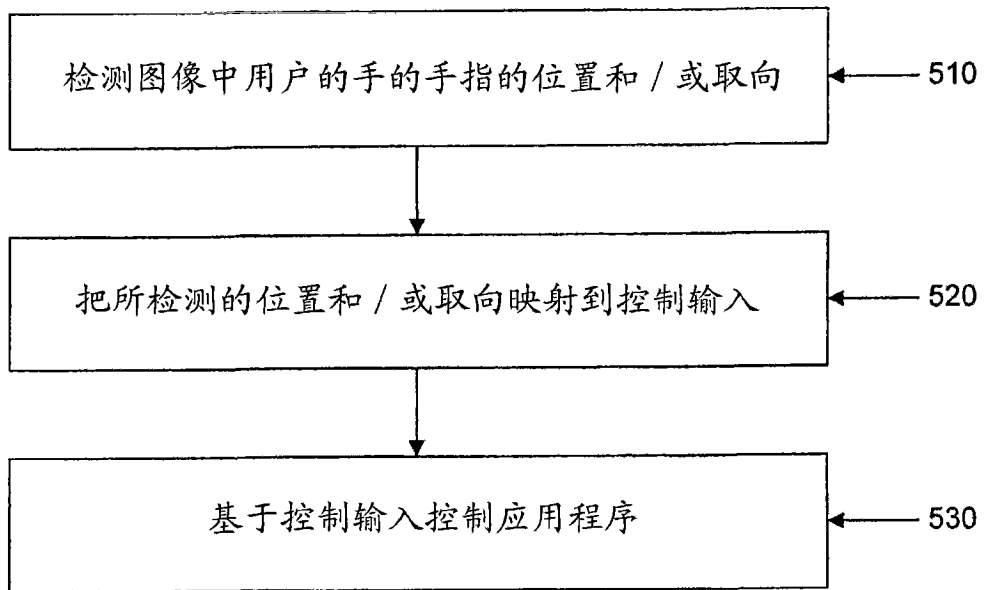


图 5

600

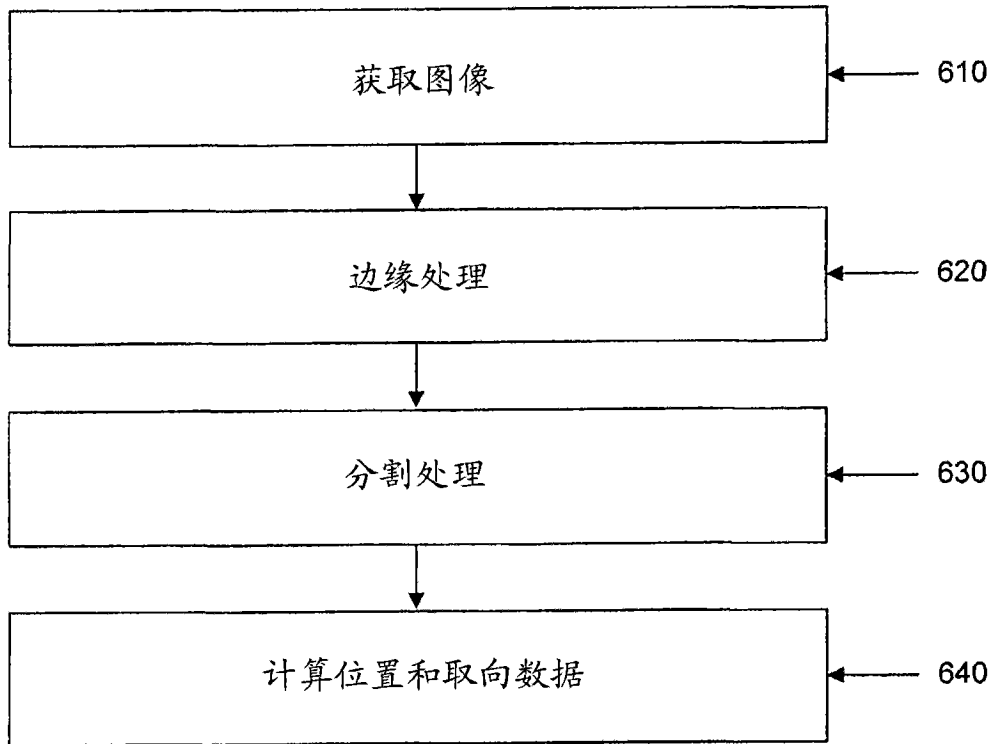


图 6

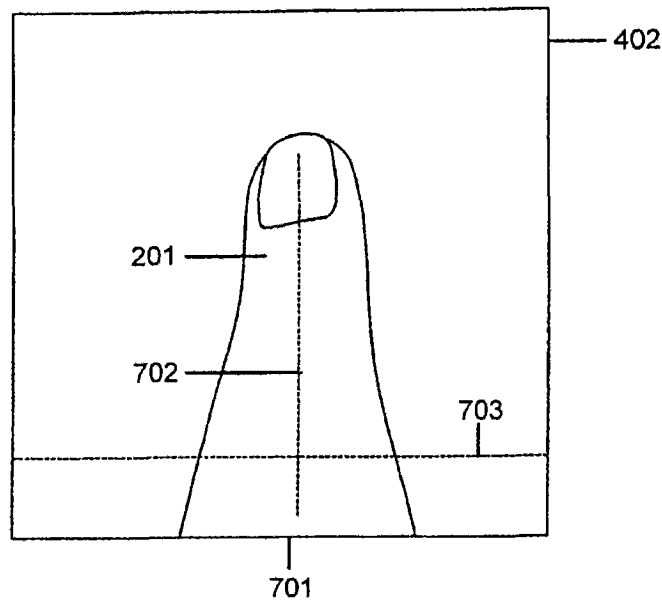


图 7A

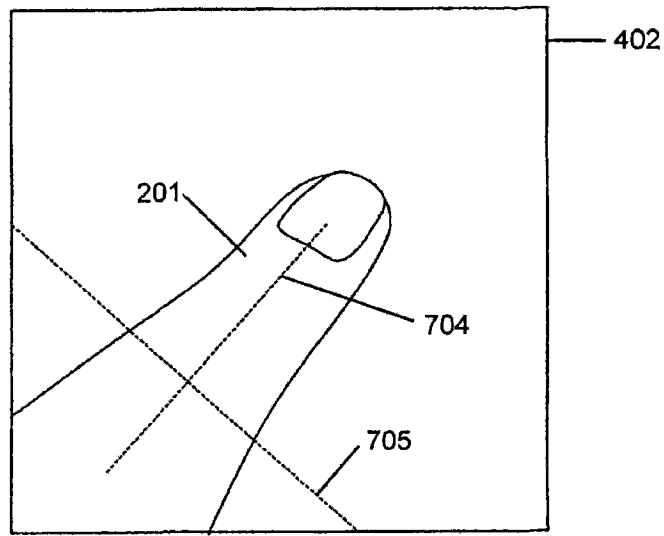


图 7B

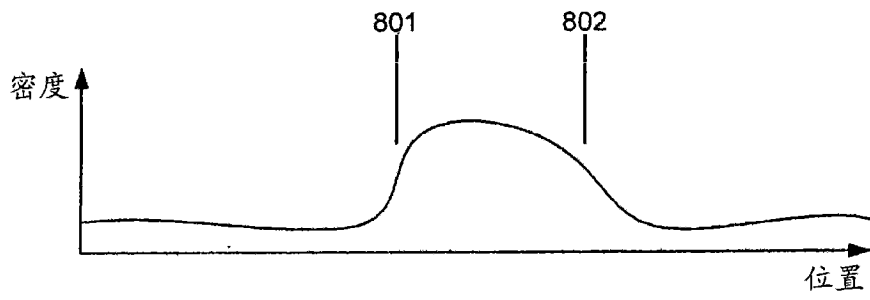


图 8A

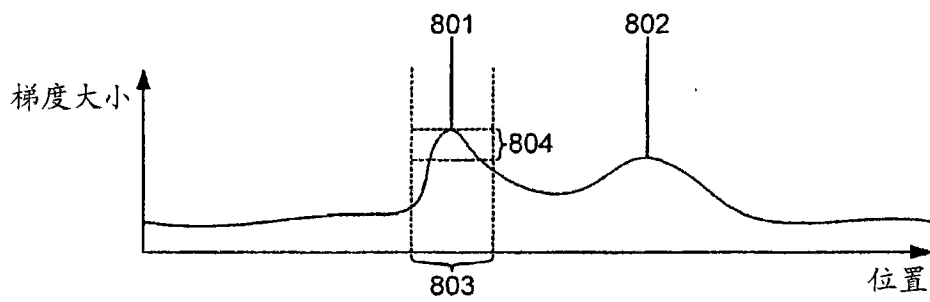


图 8B

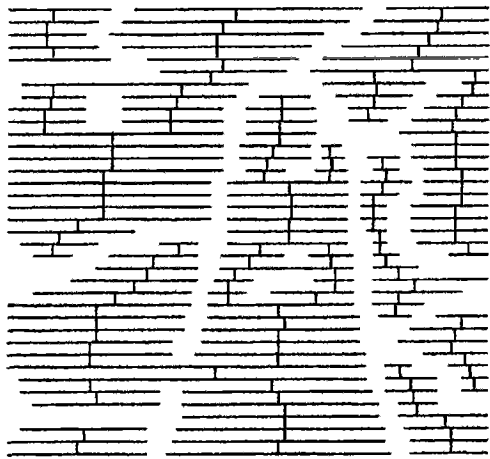


图 9A

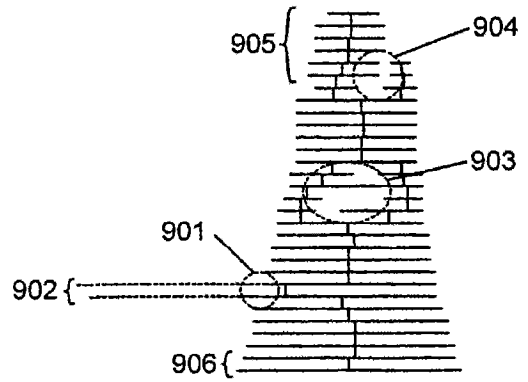


图 9B

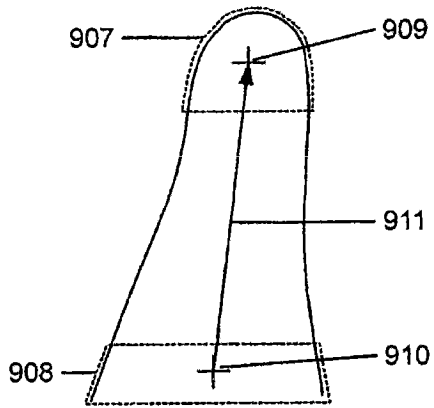


图 9C

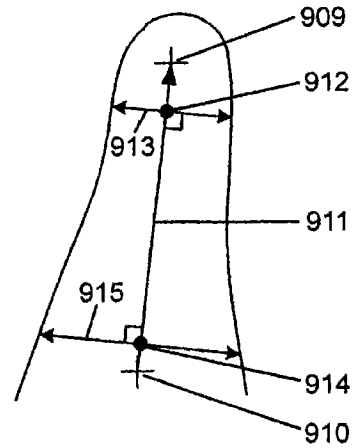


图 9D

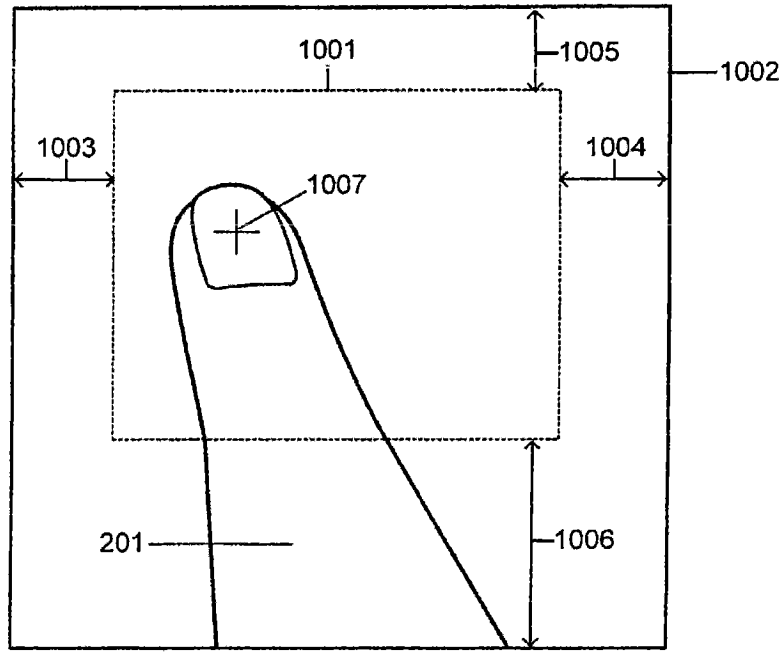


图 10A

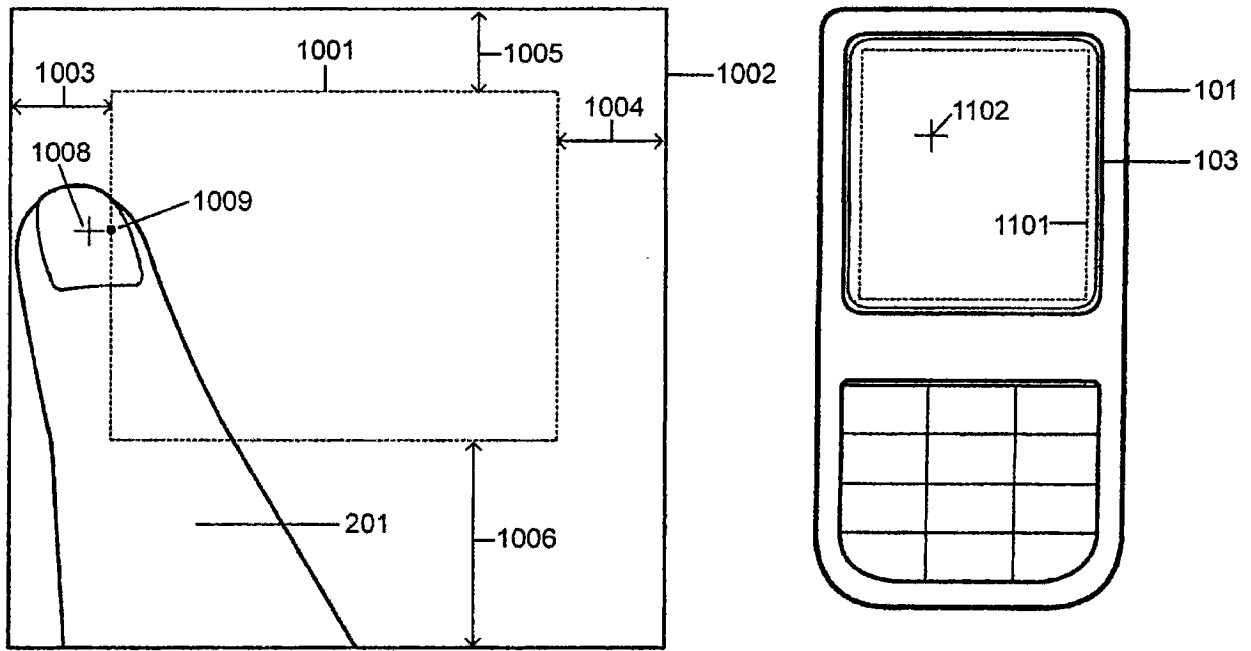


图 10B

图 11A

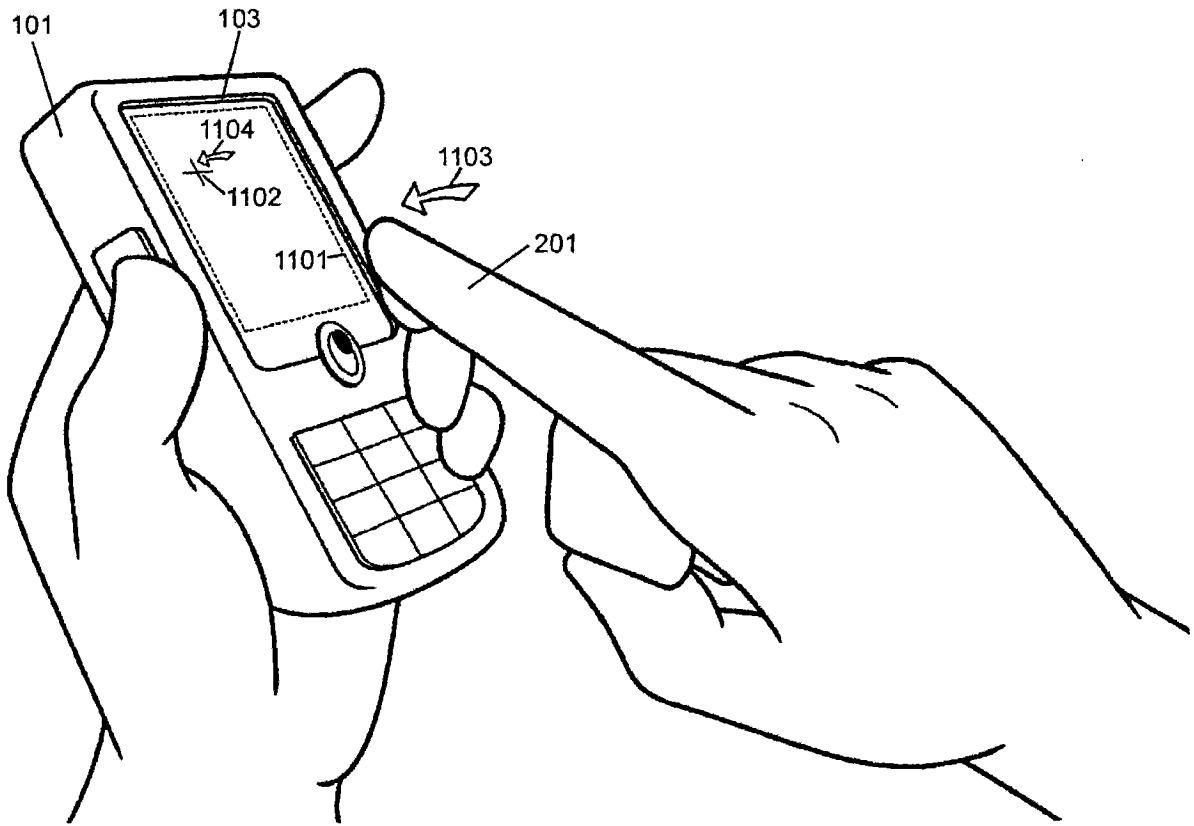


图 11B

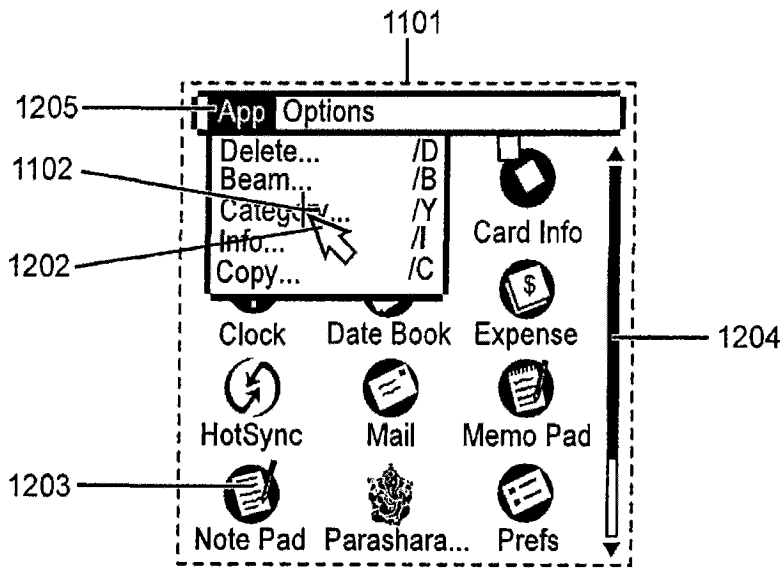


图 12

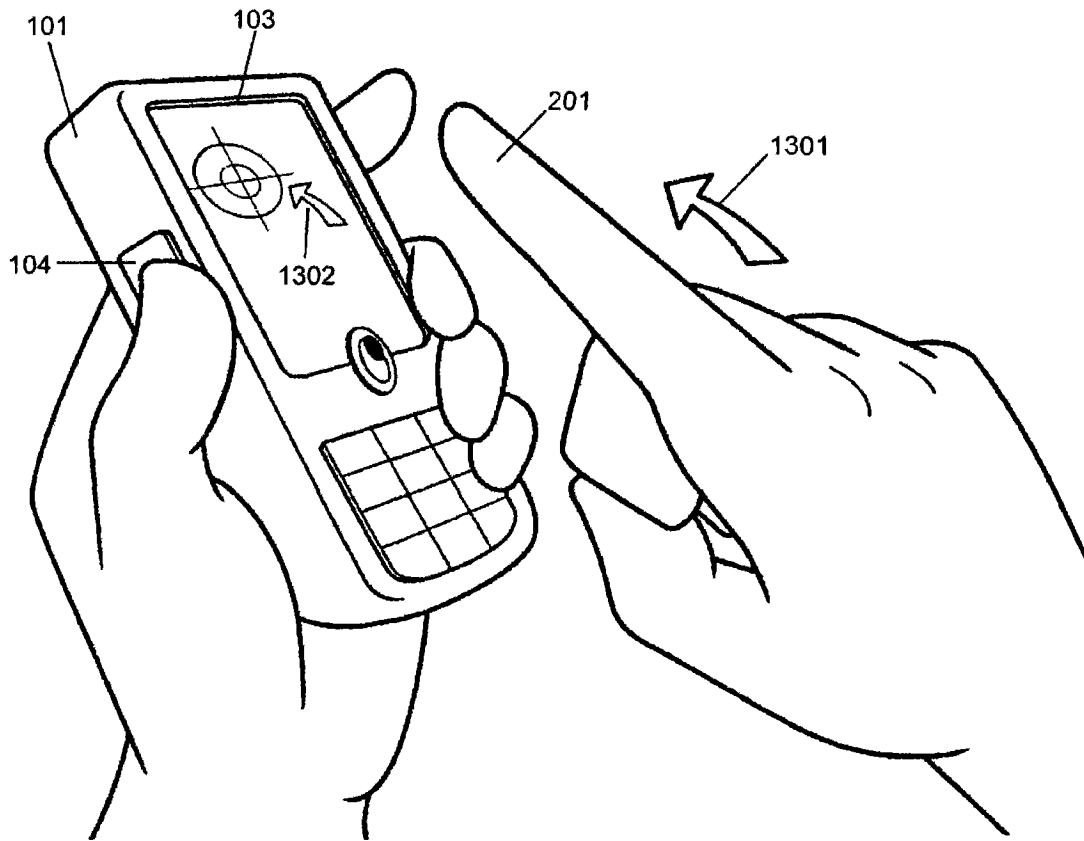


图 13A

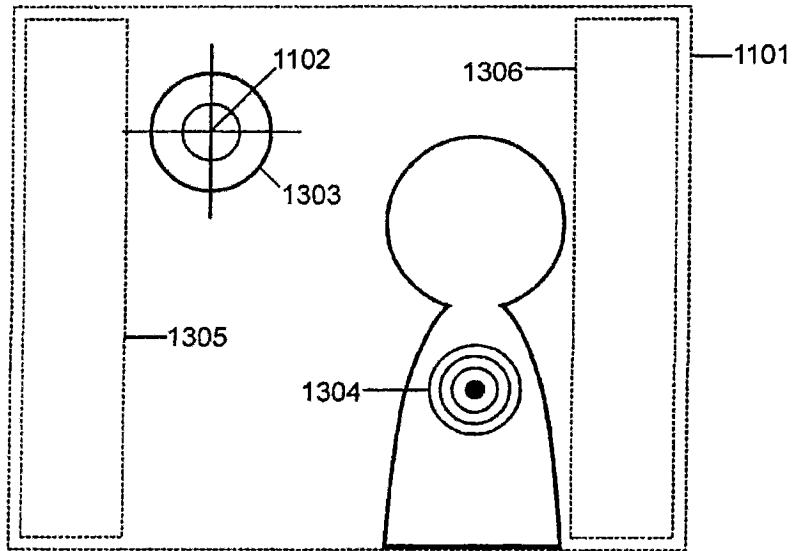


图 13B

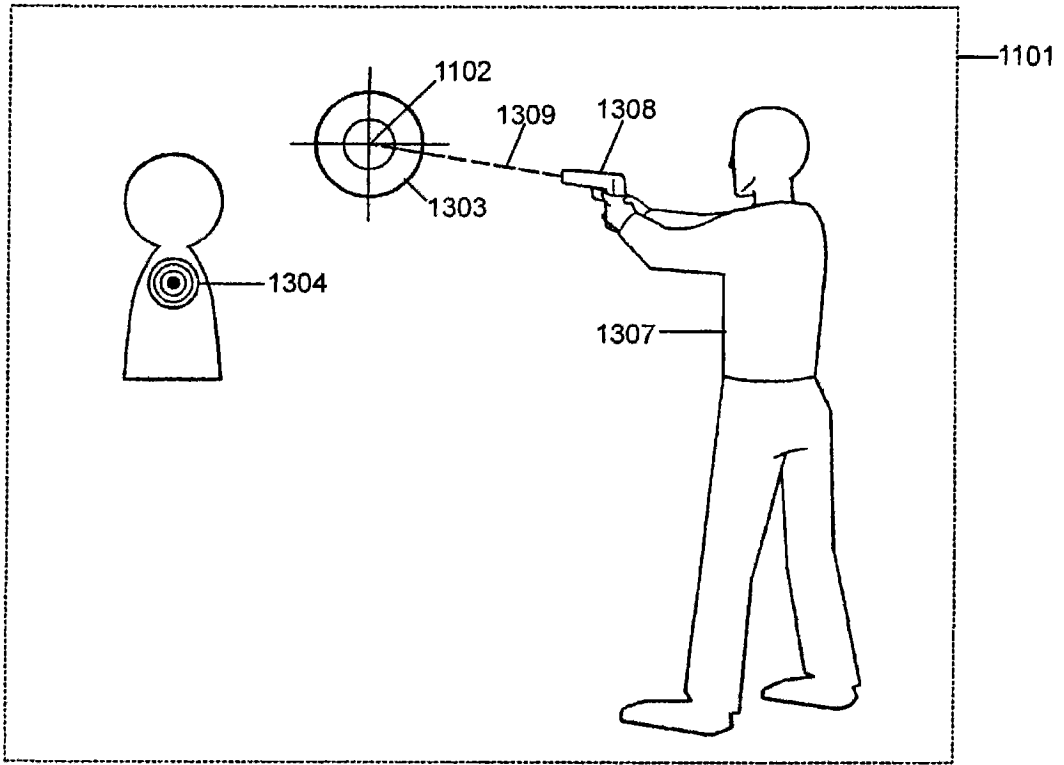


图 13C