(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 116798069 A
(43)申请公布日 2023.09.22

(21)申请号 202310283575.8
(22)申请日 2023.03.22
(30)优先权数据
2022-045114 2022.03.22 JP
(71)申请人 佳能株式会社
地址 日本东京都大田区下丸子3丁目30番2号
(72)发明人 田中裕树
(74)专利代理机构 北京魏启学律师事务所 11398
专利代理师 魏启学 王小香
(51)Int.Cl.
G06V 40/10(2022.01)
G06V 10/764(2022.01)
G06N 3/04(2023.01)
G06N 3/08(2023.01)

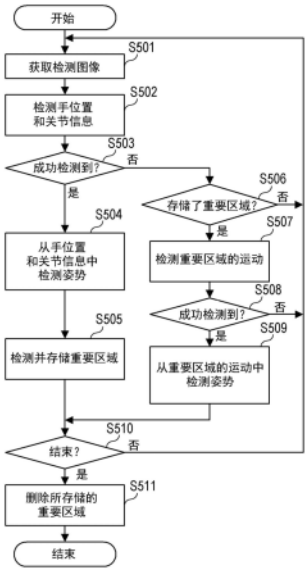
权利要求书2页 说明书5页 附图6页

(54)发明名称

姿势识别设备及方法、头戴式显示设备和计算机可读介质

(57)摘要

本公开提供了姿势识别设备及方法、头戴式显示设备和计算机可读介质。根据本发明的姿势识别设备包括:部位检测单元,其被配置为从拍摄图像中检测做出姿势的部位;识别单元,其被配置为基于所述部位检测单元所检测到的所述部位的运动来识别所述姿势;区域检测单元,其被配置为从所述部位中检测要用于所述姿势的区域;以及运动检测单元,其被配置为对所述区域检测单元所检测到的所述区域的运动进行检测,其中,所述识别单元基于所述运动检测单元所检测到的所述运动来识别所述姿势。



CN 116798069 A

1.一种姿势识别设备,包括:

部位检测单元,其被配置为从拍摄图像中检测做出姿势的部位;

识别单元,其被配置为基于所述部位检测单元所检测到的所述部位的运动来识别所述姿势;

区域检测单元,其被配置为从所述部位中检测要用于所述姿势的区域;以及

运动检测单元,其被配置为对所述区域检测单元所检测到的所述区域的运动进行检测,

其中,所述识别单元基于所述运动检测单元所检测到的所述运动来识别所述姿势。

2.根据权利要求1所述的姿势识别设备,其中,

在所述部位检测单元没有检测到所述部位的一部分的情况下,所述识别单元基于所述运动检测单元所检测到的所述运动来识别所述姿势。

3.根据权利要求1或2所述的姿势识别设备,其中,

所述区域检测单元根据所述识别单元所识别出的所述姿势的类型来改变要检测的区域。

4.根据权利要求3所述的姿势识别设备,其中,

所述部位检测单元从拍摄图像中检测手,以及

在所述识别单元所识别出的所述姿势是捏操作的情况下,所述区域检测单元检测拇指的尖端的区域和食指的尖端的区域。

5.根据权利要求3所述的姿势识别设备,其中,

所述部位检测单元从拍摄图像中检测手,以及

在所述识别单元所识别出的所述姿势是拖动操作的情况下,所述区域检测单元检测食指的尖端的区域。

6.根据权利要求1或2所述的姿势识别设备,其中,

所述区域检测单元根据所述部位检测单元所检测到的所述部位的位置来改变要检测的区域。

7.根据权利要求1或2所述的姿势识别设备,其中,

在所述部位检测单元检测到多个部位的情况下,所述区域检测单元检测所述多个部位中的要最频繁用于所述姿势的部位的区域。

8.根据权利要求1或2所述的姿势识别设备,还包括:

生成单元,其被配置为生成与所述识别单元所识别出的所述姿势相对应的控制信号。

9.一种头戴式显示设备,包括:

显示控制单元,其被配置为基于根据权利要求1至8中任一项所述的姿势识别设备的识别结果来控制显示。

10.一种姿势识别方法,包括:

部位检测步骤,用于从拍摄图像中检测做出姿势的部位;

识别步骤,用于基于所述部位检测步骤中所检测到的所述部位的运动来识别所述姿势;

区域检测步骤,用于从所述部位中检测要用于所述姿势的区域;以及

运动检测步骤,用于对所述区域检测步骤中所检测到的所述区域的运动进行检测,

其中，在所述识别步骤中，基于所述运动检测步骤中所检测到的所述运动来识别所述姿势。

11.一种存储程序的计算机可读介质，所述程序使得计算机用作根据权利要求1至8中任一项所述的姿势识别设备的各个单元。

姿势识别设备及方法、头戴式显示设备和计算机可读介质

技术领域

[0001] 本发明涉及用于识别姿势的技术。

背景技术

[0002] 已经提出用于识别利用用户的手或手指等的姿势并根据所识别出的姿势来进行处理的技术。利用这样的技术，使得用户能够在不触摸电子设备(电子设备的输入装置)的情况下通过做出姿势来操作该电子设备。

[0003] 在日本特开2015-172887中公开了从通过利用照相机对用户的手或手指等进行摄像而获得的图像中提取用作做出姿势的部位的手、以及指定所提取的手的形状以识别姿势。

[0004] 在日本特开2019-71048中公开了将深度学习用于识别姿势。

[0005] 然而，诸如用于AR(增强现实)的智能眼镜等的电子设备是便携式的，并且用于各种状况。例如，当由智能眼镜辅助操作时，存在操作场所是黑暗的或者戴着智能眼镜的操作者戴手套的可能性。因此，当由这样的电子设备识别姿势时，取决于电子设备的使用状况，可能无法高精度地识别姿势(姿势的识别变得不稳定)。例如，可能无法高精度地进行关节等(用于识别姿势的关键点)的检测或手等(做出姿势的部位)的形状的指定(识别)等。

发明内容

[0006] 本发明提供了能够在各种状况下高精度地识别姿势的技术。

[0007] 本发明在其第一方面提供了一种姿势识别设备，包括：部位检测单元，其被配置为从拍摄图像中检测做出姿势的部位；识别单元，其被配置为基于所述部位检测单元所检测到的所述部位的运动来识别所述姿势；区域检测单元，其被配置为从所述部位中检测要用于所述姿势的区域；以及运动检测单元，其被配置为对所述区域检测单元所检测到的所述区域的运动进行检测，其中，所述识别单元基于所述运动检测单元所检测到的所述运动来识别所述姿势。

[0008] 本发明在其第二方面提供了一种头戴式显示设备，包括：显示控制单元，其被配置为基于上述的姿势识别设备的识别结果来控制显示。

[0009] 本发明在其第三方面提供了一种姿势识别方法，包括：部位检测步骤，用于从拍摄图像中检测做出姿势的部位；识别步骤，用于基于所述部位检测步骤中所检测到的所述部位的运动来识别所述姿势；区域检测步骤，用于从所述部位中检测要用于所述姿势的区域；以及运动检测步骤，用于对所述区域检测步骤中所检测到的所述区域的运动进行检测，其中，在所述识别步骤中，基于所述运动检测步骤中所检测到的所述运动来识别所述姿势。

[0010] 本发明在其第四方面提供了一种存储程序的计算机可读介质，所述程序使得计算机用作上述的姿势识别设备的各个单元。

[0011] 通过以下参考附图对示例性实施例的描述，本发明的进一步的特征将变得明显。

附图说明

[0012] 图1是示出姿势识别设备的配置示例的框图；

[0013] 图2是示出手和关节的检测示例的图；

[0014] 图3A和图3B是用于描述根据姿势的重要区域的示例的图；

[0015] 图4是用于描述运动检测单元的检测结果的示例的图；

[0016] 图5是示出姿势识别处理的示例的流程图；以及

[0017] 图6是用于描述姿势识别处理的示例的图。

具体实施方式

[0018] 将描述本发明的实施例。图1是示出本实施例中的姿势识别设备101的配置示例的框图。姿势识别设备101具有摄像单元102、系统控制单元103、非易失性存储器104、系统存储器105和姿势识别单元110。

[0019] 摄像单元102具有包括多个透镜的光学系统和图像传感器(例如,CCD或CMOS传感器),并且被布置成面向能够拍摄用户的姿势的图像的方向。在本实施例中使用图像传感器,但是可以仅使用可用于姿势的识别的任何传感器。可以这样使用测距传感器(例如,使用诸如LiDAR(光检测和测距)等的激光的传感器)。

[0020] 非易失性存储器104是电可擦除/可记录存储器,并且例如是Flash-ROM等。在非易失性存储器104中,存储系统控制单元103的操作常数或程序等。这里,程序表示例如用于进行将在本实施例中稍后描述的流程图的各个处理的程序。

[0021] 系统存储器105例如是RAM等,并且展开系统控制单元103的操作变量或从非易失性存储器104中读取的常数或程序等。

[0022] 系统控制单元103是由至少一个处理器或电路构成的控制单元,并控制整个姿势识别设备101。系统控制单元103通过运行上述非易失性存储器104中所存储的程序来实现稍后将描述的流程图的各个处理。系统控制单元103通过将诸如降噪和调整大小等的图像处理应用于从摄像单元102中获得的拍摄图像来生成检测图像,并将该检测图像记录在系统存储器105上。此外,系统控制单元103根据由姿势识别单元110识别出的姿势来进行处理。例如,系统控制单元103生成与所识别出的姿势相对应的控制信号,并使用该控制信号来控制姿势识别设备101的各个单元。

[0023] 姿势识别单元110具有对象部位检测单元111、重要区域检测单元112、运动检测单元113和姿势检测单元114,并且基于从摄像单元102中获得的拍摄图像(具体是上述检测图像)来识别姿势。

[0024] 将使用图2至图4详细描述姿势识别单元110。图2是示出本实施例中的手和关节的检测示例的图。图3A和图3B是用于描述根据本实施例中的姿势的重要区域的示例的图。图4是用于描述本实施例中的运动检测单元113的检测结果的示例的图。

[0025] 图2中所示的图像200是通过将图像处理应用于从摄像单元102中获得的拍摄图像而获得的检测图像。在检测图像200中,反映了做出姿势(姿势操作)的用户的左手211和右手212。图2中所示的图像201是通过将对象部位检测单元111的检测结果叠加在检测图像200上而获得的图像。

[0026] 对象部位检测单元111从检测图像200中检测(提取)作为做出姿势的部位(对象部

位)的手和关节。在图2中,针对左手211和右手212中的各个,检测包括手位置(手的位置)和从手腕到各个指尖的21个点的关节位置(关节的位置)的总共22个点的位置。手位置例如是手的中心位置。对象部位检测单元111基于21个点的关节位置来获得关节信息。针对左手211获得手位置221和关节信息231,并且针对右手212获得手位置222和关节信息232。关节信息231和232示出了21个点的关节位置、以及用于将21个点的关节位置连接在一起以形成骨骼框架的多个线段。注意,在本实施例中,用户利用手或手指做出姿势,但是也可以利用他/她的手臂、脚、眼睛或嘴等做出姿势。

[0027] 重要区域检测单元112从由对象部位检测单元111检测到的对象部位(手和关节)中检测作为要(主要)用于姿势的区域的重要区域。例如,重要区域检测单元112从由对象部位检测单元111检测到的多个对象部位中,选择作为要(主要)用于姿势的部位的重要部位,并将所选择的重要部位的区域确定为重要区域。重要区域根据姿势的类型而不同。图3A和图3B示出了姿势的示例。图3A示出了捏操作(利用拇指和食指的捏操作),并且图3B示出了利用食指的拖动操作(移动操作)。捏操作主要通过拇指和食指的尖端的移动来进行。因此,重要区域检测单元112在捏操作期间将拇指和食指的尖端的区域检测为重要区域。拖动操作(利用食指的拖动操作)主要通过食指的尖端的移动来进行。因此,重要区域检测单元112在拖动操作期间将食指的尖端的区域检测为重要区域。注意,姿势不限于捏操作或拖动操作。通过预先设置与姿势和重要区域之间的对应关系有关的信息并将该信息存储在非易失性存储器104中,重要区域检测单元112能够使用与对应关系有关的信息,来检测与所做出的姿势相对应的重要区域。

[0028] 运动检测单元113基于重要区域检测单元112的检测结果,检测由重要区域检测单元112检测到的重要区域的运动。在图4中,在过去,手411的食指的尖端的区域421被检测为重要区域。例如,运动检测单元113使用诸如块匹配等的各种方法从当前检测图像中检测与重要区域421相对应的移动区域422,以检测从重要区域421到移动区域422的运动430。作为移动区域422,检测(与手411相同的)手412的食指的尖端的区域。运动检测单元113使得能够通过计算重要区域中的特征点(例如,关节位置)的运动向量来检测移动区域或重要区域的运动。

[0029] 姿势检测单元114基于由对象部位检测单元111检测到的对象区域(手和关节)的运动来检测(识别)姿势。在本实施例中,姿势检测单元114使用对象部位检测单元111的检测结果(手位置和关节信息)或运动检测单元113的检测结果(重要区域的运动)来检测姿势。姿势检测单元114通过例如将对象部位检测单元111或运动检测单元113的检测结果与预先存储在非易失性存储器104中的姿势模型进行比较来检测姿势。注意,用于检测姿势的方法不限于此,并且可以使用已经通过深度学习等学习了的检测器来检测姿势。在使用深度学习的情况下,可以通过使用RNN(递归神经网络),从关节信息的时间序列数据中检测到不能从一帧的数据中检测到的姿势(诸如用手指画圆的姿势等)。姿势检测单元114可检测的姿势不限于一个,而是姿势检测单元114可以检测多个姿势。

[0030] 将描述本实施例中的姿势识别处理。图5是示出本实施例中的姿势识别处理的示例的流程图。当系统控制单元103将非易失性存储器104中所存储的程序展开到系统存储器105中并运行该程序时,实现该处理。例如,当姿势识别设备101启动时,图5的处理开始。图6是用于描述本实施例中的姿势识别处理的示例的图。在图6中,检测图像601至603按照时间

顺序布置。检测图像601是最旧的，并且检测图像603是最新的。手611被反映在检测图像601中、手621被反映在检测图像602中、以及手631被反映在检测图像603中。手611、621和631是相同的。

[0031] 在步骤S501中，系统控制单元103从摄像单元102中获取拍摄图像，并将图像处理应用于拍摄图像，以生成(获取)检测图像601。

[0032] 在步骤S502中，系统控制单元103使用对象部位检测单元111从步骤S501中所生成的检测图像601中获取手位置612和关节信息613。

[0033] 在步骤S503中，系统控制单元103判断是否已经在步骤S502中成功获取手位置612和关节信息613。当已经成功获取手位置612和关节信息613时，处理进入到步骤S504。否则，处理进入到步骤S506。当尚未检测到对象部位的一部分(例如，关节的一部分)时，获得不完全的关节信息。因此，处理进入到步骤S506。这里，假设已经成功获取手位置612和关节信息613，并且处理进入到步骤S504。

[0034] 在步骤S504中，系统控制单元103使用姿势检测单元114从步骤S502中所获取的手位置612和关节信息613中检测(识别)姿势。注意，系统控制单元103可以仅使用从一个检测图像获取的手位置和关节信息来检测姿势，或者使用从直到目前为止所获取的多个检测图像中的各个检测图像中获取的手位置和关节信息来检测姿势。

[0035] 在步骤S505中，系统控制单元103使用重要区域检测单元112从手611中检测重要区域614和615，并且将重要区域614和615的数据存储在系统存储器105中。系统存储器105中所存储的重要区域的数据可以是从检测图像中切出的重要区域的图像的数据，或者是重要区域的图像的特征(特征点)的数据。在步骤S505中，系统控制单元103检测与步骤S504中所检测到的姿势相对应的重要区域。当在步骤S504中未能检测到姿势时，系统控制单元103检测与在手位置612处能够做出的姿势相对应的重要区域。因此，所检测到的重要区域根据手位置而不同。当在手位置612处能够做出多个姿势时，系统控制单元103可以将手611的区域检测为重要部位，或者可以将多个对象部位(例如，多个关节)中的要最频繁用于姿势的对象部位的区域检测为重要部位。

[0036] 在步骤S510中，系统控制单元103判断是否结束姿势识别处理。当系统控制单元103结束姿势识别处理时，处理进入到步骤S511。在步骤S511中，系统控制单元103删除系统存储器105中所存储的重要区域的数据。然后，系统控制单元103结束姿势识别处理。当系统控制单元103没有结束姿势识别处理时，处理返回到步骤S501。这里，假设处理返回到步骤S501。系统控制单元103在步骤S501中获取新的检测图像602，并在步骤S502中获取手位置622和关节信息623。然后，处理进入到步骤S503。

[0037] 在步骤S503中，系统控制单元103判断是否已经在步骤S502中成功获取手位置622和关节信息623。当系统控制单元103已经成功获取手位置622和关节信息623时，处理进入到步骤S504。否则，处理进入到步骤S506。这里，假设尚未成功获取信息(手位置622和关节信息623)的一部分，并且处理进入到步骤S506。

[0038] 在步骤S506中，系统控制单元103判断重要区域的数据是否被存储在系统存储器105中。当重要区域的数据被存储时，处理进入到步骤S507。否则，处理返回到步骤S501。这里，假设重要区域614和615的数据被存储，并且处理进入到步骤S507。

[0039] 在步骤S507中，系统控制单元103使用步骤S506中所获取的重要区域614和615的

数据，从检测图像602中检测分别与重要区域614和615相对应的两个移动区域，并且检测重要区域614和615的运动。系统控制单元103可以从检测图像602的整个区域中检测移动区域，但是可以基于重要区域614和615等的位置，将检测到移动区域的区域限制为作为检测图像602的一部分的特定区域626等。

[0040] 在步骤S508中，系统控制单元103判断是否已经在步骤S507中成功检测到重要区域614和615的运动。当已经成功检测到重要区域614和615的运动时，处理进入到步骤S509。否则，处理返回到步骤S501。如果与重要区域614和615相对应的部分位于检测图像602之外或者隐藏在其他被摄体的后面，则没有成功检测到重要区域614和615的运动。在该情况下，处理返回到步骤S501。这里，假设已经成功检测到重要区域614和615的运动，并且处理进入到步骤S509。

[0041] 在步骤S509中，系统控制单元103使用姿势检测单元114从步骤S507中所检测到的重要区域614和615的运动中检测姿势。此时，系统控制单元103可以附加地使用在过去获取的手位置612和关节信息613。

[0042] 在步骤S510中，系统控制单元103判断是否结束姿势识别处理。当系统控制单元103结束姿势识别处理时，处理进入到步骤S511。否则，处理返回到步骤S501。这里，假设处理返回到步骤S501。系统控制单元103在步骤S501中获取新的检测图像603，并且在步骤S502中获取手位置632和关节信息633。然后，处理进入到步骤S503。

[0043] 如上所述，根据本实施例，基于做出姿势的部位中的要用于姿势的区域的运动来识别姿势。以该方式，可以在各种环境下高精度地识别姿势。

[0044] 注意，上面的实施例仅为示例，并且通过在本发明的要点的范围内对上述配置进行适当变形或修改而获得的配置也包括在本发明中。例如，姿势识别设备101可以被设置在诸如用于AR(增强现实)的智能眼镜等的电子设备(头戴式显示设备)中。在该情况下，设置在头戴式显示设备中的显示控制单元基于姿势识别设备101的识别结果来控制显示。姿势识别设备101(姿势识别单元110)可以是与头戴式显示设备分开的电子设备，并且可以例如是云上的计算机(服务器)。

[0045] 根据本发明，可在各种环境下高精度地识别姿势。

[0046] 其他实施例

[0047] 本发明的实施例还可以通过如下的方法来实现，即，通过网络或者各种存储介质将执行上述实施例的功能的软件(程序)提供给系统或装置，该系统或装置的计算机或是中央处理单元(CPU)、微处理单元(MPU)读出并执行程序的方法。

[0048] 尽管已经参考示例性实施例描述了本发明，但是应当理解，本发明不限于所公开的示例性实施例。所附权利要求的范围应符合最广泛的解释，以涵盖所有这样的修改、等同结构和功能。

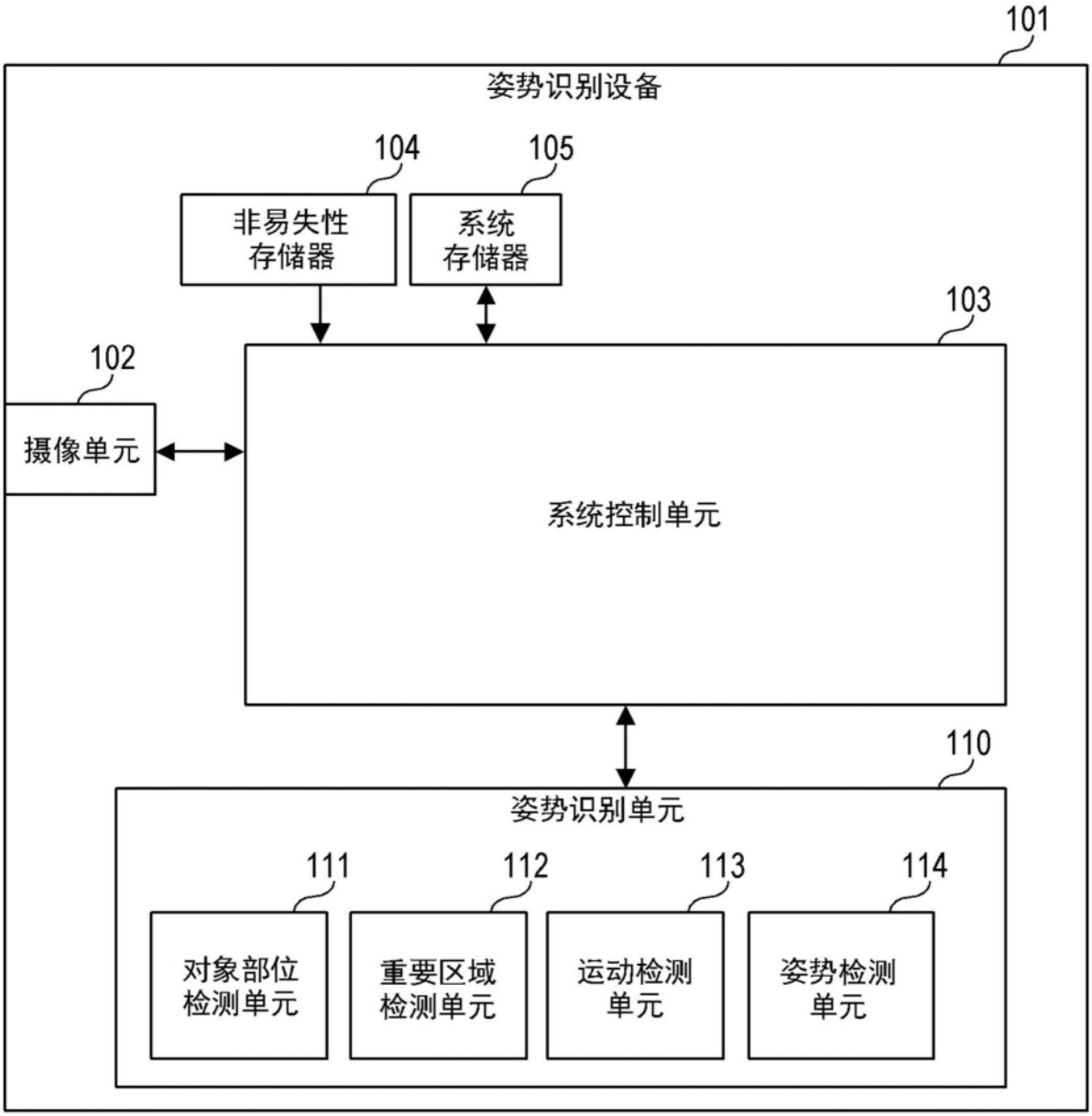
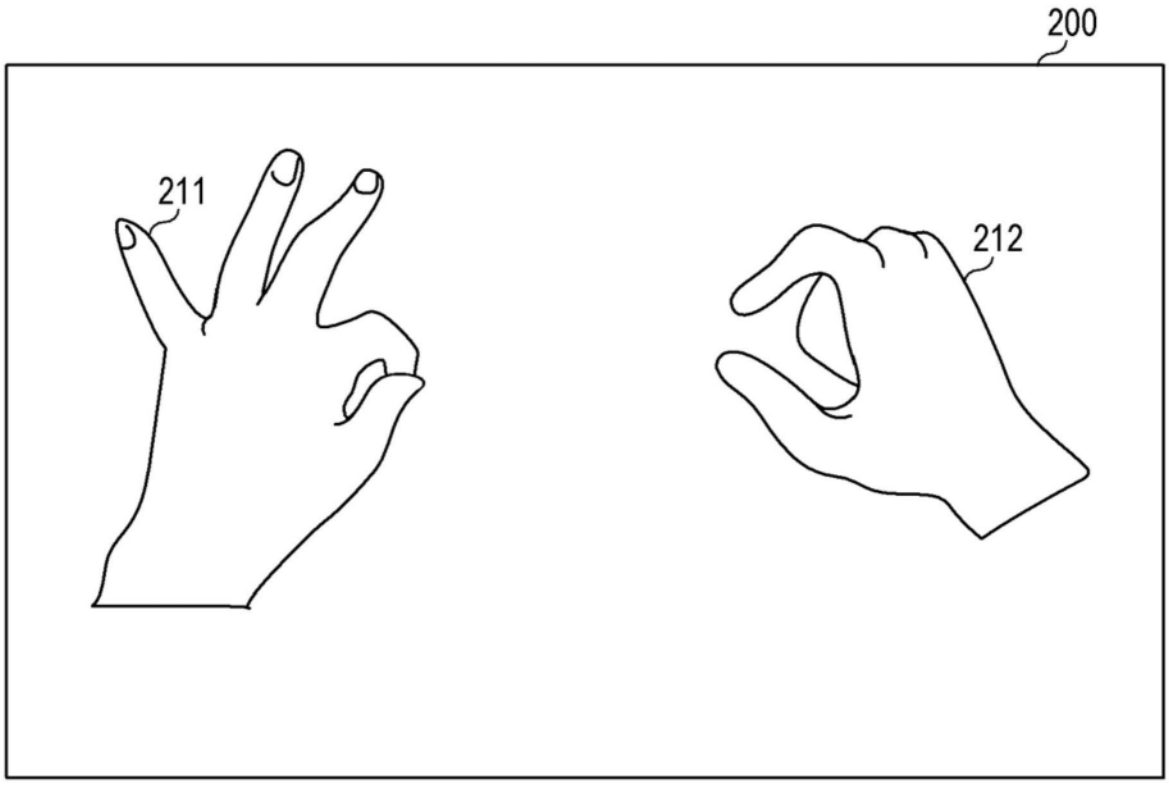

101
姿势识别设备
104
105
非易失性
存储器
系统
存储器
103
102
摄像单元
系统控制单元
110
姿势识别单元
111
112
113
114
对象部位
检测单元
重要区域
检测单元
运动检测
单元
姿势检测
单元

图1



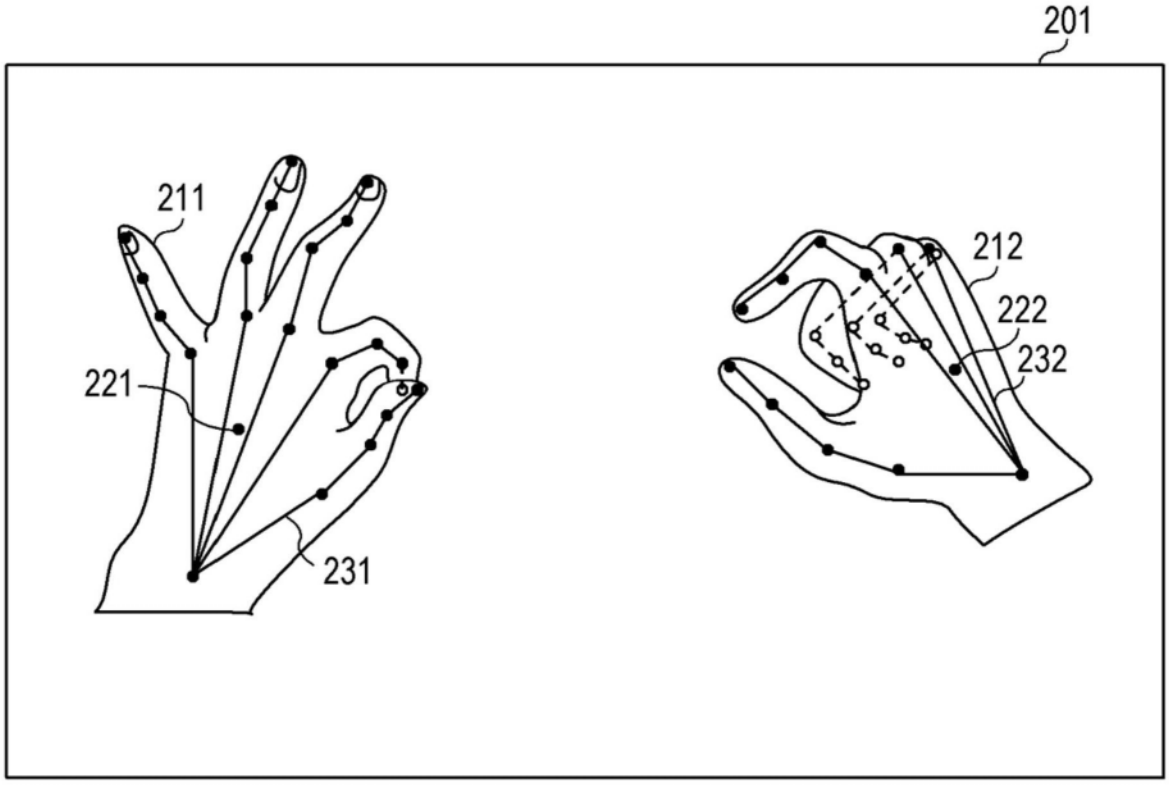


图2

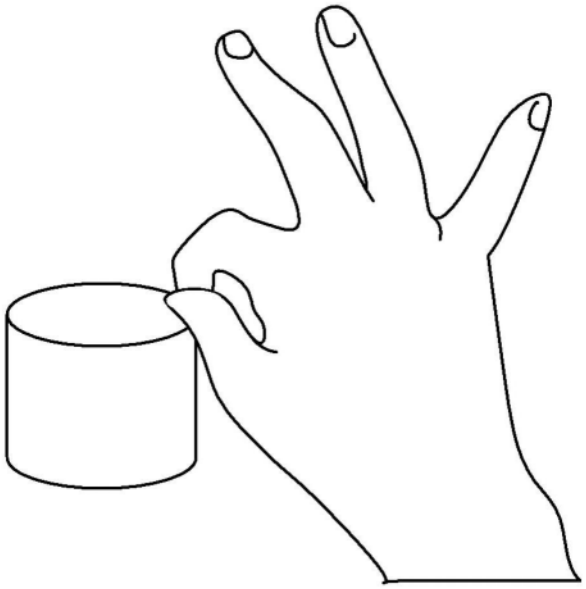


图3A

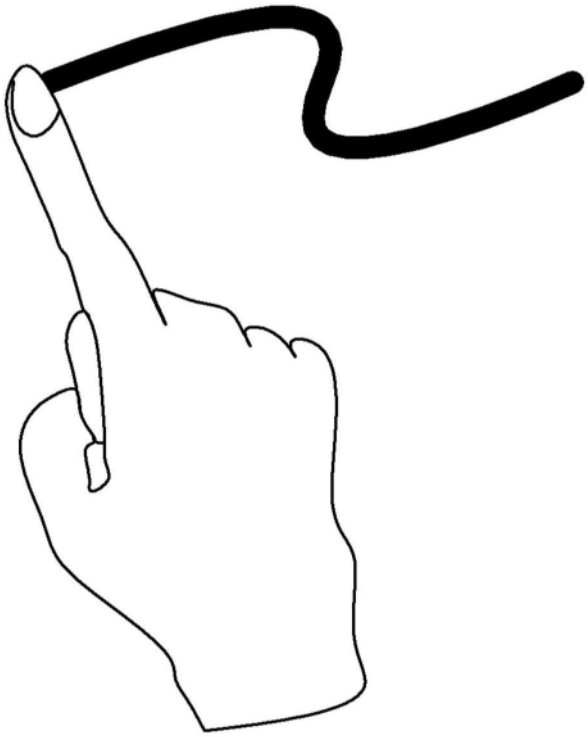


图3B

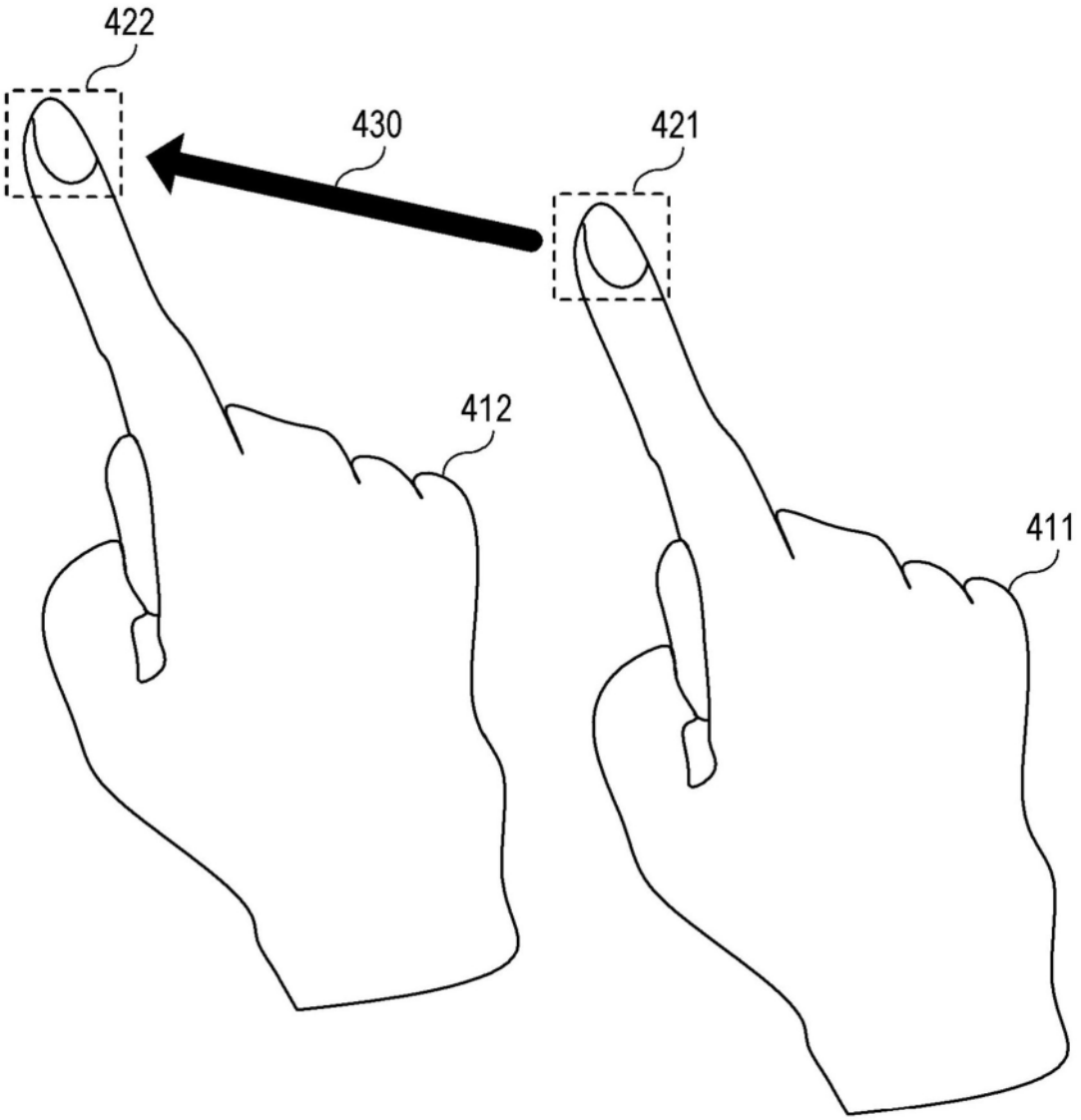


图4

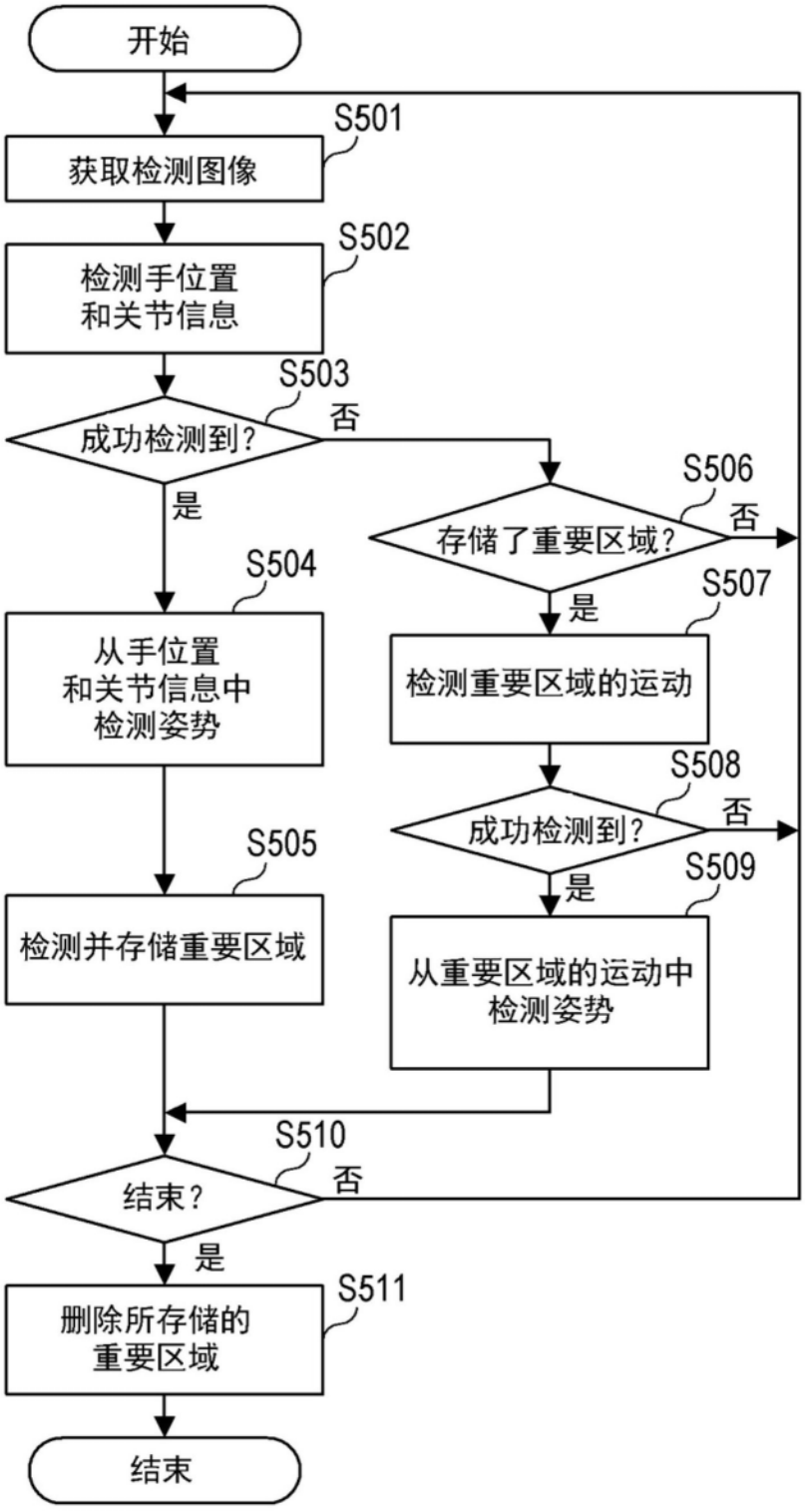

开始
S501
获取检测图像
S502
检测手位置
和关节信息
S503
成功检测到?
否
是
S504
从手位置
和关节信息中
检测姿势
S505
检测并存储重要区域
S506
存储了重要区域?
否
是
S507
检测重要区域的运动
S508
成功检测到?
否
是
S509
从重要区域的运动中
检测姿势
S510
结束?
否
是
S511
删除所存储的
重要区域
结束

图5

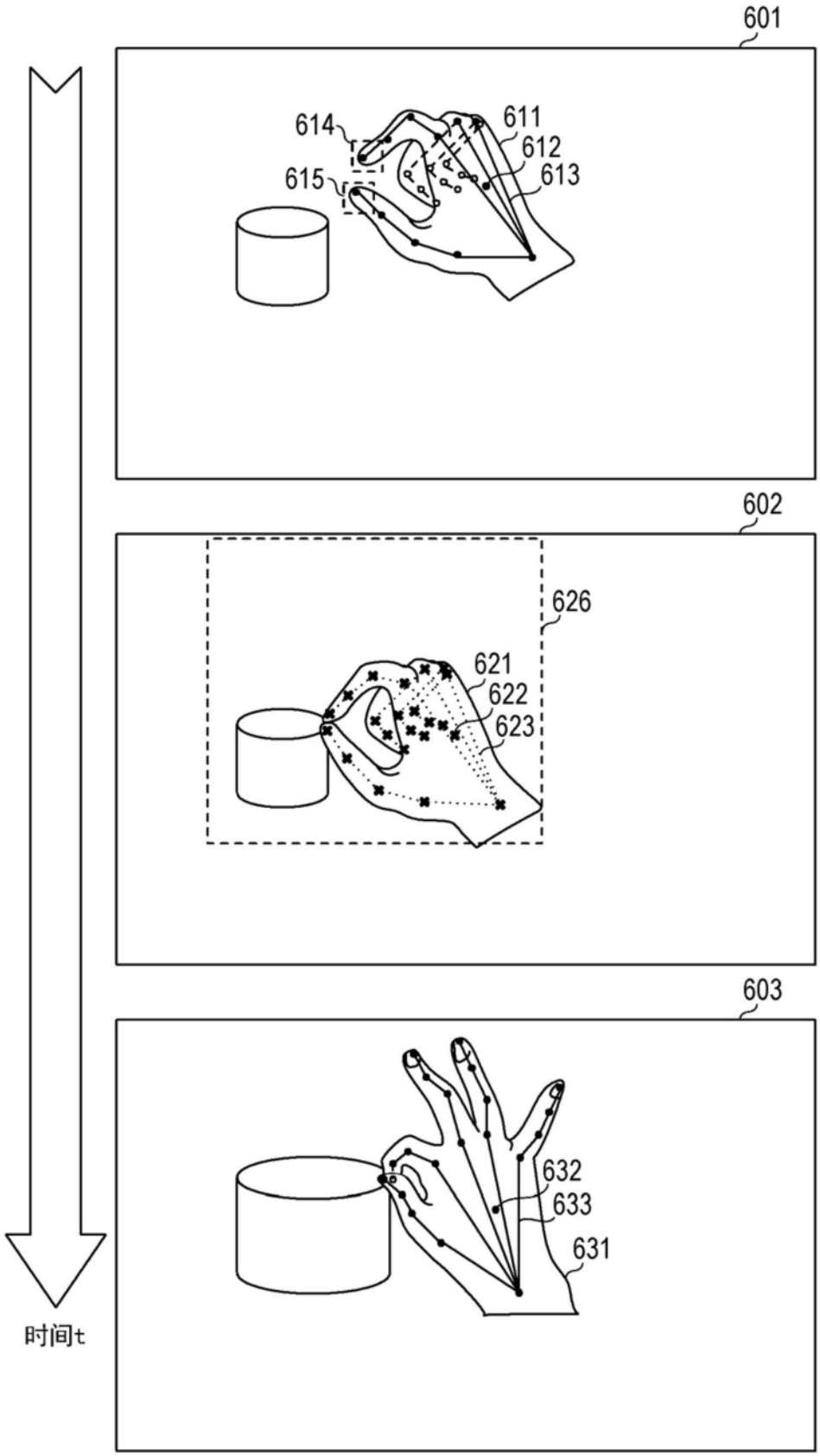


图6