



(21) 申请号 202010582730.2

G06F 3/01 (2006.01)

(22) 申请日 2020.06.23

(56) 对比文件

(65) 同一申请的已公布的文献号

CN 108921101 A, 2018.11.30

申请公布号 CN 113473198 A

CN 109991859 A, 2019.07.09

(43) 申请公布日 2021.10.01

CN 110688914 A, 2020.01.14

(73) 专利权人 青岛海信电子产业控股股份有限公司

US 2012069168 A1, 2012.03.22

US 2012119985 A1, 2012.05.17

地址 266555 山东省青岛市经济技术开发区前湾港路218号

US 2013222232 A1, 2013.08.29

US 2015293595 A1, 2015.10.15

(72) 发明人 陈维强 冯谨强 孟祥奇 高雪松

US 2017285760 A1, 2017.10.05

US 2020012351 A1, 2020.01.09

(74) 专利代理机构 北京同达信恒知识产权代理有限公司 11291

CN 110716648 A, 2020.01.21

CN 102801409 A, 2012.11.28

专利代理师 宋正伟

CN 106354252 A, 2017.01.25

US 2011111384 A1, 2011.05.12

US 2019369742 A1, 2019.12.05

(51) Int. Cl.

审查员 姜孜铨

H04N 21/422 (2011.01)

权利要求书2页 说明书14页 附图8页

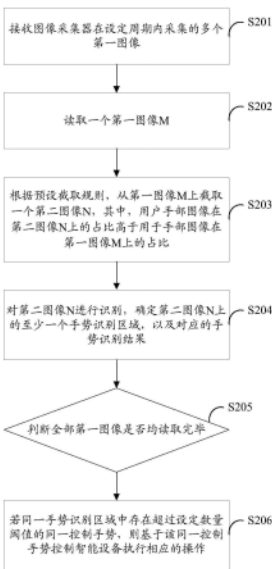
H04N 21/472 (2011.01)

(54) 发明名称

一种智能设备的控制方法及智能设备

(57) 摘要

本申请涉及智能设备技术领域,提供一种智能设备的控制方法及智能设备,该方法包括:接收图像采集器在设定周期内采集的多个第一图像,每读取一个第一图像,按照预设截取规则从所述一个第一图像上截取一个第二图像,并对所述第二图像进行识别,确定至少一个手势识别区域以及对应的手势识别结果;若同一手势识别区域中存在超过设定数量阈值的同一控制手势,则基于该控制手势控制智能设备执行相应的操作。按照预设截取规则对第一图像进行预处理,以降低手部检测难度,采集用户的静态手势不仅可以缩短手势识别花费的时间,还可以提高用户体验感。



1. 一种智能设备,其特征在于,包括:

显示器,被配置为显示画面;

图像采集器,被配置为在设定周期采集多个第一图像,并将所述多个第一图像传输至控制器中;

所述控制器,被配置为接收所述图像采集器在设定周期内采集的所述多个第一图像;

针对各个第一图像分别执行以下处理,每读取一个第一图像,根据预设截取规则,从所述一个第一图像上截取一个第二图像,其中,用户手部图像在所述一个第二图像上的占比高于所述用户手部图像在所述一个第一图像上的占比;对所述一个第二图像进行静态手势识别,确定所述一个第二图像上的多个手势识别区域,以及至少一个手势识别区域中每个手势识别区域对应的手势识别结果;其中,所述静态手势识别用于对一幅图像进行手势识别操作,输出手势识别结果;所述手势识别区域对应的手势识别结果是基于所述手势识别区域对应的图像确定的,所述手势识别区域为手部所在的区域;

基于每个第一图像截取的所述第二图像的至少一个手势识别区域以及各手势识别区域对应的手势识别结果,若同一手势识别区域中存在连续M个唤醒手势,则将所述同一手势识别区域确定为手势控制区域;以及若所述手势控制区域在所述连续M个唤醒手势之后,存在连续N个第一控制手势,则基于所述第一控制手势控制所述智能设备执行相应的操作;其中,M、N均为正整数。

2. 如权利要求1所述的智能设备,其特征在于,所述控制器被配置为:

根据所述图像采集器的视场角、设定视场角阈值,以及所述一个第一图像的尺寸信息,确定所述手部图像的尺寸信息和角点坐标信息;

根据所述手部图像的尺寸信息和所述角点坐标信息,从所述一个第一图像中截取出所述一个第二图像。

3. 如权利要求2所述的智能设备,其特征在于,所述控制器被进一步配置为:

若所述图像采集器在X方向上的视场角,与所述图像采集器在Y方向上的视场角均低于所述视场角阈值时,则将所述一个第一图像确定为所述一个第二图像。

4. 如权利要求1所述的智能设备,其特征在于,所述控制器被配置为:

若所述手势控制区域在所述连续N个第一控制手势之后,存在连续N个第二控制手势,则基于所述第二控制手势控制所述智能设备执行相应的操作;其中,N为正整数。

5. 一种控制智能设备的方法,其特征在于,包括:

接收图像采集器在设定周期内采集的多个第一图像;

针对各个第一图像分别执行以下处理,每读取一个第一图像,根据预设截取规则,从所述一个第一图像上截取一个第二图像,其中,用户手部图像在所述一个第二图像上的占比高于所述用户手部图像在所述一个第一图像上的占比;对所述一个第二图像进行静态手势识别,确定所述一个第二图像上的多个手势识别区域,以及至少一个手势识别区域中每个手势识别区域对应的手势识别结果;其中,所述静态手势识别用于对一幅图像进行手势识别操作,输出手势识别结果;所述手势识别区域对应的手势识别结果是基于所述手势识别区域对应的图像确定的,所述手势识别区域为手部所在的区域;

基于每个第一图像截取的所述第二图像的至少一个手势识别区域以及各手势识别区域对应的手势识别结果,若同一手势识别区域中存在连续M个唤醒手势,则将所述同一手势识别

区域确定为手势控制区域;以及若所述手势控制区域在所述连续M个唤醒手势之后,存在连续N个第一控制手势,则基于所述第一控制手势控制所述智能设备执行相应的操作;其中,M、N均为正整数。

6.如权利要求5所述的方法,其特征在于,根据预设截取规则,从所述一个第一图像上截取一个第二图像,包括:

根据所述图像采集器的视场角、设定视场角阈值,以及所述一个第一图像的尺寸信息,确定所述手部图像的尺寸信息和角点坐标信息;

根据所述手部图像的尺寸信息和所述角点坐标信息,从所述一个第一图像中截取出所述一个第二图像。

7.如权利要求6所述的方法,其特征在于,在确定所述手部图像的尺寸信息和角点坐标信息之前,进一步包括:

若所述图像采集器在X方向上的视场角,与所述图像采集器在Y方向上的视场角均低于所述视场角阈值时,则将所述一个第一图像确定为所述一个第二图像。

8.如权利要求5所述的方法,其特征在于,在基于所述第一控制手势控制所述智能设备执行相应的操作之后,进一步包括:

若所述手势控制区域在所述连续N个第一控制手势之后,存在连续N个第二控制手势,则基于所述第二控制手势控制所述智能设备执行相应的操作;其中,N为正整数。

一种智能设备的控制方法及智能设备

技术领域

[0001] 本申请涉及智能设备技术领域,提供了一种智能设备的控制方法及智能设备。

背景技术

[0002] 随着产品设备的智能化升级,用户用手做不同的连续动作(即动态手势)操作屏幕,实现调节音量、换选频道、快进快退等功能,使得用户控制智能设备的操作越来越简单。

[0003] 但采用上述方式控制智能设备时,会产生以下问题:用户距离摄像头越远,在采集到的图像中用户的手部图像占比越小,导致手部检测难度大幅度提升;在用户做动态手势时,摄像头会采集到连续的多个图像,并将连续的多个图像输入到神经网络中进行识别,由于输入的图像较多,神经网络的识别过程会花费较长时间,无法满足实时响应需求;最后,用户需要做动态手势才能满足响应,用户体验感较差。

[0004] 有鉴于此,本申请实施例提供了一种新的智能设备的控制方法及智能设备。

发明内容

[0005] 本申请实施例提供了一种智能设备的控制方法及智能设备,用以降低手部检测难度,缩短手势识别花费的时间,提高用户体验感。

[0006] 第一方面,本申请实施例提供了一种智能设备,包括:

[0007] 显示器,被配置为显示画面;

[0008] 图像采集器,被配置为在设定周期采集多个第一图像,并将所述多个第一图像传输至控制器中;

[0009] 所述控制器,被配置为接收所述图像采集器在设定周期内采集的所述多个第一图像;

[0010] 针对各个第一图像分别执行以下处理,每读取一个第一图像,根据预设截取规则,从所述一个第一图像上截取一个第二图像,其中,用户手部图像在所述一个第二图像上的占比高于所述用户手部图像在所述一个第一图像上的占比;对所述一个第二图像进行识别,确定所述一个第二图像上的至少一个手势识别区域,以及对应的手势识别结果;

[0011] 若同一手势识别区域中存在超过设定数量阈值的同一控制手势,则基于所述同一控制手势控制智能设备执行相应的操作。

[0012] 可选的,所述控制器被配置为:

[0013] 根据所述图像采集器的视场角、设定视场角阈值,以及所述一个第一图像的尺寸信息,确定所述手部图像的尺寸信息和角点坐标信息;

[0014] 根据所述手部图像的尺寸信息和所述角点坐标信息,从所述一个第一图像中截取所述一个第二图像。

[0015] 可选的,所述控制器被进一步配置为:

[0016] 若所述图像采集器在X方向上的视场角,与所述图像采集器在Y方向上的视场角均低于所述视场角阈值时,则将所述一个第一图像确定为所述一个第二图像。

[0017] 可选的,所述控制器被配置为:

[0018] 若同一手势识别区域中存在连续M个唤醒手势,则将所述同一手势识别区域确定为手势控制区域;

[0019] 若所述手势控制区域在所述连续M个唤醒手势之后,存在连续N个第一控制手势,则基于所述第一控制手势控制所述智能设备执行相应的操作;

[0020] 其中,M、N均为正整数。

[0021] 可选的,所述控制器被配置为:

[0022] 若所述手势控制区域在所述连续N个第一控制手势之后,存在连续N个第二控制手势,则基于所述第二控制手势控制所述智能设备执行相应的操作;其中,N为正整数。

[0023] 第二方面,本申请实施例还提供了一种控制智能设备的方法,包括:

[0024] 接收图像采集器在设定周期内采集的多个第一图像;

[0025] 针对各个第一图像分别执行以下处理,每读取一个第一图像,根据预设截取规则,从所述一个第一图像上截取一个第二图像,其中,用户手部图像在所述一个第二图像上的占比高于所述用户手部图像在所述一个第一图像上的占比;对所述一个第二图像进行识别,确定所述一个第二图像上的至少一个手势识别区域,以及对应的手势识别结果;

[0026] 若同一手势识别区域中存在超过设定数量阈值的同一控制手势,则基于所述同一控制手势控制智能设备执行相应的操作。

[0027] 可选的,根据预设截取规则,从所述一个第一图像上截取一个第二图像,包括:

[0028] 根据所述图像采集器的视场角、设定视场角阈值,以及所述一个第一图像的尺寸信息,确定所述手部图像的尺寸信息和角点坐标信息;

[0029] 根据所述手部图像的尺寸信息和所述角点坐标信息,从所述一个第一图像中截取出所述一个第二图像。

[0030] 可选的,在确定所述手部图像的尺寸信息和角点坐标信息之前,进一步包括:

[0031] 若所述图像采集器在X方向上的视场角,与所述图像采集器在Y方向上的视场角均低于所述视场角阈值时,则将所述一个第一图像确定为所述一个第二图像。

[0032] 可选的,若同一手势识别区域中存在超过设定数量阈值的同一控制手势,则基于所述同一控制手势控制智能设备执行相应的操作,包括:

[0033] 若同一手势识别区域中存在连续M个唤醒手势,则将所述同一手势识别区域确定为手势控制区域;

[0034] 若所述手势控制区域在所述连续M个唤醒手势之后,存在连续N个第一控制手势,则基于所述第一控制手势控制所述智能设备执行相应的操作;

[0035] 其中,M、N均为正整数。

[0036] 可选的,在基于所述第一控制手势控制所述智能设备执行相应的操作之后,进一步包括:

[0037] 若所述手势控制区域在所述连续N个第一控制手势之后,存在连续N个第二控制手势,则基于所述第二控制手势控制所述智能设备执行相应的操作;其中,N为正整数。

[0038] 第三方面,本申请实施例还提供了一种计算机可读存储介质,其包括程序代码,当所述程序产品在终端上运行时,所述程序代码用于使所述终端执行上述任意一种智能设备的控制方法的步骤。

[0039] 本申请有益效果如下：

[0040] 本申请提供了一种智能设备的控制方法及智能设备，该方法包括：接收图像采集器在设定周期内采集的多个第一图像，每读取一个第一图像，按照预设截取规则从所述一个第一图像上截取一个第二图像，并对所述第二图像进行识别，确定至少一个手势识别区域以及对应的手势识别结果；若同一手势识别区域中存在超过设定数量阈值的同一控制手势，则基于该控制手势控制智能设备执行相应的操作。按照预设截取规则对第一图像进行预处理，以降低手部检测难度，采集用户的静态手势不仅可以缩短手势识别花费的时间，还可以提高用户体验感。

[0041] 本申请的其它特征和优点将在随后的说明书中阐述，并且，部分地从说明书中变得显而易见，或者通过实施本申请而了解。本申请的目的和其他优点可通过在所写的说明书、权利要求书、以及附图中所特别指出的结构来实现和获得。

附图说明

[0042] 此处所说明的附图用来提供对本申请的进一步理解，构成本申请的一部分，本申请的示意性实施例及其说明用于解释本申请，并不构成对本申请的不当限定。在附图中：

[0043] 图1a中示例性示出了智能设备与控制装置之间操作场景的示意图；

[0044] 图1b中示例性示出了图1a中控制装置100的配置框图；

[0045] 图1c中示例性示出了图1a中智能设备200的配置框图；

[0046] 图1d中示例性示出了智能设备200存储器中操作系统的架构配置框图；

[0047] 图2示例性示出了控制智能设备的流程示意图；

[0048] 图3a示例性示出了视场角示意图；

[0049] 图3b示例性示出了X方向的视场角的俯视图；

[0050] 图3c示例性示出了Y方向的视场角的侧视图；

[0051] 图4a示例性示出了唤醒手势示意图；

[0052] 图4b示例性示出了调高音量的控制手势示意图；

[0053] 图4c示例性示出了调低音量的控制手势示意图；

[0054] 图4d示例性示出了视频后退的控制手势示意图；

[0055] 图4e示例性示出了视频快进的控制手势示意图；

[0056] 图4f示例性示出了确定的控制手势示意图；

[0057] 图4g示例性示出了取消的控制手势示意图；

[0058] 图4h示例性示出了静音/结束的控制手势示意图。

具体实施方式

[0059] 为使本申请实施例的目的、技术方案和优点更加清楚，下面将结合本申请实施例中的附图，对本申请的技术方案进行清楚、完整地描述，显然，所描述的实施例是本申请技术方案的一部分实施例，而不是全部的实施例。基于本申请文件中记载的实施例，本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例，都属于本申请技术方案保护的范围。

[0060] 图1a中示例性示出了智能设备与控制装置之间操作场景的示意图。如图1a所示，

控制装置100和智能设备200之间可以有线或无线方式进行通信。

[0061] 其中,控制装置100被配置为控制智能设备200,其可接收用户输入的操作指令,且将操作指令转换为智能设备200可识别和响应的指令,起着用户与智能设备200之间交互的中介作用。如:用户通过操作控制装置100上频道加减键,智能设备200响应频道加减的操作。

[0062] 控制装置100可以是遥控器100A,包括红外协议通信或蓝牙协议通信,及其他短距离通信方式等,通过无线或其他有线方式来控制智能设备200。用户可以通过遥控器上按键、语音输入、控制面板输入等输入用户指令,来控制智能设备200。如:用户可以通过遥控器上音量加减键、频道控制键、上/下/左/右的移动按键、语音输入按键、菜单键、开关机按键等输入相应控制指令,来实现控制智能设备200的功能。

[0063] 控制装置100也可以是移动终端100B、平板电脑、计算机、笔记本电脑等。例如,使用在移动终端100B上运行的应用程序控制智能设备200。该应用程序通过配置可以在与移动终端100B关联的屏幕上,通过直观的用户界面(UI)为用户提供各种控制。

[0064] 本申请实施例中的用户界面,是应用程序或操作系统与用户之间进行交互和信息交换的介质接口,它实现信息的内部形式与用户可以接受形式之间的转换。用户界面常用的表现形式是图形用户界面(graphic user interface, GUI),是指采用图形方式显示的与计算机操作相关的用户界面。它可以是在智能设备200的显示屏中显示的一个图标、窗口、控件等界面元素,其中控件可以包括图标、按钮、菜单、选项卡、文本框、对话框、状态栏、导航栏、Widget等可视的界面元素。

[0065] 示例性的,移动终端100B可与智能设备200安装软件应用,通过网络通信协议实现连接通信,实现一对一控制操作的和数据通信的目的。如:可以使移动终端100B与智能设备200建立控制指令协议,通过操作移动终端100B上提供的用户界面的各种功能键或虚拟按钮,来实现如遥控器100A布置的实体按键的功能。也可以将移动终端100B上显示的音视频内容传输到智能设备200的显示器上,实现同步显示功能。

[0066] 在其他一些示例性实施例中,智能设备200还可以调用内部配置的图像采集器,如相机、摄像头等,用于采集外部环境场景,以自适应变化智能设备200的显示参数;以及用于在设定周期采集多个第一图像,所述第一图像中可能包含用户的属性或用户的交互手势,以实现智能设备与用户之间互动的功能。

[0067] 智能设备200可提供广播接收功能和计算机支持功能的网络电视功能。智能设备可以是数字电视、网络电视、互联网协议电视(IPTV)等。智能设备200的显示器可以是液晶显示器、有机发光显示器、投影设备。具体显示器类型、尺寸大小和分辨率等不作限定。

[0068] 智能设备200还与服务器300通过多种通信方式进行数据通信。这里可允许智能设备200通过局域网(LAN)、无线局域网(WLAN)和其他网络进行通信连接。服务器300可以向智能设备200提供各种内容和互动。示例的,智能设备200可以发送和接收信息,例如:接收电子节目指南(EPG)数据、接收软件程序更新、或访问远程储存的数字媒体库。服务器300可以一组,也可以多组,可以一类或多类服务器。通过服务器300提供视频点播和广告服务等其他网络服务内容。

[0069] 图1b中示例性示出了控制装置100的配置框图。如图1b所示,控制装置100包括控制器110、存储器120、通信器130、用户输入接口140、输出接口150、供电电源160。

[0070] 控制器110包括随机存取存储器(RAM)111、只读存储器(ROM)112、处理器113、通信接口以及通信总线。控制器110用于控制控制装置100的运行和操作,以及内部各部件之间的通信协作、外部和内部的数据处理功能。

[0071] 示例性的,当检测到用户按压在遥控器100A上布置的按键的交互或触摸在遥控器100A上布置的触摸面板的交互时,控制器110可控制产生与检测到的交互相应的信号,并将该信号发送到智能设备200。

[0072] 存储器120,用于在控制器110的控制下存储驱动和控制控制装置100的各种运行程序、数据和应用。存储器120,可以存储用户输入的各类控制信号指令。

[0073] 通信器130在控制器110的控制下,实现与智能设备200之间控制信号和数据信号的通信。如:控制装置100经由通信器130将控制信号(例如触摸信号或按钮信号)发送至智能设备200上,控制装置100可经由通信器130接收由智能设备200发送的信号。通信器130可以包括红外信号接口131和射频信号接口132。例如:红外信号接口时,需要将用户输入指令按照红外控制协议转化为红外控制信号,经红外发送模块进行发送至智能设备200。再如:射频信号接口时,需将用户输入指令转化为数字信号,然后按照射频控制信号调制协议进行调制后,由射频发送端子发送至智能设备200。

[0074] 用户输入接口140,可包括麦克风141、触摸板142、传感器143、按键144等中至少一者,从而用户可以通过语音、触摸、手势、按压等将关于控制智能设备200的用户指令输入到控制装置100。

[0075] 输出接口150,通过将用户输入接口140接收的用户指令输出至智能设备200,或者,输出由智能设备200接收的图像或语音信号。这里,输出接口150可以包括LED接口151、产生振动的振动接口152、输出声音的声音输出接口153和输出图像的显示器154等。例如,遥控器100A可从输出接口150接收音频、视频或数据等输出信号,并且将输出信号在显示器154上显示为图像形式、在声音输出接口153输出为音频形式或在振动接口152输出为振动形式。

[0076] 供电电源160,用于在控制器110的控制下为控制装置100各元件提供运行电力支持。形式可以为电池及相关控制电路。

[0077] 图1c中示例性示出了智能设备200的硬件配置框图。如图1c所示,智能设备200中可以包括调谐解调器210、通信器220、检测器230、外部装置接口240、控制器250、存储器260、用户接口265、视频处理器270、显示器275、音频处理器280、音频输出接口285、供电电源290。

[0078] 调谐解调器210,通过有线或无线方式接收广播电视信号,可以进行放大、混频和谐振等调制解调处理,用于从多个无线或有线广播电视信号中解调出用户所选择的电视频道的频率中所携带的音视频信号,以及附加信息(例如EPG数据)。

[0079] 调谐解调器210,可根据用户选择,以及由控制器250控制,响应用户选择的电视频道的频率以及该频率所携带的电视信号。

[0080] 调谐解调器210,根据电视信号的广播制式不同,可以接收信号的途径有很多种,诸如:地面广播、有线广播、卫星广播或互联网广播等;以及根据调制类型不同,可以数字调制方式或模拟调制方式;以及根据接收电视信号的种类不同,可以解调模拟信号和数字信号。

[0081] 在其他一些示例性实施例中,调谐解调器210也可在外部设备中,如外部机顶盒等。这样,机顶盒通过调制解调后输出电视信号,经过外部装置接口240输入至智能设备200中。

[0082] 通信器220,是用于根据各种通信协议类型与外部设备或外部服务器进行通信的组件。例如智能设备200可将内容数据发送至经由通信器220连接的外部设备,或者,从经由通信器220连接的外部设备浏览和下载内容数据。通信器220可以包括WIFI模块221、蓝牙通信协议模块222、有线以太网通信协议模块223等网络通信协议模块或近场通信协议模块,从而通信器220可根据控制器250的控制接收控制装置100的控制信号,并将控制信号实现为WIFI信号、蓝牙信号、射频信号等。

[0083] 检测器230,是智能设备200用于采集外部环境或与外部交互的信号组件。检测器230可以包括声音采集器231,如麦克风,可以用于接收用户的语音信号,如用户控制智能设备200的控制指令的语音信号;或者,可以采集用于识别环境场景类型的环境声音,实现智能设备200可以自适应环境噪声。

[0084] 在其他一些示例性实施例中,检测器230,还可以包括图像采集器232,如相机、摄像头等,可以用于采集外部环境场景,以自适应变化智能设备200的显示参数;以及用于在设定周期采集多个第一图像,所述第一图像中可能包含用户的属性或与用户交互手势,以实现智能设备与用户之间互动的功能。

[0085] 在其他一些示例性实施例中,检测器230,还可以包括光接收器,用于采集环境光线强度,以自适应智能设备200的显示参数变化等。

[0086] 在其他一些示例性实施例中,检测器230,还可以包括温度传感器,如通过感测环境温度,智能设备200可自适应调整图像的显示色温。示例性的,当温度偏高的环境时,可调整智能设备200显示图像色温偏冷色调;当温度偏低的环境时,可以调整智能设备200显示图像色温偏暖色调。

[0087] 外部装置接口240,是提供控制器250控制智能设备200与外部设备间数据传输的组件。外部装置接口240可按照有线/无线方式与诸如机顶盒、游戏装置、笔记本电脑等外部设备连接,可接收外部设备的诸如视频信号(例如运动图像)、音频信号(例如音乐)、附加信息(例如EPG)等数据。

[0088] 其中,外部装置接口240可以包括:高清多媒体接口(HDMI)端子241、复合视频消隐同步(CVBS)端子242、模拟或数字分量端子243、通用串行总线(USB)端子244、组件(Component)端子(图中未示出)、红绿蓝(RGB)端子(图中未示出)等任一个或多个。

[0089] 控制器250,通过运行存储在存储器260上的各种软件控制程序(如操作系统和各种应用程序),来控制智能设备200的工作和响应用户的操作。控制器250还可以针对各个第一图像分别执行以下处理,每读取一个第一图像M,根据预设截取规则,从第一图像M上截取一个第二图像N,其中,用户手部图像在第二图像N上的占比高于所述用户手部图像在第一图像M上的占比;对第二图像N进行识别,确定第二图像N上的至少一个手势识别区域,以及对应的手势识别结果;若同一手势识别区域中存在超过设定阈值的同一控制手势,则基于该同一控制手势控制智能设备执行相应的操作。

[0090] 如图1c所示,控制器250包括随机存取存储器(RAM) 251、只读存储器(ROM) 252、图形处理器253、CPU处理器254、通信接口255、以及通信总线256。其中,RAM251、ROM252以及图

形处理器253、CPU处理器254通信接口255通过通信总线256相连接。

[0091] ROM252,用于存储各种系统启动指令。如在接收到开机信号时,智能设备200电源开始启动,CPU处理器254运行ROM252中的系统启动指令,将存储在存储器260的操作系统拷贝至RAM251中,以开始运行启动操作系统。当操作系统启动完成后,CPU处理器254再将存储器260中各种应用程序拷贝至RAM251中,然后,开始运行启动各种应用程序。

[0092] 图形处理器253,用于产生各种图形对象,如图标、操作菜单、以及用户输入指令显示图形等。图形处理器253可以包括运算器,用于通过接收用户输入各种交互指令进行运算,进而根据显示属性显示各种对象;以及包括渲染器,用于产生基于运算器得到的各种对象,将进行渲染的结果显示在显示器275上。

[0093] CPU处理器254,用于执行存储在存储器260中的操作系统和应用程序指令。以及根据接收的用户输入指令,来执行各种应用程序、数据和内容的处理,以便最终显示和播放各种音视频内容。

[0094] 在一些示例性实施例中,CPU处理器254,可以包括多个处理器。多个处理器可包括一个主处理器以及多个或一个子处理器。主处理器,用于在智能设备预加载模式中执行智能设备200的一些初始化操作,和/或,在正常模式下显示画面的操作。多个或一个子处理器,用于执行在智能设备待机模式等状态下的一种操作。

[0095] 通信接口255,可包括第一接口到第n接口。这些接口可以是经由网络被连接到外部设备的网络接口。

[0096] 控制器250可以控制智能设备200的整体操作。例如:响应于接收到用于选择在显示器275上显示的GUI对象的用户输入命令,控制器250便可以执行与由用户输入命令选择的对象有关的操作。

[0097] 其中,该对象可以是可选对象中的任何一个,例如超链接或图标。该与所选择的对象有关的操作,例如显示连接到超链接页面、文档、图像等操作,或者执行与对象相对应的程序的操作。该用于选择GUI对象的用户输入命令,可以通过连接到智能设备200的各种输入装置(例如,鼠标、键盘、触摸板等)输入命令或者与由用户说出语音相对应的语音命令。

[0098] 存储器260,用于存储驱动和控制智能设备200运行的各种类型的数据、软件程序或应用程序。存储器260可以包括易失性和/或非易失性存储器。而术语“存储器”包括存储器260、控制器250的RAM251和ROM252、或智能设备200中的存储卡。

[0099] 在一些实施例中,存储器260具体用于存储驱动智能设备200中控制器250的运行程序;存储智能设备200内置的和用户从外部设备下载的各种应用程序;存储用于配置由显示器275提供的各种GUI、与GUI相关的各种对象及用于选择GUI对象的选择器的视觉效果图像等数据。

[0100] 在一些实施例中,存储器260具体用于存储调谐解调器210、通信器220、检测器230、外部装置接口240、视频处理器270、显示器275、音频处理器280等的驱动程序和相关数据,例如从外部装置接口接收的外部数据(例如音视频数据)或用户接口接收的用户数据(例如按键信息、语音信息、触摸信息等)。

[0101] 在一些实施例中,存储器260具体存储用于表示操作系统(OS)的软件和/或程序,这些软件和/或程序可包括,例如:内核、中间件、应用编程接口(API)和/或应用程序。示例

性的,内核可控制或管理系统资源,以及其它程序所实施的功能(如所述中间件、API或应用程序);同时,内核可以提供接口,以允许中间件、API或应用程序访问控制器,以实现控制或管理系统资源。

[0102] 图1d中示例性示出了智能设备200存储器中操作系统的架构配置框图。该操作系统架构从上到下依次是应用层、中间件层和内核层。

[0103] 应用层,系统内置的应用程序以及非系统级的应用程序都是属于应用层。负责与用户进行直接交互。应用层可包括多个应用程序,如设置应用程序、电子帖应用程序、媒体中心应用程序等。这些应用程序可被实现为Web应用,其基于WebKit引擎来执行,具体可基于HTML5、层叠样式表(CSS)和JavaScript来开发并执行。

[0104] 这里,HTML,全称为超文本标记语言(HyperText Markup Language),是一种用于创建网页的标准标记语言,通过标记标签来描述网页,HTML标签用以说明文字、图形、动画、声音、表格、链接等,浏览器会读取HTML文档,解释文档内标签的内容,并以网页的形式显示出来。

[0105] CSS,全称为层叠样式表(Cascading Style Sheets),是一种用来表现HTML文件样式的计算机语言,可以用来定义样式结构,如字体、颜色、位置等的语言。CSS样式可以直接存储与HTML网页或者单独的样式文件中,实现对网页中样式的控制。

[0106] JavaScript,是一种应用于Web网页编程的语言,可以插入HTML页面并由浏览器解释执行。其中Web应用的交互逻辑都是通过JavaScript实现。JavaScript可以通过浏览器,封装JavaScript扩展接口,实现与内核层的通信。

[0107] 中间件层,可以提供一些标准化的接口,以支持各种环境和系统的操作。例如,中间件层可以实现为与数据广播相关的中间件的多媒体和超媒体信息编码专家组(MHEG),还可以实现为与外部设备通信相关的中间件的DLNA中间件,还可以实现为提供智能设备内各应用程序所运行的浏览器环境的中间件等。

[0108] 内核层,提供核心系统服务,例如:文件管理、内存管理、进程管理、网络管理、系统安全权限管理等服务。内核层可以被实现为基于各种操作系统的内核,例如,基于Linux操作系统的内核。

[0109] 内核层也同时提供系统软件和硬件之间的通信,为各种硬件提供设备驱动服务,例如:为显示器提供显示驱动程序、为摄像头提供摄像头驱动程序、为遥控器提供按键驱动程序、为WIFI模块提供WiFi驱动程序、为音频输出接口提供音频驱动程序、为电源管理(PM)模块提供电源管理驱动等。

[0110] 用户接口265,接收各种用户交互。具体的,用于将用户的输入信号发送给控制器250,或者,将从控制器250的输出信号传送给用户。示例性的,遥控器100A可将用户输入的诸如电源开关信号、频道选择信号、音量调节信号等输入信号发送至用户接口265,再由用户接口265转送至控制器250;或者,遥控器100A可接收经控制器250处理从用户接口265输出的音频、视频或数据等输出信号,并且显示接收的输出信号或将接收的输出信号输出为音频或振动形式。

[0111] 在一些实施例中,用户可在显示器275上显示的图形用户界面(GUI)输入用户命令,则用户接口265通过GUI接收用户输入命令。确切的说,用户接口265可接收用于控制选择器在GUI中的位置以选择不同的对象或项目的用户输入命令。

[0112] 或者,用户可通过输入特定的声音或手势进行输入用户命令,则用户接口265通过传感器识别出声音或手势,来接收用户输入命令。视频处理器270,用于接收外部的视频信号,根据输入信号的标准编解码协议,进行解压缩、解码、缩放、降噪、帧率转换、分辨率转换、图像合成等视频数据处理,可得到直接在显示器275上显示或播放的视频信号。

[0113] 示例的,视频处理器270,包括解复用模块、视频解码模块、图像合成模块、帧率转换模块、显示格式化模块等。

[0114] 其中,解复用模块,用于对输入音视频数据流进行解复用处理,如输入MPEG-2流(基于数字存储媒体运动图像和语音的压缩标准),则解复用模块将其进行解复用成视频信号和音频信号等。

[0115] 视频解码模块,用于对解复用后的视频信号进行处理,包括解码和缩放处理等。

[0116] 图像合成模块,如图像合成器,其用于将图形生成器根据用户输入或自身生成的GUI信号,与缩放处理后视频图像进行叠加混合处理,以生成可供显示的图像信号。

[0117] 帧率转换模块,用于对输入视频的帧率进行转换,如将输入的60Hz视频的帧率转换为120Hz或240Hz的帧率,通常的格式采用如插帧方式实现。

[0118] 显示格式化模块,用于将帧率转换模块输出的信号,改变为符合诸如显示器显示格式的的信号,如将帧率转换模块输出的信号进行格式转换以输出RGB数据信号。

[0119] 显示器275,用于接收源自视频处理器270输入的图像信号,进行显示视频内容、图像以及菜单操控界面。显示视频内容,可以来自调谐解调器210接收的广播信号中的视频内容,也可以来自通信器220或外部装置接口240输入的视频内容。显示器275,同时显示智能设备200中产生且用于控制智能设备200的用户操控界面UI。

[0120] 以及,显示器275可以包括用于呈现画面的显示屏组件以及驱动图像显示的驱动组件。或者,倘若显示器275为一种投影显示器,还可以包括一种投影装置和投影屏幕。

[0121] 音频处理器280,用于接收外部的音频信号,根据输入信号的标准编解码协议,进行解压缩和解码,以及降噪、数模转换、和放大处理等音频数据处理,得到可以在扬声器286中播放的音频信号。

[0122] 示例性的,音频处理器280可以支持各种音频格式。例如MPEG-2、MPEG-4、高级音频编码(AAC)、高效AAC(HE-AAC)等格式。

[0123] 音频输出接口285,用于在控制器250的控制下接收音频处理器280输出的音频信号,音频输出接口285可包括扬声器286,或输出至外接设备的发生装置的外接音响输出端子287,如耳机输出端子。

[0124] 在其他一些示例性实施例中,视频处理器270可以包括一个或多个芯片组成。音频处理器280,也可以包括一个或多个芯片组成。

[0125] 以及,在其他一些示例性实施例中,视频处理器270和音频处理器280,可以为单独的芯片,也可以与控制器250一起集成在一个或多个芯片中。

[0126] 供电电源290,用于在控制器250的控制下,将外部电源输入的电力为智能设备200提供电源供电支持。供电电源290可以是安装在智能设备200内部的内置电源电路,也可以是安装在智能设备200外部的电源。

[0127] 在现有技术中,需要采集并识别用户的动态手势图像,根据对应的识别结果控制智能设备执行相应的操作,但采用上述方式控制智能设备时,会产生以下问题:用户距离摄

像头越远,在采集到的图像中用户的手部图像占比越小,导致手部检测难度大幅度提升;神经网络的识别过程会花费较长时间,识别连续的多张图像,无法满足实时响应需求;以及,用户需要做动态手势才能满足响应,用户体验感较差。为了解决前述问题,本申请实施例提供了一种新的控制智能设备的方法,参阅图2所示,包括以下步骤:

[0128] S201:接收图像采集器在设定周期内采集的多个第一图像。

[0129] 智能设备处于开机状态时,调用图像采集器在设定周期内,按照预设采集频率拍摄当前场景,得到连续的多个第一图像。由于图像采集器是根据设定周期和预设采集频率进行图像采集操作,可能会采集到无用户的纯背景图像,也可能采集到包含一个用户或者多个用户的图像。例如,摄像头在1秒内采集到20张图像,其中,前5张图像是不包含用户的纯背景图像,第6-20张图像是包含两个用户的图像。

[0130] S202:读取一个第一图像M。

[0131] S203:根据预设截取规则,从第一图像M上截取一个第二图像N,其中,用户手部图像在第二图像N上的占比高于用于手部图像在第一图像M上的占比。

[0132] 由于本申请实施例中需要对用户的手部图像进行检测,但手部图像在采集到的第一图像中占据的像素数一般较少,而且随着用户距离摄像头越远,在采集到的第一图像中手部图像的占比越小,导致手部检测难度大幅度提升,因此,本申请实施例在检测手部图像之前,需要对第一图像进行预处理操作,以截取出手部图像占比率高的第二图像,降低检测难度。具体地,从第一图像M上截取第二图像N的步骤如下:

[0133] A1:根据图像采集器的视场角、设定视场角阈值,以及第一图像M的尺寸信息,确定手部图像的尺寸信息和角点坐标信息。

[0134] 在图像采集器中,以图像采集器的镜头为顶点,以被测目标的物象可通过镜头的最大范围的两条边缘构成的夹角称为视场角,参阅图3a所示,示出了视场角示意图。

[0135] 参阅图3b所示,当通过镜头中心的线(即镜头光轴)与地面平行时,此时被测目标的物象可通过镜头的最大范围的两条边缘构成的夹角,称为在X方向的视场角(即 V_x);参阅图3c所示,当镜头光轴垂直于地面时,此时被测目标的物象可通过镜头的最大范围的两条边缘构成的夹角,称为在Y方向的视场角(即 V_y)。

[0136] 若 V_x 和 V_y 均小于视场角阈值,说明第一图像M的尺寸较小,手部图像在该第一图像M的占比率较高,则将整幅第一图像M作为第二图像N输入到训练完毕的手部检测网络中,检测第二图像N上包含的至少一个手势识别区域。

[0137] 若 V_x 大于视场角阈值,采用公式(1)计算包含手部图像的感兴趣区域(Region Of Interest,ROI)的待截取宽度, V_x 表征在X方向的视场角, T_v 表征视场角阈值, L_x 表征ROI的待截取宽度,W表征第一图像M的宽度。

$$[0138] \quad \frac{T_v}{V_x} = \frac{L_x}{W} \quad \text{公式(1);}$$

[0139] 若 V_y 大于视场角阈值,采用公式(2)计算包含手部图像的感兴趣区域(Region Of Interest,ROI)的待截取高度, V_y 表征在Y方向的视场角, T_v 表征视场角阈值, L_y 表征ROI的待截取高度,H表征第一图像M的高度。

$$[0140] \quad \frac{T_v}{V_y} = \frac{L_y}{H} \quad \text{公式(2);}$$

[0141] 进一步地,本申请实施例中将ROI的左上角确定为角点,采用公式(3)计算ROI的角点坐标信息。最终ROI的尺寸信息和角点坐标信息可表示为 (x,y,w,h) , (x,y) 表征ROI的角点坐标信息, w 等同于 L_x ,表征ROI的待截取宽度; h 等同于 L_y ,表征ROI的待截取高度。

$$[0142] \quad \begin{cases} x = \frac{W}{2} - \frac{L_x}{2} \\ y = \frac{H}{2} - \frac{L_y}{2} \end{cases} \quad \text{公式(3);}$$

[0143] A2:根据手部图像的尺寸信息和角点坐标信息,从第一图像M中截取出第二图像N。

[0144] S204:对第二图像N进行识别,确定第二图像N上的至少一个手势识别区域,以及对应的手势识别结果。

[0145] B1:将第二图像N输入到训练完毕的手部检测网络中,得到第二图像N上包含的手势识别区域的数量,以及各个手势识别区域在第二图像N上的坐标信息,和各个手势识别区域中存在手部图像的概率值(即手势识别区域的置信度)。

[0146] 为了保证手部检测的准确率,本申请实施例采用了多尺度SSD算法,设计了一个端到端的手部检测网络。手部检测网络的架构设计和训练过程如下:

[0147] 手部检测网络由输入层、卷积层、池化层和输出层构成,其中,卷积层用于从输入层传输的第二图像N提取有用的特征,比如,水平、垂直、边缘或对角线等特征;池化层用于提高提取到的特征的感受野和降低优化难度,所谓感受野指的是一个像素对应回第一图像M的区域大小。

[0148] 为了快速缩小第二图像N的尺寸,本申请实施例为卷积层和池化层配置了较大的采样步长。例如,卷积层1的采样步长为4,卷积层2、池化层1、池化层2的采样步长均为2,则第二图像N在经过两个卷积层和2个池化层后,尺寸缩小了32倍。进一步地,本申请实施例采用包含多种卷积分支的Inception模块,Inception模块是由不同结构的池化层和卷积层组成的,一方面增加了网络的宽度,另一方面增加了网络对尺度的适应性,有效提高了感受野的多样性。

[0149] 对各个样本图像进行标注,将样本图像上包含手部图像的矩形框称为正类矩形框,将仅包含纯背景的矩形框称为负类矩形框,并确定正类矩形框与负类矩形框对应的坐标信息;

[0150] 手部检测网络每读取一个标注后的样本图像,得到手部检测网络输出的预测矩形框的坐标信息和类别信息,由于样本图像上负类矩形框的数量远高于正类矩形框的数量,若基于全部的实际矩形框和其对应的预测矩形框之间的类别误差值,重新调整手部检测网络的参数,会使得训练完毕的手部检测网络偏向于输出负类矩形框的结果,为了解决这个问题,本申请实施例中只选取全部的正类矩形框和较难检测的多个负类矩形框,调用分类损失函数,计算所述多个负类矩形框与对应的预测矩形框之间的类别误差值,以及计算全部的正类矩形框与对应的预测矩形框之间的类别误差值;调用位置损失函数,仅计算预测矩形框坐标位置与对应的正类矩形框坐标位置之间的位置误差值;最后,基于上述类别误差值和位置误差值,重新调整手部检测网络的参数;

[0151] 重复上述训练过程,直到达到设定迭代次数为止,或者所有样本图像均读取完毕为止,或者误差值低于设定误差阈值为止,输出训练完毕的手部检测网络。

[0152] B2:将各个手势识别区域对应的图像,分别输入到训练完毕的手势分类网络中,得

到对应的手势识别结果。

[0153] 手势分类网络由基础网络、全连接层和SoftMax分类网络两部分组成,将图像输入到基础网络中,最终由SoftMax分类网络输出该图像的手势类别和手势类别的置信度(即手势识别结果)。其中,基础网络可为以下任意一种视觉几何组网络(Visual Geometry Group Network,VGGNet)、残差网络(Residual Network,ResNet)、AlexNet或者其他卷积神经网络(Convolutional Neural Network,CNN)。

[0154] 手势类别可分为无效手势、唤醒手势和控制手势。本申请实施例的唤醒手势如图4a所示,设置唤醒手势既可以告知智能设备接下来的动作才需要执行相应的操作,避免误操作,还可以解决智能设备在多人交互场景下确定手势控制区域,保证在多人交互场景下只允许一个用户控制智能设备。本申请实施例中,图4b所示的控制手势表征调高音量,图4c所示的控制手势表征调低音量,图4d所示的控制手势表征视频后退,图4e所示的控制手势表征视频快进,图4f所示的控制手势表征确定,图4g所示的控制手势表征取消,图4h所示的控制手势表征静音/结束。将不属于预设手势的其它手势定义为无效手势。

[0155] S205:判断全部第一图像是否均读取完毕,若是,执行步骤206;否则,返回步骤202。

[0156] S206:若同一手势识别区域中存在超过设定数量阈值的同一控制手势,则基于该同一控制手势控制智能设备执行相应的操作。

[0157] 本申请实施例采用静态手势识别,即对一幅图像进行手势识别操作,输出手势识别结果,由于图像采集器在设定周期内采集了多个第一图像,若智能设备按照每一个第一图像的手势识别结果,控制自身执行相应的操作,智能设备可能会在很短的时间内多次执行同一操作,不利于智能设备的控制。为了解决前述问题,本申请实施例的智能设备会根据多帧的手势识别结果,综合判定最终需执行的操作。

[0158] 首先,每读取一个第一图像,对其执行步骤203-204,得到该图像上包含的手势识别区域的手势识别结果,也就是说,当全部第一图像均读取完毕时,会得到至少一个手势识别区域,及其对应的手势识别结果集合,为了解决智能设备在多人交互场景下确定手势控制区域,本申请实施例中设置了唤醒手势。因此,若同一手势识别区域中存在连续M个唤醒手势,则将该同一手势识别区域确定为手势控制区域,其中M为正整数。

[0159] 其次,若该手势控制区域在所述连续M个唤醒手势之后,存在连续N个第一操作手势,则基于该第一控制手势控制智能设备执行相应的操作,其中,N为正整数。

[0160] 例如,手势识别区域1包含20个手势识别结果,其中,第3-10个均为唤醒手势,第11-15均为如图4b所示的控制手势,则智能设备将按照该控制手势的指示,调高视频播放的音量。

[0161] 进一步地,在基于第一控制手势控制智能设备执行相应的操作之后,若该手势控制区域在连续N个第一控制手势之后,存在连续N个第二控制手势,则基于该第二控制手势控制智能设备执行相应的操作。其中,N为正整数,第一控制手势和第二控制手势可以为相同的控制手势,也可以为不同的控制手势。

[0162] 承接上述例子,智能设备在执行完调高视频播放的音量操作后,若第16-20均为如图4e所示的控制手势,则智能设备将按照该控制手势的指示,对当前播放的视频执行快进操作。

[0163] 进一步地,在基于第一控制手势控制智能设备执行相应的操作之后,若该手势控制区域在连续N个第一控制手势之后,不存在连续N个第二控制手势,在确定其他手势识别区域中存在连续M个唤醒手势时,将控制焦点由当前手势识别区域转移到前述的其他手势识别区域,并将其他手势识别区域确定为新的手势控制区域,以便智能设备根据新的手势控制区域的手势识别结果集合,执行新的操作。

[0164] 进一步地,在基于第一控制手势控制智能设备执行相应的操作之后,若该手势控制区域在连续N个第一控制手势之后,不存在连续N个第二控制手势,且其他手势识别区域中均不存在连续M个唤醒手势时,不转移控制焦点,直到该手势控制区域中出现符合要求的控制手势,并根据该控制手势执行相应的操作。但是,当智能设备重新启动后,默认不存在有效的手势控制区域,需要根据符合要求的唤醒手势,重新确定手势控制区域。

[0165] 在一些可能的实施方式中,本申请提供的智能设备的控制方法的各个方面还可以实现为一种程序产品的形式,其包括程序代码,当程序产品在计算机设备上运行时,程序代码用于使计算机设备执行本说明书上述描述的根据本申请各种示例性实施方式的业务控制方法中的步骤,例如,计算机设备可以执行如图2中所示的步骤。

[0166] 程序产品可以采用一个或多个可读介质的任意组合。可读介质可以是可读信号介质或者可读存储介质。可读存储介质例如可以是但不限于电、磁、光、电磁、红外线、或半导体的系统、装置或器件,或者任意以上的组合。可读存储介质的更具体的例子(非穷举的列表)包括:具有一个或多个导线的电连接、便携式盘、硬盘、随机存取存储器(RAM)、只读存储器(ROM)、可擦式可编程只读存储器(EPROM或闪存)、光纤、便携式紧凑盘只读存储器(CD-ROM)、光存储器件、磁存储器件、或者上述的任意合适的组合。

[0167] 本申请的实施方式的用于业务控制的程序产品可以采用便携式紧凑盘只读存储器(CD-ROM)并包括程序代码,并可以在计算装置上运行。然而,本申请的程序产品不限于此,在本文件中,可读存储介质可以是任何包含或存储程序的有形介质,该程序可以被命令执行系统、装置或者器件使用或者与其结合使用。

[0168] 可读信号介质可以包括在基带中或者作为载波一部分传播的数据信号,其中承载了可读程序代码。这种传播的数据信号可以采用多种形式,包括但不限于电磁信号、光信号或上述的任意合适的组合。可读信号介质还可以是可读存储介质以外的任何可读介质,该可读介质可以发送、传播或者传输用于由命令执行系统、装置或者器件使用或者与其结合使用的程序。

[0169] 可读介质上包含的程序代码可以用任何适当的介质传输,包括但不限于无线、有线、光缆、RF等,或者上述的任意合适的组合。

[0170] 可以以一种或多种程序设计语言的任意组合来编写用于执行本申请操作的程序代码,程序设计语言包括面向对象的程序设计语言—诸如Java、C++等,还包括常规的过程式程序设计语言—诸如“C”语言或类似的程序设计语言。程序代码可以完全地在用户计算装置上执行、部分地在用户设备上执行、作为一个独立的软件包执行、部分在用户计算装置上部分在远程计算装置上执行、或者完全在远程计算装置或服务器上执行。在涉及远程计算装置的情形中,远程计算装置可以通过任意种类的网络包括局域网(LAN)或广域网(WAN)连接到用户计算装置,或者,可以连接到外部计算装置(例如利用因特网服务提供商来通过因特网连接)。

[0171] 尽管已描述了本申请的优选实施例,但本领域内的技术人员一旦得知了基本创造性概念,则可对这些实施例做出另外的变更和修改。所以,所附权利要求意欲解释为包括优选实施例以及落入本申请范围的所有变更和修改。

[0172] 显然,本领域的技术人员可以对本申请进行各种改动和变型而不脱离本申请的精神和范围。这样,倘若本申请的这些修改和变型属于本申请权利要求及其等同技术的范围之内,则本申请也意图包含这些改动和变型在内。

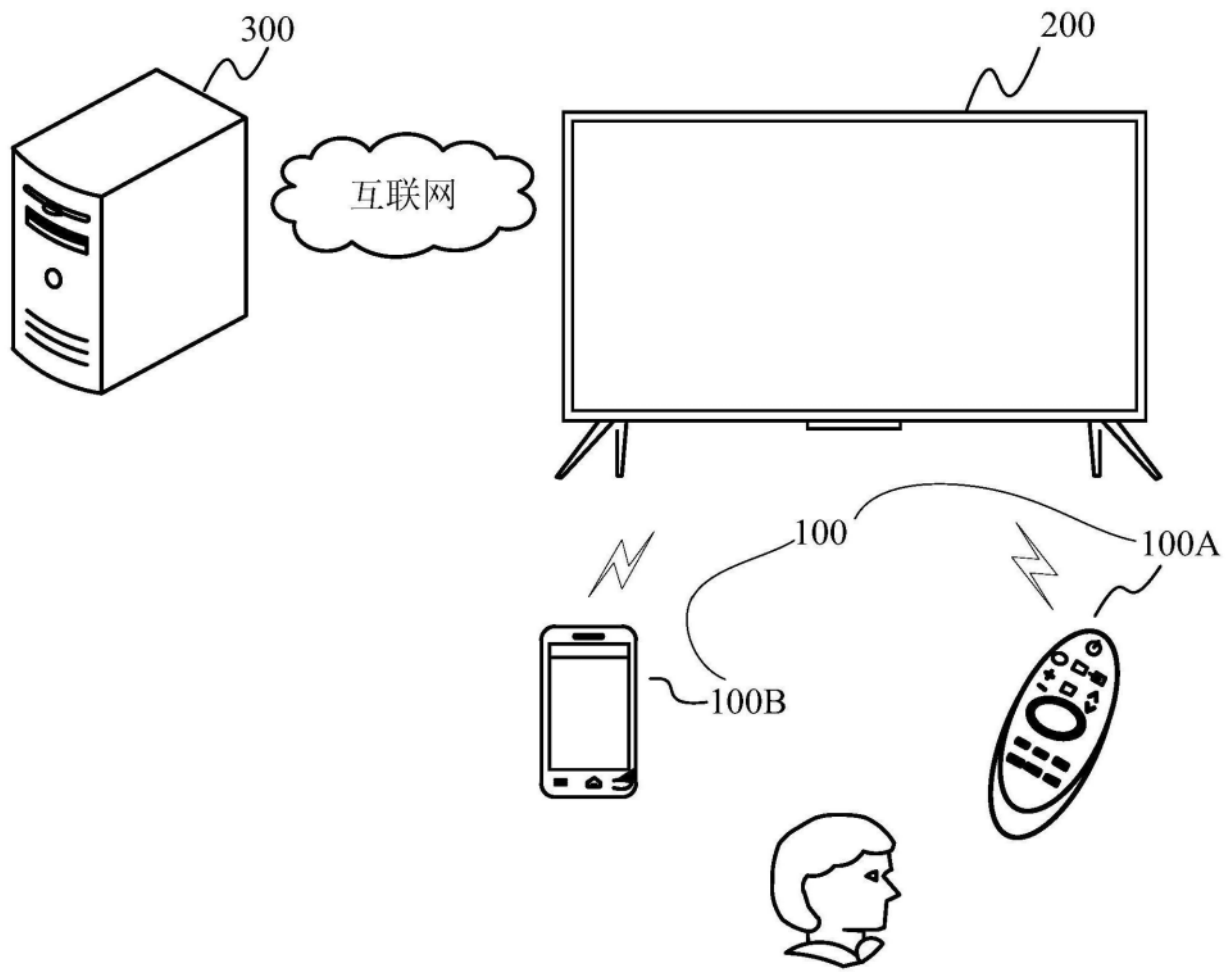


图1a

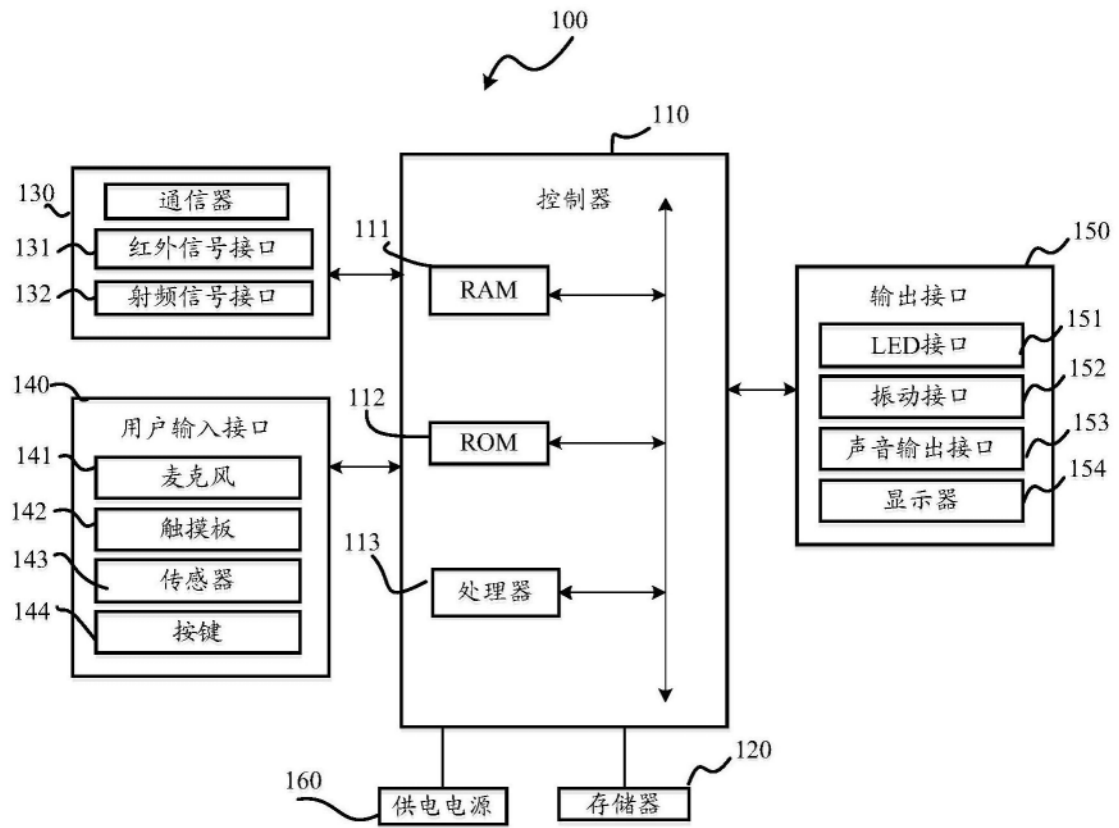


图1b

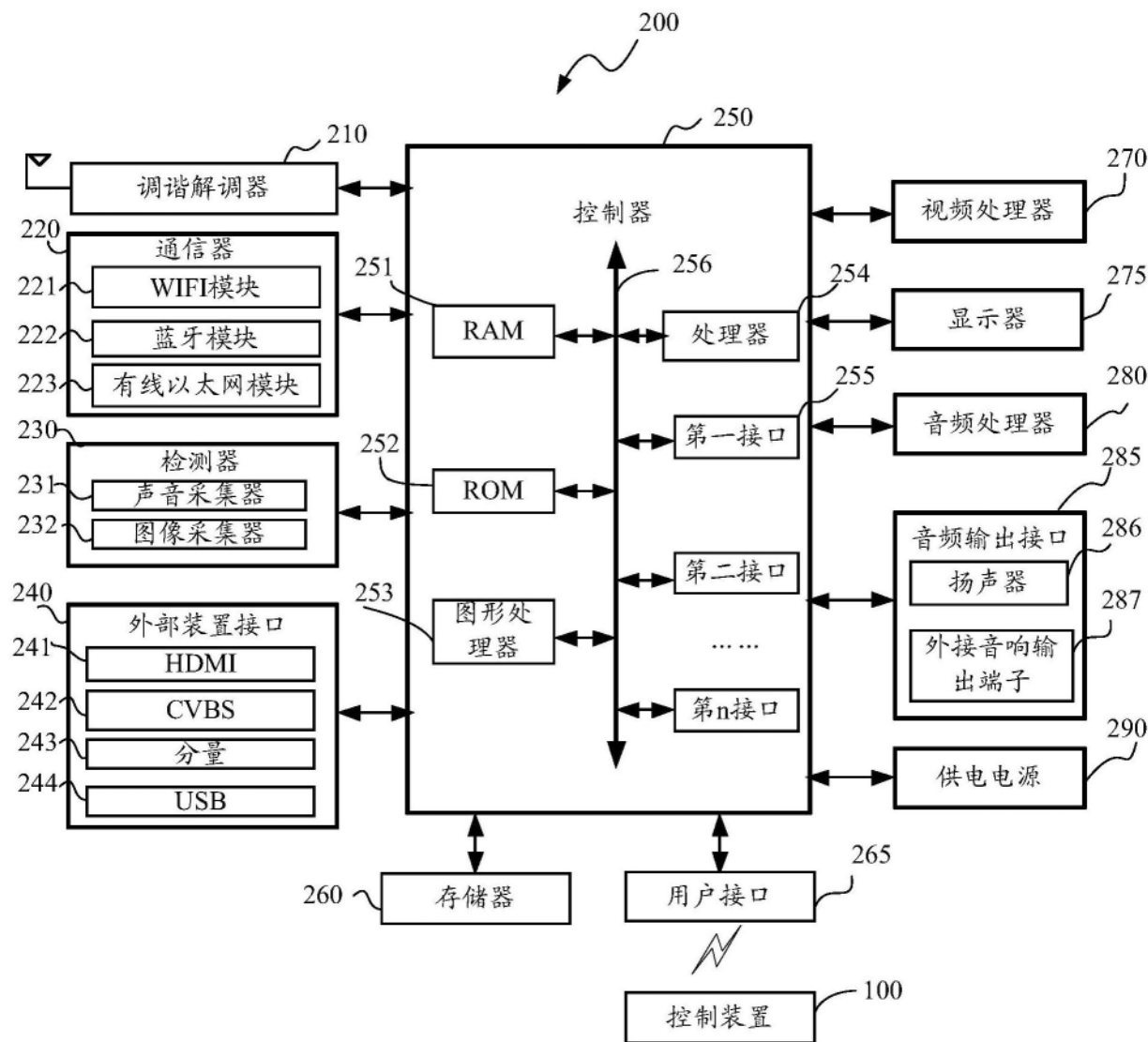


图1c

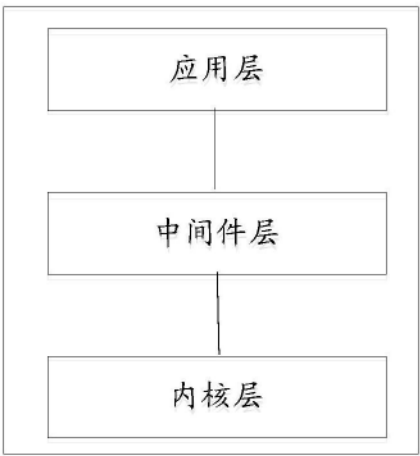


图1d

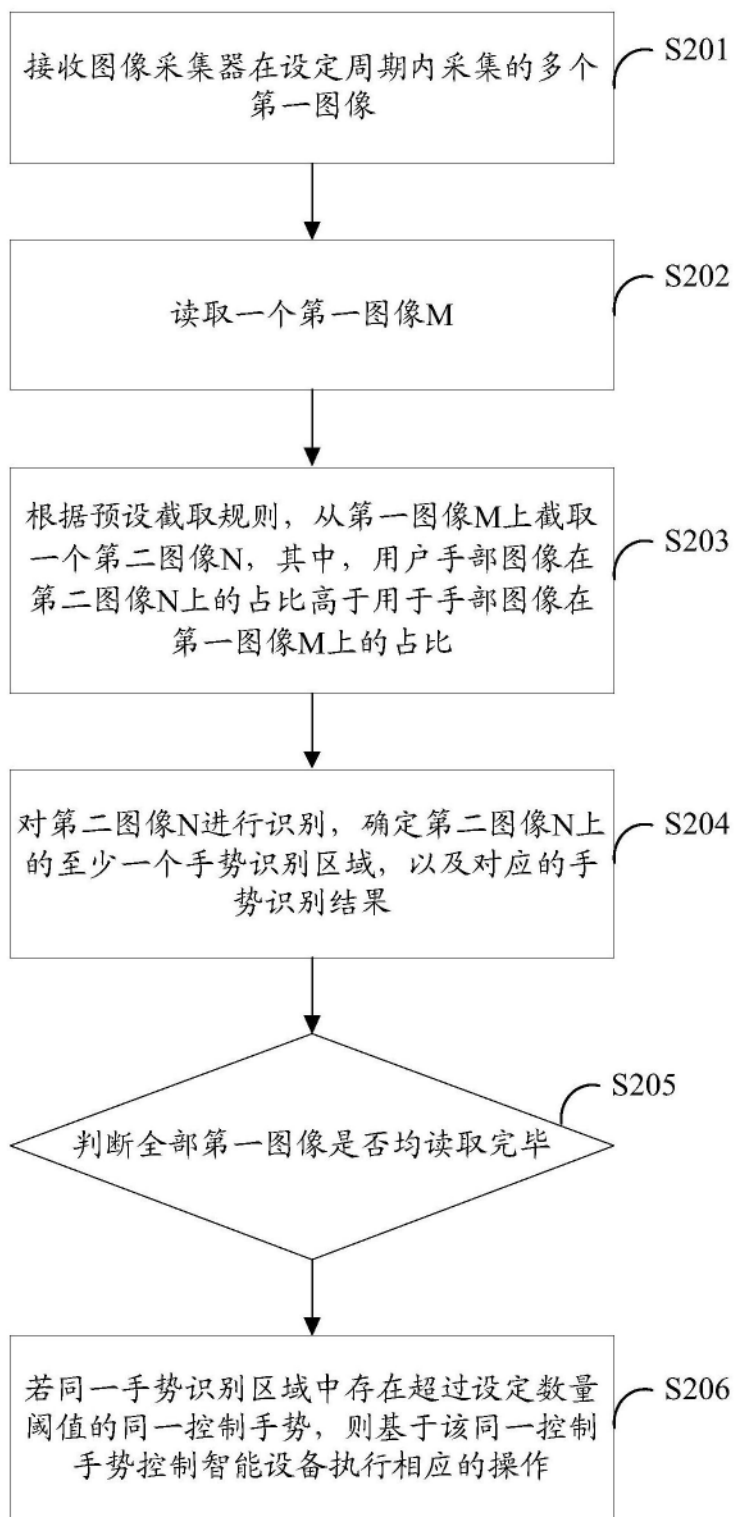


图2

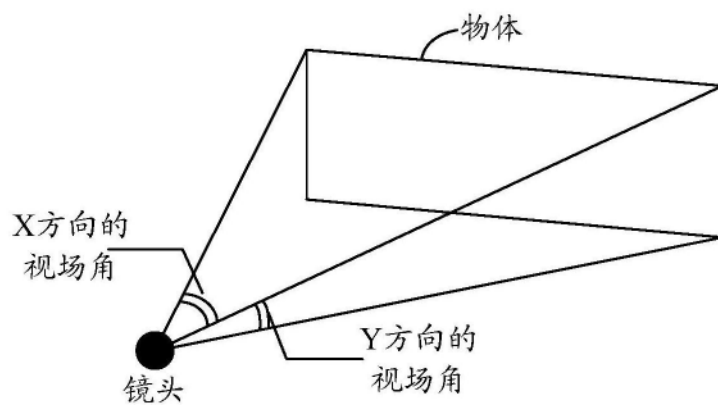


图3a

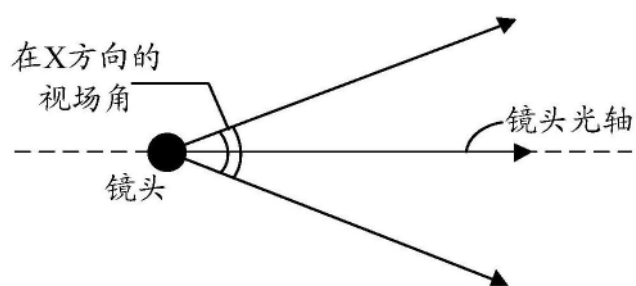


图3b

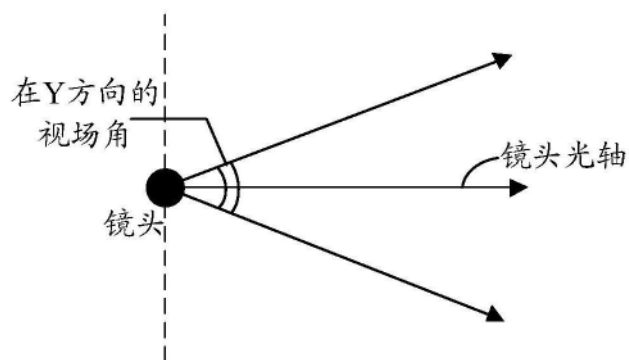


图3c



图4a



图4b



图4c



图4d



图4e



图4f



图4g

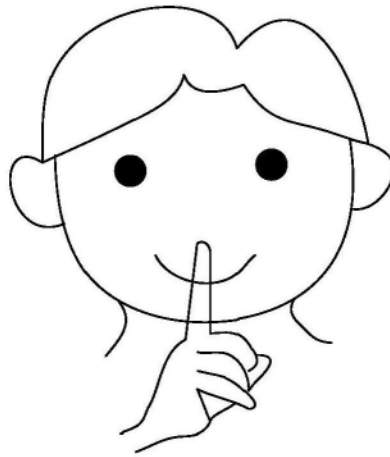


图4h