



(12) 发明专利申请

(10) 申请公布号 CN 113741770 A

(43) 申请公布日 2021. 12. 03

(21) 申请号 202010474493.8

(22) 申请日 2020.05.29

(71) 申请人 比亚迪股份有限公司

地址 518118 广东省深圳市坪山新区比亚迪路3009号

(72) 发明人 叶将涛 阮洲 杨世琨 刘柯

(51) Int. Cl.

G06F 3/0484 (2013.01)

G06F 3/16 (2006.01)

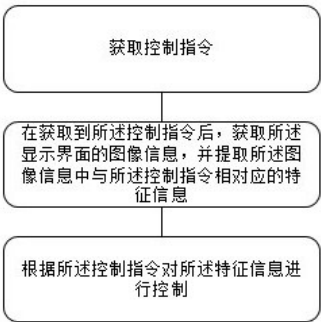
权利要求书4页 说明书16页 附图3页

(54) 发明名称

基于图像识别的控制方法和系统及车辆、存储介质

(57) 摘要

本发明公开了一种基于图像识别的控制方法和系统及车辆、存储介质,所述方法,包括:获取控制指令;获取所述显示界面的图像信息并提取所述图像信息中与所述控制指令相对应的特征信息;根据所述控制指令对所述特征信息进行控制。由此,不需要对软件进行适配性调试,用户也可以对软件进行语音操控,提升了用户的使用体验。



1. 一种基于图像识别的控制方法,其特征在于,包括:

获取控制指令;

在获取到所述控制指令后,获取所述显示界面的图像信息,并提取所述图像信息中与所述控制指令相对应的特征信息;

根据所述控制指令对所述特征信息进行控制。

2. 根据权利要求1所述的基于图像识别的控制方法,其特征在于,所述在获取到所述控制指令后,获取所述显示界面的图像信息,并提取所述图像信息中与所述控制指令相对应的特征信息包括,

在获取到所述控制指令后,获取显示界面的状态信息并进行判断;

如果所述状态信息发生变化,并且无法从状态信息变化后的所述显示界面的图像信息中提取到与所述控制指令相对应的特征信息,则根据状态信息变化前的所述显示界面的图像信息提取与所述控制指令相对应的特征信息。

3. 根据权利要求1所述的基于图像识别的控制方法,其特征在于,如果无法从所述显示界面的图像信息中提取到至少部分与所述控制指令相对应的特征信息,则根据所述控制指令对已提取到的与所述控制指令相对应的特征信息进行控制;

所述方法还包括,

继续获取所述显示界面的图像信息并提取所述至少部分未提取到的与所述控制指令相对应的特征信息;

根据所述控制指令对新提取到的与所述控制指令相对应的特征信息进行控制,直至与所述控制指令相对应的特征信息全部被提取完毕。

4. 根据权利要求3所述的基于图像识别的控制方法,其特征在于,所述继续获取所述显示界面的图像信息并提取所述至少部分未提取到的与所述控制指令相对应的特征信息,包括,

继续获取所述显示界面的图像信息,并与之前获取的图像信息进行比对,识别所述图像信息中发生变化的图像区域,并从所述图像区域中提取所述至少部分未提取到的与所述控制指令相对应的特征信息。

5. 根据权利要求1所述的基于图像识别的控制方法,其特征在于,所述根据所述控制指令获取所述显示界面的图像信息,并提取所述图像信息中与所述控制指令相对应的特征信息,包括,

根据所述控制指令对所述显示界面进行视频录制,并提取所述视频中与所述控制指令相对应的特征信息。

6. 根据权利要求5所述的基于图像识别的控制方法,其特征在于,所述提取所述视频中与所述控制指令相对应的特征信息包括,提取所述视频的视频关键帧中与所述控制指令相对应的特征信息;所述方法还包括,

当从所述视频关键帧中提取到所有所述控制指令相对应的特征信息时,停止对所述显示界面的视频录制。

7. 根据权利要求1所述的基于图像识别的控制方法,其特征在于,如果无法从所述显示界面的图像信息中提取到至少部分与所述控制指令相对应的特征信息,则继续获取所述显示界面的图像信息,并提取所述图像信息中与所述控制指令相对应的特征信息,直至提取

到全部与所述控制指令相对应的特征信息；根据所述控制指令对所述特征信息进行控制。

8. 根据权利要求2所述的基于图像识别的控制方法，其特征在于，所述获取显示界面的状态信息包括，获取应用程序在所述显示界面的显示状态信息；所述状态信息发生变化包括，所述显示界面显示的应用程序发生了变化或者应用程序的当前显示界面发生了变化。

9. 根据权利要求1所述的基于图像识别的控制方法，其特征在于，所述特征信息包括，文字控制按钮区域、图形控制按钮区域、文字输入区域；

所述获取对所述显示界面的控制指令包括，获取语音指令、服务器下发的指令、第三方传输的指令或者系统自动生成的指令；

所述根据所述控制指令对所述特征信息进行控制包括，对所述特征信息进行点击、滑动、文字输入操作；

所述根据所述控制指令对所述特征信息进行控制之后，所述方法还包括，向用户反馈控制结果。

10. 根据权利要求1所述的基于图像识别的控制方法，其特征在于，所述特征信息还包括在所述显示界面的坐标位置信息；

所述方法还包括，获取用户在控制所述特征信息时的操作信息，并在根据所述控制指令对所述特征信息进行控制时模拟用户的操作信息；

其中，所述操作信息包括点击动作信息、滑动动作信息、文字输入动作信息。

11. 一种计算机可读存储介质，其上存储有计算机程序，其特征在于，所述计算机程序被执行时实现权利要求1-10任一项所述的基于图像识别的控制方法。

12. 一种基于图像识别的控制系统，其特征在于，包括：

控制指令获取模块，用于获取控制指令；

图像信息获取模块，用于在所述控制指令获取模块获取到所述控制指令后，获取所述显示界面的图像信息，并提取所述图像信息中与所述控制指令相对应的特征信息；

特征信息控制模块，用于根据所述控制指令对所述特征信息进行控制。

13. 根据权利要求12所述的基于图像识别的控制系统，其特征在于，所述系统还包括，状态信息获取模块；所述图像信息获取模块获取所述显示界面的图像信息，并提取所述图像信息中与所述控制指令相对应的特征信息包括，

所述状态信息获取模块获取显示界面的状态信息并进行判断；

如果所述状态信息发生变化，并且无法从状态信息变化后的所述显示界面的图像信息中提取到与所述控制指令相对应的特征信息，则所述图像信息获取模块根据状态信息变化前的所述显示界面的图像信息提取与所述控制指令相对应的特征信息。

14. 根据权利要求12所述的基于图像识别的控制系统，其特征在于，如果所述图像信息获取模块无法从所述显示界面的图像信息中提取到至少部分与所述控制指令相对应的特征信息，则所述特征信息控制模块根据所述控制指令对已提取到的与所述控制指令相对应的特征信息进行控制；

所述图像信息获取模块继续获取所述显示界面的图像信息并提取所述至少部分未提取到的与所述控制指令相对应的特征信息；

所述特征信息控制模块根据所述控制指令对新提取到的与所述控制指令相对应的特征信息进行控制，直至与所述控制指令相对应的特征信息全部被提取完毕。

15. 根据权利要求14所述的基于图像识别的控制系统,其特征在于,所述图像信息获取模块继续获取所述显示界面的图像信息并提取所述至少部分未提取到的与所述控制指令相对应的特征信息,包括,

所述图像信息获取模块继续获取所述显示界面的图像信息,并与之前获取的图像信息进行比对,识别所述图像信息中发生变化的图像区域,并从所述图像区域中提取所述至少部分未提取到的与所述控制指令相对应的特征信息。

16. 根据权利要求12所述的基于图像识别的控制系统,其特征在于,所述图像信息获取模块根据所述控制指令获取所述显示界面的图像信息,并提取所述图像信息中与所述控制指令相对应的特征信息,包括,

所述图像信息获取模块根据所述控制指令对所述显示界面进行视频录制,并提取所述视频中与所述控制指令相对应的特征信息。

17. 根据权利要求16所述的基于图像识别的控制系统,其特征在于,所述提取所述视频中与所述控制指令相对应的特征信息包括,提取所述视频的视频关键帧中与所述控制指令相对应的特征信息;所述系统还包括,

当从所述视频关键帧中提取到所有所述控制指令相对应的特征信息时,停止对所述显示界面的视频录制。

18. 根据权利要求12所述的基于图像识别的控制系统,其特征在于,如果所述图像信息获取模块无法从所述显示界面的图像信息中提取到至少部分与所述控制指令相对应的特征信息,则继续获取所述显示界面的图像信息,并提取所述图像信息中与所述控制指令相对应的特征信息,直至提取到全部与所述控制指令相对应的特征信息;根据所述控制指令对所述特征信息进行控制。

19. 根据权利要求13所述的基于图像识别的控制系统,其特征在于,所述状态信息获取模块获取显示界面的状态信息包括,所述状态信息获取模块获取应用程序在所述显示界面的显示状态信息;所述状态信息发生变化包括,所述显示界面显示的应用程序发生了变化或者应用程序的当前显示界面发生了变化。

20. 根据权利要求12所述的基于图像识别的控制系统,其特征在于,所述特征信息包括,文字控制按钮区域、图形控制按钮区域、文字输入区域;

所述控制指令获取模块获取对所述显示界面的控制指令包括,获取语音指令、服务器下发的指令、第三方传输的指令或者系统自动生成的指令;

所述特征信息控制模块根据所述控制指令对所述特征信息进行控制包括,对所述特征信息进行点击、滑动、文字输入操作;

所述特征信息控制模块根据所述控制指令对所述特征信息进行控制之后,所述系统还包括,反馈模块,用于向用户反馈控制结果。

21. 根据权利要求12所述的基于图像识别的控制系统,其特征在于,所述特征信息还包括在所述显示界面的坐标位置信息;

所述系统还包括,用户操作信息获取模块,用于获取用户在控制所述特征信息时的操作信息,所述特征信息控制模块在根据所述控制指令对所述特征信息进行控制时模拟用户的操作信息;

其中,所述操作信息包括点击动作信息、滑动动作信息、文字输入动作信息。

22. 一种车辆, 其特征在于, 包括显示装置和如权利要求12-21所述的基于图像识别的控制系统。

基于图像识别的控制方法和系统及车辆、存储介质

技术领域

[0001] 本发明涉及应用控制技术领域,尤其是涉及一种基于图像识别的控制方法,以及计算机可读存储介质和基于图像识别的控制系统以及车辆。

背景技术

[0002] 近些年来,随着技术的不断发展,图像识别在生活中的应用也日趋广泛。但是,发明人发现,现有技术中,图像识别的用户使用体验不够完善。

发明内容

[0003] 本发明旨在至少解决现有技术中存在的技术问题之一。为此,本发明的一个目的在于提出一种基于图像识别的控制方法,该方法旨在一定程度上解决现有技术存在的不足。

[0004] 本发明的目的之二在于提出一种计算机可读存储介质。

[0005] 本发明的目的之三在于提出一种基于图像识别的控制系统。

[0006] 本发明的目的之四在于提出一种车辆。

[0007] 为了解决上述问题,本发明第一方面实施例的基于图像识别的控制方法,包括:获取控制指令;在获取到所述控制指令后,获取所述显示界面的图像信息,并提取所述图像信息中与所述控制指令相对应的特征信息;根据所述控制指令对所述特征信息进行控制。

[0008] 根据本发明实施例提供的基于图像识别的控制方法,在获取到控制指令后,获取显示界面的图像信息并提取图像信息中与控制指令相对应的特征信息,进而根据控制指令对特征信息进行控制,由此,在一定程度上提升了图像识别的用户使用体验。。

[0009] 在一些实施例中,

所述在获取到所述控制指令后,获取所述显示界面的图像信息,并提取所述图像信息中与所述控制指令相对应的特征信息包括,

在获取到所述控制指令后,获取显示界面的状态信息并进行判断;

如果所述状态信息发生变化,并且无法从状态信息变化后的所述显示界面的图像信息中提取到与所述控制指令相对应的特征信息,则根据状态信息变化前的所述显示界面的图像信息提取与所述控制指令相对应的特征信息。

[0010] 在一些实施例中,如果无法从所述显示界面的图像信息中提取到至少部分与所述控制指令相对应的特征信息,则根据所述控制指令对已提取到的与所述控制指令相对应的特征信息进行控制;

所述方法还包括,

继续获取所述显示界面的图像信息并提取所述至少部分未提取到的与所述控制指令相对应的特征信息;

根据所述控制指令对新提取到的与所述控制指令相对应的特征信息进行控制,直至与所述控制指令相对应的特征信息全部被提取完毕。

[0011] 在一些实施例中,所述继续获取所述显示界面的图像信息并提取所述至少部分未提取到的与所述控制指令相对应的特征信息,包括,

继续获取所述显示界面的图像信息,并与之前获取的图像信息进行比对,识别所述图像信息中发生变化的图像区域,并从所述图像区域中提取所述至少部分未提取到的与所述控制指令相对应的特征信息。

[0012] 在一些实施例中,所述根据所述控制指令获取所述显示界面的图像信息,并提取所述图像信息中与所述控制指令相对应的特征信息,包括,

根据所述控制指令对所述显示界面进行视频录制,并提取所述视频中与所述控制指令相对应的特征信息。

[0013] 在一些实施例中,所述提取所述视频中与所述控制指令相对应的特征信息包括,提取所述视频的视频关键帧中与所述控制指令相对应的特征信息;所述方法还包括,

当从所述视频关键帧中提取到所有所述控制指令相对应的特征信息时,停止对所述显示界面的视频录制。

[0014] 在一些实施例中,如果无法从所述显示界面的图像信息中提取到至少部分与所述控制指令相对应的特征信息,则继续获取所述显示界面的图像信息,并提取所述图像信息中与所述控制指令相对应的特征信息,直至提取到全部与所述控制指令相对应的特征信息;根据所述控制指令对所述特征信息进行控制。

[0015] 在一些实施例中,所述获取显示界面的状态信息包括,获取应用程序在所述显示界面的显示状态信息;所述状态信息发生变化包括,所述显示界面显示的应用程序发生了变化或者应用程序的当前显示界面发生了变化。

[0016] 在一些实施例中,所述特征信息包括,文字控制按钮区域、图形控制按钮区域、文字输入区域;

所述获取对所述显示界面的控制指令包括,获取语音指令、服务器下发的指令、第三方传输的指令或者系统自动生成的指令;

所述根据所述控制指令对所述特征信息进行控制包括,对所述特征信息进行点击、滑动、文字输入操作;

所述根据所述控制指令对所述特征信息进行控制之后,所述方法还包括,向用户反馈控制结果。

[0017] 在一些实施例中,所述特征信息还包括在所述显示界面的坐标位置信息;

所述方法还包括,获取用户在控制所述特征信息时的操作信息,并在根据所述控制指令对所述特征信息进行控制时模拟用户的操作信息;

其中,所述操作信息包括点击动作信息、滑动动作信息、文字输入动作信息。

[0018] 本发明第二方面实施例的计算机可读存储介质,其上存储有计算机程序,所述计算机程序被执行时实现上述实施例所述的基于图像识别的控制方法。

[0019] 本发明第三方面实施例的基于图像识别的控制系统,包括:控制指令获取模块,用于获取控制指令;图像信息获取模块,在所述控制指令获取模块获取到所述控制指令后,获取所述显示界面的图像信息,并提取所述图像信息中与所述控制指令相对应的特征信息;特征信息控制模块,用于根据所述控制指令对所述特征信息进行控制。

[0020] 根据本发明实施例提供的基于图像识别的控制系统,在获取到控制指令后,获取

显示界面的图像信息并提取图像信息中与控制指令相对应的特征信息,进而根据控制指令对特征信息进行控制,由此,在一定程度上提升了图像识别的用户使用体验。

[0021] 本发明第四方面实施例的车辆,包括显示装置和上述实施例所述的基于图像识别的控制系统。

[0022] 根据本发明实施例的车辆,通过采用上面实施例的基于图像识别的控制系统,在一定程度上提升了图像识别的用户使用体验。

[0023] 本发明的附加方面和优点将在下面的描述中部分给出,部分将从下面的描述中变得明显,或通过本发明的实践了解到。

附图说明

[0024] 本发明的上述和/或附加的方面和优点从结合下面附图对实施例的描述中将变得明显和容易理解,其中:

图1是根据本发明的一个实施例的基于图像识别的控制方法流程图;

图2是根据本发明的一个实施例的应用于车载中控显示屏的控制方法流程图;

图3是根据本发明的一个实施例的坐标位置示意图;

图4是根据本发明的一个实施例的基于图像识别的控制系统示意图;

图5是根据本发明的一个实施例的应用于车载中控显示屏的控制系统示意图;

图6是根据本发明的一个实施例车辆示意图。

具体实施方式

[0025] 为使本发明实施例的目的、技术方案和优点更加清楚,下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整的描述,显然,所描述的实施例是本发明一部分实施例,而不是全部实施例。基于本发明中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其它实施例,都属于本发明保护的范围。

[0026] 需要说明的是,在不冲突的情况下,本申请中的实施例及实施例中的特征可以相互组合。

[0027] 本发明可以在有计算机执行的计算机可执行指令的一般上下文中描述,例如程序模块,一般的,程序模块包括执行特定任务或实现特定抽象数据类型的例程、程序、对象、元件、数据结构等等。也可以在分布式计算环境中实践本发明,在这些分布式计算环境中,通过通信网络而被连接的远程处理设备来执行任务。在分布式计算环境中,程序模块可以位于包括存储设备在内的本地和远程计算机存储介质中。

[0028] 在本发明中,“模块”、“装置”、“系统”等指应用于计算机的相关实体,如硬件、硬件和软件的结合、软件或执行中的软件等。详细的说,例如,元件可以,但不限于运行于处理器的过程、处理器、对象、可执行元件、执行线程、程序和/或计算机。还有,运行于服务器上的应用程序或脚本程序、服务器都可以是元件。一个或多个元件可在执行的过程和/或线程中,并且元件可以在一台计算机上本地化和/或分布在两台或多台计算机之间,并可以由各种计算机可读介质运行,元件还可以根据具有一个或多个数据包的信号,例如,来自一个与本地系统、分布式系统中另一元件交互的,和/或在因特网的网络通过信号与其它系统交互的数据的信号通过本地和/或远程过程来进行通信。

[0029] 最后,还需要说明的是,在本文中,诸如第一和第二等之类的关系术语仅仅用来将一个实体或者操作与另一个实体或操作区分开来,而不一定要求或者暗示这些实体或操作之间存在任何这种实际的关系或者顺序。而且,术语“包括”、“包含”,不仅包括那些要素,而且还包括没有明确列出的其他要素,或者是还包括为这种过程、方法、物品或者设备所固有的要素。在没有更多限制的情况下,由语句“包括……”限定的要素,并不排除在包括要素的过程、方法、物品或者设备中还存在另外的相同要素。

[0030] 本发明实施例中的基于图像识别的控制方法,对应于一种计算机程序产品,该计算机程序产品安装于智能终端设备上,用于实现对安装于智能终端设备上的第三方应用程序的语音控制(在无需对第三方应用程序进行定制或适配性调试或修改的情况下即可实现对其的语音控制),该智能终端上配置有显示屏或者该终端设备能够投影出显示界面用于用户进行交互操作,例如,智能手机、平板电脑、PC、车载终端、智能家居、投影仪等任何智能硬件,本发明对此不作限定。

[0031] 本发明提供的基于图像识别的控制技术方案可以应用于对显示界面应用软件的控制,例如控制音乐播放软件、地图导航软件等的打开、界面切换、应用移除等操作,也可以应用于显示界面内容的控制,例如图片的处理等。

[0032] 如图1所示为本发明的一个实施例提供的一种基于图像识别的控制方法流程图。本发明实施例的基于图像识别的控制方法至少包括步骤S101-S104。

[0033] 步骤S101,获取控制指令。

[0034] 具体的,本方案提供的实施例中,首先获取对于显示界面的控制指令。其中,显示界面可以是显示屏幕的显示界面,也可以是投影的显示界面。以车载中控显示屏为例,显示界面可以为地图导航界面、音乐广播界面、游戏界面等,也可以是应用程序的当前界面,例如音乐播放器的音乐列表界面、音乐搜索界面、音乐播放界面等。获取控制指令,该控制指令可以来自于用户的语音控制指令,也可以是来自于服务器等发送的控制指令。控制指令例如:导航到M景点,下载邮件附件N等。例如,从用户发送的一段语音指令中提取到对特征信息(点击、滑动等)的控制指令。

[0035] 步骤S102,根据控制指令获取显示界面的图像信息并提取图像信息中与控制指令相对应的特征信息。

[0036] 具体的,在获取到对显示界面的控制指令后,根据控制指令获取显示界面的图像信息,并提取图像信息中与控制指令相对应的特征信息。获取显示界面的图像信息可以采用的方式为对当前显示界面进行截图,也可以是只针对部分显示界面进行截图。例如,显示界面可能显示有多种不同类型的信息、不同应用程序的信息。例如,对于车载中控显示屏,其可以同时显示天气信息、车辆信息(包括车内温度、续航等)、导航信息、音乐播放信息以及多媒体信息(例如微信信息、微博信息等)等,有些显示内容属于常驻显示内容,例如显示车辆信息等,有些应用程序不会在显示界面全屏显示,只占用部分显示界面。因此,当显示界面的状态信息发生变化时,例如打开了一个新的应用程序,但是该应用程序只占用了部分显示界面,这时,如果依然采用获取全部显示界面的图像信息并提取特征信息,会导致信息的冗余处理,造成算力的浪费,同时,可能造成信息处理时间的不必要延长,耽误用户的时间。

[0037] 其中,提取图像信息中与控制指令相对应的图像信息的特征信息,特征信息可以

包括文字形式的控制按钮、图形形式的控制按钮、文字输入区域等等。例如,提取图像信息中的返回图标控制按钮信息、上一页/下一页控制按钮信息、快进/快退按钮信息、进度条、文字输入框等信息。其中,可以理解的,这里的按钮可以是圆形、矩形、其它不规则按钮等形式,或者直接为文字形式;也可以是非常规按钮形式,例如,现在普遍采用的左滑进入上一页、右滑进入下一页、左侧上/下滑调整亮度、右侧上/下滑调整音量、双击播放/暂停等,在这种情况下,本实施例的按钮也可以是非常规按钮形式的控制按钮,非常规按钮形式的控制按钮信息可以通过如下方式获得:分析应用程序或者应用程序在显示界面的显示坐标区域等信息,按照用户常规的左右滑动、双击操作、上下滑动等操作动作,确定左右滑动、双击操作、上下滑动等的操作坐标区域,并将该操作坐标区域及操作动作作为图像信息的特征信息。由此,实现了对于显示界面的图像信息中与控制指令相对应的特征信息的全面提取。

[0038] 其中,提取与图像信息中与控制指令相对应的特征信息,可以通过大数据自学习的方式来提高效率。例如,控制指令为返回上一页的指令,则可以通过大数据自学习的方式,获取返回操作的常规按钮形式、常规文字形式或者常规操作区域形式,由此,在对图像信息进行处理并提取与该指令相对应的特征信息时,可以更具针对性的对图像中特定区域、特定符号标示的内容进行识别,由此,可以更快速更精确的识别到与控制指令相对应的特征信息,避免了毫无目的的对图像进行全面分析、全面识别,数据处理量更少,效率更高。

[0039] 步骤S103,根据控制指令对特征信息进行控制。

[0040] 具体的,根据控制指令对提取到的特征信息进行控制。例如,用户的语音控制指令为播放A歌星的B歌曲,此时,将根据控制指令对特征信息中的音乐搜索框进行文字输入并检索,并播放检索到的B歌曲。

[0041] 根据本发明实施例提供的基于图像识别方法,通过图像识别,寻找与控制指令相对应的特征信息并进行控制,不需要对应用程序进行定制或者适配性调试或者修改,改善了用户的应用体验。

[0042] 其中,可以理解的,本发明部分实施例是以用户语音控制为例进行说明,但是本发明的应用领域并不局限于用户语音控制技术领域,还可以应用于多种可以代替用户操作的领域,例如,购票时,可以帮助用户自动刷新、自动输入验证码、自动下单并退出等一系列模拟用户的操作。本发明的基于图像识别的控制方法,通过图像识别,寻找特征信息并进行控制,不需要对应用程序进行定制或者适配性调试或者修改,即可根据接收到的控制指令对特征信息进行控制,提升了智能化体验。

[0043] 其中,本发明实施例提供的基于图像识别的控制方法,根据接收到的控制指令对显示界面的图像信息进行获取。即显示界面图像信息的获取响应于下达的控制指令,相较于普通的获取显示界面的图像信息并提取特征信息,更具有针对性,且时效性更高,避免信息的无效冗余处理。同时,对图像信息的处理以及特征信息的提取与控制指令的内容密切相关,即只提取与控制指令相对应的特征信息,针对性更强,能进一步避免数据的冗余处理,进一步提升了处理效率。

[0044] 在一些实施例中,获取显示界面的图像信息并提取图像信息中与控制指令相对应的特征信息包括,

获取显示界面的状态信息并进行判断;

如果状态信息发生变化,并且无法从状态信息变化后的显示界面的图像信息中提取到

与控制指令相对应的特征信息,则根据状态信息变化前的显示界面的图像信息提取与控制指令相对应的特征信息。

[0045] 具体的,在获取到对显示界面的控制指令后,开始获取对显示界面的图像信息。如果在获取显示界面的图像信息时,发现显示界面的状态信息发生了变化。例如,显示界面显示的应用程序由导航应用切换为了音乐播放应用,则可以判断状态信息发生了变化;又或者音乐播放应用的当前显示界面由播放界面变为音乐搜索界面,则可以判断为状态信息发生了变化。当判断到显示界面的状态信息发生变化后,例如,由原来的导航应用变为音乐应用时,获取的图像信息变成了音乐应用界面的图像信息,如果控制指令是针对原来的导航应用的,比如导航到M景点。但是,在音乐应用界面的图像信息中无法提取与该控制指令相对应的特征信息。此时,将根据状态信息变化之前的显示界面图像信息来提取与该控制指令相对应的特征信息,即提取导航应用界面的特征信息。具体的,获取该状态信息变化之前的显示界面的图像信息,可以是状态信息变化之前获取的。这是因为图像信息的获取速度一般比较快,例如采用截图等方式,在显示界面状态变化时,截取到了变化之前的图像信息。也可以采用将显示界面的状态信息调整到变化之前的状态,并获取图像信息。例如,将此时显示界面显示的内容重新调整到导航应用界面,此时,获取导航应用界面的图像信息,并提取图像信息中与控制指令(导航到M景点)相对应的特征信息。满足了用户的需求。其中,可以理解的,将显示界面的状态信息调整到变化之前的状态,也可以采用后台调整的方式,例如,显示界面当前显示的音乐应用界面保持不变,只是在后台虚拟一个显示界面,并在该虚拟的显示界面显示导航应用,进而获取该虚拟显示界面的导航应用的图像信息。由此,既保证了对导航应用的图像信息的获取,又不影响用户当前对音乐应用的使用。在获取了导航应用的图像信息后,获取该图像信息与控制指令(导航到M景点)相对应的控制指令,并在后台根据控制指令对导航应用进行控制。由此,既实现了用户对导航应用控制的需求,又不影响用户对当前音乐界面的使用。提升了用户的使用体验。

[0046] 现有技术中,用户希望通过语音下达指令来对应用程序做出控制。但是,应用软件需要进行适配等修改才能实现语音控制的功能。根据本发明实施例提供的技术方案,通过图像识别技术,提取图像信息中的特征信息,进而根据控制指令对特征信息做出控制。由此,在不需要对软件进行适配性修改等的情况下,能够对绝大部分常规软件应用进行图像识别特征信息,进而根据语音控制指令做出控制。

[0047] 在一些实施例中,如果无法从显示界面的图像信息中提取到至少部分与控制指令相对应的特征信息,则根据控制指令对已提取到的与控制指令相对应的特征信息进行控制;

方法还包括,

继续获取显示界面的图像信息并提取至少部分未提取到的与控制指令相对应的特征信息;

根据控制指令对新提取到的与控制指令相对应的特征信息进行控制,直至与控制指令相对应的特征信息全部被提取完毕。

[0048] 具体的,本发明提供的实施例中,在获取到对显示界面的控制指令后,获取显示界面的图像信息并提取图像信息中与控制指令相对应的特征信息,并根据控制指令对特征信息进行控制。在实际中,有可能会出现无法从图像信息中提取全部与控制指令相对应的特

征信息。例如,如果控制指令涉及到多个步骤,而该多个步骤分别对应多个不同的界面,此时,将根据控制指令对已提取到的特征信息进行控制,并继续获取显示界面的图像信息以及从图像信息中提取与前述未提取到的与控制指令相对应的特征信息。进而根据控制指令对新提取到的特征信息进行控制,直至所有与控制指令相对应的特征信息全部被提取完毕。

[0049] 例如,控制指令为打开K视频应用搜索并播放L电影。此时,将根据该控制指令获取显示界面的图像信息,从图像信息中提取K视频应用的位置坐标信息,但是与控制指令“播放L电影”相对应的特征信息此时无法提取到。此时,将根据获取的特征信息“K视频应用的位置坐标信息”,模拟用户点击该位置坐标,从而打开K视频应用。打开K视频应用后,此时将继续获取显示界面的图像信息,即K视频应用的显示界面图像信息,并从中提取到搜索框的位置坐标信息,此时执行搜索指令,即模拟用户点击搜索框输入“L电影”并点击检索,此时,K视频应用进入检索结果页面,此时继续获取显示界面的图像信息并从中提取L电影所在的位置信息,模拟用户点击,完成播放动作。由此,完成用户的全部指令。可以理解的,用户的指令可能是简短的,即部分操作步骤被跳过,此时,将根据大数据自学习进行自适应的步骤补全,并执行该步骤指令。例如,用户的指令为打开K视频并播放L电影,此时将自动补充搜索L电影的步骤,以在打开K视频步骤执行完后执行搜索L电影的步骤,从而保证播放L电影的步骤得以完成。

[0050] 在一些实施例中,继续获取显示界面的图像信息并提取至少部分未提取到的与控制指令相对应的特征信息,包括,

继续获取显示界面的图像信息,并与之前获取的图像信息进行比对,识别图像信息中发生变化的图像区域,并从图像区域中提取至少部分未提取到的与控制指令相对应的特征信息。

[0051] 具体的,为提升信息处理效率,避免信息的重复冗余处理,本发明实施例还提供了一种方案,包括,继续获取显示界面的图像信息并与之前获取的图像信息进行比对,识别图像信息中发生变化的图像区域,提取该图像区域中未提取到的与控制指令相对应的特征信息。这是由于很多情况下,显示界面只有部分显示内容发生了变化,此时,只需要对发生变化的部分进行处理,并提取图像区域中未提取到的与控制指令相对应的特征信息。由此可以大大降低了信息处理量,节约数据处理时间,减少用户的等待时间,进一步提升了用户的使用体验。

[0052] 在一些实施例中,根据控制指令获取显示界面的图像信息,并提取图像信息中与控制指令相对应的特征信息,包括,

根据控制指令对显示界面进行视频录制,并提取视频中与控制指令相对应的特征信息。

[0053] 具体的,获取显示界面的图像信息可以采用对显示界面进行视频录制。例如,在获取到对显示界面的控制指令后,开始对显示界面进行视频录制,并提取视频中的特征信息。视频相比于图片的形式,可以获取更多的图像信息,并且图像信息为连续的,由此可以获取更多的更全面的并且连续的特征信息。例如,用户打开了视频应用C,视频应用C加载完成后进入了视频应用C的首页,用户操作视频应用C进入到视频应用C的个人中心,打开播放历史,滑动到上次的播放内容并点击了继续播放。由此,当用户下达继续播放的指令时,或者

根据大数据判断用户需要继续播放时,将根据获取的连续特征信息进行自动连续操作,不需要用户或者系统等一步步的下达操作指令,改善了用户的操作体验,提升了智能化程度。

[0054] 在一些实施例中,提取视频中与控制指令相对应的特征信息包括,提取视频的视频关键帧中与控制指令相对应的特征信息;方法还包括,

当从视频关键帧中提取到所有控制指令相对应的特征信息时,停止对显示界面的视频录制。

[0055] 具体的,在获取到对显示界面的控制指令时,开始进行视频录制,并提取视频中与控制指令相对应的特征信息。其中,可以仅对视频中的关键帧进行特征信息的提取。关键帧的确定可以是系统设定的,比如每隔一段时间提取视频中的一帧并将其确定为关键帧。也可以是大数据自动分析,当画面中关键部分出现了变动,即可将此时的视频画面作为关键帧。通过选取关键帧进行特征信息提取,可以显著提升处理效率,降低对算力的占用,节约处理时间。

[0056] 在一些实施例中,如果无法从显示界面的图像信息中提取到至少部分与控制指令相对应的特征信息,则继续获取显示界面的图像信息,并提取图像信息中与控制指令相对应的特征信息,直至提取到全部与控制指令相对应的特征信息;根据控制指令对特征信息进行控制。

[0057] 具体的,本发明提供的一个实施例中,如果在获取到控制指令后,开始获取显示界面的图像信息并提取与控制指令相对应的特征信息,如果发现无法从图像信息中提取到至少部分与控制指令相对应的特征信息,此时将继续获取图像信息,并从中提取特征信息,直至与控制指令相对应的特征信息已全部被提取完毕。由此,可以保证控制指令被完整正确的执行。例如,控制指令为导航到M景点,但是此时导航应用正在加载,没有完全被加载完毕,此时,与控制指令相对应的特征信息可能无法被全部提取完毕。此时,将待导航应用全部加载完毕时,可以从图像信息中提取到全部与控制指令相对应的特征信息,然后再根据控制指令对特征信息进行控制。由此,保证了指令的全部正确执行,避免出错给用户带来不好的使用体验。

[0058] 在一些实施例中,获取显示界面的状态信息包括,获取应用程序在显示界面的显示状态信息;状态信息发生变化包括,显示界面显示的应用程序发生了变化或者应用程序的当前显示界面发生了变化。

[0059] 具体的,显示界面的状态信息可以是显示界面当前所展示的应用程序界面。例如,本发明的显示界面为车载中控显示屏的显示界面,显示界面的状态信息为当前显示界面展示的是应用程序界面,可以是地图导航界面、音乐广播界面、游戏界面等,也可以是应用程序的当前界面,例如音乐播放器的音乐列表界面、音乐搜索界面、音乐播放界面等。获取显示界面的状态信息并进行判断,可以通过获取当前正在运行的应用程序信息,应用程序在显示界面的显示信息,应用程序当前显示界面的信息等。

[0060] 在获取显示界面的状态信息后分析状态信息是否发生变化。例如,在获取到显示界面当前正在显示的应用程序或者应用程序当前显示界面信息后,与之前获取的状态信息进行分析比对,判断显示的应用程序是否发生了变化或者应用程序的显示界面是否发生了变化。例如,显示界面显示的应用程序由导航应用切换为了音乐播放应用,则可以判断状态信息发生了变化;又或者音乐播放应用的当前显示界面由播放界面变为音乐搜索界面,则

可以判断为状态信息发生了变化。

[0061] 在一些实施例中,特征信息包括,文字控制按钮区域、图形控制按钮区域、文字输入区域;

获取控制指令包括,获取语音指令、服务器下发的指令、第三方传输的指令或者系统自动生成的指令;

根据控制指令对特征信息进行控制包括,对特征信息进行点击、滑动、文字输入操作;

根据控制指令对特征信息进行控制之后,方法还包括,向用户反馈控制结果。

[0062] 具体的,特征信息可以是操作控制区域信息,例如点击、滑动等操作区域信息,还可以包括文字控制按钮区域、图形控制按钮区域、文字输入区域等。

[0063] 控制指令可以是用户的语音控制指令、服务器下发的指令、通过第三方传输的指令(例如通过网络或者U盘等传输的控制指令)以及系统自动生成的控制指令等。

[0064] 根据控制指令对特征信息进行控制包括点击、滑动、文字输入等操作形式。

[0065] 其中,本发明提供的实施例,还包括对控制指令的执行结果进行反馈。

[0066] 在一些实施例中,特征信息还包括在显示界面的坐标位置信息;

方法还包括,获取用户在控制特征信息时的操作信息,并在根据控制指令对特征信息进行控制时模拟用户的操作信息;

其中,操作信息包括点击动作信息、滑动动作信息、文字输入动作信息。

[0067] 具体的,特征信息包括在显示界面的坐标位置信息,例如,返回按钮在显示界面的坐标位置信息,由此,通过提取的坐标位置信息,可以模拟用户的操作,例如,模拟用户的滑动、点击、文字输入等操作信息。

[0068] 本发明还提供了一种应用于车载中控显示屏的具体实施例,如图2所示,包括步骤S201-S204:

步骤S201、车载中控显示屏的显示界面进入系统某界面。

[0069] 具体的,由车载中控显示屏的控制系统根据应用程序的运行情况来判断当前显示屏的显示界面是否发生改变,是否进入到系统某界面,并以此作为触发条件,触发后续的识别控制等操作。

[0070] 步骤S202、当车载中控显示屏的控制系统检测到车载中控显示屏的当前显示界面已经进入某界面并且停留时间超过2秒,可以认为用户已完整获取了界面内容,此时开始进入视频录制,否则则忽略当前界面的录制,以接下来进入的界面重新作为触发条件。

[0071] 具体的,由于不同的用户对于界面的感知情况有差异,同时界面的刷新加载速度也有差异,所以视频录制的触发点,会进一步准确到当前界面的内容绘制完毕才开始录制,减少无效录制时间。

[0072] 步骤S203、视频实时录制并同步上传服务器,服务器获取到本地上传的视频文件后,对每一秒的视频中的图片进行识别,包括确定界面文字或图标内容以及对应文字或图标所处的显示界面中的位置坐标。

[0073] 具体的,虽然进入某界面可能整体上的内容不会发生太多的变化,但有时会因为后台操作或者界面部分区域进行刷新,所以需要录制整个界面的所有时长的视频并进行文字或图标识别,保证即使出现上述情况,也可以保证新出现的内容也可以进行识别。其中,图标的识别可以采用识别后转换为文字的形式,以便与用户的语音控制指令相匹配。例如,

表示返回的图标可能有多种展现形式,识别后将统一转换为“返回”的文本,当用户下达返回指令时,将与该“返回”的文本进行匹配,并进行控制。

[0074] 步骤S204、如果用户进行语音操作,则分析语义结果,并与视频的识别结果进行匹配,若匹配成功则执行对应语义,并模拟用户对显示界面进行操作控制;当系统获知当前已经进入某页面超过60秒,且用户未进行语音控制操作,则表示用户已经无需进行进一步的界面控制,此时停止视频录制

具体的,如当用户说出“播放歌曲”时,如果以目前的技术而言需要对音乐应用进行API的适配或者调试,才能在语音控制的时候进行音乐应用的控制。根据本发明实施例提供的方法,可以通过识别当前显示界面中的歌曲播放按钮及其对应的显示界面坐标,进而通过系统模拟用户点击则可完成语音操作。

[0075] 具体的,如图3所示,获取到目标识别区域的左上角以及右下角坐标,进行该区域的中点位置计算,得到坐标 $((x_2+x_1)/2, (y_2+y_1)/2)$,然后通过系统模拟用户点击则可完成语音操作。

[0076] 本发明第二方面实施例提出一种计算机可读存储介质,其上存储有计算机程序,其中,该计算机程序被执行时实现上面实施例的基于图像识别的控制方法。

[0077] 基于上面实施例的基于图像识别的控制系统,下面描述本发明实施例第三方面提出的基于图像识别的控制系统。

[0078] 如图4所示,本发明实施例的基于图像识别的控制系统,包括,
控制指令获取模块,用于获取控制指令;

具体的,本方案提供的实施例中,首先通过控制指令获取模块获取对于显示界面的控制指令。其中,显示界面可以是显示屏幕的显示界面,也可以是投影的显示界面。以车载中控显示屏为例,显示界面可以为地图导航界面、音乐广播界面、游戏界面等,也可以是应用程序的当前界面,例如音乐播放器的音乐列表界面、音乐搜索界面、音乐播放界面等。获取控制指令,该控制指令可以来自于用户的语音控制指令,也可以是来自于服务器等发送的控制指令。控制指令例如:导航到M景点,下载邮件附件N等。例如,从用户发送的一段语音指令中提取到对特征信息(点击、滑动等)的控制指令。

[0079] 图像信息获取模块,用于根据控制指令获取显示界面的图像信息并提取图像信息中与控制指令相对应的特征信息;

具体的,在获取到对显示界面的控制指令后,图像信息获取模块获取显示界面的图像信息,并提取图像信息中与控制指令相对应的特征信息。图像信息获取模块获取显示界面的图像信息可以采用的方式为对当前显示界面进行截图,也可以是只针对部分显示界面进行截图。例如,显示界面可能显示有多种不同类型的信息、不同应用程序的信息。例如,对于车载中控显示屏,其可以同时显示天气信息、车辆信息(包括车内温度、续航等)、导航信息、音乐播放信息以及多媒体信息(例如微信信息、微博信息等)等,有些显示内容属于常驻显示内容,例如显示车辆信息等,有些应用程序不会在显示界面全屏显示,只占用部分显示界面。因此,当显示界面的状态信息发生变化时,例如打开了一个新的应用程序,但是该应用程序只占用了部分显示界面,这时,如果依然采用获取全部显示界面的图像信息并提取特征信息,会导致信息的冗余处理,造成算力的浪费,同时,可能造成信息处理时间的不必要延长,耽误用户的时间。

[0080] 其中,提取图像信息中与控制指令相对应的图像信息的特征信息,特征信息可以包括文字形式的控制按钮、图形形式的控制按钮、文字输入区域等等。例如,提取图像信息中的返回图标控制按钮信息、上一页/下一页控制按钮信息、快进/快退按钮信息、进度条、文字输入框等信息。其中,可以理解的,这里的按钮可以是圆形、矩形、其它不规则按钮等形式,或者直接为文字形式;也可以是非常规按钮形式,例如,现在普遍采用的左滑进入上一页、右滑进入下一页、左侧上/下滑调整亮度、右侧上/下滑调整音量、双击播放/暂停等,在这种情况下,本实施例的按钮也可以是非常规按钮形式的控制按钮,非常规按钮形式的控制按钮信息可以通过如下方式获得:分析应用程序或者应用程序在显示界面的显示坐标区域等信息,按照用户常规的左右滑动、双击操作、上下滑动等操作动作,确定左右滑动、双击操作、上下滑动等的操作坐标区域,并将该操作坐标区域及操作动作作为图像信息的特征信息。由此,实现了对于显示界面的图像信息中与控制指令相对应的特征信息的全面提取。

[0081] 其中,提取与图像信息中与控制指令相对应的特征信息,可以通过大数据自学习的方式来提高效率。例如,控制指令为返回上一页的指令,则可以通过大数据自学习的方式,获取返回操作的常规按钮形式、常规文字形式或者常规操作区域形式,由此,在对图像信息进行处理并提取与该指令相对应的特征信息时,可以更具针对性的对图像中特定区域、特定符号标示的内容进行识别,由此,可以更快速更精确的识别到与控制指令相对应的特征信息,避免了毫无目的的对图像进行全面分析、全面识别,数据处理量更少,效率更高。

[0082] 特征信息控制模块,根据控制指令对特征信息进行控制。

[0083] 具体的,特征信息控制模块根据控制指令对提取到的特征信息进行控制。例如,用户的语音控制指令为播放A歌星的B歌曲,此时,将根据控制指令对特征信息中的音乐搜索框进行文字输入并检索,并播放检索到的B歌曲。

[0084] 根据本发明实施例提供的基于图像识别系统,通过图像识别,寻找与控制指令相对应的特征信息并进行控制,不需要对应用程序进行定制或者适配性调试或者修改,改善了用户的应用体验。

[0085] 其中,可以理解的,本发明部分实施例是以用户语音控制为例进行说明,但是本发明的应用领域并不局限于用户语音控制技术领域,还可以应用于多种可以代替用户操作的领域,例如,购票时,可以帮助用户自动刷新、自动输入验证码、自动下单并退出等一系列模拟用户的操作。本发明的基于图像识别的控制系统,通过图像识别,寻找特征信息并进行控制,不需要对应用程序进行定制或者适配性调试或者修改,即可根据接收到的控制指令对特征信息进行控制,提升了智能化体验。

[0086] 其中,本发明实施例提供的基于图像识别的控制系统,根据接收到的控制指令对显示界面的图像信息进行获取。即显示界面图像信息的获取响应于下达的控制指令,相较于普通的获取显示界面的图像信息并提取特征信息,更具有针对性,且时效性更高,避免信息的无效冗余处理。同时,对图像信息的处理以及特征信息的提取与控制指令的内容密切相关,即只提取与控制指令相对应的特征信息,针对性更强,能进一步避免数据的冗余处理,进一步提升了处理效率。

[0087] 本发明提供的基于图像识别的控制技术方案可以应用于对显示界面应用软件的控制,例如控制音乐播放软件、地图导航软件等的打开、界面切换、应用移除等操作,也可以应用于显示界面界面内容的控制,例如图片的处理等。

[0088] 现有技术中,用户希望通过语音下达指令来对应用程序做出控制。但是,应用软件需要进行适配等修改才能实现语音控制的功能。根据本发明实施例提供的技术方案,通过图像识别技术,提取图像信息中的特征信息,进而根据控制指令对特征信息做出控制。由此,在不需要对软件进行适配性修改等的情况下,能够对绝大部分常规软件应用进行图像识别特征信息,进而根据语音控制指令做出控制。

[0089] 在一些实施例中,系统还包括,状态信息获取模块;图像信息获取模块根据控制指令获取显示界面的图像信息,并提取图像信息中与控制指令相对应的特征信息包括,

状态信息获取模块获取显示界面的状态信息并进行判断;

如果状态信息发生变化,并且无法从状态信息变化后的显示界面的图像信息中提取到与控制指令相对应的特征信息,则图像信息获取模块根据状态信息变化前的显示界面的图像信息提取与控制指令相对应的特征信息。

[0090] 具体的,控制指令获取模块在获取到对显示界面的控制指令后,图像信息获取模块开始获取对显示界面的图像信息。如果在获取显示界面的图像信息时,状态信息获取模块发现显示界面的状态信息发生了变化。例如,显示界面显示的应用程序由导航应用切换为了音乐播放应用,则可以判断状态信息发生了变化;又或者音乐播放应用的当前显示界面由播放界面变为音乐搜索界面,则可以判断为状态信息发生了变化。当判断到显示界面的状态信息发生变化后,例如,由原来的导航应用变为音乐应用时,获取的图像信息变成了音乐应用界面的图像信息,如果控制指令是针对原来的导航应用的,比如导航到M景点。但是,在音乐应用界面的图像信息中无法提取与该控制指令相对应的特征信息。此时,将根据状态信息变化之前的显示界面图像信息来提取与该控制指令相对应的特征信息,即提取导航应用界面的特征信息。具体的,获取该状态信息变化之前的显示界面的图像信息,可以是状态信息变化之前获取的。这是因为图像信息的获取速度一般比较快,例如采用截图等方式,在显示界面状态变化时,截取到了变化之前的图像信息。也可以采用将显示界面的状态信息调整到变化之前的状态,并获取图像信息。例如,将此时显示界面显示的内容重新调整到导航应用界面,此时,获取导航应用界面的图像信息,并提取图像信息中与控制指令(导航到M景点)相对应的特征信息。满足了用户的需求。其中,可以理解的,将显示界面的状态信息调整到变化之前的状态,也可以采用后台调整的方式,例如,显示界面当前显示的音乐应用界面保持不变,只是在后台虚拟一个显示界面,并在该虚拟的显示界面显示导航应用,进而获取该虚拟显示界面的导航应用的图像信息。由此,既保证了对导航应用的图像信息的获取,又不影响用户当前对音乐应用的使用。在获取了导航应用的图像信息后,获取该图像信息与控制指令(导航到M景点)相对应的控制指令,并在后台根据控制指令对导航应用进行控制。由此,既实现了用户对导航应用控制的需求,又不影响用户对当前音乐界面的使用。提升了用户的使用体验。

[0091] 在一些实施例中,如果图像信息获取模块无法从显示界面的图像信息中提取到至少部分与控制指令相对应的特征信息,则特征信息控制模块根据控制指令对已提取到的与控制指令相对应的特征信息进行控制;

图像信息获取模块继续获取显示界面的图像信息并提取至少部分未提取到的与控制指令相对应的特征信息;

特征信息控制模块根据控制指令对新提取到的与控制指令相对应的特征信息进行控

制,直至与控制指令相对应的特征信息全部被提取完毕。

[0092] 具体的,本发明提供的实施例中,在控制指令获取模块获取到对显示界面的控制指令后,图像信息获取模块获取显示界面的图像信息并提取图像信息中与控制指令相对应的特征信息,特征信息控制模块根据控制指令对特征信息进行控制。在实际中,有可能会无法从图像信息中提取全部与控制指令相对应的特征信息。例如,如果控制指令涉及到多个步骤,而该多个步骤分别对应多个不同的界面,此时,将根据控制指令对已提取到的特征信息进行控制,并继续获取显示界面的图像信息以及从图像信息中提取与前述未提取到的与控制指令相对应的特征信息。进而根据控制指令对新提取到的特征信息进行控制,直至所有与控制指令相对应的特征信息全部被提取完毕。

[0093] 例如,控制指令为打开K视频应用搜索并播放L电影。此时,将根据该控制指令获取显示界面的图像信息,从图像信息中提取K视频应用的位置坐标信息,但是与控制指令“播放L电影”相对应的特征信息此时无法提取到。此时,将根据获取的特征信息“K视频应用的位置坐标信息”,模拟用户点击该位置坐标,从而打开K视频应用。打开K视频应用后,此时将继续获取显示界面的图像信息,即K视频应用的显示界面图像信息,并从中提取到搜索框的位置坐标信息,此时执行搜索指令,即模拟用户点击搜索框输入“L电影”并点击检索,此时,K视频应用进入检索结果页面,此时继续获取显示界面的图像信息并从中提取L电影所在的位置信息,模拟用户点击,完成播放动作。由此,完成用户的全部指令。可以理解的,用户的指令可能是简短的,即部分操作步骤被跳过,此时,将根据大数据自学习进行自适应的步骤补全,并执行该步骤指令。例如,用户的指令为打开K视频并播放L电影,此时将自动补充搜索L电影的步骤,以在打开K视频步骤执行完后执行搜索L电影的步骤,从而保证播放L电影的步骤得以完成。

[0094] 在一些实施例中,图像信息获取模块继续获取显示界面的图像信息并提取至少部分未提取到的与控制指令相对应的特征信息,包括,

图像信息获取模块继续获取显示界面的图像信息,并与之前获取的图像信息进行比对,识别图像信息中发生变化的图像区域,并从图像区域中提取至少部分未提取到的与控制指令相对应的特征信息。

[0095] 具体的,为提升信息处理效率,避免信息的重复冗余处理,本发明实施例还提供了一种方案,包括,图像信息获取模块继续获取显示界面的图像信息并与之前获取的图像信息进行比对,识别图像信息中发生变化的图像区域,提取该图像区域中未提取到的与控制指令相对应的特征信息。这是由于很多情况下,显示界面只有部分显示内容发生了变化,此时,只需要对发生变化的部分进行处理,并提取图像区域中未提取到的与控制指令相对应的特征信息。由此可以大大降低了信息处理量,节约数据处理时间,减少用户的等待时间,进一步提升了用户的使用体验。

[0096] 在一些实施例中,图像信息获取模块根据控制指令获取显示界面的图像信息,并提取图像信息中与控制指令相对应的特征信息,包括,

图像信息获取模块根据控制指令对显示界面进行视频录制,并提取视频中与控制指令相对应的特征信息。

[0097] 具体的,图像信息获取模块获取显示界面的图像信息可以采用对显示界面进行视频录制。例如,在获取到对显示界面的控制指令后,开始对显示界面进行视频录制,并提取

视频中的特征信息。视频相比于图片的形式,可以获取更多的图像信息,并且图像信息为连续的,由此可以获取更多的更全面的并且连续的特征信息。例如,用户打开了视频应用C,视频应用C加载完成后进入了视频应用C的首页,用户操作视频应用C进入到视频应用C的个人中心,打开播放历史,滑动到上次的播放内容并点击了继续播放。由此,当用户下达继续播放的指令时,或者根据大数据判断用户需要继续播放时,将根据获取的连续特征信息进行自动连续操作,不需要用户或者系统等一步步的下达操作指令,改善了用户的操作体验,提升了智能化程度。

[0098] 在一些实施例中,提取视频中与控制指令相对应的特征信息包括,提取视频的视频关键帧中与控制指令相对应的特征信息;系统还包括,

当从视频关键帧中提取到所有控制指令相对应的特征信息时,停止对显示界面的视频录制。

[0099] 具体的,在控制指令获取模块获取到对显示界面的控制指令时,图像信息获取模块开始进行视频录制,并提取视频中与控制指令相对应的特征信息。其中,可以仅对视频中的关键帧进行特征信息的提取。关键帧的确定可以是系统设定的,比如每隔一段时间提取视频中的一帧并将其确定为关键帧。也可以是大数据自动分析,当画面中关键部分出现了变动,即可将此时的视频画面作为关键帧。通过选取关键帧进行特征信息提取,可以显著提升处理效率,降低对算力的占用,节约处理时间。

[0100] 在一些实施例中,如果图像信息获取模块无法从显示界面的图像信息中提取到至少部分与控制指令相对应的特征信息,则继续获取显示界面的图像信息,并提取图像信息中与控制指令相对应的特征信息,直至提取到全部与控制指令相对应的特征信息;根据控制指令对特征信息进行控制。

[0101] 具体的,本发明提供的一个实施例中,如果在控制指令获取模块获取到控制指令后,图像信息获取模块开始获取显示界面的图像信息并提取与控制指令相对应的特征信息,如果发现无法从图像信息中提取到至少部分与控制指令相对应的特征信息,此时将继续获取图像信息,并从中提取特征信息,直至与控制指令相对应的特征信息已全部被提取完毕。由此,可以保证控制指令被完整正确的执行。例如,控制指令为导航到M景点,但是此时导航应用正在加载,没有完全被加载完毕,此时,与控制指令相对应的特征信息可能无法被全部提取完毕。此时,将待导航应用全部加载完毕时,可以从图像信息中提取到全部与控制指令相对应的特征信息,然后再根据控制指令对特征信息进行控制。由此,保证了指令的全部正确执行,避免出错给用户带来不好的使用体验。

[0102] 在一些实施例中,状态信息获取模块获取显示界面的状态信息包括,状态信息获取模块获取应用程序在显示界面的显示状态信息;状态信息发生变化包括,显示界面显示的应用程序发生了变化或者应用程序的当前显示界面发生了变化。

[0103] 具体的,显示界面的状态信息可以是显示界面当前所展示的应用程序界面。例如,本发明的显示界面为车载中控显示屏的显示界面,显示界面的状态信息为当前显示界面展示的是应用程序界面,可以是地图导航界面、音乐广播界面、游戏界面等,也可以是应用程序的当前界面,例如音乐播放器的音乐列表界面、音乐搜索界面、音乐播放界面等。获取显示界面的状态信息并进行判断,可以通过获取当前正在运行的应用程序信息,应用程序在显示界面的显示信息,应用程序当前显示界面的信息等。

[0104] 在状态信息获取模块获取显示界面的状态信息后分析状态信息是否发生变化。例如,在获取到显示界面当前正在显示的应用程序或者应用程序当前显示界面信息后,与之前获取的状态信息进行分析比对,判断显示的应用程序是否发生了变化或者应用程序的显示界面是否发生了变化。例如,显示界面显示的应用程序由导航应用切换为了音乐播放应用,则可以判断状态信息发生了变化;又或者音乐播放应用的当前显示界面由播放界面变为音乐搜索界面,则可以判断为状态信息发生了变化。

[0105] 在一些实施例中,特征信息包括,文字控制按钮区域、图形控制按钮区域、文字输入区域;

控制指令获取模块获取控制指令包括,获取语音指令、服务器下发的指令、第三方传输的指令或者系统自动生成的指令;

特征信息控制模块根据控制指令对特征信息进行控制包括,对特征信息进行点击、滑动、文字输入操作;

特征信息控制模块根据控制指令对特征信息进行控制之后,系统还包括,反馈模块,用于向用户反馈控制结果。

[0106] 具体的,特征信息可以是操作控制区域信息,例如点击、滑动等操作区域信息,还可以包括文字控制按钮区域、图形控制按钮区域、文字输入区域等。

[0107] 控制指令可以是用户的语音控制指令、服务器下发的指令、通过第三方传输的指令(例如通过网络或者U盘等传输的控制指令)以及系统自动生成的控制指令等。

[0108] 根据控制指令对特征信息进行控制包括点击、滑动、文字输入等操作形式。

[0109] 其中,本发明提供的实施例,还包括对控制指令的执行结果进行反馈。

[0110] 在一些实施例中,特征信息还包括在显示界面的坐标位置信息;

系统还包括,用户操作信息获取模块,用于获取用户在控制特征信息时的操作信息,特征信息控制模块在根据控制指令对特征信息进行控制时模拟用户的操作信息;

其中,操作信息包括点击动作信息、滑动动作信息、文字输入动作信息。

[0111] 具体的,特征信息包括在显示界面的坐标位置信息,例如,返回按钮在显示界面的坐标位置信息,由此,通过提取的坐标位置信息,可以模拟用户的操作,例如,模拟用户的滑动、点击、文字输入等操作信息。

[0112] 本发明还提供了一种应用于车载中控显示屏具体实施例,如图5所示,包括:

界面状态信息获取单元,获取显示界面的界面变化,并根据界面的加载完成情况,通知视频录制单元进行视频录制的开始以及停止。

[0113] 具体的,界面状态信息获取单元不仅仅监听界面的变化情况,同时提供给用户对视频录制的响应速度的设置界面,包括,开始录制视频的时间点,以及视频录制的分辨率(分辨率过高,数据量太大,处理时间太长;分辨率过低,识别准确度不高)。

[0114] 视频录制单元,接收界面状态信息获取单元的视频录制控制信号,同时与服务器进行通信(包括将录制的视频即时上传服务器)。

[0115] 此外,视频录制单元,不仅负责系统界面的视频录制,而且还承担着本地视频文件的管理工作,定时删除本地视频缓存,避免对本地存储空间的过度占用。

[0116] 视频录制单元还用于接收服务器的图像识别结果,并将对应结果存储到本地进行备用。

[0117] 当用户进行语音操作的时候,语音指令获取单元获取用户的语音指令并进行语义识别,语音指令执行单元调取图片识别结果,并根据识别结果执行用户的语音指令。

[0118] 具体的,若当前语音指令与图片识别结果有匹配的部分,则语音指令执行单元执行对应的语音控制指令。

[0119] 其中,语音指令执行单元在执行用户的语音控制指令时,根据图像识别结果中的坐标信息,模拟用户操作。例如模拟用户点击、滑动等操作。

[0120] 其中,如果当前的坐标信息有误差无法执行操作,则会反馈到服务器端进行视频图片的重新标记训练。

[0121] 反馈单元,负责接收系统执行语音命令的执行结果,以文字、图片、语音、铃声等提示手段进行用户提示。

[0122] 具体的,会根据不同的场景进行反馈提示,如果当前处于音乐播放场景,有一定的背景音乐,且当前的语音指令会带来听觉或者视觉上的变化,则语音反馈提醒以文字为主,不进行声音提醒;如果当前语音指令会带来体感(如温度,风量)的变化也是一样,以文字提醒为主,用户无需声音即可感知到语音控制的变化。如此,避免了对用户的打扰,提升了用户体验。

[0123] 本发明实施例提供的基于图像识别的控制系统,根据视频录制获取视频流中的图片数据,全程无需用户点击,自动识别文字以及对应文字的坐标,在用户语音命令发出后,对命令中和当前界面内容直接相同或者相关的进行识别,执行相应语音操作,如打开视频应用后界面中包含“电视剧”的文字,用户可直接发出语音指令“打开电视剧”即可打开,模拟用户点击操作,方便用户操作,在用户开车过程中尤为实用,同时可适配各类应用,无须重新开放适配。

[0124] 本发明实施例第四方面提出一种车辆,如图6所示,本发明实施例的车辆包括显示装置和实施例中基于图像识别的控制系统。例如,显示装置可以包括车载中控显示屏、抬头显示HUD等。

[0125] 根据本发明实施例的车辆,通过采用上面实施例的基于图像识别的控制系统,不需要对软件进行适配性调试,用户也可以对软件进行语音控制,提升了用户的使用体验。

[0126] 在本说明书的描述中,参考术语“一个实施例”、“一些实施例”、“示意性实施例”、“示例”、“具体示例”、或“一些示例”等的描述意指结合该实施例或示例描述的具体特征、结构、材料或者特点包含于本发明的至少一个实施例或示例中。在本说明书中,对上述术语的示意性表述不一定指的是相同的实施例或示例。

[0127] 尽管已经示出和描述了本发明的实施例,本领域的普通技术人员可以理解:在不脱离本发明的原理和宗旨的情况下可以对这些实施例进行多种变化、修改、替换和变型,本发明的范围由权利要求及其等同物限定。

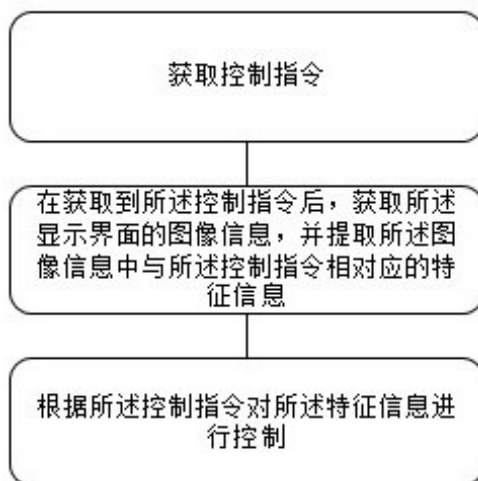


图1

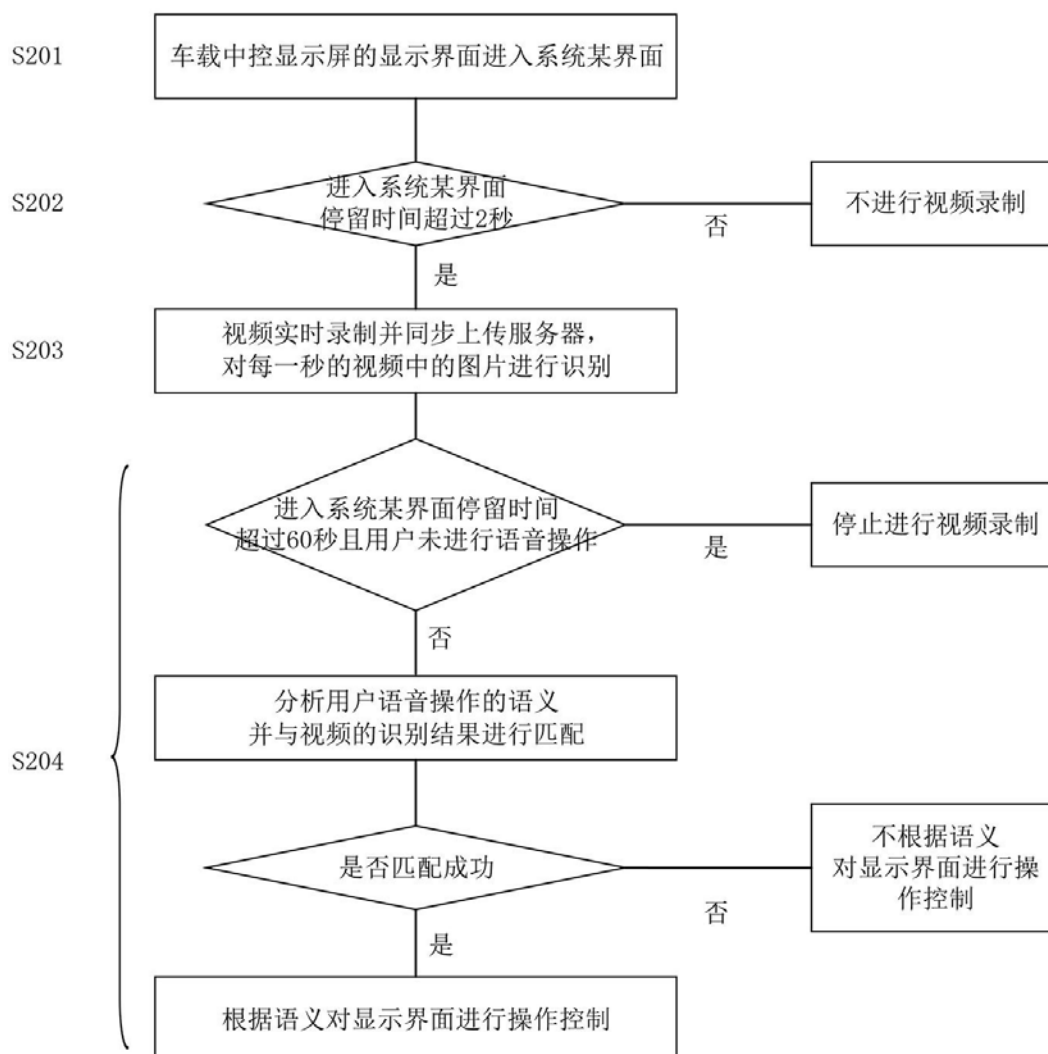


图2

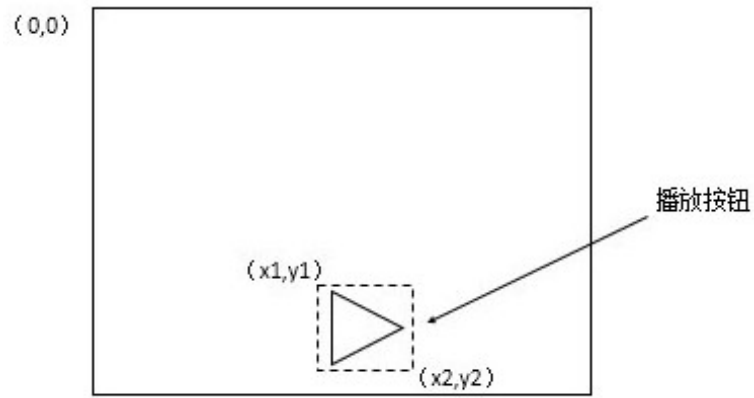


图3

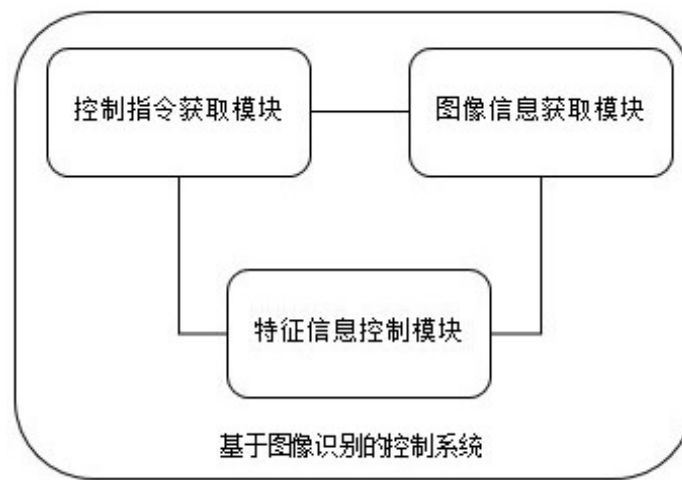


图4

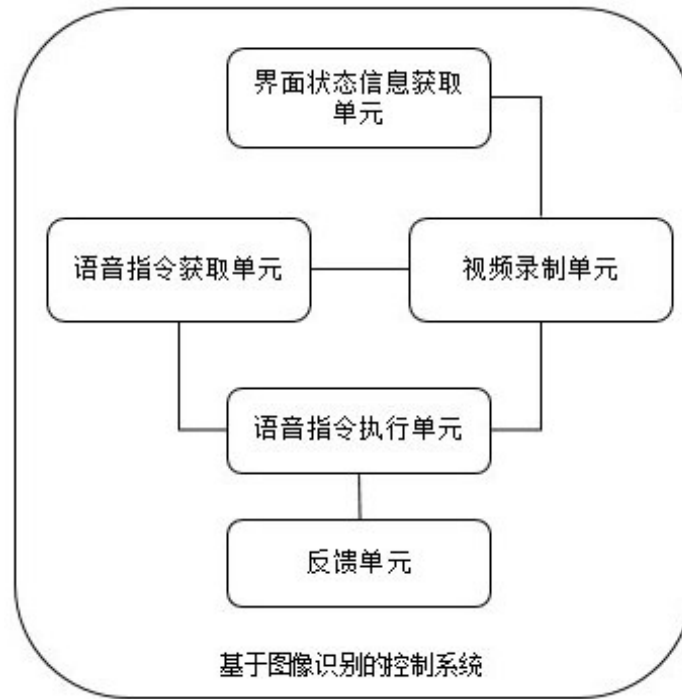


图5

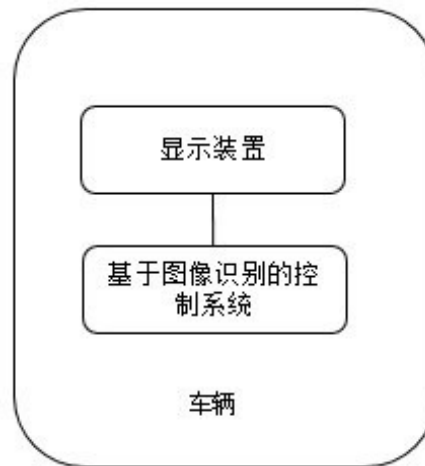


图6