



(12) 发明专利

(10) 授权公告号 CN 108198552 B

(45) 授权公告日 2021.02.02

(21) 申请号 201810048538.8

G10L 15/18 (2013.01)

(22) 申请日 2018.01.18

G02B 27/01 (2006.01)

(65) 同一申请的已公布的文献号

(56) 对比文件

申请公布号 CN 108198552 A

CN 101149922 A, 2008.03.26

(43) 申请公布日 2018.06.22

CN 103247291 A, 2013.08.14

(73) 专利权人 深圳市大疆创新科技有限公司

CN 103646646 A, 2014.03.19

地址 518057 广东省深圳市南山区高新区

CN 106469552 A, 2017.03.01

南区粤兴一道9号香港科大深圳产学研

CN 1659624 A, 2005.08.24

研大楼6楼

CN 107112010 A, 2017.08.29

(72) 发明人 庞磊 刘平

JP 2014119559 A, 2014.06.30

(74) 专利代理机构 广州三环专利商标代理有限公司

US 2016196821 A1, 2016.07.07

公司 44202

US 9697822 B1, 2017.07.04

代理人 郝传鑫 熊永强

US 8983823 B1, 2015.03.17

(51) Int. Cl.

EP 3125236 A1, 2017.02.01

G10L 15/22 (2006.01)

审查员 徐联微

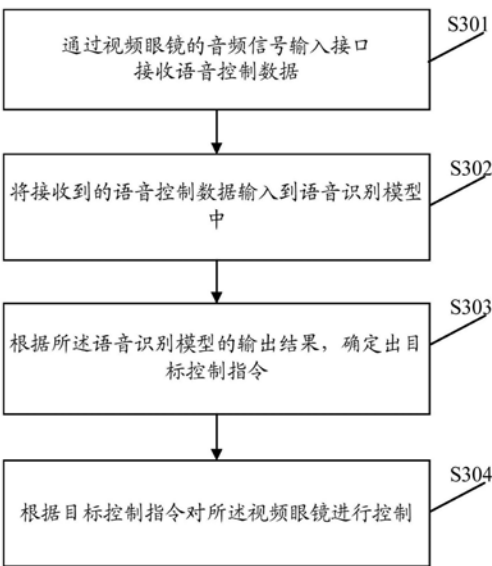
权利要求书2页 说明书10页 附图6页

(54) 发明名称

一种语音控制方法及视频眼镜

(57) 摘要

本发明实施例提供了一种语音控制方法及视频眼镜,其中方法包括:通过所述视频眼镜的音频信号输入接口接收语音控制数据;将接收到的语音控制数据输入到语音识别模型中,所述语音识别模型用于对所述语音控制数据进行分析识别;根据所述语音识别模型的输出结果,确定出目标控制指令;根据目标控制指令对所述视频眼镜进行控制,可以在一定程度上提高视频眼镜的使用方便性。



1. 一种语音控制方法,其特征在于,应用于视频眼镜,所述方法包括:
通过所述视频眼镜的音频信号输入接口接收第二语音训练数据;
输出语音识别结果预设界面,将在所述语音识别结果预设界面上选择的语音识别结果作为预设语音识别结果;
将所述第二语音训练数据输入到预置的语音识别初始模型中,所述语音识别初始模型用于对所述第二语音训练数据进行分析识别;
如果所述语音识别初始模型的输出结果与所述预设语音识别结果不相同,则对所述语音识别初始模型进行优化更新,以便于得到语音识别模型;
通过所述视频眼镜的音频信号输入接口接收语音控制数据;
将接收到的语音控制数据输入到所述语音识别模型中,所述语音识别模型用于对所述语音控制数据进行分析识别;
根据所述语音识别模型的输出结果,确定出目标控制指令;
根据目标控制指令对所述视频眼镜进行控制。
2. 根据权利要求1所述的方法,其特征在于,所述输出结果为关键字;所述根据所述语音识别模型的输出结果,确定出目标控制指令,包括:
根据所述语音识别模型输出的关键字以及预置的关键字与控制指令之间的对应关系,确定出所述语音识别模型输出的关键字对应的目标控制指令。
3. 根据权利要求1所述的方法,其特征在于,所述视频眼镜包括语音控制触发按钮;在所述通过所述视频眼镜的音频信号输入接口接收语音控制数据之前,所述方法还包括:
当所述视频眼镜处于待机模式时,接收对所述语音控制触发按钮的第一用户操作;
根据所述第一用户操作,控制所述视频眼镜进入语音控制模式。
4. 根据权利要求1所述的方法,其特征在于,在所述通过所述视频眼镜的音频信号输入接口接收语音控制数据之前,所述方法还包括:
当所述视频眼镜处于待机模式时,接收第二用户操作;
如果所述第二用户操作是对所述视频眼镜的开关按钮的操作,则在所述视频眼镜开启后,控制所述视频眼镜进入语音控制模式。
5. 根据权利要求1所述的方法,其特征在于,在所述根据所述语音识别模型的输出结果,确定出目标控制指令之后,所述方法还包括:
检测所述视频眼镜当前所处的工作状态;
判断所述目标控制指令所指示的控制状态是否与所述工作状态相匹配;
若是,则执行所述根据目标控制指令对所述视频眼镜进行控制。
6. 一种视频眼镜,其特征在于,所述视频眼镜包括:存储器、处理器;
所述存储器,用于存储程序指令;
所述处理器,用于执行所述存储器存储的程序指令,当程序指令被执行时,用于:
通过所述视频眼镜的音频信号输入接口接收第二语音训练数据;
输出语音识别结果预设界面,将在所述语音识别结果预设界面上选择的语音识别结果作为预设语音识别结果;
将所述第二语音训练数据输入到预置的语音识别初始模型中,所述语音识别初始模型用于对所述第二语音训练数据进行分析识别;

如果所述语音识别初始模型的输出结果与所述预设语音识别结果不相同,则对所述语音识别初始模型进行优化更新,以便于得到语音识别模型;

通过所述视频眼镜的音频信号输入接口接收语音控制数据;

将接收到的语音控制数据输入到所述语音识别模型中,所述语音识别模型用于对所述语音控制数据进行分析识别;

根据所述语音识别模型的输出结果,确定出目标控制指令;

根据目标控制指令对所述视频眼镜进行控制。

7. 根据权利要求6所述的视频眼镜,其特征在于,所述输出结果为关键字;所述处理器用于根据所述语音识别模型的输出结果,确定出目标控制指令时,具体用于:

根据所述语音识别模型输出的关键字以及预置的关键字与控制指令之间的对应关系,确定出所述语音识别模型输出的关键字对应的目标控制指令。

8. 根据权利要求6所述的视频眼镜,其特征在于,所述视频眼镜包括语音控制触发按钮;

所述处理器还用于:

当所述视频眼镜处于待机模式时,接收对所述语音控制触发按钮的第一用户操作;

根据所述第一用户操作,控制所述视频眼镜进入语音控制模式。

9. 根据权利要求6所述的视频眼镜,其特征在于,所述处理器还用于:

当所述视频眼镜处于待机模式时,接收第二用户操作;

如果所述第二用户操作是对所述视频眼镜的开关按钮的操作,则在所述视频眼镜开启后,控制所述视频眼镜进入语音控制模式。

10. 根据权利要求6所述的视频眼镜,其特征在于,所述处理器还用于:

检测所述视频眼镜当前所处的工作状态;

判断所述目标控制指令所指示的控制状态是否与所述工作状态相匹配;

若是,则执行所述根据目标控制指令对所述视频眼镜进行控制。

一种语音控制方法及视频眼镜

技术领域

[0001] 本发明涉及电子技术领域,尤其涉及一种语音控制方法及视频眼镜。

背景技术

[0002] 视频眼镜,是一种可以直接穿戴在使用者身上的便携式设备,它可以通过软件支持、数据交互以及云端交互等来实现强大的功能,大大方便了人们的生活、学习与感知。

[0003] 一些配置了遮目装置的视频眼镜,例如增强现实 (Augmented Reality, AR) 眼镜、虚拟现实 (Virtual Reality, VR) 眼镜、飞行控制眼镜、智能头盔、智能头带等,使用者在使用这类设备的过程中,由于外界光线被遮挡,无法看到视频眼镜上控制按钮的位置,在操作视频眼镜时十分不便。

[0004] 因此,如何有效的提高视频眼镜的使用方便性成为了一个亟待解决的问题。

发明内容

[0005] 本发明实施例公开了一种语音控制方法以及视频眼镜,可以在一定程度上提高视频眼镜的使用方便性。

[0006] 本发明实施例第一方面公开了一种语音控制方法,应用于视频眼镜,所述方法包括:

[0007] 通过所述视频眼镜的音频信号输入接口接收语音控制数据;

[0008] 将接收到的语音控制数据输入到语音识别模型中,所述语音识别模型用于对所述语音控制数据进行分析识别;

[0009] 根据所述语音识别模型的输出结果,确定出目标控制指令;

[0010] 根据目标控制指令对所述视频眼镜进行控制。

[0011] 本发明实施例第二方面公开了一种视频眼镜,包括:处理器、存储器;

[0012] 所述存储器,用于存储程序指令;

[0013] 所述处理器,用于执行所述存储器存储的程序指令,当程序指令被执行时,用于:

[0014] 通过所述视频眼镜的音频信号输入接口接收语音控制数据;

[0015] 将接收到的语音控制数据输入到语音识别模型中,所述语音识别模型用于对所述语音控制数据进行分析识别;

[0016] 根据所述语音识别模型的输出结果,确定出目标控制指令;

[0017] 根据目标控制指令对所述视频眼镜进行控制。

[0018] 本发明实施例中,视频眼镜可以接收语音控制数据,并将接收到的语音控制数据输入到语音识别模型中,根据语音识别模型的输出结果,确定出目标控制指令,最后根据目标控制指令对该视频眼镜进行控制,可以通过语音识别的方式实现对视频眼镜的控制,在一定程度上提高视频眼镜的使用方便性。

附图说明

[0019] 为了更清楚地说明本发明实施例中的技术方案,下面将对实施例中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本发明的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动性的前提下,还可以根据这些附图获得其他的附图。

[0020] 图1a为本发明实施例提供的一种用于语音控制的情景示意图;

[0021] 图1b为本发明实施例提供的视频眼镜的一种结构示意图;

[0022] 图1c为本发明实施例提供的视频眼镜的另一种结构示意图;

[0023] 图2为本发明实施例提供的一种用于语音控制的流程示意图;

[0024] 图3为本发明实施例所提供的另一种语音控制方法的流程示意图;

[0025] 图4为本发明实施例所提供的又一种语音控制方法的流程示意图;

[0026] 图5为本发明实施例提供的一种视频眼镜的结构示意图。

具体实施方式

[0027] 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、完整地描述。

[0028] 针对配置了遮目装置的视频眼镜(例如AR眼镜、VR眼镜、飞行控制眼镜等),目前需要通过使用者用手来控制视频眼镜上的控制按钮来对该视频眼镜进行控制,但在视频眼镜的使用过程中,使用者由于外界光线被遮挡,导致使用者在其视线范围内难以看到控制按钮的位置,尤其是在视频眼镜的控制按钮都相对较小的情况下,使用者及其容易选错按钮,这极大的限制了使用者触发功能的便利性,视频眼镜的使用方便性被极大的降低。

[0029] 为了解决上述技术问题,本发明实施例提供了一种语音控制方法及视频眼镜。

[0030] 下面请参阅图1a,为本发明实施例提供的一种用于语音控制的情景示意图。使用者可以通过头戴视频眼镜10来显示虚拟场景,或者使虚拟场景与现实场景相融合。

[0031] 在一个实施例中,该视频眼镜10可以配置有遮目装置,该视频眼镜10可以采用放大镜成像方式或光场成像方式来实现屏幕内容的显示。

[0032] 在一个实施例中,本发明实施例可以基于语音识别技术。语音识别技术可以使用人工设计特征或深度学习等方法来完成对语音的分类,使视频眼镜可以确定出使用者的指令,并对使用者的指令做出辨析。

[0033] 在一个实施例中,该视频眼镜10可以基于深度学习来对语音数据进行训练。例如,可以通过采集大量各个国家和地区的数据,将这些数据进行训练,得到语音识别模型,通过该语音识别模型对使用者输入的语音控制数据进行分析识别,可以准确识别出使用者的控制指令(也即目标控制指令),满足各种人群在各种场景下的准确识别。

[0034] 下面请参阅图1b以及图1c,图1b为本发明实施例提供的视频眼镜的一种结构示意图,图1c为图1b中的视频眼镜的俯视图。如图1b所示,本发明实施例提供的视频眼镜包括第一显示屏10a、第二显示屏10b、第一光学镜片20a、第二光学镜片20b、以及设置在第一显示屏10a与第一光学镜片20a之间的第一目镜30a、设置在第二显示屏10b与第二光学镜片20b之间的第二目镜30b。

[0035] 其中,第一显示屏10a的显示面11a与第一目镜30a的光轴方向平行。第二显示屏

10b的显示面11b与第二目镜30b的光轴方向平行。

[0036] 第一光学镜片20a上和第二光学镜片20b上分别形成有反射面(21a、21b)。第一光学镜片20a的反射面21a朝向第一显示屏10a,且第一光学镜片20a的反射面21a与第一显示屏10a之间形成第一预设夹角 α ,以使第一显示屏10a投射到第一光学镜片20a上所成物像101a与第一目镜30a的光轴方向垂直。第二光学镜片20b的反射面朝向第二显示屏10b,且第二光学镜片20b的反射面21b与第二显示屏10b之间形成第二预设夹角 β ,以使第二显示屏10b投射到第二光学镜片20b上所成物像与第二目镜30b的光轴方向垂直。

[0037] 在本发明实施例中,第一光学镜片20a的反射面21a和第二光学镜片20b的反射面21b上可以镀有反射膜或半透半反膜,例如,当第一光学镜片20a、第二光学镜片20b只需要全反射时,该反射面可以镀有反射膜;而当第一光学镜片20a和第二光学镜片20b既需要反射又需要透射时,第一光学镜片20a的反射面21a和第二光学镜片20b的反射面21b可以镀有半反半透膜,例如,在增强现实模式下,第一光学镜片20a和第二光学镜片20b应均能部分透射部分反射,因此,第一光学镜片20a的反射面21a和第二光学镜片20b的反射面21b可以镀有半透半反膜。

[0038] 本领域技术人员能够理解的是,在光学系统中,目镜一般为将物镜所成的像放大后供眼睛观察用的光学部件,用户通过第一目镜30a和第二目镜30b看到的图像是呈放大后的图像。在本发明实施例中,优选的,第一目镜30a和第二目镜30b均为由至少一个凸透镜和至少一个凹透镜层叠布置形成的透镜组。凸透镜可以起到进一步放大图像,提高视场角,提高沉浸感的作用。而凹透镜可以起到限制视场,只允许一定范围内的光线通过透镜组的作用。

[0039] 第一显示屏10a和第二显示屏10b是具有显示功能的显示器,其具体的形状和结构并不限定,例如第一显示屏10a和第二显示屏10b可以为LCD显示屏(液晶显示屏),或者,OLED显示屏(有机发光显示屏),当然,第一显示屏10a和第二显示屏10b的具体类型也可以为其他类型,并且第一显示屏10a与第二显示屏10b两者的类型也可以互不相同。第一显示屏10a和第二显示屏10b内所显示的画面经过第一光学镜片20a和第二光学镜片20b的反射后进入人眼中,使得用户能够观看到第一显示屏10a和第二显示屏10b中的画面内容,例如,左眼看到第一显示屏10a上显示的内容,右眼看到第二显示屏10b上显示的内容,由于两眼视差,可以产生3D的感觉。

[0040] 该视频眼镜还包括遮目装置40,在一个实施例中,遮目装置40可以朝向第一光学镜片20a上与反射面21a相对的表面以及第二光学镜片20b上与反射面21b相对的表面,遮目装置40用于阻挡外界真实场景(非虚拟场景)的光线投射到第一光学镜片20a和第二光学镜片20b上。

[0041] 在一个实施例中,该视频眼镜本体50可以为如图1b所示的盒状结构,遮目装置40可以为薄板类结构,或者其他结构。当外界光线需要投射到第一光学镜片20a上与反射面21a相对的表面,以及第二光学镜片20b上与反射面21b相对的表面,则首先需要穿透遮目装置40,因此,当遮目装置40阻挡外界光线射入时,用户便无法看到眼前真实景象而沉浸在虚拟现实模式下的虚拟场景中。当然,视频眼镜本体50的结构以及遮目装置40的结构并不限于此,本领域技术人员可以根据实际需求而具体设计不同的可行结构,在此,本实施例不作赘述。

[0042] 在一个实施例中,该遮目装置40可以为透光率可调件,调节装置可以与遮目装置40电连接的调节装置(图中未示出),调节装置用于调节对遮目装置40施加的电压大小,以改变遮目装置40的透光率。在本实施例中,优选的,该遮目装置40可以为LCD液晶屏,当然,遮目装置40还可以为其他类型,例如,TN LCD液晶屏,或者电致变色玻璃等等。

[0043] 在使用过程中,第一目镜30a可以供用户的左眼观看,第二目镜30b可以供用户右眼观看。第一显示屏10a的显示面11a与第一目镜30a的光轴方向平行,即,第一显示屏10a的显示面11a与人的左眼光轴方向平行。第二显示屏10b的显示面11b与第二目镜30b的光轴方向平行,即,第二显示屏10b的显示面11b与人的右眼光轴方向平行。第一光学镜片20a的反射面21a与第一显示屏10a之间形成第一预设夹角 α ,第二光学镜片20b的反射面21b与第二显示屏10b之间形成第二预设夹角 β ,第一预设夹角 α 和第二预设夹角 β 的具体值可以具体设定,本发明实施例不具体限定,但需保证第一显示屏10a投射到第一光学镜片20a上所成物像与第一目镜30a的光轴方向垂直,第二显示屏10b投射到第二光学镜片20b上所成物像与所述第二目镜30b的光轴方向垂直,即,物像各个点到目镜平面的距离相同,从而避免物像呈梯形,影响观看效果。

[0044] 在本发明实施例中,第一目镜30a和第二目镜30b均为至少一个凸透镜和至少一个凹透镜层叠布置形成的透镜组,则第一显示屏10a和第二显示屏10b上显示的画面内容经过第一光学镜片20a和第二光学镜片20b反射后形成的物像为虚像。

[0045] 本发明实施例提供的视频眼镜,由于通过双显示屏(第一显示屏10a和第二显示屏10b)显示内容,并通过具有反射功能的第一光学镜片20a和第二光学镜片20b分别将第一显示屏10a和第二显示屏10b上的显示内容分别入射至人的左右眼中,可实现 70° 的视场角,能够实现很强的沉浸感,并且由于单眼的有效显示区域较大,目镜的放大倍数不需要很高也能达到较好的沉浸感,且用户在使用时视觉效果上的颗粒感较弱,显示效果较为细腻。而对于头盔产品来说,则可采用更大的尺寸更高分辨率的显示屏来达到更好的显示效果。

[0046] 在一个实施例中,第一显示屏10a与第二显示屏10b可以平行相对设置。更具体的,如图1b和图1c所示,第一显示屏10a与第二显示屏10b之间可以形成预设距离,且第一光学镜片20a、第二光学镜片20b、第一目镜30a以及第二目镜30b设置在第一显示屏10a与第二显示屏10b之间。第一显示屏10a与第二显示屏10b之间的放置位置大体上可以根据人脸部或者头部的宽度而设定,以使得在使用时,第一显示屏10a和第二显示屏10b能够基本平行相对地设置在人的眼睛两侧,一方面能够保证用户在使用时具有较好的观看感受,另一方面能够节省视频眼镜的结构空间,保证整个视频眼镜的体积较小。

[0047] 下面请参阅图2,为本发明实施例提供的一种用于语音控制的流程示意图。

[0048] 在101中,该视频眼镜可以处于待机模式。在一个实施例中,该待机模式可以是电源开启,但该视频眼镜处于休眠状态的模式。使用者可以佩戴上该视频眼镜,并可以说出预先设定的语音数据a。

[0049] 在102中,该视频眼镜可以通过音频信号输入接口接收到该语音数据a,并基于语音识别基础技术识别出该语音数据a的内容,确定出语音指令a。

[0050] 在一个实施例中,该语音识别基础技术可以是基于人工设计特征的语音识别技术。

[0051] 在103中,如果该语音指令a为预设的可用于切换到语音控制模式的指令,则该视

频眼镜从待机模式切换到语音控制模式。

[0052] 在一个实施例中,该视频眼镜也可以配置语音控制触发按钮。当该视频眼镜处于待机模式时,可以接收对该语音控制触发按钮的第一用户操作,根据该第一用户操作控制该视频眼镜从待机模式切换到语音控制模式。

[0053] 该第一用户操作可以是对该语音控制按钮的拨动、触控、长按、短按等等,本发明对此不作限制。

[0054] 在一个实施例中,当该视频眼镜处于待机模式时,也可以接收第二用户操作;如果该第二用户操作是对该视频眼镜的开关按钮的操作,则可以在该视频眼镜开启之后,控制该视频眼镜进入该语音控制模式。

[0055] 在104中,该视频眼镜在进入语音控制模式之后,可以接收使用者发出的语音控制数据b,并将该语音控制数据b输入到语音识别模式中,并根据语音识别模式的输出结果,识别出目标控制指令b。

[0056] 该语音识别模型可以用于对该语音控制数据进行分析识别。在一个实施例中,该语音识别模型可以是利用深度学习的方式训练出的模型,例如,该语音识别模型可以为受限玻尔兹曼机(Restricted Boltzman Machine,RBM)模型、深度神经网络(Deep Neural Network,DNN)模型、卷积神经网络(Convolutional Neural Network,CNN)模型等等,本发明实施例对此不作限制。

[0057] 在105中,该视频眼镜可以根据该目标控制指令b对该视频眼镜进行控制。

[0058] 举例来说,如果该语音识别模型识别出语音控制数据b的关键字为“弹出菜单”,则该视频眼镜可以确定出目标控制指令b为弹出菜单选项的指令,该视频眼镜可以控制弹出菜单选项。进一步的,使用者可以说出菜单中对应的内容,该视频眼镜可以识别使用者说出的内容来选择执行相应的功能。

[0059] 又举例来说,如果该语音识别模型识别出语音控制数据b的关键字为“关闭”,则该视频眼镜可以确定出目标控制指令b为关闭视频眼镜的指令,该视频眼镜可以控制进行关闭处理。

[0060] 在一些可行的实施方式中,该视频眼镜还可以通过手势识别来确定出目标控制指令,并根据目标控制指令对该视频眼镜进行控制。

[0061] 例如,该视频眼镜可以预置一个或者多个手势动作,并将手势动作与控制指令进行对应。该视频眼镜可以通过摄头实时检测是否有该手势动作,如果有,则将检测到手势动作与预置的手势动作进行匹配,确定出与该手势动作匹配的目标控制指令,并根据该目标控制指令对该视频眼镜进行控制。

[0062] 在一个实施例中,一个手势动作可以与一个控制指令对应,也可以与多个控制指令对应,本发明实施例对此不作任何限制。

[0063] 下面介绍本申请的方法实施例。需要说明的是,本申请的方法实施例可以由视频眼镜执行,该视频眼镜可以是VR眼镜、AR眼镜、飞行控制眼镜、智能头盔、智能头带等等,本发明实施例对此不作任何限制。

[0064] 请参阅图3,为本发明实施例提供的一种语音控制方法的流程示意图。如图3所示的语音控制方法可包括:

[0065] S301、通过视频眼镜的音频信号输入接口接收语音控制数据。

[0066] 该视频眼镜的音频信号输入接口,可以是麦克风、AUX(Auxiliary)接口等等可用于输入语音控制数据的接口,本发明实施例对此不作任何限制。

[0067] 在一个实施例中,该语音控制数据为使用者发出的语音数据。该视频眼镜可以实时接收该语音控制数据,或者,该视频眼镜也可以在进入语音控制模式时,接收该语音控制数据。

[0068] S302、将接收到的语音控制数据输入到语音识别模型中。

[0069] 所述语音识别模型用于对所述语音控制数据进行分析识别。

[0070] 在一个实施例中,该语音识别模型可以是利用深度学习的方式训练出的模型,例如,该语音识别模型可以为RBM模型、DNN模型、CNN模型等等,本发明实施例对此不作限制。

[0071] 在一个实施例中,训练得到该语音识别模型,可以通过以下方式:获取第一语音训练数据,并查询出所述第一语音训练数据对应的预设语音识别结果;将所述第一语音训练数据输入到预置的语音识别初始模型中,所述语音识别初始模型用于对所述第一语音训练数据进行分析识别;如果所述语音识别初始模型的输出结果与所述预设语音识别结果不相同,则对所述语音识别初始模型进行优化更新,以便于得到所述语音识别模型。

[0072] 该第一语音训练数据可以包括各个国家和地区,且拥有不同口音的使用者发出的语音数据,或者,该第一语音训练数据也可以是来自同一地区的使用者们的语音数据,本发明实施例对此不作任何限制。

[0073] 举例来说,该视频眼镜可以对该第一语音训练数据以预设的采样率进行声波采样(sampling),即可以记录等时间点的声波的高度。通过该声波采样,该视频眼镜可以生成音频文件,该音频文件可以包括每个采样时间点的声波高度的数字。

[0074] 该视频眼镜可以将该音频文件输入到预置的语音识别初始模型中,该语音识别初始模型可以根据音频文件中的频率识别出字符,识别出的字符可以作为该语音识别初始模型的输出结果。

[0075] 另一方面,该视频眼镜可以预先存储语音训练数据与语音识别结果的对应关系,并根据获取到的第一语音训练数据以及该对应关系查询出与该第一语音训练数据对应的预设语音识别结果。

[0076] 如果该语音识别初始模型的输出结果与该预设语音识别结果不相同,即可以对该语音识别初始模型的输出结果进行更正,并对该语音识别初始模型进行优化更新,通过反复进行上述训练,可以得到语音识别模型。

[0077] 在一个实施例中,该采样率可以为16k赫兹,17k赫兹,15k赫兹等可以覆盖人类语音的频率范围,本发明实施例对此不作任何限制。

[0078] 在一个实施例中,该语音识别初始模型可以保存先前预测的数据,通过先前预测的数据可以获取到语音训练数据与预设语音识别结果的对应关系。或者,该视频眼镜也可以通过其他方式保存该对应关系,本发明实施例对此不作任何限制。

[0079] 在一个实施例中,训练得到该语音识别模型,也可以通过以下方式:通过所述视频眼镜的音频信号输入接口接收第二语音训练数据;输出语音识别结果预设界面,将在所述语音识别结果选择界面上选择的语音识别结果作为预设语音识别结果;将所述第二语音训练数据输入到预置的语音识别初始模型中,所述语音识别初始模型用于对所述第二语音训练数据进行分析识别;如果所述语音识别初始模型的输出结果与所述预设语音识别结果不

相同,则对所述语音识别初始模型进行优化更新,以便于得到所述语音识别模型。

[0080] 举例来说,使用者发出第二语音训练数据,该视频眼镜可以通过音频信号输入接口接收到该第二语音训练数据,然后输出语音识别结果预设界面,该语音识别结果预设界面上可以包括多个语音识别结果,使用者可以选择其中一个语音识别结果,该视频眼镜将使用者选择出的语音识别结果作为预设语音识别结果。

[0081] 另一方面,该视频眼镜可以将该第二语音训练数据输入到预置的语音识别初始模型中进行分析识别,得到该语音识别初始模型的输出结果。

[0082] 如果该语音识别初始模型的输出结果与该预设语音识别结果不相同,即可以对该语音识别初始模型的输出结果进行更正,并对该语音识别初始模型进行优化更新,通过反复进行上述训练,可以得到语音识别模型。

[0083] S303、根据所述语音识别模型的输出结果,确定出目标控制指令。

[0084] 在一个实施例中,所述语音识别模型的输出结果为关键字。

[0085] 所述根据所述语音识别模型的输出结果,确定出目标控制指令,包括:根据所述语音识别模型输出的关键字以及预置的关键字与控制指令之间的对应关系,确定出所述语音识别模型输出的关键字对应的目标控制指令。

[0086] 举例来说,关键字“启动”可以对应启动该视频眼镜的控制指令;关键字“菜单”可以对应弹出菜单选项的控制指令;关键字“关闭”可以对应关闭该视频眼镜的控制指令。

[0087] 又举例来说,如果该语音识别模型输出的关键字为“菜单”,那么该语音识别模型输出的关键字对应的目标控制指令则为弹出菜单选项的控制指令;如果该语音识别模型输出的关键字为“关闭”,那么该语音识别模型输出的关键字对应的目标控制指令则为关闭该视频眼镜的控制指令。

[0088] S304、根据目标控制指令对所述视频眼镜进行控制。

[0089] 举例来说,如果该目标控制指令为弹出菜单选项的控制指令,则该视频眼镜可以控制弹出该菜单选项,并可以通过接收使用者针对菜单选择中的内容发出的语音控制数据,来选择执行菜单选项中对应的功能。

[0090] 又举例来说,如果该目标控制指令为关闭该视频眼镜的控制指令,那么该视频眼镜可以进行对应的关闭处理,例如控制电源进行关闭,或者控制该视频眼镜进入待机模式等,本发明实施例对此不作任何限制。

[0091] 可见,通过本发明实施例,该视频眼镜可通过音频信号输入接口接收语音控制数据,然后将接收到的语音控制数据输入到语音识别模型中,并根据语音识别模型的输出结果,确定出目标控制指令,最后根据该目标控制指令对该视频眼镜进行控制,通过语音识别技术,可以快速地辨识出使用者的控制指令,从而可以自然的为使用者提供需要的功能,在一定程度上提高了视频眼镜的使用方便性。

[0092] 下面请参阅图4,为本发明实施例提供的另一种语音控制方法的流程示意图。如图4所示的方法可包括:

[0093] S401、当视频眼镜处于待机模式时,接收对语音控制触发按钮的第一用户操作。

[0094] 在一个实施例中,该待机模式可以是电源开启,但该视频眼镜处于休眠状态的模式。

[0095] 在一个实施例中,该待机模式也可以是电源未开启的模式,本发明实施例对此不

作任何限制。

[0096] 该语音控制触发按钮可以是配置在该视频眼镜上,用于启动语音控制模式的按钮。该第一用户操作可以是使用者针对该语音控制触发按钮的拨动、触控、长按、短按等等操作,本发明实施例对此不作任何限制。

[0097] 举例来说,当该视频眼镜处于待机模式时,使用者长按该语音控制触发按钮,该视频眼镜即可接收到使用者的长按操作。

[0098] S402、根据所述第一用户操作,控制所述视频眼镜进入语音控制模式。

[0099] 在一个实施例中,该语音控制模式可以是指通过语音数据实现对该视频眼镜的控制的模式。

[0100] 举例来说,如果该第一用户操作为使用者长按该语音控制触发按钮的操作,则该视频眼镜可以根据该第一用户操作,将该视频眼镜从待机模式切换到该语音控制模式。

[0101] 在一个实施例中,当该视频眼镜处于待机模式时,也可以接收第二用户操作;如果该第二用户操作是对该视频眼镜的开关按钮的操作,则可以在该视频眼镜开启之后,控制该视频眼镜进入该语音控制模式。

[0102] 在一个实施例中,该开关按钮可以配置在该视频眼镜上,可以用于控制该视频眼镜开启。

[0103] 举例来说,如果该视频眼镜在待机模式时接收到针对该开关按钮的第二用户操作,则可以将该视频眼镜开启,并在开启之后,将该视频眼镜的模式切换为语音控制模式。

[0104] 在一个实施例中,当该视频眼镜处于待机模式时,也可以接收使用者的语音数据,并基于语音识别基础技术识别出该语音数据的内容,确定出目标语音指令,如果该语音指令为预设的可用于切换到语音控制模式的指令,则该视频眼镜从待机模式切换到语音控制模式。

[0105] 在一个实施例中,该语音识别基础技术可以是基于人工设计特征的语音识别技术。

[0106] S403、通过所述视频眼镜的音频信号输入接口接收语音控制数据。

[0107] S404、将接收到的语音控制数据输入到语音识别模型中。

[0108] S405、根据所述语音识别模型的输出结果,确定出目标控制指令。

[0109] 需要说明的是,上述S403至S405的具体实现过程可参考前述方法实施例的S301至S303步骤中对应的描述,在此不作赘述。

[0110] S406、检测所述视频眼镜当前所处的工作状态。

[0111] S407、判断所述目标控制指令所指示的控制状态是否与所述工作状态相匹配。

[0112] 举例来说,该目标控制指令所述指示的控制状态为开启状态,而该视频眼镜当前的工作状态也为开启状态,那么该目标控制指令所指示的控制状态与所述工作状态相匹配。

[0113] 又举例来说,该目标控制指令所指示的控制状态为开启状态,而该视频眼镜当前的工作状态为待机状态,那么该目标控制指令所指示的控制状态与所述工作状态不匹配。

[0114] S408、若是,则执行根据目标控制指令对所述视频眼镜进行控制。

[0115] 当该目标控制指令所指示的控制状态与工作状态相匹配时,该视频眼镜可以执行根据该目标控制指令对该视频眼镜进行控制。

[0116] 当该目标控制指令所指示的控制状态与工作状态不匹配时,该视频眼镜可以不执行根据该目标控制指令对该视频眼镜进行控制。

[0117] 在一个实施例中,如果该目标控制指令为控制该视频眼镜进入对应的控制状态(如开启状态),而该视频眼镜当前的工作状态即为该对应的控制状态(如当前的工作状态即为开启状态),那么该视频眼镜也可以不执行根据该目标控制指令对该视频眼镜进行控制。

[0118] 可见,通过本发明实施例,该视频眼镜处于待机模式时,接收对语音控制触发按钮的第一用户操作,根据该第一用户操作,控制该视频眼镜进入语音控制模式,然后通过该视频眼镜的音频信号输入接口接收语音控制数据,将接收到的语音控制数据输入到语音识别模型中,根据该语音识别模型的输出结果,确定出目标控制指令,在该目标控制指令所指示的控制状态与该工作状态相匹配的情况下,则根据该目标控制指令对该视频眼镜进行控制,可以在该视频眼镜进入语音控制模式时,利用语音识别技术来快速的辨识出使用者的控制指令,从而可以自然的为用户提供需要的功能,在一定程度上提高了视频眼镜的使用方便性。

[0119] 下面请参阅图5,为本发明实施例提供的一种视频眼镜的结构示意图。

[0120] 如图5所示的视频眼镜包括:存储器501以及处理器502。

[0121] 所述存储器501,用于存储程序指令;

[0122] 所述处理器502,用于执行所述存储器501存储的程序指令,当程序指令被执行时,用于:

[0123] 通过所述视频眼镜的音频信号输入接口接收语音控制数据;

[0124] 将接收到的语音控制数据输入到语音识别模型中,所述语音识别模型用于对所述语音控制数据进行分析识别;

[0125] 根据所述语音识别模型的输出结果,确定出目标控制指令;

[0126] 根据目标控制指令对所述视频眼镜进行控制。

[0127] 在一个实施例中,所述输出结果为关键字;所述处理器502用于根据所述语音识别模型的输出结果,确定出目标控制指令时,具体用于:

[0128] 根据所述语音识别模型输出的关键字以及预置的关键字与控制指令之间的对应关系,确定出所述语音识别模型输出的关键字对应的目标控制指令。

[0129] 在一个实施例中,所述视频眼镜包括语音控制触发按钮;

[0130] 所述处理器502还用于:当所述视频眼镜处于待机模式时,接收对所述语音控制触发按钮的第一用户操作;根据所述第一用户操作,控制所述视频眼镜进入语音控制模式。

[0131] 在一个实施例中,所述处理器502还用于:当所述视频眼镜处于待机模式时,接收第二用户操作;如果所述第二用户操作是对所述视频眼镜的开关按钮的操作,则在所述视频眼镜开启后,控制所述视频眼镜进入语音控制模式。

[0132] 在一个实施例中,所述处理器502还用于:检测所述视频眼镜当前所处的工作状态;判断所述目标控制指令所指示的控制状态是否与所述工作状态相匹配;若是,则执行所述根据目标控制指令对所述视频眼镜进行控制。

[0133] 在一个实施例中,所述处理器502还用于:获取第一语音训练数据,并查询出所述第一语音训练数据对应的预设语音识别结果;将所述第一语音训练数据输入到预置的语音

识别初始模型中,所述语音识别初始模型用于对所述第一语音训练数据进行分析识别;如果所述语音识别初始模型的输出结果与所述预设语音识别结果不相同,则对所述语音识别初始模型进行优化更新,以便于得到所述语音识别模型。

[0134] 在一个实施例中,所述处理器502还用于:通过所述视频眼镜的音频信号输入接口接收第二语音训练数据;输出语音识别结果预设界面,将在所述语音识别结果选择界面上选择的语音识别结果作为预设语音识别结果;将所述第二语音训练数据输入到预置的语音识别初始模型中,所述语音识别初始模型用于对所述第二语音训练数据进行分析识别;如果所述语音识别初始模型的输出结果与所述预设语音识别结果不相同,则对所述语音识别初始模型进行优化更新,以便于得到所述语音识别模型。

[0135] 需要说明的是,对于前述的各个方法实施例,为了简单描述,故将其都表述为一系列的动作组合,但是本领域技术人员应所述知悉,本发明并不受所描述的动作顺序的限制,因为依据本发明,某一些步骤可以采用其他顺序或者同时进行。其次,本领域技术人员也应所述知悉,说明书中所描述的实施例均属于优选实施例,所涉及的动作和模块并不一定是本发明所必须的。

[0136] 本领域普通技术人员可以理解上述实施例的各种方法中的全部或部分步骤是可以通程序来指令相关的硬件来完成,所述程序可以存储于一计算机可读存储介质中,存储介质可以包括:闪存盘、只读存储器(Read-Only Memory,ROM)、随机存取器(Random Access Memory,RAM)、磁盘或光盘等。

[0137] 以上对本发明实施例所提供的一种语音控制方法及视频眼镜进行了详细介绍,本文中应用了具体个例对本发明的原理及实施方式进行了阐述,以上实施例的说明只是用于帮助理解本发明的方法及其核心思想;同时,对于本领域的一般技术人员,依据本发明的思想,在具体实施方式及应用范围上均会有改变之处,综上所述,本说明书内容不应理解为对本发明的限制。

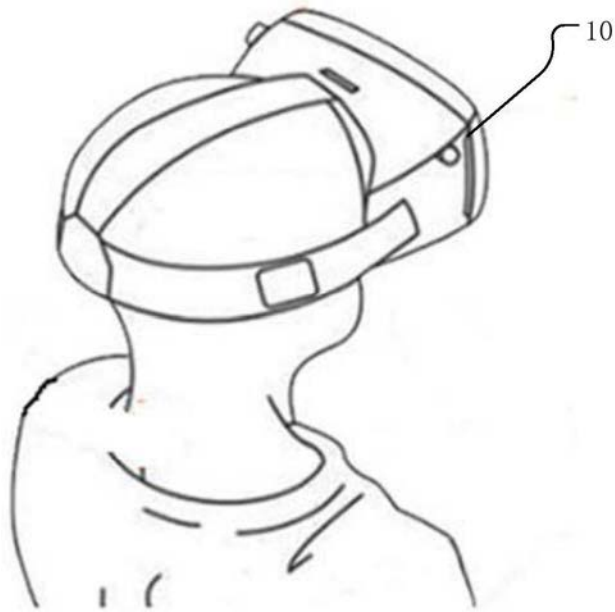


图1a

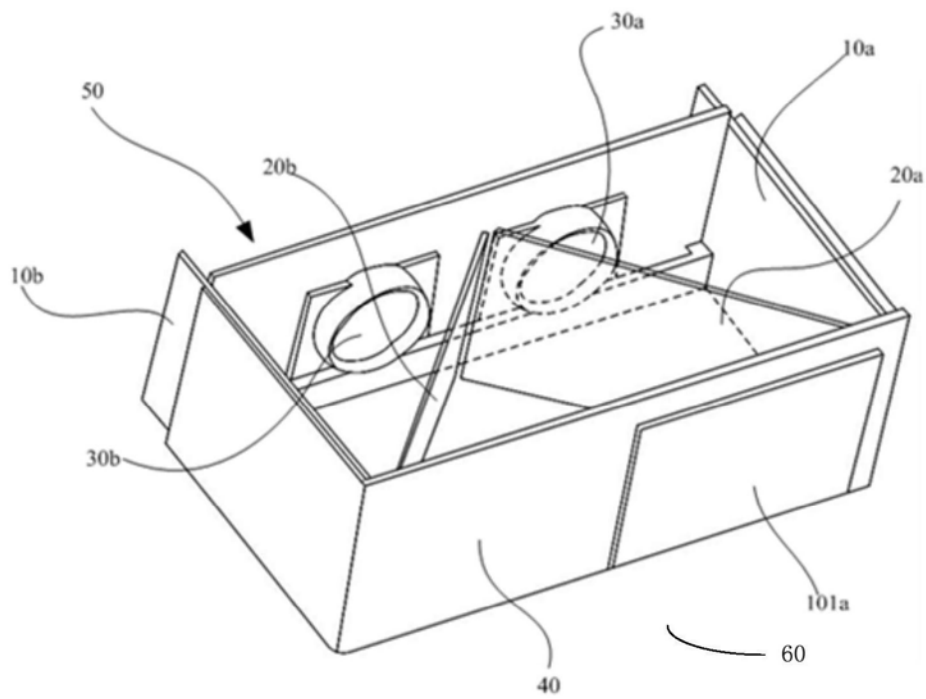


图1b

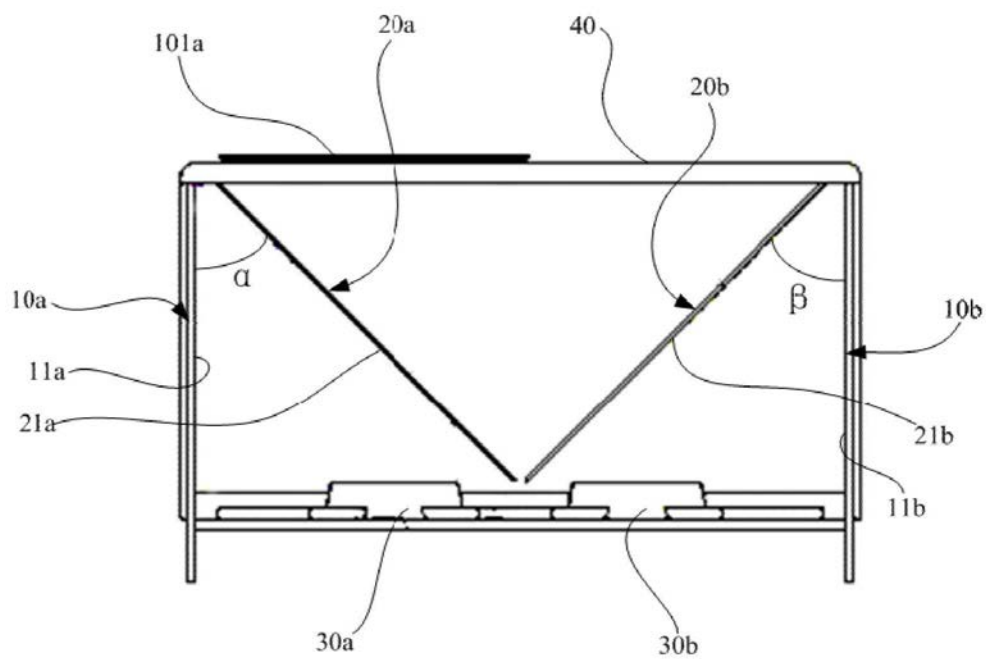


图1c

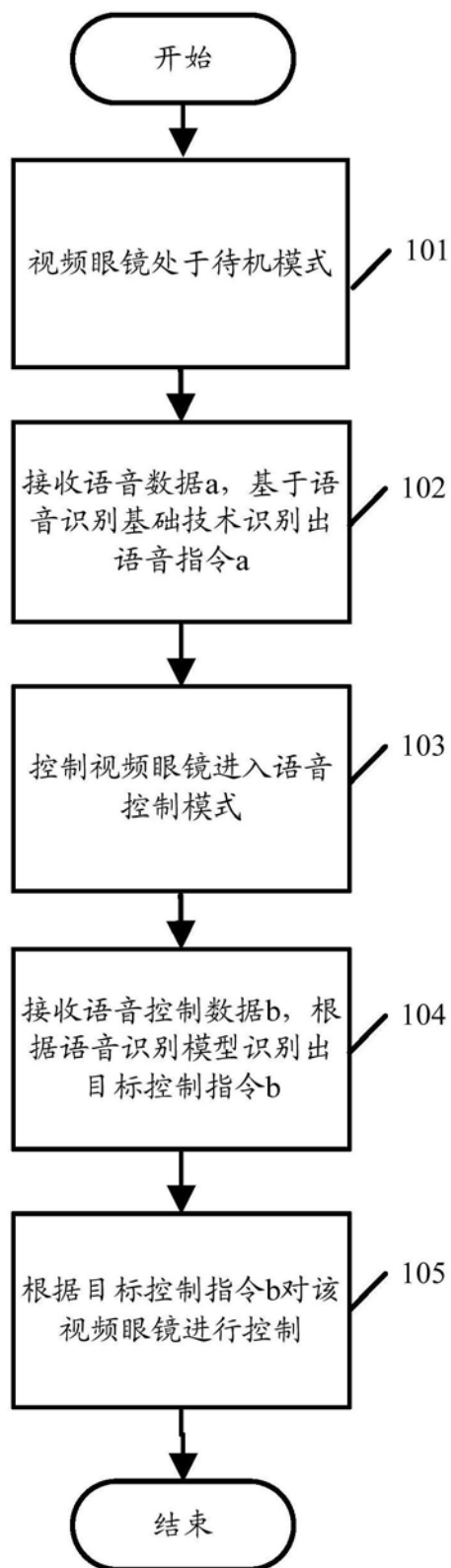


图2

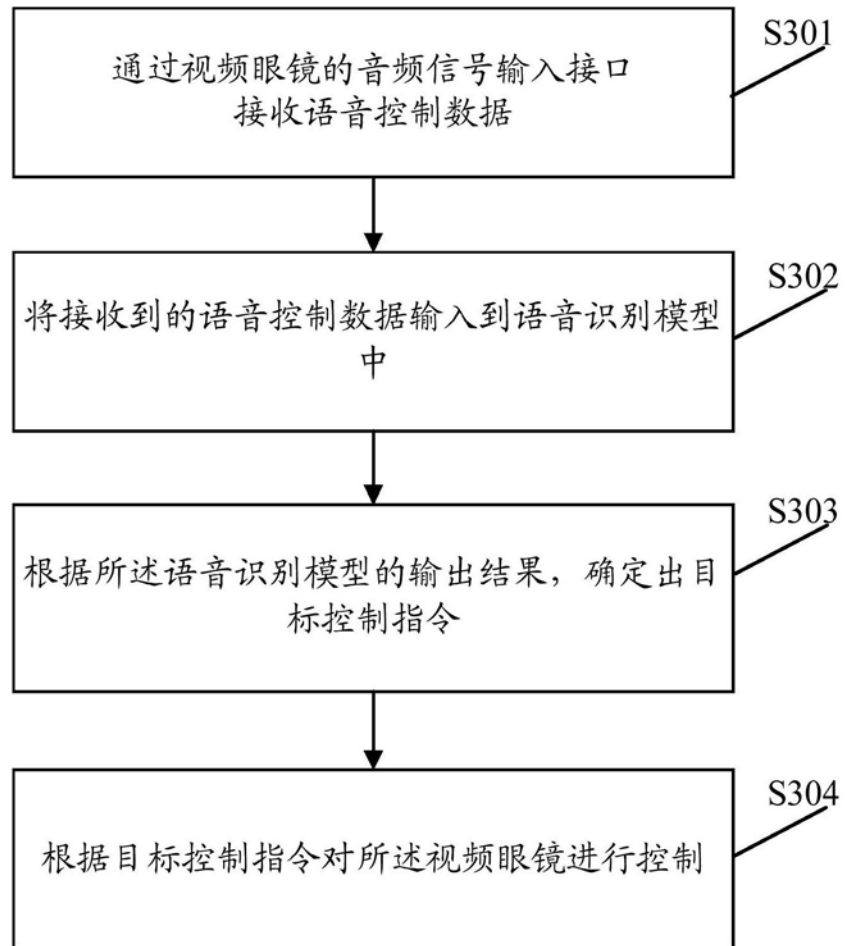


图3

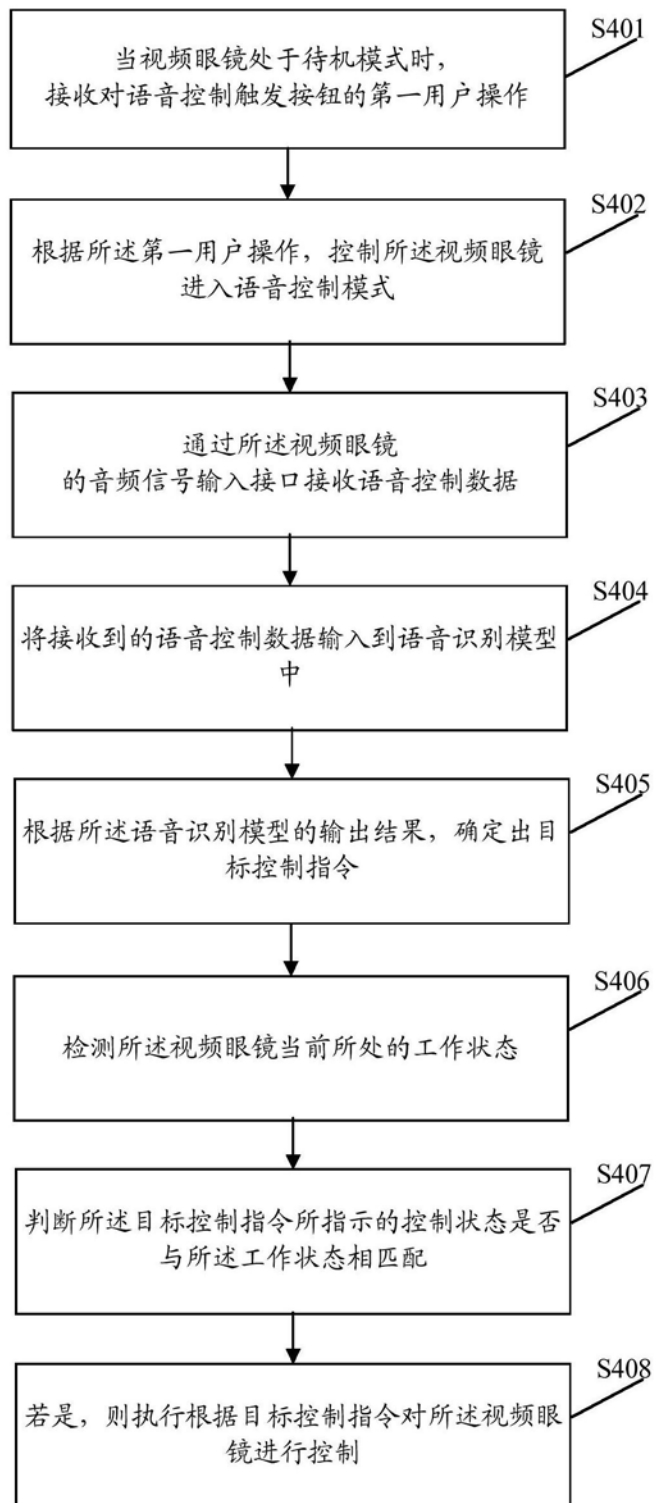


图4

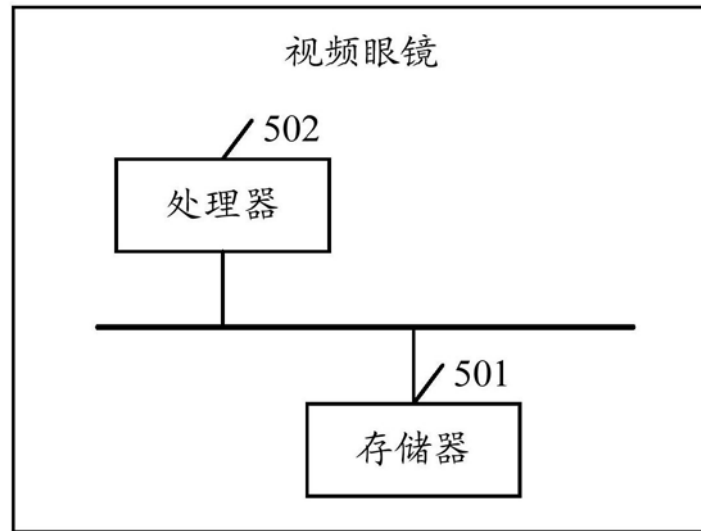


图5