



(12)发明专利申请

(10)申请公布号 CN 111063024 A  
(43)申请公布日 2020.04.24

(21)申请号 201911267521.2

(22)申请日 2019.12.11

(71)申请人 腾讯科技(深圳)有限公司  
地址 518000 广东省深圳市南山区高新区  
科技中一路腾讯大厦35层

(72)发明人 黄生辉 林祥凯 欧阳才晟

(74)专利代理机构 深圳市隆天联鼎知识产权代  
理有限公司 44232  
代理人 王鹏健

(51) Int. Cl.  
G06T 17/00(2006.01)  
G06T 19/00(2011.01)

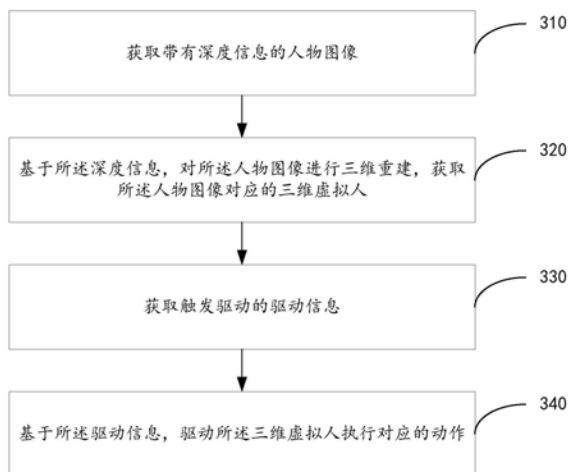
权利要求书2页 说明书17页 附图5页

(54)发明名称

三维虚拟人驱动方法、装置、电子设备及存储介质

(57)摘要

本公开提供了一种三维虚拟人驱动方法、装置、电子设备及存储介质,所述方法包括:获取带有深度信息的人物图像;基于所述深度信息,对所述人物图像进行三维重建,获取所述人物图像对应的三维虚拟人;获取触发驱动的驱动信息;基于所述驱动信息,驱动所述三维虚拟人执行对应的动作。本公开实施例能够提高三维虚拟人执行动作的逼真度。



1. 一种三维虚拟人驱动方法,其特征在于,所述方法包括:  
获取带有深度信息的人物图像;  
基于所述深度信息,对所述人物图像进行三维重建,获取所述人物图像对应的三维虚拟人;  
获取触发驱动的驱动信息;  
基于所述驱动信息,驱动所述三维虚拟人执行对应的动作。
2. 根据权利要求1所述的方法,其特征在于,事先设置各拍摄角度分别对应的采集槽,获取带有深度信息的人物图像,包括:从各所述采集槽分别获取各拍摄角度的带有深度信息的人物图像。
3. 根据权利要求1所述的方法,其特征在于,获取带有深度信息的人物图像,包括:  
获取包含有人物图像的视频帧、以及所述视频帧的深度信息;  
基于预设的关键点提取算法,提取所述视频帧中的人物关键点;  
基于所述人物关键点,获取所述视频帧中的人物图像的拍摄角度;  
若所述拍摄角度符合预设条件,则获取所述视频帧中的人物图像、以及所述人物图像的深度信息。
4. 根据权利要求1所述的方法,其特征在于,基于所述深度信息,对所述人物图像进行三维重建,获取所述人物图像对应的三维虚拟人,包括:  
基于所述深度信息,对所述人物图像进行三维重建,获取所述人物图像对应的三维模型;  
基于所述人物图像的纹理信息,对所述三维模型进行图形渲染,获取所述人物图像对应的三维虚拟人。
5. 根据权利要求1所述的方法,其特征在于,获取触发驱动的驱动信息,包括:获取预设的触发驱动的驱动信息;  
基于所述驱动信息,驱动所述三维虚拟人执行对应的动作,包括:基于所述驱动信息,驱动所述三维虚拟人执行对应的预设的动作。
6. 根据权利要求5所述的方法,其特征在于,基于所述驱动信息,驱动所述三维虚拟人执行对应的预设的动作,包括:  
获取所述驱动信息对应的预设的口型数据、以及对应的预设的语音数据;  
基于所述语音数据以及所述口型数据,驱动所述三维虚拟人以所述口型数据对应的口型发出所述语音数据对应的语音。
7. 根据权利要求5所述的方法,其特征在于,基于所述驱动信息,驱动所述三维虚拟人执行对应的预设的动作,包括:  
获取所述驱动信息对应的预设的姿势数据;  
基于所述姿势数据,驱动所述三维虚拟人执行所述姿势数据对应的姿势。
8. 根据权利要求1所述的方法,其特征在于,获取触发驱动的驱动信息,包括:获取文本形式的触发驱动的驱动信息;  
基于所述驱动信息,驱动所述三维虚拟人执行对应的动作,包括:  
基于预设的自然语言理解技术,对所述驱动信息进行处理,获取所述驱动信息对应的驱动数据;

基于所述驱动数据,驱动所述三维虚拟人执行对应的动作。

9. 根据权利要求1所述的方法,其特征在于,获取触发驱动的驱动信息,包括:获取语音形式的触发驱动的驱动信息;

基于所述驱动信息,驱动所述三维虚拟人执行对应的动作,包括:

对所述驱动信息进行语音识别,获取对应的文本形式的所述驱动信息;

基于预设的自然语言理解技术,对所述驱动信息进行处理,获取所述驱动信息对应的驱动数据;

基于所述驱动数据,驱动所述三维虚拟人执行对应的动作。

10. 根据权利要求8或9所述的方法,其特征在于,基于预设的自然语言理解技术,对所述驱动信息进行处理,获取所述驱动信息对应的驱动数据,包括:

基于预设的自然语言理解技术,对所述驱动信息进行处理,获取所述驱动信息对应的驱动意图;

基于预设的语音合成技术,对所述驱动意图进行处理,获取所述驱动意图对应的语音数据以及口型数据;

基于所述驱动数据,驱动所述三维虚拟人执行对应的动作,包括:基于所述语音数据以及口型数据,驱动所述三维虚拟人以所述口型数据对应的口型发出所述语音数据对应的语音。

11. 根据权利要求8或9所述的方法,其特征在于,基于预设的自然语言理解技术,对所述驱动信息进行处理,获取所述驱动信息对应的驱动数据,包括:

基于预设的自然语言理解技术,对所述驱动信息进行处理,获取所述驱动信息对应的驱动意图;

基于所述驱动意图与预设的姿势数据集的匹配结果,获取所述驱动意图对应的姿势数据;

基于所述驱动数据,驱动所述三维虚拟人执行对应的动作,包括:基于所述姿势数据,驱动所述三维虚拟人执行所述姿势数据对应的姿势。

12. 一种三维虚拟人驱动装置,其特征在于,所述装置包括:

第一获取模块,配置为获取带有深度信息的人物图像;

第二获取模块,配置为基于所述深度信息,对所述人物图像进行三维重建,获取所述人物图像对应的三维虚拟人;

第三获取模块,配置为获取触发驱动的驱动信息;

驱动模块,配置为基于所述驱动信息,驱动所述三维虚拟人执行对应的动作。

13. 一种三维虚拟人驱动电子设备,其特征在于,包括:

存储器,存储有计算机可读指令;

处理器,读取存储器存储的计算机可读指令,以执行权利要求1-11中的任一个所述的方法。

14. 一种计算机可读存储介质,其特征在于,其上存储有计算机可读指令,当所述计算机可读指令被计算机的处理器执行时,使计算机执行权利要求1-11中的任一个所述的方法。

## 三维虚拟人驱动方法、装置、电子设备及存储介质

### 技术领域

[0001] 本公开涉及人工智能领域,具体涉及一种三维虚拟人驱动方法、装置、电子设备及存储介质。

### 背景技术

[0002] 随着互联网技术的发展,三维虚拟人在越来越多的用户的生活中发挥的影响越来越大。具体的,用户可以与三维虚拟人进行交互,以满足用户自身信息的需求、情感的需求、娱乐的需求。现有技术中,三维虚拟人驱动的过程较为僵硬,逼真度低,从而会导致用户与三维虚拟人进行交互时难以融入情绪,用户体验较差。

### 发明内容

[0003] 本公开的一个目的在于提出一种三维虚拟人驱动方法、装置、电子设备及存储介质,能够提高三维虚拟人执行动作的逼真度。

[0004] 根据本公开实施例的一方面,公开了一种三维虚拟人驱动方法,所述方法包括:

[0005] 获取带有深度信息的人物图像;

[0006] 基于所述深度信息,对所述人物图像进行三维重建,获取所述人物图像对应的三维虚拟人;

[0007] 获取触发驱动的驱动信息;

[0008] 基于所述驱动信息,驱动所述三维虚拟人执行对应的动作。

[0009] 根据本公开实施例的一方面,公开了一种三维虚拟人驱动装置,所述装置包括:

[0010] 第一获取模块,配置为获取带有深度信息的人物图像;

[0011] 第二获取模块,配置为基于所述深度信息,对所述人物图像进行三维重建,获取所述人物图像对应的三维虚拟人;

[0012] 第三获取模块,配置为获取触发驱动的驱动信息;

[0013] 驱动模块,配置为基于所述驱动信息,驱动所述三维虚拟人执行对应的动作。

[0014] 根据本公开实施例的一方面,公开了一种三维虚拟人驱动电子设备,包括:存储器,存储有计算机可读指令;处理器,读取存储器存储的计算机可读指令,以执行以上权利要求中的任一个所述的方法。

[0015] 根据本公开实施例的一方面,公开了一种计算机可读存储介质,其上存储有计算机可读指令,当所述计算机可读指令被计算机的处理器执行时,使计算机执行以上权利要求中的任一个所述的方法。

[0016] 本公开实施例中,基于人物图像的深度信息对人物图像进行三维重建,从而使得生成的三维虚拟人能够全面地还原出该人物图像对应的用户的外貌。进而获取触发驱动的驱动信息,驱动该三维虚拟人执行对应的动作,提高了三维虚拟人动作执行的逼真度。

[0017] 本公开的其他特性和优点将通过下面的详细描述变得显然,或部分地通过本公开的实践而习得。

[0018] 应当理解的是,以上的一般描述和后文的细节描述仅是示例性的,并不能限制本公开。

### 附图说明

[0019] 通过参考附图详细描述其示例实施例,本公开的上述和其它目标、特征及优点将变得更加显而易见。

[0020] 图1A-1D示出了本公开实施例的可以应用的体系构架图。

[0021] 图2示出了根据本公开一个实施例的三维虚拟人的生成过程中客户端的相关界面。

[0022] 图3示出了根据本公开一个实施例的三维虚拟人驱动方法的流程图。

[0023] 图4示出了根据本公开一个实施例的三维虚拟人驱动的实施逻辑。

[0024] 图5示出了根据本公开一个实施例的三维虚拟人驱动的实施逻辑。

[0025] 图6示出了根据本公开一个实施例的三维虚拟人驱动装置的框图。

[0026] 图7示出了根据本公开一个实施例的三维虚拟人驱动电子设备的硬件图。

### 具体实施方式

[0027] 现在将参考附图更全面地描述示例实施方式。然而,示例实施方式能够以多种形式实施,且不应被理解为限于在此阐述的范例;相反,提供这些示例实施方式使得本公开的描述将更加全面和完整,并将示例实施方式的构思全面地传达给本领域的技术人员。附图仅为本公开的示意性图解,并非一定是按比例绘制。图中相同的附图标记表示相同或类似的部分,因而将省略对它们的重复描述。

[0028] 此外,所描述的特征、结构或特性可以以任何合适的方式结合在一个或更多示例实施方式中。在下面的描述中,提供许多具体细节从而给出对本公开的示例实施方式的充分理解。然而,本领域技术人员将意识到,可以实践本公开的技术方案而省略所述特定细节中的一个或更多,或者可以采用其它的方法、组元、步骤等。在其它情况下,不详细示出或描述公知结构、方法、实现或者操作以避免喧宾夺主而使得本公开的各方面变得模糊。

[0029] 附图中所示的一些方框图是功能实体,不一定必须与物理或逻辑上独立的实体相对应。可以采用软件形式来实现这些功能实体,或在一个或多个硬件模块或集成电路中实现这些功能实体,或在不同网络和/或处理器装置和/或微控制器装置中实现这些功能实体。

[0030] 人工智能(Artificial Intelligence, AI)是利用数字计算机或者数字计算机控制的机器模拟、延伸和扩展人的智能,感知环境、获取知识并使用知识获得最佳结果的理论、方法、技术及应用系统。换句话说,人工智能是计算机科学的一个综合技术,它企图了解智能的实质,并生产出一种新的能以人类智能相似的方式做出反应的智能机器。人工智能也就是研究各种智能机器的设计原理与实现方法,使机器具有感知、推理与决策的功能。

[0031] 人工智能技术是一门综合学科,涉及领域广泛,既有硬件层面的技术也有软件层面的技术。人工智能基础技术一般包括如传感器、专用人工智能芯片、云计算、分布式存储、大数据处理技术、操作/交互系统、机电一体化等技术。人工智能软件技术主要包括计算机视觉技术、语音处理技术、自然语言处理技术以及机器学习/深度学习等几大方向。

[0032] 计算机视觉技术 (Computer Vision, CV) 计算机视觉是一门研究如何使机器“看”的科学,更进一步的说,就是指用摄影机和电脑代替人眼对目标进行识别、跟踪和测量等机器视觉,并进一步做图形处理,使电脑处理成为更适合人眼观察或传送给仪器检测的图像。作为一个科学学科,计算机视觉研究相关的理论和技术,试图建立能够从图像或者多维数据中获取信息的人工智能系统。计算机视觉技术通常包括图像处理、图像识别、图像语义理解、图像检索、OCR、视频处理、视频语义理解、视频内容/行为识别、三维物体重建、3D技术、虚拟现实、增强现实、同步定位与地图构建等技术,还包括常见的人脸识别、指纹识别等生物特征识别技术。

[0033] 语音技术 (Speech Technology) 的关键技术有自动语音识别技术 (ASR) 和语音合成技术 (TTS) 以及声纹识别技术。让计算机能听、能看、能说、能感觉,是未来人机交互的发展方向,其中语音成为未来最被看好的人机交互方式之一。

[0034] 自然语言处理 (Nature Language processing, NLP) 是计算机科学领域与人工智能领域中的一个重要方向。它研究能实现人与计算机之间用自然语言进行有效通信的各种理论和方法。自然语言处理是一门融语言学、计算机科学、数学于一体的科学。因此,这一领域的研究将涉及自然语言,即人们日常使用的语言,所以它与语言学的研究有着密切的联系。自然语言处理技术通常包括文本处理、语义理解、机器翻译、机器人问答、知识图谱等技术。

[0035] 本公开实施例提供的方案涉及人工智能的计算机视觉技术、语音技术、自然语言处理技术,具体通过如下实施例进行说明。

[0036] 下面首先参考图1A-1D描述本公开实施例可以应用的体系构架。

[0037] 图1A示出了本公开一实施例应用的体系构架。该实施例中,三维虚拟人的生成、驱动由客户端10进行控制。具体的,客户端10获取到用户的带有深度信息的人物图像,基于该深度信息,对人物图像进行三维重建,于客户端10内部,得到该人物图像对应的三维虚拟人,即,得到该用户对应的三维虚拟人。三维虚拟人生成后,用户可与该三维虚拟人进行交互。具体的,客户端10若获取到触发驱动的驱动信息,则驱动该三维虚拟人执行对应的动作,以与用户进行交互。

[0038] 图1B示出了本公开一实施例应用的体系构架。该实施例中,三维虚拟人的生成、驱动主要由云端服务器20进行控制。具体的,客户端10获取到用户的带有深度信息的人物图像后,将该带有深度信息的人物图像上传至云端服务器20(图1B中虚线所示过程)。云端服务器20获取到该带有深度信息的人物图像后,基于该深度信息,对人物图像进行三维重建,于云端服务器20内部,得到该人物图像对应的虚拟人,即,得到该用户对应的三维虚拟人。三维虚拟人生成后,云端服务器20将三维虚拟人的相关模型数据发送给客户端10,使得用户可以与客户端10内的三维虚拟人进行交互。具体的,云端服务器20若获取到触发驱动的驱动信息,则向客户端10发送对应的驱动指令,使得客户端10根据该驱动指令驱动客户端10内部的三维虚拟人执行对应的动作,以与用户进行交互。

[0039] 图1C示出了本公开一实施例应用的体系构架。该实施例中,三维虚拟人的生成、驱动主要由云端服务器20进行控制。具体的,主播端10a获取到主播的带有深度信息的人物图像后,将该带有深度信息的人物图像上传至云端服务器20(图1C中虚线所示过程)。云端服务器20获取到该带有深度信息的人物图像后,基于该深度信息,对人物图像进行三维重建,

于云端服务器20内部,得到该人物图像对应的虚拟人,即,得到主播对应的三维虚拟人。

[0040] 三维虚拟人生成后,云端服务器20将三维虚拟人的相关模型数据发送给主播端10a,使得主播可以与主播端10a内的三维虚拟人进行交互,具体的,使得主播端10a内的三维虚拟人可以模仿主播的行为;云端服务器20若获取到主播的行为数据,则向主播端10a发送对应的驱动指令,使得主播端10a根据该驱动指令驱动主播端10a内部的三维虚拟人执行对应的动作,以模仿主播的行为。

[0041] 三维虚拟人生成后,云端服务器20将三维虚拟人的相关模型数据发送给观众端10b,使得观众端10b能够播放出模仿主播行为的三维虚拟人的视频。具体的,云端服务器20将指示主播端10a内的三维虚拟人模仿主播的行为的驱动指令发送给观众端10b,使得观众端10b能够根据该驱动指令驱动观众端10b中的三维虚拟人,向观众播放出模仿主播行为的三维虚拟人的视频。

[0042] 图1D示出了本公开一实施例应用的体系构架。该实施例中,三维虚拟人的生成、驱动主要由云端服务器20进行控制。具体的,第一用户端10a获取到第一用户的带有深度信息的人物图像后,将该带有深度信息的人物图像上传至云端服务器20(图1D中虚线所示过程)。云端服务器20获取到该带有深度信息的人物图像后,基于该深度信息,对人物图像进行三维重建,于云端服务器20内部,得到该人物图像对应的虚拟人,即,得到第一用户对应的三维虚拟人。

[0043] 三维虚拟人生成后,云端服务器20将三维虚拟人的相关模型数据发送给第二用户端10b,使得第二用户能够与第二用户端10b中的第一用户的三维虚拟人进行交互。具体的,云端服务器20若获取到触发驱动的驱动信息,则向第二用户端10b发送对应的驱动指令,使得第二用户端10b根据该驱动指令驱动第二用户端10b内部的第一用户的三维虚拟人执行对应的动作,以与第二用户进行交互。

[0044] 在更具体的应用场景中,该实施例可以应用到聊天机器人的场景中。例如:第一用户为小张,第二用户为小王,小张与小王互为朋友,但小张由于工作繁忙无法经常与小王进行联系。小张通过其手机终端将其带有深度信息的人物图像上传至云端服务器的聊天机器人系统中。聊天机器人系统根据小张的带有深度信息的人物图像,生成小张的三维虚拟人,并部署到小王的手机终端。从而小王可以在其手机终端的聊天界面中,随时与小张的三维虚拟人进行聊天。

[0045] 需要说明的是,有关图1A-1D所展示的本公开可应用的体系构架的说明仅仅是示例性的,不应对本公开的功能和使用范围造成限制。

[0046] 需要说明的是,本公开所提供的方法可以部署至个人电脑终端、也可以部署至手机终端、也可以部署至其他具有足够运算能力的计算终端上。

[0047] 下面参考图2对本公开一实施例中,三维虚拟人的生成过程中客户端的相关界面进行描述,以示例性地展示本公开在生成三维虚拟人这一过程中,用户与客户端的大致交互过程。

[0048] 该实施例中,客户端可以为手机终端。用户开启手机终端的摄像头后,根据手机终端的图像采集界面中显示的头部虚线框进行头部的校准。用户校准头部后,点击图像采集界面中的“开始扫描”按钮,根据手机终端的提示进行头部图像的采集。

[0049] 具体的,该实例中,手机终端会于图像采集界面进行头部角度指定的提示。例如:

随着时间的推移,手机终端于图像采集界面依次显示出“请正对摄像头”的文字信息、“请向左转动头部90度”的文字信息、“请正对摄像头,再向右转动头部90度”的文字信息、“请正对摄像头,再向上仰起头部”的文字信息。

[0050] 从而手机终端能够从对用户头部扫描得到的视频帧中选择出预定角度带有深度信息的头部图像:正面的头部图像、朝左的头部图像、朝右的头部图像、朝上的头部图像。进而手机终端根据得到的带有深度信息的头部图像,对用户的头部进行三维重建,从而在手机终端的交互界面中显示出头部重建结果——该用户的三维虚拟头部。

[0051] 需要说明的是,图2所示的实施例只是示例性的说明,不应对本公开的功能和使用范围造成限制。

[0052] 下面对本公开的具体实施过程进行详细描述。

[0053] 参考图3所示,一种三维虚拟人驱动方法,包括:

[0054] 步骤310、获取带有深度信息的人物图像;

[0055] 步骤320、基于所述深度信息,对所述人物图像进行三维重建,获取所述人物图像对应的三维虚拟人;

[0056] 步骤330、获取触发驱动的驱动信息;

[0057] 步骤340、基于所述驱动信息,驱动所述三维虚拟人执行对应的动作。

[0058] 本公开实施例中,基于人物图像的深度信息对人物图像进行三维重建,从而使得生成的三维虚拟人能够全面地还原出该人物图像对应的用户的外貌。进而获取触发驱动的驱动信息,驱动该三维虚拟人执行对应的动作,提高了三维虚拟人动作执行的逼真度。

[0059] 可以理解的,获取到的带有深度信息的人物图像可以为用户的全身图像,对应的,三维重建时,可以获取到与用户全身上下基本一致的三维虚拟人。出于简要说明的目的,在本公开的示例性说明中:获取到的带有深度信息的人物图像为用户的头部图像,对应的,三维重建时,将获取到的与用户头部基本一致的三维虚拟头部视为该用户对应的三维虚拟人。该种说明方式仅仅是示例性的,不应对本公开的功能和使用范围造成限制。

[0060] 在步骤310中,获取带有深度信息的人物图像。

[0061] 本公开实施例中,为获取带有深度信息的人物图像,要求负责采集人物图像的图像采集终端通过能够采集深度信息的摄像头进行人物图像的采集。

[0062] 在一实施例中,事先设置各拍摄角度分别对应的采集槽,获取带有深度信息的人物图像,包括:从各所述采集槽分别获取各拍摄角度的带有深度信息的人物图像。

[0063] 该实施例中,事先于客户端设置各采集槽,用于采集相应拍摄角度的人物图像。当用户触发图像采集后,客户端显示出各采集槽,用于指示用户具体进行哪些拍摄角度的人物图像的采集。在用户根据采集槽的指示,通过各采集槽进行相应拍摄角度的人物图像的拍摄后,即可获取到相应拍摄角度的带有深度信息的人物图像,从而能够在此基础上对人物图像进行三维重建。

[0064] 例如:事先于客户端的图像采集界面设置4个采集槽——用于拍摄正面的头部图像的槽1、用于拍摄朝左的头部图像的槽2、用于拍摄朝右的头部图像的槽3、用于拍摄朝上的头部图像的槽4。

[0065] 当用户触发客户端进入图像采集界面后,客户端在图像采集界面中显示出这4个采集槽。当用户点击槽1的按钮后,客户端弹出相机窗口,并于相机窗口显示出“请正对摄像



头”的文字信息。从而用户可以根据该文字信息的提醒,于相机窗口拍摄正面的头部图像。拍摄完成后,客户端退出相机窗口,回到图像采集界面,并在槽1中显示出待上传到云端服务器、采集到的正面的头部图像。

[0066] 同理,用户点击槽2,进行朝左的头部图像的拍摄,客户端在图像采集界面的槽2中显示出待上传到云端服务器、采集到的朝左的头部图像;用户点击槽3,进行朝右的头部图像的拍摄,客户端在图像采集界面的槽3中显示出待上传到云端服务器、采集到的朝右的头部图像;用户点击槽4,进行朝上的头部图像的拍摄,客户端在图像采集界面的槽4中显示出待上传到云端服务器、采集到的朝上的头部图像。

[0067] 当这4个采集槽均采集到相应拍摄角度的头部图像后,客户端自动将这4张头部图像上传至云端服务器,从而使得云端服务器能够根据这4张头部图像对用户的头部进行三维重建,得到对应的三维虚拟人。

[0068] 或者,当这4个采集槽均采集到相应拍摄角度的头部图像后,用户可以点击图像采集界面中的“上传”按钮。从而客户端将这4张头部图像上传至云端服务器,从而使得云端服务器能够根据这4张头部图像对用户的头部进行三维重建,得到对应的三维虚拟人。

[0069] 该实施例的优点在于,通过事先设置各拍摄角度对应的采集槽,从而能够根据采集槽与人物图像的对应关系,快速确定各人物图像对应的拍摄角度,节省了对人物图像的处理时间。

[0070] 需要说明的是,该实施例只是示例性的说明,不应对本公开的功能和使用范围造成限制。

[0071] 在一实施例中,获取带有深度信息的人物图像,包括:

[0072] 获取包含有人物图像的视频帧、以及该视频帧的深度信息;

[0073] 基于预设的关键点提取算法,提取该视频帧中的人物关键点;

[0074] 基于该人物关键点,获取该视频帧中的人物图像的拍摄角度;

[0075] 若该拍摄角度符合预设条件,则获取该视频帧中的人物图像、以及该人物图像的深度信息。

[0076] 该实施例中,当用户触发图像采集后,客户端弹出相机窗口,并指示用户根据对应的提示信息进行视频的拍摄。在用户根据提示信息完成视频的拍摄后,即可获取到包含有人物图像的视频帧、以及视频帧的深度信息。进而基于预设的关键点提取算法,提取出视频帧的人物关键点;基于视频帧的人物关键点,确定视频帧的人物图像的拍摄角度;挑选出拍摄角度符合预设条件的视频帧,得到该视频帧中的人物图像、该人物图像的深度信息,从而能够在此基础上对人物图像进行三维重建。

[0077] 例如:当用户触发客户端进入图像采集界面后,客户端在图像采集界面中弹出相机窗口,并随着时间的推移,以语音信息的形式发出如下提示信息:“请正对摄像头”的语音信息、“请向左转动头部90度”的语音信息、“请正对摄像头,再向右转动头部90度”的语音信息、“请正对摄像头,再向上仰起头部”的语音信息。

[0078] 用户根据提示信息完成视频的拍摄后,客户端将视频上传至云端服务器。云端服务器按照预设的视频帧挑选规则从该视频中挑选出视频帧(例如:每隔5帧挑选出一视频帧),基于预设的关键点提取算法,提取出视频帧的人物关键点;基于视频帧的人物关键点,确定视频帧的人物图像的拍摄角度;根据拍摄角度,挑选出用户头部正对摄像头的视频帧、

用户头部朝左的视频帧、用户头部朝右的视频帧、用户头部朝上的视频帧,并获取这些视频帧中的人物图像、以及这些人物图像的深度信息,从而云端服务器能在此基础上对用户的头部进行三维重建,得到对应的三维虚拟人。

[0079] 或者,用户根据提示信息完成视频的拍摄后,客户端按照预设的视频帧挑选规则从该视频中挑选出视频帧(例如:每隔5帧挑选出一视频帧),基于预设的关键点提取算法,提取出视频帧的人物关键点;基于视频帧的人物关键点,确定视频帧的人物图像的拍摄角度;根据拍摄角度,挑选出用户头部正对摄像头的视频帧、用户头部朝左的视频帧、用户头部朝右的视频帧、用户头部朝上的视频帧,并获取这些视频帧中的人物图像、以及这些人物图像的深度信息。客户端将这些视频帧中的人物图像、以及这些人物图像的深度信息上传至云端服务器,从而云端服务器能在此基础上对用户的头部进行三维重建,得到对应的三维虚拟人。

[0080] 该实施例的优点在于,通过对大量连续视频帧进行基于拍摄角度的选取,使得选取出的人物图像更符合对拍摄角度的要求,尽量减小了用户对拍摄角度的主观判断所造成的误差。

[0081] 需要说明的是,该实施例只是示例性的说明,不应对本公开的功能和使用范围造成限制。

[0082] 下面详细描述对人物图像进行三维重建,获取人物图像对应的三维虚拟人的过程。

[0083] 在步骤320中,基于该深度信息,对该人物图像进行三维重建,获取该人物图像对应的三维虚拟人。

[0084] 在一实施例中,基于该深度信息,对该人物图像进行三维重建,获取该人物图像对应的三维虚拟人,包括:

[0085] 基于所述深度信息,对所述人物图像进行三维重建,获取所述人物图像对应的三维模型;

[0086] 基于所述人物图像的纹理信息,对所述三维模型进行图形渲染,获取所述人物图像对应的三维虚拟人。

[0087] 可以理解的,深度信息反映的是像素点的三维空间深度。该实施例中,基于深度信息还原出人物图像中像素点在三维空间的位置,从而获取人物图像对应的三维模型。得到人物图像对应的三维模型后,基于人物图像的纹理信息,将人物图像的表面纹理对应地渲染到三维模型表面,从而得到人物图像对应的三维虚拟人。

[0088] 例如:获取到小明的带有深度信息的头部图像后,基于深度信息对头部图像进行三维重建,得到小明的头部的三维模型;基于头部图像的纹理信息,将小明的头部的表面纹理对应地渲染到小明的头部的三维模型表面,从而得到与小明的面部外表一致的三维虚拟人。

[0089] 需要说明的是,该实施例只是示例性的说明,不应对本公开的功能和使用范围造成限制。

[0090] 在一实施例中,基于该深度信息,对该人物图像进行三维重建,获取该人物图像对应的三维模型,包括:使用预设的RGBD算法对该深度信息进行处理,得到该人物图像对应的三维模型。

[0091] RGBD算法是指在RGB (R红;G绿;B蓝) 数据与深度图 (D:Depth map) 数据的基础上,对二维的图像进行三维重建的算法。

[0092] 该实施例中,用以对深度信息进行处理的是预设的RGBD算法。具体的,获取到人物图像、以及该人物图像的深度图后,使用RGBD算法对该二维的人物图像进行三维重建,得到该人物图像对应的三维模型。

[0093] 需要说明的是,该实施例只是示例性的说明,不应对本公开的功能和使用范围造成限制。

[0094] 下面对驱动三维虚拟人的过程进行详细描述。

[0095] 在步骤330中,获取触发驱动的驱动信息。

[0096] 在步骤340中,基于该驱动信息,驱动该三维虚拟人执行对应的动作。

[0097] 在一实施例中,获取触发驱动的驱动信息,包括:获取预设的触发驱动的驱动信息。

[0098] 基于该驱动信息,驱动该三维虚拟人执行对应的动作,包括:基于该驱动信息,驱动该三维虚拟人执行对应的预设的动作。

[0099] 该实施例中,触发驱动的驱动信息是事先于驱动三维虚拟人的系统内部设置好的。若三维虚拟人的驱动主要由客户端控制,则该系统位于客户端;若三维虚拟人的驱动主要由云端服务器控制,则该系统位于云端服务器。对应的,与驱动信息相对应的三维虚拟人应执行的动作也是事先设置好的。

[0100] 随着系统内部业务逻辑的进行,触发对应的事件时,系统即会获取到内部预设的触发驱动的驱动信息,从而驱动三维虚拟人执行预设的动作。

[0101] 例如:三维虚拟人的驱动主要由云端服务器中的RPG (Role-playing game) 游戏系统控制。通过本公开提供的方法,小明在手机终端游玩该RPG游戏时,于手机终端生成了小明对应的三维虚拟人,并通过该游戏自带的系统功能,将该三维虚拟人替换掉了该游戏中的角色甲。事先于云端服务器中的该游戏系统内部,设置好触发驱动角色甲执行相应动作的驱动信息——触发驱动角色甲进行奔跑的驱动信息为“对选项A的点击”;触发驱动角色甲进行喝水的驱动信息为“对选项B的点击”。从而,当检测到对选项A的点击时,控制角色甲进行奔跑;当检测到对选项B的点击时,控制角色甲进行喝水。

[0102] 小明对应的三维虚拟人替换掉角色甲后,对该三维虚拟人的控制逻辑与对角色甲的控制逻辑保持一致。因此,若云端服务器中的该游戏系统检测到对选项A的点击,则控制手机终端中的该三维虚拟人于手机终端的游戏界面进行奔跑;若云端服务器中的该游戏系统检测到对选项B的点击,则控制手机终端中的该三维虚拟人于手机终端的游戏界面进行喝水。

[0103] 可以理解,预设的触发驱动的驱动信息并不一定需要特定的触发动作才可以触发。例如:角色甲在该游戏中所执行的动作事先于云端服务器中的该游戏系统内部设置好——每过5分钟,控制角色甲于手机终端的游戏界面进行大笑。小明对应的三维虚拟人替换掉角色甲后,每过5分钟,云端服务器中的该游戏系统即会根据时间信息获取到驱动信息,进而控制手机终端中的该三维虚拟人于手机终端的游戏界面进行大笑。

[0104] 该实施例的优点在于,通过与较为封闭的系统进行结合,使得系统内的原本行为较为固定的虚拟客体(例如:游戏角色)能够以用户的外表执行对应的行为(例如:以用户的

外表进行奔跑、大笑),提高了三维虚拟人动作执行的逼真度,进而提高了用户与系统的交互亲密度。

[0105] 需要说明的是,该实施例只是示例性的说明,不应对本公开的功能和使用范围造成限制。

[0106] 在一实施例中,基于该驱动信息,驱动该三维虚拟人执行对应的预设的动作,包括:

[0107] 获取该驱动信息对应的预设的口型数据、以及对应的预设的语音数据;

[0108] 基于该语音数据以及该口型数据,驱动该三维虚拟人以该口型数据对应的口型发出该语音数据对应的语音。

[0109] 该实施例中,获取到的驱动信息所对应的用于驱动三维虚拟人动作执行的数据为预设的口型数据以及预设的语音数据,对应的,在该口型数据以及语音数据的基础上,驱动三维虚拟人以对应的口型,发出对应的语音。

[0110] 例如:预先对驱动信息“对选项B的点击”设置对应的口型数据——驱动口型规律张合的数据、以及对应的语音数据——驱动发出喝水声的数据。当检测到对选项B的点击后,根据该口型数据以及语音数据,驱动三维虚拟人的口型规律张合,并发出喝水声。具体的,可以通过驱动三维虚拟人的口型规律张合的同时,驱动终端的喇叭发出喝水声,从而达到三维虚拟人在口型规律张合的同时发出喝水声的效果。

[0111] 需要说明的是,该实施例只是示例性的说明,不应对本公开的功能和使用范围造成限制。

[0112] 在一实施例中,基于该驱动信息,驱动该三维虚拟人执行对应的预设的动作,包括:

[0113] 获取该驱动信息对应的预设的姿势数据;

[0114] 基于该姿势数据,驱动该三维虚拟人执行该姿势数据对应的姿势。

[0115] 该实施例中,获取到的驱动信息所对应的用于驱动三维虚拟人动作执行的数据为预设的姿势数据,对应的,在该姿势数据的基础上,驱动三维虚拟人执行对应的姿势。

[0116] 例如:预先对驱动信息“对选项A的点击”设置对应的姿势数据——驱动跑步姿势的数据。当检测到对选项A的点击后,根据该姿势数据,驱动三维虚拟人做出跑步的姿势。

[0117] 需要说明的是,该实施例只是示例性的说明,不应对本公开的功能和使用范围造成限制。

[0118] 在一实施例中,触发驱动的驱动信息是来自驱动三维虚拟人的系统外部的,主要的,是来自外部的用户。由于外部的触发驱动的信息具体内容难以穷举,因此,难以在内部预设对应的动作。故该实施例中,当获取到来自外部的触发驱动的驱动信息,需要对该驱动信息进行处理、理解,进而驱动三维虚拟人执行对应的动作。

[0119] 在一实施例中,基于该驱动信息,驱动该三维虚拟人执行对应的动作,包括:基于该驱动信息,驱动该三维虚拟人执行对应的用于模仿的动作。

[0120] 该实施例中,获取到来自外部的触发驱动的驱动信息后,对该驱动信息进行处理,驱动该三维虚拟人执行对应的用于模仿的动作,使得该三维虚拟人能够对该驱动信息的来源用户进行模仿。

[0121] 例如:三维虚拟人的驱动主要由云端服务器中的虚拟主播系统控制。通过本公开

提供的方法,小红在主播端生成了与其面部外表一致的三维虚拟人,并通过该虚拟主播系统向各观众的观众端推送视频流、语音流,使得各观众能够观看由该三维虚拟人为动作载体所呈现的内容。具体的,主播端将拍摄到的小红的动作、采集到的语音上传至云端服务器中的该虚拟主播系统。虚拟主播系统接收到来自小红的这些动作信息、语音信息后,驱动各观众端中的三维虚拟人于播放界面做出这些动作信息所对应的动作、发出这些语音信息所对应的语音。

[0122] 该实施例的优点在于,驱动与用户外表基本一致的三维虚拟人模仿用户的行为,提高了三维虚拟人动作执行的逼真度,从而使得用户更容易投入情绪,提高了用户与三维虚拟人交互的体验。

[0123] 需要说明的是,该实施例只是示例性的说明,不应对本公开的功能和使用范围造成限制。

[0124] 在一实施例中,基于该驱动信息,驱动该三维虚拟人执行对应的动作,包括:基于该驱动信息,驱动该三维虚拟人执行对应的用于应答的动作。该实施例中,获取到来自外部的触发驱动的驱动信息后,对该驱动信息进行处理、理解,驱动该三维虚拟人执行对应的用于应答的动作,使得该三维虚拟人能够对该驱动信息的来源用户进行应答。

[0125] 例如:三维虚拟人的驱动主要由云端服务器中的聊天机器人系统控制。通过本公开提供的方法,小刚在其手机终端的聊天界面生成了与其面部外表一致的三维虚拟人,该三维虚拟人可以作为聊天机器人与小刚进行对话、聊天。具体的,小刚对手机终端的聊天界面中的三维虚拟人进行语音提问。手机终端采集到小刚的语音后,将采集到的语音上传至云端服务器中的该聊天机器人系统。聊天机器人系统接收到来自小刚的这些语音信息后,对这些语音信息进行处理、理解,进而驱动聊天界面中的三维虚拟人以对应的口型发出对应的应答语音。

[0126] 该实施例的优点在于,驱动与用户外表基本一致的三维虚拟人对用户进行应答,提高了三维虚拟人动作执行的逼真度,从而提高了用户与三维虚拟人互动的体验。

[0127] 需要说明的是,该实施例只是示例性的说明,不应对本公开的功能和使用范围造成限制。

[0128] 下面对响应外部驱动三维虚拟人的过程进行详细描述。

[0129] 在一实施例中,获取触发驱动的驱动信息,包括:获取文本形式的触发驱动的驱动信息。

[0130] 基于该驱动信息,驱动该三维虚拟人执行对应的动作,包括:

[0131] 基于预设的自然语言理解技术,对该驱动信息进行处理,获取该驱动信息对应的驱动数据;

[0132] 基于该驱动数据,驱动该三维虚拟人执行对应的动作。

[0133] 自然语言理解技术,即NLU(Natural Language Understanding)技术,是一种使用计算机进行自然语言处理、理解的技术。

[0134] 该实施例中,三维虚拟人可以由文本形式的驱动信息进行驱动。当获取到文本形式的驱动信息后,基于预设的自然语言理解技术对该驱动信息进行处理,进而得到该驱动信息对应的驱动数据,进而基于该驱动数据,驱动该三维虚拟人执行对应的动作。

[0135] 具体的,基于预设的自然语言理解技术对该驱动信息进行处理后,可以得到对应

的用于模仿的驱动数据,进而基于该驱动数据,驱动该三维虚拟人执行对应的用于模仿的动作;也可以得到对应的用于应答的驱动数据,进而基于该驱动数据,驱动该三维虚拟人执行对应的用于应答的动作。

[0136] 例如:小刚通过手机终端的键盘输入文本信息——“今天的天气有点冷啊”,以与聊天界面中与其面部外表一致的、作为聊天机器人的三维虚拟人进行聊天。云端服务器中的聊天机器人系统接收到该文本信息后,基于预设的自然语言理解技术对该文本信息进行处理、理解,生成能够驱动三维虚拟人作出相应应答的驱动数据;基于该驱动数据,驱动聊天界面中的三维虚拟人做出对应的动作,以使得三维虚拟人对用户进行用以表达“最近寒潮来了,需要多加一些衣服”的应答。

[0137] 需要说明的是,该实施例只是示例性的说明,不应对本公开的功能和使用范围造成限制。

[0138] 在一实施例中,获取触发驱动的驱动信息,包括:获取语音形式的触发驱动的驱动信息。

[0139] 基于该驱动信息,驱动该三维虚拟人执行对应的动作,包括:

[0140] 对该驱动信息进行语音识别,获取对应的文本形式的该驱动信息;

[0141] 基于预设的自然语言理解技术,对该驱动信息进行处理,获取该驱动信息对应的驱动数据;

[0142] 基于该驱动数据,驱动该三维虚拟人执行对应的动作。

[0143] 进行语音识别可以采用预设的语音识别技术,即,ASR(Automatic Speech Recognition)技术。ASR技术能够对语音进行识别,将语音转化为对应的文本。

[0144] 该实施例中,三维虚拟人可以由语音形式的驱动信息进行驱动。当获取到语音形式的驱动信息后,对该驱动信息进行语音识别,获取对应的文本形式的该驱动信息。基于预设的自然语言理解技术对该驱动信息进行处理,进而得到该驱动信息对应的驱动数据,进而基于该驱动数据,驱动该三维虚拟人执行对应的动作。

[0145] 具体的,基于预设的自然语言理解技术对该驱动信息进行处理后,可以得到对应的用于模仿的驱动数据,进而基于该驱动数据,驱动该三维虚拟人执行对应的用于模仿的动作;也可以得到对应的用于应答的驱动数据,进而基于该驱动数据,驱动该三维虚拟人执行对应的用于应答的动作。

[0146] 例如:小刚通过手机终端的麦克风输入语音信息——“今天的天气有点冷啊”,以与聊天界面中与其面部外表一致的、作为聊天机器人的三维虚拟人进行聊天。云端服务器中的聊天机器人系统接收到语音信息,对该语音信息进行语音识别,得到该语音信息对应的文本信息;基于预设的自然语言理解技术对该文本信息进行处理、理解,生成能够驱动三维虚拟人作出相应应答的驱动数据;基于该驱动数据,驱动聊天界面中的三维虚拟人做出对应的动作,以使得三维虚拟人对用户进行用以表达“最近寒潮来了,需要多加一些衣服”的应答。

[0147] 需要说明的是,该实施例只是示例性的说明,不应对本公开的功能和使用范围造成限制。

[0148] 在一实施例中,基于预设的自然语言理解技术,对该驱动信息进行处理,获取该驱动信息对应的驱动数据,包括:

[0149] 基于预设的自然语言理解技术,对该驱动信息进行处理,获取该驱动信息对应的驱动意图;

[0150] 基于预设的语音合成技术,对该驱动意图进行处理,获取该驱动意图对应的语音数据以及口型数据。

[0151] 基于该驱动数据,驱动该三维虚拟人执行对应的动作,包括:基于该语音数据以及口型数据,驱动该三维虚拟人以该口型数据对应的口型发出该语音数据对应的语音。

[0152] 驱动意图是指与具体场景结合后,得到的驱动信息指示的内容所对应的意图。该实施例中,获取到的驱动意图是文本形式的信息。

[0153] 例如:在虚拟主播的场景中,获取到的驱动信息为主播说出的“今天的天气有点冷啊”这句语音,则结合该场景、经过自然语言理解技术处理后,得到的驱动意图为——“今天的天气有点冷啊”。即,该驱动信息在该场景中的意图在于驱动三维虚拟人说出“今天的天气有点冷啊”这句语音。

[0154] 又例如:在聊天机器人的场景中,获取到的驱动信息为用户说出的“今天的天气有点冷啊”这句语音,则结合该场景、经过自然语言理解技术处理后,得到的驱动意图为——“最近寒潮来了,需要多加一些衣服”。即,该驱动信息在该场景中的意图在于驱动三维虚拟人说出“最近寒潮来了,需要多加一些衣服”这句语音。

[0155] 语音合成技术,即TTS(Text To Speech)技术,是一种能够将文本转化为语音的技术。

[0156] 该实施例中,驱动三维虚拟人以对应的口型发出对应的语音。具体的,获取到驱动信息后,基于自然语言理解技术获取对应的驱动意图;进而基于语音合成技术得到驱动意图对应的语音数据以及口型数据;进而驱动三维虚拟人做出该口型数据对应的口型,同时,驱动客户端的喇叭发出该语音数据对应的语音,从而达到三维虚拟人以相应口型发出相应语音的效果。

[0157] 例如:得到驱动意图——“最近寒潮来了,需要多加一些衣服”后,基于语音合成技术得到“最近寒潮来了,需要多加一些衣服”这句话的语音数据以及口型数据。通过该语音数据驱动手机终端的喇叭播放这句话的语音的同时,基于该口型数据,驱动聊天界面中的三维虚拟人做出对应的口型,从而达到三维虚拟人以相应口型说出“最近寒潮来了,需要多加一些衣服”的效果。

[0158] 需要说明的是,该实施例只是示例性的说明,不应对本公开的功能和使用范围造成限制。

[0159] 在一实施例中,基于预设的自然语言理解技术,对该驱动信息进行处理,获取该驱动信息对应的驱动数据,包括:

[0160] 基于预设的自然语言理解技术,对该驱动信息进行处理,获取该驱动信息对应的驱动意图;

[0161] 基于该驱动意图与预设的姿势数据集的匹配结果,获取该驱动意图对应的姿势数据。

[0162] 基于该驱动数据,驱动该三维虚拟人执行对应的用于应答的交互动作,包括:驱动该三维虚拟人执行该姿势数据对应的姿势。

[0163] 该实施例中,驱动三维虚拟人作出对应的姿势。具体的,获取到驱动信息后,基于

自然语言理解技术获取对应的驱动意图；进而将该驱动意图与预设的姿势数据集进行匹配，进而根据匹配结果，获取该驱动意图对应的姿势数据；进而驱动三维虚拟人执行该姿势数据对应的姿势。

[0164] 例如：云端服务器中的聊天机器人系统内预设姿势数据集，该姿势数据集中存储着的每一姿势数据均可驱动三维虚拟人作出对应的姿势。具体的，该姿势数据集中存储着用驱动三维虚拟人做出“胜利姿势”的姿势数据、用以驱动三维虚拟人做出“跑步姿势”的姿势数据、用以驱动三维虚拟人做出“青蛙跳姿势”的姿势数据。

[0165] 小刚通过手机终端的麦克风输入语音信息——“学青蛙跳”。该语音信息经过云端服务器中的聊天机器人系统的自然语言理解处理后，得到驱动意图——“青蛙跳”。将该驱动意图与姿势数据集进行匹配，得到匹配结果——用以驱动三维虚拟人做出“青蛙跳姿势”的姿势数据；进而使用该姿势数据驱动聊天界面中的三维虚拟人做出青蛙跳的姿势。

[0166] 需要说明的是，该实施例只是示例性的说明，不应对本公开的功能和使用范围造成限制。

[0167] 下面参考图4对本公开一实施例的三维虚拟人驱动的实施逻辑进行描述，以展示本公开在实际应用中的大致实施逻辑。

[0168] 图4示出了本公开一实施例的三维虚拟人驱动的实施逻辑：以包含有用户的人物图像的视频为基础，进行三维重建，得到用户的三维虚拟人的模型。

[0169] 通过语音检测技术(VAD, Voice Activity Detection, 一种用于检测语音信号是否存在的技术)检测到用户的语音后，将语音发送给云端服务器。由云端服务器对语音进行语音识别，得到语音对应的文本；进而对语音对应的文本进行自然语言理解，得到应答结果对应的文本；进而对应答结果对应的文本进行语音合成，驱动三维虚拟人的模型的面部表情。必要时，驱动三维虚拟人的模型的面部表情时，驱动喇叭播放应答结果对应的语音，从而达到三维虚拟人以相应口型与用户进行语音交互的效果。

[0170] 需要说明的是，该实施例只是示例性的说明，不应对本公开的功能和使用范围造成限制。

[0171] 下面参考图5对本公开一实施例的三维虚拟人驱动的实施逻辑进行详细描述，以展示本公开在实际应用时的详细实施逻辑。

[0172] 图5示出了本公开一实施例的三维虚拟人驱动的实施逻辑：在对三维虚拟人的面部进行控制，以使得三维虚拟人与交互的过程中，实施逻辑可以分为如下两条。

[0173] 1、视频侧：该实施例中，通过手机摄像头获取到包含有人物图像的视频帧、以及深度图；进而根据人物图像的拍摄角度进行视频帧和深度图的选择；选择出拍摄角度符合预设条件的视频帧和深度图后，在此基础上由服务器端基于RGBD算法对人物图像进行三维重建、模型渲染，从而得到了三维虚拟人的模型。

[0174] 2、音频侧：该实施例中，通过手机麦克风进行音频的采集；对采集到的音频进行语音检测，确定是否有用户的语音；获取到用户语音后，对用户语音进行语音识别，得到用户语音对应的文本；对用户语音对应的文本进行自然语言理解，得到应答结果；对应答结果进行语音合成，得到应答结果对应的语音数据、口型数据。

[0175] 进而根据音频侧得到的应答结果对应的口型数据对视频侧得到的三维虚拟人的模型的面部进行控制，同时使得根据音频侧得到的应答结果对应的语音数据驱动应用语音



的播放,从而达到三维虚拟人以相应口型与用户进行语音交互的效果。

[0176] 需要说明的是,该实施例只是示例性的说明,不应对本公开的功能和使用范围造成限制。

[0177] 根据本公开一实施例,如图6所示,还提供了一种三维虚拟人驱动装置,所述装置包括:

[0178] 第一获取模块410,配置为获取带有深度信息的人物图像;

[0179] 第二获取模块420,配置为基于所述深度信息,对所述人物图像进行三维重建,获取所述人物图像对应的三维虚拟人;

[0180] 第三获取模块430,配置为获取触发驱动的驱动信息;

[0181] 驱动模块440,配置为基于所述驱动信息,驱动所述三维虚拟人执行对应的动作。

[0182] 在本公开的一示例性实施例中,事先设置各拍摄角度分别对应的采集槽,所述第一获取模块410配置为:从各所述采集槽分别获取各拍摄角度的带有深度信息的人物图像。

[0183] 在本公开的一示例性实施例中,所述第一获取模块410配置为:

[0184] 获取包含有人物图像的视频帧、以及所述视频帧的深度信息;

[0185] 基于预设的关键点提取算法,提取所述视频帧中的人物关键点;

[0186] 基于所述人物关键点,获取所述视频帧中的人物图像的拍摄角度;

[0187] 若所述拍摄角度符合预设条件,则获取所述视频帧中的人物图像、以及所述人物图像的深度信息。

[0188] 在本公开的一示例性实施例中,所述第二获取模块420配置为:

[0189] 基于所述深度信息,对所述人物图像进行三维重建,获取所述人物图像对应的三维模型;

[0190] 基于所述人物图像的纹理信息,对所述三维模型进行图形渲染,获取所述人物图像对应的三维虚拟人。

[0191] 在本公开的一示例性实施例中,所述第三获取模块430配置为:获取预设的触发驱动的驱动信息;所述驱动模块440配置为:

[0192] 获取所述驱动信息对应的预设的口型数据、以及对应的预设的语音数据;

[0193] 基于所述语音数据以及所述口型数据,驱动所述三维虚拟人以所述口型数据对应的口型发出所述语音数据对应的语音。

[0194] 在本公开的一示例性实施例中,所述驱动模块440配置为:

[0195] 获取所述驱动信息对应的预设的姿势数据;

[0196] 基于所述姿势数据,驱动所述三维虚拟人执行所述姿势数据对应的姿势。

[0197] 在本公开的一示例性实施例中,所述第三获取模块430配置为:获取文本形式的触发驱动的驱动信息;所述驱动模块440配置为:

[0198] 基于预设的自然语言理解技术,对所述驱动信息进行处理,获取所述驱动信息对应的驱动数据;

[0199] 基于所述驱动数据,驱动所述三维虚拟人执行对应的动作。

[0200] 在本公开的一示例性实施例中,所述第三获取模块430配置为:获取语音形式的触发驱动的驱动信息;所述驱动模块440配置为:

[0201] 对所述驱动信息进行语音识别,获取对应的文本形式的所述驱动信息;

[0202] 基于预设的自然语言理解技术,对所述驱动信息进行处理,获取所述驱动信息对应的驱动数据;

[0203] 基于所述驱动数据,驱动所述三维虚拟人执行对应的动作。

[0204] 在本公开的一示例性实施例中,所述驱动模块440配置为:

[0205] 基于预设的自然语言理解技术,对所述驱动信息进行处理,获取所述驱动信息对应的驱动意图;

[0206] 基于预设的语音合成技术,对所述驱动意图进行处理,获取所述驱动意图对应的语音数据以及口型数据;

[0207] 基于所述语音数据以及口型数据,驱动所述三维虚拟人以所述口型数据对应的口型发出所述语音数据对应的语音。

[0208] 在本公开的一示例性实施例中,所述驱动模块440配置为:

[0209] 基于预设的自然语言理解技术,对所述驱动信息进行处理,获取所述驱动信息对应的驱动意图;

[0210] 基于所述驱动意图与预设的姿势数据集的匹配结果,获取所述驱动意图对应的姿势数据;

[0211] 基于所述姿势数据,驱动所述三维虚拟人执行所述姿势数据对应的姿势。

[0212] 下面参考图7来描述根据本公开实施例的三维虚拟人驱动电子设备50。图7显示的三维虚拟人驱动电子设备50仅仅是一个示例,不应对本公开实施例的功能和使用范围带来任何限制。

[0213] 如图7所示,三维虚拟人驱动电子设备50以通用计算设备的形式表现。三维虚拟人驱动电子设备50的组件可以包括但不限于:上述至少一个处理单元510、上述至少一个存储单元520、连接不同系统组件(包括存储单元520和处理单元510)的总线530。

[0214] 其中,所述存储单元存储有程序代码,所述程序代码可以被所述处理单元510执行,使得所述处理单元510执行本说明书上述示例性方法的描述部分中描述的根据本发明各种示例性实施方式的步骤。例如,所述处理单元510可以执行如图3中所示的各个步骤。

[0215] 存储单元520可以包括易失性存储单元形式的可读介质,例如随机存取存储单元(RAM) 5201和/或高速缓存存储单元5202,还可以进一步包括只读存储单元(ROM) 5203。

[0216] 存储单元520还可以包括具有一组(至少一个)程序模块5205的程序/实用工具5204,这样的程序模块5205包括但不限于:操作系统、一个或者多个应用程序、其它程序模块以及程序数据,这些示例中的每一个或某种组合中可能包括网络环境的实现。

[0217] 总线530可以为表示几类总线结构中的一种或多种,包括存储单元总线或者存储单元控制器、外围总线、图形加速端口、处理单元或者使用多种总线结构中的任意总线结构的局域总线。

[0218] 三维虚拟人驱动电子设备50也可以与一个或多个外部设备600(例如键盘、指向设备、蓝牙设备等)通信,还可与一个或者多个使得用户能与该三维虚拟人驱动电子设备50交互的设备通信,和/或与使得该三维虚拟人驱动电子设备50能与一个或多个其它计算设备进行通信的任何设备(例如路由器、调制解调器等等)通信。这种通信可以通过输入/输出(I/O)接口550进行。并且,三维虚拟人驱动电子设备50还可以通过网络适配器560与一个或者多个网络(例如局域网(LAN),广域网(WAN)和/或公共网络,例如因特网)通信。如图所示,

网络适配器560通过总线530与三维虚拟人驱动电子设备50的其它模块通信。应当明白,尽管图中未示出,可以结合三维虚拟人驱动电子设备50使用其它硬件和/或软件模块,包括但不限于:微代码、设备驱动器、冗余处理单元、外部磁盘驱动阵列、RAID系统、磁带驱动器以及数据备份存储系统等。

[0219] 通过以上的实施方式的描述,本领域的技术人员易于理解,这里描述的示例实施方式可以通过软件实现,也可以通过软件结合必要的硬件的方式来实现。因此,根据本公开实施方式的技术方案可以以软件产品的形式体现出来,该软件产品可以存储在一个非易失性存储介质(可以是CD-ROM,U盘,移动硬盘等)中或网络上,包括若干指令以使得一台计算设备(可以是个人计算机、服务器、终端装置、或者网络设备等)执行根据本公开实施方式的方法。

[0220] 在本公开的示例性实施例中,还提供了一种计算机可读存储介质,其上存储有计算机可读指令,当所述计算机可读指令被计算机的处理器执行时,使计算机执行上述方法实施例部分描述的方法。

[0221] 根据本公开的一个实施例,还提供了一种用于实现上述方法实施例中的方法的程序产品,其可以采用便携式紧凑盘只读存储器(CD-ROM)并包括程序代码,并可以在终端设备,例如个人电脑上运行。然而,本发明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,该程序可以被指令执行系统、装置或者器件使用或者与其结合使用。

[0222] 所述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为但不限于电、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(RGM)、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。

[0223] 计算机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指令执行系统、装置或者器件使用或者与其结合使用的程序。

[0224] 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等等,或者上述的任意合适的组合。

[0225] 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发明操作的程序代码,所述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”语言或类似的设计语言。程序代码可以完全地在用户计算设备上执行、部分地在用户设备上执行、作为一个独立的软件包执行、部分在用户计算设备上部分在远程计算设备上执行、或者完全在远程计算设备或服务器上执行。在涉及远程计算设备的情形中,远程计算设备可以通过任意种类的网络,包括局域网(KGN)或广域网(WGN),连接到用户计算设备,或者,可以连接到外部计算设备(例如利用因特网服务提供商来通过因特网连接)。

[0226] 应当注意,尽管在上文详细描述中提及了用于动作执行的设备的若干模块或者单元,但是这种划分并非强制性的。实际上,根据本公开的实施方式,上文描述的两个或更多模块或者单元的特征和功能可以在一个模块或者单元中具体化。反之,上文描述的一个模块或者单元的特征和功能可以进一步划分为由多个模块或者单元来具体化。

[0227] 此外,尽管在附图中以特定顺序描述了本公开中方法的各个步骤,但是,这并非要求或者暗示必须按照该特定顺序来执行这些步骤,或是必须执行全部所示的步骤才能实现期望的结果。附加的或备选的,可以省略某些步骤,将多个步骤合并为一个步骤执行,以及/或者将一个步骤分解为多个步骤执行等。

[0228] 通过以上的实施方式的描述,本领域的技术人员易于理解,这里描述的示例实施方式可以通过软件实现,也可以通过软件结合必要的硬件的方式来实现。因此,根据本公开实施方式的技术方案可以以软件产品的形式体现出来,该软件产品可以存储在一个非易失性存储介质(可以是CD-ROM,U盘,移动硬盘等)中或网络上,包括若干指令以使得一台计算设备(可以是个人计算机、服务器、移动终端、或者网络设备等)执行根据本公开实施方式的方法。

[0229] 本领域技术人员在考虑说明书及实践这里公开的发明后,将容易想到本公开的其它实施方案。本申请旨在涵盖本公开的任何变型、用途或者适应性变化,这些变型、用途或者适应性变化遵循本公开的一般性原理并包括本公开未公开的本技术领域中的公知常识或惯用技术手段。说明书和实施例仅被视为示例性的,本公开的真正范围和精神由所附的权利要求指出。

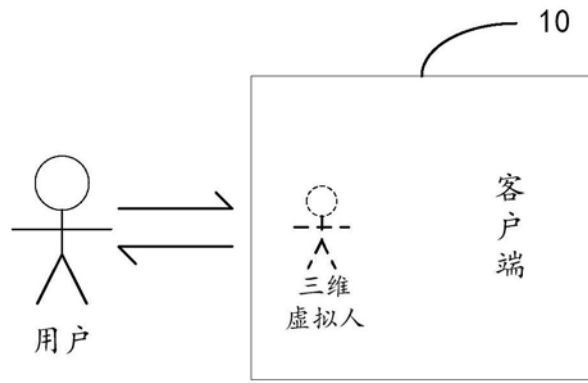


图1A

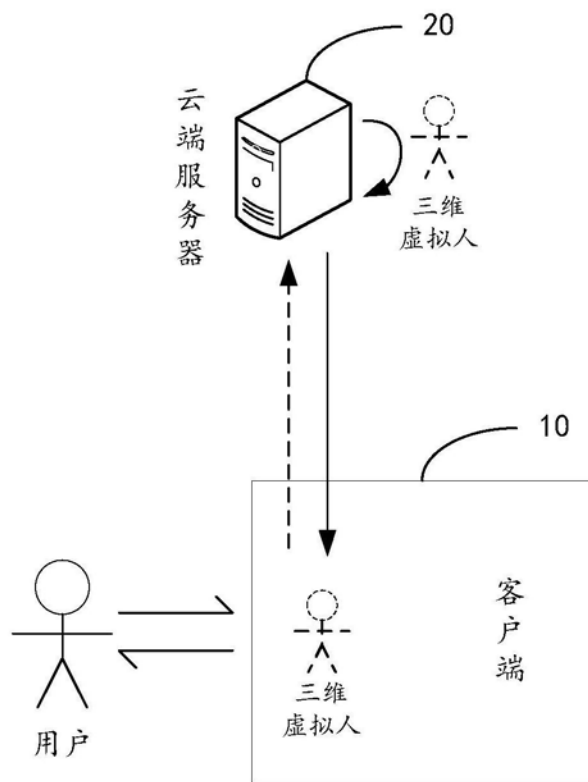


图1B

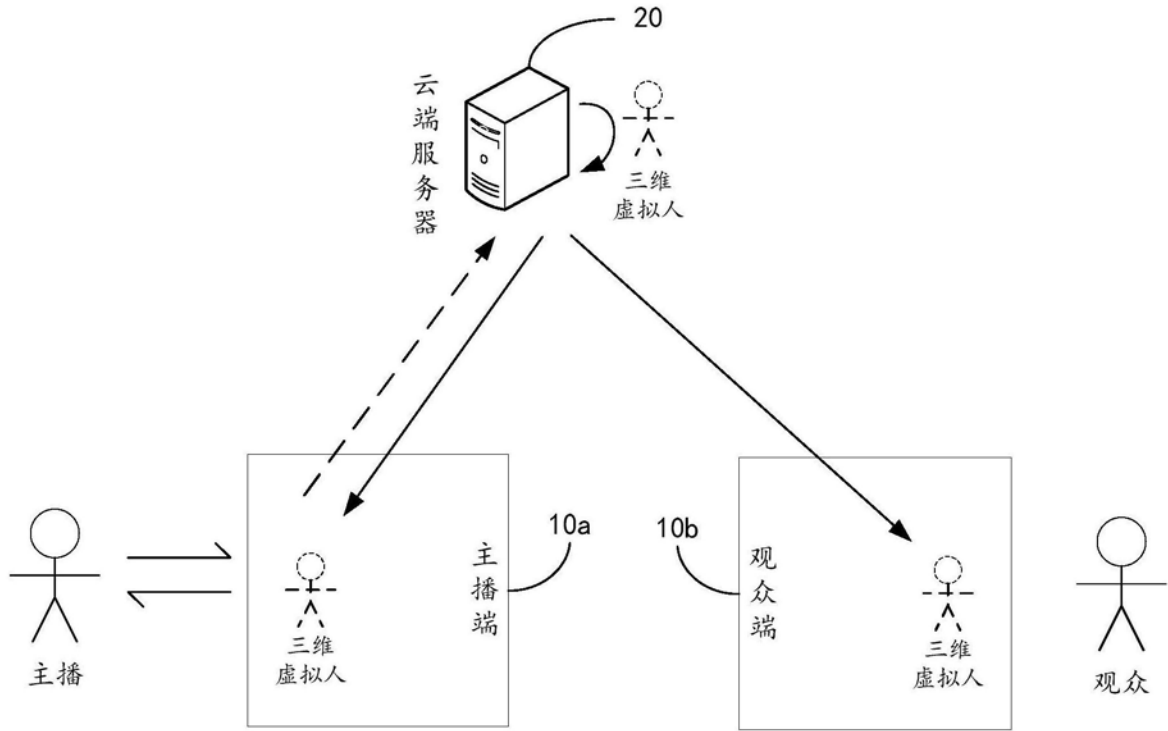


图1C

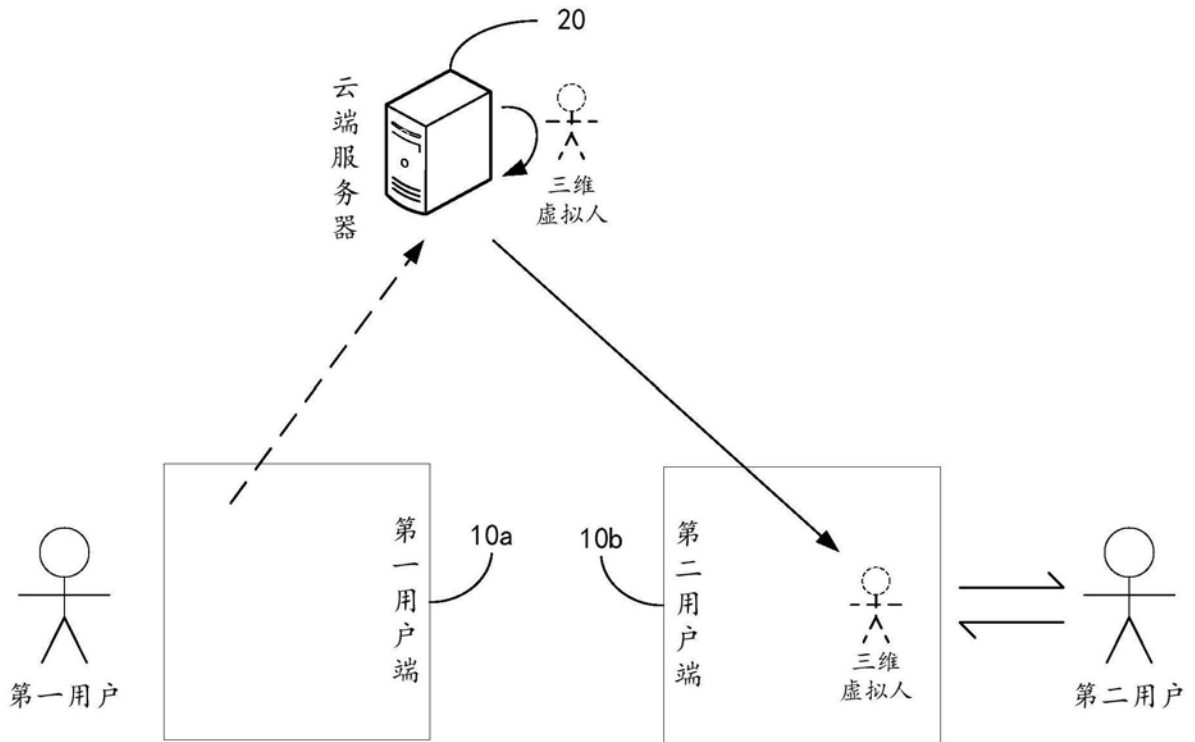


图1D

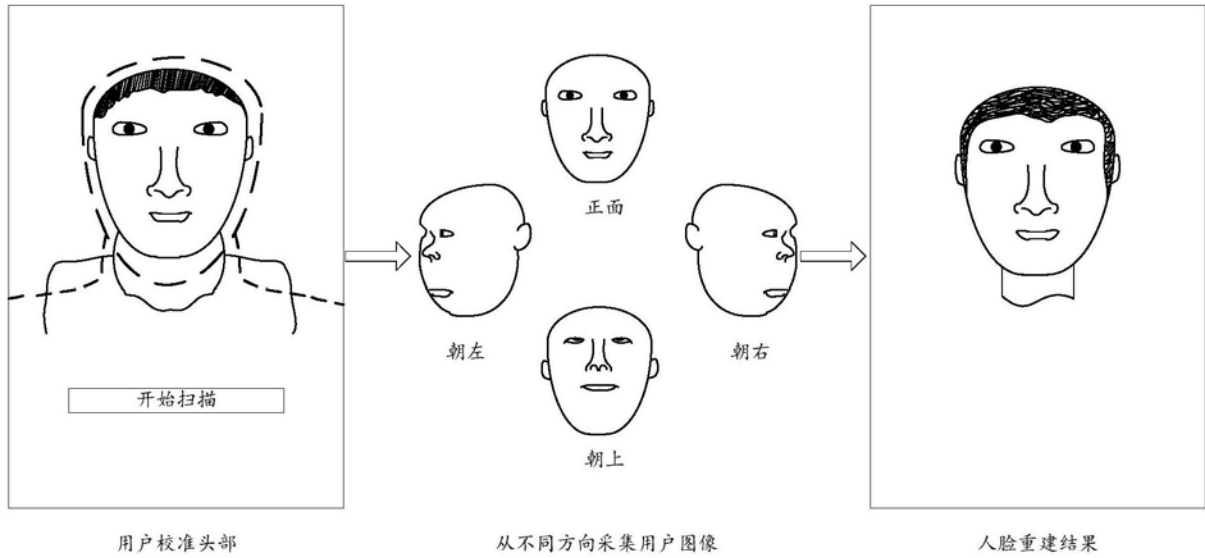


图2

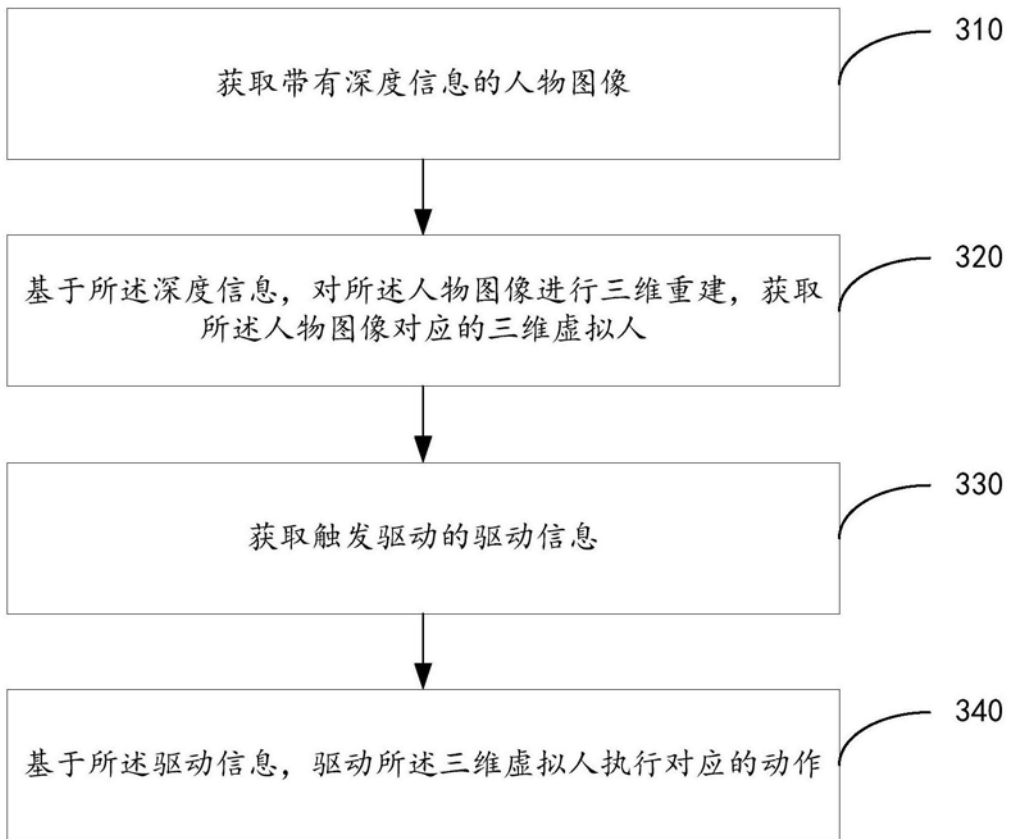


图3

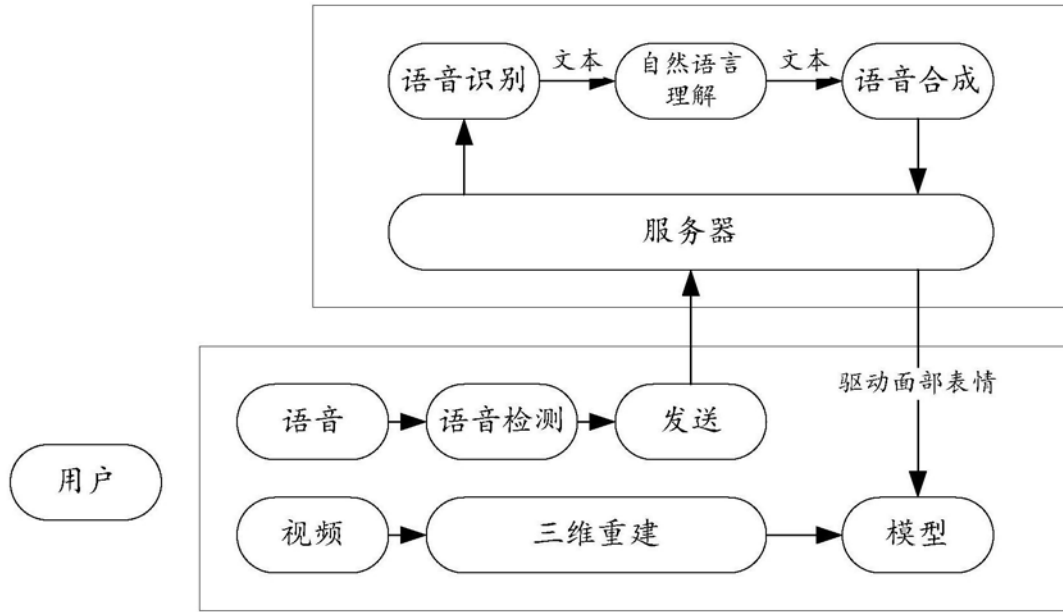


图4

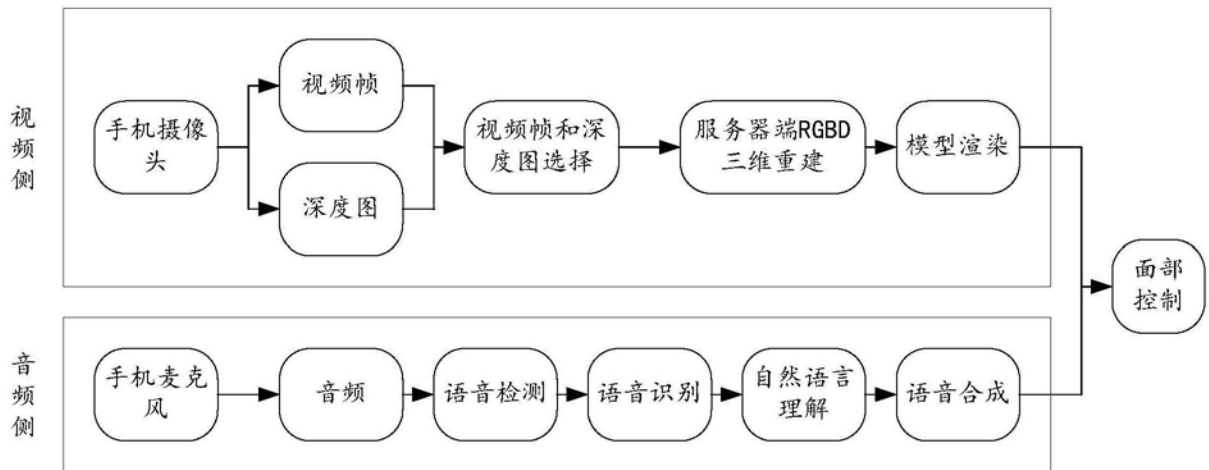


图5

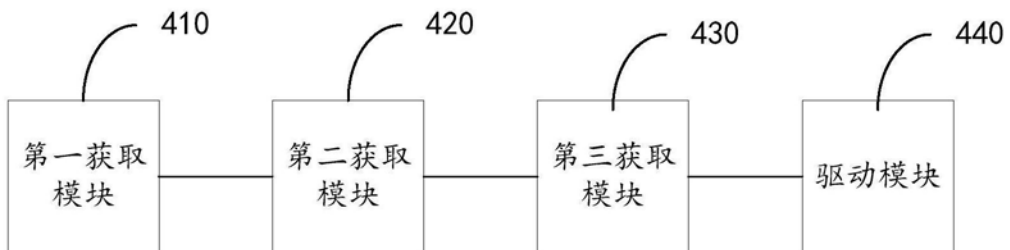


图6



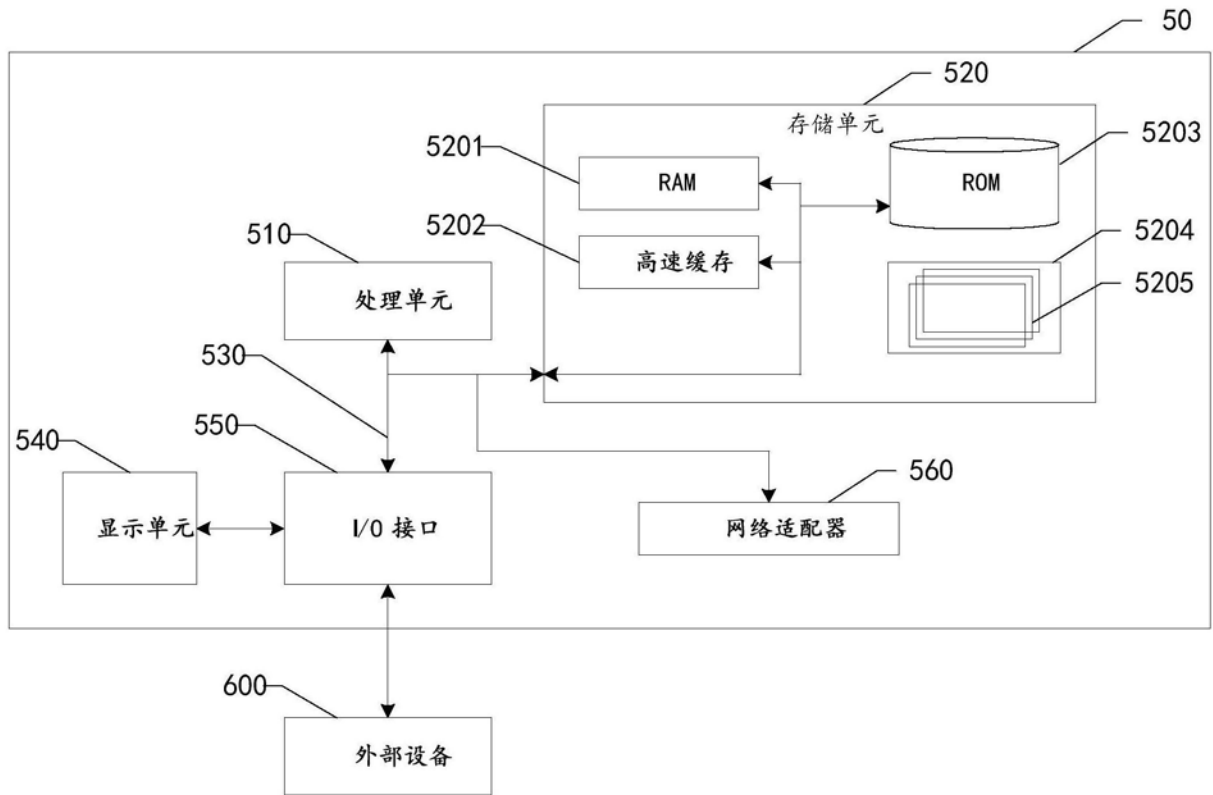


图7