



(12) 发明专利

(10) 授权公告号 CN 111833418 B

(45) 授权公告日 2024. 03. 29

(21) 申请号 202010676929.1

G06T 15/00 (2011.01)

(22) 申请日 2020.07.14

(56) 对比文件

(65) 同一申请的已公布的文献号

申请公布号 CN 111833418 A

CN 108573527 A, 2018.09.25

CN 110189754 A, 2019.08.30

CN 110262665 A, 2019.09.20

(43) 申请公布日 2020.10.27

CN 110362666 A, 2019.10.22

(73) 专利权人 北京百度网讯科技有限公司

CN 110674398 A, 2020.01.10

地址 100085 北京市海淀区上地十街10号

CN 111028330 A, 2020.04.17

百度大厦2层

CN 111145322 A, 2020.05.12

(72) 发明人 陈睿智 彭昊天

CN 111383642 A, 2020.07.07

US 2019197755 A1, 2019.06.27

(74) 专利代理机构 北京英赛嘉华知识产权代理

有限责任公司 11204

审查员 赵鹏翔

专利代理师 王达佐 马晓亚

(51) Int. Cl.

G06T 13/20 (2011.01)

G06T 13/40 (2011.01)

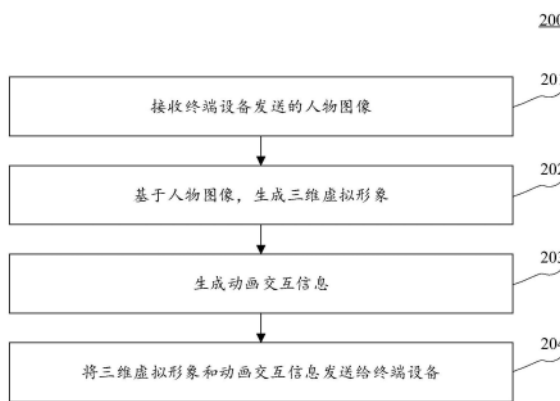
权利要求书3页 说明书11页 附图5页

(54) 发明名称

动画交互方法、装置、设备以及存储介质

(57) 摘要

本申请公开了动画交互方法、装置、设备以及存储介质,涉及图像处理、三维建模、增强现实技术领域。该方法的一具体实施方式包括:接收终端设备发送的人物图像;基于人物图像,生成三维虚拟形象,其中,三维虚拟形象与人物图像中的人物相似;生成动画交互信息,其中,动画交互信息包括交互表情序列帧;将三维虚拟形象和动画交互信息发送给终端设备。该实施方式将人物图像中的人物替换为与其相似的三维虚拟形象,并利用动画交互信息驱动三维虚拟形象陪伴用户,从而使虚拟陪伴的呈现形态更加多样,提升了虚拟陪伴的呈现效果质量和整体互动质量。



1. 一种动画交互方法,包括:
 - 接收终端设备发送的人物图像;
 - 基于所述人物图像,生成三维虚拟形象,其中,所述三维虚拟形象与所述人物图像中的人物相似;
 - 生成动画交互信息,其中,所述动画交互信息包括交互表情序列帧;
 - 将所述三维虚拟形象和所述动画交互信息发送给所述终端设备,
 - 所述生成动画交互信息,包括:
 - 识别所述人物图像中的人物数量和环境信息,以得到所述人物图像所处的场景信息;
 - 基于所述人物图像中的人物数量和环境信息,生成所述人物图像中的人物之间交互的与所述人物图像所处的场景相匹配的动画交互信息。
2. 根据权利要求1所述的方法,其中,所述动画交互信息还包括交互语音。
3. 根据权利要求1所述的方法,其中,所述基于所述人物图像,生成三维虚拟形象,包括:
 - 对所述人物图像中的人物的面部器官进行分类,得到所述面部器官属于多种预设虚拟面部器官类型的权值;
 - 基于所述权值,对所述多种预设虚拟面部器官类型对应的虚拟面部器官进行加权,生成所述人物图像中的人物的虚拟面部器官;
 - 基于所述人物图像中的人物的虚拟面部器官,生成所述三维虚拟形象。
4. 根据权利要求1或2所述的方法,其中,所述生成动画交互信息,包括:
 - 接收所述终端设备发送的用户语音;
 - 识别所述用户语音的内容和/或用户心情;
 - 基于所述用户语音的内容和/或用户心情,生成与所述用户交互的动画交互信息。
5. 一种动画交互方法,包括:
 - 向服务器发送人物图像,以及接收所述服务器返回的三维虚拟形象和动画交互信息,其中,所述三维虚拟形象与所述人物图像中的人物相似,所述动画交互信息包括交互表情序列帧;
 - 基于所述交互表情序列帧对所述三维虚拟形象进行渲染,生成所述三维虚拟形象的交互动画;
 - 将所述交互动画融合到所述人物图像中进行显示,
 - 生成动画交互信息,包括:
 - 识别所述人物图像中的人物数量和环境信息,以得到所述人物图像所处的场景信息;
 - 基于所述人物图像中的人物数量和环境信息,生成所述人物图像中的人物之间交互的与所述人物图像所处的场景相匹配的动画交互信息。
6. 根据权利要求5所述的方法,其中,所述动画交互信息还包括交互语音;以及在所述将所述交互动画融合到所述人物图像中进行显示的同时,还包括:
 - 同步播放所述交互语音。
7. 根据权利要求5或6所述的方法,其中,所述方法还包括:
 - 采集用户输入的用户语音,以及向所述服务器发送所述用户语音;以及
 - 所述接收所述服务器返回的动画交互信息,包括:

接收所述服务器返回的基于所述用户语音生成的与所述用户交互的动画交互信息。

8. 一种动画交互装置,包括:

接收模块,被配置成接收终端设备发送的人物图像;

第一生成模块,被配置成基于所述人物图像,生成三维虚拟形象,其中,所述三维虚拟形象与所述人物图像中的人物相似;

第二生成模块,被配置成生成动画交互信息,其中,所述动画交互信息包括交互表情序列帧;

发送模块,被配置成将所述三维虚拟形象和所述动画交互信息发送给所述终端设备,所述第二生成模块进一步被配置成:

识别所述人物图像中的人物数量和环境信息,以得到所述人物图像所处的场景信息;

基于所述人物图像中的人物数量和环境信息,生成所述人物图像中的人物之间交互的与所述人物图像所处的场景相匹配的动画交互信息。

9. 根据权利要求8所述的装置,其中,所述动画交互信息还包括交互语音。

10. 根据权利要求8所述的装置,其中,所述第一生成模块进一步被配置成:

对所述人物图像中的人物的面部器官进行分类,得到所述面部器官属于多种预设虚拟面部器官类型的权值;

基于所述权值,对所述多种预设虚拟面部器官类型对应的虚拟面部器官进行加权,生成所述人物图像中的人物的虚拟面部器官;

基于所述人物图像中的人物的虚拟面部器官,生成所述三维虚拟形象。

11. 根据权利要求8或9所述的装置,其中,所述第二生成模块进一步被配置成:

接收所述终端设备发送的用户语音;

识别所述用户语音的内容和/或用户心情;

基于所述用户语音的内容和/或用户心情,生成与所述用户交互的动画交互信息。

12. 一种动画交互装置,包括:

发送及接收模块,被配置成向服务器发送人物图像,以及接收所述服务器返回的三维虚拟形象和动画交互信息,其中,所述三维虚拟形象与所述人物图像中的人物相似,所述动画交互信息包括交互表情序列帧;

渲染及生成模块,被配置成基于所述交互表情序列帧对所述三维虚拟形象进行渲染,生成所述三维虚拟形象的交互动画;

显示模块,被配置成将所述交互动画融合到所述人物图像中进行显示,

所述渲染及生成模块进一步被配置成:

识别所述人物图像中的人物数量和环境信息,以得到所述人物图像所处的场景信息;

基于所述人物图像中的人物数量和环境信息,生成所述人物图像中的人物之间交互的与所述人物图像所处的场景相匹配的动画交互信息。

13. 根据权利要求12所述的装置,其中,所述动画交互信息还包括交互语音;以及所述装置还包括:

播放模块,被配置成同步播放所述交互语音。

14. 根据权利要求12或13所述的装置,其中,所述装置还包括:

采集及发送模块,被配置成采集用户输入的用户语音,以及向所述服务器发送所述用

户语音;以及

所述发送及接收模块进一步被配置成:

接收所述服务器返回的基于所述用户语音生成的与所述用户交互的动画交互信息。

15. 一种电子设备,包括:

一个或多个处理器;

存储装置,其上存储有一个或多个程序,

当所述一个或多个程序被所述一个或多个处理器执行,使得所述一个或多个处理器实现如权利要求1-4中任一所述的方法或者实现如权利要求5-7中任一所述的方法。

16. 一种计算机可读介质,其上存储有计算机程序,其中,所述计算机程序被处理器执行时实现如权利要求1-4中任一所述的方法或者实现如权利要求5-7中任一所述的方法。

动画交互方法、装置、设备以及存储介质

技术领域

[0001] 本申请实施例涉及计算机技术领域,具体涉及图像处理、三维建模、增强现实技术领域,尤其涉及动画交互方法、装置、设备以及存储介质。

背景技术

[0002] 人工智能(Artificial Intelligence, AI)是研究、开发用于模拟、延伸和扩展人的智能的理论、方法、技术及应用系统的一门新的技术科学。人工智能是计算机科学的一个分支,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,该领域的研究包括机器人、语言识别、图像识别、自然语言处理和专家系统等。人工智能从诞生以来,理论和技术日益成熟,应用领域也不断扩大。

[0003] 虚拟陪伴作为人工智能的重要应用技术,在越来越多的物联网场景中有了深入的应用。但现有的虚拟陪伴多以语音形式为主,呈现形式单调。

发明内容

[0004] 本申请实施例提出了动画交互方法、装置、设备以及存储介质。

[0005] 第一方面,本申请实施例提出了一种动画交互方法,包括:接收终端设备发送的人物图像;基于人物图像,生成三维虚拟形象,其中,三维虚拟形象与人物图像中的人物相似;生成动画交互信息,其中,动画交互信息包括交互表情序列帧;将三维虚拟形象和动画交互信息发送给终端设备。

[0006] 第二方面,本申请实施例提出了一种动画交互方法,包括:向服务器发送人物图像,以及接收服务器返回的三维虚拟形象和动画交互信息,其中,三维虚拟形象与人物图像中的人物相似,动画交互信息包括交互表情序列帧;基于交互表情序列帧对三维虚拟形象进行渲染,生成三维虚拟形象的交互动画;将交互动画融合到人物图像中进行显示。

[0007] 第三方面,本申请实施例提出了一种动画交互装置,包括:接收模块,被配置成接收终端设备发送的人物图像;第一生成模块,被配置成基于人物图像,生成三维虚拟形象,其中,三维虚拟形象与人物图像中的人物相似;第二生成模块,被配置成生成动画交互信息,其中,动画交互信息包括交互表情序列帧;发送模块,被配置成将三维虚拟形象和动画交互信息发送给终端设备。

[0008] 第四方面,本申请实施例提出了一种动画交互装置,包括:发送及接收模块,被配置成向服务器发送人物图像,以及接收服务器返回的三维虚拟形象和动画交互信息,其中,三维虚拟形象与人物图像中的人物相似,动画交互信息包括交互表情序列帧;渲染及生成模块,被配置成基于交互表情序列帧对三维虚拟形象进行渲染,生成三维虚拟形象的交互动画;显示模块,被配置成将交互动画融合到人物图像中进行显示。

[0009] 第五方面,本申请实施例提出了一种电子设备,包括:至少一个处理器;以及至少一个处理器通信连接的存储器;其中,存储器存储有可被至少一个处理器执行的指令,指令被至少一个处理器执行,以使至少一个处理器能够执行如第一方面中任一实现方式描述

的方法或如第二方面中任一实现方式描述的方法。

[0010] 第六方面,本申请实施例提出了一种存储有计算机指令的非瞬时计算机可读存储介质,计算机指令用于使计算机执行如第一方面中任一实现方式描述的方法或者如第二方面中任一实现方式描述的方法。

[0011] 第七方面,本申请实施例提供了一种计算机程序产品,包括计算机程序,计算机程序在被处理器执行时实现根据如第一方面中任一实现方式描述的方法或者实现根据如第二方面中任一实现方式描述的方法。

[0012] 本申请实施例提供的动画交互方法、装置、设备以及存储介质,首先接收终端设备发送的人物图像;然后基于人物图像,生成与人物图像中的人物相似的三维虚拟形象,并生成动画交互信息;最后将三维虚拟形象和动画交互信息发送给终端设备。将人物图像中的人物替换为与其相似的三维虚拟形象,并利用动画交互信息驱动三维虚拟形象陪伴用户,从而使虚拟陪伴的呈现形态更加多样,提升了虚拟陪伴的呈现效果质量和整体互动质量。进而极大地提升了用户的参与度和认同感,从而增加应用动画交互方法的产品的竞争力和影响力。

[0013] 应当理解,本部分所描述的内容并非旨在标识本公开的实施例的关键或重要特征,也不用于限制本公开的范围。本公开的其它特征将通过以下的说明书而变得容易理解。

附图说明

[0014] 通过阅读参照以下附图所作的对非限制性实施例所作的详细描述,本申请的其它特征、目的和优点将会变得更明显。附图用于更好地理解本方案,不构成对本申请的限定。其中:

[0015] 图1是本申请可以应用于其中的示例性系统架构;

[0016] 图2是根据本申请的动画交互方法的一个实施例的流程图;

[0017] 图3是根据本申请的动画交互方法的又一个实施例的流程图;

[0018] 图4是根据本申请的动画交互方法的另一个实施例的流程图;

[0019] 图5是可以实现本申请实施例的动画交互方法的场景图;

[0020] 图6是根据本申请的动画交互装置的一个实施例的结构示意图;

[0021] 图7是根据本申请的动画交互装置的又一个实施例的结构示意图;

[0022] 图8是用来实现本申请实施例的动画交互方法的电子设备的框图。

具体实施方式

[0023] 以下结合附图对本申请的示范性实施例做出说明,其中包括本申请实施例的各种细节以助于理解,应当将它们认为仅仅是示范性的。因此,本领域普通技术人员应当认识到,可以对这里描述的实施例做出各种改变和修改,而不会背离本申请的范围和精神。同样,为了清楚和简明,以下的描述中省略了对公知功能和结构的描述。

[0024] 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。下面将参考附图并结合实施例来详细说明本申请。

[0025] 图1示出了可以应用本申请的动画交互方法或动画交互装置的实施例的示例性系统架构100。

[0026] 如图1所示,系统架构100中可以包括终端设备101、网络102和服务器103。网络102用以在终端设备101和服务器103之间提供通信链路的介质。网络102可以包括各种连接类型,例如有线、无线通信链路或者光纤电缆等等。

[0027] 用户可以使用终端设备101通过网络102与服务器103交互,以接收或发送消息等。终端设备101上可以安装有各种客户端应用,例如3D捏脸软件、智能相框软件等。终端设备101可以对从服务器101接收到的三维虚拟形象和动画交互信息等数据进行渲染等处理,并呈现处理结果(例如融合显示交互动画,以及同步播放交互语音)。

[0028] 终端设备101可以是硬件,也可以是软件。当终端设备101为硬件时,可以是各种电子设备,包括但不限于电子相框、智能手机和平板电脑等等。当终端设备101为软件时,可以安装在上述电子设备中。其可以实现成多个软件或软件模块,也可以实现成单个软件或软件模块。在此不做具体限定。

[0029] 服务器103可以是提供各种服务的服务器,例如3D捏脸软件或智能相框软件的后台服务器。后台服务器可以对从终端设备101接收到的人物图像等数据进行分析等处理,并将处理结果(例如三维虚拟形象和动画交互信息)反馈给终端设备101。

[0030] 需要说明的是,服务器103可以是硬件,也可以是软件。当服务器103为硬件时,可以实现成多个服务器组成的分布式服务器集群,也可以实现成单个服务器。当服务器103为软件时,可以实现成多个软件或软件模块(例如用来提供分布式服务),也可以实现成单个软件或软件模块。在此不做具体限定。

[0031] 需要说明的是,本申请实施例所提供的动画交互方法可以由服务器103执行,相应地,动画交互装置设置于服务器103中;本申请实施例所提供的动画交互方法也可以由终端设备101执行,相应地,动画交互装置设置于终端设备101。

[0032] 应该理解,图1中的终端设备、网络和服务器的数目仅仅是示意性的。根据实现需要,可以具有任意数目的终端设备、网络和服务器。

[0033] 继续参考图2,其示出了根据本申请的动画交互方法的一个实施例的流程200。该动画交互方法包括以下步骤:

[0034] 步骤201,接收终端设备发送的人物图像。

[0035] 在本实施例中,动画交互方法的执行主体(例如图1所示的服务器103)可以接收终端设备(例如图1所示的终端设备101)发送的人物图像。

[0036] 实践中,终端设备可以包括但不限于电子相框、智能手机和平板电脑等等。终端设备上可以安装有3D捏脸软件或智能相框软件。用户可以通过3D捏脸软件或智能相框软件将人物图像上传至服务器。其中,人物图像通常是真实世界中的人物的二维图像。在一些实施例中,3D捏脸软件或智能相框软件可以预先获取终端设备的摄像头权限,通过终端设备的摄像头拍摄人物图像。在一些实施例中,3D捏脸软件或智能相框软件可以预先获取终端设备的相册读取权限,读取终端设备的相册中存储的人物图像。

[0037] 步骤202,基于人物图像,生成三维虚拟形象。

[0038] 在本实施例中,上述执行主体可以基于人物图像,生成三维虚拟形象。其中,三维虚拟形象与人物图像中的人物相似,可以是对人物图像中的人物进行风格化,突显其个人特点的三维动画人物。

[0039] 在一些实施例中,上述执行主体可以预先存储大量三维虚拟形象,上述执行主体

可以提取人物图像中的人物的特征,并与预先存储的每个三维虚拟形象匹配,将特征高度匹配的三维虚拟形象作为人物图像中的人物的三维虚拟形象。

[0040] 在一些实施例中,上述执行主体可以利用PTA(photo to avatar,虚拟形象自动生成)技术对人物图像中的人物进行3D捏脸,生成对应的三维虚拟形象。具体地,上述执行主体可以首先对人物图像中的人物的面部器官进行分类,得到面部器官属于多种预设虚拟面部器官类型的权值;然后基于权值,对多种预设虚拟面部器官类型对应的虚拟面部器官进行加权,生成人物图像中的人物的虚拟面部器官;最后基于人物图像中的人物的虚拟面部器官,生成三维虚拟形象。通过对预先存储的多种类型的虚拟面部器官进行融合,能够融合出任意类型的虚拟面部器官。进一步地,基于相似度权值对预先存储的多种类型的虚拟面部器官进行融合,能够融合出与人物图像中的人物的面部器官高度相似的虚拟面部器官。其中,预设虚拟面部器官可以包括但不限于眼睛、鼻子、嘴巴、眉毛、耳朵等等。例如,对于眼睛,可以预先存储多种虚拟眼型(包括但不限于桃花眼、瑞凤眼、睡凤眼、柳叶眼和杏眼等等),计算人物图像中的人物的眼睛与这些虚拟眼型的相似度权值,基于相似度权值对这些虚拟眼型进行融合,即可得到人物图像中的人物的虚拟眼睛。

[0041] 步骤203,生成动画交互信息。

[0042] 在本实施例中,上述执行主体可以生成动画交互信息。通常,上述执行主体可以预先存储一套通用的表情基底,包括各种各样的表情帧。上述执行主体可以对至少部分表情帧进行组合,生成交互表情序列帧。此时,动画交互信息可以包括交互表情序列帧。可选的,上述执行主体还可以为表情序列帧设计匹配的交互语音。此时,动画交互信息还可以包括交互语音。

[0043] 步骤204,将三维虚拟形象和动画交互信息发送给终端设备。

[0044] 在本实施例中,上述执行主体可以将三维虚拟形象和动画交互信息发送给终端设备。这样,终端设备可以基于交互表情序列帧对三维虚拟形象进行渲染,生成交互动画,并融合到人物图像中进行显示,把虚拟世界中的三维虚拟形象叠加在现实世界中的人物图像进行互动,实现了人物图像的增强现实。可选的,在动画交互信息还包括交互语音的情况下,在将交互动画融合到人物图像中进行显示的同时,上述执行主体还可以同步播放交互语音,实现了多呈现形态的虚拟陪伴。其中,交互动画中的三维虚拟形象依次作出交互表情序列帧中的表情。

[0045] 本申请实施例提供的动画交互方法,首先接收终端设备发送的人物图像;然后基于人物图像,生成与人物图像中的人物相似的三维虚拟形象,并生成动画交互信息;最后将三维虚拟形象和动画交互信息发送给终端设备。将人物图像中的人物替换为与其相似的三维虚拟形象,并利用动画交互信息驱动三维虚拟形象陪伴用户,从而使虚拟陪伴的呈现形态更加多样,提升了虚拟陪伴的呈现效果质量和整体互动质量。进而极大地提升了用户的参与度和认同感,从而增加应用动画交互方法的产品竞争力和影响力。

[0046] 进一步参考图3,其示出了根据本申请的动画交互方法的又一个实施例的流程300。该动画交互方法包括以下步骤:

[0047] 步骤301,接收终端设备发送的人物图像。

[0048] 步骤302,基于人物图像,生成三维虚拟形象。

[0049] 在本实施例中,步骤301-302具体操作已在图2所示的实施例中步骤201-202进行

了详细的介绍,在此不再赘述。

[0050] 步骤303,识别人物图像中的人物数量和环境信息。

[0051] 在本实施例中,动画交互方法的执行主体(例如图1所示的服务器103)可以生成默认的动画交互信息。

[0052] 在一些实施例中,无论用户是否输入语音,上述执行主体均可以生成的默认动画交互信息存储备用。

[0053] 在一些实施例中,只有用户未输入语音的情况下,上述执行主体才会生成默认的动画交互信息。

[0054] 通常,默认的动画交互信息与人物图像所处的场景相匹配。具体地,上述执行主体可以识别人物图像中的人物数量和环境信息,以得到人物图像所处的场景信息。例如,上述执行主体可以利用目标检测模型检测人物图像中的人体框,并基于检测出的人体框的数量,确定人物图像中的人物数量。上述执行主体可以利用目标识别模型识别人物图像的背景中的物体,并基于识别出的物体确定人物图像中的环境信息。其中,目标检测模型和目标识别模型可以通过深度学习方式,预先训练得到的神经网络模型。

[0055] 步骤304,基于人物图像中的人物数量和环境信息,生成人物图像中的人物之间交互的动画交互信息。

[0056] 在本实施例中,上述执行主体可以基于人物图像中的人物数量和环境信息,生成人物图像中的人物之间交互的动画交互信息。其中,基于人物图像中的人物数量可以确定交互参与人数,基于人物图像中的环境信息可以生成与其匹配的交互内容。其中,交互参与人数不大于人物图像中的人物数量,通常情况下等于人物图像中的人物数量。例如,若人物图像中存在3个人,且处于商场中,则动画交互信息可以是3个人讨论在商场购物的交互信息。

[0057] 步骤305,接收终端设备发送的用户语音。

[0058] 在本实施例中,在用户输入语音的情况下,上述执行主体可以生成与用户交互的动画交互信息。具体地,接收终端设备(例如图1所示的终端设备101)发送的用户语音,生成与用户语音匹配的动画交互信息。

[0059] 实践中,终端设备可以包括但不限于电子相框、智能手机和平板电脑等等。终端设备上可以安装有3D捏脸软件或智能相框软件。3D捏脸软件或智能相框软件可以预先获取终端设备的录音权限,通过终端设备的麦克风采集用户输入的用户语音。

[0060] 步骤306,识别用户语音的内容和/或用户心情。

[0061] 在本实施例中,上述执行主体可以识别用户语音的内容和/或用户心情。其中,用户语音的内容可以通过对用户语音进行语音转换文字来得到。用户的心情可以从用户语音和/或用户语音的内容中提取情绪特征信息来确定。

[0062] 对于仅识别用户语音的内容的方案,上述执行主体可以对用户语音进行语音转换文字,得到用户语音的内容。

[0063] 对于仅识别用户心情的方案,上述执行主体可以直接从用户语音中提取用户的发音特点,分析对应的情绪特征信息。其中,发音特点可以包括但不限于韵律、节奏、语速、语调、修辞、声音强度等。例如,若用户语音的语调较为欢快,则确定用户心情比较高兴。

[0064] 对于同时识别用户语音的内容和用户心情的方案,上述执行主体可以对用户语音

进行语音转换文字,得到用户语音的内容。并且,上述执行主体不仅可以从用户语音中提取用户的发音特点,分析对应的情绪特征信息,还可以从用户的语音内容中提取带有情绪信息的词语,分析对应的情绪特征信息。

[0065] 步骤307,基于用户语音的内容和/或用户心情,生成与用户交互的动画交互信息。

[0066] 在本实施例中,上述执行主体可以基于用户语音的内容和/或用户心情,生成与用户交互的动画交互信息。其中,基于用户心情可以确定与其匹配的表情。基于用户语音的内容可以生成与其匹配的交互内容。这里,基于与用户心情匹配的表情和/或与用户心情匹配的交互内容,可以生成与用户交互的动画交互信息。

[0067] 对于仅基于与用户心情匹配的表情生成动画交互信息的方案,动画交互信息可以是描述人物作出一系列与用户心情匹配的表情面部动作的信息。

[0068] 对于仅基于与用户心情匹配的交互内容生成动画交互信息的方案,动画交互信息可以是描述人物说出一系列与用户心情匹配的交互内容的口型动作的信息。

[0069] 对于同时基于与用户心情匹配的表情和与用户心情匹配的交互内容生成动画交互信息的方案,动画交互信息不仅可以包括描述人物作出一系列与用户心情匹配的表情面部动作的信息,还可以包括描述人物说出一系列与用户心情匹配的交互内容的口型动作的信息。

[0070] 步骤308,将三维虚拟形象和动画交互信息发送给终端设备。

[0071] 在本实施例中,步骤308具体操作已在图2所示的实施例中步骤204进行了详细的介绍,在此不再赘述。

[0072] 从图3中可以看出,与图2对应的实施例相比,本实施例中的动画交互方法的流程300突出了生成动画交互信息的步骤。由此,在本实施例描述的方案中,在用户未输入语音的情况下,生成人物图像中的人物之间交互的动画交互信息发送给终端设备,驱动人物图像中的不同人物之间交互,且交互内容与人物图像所处的场景匹配;在用户输入语音的情况下,生成与用户交互的动画交互信息发送给终端设备,驱动人物图像中的人物与用户交互,且交互内容与用户语音匹配。针对不同的情况,生成不同的动画交互信息,使得交互更加具有针对性。

[0073] 进一步参考图4,其示出了根据本申请的动画交互方法的另一个实施例的流程400。该动画交互方法包括以下步骤:

[0074] 步骤401,向服务器发送人物图像,以及接收服务器返回的三维虚拟形象和动画交互信息。

[0075] 在本实施例中,动画交互方法的执行主体(例如图1所示的终端设备101)可以向服务器(例如图1所示的服务器103)发送人物图像,以及接收服务器返回的三维虚拟形象和动画交互信息。

[0076] 实践中,终端设备可以包括但不限于电子相框、智能手机和平板电脑等等。终端设备上可以安装有3D捏脸软件或智能相框软件。用户可以通过3D捏脸软件或智能相框软件将人物图像上传至服务器。其中,人物图像通常是真实世界中的人物的二维图像。在一些实施例中,3D捏脸软件或智能相框软件可以预先获取终端设备的摄像头权限,通过终端设备的摄像头拍摄人物图像。在一些实施例中,3D捏脸软件或智能相框软件可以预先获取终端设备的相册读取权限,读取终端设备的相册中存储的人物图像。

[0077] 服务器可以基于人物图像生成三维虚拟形象和动画交互信息。其中,三维虚拟形象与人物图像中的人物相似,可以是对人物图像中的人物进行风格化,突显其个人特点的三维动画人物。动画交互信息可以包括交互表情序列帧。可选的,动画交互信息还可以包括交互语音。

[0078] 在一些实施例中,动画交互信息可以与人物图像所处的场景相匹配。具体地,服务器可以首先识别人物图像中的人物数量和环境信息;然后基于人物图像中的人物数量和环境信息,生成人物图像中的人物之间交互的动画交互信息。在用户未输入语音的情况下,生成人物图像中的人物之间交互的动画交互信息发送给终端设备,驱动人物图像中的不同人物之间交互,且交互内容与人物图像所处的场景匹配。

[0079] 在一些实施例中,动画交互信息可以与用户语音相匹配。具体地,3D捏脸软件或智能相框软件还可以预先获取终端设备的录音权限,通过终端设备的麦克风采集用户输入的用户语音,以及向服务器发送用户语音。服务器可以首先识别用户语音的内容和/或用户心情;然后基于用户语音的内容和/或用户心情,生成与用户交互的动画交互信息。在用户输入语音的情况下,生成与用户交互的动画交互信息发送给终端设备,驱动人物图像中的人物与用户交互,且交互内容与用户语音匹配。针对不同的情况,生成不同的动画交互信息,使得交互更加具有针对性。

[0080] 步骤402,基于交互表情序列帧对三维虚拟形象进行渲染,生成三维虚拟形象的交互动画。

[0081] 在本实施例中,上述执行主体可以基于交互表情序列帧对三维虚拟形象进行渲染,生成三维虚拟形象的交互动画。其中,交互动画中的三维虚拟形象依次作出交互表情序列帧中的表情。

[0082] 步骤403,将交互动画融合到人物图像中进行显示。

[0083] 在本实施例中,上述执行主体可以将交互动画融合到人物图像中进行显示,把虚拟世界中的三维虚拟形象叠加在现实世界中的人物图像进行互动,实现了人物图像的增强现实。可选的,在动画交互信息还包括交互语音的情况下,在将交互动画融合到人物图像中进行显示的同时,上述执行主体可以同步播放交互语音,实现了多呈现形态的虚拟陪伴。

[0084] 本申请实施例提供的动画交互方法,首先向服务器发送人物图像,以及接收服务器返回的与人物图像中的人物相似的三维虚拟形象和动画交互信息;然后基于交互表情序列帧对三维虚拟形象进行渲染,生成三维虚拟形象的交互动画;最后将交互动画融合到人物图像中进行显示,以及同步播放交互语音。将人物图像中的人物替换为与其相似的三维虚拟形象,并利用动画交互信息驱动三维虚拟形象陪伴用户,从而使虚拟陪伴的呈现形态更加多样,提升了虚拟陪伴的呈现效果质量和整体互动质量。进而极大地提升了用户的参与度和认同感,从而增加应用动画交互方法的产品竞争力和影响力。

[0085] 为了便于理解,图5示出了可以实现本申请实施例的动画交互方法的场景图。如图5所示,电子相框501包括麦克风5011、显示器5012、扬声器5013、图像存储器5014、三维虚拟形象存储器5015、动画交互信息存储器5016、三维虚拟形象驱动器5017和图像合成器5018。在用户将人物图像存储到图像存储器5014之后,会触发人物图像上传服务器502的操作。当人物图像首次上传服务器502后,服务器502可以根据人物图像中的所有人物,利用PTA技术生成与其相对应的三维虚拟形象,并下载到三维虚拟形象存储器5015中。随后,服务器502

可以根据人物图像中人物数量和环境信息,生成与人物图像所处的场景相匹配的动画交互信息(包括表情序列帧和交互语音),并下载到动画交互信息存储器5016,作为默认的动画交互信息。在运行过程中,若麦克风5011未采集到用户输入的用户语音,会直接按照默认的动画交互信息完成后续的驱动和合成操作。在运行过程中,若麦克风5011采集到用户输入的用户语音,麦克风5011可以将采集到的用户语音上传至服务器502。服务器502可以根据用户语音的内容和用户心情,生成与用户交互的临时的动画交互信息,并下载到动画交互信息存储器5016。此时,会按照临时的动画交互信息完成后续的驱动和合成操作。具体地,在三维虚拟形象驱动器5017中按照动画交互信息驱动三维虚拟形象,生成交互动画。在图像合成器5018中将交互动画融合到人物图像中,并利用显示器5012进行显示。同时,利用扬声器5013同步播放交互语音。

[0086] 进一步参考图6,作为对上述各图所示方法的实现,本申请提供了一种动画交互装置的一个实施例,该装置实施例与图2所示的方法实施例相对应,该装置具体可以应用于各种电子设备中。

[0087] 如图6所示,本实施例的动画交互装置600可以包括:接收模块601、第一生成模块602、第二生成模块603和发送模块604。其中,接收模块601,被配置成接收终端设备发送的人物图像;第一生成模块602,被配置成基于人物图像,生成三维虚拟形象,其中,三维虚拟形象与人物图像中的人物相似;第二生成模块603,被配置成生成动画交互信息,其中,动画交互信息包括交互表情序列帧;发送模块604,被配置成将三维虚拟形象和动画交互信息发送给终端设备。

[0088] 在本实施例中,动画交互装置600中:接收模块601、第一生成模块602、第二生成模块603和发送模块604的具体处理及其所带来的技术效果可分别参考图2对应实施例中的步骤201-204的相关说明,在此不再赘述。

[0089] 在本实施例的一些可选的实现方式中,动画交互信息还包括交互语音。

[0090] 在本实施例的一些可选的实现方式中,第一生成模块602进一步被配置成:对人物图像中的人物的面部器官进行分类,得到面部器官属于多种预设虚拟面部器官类型的权值;基于权值,对多种预设虚拟面部器官类型对应的虚拟面部器官进行加权,生成人物图像中的人物的虚拟面部器官;基于人物图像中的人物的虚拟面部器官,生成三维虚拟形象。

[0091] 在本实施例的一些可选的实现方式中,第二生成模块603进一步被配置成:识别人物图像中的人物数量和环境信息;基于人物图像中的人物数量和环境信息,生成人物图像中的人物之间交互的动画交互信息。

[0092] 在本实施例的一些可选的实现方式中,第二生成模块603进一步被配置成:接收终端设备发送的用户语音;识别用户语音的内容和/或用户心情;基于用户语音的内容和/或用户心情,生成与用户交互的动画交互信息。

[0093] 进一步参考图7,作为对上述各图所示方法的实现,本申请提供了一种动画交互装置的一个实施例,该装置实施例与图4所示的方法实施例相对应,该装置具体可以应用于各种电子设备中。

[0094] 如图7所示,本实施例的动画交互装置700可以包括:发送及接收模块701、渲染及生成模块702和显示及播放模块703。其中,发送及接收模块701,被配置成向服务器发送人物图像,以及接收服务器返回的三维虚拟形象和动画交互信息,其中,三维虚拟形象与人物

图像中的人物相似,动画交互信息包括交互表情序列帧;渲染及生成模块702,被配置成基于交互表情序列帧对三维虚拟形象进行渲染,生成三维虚拟形象的交互动画;显示模块703,被配置成将交互动画融合到人物图像中进行显示。

[0095] 在本实施例中,动画交互装置700中:发送及接收模块701、渲染及生成模块702和显示模块703的具体处理及其所带来的技术效果可分别参考图4对应实施例中的步骤401-403的相关说明,在此不再赘述。

[0096] 在本实施例的一些可选的实现方式中,动画交互信息还包括交互语音;以及动画交互装置700还包括:播放模块(图中未示出),被配置成同步播放所述交互语音。

[0097] 在本实施例的一些可选的实现方式中,动画交互装置700还包括:采集及发送模块(图中未示出),被配置成采集用户输入的用户语音,以及向服务器发送用户语音;以及发送及接收模块701进一步被配置成:接收服务器返回的基于用户语音生成的与用户交互的动画交互信息。

[0098] 根据本申请的实施例,本申请还提供了一种电子设备和一种可读存储介质。

[0099] 如图8所示,是根据本申请实施例动画交互方法的电子设备的框图。电子设备旨在表示各种形式的数字计算机,诸如,膝上型计算机、台式计算机、工作台、个人数字助理、服务器、刀片式服务器、大型计算机、和其它适合的计算机。电子设备还可以表示各种形式的移动装置,诸如,个人数字处理、蜂窝电话、智能电话、可穿戴设备和其它类似的计算装置。本文所示的部件、它们的连接和关系、以及它们的功能仅仅作为示例,并且不意在限制本文中描述的和/或者要求的本申请的实现。

[0100] 如图8所示,该电子设备包括:一个或多个处理器801、存储器802,以及用于连接各部件的接口,包括高速接口和低速接口。各个部件利用不同的总线互相连接,并且可以被安装在公共主板上或者根据需要以其它方式安装。处理器可以对在电子设备内执行的指令进行处理,包括存储在存储器中或者存储器上以在外部输入/输出装置(诸如,耦合至接口的显示设备)上显示GUI的图形信息的指令。在其它实施方式中,若需要,可以将多个处理器和/或多条总线与多个存储器和多个存储器一起使用。同样,可以连接多个电子设备,各个设备提供部分必要的操作(例如,作为服务器阵列、一组刀片式服务器、或者多处理器系统)。图8中以一个处理器801为例。

[0101] 存储器802即为本申请所提供的非瞬时计算机可读存储介质。其中,所述存储器存储有可由至少一个处理器执行的指令,以使所述至少一个处理器执行本申请所提供的动画交互方法。本申请的非瞬时计算机可读存储介质存储计算机指令,该计算机指令用于使计算机执行本申请所提供的动画交互方法。

[0102] 存储器802作为一种非瞬时计算机可读存储介质,可用于存储非瞬时软件程序、非瞬时计算机可执行程序以及模块,如本申请实施例中的动画交互方法对应的程序指令/模块(例如,附图6所示的接收模块601、第一生成模块602、第二生成模块603和发送模块604,或者附图7所示的发送及接收模块701、渲染及生成模块702和显示模块703)。处理器801通过运行存储在存储器802中的非瞬时软件程序、指令以及模块,从而执行服务器的各种功能应用以及数据处理,即实现上述方法实施例中的动画交互方法。

[0103] 存储器802可以包括存储程序区和存储数据区,其中,存储程序区可存储操作系统、至少一个功能所需要的应用程序;存储数据区可存储根据动画交互方法的电子设备的

使用所创建的数据等。此外,存储器802可以包括高速随机存取存储器,还可以包括非瞬时存储器,例如至少一个磁盘存储器件、闪存器件、或其他非瞬时固态存储器件。在一些实施例中,存储器802可选包括相对于处理器801远程设置的存储器,这些远程存储器可以通过网络连接至动画交互方法的电子设备。上述网络的实例包括但不限于互联网、企业内部网、局域网、移动通信网及其组合。

[0104] 动画交互方法的电子设备还可以包括:输入装置803和输出装置804。处理器801、存储器802、输入装置803和输出装置804可以通过总线或者其他方式连接,图8中以通过总线连接为例。

[0105] 输入装置803可接收输入的数字或字符信息,以及产生与动画交互方法的电子设备的用户设置以及功能控制有关的键信号输入,例如触摸屏、小键盘、鼠标、轨迹板、触摸板、指示杆、一个或者多个鼠标按钮、轨迹球、操纵杆等输入装置。输出装置804可以包括显示设备、辅助照明装置(例如,LED)和触觉反馈装置(例如,振动电机)等。该显示设备可以包括但不限于,液晶显示器(LCD)、发光二极管(LED)显示器和等离子体显示器。在一些实施方式中,显示设备可以是触摸屏。

[0106] 此处描述的系统和技术各种实施方式可以在数字电子电路系统、集成电路系统、专用ASIC(专用集成电路)、计算机硬件、固件、软件、和/或它们的组合中实现。这些各种实施方式可以包括:实施在一个或者多个计算机程序中,该一个或者多个计算机程序可在包括至少一个可编程处理器的可编程系统上执行和/或解释,该可编程处理器可以是专用或者通用可编程处理器,可以从存储系统、至少一个输入装置、和至少一个输出装置接收数据和指令,并且将数据和指令传输至该存储系统、该至少一个输入装置、和该至少一个输出装置。

[0107] 这些计算程序(也称作程序、软件、软件应用、或者代码)包括可编程处理器的机器指令,并且可以利用高级过程和/或面向对象的编程语言、和/或汇编/机器语言来实施这些计算程序。如本文使用的,术语“机器可读介质”和“计算机可读介质”指的是用于将机器指令和/或数据提供给可编程处理器的任何计算机程序产品、设备、和/或装置(例如,磁盘、光盘、存储器、可编程逻辑装置(PLD)),包括,接收作为机器可读信号的机器指令的机器可读介质。术语“机器可读信号”指的是用于将机器指令和/或数据提供给可编程处理器的任何信号。

[0108] 为了提供与用户的交互,可以在计算机上实施此处描述的系统和技术,该计算机具有:用于向用户显示信息的显示装置(例如,CRT(阴极射线管)或者LCD(液晶显示器)监视器);以及键盘和指向装置(例如,鼠标或者轨迹球),用户可以通过该键盘和该指向装置来将输入提供给计算机。其它种类的装置还可以用于提供与用户的交互;例如,提供给用户的反馈可以是任何形式的传感反馈(例如,视觉反馈、听觉反馈、或者触觉反馈);并且可以用任何形式(包括声输入、语音输入或者、触觉输入)来接收来自用户的输入。

[0109] 可以将此处描述的系统和技术实施在包括后台部件的计算系统(例如,作为数据服务器)、或者包括中间件部件的计算系统(例如,应用服务器)、或者包括前端部件的计算系统(例如,具有图形用户界面或者网络浏览器的用户计算机,用户可以通过该图形用户界面或者该网络浏览器来与此处描述的系统和技术实施方式交互)、或者包括这种后台部件、中间件部件、或者前端部件的任何组合的计算系统中。可以通过任何形式或者介质的数

字数据通信(例如,通信网络)来将系统的部件相互连接。通信网络的示例包括:局域网(LAN)、广域网(WAN)和互联网。

[0110] 计算机系统可以包括客户端和服务端。客户端和服务端一般远离彼此并且通常通过通信网络进行交互。通过在相应的计算机上运行并且彼此具有客户端-服务器关系的计算机程序来产生客户端和服务端的关系。

[0111] 根据本申请的技术方案,首先接收终端设备发送的人物图像;然后基于人物图像,生成与人物图像中的人物相似的三维虚拟形象,并生成动画交互信息;最后将三维虚拟形象和动画交互信息发送给终端设备。将人物图像中的人物替换为与其相似的三维虚拟形象,并利用动画交互信息驱动三维虚拟形象陪伴用户,从而使虚拟陪伴的呈现形态更加多样,提升了虚拟陪伴的呈现效果质量和整体互动质量。进而极大地提升了用户的参与度和认同感,从而增加应用动画交互方法的产品竞争力和影响力。

[0112] 应该理解,可以使用上面所示的各种形式的流程,重新排序、增加或删除步骤。例如,本发申请中记载的各步骤可以并行地执行也可以顺序地执行也可以不同的次序执行,只要能够实现本申请公开的技术方案所期望的结果,本文在此不进行限制。

[0113] 上述具体实施方式,并不构成对本申请保护范围的限制。本领域技术人员应该明白的是,根据设计要求和因素,可以进行各种修改、组合、子组合和替代。任何在本申请的精神和原则之内所作的修改、等同替换和改进等,均应包含在本申请保护范围之内。

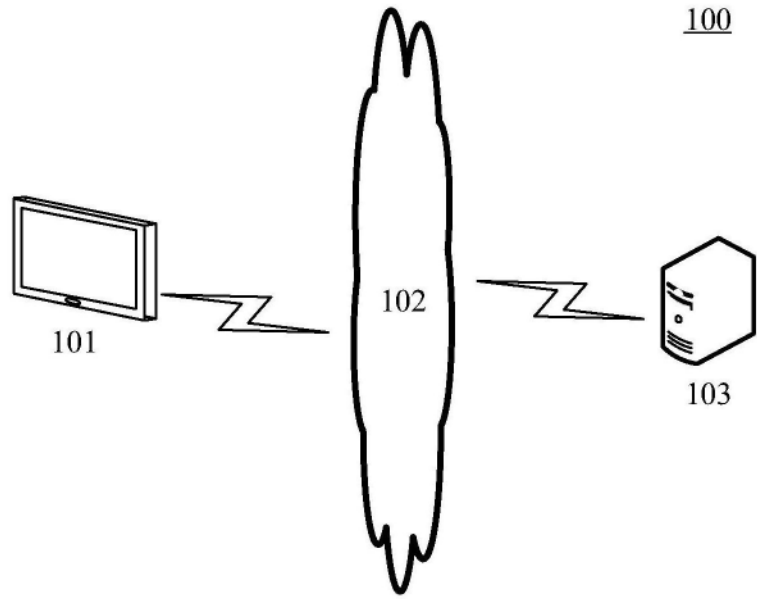


图1

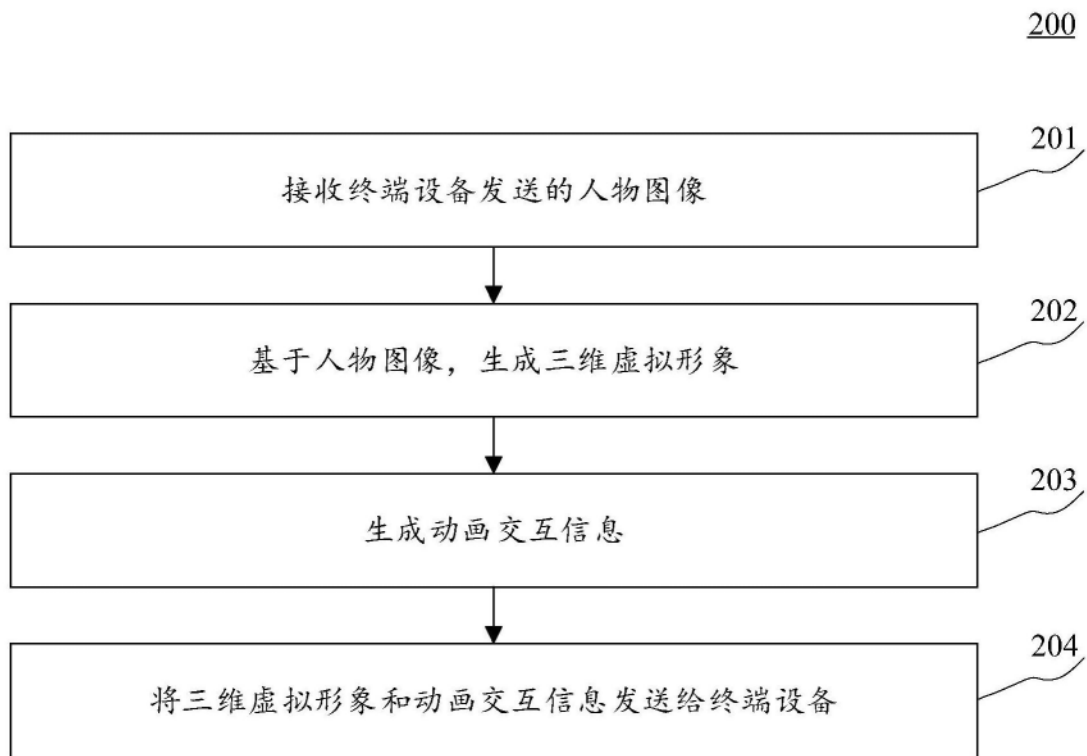


图2

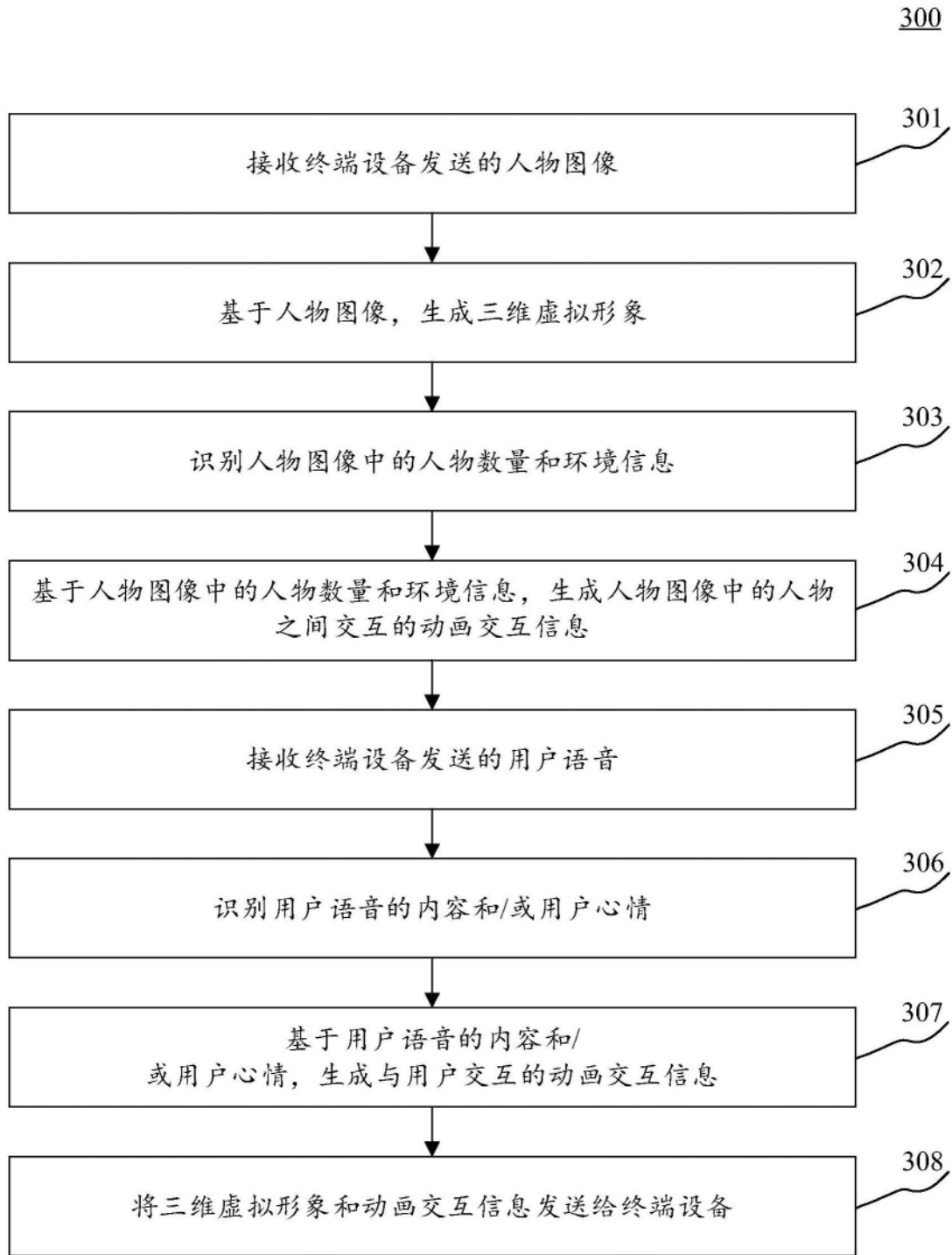


图3

400

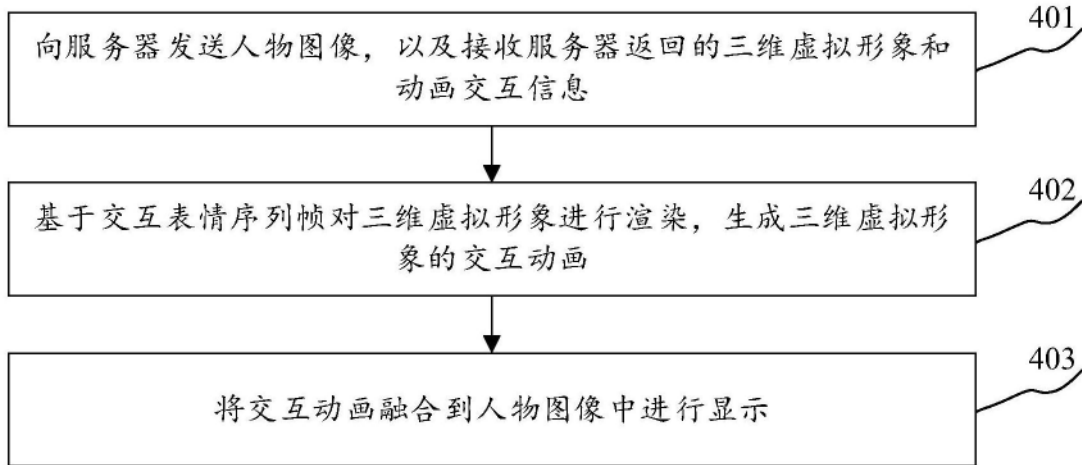


图4

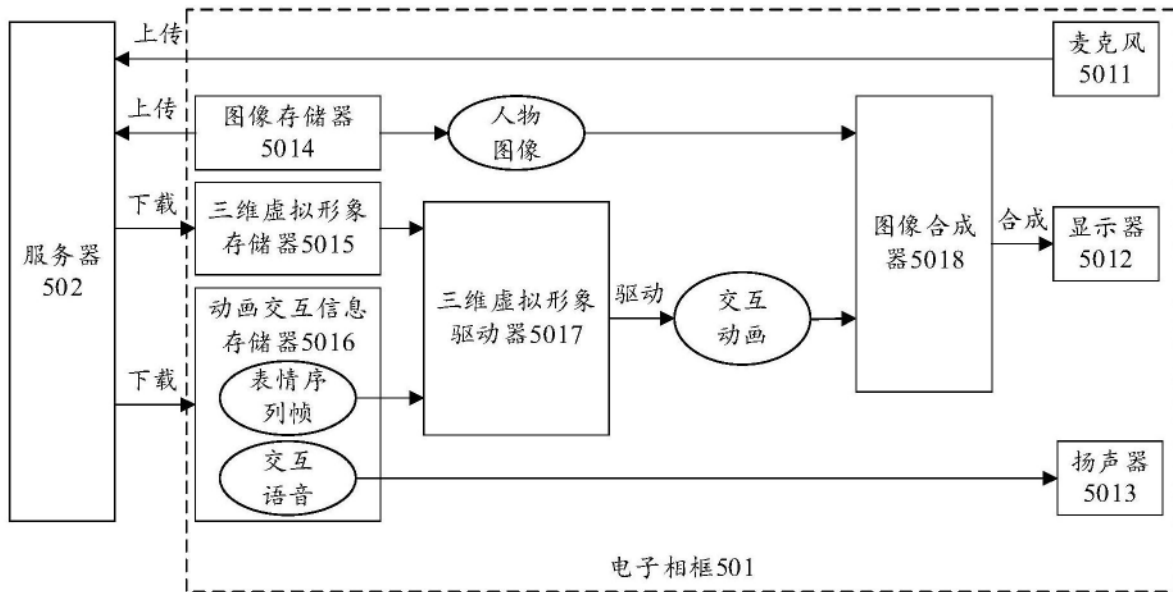


图5

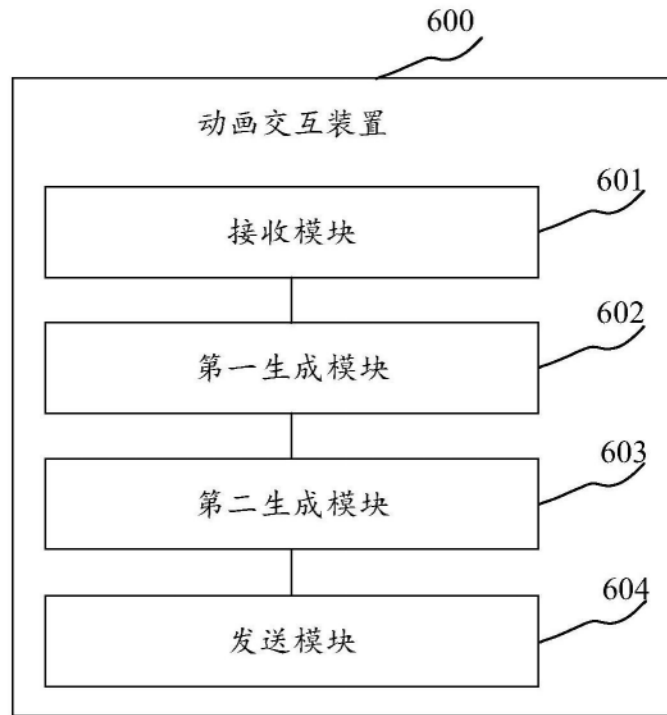


图6

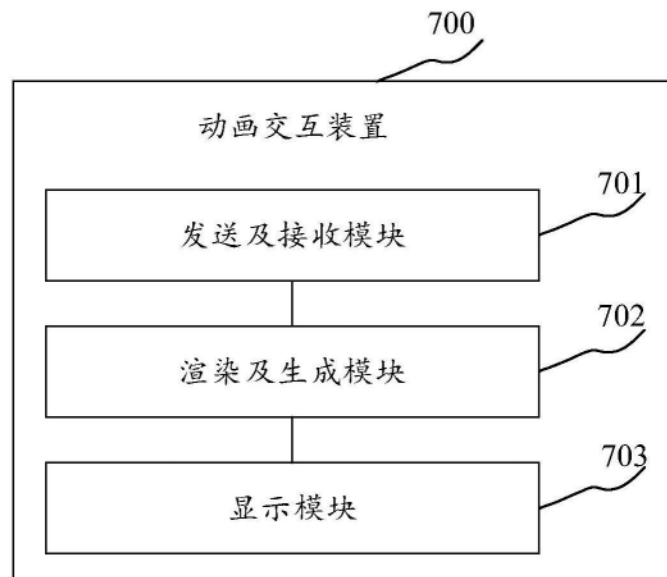


图7

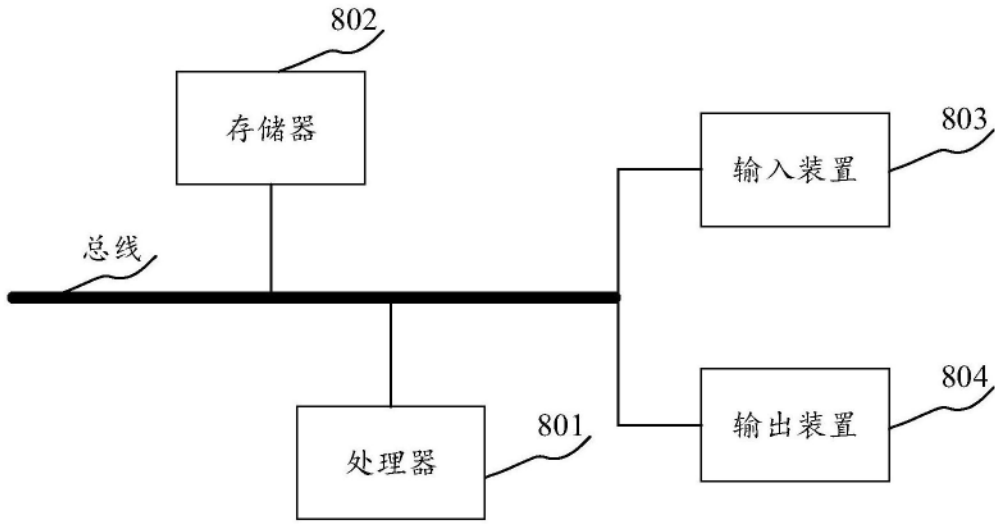


图8