



(12) 发明专利申请

(10) 申请公布号 CN 117643048 A

(43) 申请公布日 2024.03.01

(21) 申请号 202180100603.7

(51) Int.Cl.

(22) 申请日 2021.07.15

H04N 7/14 (2006.01)

(85) PCT国际申请进入国家阶段日

2024.01.15

(86) PCT国际申请的申请数据

PCT/EP2021/069675 2021.07.15

(87) PCT国际申请的公布数据

W02023/284958 EN 2023.01.19

(71) 申请人 瑞典爱立信有限公司

地址 瑞典斯德哥尔摩

(72) 发明人 A·厄尔厄赛利 N·图迪纳

E·阿坎 J·C·埃沃特 张赛

(74) 专利代理机构 中国专利代理(香港)有限公

司 72001

专利代理师 徐予红 陈岚

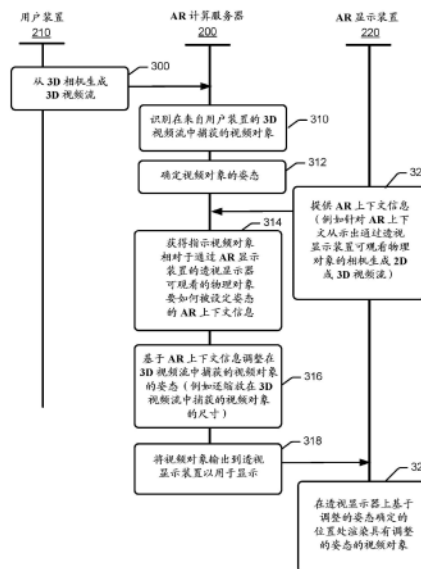
权利要求书5页 说明书9页 附图5页

(54) 发明名称

基于来自增强现实显示装置的增强现实上下文信息调整来自用户装置的3D视频流中的视频对象的姿态

(57) 摘要

一种增强现实AR计算服务器(200)包括网络接口(202)、处理器(204)和存储由处理器可执行以执行操作的指令的存储器(206)。网络接口配置成在会议会话期间通过网络从用户装置接收三维(3D)视频流。所述操作识别在3D视频流中捕获的视频对象,并且确定在3D视频流中捕获的视频对象的姿态。所述操作从AR显示装置获得指示视频对象相对于通过AR显示装置的透视显示器可观看的物理对象要如何被设定姿态的AR上下文信息,并且基于AR上下文信息调整在3D视频流中捕获的视频对象的姿态。所述操作将视频对象输出到透视显示器以用于显示。公开了相关方法和计算机程序产品。



1. 一种增强现实AR计算服务器(200),包括:
配置成在会议会话期间通过网络从用户装置(210)接收三维3D视频流的网络接口(202);
连接到所述网络接口(202)的处理器(204);以及
存储能够由所述处理器(204)执行以执行以下操作的指令的存储器(206):
识别(310)在所述3D视频流中捕获的视频对象;
确定(312)在所述3D视频流中捕获的所述视频对象的姿态;
从AR显示装置(220)获得(314)指示所述视频对象相对于通过所述AR显示装置(220)的透视显示器(234)能够观看的物理对象要如何被设定姿态的AR上下文信息;
基于所述AR上下文信息调整(316)在所述3D视频流中捕获的所述视频对象的姿态;以及
将所述视频对象输出(318)到所述透视显示器(234)以用于显示。
2. 根据权利要求1所述的AR计算服务器(200),其中:
确定(312)在所述3D视频流中捕获的所述视频对象的所述姿态的所述操作包括确定在所述3D视频流中捕获的脸部的特征的姿态;以及
基于所述AR上下文信息调整(316)在所述3D视频流中捕获的所述视频对象的姿态的所述操作包括基于在所述3D视频流中捕获的所述脸部的所述特征的所述姿态与所述脸部的所述特征相对于通过所述AR显示装置(220)的所述透视显示器(234)能够观看的所述物理对象要如何被设定姿态的所述AR上下文信息指示的比较来旋转和/或平移在所述3D视频流中捕获的所述脸部的所述特征。
3. 根据权利要求1至2中任一项所述的AR计算服务器(200),其中,所述AR上下文信息指示所述物理对象的姿态,并且基于所述AR上下文信息来调整(316)在所述3D视频流中捕获的所述视频对象的姿态的所述操作包括基于所述视频对象的所述姿态与所述物理对象的所述姿态的比较来调整在所述3D视频流中捕获的所述视频对象的姿态。
4. 根据权利要求1至3中任一项所述的AR计算服务器(200),其中:
获得(314)所述AR上下文信息的所述操作包括确定所述AR显示装置(220)的所述透视显示器(234)相对于在来自所述AR显示装置(220)的相机(232)的视频流中捕获的所述物理对象的姿态;以及
调整(316)在所述3D视频流中捕获的所述视频对象的姿态的所述操作包括基于所述视频对象的所述姿态与所述AR显示装置(220)的所述透视显示器(234)相对于在来自所述AR显示装置(220)的相机(232)的视频流中捕获的所述物理对象的所述姿态的比较来调整在所述3D视频流中捕获的所述视频对象的姿态。
5. 根据权利要求4所述的AR计算服务器(200),其中,确定所述AR显示装置(220)的所述透视显示器(234)相对于在来自所述AR显示装置(220)的所述相机(232)的所述视频流中捕获的所述物理对象的所述姿态的所述操作包括:
识别在来自所述AR显示装置(220)的所述相机(232)的所述视频流中捕获的多个物理对象的姿态;
从所述多个物理对象之中选择所述物理对象中的一个,这基于所述物理对象中的所选择的一个满足上下文选择规则;以及

执行相对于所述物理对象中的所述选择的一个的所述姿态的所述透视显示器 (234) 的所述姿态的所述确定。

6. 根据权利要求5所述的AR计算服务器 (200), 其中, 所述操作进一步包括基于所述物理对象中的一个具有匹配以下中的一个的定义形状的形状来确定在来自所述AR显示装置的所述相机的所述视频流中捕获的所述物理对象中的所述一个满足所述上下文选择规则:

座位, 在所述座位上在所述3D视频流中捕获的所述视频对象要在所述透视显示器 (234) 上被显示具有被观看为看起来由所述座位支撑的姿态;

桌子, 在所述桌子上在所述3D视频流中捕获的所述视频对象要在所述透视显示器 (234) 上被显示具有被观看为看起来由所述桌子支撑的姿态; 以及

地板, 在所述地板上在所述3D视频流中捕获的所述视频对象要在所述透视显示器 (234) 上被显示具有被观看为看起来由所述地板支撑的姿态。

7. 根据权利要求4至6中任一项所述的AR计算服务器 (200), 其中, 所述操作进一步包括:

基于在来自所述AR显示装置 (220) 的所述相机 (232) 的所述视频流中捕获的所述物理对象的颜色和/或阴影, 调整被输出到所述透视显示器 (234) 以用于显示的所述物理对象的颜色和/或阴影。

8. 根据权利要求1至7中任一项所述的AR计算服务器 (200), 其中, 调整 (316) 在所述3D视频流中捕获的所述视频对象的姿态的所述操作进一步包括:

基于在所述3D视频流中捕获的所述视频对象的所述姿态与所述视频对象相对于通过所述AR显示装置 (220) 的所述透视显示器 (234) 能够观看的所述物理对象要如何被设定姿态的所述AR上下文信息指示的比较来旋转和/或平移在所述3D视频流中捕获的所述视频对象的姿态 (400)。

9. 根据权利要求1至8中任一项所述的AR计算服务器 (200), 其中, 调整 (316) 在所述3D视频流中捕获的所述视频对象的姿态的所述操作进一步包括:

基于在所述3D视频流中捕获的所述视频对象的尺寸与通过所述AR显示装置 (220) 的所述透视显示器 (234) 能够观看的所述物理对象的尺寸的所述AR上下文信息指示的比较来缩放在所述3D视频流中捕获的所述视频对象的尺寸 (402)。

10. 根据权利要求1至9中任一项所述的AR计算服务器 (200), 其中, 所述操作进一步包括:

提取 (500) 在所述会议会话期间在较早时间在所述3D视频流中捕获的或者来自另一会议会话的另一3D视频流的所述视频对象的扩展部分的图像, 其中在所述视频对象的所述姿态的所述确定 (312) 的时间, 所述视频对象的所述扩展部分不在所述3D视频流中捕获;

将所述视频对象的所述扩展部分的所述图像存储 (502) 在所述存储器中;

基于在所述3D视频流中捕获的所述视频对象的所述姿态与从所述存储器检索的所述视频对象的所述扩展部分的所述图像的所述姿态的比较, 调整 (504) 从所述存储器检索的所述视频对象的所述扩展部分的所述图像的姿态和/或在所述3D视频流中捕获的所述视频对象的姿态;

基于在所述3D视频流中捕获的所述视频对象的尺寸与从所述存储器检索的所述视频对象的所述扩展部分的所述图像的尺寸的比较, 缩放 (506) 从所述存储器检索的所述视频

对象的所述扩展部分的所述图像的尺寸和/或在所述3D视频流中捕获的所述视频对象的尺寸;以及

将所述视频对象的所述扩展部分的所述图像与在所述3D视频流中捕获的所述视频对象组合(508),以生成被输出(318)到所述透视显示器(234)以用于显示的组合视频对象。

11. 根据权利要求1至10中任一项所述的AR计算服务器(200),其中:

所述AR计算服务器(200)包括网络计算服务器;以及

所述网络接口(202)进一步配置成通过所述网络与所述AR显示装置(220)的所述透视显示器(234)通信。

12. 根据权利要求1至11中任一项所述的AR计算服务器(200),其中,所述视频对象是在所述3D视频流中捕获的场景的多个分量中的一个,以及

调整(316)在所述3D视频流中捕获的所述视频对象的姿态的所述操作包括从所述3D视频流提取所述视频对象而没有所述场景的其它分量;以及

将所述视频对象输出(318)到所述透视显示器(234)以用于显示的所述操作包括输出具有所调整的姿态的所提取的视频对象。

13. 一种由增强现实AR计算服务器进行的方法,包括:

识别(310)在会议会话期间在从用户装置接收的三维3D视频流中捕获的视频对象;

确定(312)在所述3D视频流中捕获的所述视频对象的姿态;

从AR显示装置获得(314)指示所述视频对象相对于通过所述AR显示装置的透视显示器能够观看的物理对象要如何被设定姿态的AR上下文信息;

基于所述AR上下文信息调整(316)在所述3D视频流中捕获的所述视频对象的姿态;以及

将所述视频对象输出(318)到所述透视显示器以用于显示。

14. 根据权利要求13所述的方法,其中:

在所述3D视频流中捕获的所述视频对象的所述姿态的所述确定(312)包括确定在所述3D视频流中捕获的脸部的特征的姿态;以及

基于所述AR上下文信息的所述调整(316)在所述3D视频流中捕获的所述视频对象的所述姿态包括基于在所述3D视频流中捕获的所述脸部的所述特征的所述姿态与所述脸部的所述特征相对于通过所述AR显示装置(220)的所述透视显示器(234)能够观看的所述物理对象要如何被设定姿态的所述AR上下文信息指示的比较来旋转和/或平移在所述3D视频流中捕获的所述脸部的所述特征。

15. 根据权利要求13至14中任一项所述的方法,其中,所述AR上下文信息指示所述物理对象的姿态,并且基于所述AR上下文信息的所述调整(316)在所述3D视频流中捕获的所述视频对象的所述姿态包括基于所述视频对象的所述姿态与所述物理对象的所述姿态的比较来调整在所述3D视频流中捕获的所述视频对象的姿态。

16. 根据权利要求13至15中任一项所述的方法,其中:

所述获得(314)所述AR上下文信息包括确定所述AR显示装置(220)的所述透视显示器(234)相对于在来自所述AR显示装置(220)的相机(232)的视频流中捕获的所述物理对象的姿态;以及

所述调整(316)在所述3D视频流中捕获的所述视频对象的所述姿态包括基于所述视频

对象的所述姿态与所述AR显示装置(220)的所述透视显示器(234)相对于在来自所述AR显示装置(220)的相机(232)的视频流中捕获的所述物理对象的所述姿态的比较来调整在所述3D视频流中捕获的所述视频对象的姿态。

17. 根据权利要求16所述的方法,其中,所述确定(312)所述AR显示装置(220)的所述透视显示器(234)相对于在来自所述AR显示装置(220)的所述相机(232)的所述视频流中捕获的所述物理对象的所述姿态包括:

识别在来自所述AR显示装置(220)的所述相机(232)的所述视频流中捕获的多个物理对象的姿态;

从所述多个物理对象之中选择所述物理对象中的一个,这基于所述物理对象中的所选择的一个满足上下文选择规则;以及

执行相对于所述物理对象中的所述选择的一个的所述姿态的所述透视显示器(234)的所述姿态的所述确定。

18. 根据权利要求17所述的方法,进一步包括基于所述物理对象中的一个具有匹配以下中的一个的定义形状的形状来确定在来自所述AR显示装置的所述相机的所述视频流中捕获的所述物理对象中的所述一个满足所述上下文选择规则:

座位,在所述座位上在所述3D视频流中捕获的所述视频对象要在所述透视显示器(234)上被显示具有被观看为看起来由所述座位支撑的姿态;

桌子,在所述桌子上在所述3D视频流中捕获的所述视频对象要在所述透视显示器(234)上被显示具有被观看为看起来由所述桌子支撑的姿态;以及

地板,在所述地板上在所述3D视频流中捕获的所述视频对象要在所述透视显示器(234)上被显示具有被观看为看起来由所述地板支撑的姿态。

19. 根据权利要求16至18中任一项所述的方法,进一步包括:

基于在来自所述AR显示装置(220)的所述相机(232)的所述视频流中捕获的所述物理对象的颜色和/或阴影,调整被输出到所述透视显示器(234)以用于显示的所述物理对象的颜色和/或阴影。

20. 根据权利要求13至19中任一项所述的方法,其中,所述调整(316)在所述3D视频流中捕获的所述视频对象的所述姿态进一步包括:

基于在所述3D视频流中捕获的所述视频对象的所述姿态与所述视频对象相对于通过所述AR显示装置(220)的所述透视显示器(234)能够观看的所述物理对象要如何被设定姿态的所述AR上下文信息指示的比较来旋转和/或平移在所述3D视频流中捕获的所述视频对象的姿态(400)。

21. 根据权利要求13至20中任一项所述的方法,其中,所述调整(316)在所述3D视频流中捕获的所述视频对象的所述姿态进一步包括:

基于在所述3D视频流中捕获的所述视频对象的尺寸与通过所述AR显示装置(220)的所述透视显示器(234)能够观看的所述物理对象的尺寸的所述AR上下文信息指示的比较,缩放在所述3D视频流中捕获的所述视频对象的尺寸(402)。

22. 根据权利要求13至21中任一项所述的方法,进一步包括:

提取(500)在所述会议会话期间在较早时间在所述3D视频流中捕获的或者来自另一会议会话的另一3D视频流的所述视频对象的扩展部分的图像,其中在所述视频对象的所述姿

态的所述确定(312)的时间,所述视频对象的所述扩展部分不在所述3D视频流中捕获;

将所述视频对象的所述扩展部分的所述图像存储(502)在所述存储器中;

基于在所述3D视频流中捕获的所述视频对象的所述姿态与从所述存储器检索的所述视频对象的所述扩展部分的所述图像的姿态的比较,调整(504)从所述存储器检索的所述视频对象的所述扩展部分的所述图像的姿态和/或在所述3D视频流中捕获的所述视频对象的姿态;

基于在所述3D视频流中捕获的所述视频对象的尺寸与从所述存储器检索的所述视频对象的所述扩展部分的所述图像的尺寸的比较,缩放(506)从所述存储器检索的所述视频对象的所述扩展部分的所述图像的尺寸和/或在所述3D视频流中捕获的所述视频对象的尺寸;以及

将所述视频对象的所述扩展部分的所述图像与在所述3D视频流中捕获的所述视频对象组合(508),以生成被输出(318)到所述透视显示器(234)以用于显示的组合视频对象。

23. 根据权利要求13至22中任一项所述的方法,其中,所述视频对象是在所述3D视频流中捕获的场景的多个分量中的一个,以及

所述调整(316)在所述3D视频流中捕获的所述视频对象的所述姿态包括从所述3D视频流提取所述视频对象而没有所述场景的其它分量;以及

将所述视频对象所述输出(318)到所述透视显示器(234)以用于显示包括输出具有所调整的姿态的所提取的视频对象。

24. 一种计算机程序产品,包括存储指令的非暂时性计算机可读介质,所述指令能够由增强现实AR计算服务器的处理器执行以执行包括以下的操作:

识别(310)在会议会话期间在从用户装置接收的三维3D视频流中捕获的视频对象;

确定(312)在所述3D视频流中捕获的所述视频对象的姿态;

从AR显示装置获得(314)指示所述视频对象相对于通过所述AR显示装置的透视显示器能够观看的物理对象要如何被设定姿态的AR上下文信息;

基于所述AR上下文信息调整(316)在所述3D视频流中捕获的所述视频对象的姿态;以及

将所述视频对象输出(318)到所述透视显示器以用于显示。

25. 根据权利要求24所述的计算机程序产品,进一步包括执行权利要求14至23中任一项所述的方法。

26. 一种增强现实AR计算服务器(200),配置成:

识别(310)在会议会话期间在从用户装置接收的三维3D视频流中捕获的视频对象;

确定(312)在所述3D视频流中捕获的所述视频对象的姿态;

从AR显示装置获得(314)指示所述视频对象相对于通过所述AR显示装置的透视显示器能够观看的物理对象要如何被设定姿态的AR上下文信息;

基于所述AR上下文信息调整(316)在所述3D视频流中捕获的所述视频对象的姿态;以及

将所述视频对象输出(318)到所述透视显示器以用于显示。

27. 根据权利要求26所述的AR计算服务器(200),进一步配置成执行权利要求14至23中任一项所述的方法。

基于来自增强现实显示装置的增强现实上下文信息调整来自用户装置的3D视频流中的视频对象的姿态

技术领域

[0001] 本公开涉及渲染增强现实 (AR) 环境与关联的AR计算服务器 (诸如网络服务器) 和AR显示装置,以及用于通过AR显示装置来显示视频对象的相关操作。

背景技术

[0002] 已经开发了为在线举行会议提供VR环境的沉浸式虚拟现实 (VR) 环境,其中计算机生成的化身 (avatar) 表示会议中的人类参与者的位置。为在线举行会议提供VR环境的示例软件产品包括MeetinVR、Glue、FrameVR、Engage、BigScreen VR、Mozilla Hubs、AltSpace、Rec Room、Spatial和Immersed。可以向参与者显示VR环境的示例用户装置包括Oculus Quest VR头戴式设备、Oculus Go VR头戴式设备以及运行各种VR应用的个人计算机和智能电话。

[0003] 与其中人类参与者仅看到计算机生成的图形渲染的VR环境相反,使用增强现实 (AR) 环境的人类参与者通过例如透视显示屏看到覆盖在物理现实世界的视图上的计算机生成的图形渲染的组合。AR环境也被称为混合现实环境,因为参与者看到混合的物理和数字渲染世界。可以显示AR环境的示例用户装置包括Google Glass、Microsoft HoloLens、Vuzix以及运行各种AR应用的个人计算机和智能电话。存在对在AR环境中提供在线举行会议能力的需要。

发明内容

[0004] 本文中公开的一些实施例针对AR计算服务器,所述AR计算服务器包括网络接口、处理器和存储由处理器可执行以执行操作的指令的存储器。网络接口配置成在会议会话期间通过网络从用户装置接收三维 (3D) 视频流。所述操作识别在3D视频流中捕获的视频对象,并且确定在3D视频流中捕获的视频对象的姿态。所述操作从AR显示装置获得指示视频对象相对于通过AR显示装置的透视显示器可观看的物理对象要如何被设定姿态的AR上下文信息,并且基于AR上下文信息调整在3D视频流中捕获的视频对象的姿态。所述操作将视频对象输出到透视显示器以用于显示。

[0005] 在一些另外的实施例中,确定在3D视频流中捕获的视频对象的姿态的操作包括确定在3D视频流中捕获的脸部的特征的姿态。基于AR上下文信息调整在3D视频流中捕获的视频对象的姿态的操作包括基于在3D视频流中捕获的脸部的特征的姿态与脸部的特征相对于通过AR显示装置的透视显示器可观看的物理对象要如何被设定姿态的AR上下文信息指示的比较来旋转和/或平移在3D视频流中捕获的脸部的特征。

[0006] 一些其它相关实施例针对由AR计算服务器进行的对应的方法。所述方法包括识别在会议会话期间在从用户装置接收的3D视频流中捕获的视频对象,以及确定视频对象的姿态。所述方法从AR显示装置获得指示视频对象相对于通过AR显示装置的透视显示器可观看的物理对象要如何被设定姿态的AR上下文信息,以及基于AR上下文信息调整在3D视频流中

捕获的视频对象的姿态。所述方法将视频对象输出到透视显示器以用于显示。

[0007] 一些其它相关实施例针对包括存储指令的非暂时性计算机可读介质的对应的计算机程序产品,所述指令由AR计算服务器的至少一个处理器可执行以执行操作。所述操作识别在会议会话期间在从用户装置接收的3D视频流中捕获的视频对象以及确定视频对象的姿态。所述操作从AR显示装置获得指示视频对象相对于通过AR显示装置的透视显示器可观看的物理对象要如何被设定姿态的AR上下文信息,以及基于AR上下文信息调整在3D视频流中捕获的视频对象的姿态。所述操作将视频对象输出到透视显示器以用于显示。

[0008] 一些其它相关实施例针对对应的AR计算服务器,所述AR计算服务器配置成识别在会议会话期间在从用户装置接收的3D视频流中捕获的视频对象以及确定视频对象的姿态。AR计算服务器进一步配置成从AR显示装置获得指示视频对象相对于通过AR显示装置的透视显示器可观看的物理对象要如何被设定姿态的AR上下文信息,以及基于AR上下文信息调整在3D视频流中捕获的视频对象的姿态。AR计算服务器进一步配置成将视频对象输出到透视显示器以用于显示。

[0009] 这些实施例的一些潜在优点是它们使能人类参与者在会议期间通过AR显示装置的透视显示器来观看视频对象(诸如另一参与者),所述视频对象被显示具有基于AR上下文信息确定的姿态。AR计算服务器可以使用AR上下文信息的各种特性来确定如何对视频对象的图像设定姿态和缩放视频对象的图像,诸如在房间内在哪里对其它参与者的视频图像设定姿态。

[0010] 在检阅以下附图和详细描述时,根据实施例的其它AR计算服务器、方法和计算机程序产品对于本领域技术人员将是或者变得显而易见。意图是全部这样的附加AR计算服务器、方法和计算机程序产品被包括在本描述中并通过所附权利要求保护。

附图说明

[0011] 本公开的方面通过示例来说明,并且不被附图限制。在附图中:

[0012] 图1A-1C说明了根据本公开的一些实施例的一系列情景,其中在视频会议期间在线会议的本地参与者正通过AR显示装置观看远程参与者,并且其中远程参与者的视频图像基于AR上下文信息被设定姿态;

[0013] 图2说明了根据本公开的一些实施例的包括用户装置的AR系统,所述用户装置向AR计算服务器提供用户的3D视频流,所述AR计算服务器对远程参与者的图像设定姿态以用于通过AR显示装置来显示;

[0014] 图3说明了根据本公开的一些实施例的由用户装置、AR计算服务器和AR显示装置执行的操作的流程图和组合数据流图;以及

[0015] 图4和图5说明了根据本公开的一些实施例的可以由图2和图3的AR计算服务器执行的操作的流程图。

具体实施方式

[0016] 现在将在下文中参考附图更充分地描述发明概念,在所述附图中示出了发明概念的实施例的示例。然而,发明概念可以以许多不同的形式来体现,并且不应该被解释为限于本文中阐述的实施例。而是,提供这些实施例使得这个公开将是透彻和完整的,并且将向本

领域技术人员充分传达各种本发明概念的范围。还应该注意到,这些实施例不是相互排斥的。来自一个实施例的组件可以不言而喻地被假设成在另一实施例中存在/使用。

[0017] 本公开的实施例针对在AR环境中提供在线举行会议能力。AR环境可以使能会议中的本地参与者视觉上体验远程参与者的沉浸式存在,所述远程参与者的视频图像相对于本地参与者通过AR显示装置(例如,由本地参与者佩戴的AR眼镜)的透视显示器观看的现实世界物理对象被设定姿态。

[0018] 图1A-1C说明了根据本公开的一些实施例的一系列情景,其中在线会议的本地参与者正通过AR显示装置观看远程参与者,并且其中远程参与者的视频图像基于AR上下文信息被设定姿态。

[0019] 参考图1A,在线会议会话期间的本地参与者100正佩戴被说明为AR眼镜或其它AR头戴式设备的AR显示装置220,并且观看由AR计算服务器200(图2中示出的)生成的在线会议会话的远程参与者的视频图像110a。远程参与者的视频图像110a通过AR显示装置220被显示具有由AR计算服务器200基于由AR显示装置220获得的AR上下文信息调整的姿态。如下面将进一步详细解释的,AR上下文信息可以指示现实世界物理对象,所述现实世界物理对象被本地参与者100观看,并且远程参与者的视频图像110a相对于所述现实世界物理对象要被设定姿态(例如,旋转、缩放和/或锚定)。在图1A中,AR上下文信息可以指示远程参与者的视频图像110a相对于房间中的床或其它家具要被设定姿态。视频图像110a可以由AR计算服务器200相对于床或其它家具来锚定,使得当本地参与者100的观看变为朝向床或家具旋转时,视频图像110a变为被显示具有叠加在现实世界上的姿态,诸如被设定姿态靠在床或家具上。AR上下文信息可以选择在来自AR显示装置220的相机的视频流中捕获的多个现实世界物理对象中的一个(例如,图1A中的床)。多个现实世界物理对象中的所选择的一个通过AR计算服务器200与远程参与者的视频图像110a关联。然后,远程参与者的视频图像110a利用由AR计算服务器200进行的操作被设定姿态(例如,通过调整远程参与者的头部的显示视频图像110a的位置和角度定向)和缩放尺寸(例如,通过调整远程参与者的头部的显示视频图像110a的尺寸),使得当显示在AR显示装置220的透视显示器上时,远程参与者的头部的视频图像110a对于本地参与者100来说看起来相对于选择的现实世界物理对象自然地被设定姿态,就好像远程参与者物理上在该位置处存在一样。

[0020] 参考图1B,本地参与者100已经移动而更靠近其中远程参与者的视频图像110a被设定姿态的物理对象(例如,床或相邻的座位),并且已经改变了他的朝向远程参与者的视频图像的观看方向。相应地,AR上下文信息由AR显示装置220负责更新,以指示本地参与者100和远程参与者的视频图像之间的距离和相对姿态。由AR计算服务器200进行的操作通过调整姿态(例如,调整所显示的远程参与者的头部的的位置和角度定向)和缩放尺寸(例如,调整所显示的远程参与者的头部的尺寸)来响应更新的上下文信息,使得当通过AR显示装置220的透视显示器显示时,远程参与者的调整的视频图像110b对于本地参与者100来说看起来相对于选择的现实世界物理对象自然地被设定姿态,就好像远程参与者物理上在该位置处存在一样。

[0021] 参考图1C,本地参与者100已移动到不同的房间,并且更新的AR上下文信息指示其中远程参与者要被设定姿态的该房间中的位置。AR上下文信息可以由AR显示装置220例如通过使用运动传感器(诸如加速度计)跟踪其移动和姿态来生成,和/或可以由AR上下文服

务器200诸如通过基于来自AR显示装置220的相机的视频流跟踪AR显示装置220相对于现实世界物理对象的移动来生成。位置可以由本地参与者100指定(诸如通过在通过AR显示装置220观看的同时选择位置),和/或位置可以被编程地选择,诸如将在下面进一步详细解释的。由AR上下文服务器200进行的操作通过调整姿态(例如调整所显示的远程参与者的头部的位置和角度定向)和缩放尺寸(例如调整所显示的远程参与者的头部的尺寸)来响应更新的上下文信息,使得当通过AR显示装置220的透视显示器显示时,远程参与者的调整的视频图像110c对于本地参与者100来说看起来相对于选择的现实世界物理对象自然地设定姿态,就好像远程参与者物理上在该位置处存在一样,所述位置被说明为与厨房中的桌子相邻。

[0022] 图2说明了包括用户装置210的AR系统,所述用户装置210向AR计算服务器200提供用户(诸如图1A-1C中提及的远程参与者)的3D视频流。根据本公开的一些实施例,AR计算服务器200对用户的图像设定姿态以用于通过AR显示装置220显示。图3说明了根据本公开的一些实施例的由用户装置210、AR计算服务器200和AR显示装置220执行的操作的流程图和组合数据流图。

[0023] 参考图2和图3,用户装置210在会议会话期间使用3D相机212来生成3D视频流。用户装置210可以包括但不限于移动电话、膝上型计算机、平板计算机、桌上型计算机、独立网络相机等。3D相机212可以包括但不限于一对立体相机、激光雷达传感器(其使用激光将距离映射到对象上的点并测量反射光返回到接收器的时间)或另一3D相机装置。3D视频流可以包括来自立体相机的一对视频流和/或可以包括来自立体相机或激光雷达传感器的被处理的信息。这样的被处理的信息可以包括点云(例如,表示3D形状或特征的点的集合)、网格(例如,多边形网格、三角形网格、或从点云转换的其它形状的网格)、或颜色和深度信息。

[0024] 3D视频流经由例如无线电接入网络240和网络250(例如,私用网络和/或诸如因特网的公共网络)被提供到AR计算服务器200以用于处理。AR计算服务器200可以通过网络250与用户装置210和AR显示装置220通信的边缘计算服务器、网络计算服务器、云计算服务器等。

[0025] AR计算服务器200包括至少一个处理器电路204(本文中称为“处理器”)、至少一个存储器206(本文中称为“存储器”)和至少一个网络接口202(本文中称为“网络接口”)。尽管网络接口202被说明为与RAN 240通信的无线收发器,但它可以此外或备选地是有线网络接口,例如以太网。处理器204可以包括一个或多个数据处理电路,诸如通用和/或专用处理器(例如,微处理器和/或数字信号处理器),其可以跨网络250并置或分布。处理器204操作地连接到这些各种组件。下面描述为计算机可读介质的存储器206存储由处理器204执行以执行操作的可执行指令208。

[0026] 由AR计算服务器200进行的操作包括识别310在3D视频流中捕获的视频对象,以及确定312在3D视频流中捕获的视频对象的姿态。视频对象的识别及其姿态的确定可以对应于识别3D视频流中的各种类型的现实世界物理对象的存在和姿态。例如,确定操作312可以识别在3D视频流中捕获的远程参与者的脸部和/或身体的特征和/或脸部、身体的姿态,诸如通过识别头部、眼睛、嘴唇、耳朵、颈部、躯干、手臂、手等的姿态。此外或备选地,确定操作312可以识别在3D视频流中捕获的家具对象(诸如图1A-1C中说明的房间中的床、座位、桌

子、地板等)的姿态。

[0027] 由AR计算服务器200进行的操作进一步包括从AR显示装置220获得314AR上下文信息,所述AR上下文信息指示视频对象相对于通过AR显示装置220的透视显示器234可观看的物理对象要如何被设定姿态。操作基于AR上下文信息调整316在3D视频流中捕获的视频对象的姿态,并且将视频对象输出318到AR显示装置220的透视显示器234以用于显示。AR显示装置220配置成在透视显示器234上基于调整的姿态(操作316)确定的位置处渲染322视频对象。

[0028] 在一个实施例中,从AR显示装置220获得的AR上下文信息可以指示例如视频对象(例如,图1A-1C中的远程参与者的视频图像)要通过透视显示器234在其上设定姿态的椅子、桌子、地板等的姿态。AR上下文信息可以指示物理对象的姿态,并且由AR计算服务器200基于AR上下文信息来调整316在3D视频流中捕获的视频对象的姿态的操作可以包括基于视频对象的姿态与物理对象的姿态的比较来调整在3D视频流中捕获的视频对象的姿态。

[0029] 在一个实施例中,由AR显示装置220提供320的AR上下文信息指示AR显示装置220的用户已经指定具有调整的姿态的视频对象要被显示在哪里。例如,用户可以指定房间中的诸如座位、桌子、床、地板等的现实世界物理对象,在所述房间中视频对象要被显示并且相对于现实世界物理对象锚定。参考图1A的说明性示例,用户可以指定其中远程参与者的上身的视频图像110a要被显示和锚定的床旁边的物理椅子。AR显示装置220可以提供来自捕获指定的物理椅子的相机232(例如,2D或3D相机)的视频流。然后,AR计算服务器200操作以基于在来自相机232的视频流中的物理椅子的姿态和/或基于其它AR上下文信息(例如,由用户输入的和/或由AR显示装置220生成的)来调整316在3D视频流中捕获的远程参与者的上身的姿态,使得远程参与者的上身被用户通过透视显示器234观看为虚拟地坐在床旁边的指定物理椅子上。

[0030] 在一个实施例中,可以通过确定来自AR显示装置220的相机232的视频流中的物理对象的姿态来获得AR上下文信息。由AR计算服务器200获得314AR上下文信息的操作可以包括确定AR显示装置220的透视显示器234相对于在来自AR显示装置220的相机232的视频流中捕获的物理对象的姿态。调整316在3D视频流中捕获的视频对象的姿态的操作可以包括基于视频对象的姿态与AR显示装置220的透视显示器234相对于在来自AR显示装置220的相机232的视频流中捕获的物理对象的姿态的比较来调整在3D视频流中捕获的视频对象的姿态。

[0031] 在图2的示例中,AR显示装置220包括至少一个处理器电路224(本文中称为“处理器”)、至少一个存储器226(本文中称为“存储器”)、至少一个网络接口222(本文中称为“网络接口”)和显示装置230。AR显示装置220可以包括相机232,所述相机232配置成输出捕获用户(例如,本地参与者)当前正观看的内容的图像的视频流。尽管网络接口222被说明为与RAN 240通信的无线收发器,但是它可以此外或备选地是有线网络接口,例如以太网。处理器224可以包括一个或多个数据处理电路,诸如通用和/或专用处理器(例如,微处理器和/或数字信号处理器),其可以跨网络250并置或分布。处理器224操作地连接到这些各种组件。下面被描述为计算机可读介质的存储器226存储由处理器224执行以执行操作的可执行指令228。

[0032] 在所说明的示例中,显示装置230是由相对于透视显示屏234定向的可头戴式框架

238可释放地持有的移动电子装置236的一部分。显示装置236被布置成显示投影在透视显示屏234上以用于直接或间接地朝向用户的眼睛反射的信息,即在佩戴框架238的同时。尽管未示出,但是框架238可以包括定位在透视显示屏234和用户的眼睛之间的中间镜子,并且因此光可以直接或间接地朝向用户的眼睛反射。

[0033] 在一些其它实施例中,透视显示器是显示装置230的一部分,其操作以将从AR计算服务器200接收的调整的姿态视频图像叠加在由相机232捕获的现实世界的视频流上。例如,持有移动电子装置236的用户可以通过显示装置230观看来自房间(例如包括图1A中示出的椅子和床)的相机232的视频流。处理器224可以操作以将来自相机232的视频流与具有从AR计算服务器200接收的调整的姿态(操作316)的视频对象(例如,远程参与者的身体的视频图像110a)组合(叠加)。因此,在图1A的上下文中,持有移动电子装置236的用户可以在显示装置230上观看来自房间的相机232的视频流,并且当用户看物理椅子(锚定到视频图像110a)时,远程参与者的身体的视频图像110a被叠加在物理椅子上。因此,本文中所参考的透视显示器可以是例如部分反射屏幕,诸如图2中的显示器234,或者可以是显示装置,在所述显示装置上,由用户装置210的远程相机所捕获的视频对象被叠加在由AR显示装置220的本地相机所捕获的现实世界的视频流上。

[0034] 如本文中所使用的,术语“姿态”指视频对象相对于定义的坐标系(例如,来自3D相机212或用户装置210的视频帧)的定位和/或定向,或者可以相对于另一装置(例如,AR显示装置220)。因此,可以仅基于一个装置相对于另一装置或相对于定义的坐标系的多维定位、仅基于装置相对于另一装置或相对于定义的坐标系的多维定向、或基于多维定位和多维定向的组合来定义姿态。

[0035] 图4说明了根据本公开的一些实施例的可以由图2和图3的AR计算服务器200执行的操作的流程图。

[0036] 参考图3和图4,调整316在3D视频流中捕获的视频对象的姿态的操作可以包括基于在3D视频流中捕获的视频对象的姿态与视频对象相对于通过AR显示装置220的透视显示器234可观看的物理对象要如何被设定姿态的AR上下文信息指示的比较来旋转和/或平移在3D视频流中捕获的视频对象的姿态400(例如,旋转和/或平移图1A中的远程参与者的身体的视频图像110a的位置)。在另外的实施例中,调整316在3D视频流中捕获的视频对象的姿态的操作进一步包括基于在3D视频流中捕获的视频对象的尺寸与通过AR显示装置220的透视显示器234可观看的物理对象的尺寸的AR上下文信息指示的比较来缩放在3D视频流中捕获的视频对象的尺寸402。

[0037] 在另外的示例实施例中,确定312在3D视频流中捕获的视频对象的姿态的操作可以包括确定在3D视频流中捕获的脸部的特征的姿态。在图1A的示例中,可以确定312在3D视频流中捕获的远程参与者的头部、眼睛、耳朵、嘴唇等的姿态。基于AR上下文信息调整316在3D视频流中捕获的视频对象的姿态的操作可以包括基于在3D视频流中捕获的脸部的特征的姿态与脸部的特征相对于通过AR显示装置220的透视显示器234可观看的物理对象要如何被设定姿态的AR上下文信息指示的比较来旋转和/或平移在3D视频流中捕获的脸部的特征。

[0038] 如上面解释的,可以通过确定来自AR显示装置220的相机232的视频流中的物理对象的姿态来获得AR上下文信息。AR计算服务器200可以配置成使用上下文选择规则来自动

地选择在来自AR显示装置220的相机232的视频流中捕获的多个物理对象之中的哪个物理对象。确定312(图3)AR显示装置220的透视显示器234相对于在来自AR显示装置(220)的相机232的视频流中捕获的物理对象的姿态的操作包括识别在来自AR显示装置220的相机232的视频流中捕获的多个物理对象的姿态,并且从多个物理对象之中选择物理对象中的一个,这基于物理对象中的所选择的一个满足上下文选择规则。然后,操作执行相对于物理对象中所选择的一个的姿态的透视显示器234的姿态的确定。

[0039] 解释了上下文选择规则操作的一些说明性非限制性示例。在一个实施例中,由AR计算服务器200进行的操作包括基于物理对象中的一个具有匹配以下中的一个的定义形状的形状来确定在来自AR显示装置220的相机232的视频流中捕获的物理对象中的一个满足上下文选择规则:座位,在所述座位上在3D视频流中捕获的视频对象要在透视显示器234上被显示具有被观看为看起来由座位支撑的姿态;桌子,在所述桌子上在3D视频流中捕获的视频对象要在透视显示器234上被显示具有被观看为看起来由桌子支撑的姿态;以及地板,在所述地板上在3D视频流中捕获的视频对象要在透视显示器234上被显示具有被观看为看起来由地板支撑的姿态。

[0040] 在一些另外的操作实施例中,AR计算服务器200操作以基于由操作AR显示装置220的用户观看的现实世界物理对象结合所显示的具有调整的姿态的视频对象的颜色和/或阴影,调整在来自用户装置210的视频流中的视频对象的颜色和/或阴影。在一个实施例中,由AR计算服务器200进行的操作包括基于在来自AR显示装置220的相机232的视频流中捕获的物理对象的颜色和/或阴影,调整被输出到透视显示器234以用于显示的物理对象的颜色和/或阴影。

[0041] 随着本地参与者围绕房间移动同时通过AR显示装置220观看相对于现实世界物理对象来设定姿态的远程参与者的视频图像,本地参与者的位置与远程参与者的设定姿态的视频图像的虚拟位置之间的相对定位可以导致对姿态进行相当大范围的调整(例如,旋转和平移),以及缩放被观看的远程参与者的身体的尺寸。一些姿态可能导致要通过AR显示装置220观看远程参与者的上躯干和头部,而一些其它姿态可能导致仅头部或头部的一部分被观看。此外,例如由于远程参与者相对于用户装置210的相机212移动,在来自用户装置210的3D视频流中捕获了远程参与者的身体中的多少可能随着时间而改变。为了促进远程参与者的视频图像的缩放和任何期望姿态的生成,AR计算服务器200的一些其它操作实施例将先前存储的较早视频对象的扩展部分(例如,远程参与者的身体的一部分)的图像与当前在3D视频流中捕获的视频对象(例如,远程参与者的头部)组合。扩展部分可以存储在如图2中示出的AR计算服务器200的存储器206中的图像部分储存库209中。例如,这些操作可以将图1A-1C中的远程参与者的身体的较早图像附加到当前在3D视频流中捕获的远程参与者的脸部的图像。

[0042] 图5说明了根据一些实施例的可以由AR计算服务器200执行的对应操作的流程图。参考图5,操作提取500在会议会话期间在较早时间在3D视频流中捕获的或者来自另一会议会话的另一3D视频流的视频对象的扩展部分的图像。可以通过仅将视频对象的扩展部分复制到存储器中(而不复制3D视频流的视频帧中的其它对象、背景等)来提取视频对象的扩展部分。在确定312视频对象的姿态的时间,视频对象的扩展部分不在3D视频流中捕获。视频对象的示例扩展部分可以对应于例如远程参与者的颈部、躯干、手臂等的视频图像。操作将

视频对象的扩展部分的图像存储502在存储器中以供后续使用。视频对象的扩展部分的图像可以被存储502在如图2中示出的AR计算服务器200的图像部分储存库209中。基于在3D视频流中捕获的视频对象的姿态与从存储器(例如,图像部分储存库209)检索的视频对象的扩展部分的图像的姿的比较,操作调整504从存储器(例如,图2中示出的AR计算服务器200的图像部分储存库209)检索的视频对象的扩展部分的图像的姿和/或在3D视频流中捕获的视频对象的姿。基于在3D视频流中捕获的视频对象的尺寸与从存储器检索的视频对象的扩展部分的图像的尺寸的比,操作缩放506从存储器(例如,图像部分储存库209)检索的视频对象的扩展部分的图像的尺和/或在3D视频流中捕获的视频对象的尺。然后,操作将视频对象的扩展部分的图像与在3D视频流中捕获的视频对象组合508,以生成被输出318到AR显示装置220的透视显示器234以用于显示的组合的视频对象。

[0043] AR计算服务器200可以提取在来自用户装置210的3D视频流中捕获的视频对象,以生成提取的视频流,所述提取的视频流被输出到AR显示装置220以用于通过透视显示器234显示。在说明性实施例中,视频对象是由用户装置210的3D相机212在3D视频流中捕获的场景的多个分量中的一个。由AR计算服务器200调整316在3D视频流中捕获的视频对象的姿的操作包括从3D视频流提取视频对象而没有场景的其它分量。由AR计算服务器200将视频对象输出318到透视显示器234以用于显示的操作包括输出具有调整的姿的提取的视频对象。

[0044] 尽管AR计算服务器200在图2中以及在别处被说明为与AR显示装置220分离,但是在一些其它实施例中,AR计算服务器200被实现为AR显示装置220的组件和/或在另一计算装置中实现。例如,本文中描述为由AR计算服务器200执行的操作中的一些可以备选地或此外由AR显示装置220、用户装置210和/或另一计算装置执行。

[0045] 进一步的定义和实施例:

[0046] 在本发明概念的各种实施例的上面描述中,要理解到,本文中使用的术语仅为了描述特定实施例的目的,并且不意图限制本发明概念。除非另有定义,本文中所使用的全部术语(包括技术术语和科学术语)具有与由本发明概念所属领域的普通技术人员通常理解的含义相同的含义。将进一步理解到,诸如通常使用的词典中定义的那些术语的术语应该被解释为具有与它们在本说明书的上下文和相关领域中的含义一致的含义,并且将不会以理想化的或过度正式的意义来解释,除非在本文中明确地如此定义。

[0047] 当元件被称为被“连接”到、被“耦合”到、“响应”于或其变体到另一元件时,它可以被直接连接到、被直接耦合到或直接响应于另一元件或者可以存在中间元件。相反,当元件被称为被“直接连接”到、被“直接耦合”到、“直接响应”于或其变体到另一元件时,不存在中间元件。相同的数字通篇指相同的元件。此外,如在本文中使用的“耦合的”、“连接的”、“响应的”或其变体可以包括无线耦合的、连接的或响应的。如本文中所使用的,单数形式“(a)”、“一个(an)”和“所述(the)”意图也包括复数形式,除非上下文另有明确指示。为了简洁和/或清晰起见,可能没有详细描述众所周知的功能或构造。术语“和/或”包括关联的列示项目中的一个或多个的任何和全部组合。

[0048] 将理解到,尽管术语第一、第二、第三等可以在本文中用来描述各种元件/操作,但是这些元件/操作不应该被这些术语限制。这些术语仅用来将一个元件/操作与另一个元件/操作区分开。因此,一些实施例中的第一元件/操作在其他实施例中可以被称为第二元

件/操作而不会背离本发明概念的教导。在整个说明书中,相同的附图标记或相同的参考标志符表示相同的或类似的元件。

[0049] 如本文中所使用的,术语“包括(comprise)”、“包括(comprising)”、“包括(comprises)”、“包含(include)”、“包含(including)”、“包含(includes)”、“具有(have)”、“具有(has)”、“具有(having)”或其变体是开放式的,并且包括一个或多个陈述的特征、整数、元件、步骤、组件或功能,但是不排除一个或多个其他特征、整数、元件、步骤、组件、功能或其组的存在或添加。此外,如本文中所使用的,从拉丁短语“例如(exempli gratia)”导出的常见缩写“例如(e.g.)”可以被用来介绍或指定先前提到的项目的一个或多个一般示例并且未规定是这样的项目的限制。从拉丁短语“也就是(id est)”导出的常见缩写“也就是(i.e.)”可以被用来指定来自更一般叙述的特定项目。

[0050] 在本文中参考计算机实现的方法、设备(系统和/或装置)和/或计算机程序产品的框图和/或流程图说明描述了示例实施例。当然,可以通过由一个或多个计算机电路执行的计算机程序指令来实现框图和/或流程图说明的框以及框图和/或流程图说明中的框的组合。这些计算机程序指令可以被提供给通用计算机电路、专用计算机电路和/或其他可编程数据处理电路的处理器电路以产生机器,使得经由计算机和/或其他可编程数据处理设备的处理器执行的指令变换和控制晶体管、存储在存储器位置中的值以及这样的电路内的其他硬件组件以实现在框图和/或流程图的一个或多个框中指定的功能/动作并且从而创建用于实现在框图和/或流程图的一个或多个框中指定的功能/动作的部件(功能性)和/或结构。

[0051] 还可以将可以引导计算机或其他可编程数据处理设备以特定方式运行的这些计算机程序指令存储在有形计算机可读介质中,使得存储在计算机可读介质中的指令产生包括实现在框图和/或流程图框的一个或多个框中指定的功能/动作的指令的制品。因此,本发明概念的实施例可以体现在硬件中和/或诸如数字信号处理器的处理器上运行的软件(包括固件、常驻软件、微代码等)中,它们可以被统称为“电路”、“模块”或其变体。

[0052] 还应该注意,在一些备选的实现中,框中注释的功能/动作可以不按流程图中注释的顺序发生。例如,取决于所涉及的功能性/动作,实际上可以基本上同时执行连续示出的两个框或者有时可以以相反的顺序执行所述框。此外,流程图和/或框图的给定框的功能性可以被分成多个框和/或可以至少部分地集成流程图和/或框图的两个或多于两个框的功能性。最后,在没有背离发明概念的范围的情况下,可以在说明的框之间添加/插入其他框,和/或可以省略框/操作。此外,尽管附图中的一些附图包括通信路径上的箭头以示出通信的主要方向,但是要理解,通信可以在与描绘的箭头相反的方向上发生。

[0053] 在基本上没有背离本发明概念的原理的情况下,可以对实施例进行许多变化和修改。所有这样的变化和修改意图是在本文中被包括在本发明概念的范围之内。因此,上面公开的主题要被认为是说明性的,而不是限制性的,并且实施例的所附示例意图覆盖属于本发明概念的精神和范围的所有这样的修改、增强和其他实施例。因此,在由法律允许的最大程度上,要由包括实施例和它们的等效物的以下示例的本公开的最广泛的可允许的解释来确定本发明概念的范围,并且本发明概念的范围不应被前面详细的描述约束或限制。

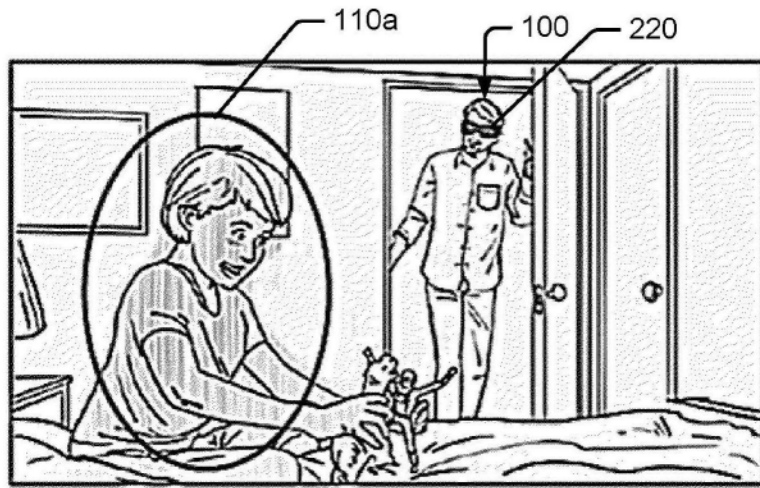


图1A

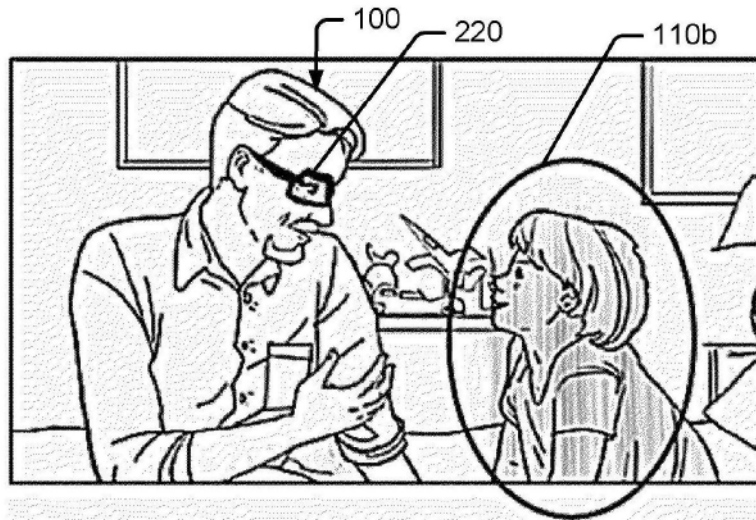


图1B

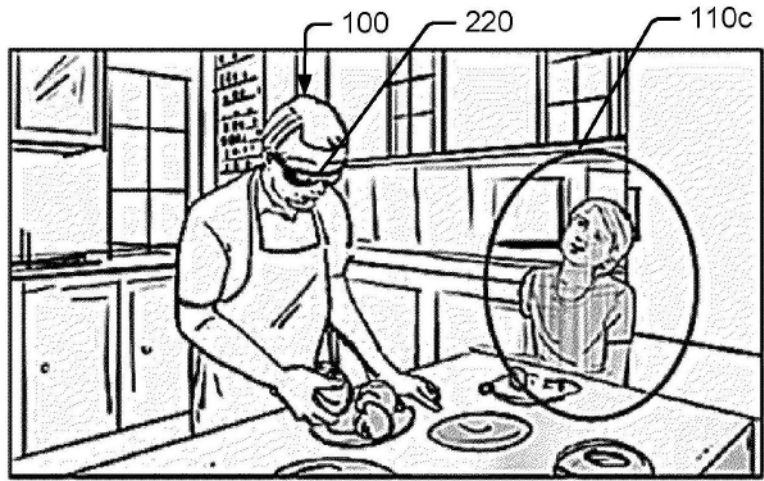


图1C

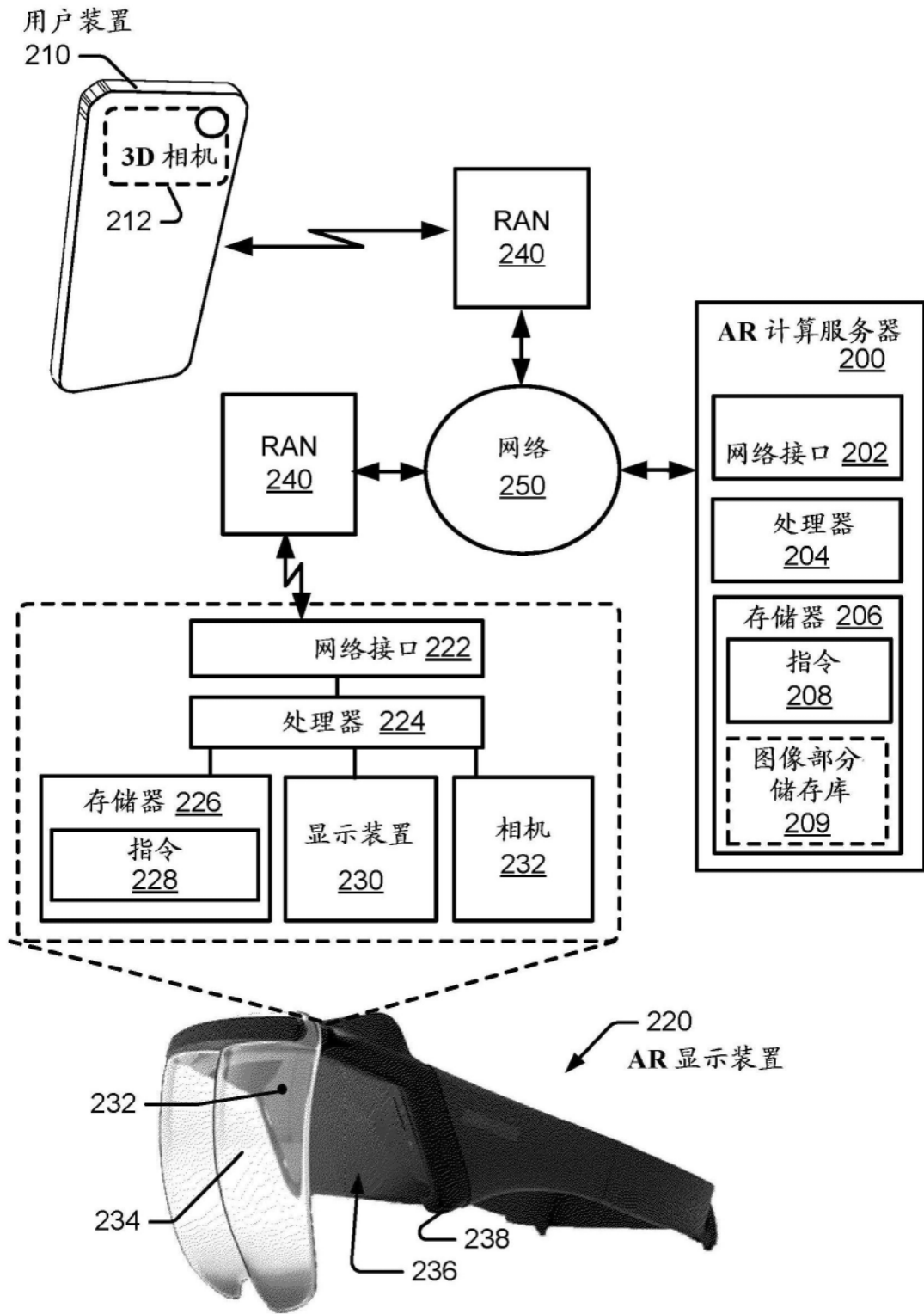


图2

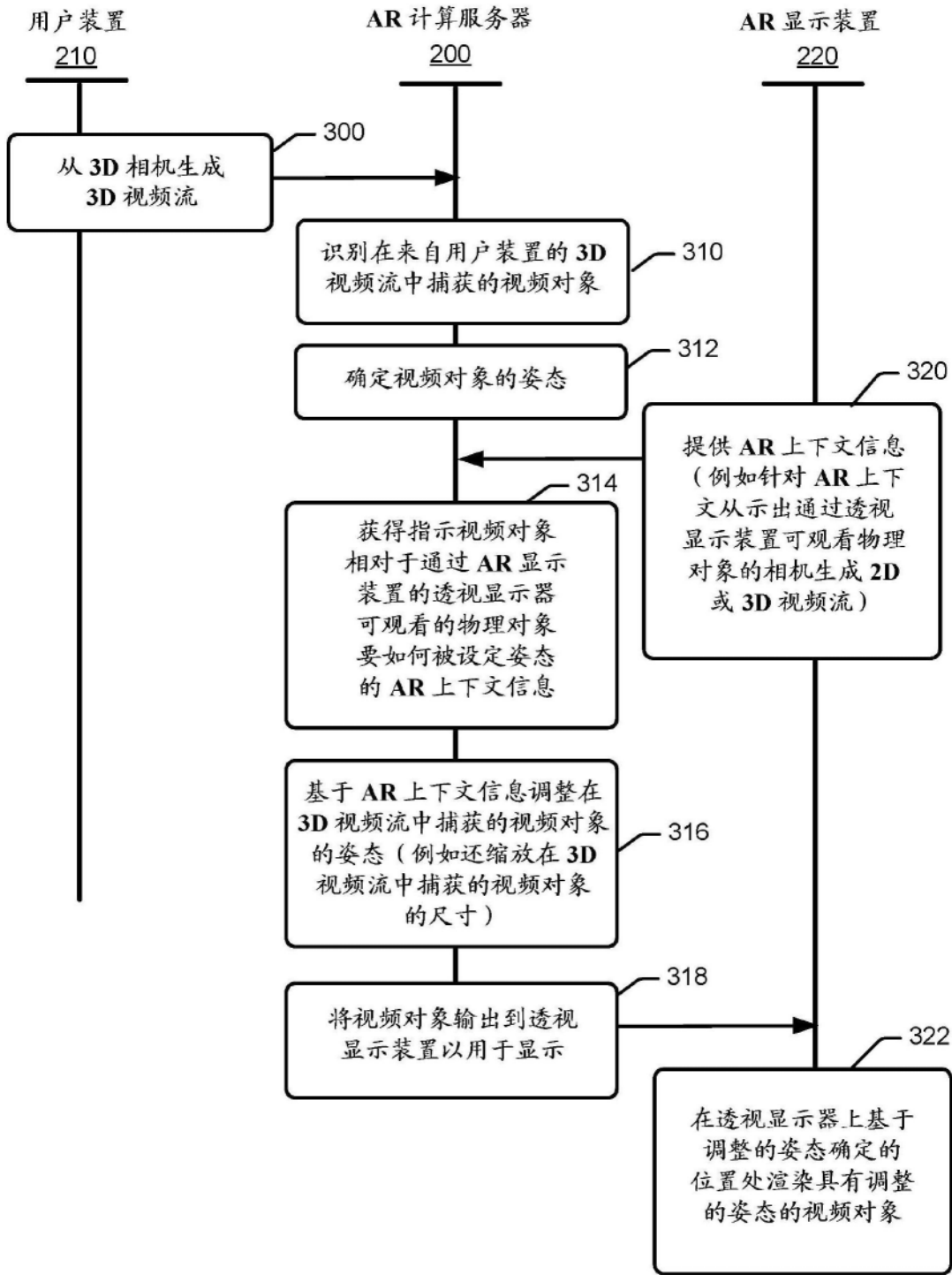


图3

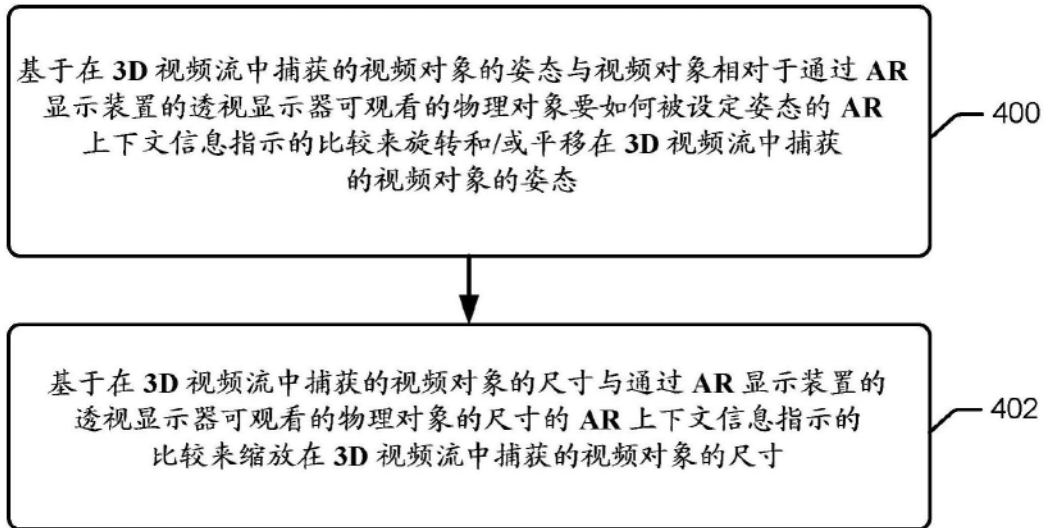


图4

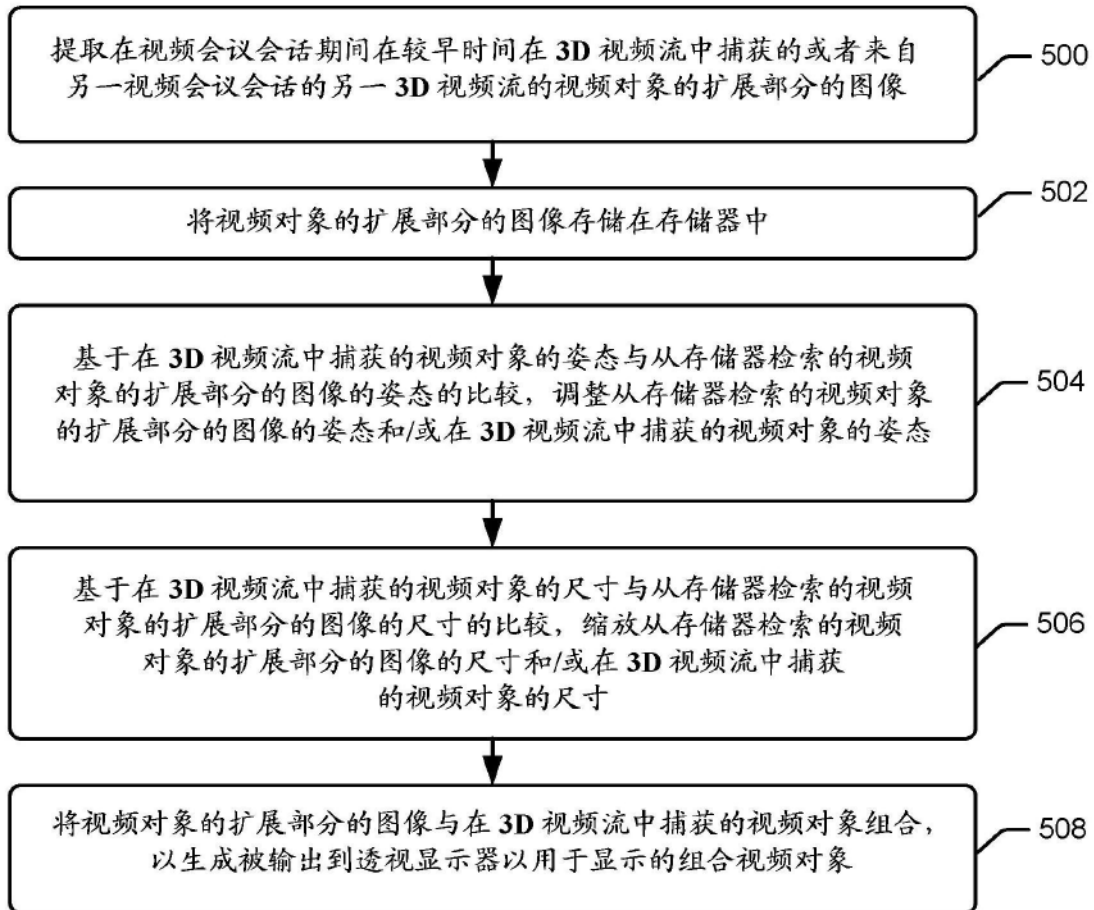


图5