

(12) 发明专利申请

(10) 申请公布号 CN 102253712 A

(43) 申请公布日 2011. 11. 23

(21) 申请号 201110159087. 3

(22) 申请日 2011. 06. 01

(30) 优先权数据

12/792, 549 2010. 06. 02 US

(71) 申请人 微软公司

地址 美国华盛顿州

(72) 发明人 G·J·布班

(74) 专利代理机构 上海专利商标事务所有限公司 31100

代理人 陈斌

(51) Int. Cl.

G06F 3/01 (2006. 01)

G06F 3/033 (2006. 01)

H04L 29/06 (2006. 01)

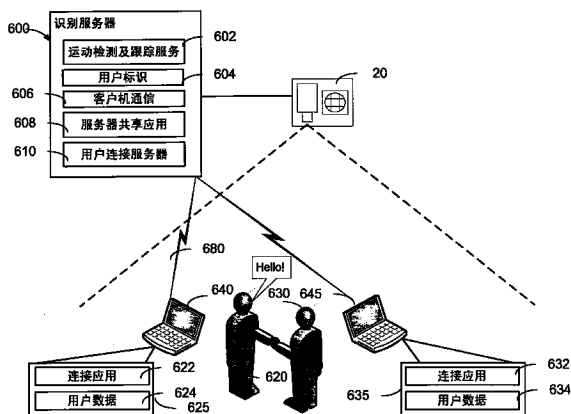
权利要求书 1 页 说明书 19 页 附图 18 页

(54) 发明名称

用于共享信息的识别系统

(57) 摘要

本发明涉及共享信息的识别系统。一种用于在用户之间基于对场景中的用户及其相关联的处理设备的识别来共享信息的系统和方法。交互可以是身体交互、口头交互、或身体和口头姿势的组合。跟踪用户和其他用户的移动并检测他们之间的交互。通过检测捕捉设备的视野内的用户来连接用户处理设备, 捕捉设备检测用户的运动跟踪数据, 如骨架模型。信息共享可由处理设备直接控制、由中间服务器控制、或由处理设备和中间服务器的组合来控制。



1. 一种计算机实现的方法,所述方法包括:  
使用捕捉设备所提供的深度及识别信息来监视场景;  
使用场景中的每一目标的骨架模型来跟踪包括所述场景中的各用户在内的目标的移动;  
将所述场景中的用户与骨架模型和处理设备进行关联;  
通过至少参考所述用户的骨架模型相对于其他用户的移动来检测所述场景中的一个或多个用户之间的交互;以及  
在检测到交互后,在所述场景中的用户之间共享信息。
2. 如权利要求 1 所述的计算机实现的方法,其特征在于,监视是由第一计算机执行的,并且跟踪移动是由与所述用户相关联的处理设备执行的。
3. 如权利要求 1 所述的计算机实现的方法,其特征在于,在用户之间共享信息包括执行预定义的动作。
4. 如权利要求 1 所述的计算机实现的方法,其特征在于,所述交互包括检测第一骨架模型与第二骨架模型之间的关系。
5. 如权利要求 1 所述的计算机实现的方法,其特征在于,进行关联包括:  
从所述处理设备接收对所述场景中的用户的标识的请求;  
响应于所述请求向所述处理设备提供与所述骨架模型相关联的骨架数据和识别信息;  
接收对所述骨架模型的所有权的请求;以及  
将所述处理设备与所述骨架模型进行链接。
6. 如权利要求 3 所述的计算机实现的方法,其特征在于,执行预定义的动作包括在检测到所述交互之后自动地共享用户联系人信息。
7. 如权利要求 1 所述的计算机实现的方法,其特征在于,检测包括检测至少第一用户和第二用户,所述第一用户具有第一骨架模型并且所述第二用户具有第二骨架模型,并且所述方法还包括通过检测所述第一骨架模型与所述第二骨架模型之间的关系来检测所述第一用户与所述第二用户之间的交互。
8. 如权利要求 7 所述的计算机实现的方法,其特征在于,检测所述交互包括检测与所述第一用户和所述第二用户之间的关系相关联的、在时间上与该关系相关的声音。
9. 如权利要求 1 所述的计算机实现的方法,其特征在于,检测包括检测至少第一用户和第二用户,所述第一用户具有第一骨架模型并且所述第二用户具有第二骨架模型,并且所述方法还包括将所述第一骨架模型和所述第二骨架模型的跟踪信息提供给至少所述处理设备。
10. 如权利要求 1 所述的计算机实现的方法,其特征在于,还包括以下步骤:  
检测到检测及跟踪服务器具有包括所述用户中的每一个的场景的视图,所述检测及跟踪服务器创建所述场景中的用户中的每一个用户的骨架模型。

## 用于共享信息的识别系统

### 技术领域

[0001] 本发明涉及用于识别用户并在这些用户之间共享信息的系统和方法。

### 背景技术

[0002] 存在着允许用户共享信息的许多系统。通常,信息共享有点麻烦并且不是基于用户之间的自然交互的。信息共享一般不绑定到物理位置。从一个计算机到另一个计算机的共享数据通常涉及具有数据的用户将计算机以某种方式配置成服务器或将信息提供给共享环境,并且另一用户配置他们的计算机来连接到该服务器或共享环境。这可能是复杂的和/或耗时的过程。用户通常感到简单地用电子邮件发送数据更容易,因为客户机/服务器关系已经建立。

[0003] 类似地,虽然存在用于在设备之间进行邻近共享的至少一个应用,但交换诸如电子邮件地址等联系人信息通常是手动的过程。bump™应用允许移动电话用户在他们将他们的电话“bump(碰)”在一起时共享信息。

### 发明内容

[0004] 概略而言,本发明的技术是用于识别用户并在这些用户之间共享信息的系统和方法。本发明的技术允许具有相关联的处理设备的用户基于用户之间的自然交互来与该系统的视野内的其他用户共享信息。交互可以是身体交互、口头交互、或身体和口头交互的组合。通过检测捕捉设备视野内的用户的用户运动跟踪数据(如骨架模型),来连接用户处理设备。跟踪该用户和其他用户的移动并检测他们之间的交互。在发生交互时,该交互可以触发处理设备之间的信息共享。信息共享可由处理设备直接控制、由中间服务器控制、或由处理设备和中间服务器的组合来控制。

[0005] 提供本发明内容是为了以精简的形式介绍将在以下详细描述中进一步描述的一些概念。本发明内容并不旨在标识出所要求保护的主题的关键特征或必要特征,也不旨在用于帮助确定所要求保护的主题的范围。在本发明内容和以下具体实施方式中描述的特征不是包括一切的,并且具体地,对本领域技术人员而言,结合附图、说明书、以及权利要求书,许多附加特征和优点将变得显而易见。

### 附图说明

[0006] 参考附图来进一步描述根据本说明书的用于限制化身的姿势显示的系统、方法和计算机可读介质,在附图中:

[0007] 图1示出了当用户玩游戏时的目标识别、分析和跟踪系统的示例性实施例。

[0008] 图2示出用于系统识别共享系统和捕捉设备的功能性的计算机实现的体系结构的说明性实施例。

[0009] 图3示出计算环境的示例实施例。

[0010] 图4示出游戏环境的替换示例。

- [0011] 图 5 示出可由一个或多个实施例使用的从正面看的用户的示例性骨架模型。
- [0012] 图 6 是示出识别及共享系统的各组件的交互的框图。
- [0013] 图 7 示出图 6 的框图,示出了识别及共享系统的两个用户的骨架数据。
- [0014] 图 8 示出图 6 的框图,示出了两个用户在识别及共享系统的视野内交互。
- [0015] 图 9 示出在图 8 中交互的用户的骨架模型表示。
- [0016] 图 10 是示出在识别及共享服务器上执行的方法的流程图。
- [0017] 图 11 是示出在识别及共享系统中的用户处理设备上的执行的方法的流程图。
- [0018] 图 12 是示出本发明的在两个处理设备和连接服务器之间的共享示例的 UML 图。
- [0019] 图 13 是示出识别及共享系统中的替换客户机设备交互的框图。
- [0020] 图 14 是示出使用识别及共享系统中的基于客户机的 RGB 捕捉设备的 RGB 数据源的框图。
- [0021] 图 15 是示出与企业环境集成在一起的识别及共享系统的框图。
- [0022] 图 16 是示出诸如移动设备或电话等客户机设备上的连接应用的示例性功能的流程图。
- [0023] 图 17 是示出诸如移动计算机等客户机设备上的连接应用的示例性功能的流程图。
- [0024] 图 18 是示出识别服务器上的连接应用的示例性功能的流程图。

### 具体实施方式

[0025] 呈现了用于识别及共享系统的技术,该系统允许具有处理设备的用户与在该系统视野内的其他用户来基于用户之间的自然交互共享信息。交互可以是身体交互、口头交互、或身体和口头交互的组合。在用户进入该系统的视野时,该用户的处理设备检测该系统并将该用户的运动跟踪数据(如,骨架模型)与该用户进行关联。该系统跟踪该用户对于具有向该系统注册了的处理设备的其他用户的移动,并检测他们之间的交互。在发生交互时,该交互可以触发处理设备之间的信息共享。信息共享可由处理设备直接控制、由中间服务器控制、或由处理设备和中间服务器的组合来控制。

[0026] 图 1 示出用户与目标识别、分析、及跟踪系统 10 的示例实施例的交互,该系统 10 在人类没有附连特殊传感设备的情况下来在他们的自然环境中识别这些对象,在三维空间中唯一地标识他们并跟踪他们。然而,本发明的技术可以适用于其他化身控制机制,如其中用户佩戴传感器的基于传感器的系统或基于设备的化身控制系统,这些设备的一些示例是键盘、鼠标、跟踪球、游戏垫、或操纵杆。

[0027] 根据该示例实施例,目标可以是人类目标(例如,用户 18)、人类目标与一物体、两个或更多人类目标等,这些目标可被扫描来生成诸如骨架模型、网格人类模型、或其任何其他合适的表示等模型。可以跟踪该模型,以使得目标的物理移动或运动可以担当调整和/或控制诸如电子游戏等应用的参数的实时用户界面。此外,该模型可作为模型呈现给应用并实时地传递给它们。例如,所跟踪的用户的运动可用于移动电子角色扮演游戏中的屏幕上的人物或化身。

[0028] 在其中模型是多点骨架模型的一个示例中,目标识别、分析、及跟踪系统 10 通过理解人类的肌肉-骨架系统的自然力学和能力来高效地跟踪人类及其自然移动。示例系统

50 还唯一地识别个人以允许多人经由其肢体的自然移动来与该系统交互。

[0029] 具体而言,图 1 示出其中用户 18 玩拳击游戏的目标识别、分析、及跟踪系统 10 的配置的示例实施例。如下所述,在计算机系统 12 上执行的对通信地耦合的捕捉设备 20 和视听显示单元 16 的其他计算机系统上的软件进行控制或进行交互的软件跟踪用户 18 的移动、分析它们、并将这些移动映射到用户的化身。因此,在该示例中,用户 18 可以移动他的身体来控制他在显示屏 14 上的化身 24,以在拳击游戏中对抗他的对手化身 22。

[0030] 姿势包括运动或姿态。姿势可以由化身或由用户执行。用户可通过移动他的身体创建姿势。例如,可在图像数据中捕捉用户。可以对用户的所标识的姿势分析意义来用作对要执行的应用或动作的控制。例如,用户 18 在图 1A 和 1B 的拳击游戏中挥出一记刺拳。游戏应用在拳击比赛中表示它,确定它是否接触了化身 62,并且如果接触则增加用户的分数。姿势也可以是作为有意义的表情的运动或姿态。例如,它可以表达情绪或想法或意图。姿势可以是静态姿态,如在一个人的躯干前面交叉握住他的前臂,或者它可以是一个或多个移动。此外,姿势可包括一个以上身体部位,如双手鼓掌。

[0031] 例如,目标识别、分析、及跟踪系统 10 可被用来识别并分析用户 18 在捕捉区域(例如,他的起居室中处于图像捕捉系统 60 的视野中的部分)中的挥拳以使得这一挥拳可被解释成姿势,在这种情况下该姿势是他的玩家化身 24 在游戏空间中执行的挥拳的游戏控制。用户 58 的其他姿势也可被解释为其他控制命令或动作,诸如上下快速摆动、闪避、滑步、格挡、刺拳或挥动各种不同力度的重拳等控制。在本申请的上下文中,姿势可包括用户之间的交互。

[0032] 在其他示例实施例中,诸如用户 18 等的人类目标可具有一物体。姿势还可包括道具。在这样的实施例中,电子游戏的用户可能在参与游戏时把持并使用该物体。在本发明的技术的上下文中,目标识别、分析、及跟踪系统可以在相对于用户的处理设备的场景中跟踪用户,如下所述。目标识别及分析技术在 2009 年 5 月 29 日提交的题为“Environment and/or Target Segmentation(环境和/或目标分割)”的美国专利申请 12/475,094 中描述并且其通过引用整体结合于此;在 2009 年 10 月 21 日提交的题为“Pose Tracking Pipeline(姿势跟踪流水线)”的美国专利申请 12/603,437 中描述并且其通过引用整体结合于此;在 2009 年 5 月 29 日提交的题为“Device for Identifying and Tracking Multiple Humans Over Time(用于随时间标识并跟踪多个人类的设备)”的美国专利申请 12/475,308 中描述并且其通过引用整体结合于此;在 2009 年 12 月 18 日提交的题为“Motion Detection Using Depth Images(使用深度图像的运动检测)”中描述并且其通过引用整体结合于此;在 2009 年 10 月 7 日提交的题为“Human Tracking System(人类跟踪系统)”的美国专利申请 12/575,388 中描述并且其通过引用整体结合于此;在 2009 年 4 月 13 日提交的题为“Gesture Recognizer System Architecture(姿势识别系统体系结构)”的美国专利申请 12/422,661 中描述并且其通过引用整体结合于此;以及在 2009 年 7 月 29 日提交的题为“Auto Generating a Visual Representation(自动生成视觉表示)”的美国专利申请 12/511,850 中描述并且其通过引用整体结合于此。

[0033] 图 2 示出用于识别及共享系统 200 的功能性的计算机实现的体系结构的说明性实施例。该识别及共享系统可包括例如捕捉设备 20 和计算系统 212。这样的体系结构系统可被实现成可由在一个或多个处理器和/或计算机硬件上执行的软件来操作的一个或多个

处理模块、或者实现成硬件或固件。

[0034] 捕捉设备 20 可被用于一场景中的目标识别、分析、及跟踪,其中目标可以是用户或物体。根据一示例实施例,捕捉设备 20 可被配置成经由任何合适的技术,包括例如飞行时间、结构化光、立体图像等来捕捉包括深度图像的带有深度信息的视频,该深度图像可包括深度值。根据一实施例,捕捉设备 20 可将所计算的深度信息组织为“Z 层”,即可与从深度相机沿其视线延伸的 Z 轴垂直的层。

[0035] 如图 2 所示,捕捉设备 20 可包括图像相机组件 23。根据一个示例实施例,图像相机组件 23 可以是捕捉一个场景的深度图像的深度相机。深度图像可包括所捕捉的场景的二维 (2-D) 像素区域,其中 2-D 像素区域中的每一像素可表示诸如按照以厘米、毫米等单位的所捕捉的场景中的物体离相机的长度或距离等的深度值。

[0036] 如图 2 所示,根据一个示例实施例,图像相机组件 23 可包括可以被用来捕捉场景的深度图像的 IR 光组件 24、诸如三维 (3-D) 相机 26 等第一传感器、以及诸如 RGB 相机 28 等第二传感器。这些组件中的每一个被聚焦在场景上。例如,在飞行时间分析中,捕捉设备 20 的 IR 光组件 24 可以将红外光发射到场景上,然后,可以使用传感器(未示出),用例如 3-D 相机 26 和 / 或 RGB 相机 28,来检测从场景中的一个或多个目标和对象的表面反向散射的光。在某些实施例中,可以使用脉冲式红外光从而可以测量出射光脉冲和相应的入射光脉冲之间的时间差并将其用于确定从捕捉设备 20 到场景中的目标或物体上的特定位置的物理距离。另外,在其他示例实施例中,可将入射光波的相位与出射光波的相位进行比较来确定相移。然后可以使用该相移来确定从捕捉设备 20 到目标或物体上的特定位置的物理距离。

[0037] 根据另一示例实施例,可使用飞行时间分析,通过经由包括例如快门式光脉冲成像的各种技术来分析反射光束随时间的强度变化以间接地确定从捕捉设备 20 到目标或物体上的特定位置的物理距离。

[0038] 在另一示例实施例中,捕捉设备 20 可使用结构化光来捕捉深度信息。在该分析中,图案化光(即,被显示为诸如网格图案或条纹图案等已知图案的光)可经由例如 IR 光组件 34 被投影到场景上。在撞击到场景中的一个或多个目标或物体的表面时,作为响应,图案可变形。图案的这种变形可由例如 3-D 相机 36 和 / 或 RGB 相机 38 来捕捉,然后可被分析以确定从捕捉设备 20 到目标或物体上的特定位置的物理距离。

[0039] 捕捉设备 20 还可包括话筒 30 或话筒阵列。话筒 30 可包括可接收声音并将其转换成电信号的换能器或传感器。根据一个实施例,麦克风 30 可以被用来减少目标识别、分析,及跟踪系统 10 中的捕捉设备 20 和计算环境 212 之间的反馈。另外,话筒 30 可用于接收也可由用户提供的音频信号,以控制可由计算环境 212 执行的诸如游戏应用、非游戏应用等应用。

[0040] 在一示例实施例中,捕捉设备 20 还可以包括可以与图像相机组件 23 进行可操作的通信的处理器或微控制器 32。处理器 32 可包括可执行指令的标准处理器、专用处理器、微处理器等,这些指令可包括用于接收深度图像的指令、用于确定合适的目标是否可被包括在深度图像中的指令、用于将合适的目标转换成该目标的骨架表示或模型的指令、或任何其他合适的指令。

[0041] 捕捉设备 20 还可包括存储器组件 34,存储器组件 34 可存储可由微控制器 32 执行

的指令、3-D 相机 25 或 RGB 相机 28 所捕捉的图像或图像的帧、或任何其他合适的信息、图像等等。根据一个示例实施例，存储器组件 34 可包括随机存取存储器 (RAM)、只读存储器 (ROM)、缓存、闪存、硬盘或任何其他合适的存储组件。微控制器 32 和存储器可一起被统称为微控制器。

[0042] 如图 2 所示，在一实施例中，存储器组件 34 可以是与图像捕捉组件 23 和处理器 32 进行通信的单独的组件。根据另一实施例，存储器组件 34 可被集成到处理器 32 和 / 或图像捕捉组件 23 中。

[0043] 如图 2 所示，捕捉设备 20 可经由通信链路 40 与计算环境 12 进行通信。通信链路 46 可以是包括例如 USB 连接、火线连接、以太网电缆连接等的有线连接和 / 或诸如无线 802. 11b、802. 11g、802. 11a 或 802. 11n 连接等无线连接。根据一个实施例，计算环境 12 可以向捕捉设备 20 提供时钟，可以使用该时钟来通过通信链路 46 确定何时捕捉，例如，场景。

[0044] 另外，捕捉设备 20 可以通过通信链路 36 向计算环境 12 提供深度信息和由，例如 3-D 相机 26 和 / 或 RGB 相机 28 捕捉到的图像，并可以由捕捉设备 20 所生成的骨架模型。计算环境 12 然后可使用该骨架模型、深度信息和捕捉的图像来例如控制诸如游戏或文字处理程序等应用。

[0045] 捕捉设备 20 能以交互式速率来捕捉数据，从而提高了数据的保真度并允许本发明的技术处理原始深度数据，数字化场景中的对象，提取对象的表面和纹理，以及实时地执行这些技术中的任一种，以使得显示器（例如，56）可以在其显示屏幕（例如，54）上提供该场景的实时描绘。

[0046] 在图 2 的系统实施例中，捕捉设备 20 在通信上耦合 36 到诸如图 3 和 4 中的计算机系统示例等计算环境 212。

[0047] 计算系统 212 包括运动检测及跟踪服务 602，该服务 602 可包括例如深度及图像检测模块 202、运动跟踪模块 204、骨架跟踪器 207、骨架数据 216。运动检测及跟踪服务 602 向一个或多个应用提供运动数据。将理解，应用 250 可以在计算系统 212 上提供，或在诸如以下讨论的替换实施例中提供，在经由网络 203 耦合到计算系统 212 的一个或多个设备 212a 上提供。网络 203 可包括一个或多个公共和 / 或专用网络，包括因特网。

[0048] 深度和图像数据被计算系统 212 经由链路 36 来接收并由深度模块 202 处理。骨架跟踪器使用数据来创建可以与场景中的不同目标或用户相关联的骨架模型数据 216。场景是处于捕捉设备 20 的视野内的区域。骨架跟踪器 207 可以使用该骨架数据来检测目标在场景内的移动，并将该移动信息提供给一个或多个应用 250。客户机通信服务 606 提供到通过在设备之间协商的标准通信协议经由网络 213 耦合的一个或多个设备 212a 的网络连接。

[0049] 模型数据 216 包括场景中的每一目标的模型，这可以是例如骨架模型。运动模块 204 参考该场景的模型数据来处理输入数据。骨架模型可被实现成表示身体部位及其相对于参考点的在各维度的位置和 / 或旋转角度的一个或多个数据结构。模型数据 216 可按照绝对位置来更新或使用位置和旋转的变化来更新。位置和旋转的变化可被表示成向量和角度。

[0050] 在所示例实施例中，化身控制系统 202 从视听数据捕捉系统 20 本地接收运动跟踪数据 205。另外，化身控制系统 202 可以通过因特网 203 或其他网络远程地接收运动跟踪

数据。对于用户,运动跟踪数据可包括图像数据本身或该数据的降采样版本。

[0051] 一些运动和姿态(例如,姿势)在应用 250 的上下文中可被赋予特殊意义。各应用能访问与姿势相关联的过滤器 206。在一个示例中,运动模块 204 可以基于相关联的数据索引(如身体部位索引)来选择一个或多个姿势过滤器 206。例如,在化身控制系统 202 接收到运动跟踪数据集更新并且特定身体部位的运动变化被指示时,运动模块 204 对与这些特定身体部位相关联的姿势过滤器 206 进行索引。

[0052] 姿势过滤器 206 基于参数数据来执行指令,这些参数数据定义用于基于运动跟踪数据 205 确定特定姿势是否已被执行的准则。在一个实施例中,每一姿势过滤器 206 与姿势库 208 中的特定姿势的库模块相链接。与姿势相关联的每一库模块 208 包括用于响应于姿势来执行处理的可执行指令。这一处理通常涉及更新化身的运动或图像来以某种形式反映该姿势。

[0053] 对用户运动的跟踪可以实时地执行,使得用户可以实时地与正在执行的应用进行交互。实时显示指的是姿势的视觉表示的显示,其中该显示与姿势在物理空间中的执行同时或几乎同时显示。例如,系统可以提供对用户进行回应的显示画面的显示画面更新速率可以是 20Hz 的速率或更高,其中无关紧要的处理延迟造成最小的显示延迟或对用户而言根本不可见。因此,实时包括与被自动数据处理所需的时间延迟了的数据的及时性有关的任何无关紧要的延迟。

[0054] 目标识别、分析、及跟踪系统 200 可确定深度图像是否包括人类目标。在一个实施例中,可确定深度图像的所捕捉的场景中的诸如人类目标和非人类目标等每一目标的边。如上所述,深度值中的每一个能以例如厘米、毫米等单位来表示诸如所捕捉的场景中的物体距捕捉设备 20 的长度和距离等深度值。根据一示例实施例,边可以通过将与例如深度图像中的相邻或邻近的像素相关联的各种深度值进行比较来确定。如果正被比较的各种深度值大于预定边容差,则这些像素可定义一条边。在一实施例中,预定边容差可以是例如 100 毫米。如果表示 1000 毫米的深度值的像素与表示 1200 毫米的深度值的相邻像素进行比较,则这些像素可定义目标的一条边,因为像素之间的长度或距离之差大于 100mm 的预定边容差。

[0055] 根据另一实施例,深度图像上的预定点或区域可被进行泛色填充(floodfill)来判定该深度图像是否包括人类目标。例如,可以比较深度图像的所选区域或点中的像素的各种深度值来确定可定义如上所述的目标或对象的边。在一示例实施例中,预定点或区域可均匀地分布在深度图像上。例如,预定点或区域可包括深度图像中心的点或区域、深度图像的左侧边与中心之间的两个点或区域、深度图像的右侧边与中心之间的两个点或区域,等等。

[0056] 可基于所确定的边来对 Z 层的 Z 值进行泛色填充。例如,与所确定的边相关联的像素以及该所确定的边内的区域的像素可互相关联来定义可与模式进行比较的捕捉区域中的目标或对象。

[0057] 根据一示例实施例,经泛色填充的目标(人类和非人类)中的每一个可与一模式进行匹配来确定捕捉区域中的目标是否包括和/或哪一目标包括人类。该模式可包括例如与各种位置或姿态(诸如手臂在两侧的典型站立姿态)中的人类相关联的预定身体模型的机器表示。



[0058] 在一示例实施例中,可隔离人类目标,并且可创建该人类目标的位掩模来扫描一个或多个身体部位。例如,当在深度图像中找到有效的人类目标后,该深度图像中与该人类目标不匹配的背景或区域可被移除。然后可为人类目标生成位掩模,该位掩模可包括该人类目标沿着例如 X、Y 和 Z 轴的值。根据一示例实施例,可从例如头部开始来对人类目标的位掩模扫描各个身体部位,以生成该人类目标的模型。位掩模的顶部可以与头的顶部的位置相关联。在确定了头的顶部之后,可以向下扫描该位掩模来随后确定颈的位置、肩的位置等等。深度图或深度图像数据可被更新以包括一像素与模型中的特定虚拟身体部位相关联的概率。

[0059] 根据一个实施例,在确定身体部位的值后,可以创建可包括诸如与人类目标的位掩模相关联的身体部位的长度、宽度等的度量值的数据结构。在一个实施例中,用于身体部位的数据结构可包括对捕捉系统 60 以某一帧率在各帧中捕捉的多个深度图像求平均的结果。该模型能以某一数量的帧来迭代地调整。根据另一实施例,可调整所确定的身体部位的度量值,如放大、缩小等,使得数据结构中的度量值更接近地对应于典型的人类身体模型。身体模型可包含任何数量的身体部位,其每一个都可以是所建模的目标的对应部分的任何机器可理解表示。

[0060] 在包括两个或更多个身体部位的模型示例中,该模型的每一身体部位可包括一个或多个结构件(即,“骨骼”),且关节位于相邻骨骼的交叉点处。例如,可使用由位掩模确定的度量值来定义骨架模型中的一个或多个关节。该一个或多个关节可用于定义可对应于人类的身体部位的一根或多根骨骼。每一关节可允许一个或多个身体部位相对于一个或多个其他身体部位移动。例如,表示人类目标的模型可包括多个刚性和/或可变形身体部位,其中某些身体部位可表示人类目标的对应的解剖学身体部位。每个身体部位可被表征为定义骨架模型的关节和骨骼的数学向量。可以理解,某些骨骼可对应于人类目标中的解剖学骨骼,和/或某些骨骼可能不具有人类目标中的对应的解剖学骨骼。

[0061] 骨骼和关节可共同构成骨架模型,其可以是另一模型的构成元素。骨架模型可包括对应于每一身体部位的一个或多个骨架构件以及相邻骨架构件之间的关节。示例性的三维骨架模型 80 分别在图 3 中示出。

[0062] 图 3 示出了从前方看的具有关节 j1 到 j33 的骨架模型 80。骨架模型 80 可包括一个或多个关节 j1-j33。根据一示例实施例,关节 j1-j33 中的每一个可使得在这些关节之间定义的一个或多个身体部位能相对于一个或多个其他身体部位移动并彼此独立地移动。例如,关节 j17 与 j19 之间定义的骨骼对应于前臂,该前臂可独立于例如关节 j23 与 j25 之间定义的骨骼——其对应于小腿——移动。

[0063] 随着捕捉系统 20 捕捉到用户在物理空间中移动,所得图像数据可被用来调整骨架模型,使得骨架模型可以准确地表示该用户。根据一示例实施例,该模型可被光栅化成合成深度图像。光栅化允许由数学向量、多边形网格或其他对象描述的模型被转换成按照像素来描述的合成深度图像。捕捉系统所检索的目标的所观察到的图像与模型的光栅化(即,合成)图像之间的差别可用于确定施加于该模型以将身体调整到不同的姿态的力向量。在一个实施例中,可向该模型的一个或多个受力方面施加一个或多个力向量,来将该模型调整成更接近地对应于目标在捕捉区域的物理空间中的姿态的姿态。在捕捉各帧时,该模型可被迭代地调整。取决于所使用的模型的类型,可向模型的关节、身体部位的质心、三

角形的顶点、或任何其他合适的受力方面施加力向量。此外,在某些实施例中,在确定力的方向和 / 或大小时可使用两个或更多不同的计算。

[0064] 作为这些元件所提供的协作的示例,考虑 IR 光组件 24 和 3-D 相机 26 可提供捕捉区域的深度图像,但是在某些情形中,单独的深度图像可能不足以辨别人类目标的位置或移动。在这些情形中,RGB 相机 28 可“接管”或补充来自 3-D 相机的信息以便能够对人类目标的移动或位置作出更完整的识别。例如,RGB 相机可用于识别与一个或多个目标相关联的色彩等等。如果用户正穿着有图案的衬衫而深度相机不能检测该图案,则可使用 RGB 相机来跟踪该图案并提供关于用户正在作出的移动的信息。作为另一示例,如果用户扭曲,则可使用 RGB 相机来补充来自一个或多个其他传感器的信息来确定用户的运动。作为又一个示例,如果用户在诸如墙或第二目标等另一对象旁边,则可使用 RGB 数据来在两个对象之间进行区分。RGB 相机还能够确定用户的精细特征,如面部识别、发色等,这可用于提供附加信息。例如,如果用户向后转,则 RGB 相机可使用发色和 / 或面部特征的缺乏来确定用户没有面对捕捉系统。

[0065] 图像的具有深度值的像素数据被称为深度图像。根据一个实施例,深度图像可包括所捕捉的场景的二维 (2-D) 像素区域,其中该 2-D 像素区域中的每一像素可具有相关联的深度值,如按照以厘米、毫米等为单位的、所捕捉的场景中的物体距某一参考点(如相对于捕捉设备 20 的某一方面)的长度或距离。例如,像素的深度值可以在“Z 层”中表示,这些层是与从深度相机 70 沿其视线延伸的 Z 轴正交的层。这些深度值可被统称为深度图。

[0066] 图 4 示出可用于诸如图 1 中的游戏控制台或图 2 中的计算系统 212 等游戏控制台中的本发明的技术的一个或多个实施例可在其中操作的计算环境的实施例的详细示例。如图 4 所示,多媒体控制台 100 具有含有一级高速缓存 102、二级高速缓存 104 和闪存 ROM(只读存储器)106 的中央处理单元 (CPU) 101。一级高速缓存 102 和二级高速缓存 104 临时存储数据并因此减少存储器访问周期数,由此改进处理速度和吞吐量。CPU 101 可以设置成具有一个以上的核,以及由此具有附加的一级和二级高速缓存 102 和 104。闪存 ROM 106 可存储在当多媒体控制台 100 通电时的引导过程的初始阶段期间加载的可执行代码。

[0067] 图形处理单元 (GPU) 108 和视频编码器 / 视频编解码器 (编码器 / 解码器) 114 形成用于高速、高分辨率图形处理的视频处理流水线。数据经由总线从图形处理单元 108 输送到视频编码器 / 视频编解码器 114。视频处理流水线将数据输出到 A/V(音频 / 视频) 端口 140 以传输到电视机或其他显示器。存储器控制器 110 连接到 GPU 108 以方便处理器访问各种类型的存储器 112, 诸如但不限于 RAM(随机存取存储器)。

[0068] 多媒体控制台 100 包括一般在模块 118 上实现的 I/O 控制器 120、系统管理控制器 122、音频处理单元 123、网络接口控制器 124、第一 USB 主控制器 126、第二 USB 控制器 128 和前面板 I/O 子部件 130。USB 控制器 126 和 128 用作外围控制器 142(1)-142(2)、无线适配器 148、和外置存储器设备 146(例如闪存、外置 CD/DVD ROM 驱动器、可移动介质等)的主机。网络接口 124 和 / 或无线适配器 148 提供对网络(例如,因特网、家庭网络等)的访问并且可以是包括以太网卡、调制解调器、蓝牙模块、电缆调制解调器等各种不同的有线或无线适配器组件中任何一种。

[0069] 提供系统存储器 143 来存储在引导进程期间加载的应用数据。提供媒体驱动器 144, 且其可包括 DVD/CD 驱动器、硬盘驱动器、或其他可移动媒体驱动器等。媒体驱动器 144

可以是对多媒体控制器 100 内置的或外置的。应用数据可经由媒体驱动器 144 访问,以由多媒体控制台 100 执行、回放等。介质驱动器 144 经由诸如串行 ATA 总线或其他高速连接(例如 IEEE 1394)等总线连接到 I/O 控制器 120。

[0070] 系统管理控制器 122 提供涉及确保多媒体控制台 100 的可用性的各种服务功能。音频处理单元 123 和音频编解码器 132 形成具有高保真度和立体声处理的对应的音频处理流水线。音频数据经由通信链路在音频处理单元 123 与音频编解码器 132 之间传输。音频处理流水线将数据输出到 A/V 端口 140 以供外置音频播放器或具有音频能力的设备再现。

[0071] 前板 I/O 子部件 130 支持暴露在多媒体控制台 100 的外表面上的电源按钮 150 和弹出按钮 152 以及任何 LED(发光二极管)或其他指示器的功能。系统供电模块 136 向多媒体控制台 100 的组件供电。风扇 138 冷却多媒体控制台 100 内的电路。

[0072] CPU 101、GPU 108、存储器控制器 110、和多媒体控制台 100 内的各个其他组件经由一条或多条总线互连,包括串行和并行总线、存储器总线、外围总线、和使用各种总线架构中任一种的处理器或局部总线。作为示例,这些架构可以包括外围部件互连(PCI)总线、PCI-Express 总线等。

[0073] 当多媒体控制台 100 通电时,应用数据可从系统存储器 143 加载到存储器 112 和/或高速缓存 102、104 中并在 CPU 101 上执行。应用可呈现在导航到多媒体控制台 100 上可用的不同媒体类型时提供一致的用户体验的图形用户界面。在操作中,媒体驱动器 144 中包含的应用和/或其他媒体可从媒体驱动器 144 启动或播放,以向多媒体控制台 100 提供附加功能。

[0074] 多媒体控制台 100 可通过将该系统简单地连接到电视机或其他显示器而作为独立系统来操作。在该独立模式中,多媒体控制台 100 允许一个或多个用户与该系统交互、看电影、或听音乐。然而,随着通过网络接口 124 或无线适配器 148 可用的宽带连接的集成,多媒体控制台 100 还可作为较大网络社区中的参与者来操作。

[0075] 当多媒体控制台 100 通电时,可以保留设定量的硬件资源以供多媒体控制台操作系统使用。这些资源可包括存储器的保留量(诸如,16MB)、CPU 和 GPU 周期的保留量(诸如,5%)、网络带宽的保留量(诸如,8kbs),等等。因为这些资源是在系统引导时间保留的,所保留的资源对于应用程序视角而言是不存在的。

[0076] 具体地,存储器保留一般足够大,以包含启动内核、并发系统应用和驱动程序。CPU 保留一般是恒定的,使得若所保留的 CPU 用量不被系统应用使用,则空闲线程将消耗任何未使用的周期。

[0077] 对于 GPU 保留,通过使用 GPU 中断来显示由系统应用生成的轻量消息(例如,弹出窗口),以调度代码来将弹出窗口呈现为覆盖图。覆盖图所需的存储器量取决于覆盖区域大小,并且覆盖图一般与屏幕分辨率成比例缩放。在并发系统应用使用完整用户界面的情况下,优选使用独立于应用分辨率的分辨率。定标器可用于设置该分辨率,从而无需改变频率,也就不会引起 TV 重新同步。

[0078] 在多媒体控制台 100 引导且系统资源被保留之后,就执行并发系统应用来提供系统功能。系统功能被封装在一组在上述所保留的系统资源中执行的系统应用中。操作系统内核标识是系统应用线程而非游戏应用线程的线程。系统应用一般被调度为在预定时间并以预定时间间隔在 CPU 101 上运行,以便为应用提供一致的系统资源视图。进行调度是为

了把由在控制台上运行的游戏应用所引起的高速缓存分裂最小化。

[0079] 当并发系统应用需要音频时,则由于时间敏感性而异步调度音频处理给游戏应用。多媒体控制台应用管理器(如下所述)在系统应用活动时控制游戏应用的音频水平(例如,静音、衰减)。

[0080] 输入设备(例如,控制器 142(1)和 142(2))由游戏应用和系统应用共享。输入设备不是所保留的资源,但却在系统应用和游戏应用之间切换以使其各自具有设备的焦点。应用管理器一般控制输入流的切换,而无需知晓游戏应用的知识,并且驱动程序维护有关焦点切换的状态信息。图像捕捉系统 60 可为控制台 100(例如,为其相机系统)定义附加输入设备。

[0081] 图 5 示出了诸如个人计算机等合适的计算系统环境 500 的示例。

[0082] 参考图 5,用于实现本发明技术的一个示例性系统包括计算机 510 形式的通用计算设备。计算机 510 的组件可以包括,但不限于,处理单元 520、系统存储器 530 和将包括系统存储器在内的各种系统组件耦合至处理单元 520 的系统总线 521。系统总线 521 可以是若干类型的总线结构中的任何一种,包括使用各种总线体系结构中的任何一种的存储器总线或存储器控制器、外围总线、以及局部总线。作为示例而非限制,这样的体系结构包括工业标准体系结构 (ISA) 总线、微通道体系结构 (MCA) 总线、增强型 ISA (EISA) 总线、视频电子标准协会 (VESA) 局部总线、以及也称为夹层 (Mezzanine) 总线的外围组件互连 (PCI) 总线。

[0083] 计算机 510 通常包括各种计算机可读介质。计算机可读介质可以是能由计算机 510 访问的任何可用介质,而且包含易失性和非易失性介质、可移动和不可移动介质。作为示例而非限制,计算机可读介质可以包括计算机存储介质和通信介质。计算机存储介质包括以用于存储诸如计算机可读指令、数据结构、程序模块或其他数据等信息的任何方法或技术实现的易失性和非易失性、可移动和不可移动介质。计算机存储介质包括,但不限于, RAM、ROM、EEPROM、闪存或其他存储器技术、CD-ROM、数字多功能盘 (DVD) 或其他光盘存储设备、磁带盒、磁带、磁盘存储设备或其他磁存储设备、或能用于存储所需信息且可以由计算机 510 访问的任何其他介质。通信介质通常以诸如载波或其他传输机制等已调制数据信号来体现计算机可读指令、数据结构、程序模块或其他数据,并包括任意信息传送介质。术语“已调制数据信号”指的是一个或多个特征以在信号中编码信息的方式被设定或更改的信号。作为示例而非限制,通信介质包括有线介质,如有线网络或直接线连接,以及诸如声学、RF、红外及其他无线介质之类的无线介质。上述中任一组合也应包括在计算机可读介质的范围之内。

[0084] 系统存储器 530 包括易失性和 / 或非易失性存储器形式的计算机存储介质,如只读存储器 (ROM) 531 和随机存取存储器 (RAM) 532。基本输入 / 输出系统 533 (BIOS) 包括如在启动时帮助在计算机 510 内的元件之间传输信息的基本例程,它通常储存在 ROM 531 中。RAM 532 通常包含处理单元 520 可以立即访问和 / 或目前正在操作的数据和 / 或程序模块。作为示例而非限制,图 5 示出了操作系统 534、应用程序 535、其他程序模块 536 和程序数据 537。

[0085] 计算机 510 还可以包括其他可移动 / 不可移动、易失性 / 非易失性计算机存储介质。仅作为示例,图 5 示出了从不可移动、非易失性磁介质中读取或向其写入的硬盘驱动器

540,从可移动、非易失性磁盘 552 中读取或向其写入的磁盘驱动器 551,以及从诸如 CD ROM 或其他光学介质等可移动、非易失性光盘 556 中读取或向其写入的光盘驱动器 555。可以在示例性操作环境中使用的其他可移动 / 不可移动、易失性 / 非易失性计算机存储介质包括但不限于,磁带盒、闪存卡、数字多功能盘、数字录像带、固态 RAM、固态 ROM 等等。硬盘驱动器 541 通常通过诸如接口 540 等不可移动存储器接口连接至系统总线 521,并且磁盘驱动器 551 和光盘驱动器 555 通常由诸如接口 550 等可移动存储器接口连接至系统总线 521。

[0086] 上文讨论并在图 5 中示出的驱动器及其相关联的计算机存储介质为计算机 510 提供了对计算机可读指令、数据结构、程序模块和其他数据的存储。例如,在图 5 中,硬盘驱动器 541 被示为存储操作系统 544、应用程序 545、其他程序模块 546 和程序数据 547。注意,这些组件可以与操作系统 534、应用程序 535、其他程序模块 536 和程序数据 537 相同,也可以与它们不同。在此给操作系统 544、应用程序 545、其他程序模块 546、以及程序数据 547 提供了不同的编号,以说明至少它们是不同的副本。用户可以通过输入设备,诸如键盘 562 和定点设备 561——通常被称为鼠标、跟踪球或触摸垫——向计算机 510 输入命令和信息。其他输入设备(未示出)可以包括话筒、操纵杆、游戏手柄、圆盘式卫星天线、扫描仪等等。这些和其他输入设备通常通过耦合至系统总线的用户输入接口 560 连接至处理单元 520,但也可以由其他接口和总线结构,诸如并行端口、游戏端口或通用串行总线(USB)来连接。监视器 591 或其他类型的显示设备也通过接口,诸如视频接口 590,连接至系统总线 521。除监视器以外,计算机也可以包括其他外围输出设备,诸如扬声器 597 和打印机 596,它们可以通过输出外围接口 590 连接。

[0087] 计算机 510 可以使用到诸如远程计算机 580 之类的一个或多个远程计算机的逻辑连接在联网环境中操作。远程计算机 580 可以是个人计算机、服务器、路由器、网络 PC、对等设备或其他常见网络节点,并且通常包括上面相对于计算机 510 所描述的许多或所有元件,但在图 5 中只示出存储器存储设备 581。图 5 中所示的逻辑连接包括局域网(LAN)571 和广域网(WAN)573,但也可以包括其他网络。这样的联网环境在办公室、企业范围计算机网络、内联网和因特网中是常见的。

[0088] 当在 LAN 联网环境中使用时,计算机 510 通过网络接口或适配器 570 连接至 LAN 571。当在 WAN 联网环境中使用时,计算机 510 通常包括调制解调器 572 或用于通过诸如因特网等 WAN 573 建立通信的其他装置。调制解调器 572 可以是内置或外置的,它可以通过用户输入接口 560 或其他适当的机制连接至系统总线 521。在联网环境中,相对于计算机 510 所描述的程序模块或其部分可被存储在远程存储器存储设备中。作为示例而非限制,图 5 示出了远程应用程序 585 驻留在存储器设备 581 上。能够理解,所示的网络连接是示例性的,并且可以使用在计算机之间建立通信链路的其他手段。

[0089] 图 5 的设备只是可在其上实现本发明的技术的一种类型的处理设备。计算系统环境 500 只是合适计算环境的一个示例,而非意在暗示对本技术的使用范围或功能有任何限制。也不应该将计算环境 500 解释为对示例性操作环境 500 中示出的任一组件或其组合有任何依赖性 or 要求。

[0090] 图 6 示出根据本发明的技术的识别及共享系统的第一实施例。图 6 示出耦合到以上参考图 2 描述的捕捉设备 20 的识别服务器 600。识别服务器 600 包括也在上文描述了运动检测及跟踪服务 602、用户标识服务 604、客户机通信服务 606、以及用户连接服务器

610。可任选地,基于服务器的信息共享应用 608 也可在识别服务器 600 上提供。识别服务器 600 在以上参考图 3 和 4 示出的计算环境中的一个或多个中实现。识别服务器 600 使用户 620 和 630 能够容易地连接相关联的客户机设备(在这一情况下,相关联的客户机设备分别是笔记本 640 和 645),并且基于用户与他们的笔记本的交互、彼此的交互、或与识别服务器 600 的交互来执行各种信息共享任务。尽管本讨论关注两用户之间的交互,但信息的交互和共享可发生在任何数量的用户之间。

[0091] 识别服务器 600 提供捕捉设备 20 在给定视野 660 中的单个立体图。具有相关联设备 675 或视野 660 之外的其他用户 670 不能与识别服务器和用户 620、630 进行交互。在一替换实施例中,先前处于视野 660 内并且已经走出该视野的用户可经由识别服务器 600 和该系统所跟踪的其他用户来保留某种交互性,如下所述。

[0092] 如上所述,运动和跟踪服务 602 用于启用对捕捉设备 20 的视野 660 内的一个或多个目标或用户的检测。在用户 630 进入捕捉设备 20 的视野 660 内时,运动检测及跟踪服务 602 检测每一用户 620、630,构造每一用户的骨架模型,并且跟踪用户在一场景中在视野 660 内的移动。如下所述,这一跟踪信息可被提供给用户相关联设备 640、645 上的连接应用 622、632 或提供给基于服务器的共享应用 608。

[0093] 用户连接服务器 610 接收来自客户机设备 640、645 的连接请求。连接请求可由连接应用 622、632 通过轮询由用户连接服务器 610 监视的已知通信端口来生成。客户机通信 606 使网络设备 640、645、以及 675 能够使用任何数量的已知通信协议和技术(包括但不限于 TCP、UDP、以及 apple talk 协议)来经由无线、有线、或其他公共或专用网络与识别服务器进行通信。还可实现用于确定连接应用 622、632 何时可找出识别服务器的其他技术,如与服务器的邻近度或基于经由已知在同一子网上具有服务器 600 的无线网络 680 的连接。

[0094] 一旦连接了用户设备 640、645,用户标识模块 604、连接服务器 610、以及运动检测及跟踪服务 602 确定用户 620 和诸如设备 640 等设备之间的关联。

[0095] 一旦用户 620 进入视野 660,服务 602 就创建并跟踪该用户的骨架模型。骨架模型被唯一地标识并跟踪,但没有相关联的计算设备或用户被链接到该模型。在连接应用 622 被连接到连接服务器 610 时,它将请求与该系统所跟踪的哪一骨架模型涉及该设备的相关联用户有关的信息。在一个实施例中,房间中的各用户的 RGB 数据和骨架模型数据被提供给客户机连接应用,并且客户机应用标识它们相关联的用户。客户机应用随后请求该应用在服务器上的所有权,由此将该处理设备与该用户进行关联。或者,识别服务器可基于由客户机应用提供给该识别服务器的信息来作出关联。

[0096] 一旦确定了用户的骨架模型,并且用户被标识为与特定骨架模型和设备相关联,则用户之间的交互可以启用不同的信息共享功能。信息共享可包括但不限于,文件共享、个人信息交换,以及在一个实施例中包括一个或多个预定义的功能。

[0097] 在图 6 所示的实施例中,每一客户机设备 640 和 645 分别包括连接应用 622、632,并且可任选地在每一设备内包括用户数据 624、634。每一用户 620、630、670 控制他们各自的设备 640、645、675。一旦客户机设备 640 处于服务器 600 的通信范围之内,相应的连接应用 622、632、672 就与客户机通信服务 606 和用户连接服务器 610 进行通信。在一个实施例中,即使用户 620、630、670 在视野 660 之外(如图 6 的用户 670),每一应用 622、632、672 也可以与客户机通信和用户连接服务器进行通信。然而,因为用户 670 在视野 660 之外,所

以运动检测及跟踪服务 602 不能获得并跟踪该用户的骨架跟踪数据一旦用户处于视野内，则用户之间的交互可以启用信息共享功能。以下参考图 10、11 和图 12 中的示例性共享实施例来描述连接应用 622 以及用户连接服务器 610 的特定操作。

[0098] 在图 6 所示的示例中，一旦用户 620 带着客户机设备 640 进入房间，连接应用 622 将经由通信服务 606 来搜索以检测识别服务器 600。用户连接服务器将注意到客户机设备 640 的连接，并且客户机设备 640 将请求与客户机设备 640 相关联的用户（在这种情况下是用户 620）是否已被运动检测及跟踪服务 602 标识。如果是，则连接应用 622 将请求与该用户相关联的骨架模型的所有权，并且用户标识模块 604 将该骨架模型的所有权分配给设备 640 和用户 620。一旦获得了骨架模型的所有权，该骨架模型（以及用户 620）与视野内的并且其设备已经通过用户连接服务器连接的其他用户的交互可以根据本文描述的技术的描述来发生。

[0099] 图 7 示出图 6 中示出的用户 620 和 630 的骨架模型 620A 和 630A 的一个示例。在图 7 中，用户 620 与骨架模型 620A 相关联，并且用户 630 与骨架模型 630A 相关联。用户标识模块 604 将把骨架模型 620A 与设备 640 相关联，并把骨架模型 630A 与设备 645 相关联。将跟踪两个骨架模型在视野内的移动，并且检测他们之间的交互。

[0100] 用户之间的交互可包括骨架模型之间的身体交互，直接与用户连接服务器或其他用户的可听交互，或与骨架模型的身体交互和可听交互的组合。还可以检测特定姿势，并且这些特定姿势包括相对于特定视野 660 中的其他骨架模型的交互。

[0101] 图 8 示出用户 620 与用户 630 之间的示例性交互。在图 8 中，用户 620 和 630 正在握手，并且用户 620 正在用可听问候语向用户 630 问候。在该示例中，可以通过在两个骨架模型紧密靠近并被识别为握手时所做出的姿势，或通过骨架的邻近度（握手或不握手）结合口头指示符“hello(你好)”来识别该交互。任何数量的其他类型的交互可被识别为定义被识别为交互的姿势。基于这一交互，连接服务器可以识别出用户希望共享信息并将用户之间的连接通知给连接应用 622、632（和 / 或可任选地通知给服务器共享应用 608）。在一个实施例中，共享功能由连接应用基于识别服务器所检测的交互来定义。在替换实施例中，提供服务器共享应用 608 并且用户之间的连接在服务器级处理。

[0102] 图 9 是图 8 中示出的交互的骨架模型表示。在两骨架模型彼此具有由给定距离定义的邻近关系时，在模型 620A 的特定跟踪点 j22 紧密靠近模型 630A 上的对面用户的相反手臂的类似点 j21 并且点 j20、j18 以及 j2 在空间上与类似点 j19、j17、j1 相关时，和 / 或在图 9 中示出的模型交互与诸如问候等可听命令、在服务器处指示的特定命令、或可听地检测到的与图 9 中示出的骨架的关系呈时间关系的任何交互相关联时，可检测到两用户之间的握手动作。一旦在两用户 620 和 630 之间检测到这一交互，用户连接服务器可以启用共享服务 608，或将这一信息提供给客户机应用 622、632，从而允许客户机应用 622、632 直接通信。

[0103] 图 10 示出在识别服务器 600 上发生的监视房间以寻找具有相应相关联的处理设备上的连接应用的用户的过程。在步骤 1050，使用捕捉设备 20 扫描房间以寻找目标。在 1052 处，持续扫描直至检测到目标为止。扫描可以由运动检测及跟踪服务 602 来执行。一旦在 1052 处检测到目标，在 1054 处就为每一所标识的目标创建骨架模型并分配标识符。目标及其相关联的骨架模型被存储，直至 1062 处接收到所有权请求为止。一旦创建了骨架

模型,就作出与该骨架模型相关联的图像数据是否可获得的判定。图像数据可以从捕捉设备 620 获得,或如下所述,从诸如与用户相关联的处理设备上的相机等备选源获得,或从具有用户图像信息的其他数据库获得。如果在 1056 处可获得图像数据,则图像数据将与骨架模型相关联并被存储,直至接收到对骨架模型的所有权请求 1062 为止。在 1060,应用继续扫描以在附加目标进入房间时寻找这些目标。当目标离开房间时,如连接服务器的管理员所确定的,关于骨架模型所存储的数据可被释放,或被保留固定的一段时间。

[0104] 如上所述,在识别服务器可用时,客户机连接应用 622、632 可找出这些服务器。在 1061 处,识别服务器等待客户机连接应用进行连接。一旦设备连接了,则在 1062 处,用户连接服务器等待,直至从客户机连接应用接收到所有权请求为止。当在 1062 处接收到请求之后,在 1064 处,在处理设备和特定骨架模型之间作出关联。在一个实施例中,所有权与特定骨架模型之间的关联由客户机应用来执行。在替换实施例中,连接应用可以使用它从先前标识中所保留的识别数据或使用来自包括 RGB 数据的联系人数据库的数据来作出关联。如下所述,用户标识与特定骨架模型之间的关联是通过使用图像比较技术将用户身份的 RGB 数据与该用户的已知 RGB 数据进行关联来执行的。这一比较可以在客户机设备 640、645 或在服务器 600 上执行。在这种情况下,客户机应用将拥有与同该特定处理设备相关联的用户如何显示有关的数据。在步骤 1064,响应于客户机应用的所有权请求来传送 RGB 数据。一旦客户机应用将其用户标识为与特定骨架相关联,它就可以从服务器请求所有权授权,并且服务器可以将该客户机和该客户机的处理设备与特定用户进行关联。虽然在一个实施例中使用 RGB 数据来进行识别,但可以使用其他类型的识别技术,包括例如语音识别,骨架识别,或 RGB、骨架和 / 或语音识别的任何组合。

[0105] 在 1066,对骨架模型持续跟踪它们在整个房间的移动。在 1070,将用户交互通知给共享应用。在一个实施例中,通知共享应用可包括向共享应用提供该应用所拥有的骨架的骨架模型数据并允许该应用确定与另一用户的交互。或者,服务器可以监视交互,并在包括交互的显著姿势发生时通知共享应用。在步骤 1070 处通知连接应用可包括通知客户机连接应用或服务器连接应用 608。以下描述连接应用中的每一个如何基于被通知了交互来操作的各实施例。可任选地,在 1072 处,如果在服务器 600 上启用了共享应用 608,则信息共享功能可根据本文描述的各实施例如下所述地发生。如上所述,这可包括仅仅连接在彼此的机器上具有授权特权的两个不同的用户,提供共享工作空间,在设备之间提供对等连接,或基于在 1068 处检测到的交互的任何数量的变型或动作。或者,步骤 1072 由客户机共享应用 622、632 来执行。

[0106] 图 11 示出在与用户相关联的处理设备进入连接服务器 600 的领域时诸如应用 622 或 632 等连接应用可发生的过程。在 1002 处,客户机应用寻找连接服务器。当在 1004 处检测到连接服务器时,在 1005 处建立连接,并且在 1006 处客户机作出对房间中的不同骨架模型的骨架模型信息和 RGB 数据的请求。如上所述,对骨架模型信息的请求可以由服务器作为图 10 中的步骤 1064 的一部分来提供。在从服务器接收到骨架信息和 RGB 数据之后,在 1008 处,客户机应用标识用户和对应的骨架,并且在 1010 处,从服务器请求骨架模型的所有权。应当注意,在替换实施例中,标识用户和对应的骨架模型的步骤 1008 可以在服务器处而非在客户机处执行。另外,步骤 1008 是通过将在 1006 处由服务器提供的 RGB 数据与客户机上所存储的已知的用户身份 RGB 数据进行匹配来执行的。在步骤 1012,客户机接



收服务器所分配的骨架模型的所有权授权。

[0107] 为了确定标识用户的骨架模型的所有权,客户机应用将接收连接服务器的视野 660 内的一个或多个骨架模型。骨架模型一般包括三维点空间和向量数据,如上所述,这些信息的量与 RGB 图像用来标识用户的数据量相比相对小。如果连接服务器标识出房间中有多个骨架,则服务器可以基于若干决定性特征来提供与所有骨架模型有关的信息和相关联的 RGB 数据或这些数据的子集。例如,如果服务器检测到请求来自最近进入房间的设备并且服务器知道特定骨架最近进入了房间,则服务器可以选择转发它最近创建的骨架模型,因为与已经处于房间内的其他客户机相比,它更可能与作出该请求的服务器相关联。另外,客户机更可能以进入视野的骨架的出现时间相关的方式来请求骨架的所有权。此外,服务器可以阻塞不处于房间中的骨架的连接。

[0108] 应当注意,服务器可能从不同的客户机接收到对相同的骨架模型的相冲突的请求。服务器可以用任何数量的方式来解决这些问题,包括向第一客户机给出请求该骨架模型的优先级或基于哪一骨架模型更可能与特定设备相关联来执行单独的判断。例如,如果服务器知道相冲突的请求来自不同的设备,其中设备之一是笔记本且另一个是处理设备,则服务器可以扫描房间以寻找靠近相冲突的骨架模型的类似笔记本的物体,并将更靠近笔记本的骨架模型的所有权授予笔记本请求设备。

[0109] 在一个实施例中,在 1012 处,一旦接收到所有权授权,服务器应用可以既检测骨架交互又控制客户机设备之间的连接。在使用基于服务器的共享应用的情况下,客户机共享应用只需要执行步骤 1002-1012。

[0110] 在图 11 所示的实施例中,在 1014 处,客户机连接应用将从服务器接收骨架模型数据,并且在 1016 处,该连接应用将作出与特定客户机应用所关联且为其拥有的骨架模型是否参与了与另一骨架模型的交互有关的判定。一旦在 1016 处作出判定,连接应用就可在 1018 处基于交互的性质和上下文来采取任何数量的动作。如上所述,不同类型的交互可造成相应客户机设备之间的不同类型的连接或关联。以下阐明这些不同交互的示例。

[0111] 图 12 是两个客户机共享应用(客户机 A 和客户机 B)与连接服务器 600 之间的一个这样的交互的图示。图 12 示出使用连接服务器 600 作为中介来直接共享信息的两个客户机应用之间的连接。

[0112] 连接服务器 600 将处于以下状态:它在 1202 处监听例如 1202 处的已知 TCP 端口上的客户机连接。客户机 A 在 1204 处且客户机 B 在 1208 处通过例如重复轮询该端口或在它们连接到包括连接服务器的新网络时,将作出连接到该 TCP 端口的尝试。连接服务器对连接尝试进行确认并分别在 1206 和 1212 处连接到每一客户机,客户机 A 在 1210 处且客户机 B 在 1216 处将各自请求它们自己的骨架模型数据。同样,在 1214 和 1218 处,每一客户机将请求可从捕捉设备获得的 RGB 数据。RGB 数据可以是整个场景的视图。骨架数据在 1220 处返回给客户机 A 并在 1222 处返回给客户机 B。在一个实施例中,在 1220 处返回的数据将是在 1222 处返回的数据之前出现的骨架模型的数据,因为在图 12 所示的图示中,对骨架数据的请求 1210 在时间上处于对骨架数据的请求 1216 之前。或者,将返回可从跟踪系统获得的所有骨架模型。同样,该场景的 RGB 数据 1224 被返回给客户机 A 并且 RGB 数据在 1226 处被返回给客户机 B。在 1228 和 1230,客户机 A 将把场景中的各用户的 RGB 图像与用户的已知图像信息进行匹配。随后在 1230 处确定与 RGB 图像中的特定用户相关联的骨架数据,

并且在 1231 处返回对骨架模型的所有权的请求。骨架数据和 RGB 图像表示相同的物理空间。如此,客户机能够将在 RGB 图像中找到的用户与对应的骨架数据进行相关。同样,客户机 B 将在 RGB 图像中搜索它的用户并在 1229 处将确定它的已知用户的对应骨架模型。在 1232,它将发起它自己的对骨架 1 的所有权的请求。一旦两客户机在 1232 和 1234 处发起了对它们各自的骨架的所有权的请求,这两者将发起 RGB 停止命令 1233、1234。这将限制提供给相应客户机中的每一个的 RGB 数据的量。在 1235 和 1236 处,骨架模型数据将被提供给两客户机。(例如,图 11 中的步骤 1016),并且相应客户机两者都监视 (1237,1238) 所拥有的骨架模型与房间中的其他骨架模型之间的有意义的交互。在 1239 处,客户机 A 确定它自己的骨架模型(骨架 0)与客户机 B 所拥有的骨架(骨架模型 1)之间发生了有意义的交互。在 1242 处,客户机 A 将请求骨架模型 1 的信息。同样,客户机 B 将检测到相同的有意义的交互并且在 1244 处请求骨架模型 0 的信息。在 1246 处,骨架模型 0 的信息将被返回给客户机 B,并且在 1248 处,骨架 1 的信息将被返回给客户机 A。在 1250 处,客户机 A 可以直接发起与客户机 B 的连接请求,或在 1252 处,客户机 B 可以向客户机 A 发起连接请求。当在 1254 和 1256 处建立了标准协议连接之后,客户机 A 和客户机 B 之间将存在连接,使得信息传输可直接发生。应用可以允许传输文件、联系人信息、或在应用内定义的或在作出连接之后由用户定义的其他这样的信息。

[0113] 图 13 示出系统的替换实现,其中视野 660 内的客户机设备是笔记本 640 和移动设备 647。移动设备可能具有比笔记本计算机 640 更有限的处理能力和有限的用户数据。在这种情况下,如果移动设备用户 631 之前从未遇到笔记本用户 620,则一旦用户如图 9 所示的那样握手,预定义的动作可包括将联系人信息从电话传送到笔记本。可以发生其他预定义的动作。移动设备 647 将同样具有以参考图 11 描述的方式来操作的连接应用 632A 和用户数据 634A。

[0114] 图 14 示出本发明技术的另一替换方案,其中提供给连接服务器的 RGB 数据是由耦合到用户设备 640 和 655 的捕捉设备 20A、20B 提供的。RGB 数据可以在连接时基于在两设备 640、655 中的任一个上所存储的信息来提供,或如下所述从连接到连接服务器 600 的数据库提供。在又一实施例中,相机 20A 和 20B 可以向运动检测及跟踪服务 602 提供附加骨架模型和附加跟踪信息。如通过参考本文公开的共同待审的申请将理解的,因为相机 20A 和 20B 分别被定位得更靠近用户 620 和 630,所以捕捉设备 20A 和 20B 能以比其视野 660 中的捕捉设备 20 更大的分辨率来检测脸部图案以及其他更精细的姿势,如手势的分辨率。如此,附加姿势和附加交互式运动和声音可以经由捕捉设备 20A 和 20B 来提供给运动检测及跟踪服务。

[0115] 图 15 示出又一实施例,其中用户标识数据可以经由任何数量的多个源之一来提供。用户标识数据可以由与用户连接服务器进行接合的移动设备 640 和 655 中的每一个来提供。作为替换或补充,连接服务器 600 耦合到企业环境 1500 或是其一部分。企业环境可包括授权服务器 1502、用户标识数据 1504、以及用户共享数据 1510。识别服务器 600 可以访问用户 ID 数据以使用作为企业 1500 的授权用户的所有用户的用户 ID 数据中的已知成像信息来标识进入捕捉设备 20 的视野的用户。企业 1500 可包括基于授权用户结构来标识用户的授权及许可服务,从而授予对诸如文件、联系人信息、个人信息、网络共享等用户资源的许可。用户共享数据 1510 包括用户使之可用于在企业内共享的用户信息,并且可结合

共享应用 608 来使用以授权用户 620 和 630 访问这样的数据 1510。

[0116] 图 16 示出可在诸如在移动设备上进行的应用 632A 等连接应用中进行的过程。在 1602 处, 该应用将在 1602 处接收发生了与其骨架模型的交互的通知。可任选地, 在 1604 处, 基于该交互的预定义的动作可以发生。在一个实施例中, 预定义的动作可以是一旦建立了两用户之间的连接则不管所发生的交互的类型而自动发生的动作。在移动电话的情况下, 预定义的动作可以是与另一用户共享个人联系人信息、输出与可用于在电话上与该另一用户共享的目录有关的共享信息、或使存储在电话或其他位置上的公共文件可用的动作。1604 中的预定义的动作可以在用户以任何方式与任何其他骨架进行交互的任何时候发生。

[0117] 作为预定义的动作的替换或补充, 在 1606 处, 该应用可以基于接收到的相关联的骨架数据和 / 或音频数据来将交互进行分类。一旦进行了分类, 交互 1608、1614、1624、1632 将造成该应用进行的动作。在 1606 处检测到的每一交互可包括用于执行对该电话可用的一类信息共享的一个或多个命令。在 1608 处, 如果该数据指示第一交互, 则第一交互可表示与用户共享联系人信息的命令。这可涉及与房间中的另一人握手, 并且检测到相应用户的两骨架模型在握手。或者, 该交互可包括握手与诸如“共享联系人信息”等音频命令。如果在步骤 1608 处检测到该交互, 则在 1610 处建立连接, 并且在 1612 处, 为在图 16 中执行的应用的定义的用户联系人信息将经由该连接转发到第二用户。在 1614 处, 如果检测到第二交互或替换交互, 则该交互可表示共享日历信息或特定日历项。如果在 1614 处检测到第二交互, 则在 1616 处, 可以提示用户指定日历项 (或者, 可以共享整个日历数据), 在 1618 处将建立连接, 并且在 1620 处, 所选择或所定义的信息将被发送到第二用户的设备。或者, 在 1624 处, 第三交互可包括共享文件, 并且如果在 1606 处基于姿势和 / 或音频数据检测到第三交互, 则可以提示用户选择共享的文件或目录, 在 1628 处可建立连接, 并且在 1630 处将文件传送到第二设备。可以向任何数量的不同类型的交互 1632 提供与 1608、1614 或 1624 相类似的任何数量的不同所得动作 1634。

[0118] 图 17 示出用于诸如用于 1630 等在笔记本上执行的应用的类似过程。当在 1702 处接收到通知后, 可任选地, 在 1704 处, 该应用可基于设备类型和设备位置来采取预定义的动作。这样的预定义的动作可以是共享个人联系人信息、共享预定义的文件、共享目录、或共享特定媒体。在一个示例中, 连接服务器可被安装在用户正在其中进行演示的房间中的笔记本上。在该示例中, 在用户进入预定义的位置时, 基于骨架模型和用户的身份, 预定义的动作可以是呈现用户的处理设备存在的文件 (如, 演示), 并将共享的文件投影到更大的演示系统, 如用于供观众观看的投影仪。另外, 在该示例中, 用户的演示可用于与房间中的其他用户共享, 以允许演示的观看者在用户呈现该演示时下载该演示的副本。作为 1704 处的预定义的动作的替换或补充, 在 1706 处对姿势进行分类。以与参考图 16 定义的方式相类似的方式, 在 1708、1714、和 1724 处的不同的交互造成连接应用 630 的不同动作。

[0119] 在 1708 处的第一类型的交互可以是被分类成造成共享用户信息的交互。一旦在 1709 处标识了第二用户, 则在 1710 处建立连接, 并且在 1712 处与第二用户共享用户联系人信息。第二类型的交互 1714 可以指示共享文件的命令。如果确定了交互 2, 则在 1716 处可以提示用户输入要共享的文件或目录项, 在 1718 处建立连接, 并且在 1720 处发送信息。在 1724 处的另一示例交互可造成 1724 处的服务器或企业登录和共享命令。用户凭证可由连

接应用存储,使得用户不必输入用户标识符或口令。用户登录可由连接应用或处理设备内的其他安全存储装置来存储。一旦对企业登录进行了存档,则在 1726 处,可作出对文件选择的提示,以允许用户选择一个或多个文件或目录来进行共享。或者,可以创建自动共享。文件的位置可以处于客户机设备本身上,或处于参考图 15 讨论的共享存储位置上。在 1728 处,可建立到企业服务器的连接,或不进行连接,其中要与其共享信息的第二设备使其自己直接连接到共享存储位置。如果文件需要被传送到共享位置,则在 1730 处可传送文件。

[0120] 图 18 示出处理客户机的连接和动作的连接服务器上的共享服务应用 608 可发生的过程。在基于服务器的共享应用中,交互检测可由服务器来执行。与基于客户机的应用一样,在 1804 处,服务器可以基于骨架的位置和与骨架相关联的设备来采取预定义的动作。这样的动作可包括加载媒体偏好、显示公共文件、以及加载通知偏好。在其中可以使用基于服务器的连接应用的一个示例中,例如,诸如家庭起居室等共享媒体环境。在该示例中,连接应用可被用来确定具有连接客户机的哪些用户位于媒体中心前方,其中连接服务器在控制台上提供并且相机在诸如图 1 所示的显示设备等显示设备上提供。在第一用户带着诸如移动电话上的应用 630A 等客户机应用进入房间时,该系统可以确定用户是例如家庭中的父母并加载该用户的频道偏好、评级偏好及限制、以及该用户专用的其他信息来经由显示设备进行呈现。同样,当诸如孩子等第二用户进入房间时,附加评级系统或评级锁可被加到效果上。即时消息收发和其他通信协议可以通过以上参考图 1 到 4 示出的处理设备来路由。如此,当在 1802 处接收到交互数据时,在 1806 处对动作进行分类。一个这样的交互可以是 1808 处的传入的电话呼叫或传入的即时消息。取决于在 1810 处房间中是否存在其他用户,连接应用 608 可以在 1811 处显示呼叫者的完整标识或在 1812 处只显示部分标识。同样,如上所述,在用户进入和退出特定视野 660 时,在用户退出房间之后的固定时间段期间维持与哪些用户曾在任何特定时间处于特定视野中有关的信息。例如,当在特定时间全家都在捕捉设备 20 前方但各用户退出房间的情况下,房间中的用户的一个交互可以是指示该用户希望在最近 5 或 10 分钟内曾处于房间中的所有其他用户返回。当在 1814 处确定了这一交互时,在 1816 处作出关于哪些用户在先前时间段内曾处于视野 660 中的判定,建立到这些用户的处理设备的连接,并且向这些处理设备发送将这些用户找回视野 660 中的通知。又一实现可以在 1824 处请求登录到企业或家庭服务器。在 1826 处,检查用户对服务器的凭证,并且在 1828 处建立到服务器的连接。在 1832 处,可设置共享许可。任何数量的不同交互 1834 可以基于在 1806 处定义的交互式数据来以此方式进行定义。

[0121] 本文公开的技术可用作其上可利用其他技术的平台。以上示出的连接服务器 600 允许连接到服务器的处理设备和基于它们的移动来标识用户的相关联的骨架数据之间的各种不同的连接方法。交互可被定义成特定姿势、或与音频命令相组合的姿势、或仅音频命令。在本发明技术的一独特方面,信息从单个观点提供给各连接应用。即,捕捉设备 20 给出整个视野 660 的单个立体图。因为使用了单个视野和单个观点,所以交互的移动以及房间中各用户之间的关系可以得到维持。

[0122] 前面的详细描述只是为了说明和描述。并非旨在穷举本技术或将其限于所公开的精确形式。鉴于上述教导,许多修改和变型都是可能的。所描述的实施例只是为了最好地说明本技术的原理以及其实际应用,从而使精通本技术的其他人在各种实施例中最佳地利用本技术,适合于特定用途的各种修改也是可以的。此外,应当注意,在说明书中使用的语

言主要是出于可读性和指导目的来选择的。尽管用专门描述结构特征和 / 或方法动作的语言描述了主题,但是应当理解,在所附权利要求书中限定的主题并不一定局限于上述特定的特征或动作。更确切而言,上述具体特征和动作是作为实现权利要求的示例形式公开的。

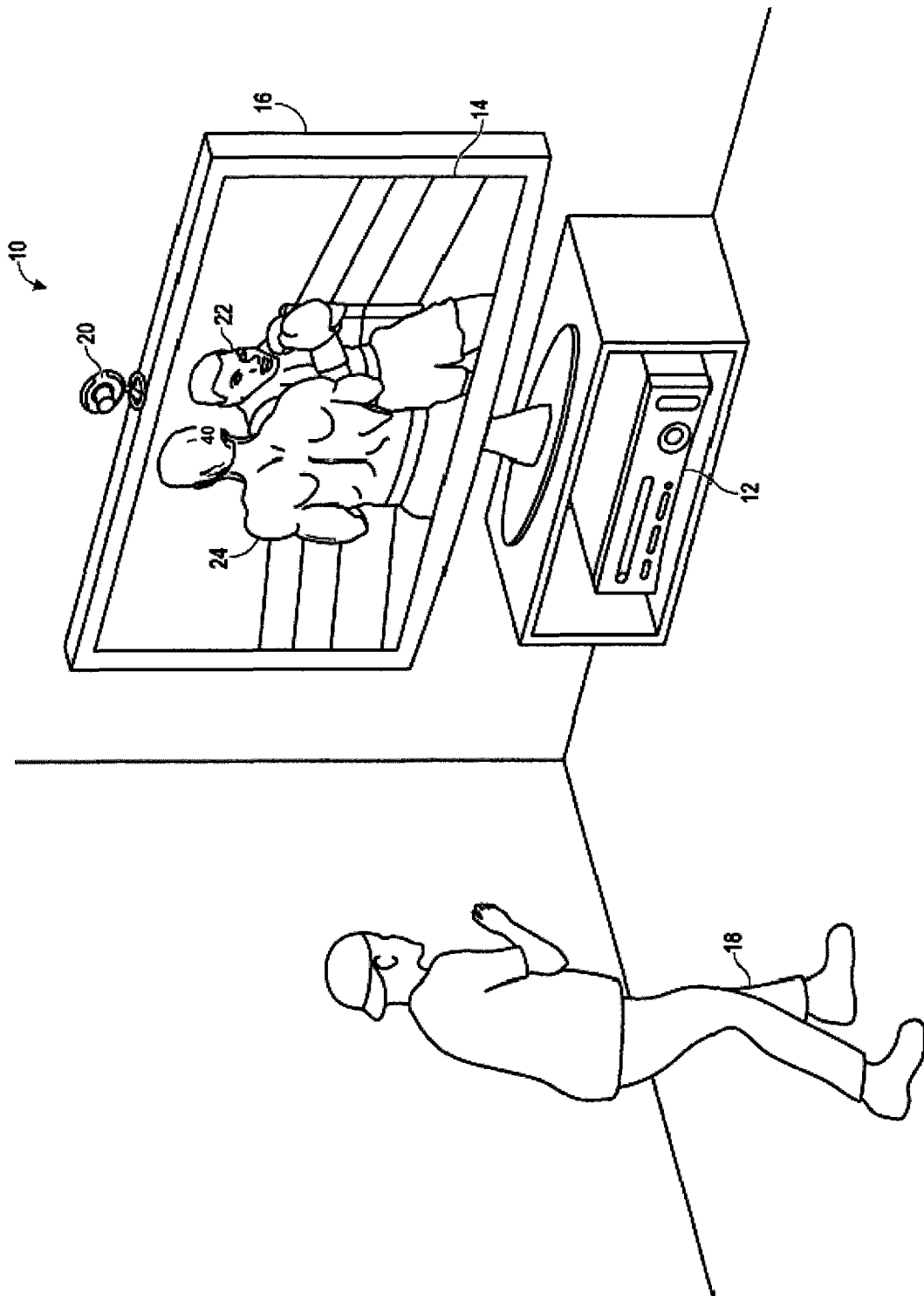


图 1

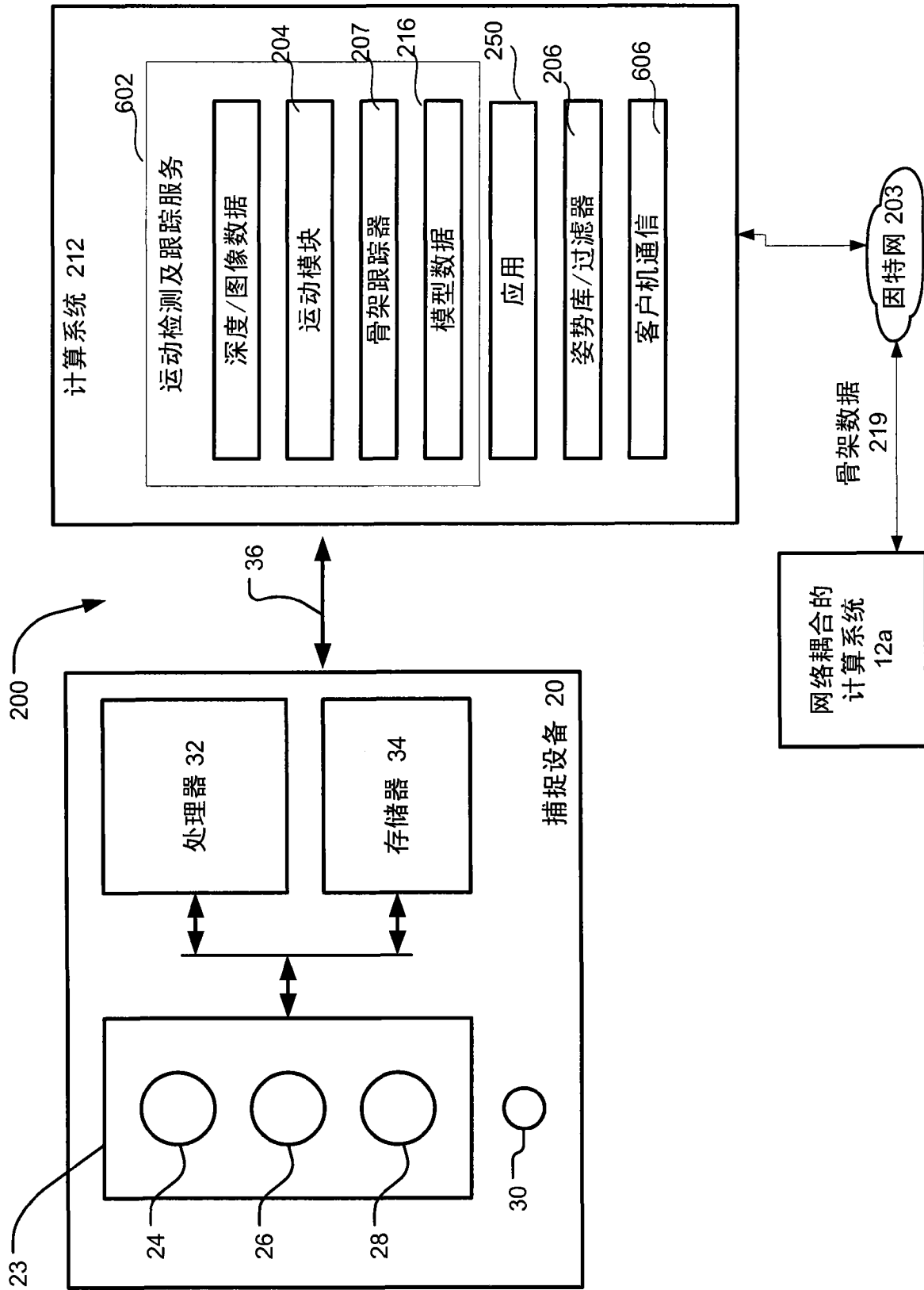


图 2

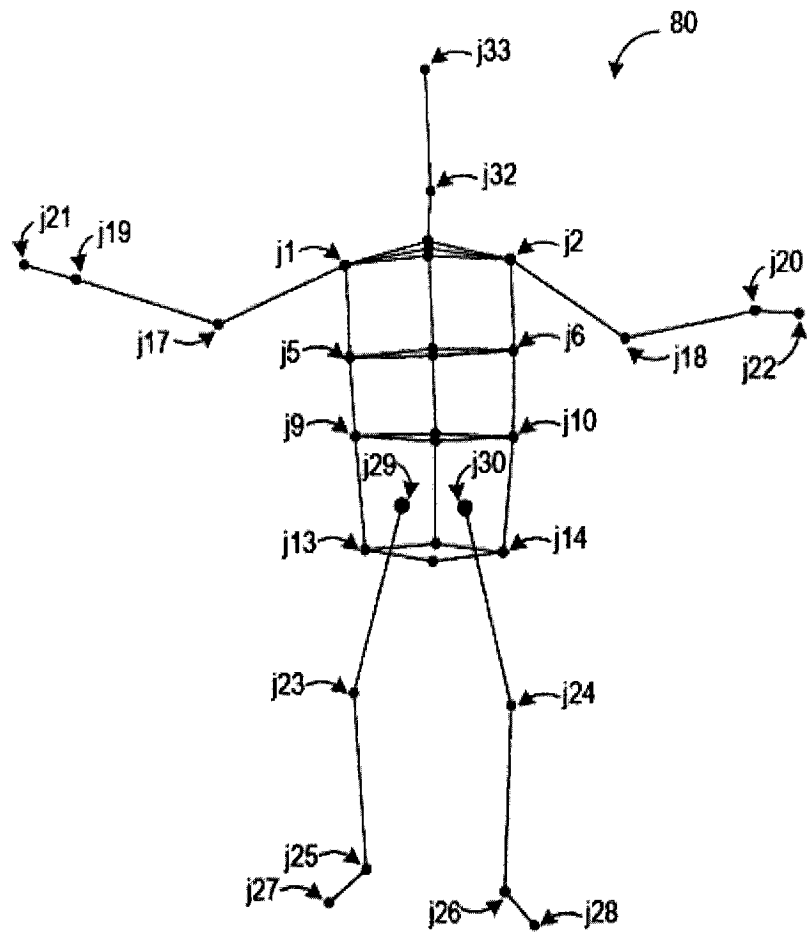


图 3



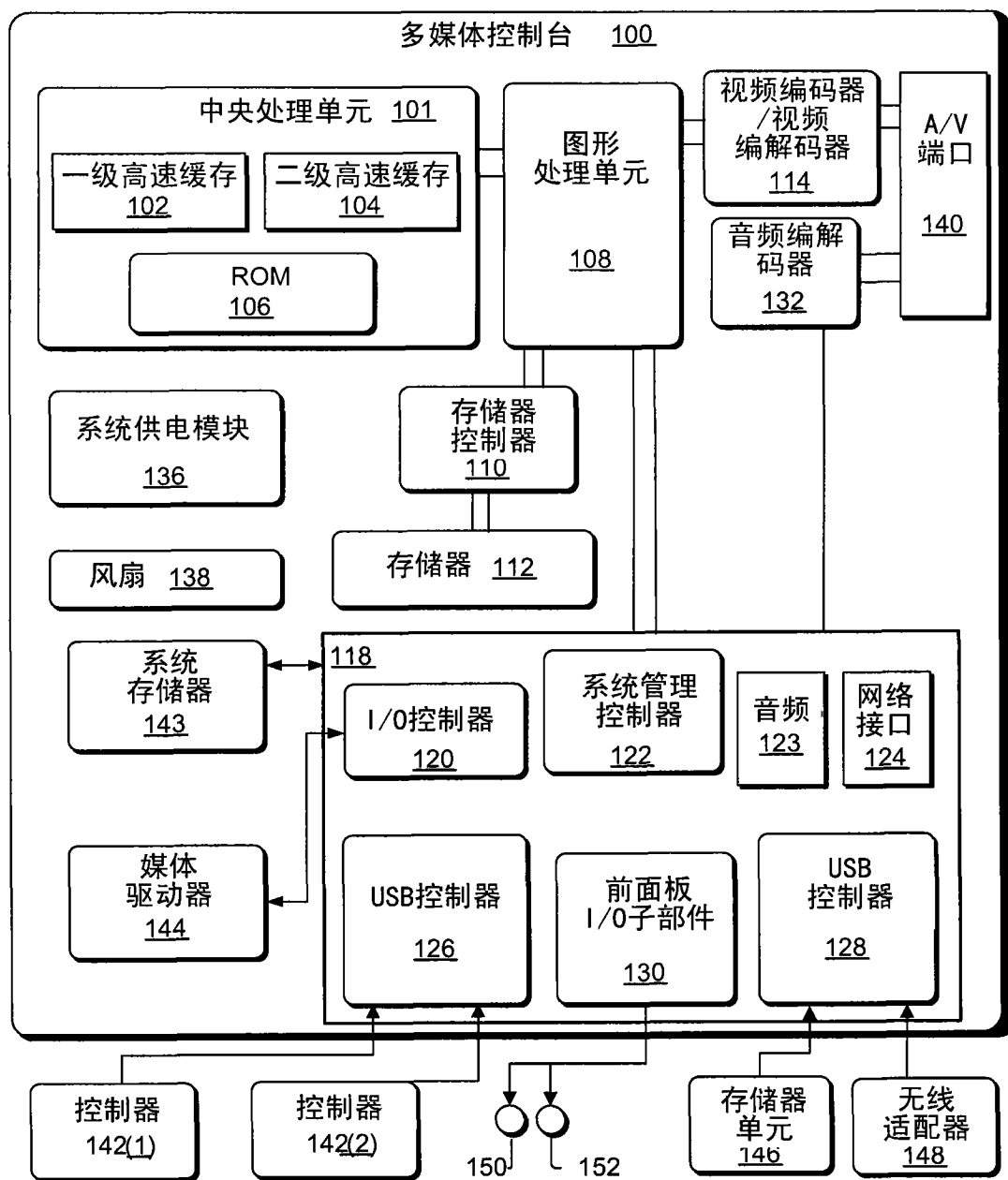


图 4

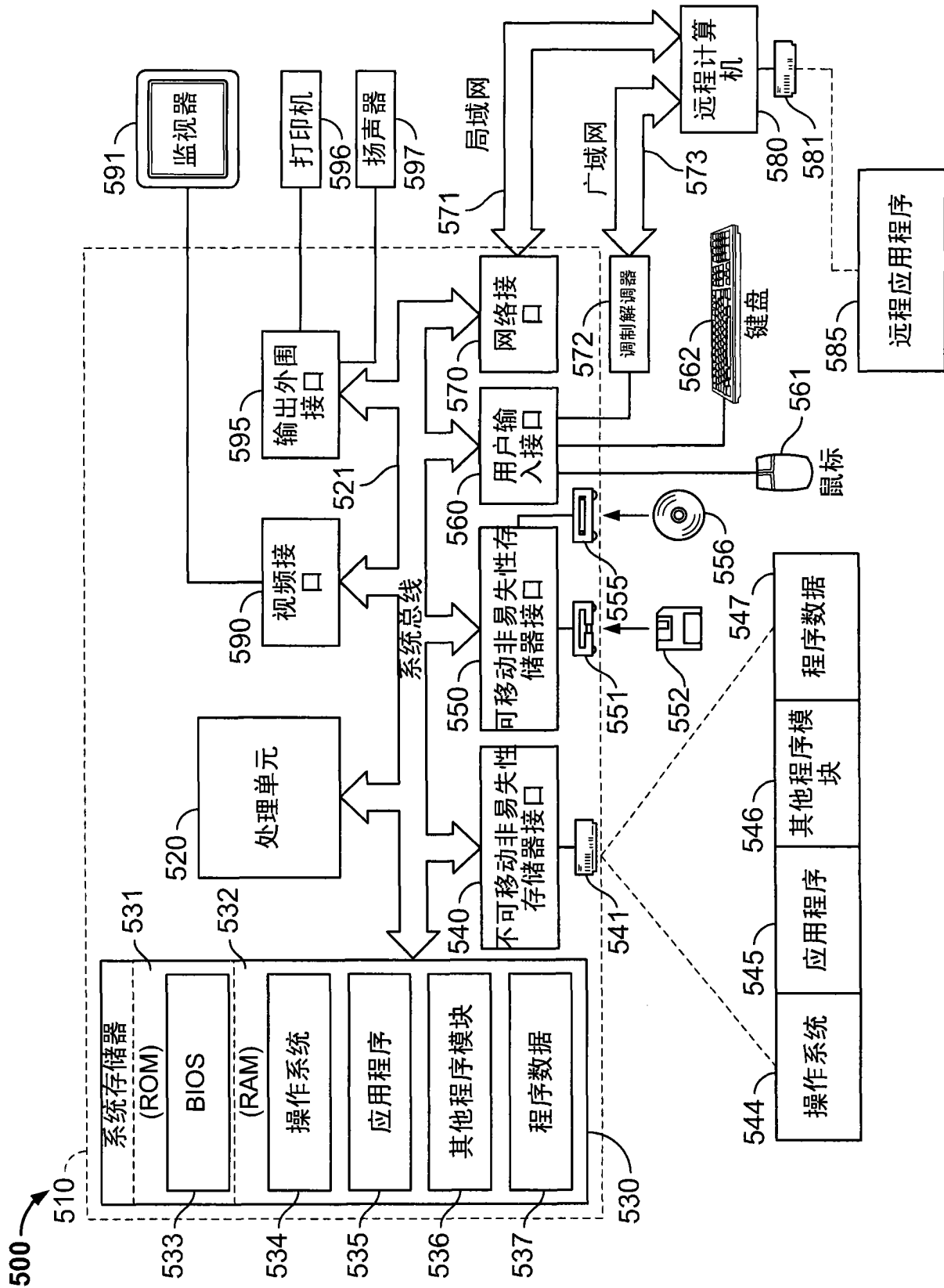


图 5

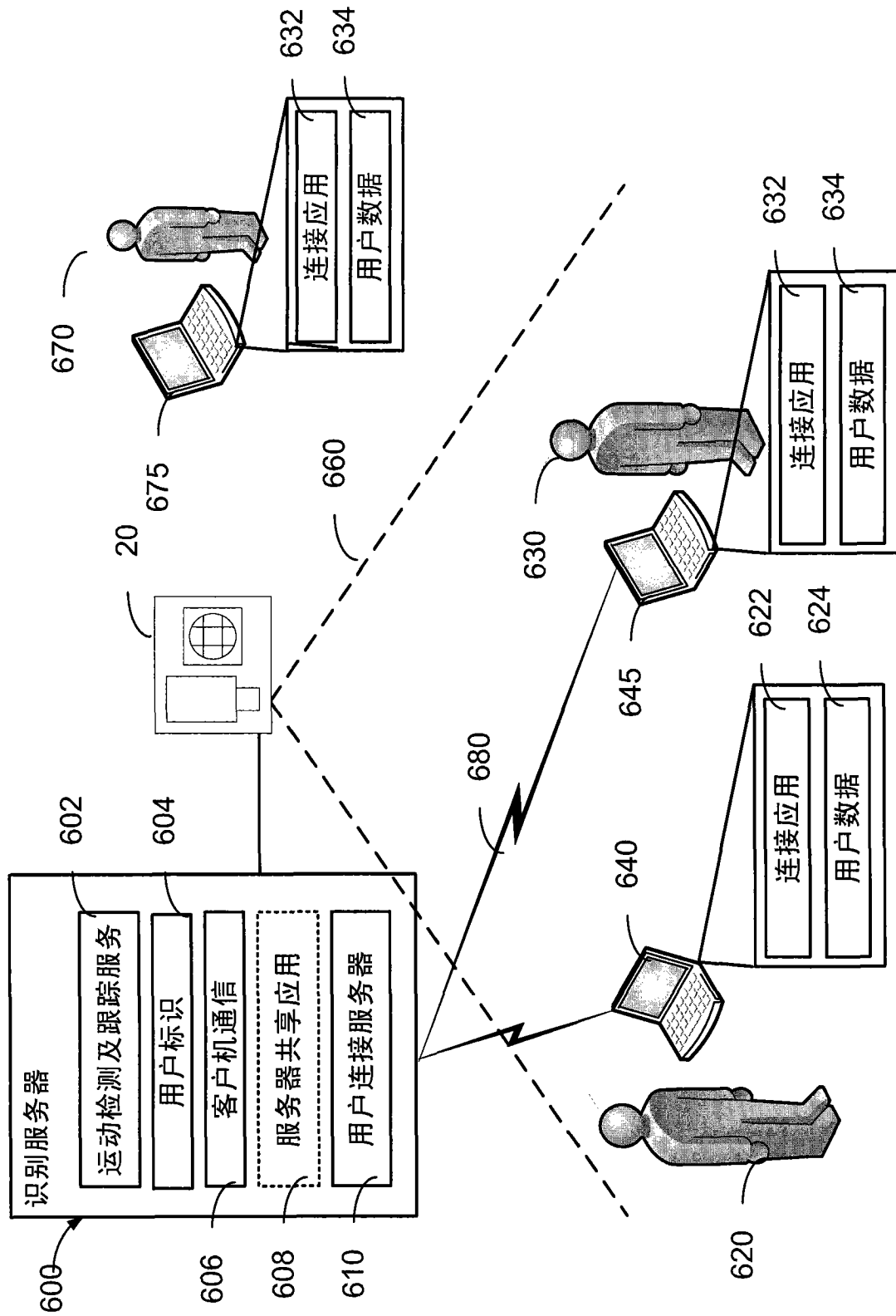


图 6

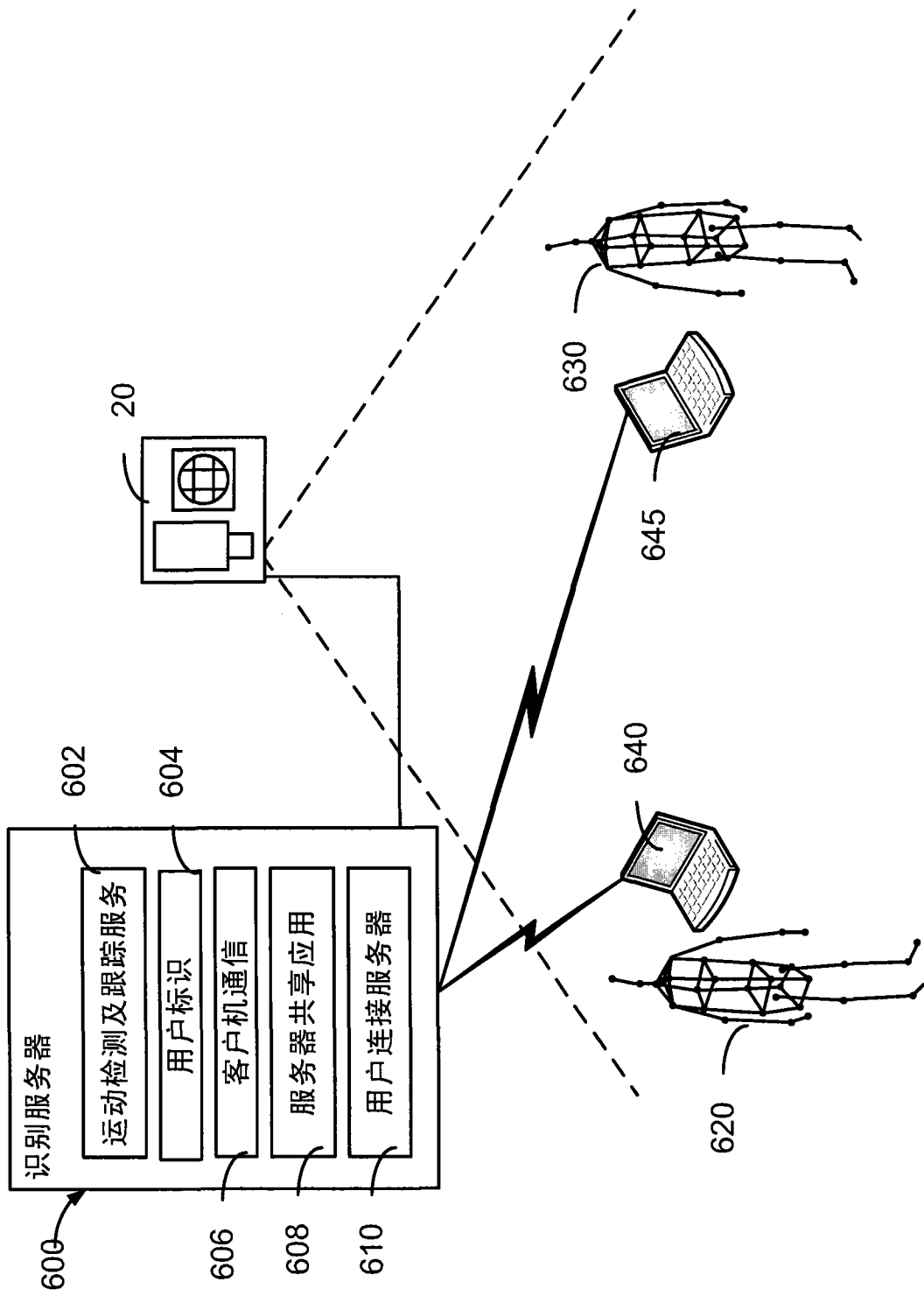


图 7

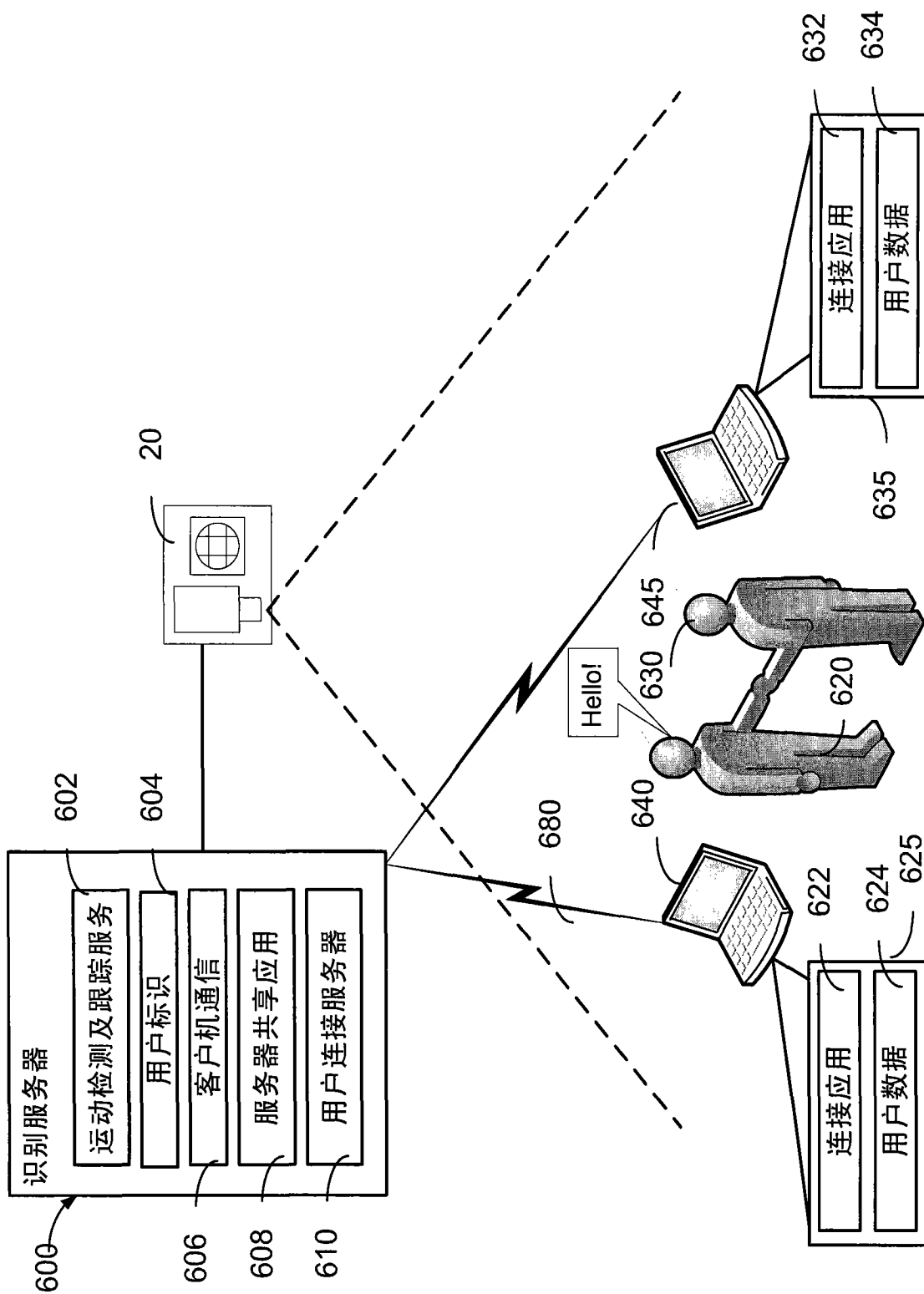


图 8

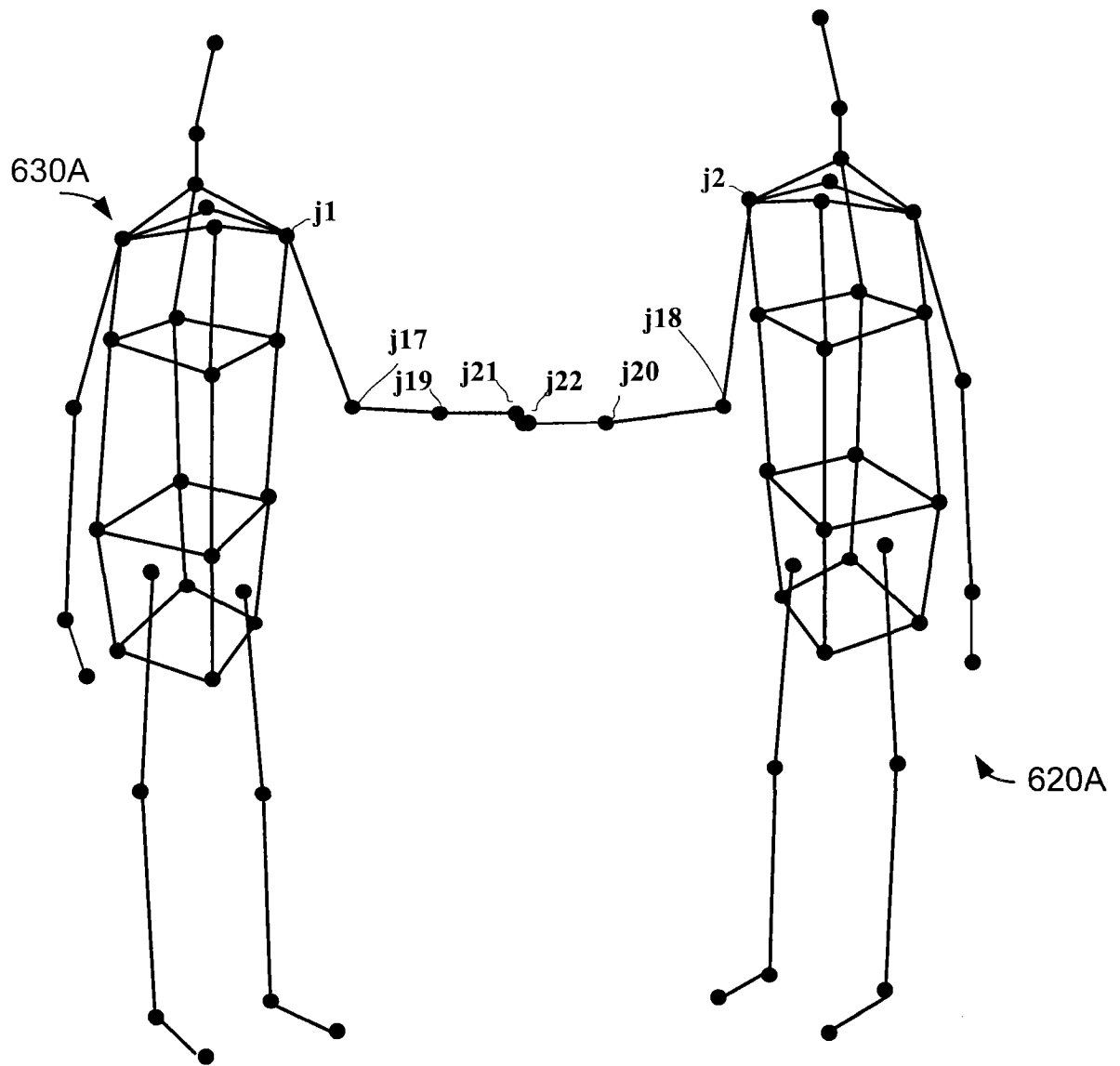


图 9

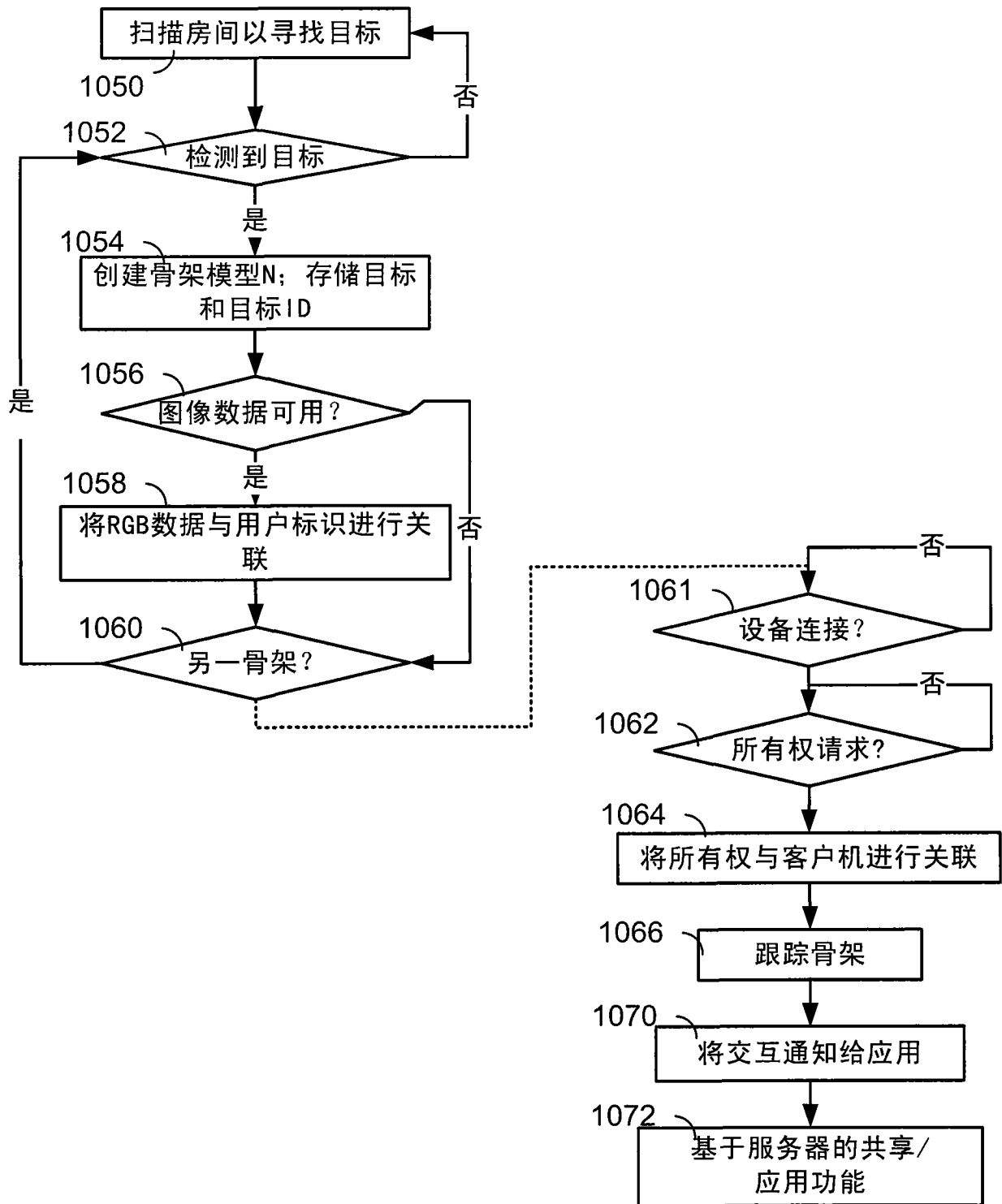


图 10

在一个实施例中，只执行这些步骤并且服务器上的应用接管

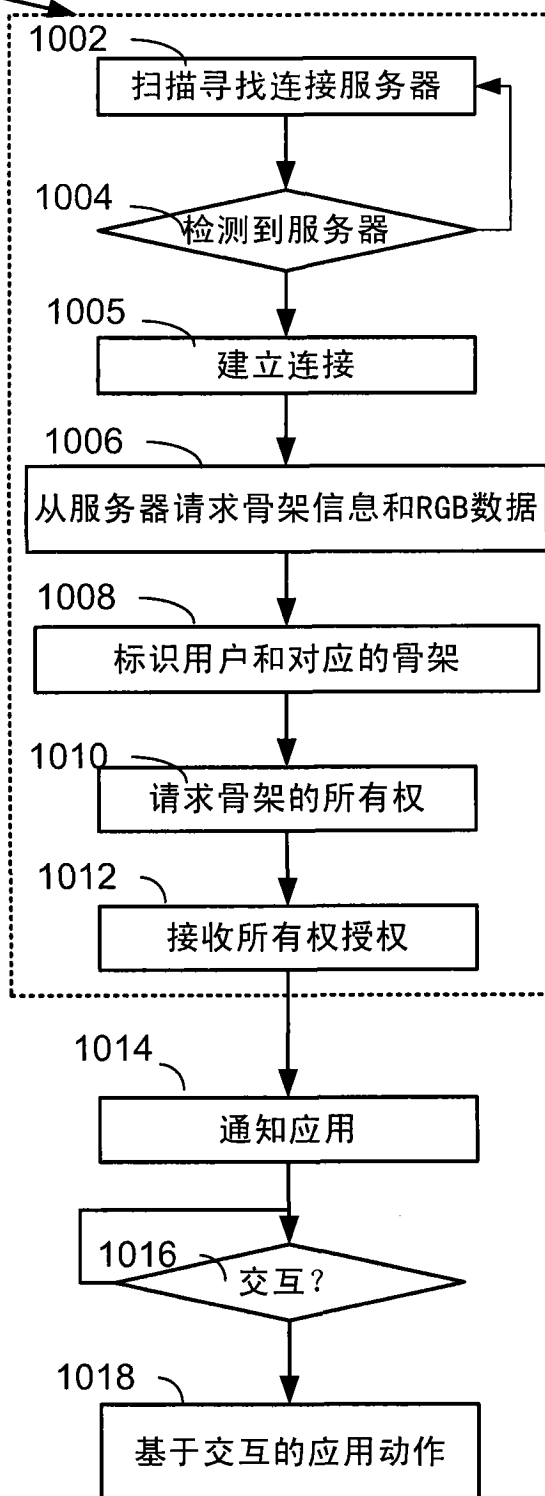


图 11



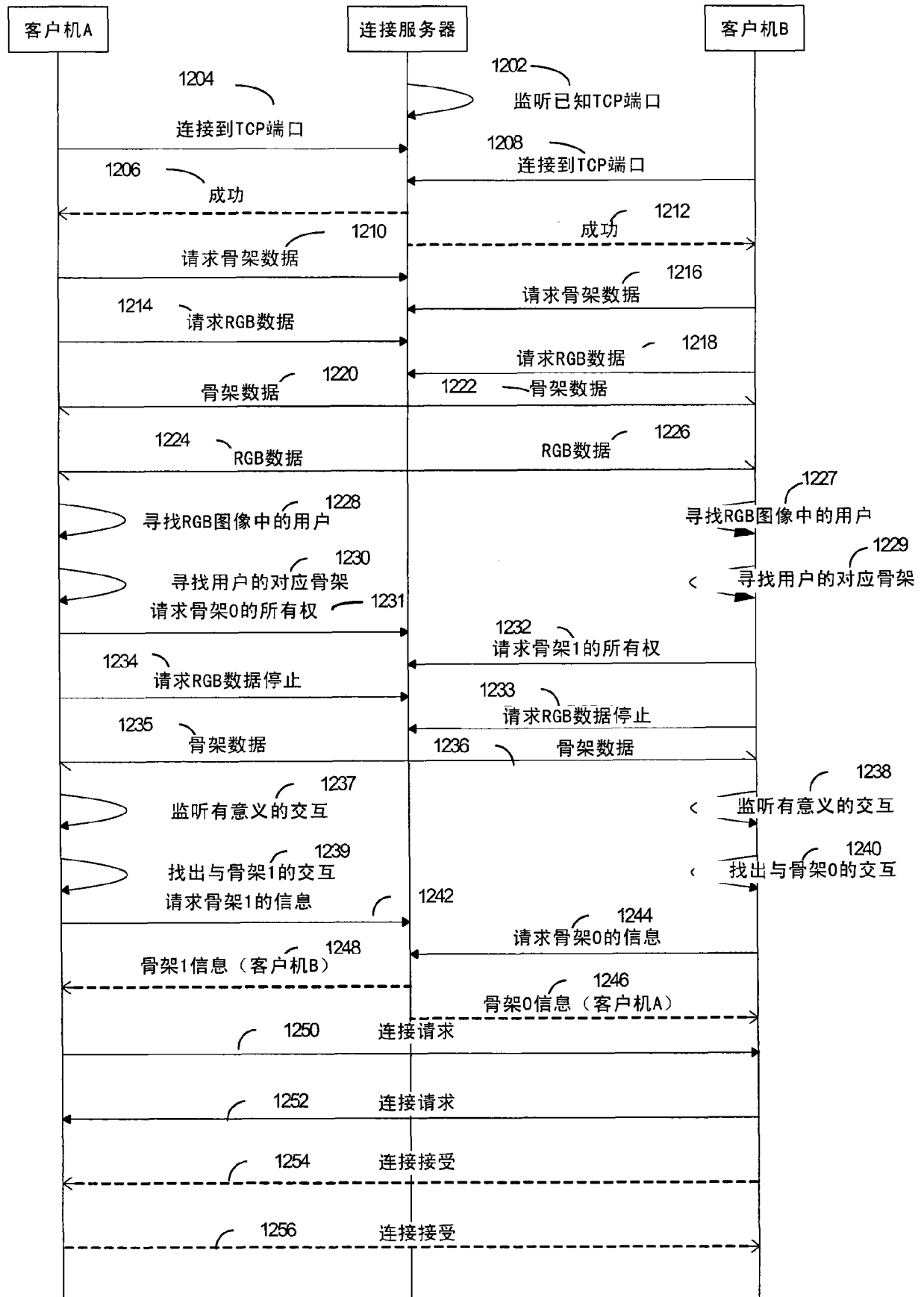


图 12

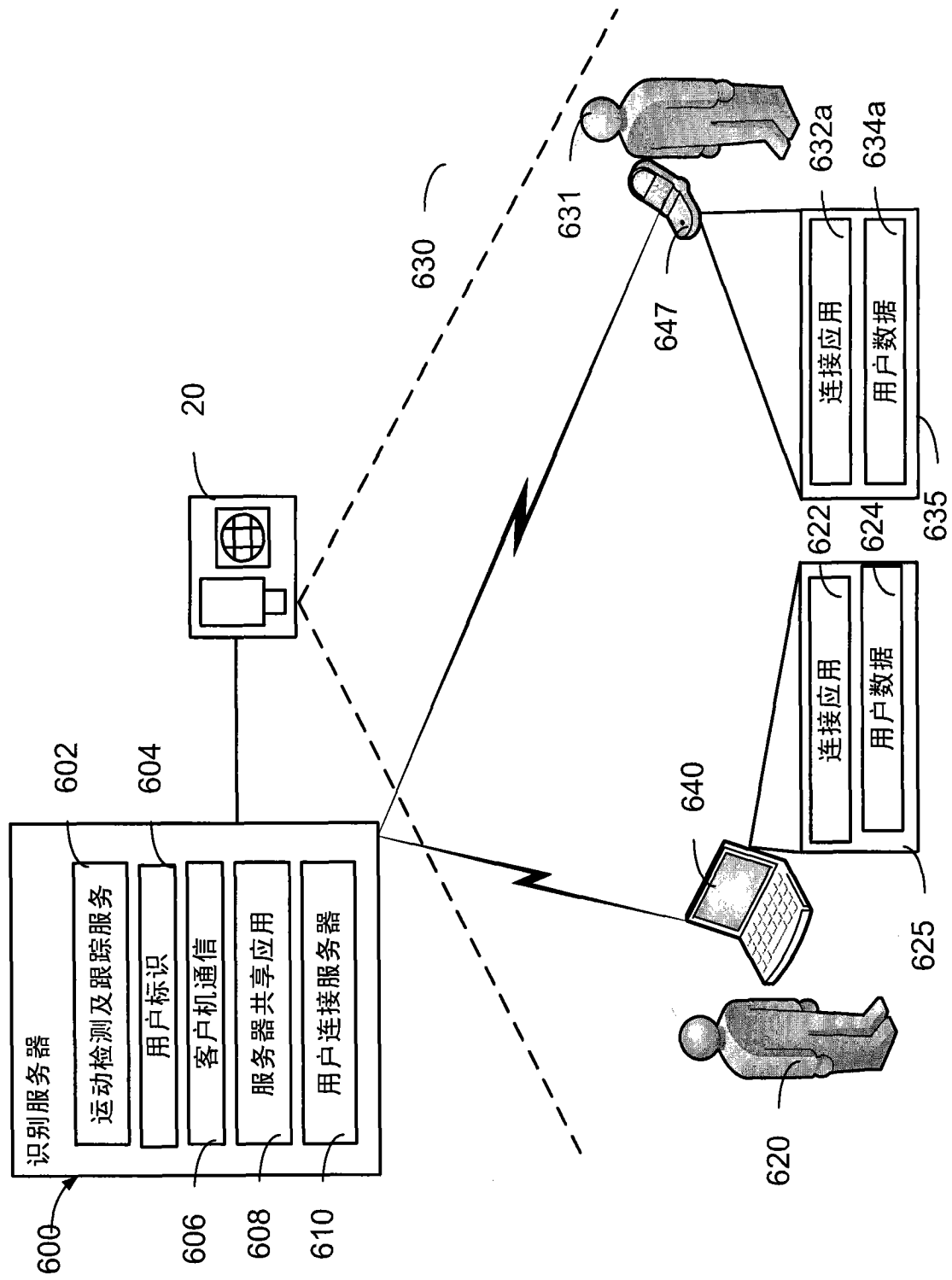


图 13

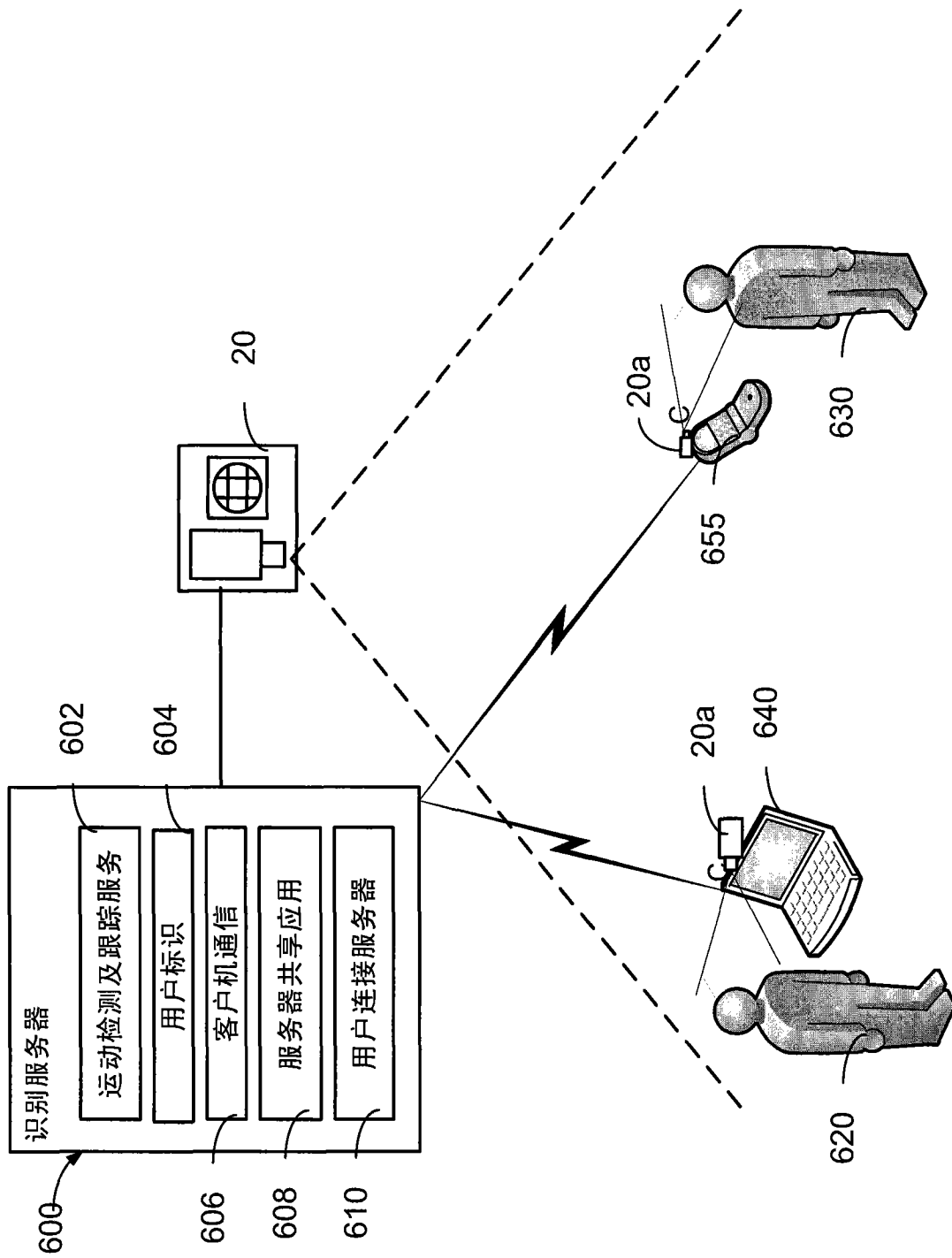


图 14

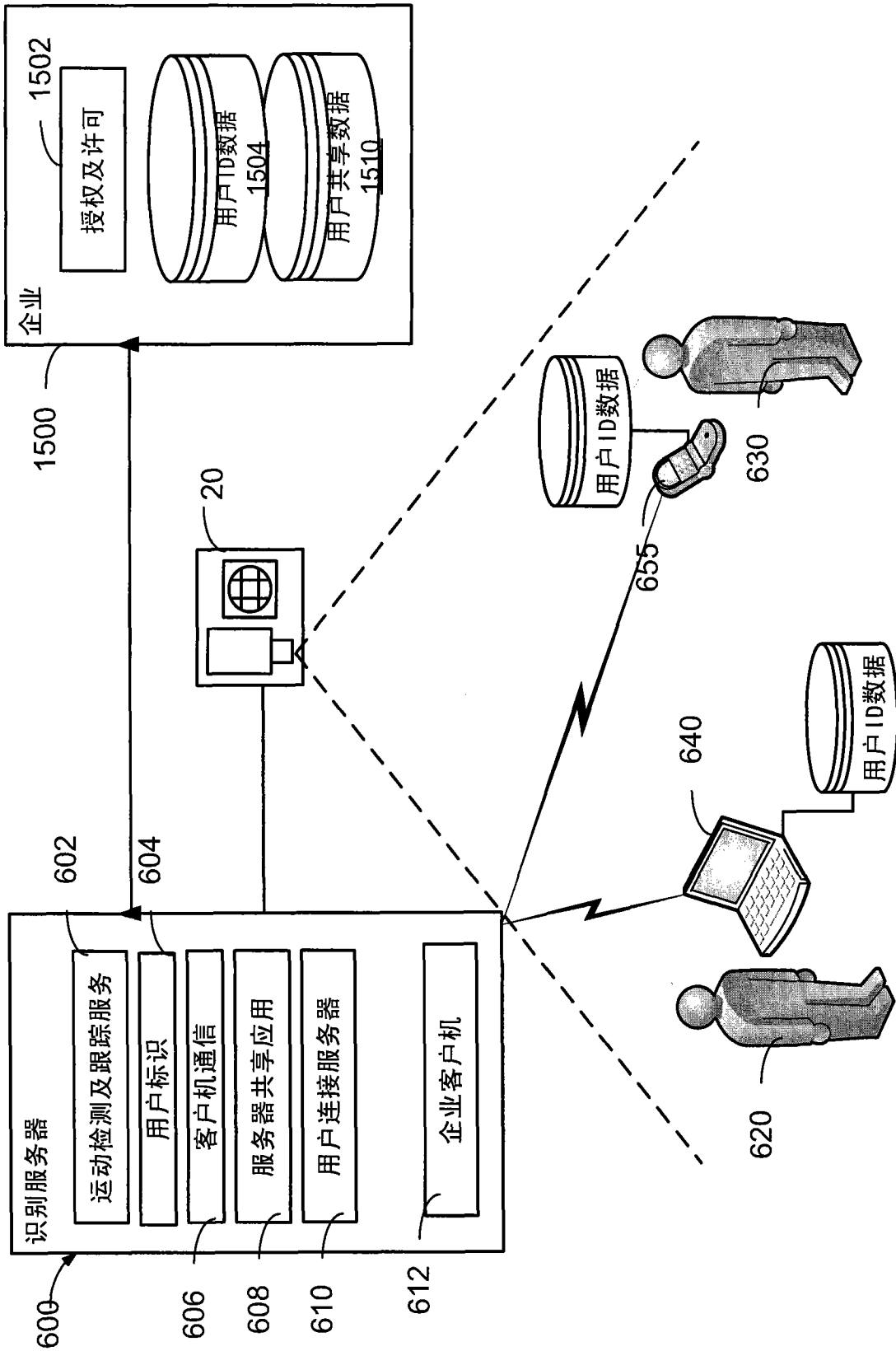


图 15

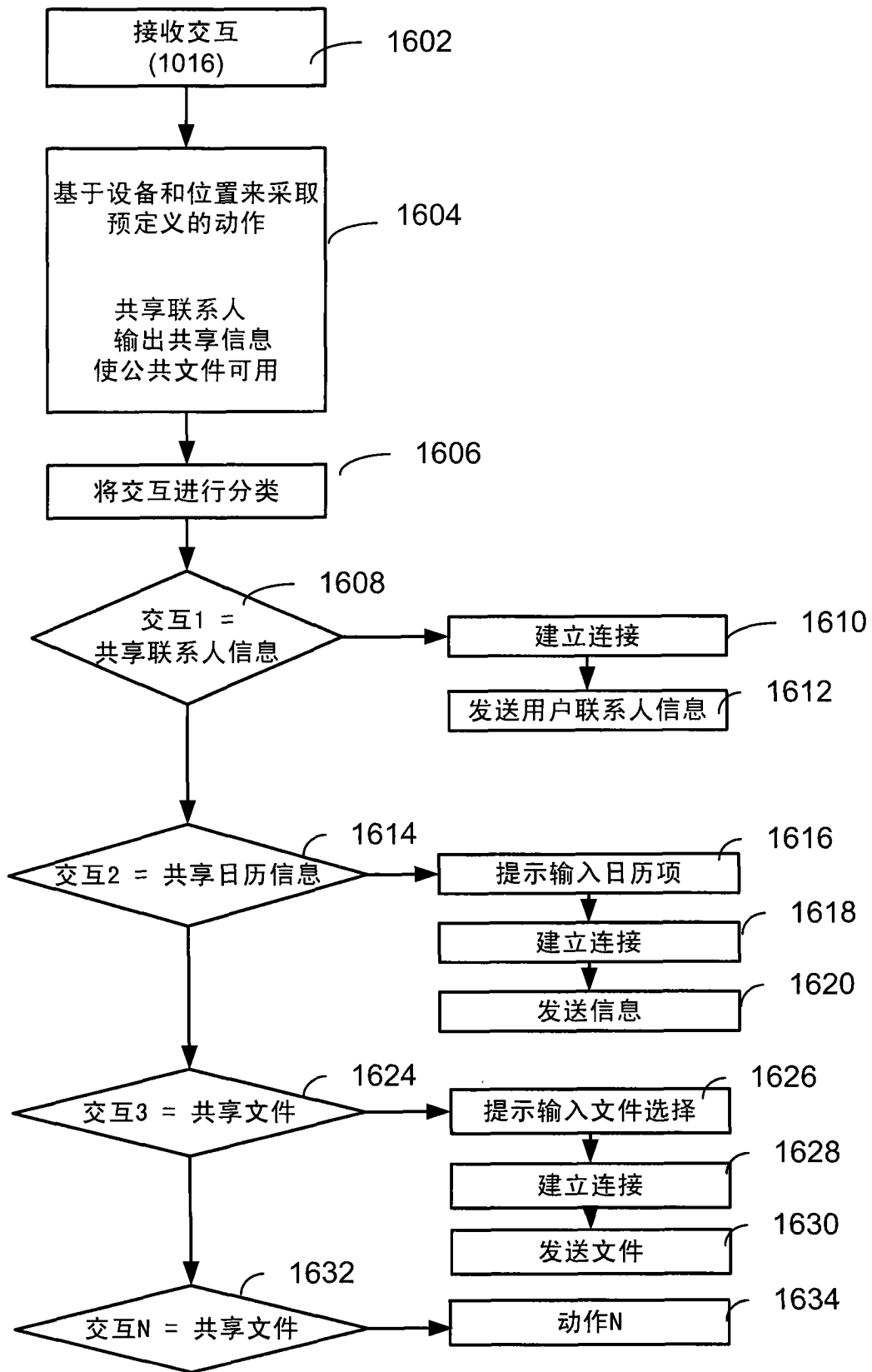


图 16

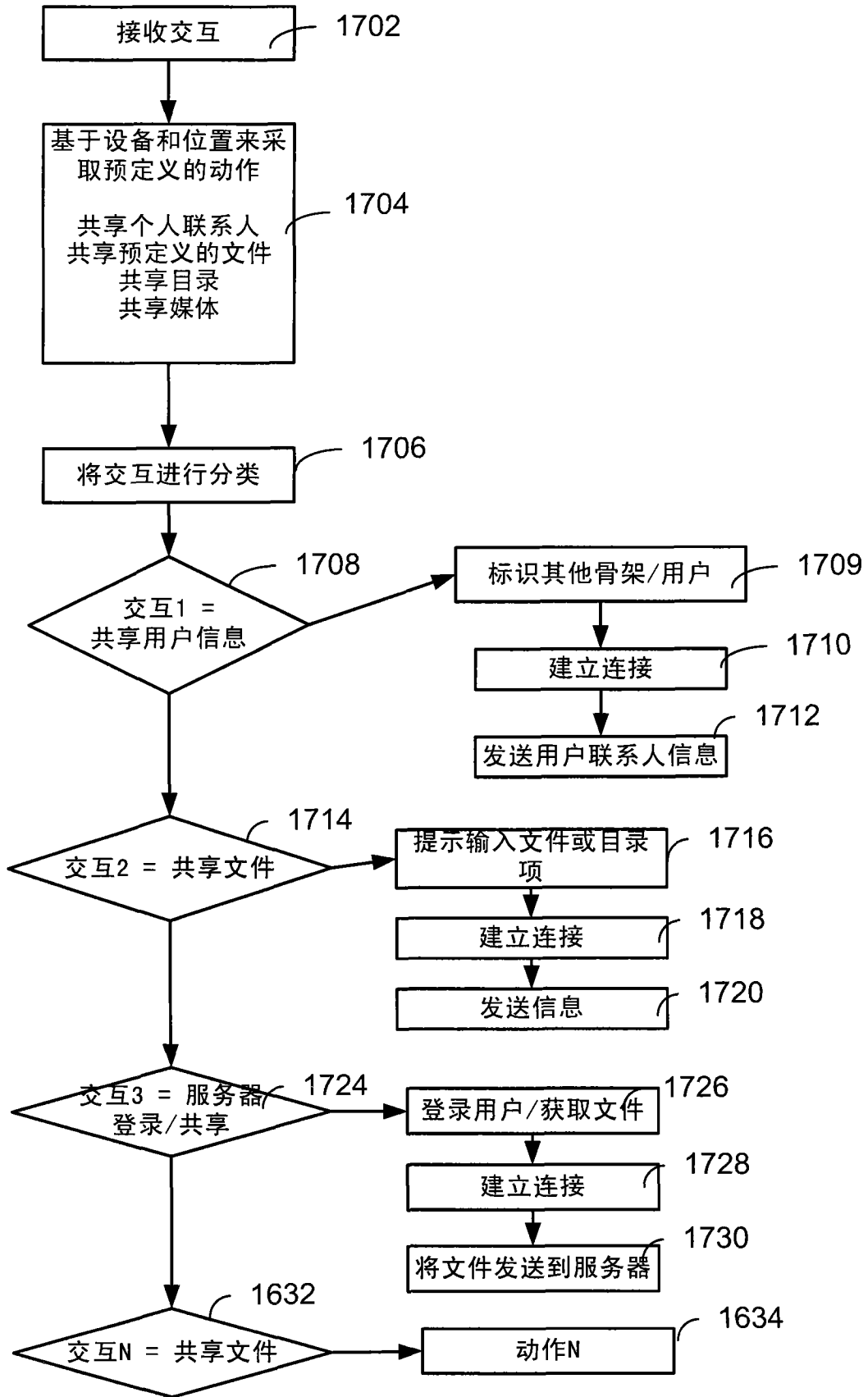


图 17

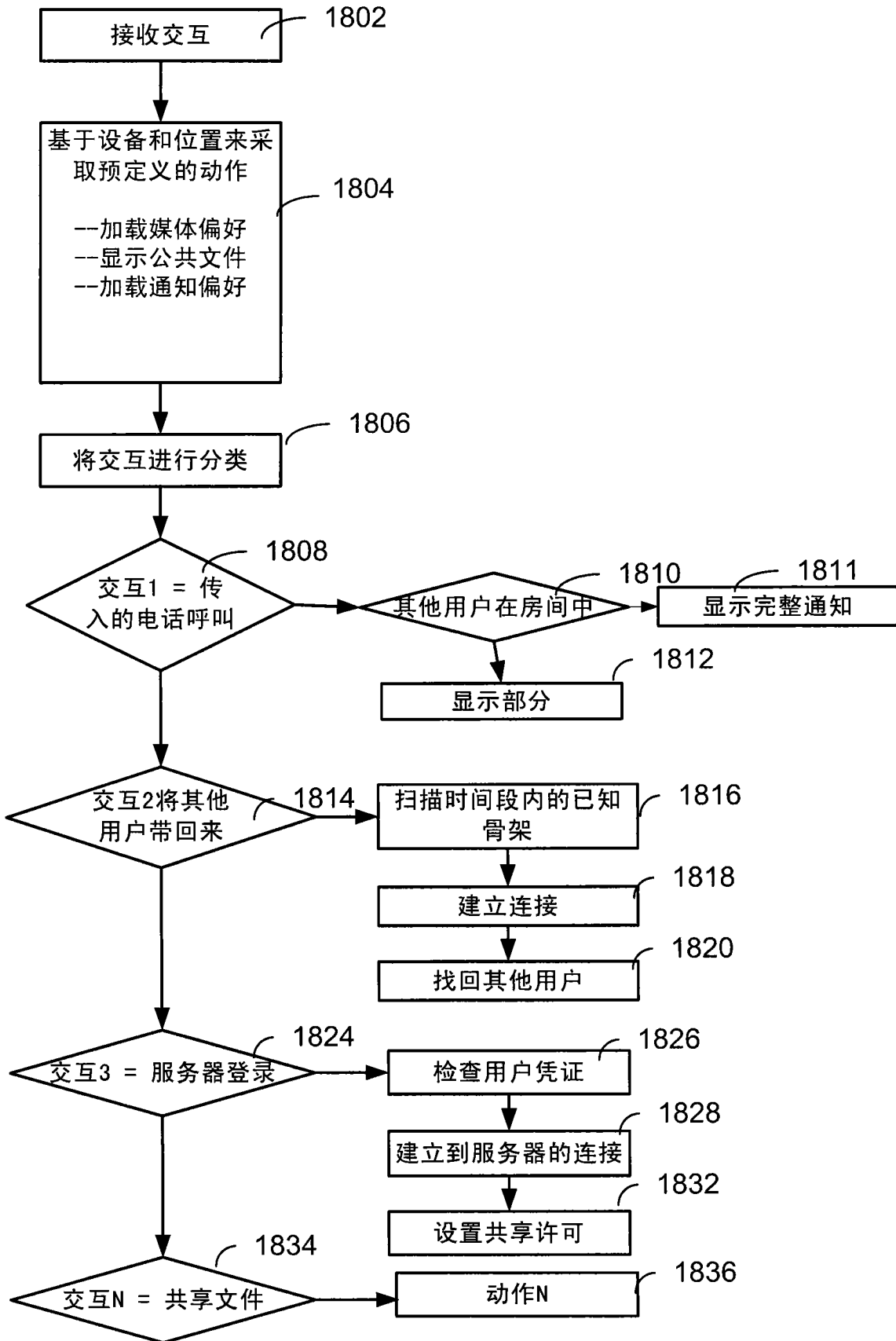


图 18