(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 101895556 B
(45)授权公告日 2013.01.16

(21)申请号 201010242542.1
(22)申请日 2006.06.23
(30)优先权数据
2005903357 2005.06.24 AU
(62)分案原申请数据
200680027097.9 2006.06.23
(73)专利权人 杜比实验室特许公司
地址 美国加利福尼亚州
(72)发明人 保罗·安德鲁·鲍斯特德
法尔扎德·撒菲伊
(74)专利代理机构 北京三友知识产权代理有限公司 11127
代理人 李辉
(51)Int.Cl.
H04L 29/06(2006.01)
A63F 13/10(2006.01)
(56)对比文件
US 6106399 A,2000.08.22, 全文.
US 6816846 B1,2004.11.09, 全文.
CN 1419796 A,2003.05.21, 全文.
审查员 阎赛
权利要求书 2 页 说明书 6 页 附图 4 页

(54)发明名称

沉浸式音频通信

(57)摘要

一种沉浸式音频通信。本发明提供一种将空间音频用于虚拟环境的系统和方法，所述虚拟环境能运行在具有有限处理能力的便携式设备上，并且利用低带宽通信连接。所述系统识别虚拟环境中的第一化身；并且确定所述化身是否满足重用准则，所述重用准则使得系统能够对所述第一化身重用已经为虚拟环境中的第二化身生成的音频信息。

CN 101895556 B

1. 一种将音频用于虚拟环境或音频会议系统的方法，该方法包括以下步骤：

识别与虚拟环境或音频会议系统相关联的第一化身，其中所述虚拟环境包括分布式虚拟环境，该分布式虚拟环境包括以下虚拟环境的一种或更多种：

多人在线游戏；

商务环境的虚拟环境；或者

教育环境的虚拟环境；

对于为第二化身创建的再现音频信息，确定是否满足音频重用准则；以及

在确定满足所述音频重用准则时，对所述第一化身重用所述再现音频信息。

2. 如权利要求 1 所述的方法，该方法还包括以下步骤：基于所述第一化身和第三化身之间的距离处理所述再现音频信息，

其中所述第一化身或第二化身中的一个或更多个包括虚拟环境或音频会议系统中的参与者。

3. 如权利要求 1 或权利要求 2 所述的方法，该方法还包括以下步骤：在确定不满足所述音频重用准则时，为所述第一化身创建替代的音频信息。

4. 如权利要求 2 所述的方法，其中确定是否满足所述音频重用准则的步骤包括以下步骤：确定所述第一化身是否会将所述再现音频信息感知成是从所述虚拟环境中所述第三化身所在的预定区域发出的。

5. 如权利要求 1 所述的方法，其中所述再现音频信息包括空间音频信息。

6. 一种将音频用于虚拟环境或音频会议系统的系统，该系统包括：

用于识别与虚拟环境或音频会议系统相关联的第一化身的装置，其中所述虚拟环境包括分布式虚拟环境，该分布式虚拟环境包括以下虚拟环境的一种或更多种：

多人在线游戏；

商务环境的虚拟环境；或者

教育环境的虚拟环境；

用于对于为第二化身创建的再现音频信息，确定是否满足音频重用准则的装置；以及

用于在确定满足所述音频重用准则时，对所述第一化身重用所述再现音频信息的装置。

7. 如权利要求 6 所述的系统，还包括：用于基于所述第一化身和第三化身之间的距离处理所述再现音频信息的装置，

其中所述第一化身或第二化身中的一个或更多个包括虚拟环境或音频会议系统中的参与者。

8. 如权利要求 6 或权利要求 7 所述的系统，还包括：用于在确定不满足所述音频重用准则时，为所述第一化身创建替代的音频信息的装置。

9. 如权利要求 7 所述的系统，其中确定是否满足所述音频重用准则的装置包括：用于确定所述第一化身是否会将所述再现音频信息感知成是从所述虚拟环境中所述第三化身所在的预定区域发出的装置。

10. 如权利要求 6 所述的系统，其中所述再现音频信息包括空间音频信息。

11. 一种将音频用于虚拟环境的计算机系统，该系统包括：

虚拟环境系统，其用于所述计算机系统以维护与虚拟环境有关的状态信息，其中识别

与虚拟环境相关联的第一化身，并且所述虚拟环境包括分布式虚拟环境，并且其中对于为第二化身创建的再现音频信息，确定是否满足音频重用准则，并且在确定满足所述音频重用准则时，对所述第一化身重用所述再现音频信息，该分布式虚拟环境包括以下虚拟环境的一种或更多种：

多人在线游戏；

商务环境的虚拟环境；或者

教育环境的虚拟环境；

用户计算设备，其用于：

将音频信息转发到虚拟环境系统；

接收与虚拟环境的音频场景有关的信息；以及

将所述音频场景再现给用户计算设备的用户，该用户计算设备通信地连接到所述虚拟环境系统。

沉浸式音频通信

[0001] 本申请是申请号为 200680027097.9(PCT/AU2006/000881)、申请日为 2006 年 6 月 23 日、发明名称为“沉浸式音频通信”的发明专利申请的分案申请。

技术领域

[0002] 本发明总体上涉及沉浸式音频通信 (immersive audio communication) 领域，并且具体地但决不排他地适用于与虚拟环境的化身 (avatar) 相关联的沉浸式音频场景。

背景技术

[0003] 传统地，分布式虚拟环境（如多人在线游戏）使用基本的装置以使得用户能够相互通信。例如，一些分布式虚拟环境已经采用了基于文本的通信，而其他的虚拟环境已经支持在因特网协议技术上的简单的单声道语音。

[0004] 已经提出，可以通过在虚拟环境中增加介质或高质量音频通信来改进用户对虚拟环境的感知。因此，已经进行了相当多的研究来为分布式虚拟环境提供实时沉浸式音频通信。大多数研究集中在为用户提供表达方向感和距离感的空间音频；即，音频信息听起来好像是从特定位置发出并且已经传播了一定距离。

[0005] 为分布式虚拟环境提供沉浸式音频通信的研究通常假设虚拟环境的用户使用计算设备（如，台式计算机），该计算设备具有合理量的处理能力，并且经由相对高带宽的数据链路连接到虚拟环境。这样，当前的沉浸式音频通信技术通常涉及跨越一个或更多中心服务器和用户的计算设备而进行分布式沉浸式音频场景计算。

[0006] 便携式无线计算设备的最近进展（如诺基亚 N-Gage 和 PlayStationPortable）意味着数量日益增长的用户将潜在地期望经由便携式无线计算设备访问虚拟环境。与台式计算设备相比，便携式无线计算设备具有有限处理能力，并且将经由合理低带宽的无线链路连接到虚拟环境。不幸的是，因为无线计算设备可能没有足够的处理能力来参与到分布式音频场景计算中，所以当今的沉浸式音频通信技术可能无法有效支持便携式无线计算设备（或其他具有相对低计算能力的计算设备）。其结果是，大部分计算密集的音频场景处理将被推回到虚拟环境服务器上。这基本意味着如果要支持便携式无线计算设备并为其提供沉浸式音频通信，可能需要采用更强大的服务器。

发明内容

[0007] 根据本发明的第一方面，提供了一种将音频用于虚拟环境的方法，该方法包括以下步骤：

[0008] 识别与所述虚拟环境相关联的第一化身；

[0009] 对于为与所述虚拟环境相关联的第二化身创建的音频信息，确定是否满足音频重用准则；以及

[0010] 在确定满足所述音频重用准则时，对所述第一化身重用所述音频信息。

[0011] 该方法的优点在于，与现有的沉浸式音频技术不同，其能够有利于便携式无线通

信设备（或具有相对低处理能力的任何其他设备）进行沉浸式音频通信。该优点源自所述方法涉及尽可能（即，当满足音频重用准则时）重用音频信息这一事实，这可以减少创建沉浸式音频的处理密集型任务。借助于示例实现，该方法可以由一个或更多个中央音频场景创建服务器来执行，以降低（多个）中央服务器的音频场景处理要求。与本发明相反，现有的沉浸式音频通信技术不采用音频信息重用，并且将因此而比采用本发明的系统需要更多处理资源。

[0012] 该方法优选地还包括以下步骤：基于所述第一化身和与虚拟环境相关联的第三化身之间的距离处理所述音频信息。

[0013] 执行处理音频的步骤的优点在于，其可以被用来将对距离的感知添加到重用的音频信息中。例如，该处理步骤可以涉及将所述音频信息衰减以表示该音频信息已经传播了一定距离。

[0014] 该方法优选地还包括以下步骤：在确定不满足所述音频重用准则时，为所述第一化身创建替代音频信息。

[0015] 创建替代音频信息的优点在于，在不满足重用准则的情况下提供了获得用于所述第一化身的音频的机制。

[0016] 确定是否满足音频重用准则的步骤优选地包括以下步骤：确定所述第一化身是否会将所述音频信息感知成是从所述虚拟环境中所述第三化身所在的预定区域发出的。

[0017] 确定第一化身是否会将所述音频信息感知成是从所述预定区域发出的优点在于，其提供了用于确保重用的音频信息具有合理地靠近精确创建的音频信息的空间方面的空间特性的技术，该精确创建的音频信息可以在创建替代音频信息的前述步骤期间创建。

[0018] 所述音频信息优选地包括空间音频信息。

[0019] 根据本发明的第二方面，提供了一种将音频用于虚拟环境的系统，该系统包括一处理装置，该处理装置被设置为执行以下步骤：

[0020] 识别与所述虚拟环境相关联的第一化身；

[0021] 对于为与所述虚拟环境相关联的第二化身创建的音频信息，确定是否满足音频重用准则；以及

[0022] 在确定满足所述音频重用准则时，对所述第一化身重用所述音频信息。

[0023] 所述处理装置优选地还被设置为执行以下步骤：基于所述第一化身和与所述虚拟环境相关联的第三化身之间的距离处理所述音频信息。

[0024] 所述处理装置优选地还被设置为执行以下步骤：在确定不满足所述音频重用准则时，为所述第一化身创建替代音频信息。

[0025] 将所述处理装置优选地设置为确定是否满足所述音频重用准则的所述步骤包括以下步骤：确定所述第一化身是否会将所述音频信息感知成是从所述虚拟环境中所述第三化身所在的预定区域发出的。

[0026] 所述音频信息优选地包括空间音频信息。

[0027] 根据本发明的第三方面，提供了一种包括至少一个指令的计算机程序，该指令被设置为使得计算设备执行根据本发明第一方面的方法。

[0028] 根据本发明的第四方面，提供了一种包括根据本发明第三方面的计算机程序的计算机可读介质。

附图说明

[0029] 尽管存在可能落入本发明的范围内的任何其他实施方式，下面仍将参照附图来描述仅作为实施例的本发明的实施方式，在附图中：

[0030] 图 1 示出了包括本发明的一实施方式的系统；

[0031] 图 2(a) 例示了由现有的沉浸式音频技术使用的技术；

[0032] 图 2(b) 例示了根据本发明一实施方式由图 1 的系统创建沉浸式音频的技术；以及

[0033] 图 3 是应用了图 1 的系统的虚拟环境的表示。

具体实施方式

[0034] 参照图 1，系统 100 包括虚拟环境系统 102；用户计算设备 104；以及通信系统 106。

[0035] 虚拟环境系统 102 的主要功能是维护虚拟环境的状态信息并为虚拟环境的参与者提供沉浸式音频通信。在本发明的该实施方式中，虚拟环境是多人在线游戏的战场，并且化身表示虚拟环境的参与者（即，用户计算设备 104 的用户）并且是动画战士的形式。由虚拟环境系统 102 维护的状态信息包括例如虚拟环境中化身的位置；即，战场中战士的位置。

[0036] 应注意到，本发明并不限于用于多人在线游戏的虚拟环境和化身。本发明适用的虚拟环境的范围包括例如商务环境（如虚拟员工会议）或教育环境（如虚拟讲座）的虚拟环境。

[0037] 为了执行维护状态信息和提供沉浸式音频通信的主要功能，虚拟环境系统 102 包括传统的计算机硬件，该计算机硬件包括主板、中央处理单元、随机存取存储器、硬盘、联网硬件和电源。除了硬件外，虚拟环境系统 102 包括操作系统（如 Linux，其可从 www.redhat.com 获得），该操作系统驻留在系统 102 的硬盘上，并且与硬件协同以提供其中可以执行软件应用的环境。在这点上，虚拟环境系统 102 的硬盘加载有用于维护状态信息的虚拟环境服务器应用（如 Quake 引擎，其可从 www.idsoftware.com 获得）、以及用于为虚拟环境的参与者提供沉浸式音频通信的沉浸式音频通信服务器应用。

[0038] 用户计算设备 104 是以便携式无线通信设备的形式，如诺基亚 N-Gage 和 PlayStation Portable。然而，容易理解到，本发明并不限于便携式无线通信设备。预期在本发明的另选实施方式中，用户计算设备 104 可以是例如膝上型或台式计算机。每个用户计算设备 104 包括传统的计算机硬件，该计算机硬件包括主板、中央处理单元、随机存取存储器、硬盘、电源、监视器和用户信息输入装置（例如，键盘）。除了硬件外，每个用户计算设备 104 的硬盘还加载有黑莓 (BlackBerry) 操作系统。该操作系统与计算设备 104 的硬件交互，以提供其中可以执行软件应用的环境。在这点上，每个用户计算设备 104 的硬盘加载有虚拟环境客户端应用和沉浸式音频通信客户端。

[0039] 虚拟环境客户端应用被设置为向虚拟环境服务器应用发送和从虚拟环境服务器应用接收用于虚拟环境的虚拟环境状态信息，该虚拟环境服务器应用加载在虚拟环境系统 102 上。沉浸式音频通信客户端被设置为向虚拟环境系统 102 上的沉浸式音频通信服务器发送并从其接收音频信息。先前描述了每个用户计算设备 104 加载有操作系统，如用于诺基亚 N-Gage 的 Symbian 操作系统。本发明并不限于 Symbian 操作系统，并且同样，本发明适用于加载在用户计算设备 104 上的不同操作系统，包括例如微软 Windows XP 或 Linux（当

计算设备 104 是台式计算机的形式时，通常使用这两者）。

[0040] 通信系统 106 使得每个用户计算设备 104 的虚拟环境客户端应用和虚拟环境系统 102 的虚拟环境服务器应用能够彼此交换数据（更具体地说，即状态信息）。通信系统 106 还使得每个用户计算设备 104 的沉浸式音频通信客户端应用和虚拟环境系统 102 的沉浸式音频通信服务器应用能够彼此交换数据（更具体地说，即音频信息）。

[0041] 为了支持数据交换，通信系统 106 包括移动电话网络 108，用于向用户计算设备 104 发送数据并从其接收数据。本发明并不限于使用移动电话网络 108，并且本发明的另选实施方式可以采用例如基于 802.11 的无线网络。除了移动电话网络 108 外，通信系统 106 包括用于向虚拟环境系统 102 发送数据并从其接收数据的基于公共访问因特网协议的分组交换网络 110，如因特网。移动电话网络 108 和通信系统 106 经由通信链路 112 彼此互连，该通信链路 112 是以高速光纤链路的形式。通信链路 112 使得能够在移动电话网络 108 和分组交换网络 110 之间交换数据。

[0042] 为了使得虚拟环境系统 102 和用户计算设备 104 能够经由通信系统 106 交换数据，虚拟环境服务器 102 经由高速数据链路形式的通信链路 114 连接到分组交换网络 110。此外，各用户计算设备 104 经由无线 GPRS 链路 116 连接到移动电话网络 108。

[0043] 本发明的实施方式主要涉及提供沉浸式音频通信。这样，下面对实施方式的描述集中在由加载在用户计算设备 104 上的沉浸式音频通信客户端应用、和加载在虚拟环境系统 102 上的沉浸式音频通信服务器应用所执行的功能。

[0044] 有两种由沉浸式语音通信客户端应用执行的主要功能。第一个功能是将由用户计算设备 104 捕获的音频信息（例如，用户的语音）转发到加载在虚拟环境系统 102 上的沉浸式音频通信服务器。第二个功能是从加载在虚拟环境系统 102 上的沉浸式音频通信服务器接收音频场景（沉浸式音频通信信息），并且将音频场景再现给用户计算设备 104 的用户。对音频场景的再现基本涉及通过扬声器或耳机播放音频场景。

[0045] 加载在虚拟环境系统 102 上的沉浸式音频通信服务器应用还执行两个主要功能。首先，服务器应用处理从沉浸式音频通信客户端接收的音频信息，从而为虚拟环境中的每个化身创建音频场景。第二，服务器应用将音频场景转发到用户计算设备 104。与创建音频场景的第一功能相关，预期可以采用任何适当的空间音频处理算法。然而，在本实施方式的情况下，沉浸式音频通信服务器应用被设置为使用头相关传输函数 (HRTF) 算法，在题为“3-D Sound for Virtual Reality and Multimedia”的出版物 (Academic Press Professional，Cambridge，MA，USA，1994，作者 Durand R. Begault) 中描述了其细节。

[0046] 除了执行头相关传输函数外，沉浸式音频通信服务器应用被设置为执行音频场景重用算法。音频场景重用算法的目的在于有可能最小化由沉浸式音频服务器应用执行的头相关传输函数操作，从而这样又最小化了创建音频场景需要的处理资源。正是重用算法实质上使得本发明适于为低处理能力设备（如无线便携式计算设备）提供用于虚拟环境的沉浸式音频。以下参照图 2 描述了音频场景重用算法，图 2 表示包括三个化身 A1、A2 和 A3 的虚拟环境，这三个化身面向相同的绝对方向（例如，北方）。做出该假设是为了简化对本发明的实施方式的描述。然而实际上化身可以朝向不同方向。

[0047] 图 2 中示出重用的实施例。图 2(a) 示出现有沉浸式音频通信技术如何将来自化身 A1 的实时语音精确地放置在 A2 和 A3 的音频场景中。这需要两个空间放置操作，这两个

空间放置操作涉及两个独立的头相关变换函数计算 - 每个听众 (A1 和 A3) 一个。执行精确的放置可能是计算密集的，特别是涉及很多化身时。相反，图 2(b) 示出了重用算法的结果。A1 是使用头相关传输函数在 A3 的音频场景中精确再现的。A3 的音频场景被重用为化身 A2 的音频场景。结果，A2 将不精确地听到 A1，就好像 A1 来自 A1'一样。因为距 A1 比距 A3 更远，所以重用音频场景还将衰减。该重用场景的代价是一个空间放置操作（即，一个头相关传输函数）和一个衰减操作。

[0048] 应注意到，场景算法是这样的，即在图 2(b) 中描述的场景中，假设重用音频场景不是特别不精确，如果可以增加重用音频场景的数量，则该算法可能对 A2 和 A3 两者不精确地再现 A1（即，重用音频场景）。

[0049] 重用算法并不仅仅是无区别地重用音频场景。在重用音频场景前，重用算法确定是否满足音频重用准则。在这点上，本发明的该实施方式使用角度差错的概念作为重用准则。参照图 2(b)，角度差错是由从化身 A1 到化身 A2 和 A3 延伸的矢量所定义的角度。重用算法是这样的，即如果由所述矢量定义的角度等于或小于预定角度（其也称作“角度误差”），则对 A2 重用 A3 的音频场景。

[0050] 应注意到，重用准则不限于先前描述的角度误差的概念。在本发明的另选实施方式中，重用准则可以基于化身之间的距离。另选的是，角度差错概念可以不固定，并且可以例如是在对特定距离（例如 10m) 内的化身的相对小的角度误差和对于大于该距离的化身的较大的角度误差。

[0051] 由沉浸式音频服务器应用执行的功能的更正式的描述如下：

[0052] 给定具有 N 个化身 (A1，A2，…An) 的虚拟世界，每个化身由客户端 (C1，C2，…CN) 控制。从这些客户端中的每一个接收音频采样 (Vi)。

[0053] 1. 计算每个化身 (i) 可以听到的说话的化身的集合，并且将该集合放置在 Si 中。

[0054] 2. 重复，直到每个集合 S_i 为空，或已经达到一处理限度

[0055] a. 对于 i = 1 到 N，

[0056] i. 找到集合 S_i 中最接近的说话者 (A_j)

[0057] ii. 如果以前在可接受的误差内尚未计算出从化身 i 的角度来自 A_j 的音频的空间化版本 (H_{ji})，则：

[0058] 1. 根据从化身 A_i 的角度精确再现的化身 A_j，创建音频 V_j 的空间版本 H_{ji}。

[0059] 2. 将 H_{ji} 的加权副本添加到集合 R_i，该集合 R_i 是针对听众 i 在音频场景中再现的对象的集合。所应用的加权是在虚拟世界中 i 和 j 之间的相对距离的函数。

[0060] iii. 其他

[0061] 1. 将可接受的误差内再现的对象 H_{ji} 的加权副本添加到集合 R_i。所应用的加权是在虚拟世界中 i 和 j 之间的相对距离的函数。

[0062] iv. 从集合 S_i 中移除 A_j

[0063] b. 结束循环

[0064] 3. 结束

[0065] 作为沉浸式音频服务器如何比现有沉浸式音频技术需要更少处理能力的示例，考虑下面的例子，其中可接受的角度误差在距听众 0 到 10 米之间为 10 度，可接受的角度误差从 10 米到每个化身听觉范围的边缘为 30 度。使用在图 3 中描绘的虚拟世界，假设所有化

身可以互相听到并且围绕每个化身的圆半径为 10m。解出重用算法的以上正式定义：

[0066] A1 需要 A2 的精确再现 (H2,1)

[0067] A2 需要 A1 的精确再现 (H1,1)

[0068] A3 需要 A2 的精确再现 (H2,3)

[0069] A1 需要 A3 的精确再现 (H3,1)

[0070] A2 可以重用 H3,1,因为它将引入小于 30 度的角度误差

[0071] A3 可以重用 H1,2,因为它将引入小于 30 度的角度误差

[0072] 使用重用机制,服务器中需要的再现计算量从(使用现有技术将可能出现的)6 减少到 4。

[0073] 容易理解的是,即使已经在用来支持虚拟环境的上下文中描述了本发明的实施方式,但是本发明不限于使用虚拟环境。因而,本发明可以被用于提供用作例如仅有音频的会议系统的沉浸式音频系统。

[0074] 尽管已经参照本实施方式描述了本发明,但是本领域技术人员将理解,在不背离本发明的范围的情况下,可以做出改变、变化和改进,并且其元件和其步骤可以由等同物替代。此外,在不背离本发明的主要范围的情况下,可以做出许多修改,以使本发明适于特定情况或使材料适于本发明的教导。尽管上面没有明确的描述,但是旨在将这样的改变、变化、修改和改进包含在本发明的范围和精神内。因此,本发明不是要限于这里描述的特定实施方式,而是将包括落入独立权利要求范围内的所有实施方式。

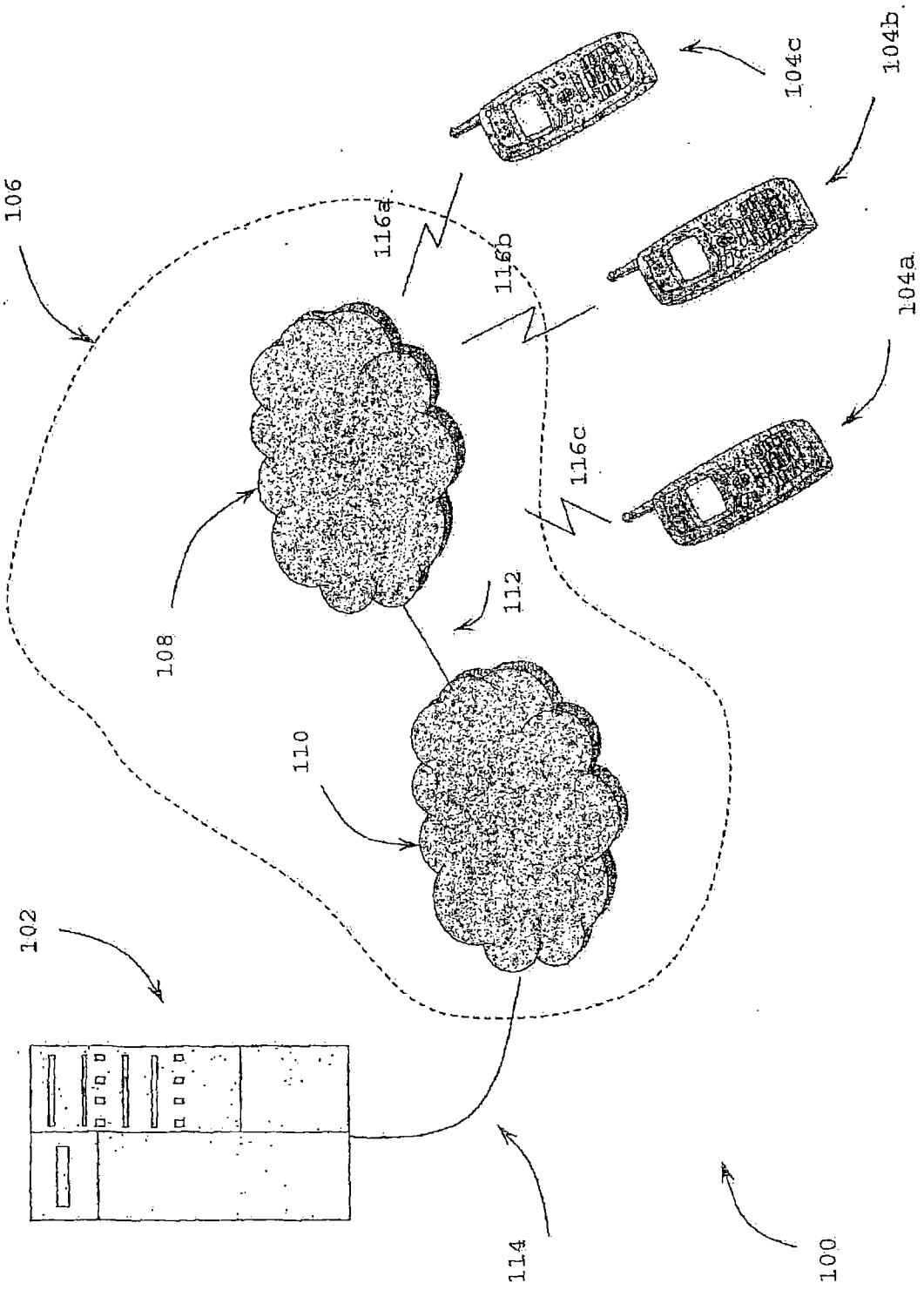


图 1

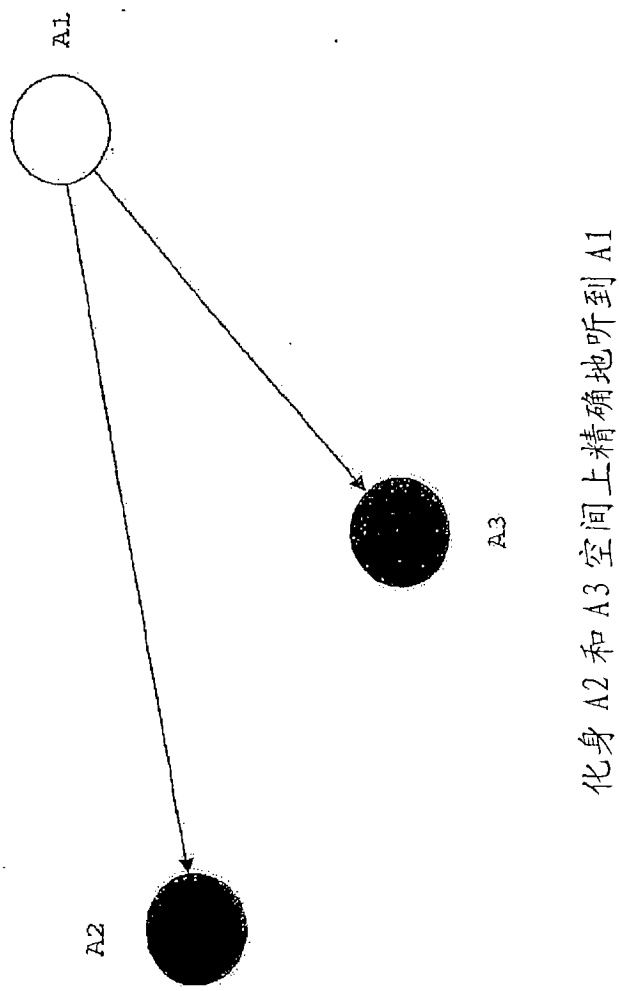


图 2(a)

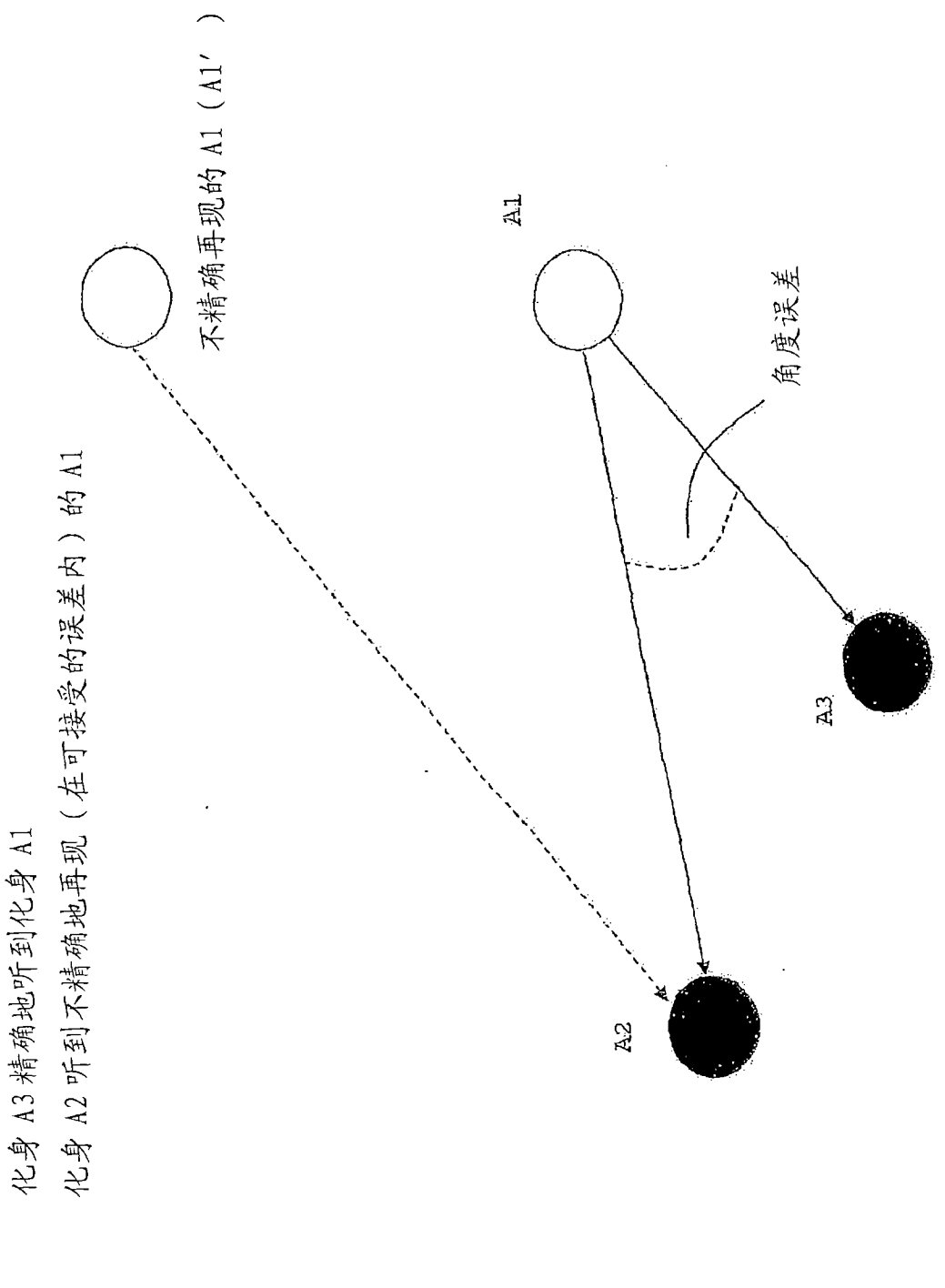

化身 A3 精确地听到化身 A1
化身 A2 听到不精确地再现（在可接受的误差内）的 A1
不精确再现的 A1（A1′）
A1
角度误差
A3
A2

图 2(b)

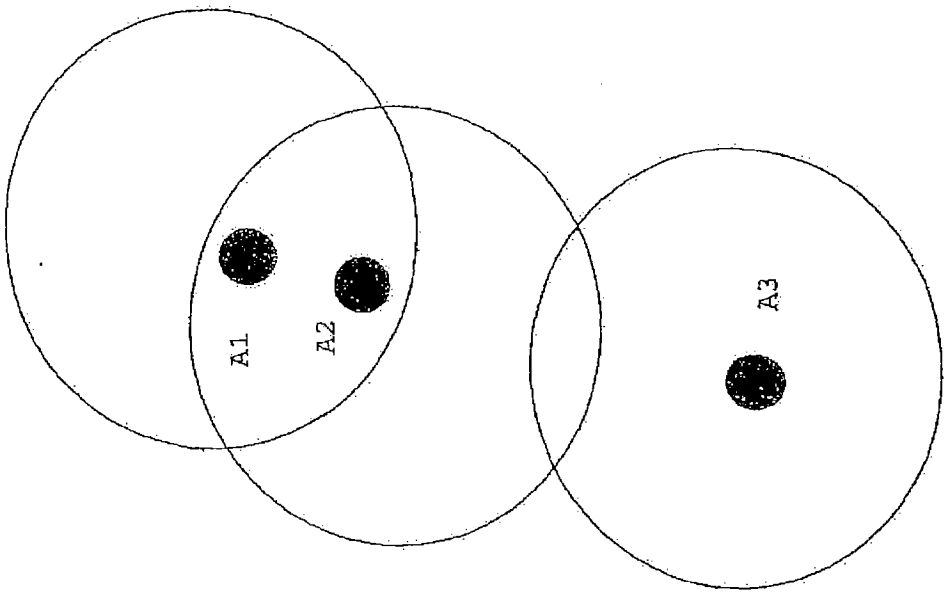


图 3