

(12) 发明专利

(10) 授权公告号 CN 101132839 B

(45) 授权公告日 2011. 09. 07

(21) 申请号 200680006438. 4

(22) 申请日 2006. 04. 28

(30) 优先权数据

60/678, 413 2005. 05. 05 US

(85) PCT申请进入国家阶段日

2007. 08. 29

(86) PCT申请的申请数据

PCT/US2006/016670 2006. 04. 28

(87) PCT申请的公布数据

W02006/121681 EN 2006. 11. 16

(73) 专利权人 索尼计算机娱乐公司

地址 日本东京都

(72) 发明人 R·L·马克斯 毛夏东

(74) 专利代理机构 北京泛华伟业知识产权代理

有限公司 11280

代理人 王勇

(51) Int. Cl.

A63F 13/02(2006. 01)

(56) 对比文件

US 5993314 A, 1999. 11. 30, 全文.

US 20020048376 A1, 2002. 04. 25, 说明书第 0003-0172 段, 附图 1-18.

US 20040046736 A1, 2004. 03. 11, 摘要和说明书第 0467 段.

审查员 张扬

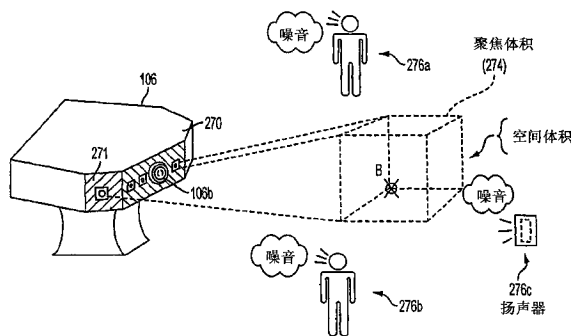
权利要求书 2 页 说明书 7 页 附图 7 页

(54) 发明名称

结合计算机交互处理的选择性声源监听

(57) 摘要

本发明提供一种在与计算机程序交互期间捕捉图像和声音的方法及设备。该设备包括一图像捕捉单元, 被配置为捕捉一个或多个图像帧。还提供一声音捕捉单元。该声音捕捉单元被配置为识别一个或多个声源。该声音捕捉单元产生能够被解析以确定一个聚焦区域的数据, 在此聚焦区域可处理声音以至基本上排除该聚焦区域之外的声音。以此方式, 从聚焦区域捕捉和处理的声



1. 一种在与计算机程序交互期间捕捉图像和声音的设备,包括:  
被配置为捕捉一个或多个图像帧的图像捕捉单元;  
声音捕捉单元,该声音捕捉单元被配置为识别一个或多个声源,该声音捕捉单元产生能够被解析以确定一个聚焦区域的数据;在所述聚焦区域中,处理声音以基本上排除该聚焦区域之外的声音,所述聚焦区域至少部分通过处理从所捕捉的一个或多个图像帧中获得的图像数据而被识别出,其中为该聚焦区域捕捉并处理的声音被用于与该计算机程序的交互通信。
2. 根据权利要求1所述的在与计算机程序交互期间捕捉图像和声音的设备,其中所述声音捕捉单元包括一麦克风阵列,该麦克风阵列被配置为接收来自一个或多个声源的声音,该一个或多个声源的声音限定到达每个麦克风的声源路径。
3. 根据权利要求2所述的在与计算机程序交互期间捕捉图像和声音的设备,其中所述声音路径包括特定延迟,该延迟使得能够计算出相对于所述捕捉图像和声音的设备的每一个或多个声源中的每一个的方向。
4. 根据权利要求1所述的在与计算机程序交互期间捕捉图像和声音的设备,该设备进一步包括:  
计算系统,用于与所述捕捉图像和声音的设备相接口,该计算系统包括:  
处理器,以及  
存储器,该存储器被配置为存储至少部分计算机程序和选择性声源监听代码,该选择性声源监听代码能够识别所述一个或多个声源中的哪一个为聚焦区域。
5. 根据权利要求1所述的在与计算机程序交互期间捕捉图像和声音的设备,其中所述声音捕捉单元包括至少四个麦克风,并且四个麦克风中的一个麦克风与其它麦克风不在同一平面上。
6. 根据权利要求5所述的在与计算机程序交互期间捕捉图像和声音的设备,其中所述四个麦克风限定一个空间体积。
7. 根据权利要求6所述的在与计算机程序交互期间捕捉图像和声音的设备,其中所述空间体积被限定为与计算机程序交互期间进行监听的一个聚焦体积。
8. 根据权利要求7所述的在与计算机程序交互期间捕捉图像和声音的设备,其中所述计算机程序是游戏程序。
9. 根据权利要求1所述的在与计算机程序交互期间捕捉图像和声音的设备,其中所述计算机程序是游戏程序。
10. 根据权利要求9所述的在与计算机程序交互期间捕捉图像和声音的设备,其中所述图像捕捉单元是摄像机,并且所述声音捕捉单元由两个或更多个麦克风组成的阵列所限定。
11. 在与计算机程序交互期间的选择性声源监听方法,包括:  
以两个或更多的声源捕捉麦克风接收来自于一个或多个声源的输入;  
确定来自于每个声源的延迟路径;  
对于接收到的每个输入,识别一个或多个声源中每个声源的方向;以及  
滤除不在所识别的一个聚焦区域的方向上的声源,所述聚焦区域为与计算机程序的交互提供声源;

其中滤除步骤接收在图像捕捉单元解析之后所处理的输入数据,所述图像捕捉单元产生图像数据以至少部分识别出所述聚焦区域,且所述图像捕捉单元被定向设置,以为计算机程序接收图像输入。

12. 根据权利要求 11 所述的在与计算机程序交互期间的选择性声源监听方法,其中所述计算机程序是一个游戏,并且该游戏接收来自于图像数据和声音数据的交互输入,所述声音数据来自于所述聚焦区域的声源。

13. 根据权利要求 11 所述的在与计算机程序交互期间的选择性声源监听方法,其中所述两个或更多的声音捕捉麦克风包括至少 4 个麦克风,并且四个麦克风中的至少一个麦克风与其它麦克风在不同的平面上。

14. 根据权利要求 13 所述的在与计算机程序交互期间的选择性声源监听方法,其中所述识别所述一个或多个声源的每个接收到的输入的方向包括处理一个三角剖分算法,该三角剖分算法限定相对于一个位置的方向,在所述位置中,以两个或更多的声源捕捉麦克风接收来自于所述一个或多个声源的输入。

15. 根据权利要求 14 所述的在与计算机程序交互期间的选择性声源监听方法,该方法进一步包括:

缓存从一个或多个声源的接收到的输入,该一个或多个声源与所述两个或更多声源捕捉麦克风相关;以及

对所接收的缓存的输入进行延迟处理;

所述滤除,包括,

选择声源中的一个,该选择的声源输出是来自于每个声源捕捉麦克风的聲音的叠加。

16. 一个游戏系统,包括:

一个图像-声音捕捉装置,该图像-声音捕捉装置被配置为与能够执行交互计算机游戏的计算系统相接口,所述图像捕捉装置包括,

视频捕捉硬件,该视频捕捉硬件能够被置于捕捉来自于一个聚焦区域的视频的位置,以及

一个麦克风阵列,该麦克风阵列捕捉来自于一个或多个声源的声音,每个声源均被识别并与相对于图像-声音捕捉装置的一个方向相关联,与视频捕捉硬件相关的所述聚焦区域被配置为用于至少部分识别出位于聚焦区域附近的方向的一个声源。

17. 根据权利要求 16 所述的游戏系统,其中所述视频捕捉硬件接收视频数据以使得能够与计算机游戏中的角色交互。

18. 根据权利要求 16 所述的游戏系统,其中在所述聚焦区域附近的所述声源使得能够与计算机游戏交互或与其他游戏用户语音通信。

19. 根据权利要求 18 所述的游戏系统,其中在所述聚焦区域外的声源被滤除于与计算机游戏的交互之外。

## 结合计算机交互处理的选择性声源监听

### 背景技术

[0001] 近年来,可看到视频游戏工业已有许多变化,随着计算能力的提高,视频游戏的开发者们同样制作出能够利用计算能力的增长的游戏软件。为此,视频游戏的开发者们已经开发出结合复杂的运算和数学操作的游戏来产生一种非常真实的游戏体验。

[0002] 一些示例性的游戏平台,如索尼 Playstation 或索尼 Playstation2(PS2),都以游戏控制台的形式销售。众所周知,游戏控制台被设计成与一个监视器(通常是一台电视机)相连接,并能够通过手持式控制器与用户交互。游戏控制台被设计成具有专门的处理硬件,包括一个中央处理器,一个用以处理密集图形运算的图形合成器,一个用以进行几何变换的矢量单元,以及其它配合的硬件,固件和软件。游戏控制台还被进一步设计成具有一用以接收游戏光盘的光盘托架,以通过游戏控制台进行本地游戏。在线游戏也是可能的,用户可以通过因特网交互地与其它用户协同或作为对手进行游戏。

[0003] 由于游戏的复杂性不断激起玩家的兴趣,游戏和硬件制造商不断改进以提供更多的交互性。但事实上,近年来用户与游戏交互的方式并没有很大地改变。

[0004] 考虑到前述问题,现在需要一种方法与系统,以使得用户与游戏之间可以进行更高级的交互。

### 发明内容

[0005] 广义地说,本发明通过提供一促进与计算机程序交互的设备和方法满足了这些需求。在一实施例中,计算机程序是一游戏程序,但是并不仅限于此,该设备和方法可以适用于任何可以接受声音输入以触发控制、输入或进行通信的计算机环境。更具体地,如果声音被用于触发控制或输入,本发明的这些实施例将能够对特定声源进行过滤的输入,并且该过滤的输入被配置为忽略或不理睬那些不关注的声源。在视频游戏环境中,依靠选定的声源,视频游戏能够在处理过关注的声源后以特定的响应应答,而不存在其它可能无关的声音带来的失真或噪音。通常地,一个游戏环境将暴露在许多背景噪声中,如音乐、其他人和物体的移动。一旦这些不关注的声音被充分地滤除,则计算机程序能够更好地应答关注的声音。该响应可以是任何形式,如一个命令、一个动作的起始、一个选择、游戏状况或状态的一种变化、功能(feature)的解锁等等。

[0006] 在一实施例中,提供了一个在与计算机程序的交互中捕捉图像和声音的设备。该设备包括一被配置为捕捉一个或多个图像帧的图像捕捉单元。还包括一声音捕捉单元。该声音捕捉单元被配置为识别一个或多个声源。该声音捕捉单元产生可被解析以确定一个聚焦区域的数据,在此聚焦区域可处理声音以至基本上排除该聚焦区域之外的声音。以此方式,从聚焦区域捕捉和处理的声音被用于与计算机程序的交互。

[0007] 在另一实施例中,揭示了一种在与计算机程序交互期间的选择性声源监听方法。该方法包括以两个或更多声源捕捉麦克风从一个或多个声源接收输入。然后,该方法包括确定每个声源的延迟路径并对于接收到的每个输入识别一个或多个声源中每个声源的方向。该方法然后包括滤除不在所识别的聚焦区域方向上的声源。该聚焦区域被配置成与

算机程序交互提供声源。

[0008] 在另一实施例中,提供了一个游戏系统。该游戏系统包括一个图像-声音捕捉设备,该设备被配置为与能够执行交互式计算机游戏的计算系统相接口。该图像捕捉设备包括能够被置于一定位置以从一聚焦区域捕捉视频的视频捕捉硬件。提供一个麦克风阵列以从一个或多个声源捕捉声音。每个声源都被识别并与一个相对于图像-声音捕捉设备的方向相关联。与视频捕捉硬件相关的聚焦区域被配置为用于识别出声源中的在聚焦区域附近方向上的一个声源。

[0009] 总的来说,交互声音识别和追踪可应用在与任何计算设备的任何计算机程序的接口上。一旦声源被识别,该声源的内容可以被进一步处理,以触发、驱动、指导或控制由一计算机程序所生成的角色或物体。

[0010] 通过下面结合附图、以例子的方式示出本发明的原则的详细描述,本发明的其它方面和优点将变得更加清楚。

### 附图说明

[0011] 通过与附图相结合的下列描述,本发明及其更多的优点可以被更好地理解。

[0012] 图 1 示出了依照本发明一个实施例的一个游戏环境,在该环境中可以执行一个视频游戏程序以与一个或多个用户进行交互。

[0013] 图 2 示出了依照本发明一个实施例的示例性的图像-声音捕捉装置的三维图示。

[0014] 图 3A 和 3B 示出了依照本发明一个实施例的不同麦克风处的声路处理以及用于输出选出的声源的逻辑,所述不同麦克风被设计为接收输入。

[0015] 图 4 示出了依照本发明一个实施例的一个示例性的计算系统,该计算系统与一个图像-声音捕捉装置相接口,以处理输入声源。

[0016] 图 5 示出了依照本发明一个实施例的一个示例,该示例中,多个麦克风被用于增加对特定声源的方向识别的精确性。

[0017] 图 6 示出了依照本发明一个实施例的一个示例,该示例中,使用位于不同平面上的麦克风在一个特定的空间体积内识别声音。

[0018] 图 7 和 8 示出了依照本发明一个实施例的示例性的方法步骤,可以在声源识别以及非聚焦声源的排除中执行这些步骤。

### 具体实施方式

[0019] 本发明揭示了便于识别特定声源并在声音被用作与计算机程序交互的工具时滤除不需要的声源的方法和设备。

[0020] 在以下的描述中,大量的具体细节被阐明以为本发明提供一个透彻的理解。然而,显而易见地是,对于本领域技术人员来说,本发明也可以在缺少其中一些或全部的具体细节的情况下实施。在其它例子中,没有详述公知的处理步骤以免模糊本发明。

[0021] 依照本发明的一个实施例,图 1 示出了一个游戏环境 100,该游戏环境中可以执行一个视频游戏程序以与一个或多个用户交互。如图所示,玩家 102 出现在包含一个显示器 110 的监视器 108 前。该监视器 108 与一个计算系统 104 相互连接。该计算系统可以是一个标准计算机系统,一个游戏控制台或一个便携式计算机系统。在一个具体例子中,游戏控制

台可以是索尼计算机娱乐公司 (Sony Computer Entertainment Inc.)、微软 (Microsoft) 或其它制造商制造的产品,但本发明不限于任何品牌。

[0022] 计算系统 104 被示出与一个图像 - 声音捕捉装置 106 相互连接。该图像 - 声音捕捉装置 106 包括一个声音捕捉单元 106a 和一个图像捕捉单元 106b。玩家 102 被示出正与显示器 110 上的一个游戏人物 112 交互地通信。在正在执行的视频游戏中,其输入至少部分地由玩家 102 经由图像捕捉单元 106b 和声音捕捉单元 106a 提供。如图所示,玩家 102 可以移动他的手以在显示 110 上选择交互图标 114。一旦被图像捕捉单元 106b 捕捉到,一个玩家 102' 的半透明图像将被投射到显示器 110 上。如此,玩家 102 知道将手移动到哪里以进行图标的选择或与游戏人物 112 接口。捕捉这些运动和交互的技术可以不同,但英国专利申请 GB0304024.3 (PCT/GB2004/000693) 和 GB0304022.7 (PCT/GB2004/000703) 中描述了示例性的技术,且二者均提交于 2003 年 2 月 21 日,二者均被引用包括进本申请中。

[0023] 在示出的例子中,交互图标 114 是一个允许玩家选择“挥动”的图标以使游戏人物 112 挥动手中的物体。另外,玩家 102 可以提供语音命令,该命令能够被声音捕捉单元 106a 捕捉,然后由计算系统 104 处理以为正在执行的视频游戏提供交互性。如图所示,声源 116a 是语音命令“跳!”。然后声源 116a 将被声音捕捉单元 106a 捕捉并由计算系统 104 处理,然后促使游戏人物 112 跳起。可以使用语音识别以进行语音命令的识别。作为选择,玩家 102 可以与接入因特网或网络的远程用户进行通讯,但这些远程用户同样直接地或部分地卷入游戏交互中。

[0024] 依照本发明的一个实施例,声音捕捉单元 106a 被配置为包含至少两个麦克风,该麦克风使得计算系统 104 能够选出从特定方向传来的声音。通过使计算系统 104 能够滤除对游戏进行不重要(或非聚焦点)的方向,当玩家 102 发出明确的命令时游戏环境 100 中的杂音就不会干扰或扰乱游戏的执行。例如,游戏玩家 102 可能轻轻跺脚,并引发一个敲击噪音,该敲击噪音是一个非语言的声音 117。这种声音可以被声音捕捉单元 106a 捕捉,但随后被滤除,这是由于从玩家 102 脚部传来的声音不在该视频游戏的聚焦区域内。

[0025] 就像下文中将描述的一样,聚焦区域最好由活动的图像区域标识,该活动的图像区域是图像捕捉单元 106b 的聚焦点。在另一种方式中,在初始化阶段后聚焦区域可以手动地从提供给用户的区域选项中选择。继续图 1 的例子,一个游戏观察者 103 可以产生一个声源 116b,该声源可能在交互游戏进行期间扰乱计算系统的处理。然而,游戏观察者 103 不在图像捕捉单元 106b 的活动的图像区域内,因此从游戏观察者 103 方向传来的声音将被滤除。这样,就像声源 116a 一样,计算系统 104 不会错误地将来自声源 116b 的命令与来自于玩家 102 的声源的命令混淆。

[0026] 图像 - 声音捕捉装置 106 包括一个图像捕捉单元 106b 和声音捕捉单元 106a。该图像 - 声音捕捉装置 106 最好能够数字化地捕捉图像帧并将这些图像帧传输给计算系统 104 以进一步处理。一个图像捕捉单元 106b 的例子是网络摄像机,该摄像机通常被使用于期望捕捉视频图像并将其数字化地传输给一个计算装置以便随后的存储或通过网络(例如因特网)通信的时候。其它类型的图像捕捉装置也可以使用,不管是模拟的还是数字的,只要图像数据被数字化处理以使其能够识别和过滤。在一个优选实施例中,在输入数据被接收后,数字化处理以进行过滤的过程是用软件实现的。声音捕捉单元 106a 被图示为包括一对麦克风(麦克风 1 和麦克风 2)。该对麦克风是标准麦克风,可以与外壳集成为一体以

组成图像 - 声音捕捉装置 106。

[0027] 图 3A 表示声音捕捉单元 106a 面临来自于声音 A 和声音 B 的声源 116。如图所示, 声音 A 射出可闻声并沿着声路 201a 和 201b 被麦克风 1 和麦克风 2 探测到。声音 B 沿着声路 202a 和 202b 射向麦克风 1 和麦克风 2。如图所示, 声音 A 的声路有不同的长度, 因而在比较声路 202a 和 202b 时提供了一个相对的延迟。然后, 来自于声音 A 和声音 B 中每一个的声音将使用一个标准三角剖分算法 (triangulation algorithm) 进行处理, 以使方向选择在 (如图 3B 所示的) 方框 216 中出现。来自于麦克风 1 和麦克风 2 的声音都将被缓存于缓存器 1 和 2 (210a、210b), 并通过延迟线 (212a、212b) 传输。在一个实施例中, 缓存和延时处理可以由软件控制, 但也可定制设计硬件以处理这些操作。基于三角剖分算法, 方向选择 126 将触发识别和选择声源 116 中的一个。

[0028] 来自于麦克风 1 和麦克风 2 中每一个的声音将在方框 214 中被累加, 然后将被作为选择的声源输出被输出。如此, 不是来自于活动图像区域方向的声音被滤除, 以使得那些声源不会扰乱计算系统 104 的处理或者扰乱与其它用户的通信, 而这些用户可通过网络或因特网交互地进行一个视频游戏。

[0029] 图 4 示出了依照本发明一个实施例的一个计算系统 250, 该计算系统 250 可以被用于与图像 - 声音捕捉装置 106 协同工作。该计算系统 250 包括一个处理器 252 和存储器 256。总线 254 将处理器和存储器 256 与图像 - 声音捕捉装置 106 相互连接。存储器 256 包括至少部分的交互程序 258, 并且还包括选择性声源监听逻辑或代码 260 以处理接收到的声源数据。基于图像捕捉单元 106b 确定的聚焦区域, 在聚焦区域之外的声源将通过 (例如: 由处理器和至少部分地存储于存储器 256) 所执行的选择性声源监听逻辑 260 选择性地过滤。该计算系统被图示为其最简单的形式, 但需强调的事实是, 只要硬件可以处理指令以实现输入声源的处理并因而能够进行选择性的监听, 任何硬件配置都可以被使用。

[0030] 计算系统 250 同样被示出通过总线与显示器 110 相互连接。在这个例子中, 聚焦区域被朝向声源 B 聚焦的图像捕捉单元所识别。当声音被声音捕捉单元 106a 捕捉并被传输至计算系统 250 时, 来自于其它声源 (例如声源 A) 的声音, 将基本上被选择性声源监听逻辑 260 滤除。

[0031] 在一个具体例子中, 一个玩家可以与另一个用户参与到因特网或网络视频游戏竞赛中, 网络中每个用户的主要音频体验将由扬声器的方式提供。扬声器可以是计算系统的一部分或是监视器 108 的一部分。因此, 假定本地扬声器产生如图 4 所示的声源 A。为了使本地扬声器作为声源 A 产生的声音不被反馈给该竞赛用户, 选择性声源监听逻辑 260 将滤除声源 A 的声音以阻止竞赛用户的声音或话语的反馈被提供给他或她自己。通过提供这种过滤, 有可能在与视频游戏接口时通过网络进行交互通信, 同时有利地避免这个过程期间的有害反馈。

[0032] 图 5 示出了一个例子, 在此图像 - 声音捕捉装置 106 包括至少四个麦克风 (麦克风 1 至麦克风 4)。因此, 声音捕捉单元 106a 能够以更好的粒度做三角剖分运算以识别出声源 116 (A 和 B) 的位置。也就是说, 通过提供额外的麦克风, 有可能更准确地确定声源的位置并因而排除和滤除不关注的或可能对游戏或对与计算系统进行的交互有害的声源。如图 5 所示, 声源 116 (B) 是由图像捕捉单元 106b 所识别的感兴趣的声源。继续图 5 的例子, 图 6 标识出了声源 B 如何被识别为一空间体积 (spatial volume)。

[0033] 声源 B 所在的空间体积将限定聚焦体积 274。通过识别出一个聚焦体积 (volume of focus), 有可能排除或滤除不在一个具体空间体积内 (即, 那些正好不在一个方向上的) 的噪声。为了便于聚焦体积 274 的选择, 图像 - 声音捕捉装置 106 最好包括至少四个麦克风。至少其中一个麦克风与其它三个麦克风在不同的平面上。在图像 - 声音捕捉装置 106 上, 通过保持四个麦克风中的一个在平面 271 上而剩余的在平面 270 上, 有可能限定一个空间体积。

[0034] 因此, 来自于附近其他人 (如图中示出的 276a 和 276b) 的噪声将被滤除, 这是因为它们不在被聚焦体积 274 定义的空间体积内。另外, 可能由刚好位于空间体积外的如图所示的扬声器 276c 产生的噪声, 也将因其落于空间体积之外而被滤除。

[0035] 图 7 示出了根据本发明一个实施例的流程图。该方法始于步骤 302, 以两个或更多的声音捕捉麦克风接收来自于一个或多个声源的输入。在一个例子中, 两个或更多的声音捕捉麦克风被集成在图像 - 声音捕捉装置 106 中。作为选择, 两个或更多的声音捕捉麦克风可以是与图像捕捉单元 106b 接口的第二模块 / 外壳的一部分。作为选择, 声音捕捉单元 106a 可以包括任何数量的声音捕捉麦克风, 且声音捕捉麦克风被置于特殊的位置, 该特殊的位置被设计为能够捕捉与计算系统接口的用户的语音。

[0036] 该方法进入步骤 304, 在此确定每一个声源的延迟路径。图 3A 中的声路 201 和 202 限定了示例性的延迟路径。众所周知, 延迟路径定义了声波从声源传播到被设置以捕捉该声音的特定麦克风所需的时间。基于声音从特定声源 116 传播的延迟, 使用标准三角剖分算法, 麦克风可以确定延迟是多少以及声音发出的大概位置。

[0037] 该方法然后继续接下来的步骤 306, 在此对于每个接收到的输入识别出一个或多个声源的方向。即识别出源自声源 116 的声音相对于图像 - 声音捕捉装置 (包括声音捕捉单元 106a) 位置的方向。基于识别出的方向, 不在所识别的一个聚焦区域 (或体积) 的方向上的声源都在步骤 308 中被滤除。通过滤除不在聚焦区域附近方向上的声源, 有可能使用未被滤除的声源与计算机程序进行交互, 如步骤 310 所示。

[0038] 例如, 交互程序可以是一视频游戏, 该游戏中用户可以与视频游戏中的角色进行交互通信, 或者与正在和主要玩家对抗的玩家进行交互通信。对抗玩家既可以是本地的也可以是远程的, 并通过网络 (如因特网) 与主要用户进行通信。另外, 视频游戏也可以在一个用户组中的多个用户之间进行, 该用户组是为用户在特定的与游戏有关的竞赛中交互地挑战彼此的技能而设计的。

[0039] 图 8 示出了一个流程图, 在此将图像 - 声音捕捉装置的操作 320 与对所接收的输入进行的软件的操作 340 分开示出。如此, 在步骤 302 中, 一旦以两个或更多的声音捕捉麦克风接收到来自于一个或多个声源的输入, 该方法进入步骤 304, 在此以软件形式确定每个声源的延迟路径。如上所述, 基于延迟路径, 在步骤 306 中, 对于每个接收到的输入, 识别出一个或多个声源中的每个声源的方向。

[0040] 此时, 该方法进入步骤 312, 在此确定所识别的视频捕捉附近的方向。例如, 将视频捕捉针对如图 1 所示的活动图像区域。这样, 视频捕捉附近的区域将在这一活动图像区域 (或体积) 之内, 并且将确定与在此图像活动区域之内或附近的声源相关的任意一个方向。基于这个确定, 该方法进入步骤 314, 在此不在视频捕捉附近的方向 (或体积) 被滤除。因此, 可能扰乱主要玩家玩视频游戏的干扰、噪声和其它无关的输入均在游戏过程中利用软



件进行处理而被滤除。

[0041] 从而,该主要用户可以与视频游戏交互,与正在使用该视频游戏的游戏中其它用户交互,或者与网络上的可能对同一游戏有兴趣并为之进行登陆或进行相关业务的其他用户进行通信。如此,这样的视频游戏通信、交互和控制将不被外部的噪声和 / 或不打算交互通信或参与到一个特定游戏或交互程序中的旁观者所干扰。

[0042] 应当理解,于此描述的实施例同样可以应用于在线游戏应用程序。即,上面描述的实施例可以出现在服务器中,该服务器通过分布式网络,如因特网,向多个用户传送视频信号,使得玩家能够在远端嘈杂的位置相互通信。还应当理解,于此描述的实施例既可以硬件实现也可以软件实现。即,以上讨论的功能性描述可以综合到一个具有一定逻辑的微芯片,该逻辑被配置为对与噪声消除方案相关联的每个模块执行功能性任务。

[0043] 同样,声源的选择性过滤也可以有其它应用,如电话。在电话使用环境中,通常有一个主要人物(例如打电话者),他期望与一个第三方(例如接电话者)交谈。然而,在通信期间,附近可能存在其它说话或制造噪音的人。使电话针对主用户(例如根据话筒的方向),可以使得来自于主用户嘴巴的声音成为聚焦区域,并因而能够选择只听主用户声音。因此,这种选择性监听将使得能够基本上滤除与主用户无关的话语和噪音,并且接收方因而能够从使用电话的主用户处接收一个更加清晰的通信。

[0044] 其它技术也可以包括其它一些电子设备,这些电子装备可以受益于接收声音作为用于控制或通信的输入。例如,一个用户可以使用语音命令控制汽车的设置,同时避免其它路过者干扰这些命令。其它应用可以包括诸如浏览应用软件、文件制作或通信等应用的计算机控制。通过使能这个过滤,有可能更有效地发出不被环境声音干扰的语音或声音命令。同样地,使用任何电子设备均可。

[0045] 进一步地,本发明的实施例具有广泛的一系列的应用,并且权利要求的范围应被解读为包括任何受益于这些实施例的任何应用。

[0046] 例如,在一个类似应用中,可能使用声分析滤除声源。如果声分析被使用,可能使用少达一个的麦克风。被单个麦克风捕捉的声音可以被数字解析(用软件或硬件)以确定哪个语音或声音是所关注的。在一些环境中,如游戏,主用户可能可以将他或她的语音录制一遍以训练系统识别特定的语音。如此,将便于排除其它语音或声音。因此,识别方向并不是必需的,因为过滤可以基于声音的音调和 / 或频率进行。

[0047] 当考虑方向和体积时,上述所有与声音过滤有关的优点是同样可以适用的。

[0048] 考虑到上述实施例,应当理解本发明可以使用涉及计算机系统存储数据的多种计算机实施的步骤。这些步骤包括需要对物理量进行物理操作的步骤。通常,虽然不是必须的,这些物理量以电或磁信号的形式存在,可以被存储、传输、组合、比较以及进行其它操作。进一步地,所进行的操作经常被(作为术语)称为诸如产生、识别、确定或比较等。

[0049] 上述发明可能以其它计算机系统配置实施,包括手持式装置、微处理器系统、基于微处理器的或可编程序的消费电子产品、小型计算机、大型计算机以及类似装置。本发明也可以实施于分布式计算环境中,在该环境中,任务由通过通信网络连接的远程处理装置执行。

[0050] 本发明还可以作为计算机可读代码在一个计算机可读媒介中实现。该计算机可读媒介可以是任何能够存储可以被计算机系统随后阅读的数据的数据存储装置,包括电磁载波。示例性的计算机可读媒介包括硬盘驱动器、网络存储器(NAS)、只读存储器、随机存储

器、CD-ROM、CD-R、CD-RW、磁带和其它光学和非光学数据存储装置。计算机可读媒介也可以分布于一个网络耦合的计算机系统,以使计算机可读代码以分布式方式被存储和执行。

[0051] 虽然为清晰理解的目的对本发明的一些细节进行了详细描述,但显然,可以在后附的权利要求的范围内进行一些变化和修改。因此,本发明的实施例可以被认为是示意性的而不是限制性的,并且本发明不被于此给出的细节所限制,而是可以在后附的权利要求的范围和其等同物内进行更改。

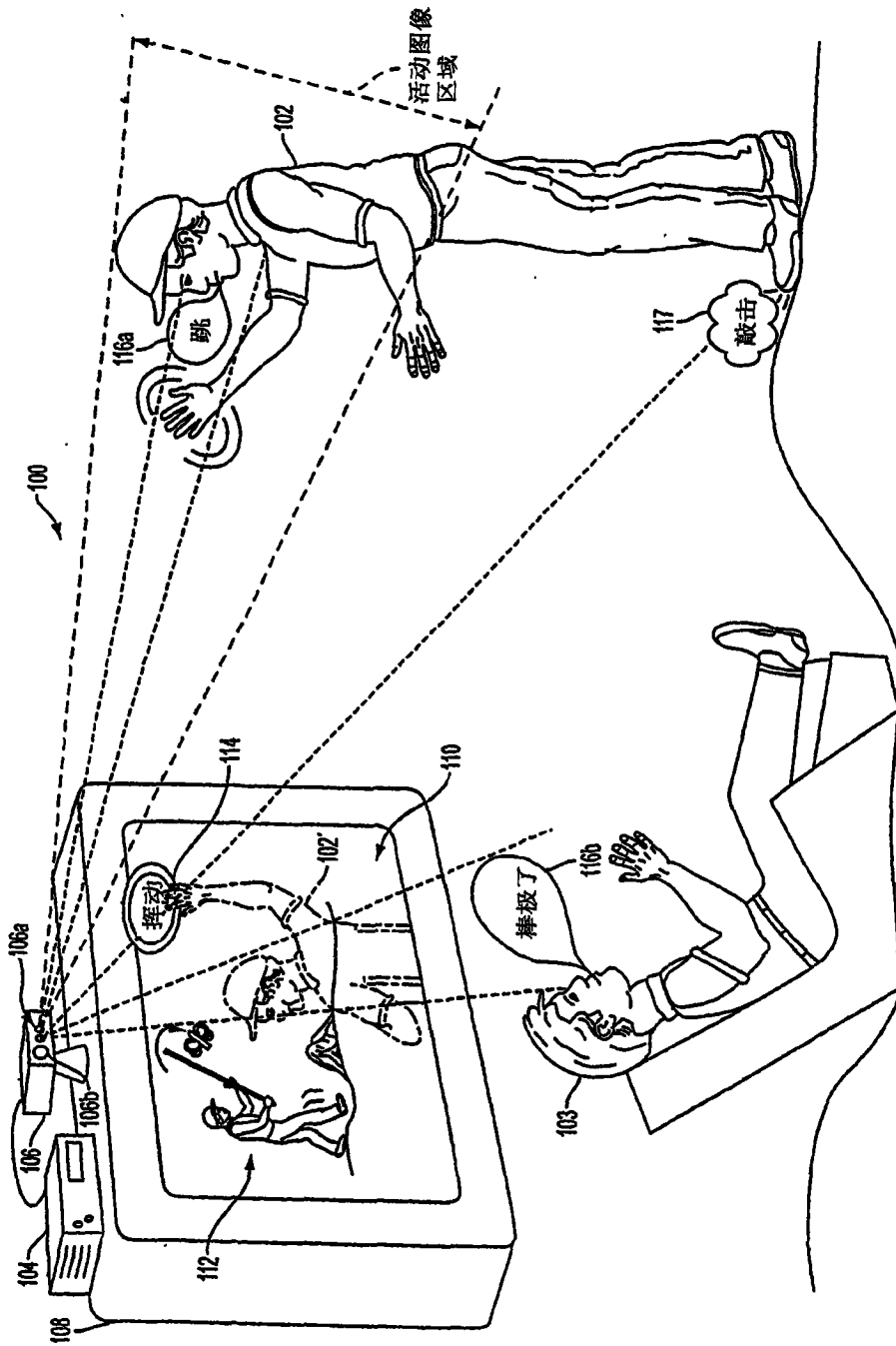


图 1

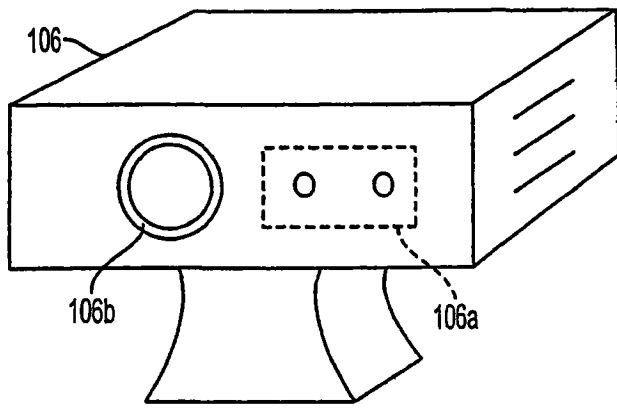


图 2

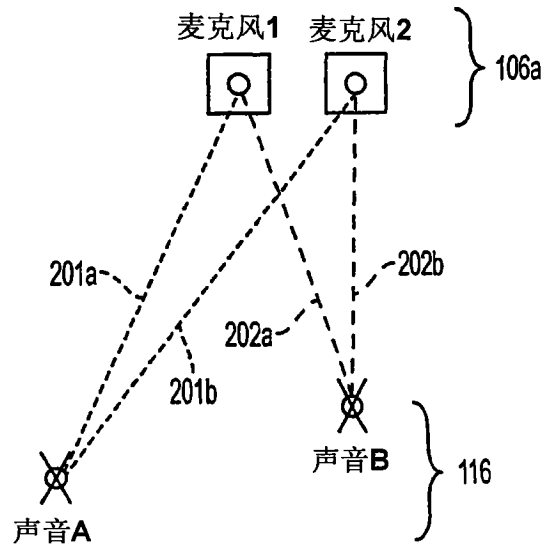


图 3A

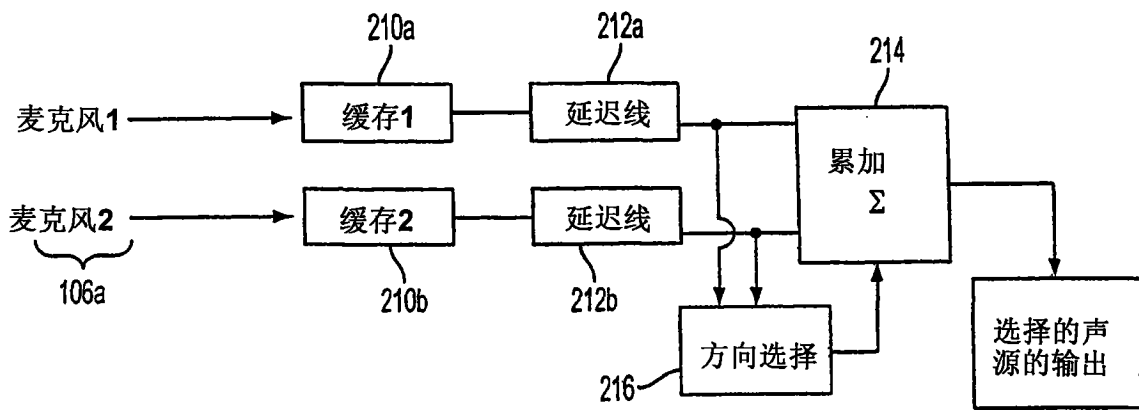


图 3B

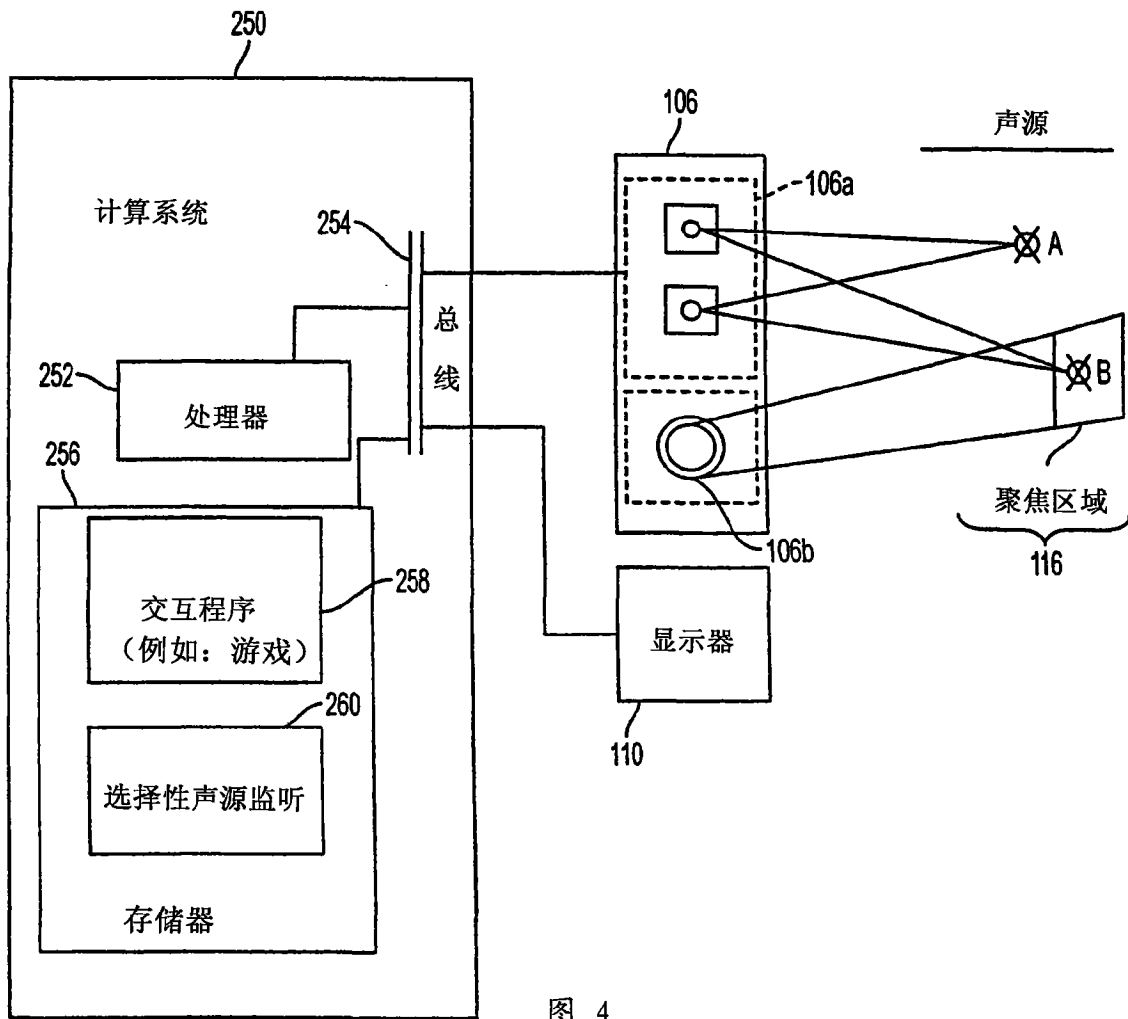


图 4

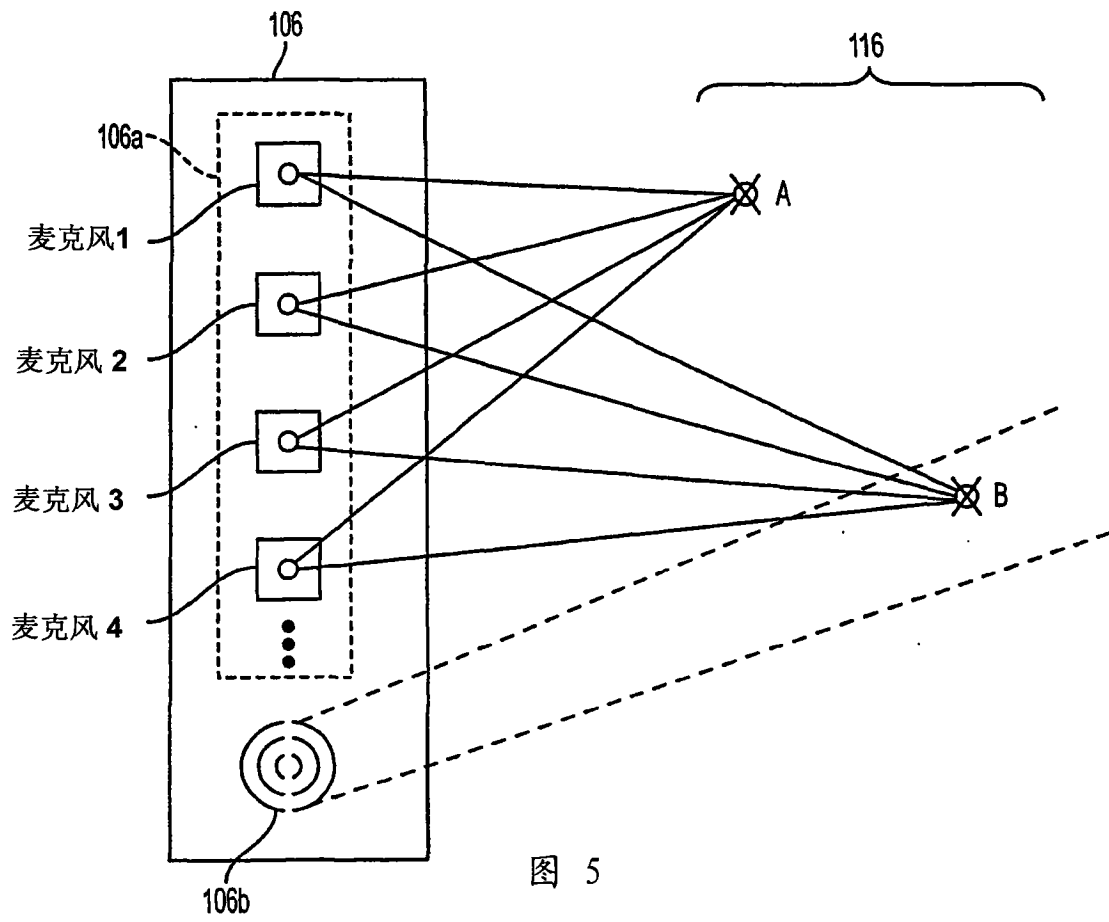


图 5

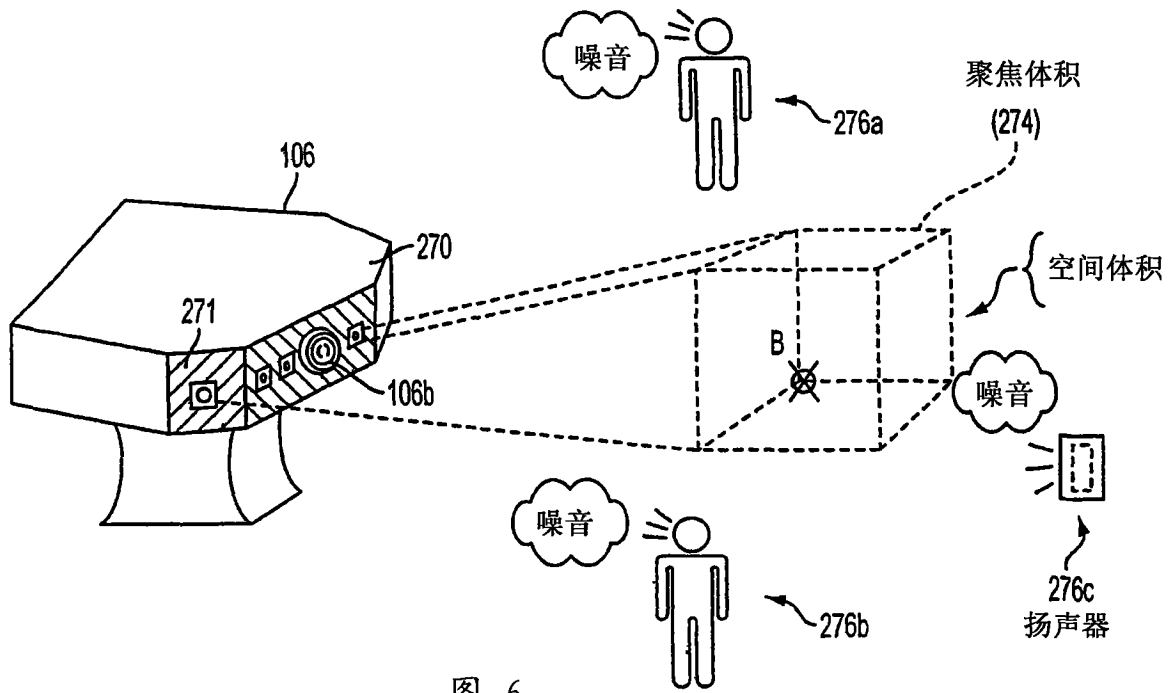


图 6

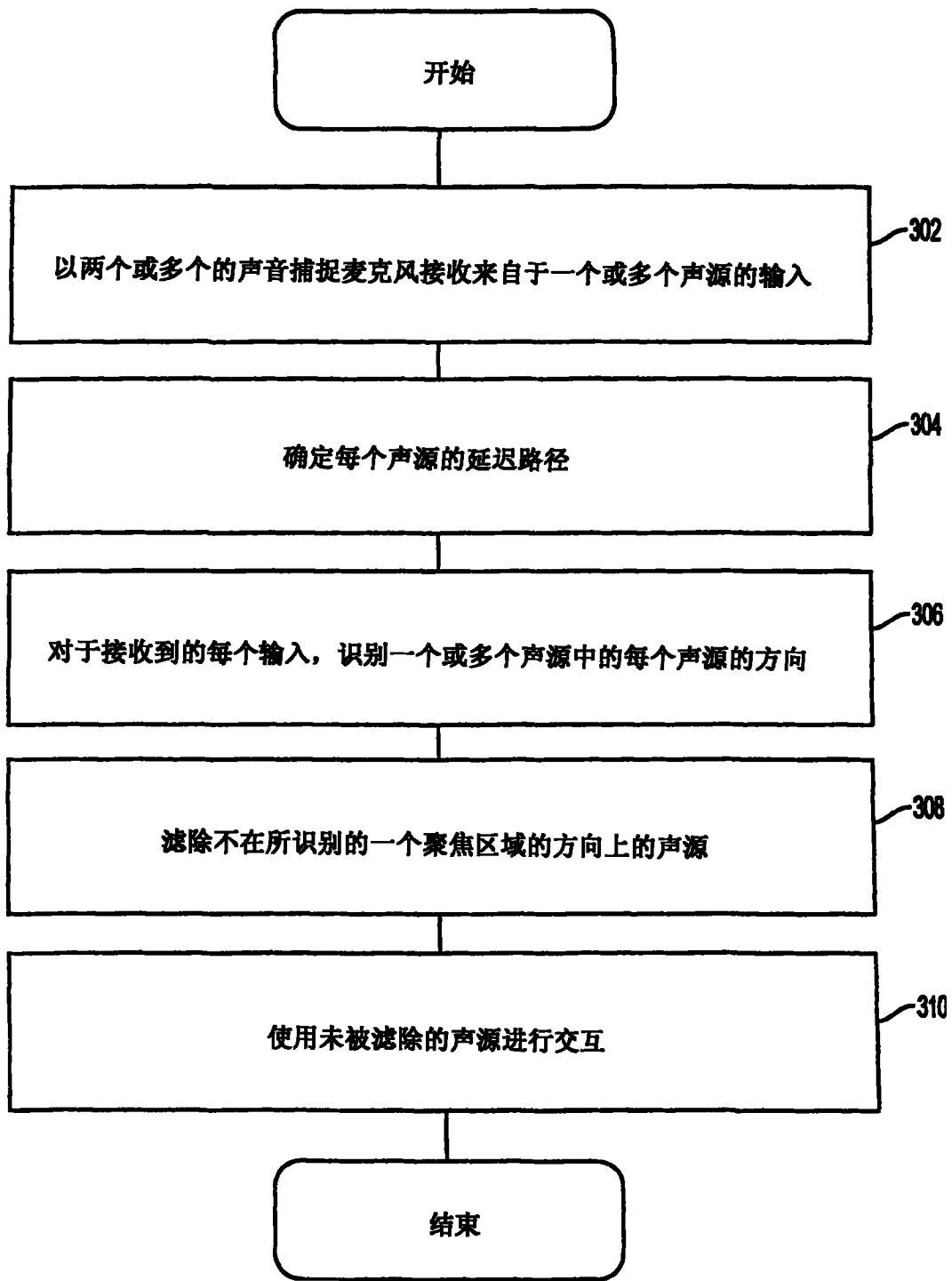


图 7



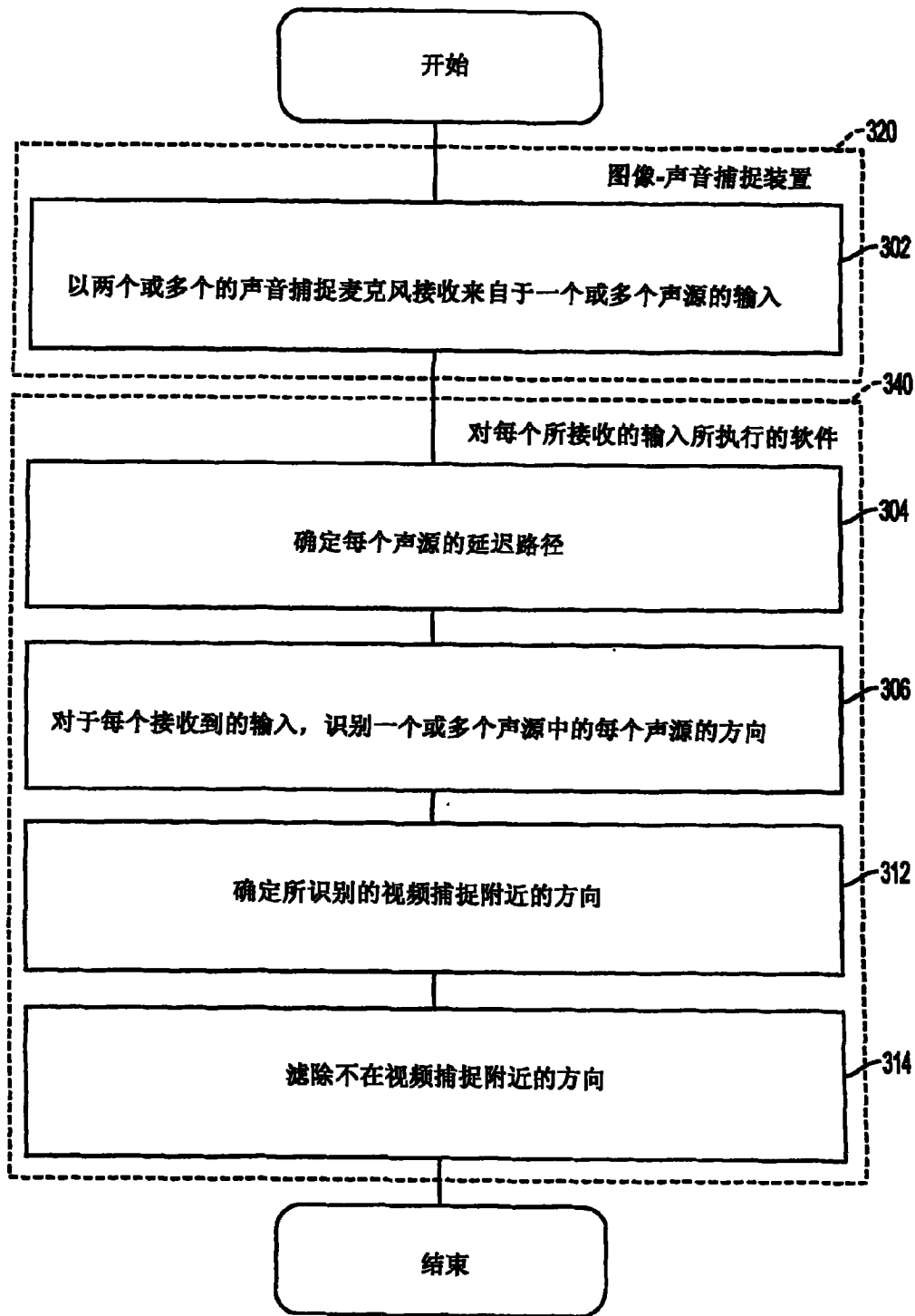


图 8