



(12)发明专利

(10)授权公告号 CN 103827966 B

(45)授权公告日 2018.05.08

(21)申请号 201280043129.X

(22)申请日 2012.07.05

(65)同一申请的已公布的文献号

申请公布号 CN 103827966 A

(43)申请公布日 2014.05.28

(30)优先权数据

1111474.1 2011.07.05 GB

13/212688 2011.08.18 US

(85)PCT国际申请进入国家阶段日

2014.03.05

(86)PCT国际申请的申请数据

PCT/US2012/045556 2012.07.05

(87)PCT国际申请的公布数据

W02013/006700 EN 2013.01.10

(73)专利权人 微软技术许可有限责任公司

地址 美国华盛顿州

(72)发明人 S.斯特罗默 K.V.索伦森

(74)专利代理机构 永新专利商标代理有限公司
72002

代理人 王英

(51)Int.Cl.

G10L 21/0208(2006.01)

H04R 3/00(2006.01)

(56)对比文件

CN 101625871 A, 2010.01.13,

CN 1953059 A, 2007.04.25,

CN 101727909 A, 2010.06.09,

US 2011038489 A1, 2011.02.17,

WO 2010098546 A2, 2010.09.02,

US 2004213419 A1, 2004.10.28,

CN 1753084 A, 2006.03.29,

审查员 李晓惠

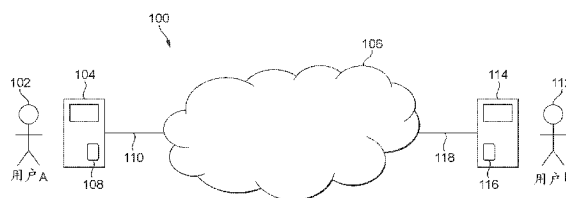
权利要求书2页 说明书10页 附图5页

(54)发明名称

处理音频信号

(57)摘要

描述了一种用于改善移动用户之间的实时视频会话的QoE的计算机实现的系统和方法。例如,根据本发明的一个实施例的方法包括:在服务提供商网络的周界配置一个或多个服务器;从第一移动设备接收与第二移动设备建立实时通信会话的请求;给所述第一和第二移动设备提供用于连接到所述服务器的联网信息;以及通过所述服务器建立所述实时通信会话。



1. 一种在用户设备与远程节点之间的通信会话期间处理音频信号的方法,所述方法包括:

在所述用户设备处的音频输入装置处接收包括至少一个基本音频信号和不需要的信号的多个音频信号;

在噪声抑制装置处接收所述音频信号的到达方向的信息;

将表示所述不需要的信号中的至少一些的已知的到达方向的信息提供给所述噪声抑制装置;以及

在所述噪声抑制装置处处理所述音频信号以基于所述音频信号的到达方向的信息与已知的到达方向的信息之间的比较,将识别为不需要的所述信号的部分对待为噪声。

2. 根据权利要求1的方法,其中所述音频输入装置包括波束形成器,所述波束形成器被布置成:

估计所述至少一个基本音频信号在所述音频输入装置处被从其接收到的至少一个主方向;以及

处理所述多个音频信号以通过在所述至少一个主方向上形成波束并且基本上抑制来自除所述主方向以外的任何方向的音频信号来生成单通道音频输出信号,其中所述单通道音频输出信号包括帧的序列,所述噪声抑制装置处理序列中的所述帧中的每一个。

3. 根据权利要求2的方法,其中针对正被处理的当前帧的主信号分量的信息的到达的方向在所述噪声抑制装置处被接收到,所述方法进一步包括:

比较针对所述当前帧的所述主信号分量的信息的到达的方向和已知的到达方向的信息,其中已知的到达方向的信息包括以下各项中的至少一个:(i) 远端信号在所述音频输入装置处被从其接收到的至少一个方向;(ii) 至少一个已分类的方向,所述至少一个已分类的方向是至少一个不需要的音频信号从其到达所述音频输入装置并且基于所述至少一个不需要的音频信号的信号特性被识别的方向;(iii) 所述至少一个基本音频信号在所述音频输入装置处被从其接收到的至少一个主方向;以及(iv) 所述波束形成器的波束方向图。

4. 根据权利要求3的方法,进一步包括:

基于所述比较来确定所述当前帧的所述主信号分量是否是不需要的信号;

如果确定所述当前帧的所述主信号分量是不需要的信号则对正被处理的所述当前帧应用最大衰减;并且如果发生以下情况则确定所述当前帧的所述主信号分量是不需要的信号:

所述主信号分量在所述音频输入装置处被接收自远端信号在所述音频输入装置处被从其接收到的所述至少一个方向;或者

所述主信号分量在所述音频输入装置处被从所述至少一个已分类的方向接收到;或者所述主信号分量在所述音频输入装置处未被从所述至少一个主方向接收到。

5. 根据权利要求3的方法,进一步包括:

在信号处理装置处接收所述多个音频信号和关于所述至少一个主方向的信息;

在所述信号处理装置处使用关于所述至少一个主方向的所述信息来处理所述多个音频信号以将附加的信息提供给所述噪声抑制装置;以及

取决于所述附加的信息和所述比较对在所述噪声抑制装置处正被处理的所述当前帧应用一定水平的衰减,其中所述附加的信息包括以下各项中的一个:(i) 关于所述当前帧的

所述主信号分量的合意性的指示, (ii) 所述当前帧的所述主信号分量相对于所述至少一个基本音频信号的平均功率电平的功率电平; (iii) 所述当前帧的所述主信号分量的信号分类; 以及 (iv) 所述当前帧的所述主信号分量在所述音频输入装置处被从其接收到的至少一个方向。

6. 根据权利要求3的方法, 进一步包括:

在信号处理装置处接收所述单通道音频输出信号和关于所述至少一个主方向的信息;

在所述信号处理装置处使用关于所述至少一个主方向的所述信息来处理所述单通道音频输出信号以将附加的信息提供给所述噪声抑制装置; 以及

取决于所述附加的信息和所述比较对在所述噪声抑制装置处正被处理的所述当前帧应用一定水平的衰减, 其中所述附加的信息包括以下各项中的一个: (i) 关于所述当前帧的所述主信号分量的合意性的指示, (ii) 所述当前帧的所述主信号分量相对于所述至少一个基本音频信号的平均功率电平的功率电平; (iii) 所述当前帧的所述主信号分量的信号分类; 以及 (iv) 所述当前帧的所述主信号分量在所述音频输入装置处被从其接收到的至少一个方向。

7. 根据权利要求2至6中任一项的方法, 其中所述至少一个主方向通过以下各项来确定:

确定使在所述音频输入装置处正被接收的所述音频信号之间的交叉相关最大化的时间延迟; 以及

用最大交叉相关的所述时间延迟来检测在所述音频输入装置处接收到的所述音频信号中的语音特性。

8. 根据权利要求1至6中的任一项的方法, 其中所述不需要的信号由在所述用户设备外部的源或在所述用户设备处的源来生成, 所述源包括以下各项中的至少一个: 所述用户设备的音频输出装置; 在所述用户设备处的活动的源, 其中所述活动包括点击活动, 所述点击活动包括按钮点击活动、键盘点击活动以及鼠标点击活动。

9. 一种用于在用户设备与远程节点之间的通信会话期间处理音频信号的用户设备, 所述用户设备包括:

音频输入装置, 其用于接收包括至少一个基本音频信号和不需要的信号的多个音频信号; 以及

噪声抑制装置, 其用于接收所述音频信号的到达方向的信息和表示所述不需要的信号中的至少一些的已知的到达方向的信息, 所述噪声抑制装置被配置成通过基于所述音频信号的到达方向的信息与已知的到达方向的信息之间的比较将识别为不需要的所述信号的部分对待为噪声来处理所述音频信号。

10. 一种计算机可读介质, 其包括用于由在用户设备处的计算机处理装置执行以用于在所述用户设备与远程节点之间的通信会话期间处理音频信号的计算机可读指令, 所述指令包括用于完成根据权利要求1至8中任一项的方法的指令。

处理音频信号

技术领域

[0001] 本发明涉及在通信会话期间处理音频信号。

背景技术

[0002] 通信系统允许用户通过网络与彼此进行通信。网络可以是例如因特网或公共交换电话网络 (PSTN)。音频信号能够在网络的节点之间被传送,以便从而允许用户通过通信系统在通信会话中向彼此传送和接收音频数据(诸如,语音数据)。

[0003] 用户设备可以具有音频输入装置,诸如能够被用来接收诸如来自用户的语音之类的音频信号的麦克风。用户可以进入与另一用户的通信会话,诸如私人呼叫(在呼叫中仅有两个用户)或会议呼叫(在呼叫中有两个以上的用户)。用户的语音在麦克风处被接收、处理并且然后通过网络被传送到呼叫中的其他(一个或多个)用户。

[0004] 和来自用户的音频信号一样,麦克风还可以接收可能干扰从用户接收到的音频信号的其他音频信号,诸如背景噪声。

[0005] 用户设备还可以具有音频输出装置,诸如用于将音频信号输出到用户的扬声器,所述音频信号在呼叫期间从(一个或多个)用户通过网络被接收到。然而,扬声器还可以被用来输出来自被在用户设备处执行的其他应用的音频信号。例如,用户设备可以是执行诸如用于通过网络进行通信的通信客户端之类的应用的TV。当用户设备正参与呼叫时,连接到用户设备的麦克风旨在接收由意在传送到呼叫中的其他(一个或多个)用户的用户所提供的语音或其他音频信号。然而,麦克风可以拾取从用户设备的扬声器输出的不需要的音频信号。从用户设备输出的不需要的音频信号可以导致对在麦克风处从用户接收以用于在呼叫中传送的音频信号的干扰。

[0006] 为了改善信号的质量,诸如用于在呼叫中使用,期望抑制在用户设备的音频输入装置处被接收到的不需要的音频信号(背景噪声和从用户设备输出的不需要的音频信号)。

[0007] 立体声麦克风和其中多个麦克风作为单个设备操作的麦克风阵列的使用正变得愈加常见。这些使得除了能够在单个麦克风中实现的东西之外还能够实现提取的空间信息的使用。当使用这种设备时抑制不需要的音频信号的一个办法是应用波束形成器 (beamformer)。波束形成是设法通过应用信号处理以增强来自一个或多个期望的方向的声音来聚焦由麦克风阵列所接收到的信号的过程。为了简单我们将在下文中描述仅具有单个期望的方向的情况,但当存在更多感兴趣方向时相同的方法将适用。波束形成通过首先估计所需信号在麦克风处被从其接收到的角度(所谓的到达方向 (“DOA”) 信息) 来实现。自适应波束形成器使用DOA信息来对来自阵列中的麦克风的信号进行滤波,以便形成在所需信号在麦克风阵列处被从其接收到的方向上具有高增益并且在任何其他方向上具有低增益的波束。

[0008] 虽然波束形成器将试图抑制来自不需要的方向的不需要的音频信号,但是麦克风的数目以及麦克风阵列的形状和尺寸将限制波束形成器的效果,并且结果不需要的音频信号被抑制,但仍然是可听的。

[0009] 对于后续的单通道处理,波束形成器的输出通常被作为输入信号供应给单通道噪声降低级。先前已经提出了实现单通道噪声降低的各种方法。使用中的大多数单通道噪声降低方法是谱相减(spectral subtraction)方法的变体。

[0010] 谱相减方法试图使噪声从语音加噪声信号分离。谱相减牵涉计算语音加噪声信号的功率谱并且获得噪声谱的估计。语音加噪声信号的功率谱与所估计的噪声谱相比较。噪声降低能够例如通过语音加噪声谱的幅度减去噪声谱的幅度来实现。如果语音加噪声信号具有高的信号加噪声与噪声之比(SNNR)则仅非常小的噪声降低被应用。然而如果语音加噪声信号具有低的SNNR,则噪声降低将显著地降低噪声能量。

[0011] 谱相减的问题是它通常使语音失真并且产生临时地且可怕地波动增益改变从而引起常常被称为乐音(musical tone)的一种类型的残留噪声的出现,这可以影响呼叫中已传送的语音质量。不同程度的这个问题同样在实现单通道噪声降低的其他已知方法中发生。

发明内容

[0012] 根据本发明的第一方面提供了在用户设备与远程节点之间的通信会话期间处理音频信号的方法,所述方法包括:在用户设备处的音频输入装置处接收包括至少一个基本音频信号和不需要的信号的多个音频信号;在噪声抑制装置处接收音频信号的到达方向的信息;将表示所述不需要的信号中的至少一些的已知的到达方向的信息提供给噪声抑制装置;以及在噪声抑制装置处处理音频信号以基于音频信号的到达方向的信息与已知的到达方向的信息之间的比较将识别为不需要的信号的部分对待为噪声识别。

[0013] 优选地,音频输入装置包括波束形成器,所述波束形成器被布置成:估计至少一个基本音频信号在音频输入装置处被从其接收到的至少一个主方向;以及处理所述多个音频信号以便通过在至少一个主方向上形成波束并且基本上抑制来自除主方向以外的任何方向的音频信号来生成单通道音频输出信号。

[0014] 优选地,单通道音频输出信号包括帧的序列,噪声抑制装置处理序列中的所述帧中的每一个。

[0015] 优选地,针对正被处理的当前帧的主信号分量的到达方向的信息在噪声抑制装置处被接收到,所述方法进一步包括:比较针对当前帧的主信号分量的信息的到达的方向和已知的到达方向的信息。

[0016] 已知的到达方向的信息包括远端信号在音频输入装置处被从其接收到的至少一个方向。替换地,或者此外,已知的到达方向的信息包括至少一个已分类的方向,所述至少一个已分类的方向是至少一个不需要的音频信号从其到达音频输入装置并且基于所述至少一个不需要的音频信号的信号特性被识别的方向。替换地,或者此外,已知的到达方向的信息包括至少一个基本音频信号在音频输入装置处被从其接收到的至少一个主方向。替换地,或者此外,已知的到达方向的信息进一步包括波束形成器的波束方向图。

[0017] 在一个实施例中,所述方法进一步包括:基于所述比较来确定当前帧的主信号分量是否是不需要的信号;以及如果确定当前帧的主信号分量是不需要的信号,则对正被处理的当前帧应用最大衰减。如果发生以下情况则当前帧的主信号分量可以被确定为不需要的信号:主信号分量在音频输入装置处被接收自远端信号在音频输入装置处被从其接收到

的至少一个方向;或者主信号分量在音频输入装置处被从至少一个已分类的方向接收到;或者主信号分量在音频输入装置处未被从至少一个主方向接收到。

[0018] 所述方法可以进一步包括:在信号处理装置处接收多个音频信号和关于至少一个主方向的信息;使用关于至少一个主方向的所述信息在信号处理装置处处理所述多个音频信号以便将附加的信息提供给噪声抑制装置;以及取决于所述附加的信息和所述比较对在噪声抑制装置处正被处理的当前帧应用一定水平的衰减。

[0019] 替换地,所述方法可以进一步包括:在信号处理装置处接收单通道音频输出信号和关于至少一个主方向的信息;使用关于至少一个主方向的所述信息在信号处理装置处处理单通道音频输出信号以便将附加的信息提供给噪声抑制装置;以及取决于所述附加的信息和所述比较对在噪声抑制装置处正被处理的当前帧应用一定水平的衰减。

[0020] 所述附加的信息可以包括:关于当前帧的主信号分量的合意性(desirability)的指示、或当前帧的主信号分量相对于至少一个基本音频信号的平均功率电平的功率电平、或当前帧的主信号分量的信号分类、或当前帧的主信号分量在音频输入装置处从其被接收到的至少一个方向。

[0021] 优选地,所述至少一个主方向通过以下各项来确定:确定最大化在音频输入装置处正被接收的音频信号之间的交叉相关的时间延迟;以及用最大交叉相关的所述时间延迟来检测在音频输入装置处接收到的音频信号中的语音特性。

[0022] 优选地,在用户设备处从通信会话中的远程节点接收到的音频数据被从用户设备的音频输出装置输出。

[0023] 不需要的信号可以由在用户设备处的源来生成,所述源包括以下各项中的至少一个:用户设备的音频输出装置;在用户设备处的活动的源,其中所述活动包括点击活动,所述点击活动包括按钮点击活动、键盘点击活动以及鼠标点击活动。替换地,不需要的信号由在用户设备外部的源生成。

[0024] 优选地,所述至少一个基本音频信号是在音频输入装置处接收到的语音信号。

[0025] 根据本发明的第二方面提供了用于在用户设备与远程节点之间的通信会话期间处理音频信号的用户设备,所述用户终端包括:音频输入装置,其用于接收包括至少一个基本音频信号和不需要的信号的多个音频信号;以及噪声抑制装置,其用于接收音频信号的到达方向的信息和表示所述不需要的信号中的至少一些的已知的到达方向的信息,所述噪声抑制装置被配置成通过基于音频信号的到达方向的信息与已知的到达方向的信息之间的比较将识别为不需要的信号的部分对待为噪声来处理音频信号识别。

[0026] 根据本发明的第三方面提供了计算机程序产品,所述计算机程序产品包括由在用户设备处的计算机处理装置执行以用于在用户设备与远程节点之间的通信会话期间处理音频信号的计算机可读指令,所述指令包括用于完成根据本发明的第一方面的方法的指令。

[0027] 在以下描述的实施例中,到达方向的信息被用来细化在后续单通道噪声降低方法中要应用多少抑制的决策。因为大部分单通道噪声降低方法具有被应用于输入信号以便确保自然发声但使背景噪声衰减的最大抑制因子,所以到达方向的信息将被用来确保当声音正从除波束形成器集中于的角度外的任何其他角度到达时最大抑制因子被应用。例如,在TV通过与被用于播出远端语音相同的扬声器播出(可能以下降的音量)的情况下,一个问题

是输出将被麦克风拾取。采用本发明的描述的实施例,将检测到的是,音频正从扬声器的角度到达并且除通过波束形成器的所尝试的抑制之外最大噪声降低将被应用。结果,不希望有的信号将是不太可听的,并且因此对远端扬声器干扰较少,并且由于降低的能量它将使用于将信号传送到远端的平均位速率下降。

附图说明

[0028] 为了更好地理解本发明并且以便示出本发明可以如何被付诸实施,现通过例子对以下图进行参考,在图中:

[0029] 图1示出了根据优选实施例的通信系统;

[0030] 图2示出了根据优选实施例的用户终端的示意视图;

[0031] 图3示出了用户终端的示例环境;

[0032] 图4示出了根据一个实施例的在用户终端处的音频输入装置的示意图;

[0033] 图5示出了表示在一个实施例中DOA信息如何被估计的图。

具体实施方式

[0034] 在本发明的以下实施例中,描述了一种技术,其中,不是完全地依赖于波束形成器来使不来自焦点的方向的声音衰减,而是在后续单通道噪声降低方法中使用DOA信息确保来自除波束形成器被集中于的方向外的任何其他方向的声音的最大单通道噪声抑制。当不希望有的信号能够通过使用空间信息从所期望的近端语音信号区分时这是显著的优点。这种源的示例是播放音乐的扩音器、吹风的风扇以及正关闭的门。

[0035] 通过使用信号分类其他源的方向也能够被找到。这种源的示例可以是例如冷却风扇/空调系统、背景中播放的音乐以及键盘敲击。

[0036] 能够采取两个办法:第一,正从某些方向到达的不希望有的源能够被识别并且角度被从其中高于用于最大抑制的噪声抑制增益的噪声抑制增益被允许的角度中排除。例如确保来自某个不希望有的方向的音频的段被按比例缩小得就像该信号仅含有噪声那样将是可能的。在实践中对于这样的段噪声估计能够被设置为等于输入信号,并且因此噪声降低方法然后将应用最大衰减。

[0037] 第二,能够在除我们预期近端语音从其到达的那些方向外的任何其他方向上使得噪声降低对语音不太敏感。也就是说,当根据信号加噪声与噪声之比来计算增益以应用于噪声信号时,作为信号加噪声与噪声之比的函数的增益还将取决于我们将传入语音的角度将看作是多么期望的。对于期望的方向作为给定信号加噪声与噪声之比的函数的增益将高于对于不太期望的方向。第二方法将确保我们不基于移动噪声源来调整,所述移动噪声源不从与(一个或多个)基本扬声器相同的方向到达,并且所述移动噪声源还尚未被检测为噪声的源。

[0038] 本发明的实施例在具有单通道的单道声再现(常常被称为单声道)应用中是特别相关的。立体声应用中的噪声降低(其中存在两个或更多个独立音频通道)典型地不通过独立单通道噪声降低方法来完成,而是通过确保噪声降低方法不使立体图像失真的方法来完成。

[0039] 首先参考图1,图1图示了优选实施例的通信系统100。通信系统的第一用户(用户A

102) 操作用户设备104。用户设备104可以是例如移动电话、电视、个人数字助理 (“PDA”)、个人计算机 (“PC”) (包括例如Windows[™]、Mac OS[™] 以及Linux[™] PC)、游戏设备或能够通过通信系统100进行通信的其他嵌入式设备。

[0040] 用户设备104包括中央处理单元 (CPU) 108,其可以被配置成执行诸如用于通过通信系统100进行通信的通信客户端之类的应用。该应用允许用户设备104通过通信系统100参与呼叫和其他通信会话 (例如,即时消息传递通信会话)。用户设备104能够经由网络106通过通信系统100进行通信,所述网络106可以是例如因特网或公共交换电话网络 (PSTN)。用户设备104能够通过链路110将数据传送到网络106,以及从网络106接收数据。

[0041] 图1同样示出了远程节点,用户设备104能够通过通信系统100与该远程节点进行通信。在图1中所示出的示例中,远程节点是可由第二用户112使用并且包括CPU 116的第二用户设备114,所述CPU 116能够执行应用 (例如,通信客户端) 以便与用户设备104在通信系统100中通过通信网络106进行通信的方式相同的方式通过通信网络106通信。用户设备114可以是例如移动电话、电视、个人数字助理 (“PDA”)、个人计算机 (“PC”) (包括,例如Windows[™]、Mac OS[™] 以及Linux[™] PC)、游戏设备或能够通过通信系统100进行通信的其他嵌入式设备。用户设备114能够通过链路118将数据传送到网络106,以及从网络106接收数据。因此,用户A 102和用户B 112能够通过通信网络106与彼此进行通信。

[0042] 图2图示了客户端在其上被执行的用户终端104的示意视图。用户终端104包括CPU 108,诸如屏幕之类的显示器204、诸如键盘214之类的输入设备以及诸如鼠标212之类的定点设备被连接到所述CPU 108。显示器204可以包括用于将数据输入到CPU 108的触摸屏。输出音频设备206 (例如扬声器) 被连接到CPU 108。诸如麦克风208之类的输入音频设备经由噪声抑制装置227被连接到CPU 108。尽管噪声抑制装置227在图2中被表示为独立式硬件设备,但是噪声抑制装置227能够被实现在软件中。例如噪声抑制装置227能够被包括在客户端中。

[0043] CPU 108被连接到诸如调制解调器之类的网络接口226以用于与网络106通信。

[0044] 现参考图3,图3图示了用户终端104的示例环境300。

[0045] 当在麦克风208处已被接收到的音频信号被处理时期望的音频信号被识别。在处理期间,期望的音频信号基于像质量这样的语音的检测被识别并且主要扬声器的主方向被确定。这在图3中被示出,其中主要扬声器 (用户102) 被示出为从主方向d1到达麦克风208的期望的音频信号的源302。虽然为了简单单个主要扬声器在图3中被示出,但是应领会的是,所需音频信号的任何数目的源可以存在于环境300中。

[0046] 不需要的噪声信号的源可以存在于环境300中。图3示出了可以从方向d3到达麦克风208的环境300中的不需要的噪声信号的噪声源304。不需要的噪声信号的源包括例如冷却风扇、空调系统以及播放音乐的设备。

[0047] 不需要的噪声信号还可以从在用户终端104处的噪声源到达麦克风208,所述噪声源例如鼠标212的点击、键盘214的敲击以及从扬声器206输出的音频信号。图3示出了连接到麦克风208和扬声器206的用户终端104。在图3中,扬声器206是可以从方向d2到达麦克风208的不需要的音频信号的源。

[0048] 虽然麦克风208和扬声器206已经被示出为连接到用户终端的外部设备,但是将领会的是,麦克风208和扬声器206可以被集成到用户终端104中。

[0049] 现参考图4,图4图示了根据一个实施例的麦克风208和噪声抑制装置227的更详细的视图。

[0050] 麦克风208包括包含多个麦克风的麦克风阵列402以及波束形成器404。麦克风阵列402中的每个麦克风的输出被耦合到波束形成器404。本领域的技术人员将领会,为了实现波束形成多个输入是需要的。麦克风阵列402在图4中被示出为具有三个麦克风,但是将理解的是,麦克风的这个数目仅仅是示例并且无论如何不是限制性的。

[0051] 波束形成器404包括从麦克风阵列402接收音频信号的处理块409。处理块409包括话音活动检测器(VAD) 411和DOA估计块413 (其操作稍后将被描述)。处理块409明确由麦克风阵列402所接收到的音频信号的性质,并且基于像由VAD 411所检测到的质量这样的语音的检测和在块413中估计的DOA信息, (一个或多个) 主要扬声器的一个或多个主方向被确定。波束形成器404使用DOA信息通过形成这样的波束来处理音频信号:所述波束在来自所需信号在麦克风阵列处被从其接收到的一个或多个主方向的方向上具有高增益并且在任何其他方向上具有低增益。虽然上面已经描述了处理块409能够确定任何数目的主方向,但是被确定的主方向的数目影响波束形成器的属性,例如在麦克风阵列处从其他(不需要的)方向接收到的信号比如果仅单个主方向被确定的情况衰减得更少。波束形成器404的输出在以待处理的单通道的形式的线路406上被提供给噪声降低级227并且然后给自动增益控制装置(在图4中未示出)。

[0052] 优选地,在增益的电平被自动增益控制装置应用之前噪声抑制被应用于波束形成器的输出。这是因为噪声抑制理论上略微能够降低语音电平(非故意地),并且自动增益控制装置将在噪声抑制之后提高语音电平并且补偿由噪声抑制所引起的语音电平中的略微降低。

[0053] 在波束形成器404中估计的DOA信息被供应给噪声降低级227并且供应给信号处理电路420。

[0054] 在波束形成器404中估计的DOA信息还可以被供应给自动增益控制装置。自动增益控制装置对噪声降低级227的输出应用一定水平的增益。应用于来自噪声降低级227的通道输出的增益的水平取决于在自动增益控制装置处被接收到的DOA信息。自动增益控制装置的操作在英国专利申请No. 1108885.3中被描述并且将不在本文中进一步详细地讨论。

[0055] 噪声降低级227对单通道信号应用噪声降低。噪声降低能够被以许多不同的方式完成,仅作为示例,所述不同的方式包括谱相减(例如,如Boll, S在IEEE会报Acoustics, Speech and Signal Processing, 1979年4月第27卷第2期第113 - 120页的论文“Suppression of acoustic noise in speech using spectral subtraction(使用谱相减对语音中的声学噪声的抑制)”中所描述的那样)。

[0056] 这个技术(以及其他已知技术)抑制识别为噪声的信号的分量以便提高信噪比,其中信号是意中的有用的信号,诸如在这种情况下的语音。

[0057] 如稍后更详细地描述的那样,到达方向的信息在噪声降低级中被用来改善噪声降低并且因此提高信号的质量。

[0058] 现将参考图5更详细地描述DOA估计块413的操作。

[0059] 在DOA估计块413中,DOA信息通过例如使用相关方法来估计在多个麦克风处的接收到的音频信号之间的时间延迟并且使用关于所述多个麦克风的位置的先验知识来估计

音频信号的源而被估计。

[0060] 图5示出了从音频源516接收音频信号的麦克风403和405。能够使用等式(1)估计被分隔开距离d的麦克风403和405处的音频信号的到达的方向：

$$[0061] \quad \theta = \arcsin\left(\frac{v\tau_D}{d}\right) \quad (1)$$

[0062] 其中v是声音的速度,并且 τ_D 是来自源516的音频信号到达麦克风403和405的时间之间的差 - 也就是说,时间延迟。该时间延迟被获得作为使在麦克风403和405的输出处的信号之间的交叉相关最大化的时滞。然后可以找到对应于这个时间延迟的角度 θ 。

[0063] 将领会的是,计算信号的交叉相关是信号处理领域中的常见技术并且将不在本文中更详细地描述。

[0064] 现在在下面更详细地描述噪声降低级227的操作。在本发明的所有实施例中噪声降低级227使用在用户终端处已知的以及由DOA块227所表示的DOA信息并且接收待处理的音频信号。噪声降低级227逐帧地处理音频信号。帧在长度上例如可以是在5与20毫秒之间,并且根据一种噪声抑制技术被划分成谱区间(bin),例如,每帧介于64个与256个区间之间。

[0065] 在噪声降低级227中施行的处理包括对输入到噪声降低级227的音频信号的每个帧应用一定水平的噪声抑制。由噪声降低级227应用于音频信号的每个帧的噪声抑制的水平取决于正被处理的当前帧的所提取的DOA信息与针对在用户终端处已知的各种音频源的DOA信息的组合知识之间的比较。所提取的DOA信息并排地在帧上被传递,使得它被用作除帧自身之外的噪声降低级227的输入参数。

[0066] 由噪声降低级227应用于输入音频信号的噪声抑制的水平可以以许多方式受到DOA信息影响。

[0067] 从已被识别为来自不需要的源的方向到达麦克风208的音频信号可以基于像特性这样的语音的检测被识别并且识别为来自主要扬声器的主方向。

[0068] 在用户终端处已知的DOA信息427可以包括波束形成器的波束方向图。噪声降低级227逐帧地处理音频输入信号。在帧的处理期间,噪声降低级227读取帧的DOA信息以找到该帧中音频信号的主要分量在麦克风208处被从其接收到的角度。帧的DOA信息被与在用户终端处已知的DOA信息427相比较。这个比较确定正被处理的帧中的音频信号的主要分量是否在麦克风208处从所需源的方向被接收到。

[0069] 替换地,或者此外,在用户终端处已知的DOA信息427可以包括以其远端信号在麦克风208处从在用户终端处的扬声器(诸如206)被接收(以供应给噪声降低级227线路407)的角度 ϕ 。

[0070] 替换地,或者此外,在用户终端处已知的DOA信息427可以得自功能425,所述功能425对来自不同方向的音频进行分类以定位可能作为固定噪声源的结果的非常嘈杂的特定方向。

[0071] 当DOA信息427表示主所需方向时,并且通过比较确定正被处理的帧的主要分量在麦克风208处从主方向被接收到。噪声降低级227使用上面所描述的常规方法来确定噪声抑制的水平。

[0072] 在第一办法中,如果确定正被处理的帧的主要分量在麦克风208处从除主方向以外的方向被接收到,则与该帧相关联的区间全部被对待为好像它们是噪声那样(即使正常

的噪声降低技术将识别良好的信号加噪声与噪声之比并且因此不显著地抑制噪声)。这可以通过针对这样的帧将噪声估计设置为等于输入信号完成,并且因此噪声降低级然后将对该帧应用最大衰减。以这种方式,从除所需方向以外的方向到达的帧能够作为噪声被抑制并且信号的质量得以改善。

[0073] 如在上面所提到的那样,噪声降低级227可以从功能425接收DOA信息,所述功能425识别从在不同方向上的(一个或多个)噪声源到达麦克风208的不需要的音频信号。这些不需要的音频信号被从它们的特性识别,例如来自键盘上的键敲击或风扇的音频信号具有不同于人类语音的特性。在高于用于最大抑制的噪声抑制增益的噪声抑制增益被允许的情况下,以其不需要的音频信号到达麦克风208的角度可以被排除。因此当正被处理的帧中的音频信号的主要分量在麦克风208处从排除的方向被接收到时,噪声降低级227对该帧应用最大衰减。

[0074] 可以进一步包括验证装置423。例如,一旦一个或多个主方向已经被检测到(例如在波束形成器的情况下基于波束方向图),客户端就经由客户端用户接口通知用户102所检测到的主方向,并且询问用户102所检测到的主方向是否是正确的。这个验证正如由图4中的虚线所指示的那样是可选的。

[0075] 如果用户102确认所检测到的主方向是正确的,则所检测到的主方向被发送到噪声降低级227并且噪声降低级227如上面所描述的那样操作。通信客户端可以将所检测到的主方向存储在存储器210中,一旦用户102登录到客户端并且已经确认所检测到的主方向是正确的,紧跟后续登录到客户端之后如果检测到的主方向和存储器中的确认的正确主方向匹配,则所检测到的主方向被认为是正确的。这使用户102避免了每当他登录到客户端中时不得不确认主方向。

[0076] 如果用户指示所检测到的主方向是不正确的,则所检测到的主方向不被作为DOA信息发送到噪声降低级227。在这种情况下,基于相关的方法(上面参考图5来描述)就将继续检测主方向并且将仅仅在用户102确认所检测到的主方向正确时发送所检测到的一个或多个主方向。

[0077] 在第一办法中,操作的模式是使得最大衰减能够基于帧的DOA信息被应用于正被处理的帧。

[0078] 在第二办法中,噪声降低级227不在这样的严格的操作模式下操作。

[0079] 在第二办法中,当根据信号加噪声与噪声之比来计算要应用于帧中的音频信号的增益时,作为信号加噪声与噪声之比的函数的增益取决于附加的信息。能够在信号处理块(图4中未示出)中计算这个附加的信息。

[0080] 在第一实施方式中信号处理块可以被实现在麦克风208中。信号处理块(在音频信号已被应用于波束形成器404之前)从麦克风阵列402接收远端音频信号来作为输入,并且同样接收关于从相关方法获得的(一个或多个)主方向的信息。在这个实施方式中,信号处理块将附加的信息输出到噪声降低级227。

[0081] 在第二实施方式中信号处理块可以被实现在噪声降低级227它本身中。信号处理块从波束形成器404接收单通道输出信号来作为输入,并且同样接收关于从相关方法获得的(一个或多个)主方向的信息。在这个实施方式中噪声降低级227可以接收指示扬声器206是活动的信息,并且能够确保在正被处理的帧中的主信号分量仅仅作为噪声被处置,只要

它不同于期望的语音的角度。

[0082] 在这两个实施方式中在信号处理块中计算的附加的信息被噪声降低级227用来根据信号加噪声与噪声之比来计算要应用于正被处理的帧中的音频信号的增益。

[0083] 附加的信息可以包括例如期望的语音将从特定方向/角度到达的可能性。

[0084] 在这个场景中信号处理块提供这样的值作为输出:所述值指示当前正被噪声降低级277处理的帧有多大可能包含噪声降低级应该保留的期望的分量。信号处理块量化传入语音在麦克风208处被从其接收到的角度的合意性。例如如果音频信号在回声期间在麦克风208处被接收到,则以其这些音频信号在麦克风208处被接收到的角度很可能是不希望有的角度,因为不希望保留从在用户终端处的扬声器(诸如206)接收到的任何远端信号。

[0085] 在这个场景中,由噪声降低级227应用于帧的作为信号加噪声与噪声之比的函数的噪声抑制增益取决于合意性的这个量化的量度。对于期望的方向作为给定信号加噪声与噪声之比的函数的增益将比对于不太期望的方向的更高,即较少衰减被噪声降低级227针对较期望的方向来应用。

[0086] 附加的信息可以替换地包括当前帧的主信号分量相对于从所期望的(一个或多个)方向接收到的音频信号的平均功率的功率。在这个场景中,由噪声降低级227应用于帧的作为信号加噪声与噪声之比的函数的噪声抑制增益取决于这个量化的功率比。主信号分量的功率相对于来自主方向的平均功率越接近,由噪声降低级227应用的作为给定信号加噪声与噪声之比的函数的增益越高,即较少衰减被应用。

[0087] 附加的信息可以替换地是提供当前帧的主信号分量的信号分类的信号分类器输出。在这个场景中,噪声降低级227可以对帧应用变化水平的衰减,其中帧的主要分量在麦克风阵列402处被从取决于信号分类器输出的特定方向接收到。因此如果角度被确定为非期望的方向,则噪声降低级227可以较之来自相同的非期望的方向的语音更多地降低来自该非期望的方向的噪声。如果期望的语音被预期从非期望的方向到达则这是可能的并且的确是实用的。然而,它具有严重的缺点,即,噪声将被调制,即当所期望的扬声器是活动的时噪声将是较高的,并且当不希望有的扬声器是活动的时噪声将是较低的。替代地,优选略微降低来自这个方向的信号中的语音的电平。如果通过确信应用相同量的衰减未将它确切地处置为噪声,则通过将它处置为介于期望的语音与噪声之间的某物。这能够通过对于非期望的方向使用略微不同的衰减函数来实现。

[0088] 附加的信息可以替换地是角度它本身,从所述角度当前帧的主信号分量在音频输入装置处被接收到,即在线路407上被供应给噪声降低级227的 ϕ 。随着音频源远离(一个或多个)主方向移动这使得噪声降低级能够应用较多衰减。

[0089] 在这个第二办法中,当噪声降低级227能够在仅将帧处置为噪声和将帧处置为如单通道噪声降低方法中照惯例完成的那样的两个极端之间操作时更多粒度被提供。因此对于从不希望有的方向到达的音频信号,能够使噪声降低级227略微更进取些,而不用完全地处置它好像它只是噪声。也就是说,在我们例如将对语音信号应用一些衰减的意义上是进取的。

[0090] 虽然上面所描述的已经涉及了从单个用户102接收音频信号的麦克风208,但是将理解的是,麦克风可以例如在会议呼叫中从多个用户接收音频信号。在这个场景中所需音频信号的多个源到达麦克风208。

[0091] 虽然已经参考优选实施例具体地示出并且描述了本发明,但是本领域的技术人员将理解,在不背离如由所附权利要求所限定的本发明的范围的情况下可以做出形式和细节上的各种改变。

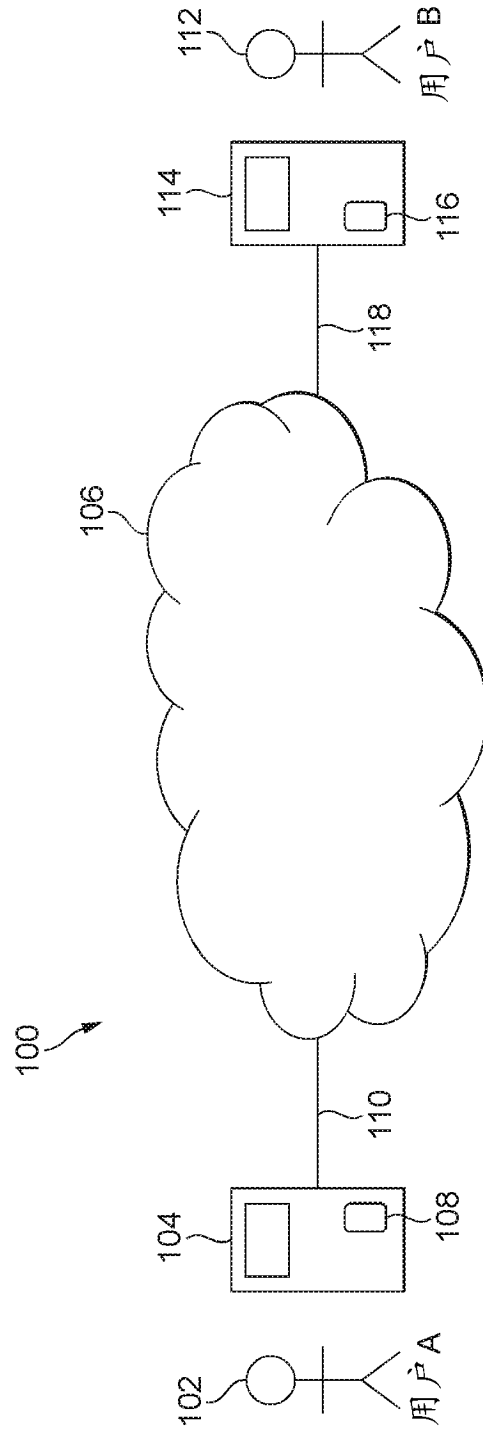


图 1

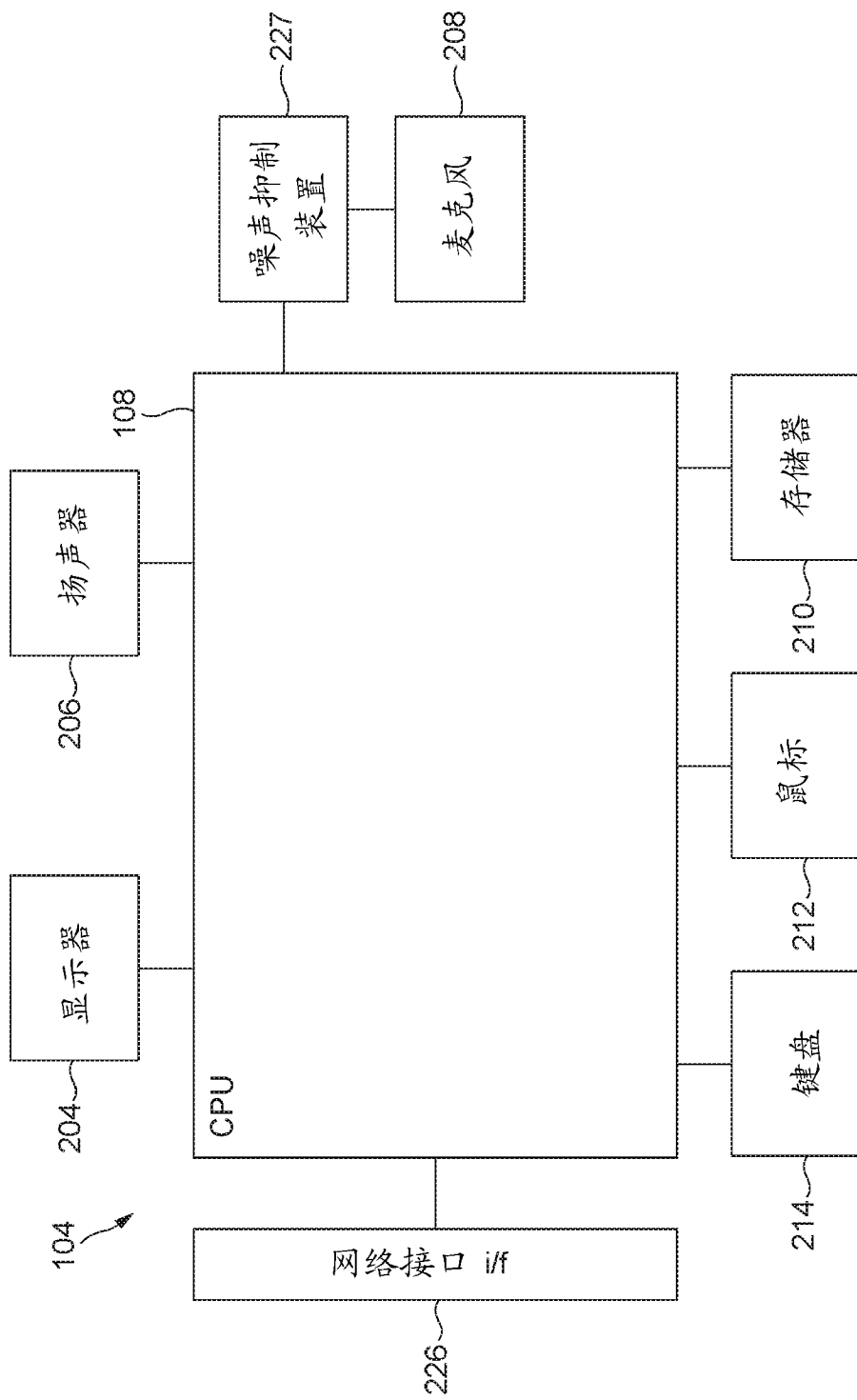


图 2

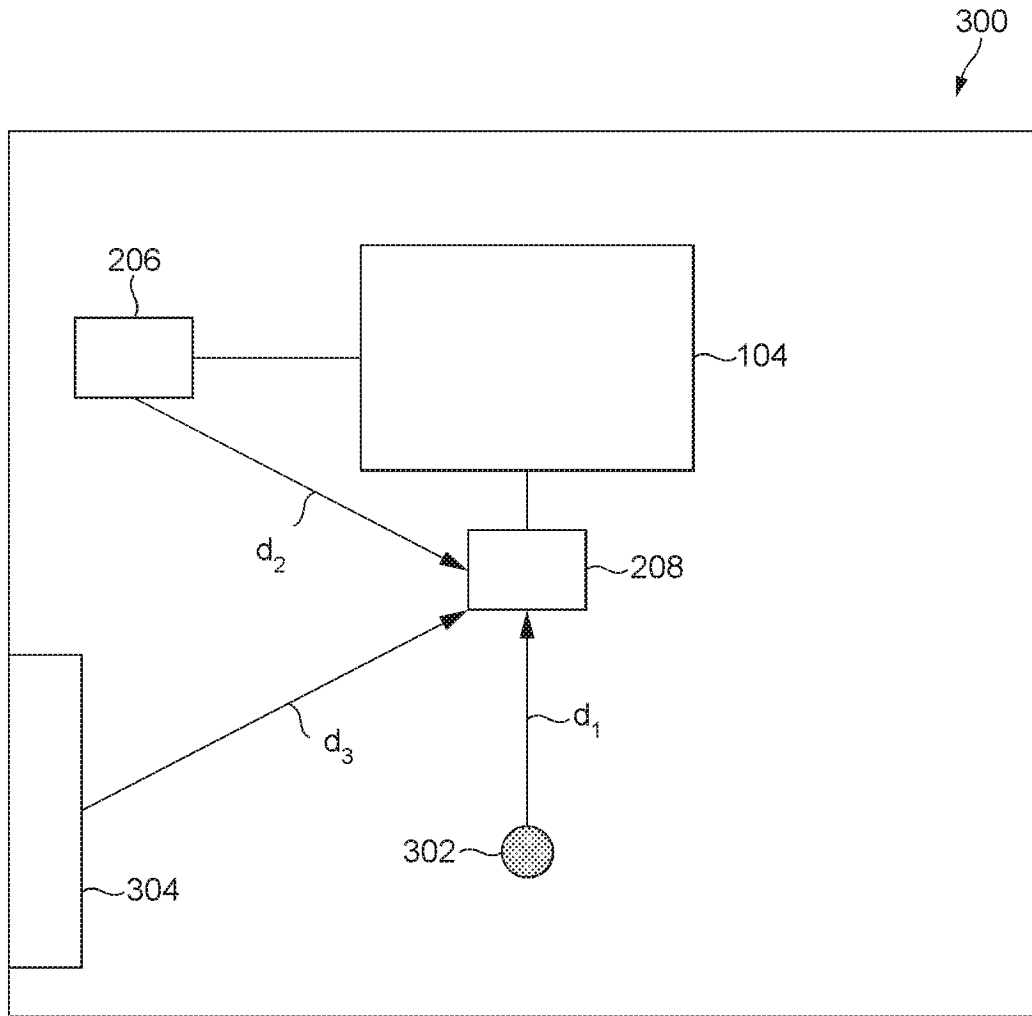


图 3

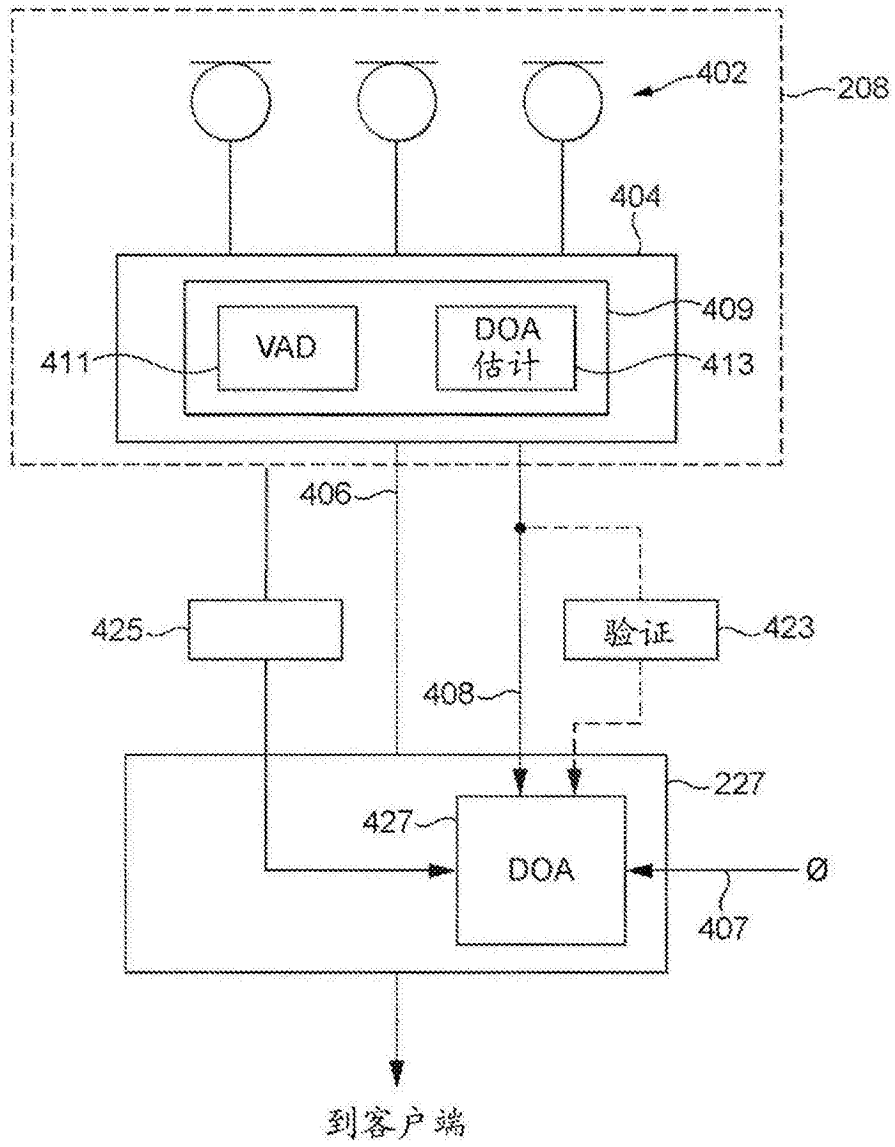


图 4

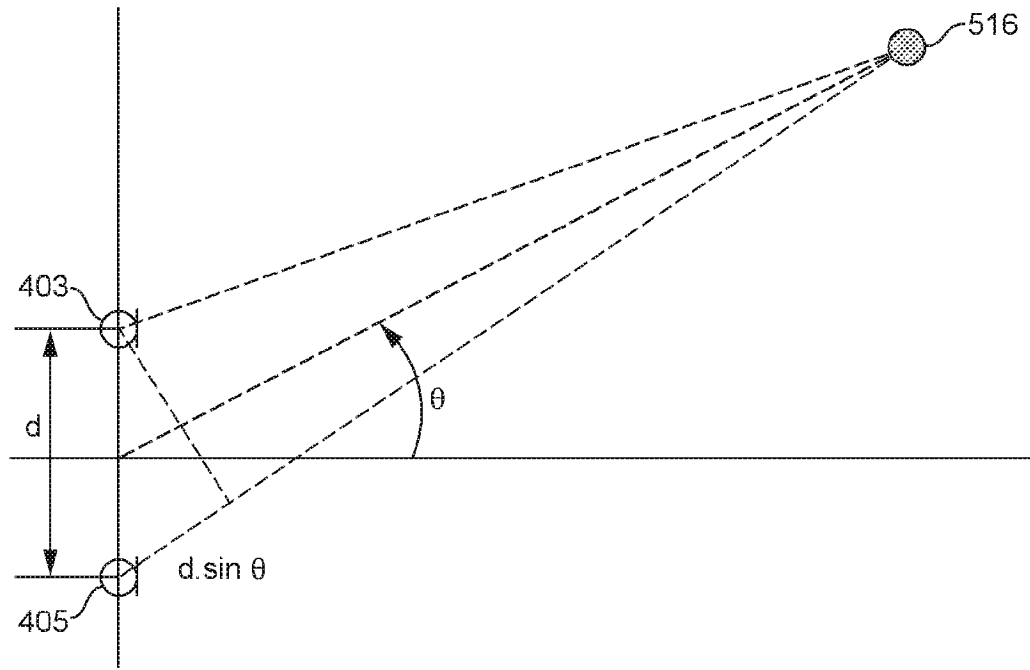


图 5