



(12) 发明专利

(10) 授权公告号 CN 108540912 B

(45) 授权公告日 2021.01.26

(21) 申请号 201810174138.1

(22) 申请日 2018.03.02

(65) 同一申请的已公布的文献号

申请公布号 CN 108540912 A

(43) 申请公布日 2018.09.14

(30) 优先权数据

17158989.8 2017.03.02 EP

(73) 专利权人 大北欧听力公司

地址 丹麦,巴勒鲁普

(72) 发明人 C·瑟伦森 J·B·博尔特

A·克塞纳基 M·G·克里斯坦森

(74) 专利代理机构 北京尚诚知识产权代理有限公司 11322

代理人 顾小曼

(51) Int.Cl.

H04R 25/00 (2006.01)

(56) 对比文件

CN 105872923 A, 2016.08.17

US 2011224976 A1, 2011.09.15

CN 105122843 A, 2015.12.02

Jesper Rindom Jensen等. Statistically efficient methods for pitch and DOA estimation.《2013 IEEE International Conference on Acoustics, Speech and Signal Processing》. IEEE, 2013,

Charlotte Sorensen. Semi-non-intrusive objective intelligibility measure using spatial filtering in hearing aids.《2016 24th European Signal Processing Conference (EUSIPCO)》. IEEE, 2016,

审查员 杨丹

权利要求书2页 说明书12页 附图3页

(54) 发明名称

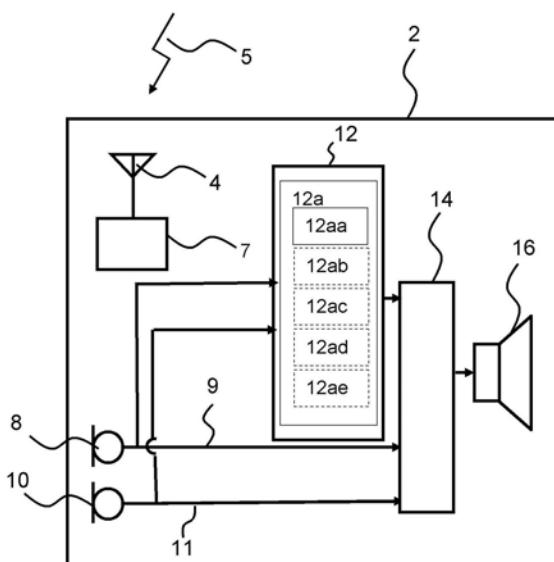
听力设备、方法和听力系统

(57) 摘要

本发明提供听力设备、方法和听力系统。本发明提供了一种听力设备，包括：麦克风组，包括用于提供第一麦克风输入信号的第一麦克风；以及处理器，用于处理输入信号并基于输入信号来提供电输出信号。听力设备包括：接收器，用于将电输出信号转换为音频输出信号；以及控制器，操作地连接至麦克风组，控制器包括语音可识度估计器，用于基于一个或多个麦克风输入信号来估计指示语音可识度的语音可识度指示符。控制器配置为基于语音可识度指示符来控制处理器。语音可识度估计器包括音高估计器，用于估计第一音频源的音高参数。语音可识度指示符基于第一音频源的音高参数和方向。

B

CN 108540912



1.一种听力设备(2),包括:

-用于提供一个或多个麦克风输入信号的麦克风组,包括用于提供第一麦克风输入信号(9)的第一麦克风(8);

-处理器(14),用于处理所述一个或多个麦克风输入信号并基于所述一个或多个麦克风输入信号来提供电输出信号;

-接收器(16),用于将所述电输出信号转换为音频输出信号;以及

-控制器(12),操作地连接至所述麦克风组,所述控制器(12)包括语音可识度估计器(12a),用于基于所述一个或多个麦克风输入信号来估计指示语音可识度的语音可识度指示符,其中,所述控制器(12)配置为基于所述语音可识度指示符来控制所述处理器,

其特征在于,所述语音可识度估计器(12a)包括用于估计第一音频源的音高参数的音高估计器(12aa),并且其中,所述语音可识度指示符基于所述第一音频源的音高参数和方向。

2.根据权利要求1所述的听力设备,其中,所述语音可识度估计器(12a)包括语音合成器,用于基于所述音高参数来生成重建的语音信号,并且其中,所述语音可识度指示符基于所述重建的语音信号。

3.根据权利要求2所述的听力设备,其中,所述语音可识度估计器(12a)包括短时客观可识度估计器,其中,所述短时客观可识度估计器(12ac)配置为比较所述重建的语音信号和基于所述一个或多个麦克风输入信号的基本语音信号并提供所述语音可识度指示符。

4.根据权利要求1至3中任一项所述的听力设备,其中,所述语音可识度估计器(12a)包括谐波模型估计器(12ad),所述谐波模型估计器操作地连接至所述音高估计器,用于提供所述麦克风输入信号的谐波模型参数,并且其中,所述音高参数基于所述谐波模型参数。

5.根据权利要求1所述的听力设备,其中,所述麦克风组包括用于提供第二麦克风输入信号(11)的第二麦克风(10);并且其中,所述语音可识度估计器(12a)包括方向估计器(12ae),用于基于所述第一麦克风输入信号和所述第二麦克风输入信号来估计所述第一音频源的方向,并且其中,所述语音可识度指示符基于所述第一音频源的方向。

6.根据权利要求1所述的听力设备,其中,所述听力设备(2)配置为:如果所述语音可识度指示符满足第一标准,则选择第一处理方案并将所述第一处理方案应用于所述麦克风输入信号。

7.根据权利要求1所述的听力设备,其中,所述听力设备(2)配置为:如果所述语音可识度指示符不满足第一标准,则继续对所述麦克风输入信号应用与之前应用的相同的处理方案。

8.根据权利要求6所述的听力设备,其中,所述第一标准基于第一可识度阈值。

9.根据权利要求8所述的听力设备,其中,当所述语音可识度指示符低于所述第一可识度阈值时,所述语音可识度指示符满足所述第一标准。

10.根据权利要求1所述的听力设备,其中,所述听力设备(2)配置为:如果所述语音可识度指示符满足第二标准,则选择第二处理方案并将所述第二处理方案应用于所述第一麦克风输入信号。

11.根据权利要求6所述的听力设备,其中,所述第一处理方案包括波束成形方案、降噪方案、增益控制方案和压缩方案中的一个或多个。

12. 根据权利要求10所述的听力设备,其中,所述第二处理方案包括波束成形方案、降噪方案、增益控制方案和压缩方案中的一个或多个。

13. 一种操作听力设备的方法,所述方法包括:

-转换(102)音频以提供包括第一麦克风输入信号的一个或多个麦克风输入信号;和

-获得(104)指示至少与所述第一麦克风输入信号相关的语音可识度的语音可识度指示符,其特征在于,获得(104)所述语音可识度指示符包括获得(104a)第一音频源的音高参数,并且其中,所述语音可识度指示符基于所述第一音频源的音高参数和方向;以及

-基于所述语音可识度指示符来控制(106)所述听力设备。

14. 根据权利要求13所述的方法,其中,获得(104a)第一音频源的音高参数包括估计(104aa)所述音高参数,并且其中,获得(104)所述语音可识度指示符包括基于所述一个或多个麦克风输入信号来估计(104b)语音可识度指示符。

15. 根据权利要求13至14中任一项所述的方法,其中,获得(104)所述语音可识度指示符包括:基于所述音高参数来生成(104c)重建的语音信号,并且基于所述重建的语音信号来确定(104d)所述语音可识度指示符。

16. 一种听力系统(200),包括附属设备(202)和根据权利要求1所述的听力设备(2),所述附属设备包括处理器(204)、接口(206)和存储器(208)。

听力设备、方法和听力系统

技术领域

[0001] 本发明涉及一种听力设备、操作听力设备的方法以及听力系统。

背景技术

[0002] 助听器(HA)使用者遇到的主要问题之一是在诸如“鸡尾酒会问题”的嘈杂的多人谈话环境中严重降低的语音可识度(speech intelligibility)。通常，助听设备使用者的语音可识度很大程度上取决于特定的听音环境。如此，语音增强处理方案可能在一些听音环境中是有益的而在其他听音环境中是有害的。语音增强处理方案不一定在任何环境中都改善语音可识度。

发明内容

[0003] 因此，需要一种能够克服背景技术中的缺陷的听力设备、方法和听力系统。

[0004] 公开了一种听力设备，其包括：麦克风组，包括用于提供第一麦克风输入信号的第一麦克风；以及处理器，用于处理输入信号并基于输入信号来提供电输出信号。听力设备包括：接收器，用于将电输出信号转换为音频输出信号；以及控制器，操作地连接至麦克风组，控制器包括语音可识度估计器，用于基于一个或多个麦克风输入信号来估计指示语音可识度的语音可识度指示符。控制器配置为基于语音可识度指示符来控制处理器。语音可识度估计器包括音高(pitch)估计器，用于估计第一音频源的音高参数。语音可识度指示符基于第一音频源的音高参数和方向。

[0005] 此外，本发明涉及一种操作听力设备的方法。方法可以在听力设备或听力系统中执行。方法包括将音频转换为包括第一麦克风输入信号的一个或多个麦克风输入信号。方法包括获得指示与第一麦克风输入信号相关的语音可识度的语音可识度指示符。获得语音可识度指示符可以包括获得第一音频源的音高参数。语音可识度指示符基于第一音频源的音高参数和方向。方法包括基于语音可识度指示符来控制听力设备。

[0006] 本发明涉及一种包括附属设备和听力设备的听力系统，附属设备包括处理器、接口和存储器。听力系统包括：麦克风组，布置在听力设备中，麦克风组包括用于提供第一麦克风输入信号的第一麦克风；以及处理器，用于处理输入信号并基于输入信号来提供电输出信号。听力系统包括：接收器，布置在听力设备中，用于将电输出信号转换为音频输出信号；以及控制器，操作地连接至麦克风组。控制器配置为基于语音可识度指示符来控制处理器。听力系统配置为基于一个或多个麦克风输入信号来估计指示语音可识度的语音可识度指示符。听力系统配置为估计第一音频源的音高参数。语音可识度指示符基于第一音频源的音高参数和方向。

[0007] 本发明的优点在于其允许在没有可用参考语音信号的情况下评估语音可识度。在本发明中有利地使用语音可识度来检测听音环境的状态并相应地适用语音增强方案。具体而言，本发明利用诸如音高的人类语音的基本方面来改善语音可识度估计的准确性。

附图说明

[0008] 通过参考附图对本发明的示例性实施例的以下详细描述,本发明的上述和其他特征和优点对于本领域技术人员将是显而易见的,其中:

[0009] 图1示意性地示出根据本发明的示例性听力设备,

[0010] 图2是根据本发明的示例性方法的流程图,

[0011] 图3示意性地示出根据本发明的示例性听力系统。

具体实施方式

[0012] 以下参考相关附图描述各种示例性实施例和细节。应当注意,附图可以或不可以不按比例绘制,并且在整个附图中,类似结构或功能的元件由类似附图标记表示。还应当理解,附图仅旨在便于实施例的描述。它们不旨在作为对本发明的详尽描述或作为对本发明的范围的限制。另外,所示实施例不需要示出所有方面或优点。结合特定实施例描述的方面或优点不一定限于该实施例,并且即使未示出或未明确描述,也可以在任何其它实施例中实践。

[0013] 发明人已经认识到,听音环境的自动可识度评估对于听力设备的使用者可能是有益的,使得可以基于听音环境的评估的语音可识度来控制语音增强方案,并且因此仅在必要时应用。因此,发明人已经发现在听力设备的处理中使用语音可识度指示符是有益的。为了评估语言可识度,存在用于以可接受的可靠性来预测语音可识度的各种侵入性方法,诸如短时客观可识度(STOI)度量和归一化协方差度量(NCM)。

[0014] 然而,STOI方法和NCM方法是侵入性的,即它们都需要获得作为参考语音信号的“干净的”语音信号。“干净的”语音信号是参考语音信号,其表现出与由音频源发出的信号相似的特性,诸如关于语音可识度的足够的信息。例如,当音频源配备有配偶麦克风设备时,“干净的”语音信号可以由音频源提供。然而,在大部分现实生活场景下,诸如鸡尾酒会,很难获得作为参考语音信号的“干净的”的语音信号。

[0015] EP3 057 335 A1描述了一种双耳系统,包括:左耳和右耳听力设备,适于放置于或放置在使用者的左耳和右耳;以及双耳语音可识度预测单元,用于基于来自相应左耳和右耳听力设备的信号处理单元的处理信号 $y_1(t)$ 、 $y_r(t)$,在暴露于所述输出刺激时,提供使用者的预测语音可识度的双耳SI测量。

[0016] 然而,EP3 057 335 A1的系统也是侵入性的,因为其需要获得来自左耳听力设备的处理语音信号作为参考语音信号。EP3 057 335 A1的双耳系统依赖于来自其中一个听力设备的处理信号在预测语音可识度时用作参考语音信号。这种技术不适用于单耳听力系统,因为不可能从单耳听力系统中的另一听力设备获得处理信号,并且因此,在单耳听力系统中没有轻易可用的参考语音信号。此外,EP3 057 335 A1的系统提供了一种次优的语音可识度估计方案,因为来自例如左耳听力设备的处理信号可能遭受与来自右耳听力设备的信号相同的语音可识度缺陷,并且因此不能被视为可靠的参考语音信号。

[0017] 本发明提供一种听力设备,其在听音环境中通过基于麦克风输入信号和音频源的音高参数估计语音可识度指示符来非侵入式地估计听音环境的语音可识度。本发明提出使用估计的语音可识度指示符来控制麦克风输入信号的处理。

[0018] 本发明的优点在于,本发明中不需要获得参考语音信号来估计语音可识度指示

符。本发明提出一种听力设备、方法和听力系统，其能够基于音高参数和麦克风输入信号来重建参考语音信号（即，表示语音信号的可识度的参考语音信号）。本发明通过利用麦克风输入信号、音高参数和到达方向来克服不可用或不能获得参考语音信号的缺陷。

[0019] 本文公开了一种听力设备。听力设备可以是助听器，其中，处理器配置为补偿使用者的听力损失。听力设备可以是助听器，例如耳后（BTE）式、耳内（ITE）式、入耳（ITC）式、耳道内接收器（RIC）式或耳内接收器（RIE）式。

[0020] 听力设备包括麦克风组。麦克风组可以包括一个或多个麦克风。麦克风组包括用于提供第一麦克风输入信号的第一麦克风和/或用于提供第二麦克风输入信号的第二麦克风。麦克风组可以包括用于提供N个麦克风信号的N个麦克风，其中，N是从1到10范围内的整数。在一个或多个示例性听力设备中，麦克风的数量N是2、3、4、5或以上。麦克风组可以包括用于提供第三麦克风输入信号的第三麦克风。

[0021] 听力设备包括用于处理输入信号（诸如麦克风输入信号）的处理器。处理器基于至处理器的输入信号来提供电输出信号。处理器的输入端子可选地连接至麦克风的相应输出端子。处理器的一个或多个麦克风输入端子可以连接至麦克风的相应的一个或多个麦克风输出端子。处理器可以配置为补偿使用者的听力损失并基于输入信号提供电输出信号。

[0022] 听力设备包括用于将电输出信号转换为音频输出信号的接收器。接收器可以配置为将电输出信号转换为音频输出信号以被引向听力设备使用者的鼓膜。

[0023] 听力设备可选地包括用于将一个或多个无线输入信号（例如，第一无线输入信号和/或第二无线输入信号）转换为天线输出信号的天线。无线输入信号源自外部源，诸如配偶麦克风设备、无线TV音频发射器和/或与无线发射器相关联的分布式麦克风阵列。

[0024] 听力设备可选地包括耦合至天线的无线电收发器，用于将天线输出信号转换为收发器输入信号。来自不同外部源的无线信号可以在无线电收发器中被多路复用为收发器输入信号，或者在无线电收发器的分离的收发器输出端子上被提供为分离的收发器输入信号。听力设备可以包括多个天线和/或天线可以配置为以一个或多个天线模式操作。收发器输入信号包括表示来自第一外部音频源的第一无线信号的第一收发器输入信号。

[0025] 听力设备包括控制器。控制器可以操作地连接至第一麦克风和处理器。控制器可以操作地连接至第二麦克风（如果存在）。控制器可以包括语音可识度估计器，用于基于一个或多个麦克风输入信号来估计指示语音可识度的语音可识度指示符。控制器可以配置为基于一个或多个麦克风输入信号来估计指示语音可识度的语音可识度指示符。控制器配置为基于语音可识度指示符来控制处理器。

[0026] 语音可识度估计器可以包括音高估计器，用于估计第一音频源的音高参数。语音可识度指示符基于第一音频源的音高参数和方向。第一音频源的方向例如是麦克风上接收到的来自第一音频源的麦克风输入信号的到达方向。例如，控制器或语音可识度估计器可以配置为基于第一音频源的音高参数和方向来估计语音可识度指示符。换言之，基于第一音频源的音高参数和方向，由控制器或语音可识度估计器来预测语音可识度指示符。根据本发明，方向可以是已知的（例如，假设面向使用者的鼻子的正面方向）或与音高参数相结合地估计该方向。

[0027] 在一个或多个示例性听力设备中，处理器包括控制器。在一个或多个示例性听力设备中，控制器与处理器并置。

[0028] 在一个或多个示例性听力设备中,语音可识度估计器包括语音合成器,用于基于音高参数来生成重建的语音信号。语音可识度指示符可以基于重建的语音信号。重建的语音信号可以被认为是表示麦克风输入信号的可识度的参考语音信号。换句话说,语音合成器配置为基于音高参数来重建语音信号,并且合成重建的语音信号。例如,控制器或语音可识度估计器可以配置为基于音高参数来生成重建的语音信号,并且基于重建的语音信号、第一音频源的音高参数和方向来估计语音可识度指示符。可以看出,基于合成和重建的语音信号,由控制器或语音可识度估计器来预测语音可识度指示符。重建的语音信号可以被看作重建的参考语音信号。由于所公开的技术解决了例如由于混响或竞争演讲者引起的多义性,所以第一音频源的方向(即,空间提示)和音高参数(即,时间提示)的组合改善了参考语音信号的重建的准确性。

[0029] 在一个或多个示例性听力设备中,语音可识度估计器包括短时客观可识度(STOI)估计器。短时客观可识度估计器可以配置为比较重建的语音信号和基于一个或多个麦克风输入信号的基本语音信号。短时客观可识度估计器可以配置为基于比较来提供语音可识度指示符。基本语音信号是指从一个或多个麦克风获得并提供给接收器的带噪声的语音信号。基本语音信号可以由单个麦克风(其是全向的)或由多个麦克风(例如,使用波束成形)捕获。例如,可以由控制器或语音可识度估计器通过使用STOI估计器比较重建的语音信号和基本语音来预测语音可识度指示符,诸如通过使用STOI估计器比较重建的语音信号和基本语音的相关性。

[0030] 在一个或多个示例性听力设备中,语音可识度估计器包括操作地连接至音高估计器的谐波模型估计器,用于提供麦克风输入信号的谐波模型参数。音高参数可以基于谐波模型参数。例如,谐波模型参数包括基频、采样频率、从第一音频源到给出到达方向的一个或多个麦克风的信号的延迟、来自第一音频源的信号的衰减、振幅(例如,复振幅)、谐波的数量、谐波的实振幅(real amplitude)和/或谐波的相位。

[0031] 在一个或多个示例性听力设备中,谐波模型估计器配置为基于多信道信号的谐波模型结构来提供麦克风输入信号的谐波模型参数,其中,信道对应于麦克风。谐波模型结构可以被看作空间-时间谐波模型结构。第一音频源可以被认为是期望的音频源,并被假设为周期性的。例如,使用空间-时间谐波模型,基于将从第一音频源接收的麦克风输入信号视为具有谐波相关载波频率的多个窄带信号,通过估计信号特性来生成重建的语音。

[0032] 在一个或多个示例性听力设备中,音高估计器配置为接收谐波模型参数并且基于谐波模型参数和对数似然函数来估计音高参数。

[0033] 在一个或多个示例性听力设备中,音高估计器包括最大似然估计器,用于基于对数似然函数和谐波模型参数来估计音高参数。

[0034] 在应用所公开的技术的说明性示例中,应用多信道空间-时间谐波模型以便生成重建的语音信号作为STOI估计器的输入。在示出的示例中,假设K个麦克风用于获得在帧长度为N的情况下添加到干扰源和背景噪声的混合的期望麦克风输入信号。对于第k个麦克风,由麦克风获得的麦克风输入信号可以由 $k=0, \dots, K-1$ 的数据向量 $x_k = [x_k(0) \ x_k(1) \ \dots \ x_k(N-1)]^T$ 表示。假设期望的音频源是周期性的,这对于有声语音的短段是合适的假设。这样,数据向量 x_k 可以被建模为:

$$[0035] x_k = \beta_k ZD(k) \alpha + e_k \quad (1)$$

- [0036] 这里, $Z = [z(\omega_0) \cdots z(L\omega_0)]$, $\mathbf{z}(l\omega_0) = [1 \cdots e^{jl\omega_0(N-1)}]$, $n=0, \dots, N-1$,
- [0037] $\mathbf{D}(k) = \text{diag}([e^{-j\omega_0 f_s \tau_k} \cdots e^{-jL\omega_0 f_s \tau_k}])$, $l=1, \dots, L$, 所有其他项都等于零;
- [0038] e_k 表示记录的噪声和干扰的总和;
- [0039] ω_0 是基频, f_s 是采样频率;
- [0040] T_k 是麦克风 0 和给出到达方向 (DOA) 的第 k 个麦克风之间的输入信号的延迟;
- [0041] β_k 是第 k 个麦克风处麦克风输入信号的衰减,
- [0042] $a = [a_1 \cdots a_L]^T$ 表示第 1 个复振幅 $\alpha = A_l e^{j\varphi_l}$ 给出的复振幅;
- [0043] L 是谐波的数量;
- [0044] $A_l > 0$ 并且 φ_l 分别是第 1 次谐波的实振幅和相位。
- [0045] 在应用所公开的技术的说明性示例中, 假设噪声是在每个麦克风或每个麦克风信道中具有方差 σ_k^2 的不相关高斯白噪声, 则复数据向量 x_k 的对数似然函数可以写为:
- [0046]
$$\ln p(x_k; \psi) = -NK \ln \pi - N \sum_{k=0}^{K-1} \ln \sigma_k^2 - \sum_{k=0}^{K-1} \frac{\|e_k\|^2}{\sigma_k^2} \quad (2)$$
- [0047] 这里, ψ 表示包含用于 x_k 的信号参数的向量。在本示例中, 高斯白噪声分布使噪声的熵最大化, 并且因此是噪声概率密度函数的一个很好的选择。可以通过分别关于振幅 $\hat{\alpha}$ 、衰减因子 β_k 和噪声方差 σ_k^2 进行微分来最大化对数似然函数, 从而估计音高。这些参数是互相依赖的, 并且因此通过将 β_k 和 σ_k^2 初始设定为 1 (initially setting β_k 's and σ_k^2 's to 1) 并迭代公式 (3)、(4) 和 (5) 中的表达式来估计这些参数。估计的复振幅由下式给出:
- [0048]
$$\hat{\alpha} = \left[\sum_{k=0}^{K-1} \frac{\beta_k^2}{\sigma_k^2} D^H(k) \mathbf{Z}^H \mathbf{Z} \mathbf{D}(k) \right]^{-1} \sum_{k=0}^{K-1} \frac{\beta_k^2}{\sigma_k^2} D^H(k) \mathbf{Z}^H \mathbf{x}_k \quad (3)$$
- [0049] 第 k 个麦克风处期望的音频源的估计衰减可以如下获得:
- [0050]
$$\hat{\beta}_k = \frac{\text{Re}\{\alpha^H \mathbf{D}^H(k) \mathbf{Z}^H \mathbf{x}_k\}}{\alpha^H \mathbf{D}^H(k) \mathbf{Z}^H \mathbf{Z} \mathbf{D}(k) \alpha} \quad (4)$$
- [0051] 此外, 噪声方差可以如下所示:
- [0052]
$$\hat{\sigma}_k^2 = N^{-1} \|\mathbf{x}_k - \hat{\beta}_k \mathbf{Z} \mathbf{D}(k) \hat{\alpha}\|^2 \quad (5)$$
- [0053] 这里, $\hat{e}_k = \mathbf{x}_k - \hat{\beta}_k \mathbf{Z} \mathbf{D}(k) \hat{\alpha}$ 。
- [0054] 然后可以使用最大似然估计器来估计音高参数, 记为:
- [0055]
$$\hat{\omega}_0 = \arg \min_{\omega_0 \in \Omega_0} \sum_{k=0}^{K-1} \ln \|\mathbf{x}_k - \hat{\beta}_k \mathbf{Z} \mathbf{D}(k) \hat{\alpha}\|^2 \quad (6)$$
- [0056] 这里, Ω_0 是可能的音高参数候选的集合。在本示例中, 第一音频源的方向被假设为已知的, 使得在一维搜索上执行音高参数的估计。这附加地限制了计算的复杂度, 并且提供了一种对来自其他方向的较强的干扰谐波源具有鲁棒性的技术。鉴于估计的音高 ω_0 和延迟 T_k , 可以获得第 k 个麦克风的重建的语音信号:
- [0057]
$$\hat{s}_k = \prod_{\mathbf{Z} \mathbf{D}(k)} \mathbf{x}_k \quad (7)$$

[0058] 其中,投影矩阵为 $\Pi_A = A(A^H A)^{-1} A^H$ 。然后通过在所有麦克风信道上对重建的语音信号进行求和来获得要被用作短时客观可识度估计器的输入的重建的语音信号:

$$[0059] \hat{\mathbf{s}}_k = \frac{1}{K} \sum_{k=0}^{K-1} \hat{\mathbf{s}}_k \quad (8)$$

[0060] 替代地或附加地,公式(5)中的方差估计可以用于形成重建的语音信号的加权估计。

[0061] 在一个或多个示例性听力设备中,麦克风组包括用于提供第二麦克风输入信号的第二麦克风。语音可识度估计器可以包括方向估计器,用于基于第一麦克风输入信号和第二麦克风输入信号来估计第一音频源的方向。语音可识度指示符可以基于第一音频源的方向。例如,通过利用由一个或多个麦克风接收到的期望的周期性麦克风输入信号的空间-时间谐波模型,在语音可识度估计器处或在控制器处联合估计第一音频源的音高参数和方向(例如,麦克风输入信号的到达方向)。

[0062] 在一个或多个示例性听力设备中,听力设备配置为:如果语音可识度指示符满足第一标准,则选择第一处理方案并将其应用于麦克风输入信号。例如,当语音可识度指示符满足第一标准(例如,语音可识度指示符指示语音可识度不足)时,需要将第一处理方案应用于麦克风输入信号以改善语音可识度。第一处理方案可以包括一个或多个语音增强处理方案。第一处理方案可以包括配置为补偿使用者听力损失的一个或多个语音增强处理。例如,控制器可以将语音可识度指示符提供给处理器,该处理器可以配置为:如果语音可识度指示符满足第一标准,则选择第一处理方案并将其应用于麦克风输入信号。在一个或多个示例性听力设备中,控制器可以配置为:如果语音可识度指示符满足第一标准,则选择第一处理方案并将其应用于麦克风输入信号。

[0063] 在一个或多个示例性听力设备中,听力设备配置为:如果语音可识度指示符不满足第一标准,则继续向麦克风输入信号应用与先前应用的相同的处理方案。例如,当语音可识度指示符不满足第一标准(即,语音可识度指示符指示语音可识度足够)时,与先前应用的相同的处理方案不需要改变,并且听力设备或处理器可以继续对麦克风输入信号应用相同处理方案。

[0064] 第一标准可以基于第一可识度阈值。在一个或多个示例性听力设备中,当语音可识度指示符低于第一可识度阈值时,语音可识度指示符满足第一标准。例如,第一可识度阈值可以在75-95%的范围内,诸如80%、85%。例如,当语音可识度指示符低于85%时,听力设备选择并应用第一处理方案,诸如波束成形方案。当语音可识度指示符等于或高于85%时,听力设备继续应用与先前应用的处理方案相同的方案。

[0065] 在一个或多个示例性听力设备中,听力设备配置为:如果语音可识度指示符满足第二标准,则选择第二处理方案并将其应用于第一麦克风输入信号。第二标准可以基于第二可识度阈值和第三可识度阈值。例如,第二可识度阈值可以在60-75%的范围内,诸如65%、70%。例如,第三可识度阈值可以在77-84%的范围内,诸如80%、84%。例如,当语音可识度指示符落在第二和第三可识度阈值之间时,语音可识度指示符满足第二标准。例如,当语音可识度指示符等于或大于70%但不大于80%时,语音可识度指示符满足第二标准。另一示例是当语音可识度指示符满足第三标准时,其中,第三标准基于零阈值。例如,当语音可识度指示符达到0%时,选择并应用较窄的波束成形(具有较高的指向性指数)。

[0066] 可以设想,听力设备可以使用各种标准(例如第一标准、第二标准、第三标准、第四标准等)来识别哪个处理方案将被选择并应用于麦克风输入信号。

[0067] 第一和/或第二处理方案可以包括波束成形方案、降噪方案、增益控制方案和压缩方案中的一个或多个。例如,由听力设备基于语音可识度指示符选择的第一处理方案可以包括波束成形方案和降噪方案的组合。

[0068] 在一个或多个示例性听力设备中,第一和/或第二处理方案可以包括具有第一组波束成形系数的第一波束成形方案和/或具有第二组波束成形系数的第二波束成形方案。

[0069] 在一个或多个示例性听力设备中,第一和/或第二处理方案可以包括提供一个或多个降噪功能的降噪方案,从而获得改善的信噪比。

[0070] 在一个或多个示例性听力设备中,第一和/或第二处理方案可以包括压缩方案和/或增益控制方案,其中,基于听力损失补偿来控制应用于麦克风输入信号的增益。本发明涉及一种操作听力设备的方法。方法可以在听力设备或听力系统中执行。方法包括将音频转换为包括第一麦克风输入信号的一个或多个麦克风输入信号。方法包括获得指示与第一麦克风输入信号相关的语音可识度的语音可识度指示符。获得语音可识度指示符包括获得第一音频源的音高参数。语音可识度指示符基于第一音频源的音高参数和方向。方法包括基于语音可识度指示符来控制听力设备。

[0071] 在一个或多个示例性方法中,获得第一音频源的音高参数包括估计音高参数。获得语音可识度指示符可以包括基于一个或多个麦克风输入信号来估计语音可识度指示符。

[0072] 在一个或多个示例性方法中,获得语音可识度指示符包括基于音高参数来生成重建的语音信号,并且基于重建的语音信号来确定语音可识度指示符。获得语音可识度指示符可以包括比较重建的语音信号和基本语音信号,例如,使用短时客观可识度估计器。获得语音可识度指示符可以包括获得麦克风输入信号的谐波模型参数并且基于谐波模型参数来导出音高参数。获得语音可识度指示符可以包括基于麦克风输入信号来估计第一音频源的方向。

[0073] 在一个或多个示例性方法中,基于语音可识度指示符来控制听力设备包括:如果可识度指示符满足第一标准,则选择第一处理方案并将其应用于麦克风输入信号。第一标准可以基于第一可识度阈值。第一处理方案可以包括波束成形方案、降噪方案、增益控制方案和压缩方案中的一个或多个。

[0074] 在一个或多个示例性方法中,获得第一音频源的音高参数可以包括从外部设备(诸如附属设备)接收音高参数。可以设想,在外部设备处(例如在线)执行语音可识度指示符的估计,其中,听力设备配置为将麦克风输入信号提供给外部设备以便获得语音可识度回馈。

[0075] 可以设想,在外部设备处(例如在线)执行语音可识度指示符的估计和根据语音可识度指示符的麦克风输入信号的处理,其中,听力设备配置为将麦克风输入信号提供给外部设备以便获得经处理的(例如,降噪的)麦克风输入信号或波束成形参数。

[0076] 本发明涉及一种包括附属设备和听力设备的听力系统,附属设备包括处理器、接口和存储器。听力系统包括:麦克风组,布置在听力设备中,麦克风组包括用于提供第一麦克风输入信号的第一麦克风;以及处理器,用于处理输入信号并基于输入信号来提供电输出信号。听力系统包括:接收器,布置在听力设备中,用于将电输出信号转换为音频输出信

号；以及控制器，操作地连接至麦克风组。控制器配置为基于语音可识度指示符来控制处理器。听力系统配置为基于一个或多个麦克风输入信号来估计指示语音可识度的语音可识度指示符。听力系统配置为估计第一音频源的音高参数。语音可识度指示符基于第一音频源的音高参数和方向。

[0077] 附属设备可以被视为听力设备的附件。附属设备可以配对或以其他方式无线耦合至听力设备。听力系统可以由听力设备使用者拥有和控制。附属设备可以是智能电话、智能手表或平板计算机。

[0078] 在一个或多个示例性听力系统中，附属设备包括由听力设备的处理器远程访问的控制器，并且从听力设备远程（例如在附属设备处）执行语音可识度估计。

[0079] 本文公开的听力设备、系统和方法允许预测语音可识度指示符，并且允许根据预测的语音可识度指示符而应用于输入信号的处理的自适应。

[0080] 为了清楚起见，附图是示意性的和简化的，并且它们仅仅示出对于理解本发明所必需的细节，而其余细节已被省略。在整个附图中，相同的附图标记用于相同或对应的部分。

[0081] 图1是根据本发明的示例性听力设备2的框图。

[0082] 听力设备2包括麦克风组。麦克风组可以包括一个或多个麦克风。麦克风组包括用于提供第一麦克风输入信号9的第一麦克风8和/或用于提供第二麦克风输入信号11的第二麦克风10。

[0083] 听力设备2可选地包括用于将第一外部源（图1中未示出）的第一无线输入信号5转换为天线输出信号的天线4。听力设备2可选地包括耦合至天线4的无线电收发器7，用于将天线输出信号转换为一个或多个收发器输入信号，并且该无线电收发器耦合至包括第一麦克风8和可选的第二麦克风10的麦克风组，用于提供相应的第一麦克风输入信号9和第二麦克风输入信号11。

[0084] 听力设备2包括用于处理输入信号（诸如麦克风输入信号）的处理器14。处理器14基于至处理器14的输入信号来提供电输出信号。

[0085] 听力设备包括用于将电输出信号转换为音频输出信号的接收器16。

[0086] 听力设备包括控制器12。控制器12操作地连接至第一麦克风8和处理器16。控制器12可以操作地连接至第二麦克风10。控制器12配置为基于一个或多个麦克风输入信号来估计指示语音可识度的语音可识度指示符。控制器12包括语音可识度估计器12a，用于基于一个或多个麦克风输入信号来估计指示语音可识度的语音可识度指示符。控制器12配置为基于语音可识度指示符来控制处理器14。

[0087] 语音可识度估计器12a包括音高估计器12aa，用于估计第一音频源的音高参数。语音可识度指示符基于第一音频源的音高参数和方向。语音可识度估计器12a配置为基于第一音频源的音高参数和方向来估计语音可识度指示符。

[0088] 处理器14配置为补偿使用者的听力损失并基于输入信号提供电输出信号15。接收器16将电输出信号15转换为音频输出信号以引向听力设备使用者的鼓膜。

[0089] 语音可识度估计器12a可以包括语音合成器12ab，用于基于音高参数来生成重建的语音信号。语音可识度估计器12a可以配置为基于由语音合成器提供的重建的语音信号来估计语音可识度指示符。

[0090] 语音可识度估计器12a可以包括短时客观可识度(STOI)估计器12ac。短时客观可识度估计器12ac配置为比较重建的语音信号和基于一个或多个麦克风输入信号的基本语音信号，并且基于比较来提供语音可识度指示符。例如，短时客观可识度估计器12ac比较重建的语音信号(例如，重建的参考语音信号)和基于麦克风输入信号获得的基本语音信号(例如，带噪声的语音信号)。换句话说，短时客观可识度估计器12ac评估重建的语音信号与基本语音之间的相关性，并且使用评估的相关性来向控制器12或向处理器14提供语音可识度指示符。

[0091] 语音可识度估计器12a可以包括操作地连接至音高估计器12aa的谐波模型估计器12ad，用于提供麦克风输入信号的谐波模型参数。音高估计器12aa可以配置为使用谐波模型参数以及可选的对数似然函数来导出音高参数。

[0092] 在一个或多个示例性听力设备中，麦克风组8、10包括用于提供第二麦克风输入信号11的第二麦克风10。语音可识度估计器12a可以包括方向估计器12ae，用于基于第一麦克风输入信号9和第二麦克风输入信号11来估计第一音频源的方向。语音可识度估计器12a可以配置为基于第一音频源的方向来导出语音可识度指示符。例如，通过利用由一个或多个麦克风中的一个接收到的期望的麦克风输入信号的空间-时间谐波模型，在语音可识度估计器12a处联合估计第一音频源的音高参数和方向(例如，麦克风输入信号的到达方向)。

[0093] 在一个或多个示例性听力设备中，听力设备2配置为：如果语音可识度指示符满足第一标准，则选择第一处理方案并将其应用于麦克风输入信号9、11。否则，如果语音可识度指示符不满足第一标准，则听力设备2配置为继续向麦克风输入信号应用与先前应用的相同的处理方案。第一标准可以基于第一可识度阈值。例如，当语音可识度指示符低于第一可识度阈值时，语音可识度指示符满足第一标准。替代地，可以设想，当语音可识度指示符等于或高于第一可识度阈值时，语音可识度指示符满足第一标准。

[0094] 在一个或多个示例性听力设备中，听力设备2可以配置为：如果语音可识度指示符满足第二标准，则选择第二处理方案并将其应用于第一麦克风输入信号。在一个或多个示例性听力设备中，第一和/或第二处理方案包括波束成形方案、降噪方案、增益控制方案和压缩方案中的一个或多个。

[0095] 听力设备2可以配置为使用处理器14和/或控制器12来选择第一和/或第二处理方案并将其应用于第一麦克风输入信号9。

[0096] 图2是根据本发明的操作听力设备的示例性方法的流程图。操作听力设备的方法100可以在根据本发明的听力设备或听力系统中执行。方法100包括将音频转换102为包括第一麦克风输入信号的一个或多个麦克风输入信号。方法包括获得104指示与第一麦克风输入信号相关的语音可识度的语音可识度指示符。获得104语音可识度指示符包括获得104a第一音频源的音高参数和可选地获得第一音频源的方向。语音可识度指示符基于第一音频源的音高参数和方向。方法包括基于语音可识度指示符来控制106听力设备。

[0097] 在一个或多个示例性方法中，获得104a第一音频源的音高参数包括估计104aa音高参数。获得语音可识度指示符104可以包括基于一个或多个麦克风输入信号、第一音频源的音高参数和方向来估计104b语音可识度指示符。

[0098] 在一个或多个示例性方法中，获得104语音可识度指示符包括基于音高参数来生成104c重建的语音信号，并且基于重建的语音信号来确定104d语音可识度指示符。获得104

语音可识度指示符可以包括比较104e重建的语音信号和基本语音信号,例如,使用短时客观可识度估计器。

[0099] 获得104语音可识度指示符可以包括获得麦克风输入信号的谐波模型参数并且基于谐波模型参数来导出音高参数。获得104语音可识度指示符可以包括基于麦克风输入信号来估计第一音频源的方向。

[0100] 在一个或多个示例性方法中,基于语音可识度指示符来控制106听力设备包括:如果可识度指示符满足第一标准,则选择第一处理方案并将其应用于麦克风输入信号。第一标准可以基于第一可识度阈值。第一处理方案可以包括波束成形方案、降噪方案、增益控制方案和压缩方案中的一个或多个。

[0101] 在一个或多个示例性方法中,获得104a第一音频源的音高参数可以包括从外部设备(诸如附属设备)接收音高参数。可以设想,在外部设备处(例如在线)执行语音可识度指示符的估计,其中,听力设备配置为将麦克风输入信号提供给外部设备以便获得语音可识度回馈。

[0102] 可以设想,在外部设备处(例如在线)执行语音可识度指示符的估计和根据语音可识度指示符的麦克风输入信号的处理,其中,听力设备配置为将麦克风输入信号提供给外部设备以便获得经处理的(例如,降噪的)麦克风输入信号或波束成形参数。

[0103] 图3是根据本发明的示例性听力系统200的框图。听力系统200包括附属设备202和听力设备210,该附属设备包括处理器204、接口206和存储器208。

[0104] 听力系统200包括:麦克风组,布置在听力设备中,麦克风组包括用于提供第一麦克风输入信号的第一麦克风212(和可选的第二麦克风214);以及处理器216,用于处理输入信号并基于输入信号来提供电输出信号。听力系统包括:接收器218,布置在听力设备中,用于将电输出信号转换为音频输出信号;以及控制器12,操作地连接至麦克风组212、214。控制器12配置为基于语音可识度指示符来控制处理器,如关于图1所述。图3中的控制器12布置在听力设备210中。在一个或多个示例性听力系统中,控制器12可以布置在附属设备202中。

[0105] 听力系统200配置为基于一个或多个麦克风输入信号来估计指示语音可识度的语音可识度指示符。例如,听力设备210可以配置为使用控制器12基于一个或多个麦克风输入信号来估计指示语音可识度的语音可识度指示符。在一个或多个示例性听力系统中,附属设备202可以配置为接收由听力设备210(例如,经由天线222和无线电收发器224并且通过通信链路230)发送的麦克风输入信号并且经由布置在附属设备中的控制器12基于一个或多个麦克风输入来估计指示语音可识度的语音可识度指示符。

[0106] 听力系统200配置为估计第一音频源的音高参数。语音可识度指示符基于第一音频源的音高参数和方向。例如,听力设备210可以配置为使用控制器12来估计第一音频源的音高参数,并且基于音高参数和方向来导出语音可识度。

[0107] 在一个或多个示例性听力系统中,附属设备202可以配置为接收由听力设备210(例如,经由天线222和无线电收发器224,并且通过通信链路230)发送的麦克风输入信号并且使用控制器12来估计第一音频源的音高参数,并且基于音高参数和方向来导出语音可识度。

[0108] 使用术语“第一”、“第二”、“第三”和“第四”等并不意味着任何特定顺序,而是被包

括以用于识别各个元件。此外，术语第一、第二等的使用不表示任何顺序或重要性，而是使用术语第一、第二等来区分一个元件与另一元件。请注意，这里和其他地方使用术语第一和第二，仅用于标注目的，并不旨在表示任何特定的空间或时间排序。此外，第一元件的标注并不意味着存在第二元件，反之亦然。

[0109] 虽然已经示出和描述了具体特征，但是应当理解，它们并不旨在限制所要求保护的发明，并且对于本领域技术人员显而易见的是，可以在不脱离所要求保护的发明的精神和范围的情况下进行各种改变和修改。因此，说明书和附图被认为是说明性的而不是限制性的。所要求保护的发明旨在涵盖所有替代方案、修改和等同物。

[0110] 附图标记列表

- [0111] 2 听力设备
- [0112] 4 天线
- [0113] 5 第一无线输入信号
- [0114] 7 无线电收发器
- [0115] 8 第一麦克风
- [0116] 9 第一麦克风输入信号
- [0117] 10 第二麦克风
- [0118] 11 第二麦克风输入信号
- [0119] 12 控制器
- [0120] 12a 语音可识度估计器
- [0121] 12aa 音高估计器
- [0122] 12ab 语音合成器
- [0123] 12ac 短时客观可识度(STOI)估计器
- [0124] 12ad 谐波模型估计器
- [0125] 12ae 方向估计器
- [0126] 14 处理器
- [0127] 16 接收器
- [0128] 100 操作听力设备的方法
- [0129] 102 将音频转换为一个或多个麦克风输入信号
- [0130] 104 获得语音可识度指示符
- [0131] 104a 获得第一音频源的音高参数
- [0132] 104aa 估计音高参数
- [0133] 104b 基于一个或多个麦克风输入信号来估计语音可识度指示符
- [0134] 104c 基于音高参数来生成重建的语音信号
- [0135] 104d 基于重建的语音信号来确定语音可识度指示符
- [0136] 104e 比较重建的语音信号和基本语音信号
- [0137] 106 基于语音可识度指示符来控制听力设备
- [0138] 200 听力系统
- [0139] 202 附属设备
- [0140] 204 处理器

- [0141] 206 接口
- [0142] 208 存储器
- [0143] 210 听力设备
- [0144] 212 第一麦克风
- [0145] 214 第二麦克风
- [0146] 216 处理器
- [0147] 218 接收器
- [0148] 222 天线
- [0149] 224 无线电收发器
- [0150] 230 通信链路

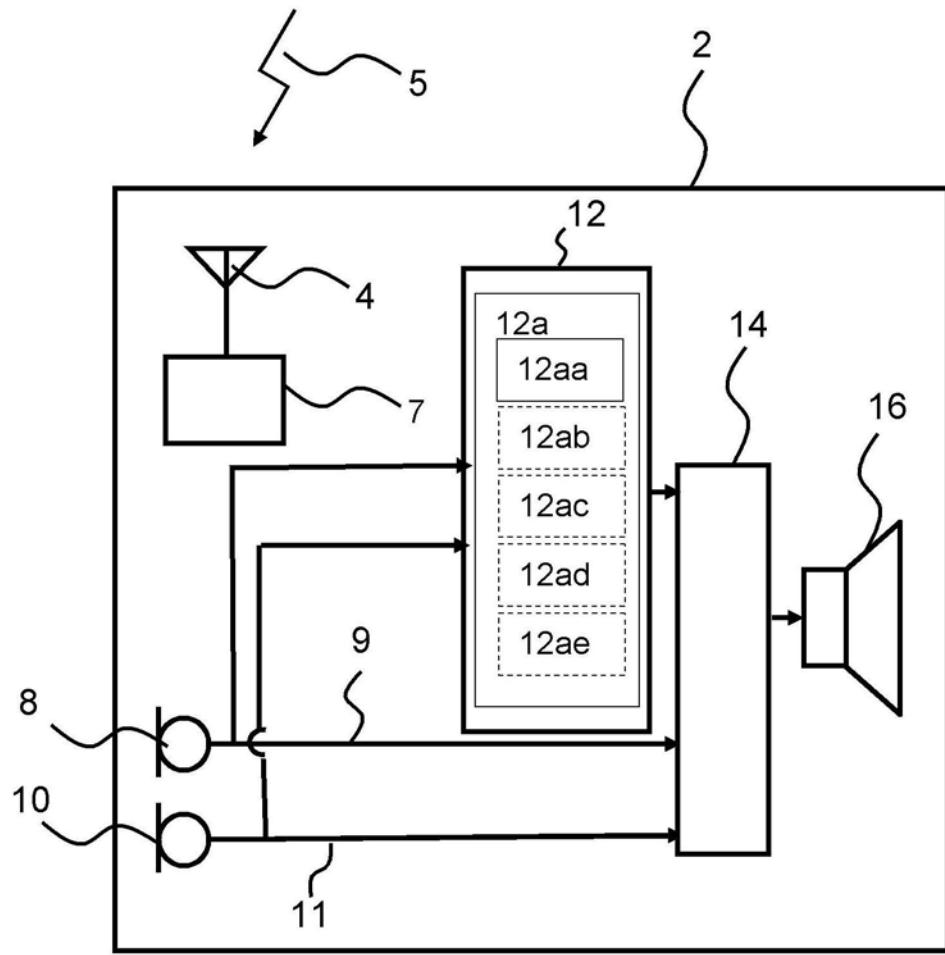


图1

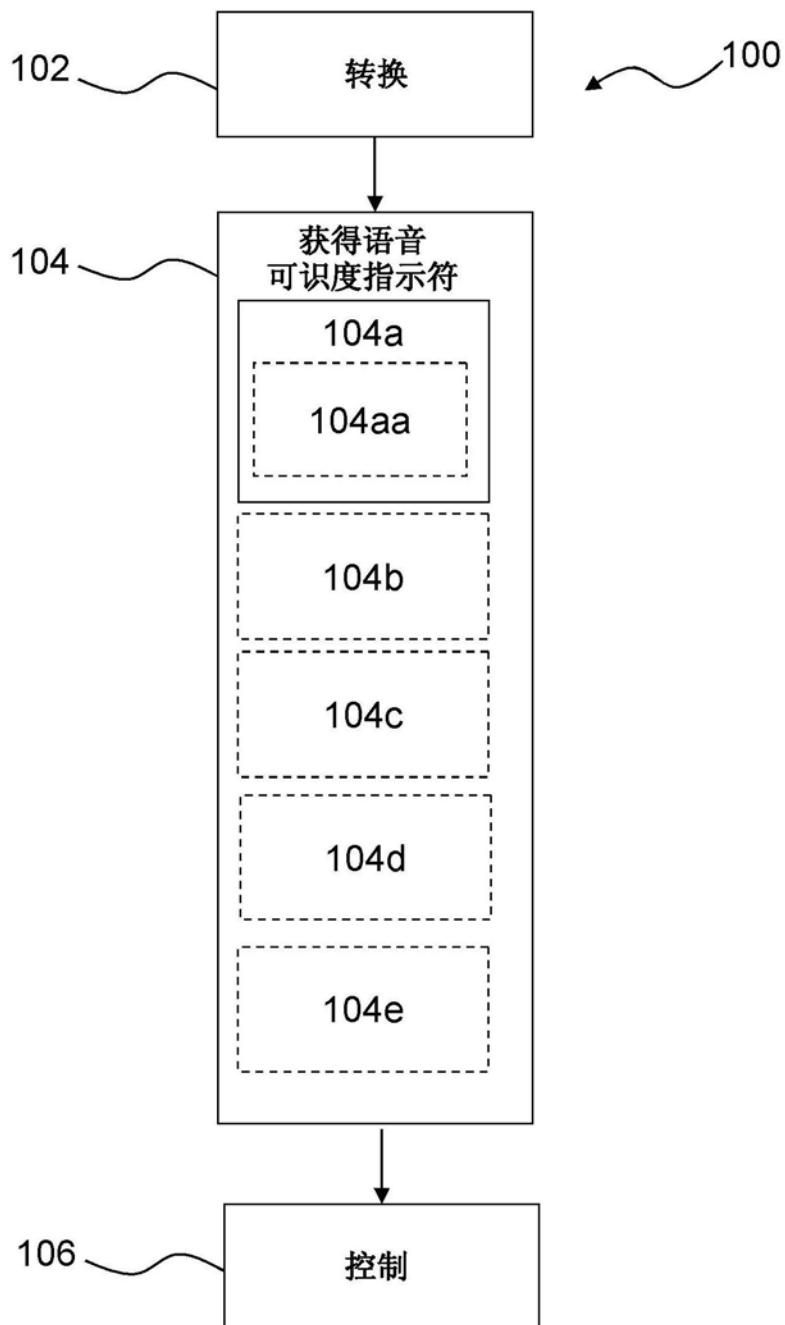


图2

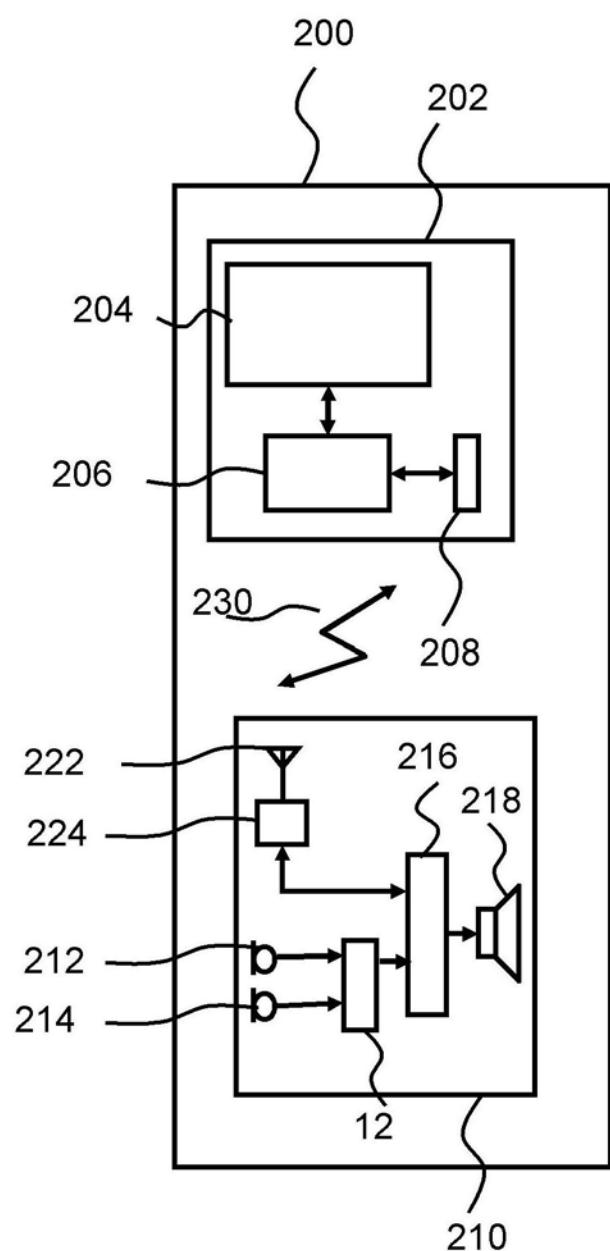


图3