



(12)发明专利

(10)授权公告号 CN 105532017 B

(45)授权公告日 2018.11.27

(21)申请号 201480019733.8

(22)申请日 2014.02.03

(65)同一申请的已公布的文献号
申请公布号 CN 105532017 A

(43)申请公布日 2016.04.27

(30)优先权数据
61/776,793 2013.03.12 US
61/798,097 2013.03.15 US
61/827,799 2013.05.28 US
13/955,723 2013.07.31 US

(85)PCT国际申请进入国家阶段日
2015.09.30

(86)PCT国际申请的申请数据
PCT/US2014/014375 2014.02.03

(87)PCT国际申请的公布数据
W02014/143439 EN 2014.09.18

(73)专利权人 谷歌技术控股有限责任公司
地址 美国加利福尼亚州

(72)发明人 凯文·J·巴什迪尔
贾尔斯·T·戴维斯

(74)专利代理机构 中原信达知识产权代理有限
责任公司 11219
代理人 李佳 穆德骏

(51)Int.Cl.
H04R 3/00(2006.01)
G10L 21/0208(2006.01)
G10L 15/01(2006.01)
G10L 25/78(2006.01)
G10L 21/0216(2006.01)

(56)对比文件
CN 102903360 A,2013.01.30,
WO 9416517 A1,1994.07.21,
US 2002193130 A1,2002.12.19,
US 2013034241 A1,2013.02.07,
CN 102708874 A,2012.10.03,
US 2010217590 A1,2010.08.26,
US 7359504 B1,2008.04.15,
US 2005049864 A1,2005.03.03,

审查员 何德超

(续)

权利要求书3页 说明书14页 附图6页

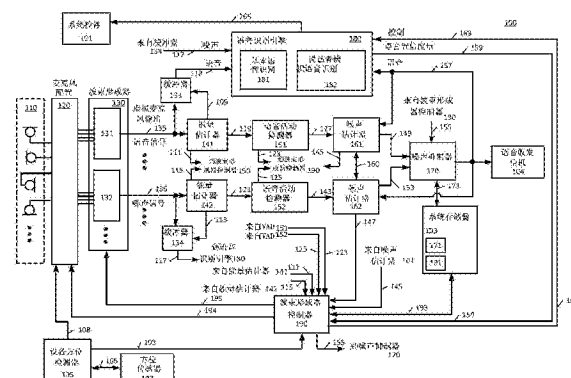
(54)发明名称

用于波束形成以获得语音和噪声信号的装置和方法

(57)摘要

一种操作方法包括波束形成多个麦克风输出以获得多个虚拟麦克风音频通道。每个虚拟麦克风音频通道对应于一个波束形成。虚拟麦克风音频通道包括至少一个语音通道(135)和至少一个噪声通道(136)。所述方法包括对至少一个语音通道(135)执行语音活动检测(151),并且调整对应的语音波束形成,直到语音活动检测(151)指示语音存在于至少一个语音通道(135)上。另一方法波束形成多个麦克风输出以获得多个虚拟麦克风音频通道,其中,每个虚拟麦克风音频通道对应于一个波束形成,并且具有至少一个语

音通道(135)和至少一个噪声通道(136)。所述方法对至少一个语音通道(135)执行语音识别,并且调整对应的语音波束形成,以提高语音识别置信度量(159)。



[转续页]

[接上页]

(72)发明人 普拉门·A·伊万诺夫
里瓦纳尔多·S·奥利韦拉

坦卡西·V·拉玛巴德兰
斯内海特哈·辛加拉朱

1. 一种用于波束形成的方法,包括:

将多个麦克风输出进行波束形成,以获得多个虚拟麦克风音频通道,每个虚拟麦克风音频通道对应于一个波束形成,包括至少一个语音通道和至少一个噪声通道;

对所述至少一个语音通道执行语音活动检测;

调整对应的语音波束形成,直到语音活动检测指示在所述至少一个语音通道上存在语音,以及

调整至少第二波束形成,直到在所述至少一个噪声通道中辨识了干扰机语音或背景噪声。

2. 如权利要求1所述的方法,还包括:

对所述至少一个噪声通道执行语音活动检测;以及

调整对应的噪声波束形成,直到语音活动检测指示在所述至少一个噪声通道上不存在语音。

3. 如权利要求2所述的方法,还包括:

对所述至少一个噪声通道执行能量估计;以及

调整对应的噪声波束形成,直到能量估计指示所述至少一个噪声通道正在从主导音频能量源接收音频。

4. 如权利要求1所述的方法,还包括:

对所述至少一个语音通道执行语音识别;以及

调整对应的语音波束形成,以提高所述语音识别的语音识别置信度量。

5. 如权利要求1所述的方法,还包括:

对所述至少一个噪声通道执行语音识别;以及

调整对应的噪声波束形成,以降低所述语音识别的语音识别置信度量。

6. 如权利要求5所述的方法,其中,对所述至少一个噪声通道执行语音识别,包括:

使用被训练成辨识特定说话者的经训练的语音识别,对所述至少一个噪声通道执行语音识别。

7. 如权利要求1所述的方法,还包括:

初始时基于对应麦克风组的检测方位,来配置所述多个麦克风输出。

8. 一种用于波束形成的方法,包括:

将多个麦克风输出进行波束形成,以获得多个虚拟麦克风音频通道,其中,每个虚拟麦克风音频通道对应于一个波束形成,包括至少一个语音通道和至少一个噪声通道;

对所述至少一个语音通道执行语音识别;

调整对应的语音波束形成,以提高所述语音识别的语音识别置信度量,以及

调整对应的噪声波束形成,直到在所述至少一个噪声通道中辨识到干扰机语音或背景噪声。

9. 如权利要求8所述的方法,其中,对所述至少一个语音通道执行语音识别,包括:

使用被训练成辨识特定说话者的经训练的语音识别,对所述至少一个语音通道执行语音识别。

10. 如权利要求8所述的方法,还包括:

对所述至少一个噪声通道执行语音活动检测;以及

调整对应的噪声波束形成,直到语音活动检测指示在所述至少一个噪声通道上不存在语音。

11.如权利要求10所述的方法,还包括:

对所述至少一个噪声通道执行能量估计;以及

调整对应的噪声波束形成,直到能量估计指示所述至少一个噪声通道正在从主导音频能量源接收音频。

12.如权利要求8所述的方法,还包括:

对所述至少一个噪声通道执行语音活动检测;以及

调整对应的噪声波束形成,直到语音活动检测指示在所述至少一个噪声通道上存在语音。

13.如权利要求8所述的方法,还包括:

对所述至少一个噪声通道执行语音识别;以及

调整对应的噪声波束形成,以降低所述语音识别的语音识别置信度量。

14.如权利要求12所述的方法,还包括:

响应于语音活动检测指示所述至少一个噪声通道上存在语音,对所述至少一个噪声通道执行语音识别。

15.如权利要求14所述的方法,其中,执行语音识别包括:

使用被训练成辨识特定说话者的经训练的语音识别,对所述至少一个噪声通道执行语音识别。

16.如权利要求15所述的方法,还包括:

调整对应的噪声波束形成,以降低所述经训练的语音识别的语音识别置信度量。

17.一种用于波束形成的装置,包括:

波束形成器,所述波束形成器操作地耦合到多个麦克风输出,所述波束形成器操作用于提供多个虚拟麦克风音频通道作为波束形成器输出,每个虚拟麦克风音频通道对应于一个波束形成,包括至少一个语音通道和至少一个噪声通道;

波束形成器控制器,所述波束形成器控制器操作地耦合到所述波束形成器,所述波束形成器控制器操作用于:

监视所述至少一个语音通道和所述至少一个噪声通道,以确定在所述至少一个语音通道或所述至少一个噪声通道中的任一个上是否存在语音;

控制波束形成器,以调整与所述至少一个语音通道相对应的波束形成,直到在所述至少一个语音通道上存在语音,以及

控制所述波束形成器,以调整与所述至少一个噪声通道相对应的波束形成,直到在所述至少一个噪声通道上不存在语音。

18.如权利要求17所述的装置,还包括:

语音活动检测器,所述语音活动检测器操作地耦合到所述波束形成器以接收所述至少一个语音通道,以及操作地耦合到所述波束形成器控制器;以及

其中,所述波束形成器控制器操作用于通过监视从所述语音活动检测器接收到的输入,来监视所述至少一个语音通道,以确定语音是否存在。

19.如权利要求17所述的装置,还包括:

语音识别引擎,所述语音识别引擎操作地耦合到所述波束形成器以接收所述至少一个语音通道,以及操作地耦合到所述波束形成器控制器,所述语音识别引擎操作用于对所述至少一个语音通道执行语音识别,以检测语音;以及

其中,所述波束形成器控制器操作用于通过监视从所述语音识别引擎接收到的输入,来监视所述至少一个语音通道,以确定语音是否存在。

用于波束形成以获得语音和噪声信号的装置和方法

技术领域

[0001] 本公开通常涉及语音处理,更具体地,涉及应用双重或多重输入噪声抑制的波束形成系统和方法。

背景技术

[0002] 诸如但不限于移动电话、智能电话、个人数字助理(PDA)、平板电脑、笔记本电脑或其他电子设备等的移动设备越来越多地包括语音识别系统,以提供设备的免提语音控制。虽然语音识别技术已经改进,但是当感兴趣的声音存在于其它讲话者或环境噪声中时,准确的声音识别仍然是一个技术挑战。这些技术挑战不仅存在于语音识别技术,也存在于语音处理,诸如在今天可以使用具有合适的电话应用几乎任何电子设备进行的电话中使用,尽管流行移动电话和智能电话。

[0003] 在移动设备上实现语音传输或语音识别系统时的一个特别挑战是许多类型的移动设备支持这样的使用情况,用户(因此用户的语音)根据使用情况相对于移动设备处于不同位置。而另外一个挑战是,包括其他讲话者的各种噪声源(即干扰机语音)也可能位于相对于移动设备的不同位置。这些噪声源中的一些根据位置和幅度可能随时间的函数而变化。所有这些因素构成移动设备操作和影响移动设备的麦克风拾取的声音的声学环境。此外,随着移动或以特定方式定位移动设备,移动设备的声学环境也随之改变,从而也改变移动设备的麦克风拾取的声音。可以通过语音识别系统或语音传输系统的接收侧的收听者在一个声学环境下识别的语音声音由于移动设备的运动、定位或环境噪声水平而可能在特定改变的条件不能识别。在周围环境中的各种其他条件可以添加噪声、回声或导致也对语音识别系统或语音传输系统产生不利影响的其他不希望的声学条件。

[0004] 更具体地,移动设备声学环境影响诸如麦克风阵列、噪声抑制器、回波消除系统的信号处理组件的操作和用于提高语音识别和语音呼叫性能两者的信号调节。对于移动设备,也对于固定设备,扬声器和其它干扰机扬声器或其它噪声源也可能相对于设备麦克风改变位置。这也导致对声学环境产生不利影响,并且可能导致由于干扰机扬声器或其它噪声源引起的噪声干扰而使语音不被语音识别系统或收听者识别。

附图说明

[0005] 图1是根据实施例的装置的示意框图。

[0006] 图2是提供根据各种实施例的图1的装置的操作的示例方法流程图。

[0007] 图3是示出根据各种实施例的图1的装置的操作的另一示例方法的流程图。

[0008] 图4是示出根据各种实施例的与形成虚拟麦克风以获得语音信号相关的操作的示例方法的流程图。

[0009] 图5是示出根据各种实施例的与形成虚拟麦克风以获得语音信号相关的操作的另一示例方法的流程图。

[0010] 图6是示出根据各种实施例的与形成虚拟麦克风以获得具有干扰机语音的噪声信

号的操作的示例方法的流程图。

[0011] 图7是示出根据各种实施例的与形成虚拟麦克风以获得噪声信号相关的操作的另一示例方法的流程图。

[0012] 图8是示出根据各种实施例的与形成虚拟麦克风以获得噪声信号的操作的另一示例方法的流程图。

具体实施方式

[0013] 简要地,所公开的实施例的操作的方法包括:波束形成多个麦克风输出以获得多个虚拟麦克风音频通道。每个虚拟麦克风音频通道对应于一个波束形成。虚拟麦克风音频通道包括至少一个语音通道和至少一个噪声通道。该方法包括对所述至少一个语音通道执行语音活动检测和调整相应的语音波束形成,直到语音活动检测指示声音存在于所述至少一个语音通道上。

[0014] 所述方法还可以包括对所述至少一个噪声通道执行语音活动检测,以及调整对应的噪声波束形成,直到语音活动检测指示所述至少一个噪声通道基本不存在语音。所述方法还可以包括对所述至少一个噪声通道执行能量估计,以及调整对应的噪声波束形成,直到能量估计指示所述至少一个噪声通道从主导音频能量源接收音频。所述方法还可以包括对所述至少一个噪声通道执行语音识别,以及调整对应的噪声波束形成,以提高语音识别的语音识别置信度量。所述方法还可以包括对所述至少一个噪声通道执行语音识别,以及调整对应的噪声波束形成,以降低噪声波束上执行的语音视频的语音识别置信度量。

[0015] 在一些实施例中,对所述至少一个噪声通道执行语音识别可以包括使用被训练成辨识特定说话者的经训练的语音识别对所述至少一个噪声通道执行语音识别。所述方法还可以包括初始基于检测到的对应麦克风组的方位来配置所述多个麦克风输出。

[0016] 所公开的实施例的另一种操作方法包括波束形成多个麦克风输出以获得多个虚拟麦克风音频通道,其中,每一个虚拟麦克风音频通道对应于一个波束形成,并且具有至少一个语音通道和至少一个噪声通道。所述方法包括对所述至少一个语音通道执行语音识别,以及调整对应的语音波束形成,以提高语音识别的语音识别置信度量。

[0017] 在一些实施例中,对所述至少一个语音通道执行语音识别可以包括使用被训练成辨识特定说话者的经训练的语音识别对所述至少一个语音通道执行语音识别。所述方法还可以包括对所述至少一个噪声通道执行语音活动检测,以及调整对应的噪声波束形成,直到语音活动检测指示在所述至少一个噪声通道上基本不存在语音。所述方法还可以包括对所述至少一个噪声通道执行能量估计,以及调整对应的噪声波束形成,直到能量估计指示所述至少一个噪声通道从主导音频能量源接收音频。所述方法还可以包括对所述至少一个噪声通道执行语音活动检测,以及调整对应的噪声波束形成,直到语音活动检测指示所述至少一个噪声通道上存在语音。所述方法还可以包括对所述至少一个噪声通道执行语音识别,以及调整对应的噪声波束形成,以降低语音识别的语音识别置信度量。所述方法还可以包括使用被训练成辨识特定说话者的经训练的语音识别对所述至少一个噪声通道执行语音识别。所述方法还可以包括响应于指示所述至少一个噪声通道上存在语音的语音活动检测对所述至少一个噪声通道执行语音识别。所述方法还可以包括调整对应的噪声波束形成,以降低经训练的语音识别的语音识别置信度量。

[0018] 所公开的实施例还提供了一种装置,包括波束形成器,该波束形成器操作地耦合到多个麦克风输出。波束形成器操作以提供多个虚拟麦克风音频通道作为波束形成器输出,其中,每一个虚拟麦克风音频通道对应于一个波束形成并具有至少一个语音通道和至少一个噪声通道。波束形成器控制器操作地耦合到波束形成器,并且操作以监视所述至少一个语音通道和所述至少一个噪声通道,以确定所述至少一个语音通道或所述至少一个噪声通道中的任一个上是否存在语音。波束形成器控制器还操作以控制波束形成器来调整与所述至少一个语音通道相对应的波束形成,直到所述至少一个语音通道上存在语音。在一些实施例中,波束形成器控制器还操作成控制波束形成器来调整与所述至少一个噪声通道相对应的波束形成,直到所述至少一个噪声通道上基本不存在语音。

[0019] 在一个实施例中,语音活动检测器操作地耦合到波束形成器以接收所述至少一个语音通道,和操作地耦合到所述波束形成器控制器。本实施例的波束形成器控制器操作来监视所述至少一个语音通道,以通过监视从语音活动检测器接收到的输入确定是否存在语音。在另一个实施例中,语音识别引擎操作地耦合到波束形成器来接收所述至少一个语音通道,和操作地耦合到所述波束形成器控制器。语音识别引擎操作以对所述至少一个语音通道执行语音识别来检测语音,并且波束形成器控制器操作来监视所述至少一个语音通道,以通过监视从语音识别引擎接收到的输入确定是否存在语音。例如,该输入可以是语音置信度量。

[0020] 在另一个实施例中,语音识别引擎操作地耦合到波束形成器来接收所述至少一个语音通道和至少一个噪声通道。语音识别引擎操作以对所述至少一个语音通道和至少一个噪声通道执行语音识别,以检测语音。波束形成器控制器操作地耦合到波束形成器、语音活动检测器、和语音识别引擎。除其他事项外,波束形成器控制器操作来监视语音活动检测器,以确定所述至少一个语音通道或所述至少一个噪声通道中的任一个上是否存在语音,并控制波束形成器以调整对应的语音波束形成,直到语音活动检测或语音识别引擎指示所述至少一个语音通道上存在语音,以及调整对应的噪声波束形成,直到语音活动检测或语音识别引擎指示所述至少一个噪声通道上基本不存在语音。

[0021] 在一些实施例中,所述装置还可以包括能量估计器,该能量估计器操作地耦合到波束形成器和语音活动检测器。在一些实施例中,所述装置还可以包括麦克风配置逻辑,麦克风配置逻辑操作地耦合到波束形成器。麦克风配置逻辑可以包括切换逻辑,该切换逻辑操作以接通或断开多个麦克风输出的任何麦克风输出。在一些实施例中,所述装置还可以包括噪声估计器,该噪声估计器操作地耦合到语音活动检测器。

[0022] 在另一个实施例中,一种操作方法包括:波束形成多个麦克风输出,以获得至少一个虚拟麦克风通道,对切换逻辑至少一个虚拟麦克风通道执行语音识别,以及调整对应的波束形成,直到语音识别指示以下中的一个:切换逻辑至少一个虚拟麦克风通道上存在语音、或所述至少一个虚拟麦克风通道上基本不存在语音。在一些实施例中,执行语音识别可以包括使用被训练成辨识特定说话者的经训练的语音识别对所述至少一个虚拟麦克风通道执行语音识别。

[0023] 现在转到附图,图1是根据各种实施例的装置100的示意性框图。装置100可以并入和用于采用语音识别、语音传输、或语音捕获的任何电子设备中。装置100的一个应用可以用于各种移动设备中的任何一个,诸如但不限于,移动电话、智能电话、照相机、摄像机、平

板电脑、笔记本电脑、或一些其他电池供电的电子设备等。然而，装置100不限于在移动设备中使用。例如，装置100可以用于语音控制电视机、数字视频录像机、汽车控制系统、或采用语音识别或语音通信的任何其它设备或系统，诸如便携式或非便携式电话、扬声电话等。

[0024] 可以理解，为了清楚的目的，图1受限于仅示出向本领域普通技术人员描述各种实施例的特征和优点以及描述如何作出和使用各种实施例有用的那些组件。因此，可以理解，可以存在各种其他组件、电路和设备等，以便实现装置，并且本领域普通技术人员可以理解存在那些各种其它组件、电路、设备等。例如，装置可以包括用于从电源、可以连接到电池或容纳在装置100并入的电子设备或系统中的一个中的其它电源的配电总线接收电力的输入，以向装置100提供电力，或者将电力分配给装置100的各种组件。在另一示例中，装置可以包括一个或多个通信总线，用于发送控制信号或操作地耦合的组件之间的其他信息等。因此，可以理解，为了清楚的目的，省略了这样的各种其它组件、电路或设备。

[0025] 另一示例是，装置100还可以包括内部通信总线，用于在各种组件、电路和设备之间提供操作地耦合。本文所使用的术语“操作地耦合”是指实现各种组件、电路和设备之间操作和/或功能通信的耦合以及被描述为操作地耦合的各种组件、电路和设备之间的关系，并且可以包括用于实现这样通信的任何中间项目（即总线、连接器、其他组件、电路、设备等），例如，内部通信总线，诸如数据通信总线或普通技术人员将理解到存在的任何其它中介项目。此外，可以理解，“操作地耦合”的项目之间可以存在其它中介项目，即使这样的其它中介项目对操作地耦合促进的功能通信是没有必要的。例如，数据通信总线可以是存在于装置100的各种实施例中，并且数据通信总线可以沿操作地耦合的两个或多个项目的路径向若干个项目提供数据。通常，在本文中所描述的图1中示出这样的操作耦合。

[0026] 在图1中，装置100可以包括麦克风组110，该麦克风组110提供麦克风输出且被操作地耦合到麦克风配置逻辑120。虽然图1的示例示出四个麦克风，每个位于不同方向上，但是实施例不限于所示的四个麦克风或示例性方位，并且在实施例中可以使用任何数目的麦克风和麦克风方位。应该理解，图1中使用虚线示出麦克风组110，原因是麦克风组110不一定是装置100的一部分。换句话说，麦克风组110可以是装置100并入其中的移动设备或其他电子设备或系统的一部分。在此情况下，装置100通过合适的通信总线或合适的连接器等操作地耦合到位于移动设备内的麦克风组110，使得麦克风组110操作地耦合到麦克风配置逻辑120。

[0027] 麦克风配置逻辑120可以包括各种前端处理，诸如但不限于，信号放大、模-数转换/数字音频采样、回音消除等，它们可以在执行附加较小功率效率信号处理（诸如噪声抑制）之前被应用于麦克风组110输出。在一些实施例中，麦克风配置逻辑120还可以包括操作地耦合到麦克风组110的切换逻辑，并且可操作响应控制信号，以单独打开或关闭每个麦克风来以各种方式配置麦克风。替代地，在一些实施例中，可以通过调整与对应的麦克风输出相关联的增益或放大器打开或关闭麦克风。例如，可以通过对对应麦克风输出将增益值降低到零来关闭麦克风。另外，在一些实施例中，麦克风配置逻辑120可以操作以接收来自装置100的其他组件的控制信号来调整前端处理参数，诸如放大器增益。

[0028] 麦克风配置逻辑120操作地耦合到波束形成器130。在一些实施例中，波束形成器130可以被实现为具有多个输出的单个波束形成器。波束形成器130的每个输出表示虚拟麦克风信号，其中，通过对来自麦克风组110的一个或多个物理麦克风波束形成来创建虚拟麦

克风。在图1所示的示例实施例中,波束形成器130被实现为两个或更多单独的波束形成器,波束形成器131和波束形成器132,以便增加操作的速度。波束形成器131和132的每一个基于麦克风配置逻辑120或通过由波束形成器控制器190或独立于波束形成器选择给定的波束形成样式所需要的麦克风输出,从麦克风组110接收输入。也就是说,在一些实施例中,在各种组合和配置中,麦克风配置逻辑120的切换逻辑可以将部分或全部麦克风输出切换到波束形成器131,并且将部分或全部麦克风输出切换到波束形成器132,或者在波束形成器控制器190的其他实施例中,或独立于波束形成器,可以控制给定的波束形成样式需要使用哪些麦克风,并且可以通过调整波束形成器内施加的增益打开或关闭麦克风。例如,在装置100的移动设备应用中,可以基于移动设备的方位通过切换逻辑、通过波束形成器控制器190、或通过波束形成器配置麦克风。

[0029] 在一些实施例中,设备方位检测器105操作地耦合到麦克风配置逻辑120和一个或多个方位传感器107。方位传感器的一个示例是陀螺仪,设备方位检测器105可以通过连接106从方位传感器接收传感器数据,并且确定移动设备的定位。对于给定的方位,设备方位检测器105可以向麦克风配置逻辑120发送控制信号108以关闭或打开麦克风组110的特定麦克风。换句话说,各种移动设备使用情况或移动设备方位可能与特定麦克风配置相关联,并且可以通过结合设备方位对设备采取的动作触发这样的麦克风配置。这可以基于对于一些实施例中的给定方位的预先确定的配置设置,或者可以基于其它实施例中其它或附加准则。例如,将设备置于坞站可以触发接合预先确定的麦克风配置。在另一示例中,将设备置于扬声器电话模式和将设备放置在桌面可以触发另一预先确定的麦克风配置。因此,在一些实施例中,设备方位检测器105当存在时可以向波束形成器控制器190发送方位信息102,是的波束形成器控制器190可以控制或覆盖这样的使用情况或方位有关的麦克风配置逻辑120的设置。

[0030] 图1的示例装置100实施例包括两个语音检测路径,一个用于每个波束形成器131和132的每个虚拟麦克风输出。虽然图1的示例示出了两个虚拟麦克风输出、语音信号135和噪声信号136,但是在各种实施例中可以生成任何数目的虚拟语音或噪声信号。在本实施例中,两个虚拟麦克风输出的每一个在需要时提供给双输入噪声抑制器170。在利用多个语音和/或噪声信号的其它实施例中,可以使用多输入噪声抑制器。在另一个实施例中,可以使用多个双输入噪声抑制器串联,以产生单一的去噪输出信号。在另外其他实施例中,可以并联使用多个双输入噪声抑制器或多个多输入噪声抑制器,并且每一个输出可以被发送到语音识别引擎180。在这样的实施例中,可以利用产生最佳训练的或未经训练的语音置信度量的输出。

[0031] 两个对称路径存在于各自的波束形成器131和132与噪声抑制器170之间,一个用于虚拟麦克风语音信号135,并且一个用于虚拟麦克风噪声信号136。两条路径是对称的,因为它们每一个采用操作地耦合到波束形成器131和132的相应的能量估计器141和142、操作地耦合到能量估计器141和142的相应的语音活动检测器(VAD) 151和152、和操作地耦合到VAD的151和152噪声估计器161和162。两个噪声估计器161和162操作地耦合到噪声抑制器170,以提供相应的控制信号149和153。噪声估计器162从VAD 152接收控制信号143。包括上述所有组件的两个通路可以被认为是“语音通道”和“噪声通道”。也就是说,适当时沿着相应通路通过各个组件发送语音信号和噪声信号连同组件之间的控制信号。在一些实施例

中,可以沿着通路、并通过组件中的一些不经任何处理或该组件所采取的其他动作来传递语音信号或噪声信号。语音通道和噪声通道是与对应的语音波束形成和噪声波束形成相关的虚拟通道。可以由波束形成器131创建语音波束形成,并且可以由波束形成器132创建噪声波束形成。语音信号135可以被认为语音通道,语音通道还可以被认为是虚拟麦克风输出中的一个。噪声信号136可以被认为噪声通道,噪声通道还可以被认为是虚拟麦克风输出中的另一个。“虚拟麦克风”对应于可以并入来自麦克风组110的一个或多个物理麦克风的音频的波束形成。虽然图1提供一个“语音通道”和一个“噪声通道”的示例,但是在各种实施例中可以创建和使用任何数目的语音通道或噪声通道。而且,在一些实施例中,各种通道组件可以是执行一个或多个通道的操作的单个集成组件。例如,能量估计器141和能量估计器142可以集成为单个能量估计器,该单个能量估计器通过提供双输入或在时域多址接入方法或某个其它合适的方法中供应语音通道和噪声通道两者。在一些实施例中,VAD 151和VAD 152或噪声估计器161和噪声估计器162也可以以集成方式来实现。

[0032] 每个虚拟麦克风输出操作地耦合到相应缓冲器133和134,缓冲器133和134可以是循环缓冲器,以在通路上正在发生信号检查时存储语音数据或噪声数据。也就是说,当信号被检查以确定语音是否实际存在于信号中时可以存储信号数据。因此,信号被缓冲作为感兴趣的信号,使得如果语音或噪声被确定为存在,则信号可以被相应地处理或使用。例如,在一些实施例中,来自波束形成器130的语音和噪声信号可以被缓冲并发送到语音识别引擎180,同时波束形成器130继续调整波束形成样式,以改善语音和噪声信号。

[0033] 出于解释的目的,将首先详细描述语音信号135通路。噪声信号136的对称通路以类似的方式操作,并且下面将讨论任何差别。因此,从语音信号135开始,能量估计器141操作地耦合到缓冲器133和VAD151。能量估计器141向缓冲器133提供控制信号109,向VAD 151提供语音和控制信号119,并且向波束形成器控制器190提供控制信号111。噪声信号136能量估计器142向缓冲器134提供控制信号113。在一些实施例中,缓冲器133和缓冲器134的每一个可以分别由VAD151和VAD 152控制,并且能量估计器141和能量估计器142可以不存在。也就是说,在一些实施例中,VAD 151和VAD 152用于检测由波束形成器130生成的相应波束形成样式中的语音能量,而不是如使用能量估计器时最初寻找非特定音频能量。在其他实施例中,可以省略VAD,而是仅语音识别引擎180和语音置信度量(没有VAD)可以用作语音存在于信号中的指示符。对于各种实施例和各种操作的相关方法,下面进一步讨论这些操作。

[0034] VAD 151进一步操作地耦合到噪声估计器161,并提供语音和控制信号127。VAD 151操作地耦合到波束形成器控制器190,并且提供当VAD 151检测到语音时通知波束形成器控制器190的控制信号123。在一些实施例中,噪声估计器161可以是信噪比(SNR)估计器或者可以是某个其它类型的噪声估计器。噪声估计器161操作地耦合到波束形成器控制器190,并且提供当需要对语音信号135噪声抑制时通知波束形成器控制器190的控制信号145。换句话说,控制信号145向波束形成器控制器190提供信息,波束形成器控制器190进而控制波束生成器131,使得波束形成器131可以继续扫描,或者可以调整波束形成样式,以便减少语音信号中包含的一些噪声。

[0035] 组件VAD 151和152和噪声估计器161和162的每一个可以都操作地耦合到相应的缓冲133和缓冲器134,以分别接收缓冲的语音信号118或缓冲的噪声信号117。噪声抑制器170可以操作地耦合到缓存器133和缓冲器134两者,以接收缓冲的语音信号118和缓冲的噪

声信号117两者。为了在示出各种其他控制连接等时清楚,没有在图1中示出这些连接。

[0036] 因此,噪声估计器161可以从缓冲器133接收缓冲的语音信号118,并且向波束形成器控制器190提供控制信号145,以及向噪声抑制器170提供语音和控制信号149。噪声估计器161还通过控制和数据连接160操作地耦合到噪声估计器162,使得两个噪声估计器可以获得和使用来自其它通道的信息,以在一些实施例中执行各种噪声估计操作。噪声抑制器170操作地耦合到语音识别引擎180以提供噪声抑制的声音信号157,操作地耦合到波束形成器控制器190以接收控制信号155,以及通过读写连接173操作地耦合到系统存储器103。噪声抑制器170可以访问系统存储器103,以读取和检索存储在噪声抑制算法数据库171中由噪声抑制器170执行的噪声抑制算法。波束形成器控制器190通过读-写连接193操作地耦合到系统存储器103,以访问存储在波束形成样式数据库191中的预先确定的波束形成样式。系统存储器103是非易失性、非短暂性存储器。

[0037] 在一些实施例中,噪声抑制器170可以从缓冲器133接收缓冲的语音信号118,并且向语音识别引擎180和/或一个或多个语音收发信机104提供噪声抑制的语音信号157。在一些实施例中,可以不使用且可以不存在语音识别引擎180。也就是说,在一些实施例中,如果装置100并入的设备采用互联网协议语音(VoIP)系统,则噪声抑制的语音信号157可以通过有线或无线电信通道或通过有线或无线网络连接仅提供给一个或多个语音收发信机104。在具有语音识别引擎180的实施例中,语音识别引擎180可以操作地耦合到系统控件101,系统控件101可以是任何类型的语音可控系统控件,取决于装置100并入的设备,诸如但不限于,移动电话的语音控制拨号器、视频记录系统控件、移动电话的应用控件、智能电话、平板电脑、笔记本电脑、车内控制系统等,或任何其他类型的语音可控系统控件。然而,系统控件101可以不存在于所有实施例。语音识别引擎包括识别人的语音的基本语音识别(VR)逻辑181。在一些实施例中,语音识别引擎180可以另外地或替代地包括被训练成识别诸如特定用户的语音的特定人的语音的说话者辨识语音识别逻辑(SI-VR)182。

[0038] 由波束形成器控制器190发送的控制信号163可以调用VR逻辑181或SI-VR逻辑182。响应于控制信号163指令,VR逻辑181或SI-VR逻辑182将读取缓冲的噪声信号117或缓冲的语音信号118中的一个或两者。语音识别引擎180将向波束形成器控制器190提供每个短语或单词组具有对应的语音置信度量的语音至文本流,作为识别的人类语音的可能性或特定用户的语音的可能性(如果已经调用SI-VR逻辑182)的指示(即,置信得分)。图1示出这个指示作为语音置信度量159。在语音识别引擎180还用作装置100的控制功能的一些实施例中,语音识别引擎180还可以响应于语音信号157上或缓冲的语音信号118上接收到的检测的命令字,命令短语或其它语音(诸如用于语音至文本应用)向系统控件101发送控制信号165。

[0039] 在各种实施例中,波束形成器控制器190是可操作以监视对语音信号135和噪声信号136提供各种条件指示的各种控制信号。响应于所述条件,波束形成器控制器190可操作来调整波束形成器130,以改变波束形成方向性。例如,波束形成器控制器190尝试调整波束形成器131,直到语音信号135基本上是用户的语音。此外,波束形成器控制器190尝试调整波束形成器132,直到噪声信号136被联接到用户的声学环境的噪声和声音而不是用户的语音,诸如一个或多个干扰机语音或其他环境背景噪声。

[0040] 在一些实施例中,在一些情况下,形成单一的波束形成可能足够。例如,对于给定

应用或对于给定的声学环境,通过使用VAD、VR逻辑181或SI-VR逻辑182(即,经过训练的VR)以形成语音波束形成通道连同使用噪声抑制器可以提供足够的保真度和去噪声。此外,对于给定应用或对于给定的声学环境,使用已训练的VR以基本上消除用户的语音并使用噪声抑制器的噪声波束形成通道也可以提供足够的保真度和去噪。

[0041] 在一些实施例中,波束形成器控制器190可操作以配置麦克风组110,这可以通过控制麦克风配置逻辑120根据由设备方位检测器105检测到的设备方位或其它条件打开或关闭麦克风来实现。在一些实施例中,波束形成器控制器190可以生成语音或噪声信号路径的随机波束形成,其中,适当的信号路径组件检查每个的结果。在其他实施例中,波束形成器控制器190可以使得虚拟麦克风波束形成改变,使得波束形成平移或扫描音频环境,直到获得期望的条件。在另外的其它实施例中,波束形成器控制器190可以使用系统存储器103中存储的波束形成样式数据库191中存储的预先确定的波束形成样式配置波束形成器130。在另外的其它实施例中,波束形成器131和波束形成器132可以是自适应波束形成器,所述自适应波束形成器可操以确定组合麦克风组110的麦克风输出所需要的幅度和相位系数,以便在期望的方向上引导光束或空(null)。在各种实施例中,波束形成器控制器190可操作并且可以监视来自任何下列分量任意组合的控制信号,诸如从能量估计器141接收到的控制信号111、来自能量估计器142的控制信号115、来自VAD 151的控制信号123、来自VAD 152的控制信号125、来自噪声估计器161的控制信号145、和/或来自噪声估计器162的控制信号147。波束形成器控制器190还可以从语音识别引擎180接收语音置信度量159。在本文描述的特定条件下,波束形成器可操作向噪声抑制器170发送控制信号155来调用噪声抑制。在一些实施例中,波束形成器控制器190可以被集成到波束形成器130中,使得波束形成器130包括波束形成器控制器的所有特征。

[0042] 所公开的实施例采用VAD 151和VAD 152以从噪声区分语音活动(反之亦然),并相应地向波束形成器控制器190发送相应控制信号123和125。实施例还利用噪声估计器161和噪声估计器162,以确定如果不能从信号正确地区别语音何时启用或禁用降噪。

[0043] 波束形成器190基于由能量估计器141和能量估计器142检测到的能量水平,如通过VAD 151或VAD 152以及噪声估计器161和162确定的语音活动,相应地调整波束形成器131和波束形成器132的波束形成方向性。也就是说,如果检测到的能量水平超过阈值,则VAD查找语音。如果没有检测到语音,则波束形成器190可以调整相应波束形成样式。如果检测到语音,则噪声估计器确定是否需要噪声抑制或者信号本身足够。如果需要噪声抑制,则波束形成器190可以发送控制信号155以激活噪声抑制器170并且由语音识别引擎180来执行对语音信号157的语音置信度量测试。

[0044] 因此,能量估计器141和142可操作,以检测与可以作为接收到的音频信号中存在语音的指示的基线的偏差,或辨识波束形成器131和132在主导能量源(可能是主背景噪声)的方向上是否具有它们相应波束形成的高灵敏度部分。如果检测到这样的偏差,则能量估计器141可以发送控制信号119来激活VAD 151,以确定语音实际上是否存在于接收到的音频信号。超过阈值的短期偏差也可以将发送控制信号109调用到缓冲器133来调用缓冲信号。

[0045] 根据图2的流程图可以理解装置100的操作的示例方法。操作放入方法开始于操作块201,其中,装置100使用波束形成来创建至少两个虚拟麦克风。一个虚拟麦克风用于用户

的语音,并且另一个虚拟麦克风用于噪声。例如,如图1所示,波束形成器131输出虚拟麦克风语音信号135,并且波束形成器132输出虚拟麦克风噪声信号136。在操作块203,波束形成器控制器190调整波束形成的一个或两个,以定位主导能量方向。例如,在一些实施例中,能量估计器141可以检测能量水平高于阈值,并相应地将控制信号111发送到波束形成器190以向波束形成器控制器190通知已经检测到高能量水平。然而,在不需要能量估计器141的实施例中,VAD 151用于首先检测语音活动。此外,在一些实施例中,可以使用超时定时器,使得如果在给定的时间段内能量估计器没有检测到能量,则波束形成器控制器190可以进行到改变波束形成,以便搜索主导能量源,例如,通过采用自适应波束形成器来确定幅度和相位系数来朝向主导能量源引导光束或空。在操作的示例中,可以在用户的语音的方向上引导一个波束形成,以形成虚拟麦克风语音通道,并且可以在用户的语音的方向上引导空,以形成虚拟麦克风噪声通道。

[0046] 用于差分双麦克风阵列的声学教材波束样式包括双向、超心形和心形的形状,它的极性样式具有无限的深度空。在典型的物理系统中,麦克风信号之间的相位和幅度失配受各种因素影响,诸如硬件、A/D转换精度、时钟局限性等。麦克风和其周围结构之间的物理间隔距离进一步减小这些空的深度。在典型实现的宽频带信号系统中,心形样式的空深度可以是低至10dB或高达36dB。因此,如果空仅朝向存在的干扰机讲话者或噪声源,则该噪声源或干扰机的预期衰减可以至少是10至12dB。注意,使用完美匹配的麦克风和信号处理通道,衰减可以高得多。如果存在面向多个方向的多个干扰机讲话者或噪声源,则只用一个可引导空可实现的最大衰减将小于这个10至12dB的值。在一个实施例中,为了形成噪声波束,波束形成器控制器(190)可以在期望的语音引导空。期望的语音将被衰减上述量,因此噪声波束将基本上是噪声。在另一个实施例中,为了形成语音波束,波束形成器控制器(190)可以在干扰机讲话者源处引导空。然后,得到的信号将基本上是语音,仅具有干扰信号的一小部分,因为它衰减了上述量。在又一个实施例中,在扩散声场的情况下,波束形成器控制器(190)可以在期望讲话者的方向上定向超心形波束形成,从而形成信号,由于相对于全向麦克风波束样式的-6dB随机能量效率,该信号基本上是语音。

[0047] 在操作块205,波束形成器控制器190调整至少一个波束形式,直到基于语音活动检测和/或语音识别置信度量的验证在至少一个语音虚拟麦克风信号上辨识了语音。在一个示例中,VAD 151或VAD 152将被调用,以确定语音是否存在于信号中。例如,如果VAD 151在信号中没有检测到语音,则VAD 151可以向波束形成器控制器190发送控制信号123以指示波束形成器控制器190应该重新适应,或通过相应地改变波束形成以某种其它方式继续搜索语音。

[0048] 在操作块207,波束形成器控制器190调整至少第二波束形成,直到在至少一个噪声虚拟麦克风信号中辨识了干扰机语音或背景噪声。例如,在一个实施例中,VAD 152可以用于确定语音是否存在于噪声信号136中。在一些实施例中,对于VAD 152检测到语音存在的情况,VAD 152可以向波束形成器控制器190发送控制信号125来调用语音识别引擎180的使用,以进一步精炼语音检测。例如,波束形成器控制器190可以向语音识别引擎180发送控制信号163来命令SI-VR182逻辑分析缓冲的噪声信号117,并确定检测到任何语音是否是用户的语音。如果基于返回到波束形成器控制器190的语音置信度量159检测到用户的语音,则波束形成器控制器190可以改变波束形成来寻找另一个主导能量源(即,继续搜索噪声)。如

果SI-VR 182逻辑没有检测到用户的语音,则在一些实施例中,VAD 152检测到的语音活动可以被假定为干扰机语音(即,噪声源)。此外,如果语音活动检测器VAD152没有检测到语音,则控制信号125可以向波束形成器控制器190指示在噪声信号136中仅检测到背景噪声,并且因此,在任一上述示例情况下,搜索噪声源(环境噪声,干扰机语音,或两者)成功。

[0049] 在操作块209,第一和第二虚拟麦克风信号被发送到双输入噪声抑制器。在特定条件下,虚拟麦克风输出将被发送到噪声抑制器170。换句话说,在一些情况下,语音信号135的波束形成可以产生充分去噪声的语音信号,使得不需要进一步的噪声抑制。噪声估计器161和162进行是否需要噪声抑制的确定。也就是说,噪声估计器161和162确定对于语音识别引擎180正常工作是否需要噪声抑制,或者用户的语音是否将被远端听众充分理解(因为它具有足够小的背景噪声)。例如,如果对于语音信号语音置信度量太低,则需要应用噪声抑制器170。根据实施例,波束形成的虚拟麦克风语音信号和波束形成的虚拟麦克风噪声信号因此用作噪声抑制器的输入。也就是说,一旦如上面描述噪声信号136被确定为仅含有背景噪声,或者发现含有干扰机的语音,则噪声信号136可以被认为足够用作噪声抑制器的输入,并且波束形成器控制器190将向噪声抑制器170发送控制信号155,以进行双输入噪声抑制过程。然后,如图所示,操作的方法结束。

[0050] 图3是示出根据各种实施例的图1的装置的操作的另一示例方法的流程图。操作方法开始,并在操作块301中,波束形成用于创建至少两个虚拟麦克风,一个用于用户语音,另一个用于噪声。在操作块303,波束形式被调整以定位主导能源方向。在操作块305,语音活动检测器用于辨识语音是否存在于上述任一信号中。在操作块307,语音识别置信度量用于辨识检测到的任何语音是用户的语音还是干扰机语音,使得信号可以被认为是噪声。

[0051] 在操作块309,对至少一个波束进行调整,直到基于语音识别置信度量在至少一个语音虚拟麦克风信号中辨识到语音。在操作块311,对至少一个第二波束进行调整,直到在至少一个噪声虚拟麦克风信号中辨识到干扰机语音或背景噪声。在操作块313,第一和第二虚拟麦克风信号被发送到双输入噪声抑制器,并且如图所示,操作的方法结束。

[0052] 图4至图8中图示了用于获得语音和噪声麦克风虚拟信号以及相关波束形成的操作的进一步细节。从图4开始,流程图示出根据各种实施例与形成虚拟麦克风和相关波束形成以获得语音信号有关的操作的示例方法。首先,装置100可以确定并入该装置100的电子设备或系统的方位。对于相对固定的一些系统,由于设备的物理位置相对恒定,因此这些操作可被省略。例如,位于房间内特定位置中的数字录像机或电视机可以保持相对恒定。然而,对于装置100并入移动设备的应用,移动设备的方位将改变由麦克风组110所感知的声学环境。因此,可以通过根据移动设备方位改变麦克风110的配置来获得优点。因此,在一些实施例中,操作的方法开始如判定框401中所示,其中,设备方位检测器105可以与方位传感器107通信,并获得设备的方位。方位信息可以作为方位信息102被发送到波束形成器控制器190。在一些实施例中,设备方位检测器105可以将控制信号108发送到麦克风配置逻辑120,并相应地调整麦克风配置。然而,在其他实施例中,波束形成器控制器190将充当此角色,并将控制信号194发送到麦克风配置逻辑120,并根据接收带的方位信息102改变麦克风配置。在操作块403中示出这些操作。

[0053] 如果方位信息不可用,或者装置100并入的特定设备不相关,则操作的方法进行到操作块405。在操作块405,麦克风组110的一些或全部麦克风通过波束形成器130相组合。在

操作块403或操作块405中已经选择麦克风配置之后,操作的方法进行到判定块407。在判定块407,是否需要噪声抑制的判定是基于噪声估计器161的估计结果,噪声估计器161对语音信号135的噪声水平或用户的声学环境中用户语音的信号噪声比的用户环境的噪声水平进行评估。如果噪声估计器161在判定块407确定不需要噪声抑制,则将控制信号145发送到波束形成器控制器190,以指示当前波束形成足够。在一些实施例中,语音信号因此可以按原样不经进一步的噪声抑制用于各种应用,并且操作的方法结束。然而,如果判定块407中需要噪声抑制,则在操作块409将所得的噪声和语音虚拟麦克风信号发送到噪声抑制器170。

[0054] 更具体地,噪声估计器161将语音和控制信号149发送到噪声抑制器170。噪声抑制器170可以从缓冲区133获得缓冲的语音信号118,并且可以从缓冲器134获得缓冲的噪声信号117。噪声抑制器170可以通过读-写连接173访问系统存储器103,并从噪声抑制器算法数据库171获得相关的噪声抑制器算法。在一些实施例中,波束形成器控制器190可以将控制信号155发送到噪声抑制器170,以指示噪声抑制器170应该执行的来自噪声抑制器算法数据库171的噪声抑制器算法。

[0055] 噪声估计器161可以检查噪声抑制器170语音信号157,以确定应用的噪声抑制算法是否足够。如果噪声抑制足够,并且因此如果在判定块411不再需要噪声抑制,则操作的方法结束。然而,如果在判定块411仍然需要噪声抑制,则可以将语音信号157发送到语音识别引擎180。作为响应,语音识别引擎会将语音置信度量159发送到波束形成器控制器190。如果置信分数过低,则波束形成器控制器190可以在判定块415中确定仍然需要噪声抑制。如果在判定块415置信分数足够高,则不再需要噪声抑制,且操作的方法结束。如果在判定块415仍然需要噪声抑制,则控制信号163可以调用SI-VR 182,以确定用户的语音是否存在于信号中。然后,操作的方法结束。

[0056] 在一些实施例中,可以通过省略操作块413和判定块415来缩短图4所示的操作的方法,并且从判定块411直接进行到操作块417。换句话说,在一些实施例中,在尝试辨识语音信号中存在用户语音时仅利用经训练的语音识别逻辑SI-VR 182。而且,如以上针对图2所讨论的,经训练的语音识别逻辑SI-VR 182也可以被应用到噪声信号,以验证噪声信号中存在的任何语音主导是干扰机语音,而不是用户的语音。

[0057] 图5是示出根据各种实施例与形成虚拟麦克风和相关波束形成以获得语音信号有关的操作的另一示例方法的流程图。判定块501、操作块503和操作块505涉及与图4的操作块401、403和405类似的操作,因此在此不需要详细讨论。因此,操作的方法进行到操作块507,其中,噪声和语音虚拟麦克风信号被立即发送到噪声抑制器170。在操作块509,所得到的噪声抑制的语音信号157被发送到SI-VR 182逻辑。波束形成器控制器190相应地接收语音置信度量159,并且如判定块511所示确定是否需要进一步的噪声抑制。如果语音置信度量足够高,则操作的方法结束,并且语音波束形成可以被认为足够。然而,如果语音置信度量159太低,则这表明将需要进一步的噪声抑制。因此操作的方法进行到操作块513。在操作块513,波束形成器控制器将控制信号194发送到麦克风配置逻辑120,并且如果适当,从麦克风组110选择不同的物理麦克风集合。也就是说,在操作块503或操作块505之后,所有麦克风可能已经在使用。在操作块515,波束形成器控制器190将控制信号195发送到波束形成器132,并且平移或适应波束形成,或者可以从系统存储器103在存储的预定的波束形成样式数据库191中选择预定的波束形成样式。这在尝试引导波束灵敏度峰值朝向可以检测语

音的另一位置时完成。因此,在操作块515之后,操作的方法循环回到操作块507,并且如所示操作的方法重复直到成功。

[0058] 图6是示出根据各种实施例与形成虚拟麦克风和相关波束形成以获得具有干扰机语音的噪声信号有关的操作的示例方法的流程图。操作块601、603和605再次涉及确定移动设备方位,并且之前参照图4已经讨论,因此在此不需要详细讨论。操作的方法进行到操作块607,其中,虚拟麦克风信号的一部分或全部被直接发送到语音识别。因此,波束形成器控制器190可以将控制信号163发送到语音识别引擎180,以指令语音识别引擎180读取缓冲的噪声信号117。在判定块609,波束形成器控制器190检查语音置信度量159,以确定语音是否存在于任何信号中。此外,波束形成器控制器190可以从VAD 152检查控制信号125,以确定语音活动检测是否已经确定语音可以存在。如果在判定块609语音似乎存在,则操作的方法进行到操作块611,并且将信号发送到SI-VR 182逻辑。如果在判定块613基于足够低的语音置信度量159没有检测到用户的语音,则如图所示,操作的方法结束。然而,如果语音置信度量159高,使得用户的语音可能存在,则操作的方法进行到操作块615,其中,如上所述,如果适当,可以选择不同的物理麦克风集合,也就是说,假设附加麦克风可用(即,在最初仅采用一些可用麦克风的情况下)。在操作块617,波束形成器控制器190再次控制波束合成器132平移或适应波束形成,或者选择预定的波束形成样式,以便继续搜索干扰机语音。

[0059] 图7是示出根据各种实施例与形成虚拟麦克风和相关波束形成以获得噪声信号有关的操作的另一示例方法的流程图。判定块701以及操作块703和705也类似于其他流程图,因为他们涉及确定移动设备的方位,因此在此将不需要详细讨论。操作的方法进行到操作块707,其中,虚拟麦克风信号的一些或全部被直接发送到语音识别,也就是说,发送到语音识别引擎180。在操作块709,虚拟麦克风信号的一些或全部还被发送到SI-VR逻辑182。在判定块711,检查语音活动检测器与语音置信度量159,以确定是否有任何信号包含语音。如果没有,则波束形成可以被认为已成功地形成充分捕获环境噪声的波束,并且如图所示,操作的方法结束。然而,如果在判定块711检测到语音,则操作的方法进行到操作块713,并且如果适当选择不同的物理麦克风集合。在操作块715,波束形成器控制器190控制波束形成器131并平移或适应波束形成器,或者从存储的系统存储器103的波束形成样式数据库191的数据库选择预定的波束形成样式。然后,操作的方法循环回到操作块705并且继续,直到已经确定成功的噪声波束。

[0060] 图8是示出根据各种实施例与形成虚拟麦克风和相关波束形成以获得噪声信号有关的操作的另一示例方法的流程图。判定块801以及操作块803和操作块805再次涉及确定移动设备的方位,因此在此将不进行详细讨论。操作的方法进行到判定块807,其中,检查能量估计器以确定虚拟麦克风信号是否在主导能量的方向上。如果不是,则操作的方法进行到操作块815,并且如果适当,可以选择麦克风组110的不同物理麦克风集合,也就是说,假设附加麦克风可用(即,在最初仅采用一些可用麦克风的情况下)。在操作块817,如上所述,波束形成器控制器190控制波束形成器132平移或适应波束形成,或者选择预定的波束形成样式,并且操作的方法循环回到操作块805,并且继续搜索噪声波束。如果能量估计器确定在判定块807发现主导能量源,则操作的方法进行到操作块809,其中,虚拟麦克风信号的一些或全部被发送到语音识别引擎180。在操作块811,虚拟麦克风信号的一些或全部也被发送到SI-VR逻辑182。在判定块813,如果语音由语音活动检测器VAD 152或语音置信度量159

指示,则操作的方法再次进入到操作块815,其中,如上所述,在一些情况下可以选择不同的物理麦克风集合,并且在块817,在环境噪声源的继续搜索中波束形成器可以平移或适应。每当在判定块813未指示语音时,波束形成可以被认为已成功捕获环境噪声,并且如图所示,操作的方法结束。当语音信号和噪声信号都被发送到噪声抑制器170时,噪声估计器162可以将控制信号153发送到噪声抑制器170。噪声估计器162从VAD 152接收控制信号143。

[0061] 因此,鉴于上面参照图1和图2至图8的流程图详细描述的实施例,可以理解,图1所示的组件的各种组合,诸如能量估计器、VAD、噪声估计器、语音识别或说话者辨识语音识别可以用于获得本公开提供的特征和优点,并且本公开考虑了这样的各种组合。此外,可以理解,在一些实施例中,在任何特定实施例中,可以不使用或者可以不存在上述组件中的一些。在一个示例中,如果VAD用于检测语音通道或噪声通道中的语音,则在该实施例中可以不使用或者可以不存在语音识别引擎180。在另一个示例中,如果语音识别用于检测语音信号或噪声信号中的语音,则在该实施例中可以不使用或者可以不存在VAD151或VAD 152。在任一上述两个例子中,在任一示例实施例中可以不使用或者可以不存在能量估计器141和能量估计器142。因此,基于本文上面提供的实施例和各种示例的描述,本领域的普通技术人员将理解,图1鉴于本公开内容而涵盖所有这样的各种实施例。因此根据本文提供的示例和公开内容,其它这样的考虑的实施例示例对于本领域的普通技术人员将变得显而易见。

[0062] 可以理解,参照图1描述的各种组件、电路、设备等和各种流程图,包括但不限于使用术语“逻辑”描述的那些,诸如麦克风配置逻辑120、波束形成器130、缓冲器133和134、能量估计器141和142、VAD 151和152、噪声估计器161和162、噪声抑制器170、语音识别引擎180、波束形成器控制器190、或系统控件101可以以各种方式来实现,诸如通过一个或多个可编程处理器上执行的软件和/或固件,诸如中央处理单元(CPU)等,或通过ASIC、DSP、FPGA、硬连线电路(逻辑电路),或者任何它们的组合。

[0063] 此外,可以理解,本文参照图1描述的各种“控制信号”和上述各种组件可以以各种方式来实现,诸如在各种组件之间使用应用程序编程接口(API)。因此,在一些实施例中,如果这样的组件被实现为一个或多个可编程处理器上执行的软件和/或固件,则可以使用API而不是硬件通信总线操作地耦合组件。例如,波束形成器控制器190和噪声抑制器170可以是在单个处理器上执行的软件和/或固件,并且可以使用API彼此通信和交互。在另一示例中,波束形成器130和波束形成器控制器190可以是在单个处理器上执行的软件和/或固件,并且可以使用API彼此通信和交互。根据本文描述的示例和描述,另外的类似示例将是显而易见的。

[0064] 此外,可以使用指示符实现涉及系统存储器103的操作,其中组件,诸如但不限于,波束形成器控制器190或噪声抑制器170按照控制信号的指示访问系统存储器103,可以包括到存储器位置的指示符或分别访问预定的波束形成样式数据库191或噪声抑制算法数据库171等的数据库访问命令。

[0065] 可以理解,除了使用语音识别控制的设备和系统之外,各种应用可以从公开的实施例中获益。例如,可以使用本文所公开的操作的波束形成方法,以确定用于辨识移动电话的语音上行链路通道的用户和/或用于对移动电话的语音上行链路通道应用双或多输入噪声抑制为目的的语音和噪声信号。在另一个示例应用中,固定电话会议系统可以结合本文中所描述的装置和方法。根据本文所描述的说明书和各示例性实施例,各种公开的实施例

的其它应用对于本领域的普通技术人员将是显而易见的。

[0066] 尽管已经图示和描述了各种实施例,但是可以理解,本发明并不局限于此。在不脱离由所附权利要求所限定的本发明的范围的情况下,本领域技术人员将想到许多修改、改变、变化、替换和等同物。

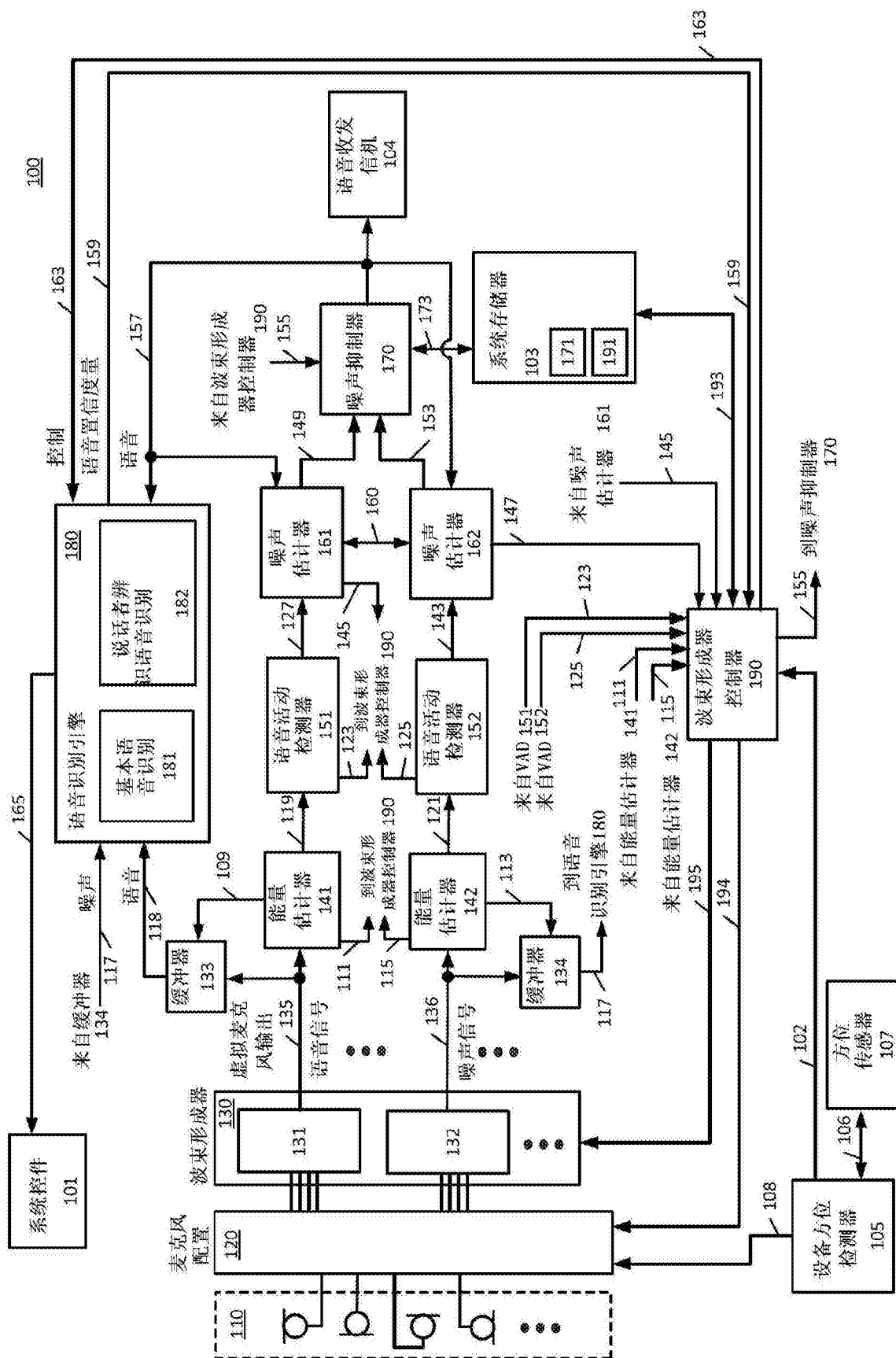


图1

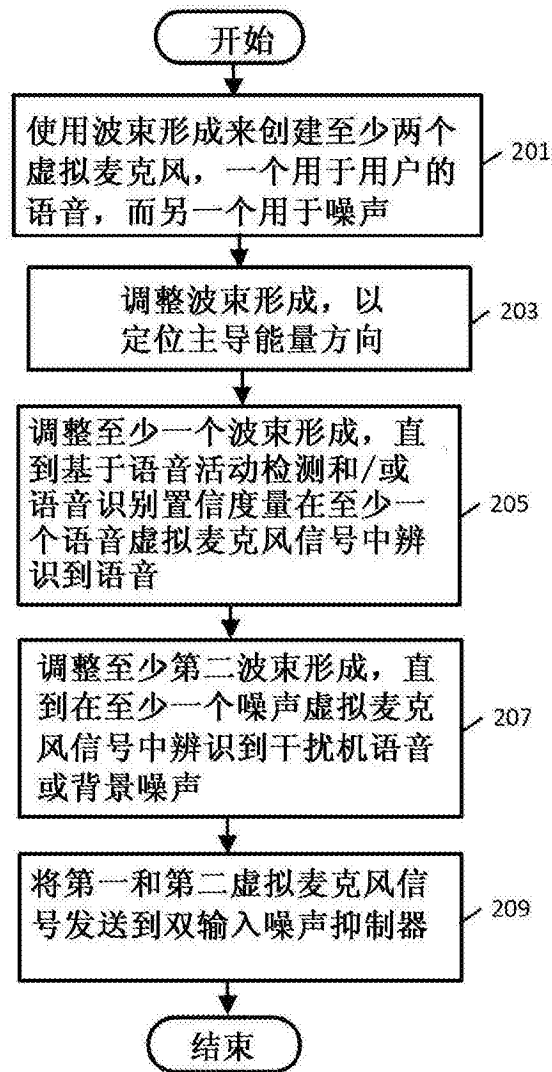


图2

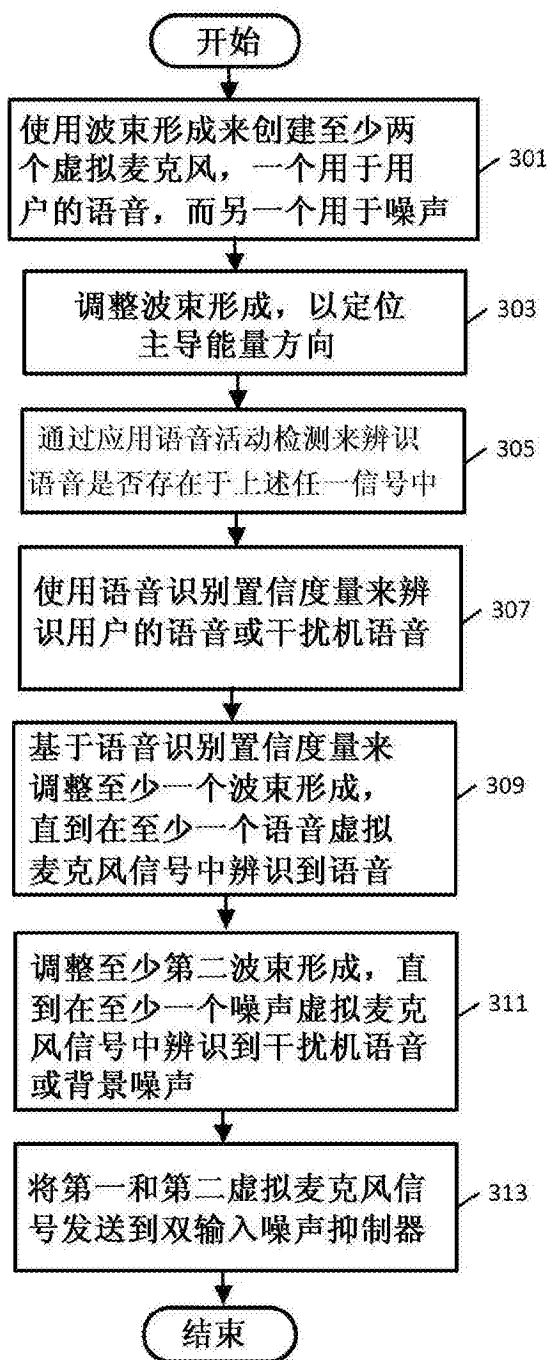
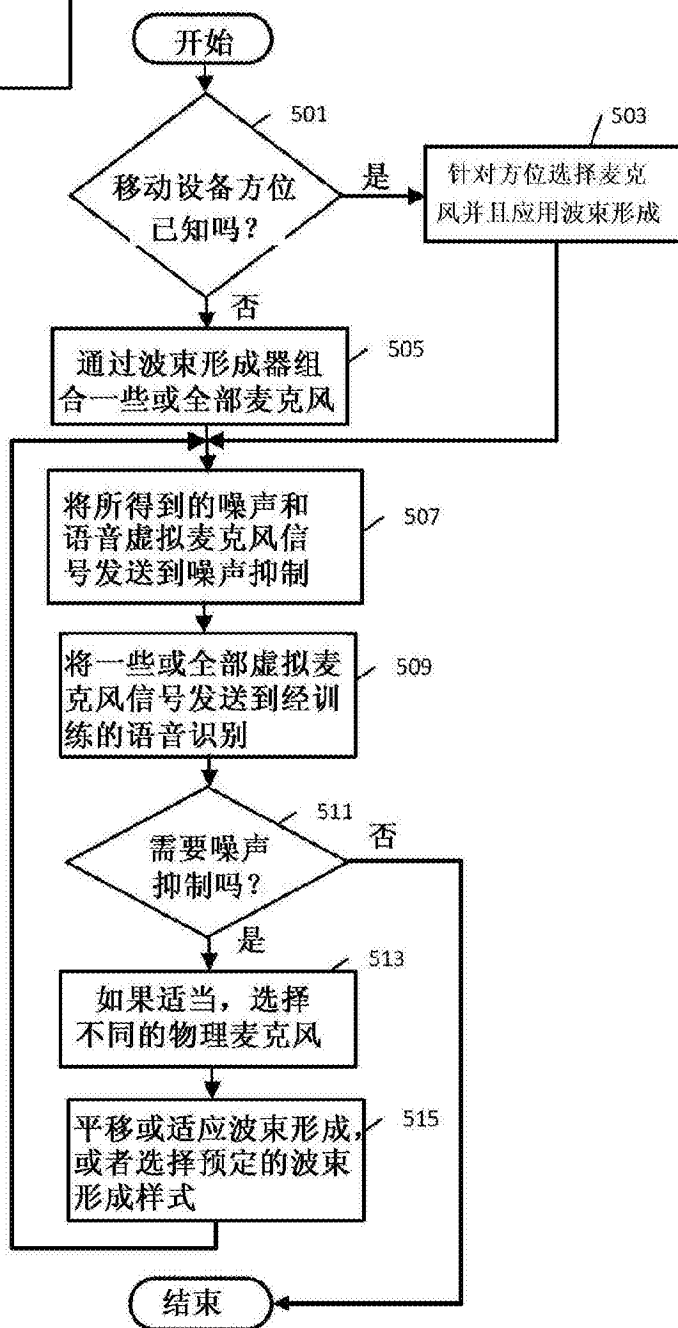
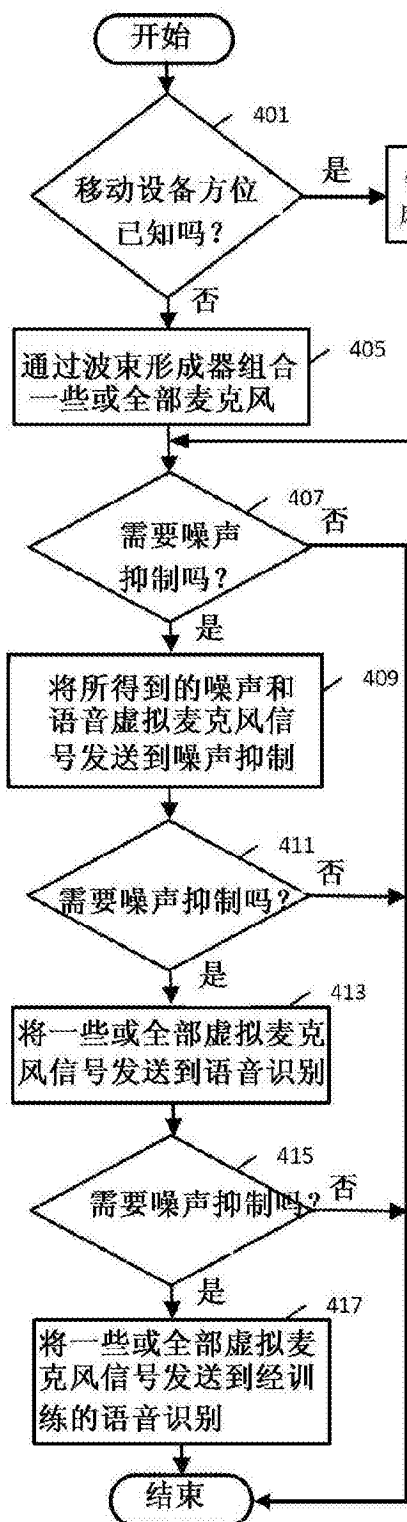


图3



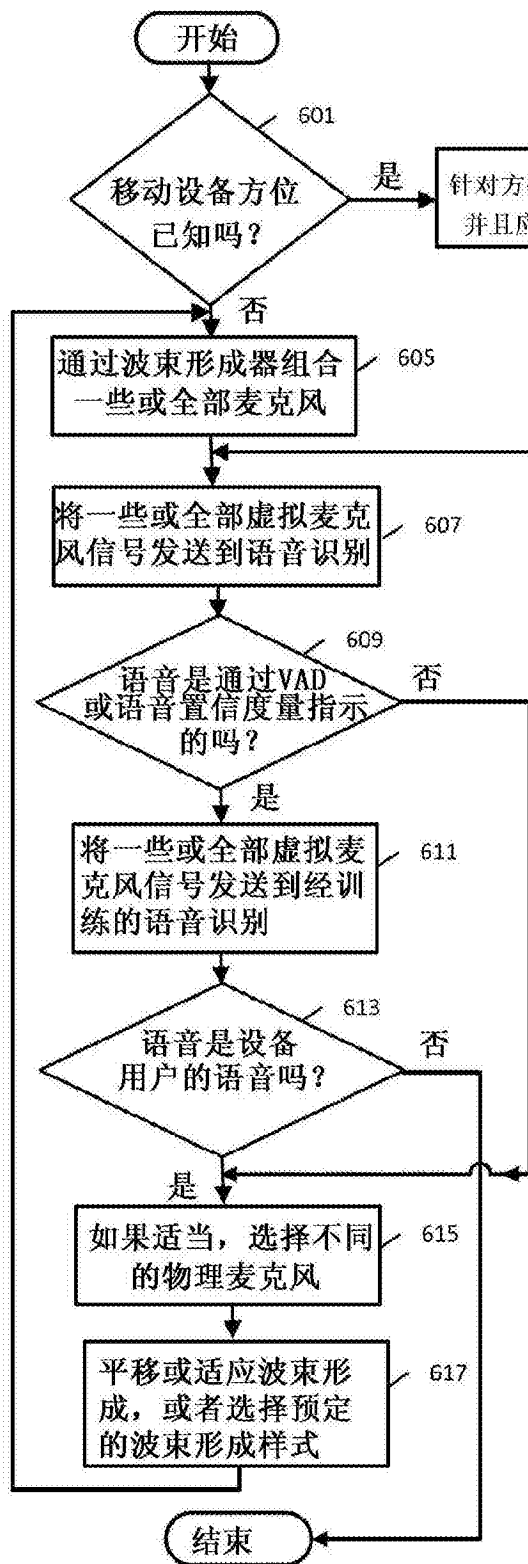


图6

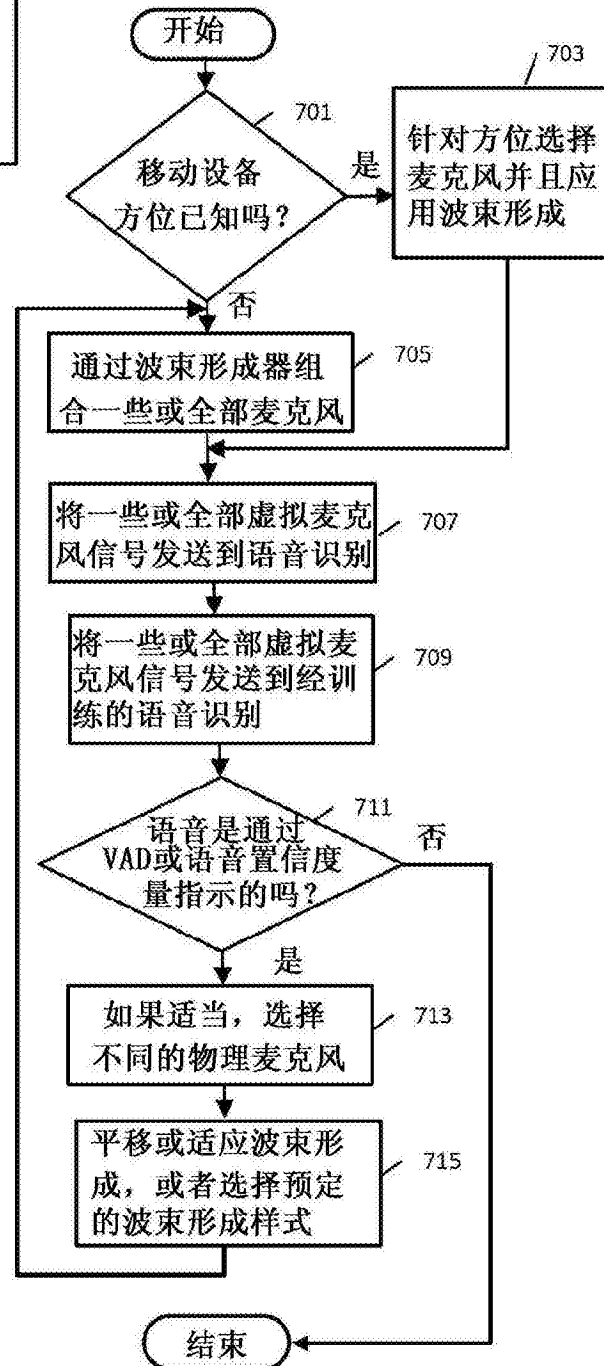


图7

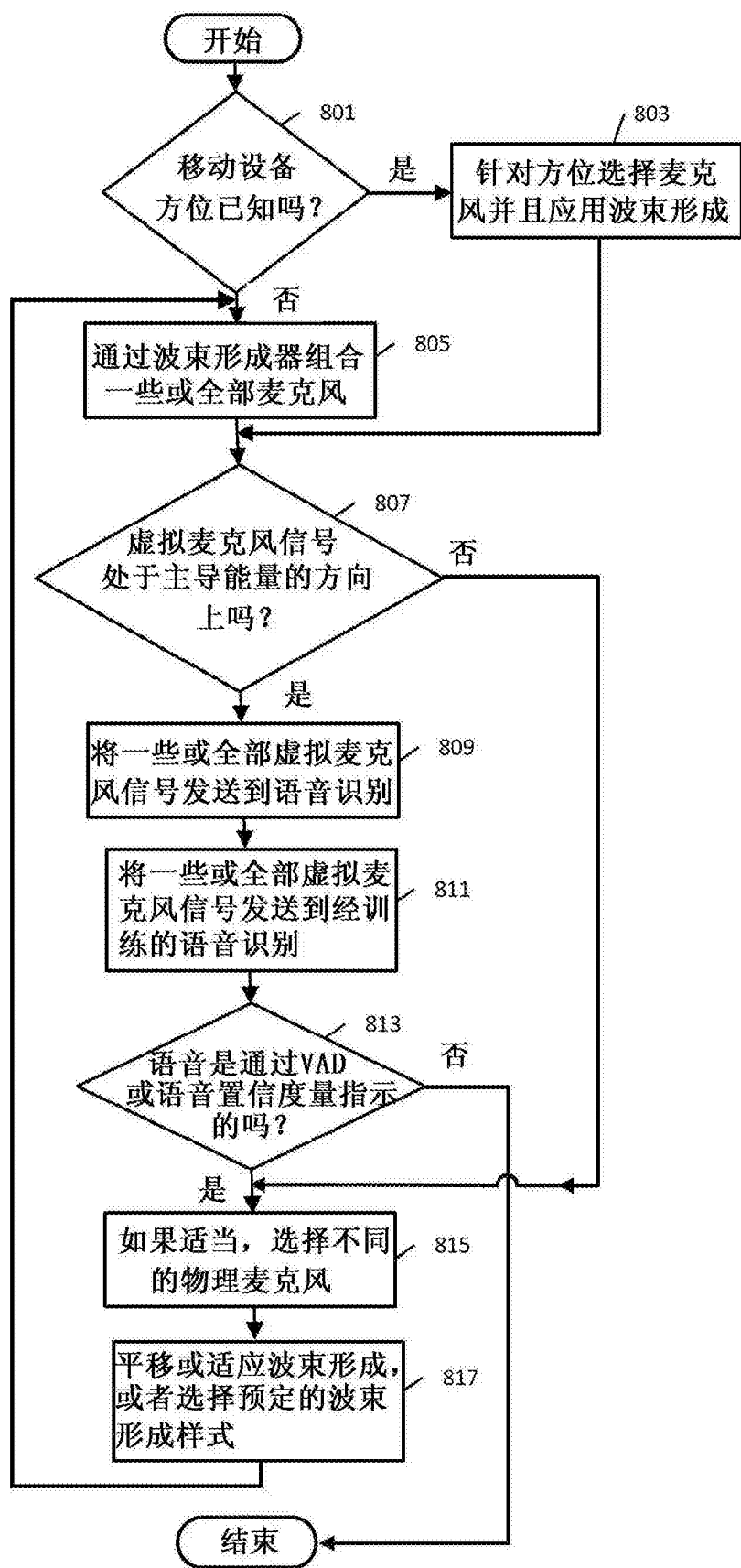


图8