

[19] 中华人民共和国国家知识产权局



[12] 发明专利说明书

专利号 ZL 200480016534.8

[51] Int. Cl.

G10L 11/02 (2006.01)

H04R 3/00 (2006.01)

H04R 1/40 (2006.01)

[45] 授权公告日 2009 年 11 月 11 日

[11] 授权公告号 CN 100559461C

[22] 申请日 2004.6.8

[21] 申请号 200480016534.8

[30] 优先权

[32] 2003.6.17 [33] EP [31] 03445076.7

[32] 2003.6.24 [33] US [31] 60/480,876

[86] 国际申请 PCT/EP2004/051059 2004.6.8

[87] 国际公布 WO2004/111995 英 2004.12.23

[85] 进入国家阶段日期 2005.12.13

[73] 专利权人 索尼爱立信移动通讯股份有限公司

地址 瑞典隆德

[72] 发明人 S·古斯塔夫松

[56] 参考文献

CN1288223A 2001.3.21

US20020009203A1 2002.1.24

CN1174457A 1998.2.25

US20030027600A1 2003.2.6

EP1206161A1 2002.5.15

审查员 何晓兰

[74] 专利代理机构 中国专利代理(香港)有限公司

代理人 杨凯 陈景峻

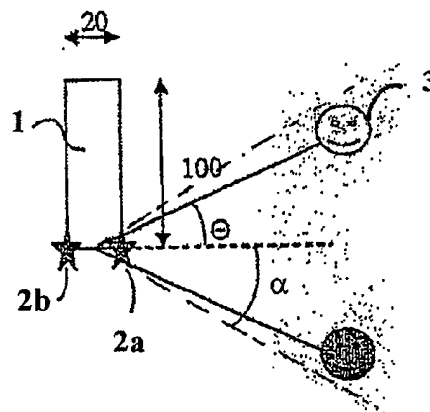
权利要求书 3 页 说明书 6 页 附图 1 页

[54] 发明名称

语音活动检测的装置和方法

[57] 摘要

本发明涉及一种具体为在移动电话中利用麦克风系统的方向敏感性和有关语音源空间方位的知识,进行语音活动检测的装置、结合该装置的移动设备及其附件以及方法。该装置包括设为判断声音信号中是否含有语音的声音信号分析器。根据本发明,所述装置还包括麦克风系统(2a, 2b, 2c, 2d, 2e),其设为区分从位于麦克风系统不同方向上的声源发出的声音,以便仅一定方向域发出的声音被作为可能含有语音的信号包括。



1. 一种用于语音活动检测的装置，包括设为确定声音信号中是否含有语音的声音信号分析器，该语音活动检测的装置包括麦克风系统（2a, 2b, 2c, 2d, 2e），该麦克风系统设为区分从位于该麦克风系统不同方向上的声源发出的声音，其特征在于：所述装置适于确定产生声音信号的声源的方向；以及

适于在所述声音信号从第一方向域发出的情况下，还对所述声音信号进行分析以确定所述声音信号是否包括语音；

而如果所述声音信号从不同的第二方向域发出，则确定所述声音信号不包括语音。

2. 如权利要求 1 所述的装置，其特征在于，所述第一方向域指向预定用户嘴（3）的方向。

3. 如权利要求 2 所述的装置，其特征在于，所述麦克风系统包括两个分开一定距离且位于指向预定用户嘴（3）的方向的线条上的麦克风元件（2a, 2b）。

4. 如权利要求 3 所述的装置，其特征在于，所述第一方向域定义为落在具有锥形角度 α ，其中 $10^\circ < \alpha < 30^\circ$ 内的锥形区域内的所有声音。

5. 如权利要求 4 所述的装置，其特征在于， α 为 25° 。

6. 如权利要求 2 所述的装置，其特征在于，所述麦克风系统包括三个分开一定距离且位于指向预定用户嘴（3）的方向的平面上的麦克风元件（2b, 2c, 2d）。

7. 如权利要求 6 所述的装置，其特征在于，所述三个麦克风元件中的两个（2c, 2d）分开一定距离且位于与预定用户嘴（3）的方向垂直的线条上。

8. 如权利要求 2 所述的装置，其特征在于，所述麦克风系统包括四个麦克风元件（2b, 2c, 2d, 2e），其设置为使其中第四个麦克风（2e）

与其它三个(2b, 2c, 2d)不在同一个平面上。

9. 权利要求1至8中任一项所述的装置, 其特征在于, 所述麦克风元件(2a, 2b, 2c, 2d, 2e)是具有在预定用户嘴(3)方向上有最大灵敏度的模式的定向麦克风元件。

10. 如权利要求1所述的装置, 其特征在于, 所述麦克风系统包括一个定向麦克风元件以及适于消除所述声源方向不确定性的一个或多个其它麦克风元件。

11. 如权利要求10所述的装置, 其特征在于, 所述定向麦克风元件适于测量相对于所述其它麦克风元件的声压级。

12. 一种移动设备, 其特征在于, 它包括如权利要求1至11中任一项所述的装置。

13. 如权利要求12所述的移动设备, 其特征在于, 所述麦克风元件(2a, 2b, 2c, 2d)位于所述装置的下边缘。

14. 如权利要求12所述的移动设备, 其特征在于, 多个麦克风元件(2a, 2b, 2c, 2d)位于所述装置的下边缘以及至少另一个麦克风元件(2e)位于与所述下边缘相距一定距离的位置上。

15. 如权利要求12至14中任一项所述的移动设备, 其特征在于, 所述移动设备是从由移动电话(1)、寻呼机、通信器、电子组织器和智能电话构成的组中选择的移动无线电终端。

16. 一种用于移动设备的附件, 其特征在于, 它包括如权利要求1至11中任一项所述的麦克风系统(2a, 2b, 2c, 2d, 2e)。

17. 如权利要求16所述的附件, 其特征在于, 所述第一方向域的方向是可调的。

18. 如权利要求16或17所述的附件, 其特征在于, 它是免提套件。

19. 如权利要求16或17所述的附件, 其特征在于, 它是电话会议麦克风。

20. 一种用于语音活动检测的方法, 其特征在于, 所述方法包括

如下步骤:

从麦克风系统 (2a, 2b, 2c, 2d, 2e) 接收声音信号, 所述麦克风系统设为区分从位于所述麦克风系统不同方向上的声源发出的声音;

确定产生所述声音信号的所述声源的方向;

如果所述声音信号是从第一方向域发出的, 则还分析所述声音信号以确定所述声音信号是否包含语音;

但如果所述声音信号是从第二方向域发出的, 则确定所述声音信号不包含语音。

21. 如权利要求 20 所述的方法, 其特征在于, 所述第一方向域指向预定用户嘴 (3) 的方向。

22. 如权利要求 21 所述的方法, 其特征在于, 所述第一方向域定义为落在具有锥形角度 α , 其中 $10^\circ < \alpha < 30^\circ$ 的锥形区域内的所有声音。

23. 如权利要求 22 所述的方法, 其特征在于, α 为 25° 。

24. 权利要求 22 或 23 中任一项所述的方法, 其特征在于, 所述麦克风系统至少包括两个彼此相距一定距离且位于指向预定用户嘴 (3) 的方向的线条上的麦克风元件 (2a, 2b), 所述两个麦克风元件间隔距离 d , 其中所述声源的方向角 θ 按如下公式计算:

$$\theta = \arccos \frac{\Delta t \cdot v}{2 \cdot d}$$

其中

Δt 是来自所述两个麦克风元件的声音之间的时差,

v 是声音的速度。

25. 如权利要求 20 所述的方法, 其特征在于, 一个定向麦克风元件配合一个或多个其它麦克风元件一起使用, 以消除所述声源方向上的不确定性。

26. 如权利要求 25 所述的方法, 其特征在于, 所述定向麦克风元件用于测量相对于所述其它麦克风元件的声压级。

语音活动检测的装置和方法

发明领域

本发明涉及一种具体为在移动电话中利用麦克风系统的方向敏感性和有关语音源空间方位的知识，进行语音活动性检测的装置、结合该装置的移动设备及其附件以及方法。所述装置协助现有语音活动检测来取得更高的灵敏度且需要较少处理器资源。

现有技术

语音活动检测器用于例如移动电话，以在某些情况下提高性能。构建语音活动检测器的最常用方法是检查输入信号子频带的电平。然后估计背景噪声电平和语音电平，并与阈值比较以判断是否存在语音。美国专利 6427134 中公开了一种语音活动检测器的实例。

例如，在噪声环境中，难以为语音活动检测器作统一的参数设置。因此，需要多种语音活动检测器，以针对特定情况进行相应调整。例如在某些模块中，需要确保是否有应该检测的语音（回波消除器），但在其它情况下，最好在信噪比等级太低的情况下指示没有任何语音。多个语音活动检测器对数字信号处理器造成负担，它必须负责执行各种语音活动检测算法。

发明概述

本发明的目的在于将声源方向纳入考虑来完善现有的语音活动检测。

在第一方面，本发明提供一种用于语音活动检测的装置，它包括设为判断声音信号是否包含语音的声音信号分析器。

根据本发明，所述装置还包括麦克风系统，其设为区分从位于

麦克风系统不同方向上的声源发出的声音,以便仅从一定方向域(range of directions)发出的声音被作为可能含有语音的信号包括。

最好,所述方向域指向预定用户嘴的方向。

在一个实施例中,所述麦克风系统包括两个分开一定距离且位于指向预定用户嘴的方向的线条上的麦克风元件。

所述方向域可以定义为落在具有锥形角度 α (其中 $10^\circ < \alpha < 30^\circ$)的锥形区域内的所有声音,最好 α 近似为 25° 。

在另一个实施例中,所述麦克风系统包括三个分开一定距离且位于指向预定用户嘴的方向的平面上的麦克风元件。

最好,所述三个麦克风元件中的两个分开一定距离且位于与预定用户嘴的方向垂直的线条上。

在另一个实施例中,所述麦克风系统包括四个麦克风元件,其中第四个麦克风与其它三个不在同一个平面上。

所述麦克风元件可以是具有在预定用户嘴方向上有最大灵敏度的模式的定向麦克风元件。

在又一个实施例中,所述麦克风系统包括一个定向麦克风元件以及一个或多个其它麦克风元件,以消除声源方向上的不确定性。所述定向麦克风元件可用于测量相对于其它麦克风元件的声压级。

在第二方面,本发明提供一种移动设备,它包括如上所述的装置。

最好所述麦克风元件位于所述装置的下边缘。

在一个实施例中,多个麦克风元件位于所述装置的下边缘以及至少另一个麦克风元件设在距所述下边缘一定距离的位置上。

所述移动设备可以是移动无线电终端,例如移动电话、寻呼机、通信器、电子组织器(electronic organiser)或智能电话。

在第三方面,本发明提供一种用于移动设备的附件,其包括如上所述的麦克风系统。

最好所述方向域的方向是可调的。

所述附件可以是免提套件或电话会议麦克风。

在第四方面，本发明提供一种用于语音活动检测的方法，包括如下步骤：

从麦克风系统接收声音信号，所述麦克风系统设为区分从位于所述麦克风系统不同方向上的声源发出的声音；确定产生所述声音信号的所述声源的方向；如果所述声音是从第一方向域发出的，则还分析所述声音以确定所述声音信号是否包含语音；但如果所述声音是从第二方向域发出的，则确定所述声音信号不包含语音。

最好所述第一方向域指向预定用户嘴的方向。

所述第一方向域可以定义为落在具有锥形角度 α （其中 $10^\circ < \alpha < 30^\circ$ ）的锥形区域内的所有声音，最好 α 近似为 25° 。

在一个实施例中，所述麦克风系统至少包括两个彼此相距一定距离且位于指向预定用户嘴的方向的线条上的麦克风元件，所述两个麦克风元件间隔距离 d ，其中至声源的方向角 θ 按如下公式计算：

$$\theta = \arccos \frac{\Delta t \cdot v}{2 \cdot d}$$

其中

Δt 是来自上述两个麦克风元件的声音之间的时差，

v 是声音的速度。

在另一个实施例中，一个定向麦克风元件配合一个或多个其它麦克风元件一起使用，以消除声源方向上的不确定性。

所述定向麦克风元件可用于测量相对于其它麦克风元件的声压级。

本发明在所附独立权利要求1、12、16和20中定义，而优选实施例在从属权利要求项中陈述。

附图简介

下面将参考如下附图更详细地说明本发明，附图中：

图 1 是集成本发明的移动电话的透视图；以及
图 2 是本发明实施例的接收角度的示意图。

优选实施例的详细说明

如前言中所简述，电话和免提套件中所用的许多信号处理算法，如回波消除和背景噪声合成基于用户在发话或未在发话的情况进行。例如，当近端用户正在发话时，语音编解码器处于活动状态，而当近端用户沉默时，背景合成处于活动状态。所有这些算法需要良好的语音活动检测器（VAD）才能较好地执行。检测操作中的错误可能导致由算法发散或其它问题引起的缺陷或故障。

现有语音活动检测器用于判断声音信号中是否存在语音。但是，实际上并非所有语音都是感兴趣或相关的，而仅有用户语音是感兴趣的或相关的。例如在若干人在讲话的噪声环境中的所有其它语音可以被忽略并视为噪声。

本发明人认识到可以利用具有某种方向灵敏度的麦克风系统来区分从位于不同方向上的声源发出的声音。非用户发出的声音可以视为非语音，这些信号无需利用常规语音活动检测器进行分析。

现有语音活动检测器会方便实施，且在本申请中仅称为声音信号分析器。

一般而言，可以采用具有某种方向灵敏度的麦克风系统。图 1 显示了具有至少两个分设的麦克风元件的实例。

图中 1 示出一般的移动电话。本发明同样适用于其它设备，如移动无线电终端、寻呼机、通信器、电子组织器（electronic organiser）或智能电话。其共同特征是，采用了语音活动检测，例如结合传送语音或通过语音识别接收语音命令。

在最简单的形式下，麦克风系统包括两个麦克风 2a 和 2b。最好将它们设在指向预定用户嘴的计算方向的线条上。最好所述麦克风元件设在所述移动设备 1 的下边缘。

图 2 显示计算声源（通常为用户嘴 3）的方向的示意图。在两个麦克风的情况下，可以只确定与麦克风元件所在线条的角度。换言之，声源的方向在具有锥形角度 θ 的锥形区域上。为计算角度 θ ，首先确定来自麦克风 2a 和 2b 的信号之间的互相关。其最大值指示两个麦克风 2a 和 2b 之间的时差 Δt 。两个麦克风 2a 和 2b 之间的距离为例如 20 毫米。角度 θ 按如下公式计算：

$$\theta = \arccos \frac{\Delta t \cdot v}{2 \cdot d}$$

注意， $\arccos$ 仅对 -1 和 1 之间的自变量有定义。如果时差为负，这意味着角度大于 90° 且声音从装置后发出。

最好该装置适于确定所有角度 θ 小于固定角度 α 的声音发自用户。阈值角度 α 可以设在例如 10° 到 30° 的范围内，最好设为 25° 。

在三个麦克风的情况下，还可以将声源的方向进一步确定为在两点（例如在上述锥形区域上）。三个麦克风元件最好设在指向用户嘴的大致方向的平面内。在图 1 中，麦克风元件 2b、2c 和 2d 是可能的设置。在前方的两个麦克风 2c 和 2d 位于垂直于用户嘴方向的线条上，而第三个麦克风 2b 位于后侧。

在四个麦克风（或更多）的情况下，可以计算所有方向的方向角，只要四个麦克风元件设置为使其中第四个麦克风与其它三个不在同一个平面上，例如设在四面体上。一种可能的设置是，前方的两个麦克风 2c 和 2d 设在下边缘，而第三个麦克风 2b 设在后侧，以及第四个麦克风 2e 设在与下边缘相距一定距离的前方。

一个类似的麦克风布置可用于移动设备的附件，如免提套件或打算放置在台面上的电话会议麦克风系统。除了麦克风元件，逻辑电路也可位于主/移动设备中。在此情况下，麦克风系统的接收角度可加以调整。这在例如当麦克风系统设置在汽车中时有用，其中用户可以坐在驾驶座位上或乘客座位上或驾驶和乘客均可以是同一呼叫过程中的发话人。接收角度的调整可以机械方式或电子方式实现，

例如通过波束成形或调整麦克风系统的方向灵敏度。

为了进一步提高麦克风系统的灵敏度，可以采用具有在用户嘴的方向上有最大灵敏度的模式的定向麦克风元件。

在另一个实施例中，一个定向麦克风元件配合一个或两个其它麦克风元件一起使用（可以是无方向的）。该定向麦克风元件用于测量相对于其它麦克风元件的声压级，由此消除声源方向上的不确定性。定向麦克风元件和无定向麦克风元件的各种组合都是可能的。

本发明可得到增强性能的语音活动检测器。利用本发明，整个信号路径上可能只需一个语音活动检测器。这将降低计算复杂性，减轻数字信号处理器上的负载并提高性能。它特别适用于具有高背景噪声和类似语音的频谱特性的噪声的环境。

本领域技术人员会意识到，本发明可以通过硬件和软件的各种组合来实现。本发明的范围仅由所附权利要求限定。

图 1

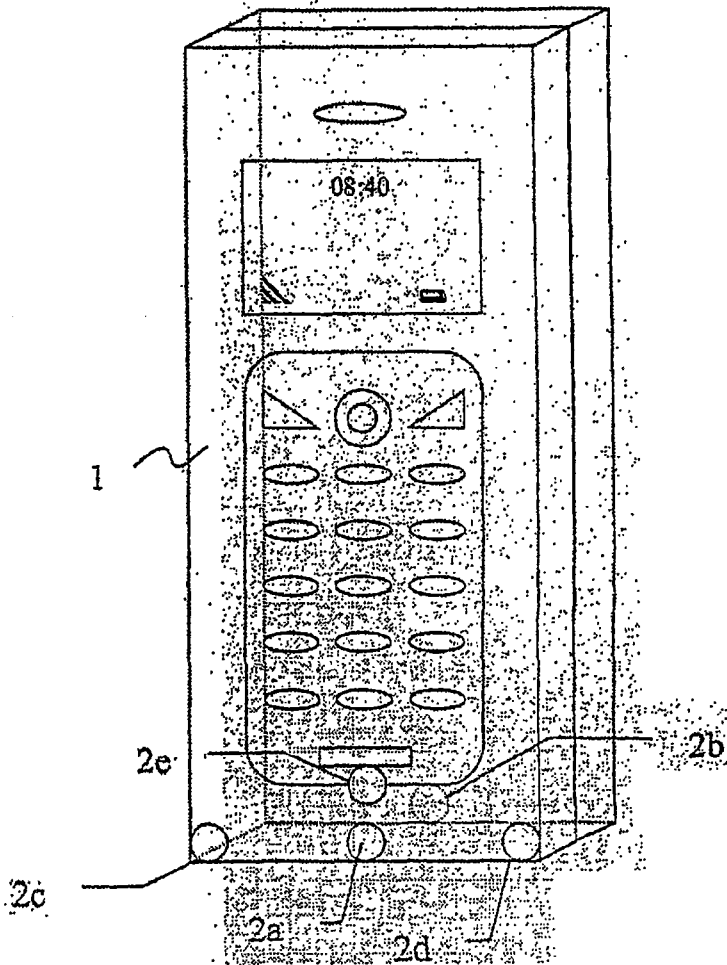
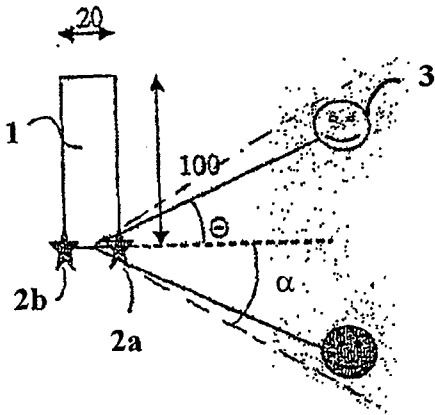


图 2