



(12)发明专利

(10)授权公告号 CN 104700832 B

(45)授权公告日 2018.05.25

(21)申请号 201310659840.4

(22)申请日 2013.12.09

(65)同一申请的已公布的文献号

申请公布号 CN 104700832 A

(43)申请公布日 2015.06.10

(73)专利权人 联发科技股份有限公司

地址 中国台湾新竹科学工业园区新竹市笃行一路一号

(72)发明人 许肇凌 郑尧文 孙良哲 王莞莞

(74)专利代理机构 北京市万慧达律师事务所

11111

代理人 张金芝 杨颖

(51)Int.Cl.

G10L 15/08(2006.01)

G10L 25/48(2013.01)

(56)对比文件

CN 102428440 A,2012.04.25,

WO 2009045861 A,2009.04.09,

US 2010277579 A1,2010.11.04,

CN 1764946 A,2006.04.26,

审查员 谭雪艳

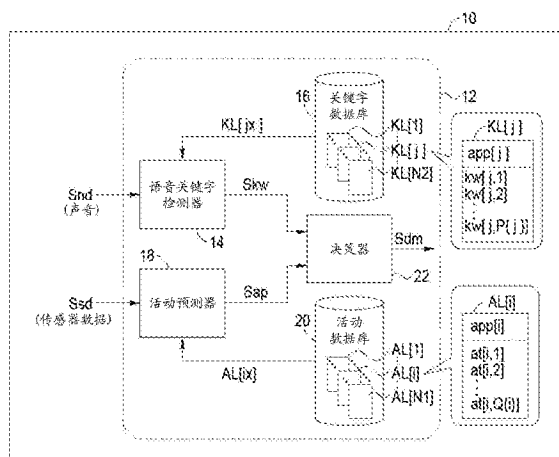
权利要求书4页 说明书13页 附图11页

(54)发明名称

语音关键字检测系统和方法

(57)摘要

本发明提供一种语音关键字检测系统和方法。该语音关键字检测系统包括,活动预测器,获得由多个传感器提供的传感器数据,以及处理该传感器数据以提供指示使用者是否将给出语音关键字的概率的活动预测结果;语音关键字检测器,用于产生初步的关键字检测结果;以及决策器,耦接于该活动预测器和该语音关键字检测器,并处理该活动预测结果和该初步的关键字检测结果以提供关键字检测结果。本发明所提供的语音关键字检测系统和方法,可更加方便、友好、可靠和准确地实现语音控制。



1. 一种语音关键字检测系统,其特征在于,包括:

活动预测器,获得由多个传感器提供的传感器数据,以及处理该传感器数据以提供指示使用者是否将给出语音关键字的概率的活动预测结果;

语音关键字检测器,用于产生初步的关键字检测结果;以及

决策器,耦接于该活动预测器和该语音关键字检测器,并处理该活动预测结果和该初步的关键字检测结果以提供关键字检测结果;

语音检测器,评估传入的声音的信息量以选择性地使能该语音关键字检测器;

其中该语音检测器包括:

第一估计器,通过该传入的声音的当前音量和第一序列的先前采样的加权和来计算该第一序列的当前采样;

第二估计器,通过该传入的声音的当前音量和第二序列的先前采样的加权和来计算该第二序列的当前采样;以及

比较器,耦接于该第一估计器和该第二估计器,将该第一序列与该第二序列进行比较以指示该传入的声音的该信息量。

2. 根据权利要求1所述的语音关键字检测系统,其特征在于,进一步包括:

活动数据库,包括多个活动清单;每一个活动清单与一个目标应用相关,并且该每一个活动清单包括多个活动模板,

其中,

该活动预测器从该多个活动清单中获得一个选择的活动清单,其中与该选择的活动清单相关的目标应用与装置当前运行的应用相匹配;以及

该活动预测器将该传感器数据与该选择的活动清单的每一个活动模板进行比较以提供该活动预测结果。

3. 根据权利要求1所述的语音关键字检测系统,其特征在于,进一步包括:

活动数据库,包括多个活动清单;每一个活动清单与一个目标应用相关,并且该每一个活动清单包括多个活动模板,

其中,

该活动预测器从该多个活动清单中获得一个选择的活动清单,其中与该选择的活动清单相关的目标应用与装置当前运行的应用相匹配;

该活动预测器通过该传感器数据提取特征以提供已提取的数据;以及

该活动预测器将该已提取的数据与该选择的活动清单的每一个活动模板进行比较以提供该活动预测结果。

4. 根据权利要求2或3所述的语音关键字检测系统,其特征在于,

该语音关键字检测系统通过增加新的活动模板或用新的活动模板代替该多个活动模板中的至少一个活动模板而更新该活动数据库。

5. 根据权利要求1所述的语音关键字检测系统,其特征在于,进一步包括:

关键字数据库,包括多个关键字清单;每一个关键字清单与一个目标应用相关,并且该每一个关键字清单包括多个候选关键字,

其中,

该语音关键字检测器从该多个关键字清单中获得一个选择的关键词清单,其中与该选

择的关键字清单相关的目标应用与装置当前运行的应用相匹配;以及

该语音关键字检测器将传入的声音与该选择的关键字清单的每一个候选关键字进行比较以提供该初步的关键字检测结果。

6. 根据权利要求5所述的语音关键字检测系统,其特征在于,

该语音关键字检测器将该传入的声音与该多个候选关键字进行比较而获得多个比较结果;

其中,该多个比较结果在数值上指示该传入的声音与该多个候选关键字之间的相似性,及/或指示该传入的声音与该多个候选关键字匹配的概率或可能。

7. 根据权利要求1所述的语音关键字检测系统,其特征在于,该语音检测器包括:

第一估计器,通过将第一乘积和第二乘积求和计算第一序列的当前采样,其中,该第一乘积为该第一序列的先前采样与第一权重的乘积,以及该第二乘积为该传入的声音的当前音量与第一互补权重的乘积;

第二估计器,通过将第三乘积和第四乘积求和计算第二序列的当前采样,其中,该第三乘积为该第二序列的先前采样与第二权重的乘积,以及该第四乘积为该传入的声音的当前音量与第二互补权重的乘积;以及

比较器,耦接于该第一估计器和该第二估计器,将该第一序列与该第二序列进行比较以指示该传入的声音的该信息量。

8. 根据权利要求7所述的语音关键字检测系统,其特征在于,该第一权重小于该第二权重。

9. 根据权利要求1所述的语音关键字检测系统,其特征在于,

该决策器采用概率融合算法或采用机器学习算法处理该活动预测结果和该初步的关键字检测结果以提供该关键字检测结果。

10. 根据权利要求1所述的语音关键字检测系统,其特征在于,

该决策器计算该活动预测结果和该初步的关键字检测结果的线性组合而作为该关键字检测结果。

11. 根据权利要求1所述的语音关键字检测系统,其特征在于,

该决策器通过该初步的关键字检测结果是否满足第一条件和该活动预测结果是否满足第二条件而获得关键字检测结果。

12. 根据权利要求1所述的语音关键字检测系统,其特征在于,该多个传感器包括加速度计、陀螺仪、磁强计、气压计、接近传感器、光传感器、触屏、定位系统的接收器、无线接收器以及照相机中的至少一个。

13. 一种语音关键字检测方法,其特征在于,包括:

获得由多个传感器提供的传感器数据;

根据该传感器数据,计算指示使用者是否将给出语音关键字的概率的活动预测结果;

通过语音关键字检测器产生初步的关键字检测结果;以及

根据该活动预测结果和该初步的关键字检测结果计算关键字检测结果;

其中,

通过评估传入的声音的信息量而执行语音检测以选择性地使能该语音关键字检测器;
该执行语音检测的步骤包括:

通过该传入的声音的当前音量和第一序列的先前采样的加权和来计算该第一序列的当前采样;

通过该传入的声音的当前音量和第二序列的先前采样的加权和来计算该第二序列的当前采样;以及

将该第一序列与该第二序列进行比较以指示该传入的声音的该信息量。

14. 根据权利要求13所述的语音关键字检测方法,其特征在于,进一步包括:

提供包括多个活动清单的活动数据库;每一个活动清单与一个目标应用相关,并且该每一个活动清单包括多个活动模板;

识别装置当前运行的应用;

通过将该当前运行的应用与该每一个活动清单的每一个目标应用进行比较选择该多个活动清单中的一个;以及

将该传感器数据与选择的活动清单的每一个活动模板进行比较以提供该活动预测结果。

15. 根据权利要求13所述的语音关键字检测方法,其特征在于,进一步包括:

提供包括多个活动清单的活动数据库;每一个活动清单与一个目标应用相关,并且该每一个活动清单包括多个活动模板;

识别装置当前运行的应用;

通过将该当前运行的应用与该每一个活动清单的每一个目标应用进行比较选择该多个活动清单中的一个;以及

通过该传感器数据提取特征以提供已提取的数据;

将该已提取的数据与选择的活动清单的每一个活动模板进行比较并相应提供包括于活动预测结果中的活动匹配结果。

16. 根据权利要求14或15所述的语音关键字检测方法,其特征在于,

通过增加新的活动模板或用新的活动模板代替该多个活动模板中的至少一个活动模板而更新该活动数据库。

17. 根据权利要求13所述的语音关键字检测方法,其特征在于,进一步包括:

提供包括多个关键字清单的关键字数据库;每一个关键字清单与一个目标应用相关,并且该每一个关键字清单包括多个候选关键字;

识别装置当前运行的应用;

通过将该当前运行的应用与该每一个关键字清单的每一个目标应用进行比较选择该多个关键字清单中的一个;以及

将传入的声音与选择的关键字清单的每一个候选关键字进行比较以提供该初步的关键字检测结果。

18. 根据权利要求17所述的语音关键字检测方法,其特征在于,该将传入的声音与选择的关键字清单的每一个候选关键字进行比较以提供该初步的关键字检测结果的步骤包括:

将该传入的声音与该多个候选关键字进行比较而获得多个比较结果;

其中,该多个比较结果在数值上指示该传入的声音与该多个候选关键字之间的相似性,及/或指示该传入的声音与该多个候选关键字匹配的概率或可能。

19. 根据权利要求13所述的语音关键字检测方法,其特征在于,该执行语音检测的步骤

包括：

通过将第一乘积和第二乘积求和以计算第一序列的当前采样，该第一乘积为该第一序列的先前采样与第一权重的乘积，以及该第二乘积为该传入的声音的当前音量与第一互补权重的乘积；

通过将第三乘积和第四乘积求和以计算第二序列的当前采样，该第三乘积为该第二序列的先前采样与第二权重的乘积，以及该第四乘积为该传入的声音的当前音量与第二互补权重的乘积；以及

将该第一序列的该当前采样与该第二序列的该当前采样进行比较以指示该传入的声音的该信息量。

20. 根据权利要求19所述的语音关键字检测方法，其特征在于，该第一权重小于该第二权重。

21. 根据权利要求13所述的语音关键字检测方法，其特征在于，该计算关键字检测结果的步骤包括：

采用概率融合算法或采用机器学习算法处理该活动预测结果和该初步的关键字检测结果。

22. 根据权利要求13所述的语音关键字检测方法，其特征在于，
计算该活动预测结果和该初步的关键字检测结果的线性组合而作为该关键字检测结果。

23. 根据权利要求13所述的语音关键字检测方法，其特征在于，
通过该初步的关键字检测结果是否满足第一条件和该活动预测结果是否满足第二条件而获得关键字检测结果。

24. 根据权利要求13所述的语音关键字检测方法，其特征在于，该多个传感器包括加速度计、陀螺仪、磁强计、气压计、接近传感器、光传感器、触屏、定位系统的接收器、无线接收器以及照相机中的至少一个。

语音关键字检测系统及方法

技术领域

[0001] 本发明有关于语音关键字检测系统及方法,特别有关于通过利用传感器检测使用者的活动以提高语音关键字检测的系统及相关方法。

背景技术

[0002] 由于具有语音控制及/或互动能力(interaction capability)的电子装置能够提供免提(hand-free)的使用者界面,这种电子装置越来越流行。识别在语音中的关键字(包括命令)的语音辨识是施行语音控制及/或互动能力所必须的技术。

发明内容

[0003] 有鉴于此,本发明提供一种语音关键字检测系统及方法。

[0004] 依据本发明一实施方式,提供一种语音关键字检测系统。该语音关键字检测系统包括:活动预测器,获得由多个传感器提供的传感器数据,以及处理该传感器数据以提供指示使用者是否将给出语音关键字的概率的活动预测结果;语音关键字检测器,用于产生初步的关键字检测结果;以及决策器,耦接于该活动预测器和该语音关键字检测器,并处理该活动预测结果和该初步的关键字检测结果以提供关键字检测结果。

[0005] 依据本发明另一实施方式,提供一种语音关键字检测方法。该语音关键字检测方法包括:获得由多个传感器提供的传感器数据;根据该传感器数据,计算指示使用者是否将给出语音关键字的概率的活动预测结果;通过该语音关键字检测产生初步的关键字检测结果;以及根据该活动预测结果和该初步的关键字检测结果计算关键字检测结果。

[0006] 本发明所提出的语音关键字检测系统及方法,可更加方便、友好、可靠和准确地实现语音控制。

附图说明

[0007] 图1为根据本发明实施方式的语音关键字检测系统的示意图。

[0008] 图2为传感器数据的检测结果和提取的检测结果的曲线。

[0009] 图3为根据本发明实施方式的传感器协助语音关键字检测的流程图。

[0010] 图4~图7为根据本发明实施方式的传感器协助语音关键字检测的操作示例。

[0011] 图8为根据本发明实施方式的修改图1中的活动数据库的流程图。

[0012] 图9为根据本发明实施方式的语音关键字检测系统的示意图。

[0013] 图10为图9中的语音检测器的实施方式的示意图。

[0014] 图11为低功率传感器协助语音关键字检测的流程图。

具体实施方式

[0015] 为增强使用者对周围环境的认知(awareness)和互动(interactivity),现代电子装置具有许多传感器。举例来说,通讯装置(例如,智能手机)的传感器包括加速度计

(accelerometer)、陀螺仪(gyroscope)、磁强计(magnetometer)、气压计(barometer)、接近传感器(proximity sensor)、光传感器(light sensor)、触屏、定位系统的接收器、无线接收器及/或照相机等等。

[0016] 由装置的传感器提供的传感器数据有利于得到人机交互(user-device interaction)、使用者状态(user status)及/或装置周围环境情况的活动信息(activity information)。人机交互的活动信息包括:(a)装置是否被抬高、降低、举起、放下、翻转、扔下(drop)、摇晃、稳固地拿着、倾斜、平放、靠近某物(例如,使用者)、远离某物、及/或放置于黑暗的环境中(例如,放置于袋子或背包中)或有光的环境中,等等;(b)传入事件(incoming event)代表装置是否需要与使用者交互。例如,装置是否接收到来电(incoming call)、短消息、及/或电子邮件;及/或装置是否将要报警预定义时刻,例如,用于即将到来的待办事项,日历、记事册、及/或时刻表中的会议清单的叫醒业务、叫醒电话、闹铃、提示、弹出式屏幕(screen pop-up)。使用者状态的活动信息包括使用者是否步行、跑步、及/或驾驶等。装置周围环境的活动信息包括:环境温度、噪声、亮度、位置、状态和当前时间。

[0017] 在此实施方式中,使用者的语音关键字(命令)通常发生在可识别模式(recognizable pattern)的活动之后,及/或使用者的语音关键字(命令)伴随可识别模式的活动而发生。以移动电话为例:当由于来电电话响铃时,使用者的自然反应是首先举起移动电话以观看相关信息(例如,来电者是谁),然后决定如何答复(例如,接电话或不理会/拒绝电话)。因此,举起的活动是提示移动电话等待(expect)语音答复(responding)关键字(命令)的线索。在另一实施方式中,当使用者想要通过移动电话的照相功能来照相时,使用者的自然反应是首先保持移动电话稳定,然后指示移动电话拍照。因此,保持稳定的活动提供关于何时等待语音拍照关键字的线索。

[0018] 较佳地,可利用传感器数据以指示已知模式的活动是否发生,以及通过提供附加的信息相应地提高语音关键字检测,例如,通过预测使用者将要说出语音关键字的时间。举例来说,关键字是用于询问搜索引擎(querying search engine)及/或人工智能引擎(artificial intelligence engine)的命令、指示、术语,及/或关键字是告知性的(informative)语音,例如,“耶哈!(Yee-Ha!)”。尽管“耶哈!”不是官方词汇。

[0019] 举例来说,当由于来电电话响铃时以及侦测到举起电话的活动时,电话可以相应地预测到使用者将要发出语音答复关键字(例如,“接听”或“拒绝”),并因此调整语音关键字检测的灵敏度以容易识别之后发出的语音答复关键字(例如,“接听”或“拒绝”)。举例来说,当移动电话被切换至照相机功能以及侦测到保持稳定的活动时,移动电话可以期望语音拍照关键字(例如“拍照”),以触发照片拍照,然后调整语音关键字检测的灵敏度以容易识别语音拍照关键字。

[0020] 因此,根据本发明的活动预测,可以加强语音关键字检测,其中活动预测可设计为利用传感器数据并相应地检测指示活动的发生,指示活动将在使用者将要发出语音关键字时或之前发生。此外,语音关键字检测和活动预测可执行于应用方案(application scenario)的情境中。在此实施方式中,当移动电话运行通信应用以用于来电响铃时,活动预测检测第一套相关指示活动(indicative activities)(例如,举起移动电话)的发生,以及语音关键字检测识别第一套相关语音关键字。举例来说,语音答复关键字(例如,“接听”或“拒绝”)。当移动电话运行照相机应用时,活动预测检测第二套相关指示活动(例如,保持

稳定)的发生,以及语音关键字检测识别第二套相关语音关键字,举例来说,语音拍照关键字(例如,“拍照”)。

[0021] 存在有两种降低语音关键字检测的性能的错误,这两种错误包括遗漏错误(miss error)和假报警错误(false alarm error)。当使用者发出语音关键字但是语音关键字检测未能够识别说出的语音关键字时,遗漏错误发生。当使用者没有发出任何语音关键字但是语音关键字检测错误地确定已经发出语音关键字时,假报警错误发生。

[0022] 语音关键字检测可以调整灵敏度(或识别趋势)以协调(trade-off)遗漏错误和假报警错误。增加灵敏度可使得语音关键字检测趋向于将传入的声音解释为语音关键字,甚至当传入的声音不太可能包括语音关键字时也如此。因此,增加灵敏度降低了遗漏错误的概率而提高了假报警错误的概率。另一方面,减少灵敏度可降低语音关键字检测将传入的声音解释为语音关键字的倾向,甚至在传入的声音很有可能包括语音关键字的时候。因此,减少灵敏度提高了遗漏错误的概率而降低了假报警错误的概率。

[0023] 在此实施方式中,语音关键字检测的灵敏度是自适应的并根据活动预测动态调整,因此,可以抑制遗漏错误和假报警错误而不用折衷(compromise)处理。当检测到指示活动时,可以提高识别相关语音关键字的灵敏度,以使得传入的声音更有可能地被识别为相关语音关键字,即使说出的关键字很模糊、不清楚、及/或充满噪声;因此,可以抑制遗漏错误。另一方面,在不存在指示活动时,可以降低语音关键字检测的灵敏度,以大幅降低传入的声音错误地识别为语音关键字而抑制假报警错误的可能性。

[0024] 请参考图1。图1为根据本发明实施方式的语音关键字检测系统12的示意图。语音关键字检测系统12可集成于装置10中。装置10可以是消费电子。例如通信装置(例如,移动电话、手机、智能机)、便携式计算机(hand-held or portable computer)、声学互动(acoustically interactive)及/或可控的个人保健(health-care)装置、玩具、可佩戴装置(例如,手表)、电视、媒体播放器、媒体记录器(例如,数字相机或摄影机)、导航器或定位装置等。在此实施方式中,甚至可在没有使用者提示的情况下,语音关键字检测器14自动监控传入的声音是否包括语音关键字。

[0025] 为改善语音关键字检测器14的性能,语音关键字检测系统12进一步包括关键字数据库16、活动预测器18、活动数据库20和决策器22。关键字数据库16耦接于语音关键字检测器14,并且关键字数据库16包括关键字清单KL[1]~关键字清单KL[N2],其中数目N2大于或等于1。较佳地,每一个关键字清单KL[j](其中,j=1~N2)与目标应用app[j]相关,并且每一个关键字清单KL[j]包括候选关键字kw[j,1]~候选关键字kw[j,P{j}],其中,数目P{j}大于或等于1。不同的关键字清单与不同的目标应用相关,并且不同的关键字清单的候选关键字的数目可以不同。也就是说,由于索引j1不等于索引j2,关键字清单KL[j1]的目标应用app[j1]与关键字清单KL[j2]的目标应用app[j2]不同;关键字清单KL[j1]的数目P{j1}与关键字清单KL[j2]的数目P{j2}可以不同,或者关键字清单KL[j1]的数目P{j1}与关键字清单KL[j2]的数目P{j2}也可以相等。

[0026] 活动数据库20耦接于活动预测器18,并且活动数据库20包括活动清单AL[1]~活动清单AL[N1],其中数目N1大于或等于1。每一个活动清单AL[i](其中,i=1~N1)与目标应用app[j]相关,并且每一个活动清单AL[i]包括活动模板at[i,1]~活动模板at[i,Q{i}](其中,数目Q{i}大于或等于1)。不同的活动清单与不同的目标应用相关,并且不同的活动

清单的活动模板的数目可以不同。

[0027] 语音关键字检测器14接收语音信号Snd。举例来说,装置10包括麦克风及/或麦克风阵列(图未示)以收集声音,并且通过处理(例如,模拟-数字转换)收集的声音而相应地提供数字的语音信号Snd。根据本发明的另一实施方式,装置10从另一个远程装置(例如,无线麦克风,图未示)接收远程提供的信号(图未示),该远程提供的信号包括编码或未编码的声音,以及通过处理远程提供的信号相应地提供语音信号Snd。

[0028] 根据装置10当前的运行应用,语音关键字检测器14也可获得关键字数据库16中的关键字清单KL[1]~关键字清单KL[N2]选择出的选择的关键字清单KL[jx](其中,jx=1~N2),其中选择的关键字清单KL[jx]的目标应用app[jx]与装置10当前运行的应用匹配。举例来说,语音关键字检测器14及/或关键字数据库16可以从装置10及/或装置10的操作系统(operation system,OS)存取指示装置10当前运行应用的状态,然后查找目标应用app[jx]等于装置10当前运行应用的关键字清单KL[jx]。装置10运行的应用涉及由操作系统控制而执行的实用程序、服务、过程、及/或子程序。当前运行的应用可涉及当前活动的应用、前景应用、背景应用。

[0029] 通过选择与当前运行的应用相对应的关键字清单,可以参照情境自适应地执行语音关键字检测。举例来说,关于负责处理来电的通信应用,对应的关键字清单可以包括答复关键字的候选,例如“接听”或“拒绝”。对于照相机应用,它的对应的关键字清单可以包括拍照关键字的候选,例如“拍照”。

[0030] 响应于语音信号Snd,语音关键字检测器14根据选择的关键字清单KL[jx]提供初步的关键字检测结果Skw。举例来说,语音关键字检测器14将语音信号Snd中的传入的声音与选择的关键字清单KL[jx]的候选关键字kw[jx,1]~候选关键字kw[jx,P{jx}]中的每一个进行比较以提供初步的关键字检测结果。

[0031] 活动预测器18接收由多个传感器提供的传感器数据信号Ssd。举例来说,装置10包括传感器以检测移动、加速、位置、状态、角度方向(angular direction)/姿态(例如,翻转或倾斜)、周围的音量、亮度及/或施加在装置10上的力场(例如,磁、电磁及/或重力场)并输出传感器数据信号Ssd。根据本发明的另一实施方式,装置10可以从另一个远程装置(例如,远程传感器,图未示)接收远程提供的信号(图未示),其中远程提供的信号包含传感器数据、嵌入传感器数据及/或编码传感器数据,并通过处理远程提供的信号而相应提供传感器数据信号Ssd。

[0032] 根据装置10当前运行的应用程序,活动预测器18可以从活动数据库20的活动清单AL[1]~活动清单AL[N1]中获得选择的活动的活动清单AL[ix](其中,ix=1~N1),其中选择的活动的活动清单AL[ix]的目标应用app[ix]表示装置10当前正在运行的应用。举例来说,活动预测器18及/或活动数据库20从装置10及/或装置10的操作系统获得指示装置10当前运行应用的状态,然后选择与指示装置10当前运行的应用的目标应用app[ix]相关的活动清单AL[ix]。通过选择与当前运行的应用相对应的活动清单,可以根据情境自适应地执行活动检测。举例来说,关于负责处理来电的通信应用,对应的活动清单可以包括举起电话的活动模板;对于照相机应用,它的对应的活动清单可以包括保持稳定的活动模板。

[0033] 根据选择的活动的活动清单AL[ix],活动预测器18处理传感器数据信号Ssd以提供指示使用者是否将给出语音关键字的概率的活动预测结果Sap。举例来说,根据本发明的一个实

施方式,活动预测器18将传感器数据信号Ssd与记录在选择的活动清单AL[ix]的活动模板at[ix,1]~活动模板at[ix,Q{ix}]中的每一个进行比较,并相应地提供活动匹配结果作为活动预测结果Sap。根据本发明的另一实施方式,活动预测器18通过传感器数据信号Ssd提取特征以提供已提取的数据,以及活动预测器18将已提取的数据与选择的活动清单AL[ix]的活动模板at[ix,1]~活动模板at[ix,Q{ix}]中的每一个进行比较以提供该活动预测结果Sap。

[0034] 根据本发明的一个实施方式,每一个活动模板at[i,q](其中,q=1~Q{i})可以包括指示活动(动作或状态)的标准的、典型的、有代表性的及/或最频繁的检测结果,其中指示活动在使用者将要发出语音关键字的时候或之前产生。每一个检测结果(sensed result)与传感器相关,并且每一个检测的结果作为参考资料(reference)而在活动模板at[i,q]中记录下来。举例来说,当活动预测器18通过将传感器数据信号Ssd与选择的活动清单AL[ix]的每一个活动模板at[ix,q]进行比较以产生活动预测结果Sap时,对给定种类的传感器的包括于传感器数据信号Ssd中的每一个检测结果来说,活动预测器18检测活动模板at[ix,q]是否包括与该种类的传感器相关的参考资料;若活动模板at[ix,q]包括与该种类的传感器相关的参考资料,则活动预测器18将传感器数据信号Ssd中的检测结果与活动模板at[ix,q]中的参考资料进行比较,然后将比较结果在活动预测结果Sap中反映。

[0035] 根据本发明的实施方式,每一个活动模板at[i,q](其中,q=1~Q{i})包括提取的参考资料(extracted reference),每一个提取的参考资料与一个传感器相关,代表指示活动的检测结果的提取特征。当活动预测器18通过将传感器数据信号Ssd与选择的活动清单AL[ix]的每一个活动模板at[ix,q]进行比较以产生活动预测结果Sap时,活动预测器18首先提取包括于传感器数据信号Ssd中的每一个检测结果的特征以相应地产生提取的检测结果(图未示);然后,对给定种类的传感器的包括于传感器数据信号Ssd中的每一个提取的检测结果来说,活动预测器18查找活动模板at[ix,q]是否包括该种类的传感器的提取的参考资料;若活动模板at[ix,q]包括该种类的传感器的提取的参考资料,则活动预测器18将传感器数据信号Ssd中的提取的检测结果与活动模板at[ix,q]中的提取的参考资料进行比较,然后将比较结果输出于活动预测结果Sap中。

[0036] 通过滤波(例如,低通滤波)检测结果,计算检测结果的统计数字,及/或将检测结果转换至频谱域(spectrum domain)可提取检测结果的特征。请参考图2。图2为传感器数据的检测结果和提取的检测结果的曲线。通过加速度计检测检测结果。如图2所示,传感器数据曲线s1x、传感器数据曲线s1y和传感器数据曲线s1z分别代表在翻转活动期间检测的x-轴、y-轴和z-轴的加速度的检测结果,并共同形成加速度计的检测结果。提取传感器数据曲线s1x、传感器数据曲线s1y和传感器数据曲线s1z的特征(例如,低通滤波)分别获得提取的传感器数据曲线e1x、提取的传感器数据曲线e1y和提取的传感器数据曲线e1z以共同形成提取的检测结果。相似地,传感器数据曲线s2x、传感器数据曲线s2y和传感器数据曲线s2z分别代表在保持稳定的活动期间检测的x-轴、y-轴和z-轴的加速度的检测结果。通过提取传感器数据曲线s2x、传感器数据曲线s2y和传感器数据曲线s2z的特征,可以分别获得提取的传感器数据曲线e2x、提取的传感器数据曲线e2y和提取的传感器数据曲线e2z。每一个检测结果和提取的检测结果的特征可以用于区别不同的活动。举例来说,可以在传感器数据曲线s1z或提取的传感器数据曲线e1z中观察到幅度约+9.8至幅度-9.8的变化穿越

(transition);相应地,例如转换持续时间(例如,时间间隔Td或时间间隔Te)、过零时间及/或变化的幅度差异的特征可以用于识别装置是否被使用者扔下(drop)或翻转。

[0037] 请重新参考图1。决策器22耦接于活动预测器18和语音关键字检测器14,能处理活动预测结果Sap和初步的关键字检测结果Skw以提供增强的关键字检测结果Sdm,使得装置10根据增强的关键字检测结果Sdm进行回应。由于增强的关键字检测结果Sdm是通过处理活动预测结果Sap和初步的语音关键字检测结果Skw来获得,增强的关键字检测结果Sdm比初步的关键字检测结果Skw更可靠,并因此提高了语音关键字检测器14的语音关键字检测。

[0038] 请一并参考图1和图3。图3为根据本发明实施方式的传感器协助语音关键字检测的流程图。图1中的语音关键字检测系统12可以采用流程100来提高语音关键字检测器14的性能。流程100包括如下步骤。

[0039] 步骤102:识别装置10当前运行的应用。如上所述,语音关键字检测系统12可以存取装置10的状态以识别当前运行的应用。举例来说,可以通过装置10的操作系统,及/或通过控制装置10的中央处理器的(central processing unit,CPU,图未示)的寄存器来提供状态。

[0040] 步骤104:分别从活动数据库20和关键字数据库16中选择相应的活动清单AL[ix]和相应的关键字清单KL[jx]。举例来说,通过将当前运行的应用与每一个活动清单AL[i](其中,i=1~N1)的每一个目标应用app[i]进行比较选择相应的活动清单AL[ix],以及通过将当前运行的应用与每一个关键字清单KL[j](其中,j=1~N2)的每一个目标应用app[j]进行比较选择相应的关键字清单KL[jx]。

[0041] 步骤106:通过语音关键字检测器14,根据语音信号Snd与关键字清单KL[jx]执行初步的语音关键字检测以提供初步的关键字检测结果Skw。举例来说,语音关键字检测器14可以分别将图1中的语音信号Snd与候选关键字kw[jx,1]~候选关键字kw[jx,P{jx}](其中,jx=1~N2,数目P{jx}大于或等于1)进行比较以获得比较结果scr[1]~比较结果scr[P{jx}],每一个比较结果scr[p](其中,p=1~P{jx})在数值上指示语音信号Snd与候选关键字kw[jx,1]~候选关键字kw[jx,P{jx}]之间的相似性,及/或指示语音信号Snd与候选关键字kw[jx,1]~候选关键字kw[jx,P{jx}]匹配的概率或可能。语音关键字检测器14也可以获得额外的比较结果scr[P{jx}+1]以在数值上指示语音信号Snd与候选关键字kw[jx,1]~候选关键字kw[jx,P{jx}]匹配失败的概率或可能。即,指示语音信号Snd为“无关键字”的概率。然后,语音关键字检测器14可以找出比较结果scr[1]~比较结果scr[P{jx}+1]的最大值并相应提供初步的关键字检测结果Skw。为方便详述,比较结果scr[1]~比较结果scr[P{jx}+1]的最大值可以记作比较结果scr[p_max],其中索引p_max为1~(P{jx}+1)中的一个,并且对应于比较结果scr[p_max]的候选关键字kw[jx,p_max]是指最可能的关键字;若索引p_max等于(P{jx}+1),则最可能的候选关键字kw[jx,p_max]是指“无关键字”的情况。

[0042] 为获得比较结果scr[1]~比较结果scr[P{jx}+1],图1所示的语音关键字检测器14可采用语音处理算法的关键字映射器(mapper),例如,隐马尔可夫模型(hidden Markov model,HMM)等。每一个比较结果scr[1]~比较结果scr[P{jx}+1]具有线性分数(linear score)或似然对数(logarithm likelihood)。

[0043] 步骤108:图1所示的活动预测器18根据选择的的活动清单AL[ix]和传感器数据信号Ssd执行活动预测。在此实施方式中,活动预测器18根据选择的的活动清单AL[ix]和传感器数

据信号Ssd中的传感器数据,计算活动预测结果Sap以指示使用者是否将要发出语音关键字的概率或可能性。举例来说,活动预测器18将传感器数据信号Ssd与选择的清单AL[ix]的每一个活动模板at[ix,q] (其中, $q=1 \sim Q\{ix\}$) 进行比较,并相应提供活动匹配比较结果acr[q]以在数值上指示活动模板at[ix,q]与传感器数据信号Ssd的相似性,及/或传感器数据信号Ssd与活动模板at[ix,q]匹配的概率或可能性。然后,收集比较结果acr[1]~比较结果acr[Q{ix}]以形成活动预测结果Sap,例如,活动预测结果Sap包括比较结果acr[1]~比较结果acr[Q{ix}]。由于传感器数据信号Ssd包括不同传感器的检测结果(或可以提供不同传感器的提取的检测结果),以及每一个活动模板at[ix,q]包括不同的传感器的参考资料(references) (或提取的检测参考资料),活动预测器18可将传感器的每一个检测结果(或每一个提取的检测结果)与传感器的参考资料(或提取的参考资料)进行比较以获得各自的比较结果,并可根据不同传感器的比较结果产生匹配比较结果acr[q]。

[0044] 为获得比较结果acr[1]~比较结果acr[Q{ix}],活动预测器18可采用基于高斯混合模型(Gaussian mixture model)、隐马尔可夫模型(hidden Markov model)、支持向量机(support vector machine)、及/或神经网络(neural network)等的基于规则(ruled-based)的算法或更加复杂(sophisticated)的算法。根据本发明的另一实施方式,活动预测器18采用基于动态时间规整(dynamic time warping)的相似的测量算法(measurement algorithm)等。步骤106和步骤108可以同时执行或顺序执行。

[0045] 步骤110:通过图1所示的决策器22,根据初步的关键字检测结果Skw和活动预测结果Sap做出最终决策,以提供增强的关键字检测结果。增强的关键字检测结果Sdm可以是测量步骤106中最可能的候选关键字kw[jx,p_max]在当前运行应用和使用者活动的内容中的可能性的概率、百分数(score)及/或似然对数(logarithm likelihood)。

[0046] 装置10周期性的重复流程100以执行传感器协助语音关键字检测。在另一实施方式中,当必要时(例如当使用者指示时),装置10执行流程100。

[0047] 根据本发明的一个实施方式,通过初步的关键字检测结果Skw是否满足第一条件和活动预测结果Sap是否满足第二条件而获得增强的关键字检测结果Sdm。举例来说,若初步的关键字检测结果Skw大于第一阈值,则满足第一条件;以及在活动预测结果Sap中的比较结果acr[1]~比较结果acr[Q{ix}]中的每一个大于第二阈值,则满足第二条件。根据本发明的另一实施方式,第二条件是:若比较结果acr[1]~比较结果acr[Q{ix}]的和(或线性组合)大于第二阈值。根据本发明的另一实施方式,第二条件是:若比较结果acr[1]~比较结果acr[Q{ix}]的统计性能(例如,最大值、最小值、平均值等)大于第二阈值。较佳地,当第一条件和第二条件均满足时,图1中的决策器22确定最可能的候选关键字kw[jx,p_max]表明具有肯定的增强的关键字检测结果Sdm(即确定具有最可能的关键字),这样装置10相应做出答复,例如,执行与最可能的候选关键字kw[jx,p_max]有关的操作。另一方面,若第一条件和第二条件中的一个不满足(确定的最可能的候选关键字kw[jx,p_max]代表“无关键字”),决策器22则否定增强的关键字检测结果Sdm,因此,装置10的操作不受影响。

[0048] 根据本发明的一个实施方式,通过计算比较结果acr[1]~比较结果acr[Q{ix}]和初步的关键字检测结果Skw的线性组合,以及比较该线性组合是否大于预定阈值,而获得增强的关键字检测结果Sdm。若该线性组合大于预定阈值,则决策器22确定得到最可能的候选关键字kw[jx,p_max],否则,决策器22确定没有识别出最可能的候选关键字kw[jx,p_max]。

[0049] 根据本发明的一个实施方式,决策器22采用概率融合算法,例如,D-S证据理论(Dempster-Shafer theory),或机器学习算法(例如,高斯混合模型(Gaussian mixture model)),以处理初步的关键字检测结果Skw和比较结果acr[1]~比较结果acr[Q{ix}]并相应地取得更加可靠的增强的关键字检测结果Sdm。根据本发明的另一实施方式,计算比较结果acr[1]~比较结果acr[Q{ix}]和初步的关键字检测结果Skw的线性组合而作为增强的关键字检测结果Sdm。由于数目Q{ix}在不同的应用内容下不同,上述的复杂的算法应用任意随机数(arbitrary number of probabilities)作为输入并相应地提供确定的(conclusive)概率作为输出,以使得能提供分别由初步的关键字检测结果Skw和比较结果acr[1]~比较结果acr[Q{ix}]提供的整体信息的灵活解(flexible solution)。根据本发明的另一实施方式,可以结合不同的算法以产生增强的关键字检测结果Sdm。举例来说,决策器22采用第一算法处理比较结果acr[1]~比较结果acr[Q{ix}]以相应获得第一结果,并采用第二算法处理第一结果和初步的关键字检测结果Skw以相应获得增强的关键字检测结果Sdm。

[0050] 请一并参考图1、图4~图7。图4~图7为根据本发明实施方式的传感器协助语音关键字检测的操作示例。传感器协助语音关键字检测实现于图1所示的装置10中。将装置10为用于图4~图7的智能手机。图4显示拍照的方案。当使用者需要用装置10来拍照时,使用者自然、方便、直观的动作序列(action sequence)为:激活装置10的照相机应用,保持装置10稳定,并发出语音拍照关键字,例如,“拍照”。相应地,为实现传感器协助语音控制照片拍照,图1所示的对应于照相机应用的活动清单包括代表保持稳定的指示活动的活动模板,以及对应于照相机应用的关键字清单包括候选的语音拍照关键字。

[0051] 传感器协助语音控制照片拍照操作将做如下详述。当使用者激活装置10的照相机应用以准备照相时,图1所示的装置10的活动预测器18利用传感器数据侦测保持稳定的指示活动是否发生。当使用者保持装置10稳定,活动预测器18会反映指示活动的发生,这样装置10可以预测使用者将要发出语音拍照关键字以触发照相,并且装置10允许语音拍照关键字更加容易地识别。举例来说,图1所示的决策器22增加识别语音拍照关键字的灵敏度,以使得即使当语音拍照关键字发音不清晰或者智能手机处于噪声环境中时,也能识别语音拍照关键字。一旦决策器22识别出语音拍照关键字,装置10回应以进行拍照动作。相反地,当没有检测到保持稳定的指示活动,使用者不太可能发出语音拍照关键字,这样装置10可以避免语音拍照关键字的错误识别。举例来说,决策器22降低识别语音拍照关键字的灵敏度。

[0052] 当装置10接收来电并且装置10响铃以引起使用者注意时,使用者自然的动作序列为:举起装置10以观看来电信息(例如,谁来电),确定如何答复(例如,接电话或不理会/拒绝电话),并相应发出语音答复关键字,例如,“接听”、“拒绝”、“忽视”或静音”。图5显示来电后使用者发出语音答复关键字“接听”的实施例。根据本发明的实施方式,为实现传感器协助语音控制来电答复,对应于负责处理来电的通信应用的活动清单包括记录举起电话的指示活动的活动模板,以及对应于通信应用的关键字清单包括候选语音答复关键字,例如,“接听”、“拒绝”、“忽视”或“静音”。举例来说,当使用者在开会并且甚至在发言时,他可以说“静音”以快速地在整个会议中静音装置10。

[0053] 传感器协助语音控制来电答复操作将做如下详述。当处理来电的应用接收来电时,活动预测器18指示检测举起电话的指示活动是否发生。当使用者举起装置10时,活动预

测器18反映指示活动的发生,这样装置10可以预测使用者将要发出语音答复关键字,以及允许更容易地识别语音答复关键字。举例来说,决策器22降低确认语音答复关键字的肯定识别的阈值,这样即使当语音拍照关键字发音不清晰(said faintly)或者智能手机处于噪声环境中时,也能识别语音答复关键字。一旦决策器22反映识别出语音答复关键字,装置10可以相应地反应,例如,通过转接(patch through)、拒绝、忽视、或静音电话。相反地,当没有检测到举起电话的指示活动时,使用者不太可能发出语音答复关键字,这样装置10可以避免语音答复关键字的错误识别。举例来说,决策器22增加确认识别语音答复关键字的阈值。

[0054] 图6显示呼出通话的方案。当使用者想要进行呼出通话时,直觉、方便的动作序列为:举起装置10至耳边,并给出受话人的姓名,例如,“打给玛丽”。为实现传感器协助语音控制拨打电话,对应于负责处理呼出通话的通信应用的活动清单包括记录举起电话至耳边的指示活动的活动模板,以及对应于通信应用的关键字清单包括与可能的被叫对象的姓名有关的候选语音呼叫关键字,例如,在地址簿及/或联系清单中列出的姓名。

[0055] 传感器协助语音控制呼叫的操作将做如下详述。当使用者激活通信应用准备呼出通话时,或当没有其他应用运行并且装置10自动运行通信应用作为默认应用时,活动预测器18指示检测举起电话的指示活动是否发生。当使用者确实将装置10举起至耳边时,活动预测器18反映指示活动的发生,这样装置10可以预测使用者将要发出语音呼叫关键字,以及允许更容易地识别语音答复关键字。举例来说,决策器22增加确认语音答复关键字的肯定识别的趋势,这样即使当语音答复关键字发音不清晰或者智能手机处于噪声环境中时,语音呼叫关键字也可以被识别。一旦决策器22反映识别出语音呼叫关键字,装置10根据语音呼叫关键字进行呼叫。另一方面,当没有检测到指示活动时,使用者不太可能发出语音呼叫关键字,这样装置10可以避免语音呼叫关键字的错误识别。即决策器22可以降低确认识别语音呼叫关键字的趋势。

[0056] 图7显示唤醒睡眠模式的电话及/或解锁锁定电话的方案。为节省功率,现代智能电话可以进入睡眠模式。在睡眠模式中,仅维持基本功能,以及关闭其他功能(例如,屏幕显示)。为方便使用者,希望通过使用者的语音而不是触摸屏幕或者按键而将电话从睡眠模式中唤醒。因此,使用者可以通过发出语音唤醒关键字(例如,“醒来”)唤醒睡眠的电话。为相应地实现传感器协助语音控制电话唤醒,对应于负责处理电话唤醒的待机应用的活动清单包括记录指示装置10保持稳定、安静和手持的状态的活动模板;以及对应于待机应用的关键字清单包括候选语音唤醒关键字,例如,“醒来”。

[0057] 传感器协助语音控制唤醒电话操作将做如下详述。当装置10进入睡眠模式并且待机应用在运行时,活动预测器18指示检测任何指示状态(indicative states)是否发生。当装置10进入一个指示状态时,活动预测器18反映进入指示状态,这样装置10等待语音唤醒关键字,以及允许更加容易识别语音唤醒。举例来说,决策器22增加确认语音唤醒关键字的肯定识别的趋势,这样即使当语音唤醒关键字发音不清晰或者智能手机处于噪声环境中时,语音唤醒关键字可以被识别。一旦决策器22反映识别出语音唤醒关键字,装置10离开唤醒模式。另一方面,当没有检测到指示状态时,例如,当装置10在背包中时,使用者不太可能发出语音唤醒关键字,这样,装置10可以避免语音唤醒关键字的错误识别。即决策器22可以增加拒绝或忽视语音唤醒关键字的趋势。

[0058] 虽然图4~图7仅解释了有限的实施方式,但是需要注意的是,本发明的传感器协助语音关键字检测实际上提供了一种完美融合在使用者日常生活中的自动(unattended)、自然、方便、直观、友好、准确并可靠的语音控制功能。举例来说,使用者通过图7所示的语音唤醒装置,然后通过图4所示的语音拍照,通过图6所示的语音呼出电话,及/或通过图5所示的语音答复来电。

[0059] 在图1所示的活动数据库20中的活动清单和对应的活动模板能够通过装置10的制造商和设计者预先建立。进一步,活动数据库20可以根据使用者的个人需要而进行更新。举例来说,语音关键字检测系统12可以允许使用者增加、删除、编辑、移动及/或替代活动模板和活动清单。请一并参考图1和图8。图8为根据本发明实施方式的修改图1中的活动数据库20的流程200。装置10可以采用流程200以通过增加新的活动模板或修改已经存在的活动模板来更新活动数据库20。流程200包括如下步骤。

[0060] 步骤202:当使用者需要手动更新活动数据库20时,通过装置10进入训练模式(training mode)以准备使用者修改操作。然后装置10提示使用者指定一个需要修改的元素(例如,活动清单及/或活动模板),并提示使用者指定怎样修改(例如,增加或替换)元素。当装置10进入训练模式时,装置10可以先停止图3所示的流程100。

[0061] 步骤204:当使用者准备将新的活动增加至活动数据库20中作为新的活动模板时,或使用者准备将新的活动代替存在的活动模板(即至少一个活动模板)时,通过装置10收集传感器数据。根据本发明的实施方式,装置10可以提取传感器数据的特征,例如,通过活动预测器18从传感器数据信号Ssd中提取。

[0062] 步骤206:检测是否建立可靠的新的活动模板。为建立统计上的可靠的活动模板,装置10提示使用者设定重复新的活动的数个时间;当使用者重复新的活动时,装置10重复步骤204。当装置10收集到充分的传感器数据以建立可靠的新的活动模板时,装置10进入步骤208。若收集的数据不符合要求,则流程200重复步骤204。

[0063] 步骤208:装置10根据收集的传感器数据更新活动数据库,例如,增加新的活动模板或用新的活动模板替换存在的活动模板。

[0064] 步骤210:退出训练模式,然后装置10恢复停止的流程100,或进入其他模式。

[0065] 除活动数据库20之外,使用者也可以根据与流程200类似的流程而修改关键字数据库16。

[0066] 由于期望语音关键字检测不需要使用者的暗示而在任何时候监控关键字,也需要考虑功率消耗,尤其是依靠电池供电的移动装置。语音检测可用于评估传入的声音的信息性如何,以使得当传入的声音看起来信息丰富时,使能语音关键字检测;否则禁能语音关键字检测以省电。

[0067] 请一并参考图1和图9。图9为根据本发明实施方式的语音关键字检测系统32的示意图。在图9中,语音关键字检测系统32为低功率传感器协助语音关键字检测系统,并集成于装置30中。与图1所示的语音关键字检测系统12相似,图9中的语音关键字检测系统32包括语音关键字检测器44、关键字数据库46、活动预测器48、活动数据库50以及决策器52。进一步,语音关键字检测系统32包括语音检测器34。装置30包括多路复用器36、模数转换器38(在图9中标示为ADC)、端点40a和端点40b。若装置30具有模拟麦克风(图未示),模拟麦克风可连接至端点40a,这样由模拟麦克风提供的模拟声音信号可以传送至模数转换器38以将

模拟声音信号转换为数字声音信号,然后通过多路复用器36输出数字的语音信号Snd。另一方面,若装置30具有数字麦克风,麦克风可连接至端点40b。如此一来,数字麦克风提供的数字声音信号,并通过多路复用器36输出数字的语音信号Snd。语音关键字检测系统32也可以接收由装置30的传感器提供的传感器数据信号Ssd。

[0068] 语音关键字检测器44、活动预测器48和决策器52的操作与图1所示的语音关键字检测器14、活动预测器18和决策器22相似。通过图3所示流程100,语音关键字检测器44、活动预测器48、决策器52、关键字数据库46和活动数据库50之间的协作可以实现传感器协作语音关键字检测。

[0069] 关键字数据库46耦接至语音关键字检测器44,以及关键字数据库46包括关键字清单KL[1]~关键字清单KL[N2],其中关键字清单的数目为N2。关键字清单KL[j] (其中,j=1~N2) 与目标应用app[j]相关,以及目标应用app[j]包括候选关键字kw[j,1]~候选关键字kw[j,P{j}],其中候选关键字的数目为P{j}。

[0070] 活动数据库50耦接至活动预测器48,以及活动数据库50包括活动清单AL[1]~活动清单AL[N1],其中活动清单的数目为N1。活动清单AL[i] (其中,i=1~N1) 与目标应用app[i]相关,以及目标应用app[i]包括活动模板at[i,1]~活动模板at[i,Q{i}],其中活动模板的数目为Q{i}。

[0071] 语音关键字检测器44从关键字数据库46的关键字清单KL[1]~关键字清单KL[N2]中获得选择的关键字清单KL[jx],其中选择的关键字清单KL[jx]与目标应用app[jx]相关,并且目标应用app[jx]与装置30的当前运行的应用相匹配。为响应语音信号Snd,语音关键字检测器44根据选择的关键字清单KL[jx]提供初步的关键字检测结果Skw。

[0072] 活动预测器48从活动数据库50的活动清单AL[1]~活动清单AL[N1]中获得选择的清单AL[ix],其中选择的清单AL[ix]与目标应用app[ix]相关,并且与目标应用app[ix]相关装置30的当前运行的应用相匹配。基于选择的清单AL[ix],活动预测器48处理传感器数据信号Ssd以提供指示使用者是否将要发出语音关键字的概率的活动预测结果Sap。

[0073] 决策器52耦接至活动预测器48和语音关键字检测器44,用于处理活动预测结果Sap和初步的关键字检测结果Skw以提供增强的关键字检测结果Sdm,这样装置30可以根据增强的关键字检测结果Sdm作出反应。

[0074] 语音检测器34耦接至语音关键字检测器44,以及语音检测器34能够基于语音信号Snd评估信息量(例如,信噪比(signal-to-noise ratio,SNR)),并相应地确定是否使能语音关键字检测器44。举例来说,若语音信号Snd的信息量较低(例如,低于信息量阈值),语音检测器34禁能语音关键字检测器34。举例来说,保持语音关键字检测器44处于低功率(空闲)模式。另一方面,若语音信号Snd的信息量较高(例如,大于信息量阈值),则语音检测器34使能(激活(activate))语音关键字检测器44,举例来说,唤醒语音关键字检测器44操作于正常模式。如图9所示,语音检测器34提供信号Svd以选择性地使能语音关键字检测器44。

[0075] 当禁能语音关键字检测器44时,活动预测器48和决策器52较佳地禁能关键字数据库46和活动数据库50。当使能语音关键字检测器44时,使能活动预测器48和决策器52(以及关键字数据库46和活动数据库50)与语音关键字检测器44协作以用于传感器协助语音关键字检测。根据本发明的另一实施方式,活动预测器48和决策器52也可以接收信号Svd而处于

使能或禁能状态。

[0076] 请参考图10。图10为图9中的语音检测器34的实施方式的示意图。语音检测器34包括语音估计器56a、噪声估计器56b和耦接至语音估计器56a和噪声估计器56b的比较器58。语音估计器56a提供序列 $S[.]$ 以响应语音信号Snd的音量；由于语音信号Snd的每一个采样被标记为数值 $x[n]$ ，语音信号Snd的音量可以被标记为绝对值 $|x[n]|$ 。噪声估计器56b提供序列 $N[.]$ 以响应语音信号Snd的音量。举例来说，可以利用以下的方程实现：

$$[0077] \quad S[n] = a_0 * S[n-1] + (1-a_0) * |x[n]|, 0 < a_0 < 1 \quad (\text{eq1})$$

$$[0078] \quad N[n] = b_0 * N[n-1] + (1-b_0) * |x[n]|, 0 < b_0 < 1 \quad (\text{eq2})$$

[0079] 其中， $S[n]$ 为序列 $S[.]$ 的当前采样， $S[n-1]$ 为序列 $S[.]$ 的先前采样， $|x[n]|$ 为语音信号Snd的当前音量， a_0 为权重； $N[n]$ 为序列 $N[.]$ 的当前采样， $N[n-1]$ 为序列 $N[.]$ 的先前采样， $|x[n]|$ 为语音信号Snd的当前音量， b_0 为权重。

[0080] 在图10的语音估计器56a中应用方程eq1，而使得语音估计器56a通过语音信号Snd的当前音量 $|x[n]|$ 和序列 $S[.]$ 的先前采样 $S[n-1]$ 的加权和而计算序列 $S[.]$ 的当前采样 $S[n]$ ，即通过将乘积 $a_0 * S[n-1]$ 与乘积 $(1-a_0) * |x[n]|$ 相加而计算序列 $S[.]$ 的当前采样 $S[n]$ ，其中乘积 $a_0 * S[n-1]$ 为先前采样 $S[n-1]$ 与权重 a_0 乘积，以及乘积 $(1-a_0) * |x[n]|$ 为当前音量 $|x[n]|$ 与互补权重(complementary weighting) $(1-a_0)$ 的乘积。根据本发明的实施方式，权重 a_0 大于0且小于1。

[0081] 在图10的噪声估计器56b中应用方程eq2，而使得可以通过当前音量 $|x[n]|$ 和序列 $N[.]$ 的先前采样 $N[n-1]$ 的加权和而计算序列 $N[.]$ 的当前采样 $N[n]$ ，即通过将乘积 $b_0 * N[n-1]$ 与乘积 $(1-b_0) * |x[n]|$ 相加而计算序列 $N[.]$ 的当前采样 $N[n]$ ，其中乘积 $b_0 * N[n-1]$ 为先前采样 $N[n-1]$ 与权重 b_0 乘积，以及乘积 $(1-b_0) * |x[n]|$ 为当前音量 $|x[n]|$ 与互补权重 $(1-b_0)$ 的乘积。根据本发明的实施方式，权重 b_0 大于0且小于1。

[0082] 在一个实施方式中，权重 a_0 小于权重 b_0 。相应地，序列 $S[.]$ 趋向于反映传入的声音的当前音量，以及序列 $N[.]$ 趋向于反映传入的声音的过去平均音量。因此，序列 $S[.]$ 指示信息量信号(informative signal)，例如，包括于语音信号Snd中的语音；而序列 $N[.]$ 指示语音信号Snd中的背景噪声。比较器58比较序列 $S[.]$ 和序列 $N[.]$ 以指示语音信号Snd的信息量，并相应提供信号Svd以控制语音关键字检测器44。举例来说，若当前采样 $S[n]$ 和当前采样 $N[n]$ 的差的绝对值 $|S[n]-N[n]|$ 大于信息量阈值(图未示)，然后语音检测器34的比较器58反映在信号Svd中的较高的信息量以使能语音关键字检测器44。相反地，若当前采样 $S[n]$ 和当前采样 $N[n]$ 的差的绝对值 $|S[n]-N[n]|$ 小于信息量阈值，然后由于对噪声执行语音关键字检测只会导致错误，语音检测器34的比较器58反映在信号Svd中的较低的信息量以禁能语音关键字检测器44。

[0083] 请一并参考图9和图11。图11为低功率传感器协助语音关键字检测的流程图。通过语音检测器34，图9中的语音关键字检测系统32采用流程300省电而不降低语音关键字检测期望的功能。流程300的主要步骤描述如下。

[0084] 步骤302：通过语音检测器34检测语音。例如，评估语音信号Snd的信息量。

[0085] 步骤304：确认是否检测到语音。若检测到语音(即信息量较高)，则进入步骤306；否则，进入步骤308。

[0086] 步骤306：使能语音关键字检测，例如，图3中的流程100的传感器协助语音关键字

检测。如图9所示,尽管语音检测器34与语音关键字检测器44,活动预测器48和决策器52一起用于传感器协助语音关键字检测,语音检测器34也可以用于简单的语音关键字检测,简单的语音关键字检测仅通过初步的语音关键字检测器(例如,语音关键字检测器44)来实现而不需要传感器数据的协助。这样的简单的语音关键字检测也可以采用流程300,以及在步骤306中使能语音关键字检测来实现初步的语音关键字检测器的操作。在步骤306之后,流程300重复步骤302。

[0087] 步骤308:禁能语音关键字检测,例如,停止图9中的语音关键字检测器44的操作。活动预测器48和决策器52的操作连同关键字数据库46和活动数据库50的操作一起停止。在步骤308之后,流程300重复步骤302。

[0088] 根据本发明的实施方式,图10中的语音检测器34通过专用硬件来实现。例如,电路。图9中的语音关键字检测器44、活动预测器48及/或决策器52也可以通过专用硬件或通过数字信号处理器(digital signal processor,DSP))来实现。数字信号处理器能够执行软件及/或固件以实现语音关键字检测器44、活动预测器48及/或决策器52的功能。关键字数据库46和活动数据库50通过可编程非易失性存储器(nonvolatile memory)来实现。现代的装置可以包括中央处理器(CPU)和数字信号处理器(DSP)。中央处理器负责执行操作系统(OS)和应用,以及数字信号处理器负责处理多媒体(例如,编码及/或解码音频及/或视频)。在此种架构下,协助语音关键字检测系统的传感器组件可以实现于数字信号处理器之中,这样协助语音关键字检测的传感器可以不需要中央处理器的参与而操作,以及相应地协助语音关键字检测的传感器具有较低的功率消耗。根据本发明的另一实施方式,仅通过专用硬件实现的语音检测器34在大多数的时候运行,以选择性地使能需要的电路。举例来说,当智能手机通过本发明揭露的语音检测增加灵敏度时,智能手机通过具有不同颜色(例如,黄色)的发光二极管在黑暗的屏幕下提示使用者可通过语音控制完成想要的操作。根据本发明的另一实施方式,屏幕可以在全屏的一部分区域中显示指示标志以提示使用者。根据本发明的另一实施方式,利用陀螺仪传感器耗能很小的特性,因此仅陀螺仪传感器可以始终或周期性地操作以选择性地使能要求的电路。根据由活动数据库和关键字数据库定义的多个操作,智能手机可以在黑屏下操作或不需要全部照亮屏幕。因此,整体系统的功率消耗不至于明显增加,甚至可以用本发明揭露的实施方式来减少功率消耗。

[0089] 总之,本发明利用通过装置的传感器提供的传感器数据来获得附加的信息以提高语音关键字检测,以使得更加方便、友好、可靠和准确地实现语音控制。本领域技术人员可能在不脱离本发明的范围内做出修改。举例来说,黑屏触摸控制可以被纳入本发明的范围以进一步提高装置控制的便利。

[0090] 本发明可以在不脱离其精神和本质特征的情况下以其他特殊形式来实现。以上描述的实施方式仅用于解释本发明,而并非限制本发明。本发明的保护范围应以权利要求界定的范围为准,而不以上述描述的实施方式为限。在与权利要求相等的含义和范围内的所有变化均包括于本发明的保护范围。

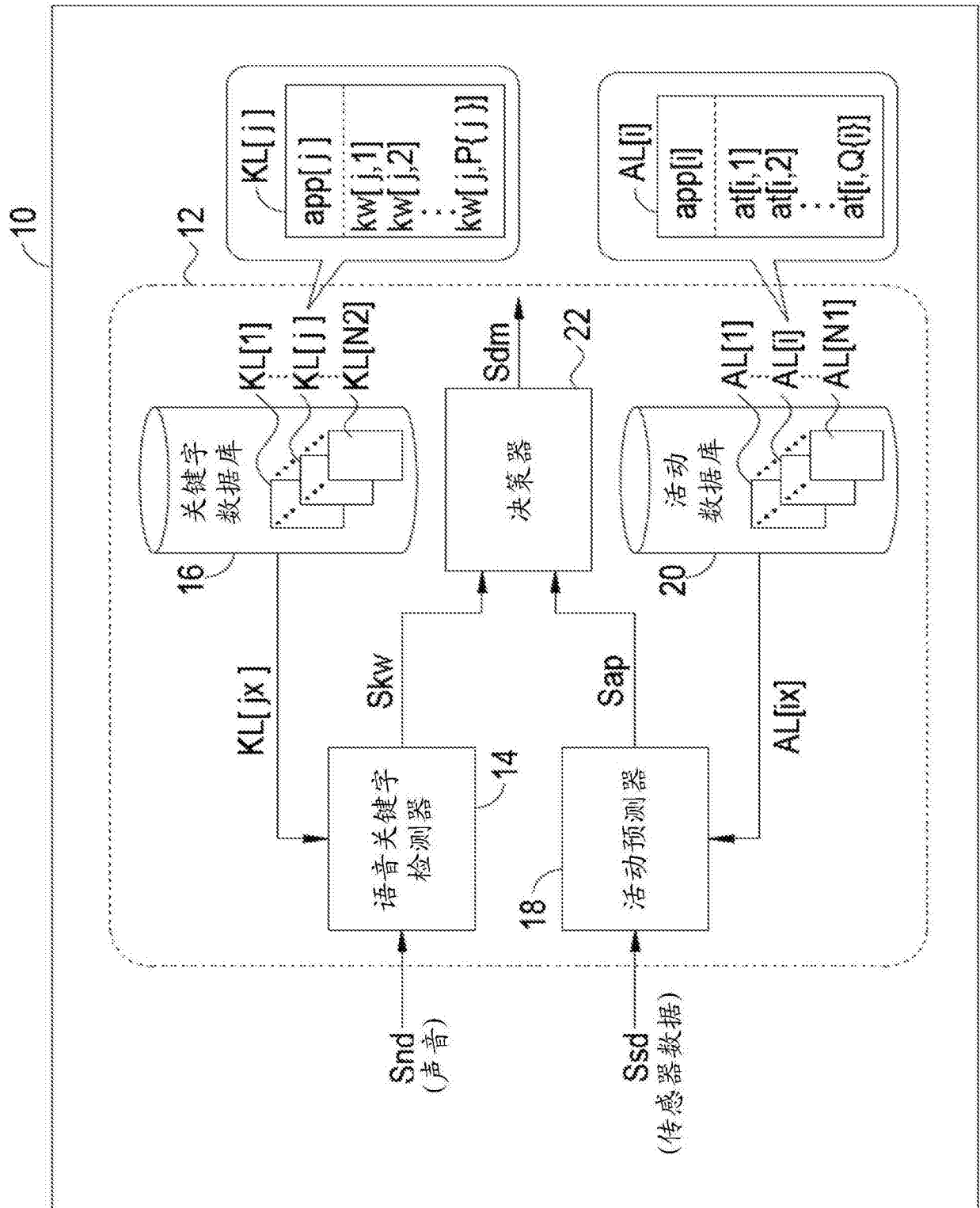


图1

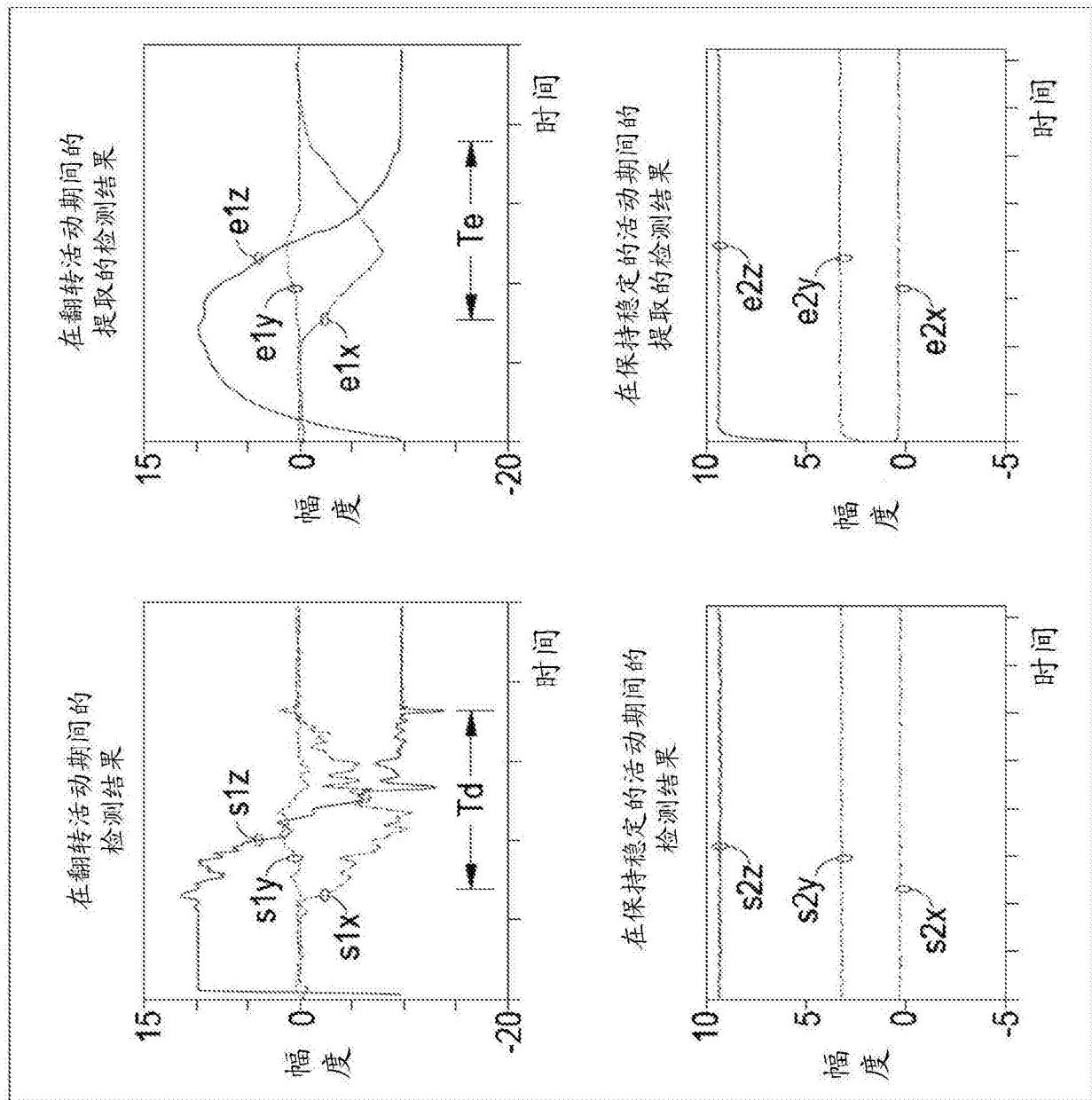


图2

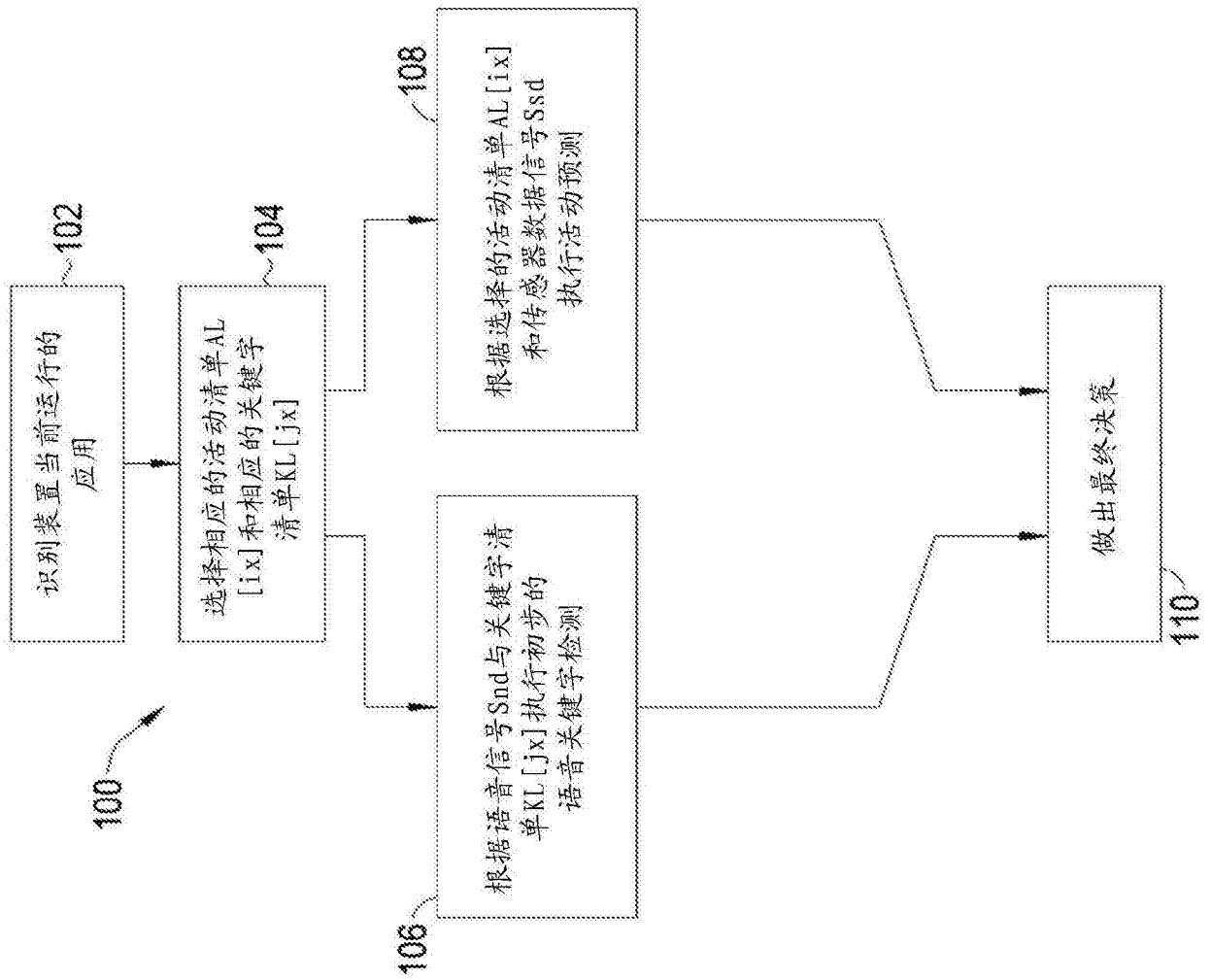


图3

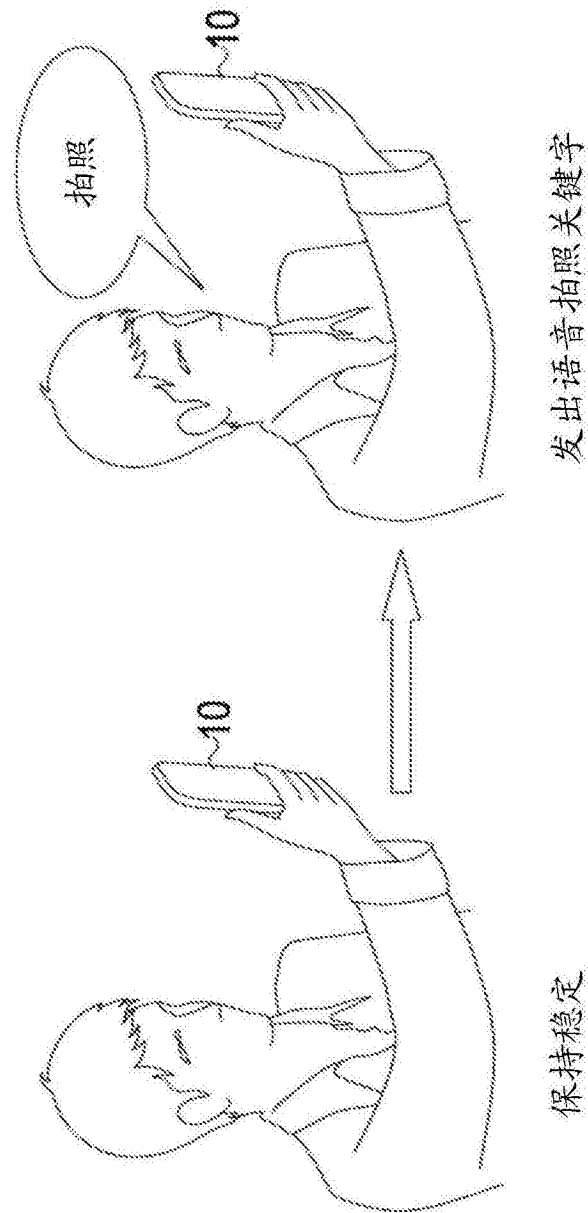


图4

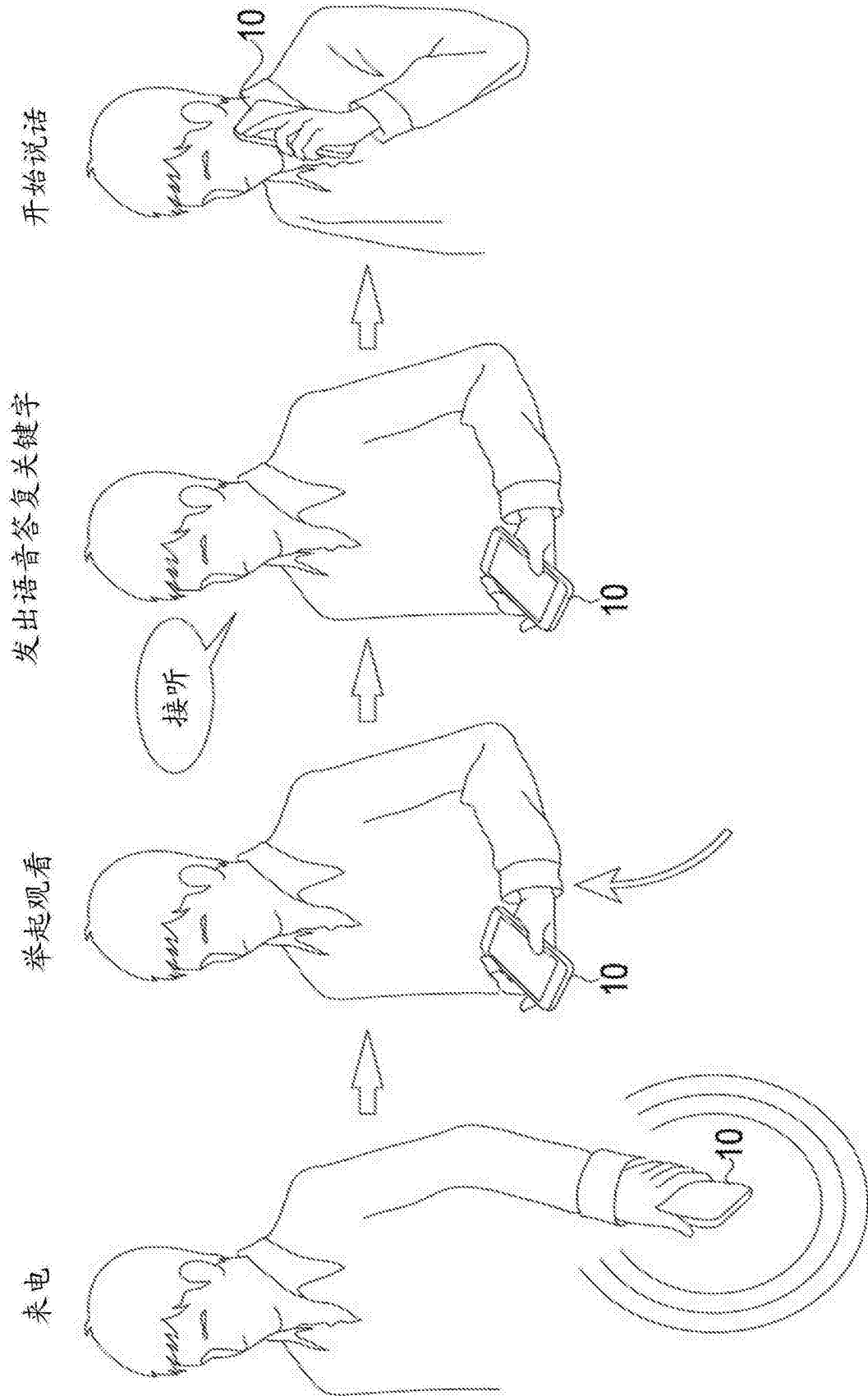


图5

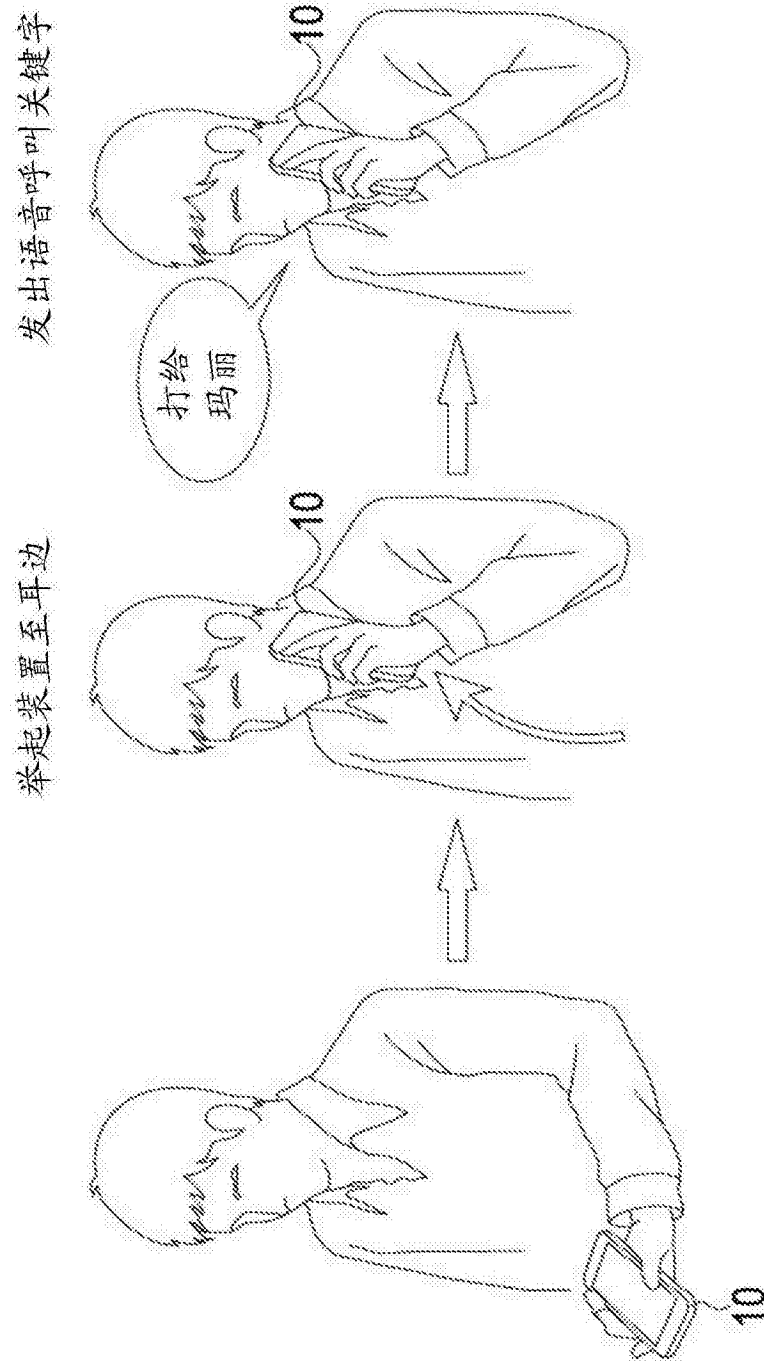


图6

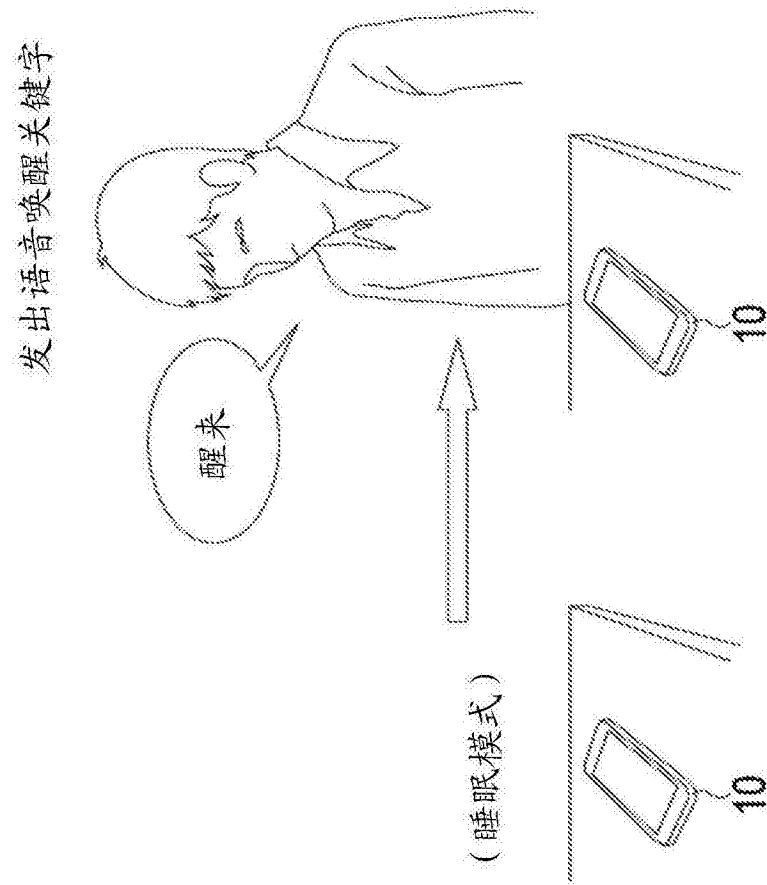


图7

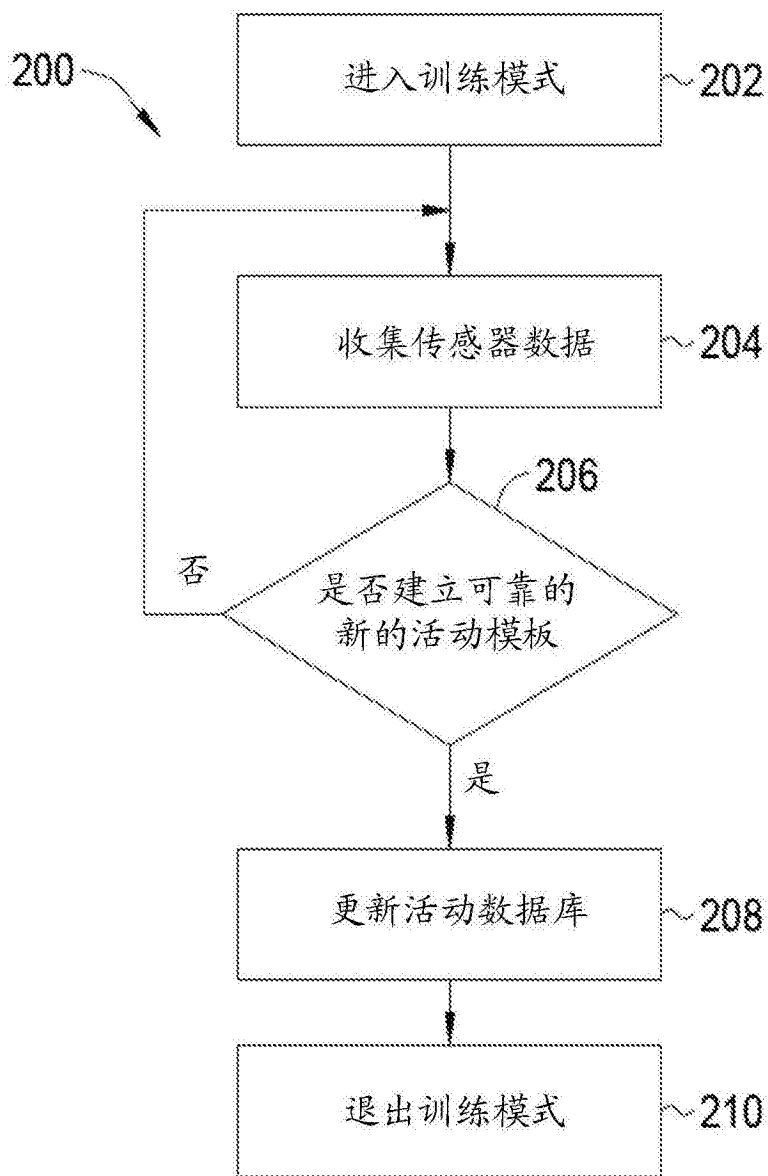


图8

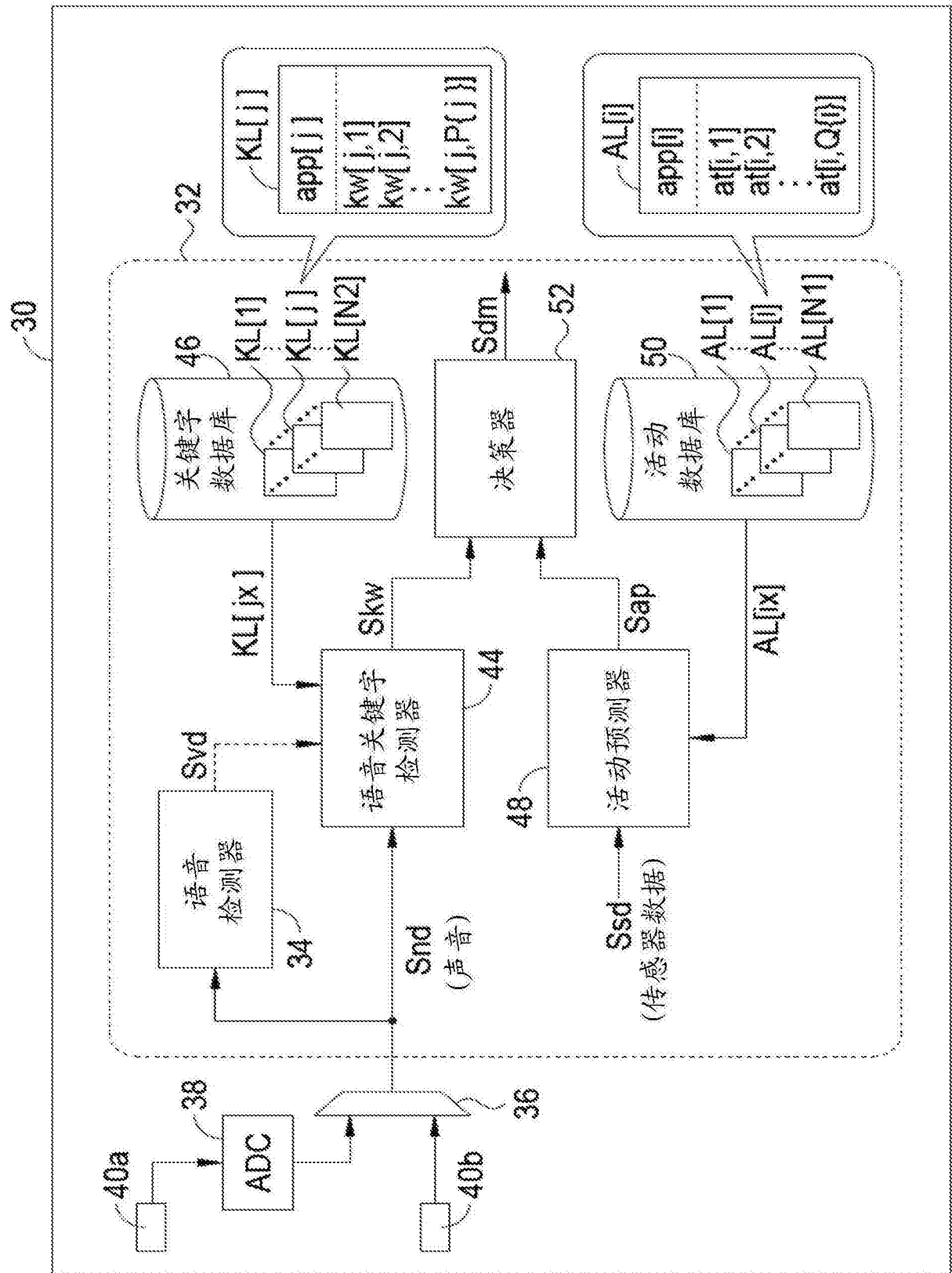


图9

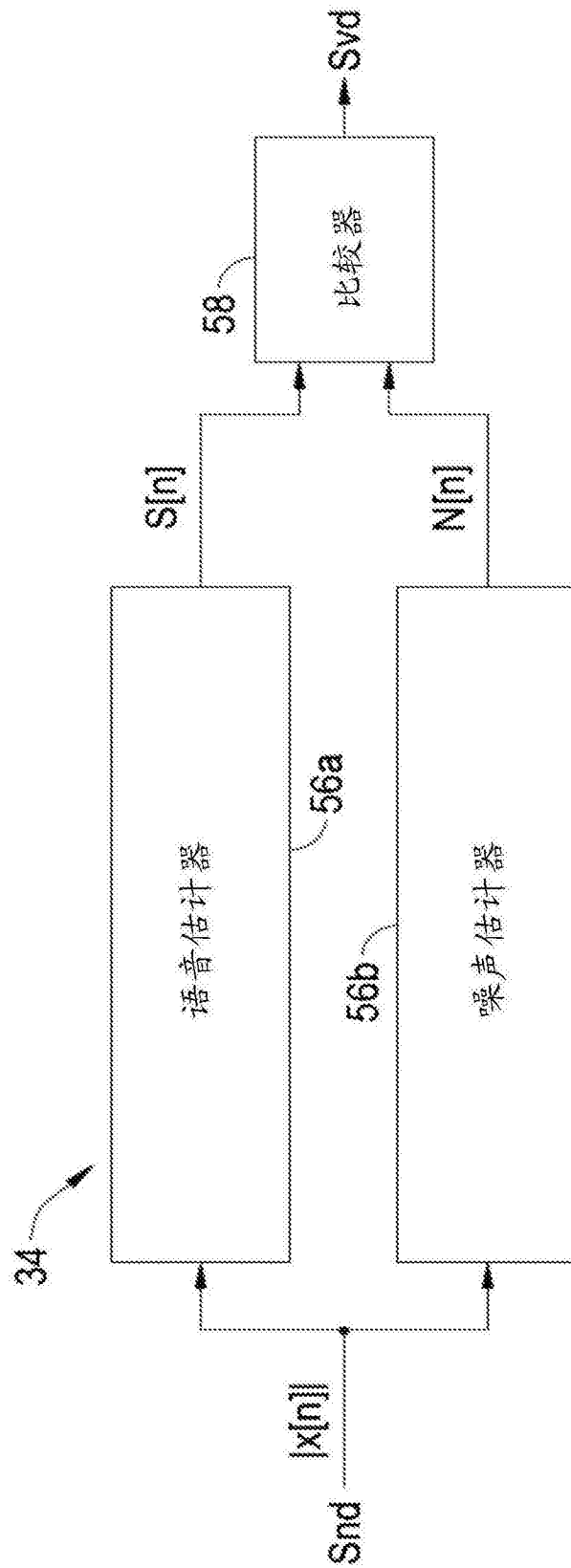


图10

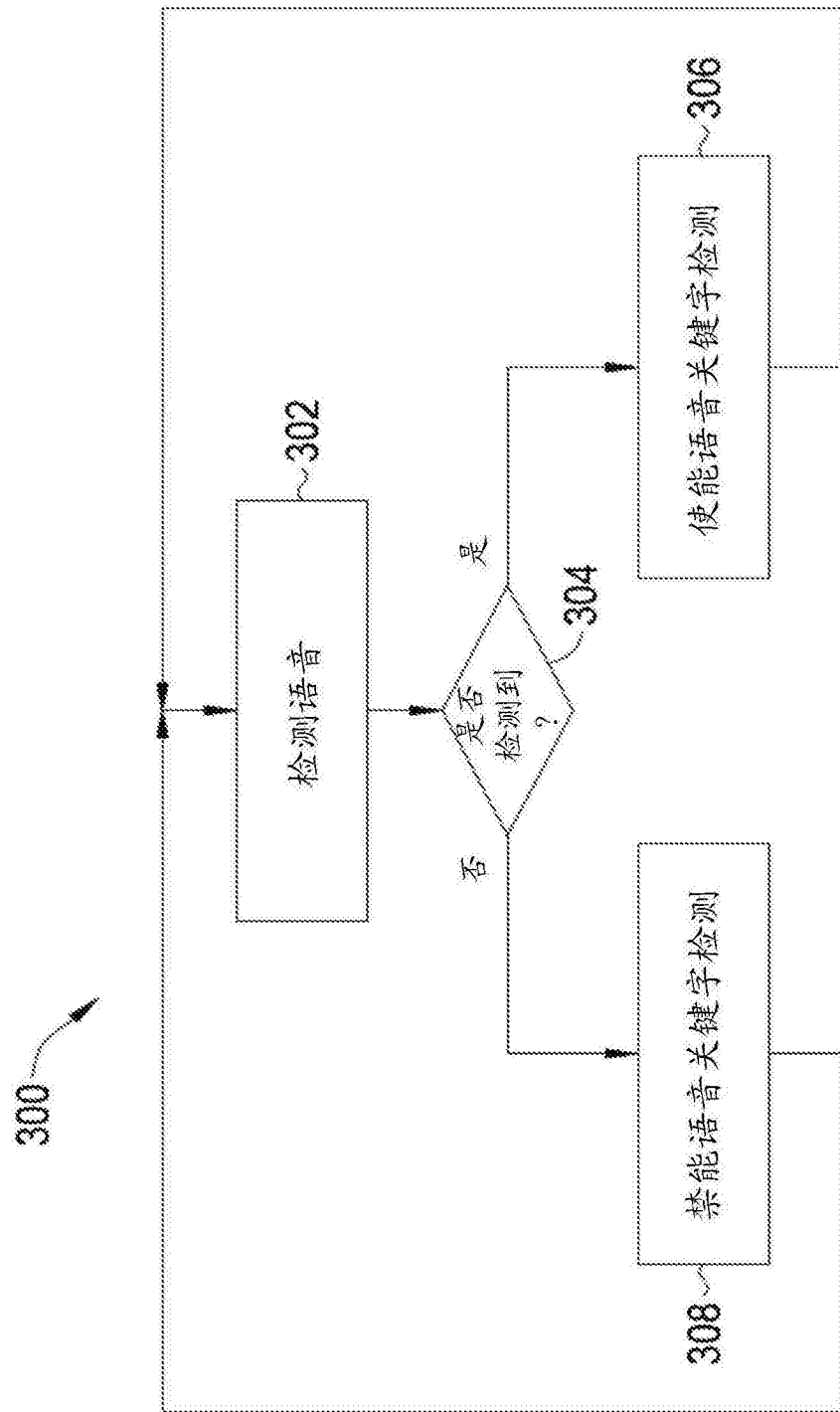


图11