

语音的功能控制方法和装置

技术领域

本申请涉及网络通信技术领域，尤其涉及一种语音的功能控制方法和装置。

5 背景技术

语音助手是运行在终端上的一种软件，能够和用户进行语音交流，以及协助用户实现其指定的各项功能，如进行信息搜索、终端操作等。自从苹果推出其语音助手 Siri 以后，用户对语音助手类的软件的关注程度和使用率有了极大的提高，也给语音助手的发展带来了动力。

10 目前语音助手能够与安装在终端上的应用程序相结合，用户向语音助手发出执行某项任务的指令，由语音助手调用相应的应用程序来完成该项任务，从而极大的丰富了用户能够以语音助手为单一入口来完成的功能。

现有技术中，以 Siri 为例，Siri 可以与六类应用程序（打车、通讯、照片搜索、支付、网络电话、健身）协作，当 Siri 接收到用户的语音输入时，
15 判断用户的意图，并决定是由自己处理还是调用应用程序处理。如果决定调用应用程序，Siri 从自己对用户语音的识别结果中提取相关信息交给应用程序。应用程序按照 Siri 提供的信息，执行该信息指定的功能。

可见，Siri 提供的信息是用户指定的任务是否能被正确完成的基础。当 Siri 识别出的输入语音不够准确时，用户通过语音进行的功能控制就难以达
20 到令人满意的水平。

发明内容

有鉴于此，本申请提供一种语音的功能控制方法，应用在终端的语音助手上，包括：

根据对用户输入语音的识别结果，确定关联应用程序；所述关联应用程

序用来实现用户要使用的功能；

将用户的输入语音传递给所述关联应用程序，供所述关联应用程序对用户的输入语音进行识别，并根据识别结果进行所述功能的实现。

本申请提供的一种语音的功能控制方法，应用在用来实现除语音助手外
5 其他功能的终端应用程序上，包括：

接收来自语音助手的用户的输入语音；

识别用户的输入语音，根据识别结果进行用户要使用功能的实现。

本申请还提供了一种语音的功能控制装置，应用在终端的语音助手上，
包括：

10 关联应用程序单元，用于根据对用户输入语音的识别结果，确定关联应用程序；所述关联应用程序用来实现用户要使用的功能；

输入语音传递单元，用于将用户的输入语音传递给所述关联应用程序，供所述关联应用程序对用户的输入语音进行识别，并根据识别结果进行所述功能的实现。

15 本申请提供的一种语音的功能控制装置，应用在用来实现除语音助手外其他功能的终端应用程序上，包括：

输入语音接收单元，用于接收来自语音助手的用户的输入语音；

功能实现单元，用于识别用户的输入语音，根据识别结果进行用户要使用功能的实现。

20 由以上技术方案可见，本申请的实施例中，在语音助手根据对用户输入语音的识别结果确定关联应用程序后，将用户的输入语音传递给关联应用程序，由关联应用程序自行识别输入语音后执行用户的指令；由于关联应用程序都有其专属的应用领域，比如高德地图是地图及导航领域，因此在其所属的功能类型的语音识别准确度在绝大多数情况下高于通用于所有功能类型的
25 语音助手的准确度，从而能够更为准确和快速的完成用户需要的功能，提高语音功能控制的效率。

附图说明

图 1 是本申请实施例中一种应用在终端的语音助手上、语音的功能控制方法的流程图；

图 2 是本申请实施例中一种应用在终端的应用程序上、语音的功能控制方法的流程图；

图 3 是本申请应用示例的工作原理示意图；

图 4 是终端的一种硬件结构图；

图 5 是本申请实施例中一种应用在终端的语音助手上、语音的功能控制装置的逻辑结构图；

图 6 是本申请实施例中一种应用在终端的应用程序上、语音的功能控制装置的逻辑结构图。

具体实施方式

现有技术中，终端上安装的很多非语音助手类的应用程序本身也具有接收用户语音输入，识别出并执行用户的语音指令，实现用户要使用功能的能力。简便起见，以下将语音助手类的应用程序称为语音助手，将非语音助手类、用来实现除语音助手外的其他功能的应用程序称为应用程序。

语音助手的设计初衷是作为与用户进行语音交互的统一入口，除了识别用户可能涉及所有功能类型的指令，还要陪用户聊天，因此语音助手词库的词汇范围极其广泛，并且其语音识别算法的优化也必须兼顾所有类型词汇的识别率。而应用程序往往专注于几项主要功能的实现，如地图类型应用程序专注于寻址、定位和导航，购物类型的应用程序专注于商品和交易等等，用户在采用这些应用程序时，意图基本确定是要使用应用程序专注的功能。因此应用程序中用于语音识别的词库的词汇范围同样集中在其专注的功能领域，语音识别算法的优化也是如此。

从发出语音指令后到等待终端给出响应的等待时长会极大的影响用户体

验，因而可供进行语音识别的时间非常有限。在有限的时间内，在任何一个特定的功能领域，语音助手对用户意图使用该领域某个功能的输入语音的识别，在准确度上很难赶上属于该功能领域的应用程序进行的识别。

比如在地点名称上，高德地图的词库比 Siri 的词库更为全面准确；同时，
5 由于有长期的关于地点方面的搜索的积累，高德地图的识别算法在识别地名上也比 Siri 更为精准，Siri 在导航这个应用场景中没有更为可靠的地名、位置等数据作为参考，来优化其识别算法。

可见，现有技术中由语音助手将识别结果传递给应用程序的做法，实际上是由不擅长某个工作的一方来完成工作，将不准确的工作结果交给擅长该
10 工作的另一方，另一方不得不以较差的工作结果为基础来执行任务，自然难以达到好的结果。

基于上述思路，本申请的实施例提出一种新的语音功能控制方法，由语音助手通过对用户输入语音的识别，确定实现用户意图的关联应用程序，再将输入语音传递给关联应用程序，由关联应用程序自行识别语音后实现用户
15 要使用的功能，使得应用程序基于原始的输入语音执行用户指令，可以以自行识别的结果而不必以语音助手不理想的识别结果为基础来执行，从而能够更为准确和快速的完成用户需要的功能，以解决现有技术中存在的问题。

本申请的实施例中，语音助手和应用程序都运行在用户的终端上。语音助手可以运行在终端的操作系统层面，也可以是作为操作系统上层的应用来
20 运行，不做限定。用户的终端可以是任何具有语音输入、计算和存储功能的设备，如手机、平板电脑、PC（Personal Computer，个人电脑）、笔记本、服务器等，同样不做限定。

本申请实施例中，语音的功能控制方法应用在语音助手中的流程如图 1 所示，应用在应用程序中的流程如图 2 所示。

25 在语音助手上，步骤 110，根据对用户输入语音的识别结果，确定关联应用程序。关联应用程序用来实现用户要使用的功能。

当语音助手接收到用户的语音输入后，识别用户的语音，如果用户的指

令不涉及对哪个功能的使用、或者用户要使用的功能由语音助手来执行，则语音助手依据识别结果回复用户的输入，或者执行用户的指令。如果语音助手的识别结果是用户要使用某种由应用程序来执行的功能，则语音助手确定实现该功能的关联应用程序。

5 用户可能在输入语音中指定想使用的应用程序。在这种情形下，语音助手可以从对用户输入语音的识别结果中提取出应用程序名称，将该应用程序（即用户在输入语音中指定的应用程序）作为关联应用程序。例如，用户对语音助手说，“用滴滴帮我叫车回家”，语音助手识别出应用程序名称“滴滴”，则将应用程序滴滴作为关联应用程序。

10 如果用户没有在输入语音中指定想使用的应用程序，语音助手可以根据对用户输入语音的识别结果，确定用户要使用的功能，再按照用户要使用的功能在终端上安装的应用程序中确定关联应用程序。语音助手从识别结果中确定用户要使用的功能的方法可以参照各种现有技术实现，例如，可以预置各个功能的若干个关键词，如果对用户语音的识别结果命中了某个功能关键词，则可获知用户想要使用的功能是哪一个。

15 由用户要使用的功能确定关联应用程序的方式可以根据实际应用场景的需要来决定，以下以两个实现方式为例进行说明。

20 在第一个实现方式中，语音助手在识别出用户要使用的功能后，可以将终端上安装的应用程序中，能够实现用户要使用的功能、并且支持语音输入的一个到多个应用程序作为备选应用程序，将备选应用程序的名称显示给用户，供用户选择。在收到用户的选择结果后，语音助手将用户选定的应用程序作为关联应用程序。

25 在第二个实现方式中，可以在终端上维护要使用的功能与应用程序的映射关系。在识别出用户要使用的功能后，语音助手可以将与用户要使用的功能具有映射关系的应用程序作为关联应用程序。在前述以功能关键词来反映用户要使用的功能的例子中，可以在终端上保存一张功能关键词与应用程序的映射关系表，语音助手从对用户输入语音的识别结果中提取出功能关键词

后，即可将与该功能关键词具有映射关系的应用程序作为关联应用程序。

在第二个实现方式中，要使用的功能与应用程序的映射关系可以由用户设置和/或修改，也可以由语音助手或操作系统生成，还可以同时应用这些方式，本申请的实施例不做限定。一个例子中，用户可以在语音助手或操作系统提供的设置项中设置语音输入时，要使用的功能和关联应用程序的映射关系（一个到多个功能对应于一个应用程序）。另一个例子中，可以把要实现某个功能时，用户最频繁采用的应用程序作为与该功能有映射关系的应用程序；具体而言，如果安装在终端上的实现某功能的应用程序只有一个，则将该应用程序作为与该功能有映射关系的应用程序；如果超过一个，则可以按照操作系统统计的用户使用每个实现该功能的应用程序的频率，将频率最高的一个作为与该功能有映射关系的应用程序。第三个例子中，用户还可以在实现某个功能的应用程序中，将该应用程序设置为语音输入时、与该功能有映射关系的应用程序，应用程序在收到用户的设置指令后，向语音助手提交本应用程序与要使用功能的映射关系；如用户在高德地图中设置语音输入时、与功能关键词“导航”具有映射关系，高德地图按照用户的操作将该设置项提交给 Siri，Siri 将这一映射关系保存起来，后续用户语音输入“导航去哪里”时，Siri 将按照映射关系，以高德地图作为关联应用程序。

需要说明的是，上述两个确定关联应用程序的方式也可以结合应用。例如，当语音助手确定用户要使用的功能后，查询保存的功能与应用程序的映射关系，如果能查到与要使用的功能有映射关系的应用程序，则按照映射关系来确定关联应用程序；如果查不到，则将终端上能够实现要使用的功能、并且支持语音输入的应用程序给用户选择，并且按照用户选择确定关联应用程序。在用户选择后可以请用户设置该功能的默认关联应用程序，如果用户进行设置，则保存该功能与用户设置的应用程序的映射关系；如果没有进行设置，语音助手也可以在用户选择一个应用程序实现某个功能的次数和频率都超过设定阈值后，保存该应用程序与该功能的映射关系。例如，Siri 维护的功能与应用程序的映射关系表中不包括“导航”这一功能的映射关系，在 5

- 次用户给出语音指令“导航去哪里”后，Siri 将终端上安装的高德地图、百度地图和搜狗地图的应用程序名称显示给用户，由用户选择要使用哪个应用程序导航；如果用户有 4 次选择了高德地图，则 Siri 将“导航”这一功能与高德地图的映射关系保存至该映射关系表中；之后用户给出导航的语音指令时，
- 5 Siri 将直接以高德地图作为关联应用程序。

在语音助手上，步骤 120，将用户的输入语音传递给关联应用程序，供关联应用程序对用户的输入语音进行识别，并根据识别结果进行该功能的实现。

- 在关联应用程序上，步骤 210，接收来自语音助手的用户的输入语音。
- 10 语音助手在确定执行用户要使用功能的关联应用程序后，打开关联应用程序（包括启动该关联应用程序、将该关联应用程序置于前台运行等），将用户的输入语音传递给关联应用程序。

在关联应用程序上，步骤 220，识别用户的输入语音，根据识别结果进行用户要使用功能的实现。

- 15 关联应用程序自行识别来自语音助手的用户的输入语音，按照识别结果，运行关联应用程序的业务处理逻辑，来实现用户要使用的功能。关联应用程序可以按照现有技术来进行语音识别和功能实现，不再赘述。

- 在一个例子中，语音助手可以将自己对用户输入语音的识别结果和用户的输入语音一并传递给关联应用程序。关联应用程序自行识别用户的输入语音，根据关联应用程序自己的识别结果和来自语音助手的识别结果，来实现
- 20 用户要使用功能的实现。语音助手的识别结果可以作为供关联应用程序在语音识别时的参考，进一步增加识别的准确程度。

- 可见，本申请的实施例，语音助手通过对用户输入语音的识别，确定实现用户要使用功能的关联应用程序，并将输入语音传递给关联应用程序，
- 25 由关联应用程序自行识别输入语音后执行用户的指令，从而能够利用应用程序其在所属功能类型更为准确的语音识别结果，更为准确和快速的完成用户需要的功能，提高语音功能控制的效率。

在本申请的一个应用示例中，用户的苹果终端上安装有语音助手 Siri 和若干个能够完成各种类型功能的应用程序。Siri 中保存有功能关键词和应用程序的映射关系表。一种映射关系表的示例如表 1：

序号	功能关键词	应用程序名称
1	导航	高德地图
2	约车	滴滴出行
3	叫出租	滴滴出行
4	转账	支付宝
⋮	⋮	⋮

表 1

5 本应用示例的工作原理如图 3 所示，在收到用户的输入语音后，Siri 对输入语音进行识别，假设用户要使用的功能需要借助于其他应用程序来实现，Siri 从识别结果中提取描述用户要使用功能的功能关键词，用功能关键词查找映射关系表。如果找到对应于该功能关键词的应用程序，则以该应用程序为关联应用程序。如果没有在表中找到该功能关键词，Siri 将终端上安装的
10 所有能够实现该功能、并且支持语音输入的应用程序名称显示给用户，请用户选择想用哪个应用程序。Siri 将用户选定的应用程序作为关联应用程序。

Siri 将关联应用程序置于前台运行，并把用户的输入语音通过操作系统传输给关联应用程序。关联应用程序识别用户的输入语音，按照自己的识别结果和业务流程来完成用户指令的任务。

15 例如，用户对 Siri 说，“转账 2000 给张三”。Siri 识别出功能关键词“转账”，从表 1 查询到关联应用程序为支付宝。Siri 打开支付宝，将用户的输入语音传递给支付宝。支付宝识别输入语音，启动转账业务流程，向用户显示“收款方：张三”、“转账金额：2000”等内容，在用户输入密码或验证指纹后即可完成转账。

20 与上述流程实现对应，本申请的实施例还提供了一种应用在终端语音助

手上的语音的功能控制装置，和一种应用在终端应用程序上的语音的功能控制装置。这两种装置均可以通过软件实现，也可以通过硬件或者软硬件结合的方式实现。以软件实现为例，作为逻辑意义上的装置，是通过终端的 CPU（Central Process Unit，中央处理器）将对应的计算机程序指令读取到内存中运行形成的。从硬件层面而言，除了图 4 所示的 CPU、内存以及非易失性存储器之外，语音的功能控制装置所在的终端通常还包括用于进行无线信号收发

5 芯片等其他硬件，和/或用于实现网络通信功能的板卡等其他硬件。

图 5 所示为本申请实施例提供的一种语音的功能控制装置，应用在终端的语音助手上，包括关联应用程序单元和输入语音传递单元，其中：关联应用程序单元用于根据对用户输入语音的识别结果，确定关联应用程序；所述

10 关联应用程序用来实现用户要使用的功能；输入语音传递单元用于将用户的输入语音传递给所述关联应用程序，供所述关联应用程序对用户的输入语音进行识别，并根据识别结果进行所述功能的实现。

一个例子中，所述终端上维护有要使用的功能与应用程序的映射关系；

15 所述关联应用程序单元具体用于：根据对用户输入语音的识别结果，确定用户要使用的功能，将与用户要使用的功能具有映射关系的应用程序作为关联应用程序。

上述例子中，所述要使用的功能与应用程序的映射关系，包括：功能关键词与应用程序的映射关系；所述关联应用程序单元具体用于：提取对用户

20 输入语音的识别结果中的功能关键词，将与所述功能关键词具有映射关系的应用程序作为关联应用程序。

上述例子中，所述要使用的功能与应用程序的映射关系，包括：由用户设置的要使用的功能与应用程序的映射关系；和/或，以实现所述要使用的功能时，用户最频繁采用的应用程序作为与所述要使用的功能具有映射关系的

25 应用程序；和/或，由某个应用程序提交的、所述应用程序与要使用功能的映射关系。

可选的，所述关联应用程序单元具体用于：根据对用户输入语音的识别

结果，确定用户要使用的功能，将所述终端上能够实现所述功能、并且支持语音输入的若干个应用程序名称显示给用户供其选择，将用户选定的应用程序作为关联应用程序。

可选的，所述关联应用程序单元具体用于：提取对用户输入语音的识别结果中的应用程序名称，将输入语音中指定的应用程序作为关联应用程序。

可选的，所述输入语音传递单元具体用于：将所述识别结果和用户的输入语音传递给所述关联应用程序，供所述关联应用程序对用户的输入语音进行识别，并根据关联应用程序的识别结果和语音助手的识别结果进行所述功能的实现。

10 图6所示为本申请实施例提供的一种语音的功能控制装置，应用在用来实现除语音助手外其他功能的终端应用程序上，包括输入语音接收单元和功能实现单元，其中：输入语音接收单元用于接收来自语音助手的用户的输入语音；功能实现单元用于识别用户的输入语音，根据识别结果进行用户要使用功能的实现。

15 可选的，所述输入语音接收单元具体用于：接收来自语音助手的用户的输入语音，以及语音助手对所述输入语音的识别结果；所述功能实现单元具体用于：识别用户的输入语音，根据自己的识别结果和来自语音助手的识别结果，进行用户要使用功能的实现。

20 一个例子中，所述装置还包括：映射关系提交单元，用于根据用户的指令，向语音助手提交本应用程序与要使用功能的映射关系。

上述例子中，所述本应用程序与要使用功能的映射关系，包括：本应用程序与功能关键词的映射关系。

25 以上所述仅为本申请的较佳实施例而已，并不用以限制本申请，凡在本申请的精神和原则之内，所做的任何修改、等同替换、改进等，均应包含在本申请保护的范围之内。

在一个典型的配置中，计算设备包括一个或多个处理器（CPU）、输入/输出接口、网络接口和内存。

内存可能包括计算机可读介质中的非永久性存储器,随机存取存储器(RAM)和/或非易失性内存等形式,如只读存储器(ROM)或闪存(flash RAM)。内存是计算机可读介质的示例。

计算机可读介质包括永久性和非永久性、可移动和非可移动媒体可以由任何方法或技术来实现信息存储。信息可以是计算机可读指令、数据结构、程序的模块或其他数据。计算机的存储介质的例子包括,但不限于相变内存(PRAM)、静态随机存取存储器(SRAM)、动态随机存取存储器(DRAM)、其他类型的随机存取存储器(RAM)、只读存储器(ROM)、电可擦除可编程只读存储器(EEPROM)、快闪记忆体或其他内存技术、只读光盘只读存储器(CD-ROM)、数字多功能光盘(DVD)或其他光学存储、磁盒式磁带,磁带磁磁盘存储或其他磁性存储设备或任何其他非传输介质,可用于存储可以被计算设备访问的信息。按照本文中的界定,计算机可读介质不包括暂存电脑可读媒体(transitory media),如调制的数据信号和载波。

还需要说明的是,术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含,从而使得包括一系列要素的过程、方法、商品或者设备不仅包括那些要素,而且还包括没有明确列出的其他要素,或者是还包括为这种过程、方法、商品或者设备所固有的要素。在没有更多限制的情况下,由语句“包括一个.....”限定的要素,并不排除在包括所述要素的过程、方法、商品或者设备中还存在另外的相同要素。

本领域技术人员应明白,本申请的实施例可提供为方法、系统或计算机程序产品。因此,本申请可采用完全硬件实施例、完全软件实施例或结合软件和硬件方面的实施例的形式。而且,本申请可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质(包括但不限于磁盘存储器、CD-ROM、光学存储器等)上实施的计算机程序产品的形式。