



(12) 发明专利

(10) 授权公告号 CN 109215640 B

(45) 授权公告日 2021.06.01

(21) 申请号 201710519007.8

(22) 申请日 2017.06.30

(65) 同一申请的已公布的文献号
申请公布号 CN 109215640 A

(43) 申请公布日 2019.01.15

(73) 专利权人 深圳大森智能科技有限公司
地址 518000 广东省深圳市罗湖区桂园街
道书城路1号都市名园裙楼5楼C区003
室

(72) 发明人 邴会

(74) 专利代理机构 北京华夏泰和知识产权代理
有限公司 11662
代理人 姚金金

(51) Int. Cl.
G10L 15/22 (2006.01)

(56) 对比文件

CN 106161155 A, 2016.11.23

CN 106409317 A, 2017.02.15

CN 106504744 A, 2017.03.15

CN 103902728 A, 2014.07.02

US 9536517 B2, 2017.01.03

CN 102237087 A, 2011.11.09

CN 106384591 A, 2017.02.08

CN 106356053 A, 2017.01.25

审查员 宋玉香

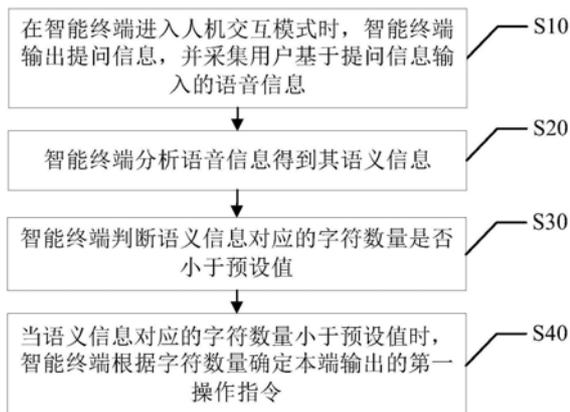
权利要求书1页 说明书6页 附图2页

(54) 发明名称

语音识别方法、智能终端及计算机可读存储
介质

(57) 摘要

本发明公开了一种语音识别方法,包括:在智能终端进入人机交互模式时,智能终端输出提问信息,并采集用户基于提问信息输入的语音信息;智能终端分析语音信息得到其语义信息;智能终端判断语义信息对应的字符数量是否小于预设值;当语义信息对应的字符数量小于预设值时,智能终端根据字符数量确定本端输出的第一操作指令。本发明还公开了一种智能终端、计算机可读存储介质。本发明在语音识别的误差率的同时,提高了语音识别效率。



1. 一种语音识别方法,其特征在于,所述语音识别方法包括以下步骤:

在智能终端进入人机交互模式时,所述智能终端输出提问信息,并采集用户基于所述提问信息输入的语音信息;

所述智能终端分析所述语音信息得到其语义信息;

所述智能终端判断所述语义信息对应的字符数量是否小于预设值;

当所述语义信息对应的字符数量小于预设值时,所述智能终端根据所述字符数量确定本端输出的第一操作指令;

当所述语义信息对应的字符数量大于预设值时,所述智能终端提取所述语义信息中的关键词,并根据所述关键词确定本端输出的第二操作指令。

2. 如权利要求1所述的语音识别方法,其特征在于,所述智能终端判断所述语义信息对应的字符数量是否小于预设值的步骤之前还包括:

所述智能终端判断所述语义信息与所述提问信息对应的预设回答信息是否相匹配;

当所述语义信息与所述预设回答信息不匹配时,则执行所述智能终端判断所述语义信息对应的字符数量是否小于预设值的步骤;

当所述语义信息与所述预设回答信息相匹配时,所述智能终端根据所述语义信息输出第三操作指令。

3. 如权利要求1-2任一项所述的语音识别方法,其特征在于,所述预设值为N,N为正整数。

4. 一种智能终端,其特征在于,所述智能终端包括:存储器、处理器及存储在所述存储器上并可在所述处理器上运行的语音识别程序,所述语音识别程序被所述处理器执行时实现如权利要求1至3中任一项所述的语音识别方法的步骤。

5. 一种计算机可读存储介质,其特征在于,所述计算机可读存储介质上存储有语音识别程序,所述语音识别程序被处理器执行时实现如权利要求1至3中任一项所述的语音识别方法的步骤。

语音识别方法、智能终端及计算机可读存储介质

技术领域

[0001] 本发明涉及语音识别技术领域,尤其涉及一种语音识别方法及装置、计算机可读存储介质。

背景技术

[0002] 在现有技术中,语音识别技术给人机交互带来方便的同时,存在以下不足:由于语言中存在大量同音字,使得在智能设备在语音识别过程中,对于长度较短的语音信息的识别误差率较高,而语音识别错误会导致智能设备的语意理解错误,从而进一步导致智能设备的做出错误判断,执行了错误的动作。

发明内容

[0003] 本发明的主要目的在于提供一种语音识别方法及装置、计算机可读存储介质,旨在降低语音识别的误差率的同时,提高语音识别效率。

[0004] 为实现上述目的,本发明提供一种语音识别方法,所述语音识别方法包括以下步骤:

[0005] 在智能终端进入人机交互模式时,所述智能终端输出提问信息,并采集用户基于所述提问信息输入的语音信息;

[0006] 所述智能终端分析所述语音信息得到其语义信息;

[0007] 所述智能终端判断所述语义信息对应的字符数量是否小于预设值;

[0008] 当所述语义信息对应的字符数量小于预设值时,所述智能终端根据所述字符数量确定本端输出的第一操作指令。

[0009] 优选地,所述智能终端判断所述语义信息对应的字符数量是否小于预设值的步骤之后还包括:

[0010] 当所述语义信息对应的字符数量大于预设值时,所述智能终端提取所述语义信息中的关键词,并根据所述关键词确定本端输出的第二操作指令。

[0011] 优选地,所述智能终端判断所述语义信息对应的字符数量是否小于预设值的步骤之前还包括:

[0012] 所述智能终端判断所述语义信息与所述提问信息对应的预设回答信息是否相匹配;

[0013] 当所述语义信息与所述预设回答信息不相匹配时,则执行所述智能终端判断所述语义信息对应的字符数量是否小于预设值的步骤;

[0014] 当所述语义信息与所述预设回答信息相匹配时,所述智能终端根据所述语义信息输出第三操作指令。

[0015] 优选地,所述预设值为N,N为正整数。

[0016] 为实现上述目的,本发明还提供一种智能终端,所述智能终端包括:存储器、处理器及存储在所述存储器上并可在所述处理器上运行的语音识别程序,所述语音识别程序被

所述处理器执行时实现如上述任一项所述的语音识别方法的步骤。

[0017] 此外,为实现上述目的,本发明还提供一种计算机可读存储介质,所述计算机可读存储介质上存储有语音识别程序,所述语音识别程序被处理器执行时实现如上述任一项所述的语音识别方法的步骤。

[0018] 本发明通过智能终端分析语音信息得到其语义信息,并判断该语义信息对应的字符数量是否小于预设值,当语义信息对应的字符数量小于预设值时,该智能终端根据所述字符数量确定本端输出的第一操作指令。通过上述方案,使得在确定语音信息对应的语义信息的字符数量小于预设值时,进一步根据字符数量来确定该智能终端需输出的第一操作指令,从而使得智能终端即使在无法识别出语音信息所对应的正确语义信息的情况下,通过语义信息的字符数量来理解该语音信息所对应的正确语义信息,并基于该字符数量来确定该智能终端需要输出的操作指令,从而实现了在降低语音识别的误差率的同时,提高语音识别效率。

附图说明

[0019] 图1是本发明实施例方案涉及的硬件运行环境的终端\装置结构示意图;

[0020] 图2为本发明语音识别方法第一实施例的流程示意图;

[0021] 图3为本发明语音识别方法第二实施例的流程示意图;

[0022] 图4为本发明语音识别方法第三实施例的流程示意图。

[0023] 本发明目的的实现、功能特点及优点将结合实施例,参照附图做进一步说明。

具体实施方式

[0024] 应当理解,此处所描述的具体实施例仅仅用以解释本发明,并不用于限定本发明。

[0025] 如图1所示,图1是本发明实施例方案涉及的硬件运行环境的终端结构示意图。

[0026] 本发明实施例终端可以是PC,也可以是智能手机、平板电脑、电子书阅读器、MP3 (Moving Picture Experts Group Audio Layer III,动态影像专家压缩标准音频层面3)播放器、MP4 (MovingPictureExperts Group Audio Layer IV,动态影像专家压缩标准音频层面3)播放器、便携计算机等具有显示功能的可移动式终端设备。

[0027] 如图1所示,该终端可以包括:处理器1001,例如CPU,网络接口1004,用户接口1003,存储器1005,通信总线1002。其中,通信总线1002用于实现这些组件之间的连接通信。用户接口1003可以包括显示屏(Display)、输入单元比如键盘(Keyboard),可选用户接口1003还可以包括标准的有线接口、无线接口。网络接口1004可选的可以包括标准的有线接口、无线接口(如WI-FI接口)。存储器1005可以是高速RAM存储器,也可以是稳定的存储器(non-volatile memory),例如磁盘存储器。存储器1005可选的还可以是独立于前述处理器1001的存储装置。

[0028] 可选地,终端还可以包括摄像头、RF (Radio Frequency,射频)电路,传感器、音频电路、WiFi模块等等。其中,传感器比如光传感器、运动传感器以及其他传感器。具体地,光传感器可包括环境光传感器及接近传感器,其中,环境光传感器可根据环境光线的明暗来调节显示屏的亮度,接近传感器可在移动终端移动到耳边时,关闭显示屏和/或背光。作为运动传感器的一种,重力加速度传感器可检测各个方向上(一般为三轴)加速度的大小,静

止时可检测出重力的大小及方向,可用于识别移动终端姿态的应用(比如横竖屏切换、相关游戏、磁力计姿态校准)、振动识别相关功能(比如计步器、敲击)等;当然,移动终端还可配置陀螺仪、气压计、湿度计、温度计、红外线传感器等其他传感器,在此不再赘述。

[0029] 本领域技术人员可以理解,图1中示出的终端结构并不构成对终端的限定,可以包括比图示更多或更少的部件,或者组合某些部件,或者不同的部件布置。

[0030] 如图1所示,作为一种计算机存储介质的存储器1005中可以包括操作系统、网络通信模块、用户接口模块以及语音识别程序。

[0031] 在图1所示的终端中,网络接口1004主要用于连接后台服务器,与后台服务器进行数据通信;用户接口1003主要用于连接客户端(用户端),与客户端进行数据通信;而处理器1001可以用于调用存储器1005中存储的语音识别程序,并执行以下操作:

[0032] 【在智能终端进入人机交互模式时,所述智能终端输出提问信息,并采集用户基于所述提问信息输入的语音信息;

[0033] 所述智能终端分析所述语音信息得到其语义信息;

[0034] 所述智能终端判断所述语义信息对应的字符数量是否小于预设值;

[0035] 当所述语义信息对应的字符数量小于预设值时,所述智能终端根据所述字符数量确定本端输出的第一操作指令。】

[0036] 进一步地,处理器1001可以调用存储器1005中存储的语音识别程序,还执行以下操作:

[0037] 【当所述语义信息对应的字符数量大于预设值时,所述智能终端提取所述语义信息中的关键词,并根据所述关键词确定本端输出的第二操作指令。】

[0038] 进一步地,处理器1001可以调用存储器1005中存储的语音识别程序,还执行以下操作:

[0039] 【所述智能终端判断所述语义信息与所述提问信息对应的预设回答信息是否相匹配;

[0040] 当所述语义信息与所述预设回答信息不匹配时,则执行所述智能终端判断所述语义信息对应的字符数量是否小于预设值的步骤;

[0041] 当所述语义信息与所述预设回答信息相匹配时,所述智能终端根据所述语义信息输出第三操作指令。】

[0042] 进一步地,处理器1001可以调用存储器1005中存储的语音识别程序,还执行以下操作:

[0043] 【所述预设值为N,N为正整数。】

[0044] 参照图2,本发明第一实施例提供一种语音识别方法,所述语音识别方法包括:

[0045] 步骤S10,在智能终端进入人机交互模式时,智能终端输出提问信息,并采集用户基于提问信息输入的语音信息;

[0046] 步骤S20,智能终端分析语音信息得到其语义信息;

[0047] 步骤S30,智能终端判断语义信息对应的字符数量是否小于预设值;

[0048] 步骤S40,当语义信息对应的字符数量小于预设值时,智能终端根据字符数量确定本端输出的第一操作指令。

[0049] 在本实施例中,面对一个二选一的问题时,用户回答问题的自然习惯是如果答复

是肯定的,那么可能回答得很简单就是“对”,“是”,“嗯”这样,如果问题的回答是否定的,那么往往会稍作解释。因此,通过用识别出语义信息的字符数量来模糊判断,以使智能终端输出第一操作指令。其中,该预设值为N,N为正整数。

[0050] 需要说明的是,该智能终端对该提问信息可以进行语音播报,也可以通过屏幕输出。

[0051] 下面通过一个具体的实施例来详细说明上述方法是如何实现的。

[0052] 在智能终端进入人机交互模式时,接收到用户输入的语音信息“帮我联系张丽”,智能终端基于该“帮我联系张丽”的语音信息,输出提问信息“您是希望帮您拨通张丽的手机吗”,此时若采集到用户基于该提问信息输入的语音信息,且分析得到该语音信息的语义信息为“队”(或者“兑”)时,判断该语义信息的字符数量是否小于3,在确定该语义信息的字符数量小于3时,该智能终端将为“队”的该语义信息确定为“对”,并执行相应的第一操作指令,如输出语音信息“好的主人,即将帮您拨打电话”。

[0053] 本实施例通过智能终端分析语音信息得到其语义信息,并判断该语义信息对应的字符数量是否小于预设值,当语义信息对应的字符数量小于预设值时,该智能终端根据所述字符数量确定本端输出的第一操作指令。通过上述方案,使得在确定语音信息对应的语义信息的字符数量小于预设值时,进一步根据字符数量来确定该智能终端需输出的第一操作指令,从而使得智能终端即使在无法识别出语音信息所对应的正确语义信息的情况下,通过语义信息的字符数量来理解该语音信息所对应的正确语义信息,并基于该字符数量来确定该智能终端需要输出的操作指令,从而实现了在降低语音识别的误差率的同时,提高语音识别效率。

[0054] 进一步的,参照图3,本发明第二实施例提供一种语音识别方法,基于上述图1所示的实施例,所述步骤S40之后还包括:

[0055] 步骤S50,当语义信息对应的字符数量大于预设值时,智能终端提取语义信息中的关键词,并根据关键词确定本端输出的第二操作指令。

[0056] 下面通过一个具体的实施例来详细说明上述方法是如何实现的。

[0057] 在智能终端进入人机交互模式时,接收到用户输入的语音信息“帮我联系张丽”,智能终端基于该“帮我联系张丽”的语音信息,输出提问信息“您是希望帮您拨通张丽的手机吗”,此时若采集到用户基于该提问信息输入的语音信息,且分析得到该语音信息的语义信息为“不是,我是想跟张丽视频”,判断该语义信息的字符数量是否小于3,在确定该语义信息的字符数量大于或者等于3时,该智能终端提取该语义信息中的关键词(如“不是”、“张丽”、“视频”),并根据该关键词确定本端输出的第二操作指令,如输出语音信息“好的主人,即将帮您打开视频聊天”。

[0058] 本实施例当语义信息对应的字符数量大于预设值时,智能终端通过提取语义信息中的关键词,并根据关键词确定本端输出的第二操作指令。由于在语义信息对应的字符数量大于预设值的情况下,语义信息的识别正确率较高,因此通过提取语义信息中的关键词可降低语音识别的误差率。

[0059] 进一步的,参照图4,本发明第三实施例提供一种语音识别方法,基于上述图1所示的实施例,所述步骤S30之前还包括:

[0060] 步骤S60,智能终端判断语义信息与提问信息对应的预设回答信息是否相匹配;

[0061] 当语义信息与预设回答信息不相匹配时,则执行所述步骤S30,即智能终端判断语义信息对应的字符数量是否小于预设值;

[0062] 步骤S70,当语义信息与预设回答信息相匹配时,智能终端根据语义信息输出第三操作指令。

[0063] 在本实施例中,为了提高智能终端的语音控制效率,当语义信息与预设回答信息相匹配时,表明该智能终端对该语音信息的识别是正确的,此时该智能终端只需输出预存的与该语义信息关联的第三操作指令即可。

[0064] 本发明实施例还提出一种智能终端,智能终端包括:存储器、处理器及存储在所述存储器上并可在所述处理器上运行的语音识别程序,所述语音识别程序被所述处理器执行时实现如下操作:

[0065] 【在智能终端进入人机交互模式时,所述智能终端输出提问信息,并采集用户基于所述提问信息输入的语音信息;

[0066] 所述智能终端分析所述语音信息得到其语义信息;

[0067] 所述智能终端判断所述语义信息对应的字符数量是否小于预设值;

[0068] 当所述语义信息对应的字符数量小于预设值时,所述智能终端根据所述字符数量确定本端输出的第一操作指令。】

[0069] 进一步地,所述语音识别程序被处理器执行时还实现如下操作:

[0070] 【当所述语义信息对应的字符数量大于预设值时,所述智能终端提取所述语义信息中的关键词,并根据所述关键词确定本端输出的第二操作指令。】

[0071] 进一步地,所述语音识别程序被处理器执行时还实现如下操作:

[0072] 【所述智能终端判断所述语义信息与所述提问信息对应的预设回答信息是否相匹配;

[0073] 当所述语义信息与所述预设回答信息不相匹配时,则执行所述智能终端判断所述语义信息对应的字符数量是否小于预设值的步骤;

[0074] 当所述语义信息与所述预设回答信息相匹配时,所述智能终端根据所述语义信息输出第三操作指令。】

[0075] 进一步地,所述语音识别程序被处理器执行时还实现如下操作:

[0076] 【所述预设值为N,N为正整数。】

[0077] 此外,本发明实施例还提出一种计算机可读存储介质,所述计算机可读存储介质上存储有语音识别程序,所述语音识别程序被处理器执行时实现如下操作:

[0078] 【在智能终端进入人机交互模式时,所述智能终端输出提问信息,并采集用户基于所述提问信息输入的语音信息;

[0079] 所述智能终端分析所述语音信息得到其语义信息;

[0080] 所述智能终端判断所述语义信息对应的字符数量是否小于预设值;

[0081] 当所述语义信息对应的字符数量小于预设值时,所述智能终端根据所述字符数量确定本端输出的第一操作指令。】

[0082] 进一步地,所述语音识别程序被处理器执行时还实现如下操作:

[0083] 【当所述语义信息对应的字符数量大于预设值时,所述智能终端提取所述语义信息中的关键词,并根据所述关键词确定本端输出的第二操作指令。】

[0084] 进一步地,所述语音识别程序被处理器执行时还实现如下操作:

[0085] 【所述智能终端判断所述语义信息与所述提问信息对应的预设回答信息是否相匹配;

[0086] 当所述语义信息与所述预设回答信息不匹配时,则执行所述智能终端判断所述语义信息对应的字符数量是否小于预设值的步骤;

[0087] 当所述语义信息与所述预设回答信息相匹配时,所述智能终端根据所述语义信息输出第三操作指令。】

[0088] 进一步地,所述语音识别程序被处理器执行时还实现如下操作:

[0089] 【所述预设值为N,N为正整数。】

[0090] 需要说明的是,在本文中,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、物品或者系统不仅包括那些要素,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、物品或者系统所固有的要素。在没有更多限制的情况下,由语句“包括一个……”限定的要素,并不排除在包括该要素的过程、方法、物品或者系统中还存在另外的相同要素。

[0091] 上述本发明实施例序号仅仅为了描述,不代表实施例的优劣。

[0092] 通过以上的实施方式的描述,本领域的技术人员可以清楚地了解到上述实施例方法可借助软件加必需的通用硬件平台的方式来实现,当然也可以通过硬件,但很多情况下前者是更佳的实施方式。基于这样的理解,本发明的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来,该计算机软件产品存储在如上所述的一个存储介质(如ROM/RAM、磁碟、光盘)中,包括若干指令用以使得一台终端设备(可以是手机,计算机,服务器,空调器,或者网络设备等)执行本发明各个实施例所述的方法。

[0093] 以上仅为本发明的优选实施例,并非因此限制本发明的专利范围,凡是利用本发明说明书及附图内容所作的等效结构或等效流程变换,或直接或间接运用在其他相关的技术领域,均同理包括在本发明的专利保护范围内。

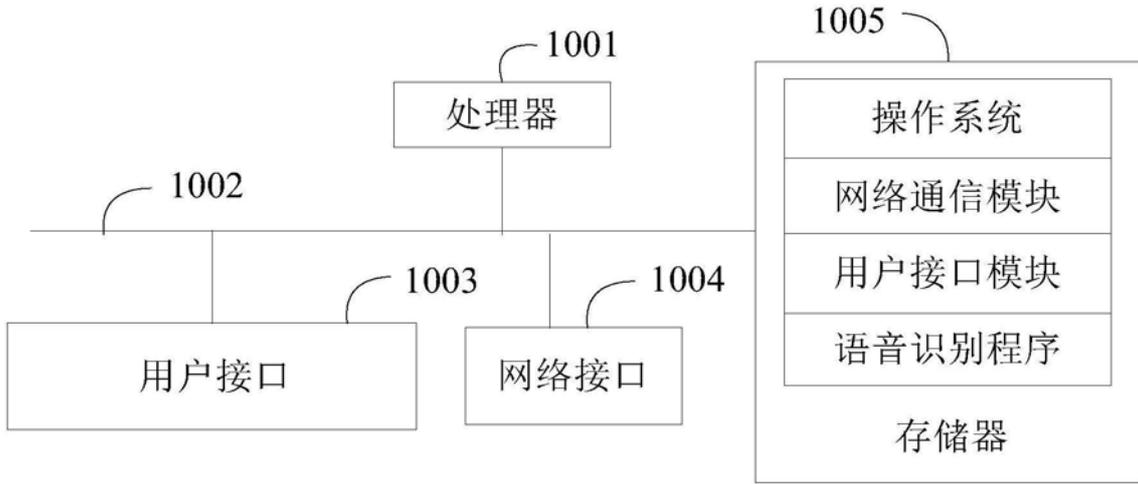


图1

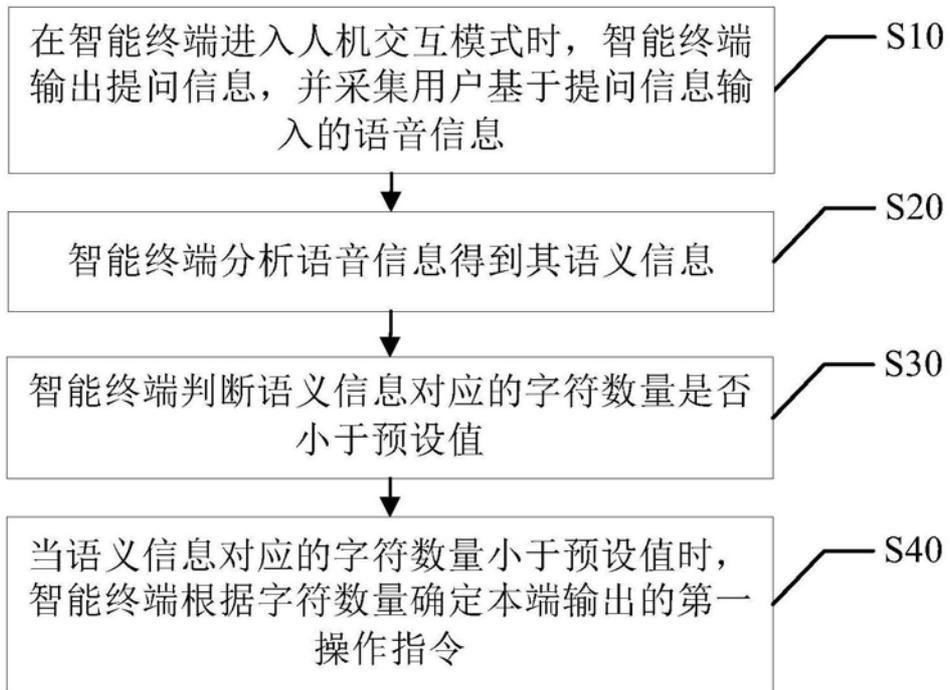


图2

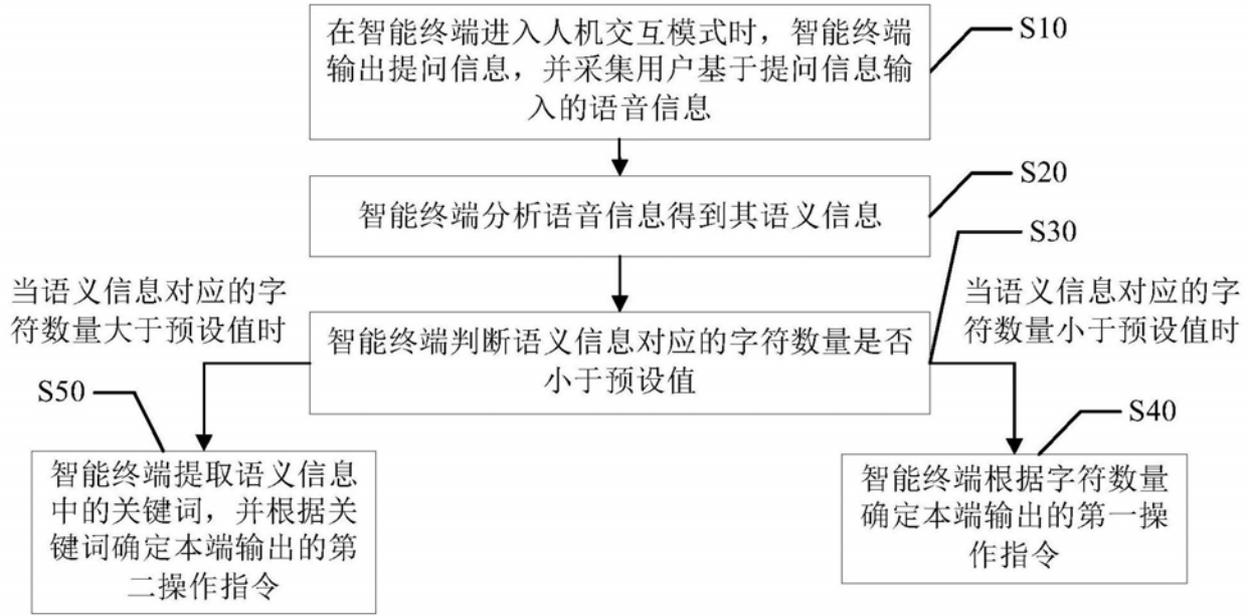


图3

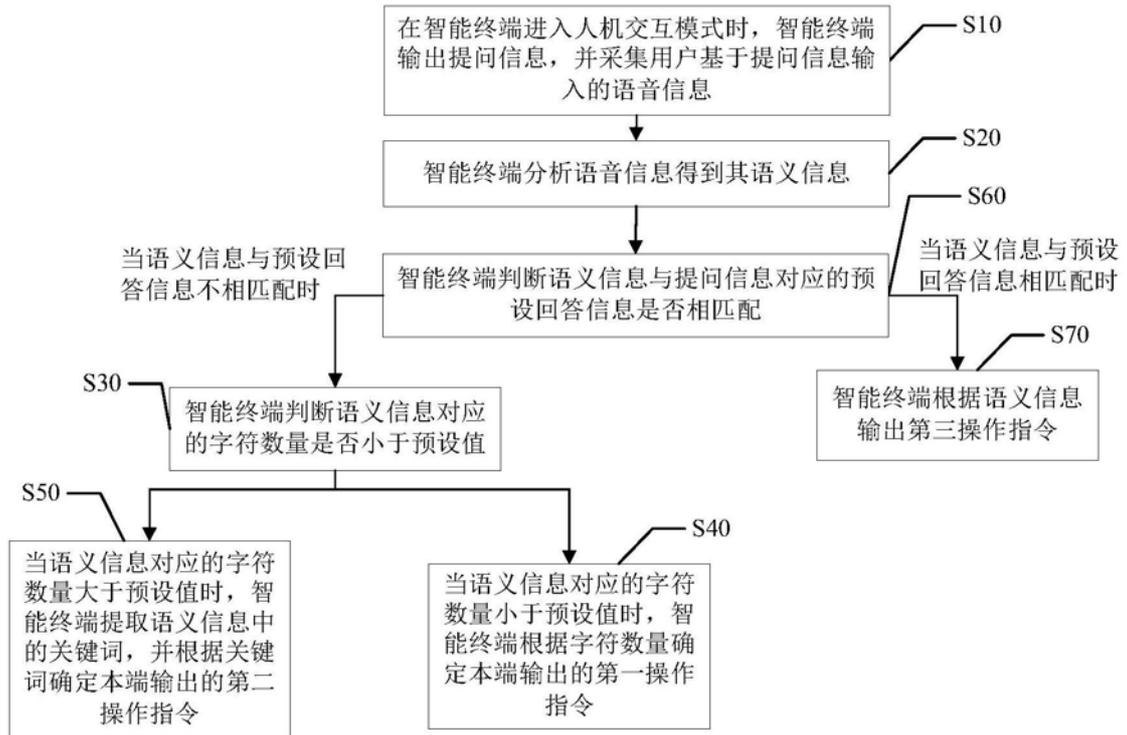


图4