



(12)发明专利申请

(10)申请公布号 CN 106504743 A

(43)申请公布日 2017. 03. 15

(21)申请号 201611025399.4

(22)申请日 2016.11.14

(71)申请人 北京光年无限科技有限公司

地址 100000 北京市石景山区石景山路3号
玉泉大厦四层常青藤青年创业工作室
193号

(72)发明人 石琰

(74)专利代理机构 北京聿华联合知识产权代理
有限公司 11611

代理人 朱绘 张文娟

(51)Int.Cl.

G10L 13/02(2013.01)

G10L 13/08(2013.01)

G10L 15/02(2006.01)

G10L 17/22(2013.01)

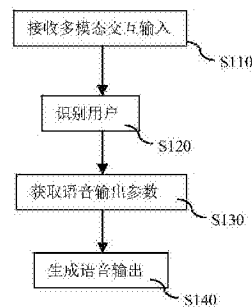
权利要求书2页 说明书7页 附图4页

(54)发明名称

一种用于智能机器人的语音交互输出方法及机器人

(57)摘要

本发明公开了一种用于智能机器人的语音交互输出方法以及一种智能机器人。本发明的方法包括:接收来自用户的多模态交互输入信息,识别当前的用户;获取所述用户对应的语音输出参数,所述语音输出参数为根据所述用户的说话习惯生成;结合所述语音输出参数进行交互数据处理以生成语音输出。根据本发明的方法,可以使得机器人的语音输出符合用户语音交谈习惯的,不仅便于用户理解,而且大大增强了用户亲切感。根据本发明的方法,机器人的拟人度被大大提高,机器人的用户体验得到有效提升。



1. 一种用于智能机器人的语音交互输出方法,其特征在于,所述方法包括:
接收来自用户的多模态交互输入信息,识别当前的用户;
获取所述用户对应的语音输出参数,所述语音输出参数为根据所述用户的说话习惯生成;
结合所述语音输出参数进行交互数据处理以生成语音输出。
2. 根据权利要求1所述的方法,其特征在于,获取当前的所述用户对应的语音输出参数,其中,当不存在与所述用户对应的语音输出参数时:
基于默认设置的语音输出参数或上一次对话被采用的语音输出参数开始与所述用户的语音交互;
在所述语音交互的过程中根据所述用户的语音输入调节当前的语音输出参数并基于调节后的语音输出参数进行下一步的语音交互。
3. 根据权利要求2所述的方法,其特征在于,根据所述用户的语音输入调节当前的语音输出参数,其中:
解析所述用户的语音输入以获取所述用户的语速和/或对话等待时间;
调节当前的语音输出参数使得语音输出的语速和/或对话等待时间与所述用户一致。
4. 根据权利要求3所述的方法,其特征在于,调节当前的语音输出参数使得语音输出的语速和/或对话等待时间与所述用户一致,其中,多次逐步调节语音输出参数使得语音输出的语速和/或对话等待时间与所述用户一致。
5. 根据权利要求2所述的方法,其特征在于,所述方法还包括:
将最终调节完毕的语音输出参数保存为所述用户对应的语音输出参数。
6. 根据权利要求1-5中任一项所述的方法,其特征在于,所述语音输出参数包括语速参量和/或对话等待时间参量。
7. 一种智能机器人,其特征在于,所述机器人包括:
用户识别模块,其配置为接收来自用户的多模态交互输入信息,识别当前的用户;
语音输出参数获取模块,其配置为获取并输出所述用户对应的语音输出参数,所述语音输出参数为根据所述用户的说话习惯生成;
语音输出模块,其配置为结合所述语音输出参数进行交互数据处理以生成语音输出。
8. 根据权利要求7所述的机器人,其特征在于,所述语音输出参数获取模块包含:
语音输出参数确认模块,其配置为确认当前是否存在与所述用户对应的语音输出参数;
语音输出参数调用模块,其配置为当不存在与所述用户对应的语音输出参数时调用默认设置的语音输出参数或上一次对话被采用的语音输出参数;
语音输出参数调节模块,其配置为根据所述用户的语音输入调节所述语音输出参数并输出调节后的所述语音输出参数。
9. 根据权利要求8所述的机器人,其特征在于,所述语音输出参数调节模块配置为:
解析所述用户的语音输入以获取所述用户的语速和/或对话等待时间;
调节当前的语音输出参数使得语音输出的语速和/或对话等待时间与所述用户一致。
10. 根据权利要求8所述的机器人,其特征在于,所述语音输出参数获取模块还包含:
语音输出参数保存模块,其配置为将最终调节完毕的语音输出参数保存为所述用户对

应的语音输出参数。

一种用于智能机器人的语音交互输出方法及机器人

技术领域

[0001] 本发明涉及机器人领域,具体说涉及一种用于智能机器人的语音交互输出方法及机器人。

背景技术

[0002] 随着机器人技术的不断发展,智能机器人越来越多的被应用与人类的家庭生活中。随着智能机器人在人类的家庭生活中的应用不断深入,对智能机器人的拟人化水平的要求也不断提高。

[0003] 当前,很多智能机器人具备了语音功能,可以与用户进行语音交互。但是,在现有技术中,智能机器人的语音输出通常是具体语义的直接体现,其并不能很好的适应于人类的语音交互习惯。这就导致虽然机器人通过语音对用户做出了正确的交互回应,但是用户体验仍然不好。

[0004] 因此,为了进一步提高智能机器人的拟人化水平,提升用户体验,需要一种用于智能机器人的语音交互输出方法。

发明内容

[0005] 本发明提供了一种用于智能机器人的语音交互输出方法,所述方法包括:

[0006] 接收来自用户的多模态交互输入信息,识别当前的用户;

[0007] 获取所述用户对应的语音输出参数,所述语音输出参数为根据所述用户的说话习惯生成;

[0008] 结合所述语音输出参数进行交互数据处理以生成语音输出。

[0009] 在一实施例中,获取当前的所述用户对应的语音输出参数,其中,当不存在与所述用户对应的语音输出参数时:

[0010] 基于默认设置的语音输出参数或上一次对话被采用的语音输出参数开始与所述用户的语音交互;

[0011] 在所述语音交互的过程中根据所述用户的语音输入调节当前的语音输出参数并基于调节后的语音输出参数进行下一步的语音交互。

[0012] 在一实施例中,根据所述用户的语音输入调节当前的语音输出参数,其中:

[0013] 解析所述用户的语音输入以获取所述用户的语速和/或对话等待时间;

[0014] 调节当前的语音输出参数使得语音输出的语速和/或对话等待时间与所述用户一致。

[0015] 在一实施例中,调节当前的语音输出参数使得语音输出的语速和/或对话等待时间与所述用户一致,其中,多次逐步调节语音输出参数使得语音输出的语速和/或对话等待时间与所述用户一致。

[0016] 在一实施例中,所述方法还包括:

[0017] 将最终调节完毕的语音输出参数保存为所述用户对应的语音输出参数。

- [0018] 在一实施例中,所述语音输出参数包括语速参量和/或对话等待时间参量。
- [0019] 本发明还提出了一种智能机器人,所述机器人包括:
- [0020] 用户识别模块,其配置为接收来自用户的多模态交互输入信息,识别当前的用户;
- [0021] 语音输出参数获取模块,其配置为获取并输出所述用户对应的语音输出参数,所述语音输出参数为根据所述用户的说话习惯生成;
- [0022] 语音输出模块,其配置为结合所述语音输出参数进行交互数据处理以生成语音输出。
- [0023] 在一实施例中,所述语音输出参数获取模块包含:
- [0024] 语音输出参数确认模块,其配置为确认当前是否存在与所述用户对应的语音输出参数;
- [0025] 语音输出参数调用模块,其配置为当不存在与所述用户对应的语音输出参数时调用默认设置的语音输出参数或上一次对话被采用的语音输出参数;
- [0026] 语音输出参数调节模块,其配置为根据所述用户的语音输入调节所述语音输出参数并输出调节后的所述语音输出参数。
- [0027] 在一实施例中,所述语音输出参数调节模块配置为:
- [0028] 解析所述用户的语音输入以获取所述用户的语速和/或对话等待时间;
- [0029] 调节当前的语音输出参数使得语音输出的语速和/或对话等待时间与所述用户一致。
- [0030] 在一实施例中,所述语音输出参数获取模块还包含:
- [0031] 语音输出参数保存模块,其配置为将最终调节完毕的语音输出参数保存为所述用户对应的语音输出参数。
- [0032] 根据本发明的方法,可以使得机器人的语音输出符合用户语音交谈习惯的,不仅便于用户理解,而且大大增强了用户亲切感。根据本发明的方法,机器人的拟人度被大大提高,机器人的用户体验得到有效提升。
- [0033] 本发明的其它特征或优点将在随后的说明书中阐述。并且,本发明的部分特征或优点将通过说明书而变得显而易见,或者通过实施本发明而被了解。本发明的目的和部分优点可通过在说明书、权利要求书以及附图中所特别指出的步骤来实现或获得。

附图说明

- [0034] 附图用来提供对本发明的进一步理解,并且构成说明书的一部分,与本发明的实施例共同用于解释本发明,并不构成对本发明的限制。在附图中:
- [0035] 图1~图3是根据本发明实施例的方法流程图;
- [0036] 图4是根据本发明一实施例的方法部分流程图
- [0037] 图5和图6是根据本发明实施例的机器人系统结构框图。

具体实施方式

- [0038] 以下将结合附图及实施例来详细说明本发明的实施方式,借此本发明的实施人员可以充分理解本发明如何应用技术手段来解决技术问题,并达成技术效果的实现过程并依据上述实现过程具体实施本发明。需要说明的是,只要不构成冲突,本发明中的各个实施例

以及各实施例中的各个特征可以相互结合,所形成的技术方案均在本发明的保护范围之内。

[0039] 随着机器人技术的不断发展,智能机器人越来越多的被应用与人类的家庭生活中。随着智能机器人在人类的家庭生活中的应用不断深入,对智能机器人的拟人化水平的要求也不断提高。

[0040] 当前,很多智能机器人具备了语音功能,可以与用户进行语音交互。但是,在现有技术中,智能机器人的语音输出通常是具体语义的直接体现,其并不能很好的适应于人类的语音交互习惯。这就导致虽然机器人通过语音对用户做出了正确的交互回应,但是用户体验仍然不好。

[0041] 因此,为了进一步提高智能机器人的拟人化水平,提升用户体验,本发明提出了一种用于智能机器人的语音交互输出方法。在本发明的方法中,通过分析人类的语音交互习惯归纳人类语音输出的特征,并尽量使得机器人的语音输出具备人类语音输出的特征,从而提高机器人的拟人化水平,提升用户体验。进一步的,由于不同人具备不同的语音输出特征,而人类在理解语音输入时比较容易理解和自己语音输出特征类似的语音输出,因此在本发明的方法中,尽量使得机器人的语音输出特征与当前交互对象的语音输出特征匹配,从而进一步提高用户体验。

[0042] 接下来基于附图详细描述本发明实施例的具体执行过程。附图的流程图中示出的步骤可以在包含诸如一组计算机可执行指令的计算机系统中执行。虽然在流程图中示出了各步骤的逻辑顺序,但是在某些情况下,可以以不同于此处的顺序执行所示出或描述的步骤。

[0043] 如图1所示,在一实施例中,机器人首先接收来自用户的多模态交互输入信息(步骤S110);根据接收到的多模态交互输入信息识别当前的用户(步骤S120);然后根据用户身份获取对应当前用户的语音输出参数(语音输出特征)(步骤S130);然后结合获取到的语音输出参数进行交互数据处理生成并输出包含上述语音输出特征的语音输出(步骤S140)。

[0044] 在步骤S130中,获取的语音输出参数与用户的语音输出特征相匹配。进一步的,该语音输出参数是根据用户的说话习惯生成的。具体的,在一实施例中,根据用户的说话习惯预先生成对应的语音输出参数(针对不同的用户生成多个不同的语音输出参数),然后在交互过程中根据用户身份调用对应的语音输出参数。

[0045] 进一步的,在人类的日常语音交流中,语音输出的特征包含语种(其包含语系归属、方言的划分)、口音、语速以及对话等待时间等等。按照程序执行的习惯,从逻辑上可以将人类语音输出过程简单归纳为:

[0046] 确定自身想要表达的意思;

[0047] 结合自身的语种(语系、方言)设定将想要表达的意思转化为具体的语言;

[0048] 以自身习惯的语速输出(说出)具体的语言并在输出(说出)的同时对输出(说出)的语言进行口音修正;

[0049] 在与交互对象的语音交互过程中,在两组语言输出的间歇使用自身习惯的对话等待时间(对方说话后等待对话等待时间后作出回应,或,自己说话时如果需要连续说出多组内容,在两组内容间等待对话等待时间)。

[0050] 基于上述流程的分析,在一实施例中,将机器人对话输出的过程划分为:

[0051] 确定语音内容；

[0052] 基于语音输出参数对语音内容进行交互数据处理生成语音输出。

[0053] 其中,确定想要表达的意思并结合语种设定将想要表达的意思转化为具体的语言属于确定语音内容的过程。因此在一实施例中,语音输出参数包括口音、语速参量和/或对话等待时间参量。进一步的,由于口音包含的语音细节太多且在实际环境中口音的种类复杂多变,其很难具体分析并量化确定。因此,在一实施例中,语音输出参数包括语速参量和/或对话等待时间参量。

[0054] 如图2所示,在一实施例中,机器人首先接收来自用户的多模态交互输入信息(步骤S210);根据接收到的多模态交互输入信息识别当前的用户(步骤S220);然后根据用户身份获取对应当前用户的语速参数以及对话等待时间(步骤S230);确定语音输出的具体内容(步骤S240);按照对应当前用户的语速参数以及对话等待时间确定步骤S240确定的语音输出内容的输出语速以及输出时机从而完成语音输出(步骤S250)。

[0055] 这样,根据图2所示的实施例,机器人“说话”的语速以及对话等待时间就可以符合当前用的说话习惯,不仅便于用户理解对话内容,而且增强了用户的亲切感,从而大大提高了机器人的用户体验。

[0056] 在图1以及图2所示的实施例中,关键步骤之一在于机器人必须调用对应当前用户的语音输出参数(语速参量和/或对话等待时间参量)。在一实施例中,机器人的内部存储系统或者机器人的云端服务器中预先存储有多个不同用户的语音输出参数。当机器人识别用户身份后,基于用户身份从已保存的语音输出参数调出对应的语音输出参数。

[0057] 但是,在实际应用环境中,不可能预知机器人将要面对的所有用户,也就不能在存储系统/云端中预存所有机器人将要面对的用户语音输出参数。进一步的,在某些情况下,机器人刚投入使用时其内部存储为初始状态,并未预存任何用户数据。或者,在某些应用环境中,机器人无法联网获取云端数据。

[0058] 在交互过程中,上述种种情况最终会导致机器人无法调出与当前用户对应的语音输出参数。针对这种情况,在本发明一实施例中,机器人在无法调用与当前用户对应的语音输出参数(不存在与当前用户对应的语音输出参数或调用过程出现错误无法调用)时,采用现有的(可以调用)语音输出参数进行语音输出,并在在交互过程中根据用户的语音输入调整当前采用的语音输出参数,最后根据调整后的语音输出参数进行下一步的语音交互,从而使得机器人最终的语音输出的语音输出参数与用户的语音对话习惯匹配。

[0059] 具体的,如图3所示,在一实施例中,机器人首先接收来自用户的多模态交互输入信息(步骤S310);根据接收到的多模态交互输入信息识别当前的用户(步骤S320);然后确认是否存在对应当前用户的语音输出参数(步骤S330);当存在对应当前用户的语音输出参数时根据用户身份获取对应当前用户的语音输出参数(步骤S351);然后结合获取到的语音输出参数进行交互数据处理生成并输出包含上述语音输出特征的语音输出(步骤S352)。

[0060] 当不存在对应当前用户的语音输出参数(或者无法调用对应当前用户的语音输出参数)时,机器人调用可用的语音输出参数(步骤S340),根据可用的语音输出参数进行语音输出(步骤S341),从而开始与用户的语音交互。在与用户的语音交互过程中,机器人根据用户的语音输入调整机器人当前采用的语音输出参数(步骤S342),并根据调整后的语音输出参数进行语音输出(步骤S343)从而使得机器人的语音输出与用户的说话习惯匹配。

[0061] 进一步的,在一实施例中,在步骤S340中,机器人调用可用的语音输出参数,具体的,在一实施例中,机器人调用默认的(预设的)语音输出参数。或者,在另一实施例中,机器人根据交互记录,调用上一次语音对话被采用的语音输出参数。

[0062] 进一步的,在一实施例中,机器人在根据用户的语音输入调节当前的语音输出参数的过程中:

[0063] 解析用户的语音输入以获取用户的语速和/或对话等待时间;

[0064] 调节当前的语音输出参数使得语音输出的语速和/或对话等待时间与用户一致。

[0065] 具体的,如图4所示,当不存在对应当前用户的语音输出参数(或者无法调用对应当前用户的语音输出参数)时,机器人调用可用的语音输出参数(步骤S410),根据可用的语音输出参数进行语音输出(步骤S411),从而开始与用户的语音交互。在与用户的语音交互过程中,机器人接收用户的语音输入(步骤S420),获取用户语音输入的语速以及对话等待时间(步骤S421)。

[0066] 进一步的,考虑到有时并不能仅仅通过一轮语音对话就能完全获取用户语音输入的语速以及对话等待时间,因此在本实施例中执行步骤S422,判断是否获取到完整的用户语音输入的语速以及对话等待时间,如果没有则继续基于当前的语音输出参数进行下一轮的语音输出(返回步骤S411)。

[0067] 如果已经获取到完整的用户语音输入的语速以及对话等待时间,则根据用户语音输入的语速以及对话等待时间调节当前的语音输出参数(步骤S430)并根据调整后的语音输出参数进行新一轮的语音输出(步骤S440),使得语音输出的语速和对话等待时间与用户一致,从而使得机器人的语音输出与用户的说话习惯匹配。

[0068] 进一步的,在通常的语音交谈中,如果一方突然改变语音输出特点(口音、语速或对话等待时间)或造成另一方的不适(对方刚刚适应你的语音输出特点,突然转变会导致对方不习惯)。因此,在本发明一实施例中,在机器人根据用户语音输入调整当前的语音输出参数时,再用了逐步调整的策略。即,将对语音输出参数的调整分为多步,每步只对语音输出参数的值做小范围调整,根据小范围调整后的语音输出参数输出下一轮的语音输出,不断在上一轮语音输出的语音输出参数的基础上做小范围调整直到语音输出参数与用户的语音输入匹配。

[0069] 这样,新输出的语音输出与上一轮的语音输出在语音输出参数上的差别维持在一个较小的范围内。这样就能消除由于语音输出参数的突然转变而造成的用户的不适感。具体的,在一实施例中,多次逐步调节语音输出参数使得语速和/或对话等待时间与用户一致。

[0070] 进一步的,为了使得下一次和相同用户进行语音交互时不需要再次进行语音输出参数的获取,在一实施例中,在步骤S440之后,将最终调节完毕的语音输出参数保存为用户对应的语音输出参数(步骤S450)。

[0071] 综上,根据本发明的方法,可以使得机器人的语音输出符合用户语音交谈习惯的,不仅便于用户理解,而且大大增强了用户亲切感。根据本发明的方法,机器人的拟人度被大大提高,机器人的用户体验得到有效提升。

[0072] 基于本发明的方法,本发明还提出了一种符合用户语音交谈习惯的机器人。如图5所示,在一实施例中,机器人包含:

[0073] 用户识别模块510,其配置为接收来自用户的多模态交互输入信息,识别当前的用户;

[0074] 语音输出参数获取模块520,其配置为获取并输出当前用户对应的语音输出参数,具体的,语音输出参数为根据用户的说话习惯生成;

[0075] 语音输出模块530,其配置为结合语音输出参数获取模块520输出的语音输出参数进行交互数据处理以生成语音输出。

[0076] 这样,语音输出模块530最终输出的语音输出就是与用户的说话习惯相匹配的,不仅便于用户理解,而且可以有效提高用户的亲切感。相较于现有技术,本发明的机器人的拟人化水平得到有效提高,机器人的用户体验被大大提升。

[0077] 进一步的,在一实施例中,如图6所示,用户识别模块610配置为接收来自用户的多模态交互输入信息,识别当前的用户。

[0078] 语音输出参数获取模块620包含语音输出参数确认模块621、语音输出参数调用模块622以及语音输出参数调节模块623。其中:

[0079] 语音输出参数确认模块621,其配置为确认当前是否存在与当前用户对应的语音输出参数;

[0080] 语音输出参数调用模块622,其配置为当不存在与用户对应的语音输出参数时调用默认设置的语音输出参数或上一次对话被采用的语音输出参数(当存在与用户对应的语音输出参数时调用与用户对应的语音输出参数);

[0081] 语音输出参数调节模块623,其配置为根据用户的语音输入调节语音输出参数调用模块622调出的语音输出参数并输出调节后的语音输出参数(进一步的,当存在与用户对应的语音输出参数时、语音交互刚刚开始不存在用户语音输入或者用户的语音输入数据量不足时语音输出参数调节模块623不对调节语音输出参数调用模块622调出的语音输出参数进行调整而是直接输出)。

[0082] 语音输出模块630配置为结合语音输出参数调节模块623输出的语音输出参数进行交互数据处理以生成语音输出。

[0083] 这样,在机器人与没有对应的语音输出参数的用户进行语音交互时也能够输出与用户谈话习惯匹配的语音输出,从而保证了机器人的用户体验。

[0084] 具体的,在一实施例中,语音输出参数调节模块623配置为:

[0085] 解析用户的语音输入以获取所述用户的语速和/或对话等待时间;

[0086] 调节当前的语音输出参数使得语音输出模块630最终的语音输出的语速和/或对话等待时间与当前用户一致。

[0087] 进一步的,在一实施例中,语音输出参数获取模块620还包含:

[0088] 语音输出参数保存模块624,其配置为将语音输出参数调节模块623最终调节完毕的语音输出参数保存为当前用户对应的语音输出参数。

[0089] 综上,本发明的机器人可以输出与用户的说话习惯相匹配的语音输出,不仅便于用户理解,而且可以有效提高用户的亲切感。相较于现有技术,本发明的机器人的拟人化水平得到有效提高,机器人的用户体验被大大提升。

[0090] 虽然本发明所公开的实施方式如上,但所述的内容只是为了便于理解本发明而采用的实施方式,并非用以限定本发明。本发明所述的方法还可有其他多种实施例。在不背离

本发明实质的情况下,熟悉本领域的技术人员当可根据本发明做出各种相应的改变或变形,但这些相应的改变或变形都应属于本发明的权利要求的保护范围。

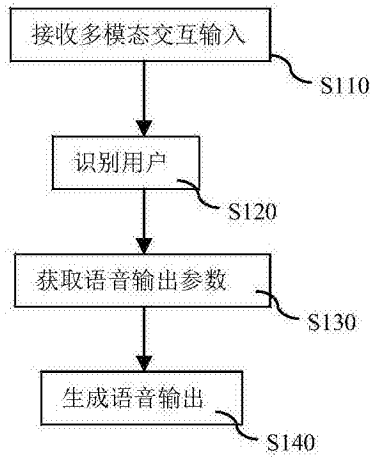


图1

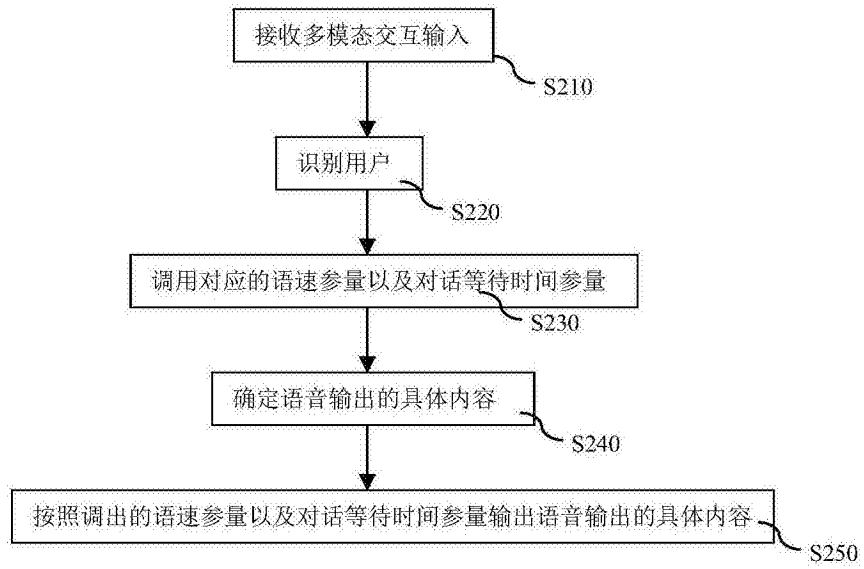


图2

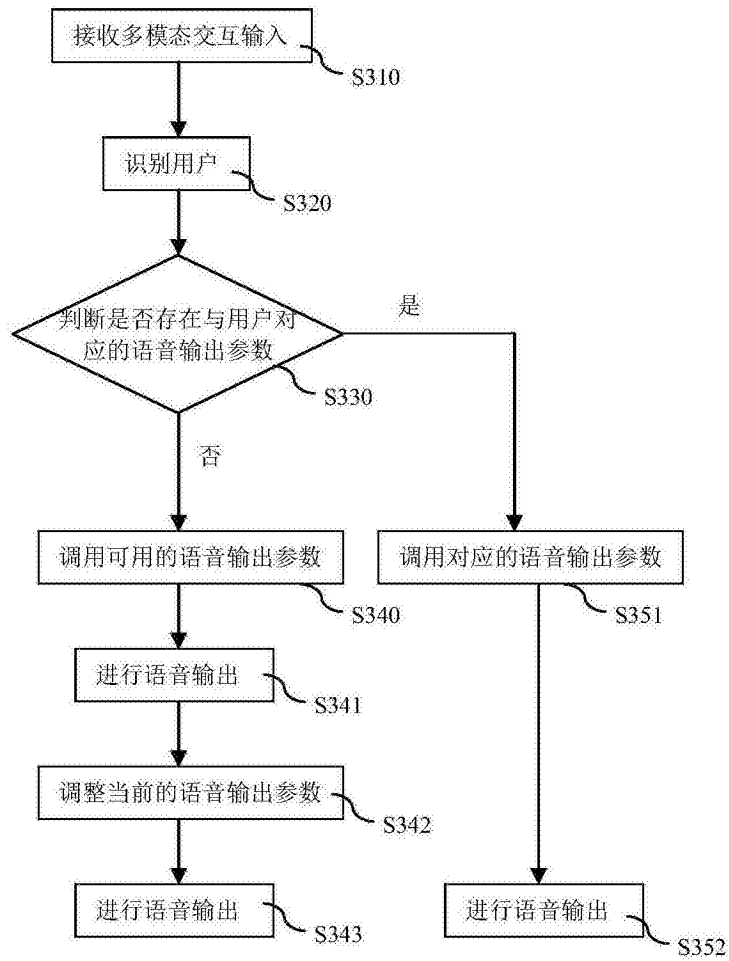


图3

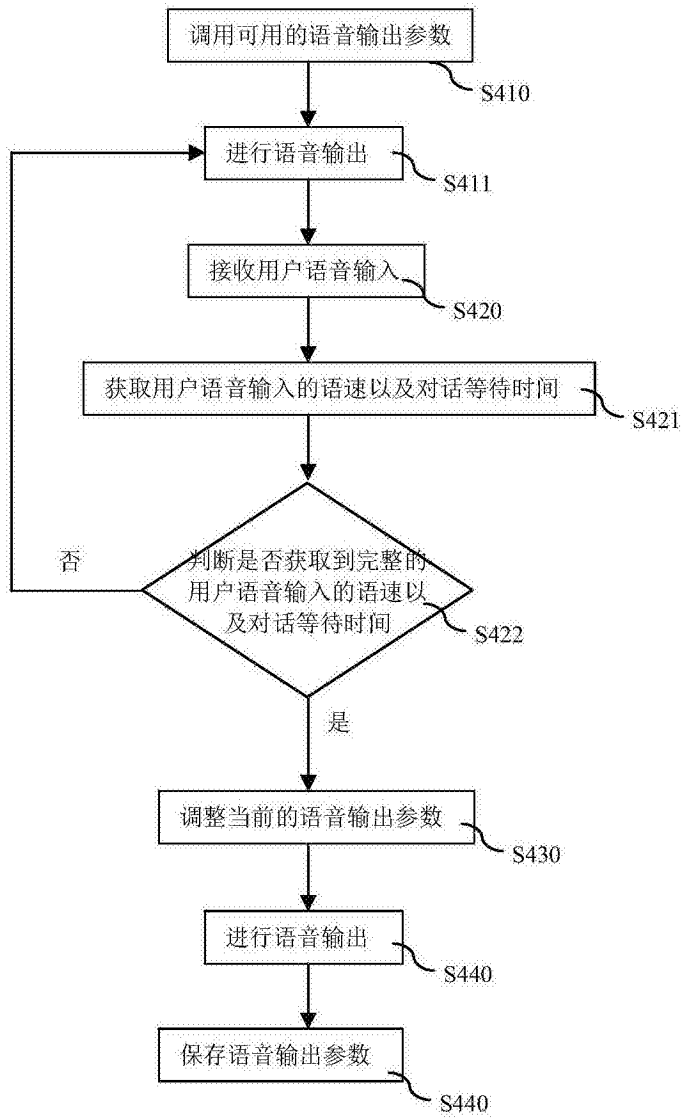


图4

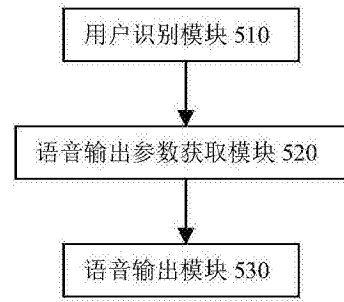


图5

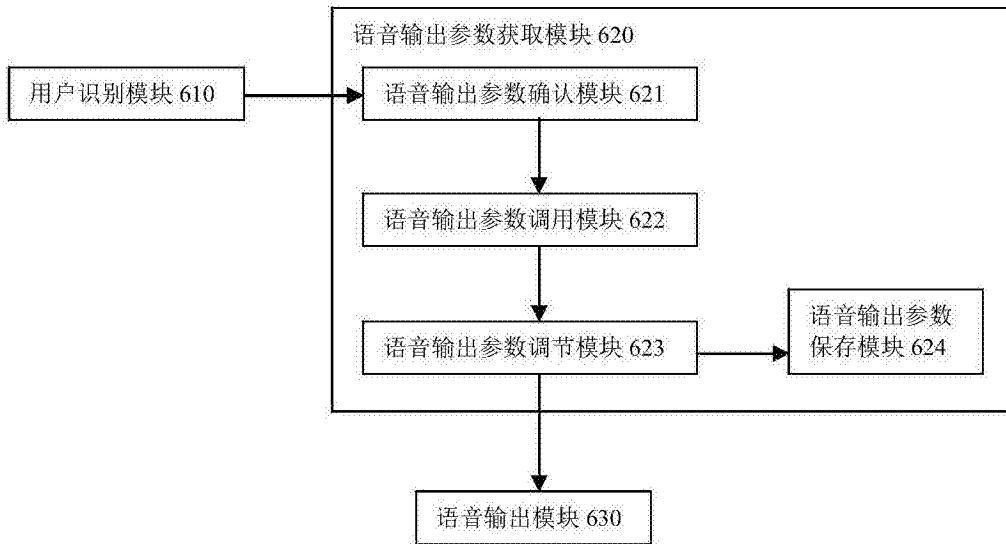


图6