



(12) 发明专利

(10) 授权公告号 CN 114299931 B

(45) 授权公告日 2025. 05. 30

(21) 申请号 202111574477.7

(22) 申请日 2021.12.21

(65) 同一申请的已公布的文献号

申请公布号 CN 114299931 A

(43) 申请公布日 2022.04.08

(73) 专利权人 广州小鹏汽车科技有限公司

地址 510000 广东省广州市天河区岑村松
岗大街8号

(72) 发明人 赵群 张天宇 王亭玉 孙文平
潘晓彤 赵恒艺

(74) 专利代理机构 北京清亦华知识产权代理事
务所(普通合伙) 11201

专利代理师 赵静

(51) Int. Cl.

G10L 15/06 (2013.01)

G10L 15/22 (2006.01)

G10L 15/26 (2006.01)

G10L 25/51 (2013.01)

G06F 40/30 (2020.01)

B60R 16/023 (2006.01)

(56) 对比文件

CN 112581955 A, 2021.03.30

CN 113806470 A, 2021.12.17

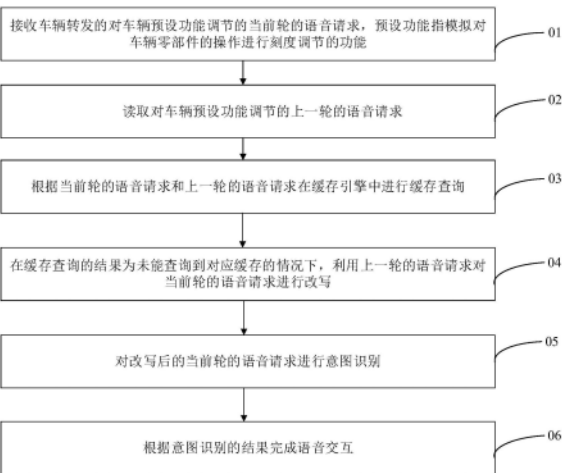
审查员 白雪旻

(54) 发明名称

语音交互方法及其装置、服务器和可读存储
介质

(57) 摘要

本发明公开了一种语音交互方法及其装置、服务器和可读存储介质。该语音交互方法包括：接收车辆转发的对车辆预设功能调节的当前轮的语音请求，预设功能指模拟对车辆零部件的操作进行刻度调节的功能；读取对车辆预设功能调节的上一轮的语音请求；根据当前轮的语音请求和上一轮的语音请求在缓存引擎中进行缓存查询；在缓存查询的结果为未能查询到对应缓存的情况下，利用上一轮的语音请求对当前轮的语音请求进行改写；对改写后的当前轮的语音请求进行意图识别；根据意图识别的结果完成语音交互。本发明结合两轮语音请求，使用高频缓存引擎与意图识别相结合的方式识别语音请求的意图，实现多轮语音请求下，用户意图的准确识别。



1. 一种语音交互方法,其特征在于,包括:

接收车辆转发的对车辆预设功能调节的当前轮的语音请求,所述预设功能指模拟对车辆零部件的操作进行刻度调节的功能;

读取对车辆预设功能调节的上一轮的语音请求;

根据所述当前轮的语音请求和所述上一轮的语音请求在缓存引擎中进行缓存查询;

在所述缓存查询的结果为未能查询到对应缓存的情况下,利用所述上一轮的语音请求对所述当前轮的语音请求进行改写;

对改写后的所述当前轮的语音请求进行意图识别;

根据所述意图识别的结果完成语音交互;

所述语音交互方法包括:

将出现频率大于预设频率的相邻两轮语音请求添加到所述缓存引擎。

2. 根据权利要求1所述的语音交互方法,其特征在于,所述语音交互方法包括:

建立当前轮的语音请求与预设意图之间的映射关系。

3. 根据权利要求2所述的语音交互方法,其特征在于,所述语音交互方法包括:

在所述缓存查询的结果为查询到对应缓存的情况下,根据所述映射关系确定所述当前轮的语音请求对应的预设意图为目标意图以完成语音交互。

4. 根据权利要求1所述的语音交互方法,其特征在于,所述利用所述上一轮的语音请求对所述当前轮的语音请求进行改写,包括:

通过改写训练数据训练得到改写模型,所述改写训练数据包括相邻两轮语音请求;

利用所述上一轮的语音请求和所述改写模型对所述当前轮的语音请求进行改写。

5. 根据权利要求1所述的语音交互方法,其特征在于,所述对改写后的所述当前轮的语音请求进行意图识别,包括:

通过意图训练数据训练得到意图识别模型,所述意图训练数据与可进行刻度调节的车辆零部件和所述车辆零部件的刻度调节范围相关;

利用所述意图识别模型对改写后的所述当前轮的语音请求进行意图识别。

6. 根据权利要求5所述的语音交互方法,其特征在于,所述根据所述意图识别的结果完成语音交互,包括:

获取所述意图识别的结果对应各个预设意图的意图判别概率;

将所述意图判别概率大于概率阈值的一个所述预设意图确定为所述当前轮的语音请求对应的目标意图以完成语音交互。

7. 根据权利要求6所述的语音交互方法,其特征在于,所述预设意图包括:音量调大、音量调小、风量调大、风量调小、温度调高、温度调低、地图放大、地图缩小、屏幕调亮、屏幕调暗、屏幕上滑、屏幕下滑、仪表调亮、仪表调暗、氛围灯调亮、氛围灯调暗、座椅向前、座椅向后、座椅升高、座椅降低、椅背向前、椅背向后、车窗上升和车窗下降中的至少一种。

8. 根据权利要求6所述的语音交互方法,其特征在于,所述语音交互方法包括:

在各个所述预设意图的所述意图判别概率均不大于概率阈值的情况下,确定所述当前轮的语音请求的意图为非刻度调节意图。

9. 一种语音交互装置,其特征在于,所述语音交互装置包括:

接收指令模块,所述接收指令模块用于接收车辆转发的对车辆预设功能调节的当前轮

的语音请求,所述预设功能指模拟对车辆零部件的操作进行刻度调节的功能;

读取指令模块,所述读取指令模块用于读取对车辆预设功能调节的上一轮的语音请求;

查询模块,所述查询模块用于根据所述当前轮的语音请求和所述上一轮的语音请求在缓存引擎中进行缓存查询;

改写模块,所述改写模块用于在所述缓存查询的结果为未能查询到对应缓存的情况下,利用所述上一轮的语音请求对所述当前轮的语音请求进行改写;

意图识别模块,所述意图识别模块用于对改写后的所述当前轮的语音请求进行意图识别;

交互模块,所述交互模块用于根据所述意图识别的结果完成语音交互;

所述查询模块还用于将出现频率大于预设频率的相邻两轮语音请求添加到所述缓存引擎。

10.一种服务器,其特征在于,所述服务器包括处理器和存储器,所述存储器上存储有计算机程序,当所述计算机程序被所述处理器执行时,实现权利要求1-8任一项所述的语音交互方法。

11.一种包含有计算机程序的非易失性计算机可读存储介质,其特征在于,当所述计算机程序被一个或多个处理器执行时,实现权利要求1-8任一项所述的语音交互方法。

语音交互方法及其装置、服务器和可读存储介质

技术领域

[0001] 本发明涉及语音技术领域,特别涉及一种语音交互方法及其装置、服务器和可读存储介质。

背景技术

[0002] 目前在智能汽车场景中,可以应用语音交互实现用户对车辆零部件的控制,比如“打开车窗”,“音量调高”等,但是,对于用户希望进行连续调节的场景,在语音场景下体现为多轮交互,用户在上一轮语音交互后很自然地省略后续每轮对话的部分内容,例如以下用户跟语音助手小P的对话:

[0003] 用户:今天天气怎么样?

[0004] 小P:广州今天晴天26-30°。

[0005] 用户:上海(的天气)呢?

[0006] 在多轮对话中,像上述第一个示例中用户字面含义是问上海,但其实是想问上海的天气,将部分内容省略符合人对话的习惯,但是这样可能会导致车辆的车载系统对某些轮次的语音请求不能准确识别,或者提示听不懂。

[0007] 进一步地,如果用户需要调整音量,可以通过对车上调整汽车音量的机械旋钮进行操作,将机械旋钮旋转到想要的音量,但是如果使用语音调整音量,只能够调大或者调小。在如下第二个示例中:

[0008] 用户:音量调大

[0009] 小P:音量已调大

[0010] 用户:大大大

[0011] 从第二个示例可以看出,当前车辆的车载系统对当前轮的大大大不能准确识别,或者提示听不懂,这样情况不能满足用户对类似机械旋钮一样精准的刻度连续调节的需求。

发明内容

[0012] 本发明实施方式提供一种语音交互方法及其装置、服务器和可读存储介质。

[0013] 本发明实施方式提供一种语音交互方法。该语音交互方法包括:接收车辆转发的对车辆预设功能调节的当前轮的语音请求,所述预设功能指模拟对车辆零部件的操作进行刻度调节的功能;读取对车辆预设功能调节的上一轮的语音请求;根据所述当前轮的语音请求和所述上一轮的语音请求在缓存引擎中进行缓存查询;在所述缓存查询的结果为未能查询到对应缓存的情况下,利用所述上一轮的语音请求对所述当前轮的语音请求进行改写;对改写后的所述当前轮的语音请求进行意图识别;根据所述意图识别的结果完成语音交互。

[0014] 如此,本发明的语音交互方法可以在接收到用户对于车辆零部件进行刻度调节的语音请求后,通过读取上一轮的语音请求,结合两轮语音请求查询是否命中缓存,在查询不

到对应缓存的情况下,利用上一轮的语音请求改写当前轮的语音请求,使得改写后的语音请求可以被车辆的车载系统识别出相应的意图,进而根据意图识别的结果以语音交互方式实现对车辆零部件进行刻度调节。使用高频缓存引擎与意图识别相结合的方式识别语音请求的意图,实现多轮语音请求下,用户意图的准确识别。

[0015] 所述语音交互方法包括:将出现频率大于预设频率的相邻两轮语音请求添加到所述缓存引擎。

[0016] 如此,本发明缓存引擎的缓存由出现频率大于预设频率相邻两轮语音请求构成,实现高频集语音请求的统计。

[0017] 所述语音交互方法包括:建立当前轮的语音请求与预设意图之间的映射关系。

[0018] 如此,本发明建立当前轮的语音请求与预设意图之间的映射关系后,将每个预设意图与对应的相邻两轮语音请求相关联,从而在缓存引擎中查询可以实现语音请求对应意图的确定。

[0019] 所述语音交互方法包括:在所述缓存查询的结果为查询到对应缓存的情况下,根据所述映射关系确定所述当前轮的语音请求对应的预设意图为目标意图以完成语音交互。

[0020] 如此,本发明在缓存引擎中查询到当前轮的语音请求和相邻的上一轮的语音请求,根据建立好的映射关系,可以直接确定当前轮的语音请求的目标意图,从而可以根据确定好的当前轮的语音请求对应的目标意图完成语音交互。

[0021] 所述利用所述上一轮的语音请求对所述当前轮的语音请求进行改写,包括:通过改写训练数据训练得到改写模型,所述改写训练数据包括相邻两轮语音请求;利用所述上一轮的语音请求和所述改写模型对所述当前轮的语音请求进行改写。

[0022] 如此,本发明通过机器学习的方式,由相邻两轮语音请求训练得到改写模型,从而根据上一轮的语音请求和改写模型可以实现对当前轮的语音请求的改写,使得改写后的语音请求可以被车辆的车载系统识别出相应的意图。

[0023] 所述对改写后的所述当前轮的语音请求进行意图识别,包括:通过意图训练数据训练得到意图识别模型,所述意图训练数据与可进行刻度调节的车辆零部件和所述车辆零部件的刻度调节范围相关;利用所述意图识别模型对改写后的所述当前轮的语音请求进行意图识别。

[0024] 如此,本发明通过机器学习的方式,由可进行刻度调节的车辆零部件和所述车辆零部件的刻度调节范围对应的训练数据训练得到意图识别模型,进而对改写后的语音请求进行意图识别,实现用户意图的准确识别。

[0025] 所述根据所述意图识别的结果完成语音交互,包括:获取所述意图识别的结果对应各个预设意图的意图判别概率;将所述意图判别概率大于概率阈值的一个所述预设意图确定为所述当前轮的语音请求对应的目标意图以完成语音交互。

[0026] 如此,本发明的语音交互方法可以获取意图识别的结果对应各个预设意图的意图判别概率,将意图判别概率大于概率阈值的一个预设意图确定为语音请求对应的目标意图,从而识别用户精准调节车辆零部件的意图。

[0027] 所述预设意图包括:音量调大、音量调小、风量调大、风量调小、温度调高、温度调低、地图放大、地图缩小、屏幕调亮、屏幕调暗、屏幕上滑、屏幕下滑、仪表调亮、仪表调暗、氛围灯调亮、氛围灯调暗、座椅向前、座椅向后、座椅升高、座椅降低、椅背向前、椅背向后、车

窗上升和车窗下降中的至少一种。

[0028] 如此,设置了多种预设意图可以进一步为识别用户的语音交互意图奠定基础。

[0029] 所述语音交互方法包括:在各个所述预设意图的所述意图判别概率均不大于概率阈值的情况下,确定所述当前轮的语音请求的意图为非刻度调节意图。

[0030] 如此,在各个预设意图的意图判别概率均不大于概率阈值的情况下,确定该语音请求为非刻度调节意图,可以排除非刻度调节意图的语音请求。

[0031] 本发明还提供一种语音交互装置。所述语音交互装置包括:接收指令模块、读取指令模块、查询模块、改写模块、意图识别模块和交互模块。所述接收指令模块用于接收车辆转发的对车辆预设功能调节的当前轮的语音请求,所述预设功能指模拟对车辆零部件的操作进行刻度调节的功能;所述读取指令模块用于读取对车辆预设功能调节的上一轮的语音请求;所述查询模块用于根据所述当前轮的语音请求和所述上一轮的语音请求在缓存引擎中进行缓存查询;所述改写模块用于在所述缓存查询的结果为未能查询到对应缓存的情况下,利用所述上一轮的语音请求对所述当前轮的语音请求进行改写;所述意图识别模块用于对改写后的所述当前轮的语音请求进行意图识别;所述交互模块用于根据所述意图识别的结果完成语音交互。

[0032] 如此,本发明的语音交互装置可以在接收到用户对于车辆零部件进行刻度调节的语音请求后,通过读取上一轮的语音请求,结合两轮语音请求查询是否命中缓存,在查询不到对应缓存的情况下,利用上一轮的语音请求改写当前轮的语音请求以使得改写后的语音请求可以被车辆的车载系统识别出相应的意图,进而根据意图识别结果以语音交互方式实现对车辆零部件进行刻度调节。使用高频缓存引擎与意图识别相结合的方式识别语音请求的意图,实现多轮语音请求下,用户意图的准确识别。

[0033] 本发明提供一种服务器。所述服务器包括处理器和存储器,所述存储器上存储有计算机程序,当所述计算机程序被所述处理器执行时,实现上述任意一项实施方式所述的语音交互方法。

[0034] 如此,本发明的服务器通过处理器执行计算机程序,可以在接收到用户对于车辆零部件进行刻度调节的语音请求后,通过读取上一轮的语音请求,结合两轮语音请求查询是否命中缓存,在查询不到对应缓存的情况下,利用上一轮的语音请求改写当前轮的语音请求以使得改写后的语音请求可以被车辆的车载系统识别出相应的意图,进而根据意图识别结果以语音交互方式实现对车辆零部件进行刻度调节。使用高频缓存引擎与意图识别相结合的方式识别语音请求的意图,实现多轮语音请求下,用户意图的准确识别。

[0035] 本发明实施方式还提供一种包含有计算机程序的非易失性计算机可读存储介质。当所述计算机程序被一个或多个处理器执行时,实现上述任意一项实施方式所述的语音交互方法。

[0036] 如此,本发明的可读存储介质存储的计算机程序被处理器执行时,可以在接收到用户对于车辆零部件进行刻度调节的语音请求后,通过读取上一轮的语音请求,结合两轮语音请求查询是否命中缓存,在查询不到对应缓存的情况下,利用上一轮的语音请求改写当前轮的语音请求以使得改写后的语音请求可以被车辆的车载系统识别出相应的意图,进而根据意图识别结果以语音交互方式实现对车辆零部件进行刻度调节。使用高频缓存引擎与意图识别相结合的方式识别语音请求的意图,实现多轮语音请求下,用户意图的准确识

别。

[0037] 本发明实施方式的附加方面和优点将在下面的描述中部分给出,部分将从下面的描述中变得明显,或通过本发明的实践了解到。

附图说明

[0038] 本发明的上述和/或附加的方面和优点可以从结合下面附图对实施方式的描述中将变得明显和容易理解,其中:

- [0039] 图1是本发明的语音交互方法的流程示意图;
- [0040] 图2是本发明的语音交互装置的结构示意图;
- [0041] 图3是本发明的语音交互方法的流程示意图;
- [0042] 图4是本发明的语音交互装置的结构示意图;
- [0043] 图5是本发明的语音交互方法的流程示意图;
- [0044] 图6是本发明的语音交互方法的流程示意图;
- [0045] 图7是本发明的语音交互方法的流程示意图;
- [0046] 图8是本发明的语音交互装置的结构示意图;
- [0047] 图9是本发明的语音交互方法的流程示意图;
- [0048] 图10是本发明的语音交互装置的结构示意图;
- [0049] 图11是本发明的语音交互方法的流程示意图;
- [0050] 图12是本发明的语音交互装置中的交互模块的结构示意图;
- [0051] 图13是本发明的服务器的结构示意图;
- [0052] 图14是本发明的计算机可读存储介质的结构示意图。

具体实施方式

[0053] 下面详细描述本发明的实施方式,所述实施方式的示例在附图中示出,其中,相同或类似的标号自始至终表示相同或类似的元件或具有相同或类似功能的元件。下面通过参考附图描述的实施方式是示例性的,仅用于解释本发明的实施方式,而不能理解为对本发明的实施方式的限制。

[0054] 目前,车辆的语音交互系统在用户进行多轮语音请求的情况下,例如,用户第一轮语音请求为“屏幕调亮”,而第二轮语音请求采用精简语音请求“亮亮亮”的情况下,语音交互系统不能根据用户的语音请求精准识别出用户第二轮的需求为屏幕亮度提高3个刻度,不能正确的下发车辆控制指令,使得屏幕亮度准确提高用户需要的三个亮度,用户体验不佳。

[0055] 为了解决上述问题,请参阅图1,本发明提供了一种语音交互方法。该语音交互方法包括:

- [0056] 01,接收车辆转发的对车辆预设功能调节的当前轮的语音请求,预设功能指模拟对车辆零部件的操作进行刻度调节的功能;
- [0057] 02,读取对车辆预设功能调节的上一轮的语音请求;
- [0058] 03,根据当前轮的语音请求和上一轮的语音请求在缓存引擎中进行缓存查询;
- [0059] 04,在缓存查询的结果为未能查询到对应缓存的情况下,利用上一轮的语音请求

对当前轮的语音请求进行改写；

[0060] 05,对改写后的当前轮的语音请求进行意图识别；

[0061] 06,根据意图识别的结果完成语音交互。

[0062] 请参阅图2,本发明还提供一种语音交互装置10。语音交互装置10包括:接收指令模块11、读取指令模块12、查询模块13、改写模块14、意图识别模块15和交互模块16。

[0063] 步骤01可以由接收指令模块11实现,步骤02可以由读取指令模块12实现,步骤03可以由查询模块13实现,步骤04可以由改写模块14实现,步骤05可以由意图识别模块15实现和步骤06可以由交互模块16实现。也即是说,接收指令模块11可以用于接收车辆转发的对车辆预设功能调节的当前轮的语音请求,预设功能指模拟对车辆零部件的操作进行刻度调节的功能;读取指令模块12可以用于读取对车辆预设功能调节的上一轮的语音请求;查询模块13可以用于根据当前轮的语音请求和上一轮的语音请求在缓存引擎中进行缓存查询;改写模块14可以用于在缓存查询的结果为未能查询到对应缓存的情况下,利用上一轮的语音请求对当前轮的语音请求进行改写;意图识别模块15可以用于对改写后的当前轮的语音请求进行意图识别;交互模块16可以用于根据意图识别的结果完成语音交互。

[0064] 用户使用语音交互模拟对车辆零部件进行刻度调节的过程,对应的语音请求可以包括但不限于“屏幕亮亮亮”、“音量大大大”、“座椅后后后”。其中,预设功能是指通过车辆零部件完成刻度调节的功能,其中的车辆零部件可以指机械旋钮或按钮等实体部件,这些是可以进行调节刻度的部件。目前在智能汽车中,对于用户希望进行连续调节的场景,在语音场景下体现为多轮交互,比如用户上一轮的语音请求为“音量大大大”,系统音量调大后,用户又发出当前轮的语音请求“小一点小一点”,这时候系统针对当前轮的语音请求会不识别,或者提示听不懂,不能满足用户对类似机械旋钮一样精准的刻度连续调节的需求。

[0065] 本发明可以在接收到用户对于车辆预设功能的语音请求后,通过读取上一轮的语音请求,结合两轮语音请求查询是否命中缓存,在查询不到对应缓存的情况下,利用上一轮的语音请求改写当前轮的语音请求以使得改写后的语音请求可以被系统识别出相应的意图,从而对改写后的当前轮的语音请求进行意图识别后,可以准确识别出用户的意图,进而可以根据意图识别结果下发控制指令,以控制对应的车辆零部件,完成语音交互。使用高频缓存引擎与意图识别相结合的方式识别语音请求的意图,实现多轮语音请求下,准确识别出通过语音交互来模拟对车辆零部件的操作实现刻度调节的用户意图。

[0066] 需要说明的是,在接收到用户对于车辆预设功能的当前轮的语音请求后,对接收到的当前轮的语音请求进行语音识别得到当前轮语音识别文本以便后续处理,例如,对用户输入的当前轮的语音请求“屏幕亮亮亮”进行语音识别,得到的当前轮识别文本即为“屏幕亮亮亮”。

[0067] 在实际情况下可能受到车辆硬件限制,或者因为网络不稳定性,用户表述口语化或者方言化等原因导致ASR识别后的文本指令不够清晰准确,可以对接收到的当前轮的语音请求进行预处理,预处理包括对一些常规文本纠错,比如“音量深深深深深”纠正为“音量增增增增增”,以及一些无意义词语的去除等,比如“啊”,“请”等。

[0068] 请结合图3,在步骤01之前,语音交互方法可以包括:

[0069] 011,确定车辆零部件的控制范围及非控制范围。

[0070] 请结合图4,语音交互装置10还包括第一确定模块111。

[0071] 步骤011可以由第一确定模块111实现。也即是说,第一确定模块111可以用于确定车辆零部件的控制范围及非控制范围。

[0072] 可以理解地,车辆并非所有功能的调节都可以、能够或有需要进行精准的刻度调节。例如,座椅在各个方向上的移动可以通过车辆零部件进行调节。而车门则没有类似旋钮、按键等车辆零部件来实现刻度调节,而通常仅通过车门把手进行开关。因此,座椅调节是属于车辆零部件的控制范围、而车门调节则属于车辆零部件的非控制范围。

[0073] 获取车辆零部件的信息,根据车辆零部件的信息,确定可通过零部件进行刻度调节的硬件,确定为车辆零部件的控制范围,将不可通过车辆零部件进行调节的硬件确定为非控制范围。

[0074] 首先,确定在车辆上可以进行刻度调节的零部件,例如:“音量旋钮”,“屏幕亮度按钮”,“空调风量旋钮/按钮”,“座椅调节旋钮/按钮”等。进一步,确定车辆零部件的控制范围可包括:车载音响、车辆内的屏幕、车辆空调、车辆座椅、车内的氛围灯、车辆外部的车灯、或车窗等。车辆零部件的非控制范围可包括:车门、后视镜、后备箱等。

[0075] 在后续语音交互的过程中,可在语音请求针对车辆零部件的非控制范围的情况下进行语音提示。

[0076] 如此,通过收集车辆零部件信息,确认可通过零部件进行刻度调节的功能,从而确定车辆零部件的控制范围,也即是可通过语音交互进行刻度调节的控制范围。

[0077] 语音交互方法还包括:

[0078] 012,确定车辆零部件的可调节范围。

[0079] 语音交互装置10还包括第二确定模块112。

[0080] 步骤012可以由第二确定模块112实现。也即是说,第二确定模块112用于确定车辆零部件的可调节范围。

[0081] 可以理解,在确定车辆零部件的控制范围和非控制范围后,需要针对控制范围中的每一个车辆零部件确定的可调节范围。车辆零部件的可调节范围与通过操作该车辆零部件进行调节的刻度范围相对应。对应不同车辆零部件,可调节范围可以是档位或量程。例如,屏幕亮度按钮累计连续按压5次,屏幕亮度依次调整1至5个档位的亮度至最大亮度,则该屏幕亮度按钮的可调节范围为1至5个档位。又如,对座椅进行前后调节的旋钮的总刻度值为90,则该座椅调节旋钮的可调节范围为刻度值1~90。

[0082] 语音交互方法还包括:

[0083] 013,将控制范围和可调节范围映射到预设意图。

[0084] 语音交互装置100还包括映射模块113。

[0085] 步骤013可以由映射模块113实现。也即是说,映射模块113可以用于将控制范围和可调节范围映射到预设意图。

[0086] 将车辆零部件的控制范围和每个车辆零部件的可调节范围,映射到意图识别模型所能够理解的意图体系。针对车辆零部件的控制范围中的对象和对应的车辆零部件的可调节范围均制定一个相应的预设意图。例如:system_volume_up代表着预设意图“音量调大”和system_volume_down代表着预设意图“音量调小”,并且包含了所有可调节的范围说法,例如“音量大大大”为预设意图对应的system_volume_up,“音量大大大大”也会对此意图。如此,针对零部件控制范围和车辆零部件的可调节范围制定了一套具体的意图映射体系。

[0087] 预设意图可以包括：音量调大、音量调小、风量调大、风量调小、温度调高、温度调低、地图放大、地图缩小、屏幕调亮、屏幕调暗、屏幕上滑、屏幕下滑、仪表调亮、仪表调暗、氛围灯调亮、氛围灯调暗、座椅向前、座椅向后、座椅升高、座椅降低、椅背向前、椅背向后、车窗上升和车窗下降中的至少一种。

[0088] 如此，设置了多种预设意图可以进一步为识别用户的语音交互意图奠定基础，根据用户提供的带有精简词的语音请求进行不同意图的识别，从而实现相应的目标意图。

[0089] 请结合图5，语音交互方法包括：

[0090] 014，将出现频率大于预设频率的相邻两轮语音请求添加到缓存引擎。

[0091] 步骤014可以由查询模块13实现。也即是说，查询模块13可以用于将出现频率大于预设频率的相邻两轮语音请求添加到缓存引擎。

[0092] 如此，本发明缓存引擎的缓存由出现频率大于预设频率相邻两轮语音请求构成，实现高频集语音请求的统计。

[0093] 首先服务器可以在用户允许的情况下收集一段时间内用户的历史语音请求，此处收集到的语音请求需要至少包含两轮语音请求。其中，预计需要收集超过1万条历史语音请求。

[0094] 其次，服务器可以对收集到的历史语音请求进行简单的筛选，以筛选掉明显语义不明确的语音请求，以及一些只包含语气词，比如“啊”，“哦”等较短的语音请求，留下语义明确同时包含具体目的的语音请求，比如“导航到公司”，“帮我打开空调”，“搜一下附近的医院”，“播放歌手A的歌曲”，“今天天气怎么样”等；并且在筛选中去掉只有一轮的语音请求。

[0095] 服务器可以对筛选后的语音请求进行高频统计，统计相邻两轮语音请求的出现频率。其中，统计相邻两轮语音请求作为唯一值出现的次数，当出现次数大于一定次数时，可以认为对应的相邻两轮语音请求的出现频率大于预设频率。

[0096] 比如，上一轮的语音请求为“音量大一点大一点”，当前轮的语音请求为“小一点”的情况，在筛选出来的语音请求中出现的次数超过预定次数，则可以将“音量大一点大一点”和“小一点”的相邻两轮语音请求加入到缓存引擎中。

[0097] 语音交互方法还包括：

[0098] 015，建立当前轮的语音请求与预设意图之间的映射关系。

[0099] 步骤015可以由查询模块13实现。也即是说，查询模块13可以用于建立当前轮的语音请求与预设意图之间的映射关系。

[0100] 如此，本发明建立当前轮的语音请求与预设意图之间的映射关系后，将每个预设意图与对应的相邻两轮语音请求相关联，从而在缓存引擎中查询可以实现语音请求对应意图的确定。

[0101] 应当理解，上一轮的语音请求和当前轮的语音请求为相邻的两轮，其中确定缓存引擎的相邻两轮语音请求中，当前轮的语音请求与预设意图之间的映射关系后，即可确定上一轮的语音请求和当前轮的语音请求结合是否属于高频集指令，以及是否可以根据高频集指令对应的预设意图确定当前轮的语音请求的目标意图。

[0102] 比如，上一轮的语音请求为“音量大一点大一点”，当前轮的语音请求为“小一点”，那么可以关联“音量大一点大一点”和“小一点”这相邻两轮语音请求中，当前轮的语音请求

“小一点”对应的预设意图为“音量调小”。

[0103] 请结合图6,语音交互方法包括:

[0104] 07,在缓存查询的结果为查询到对应缓存的情况下,根据映射关系确定当前轮的语音请求对应的预设意图为目标意图以完成语音交互。

[0105] 步骤07可以由交互模块16实现。也即是说,交互模块16可以用于在缓存查询的结果为查询到对应缓存的情况下,根据映射关系确定当前轮的语音请求对应的预设意图为目标意图以完成语音交互。

[0106] 本发明在缓存引擎中查询到与当前轮的语音请求和上一轮的语音请求相应的相邻两轮语音请求的情况下,根据建立好的当前轮的语音请求与预设意图之间的映射关系,可以直接确定当前轮的语音请求对应的目标意图,从而可以根据确定好的当前轮的语音请求对应的目标意图完成语音交互。

[0107] 例如,上一轮的语音请求为“音量大一点大一点”,当前轮的语音请求为“小一点”,若在缓存引擎中查询到缓存的相邻两轮语音请求为“音量大一点大一点”和“小一点”,且“小一点”对应的预设意图为“音量调小”,则可以直接确定当前轮的语音请求的目标意图为查询到的预设意图“音量调小”,从而可以根据“音量调小”的意图通过语音交互模拟对车辆零部件的操作,实现多轮语音请求下,用户意图的准确识别。

[0108] 请结合图7,步骤04包括:

[0109] 041,通过改写训练数据训练得到改写模型,改写训练数据包括相邻两轮语音请求;

[0110] 042,利用上一轮的语音请求和改写模型对当前轮的语音请求进行改写。

[0111] 请结合图8,语音交互装置10包括改写训练模块114。

[0112] 步骤041可以用于改写训练模块114实现,步骤042可以由改写模块14实现。也即是说,改写训练模块114可以用于通过改写训练数据训练得到改写模型。改写模块14可以用于利用上一轮的语音请求和改写模型对当前轮的语音请求进行改写。

[0113] 本发明通过机器学习的方式,由相邻两轮语音请求训练得到改写模型,从而根据上一轮的语音请求和改写模型可以实现对当前轮的语音请求的改写,使得改写后的语音请求可以被车辆的车载系统识别出相应的意图。其中,对于改写模型,可使用BERT (Bidirectional Encoder Representation from Transformers)和序列标注的方式进行模型训练,从而获取训练好的改写模型。

[0114] 其中,改写数据可以是上述筛选后的语音请求中的相邻两轮语音请求经过标注后得到,可以由人工对相邻两轮语音请求中的当前轮的语音请求进行改写标注,例如,上一轮的语音请求为“音量大一点大一点”,当前轮的语音请求为“小一点”,那么可以将当前轮的语音请求改写标注为“音量小一点”。如此,将标注好的相邻两轮语音请求输入到建立好的改写模型,在训练过程中,改写模型可以通过特征提取的方式学习如何通过相邻两轮语音请求将标注前的当前轮的语音请求改写成标注后的当前轮的语音请求。

[0115] 在训练过程中,标注后的语音请求中的相邻两轮语音请求划分为改写训练集和改写验证集,划分比例可根据需求设定,在此不作限定。例如改写训练集80%,改写验证集为20%。对于建立好的改写模型,先通改写训练集中的至少部分数据用于训练改写模型,然后利用改写验证集的至少部分数据对训练后的改写模型的准确率进行改写验证。在改写验证

的准确率没有达到改写准确率阈值的情况下,再次通过改写训练集的至少另一部分数据对改写模型进行训练,以及再次利用改写验证集的至少另一部分数据对再次训练后的改写模型的准确率进行改写验证,如此重复训练和改写验证的过程,直到改写验证的准确率达到改写准确率阈值时,可以认为改写模型已经达标,完成改写模型的训练。

[0116] 需要说明的是,改写训练集和改写验证集中的每个数据均只使用一次,在改写模型遍历改写训练集和改写验证集的所有数据均未能训练达标的情况下,可以再次在用户允许的情况下收集更多的语音请求,从而筛选并标注得到更多的改写训练数据对改写模型进行训练,从而保证改写模型能够准确对语音请求进行改写。

[0117] 请结合图9,步骤05包括:

[0118] 051,通过意图训练数据训练得到意图识别模型,意图训练数据与可进行刻度调节的车辆零部件和车辆零部件的刻度调节范围相关;

[0119] 052,利用意图识别模型对改写后的当前轮的语音请求进行意图识别。

[0120] 请结合图10,语音交互装置10包括意图训练模块115。

[0121] 步骤051可以由意图训练模块115实现,步骤052可以由意图识别模块15实现。也就是说,意图训练模块115可以用于通过意图训练数据训练得到意图识别模型。意图识别模块15可以用于利用意图识别模型对改写后的当前轮的语音请求进行意图识别。

[0122] 本发明通过机器学习的方式,由可进行刻度调节的车辆零部件和车辆零部件的刻度调节范围对应的训练数据训练得到意图识别模型,进而对改写后的当前轮的语音请求进行意图识别,实现用户意图的准确识别。其中,模型训练可以利用BERT、ALBERT、XLNet、RoBERTa等模型。

[0123] 其中,意图训练数据与可进行刻度调节的车辆零部件和零部件的刻度调节范围相关。车辆零部件指的是在智能汽车上可以进行刻度调节的零部件,例如:“音量旋钮”,“屏幕亮度按钮”,“空调风量旋钮/按钮”,“座椅调节旋钮/按钮”等。车辆零部件的可调节范围与通过操作该车辆零部件进行调节的刻度范围相对应。对应不同车辆零部件,可调节范围可以是档位或量程。

[0124] 其中,意图训练数据可以是上述筛选后的语音请求中的相邻两轮语音请求中上一轮的语音请求经过标注后得到,可以由人工对相邻两轮语音请求中的上一轮的语音请求进行意图标注,可以理解,上一轮的语音请求中应当包括用户需要调节的意图相关的内容,例如,上一轮的语音请求为“音量大一点大一点”,用户需要调节的是音量调大2次,此时,可以通过人工将上一轮的语音请求对应的意图标注为“音量调大”。如此,将标注好的上一轮的语音请求给到建立好的意图识别模型,在训练过程中,意图识别模型可以通过特征提取学习如何通过输入的语音请求识别出用户想要实现的目标意图。

[0125] 在训练过程中,可以将标注后的上一轮的语音请求划分为意图训练集和意图验证集,划分比例可根据需求设定,在此不作限定。例如意图训练集80%,意图验证集为20%。对于建立好的意图识别模型,先利用意图训练集中的至少部分数据用于训练意图识别模型,然后利用意图验证集的至少部分数据对训练后的意图识别模型的准确率进行意图验证。在意图验证的准确率没有达到意图准确率阈值的情况下,再次通过意图训练集的至少另一部分数据对意图识别模型进行训练,以及再次利用意图验证集的另一部分数据对再次训练后的意图识别模型的准确率进行意图验证,如此重复训练和意图验证的过程,直到意图验证

的准确率达到意图准确率阈值时,可以认为意图识别模型已经达标,完成意图识别模型的训练。

[0126] 需要说明的是,意图训练集和意图验证集中的每个数据均只使用一次,在意图识别模型遍历意图训练集和意图验证集的所有数据均未能训练达标的情况下,可以再次在用户允许的情况下收集更多的语音请求,从而筛选并标注得到更多的意图训练数据对意图识别模型进行训练,从而保证意图识别模型能够准确识别输入的语音请求对应的意图。

[0127] 可以理解,上述改写模型和意图识别模型的训练可以离线进行训练,将离线训练好的改写模型和意图识别模型部署到服务器后,服务器可以在接收到当前轮的语音请求后利用上一轮的语音请求改写模型对当前轮的语音请求进行改写,以及利用意图识别模型对改写后的当前轮的语音请求进行意图识别。

[0128] 请结合图11,步骤06包括:

[0129] 061,获取意图识别的结果对应各个预设意图的意图判别概率;

[0130] 062,将意图判别概率大于概率阈值的一个预设意图确定为当前轮的语音请求对应的目标意图以完成语音交互。

[0131] 请结合图12,交互模块16包括获取单元161和意图确定单元162。

[0132] 步骤061可以由获取单元161实现,步骤062可以由意图确定单元162实现。也即是说,获取单元161可以用于获取意图识别的结果对应各个预设意图的意图判别概率。意图确定单元162可以用于将意图判别概率大于概率阈值的一个预设意图确定为语音请求对应的目标意图以完成语音交互。

[0133] 根据多个类别的预设意图对应的每个预设意图类别的识别结果,意图识别模块15可以给出各个预设意图相匹配的意图判别概率,则可以得到多个意图判别概率。若概率阈值为0.9,则意图识别的结果为某个类别的预设意图的意图判别概率超过0.9,那么服务端认为该类别的预设意图就是当前用户的语音请求的目标意图。概率阈值也可以为其他数值,概率阈值可以为默认设置的数值,也可以根据用户需要自行设定,在此不作限制。

[0134] 如此,本发明的语音交互方法可以获取意图识别的结果对应各个预设意图的意图判别概率,将意图判别概率大于概率阈值的一个预设意图确定为语音请求对应的目标意图,从而实现识别用户精准调节车辆零部件的意图的需求。

[0135] 语音交互方法包括:

[0136] 063,在各个预设意图的意图判别概率均不大于概率阈值的情况下,确定当前轮的语音请求的意图为非刻度调节意图。

[0137] 步骤063可以由意图确定单元162实现。也即是说,意图确定单元162可以用于在各个预设意图的意图判别概率均不大于概率阈值的情况下,确定语音请求为非刻度调节意图。

[0138] 例如,当根据各个类别的预设意图对应得到的意图判别概率均不大于概率阈值的情况,即根据语音请求得到用户的意图识别结果与各个类别预设意图匹配的概率比较低,均低于概率阈值,例如概率阈值可以为0.9,则确定该语音请求为非刻度调节意图,非刻度调节意图指的是对不能用带有刻度的旋钮或按钮调节的零部件的用户意图,例如,用户输入的语音请求为“车门开开开”,因为车门不是用带有刻度的旋钮或按钮进行调节的零部件,因此,该语音请求“车门开开开”是非刻度调节意图。

[0139] 如此,在各个预设意图的意图判别概率均不大于概率阈值的情况下,确定该语音请求为非刻度调节意图,可以排除非刻度调节意图的语音请求。

[0140] 请参阅图13,本发明还提供一种服务器20。该服务器20包括处理器21和存储器22,存储器22上存储有计算机程序221,当计算机程序221被处理器21执行时,实现上述任意一个实施例中所述的语音交互方法。

[0141] 本发明的服务器可以通过处理器21执行计算机程序221,可以在接收到用户对于车辆预设功能的语音请求后,通过读取上一轮的语音请求,结合两轮语音请求查询是否命中缓存,在查询不到对应缓存的情况下,利用上一轮的语音请求改写当前轮的语音请求以使得改写后的语音请求可以被系统识别出相应的意图,从而对改写后的当前轮的语音请求进行意图识别后,可以准确识别出用户的意图,进而可以根据意图识别结果下发控制指令,以控制对应的车辆零部件,完成语音交互。使用高频缓存引擎与意图识别相结合的方式识别语音请求的意图,实现多轮语音请求下,准确识别出通过语音交互来模拟对车辆零部件的操作实现刻度调节的用户意图。

[0142] 请参阅图14,本发明还提供一种包含有计算机程序31的非易失性计算机可读存储介质30。当计算机程序31被一个或多个处理器40执行时,实现上述任意实施条例的语音交互方法。

[0143] 例如,计算机程序31被处理器40执行时实现以下数据处理方法的步骤:

[0144] 01,接收车辆转发的对车辆预设功能调节的当前轮的语音请求,预设功能指模拟对车辆零部件的操作进行刻度调节的功能;

[0145] 02,读取对车辆预设功能调节的上一轮的语音请求;

[0146] 03,根据当前轮的语音请求和上一轮的语音请求在缓存引擎中进行缓存查询;

[0147] 04,在缓存查询的结果为未能查询到对应缓存的情况下,利用上一轮的语音请求对当前轮的语音请求进行改写;

[0148] 05,对改写后的当前轮的语音请求进行意图识别;

[0149] 06,根据意图识别的结果完成语音交互。

[0150] 可以理解地,计算机程序包括计算机程序代码。计算机程序代码可以为源代码形式、对象代码形式、可执行文件或某些中间形式等。计算机可读存储介质可以包括:能够携带计算机程序代码的任何实体或装置、记录介质、U盘、移动硬盘、磁碟、光盘、计算机存储器、只读存储器(ROM,Read-Only Memory)、随机存取存储器(RAM,Random Access Memory)、以及软件分发介质等。

[0151] 本发明的计算机可读存储介质30存储的计算机程序31被处理器40执行时,可以在接收到用户对于车辆预设功能的语音请求后,通过读取上一轮的语音请求,结合两轮语音请求查询是否命中缓存,在查询不到对应缓存的情况下,利用上一轮的语音请求改写当前轮的语音请求以使得改写后的语音请求可以被系统识别出相应的意图,从而对改写后的当前轮的语音请求进行意图识别后,可以准确识别出用户的意图,进而可以根据意图识别结果下发控制指令,以控制对应的车辆零部件,完成语音交互。使用高频缓存引擎与意图识别相结合的方式识别语音请求的意图,实现多轮语音请求下,准确识别出通过语音交互来模拟对车辆零部件的操作实现刻度调节的用户意图。

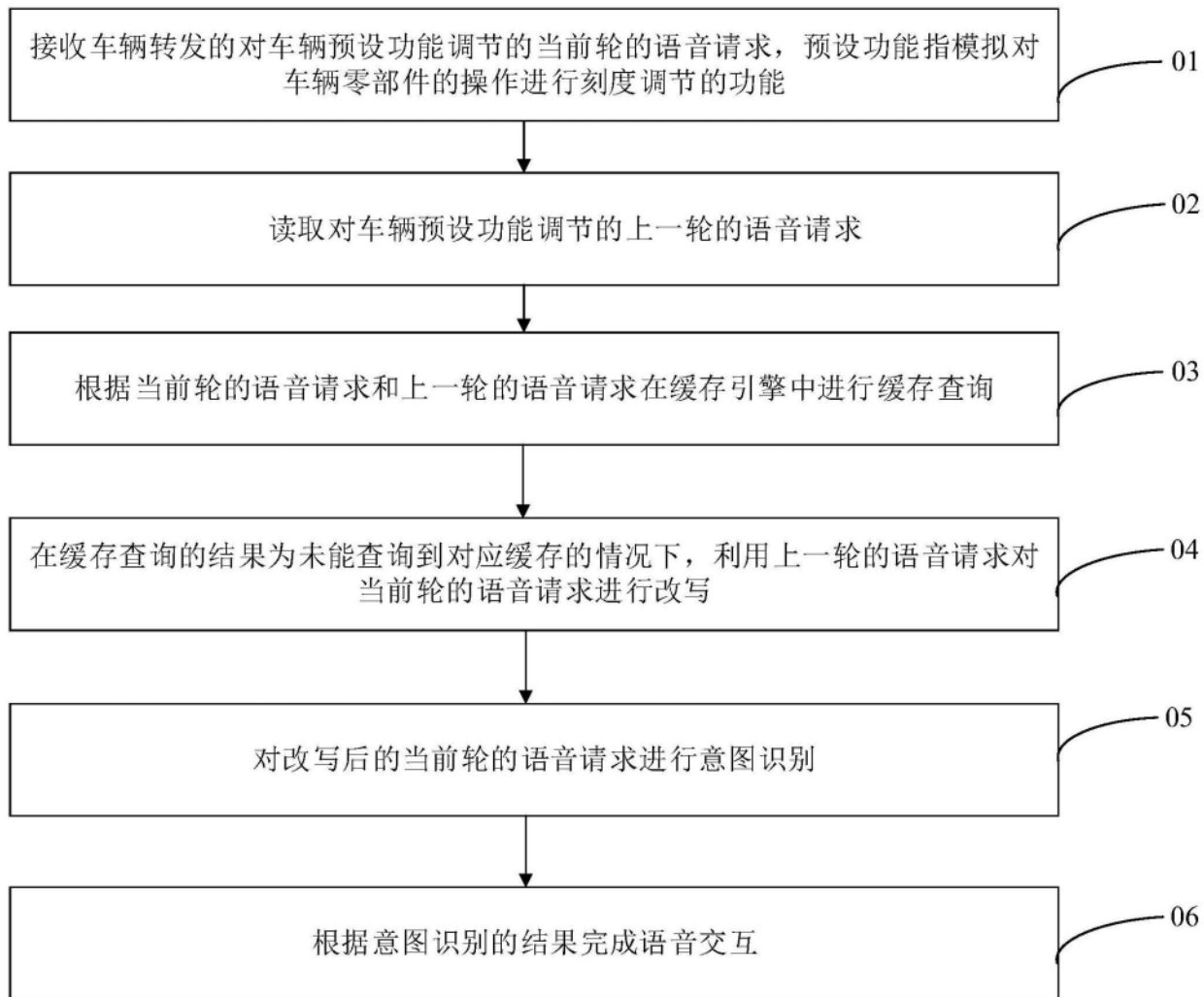


图1

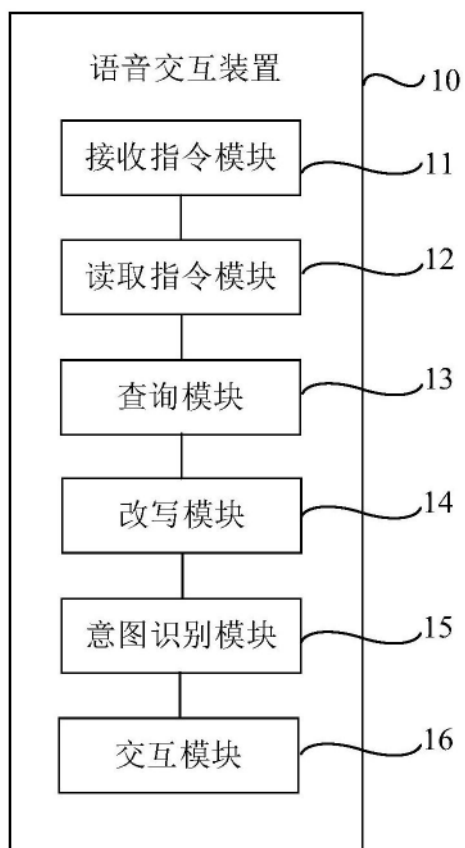


图2

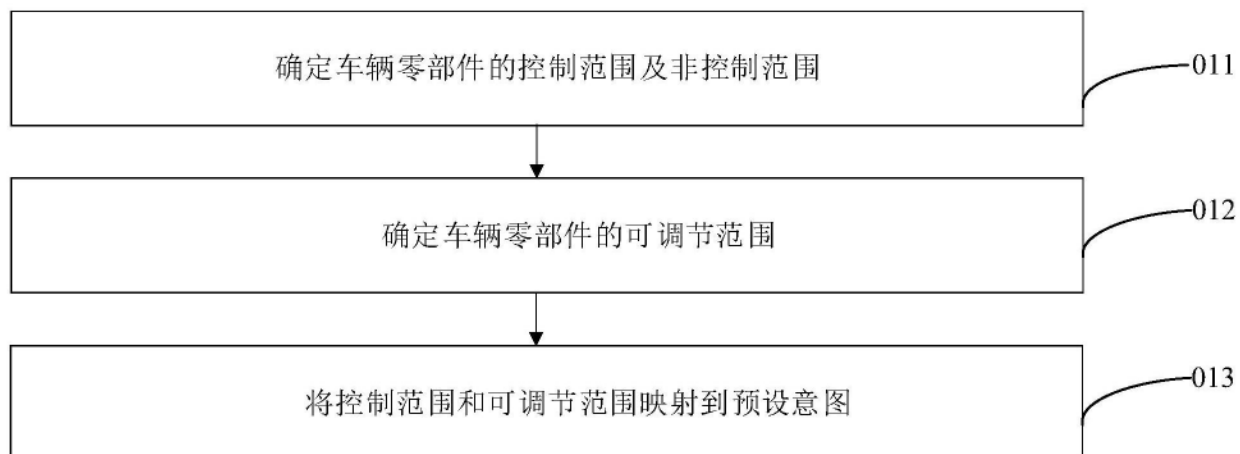


图3

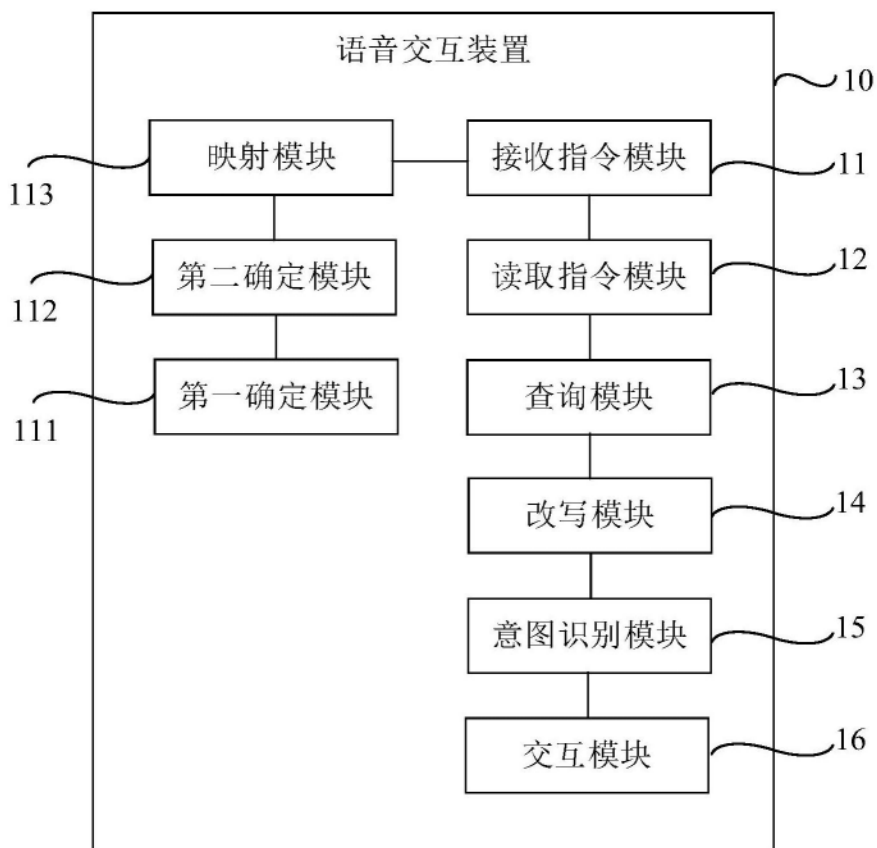


图4

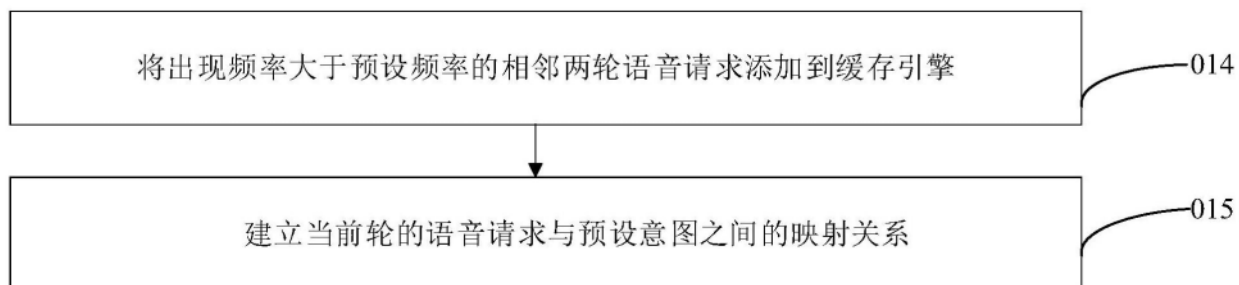


图5

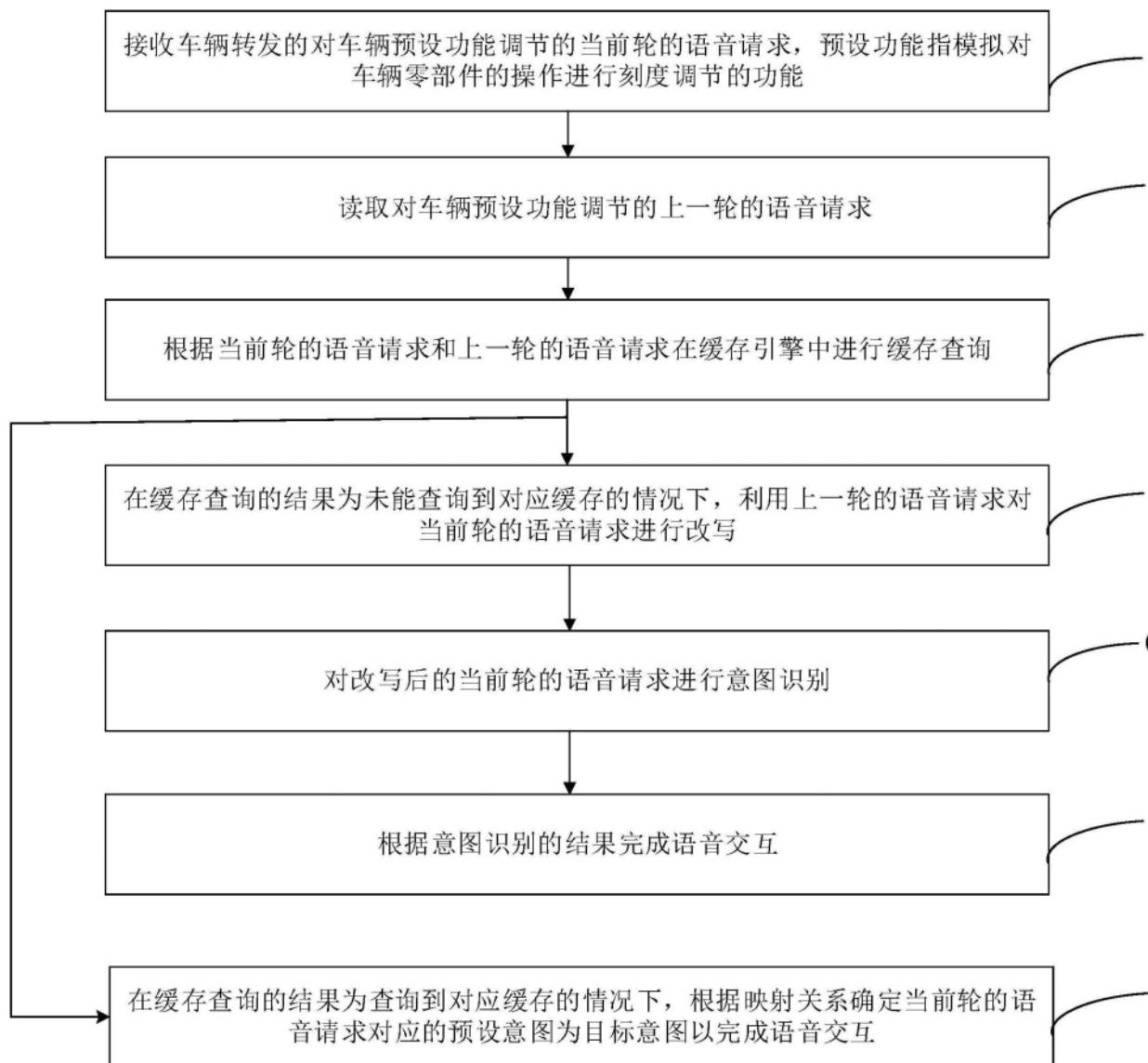


图6

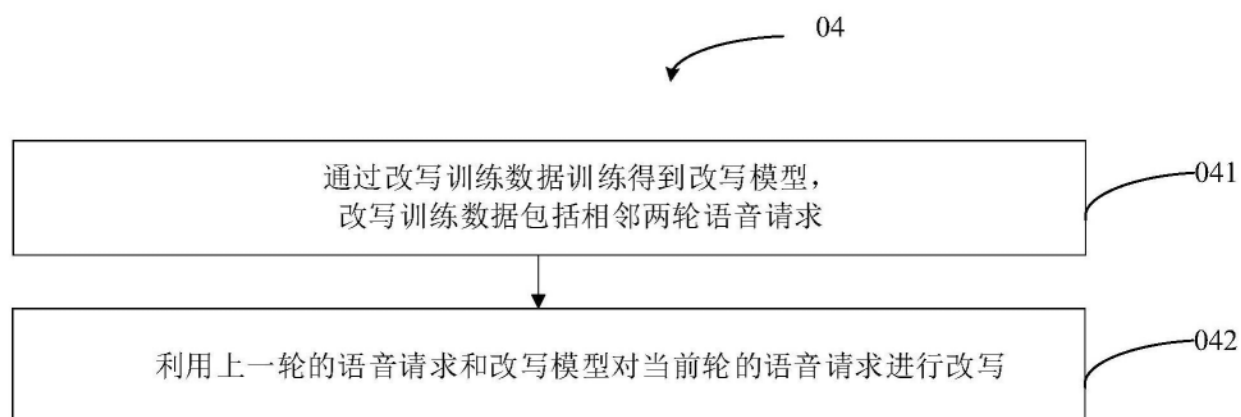


图7

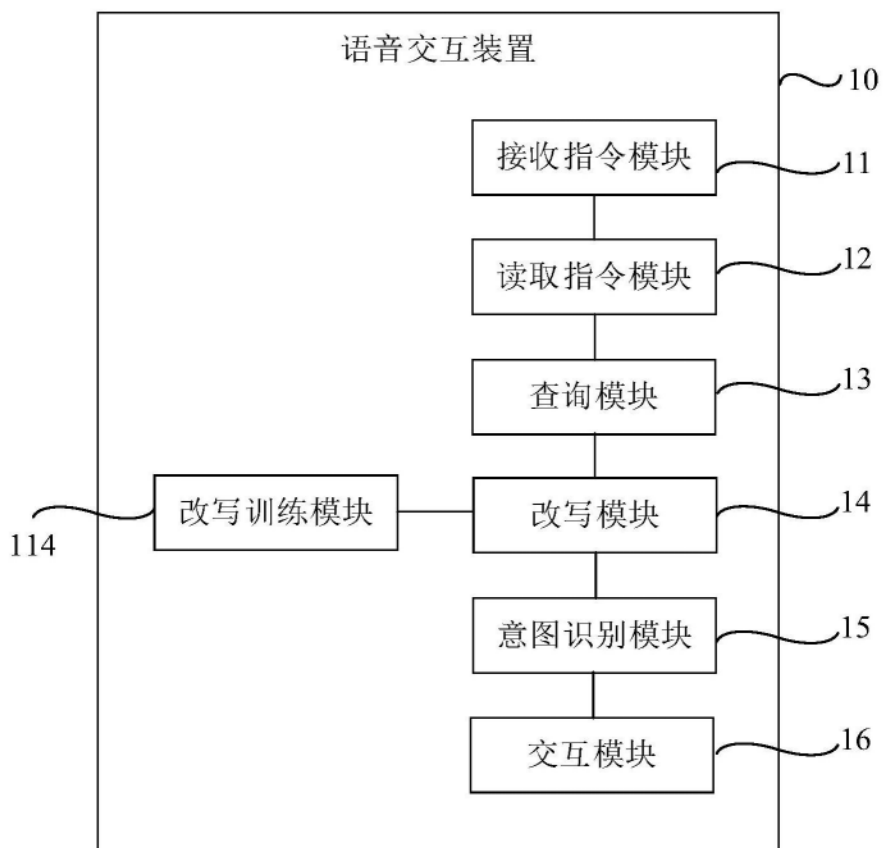


图8

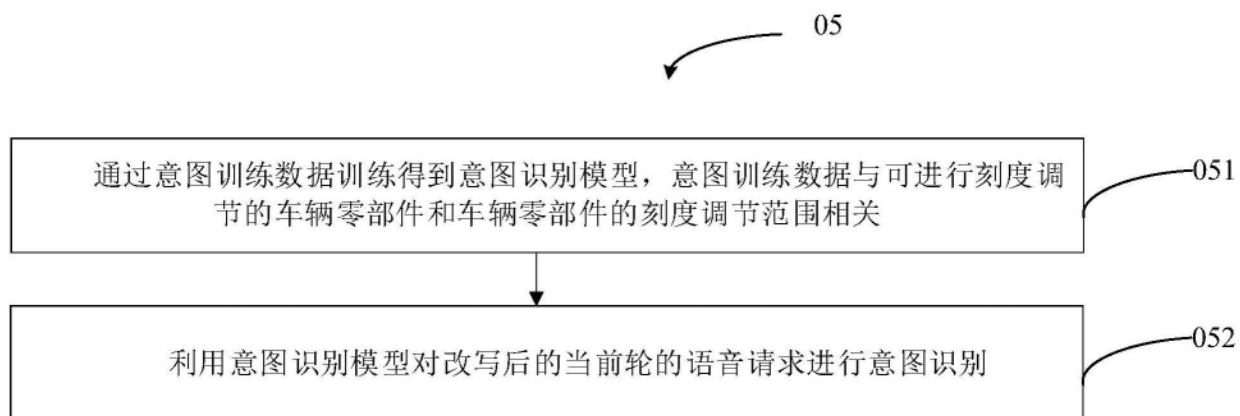


图9

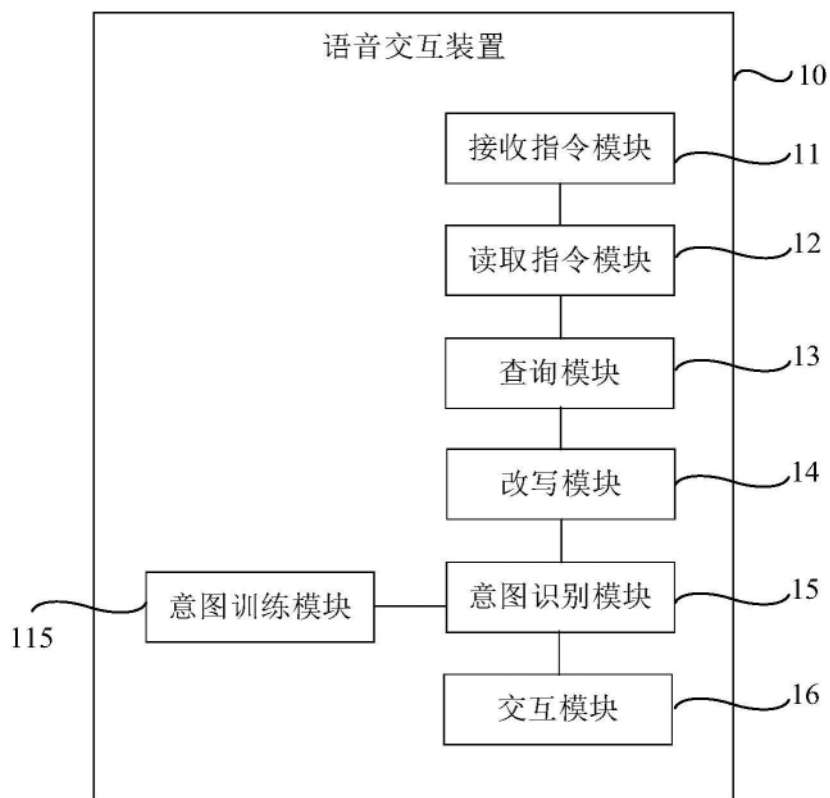


图10

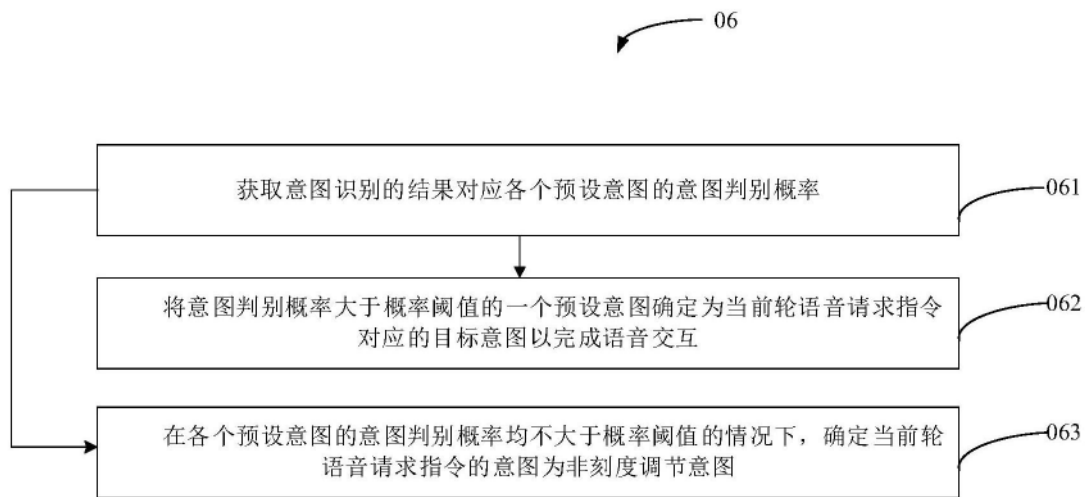


图11

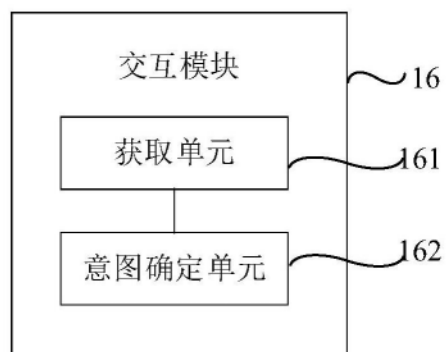


图12

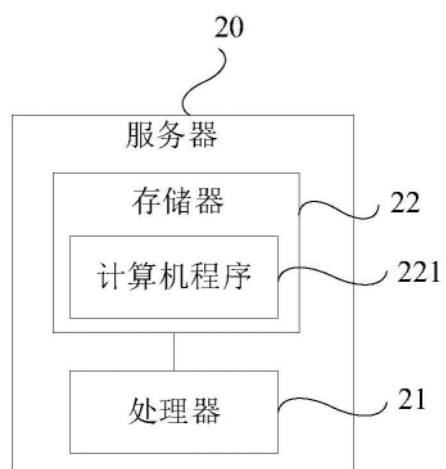


图13

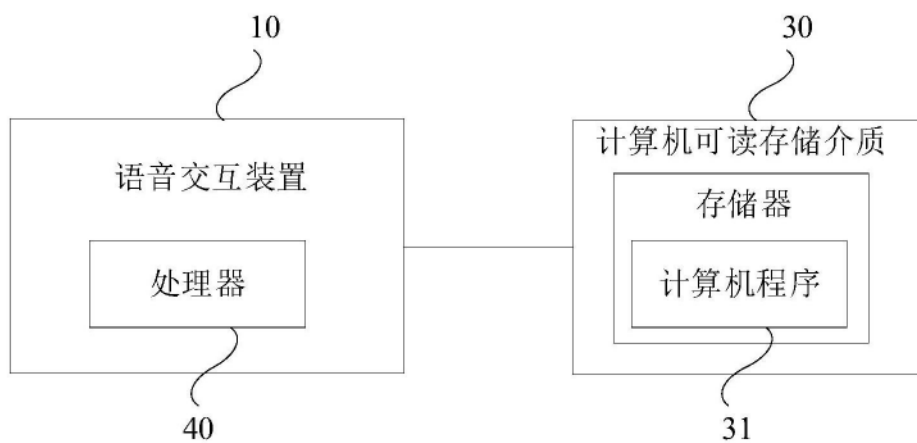


图14