



(12)发明专利申请

(10)申请公布号 CN 110890089 A

(43)申请公布日 2020.03.17

(21)申请号 201810942411.0

(22)申请日 2018.08.17

(71)申请人 珠海格力电器股份有限公司
地址 519070 广东省珠海市前山金鸡西路
六号

(72)发明人 秦萍 连园园 陈浩广 彭磊

(74)专利代理机构 北京康信知识产权代理有限
责任公司 11240
代理人 赵囡囡 曾红芳

(51) Int. Cl.

G10L 15/22(2006.01)

G10L 15/06(2013.01)

G10L 15/18(2013.01)

G10L 25/63(2013.01)

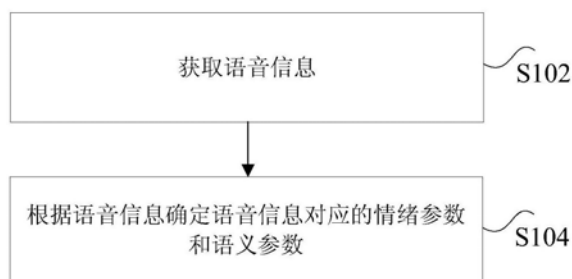
权利要求书2页 说明书6页 附图1页

(54)发明名称

语音识别方法及装置

(57)摘要

本发明公开了一种语音识别方法及装置。其中,该方法包括:获取语音信息;根据语音信息确定语音信息对应的情绪参数和语义参数。本发明解决了相关技术中由于车载人机语言交互方式单一造成的语音识别准确性低的技术问题。



1. 一种语音识别方法,其特征在于,包括:

获取语音信息;

根据所述语音信息确定所述语音信息对应的情绪参数和语义参数;

其中,根据所述语音信息确定所述语音信息对应的情绪参数时,根据语音信息输入情绪识别模型,由情绪识别模型输出所述语音信息对应的情绪参数,其中,所述情绪识别模型为使用多组数据通过机器学习训练得出的,所述多组数据中的每组数据均包括:语音信息和所述语音信息对应的情绪参数;

根据所述语音信息确定所述语音信息对应的语义参数时,根据语音信息输入语义识别模型,由语义识别模型输出所述语音信息对应的语义参数,其中所述语义识别模型为使用多组数据通过机器学习训练得出的,所述多组数据中的每组数据均包括:语音信息和所述语音信息对应的语义参数。

2. 根据权利要求1所述的方法,其特征在于,在根据所述语音信息确定所述语音信息对应的情绪参数和语义参数之后,还包括:

根据所述情绪参数和语义参数确定用于应答所述语音信息的应答信息;

根据所述应答信息进行对所述语音信息的应答。

3. 根据权利要求2所述的方法,其特征在于,根据所述应答信息进行对所述语音信息的应答包括:

将所述应答信息发送给语音播放器,由所述语音播放器对所述应答信息进行播放。

4. 根据权利要求3所述的方法,其特征在于,根据所述情绪参数和语义参数确定用于应答所述语音信息的应答信息包括:

根据所述情绪参数生成与所述情绪参数对应的情绪信息;

根据所述语义参数生成与所述语义参数对应的语音信息;

根据所述情绪信息和所述语音信息生成应答信息。

5. 根据权利要求1所述的方法,其特征在于,在根据所述语音信息确定所述语音信息对应的情绪参数和语义参数之后,还包括:

根据所述情绪参数和语义参数确定与所述语音信息对应的音乐参数;

根据所述音乐参数播放音乐。

6. 根据权利要求5所述的方法,其特征在于,根据所述音乐参数播放音乐包括:

将所述音乐参数发送至音乐播放器,由所述音乐播放器播放与音乐参数对应的音乐;

其中,所述音乐参数至少包括下列之一:音乐类型,音乐播放方式。

7. 根据权利要求1至6中任意一项所述的方法,其特征在于,获取语音信息包括:

向语音接收装置发送获取请求;

接收所述语音接收装置发送的语音信息。

8. 一种语音识别装置,其特征在于,包括:

获取模块,用于获取语音信息;

确定模块,用于根据所述语音信息确定所述语音信息对应的情绪参数和语义参数;

其中,根据所述语音信息确定所述语音信息对应的情绪参数时,根据语音信息输入情绪识别模型,由情绪识别模型输出所述语音信息对应的情绪参数,其中,所述情绪识别模型为使用多组数据通过机器学习训练得出的,所述多组数据中的每组数据均包括:语音信息

和所述语音信息对应的情绪参数；

根据所述语音信息确定所述语音信息对应的语义参数时，根据语音信息输入语义识别模型，由语义识别模型输出所述语音信息对应的语义参数，其中所述语义识别模型为使用多组数据通过机器学习训练得出的，所述多组数据中的每组数据均包括：语音信息和所述语音信息对应的语义参数。

9. 一种存储介质，其特征在于，所述存储介质存储有程序指令，其中，在所述程序指令运行时控制所述存储介质所在设备执行权利要求1至7中任意一项所述的方法。

10. 一种处理器，其特征在于，所述处理器用于运行程序，其中，所述程序运行时执行权利要求1至7中任意一项所述的方法。

语音识别方法及装置

技术领域

[0001] 本发明涉及汽车驾驶人机交互领域,具体而言,涉及一种语音识别方法及装置。

背景技术

[0002] 在汽车驾驶领域,驾驶者由于双手控制方向盘,双眼注视前方道路,故无法在开车时候进行道路规划、音乐选择、电话拨出等操作。但是在车中只有驾驶员一个人时,时常需要上述操作或相关信息的提示,或是与驾驶员进行简单对话以避免驾驶员长时间独自驾车造成的疲劳驾驶。而在现有技术中,车载人机语言交互方式单一,大多是通过固定的语音指令对汽车发出命令时,汽车才能识别,对于用户常用的自然语言,汽车交互系统多不能识别,更无法进行富含感情色彩的语言交流。

[0003] 针对上述的问题,目前尚未提出有效的解决方案。

发明内容

[0004] 本发明实施例提供了一种语音识别方法及装置,以至少解决相关技术中由于车载人机语言交互方式单一造成的语音识别准确性低的技术问题。

[0005] 根据本发明实施例的一个方面,提供了一种语音识别方法,包括:获取语音信息;根据所述语音信息确定所述语音信息对应的情绪参数和语义参数;其中,根据所述语音信息确定所述语音信息对应的情绪参数时,根据语音信息输入情绪识别模型,由情绪识别模型输出所述语音信息对应的情绪参数,其中,所述情绪识别模型为使用多组数据通过机器学习训练得出的,所述多组数据中的每组数据均包括:语音信息和所述语音信息对应的情绪参数;根据所述语音信息确定所述语音信息对应的语义参数时,根据语音信息输入语义识别模型,由语义识别模型输出所述语音信息对应的语义参数,其中所述语义识别模型为使用多组数据通过机器学习训练得出的,所述多组数据中的每组数据均包括:语音信息和所述语音信息对应的语义参数。

[0006] 可选的,在根据所述语音信息确定所述语音信息对应的情绪参数和语义参数之后,还包括:根据所述情绪参数和语义参数确定用于应答所述语音信息的应答信息;根据所述应答信息进行对所述语音信息的应答。

[0007] 可选的,根据所述应答信息进行对所述语音信息的应答包括:将所述应答信息发送给语音播放器,由所述语音播放器对所述应答信息进行播放。

[0008] 可选的,根据所述情绪参数和语义参数确定用于应答所述语音信息的应答信息包括:根据所述情绪参数生成与所述情绪参数对应的情绪信息;根据所述语义参数生成与所述语义参数对应的语音信息;根据所述情绪信息和所述语音信息生成应答信息。

[0009] 可选的,在根据所述语音信息确定所述语音信息对应的情绪参数和语义参数之后,还包括:根据所述情绪参数和语义参数确定与所述语音信息对应的音乐参数;根据所述音乐参数播放音乐。

[0010] 可选的,根据所述音乐参数播放音乐包括:将所述音乐参数发送至音乐播放器,由

所述音乐播放器播放与音乐参数对应的音乐;其中,所述音乐参数至少包括下列之一:音乐类型,音乐播放方式。

[0011] 可选的,获取语音信息包括:向语音接收装置发送获取请求;接收所述语音接收装置发送的语音信息。

[0012] 根据本发明实施例的另一方面,还提供了一种语音识别装置,包括:获取模块,用于获取语音信息;确定模块,用于根据所述语音信息确定所述语音信息对应的情绪参数和语义参数;其中,根据所述语音信息确定所述语音信息对应的情绪参数时,根据语音信息输入情绪识别模型,由情绪识别模型输出所述语音信息对应的情绪参数,其中,所述情绪识别模型为使用多组数据通过机器学习训练得出的,所述多组数据中的每组数据均包括:语音信息和所述语音信息对应的情绪参数;根据所述语音信息确定所述语音信息对应的语义参数时,根据语音信息输入语义识别模型,由语义识别模型输出所述语音信息对应的语义参数,其中所述语义识别模型为使用多组数据通过机器学习训练得出的,所述多组数据中的每组数据均包括:语音信息和所述语音信息对应的语义参数。

[0013] 根据本发明实施例的另一方面,还提供了一种存储介质,所述存储介质存储有程序指令,其中,在所述程序指令运行时控制所述存储介质所在设备执行上述任意一项所述的方法。

[0014] 根据本发明实施例的另一方面,还提供了一种处理器,所述处理器用于运行程序,其中,所述程序运行时执行上述任意一项所述的方法。

[0015] 在本发明实施例中,采用获取语音信息;根据所述语音信息确定所述语音信息对应的情绪参数和语义参数的方式,通过情绪识别模型和语义识别模型,达到了根据语音信息确定语音信息对应的情绪参数和语义参数的目的,从而实现了提高语音识别准确性的技术效果,进而解决了相关技术中由于车载人机语言交互方式单一造成的语音识别准确性低的技术问题。

附图说明

[0016] 此处所说明的附图用来提供对本发明的进一步理解,构成本申请的一部分,本发明的示意性实施例及其说明用于解释本发明,并不构成对本发明的不当限定。在附图中:

[0017] 图1是根据本发明实施例的一种语音识别方法的流程图;

[0018] 图2是根据本发明实施例的一种语音识别装置的结构示意图。

具体实施方式

[0019] 为了使本技术领域的人员更好地理解本发明方案,下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本发明一部分的实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都应当属于本发明保护的范围。

[0020] 需要说明的是,本发明的说明书和权利要求书及上述附图中的术语“第一”、“第二”等是用于区别类似的对象,而不必用于描述特定的顺序或先后次序。应该理解这样使用的数据在适当情况下可以互换,以便这里描述的本发明的实施例能够以除了在这里图示或

描述的那些以外的顺序实施。此外，术语“包括”和“具有”以及他们的任何变形，意图在于覆盖不排他的包含，例如，包含了一系列步骤或单元的过程、方法、系统、产品或设备不必限于清楚地列出的那些步骤或单元，而是可包括没有清楚地列出的或对于这些过程、方法、产品或设备固有的其它步骤或单元。

[0021] 根据本发明实施例，提供了一种语音识别方法的实施例，需要说明的是，在附图的流程图示出的步骤可以在诸如一组计算机可执行指令的计算机系统中执行，并且，虽然在流程图中示出了逻辑顺序，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。

[0022] 图1是根据本发明实施例的一种语音识别方法的流程图，如图1所示，该方法包括如下步骤：

[0023] 步骤S102，获取语音信息；

[0024] 步骤S104，根据语音信息确定语音信息对应的情绪参数和语义参数；其中，根据语音信息确定语音信息对应的情绪参数时，根据语音信息输入情绪识别模型，由情绪识别模型输出语音信息对应的情绪参数，其中，情绪识别模型为使用多组数据通过机器学习训练得出的，多组数据中的每组数据均包括：语音信息和语音信息对应的情绪参数；根据语音信息确定语音信息对应的语义参数时，根据语音信息输入语义识别模型，由语义识别模型输出语音信息对应的语义参数，其中语义识别模型为使用多组数据通过机器学习训练得出的，多组数据中的每组数据均包括：语音信息和语音信息对应的语义参数。

[0025] 在本发明实施例中，采用获取语音信息；根据语音信息确定语音信息对应的情绪参数和语义参数的方式，通过情绪识别模型和语义识别模型，达到了根据语音信息确定语音信息对应的情绪参数和语义参数的目的，从而实现了提高语音识别准确性的技术效果，进而解决了相关技术中由于车载人机语言交互方式单一造成的语音识别准确性低的技术问题。

[0026] 上述获取语音信息，通过语音采集装置获取语音信息，可以是汽车配置的采集装置，例如，车载麦克风、行车记录仪等，也可以是非汽车配置的采集装置，例如，手机、录音笔等。在获取语音信息后，需要对语音信息进行预处理，将一些对语音识别无用的语音信息滤除，例如，干扰噪声、重复表达等。通过预处理可以将语音信息中的关键特征保留，不仅提高了语音信息的精度，还对大大提升了语音识别的准确性。

[0027] 在根据语音信息确定语音信息对应的情绪参数时，根据语音信息输入情绪识别模型，由情绪识别模型输出语音信息对应的情绪参数，其中，情绪识别模型为使用多组数据通过机器学习训练得出的，多组数据中的每组数据均包括：语音信息和语音信息对应的情绪参数。需要说明的是，语音信息对应的情绪参数包括情绪种类和情绪等级，不同的情绪种类或者不同的情绪等级，以及情绪种类和情绪等级的不同组合，对应的情绪参数也不同。其中，情绪种类包括喜、怒、哀、乐等，情绪等级可以分为高、中、低，也可以根据用户的具体情形做出相应的设置，例如，根据用户生活习惯等将情绪划分等级。通过对情绪参数的细致划分，可以根据情绪种类和情绪等级的不同，进行组合，使得情绪参数更加精确，更能反映语音信息的真实内容，进而提高语音识别的准确性。

[0028] 在根据语音信息确定语音信息对应的语义参数时，根据语音信息输入语义识别模型，由语义识别模型输出语音信息对应的语义参数，其中语义识别模型为使用多组数据通

过机器学习训练得出的,多组数据中的每组数据均包括:语音信息和语音信息对应的语义参数。采用语义识别模型可以使汽车识别出以自然语言形式发出的语音信息,且该模型是对大量的语音信息进行机器学习得到的,相比现有技术基于统计概率的语音识别方法,该语义识别模型具备用户自适应能力,即新用户不需要进行全部语音词汇的训练,能够在使用中不断学习,以此提高识别效率。

[0029] 可选的,在根据语音信息确定语音信息对应的情绪参数和语义参数之后,还包括:根据情绪参数和语义参数确定用于应答语音信息的应答信息;根据应答信息进行对语音信息的应答。

[0030] 在用户与汽车进行语音交流中,获取的用户语音信息经过情绪识别模型和语义识别模型分别生成对应的情绪参数和语义参数,根据情绪参数和语义参数生成应答信息,对语音信息进行应答。例如,在汽车导航路线选择时,针对用户的需求,汽车可以推荐出几种方案,并将各自特点以语音形式告知用户,通过对用户回复的感情色彩辨识,自动选择用户最青睐的方案。如对用户进行路径推荐时,列出三种不同的行驶方案,并配以各自特点“优先选择高速”、“路径最短”、“绕过易堵点”等,如果用户回复“高速费太贵,绕路又会多行驶太多了,都不太满意”,则系统自动选定“路径最短”行驶方案为用户进行路径指引。

[0031] 可选的,根据应答信息进行对语音信息的应答包括:将应答信息发送给语音播放器,由语音播放器对应答信息进行播放。

[0032] 在对语音信息进行应答时,与语音信息对应的应答信息会被发送给语音播放器,并通过语音播放器进行播放。语音播放器可以是汽车扬声器、车载音响,也可以是与汽车连接的手机、蓝牙耳机等。例如,汽车在行驶过程中,用户需要启动汽车导航,此时用户发出请求导航的语音,从语音中提取语音信息,针对该语音信息生成对应的应答信息,该应答信息会通过汽车扬声器播放,以告知用户需要的导航信息。除此以外,由于能够实现人机交互,在用户长期驾驶过程中,为了减少用户的长时间乏味的驾驶操作,可以通过该语音播放器提醒用户安全驾驶或者与用户语音互动,消除驾驶的疲劳感。

[0033] 可选的,根据情绪参数和语义参数确定用于应答语音信息的应答信息包括:根据情绪参数生成与情绪参数对应的情绪信息;根据语义参数生成与语义参数对应的语音信息;根据情绪信息和语音信息生成应答信息。

[0034] 在本发明实施例中,情绪参数生成对应的情绪信息,语义参数生成对应的语音信息,根据情绪信息和语音信息生成应答信息。传统的语音识别只能对语义进行识别,并未考虑人的情感,识别结果缺乏感情色彩和人文关怀。相比之下,本发明将语音中提取的语音信息通过不同的识别模型得到不同的识别结果。例如,通过情绪识别模型识别语音信息得到情绪参数,由情绪参数生成对应的情绪信息,以及通过语义识别模型识别语音信息得到语义参数,由语义参数生成对应的语音信息,将情绪信息与语音信息结合生成应答信息。因此,应答信息是对语音的语义以及感情色彩辨识的结果,更能够反映用户的真实的想法,使得车载人机交互更加智能化。

[0035] 可选的,在根据语音信息确定语音信息对应的情绪参数和语义参数之后,还包括:根据情绪参数和语义参数确定与语音信息对应的音乐参数;根据音乐参数播放音乐。

[0036] 在本发明实施例中,可以根据情绪参数和语义参数确定与语音信息对应的音乐参数,进而播放与该语音参数对应的音乐。例如,汽车在与用户进行沟通对话时,不仅通过用

户语音中的感情色彩自动匹配适宜的回 复,还可以根据用户语音中的感情色彩推荐适宜的音乐,该音乐与获取用户当前的语音信息紧密相关,进而通过提高驾驶过程中的便利性和趣味性,提高用户的驾驶体验。

[0037] 可选的,根据音乐参数播放音乐包括:将音乐参数发送至音乐播放器,由音乐播放器播放与音乐参数对应的音乐;其中,音乐参数至少包括下列之一:音乐类型,音乐播放方式。

[0038] 需要说明的是,音乐播放器在接收到音乐参数后,就会播放与音乐参数对应的音乐,其中,音乐参数包括音乐类型、音乐播放方式、音乐播放时间等。具体的,音乐类型包括轻音乐、伤感、情歌、网络歌曲等,音乐播放方式包括随机播放、顺序播放、单曲循环等。除此以外,还可以根据用户的听歌的频率、时长、歌手等,优先选取播放常听音乐或者推荐类似风格的音乐。

[0039] 可选的,获取语音信息包括:向语音接收装置发送获取请求;接收语音接收装置发送的语音信息。

[0040] 在获取语音信息时,需要向语音接收装置发送获取请求,该获取请求可以是用户主动发起,例如,用户可以通过物理按压或者触摸语音接收装置、语音命令,也可以通过其他能够开启语音接收装置的方法;获取请求也可以是安装在汽车上的感应装置得到,例如,感应装置检测到用户在车上的当前状态,该当前状态为用户的行为动作,包括用户打开或关闭车门、用户驾驶的换挡等。上述方法都可以触发相语音接收装置发送获取请求的命令。当然,用户可以根据自己的习惯进行设置。语音接收装置接收到获取请求,对用户语音进行采集,将采集的语音信息输入相应的识别模型进行处理。语音接收装置能够通过多种方式、灵活的接收获取语音的请求,而且语音接收装置内置的存储装置可以将采集的语音信息存储起来,可以根据需要将采集的语音信息进一步处理与应用。

[0041] 图2是根据本发明实施例的一种语音识别装置的结构示意图;如图2所示,该语音识别装置20,包括:获取模块22,确定模块24。下面对该语音识别装置20进行详细说明。

[0042] 获取模块22,用于获取语音信息;确定模块24,与上述获取模块22连接,用于根据语音信息确定语音信息对应的情绪参数和语义参数;其中,根据语音信息确定语音信息对应的情绪参数时,根据语音信息输入情绪识别模型,由情绪识别模型输出语音信息对应的情绪参数,其中,情绪识别模型为使用多组数据通过机器学习训练得出的,多组数据中的每组数据均包括:语音信息和语音信息对应的情绪参数;根据语音信息确定语音信息对应的语义参数时,根据语音信息输入语义识别模型,由语义识别模型输出语音信息对应的语义参数,其中语义识别模型为使用多组数据通过机器学习训练得出的,多组数据中的每组数据均包括:语音信息和语音信息对应的语义参数。

[0043] 根据本发明实施例的另一方面,还提供了一种存储介质,存储介质存储有程序指令,其中,在程序指令运行时控制存储介质所在设备执行上述任意一项的方法。

[0044] 根据本发明实施例的另一方面,还提供了一种处理器,处理器用于运行程序,其中,程序运行时执行上述任意一项的方法。

[0045] 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。

[0046] 在本发明的上述实施例中,对各个实施例的描述都各有侧重,某个实施例中并没有详述的部分,可以参见其他实施例的相关描述。

[0047] 在本申请所提供的几个实施例中,应该理解到,所揭露的技术内容,可通过其它的方式实现。其中,以上所描述的装置实施例仅仅是示意性的,例如单元的划分,可以为一种逻辑功能划分,实际实现时可以有另外的划分方式,例如多个单元或组件可以结合或者可以集成到另一个系统,或一些特征可以忽略,或不执行。另一点,所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口,单元或模块的间接耦合或通信连接,可以是电性或其它的形式。

[0048] 作为分离部件说明的单元可以是或者也可以不是物理上分开的,作为单元显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者也可以分布到多个单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

[0049] 另外,在本发明各个实施例中的各功能单元可以集成在一个处理单元中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现,也可以采用软件功能单元的形式实现。

[0050] 集成的单元如果以软件功能单元的形式实现并作为独立的产品销售或使用,可以存储在一个计算机可读取存储介质中。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或者该技术方案的全部或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一个存储介质中,包括若干指令用以使得一台计算机设备(可为个人计算机、服务器或者网络设备)执行本发明各个实施例方法的全部或部分步骤。而前述的存储介质包括:U盘、只读存储器(ROM,Read-Only Memory)、随机存取存储器(RAM,Random Access Memory)、移动硬盘、磁碟或者光盘等各种可以存储程序代码的介质。

[0051] 以上仅是本发明的优选实施方式,应当指出,对于本技术领域的普通技术人员来说,在不脱离本发明原理的前提下,还可以做出若干改进和润饰,这些改进和润饰也应视为本发明的保护范围。

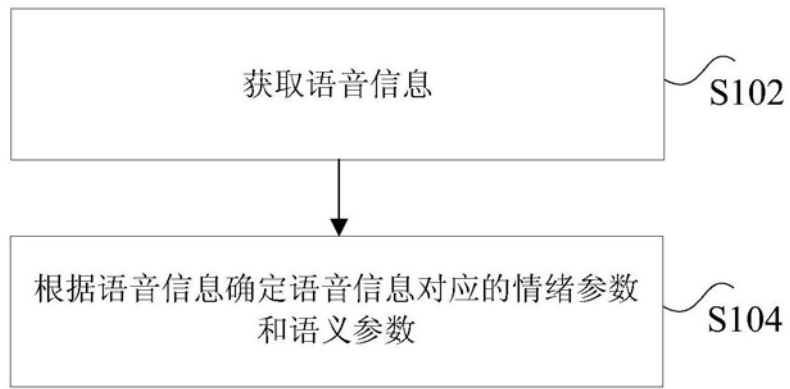


图1

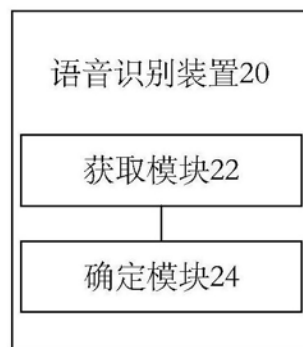


图2