



(12)发明专利申请

(10)申请公布号 CN 110265021 A

(43)申请公布日 2019.09.20

(21)申请号 201910665234.0

(22)申请日 2019.07.22

(71)申请人 深圳前海微众银行股份有限公司
地址 518000 广东省深圳市前海深港合作区前湾一路1号A栋201室(入驻深圳市前海商务秘书有限公司)

(72)发明人 蔡杭 于夕畔 周楠楠

(74)专利代理机构 深圳市世纪恒程知识产权代理事务所 44287

代理人 王韬

(51)Int.Cl.

G10L 15/22(2006.01)

G10L 15/02(2006.01)

G10L 15/18(2013.01)

G10L 25/12(2013.01)

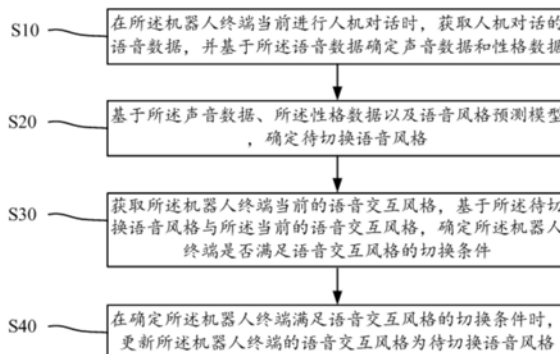
权利要求书2页 说明书10页 附图2页

(54)发明名称

个性化语音交互方法、机器人终端、装置及可读存储介质

(57)摘要

本发明公开了一种个性化语音交互方法,包括以下步骤:在机器人终端当前进行人机对话时,获取人机对话的语音数据,并基于语音数据确定声音数据和性格数据,基于声音数据、性格数据以及语音风格预测模型,确定待切换语音风格,获取所述机器人终端当前的语音交互风格,基于所述待切换语音风格与所述当前的语音交互风格,确定所述机器人终端是否满足语音交互风格的切换条件,在确定机器人终端满足语音交互风格的切换条件时,更新机器人终端的语音交互风格为待切换语音风格。本发明还公开了一种装置、机器人终端及可读存储介质。实现了机器人终端自适应对话人物风格,从而带来更加符合用户对话习惯的语音交互体验。



1. 一种个性化语音交互方法,应用于机器人终端,其特征在于,所述的个性化语音交互方法包括以下步骤:

在所述机器人终端当前进行人机对话时,获取人机对话的语音数据,并基于所述语音数据确定声音数据和性格数据;

基于所述声音数据、所述性格数据以及语音风格预测模型,确定待切换语音风格;

获取所述机器人终端当前的语音交互风格,基于所述待切换语音风格与所述当前的语音交互风格,确定所述机器人终端是否满足语音交互风格的切换条件;

在确定所述机器人终端满足语音交互风格的切换条件时,更新所述机器人终端的语音交互风格为待切换语音风格。

2. 如权利要求1所述的个性化语音交互方法,其特征在于,所述在确定所述机器人终端满足语音交互风格的切换条件时,更新所述机器人终端的语音交互风格为待切换语音风格的步骤之后,还包括:

获取人机对话的语音数据以及人机对话的视频信息;

基于所述语音数据以及所述视频信息确定所述机器人终端是否满足语音风格保持条件;

在确定所述机器人终端不满足语音风格保持条件时,更新所述机器人终端的语音交互风格为默认风格。

3. 如权利要求1所述的个性化语音交互方法,其特征在于,所述在确定所述机器人终端满足语音交互风格的切换条件时,更新所述机器人终端的语音交互风格为待切换语音风格的步骤之后,还包括:

获取人机对话的语音数据;

确定所述语音数据中是否存在切换语音风格的信息;

在所述语音数据中存在切换语音风格的信息时,基于所述切换语音风格的信息执行语音切换操作。

4. 如权利要求1所述的个性化语音交互方法,其特征在于,所述在所述机器人终端当前进行人机对话时,获取人机对话的语音数据,并基于所述语音数据确定声音数据和性格数据的步骤之前,还包括:

在检测到所述机器人终端启动人机对话时,基于摄像头获取人脸信息;

确定所述人脸信息是否满足预设条件;

在所述人脸信息满足预设条件时,基于人物与语音风格匹配库获取所述人脸信息对应的个性化语音风格;

更新所述机器人终端的语音交互风格为所述个性化语音风格。

5. 如权利要求4所述的个性化语音交互方法,其特征在于,所述确定所述人脸信息是否满足预设条件的步骤之后,还包括:

在所述人脸信息不满足预设条件时,更新所述机器人终端的语音交互风格为默认风格。

6. 如权利要求4所述的个性化语音交互方法,其特征在于,所述在检测到所述机器人终端启动人机对话时,基于摄像头获取人脸信息的步骤之前,还包括:

基于预设数量的人物信息样本以及所述人物信息样本对应的语音风格生成所述人物

与语音风格匹配库,其中,所述人物信息样本包括人脸信息、声音数据以及性格数据。

7.如权利要求6所述的个性化语音交互方法,其特征在于,所述在检测到所述机器人终端启动人机对话时,基于摄像头获取人脸信息的步骤之前,还包括:

基于所述人物与语音风格匹配库,定时获取人物信息样本以及所述人物信息样本对应的语音风格;

基于所述人物信息样本以及所述人物信息样本对应的语音风格训练初始语音风格预测模型,得到所述语音风格预测模型。

8.如权利要求7所述的个性化语音交互方法,其特征在于,所述确定所述人脸信息是否满足预设条件的步骤包括:

确定所述人物与语音风格匹配库是否存在所述人脸信息,其中,在所述人物与语音风格匹配库存在所述人脸信息时,确定所述人脸信息满足预设条件。

9.如权利要求1至8中任一项所述的个性化语音交互方法,其特征在于,所述在确定所述机器人终端满足语音交互风格的切换条件时,更新所述机器人终端的语音交互风格为待切换语音风格的步骤之后,还包括:

在检测到人机对话结束指令时,获取所述人机对话结束指令对应的目标语音交互风格以及所述人机对话结束指令对应的目标人物信息,其中,所述目标人物信息包括:人脸信息、声音数据以及性格数据;

关联保存所述目标语音交互风格和所述目标人物信息至所述人物与语音风格匹配库。

10.一种个性化语音交互装置,其特征在于,所述个性化语音交互装置包括:

获取模块,在所述机器人终端当前进行人机对话时,获取人机对话的语音数据,并基于所述语音数据确定声音数据和性格数据;

确定模块,基于所述声音数据、所述性格数据以及语音风格预测模型,确定待切换语音风格;

判断模块,获取所述机器人终端当前的语音交互风格,基于所述待切换语音风格与所述当前的语音交互风格,确定所述机器人终端是否满足语音交互风格的切换条件;

更新模块,在确定所述机器人终端满足语音交互风格的切换条件时,更新所述机器人终端的语音交互风格为待切换语音风格。

11.一种机器人终端,其特征在于,所述机器人终端包括: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个性化语音交互程序,所述个性化语音交互程序被所述处理器执行时实现如权利要求1至9中任一项所述的个性化语音交互方法的步骤。

12.一种可读存储介质,其特征在于,所述可读存储介质上存储有所述个性化语音交互程序,所述个性化语音交互程序被处理器执行时实现如权利要求1至9中任一项所述的个性化语音交互方法的步骤。

个性化语音交互方法、机器人终端、装置及可读存储介质

技术领域

[0001] 本发明涉及机器学习技术领域,尤其涉及一种个性化语音交互方法、机器人终端、装置及可读存储介质。

背景技术

[0002] 近年来,随着语音识别技术在机器人终端控制中的应用,机器人终端的应用领域不断扩大,广泛应用于工业、家电、通信、汽车电子、医疗、家庭服务和消费电子类产品等领域。智能机器人终端语音交互是基于语音输入的新一代交互模式,通过说话就可以得到反馈结果。

[0003] 目前,语音交互对话形式,尽管通过人工预先编辑应答格式,改善音色等,从表面上使得对话更接近人类对话,具备一定亲和感。但是智能机器人终端与人对话时都是一种口吻,一种音调,一种风格,因此,导致了与智能机器人终端对话缺乏个性化,显得无聊。

[0004] 上述内容仅用于辅助理解本发明的技术方案,并不代表承认上述内容是现有技术。

发明内容

[0005] 本发明的主要目的在于提供一种个性化语音交互方法、机器人终端、装置及可读存储介质,旨在解决现有智能机器人终端与人对话时是一种口吻,一种音调,一种风格,无法自适应对话人物风格的技术问题。

[0006] 为实现上述目的,本发明提供一种个性化语音交互方法,所述的个性化语音交互方法包括以下步骤:

[0007] 在所述机器人终端当前进行人机对话时,获取人机对话的语音数据,并基于所述语音数据确定声音数据和性格数据;

[0008] 基于所述声音数据、所述性格数据以及语音风格预测模型,确定待切换语音风格;

[0009] 获取所述机器人终端当前的语音交互风格,基于所述待切换语音风格与所述当前的语音交互风格,确定所述机器人终端是否满足语音交互风格的切换条件;

[0010] 在确定所述机器人终端满足语音交互风格的切换条件时,更新所述机器人终端的语音交互风格为待切换语音风格。

[0011] 进一步地,在一实施方式中,所述在确定所述机器人终端满足语音交互风格的切换条件时,更新所述机器人终端的语音交互风格为待切换语音风格的步骤之后,还包括:

[0012] 获取人机对话的语音数据以及人机对话的视频信息;

[0013] 基于所述语音数据以及所述视频信息确定所述机器人终端是否满足语音风格保持条件;

[0014] 在确定所述机器人终端不满足语音风格保持条件时,更新所述机器人终端的语音交互风格为默认风格。

[0015] 进一步地,在一实施方式中,所述在确定所述机器人终端满足语音交互风格的切

换条件时,更新所述机器人终端的语音交互风格为待切换语音风格的步骤之后,还包括:

[0016] 获取人机对话的语音数据;

[0017] 确定所述语音数据中是否存在切换语音风格的信息;

[0018] 在所述语音数据中存在切换语音风格的信息时,基于所述切换语音风格的信息执行语音切换操作。

[0019] 进一步地,在一实施方式中,所述在所述机器人终端当前进行人机对话时,获取人机对话的语音数据,并基于所述语音数据确定声音数据和性格数据的步骤之前,还包括:

[0020] 在检测到所述机器人终端启动人机对话时,基于摄像头获取人脸信息;

[0021] 确定所述人脸信息是否满足预设条件;

[0022] 在所述人脸信息满足预设条件时,基于人物与语音风格匹配库获取所述人脸信息对应的个性化语音风格;

[0023] 更新所述机器人终端的语音交互风格为所述个性化语音风格。

[0024] 进一步地,在一实施方式中,所述确定所述人脸信息是否满足预设条件的步骤之后,还包括:

[0025] 在所述人脸信息不满足预设条件时,更新所述机器人终端的语音交互风格为默认风格。

[0026] 进一步地,在一实施方式中,所述在检测到所述机器人终端启动人机对话时,基于摄像头获取人脸信息的步骤之前,还包括:

[0027] 基于预设数量的人物信息样本以及所述人物信息样本对应的语音风格生成所述人物与语音风格匹配库,其中,所述人物信息样本包括人脸信息、声音数据以及性格数据。

[0028] 进一步地,在一实施方式中,所述在检测到所述机器人终端启动人机对话时,基于摄像头获取人脸信息的步骤之前,还包括:

[0029] 基于所述人物与语音风格匹配库,定时获取人物信息样本以及所述人物信息样本对应的语音风格;

[0030] 基于所述人物信息样本以及所述人物信息样本对应的语音风格训练初始语音风格预测模型,得到所述语音风格预测模型。

[0031] 进一步地,在一实施方式中,所述确定所述人脸信息是否满足预设条件的步骤包括:

[0032] 确定所述人物与语音风格匹配库是否存在所述人脸信息,其中,在所述人物与语音风格匹配库存在所述人脸信息时,确定所述人脸信息满足预设条件。

[0033] 进一步地,在一实施方式中,所述在确定所述机器人终端满足语音交互风格的切换条件时,更新所述机器人终端的语音交互风格为待切换语音风格的步骤之后,还包括:

[0034] 在检测到人机对话结束指令时,获取所述人机对话结束指令对应的目标语音交互风格以及所述人机对话结束指令对应的目标人物信息,其中,所述目标人物信息包括:人脸信息、声音数据以及性格数据;

[0035] 关联保存所述目标语音交互风格和所述目标人物信息至所述人物与语音风格匹配库。

[0036] 进一步地,在一实施方式中,所述个性化语音交互装置包括:

[0037] 获取模块,在所述机器人终端当前进行人机对话时,获取人机对话的语音数据,并

基于所述语音数据确定声音数据和性格数据；

[0038] 确定模块,基于所述声音数据、所述性格数据以及语音风格预测模型,确定待切换语音风格；

[0039] 判断模块,获取所述机器人终端当前的语音交互风格,基于所述待切换语音风格与所述当前的语音交互风格,确定所述机器人终端是否满足语音交互风格的切换条件；

[0040] 更新模块,在确定所述机器人终端满足语音交互风格的切换条件时,更新所述机器人终端的语音交互风格为待切换语音风格。

[0041] 此外,为实现上述目的,本发明还提供一种机器人终端,所述机器人终端包括: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个性化语音交互程序,所述个性化语音交互程序被所述处理器执行时实现上述任一项所述的个性化语音交互方法的步骤。

[0042] 此外,为实现上述目的,本发明还提供一种可读存储介质,所述可读存储介质上存储有个性化语音交互程序,所述个性化语音交互程序被处理器执行时实现上述任一项所述的个性化语音交互方法的步骤。

[0043] 本发明在所述机器人终端当前进行人机对话时,获取人机对话的语音数据,并基于所述语音数据确定声音数据和性格数据,而后基于所述声音数据、所述性格数据以及语音风格预测模型,确定待切换语音风格,然后获取所述机器人终端当前的语音交互风格,基于所述待切换语音风格与所述当前的语音交互风格,确定所述机器人终端是否满足语音交互风格的切换条件,最后在确定所述机器人终端满足语音交互风格的切换条件时,更新所述机器人终端的语音交互风格为待切换语音风格。通过人机对话的语音数据和语音风格预测模型确定机器人终端的语音风格,实现了机器人终端自适应对话人物风格,从而带来更加符合用户对话习惯的语音交互体验。

附图说明

[0044] 图1是本发明实施例方案涉及的硬件运行环境中机器人终端的结构示意图；

[0045] 图2为本发明个性化语音交互方法第一实施例的流程示意图；

[0046] 图3为本发明个性化语音交互方法第二实施例的流程示意图；

[0047] 图4为本发明个性化语音交互装置实施例的功能模块示意图。

[0048] 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

[0049] 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。

[0050] 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中机器人终端的结构示意图。

[0051] 如图1所示,该机器人终端可以包括:处理器1001,例如CPU,网络接口1004,客户接口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。客户接口1003可以包括显示屏(Display)、输入单元比如键盘(Keyboard),可选客户接口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。

[0052] 可选地,机器人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等,在此不再赘述。

[0053] 本领域技术人员可以理解,图1中示出的系统结构并不构成对机器人终端的限定,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。

[0054] 如图1所示,作为一种可读存储介质的存储器1005中可以包括操作系统、网络通信模块、客户接口模块以及个性化语音交互程序。

[0055] 在图1所示的系统中,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;客户接口1003主要用于连接客户端(客户端),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中存储的个性化语音交互程序。

[0056] 在本实施例中,机器人终端包括:存储器1005、处理器1001及存储在所述存储器1005上并可在所述处理器1001上运行的个性化语音交互程序,其中,处理器1001调用存储器1005中存储的个性化语音交互程序时,执行本申请各个实施例提供的个性化语音交互方法的步骤。

[0057] 本发明还提供一种个性化语音交互方法,参照图2,图2为本发明个性化语音交互方法第一实施例的流程示意图。

[0058] 本发明实施例提供了个性化语音交互方法的实施例,需要说明的是,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。

[0059] 在本实施例中,该个性化语音交互方法包括:

[0060] 步骤S10,在所述机器人终端当前进行人机对话时,获取人机对话的语音数据,并基于所述语音数据确定声音数据和性格数据;

[0061] 在本实施例中,语音作为人类所特有的能力,是人类之间交流及获取外界信息资源的重要的工具和渠道,对于人类文明的发展具有重要的意义。语音识别技术作为人机交互分支的重要组成,是人机交互的重要接口,对于人工智能的发展具有重要的实际意义。语音识别技术经过数十年的发展,已经取得了显著的进步,逐步开始从实验室慢慢走向市场。目前,针对特定说话人的语音识别系统已经有较高的识别精度。

[0062] 目前,人机对话进入第三代,人机交流的内容主要是人习惯的自然交流语言,但是,机器人终端与人对话时过程中通常都是一种口吻和风格,无法满足用户对话习惯的语音交互风格。有鉴于此,本发明通过人机对话的语音数据和语音风格预测模型确定机器人终端的语音风格,实现了机器人终端自适应对话人物风格,从而带来更加符合用户对话习惯的语音交互体验。

[0063] 具体地,在机器人终端与用户进行人机对话时,机器人终端可以实时获取人机对话的语音数据,根据语音数据,利用现有的语音识别技术可以确定该用户的声音数据和性格数据,其中,声音数据包括在语言学中被称为语音“四要素”的音高、音强、音长和音色,不同用户的语声是不同的,因为人的发声器官实际上存在着大小、形态及功能上的差异。发声控制器官包括声带、软腭、舌头、牙齿、唇等;发声共鸣器包括咽腔、口腔、鼻腔。这些器官的

微小差异都会导致发声气流的改变,造成音质、音色的差别。此外,人发声的习惯亦有快有慢,用力有大有小,也造成音强、音长的差别;性格数据是用户的性格特点,根据语音数据的对话内容,利用NLP (Natural Language Processing,自然语言处理) 技术对对话内容进行解析,从而分析出用户的性格,用户的性格可包括:活跃型、稳重型、幽默型、可爱型、较真型等。可选地,在人机对话时,机器人终端还可以通过提问的方式询问用户的性格,以及用户喜欢的语音风格,在进行对话内容解析时,就可以直接解析出用户的性格数据。

[0064] 步骤S20,基于所述声音数据、所述性格数据以及语音风格预测模型,确定待切换语音风格;

[0065] 在本实施例中,预测模型是指用于预测的,用数学语言或公式所描述的事物间的数量关系。它在一定程度上揭示了事物间的内在规律性,预测时把它作为计算预测值的直接依据。语音风格预测模型可以根据输入的用户的声音数据和性格数据得出该用户喜欢的语音风格,其中,语音风格预测模型是现有技术中的一种预测模型,在本发明中不做限定,该语音风格预测模型是利用大量的训练样本以及训练样本对应的语音风格训练的,训练样本包括人的人脸信息、声音数据和性格数据。具体地,根据人机对话的语音数据确定声音数据和性格数据,然后再根据声音数据、性格数据以及语音风格预测模型,确定待切换语音风格。

[0066] 步骤S30,获取所述机器人终端当前的语音交互风格,基于所述待切换语音风格与所述当前的语音交互风格,确定所述机器人终端是否满足语音交互风格的切换条件;

[0067] 在本实施例中,机器人终端与用户正在进行人机对话,机器人终端以一种语音交互风格与用户对话,故获取机器人终端当前的语音交互风格,并将机器人终端当前的语音交互风格与待切换语音风格进行比较,进而确定机器人终端是否需要语音交互风格切换。

[0068] 具体地,当机器人终端当前的语音交互风格与待切换语音风格一致时,则说明不需要进行语音交互风格切换,反之,当机器人终端当前的语音交互风格与待切换语音风格不一致时,则说明需要进行语音交互风格切换,并将机器人终端的语音交互风格切换为待切换语音风格。

[0069] 步骤S40,在确定所述机器人终端满足语音交互风格的切换条件时,更新所述机器人终端的语音交互风格为待切换语音风格;

[0070] 在本实施例中,当机器人终端与用户正在进行人机对话时,实时获取的语音数据,可能因为用户的情绪、谈论的话题而引起用户当前语音风格的变化,机器人终端根据语音数据确定的声音数据和性格数据也会发生变化,故将有变化的声音数据和性格数据作为语音风格预测模型的输入后,该语音风格预测模型输出的待切换语音风格可能与机器人终端当前的语音风格是不同的,如果二者不同,则需要语音风格切换。具体地,当机器人终端当前的语音交互风格与待切换语音风格不一致时,则说明需要进行语音交互风格切换,故将机器人终端的语音交互风格切换为待切换语音风格。

[0071] 举例说明,当机器人终端与用户正在进行人机对话时,获取人机对话的语音数据,并根据语音数据确定声音数据和性格数据,将声音数据和性格数据输入语音风格预测模型中,对该用户风格进行预测判断,输出待切换语音风格是可爱型,而获取机器人终端当前的语音交互风格是稳重型,即机器人终端当前的语音交互风格与待切换语音风格不一致,则

将机器人终端的语音交互风格切换为可爱型。

[0072] 进一步地,在一实施例中,步骤S50之后还包括:

[0073] 步骤S50,获取人机对话的语音数据以及人机对话的视频信息;

[0074] 在本实施例中,当机器人终端当前的语音交互风格与待切换语音风格不一致而进行语音交互风格切换后,机器人终端的交互风格发生变化,需要继续监测人机对话的对话内容,以及利用摄像头监测该用户的表情,根据人机对话的对话内容以及用户的表情进一步确定用户是否对切换后的语音交互风格满意。

[0075] 步骤S60,基于所述语音数据以及所述视频信息确定所述机器人终端是否满足语音风格保持条件;

[0076] 在本实施例中,根据获取的语音数据,利用NLP技术进行解析,进而确定用户与机器人终端对话过程中,用户没有出现不愉悦的话语,同时,利用摄像头监测该用户的有没有出现不愉悦的表情,进而确定是否保持当前语音风格,具体地,如果该用户未出现不愉悦的话语与表情,则机器人终端保持当前语音风格,如果该用户出现不愉悦的话语与表情,则机器人终端不保持当前语音风格。

[0077] 步骤S70,在确定所述机器人终端不满足语音风格保持条件时,更新所述机器人终端的语音交互风格为默认风格。

[0078] 在本实施例中,当确定机器人终端保持当前语音风格,即用户未出现不愉悦的话语与表情,则继续以该语音交互风格进行人机对话;当确定机器人终端不保持当前语音风格,即用户出现不愉悦的话语与表情,则机器人终端的语音交互风格切换到默认风格。

[0079] 进一步地,在一实施例中,步骤S50之后还包括:

[0080] 步骤a,获取人机对话的语音数据;

[0081] 步骤b,确定所述语音数据中是否存在切换语音风格的信息;

[0082] 步骤c,在所述语音数据中存在切换语音风格的信息时,基于所述切换语音风格的信息执行语音切换操作。

[0083] 在本实施例中,获取人机对话的语音数据,根据获取的语音数据,利用NLP技术进行解析,进而确定用户与机器人终端对话过程中,用户有没有出现切换语音风格的内容。机器人终端不仅可以人物相关信息进行语音风格设定,也支持用户自主切换语音对话风格。

[0084] 具体地,在人机对话过程中,获取人机对话的语音数据,进一步确定语音数据中是否存在切换语音风格的内容,当语音数据中存在切换语音风格的内容时,则机器人终端进行语音风格的切换操作,切换至用户希望的语音风格。

[0085] 进一步地,在一实施例中,步骤S50之后还包括:

[0086] 步骤d,在检测到人机对话结束指令时,获取所述人机对话结束指令对应的目标语音交互风格以及所述人机对话结束指令对应的目标人物信息,其中,所述目标人物信息包括:人脸信息、声音数据以及性格数据;

[0087] 在本实施例中,当机器人终端检测到人机对话结束指令时,则结束人机对话,此时,获取机器人终端当前的语音交互风格,以及在人机对话结束前,与机器人终端进行对话的用户的人脸信息、声音数据以及性格数据。其中,人脸信息是在机器人终端启动人机对话时,基于摄像头获取的人脸信息;声音数据和性格数据是在机器人终端进行人机对话时,获

取人机对话的语音数据,然后基于语音数据分析得到的声音数据以及性格数据。

[0088] 步骤e,关联保存所述目标语音交互风格和所述目标人物信息至所述人物与语音风格匹配库。

[0089] 在本实施例中,关联保存目标语音交互风格和目标人物信息至人物与语音风格匹配库。如果人物与语音风格匹配库之前已经保存该目标人物信息的历史数据,则将目标语音交互风格和目标人物信息替换该目标人物信息的历史数据,如果人物与语音风格匹配库之前未保存该目标人物信息的历史数据,则在人物与语音风格匹配库新建一条记录,保存目标语音交互风格和目标人物信息。

[0090] 本实施例提出的个性化语音交互方法,通过在所述机器人终端当前进行人机对话时,获取人机对话的语音数据,并基于所述语音数据确定声音数据和性格数据,而后基于所述声音数据、所述性格数据以及语音风格预测模型,确定待切换语音风格,然后获取所述机器人终端当前的语音交互风格,基于所述待切换语音风格与所述当前的语音交互风格,确定所述机器人终端是否满足语音交互风格的切换条件,最后在确定所述机器人终端满足语音交互风格的切换条件时,更新所述机器人终端的语音交互风格为待切换语音风格。通过人机对话的语音数据和语音风格预测模型确定机器人终端的语音风格,实现了机器人终端自适应对话人物风格,从而带来更加符合用户对话习惯的语音交互体验。

[0091] 基于第一实施例,参照图3,提出本发明个性化语音交互方法的第二实施例,在本实施例中,步骤S10之后,还包括:

[0092] 步骤S80,在检测到所述机器人终端启动人机对话时,基于摄像头获取人脸信息;

[0093] 在本实施例中,当有用户与机器人终端刚刚启动人机对话时,机器人终端通过摄像头采集用户面部图像,并基于现有技术中的图像识别技术,从用户面部图像中捕捉该用户的人脸信息,其中,摄像头可以是机器人终端的内置摄像头,也可以是外置摄像头,在本发明中不做限定。

[0094] 步骤S90,确定所述人脸信息是否满足预设条件;

[0095] 在本实施例中,在获取到用户的人脸信息后,进一步确定该人脸信息是否满足预设条件,人脸信息用于用户身份的识别,进而确定机器人终端与该用户进行人机对话时的初始语音交互风格。

[0096] 具体地,步骤S90包括:确定所述人物与语音风格匹配库是否存在所述人脸信息,其中,在所述人物与语音风格匹配库存在所述人脸信息时,确定所述人脸信息满足预设条件

[0097] 在本实施例中,人物与语音风格匹配库关联保存了人物信息以及该人物信息对应的语音风格,其中,人物信息包括人脸,常用音色、常用音调、声纹图像、人物性格、爱好等。需要说明的是,人物信息和人物信息对应的语音风格支持更新,同理,该人物与语音风格匹配库也支持更新,即可增加记录、删除记录、修改记录等。当机器人终端与用户停止人机对话后,将该用户的对应的语音交互风格和人物信息更新至人物与语音风格匹配库。

[0098] 具体地,判断人物与语音风格匹配库是否存在当前获取到的人脸信息,即人物与语音风格匹配库是否保存了该用户的历史记录,如果人物与语音风格匹配库存在该人脸信息时,则确定人脸信息满足预设条件。

[0099] 步骤S100,在所述人脸信息满足预设条件时,基于人物与语音风格匹配库获取所

述人脸信息对应的个性化语音风格；

[0100] 步骤S110,更新所述机器人终端的语音交互风格为所述个性化语音风格。

[0101] 在本实施例中,当人物与语音风格匹配库存在当前获取到的人脸信息,即人物与语音风格匹配库保存了该用户的历史记录,则确定人脸信息满足预设条件,此时,可以直接从人物与语音风格匹配库获取该人脸信息对应的语音风格即个性化语音风格,故确定机器人终端与该用户进行人机对话时的初始语音交互风格为个性化语音风格,其中,个性化语音风格为活跃型、稳重型、幽默型、可爱型或较真型等。

[0102] 而后,当确定了机器人终端与用户进行人机对话时的初始语音交互风格为个性化语音风格,则将机器人终端的语音交互风格设置为个性化语音风格。

[0103] 进一步地,在一实施例中,步骤S90之后还包括:

[0104] 步骤S120,在所述人脸信息不满足预设条件时,更新所述机器人终端的语音交互风格为默认风格;

[0105] 在本实施例中,当人物与语音风格匹配库不存在当前获取到的人脸信息,即人物与语音风格匹配库未保存过该用户的历史记录,则确定人脸信息不满足预设条件,此时,确定机器人终端与该用户进行人机对话时的初始语音交互风格为默认风格,其中,默认风格是机器人终端一种标准的人机对话风格,在出厂时设定好。

[0106] 进一步地,在一实施例中,步骤S80之前还包括:基于预设数量的人物信息样本以及所述人物信息样本对应的语音风格生成所述人物与语音风格匹配库,其中,所述人物信息样本包括人脸信息、声音数据以及性格数据。

[0107] 在本实施例中,获取预设数量的人物信息样本,其中,人物信息样本包括人脸信息、声音数据以及性格数据,具体地,声音数据至少包括常用音色、常用音调,还可以包括音强、音量等;性格数据至少包括人物性格,还可以包括爱好、习惯等。另外,人物信息样本对应的语音风格是基于历史数据已经确定的。进一步地,根据人物信息样本以及人物信息样本对应的语音风格生成人物与语音风格匹配库,用于保存已经与机器人终端进行过人机交互的用户的相关信息。需要说明的是,该人物与语音风格匹配库支持更新,机器人终端在进行人机对话的时,如果发现该用户的信息未在数据库中,则在人机对话结束时,机器人终端会将该用户对应的人物信息和语音风格保存到人物与语音风格匹配库。

[0108] 进一步地,在一实施例中,步骤S80之前,还包括:

[0109] 步骤f,基于所述人物与语音风格匹配库,定时获取人物信息样本以及所述人物信息样本对应的语音风格;

[0110] 步骤g,基于所述人物信息样本以及所述人物信息样本对应的语音风格训练初始语音风格预测模型,得到所述语音风格预测模型。

[0111] 在本实施例中,在人物与语音风格匹配库中获取数据库中保存的人物信息和该人物信息对应的语音风格,其中,人物信息包括人脸信息、声音数据以及性格数据,具体地,声音数据至少包括常用音色、常用音调,还可以包括音强、音量等;性格数据至少包括人物性格,还可以包括爱好、习惯等。人物信息样本对应的语音风格是基于历史数据已经确定的,将人物信息样本以及人物信息样本对应的语音风格作为训练样本,训练初始语音风格预测模型,最终得到一个语音风格预测模型,其中,初始语音风格预测模型是现有技术中的一种预测模型,在本发明中不做限定。

[0112] 需要说明的是,机器人终端在进行人机对话的时,如果发现该用户的信息未在数据库中,则在人机对话结束时,机器人终端会将该用户对应的人物信息和语音风格保存到人物与语音风格匹配库,该匹配库的样本数量越来越多,样本数量越多,得到的语音风格预测模型就越精确,故需要定期利用人物与语音风格匹配库的样本进行模型训练,以使语音风格预测模型的预测效果更好。

[0113] 本实施例提出的个性化语音交互方法,通过在检测到所述机器人终端启动人机对话时,基于摄像头获取人脸信息,而后确定所述人脸信息是否满足预设条件,然后在所述人脸信息满足预设条件时,基于人物与语音风格匹配库获取所述人脸信息对应的个性化语音风格,最后更新所述机器人终端的语音交互风格为所述个性化语音风格。通过确定人物与语音风格匹配库是否存在所述人脸信息,进而确定机器人终端与该用户进行人机对话时的初始语音交互风格,满足语音风格的个性化设置,提升用户体验。

[0114] 本发明进一步提供一种个性化语音交互装置,参照图4,图4为本发明个性化语音交互装置实施例的功能模块示意图。

[0115] 获取模块10,在所述机器人终端当前进行人机对话时,获取人机对话的语音数据,并基于所述语音数据确定声音数据和性格数据;

[0116] 确定模块20,基于所述声音数据、所述性格数据以及语音风格预测模型,确定待切换语音风格;

[0117] 判断模块30,获取所述机器人终端当前的语音交互风格,基于所述待切换语音风格与所述当前的语音交互风格,确定所述机器人终端是否满足语音交互风格的切换条件;

[0118] 更新模块40,在确定所述机器人终端满足语音交互风格的切换条件时,更新所述机器人终端的语音交互风格为待切换语音风格。

[0119] 进一步地,所述个性化语音交互装置还用于:

[0120] 第一采集模块,获取人机对话的语音数据以及人机对话的视频信息;

[0121] 第一判别模块,基于所述语音数据以及所述视频信息确定所述机器人终端是否满足语音风格保持条件;

[0122] 第一切换模块,在确定所述机器人终端不满足语音风格保持条件时,更新所述机器人终端的语音交互风格为默认风格。

[0123] 进一步地,所述个性化语音交互装置还用于:

[0124] 第二采集模块,获取人机对话的语音数据;

[0125] 第二判别模块,确定所述语音数据中是否存在切换语音风格的信息;

[0126] 第二切换模块,在所述语音数据中存在切换语音风格的信息时,基于所述切换语音风格的信息执行语音切换操作。

[0127] 进一步地,所述个性化语音交互装置还用于:

[0128] 第三采集模块,在检测到所述机器人终端启动人机对话时,基于摄像头获取人脸信息;

[0129] 第三判别模块,确定所述人脸信息是否满足预设条件;

[0130] 筛选模块,在所述人脸信息满足预设条件时,基于人物与语音风格匹配库获取所述人脸信息对应的个性化语音风格;

[0131] 第三切换模块,更新所述机器人终端的语音交互风格为所述个性化语音风格。

[0132] 进一步地,所述个性化语音交互装置还用于:

[0133] 第三切换模块,在所述人脸信息不满足预设条件时,更新所述机器人终端的语音交互风格为默认风格。

[0134] 进一步地,所述个性化语音交互装置还用于:

[0135] 生成模块,基于预设数量的人物信息样本以及所述人物信息样本对应的语音风格生成所述人物与语音风格匹配库,其中,所述人物信息样本包括人脸信息、声音数据以及性格数据。

[0136] 进一步地,所述个性化语音交互装置还用于:

[0137] 第四采集模块,基于所述人物与语音风格匹配库,定时获取人物信息样本以及所述人物信息样本对应的语音风格;

[0138] 训练模块,基于所述人物信息样本以及所述人物信息样本对应的语音风格训练初始语音风格预测模型,得到所述语音风格预测模型。

[0139] 进一步地,所述第三判别模块还用于:

[0140] 确定所述人物与语音风格匹配库是否存在所述人脸信息,其中,在所述人物与语音风格匹配库存在所述人脸信息时,确定所述人脸信息满足预设条件。

[0141] 进一步地,所述个性化语音交互装置还用于:

[0142] 第五采集模块,在检测到人机对话结束指令时,获取所述人机对话结束指令对应的目标语音交互风格以及所述人机对话结束指令对应的目标人物信息,其中,所述目标人物信息包括:人脸信息、声音数据以及性格数据;

[0143] 保存模块,关联保存所述目标语音交互风格和所述目标人物信息至所述人物与语音风格匹配库。

[0144] 此外,本发明实施例还提出一种可读存储介质,所述可读存储介质上存储有个性化语音交互程序,所述个性化语音交互程序被处理器执行时实现上述各个实施例中个性化语音交互方法的步骤。

[0145] 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。

[0146] 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。

[0147] 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可读存储介质(如ROM/RAM、磁碟、光盘)中,包括若干指令用以使得一台系统设备(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。

[0148] 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。

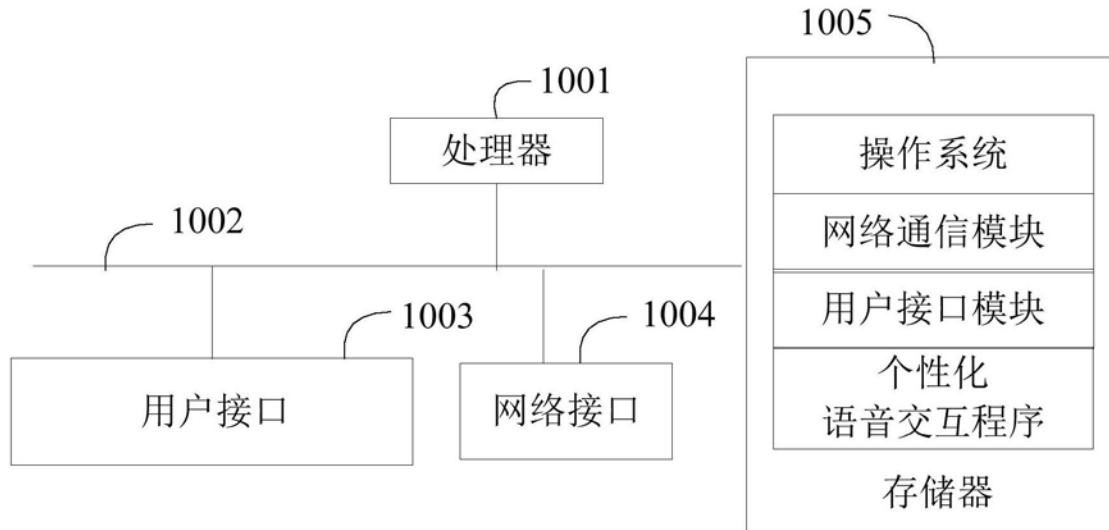


图1

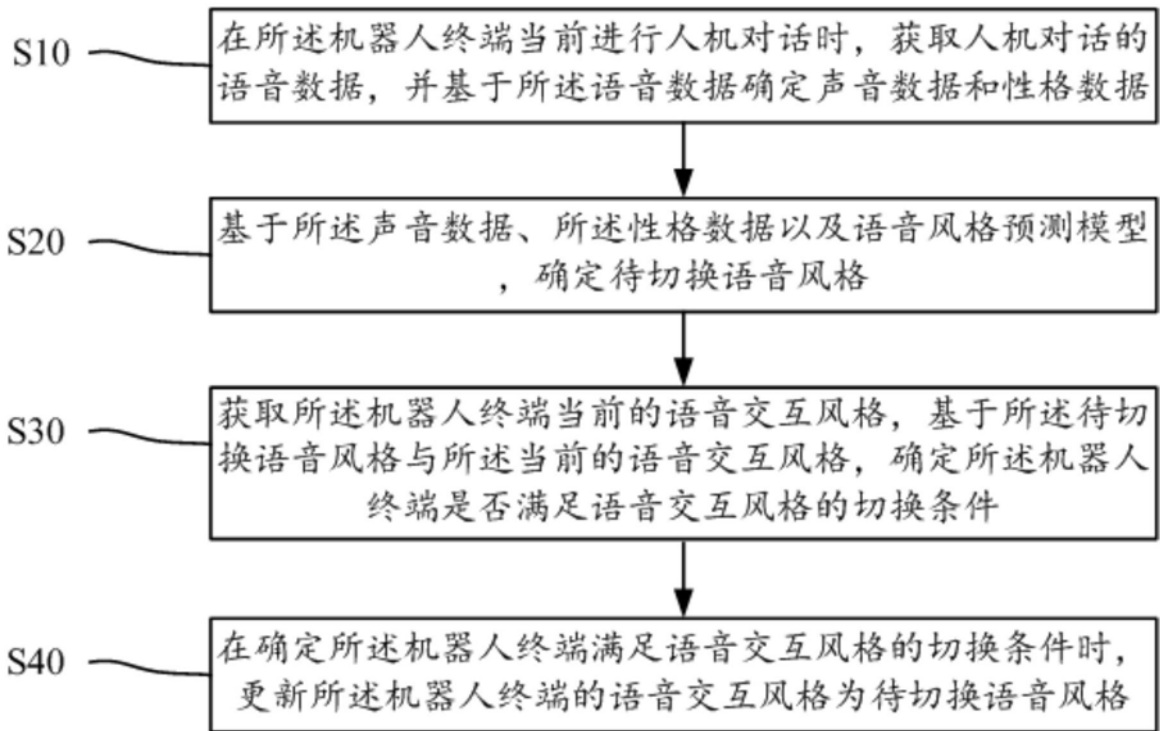


图2

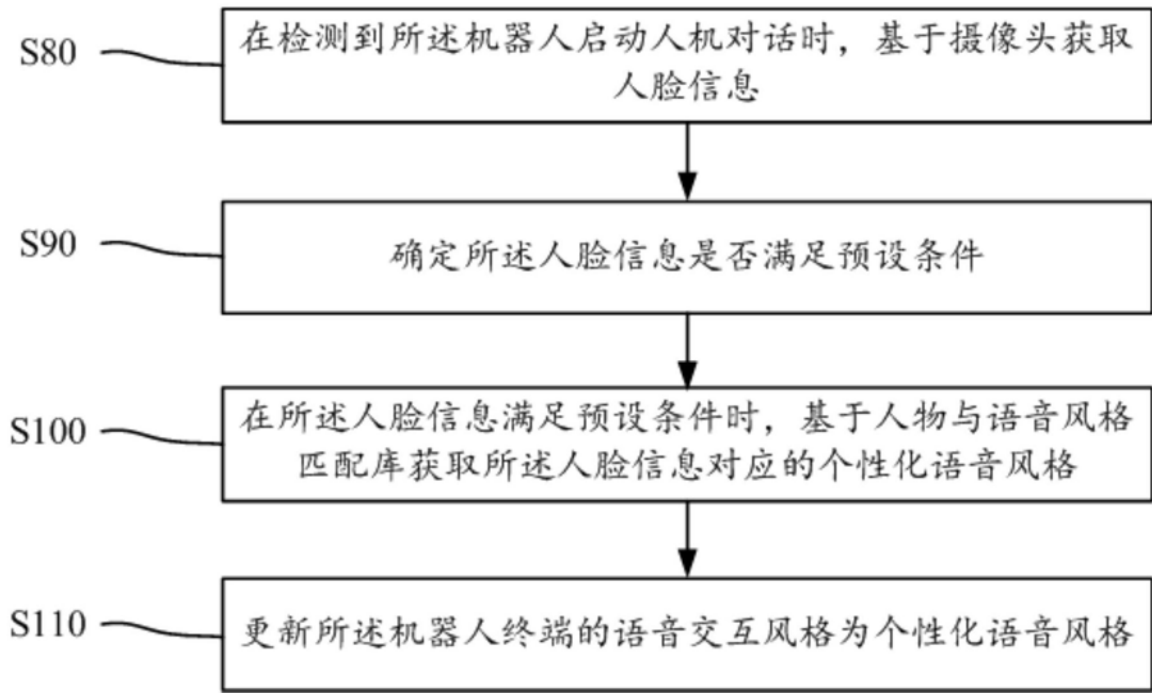


图3

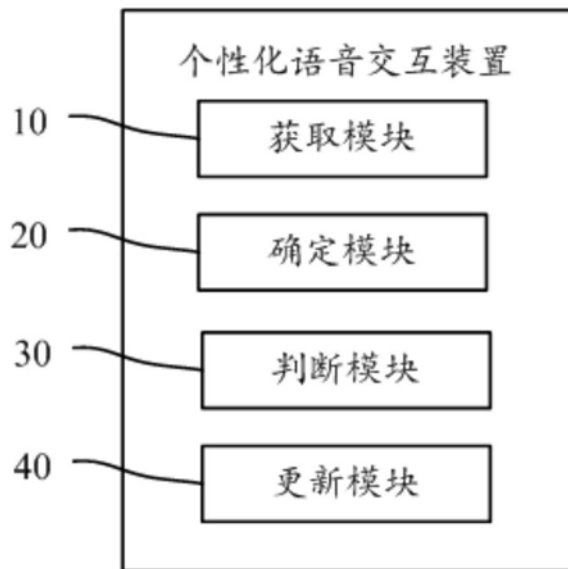


图4