



(12)发明专利申请

(10)申请公布号 CN 107274891 A

(43)申请公布日 2017. 10. 20

(21)申请号 201710368669.X

(22)申请日 2017.05.23

(71)申请人 武汉秀宝软件有限公司

地址 430000 湖北省武汉市东湖新技术开发区武大园路8号武大科技园一号楼生产楼宏业楼1楼C-6号房(部位)

(72)发明人 胡德志 孙碧亮 袁超飞

(74)专利代理机构 北京轻创知识产权代理有限公司 11212

代理人 陈薇

(51)Int.Cl.

G10L 15/06(2013.01)

G10L 15/22(2006.01)

G06F 3/01(2006.01)

G06F 9/44(2006.01)

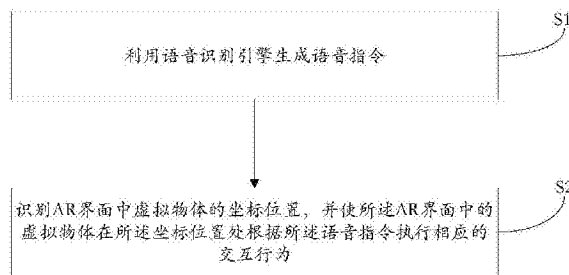
权利要求书2页 说明书5页 附图2页

(54)发明名称

一种基于语音识别引擎的AR界面交互方法及系统

(57)摘要

本发明涉及一种基于语音识别引擎的AR界面交互方法及系统,其方法包括以下步骤,S1,利用语音识别引擎生成语音指令;S2,识别AR界面中虚拟物体的坐标位置,并使所述AR界面中的虚拟物体在所述坐标位置处根据所述语音指令执行相应的交互行为。本发明一种基于语音识别引擎的AR界面交互方法解决了现有增强现实技术对玩家的语音无法进行识别的问题,丰富了AR应用中人与游戏的交互性,不仅可以通过图片与虚拟物体进行交互,还能通过语音达到对虚拟物体一定行为的控制功能,增强了AR游戏的趣味性与可玩性。



1. 一种基于语音识别引擎的AR界面交互方法,其特征在于:包括以下步骤,
 - S1,利用语音识别引擎生成语音指令;
 - S2,识别AR界面中虚拟物体的坐标位置,并使所述AR界面中的虚拟物体在所述坐标位置处根据所述语音指令执行相应的交互行为。
2. 根据权利要求1所述的一种基于语音识别引擎的AR界面交互方法,其特征在于:所述S1具体为,
 - S11,对需要语音识别的词汇进行离线学习,得到语音识别库;
 - S12,将所述语音识别库导入到语音识别引擎中,
 - S13,将语音信号输入至所述语音识别引擎中,并将所述语音信号与所述语音识别库中的数据进行匹配,所述语音识别引擎根据匹配结果生成对应的语音指令。
3. 根据权利要求2所述的一种基于语音识别引擎的AR界面交互方法,其特征在于:所述S11具体为,
 - S111,对需要语音识别的词汇进行标准发音,生成指令音频文件;
 - S112,通过语音识别引擎离线学习工具对所述指令音频文件进行训练,并提取所述指令音频文件中的特征信息,生成指令文本文件;
 - S113,对所述指令文本文件进行统计,并提取不同指令文本文件之间先后发生的统计关系;
 - S114,根据所述统计关系构建语音识别模型,并输出语音识别模型的初始精度参数,
 - S115,利用参数精度参数对所述语音识别模型进行反复测试,并根据测试结果调节所述初始精度参数,得出最终精度参数;
 - S116,将所述最终精度参数与所述语音识别模型结合生成语音识别库。
4. 根据权利要求1至3任一项所述的一种基于语音识别引擎的AR界面交互方法,其特征在于:所述语音识别引擎是基于PocketSphinx语音识别系统开发。
5. 根据权利要求2或3所述的一种基于语音识别引擎的AR界面交互方法,其特征在于:在所述S13中,将所述语音信号与所述语音识别库中的数据进行匹配前还包括:对所述语音信号进行过滤处理。
6. 一种基于语音识别引擎的AR界面交互系统,其特征在于:包括语音识别引擎和AR引擎,
 - 所述语音识别引擎,其用于生成语音指令;
 - 所述AR引擎,其用于识别AR界面中虚拟物体的坐标位置,并使所述AR界面中的虚拟物体在所述坐标位置处根据所述语音指令执行相应的交互行为。
7. 根据权利要求6所述的一种基于语音识别引擎的AR界面交互系统,其特征在于:所述语音识别引擎具体用于,
 - 对需要语音识别的词汇进行离线学习,得到语音识别库;
 - 将所述语音识别库导入到语音识别引擎中,
 - 将语音信号输入至所述语音识别引擎中,并将所述语音信号与所述语音识别库中的数据进行匹配,所述语音识别引擎根据匹配结果生成对应的语音指令。
8. 根据权利要求7所述的一种基于语音识别引擎的AR界面交互系统,其特征在于:所述语音识别引擎具体用于,

对需要语音识别的词汇进行标准发音,生成指令音频文件;

通过语音识别引擎离线学习工具对所述指令音频文件进行训练,并提取所述指令音频文件中的特征信息,生成指令文本文件;

对所述指令文本文件进行统计,并提取不同指令文本文件之间先后发生的统计关系;

根据所述统计关系构建语音识别模型,并输出语音识别模型的初始精度参数,

利用参数精度参数对所述语音识别模型进行反复测试,并根据测试结果调节所述初始精度参数,得出最终精度参数;

将所述最终精度参数与所述语音识别模型结合生成语音识别库。

9. 根据权利要求6至8任一项所述的一种基于语音识别引擎的AR界面交互系统,其特征在于:所述语音识别引擎是基于PocketSphinx语音识别系统开发。

10. 根据权利要求7或8所述的一种基于语音识别引擎的AR界面交互系统,其特征在于:在所述语音识别引擎中,将所述语音信号与所述语音识别库中的数据进行匹配前还包括:对所述语音信号进行过滤处理。

一种基于语音识别引擎的AR界面交互方法及系统

技术领域

[0001] 本发明涉及AR技术领域,具体的涉及一种基于语音识别引擎的AR界面交互方法及系统。

背景技术

[0002] 现有的关于增强现实技术 (Augmented Reality,简称AR) 方面的游戏的玩法是玩家通过操作设备界面上的按钮来执行对虚拟角色的行为操作,而这种玩法无法快速达到指挥AR界面上虚拟物体的目的;同时,现有的语音识别技术对大词汇量的识别无法达到很高的识别精度,而在AR游戏中,是不能容许语音指令的出现较高的误检率的,如果出现较高的误检率会降低游戏的可玩性。

发明内容

[0003] 本发明所要解决的技术问题是提供一种基于语音识别引擎的AR界面交互方法及系统,解决了现有增强现实技术对玩家的语音无法进行识别的问题,不仅能提高AR游戏应用的可玩性,也丰富了人与虚拟现实之间的互动,增强了AR技术产品的趣味性。

[0004] 本发明解决上述技术问题的技术方案如下:一种基于语音识别引擎的AR界面交互方法,包括以下步骤,

[0005] S1,利用语音识别引擎生成语音指令;

[0006] S2,识别AR界面中虚拟物体的坐标位置,并使所述AR界面中的虚拟物体在所述坐标位置处根据所述语音指令执行相应的交互行为。

[0007] 本发明的有益效果是:本发明一种基于语音识别引擎的AR界面交互方法利用语音识别引擎生成语音指令;识别AR界面中虚拟物体的坐标位置,并使所述AR界面中的虚拟物体在所述坐标位置处根据所述语音指令执行相应的交互行为;该方法解决了现有增强现实技术对玩家的语音无法进行识别的问题,丰富了AR应用中人与游戏的交互性,不仅可以图片与虚拟物体进行交互,还能通过语音达到对虚拟物体一定行为的控制功能,增强了AR游戏的趣味性与可玩性。

[0008] 在上述技术方案的基础上,本发明还可以做如下改进。

[0009] 进一步,所述S1具体为,

[0010] S11,对需要语音识别的词汇进行离线学习,得到语音识别库;

[0011] S12,将所述语音识别库导入到语音识别引擎中,

[0012] S13,将语音信号输入至所述语音识别引擎中,并将所述语音信号与所述语音识别库中的数据进行匹配,所述语音识别引擎根据匹配结果生成对应的语音指令。

[0013] 进一步,所述S11具体为,

[0014] S111,对需要语音识别的词汇进行标准发音,生成指令音频文件;

[0015] S112,通过语音识别引擎离线学习工具对所述指令音频文件进行训练,并提取所述指令音频文件中的特征信息,生成指令文本文件;

[0016] S113,对所述指令文本文件进行统计,并提取不同指令文本文件之间先后发生的统计关系;

[0017] S114,根据所述统计关系构建语音识别模型,并输出语音识别模型的初始精度参数,

[0018] S115,利用参数精度参数对所述语音识别模型进行反复测试,并根据测试结果调节所述初始精度参数,得出最终精度参数;

[0019] S116,将所述最终精度参数与所述语音识别模型结合生成语音识别库。

[0020] 采用上述进一步方案的有益效果是:本发明一种基于语音识别引擎的AR界面交互方法基于语音识别引擎,先对需要离线学习的语音指令的音频文件进行反复的训练和对识别参数的反复调试,以获取到高准确率率的语音识别库,提高语音识别的精度。

[0021] 进一步,所述语音识别引擎是基于PocketSphinx语音识别系统开发。

[0022] 采用上述进一步方案的有益效果是:基于PocketSphinx语音识别系统开发的语音识别引擎是一个计算量和体积都很小的语音识别引擎,它对小词汇量的识别精度是很高的,并且对性能的销毁很小,反应很快,可以进一步解决现有增强现实技术对玩家的语音无法进行识别的问题。

[0023] 进一步,在所述S13中,将所述语音信号与所述语音识别库中的数据进行匹配前还包括:对所述语音信号进行过滤处理。

[0024] 采用上述进一步方案的有益效果是:对语音信号进行过滤处理后再与语音识别库中的数据进行匹配,可以去除语音信号中的噪声,避免干扰,提高匹配的准确率。

[0025] 基于上述一种基于语音识别引擎的AR界面交互方法,本发明还提供一种基于语音识别引擎的AR界面交互系统。

[0026] 一种基于语音识别引擎的AR界面交互系统,包括语音识别引擎和AR引擎,

[0027] 所述语音识别引擎,其用于生成语音指令;

[0028] 所述AR引擎,其用于识别AR界面中虚拟物体的坐标位置,并使所述AR界面中的虚拟物体在所述坐标位置处根据所述语音指令执行相应的交互行为。

[0029] 本发明的有益效果是:本发明一种基于语音识别引擎的AR界面交互系统将语音信号在语音识别引擎中生成语音指令并传送到AR引擎中,通过结合AR引擎识别追踪得到AR界面中虚拟物体的坐标位置,对AR界面中虚拟物体进行相应的行为控;该方法解决了现有增强现实技术对玩家的语音无法进行识别的问题,丰富了AR应用中人与游戏的交互性,不仅可以通过图片与虚拟物体进行交互,还能通过语音达到对虚拟物体一定行为的控制功能,增强了AR游戏的趣味性与可玩性。

[0030] 在上述技术方案的基础上,本发明还可以做如下改进。

[0031] 进一步,所述语音识别引擎具体用于,

[0032] 对需要语音识别的词汇进行离线学习,得到语音识别库;

[0033] 将所述语音识别库导入到语音识别引擎中,

[0034] 将语音信号输入至所述语音识别引擎中,并将所述语音信号与所述语音识别库中的数据匹配,所述语音识别引擎根据匹配结果生成对应的语音指令。

[0035] 进一步,所述语音识别引擎具体用于,

[0036] 对需要语音识别的词汇进行标准发音,生成指令音频文件;

- [0037] 通过语音识别引擎离线学习工具对所述指令音频文件进行训练,并提取所述指令音频文件中的特征信息,生成指令文本文件;
- [0038] 对所述指令文本文件进行统计,并提取不同指令文本文件之间先后发生的统计关系;
- [0039] 根据所述统计关系构建语音识别模型,并输出语音识别模型的初始精度参数,
- [0040] 利用参数精度参数对所述语音识别模型进行反复测试,并根据测试结果调节所述初始精度参数,得出最终精度参数;
- [0041] 将所述最终精度参数与所述语音识别模型结合生成语音识别库。
- [0042] 采用上述进一步方案的有益效果是:本发明一种基于语音识别引擎的AR界面交互系统基于语音识别引擎,先对需要离线学习的语音指令的音频文件进行反复的训练和对识别参数的反复调试,以获取到高准确率的语音识别库,提高语音识别的精度。
- [0043] 进一步,所述语音识别引擎是基于PocketSphinx语音识别系统开发。
- [0044] 采用上述进一步方案的有益效果是:基于PocketSphinx语音识别系统开发的语音识别引擎是一个计算量和体积都很小的语音识别引擎,它对小词汇量的识别精度是很高的,并且对性能的销毁很小,反应很快,可以进一步解决现有增强现实技术对玩家的语音无法进行识别的问题。
- [0045] 进一步,在所述语音识别引擎中,将所述语音信号与所述语音识别库中的数据进行匹配前还包括:对所述语音信号进行过滤处理。
- [0046] 采用上述进一步方案的有益效果是:对语音信号进行过滤处理后再与语音识别库中的数据进行匹配,可以去除语音信号中的噪声,避免干扰,提高匹配的准确率。

附图说明

- [0047] 图1为本发明一种基于语音识别引擎的AR界面交互方法的整体流程图;
- [0048] 图2为本发明一种基于语音识别引擎的AR界面交互方法中生成语音指令的流程图;
- [0049] 图3为本发明一种基于语音识别引擎的AR界面交互方法中对需要语音识别的词汇进行离线学习得到语音识别库的流程图;
- [0050] 图4为本发明一种基于语音识别引擎的AR界面交互系统的结构框图。

具体实施方式

- [0051] 以下结合附图对本发明的原理和特征进行描述,所举实例只用于解释本发明,并非用于限定本发明的范围。
- [0052] 如图1所示,一种基于语音识别引擎的AR界面交互方法,包括以下步骤,
- [0053] S1,利用语音识别引擎生成语音指令;
- [0054] S2,识别AR界面中虚拟物体的坐标位置,并使所述AR界面中的虚拟物体在所述坐标位置处根据所述语音指令执行相应的交互行为。
- [0055] 在本具体实施例中,所述S1具体为,如图2所示,
- [0056] S11,对需要语音识别的词汇进行离线学习,得到语音识别库;
- [0057] S12,将所述语音识别库导入到语音识别引擎中,

[0058] S13,将语音信号输入至所述语音识别引擎中,并将所述语音信号与所述语音识别库中的数据进行匹配,所述语音识别引擎根据匹配结果生成对应的语音指令。

[0059] 在本具体实施例中,所述S11中,对需要语音识别的词汇进行离线学习得到语音识别库的具体步骤为,如图3所示,

[0060] S111,对需要语音识别的词汇进行标准发音,生成指令音频文件;

[0061] S112,通过语音识别引擎离线学习工具对所述指令音频文件进行训练,并提取所述指令音频文件中的特征信息,生成指令文本文件;

[0062] S113,对所述指令文本文件进行统计,并提取不同指令文本文件之间先后发生的统计关系;

[0063] S114,根据所述统计关系构建语音识别模型,并输出语音识别模型的初始精度参数,

[0064] S115,利用参数精度参数对所述语音识别模型进行反复测试,并根据测试结果调节所述初始精度参数,得出最终精度参数;

[0065] S116,将所述最终精度参数与所述语音识别模型结合生成语音识别库。

[0066] 在本具体实施例中,所述语音识别引擎是基于PocketSphinx语音识别系统开发。

[0067] 在本具体实施例中,在所述S13中,将所述语音信号与所述语音识别库中的数据进行匹配前还包括对所述语音信号进行过滤处理。

[0068] 下面对本发明一种基于语音识别引擎的AR界面交互方法以AR游戏为例进行说明。

[0069] 具体的例如:

[0070] 利用本发明的方法开发的一款AR游戏应用中,可以实现对玩家的士兵进行简单指挥操作的功能,依次为命令士兵优先攻击前方的敌人,优先攻击左边的敌人,优先攻击右边的敌人,向后防御这4条指令,而拿着AR道具(如开发的玩具枪)体验AR游戏的玩家无法快速的通过手动的操作移动端的相应按钮来达到指挥的士兵功能。利用本发明的方法开发的语音识别引擎可以快速精确的响应向前攻击,向左攻击,向右攻击,向后防御这4条简单的语音指令,做到这一点就需要,先通过测试人员对语音指令的词汇的进行标准发音,生成指令音频文件;通过语音识别引擎离线学习工具对所述指令音频文件进行训练,并提取所述指令音频文件中的特征信息,生成指令文本文件;对所述指令文本文件进行统计,并提取不同指令文本文件之间先后发生的统计关系;根据所述统计关系构建语音识别模型,并输出语音识别模型的初始精度参数,利用参数精度参数对所述语音识别模型进行反复测试,并根据测试结果调节所述初始精度参数,得出最终精度参数;将所述最终精度参数与所述语音识别模型结合生成语音识别库。

[0071] 把学习得到语音识别库导入到语音识别引擎中,从而实现对人发出的语音指令进行识别的功能;其具体方法是玩家标准的说出对应的语音指令如向前攻击,移动设备获取到相应的音频数据,输入到语音识别引擎中去,进行去噪等相应的过滤处理后,把音频数据与语音识别库里的数据做匹配,得到匹配结果后传送到AR引擎中,通过AR引擎对其做响应。

[0072] AR系统通过对指定图片的识别追踪可以稳定的获取到虚拟物体向现实世界中的位置信息,语音识别引擎传入的语音识别结果进行分析判断后转化为对游戏士兵的指挥操作指令。

[0073] 本发明的方法使得增强现实技术获得了语音识别功能系统,增强人与虚拟现实之

间的互动,增强了AR游戏的趣味性和可玩性;如玩家在拿着玩具枪体验AR游戏时,不再需要像传统游戏那样,必须通过操作移动端界面上的按钮来执行对虚拟角色的行为操作,而是通过自身的移动与语音指令来执行对AR游戏中虚拟角色的行为操作。

[0074] 基于上述一种基于语音识别引擎的AR界面交互方法,本发明还提供一种基于语音识别引擎的AR界面交互系统。

[0075] 如图4所示,一种基于语音识别引擎的AR界面交互系统,包括语音识别引擎和AR引擎,

[0076] 所述语音识别引擎,其用于生成语音指令;

[0077] 所述AR引擎,其用于识别AR界面中虚拟物体的坐标位置,并使所述AR界面中的虚拟物体在所述坐标位置处根据所述语音指令执行相应的交互行为。

[0078] 在本具体实施例中,所述语音识别引擎具体用于,

[0079] 对需要语音识别的词汇进行离线学习,得到语音识别库;

[0080] 将所述语音识别库导入到语音识别引擎中,

[0081] 将语音信号输入至所述语音识别引擎中,并将所述语音信号与所述语音识别库中的数据进行匹配,所述语音识别引擎根据匹配结果生成对应的语音指令。

[0082] 在本具体实施例中,所述语音识别引擎具体用于,

[0083] 对需要语音识别的词汇进行标准发音,生成指令音频文件;

[0084] 通过语音识别引擎离线学习工具对所述指令音频文件进行训练,并提取所述指令音频文件中的特征信息,生成指令文本文件;

[0085] 对所述指令文本文件进行统计,并提取不同指令文本文件之间先后发生的统计关系;

[0086] 根据所述统计关系构建语音识别模型,并输出语音识别模型的初始精度参数,

[0087] 利用参数精度参数对所述语音识别模型进行反复测试,并根据测试结果调节所述初始精度参数,得出最终精度参数;

[0088] 将所述最终精度参数与所述语音识别模型结合生成语音识别库。

[0089] 在本具体实施例中,所述语音识别引擎是基于PocketSphinx语音识别系统开发。

[0090] 在本具体实施例中,在所述语音识别引擎中,将所述语音信号与所述语音识别库中的数据进行匹配前还包括:对所述语音信号进行过滤处理。

[0091] 本发明一种基于语音识别引擎的AR界面交互系统将语音信号在语音识别引擎中生成语音指令并传送到AR引擎中,通过结合AR引擎识别追踪得到AR界面中虚拟物体的坐标位置,对AR界面中虚拟物体进行相应的行为控;该方法解决了现有增强现实技术对玩家的语音无法进行识别的问题,丰富了AR应用中人与游戏的交互性,不仅可以通过图片与虚拟物体进行交互,还能通过语音达到对虚拟物体一定行为的控制功能,增强了AR游戏的趣味性与可玩性。

[0092] 以上所述仅为本发明的较佳实施例,并不用以限制本发明,凡在本发明的精神和原则之内,所作的任何修改、等同替换、改进等,均应包含在本发明的保护范围之内。

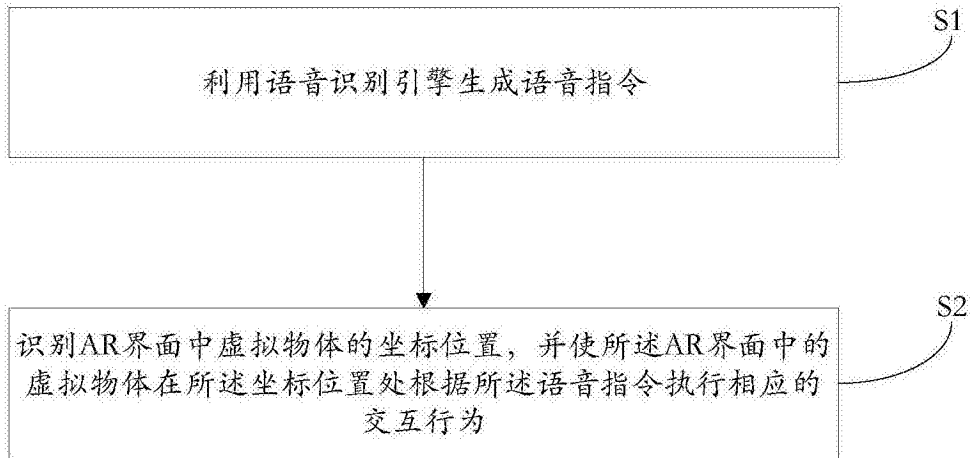


图1

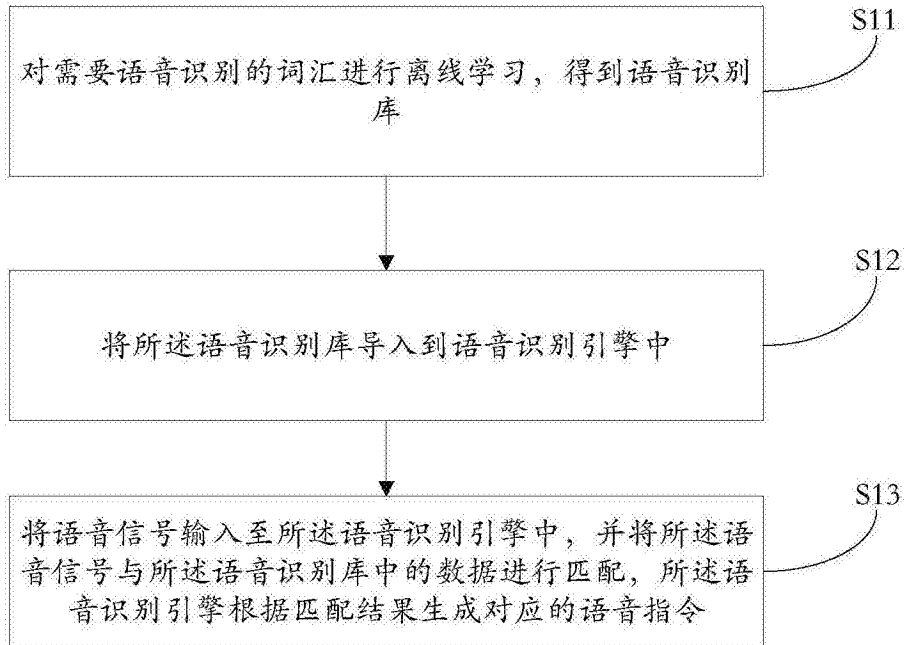


图2

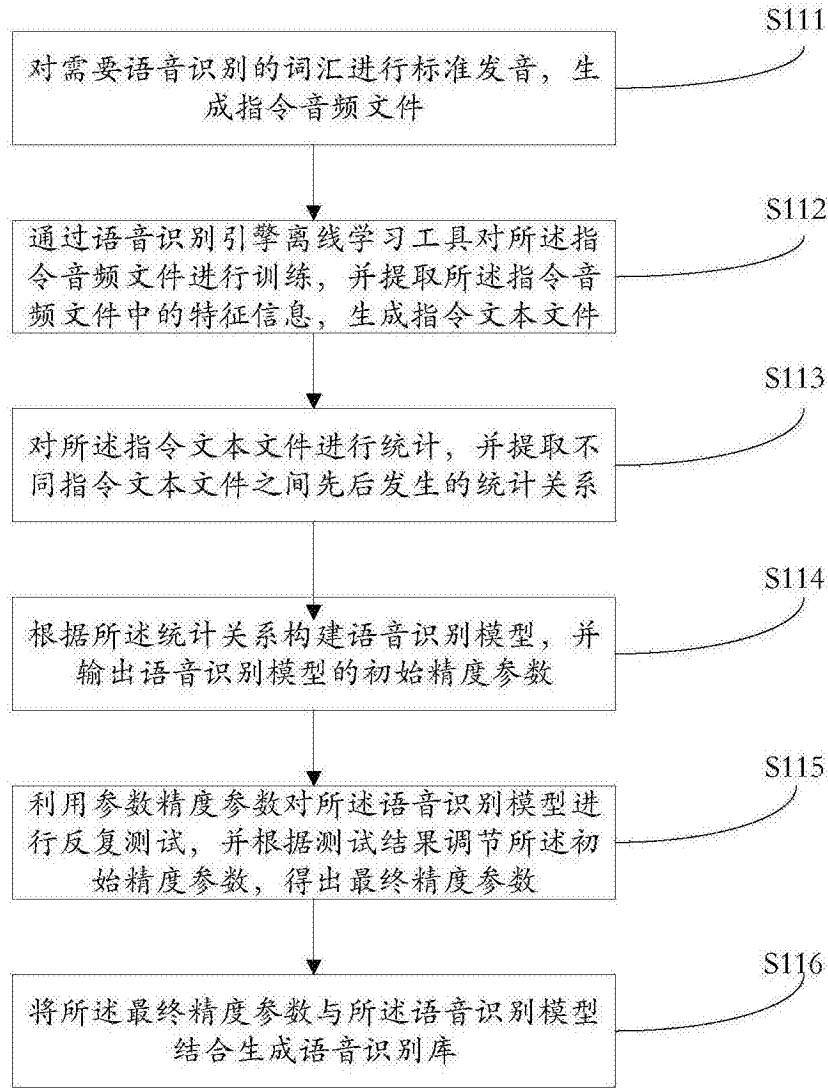


图3

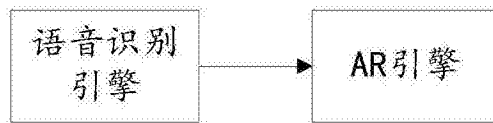


图4