



(12)发明专利

(10)授权公告号 CN 105488142 B

(45)授权公告日 2019.07.30

(21)申请号 201510836538.0

(51)Int.CI.

(22)申请日 2015.11.24

G06F 16/23(2019.01)

(65)同一申请的已公布的文献号

G06Q 50/20(2012.01)

申请公布号 CN 105488142 A

(56)对比文件

(43)申请公布日 2016.04.13

CN 104380375 A, 2015.02.25,

(73)专利权人 科大讯飞股份有限公司

CN 104732977 A, 2015.06.24,

地址 230088 安徽省合肥市高新区望
江西路666号

CN 104199956 A, 2014.12.10,

(72)发明人 李扬 钟锟 王丽红 徐飞龙

CN 104346389 A, 2015.02.11,

王国庆 王颖 徐瑞祥 梁昭
段忠苗

US 2015161271 A1, 2015.06.11,

(74)专利代理机构 北京维澳专利代理有限公司

US 2015081678 A1, 2015.03.19,

11252

赵玲等.应用计算机技术提高教学管理效

代理人 刘路尧 逢京喜

率.《中国教育信息化》.2007, 第80页-第81页.

审查员 唐文俊

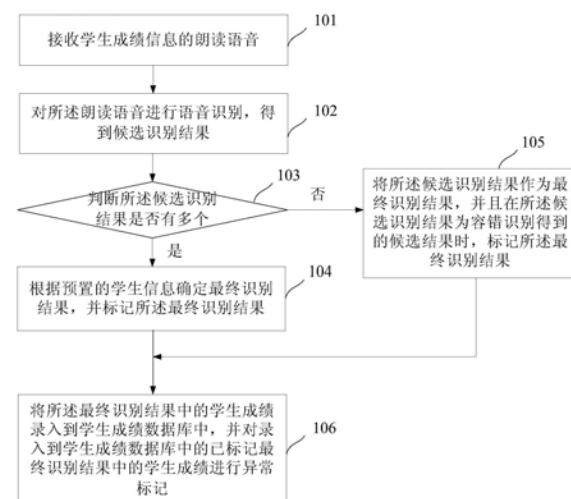
权利要求书3页 说明书9页 附图2页

(54)发明名称

成绩信息录入方法及系统

(57)摘要

本发明公开了一种成绩信息录入方法及系统,该方法包括:接收学生成绩信息的朗读语音,所述学生成绩信息包括:学生成绩和学生属性信息;对所述朗读语音进行语音识别,得到候选识别结果;如果所述候选识别结果有多个,则根据预置的学生信息确定最终识别结果,并标记所述最终识别结果;否则将所述候选识别结果作为最终识别结果,并且在所述候选识别结果为容错识别得到的候选结果时,标记所述最终识别结果;将所述最终识别结果中的学生成绩录入到学生成绩数据库中,并对录入到学生成绩数据库中的已标记最终识别结果中的学生成绩进行异常标记。利用本发明,可以更方便快捷、准确地录入学生成绩,减轻成绩录入的工作量。



1. 一种成绩信息录入方法,其特征在于,包括:

接收学生成绩信息的朗读语音,所述学生成绩信息包括:学生成绩和学生属性信息;

对所述朗读语音进行语音识别,得到候选识别结果;

如果所述候选识别结果有多个,则根据预置的学生信息确定最终识别结果,并标记所述最终识别结果,所述学生信息包括:学生名册、以及各学生整体能力得分和本次考试的预测得分这两者中其一或全部;否则将所述候选识别结果作为最终识别结果,并且在所述候选识别结果为容错识别得到的候选结果时,标记所述最终识别结果;

将所述最终识别结果中的学生成绩录入到学生成绩数据库中,并对录入到学生成绩数据库中的已标记最终识别结果中的学生成绩进行异常标记。

2. 根据权利要求1所述的方法,其特征在于,所述学生属性信息包括以下任意一种或多种:姓名、学号、准考证号。

3. 根据权利要求1所述的方法,其特征在于,所述方法还包括:

基于各学生历史考试成绩,应用IRT模型计算得到所述学生整体能力得分及试题难度得分;

根据各学生历史考试成绩,计算得到所述学生的历史得分率;

根据各学生整体能力得分、历史得分率、以及试题难度得分,采用数据挖掘算法构建得分预测模型;

根据所述得分预测模型对本次考试进行得分预测,得到各学生本次考试的预测得分。

4. 根据权利要求1所述的方法,其特征在于,所述方法还包括:

在得到最终识别结果后,判断所述最终识别结果中的学生属性信息与之前的朗读语音的最终识别结果中的学生属性信息是否相同;

如果相同,并且学生成绩数据库中对应所述学生属性信息的学生成绩已录入,则删除已录入的对应所述学生属性信息的学生成绩,然后将当前最终识别结果中的学生成绩录入到学生成绩数据库中,并对该学生成绩进行异常标记;

否则,执行将所述最终识别结果中的学生成绩录入到学生成绩数据库中的步骤。

5. 根据权利要求1所述的方法,其特征在于,所述方法还包括:

预先根据学生姓名中的生僻字及易错字构建容错发音词集;

检查所述候选识别结果中是否包含所述容错发音词集中的词;

如果是,则对所述朗读语音进行容错识别,得到新的候选识别结果。

6. 根据权利要求1至5任一项所述的方法,其特征在于,所述方法还包括:

对所述最终识别结果进行异常检测,确定异常识别结果;

对录入到学生成绩数据库中的异常识别结果中的学生成绩进行异常标记。

7. 根据权利要求6所述的方法,其特征在于,所述对所述最终识别结果进行异常检测包括:

根据所述最终识别结果的置信度对所述最终识别结果进行异常检测;和/或

根据所述最终识别结果中的学生成绩与对应的学生整体能力得分和/或本次考试的预测得分,对所述最终识别结果进行异常检测。

8. 根据权利要求1至5任一项所述的方法,其特征在于,所述方法还包括:

对学生成绩数据库中异常标记的学生成绩进行异常录入提醒。

9. 根据权利要求8所述的方法,其特征在于,所述异常录入提醒包括:

视觉形式的提醒,和/或听觉形式的提醒。

10. 一种成绩信息录入系统,其特征在于,包括:

语音接收模块,用于接收学生成绩信息的朗读语音,所述学生成绩信息包括:学生成绩和学生属性信息;

语音识别模块,用于对所述朗读语音进行语音识别,得到候选识别结果;

候选识别结果判断模块,用于判断所述候选识别结果是否有多个;

最终识别结果确定模块,用于在所述候选识别结果有多个时,根据预置的学生信息确定最终识别结果,并标记所述最终识别结果,所述学生信息包括:学生名册、以及学生整体能力得分和本次考试各学生的预测得分这两者中其一或全部;否则将所述候选识别结果作为最终识别结果,并且在所述候选识别结果为容错识别得到的候选结果时,标记所述最终识别结果;

录入模块,用于将所述最终识别结果中的学生成绩录入到学生成绩数据库中,并对录入到学生成绩数据库中的已标记最终识别结果中的学生成绩进行异常标记。

11. 根据权利要求10所述的系统,其特征在于,所述系统还包括:

学生信息获取模块,用于获取所述学生信息;所述学生信息获取模块包括:

学生名册获取单元,用于获取学生名册;

整体能力得分及试题难度得分获取单元,用于基于学生历史考试成绩,应用IRT模型计算得到学生整体能力得分及试题难度得分;

历史得分率计算单元,用于根据学生历史考试成绩,计算历史得分率;

模型训练单元,用于根据各学生整体能力得分、历史得分率、以及试题难度得分,采用数据挖掘算法构建得分预测模型;

得分预测单元,用于根据所述得分预测模型对本次考试进行得分预测,得到各学生本次考试的预测得分。

12. 根据权利要求10所述的系统,其特征在于,所述系统还包括:

重复录入处理模块,用于在所述最终识别结果确定模块得到最终识别结果后,判断所述最终识别结果中的学生属性信息与之前的朗读语音的最终识别结果中的学生属性信息是否相同;如果相同,并且学生成绩数据库中对应所述学生属性信息的学生成绩已录入,则触发所述录入模块删除已录入的对应所述学生属性信息的学生成绩,然后将当前最终识别结果中的学生成绩录入到学生成绩数据库中,并对该学生成绩进行异常标记;否则,触发所述录入模块将所述最终识别结果中的学生成绩录入到学生成绩数据库中。

13. 根据权利要求10所述的系统,其特征在于,所述系统还包括:

容错发音词集构建模块,用于预先根据学生姓名中的生僻字及易错字构建容错发音词集;

容错检查模块,用于检查所述候选识别结果中是否包含所述容错发音词集中的词;

容错识别模块,用于在所述候选识别结果中包含所述容错发音词集中的词时,对所述朗读语音进行容错识别,得到新的候选识别结果。

14. 根据权利要求10至13任一项所述的系统,其特征在于,所述系统还包括:

异常检测模块,用于对所述最终识别结果进行异常检测,确定异常识别结果;

异常标记模块,用于对录入到学生成绩数据库中的异常识别结果中的学生成绩进行异常标记。

15. 根据权利要求14所述的系统,其特征在于,所述异常检测模块具体用于:

根据所述最终识别结果的置信度对所述最终识别结果进行异常检测;和/或

根据所述最终识别结果中的学生成绩与对应的学生整体能力得分和/或本次考试的预测得分,对所述最终识别结果进行异常检测。

16. 根据权利要求10至13任一项所述的系统,其特征在于,所述系统还包括:

提醒模块,用于对学生成绩数据库中异常标记的学生成绩进行异常录入提醒。

17. 根据权利要求16所述的系统,其特征在于,

所述提醒模块,具体用于对学生成绩数据库中异常标记的学生成绩进行视觉形式的提醒,和/或听觉形式的提醒。

成绩信息录入方法及系统

技术领域

[0001] 本发明涉及信息处理技术领域,具体涉及一种成绩信息录入方法及系统。

背景技术

[0002] 在国内现有的基础教育阶段,学生学习状况的主要考察形式仍然是各种类型的考试。且随着教育电子信息化的不断发展,在学生进行各类考试后,多需要将学生成绩录入到计算机的数字化考生的表格中,因而教师面临着很大的录入成绩所带来的工作量,而且通过键盘录入的方式也易出错。

发明内容

[0003] 本发明提供一种成绩信息录入方法及系统,以更方便快捷、准确地录入学生成绩,减轻成绩录入的工作量。

[0004] 为此,本发明提供如下技术方案:

[0005] 一种成绩信息录入方法,包括:

[0006] 接收学生成绩信息的朗读语音,所述学生成绩信息包括:学生成绩和学生属性信息;

[0007] 对所述朗读语音进行语音识别,得到候选识别结果;

[0008] 如果所述候选识别结果有多个,则根据预置的学生信息确定最终识别结果,并标记所述最终识别结果,所述学生信息包括:学生名册、以及各学生整体能力得分和本次考试的预测得分这两者中其一或全部;否则将所述候选识别结果作为最终识别结果,并且在所述候选识别结果为容错识别得到的候选结果时,标记所述最终识别结果;

[0009] 将所述最终识别结果中的学生成绩录入到学生成绩数据库中,并对录入到学生成绩数据库中的已标记最终识别结果中的学生成绩进行异常标记。

[0010] 优选地,所述学生属性信息包括以下任意一种或多种:姓名、学号、准考证号。

[0011] 优选地,所述方法还包括:

[0012] 基于各学生历史考试成绩,应用IRT模型计算得到所述学生整体能力得分及试题难度得分;

[0013] 根据各学生历史考试成绩,计算得到所述学生的历史得分率;

[0014] 根据各学生整体能力得分、历史得分率、以及试题难度得分,采用数据挖掘算法构建得分预测模型;

[0015] 根据所述得分预测模型对本次考试进行得分预测,得到各学生本次考试的预测得分。

[0016] 优选地,所述方法还包括:

[0017] 在得到最终识别结果后,判断所述最终识别结果中的学生属性信息与之前的朗读语音的最终识别结果中的学生属性信息是否相同;

[0018] 如果相同,并且学生成绩数据库中对应所述学生属性信息的学生成绩已录入,则

删除已录入的对应所述学生属性信息的学生成绩,然后将当前最终识别结果中的学生成绩录入到学生成绩数据库中,并对该学生成绩进行异常标记;

- [0019] 否则,执行将所述最终识别结果中的学生成绩录入到学生成绩数据库中的步骤。
- [0020] 优选地,所述方法还包括:
 - [0021] 预先根据学生姓名中的生僻字及易错字构建容错发音词集;
 - [0022] 检查所述候选识别结果中是否包含所述容错发音词集中的词;
 - [0023] 如果是,则对所述朗读语音进行容错识别,得到新的候选识别结果。
- [0024] 优选地,所述方法还包括:
 - [0025] 对所述最终识别结果进行异常检测,确定异常识别结果;
 - [0026] 对录入到学生成绩数据库中的异常识别结果中的学生成绩进行异常标记。
- [0027] 优选地,所述对所述最终识别结果进行异常检测包括:
 - [0028] 根据所述最终识别结果的置信度对所述最终识别结果进行异常检测;和/或
 - [0029] 根据所述最终识别结果中的学生成绩与对应的学生整体能力得分和/或本次考试的预测得分,对所述最终识别结果进行异常检测。
- [0030] 优选地,所述方法还包括:
 - [0031] 对学生成绩数据库中异常标记的学生成绩进行异常录入提醒。
 - [0032] 优选地,所述异常录入提醒包括:
 - [0033] 视觉形式的提醒,和/或听觉形式的提醒。
 - [0034] 一种成绩信息录入系统,包括:
 - [0035] 语音接收模块,用于接收学生成绩信息的朗读语音,所述学生成绩信息包括:学生成绩和学生属性信息;
 - [0036] 语音识别模块,用于对所述朗读语音进行语音识别,得到候选识别结果;
 - [0037] 候选识别结果判断模块,用于判断所述候选识别结果是否有多个;
 - [0038] 最终识别结果确定模块,用于在所述候选识别结果有多个时,根据预置的学生信息确定最终识别结果,并标记所述最终识别结果,所述学生信息包括:学生名册、以及学生整体能力得分和本次考试各学生的预测得分这两者中其一或全部;否则将所述候选识别结果作为最终识别结果,并且在所述候选识别结果为容错识别得到的候选结果时,标记所述最终识别结果;
 - [0039] 录入模块,用于将所述最终识别结果中的学生成绩录入到学生成绩数据库中,并对录入到学生成绩数据库中的已标记最终识别结果中的学生成绩进行异常标记。
 - [0040] 优选地,所述系统还包括:
 - [0041] 学生信息获取模块,用于获取所述学生信息;所述学生信息获取模块包括:
 - [0042] 学生名册获取单元,用于获取学生名册;
 - [0043] 整体能力得分及试题难度得分获取单元,用于基于学生历史考试成绩,应用IRT模型计算得到学生整体能力得分及试题难度得分;
 - [0044] 历史得分率计算单元,用于根据学生历史考试成绩,计算历史得分率;
 - [0045] 模型训练单元,用于根据各学生整体能力得分、历史得分率、以及试题难度得分,采用数据挖掘算法构建得分预测模型;
 - [0046] 得分预测单元,用于根据所述得分预测模型对本次考试进行得分预测,得到各学

生本次考试的预测得分。

[0047] 优选地,所述系统还包括:

[0048] 重复录入处理模块,用于在所述最终识别结果确定模块得到最终识别结果后,判断所述最终识别结果中的学生属性信息与之前的朗读语音的最终识别结果中的学生属性信息是否相同;如果相同,并且学生成绩数据库中对应所述学生属性信息的学生成绩已录入,则触发所述录入模块删除已录入的对应所述学生属性信息的学生成绩,然后将当前最终识别结果中的学生成绩录入到学生成绩数据库中,并对该学生成绩进行异常标记;否则,触发所述录入模块将所述最终识别结果中的学生成绩录入到学生成绩数据库中。

[0049] 优选地,所述系统还包括:

[0050] 容错发音词集构建模块,用于预先根据学生姓名中的生僻字及易错字构建容错发音词集;

[0051] 容错检查模块,用于检查所述候选识别结果中是否包含所述容错发音词集中的词;

[0052] 容错识别模块,用于在所述候选识别结果中包含所述容错发音词集中的词时,对所述朗读语音进行容错识别,得到新的候选识别结果。

[0053] 优选地,所述系统还包括:

[0054] 异常检测模块,用于对所述最终识别结果进行异常检测,确定异常识别结果;

[0055] 异常标记模块,用于对录入到学生成绩数据库中的异常识别结果中的学生成绩进行异常标记。

[0056] 优选地,所述异常检测模块具体用于:

[0057] 根据所述最终识别结果的置信度对所述最终识别结果进行异常检测;和/或

[0058] 根据所述最终识别结果中的学生成绩与对应的学生整体能力得分和/或本次考试的预测得分,对所述最终识别结果进行异常检测。

[0059] 优选地,所述系统还包括:

[0060] 提醒模块,用于对学生成绩数据库中异常标记的学生成绩进行异常录入提醒。

[0061] 优选地,所述提醒模块,具体用于对学生成绩数据库中异常标记的学生成绩进行视觉形式的提醒,和/或听觉形式的提醒。

[0062] 本发明实施例提供的成绩信息录入方法及系统,采用语音录入方式,对接收朗读语音进行语音识别,得到候选识别结果,并且在候选识别结果有多个的情况下,根据预置的学生信息确定最终识别结果,将最终识别结果中的学生成绩录入到学生成绩数据库中,并对录入到学生成绩数据库中有错误可能的数据进行异常标记,以提示录入成员。从而大大减轻了成绩录入的工作量,并保证了录入数据的准确性。

附图说明

[0063] 为了更清楚地说明本申请实施例或现有技术中的技术方案,下面将对实施例中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本发明中记载的一些实施例,对于本领域普通技术人员来讲,还可以根据这些附图获得其他的附图。

[0064] 图1是本发明实施例成绩信息录入方法的流程图;

[0065] 图2是本发明实施例成绩信息录入系统的一种结构示意图;

- [0066] 图3是本发明实施例成绩信息录入系统的另一种结构示意图；
[0067] 图4是本发明实施例成绩信息录入系统的另一种结构示意图。

具体实施方式

[0068] 为了使本技术领域的人员更好地理解本发明实施例的方案，下面结合附图和实施方式对本发明实施例作进一步的详细说明。

[0069] 如图1所示，是本发明实施例成绩信息录入方法的流程图，包括以下步骤：

[0070] 步骤101，接收学生成绩信息的朗读语音。

[0071] 所述学生成绩信息不仅包含学生成绩，还包含学生属性信息，所述学生属性信息包括以下任意一种或多种：姓名、学号、准考证号，在有多个班级时，所述学生属性信息还可以包含班级信息等。

[0072] 步骤102，对所述朗读语音进行语音识别，得到候选识别结果。

[0073] 语音识别的具体过程与现有技术相同，即利用预先训练的声学模型及语言模型构建解码网络；提取语音数据的特征参数，比如，线性预测参数 (LPCC)、和/或Mell频率倒谱系数 (MFCC) 参数，然后基于所述解码网络及所述特征参数对所述朗读语音进行解码，得到各解码路径对应的识别文本，即候选识别结果。

[0074] 需要说明的是，在该应用中，语言模型是根据学生名册构建的，因此所述候选识别结果只会出现学生名册中相对应的识别结果。

[0075] 步骤103，判断所述候选识别结果是否多个；如果是，则执行步骤104；否则执行步骤105。

[0076] 步骤104，根据预置的学生信息确定最终识别结果，并标记所述最终识别结果。然后，执行步骤106。

[0077] 所述学生信息包括：学生名册、以及各学生整体能力得分和本次考试的预测得分这两者中其一或全部。

[0078] 其中，学生整体能力得分可以基于学生历史考试成绩，应用IRT (Item Response Theory，项目反映理论) 模型计算得到。

[0079] 所述IRT的基本思想是：对于任一学生，一道试题(即项目)能否正确回答(即反映)的概率，取决于两个方面：一是自身能力，此为主观因素；二是用难度与区分度等指标反映的试题情况，此为客观因素。

[0080] 比如，针对客观题，可以将答题结果标注为0和1，0表示答错，1表示答对，得到答题矩阵作为IRT模型输入，然后可基于极大似然估计算法训练一级IRT模型。针对主观题，可以按答题分数标注为不同等级，如一个题目总分是10分，答题分数在0-3分标注为等级一，4-7分标注为等级二，8-10分标注为等级三，然后采用多级IRT模型进行训练，IRT模型的输出为学生参数 θ (表示学生整体能力得分) 和试题参数 (a, b) ，其中 (a, b) 表示 (试题难度得分, 试题区分度)。

[0081] 在该实施例中，基于学生历史考试成绩得到该学生的答题矩阵，将该答题矩阵作为IRT模型的输入，根据IRT模型的输出即可得到该学生的整体能力得分及试题难度得分。

[0082] 各学生本次考试的预测得分可以按照以下方式获得：

[0083] (1) 基于学生历史考试成绩计算各学生的历史得分率。

[0084] 所述历史得分率具体可以是所有试题的得分之和除以所有试题实际分值之和,也可以是历次考试得分率的均值。

[0085] (2) 根据各学生整体能力得分、历史得分率、以及试题难度得分等,采用数据挖掘算法(如分类/回归算法、协同过滤算法等)构建得分预测模型。

[0086] (3) 根据所述得分预测模型对本次考试进行得分预测,得到各学生本次考试的预测得分。

[0087] 比如得分预测模型是一个分类/回归模型,进行得分预测时,该分类/回归模型的输入是学生特征和试题特征,输出就是学生的答题情况,如答对或答错。

[0088] 具体地,首先根据新的试卷,抽取学生特征与试题特征,学生特征包括:学生整体能力、历史得分率等;试题特征包括:试题难度得分等。然后,根据每个学生做的每道试题形成一条对应的数据,该数据中包括:学生特征,试题特征,试题得分。最后,针对客观题,采用分类思想预测学生在新题上是否答对(即类别为0或1),分类算法可采用决策树、SVM(Support Vector Machine,支持向量机)等;针对主观题,采用回归思想预测学生在新题上的得分,回归算法可采用线性回归、逻辑回归等。

[0089] 前面提到,学生成绩信息的朗读语音中,所述学生成绩信息包括:学生成绩和学生属性信息,所述学生属性信息包括以下任意一种或多种:姓名、学号、准考证号。

[0090] 通常情况下,学号、准考证号可以唯一标识一个学生,而如果用姓名来标识一个学生,有时还会碰到同名同姓、发音相同的姓名等情况。因此,对于同名同姓的情况,在实际应用中,可以借助于其他属性信息,比如,性别、班级、学号等来唯一标识该学生,相应地,在通过语音录入学生成绩信息时,朗读语音中可以包含这些属性信息,这样,可以得到唯一的候选识别结果。

[0091] 当然,为了进一步提高语音录入效率,朗读语音中也可以不包含这些属性信息,这样,根据学生名册,会得到多个候选识别结果。同样,对于学生相同发音的姓名,比如学生“张杨”与学生“张扬”,发音相同,对其朗读语音进行语音识别时,也会出现多个候选识别结果。

[0092] 在本发明实施例中,在出现多个候选识别结果的情况下,可以根据学生整体能力得分和/或本次考试的预测得分来确定最终识别结果。

[0093] 比如,对于朗读语音“zhangyang 99分”,得到候选识别结果如下:

[0094] 张杨99分;

[0095] 张扬99分。

[0096] 根据预置的学生信息可知,学生“张杨”整体能力水平不高,如以往成绩均值在60分左右,而学生“张扬”整体能力水平较高,以往成绩均值在95分左右,则将“张扬99分”作为最终识别结果。

[0097] 同样,也可以基于各学生本次考试的预测得分来确定最终识别结果。

[0098] 当然,为了使最终识别结果更准确,还可以同时基于各学生整体能力得分和本次考试的预测得分这两者来确定最终识别结果。

[0099] 需要说明的是,在有多个候选识别结果时,也可以对这些候选识别结果给出提示,即标记所述最终识别结果,比如高亮标记候选识别结果、或者给出提示信息(比如“请选择”)等,然后由人工选择其中一个作为最终识别结果。

[0100] 步骤105,将所述候选识别结果作为最终识别结果,并且在所述候选识别结果为容错识别得到的候选结果时,标记所述最终识别结果。

[0101] 也就是说,候选识别结果只有一个时,将该候选识别结果作为最终识别结果。而且,如果候选识别结果不是容错识别得到的结果时,无需标记相应的最终识别结果,否则还需要标记该最终识别结果。

[0102] 步骤106,将所述最终识别结果中的学生成绩录入到学生成绩数据库中,并对录入到学生成绩数据库中的已标记最终识别结果中的学生成绩进行异常标记。

[0103] 需要说明的是,在人工朗读学生成绩信息时,可能会出现重复朗读的情况,比如,朗读语音意外中断、或者朗读错误的更正等,因此,在本发明方法另一实施例中,还可以在得到最终识别结果后,进一步判断所述最终识别结果中的学生属性信息与之前的朗读语音的最终识别结果中的学生属性信息是否相同;如果相同,并且学生成绩数据库中对应所述学生属性信息的学生成绩已录入,则删除已录入的对应所述学生属性信息的学生成绩,然后将当前最终识别结果中的学生成绩录入到学生成绩数据库中,也就是说,将后一次语音录入作为最终录入结果,并对该学生成绩进行异常标记;否则,直接将所述最终识别结果中的学生成绩录入到学生成绩数据库中。

[0104] 另外,学生姓名中可能会存在生僻字、易错字等,针对这种情况,可以收集这些字的易错发音(如“栗”与“粟”、“弋”与“戈”容易混淆读错),构建容错发音词集,将正常语音识别得到的候选识别结果作为预选识别结果,然后检查所述预选识别结果中是否包含所述容错发音词集中的词;如果是,则对所述朗读语音进行容错识别,得到新的候选识别结果。

[0105] 所述容错识别可以是在语言模型中包含容错发音词集信息,在识别解码的过程中,即可以得到容错的解码路径。

[0106] 进一步地,还可以对用户发音进行纠错,以引导用户下次发音正确。如班级有个学生叫“栗冰”,成绩录入时,读的是“subing”,通过容错识别,得到最终识别结果为“栗冰”,此时可以提醒用户发音错误,并给出正确发音提示:“libing”。

[0107] 为了进一步保证成绩录入的正确性,在本发明方法另一实施例中,还可以对所述最终识别结果进行异常检测,以确定异常识别结果,并对录入到学生成绩数据库中的异常识别结果中的学生成绩进行异常标记。在进行异常检测时,可以根据语音识别置信度,还可以结合学生整体能力及本次考试各学生的预测得分等来进行。比如,语音识别结果的置信度小于一定阈值,或者语音识别结果中的成绩与预测的成绩或者相对应的学生整体能力相差大于一定阈值时,确定所述语音识别结果异常。

[0108] 对于异常识别结果,仍需将其包含的学生成绩录入到学生成绩数据库中,但需要在学生成绩数据库中对该学生成绩进行异常标记,以便后续对这些异常数据做进一步的处理。

[0109] 比如,可以对学生成绩数据库中异常标记的学生成绩进行异常录入提醒,异常录入的提醒可以是在整个成绩录入结束后进行,也可以是在录入过程中提醒,对此本发明实施例不做限定。而且,具体的提醒方式可以是视觉形式的提醒、和/或听觉形式的提醒。

[0110] 在实际应用中,可以由人工对这些提醒的相应数据进行校对、纠正,以进一步保证录入成绩的正确性。

[0111] 本发明实施例提供的成绩信息录入方法,采用语音录入方式,对接收朗读语音进

行语音识别,得到候选识别结果,并且在候选识别结果有多个的情况下,根据预置的学生信息确定最终识别结果,将最终识别结果中的学生成绩录入到学生成绩数据库中,并对录入到学生成绩数据库中有错误可能的数据进行异常标记,以提示录入人员。从而大大减轻了成绩录入的工作量,并保证了录入数据的准确性。

[0112] 相应地,本发明实施例还提供一种成绩信息录入系统,如图2所示,是该系统的一种结构示意图。

[0113] 在该实施例中,所述系统包括:

[0114] 语音接收模块201,用于接收学生成绩信息的朗读语音,所述学生成绩信息包括:学生成绩和学生属性信息;

[0115] 语音识别模块202,用于对所述朗读语音进行语音识别,得到候选识别结果;

[0116] 候选识别结果判断模块203,用于判断所述候选识别结果是否多个;

[0117] 最终识别结果确定模块204,用于在所述候选识别结果有多个时,根据预置的学生信息确定最终识别结果,并标记所述最终识别结果,所述学生信息包括:学生名册、以及学生整体能力得分和本次考试各学生的预测得分这两者中其一或全部;否则将所述候选识别结果作为最终识别结果,并且在所述候选识别结果为容错识别得到的候选结果时,标记所述最终识别结果;

[0118] 录入模块205,用于将所述最终识别结果中的学生成绩录入到学生成绩数据库中,并对录入到学生成绩数据库中的已标记最终识别结果中的学生成绩进行异常标记。

[0119] 需要说明的是,在本发明实施例中,所述学生成绩信息不仅包含学生成绩,还包含学生属性信息,所述学生属性信息包括以下任意一种或多种:姓名、学号、准考证号,在有多个班级时,所述学生属性信息还可以包含班级信息等。所述学生信息包括:学生名册、以及学生整体能力得分和本次考试各学生的预测得分这两者中其一或全部。其中,学生信息可以由其它独立系统预先基于学生历史考试成绩来获得,也可以由本发明系统预先获得,比如在本发明系统中设置学生信息获取模块来获取所述学生信息。所述学生信息获取模块的一种具体结构可以包括以下各单元:

[0120] 学生信息获取模块,用于获取所述学生信息;所述学生信息获取模块包括:

[0121] 学生名册获取单元,用于获取学生名册;

[0122] 整体能力得分及试题难度得分获取单元,用于基于学生历史考试成绩,应用IRT模型计算得到学生整体能力得分及试题难度得分;

[0123] 历史得分率计算单元,用于根据学生历史考试成绩,计算历史得分率;

[0124] 模型训练单元,用于根据各学生整体能力得分、历史得分率、以及试题难度得分,采用数据挖掘算法构建得分预测模型;

[0125] 得分预测单元,用于根据所述得分预测模型对本次考试进行得分预测,得到各学生本次考试的预测得分。

[0126] 上述各单元获取相应参数的具体过程可参见前面本发明方法实施例中的描述,在此不再赘述。

[0127] 当然,在根据预置的学生信息确定最终识别结果时,如果仅需要学生整体能力得分和本次考试各学生的预测得分这两者中的一种得分时,上述学生信息获取模块中的一些单元需要进行适应性地调整,对此本发明实施例不做限定。

[0128] 在人工朗读学生成绩信息时,可能会出现重复朗读的情况,比如,朗读语音意外中断、或者朗读错误的更正等,因此,在本发明系统另一实施例中,如图3所示,所述系统还包括:重复录入处理模块301,用于在所述最终识别结果确定模块204得到最终识别结果后,判断所述最终识别结果中的学生属性信息与之前的朗读语音的最终识别结果中的学生属性信息是否相同;如果相同,并且学生成绩数据库中对应所述学生属性信息的学生成绩已录入,则触发所述录入模块205删除已录入的对应所述学生属性信息的学生成绩,然后将当前最终识别结果中的学生成绩录入到学生成绩数据库中,并对该学生成绩进行异常标记;否则,触发所述录入模块205将所述最终识别结果中的学生成绩录入到学生成绩数据库中。

[0129] 另外,学生姓名中可能会存在生僻字、易错字等,针对这种情况,在本发明系统另一实施例中,可以收集这些字的易错发音,构建容错发音词集,将正常语音识别得到的候选识别结果作为预选识别结果,然后检查所述预选识别结果中是否包含所述容错发音词集中的词;如果是,则对所述朗读语音进行容错识别,得到新的候选识别结果。相应地,该系统实施例的一种结构如图4所示。

[0130] 与图2所示实施例相比,在该实施例中,所述系统还包括:

[0131] 容错发音词集构建模块401,用于预先根据学生姓名中的生僻字及易错字构建容错发音词集;

[0132] 容错检查模块402,用于检查所述候选识别结果中是否包含所述容错发音词集中的词;

[0133] 容错识别模块403,用于在所述候选识别结果中包含所述容错发音词集中的词时,对所述朗读语音进行容错识别,得到新的候选识别结果。

[0134] 在该实施例中,如果语音识别模块202得到的候选识别结果中如果不包含所述容错发音词集中的词,则候选识别结果判断模块203直接对语音识别模块202得到的候选识别结果进行判断,否则候选识别结果判断模块203需要对容错识别模块403得到的新的候选识别结果进行判断。

[0135] 当然,在该实施例中,同样也可以包括上述重复录入处理模块301。

[0136] 为了进一步保证成绩录入的正确性,在本发明系统上述各实施例中,还可包括:

[0137] 异常检测模块,用于对所述最终识别结果进行异常检测,确定异常识别结果,比如根据所述最终识别结果的置信度对所述最终识别结果进行异常检测;和/或根据所述最终识别结果中的学生成绩与对应的学生整体能力得分和/或本次考试的预测得分,对所述最终识别结果进行异常检测;

[0138] 异常标记模块,用于对录入到学生成绩数据库中的异常识别结果中的学生成绩进行异常标记。

[0139] 另外,所述系统还可进一步包括:提醒模块,用于对学生成绩数据库中异常标记的学生成绩进行异常录入提醒,比如采用视觉形式的提醒,和/或听觉形式的提醒。这样,可以有效提示录入人员,由人工对这些提醒的相应数据进行校对、纠正,以进一步保证录入成绩的正确性。

[0140] 需要说明的是,在实际应用中,上述各模块可以根据应用场景需要来选用,对此本发明实施例不做限定。

[0141] 本发明实施例提供的成绩信息录入系统,采用语音录入方式,对接收朗读语音进

行语音识别,得到候选识别结果,并且在候选识别结果有多个的情况下,根据预置的学生信息确定最终识别结果,将最终识别结果中的学生成绩录入到学生成绩数据库中,并对录入到学生成绩数据库中有错误可能的数据进行异常标记,以提示录入成员。从而大大减轻了成绩录入的工作量,并保证了录入数据的准确性。

[0142] 本说明书中的各个实施例均采用递进的方式描述,各个实施例之间相同相似的部分互相参见即可,每个实施例重点说明的都是与其他实施例的不同之处。尤其,对于系统实施例而言,由于其基本相似于方法实施例,所以描述得比较简单,相关之处参见方法实施例的部分说明即可。以上所描述的系统实施例仅仅是示意性的,其中所述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网络单元上。可以根据实际的需要选择其中的部分或者全部模块来实现本实施例方案的目的。本领域普通技术人员在不付出创造性劳动的情况下,即可以理解并实施。

[0143] 以上对本发明实施例进行了详细介绍,本文中应用了具体实施方式对本发明进行了阐述,以上实施例的说明只是用于帮助理解本发明的方法及系统;同时,对于本领域的一般技术人员,依据本发明的思想,在具体实施方式及应用范围上均会有改变之处,综上所述,本说明书内容不应理解为对本发明的限制。

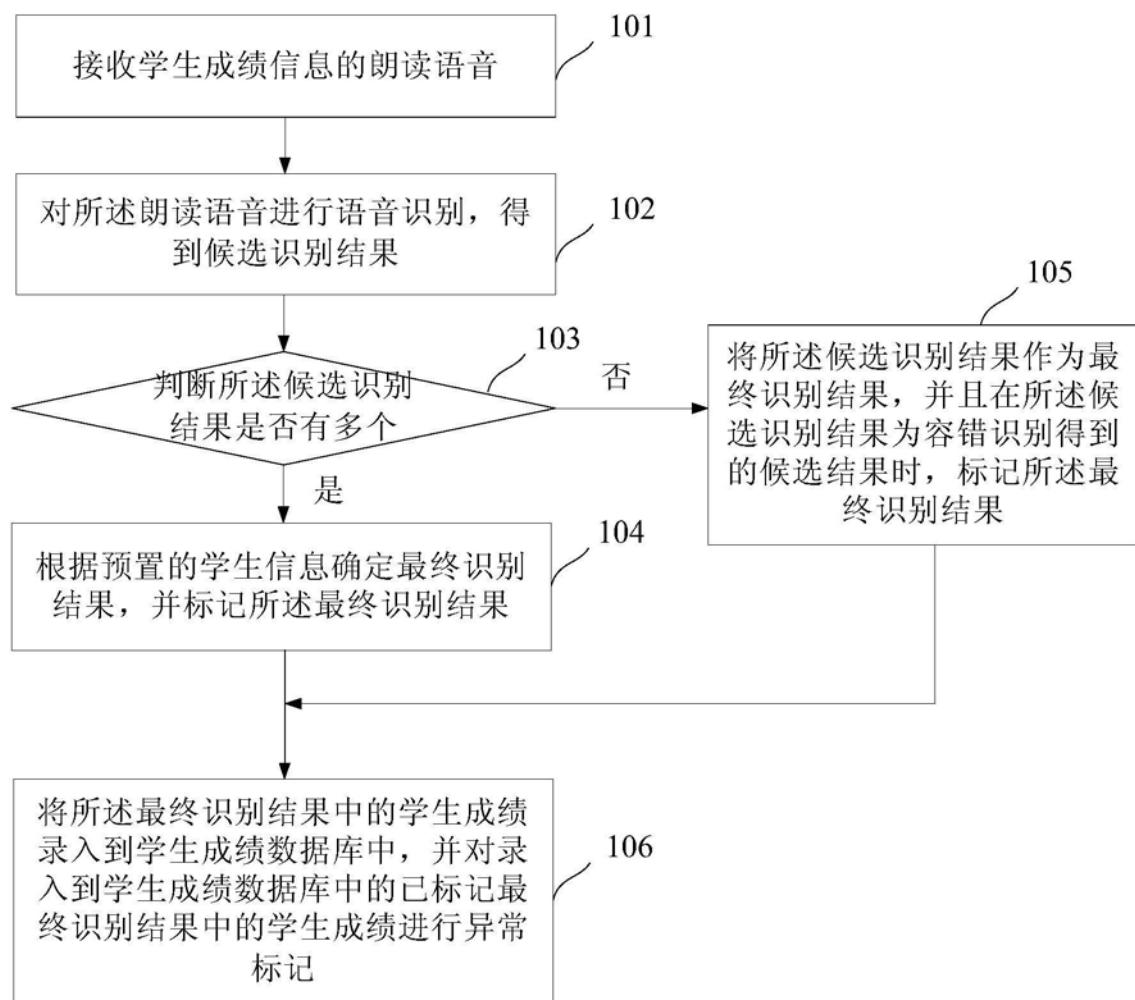


图1

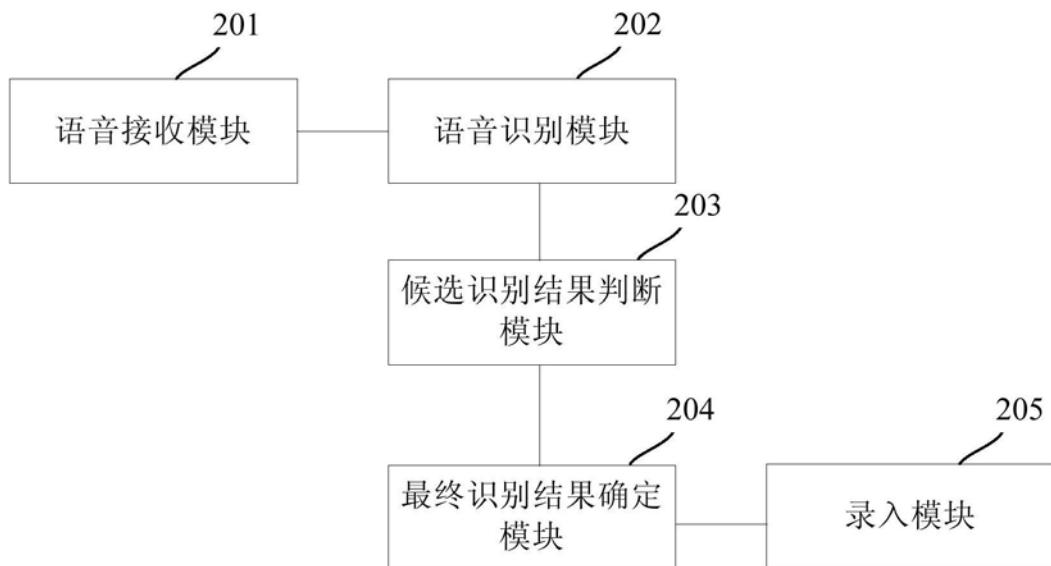


图2

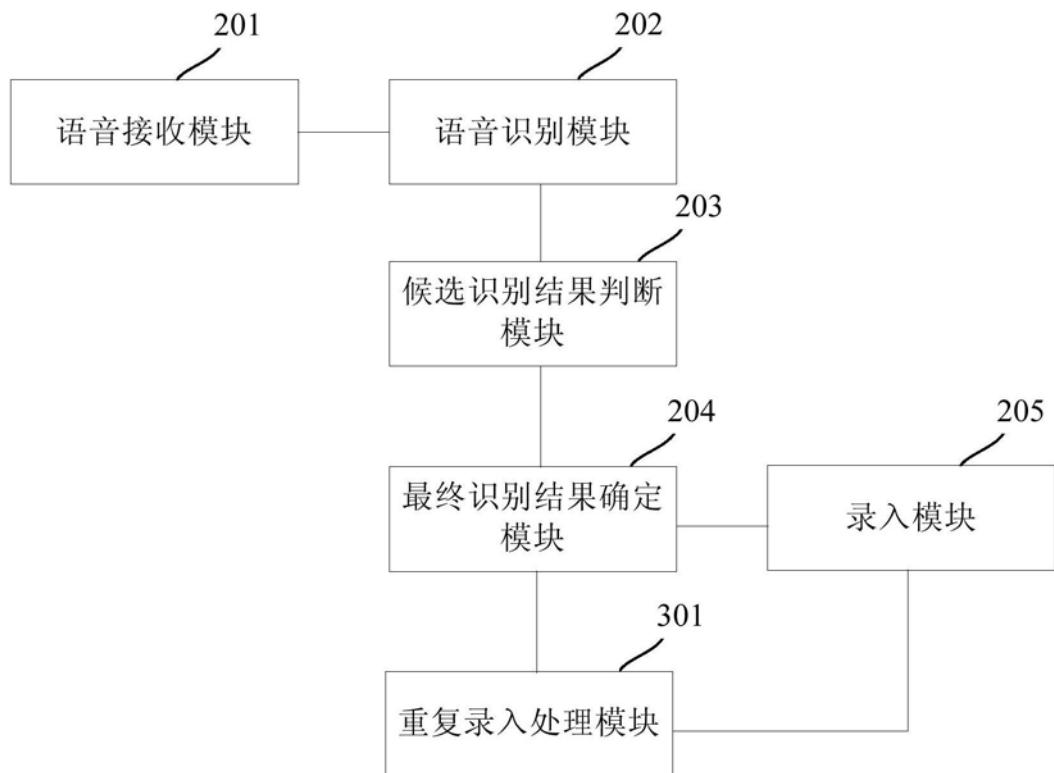


图3

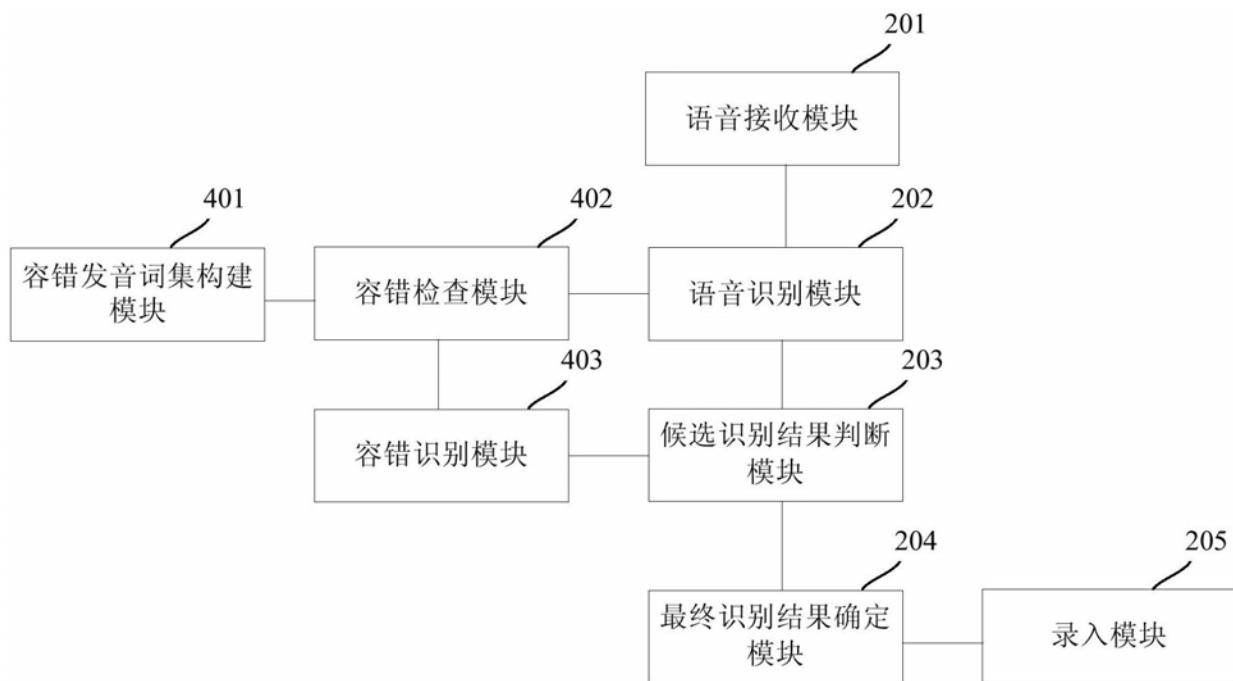


图4