



(12) 发明专利申请

(10) 申请公布号 CN 114203169 A

(43) 申请公布日 2022.03.18

(21) 申请号 202210094152.7

(22) 申请日 2022.01.26

(71) 申请人 合肥讯飞数码科技有限公司

地址 230088 安徽省合肥市中国(安徽)自由贸易试验区合肥市高新区望江西路666号人工智能云服务平台研发楼

(72) 发明人 桑宏报 方磊 杨军 何雪 方四安 柳林

(74) 专利代理机构 北京集佳知识产权代理有限公司 11227

代理人 王云晓

(51) Int. Cl.

G10L 15/08 (2006.01)

G10L 25/72 (2013.01)

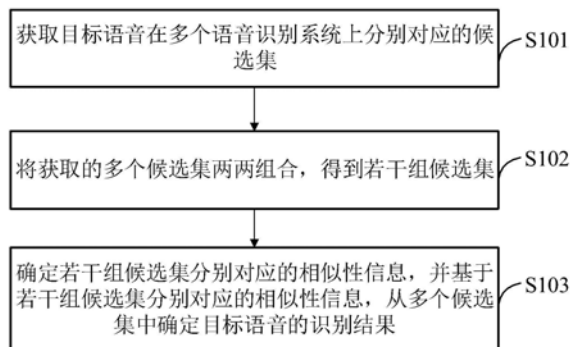
权利要求书3页 说明书16页 附图3页

(54) 发明名称

一种语音识别结果确定方法、装置、设备及存储介质

(57) 摘要

本申请提供了一种语音识别结果确定方法、装置、设备及存储介质,其中,方法包括:获取目标语音在多个语音识别系统上分别对应的候选集,其中,目标语音在一个语音识别系统上对应的候选集包括基于该语音识别系统获得的若干候选识别结果;将获取的多个候选集两两组合,以得到若干组候选集;确定若干组候选集分别对应的相似性信息,并基于若干组候选集分别对应的相似性信息,从多个候选集中确定目标语音的识别结果,其中,相似性信息为对应的两个候选集中的第一候选集包含的候选识别结果与第二候选集包含的候选识别结果的句子级别的相似性信息。通过本申请提供的语音识别结果确定方法能够较快地确定出准确度较高的语音识别结果。



1. 一种语音识别结果确定方法,其特征在于,包括:

获取目标语音在多个语音识别系统上分别对应的候选集,其中,所述目标语音在一个语音识别系统上对应的候选集包括基于该语音识别系统获得的若干候选识别结果;

将获取的多个候选集两两组合,以得到若干组候选集;

确定所述若干组候选集分别对应的相似性信息,并基于所述若干组候选集分别对应的相似性信息,从所述多个候选集中确定目标语音的识别结果,其中,所述相似性信息为对应的两个候选集中的第一候选集包含的候选识别结果与第二候选集包含的候选识别结果的句子级别的相似性信息。

2. 根据权利要求1所述的语音识别结果确定方法,其特征在于,获取目标语音在一个语音识别系统上对应的候选集,包括:

获取该语音识别系统对目标语音进行识别后输出的多个候选识别结果;

确定所述多个候选识别结果中是否存在解码异常的候选识别结果;

若是,则将解码异常的候选识别结果剔除,由剩余的候选识别结果组成所述目标语音在该语音识别系统上对应的候选集;

若否,则由所述多个候选识别结果组成所述目标语音在该语音识别系统上对应的候选集。

3. 根据权利要求2所述的语音识别结果确定方法,其特征在于,所述确定所述多个候选识别结果中是否存在解码异常的候选识别结果,包括:

确定所述多个候选识别结果中是否存在每秒解码词数大于预设词数阈值的候选识别结果;

若是,则确定所述多个候选识别结果中存在解码异常的候选识别结果;

若否,则确定所述多个候选识别结果中不存在解码异常的候选识别结果;

所述将解码异常的候选识别结果剔除,包括:

将每秒解码词数大于所述预设词数阈值的候选识别结果剔除。

4. 根据权利要求1~3中任一项所述的语音识别结果确定方法,其特征在于,所述基于所述若干组候选集分别对应的相似性信息,从所述多个候选集中确定目标语音的识别结果,包括:

基于所述若干组候选集分别对应的相似性信息,确定所述若干组候选集分别对应的候选识别结果对,以得到若干候选识别结果对,其中,所述候选识别结果对包括对应的两个候选集中最相似的两个候选识别结果;

基于所述若干候选识别结果对分别对应的相似性信息,确定所述目标语音的识别结果,其中,一候选识别结果对对应的相似性信息为该候选识别结果对包含的两个候选识别结果的相似性信息。

5. 根据权利要求4所述的语音识别结果确定方法,其特征在于,所述相似性信息为相似性度量指标;

所述基于所述若干候选识别结果对分别对应的相似性信息,确定所述目标语音的识别结果,包括:

基于所述若干候选识别结果对分别对应的相似性度量指标,从所述若干候选识别结果对中确定最相似的候选识别结果对,作为目标候选识别结果对;

至少基于所述目标候选识别结果对对应的相似性度量指标,确定所述目标语音的识别结果。

6. 根据权利要求5所述的语音识别结果确定方法,其特征在于,所述至少基于所述目标候选识别结果对对应的相似性度量指标,确定所述目标语音的识别结果,包括:

若所述目标候选识别结果对对应的相似性度量指标为表征两个候选识别结果相同的目标值,则将所述目标候选识别结果对中的任一候选识别结果确定为所述目标语音的识别结果;

若所述目标候选识别结果对对应的相似性度量指标不为所述目标值,则基于所述目标候选识别结果对对应的相似性度量指标以及所述目标候选识别结果对中每个候选识别结果的词数和综合得分,确定所述目标语音的识别结果,其中,一候选识别结果的综合得分基于该候选识别结果的声学得分和语言得分确定。

7. 根据权利要求6所述的语音识别结果确定方法,其特征在于,确定一候选识别结果的综合得分,包括:

根据该候选识别结果的词数和该候选识别结果的语言得分,确定该候选识别结果的合理程度表征值;

将该候选识别结果的合理程度表征值归一化为目标范围内的值,以得到归一化后的合理程度表征值,其中,该候选识别结果的声学得分处于所述目标范围内;

将所述归一化后的合理程度表征值与该候选识别结果的声学得分融合,融合后的得分作为该候选识别结果的综合得分。

8. 根据权利要求6所述的语音识别结果确定方法,其特征在于,所述基于所述目标候选识别结果对对应的相似性度量指标以及所述目标候选识别结果对中每个候选识别结果的词数和综合得分,确定所述目标语音的识别结果,包括:

在所述目标候选识别结果对的相似性度量指标与预设相似性阈值满足预设关系的情况下:

若所述目标候选识别结果对中第一候选识别结果的词数与第二候选识别结果的词数相同,则将所述目标候选识别结果对中综合得分大的候选识别结果确定为所述目标语音的识别结果;

若所述第一候选识别结果的词数大于所述第二候选识别结果的词数,且所述第二候选识别结果的综合得分大于所述第一候选识别结果的综合得分,则将所述第二候选识别结果确定为所述目标语音的识别结果;

若所述第二候选识别结果的词数大于所述第一候选识别结果的词数,且所述第一候选识别结果的综合得分大于所述第二候选识别结果的综合得分,则将所述第一候选识别结果确定为所述目标语音的识别结果。

9. 根据权利要求8所述的语音识别结果确定方法,其特征在于,还包括:

若基于所述目标候选识别结果对的相似性度量指标以及所述目标候选识别结果对中每个候选识别结果的词数和综合得分,确定无法从所述目标候选识别结果对中确定所述目标语音的识别结果,则从每个候选集中获取识别得分最高的候选识别结果;

将从各个候选集中分别获取的识别得分最高的候选识别结果中综合得分最高的候选识别结果,确定为所述目标语音的识别结果。

10. 一种语音识别结果确定装置,其特征在于,包括:候选集获取模块、候选集组合模块、相似性信息确定模块和识别结果确定模块;

所述候选集获取模块,用于获取目标语音在多个语音识别系统上分别对应的候选集,其中,所述目标语音在一个语音识别系统上对应的候选集包括基于该语音识别系统获得的若干候选识别结果;

所述候选集组合模块,用于将获取的多个候选集两两组合,以得到若干组候选集;

所述相似性信息确定模块,用于确定所述若干组候选集分别对应的相似性信息,其中,所述相似性信息为对应的两个候选集中的第一候选集包含的候选识别结果与第二候选集包含的候选识别结果的句子级别的相似性信息;

所述识别结果确定模块,用于基于所述若干组候选集分别对应的相似性信息,从所述多个候选集中确定目标语音的识别结果。

11. 一种语音识别结果确定设备,其特征在于,包括:存储器和处理器;

所述存储器,用于存储程序;

所述处理器,用于执行所述程序,实现如权利要求1~9中任一项所述的语音识别结果确定方法的各个步骤。

12. 一种计算机可读存储介质,其上存储有计算机程序,其特征在于,所述计算机程序被处理器执行时,实现如权利要求1~9中任一项所述的语音识别结果确定方法的各个步骤。

## 一种语音识别结果确定方法、装置、设备及存储介质

### 技术领域

[0001] 本申请涉及语音识别技术领域,尤其涉及一种语音识别结果确定方法、装置、设备及存储介质。

### 背景技术

[0002] 语音识别是一种利用机器自动地将人类语音内容转换成文本的技术。近年来,随着深度神经网络的广泛应用,语音识别技术取得了长足的进步。

[0003] 目前的确定语音识别结果的方案为:基于一个语音识别系统对目标语音进行识别,以得到若干候选识别结果,然后将若干候选识别结果中识别得分最高的候选识别结果,确定为目标语音的识别结果。

[0004] 然而,受某些因素影响(比如复杂的口音、方言、远讲、噪声、混响环境、语义理解等),语音识别系统的识别性能不稳定,这导致上述方案稳定性不足,有时无法获得准确度较高的识别结果。

### 发明内容

[0005] 有鉴于此,本申请提供了一种语音识别结果确定方法、装置、设备及存储介质,用以解决现有的语音识别结果确定方法稳定性不足,有时无法获得准确度较高的识别结果的问题,其技术方案如下:

[0006] 一种语音识别结果确定方法,包括:

[0007] 获取目标语音在多个语音识别系统上分别对应的候选集,其中,所述目标语音在一个语音识别系统上对应的候选集包括基于该语音识别系统获得的若干候选识别结果;

[0008] 将获取的多个候选集两两组合,以得到若干组候选集;

[0009] 确定若干组候选集分别对应的相似性信息,并基于所述若干组候选集分别对应的相似性信息,从所述多个候选集中确定目标语音的识别结果,其中,所述相似性信息为对应的两个候选集中的第一候选集包含的候选识别结果与第二候选集包含的候选识别结果的句子级别的相似性信息。

[0010] 可选的,获取目标语音在一个语音识别系统上对应的候选集,包括:

[0011] 获取该语音识别系统对目标语音进行识别后输出的多个候选识别结果;

[0012] 确定所述多个候选识别结果中是否存在解码异常的候选识别结果;

[0013] 若是,则将解码异常的候选识别结果剔除,由剩余的候选识别结果组成所述目标语音在该语音识别系统上对应的候选集;

[0014] 若否,则由所述多个候选识别结果组成所述目标语音在该语音识别系统上对应的候选集。

[0015] 可选的,所述确定所述多个候选识别结果中是否存在解码异常的候选识别结果,包括:

[0016] 确定所述多个候选识别结果中是否存在每秒解码词数大于预设词数阈值的候选

识别结果；

[0017] 若是，则确定所述多个候选识别结果中存在解码异常的候选识别结果；

[0018] 若否，则确定所述多个候选识别结果中不存在解码异常的候选识别结果；

[0019] 所述将解码异常的候选识别结果剔除，包括：

[0020] 将每秒解码词数大于所述预设词数阈值的候选识别结果剔除。

[0021] 可选的，所述基于所述若干组候选集分别对应的相似性信息，从所述多个候选集中确定目标语音的识别结果，包括：

[0022] 基于所述若干组候选集分别对应的相似性信息，确定所述若干组候选集分别对应的候选识别结果对，以得到若干候选识别结果对，其中，所述候选识别结果对包括对应的两个候选集中最相似的两个候选识别结果；

[0023] 基于所述若干候选识别结果对分别对应的相似性信息，确定所述目标语音的识别结果，其中，一候选识别结果对对应的相似性信息为该候选识别结果对包含的两个候选识别结果的相似性信息。

[0024] 可选的，所述相似性信息为相似性度量指标；

[0025] 所述基于所述若干候选识别结果对分别对应的相似性信息，确定所述目标语音的识别结果，包括：

[0026] 基于所述若干候选识别结果对分别对应的相似性度量指标，从所述若干候选识别结果对中确定最相似的候选识别结果对，作为目标候选识别结果对；

[0027] 至少基于所述目标候选识别结果对对应的相似性度量指标，确定所述目标语音的识别结果。

[0028] 可选的，所述至少基于所述目标候选识别结果对对应的相似性度量指标，确定所述目标语音的识别结果，包括：

[0029] 若所述目标候选识别结果对对应的相似性度量指标为表征两个候选识别结果相同的目标值，则将所述目标候选识别结果对中的任一候选识别结果确定为所述目标语音的识别结果；

[0030] 若所述目标候选识别结果对对应的相似性度量指标不为所述目标值，则基于所述目标候选识别结果对对应的相似性度量指标以及所述目标候选识别结果对中每个候选识别结果的词数和综合得分，确定所述目标语音的识别结果，其中，一候选识别结果的综合得分基于该候选识别结果的声学得分和语言得分确定。

[0031] 可选的，确定一候选识别结果的综合得分，包括：

[0032] 根据该候选识别结果的词数和该候选识别结果的语言得分，确定该候选识别结果的合理程度表征值；

[0033] 将该候选识别结果的合理程度表征值归一化为目标范围内的值，以得到归一化后的合理程度表征值，其中，该候选识别结果的声学得分处于所述目标范围内；

[0034] 将所述归一化后的合理程度表征值与该候选识别结果的声学得分融合，融合后的得分作为该候选识别结果的综合得分。

[0035] 可选的，所述基于所述目标候选识别结果对对应的相似性度量指标以及所述目标候选识别结果对中每个候选识别结果的词数和综合得分，确定所述目标语音的识别结果，包括：

[0036] 在所述目标候选识别结果对的相似性度量指标与预设相似性阈值满足预设关系的情况下：

[0037] 若所述目标候选识别结果对中第一候选识别结果的词数与第二候选识别结果的词数相同，则将所述目标候选识别结果对中综合得分大的候选识别结果确定为所述目标语音的识别结果；

[0038] 若所述第一候选识别结果的词数大于所述第二候选识别结果的词数，且所述第二候选识别结果的综合得分大于所述第一候选识别结果的综合得分，则将所述第二候选识别结果确定为所述目标语音的识别结果；

[0039] 若所述第二候选识别结果的词数大于所述第一候选识别结果的词数，且所述第一候选识别结果的综合得分大于所述第二候选识别结果的综合得分，则将所述第一候选识别结果确定为所述目标语音的识别结果。

[0040] 可选的，所述语音识别结果确定方法还包括：

[0041] 若基于所述目标候选识别结果对的相似性度量指标以及所述目标候选识别结果对中每个候选识别结果的词数和综合得分，确定无法从所述目标候选识别结果对中确定所述目标语音的识别结果，则从每个候选集中获取识别得分最高的候选识别结果；

[0042] 将从各个候选集中分别获取的识别得分最高的候选识别结果中综合得分最高的候选识别结果，确定为所述目标语音的识别结果。

[0043] 一种语音识别结果确定装置，包括：候选集获取模块、候选集组合模块、相似性信息确定模块和识别结果确定模块；

[0044] 所述候选集获取模块，用于获取目标语音在多个语音识别系统上分别对应的候选集，其中，所述目标语音在一个语音识别系统上对应的候选集包括基于该语音识别系统获得的若干候选识别结果；

[0045] 所述候选集组合模块，用于将获取的多个候选集两两组合，以得到若干组候选集；

[0046] 所述相似性信息确定模块，用于确定所述若干组候选集分别对应的相似性信息，其中，所述相似性信息为对应的两个候选集中的第一候选集包含的候选识别结果与第二候选集包含的候选识别结果的句子级别的相似性信息；

[0047] 所述识别结果确定模块，用于基于所述若干组候选集分别对应的相似性信息，从所述多个候选集中确定目标语音的识别结果。

[0048] 一种语音识别结果确定设备，包括：存储器和处理器；

[0049] 所述存储器，用于存储程序；

[0050] 所述处理器，用于执行所述程序，实现上述第1~9项中任一项所述的语音识别结果确定方法的各个步骤。

[0051] 一种计算机可读存储介质，其上存储有计算机程序，所述计算机程序被处理器执行时，实现上述第1~9项中任一项所述的语音识别结果确定方法的各个步骤。

[0052] 本申请提供的语音识别结果确定方法、装置、设备及存储介质，首先获取目标语音在多个语音识别系统上分别对应的候选集，然后将获取的多个候选集两两组合，以得到若干组候选集，接着确定若干组候选集分别对应的相似性信息，最后基于若干组候选集分别对应的相似性信息，从多个候选集中确定目标语音的识别结果。考虑到多个语音识别系统具有差异性和互补性，本申请从目标语音在多个语音识别系统上分别对应的候选集中确定

目标语音的识别结果,相比于从基于单个语音识别系统获得的候选识别结果中确定目标语音的识别结果,本申请提供的语音识别结果确定方法更加稳定,且能够提升语音识别结果的准确率,在从目标语音在多个语音识别系统上分别对应的候选集中确定目标语音的识别结果时,为了实现不同语音识别系统的候选识别结果互补,同时为了能够更快的确定出语音识别结果,本申请将目标语音在多个语音识别系统上分别对应的候选集两两组合,基于各组候选集分别对应的句子级别的相似性信息,从多个目标语音在多个语音识别系统上分别对应的候选集中确定目标语音的识别结果。经由本申请提供的语音识别结果确定方法能够较快地确定出准确率较高的语音识别结果。

### 附图说明

[0053] 为了更清楚地说明本发明实施例或现有技术中的技术方案,下面将对实施例或现有技术描述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的实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据提供的附图获得其他的附图。

[0054] 图1为本申请实施例提供的语音识别结果确定方法的流程示意图;

[0055] 图2为本申请实施例提供的基于若干组候选集分别对应的相似性信息,从多个候选集中确定目标语音的识别结果的流程示意图;

[0056] 图3为本申请实施例提供的基于若干候选识别结果对分别对应的相似性信息,确定目标语音的识别结果的流程示意图;

[0057] 图4为本申请实施例提供的语音识别结果确定装置的结构示意图;

[0058] 图5为本申请实施例提供的语音识别结果确定设备的结构示意图。

### 具体实施方式

[0059] 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,都属于本发明保护的范围。

[0060] 为了解决目前的语音识别结果确定方案不稳定,有时无法获得准确度较高的识别结果的问题,申请人想到,对于同一条语音,不同语音识别系统输出的识别结果有较大的差异,在多个语音识别系统的识别结果具有互补性(比如,A系统不稳定,没输出优秀的识别结果,而B系统比较稳定,输出了优秀的识别结果)时,将这些系统的识别结果进行融合,能够提高语音识别结果的准确率,沿着上述思路,申请人进行了研究,起初,申请人想到了如下的语音识别结果确定方法:

[0061] 为不同语音识别系统的输出建立一个线性的词转移网络WTN,将WTN按照词错误率WER从小到大排序,并选择最小WER对应的WTN作为基准WTN,运用基准WTN不断地将剩余的WTN进行对齐与合并,直到所有的WTN都合并入基准WTN为止,完成所有WTN的对齐与合并后,使用一种投票方案,比如词频、置信度等,对WTN的每个分支点进行评估,然后挑选最高分的词(拥有最高的票数)作为最终的识别结果。

[0062] 申请人对上述方法进行研究发现,上述方法存在一些问题,主要体现在:其一,上



述方法是在词级别上进行的,需要进行复杂繁琐的对齐、合并、投票、重新解码、词序连接等操作,这些操作需要耗费较长的时间,因此,语音识别结果的确定效率较低,如果融合系统数量增多,将进一步降低语音识别结果的确定效率,这对于响应速度要求较高的应用场景是非常不利的;其二,上述方法只能应用于不同语音识别系统输出的1-Best候选结果,即最终的语音识别结果从各语音识别系统输出的1-Best候选结果中确定,然而,在某些时候,更优秀的识别结果可能会出现在N-Best中除1-Best外的其它结果中,在这种情况下,上述方法无法获得准确度较高的识别结果。

[0063] 针对上述方法存在的问题,申请人继续进行研究,通过不断研究,最终提出了一种效果较好的语音识别结果确定方法,该方法比较稳定,能够比较快速地确定出准确的语音识别结果,该方法可应用于具有数据处理能力的电子设备,该电子设备可以为用户侧使用的终端,比如智能手机、PC、笔记本、PAD、智能家电、车载终端等等,该电子设备还可以为网络侧的服务器,服务器可以是一台服务器,也可以是由多台服务器组成的服务器集群,或者是一个云计算服务器中心,电子设备可按本申请提供的语音识别结果确定方法快速地确定出准确的语音识别结果。

[0064] 本领域技术人员应能理解上述电子设备仅为举例,其它现有的或今后可能出现的电子设备或服务器如可适用于本申请,也应包含在本申请保护范围以内,并在此以引用方式包含于此。

[0065] 接下来通过下述实施例对本申请提供的语音识别结果确定方法进行介绍。

[0066] 第一实施例

[0067] 请参阅图1,示出了本申请实施例提供的语音识别结果确定方法的流程示意图,可以包括:

[0068] 步骤S101:获取目标语音在多个语音识别系统上分别对应的候选集。

[0069] 其中,目标语音为待确定语音识别结果的语音,目标语音在一语音识别系统上对应的候选集包括基于该语音识别系统获得的若干候选识别结果。

[0070] 经由步骤S101可获得多个候选集,示例性的,语音识别系统为3个,则经由步骤S101可获得3个候选集。

[0071] 由于获取各个候选集的实现方式相同,本实施例以获取一个候选集 $R_i$ (目标语音在第 $i$ 个语音识别系统上对应的候选集)为例,对获取目标语音在第 $i$ 个语音识别系统 $S_i$ 上对应的候选集 $R_i$ 的实现方式进行介绍。

[0072] 获取目标语音在第 $i$ 个语音识别系统 $S_i$ 上对应的候选集 $R_i$ 的实现方式有多种,本实施例提供如下两种可选实现方式:

[0073] 第一种实现方式:获取语音识别系统 $S_i$ 对目标语音进行识别后输出的多个候选识别结果,将语音识别系统 $S_i$ 输出的多个候选识别结果组成的候选集作为目标语音在语音识别系统 $S_i$ 上对应的候选集 $R_i$ 。

[0074] 其中,语音识别系统 $S_i$ 对目标语音进行识别后输出的多个候选识别结果可以为所有的候选识别结果,也可以为所有的候选识别结果中的N-best,通常而言,语音识别系统 $S_i$ 对目标语音进行识别后会输出N-best,即排名前N(N为大于或等于1的整数)的候选识别结果,需要说明的是,N根据具体应用情况设定,不同语音识别系统的N可以相同,也可以不同,另外,语音识别系统输出的多个候选识别结果是排序好,比如,语音识别系统将识别得分最

高的排在第一位,将识别得分次高的排在第二位,其它以此类推。

[0075] 假设语音识别系统 $S_i$ 对目标语音进行识别后输出 $N_i$ 个候选识别结果,上述第一实现方式直接将 $N_i$ 个候选识别结果组成的候选集作为目标语音在语音识别系统 $S_i$ 上对应的候选集 $R_i$ 。可见,若通过上述第一种实现方式获取目标语音在语音识别系统 $S_i$ 上对应的候选集 $R_i$ ,则候选集 $R_i$ 中包含语音识别系统 $S_i$ 对目标语音进行识别后输出的所有候选识别结果。

[0076] 考虑到语音识别系统 $S_i$ 对目标语音进行识别后输出的多个候选识别结果中可能存在解码异常的候选识别结果,为了避免解码异常的候选识别结果对最终识别结果确定的影响,同时提高识别结果的确定效率,本实施例提供如下第二种实现方式:

[0077] 首先,获取语音识别系统 $S_i$ 对目标语音进行识别后输出的多个候选识别结果(比如N-best候选识别结果);然后确定语音识别系统 $S_i$ 输出的多个候选识别结果中是否存在解码异常的识别结果;若 $S_i$ 输出的多个候选识别结果中存在解码异常的识别结果,则将解码异常的候选识别结果剔除,由剩余的候选识别结果组成目标语音在语音识别系统 $S_i$ 上对应的候选集 $R_i$ ;若语音识别系统 $S_i$ 输出的多个候选识别结果中不存在解码异常的候选识别结果,则由语音识别系统 $S_i$ 输出的多个候选识别结果组成目标语音在语音识别系统 $S_i$ 上对应的候选集 $R_i$ 。

[0078] 示例性的,语音识别系统 $S_i$ 对目标语音进行识别后输出 $N_i$ 个候选识别结果,若 $N_i$ 个候选识别结果中存在3个解码异常的候选识别结果,则将3个解码异常的候选识别结果剔除,将剩余的 $N_i - 3$ 个候选识别结果组成目标语音在语音识别系统 $S_i$ 上对应的候选集 $R_i$ ,若 $N_i$ 个候选识别结果中不存在解码异常的候选识别结果,则将 $N_i$ 个候选识别结果组成目标语音在语音识别系统 $S_i$ 上对应的候选集 $R_i$ 。

[0079] 可见,若通过上述第二种实现方式获取目标语音在语音识别系统 $S_i$ 上对应的候选集 $R_i$ ,则候选集 $R_i$ 中包含语音识别系统 $S_i$ 对目标语音进行识别后输出的所有(输出的多个候选识别结果中不存在解码异常的候选识别结果)或部分(输出的多个候选识别结果中存在解码异常的候选识别结果)候选识别结果。

[0080] 上述第二种实现方式中,在获得语音识别系统 $S_i$ 对目标语音进行识别后输出的多个候选识别结果后,需要确定语音识别系统 $S_i$ 输出的多个候选识别结果中是否存在解码异常的识别结果,接下来对这一过程进行介绍。

[0081] 确定语音识别系统 $S_i$ 输出的多个候选识别结果中是否存在解码异常的识别结果可以包括:确定语音识别系统 $S_i$ 输出的多个候选识别结果中是否存在单位时间解码词数大于预设词数阈值 $e_{thred}$ 的候选识别结果;若是,则确定语音识别系统 $S_i$ 输出的多个候选识别结果中存在解码异常的候选识别结果;若否,则确定语音识别系统 $S_i$ 输出的多个候选识别结果中不存在解码异常的候选识别结果。其中,预设词数阈值 $e_{thred}$ 可根据具体应用情况设定。

[0082] 其中,一候选识别结果 $r_{ij}$ 的单位时间解码词数根据候选识别结果 $r_{ij}$ 的解码时间 $t_{ij}$ 和候选识别结果 $r_{ij}$ 的词数 $q_{ij}$ 确定,具体的,候选识别结果 $r_{ij}$ 的单位时间解码词数 $e_{ij}$ 可表示为 $q_{ij}/t_{ij}$ 。

[0083] 需要说明的是,若一候选识别结果的单位时间解码词数大于预设词数阈值 $e_{thred}$ ,则可确定该候选识别结果解码异常,上述第二种实现方式中,将解码异常的候选识别结果剔除,即,将单位时间解码词数大于预设词数阈值 $e_{thred}$ 的候选识别结果剔除。

[0084] 步骤S102:将获取的多个候选集两两组合,得到若干组候选集。

[0085] 假设经由步骤S101获取了M个候选集,则对M个候选集进行两两组合,会得到 $M(M-1)/2$ 组候选集。

[0086] 示例性的,经由步骤S101获取了3个候选集,分别为 $R_1$ 、 $R_2$ 和 $R_3$ ,将 $R_1$ 、 $R_2$ 和 $R_3$ 两两组合会得到3组候选集,分别为 $(R_1, R_2)$ 、 $(R_1, R_3)$ 以及 $(R_2, R_3)$ 。

[0087] 步骤S103:确定若干组候选集分别对应的相似性信息。

[0088] 其中,相似性信息为对应的两个候选集中的第一候选集包含的候选识别结果与第二候选集包含的候选识别结果的句级级别的相似性信息。

[0089] 示例性的,一组候选集为 $(R_i, R_j)$ ,其中, $R_i = \{r_{i1}, r_{i2}, r_{i3}\}$ , $R_j = \{r_{j1}, r_{j2}, r_{j3}\}$ ,则 $(R_i, R_j)$ 对应的相似性信息可以包括 $r_{i1}$ 分别与 $r_{j1}$ 、 $r_{j2}$ 、 $r_{j3}$ 的句子级别的相似性信息, $r_{i2}$ 分别与 $r_{j1}$ 、 $r_{j2}$ 、 $r_{j3}$ 的句子级别的相似性信息, $r_{i3}$ 分别与 $r_{j1}$ 、 $r_{j2}$ 、 $r_{j3}$ 的句子级别的相似性信息。

[0090] 步骤S104:基于若干组候选集分别对应的相似性信息,从多个候选集中确定目标语音的识别结果。

[0091] 考虑到多个语音识别系统具有差异性和互补性,本实施例从目标语音在多个语音识别系统上分别对应的候选集中确定目标语音的识别结果,相比于从基于单个语音识别系统获得的候选识别结果中确定目标语音的识别结果,本申请提供的语音识别结果确定方法更加稳定,且能够提升语音识别结果的准确率,在从目标语音在多个语音识别系统上分别对应的候选集中确定目标语音的识别结果时,为了实现不同语音识别系统的候选识别结果互补,同时为了能够更快的确定出语音识别结果,本申请将目标语音在多个语音识别系统上分别对应的候选集两两组合,基于各组候选集分别对应的句子级别的相似性信息,从多个目标语音在多个语音识别系统上分别对应的候选集中确定目标语音的识别结果。经由本实施例提供的语音识别结果确定方法能够较快地确定出准确率较高的语音识别结果。

[0092] 第二实施例

[0093] 本实施例对上述实施例中的“步骤S103:确定若干组候选集分别对应的相似性信息”的具体实现过程进行介绍。

[0094] 可选的,相似性信息可以为相似性度量指标,相似性度量指标为可以度量相似程度的指标,假设一组候选集 $(R_i, R_j)$ ,则 $(R_i, R_j)$ 对应的相似性度量指标包括 $R_i$ 中每个候选识别结果与 $R_j$ 中每个候选识别结果的相似性度量指标。

[0095] 示例性的, $R_i = \{r_{i1}, r_{i2}\}$ , $R_j = \{r_{j1}, r_{j2}\}$ ,则 $(R_i, R_j)$ 对应的相似性度量指标包括 $R_i$ 中的候选识别结果 $r_{i1}$ 与 $R_j$ 中的候选识别结果 $r_{j1}$ 的相似性度量指标、 $R_i$ 中的候选识别结果 $r_{i1}$ 与 $R_j$ 中的候选识别结果 $r_{j2}$ 的相似性度量指标、 $R_i$ 中的候选识别结果 $r_{i2}$ 与 $R_j$ 中的候选识别结果 $r_{j1}$ 的相似性度量指标、 $R_i$ 中的候选识别结果 $r_{i2}$ 与 $R_j$ 中的候选识别结果 $r_{j2}$ 的相似性度量指标。

[0096] 需要说明的是,相似性度量指标可以采用第一度量指标,也可以采用第二度量指标,其中,两个候选识别结果的第一度量指标越小,则两个候选识别结果越相似,两个候选识别结果的第二度量指标越大,则两个候选识别结果越相似。可选的,第一度量指标可以但不限定为编辑距离、欧式距离等,第二度量指标可以但不限定为余弦相似度等。

[0097] 其中,两个候选识别结果(即一组候选集)的编辑距离可通过如下方式确定:

[0098] 定义 $d[k_1][k_2]$ , $k_1$ 对应两个候选识别结果中的第一候选识别结果, $k_1 \in [1, n_1]$ , $n_1$

表示第一候选识别结果的词数,  $k_2$  对应两个候选识别结果中的第二候选识别结果,  $k_2 \in [1, n_2]$ ,  $n_2$  表示第二候选识别结果的词数, 对于  $k_1$  从 1 到  $n_1$ ,  $k_2$  从 1 到  $n_2$ , 依次计算  $d[k_1][k_2]$ , 计算方式如下:

[0099] 当第一候选识别结果在  $k_1-1$  处与第二候选识别结果在  $k_2-1$  处相等时,  $d[k_1][k_2] = d[k_1-1][k_2-1]$ , 否则:

[0100]  $d[k_1][k_2] = \min \{d[k_1-1][k_2], d[k_1][k_2-1], d[k_1-1][k_2-1]\} + 1$

[0101] 经过计算, 最终可得到  $d[n_1][n_2]$ ,  $d[n_1][n_2]$  即为两个候选识别结果的编辑距离。

[0102] 示例性的, 两个候选识别结果中的第一候选识别结果为“my name is jack”, 第二候选识别结果为“my nick name tom”, 则按上述方式计算可获得:

		k1	my	name	is	jack
	k2	0	1	2	3	4
[0103]	my	1	0	1	2	3
	nick	2	1	1	1	2
	name	3	2	2	2	2
	tom	4	3	3	3	3

[0104] 由于  $d[4][4] = 3$ , 因此两个候选识别结果的编辑距离为 3。

[0105] 其中, 两个候选识别结果的余弦相似度可通过如下方式确定:

[0106] 确定每个候选识别结果的表征向量, 计算两个候选识别结果的表征向量的余弦相似度, 计算得到的余弦相似度作为两个候选识别结果的余弦相似度。其中, 一候选识别结果的表征向量可通过将该候选识别结果包含的每个词转换为对应的词向量获得。

[0107] 两个候选识别结果的欧式距离的确定计算方式与余弦相似度的确定方式类似, 均是先确定每个候选识别结果的表征向量, 然后确定两个候选识别结果的表征向量的欧式距离, 计算得到的欧式距离作为两个候选识别结果的欧式距离。

[0108] 需要说明的是, 上述给出的相似性度量指标 (即编辑距离、余弦相似度、欧式距离) 仅为示例, 其它可以对两个候选识别结果的相似性进行度量的指标也适用于本申请。

[0109] 第三实施例

[0110] 本实施例对上述实施例中的“步骤S104: 基于若干组候选集分别对应的相似性信息, 从多个候选集中确定目标语音的识别结果”的具体实现过程进行介绍。

[0111] 请参阅图2, 示出了基于若干组候选集分别对应的相似性信息, 从多个候选集中确定目标语音的识别结果的流程示意图, 可以包括:

[0112] 步骤S201: 基于若干组候选集分别对应的相似性信息, 确定若干组候选集分别对应的候选识别结果对, 以得到若干候选识别结果对。

[0113] 其中, 候选识别结果对包括对应的两个候选集中最相似的两个候选识别结果。

[0114] 示例性的, 一组候选集为  $(R_1, R_2)$ , 若  $R_1$  中的候选识别结果  $r_{13}$  与  $R_2$  中的候选识别结果  $r_{21}$  最相似, 则将  $(r_{13}, r_{21})$  确定为  $(R_1, R_2)$  对应的目标候选识别结果对。

[0115] 具体的, 基于若干组候选集分别对应的相似性信息, 确定若干组候选集分别对应的候选识别结果对的过程可以包括: 针对每组候选集, 若相似性度量指标采用第一度量指标, 则将相似性度量指标最小 (比如, 编辑距离最小) 的两个候选识别结果组成的候选识别

结果对,确定为该组候选集对应的候选识别结果对,若相似性度量指标采用第二度量指标,则将相似性度量指标最大(比如,余弦相似度最大)的两个候选识别结果组成的候选识别结果对,确定为该组候选集对应的候选识别结果对。

[0116] 示例性的,相似性度量指标采用编辑距离,有3组候选集,分别为 $(R_1, R_2)$ 、 $(R_1, R_3)$ 和 $(R_2, R_3)$ ,假设 $R_1 = \{r_{11}, r_{12}\}$ ,  $R_2 = \{r_{21}, r_{22}\}$ ,  $R_3 = \{r_{31}, r_{32}, r_{33}\}$ ,对于 $(R_1, R_2)$ 而言, $(R_1, R_2)$ 对应的相似性信息包括 $r_{11}$ 与 $r_{21}$ 的编辑距离、 $r_{11}$ 与 $r_{22}$ 的编辑距离、 $r_{12}$ 与 $r_{21}$ 的编辑距离、 $r_{12}$ 与 $r_{22}$ 的编辑距离,假设 $r_{12}$ 与 $r_{21}$ 的编辑距离最小,则将候选识别结果对 $(r_{12}, r_{21})$ 确定为候选集为 $(R_1, R_2)$ 对应的候选识别结果对,同样的方式可确定除 $(R_1, R_3)$ 对应的候选识别结果对以及 $(R_2, R_3)$ 以及 $(R_2, R_3)$ 对应的候选识别结果对,最终会获得3个候选识别结果对。

[0117] 需要说明的是,对于一组候选集,若其包含的两组候选集中最相似的候选识别结果对有多个(比如,编辑距离最小的候选识别结果对有多个),则将综合得分和最大的候选识别结果对作为该组候选集对应的候选识别结果对。示例性的,一组候选集为 $(R_1, R_2)$ ,  $R_1 = \{r_{11}, r_{12}\}$ ,  $R_2 = \{r_{21}, r_{22}\}$ ,其中, $r_{11}$ 与 $r_{21}$ 的编辑距离为 $d_1$ , $r_{11}$ 与 $r_{22}$ 的编辑距离为 $d_2$ , $r_{12}$ 与 $r_{21}$ 的编辑距离为 $d_3$ , $r_{12}$ 与 $r_{22}$ 的编辑距离为 $d_4$ ,假设 $d_1 = d_2 < d_3 < d_4$ ,则将 $r_{11}$ 与 $r_{21}$ 的综合得分求和,求和后得分表示为 $s_1$ ,将 $r_{11}$ 与 $r_{22}$ 的综合得分求和,求和后得分表示为 $s_2$ ,若 $s_1 > s_2$ ,则将 $(r_{11}, r_{21})$ 确定为 $(R_1, R_2)$ 对应的候选识别结果对,若 $s_1 < s_2$ ,则将 $(r_{11}, r_{22})$ 确定为 $(R_1, R_2)$ 对应的候选识别结果对,若 $s_1 = s_2$ ,则将 $(r_{11}, r_{21})$ 与 $(r_{11}, r_{22})$ 中的任一个确定为 $(R_1, R_2)$ 对应的候选识别结果对。

[0118] 步骤S202:基于若干候选识别结果对分别对应的相似性信息,确定目标语音的识别结果。

[0119] 其中,一候选识别结果对对应的相似性信息为该候选识别结果对包含的两个候选识别结果的相似性信息。

[0120] 请参阅图3,示出了基于若干候选识别结果对分别对应的相似性信息,确定目标语音的识别结果的流程示意图,可以包括:

[0121] 步骤S301:基于若干候选识别结果对分别对应的相似性度量指标,从若干候选识别结果对中确定最相似的候选识别结果对,作为目标候选识别结果对。

[0122] 其中,一候选识别结果对对应的相似性度量指标为该候选识别结果对包含的两个候选识别结果的相似性度量指标。示例性的,相似性度量指标度量采用编辑距离,则一候选识别结果对对应的编辑距离为该候选识别结果对包含的两个候选识别结果的编辑距离。

[0123] 示例性的,有3组候选集,则相应获得3个候选识别结果对,其中,第1个候选识别结果对对应的编辑距离为 $d_1$ ,第2个候选识别结果对对应的编辑距离为 $d_2$ ,第3个候选识别结果对对应的编辑距离为 $d_3$ ,若 $d_3 < d_2 < d_1$ ,则说明第3个候选识别结果对为最相似的识别结果对,将第3个候选识别结果对确定为目标候选识别结果对。

[0124] 需要说明的是,若若干候选识别结果对中最相似的候选识别结果对有多个,则将综合得分和最高的候选识别结果对确定为目标候选识别结果对。示例性的,有3组候选集,则相应获得3个候选识别结果对,其中,第1个候选识别结果对对应的编辑距离为 $d_1$ ,第2个候选识别结果对对应的编辑距离为 $d_2$ ,第3个候选识别结果对对应的编辑距离为 $d_3$ ,若 $d_3 = d_2 < d_1$ ,则将第2个候选识别结果对包含的两个候选识别结果的综合得分求和,并将第3个候选识别结果对包含的两个候选识别结果的综合得分求和,若第2个候选识别结果对包含的

两个候选识别结果的综合得分和大于第3个候选识别结果对包含的两个候选识别结果的综合得分和,则将第2个候选识别结果对确定为目标候选识别结果对,若第3个候选识别结果对包含的两个候选识别结果的综合得分和大于第2个候选识别结果对包含的两个候选识别结果的综合得分和,则将第3个候选识别结果对确定为目标候选识别结果对,若第3个候选识别结果对包含的两个候选识别结果的综合得分和与第2个候选识别结果对包含的两个候选识别结果的综合得分和相同,则将第2个候选识别结果对与第3个候选识别结果对中的任意一对确定为目标候选识别结果对。

[0125] 步骤S302:至少基于目标候选识别结果对对应的相似性度量指标,确定目标语音的识别结果。

[0126] 具体的,至少基于目标候选识别结果对对应的相似性度量指标,确定目标语音的识别结果包括:

[0127] 步骤S302a、若目标候选识别结果对对应的相似性度量指标为表征两个候选识别结果相同的目标值,则将目标候选识别结果对中的任意一候选识别结果确定为目标语音的识别结果。

[0128] 示例性的,相似性度量指标采用编辑距离,则表征两个候选识别结果相同的目标值为0,若目标候选识别结果对对应的编辑距离为0,则说明两个候选识别结果相同,则可将目标候选识别结果对中的任意一候选识别结果确定为目标语音的识别结果。

[0129] 示例性的,相似性度量指标采用余弦相似度,则表征两个候选识别结果相同的目标值为1,若目标候选识别结果对对应的余弦相似度为1,则说明两个候选识别结果相同,则可将目标候选识别结果对中的任意一候选识别结果确定为目标语音的识别结果。

[0130] 步骤S302b、若目标候选识别结果对对应的相似性度量指标不为目标值,则基于目标候选识别结果对对应的相似性度量指标以及目标候选识别结果对中每个候选识别结果的词数和综合得分,确定目标语音的识别结果。

[0131] 其中,一候选识别结果的综合得分基于该候选识别结果的声学得分和语言得分确定。需要说明的是,每个候选识别结果均对应有两个得分,即声学得分(声学模型的打分)和语言得分(语言模型的打分),声学得分和语言得分由对应的语音识别系统在输出候选识别结果时一并输出,示例性的,一语音识别系统输出了10个候选识别结果,其在输出这10个候选识别结果时,还会输出每个候选识别结果的声学得分和语言得分,

[0132] 基于一候选识别结果的声学得分和语言得分确定该候选识别结果的综合得分的实现方式有多种,本实施例提供如下两种可选的实现方式:

[0133] 第一种实现方式:

[0134] 步骤a1、将该候选识别结果的语言得分处理为目标范围内的值,以得到处理后的语言得分。

[0135] 其中,该候选识别结果的声学得分处于目标范围内。

[0136] 步骤a2、将处理后的语言得分与该候选识别结果的声学得分融合,融合后的得分作为该候选识别结果的综合得分。

[0137] 将处理后的语言得分与该候选识别结果的声学得分融合的方式有多种,比如,可将处理后的语言得分与该候选识别结果的声学得分直接求和,也可将处理后的语言得分与该候选识别结果的声学得分加权求和。

[0138] 第二种实现方式:

[0139] 步骤b1、根据该候选识别结果的词数和该候选识别结果的语言得分,确定该候选识别结果的合理程度表征值。

[0140] 假设该候选识别结果为目标语音在第*i*个语音识别系统上对应的候选集中的第*j*个候选识别结果 $r_{ij}$ ,则可根据下式示出的方式确定该候选识别结果 $r_{ij}$ 的合理程度表征值

$z_{ij}$ :

$$[0141] \quad z_{ij} = q_{ij} / \ln(ppl_{ij} + \epsilon) \quad (1)$$

[0142] 其中, $q_{ij}$ 表示该候选识别结果 $r_{ij}$ 的词数,即该候选识别结果 $r_{ij}$ 包含的词的数量, $ppl_{ij}$ 表示该候选识别结果 $r_{ij}$ 的语言得分, $\epsilon$ 取 $10^{-6}$ 。

[0143] 步骤b2、将该候选识别结果的合理程度表征值归一化为目标范围内的值,以得到归一化后的合理程度表征值。

[0144] 其中,该候选识别结果的声学得分处于目标范围内。

[0145] 具体的,按下式示出的归一化方式将候选识别结果 $r_{ij}$ 的合理程度表征值 $z_{ij}$ 归一化为目标范围内的值:

$$[0146] \quad l_{ij} = \tanh(cz_{ij}) = (e^{cz_{ij}} - e^{-cz_{ij}}) / (e^{cz_{ij}} + e^{-cz_{ij}}) \quad (2)$$

[0147] 其中, $c$ 表示平滑化因子,其为可调参数, $l_{ij}$ 即表示归一化后的合理程度表征值。

[0148] 步骤b3、将归一化后的合理程度表征值与该候选识别结果的声学得分融合,融合后的得分作为该候选识别结果的综合得分。

[0149] 在一种可能的实现方式中,可直接将归一化后的合理程度表征值与该候选识别结果的声学得分求和,求和后的得分作为该候选识别结果的综合得分;在另一种可能的实现方式中,可将归一化后的合理程度表征值与该候选识别结果的声学得分加权求和,加权求和后的得分作为该候选识别结果的综合得分,即:

$$[0150] \quad s_{ij} = \alpha l_{ij} + \beta a_{ij} \quad (3)$$

[0151] 其中, $s_{ij}$ 即表示候选识别结果 $r_{ij}$ 的综合得分, $a_{ij}$ 表示候选识别结果 $r_{ij}$ 的声学得分, $\alpha$ 表示语言模型得分因子, $\beta$ 表示声学模型得分因子, $\alpha + \beta = 1$ , $\alpha$ 和 $\beta$ 为可调参数。

[0152] 相比于第一种实现方式,第二种实现方式还考虑了候选识别结果的词数,由于候选识别结果的优劣与词数相关,因此,结合候选识别结果的词数确定的综合得分能更好地反映候选识别结果的优劣。有鉴于此,优选上述第二种实现方式确定候选识别结果的综合得分。

[0153] 接下来对步骤S302b中“基于目标候选识别结果对对应的相似性度量指标以及目标候选识别结果对中每个候选识别结果的词数和综合得分,确定目标语音的识别结果”的实现过程进行介绍。

[0154] 基于目标候选识别结果对对应的相似性度量指标以及目标候选识别结果对中每个候选识别结果的词数和综合得分,确定目标语音的识别结果的过程可以包括:

[0155] 在目标候选识别结果对的相似性度量指标与预设相似性阈值满足预设关系的情况下:

[0156] (1) 若目标候选识别结果对中第一候选识别结果的词数与第二候选识别结果的词数相同,则将目标候选识别结果对中综合得分大的候选识别结果确定为目标语音的识别结果。

[0157] 需要说明的是,目标候选识别结果对中的第一候选识别结果的词数与第二候选识别结果的词数相同,说明第一候选识别结果的词数与第二候选识别结果可能发生的替换、删除、插入的数量是相同的,此种情况下,将第一候选识别结果与第二候选识别结果中综合得分大的候选识别结果确定为目标语音的识别结果。

[0158] (2)若第一候选识别结果的词数大于第二候选识别结果的词数,且第二候选识别结果的综合得分大于第一候选识别结果的综合得分,则将第二候选识别结果确定为目标语音的识别结果。

[0159] 需要说明的是,目标候选识别结果对中的第一候选识别结果的词数大于第二候选识别结果的词数,说明第一候选识别结果可能发生的插入错误比第二候选识别结果更多,则取插入错误更少且综合得分更高的第二候选识别结果作为目标语音的识别结果。

[0160] (3)若第二候选识别结果的词数大于第一候选识别结果的词数,且第一候选识别结果的综合得分大于第二候选识别结果的综合得分,则将第一候选识别结果确定为目标语音的识别结果。

[0161] 需要说明的是,目标候选识别结果对中的第二候选识别结果的词数大于第二候选识别结果的词数,说明第二候选识别结果可能发生的插入错误比第一候选识别结果更多,则取插入错误更少且综合得分更高的第一候选识别结果作为目标语音的识别结果。

[0162] 示例性的,目标候选识别结果对为 $(r_{ip}, r_{jq})$ , $(r_{ip}, r_{jq})$ 对应的相似性度量指标为 $(r_{ip}, r_{jq})$ 对应的编辑距离,即 $r_{ip}$ 与 $r_{jq}$ 的编辑距离,则基于 $(r_{ip}, r_{jq})$ 对应的编辑距离以及目标候选识别结果对中每个候选识别结果的词数和综合得分,确定目标语音的识别结果的过程包括:在 $(r_{ip}, r_{jq})$ 对应的编辑距离小于预设的编辑距离阈值的情况下,若 $r_{ip}$ 的词数与 $r_{jq}$ 的词数相同,则将 $r_{ip}, r_{jq}$ 中的任一个确定为目标语音的识别结果;若 $r_{ip}$ 的词数大于 $r_{jq}$ 的词数,且 $r_{jq}$ 的综合得分大于 $r_{ip}$ 的综合得分,则将插入错误更少且综合得分更高的 $r_{jq}$ 确定为目标语音的识别结果;若 $r_{jq}$ 的词数大于 $r_{ip}$ 的词数, $r_{ip}$ 的综合得分大于 $r_{jq}$ 的综合得分,则将插入错误更少且综合得分更高的 $r_{ip}$ 确定为目标语音的识别结果。

[0163] 需要说明的是:若目标候选识别结果对的相似性度量指标不为目标值,且目标候选识别结果对的相似性度量指标与预设相似性阈值不满足预设条件(比如,目标候选识别结果对的编辑距离不小于预设的编辑距离阈值),则确定无法从目标候选识别结果对中确定目标语音的识别结果;在目标候选识别结果对的相似性度量指标不为目标值,且目标候选识别结果对的相似性度量指标与预设相似性阈值满足预设条件的情况下,若第一候选识别结果的词数大于第二候选识别结果的词数,且第一候选识别结果的综合得分大于或等于第二候选识别结果的综合得分,则确定无法从目标候选识别结果对中确定目标语音的识别结果;在目标候选识别结果对的相似性度量指标不为目标值,且目标候选识别结果对的相似性度量指标与预设相似性阈值满足预设条件的情况下,若第二候选识别结果的词数大于第一候选识别结果的词数,且第二候选识别结果的综合得分大于或等于一候选识别结果的综合得分,则确定无法从目标候选识别结果对中确定目标语音的识别结果。

[0164] 在确定无法从目标候选识别结果对中确定目标语音的识别结果时,本申请采用如下策略确定目标语音的识别结果:从每个候选集中获取识别得分最高的候选识别结果(比如每个语音识别系统输出的1-best);将从各个候选集中分别获取的识别得分最高的候选识别结果中综合得分最高的候选识别结果,确定为目标语音的识别结果。



[0165] 经由上述第一实施例和第二实施例可知,本申请实施例提供的语音识别结果的确定方法可获取多个不同语音识别系统分别输出的N-Best候选识别结果,可将每个语音识别系统输出的N-Best候选识别结果中解码异常的候选识别结果剔除,以得到目标语音在不同语音识别系统上分别对应的候选集,还可将目标语音在不同语音识别系统上分别对应的候选集两两组合,针对每组候选集,从两个候选集中找出最相似的两个候选识别结果,从而得到各组候选集分别对应的候选识别结果对,然后从各组候选集分别对应的候选识别结果对中最相似的候选识别结果对,最后基于最相似的候选识别结果对的相关信息(比如最相似的候选识别结果对对应的相似性度量指标、最相似的候选识别结果对包含的每个候选识别结果的次数和综合得分)确定目标语音的识别结果。本申请从目标语音在不同语音识别系统上分别对应的候选集中确定目标语音的识别结果,由于不同语音识别系统输出的候选识别结果具有互补性,因此,相比于从基于单个语音识别系统输出的候选识别结果中确定目标语音的识别结果,本申请提供的语音识别结果确定方法更加稳定,且能够提升语音识别结果的准确率,在从目标语音在多个语音识别系统上分别对应的候选集中确定目标语音的识别结果时,为了实现不同语音识别系统的候选识别结果互补,同时为了能够更快的确定出语音识别结果,本申请将目标语音在多个语音识别系统上分别对应的候选集两两组合,基于各组候选集分别对应的句子级别的相似性信息,从多个目标语音在多个语音识别系统上分别对应的候选集中确定目标语音的识别结果。经由本申请提供的语音识别结果确定方法能够较快地确定出准确率较高的语音识别结果。

#### [0166] 第四实施例

[0167] 本申请实施例还提供了一种语音识别结果确定装置,下面对本申请实施例提供的语音识别结果确定装置进行描述,下文描述的语音识别结果确定装置与上文描述的语音识别结果确定方法可相互对应参照。

[0168] 请参阅图4,示出了本申请实施例提供的语音识别结果确定装置的结构示意图,可以包括候选集获取模块401、候选集组合模块402、相似性信息确定模块403和识别结果确定模块404。其中:

[0169] 候选集获取模块401,用于获取目标语音在多个语音识别系统上分别对应的候选集。其中,所述目标语音在一个语音识别系统上对应的候选集包括基于该语音识别系统获得的若干候选识别结果。

[0170] 候选集组合模块402,用于将获取的多个候选集两两组合,以得到若干组候选集。

[0171] 相似性信息确定模块403,用于确定所述若干组候选集分别对应的相似性信息。其中,所述相似性信息为对应的两个候选集中的第一候选集包含的候选识别结果与第二候选集包含的候选识别结果的句子级别的相似性信息。

[0172] 识别结果确定模块404,用于基于所述若干组候选集分别对应的相似性信息,从所述多个候选集中确定目标语音的识别结果。

[0173] 可选的,候选集获取模块401在获取目标语音在多个语音识别系统上分别对应的候选集时,具体用于:

[0174] 获取该语音识别系统对目标语音进行识别后输出的多个候选识别结果;

[0175] 确定所述多个候选识别结果中是否存在解码异常的候选识别结果;

[0176] 若是,则将解码异常的候选识别结果剔除,由剩余的候选识别结果组成所述目标

语音在该语音识别系统上对应的候选集；

[0177] 若否，则由所述多个候选识别结果组成所述目标语音在该语音识别系统上对应的候选集。

[0178] 可选的，候选集获取模块401在确定所述多个候选识别结果中是否存在解码异常的候选识别结果时，具体用于：

[0179] 确定所述多个候选识别结果中是否存在每秒解码词数大于预设词数阈值的候选识别结果；

[0180] 若是，则确定所述多个候选识别结果中存在解码异常的候选识别结果；

[0181] 若否，则确定所述多个候选识别结果中不存在解码异常的候选识别结果；

[0182] 候选集获取模块401在将解码异常的候选识别结果剔除时，具体用于：

[0183] 将每秒解码词数大于所述预设词数阈值的候选识别结果剔除。

[0184] 可选的，识别结果确定模块404可以包括：候选识别结果对确定子模块和识别结果确定子模块。

[0185] 候选识别结果对确定子模块，用于基于所述若干组候选集分别对应的相似性信息，确定所述若干组候选集分别对应的候选识别结果对，以得到若干候选识别结果对，其中，所述候选识别结果对包括对应的两个候选集中最相似的两个候选识别结果；

[0186] 识别结果确定子模块，用于基于所述若干候选识别结果对分别对应的相似性信息，确定所述目标语音的识别结果，其中，一候选识别结果对对应的相似性信息为该候选识别结果对包含的两个候选识别结果的相似性信息。

[0187] 可选的，所述相似性信息为相似性度量指标，则识别结果确定子模块在基于所述若干候选识别结果对分别对应的相似性信息，确定所述目标语音的识别结果时，具体用于：

[0188] 基于所述若干候选识别结果对分别对应的相似性度量指标，从所述若干候选识别结果对中确定最相似的候选识别结果对，作为目标候选识别结果对；

[0189] 至少基于所述目标候选识别结果对对应的相似性度量指标，确定所述目标语音的识别结果。

[0190] 可选的，识别结果确定子模块在至少基于所述目标候选识别结果对对应的相似性度量指标，确定所述目标语音的识别结果时，具体用于：

[0191] 若所述目标候选识别结果对对应的相似性度量指标为表征两个候选识别结果相同的目标值，则将所述目标候选识别结果对中的任一候选识别结果确定为所述目标语音的识别结果；

[0192] 若所述目标候选识别结果对对应的相似性度量指标不为所述目标值，则基于所述目标候选识别结果对对应的相似性度量指标以及所述目标候选识别结果对中每个候选识别结果的词数和综合得分，确定所述目标语音的识别结果，其中，一候选识别结果的综合得分基于该候选识别结果的声学得分和语言得分确定。

[0193] 可选的，本申请提供的语音识别结果确定装置还可以包括综合得分确定模块。综合得分确定模块在确定一候选识别结果的综合得分时，具体用于：

[0194] 根据该候选识别结果的词数和该候选识别结果的语言得分，确定该候选识别结果的合理程度表征值；

[0195] 将该候选识别结果的合理程度表征值归一化为目标范围内的值，以得到归一化后

的合理程度表征值,其中,该候选识别结果的声学得分处于所述目标范围内;

[0196] 将所述归一化后的合理程度表征值与该候选识别结果的声学得分融合,融合后的得分作为该候选识别结果的综合得分。

[0197] 可选的,识别结果确定子模块在基于所述目标候选识别结果对对应的相似性度量指标以及所述目标候选识别结果对中每个候选识别结果的词数和综合得分,确定所述目标语音的识别结果时,具体用于:

[0198] 在所述目标候选识别结果对的相似性度量指标与预设相似性阈值满足预设关系的情况下:

[0199] 若所述目标候选识别结果对中第一候选识别结果的词数与第二候选识别结果的词数相同,则将所述目标候选识别结果对中综合得分大的候选识别结果确定为所述目标语音的识别结果;

[0200] 若所述第一候选识别结果的词数大于所述第二候选识别结果的词数,且所述第二候选识别结果的综合得分大于所述第一候选识别结果的综合得分,则将所述第二候选识别结果确定为所述目标语音的识别结果;

[0201] 若所述第二候选识别结果的词数大于所述第一候选识别结果的词数,且所述第一候选识别结果的综合得分大于所述第二候选识别结果的综合得分,则将所述第一候选识别结果确定为所述目标语音的识别结果。

[0202] 可选的,识别结果确定子模块还用于:

[0203] 若基于所述目标候选识别结果对的相似性度量指标以及所述目标候选识别结果对中每个候选识别结果的词数和综合得分,确定无法从所述目标候选识别结果对中确定所述目标语音的识别结果,则从每个候选集中获取识别得分最高的候选识别结果;

[0204] 将从各个候选集中分别获取的识别得分最高的候选识别结果中综合得分最高的候选识别结果,确定为所述目标语音的识别结果。

[0205] 考虑到多个语音识别系统具有差异性和互补性,本实施例提供的语音识别确定装置从目标语音在多个语音识别系统上分别对应的候选集中确定目标语音的识别结果,相比于从基于单个语音识别系统获得的候选识别结果中确定目标语音的识别结果,本申请提供的语音识别确定装置确定语音识别结果的方式更加稳定,且能够提升语音识别结果的准确率,在从目标语音在多个语音识别系统上分别对应的候选集中确定目标语音的识别结果时,为了实现不同语音识别系统的候选识别结果互补,同时为了能够更快的确定出语音识别结果,本申请实施例提供的语音识别确定装置将目标语音在多个语音识别系统上分别对应的候选集两两组合,基于各组候选集分别对应的句子级别的相似性信息,从多个目标语音在多个语音识别系统上分别对应的候选集中确定目标语音的识别结果。经由本实施例提供的语音识别结果确定装置能够较快地确定出准确率较高的语音识别结果。

[0206] 第五实施例

[0207] 本申请实施例还提供了一种语音识别结果确定设备,请参阅图5,示出了该语音识别结果确定设备的结构示意图,该舆情分析设备可以包括:至少一个处理器501,至少一个通信接口502,至少一个存储器503和至少一个通信总线504;

[0208] 在本申请实施例中,处理器501、通信接口502、存储器503、通信总线504的数量为至少一个,且处理器501、通信接口502、存储器503通过通信总线504完成相互间的通信;

[0209] 处理器501可能是一个中央处理器CPU,或者是特定集成电路ASIC (Application Specific Integrated Circuit),或者是被配置成实施本发明实施例的一个或多个集成电路等;

[0210] 存储器503可能包含高速RAM存储器,也可能还包括非易失性存储器 (non-volatile memory) 等,例如至少一个磁盘存储器;

[0211] 其中,存储器存储有程序,处理器可调用存储器存储的程序,所述程序用于:

[0212] 获取目标语音在多个语音识别系统上分别对应的候选集,其中,所述目标语音在一个语音识别系统上对应的候选集包括基于该语音识别系统获得的若干候选识别结果;

[0213] 将获取的多个候选集两两组合,以得到若干组候选集;

[0214] 确定所述若干组候选集分别对应的相似性信息,并基于所述若干组候选集分别对应的相似性信息,从所述多个候选集中确定目标语音的识别结果,其中,所述相似性信息为对应的两个候选集中的第一候选集包含的候选识别结果与第二候选集包含的候选识别结果的句子级别的相似性信息。

[0215] 可选的,所述程序的细化功能和扩展功能可参照上文描述。

[0216] 第六实施例

[0217] 本申请实施例还提供一种计算机可读存储介质,该计算机可读存储介质可存储有适于处理器执行的程序,所述程序用于:

[0218] 获取目标语音在多个语音识别系统上分别对应的候选集,其中,所述目标语音在一个语音识别系统上对应的候选集包括基于该语音识别系统获得的若干候选识别结果;

[0219] 将获取的多个候选集两两组合,以得到若干组候选集;

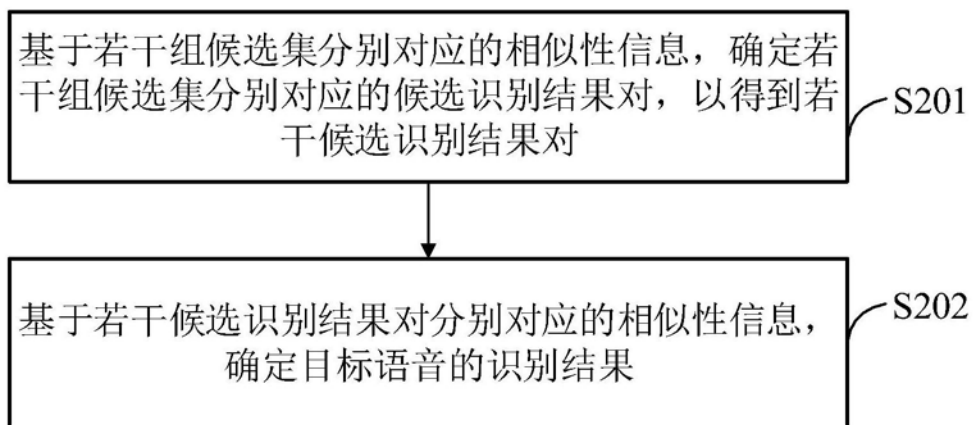
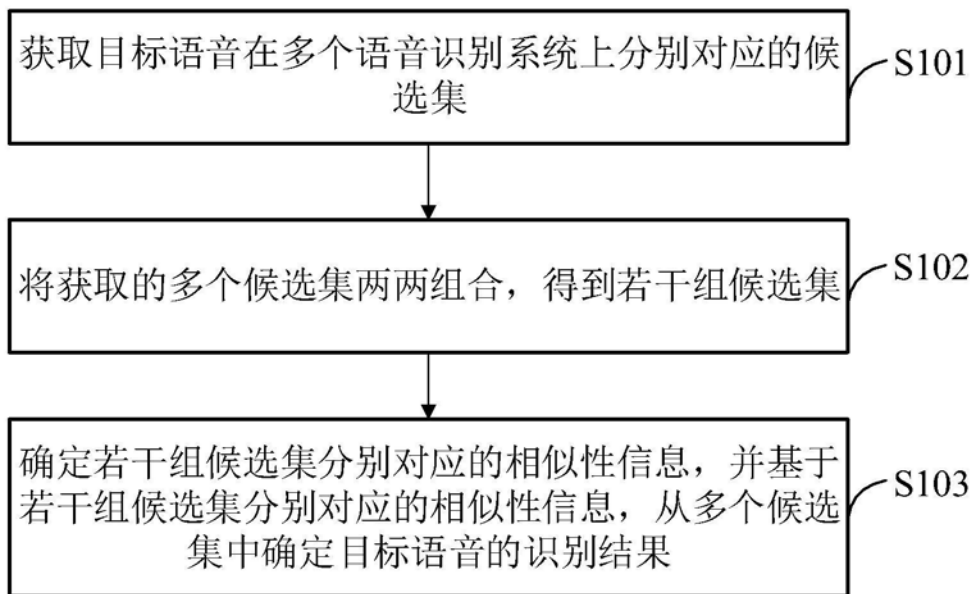
[0220] 确定所述若干组候选集分别对应的相似性信息,并基于所述若干组候选集分别对应的相似性信息,从所述多个候选集中确定目标语音的识别结果,其中,所述相似性信息为对应的两个候选集中的第一候选集包含的候选识别结果与第二候选集包含的候选识别结果的句子级别的相似性信息。

[0221] 可选的,所述程序的细化功能和扩展功能可参照上文描述。

[0222] 最后,还需要说明的是,在本文中,诸如第一和第二等之类的关系术语仅仅用来将一个实体或者操作与另一个实体或操作区分开来,而不一定要求或者暗示这些实体或操作之间存在任何这种实际的关系或者顺序。而且,术语“包括”、“包含”或者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、物品或者设备不仅包括那些要素,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、物品或者设备所固有的要素。在没有更多限制的情况下,由语句“包括一个……”限定的要素,并不排除在包括所述要素的过程、方法、物品或者设备中还存在另外的相同要素。

[0223] 本说明书中各个实施例采用递进的方式描述,每个实施例重点说明的都是与其他实施例的不同之处,各个实施例之间相同相似部分互相参见即可。

[0224] 对所公开的实施例的上述说明,使本领域专业技术人员能够实现或使用本发明。对这些实施例的多种修改对本领域的专业技术人员来说将是显而易见的,本文中所定义的一般原理可以在不脱离本发明的精神或范围的情况下,在其它实施例中实现。因此,本发明将不会被限制于本文所示的这些实施例,而是要符合与本文所公开的原理和新颖特点相一致的最宽的范围。



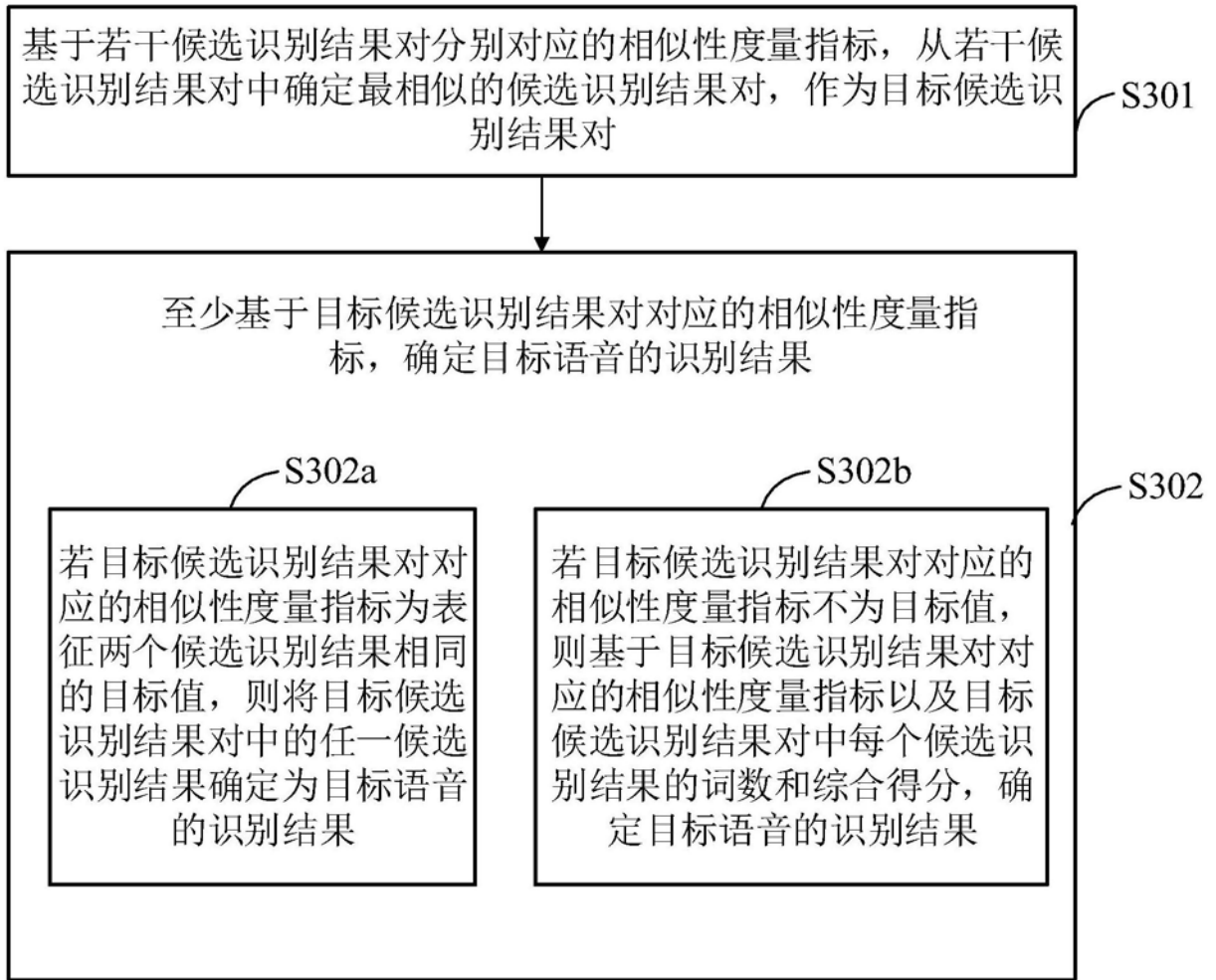


图3

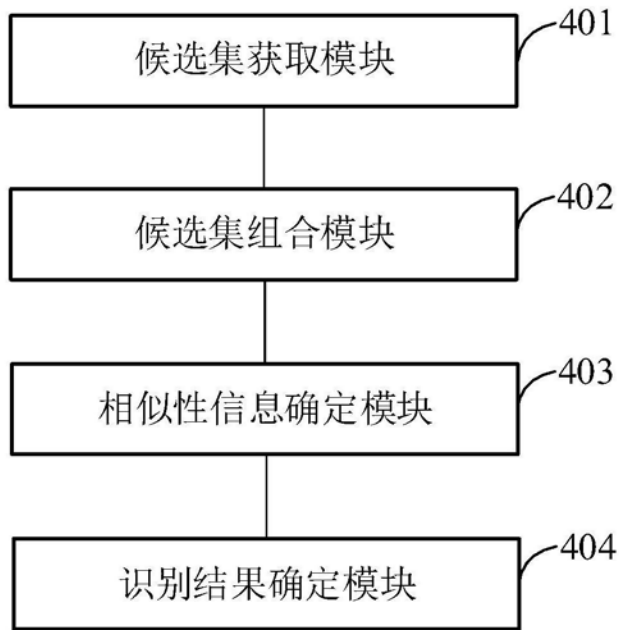


图4

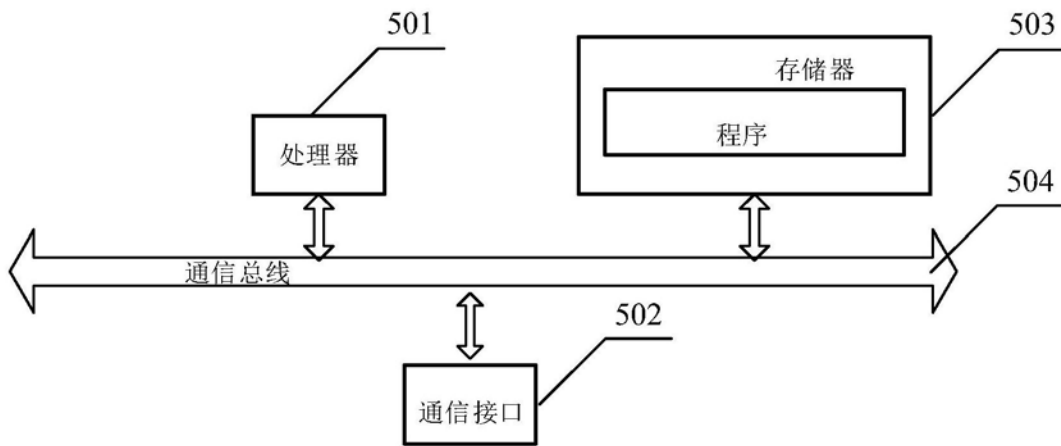


图5