



# (12) 发明专利申请

(10) 申请公布号 CN 116245103 A

(43) 申请公布日 2023. 06. 09

(21) 申请号 202211091272.8

(22) 申请日 2022.09.07

(71) 申请人 京东科技信息技术有限公司

地址 101116 北京市大兴区北京经济技术  
开发区科创十一街18号院2号楼6层  
601

(72) 发明人 王泽勋 冯明超 陈蒙

(74) 专利代理机构 北京唯智勤实知识产权代理  
事务所(普通合伙) 11557

专利代理师 史立状

(51) Int.Cl.

G06F 40/295 (2020.01)

G06N 3/08 (2023.01)

G06N 3/0464 (2023.01)

G10L 15/26 (2006.01)

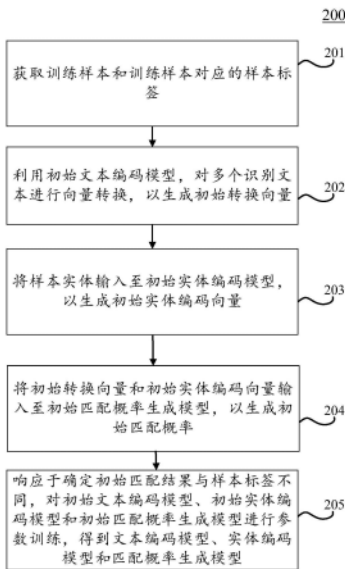
权利要求书3页 说明书15页 附图6页

## (54) 发明名称

模型训练方法、实体确定方法、装置、电子设  
备和介质

## (57) 摘要

本公开的实施例公开了模型训练方法、实体确定方法、装置、电子设备和介质。该方法的一具体实施方式包括：获取训练样本和训练样本对应的样本标签；利用初始文本编码模型，对多个识别文本进行向量转换，以生成初始转换向量；将样本实体输入至初始实体编码模型，以生成初始实体编码向量；将初始转换向量和初始实体编码向量输入至初始匹配概率生成模型，以生成初始匹配概率；响应于确定初始匹配结果与样本标签不同，对初始文本编码模型、初始实体编码模型和初始匹配概率生成模型进行参数训练。该实施方式与人工智能有关，可以生成实体检索精准度更高的文本编码模型、实体编码模型和匹配概率生成模型。



1. 一种模型训练方法,包括:

获取训练样本和所述训练样本对应的样本标签,其中,所述训练样本包括:针对音频样本的多个识别文本和所述音频样本中涉及的样本实体;

利用初始文本编码模型,对所述多个识别文本进行向量转换,以生成初始转换向量,其中,所述初始转换向量表征所述多个识别文本的文本特征信息;

将所述样本实体输入至初始实体编码模型,以生成初始实体编码向量;

将所述初始转换向量和所述初始实体编码向量输入至初始匹配概率生成模型,以生成初始匹配概率,其中,所述初始匹配概率表征所述初始转换向量与所述初始实体编码向量之间的匹配关系;

响应于确定初始匹配结果与所述样本标签不同,对所述初始文本编码模型、所述初始实体编码模型和所述初始匹配概率生成模型进行参数训练,得到文本编码模型、实体编码模型和匹配概率生成模型,其中,所述初始匹配结果是基于所述初始匹配概率生成的。

2. 根据权利要求1所述的方法,其中,所述利用初始文本编码模型,对所述多个识别文本进行向量转换,以生成初始转换向量,包括:

对所述多个识别文本进行文本拼接,得到拼接后文本;

将所述拼接后文本输入至所述初始文本编码模型,得到所述初始转换向量。

3. 根据权利要求1所述的方法,其中,所述利用初始文本编码模型,对所述多个识别文本进行向量转换,以生成初始转换向量,包括:

将所述多个识别文本输入至所述初始文本编码模型,得到多个初始文本编码向量;

对所述多个初始文本编码向量进行向量组合,得到组合矩阵;

对所述组合矩阵进行池化操作,得到所述初始转换向量。

4. 一种实体确定方法,包括:

获取针对目标音频的多个识别文本;

利用预先训练的文本编码模型,对所述多个识别文本进行向量转换,以生成转换向量,其中,所述转换向量表征所述多个识别文本的特征信息,所述文本编码模型是根据权利要求1-3中任一所述的方法训练的;

利用预先训练的实体编码模型,对实体库中的各个实体进行实体编码,得到实体编码向量集,其中,所述实体编码模型是根据权利要求1-3中任一所述的方法训练的;

利用预先训练的匹配概率生成模型,从所述实体库中筛选出对应匹配概率满足预设条件的实体,作为目标实体,得到至少一个目标实体,其中,目标实体对应的匹配概率是基于对应实体编码向量和所述转换向量生成的,所述匹配概率生成模型是根据权利要求1-3中任一所述的方法训练的。

5. 根据权利要求4所述的方法,其中,所述利用预先训练的匹配概率生成模型,从所述实体库中筛选出对应匹配概率满足预设条件的实体,作为目标实体,包括:

确定所述多个识别文本对应的关键词集;

从所述实体库中筛选出与关键词集存在关联关系的实体,作为关联实体,得到关联实体集;

从所述实体编码向量集中筛选出与所述关联实体集中各个关联实体相对应的实体编码向量,作为目标实体编码向量,得到目标实体编码向量集;

利用所述匹配概率生成模型,确定所述目标实体编码向量集中每个目标实体编码向量与所述转换向量之间的匹配概率;

从所述关联实体集中筛选出对应匹配概率满足所述预设条件的关联实体,作为目标实体。

6. 根据权利要求4所述的方法,其中,所述利用预先训练的匹配概率生成模型,从所述实体库中筛选出对应匹配概率满足预设条件的实体,作为目标实体,包括:

利用所述匹配概率生成模型,确定所述实体编码向量集中每个实体编码向量与所述转换向量之间的匹配概率;

从所述实体库中筛选出对应匹配概率满足所述预设条件的实体,作为目标实体。

7. 一种模型训练装置,包括:

第一获取单元,被配置成获取训练样本和所述训练样本对应的样本标签,其中,所述训练样本包括:针对音频样本的多个识别文本和所述音频样本中涉及的样本实体;

第一向量转换单元,被配置成利用初始文本编码模型,对所述多个识别文本进行向量转换,以生成初始转换向量,其中,所述初始转换向量表征所述多个识别文本的文本特征信息;

第一输入单元,被配置成将所述样本实体输入至初始实体编码模型,以生成初始实体编码向量;

第二输入单元,被配置成将所述初始转换向量和所述初始实体编码向量输入至初始匹配概率生成模型,以生成初始匹配概率,其中,所述初始匹配概率表征所述初始转换向量与所述初始实体编码向量之间的匹配关系;

参数训练单元,被配置成响应于确定初始匹配结果与所述样本标签不同,对所述初始文本编码模型、所述初始实体编码模型和所述初始匹配概率生成模型进行参数训练,得到文本编码模型、实体编码模型和匹配概率生成模型,其中,所述初始匹配结果是基于所述初始匹配概率生成的。

8. 一种实体确定装置,包括:

第二获取单元,被配置成获取针对目标音频的多个识别文本;

第二向量转换单元,被配置成利用预先训练的文本编码模型,对所述多个识别文本进行向量转换,以生成转换向量,其中,所述转换向量表征所述多个识别文本的特征信息,所述文本编码模型是根据权利要求1-3中任一所述的方法训练的;

实体编码单元,被配置成利用预先训练的实体编码模型,对实体库中的各个实体进行实体编码,得到实体编码向量集,其中,所述实体编码模型是根据权利要求1-3中任一所述的方法训练的;

实体筛选单元,被配置成利用预先训练的匹配概率生成模型,从所述实体库中筛选出对应匹配概率满足预设条件的实体,作为目标实体,得到至少一个目标实体,其中,目标实体对应的匹配概率是基于对应实体编码向量和所述转换向量生成的,所述匹配概率生成模型是根据权利要求1-3中任一所述的方法训练的。

9. 一种电子设备,包括:

一个或多个处理器;

存储装置,其上存储有一个或多个程序,

当所述一个或多个程序被所述一个或多个处理器执行,使得所述一个或多个处理器实现如权利要求1-6中任一所述的方法。

10.一种计算机可读介质,其上存储有计算机程序,其中,所述程序被处理器执行时实现如权利要求1-6中任一所述的方法。

## 模型训练方法、实体确定方法、装置、电子设备和介质

### 技术领域

[0001] 本公开的实施例涉及计算机技术领域，具体涉及模型训练方法、实体确定方法、装置、电子设备和介质。

### 背景技术

[0002] 目前，用户语音对话对应的实体检索技术为从实体库中检索出用户所提及的各个实体的技术。对于从目标音频中检索各个实体，通常采用的方式为：首先，将目标音频输入至语音识别模型，得到识别文本序列。然后，从识别文本序列中确定出文本置信度率最高的识别文本。最后，利用实体检索模型，对文本置信度最高的识别文本进行实体检索，得到与目标音频相关联的各个实体。

[0003] 然而，当采用上述方式来训练实体检索模型，经常会存在如下技术问题：

[0004] 语音识别模型输出的识别文本序列中识别率最高的识别文本可能存在文本识别错误的问题。由此，使用上述实体检索模型对识别率最高的识别文本进行实体检索，会存在检索错误的问题。

### 发明内容

[0005] 本公开的内容部分用于以简要的形式介绍构思，这些构思将在后面的具体实施方式部分被详细描述。本公开的内容部分并不旨在标识要求保护的技术方案的关键特征或必要特征，也不旨在用于限制所要求的保护的技术方案的范围。

[0006] 本公开的一些实施例提出了模型训练方法、实体确定方法、装置、电子设备和介质，来解决以上背景技术部分提到的技术问题。

[0007] 第一方面，本公开的一些实施例提供了一种模型训练方法，包括：获取训练样本和上述训练样本对应的样本标签，其中，上述训练样本包括：针对音频样本的多个识别文本和上述音频样本中涉及的样本实体；利用初始文本编码模型，对上述多个识别文本进行向量转换，以生成初始转换向量，其中，上述初始转换向量表征上述多个识别文本的文本特征信息；将上述样本实体输入至初始实体编码模型，以生成初始实体编码向量；将上述初始转换向量和上述初始实体编码向量输入至初始匹配概率生成模型，以生成初始匹配概率，其中，上述初始匹配概率表征上述初始转换向量与上述初始实体编码向量之间的匹配关系；响应于确定初始匹配结果与上述样本标签不同，对上述初始文本编码模型、上述初始实体编码模型和上述初始匹配概率生成模型进行参数训练，得到文本编码模型、实体编码模型和匹配概率生成模型，其中，上述初始匹配结果是基于上述初始匹配概率生成的。

[0008] 可选地，上述利用初始文本编码模型，对上述多个识别文本进行向量转换，以生成初始转换向量，包括：对上述多个识别文本进行文本拼接，得到拼接后文本；将上述拼接后文本输入至上述初始文本编码模型，得到上述初始转换向量。

[0009] 可选地，上述利用初始文本编码模型，对上述多个识别文本进行向量转换，以生成初始转换向量，包括：将上述多个识别文本输入至上述初始文本编码模型，得到多个初始文

本编码向量;对上述多个初始文本编码向量进行向量组合,得到组合矩阵;对上述组合矩阵进行池化操作,得到上述初始转换向量。

[0010] 第二方面,本公开的一些实施例提供了一种实体确定方法,包括:获取针对目标音频的多个识别文本;利用预先训练的文本编码模型,对上述多个识别文本进行向量转换,以生成转换向量,其中,上述转换向量表征上述多个识别文本的特征信息,上述文本编码模型是根据第一方面中任一上述的方法训练的;利用预先训练的实体编码模型,对实体库中的各个实体进行实体编码,得到实体编码向量集,其中,上述实体编码模型是根据第一方面中任一上述的方法训练的;利用预先训练的匹配概率生成模型,从上述实体库中筛选出对应匹配概率满足预设条件的实体,作为目标实体,得到至少一个目标实体,其中,目标实体对应的匹配概率是基于对应实体编码向量和上述转换向量生成的,上述匹配概率生成模型是根据第一方面中任一上述的方法训练的。

[0011] 可选地,上述利用预先训练的匹配概率生成模型,从上述实体库中筛选出对应匹配概率满足预设条件的实体,作为目标实体,包括:确定上述多个识别文本对应的关键词集;从上述实体库中筛选出与关键词集存在关联关系的实体,作为关联实体,得到关联实体集;从上述实体编码向量集中筛选出与上述关联实体集中各个关联实体相对应的实体编码向量,作为目标实体编码向量,得到目标实体编码向量集;利用上述匹配概率生成模型,确定上述目标实体编码向量集中每个目标实体编码向量与上述转换向量之间的匹配概率;从上述关联实体集中筛选出对应匹配概率满足上述预设条件的关联实体,作为目标实体。

[0012] 可选地,上述利用预先训练的匹配概率生成模型,从上述实体库中筛选出对应匹配概率满足预设条件的实体,作为目标实体,包括:利用上述匹配概率生成模型,确定上述实体编码向量集中每个实体编码向量与上述转换向量之间的匹配概率;从上述实体库中筛选出对应匹配概率满足上述预设条件的实体,作为目标实体。

[0013] 第三方面,本公开的一些实施例提供了一种模型训练装置,包括:第一获取单元,被配置成获取训练样本和上述训练样本对应的样本标签,其中,上述训练样本包括:针对音频样本的多个识别文本和上述音频样本中涉及的样本实体;第一向量转换单元,被配置成利用初始文本编码模型,对上述多个识别文本进行向量转换,以生成初始转换向量,其中,上述初始转换向量表征上述多个识别文本的文本特征信息;第一输入单元,被配置成将上述样本实体输入至初始实体编码模型,以生成初始实体编码向量;第二输入单元,被配置成将上述初始转换向量和上述初始实体编码向量输入至初始匹配概率生成模型,以生成初始匹配概率,其中,上述初始匹配概率表征上述初始转换向量与上述初始实体编码向量之间的匹配关系;参数训练单元,被配置成响应于确定初始匹配结果与上述样本标签不同,对上述初始文本编码模型、上述初始实体编码模型和上述初始匹配概率生成模型进行参数训练,得到文本编码模型、实体编码模型和匹配概率生成模型,其中,上述初始匹配结果是基于上述初始匹配概率生成的。

[0014] 可选地,第一向量转换单元进一步被配置成:对上述多个识别文本进行文本拼接,得到拼接后文本;将上述拼接后文本输入至上述初始文本编码模型,得到上述初始转换向量。

[0015] 可选地,第一向量转换单元进一步被配置成:将上述多个识别文本输入至上述初始文本编码模型,得到多个初始文本编码向量;对上述多个初始文本编码向量进行向量组

合,得到组合矩阵;对上述组合矩阵进行池化操作,得到上述初始转换向量。

[0016] 第四方面,本公开的一些实施例提供了一种实体确定装置,包括:第二获取单元,被配置成获取针对目标音频的多个识别文本;第二向量转换单元,被配置成利用预先训练的文本编码模型,对上述多个识别文本进行向量转换,以生成转换向量,其中,上述转换向量表征上述多个识别文本的特征信息,上述文本编码模型是根据第一方面中任一上述的方法训练的;实体编码单元,被配置成利用预先训练的实体编码模型,对实体库中的各个实体进行实体编码,得到实体编码向量集,其中,上述实体编码模型是根据第一方面中任一上述的方法训练的;实体筛选单元,被配置成利用预先训练的匹配概率生成模型,从上述实体库中筛选出对应匹配概率满足预设条件的实体,作为目标实体,得到至少一个目标实体,其中,目标实体对应的匹配概率是基于对应实体编码向量和上述转换向量生成的,上述匹配概率生成模型是根据第一方面中任一上述的方法训练的。

[0017] 可选地,实体筛选单元进一步被配置成:确定上述多个识别文本对应的关键词集;从上述实体库中筛选出与关键词集存在关联关系的实体,作为关联实体,得到关联实体集;从上述实体编码向量集中筛选出与上述关联实体集中各个关联实体相对应的实体编码向量,作为目标实体编码向量,得到目标实体编码向量集;利用上述匹配概率生成模型,确定上述目标实体编码向量集中每个目标实体编码向量与上述转换向量之间的匹配概率;从上述关联实体集中筛选出对应匹配概率满足上述预设条件的关联实体,作为目标实体。

[0018] 可选地,实体筛选单元进一步被配置成:利用上述匹配概率生成模型,确定上述实体编码向量集中每个实体编码向量与上述转换向量之间的匹配概率;从上述实体库中筛选出对应匹配概率满足上述预设条件的实体,作为目标实体。

[0019] 第五方面,本公开的一些实施例提供了一种电子设备,包括:一个或多个处理器;存储装置,其上存储有一个或多个程序,当一个或多个程序被一个或多个处理器执行,使得一个或多个处理器实现如第一方面或第二方面中任一实现方式描述的方法。

[0020] 第六方面,本公开的一些实施例提供了一种计算机可读介质,其上存储有计算机程序,其中,程序被处理器执行时实现如第一方面或第二方面中任一实现方式描述的方法。

[0021] 本公开的上述各个实施例具有如下有益效果:通过本公开的一些实施例的模型训练方法得到的文本编码模型、实体编码模型和匹配概率生成模型,检索实体的精确度有所提高。具体来说,造成相关的实体检索不够精确的原因在于:语音识别模型可能存在文本识别错误的问题,导致后续实体检索会存在检索错误的问题。基于此,本公开的一些实施例的模型训练方法可以首先获取训练样本和样本标签。其中,上述训练样本包括:针对音频样本的多个识别文本和上述音频样本中涉及的样本实体。在这里,多个识别文本为音频样本的多个识别结果。多个识别结果中对应识别概率最高的识别结果可能也存在识别错误的文本。而对于识别概率最高的识别结果的错误信息,其余识别结果集中可能存在同样位置文本识别正确的识别结果。由此,将多个识别文本和音频样本中所涉及的样本实体作为训练样本,可以使得后续匹配概率生成模型的训练过程中考量多个识别文本中更多的文本特征信息。使得后续基于匹配概率生成模型的实体识别更为精准。然后,利用初始文本编码模型,对上述多个识别文本进行向量转换,以生成初始转换向量。其中,上述初始转换向量表征上述多个识别文本的文本特征信息。在这里,通过多个识别文本的向量转换,使得初始转换向量包括多个识别文本的文本特征信息。便于后续匹配概率生成模型的训练过程中考量

多个识别文本中更多的文本特征信息。接着,利用初始实体编码模型,样本实体对应的初始实体编码向量。以便于后续输入初始匹配概率生成模型进行初始匹配概率计算。进而,将上述初始转换向量和上述初始实体编码向量输入至初始匹配概率生成模型,以使得初始匹配概率生成模型依据多个识别文本的文本特征信息来生成更为精准地初始匹配概率。其中,上述初始匹配概率表征上述初始转换向量与上述初始实体编码向量之间的匹配关系。最后,响应于确定初始匹配结果与上述样本标签不同,对上述初始文本编码模型、上述初始实体编码模型和上述初始匹配概率生成模型进行参数训练,得到文本编码模型、实体编码模型和匹配概率生成模型,以使得文本编码模型、实体编码模型和匹配概率生成模型的输出越来越精准。其中,上述初始匹配结果是基于上述初始匹配概率生成的。由此,基于上述模型训练方法训练出的文本编码模型、实体编码模型和匹配概率生成模型来确定实体更为精准。

### 附图说明

[0022] 结合附图并参考以下具体实施方式,本公开各实施例的上述和其他特征、优点及方面将变得更加明显。贯穿附图中,相同或相似的附图标记表示相同或相似的元素。应当理解附图是示意性的,元件和元素不一定按照比例绘制。

[0023] 图1是根据本公开的一些实施例的模型训练方法的一个应用场景的示意图;

[0024] 图2是根据本公开的模型训练方法的一些实施例的流程图;

[0025] 图3是根据本公开的模型训练方法的另一些实施例的流程图;

[0026] 图4是根据本公开的一些实施例的实体确定方法的一个应用场景的示意图;

[0027] 图5是根据本公开的实体确定方法的一些实施例的流程图;

[0028] 图6是根据本公开的模型训练装置的一些实施例的结构示意图;

[0029] 图7是根据本公开的实体确定装置的一些实施例的结构示意图;

[0030] 图8是适于用来实现本公开的一些实施例的电子设备的结构示意图。

### 具体实施方式

[0031] 下面将参照附图更详细地描述本公开的实施例。虽然附图中显示了本公开的某些实施例,然而应当理解的是,本公开可以通过各种形式来实现,而且不应该被解释为限于这里阐述的实施例。相反,提供这些实施例是为了更加透彻和完整地理解本公开。应当理解的是,本公开的附图及实施例仅用于示例性作用,并非用于限制本公开的保护范围。

[0032] 另外还需要说明的是,为了便于描述,附图中仅示出了与有关发明相关的部分。在不冲突的情况下,本公开中的实施例及实施例中的特征可以相互组合。

[0033] 需要注意,本公开中提及的“第一”、“第二”等概念仅用于对不同的装置、模块或单元进行区分,并非用于限定这些装置、模块或单元所执行的功能的顺序或者相互依存关系。

[0034] 需要注意,本公开中提及的“一个”、“多个”的修饰是示意性而非限制性的,本领域技术人员应当理解,除非在上下文另有明确指出,否则应该理解为“一个或多个”。

[0035] 本公开实施方式中的多个装置之间所交互的消息或者信息的名称仅用于说明性的目的,而并不是用于对这些消息或信息的范围进行限制。

[0036] 下面将参考附图并结合实施例来详细说明本公开。

[0037] 图1是根据本公开一些实施例的模型训练方法的一个应用场景的示意图。

[0038] 在图1的应用场景中,首先,电子设备101可以获取训练样本102和上述训练样本102对应的样本标签113。其中,上述训练样本102包括:针对音频样本103的多个识别文本104和上述音频样本103中涉及的样本实体105。在本应用场景中,多个识别文本104可以包括:识别文本1041、识别文本1042、识别文本1043。其中,上述识别文本1041可以是“播放歌曲成都”。上述识别文本1042可以是“播放歌曲尘土”。上述识别文本1043可以是“播放歌曲程度”。上述样本实体105可以是“成都”。样本标签113可以是“1”。然后,电子设备101可以利用初始文本编码模型106,对上述多个识别文本104进行向量转换,以生成初始转换向量107。其中,上述初始转换向量107表征上述多个识别文本104的文本特征信息。接着,电子设备101可以将上述样本实体105输入至初始实体编码模型108,以生成初始实体编码向量109。进而,电子设备101可以将上述初始转换向量107和上述初始实体编码向量109输入至初始匹配概率生成模型110,以生成初始匹配概率111。其中,上述初始匹配概率111表征上述初始转换向量107与上述初始实体编码向量109之间的匹配关系。在本应用场景中,上述初始匹配概率111可以是“80%”。最后,响应于确定初始匹配结果112与上述样本标签113不同,电子设备101可以对上述初始文本编码模型106、上述初始实体编码模型108和上述初始匹配概率生成模型110进行参数训练,得到文本编码模型、实体编码模型和匹配概率生成模型。其中,上述初始匹配结果112是基于上述初始匹配概率111生成的。在本应用场景中,上述初始匹配结果112可以是“0”。

[0039] 需要说明的是,上述电子设备101可以是硬件,也可以是软件。当电子设备为硬件时,可以实现成多个服务器或终端设备组成的分布式集群,也可以实现成单个服务器或单个终端设备。当电子设备体现为软件时,可以安装在上述所列举的硬件设备中。其可以实现成例如用来提供分布式服务的多个软件或软件模块,也可以实现成单个软件或软件模块。在此不做具体限定。

[0040] 应该理解,图1中的电子设备的数目仅仅是示意性的。根据实现需要,可以具有任意数目的电子设备。

[0041] 继续参考图2,示出了根据本公开的模型训练方法的一些实施例的流程200。该模型训练方法,包括以下步骤:

[0042] 步骤201,获取训练样本和上述训练样本对应的样本标签。

[0043] 在一些实施例中,上述模型训练方法的执行主体(例如图1所示的电子设备101)可以通过有线连接方式或者无线连接方式来获取训练样本和上述训练样本对应的样本标签。其中,上述训练样本包括:针对音频样本的多个识别文本和上述音频样本中涉及的样本实体。上述训练样本可以是用于后续训练初始文本编码模型、初始实体编码模型和初始匹配概率生成模型的样本。上述音频样本可以是待文字识别的音频。上述多个识别文本为音频样本的多个文字识别结果。上述样本实体为音频样本对应音频内容中所提及的实体。其中,实体(Entity)是音频样本对应音频内容中所涉及的命名实体。上述命名实体可以是音频内容中被识别的目标。例如,被识别的目标可以是音频内容中的人名,地名,组织名称,日期等。例如,音频内容为“播放歌曲成都”。样本实体可以为“成都”。上述样本标签可以是预先标注的标签。样本标签可以是各种形式的标识。例如,样本标签可以是数字标识“1”或“0”。样本标签为1时,可以表征样本实体为音频样本对应音频内容中所提及的实体。相应的,样

本标签为0时,可以表征样本实体不是音频样本对应音频内容中所提及的实体。

[0044] 可选地,将上述音频样本输入至音频识别模块,可以得到多个识别文本。即,音频识别模块可以输出针对音频样本的多个识别文本。其中,上述音频识别模块为将音频转换为识别文本的模块。即,音频识别模块为采用了自动语音识别技术(ASR, Automatic Speech Recognition)的模块。

[0045] 每个识别文本存在对应的置信度。置信度可以表征音频样本与识别文本之间的关联关系。置信度越高,表征识别文本更能代表音频样本的音频内容。相反,置信度越低,表征识别文本代表音频样本的音频内容的程度越低。

[0046] 需要说明的是,上述多个识别文本中可能存在识别音频样本的音频内容正确的文本,也可能都是识别错误的样本。

[0047] 可选地,上述多个识别文本还可以是多个音频识别模块的输出文本。即,将音频样本输入至多个音频识别模块,输出多个识别文本。

[0048] 步骤202,利用初始文本编码模型,对上述多个识别文本进行向量转换,以生成初始转换向量。

[0049] 在一些实施例中,上述执行主体可以利用初始文本编码模型,对上述多个识别文本进行向量转换,以生成初始转换向量。上述初始文本编码模型可以是对文本进行向量编码的模型。其中,上述初始转换向量表征上述多个识别文本的文本特征信息。上述初始文本编码模型可以是但不限于以下之一:BERT(Bidirectional Encoder Representation from Transformers)模型和TextCNN模型。上述初始文本编码模型为参数初始化后的文本编码模型。

[0050] 作为示例,首先,上述执行主体可以将多个识别文本输入至初始文本编码模型,得到多个初始文本编码向量。然后,上述执行主体可以将多个初始文本编码向量进行向量对应元素求平均值处理,得到处理后向量,作为初始转换向量。

[0051] 在一些实施例的一些可选的实现方式中,上述利用初始文本编码模型,对上述多个识别文本进行向量转换,以生成初始转换向量,可以包括以下步骤:

[0052] 第一步,上述执行主体可以对上述多个识别文本进行文本拼接,得到拼接后文本。

[0053] 实践中,上述执行主体可以通过插入文本标识符,来实现对多个识别文本的文本拼接,得到拼接后文本。

[0054] 作为示例,多个识别文本可以为{“播放歌曲程度”,“播放歌曲成都”,“播放歌曲尘土”。文本标识符可以是“[SEP]”。拼接后文本可以是“播放歌曲程度[SEP]播放歌曲程度[SEP]播放歌曲尘土”。

[0055] 第二步,上述执行主体可以将上述拼接后文本输入至上述初始文本编码模型,得到上述初始转换向量。

[0056] 步骤203,将上述样本实体输入至初始实体编码模型,以生成初始实体编码向量。

[0057] 在一些实施例中,上述执行主体可以将上述样本实体输入至初始实体编码模型,以生成初始实体编码向量。其中,上述初始实体编码模型可以是对实体进行向量编码的模型。上述初始实体编码模型为参数初始化后的实体编码模型。上述初始实体编码向量可以表征样本实体的实体特征信息。实体编码模型可以是但不限于以下之一:word2vec模型和TextCNN模型。

[0058] 步骤204,将上述初始转换向量和上述初始实体编码向量输入至初始匹配概率生成模型,以生成初始匹配概率。

[0059] 在一些实施例中,上述执行主体可以将上述初始转换向量和上述初始实体编码向量输入至初始匹配概率生成模型,以生成初始匹配概率。其中,上述初始匹配概率表征上述初始转换向量与上述初始实体编码向量之间的匹配度。初始匹配概率越高,表征初始转换向量与初始实体编码向量更匹配。上述初始匹配概率生成模型可以是用于确定初始转换向量和初始实体编码向量的匹配度,以生成匹配概率的模型。上述初始实体编码模型为参数初始化后的实体编码模型。匹配概率生成模型可以是但不限于以下之一:深度残差网络(Deep residual network,ResNet)和VGG(Visual Geometry Group)模型。

[0060] 步骤205,响应于确定初始匹配结果与上述样本标签不同,对上述初始文本编码模型、上述初始实体编码模型和上述初始匹配概率生成模型进行参数训练。

[0061] 在一些实施例中,响应于确定初始匹配结果与上述样本标签不同,上述执行主体可以对上述初始文本编码模型、上述初始实体编码模型和上述初始匹配概率生成模型进行参数训练,得到文本编码模型、实体编码模型和匹配概率生成模型。其中,上述初始匹配结果是基于上述初始匹配概率生成的。

[0062] 作为示例,初始匹配结果可以通过以下步骤生成:

[0063] 第一步,确定初始匹配概率是否大于等于预设概率值。

[0064] 第二步,响应于确定初始匹配概率大于等于预设概率值,生成表征初始实体编码向量与初始转换向量相匹配的初始匹配结果。

[0065] 第三步,响应于确定初始匹配概率小于预设概率值,生成表征初始实体编码向量与初始转换向量不匹配的初始匹配结果。

[0066] 本公开的上述各个实施例具有如下有益效果:通过本公开的一些实施例的模型训练方法得到的文本编码模型、实体编码模型和匹配概率生成模型,检索实体的精确度有所提高。具体来说,造成相关的实体检索不够精确的原因在于:语音识别模型可能存在文本识别错误的问题,导致后续实体检索会存在检索错误的问题。基于此,本公开的一些实施例的模型训练方法可以首先获取训练样本和样本标签。其中,上述训练样本包括:针对音频样本的多个识别文本和上述音频样本中涉及的样本实体。在这里,多个识别文本为音频样本的多个识别结果。多个识别结果中对应识别概率最高的识别结果可能也存在识别错误的文本。而对于识别概率最高的识别结果的错误信息,其余识别结果集中可能存在同样位置文本识别正确的识别结果。由此,将多个识别文本和音频样本中所涉及的样本实体作为训练样本,可以使得后续匹配概率生成模型的训练过程中考量多个识别文本中更多的文本特征信息。使得后续基于匹配概率生成模型的实体识别更为精准。然后,利用初始文本编码模型,对上述多个识别文本进行向量转换,以生成初始转换向量。其中,上述初始转换向量表征上述多个识别文本的文本特征信息。在这里,通过多个识别文本的向量转换,使得初始转换向量包括多个识别文本的文本特征信息。便于后续匹配概率生成模型的训练过程中考量多个识别文本中更多的文本特征信息。接着,利用初始实体编码模型,样本实体对应的初始实体编码向量。以便于后续输入初始匹配概率生成模型进行初始匹配概率计算。进而,将上述初始转换向量和上述初始实体编码向量输入至初始匹配概率生成模型,以使得初始匹配概率生成模型依据多个识别文本的文本特征信息来生成更为精准地初始匹配概率。其中,

上述初始匹配概率表征上述初始转换向量与上述初始实体编码向量之间的匹配关系。最后,响应于确定初始匹配结果与上述样本标签不同,对上述初始文本编码模型、上述初始实体编码模型和上述初始匹配概率生成模型进行参数训练,得到文本编码模型、实体编码模型和匹配概率生成模型,以使得文本编码模型、实体编码模型和匹配概率生成模型的输出越来越精准。其中,上述初始匹配结果是基于上述初始匹配概率生成的。由此,基于上述模型训练方法训练出的文本编码模型、实体编码模型和匹配概率生成模型来确定实体更为精准。

[0067] 进一步参考图3,示出了根据本公开的模型训练方法的另一些实施例的流程300。该模型训练方法,包括以下步骤:

[0068] 步骤301,获取训练样本和上述训练样本对应的样本标签。

[0069] 步骤302,将上述多个识别文本输入至上述初始文本编码模型,得到多个初始文本编码向量。

[0070] 在一些实施例中,执行主体(例如图1所示的电子设备101)可以将上述多个识别文本输入至上述初始文本编码模型,得到多个初始文本编码向量。

[0071] 步骤303,对上述多个初始文本编码向量进行向量组合,得到组合矩阵。

[0072] 在一些实施例中,上述执行主体可以对上述多个初始文本编码向量进行向量组合,得到组合矩阵。

[0073] 作为示例,初始文本编码的向量维度可以是 $1*d$ 。其中, $d$ 为初始文本编码向量的列数。上述执行主体可以依据识别文本对应置信度由大到小的顺序,自上而下的对多个初始文本向量进行向量组合,所得到的组合矩阵的矩阵为 $n*d$ 。其中, $n$ 为多个初始文本向量中的初始文本向量的个数。即,组合矩阵中每行向量为初始文本编码向量。

[0074] 步骤304,对上述组合矩阵进行池化操作,得到上述初始转换向量。

[0075] 在一些实施例中,上述执行主体可以对上述组合矩阵进行池化操作,得到上述初始转换向量。上述池化操作可以是但不限于以下之一:平均池化操作和最大池化操作。

[0076] 步骤305,将上述样本实体输入至初始实体编码模型,以生成初始实体编码向量。

[0077] 步骤306,将上述初始转换向量和上述初始实体编码向量输入至初始匹配概率生成模型,以生成初始匹配概率。

[0078] 步骤307,响应于确定初始匹配结果与上述样本标签不同,对上述初始文本编码模型、上述初始实体编码模型和上述初始匹配概率生成模型进行参数训练,得到文本编码模型、实体编码模型和匹配概率生成模型。

[0079] 在一些实施例中,步骤301、305-307的具体实现及其所带来的技术效果,可以参考图2对应的实施例中的步骤201、203-205,在此不再赘述。

[0080] 从图3中可以看出,与图2对应的一些实施例的描述相比,图3对应的一些实施例中的模型训练方法的流程300更加突出了利用初始文本编码模型来生成初始转换向量的具体步骤。由此,这些实施例描述的方案可以更为精准的生成更能代表多个识别文本的文本特征信息的初始转换向量,以使得后续模型的训练更为精准。

[0081] 图4是根据本公开一些实施例的实体确定方法的一个应用场景的示意图。

[0082] 在图4的应用场景中,首先,电子设备401可以获取针对目标音频402的多个识别文本403。在本应用场景中,上述多个识别文本403可以包括:识别文本4031、识别文本4032和

识别文本4033。其中,上述识别文本4031可以是“播放歌曲成都”。上述识别文本4032可以是“播放歌曲尘土”。上述识别文本4033可以是“播放歌曲程度”。然后,电子设备401可以利用预先训练的文本编码模型404,对上述多个识别文本403进行向量转换,以生成转换向量405。其中,上述转换向量405表征上述多个识别文本403的特征信息,上述文本编码模型404是通过本公开一些实施例的模型训练方法生成的。接着,电子设备401可以利用预先训练的实体编码模型407,对实体库406中的各个实体进行实体编码,得到实体编码向量集408。其中,上述实体编码模型408是通过本公开一些实施例的模型训练方法生成的。在本应用场景中,实体库406中的各个实体包括:实体4061、实体4062和实体4063。实体4061可以是“程度”。实体4062可以是“成都”。实体4063可以是“尘土”。实体编码向量集408可以包括:实体4061对应的实体编码向量4081、实体4062对应的实体编码向量4082和实体4063对应的实体编码向量4083。最后,电子设备401可以利用预先训练的匹配概率生成模型409,从上述实体库406中筛选出对应匹配概率满足预设条件的实体,作为目标实体,得到至少一个目标实体。其中,目标实体对应的匹配概率是基于对应实体编码向量和上述转换向量生成的,上述匹配概率生成模型是通过本公开一些实施例的模型训练方法生成的。在本应用场景中,将实体编码向量4081和转换向量405输入至匹配概率生成模型409,得到匹配概率4101。匹配概率4101可以是50%。将实体编码向量4082和转换向量405输入至匹配概率生成模型409,得到匹配概率4102。匹配概率4102可以是90%。将实体编码向量4083和转换向量405输入至匹配概率生成模型409,得到匹配概率4103。匹配概率4103可以是60%。将实体库406中对应匹配概率大于70%的实体确定为目标实体。即目标实体为实体4062。

[0083] 需要说明的是,上述电子设备401可以是硬件,也可以是软件。当电子设备为硬件时,可以实现成多个服务器或终端设备组成的分布式集群,也可以实现成单个服务器或单个终端设备。当电子设备体现为软件时,可以安装在上述所列举的硬件设备中。其可以实现成例如用来提供分布式服务的多个软件或软件模块,也可以实现成单个软件或软件模块。在此不做具体限定。

[0084] 应该理解,图4中的电子设备的数目仅仅是示意性的。根据实现需要,可以具有任意数目的电子设备。

[0085] 继续参考图5,示出了根据本公开的实体确定方法的一些实施例的流程500。该实体确定方法,包括以下步骤:

[0086] 步骤501,获取针对目标音频的多个识别文本。

[0087] 在一些实施例中,上述实体确定方法的执行主体(例如图4所示的电子设备401)可以通过有线连接方式或者无线连接方式来获取针对目标音频的多个识别文本。其中,目标音频为待识别实体的音频。目标音频的多个识别文本在此不再赘述,具体解释可以参见步骤201中音频样本的多个识别样本。

[0088] 步骤502,利用预先训练的文本编码模型,对上述多个识别文本进行向量转换,以生成转换向量。

[0089] 在一些实施例中,上述执行主体可以利用预先训练的文本编码模型,对上述多个识别文本进行向量转换,以生成转换向量。其中,上述转换向量表征上述多个识别文本的特征信息,上述文本编码模型是通过本公开一些实施例的模型训练方法生成的。上述预先训练的文本编码模型可以是模型参数更新后的模型。

[0090] 作为示例,上述执行主体利用预先训练的文本编码模型,对上述多个识别文本进行向量转换,以生成转换向量,可以包括以下步骤:

[0091] 第一步,上述执行主体可以对上述多个识别文本进行文本拼接,得到拼接后文本。

[0092] 第二步,上述执行主体可以将上述拼接后文本输入至上述预先训练的文本编码模型,得到上述转换向量。

[0093] 作为又一个示例,上述执行主体利用预先训练的文本编码模型,对上述多个识别文本进行向量转换,以生成转换向量,可以包括以下步骤:

[0094] 第一步,上述执行主体可以将上述多个识别文本输入至上述初始文本编码模型,得到多个初始文本编码向量。

[0095] 第二步,上述执行主体可以对上述多个初始文本编码向量进行向量组合,得到组合矩阵。

[0096] 第三步,上述执行主体可以对上述组合矩阵进行池化操作,得到上述初始转换向量。

[0097] 步骤503,利用预先训练的实体编码模型,对实体库中的各个实体进行实体编码,得到实体编码向量集。

[0098] 在一些实施例中,上述执行主体可以利用预先训练的实体编码模型,对实体库中的各个实体进行实体编码,得到实体编码向量集。其中,上述实体编码模型是通过本公开一些实施例的模型训练方法生成的。上述预先训练的实体编码模型可以是模型参数更新后的模型。

[0099] 步骤504,利用预先训练的匹配概率生成模型,从上述实体库中筛选出对应匹配概率满足预设条件的实体,作为目标实体,得到至少一个目标实体。

[0100] 在一些实施例中,上述执行主体可以利用预先训练的匹配概率生成模型,从上述实体库中筛选出对应匹配概率满足预设条件的实体,作为目标实体,得到至少一个目标实体。其中,目标实体对应的匹配概率是基于对应实体编码向量和上述转换向量生成的。上述匹配概率生成模型是通过本公开一些实施例的模型训练方法生成的。上述目标实体为模型预测出的、上述目标音频的音频内容中所涉及的实体。

[0101] 在一些实施例的一些可选的实现方式中,上述利用预先训练的匹配概率生成模型,从上述实体库中筛选出对应匹配概率满足预设条件的实体,作为目标实体,可以包括以下步骤:

[0102] 第一步,上述执行主体可以确定上述多个识别文本对应的关键词集。

[0103] 作为示例,上述执行主体可以对多个识别文本进行文本分词处理,得到词集。然后,从上述词集中确定出与目标领域(例如,电商领域)相关联的关键词,得到关键词集。上述目标领域可以是预先确定的行业领域。

[0104] 第二步,上述执行主体可以从上述实体库中筛选出与关键词集存在关联关系的实体,作为关联实体,得到关联实体集。

[0105] 作为示例,首先,上述执行主体可以确定关键词集中每个关键词对应的至少一个同音词,得到同音词集。然后,上述执行主体可以将同音词集和关键词集进行合并,得到合并后词集,作为关联实体集。

[0106] 第三步,上述执行主体可以从上述实体编码向量集中筛选出与上述关联实体集中

各个关联实体相对应的实体编码向量,作为目标实体编码向量,得到目标实体编码向量集。

[0107] 第四步,上述执行主体可以利用上述匹配概率生成模型,确定上述目标实体编码向量集中每个目标实体编码向量与上述转换向量之间的匹配概率。

[0108] 作为示例,上述执行主体可以将目标实体编码向量集中每个目标实体编码向量和上述转换向量输入至匹配概率生成模型,得到匹配概率。

[0109] 第五步,上述执行主体可以从上述关联实体集中筛选出对应匹配概率满足上述预设条件的关联实体,作为目标实体。其中,上述预设条件可以是匹配概率大于等于预设概率值。例如,预设概率值为70%。

[0110] 在一些实施例的一些可选的实现方式中,上述利用预先训练的匹配概率生成模型,从上述实体库中筛选出对应匹配概率满足预设条件的实体,作为目标实体,可以包括以下步骤:

[0111] 第一步,上述执行主体可以利用上述匹配概率生成模型,确定上述实体编码向量集中每个实体编码向量与上述转换向量之间的匹配概率。

[0112] 第二步,上述执行主体可以从上述实体库中筛选出对应匹配概率满足上述预设条件的实体,作为目标实体。其中,上述预设条件可以是匹配概率大于等于预设概率值。例如,预设概率值为70%。

[0113] 本公开的上述各个实施例具有如下有益效果:通过本公开的一些实施例的实体确定方法通过利用多个识别文本的特征信息、预先训练的文本编码模型、实体检测模型和匹配概率生成模型可以精准地确定出目标音频对应音频内容所涉及的至少一个实体。

[0114] 进一步参考图6,作为对上述各图所示方法的实现,本公开提供了一种模型训练装置的一些实施例,这些装置实施例与图2所示的那些方法实施例相对应,该装置具体可以应用于各种电子设备中。

[0115] 如图6所示,一种模型训练装置600包括:第一获取单元601、第一向量转换单元602、第一输入单元603、第二输入单元604和参数训练单元605。其中,第一获取单元601,被配置成获取训练样本和上述训练样本对应的样本标签,其中,上述训练样本包括:针对音频样本的多个识别文本和上述音频样本中涉及的样本实体;第一向量转换单元602,被配置成利用初始文本编码模型,对上述多个识别文本进行向量转换,以生成初始转换向量,其中,上述初始转换向量表征上述多个识别文本的文本特征信息;第一输入单元603,被配置成将上述样本实体输入至初始实体编码模型,以生成初始实体编码向量;第二输入单元604,被配置成将上述初始转换向量和上述初始实体编码向量输入至初始匹配概率生成模型,以生成初始匹配概率,其中,上述初始匹配概率表征上述初始转换向量与上述初始实体编码向量之间的匹配关系;参数训练单元605,被配置成响应于确定初始匹配结果与上述样本标签不同,对上述初始文本编码模型、上述初始实体编码模型和上述初始匹配概率生成模型进行参数训练,得到文本编码模型、实体编码模型和匹配概率生成模型,其中,上述初始匹配结果是基于上述初始匹配概率生成的。

[0116] 在一些实施例的一些可选的实现方式中,上述第一向量转换单元602可以进一步被配置成:对上述多个识别文本进行文本拼接,得到拼接后文本;将上述拼接后文本输入至上述初始文本编码模型,得到上述初始转换向量。

[0117] 在一些实施例的一些可选的实现方式中,上述第一向量转换单元602可以进一步

被配置成:将上述多个识别文本输入至上述初始文本编码模型,得到多个初始文本编码向量;对上述多个初始文本编码向量进行向量组合,得到组合矩阵;对上述组合矩阵进行池化操作,得到上述初始转换向量。

[0118] 可以理解的是,该装置600中记载的诸单元与参考图2描述的方法中的各个步骤相对应。由此,上文针对方法描述的操作、特征以及产生的有益效果同样适用于装置600及其中包含的单元,在此不再赘述。

[0119] 进一步参考图7,作为对上述各图所示方法的实现,本公开提供了一种实体确定装置的一些实施例,这些装置实施例与图5所示的那些方法实施例相对应,该装置具体可以应用于各种电子设备中。

[0120] 如图7所示,一种实体确定装置700包括:第二获取单元701、第二向量转换单元702、实体编码单元703和实体筛选单元704。其中,第二获取单元701,被配置成获取针对目标音频的多个识别文本;第二向量转换单元702,被配置成利用预先训练的文本编码模型,对上述多个识别文本进行向量转换,以生成转换向量,其中,上述转换向量表征上述多个识别文本的特征信息,上述文本编码模型是通过本公开一些实施例的模型训练方法生成的;实体编码单元703,被配置成利用预先训练的实体编码模型,对实体库中的各个实体进行实体编码,得到实体编码向量集,其中,上述实体编码模型是通过本公开一些实施例的模型训练方法生成的;实体筛选单元704,被配置成利用预先训练的匹配概率生成模型,从上述实体库中筛选出对应匹配概率满足预设条件的实体,作为目标实体,得到至少一个目标实体,其中,目标实体对应的匹配概率是基于对应实体编码向量和上述转换向量生成的,上述匹配概率生成模型是通过本公开一些实施例的模型训练方法生成的。

[0121] 在一些实施例的一些可选的实现方式中,上述实体筛选单元704可以进一步被配置成:确定上述多个识别文本对应的关键词集;从上述实体库中筛选出与关键词集存在关联关系的实体,作为关联实体,得到关联实体集;从上述实体编码向量集中筛选出与上述关联实体集中各个关联实体相对应的实体编码向量,作为目标实体编码向量,得到目标实体编码向量集;利用上述匹配概率生成模型,确定上述目标实体编码向量集中每个目标实体编码向量与上述转换向量之间的匹配概率;从上述关联实体集中筛选出对应匹配概率满足上述预设条件的关联实体,作为目标实体。

[0122] 在一些实施例的一些可选的实现方式中,上述实体筛选单元704可以进一步被配置成:利用上述匹配概率生成模型,确定上述实体编码向量集中每个实体编码向量与上述转换向量之间的匹配概率;从上述实体库中筛选出对应匹配概率满足上述预设条件的实体,作为目标实体。

[0123] 可以理解的是,该装置700中记载的诸单元与参考图5描述的方法中的各个步骤相对应。由此,上文针对方法描述的操作、特征以及产生的有益效果同样适用于装置700及其中包含的单元,在此不再赘述。

[0124] 下面参考图8,其示出了适于用来实现本公开的一些实施例的电子设备(例如图1中的电子设备101)800的结构示意图。图8示出的电子设备仅仅是一个示例,不应对本公开的实施例的功能和使用范围带来任何限制。

[0125] 如图8所示,电子设备800可以包括处理装置(例如中央处理器、图形处理器等)801,其可以根据存储在只读存储器(ROM)802中的程序或者从存储装置808加载到随机访问

存储器(RAM)803中的程序而执行各种适当的动作和处理。在RAM 803中,还存储有电子设备800操作所需的各种程序和数据。处理装置801、ROM 802以及RAM 803通过总线804彼此相连。输入/输出(I/O)接口805也连接至总线804。

[0126] 通常,以下装置可以连接至I/O接口805:包括例如触摸屏、触摸板、键盘、鼠标、摄像头、麦克风、加速度计、陀螺仪等的输入装置806;包括例如液晶显示器(LCD)、扬声器、振动器等的输出装置807;包括例如磁带、硬盘等的存储装置808;以及通信装置809。通信装置809可以允许电子设备800与其他设备进行无线或有线通信以交换数据。虽然图8示出了具有各种装置的电子设备800,但是应理解的是,并不要求实施或具备所有示出的装置。可以替代地实施或具备更多或更少的装置。图8中示出的每个方框可以代表一个装置,也可以根据需要代表多个装置。

[0127] 特别地,根据本公开的一些实施例,上文参考流程图描述的过程可以被实现为计算机软件程序。例如,本公开的一些实施例包括一种计算机程序产品,其包括承载在计算机可读介质上的计算机程序,该计算机程序包含用于执行流程图所示的方法的程序代码。在这样的一些实施例中,该计算机程序可以通过通信装置809从网络上被下载和安装,或者从存储装置808被安装,或者从ROM 802被安装。在该计算机程序被处理装置801执行时,执行本公开的一些实施例的方法中限定的上述功能。

[0128] 需要说明的是,本公开的一些实施例上述的计算机可读介质可以是计算机可读信号介质或者计算机可读存储介质或者是上述两者的任意组合。计算机可读存储介质例如可以是一——但不限于——电、磁、光、电磁、红外线、或半导体的系统、装置或器件,或者任意以上的组合。计算机可读存储介质的更具体的例子可以包括但不限于:具有一个或多个导线的电连接、便携式计算机磁盘、硬盘、随机访问存储器(RAM)、只读存储器(ROM)、可擦式可编程只读存储器(EPROM或闪存)、光纤、便携式紧凑磁盘只读存储器(CD-ROM)、光存储器件、磁存储器件、或者上述的任意合适的组合。在本公开的一些实施例中,计算机可读存储介质可以是任何包含或存储程序的有形介质,该程序可以被指令执行系统、装置或者器件使用或者与其结合使用。而在本公开的一些实施例中,计算机可读信号介质可以包括在基带中或者作为载波一部分传播的数据信号,其中承载了计算机可读的程序代码。这种传播的数据信号可以采用多种形式,包括但不限于电磁信号、光信号或上述的任意合适的组合。计算机可读信号介质还可以是计算机可读存储介质以外的任何计算机可读介质,该计算机可读信号介质可以发送、传播或者传输用于由指令执行系统、装置或者器件使用或者与其结合使用的程序。计算机可读介质上包含的程序代码可以用任何适当的介质传输,包括但不限于:电线、光缆、RF(射频)等等,或者上述的任意合适的组合。

[0129] 在一些实施方式中,客户端、服务器可以利用诸如HTTP(HyperText Transfer Protocol,超文本传输协议)之类的任何当前已知或未来研发的网络协议进行通信,并且可以与任意形式或介质的数字数据通信(例如,通信网络)互连。通信网络的示例包括局域网(“LAN”),广域网(“WAN”),网际网(例如,互联网)以及端对端网络(例如,ad hoc端对端网络),以及任何当前已知或未来研发的网络。

[0130] 上述计算机可读介质可以是上述电子设备中所包含的;也可以是单独存在,而未装配入该电子设备中。上述计算机可读介质承载有一个或者多个程序,当上述一个或者多个程序被该电子设备执行时,使得该电子设备:获取训练样本和上述训练样本对应的样本

标签,其中,上述训练样本包括:针对音频样本的多个识别文本和上述音频样本中涉及的样本实体;利用初始文本编码模型,对上述多个识别文本进行向量转换,以生成初始转换向量,其中,上述初始转换向量表征上述多个识别文本的文本特征信息;将上述样本实体输入至初始实体编码模型,以生成初始实体编码向量;将上述初始转换向量和上述初始实体编码向量输入至初始匹配概率生成模型,以生成初始匹配概率,其中,上述初始匹配概率表征上述初始转换向量与上述初始实体编码向量之间的匹配关系;响应于确定初始匹配结果与上述样本标签不同,对上述初始文本编码模型、上述初始实体编码模型和上述初始匹配概率生成模型进行参数训练,得到文本编码模型、实体编码模型和匹配概率生成模型,其中,上述初始匹配结果是基于上述初始匹配概率生成的。获取针对目标音频的多个识别文本;利用预先训练的文本编码模型,对上述多个识别文本进行向量转换,以生成转换向量,其中,上述转换向量表征上述多个识别文本的特征信息,上述文本编码模型是通过本公开一些实施例的模型训练方法生成的;利用预先训练的实体编码模型,对实体库中的各个实体进行实体编码,得到实体编码向量集,其中,上述实体编码模型是通过本公开一些实施例的模型训练方法生成的;利用预先训练的匹配概率生成模型,从上述实体库中筛选出对应匹配概率满足预设条件的实体,作为目标实体,得到至少一个目标实体,其中,目标实体对应的匹配概率是基于对应实体编码向量和上述转换向量生成的,上述匹配概率生成模型是通过本公开一些实施例的模型训练方法生成的。

[0131] 可以以一种或多种程序设计语言或其组合来编写用于执行本公开的一些实施例的操作的计算机程序代码,上述程序设计语言包括面向对象的程序设计语言—诸如Java、Smalltalk、C++,还包括常规的过程式程序设计语言—诸如“C”语言或类似的程序设计语言。程序代码可以完全地在用户计算机上执行、部分地在用户计算机上执行、作为一个独立的软件包执行、部分在用户计算机上部分在远程计算机上执行、或者完全在远程计算机或服务器上执行。在涉及远程计算机的情形中,远程计算机可以通过任意种类的网络——包括局域网(LAN)或广域网(WAN)——连接到用户计算机,或者,可以连接到外部计算机(例如利用因特网服务提供商来通过因特网连接)。

[0132] 附图中的流程图和框图,图示了按照本公开各种实施例的系统、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上,流程图或框图中的每个方框可以代表一个模块、程序段、或代码的一部分,该模块、程序段、或代码的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。也应当注意,在有些作为替换的实现中,方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如,两个接连地表示的方框实际上可以基本并行地执行,它们有时也可以按相反的顺序执行,这依所涉及的功能而定。也要注意,框图和/或流程图中的每个方框、以及框图和/或流程图中的方框的组合,可以用执行规定的功能或操作的专用的基于硬件的系统来实现,或者可以用专用硬件与计算机指令的组合来实现。

[0133] 描述于本公开的一些实施例中的单元可以通过软件的方式实现,也可以通过硬件的方式来实现。所描述的单元也可以设置在处理器中,例如,可以描述为:一种处理器包括:第一获取单元、第一向量转换单元、第一输入单元、第二输入单元和参数训练单元。其中,这些单元的名称在某种情况下并不构成对该单元本身的限定,例如,第一获取单元还可以被描述为“获取训练样本和上述训练样本对应的样本标签的单元”。

[0134] 本文中以上描述的功能可以至少部分地由一个或多个硬件逻辑部件来执行。例如,非限制性地,可以使用的示范类型的硬件逻辑部件包括:现场可编程门阵列(FPGA)、专用集成电路(ASIC)、专用标准产品(ASSP)、片上系统(SOC)、复杂可编程逻辑设备(CPLD)等等。

[0135] 以上描述仅为本公开的一些较佳实施例以及对所运用技术原理的说明。本领域技术人员应当理解,本公开的实施例中所涉及的发明范围,并不限于上述技术特征的特定组合而成的技术方案,同时也应涵盖在不脱离上述发明构思的情况下,由上述技术特征或其等同特征进行任意组合而形成的其它技术方案。例如上述特征与本公开的实施例中公开的(但不限于)具有类似功能的技术特征进行互相替换而形成的技术方案。

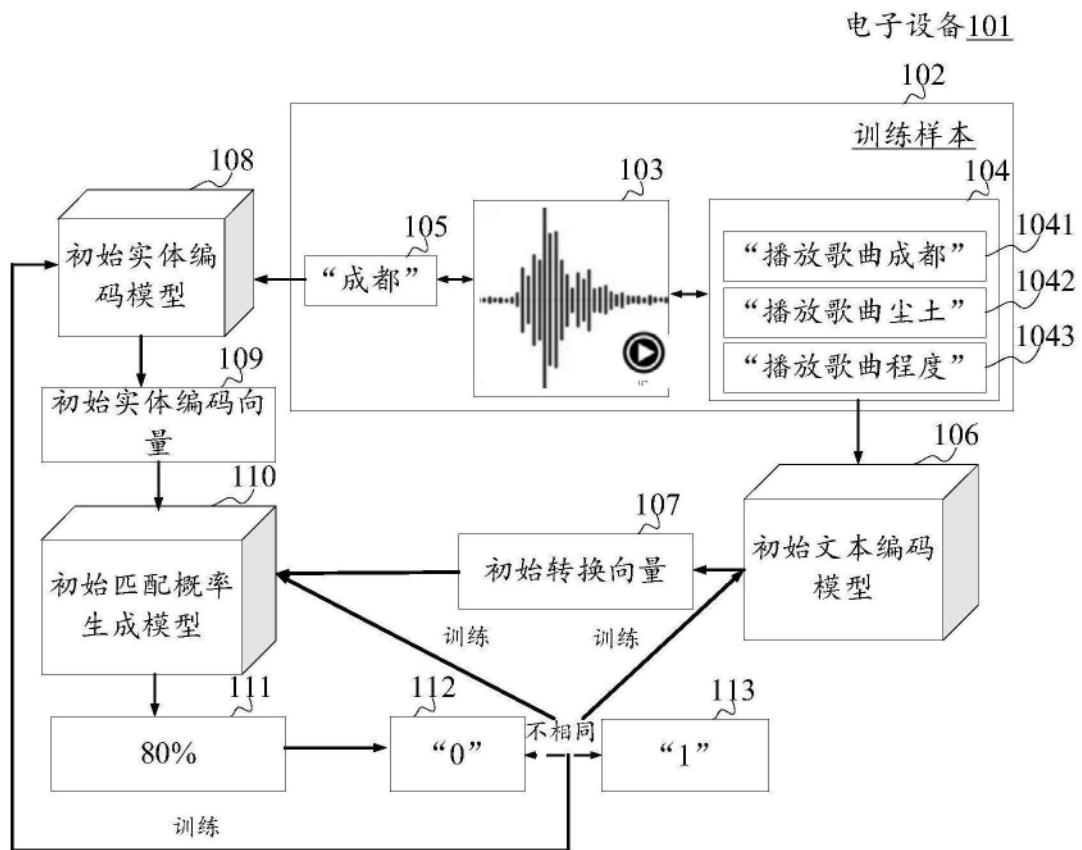


图1

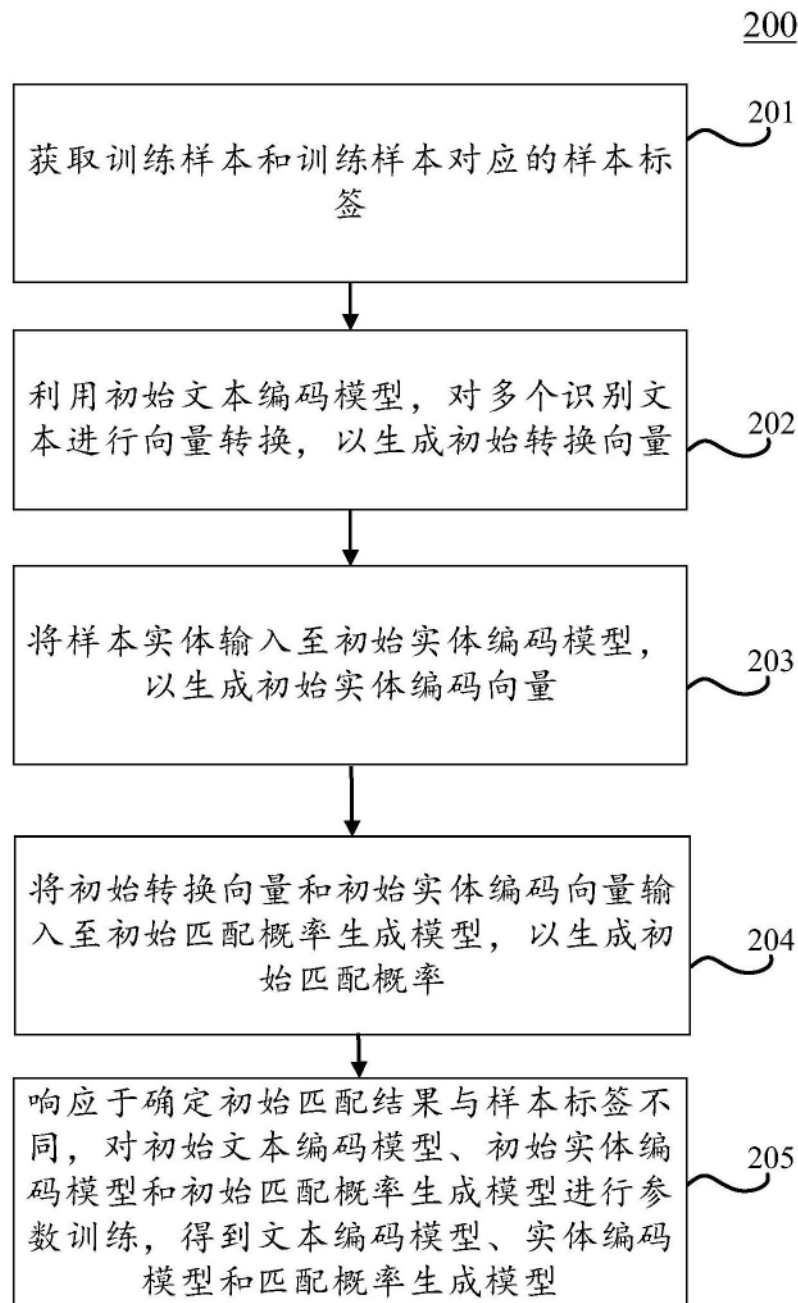


图2

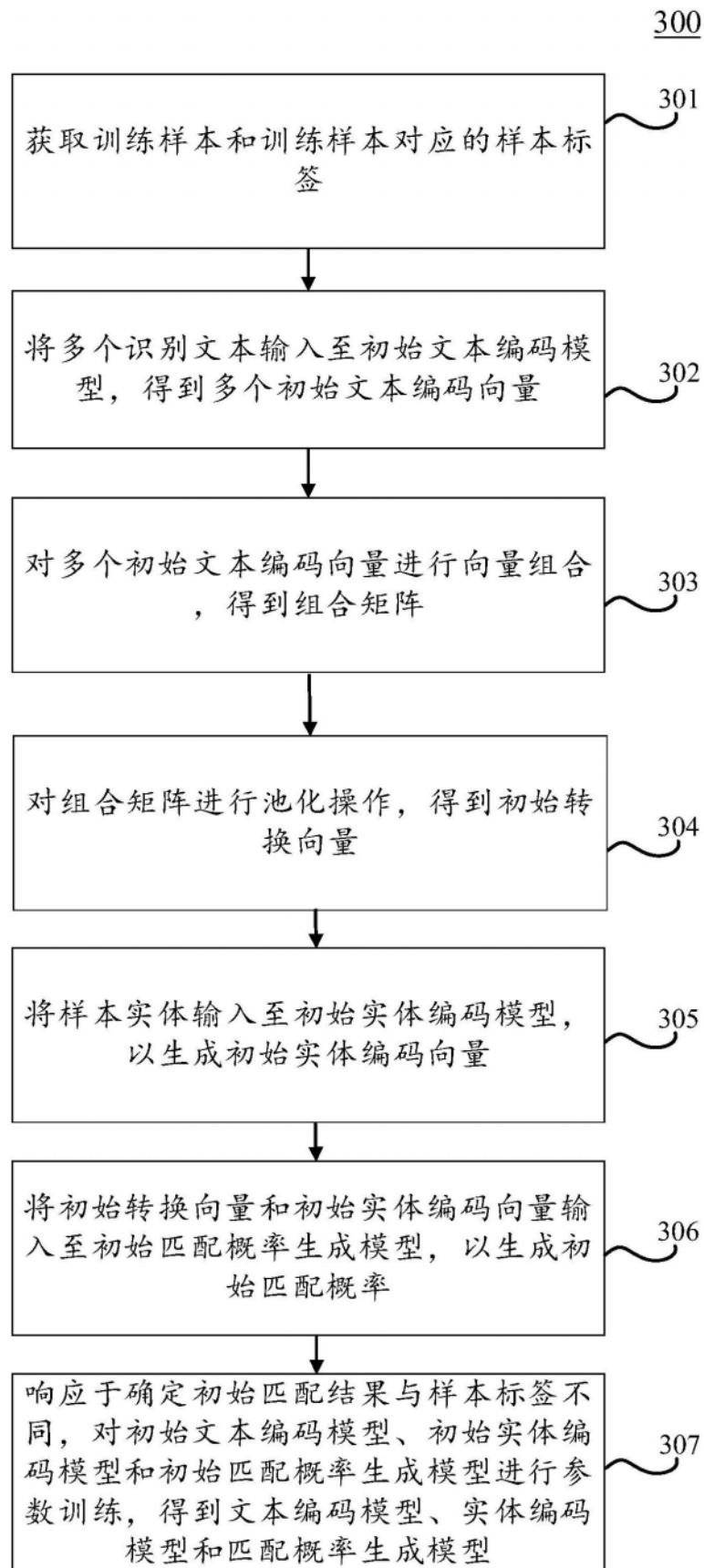


图3

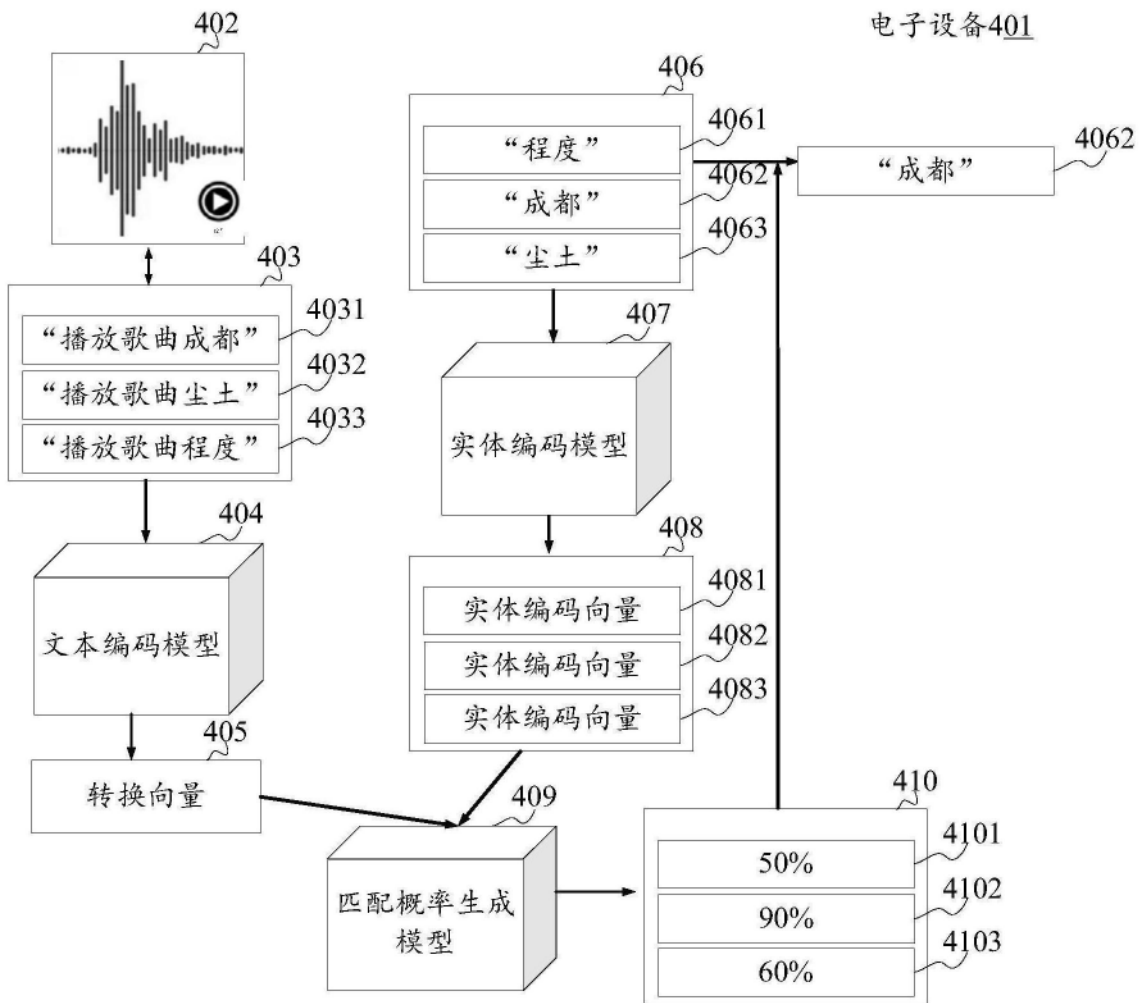


图4

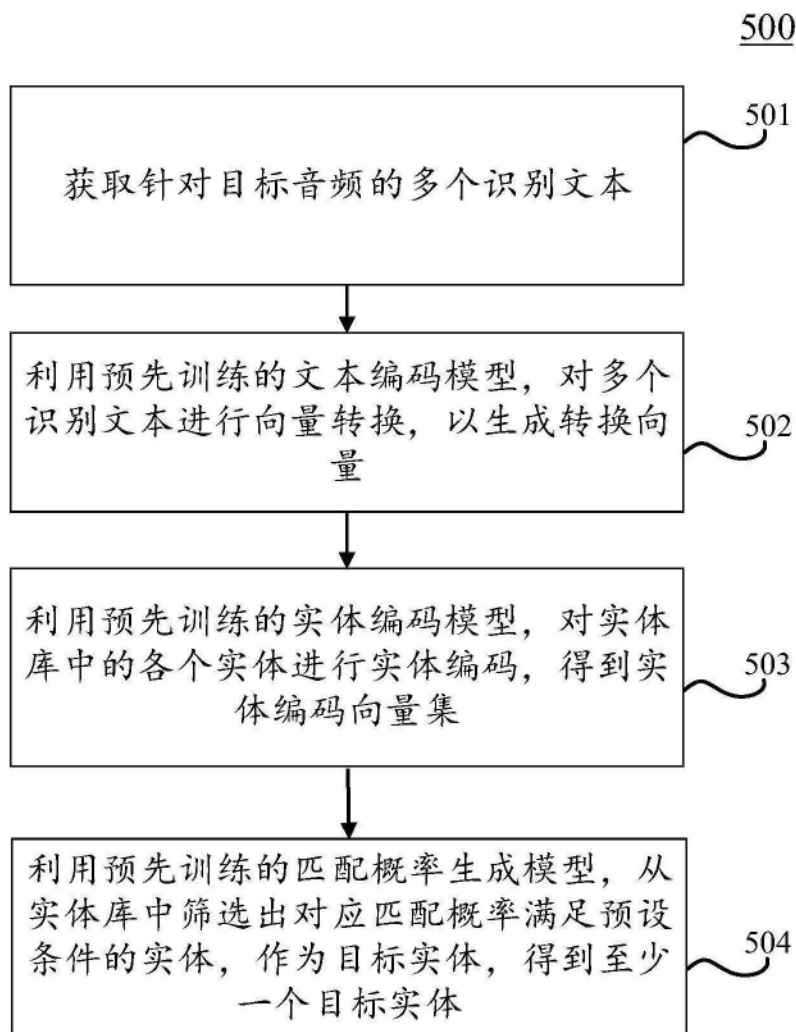


图5

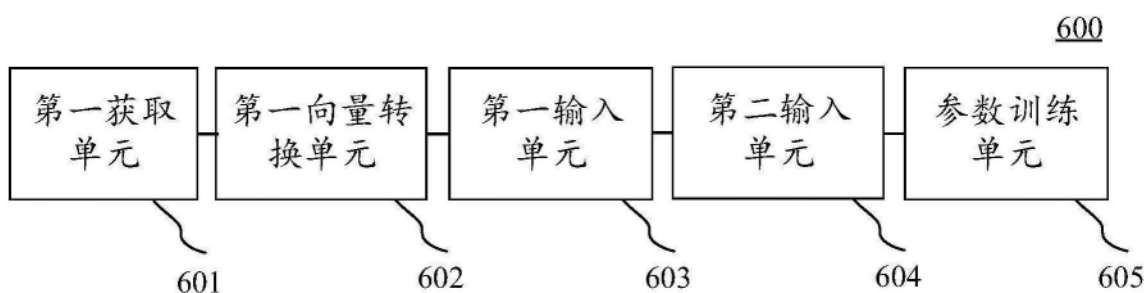


图6

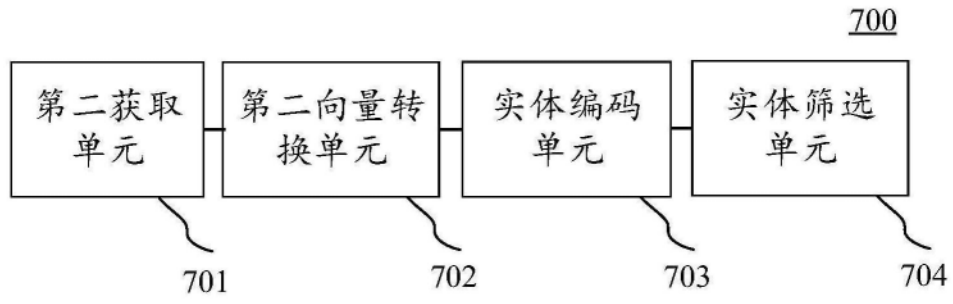


图7

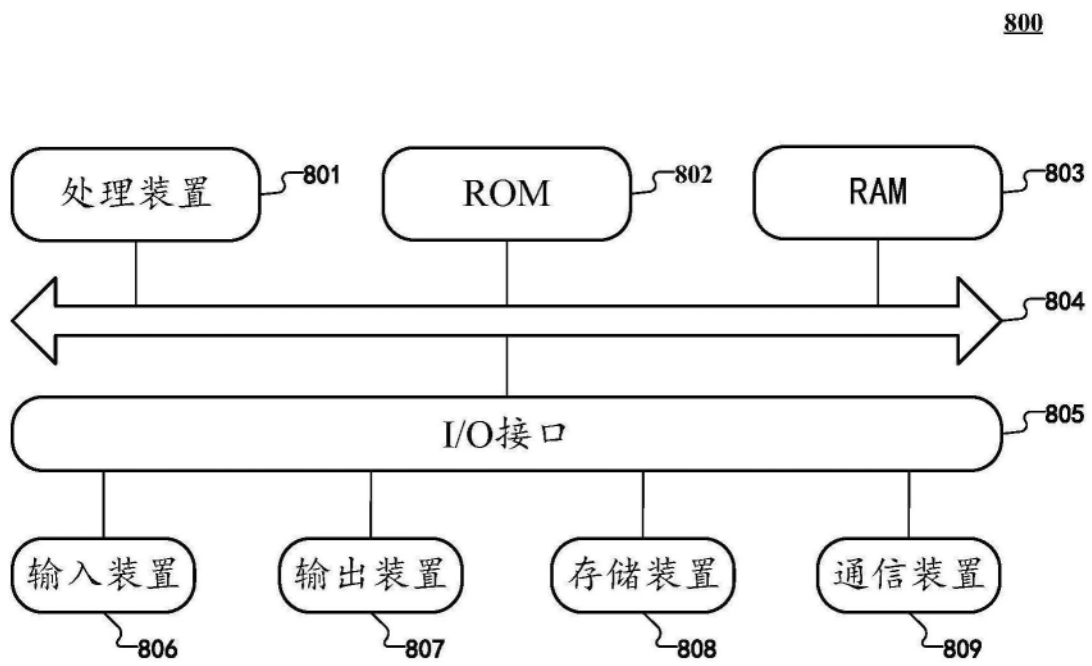


图8