

模型训练方法和装置及语音含义的理解方法和装置

技术领域

[01] 本申请涉及电子信息技术，尤其涉及模糊音识别模型的训练方法和装置、语音含义的理解方法和装置。

背景技术

[02] 目前，语音识别技术被广泛应用。在应用语音识别技术时，通常都会先对用户说出的语音进行识别，从语音转换为文本，然后再对文本的含义进行理解，从而得到语音的含义，并进行相关的处理。

[03] 然而，目前的语音识别技术尚未发展成熟，在从语音转换为文本时，经常出现识别错误，根据错误的文本，则无法准确地理解出语音的含义。

发明内容

[04] 本说明书一个或多个实施例描述了模糊音识别模型的训练方法和装置、语音含义的理解方法和装置，能够更加准确地理解出语音的含义。

[05] 根据第一方面，提供了一种模糊音识别模型的训练方法，包括：得到包括多个字符的具有语义的样本文本；对样本文本中的每一个字符，获取该字符的拼音；根据样本文本中的每一个字符的拼音，得到每一个字符对应的模糊音；利用样本文本、该样本文本中每一个字符对应的模糊音以及该样本文本的标签，训练所述模糊音识别模型。

[06] 所述根据样本文本中的每一个字符的拼音得到每一个字符对应的模糊音，包括：针对样本文本中每一个字符的拼音，判断该拼音中是否包括第一发音；该第一发音符合：一个第二发音的读音会被混淆为第一发音的读音；如果否，则将该拼音直接作为该字符对应的模糊音；如果是，则将该拼音中的第一发音替换为第二发音，替换后得到的拼音作为该字符对应的模糊音。

[07] 所述根据样本文本中的每一个字符的拼音得到每一个字符对应的模糊音，包括：将样本文本中每一个字符的拼音拆分成声母以及韵母；针对拆分出的声母，判断声母中是否包括第一发音；如果否，则将该声母直接作为该字符的声母对应的模糊音；如果是，则将该声母中的第一发音替换为该第一发音对应的第二发音，以得到该字符的声母对应

的模糊音；针对拆分出的韵母，判断韵母中是否包括第一发音；如果否，则将该韵母直接作为该字符的韵母对应的模糊音；如果是，则将该韵母中的第一发音替换为该第一发音对应的第二发音，以得到该字符的韵母对应的模糊音。

[08]针对拆分出的韵母，在判断韵母中是否包括第一发音之前，进一步包括：判断韵母是否包括韵头及韵尾，如果是，则删除该韵母中的韵头；所述判断韵母中是否包括第一发音包括：判断删除韵头后的韵母中是否包括第一发音。

[09]所述训练所述模糊音识别模型，包括：针对该样本文本中的每一个字符，生成对应该字符的三元组，该三元组包括：该字符、该字符的声母对应的模糊音及该字符的韵母对应的模糊音；按照每一个字符在样本文本中的顺序，依次将各字符对应的三元组以及所述标签输入待训练的所述模糊音识别模型。

[10]所述样本文本的标签包括：从情绪维度、所属领域维度、标的物维度、文本含义维度中的至少一个维度给出的标签。

[11]根据第二方面，提供了一种语音含义的理解方法，包括：得到第一文本；该第一文本是对语音进行语音识别后生成的；对第一文本中的每一个字符，获取该字符的拼音；根据每一个字符的拼音，得到每一个字符对应的模糊音；将第一文本、该第一文本中每一个字符对应的模糊音输入模糊音识别模型，得到模糊音识别模型输出的第二文本；对第二文本进行理解，得到所述语音的含义。

[12]所述根据第一文本中每一个字符的拼音得到每一个字符对应的模糊音，包括：针对第一文本中每一个字符的拼音，判断该拼音中是否包括第一发音；该第一发音符合：一个第二发音的读音会被混淆为第一发音的读音；如果否，则将该拼音直接作为该字符对应的模糊音；如果是，则将该拼音中的第一发音替换为第二发音，替换后得到的拼音作为该字符对应的模糊音。

[13]所述根据每一个字符的拼音得到每一个字符对应的模糊音，包括：将第一文本中每一个字符的拼音拆分成声母以及韵母；针对拆分出的声母，判断该声母中是否包括第一发音；如果否，则将该声母直接作为该字符的声母对应的模糊音；如果是，则将该声母中的第一发音替换为该第一发音对应的第二发音，以得到该字符的声母对应的模糊音；针对拆分出的韵母，判断韵母中是否包括第一发音；如果否，则将该韵母直接作为该字符的韵母对应的模糊音；如果是，则将该韵母中的第一发音替换为该第一发音对应的第二发音，以得到该字符的韵母对应的模糊音。

[14]针对拆分出的韵母，在判断韵母中是否包括第一发音之前，进一步包括：判断韵母是否包括韵头及韵尾，如果是，则删除该韵母中的韵头；判断韵母中是否包括第一发音包括：判断该删除韵头后的韵母中是否包括第一发音。

[15]所述将第一文本、该第一文本中每一个字符对应的模糊音输入模糊音识别模型，包括：针对该第一文本中的每一个字符，生成对应该字符的三元组，该三元组包括：该字符、该字符的声母对应的模糊音及该字符的韵母对应的模糊音；按照每一个字符在第一文本中的顺序，依次将各字符对应的三元组输入所述模糊音识别模型。

[16]根据第三方面，提供了一种模糊音识别模型的训练装置，包括：样本文本获取模块，配置为得到包括多个字符的具有语义的样本文本；拼音获取模块，配置为对样本文本中的每一个字符，获取该字符的拼音；模糊音生成模块，配置为根据样本文本中的每一个字符的拼音，得到每一个字符对应的模糊音；训练执行模块，配置为利用样本文本、该样本文本中每一个字符对应的模糊音以及该样本文本的标签，训练所述模糊音识别模型。

[17]根据第四方面，提供了一种语音含义的理解装置，包括：语音识别结果接收模块，配置为得到第一文本；该第一文本是对语音进行语音识别后生成的；字符拼音生成模块，配置为对样本文本中的每一个字符，获取该字符的拼音；字符模糊音生成模块，配置为根据每一个字符的拼音，得到每一个字符对应的模糊音；输入模块，配置为将第一文本、该第一文本中每一个字符对应的模糊音输入模糊音识别模型，得到模糊音识别模型输出的第二文本；语音含义理解模块，配置为对第二文本进行理解，得到所述语音的含义。

[18]根据第五方面，提供了一种计算设备，包括存储器和处理器，所述存储器中存储有可执行代码，所述处理器执行所述可执行代码时，实现本说明书任一实施例所述的方法。

[19]本说明书实施例能够训练出可以纠正语音识别的文字错误的模糊音识别模型，并基于该模型能够更加准确地理解语音的含义。

附图说明

[20]为了更清楚地说明本说明书实施例中的技术方案，下面将对实施例描述中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图是本说明书的一些实施例，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，还可以根据这些附图获得其他的附图。

[21]图1是本说明书一个实施例中模糊音识别模型的训练方法的流程图。

[22]图 2 是本说明书一个实施例中语音含义的理解方法的流程图。

[23]图 3 是本说明书一个实施例中模糊音识别模型的训练装置的结构示意图。

[24]图 4 是本说明书一个实施例中语音含义的理解装置的结构示意图。

具体实施方式

[25]如前所述，目前在从语音转换为文本时，经常出现识别错误，根据错误的文本，则无法准确地理解出语音所要表达的含义。举例说明：比如机器实现的智能客服系统向用户提问：请问您买的是实物还是虚拟物品？用户使用语音进行回答，用户原本希望回答的是实物，但是因为用户使用的是方言，因此导致语音识别出错，识别出的文本为：四五。这样，智能客服系统则无法根据识别出的错误的文本，来理解用户语音所要表达的含义，从而导致业务出错。

[26]下面结合附图，对本说明书提供的方案进行描述。

[27]针对语音识别的过程进行分析可知，语音识别出错的一个重要原因是：用户的语音发音不标准，不同地区的人可能使用不同的方言，将一个发音与另一个发音相混。比如，平舌音跟翘舌音相混，一些地方的人，会把拼音中的“z、c、s”的音发成“zh、ch、sh”，如把“自行车”读成“zhi xing che”行车。再如，一些地方的人，会把拼音中的“j、q、x”的音发成“z、c、s”，如把“进修”读作“zin siu”。又如，一些地方的人，会把拼音中的“f”与“h”的发音相混，如将“反对”读作“hu ǎ n dui”等。

[28]因此，如果能够将语音识别出的文本中的每一个字符的发音都利用发音相混的特点来进行修正，则可以有效解决因为用户发音不标准导致的语音识别出错的问题。

[29]为了能够利用发音相混的特点修正语音识别的错误，在本说明书一个实施例中，可以预先训练一个模糊音识别模型，在实际业务应用中，则可以利用该模糊音识别模型来修正语音识别出的文本，从而更为正确地理解语音的含义。

[30]本说明书实施例的方法可以适用于各种应用语音识别的场景中。比如，包括以下场景一至场景三。

场景一、智能客服系统

[31]在用户通过电话或者网络输入一段语音后，智能客服系统（比如支付宝平台的机器客服）会进行语音识别，识别出一段文本，应用本说明书实施例提供的模糊音识别模型

及语音含义的理解方法，则能够纠正语音识别出的文本的错误，得到正确的更加符合用户意图的文本，从而让机器更加正确地理解出用户语音的含义，比如用户购买的是实物还是虚拟物品，需要退货还是换货等。

场景二、人工智能系统

[32]在用户通过现场对话方式、电话或者网络方式发出一段语音后，人工智能系统（比如机器人）会进行语音识别，识别出一段文本，应用本说明书实施例提供的模糊音识别模型及语音含义的理解方法，人工智能系统（比如机器人）则能够纠正语音识别出的文本的错误，得到正确的更加符合用户意图的文本，从而让机器更加正确地理解出用户语音的含义，比如命令机器人改变行走路线等。

场景三、基于物联网的智能家居系统

[33]在用户通过现场对话方式、电话或者网络方式发出一段语音后，智能家居系统（比如智能电视）会进行语音识别，识别出一段文本，应用本说明书实施例提供的模糊音识别模型及语音含义的理解方法，智能家居系统（比如智能电视）则能够纠正语音识别出的文本的错误，得到正确的更加符合用户意图的文本，从而让机器更加正确地理解出用户语音的含义，比如命令智能电视开机录制某一时刻的电视节目等。

[34]下面将分两个方面说明本说明书实施例的实现方式，第一方面说明模糊音识别模型的训练方法，第二方面说明语音含义的理解方法。

[35]首先，在第一方面，说明模糊音识别模型的训练方法。

[36]图 1 是本说明书一个实施例中模糊音识别模型的训练方法的流程图。该方法的执行主体为模糊音识别模型的训练装置。可以理解，该方法也可以通过任何具有计算、处理能力的装置、设备、平台、设备集群来执行。参见图 1，该方法包括：步骤 101：得到包括多个字符的具有语义的样本文本；步骤 103：对样本文本中的每一个字符，获取该字符的拼音；步骤 105：根据样本文本中的每一个字符的拼音，得到每一个字符对应的模糊音；步骤 107：利用样本文本、该样本文本中每一个字符对应的模糊音以及该样本文本的标签，训练模糊音识别模型。

[37]可见，在上述图 1 所示的过程中，在对模糊音识别模型进行训练时，考虑到用户的发音不标准，对一个字符比如汉字的发音会与其他汉字的发音相混，因此，设计了模糊音的概念，这样，无论用户使用何种口音或者发音方式，都可以通过模糊音，将因为不同发音而导致相混的两个字符统一为同一个模糊音的发音，从而让模糊音识别模型学习

到被混读的字符的发音方式，并结合样本文本的上下文，得到正确的字符，从而使得此种模糊音识别模型能够用于后续纠正语音识别的文本的错误。

[38]下面结合具体的例子对图 1 所示的每一个步骤分别进行说明。

[39]首先，在步骤 101：得到包括多个字符的具有语义的样本文本。

[40]为了对模糊音识别模型进行训练，需要样本文本。样本文本可以是任意一种类型的具有语义的文本，比如一篇文章、一段用户投诉文字、一段产品说明文字等。为了能让模糊音识别模型学习用户口音导致的发音错误及对应的字符错误的各种情况，该样本文本中应该包括至少一个非标准发音（用户口音或者方言）形成的字符。比如，样本文本中包括“……至行车……，…发路”，因为用户口音，自行车中的自“zi”会发音成“zhi”，“zhi”对应出“至”这个非标准发音形成的字符，同时，因为用户口音，发怒中的怒“nu”会发音成“lu”，“lu”对应出“路”这个非标准发音形成的字符。

[41]在本说明书的实施例中，字符可以包括：汉字、英文字母、标点符号中的至少一个。

[42]在本说明书的实施例中，样本文本具有标签，该标签可以是情绪维度、所属领域维度、标的物维度、文本含义维度中的至少一个维度所给出的标签。比如，标签表征了样本文本所表达的情绪是愤怒；标签表征该样本文本属于用户投诉领域；标签表征样本文本的含义是用户购买了实体物品等，以便模糊音识别模型根据该标签对样本文本中的每一个字符及其模糊音是否正确进行学习。

[43]接下来对于步骤 103：对样本文本中的每一个字符，获取该字符的拼音。

[44]这里，可以从字典中获取每一个字符的标准的拼音。

[45]比如，样本文本中包括“……至行车……，…发路”，在本步骤 103 中，对于其中的字符“至”，获取其拼音为“zhi”；对于字符“行”，获取其拼音为“xing”；对于字符“车”，获取其拼音为“che”；对于字符“路”，获取其拼音为“lu”。

[46]接下来对于步骤 105：根据样本文本中的每一个字符的拼音，得到每一个字符对应的模糊音。

[47]在本说明书实施例中，设计了模糊音的概念，一个字符的拼音对应的模糊音符合：该拼音的发音不包括读音容易被混淆的第一发音时，拼音与该模糊音相同，而该拼音的发音包括第一发音时，该拼音的发音与模糊音的发音相混，这样，无论用户使用何种口音或者发音方式，都可以通过模糊音，将因为不同发音而导致相混的两个字符统一为同

一个模糊音的发音，从而让模糊音识别模型学习到被混读的字符的发音方式。

[48]如前所述，由于用户口音或者方言等原因，用户经常会将一个发音的读音混淆为另一个发音的读音，比如把发音“z、c、s”的读音分别混淆为发音“zh、ch、sh”的读音，把发音“ing”的读音混淆为“in”的读音，把发音“f”的读音混淆成“h”的读音等。因此，可以预先设置在读音上会相互混淆的第一发音与第二发音的对应关系，明确出哪些对发音会相互混淆。比如通过如下表 1 记录第一发音与第二发音的对应关系。

第一发音	第二发音
zh	z
ch	c
sh	s
l	n
h	f
in	ing
en	eng
an	ang

表 1

[49]在上述表 1 中，第一发音通常是带有口音或者使用方言的用户所发出的读音，第二发音是字符原本的读音。可以理解，上述表 1 只是示意性的，在实际的业务中，可以根据应用地点的不同，即用户口音特点的不同，设置不同的第一发音与第二发音的对应关系。在设置了诸如表 1 所示的对应关系之后，则可以利用该对应关系，对样本文本中的字符的拼音进行模糊音的替换。

[50]步骤 105 包括如下两种实现方式：方式一、以一个拼音为单位，进行拼音的模糊音的替换。方式二、以一个声母、韵母分别为一个单位，进行拼音的模糊音的替换。

[51]首先说明方式一：在本说明书一个实施例中，基于方式一的步骤 105 的具体实现过程包括：步骤 1051A：针对样本文本中每一个字符的拼音，判断该拼音中是否包括第一发音；该第一发音符合：一个第二发音的读音会被混淆为第一发音的读音，如果否，则执行步骤 1053A，如果是，则执行步骤 1055A。

[52]步骤 1053A：将该拼音直接作为该字符对应的模糊音。

[53]步骤 1055A: 将该拼音中的第一发音替换为第二发音, 替换后得到的拼音作为该字符对应的模糊音。

[54]举例说明上述步骤 1051A 至步骤 1055A 的过程。比如, 在上述步骤 103 中, 对于样本文本中包括的各个字符“...至行车..., ...发路”分别获取该各个字符对应的拼音为“zhi”、“xing”、“che”、“lu”。这样, 在步骤 1051A 至步骤 1055A 的过程中, 首先对于字符“至”的拼音“zhi”, 因为该拼音“zhi”中包括表 1 中的一个第一发音“zh”, 因此, 将该拼音中的第一发音“zh”替换为其对应的第二发音“z”, 替换后得到的拼音“zi”作为该字符“至”对应的模糊音; 接下来对于字符“行”的拼音“xing”, 因为该拼音“xing”中不包括表 1 中的任何第一发音, 因此直接将该拼音“xing”作为字符“行”对应的模糊音; 以此类推, 对于字符“路”的拼音“lu”, 因为该拼音“lu”中包括表 1 中的第一发音“l”, 因此, 将该拼音中的第一发音“l”替换为其对应的第二发音“n”, 替换后得到的拼音“nu”作为字符“路”对应的模糊音。

[55]这样, 经过步骤 1051A 至步骤 1055A 的处理, 得到的各字符对应的模糊音依次包括: “...zi xing che, ...fa nu”。

[56]下面说明方式二: 为了能够提高训练效率, 降低训练难度, 在本说明书一个实施例中, 可以采用方式二, 即, 对拼音进行拆分, 拆分成声母以及韵母, 然后分别判断声母中是否包括第一发音, 以及判断韵母中是否包括第一发音, 然后分别替换模糊音。此时, 步骤 105 具体包括: 步骤 1051B: 将样本文本中每一个字符的拼音拆分成声母以及韵母; 步骤 1053B: 针对拆分出的声母, 判断声母是否包括第一发音; 如果否, 则将该声母直接作为该字符的声母对应的模糊音; 如果是, 则将该声母中的第一发音替换为该第一发音对应的第二发音, 以得到该字符的声母对应的模糊音; 步骤 1055B: 针对拆分出的韵母, 判断韵母中是否包括第一发音; 如果否, 则将该韵母直接作为该字符的韵母对应的模糊音; 如果是, 则将该韵母中的第一发音替换为该第一发音对应的第二发音, 以得到该字符的韵母对应的模糊音。

[57]在本说明书一个实施例中, 为了进一步提高训练效率, 降低训练难度, 还可以进一步对韵母进行简化, 即删除韵母中的韵头部分。比如对于拼音“guang”, 其中, 韵母包括: “uang”, 韵母中的韵头“u”对于拼音“guang”的发音的贡献度比较小, 而韵母“ang”对于拼音“guang”的发音的贡献度比较大, 是韵母发音的关键部分, 因此, 为了提高效率, 可以在训练过程中不考虑该韵头对发音的影响。这样, 在上述步骤 1055B 中, 在判断韵母中是否包括第一发音之前, 进一步包括: 判断韵母是否包括韵头及韵尾,

如果是，则删除该韵母中的韵头；然后，在步骤 1055B 中，判断删除韵头后的韵母中是否包括第一发音；如果否，则将该删除韵头后的韵母直接作为该字符的韵母对应的模糊音；如果是，则将该删除韵头后的韵母中的第一发音替换为该第一发音对应的第二发音，以得到该字符的韵母对应的模糊音。

[58]接下来对于步骤 107：利用样本文本、该样本文本中每一个字符对应的模糊音以及该样本文本的标签，训练所述模糊音识别模型。

[59]仍以上述例子说明，如果步骤 105 采用上述方式一，则本步骤 107 中输入模糊音识别模型的信息包括：“...至 (zi) 行 (xing) 车 (che) ...，...发 (fa) 路 (nu)”以及样本文本的标签如“交通事故纠纷”。

[60]如果步骤 105 采用上述方式二，则本步骤 107 的具体实现过程包括：步骤 1071：针对该样本文本中的每一个字符，生成对应该字符的三元组，该三元组包括：该字符、该字符的声母对应的模糊音及该字符的韵母对应的模糊音；步骤 1073：按照每一个字符在样本文本中的顺序，依次将各字符对应的三元组以及标签输入待训练的模糊音识别模型。

[61]对比方式一及方式二，正常来说，声母有 23 个，韵母有 24 个，那么，采用方式一实现时，对于模糊音识别模型来说，为了学习模糊音，总共有 23×24 个未知数需要学习。而采用方式二实现时，对于模糊音识别模型来说，为了学习模糊音，总共有 $23 + 24$ 个未知数需要学习。可见，方式二，能够大大提高模糊音识别模型的训练效率，降低训练难度。

[62]无论采用方式一还是方式二，因为模糊音识别模型从训练过程的其他字符中已经学习到“自”的发音为“zi”，“怒”的发音为“nu”，再结合样本文本的上下文以及标签，该模糊音识别模型根据输入的信息则可以对样本文本进行纠正，比如纠正为...自行车...，...发怒”，从而得到正确的含义。

[63]可以理解，对于模糊音识别模型的训练会进行多轮，使用多个样本文本进行训练，每一轮的训练过程可以参见上述实施例，直至模糊音识别模型收敛。

[64]在训练完毕模糊音识别模型之后，则可以利用该模糊音识别模型进行语音含义的理解。

[65]下面说明第二方面，语音含义的理解方法。

[66]图 2 是本说明书一个实施例中语音含义的理解方法的流程图。该方法的执行主体为语音含义的理解装置。可以理解，该方法也可以通过任何具有计算、处理能力的装置、

设备、平台、设备集群来执行。参见图 2，该方法包括步骤 201 至步骤 209。

[67] 步骤 201：得到第一文本；该第一文本是对语音进行语音识别后生成的。

[68] 步骤 203：对第一文本中的每一个字符，获取该字符的拼音。

[69] 步骤 205：根据第一文本中的每一个字符的拼音，得到第一文本中的每一个字符对应的模糊音。

[70] 步骤 207：将第一文本、该第一文本中每一个字符对应的模糊音输入模糊音识别模型，得到模糊音识别模型输出的第二文本；步骤 209：对第二文本进行理解，得到语音的含义。

[71] 可见，在上述图 2 所示的过程中，无论用户使用何种口音或者发音方式，都可以通过模糊音，将因为不同发音而导致相混的两个字符统一为同一个模糊音的发音，这样让模糊音识别模型结合第一文本的上下文，就可以得到用户的语音真正需要的字符，即得到反映真实语义的第二文本，从而能够纠正语音识别的第一文本的错误，根据第二文本则可以让机器更加准确地理解用户语音的含义。

[72] 下面对图 2 中的每一个步骤进行说明。

[73] 首先对于步骤 201：得到第一文本；该第一文本是对语音进行语音识别后生成的。

[74] 这里，第一文本是在实际应用场景中，对用户语音进行语音识别后生成的，比如，用户向智能客服系统输入一段语音，由语音识别系统对该语音进行识别，从而得到了第一文本。

[75] 接下来对于步骤 203：对第一文本中的每一个字符，获取该字符的拼音。

[76] 这里，可以查找每一个字符在字典中的拼音。

[77] 接下来对于步骤 205：根据第一文本中的每一个字符的拼音，得到第一文本中的每一个字符对应的模糊音。

[78] 本步骤 205 也可以采用上述方式一及方式二两种方式来实现。

[79] 当采用方式一时，在本说明书一个实施例中，本步骤 205 的实现过程包括：针对第一文本中每一个字符的拼音，判断该拼音中是否包括第一发音；该第一发音符合：一个第二发音的读音会被混淆为第一发音的读音；如果否，则将该拼音直接作为该字符对应的模糊音；如果是，则将该拼音中的第一发音替换为第二发音，替换后得到的拼音作为该字符对应的模糊音。

[80]当采用方式二时，在本说明书一个实施例中，本步骤 205 的实现过程包括：将第一文本中每一个字符的拼音拆分成声母以及韵母；针对拆分出的声母，判断声母是否包括第一发音；如果否，则将该声母直接作为该字符的声母对应的模糊音；如果是，则将该声母中的第一发音替换为该第一发音对应的第二发音，以得到该字符的声母对应的模糊音；针对拆分出的韵母，判断韵母中是否包括第一发音；如果否，则将该韵母直接作为该字符的韵母对应的模糊音；如果是，则将该韵母中的第一发音替换为该第一发音对应的第二发音，以得到该字符的韵母对应的模糊音。

[81]在采用方式二时，针对拆分出的韵母，在判断韵母中是否包括第一发音之前，进一步包括：判断韵母是否包括韵头及韵尾，如果是，则删除该韵母中的韵头；判断韵母中是否包括第一发音包括：判断该删除韵头后的韵母中是否包括第一发音。

[82]本步骤 205 的具体实现过程也可以参见上述对步骤 105 的所有相关描述，其处理的思路是相同的。

[83]接下来对于步骤 207：将第一文本、该第一文本中每一个字符对应的模糊音输入模糊音识别模型，得到模糊音识别模型输出的第二文本；当步骤 205 采用方式二实现时，本步骤 207 的过程包括：针对该第一文本中的每一个字符，生成对应该字符的三元组，该三元组包括：该字符、该字符的声母对应的模糊音及该字符的韵母对应的模糊音；按照每一个字符在第一文本中的顺序，依次将各字符对应的三元组输入待训练的模糊音识别模型。

[84]对于步骤 207 的相关说明及理解可以参见上述对步骤 107 的相关说明，其处理思路是相同的。

[85]在本说明书的一个实施例中，提供了一种模糊音识别模型的训练装置，参见图 3，包括：样本文本获取模块 301，配置为得到包括多个字符的具有语义的样本文本；

[86]拼音获取模块 302，配置为对样本文本中的每一个字符，获取该字符的拼音；模糊音生成模块 303，配置为根据样本文本中的每一个字符的拼音，得到每一个字符对应的模糊音；训练执行模块 304，配置为利用样本文本、该样本文本中每一个字符对应的模糊音以及该样本文本的标签，训练所述模糊音识别模型。

[87]在本说明书装置的一个实施例中，模糊音生成模块 303 被配置为执行如下操作：针对样本文本中每一个字符的拼音，判断该拼音中是否包括第一发音；该第一发音符合：一个第二发音的读音会被混淆为第一发音的读音；如果否，则将该拼音直接作为该字符

对应的模糊音；如果是，则将该拼音中的第一发音替换为第二发音，替换后得到的拼音作为该字符对应的模糊音。

[88]在本说明书装置的另一个实施例中，模糊音生成模块 303 被配置为执行如下操作：将样本文本中每一个字符的拼音拆分成声母以及韵母；针对拆分出的声母，判断声母中是否包括第一发音；如果否，则将该声母直接作为该字符的声母对应的模糊音；如果是，则将该声母中的第一发音替换为该第一发音对应的第二发音，以得到该字符的声母对应的模糊音；针对拆分出的韵母，判断韵母中是否包括第一发音；如果否，则将该韵母直接作为该字符的韵母对应的模糊音；如果是，则将该韵母中的第一发音替换为该第一发音对应的第二发音，以得到该字符的韵母对应的模糊音。

[89]在本说明书装置的一个实施例中，模糊音生成模块 303 被配置为执行如下操作：针对拆分出的韵母，在判断韵母中是否包括第一发音之前，判断韵母是否包括韵头及韵尾，如果是，则删除该韵母中的韵头；判断删除韵头后的韵母中是否包括第一发音。

[90]在本说明书装置的一个实施例中，训练执行模块 304 配置为执行：针对该样本文本中的每一个字符，生成对应该字符的三元组，该三元组包括：该字符、该字符的声母对应的模糊音及该字符的韵母对应的模糊音；按照每一个字符在样本文本中的顺序，依次将各字符对应的三元组以及所述标签输入待训练的所述模糊音识别模型

[91]在本说明书装置的一个实施例中，样本文本的标签包括：从情绪维度、所属领域维度、标的物维度、文本含义维度中的至少一个维度给出的标签。

[92]在本说明书的一个实施例中，提供了一种语音含义的理解装置，参见图 4，包括：语音识别结果接收模块 401，配置为得到第一文本；该第一文本是对语音进行语音识别后生成的；字符拼音生成模块 402，配置为对第一文本中的每一个字符，获取该字符的拼音；字符模糊音生成模块 403，配置为根据每一个字符的拼音，得到每一个字符对应的模糊音；输入模块 404，配置为将第一文本、该第一文本中每一个字符对应的模糊音输入模糊音识别模型，得到模糊音识别模型输出的第二文本；语音含义理解模块 405，配置为对第二文本进行理解，得到所述语音的含义。

[93]在本说明书一个实施例的语音含义的理解装置中，字符模糊音生成模块 403 配置为执行如下操作：针对第一文本中每一个字符的拼音，判断该拼音中是否包括第一发音；该第一发音符合：一个第二发音的读音会被混淆为第一发音的读音；如果否，则将该拼音直接作为该字符对应的模糊音；如果是，则将该拼音中的第一发音替换为第二发音，

替换后得到的拼音作为该字符对应的模糊音。

[94]在本说明书另一个实施例的语音含义的理解装置中，字符模糊音生成模块 403 配置为执行如下操作：将第一文本中每一个字符的拼音拆分成声母以及韵母；针对拆分出的声母，判断该声母中是否包括第一发音；如果否，则将该声母直接作为该字符的声母对应的模糊音；如果是，则将该声母中的第一发音替换为该第一发音对应的第二发音，以得到该字符的声母对应的模糊音；针对拆分出的韵母，判断韵母中是否包括第一发音；如果否，则将该韵母直接作为该字符的韵母对应的模糊音；如果是，则将该韵母中的第一发音替换为该第一发音对应的第二发音，以得到该字符的韵母对应的模糊音。

[95]在本说明书实施例的语音含义的理解装置中，字符模糊音生成模块 403 配置为执行如下操作：针对拆分出的韵母，在判断韵母中是否包括第一发音之前，判断韵母是否包括韵头及韵尾，如果是，则删除该韵母中的韵头；判断该删除韵头后的韵母中是否包括第一发音。

[96]在本说明书实施例的语音含义的理解装置中，输入模块 404 配置为执行如下操作：针对该第一文本中的每一个字符，生成对应该字符的三元组，该三元组包括：该字符、该字符的声母对应的模糊音及该字符的韵母对应的模糊音；按照每一个字符在第一文本中的顺序，依次将各字符对应的三元组输入所述模糊音识别模型。

[97]本说明书一个实施例提供了一种计算机可读存储介质，其上存储有计算机程序，当所述计算机程序在计算机中执行时，令计算机执行说明书中任一个实施例中的方法。

[98]本说明书一个实施例提供了一种计算设备，包括存储器和处理器，所述存储器中存储有可执行代码，所述处理器执行所述可执行代码时，实现执行说明书中任一个实施例中的方法。

[99]可以理解的是，本说明书实施例示意的结构并不构成对本说明书实施例的装置的具体限定。在说明书的另一些实施例中，上述装置可以包括比图示更多或者更少的部件，或者组合某些部件，或者拆分某些部件，或者不同的部件布置。图示的部件可以以硬件、软件或者软件和硬件的组合来实现。

[100] 上述装置、系统内的各模块之间的信息交互、执行过程等内容，由于与本说明书方法实施例基于同一构思，具体内容可参见本说明书方法实施例中的叙述，此处不再赘述。

[101] 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，对于装置实施例而言，由于其基本相似于方法实施例，所以描述的比较简单，相关之处参见方法实施例的部分说明即可。

[102] 本领域技术人员应该可以意识到，在上述一个或多个示例中，本公开所描述的功能可以用硬件、软件、挂件或它们的任意组合来实现。当使用软件实现时，可以将这些功能存储在计算机可读介质中或者作为计算机可读介质上的一个或多个指令或代码进行传输。

[103] 以上所述的具体实施方式，对本公开的目的、技术方案和有益效果进行了进一步详细说明，所应理解的是，以上所述仅为本公开的具体实施方式而已，并不用于限定本公开的保护范围，凡在本公开的技术方案的基础之上，所做的任何修改、等同替换、改进等，均应包括在本公开的保护范围之内。

权利要求书

- 1、一种模糊音识别模型的训练方法，包括：
得到包括多个字符的具有语义的样本文本；
对样本文本中的每一个字符，获取该字符的拼音；
根据样本文本中的每一个字符的拼音，得到每一个字符对应的模糊音；
利用样本文本、该样本文本中每一个字符对应的模糊音以及该样本文本的标签，训练所述模糊音识别模型。
- 2、根据权利要求1所述的方法，其中，所述根据样本文本中的每一个字符的拼音得到每一个字符对应的模糊音，包括：
针对样本文本中每一个字符的拼音，判断该拼音中是否包括第一发音；该第一发音符合：一个第二发音的读音会被混淆为第一发音的读音；
如果否，则将该拼音直接作为该字符对应的模糊音；
如果是，则将该拼音中的第一发音替换为第二发音，替换后得到的拼音作为该字符对应的模糊音。
- 3、根据权利要求1所述的方法，其中，所述根据样本文本中的每一个字符的拼音得到每一个字符对应的模糊音，包括：
将样本文本中每一个字符的拼音拆分成声母以及韵母；
针对拆分出的声母，判断声母中是否包括第一发音；如果否，则将该声母直接作为该字符的声母对应的模糊音；如果是，则将该声母中的第一发音替换为该第一发音对应的第二发音，以得到该字符的声母对应的模糊音；
针对拆分出的韵母，判断韵母中是否包括第一发音；如果否，则将该韵母直接作为该字符的韵母对应的模糊音；如果是，则将该韵母中的第一发音替换为该第一发音对应的第二发音，以得到该字符的韵母对应的模糊音。
- 4、根据权利要求3所述的方法，其中，针对拆分出的韵母，在判断韵母中是否包括第一发音之前，进一步包括：判断韵母是否包括韵头及韵尾，如果是，则删除该韵母中的韵头；
所述判断韵母中是否包括第一发音包括：判断删除韵头后的韵母中是否包括第一发音。
- 5、根据权利要求3所述的方法，其中，所述训练所述模糊音识别模型，包括：
针对该样本文本中的每一个字符，生成对应该字符的三元组，该三元组包括：该字符、该字符的声母对应的模糊音及该字符的韵母对应的模糊音；

按照每一个字符在样本文本中的顺序,依次将各字符对应的三元组以及所述标签输入待训练的所述模糊音识别模型。

6、根据权利要求1所述的方法,所述样本文本的标签包括:从情绪维度、所属领域维度、标的物维度、文本含义维度中的至少一个维度给出的标签。

7、一种语音含义的理解方法,包括:

得到第一文本;该第一文本是对语音进行语音识别后生成的;

对第一文本中的每一个字符,获取该字符的拼音;

根据每一个字符的拼音,得到每一个字符对应的模糊音;

将第一文本、该第一文本中每一个字符对应的模糊音输入模糊音识别模型,得到模糊音识别模型输出的第二文本;

对第二文本进行理解,得到所述语音的含义。

8、根据权利要求7所述的方法,其中,所述根据第一文本中每一个字符的拼音得到每一个字符对应的模糊音,包括:

针对第一文本中每一个字符的拼音,判断该拼音中是否包括第一发音;该第一发音符合:一个第二发音的读音会被混淆为第一发音的读音;

如果否,则将该拼音直接作为该字符对应的模糊音;

如果是,则将该拼音中的第一发音替换为第二发音,替换后得到的拼音作为该字符对应的模糊音。

9、根据权利要求7所述的方法,其中,所述根据每一个字符的拼音得到每一个字符对应的模糊音,包括:

将第一文本中每一个字符的拼音拆分成声母以及韵母;

针对拆分出的声母,判断该声母中是否包括第一发音;如果否,则将该声母直接作为该字符的声母对应的模糊音;如果是,则将该声母中的第一发音替换为该第一发音对应的第二发音,以得到该字符的声母对应的模糊音;

针对拆分出的韵母,判断韵母中是否包括第一发音;如果否,则将该韵母直接作为该字符的韵母对应的模糊音;如果是,则将该韵母中的第一发音替换为该第一发音对应的第二发音,以得到该字符的韵母对应的模糊音。

10、根据权利要求9所述的方法,其中,针对拆分出的韵母,在判断韵母中是否包括第一发音之前,进一步包括:判断韵母是否包括韵头及韵尾,如果是,则删除该韵母中的韵头;

判断韵母中是否包括第一发音包括:判断该删除韵头后的韵母中是否包括第一发音。

11、根据权利要求9所述的方法，其中，所述将第一文本、该第一文本中每一个字符对应的模糊音输入模糊音识别模型，包括：

针对该第一文本中的每一个字符，生成对应该字符的三元组，该三元组包括：该字符、该字符的声母对应的模糊音及该字符的韵母对应的模糊音；

按照每一个字符在第一文本中的顺序，依次将各字符对应的三元组输入所述模糊音识别模型。

12、一种模糊音识别模型的训练装置，包括：

样本文本获取模块，配置为得到包括多个字符的具有语义的样本文本；

拼音获取模块，配置为对样本文本中的每一个字符，获取该字符的拼音；

模糊音生成模块，配置为根据样本文本中的每一个字符的拼音，得到每一个字符对应的模糊音；

训练执行模块，配置为利用样本文本、该样本文本中每一个字符对应的模糊音以及该样本文本的标签，训练所述模糊音识别模型。

13、一种语音含义的理解装置，包括：

语音识别结果接收模块，配置为得到第一文本；该第一文本是对语音进行语音识别后生成的；

字符拼音生成模块，配置为对样本文本中的每一个字符，获取该字符的拼音；

字符模糊音生成模块，配置为根据每一个字符的拼音，得到每一个字符对应的模糊音；

输入模块，配置为将第一文本、该第一文本中每一个字符对应的模糊音输入模糊音识别模型，得到模糊音识别模型输出的第二文本；

语音含义理解模块，配置为对第二文本进行理解，得到所述语音的含义。

14、一种计算机可读存储介质，其上存储有计算机程序，当所述计算机程序在计算机中执行时，令计算机执行权利要求1-11中任一项所述的方法。

15、一种计算设备，包括存储器和处理器，所述存储器中存储有可执行代码，所述处理器执行所述可执行代码时，实现权利要求1-11中任一项所述的方法。