



(12) 发明专利

(10) 授权公告号 CN 108304376 B

(45) 授权公告日 2021.09.10

(21) 申请号 201711362380.3

(22) 申请日 2017.12.15

(65) 同一申请的已公布的文献号  
申请公布号 CN 108304376 A

(43) 申请公布日 2018.07.20

(73) 专利权人 腾讯科技(深圳)有限公司  
地址 518000 广东省深圳市南山区高新区  
科技中一路腾讯大厦35层

(72) 发明人 宋彦 史树明

(74) 专利代理机构 北京康信知识产权代理有限  
责任公司 11240

代理人 赵囡囡

(51) Int. Cl.  
G06F 40/279 (2020.01)

(56) 对比文件  
CN 107273355 A, 2017.10.20  
CN 105930318 A, 2016.09.07  
Jinxing Yu等. Joint Embeddings of

Chinese Words, Characters, and Fine-grained Subcharacter Components. 《Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing》. 2017, 第1-4节.

Yaming Sun等. Radical-Enhanced Chinese Character Embedding. 《ICONIP 2014: Neural Information Processing》. 2014, 第1-6节.

Rongchao Yin等. Multi-Granularity Chinese Word Embedding. 《Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing》. 2016, 第1-4节.

Xinxiong Chen等. Joint Learning of Character and Word Embeddings. 《PROCEEDINGS OF THE TWENTY-FOURTH INTERNATIONAL JOINT CONFERENCE ON ARTIFICIAL INTELLIGENCE (IJCAI) 》. 2015, 第1-5节.

审查员 崔英秀

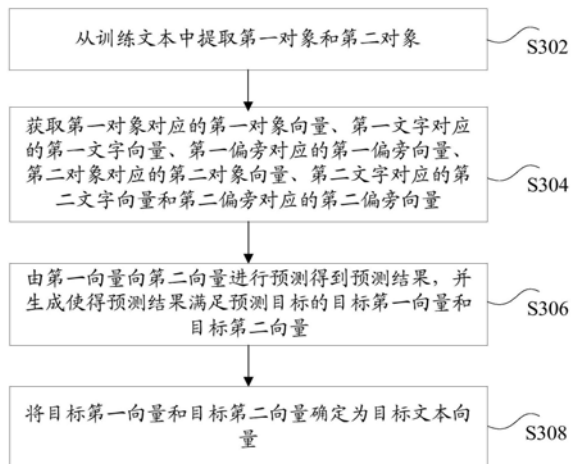
权利要求书4页 说明书18页 附图7页

(54) 发明名称

文本向量的确定方法、装置、存储介质及电子装置

(57) 摘要

本发明公开了一种文本向量的确定方法、装置、存储介质及电子装置。其中，该方法包括：从训练文本中提取第一对象和第二对象；获取第一对象对应的第一对象向量、第一文字对应的第一文字向量、第一偏旁对应的第一偏旁向量、第二对象对应的第二对象向量、第二文字对应的第二文字向量和第二偏旁对应的第二偏旁向量；由第一向量向第二向量进行预测得到预测结果，并生成使得预测结果满足预测目标的目标第一向量和目标第二向量；将目标第一向量和目标第二向量确定为目标文本向量。本发明解决了现有技术中生成文本向量时无语义信息以及对文本的语言资源利用率较低的技术问题。



CN 108304376 B

1. 一种文本向量的确定方法,其特征在于,包括:

从训练文本中提取第一对象和第二对象,其中,所述第一对象和所述第二对象为所述训练文本中不同的对象,所述对象按照不同粒度分为对象层、文字层及偏旁层三个层级,所述第一对象包括第一文字、所述第一文字包括第一偏旁,所述第二对象包括第二文字、所述第二文字包括第二偏旁;

获取所述第一对象对应的第一对象向量、所述第一文字对应的第一文字向量、所述第一偏旁对应的第一偏旁向量和所述第二对象对应的第二对象向量、所述第二文字对应的第二文字向量和所述第二偏旁对应的第二偏旁向量;

由第一向量向第二向量进行预测得到预测结果,并生成使得所述预测结果满足预测目标的目标第一向量和目标第二向量,其中,所述第一向量和所述第二向量为所述第一对象向量、所述第一文字向量、所述第一偏旁向量、所述第二对象向量、所述第二文字向量和所述第二偏旁向量中的任一向量,在基于各个层级各自对应的向量进行预测的过程中,包括:所述偏旁层的向量预测依赖于所述文字层的向量或所述对象层的向量,所述文字层的向量预测依赖于所述偏旁层的向量或所述对象层的向量,所述对象层的向量预测依赖于所述文字层的向量或所述偏旁层的向量;

将所述目标第一向量和所述目标第二向量确定为目标文本向量。

2. 根据权利要求1所述的方法,其特征在于,所述由第一向量向第二向量进行预测得到预测结果,并生成使得所述预测结果满足预测目标的目标第一向量和目标第二向量包括:

重复执行以下步骤,直至得到的预测准确率收敛,将使得所述预测准确率收敛的第一向量确定为所述目标第一向量并将使得所述预测准确率收敛的第二向量确定为所述目标第二向量,其中,所述预测结果满足所述预测目标包括所述预测准确率收敛:

将第一向量作为预测函数的输入参数输入所述预测函数,得到预测向量;

确定所述预测向量与第二向量之间的所述预测准确率;

在所述预测准确率未收敛的情况下,修正第一向量和第二向量,生成当前新的第一向量和当前新的第二向量。

3. 根据权利要求1或2所述的方法,其特征在于,由所述第一向量向所述第二向量进行预测包括以下之一:

由所述第一偏旁向量预测所述第一文字向量,由所述第一文字向量预测所述第一对象向量,由所述第一偏旁向量预测所述第二偏旁向量,由所述第一文字向量预测所述第二偏旁向量,由所述第一文字向量预测所述第二文字向量,由所述第一对象向量预测所述第二文字向量,以及由所述第一对象向量预测所述第二对象向量;

由所述第一偏旁向量预测所述第一文字向量,由所述第一文字向量预测所述第一对象向量,由所述第一偏旁向量预测所述第二偏旁向量,由所述第一偏旁向量预测所述第二文字向量,由所述第一文字向量预测所述第二文字向量,由所述第一文字向量预测所述第二对象向量,以及由所述第一对象向量预测所述第二对象向量;

由所述第一偏旁向量预测所述第一文字向量,由所述第一偏旁向量预测所述第二偏旁向量,由所述第一文字向量预测所述第二偏旁向量,以及由所述第一文字向量预测所述第二文字向量;

由所述第一偏旁向量预测所述第一对象向量,由所述第一偏旁向量预测所述第二偏旁

向量,由所述第一对象向量预测所述第二偏旁向量,以及由所述第一对象向量预测所述第二对象向量;

由所述第一文字向量预测所述第一对象向量,由所述第一文字向量预测所述第二文字向量,由所述第一对象向量预测所述第二文字向量,以及由所述第一对象向量预测所述第二对象向量。

4. 根据权利要求1所述的方法,其特征在于,所述第一偏旁为第一部首,所述第二偏旁为第二部首,其中,将所述目标第一向量和所述目标第二向量确定为所述目标文本向量包括:

从所述目标第一向量和所述目标第二向量中确定所述目标文本向量,其中,所述目标文本向量包括以下至少之一:第一目标对象向量、第一目标文字向量、第一目标部首向量、第二目标对象向量、第二目标文字向量、第二目标部首向量,其中,所述第一目标对象向量为所述第一对象对应的使得所述预测结果满足所述预测目标的向量、所述第一目标文字向量为所述第一文字对应的使得所述预测结果满足所述预测目标的向量、所述第一目标部首向量为所述第一部首对应的使得所述预测结果满足所述预测目标的向量,所述第二目标对象向量为所述第二对象对应的使得所述预测结果满足所述预测目标的向量、所述第二目标文字向量为所述第二文字对应的使得所述预测结果满足所述预测目标的向量、所述第二目标部首向量为所述第二部首对应的使得所述预测结果满足所述预测目标的向量;

将所述目标文本向量存储在数据库中。

5. 根据权利要求1所述的方法,其特征在于,从所述训练文本中提取所述第一对象和所述第二对象包括:

从所述训练文本中随机提取所述第一对象,并从所述训练文本中包括的除所述第一对象之外的其他对象中随机提取所述第二对象,其中,所述训练文本中的对象包括以下至少之一:短语,词语。

6. 根据权利要求5所述的方法,其特征在于,

从所述训练文本中随机提取所述第一对象包括:重复执行以下步骤,直至生成的第一随机数落入第一阈值范围,并将使所述第一随机数落入所述第一阈值范围的第一目标对象确定为所述第一对象:从所述训练文本中随机提取第一目标对象;为所述第一目标对象生成所述第一随机数;在所述第一随机数未落入所述第一阈值范围的情况下,从所述训练文本中随机提取一个目标对象,作为新的所述第一目标对象;

从所述训练文本中包括的除所述第一对象之外的其他对象中随机提取所述第二对象包括:在确定了所述第一对象之后,重复执行以下步骤,直至第二随机数落入第二阈值范围,并将使所述第二随机数落入所述第二阈值范围的第二目标对象确定为所述第二对象:从所述训练文本中包括的除所述第一对象之外的其他对象中随机提取第二目标对象;为所述第二目标对象生成所述第二随机数;在所述第二随机数未落入所述第二阈值范围的情况下,从所述训练文本中包括的除所述第一对象和所述第二目标对象之外的其他对象中随机提取一个目标对象,作为新的所述第二目标对象。

7. 根据权利要求1所述的方法,其特征在于,在将所述目标第一向量和所述目标第二向量确定为目标文本向量之后,所述方法还包括:

提取所述目标文本向量;

使用所述目标文本向量训练自然语言处理模型；  
使用所述自然语言处理模型执行自然语言处理操作。

8. 一种文本向量的确定装置,其特征在於,包括:

第一提取模块,用于从训练文本中提取第一对象和第二对象,其中,所述第一对象和所述第二对象为所述训练文本中不同的对象,所述对象按照不同粒度分为对象层、文字层及偏旁层三个层级,所述第一对象包括第一文字、所述第一文字包括第一偏旁,所述第二对象包括第二文字、所述第二文字包括第二偏旁;

获取模块,用于获取所述第一对象对应的第一对象向量、所述第一文字对应的第一文字向量、所述第一偏旁对应的第一偏旁向量和所述第二对象对应的第二对象向量、所述第二文字对应的第二文字向量和第二偏旁对应的第二偏旁向量;

预测模块,用于由第一向量向第二向量进行预测得到预测结果,并生成使得所述预测结果满足预测目标的目标第一向量和目标第二向量,其中,所述第一向量和所述第二向量为所述第一对象向量、所述第一文字向量、所述第一偏旁向量、所述第二对象向量、所述第二文字向量和所述第二偏旁向量中的任一向量,在基于各个层级各自对应的向量进行预测的过程中,包括:所述偏旁层的向量预测依赖于所述文字层的向量或所述对象层的向量,所述文字层的向量预测依赖于所述偏旁层的向量或所述对象层的向量,所述对象层的向量预测依赖于所述文字层的向量或所述偏旁层的向量;

确定模块,用于将所述目标第一向量和所述目标第二向量确定为目标文本向量。

9. 根据权利要求8所述的装置,其特征在於,所述预测模块包括:

执行单元,用于重复执行以下步骤,直至得到的预测准确率收敛,将使得所述预测准确率收敛的第一向量确定为所述目标第一向量并将使得所述预测准确率收敛的第二向量确定为所述目标第二向量,其中,所述预测结果满足所述预测目标包括所述预测准确率收敛:

输入单元,用于将第一向量作为预测函数的输入参数输入所述预测函数,得到预测向量;

第一确定单元,用于确定所述预测向量与第二向量之间的所述预测准确率;

修正单元,用于在所述预测准确率未收敛的情况下,修正第一向量和第二向量,生成当前新的第一向量和当前新的第二向量。

10. 根据权利要求8或9所述的装置,其特征在於,所述预测模块包括以下之一:

第一预测单元,用于由所述第一偏旁向量预测所述第一文字向量,由所述第一文字向量预测所述第一对象向量,由所述第一偏旁向量预测所述第二偏旁向量,由所述第一文字向量预测所述第二偏旁向量,由所述第一文字向量预测所述第二文字向量,由所述第一对象向量预测所述第二文字向量,以及由所述第一对象向量预测所述第二对象向量;

第二预测单元,用于由所述第一偏旁向量预测所述第一文字向量,由所述第一文字向量预测所述第一对象向量,由所述第一偏旁向量预测所述第二偏旁向量,由所述第一偏旁向量预测所述第二文字向量,由所述第一文字向量预测所述第二文字向量,由所述第一文字向量预测所述第二对象向量,以及由所述第一对象向量预测所述第二对象向量;

第三预测单元,用于由所述第一偏旁向量预测所述第一文字向量,由所述第一偏旁向量预测所述第二偏旁向量,由所述第一文字向量预测所述第二偏旁向量,以及由所述第一文字向量预测所述第二文字向量;

第四预测单元,用于由所述第一偏旁向量预测所述第一对象向量,由所述第一偏旁向量预测所述第二偏旁向量,由所述第一对象向量预测所述第二偏旁向量,以及由所述第一对象向量预测所述第二对象向量;

第五预测单元,用于由所述第一文字向量预测所述第一对象向量,由所述第一文字向量预测所述第二文字向量,由所述第一对象向量预测所述第二文字向量,以及由所述第一对象向量预测所述第二对象向量。

11. 根据权利要求8所述的装置,其特征在于,所述第一偏旁为第一部首,所述第二偏旁为第二部首,其中,所述确定模块包括:

第二确定单元,用于从所述目标第一向量和所述目标第二向量中确定所述目标文本向量,其中,所述目标文本向量包括以下至少之一:第一目标对象向量、第一目标文字向量、第一目标部首向量、第二目标对象向量、第二目标文字向量、第二目标部首向量,其中,所述第一目标对象向量为所述第一对象对应的使得所述预测结果满足所述预测目标的向量、所述第一目标文字向量为所述第一文字对应的使得所述预测结果满足所述预测目标的向量、所述第一目标部首向量为所述第一部首对应的使得所述预测结果满足所述预测目标的向量,所述第二目标对象向量为所述第二对象对应的使得所述预测结果满足所述预测目标的向量、所述第二目标文字向量为所述第二文字对应的使得所述预测结果满足所述预测目标的向量、所述第二目标部首向量为所述第二部首对应的使得所述预测结果满足所述预测目标的向量;

存储单元,用于将所述目标文本向量存储在数据库中。

12. 根据权利要求8所述的装置,其特征在于,所述第一提取模块包括:

提取单元,用于从所述训练文本中随机提取所述第一对象,并从所述训练文本中包括的除所述第一对象之外的其他对象中随机提取所述第二对象,其中,所述训练文本中的对象包括以下至少之一:短语,词语。

13. 根据权利要求8所述的装置,其特征在于,所述装置还包括:

第二提取模块,用于提取所述目标文本向量;

训练模块,用于使用所述目标文本向量训练自然语言处理模型;

执行模块,用于使用所述自然语言处理模型执行自然语言处理操作。

14. 一种存储介质,其特征在于,所述存储介质包括存储的程序,其中,所述程序运行时执行上述权利要求1至7任一项中所述的方法。

15. 一种电子装置,包括存储器、处理器及存储在所述存储器上并可在所述处理器上运行的计算机程序,其特征在于,所述处理器通过所述计算机程序执行上述权利要求1至7任一项中所述的方法。

## 文本向量的确定方法、装置、存储介质及电子装置

### 技术领域

[0001] 本发明涉及计算机领域,具体而言,涉及一种文本向量的确定方法、装置、存储介质及电子装置。

### 背景技术

[0002] 随着自然语言处理技术的飞速发展,自然语言处理技术中的基础研究也越来越得到人们的重视,这其中就包括对如何生成文本向量的研究。然而,在目前的技术中,文本向量的生成方式均没有充分的利用文本中的语义资源,导致生成的文本向量对语义的表达效率较低,从而使得使用该文本向量进行自然语言处理时无法充分利用被处理文本的语义信息。

[0003] 针对上述的问题,目前尚未提出有效的解决方案。

### 发明内容

[0004] 本发明实施例提供了一种文本向量的确定方法、装置、存储介质及电子装置,以至少解决现有技术中生成文本向量时无语义信息以及对文本的语言资源利用率较低的技术问题。

[0005] 根据本发明实施例的一个方面,提供了一种文本向量的确定方法,包括:从训练文本中提取第一对象和第二对象,其中,所述第一对象和所述第二对象为所述训练文本中不同的对象;获取所述第一对象对应的第一对象向量、第一文字对应的第一文字向量、第一偏旁对应的第一偏旁向量、所述第二对象对应的第二对象向量、第二文字对应的第二文字向量和第二偏旁对应的第二偏旁向量,其中,所述第一文字为所述第一对象包括的文字,所述第一偏旁为所述第一文字包括的偏旁,所述第二文字为所述第二对象包括的文字,所述第二偏旁为所述第二文字包括的偏旁;由第一向量向第二向量进行预测得到预测结果,并生成使得所述预测结果满足预测目标的目标第一向量和目标第二向量,其中,所述第一向量和所述第二向量为所述第一对象向量、所述第一文字向量、所述第一偏旁向量、所述第二对象向量、所述第二文字向量和所述第二偏旁向量中的任一向量;将所述目标第一向量和所述目标第二向量确定为目标文本向量。

[0006] 根据本发明实施例的另一方面,还提供了一种文本向量的确定装置,包括:第一提取模块,用于从训练文本中提取第一对象和第二对象,其中,所述第一对象和所述第二对象为所述训练文本中不同的对象;获取模块,用于获取所述第一对象对应的第一对象向量、第一文字对应的第一文字向量、第一偏旁对应的第一偏旁向量、所述第二对象对应的第二对象向量、第二文字对应的第二文字向量和第二偏旁对应的第二偏旁向量,其中,所述第一文字为所述第一对象包括的文字,所述第一偏旁为所述第一文字包括的偏旁,所述第二文字为所述第二对象包括的文字,所述第二偏旁为所述第二文字包括的偏旁;预测模块,用于由第一向量向第二向量进行预测得到预测结果,并生成使得所述预测结果满足预测目标的目标第一向量和目标第二向量,其中,所述第一向量和所述第二向量为所述第一对象向量、所

述第一文字向量、所述第一偏旁向量、所述第二对象向量、所述第二文字向量和所述第二偏旁向量中的任一向量；确定模块，用于将所述目标第一向量和所述目标第二向量确定为目标文本向量。

[0007] 根据本发明实施例的另一方面，还提供了一种存储介质，所述存储介质包括存储的程序，其中，所述程序运行时执行上述任一项中所述的方法。

[0008] 根据本发明实施例的另一方面，还提供了一种电子装置，包括存储器、处理器及存储在所述存储器上并可在所述处理器上运行的计算机程序，所述处理器通过所述计算机程序执行上述任一项中所述的方法。

[0009] 在本发明实施例中，从训练文本中提取第一对象和第二对象，其中，第一对象和第二对象为训练文本中不同的对象；获取第一对象对应的第一对象向量、第一文字对应的第一文字向量、第一偏旁对应的第一偏旁向量、第二对象对应的第二对象向量、第二文字对应的第二文字向量和第二偏旁对应的第二偏旁向量，其中，第一文字为第一对象包括的文字，第一偏旁为第一文字包括的偏旁，第二文字为第二对象包括的文字，第二偏旁为第二文字包括的偏旁；由第一向量向第二向量进行预测得到预测结果，并生成使得预测结果满足预测目标的目标第一向量和目标第二向量，其中，第一向量和第二向量为第一对象向量、第一文字向量、第一偏旁向量、第二对象向量、第二文字向量和第二偏旁向量中的任一向量；将目标第一向量和目标第二向量确定为目标文本向量。也就是说，在文本向量的确定过程中，将组成词汇的组件（例如：文字和偏旁）所携带的信息加入到词汇向量的学习过程中，并能够同步得到这些组件的向量，从而充分地利用了文本的语义资源，生成的文本向量也能更充分地表达出语义信息，使得生成文本向量时对文本的语义资源利用率得到了显著地提高，进而克服现有技术中生成文本向量时无语义信息以及对文本的语言资源利用率较低的问题。

## 附图说明

[0010] 此处所说明的附图用来提供对本发明的进一步理解，构成本申请的一部分，本发明的示意性实施例及其说明用于解释本发明，并不构成对本发明的不当限定。在附图中：

[0011] 图1是根据本发明实施例的一种可选的文本向量的确定方法的应用环境示意图；

[0012] 图2是根据本发明实施例的另一种可选的文本向量的确定方法的应用环境示意图；

[0013] 图3是根据本发明实施例的一种可选的文本向量的确定方法的示意图；

[0014] 图4是根据本发明可选的实施方式的一种可选的文本向量的确定方法的示意图一；

[0015] 图5是根据本发明可选的实施方式的一种可选的文本向量的确定方法的示意图二；

[0016] 图6是根据本发明可选的实施方式的一种可选的文本向量的确定方法的示意图三；

[0017] 图7是根据本发明可选的实施方式的一种可选的文本向量的确定方法的示意图四；

[0018] 图8是根据本发明可选的实施方式的一种可选的文本向量的确定方法的示意图

五;

[0019] 图9是根据本发明可选的实施方式的一种可选的文本向量的确定方法的示意图六;

[0020] 图10是根据本发明可选的实施方式的一种可选的文本向量的确定方法的示意图七;

[0021] 图11是根据本发明实施例的一种可选的文本向量的确定装置的示意图;

[0022] 图12是根据本发明实施例的一种可选的文本向量的确定方法的应用场景示意图;

[0023] 图13是根据本发明实施例的另一种可选的文本向量的确定方法的应用场景示意图;以及

[0024] 图14是根据本发明实施例的一种可选的电子装置的示意图。

### 具体实施方式

[0025] 为了使本技术领域的人员更好地理解本发明方案,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的实施例,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,都应当属于本发明保护的范围。

[0026] 需要说明的是,本发明的说明书和权利要求书及上述附图中的术语“第一”、“第二”等是用于区别类似的对象,而不必用于描述特定的顺序或先后次序。应该理解这样使用的数据在适当情况下可以互换,以便这里描述的本发明的实施例能够以除了在这里图示或描述的那些以外的顺序实施。此外,术语“包括”和“具有”以及他们的任何变形,意图在于覆盖不排他的包含,例如,包含了一系列步骤或单元的过程、方法、系统、产品或设备不必限于清楚地列出的那些步骤或单元,而是可包括没有清楚地列出的或对于这些过程、方法、产品或设备固有的其它步骤或单元。

[0027] 在本发明实施例中,提供了一种上述文本向量的确定方法的实施例。作为一种可选的实施方式,该文本向量的确定方法可以但不限于应用于如图1所示的应用环境中,服务器102用于从训练文本中提取第一对象和第二对象,其中,第一对象和第二对象为训练文本中不同的对象;获取第一对象对应的第一对象向量、第一文字对应的第一文字向量、第一偏旁对应的第一偏旁向量、第二对象对应的第二对象向量、第二文字对应的第二文字向量和第二偏旁对应的第二偏旁向量,其中,第一文字为第一对象包括的文字,第一偏旁为第一文字包括的偏旁,第二文字为第二对象包括的文字,第二偏旁为第二文字包括的偏旁;由第一向量向第二向量进行预测得到预测结果,并生成使得预测结果满足预测目标的目标第一向量和目标第二向量,其中,第一向量和第二向量为第一对象向量、第一文字向量、第一偏旁向量、第二对象向量、第二文字向量和第二偏旁向量中的任一向量;将目标第一向量和目标第二向量确定为目标文本向量。

[0028] 可选地,在本实施例中,该文本向量的确定方法还可以但不限于应用于如图2所示的应用环境中,服务器102与数据库202连接,服务器102从训练文本中提取第一对象和第二对象,其中,第一对象和第二对象为训练文本中不同的对象,从数据库202中获取第一对象对应的第一对象向量、第一文字对应的第一文字向量、第一偏旁对应的第一偏旁向量、第二



对象对应的第二对象向量、第二文字对应的第二文字向量和第二偏旁对应的第二偏旁向量,由第一向量向第二向量进行预测得到预测结果,并生成使得预测结果满足预测目标的目标第一向量和目标第二向量,将目标第一向量和目标第二向量确定为目标文本向量,并将目标文本向量发送给数据库202。数据库202用于存储具有对应关系的第一对象和第一对象向量,具有对应关系的第一文字和第一文字向量,具有对应关系的第一偏旁和第一偏旁向量,第二对象和第二对象向量,具有对应关系的第二文字和第二文字向量,具有对应关系的第二偏旁和第二偏旁向量。数据库202还用于使用服务器102发送的目标文本向量更新已存储的文本向量,即将存储的具有对应关系的第一对象和第一对象向量更新为具有对应关系的第一对象和第一目标对象向量,将存储的具有对应关系的第一文字和第一文字向量更新为具有对应关系的第一文字和第一目标文字向量,将存储的具有对应关系的第一偏旁和第一偏旁向量更新为具有对应关系的第一偏旁和第一目标偏旁向量,将存储的具有对应关系的第二对象和第二对象向量更新为具有对应关系的第二对象和第二目标对象向量,将存储的具有对应关系的第二文字和第二文字向量更新为具有对应关系的第二文字和第二目标文字向量,将存储的具有对应关系的第二偏旁和第二偏旁向量更新为具有对应关系的第二偏旁和第二目标偏旁向量,其中,第一目标对象向量,第一目标文字向量和第一目标偏旁向量为目标输入向量,第二目标对象向量,第二目标文字向量和第二目标偏旁向量为目标输出向量。

[0029] 根据本发明实施例,提供了一种文本向量的确定方法,如图3所示,该方法包括:

[0030] S302,从训练文本中提取第一对象和第二对象,其中,第一对象和第二对象为训练文本中不同的对象;

[0031] S304,获取第一对象对应的第一对象向量、第一文字对应的第一文字向量、第一偏旁对应的第一偏旁向量、第二对象对应的第二对象向量、第二文字对应的第二文字向量和第二偏旁对应的第二偏旁向量,其中,第一文字为第一对象包括的文字,第一偏旁为第一文字包括的偏旁,第二文字为第二对象包括的文字,第二偏旁为第二文字包括的偏旁;

[0032] S306,由第一向量向第二向量进行预测得到预测结果,并生成使得预测结果满足预测目标的目标第一向量和目标第二向量,其中,第一向量和第二向量为第一对象向量、第一文字向量、第一偏旁向量、第二对象向量、第二文字向量和第二偏旁向量中的任一向量;

[0033] S308,将目标第一向量和目标第二向量确定为目标文本向量。

[0034] 可选地,在本实施例中,上述文本向量的确定方法可以但不限于应用于自然语言处理任务中生成文本向量的场景中。其中,上述自然语言处理任务可以但不限于为任何以文本为基础的任务,例如:基于深度学习的中文分词,词性标注,命名实体识别,搜索,机器翻译,文本分类,资源推荐等等。具体的,可以但不限于应用于在上述基于深度学习的中文分词任务中生成文本向量的场景中,或还可以但不限于应用于在上述机器翻译任务中生成文本向量的场景中,以提高生成文本向量时对文本的语义资源的利用率。上述仅是一种示例,本实施例中对此不做任何限定。

[0035] 可选地,在本实施例中,文字的偏旁可以为偏旁中的部首。例如:对于文字“好”,该文字的偏旁为“女”和“子”,部首为“女”,在上述步骤S304中获取的偏旁可以为该字的全部偏旁“女”和“子”,也可以为该字的部首部分“女”。

[0036] 需要说明的是,由于部首一般情况下即是一个汉字语义的载体,例如:形声字的偏

旁可分为行部和声部,其部首通常是形部,表达该字的意义,而声部表达该字的发音,因此,在本实施例中,以获取的偏旁为部首为例而不是获取所有偏旁。获取所有偏旁生成文本向量的方式与此类似,在此不再赘述。

[0037] 可选地,在本实施例中,文本中的对象可以但不限于是词语或者短语等等。例如:词语可以是“青蛙”,短语则可以是“绿色的青蛙”,那么如果对象为词语“青蛙”,该词语对象对应的文字则为“青”和“蛙”,该词语对象对应的偏旁为“青”和“虫”。如果对象为短语“绿色的青蛙”,那么该词语对象对应的文字则为“绿”、“色”、“的”、“青”和“蛙”,该词语对象对应的偏旁为“纟”、“色”、“白”、“青”和“虫”。

[0038] 可选地,在本实施例中,第一向量可以但不限于为第一偏旁向量、第一文字向量和第一对象向量,第二向量可以但不限于为第一文字向量、第一对象向量、第二偏旁向量、第二文字向量和第二对象向量。

[0039] 可选地,在本实施例中,由第一向量向第二向量进行预测可以包括:由第一偏旁向量预测第一文字向量,由第一文字向量预测第一对象向量,由第一偏旁向量预测第二偏旁向量,由第一文字向量预测第二偏旁向量,由第一文字向量预测第二文字向量,由第一对象向量预测第二文字向量,以及由第一对象向量预测第二对象向量。

[0040] 可选地,在本实施例中,上述由第一对象向量预测第二文字向量可以替换为由第一文字向量预测第二对象向量,上述由第一文字向量预测第二偏旁向量可以替换为由第一偏旁向量预测第二文字向量。

[0041] 可见,通过上述步骤,在文本向量的确定过程中,将组成词汇的组件(例如:文字和偏旁)所携带的信息加入到词汇向量的学习过程中,并能够同步得到这些组件的向量,从而充分地利用了文本的语义资源,生成的文本向量也能更充分地表达出语义信息,使得生成文本向量时对文本的语义资源的利用率得到了显著地提高,进而克服现有技术中生成文本向量时无语义信息以及对文本的语言资源利用率较低的问题。

[0042] 作为一种可选的方案,由第一向量向第二向量进行预测得到预测结果,并生成使得预测结果满足预测目标的目标第一向量和目标第二向量包括:

[0043] S1,重复执行以下步骤,直至得到的预测准确率收敛,将使得预测准确率收敛的第一向量确定为目标第一向量并将使得预测准确率收敛的第二向量确定为目标第二向量,其中,预测结果满足预测目标包括预测准确率收敛;

[0044] S2,将第一向量作为预测函数的输入参数输入预测函数,得到预测向量;

[0045] S3,确定预测向量与第二向量之间的预测准确率;

[0046] S4,在预测准确率未收敛的情况下,修正第一向量和第二向量,生成当前新的第一向量和当前新的第二向量。

[0047] 可选地,在本实施例中,预测函数可以但不限于是预先设置的。确定预测准确率的方式可以但不限于为确定预测向量与当前的输出向量之间的相似度。例如:通过计算预测向量与当前的输出向量之间的欧式距离确定相似度等等。

[0048] 可选地,在本实施例中,预测准确率收敛可以但不限于指预测准确率达到稳定值。

[0049] 在一个可选的实施方式中,如图4所示,由第一向量向第二向量进行预测得到预测结果,并生成使得预测结果满足预测目标的目标第一向量和目标第二向量包括如下步骤:

[0050] 步骤S402,将第一向量作为预测函数的输入参数输入预测函数,得到预测向量;

[0051] 步骤S404,确定预测向量与第二向量之间的预测准确率;

[0052] 步骤S406,判断预测准确率是否收敛,在预测准确率未收敛的情况下,执行步骤S408,在预测准确率收敛的情况下,执行步骤S410;

[0053] 步骤S408,修正第一向量和第二向量,生成当前新的第一向量和当前新的第二向量,并返回步骤S402;

[0054] 步骤S410,将使得预测准确率收敛的第一向量确定为目标第一向量并将使得预测准确率收敛的第二向量确定为目标第二向量。

[0055] 作为一种可选的方案,由第一向量向第二向量进行预测包括以下方式之一:

[0056] 方式1,由第一偏旁向量预测第一文字向量,由第一文字向量预测第一对象向量,由第一偏旁向量预测第二偏旁向量,由第一文字向量预测第二偏旁向量,由第一文字向量预测第二文字向量,由第一对象向量预测第二文字向量,以及由第一对象向量预测第二对象向量。

[0057] 在一个可选的实施方式中,如图5所示,预测模型包括三层,其中,第一层为偏旁层,E(1)为第一偏旁向量,D(1)为第二偏旁向量,第二层为文字层,E(2)为第一文字向量,D(2)为第二文字向量,第三层为对象层,E(3)为第一对象向量,D(3)为第二对象向量,如图5中箭头所示为方式1中的预测过程,该预测过程包括:由E(1)预测E(2),由E(2)预测E(3),由E(1)预测D(1),由E(2)预测D(1),由E(2)预测D(2),由E(3)预测D(2),以及由E(3)预测D(3)。

[0058] 方式2,由第一偏旁向量预测第一文字向量,由第一文字向量预测第一对象向量,由第一偏旁向量预测第二偏旁向量,由第一偏旁向量预测第二文字向量,由第一文字向量预测第二文字向量,由第一文字向量预测第二对象向量,以及由第一对象向量预测第二对象向量。

[0059] 在一个可选的实施方式中,如图6所示,上述方式1中的由第一文字向量预测第二偏旁向量可以替换为由第一偏旁向量预测第二文字向量,即将由E(2)预测D(1)替换为由E(1)预测D(2),由第一对象向量预测第二文字向量可以替换为由第一文字向量预测第二对象向量,即将由E(3)预测D(2)替换为E(2)预测D(3)。

[0060] 方式3,由第一偏旁向量预测第一文字向量,由第一偏旁向量预测第二偏旁向量,由第一文字向量预测第二偏旁向量,以及由第一文字向量预测第二文字向量。

[0061] 在一个可选的实施方式中,如图7所示,预测模型包括两层,其中,第一层为偏旁层,E(1)为第一偏旁向量,D(1)为第二偏旁向量,第二层为文字层,E(2)为第一文字向量,D(2)为第二文字向量,如图7中箭头所示为方式3中的预测过程,该预测过程包括:由E(1)预测E(2),由E(1)预测D(1),由E(2)预测D(1),由E(2)预测D(2)。

[0062] 方式4,由第一偏旁向量预测第一对象向量,由第一偏旁向量预测第二偏旁向量,由第一对象向量预测第二偏旁向量,以及由第一对象向量预测第二对象向量。

[0063] 在一个可选的实施方式中,如图8所示,预测模型包括两层,其中,第一层为偏旁层,E(1)为第一偏旁向量,D(1)为第二偏旁向量,第二层为对象层,E(3)为第一对象向量,D(3)为第二对象向量,如图8中箭头所示为方式4中的预测过程,该预测过程包括:由E(1)预测E(3),由E(1)预测D(1),由E(3)预测D(1),以及由E(3)预测D(3)。

[0064] 方式5,由第一文字向量预测第一对象向量,由第一文字向量预测第二文字向量,由第一对象向量预测第二文字向量,以及由第一对象向量预测第二对象向量。

[0065] 在一个可选的实施方式中,如图9所示,预测模型包括两层,其中,第一层为文字层,E(2)为第一文字向量,D(2)为第二文字向量,第二层为对象层,E(3)为第一对象向量,D(3)为第二对象向量,如图9中箭头所示为方式5中的预测过程,该预测过程包括:由E(2)预测E(3),由E(2)预测D(2),由E(3)预测D(2),以及由E(3)预测D(3)。

[0066] 作为一种可选的方案,第一偏旁为第一部首,第二偏旁为第二部首,其中,将目标第一向量和目标第二向量确定为目标文本向量包括:

[0067] S1,从目标第一向量和目标第二向量中确定目标文本向量,其中,目标文本向量包括以下至少之一:第一目标对象向量、第一目标文字向量、第一目标部首向量、第二目标对象向量、第二目标文字向量、第二目标部首向量,其中,第一目标对象向量为第一对象对应的使得预测结果满足预测目标的向量、第一目标文字向量为第一文字对应的使得预测结果满足预测目标的向量、第一目标部首向量为第一部首对应的使得预测结果满足预测目标的向量,第二目标对象向量为第二对象对应的使得预测结果满足预测目标的向量、第二目标文字向量为第二文字对应的使得预测结果满足预测目标的向量、第二目标部首向量为第二部首对应的使得预测结果满足预测目标的向量;

[0068] S2,将目标文本向量存储在数据库中。

[0069] 可选地,在本实施例中,上述数据库可以存储了具有对应关系的第一对象和第一对象向量、具有对应关系的第一文字和第一文字向量、具有对应关系的第一偏旁和第一偏旁向量、具有对应关系的第二对象和第二对象向量、具有对应关系的第二文字和第二文字向量、具有对应关系的第二偏旁和第二偏旁向量。那么,将目标文本向量存储在数据库中的过程可以但不限于是对数据库中已存储的具有对应关系的文本与文本向量进行更新。

[0070] 例如:数据库中原存储了具有对应关系的第一对象和第一对象向量、具有对应关系的第一文字和第一文字向量、具有对应关系的第一偏旁和第一偏旁向量、具有对应关系的第二对象和第二对象向量、具有对应关系的第二文字和第二文字向量以及具有对应关系的第二偏旁和第二偏旁向量,那么,在得到第一目标对象向量、第一目标文字向量、第一目标偏旁向量、第二目标对象向量、第二目标文字向量、第二目标偏旁向量后,可以将数据库中存储的上述内容更新为具有对应关系的第一对象和第一目标对象向量、具有对应关系的第一文字和第一目标文字向量、具有对应关系的第一偏旁和第一目标偏旁向量、具有对应关系的第二对象和第二目标对象向量、具有对应关系的第二文字和第二目标文字向量以及具有对应关系的第二偏旁和第二目标偏旁向量。

[0071] 作为一种可选的方案,从训练文本中提取第一对象和第二对象包括:

[0072] S1,从训练文本中随机提取第一对象,并从训练文本中包括的除第一对象之外的其他对象中随机提取第二对象,其中,训练文本中的对象包括以下至少之一:短语,词语。

[0073] 可选地,在本实施例中,提取的第一对象可以为一个短语,或者为一个词语,或者为一个短语和组成该短语的词语。如果将第一对象向量、第一文字向量和第一偏旁向量作为预测模型的输入层,对于不同形式的第一对象,输入层的形式也不相同,例如:第一对象为一个短语(例如:第一短语)的情况下,输入层为第一短语向量、第一文字向量和第一偏旁向量,第一对象为一个词语(例如:第一词语)的情况下,输入层为第一词语向量、第一文字向量和第一偏旁向量,第一对象为一个短语和组成该短语的词语(例如:第一短语和第一词语,其中,第一词语是组成第一短语的词语)的情况下,输入层为第一短语向量、第一词语向

量、第一文字向量和第一偏旁向量。需要说明的是,对于第二对象的提取形式与第一对象类似,在此不在赘述。

[0074] 可选地,在本实施例中,可以但不限于通过以下方式提取第一对象:重复执行以下步骤,直至生成的第一随机数落入第一阈值范围,并将使第一随机数落入第一阈值范围的第一目标对象确定为第一对象:从训练文本中随机提取第一目标对象;为第一目标对象生成第一随机数;在第一随机数未落入第一阈值范围的情况下,从训练文本中随机提取一个目标对象,作为新的第一目标对象。

[0075] 可选地,在本实施例中,可以但不限于通过以下方式提取第二对象:在确定了第一对象之后,重复执行以下步骤,直至第二随机数落入第二阈值范围,并将使第二随机数落入第二阈值范围的第二目标对象确定为第二对象:从训练文本中包括的除第一对象之外的其他对象中随机提取第二目标对象;为第二目标对象生成第二随机数;在第二随机数未落入第二阈值范围的情况下,从训练文本中包括的除第一对象和第二目标对象之外的其他对象中随机提取一个目标对象,作为新的第二目标对象。

[0076] 在一个可选的实施方式中,如图10所示,从训练样本中提取第一对象和第二对象的过程可以但不限于包括如下步骤:

[0077] 步骤S1002,从训练文本中随机提取第一目标对象;

[0078] 步骤S1004,为第一目标对象生成第一随机数;

[0079] 步骤S1006,判断第一随机数是否落入第一阈值范围,如果是,则执行步骤S1010,如果不是,则执行步骤S1008;

[0080] 步骤S1008,从训练文本中随机提取一个目标对象,作为新的第一目标对象,返回步骤S1002;

[0081] 步骤S1010,将使第一随机数落入第一阈值范围的第一目标对象确定为第一对象;

[0082] 步骤S1012,从训练文本中包括的除第一对象之外的其他对象中随机提取第二目标对象;

[0083] 步骤S1014,为第二目标对象生成第二随机数;

[0084] 步骤S1016,判断第二随机数是否落入第二阈值范围,如果是,则执行步骤S1020,如果不是,则执行步骤S1018;

[0085] 步骤S1018,从训练文本中包括的除第一对象和第二目标对象之外的其他对象中随机提取一个目标对象,作为新的第二目标对象,返回步骤S1012;

[0086] 步骤S1020,将使第二随机数落入第二阈值范围的第二目标对象确定为第二对象。

[0087] 作为一种可选的方案,在将目标输入向量和目标输出向量确定为目标文本向量之后,还包括:

[0088] S1,提取目标文本向量;

[0089] S2,使用目标文本向量训练自然语言处理模型;

[0090] S3,使用自然语言处理模型执行自然语言处理操作。

[0091] 可选地,在本实施例中,确定的目标文本向量可以应用于自然语言处理任务中,可以使用确定的目标文本向量训练自然语言处理模型,并使用训练好的模型执行自然语言处理操作。例如:可以使用确定的目标文本向量训练文本分类模型,并利用训练后的文本分类模型执行文本资源分类的操作。

[0092] 作为一种可选的方案,获取第一对象对应的第一对象向量、第一文字对应的第一文字向量、第一偏旁对应的第一偏旁向量、第二对象对应的第二对象向量、第二文字对应的第二文字向量和第二偏旁对应的第二偏旁向量包括:

[0093] S1,获取第一对象对应的第一对象向量和第二对象对应的第二对象向量;

[0094] S2,从第一对象中提取组成第一对象的第一文字,从第二对象中提取组成第二对象的第二文字,并获取第一文字对应的第一文字向量和第二文字对应的第二文字向量;

[0095] S3,从第一文字中提取第一偏旁,从第二文字中提取第二偏旁,并获取第一偏旁对应的第一偏旁向量和第二偏旁对应的第二偏旁向量。

[0096] 可选地,在本实施例中,如果获取的第一对象或者第二对象是首次获取的对象,那么可以为首次获取的对象随机生成一个随机向量,如果该对象不是首次获取的对象,那么可以在数据库中获得数据库记录的具有对应关系的该对象和该对象的向量。对于文字向量和偏旁向量的获取,可以首先查找数据库中是否已经存储了该文字向量或者偏旁向量,如果已经存储,则可以直接提取使用,如果未存储,则可以为其生成一个随机向量。

[0097] 在一个可选的实施方式中,各个向量的获取可以但不限于采用以下的方式:

[0098] 第一对象向量的获取方式包括:方式1,在第一对象为首次提取的对象的情况下,生成第一随机向量作为第一对象向量;或者,方式2,在第一对象为非首次提取的对象的情况下,获取数据库中存储的第一对象向量,其中,数据库中记录了具有对应关系的第一对象和第一对象向量。

[0099] 第二对象向量的获取方式包括:方式1,在第二对象为首次提取的对象的情况下,生成第二随机向量作为第二对象向量;或者,方式2,在第二对象为非首次提取的对象的情况下,获取数据库中存储的第二对象向量,其中,数据库中记录了具有对应关系的第二对象和第二对象向量。

[0100] 第一文字向量的获取方式包括:方式1,在数据库中存储了具有对应关系的第一文字和第一文字向量的情况下,从数据库获取第一文字向量;或者,方式2,在数据库中未存储具有对应关系的第一文字和第一文字向量的情况下,生成第三随机向量作为第一文字向量。

[0101] 第二文字向量的获取方式包括:方式1,在数据库中存储了具有对应关系的第二文字和第二文字向量的情况下,从数据库获取第二文字向量;或者,方式2,在数据库中未存储具有对应关系的第二文字和第二文字向量的情况下,生成第三随机向量作为第二文字向量。

[0102] 第一偏旁向量的获取方式包括:方式1,在数据库中存储了具有对应关系的第一偏旁和第一偏旁向量的情况下,从数据库获取第一偏旁向量;或者,方式2,在数据库中未存储具有对应关系的第一偏旁和第一偏旁向量的情况下,生成第四随机向量作为第一偏旁向量。

[0103] 第二偏旁向量的获取方式包括:方式1,在数据库中存储了具有对应关系的第二偏旁和第二偏旁向量的情况下,从数据库获取第二偏旁向量;或者,方式2,在数据库中未存储具有对应关系的第二偏旁和第二偏旁向量的情况下,生成第四随机向量作为第二偏旁向量。

[0104] 需要说明的是,对于前述的各方法实施例,为了简单描述,故将其都表述为一系列

的动作组合,但是本领域技术人员应该知悉,本发明并不受所描述的动作顺序的限制,因为依据本发明,某些步骤可以采用其他顺序或者同时进行。其次,本领域技术人员也应该知悉,说明书中所描述的实施例均属于优选实施例,所涉及的动作和模块并不一定是本发明所必须的。

[0105] 通过以上的实施方式的描述,本领域的技术人员可以清楚地了解到根据上述实施例的方法可借助软件加必需的通用硬件平台的方式来实现,当然也可以通过硬件,但很多情况下前者是更佳的实施方式。基于这样的理解,本发明的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来,该计算机软件产品存储在一个存储介质(如ROM/RAM、磁碟、光盘)中,包括若干指令用以使得一台终端设备(可以是手机,计算机,服务器,或者网络设备等)执行本发明各个实施例所述的方法。

[0106] 根据本发明实施例,还提供了一种用于实施上述文本向量的确定方法的文本向量的确定装置,如图11所示,该装置包括:

[0107] 1) 第一提取模块112,用于从训练文本中提取第一对象和第二对象,其中,第一对象和第二对象为训练文本中不同的对象;

[0108] 2) 获取模块114,用于获取第一对象对应的第一对象向量、第一文字对应的第一文字向量、第一偏旁对应的第一偏旁向量、第二对象对应的第二对象向量、第二文字对应的第二文字向量和第二偏旁对应的第二偏旁向量,其中,第一文字为第一对象包括的文字,第一偏旁为第一文字包括的偏旁,第二文字为第二对象包括的文字,第二偏旁为第二文字包括的偏旁;

[0109] 3) 预测模块116,用于由第一向量向第二向量进行预测得到预测结果,并生成使得预测结果满足预测目标的目标第一向量和目标第二向量,其中,第一向量和第二向量为第一对象向量、第一文字向量、第一偏旁向量、第二对象向量、第二文字向量和第二偏旁向量中的任一向量;

[0110] 4) 确定模块118,用于将目标第一向量和目标第二向量确定为目标文本向量。

[0111] 可选地,在本实施例中,上述文本向量的确定装置可以但不限于应用于自然语言处理任务中生成文本向量的场景中。其中,上述自然语言处理任务可以但不限于为任何以文本为基础的任务,例如:基于深度学习的中文分词,词性标注,命名实体识别,搜索,机器翻译,文本分类,资源推荐等等。具体的,可以但不限于应用于在上述基于深度学习的中文分词任务中生成文本向量的场景中,或还可以但不限于应用于在上述机器翻译任务中生成文本向量的场景中,以提高生成文本向量时对文本的语义资源的利用率。上述仅是一种示例,本实施例中对此不做任何限定。

[0112] 可选地,在本实施例中,文字的偏旁可以为偏旁中的部首。例如:对于文字“好”,该文字的偏旁为“女”和“子”,部首为“女”,上述获取模块获取的偏旁可以为该字的全部偏旁“女”和“子”,也可以为该字的部首部分“女”。

[0113] 需要说明的是,由于部首一般情况下即是一个汉字语义的载体,例如:形声字的偏旁可分为行部和声部,其部首通常是形部,表达该字的意义,而声部表达该字的发音,因此,在本实施例中,以获取的偏旁为部首为例而不是获取所有偏旁。获取所有偏旁生成文本向量的方式与此类似,再次不再赘述。

[0114] 可选地,在本实施例中,文本中的对象可以但不限于词语或者短语等等。例如:

词语可以是“青蛙”，短语则可以是“绿色的青蛙”，那么如果对象为词语“青蛙”，该词语对象对应的文字则为“青”和“蛙”，该词语对象对应的偏旁为“青”和“虫”。如果对象为短语“绿色的青蛙”，那么该词语对象对应的文字则为“绿”、“色”、“的”、“青”和“蛙”，该词语对象对应的偏旁为“纟”、“色”、“白”、“青”和“虫”。

[0115] 可选地，在本实施例中，第一向量可以但不限于为第一偏旁向量、第一文字向量和第一对象向量，第二向量可以但不限于为第一文字向量、第一对象向量、第二偏旁向量、第二文字向量和第二对象向量。

[0116] 可选地，在本实施例中，由第一向量向第二向量进行预测可以包括：由第一偏旁向量预测第一文字向量，由第一文字向量预测第一对象向量，由第一偏旁向量预测第二偏旁向量，由第一文字向量预测第二偏旁向量，由第一文字向量预测第二文字向量，由第一对象向量预测第二文字向量，以及由第一对象向量预测第二对象向量。

[0117] 可选地，在本实施例中，上述由第一对象向量预测第二对象向量可以替换为由第一文字向量预测第二对象向量，上述由第一文字向量预测第二偏旁向量可以替换为由第一偏旁向量预测第二文字向量。

[0118] 可见，通过上述装置，在文本向量的确定过程中，将组成词汇的组件（例如：文字和偏旁）所携带的信息加入到词汇向量的学习过程中，并能够同步得到这些组件的向量，从而充分地利用了文本的语义资源，生成的文本向量也能更充分地表达出语义信息，使得生成文本向量时对文本的语义资源的利用率得到了显著地提高，进而克服现有技术中生成文本向量时无语义信息以及对文本的语言资源利用率较低的问题。

[0119] 作为一种可选的方案，预测模块包括：

[0120] 1) 执行单元，用于重复执行以下步骤，直至得到的预测准确率收敛，将使得预测准确率收敛的第一向量确定为目标第一向量并将使得预测准确率收敛的第二向量确定为目标第二向量，其中，预测结果满足预测目标包括预测准确率收敛：

[0121] 2) 输入单元，用于将第一向量作为预测函数的输入参数输入预测函数，得到预测向量；

[0122] 3) 第一确定单元，用于确定预测向量与第二向量之间的预测准确率；

[0123] 4) 修正单元，用于在预测准确率未收敛的情况下，修正第一向量和第二向量，生成当前新的第一向量和当前新的第二向量。

[0124] 可选地，在本实施例中，预测函数可以但不限于是预先设置的。确定预测准确率的方式可以但不限于为确定预测向量与当前的输出向量之间的相似度。例如：通过计算预测向量与当前的输出向量之间的欧式距离确定相似度等等。

[0125] 可选地，在本实施例中，预测准确率收敛可以但不限于指预测准确率达到稳定值。

[0126] 作为一种可选的方案，预测模块包括以下之一：

[0127] 第一预测单元，用于由第一偏旁向量预测第一文字向量，由第一文字向量预测第一对象向量，由第一偏旁向量预测第二偏旁向量，由第一文字向量预测第二偏旁向量，由第一文字向量预测第二文字向量，由第一对象向量预测第二文字向量，以及由第一对象向量预测第二对象向量。

[0128] 在一个可选的实施方式中，如图5所示，预测模型包括三层，其中，第一层为偏旁层，E(1)为第一偏旁向量，D(1)为第二偏旁向量，第二层为文字层，E(2)为第一文字向量，D



(2)为第二文字向量,第三层为对象层,E(3)为第一对象向量,D(3)为第二对象向量,如图5中箭头所示为方式1中的预测过程,该预测过程包括:由E(1)预测E(2),由E(2)预测E(3),由E(1)预测D(1),由E(2)预测D(1),由E(2)预测D(2),由E(3)预测D(2),以及由E(3)预测D(3)。

[0129] 第二预测单元,用于由第一偏旁向量预测第一文字向量,由第一文字向量预测第一对象向量,由第一偏旁向量预测第二偏旁向量,由第一偏旁向量预测第二文字向量,由第一文字向量预测第二文字向量,由第一文字向量预测第二对象向量,以及由第一对象向量预测第二对象向量。

[0130] 在一个可选的实施方式中,如图6所示,上述方式1中的由第一文字向量预测第二偏旁向量可以替换为由第一偏旁向量预测第二文字向量,即将由E(2)预测D(1)替换为由E(1)预测D(2),由第一对象向量预测第二文字向量可以替换为由第一文字向量预测第二对象向量,即将由E(3)预测D(2)替换为E(2)预测D(3)。

[0131] 第三预测单元,用于由第一偏旁向量预测第一文字向量,由第一偏旁向量预测第二偏旁向量,由第一文字向量预测第二偏旁向量,以及由第一文字向量预测第二文字向量。

[0132] 在一个可选的实施方式中,如图7所示,预测模型包括两层,其中,第一层为偏旁层,E(1)为第一偏旁向量,D(1)为第二偏旁向量,第二层为文字层,E(2)为第一文字向量,D(2)为第二文字向量,如图7中箭头所示为方式3中的预测过程,该预测过程包括:由E(1)预测E(2),由E(1)预测D(1),由E(2)预测D(1),由E(2)预测D(2)。

[0133] 第四预测单元,用于由第一偏旁向量预测第一对象向量,由第一偏旁向量预测第二偏旁向量,由第一对象向量预测第二偏旁向量,以及由第一对象向量预测第二对象向量。

[0134] 在一个可选的实施方式中,如图8所示,预测模型包括两层,其中,第一层为偏旁层,E(1)为第一偏旁向量,D(1)为第二偏旁向量,第二层为对象层,E(3)为第一对象向量,D(3)为第二对象向量,如图8中箭头所示为方式4中的预测过程,该预测过程包括:由E(1)预测E(3),由E(1)预测D(1),由E(3)预测D(1),以及由E(3)预测D(3)。

[0135] 第五预测单元,用于由第一文字向量预测第一对象向量,由第一文字向量预测第二文字向量,由第一对象向量预测第二文字向量,以及由第一对象向量预测第二对象向量。

[0136] 在一个可选的实施方式中,如图9所示,预测模型包括两层,其中,第一层为文字层,E(2)为第一文字向量,D(2)为第二文字向量,第二层为对象层,E(3)为第一对象向量,D(3)为第二对象向量,如图9中箭头所示为方式5中的预测过程,该预测过程包括:由E(2)预测E(3),由E(2)预测D(2),由E(3)预测D(2),以及由E(3)预测D(3)。

[0137] 作为一种可选的方案,第一偏旁为第一部首,第二偏旁为第二部首,其中,确定模块包括:

[0138] 1) 第二确定单元,用于从目标第一向量和目标第二向量中确定目标文本向量,其中,目标文本向量包括以下至少之一:第一目标对象向量、第一目标文字向量、第一目标部首向量、第二目标对象向量、第二目标文字向量、第二目标部首向量,其中,第一目标对象向量为第一对象对应的使得预测结果满足预测目标的向量、第一目标文字向量为第一文字对应的使得预测结果满足预测目标的向量、第一目标部首向量为第一部首对应的使得预测结果满足预测目标的向量,第二目标对象向量为第二对象对应的使得预测结果满足预测目标的向量、第二目标文字向量为第二文字对应的使得预测结果满足预测目标的向量、第二目标部首向量为第二部首对应的使得预测结果满足预测目标的向量;

[0139] 2) 存储单元,用于将目标文本向量存储在数据库中。

[0140] 可选地,在本实施例中,上述数据库可以存储了具有对应关系的第一对象和第一对象向量、具有对应关系的第一文字和第一文字向量、具有对应关系的第一偏旁和第一偏旁向量、具有对应关系的第二对象和第二对象向量、具有对应关系的第二文字和第二文字向量、具有对应关系的第二偏旁和第二偏旁向量。那么,将目标文本向量存储在数据库中的过程可以但不限于是对数据库中已存储的具有对应关系的文本与文本向量进行更新。

[0141] 例如:数据库中原存储了具有对应关系的第一对象和第一对象向量、具有对应关系的第一文字和第一文字向量、具有对应关系的第一偏旁和第一偏旁向量、具有对应关系的第二对象和第二对象向量、具有对应关系的第二文字和第二文字向量以及具有对应关系的第二偏旁和第二偏旁向量,那么,在得到第一目标对象向量、第一目标文字向量、第一目标偏旁向量、第二目标对象向量、第二目标文字向量、第二目标偏旁向量后,可以将数据库中存储的上述内容更新为具有对应关系的第一对象和第一目标对象向量、具有对应关系的第一文字和第一目标文字向量、具有对应关系的第一偏旁和第一目标偏旁向量、具有对应关系的第二对象和第二目标对象向量、具有对应关系的第二文字和第二目标文字向量以及具有对应关系的第二偏旁和第二目标偏旁向量。

[0142] 作为一种可选的方案,第一提取模块包括:

[0143] 提取单元,用于从训练文本中随机提取第一对象,并从训练文本中包括的除第一对象之外的其他对象中随机提取第二对象,其中,训练文本中的对象包括以下至少之一:短语,词语。

[0144] 可选地,在本实施例中,提取的第一对象可以为一个短语,或者为一个词语,或者为一个短语和组成该短语的词语。如果将第一对象向量、第一文字向量和第一偏旁向量作为预测模型的输入层,对于不同形式的第一对象,输入层的形式也不相同,例如:第一对象为一个短语(例如:第一短语)的情况下,输入层为第一短语向量、第一文字向量和第一偏旁向量,第一对象为一个词语(例如:第一词语)的情况下,输入层为第一词语向量、第一文字向量和第一偏旁向量,第一对象为一个短语和组成该短语的词语(例如:第一短语和第一词语,其中,第一词语是组成第一短语的词语)的情况下,输入层为第一短语向量、第一词语向量、第一文字向量和第一偏旁向量。需要说明的是,对于第二对象的提取形式与第一对象类似,在此不在赘述。

[0145] 可选地,在本实施例中,提取单元可以但不限于通过以下方式提取第一对象:重复执行以下步骤,直至生成的第一随机数落入第一阈值范围,并将使第一随机数落入第一阈值范围的第一目标对象确定为第一对象:从训练文本中随机提取第一目标对象;为第一目标对象生成第一随机数;在第一随机数未落入第一阈值范围的情况下,从训练文本中随机提取一个目标对象,作为新的第一目标对象。

[0146] 可选地,在本实施例中,提取单元可以但不限于通过以下方式提取第二对象:在确定了第一对象之后,重复执行以下步骤,直至第二随机数落入第二阈值范围,并将使第二随机数落入第二阈值范围的第二目标对象确定为第二对象:从训练文本中包括的除第一对象之外的其他对象中随机提取第二目标对象;为第二目标对象生成第二随机数;在第二随机数未落入第二阈值范围的情况下,从训练文本中包括的除第一对象和第二目标对象之外的其他对象中随机提取一个目标对象,作为新的第二目标对象。

[0147] 作为一种可选的方案,上述装置还包括:

[0148] 1) 第二提取模块,用于提取目标文本向量;

[0149] 2) 训练模块,用于使用目标文本向量训练自然语言处理模型;

[0150] 3) 执行模块,用于使用自然语言处理模型执行自然语言处理操作。

[0151] 可选地,在本实施例中,确定的目标文本向量可以应用于自然语言处理任务中,可以使用确定的目标文本向量训练自然语言处理模型,并使用训练好的模型执行自然语言处理操作。例如:可以使用确定的目标文本向量训练文本分类模型,并利用训练后的文本分类模型执行文本资源分类的操作。

[0152] 作为一种可选的方案,获取模块用于:获取第一对象对应的第一对象向量和第二对象对应的第二对象向量;从第一对象中提取组成第一对象的第一文字,从第二对象中提取组成第二对象的第二文字,并获取第一文字对应的第一文字向量和第二文字对应的第二文字向量;从第一文字中提取第一偏旁,从第二文字中提取第二偏旁,并获取第一偏旁对应的第一偏旁向量和第二偏旁对应的第二偏旁向量。

[0153] 可选地,在本实施例中,如果获取的第一对象或者第二对象是首次获取的对象,那么可以为首次获取的对象随机生成一个随机向量,如果该对象不是首次获取的对象,那么可以在数据库中获得数据库记录的具有对应关系的该对象和该对象的向量。对于文字向量和偏旁向量的获取,可以首先查找数据库中是否已经存储了该文字向量或者偏旁向量,如果已经存储,则可以直接提取使用,如果未存储,则可以为其生成一个随机向量。

[0154] 在一个可选的实施方式中,各个向量的获取可以但不限于采用以下方式:

[0155] 第一对象向量的获取方式包括:方式1,在第一对象为首次提取的对象的情况下,生成第一随机向量作为第一对象向量;或者,方式2,在第一对象为非首次提取的对象的情况下,获取数据库中存储的第一对象向量,其中,数据库中记录了具有对应关系的第一对象和第一对象向量。

[0156] 第二对象向量的获取方式包括:方式1,在第二对象为首次提取的对象的情况下,生成第二随机向量作为第二对象向量;或者,方式2,在第二对象为非首次提取的对象的情况下,获取数据库中存储的第二对象向量,其中,数据库中记录了具有对应关系的第二对象和第二对象向量。

[0157] 第一文字向量的获取方式包括:方式1,在数据库中存储了具有对应关系的第一文字和第一文字向量的情况下,从数据库获取第一文字向量;或者,方式2,在数据库中未存储具有对应关系的第一文字和第一文字向量的情况下,生成第三随机向量作为第一文字向量。

[0158] 第二文字向量的获取方式包括:方式1,在数据库中存储了具有对应关系的第二文字和第二文字向量的情况下,从数据库获取第二文字向量;或者,方式2,在数据库中未存储具有对应关系的第二文字和第二文字向量的情况下,生成第三随机向量作为第二文字向量。

[0159] 第一偏旁向量的获取方式包括:方式1,在数据库中存储了具有对应关系的第一偏旁和第一偏旁向量的情况下,从数据库获取第一偏旁向量;或者,方式2,在数据库中未存储具有对应关系的第一偏旁和第一偏旁向量的情况下,生成第四随机向量作为第一偏旁向量。

[0160] 第二偏旁向量的获取方式包括:方式1,在数据库中存储了具有对应关系的第二偏旁和第二偏旁向量的情况下,从数据库获取第二偏旁向量;或者,方式2,在数据库中未存储具有对应关系的第二偏旁和第二偏旁向量的情况下,生成第四随机向量作为第二偏旁向量。

[0161] 本发明实施例的应用环境可以但不限于参照上述实施例中的应用环境,本实施例中对此不再赘述。本发明实施例提供了用于实施上述文本向量的确定方法的一种可选的具体应用示例。

[0162] 作为一种可选的实施例,上述文本向量的确定方法可以但不限于应用于如图12所示的对文本向量进行确定的场景中。

[0163] 考虑到汉语的特点,在本实施例中,提出的模型的原理是为了将汉语词汇的组件(汉字和偏旁)所携带的信息加入到词汇向量的学习过程中,并同步得到这些组件的向量。该模型符合汉语语言组织形式及对汉字词汇生成过程的认知,例如,在想要表达某种语义信息的时候,人们面临不同的词语选择,而选定的词语会受到上文其它词语的影响,同时也会影响接下来的下文,特别对于汉语而言,选定的词语包含的汉字和偏旁,以及这些组件与上下文其它词语及其组件的关系,携带了比英语等西方语言更丰富的细粒度语义信息,而这些信息在以往的词向量工作中并未得到充分利用。

[0164] 本实施例提出的模型框架如图12所示,在本实施例中,文本中的对象以词语为例,每个汉语词语按照不同的颗粒度分为三层,包括(1)偏旁(部首)层;(2)文字层;以及(3)词语层。每一层包括输入(E)和输出(D)两个不同的向量,分别对应输入和输出词语及其组件。

[0165] 上述模型的训练过程如下所述:

[0166] 步骤1,在训练文本中,针对每个输入样本(一组词构成的上下文),取其中一个词作为输入词(E(3)),并在其它上下文词中随机选择一个词语作为输出词(D(3)),取出他们的向量形成第3层,其中,如果一个词语是第一次输入的词语则获取一个随机向量作为其初始向量);

[0167] 步骤2,分别取出输入词和输出词的组成文字以及这些文字的组成偏旁,得到它们的向量构成第2层和第1层(如“葡萄”得到“葡”和“萄”,进而得到“艹”)。

[0168] 步骤3,按照图8所示的实线箭头学习同层或者不同层向量之间的关系,该关系的学习方式是使用箭头起始位置的向量预测箭头终止位置的向量。整个预测过程包括:E(1)预测E(2)和D(1),E(2)预测E(3)、D(1)和D(2),E(3)预测D(2)和D(3)。

[0169] 步骤4,在整个语料上重复步骤1至步骤3,直到模型收敛,即每组向量之间的预测准确率接近稳定值。

[0170] 按照上述模型的训练过程可以最终得到整个语料上所有词语及其组件的向量。尤其是,在学习各个层级的向量的过程中,考虑了其它层级的信息对该层级的影响,例如在学习词汇级的向量的过程中,引入了该词汇的构成汉字对它的影响(E(2)→E(3)),以及偏旁对汉字的影响(E(1)→E(2))、词汇对汉字的影响等等。由于不同颗粒度的语汇单元的向量互相依赖,因此当模型收敛的时候,可以认为整个模型对各个不同层级的向量之间的关系描述实现了最优化。

[0171] 本实施例中所提出的模型有一些以往工作所不具备的独特性:

[0172] 首先,整个模型从结构上模拟了汉语书写过程中构词的组件及其关系。针对任意

输入的词语对,左边为源词,右边为目标词,如图8所示的结构中左边自底向上的箭头表示了从偏旁到汉字进而到词语的生成过程,而右边的虚线箭头则示意了选定的词语所包含的组件的分解关系。汉语的书写或生成过程涉及到选词、构词、确定下文、再选词的循环过程,因此,按照这样的逻辑学习汉语词汇及其组件的关系可以有效提升所得的向量的质量。

[0173] 其次,本模型的分层结构可拆解,即模型中的某个或多个层级可以移除,剩下的层级依然可以构成完整的模型。当分别去除词层(W)、字层(C)、偏旁层(R)的时候,如图13所示的模型则形成三个双层模型,这些模型依然可以学习不同的词汇组件的向量。举例来说,图13中第二个模型去除了汉字层,因此相当于同步学习词汇和偏旁的向量,也即在向量的学习过程中仅考虑偏旁和词汇之间的关系。同样地,也可以将原模型去除两层,仅学习词或字亦或偏旁的向量。

[0174] 此外,由于使用了汉语词的内部组件,因此本模型可以认为隐式地利用了某种语义资源(汉字和偏旁所携带的信息),因此可以在较少数据上实现高质量的汉语词向量计算结果,或者在给定相等数据的情况下,较之其他模型得到更好的向量。得益于词语内部信息的有效利用,本模型可以使用较少的计算资源(例如主流台式机CPU+16G内存)在受限(例如10万词规模)数据基础上得到的中文词向量可以媲美在大规模(超过100万词规模)数据上使用其他方法得到的结果。

[0175] 可选地,在本实施例中,提出的模型可以不局限于对偏旁、字、词的三级语汇颗粒度进行向量学习,还可以扩展到短语和组块(chunk)级。

[0176] 可选地,在本实施例中,图12中的交叉学习可以更改顺序和方向,即从源词到目标词的字,以及源字到目标偏旁的预测可以替换为源字到目标词以及源偏旁到目标字等;

[0177] 可选地,在本实施例中,词到词的预测函数可以更改为CBOW形式,同时带来字和偏旁预测方式发生改变,即可以不仅使用一个词及其组件对目标词进行预测,而是使用一组词与其组件预测目标词。

[0178] 通过本实施例提出的一种同步(联合)学习中文词,字及偏旁三级不同颗粒度文本向量的模型,模拟中文构词的过程,充分利用了中文构词方法中字和偏旁带来的语义信息,由此得到的各级向量可以显著提高其语义(语法)表达能力。更重要的是,与同类方法相比,由于有效挖掘中文词语内部携带的语义信息,本模型可以在训练数据较小(或同等数据规模)的情况下学习得到较好的向量,因此在工业环境中具有非常显著的使用价值。由本模型所训练的向量可以用于各种基于深度学习的自然语言处理任务中,例如:新闻阅读、网页搜索、广告推荐、聊天系统等。

[0179] 通过本模型可以在使用较少数据的基础上得到较好的词向量结果,尤其适合自然语言处理中存在的领域内(in-domain)冷启动(cold-start)的情况。考虑到中文处理中存在的分词问题,任何新领域的语料资源相对较少甚至没有,于是在有限的人工分词的基础上,可以通过本模型有效习得较高水平的词向量,以便于增强后续任务的性能;

[0180] 此外,由于本模型所设计的学习框架的灵活性,可以针对不同的场景使用不同的资源和组合方式学习不同语汇单位的向量。例如在音译或人名较多的使用场景,可以不使用偏旁信息,以克服偏旁对这类词汇带来的负面影响(这类词汇的组合一般不按照汉语的构词特点,因此其构成单元往往无法拆分出有意义的语义信息)。

[0181] 根据本发明实施例的又一个方面,还提供了一种用于实施上述引导语的传输方法的电子装置,如图14所示,该电子装置可以包括:一个或多个(图中仅示出一个)处理器1402、存储器1404、传感器1406、编码器1408以及传输装置1410。

[0182] 其中,存储器1404可用于存储软件程序以及模块,如本发明实施例中的视频图像的播放方法及装置。

[0183] 对应的程序指令/模块,处理器1402通过运行存储在存储器1404内的软件程序以及模块,从而执行各种功能应用以及数据处理,即图像编码方法。存储器1404可包括高速随机存储器,还可以包括非易失性存储器,如一个或者多个磁性存储装置、闪存、或者其他非易失性固态存储器。在一些实例中,存储器1404可进一步包括相对于处理器1402远程设置的存储器,这些远程存储器可以通过网络连接至终端。上述网络的实例包括但不限于互联网、企业内部网、局域网、移动通信网及其组合。

[0184] 上述的传输装置1410用于经由一个网络接收或者发送数据。上述的网络具体实例可包括有线网络及无线网络。在一个实例中,传输装置1410包括一个网络适配器(Network Interface Controller, NIC),其可通过网线与其他网络设备与路由器相连从而可与互联网或局域网进行通讯。在一个实例中,传输装置1410为射频(Radio Frequency, RF)模块,其用于通过无线方式与互联网进行通讯。

[0185] 可选地,本实施例中的具体示例可以参考上述实施例中所描述的示例,本实施例在此不再赘述。

[0186] 本领域普通技术人员可以理解,图14所示的结构仅为示意,电子装置也可以是智能手机(如Android手机、iOS手机等)、平板电脑、掌上电脑以及移动互联网设备(Mobile Internet Devices, MID)、PAD等终端设备。图14其并不对上述电子装置的结构造成限定。例如,电子装置还可包括比图14中所示更多或者更少的组件(如网络接口、显示装置等),或者具有与图14所示不同的配置。

[0187] 本领域普通技术人员可以理解上述实施例的各种方法中的全部或部分步骤是可以通程序来指令终端设备相关的硬件来完成,该程序可以存储于一计算机可读存储介质中,存储介质可以包括:闪存盘、只读存储器(Read-Only Memory, ROM)、随机存取器(Random Access Memory, RAM)、磁盘或光盘等。

[0188] 本发明的实施例还提供了一种存储介质。可选地,在本实施例中,上述存储介质可以位于网络中的多个网络设备中的至少一个网络设备。

[0189] 可选地,在本实施例中,存储介质被设置为存储用于执行以下步骤的程序代码:

[0190] S1,从训练文本中提取第一对象和第二对象,其中,第一对象和第二对象为训练文本中不同的对象;

[0191] S2,获取第一对象对应的第一对象向量、第一文字对应的第一文字向量、第一偏旁对应的第一偏旁向量、第二对象对应的第二对象向量、第二文字对应的第二文字向量和第二偏旁对应的第二偏旁向量,其中,第一文字为第一对象包括的文字,第一偏旁为第一文字包括的偏旁,第二文字为第二对象包括的文字,第二偏旁为第二文字包括的偏旁;

[0192] S3,由第一向量向第二向量进行预测得到预测结果,并生成使得预测结果满足预测目标的目标第一向量和目标第二向量,其中,第一向量和第二向量为第一对象向量、第一文字向量、第一偏旁向量、第二对象向量、第二文字向量和第二偏旁向量中的任一向量;

[0193] S4,将目标第一向量和目标第二向量确定为目标文本向量。

[0194] 可选地,在本实施例中,上述存储介质可以包括但不限于:U盘、只读存储器(ROM, Read-Only Memory)、随机存取存储器(RAM,Random Access Memory)、移动硬盘、磁碟或者光盘等各种可以存储程序代码的介质。

[0195] 可选地,本实施例中的具体示例可以参考上述实施例中所描述的示例,本实施例在此不再赘述。

[0196] 上述本发明实施例序号仅仅为了描述,不代表实施例的优劣。

[0197] 上述实施例中的集成的单元如果以软件功能单元的形式实现并作为独立的产品销售或使用,可以存储在上述计算机可读的存储介质中。基于这样的理解,本发明的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的全部或部分可以以软件产品的形式体现出来,该计算机软件产品存储在存储介质中,包括若干指令用以使得一台或多台计算机设备(可为个人计算机、服务器或者网络设备)执行本发明各个实施例所述方法的全部或部分步骤。

[0198] 在本发明的上述实施例中,对各个实施例的描述都各有侧重,某个实施例中未详述的部分,可以参见其他实施例的相关描述。

[0199] 在本申请所提供的几个实施例中,应该理解到,所揭露的客户端,可通过其它的方式实现。其中,以上所描述的装置实施例仅仅是示意性的,例如所述单元的划分,仅仅为一种逻辑功能划分,实际实现时可以有另外的划分方式,例如多个单元或组件可以结合或者可以集成到另一个系统,或一些特征可以忽略,或不执行。另一点,所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口,单元或模块的间接耦合或通信连接,可以是电性或其它的形式。

[0200] 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单元来实现本实施例方案的目的。

[0201] 另外,在本发明各个实施例中的各功能单元可以集成在一个处理单元中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现,也可以采用软件功能单元的形式实现。

[0202] 以上所述仅是本发明的优选实施方式,应当指出,对于本技术领域的普通技术人员来说,在不脱离本发明原理的前提下,还可以做出若干改进和润饰,这些改进和润饰也应视为本发明的保护范围。

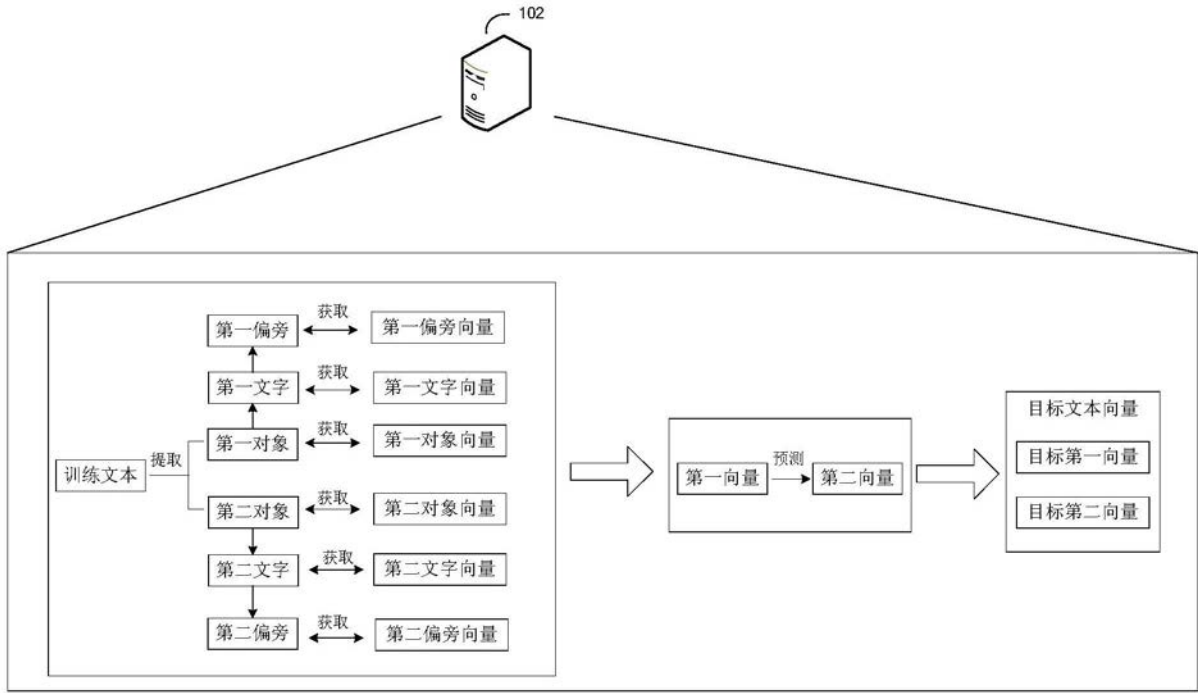


图1

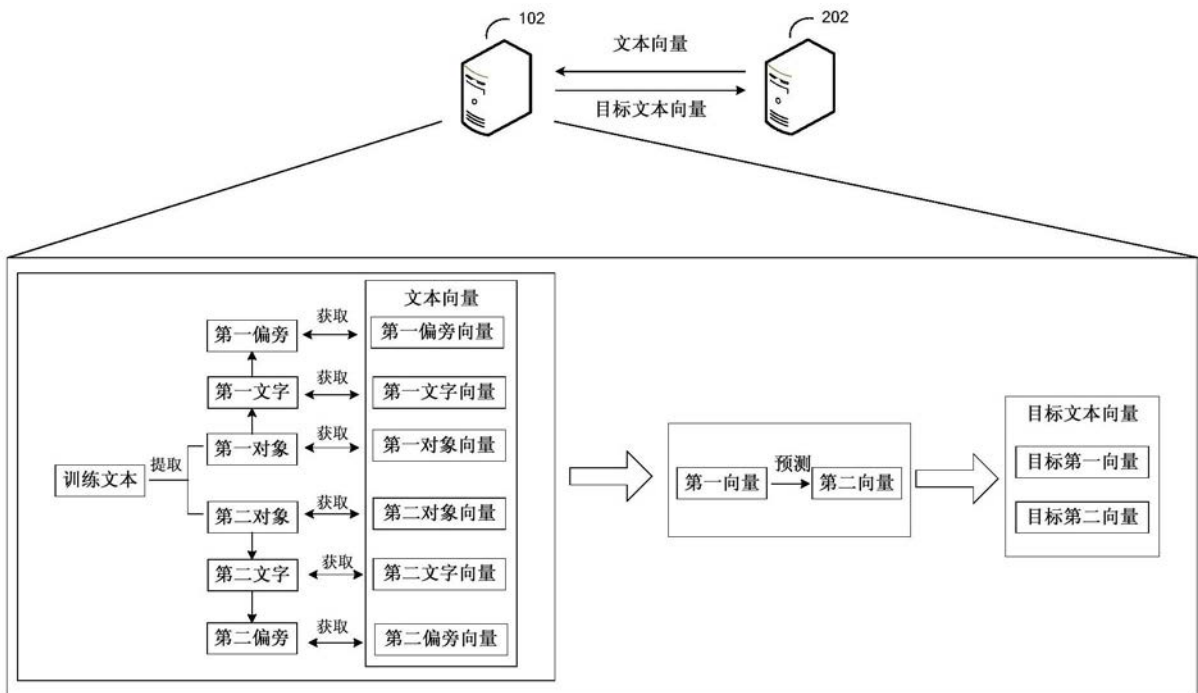


图2



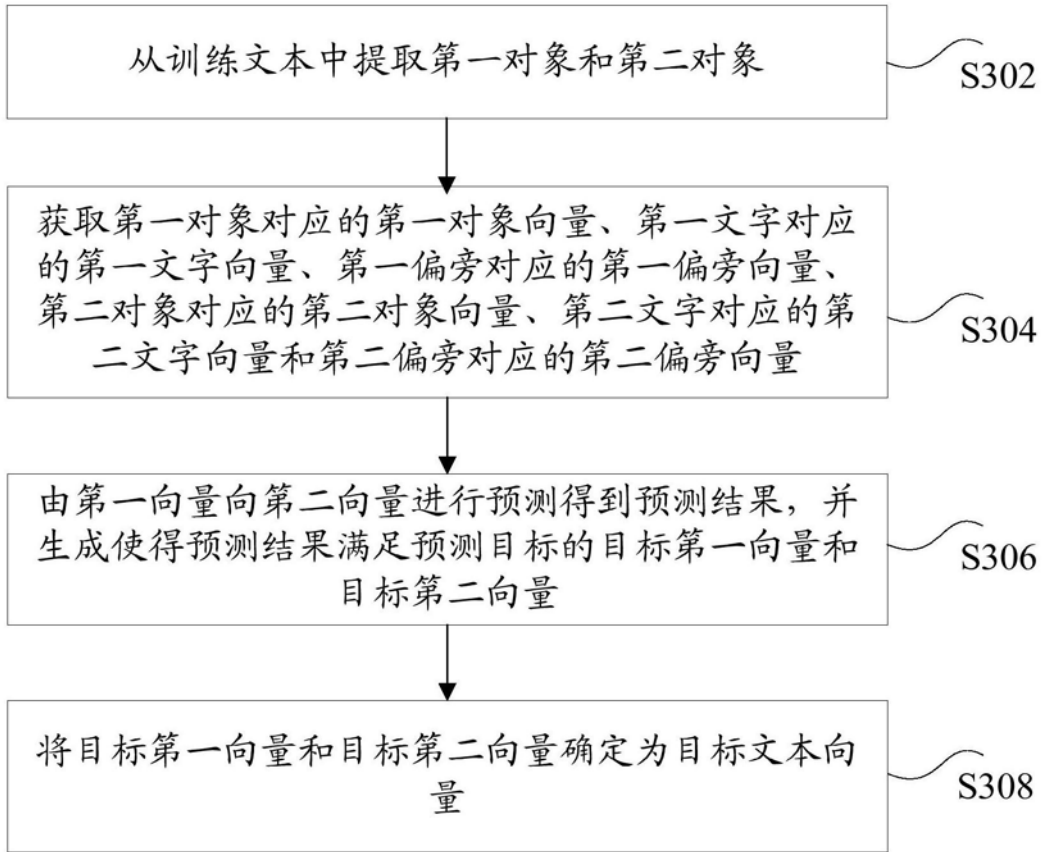


图3

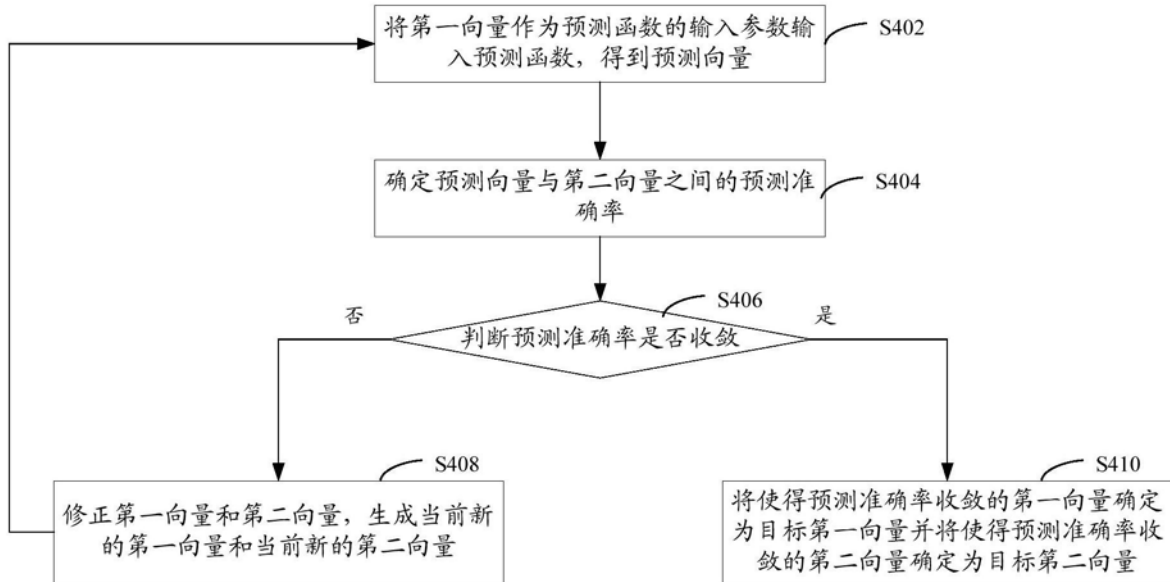


图4

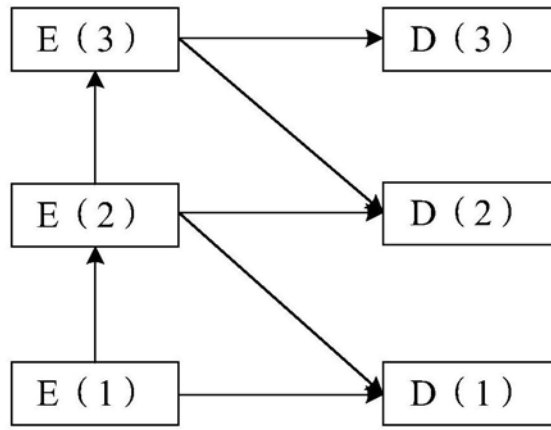


图5

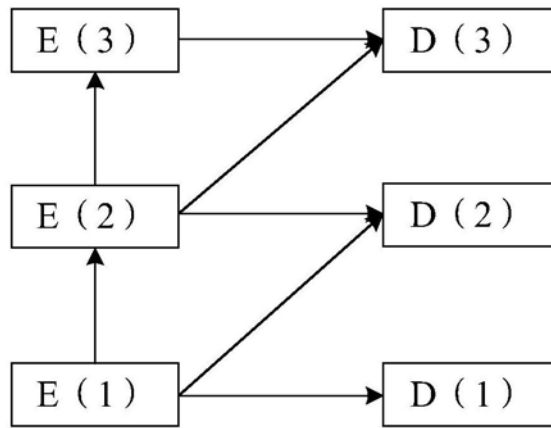


图6

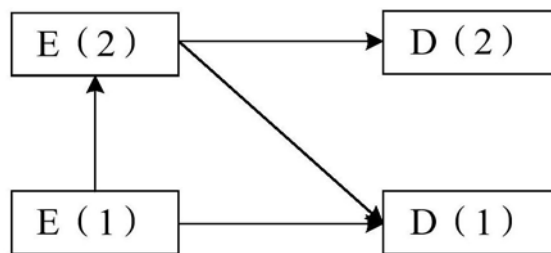


图7

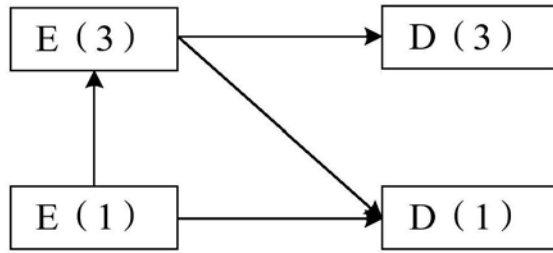


图8

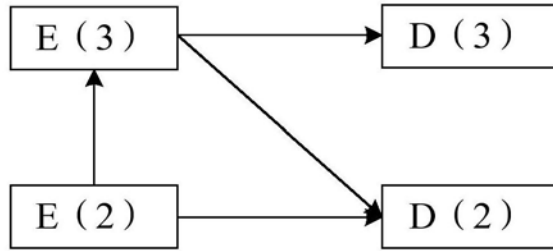


图9

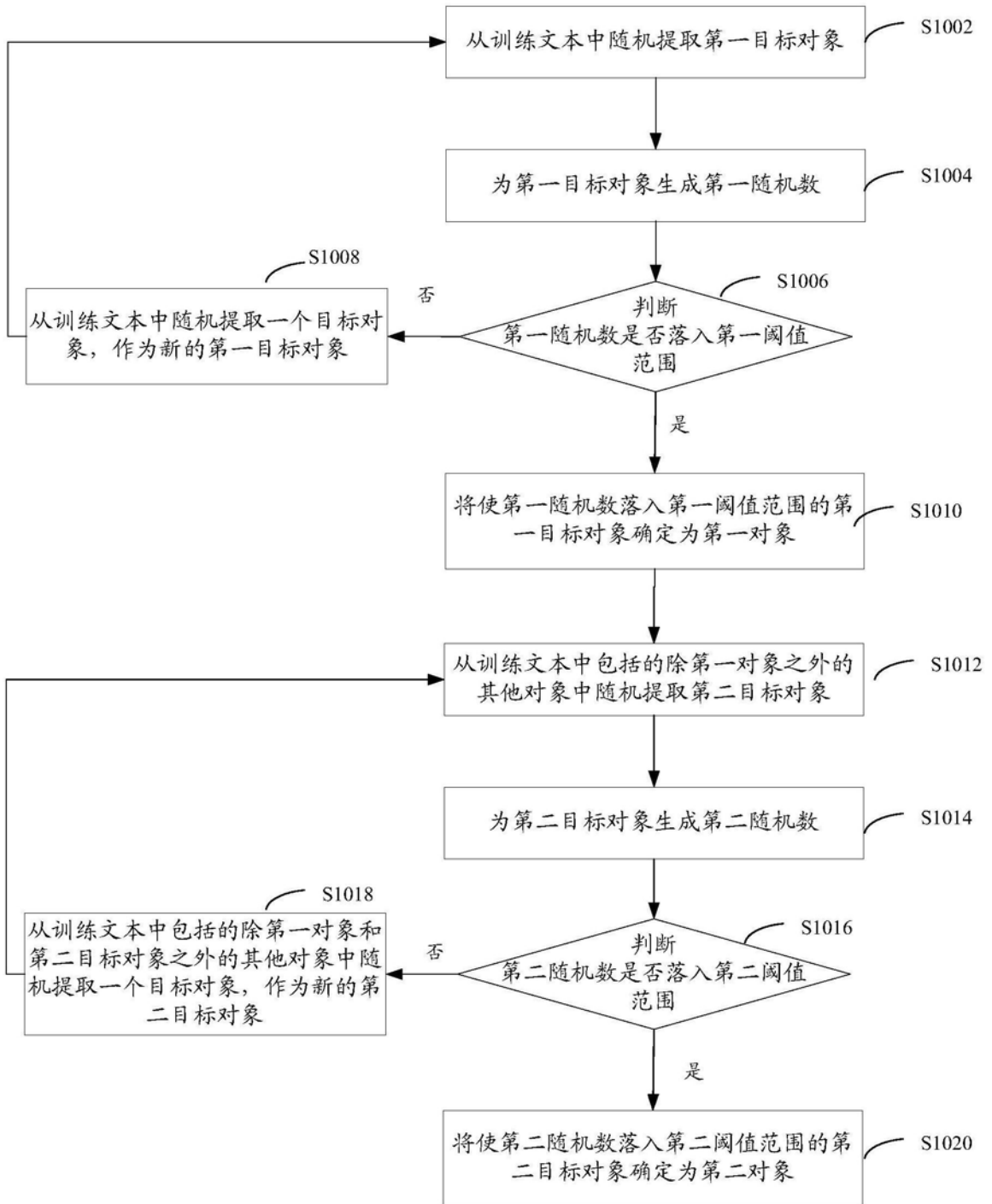


图10

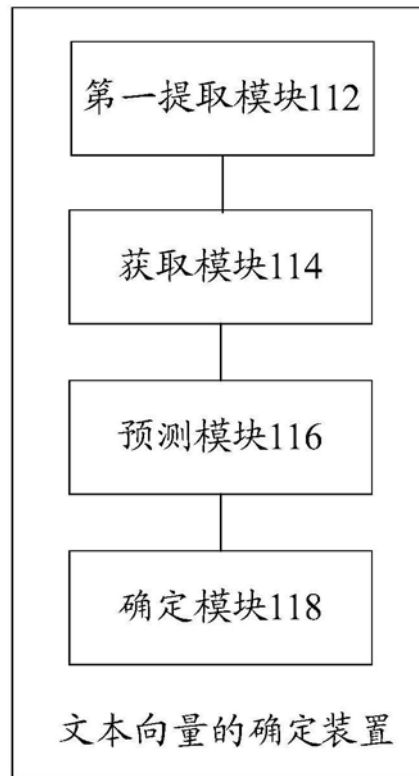


图11

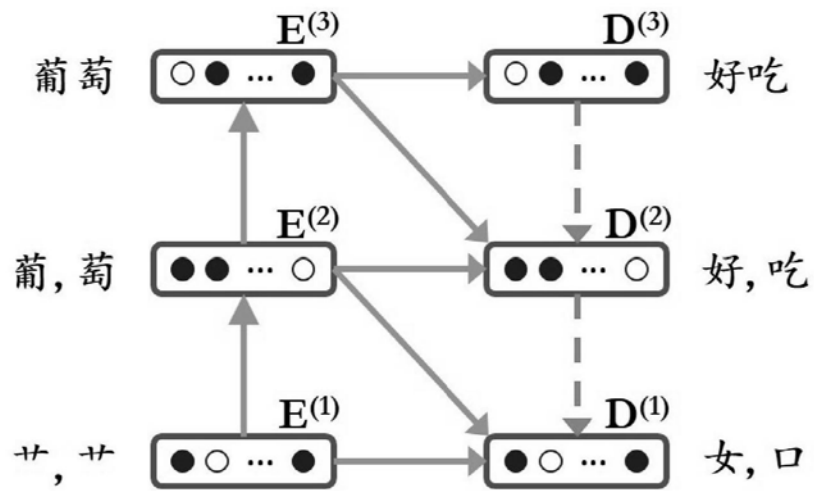


图12

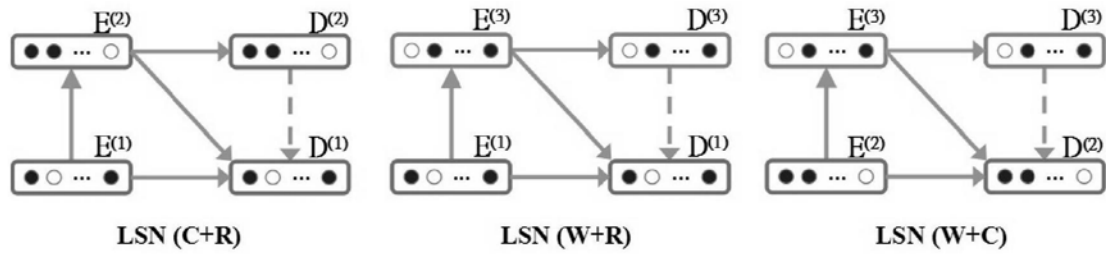


图13

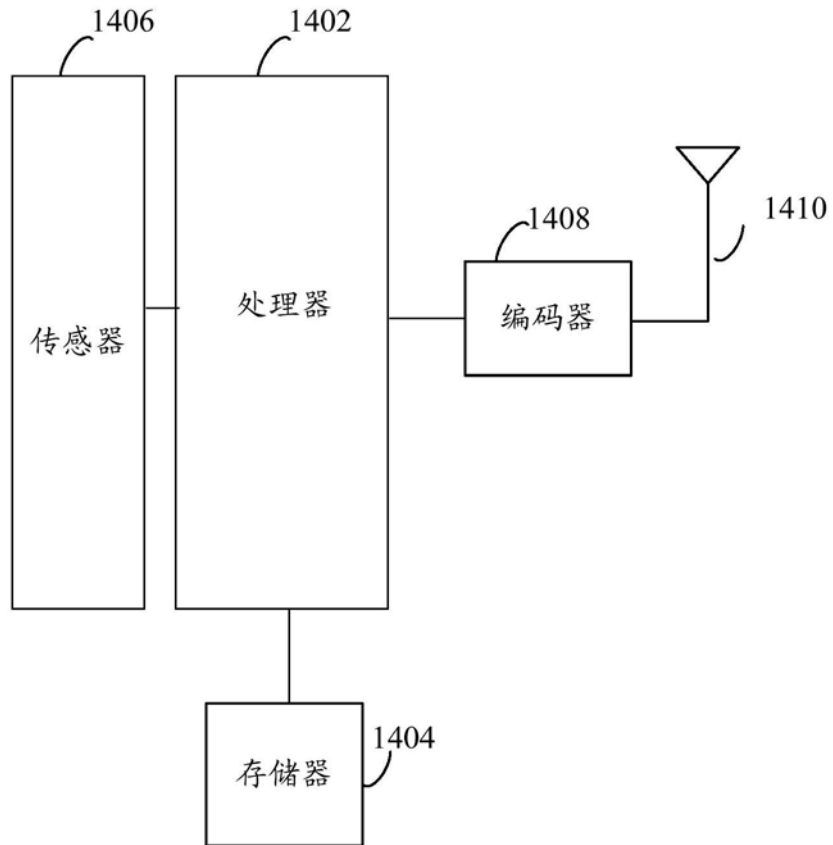


图14