



(12) 发明专利申请

(10) 申请公布号 CN 111767694 A

(43) 申请公布日 2020.10.13

(21) 申请号 201910231478.8

(22) 申请日 2019.03.26

(71) 申请人 北京京东尚科信息技术有限公司
地址 100086 北京市海淀区知春路76号8层
申请人 北京京东世纪贸易有限公司

(72) 发明人 刘瑞雪 陈佩佩 陈蒙 刘晓华

(74) 专利代理机构 中国贸促会专利商标事务所
有限公司 11038
代理人 李昊 刘剑波

(51) Int. Cl.

G06F 40/126 (2020.01)

G06N 3/04 (2006.01)

G06F 16/332 (2019.01)

G06F 16/33 (2019.01)

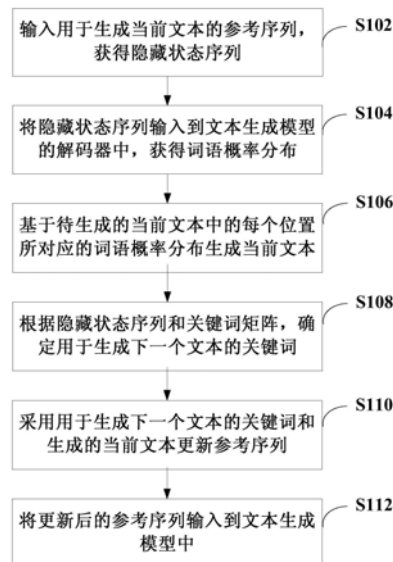
权利要求书3页 说明书12页 附图6页

(54) 发明名称

文本生成方法、装置和计算机可读存储介质

(57) 摘要

本发明公开了一种文本生成方法、装置和计算机可读存储介质,涉及自然语言处理领域。文本生成方法包括:将用于生成当前文本的参考序列输入到文本生成模型的编码器中,获得基于编码器的输出而确定的隐藏状态序列;将隐藏状态序列输入到文本生成模型的解码器中,获得待生成的当前文本中的每个位置所对应的词语概率分布;基于待生成的当前文本中的每个位置所对应的词语概率分布生成当前文本;根据隐藏状态序列和预设的关键词矩阵,确定用于生成下一个文本的关键词;采用用于生成下一个文本的关键词和生成的当前文本更新参考序列,其中,更新后的参考序列用于生成下一个文本;将更新后的参考序列输入到文本生成模型中。



1. 一种文本生成方法, 包括:

将用于生成当前文本的参考序列输入到文本生成模型的编码器中, 获得基于编码器的输出而确定的隐藏状态序列;

将所述隐藏状态序列输入到文本生成模型的解码器中, 获得待生成的当前文本中的每个位置所对应的词语概率分布;

基于所述待生成的当前文本中的每个位置所对应的词语概率分布生成当前文本;

根据所述隐藏状态序列和预设的关键词矩阵, 确定用于生成下一个文本的关键词;

采用用于生成下一个文本的关键词和生成的当前文本更新参考序列, 其中, 更新后的参考序列用于生成下一个文本;

将更新后的参考序列输入到文本生成模型中。

2. 根据权利要求1所述的文本生成方法,

还包括:

获取用于生成当前文本的输入文本;

将所述输入文本进行反向处理, 以生成反向输入文本; 以及,

生成输入文本对应的正向序列、以及反向输入文本对应的反向序列, 作为用于生成当前文本的参考序列;

所述将用于生成当前文本的参考序列输入到文本生成模型的编码器中, 获得基于编码器的输出确定的隐藏状态序列包括:

将用于生成当前文本的参考序列输入到文本生成模型的编码器中, 获得正向隐藏状态序列和反向隐藏状态序列; 以及,

将正向隐藏状态序列和反向隐藏状态序列中位于相同位置的隐藏状态进行拼接, 获得隐藏状态序列。

3. 根据权利要求1或2所述的文本生成方法, 其中, 所述根据所述隐藏状态序列和预设的关键词矩阵, 确定用于生成下一个文本的关键词包括:

将隐藏状态序列中的最后一个隐藏状态和预先获取的关键词矩阵的乘积与预设的偏置向量之和输入到激活函数中, 获得词库中每个词对应的概率值, 其中, 关键词矩阵包括词库中每个词对应的预设的向量;

将概率值最大的预设数量个词确定为用于生成下一个文本的关键词。

4. 根据权利要求1所述的文本生成方法, 其中, 所述基于所述待生成的当前文本中的每个位置所对应的词语概率分布生成当前文本包括:

获取待生成的当前文本的第 i 个位置对应的、预设数量个不同的候选序列, 其中, 每个候选序列中包括待生成的当前文本的前 $i-1$ 个位置中每个位置所对应的词语, 每个词语具有在相应的位置上出现的概率值, 其中, i 为大于1的整数;

根据待生成的当前文本的第 i 个位置对应的词语概率分布, 确定词库中的词语在第 i 个位置出现的概率值;

将词库中的词语与第 i 个位置对应的所有候选序列进行组合, 并根据每种组合中的词语的概率值, 确定每个组合的备选值;

采用备选值最大的预设数量个组合更新候选序列, 获得待生成的当前文本的第 $i+1$ 个位置对应的、预设数量个不同的候选序列, 以便采用待生成的当前文本的最后一个位置对

应的、备选值最大的候选序列生成当前文本。

5. 根据权利要求4所述的文本生成方法,所述根据每种组合中的词语的概率值,确定每个组合的备选值包括:

将组合中的每个词对应的概率值的对数值的总和与所述组合中词的数量比值,确定为所述组合的备选值。

6. 根据权利要求1所述的文本生成方法,还包括:

获取用户输入的文本、音频、图像中的至少一种,作为用户输入信息;

将用户输入信息映射为包括一个或多个向量的序列,作为用于生成首个文本的参考序列。

7. 根据权利要求6所述的文本生成方法,其中,在用户输入信息为文本的情况下,在所述将用户输入信息映射为包括多个向量的序列之前,所述文本生成方法还包括:

在用户输入信息中的词语数量大于预设值的情况下,将用户输入信息截取为预设的长度;

在用户输入信息中的词语数量小于预设值的情况下,将用户输入信息填充为预设的长度。

8. 根据权利要求1所述的文本生成方法,其中,所述文本生成模型的编码器和解码器为循环神经网络或者卷积神经网络。

9. 根据权利要求1所述的文本生成方法,还包括:

获取训练文本,其中,所述训练文本中的关键词被预先进行标记;

将用于训练的参考序列输入到文本生成模型的编码器中,获得基于编码器的输出而确定的、训练文本的隐藏状态序列;

将所述训练文本的隐藏状态序列输入到文本生成模型的解码器中,获得训练文本中的每个位置所对应的词语概率分布;

基于所述训练文本中的每个位置所对应的词语概率分布生成预测文本;

根据所述训练文本的隐藏状态序列和关键词矩阵,确定下一个文本对应的预测关键词;

根据预测文本与训练文本的差距、以及预测关键词与所述训练文本的下一个文本的关键词的差距,对文本生成模型和关键词矩阵的参数进行调整,直到预测文本与训练文本的差距、以及预测关键词与所述训练文本的下一个文本的关键词的差距小于预设值。

10. 一种文本生成装置,包括:

隐藏状态序列确定模块,被配置为将用于生成当前文本的参考序列输入到文本生成模型的编码器中,获得基于编码器的输出而确定的隐藏状态序列;

词语概率分布确定模块,被配置为将所述隐藏状态序列输入到文本生成模型的解码器中,获得待生成的当前文本中的每个位置所对应的词语概率分布;

文本生成模块,被配置为基于所述待生成的当前文本中的每个位置所对应的词语概率分布生成当前文本;

关键词确定模块,被配置为根据所述隐藏状态序列和预设的关键词矩阵,确定用于生成下一个文本的关键词;

参考信息更新模块,被配置为采用用于生成下一个文本的关键词和生成的当前文本更

新参考序列,其中,更新后的参考序列用于生成下一个文本;

输入模块,被配置为将更新后的参考序列输入到隐藏状态序列确定模块。

11.一种文本生成装置,包括:

存储器;以及

耦接至所述存储器的处理器,所述处理器被配置为基于存储在所述存储器中的指令,执行如权利要求1~9中任一项所述的文本生成方法。

12.一种计算机可读存储介质,其上存储有计算机程序,该程序被处理器执行时实现权利要求1~9中任一项所述的文本生成方法。

文本生成方法、装置和计算机可读存储介质

技术领域

[0001] 本发明涉及自然语言处理领域,特别涉及一种文本生成方法、装置和计算机可读存储介质。

背景技术

[0002] 随着人工智能技术的发展,基于文本生成(Natural Language Generation)的项目越来越多,从而节省了很多人力资源成本和经济成本。相关技术采用机器学习的方法生成文本。传统的基于机器学习的模型主要包括基于概率统计的机器翻译模型,其本质是在已知源语言的条件下,用概率统计的方式,计算出与源语言的共现概率最高的目标语言的组合。深度学习的方法则是通过模拟人脑的神经元,设计和使用深度神经网络(Neural Network)来建立不同复杂程度和深度的非线性学习模型。通过使用大量训练数据,模型能够学习到自然语言中的语义信息。通过分析源语言的特点以及源语言与目标语言的关系,模型能够对特定题材或者主题的文本进行预测。

发明内容

[0003] 发明人对相关技术进行分析后发现,相关技术生成的文本存在文本结构不合理、内容单一的问题,从而导致生成的文本可读性很差。

[0004] 本发明实施例所要解决的一个技术问题是:如何提高自动生成的文本的可读性。

[0005] 根据本发明一些实施例的第一个方面,提供一种文本生成方法,包括:将用于生成当前文本的参考序列输入到文本生成模型的编码器中,获得基于编码器的输出而确定的隐藏状态序列;将隐藏状态序列输入到文本生成模型的解码器中,获得待生成的当前文本中的每个位置所对应的词语概率分布;基于待生成的当前文本中的每个位置所对应的词语概率分布生成当前文本;根据隐藏状态序列和预设的关键词矩阵,确定用于生成下一个文本的关键词;采用用于生成下一个文本的关键词和生成的当前文本更新参考序列,其中,更新后的参考序列用于生成下一个文本;将更新后的参考序列输入到文本生成模型中。

[0006] 在一些实施例中,文本生成方法还包括:获取用于生成当前文本的输入文本;将输入文本进行反向处理,以生成反向输入文本;以及,生成输入文本对应的正向序列、以及反向输入文本对应的反向序列,作为用于生成当前文本的参考序列;将用于生成当前文本的参考序列输入到文本生成模型的编码器中,获得基于编码器的输出确定的隐藏状态序列包括:将用于生成当前文本的参考序列输入到文本生成模型的编码器中,获得正向隐藏状态序列和反向隐藏状态序列;以及,将正向隐藏状态序列和反向隐藏状态序列中位于相同位置的隐藏状态进行拼接,获得隐藏状态序列。

[0007] 在一些实施例中,根据隐藏状态序列和预设的关键词矩阵,确定用于生成下一个文本的关键词包括:将隐藏状态序列中的最后一个隐藏状态和预先获取的关键词矩阵的乘积与预设的偏置向量之和输入到激活函数中,获得词库中每个词对应的概率值,其中,关键词矩阵包括词库中每个词对应的预设的向量;将概率值最大的预设数量个词确定为用于生

成下一个文本的关键词。

[0008] 在一些实施例中,基于待生成的当前文本中的每个位置所对应的词语概率分布生成当前文本包括:获取待生成的当前文本的第 i 个位置对应的、预设数量个不同的候选序列,其中,每个候选序列中包括待生成的当前文本的前 $i-1$ 个位置中每个位置所对应的词语,每个词语具有在相应的位置上出现的概率值,其中, i 为大于1的整数;根据待生成的当前文本的第 i 个位置对应的词语概率分布,确定词库中的词语在第 i 个位置出现的概率值;将词库中的词语与第 i 个位置对应的所有候选序列进行组合,并根据每种组合中的词语的概率值,确定每个组合的备选值;采用备选值最大的预设数量个组合更新候选序列,获得待生成的当前文本的第 $i+1$ 个位置对应的、预设数量个不同的候选序列,以便采用待生成的当前文本的最后一个位置对应的、备选值最大的候选序列生成当前文本。

[0009] 在一些实施例中,根据每种组合中的词语的概率值,确定每个组合的备选值包括:将组合中的每个词对应的概率值的对数值的总和与组合中词的数量比值,确定为组合的备选值。

[0010] 在一些实施例中,文本生成方法还包括:获取用户输入的文本、音频、图像中的至少一种,作为用户输入信息;将用户输入信息映射为包括一个或多个向量的序列,作为用于生成首个文本的参考序列。

[0011] 在一些实施例中,在用户输入信息为文本的情况下,在将用户输入信息映射为包括多个向量的序列之前,文本生成方法还包括:在用户输入信息中的词语数量大于预设值的情况下,将用户输入信息截取为预设的长度;在用户输入信息中的词语数量小于预设值的情况下,将用户输入信息填充为预设的长度。

[0012] 在一些实施例中,文本生成模型的编码器和解码器为循环神经网络或者卷积神经网络。

[0013] 在一些实施例中,文本生成方法还包括:获取训练文本,其中,训练文本中的关键词被预先进行标记;将用于训练的参考序列输入到文本生成模型的编码器中,获得基于编码器的输出而确定的、训练文本的隐藏状态序列;将训练文本的隐藏状态序列输入到文本生成模型的解码器中,获得训练文本中的每个位置所对应的词语概率分布;基于训练文本中的每个位置所对应的词语概率分布生成预测文本;根据训练文本的隐藏状态序列和关键词矩阵,确定下一个文本对应的预测关键词;根据预测文本与训练文本的差距、以及预测关键词与训练文本的下一个文本的关键词的差距,对文本生成模型和关键词矩阵的参数进行调整,直到预测文本与训练文本的差距、以及预测关键词与训练文本的下一个文本的关键词的差距小于预设值。

[0014] 根据本发明一些实施例的第二个方面,提供一种文本生成装置,包括:隐藏状态序列确定模块,被配置为将用于生成当前文本的参考序列输入到文本生成模型的编码器中,获得基于编码器的输出而确定的隐藏状态序列;词语概率分布确定模块,被配置为将隐藏状态序列输入到文本生成模型的解码器中,获得待生成的当前文本中的每个位置所对应的词语概率分布;文本生成模块,被配置为基于待生成的当前文本中的每个位置所对应的词语概率分布生成当前文本;关键词确定模块,被配置为根据隐藏状态序列和预设的关键词矩阵,确定用于生成下一个文本的关键词;参考信息更新模块,被配置为采用用于生成下一个文本的关键词和生成的当前文本更新参考序列,其中,更新后的参考序列用于生成下一

个文本;输入模块,被配置为将更新后的参考序列输入到隐藏状态序列确定模块。

[0015] 根据本发明一些实施例的第三个方面,提供一种文本生成装置,包括:存储器;以及耦接至存储器的处理器,处理器被配置为基于存储在存储器中的指令,执行前述任意一种文本生成方法。

[0016] 根据本发明一些实施例的第四个方面,提供一种计算机可读存储介质,其上存储有计算机程序,其中,该程序被处理器执行时实现前述任意一种文本生成方法。

[0017] 上述发明中的一些实施例具有如下优点或有益效果:本发明的实施例能够通过编码器的隐藏状态序列并行地预测当前的文本和下一个文本的关键词,并且在生成下一个文本时利用这些信息进行预测。这种将源语言的规划和文本的生成过程同时进行的方式,能够使得生成的多个连续文本具有连贯的信息。从而,提高了自动生成的文本的可读性,提高了文本生成的效率。

[0018] 通过以下参照附图对本发明的示例性实施例的详细描述,本发明的其它特征及其优点将会变得清楚。

附图说明

[0019] 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性的前提下,还可以根据这些附图获得其他的附图。

[0020] 图1为根据本发明一些实施例的文本生成方法的流程示意图。

[0021] 图2为根据本发明一些实施例的隐藏状态确定方法的流程示意图。

[0022] 图3为根据本发明一些实施例的关键词生成方法的流程示意图。

[0023] 图4为根据本发明一些实施例的根据词语概率分布生成文本的方法的流程示意图。

[0024] 图5为根据本发明一些实施例的文本生成模型的训练方法的流程示意图。

[0025] 图6为根据本发明一些实施例的端到端文本生成方法的流程示意图。

[0026] 图7为根据本发明一些实施例的文本生成装置的结构示意图。

[0027] 图8为根据本发明另一些实施例的文本生成装置的结构示意图。

[0028] 图9为根据本发明又一些实施例的文本生成装置的结构示意图。

具体实施方式

[0029] 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。以下对至少一个示例性实施例的描述实际上仅仅是说明性的,决不作为对本发明及其应用或使用的任何限制。基于本发明中的实施例,本领域普通技术人员在没有作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。

[0030] 除非另外具体说明,否则在这些实施例中阐述的部件和步骤的相对布置、数字表达式和数值不限制本发明的范围。

[0031] 同时,应当明白,为了便于描述,附图中所示出的各个部分的尺寸并不是按照实际

的比例关系绘制的。

[0032] 对于相关领域普通技术人员已知的技术、方法和设备可能不作详细讨论,但在适当情况下,所述技术、方法和设备应当被视为授权说明书的一部分。

[0033] 在这里示出和讨论的所有示例中,任何具体值应被解释为仅仅是示例性的,而不是作为限制。因此,示例性实施例的其它示例可以具有不同的值。

[0034] 应注意到:相似的标号和字母在下面的附图中表示类似项,因此,一旦某一项在一个附图中被定义,则在随后的附图中不需要对其进行进一步讨论。

[0035] 发明人经过分析后发现,在相关技术的文本生成框架中,源语言的规划方式有多种。例如,可以通过概率统计的方法,根据语言模型统计出与用户定义的主题词的关联性最高的首句文字;又例如,可以基于语义相似度模型或者共享知识图谱扩展出与用户定义的主题词相似的关键词序列,并通过对关键词序列的规划与限定,扩展出相对应的文本。然后,在确定源语言的基础上,再进一步地对首句文字或者相关的关键词序列进行学习,从而生成多句文字。

[0036] 然而,将生成的首个语句作为源语言以生成后文的方式,既无法保证首个语句对全文的影响力,也无法保证生成的文本与主题词的关联性。而且如果第一句话的生成与用户定义的主题词有偏差,那么这种偏差也会随着后面几句文本的生成而更加明显。

[0037] 而基于主题词进行关键词序列的扩展的方式,对关键词的生成具有较高的质量要求。因为只有合理、高质量的关键词才能保证其扩展出的目标文本具有较强的可读性。相关技术所采用的方式会产生堆积相似词语的后果,例如,从“冬天”扩展到“冷”、“雪花”等,从而不能保证扩展的关键词的语义丰富度。

[0038] 发明人认识到,相关技术将源语言的规划过程和文本的生成过程分别进行,二者的相关性较小,从而无法保证生成的文本的质量和连贯性。因此,发明人经过进一步研究,提出了将源语言规划和文本生成过程同时进行的方式,以提高生成的文本的质量。下面参考图1描述本发明文本生成方法的实施例。

[0039] 图1为根据本发明一些实施例的文本生成方法的流程示意图。如图1所示,该实施例的文本生成方法包括步骤S102~S112。

[0040] 在步骤S102中,将用于生成当前文本的参考序列输入到文本生成模型的编码器中,获得基于编码器的输出而确定的隐藏状态序列。例如,可以直接将编码器的输出结果作为隐藏状态序列。

[0041] 参考序列为用于生成文本的信息所对应的序列。用于生成文本的信息例如可以包括关键词、前文文本等文本类信息,也可以是音频、图像等其他类型的信息。在一些实施例中,可以通过预设的映射关系将用于生成文本的信息映射为多维向量,由这些向量构成参考序列。

[0042] 例如,用于生成文本的信息包括“举头望明月故乡”,其中,“举头望明月”为前文文本,“故乡”为待生成的当前文本的关键词。则参考序列可以为 $\{V_{\text{举}}, V_{\text{头}}, V_{\text{望}}, V_{\text{明月}}, V_{\text{故乡}}\}$,其中, $V_{\text{举}}$ 、 $V_{\text{头}}$ 、 $V_{\text{望}}$ 、 $V_{\text{明月}}$ 、 $V_{\text{故乡}}$ 分别为“举”、“头”、“望”、“明月”、“故乡”所对应的词向量。

[0043] 文本生成模型包括编码器和解码器。编码器用于将输入的向量转换为隐藏状态,解码器用于基于编码器的输出来确定生成的文本中各个位置对应的词语概率分布。编码器的编码过程例如可以如公式(1)所示。

[0044] $h_m = \text{Encoder}(h_{m-1}, [x_m])$ (1)

[0045] 其中, m 表示参考序列中的各个元素所在的位置的标识; h_m 表示位置 m 所对应的隐藏状态; $\text{Encoder}(\cdot)$ 表示编码器; x_m 表示参考序列中的第 m 个元素所对应的向量。当 $m=1$ 时, 可以令编码器的输入为 0, 以确定参考序列的首个位置所对应的隐藏状态 h_1 。

[0046] 在一些实施例中, 文本生成模型的编码器和解码器为循环神经网络或者卷积神经网络 (Convolutional Neural Networks, 简称: CNN)。循环神经网络例如可以包括长短记忆网络 (Long Short-Term Memory, 简称: LSTM)、门控循环单元 (Gated Recurrent Unit, 简称: GRU) 等等。

[0047] 在步骤 S104 中, 将隐藏状态序列输入到文本生成模型的解码器中, 获得待生成的当前文本中的每个位置所对应的词语概率分布。

[0048] 解码器在解码过程中也会产生隐藏状态。设编码器所产生的隐藏状态为第一隐藏状态, 解码器所产生的隐藏状态为第二隐藏状态。则在一些实施例中, 可以将基于多个第一隐藏状态生成的、待生成的当前文本中的每个位置所对应的上下文向量输入到解码器中, 获得待生成的当前文本中的每个位置所对应的第二隐藏状态; 然后根据每个第二隐藏状态以及预设的权重向量矩阵, 确定待生成文本中的每个位置所对应的词语概率分布。

[0049] 在一些实施例中, 可以采用公式 (2) 确定解码器所产生的第二隐藏状态。

[0050] $s_n = \text{Decoder}(s_{n-1}, [c_n; e(y_{n-1})])$ (2)

[0051] 在公式 (2) 中, n 表示待生成的当前文本中的各个词所在的位置; s_n 表示待生成的当前文本中的第 n 个位置所对应的第二隐藏状态; $\text{Decoder}(\cdot)$ 表示解码器; c_n 表示待生成文本中的第 n 个位置所对应的上下文向量; y_{n-1} 表示待生成的当前文本中的第 $n-1$ 个位置上所对应的词语概率分布; $e(y_{n-1})$ 表示基于 y_{n-1} 确定的相应词语的词向量; $[\cdot]$ 表示将中括号中的多个项进行合并后的结果。当 $n=1$ 时, 可以令解码器输入为 0, 以确定输入信息的首个位置所对应的第二隐藏状态 s_1 。

[0052] 此外, 可以采用公式 (3) ~ (5) 计算上下文向量 c_n 。

[0053] $\alpha_{mn} = f(s_{n-1}, h_m)$ (3)

[0054] $\alpha'_{mn} = \text{softmax}(\alpha_{mn})$ (4)

[0055] $c_n = \sum_{t'=0}^M \alpha'_{mt'} * h_{t'}$ (5)

[0056] 在公式 (3) 中, m 表示参考序列中的各个元素所在的位置; n 表示待生成的当前文本中的各个词所在的位置; α_{mn} 和 α'_{mn} 为中间变量, 用于反映输入的内容中各个位置所对应的得分信息; $f(\cdot)$ 为预设的激活函数, 用于表示函数中的两个项的匹配程度, 本领域技术人员可以根据需要来进行设置; s_{n-1} 表示待生成的当前文本中的第 $n-1$ 个位置所对应的第二隐藏状态; h_m 表示参考序列中第 m 个位置所对应的第一隐藏状态; c_n 表示待生成的当前文本中的第 n 个位置所对应的上下文向量, 其作用是基于注意力机制将输出与输入进行关联; M 表示参考序列中的元素总数。

[0057] 在确定了解码器产生的第二隐藏状态后, 可以计算位置 n 所对应的词语概率分布 y_n 。 y_n 中包括了词库中每个词所对应的概率。 y_n 的计算方式例如可以参考公式 (6)。

[0058] $y_n = P(y_n | y_1, y_2, y_3, \dots, y_{n-1}, c_n) = \text{softmax}(W s_n)$ (6)

[0059] 在公式 (6) 中, $P(y_n | y_1, y_2, y_3, \dots, y_{n-1}, c_n)$ 表示在待生成的当前文本中的前 $n-1$ 个位置分别对应词语概率分布 $y_1, y_2, y_3, \dots, y_{n-1}$ 并且上下文向量为 c_n 的情况下, 第 n 个位置的

词语概率分布为 y_n 的概率; W 表示预设的权重向量; s_n 表示待生成的当前文本中的第 n 个位置所对应的第二隐藏状态。

[0060] 在步骤S106中,基于待生成的当前文本中的每个位置所对应的词语概率分布生成当前文本。例如,可以将概率分布中最大概率值所对应的词语确定为相应位置上的词语。在确定了每个位置上的词语后,则可以获得生成的文本。

[0061] 在步骤S108中,根据隐藏状态序列和预设的关键词矩阵,确定用于生成下一个文本的关键词。例如,可以根据隐藏状态序列和预设的关键词矩阵的乘积确定对关键词的解析结果。

[0062] 在步骤S110中,采用用于生成下一个文本的关键词和生成的当前文本更新参考序列,其中,更新后的参考序列用于生成下一个文本。

[0063] 由于隐藏状态序列能够反映参考序列的信息以及生成的当前文本的信息,因此通过根据隐藏状态序列来生成用于生成下一个文本的关键词,能够使得根据生成的下一个文本的关键词与当前文本的信息更连贯。

[0064] 在步骤S112中,将更新后的参考序列输入到文本生成模型中,以重复执行步骤S102~S112的过程。本发明实施例的方法可以根据需要循环进行一次或多次,从而可以生成由多个连续的文本构成的一段文本。

[0065] 通过上述实施例的方法,能够通过编码器的隐藏状态序列并行地预测当前的文本和下一个文本的关键词,并且在生成下一个文本时利用这些信息进行预测。这种将源语言的规划和文本的生成过程同时进行的方式,能够使得生成的多个连续文本具有连贯的信息。从而,提高了自动生成的文本的可读性,提高了文本生成的效率。

[0066] 本发明的实施例可以应用于多种场景。例如,自动生成网络中的产品的描述文本、自动生成对来自网络用户的问题的应答文本等等。例如,用户输入“礼服”以生成一段描述礼服的文本。文本生成模型首先生成文本“这款时髦简约的礼物裁自白色绉纱”,并预测出下一个文本的关键词为“荷叶边”;然后,根据“荷叶边”和前一个文本“这款时髦简约的礼物裁自白色绉纱”,生成文本“衣袖上优雅垂坠的荷叶边饰更是品牌的标志性特色”,并预测出下一个文本的关键词为“亲肤”;根据“亲肤”和前一个文本“衣袖上优雅垂坠的荷叶边饰更是品牌的标志性特色”,生成文本“全衬里设计令穿着更为顺滑”,并预测出下一个文本的关键词为“宴会”;以及,根据“宴会”和前一个文本“全衬里设计令穿着更为顺滑”,生成文本“正面的开衩让你以轻松翩然的姿态演绎婚礼派对上的第一支舞”。从而,根据用户一次性输入的信息,即可以预测出语义丰富的一个文本段落。

[0067] 由于本发明的实施例生成的文本具备更好的可读性,用户可以快速地理解自动生成的文本所表达的内容,降低了用户由于不清楚文本内容而与文本的提供方进行反复沟通、确认的次数,节约了网络资源。

[0068] 在生成诗歌等特定体裁的文本时,需要满足语言的平仄变化规则、押韵等问题。本发明的一些实施例可以借助Bi-LSTM等双向循环神经网络生成隐藏状态序列,更好地适应此类场景。下面参考图2描述本发明隐藏状态序列确定方法的实施例。

[0069] 图2为根据本发明一些实施例的隐藏状态确定方法的流程示意图。如图2所示,该实施例的隐藏状态确定方法包括步骤S202~S210。

[0070] 在步骤S202中,获取用于生成当前文本的输入文本。

[0071] 在步骤S204中,将输入文本进行反向处理,以生成反向输入文本。

[0072] 例如,输入文本为“好雨知时节”,则反向输入文本为“节时知雨好”。

[0073] 在步骤S206中,生成输入文本对应的正向序列、以及反向输入文本对应的反向序列,作为用于生成当前文本的参考序列。

[0074] 在步骤S208中,将用于生成当前文本的参考序列输入到文本生成模型的编码器中,获得正向隐藏状态序列和反向隐藏状态序列。

[0075] 在步骤S210中,将正向隐藏状态序列和反向隐藏状态序列中位于相同位置的隐藏状态进行拼接,获得隐藏状态序列。

[0076] 例如,设参考序列为 $\{w_1, w_2, \dots, w_{n-1}, w_n\}$,基于参考序列获得的正向隐藏状态序列为 $\{h_1, h_2, \dots, h_{n-1}, h_n\}$,反向序列为 $\{w_n, w_{n-1}, \dots, w_2, w_1\}$,基于反向序列获得的反向隐藏状态序列为 $\{h'_1, h'_2, \dots, h'_{n-1}, h'_n\}$ 。则隐藏状态序列片 $= \{H_1, H_2, \dots, H_{n-1}, H_n\} = \{[h_1; h'_1], [h_2; h'_2], \dots, [h_{n-1}; h'_{n-1}], [h_n; h'_n]\}$ 。“ $[\]$ ”表示将中括号内的向量进行拼接,例如,设 $h_1 = [a_1, a_2, \dots, a_m]$, $h'_1 = [a'_1, a'_2, \dots, a'_m]$,则 $[h_1; h'_1] = [a_1, a_2, \dots, a_m, a'_1, a'_2, \dots, a'_m]$ 。在上述示例中,m和n为正整数。

[0077] 正向隐藏状态序列还原了原始的输入文本的信息,而反向隐藏状态序列是通过从原始输入文本的最后一个词开始处理而得到的。通过上述实施例的方法,隐藏状态序列可以在不破坏输入文本的上下文关系的基础上,反映文本的原始信息和韵律信息,使得生成的文本可读性更好。

[0078] 下面参考图3示例性地描述一种生成关键词的方法。

[0079] 图3为根据本发明一些实施例的关键词生成方法的流程示意图。如图3所示,该实施例的关键词生成方法包括步骤S302~S304。

[0080] 在步骤S302中,将隐藏状态序列中的最后一个隐藏状态和预先获取的关键词矩阵的乘积与预设的偏置向量之和输入到激活函数中,获得词库中每个词对应的概率值,其中,关键词矩阵包括词库中每个词对应的预设的向量。

[0081] 在一些实施例中,可以根据公式(7)的计算结果确定关键词。

[0082]
$$P = \text{softmax}(W \cdot h_n + b) \quad (7)$$

[0083] 其中,P表示关键词的概率分布,概率分布中包括了词库中每个词对应的概率值;softmax表示冲激函数;W表示关键词矩阵; h_n 表示隐藏状态序列中的最后一个隐藏状态;b表示预设的偏置参数。W和b的值可以根据预先训练的结果确定。

[0084] 在步骤S304中,将概率值最大的预设数量个词确定为用于生成下一个文本的关键词。例如,可以将概率值最大的一个词确定为用于生成下一个文本的关键词。

[0085] 隐藏状态序列中的每一个状态是根据前一个状态确定的。隐藏状态序列中的最后一个隐藏状态能够不仅能够反映输入序列中最后一个元素及其之前的元素的信息。因此,通过根据最后一个隐藏状态生成关键词,可以使得生成的关键词能够涵盖前文文本的信息,进而使得生成的下一个文本具备更好的可读性。

[0086] 为了提升生成的文本的多样性,本发明可以采用集束搜索(Beam_Search)的方式生成文本。在一些实施例中,在确定当前文本的第一个位置所对应的词语时,可以根据第一个位置所对应的词语概率分布,将概率值最高的预设数量个词语中的每一个确定为一个候选序列。此后,确定当前文本的第二个及之后位置的词语时,可以采用以下方法。下面参考

图4描述本发明根据词语概率分布生成文本的方法的实施例。

[0087] 图4为根据本发明一些实施例的根据词语概率分布生成文本的方法的流程示意图。如图4所示,该实施例的根据词语概率分布生成文本的方法包括步骤S402~S414。

[0088] 在步骤S402中,将*i*的值初始化为2,其中,*i*为生成的当前文本中词语所在的位置的标识。

[0089] 在步骤S404中,获取待生成的当前文本的第*i*个位置对应的、预设数量个不同的候选序列,其中,每个候选序列中包括待生成的当前文本的前*i*-1个位置中每个位置所对应的词语,每个词语具有在相应的位置上出现的概率值,其中,*i*为大于1的整数。

[0090] 在步骤S406中,根据待生成的当前文本的第*i*个位置对应的词语概率分布,确定词库中的词语在第*i*个位置出现的概率值。

[0091] 在步骤S408中,将词库中的词语与第*i*个位置对应的所有候选序列进行组合,并根据每种组合中的词语的概率值,确定每个组合的备选值。

[0092] 在一些实施例中,可以将组合中的每个词对应的概率值的对数值的总和与组合中词的数量的比值,确定为组合的备选值。备选值的一个示例性的确定方法可以参考公式(8)。

$$OV_i = \frac{\sum_{n=1}^i \log p_n}{i} \quad (8)$$

[0094] 在公式(8)中, OV_i 表示第*i*个位置所对应的组合的备选值,其中,*i*为大于1的整数; p_n 表示组合中第*n*个词语的概率值。通过这种标准化处理,组合或备选序列的备选值不会受到词的数量影响。从而,可以根据语句本身的语义而非语句长短的限制来生成文本。根据需要,本领域技术人员还可以选择其他的确定备选值的方法,这里不再赘述。

[0095] 在步骤S410中,采用备选值最大的预设数量个组合更新候选序列,获得待生成的当前文本的第*i*+1个位置对应的、预设数量个不同的候选序列。

[0096] 例如,在确定待生成的当前文本的第3个位置对应的词语时,候选序列例如可以包括:{春天,带走}、{春风,拂过}。将词库中的词语与第*i*个位置对应的所有候选序列进行组合后,发现{春天,带走,冬天}的备选值以及{春风,拂过,柳树}的备选值要高于其他的组合,则可以将候选序列更新为{春天,带走,冬天}、{春风,拂过,柳树}。

[0097] 在步骤S412中,将*i*的值加1,并回到步骤S404,直到达到结束条件。结束条件例如可以为生成的词语为结束符时,结束当前文本的生成过程。

[0098] 在步骤S414中,采用待生成的当前文本的最后一个位置对应的、备选值最大的候选序列生成当前文本。

[0099] 通过上述实施例的方法,可以根据多个候选序列依次生成当前文本的每个词语,并不断更新候选序列,使得文本的生成结果具有多样性,提高了生成的文本的可读性。

[0100] 本发明可以预先训练文本生成模型。下面参考图5描述本发明文本生成模型的训练方法的实施例。

[0101] 图5为根据本发明一些实施例的文本生成模型的训练方法的流程示意图。如图5所示,该实施例的训练方法包括步骤S502~S512。

[0102] 在步骤S502中,获取训练文本,其中,训练文本中的关键词被预先进行标记。

[0103] 例如,可以预先获取包括多个训练文本的训练集。训练集中可以包括一个或多个

段落,每个段落中可以包括一个或多个文本。从而,属于一个段落的多个文本具有上下文的关系。

[0104] 在一些实施例中,可以采用以下方法确定训练文本的关键词:首先,对训练文本进行分词处理;然后,根据基于得分、词性、长度中的任意一个或多个的条件筛选关键词,其中,得分的确定方法可以通过词频-逆文档频率指数(Term Frequency-Inverse Document Frequency,简称:TF-IDF)或者文本排序算法(TextRank)等方法。

[0105] 在步骤S504中,将用于训练的参考序列输入到文本生成模型的编码器中,获得基于编码器的输出而确定的、训练文本的隐藏状态序列。

[0106] 在步骤S506中,将训练文本的隐藏状态序列输入到文本生成模型的解码器中,获得训练文本中的每个位置所对应的词语概率分布。

[0107] 在步骤S508中,基于训练文本中的每个位置所对应的词语概率分布生成预测文本。

[0108] 在步骤S510中,根据训练文本的隐藏状态序列和关键词矩阵,确定下一个文本对应的预测关键词。

[0109] 在步骤S512中,根据预测文本与训练文本的差距、以及预测关键词与训练文本的下一个文本的关键词的差距,对文本生成模型和关键词矩阵的参数进行调整,直到预测文本与训练文本的差距、以及预测关键词与训练文本的下一个文本的关键词的差距小于预设值。

[0110] 通过上述训练过程,可以基于训练文本的已有的上下文关系以及关键词标记结果,对文本生成模型和关键词矩阵的参数进行训练,从而提高了生成的文本的可读性。

[0111] 本发明的一些实施例可以为用户提供端到端的文本生成方式。用户可以根据需要输入多种格式的用于生成文本的信息,也可以输入任意数量的关键词。下面参考图6和图7描述本发明文本生成方法的实施例。

[0112] 图6为根据本发明一些实施例的端到端文本生成方法的流程示意图。如图6所示,该实施例的端到端文本生成方法包括步骤S602~S606。

[0113] 在步骤S602中,获取用户输入的文本、音频、图像中的至少一种,作为用户输入信息。

[0114] 在一些实施例中,如果用户输入信息为文本,在用户输入信息中的词语数量大于预设值的情况下,将用户输入信息截取为预设的长度;在用户输入信息中的词语数量小于预设值的情况下,将用户输入信息填充为预设的长度。从而,用户可以根据需要输入任意长度的文本,提高了用户体验。并且,针对不同长度的文本,本发明的实施例可以处理为相同长度的用户输入信息,从而适用性更广。

[0115] 在步骤S602中,将用户输入信息映射为包括一个或多个向量的序列,作为用于生成首个文本的参考序列。

[0116] 在一些实施例中,对于每种类型的用户输入信息,可以采用相应的映射模型或者映射矩阵将其转换为参考序列。映射模型、映射矩阵的参数可以通过预先训练获得。

[0117] 在步骤S606中,采用文本生成模型,根据参考序列生成下一个文本和用于生成下一个文本的关键词。

[0118] 通过上述实施例的方法,用户可以根据需要选择输入信息的类型,从而可以适用

于多种应用场景。

[0119] 下面参考图7描述本发明文本生成装置的实施例。

[0120] 图7为根据本发明一些实施例的文本生成装置的结构示意图。如图7所示,该实施例的文本生成装置70包括:隐藏状态序列确定模块710,被配置为将用于生成当前文本的参考序列输入到文本生成模型的编码器中,获得基于编码器的输出而确定的隐藏状态序列;词语概率分布确定模块720,被配置为将隐藏状态序列输入到文本生成模型的解码器中,获得待生成的当前文本中的每个位置所对应的词语概率分布;文本生成模块730,被配置为基于待生成的当前文本中的每个位置所对应的词语概率分布生成当前文本;关键词确定模块740,被配置为根据隐藏状态序列和预设的关键词矩阵,确定用于生成下一个文本的关键词;参考信息更新模块750,被配置为采用用于生成下一个文本的关键词和生成的当前文本更新参考序列,其中,更新后的参考序列用于生成下一个文本;输入模块760,被配置为将更新后的参考序列输入到隐藏状态序列确定模块。

[0121] 在一些实施例中,文本生成装置70还包括:反向处理模块770,被配置为获取用于生成当前文本的输入文本;将输入文本进行反向处理,以生成反向输入文本;以及,生成输入文本对应的正向序列、以及反向输入文本对应的反向序列,作为用于生成当前文本的参考序列。

[0122] 在一些实施例中,隐藏状态序列确定模块710进一步被配置为将用于生成当前文本的参考序列输入到文本生成模型的编码器中,获得正向隐藏状态序列和反向隐藏状态序列;以及,将正向隐藏状态序列和反向隐藏状态序列中位于相同位置的隐藏状态进行拼接,获得隐藏状态序列。

[0123] 在一些实施例中,关键词确定模块740进一步被配置为将隐藏状态序列中的最后一个隐藏状态和预先获取的关键词矩阵的乘积与预设的偏置向量之和输入到激活函数中,获得词库中每个词对应的概率值,其中,关键词矩阵包括词库中每个词对应的预设的向量;将概率值最大的预设数量个词确定为用于生成下一个文本的关键词。

[0124] 在一些实施例中,文本生成模块730进一步被配置为获取待生成的当前文本的第*i*个位置对应的、预设数量个不同的候选序列,其中,每个候选序列中包括待生成的当前文本的前*i*-1个位置中每个位置所对应的词语,每个词语具有在相应的位置上出现的概率值,其中,*i*为大于1的整数;根据待生成的当前文本的第*i*个位置对应的词语概率分布,确定词库中的词语在第*i*个位置出现的概率值;将词库中的词语与第*i*个位置对应的所有候选序列进行组合,并根据每种组合中的词语的概率值,确定每个组合的备选值;采用备选值最大的预设数量个组合更新候选序列,获得待生成的当前文本的第*i*+1个位置对应的、预设数量个不同的候选序列,以便采用待生成的当前文本的最后一个位置对应的、备选值最大的候选序列生成当前文本。

[0125] 在一些实施例中,文本生成模块730进一步被配置为将组合中的每个词对应的概率值的对数值的总和与组合中词的数量之比,确定为组合的备选值。

[0126] 在一些实施例中,文本生成装置70还包括:用户输入模块780,被配置为获取用户输入的文本、音频、图像中的至少一种,作为用户输入信息;将用户输入信息映射为包括一个或多个向量的序列,作为用于生成首个文本的参考序列。

[0127] 在一些实施例中,用户输入模块780进一步被配置为在用户输入信息为文本的情

况下,在将用户输入信息映射为包括多个向量的序列之前,文本生成方法还包括:在用户输入信息中的词语数量大于预设值的情况下,将用户输入信息截取为预设的长度;在用户输入信息中的词语数量小于预设值的情况下,将用户输入信息填充为预设的长度。

[0128] 在一些实施例中,文本生成模型的编码器和解码器为循环神经网络或者卷积神经网络。

[0129] 在一些实施例中,文本生成装置70还包括:训练模块790,被配置为获取训练文本,其中,训练文本中的关键词被预先进行标记;将用于训练的参考序列输入到文本生成模型的编码器中,获得基于编码器的输出而确定的、训练文本的隐藏状态序列;将训练文本的隐藏状态序列输入到文本生成模型的解码器中,获得训练文本中的每个位置所对应的词语概率分布;基于训练文本中的每个位置所对应的词语概率分布生成预测文本;根据训练文本的隐藏状态序列和关键词矩阵,确定下一个文本对应的预测关键词;根据预测文本与训练文本的差距、以及预测关键词与训练文本的下一个文本的关键词的差距,对文本生成模型和关键词矩阵的参数进行调整,直到预测文本与训练文本的差距、以及预测关键词与训练文本的下一个文本的关键词的差距小于预设值。

[0130] 图8为根据本发明另一些实施例的文本生成装置的结构示意图。如图8所示,该实施例的文本生成装置80包括:存储器810以及耦接至该存储器810的处理器820,处理器820被配置为基于存储在存储器810中的指令,执行前述任意一个实施例中的文本生成方法。

[0131] 其中,存储器810例如可以包括系统存储器、固定非易失性存储介质等。系统存储器例如存储有操作系统、应用程序、引导装载程序(Boot Loader)以及其他程序等。

[0132] 图9为根据本发明又一些实施例的文本生成装置的结构示意图。如图9所示,该实施例的文本生成装置90包括:存储器910以及处理器920,还可以包括输入输出接口930、网络接口940、存储接口950等。这些接口930,940,950以及存储器910和处理器920之间例如可以通过总线960连接。其中,输入输出接口930为显示器、鼠标、键盘、触摸屏等输入输出设备提供连接接口。网络接口940为各种联网设备提供连接接口。存储接口950为SD卡、U盘等外置存储设备提供连接接口。

[0133] 本发明的实施例还提供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,该程序被处理器执行时实现前述任意一种文本生成方法。

[0134] 本领域内的技术人员应当明白,本发明的实施例可提供为方法、系统、或计算机程序产品。因此,本发明可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且,本发明可采用在一个或多个其中包含有计算机可用程序代码的计算机可用非瞬时性存储介质(包括但不限于磁盘存储器、CD-ROM、光学存储器等)上实施的计算机程序产品的形式。

[0135] 本发明是参照根据本发明实施例的方法、设备(系统)、和计算机程序产品的流程图和/或方框图来描述的。应理解为可由计算机程序指令实现流程图和/或方框图中的每一流程和/或方框、以及流程图和/或方框图中的流程和/或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器以产生一个机器,使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的装置。

[0136] 这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中,使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品,该指令装置实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能。

[0137] 这些计算机程序指令也可装载到计算机或其他可编程数据处理设备上,使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理,从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的步骤。

[0138] 以上所述仅为本发明的较佳实施例,并不用以限制本发明,凡在本发明的精神和原则之内,所作的任何修改、等同替换、改进等,均应包含在本发明的保护范围之内。

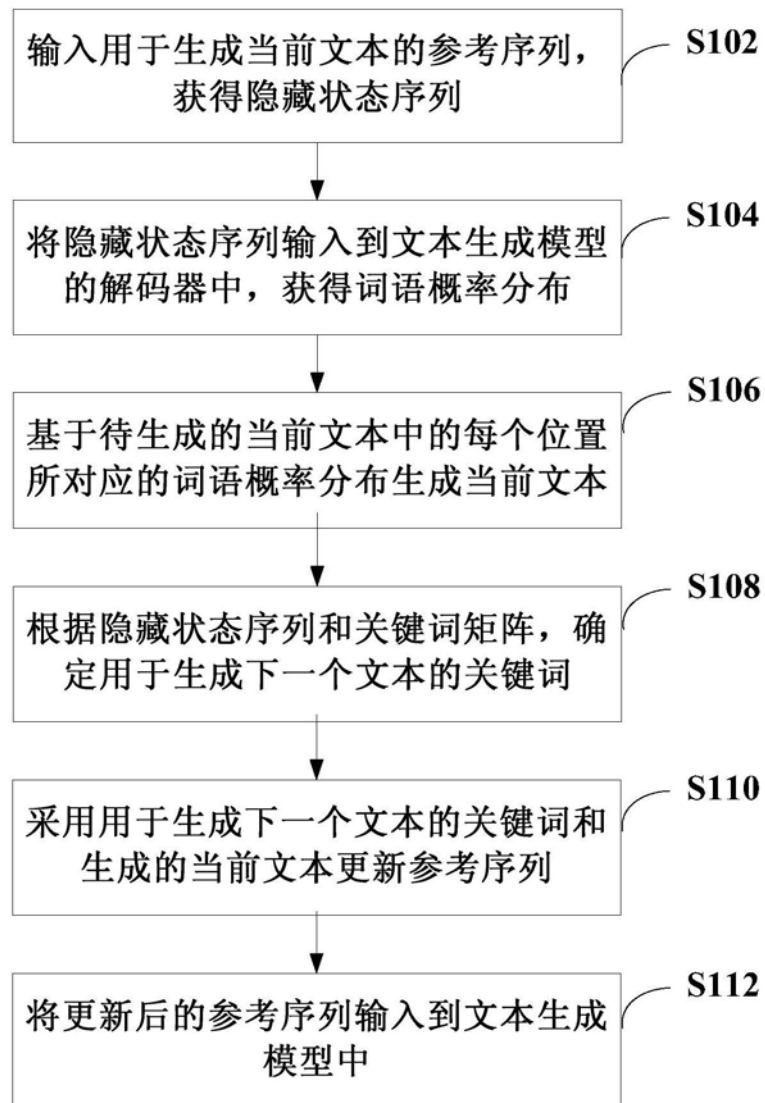


图1

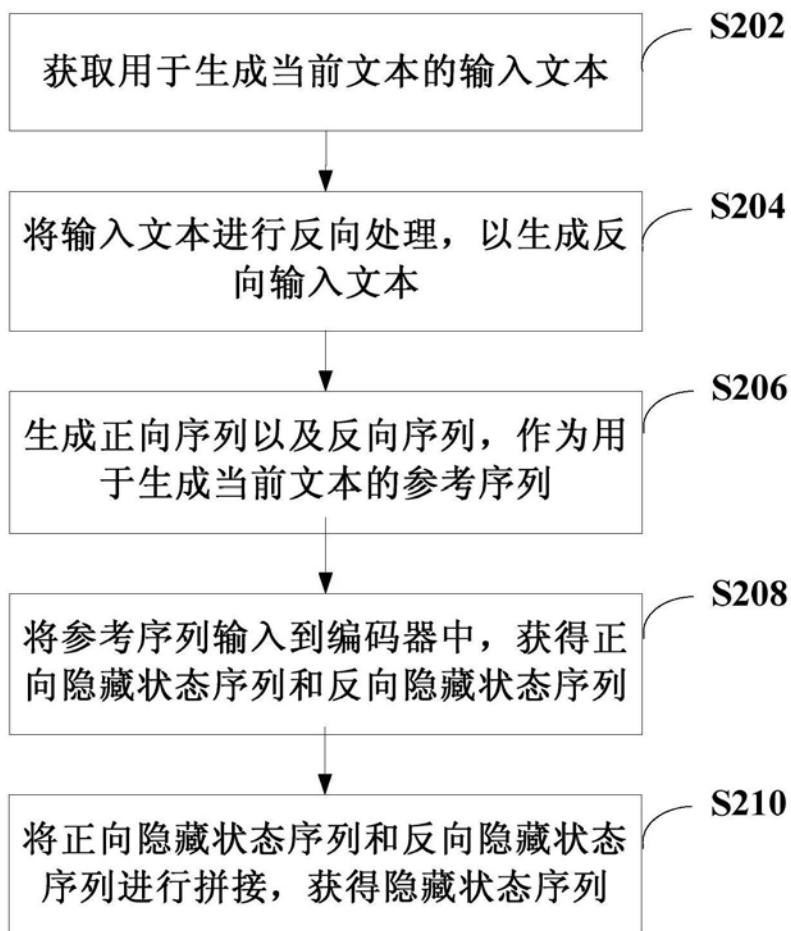


图2

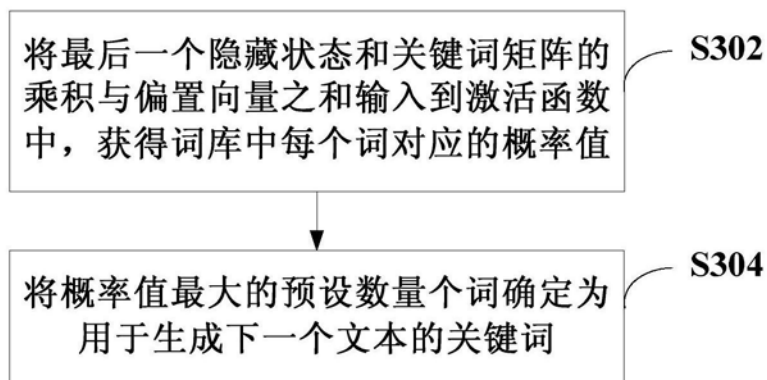


图3

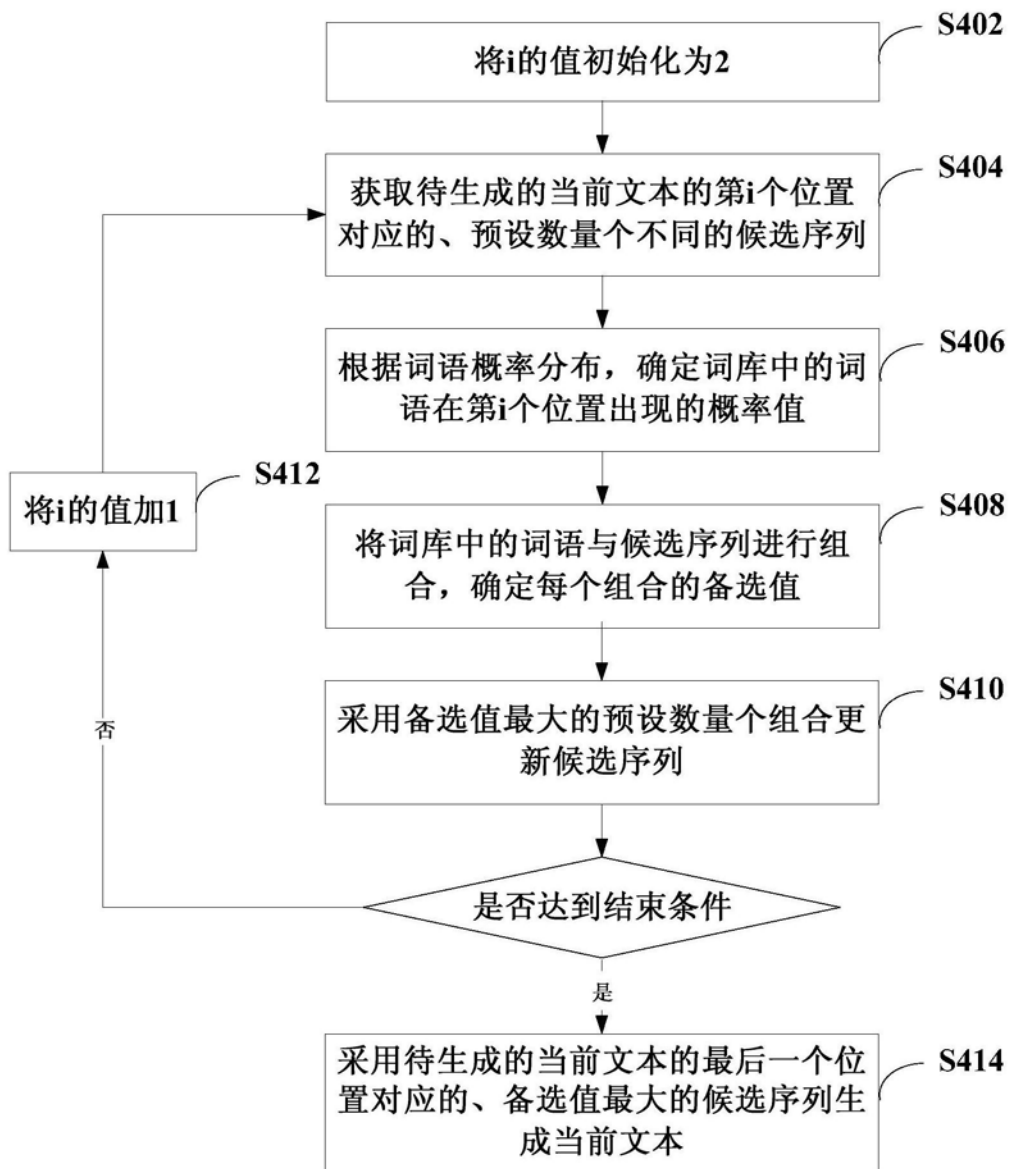


图4



图5

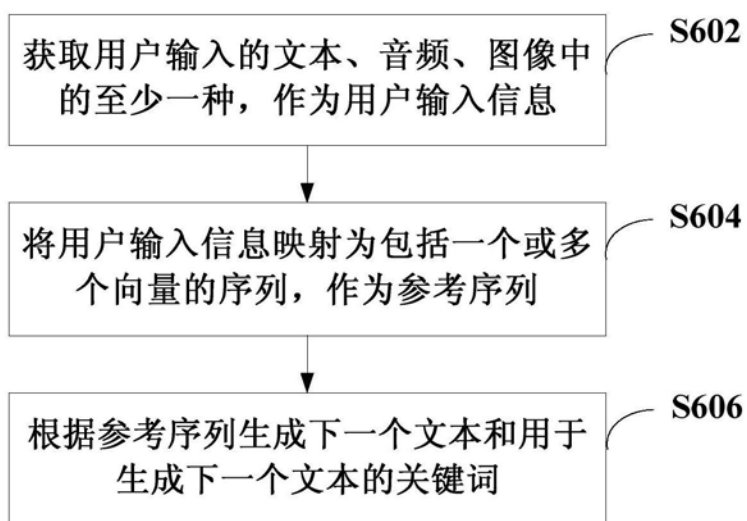


图6

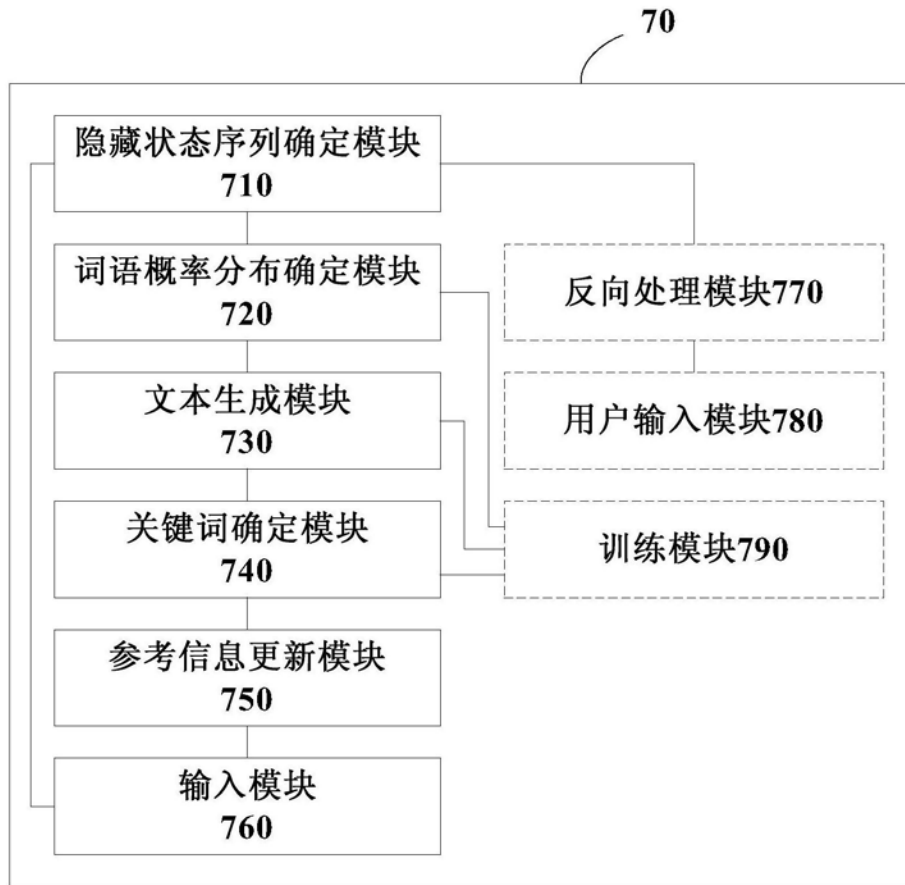


图7

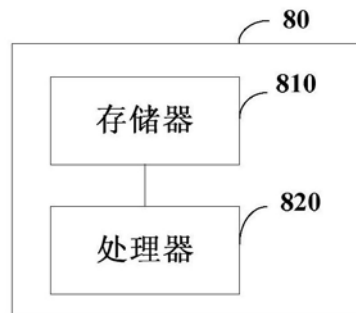


图8

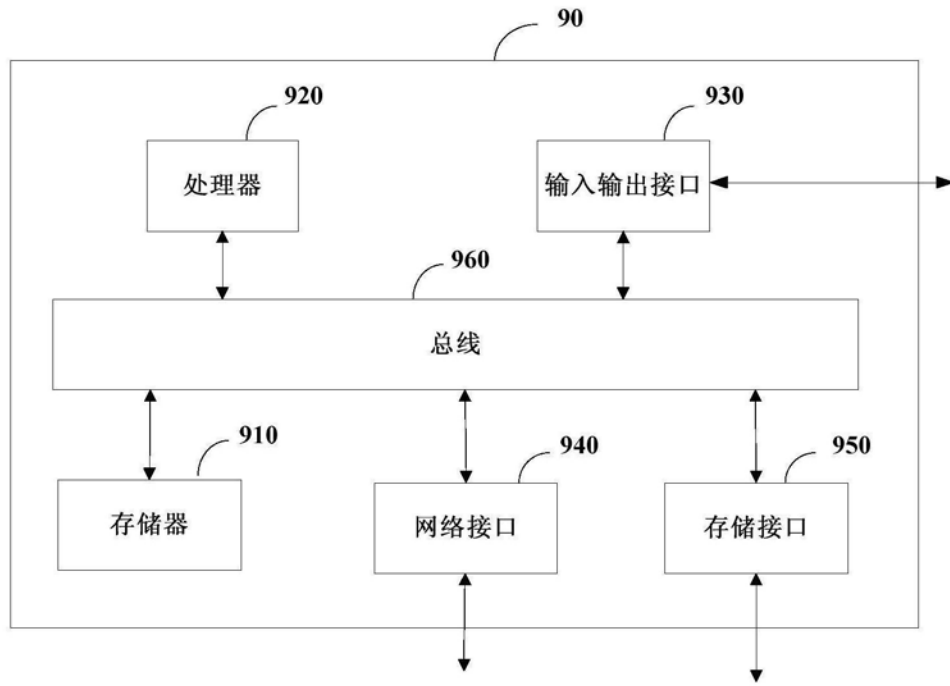


图9