(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国 际 局

(43) 国际公布日
2020 年 6 月 4 日 (04.06.2020)

WIPO | PCT

(10) 国际公布号
WO 2020/108400 A1

(51) 国际专利分类号:
G06F 17/00 (2019.01)

(21) 国际申请号: PCT/CN2019/120264

(22) 国际申请日: 2019 年 11 月 22 日 (22.11.2019)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201811448899.8 2018年11月28日 (28.11.2018) CN

(71) 申请人: 腾 讯 科 技 (深 圳) 有 限 公 司 (**TENCENT TECHNOLOGY (SHENZHEN) COMPANY LIMITED**) [CN/CN]; 中国广东省深圳市南山区高新区科技中一路腾讯大厦35层, Guangdong 518057 (CN)。

(72) 发明人: 涂兆鹏 (**TU, Zhaopeng**); 中国广东省深圳市南山区高新区科技中一路腾讯大厦35层, Guangdong 518057 (CN)。 耿昕伟 (**GENG, Xinwei**); 中国广东省深圳市南山区高新区科技中一路腾讯大厦35层, Guangdong 518057 (CN)。

王龙跃 (**WANG, Longyue**); 中国广东省深圳市南山区高新区科技中一路腾讯大厦35层, Guangdong 518057 (CN)。 王 星 (**WANG, Xing**); 中国广东省深圳市南山区高新区科技中一路腾讯大厦35层, Guangdong 518057 (CN)。

(74) 代理人: 北京三高永信知识产权代理有限责任公司 (**BEIJING SAN GAO YONG XIN INTELLECTUAL PROPERTY AGENCY CO., LTD.**); 中国北京市海淀区学院路蓟门里和景园A座1单元102室, Beijing 100088 (CN)。

(81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,

(54) **Title:** TEXT TRANSLATION METHOD AND DEVICE, AND STORAGE MEDIUM

(54) 发明名称: 一种文本翻译的方法、装置及存储介质

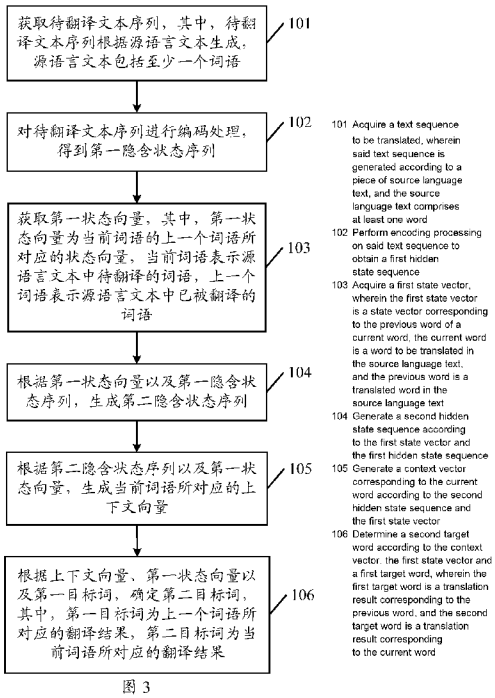

图 3

(57) **Abstract:** Disclosed is a text translation method. The method comprises: acquiring a text sequence to be translated; performing encoding processing on said text sequence to obtain a first hidden state sequence; acquiring a first state vector; generating a second hidden state sequence according to the first state vector and the first hidden state sequence; generating a context vector corresponding to a current word according to the second hidden state sequence and the first state vector; and determining a second target word according to the context vector, the first state vector and a first target word, wherein the first target word is a translation result corresponding to the previous word, and the second target word is a translation result corresponding to the current word. Further disclosed is a text translation device. In the embodiments of the present invention, in the process of encoding a text sequence to be translated which corresponds to a piece of source language text, a context vector obtained by means of decoding is introduced, such that the representation of said text sequence is enhanced, and the translation quality is thus improved.

WO 2020/108400 A1

[见续页]

SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(84) **指定国**(除另有指明，要求每一种可提供的地区保护)：ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告(条约第21条(3))。

(57) 摘要：本发明公开了一种文本翻译的方法，包括：获取待翻译文本序列；对待翻译文本序列进行编码处理，得到第一隐含状态序列；获取第一状态向量；根据第一状态向量以及第一隐含状态序列，生成第二隐含状态序列；根据第二隐含状态序列以及第一状态向量，生成当前词语所对应的上下文向量；根据上下文向量、第一状态向量以及第一目标词，确定第二目标词，其中，第一目标词为上一个词语所对应的翻译结果，第二目标词为当前词语所对应的翻译结果。本发明还公开了一种文本翻译装置。本发明实施例在对源语言文本所对应的待翻译文本序列进行编码的过程中，引入了解码得到的上下文向量，由此增强对待翻译文本序列的表示，从而提升翻译质量。

一种文本翻译的方法、装置及存储介质

本申请实施例要求于2018年11月28日提交,申请号为201811448899.8、发明名称为“一种文本翻译的方法以及相关装置”的中国专利申请的优先权，其全部内容通过引用结合在本申请实施例中。

技术领域

本发明涉及人工智能领域，尤其涉及一种文本翻译的方法、装置及存储介质。

背景技术

近年来，编码器——解码器框架在文本处理任务中取得突出的成果，文本处理任务包括机器对话、机器问答以及机器翻译等。在机器翻译这个项目中，可以对不同语种进行翻译，比如输入序列是英文句子，那么输出可以是该英文句子的中文翻译结果。

目前，利用编码器——解码器框架作为翻译模型进行翻译的过程具体为，首先将源语言句子转化成向量表示，再将向量表示的序列输入至编码器，编码后得到中间向量，最后，由解码器对该中间向量进行解码，从而生成目标语言所对应的翻译结果。

然而，采用编码器——解码器框架进行翻译，虽然可以达到翻译的目的，但是翻译质量并不高，尤其对于长句而言，更容易出现翻译上的偏差，从而导致翻译效果较差。

发明内容

本发明实施例提供了一种文本翻译的方法、装置及存储介质，在对源语言文本所对应的待翻译文本序列进行编码的过程中，引入了解码得到的上下文向量，由此增强对待翻译文本序列的表示，加强对源语言文本的理解，从而提升翻译质量，尤其对于长句而言，翻译效果更佳。

有鉴于此，本发明实施例的第一方面提供了一种文本翻译的方法，包括:

获取待翻译文本序列，其中，所述待翻译文本序列根据源语言文本生成，

所述源语言文本包括至少一个词语;

对所述待翻译文本序列进行编码处理，得到第一隐含状态序列;

获取第一状态向量，其中，所述第一状态向量为当前词语的上一个词语所对应的状态向量，所述当前词语表示所述源语言文本中待翻译的词语，所述上一个词语表示所述源语言文本中已被翻译的词语;

根据所述第一状态向量以及所述第一隐含状态序列，生成第二隐含状态序列;

根据所述第二隐含状态序列以及所述第一状态向量，生成所述当前词语所对应的上下文向量;

根据所述上下文向量、所述第一状态向量以及第一目标词，确定第二目标词，其中，所述第一目标词为所述上一个词语所对应的翻译结果，所述第二目标词为所述当前词语所对应的翻译结果。获取待翻译文本序列，其中，所述待翻译文本序列包括至少一个向量，所述待翻译文本序列为根据源语言文本生成的;

本发明实施例的第二方面提供了一种文本翻译装置，包括:

获取模块，用于获取待翻译文本序列，其中，所述待翻译文本序列根据源语言文本生成，所述源语言文本包括至少一个词语;

编码模块，用于对所述获取模块获取的所述待翻译文本序列进行编码处理，得到第一隐含状态序列;

所述获取模块，还用于获取第一状态向量，其中，所述第一状态向量为当前词语的上一个词语所对应的状态向量，所述当前词语表示所述源语言文本中待翻译的词语，所述上一个词语表示所述源语言文本中已被翻译的词语;

生成模块，用于根据所述获取模块获取的所述第一状态向量以及所述第一隐含状态序列，生成第二隐含状态序列;

所述生成模块，还用于根据所述第二隐含状态序列以及所述第一状态向量，生成所述当前词语所对应的上下文向量;

确定模块，用于根据所述生成模块生成的所述上下文向量、所述第一状态向量以及第一目标词，确定第二目标词，其中，所述第一目标词为所述上一个词语所对应的翻译结果，所述第二目标词为所述当前词语所对应的翻译结果。

本发明实施例的第三方面提供了一种文本翻译装置，包括：存储器、收发器、处理器以及总线系统;

其中，所述存储器用于存储程序；

所述处理器用于执行所述存储器中的程序，以实现如上述方面所述的方法。

本发明实施例的第四方面提供了一种计算机可读存储介质，所述计算机可读存储介质中存储有指令，当其在计算机上运行时，使得计算机执行上述各方面所述的方法。

从以上技术方案可以看出，本发明实施例具有以下优点：

本发明实施例中，提供了一种文本翻译的方法，首先获取待翻译文本序列，待翻译文本序列包括至少一个向量，待翻译文本序列为根据源语言文本生成的，对待翻译文本序列进行编码处理，得到第一隐含状态序列，然后根据第一状态向量以及第一隐含状态序列，生成第二隐含状态序列，第一状态向量为上一个词语所对应的状态向量，再根据第二隐含状态序列以及第一状态向量，生成当前词语所对应的上下文向量，最后根据上下文向量、第一状态向量以及第一目标词，确定第二目标词，第一目标词为上一个词语所对应的翻译结果，第二目标词为当前词语所对应的翻译结果。通过上述方式，在对源语言文本所对应的待翻译文本序列进行编码的过程中，引入了解码得到的上下文向量，由此增强对待翻译文本序列的表示，加强对源语言文本的理解，从而提升翻译质量，尤其对于长句而言，翻译效果更佳。

附图说明

图 1 为本发明实施例中文本翻译系统的一个架构示意图；

图 2 为本发明实施例中解码器-精炼器-解码器的一个框架示意图；

图 3 为本发明实施例中文本翻译的方法一个实施例示意图；

图 4 为本发明实施例中解码器-精炼器-解码器的一个网络结构示意图；

图 5 为本发明应用场景中不同句子长度翻译质量的一个对比示意图；

图 6 为本发明实施例中文本翻译装置一个实施例示意图；

图 7 为本发明实施例中文本翻译装置另一个实施例示意图；

图 8 为本发明实施例中文本翻译装置一个结构示意图；

图 9 为本发明实施例中文本翻译装置另一个结构示意图。

具体实施方式

本发明实施例提供了一种文本翻译的方法以及相关装置，在对源语言文

本所对应的待翻译文本序列进行编码的过程中，引入了解码得到的上下文向量，由此增强对待翻译文本序列的表示，加强对源语言文本的理解，从而提升翻译质量，尤其对于长句而言，翻译效果更佳。

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”、“第三”、“第四”等（如果存在）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例如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。

应理解，本发明实施例所提供的文本翻译方法可以应用于问答系统、对话系统、自然语言推理以及文本摘要等场景。为了提升翻译的质量，本发明提出在对源语言句子编码过程融入目标端（即目标语言）的上下文向量，从而实现对源语言表示的改进，使其尽可能多的包含与当前上下文向量相关的信息，并去除与当前目标端上下文向量无关的内容。可选的，首先在源语言表示中引入目标端的上下文向量，从而实现对源语言表示的浅层的重新理解。其次，以此浅层理解的结果为输入，引入另外一个编码器对其进行重新编码，从而得到对源端（即源语言）表示的深层理解。可选地，为了降低编码和解码的时间复杂度，我们还引入选择策略动态的决定是否需要对当前的源语言表示进行重新的编码。

以文本翻译场景为例进行介绍，请参阅图 1，图 1 为本发明实施例中文本翻译系统的一个架构示意图，如图所示，将本发明提供的文本翻译模型部署于服务器上，在终端设备向服务器发送源语言文本之后，由服务器对源语言文本进行编码处理、提炼处理以及解码处理，从而生成翻译结果，服务器再将翻译结果发送至终端设备，由终端设备展示翻译结果。可选地，在实际应用中，文本翻译模型也可以直接部署在终端设备上，即由终端设备在离线的状态下也可以采用该文本翻译模型对源语言文本进行翻译，并生成翻译结果，最后仍然由终端设备展示翻译结果。可以理解的是，终端设备包含但不仅限于平板电脑、手机、笔记本电脑、个人电脑（PersonalComputer，PC）以及掌

上电脑。

为了在编码过程中引入目标端的上下文向量，本发明提出编码器-精炼器-解码器（encoder-refiner-decoder）框架（编码器-精炼器-解码器即构成了文本翻译模型），其网络结构如图 2 所示。为了便于理解，请参阅图 2，图 2 为本发明实施例中编码器-精炼器-解码器的一个框架示意图，如图所示，首先，我们使用标准的编码器将源语言句子 x 编码成连续的序列，然后在解码的过程中，我们通过引入目标端的上下文向量实现对源端连续序列的提炼，使文本序列所包含的内容与当前的解码内容更加相关，最终解码获得一个完整的目标端文本序列，从而得到翻译结果 y。

在 encoder-refiner-decoder 框架中，其中，refiner 是新增的网络层，也是本发明提出的核心模块。refiner 的主要功能是通过考虑目标端的上下文向量，从而实现对输入序列表示的重新理解，提炼其中对于当前 decoder 比较重要的信息，而去除不相关的内容。为了达到这个目的，本发明中的 refiner 包含以下重要功能：1、以解码器的当前状态作为上下文向量，并将上下文向量与源端的连续序列为输入，实现浅层理解。2、以浅层理解为输入，通过重新编码实现浅层理解的深度理解。3、为了降低模型的时间复杂度，通过强化学习条件化的理解策略来决定是否有必要对源端表示进行精炼。

下面将对本发明实施例中文本翻译的方法进行介绍，请参阅图 3，本发明一个示例性实施例提供的文本翻译的方法包括如下步骤：

101、获取待翻译文本序列，其中，待翻译文本序列根据源语言文本生成，源语言文本包括至少一个词语；

本实施例中，首先，文本翻译装置的 encoder 获取源语言文本，其中，源语言可以是中文、英文、法文、日文以及德文等不同国家的语言，源语言文本包括至少一个词语，通常情况下包括多个词语。示意性的，当源语言为英文时，源语言文本可以表示为 “I have a pen”，该源语言文本即包括 4 个词语。

与源语言相对的就是翻译后得到的目标语言，目标语言也可以是中文、英文、法文、日文以及德文等不同国家的语言。在 encoder 获取源语言文本之后，需要对源语言文本进行词嵌入（embedding）处理，即 encoder 通过词嵌入用 N 维向量表示每个单词。相似单词具有相似词嵌入，在 N 维嵌入空间中距离相近。词嵌入基于在某种语言任务上训练的模型得到。假设使用 300 维的词嵌入。将输入句子表示为词嵌入序列后，可以传入编码器的循环层。

其中，源语言文本可以表示为 $x = x_1,...,x_j,...x_J$，相应的，目标语言文本可以表示为 $y = y_1,...,y_i,...y_I$。

102、对待翻译文本序列进行编码处理，得到第一隐含状态序列；

本实施例中，在 encoder 获取待翻译文本序列之后，开始对该待翻译文本序列进行编码处理，从而得到第一隐含状态序列。其中，可以是采用循环神经网络（Recurrent Neural Network，RNN）对待翻译文本序列进行编码，从而得到第一隐含状态序列。RNN 是一种将序列建模转变为时序建模的网络模型，它将状态在自身网络中循环传递。

在处理序列的每一步中，RNN 的隐藏状态传给接收序列的下一项，作为 RNN 的下一次迭代，并在迭代的同时为批次中的每个样本输出一个编码向量。序列处理的每一步都输出一个“矩阵”，并与 RNN 反向处理序列过程中输出的矩阵相连接。

可以理解的是，在实际应用中，还可采用其他类型的神经网络进行编码，比如，长短时记忆网络（Long Short-Term Memory，LSTM）、门控循环单元（Gated Recurrent Unit，GRU）、卷积神经网络（Convolutional Neural Network，CNN）或者自关注神经网络（Self-Attention Network，SAN），此处仅为一个示意，不应理解为对本发明的限定。

103、获取第一状态向量，其中，第一状态向量为当前词语的上一个词语所对应的状态向量，当前词语表示源语言文本中待翻译的词语，上一个词语表示源语言文本中已被翻译的词语；

本实施例中，文本翻译装置获取第一状态向量，其中，第一状态向量是上一个词语所对应的状态向量，假设源语言文本是“many airports were forced close”，当前翻译到的词语是“airports”，那么当前词语即为“airports”。“airports”的前一个已经翻译过的词语为“many”，那么“many”即为当前词语的上一个词语。

104、根据第一状态向量以及第一隐含状态序列，生成第二隐含状态序列；

本实施例中，文本翻译装置的 refiner 利用上一个词语所对应的状态向量以及第一隐含状态序列，对该第一隐含状态序列进行提炼，从而得到第二隐含状态序列。可选地，可以利用上一个词语“many”的状态向量（即第一状态向量）和第一隐含状态序列作为 refiner 的输入，由此相当于对第一隐含状态序列进行提炼，生成第二隐含状态序列、

105、根据第二隐含状态序列以及第一状态向量，生成当前词语所对应的上下文向量；

可选的，文本翻译装置的 refiner 在得到第二隐含状态序列之后，使用注意力（attention）模型，将第二隐含状态序列与第一状态向量输入至注意力模型，并输出相应的上下文向量。相应的，decoder 的输入为 refiner 输出的上下文向量，以及循环单元前一步的预测单词（即第一目标词）。

106、根据上下文向量、第一状态向量以及第一目标词，确定第二目标词，其中，第一目标词为上一个词语所对应的翻译结果，第二目标词为当前词语所对应的翻译结果。

本实施例中，文本翻译装置的 decoder 根据上下文向量、第一状态向量以及第一目标词，解码得到第二目标词。其中，第一目标词为上一个词语所对应的翻译结果，第二目标词为当前词语所对应的翻译结果。假设源语言文本是“many airports were forced close”，被翻译成中文“许多机场被迫关闭”。假设当前需要翻译“airports”这个词，那么“机场”即为第二目标词。第一目标词为“许多”。

本发明实施例中，提供了一种文本翻译的方法，首先获取待翻译文本序列，待翻译文本序列包括至少一个向量，待翻译文本序列为根据源语言文本生成的，对待翻译文本序列进行编码处理，得到第一隐含状态序列，然后根据第一状态向量以及第一隐含状态序列，生成第二隐含状态序列，第一状态向量为上一个词语所对应的状态向量，再根据第二隐含状态序列以及第一状态向量，生成当前词语所对应的上下文向量，最后根据上下文向量、第一状态向量以及第一目标词，确定第二目标词，第一目标词为上一个词语所对应的翻译结果，第二目标词为当前词语所对应的翻译结果。通过上述方式，在对源语言文本所对应的待翻译文本序列进行编码的过程中，引入了解码得到的上下文向量，由此增强对待翻译文本序列的表示，加强对源语言文本的理解，从而提升翻译质量，尤其对于长句而言，翻译效果更佳。

可选地，在上述图 3 对应的实施例的基础上，本发明实施例提供的文本翻译的方法第一个可选实施例中，根据第一状态向量以及第一隐含状态序列，生成第二隐含状态序列，可以包括：

根据目标隐含状态向量以及第一状态向量计算门函数，其中，目标隐含状态向量属于第一隐含状态序列中的一个隐含状态向量；

根据门函数以及目标隐含状态向量计算目标浅层理解向量;

根据目标浅层理解向量生成浅层理解序列，其中，浅层理解序列与第一隐含状态序列具有对应关系;

对浅层理解序列进行编码处理，得到第二隐含状态序列。

本实施例中，将介绍文本翻译装置如何生成第二隐含状态序列。为了便于介绍，请参阅图 4，图 4 为本发明实施例中编码器-精炼器-解码器的一个网络结构示意图，如图所示，在 decoder 解码的每一步中，使用上一步解码器状态 S_{i-1} 以及源端的第一隐含状态序列 h 为输入，最终获取得到提炼后的序列 h^i。

在一种可能的实现方式中，根据目标隐含状态向量 h_j（属于第一隐含状态序列 h）以及第一状态向量 S_{i-1} 计算门函数 z_j^i（Gating Function），然后根据门函数 z_j^i 和目标隐含状态向量 h_j 计算得到目标浅层理解向量 $\overline{h}_j^i$。经过一系列的计算，得出各个浅层理解向量，从而生成浅层理解序列 $\overline{h}^i$（包含 $\overline{h}_1^i, \overline{h}_2^i, \overline{h}_3^i, ..., \overline{h}_j^i$）。浅层理解序列与第一隐含状态序列具有对应关系，即第一隐含状态序列中的 h_1 与浅层理解序列中的 $\overline{h}_1^i$ 对应，第一隐含状态序列中的 h_2 与浅层理解序列中的 $\overline{h}_2^i$ 对应，以此类推。最后，refiner 对浅层理解序列进行编码处理，得到第二隐含状态序列 $\hat{h}^i$（包含 $\hat{h}_1^i, \hat{h}_2^i, \hat{h}_3^i, ..., \hat{h}_j^i$）。

其次，本发明实施例中，介绍了一种根据第一状态向量以及第一隐含状态序列，生成第二隐含状态序列的方法，首先根据目标隐含状态向量以及第一状态向量计算门函数，然后根据门函数以及目标隐含状态向量计算目标浅层理解向量，再根据目标浅层理解向量生成浅层理解序列，其中，浅层理解序列与第一隐含状态序列具有对应关系，最后对浅层理解序列进行编码处理，得到第二隐含状态序列。通过上述方式，引入门函数来控制源端编码中信息的传递，从而实现源端信息表示的动态化，由此提升模型的识别能力。

可选地，在上述图 3 对应的第一个实施例的基础上，本发明实施例提供的文本翻译的方法第二个可选实施例中，根据目标隐含状态向量以及第一状态向量计算门函数，可以包括:

根据目标隐含状态向量、第一状态向量以及 S 形（sigmoid）函数，计算门函数。

在一种可能的实施方式中，可以采用如下方式计算门函数:

$$z_j^i = \sigma\left(W_z h_j + U_z s_{i-1} + b_z\right);$$

其中，z_j^i表示门函数，$\sigma(\cdot)$表示 sigmoid 函数，W_z表示第一网络参数，U_z表示第二网络参数，b_z表示第三网络参数，h_j表示目标隐含状态向量，s_{i-1}表示第一状态向量，可以理解的是，门函数可以用于控制信息流动程度，sigmoid 函数作为非线性函数，其取值范围在 0 到 1，当然，在其他可能的实施方式中，也可以采用其他非线性函数计算门函数，本实施例对此不做限定。

根据门函数以及目标隐含状态向量计算目标浅层理解向量，可以包括:

将目标隐含状态向量与目标隐含向量进行元素级相乘，得到目标浅层理解向量。

在一种可能的实施方式中，可以采用如下方式计算目标浅层理解向量:

$$\overline{h}_j^i = z_j^i \odot h_j;$$

其中，$\overline{h}_j^i$表示目标浅层理解向量，$\odot$表示元素级相乘。

本实施例中，介绍了一种计算门函数以及计算目标浅层理解向量的方式。首先，文本翻译装置的 refiner 采用 sigmoid 函数计算门函数；接下来，利用门函数以及目标隐含状态向量计算目标浅层理解向量，这里采用元素级相乘的方式计算目标浅层理解向量。假设一组数据为[a1,a2,a3]，另一组数据为[b1,b2,b3]，元素级相乘就是得到这样的一组数据[a1b1,a2b2,a3b3]，也即是向量之间相乘得到结果。

再次，本发明实施例中，提供了一种计算门函数以及计算目标浅层理解向量的具体方式。通过上述方式，一方面为目标浅层理解向量的计算提供具体的实现依据，从而提升方案的可行性。另一方面，在实际应用中，能够更准确地生成目标浅层理解向量，从而提升方案的实用性。

可选地，在上述图 3 对应的第一个实施例的基础上，本发明实施例提供的文本翻译的方法第三个可选实施例中，对浅层理解序列进行编码处理，得到第二隐含状态序列，可以包括:

采用如下方式计算第二隐含状态序列:

$$\hat{h}^i = encoder_{re}\left(\overline{h}_1^i,...,\overline{h}_j^i,...,\overline{h}_J^i\right);$$

其中，$\hat{h}^i$表示第二隐含状态序列，$encoder_{re}(\cdot)$表示第二次编码处理，$\overline{h}_1^i$表示第一个浅层理解向量，$\overline{h}_j^i$表示目标浅层理解向量，$\overline{h}_J^i$表示第 J 个浅层理解向量。

本实施例中，在文本翻译装置完成浅层理解，即得到浅层理解序列之后，还可以额外引入一个 encoder 对浅层理解序列进行深层理解。输入的浅层理解序列 $\overline{h}_1^i,...,\overline{h}_j^i,...,\overline{h}_J^i$，使用另外的编码器 $encoder_{re}$ 来对其进行重新的编码，获得深层理解的表示，即第二隐含状态序列 $\hat{h}^i$。

其中，$encoder_{re}$ 和 $encoder$（对待翻译文本序列进行编码时使用的编码器）使用不同的参数集合。

再次，本发明实施例中，文本翻译装置还需要对浅层理解序列进行编码处理，得到第二隐含状态序列。通过上述方式，实现对浅层理解序列的深度理解。并且引入额外的编码器对浅层理解序列进行重新编码，从而提升了方案的可操作性和可行性。

可选地，在上述图 3 以及图 3 对应的第一个至第三个实施例中任一项的基础上，本发明实施例提供的文本翻译的方法第四个可选实施例中，根据上下文向量、第一状态向量以及第一目标词，确定第二目标词之后，还可以包括:

根据上下文向量、第二状态向量以及第二目标词所对应的词向量，计算目标输出概率，其中，第二状态向量为当前词语所对应的状态向量;

根据目标输出概率计算连续采样向量，其中，连续采样向量用于生成连续采样序列;

根据连续采样向量计算离散采样向量，其中，离散采样向量用于生成离散采样序列;

根据离散采样向量计算编码处理结果;

根据编码处理结果确定处理模式，其中，处理模式包括第一处理模式以及第二处理模式，第一处理模式表示采用已有的编码结果，第二处理模式表示采用第一隐含状态序列进行编码处理。

本实施例中，提出了一种条件化的提炼选择策略，如果提出的 encoder-refiner-decoder 框架在解码的每一步都对源语言文本都表示进行重新的理解，时间复杂度是非常高的。而实际上，并非每一步解码都需要对源语言文本编码进行重新提炼和理解，比如在同一个完整的语义单元（比如短语）中，由于其语义比较接近，只需要在语义单元开始的时候进行一次提炼即可，之后在整个语义单元翻译的过程中，均使用该提炼结果。因此，为了降低模型的时间复杂性，提出条件化机制来控制当前步是否需要重新提炼源端的编码。

在一种可能的实施方式中，文本翻译装置可以预测是否需要对下一个词语进行提炼，首先，根据上下文向量、第二状态向量以及第二目标词所对应的词向量，计算目标输出概率，然后根据目标输出概率计算连续采样向量，其中，连续采样向量用于生成连续采样序列，文本翻译装置再根据连续采样向量计算离散采样向量，其中，离散采样向量用于生成离散采样序列。最后，文本翻译装置根据离散采样向量计算编码处理结果，并根据编码处理结果确定处理模式，其中，处理模式包括第一处理模式以及第二处理模式，第一处理模式表示采用已有的编码结果，第二处理模式表示对第一隐含状态序列进行编码处理。

通过上述方式，可以不必对每个词语都进行提炼，从而降低了模型的复杂度，通过提出条件化的精炼策略，可以动态决定是否需要对当前的表示进行精炼，进而提升方案的灵活性和实用性。

可选地，在上述图3对应的第四个实施例的基础上，本发明实施例提供的文本翻译的方法第五个可选实施例中，根据上下文向量、第二状态向量以及第二目标词所对应的词向量，计算目标输出概率，可以包括:

根据上下文向量、第二状态向量以及第二目标词所对应的词向量，通过双曲正切函数确定策略函数的状态;

根据策略函数的状态，通过归一化指数函数计算目标输出概率。

在一种可能的实施方式中，可以采用如下方式计算目标输出概率:

$$\pi\left(a_i|m_i\right)=softmax\left(W_p m_i+b_p\right)$$

$$m_i=\tanh\left(W_p^{'}\left[s_i;E_{yi};c_i\right]+b_p^{'}\right);$$

其中，$\pi\left(a_i|m_i\right)$表示目标输出概率，a_i表示输出动作，m_i表示策略函数的状态，W_p表示第四网络参数，b_p表示第五网络参数，$W_p^{'}$表示第六网络参数，$b_p^{'}$表示第七网络参数，s_i表示第二状态向量，E_{yi}表示第二目标词所对应的词向量，c_i表示上下文向量，$softmax(\cdot)$表示归一化指数函数，$\tanh\left(\cdot\right)$表示双曲正切函数。

本实施例中，在引入条件化的提炼选择策略之后，可以设置两个输出动作，一个输出动作是对源语言文本的编码进行重新提炼，即生成$\hat{h}_{i+1}$，另一个输出动作是选择使用上一步的提炼的结果，即获取$\hat{h}_i$。

因为输出动作a_i预测的是下一时刻（下一个词语）是否进行提炼，所以使

用当前时刻（当前词语）的第二状态向量s_i、第二目标词所对应的词向量E_{yi}以及上下文向量c_i。策略函数m_i输出的是连续的概率值，而输出动作表示是否需要进行提炼的离散值，可以通过采样函数来实现从连续的概率值到离散的输出动作。这个采样函数根据概率进行采样，概率值越大，其采样对应的动作次数越多。

更进一步地，本发明实施例中，介绍了文本翻译装置根据上下文向量、第二状态向量以及第二目标词所对应的词向量，计算目标输出概率的方式。通过上述方式，一方面为目标输出概率的计算提供实现依据，从而提升方案的可行性。另一方面，在实际应用中，能够更准确地表示提炼的可能性，从而提升方案的可操作性。

可选地，在上述图3对应的第五个实施例的基础上，本发明实施例提供的文本翻译的方法第六个可选实施例中，根据目标输出概率计算连续采样向量，可以包括:

通过预设采样函数对所述目标输出概率进行采样，得到采样参数，预设采样函数为耿贝尔归一化指数函数（Gumbel-softmax），所述采样参数离散;

对采样参数进行连续化处理，得到连续采样向量。在一种可能的实施方式中，可以采用如下方式计算连续采样向量:

$$\bar{a}_i^k = \frac{\exp\left(\left(o_i^k + g_i^k\right)/\tau\right)}{\sum_{k'=1}^{K}\exp\left(\left(o_i^{k'} + g_i^{k'}\right)/\tau\right)};$$

$$g_i^k = -\log\left(-\log u_i^k\right);$$

$$u_i^k \sim Uniform(0,1);$$

其中，$\bar{a}_i^k$表示连续采样向量，$\exp(\cdot)$表示以自然常数e为底的指数函数，o_i^k表示未进行归一化的第一概率，g_i^k表示第一噪声，$o_i^{k'}$表示未进行归一化的第二概率，$g_i^{k'}$表示第二噪声，τ表示第一超参数，u_i^k表示采样参数，$\sim$表示采样操作，$Uniform(0,1)$表示在0至1的范围内均匀分布，K表示输出动作的总维度。

本实施例中，由于训练过程中需要对策略函数进行采样，以生成相应的动作序列。为了优化该网络，可以使用Gumbel-softmax将离散的

$\overline{a}_i^k = \overline{a}_i^1,...,\overline{a}_i^K$ 进行连续化处理。其中，Gumbel-softmax 可以认为是一种采样函数。采样函数可以对目标输出概率 $\pi(a_i|m_i)$ 进行采样，得到输出动作 a_i。利用离散变量 a_i 采样得到具有连续表示的分布向量 $\overline{a}_i^k$。之所以要对离散化的数值进行连续化表示，是因为模型需要计算梯度，而连续化的数值才具有梯度。通过 Gumbel-softmax 近似采样过程，使得最后结果是连续可导的。

可以理解的是，采样函数就如同掷硬币，投掷硬币的结果是正面或者反面，即表示采样函数的输出，如果出现正面的概率是 0.7，反面的概率是 0.3，概率越大表示出现相应结果的可能性越大。

更进一步地，本发明实施例中，介绍了文本翻译装置根据目标输出概率计算连续采样向量的具体方式。通过上述方式，可以将离散化的目标输出概率变得连续化，即生成连续采样向量。由于连续采样向量没有梯度，以此可以求导，也可以模拟采样过程，符合数据处理的规则，从而提升方案的可行性和可操作性。

可选地，在上述图 3 对应的第六个实施例的基础上，本发明实施例提供的文本翻译的方法第七个可选实施例中，根据连续采样向量计算离散采样向量，可以包括:

通过最大值自变量点集（argmax）函数对连续采样向量进行处理，生成离散采样向量，离散采样向量中包含的数值为 0 或 1。

在一种可能的实施方式中，可以采用如下方式计算离散采样向量:

$$\hat{a}_i^k = \begin{cases} 1 & k = \arg\max_{k'} \overline{a}_i^{k'} \\ 0 & otherwise \end{cases};$$

其中，$\hat{a}_i^k$ 表示离散采样向量，$\arg\max_{k'}$ 表示 $\overline{a}_i^{k'}$ 取最大值所对应的 k'，$otherwise$ 表示 $k \neq \arg\max_{k'} \overline{a}_i^{k'}$ 的情况。

本实施例中，文本翻译装置使用直通 Gumbel-softmax（Straight-ThroughGumbel-softmax，ST Gumbel-softmax）对连续采样向量进行离散化处理，输出离散化结果 $\hat{a} = \hat{a}_i^k,...,\hat{a}_i^K$，可选的，该计算方式如下所示:

$$\hat{a}_i^k = \begin{cases} 1 & k = \arg\max_{k'} \overline{a}_i^{k'} \\ 0 & otherwise \end{cases};$$

其中，$\hat{a}_i^{k=1}$ 表示第一离散采样向量，其中，k=1 的情况可以表示为“REUSE”，也就是使用上一步的提炼的结果 $\hat{h}^i$。k=2 的情况可以表示为

“REFINE”，也就是对源端编码进行重新的提炼。

更进一步地，本发明实施例中，介绍了文本翻译装置根据连续采样向量计算离散采样向量的方式。通过上述方式，一方面为离散采样向量的计算提供实现依据，从而提升方案的可行性。另一方面，在实际应用中，需要对连续化的连续采样向量进行离散处理，从而有利于表示编码处理结果，进而提升方案的实用性。

可选地，在上述图3对应的第七个实施例的基础上，本发明实施例提供的文本翻译的方法第八个可选实施例中，根据离散采样向量计算编码处理结果，可以包括:

根据第一离散采样向量、第二离散采样向量、第二隐含状态序列和第三隐含状态序列，计算编码处理结果，第一离散采样向量为第一处理模式对应的离散采样向量，第二离散采样向量为第二处理模式对应的离散采样向量，第三隐含状态序列是经过二次编码后生成的隐含状态序列。

在一种可能的实施方式中，可以采用如下方式计算编码处理结果:

$$\tilde{h}^{i+1}=\hat{a}_i^{k=1}\hat{h}^i+\hat{a}_i^{k=2}\hat{h}^{i+1};$$

其中，$\tilde{h}^{i+1}$表示编码处理结果，$\hat{a}_i^{k=1}$表示第一离散采样向量，第一离散采样向量与第一处理模式具有对应关系，$\hat{a}_i^{k=2}$表示第二离散采样向量，第二离散采样向量与第二处理模式具有对应关系，$\hat{h}^i$表示第二隐含状态序列，$\hat{h}^{i+1}$表示第三隐含状态序列。

更进一步地，本发明实施例中，提供了一种根据离散采样向量计算编码处理结果的具体方式。通过上述方式，一方面为编码处理结果的计算提供实现依据，从而提升方案的可行性。另一方面，在实际应用中，能够更高效地预测是否需要对下一个词语进行提炼，从而提升方案的实用性。

可选地，在上述图3对应的第四个实施例的基础上，本发明实施例提供的文本翻译的方法第九个可选实施例中，文本翻译模型（encoder-refiner-decoder）的训练过程可以包括:

获取样本源语言文本对应的预测目标语言文本和样本目标语言文本，预测目标语言文本由文本翻译模型翻译得到，样本目标语言文本用于在模型训练时作为预测目标语言文本的监督;

根据预测目标语言文本和样本目标语言文本计算模型预测损失;

根据模型预测损失和惩罚项计算模型总损失，惩罚项根据第二离散采样向

量和样本源语言文本的词语总数确定，第二离散采样向量是第二处理模式对应的离散采样向量；

根据模型总损失训练文本翻译模型。

在一种可能的实施方式中，可以采用如下方式计算模型预测损失：

$$L'(\hat{y},y)=L(\hat{y},y)+r(\hat{a});$$

$$r(\hat{a})=\alpha\left(\sum_{i=1}^{I}\hat{a}_i^{k=2}\right)/I;$$

其中，$L'(\hat{y},y)$表示模型总损失函数，$L(\hat{y},y)$表示模型预测损失函数，$\hat{y}$表示模型的预测值（即预测目标语言文本），y表示模型的真实值（即样本目标语言文本），$r(\hat{a})$表示惩罚项，$\hat{a}_i^{k=2}$表示第二离散采样向量，第二离散采样向量与第二处理模式具有对应关系，I表示源语言文本的词语总数，α表示第二超参数。

本实施例中，当在训练神经网络模型的时候，特别是一些线性分类器，往往需要定义一个损失函数，模型的训练就是要通过样本将损失函数的函数值最小化，当损失函数的函数值满足收敛条件时，则停止训练。为了对重炼动作数量进行限制，本发明在原本的模型预测损失函数基础上，增加了一个新的函数，即惩罚项。

基于模型损失函数$L(\hat{y},y)$，增加惩罚项$r(\hat{a})$后得到如下模型总损失函数：

$$L'(\hat{y},y)=L(\hat{y},y)+\alpha\left(\sum_{i=1}^{I}\hat{a}_i^{k=2}\right)/I;$$

采用上述模型总损失函数实现损失函数的最小化。

需要说明的是，损失函数的类型较多，可以是交叉熵损失函数，此处不做限定。

更进一步地，本发明实施例中，提供了一种限制提炼动作的方式，在原有的损失函数基础上，增加一个惩罚项。通过上述方式，为了训练一个更加更合适的选择策略，可以对重新精炼动作数量进行限制，从而鼓励模型重用之前的结果。以增加惩罚项实现上述约束，由此提升方案的可行性和可靠性。

为了便于理解，本发明可以应用于需要强化局部信息，并针对离散队列建

模的神经网络模型中。以机器翻译为例，在美国国家标准与技术研究院（National Institute of Standards and Technology，NIST）的中英机器翻译任务测试中，采用本发明所提供的方案可以显著提升翻译质量。请参阅表 1，表 1 为采用本发明所提供的方案在机器翻译系统上所取得的效果。

表 1

模型	参数数量	MT03	MT04	MT05	MT06	MT08	Ave.	Δ
基准	86.69M	37.26	40.50	36.67	37.10	28.54	36.01	-
浅层理解	92.69M	38.12	41.35	38.51	37.85	29.32	37.03	+1.02
深层理解	110.70M	39.23	42.72	39.90	39.24	30.68	38.35	+2.34
条件化	112.95M	38.61	41.99	38.89	38.04	29.36	37.38	+1.37

由此可见，双语评价研究（bilingual evaluation understudy，BLEU）一般提高超过 0.5 个点即是显著提高，该栏的Δ是指提高的绝对数值，参数数量的单位为百万（M），MT03、MT 04、MT 05、MT 06 和 MT 08 是 NIST 机器翻译测试集。

请参阅图 5，图 5 为本发明应用场景中不同句子长度翻译质量的一个对比示意图，如图所示，基线（baseline）表示采用标准神经网络机器翻译（Neural Machine Translation，NMT）模型训练的，浅层提炼（shallow refiner）表示添加上下文向量后的序列，深层提炼（deep refiner）表示对 shallow refiner 进行重新编码后生成的序列，条件化（contiditonal）表示提炼的条件。由此可见，本发明所提出方法在较长句子的翻译上表现出色，且经过两次编码后得到的翻译质量也更好。

下面对本发明中的文本翻译装置进行详细描述，请参阅图 6，图 6 为本发明实施例中文本翻译装置一个实施例示意图，文本翻译装置 20 包括:

获取模块 201，用于获取待翻译文本序列，其中，所述待翻译文本序列根据源语言文本生成，所述源语言文本包括至少一个词语;

编码模块 202，用于对所述获取模块 201 获取的所述待翻译文本序列进行编码处理，得到第一隐含状态序列;

所述获取模块 201，还用于获取第一状态向量，其中，所述第一状态向量为当前词语的上一个词语所对应的状态向量，所述当前词语表示所述源语言文本中待翻译的词语，所述上一个词语表示所述源语言文本中已被翻译的词语；

生成模块 203，用于根据所述获取模块 201 获取的所述第一状态向量以及所述第一隐含状态序列，生成第二隐含状态序列；

所述生成模块 203，还用于根据所述第二隐含状态序列以及所述第一状态向量，生成所述当前词语所对应的上下文向量；

确定模块 204，用于根据所述生成模块 203 生成的所述上下文向量、所述第一状态向量以及第一目标词，确定第二目标词，其中，所述第一目标词为所述上一个词语所对应的翻译结果，所述第二目标词为所述当前词语所对应的翻译结果。

在一个实施例中，所述生成模块 203，用于根据目标隐含状态向量以及所述第一状态向量计算门函数，其中，所述目标隐含状态向量属于所述第一隐含状态序列中的一个隐含状态向量；

根据所述门函数以及所述目标隐含状态向量计算目标浅层理解向量；

根据所述目标浅层理解向量生成浅层理解序列，其中，所述浅层理解序列与所述第一隐含状态序列具有对应关系；

对所述浅层理解序列进行编码处理，得到所述第二隐含状态序列。

在一个实施例中，所述生成模块 203，用于：

根据所述目标隐含状态向量、所述第一状态向量以及 sigmoid 函数，计算所述门函数；

将所述目标隐含状态向量与所述目标隐含向量进行元素级相乘，得到所述目标浅层理解向量。

在一个实施例中，在图 6 所示文本翻译的基础上，如图 7 所示，所述装置还包括：

计算模块 205，用于根据所述上下文向量、第二状态向量以及所述第二目标词所对应的词向量，计算目标输出概率，其中，所述第二状态向量为所述当前词语所对应的状态向量；

所述计算模块 205，还用于根据所述目标输出概率计算连续采样向量，其中，所述连续采样向量用于生成连续采样序列；

所述计算模块 205，还用于根据所述连续采样向量计算离散采样向量，其

中，所述离散采样向量用于生成离散采样序列；

所述计算模块 205，还用于根据所述离散采样向量计算编码处理结果；

所述确定模块 204，用于根据所述编码处理结果确定处理模式，其中，所述处理模式包括第一处理模式以及第二处理模式，所述第一处理模式表示采用已有的编码结果，所述第二处理模式表示对所述第一隐含状态序列进行编码处理。

在一个实施例中，所述计算模块 205，还用于：

根据所述上下文向量、所述第二状态向量以及所述第二目标词所对应的词向量，通过双曲正切函数确定策略函数的状态；

根据所述策略函数的状态，通过归一化指数函数计算所述目标输出概率。

在一个实施例中，所述计算模块 205，还用于：

通过预设采样函数对所述目标输出概率进行采样，得到采样参数，所述预设采样函数为 Gumbel-softmax，所述采样参数离散；

对所述采样参数进行连续化处理，得到所述连续采样向量。

在一个实施例中，所述计算模块 205，还用于：

通过 argmax 函数对所述连续采样向量进行处理，生成所述离散采样向量，所述离散采样向量中包含的数值为 0 或 1。

在一个实施例中，所述计算模块 205，还用于：

根据第一离散采样向量、第二离散采样向量、所述第二隐含状态序列和第三隐含状态序列，计算所述编码处理结果，所述第一离散采样向量为所述第一处理模式对应的离散采样向量，所述第二离散采样向量为所述第二处理模式对应的离散采样向量，所述第三隐含状态序列是经过二次编码后生成的隐含状态序列。

在一个实施例中，所述装置还包括训练模块，所述训练模块，用于：

获取样本源语言文本对应的预测目标语言文本和样本目标语言文本，所述预测目标语言文本由文本翻译模型翻译得到，所述样本目标语言文本用于在模型训练时作为所述预测目标语言文本的监督；

根据所述预测目标语言文本和所述样本目标语言文本计算模型预测损失；

根据所述模型预测损失和惩罚项计算模型总损失，所述惩罚项根据第二离散采样向量和所述样本源语言文本的词语总数确定，所述第二离散采样向量是所述第二处理模式对应的离散采样向量；

根据所述模型总损失训练所述文本翻译模型。

需要说明的是，上述文本翻译装置中各个功能模块的功能实现方式可以参考方法实施例，本实施例在此不再赘述。

图 8 是本发明实施例提供的一种服务器结构示意图，该服务器 300 可因配置或性能不同而产生比较大的差异，可以包括一个或一个以上中央处理器（central processing units，CPU）322（例如，一个或一个以上处理器）和存储器 332，一个或一个以上存储应用程序 342 或数据 344 的存储介质 330（例如一个或一个以上海量存储设备）。其中，存储器 332 和存储介质 330 可以是短暂存储或持久存储。存储在存储介质 330 的程序可以包括一个或一个以上模块（图示没标出），每个模块可以包括对服务器中的一系列指令操作。更进一步地，中央处理器 322 可以设置为与存储介质 330 通信，在服务器 300 上执行存储介质 330 中的一系列指令操作。

服务器 300 还可以包括一个或一个以上电源 326，一个或一个以上有线或无线网络接口 350，一个或一个以上输入输出接口 358，和/或，一个或一个以上操作系统 341，例如 Windows ServerTM，Mac OS XTM，UnixTM，LinuxTM，FreeBSDTM 等等。

上述实施例中由服务器所执行的步骤可以基于该图 8 所示的服务器结构。

在本发明实施例中，该服务器所包括的 CPU 322 用于实现上述方法实施例中所述的文本翻译方法。

本发明实施例还提供了另一种文本翻译装置，如图 9 所示，为了便于说明，仅示出了与本发明实施例相关的部分，具体技术细节未揭示的，请参照本发明实施例方法部分。该终端设备可以为包括手机、平板电脑、个人数字助理（personal digital assistant，PDA）、销售终端（point of sales，POS）、车载电脑等任意终端设备，以终端为手机为例：

图 9 示出的是与本发明实施例提供的终端设备相关的手机的部分结构的框图。参考图 9，手机包括：射频（radio frequency，RF）电路 410、存储器 420、输入单元 430、显示单元 440、传感器 450、音频电路 460、无线保真（wireless fidelity，WiFi）模块 470、处理器 480、以及电源 490 等部件。本领域技术人员可以理解，图 9 中示出的手机结构并不构成对手机的限定，可以包括比图示更多或更少的部件，或者组合某些部件，或者不同的部件布置。

下面结合图 9 对手机的各个构成部件进行介绍:

RF 电路 410 可用于收发信息或通话过程中，信号的接收和发送，特别地，将基站的下行信息接收后，给处理器 480 处理；另外，将设计上行的数据发送给基站。通常，RF 电路 410 包括但不限于天线、至少一个放大器、收发信机、耦合器、低噪声放大器（low noise amplifier，LNA）、双工器等。此外，RF 电路 410 还可以通过无线通信与网络和其他设备通信。上述无线通信可以使用任一通信标准或协议，包括但不限于全球移动通讯系统（global system of mobile communication，GSM）、通用分组无线服务（general packet radio service，GPRS）、码分多址（code division multiple access，CDMA）、宽带码分多址（wideband code division multiple access, WCDMA）、长期演进（long term evolution，LTE）、电子邮件、短消息服务（short messaging service，SMS）等。

存储器 420 可用于存储软件程序以及模块，处理器 480 通过运行存储在存储器 420 的软件程序以及模块，从而执行手机的各种功能应用以及数据处理。存储器 420 可主要包括存储程序区和存储数据区，其中，存储程序区可存储操作系统、至少一个功能所需的应用程序（比如声音播放功能、图像播放功能等）等；存储数据区可存储根据手机的使用所创建的数据（比如音频数据、电话本等）等。此外，存储器 420 可以包括高速随机存取存储器，还可以包括非易失性存储器，例如至少一个磁盘存储器件、闪存器件、或其他易失性固态存储器件。

输入单元 430 可用于接收输入的数字或字符信息，以及产生与手机的用户设置以及功能控制有关的键信号输入。具体地，输入单元 430 可包括触控面板 431 以及其他输入设备 432。触控面板 431，也称为触摸屏，可收集用户在其上或附近的触摸操作（比如用户使用手指、触笔等任何适合的物体或附件在触控面板 431 上或在触控面板 431 附近的操作），并根据预先设定的程式驱动相应的连接装置。可选的，触控面板 431 可包括触摸检测装置和触摸控制器两个部分。其中，触摸检测装置检测用户的触摸方位，并检测触摸操作带来的信号，将信号传送给触摸控制器；触摸控制器从触摸检测装置上接收触摸信息，并将它转换成触点坐标，再送给处理器 480，并能接收处理器 480 发来的命令并加以执行。此外，可以采用电阻式、电容式、红外线以及表面声波等多种类型实现触控面板 431。除了触控面板 431，输入单元 430 还可以包括其他输入设备 432。具体地，其他输入设备 432 可以包括但不限于物理键盘、功能键（比如

音量控制按键、开关按键等）、轨迹球、鼠标、操作杆等中的一种或多种。

显示单元440可用于显示由用户输入的信息或提供给用户的信息以及手机的各种菜单。显示单元440可包括显示面板441，可选的，可以采用液晶显示器（liquid crystal display，LCD）、有机发光二极管（organic light-emitting diode，OLED）等形式来配置显示面板441。进一步的，触控面板431可覆盖显示面板441，当触控面板431检测到在其上或附近的触摸操作后，传送给处理器480以确定触摸事件的类型，随后处理器480根据触摸事件的类型在显示面板441上提供相应的视觉输出。虽然在图9中，触控面板431与显示面板441是作为两个独立的部件来实现手机的输入和输入功能，但是在某些实施例中，可以将触控面板431与显示面板441集成而实现手机的输入和输出功能。

手机还可包括至少一种传感器450，比如光传感器、运动传感器以及其他传感器。具体地，光传感器可包括环境光传感器及接近传感器，其中，环境光传感器可根据环境光线的明暗来调节显示面板441的亮度，接近传感器可在手机移动到耳边时，关闭显示面板441和/或背光。作为运动传感器的一种，加速计传感器可检测各个方向上（一般为三轴）加速度的大小，静止时可检测出重力的大小及方向，可用于识别手机姿态的应用（比如横竖屏切换、相关游戏、磁力计姿态校准）、振动识别相关功能（比如计步器、敲击）等；至于手机还可配置的陀螺仪、气压计、湿度计、温度计、红外线传感器等其他传感器，在此不再赘述。

音频电路460、扬声器461，传声器462可提供用户与手机之间的音频接口。音频电路460可将接收到的音频数据转换后的电信号，传输到扬声器461，由扬声器461转换为声音信号输出；另一方面，传声器462将收集的声音信号转换为电信号，由音频电路460接收后转换为音频数据，再将音频数据输出处理器480处理后，经RF电路410以发送给比如另一手机，或者将音频数据输出至存储器420以便进一步处理。

WiFi属于短距离无线传输技术，手机通过WiFi模块470可以帮助用户收发电子邮件、浏览网页和访问流式媒体等，它为用户提供了无线的宽带互联网访问。虽然图9示出了WiFi模块470，但是可以理解的是，其并不属于手机的必须构成，完全可以根据需要在不改变发明的本质的范围内而省略。

处理器480是手机的控制中心，利用各种接口和线路连接整个手机的各个部分，通过运行或执行存储在存储器420内的软件程序和/或模块，以及调用存

储在存储器 420 内的数据，执行手机的各种功能和处理数据，从而对手机进行整体监控。可选的，处理器 480 可包括一个或多个处理单元；可选的，处理器 480 可集成应用处理器和调制解调处理器，其中，应用处理器主要处理操作系统、用户界面和应用程序等，调制解调处理器主要处理无线通信。可以理解的是，上述调制解调处理器也可以不集成到处理器 480 中。

手机还包括给各个部件供电的电源 490（比如电池），可选的，电源可以通过电源管理系统与处理器 480 逻辑相连，从而通过电源管理系统实现管理充电、放电、以及功耗管理等功能。

尽管未示出，手机还可以包括摄像头、蓝牙模块等，在此不再赘述。

在本发明实施例中，该终端设备所包括的处理器 480 用于实现上述方法实施例中所述的文本翻译方法。

所属领域的技术人员可以清楚地了解到，为描述的方便和简洁，上述描述的系统，装置和单元的具体工作过程，可以参考前述方法实施例中的对应过程，在此不再赘述。

在本发明所提供的几个实施例中，应该理解到，所揭露的系统，装置和方法，可以通过其它的方式实现。例如，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，装置或单元的间接耦合或通信连接，可以是电性，机械或其它的形式。

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。

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。

所述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时，可以存储在一个计算机可读取存储介质中。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一个存储介质中，包括若干指令用以使得一台计算机设备（可以是个人计算机，服务器，或者网络设备等）执行本发明各个实施例所述方法的全部或部分步骤。而前述的存储介质包括：U 盘、移动硬盘、只读存储器（read-only memory，ROM）、随机存取存储器（randomaccess memory，RAM）、磁碟或者光盘等各种可以存储程序代码的介质。

以上所述，以上实施例仅用以说明本发明的技术方案，而非对其限制；尽管参照前述实施例对本发明进行了详细的说明，本领域的普通技术人员应当理解：其依然可以对前述各实施例所记载的技术方案进行修改，或者对其中部分技术特征进行等同替换；而这些修改或者替换，并不使相应技术方案的本质脱离本发明各实施例技术方案的精神和范围。

权利要求书

1、一种文本翻译的方法，其特征在于，包括:

获取待翻译文本序列，其中，所述待翻译文本序列根据源语言文本生成，所述源语言文本包括至少一个词语;

对所述待翻译文本序列进行编码处理，得到第一隐含状态序列;

获取第一状态向量，其中，所述第一状态向量为当前词语的上一个词语所对应的状态向量，所述当前词语表示所述源语言文本中待翻译的词语，所述上一个词语表示所述源语言文本中已被翻译的词语;

根据所述第一状态向量以及所述第一隐含状态序列，生成第二隐含状态序列;

根据所述第二隐含状态序列以及所述第一状态向量，生成所述当前词语所对应的上下文向量;

根据所述上下文向量、所述第一状态向量以及第一目标词，确定第二目标词，其中，所述第一目标词为所述上一个词语所对应的翻译结果，所述第二目标词为所述当前词语所对应的翻译结果。

2、根据权利要求 1 所述的方法，其特征在于，所述根据所述第一状态向量以及所述第一隐含状态序列，生成第二隐含状态序列，包括:

根据目标隐含状态向量以及所述第一状态向量计算门函数，其中，所述目标隐含状态向量属于所述第一隐含状态序列中的一个隐含状态向量;

根据所述门函数以及所述目标隐含状态向量计算目标浅层理解向量;

根据所述目标浅层理解向量生成浅层理解序列，其中，所述浅层理解序列与所述第一隐含状态序列具有对应关系;

对所述浅层理解序列进行编码处理，得到所述第二隐含状态序列。

3、根据权利要求 2 所述的方法，其特征在于，所述根据目标隐含状态向量以及所述第一状态向量计算门函数，包括:

根据所述目标隐含状态向量、所述第一状态向量以及 S 形 sigmoid 函数，计算所述门函数;

所述根据所述门函数以及所述目标隐含状态向量计算目标浅层理解向量，包括:

将所述目标隐含状态向量与所述目标隐含向量进行元素级相乘，得到所述目标浅层理解向量。

4、根据权利要求 1 至 3 中任一项所述的方法，其特征在于，所述根据所述上下文向量、所述第一状态向量以及第一目标词，确定第二目标词之后，所述方法还包括:

根据所述上下文向量、第二状态向量以及所述第二目标词所对应的词向量，计算目标输出概率，其中，所述第二状态向量为所述当前词语所对应的状态向量;

根据所述目标输出概率计算连续采样向量，其中，所述连续采样向量用于生成连续采样序列;

根据所述连续采样向量计算离散采样向量，其中，所述离散采样向量用于生成离散采样序列;

根据所述离散采样向量计算编码处理结果;

根据所述编码处理结果确定处理模式，其中，所述处理模式包括第一处理模式以及第二处理模式，所述第一处理模式表示采用已有的编码结果，所述第二处理模式表示对所述第一隐含状态序列进行编码处理。

5、根据权利要求 4 所述的方法，其特征在于，所述根据所述上下文向量、第二状态向量以及所述第二目标词所对应的词向量，计算目标输出概率，包括:

根据所述上下文向量、所述第二状态向量以及所述第二目标词所对应的词向量，通过双曲正切函数确定策略函数的状态;

根据所述策略函数的状态，通过归一化指数函数计算所述目标输出概率。

6、根据权利要求 5 所述的方法，其特征在于，所述根据所述目标输出概率计算连续采样向量，包括:

通过预设采样函数对所述目标输出概率进行采样，得到采样参数，所述预设采样函数为耿贝尔归一化指数函数 Gumbel-softmax，所述采样参数离散;

对所述采样参数进行连续化处理，得到所述连续采样向量。

7、根据权利要求 6 所述的方法，其特征在于，所述根据所述连续采样向量计算离散采样向量，包括:

通过最大值自变量点集 argmax 函数对所述连续采样向量进行处理，生成所述离散采样向量，所述离散采样向量中包含的数值为 0 或 1。

8、根据权利要求 7 所述的方法，其特征在于，所述根据所述离散采样向量计算编码处理结果，包括:

根据第一离散采样向量、第二离散采样向量、所述第二隐含状态序列和第

三隐含状态序列，计算所述编码处理结果，所述第一离散采样向量为所述第一处理模式对应的离散采样向量，所述第二离散采样向量为所述第二处理模式对应的离散采样向量，所述第三隐含状态序列是经过二次编码后生成的隐含状态序列。

9、根据权利要求 4 所述的方法，其特征在于，所述方法还包括:

获取样本源语言文本对应的预测目标语言文本和样本目标语言文本，所述预测目标语言文本由文本翻译模型翻译得到，所述样本目标语言文本用于在模型训练时作为所述预测目标语言文本的监督;

根据所述预测目标语言文本和所述样本目标语言文本计算模型预测损失;

根据所述模型预测损失和惩罚项计算模型总损失，所述惩罚项根据第二离散采样向量和所述样本源语言文本的词语总数确定，所述第二离散采样向量是所述第二处理模式对应的离散采样向量;

根据所述模型总损失训练所述文本翻译模型。

10、一种文本翻译装置，其特征在于，包括:

获取模块，用于获取待翻译文本序列，其中，所述待翻译文本序列根据源语言文本生成，所述源语言文本包括至少一个词语;

编码模块，用于对所述获取模块获取的所述待翻译文本序列进行编码处理，得到第一隐含状态序列;

所述获取模块，还用于获取第一状态向量，其中，所述第一状态向量为当前词语的上一个词语所对应的状态向量，所述当前词语表示所述源语言文本中待翻译的词语，所述上一个词语表示所述源语言文本中已被翻译的词语;

生成模块，用于根据所述获取模块获取的所述第一状态向量以及所述第一隐含状态序列，生成第二隐含状态序列;

所述生成模块，还用于根据所述第二隐含状态序列以及所述第一状态向量，生成所述当前词语所对应的上下文向量;

确定模块，用于根据所述生成模块生成的所述上下文向量、所述第一状态向量以及第一目标词，确定第二目标词，其中，所述第一目标词为所述上一个词语所对应的翻译结果，所述第二目标词为所述当前词语所对应的翻译结果。

11、根据权利要求 10 所述的文本翻译装置，其特征在于，

所述生成模块，用于根据目标隐含状态向量以及所述第一状态向量计算门函数，其中，所述目标隐含状态向量属于所述第一隐含状态序列中的一个隐含

状态向量;

根据所述门函数以及所述目标隐含状态向量计算目标浅层理解向量;

根据所述目标浅层理解向量生成浅层理解序列，其中，所述浅层理解序列与所述第一隐含状态序列具有对应关系;

对所述浅层理解序列进行编码处理，得到所述第二隐含状态序列。

12、根据权利要求 11 所述的文本翻译装置，其特征在于，所述生成模块，用于:

根据所述目标隐含状态向量、所述第一状态向量以及 S 形 sigmoid 函数，计算所述门函数;

将所述目标隐含状态向量与所述目标隐含向量进行元素级相乘，得到所述目标浅层理解向量。

13、根据权利要求 9 至 12 中任一项所述的文本翻译装置，其特征在于，所述装置还包括:

计算模块，用于根据所述上下文向量、第二状态向量以及所述第二目标词所对应的词向量，计算目标输出概率，其中，所述第二状态向量为所述当前词语所对应的状态向量;

所述计算模块，还用于根据所述目标输出概率计算连续采样向量，其中，所述连续采样向量用于生成连续采样序列;

所述计算模块，还用于根据所述连续采样向量计算离散采样向量，其中，所述离散采样向量用于生成离散采样序列;

所述计算模块，还用于根据所述离散采样向量计算编码处理结果;

所述确定模块，用于根据所述编码处理结果确定处理模式，其中，所述处理模式包括第一处理模式以及第二处理模式，所述第一处理模式表示采用已有的编码结果，所述第二处理模式表示对所述第一隐含状态序列进行编码处理。

14、根据权利要求 13 所述的文本翻译装置，其特征在于，所述计算模块，还用于:

根据所述上下文向量、所述第二状态向量以及所述第二目标词所对应的词向量，通过双曲正切函数确定策略函数的状态;

根据所述策略函数的状态，通过归一化指数函数计算所述目标输出概率。

15、根据权利要求 14 所述的文本翻译装置，其特征在于，所述计算模块，还用于:

通过预设采样函数对所述目标输出概率进行采样，得到采样参数，所述预设采样函数为耿贝尔归一化指数函数 Gumbel-softmax，所述采样参数离散;

对所述采样参数进行连续化处理，得到所述连续采样向量。

16、根据权利要求 15 所述的文本翻译装置，其特征在于，所述计算模块，还用于:

通过最大值自变量点集 argmax 函数对所述连续采样向量进行处理，生成所述离散采样向量，所述离散采样向量中包含的数值为 0 或 1。

17、根据权利要求 16 所述的文本翻译装置，其特征在于，所述计算模块，还用于:

根据第一离散采样向量、第二离散采样向量、所述第二隐含状态序列和第三隐含状态序列，计算所述编码处理结果，所述第一离散采样向量为所述第一处理模式对应的离散采样向量，所述第二离散采样向量为所述第二处理模式对应的离散采样向量，所述第三隐含状态序列是经过二次编码后生成的隐含状态序列。

18、根据权利要求 13 所述的文本翻译装置，其特征在于，所述装置还包括训练模块，所述训练模块，用于:

获取样本源语言文本对应的预测目标语言文本和样本目标语言文本，所述预测目标语言文本由文本翻译模型翻译得到，所述样本目标语言文本用于在模型训练时作为所述预测目标语言文本的监督;

根据所述预测目标语言文本和所述样本目标语言文本计算模型预测损失;

根据所述模型预测损失和惩罚项计算模型总损失，所述惩罚项根据第二离散采样向量和所述样本源语言文本的词语总数确定，所述第二离散采样向量是所述第二处理模式对应的离散采样向量;

根据所述模型总损失训练所述文本翻译模型。

19、一种文本翻译装置，其特征在于，所述装置包括: 存储器、收发器、处理器以及总线系统;

其中，所述存储器用于存储程序;

所述处理器用于执行所述存储器中的程序，以实现如权利要求 1 至 9 任一所述的方法。

20、一种计算机可读存储介质，其特征在于，所述计算机可读存储介质中存储有指令，当所述在计算机上运行时，使得计算机执行权利要求 1 至 9 任一

所述的方法。

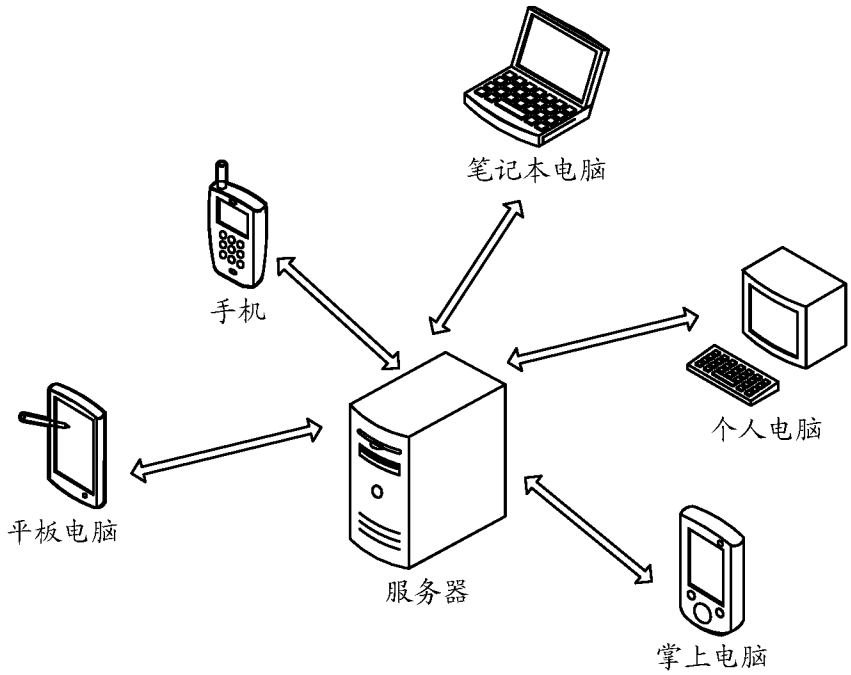
笔记本电脑
手机
个人电脑
平板电脑
服务器
掌上电脑

图 1

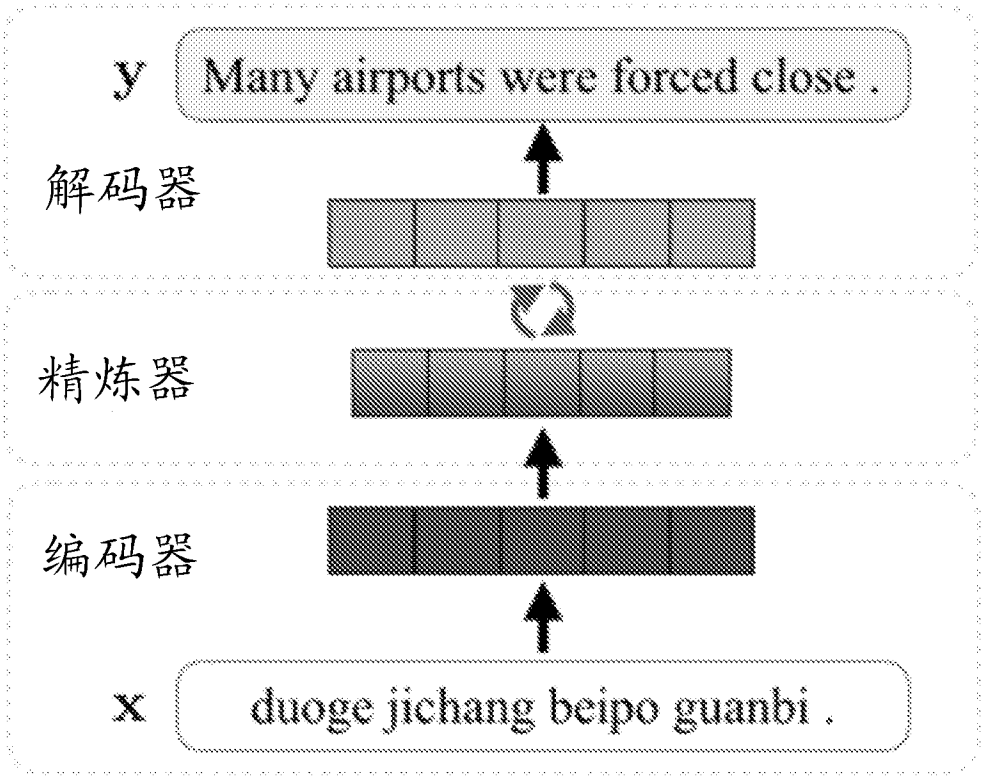
y
Many airports were forced close .
解码器
精炼器
编码器
x
duoge jichang beipo guanbi .

图 2

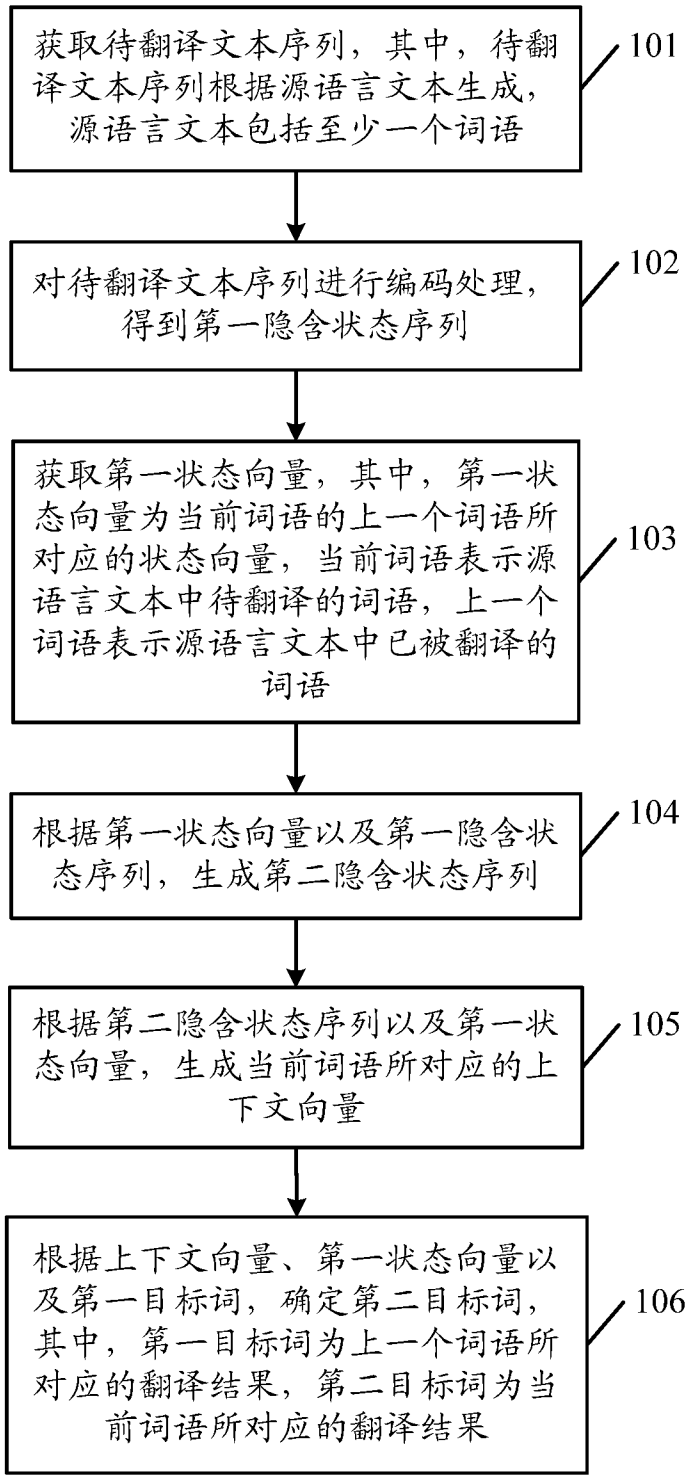
获取待翻译文本序列，其中，待翻译文本序列根据源语言文本生成，源语言文本包括至少一个词语
101
对待翻译文本序列进行编码处理，得到第一隐含状态序列
102
获取第一状态向量，其中，第一状态向量为当前词语的上一个词语所对应的状态向量，当前词语表示源语言文本中待翻译的词语，上一个词语表示源语言文本中已被翻译的词语
103
根据第一状态向量以及第一隐含状态序列，生成第二隐含状态序列
104
根据第二隐含状态序列以及第一状态向量，生成当前词语所对应的上下文向量
105
根据上下文向量、第一状态向量以及第一目标词，确定第二目标词，其中，第一目标词为上一个词语所对应的翻译结果，第二目标词为当前词语所对应的翻译结果
106

图 3

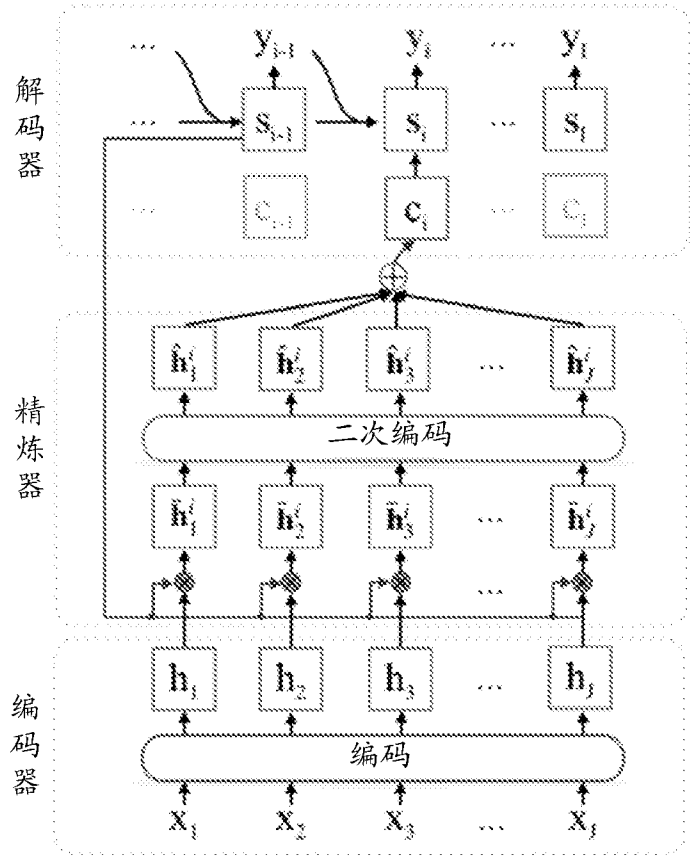
解码器
精炼器
编码器
二次编码
编码

图 4

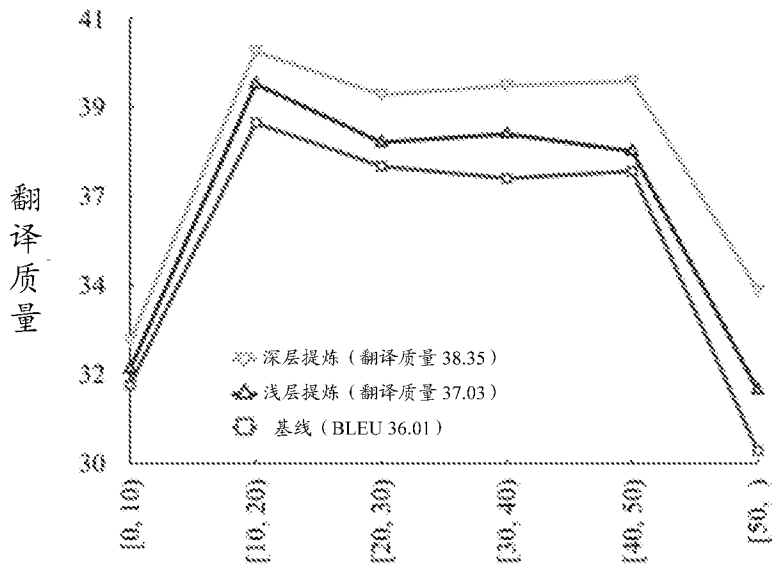
翻译质量
41
39
37
34
32
30
深层提炼（翻译质量 38.35）
浅层提炼（翻译质量 37.03）
基线（BLEU 36.01）
[0, 10)
[10, 20)
[20, 30)
[30, 40)
[40, 50)
[50,)
句子长度

图 5

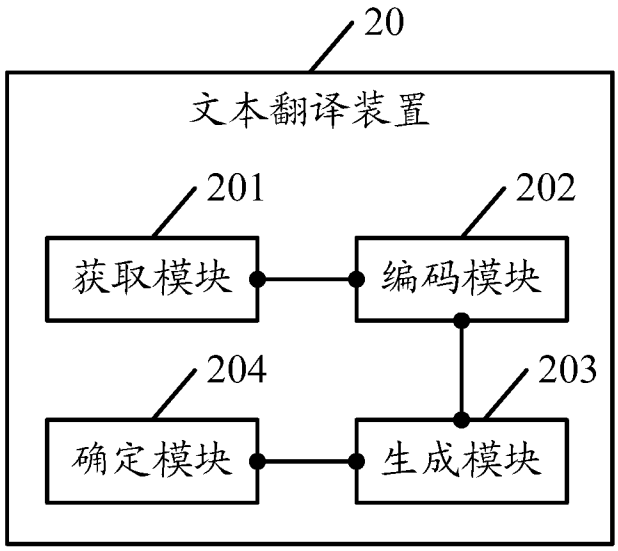
20
文本翻译装置
201
202
获取模块
编码模块
204
203
确定模块
生成模块

图 6

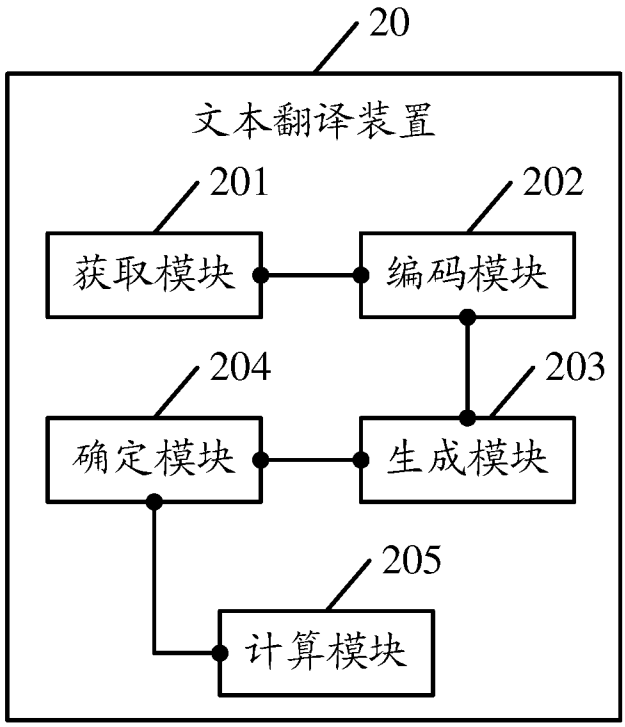
20
文本翻译装置
201
202
获取模块
编码模块
204
203
确定模块
生成模块
205
计算模块

图 7

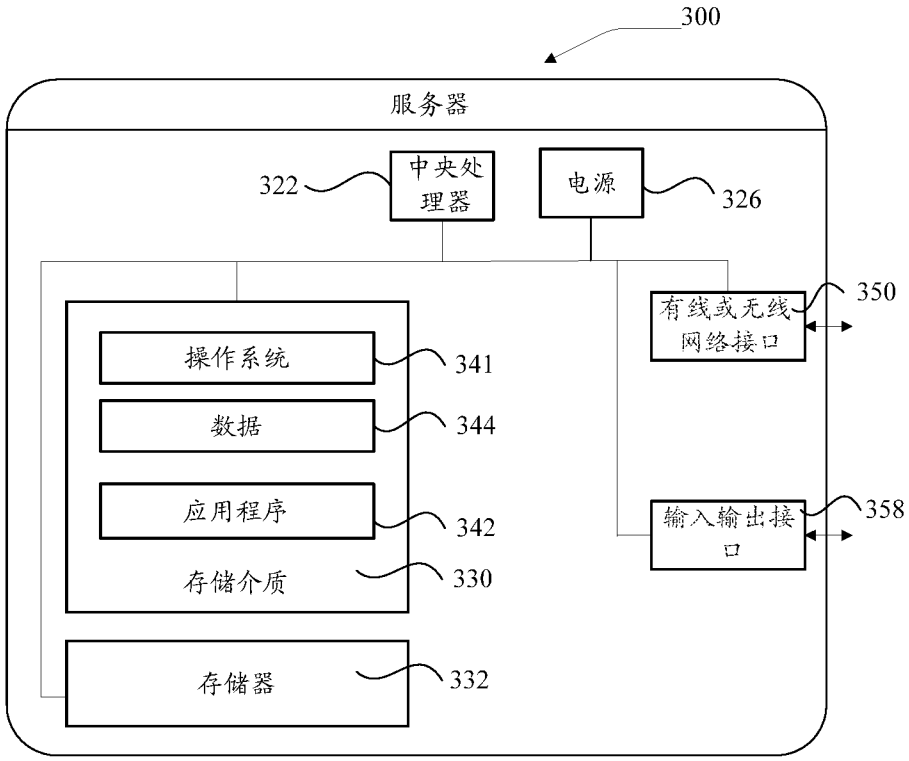
300
服务器
322
中央处理器
电源
326
有线或无线网络接口
350
操作系统
341
数据
344
应用程序
342
存储介质
330
输入输出接口
358
存储器
332

图 8

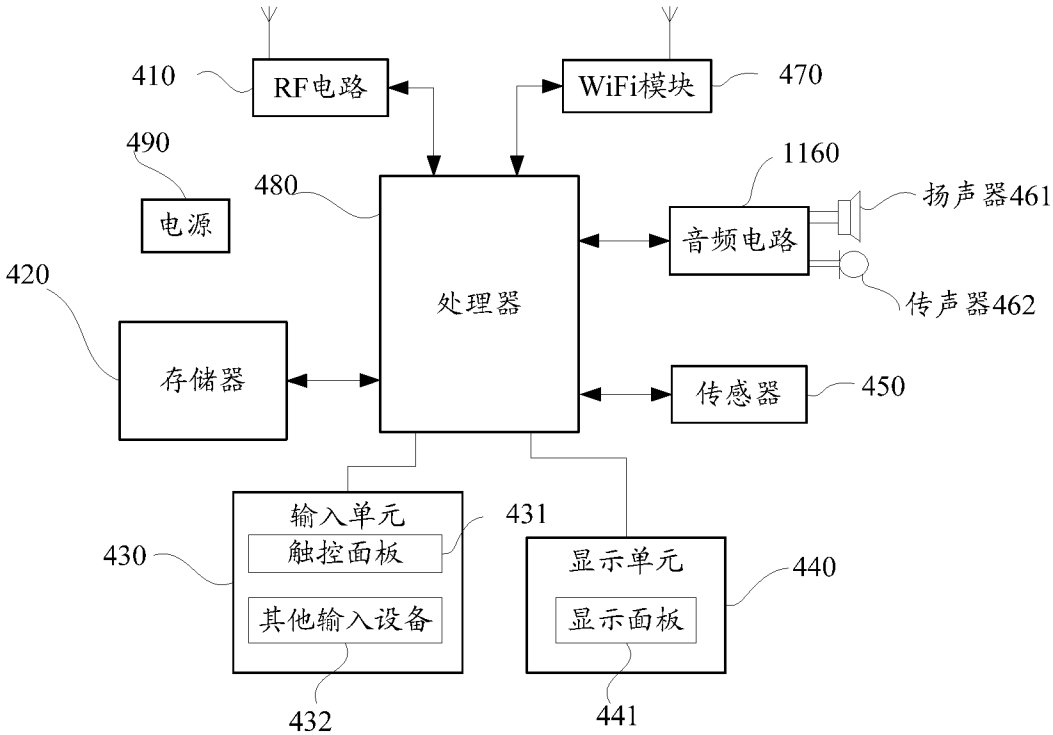
410
RF电路
WiFi模块
470
490
电源
480
处理器
1160
音频电路
扬声器461
传声器462
420
存储器
传感器
450
输入单元
431
触控面板
430
其他输入设备
432
显示单元
显示面板
440
441

图 9

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2019/120264

A. CLASSIFICATION OF SUBJECT MATTER

G06F 17/00(2019.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNPAT, CNKI, WPI, EPODOC, IEEE: 翻译, 上下文, 编码, 解码, 精炼, 提炼, 注意力, 向量, translat+, context, cod+, encod+, decod+, refin+, attention, vector

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 109543199 A (TENCENT TECHNOLOGY SHENZHEN CO., LTD.) 29 March 2019 (2019-03-29) claims 1-14, and description, paragraphs 0005-0391	1-20
X	CN 107368476 A (SHENZHEN TENCENT COMPUTER SYSTEM CO., LTD.) 21 November 2017 (2017-11-21) description, paragraphs 0021-0029 and 0052-0266, and figures 1-10	1-20
A	CN 107729329 A (SOOCHOW UNIVERSITY) 23 February 2018 (2018-02-23) entire document	1-20
A	CN 108304388 A (TENCENT TECHNOLOGY SHENZHEN CO., LTD.) 20 July 2018 (2018-07-20) entire document	1-20
A	US 2016179790 A1 (NATIONAL INSTITUTE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY) 23 June 2016 (2016-06-23) entire document	1-20

☐ Further documents are listed in the continuation of Box C. ☑ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
16 January 2020	**01 February 2020**
Name and mailing address of the ISA/CN **China National Intellectual Property Administration** **No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088** **China** Facsimile No. **(86-10)62019451**	Authorized officer Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2019/120264

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN 109543199 A	29 March 2019	None	
CN 107368476 A	21 November 2017	WO 2019019916 A1	31 January 2019
CN 107729329 A	23 February 2018	None	
CN 108304388 A	20 July 2018	US 2019384822 A1	19 December 2019
		KR 20190130636 A	22 November 2019
		WO 2019052293 A1	21 March 2019
US 2016179790 A1	23 June 2016	WO 2014196375 A1	11 December 2014
		JP 2014235599 A	15 December 2014
		EP 3007077 A1	13 April 2016
		CN 105190609 A	23 December 2015
		KR 20160016769 A	15 February 2016

国际检索报告

国际申请号
PCT/CN2019/120264

A. 主题的分类

G06F 17/00(2019.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

G06F

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称，和使用的检索词（如使用））

CNPAT, CNKI, WPI, EPODOC, IEEE: 翻译，上下文，编码，解码，精炼，提炼，注意力，向量，translat+, context, cod+, encod+, decod+, refin+, attention, vector

C. 相关文件

类 型*	引用文件，必要时，指明相关段落	相关的权利要求
PX	CN 109543199 A (腾讯科技深圳有限公司) 2019年 3月 29日 (2019-03-29) 权利要求1-14，说明书第0005-0391段	1-20
X	CN 107368476 A (深圳市腾讯计算机系统有限公司) 2017年 11月 21日 (2017-11-21) 说明书第0021-0029，0052-0266段、附图1-10	1-20
A	CN 107729329 A (苏州大学) 2018年 2月 23日 (2018-02-23) 全文	1-20
A	CN 108304388 A (腾讯科技深圳有限公司) 2018年 7月 20日 (2018-07-20) 全文	1-20
A	US 2016179790 A1 (NATIONAL INSTITUTE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY) 2016年 6月 23日 (2016-06-23) 全文	1-20

☐ 其余文件在C栏的续页中列出。

☑ 见同族专利附件。

* 引用文件的具体类型：

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件，或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件（如具体说明的）

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布，与申请不相抵触，但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件，单独考虑该文件，认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件，当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时，要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期	国际检索报告邮寄日期
2020年 1月 16日	2020年 2月 1日
ISA/CN的名称和邮寄地址 中国国家知识产权局(ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451	受权官员 贺希佳 电话号码 86-（10）-53961586

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2019/120264

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	109543199	A	2019年 3月 29日		无		
CN	107368476	A	2017年 11月 21日	WO	2019019916	A1	2019年 1月 31日
CN	107729329	A	2018年 2月 23日		无		
CN	108304388	A	2018年 7月 20日	US	2019384822	A1	2019年 12月 19日
				KR	20190130636	A	2019年 11月 22日
				WO	2019052293	A1	2019年 3月 21日
US	2016179790	A1	2016年 6月 23日	WO	2014196375	A1	2014年 12月 11日
				JP	2014235599	A	2014年 12月 15日
				EP	3007077	A1	2016年 4月 13日
				CN	105190609	A	2015年 12月 23日
				KR	20160016769	A	2016年 2月 15日