



(12) 发明专利

(10) 授权公告号 CN 110147533 B

(45) 授权公告日 2023. 08. 29

(21) 申请号 201910069752.6

G06F 16/35 (2019.01)

(22) 申请日 2019.01.24

G06F 40/284 (2020.01)

(65) 同一申请的已公布的文献号
申请公布号 CN 110147533 A

(56) 对比文件

(43) 申请公布日 2019.08.20

WO 2018058994 A1, 2018.04.05

(73) 专利权人 腾讯科技(深圳)有限公司
地址 518057 广东省深圳市南山区高新区
科技中一路腾讯大厦35层

US 2018365321 A1, 2018.12.20

CN 107832310 A, 2018.03.23

CN 108280112 A, 2018.07.13

KR 101929509 B1, 2018.12.14

US 2017024645 A1, 2017.01.26

US 2017200077 A1, 2017.07.13

CN 106919646 A, 2017.07.04

WO 2018165932 A1, 2018.09.20

CN 107229684 A, 2017.10.03

CN 108694225 A, 2018.10.23

CN 109241536 A, 2019.01.18

US 2016350653 A1, 2016.12.01

US 2018357225 A1, 2018.12.13

(续)

(72) 发明人 谭翊章 孙硕 曹杰 田乐 牛成
周杰

(74) 专利代理机构 北京三高永信知识产权代理
有限责任公司 11138
专利代理师 张所明

审查员 叶敏

(51) Int. Cl.

G06F 40/151 (2020.01)

G06F 40/205 (2020.01)

G06N 3/0442 (2023.01)

G06N 3/08 (2023.01)

G06F 16/332 (2019.01)

权利要求书3页 说明书15页 附图9页

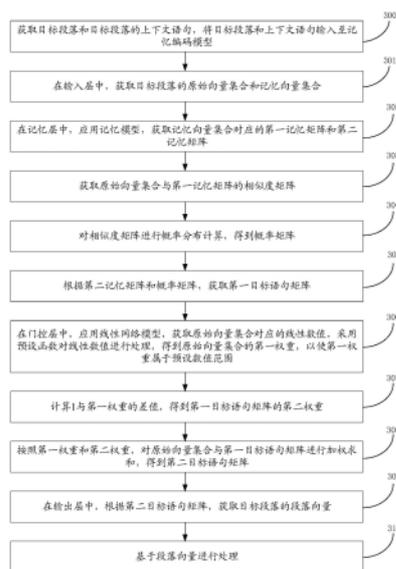
(54) 发明名称

编码方法、装置、设备及存储介质

准确率。

(57) 摘要

本发明实施例公开了一种编码方法、装置、设备及存储介质,属于自然语言处理领域。方法包括:获取目标段落和目标段落的上下文语句,将目标段落和上下文语句输入至记忆编码模型;在输入层中获取原始向量集合以及记忆向量集合,在记忆层中根据原始向量集合和记忆向量集合,获取原始向量集合的第一目标语句矩阵,在输出层中根据第一目标语句矩阵获取目标段落的段落向量,基于段落向量进行处理。以目标段落为单位对每个语句分别进行编码,提高了编码速度。并且,编码过程中不仅考虑目标段落本身,还考虑了目标段落的上下文语句,能够根据目标段落中的语句与上下文语句之间的关联关系获取到准确描述目标段落的段落向量,提高了编码



CN 110147533 B

[接上页]

(56) 对比文件

Donald McMillan.Repurposing
Conversation: Experiments with the

Continuous Speech Stream.CHI '15:
Proceedings of the 33rd Annual ACM
Conference on Human Factors in Computing
Systems.2015,全文.

1. 一种编码方法,其特征在于,所述方法包括:

获取目标段落和所述目标段落的上下文语句,将所述目标段落和所述上下文语句输入至记忆编码模型,所述目标段落包括至少一个语句,所述记忆编码模型至少包括输入层、记忆层、门控层和输出层;

在所述输入层中,获取所述目标段落的原始向量集合和记忆向量集合,所述原始向量集合包括所述目标段落中每个语句的语句向量,所述记忆向量集合包括所述目标段落的上下文语句中每个词语的词向量;

在所述记忆层中,采用注意力学习机制,从所述记忆向量集合中抽取与所述原始向量集合相关的信息,得到第一目标语句矩阵,所述第一目标语句矩阵用于根据所述原始向量集合与所述记忆向量集合之间的关联关系,对所述目标段落进行描述;

所述门控层包括线性网络模型,在所述门控层中,应用所述线性网络模型,获取所述原始向量集合对应的线性数值,采用预设函数对所述线性数值进行处理,得到所述原始向量集合的第一权重,以使所述第一权重属于预设数值范围;计算1与所述第一权重的差值,得到所述第一目标语句矩阵的第二权重;按照所述第一权重和所述第二权重,对所述原始向量集合与所述第一目标语句矩阵进行加权求和,得到第二目标语句矩阵;

在所述输出层中,对所述第二目标语句矩阵进行列向求和,得到段落向量;

基于所述段落向量进行处理。

2. 根据权利要求1所述的方法,其特征在于,所述输入层包括语句编码模型,所述在所述输入层中,获取所述目标段落的原始向量集合和记忆向量集合,包括:

根据所述目标段落中每个语句中每个词语的词向量,应用所述语句编码模型,获取所述每个语句的语句向量,得到所述原始向量集合;

根据所述上下文语句中每个词语的词向量,获取所述记忆向量集合。

3. 根据权利要求2所述的方法,其特征在于,所述语句编码模型包括第一语句编码子模型和第二语句编码子模型;所述根据所述目标段落中每个语句中每个词语的词向量,应用所述语句编码模型,获取所述每个语句的语句向量,得到所述原始向量集合,包括:

对于所述目标段落中的每个语句,

获取所述语句中每个词语的词向量,得到多个词向量;

应用所述第一语句编码子模型,对所述多个词向量进行正序编码,得到第一向量,应用所述第二语句编码子模型,对所述多个词向量进行倒序编码,得到第二向量;

根据所述第一向量和所述第二向量,获取所述语句的语句向量。

4. 根据权利要求1所述的方法,其特征在于,所述记忆层包括记忆模型,所述在所述记忆层中,采用注意力学习机制,从所述记忆向量集合中抽取与所述原始向量集合相关的信息,得到第一目标语句矩阵,包括:

应用所述记忆模型,获取所述记忆向量集合对应的第一记忆矩阵和第二记忆矩阵;

根据所述原始向量集合、所述第一记忆矩阵和所述第二记忆矩阵,获取所述第一目标语句矩阵。

5. 根据权利要求4所述的方法,其特征在于,所述记忆向量集合与所述原始向量集合相同,所述根据所述原始向量集合、所述第一记忆矩阵和所述第二记忆矩阵,获取所述第一目标语句矩阵,包括:

获取所述原始向量集合与所述第一记忆矩阵的相似度矩阵；

对所述相似度矩阵进行概率分布计算，得到概率矩阵；

根据所述第二记忆矩阵和所述概率矩阵，获取所述第一目标语句矩阵。

6. 根据权利要求1-5任一项所述的方法，其特征在于，所述方法还包括：

将所述第二目标语句矩阵作为更新后的原始向量集合和记忆向量集合，在所述记忆层和所述门控层中，重复执行根据更新后的原始向量集合和记忆向量集合获取目标语句矩阵的步骤，直至重复次数达到预设次数时，在所述门控层中，按照所述第一权重和所述第二权重，对所述原始向量集合与当前的第一目标语句矩阵进行加权求和，得到第二目标语句矩阵。

7. 一种编码装置，其特征在于，所述装置包括：

获取模块，用于获取目标段落和所述目标段落的上下文语句，将所述目标段落和所述上下文语句输入至记忆编码模型，所述目标段落包括至少一个语句，所述记忆编码模型至少包括输入层、记忆层、门控层和输出层；

输入层模块，用于获取所述目标段落的原始向量集合和记忆向量集合，所述原始向量集合包括所述目标段落中每个语句的语句向量，所述记忆向量集合包括所述目标段落的上下文语句中每个词语的词向量；

记忆层模块，用于采用注意力学习机制，从所述记忆向量集合中抽取与所述原始向量集合相关的信息，得到第一目标语句矩阵，所述第一目标语句矩阵用于根据所述原始向量集合与所述记忆向量集合之间的关联关系，对所述目标段落进行描述；

门控层模块，用于应用线性网络模型，获取所述原始向量集合对应的线性数值，采用预设函数对所述线性数值进行处理，得到所述原始向量集合的第一权重，以使所述第一权重属于预设数值范围；计算1与所述第一权重的差值，得到所述第一目标语句矩阵的第二权重；按照所述第一权重和所述第二权重，对所述原始向量集合与所述第一目标语句矩阵进行加权求和，得到第二目标语句矩阵；

输出层模块，用于对所述第二目标语句矩阵进行列向求和，得到段落向量；

处理模块，用于基于所述段落向量进行处理。

8. 根据权利要求7所述的装置，其特征在于，所述输入层模块包括：

第一获取单元，用于根据所述目标段落中每个语句中每个词语的词向量，应用语句编码模型，获取所述每个语句的语句向量，得到所述原始向量集合；

第二获取单元，用于根据所述上下文语句中每个词语的词向量，获取所述记忆向量集合。

9. 根据权利要求8所述的装置，其特征在于，所述语句编码模型包括第一语句编码子模型和第二语句编码子模型；所述第一获取单元，用于：

对于所述目标段落中的每个语句，

获取所述语句中每个词语的词向量，得到多个词向量；

应用所述第一语句编码子模型，对所述多个词向量进行正序编码，得到第一向量，应用所述第二语句编码子模型，对所述多个词向量进行倒序编码，得到第二向量；

根据所述第一向量和所述第二向量，获取所述语句的语句向量。

10. 根据权利要求7所述的装置，其特征在于，所述记忆层模块，包括：

记忆获取单元,用于应用记忆模型,获取所述记忆向量集合对应的第一记忆矩阵和第二记忆矩阵;

第一目标获取单元,用于根据所述原始向量集合、所述第一记忆矩阵和所述第二记忆矩阵,获取所述原始向量集合的第一目标语句矩阵。

11. 根据权利要求10所述的装置,其特征在于,所述记忆向量集合与所述原始向量集合相同,所述第一目标获取单元用于获取所述原始向量集合与所述第一记忆矩阵的相似度矩阵;对所述相似度矩阵进行概率分布计算,得到概率矩阵;根据所述第二记忆矩阵和所述概率矩阵,获取所述第一目标语句矩阵。

12. 根据权利要求7-11任一项所述的装置,其特征在于,所述记忆层模块,还用于将所述第二目标语句矩阵作为更新后的原始向量集合和记忆向量集合,在所述记忆层和所述门控层中,重复执行根据更新后的原始向量集合和记忆向量集合获取目标语句矩阵的步骤,直至重复次数达到预设次数时,触发所述输出层模块;

所述输出层模块,还用于在所述门控层中,按照所述第一权重和所述第二权重,对所述原始向量集合与当前的第一目标语句矩阵进行加权求和,得到第二目标语句矩阵。

13. 一种编码设备,其特征在于,所述编码设备包括一个或多个处理器和一个或多个存储器,所述一个或多个存储器中存储有至少一条指令、至少一段程序、代码集或指令集,所述指令、所述程序、所述代码集或所述指令集由所述一个或多个处理器加载并执行以实现如权利要求1至6任一权利要求所述的编码方法中所执行的操作。

14. 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至少一条指令、至少一段程序、代码集或指令集,所述指令、所述程序、所述代码集或所述指令集由一个或多个处理器加载并执行以实现如权利要求1至6任一权利要求所述的编码方法中所执行的操作。

编码方法、装置、设备及存储介质

技术领域

[0001] 本发明涉及自然语言处理领域,特别涉及一种编码方法、装置、设备及存储介质。

背景技术

[0002] 编码是将文字转换成编码值,从而得到能够准确描述该文本含义的向量的过程。通过进行编码,可以将文字转换为便于运算处理的向量形式,现已广泛应用于语句选取、语句生成等多种领域。

[0003] 目前提出了一种分级编码的方案,当要对包括多个语句的目标段落进行编码时,获取目标段落中每个语句中每个词语的词向量。对于目标段落中的每个语句,应用基于词语层面的第一编码模型,将该语句中每个词语的词向量编码成一个向量,得到该语句的语句向量,进而得到目标段落中多个语句的语句向量。再应用基于语句层面的第二编码模型,将该多个语句的语句向量编码成一个向量,得到目标段落的段落向量。

[0004] 上述方案仅能够采用串行的方式,依次对目标段落中每个语句的词向量分别进行编码,再采用串行的方式,对多个语句向量进行编码,编码速度较慢,且编码过程中未考虑目标段落的上下文之间的关联关系,准确率较低。

发明内容

[0005] 本发明实施例提供了一种编码方法、装置、设备及存储介质,解决了相关技术存在的编码速度较慢以及准确率较低的问题。所述技术方案如下:

[0006] 一方面,提供了一种编码方法,所述方法包括:

[0007] 获取目标段落和所述目标段落的上下文语句,将所述目标段落和所述上下文语句输入至记忆编码模型,所述目标段落包括至少一个语句,所述记忆编码模型至少包括输入层、记忆层和输出层;

[0008] 在所述输入层中,获取所述目标段落的原始向量集合和记忆向量集合,所述原始向量集合包括所述目标段落中每个语句的语句向量,所述记忆向量集合包括所述目标段落的上下文语句中每个词语的词向量;

[0009] 在所述记忆层中,根据所述原始向量集合和所述记忆向量集合,获取所述原始向量集合的第一目标语句矩阵,所述第一目标语句矩阵用于根据所述原始向量集合与所述记忆向量集合之间的关联关系,对所述目标段落进行描述;

[0010] 在所述输出层中,根据所述第一目标语句矩阵,获取所述目标段落的段落向量;

[0011] 基于所述段落向量进行处理。

[0012] 另一方面,提供了一种编码装置,所述装置包括:

[0013] 获取模块,用于获取目标段落和所述目标段落的上下文语句,将所述目标段落和所述上下文语句输入至记忆编码模型,所述目标段落包括至少一个语句;

[0014] 输入层模块,用于获取所述目标段落的原始向量集合和记忆向量集合,所述原始向量集合包括所述目标段落中每个语句的语句向量,所述记忆向量集合包括所述目标段落

的上下文语句中每个词语的词向量；

[0015] 记忆层模块,用于根据所述原始向量集合和所述记忆向量集合,获取所述原始向量集合的第一目标语句矩阵,所述第一目标语句矩阵用于根据所述原始向量集合与所述记忆向量集合之间的关联关系,对所述目标段落进行描述;

[0016] 输出层模块,用于根据所述第一目标语句矩阵,获取所述目标段落的段落向量;

[0017] 处理模块,用于基于所述段落向量进行处理。

[0018] 再一方面,提供了一种编码设备,所述编码设备包括一个或多个处理器和一个或多个存储器,所述一个或多个存储器中存储有至少一条指令、至少一段程序、代码集或指令集,所述指令、所述程序、所述代码集或所述指令集由所述一个或多个处理器加载并执行以实现如所述的编码方法中所执行的操作。

[0019] 再一方面,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至少一条指令、至少一段程序、代码集或指令集,所述指令、所述程序、所述代码集或所述指令集由一个或多个处理器加载并执行以实现如所述的编码方法中所执行的操作。

[0020] 本发明实施例提供的技术方案带来的有益效果是:

[0021] 本发明实施提供的编码方法,提供了一种记忆编码模型,记忆编码模型包括输入层、记忆层和输出层,获取目标段落和目标段落的上下文语句,将目标段落和上下文语句输入至记忆编码模型,在输入层中获取目标段落的原始向量集合以及目标段落的上下文语句中的记忆向量集合,在记忆层中根据原始向量集合和记忆向量集合,获取原始向量集合的第一目标语句矩阵,在输出层中根据第一目标语句矩阵获取目标段落的段落向量,基于段落向量进行处理。本发明实施例无需采用串行的方式对每个语句分别进行编码,而是以目标段落为单位,应用记忆编码模型对目标段落进行编码,因此提高了编码速度。并且,编码过程中不仅考虑目标段落本身,还考虑了目标段落的上下文语句,能够根据目标段落中的语句与上下文语句之间的关联关系获取到准确描述目标段落的段落向量,提高了编码准确率。

[0022] 本发明实施例提供的记忆编码模型具有自注意力性,将自注意力机制应用于段落的语句层面上,根据目标段落和上下文语句进行综合处理,可以保证目标段落的段落向量表达更为丰富,更能准确地描述目标段落含义。且本发明实施例可以应用于多种场景下,应用范围广泛。

附图说明

[0023] 为了更清楚地说明本发明实施例中的技术方案,下面将对实施例描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图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的前提下,还可以根据这些附图获得其他的附图。

[0024] 图1是本发明实施例提供的一种记忆编码模型的结构示意图;

[0025] 图2是本发明实施例提供的另一种记忆编码模型的结构示意图;

[0026] 图3是本发明实施例提供的一种编码方法的流程图;

[0027] 图4是本发明实施例提供的一种语句编码模型的结构示意图;

[0028] 图5是本发明实施例提供的一种语句编码模型的流程示意图;

- [0029] 图6是本发明实施例提供的一种记忆编码模型的结构示意图；
- [0030] 图7是本发明实施例提供的一种记忆层的结构示意图；
- [0031] 图8是本发明实施例提供的一种门控层的结构示意图；
- [0032] 图9是本发明实施例提供的一种记忆编码模型的结构示意图；
- [0033] 图10是本发明实施例提供的一种记忆编码模型的结构示意图；
- [0034] 图11是本发明实施例提供的一种编码装置的结构示意图；
- [0035] 图12是本发明实施例提供的一种终端的结构框图；
- [0036] 图13是本发明实施例提供的一种服务器的结构示意图。

具体实施方式

[0037] 为使本发明的目的、技术方案和优点更加清楚，下面将结合附图对本发明实施方式作进一步地详细描述。

[0038] 本发明实施例提供了一种记忆编码模型，获取目标段落和目标段落的上下文语句，输入至记忆编码模型，应用该记忆编码模型可以对该目标段落进行编码，得到该目标段落的段落向量，从而能够基于段落向量进行处理。编码过程中，可以目标段落为单位，一次性地应用该记忆编码模型对目标段落进行编码，无需采用串行的方式分别对目标段落中每个语句进行编码。而且，不仅考虑到目标段落中每个语句的含义，还考虑到目标段落中的语句与目标段落的上下文语句之间的关联联系，从而使获取到的段落向量更能准确表达目标段落含义，基于段落向量进行处理时能够提高精确度。

[0039] 参见图1，该记忆编码模型包括输入层101、记忆层102和输出层103，输入层101与记忆层102连接，记忆层102与输出层103连接。

[0040] 其中，输入层101根据目标段落中的每个语句，提取代表语句含义的语句向量，得到该目标段落的原始向量集合，将原始向量集合输入至记忆层102中。且，输入层101还会根据目标段落的上下文语句中的每个词语，获取每个词语的词向量，并将获取的多个词向量组成记忆向量集合，输入至记忆层102中。记忆层102根据输入的原始向量集合与记忆向量集合，获取第一目标语句矩阵，并将第一目标语句矩阵输入至输出层103；输出层103根据第一目标语句矩阵，获取该目标段落的段落向量。

[0041] 由于记忆层102采用了注意力学习机制，能够在记忆向量集合中抽取对原始向量集合影响较大的相关信息，因此能够考虑目标段落中的语句与上下文语句之间的关联关系，获取到更为准确的段落向量。

[0042] 在一种可能实现方式中，参见图2，该记忆编码模型还包括门控层104，输入层101与记忆层102和门控层104连接，记忆层102与门控层104连接，门控层104与输出层103连接。即本发明实施例提供了一种GSMN(Gated Self-attentive Memory Network,门控自注意力型记忆网络)模型。

[0043] 记忆层102得到第一目标语句矩阵后，输入至门控层104，门控层104对原始向量集合和第一目标矩阵进行加权求和，得到第二目标语句矩阵，将第二目标语句矩阵输入至输出层103，输出层103根据第二目标语句矩阵，获取该目标段落的段落向量。

[0044] 在一种可能实现方式中，该记忆网络模型会重复运行记忆层102和门控层104，将门控层104输出的第二目标语句矩阵作为记忆层102的原始向量集合和记忆向量集合，重新

输入至记忆层102中,重复运行记忆层102和门控层104,直至重复次数达到预设次数时,将当前的目标语句矩阵输入至输出层103,得到目标段落的段落向量。其中,重复次数可以是2次或3次,或者可以为其他数值。

[0045] 本发明实施例可以应用于对任一段落进行编码的场景下。

[0046] 例如,在智能对话的场景下,用户与聊天机器人进行对话,聊天机器人可以获取用户输入的文本消息,作为目标段落,采用本发明实施例提供的方法,对目标段落进行编码得到段落向量,将该段落向量与语料数据库中多个答复消息的向量进行匹配,得到其向量与该段落向量匹配的答复消息,展示给用户,实现了用户与聊天机器人进行对话的效果。

[0047] 由于编码过程中会考虑用户输入的文本消息以及该文本消息的上下文语句,因此生成的段落向量较为准确,能够使聊天机器人更好地理解用户想要表达的含义,根据该段落向量可以获得到更为匹配的答复消息,能够针对用户输入的文本消息给出更符合用户需求的答复,提升对话效果。

[0048] 或者,在文本分类的场景下,获取待分类的目标段落,采用本发明实施例提供的方法,对目标段落进行编码得到段落向量,根据该段落向量进行分类,可以确定该目标段落所属的类别。

[0049] 由于编码过程中会考虑目标段落的上下文语句,因此生成的段落向量较为准确,能够更好地理解目标段落的含义,根据该段落向量进行分类,可以提高分类准确度。

[0050] 或者,在段落选取的场景下,获取待选取的多个目标段落,采用本发明实施例提供的方法,对每个目标段落进行编码得到段落向量,根据多个目标段落的段落向量,从多个目标段落中选取满足要求的目标段落。

[0051] 由于编码过程中会考虑目标段落的上下文语句,因此生成的段落向量较为准确,能够更好地理解目标段落的含义,根据段落向量进行选取,可以选取到满足要求的目标段落,避免选取错误的问题。

[0052] 除上述场景之外,本发明实施例提供的方法还可以应用于阅读理解等其他场景下,本发明实施例对此不做限定。

[0053] 图3是本发明实施例提供的一种编码方法的流程图,本发明实施例对应用记忆编码模型对目标段落进行编码的过程进行说明,执行主体为编码设备,该编码设备可以为服务器或者还可以为手机、计算机等终端。参见图3,该方法包括:

[0054] 300、获取目标段落和目标段落的上下文语句,将目标段落和上下文语句输入至记忆编码模型。

[0055] 其中,该目标段落包括至少一个语句,每个语句包括至少一个词语。目标段落的上下文语句可以包括目标段落之前的一个或多个段落中的语句、目标段落之后的一个或多个段落中的语句,或者还可以包括目标段落中的一个或多个语句。例如,目标段落的上下文语句可以为目标段落的原文。

[0056] 在一种可能实现方式中,该目标段落为某一文章中的一个段落,则上下文语句可以包括该文章中在该段落之前或之后的语句,或者还可以包括该段落中的语句。或者,该目标段落为智能对话场景中用户输入的某一段文本,则上下文语句可以包括在该目标段落之前用户输入的文本,或者该目标段落中的文本,或者还可以包括在该目标段落之前聊天机器人回复用户的文本等。

[0057] 对目标段落进行编码,则获取该目标段落与该目标段落的上下文语句,将目标段落和上下文语句输入至记忆编码模型中。

[0058] 301、在输入层中,获取目标段落的原始向量集合和记忆向量集合。

[0059] 输入层为记忆编码模型中的第一层,当要对目标段落进行编码时,将目标段落和目标段落的上下文语句输入至输入层中,在输入层中对目标段落以及目标段落的上下文语句进行处理,获取目标段落的原始向量集合和记忆向量集合。

[0060] 该原始向量集合包括目标段落中每个语句的语句向量,记忆向量集合包括目标段落的上下文语句中每个词语的词向量。本发明实施例中,在对目标段落进行编码时,不仅要考虑目标段落本身,还要考虑目标段落的上下文语句,因此不仅要获取原始向量集合,还要获取记忆向量集合,后续会根据原始向量集合和记忆向量集合进行处理。

[0061] 在一种可能实现方式中,步骤301可以包括以下步骤3011-3012:

[0062] 3011、根据目标段落中每个语句中每个词语的词向量,应用语句编码模型,获取每个语句的语句向量,得到原始向量集合。

[0063] 首先对该目标段落进行预处理,该预处理过程包括:对目标段落进行语句划分,得到目标段落中的每个语句,对每个语句进行词语划分,得到每个语句中的每个词语,获取每个词语的词向量。

[0064] 其中,针对语句划分过程:可以获取目标段落中能够代表对应的语句已经结束的标点符号,如句号、问号、感叹号等,根据获取的标点符号对目标段落进行划分,即可得到目标段落中的语句。

[0065] 针对词语划分过程:可以采用分词算法对每个语句进行分词,该分词算法可以包括多种算法,如双向最大匹配法、最少切分法等。或者采用其他方式进行词语划分。

[0066] 针对获取词向量的过程:对于每个词语,可以根据词向量字典查询该词语对应的词向量,该词向量字典可以包括词语与词向量之间的对应关系,或者该词向量字典可以为词向量获取模型,如循环神经网络模型、深度学习网络模型、卷积神经网络模型等,应用该词向量获取模型可以获取词语的词向量。

[0067] 对目标段落进行预处理之后,对于每个语句,应用语句编码模型,对该语句中每个词语的词向量进行处理,获取该语句的语句向量,从而能够得到目标段落中每个语句的语句向量,根据每个语句的语句向量,组成原始向量集合。

[0068] 其中,语句编码模型用于将任一语句中多个词语的词向量压缩成一个代表该语句含义的语句向量,可以为循环神经网络模型、深度学习网络模型、卷积神经网络模型、变换神经网络模型、基于词语层面的GSMN模型等多种类型的模型。

[0069] 在一种可能实现的方式中,语句编码模型包括第一语句编码子模型和第二语句编码子模型,获取语句的语句向量的过程可以包括:对于目标段落中的每个语句,获取该语句中每个词语的词向量,得到多个词向量;应用第一语句编码子模型,对多个词向量进行正序编码,得到第一向量,应用第二语句编码子模型,对多个词向量进行倒序编码,得到第二向量;根据第一向量和第二向量,获取语句的语句向量。重复执行上述步骤,可以获取到目标段落中每个语句的语句向量。

[0070] 其中,该第一语句编码子模型为正序编码模型,该第二语句编码子模型为倒序编码模型。该语句中的多个词语的词向量按照顺序排列,则应用第一语句编码子模型,会按照

多个词向量的排列顺序对多个词向量进行正序编码,得到第一向量。而应用第二语句编码模型,会对多个词向量进行倒序处理,再按照倒序处理后的排列顺序对多个词向量进行倒序编码,得到第二向量。

[0071] 另外,获取到第一向量和第二向量之后,可以将第一向量和第二向量串联,得到语句向量,或者将第一向量和第二向量相加,得到语句向量,或者还可以采用其他方式得到语句向量。

[0072] 以语句编码模型为双向循环神经网络模型为例进行说明,如图4和图5所示,双向循环神经网络模型包括一个前向循环神经网络模型和一个后向循环神经网络模型,通过前向循环神经网络模型对语句的多个词向量进行正序编码,获取第一向量,通过后向循环神经网络模型对语句的多个词向量进行倒序编码,获取第二向量,将第一向量和第二向量串联,得到该语句的语句向量。

[0073] 3012、根据上下文语句中每个词语的词向量,获取记忆向量集合。

[0074] 对上下文语句进行词语划分,得到上下文语句中的每个词语,之后获取每个词语的词向量,根据获取的词向量组成记忆向量集合。其中,进行词语划分以及获取词语的词向量的过程与上述步骤3011类似,在此不再赘述。

[0075] 需要说明的是,如果目标段落与上下文语句相同,则只需对目标段落中的语句进行处理即可得到原始向量集合和记忆向量集合,而无需对其他语句进行处理。如图6所示,根据目标段落进行预处理后得到的词向量获取记忆向量集合。

[0076] 本发明实施例中,记忆编码模型以目标段落为单位进行编码,因此输入层将获取到的原始向量集合和记忆向量集合均输入到记忆层中进行处理。

[0077] 302、在记忆层中,应用记忆模型,获取记忆向量集合对应的第一记忆矩阵和第二记忆矩阵。

[0078] 记忆层包括记忆模型,应用该记忆模型可以获取该记忆向量集合对应的第一记忆矩阵和第二记忆矩阵,其中,第一记忆矩阵和第二记忆矩阵用于对该记忆向量集合进行描述,且第一记忆矩阵和第二记忆矩阵可以相同,也可以不同。

[0079] 针对第一记忆矩阵的获取方式:可以根据记忆向量集合获取上下文语句中每个词语的词向量,应用语句编码模型,获取每个语句的语句向量,根据每个语句的语句向量获取第一记忆矩阵。

[0080] 在一种可能实现方式中,该语句编码模型包括第三语句编码模型和第四语句编码模型,获取语句的语句向量过程可以包括:对于上下文语句中的每个语句,获取该语句中每个词语的词向量,得到多个词向量;应用第三语句编码模型,对多个词向量进行正序编码,得到第三向量,应用第四语句编码模型,对多个词向量进行倒序编码,得到第四向量;根据第三向量和第四向量,获取语句的语句向量。

[0081] 其中,获取语句向量的具体过程与上述步骤3011类似,在此不再赘述。

[0082] 获取到上下文语句中的每个语句的语句向量后,将这些语句的语句向量进行组合,得到第一记忆矩阵。

[0083] 另外,第二记忆矩阵的获取方式与第一记忆矩阵的获取方式类似,区别仅在于采用的语句编码模型可以与获取第一记忆矩阵时采用的语句编码模型相同或者不同。

[0084] 如图7所示,获取第一记忆矩阵和第二记忆矩阵时采用的语句编码模型均为双向

循环神经网络模型,应用这两个双向循环神经网络模型对记忆向量集合分别进行处理,可以得到第一记忆矩阵和第二记忆矩阵。这两个双向循环神经网络模型的参数可以相同,也可以不同,因此得到的第一记忆矩阵和第二记忆矩阵可以相同,也可以不同。

[0085] 由于第一记忆矩阵和第二记忆矩阵能够对记忆向量集合进行描述,因此根据第一记忆矩阵和第二记忆矩阵以及原始向量集合进行处理,能够考虑上下文语句与目标段落之间的关联关系,以便得到能够更加准确地描述该目标段落的段落向量。

[0086] 本发明实施例以目标段落与上下文语句相同为例,则原始向量集合与记忆向量集合相同。在此情况下,可以执行下述步骤303-305获取用于描述目标段落的第一目标语句矩阵。当然,在目标段落与上下文语句不同的情况下,还可以采用多种方式获取第一目标语句矩阵。

[0087] 303、获取原始向量集合与第一记忆矩阵的相似度矩阵。

[0088] 其中,获取相似度矩阵的方式有多种,如矩阵相乘法、矩阵相减法等。在一种可能实现方式中,参见图7,根据该原始向量集合中的语句向量进行组合,得到目标段落的原始语句矩阵,将该原始语句矩阵与第一记忆矩阵相乘,得到的矩阵作为相似度矩阵。或者,还可以将该原始语句矩阵与第一记忆矩阵的转置相乘,得到的矩阵作为相似度矩阵。

[0089] 相似度矩阵中的每一个数值代表了原始向量集合中的语句与上下文语句中对应的语句之间的相似度,相似度越高,表示两个语句关联越紧密,在后续的处理过程中越应当注意该语句。

[0090] 304、对相似度矩阵进行概率分布计算,得到概率矩阵。

[0091] 相似度矩阵中包括多个相似度,对相似度矩阵进行概率分布计算,可以得到概率矩阵,概率矩阵中包括每个相似度对应的概率,且所有相似度的概率之和为1。

[0092] 其中,概率分布计算方式可以有多种,在一种可能实现方式中,采用Softmax(归一化指数)函数对相似度矩阵进行计算,得到与相似度矩阵对应的概率矩阵。或者,对于相似度矩阵中的每个位置,获取该位置上的相似度与相似度矩阵中所有相似度之和的比值,得到该位置上的相似度对应的概率,从而获取到每个位置上的相似度对应的概率,将获取到的概率组成概率矩阵。

[0093] 305、根据第二记忆矩阵和概率矩阵,获取第一目标语句矩阵。

[0094] 根据第二记忆矩阵和概率矩阵,获取第一目标语句矩阵的方式有多种,在一种可能实现方式中,参见图7,将概率矩阵与第二记忆矩阵相乘,得到与目标段落的语句矩阵尺寸相同的第一目标语句矩阵。

[0095] 其中,第一目标语句矩阵用于根据原始向量集合与记忆向量集合之间的关联关系,对目标段落进行描述。由于目标段落中的语句与上下文语句的相似度越高,概率越大,因此将概率矩阵与第二记忆矩阵相乘,可以将与上下文语句相似度较高的语句进行记忆强化,在后续处理过程中对该语句更加注意,相当于应用了注意力机制得到第一目标语句矩阵,使第一目标语句矩阵对该目标段落的描述更加准确。

[0096] 举例来说,原始向量集合中包括目标段落的J个语句的语句向量,记忆向量集合中包括上下文语句的K个语句的词向量,J和K为正整数,则原始向量集合对应的矩阵X为J*D的矩阵,记忆向量集合对应的矩阵M为K*D的矩阵,D为语句向量的维度数量。将这两个矩阵输入至记忆层,通过执行上述步骤302-305,得到的第一目标语句矩阵为 $O = \text{Softmax}(X \Phi_1(M)^T)$

$\Phi_2(M)$, 其中, $\Phi_1(M)$ 为第一记忆矩阵, $\Phi_2(M)$ 为第二记忆矩阵。

[0097] 306、在门控层中,应用线性网络模型,获取原始向量集合对应的线性数值,采用预设函数对线性数值进行处理,得到原始向量集合的第一权重,以使第一权重属于预设数值范围。

[0098] 输入层将原始向量集合输入至门控层,记忆层将第一目标语句矩阵输入至门控层,在门控层中根据原始向量集合和第一目标语句矩阵进行处理,对经过记忆强化的第一目标语句矩阵和原始的原始向量集合所占的比重进行调整,从而对目标段落中与上下文语句相似度较高的语句所占的比重进行调整。

[0099] 参见图8,在门控层中,应用线性网络模型,获取原始向量集合对应的线性数值,其中,线性网络模型可以是线性神经网络模型,或者还可以是其他线性网络模型,对原始向量集合进行线性处理后,得到的该线性数值能够对原始向量集合进行描述。

[0100] 获取到线性数值之后,采用预设函数对线性数值进行处理,得到原始向量集合的第一权重。该预设函数用于将线性数值压缩到预设数值范围,以使得到的第一权重属于该预设数值范围。其中,该预设函数可以是sigmoid(神经元的非线性作用)函数或者其他函数,该预设数值范围可以为0至1的数值范围,则第一权重大于0小于1。

[0101] 307、计算1与第一权重的差值,得到第一目标语句矩阵的第二权重。

[0102] 其中,第一权重是原始向量集合所占的权重,第二权重是第一目标语句矩阵所占的权重,第一权重与第二权重之和为1,在得到第一权重之后,通过计算1与第一权重的差值,得到第二权重。

[0103] 308、按照第一权重和第二权重,对原始向量集合与第一目标语句矩阵进行加权求和,得到第二目标语句矩阵。

[0104] 参见图9,根据该原始向量集合中的语句向量进行组合,得到目标段落的原始语句矩阵,第一权重即为该原始语句矩阵的权重,第二权重即为第一目标语句矩阵的权重,按照第一权重与第二权重,对该原始语句矩阵和该第一目标语句矩阵进行加权求和,得到第二目标语句矩阵,使第二目标语句矩阵中的每个数值属于预设数值范围。

[0105] 在一种可能实现方式中,采用以下公式进行加权求和:

$$O' = G * X + (1 - G) * O;$$

[0107] 其中, O' 为第二目标语句矩阵, G 为第一权重, X 为目标段落的原始语句矩阵, O 为第一目标语句矩阵。

[0108] 通过门控层可以筛选经过记忆加强后学习到的信息,调整目标段落与上下文语句之间的比重,控制信息的流动,避免加入过多与目标段落不相关的信息。

[0109] 309、在输出层中,根据第二目标语句矩阵,获取目标段落的段落向量。

[0110] 在输出层中,将第二目标语句矩阵转换为一个向量,作为目标段落的段落向量。其中,获取段落向量的方式可以有多种,在一种可能实现方式中,对第二目标语句矩阵进行列向求和,即将第二目标语句矩阵划分为多个列向量,计算每个列向量中的数值之和,得到每个列向量的总数值,将该多个列向量的总数值组成一个向量,得到段落向量。

[0111] 需要说明的一点是,参见图10,本发明实施例仅是以运行一次记忆层和门控层为例进行说明,而在另一实施例中,还可以重复运行记忆层和门控层。即在门控层中获取到第二目标语句矩阵后,将第二目标语句矩阵作为更新后的原始向量集合和记忆向量集合,在

记忆层和门控层中,重复执行根据更新后的原始向量集合和记忆向量集合获取目标语句矩阵的步骤,直至重复次数达到预设次数时,将当前的目标语句矩阵输入至输出层中,在输出层中,根据当前的目标语句矩阵,获取目标段落的段落向量。其中,该预设次数可以根据需求确定,或者为通过实验确定的优选值,该预设次数可以为2或3等。

[0112] 需要说明的另一点是,本发明实施例仅是以记忆编码模型包括门控层为例进行说明,而在另一实施例中,记忆编码模型不包括门控层时,不执行步骤305至308,当在记忆层中获取到第一目标语句矩阵后,将第一目标语句矩阵输入至输出层中,在输出层中根据第一目标语句矩阵,获取目标段落的段落向量。

[0113] 在一种可能实现方式中,对第一目标语句矩阵进行列向求和,得到一个向量,作为目标段落的段落向量。

[0114] 在记忆编码模型不包括门控层的情况下,还可以重复运行记忆层。即在记忆层中获取到第一目标语句矩阵后,将第一目标语句矩阵作为更新后的原始向量集合和记忆向量集合,在记忆层中,重复执行根据更新后的原始向量集合和记忆向量集合获取目标语句矩阵的步骤,直至重复次数达到预设次数时,将当前的目标语句矩阵输入至输出层中,在输出层中,根据当前的目标语句矩阵,获取目标段落的段落向量。其中,该预设次数可以根据需求确定,或者为通过实验确定的优选值,该预设次数可以为2或3等。

[0115] 310、基于段落向量进行处理。

[0116] 在获取到目标段落的段落向量之后,会对段落向量进行处理。应用场景不同,对段落向量的处理方式也不同,具体采用何种处理方式可以根据需求确定。例如:在智能对话场景下,目标段落为用户输入的文本消息,在获取到目标段落的段落向量之后,会根据该段落向量获取到匹配的答复消息,能够针对用户输入的文本消息给出符合用户需求的答复。

[0117] 本发明实施提供的编码方法,提供了一种记忆编码模型,记忆编码模型包括输入层、记忆层和输出层,获取目标段落和目标段落的上下文语句,将目标段落和上下文语句输入至记忆编码模型,在输入层中获取目标段落的原始向量集合以及目标段落的上下文语句中的记忆向量集合,在记忆层中根据原始向量集合和记忆向量集合,获取原始向量集合的第一目标语句矩阵,在输出层中根据第一目标语句矩阵获取目标段落的段落向量,基于段落向量进行处理。本发明实施例无需采用串行的方式对每个语句分别进行编码,而是以目标段落为单位,应用记忆编码模型对目标段落进行编码,因此提高了编码速度。并且,编码过程中不仅考虑目标段落本身,还考虑了目标段落的上下文语句,能够根据目标段落中的语句与上下文语句之间的关联关系获取到准确描述目标段落的段落向量,提高了编码准确率。

[0118] 本发明实施例提供的记忆编码模型具有自注意力性,将自注意力机制应用于段落的语句层面上,根据目标段落和上下文语句进行综合处理,可以保证目标段落的段落向量表达更为丰富,更能准确地描述目标段落的含义。且本发明实施例可以应用于多种场景下,应用范围广泛。

[0119] 本发明实施例提供了一种记忆编码模型的网络架构,应用该记忆编码模型可以对目标段落进行编码。且上述实施例提供的编码方法既可以应用于编码过程,也可以应用于训练记忆编码模型的过程。

[0120] 也即是,在一种可能实现方式中,在训练记忆编码模型的过程中,获取初始化的记

忆编码模型,或者获取已经进行过一次或多次训练、但其准确率还未满足要求的记忆编码模型。并且,获取一个或多个样本段落,作为目标段落。应用当前的记忆编码模型对目标段落进行处理,处理过程中执行上述实施例提供的编码方法,即可得到目标段落的段落向量。

[0121] 之后,将目标段落的段落向量进行解码,得到与该段落向量对应的测试段落,根据目标段落与测试段落之间的误差,对记忆编码模型中的模型参数进行修正。其中,解码方式可以有多种,如可以采用解码算法对段落向量进行解码,或者应用解码模型对段落向量进行解码,该解码模型可以为循环神经网络模型、深度学习网络模型、卷积神经网络模型等。

[0122] 则采用上述方式进行一次或多次训练之后,即可确定记忆编码模型中的模型参数,得到准确率满足要求的记忆编码模型。

[0123] 而在另一种可能实现方式中,记忆编码模型已经训练完成,其准确率满足要求。则获取该记忆编码模型,当要对某一目标段落进行编码时,应用该记忆编码模型对目标段落进行处理,处理过程中执行上述实施例提供的编码方法,即可得到目标段落的段落向量。其中,该记忆编码模型可以由编码设备训练,或者由训练设备训练后发送给编码设备,该训练设备也可以为终端或者服务器等。

[0124] 图11是本发明实施例提供的一种编码装置的结构示意图。参见图11,该装置包括:获取模块1100、输入层模块1101、记忆层模块1102和输出层模块1103;

[0125] 获取模块1100,用于获取目标段落和目标段落的上下文语句,将目标段落和上下文语句输入至记忆编码模型,目标段落包括至少一个语句;

[0126] 输入层模块1101,用于获取目标段落的原始向量集合和记忆向量集合,原始向量集合包括目标段落中每个语句的语句向量,记忆向量集合包括目标段落的上下文语句中每个词语的词向量;

[0127] 记忆层模块1102,用于根据原始向量集合和记忆向量集合,获取原始向量集合的第一目标语句矩阵,第一目标语句矩阵用于根据原始向量集合与记忆向量集合之间的关联关系,对目标段落进行描述;

[0128] 输出层模块1103,用于根据第一目标语句矩阵,获取目标段落的段落向量;

[0129] 处理模块1104,用于基于段落向量进行处理。

[0130] 本发明实施提供的编码装置,获取模块获取目标段落和目标段落的上下文语句,将目标段落和上下文语句输入至记忆编码模型,输入层模块获取目标段落的原始向量集合以及目标段落的上下文语句中的记忆向量集合,记忆层模块根据原始向量集合和记忆向量集合,获取原始向量集合的第一目标语句矩阵,输出层模块根据第一目标语句矩阵获取目标段落的段落向量,处理模块基于段落向量进行处理。本发明实施例无需采用串行的方式对每个语句分别进行编码,而是以目标段落为单位,应用编码装置对目标段落进行编码,因此提高了编码速度。并且,编码过程中不仅考虑目标段落本身,还考虑了目标段落的上下文语句,能够根据目标段落中的语句与上下文语句之间的关联关系获取到准确描述目标段落的段落向量,提高了编码准确率。

[0131] 本发明实施例提供的记忆编码模型具有自注意力性,将自注意力机制应用于段落的语句层面上,根据目标段落和上下文语句进行综合处理,可以保证目标段落的段落向量表达更为丰富,更能准确地描述目标段落的含义。且本发明实施例可以应用于多种场景下,应用范围广泛。

- [0132] 在一种可能实现方式中,输入层模块1101,包括:
- [0133] 第一获取单元,用于根据目标段落中每个语句中每个词语的词向量,应用语句编码模型,获取每个语句的语句向量,得到原始向量集合;
- [0134] 第二获取单元,用于根据上下文语句中每个词语的词向量,获取记忆向量集合。
- [0135] 在一种可能实现方式中,语句编码模型包括第一语句编码子模型和第二语句编码子模型;第一获取单元用于:
- [0136] 对目标段落中的每个语句,获取语句中每个词语的词向量,得到多个词向量;
- [0137] 应用第一语句编码子模型,对多个词向量进行正序编码,得到第一向量,应用第二语句编码子模型,对多个词向量进行倒序编码,得到第二向量;
- [0138] 根据第一向量和第二向量,获取语句的语句向量。
- [0139] 在一种可能实现方式中,记忆层模块1102,包括:
- [0140] 记忆获取单元,用于应用记忆模型,获取记忆向量集合对应的第一记忆矩阵和第二记忆矩阵;
- [0141] 第一目标获取单元,用于根据原始向量集合、第一记忆矩阵和第二记忆矩阵,获取原始向量集合的第一目标语句矩阵。
- [0142] 在一种可能实现方式中,第一目标获取单元用于获取原始向量集合与第一记忆矩阵的相似度矩阵;对相似度矩阵进行概率分布计算,得到概率矩阵;根据第二记忆矩阵和概率矩阵,获取第一目标语句矩阵。
- [0143] 在一种可能实现方式中,装置还包括门控层模块;
- [0144] 门控层模块,用于对原始向量集合与第一目标语句矩阵进行加权求和,得到第二目标语句矩阵,使第二目标语句矩阵中的每个数值属于预设数值范围;
- [0145] 输出层模块1103,还用于根据第二目标语句矩阵,获取目标段落的段落向量。
- [0146] 在一种可能实现方式中,门控层模块,包括:
- [0147] 第一权重获取单元,用于应用线性网络模型,获取原始向量集合对应的线性数值,采用预设函数对线性数值进行处理,得到原始向量集合的第一权重,以使第一权重属于预设数值范围;
- [0148] 第二权重获取单元,用于计算1与第一权重的差值,得到第一目标语句矩阵的第二权重;
- [0149] 加权单元,用于按照第一权重和第二权重,对原始向量集合与第一目标语句矩阵进行加权求和,得到第二目标语句矩阵。
- [0150] 在一种可能实现方式中,输出层模块1103,包括:
- [0151] 列向求和单元,用于对第一目标语句矩阵进行列向求和,得到段落向量。
- [0152] 在一种可能实现方式中,记忆层模块1102还用于将第一目标语句矩阵作为更新后的原始向量集合和记忆向量集合,重复执行根据更新后的原始向量集合和记忆向量集合获取目标语句矩阵的步骤,直至重复次数达到预设次数时,触发输出层模块1103;
- [0153] 输出层模块1103,还用于根据当前的目标语句矩阵,获取目标段落的段落向量。
- [0154] 需要说明的是:上述实施例提供的编码装置在对段落进行编码时,仅以上述各功能模块的划分进行举例说明,实际应用中,可以根据需要而将上述功能分配由不同的功能模块完成,即将编码设备的内部结构划分成不同的功能模块,以完成以上描述的全部或者

部分功能。另外,上述实施例提供的编码装置与编码方法实施例属于同一构思,其具体实现过程详见方法实施例,这里不再赘述。

[0155] 图12是本发明实施例提供的一种终端的结构框图。该终端1200用于执行上述实施例中编码设备执行的步骤,可以是便携式移动终端,比如:智能手机、平板电脑、MP3播放器(Moving Picture Experts Group Audio Layer III,动态影像专家压缩标准音频层面3)、MP4(Moving Picture Experts Group Audio Layer IV,动态影像专家压缩标准音频层面4)播放器、笔记本电脑或台式电脑。终端1200还可能被称为用户设备、便携式终端、膝上型终端、台式终端等其他名称。

[0156] 通常,终端1200包括有:处理器1201和存储器1202。

[0157] 处理器1201可以包括一个或多个处理核心,比如4核心处理器、8核心处理器等。处理器1201可以采用DSP(Digital Signal Processing,数字信号处理)、FPGA(Field-Programmable Gate Array,现场可编程门阵列)、PLA(Programmable Logic Array,可编程逻辑阵列)中的至少一种硬件形式来实现。处理器1201也可以包括主处理器和协处理器,主处理器是用于对在唤醒状态下的数据进行处理的处理单元,也称CPU(Central Processing Unit,中央处理器);协处理器是用于对在待机状态下的数据进行处理的低功耗处理器。在一些实施例中,处理器1201可以在集成有GPU(Graphics Processing Unit,图像处理器),GPU用于负责显示屏所需要显示的内容的渲染和绘制。一些实施例中,处理器1201还可以包括AI(Artificial Intelligence,人工智能)处理器,该AI处理器用于处理有关机器学习的计算操作。

[0158] 存储器1202可以包括一个或多个计算机可读存储介质,该计算机可读存储介质可以是非暂态的。存储器1202还可包括高速随机存取存储器,以及非易失性存储器,比如一个或多个磁盘存储设备、闪存存储设备。在一些实施例中,存储器1202中的非暂态的计算机可读存储介质用于存储至少一个指令,该至少一个指令用于被处理器1201所执行以实现本申请中方法实施例提供的编码方法。

[0159] 在一些实施例中,终端1200还可选包括有:外围设备接口1203和至少一个外围设备。处理器1201、存储器1202和外围设备接口1203之间可以通过总线或信号线相连。各个外围设备可以通过总线、信号线或电路板与外围设备接口1203相连。具体地,外围设备包括:射频电路1204、触摸显示屏1205、摄像头1206、音频电路1207、定位组件1208和电源1209中的至少一种。

[0160] 外围设备接口1203可被用于将I/O(Input/Output,输入/输出)相关的至少一个外围设备连接到处理器1201和存储器1202。在一些实施例中,处理器1201、存储器1202和外围设备接口1203被集成在同一芯片或电路板上;在一些其他实施例中,处理器1201、存储器1202和外围设备接口1203中的任意一个或两个可以在单独的芯片或电路板上实现,本实施例对此不加以限定。

[0161] 射频电路1204用于接收和发射RF(Radio Frequency,射频)信号,也称电磁信号。射频电路1204通过电磁信号与通信网络以及其他通信设备进行通信。射频电路1204将电信号转换为电磁信号进行发送,或者,将接收到的电磁信号转换为电信号。可选地,射频电路1204包括:天线系统、RF收发器、一个或多个放大器、调谐器、振荡器、数字信号处理器、编解码芯片组、用户身份模块卡等等。射频电路1204可以通过至少一种无线通信协议来与其它

终端进行通信。该无线通信协议包括但不限于：万维网、城域网、内联网、各代移动通信网络(2G、3G、4G及5G)、无线局域网和/或WiFi(Wireless Fidelity,无线保真)网络。在一些实施例中,射频电路1204还可以包括NFC(Near Field Communication,近距离无线通信)有关的电路,本申请对此不加以限定。

[0162] 显示屏1205用于显示UI(User Interface,用户界面)。该UI可以包括图形、文本、图标、视频及其它们的任意组合。当显示屏1205是触摸显示屏时,显示屏1205还具有采集在显示屏1205的表面或表面上方的触摸信号的能力。该触摸信号可以作为控制信号输入至处理器1201进行处理。此时,显示屏1205还可以用于提供虚拟按钮和/或虚拟键盘,也称软按钮和/或软键盘。在一些实施例中,显示屏1205可以为一个,设置终端1200的前面板;在另一些实施例中,显示屏1205可以为至少两个,分别设置在终端1200的不同表面或呈折叠设计;在再一些实施例中,显示屏1205可以是柔性显示屏,设置在终端1200的弯曲表面上或折叠面上。甚至,显示屏1205还可以设置成非矩形的不规则图形,也即异形屏。显示屏1205可以采用LCD(Liquid Crystal Display,液晶显示屏)、OLED(Organic Light-Emitting Diode,有机发光二极管)等材质制备。

[0163] 摄像头组件1206用于采集图像或视频。可选地,摄像头组件1206包括前置摄像头和后置摄像头。通常,前置摄像头设置在终端的前面板,后置摄像头设置在终端的背面。在一些实施例中,后置摄像头为至少两个,分别为主摄像头、景深摄像头、广角摄像头、长焦摄像头中的任意一种,以实现主摄像头和景深摄像头融合实现背景虚化功能、主摄像头和广角摄像头融合实现全景拍摄以及VR(Virtual Reality,虚拟现实)拍摄功能或者其它融合拍摄功能。在一些实施例中,摄像头组件1206还可以包括闪光灯。闪光灯可以是单色温闪光灯,也可以是双色温闪光灯。双色温闪光灯是指暖光闪光灯和冷光闪光灯的组合,可以用于不同色温下的光线补偿。

[0164] 音频电路1207可以包括麦克风和扬声器。麦克风用于采集用户及环境的声波,并将声波转换为电信号输入至处理器1201进行处理,或者输入至射频电路1204以实现语音通信。出于立体声采集或降噪的目的,麦克风可以为多个,分别设置在终端1200的不同部位。麦克风还可以是阵列麦克风或全向采集型麦克风。扬声器则用于将来自处理器1201或射频电路1204的电信号转换为声波。扬声器可以是传统的薄膜扬声器,也可以是压电陶瓷扬声器。当扬声器是压电陶瓷扬声器时,不仅可以将电信号转换为人类可听见的声波,也可以将电信号转换为人类听不见的声波以进行测距等用途。在一些实施例中,音频电路1207还可以包括耳机插孔。

[0165] 定位组件1208用于定位终端1200的当前地理位置,以实现导航或LBS(Location Based Service,基于位置的服务)。定位组件1208可以是基于GPS(Global Positioning System,全球定位系统)、北斗系统或格雷纳斯系统或伽利略系统的定位组件。

[0166] 电源1209用于为终端1200中的各个组件进行供电。电源1209可以是交流电、直流电、一次性电池或可充电电池。当电源1209包括可充电电池时,该可充电电池可以支持有线充电或无线充电。该可充电电池还可以用于支持快充技术。

[0167] 在一些实施例中,终端1200还包括有一个或多个传感器1210。该一个或多个传感器1210包括但不限于:加速度传感器1211、陀螺仪传感器1212、压力传感器1213、指纹传感器1214、光学传感器1215以及接近传感器1216。

[0168] 加速度传感器1211可以检测以终端1200建立的坐标系的三个坐标轴上的加速度大小。比如,加速度传感器1211可以用于检测重力加速度在三个坐标轴上的分量。处理器1201可以根据加速度传感器1211采集的重力加速度信号,控制触摸显示屏1205以横向视图或纵向视图进行用户界面的显示。加速度传感器1211还可以用于游戏或者用户的运动数据的采集。

[0169] 陀螺仪传感器1212可以检测终端1200的机体方向及转动角度,陀螺仪传感器1212可以与加速度传感器1211协同采集用户对终端1200的3D动作。处理器1201根据陀螺仪传感器1212采集的数据,可以实现如下功能:动作感应(比如根据用户的倾斜操作来改变UI)、拍摄时的图像稳定、游戏控制以及惯性导航。

[0170] 压力传感器1213可以设置在终端1200的侧边框和/或触摸显示屏1205的下层。当压力传感器1213设置在终端1200的侧边框时,可以检测用户对终端1200的握持信号,由处理器1201根据压力传感器1213采集的握持信号进行左右手识别或快捷操作。当压力传感器1213设置在触摸显示屏1205的下层时,由处理器1201根据用户对触摸显示屏1205的压力操作,实现对UI界面上的可操作性控件进行控制。可操作性控件包括按钮控件、滚动条控件、图标控件、菜单控件中的至少一种。

[0171] 指纹传感器1214用于采集用户的指纹,由处理器1201根据指纹传感器1214采集到的指纹识别用户的身份,或者,由指纹传感器1214根据采集到的指纹识别用户的身份。在识别出用户的身份为可信身份时,由处理器1201授权该用户执行相关的敏感操作,该敏感操作包括解锁屏幕、查看加密信息、下载软件、支付及更改设置等。指纹传感器1214可以被设置终端1200的正面、背面或侧面。当终端1200上设置有物理按键或厂商Logo时,指纹传感器1214可以与物理按键或厂商标志集成在一起。

[0172] 光学传感器1215用于采集环境光强度。在一个实施例中,处理器1201可以根据光学传感器1215采集的环境光强度,控制触摸显示屏1205的显示亮度。具体地,当环境光强度较高时,调高触摸显示屏1205的显示亮度;当环境光强度较低时,调低触摸显示屏1205的显示亮度。在另一个实施例中,处理器1201还可以根据光学传感器1215采集的环境光强度,动态调整摄像头组件1206的拍摄参数。

[0173] 接近传感器1216,也称距离传感器,通常设置在终端1200的前面板。接近传感器1216用于采集用户与终端1200的正面之间的距离。在一个实施例中,当接近传感器1216检测到用户与终端1200的正面之间的距离逐渐变小时,由处理器1201控制触摸显示屏1205从亮屏状态切换为息屏状态;当接近传感器1216检测到用户与终端1200的正面之间的距离逐渐变大时,由处理器1201控制触摸显示屏1205从息屏状态切换为亮屏状态。

[0174] 本领域技术人员可以理解,图12中示出的结构并不构成对终端1200的限定,可以包括比图示更多或更少的组件,或者组合某些组件,或者采用不同的组件布置。

[0175] 图13是本发明实施例提供的一种服务器的结构示意图,该服务器1300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处理器(central processing units,CPU)1301和一个或一个以上的存储器1302,其中,存储器1302中存储有至少一条指令,至少一条指令由处理器1301加载并执行以实现上述各个方法实施例提供的方法。当然,该服务器还可以具有有线或无线网络接口、键盘以及输入输出接口等部件,以便进行输入输出,该服务器还可以包括其他用于实现设备功能的部件,在此不做赘述。

[0176] 服务器1300可以用于执行上述编码方法中编码设备所执行的步骤。

[0177] 本发明实施例还提供了一种编码设备,该编码设备包括一个或多个处理器和一个或多个存储器,该一个或多个存储器中存储有至少一条指令、至少一段程序、代码集或指令集,该指令、该程序、该代码集或该指令集由一个或多个处理器加载并执行以实现上述实施例的编码方法中所执行的操作。

[0178] 本发明实施例还提供了一种计算机可读存储介质,该计算机可读存储介质中存储有至少一条指令、至少一段程序、代码集或指令集,该指令、该程序、该代码集或该指令集由一个或多个处理器加载并执行以实现上述实施例的编码方法中所执行的操作。

[0179] 本领域普通技术人员可以理解实现上述实施例的全部或部分步骤可以通过硬件来完成,也可以通过程序来指令相关的硬件完成,所述的程序可以存储于一种计算机可读存储介质中,上述提到的存储介质可以是只读存储器,磁盘或光盘等。

[0180] 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

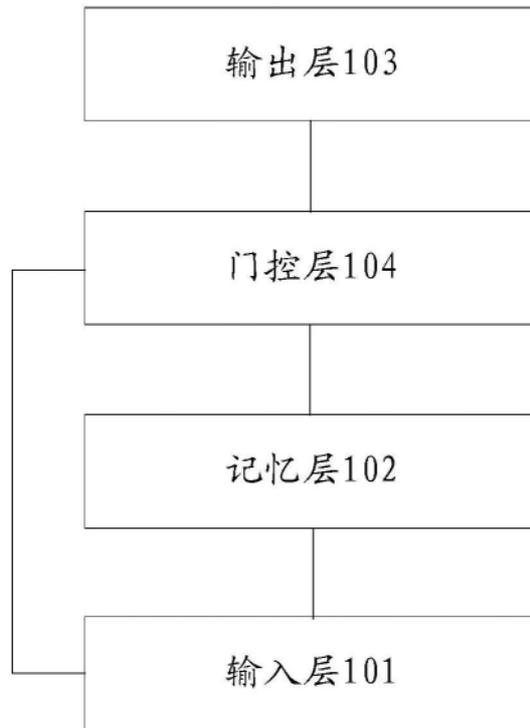


图2

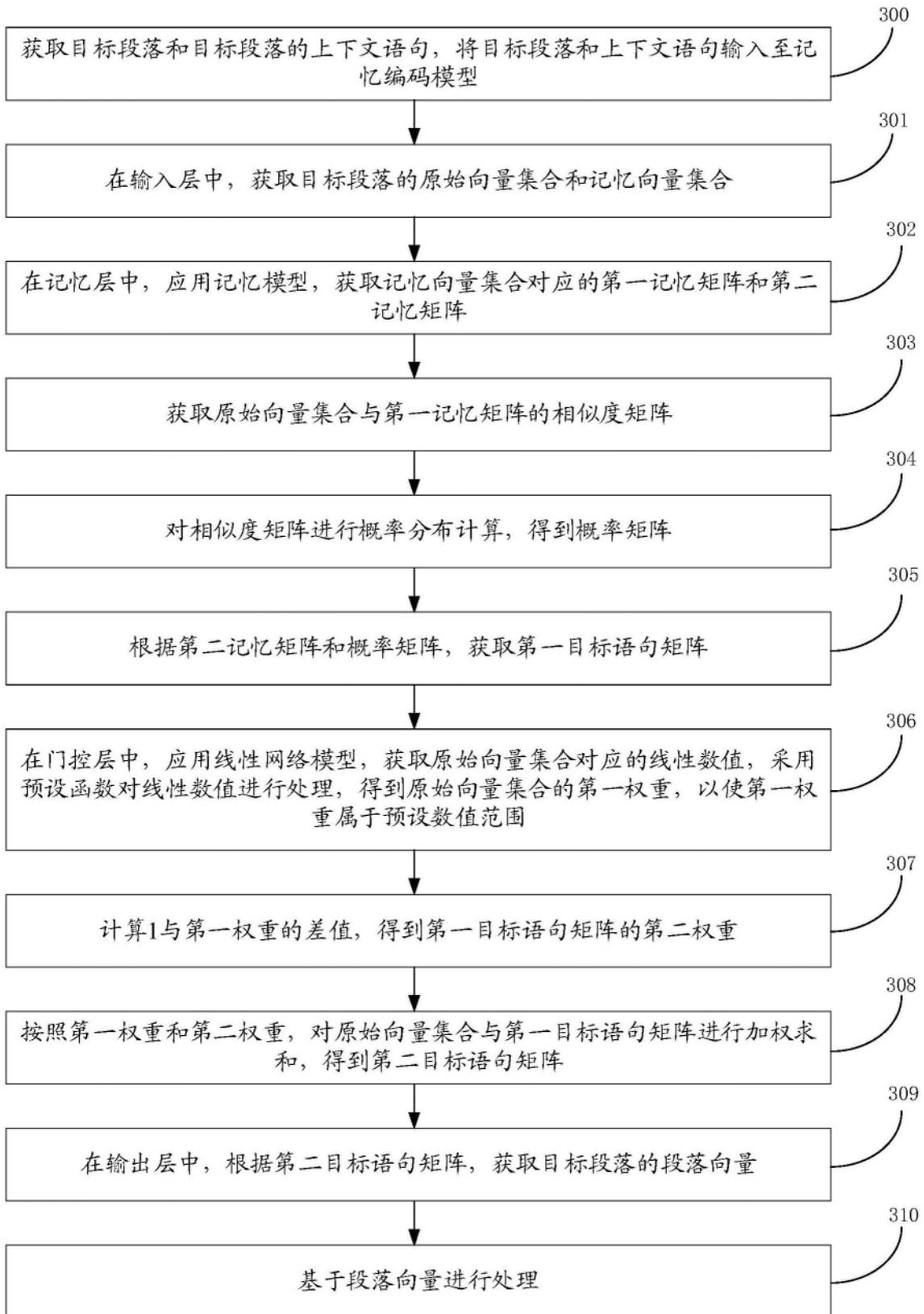


图3

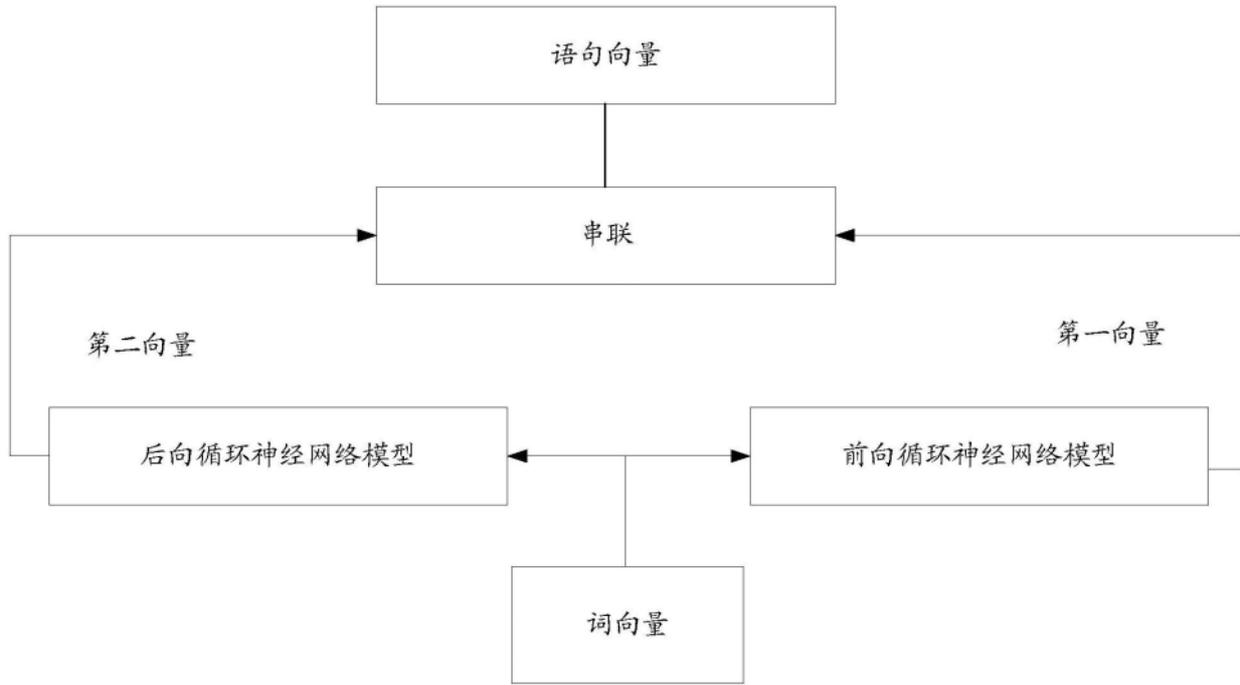


图4

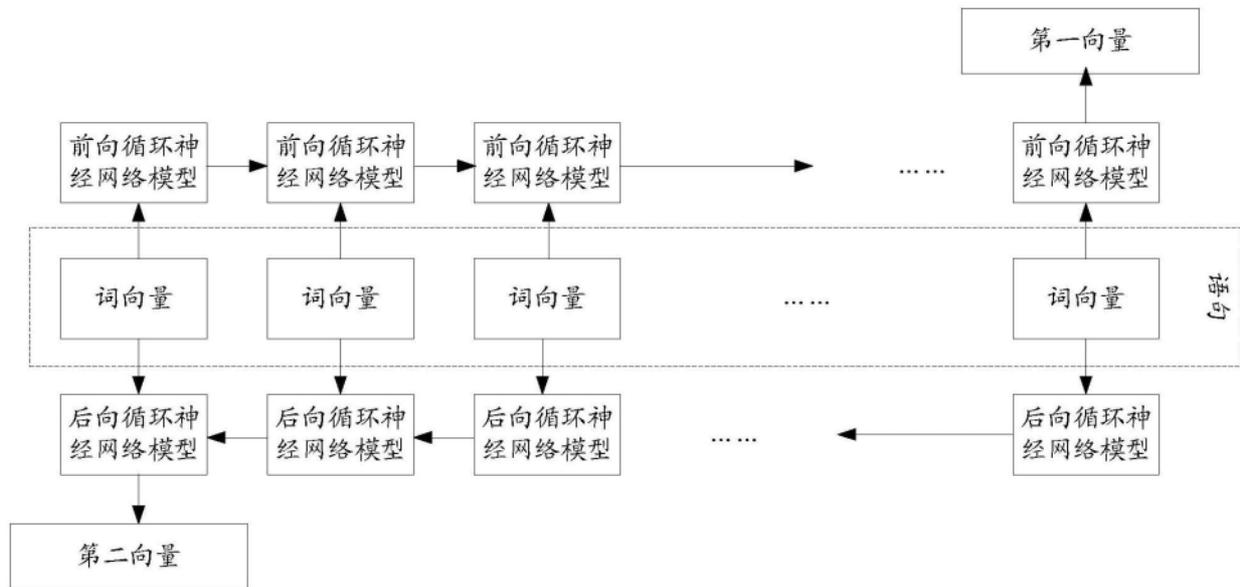


图5

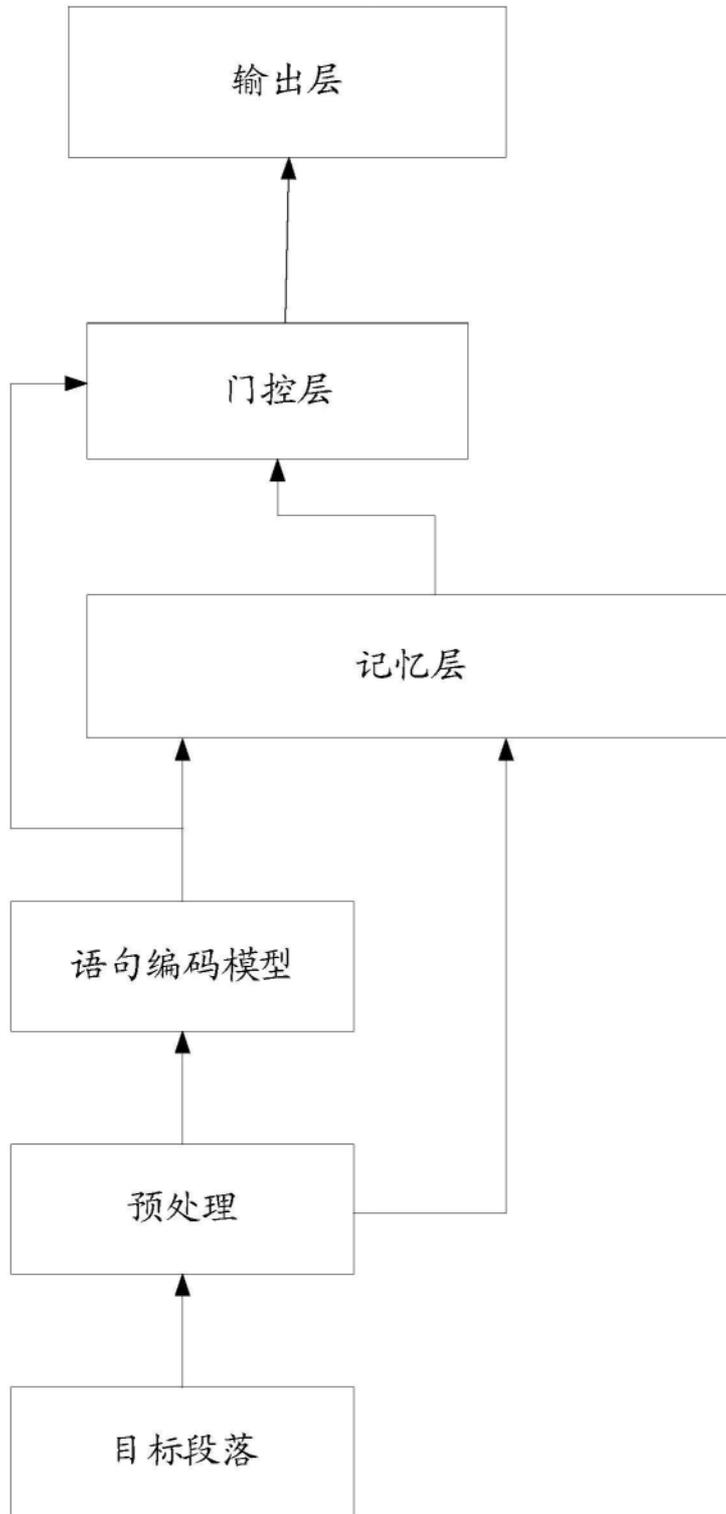


图6

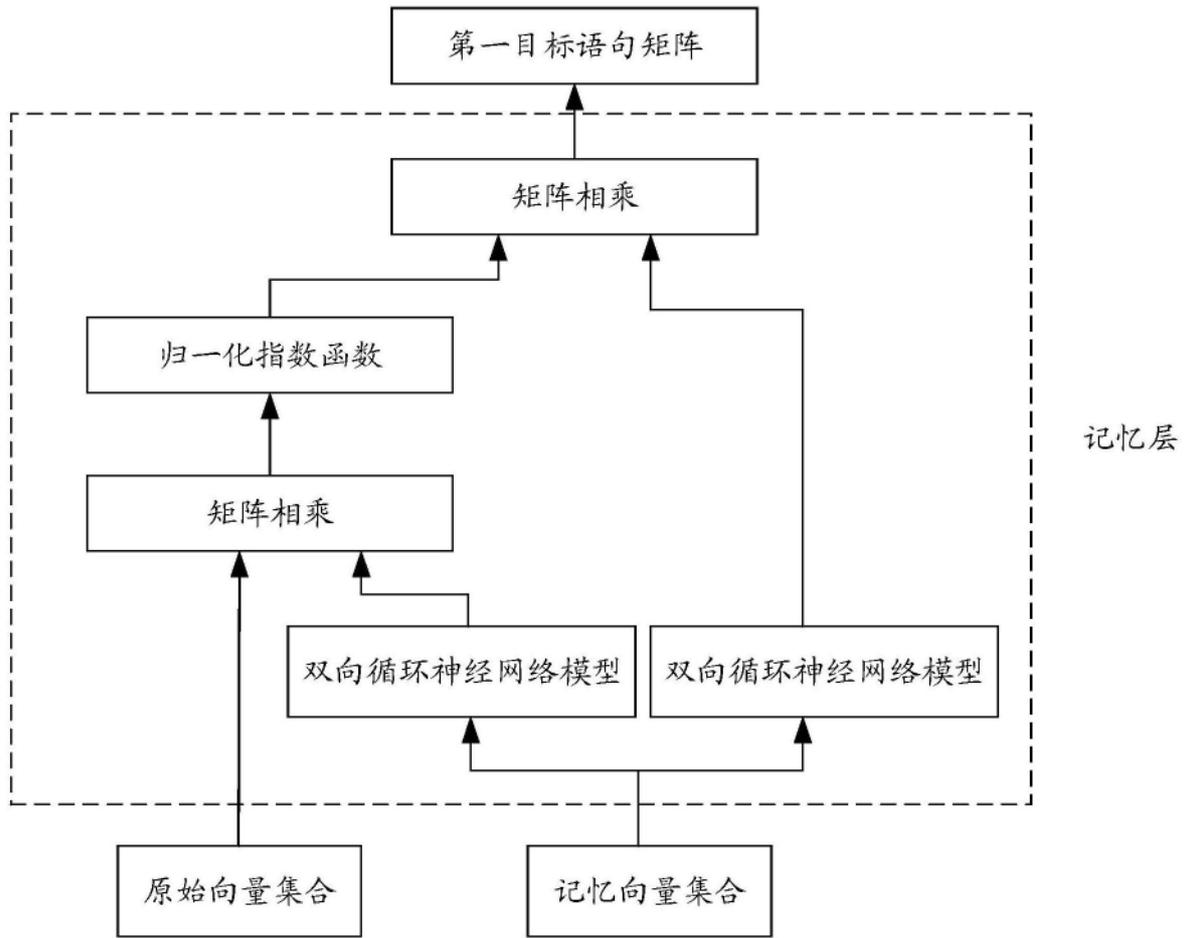


图7

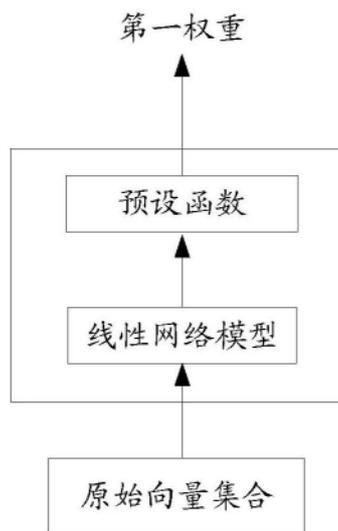


图8

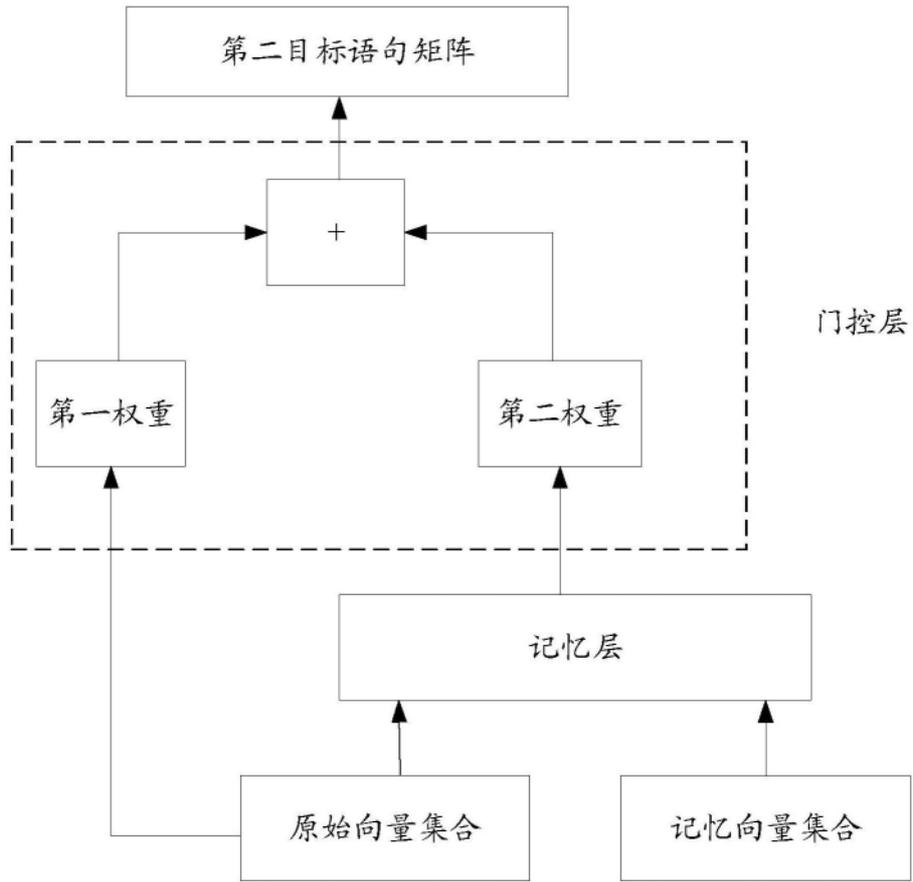


图9

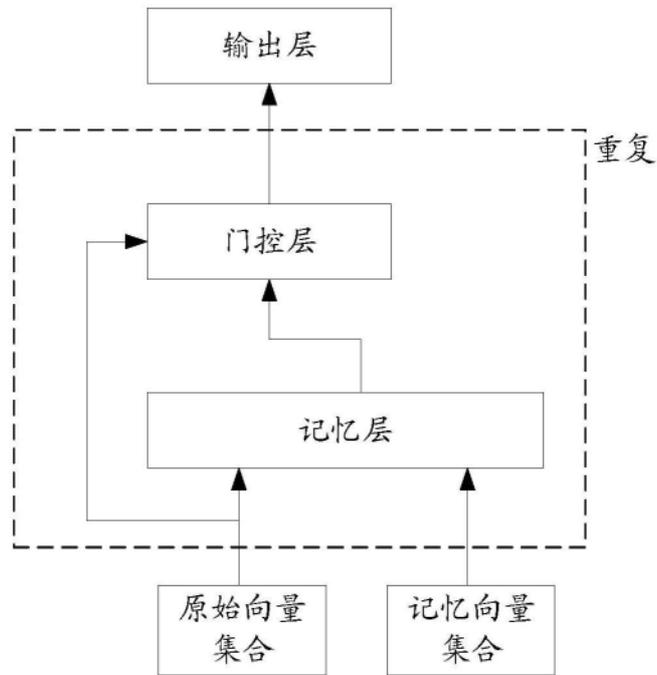


图10

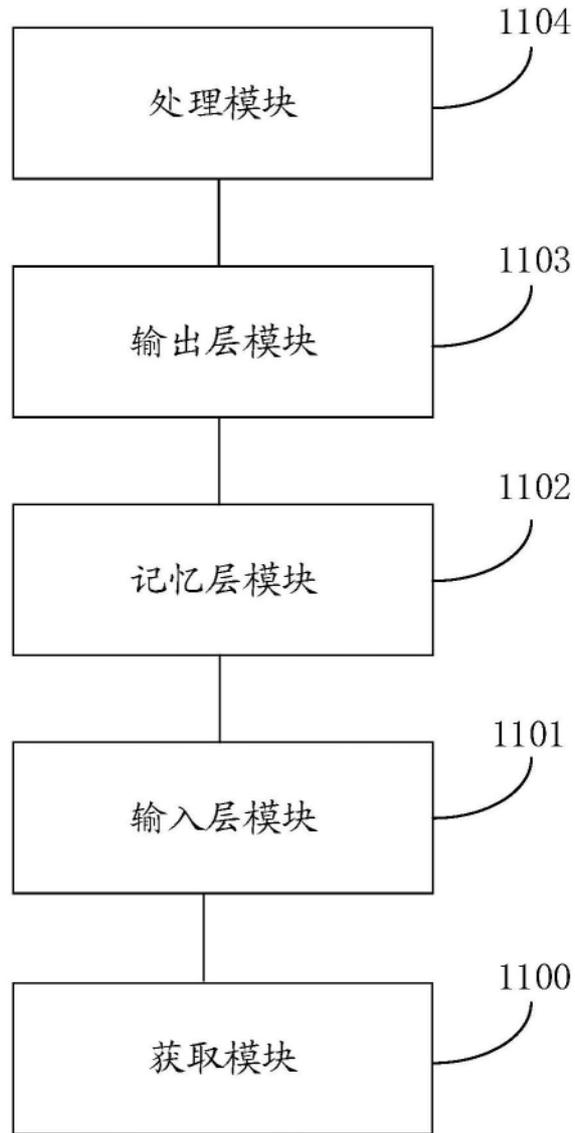


图11

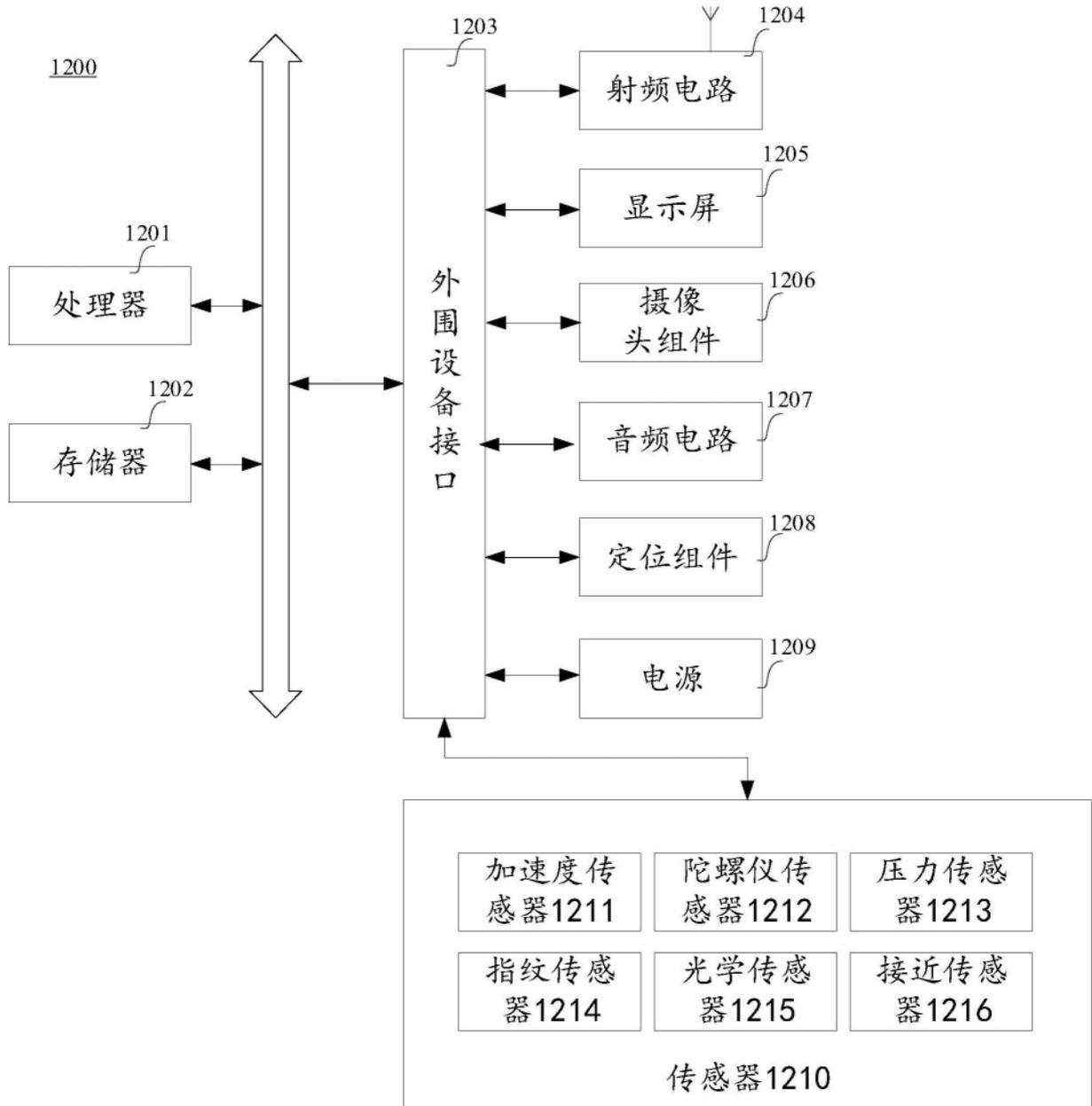


图12

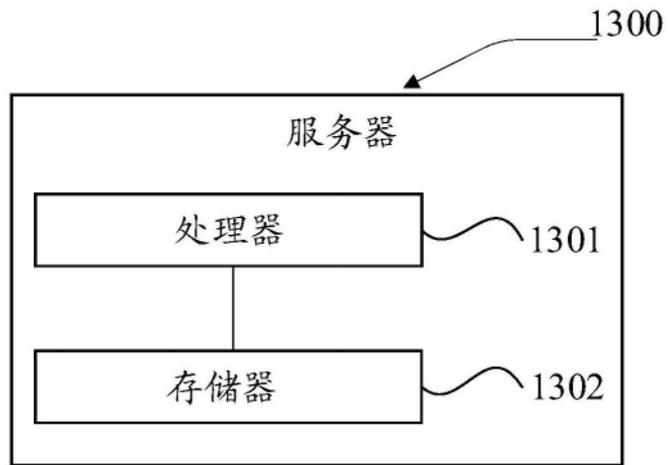


图13