

(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
2021 年 9 月 10 日 (10.09.2021)



(10) 国际公布号
WO 2021/175005 A1

(51) 国际专利分类号:
G06F 16/33 (2019.01)

(21) 国际申请号: PCT/CN2021/070585

(22) 国际申请日: 2021 年 1 月 7 日 (07.01.2021)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202010143243.6 2020 年 3 月 4 日 (04.03.2020) CN

(71) 申请人: 深圳壹账通智能科技有限公司(ONE CONNECT SMART TECHNOLOGY CO., LTD. (SHENZHEN)) [CN/CN]; 中国广东省深圳市前海深港合作区前湾一路 1 号 A 栋 201 室, Guangdong 518000 (CN)。

(72) 发明人: 王盼(WANG, Pan); 中国广东省深圳市前海深港合作区前湾一路 1 号 A 栋 201 室, Guangdong 518000 (CN)。

(74) 代理人: 深圳众鼎专利商标代理事务所(普通合伙)(SHENZHEN ZHONGDING INTELLECTUAL PROPERTY AGENCY); 中国广东省深圳市罗湖区笋岗街道笋岗东路 3012 号中民时代广场 B 座 701, Guangdong 518000 (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, IT, 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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, 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,

(54) **Title:** VECTOR-BASED DOCUMENT RETRIEVAL METHOD AND APPARATUS, COMPUTER DEVICE, AND STORAGE MEDIUM

(54) 发明名称: 基于向量的文档检索方法、装置、计算机设备及存储介质

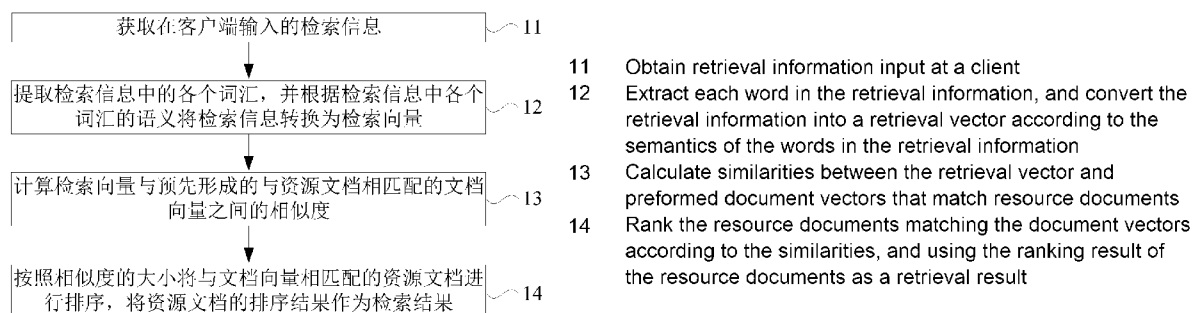


图 2

(57) **Abstract:** A vector-based document retrieval method and apparatus, a computer device, and a storage medium. The vector-based document retrieval method comprises: obtaining retrieval information input at a client (11); extracting each word in the retrieval information, and converting the retrieval information into a retrieval vector according to the semantics of the words in the retrieval information (12); calculating similarities between the retrieval vector and preformed document vectors that match resource documents (13); and ranking the resource documents matching the document vectors according to the similarities, and using the ranking result of the resource documents as a retrieval result (14). The implementation of the present technical solution, the problems of low retrieval accuracy and high retrieval difficulty using the document retrieval methods in the prior art can be solved.

(57) **摘要:** 一种基于向量的文档检索方法、装置、计算机设备及存储介质, 该基于向量的文档检索方法包括: 获取在客户端输入的检索信息(11); 提取检索信息中的各个词汇, 并根据检索信息中各个词汇的语义将检索信息转换为检索向量(12); 计算检索向量与预先形成的与资源文档相匹配的文档向量之间的相似度(13); 按照相似度的大小将与文档向量相匹配的资源文档进行排序, 将资源文档的排序结果作为检索结果(14)。通过该技术方案实施, 能够解决现有技术中的文档检索方法存在检索准确度较低、检索难度较大的问题。

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))。

基于向量的文档检索方法、装置、计算机设备及存储介质

[0001] 本申请要求于2020年03月04日提交中国专利局、申请号为202010143243.6，发明名称为“基于向量的文档检索方法、装置、计算机设备及存储介质”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中。

技术领域

[0002] 本申请涉及信息检索技术领域，尤其涉及一种基于向量的文档检索方法、装置、计算机设备及存储介质。

背景技术

[0003] 随着信息技术的不断发展，各行各业产生的信息量越来越大，传统的检索方式逐渐不能满足人们日常的检索需求。为了能够快速准确地获取到检索结果，需要对传统的文档检索方法进行改进，以使用户能够快速获得检索结果。目前，现有的文档检索方法一般是先建立包含第一关键词和逻辑运算符的布尔表达式，然后利用布尔表达式所表示的逻辑关系和第一关键词对文档进行检索。

[0004] 虽然通过上述文档检索方法能够获得检索结果，但由于布尔逻辑式的构造不易全面准确反映用户的需求，对用户的检索技能有较高要求，同时仅根据用户输入的第一关键词的文字表意进行检索，使得检索的结果准确度较低、检索难度较大。

[0005] 综上所述，发明人意识到现有技术中的文档检索方法存在检索准确度较低、检索难度较大的问题。

发明概述

技术问题

[0006] 本申请提供一种基于向量的文档检索方法、装置、计算机设备及存储介质，以解决现有的基于向量的文档检索方法存在特征数据提取难度较大、分类精确度不高的问题。

问题的解决方案

技术解决方案

- [0007] 本申请的第一实施例提供一种基于向量的文档检索方法，包括：
- [0008] 获取在客户端输入的检索信息；
- [0009] 提取检索信息中的各个词汇，并根据检索信息中各个词汇的语义将检索信息转换为检索向量；
- [0010] 计算检索向量与预先形成的与资源文档相匹配的文档向量之间的相似度；
- [0011] 按照相似度的大小将与文档向量相匹配的资源文档进行排序，将资源文档的排序结果作为检索结果。
- [0012] 本申请的第二实施例提供一种基于向量的文档检索装置，包括：
- [0013] 检索信息获取模块，用于获取在客户端输入的检索信息；
- [0014] 检索向量获取模块，用于提取检索信息中的各个词汇，并根据检索信息中各个词汇的语义将检索信息转换为检索向量；
- [0015] 相似度获取模块，用于计算检索向量与预先形成的与资源文档相匹配的文档向量之间的相似度；
- [0016] 检索结果获取模块，用于按照相似度的大小将与文档向量相匹配的资源文档进行排序，将资源文档的排序结果作为检索结果。
- [0017] 本申请的第三实施例提供一种计算机设备，包括存储器、处理器以及存储在存储器中并可在处理器上运行的计算机可读指令，处理器执行计算机可读指令时实现如下步骤：
- [0018] 获取在客户端输入的检索信息；
- [0019] 提取所述检索信息中的各个词汇，并根据所述检索信息中各个词汇的语义将所述检索信息转换为检索向量；
- [0020] 计算所述检索向量与预先形成的与资源文档相匹配的文档向量之间的相似度；
- [0021] 按照所述相似度的大小将与所述文档向量相匹配的所述资源文档进行排序，将所述资源文档的排序结果作为检索结果。
- [0022] 本申请的第四实施例提供一种一个或多个存储有计算机可读指令的可读存储介质，所述计算机可读存储介质存储有计算机可读指令，其中，所述计算机可读

指令被一个或多个处理器执行时，使得所述一个或多个处理器执行如下步骤：

- [0023] 获取在客户端输入的检索信息；
- [0024] 提取所述检索信息中的各个词汇，并根据所述检索信息中各个词汇的语义将所述检索信息转换为检索向量；
- [0025] 计算所述检索向量与预先形成的与资源文档相匹配的文档向量之间的相似度；
- [0026] 按照所述相似度的大小将与所述文档向量相匹配的所述资源文档进行排序，将所述资源文档的排序结果作为检索结果。

发明的有益效果

有益效果

- [0027] 本申请提供的基于向量的文档检索方法、装置、计算机设备及存储介质中，首先获取在客户端输入的检索信息，再提取检索信息中的各个词汇，并根据检索信息中各个词汇的语义将检索信息转换为检索向量，然后计算检索向量与预先形成的与资源文档相匹配的文档向量之间的相似度，最后按照相似度的大小将与文档向量相匹配的资源文档进行排序，将资源文档的排序结果作为检索结果。提取检索信息中的各个词汇，并根据检索信息中的各个词汇的语义将检索信息转换成检索向量，通过本申请的实施，能够解决现有技术中的文档检索方法存在检索准确度较低、检索难度较大的问题。

对附图的简要说明

附图说明

- [0028] 为了更清楚地说明本申请实施例的技术方案，下面将对本申请实施例的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申请的一些实施例，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。
- [0029] 图1是本申请的第一实施例的基于向量的文档检索方法的一应用环境示意图；
- [0030] 图2是本申请的第一实施例的基于向量的文档检索方法的流程图；
- [0031] 图3是本申请的第一实施例的基于向量的文档检索方法中步骤12的流程图；
- [0032] 图4是本申请的第一实施例的基于向量的文档检索方法中步骤122的流程图；
- [0033] 图5是本申请的第一实施例的基于向量的文档检索方法中步骤1222的流程图；

[0034] 图6是本申请的第一实施例的基于向量的文档检索方法的又一流程图；

[0035] 图7是本申请的第二实施例的基于向量的文档检索装置的模块示意图；

[0036] 图8是本申请的第二实施例的基于向量的文档检索装置的又一模块示意图；

[0037] 图9是本申请的第二实施例的基于向量的文档检索装置的又一模块示意图；

[0038] 图10是本申请的第三实施例的计算机设备的一模块示意图。

[0039] 下面将结合本申请实施例中的附图，对本申请实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例是本申请一部分实施例，而不是全部的实施例。基于本申请中的实施例，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，都属于本申请保护的范围。

[0040] 本申请的第一实施例提供的基于向量的文档检索方法，可应用于如图1所示的应用环境中，其中，客户端（计算机设备）通过网络与服务端进行通信。服务器获取在客户端输入的检索信息，提取检索信息中的各个词汇，并根据检索信息中各个词汇的语义将检索信息转换为检索向量，计算检索向量与预先形成的与资源文档相匹配的文档向量之间的相似度，按照相似度的大小将与文档向量相匹配的资源文档进行排序，将资源文档的排序结果作为检索结果，并将检索结果发送至客户端。其中，客户端（计算机设备）可以但不限于各种个人计算机、笔记本电脑、智能手机、平板电脑和便携式可穿戴设备。服务端可以用独立的服务器或者是多个服务器组成的服务器集群来实现。

[0041] 在本申请的第一实施例中，如图2所示，提供一种基于向量的文档检索方法，以该方法应用在图1中的服务端为例进行说明，包括如下步骤11至步骤15。

[0042] 步骤11：获取在客户端输入的检索信息。

[0043] 其中，检索信息包含用户想要获得目标文档而输入的指定的与目标文档相关的信息。

[0044] 步骤12：提取检索信息中的各个词汇，并根据检索信息中各个词汇的语义将检索信息转换为检索向量。

[0045] 其中，一条检索信息对应一个检索向量。

[0046] 进一步地，作为本实施例的一种实施方式，如图3所示，上述步骤12具体包括以下步骤121至步骤124：

[0047] 步骤121：获取检索信息所包含的各个词汇。

[0048] 其中，上述步骤121包括：先对检索信息进行分词处理，然后去掉检索信息中的停用词。具体是先对检索信息中的各个文字进行隔离，在词典中查询各相邻文字之间是否能够组成词组，若能够组成词组，则将该相邻文字组成词组，若不能组成词组，则将该相邻词组进行隔离，然后将未形成词组的问或符号作为停用词，将形成的各个词组作为检索信息中的各个词汇。需要注意的是，相邻文字可以是连续两个相邻的文字，也可以是连续三个相邻的文字，此处不做具体限制。

[0049] 为能够更加清楚地理解上述步骤121，列举示例：检索信息为“探索宇宙的奥秘”，先对各个检索信息中的各个文字进行隔离得到“探/索/宇/宙/的/奥/秘”，在词典中查询各相邻文字之间是否能够组成词组，得到“探索/宇宙/的/奥秘”，将词组“探索”、“宇宙”、“奥秘”作为检索信息中的各个词汇，由于“的”并未与检索信息中的其他文字形成词组，将“的”作为停用词。

[0050] 步骤122：将各个词汇中语义相近的词汇分别用同一第一关键词进行表示。

[0051] 其中，具体是将词汇中语义相近的词汇用同一关键词进行表示。当词汇的数量为多个时，应当查询获得每个词汇的近义词，以使多个词汇能够用同一第一关键词进行表示。

[0052] 步骤123：统计各个第一关键词出现的次数。

[0053] 具体可以通过TF-IDF算法（term frequency - inverse document frequency，词频统计算法）统计第一关键词出现的次数，并对各个第一关键词出现的次数进行标记。

[0054] 为例能够更加清楚地理解上述步骤122至步骤123，列举示例：其中，具体是按照各个词汇的语义获得各个词汇的第一关键词，统计各个第一关键词出现的次数。例如，存在词汇“开心”出现3次、“快乐”出现4次、“高兴”出现5次，词汇“开心”、“快乐”、“高兴”的近义词均可以为“愉快”，此时，将该近义词“愉快”作为第一关键词，用该第一关键词“愉快”来代表词汇“开心”、“快乐”、“高兴”，第一关键词“愉快”出现的次数为词汇“开心”、“快乐”、“高兴”出现次数之和12次，同时对第一关键词“愉快”进行标记

，标记次数为12。

[0055] 步骤124：将各个第一关键词和各个第一关键词出现的次数映射到向量词典，以得到检索向量。

[0056] 其中，所述检索向量是一个M维向量，M代表词典中第一关键词的个数。向量词典中的各个词汇采用树状层次的结构进行存储，在该结构中，每个词汇之间通过节点的形式连接，每个词汇之间相差的节点数量越少，表示每个词汇之间的相似度越高。另外，向量词典中的各个词汇展现方式可以是三维矩阵的形式，具体而言，向量词典具体是由各个词汇和各个词汇出现的频率组成。

[0057] 需要注意的是，在本方案中，向量词典中每个第一关键词的数量仅有一个，也就是说，在该向量词典中，各个第一关键词均各不相同。例如在向量词典中，当第一关键词A出现的次数为5，则在向量词典形成的检索向量中给关键词A赋一个值5，当关键词B出现的次数为0，则在空间向量形成的检索向量中给关键词B赋一个值0。

[0058] 步骤13：计算检索向量与预先形成的与资源文档相匹配的文档向量之间的相似度。

[0059] 其中，文档向量包含资源文档中的信息，用于代表资源文档，一个文档向量代表一个资源文档。

[0060] 具体是通过以下公式（1）计算获得检索向量与预先形成的与资源文档相匹配的文档向量之间的相似度：

$$[0061] \quad \cos \theta = (a \cdot b) / (|a| \times |b|) \quad (1)$$

[0062] 其中， $\cos \theta$ 代表检索向量与文档向量的相似度， θ 代表检索向量与文档向量之间的夹角， a 代表检索向量， b 代表文档向量， $\cdot$ 代表向量的点乘， $|a|$ 代表检索向量的模， $|b|$ 代表文档向量的模。

[0063] 需要注意的是，当存在多个资源文档时，文档向量也应当存在多个，按照上述公式能够分别计算出各个文档向量与检索向量的相似度。

[0064] 步骤14：按照相似度的大小将与文档向量相匹配的资源文档进行排序，将资源文档的排序结果作为检索结果。

[0065] 其中，相似度的数值越大，则代表文档向量与检索向量越相似。具体是按照相

似度由大到小的顺序将与文档向量相匹配的资源文档进行排序后发送至客户端。

[0066] 另外，在本实施例中，也可以将达到预设阈值的相似度对应的资源文档发送至客户端。

[0067] 通过上述步骤11至步骤14的实施，能够将检索信息按照语义转换为检索向量，并计算检索向量与用于表示资源文档的文档向量的相似度，按照相似度的大小对资源文档进行排序，无需使用检索表达式进行检索，并按照检索信息中词汇的语义对资源文档进行检索，以防止检索得到的资源文档仅满足与检索信息中的词汇文字高度吻合这一条件，而忽略资源文档和检索信息中的词汇语义，降低了检索的难度的同时，提高了对资源文档检索的精准度。

[0068] 进一步地，作为本实施例的一种实施方式，如图4所示，上述步骤122具体包括以下步骤1221至步骤1223：

[0069] 步骤1221：从同义词词林中获取与各个词汇相匹配的至少一个同义词。

[0070] 其中，同义词可以是在同义词词林中与各个词汇有任意关联的词汇。

[0071] 步骤1222：计算各个词汇与对应匹配的同义词的语义相似度。

[0072] 其中，具体是分别计算各个词汇与对应的各个同义词之间的语义相似度。

[0073] 步骤1223：当词汇与同义词的相似度达到预设第一阈值时，将同义词作为与对应词汇相匹配的第一关键词。

[0074] 通过上述步骤1221至步骤1223的实施，能够根据语义从同义词中获得第一关键词。

[0075] 进一步地，作为本实施例的一种实施方式，如图5所示，上述步骤1222具体包括以下步骤12221至步骤12223：

[0076] 步骤12221：根据词汇获取第一语义信息，根据同义词获取第二语义信息；

[0077] 步骤12222：从第一语义信息中获取第一语义关键词，以形成第一数据集，从第二语义信息中获取第二语义关键词，以形成第二数据集；

[0078] 步骤12223：计算第一数据集和第二数据集之间的相似度，并将计算的相似度作为语义相似度。

[0079] 对于上述步骤12221，具体是通过在词典或者搜索引擎中搜索词汇和同义词，

获得与词汇和同义词对应的搜索反馈结果，将与词汇对应的反馈结果作为第一语义信息，将与同义词对应的反馈结果作为第二语义信息。具体地，反馈结果一般是对词汇和同义词的解释内容。

[0080] 对于上述步骤12222，具体抽取第一语义信息中的关键词作为第一语义关键词，抽取第二语义信息中的关键词作为第二语义信息。

[0081] 对于上述步骤12223，具体是计算第一数据集中的各个词汇和第二数据集中的各个词汇之间的相似度，可以将第一数据集中的各个词汇和第二数据集中的各个词汇之间的最大相似度作为语义相似度，也可以将第一数据集中的各个词汇和第二数据集中的各个词汇的相似度平均值作为语义相似度。

[0082] 通过上述步骤12221至步骤12223的实施，能够根据各个词汇和同义词的语义判断各个词汇和同义词之间的语义相似度，有利于实现根据语义对资源文档进行检索，提高了检索的精度。

[0083] 进一步地，作为本实施例的一种实施方式，需要将资源文档转换成可供检索的文档向量，如图6所示，获取预先形成的与资源文档相匹配的文档向量具体包括以下步骤21至步骤24：

[0084] 步骤21：从资源文档中获取各个资源词汇。

[0085] 其中，资源文档代表记载信息的载体。获取资源文档的方式可以是利用爬虫技术爬取网站网页内的信息形成的文档，也可以是利用文字识别技术获取文本信息形成的文档，此处不做具体限制。

[0086] 步骤22：将各个资源词汇中语义相近的资源词汇分别用第二关键词表示。

[0087] 其中，具体是先提取资源文档中所有的资源字段，资源字段为资源文档中存在的文字、字符等，提取出来的资源字段的顺序应当与资源文档中原有的资源字段的顺序相同。

[0088] 另外，由于步骤22中的将各个资源词汇中语义相近的资源词汇分别用第二关键词表示的方法与上述步骤122中将各个词汇中语义相近的词汇分别用同一第一关键词进行表示的方法相同，此处不再赘述。

[0089] 步骤23：统计各个第二关键词出现的次数。

[0090] 其中，具体可以通过TF-IDF算法统计第二关键词出现的次数，并对各个第二关

关键词出现的次数进行标记。

[0091] 步骤24：将各个第二关键词和各个第二关键词出现的次数映射到向量词典，以得到文档向量。

[0092] 其中，由于步骤24中将各个第二关键词和各个第二关键词出现的次数映射到向量词典的方法与上述步骤124中将各个第一关键词和各个第一关键词出现的次数映射到向量词典的方法相同，此处不再赘述。

[0093] 通过上述步骤21至步骤24的实施，能够实现将资源文档转换成文档向量，以便于在上述步骤11至步骤14过程中计算文档向量与检索向量的相似度，获得与文档向量对应的资源文档。

[0094] 需要注意的是，在本实施例中，应当不对检索信息中的内容进行限定，检索信息也可以是文档，当检索信息为文档时，根据上述步骤11至步骤14的方法能够获得与文档相匹配的资源文档，以获得与文档最相似的资源文档。

[0095] 应理解，上述实施例中各步骤的序号的大小并不意味着执行顺序的先后，各过程的执行顺序应以其功能和内在逻辑确定，而不对本申请实施例的实施过程构成任何限定。

[0096] 本申请的第二实施例提供一种基于向量的文档检索装置，该基于向量的文档检索装置与上述第一实施例提供的基于向量的文档检索方法一一对应。

[0097] 进一步地，如图7所示，该基于向量的文档检索装置包括检索信息获取模块41、检索向量获取模块42、相似度获取模块43和检索结果获取模块44。各功能模块详细说明如下：

[0098] 检索信息获取模块41，用于获取在客户端输入的检索信息；

[0099] 检索向量获取模块42，用于提取检索信息中的各个词汇，并根据检索信息中各个词汇的语义将检索信息转换为检索向量；

[0100] 相似度获取模块43，用于计算检索向量与预先形成的与资源文档相匹配的文档向量之间的相似度；

[0101] 检索结果获取模块44，用于按照相似度的大小将与文档向量相匹配的资源文档进行排序，将资源文档的排序结果作为检索结果。

[0102] 进一步地，作为本实施例的一种实施方式，如图8所示，检索向量获取模块42

包括分词处理单元421、第一关键词获取单元422、统计单元423和检索向量获取单元424。各功能单元详细功能如下：

[0103] 分词处理单元421，用于获取检索信息所包含的各个词汇；

[0104] 第一关键词获取单元422，用于将各个词汇中语义相近的词汇分别用同一第一关键词进行表示；

[0105] 统计单元423，用于统计各个第一关键词出现的次数；

[0106] 检索向量获取单元424，用于将各个第一关键词和各个第一关键词出现的次数映射到向量词典，以得到检索向量。

[0107] 进一步地，作为本实施例的一种实施方式，如图9所示，第一关键词获取单元422包括同义词获取子单元4221、语义相似度获取子单元4222和第一关键词获取子单元4223。各功能子单元详细功能如下：

[0108] 同义词获取子单元4221，用于从同义词词林中获取与各个词汇相匹配的至少一个同义词；

[0109] 语义相似度获取子单元4222，用于计算各个词汇与对应匹配的同义词的语义相似度；

[0110] 第一关键词获取子单元4223，用于当词汇与同义词的相似度达到预设第一阈值时，将同义词作为与对应词汇相匹配的第一关键词。

[0111] 进一步地，作为本实施例的一种实施方式，语义相似度获取子单元4222包括语义信息获取子单元、数据集获取子单元和语义相似度计算子单元。各功能子单元详细功能如下：

[0112] 语义信息获取子单元，用于根据词汇获取第一语义信息，根据同义词获取第二语义信息；

[0113] 数据集获取子单元，用于从第一语义信息中获取第一语义关键词，以形成第一数据集，从第二语义信息中获取第二语义关键词，以形成第二数据集；

[0114] 语义相似度计算子单元，用于计算第一数据集和第二数据集之间的相似度，并将计算的相似度作为语义相似度。

[0115] 进一步地，作为本实施例的一种实施方式，基于向量的文档检索装置还包括资源词汇获取模块、第二关键词获取模块、词频统计模块和文档向量获取模块。

各功能模块详细功能如下：

- [0116] 资源词汇获取模块，用于从资源文档中获取各个资源词汇；
- [0117] 第二关键词获取模块，用于将各个资源词汇中语义相近的资源词汇分别用第二关键词表示；
- [0118] 词频统计模块，用于统计各个第二关键词出现的次数；
- [0119] 文档向量获取模块，用于将各个第二关键词和各个第二关键词出现的次数映射到向量词典，以得到文档向量。
- [0120] 关于基于向量的文档检索装置的具体限定可以参见上文中对于基于向量的文档检索方法的限定，在此不再赘述。上述基于向量的文档检索装置中的各个模块可全部或部分通过软件、硬件及其组合来实现。上述各模块可以硬件形式内嵌于或独立于计算机设备中的处理器中，也可以以软件形式存储于计算机设备中的存储器中，以便于处理器调用执行以上各个模块对应的操作。
- [0121] 本申请的第三实施例提供了一种计算机设备，该计算机设备可以是服务器，其内部结构图可以如图10所示。该计算机设备包括通过系统总线连接的处理器、存储器、网络接口和数据库。其中，该计算机设备的处理器用于提供计算和控制能力。该计算机设备的存储器包括可读存储介质、内存储器。该非易失性存储介质存储有操作系统、计算机可读指令和数据库。该内存储器为非易失性存储介质中的操作系统和计算机可读指令的运行提供环境。该计算机设备的数据库用于存储基于向量的文档检索方法中涉及到的数据。该计算机设备的网络接口用于与外部的终端通过网络连接通信。该计算机可读指令被处理器执行时以实现本申请的第一实施例提供的基于向量的文档检索方法。本示例中，可读存储介质可以是非易失性可读存储介质，也可以是易失性可读存储介质。
- [0122] 本申请的第四实施例提供了一种一个或多个存储有计算机可读指令的可读存储介质，所述计算机可读存储介质存储有计算机可读指令，其中，所述计算机可读指令被一个或多个处理器执行时，使得所述一个或多个处理器执行时实现本申请的第一实施例提供的基于向量的文档检索方法的步骤，例如图2所示的步骤11至步骤14、如图3所示的步骤121至步骤124、如图4所示的步骤1221至步骤1223、如图5所示的步骤12221至步骤12223以及如图6所示的步骤21至步骤24。或者

，计算机可读指令被处理器执行时实现上述第一实施例提供的基于向量的文档检索方法的各模块/单元的功能。为避免重复，这里不再赘述。

[0123] 本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程，是可以通过计算机可读指令来指令相关的硬件来完成，所述的计算机可读指令可存储于一非易失性计算机可读取存储介质中，该计算机可读指令在执行时，可包括如上述各方法的实施例的流程。其中，本申请所提供的各实施例中所使用的对存储器、存储、数据库或其它介质的任何引用，均可包括非易失性和/或易失性存储器。非易失性存储器可包括只读存储器（ROM）、可编程ROM（PROM）、电可编程ROM（EPROM）、电可擦除可编程ROM（EEPROM）或闪存。易失性存储器可包括随机存取存储器（RAM）或者外部高速缓冲存储器。作为说明而非局限，RAM以多种形式可得，诸如静态RAM（SRAM）、动态RAM（DRAM）、同步DRAM（SDRAM）、双数据率SDRAM（DDRSDRAM）、增强型SDRAM（ESDRAM）、同步链路（Synchlink）DRAM（SLDRAM）、存储器总线（Rambus）直接RAM（RDRAM）、直接存储器总线动态RAM（DRDRAM）、以及存储器总线动态RAM（RDRAM）等。

[0124] 所属领域的技术人员可以清楚地了解到，为了描述的方便和简洁，仅以上述各功能单元、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单元、模块完成，即将所述装置的内部结构划分成不同的功能单元或模块，以完成以上描述的全部或者部分功能。

[0125] 以上所述实施例仅用以说明本申请的技术方案，而非对其限制；尽管参照前述实施例对本申请进行了详细的说明，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申请各实施例技术方案的精神和范围，均应包含在本申请的保护范围之内。

权利要求书

- [权利要求 1] 一种基于向量的文档检索方法，其中，所述文档检索方法包括：
- 获取在客户端输入的检索信息；
- 提取所述检索信息中的各个词汇，并根据所述检索信息中各个词汇的语义将所述检索信息转换为检索向量；
- 计算所述检索向量与预先形成的与资源文档相匹配的文档向量之间的相似度；
- 按照所述相似度的大小将与所述文档向量相匹配的所述资源文档进行排序，将所述资源文档的排序结果作为检索结果。
- [权利要求 2] 根据权利要求1所述的基于向量的文档检索方法，其中，所述提取所述检索信息中的各个词汇，并根据所述检索信息中各个词汇的语义将所述检索信息转换为检索向量包括：
- 获取所述检索信息所包含的各个词汇；
- 将所述各个词汇中语义相近的所述词汇分别用同一第一关键词进行表示；
- 统计各个所述第一关键词出现的次数；
- 将各个所述第一关键词和各个所述第一关键词出现的次数映射到向量词典，以得到所述检索向量。
- [权利要求 3] 根据权利要求2所述的基于向量的文档检索方法，其中，所述将所述各个词汇中语义相近的所述词汇分别用同一第一关键词进行表示包括：
- 从同义词词林中获取与所述各个词汇相匹配的至少一个同义词；
- 计算所述各个词汇与对应匹配的所述同义词的语义相似度；
- 当所述词汇与所述同义词的相似度达到预设第一阈值时，将所述同义词作为与对应词汇相匹配的所述第一关键词。
- [权利要求 4] 根据权利要求3所述的基于向量的文档检索方法，其中，所述计算所述各个词汇与对应匹配的所述同义词的语义相似度的步骤包括：
- 根据所述词汇获取第一语义信息，根据所述同义词获取第二语义信息

;

从所述第一语义信息中获取第一语义关键词，以形成第一数据集，从所述第二语义信息中获取第二语义关键词，以形成第二数据集；
计算所述第一数据集和所述第二数据集之间的相似度，并将计算的所述相似度作为所述语义相似度。

[权利要求 5]

根据权利要求1所述的基于向量的文档检索方法，其中，所述基于向量的文档检索方法还包括：

从所述资源文档中获取各个资源词汇；

将所述各个资源词汇中语义相近的所述资源词汇分别用第二关键词表示；

统计各个所述第二关键词出现的次数；

将各个所述第二关键词和各个所述第二关键词出现的次数映射到向量词典，以得到所述文档向量。

[权利要求 6]

一种基于向量的文档检索装置，其中，包括：

检索信息获取模块，用于获取在客户端输入的检索信息；

检索向量获取模块，用于提取所述检索信息中的各个词汇，并根据所述检索信息中各个词汇的语义将所述检索信息转换为检索向量；

相似度获取模块，用于计算所述检索向量与预先形成的与资源文档相匹配的文档向量之间的相似度；

检索结果获取模块，用于按照所述相似度的大小将与所述文档向量相匹配的所述资源文档进行排序，将所述资源文档的排序结果作为检索结果。

[权利要求 7]

根据权利要求6所述的基于向量的文档检索装置，其中，所述检索向量获取模块包括：

分词处理单元，用于获取所述检索信息所包含的各个词汇；

第一关键词获取单元，用于将所述各个词汇中语义相近的所述词汇分别用同一第一关键词进行表示；

统计单元，用于统计各个所述第一关键词出现的次数；

检索向量获取单元，用于将各个所述第一关键词和各个所述第一关键词出现的次数映射到向量词典，以得到所述检索向量。

- [权利要求 8] 根据权利要求7所述的基于向量的文档检索装置，其中，所述第一关键词获取单元包括：
- 同义词获取子单元，用于从同义词词林中获取与所述各个词汇相匹配的至少一个同义词；
- 语义相似度获取子单元，用于计算所述各个词汇与对应匹配的所述同义词的语义相似度；
- 第一关键词获取子单元，用于当所述词汇与所述同义词的相似度达到预设第一阈值时，将所述同义词作为与对应词汇相匹配的所述第一关键词。

- [权利要求 9] 根据权利要求8所述的基于向量的文档检索装置，其中，所述语义相似度获取子单元，包括：
- 语义信息获取子单元，用于根据所述词汇获取第一语义信息，根据所述同义词获取第二语义信息；
- 数据集获取子单元，用于从所述第一语义信息中获取第一语义关键词，以形成第一数据集，从所述第二语义信息中获取第二语义关键词，以形成第二数据集；
- 语义相似度计算子单元，用于计算所述第一数据集和所述第二数据集之间的相似度，并将计算的所述相似度作为所述语义相似度。

- [权利要求 10] 根据权利要求6所述的基于向量的文档检索装置，其中，所述基于向量的文档检索装置还包括：
- 资源词汇获取模块，用于从所述资源文档中获取各个资源词汇；
- 第二关键词获取模块，用于将所述各个资源词汇中语义相近的所述资源词汇分别用第二关键词表示；
- 词频统计模块，用于统计各个所述第二关键词出现的次数；
- 文档向量获取模块，用于将各个所述第二关键词和各个所述第二关键词出现的次数映射到向量词典，以得到所述文档向量。

- [权利要求 11] 一种计算机设备，包括存储器、处理器以及存储在所述存储器中并可在所述处理器上运行的计算机可读指令，其中，所述处理器执行所述计算机可读指令时实现如下步骤：
- 获取在客户端输入的检索信息；
- 提取所述检索信息中的各个词汇，并根据所述检索信息中各个词汇的语义将所述检索信息转换为检索向量；
- 计算所述检索向量与预先形成的与资源文档相匹配的文档向量之间的相似度；
- 按照所述相似度的大小将与所述文档向量相匹配的所述资源文档进行排序，将所述资源文档的排序结果作为检索结果。
- [权利要求 12] 根据权利要求11所述的计算机设备，其中，所述提取所述检索信息中的各个词汇，并根据所述检索信息中各个词汇的语义将所述检索信息转换为检索向量包括：
- 获取所述检索信息所包含的各个词汇；
- 将所述各个词汇中语义相近的所述词汇分别用同一第一关键词进行表示；
- 统计各个所述第一关键词出现的次数；
- 将各个所述第一关键词和各个所述第一关键词出现的次数映射到向量词典，以得到所述检索向量。
- [权利要求 13] 根据权利要求12所述的计算机设备，其中，所述将所述各个词汇中语义相近的所述词汇分别用同一第一关键词进行表示包括：
- 从同义词词林中获取与所述各个词汇相匹配的至少一个同义词；
- 计算所述各个词汇与对应匹配的所述同义词的语义相似度；
- 当所述词汇与所述同义词的相似度达到预设第一阈值时，将所述同义词作为与对应词汇相匹配的所述第一关键词。
- [权利要求 14] 根据权利要求13所述的计算机设备，其中，所述计算所述各个词汇与对应匹配的所述同义词的语义相似度的步骤包括：
- 根据所述词汇获取第一语义信息，根据所述同义词获取第二语义信息

;

从所述第一语义信息中获取第一语义关键词，以形成第一数据集，从所述第二语义信息中获取第二语义关键词，以形成第二数据集；
计算所述第一数据集和所述第二数据集之间的相似度，并将计算的所述相似度作为所述语义相似度。

[权利要求 15] 根据权利要求11所述的计算机设备，其中，所述处理器执行所述计算机可读指令时还实现如下步骤：

从所述资源文档中获取各个资源词汇；
将所述各个资源词汇中语义相近的所述资源词汇分别用第二关键词表示；
统计各个所述第二关键词出现的次数；
将各个所述第二关键词和各个所述第二关键词出现的次数映射到向量词典，以得到所述文档向量。

[权利要求 16] 一个或多个存储有计算机可读指令的可读存储介质，所述计算机可读存储介质存储有计算机可读指令，其中，所述计算机可读指令被一个或多个处理器执行时，使得所述一个或多个处理器执行如下步骤：

获取在客户端输入的检索信息；
提取所述检索信息中的各个词汇，并根据所述检索信息中各个词汇的语义将所述检索信息转换为检索向量；
计算所述检索向量与预先形成的与资源文档相匹配的文档向量之间的相似度；
按照所述相似度的大小将与所述文档向量相匹配的所述资源文档进行排序，将所述资源文档的排序结果作为检索结果。

[权利要求 17] 根据权利要求16所述的存储有计算机可读指令的可读存储介质，其中，所述提取所述检索信息中的各个词汇，并根据所述检索信息中各个词汇的语义将所述检索信息转换为检索向量包括：

获取所述检索信息所包含的各个词汇；
将所述各个词汇中语义相近的所述词汇分别用同一第一关键词进行表

示；

统计各个所述第一关键词出现的次数；

将各个所述第一关键词和各个所述第一关键词出现的次数映射到向量词典，以得到所述检索向量。

[权利要求 18] 根据权利要求17所述的 readable 存储介质，其中，所述将所述各个词汇中语义相近的所述词汇分别用同一第一关键词进行表示包括：

从同义词词林中获取与各个所述词汇相匹配的至少一个同义词；

计算各个所述词汇与对应匹配的所述同义词的语义相似度；

当所述词汇与所述同义词的相似度达到预设第一阈值时，将所述同义词作为与对应词汇相匹配的所述第一关键词。

[权利要求 19] 根据权利要求18所述的 readable 存储介质，其中，所述计算各个所述词汇与对应匹配的所述同义词的语义相似度的步骤包括：

根据所述词汇获取第一语义信息，根据所述同义词获取第二语义信息；

从所述第一语义信息中获取第一语义关键词，以形成第一数据集，从所述第二语义信息中获取第二语义关键词，以形成第二数据集；

计算所述第一数据集和所述第二数据集之间的相似度，并将计算的所述相似度作为所述语义相似度。

[权利要求 20] 根据权利要求16所述的 readable 存储介质，其中，所述计算机可读指令被一个或多个处理器执行时，使得所述一个或多个处理器还执行如下步骤：

从所述资源文档中获取各个资源词汇；

将各个所述资源词汇中语义相近的所述资源词汇分别用第二关键词表示；

统计各个所述第二关键词出现的次数；

将各个所述第二关键词和各个所述第二关键词出现的次数映射到向量词典，以得到所述文档向量。

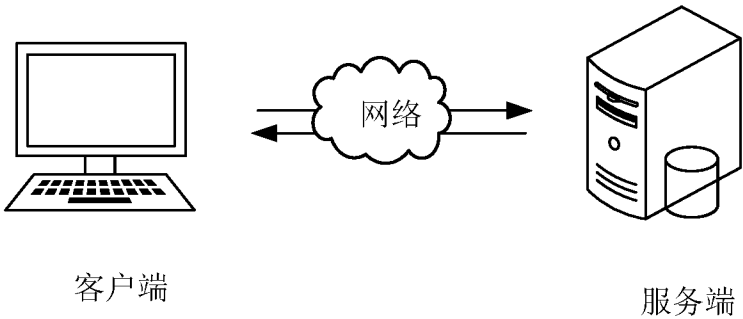


图 1

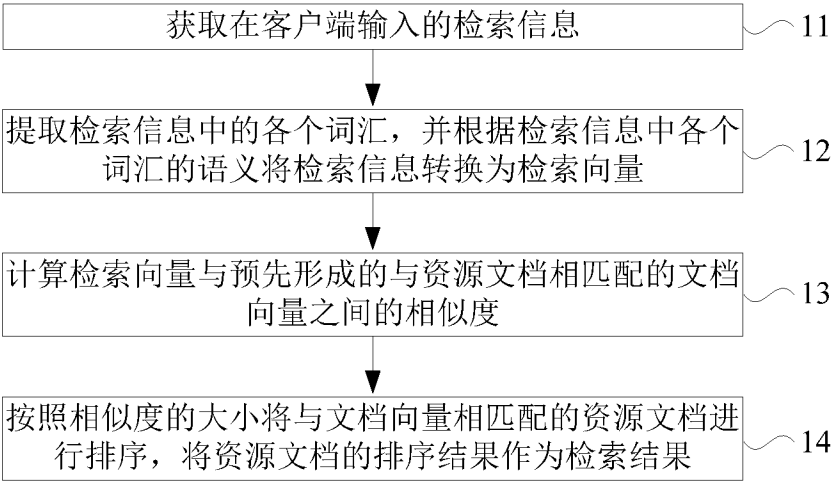


图 2

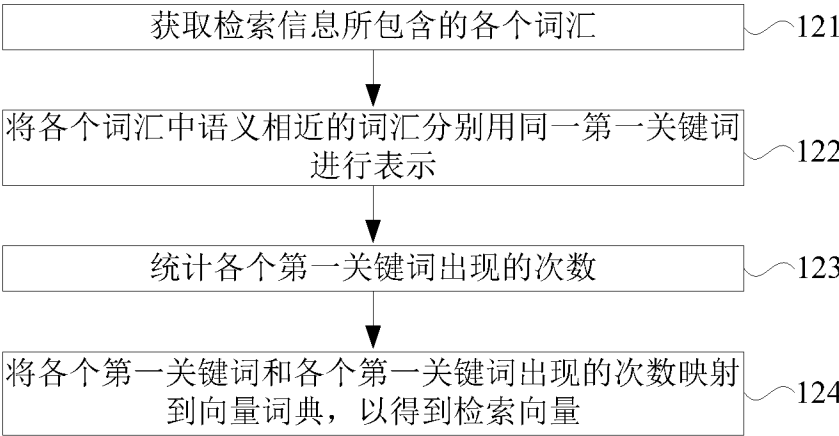


图 3

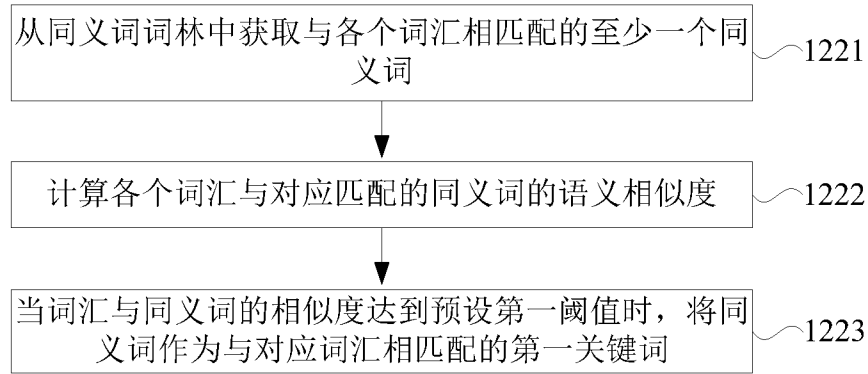


图 4

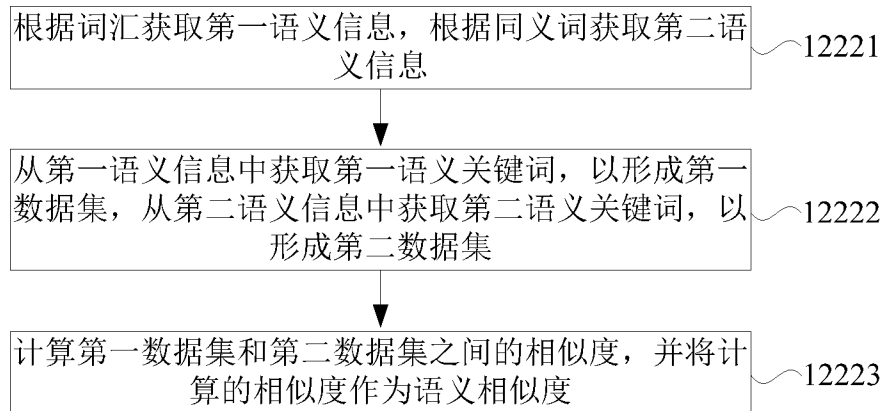


图 5

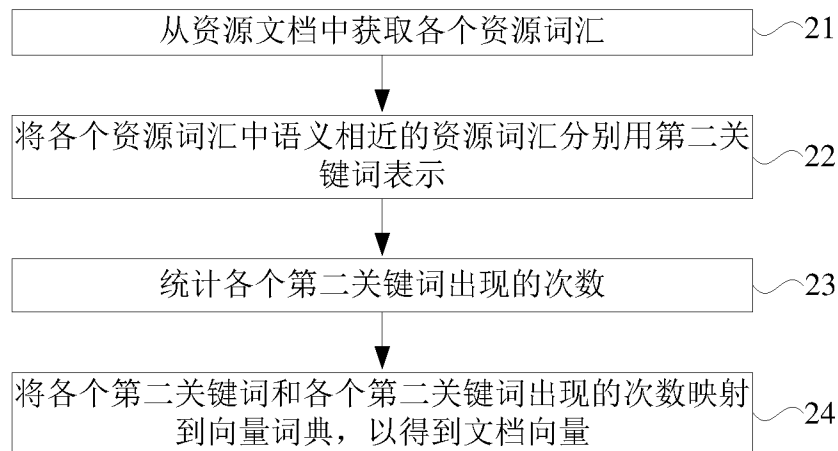


图 6

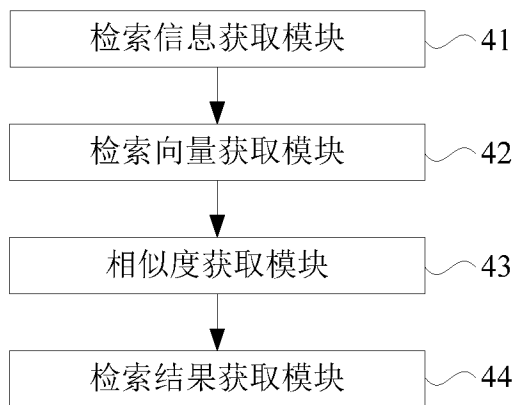


图 7

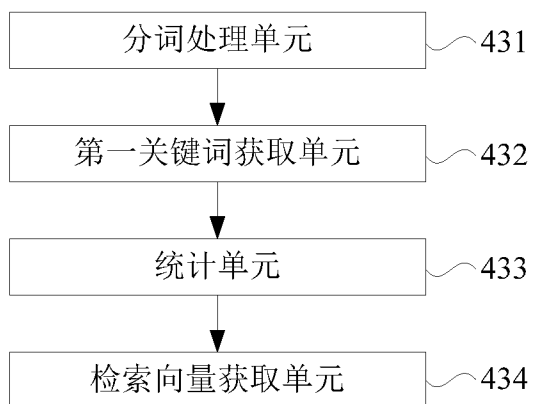


图 8

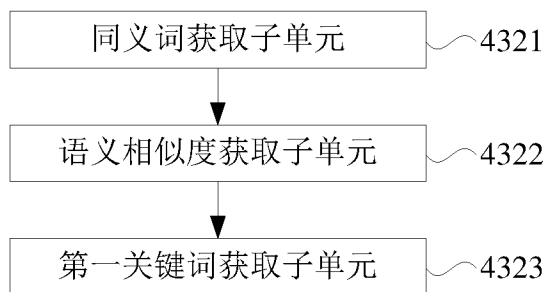


图 9

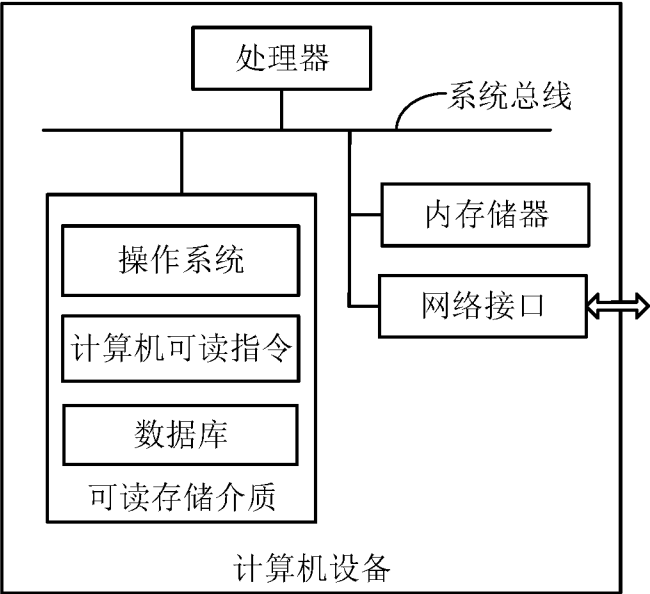


图 10

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2021/070585

A. CLASSIFICATION OF SUBJECT MATTER

G06F 16/33(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, WPI, EPODOC, CNKI, IEEE: 向量, 文档, 检索, 词汇, 相似, 排序, 同义词, 关键词, 替换, vector, document, search
+, word, similar+, rank, synonyms, keyword, replac+**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 111460090 A (SHENZHEN ONECONNECT TECHNOLOGY CO., LTD.) 28 July 2020 (2020-07-28) description, paragraphs 32-111, and figures 1-4	1-20
X	CN 110807149 A (ZALL SMART COMMERCE (WUHAN) RESEARCH INSTITUTE CO., LTD.) 18 February 2020 (2020-02-18) description, paragraphs 56-73, paragraphs 151-185, figures 1, 4	1-20
X	CN 110019669 A (BEIJING GRIDSUM TECHNOLOGY CO., LTD.) 16 July 2019 (2019-07-16) description, paragraphs 53-94, and figures 1-2	1-20
A	CN 110276071 A (ZHONGAN ONLINE P&C INSURANCE CO., LTD.) 24 September 2019 (2019-09-24) entire document	1-20
A	US 2019179910 A1 (INTERNATIONAL BUSINESS MACHINES CORPORATION) 13 June 2019 (2019-06-13) 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

11 March 2021

Date of mailing of the international search report

01 April 2021

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
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/CN2021/070585

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN	111460090	A	28 July 2020	None	
CN	110807149	A	18 February 2020	None	
CN	110019669	A	16 July 2019	None	
CN	110276071	A	24 September 2019	None	
US	2019179910	A1	13 June 2019	None	

国际检索报告

国际申请号

PCT/CN2021/070585

A. 主题的分类 G06F 16/33 (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, WPI, EPDOC, CNKI, IEEE: 向量, 文档, 检索, 词汇, 相似, 排序, 同义词, 关键词, 替换, vector, document, search+, word, similar+, rank, synonyms, keyword, replac+																				
C. 相关文件 <table border="1"> <thead> <tr> <th>类 型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>PX</td> <td>CN 111460090 A (深圳壹账通智能科技有限公司) 2020年 7月 28日 (2020 - 07 - 28) 说明书第32-111段, 图1-4</td> <td>1-20</td> </tr> <tr> <td>X</td> <td>CN 110807149 A (卓尔智联武汉研究院有限公司) 2020年 2月 18日 (2020 - 02 - 18) 说明书第56-73段、第151-185段, 图1、4</td> <td>1-20</td> </tr> <tr> <td>X</td> <td>CN 110019669 A (北京国双科技有限公司) 2019年 7月 16日 (2019 - 07 - 16) 说明书第53-94段, 图1-2</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>CN 110276071 A (众安在线财产保险股份有限公司) 2019年 9月 24日 (2019 - 09 - 24) 全文</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>US 2019179910 A1 (INTERNATIONAL BUSINESS MACHINES CORPORATION) 2019年 6月 13日 (2019 - 06 - 13) 全文</td> <td>1-20</td> </tr> </tbody> </table>			类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求	PX	CN 111460090 A (深圳壹账通智能科技有限公司) 2020年 7月 28日 (2020 - 07 - 28) 说明书第32-111段, 图1-4	1-20	X	CN 110807149 A (卓尔智联武汉研究院有限公司) 2020年 2月 18日 (2020 - 02 - 18) 说明书第56-73段、第151-185段, 图1、4	1-20	X	CN 110019669 A (北京国双科技有限公司) 2019年 7月 16日 (2019 - 07 - 16) 说明书第53-94段, 图1-2	1-20	A	CN 110276071 A (众安在线财产保险股份有限公司) 2019年 9月 24日 (2019 - 09 - 24) 全文	1-20	A	US 2019179910 A1 (INTERNATIONAL BUSINESS MACHINES CORPORATION) 2019年 6月 13日 (2019 - 06 - 13) 全文	1-20
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求																		
PX	CN 111460090 A (深圳壹账通智能科技有限公司) 2020年 7月 28日 (2020 - 07 - 28) 说明书第32-111段, 图1-4	1-20																		
X	CN 110807149 A (卓尔智联武汉研究院有限公司) 2020年 2月 18日 (2020 - 02 - 18) 说明书第56-73段、第151-185段, 图1、4	1-20																		
X	CN 110019669 A (北京国双科技有限公司) 2019年 7月 16日 (2019 - 07 - 16) 说明书第53-94段, 图1-2	1-20																		
A	CN 110276071 A (众安在线财产保险股份有限公司) 2019年 9月 24日 (2019 - 09 - 24) 全文	1-20																		
A	US 2019179910 A1 (INTERNATIONAL BUSINESS MACHINES CORPORATION) 2019年 6月 13日 (2019 - 06 - 13) 全文	1-20																		
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。																				
<table border="0"> <tr> <td> * 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 </td> <td> “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 </td> </tr> </table>			* 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																
* 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 2021年 3月 11日		国际检索报告邮寄日期 2021年 4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		授权官员 徐晓 电话号码 010-53961324																		

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2021/070585

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	111460090	A	2020年 7月 28日	无	
CN	110807149	A	2020年 2月 18日	无	
CN	110019669	A	2019年 7月 16日	无	
CN	110276071	A	2019年 9月 24日	无	
US	2019179910	A1	2019年 6月 13日	无	