

(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
2024 年 12 月 19 日 (19.12.2024)



(10) 国际公布号
WO 2024/255714 A1

(51) 国际专利分类号:
G06F 16/2458 (2019.01)

(21) 国际申请号: PCT/CN2024/098256

(22) 国际申请日: 2024 年 6 月 7 日 (07.06.2024)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202310727602.6 2023年6月16日 (16.06.2023) CN

(71) 申请人: 深圳市世强元件网络有限公司
(**SEKORM ESERVICE CO., LTD.**) [CN/CN]; 中国广东省深圳市龙岗区坂田街道雪岗路 2018 号天安云谷产业园一期 3 栋 A 座 2401、2402 单元, Guangdong 518129 (CN)。

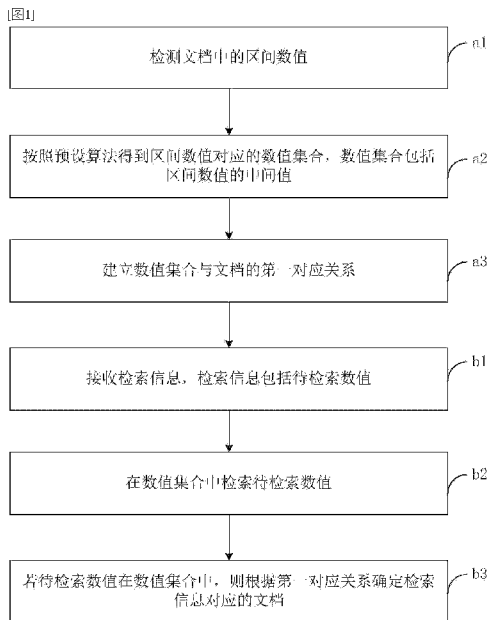
(72) 发明人: 包伟(**BAO, Wei**); 中国广东省深圳市龙岗区坂田街道雪岗路 2018 号天安云谷产业园一期 3 栋 A 座 2401、2402 单元, Guangdong 518129 (CN)。

(74) 代理人: 深圳市瑞方达知识产权事务所 (普通合伙) (**SHENZHEN REFINED INTELLECTUAL PROPERTY OFFICE (GENERAL PARTNERSHIP)**); 中国广东省深圳市南山区科兴路 11 号深南花园裙楼 B 区 413 室, Guangdong 518057 (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, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA,

(54) **Title:** RETRIEVAL METHOD FOR NUMERIC VALUES WITHIN INTERVAL, STORAGE MEDIUM, AND COMPUTER

(54) 发明名称: 一种区间数值检索方法、存储介质及计算机



- a1 Detect a numeric interval in a document
a2 Obtain, according to a preset algorithm, a numeric value set corresponding to said interval, wherein the numeric value set comprises the intermediate values of the interval
a3 Establish a first correspondence relationship between the numeric value set and the document
b1 Receive search information, said search information comprising a numeric value to be retrieved
b2 Search for said numeric value in numeric value sets
b3 If the numeric value is in the numeric value sets, determine, according to the first correspondence relationship, documents corresponding to the search information

(57) **Abstract:** The present invention relates to a retrieval method for numeric values within an interval, a storage medium, and a computer, said method comprising a procedure for establishing a numeric value set and a procedure for retrieving a numeric value. The procedure for establishing a numeric value set comprises: step a1, detecting a numeric interval in a document; step a2, obtaining, according to a preset algorithm, a numeric value set corresponding to said interval, wherein the numeric value set comprises the intermediate values of the interval; and step a3, establishing a first correspondence relationship between the numeric value set and the document. The procedure for retrieving a numeric value comprises: step b1, receiving search information, said search information comprising a numeric value to be retrieved; step b2, searching for said numeric value in numeric value sets; and step b3, if the numeric

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, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, 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, ME, 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))。

value is in the numeric value sets, determining, according to the first correspondence relationship, documents corresponding to the search information. By supplementing part of the intermediate values of an interval, the retrieval of part of the intermediate values of the interval is achieved, thus improving retrieval accuracy for numeric values within an interval.

(57) 摘要: 本发明涉及一种区间数值检索方法、存储介质及计算机。该方法包括建立数值集合过程和数值检索过程。建立数值集合包括: 步骤a1、检测文档中的区间数值; 步骤a2、按照预设算法得到区间数值对应的数值集合, 数值集合包括区间数值的中间值; 步骤a3、建立数值集合与文档的第一对应关系。数值检索过程包括: 步骤b1、接收检索信息, 检索信息包括待检索数值; 步骤b2、在数值集合中检索待检索数值; 步骤b3、若待检索数值在数值集合中, 则根据第一对应关系确定检索信息对应的文档。本发明通过补齐区间数值的部分中间值, 实现区间数值部分中间值的检索, 提高区间数值检索准确率。

说明书

发明名称: 一种区间数值检索方法、存储介质及计算机技术领域

[0001] 本发明涉及信息检索领域，更具体地说，涉及一种区间数值检索方法、存储介质及计算机。

背景技术

[0002] 利用计算机或互联网进行数据检索能帮助人们在海量数据中快速找到所需内容，其中关键字检索是常用的检索方式。现有技术中关键字检索采用“所见即所得”的匹配逻辑，即检索时只能检索数据库中已经存在的文字，判断用户输入的关键字是否与数据库中已有文字匹配。例如用户输入的关键字为数字“3”，则需要数据库中明确出现数字“3”，才能成功匹配。

[0003] 现有技术这种“所见即所得”的关键字检索匹配逻辑，无法实现对于区间数值中间值的检索，区间数值表达数值的一个区间段，中间值指区间数值最小值和最大值之间的取值，例如区间数值0至9，中间值为0至9之间的任意值，例如3。根据“所见即所得”的原理，区间数值0至9能够表现出的数值进是0和9两个值，则现有技术只能检索到0和9，而0至9之间的取值就不能被检索到。例如用户输入的关键字是3，虽然3处于0至9之间，但因为3并没有实际展现出来，所以不能被检索到，从而导致检索结果不准确。

发明内容

[0004] 本发明要解决的技术问题在于，提供一种区间数值检索方法、存储介质及计算机。

[0005] 本发明解决其技术问题所采用的技术方案是：构造一种区间数值检索方法，包括建立数值集合过程和数值检索过程；

[0006] 所述建立数值集合包括：

[0007] 步骤a1、检测文档中的区间数值；

- [0008] 步骤a2、按照预设算法得到所述区间数值对应的数值集合，所述数值集合包括所述区间数值的中间值；
- [0009] 步骤a3、建立所述数值集合与所述文档的第一对应关系；
- [0010] 所述数值检索过程包括：
- [0011] 步骤b1、接收检索信息，所述检索信息包括待检索数值；
- [0012] 步骤b2、在所述数值集合中检索所述待检索数值；
- [0013] 步骤b3、若所述待检索数值在所述数值集合中，则根据所述第一对应关系确定所述检索信息对应的文档。
- [0014] 进一步，在本发明所述的区间数值检索方法中，所述数值集合包括至少一个中间值，所述中间值大于所述区间数值的最小值且小于所述区间数值的最大值。
- [0015] 进一步，在本发明所述的区间数值检索方法中，所述步骤a2包括：
- [0016] 从所述区间数值的最小值开始按照预设步长逐步增加取值直至所述区间数值的最大值，每增加一次所述预设步长产生一个中间值，由所得所有中间值组成所述区间数值对应的数值集合；或者
- [0017] 从所述区间数值的最大值开始按照预设步长逐步减小取值直至所述区间数值的最小值，每减小一次所述预设步长产生一个中间值，由所得所有中间值组成所述区间数值对应的数值集合。
- [0018] 进一步，在本发明所述的区间数值检索方法中，在所述步骤a1之后所述步骤a2之前还包括：
- [0019] 获取所述区间数值对应的数值单位；
- [0020] 根据第二对应关系查找所述数值单位对应的预设步长，所述第二对应关系为所述数值单位和所述预设步长的对应关系。
- [0021] 进一步，在本发明所述的区间数值检索方法中，在所述步骤a1之后所述步骤a2之前还包括：
- [0022] 获取所述区间数值对应的数值单位及所述文档所属的产品信息；
- [0023] 根据第三对应关系查找所述数值单位和所述产品信息对应的预设步长，所述第三对应关系为所述预设步长与所述数值单位和产品信息的对应关系。

- [0024] 进一步，在本发明所述的区间数值检索方法中，所述步骤a1包括：检测文档中的区间数值以及所述区间数值在所述文档中的位置信息；
- [0025] 所述步骤a3包括：建立所述数值集合与所述文档的第一对应关系，以及所述数值集合与所述位置信息的第四对应关系；
- [0026] 所述步骤b3包括：若所述待检索数值在所述数值集合中，则根据所述第一对应关系确定所述检索信息对应的文档，以及根据所述第四对应关系确定所述检索信息在所述文档中的位置。
- [0027] 进一步，在本发明所述的区间数值检索方法中，所述步骤a1包括：检测文档中的区间数值及所述区间数值对应的数值单位；
- [0028] 所述步骤a2中所述数值集合包括所述区间数值的中间值和所述数值单位；
- [0029] 所述步骤b1中所述检索信息包括待检索数值和数值单位；
- [0030] 所述步骤b2包括：在所述数值集合中检索所述待检索数值和所述数值单位；
- [0031] 所述步骤b3中若所述待检索数值在所述数值集合中包括：若所述待检索数值和所述数值单位在同一所述数值集合中。
- [0032] 进一步，在本发明所述的区间数值检索方法中，在所述步骤b3之后还包括：
- [0033] 若根据所述第一对应关系确定所述检索信息对应的文档的数量超过一个，则按照预设排序规则对所得文档进行排序并显示。
- [0034] 另外，本发明还提供一种计算机可读存储介质，所述计算机可读存储介质存储有计算机程序，所述计算机程序适于处理器进行加载，以执行如上述的区间数值检索方法的步骤。
- [0035] 另外，本发明还提供一种计算机，包括存储器和处理器，所述存储器中存储有计算机程序，所述处理器通过调用所述存储器中存储的所述计算机程序，执行如上述的区间数值检索方法的步骤。
- [0036] 实施本发明的一种区间数值检索方法、存储介质及计算机，具有以下有益效果：本发明通过补齐区间数值的部分中间值，实现区间数值部分中间值的检索，提高区间数值检索准确率。

附图说明

- [0037] 下面将结合附图及实施例对本发明作进一步说明，附图中：

[0038] 图1是本发明实施例提供的区间数值检索方法的流程图；

[0039] 图2是本发明实施例提供的区间数值检索方法的流程图；

[0040] 图3是本发明实施例提供的区间数值检索方法的流程图。

具体实施方式

[0041] 为了对本发明的技术特征、目的和效果有更加清楚的理解，现对照附图详细说明本发明的具体实施方式。

[0042] 在一个优选实施例中，参考图1，本实施例的区间数值检索方法应用于检索领域，包括计算机本地数据检索、服务器数据检索和云端数据检索等。区间数值是指数值的一个区间段，即一个取值范围。区间数值可以是只包含数字的数字范围，例如0至9；也可以是包含数值单位的取值范围，例如0℃（摄氏度）至9℃，0m（米）至9m，0A（安培）至9A等，其中0℃（摄氏度）、m（米）和A（安培）为数值单位。具体的，该区间数值检索方法包括建立数值集合过程和数值检索过程，以下分别进行说明。

[0043] 首先，在检索前要建立数值集合，该数值集合建立过程包括下述步骤：

[0044] 步骤a1、检测文档中的区间数值。

[0045] 具体的，本实施例的文档是指包含一定数据的文件，尤其指包含文字的文件，以便能够用于本实施例的文字检索。可以理解，文件格式表现形式较多，本实施例对此不做限定。作为选择，文档包括但不限于word文档、TXT文档、PDF文档、EXCEL表格、PPT文档、网页等。区间数值通常有特定表达方式，可以使用中间符号“-”、“~”、“（）”、“[]”等来表示，例如使用中间符号表示的区间数值0-9、0~9，（0,9），[0,9]；还可以用表示范围的文字来表达，例如0至9，从0到9等，位于0和9之间等。基于区间数值的特定表达方式，将这些特定表达方式制作为区间数值模型，使用区间数值模型在文档中检测与区间数值模型匹配的区间数值。例如区间数值模型“数字-数字”，如果在文档中检测到符合区间数值模型“数字-数字”的文字，认为是检索到区间数值，例如0-9符合区间数值模型“数字-数字”。通过使用区间数值模型可实现文档中区间数值的自动检测，不需要人工查找文档中的区间数值，效率高，成本低。

[0046] 步骤a2、按照预设算法得到区间数值对应的数值集合，数值集合包括区间数值的中间值。

[0047] 具体的，检索到文档中的区间数值后，按照预设算法得到区间数值对应的数值集合，每个区间数值对应一个数值集合，数值集合包括区间数值的中间值。数值集合包括至少一个中间值，中间值大于区间数值的最小值且小于区间数值的最大值。可以理解，某一区间数值对应的数值集合包含的中间值越多，被检索到的概率就越大，从而提高该区间数值被检索到的概率。例如，区间数值0至9，中间值可取1、2、3、4、5、6、7、8，即由中间值1、2、3、4、5、6、7、8组成区间数值0至9对应的数值集合。但需要说明的是，数值集合并不包含区间数值最小值和最大值之间的所有数值，而是取尽可能多的数值，所取数值越多，即密度越大，区间数值被检索的概率越大；但中间值越多，对应的数据处理量就越大，所以需要在中间值的数量和被检索概率之间取得平衡。

[0048] 步骤a3、建立数值集合与文档的第一对应关系。

[0049] 具体的，得到某一区间数值对应的数值集合后，建立数值集合与文档的第一对应关系，即建立数值集合与对应区间数值所在文档的第一对应关系。若文档包含两个以上区间数值，则该文档对应多个数值集合。通常情况下，计算机或服务器或云端数据库存储有大量文档，检索出每个文档中的区间数值，并建立每个文档对应的数值集合，从而建立数据库对应的数值集合。

[0050] 可以理解，建立第一对应关系后数值集合中每个中间值与文档都对应，这样就能实现文档内区间数值的中间值的检索。需要说明的是，文档中某一区间数值对应的数值集合为文档之外新产生数据，并不属于文档本身，也不需要修改文档，而是作为单独的数据另外存储，然后建立数值集合与文档的第一对应关系。

[0051] 其次，在建立数值集合后进行数值检索，该数值检索过程包括下述步骤：

[0052] 步骤b1、接收检索信息，检索信息包括待检索数值。

[0053] 具体的，接收用户输入的检索信息，该检索信息包括待检索数值，当然还可以包括其他检索信息。需要说明的是，该待检索数值为数值或数字。

[0054] 步骤b2、在数值集合中检索待检索数值。

- [0055] 具体的，因数值集合中所有中间值都已是具体数值，能够通过“所见即所得”的文字匹配方式进行检索，所以可在数值集合中检索待检索数值。例如区间数值0至9，中间值可取1、2、3、4、5、6、7、8，即由中间值1、2、3、4、5、6、7、8组成区间数值0至9对应的数值集合。若用户输入的待检索数值为3，则在数值集合1、2、3、4、5、6、7、8中检索待检索数值3。可以理解的，待检索数值依然可用于文档中已有内容的检索，可参考现有技术，此处不再赘述。
- [0056] 步骤b3、若待检索数值在数值集合中，则根据第一对应关系确定检索信息对应的文档。
- [0057] 具体的，通过在数值集合中检索待检索数值，若待检索数值在数值集合中，则根据第一对应关系确定检索信息对应的文档，从而实现区间数值未直接显示部分的检索。例如区间数值0至9，中间值可取1、2、3、4、5、6、7、8，即由中间值1、2、3、4、5、6、7、8组成区间数值0至9对应的数值集合。若用户输入的待检索数值为3，则在数值集合1、2、3、4、5、6、7、8中检索待检索数值3，刚好能匹配到，则可确定待检索数值3对应的文档。
- [0058] 另外，假如用户输入的待检索数值是3.5，则在数值集合1、2、3、4、5、6、7、8中不能匹配成功，但3.5确实也在区间数值0至9中，此处也说明，本实施例仅能一定程度上增加区间数值被检索到的概率，但不能实现区间数值的全覆盖检索。并且也可以看出，数值集合所取数值越多，即密度越大，区间数值被检索的概率越大。
- [0059] 本实施例通过补齐区间数值的部分中间值，实现区间数值部分中间值的检索，提高区间数值检索准确率。
- [0060] 在一些实施例的区间数值检索方法中，按照预设算法得到区间数值对应的数值集合，数值集合包括区间数值的中间值。本实施例提供一种预设算法：从区间数值的最小值开始按照预设步长逐步增加取值直至区间数值的最大值，每增加一次预设步长产生一个中间值，由所得所有中间值组成区间数值对应的数值集合。例如区间数值0至9，预设步长为1，则从0开始每次增加1，所得中间值为

1、2、3、4、5、6、7、8，即由中间值1、2、3、4、5、6、7、8组成区间数值0至9对应的数值集合。

[0061] 在一些实施例的区间数值检索方法中，按照预设算法得到区间数值对应的数值集合，数值集合包括区间数值的中间值。本实施例提供一种预设算法：从区间数值的最大值开始按照预设步长逐步减小取值直至区间数值的最小值，每减小一次预设步长产生一个中间值，由所得所有中间值组成区间数值对应的数值集合。例如区间数值0至9，预设步长为1，则从9开始每次减小1，所得中间值为1、2、3、4、5、6、7、8，即由中间值1、2、3、4、5、6、7、8组成区间数值0至9对应的数值集合。

[0062] 在一些实施例的区间数值检索方法中，在步骤a1之后步骤a2之前还包括确定预设步长的过程：获取区间数值对应的数值单位，数值单位通常位于紧邻区间数值后的位置，可将区间数值后紧邻的单位作为该区间数值的数值单位。例如，区间数值0至9℃，则将区间数值0至9后紧邻的单位℃作为该区间数值的数值单位。可以理解的，区间数值和数值单位的位置关系具有多样性，可根据区间数值和数值单位的其他位置关系确定区间数值对应的数值单位，其原理与数值单位位于紧邻区间数值后类似，不再一一说明。

[0063] 进一步，确定区间数值对应的数值单位后，根据第二对应关系查找数值单位对应的预设步长，第二对应关系为数值单位和预设步长的对应关系。可以理解，每个数值单位下对应的数据精度通常是固定的，例如摄氏度℃对应的精度通常为0.1，即用户搜索数值时通常也是精确到0.1℃，此时预设步长可为0.1，在精度0.1下得到的数值集合可满足大部分用户需求。所以第二对应关系的设定需要结合数值单位在大量数据中的普遍精度要求以及工作人员的经验进行设定，每种数值单位对应一个预设步长。

[0064] 本实施例根据区间数值对应的数值单位确定预设步长，使预设步长的取值更加合理，从而使区间数值的中间值取值更加合理，进而增加区间数值检索准确率。

[0065] 在一些实施例的区间数值检索方法中，在步骤a1之后步骤a2之前还包括确定预设步长的过程：获取区间数值对应的数值单位及文档所属的产品信息，其中数

值单位通常位于紧邻区间数值后的位置，可将区间数值后紧邻的单位作为该区间数值的数值单位。例如，区间数值0至9℃，则将区间数值0至9后紧邻的单位℃作为该区间数值的数值单位。可以理解的，区间数值和数值单位的位置关系具有多样性，可根据区间数值和数值单位的其他位置关系确定区间数值对应的数值单位，其原理与数值单位位于紧邻区间数值后类似，不再一一说明。

[0066] 进一步，文档所属的产品信息是指文档所限定或描述或归属产品的产品信息，产品信息包括但不限于产品名称、产品型号、产品领域、产品用途等，能用于描述产品的信息都可以作为产品信息。通过产品信息可以进一步明确区间数值所描述的产品，进而可更加准确的确定预设步长。例如区间数值0至9mm（毫米），如果其对应的产品信息是电脑尺寸，则预设步长可为1mm；如果其对应的产品信息是电脑中内存条的尺寸，则预设步长可为0.1mm。

[0067] 进一步，确定区间数值对应的数值单位和产品信息后，根据第三对应关系查找数值单位和产品信息对应的预设步长，第三对应关系为预设步长与数值单位和产品信息的对应关系。可以理解，本实施例同时考虑数值单位和产品信息来确定预设步长，从数值单位本身和其所所述的对象来确定预设步长，可使所取预设步长更加合理，所以第三对应关系的设定需要结合数值单位在大量数据中的普遍精度要求、同时考虑产品信息对尺寸的精度要求以及工作人员的经验进行设定，每种数值单位和产品信息对应一个预设步长。

[0068] 本实施例根据区间数值对应的数值单位以及产品信息来确定预设步长，相较于只根据数值单位确定预设步长，可使预设步长的取值更加合理，从而使区间数值的中间值取值更加合理，进而增加区间数值检索准确率。

[0069] 在一个优选实施例中，参考图2，本实施例的区间数值检索方法应用于检索领域，包括计算机本地数据检索、服务器数据检索和云端数据检索等。区间数值是指数值的一个区间段，即一个取值范围。区间数值可以是只包含数字的数字范围，例如0至9；也可以是包含数值单位的取值范围，例如0℃（摄氏度）至9℃，0m（米）至9m，0A（安培）至9A等，其中0℃（摄氏度）、m（米）和A（安培）为数值单位。具体的，该区间数值检索方法包括建立数值集合过程和数值检索过程，以下分别进行说明。

[0070] 首先，在检索前要建立数值集合，该数值集合建立过程包括下述步骤：

[0071] 步骤a11、检测文档中的区间数值以及区间数值在文档中的位置信息。

[0072] 具体的，本实施例的文档是指包含一定数据的文件，尤其指包含文字的文件，以便能够用于本实施例的文字检索。可以理解，文件格式表现形式较多，本实施例对此不做限定。作为选择，文档包括但不限于word文档、TXT文档、PDF文档、EXCEL表格、PPT文档、网页等。区间数值通常有特定表达方式，可以使用中间符号“-”、“~”、“（）”、“[]”等来表示，例如使用中间符号表示的区间数值0-9,0~9, (0,9), [0,9]；还可以用表示范围的文字来表达，例如0至9，从0到9等，位于0和9之间等。基于区间数值的特定表达方式，将这些特定表达方式制作为区间数值模型，使用区间数值模型在文档中检测与区间数值模型匹配的区间数值。例如区间数值模型“数字-数字”，如果在文档中检测到符合区间数值模型“数字-数字”的文字，认为是检索到区间数值，例如0-9符合区间数值模型“数字-数字”。通过使用区间数值模型可实现文档中区间数值的自动检测，不需要人工查找文档中的区间数值，效率高，成本低。

[0073] 进一步，在检测到文档中的区间数值后获取该区间数值在文档中的位置信息，位置信息用于定位区间数值在文档中的位置，例如位置信息为第2页第3行，则在第2页第3行能够定位到该区间数值。若文档中包含多个区间数值，则获取每个区间数值对应的位置信息，并建立各个区间数值与位置信息的对应关系。

[0074] 步骤a21、按照预设算法得到区间数值对应的数值集合，数值集合包括区间数值的中间值。

[0075] 具体的，检索到文档中的区间数值后，按照预设算法得到区间数值对应的数值集合，每个区间数值对应一个数值集合，数值集合包括区间数值的中间值。数值集合包括至少一个中间值，中间值大于区间数值的最小值且小于区间数值的最大值。可以理解，某一区间数值对应的数值集合包含的中间值越多，被检索到的概率就越大，从而提高该区间数值被检索到的概率。例如，区间数值0至9，中间值可取1、2、3、4、5、6、7、8，即由中间值1、2、3、4、5、6、7、8组成区间数值0至9对应的数值集合。但需要说明的是，数值集合并不包含区间

数值最小值和最大值之间的所有数值，而是取尽可能多的数值，所取数值越多，即密度越大，区间数值被检索的概率越大。

[0076] 步骤a31、建立数值集合与文档的第一对应关系，以及数值集合与位置信息的第四对应关系。

[0077] 具体的，得到某一区间数值对应的数值集合后，建立数值集合与文档的第一对应关系，即建立数值集合与对应区间数值所在文档的第一对应关系。若文档包含两个以上区间数值，则该文档对应多个数值集合。通常情况下，计算机或服务器或云端数据库存储有大量文档，检索出每个文档中的区间数值，并建立每个文档对应的数值集合，从而建立数据库对应的数值集合。

[0078] 可以理解，建立第一对应关系后数值集合中每个中间值与文档都对应，这样就能实现文档内区间数值的中间值的检索。需要说明的是，文档中某一区间数值对应的数值集合为文档之外新产生数据，并不属于文档本身，也不需要修改文档，而是作为单独的数据另外存储，然后建立数值集合与文档的第一对应关系。

[0079] 进一步，建立数值集合与文档的第一对应关系后，根据第一对应确定数值集合对应的文档。为使用户能够进一步定位到区间数值所在文档内的具体位置，本实施例同时建立数值集合与位置信息的第四对应关系，根据第四对应关系可确定数值集合在文档中的具体位置。

[0080] 其次，在建立数值集合后进行数值检索，该数值检索过程包括下述步骤：

[0081] 步骤b11、接收检索信息，检索信息包括待检索数值。

[0082] 具体的，接收用户输入的检索信息，该检索信息包括待检索数值，当然还可以包括其他检索信息。需要说明的是，该待检索数值为数值或数字。

[0083] 步骤b21、在数值集合中检索待检索数值。

[0084] 具体的，因数值集合中所有中间值都已是具体数值，能够通过“所见即所得”的文字匹配方式进行检索，所以可在数值集合中检索待检索数值。例如区间数值0至9，中间值可取1、2、3、4、5、6、7、8，即由中间值1、2、3、4、5、6、7、8组成区间数值0至9对应的数值集合。若用户输入的待检索数值为3，则在数值集合1、2、3、4、5、6、7、8中检索待检索数值3。可以理解

的，待检索数值依然可用于文档中已有内容的检索，可参考现有技术，此处不再赘述。

[0085] 步骤b31、若待检索数值在数值集合中，则根据第一对应关系确定检索信息对应的文档，以及根据第四对应关系确定检索信息在文档中的位置。

[0086] 具体的，通过在数值集合中检索待检索数值，若待检索数值在数值集合中，则根据第一对应关系确定检索信息对应的文档，从而实现区间数值未直接显示部分的检索。例如区间数值0至9，中间值可取1、2、3、4、5、6、7、8，即由中间值1、2、3、4、5、6、7、8组成区间数值0至9对应的数值集合。若用户输入的待检索数值为3，则在数值集合1、2、3、4、5、6、7、8中检索待检索数值3，刚好能匹配到，则可确定待检索数值3对应的文档。

[0087] 进一步，根据第一对应关系确定检索信息对应的文档后，再根据第四对应关系确定检索信息在文档中的位置。例如区间数值0至9对应的位置信息为第2页第3行，则在第2页第3行能够定位到区间数值0至9。也就是说，如果用户输入的待检索数值为3，则在数值集合1、2、3、4、5、6、7、8中检索待检索数值3，刚好能匹配到，则可确定待检索数值3对应的文档。进一步，区间数值0至9对应的位置信息为第2页第3行，则进一步定位待检索数值3对应的区间数值0至9位于文档的第2页第3行。

[0088] 本实施例通过补齐区间数值的部分中间值，实现区间数值部分中间值的检索，提高区间数值检索准确率。并且在检索文档中区间数值时记录区间数值在文档的位置，从而在检索结果中直接显示区间数值在文档中的位置，方便用户查看，提高用户使用体验。

[0089] 在一个优选实施例中，参考图3，本实施例的区间数值检索方法应用于检索领域，包括计算机本地数据检索、服务器数据检索和云端数据检索等。区间数值是指数值的一个区间段，即一个取值范围。区间数值可以是只包含数字的数字范围，例如0至9；也可以是包含数值单位的取值范围，例如0℃（摄氏度）至9℃，0m（米）至9m，0A（安培）至9A等，其中0℃（摄氏度）、m（米）和A（安培）为数值单位。具体的，该区间数值检索方法包括建立数值集合过程和数值检索过程，以下分别进行说明。

[0090] 首先，在检索前要建立数值集合，该数值集合建立过程包括下述步骤：

[0091] 步骤a12、检测文档中的区间数值及区间数值对应的数值单位。

[0092] 具体的，本实施例的文档是指包含一定数据的文件，尤其指包含文字的文件，以便能够用于本实施例的文字检索。可以理解，文件格式表现形式较多，本实施例对此不做限定。作为选择，文档包括但不限于word文档、TXT文档、PDF文档、EXCEL表格、PPT文档、网页等。区间数值通常有特定表达方式，可以使用中间符号“-”、“~”、“（ ）”、“[]”等来表示，例如使用中间符号表示的区间数值0-9,0~9, (0,9), [0,9]；还可以用表示范围的文字来表达，例如0至9, 从0到9等，位于0和9之间等。基于区间数值的特定表达方式，将这些特定表达方式制作为区间数值模型，使用区间数值模型在文档中检测与区间数值模型匹配的区间数值。例如区间数值模型“数字-数字”，如果在文档中检测到符合区间数值模型“数字-数字”的文字，认为是检索到区间数值，例如0-9符合区间数值模型“数字-数字”。通过使用区间数值模型可实现文档中区间数值的自动检测，不需要人工查找文档中的区间数值，效率高，成本低。

[0093] 进一步，检测文档中的区间数值后，数值单位通常位于紧邻区间数值后的位置，可将区间数值后紧邻的单位作为该区间数值的数值单位。例如，区间数值0至9℃，则将区间数值0至9后紧邻的单位℃作为该区间数值的数值单位。可以理解的，区间数值和数值单位的位置关系具有多样性，可根据区间数值和数值单位的其他位置关系确定区间数值对应的数值单位，其原理与数值单位位于紧邻区间数值后类似，不再一一说明。

[0094] 步骤a22、按照预设算法得到区间数值对应的数值集合，数值集合包括区间数值的中间值和数值单位。

[0095] 具体的，检索到文档中的区间数值后，按照预设算法得到区间数值对应的数值集合，每个区间数值对应一个数值集合，数值集合包括区间数值的中间值和数值单位。数值集合包括至少一个中间值，中间值大于区间数值的最小值且小于区间数值的最大值。可以理解，某一区间数值对应的数值集合包含的中间值越多，被检索到的概率就越大，从而提高该区间数值被检索到的概率。例如，区间数值0至9℃，中间值可取1、2、3、4、5、6、7、8，即由中间值

1、2、3、4、5、6、7、8以及数值单位℃组成区间数值0至9℃对应的数值集合。但需要说明的是，数值集合并不包含区间数值最小值和最大值之间的所有数值，而是取尽可能多的数值，所取数值越多，即密度越大，区间数值被检索的概率越大。

[0096] 步骤a32、建立数值集合与文档的第一对应关系。

[0097] 具体的，得到某一区间数值对应的数值集合后，建立数值集合与文档的第一对应关系，即建立数值集合与对应区间数值所在文档的第一对应关系。若文档包含两个以上区间数值，则该文档对应多个数值集合。通常情况下，计算机或服务器或云端数据库存储有大量文档，检索出每个文档中的区间数值，并建立每个文档对应的数值集合，从而建立数据库对应的数值集合。

[0098] 可以理解，建立第一对应关系后数值集合中每个中间值与文档都对应，这样就能实现文档内区间数值的中间值的检索。需要说明的是，文档中某一区间数值对应的数值集合为文档之外新产生数据，并不属于文档本身，也不需要修改文档，而是作为单独的数据另外存储，然后建立数值集合与文档的第一对应关系。

[0099] 其次，在建立数值集合后进行数值检索，该数值检索过程包括下述步骤：

[0100] 步骤b12、接收检索信息，检索信息包括待检索数值和数值单位。

[0101] 具体的，接收用户输入的检索信息，该检索信息包括待检索数值和数值单位，当然还可以包括其他检索信息。需要说明的是，该待检索数值为数值或数字，数值单位为数学方面或物理方面计量事物的标准量的名称。例如户输入的检索信息为3℃，包括待检索数值3和数值单位℃。

[0102] 步骤b22、在数值集合中检索待检索数值和数值单位。

[0103] 具体的，因数值集合中所有中间值都已是具体数值以及对应的数值单位，能够通过“所见即所得”的文字匹配方式进行检索，所以可在数值集合中检索待检索数值和数值单位。例如区间数值0至9℃，中间值可取1、2、3、4、5、6、7、8，即由中间值1、2、3、4、5、6、7、8和数值单位℃组成区间数值0至9℃对应的数值集合。若用户输入的待检索数值为3℃，则在数值集合1、2、3、4、5、6、7、8和数值单位℃中检索待检索数值3℃。可以理解的，

待检索数值依然可用于文档中已有内容的检索，可参考现有技术，此处不再赘述。

[0104] 步骤b32、若待检索数值和数值单位在同一数值集合中，则根据第一对应关系确定检索信息对应的文档。

[0105] 具体的，通过在数值集合中检索待检索数值和数值单位，若待检索数值和数值单位在数值集合中，则根据第一对应关系确定检索信息对应的文档，从而实现区间数值未直接显示部分的检索。例如区间数值0至9℃，中间值可取1、2、3、4、5、6、7、8，即由中间值1、2、3、4、5、6、7、8和数值单位℃组成区间数值0至9℃对应的数值集合。若用户输入的待检索数值为3℃，则在数值集合1、2、3、4、5、6、7、8和数值单位℃中检索待检索数值3和数值单位℃，刚好能同时匹配到，则可确定待检索数值3℃对应的文档。

[0106] 另外，假如用户输入的待检索数值是3.5℃，则在数值集合1、2、3、4、5、6、7、8和数值单位℃中不能匹配成功，但3.5℃确实也在区间数值0至9中，此处也说明，本实施例仅能一定程度上增加区间数值被检索到的概率，但不能实现区间数值的全覆盖检索。并且也可以看出，数值集合所取数值越多，即密度越大，区间数值被检索的概率越大。

[0107] 本实施例通过补齐区间数值的部分中间值，实现区间数值部分中间值的检索，并且使用数值和数值单位两个参数同时进行检索，提高区间数值检索准确率。

[0108] 在一些实施例的区间数值检索方法中，在步骤b3之后还包括步骤：可以理解，因数据库中包括多个文档，每个文档都有对应的数值集合，且一个文档可能对应多个数值集合，不同数值集合之间可能存在相同中间值的情况，所以根据第一对应关系确定检索信息对应的文档的数量可能超过一个。若根据第一对应关系确定检索信息对应的文档的数量超过一个，则按照预设排序规则对所得文档进行排序并显示，预设排序规则可参考现有技术。本实施例出现多个检索结果的情况进行排序显示，方便用户查看检索结果。

[0109] 在一个优选实施例中，本实施例的计算机可读存储介质存储有计算机程序，计算机程序适于处理器进行加载，以执行如上述实施例的区间数值检索方法的步骤。

- [0110] 在一个优选实施例中，本实施例的计算机包括存储器和处理器，存储器中存储有计算机程序，处理器通过调用存储器中存储的计算机程序，执行如上述实施例的区间数值检索方法的步骤。
- [0111] 本说明书中各个实施例采用递进的方式描述，每个实施例重点说明的都是与其他实施例的不同之处，各个实施例之间相同相似部分互相参见即可。对于实施例公开的装置而言，由于其与实施例公开的方法相对应，所以描述的比较简单，相关之处参见方法部分说明即可。
- [0112] 专业人员还可以进一步意识到，结合本文中所公开的实施例描述的各示例的单元及算法步骤，能够以电子硬件、计算机软件或者二者的结合来实现，为了清楚地说明硬件和软件的可互换性，在上述说明中已经按照功能一般性地描述了各示例的组成及步骤。这些功能究竟以硬件还是软件方式来执行，取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能，但是这种实现不应认为超出本发明的范围。
- [0113] 结合本文中所公开的实施例描述的方法或算法的步骤可以直接用硬件、处理器执行的软件模块，或者二者的结合来实施。软件模块可以置于随机存储器（RAM）、内存、只读存储器（ROM）、电可编程ROM、电可擦除可编程ROM、寄存器、硬盘、可移动磁盘、CD-ROM、或技术领域内所公知的任意其它形式的存储介质中。
- [0114] 以上实施例只为说明本发明的技术构思及特点，其目的在于让熟悉此项技术的人士能够了解本发明的内容并据此实施，并不能限制本发明的保护范围。凡跟本发明权利要求范围所做的均等变化与修饰，均应属于本发明权利要求的涵盖范围。

权利要求书

- [权利要求 1] 一种区间数值检索方法，其特征在于，包括建立数值集合过程和数值检索过程；
- 所述建立数值集合包括：
- 步骤a1、检测文档中的区间数值；
- 步骤a2、按照预设算法得到所述区间数值对应的数值集合，所述数值集合包括所述区间数值的中间值；
- 步骤a3、建立所述数值集合与所述文档的第一对应关系；
- 所述数值检索过程包括：
- 步骤b1、接收检索信息，所述检索信息包括待检索数值；
- 步骤b2、在所述数值集合中检索所述待检索数值；
- 步骤b3、若所述待检索数值在所述数值集合中，则根据所述第一对应关系确定所述检索信息对应的文档。
- [权利要求 2] 根据权利要求1所述的区间数值检索方法，其特征在于，所述数值集合包括至少一个中间值，所述中间值大于所述区间数值的最小值且小于所述区间数值的最大值。
- [权利要求 3] 根据权利要求1所述的区间数值检索方法，其特征在于，所述步骤a2包括：
- 从所述区间数值的最小值开始按照预设步长逐步增加取值直至所述区间数值的最大值，每增加一次所述预设步长产生一个中间值，由所得所有中间值组成所述区间数值对应的数值集合；或者
- 从所述区间数值的最大值开始按照预设步长逐步减小取值直至所述区间数值的最小值，每减小一次所述预设步长产生一个中间值，由所得所有中间值组成所述区间数值对应的数值集合。
- [权利要求 4] 根据权利要求3所述的区间数值检索方法，其特征在于，在所述步骤a1之后所述步骤a2之前还包括：
- 获取所述区间数值对应的数值单位；

根据第二对应关系查找所述数值单位对应的预设步长，所述第二对应关系为所述数值单位和所述预设步长的对应关系。

[权利要求 5]

根据权利要求3所述的区间数值检索方法，其特征在于，在所述步骤a1之后所述步骤a2之前还包括：

获取所述区间数值对应的数值单位及所述文档所属的产品信息；
根据第三对应关系查找所述数值单位和所述产品信息对应的预设步长，所述第三对应关系为所述预设步长与所述数值单位和产品信息
的对应关系。

[权利要求 6]

根据权利要求1所述的区间数值检索方法，其特征在于，所述步骤a1包括：检测文档中的区间数值以及所述区间数值在所述文档中的位置信息；
所述步骤a3包括：建立所述数值集合与所述文档的第一对应关系，以及所述数值集合与所述位置信息的第四对应关系；
所述步骤b3包括：若所述待检索数值在所述数值集合中，则根据所述第一对应关系确定所述检索信息对应的文档，以及根据所述第四对应关系确定所述检索信息在所述文档中的位置。

[权利要求 7]

根据权利要求1所述的区间数值检索方法，其特征在于，所述步骤a1包括：检测文档中的区间数值及所述区间数值对应的数值单位；
所述步骤a2中所述数值集合包括所述区间数值的中间值和所述数值单位；
所述步骤b1中所述检索信息包括待检索数值和数值单位；
所述步骤b2包括：在所述数值集合中检索所述待检索数值和所述数值单位；
所述步骤b3中若所述待检索数值在所述数值集合中包括：若所述待检索数值和所述数值单位在同一所述数值集合中。

[权利要求 8]

根据权利要求1所述的区间数值检索方法，其特征在于，在所述步骤b3之后还包括：

若根据所述第一对应关系确定所述检索信息对应的文档的数量超过一个，则按照预设排序规则对所得文档进行排序并显示。

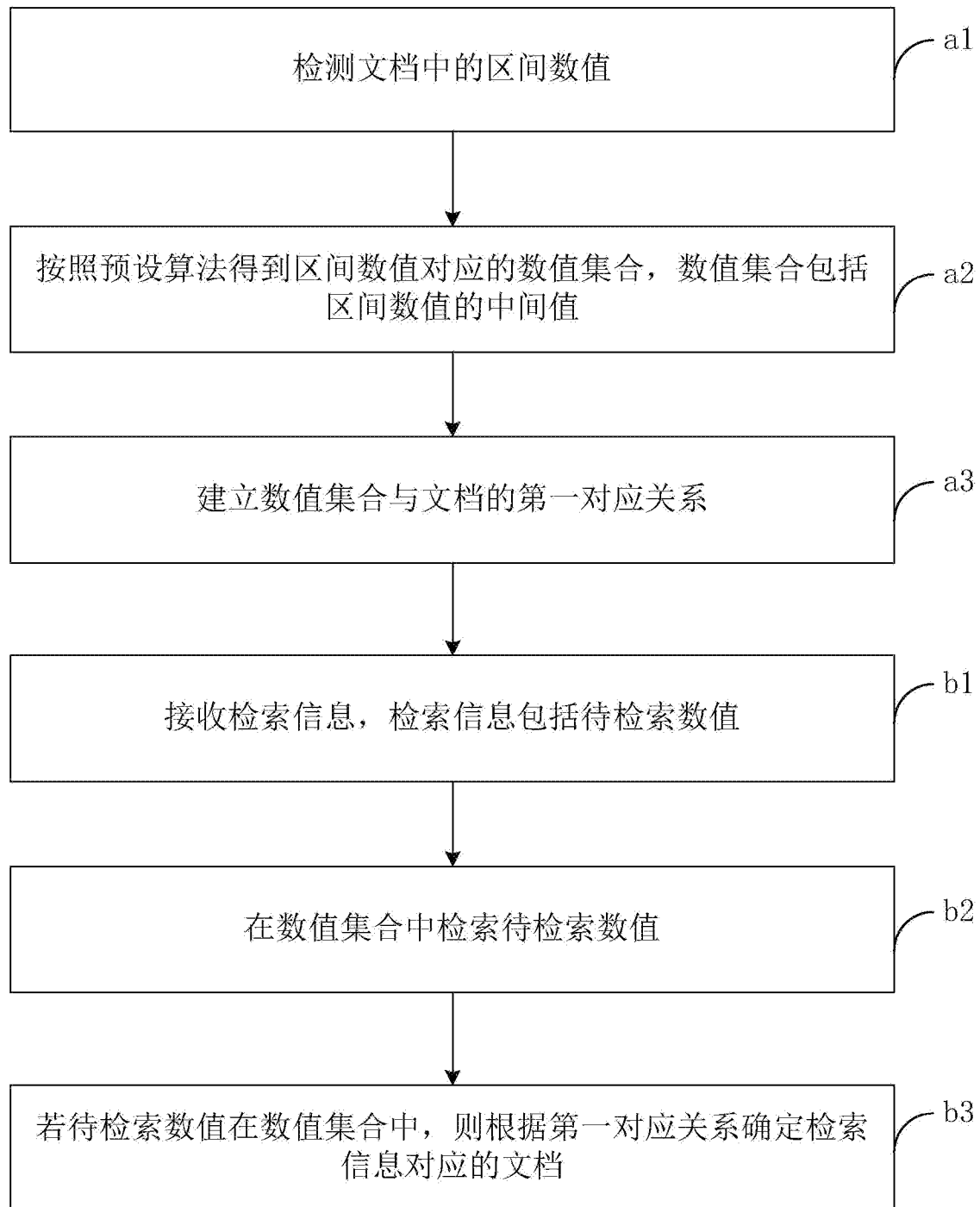
[权利要求 9]

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存储有计算机程序，所述计算机程序适于处理器进行加载，以执行如权利要求1至8任一项所述的区间数值检索方法的步骤。

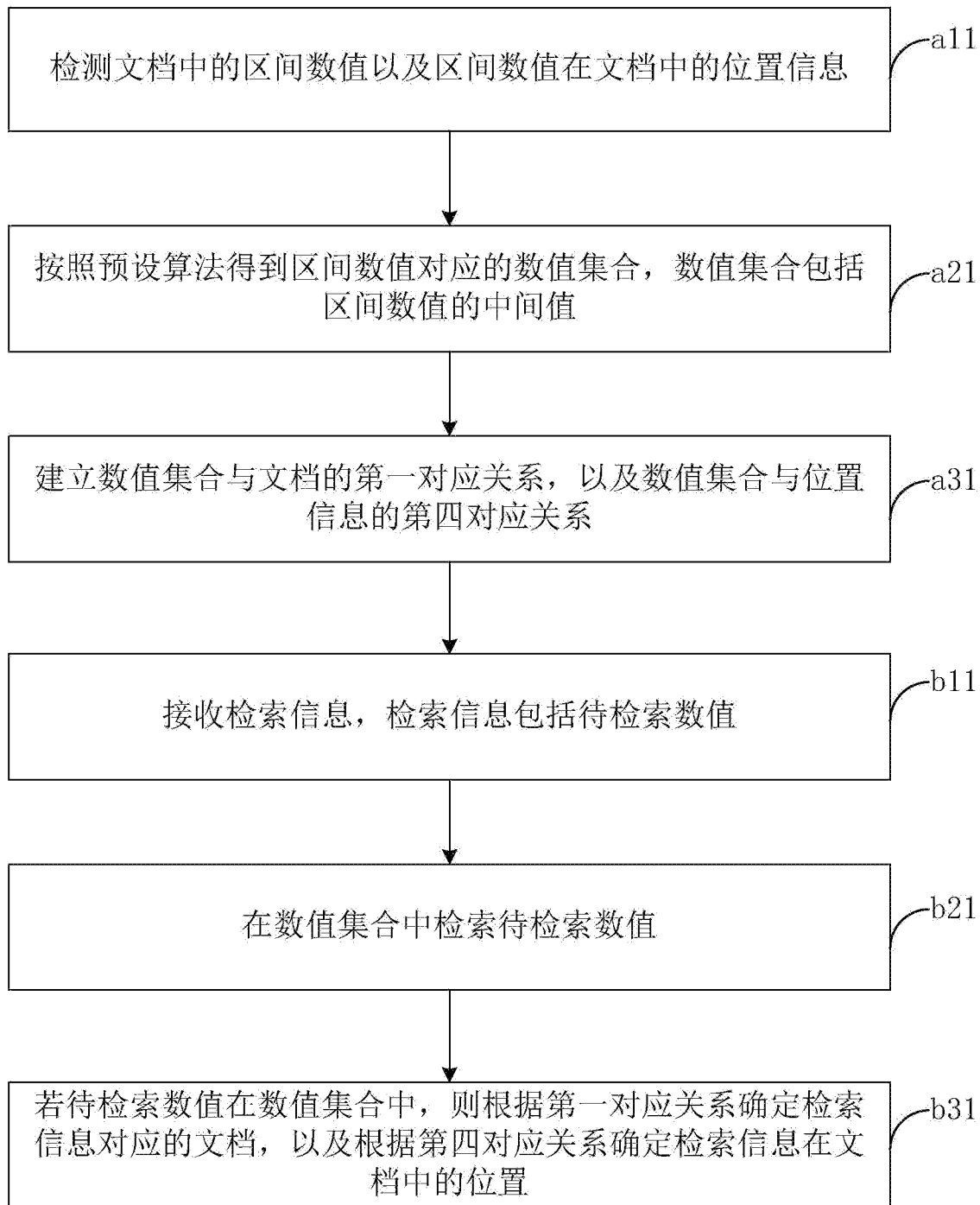
[权利要求 10]

一种计算机，其特征在于，包括存储器和处理器，所述存储器中存储有计算机程序，所述处理器通过调用所述存储器中存储的所述计算机程序，执行如权利要求1至8任一项所述的区间数值检索方法的步骤。

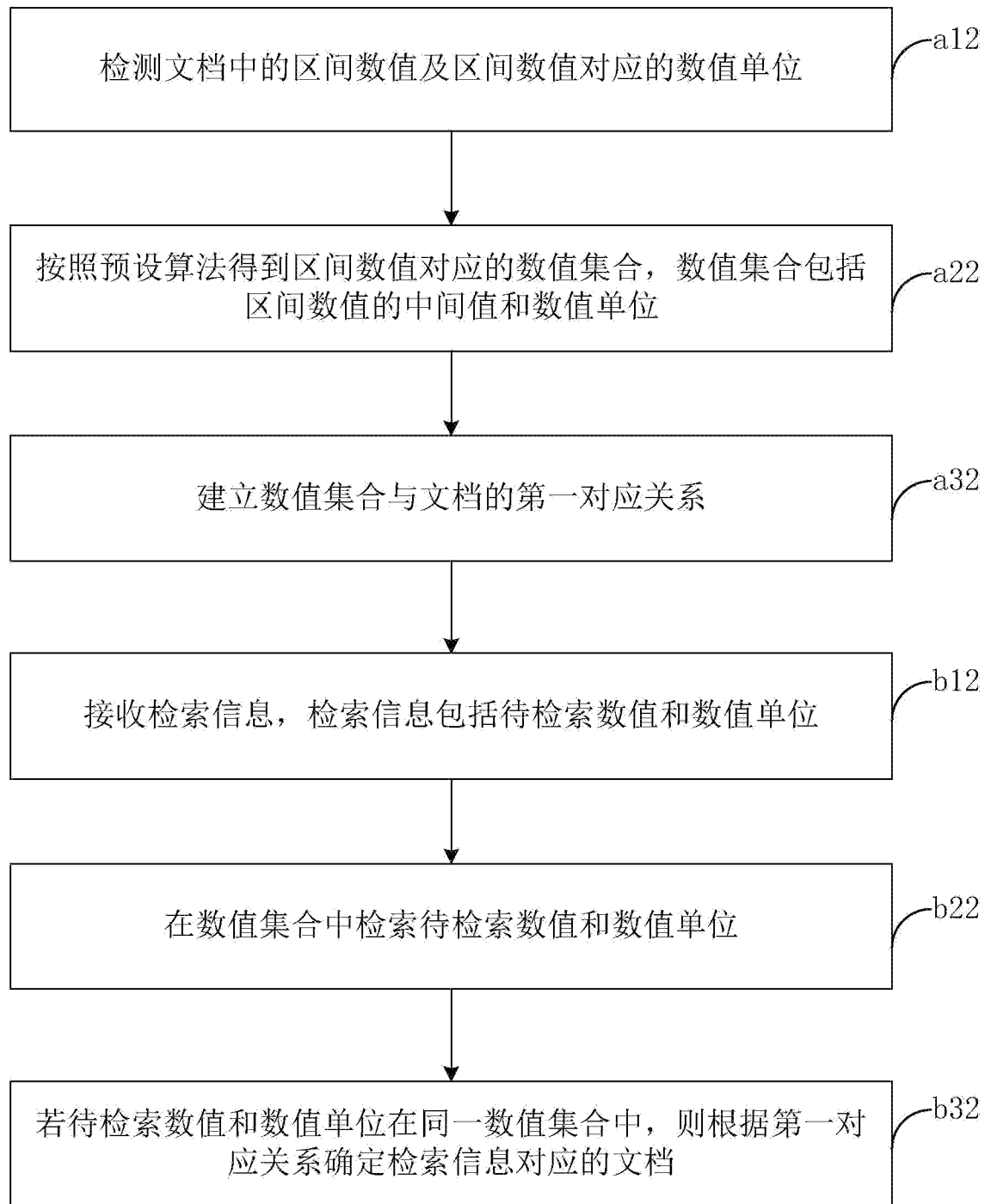
[图1]



[图2]



[图3]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2024/098256

A. CLASSIFICATION OF SUBJECT MATTER G06F 16/2458(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) IPC: 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) CNABS; CNTXT; DWPI; VEN; USTXT; WOTXT; EPTXT; CNKI; IEEE: 数值, 范围, 区间, 查询, 最大值, 最小值, number, range, interval, query, maximum, minimum		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 116860828 A (SHENZHEN SEKORM COMPONENT NETWORK CO., LTD.) 10 October 2023 (2023-10-10) claims 1-10, and description, paragraphs 43-109	1-10
X	CN 101295307 A (HITACHI, LTD.) 29 October 2008 (2008-10-29) description, pages 2-4 and 6-12, and figures 2 and 11	1-10
A	US 2005198027 A1 (OKI ELECTRIC INDUSTRY CO., LTD.) 08 September 2005 (2005-09-08) entire document	1-10
A	US 2014156670 A1 (NEC CORP.) 05 June 2014 (2014-06-05) entire document	1-10
☐ 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 “D” document cited by the applicant in the international application “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 23 August 2024		Date of mailing of the international search report 03 September 2024
Name and mailing address of the ISA/CN China National Intellectual Property Administration (ISA/CN) China No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088		Authorized officer Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2024/098256

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	116860828	A	10 October 2023	None			
CN	101295307	A	29 October 2008	US	2008270386	A1	30 October 2008
				US	8046368	B2	25 October 2011
				JP	2008276550	A	13 November 2008
				JP	5154832	B2	27 February 2013
				CN	101295307	B	04 August 2010
US	2005198027	A1	08 September 2005	US	7509314	B2	24 March 2009
				JP	2005250980	A	15 September 2005
US	2014156670	A1	05 June 2014	WO	2012176374	A1	27 December 2012
				JPWO	2012176374	A1	23 February 2015
				JP	5924339	B2	25 May 2016
				US	9465838	B2	11 October 2016

A. 主题的分类

G06F 16/2458(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)

IPC: 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))

CNABS; CNTXT; DWPI; VEN; USTXT; WOTXT; EPTXT; CNKI; IEEE: 数值, 范围, 区间, 查询, 最大值, 最小值, number, range, interval, query, maximum, minimum

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 116860828 A (深圳市世强元件网络有限公司) 2023年10月10日 (2023 - 10 - 10) 权利要求1-10, 说明书第43-109段	1-10
X	CN 101295307 A (株式会社日立制作所) 2008年10月29日 (2008 - 10 - 29) 说明书第2 -4、6-12 页, 图2、11	1-10
A	US 2005198027 A1 (OKI ELECTRIC IND CO LTD) 2005年9月8日 (2005 - 09 - 08) 全文	1-10
A	US 2014156670 A1 (NEC CORPORATION) 2014年6月5日 (2014 - 06 - 05) 全文	1-10

☐ 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

“D” 申请人在国际申请中引证的文件

“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

2024年8月23日

国际检索报告邮寄日期

2024年9月3日

ISA/CN的名称和邮寄地址

中国国家知识产权局
中国北京市海淀区蓟门桥西土城路6号 100088

授权官员

程潇杰

电话号码 (+86) 020-28957217

PCT/ISA/210 表(第2页) (2022年7月)

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2024/098256

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	116860828	A	2023年10月10日	无			
CN	101295307	A	2008年10月29日	US	2008270386	A1	2008年10月30日
				US	8046368	B2	2011年10月25日
				JP	2008276550	A	2008年11月13日
				JP	5154832	B2	2013年2月27日
				CN	101295307	B	2010年8月4日
US	2005198027	A1	2005年9月8日	US	7509314	B2	2009年3月24日
				JP	2005250980	A	2005年9月15日
US	2014156670	A1	2014年6月5日	WO	2012176374	A1	2012年12月27日
				JPWO	2012176374	A1	2015年2月23日
				JP	5924339	B2	2016年5月25日
				US	9465838	B2	2016年10月11日