

(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
2019年12月26日 (26.12.2019)

(10) 国际公布号
WO 2019/242124 A1

- (51) 国际专利分类号:
G06F 17/24 (2006.01)
- (21) 国际申请号: PCT/CN2018/105533
- (22) 国际申请日: 2018年9月13日 (13.09.2018)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201810630791.4 2018年6月19日 (19.06.2018) CN
- (71) 申请人: 平安科技(深圳)有限公司 (PING AN TECHNOLOGY(SHENZHEN)CO., LTD.) [CN/CN]; 中国广东省深圳市福田区福田街道福安社区益田路5033号平安金融中心23楼, Guangdong 518000 (CN)。
- (72) 发明人: 苏晓明(SU, Xiaoming); 中国广东省深圳市福田区福田街道福安社区益田路5033号平

安金融中心23楼, Guangdong 518000 (CN)。汪伟(WANG, Wei); 中国广东省深圳市福田区福田街道福安社区益田路5033号平安金融中心23楼, Guangdong 518000 (CN)。王晓伟(WANG, Xiaowei); 中国广东省深圳市福田区福田街道福安社区益田路5033号平安金融中心23楼, Guangdong 518000 (CN)。刘媛源(LIU, Yuanyuan); 中国广东省深圳市福田区福田街道福安社区益田路5033号平安金融中心23楼, Guangdong 518000 (CN)。肖京(XIAO, Jing); 中国广东省深圳市福田区福田街道福安社区益田路5033号平安金融中心23楼, Guangdong 518000 (CN)。

(74) 代理人: 深圳中一专利商标事务所(SHENZHEN ZHONGYI PATENT AND TRADEMARK OFFICE); 中国广东省深圳市福田区深南中路1014号老特区报社四楼(5号信箱), Guangdong 518028 (CN)。

(54) Title: SUM OF MONEY INFORMATION EXTRACTION METHOD AND APPARATUS, AND TERMINAL DEVICE AND MEDIUM

(54) 发明名称: 金额信息的提取方法、装置、终端设备及介质

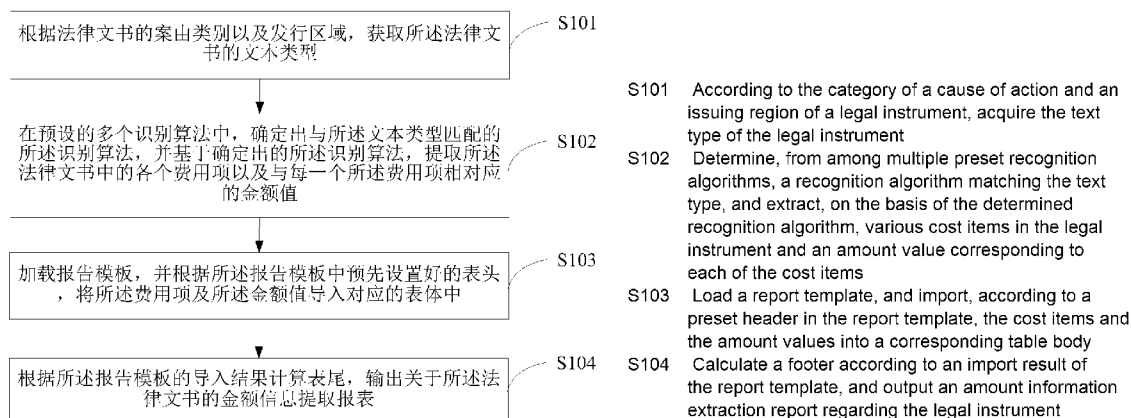


图 1

(57) Abstract: Provided are a sum of money information extraction method and apparatus, and a terminal device and a medium, wherein same are applicable in the technical field of information processing. The method comprises: according to the category of a cause of action and an issuing region of a legal instrument, acquiring the text type of the legal instrument; determining, from among multiple preset recognition algorithms, a recognition algorithm matching the text type, and extracting, on the basis of the determined recognition algorithm, various cost items in the legal instrument and a sum of money value corresponding to each of the cost items; loading a report template, and importing, according to a preset header in the report template, the cost items and the sum of money values into a corresponding table body; and calculating a footer according to an import result of the report template, and outputting a sum of money information extraction report regarding the legal instrument. The solution can use the optimal and the most targeted recognition algorithm to accurately extract sum of money data information from a legal text, and a user does not need to artificially recognize a legal instrument, thereby improving the sum of money information extraction accuracy and extraction efficiency.



WO 2019/242124 A1

(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, 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: 本方案提供了一种金额信息的提取方法、装置、终端设备及介质, 适用于信息处理技术领域, 该方法包括: 根据法律文书的案由类别以及发行区域, 获取法律文书的文本类型; 在预设的多个识别算法中, 确定出与文本类型匹配的识别算法, 并基于确定出的识别算法, 提取法律文书中的各个费用项以及与每一个费用项相对应的金额值; 加载报告模板, 并根据报告模板中预先设置好的表头, 将费用项及金额值导入对应的表体中; 根据报告模板的导入结果计算表尾, 输出关于法律文书的金额信息提取报表。本方案能够利用最佳且最具有针对性的一种识别算法来准确地提取法律文本中的金额数据信息, 用户无需对法律文书进行人工识别, 故提高了金额信息的提取准确率及提取效率。

金额信息的提取方法、装置、终端设备及介质

[0001] 本申请要求于2018年06月19日提交中国专利局、申请号为201810630791.4、发明名称为“金额信息的提取方法、终端设备及介质”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] 本申请属于信息处理技术领域，尤其涉及一种金额信息的提取方法、装置、终端设备及计算机可读存储介质。

背景技术

[0003] 司法行政机关及当事人、律师等在解决诉讼和非讼案件时所使用的文书为法律文书。法律文书往往记载有不同费用项的原告诉请金额、判决金额等金额数据信息。这些金额数据信息在自动断案、案件预判、律师画像、律师胜诉率分析以及判决合理性分析等方面均起着极为重要的参考作用。

[0004] 然而，基于法律文书的描述多样性，不同机关所对应发布的法律文书的样式差异较大，因此，若要提取法律文书中的金额数据信息，则目前用户也只能通过人工读取的方式来实现。然而，当法律文书所记载的费用项较多时，用户依次遍历法律文书中的各个字符需要耗费较多的时间，由此使得金额信息的提取效率较为低下。并且，在法律文书的读取过程中，或多或少地会出现人工识读出错的情况，因而也降低了金额信息的识别准确率。

发明概述

技术问题

[0005] 有鉴于此，本申请实施例提供了一种金额信息的提取方法、装置、终端设备及介质，以解决现有技术中金额信息的识别准确率以及提取效率均较为低下的问题。

问题的解决方案

技术解决方案

[0006] 本申请实施例的第一方面提供了一种金额信息的提取方法，包括：

- [0007] 根据法律文书的案由类别以及发行区域，获取所述法律文书的文本类型；
- [0008] 在预设的多个识别算法中，确定出与所述文本类型匹配的所述识别算法，并基于确定出的所述识别算法，提取所述法律文书中的各个费用项以及与每一个所述费用项相对应的金额值；
- [0009] 加载报告模板，并根据所述报告模板中预先设置好的表头，将所述费用项及所述金额值导入对应的表体中；
- [0010] 根据所述报告模板的导入结果计算表尾，输出关于所述法律文书的金额信息提取报表。
- [0011] 本申请实施例的第二方面提供了一种金额信息的提取装置，所述监控装置包括用于执行上述第一方面所述的金额信息的提取方法的单元。
- [0012] 本申请实施例的第三方面提供了一种终端设备，包括存储器以及处理器，所述存储器中存储有可在所述处理器上运行的计算机可读指令，所述处理器执行所述计算机可读指令时实现如第一方面所述的金额信息的提取方法的步骤。
- [0013] 本申请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存储有计算机可读指令，所述计算机可读指令被处理器执行时实现如第一方面所述的金额信息的提取方法的步骤。

发明的有益效果

有益效果

- [0014] 本申请实施例中，由于不同案由类别以及不同地区所发行的法律文书的样式差异较大，故根据法律文书的案由类别以及发行区域，确定出法律文书的文本类型后，再从预设的多种识别算法中获取与该文本类型匹配的识别算法，保证了能够利用最佳且最具有针对性的一种识别算法来准确地提取法律文本中的金额数据信息，提高了金额信息的提取准确率，同时也为案件预判、律师画像、律师胜诉率分析、判决合理性分析等应用提供了更具参考价值的基础数据；另外，由于本申请实施例能够通过预设的识别算法以及报告模板来自动导出法律文书中各费用项以及金额值的对应关系，因而对于用户而言，其能够直观地查看到自己所需的金额信息，无需再对法律文书进行人工识别，因此，提高了金额信息的提取效率。

对附图的简要说明

附图说明

- [0015] 图1是本申请实施例提供的金额信息的提取方法的实现流程图；
- [0016] 图2是本申请实施例提供的金额信息的提取方法S102的具体实现流程图；
- [0017] 图3是本申请实施例提供的金额信息的提取方法S102的另一具体实现流程图；
- [0018] 图4是本申请另一实施例提供的金额信息的提取方法S102的具体实现流程图；
- [0019] 图5是本申请又一实施例提供的金额信息的提取方法的实现流程图；
- [0020] 图6是本申请实施例提供的金额信息的提取装置的结构框图；
- [0021] 图7是本申请实施例提供的终端设备的示意图。

发明实施例

本发明的实施方式

- [0022] 为了说明本申请所述的技术方案，下面通过具体实施例来进行说明。
- [0023] 图1示出了本申请实施例提供的金额信息的提取方法的实现流程，该方法流程包括步骤S101至S104。各步骤的具体实现原理如下：
 - [0024] S101：根据法律文书的案由类别以及发行区域，获取所述法律文书的文本类型。
- [0025] 本申请实施例中，司法机关及当事人、律师等在解决诉讼和非讼案件时所使用的文书为法律文书，法律文书包括包括司法机关的规范性文件以及非规范性文件。根据预设的网站信息，定期从对应的公开网站中下载上述法律文书。
- [0026] 案由表示人民法院对诉讼案件所涉及的法律关系的性质进行概括后形成的案件名称。案由类别包括但不限于人格权纠纷、婚姻家庭继承纠纷、物权纠纷、债权债务纠纷、劳动争议与人事争议纠纷以及知识产权纠纷等。发行区域表示发行法律文书的司法机关所属的行政区域。通过预设的识别算法来对法律文书进行检测处理，可识别出该法律文书中所记载的案由类别以及发行区域。
- [0027] 本申请实施例中，在预设的信息表中，记录有案由类别、发行区域以及文本类型的映射关系。因此，根据当前法律文书的案由类别以及发行区域，读取对应该法律文书的文本类型。
- [0028] S102：在预设的多个识别算法中，确定出与所述文本类型匹配的所述识别算法

，并基于确定出的所述识别算法，提取所述法律文书中的各个费用项以及与每一个所述费用项相对应的金额值。

[0029] 由于不同文本类型的法律文书具有不同的文本样式风格，而每一种文本样式风格具有相对规律的文本特点，例如，第一文本类型的法律文书中，各个费用项的金额值通常展示于费用项的具体名称之后；第二文本类型的法律文书中，各个费用项的金额值通常展示于费用项的具体名称之前等。因此，针对每一类文本类型，获取开发人员所针对性设置的识别算法，并将各类识别算法与文本类型的匹配关系进行关联存储。

[0030] 本申请实施例中，根据上述S101中所获取到的法律文书的文本类型，查找与该文本类型匹配的识别算法，以基于查找出的识别算法，提取法律文书中的各个费用项以及与每一个费用项对应的金额值信息。

[0031] 作为本申请的一个实施例，当查找出的识别算法为第一算法时，图2示出了本申请实施例提供的金额信息的提取方法S102的具体实现流程，详述如下：

[0032] S1021：获取与费用项描述信息相关联的第一正则表达式。

[0033] 本申请实施例中，费用项描述信息包括各个费用项的名称或者与费用项名称相关的描述性信息。通过提取各费用项描述信息的模式规则，创建与费用项描述信息相关联的第一正则表达式。其中，费用项例为法律文书中原告告诉的各个赔偿项以及实际判决的各个赔偿项，例如可以是医药费、精神损失费以及衣物损失费等。每一个费用项所对应的金额值即各个赔偿项的具体金额。

[0034] S1022：在所述法律文书中，定位与所述第一正则表达式匹配的各个所述费用项。

[0035] 本申请实施例中，基于上述第一正则表达式，对法律文书所包含的各个字符进行依次检测处理。若法律文书中存在与第一正则表达式匹配的字符串，则将该字符串所处的文本位置确定为一个费用项所属的定位。

[0036] S1023：对定位出的每一个所述费用项，查找所述法律文书中出现在该费用项之后的首个金额值，并将该费用项以及该金额值的对应关系进行提取。

[0037] 以确定出的每一费用项所属的文本位置为起点，通过预设的第二正则表达式，检测出出现在该文本位置之后的金额值。其中，第二正则表达式用于描述与金

额值关联的规则模式。例如，预设的第二正则表达式为(?:[0-9]*[,]*[0-9])+\.\?[0-9]*(?:元|万元|百万元|千万元)，则基于该第二正则表达式，可识别出法律文书中包含任意数字以及价格单位的字符串，将该字符串确定出当前所需提取的一个金额值。

[0038] 本申请实施例中，在每一个费用项所属的文本位置之后，若检测到匹配第二正则表达式的字符串，则将该字符串识别为与该费用项对应的金额值，并将二者的对应关系进行记录。此后，读取上述定位出的下一个费用项，并返回执行上述步骤S1023。

[0039] 作为本申请的一个实施示例，上述步骤S1021至S1023例如可以是：

[0040] 加载预先建立的一个正则文件，其中，该正则文件包括以下多个第一正则表达式，每一个第一正则表达式与一个费用项对应：

[0041] 费用项第一正则表达式

[0042] 残疾赔偿金 (?:伤残|残疾).\{0,5\}?(?:赔偿|补偿)

[0043] 残疾器具费 (?:伤残|残疾).\{0,5\}?(?:器具|用具)

[0044] 死亡赔偿金 (?:死亡|亡故).\{0,5\}?(?:赔偿|补偿)

[0045] 医疗费 (?:医疗|医药).\{0,5\}?(?:赔偿|补偿|费)

[0046] 本实施示例中，依次利用上述各个第一正则表达式对法律文书进行检测，在检测出与“(?:伤残|残疾).\{0,5\}?(?:赔偿|补偿)”匹配的字符串时，将该字符串所属文本位置确定为“残疾赔偿金”这一费用项所属的定位。以该定位为分割点，将法律文书分成前后两个文本区域。在其中的后一文本区域中，查找出最先检测得到的阿拉伯数字。例如，若最先检测到的阿拉伯数字为“200”，则当前得到的提取结果为“交通费--200”。

[0047] 优选地，在提取得到的费用项以及金额值的对应关系中，在上述阿拉伯数字之后加上金额值的默认位置“元”。

[0048] 作为本申请的一个实施示例，图3示出了本申请实施例提供的金额信息的提取方法S102的另一具体实现流程，详述如下：

[0049] S1024：获取包含依序排列的多个费用项的信息列表以及与各类金额值关联的第二正则表达式。

- [0050] 本申请实施例中，若与文本类型匹配的识别算法为第二算法，则根据用户的金额值信息获取需求，获取用户输入的多个费用项的名称，并将其中每一个费用项的名称录入预先创建的信息列表中。
- [0051] 不同的费用项具有不同的优先级。在信息列表中，根据用户对各个费用项所设置的优先级，将优先级较高的费用项排于优先级较低的费用项之前。若未检测到用户所输入的优先级设置指令，则读取各个费用项名称的输入时间，并将输入时间较早的费用项确定为优先级较高的费用项。
- [0052] 本申请实施例中，加载预先建立的第二正则文件，第二正则文件包括与各类金额值关联的第二正则表达式：例如，该第二正则表达式为'(?:[0-9]*[,]*[0-9])+\.?[0-9]*(?:元|万元|百万元|千万元)?'。
- [0053] S1025：在所述法律文书中，定位与所述第二正则表达式匹配的各个所述金额值。
- [0054] S1026：对于每相邻的两个所述金额值之间所存在的文本区域，依序读取所述信息列表中的所述费用项，并判断所述文本区域中是否存在该费用项。
- [0055] 在法律文书中，利用上述第二正则表达式进行匹配查找，以检测出法律文本中各个金额值所属的文本位置。例如，检测的各个金额值依次为“4437元 400元 200元”。此后，以这些金额值为分割点，把法律文书分割成若干个文本区域。
- [0056] 本申请实施例中，对于每两个金额值之间所存在的文本区域，依序读取信息列表中的一个费用项，并检测该费用项是否出现于该文本区域中。
- [0057] 例如，上述示例中，由于“400元”和“200元”为依次检测得到的两个金额值，因此，在法律文书中，读取存在于“400元”与“200元”之间的各个字符。根据信息列表中的各个费用项的优先级高低次序，依次读取其中每一费用项所对应的第二正则表达式。若当前时刻所读取的一个费用项为“医疗费”，则利用“医疗费”所对应的第二正则表达式'(?:医疗|医药).[0,5]?(?:赔偿|补偿|费)'，检测上述“400元”与“200元”之间的文本区域中，是否存在与该第二正则表达式匹配的的字符串。若判断结果为否，则读取信息列表中排在“医疗费”之后的下一费用项所对应的第二正则表达式，并返回执行检测上述“400元”与“200元”之间的文本区域中，是否存在与该第二正则表达式匹配的的字符串。

- [0058] S1027: 若所述文本区域中存在该费用项, 则在所述相邻的两个所述金额值中, 选取排序在后的金额值, 并在将该费用项以及该金额值的对应关系进行提取后, 停止读取所述信息列表中的所述费用项。
- [0059] 对于当前时刻在信息列表中所读取的一个费用项, 若检测到文本区域中存在该费用项, 则停止检测操作, 并确定该费用项所对应的金额值为该文本区域后所出现的首个金额值。
- [0060] 例如, 上述“400元”与“200元”之间的文本区域中, 若存在“赔偿交通费以及医疗费共”这一字符串, 则根据上述步骤S1021至S1023来进行检测时, 将输出“交通费-200”以及“医疗费-200”这两项金额数据信息。然而, 本申请实施例中, 根据上述步骤S1024至S1027来进行检测时, 由于“200”为费用项“医疗费”之后所首个出现的金额值, 故输出的金额数据信息为“医疗费-200”, 而不再会输出“交通费-200”, 由此保证最终在计算费用项总额时, 得到的总额结果依然是正确的, 从而降低了金额数据信息对后续应用分析过程所产生的影响, 因而在一定程度上也提高了金额信息的提取准确率。
- [0061] S103: 加载报告模板, 并根据所述报告模板中预先设置好的表头, 将所述费用项及所述金额值导入对应的表体中。
- [0062] 本申请实施例中, 加载预先生成的报告模板, 所述报告模板包含各项表头, 每一表头与一表体对应, 且每一表头用于描述金额信息中的一基础信息字段。上述基础信息字段包括费用项以及金额值。每一表体用于记录一项与基础信息字段对应的匹配参数。对于报告模板中预先设置好的每一表头, 根据S102所获得的各个费用项以及金额值的对应关系中, 筛选出与基础信息字段对应的匹配参数, 并将该匹配参数导入报告模板的表体中。
- [0063] S104: 根据所述报告模板的导入结果计算表尾, 输出关于所述法律文书的金额信息提取报表。
- [0064] 根据报告模板所导入的每一费用项所对应的金额值, 计算各个金额值的总和, 以将得到的结果输出为诉请项参考总数后, 导入至报告模板的表尾, 从而输出关于上述法律文书的金额信息提取报表, 并对该金额信息提取报表进行展示。
- [0065] 本申请实施例中, 由于不同案由类别以及不同地区所发行的法律文书的样式差

异较大，故根据法律文书的案由类别以及发行区域，确定出法律文书的文本类型后，再从预设的多种识别算法中获取与该文本类型匹配的识别算法，保证了能够利用最佳且最具有针对性的一种识别算法来准确地提取法律文本中的金额数据信息，提高了金额信息的提取准确率，同时也为案件预判、律师画像、律师胜诉率分析、判决合理性分析等应用提供了更具参考价值的基础数据；另外，由于本申请实施例能够通过预设的识别算法以及报告模板来自动导出法律文书中各费用项以及金额值的对应关系，因而对于用户而言，其能够直观地查看到自己所需的金额信息，无需再对法律文书进行人工识别，因此，提高了金额信息的提取效率。

[0066] 作为本申请的另一实施例，图4示出了本申请另一实施例提供的金额信息的提取方法S102的具体实现流程。如图4所示，在上述S1026之后，还包括：

[0067] S1028：若所述文本区域中不存在所述信息列表中的任一所述费用项，则在所述相邻的两个所述金额值中，选取排序在后的金额值，并读取位于该金额值之前的N个字符；其中，所述N为大于零的整数，且所述N为预设值。

[0068] S1029：通过预设的自然语言处理算法，分别对各个所述字符进行实体标注处理。

[0069] S1030：获取用户分别对各个所述字符设置的费用项名称。

[0070] S1031：基于各个所述字符所分别对应的所述实体标注以及所述费用项名称，构建并训练神经网络模型。

[0071] S1032：在每相邻的两个所述金额值之间的文本区域中，若再次检测到所述文本区域中未存在所述信息列表中的任一所述费用项，则通过所述神经网络模型，输出与排序在后的所述金额值相对应的所述费用项。

[0072] 由于信息列表中所存储的各个费用项基于用户的输入指令的预先获取。因此，若用户所输入的费用项并未涵盖法律文书中所可能出现的各类费用项，则基于该信息列表来对上述文本区域进行检测时，会得出信息列表中的每一费用项均不存在于该文本区域的情况。例如，法律文书中可能出现“衣服损失费200元”，但信息列表中并未预设“衣物损失费”这一费用项。因此，本申请实施例中，为了自动识别出“200元”所对应的费用项名称。通过标注在该金额值“200元”之前所

出现的多个字符的实体类型，并请求用户手工输入这些字符实际所对应的费用项，保证了在基于这些字符及其实体类型来作为训练样本来训练神经网络模型后，能够得到用于自动化识别费用项名称的神经网络模型。

[0073] 当在下次再次遇到文本区域中仅存在金额值而未检测到信息列表中所存在的费用项的情况下，通过读取该金额值之前的N个字符，并识别字符所对应的实体类型，将该实体类型以及字符作为当前时刻神经网络模型的输入参数，可在神经网络模型运行处理后，自动识别该金额值所对应的费用项，因此，本申请实施例提高了对金额信息的识别准确度，也提高了整个金额信息提取系统的智能化程度，避免了每次在费用项匹配失败时，只能由用户执行人工识读。

[0074] 作为本申请的又一实施例，图5示出了本申请又一实施例提供的金额信息的提取方法的实现流程。如图5所示，在上述S104之后，还包括：

[0075] S105：获取与信息变更事件关联的关键词。

[0076] 实际场景中，法律文书可能包含多个名称相同的费用项，但这些费用项所分布的文本位置不同。例如，法律文书在第一段文本区域中，先记载原告在6月1日所诉请的各个费用项金额，而在第二段文本区域中，再记载原告在6月10日所重新诉请的费用项金额。因此，在通过上述S101至S104来对该法律文书进行检测处理后，在输出得到的关于该法律文书的金额信息提取报表中，可能会存在同一费用项对应不同金额值的情况出现。

[0077] 本申请实施例中，获取预设的与信息变更事件关联的关键词。上述关键词包括但不限于“变更为”、“修改为”以及“更改为”等。用户对预先收集得到的各类法律文书进行人工识别后，可确定出费用项变更时存在于法律文书中的常见词组，故通过人工归纳总结后，可得到上述各个关键词。因此，基于用户输入的关键词设置指令，可获取与信息变更事件关联的关键词。

[0078] S106：对所述法律文书进行检测处理，判断所述法律文书中是否存在所述关键词。

[0079] S107：若所述法律文书中存在所述关键词，则提取出与所述关键词邻近的所述费用项以及所述金额值，并在所述金额信息提取报表中，将该费用项所对应的金额值变更为当前时刻提取得到的所述金额值。

- [0080] 本申请实施例中，判断法律文书中是否包含上述任一关键词。若检测到任一关键词存在于该法律文书，则基于上述第一正则表达式以及第二正则表达式，分别识别出与该关键词所属文本位置最接近的费用项名称以及金额值。在上述S104所输出的金额信息提取报表中，检测出与该费用项名称匹配的表体数据，将该表体数据中所存在的金额值更新为当前时刻所识别得到的金额值。
- [0081] 示例性地，若预设的关键词包括“变更为”，且在法律文书中存在“原告将交通费诉请变更为300元”这一字符串，则可确定出该法律文书中存在关键词。通过与费用项描述信息相关联的第一正则表达式，在“变更为”这一关键词所属的文本位置之前以及在“变更为”这一关键词所属的文本位置之后，检测出首个出现的费用项名称为“交通费”；通过与各类金额值关联的第二正则表达式，检测出首个出现的金额值为“300”。因此，在金额信息提取报表中，将所存在的表体数据“交通费-200”更新为“交通费-300”。
- [0082] 本申请实施例中，在法律文书中存在与信息变更事件关联的关键词时，通过提取出与该关键词邻近的费用项以及所述金额值，并在金额信息提取报表中，将该费用项所对应的金额值变更为当前时刻提取得到的金额值，保证了最终用户所查看得到的金额信息为实际生效的金额信息，故提高了金额信息的参考价值，提高了金额信息的获取准确率。
- [0083] 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，而不应对本申请实施例的实施过程构成任何限定。
- [0084] 对应于本申请实施例所提供的金额信息的提取方法，图6示出了本申请实施例提供的金额信息的提取装置的结构框图。为了便于说明，仅示出了与本实施例相关的部分。
- [0085] 参照图6，该装置包括：
- [0086] 第一获取单元61，用于根据法律文书的案由类别以及发行区域，获取所述法律文书的文本类型。
- [0087] 确定单元62，用于在预设的多个识别算法中，确定出与所述文本类型匹配的所述识别算法，并基于确定出的所述识别算法，提取所述法律文书中的各个费用

项以及与每一个所述费用项相对应的金额值。

[0088] 加载单元63，用于加载报告模板，并根据所述报告模板中预先设置好的表头，将所述费用项及所述金额值导入对应的表体中。

[0089] 第一输出单元64，用于根据所述报告模板的导入结果计算表尾，输出关于所述法律文书的金额信息提取报表。

[0090] 可选地，若确定出的所述识别算法为第一算法，则所述确定单元62包括：

[0091] 第一获取子单元，用于获取与费用项描述信息相关联的第一正则表达式。

[0092] 第一定位子单元，用于在所述法律文书中，定位与所述第一正则表达式匹配的各个所述费用项。

[0093] 查找子单元，用于对定位出的每一个所述费用项，查找所述法律文书中出现在该费用项之后的首个金额值，并将该费用项以及该金额值的对应关系进行提取。

[0094] 可选地，若确定出的所述识别算法为第二算法，则所述确定单元62包括：

[0095] 第二获取子单元，用于获取包含依序排列的多个费用项的信息列表以及与各类金额值关联的第二正则表达式。

[0096] 第二定位子单元，用于在所述法律文书中，定位与所述第二正则表达式匹配的各个所述金额值。

[0097] 判断子单元，用于对于每相邻的两个所述金额值之间所存在的文本区域，依序读取所述信息列表中的所述费用项，并判断所述文本区域中是否存在该费用项。

[0098] 选取子单元，用于若所述文本区域中存在该费用项，则在所述相邻的两个所述金额值中，选取排序在后的金额值，并在将该费用项以及该金额值的对应关系进行提取后，停止读取所述信息列表中的所述费用项。

[0099] 可选地，所述金额信息的提取装置还包括：

[0100] 选取单元，用于若所述文本区域中不存在所述信息列表中的任一所述费用项，则在所述相邻的两个所述金额值中，选取排序在后的金额值，并读取位于该金额值之前的N个字符；其中，所述N为大于零的整数，且所述N为预设值。

[0101] 标注单元，用于通过预设的自然语言处理算法，分别对各个所述字符进行实体

标注处理。第二获取单元，用于获取用户分别对各个所述字符设置的费用项名称。

[0102] 构建单元，用于基于各个所述字符所分别对应的所述实体标注以及所述费用项名称，构建并训练神经网络模型。

[0103] 第二输出单元，用于在每相邻的两个所述金额值之间的文本区域中，若再次检测到所述文本区域中未存在所述信息列表中的任一所述费用项，则通过所述神经网络模型，输出与排序在后的所述金额值相对应的所述费用项。

[0104] 可选地，所述金额信息的提取装置还包括：

[0105] 第三获取单元，用于获取与信息变更事件关联的关键词。判断单元，用于对所述法律文书进行检测处理，判断所述法律文书中是否存在所述关键词。变更单元，用于若所述法律文书中存在所述关键词，则提取出与所述关键词邻近的所述费用项以及所述金额值，并在所述金额信息提取报表中，将该费用项所对应的金额值变更为当前时刻提取得到的所述金额值。

[0106] 图7是本申请一实施例提供的终端设备的示意图。如图7所示，该实施例的终端设备7包括：处理器70以及存储器71，所述存储器71中存储有可在所述处理器70上运行的计算机可读指令72，例如金额信息的提取程序。所述处理器70执行所述计算机可读指令72时实现上述各个金额信息的提取方法实施例中的步骤，例如图1所示的步骤101至104。或者，所述处理器70执行所述计算机可读指令72时实现上述各装置实施例中各模块/单元的功能，例如图6所示单元61至64的功能。

[0107] 示例性的，所述计算机可读指令72可以被分割成一个或多个模块/单元，所述一个或者多个模块/单元被存储在所述存储器71中，并由所述处理器70执行，以完成本申请。所述一个或多个模块/单元可以是能够完成特定功能的一系列计算机可读指令段，该指令段用于描述所述计算机可读指令72在所述终端设备7中的执行过程。

[0108] 所述终端设备7可以是桌上型计算机、笔记本、掌上电脑及云端服务器等计算设备。所述终端设备可包括，但不仅限于，处理器70、存储器71。本领域技术人员可以理解，图7仅仅是终端设备7的示例，并不构成对终端设备7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，例如

所述终端设备还可以包括输入输出设备、网络接入设备、总线等。

[0109] 所称处理器70可以是中央处理单元(Central Processing Unit, CPU), 还可以是其他通用处理器、数字信号处理器(Digital Signal Processor, DSP)、专用集成电路(Application Specific Integrated Circuit, ASIC)、现成可编程门阵列(Field-Programmable Gate Array, FPGA)或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等。

[0110] 所述存储器71可以是所述终端设备7的内部存储单元, 例如终端设备7的硬盘或内存。所述存储器71也可以是所述终端设备7的外部存储设备, 例如所述终端设备7上配备的插接式硬盘, 智能存储卡(Smart Media Card, SMC), 安全数字(Secure Digital, SD)卡, 闪存卡(Flash Card)等。进一步地, 所述存储器71还可以既包括所述终端设备7的内部存储单元也包括外部存储设备。所述存储器71用于存储所述计算机可读指令以及所述终端设备所需的其他程序和数据。所述存储器71还可以用于暂时地存储已经输出或者将要输出的数据。

[0111] 另外, 在本申请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。

[0112] 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, 本申请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)执行本申请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(Random Access Memory, RAM)、磁碟或者光盘等各种可以存储程序代码的介质。

[0113] 以上所述, 以上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。

权利要求书

- [权利要求 1] 一种金额信息的提取方法，其特征在于，包括：
根据法律文书的案由类别以及发行区域，获取所述法律文书的文本类型；
在预设的多个识别算法中，确定出与所述文本类型匹配的所述识别算法，并基于确定出的所述识别算法，提取所述法律文书中的各个费用项以及与每一个所述费用项相对应的金额值；
加载报告模板，并根据所述报告模板中预先设置好的表头，将所述费用项及所述金额值导入对应的表体中；
根据所述报告模板的导入结果计算表尾，输出关于所述法律文书的金额信息提取报表。
- [权利要求 2] 如权利要求1所述的金额信息的提取方法，其特征在于，若确定出的所述识别算法为第一算法，则所述提取所述法律文书中的各个费用项以及与每一个所述费用项对应的金额值，包括：
获取与费用项描述信息相关联的第一正则表达式；
在所述法律文书中，定位与所述第一正则表达式匹配的各个所述费用项；
对定位出的每一个所述费用项，查找所述法律文书中出现在该费用项之后的首个金额值，并将该费用项以及该金额值的对应关系进行提取。
- [权利要求 3] 如权利要求1所述的金额信息的提取方法，其特征在于，若确定出的所述识别算法为第二算法，则所述提取所述法律文书中的各个费用项以及与每一个所述费用项对应的金额值，包括：
获取包含依序排列的多个费用项的信息列表以及与各类金额值关联的第二正则表达式；
在所述法律文书中，定位与所述第二正则表达式匹配的各个所述金额值；
对于每相邻的两个所述金额值之间所存在的文本区域，依序读取所述

信息列表中的所述费用项，并判断所述文本区域中是否存在该费用项；

若所述文本区域中存在该费用项，则在所述相邻的两个所述金额值中，选取排序在后的金额值，并在将该费用项以及该金额值的对应关系进行提取后，停止读取所述信息列表中的所述费用项。

[权利要求 4]

如权利要求3所述的金额信息的提取方法，其特征在于，还包括：

若所述文本区域中不存在所述信息列表中的任一所述费用项，则在所述相邻的两个所述金额值中，选取排序在后的金额值，并读取位于该金额值之前的N个字符；其中，所述N为大于零的整数，且所述N为预设值；

通过预设的自然语言处理算法，分别对各个所述字符进行实体标注处理；

获取用户分别对各个所述字符设置的费用项名称；

基于各个所述字符所分别对应的所述实体标注以及所述费用项名称，构建并训练神经网络模型；

在每相邻的两个所述金额值之间的文本区域中，若再次检测到所述文本区域中未存在所述信息列表中的任一所述费用项，则通过所述神经网络模型，输出与排序在后的所述金额值相对应的所述费用项。

[权利要求 5]

如权利要求1所述的金额信息的提取方法，其特征在于，还包括：

获取与信息变更事件关联的关键词；

对所述法律文书进行检测处理，判断所述法律文书中是否存在所述关键词；

若所述法律文书中存在所述关键词，则提取出与所述关键词邻近的所述费用项以及所述金额值，并在所述金额信息提取报表中，将该费用项所对应的金额值变更为当前时刻提取得到的所述金额值。

[权利要求 6]

一种金额信息的提取装置，其特征在于，包括：

第一获取单元，用于根据法律文书的案由类别以及发行区域，获取所述法律文书的文本类型；

确定单元，用于在预设的多个识别算法中，确定出与所述文本类型匹配的所述识别算法，并基于确定出的所述识别算法，提取所述法律文书中的各个费用项以及与每一个所述费用项相对应的金额值；

加载单元，用于加载报告模板，并根据所述报告模板中预先设置好的表头，将所述费用项及所述金额值导入对应的表体中；

第一输出单元，用于根据所述报告模板的导入结果计算表尾，输出关于所述法律文书的金额信息提取报表。

[权利要求 7]

根据权利要求6所述的金额信息的提取装置，其特征在于，若确定出的所述识别算法为第一算法，则所述确定单元包括：

第一获取子单元，用于获取与费用项描述信息相关联的第一正则表达式；

第一定位子单元，用于在所述法律文书中，定位与所述第一正则表达式匹配的各个所述费用项；

查找子单元，用于对定位出的每一个所述费用项，查找所述法律文书中出现在该费用项之后的首个金额值，并将该费用项以及该金额值的对应关系进行提取。

[权利要求 8]

根据权利要求6所述的金额信息的提取装置，其特征在于，若确定出的所述识别算法为第二算法，则所述确定单元包括：

第二获取子单元，用于获取包含依序排列的多个费用项的信息列表以及与各类金额值关联的第二正则表达式；

第二定位子单元，用于在所述法律文书中，定位与所述第二正则表达式匹配的各个所述金额值；

判断子单元，用于对于每相邻的两个所述金额值之间所存在的文本区域，依序读取所述信息列表中的所述费用项，并判断所述文本区域中是否存在该费用项；

选取子单元，用于若所述文本区域中存在该费用项，则在所述相邻的两个所述金额值中，选取排序在后的金额值，并在将该费用项以及该金额值的对应关系进行提取后，停止读取所述信息列表中的所述费用

项。

[权利要求 9] 根据权利要求8所述的金额信息的提取装置，其特征在于，还包括：
选取单元，用于若所述文本区域中不存在所述信息列表中的任一所述费用项，则在所述相邻的两个所述金额值中，选取排序在后的金额值，并读取位于该金额值之前的N个字符；其中，所述N为大于零的整数，且所述N为预设值；
标注单元，用于通过预设的自然语言处理算法，分别对各个所述字符进行实体标注处理；
第二获取单元，用于获取用户分别对各个所述字符设置的费用项名称；
构建单元，用于基于各个所述字符所分别对应的所述实体标注以及所述费用项名称，构建并训练神经网络模型；
第二输出单元，用于在每相邻的两个所述金额值之间的文本区域中，若再次检测到所述文本区域中未存在所述信息列表中的任一所述费用项，则通过所述神经网络模型，输出与排序在后的所述金额值相对应的所述费用项。

[权利要求 10] 根据权利要求6所述的金额信息的提取装置，其特征在于，还包括：
第三获取单元，用于获取与信息变更事件关联的关键词；
判断单元，用于对所述法律文书进行检测处理，判断所述法律文书中是否存在所述关键词；
变更单元，用于若所述法律文书中存在所述关键词，则提取出与所述关键词邻近的所述费用项以及所述金额值，并在所述金额信息提取报表中，将该费用项所对应的金额值变更为当前时刻提取得到的所述金额值。

[权利要求 11] 一种终端设备，其特征在于，包括存储器以及处理器，所述存储器中存储有可在所述处理器上运行的计算机可读指令，所述处理器执行所述计算机可读指令时实现如下步骤：
根据法律文书的案由类别以及发行区域，获取所述法律文书的文本类

型；

在预设的多个识别算法中，确定出与所述文本类型匹配的所述识别算法，并基于确定出的所述识别算法，提取所述法律文书中的各个费用项以及与每一个所述费用项相对应的金额值；

加载报告模板，并根据所述报告模板中预先设置好的表头，将所述费用项及所述金额值导入对应的表体中；

根据所述报告模板的导入结果计算表尾，输出关于所述法律文书的金额信息提取报表。

[权利要求 12] 根据权利要求11所述的终端设备，其特征在于，若确定出的所述识别算法为第一算法，则所述提取所述法律文书中的各个费用项以及与每一个所述费用项对应的金额值，包括：

获取与费用项描述信息相关联的第一正则表达式；

在所述法律文书中，定位与所述第一正则表达式匹配的各个所述费用项；

对定位出的每一个所述费用项，查找所述法律文书中出现在该费用项之后的首个金额值，并将该费用项以及该金额值的对应关系进行提取。

。

[权利要求 13] 根据权利要求11所述的终端设备，其特征在于，若确定出的所述识别算法为第二算法，则所述提取所述法律文书中的各个费用项以及与每一个所述费用项对应的金额值，包括：

获取包含依序排列的多个费用项的信息列表以及与各类金额值关联的第二正则表达式；

在所述法律文书中，定位与所述第二正则表达式匹配的各个所述金额值；

对于每相邻的两个所述金额值之间所存在的文本区域，依序读取所述信息列表中的所述费用项，并判断所述文本区域中是否存在该费用项；

；

若所述文本区域中存在该费用项，则在所述相邻的两个所述金额值中

，选取排序在后的金额值，并在将该费用项以及该金额值的对应关系进行提取后，停止读取所述信息列表中的所述费用项。

- [权利要求 14] 根据权利要求13所述的终端设备，其特征在于，所述处理器执行所述计算机可读指令时还实现如下步骤：
- 若所述文本区域中不存在所述信息列表中的任一所述费用项，则在所述相邻的两个所述金额值中，选取排序在后的金额值，并读取位于该金额值之前的N个字符；其中，所述N为大于零的整数，且所述N为预设值；
- 通过预设的自然语言处理算法，分别对各个所述字符进行实体标注处理；
- 获取用户分别对各个所述字符设置的费用项名称；
- 基于各个所述字符所分别对应的所述实体标注以及所述费用项名称，构建并训练神经网络模型；
- 在每相邻的两个所述金额值之间的文本区域中，若再次检测到所述文本区域中未存在所述信息列表中的任一所述费用项，则通过所述神经网络模型，输出与排序在后的所述金额值相对应的所述费用项。

- [权利要求 15] 根据权利要求11所述的终端设备，其特征在于，所述处理器执行所述计算机可读指令时还实现如下步骤：
- 获取与信息变更事件关联的关键词；
- 对所述法律文书进行检测处理，判断所述法律文书中是否存在所述关键词；
- 若所述法律文书中存在所述关键词，则提取出与所述关键词邻近的所述费用项以及所述金额值，并在所述金额信息提取报表中，将该费用项所对应的金额值变更为当前时刻提取得到的所述金额值。

- [权利要求 16] 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可读指令，其特征在于，所述计算机可读指令被至少一个处理器执行时实现如下步骤：
- 根据法律文书的案由类别以及发行区域，获取所述法律文书的文本类

型；

在预设的多个识别算法中，确定出与所述文本类型匹配的所述识别算法，并基于确定出的所述识别算法，提取所述法律文书中的各个费用项以及与每一个所述费用项相对应的金额值；

加载报告模板，并根据所述报告模板中预先设置好的表头，将所述费用项及所述金额值导入对应的表体中；

根据所述报告模板的导入结果计算表尾，输出关于所述法律文书的金额信息提取报表。

[权利要求 17] 根据权利要求16所述的计算机可读存储介质，其特征在于，若确定出的所述识别算法为第一算法，则所述提取所述法律文书中的各个费用项以及与每一个所述费用项对应的金额值，包括：
获取与费用项描述信息相关联的第一正则表达式；
在所述法律文书中，定位与所述第一正则表达式匹配的各个所述费用项；
对定位出的每一个所述费用项，查找所述法律文书中出现在该费用项之后的首个金额值，并将该费用项以及该金额值的对应关系进行提取。

[权利要求 18] 根据权利要求16所述的计算机可读存储介质，其特征在于，若确定出的所述识别算法为第二算法，则所述提取所述法律文书中的各个费用项以及与每一个所述费用项对应的金额值，包括：
获取包含依序排列的多个费用项的信息列表以及与各类金额值关联的第二正则表达式；
在所述法律文书中，定位与所述第二正则表达式匹配的各个所述金额值；
对于每相邻的两个所述金额值之间所存在的文本区域，依序读取所述信息列表中的所述费用项，并判断所述文本区域中是否存在该费用项；
若所述文本区域中存在该费用项，则在所述相邻的两个所述金额值中

，选取排序在后的金额值，并在将该费用项以及该金额值的对应关系进行提取后，停止读取所述信息列表中的所述费用项。

- [权利要求 19] 根据权利要求18所述的计算机可读存储介质，其特征在于，所述计算机可读指令被至少一个处理器执行时还实现如下步骤：
- 若所述文本区域中不存在所述信息列表中的任一所述费用项，则在所述相邻的两个所述金额值中，选取排序在后的金额值，并读取位于该金额值之前的N个字符；其中，所述N为大于零的整数，且所述N为预设值；
- 通过预设的自然语言处理算法，分别对各个所述字符进行实体标注处理；
- 获取用户分别对各个所述字符设置的费用项名称；
- 基于各个所述字符所分别对应的所述实体标注以及所述费用项名称，构建并训练神经网络模型；
- 在每相邻的两个所述金额值之间的文本区域中，若再次检测到所述文本区域中未存在所述信息列表中的任一所述费用项，则通过所述神经网络模型，输出与排序在后的所述金额值相对应的所述费用项。

- [权利要求 20] 根据权利要求16所述的计算机可读存储介质，其特征在于，所述计算机可读指令被至少一个处理器执行时还实现如下步骤：
- 获取与信息变更事件关联的关键词；
- 对所述法律文书进行检测处理，判断所述法律文书中是否存在所述关键词；
- 若所述法律文书中存在所述关键词，则提取出与所述关键词邻近的所述费用项以及所述金额值，并在所述金额信息提取报表中，将该费用项所对应的金额值变更为当前时刻提取得到的所述金额值。

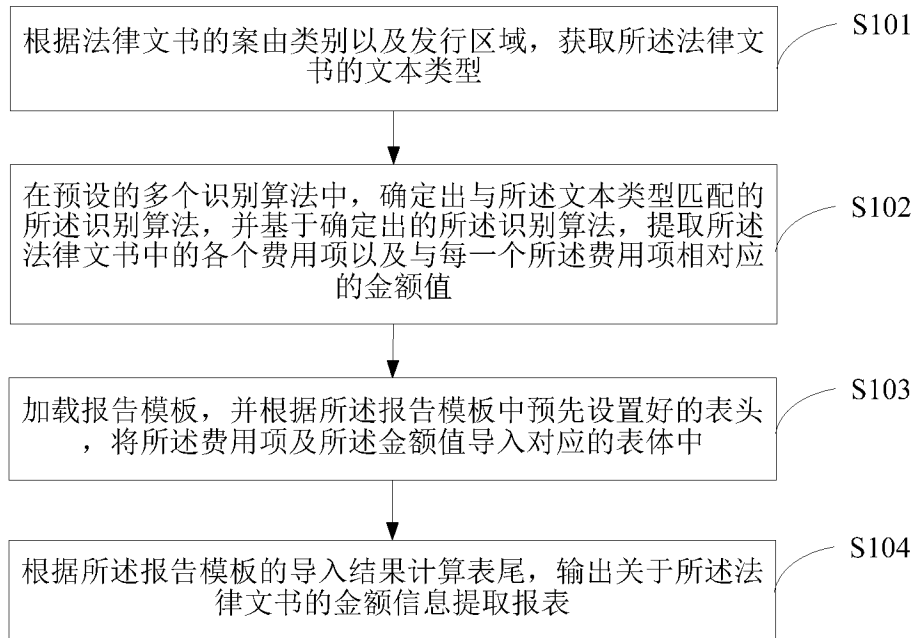


图 1

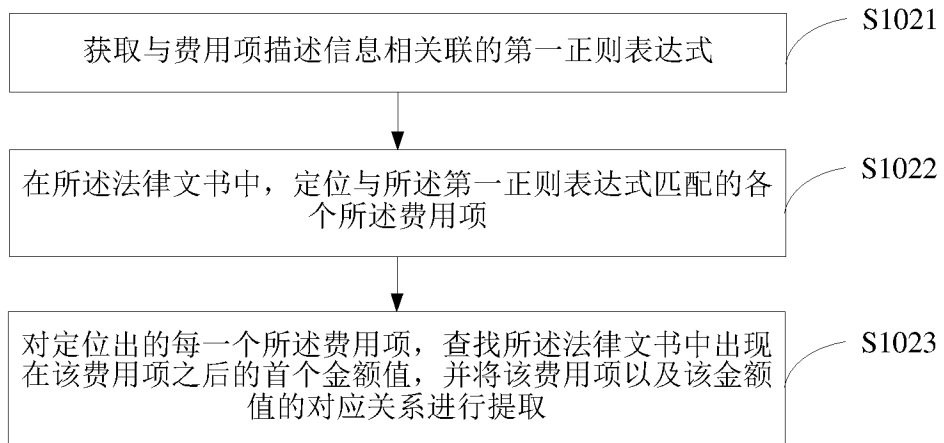


图 2

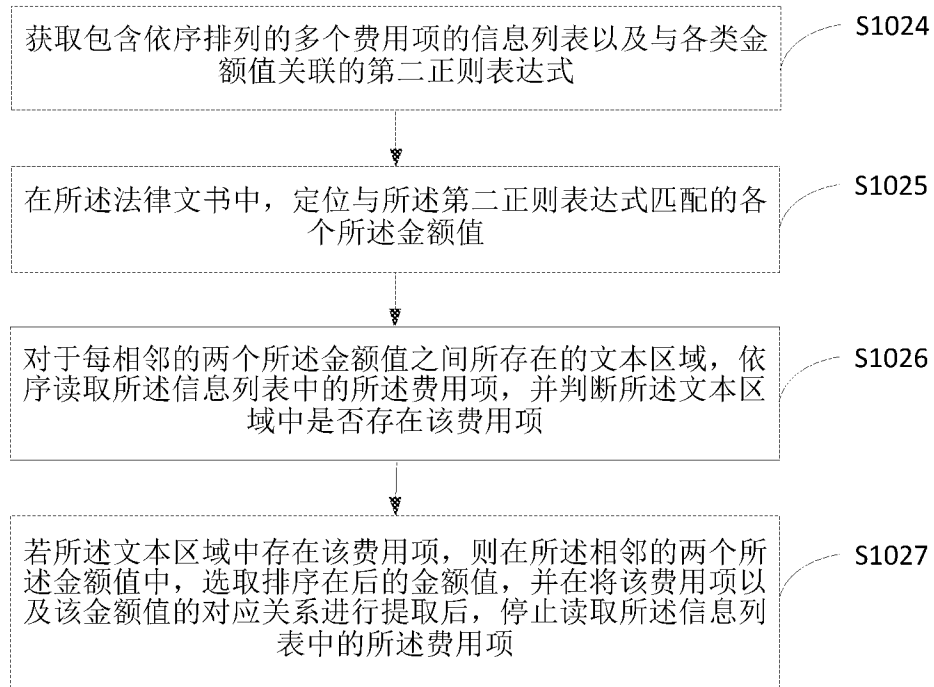


图 3



图 4



图 5

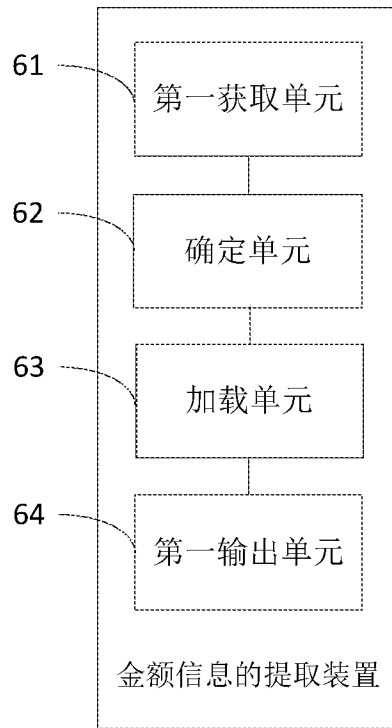


图 6

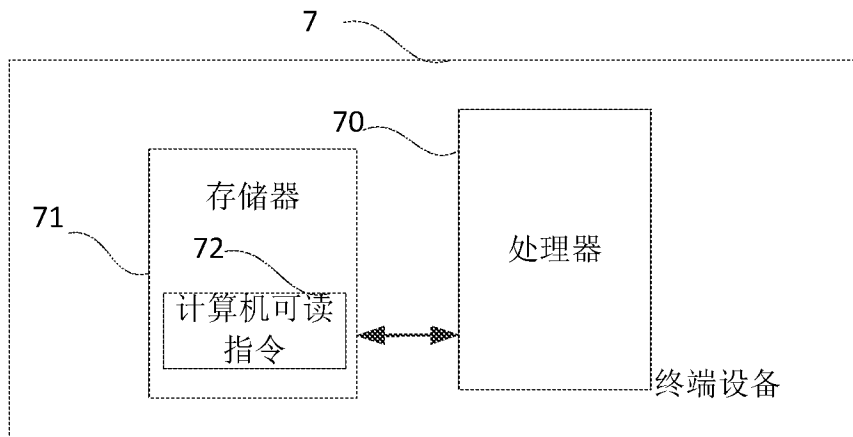


图 7

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2018/105533

A. CLASSIFICATION OF SUBJECT MATTER

G06F 17/24(2006.01)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 G06Q G06K

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)

CNKI, CNPAT, WPI, EPODOC: 法律, 司法, 案件, 涉案, 文书, 金额, 费用, 识别, 提取, 模板, 报表, judicial, case, amount, fee, sum, extract+, report, recogn+, identifi+

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 108984500 A (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) 11 December 2018 (2018-12-11) claims 1-10, and description, paragraphs 0093-0116	1-20
Y	CN 106815203 A (BEIJING GRIDSUM TECHNOLOGY CO., LTD.) 09 June 2017 (2017-06-09) description, paragraphs 0038-0063 and 0107	1-20
Y	CN 108132925 A (ZHONGYING YOUCHUANG INFORMATION TECHNOLOGY CO., LTD.) 08 June 2018 (2018-06-08) description, paragraphs 0040-0072	1-20
A	CN 106502969 A (TIANJIN HYLANDA INFORMATION TECHNOLOGY CO., LTD.) 15 March 2017 (2017-03-15) entire document	1-20
A	CN 107784027 A (BEIJING GRIDSUM TECHNOLOGY CO., LTD.) 09 March 2018 (2018-03-09) 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

12 January 2019

Date of mailing of the international search report

07 March 2019

Name and mailing address of the ISA/CN

State Intellectual Property Office of the P. R. China
No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing
100088
China

Authorized officer

Facsimile No. (86-10)62019451

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2018/105533

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	108984500	A	11 December 2018	None			
CN	106815203	A	09 June 2017	WO	2017092555	A1	08 June 2017
CN	108132925	A	08 June 2018	None			
CN	106502969	A	15 March 2017	None			
CN	107784027	A	09 March 2018	None			

国际检索报告

国际申请号

PCT/CN2018/105533

<p>A. 主题的分类 G06F 17/24(2006.01)i</p> <p>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</p>																				
<p>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 G06Q G06K</p> <p>包含在检索领域中的除最低限度文献以外的检索文献</p> <p>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)) CNKI, CNPAT, WPI, EPODOC: 法律, 司法, 案件, 涉案, 文书, 金额, 费用, 识别, 提取, 模板, 报表, judicial, case, amount, fee, sum, extract+, report, recogn+, identifi+</p>																				
<p>C. 相关文件</p> <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 108984500 A (平安科技深圳有限公司) 2018年 12月 11日 (2018 - 12 - 11) 权利要求1-10、说明书0093-0116段</td> <td>1-20</td> </tr> <tr> <td>Y</td> <td>CN 106815203 A (北京国双科技有限公司) 2017年 6月 9日 (2017 - 06 - 09) 说明书0038-0063、0107段</td> <td>1-20</td> </tr> <tr> <td>Y</td> <td>CN 108132925 A (中盈优创资讯科技有限公司) 2018年 6月 8日 (2018 - 06 - 08) 说明书0040-0072段</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>CN 106502969 A (天津海量信息技术股份有限公司) 2017年 3月 15日 (2017 - 03 - 15) 全文</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>CN 107784027 A (北京国双科技有限公司) 2018年 3月 9日 (2018 - 03 - 09) 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 108984500 A (平安科技深圳有限公司) 2018年 12月 11日 (2018 - 12 - 11) 权利要求1-10、说明书0093-0116段	1-20	Y	CN 106815203 A (北京国双科技有限公司) 2017年 6月 9日 (2017 - 06 - 09) 说明书0038-0063、0107段	1-20	Y	CN 108132925 A (中盈优创资讯科技有限公司) 2018年 6月 8日 (2018 - 06 - 08) 说明书0040-0072段	1-20	A	CN 106502969 A (天津海量信息技术股份有限公司) 2017年 3月 15日 (2017 - 03 - 15) 全文	1-20	A	CN 107784027 A (北京国双科技有限公司) 2018年 3月 9日 (2018 - 03 - 09) 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 108984500 A (平安科技深圳有限公司) 2018年 12月 11日 (2018 - 12 - 11) 权利要求1-10、说明书0093-0116段	1-20																		
Y	CN 106815203 A (北京国双科技有限公司) 2017年 6月 9日 (2017 - 06 - 09) 说明书0038-0063、0107段	1-20																		
Y	CN 108132925 A (中盈优创资讯科技有限公司) 2018年 6月 8日 (2018 - 06 - 08) 说明书0040-0072段	1-20																		
A	CN 106502969 A (天津海量信息技术股份有限公司) 2017年 3月 15日 (2017 - 03 - 15) 全文	1-20																		
A	CN 107784027 A (北京国双科技有限公司) 2018年 3月 9日 (2018 - 03 - 09) 全文	1-20																		
<p><input type="checkbox"/> 其余文件在C栏的续页中列出。</p> <p><input checked="" type="checkbox"/> 见同族专利附件。</p>																				
<p>* 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</p> <p>“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</p>																				
国际检索实际完成的日期	国际检索报告邮寄日期																			
2019年 1月 12日	2019年 3月 7日																			
ISA/CN的名称和邮寄地址	受权官员																			
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	柴华																			
传真号 (86-10)62019451	电话号码 86-(10)-53961630																			

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2018/105533

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	108984500	A	2018年 12月 11日	无			
CN	106815203	A	2017年 6月 9日	W0	2017092555	A1	2017年 6月 8日
CN	108132925	A	2018年 6月 8日	无			
CN	106502969	A	2017年 3月 15日	无			
CN	107784027	A	2018年 3月 9日	无			