



(12)发明专利申请

(10)申请公布号 CN 108681551 A

(43)申请公布日 2018.10.19

(21)申请号 201810281778.2

(22)申请日 2018.03.30

(71)申请人 北京智慧正安科技有限公司
地址 100000 北京市海淀区中关村东路18号1幢7层A-807

(72)发明人 李晓辉

(74)专利代理机构 深圳市舜立知识产权代理事务所(普通合伙) 44335
代理人 侯艺

(51) Int. Cl.
G06F 17/30(2006.01)
G06Q 50/18(2012.01)

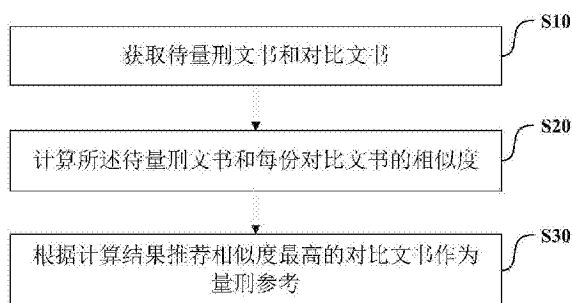
权利要求书2页 说明书7页 附图2页

(54)发明名称

文书对比方法、电子装置及计算机可读存储介质

(57)摘要

本发明公开了一种文书对比方法,该方法包括:获取待量刑文书和多份对比文书;计算所述待量刑文书和每份对比文书的相似度;根据计算结果推荐相似度最高的对比文书作为量刑参考。本发明实施例还公开了一种电子装置和计算机可读存储介质。所述文书对比方法、电子装置及计算机可读存储介质,能够辅助公检法机关做到“类案同判”,从而帮助准确、规范的量刑。



1. 一种文书对比方法,其特征在于,该方法包括步骤:
获取待量刑文书和多份对比文书;
计算所述待量刑文书和每份对比文书的相似度;及
根据计算结果推荐相似度最高的对比文书作为量刑参考。
2. 根据权利要求1所述的文书对比方法,其特征在于,所述计算所述待量刑文书和每份对比文书的相似度的步骤包括:
将每份文书分解为多个部分;
将所述待量刑文书和所述对比文书中的每个部分进行对比,计算出每个部分的相似度;
按照每个部分的预设权重和所述每个部分的相似度计算出所述待量刑文书和所述对比文书的总相似度。
3. 根据权利要求2所述的文书对比方法,其特征在于,所述根据计算结果推荐相似度最高的对比文书作为量刑参考的步骤包括:
根据所述总相似度将所有所述对比文书进行排序;
将与所述待量刑文书的总相似度最高的对比文书推荐给用户,为所述待量刑文书中的被告参考量刑。
4. 根据权利要求1-3任一项所述的文书对比方法,其特征在于,在所述获取待量刑文书和多份对比文书的步骤中,所述对比文书为从案件数据库中按设置的范围获取,或者根据所述案件数据库中记录的案件基本信息进行初步筛选后获取。
5. 根据权利要求2所述的文书对比方法,其特征在于,在所述将每份文书分解为多个部分的步骤中,根据预设的关键词对每份文书进行文本识别,将每份文书分解为案由、事实情节、被告人信息、案件基本信息四个部分。
6. 根据权利要求3所述的文书对比方法,其特征在于,所述根据所述总相似度将所有所述对比文书进行排序的步骤包括:
筛选与所述待量刑文书的总相似度超过预设阈值的对比文书;
将筛选后的对比文书按总相似度从高到低进行排序。
7. 根据权利要求3或6所述的文书对比方法,其特征在于,所述将与所述待量刑文书的总相似度最高的对比文书推荐给用户,为所述待量刑文书中的被告参考量刑的步骤包括:
在显示界面中将所述对比文书的基本信息按总相似度从高到低排序显示,排在第一位的对比文书即为与所述待量刑文书的总相似度最高的对比文书,或者将所述总相似度最高的对比文书醒目显示。
8. 根据权利要求7所述的文书对比方法,其特征在于,所述将与所述待量刑文书的总相似度最高的对比文书推荐给用户,为所述待量刑文书中的被告参考量刑的步骤还包括:
根据所述总相似度最高的对比文书或者所述总相似度排在前N名的对比文书中的量刑结果,对所述待量刑文书给出一个建议的量刑范围。
9. 一种电子装置,其特征在于,所述电子装置包括:存储器、处理器及存储在所述存储器上并可在所述处理器上运行的文书对比程序,所述文书对比程序被所述处理器执行时实现如权利要求1至8中任一项所述的文书对比方法的步骤。
10. 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上存储有文书对

比程序,所述文书对比程序被处理器执行时实现如权利要求1至8中任一项所述的文书对比方法的步骤。

文书对比方法、电子装置及计算机可读存储介质

技术领域

[0001] 本发明涉及文本识别技术领域,尤其涉及一种文书对比方法、电子装置及计算机可读存储介质。

背景技术

[0002] 目前,公检法机关对于一个案件中被告的量刑计算,通常采用根据案件文书和相关法规进行估算的方式。即使在最高法实施量刑规范化指导意见后,检察官、法官对同一案件的处理也可能不相同。因此,现有的量刑计算方式很难做到“类案同判”,导致相似案件量刑可能出现较大差别,即量刑不准确、不规范。

发明内容

[0003] 本发明的主要目的在于提出一种文书对比方法、电子装置及计算机可读存储介质,旨在解决如何准确规范地进行量刑计算的问题。

[0004] 为实现上述目的,本发明提供了一种文书对比方法,该方法包括步骤:

[0005] 获取待量刑文书和多份对比文书;

[0006] 计算所述待量刑文书和每份对比文书的相似度;及

[0007] 根据计算结果推荐相似度最高的对比文书作为量刑参考。

[0008] 可选地,所述计算所述待量刑文书和每份对比文书的相似度的步骤包括:

[0009] 将每份文书分解为多个部分;

[0010] 将所述待量刑文书和所述对比文书中的每个部分进行对比,计算出每个部分的相似度;

[0011] 按照每个部分的预设权重和所述每个部分的相似度计算出所述待量刑文书和所述对比文书的总相似度。

[0012] 可选地,所述根据计算结果推荐相似度最高的对比文书作为量刑参考的步骤包括:

[0013] 根据所述总相似度将所有所述对比文书进行排序;

[0014] 将与所述待量刑文书的总相似度最高的对比文书推荐给用户,为所述待量刑文书中的被告参考量刑。

[0015] 可选地,在所述获取待量刑文书和多份对比文书的步骤中,所述对比文书为从案件数据库中按设置的范围获取,或者根据所述案件数据库中记录的案件基本信息进行初步筛选后获取。

[0016] 可选地,在所述将每份文书分解为多个部分的步骤中,根据预设的关键词对每份文书进行文本识别,将每份文书分解为案由、事实情节、被告人信息、案件基本信息四个部分。

[0017] 可选地,所述根据所述总相似度将所有所述对比文书进行排序的步骤包括:

[0018] 筛选与所述待量刑文书的总相似度超过预设阈值的对比文书;

[0019] 将筛选后的对比文书按总相似度从高到低进行排序。

[0020] 可选地,所述将与所述待量刑文书的总相似度最高的对比文书推荐给用户,为所述待量刑文书中的被告参考量刑的步骤包括:

[0021] 在显示界面中将所述对比文书的基本信息按总相似度从高到低排序显示,排在第一位的对比文书即为与所述待量刑文书的总相似度最高的对比文书,或者将所述总相似度最高的对比文书醒目显示。

[0022] 可选地,所述将与所述待量刑文书的总相似度最高的对比文书推荐给用户,为所述待量刑文书中的被告参考量刑的步骤还包括:

[0023] 根据所述总相似度最高的对比文书或者所述总相似度排在前N名的对比文书中的量刑结果,对所述待量刑文书给出一个建议的量刑范围。

[0024] 此外,为实现上述目的,本发明还提出一种电子装置,所述电子装置包括:存储器、处理器及存储在所述存储器上并可在所述处理器上运行的文书对比程序,所述文书对比程序被所述处理器执行时实现如上述的文书对比方法的步骤。

[0025] 进一步地,为实现上述目的,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上存储有文书对比程序,所述文书对比程序被处理器执行时实现如上述的文书对比方法的步骤。

[0026] 本发明提出的文书对比方法、电子装置及计算机可读存储介质,能够通过文书内容的文本识别和对比,推荐与待量刑文书的相似度最高的对比文书给用户作为量刑参考,辅助公检法机关做到“类案同判”,从而帮助准确、规范的量刑。一方面,避免对同一案件的不同理解,同时可以对比已经判决的其他文书进行参考。另一方面,减轻用户的负担,做到判必有据。该方法可以应用于公安机关立案的自动化测算、检察院的量刑建议自动化测算、法院的量刑计算、公众的刑罚测算等,为未来刑事相关人工智能提供模型基础,为当前刑事参与者提供便捷的量刑方式。

附图说明

[0027] 图1为本发明第一实施例提出的一种电子装置的架构图;

[0028] 图2为本发明第二实施例提出的一种文书对比方法的流程图;

[0029] 图3为本发明第三实施例提出的一种文书对比方法的流程图。

[0030] 本发明目的的实现、功能特点及优点将结合实施例,参照附图做进一步说明。

具体实施方式

[0031] 应当理解,此处所描述的具体实施例仅仅用以解释本发明,并不用于限定本发明。

[0032] 实施例一

[0033] 参阅图1所示,本发明第一实施例提出一种电子装置2。所述电子装置2可以是移动电话、智能电话、笔记本电脑、PAD(平板电脑)等可移动设备,以及诸如台式计算机、服务器等固定终端。所述电子装置2包括存储器20、处理器22和文书对比程序28。

[0034] 其中,所述存储器20至少包括一种类型的可读存储介质,用于存储安装于所述电子装置2的操作系统和各类应用软件,例如文书对比程序28的程序代码等。此外,所述存储器20还可以用于暂时地存储已经输出或者将要输出的各类数据。

[0035] 所述处理器22在一些实施例中可以是中央处理器(Central Processing Unit, CPU)、控制器、微控制器、微处理器、或其他数据处理芯片。该处理器22通常用于控制所述电子装置2的总体操作。本实施例中,所述处理器22用于运行所述存储器20中存储的程序代码或者处理数据,例如运行所述文书对比程序28等。

[0036] 所述文书对比程序28被所述处理器22执行时,实现如下步骤:

[0037] a、获取待量刑文书和对比文书。

[0038] 具体地,当根据案件文书对被告进行量刑计算时,首先获取待量刑的文书及需要与所述待量刑文书进行对比的多件对比文书。所述对比文书可以从案件数据库中按设置的范围获取,例如获取某个时间和/或地区范围内的所有案件文书来对比。另外,还可以先根据所述案件数据库中记录的案件基本信息进行初步筛选,例如根据案件类型筛选,得到一批对比文书来与所述待量刑文书进行对比。

[0039] b、计算所述待量刑文书和每份对比文书的相似度。

[0040] 具体地,将所述待量刑文书和每份对比文书中的内容进行文本识别和对比,计算出所述待量刑文书和每份对比文书之间的相似度。该步骤包括:

[0041] 1、将每份文书分解为多个部分。在本实施例中,将每份文书分解为大致四部分:

[0042] (1) 案由,例如:盗窃罪、抢夺罪、故意伤害罪等;

[0043] (2) 事实情节,例如:自首、坦白、积极赔偿等;

[0044] (3) 被告人信息,例如:性别、家庭情况、社会属性等;

[0045] (4) 案件基本信息,例如:是否公开、法官法院等。

[0046] 其中,可以根据预设的关键字对每份文书进行文本识别,找到每一部分对应的内容。

[0047] 2、将所述待量刑文书和所述对比文书中的每个部分进行对比,计算出每个部分的相似度。

[0048] 具体地,针对所述待量刑文书和每份所述对比文书,将分解后的得到的每个部分的内容进行文本识别和对比,得到相似度的百分比。其中,针对每个部分的内容,可以利用预设关键词和对应的近义词库(认为与所述关键词相同)进行查找对比,然后根据查找结果判断每个部分的内容有多少相同,计算出百分比,即为所述待量刑文书和所述对比文书在该部分的相似度。例如,两份文书的案由同为诈骗罪,相似度100%;事实情节中有部分相似,相似度40%;被告人信息中性别、家庭属性、社会属性等部分相同,相似度95%;案件基本信息中的案件类型、本院案件等部分相同,相似度35%。

[0049] 3、按照每个部分的预设权重计算出所述待量刑文书和所述对比文书的总相似度。

[0050] 具体地,预先针对案件文书分解后的每个部分设置一个对应的权重,根据所述每个部分在对被告量刑时的重要性,每个部分的权重可以相同也可以不同。当得到所述待量刑文书与每份所述对比文书中每个部分的相似度之后,以每个部分的相似度乘以该部分的预设权重,然后相加得到总相似度。

[0051] 例如,上述两份文书的案由部分相似度为100%,事实情节部分相似度为40%,被告人信息部分相似度为95%,案件基本信息部分相似度为35%,且四个部分的权重均为25%,则计算得到两份文书的总相似度为 $100\%*25\%+40\%*25\%+95\%*25\%+35\%*25\%=67.5\%$ 。

[0052] c、根据计算结果推荐相似度最高的对比文书作为量刑参考。包括：

[0053] 1、根据所述总相似度将所有所述对比文书进行排序。

[0054] 具体地，当计算完所述待量刑文书与所有所述对比文书之间的总相似度之后，将所有所述对比文书按所述总相似度从高到低进行排序。或者，也可以先筛选总相似度超过预设阈值（例如80%）的对比文书，然后将筛选后的对比文书按总相似度从高到低进行排序。

[0055] 2、将与所述待量刑文书的总相似度最高的对比文书推荐给用户参考量刑。

[0056] 具体地，当按总相似度从高到低排序之后，排在第一位的对比文书即为与所述待量刑文书的总相似度最高的对比文书。在本实施例中，可以在显示界面中将所述对比文书的基本信息（例如文书名称、时间等）按总相似度从高到低排序显示。还可以另外将所述总相似度最高的对比文书用红色或粗体等方式进行醒目显示，以推荐给用户作为对所述待量刑文书中的被告的量刑参考。或者，在所述显示界面中仅显示所述总相似度最高的对比文书，推荐给用户参考。

[0057] 在一优选实施例中，所述显示界面中除了按总相似度从高到低排序显示所述对比文书的基本信息之外，还可以显示每个所述对比文书中的量刑结果，以方便用户直接查看，而不用再进入到所述对比文书中去翻阅量刑结果。

[0058] 在一优选实施例中，所述显示界面中除了排序显示所述对比文书的基本信息和量刑结果之外，还可以对所述待量刑文书给出一个建议的量刑范围。例如，根据所述总相似度最高的对比文书中的量刑结果，或者根据所述总相似度排在前N名（例如前三名）的对比文书中的量刑结果综合参考，提供建议的量刑范围。举例而言，所述总相似度排在前三名的对比文书中分别量刑3年6个月、3年、4年，则可以对所述待量刑文书的被告建议量刑3年6个月，或者建议量刑3-4年。

[0059] 本领域技术人员可以理解，图1中示出的结构并不构成对所述电子装置2的限定，所述电子装置2还可以包括其他必要部件，或者组合某些部件，或者不同的部件布置。

[0060] 实施例二

[0061] 参阅图2所示，本发明第二实施例提出一种文书对比方法。在本实施例中，根据不同的需求，图2所示的流程图中的步骤的执行顺序可以改变，某些步骤可以省略。该方法包括以下步骤：

[0062] S10，获取待量刑文书和对比文书。

[0063] 具体地，当根据案件文书对被告进行量刑计算时，首先获取待量刑的文书及需要与所述待量刑文书进行对比的多件对比文书。所述对比文书可以从案件数据库中按设置的范围获取，例如获取某个时间和/或地区范围内的所有案件文书来对比。另外，还可以先根据所述案件数据库中记录的案件基本信息进行初步筛选，例如根据案件类型筛选，得到一批对比文书来与所述待量刑文书进行对比。

[0064] S20，计算所述待量刑文书和每份对比文书的相似度。

[0065] 具体地，将所述待量刑文书和每份对比文书中的内容进行文本识别和对比，计算出所述待量刑文书和每份对比文书之间的相似度。

[0066] 关于步骤S20的具体介绍，请参阅图3中相关步骤。例如，步骤S20可以参照图3中步骤S102-S106实现，也可以涉及图3中更多其他的步骤。

[0067] S30,根据计算结果推荐相似度最高的对比文书作为量刑参考。

[0068] 具体地,当计算出所述待量刑文书和每份对比文书之间的相似度后,选出与所述待量刑文书相似度最高的对比文书,将该对比文书推荐给用户作为对所述待量刑文书中的被告的量刑参考。

[0069] 关于步骤S30的具体介绍,请参阅图3中相关步骤。例如,步骤S30可以参照图3中步骤S108-S110实现,也可以涉及图3中更多其他的步骤。

[0070] 实施例三

[0071] 参阅图3所示,本发明第三实施例提出一种文书对比方法。在本实施例中,根据不同的需求,图3所示的流程图中的步骤的执行顺序可以改变,某些步骤可以省略。该方法包括以下步骤:

[0072] S100,获取待量刑文书和对比文书。

[0073] 具体地,当根据案件文书对被告进行量刑计算时,首先获取待量刑的文书及需要与所述待量刑文书进行对比的多件对比文书。所述对比文书可以从案件数据库中按设置的范围获取,例如获取某个时间和/或地区范围内的所有案件文书来对比。另外,还可以先根据所述案件数据库中记录的案件基本信息进行初步筛选,例如根据案件类型筛选,得到一批对比文书来与所述待量刑文书进行对比。

[0074] S102,将每份文书分解为多个部分。

[0075] 具体地,一般案件文书中均会包括案由、事实情节、被告人信息、辩护人信息、判决结果、法律条文等信息。为了对被告进行量刑计算,可以从每份文书中选取与被告量刑相关的部分信息,用来进行文书之间的信息对比和相似度计算。在本实施例中,将每份文书分解为大致四部分:

[0076] (1) 案由,例如:盗窃罪、抢夺罪、故意伤害罪等;

[0077] (2) 事实情节,例如:自首、坦白、积极赔偿等;

[0078] (3) 被告人信息,例如:性别、家庭情况、社会属性等;

[0079] (4) 案件基本信息,例如:是否公开、法官法院等。

[0080] 其中,可以根据预设的关键字对每份文书进行文本识别,找到每一部分对应的内容。

[0081] S104,将所述待量刑文书和所述对比文书中的每个部分进行对比,计算出每个部分的相似度。

[0082] 具体地,针对所述待量刑文书和每份所述对比文书,将分解后的得到的每个部分的内容进行文本识别和对比,得到相似度的百分比。其中,针对每个部分的内容,可以利用预设关键词和对应的近义词库(认为与所述关键词相同)进行查找对比,然后根据查找结果判断每个部分的内容有多少相同,计算出百分比,即为所述待量刑文书和所述对比文书在该部分的相似度。例如,两份文书的案由同为诈骗罪,相似度100%;事实情节中有部分相似,相似度40%;被告人信息中性别、家庭属性、社会属性等部分相同,相似度95%;案件基本信息中的案件类型、本院案件等部分相同,相似度35%。

[0083] S106,按照每个部分的预设权重计算出所述待量刑文书和所述对比文书的总相似度。

[0084] 具体地,预先针对案件文书分解后的每个部分设置一个对应的权重,根据所述每

个部分在对被告量刑时的重要性,每个部分的权重可以相同也可以不同。当得到所述待量刑文书与每份所述对比文书中每个部分的相似度之后,以每个部分的相似度乘以该部分的预设权重,然后相加得到总相似度。

[0085] 例如,上述两份文书的案由部分相似度为100%,事实情节部分相似度为40%,被告人信息部分相似度为95%,案件基本信息部分相似度为35%,且四个部分的权重均为25%,则计算得到两份文书的总相似度为 $100\%*25\%+40\%*25\%+95\%*25\%+35\%*25\%=67.5\%$ 。

[0086] 将所述待量刑文书与所获取的所有所述对比文书按照上述步骤分别计算出总相似度,从而可以在后续统计出与所述待量刑文书最相似的对比文书。

[0087] S108,根据所述总相似度将所有所述对比文书进行排序。

[0088] 具体地,当计算完所述待量刑文书与所有所述对比文书之间的总相似度之后,将所有所述对比文书按所述总相似度从高到低进行排序。或者,也可以先筛选总相似度超过预设阈值(例如80%)的对比文书,然后将筛选后的对比文书按总相似度从高到低进行排序。

[0089] S110,将与所述待量刑文书的总相似度最高的对比文书推荐给用户参考量刑。

[0090] 具体地,当按总相似度从高到低排序之后,排在第一位的对比文书即为与所述待量刑文书的总相似度最高的对比文书。在本实施例中,可以在显示界面中将所述对比文书的基本信息(例如文书名称、时间等)按总相似度从高到低排序显示。还可以另外将所述总相似度最高的对比文书用红色或粗体等方式进行醒目显示,以推荐给用户作为对所述待量刑文书中的被告的量刑参考。或者,在所述显示界面中仅显示所述总相似度最高的对比文书,推荐给用户参考。

[0091] 在一优选实施例中,所述显示界面中除了按总相似度从高到低排序显示所述对比文书的基本信息之外,还可以显示每个所述对比文书中的量刑结果,以方便用户直接查看,而不用再进入到所述对比文书中去翻阅量刑结果。

[0092] 在一优选实施例中,所述显示界面中除了排序显示所述对比文书的基本信息和量刑结果之外,还可以对所述待量刑文书给出一个建议的量刑范围。例如,根据所述总相似度最高的对比文书中的量刑结果,或者根据所述总相似度排在前N名(例如前三名)的对比文书中的量刑结果综合参考,提供建议的量刑范围。举例而言,所述总相似度排在前三名的对比文书中分别量刑3年6个月、3年、4年,则可以对所述待量刑文书的被告建议量刑3年6个月,或者建议量刑3-4年。

[0093] 本实施例所提出的文书对比方法,可以基于法律规定的前提下,定义一套准确定义类案的计算机语言,通过文书内容的文本识别和对比,推荐与待量刑文书的相似度最高的对比文书给用户作为量刑参考,辅助公检法机关做到“类案同判”,从而帮助准确、规范量刑。一方面,避免对同一案件的不同理解,同时可以对比已经判决的其他文书进行参考。另一方面,减轻用户的负担,做到判必有据。该方法可以应用于公安机关立案的自动化测算、检察院的量刑建议自动化测算、法院的量刑计算、公众的刑罚测算等,为未来刑事相关人工智能提供模型基础,为当前刑事参与者提供便捷的量刑方式。

[0094] 实施例四

[0095] 本发明还提供了另一种实施方式,即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文书对比程序,所述文书对比程序可被至少一个处理器执行,以使所述至少一个处理器执行如上述的文书对比方法的步骤。

[0096] 需要说明的是,在本文中,术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含,从而使得包括一系列要素的过程、方法、物品或者装置不仅包括那些要素,而且还包括没有明确列出的其他要素,或者是还包括为这种过程、方法、物品或者装置所固有的要素。在没有更多限制的情况下,由语句“包括一个……”限定的要素,并不排除在包括该要素的过程、方法、物品或者装置中还存在另外的相同要素。

[0097] 上述本发明实施例序号仅仅为了描述,不代表实施例的优劣。

[0098] 通过以上的实施方式的描述,本领域的技术人员可以清楚地了解到上述实施例方法可借助软件加必需的通用硬件平台的方式来实现,当然也可以通过硬件,但很多情况下前者是更佳的实施方式。基于这样的理解,本发明的技术方案本质上或者说对现有技术做出贡献的部分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(如ROM/RAM、磁碟、光盘)中,包括若干指令用以使得一台终端(可以是手机,计算机,服务器,空调器,或者网络设备等)执行本发明各个实施例所述的方法。

[0099] 上面结合附图对本发明的实施例进行了描述,但是本发明并不局限于上述的具体实施方式,上述的具体实施方式仅仅是示意性的,而不是限制性的,本领域的普通技术人员在本发明的启示下,在不脱离本发明宗旨和权利要求所保护的范围情况下,还可做出很多形式,这些均属于本发明的保护之内。

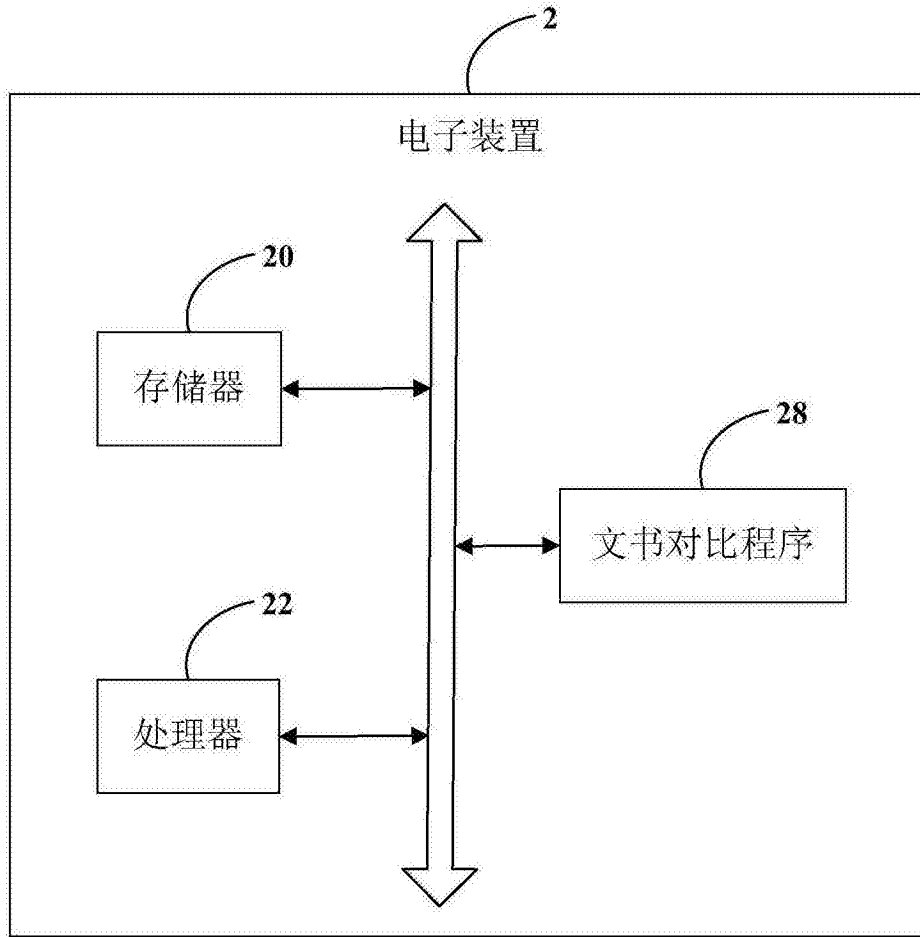


图1

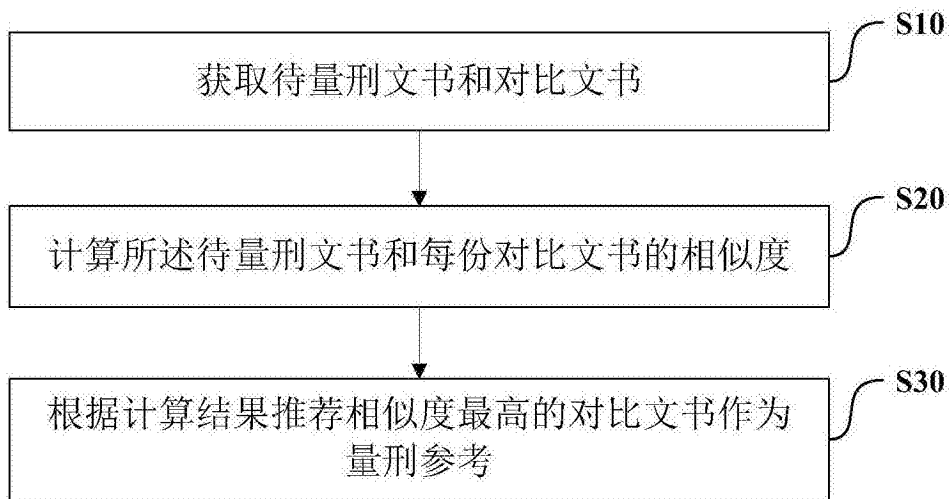


图2

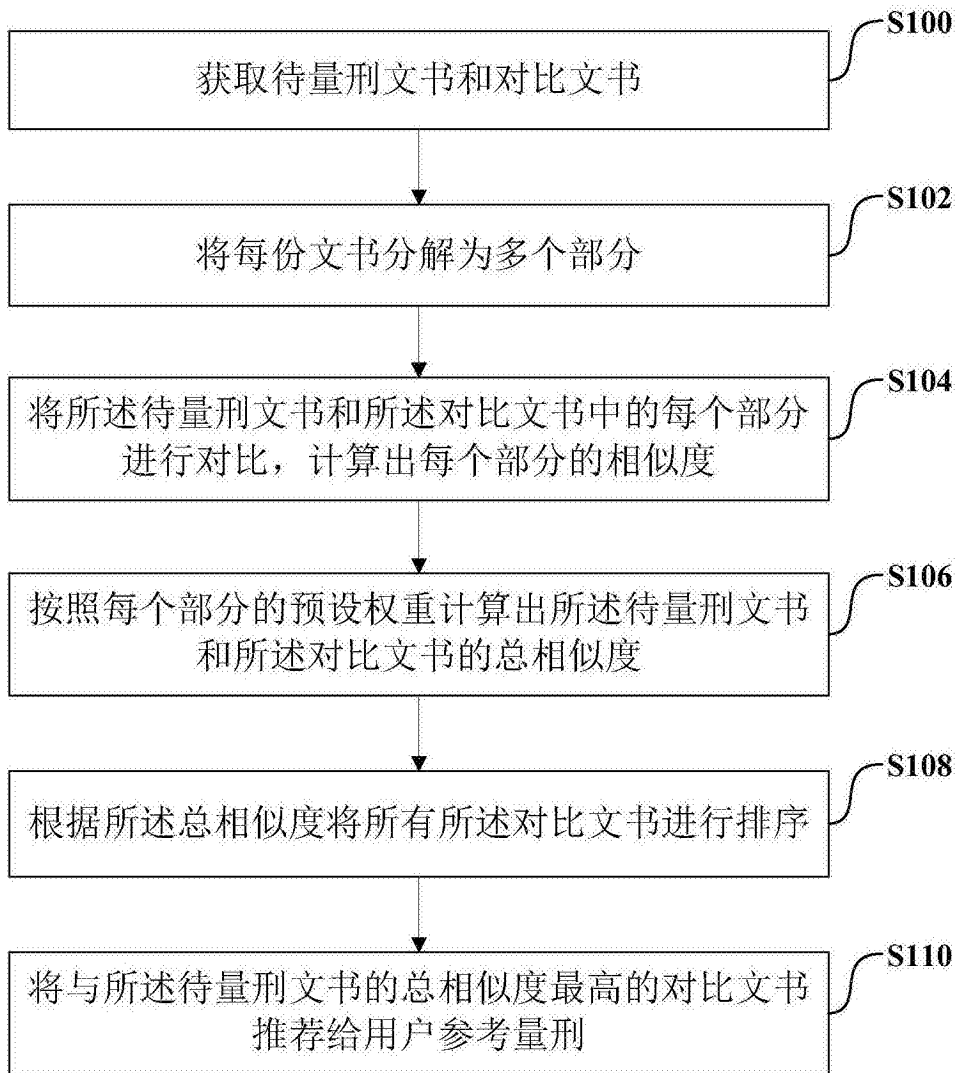


图3