

(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
2018 年 12 月 6 日 (06.12.2018)



(10) 国际公布号
WO 2018/219261 A1

(51) 国际专利分类号:
G06F 3/0481 (2013.01) **G06F 3/0488** (2013.01)

(21) 国际申请号: PCT/CN2018/088789

(22) 国际申请日: 2018 年 5 月 29 日 (29.05.2018)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201710403566.2 2017年6月1日 (01.06.2017) CN

(71) 申请人: 腾讯科技(深圳)有限公司 (TENCENT TECHNOLOGY (SHENZHEN) COMPANY LIMITED) [CN/CN]; 中国广东省深

圳市南山区高新区科技中一路腾讯大厦 35 层, Guangdong 518057 (CN)。

(72) 发明人: 阮闪闪(RUAN, Shanshan); 中国广东省深圳市南山区高新区科技中一路腾讯大厦 35 层, Guangdong 518057 (CN)。 钱成(QIAN, Cheng); 中国广东省深圳市南山区高新区科技中一路腾讯大厦 35 层, Guangdong 518057 (CN)。 罗根(LUO, Gen); 中国广东省深圳市南山区高新区科技中一路腾讯大厦 35 层, Guangdong 518057 (CN)。 蔡元锋(CAI, Yuanfeng); 中国广东省深圳市南山区高新区科技中一路腾讯大厦 35 层, Guangdong 518057 (CN)。 李杨(LI, Yang); 中国广东省深圳市南山区高新区科技中一路腾讯大厦 35 层, Guangdong 518057 (CN)。 王波(WANG, Bo); 中国广东省深

(54) Title: TEXT RECOMBINATION METHOD AND DEVICE, TERMINAL DEVICE, AND COMPUTER-READABLE STORAGE MEDIUM

(54) 发明名称: 文本重组方法、装置、终端设备及计算机可读存储介质

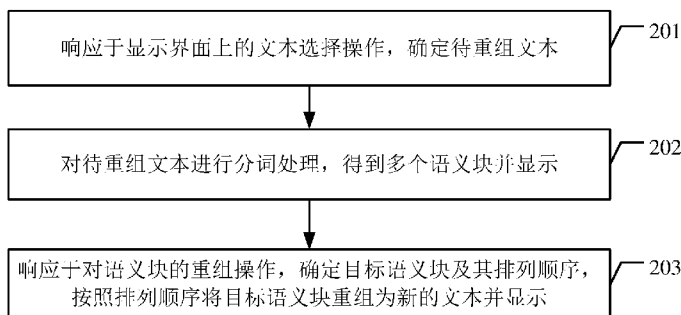


图 2

- 201 In response to a text selection operation on a display interface, determine text to be recombined
- 202 Performing word segmentation processing on the text to be recombined, to obtain a plurality of semantic blocks and display same
- 203 In response to a recombination operation on the semantic blocks, determine target semantic blocks and the arrangement sequence thereof, and recombine the target semantic blocks into new text according to the arrangement sequence and display same

(57) Abstract: Disclosed in the embodiments of the present application are a text recombination method and device, a terminal device, and a computer-readable storage medium. The method comprises: in response to a text selection operation on a display interface, determining text to be recombined; performing word segmentation processing on the text to be recombined, to obtain a plurality of semantic blocks and displaying same; and in response to a recombination operation on the semantic blocks, determining target semantic blocks and the arrangement sequence thereof, and recombining the target semantic blocks into new text according to the arrangement sequence and displaying same.

(57) 摘要: 本申请实施例公开了一种文本重组方法、装置、终端设备及计算机可读存储介质，方法包括：响应于显示界面上的文本选择操作，确定待重组文本，对该待重组文本进行分词处理，得到多个语义块并显示，响应于对语义块的重组操作，确定目标语义块及其排列顺序，根据所述排列顺序将所述目标语义块重组为新的文本并显示。

圳市南山区高新区科技中一路腾讯大厦35层,
Guangdong 518057 (CN)。 许耀峰(XU, Yaofeng);
中国广东省深圳市南山区高新区科技中一路
腾讯大厦35层, Guangdong 518057 (CN)。

(74) 代理人: 北 京 德 琦 知 识 产 权 代 理 有
限 公 司 (DEQI INTELLECTUAL PROPERTY LAW
CORPORATION); 中国北京市海淀区知春路1号
学院国际大厦7层, Beijing 100083 (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,
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))。

文本重组方法、装置、终端设备及计算机可读存储介质

本申请要求于 2017 年 6 月 1 日提交中国专利局、申请号为 201710403566.2、申请名称为“文本重组方法、装置、终端设备及计算机可读存储介质”的中国专利申请的优先权。

5 技术领域

本申请涉及文字处理技术领域，尤其涉及一种文本重组方法、装置、终端及可读存储介质。

发明背景

随着科技的发展，用户使用终端设备的时间越来越长，终端设备获得的重要信息也越多，用户也具有从喜欢的某段文字或与朋友的某段聊天内容中获得重要信息的需求，以从聊天内容中获得重要信息为例，目前通常是通过以下方式实现：对聊天内容进行选择、复制、粘贴到输入框，且粘贴到输入框之后，删除冗余文字，适当调整词语的先后顺序等等一系列的操作。操作过程复杂，容易出现误操作，用户体验不佳。

15 发明内容

本申请实施例的主要目的在于提供一种文本重组方法、装置、终端设备及计算机可读存储介质，旨在解决现有技术中从文本中提取重要信息的操作过程复杂，容易出现误操作，用户体验不佳的技术问题。

为实现上述目的，本申请实施例第一方面提供一种文本重组方法，
20 包括：

响应于显示界面上的文本选择操作，确定待重组文本；

对所述待重组文本进行分词处理，得到多个语义块并显示；

响应于对所述语义块的重组操作，确定目标语义块及其排列顺序，根据所述排列顺序将所述目标语义块重组为新的文本并显示。

为实现上述目的，本申请实施例第二方面一种文本重组装置，包括：

第一确定模块，用于响应于显示界面上的文本选择操作，确定待重组文本；

分词模块，用于对所述待重组文本进行分词处理，得到多个语义块并显示；

显示重组模块，用于响应于对所述语义块的重组操作，按确定目标语义块及其排列顺序，根据所述排列顺序将所述目标语义块重组为新的文本并显示。

为实现上述目的，本申请实施例第三方面提供一种终端设备，包括：存储器、处理器及存储在所述存储器上且在所述处理器上运行的计算机程序，所述处理器执行所述计算机程序时，实现如本申请实施例第一方面提供的文本重组方法中的各个步骤。

为实现上述目的，本申请实施例第四方面提供一种计算机可读存储介质，其上存储有计算机程序，所述计算机程序被处理器执行时，实现如本申请实施例第一方面提供的文本重组方法的各个步骤。

附图简要说明

为了更清楚的说明本申请实施例中的技术方案，下面将对实施例描述中所需要使用的附图作简单的介绍，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说，在不付出创造性劳动的前提下，还可以根据这些附图获得其它的附图。其中，

图 1 为一种终端设备的结构框图；

图 2 为本申请实施例中文本重组方法的流程示意图；

图 3 为本申请实施例中文本重组方法的另一流程示意图；

图 4a 为本申请实施例中分词处理后显示界面的示意图；

图 4b 为基于图 4a 执行选择操作后的显示界面的示意图；

图 4c 为基于图 4b 执行光标定位操作后的显示界面的示意图；

5 图 4d 为基于图 4c 执行选择操作后的显示界面的示意图；

图 5 为本申请实施例中对目标语义块的位置进行调整的方法的流程示意图；

图 6a 为基于图 4d 执行拖动操作的示意图；

图 6b 为基于图 6a 执行完拖动操作后的示意图；

10 图 7 为本申请实施例中文本重组方法的另一流程示意图；

图 8 为本申请实施例中显示界面响应长按操作的示意图；

图 9 为本申请实施例中图 6 所示步骤 603 的细化步骤的流程示意图；

图 10 为本申请实施例中图 6 所示步骤 604 的细化步骤的流程示意图；

15 图 11 为本申请实施例中文本重组装置的结构示意图；

图 12 为本申请实施例中文本重组装置的另一结构示意图；

图 13 为本申请实施例中文本重组装置的另一结构示意图。

实施方式

为使得本申请的发明目的、特征、优点能够更加的明显和易懂，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，所描述的实施例仅仅是本申请一部分实施例，而非全部实施例。基于本申请中的实施例，本领域技术人员在没有做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。

20

请参阅图 1，图 1 为一种终端设备 100 的结构示意图。本申请实施例提供的文本重组方法可应用于如图 1 所示的终端设备 100 中，终端设备 100 可以但不限于包括：需依靠电池维持正常运行且支持网络及下载功能的智能手机、笔记本电脑、平板电脑、智能穿戴设备等等。

5 如图 1 所示，终端设备 100 包括存储器 102、存储控制器 104，一个或多个（图中仅示出一个）处理器 106、外设接口 108、射频单元 110、按键单元 112、音频单元 114 以及显示单元 116。这些组件通过一条或多条通讯总线/信号线 122 相互通讯。

可以理解，图 1 所示的结构仅为示意，其并不对终端设备 100 的结
10 构造成限定。例如，终端设备 100 还可包括比图 1 所示更多或者更少的组件，或者具有与图 1 所示不同的配置。图 1 所示的各组件可以采用硬件、软件或其组合实现。

存储器 102 可用于存储计算机程序，如本申请实施例中的文本重组方法及装置对应的程序指令或模块，处理器 106 在执行存储在存储器 102
15 内的计算机程序时，实现下述图 2、图 3 及图 7 所示的文本重组方法中的各个步骤。

存储器 102，即计算机可读存储介质，可包括高速随机存储器，还可包括非易失性存储器，如一个或者多个磁性存储装置、闪存、或者其他非易失性固态存储器。在一些实例中，存储器 102 可进一步包括相对
20 于处理器 106 远程设置的存储器，这些远程存储器可以通过网络连接至终端设备 100。处理器 106 以及其他可能的组件对存储器 102 的访问可在存储控制器 104 的控制下进行。

外设接口 108 将各种输入/输出装置耦合至处理器 106 以及存储器 102。处理器 106 运行存储器 102 内的各种软件、指令以执行终端设备
25 100 的各种功能以及进行数据处理。

在一些实例中，外设接口 108，处理器 106 以及存储控制器 104 可以在单个芯片中实现。在其他一些实例中，他们可以分别由独立的芯片实现。

5 射频单元 110 用于接收以及发送电磁波，实现电磁波与电信号的相互转换，从而与通讯网络或者其他设备进行通讯。射频单元 110 可包括各种现有的用于执行这些功能的电路元件。按键单元 112 提供用户向终端设备 100 进行输入的接口，用户可以通过按下不同的按键以使终端设备 100 执行不同的功能。

音频单元 114 向用户提供音频接口，其可包括一个或多个麦克风、
10 一个或者多个扬声器以及音频电路。音频电路从外设接口 108 处接收声音数据，将声音数据转换为电信息，将电信息传输至扬声器。扬声器将电信息转换为人耳能听到的声波。音频电路还从麦克风处接收电信息，将电信号转换为声音数据，并将声音数据传输至外设接口 108 中以进行进一步的处理。音频数据可以从存储器 102 处或者通过射频单元 110 获
15 取。此外，音频数据也可以存储至存储器 102 中或者通过射频单元 110 进行发送。在一些实例中，音频单元 114 还可包括一个耳机插孔，用于向耳机或者其他设备提供音频接口。

显示单元 116 在终端设备 100 与用户之间提供一个输出界面。具体地，显示单元 116 向用户显示视频输出，这些视频输出的内容可包括文
20 字、图形、视频、及其任意组合。一些输出结果是对应于一些用户界面对象。进一步地，还在终端设备 100 与用户之间提供一个输入界面，用于接收用户的输入，例如用户的点击、滑动等手势操作，以使用户界面对象对这些用户的输入做出响应。检测用户输入的技术可以是基于电阻式、电容式或者其他任意可能的触控检测技术。

25 由于现有技术中从文本中提取重要信息的操作过程复杂、容易出现

误操作，用户体验不佳的技术问题。

为了解决上述问题，本申请实施例提出一种文本重组方法，通过将文本进行分词处理得到多个语义块，及通过重组操作的方式对语义块中的重要信息进行重新组合，得到新的文本，使得通过语义块重组的方式
5 就能够得到文本中的重要信息，而不需要通过选择、复制、粘贴、删除冗余等操作，操作更加简便，且降低误操作的概率，有效改善用户体验。

请参阅图 2，为本申请实施例中文本重组方法的流程示意图，该方法包括：

步骤 201、响应于显示界面上的文本选择操作，确定待重组文本；

10 在本申请实施例中，上述文本重组方法可以通过文本重组装置（以下简称：重组装置）实现，该重组装置可以为图 1 所示终端设备中的程序指令或模块。

其中，上述的文本重组方法可以在多种场景下使用，例如，可以应用在用户利用即时通讯应用程序聊天的场景、用户看新闻、看书、看短
15 信等等场景。

其中，用户可以通过文本选择操作触发该文本重组方法，且重组装置在检测到用户在显示界面上执行文本选择操作时，响应于该文本选择操作，确定待重组文本。可以理解的是，在不同场景下，该待重组文本是有差别的，例如，在聊天场景下，该待重组文本可以是一条或多条聊天内容，在看书场景下，该待重组文本可以是当前页面显示的所有内容，
20 或者是某一段内容，在实际应用中可以按照具体的场景设定待重组文本，此处不做限定。

步骤 202、对待重组文本进行分词处理，得到多个语义块并显示；

在本申请实施例中，重组装置将该待重组文本进行分词处理，得到
25 多个语义块并显示。

其中，语义块是具有语义的一个文字组合，可以是一个词，也可以是一个词组，例如，“鼠标”这个词组就是一个语义块。

步骤 203、响应于对语义块的重组操作，确定目标语义块及其排列顺序，按照排列顺序将目标语义块重组为新的文本并显示。

5 在本申请实施例中，显示界面上显示分词处理后的多个语义块后，用户可以选择进行重组的语义块，且还可以执行语义块的删除、位置调整等一系列的重组操作，重组装置在检测到重组操作之后，将响应于语义块的重组操作，按照重组操作确定的目标语义块的排列顺序显示目标语义块，以重组为新的文本，使得用户仅需要选择属于重要信息的语义
10 块进行重组，就能够得到所需要的重要信息。

在本申请实施例中，响应于显示界面上的文本选择操作，确定待重组文本，对该待重组文本进行分词处理，得到多个语义块并显示，响应于对语义块的重组操作，按照该重组操作确定的目标语义块的排列顺序显示目标语义块，以重组为新的文本。相对于现有技术，本申请实施例
15 中的技术方案中通过将文本进行分词处理得到多个语义块，及通过重组操作的方式对语义块中的重要信息进行重新组合，得到新的文本，使得通过语义块重组的方式就能够得到文本中的重要信息，而不需要通过选择、复制、粘贴、删除冗余等操作，操作更加简便，且降低误操作的概率，有效改善用户体验。

20 在本申请实施例中，为了方便操作，可以将显示界面划分为第一显示区域和第二显示区域，且该两个显示区域可以通过横向并列的方式排列，也可以通过纵向并列的方式排列。其中，该第一显示区域用于显示分词处理得到的多个语义块。

请参阅图 3，为本申请实施例中文本重组方法的另一流程示意图，
25 该方法包括：

步骤 301、响应于显示界面上的文本选择操作，确定待重组文本；

步骤 302、对待重组文本进行分词处理，得到多个语义块并在第一显示区域内显示；

步骤 303、响应于对第一显示区域内的语义块的选择操作，将选择的语义块作为目标语义块，并按照对目标语义块的选择顺序或光标所在位置在第二显示区域内显示目标语义块。

在本申请实施例中，在第一显示区域内显示待重组文本的多个语义块之后，用户可以对显示的语义块执行重组操作，该重组操作包括选择用于重组的目标语义块，删除已选择的目标语义块，改变目标语义块在新的文本中的位置等等。

其中，用户可对第一显示区域内的语义块执行选择操作，重组装置将响应于该选择操作，将选择的语义块作为目标语义块，并按照对目标语义块的选择顺序在第二显示区域内显示目标语义块。其中选择操作可以是点击操作，若终端设备为非触摸设备，该点击操作可以通过鼠标单击实现，若终端设备为触摸设备，则该点击操作可以通过鼠标点击实现，或者触摸点击操作实现。例如，重组装置在检测到用户对第一显示区域内的语义块 A 执行点击操作，则将语义块 A 作为目标语义块，并在第二显示区域内已有的目标语义块的后面显示语义块 A，若重组装置接着检测到用户对第一显示区域内的语义块 B 执行点击操作，则将语义块 B 作为目标语义块，并在第二显示区域中语义块 A 的后面显示语义块 B。即新选择的目标语义块排列在已有的目标语义块的后面。可以理解的是，在本申请实施例，用户可以对同一个语义块执行多次的选择操作。

请参阅图 4a，为本申请实施例中分词处理后显示界面的示意图，在图 4a 中，待重组文本为“我们已经是好友啦，一起聊天吧！”，该待重组文本分词处理后得到的语义块分别是“我们”、“已经”、“是”、“好

友”、“啦”，“，”、“一起”、“聊天”、“吧”、“！”及“ ”。且上述语义块显示的区域即为第一显示区域，横线下方的区域即为第二显示区域。

请参阅图 4b，为本申请实施例中基于图 4a 执行选择操作后的显示界面的示意图，在图 4b 中，用户对第一显示区域中的语义块“好友”
5 执行选择操作，则目标语义块“好友”在第二显示区域内显示，且接着用户对第一显示区域内的语义块“一起”执行选择操作，则目标语义块“一起”在第二显示区域中显示。

需要说明的是，在用户执行选择操作的过程中，若用户未使用到光标，则将按照上述的选择目标语义块的顺序进行显示，若用户使用到光标，即检测到用户执行光标定位操作，则将响应于该光标定位操作，在
10 第二显示区域内显示光标，且在用户执行选择操作之后，将选择的目标语义块显示在该光标所在的位置，以便通过光标定位的方式，实现目标语义块的插入。其中，光标定位操作可以是点击第二显示区域内的目标语义块，以便光标定位在该目标语义块的前面。请参阅图 4c，为本申请
15 实施例中基于图 4b 执行光标定位操作后的显示界面的示意图，在图 4b 中，用户点击“一起”，以便在目标语义块“好友”及“一起”之间定位光标“|”，若用户对第一显示区域中的语义块“聊天”执行选择操作，则将响应于该选择操作，将目标语义块“聊天”显示在光标所在的位置。如图 4d 所示，为本申请实施例中基于图 4c 执行选择操作后的显示界面
20 的示意图。

进一步的，为了实现第二显示区域内的目标语义块的删除操作，第二显示区域内显示的目标语义块包含删除标记，其中，该删除标记在显示时可以有多种不同的形式，且可以位于语义块的不同方位或角度上。如图 4b 所示，第二显示区域中的目标语义块的删除标记即为目标
25 语义块的右上角的圆圈，且该圆圈中包含“-”。且可以理解的是，图 4b

仅为可行的一个示意图，并不对本申请实施例的技术方案造成限定。

因此，还可以执行以下步骤：

步骤 304、响应于对第二显示区域内的目标语义块的删除标记的点击操作，按照点击操作删除指定的目标语义块。

- 5 用户对第二显示区域内的目标语义块的删除标记执行点击操作后，重组装置将响应于该点击操作，按照该点击操作删除指定的目标语义块，例如，若用户对目标语义块 C 的删除标记执行了点击操作，则将第二显示区域内的目标语义块 C 删除。

- 10 进一步的，还可以对第二显示区域中的目标语义块的位置进行调整，请参阅图 5，为本申请实施例中对目标语义块的位置进行调整的方法的流程示意图，该方法包括：

- 15 步骤 501、响应于对第二显示区域内的目标语义块的拖动操作，基于拖动操作拖动的目标语义块的实时位置，确定目标语义块的可插入位置，并在可插入位置显示目标语义块的虚拟语义块；其中，虚拟语义块与目标语义块具有相同的文本内容。

步骤 502、检测到拖动操作结束时，则判断目标语义块与虚拟语义块的重叠区域是否大于或等于预设值；

步骤 503、若大于或等于预设值，则利用目标语义块替换虚拟语义块；

- 20 步骤 504、若小于预设值，则取消虚拟语义块的显示，并将目标语义块还原至上一个位置处显示。

- 25 在本申请实施例中，对于第二显示区域内的至少两个目标语义块，用户还可以进一步调整目标语义块之间的位置关系，可以理解的是，目标语义块的位置关系也代表着目标语义块的排列顺序，且位置关系发生改变，则表明重组得到的新的文本也将发生改变。

响应于对第二显示区域内的目标语义块的拖动操作，基于该拖动操作拖动的目标语义块的实时位置显示被拖动的目标语义块的虚拟语义块，具体的，用户在执行拖动操作的过程中，该拖动操作拖动的目标语义块将随着拖动操作的移动而移动，且在拖动该目标语义块的过程中，

5 重组装置还将实时根据该目标语义块的实时位置确定以该实时位置为中心的预置区域内，可插入该目标语义块的位置，并在所有可插入目标语义块的位置上显示虚拟语义块，以使用户更好的拖动的目标语义块所需要插入的位置，其中，可插入目标语义块的位置可以是目标语义块的前面、后面等方位上。例如，请参阅图 4d，图 4d 中的目标语义块“一起”位于目标语义块“聊天”的后面。若需要将“一起”调整至“好友”与“聊天”之间，则可拖动“一起”，此时，第二显示区域内将显示“一起”的虚拟语义块，请参阅图 6a，为本申请实施例中基于图 4d 执行拖动操作的示意图。

10

在本申请实施例中，若拖动操作结束，则将判断被拖动的目标语义块与虚拟语义块的重叠区域是否大于或等于预设值，该预设值可以是目标语义块显示的区域大小的预置比例，若大于或等于预设值，表明用户将目标语义块拖动到其虚拟语义块的位置，将利用该目标语义块替换其虚拟语义块，请参阅图 6b，为本申请实施例中基于图 6a 执行完拖动操作后的示意图，如图 6b 所示，目标语义块“一起”替换了其虚拟语义块，完成了位置的调整。若小于预设值，则取消虚拟语义块的显示，并将被拖动的目标语义块还原至上一个位置处显示，该上一个位置指拖动操作执行之前，被拖动的目标语义块所在的位置，例如，若在图 6a 时，用户结束了拖动操作，则第二显示区域显示的内容将还原成图 4d。

15

20

可以理解的是，上述步骤 303 是针对目标语义块的选择、步骤 304 是针对目标语义块的删除、步骤 501 至 505 是针对目标语义块的位置调

25

整，在实际应用中，并限定选择、删除、位置调整之间执行的先后关系，在第二显示区域内存在目标语义块的情况下，用户可以基于其需要执行上述的选择、删除、位置调整中的任意一个或多个，且该多个操作之间也是互相独立的。

- 5 在本申请实施例中，在显示界面上的第一显示区域内显示待重组文本分词处理后得到的多个语义块之后，用户可以通过选择、删除、位置调整的方式进行文本的重组，且操作简单方便，不容易产生误操作，用户体验好。

10 请参阅图 7，为本申请实施例中文本重组方法的流程示意图，该方法包括：

步骤 701、响应于显示界面的长按操作，确定选择的文本，并显示功能项；

步骤 702、响应于对功能项中的重组项的选择操作，确定选择的文本为待重组文本；

- 15 步骤 703、利用预先设置的字典扫描待重组文本，基于扫描到的语义块构建待重组文本所有可能的有向无环图的集合；

步骤 704、利用逆向比较语义块的概率的方式从集合中选择满足预设条件的有向无环图，显示选择的有向无环图中的语义块；

- 20 步骤 705、响应于对语义块的重组操作，按照重组操作确定的目标语义块的排列顺序显示目标语义块，以重组为新的文本。

在本申请实施例中，用户可以通过特定的操作触发文本重组方法，例如，用户可以对显示界面上的文本执行长按操作，重组装置将响应该长按操作，确定选择的文本，并显示功能项，例如：请参阅图 8，为本申请实施例中显示界面响应长按操作的示意图，如图 8 所示，用户在对聊天内容“今天中午去哪吃？”执行长按操作之后，将显示功能项，该

25

功能项包括：复制、转发、编辑、收藏及撤回等等。其中，该编辑功能项即为重组项，重组装置可响应于对功能项中的编辑（重组项）的选择操作，确定选择的文本即为待重组文本。

可以理解的是，通过长按操作及选择操作，能够有效的触发文本重组过程，相对于手指大面积按压文本这种依赖于触摸屏的灵敏性和大面积按压的方式，触发效果更好。

在本申请实施例中，本地保存有字典，其中，该字典以离线包的形式存储，用户每次登陆时后台自动更新离线包，省去用户频繁更新应用程序的烦恼。

其中，字典中包含大量记录，且每一行记录中至少包含语义块及语义块的频数，此外，还可以包含该语义块的词性。其中，语义块的频数是通过统计语义块的使用次数得到的，且该频数可以用于计算语义块的概率，且语义块的概率等于其频数与字典中所有语义块的频数之和的商。

在本申请实施例中，重组装置将利用上述字典扫描待重组文本，基于扫描到的语义块构建待重组文本的所有可能的有向无环图的集合，其中，有向无环图中的节点即为语义块，其方向为语义块在待重组文本中的先后顺序。

例如，若待重组文本为“我们大概下午2点左右出发”的有向无环图的集合包括：我→们→大→概→下→午→2→点→左→右→出→发；我们→大概→下午→2→点→左→右→出发；我们→大概→下午→2点→左右→出发。

重组装置利用逆向比较语义块的概率的方式从集合中选择满足预设条件的有向无环图，显示选择的有向无环图中的语义块。

在本申请实施例中，重组装置通过利用预先设置的字典扫描待重组

文本，并构建该待重组文本的所有可能的有向无环图的集合，通过逆向比较语义块的概率的方式从集合中选择满足预设条件的有向无环图，显示选择的有向无环图，使得能够有效的实现分词处理，且提高分词的准确性及性能。

5 请参阅图 9，为本申请实施例中图 7 所示步骤 703 的细化步骤的流程示意图，该步骤 703 包括：

步骤 901、调用正则表达式扫描待重组文本，并利用扫描到的词或词组与字典中的语义块进行匹配；

10 这里，正则表达式是指对文本操作的一种逻辑公式，使用预先设置的一些词或词组组成一个“规则词串”，这个“规则词串”用来表达对待重组文本的一种过滤逻辑。

步骤 902、若匹配到，则将词或词组划分为第一类别的语义块，且利用与词或词组匹配的语义块的频数计算第一类别的语义块的概率；

15 步骤 903、若未匹配到，且词或词组的语言类型与字典的语言类型不一致，则确定词或词组为第二类别的语义块，并利用马尔科夫模型中的维比特算法计算第二类别的语义块的概率；

步骤 904、按照第一类别的语义块和/或第二类别的语义块在待重组文本中的位置，构建所有可能的有向无环图，得到有向无环图的初始集合。

20 在本申请实施例中，对于待重组文本，重组装置将调用正则表达式扫描待重组文本，并利用扫描到的词或词组与字典中的语义块进行匹配，例如，若扫描到的词为“我”，则用“我”与词典中的语义块进行匹配，以确定词典中是否含“我”这个词的语义块。

25 其中，若匹配到，这表明该词或词组是语义块，且确定该词或词组为第一类别的语义块，其中，该第一类别是指词典的语言类型的类别，

例如，若词典为汉字词典，则第一类别为汉字，若词典为英语词典，则第一类别为英语。进一步的，由于词典中包含语义块的频数，重组装置还将利用匹配的语义块的频数计算概率，例如，待重组文本中的词“我”在词典中匹配到，则表明“我”是语义块，则将利用词典中“我”的频数计算其概率。

其中，若未匹配到，则将进一步判断该词或词组的语言类型与字典的语言类型是否一致，若一致，则确定并不存在该词或词组匹配的语义块，若不一致，则确定该词或词组为第二类别的语义块，该第二类别是指不同于第一类别的类别，例如，若第一类别是汉字，则第二类别是英文、数字、符号等等。且对于第二类别的语义块，将利用马尔科夫模型中的维特比算法计算该第二类别的语义块的概率。例如，若词组“友啦”在汉字词典中未匹配到，则表明该词组并非为语义块，若词组“sorry”在汉字词典中未匹配到，则表明该词组并非为汉字。可以理解的是，词或词组里面包含的所有内容是属于同一种语言类型的，并不存在“我 ok”“是 488”这种同时包含两种不同语言类型的词或词组，以便能够更好的进行语义块的划分。

通过上述方式，重组装置能够得到待重组文本中的所有语义块，且每一个语义块的概率。并进一步的，利用该待重组文本中的所有语义块，包括第一类别的语义块和/或第二类别的语义块在待重组文本中的位置，构建所有可能的有向无环图，得到有向无环图的初始集合。

在本申请实施例中，通过上述方式能够得到待重组文本的有向无环图的初始集合，以便选择满足预设条件的有向无环图，并作为待重组文本的最终的分词结果进行显示。

请参阅图 10，为本申请实施例中图 7 所示步骤 704 的细化步骤的流程示意图，该步骤 704 包括：

步骤 1001、比较第 m 个集合中有向无环图逆向的第 i 个节点的语义块的概率；

步骤 1002、若在逆向的第 i 个节点的语义块中，仅有一个语义块的概率最大，则确定概率最大的语义块所在的有向无环图为满足预设条件的有向无环图；

步骤 1003、若在逆向的第 i 个节点的语义块中，有至少两个语义块的概率最大，则将概率最大的至少两个语义块分别所在的有向无环图作为第 $m+1$ 个集合，且令 $i=i+1$ ， $m=m+1$ ，返回步骤 1001。

在本申请实施例中，重组装置在得到包含所有可能的有向无环图的初始集合之后，将从该初始集合中选择满足预设条件的有向无环图。由于选择的过程涉及到循环，且是通过逐渐缩小有向无环图的集合的方式实现的，因此，为了描述的方便，将初始集合作为第 1 个集合，且设置编号 i 和 m ， i 和 m 的初始值均为 1 且为正整数，先比较第 m 个集合中有向无环图逆向的第 i 个节点的语义块的概率，即从最后的一个节点的语义块开始比较。

若对于第 m 个集合中的有向无环图，逆向的第 i 个节点的语义块中，仅有一个语义块的概率最大，则确定该概率最大的语义块所在的有向无环图是满足预设条件的，该有向无环图中包含的语义块是最终分词处理后的语义块。

若对于第 m 个集合中的有向无环图，逆向的第 i 个节点的语义块中，有至少两个语义块的概率最大，则将概率最大的至少两个语义块分别所在的有向无环图作为第 $m+1$ 个集合，例如，对于包含 5 个有向无环图的第 3 个集合，在比较逆向的第 2 个节点的语义块时，逆向的第 2 个节点的语义块中，有 3 个语义块的概率相同且最大，则将该 3 个语义块分别所在的有向无环图构成第 4 个集合。其中，在得到第 $m+1$ 个集合之后，

令 $i=i+1$, $m=m+1$, 返回上述的比较第 m 个集合中有向无环图逆向的第 i 个节点的语义块的概率的步骤, 即步骤 1001。例如, 继续对包含 3 个有向无环图的第 4 个集合中, 有向无环图的第 3 个节点的语义块进行比较。

- 5 在本申请实施例中, 通过逆向的循环比较的方式, 能够找出有向无环图集合中满足预设条件的有向无环图, 以实现最佳的分词处理过程。

请参阅图 11, 为本申请实施例中文本重组装置的结构示意图, 该装置包括:

- 第一确定模块 1101, 用于响应于显示界面上的文本选择操作, 确定
10 待重组文本;

 分词模块 1102, 用于对待重组文本进行分词处理, 得到多个语义块并显示;

 显示重组模块 1103, 用于响应于对语义块的重组操作, 按照重组操作确定的目标语义块的排列顺序显示目标语义块, 以重组为新的文本。

- 15 本申请实施例中的文本重组装置中各程序模块实现各自功能的具体过程, 请参见上述图 2 所示实施例中描述的内容, 此处不做赘述。

- 在本申请实施例中, 响应于显示界面上的文本选择操作, 确定待重组文本, 对该待重组文本进行分词处理, 得到多个语义块并显示, 响应于对语义块的重组操作, 按照该重组操作确定的目标语义块的排列顺序
20 显示目标语义块, 以重组为新的文本。相对于现有技术, 本申请实施例中的技术方案中通过将文本进行分词处理得到多个语义块, 及通过重组操作的方式对语义块中的重要信息进行重新组合, 得到新的文本, 使得通过语义块重组的方式就能够得到文本中的重要信息, 而不需要通过选择、复制、粘贴、删除冗余等操作, 操作更加简便, 且降低误操作的概
25 率, 有效改善用户体验。

在本申请实施例中，为了方便操作，可以将显示界面划分为第一显示区域和第二显示区域，且该两个显示区域可以通过横向并列的方式排列，也可以通过纵向并列的方式排列。其中，该第一显示区域用于显示分词处理得到的多个语义块。

5 请参阅图 12，为本申请实施例中文本重组装置的另一结构示意图，包括：如图 11 所示实施例中的第一确定模块 1101、分词模块 1102 及显示重组模块 1103，且与图 11 所示实施例中描述的内容相似，此处不做赘述。

在本申请实施例中，显示重组模块 1103 包括：

10 选择显示模块 1201，用于响应于对第一显示区域内的语义块的选择操作，将选择的语义块作为目标语义块，并按照对目标语义块的选择顺序或光标所在位置在第二显示区域内显示目标语义块。

其中，第二显示区域内显示的目标语义块包含删除标记；

显示重组模块 1103 还包括：

15 删除模块 1202，用于响应于对第二显示区域内的目标语义块的删除标记的点击操作，按照点击操作删除指定的目标语义块。

进一步的，显示重组模块 1103 还包括：

20 拖动插入模块 1203，用于响应于对第二显示区域内的目标语义块的拖动操作，基于拖动操作拖动的目标语义块的实时位置，确定目标语义块的可插入位置，并在可插入位置显示目标语义块的虚拟语义块；

判断模块 1204，用于检测到拖动操作结束时，则判断目标语义块与虚拟语义块的重叠区域是否大于或等于预设值；

替换模块 1205，用于若大于或等于预设值，则利用目标语义块替换虚拟语义块；

25 取消还原模块 1206，用于若小于预设值，则取消虚拟语义块的显示，

并将目标语义块还原至上一个位置处显示，上一个位置指拖动操作执行之前，目标语义块所在的位置。

本申请实施例中的文本重组装置中各程序模块实现各自功能的具体过程，请参见上述图 3 及图 5 所示实施例中描述的内容，此处不做赘述。

在本申请实施例中，在显示界面上的第一显示区域内显示待重组文本分词处理后得到的多个语义块之后，用户可以通过选择、删除、位置调整的方式进行文本的重组，且操作简单方便，不容易产生误操作，用户体验好。

请参阅图 13，为本申请实施例中文本重组装置的另一结构示意图，包括：

如图 11 所示实施例中的第一确定模块 1101、分词模块 1102 及显示重组模块 1103，且与图 11 所示实施例中描述的内容相似，此处不做赘述。

在本申请实施例中，第一确定模块 1101 包括：

确定显示模块 1301，用于响应于显示界面的长按操作，确定选择的文本，并显示功能项；

第二确定模块 1302，用于响应于对功能项中的重组项的选择操作，确定选择的文本为待重组文本。

在本申请实施例中，分词模块 1102 包括：

扫描构建模块 1303，用于利用预先设置的字典扫描待重组文本，基于扫描到的语义块构建待重组文本所有可能的有向无环图的集合；

选择模块 1304，用于利用逆向比较语义块的概率的方式从集合中选择满足预设条件的有向无环图，显示选择的有向无环图中的语义块，其中，语义块的概率基于语义块在字典中的频数得到。

在本申请实施例中，字典包含语义块与语义块的频数；

则扫描构建模块 1303 具体用于：

调用正则表达式扫描待重组文本，并利用扫描到的词或词组与字典中的语义块进行匹配；

- 5 若匹配到，则将词或词组划分为第一类别的语义块，且利用与词或词组匹配的语义块的频数计算第一类别的语义块的概率；

若未匹配到，且词或词组的语言类型与字典的语言类型不一致，则确定词或词组为第二类别的语义块，并利用马尔科夫模型中的维比特算法计算第二类别的语义块的概率；

- 10 按照第一类别的语义块和/或第二类别的语义块在待重组文本中的位置，构建所有可能的有向无环图，得到有向无环图的初始集合。

其中，选择模块 1304 具体用于：

比较第 m 个集合中有向无环图逆向的第 i 个节点的语义块的概率， i 、 m 为正整数，且 i 、 m 的初始值为 1，且第 1 个集合为初始集合；

- 15 若在逆向的第 i 个节点的语义块中，仅有一个语义块的概率最大，则确定概率最大的语义块所在的有向无环图为满足预设条件的有向无环图；

- 若在逆向的第 i 个节点的语义块中，有至少两个语义块的概率最大，则将概率最大的至少两个语义块分别所在的有向无环图作为第 $m+1$ 个集合，且令 $i=i+1$ ， $m=m+1$ ，返回比较第 m 个集合中有向无环图逆向的第 i 个节点的语义块的概率的步骤。
- 20

本申请实施例中的文本重组装置中各程序模块实现各自功能的具体过程，请参见上述图 3、图 5 及图 7 所示实施例中描述的内容，此处不做赘述。

- 25 在本申请实施例中，重组装置通过利用预先设置的字典扫描待重组

文本，并构建该待重组文本的所有可能的有向无环图的集合，通过逆向比较语义块的概率的方式从集合中选择满足预设条件的有向无环图，显示选择的有向无环图，使得能够有效的实现分词处理，且提高分词的准确性及性能。且通过上述方式，重组装置能够得到待重组文本中的所有语义块，且每一个语义块的概率。并进一步的，利用该待重组文本中的所有语义块，包括第一类别的语义块和/或第二类别的语义块在待重组文本中的位置，构建所有可能的有向无环图，得到有向无环图的初始集合。

需要说明的是，在本申请实施例中，上述图 11 至图 13 所示的文本重组装置可应用于图 1 所示的终端设备中，用于实现上述实施例中的文本重组方法。

且本申请实施例还提供一种计算机可读存储介质，其上存储有计算机程序，计算机程序被处理器执行时，实现上述实施例中的文本重组方法。

在本申请所提供的几个实施例中，应该理解到，所揭露的装置和方法，可以通过其它的方式实现。例如，以上所描述的装置实施例仅仅是示意性的，例如，所述模块的划分，仅仅为一种逻辑功能划分，实际实现时可以有另外的划分方式，例如多个模块或组件可以结合或者可以集成到另一个系统，或一些特征可以忽略，或不执行。另一点，所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口，装置或模块的间接耦合或通信连接，可以是电性，机械或其它的形式。

所述作为分离部件说明的模块可以是或者也可以不是物理上分开的，作为模块显示的部件可以是或者也可以不是物理模块，即可以位于一个地方，或者也可以分布到多个网络模块上。可以根据实际的需要选择其中的部分或者全部模块来实现本实施例方案的目的。

另外，在本申请各个实施例中的各功能模块可以集成在一个处理模

块中，也可以是各个模块单独物理存在，也可以两个或两个以上模块集成在一个模块中。上述集成的模块既可以采用硬件的形式实现，也可以采用软件功能模块的形式实现。

所述集成的模块如果以软件功能模块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，本申请实施例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（ROM，Read-Only Memory）、随机存取存储器（RAM，Random Access Memory）、磁碟或者光盘等各种可以存储程序代码的介质。

需要说明的是，对于前述的各方法实施例，为了简便描述，故将其都表述为一系列的动作组合，但是本领域技术人员应该知悉，本申请实施例并不受所描述的动作顺序的限制，因为依据本申请实施例，某些步骤可以采用其它顺序或者同时进行。其次，本领域技术人员也应该知悉，说明书中所描述的实施例均属于优选实施例，所涉及的动作和模块并不一定都是本申请实施例所必须的。

在上述实施例中，对各个实施例的描述都各有侧重，某个实施例中

没有详述的部分，可以参见其它实施例的相关描述。

以上为对本申请实施例所提供的一种文本重组方法、装置、终端及可读存储介质的描述，对于本领域的技术人员，依据本申请实施例的思想，在具体实施方式及应用范围上均会有改变之处，综上，本说明书内容不应理解为对本申请的限制。

权利要求书

1、一种文本重组方法，其特征在于，应用于终端设备，所述方法包括：

响应于显示界面上的文本选择操作，确定待重组文本；

5 对所述待重组文本进行分词处理，得到多个语义块并显示；

响应于对所述语义块的重组操作，确定目标语义块及其排列顺序，根据所述排列顺序将所述目标语义块重组为新的文本并显示。

2、根据权利要求 1 所述的方法，其特征在于，所述显示界面包含第一显示区域及第二显示区域，所述第一显示区域用于显示分词处理得
10 到的多个语义块；

则所述响应于对所述语义块的重组操作，确定目标语义块及其排列顺序，根据所述排列顺序将所述目标语义块重组为新的文本并显示包括：

响应于对所述第一显示区域内的语义块的选择操作，将选择的语义
15 块作为目标语义块，并按照对所述目标语义块的选择顺序在所述第二显示区域内显示所述目标语义块。

3、根据权利要求 1 所述的方法，其特征在于，所述显示界面包含第一显示区域及第二显示区域，所述第一显示区域用于显示分词处理得到的多个语义块；

20 则所述响应于对所述语义块的重组操作，确定目标语义块及其排列顺序，根据所述排列顺序将所述目标语义块重组为新的文本并显示包括：

响应于对所述第一显示区域内的语义块的选择操作，将选择的语义块作为目标语义块，并按照所述选择操作以及光标所在位置在所述第二
25 显示区域内显示所述目标语义块。

4、根据权利要求 2 或者 3 所述的方法，其特征在于，所述第二显示区域内显示的目标语义块包含删除标记；

所述方法还包括：

5 响应于对所述第二显示区域内的目标语义块的删除标记的点击操作，按照所述点击操作删除指定的目标语义块。

5、根据权利要求 2 或者 3 所述的方法，其特征在于，所述方法还包括：

10 响应于对所述第二显示区域内的目标语义块的拖动操作，基于所述拖动操作拖动的目标语义块的实时位置，确定所述目标语义块的可插入位置，并在所述可插入位置显示所述目标语义块的虚拟语义块；

检测到所述拖动操作结束时，判断所述目标语义块与所述虚拟语义块的重叠区域是否大于或等于预设值；

若大于或等于预设值，则利用所述目标语义块替换所述虚拟语义块。

15 6、根据权利要求 1 所述的方法，其特征在于，所述对所述待重组文本进行分词处理，得到多个语义块并显示的步骤包括：

利用预先设置的字典扫描所述待重组文本，基于扫描到的语义块构建所述待重组文本所有可能的有向无环图的集合；

20 利用逆向比较语义块的概率的方式从所述集合中选择满足预设条件的有向无环图，显示选择的有向无环图中的语义块，其中，所述字典包含语义块与所述语义块的频数；所述语义块的概率基于所述语义块在所述字典中的频数得到。

25 7、根据权利要求 6 所述的方法，其特征在于，所述利用预先设置的字典扫描所述待重组文本，基于扫描到的语义块构建所述待重组文本所有可能的有向无环图的集合的步骤包括：

调用正则表达式扫描所述待重组文本，并利用扫描到的词或词组与
所述字典中的语义块进行匹配；

若匹配到，则将所述词或词组划分为第一类别的语义块，且利用与
所述词或词组匹配的语义块的频数计算所述第一类别的语义块的概率；

5 按照所述第一类别的语义块在所述待重组文本中的位置，构建所有
可能的有向无环图，得到有向无环图的初始集合。

8、根据权利要求 6 所述的方法，其特征在于，所述利用预先设置
的字典扫描所述待重组文本，基于扫描到的语义块构建所述待重组文本
所有可能的有向无环图的集合的步骤包括：

10 调用正则表达式扫描所述待重组文本，并利用扫描到的词或词组与
所述字典中的语义块进行匹配；

若未匹配到，且所述词或词组的语言类型与所述字典的语言类型不
一致，则确定所述词或词组为第二类别的语义块，并计算所述第二类别
的语义块的概率；

15 按照所述第二类别的语义块在所述待重组文本中的位置，构建所有
可能的有向无环图，得到有向无环图的初始集合。

9、根据权利要求 6 所述的方法，其特征在于，所述利用预先设置
的字典扫描所述待重组文本，基于扫描到的语义块构建所述待重组文本
所有可能的有向无环图的集合的步骤包括：

20 调用正则表达式扫描所述待重组文本，并利用扫描到的词或词组与
所述字典中的语义块进行匹配；

若匹配到，则将所述词或词组划分为第一类别的语义块，且利用与
所述词或词组匹配的语义块的频数计算所述第一类别的语义块的概率；

25 若未匹配到，且所述词或词组的语言类型与所述字典的语言类型不
一致，则确定所述词或词组为第二类别的语义块，并计算所述第二类别

的语义块的概率；

按照所述第一类别的语义块和所述第二类别的语义块在所述待重组文本中的位置，构建所有可能的有向无环图，得到有向无环图的初始集合。

- 5 10、根据权利要求 7 至 9 中任一项所述的方法，其特征在于，所述利用逆向比较语义块的概率的方式从所述集合中选择满足预设条件的有向无环图，显示选择的有向无环图中的语义块的步骤包括：

比较第 m 个集合中有向无环图逆向的第 i 个节点的语义块的概率， i 、 m 为正整数，且 i 、 m 的初始值为 1，且第 1 个集合为所述初始集合；

- 10 若在逆向的所述第 i 个节点的语义块中，仅有一个语义块的概率最大，则确定概率最大的语义块所在的有向无环图为所述满足预设条件的有向无环图；

- 若在逆向的所述第 i 个节点的语义块中，有至少两个语义块的概率最大，则将概率最大的至少两个语义块分别所在的有向无环图作为第
15 $m+1$ 个集合，且令 $i=i+1$ ， $m=m+1$ ，返回所述比较第 m 个集合中有向无环图逆向的第 i 个节点的语义块的概率的步骤。

11、一种文本重组装置，其特征在于，所述装置包括：

第一确定模块，用于响应于显示界面上的文本选择操作，确定待重组文本；

- 20 分词模块，用于对所述待重组文本进行分词处理，得到多个语义块并显示；

显示重组模块，用于响应于对所述语义块的重组操作，确定目标语义块及其排列顺序，根据所述排列顺序将所述目标语义块重组为新的文本并显示。

- 25 12、根据权利要求 11 所述的装置，其特征在于，所述显示界面包

含第一显示区域及第二显示区域，所述第一显示区域用于显示分词处理得到的多个语义块；

则显示重组模块包括：

5 选择显示模块，用于响应于对所述第一显示区域内的语义块的选择操作，将选择的语义块作为目标语义块，并按照对所述目标语义块的选择顺序在所述第二显示区域内显示所述目标语义块。

13、根据权利要求 11 所述的装置，其特征在于，所述显示界面包含第一显示区域及第二显示区域，所述第一显示区域用于显示分词处理得到的多个语义块；

10 则显示重组模块包括：

选择显示模块，用于响应于对所述第一显示区域内的语义块的选择操作，将选择的语义块作为目标语义块，并按照所述选择操作以及光标所在位置在所述第二显示区域内显示所述目标语义块。

14、根据权利要求 12 或者 13 所述的装置，其特征在于，所述第二显示区域内显示的目标语义块包含删除标记；

所述显示重组模块还包括：

删除模块，用于响应于对所述第二显示区域内的目标语义块的删除标记的点击操作，按照所述点击操作删除指定的目标语义块。

15、根据权利要求 12 或者 13 所述的装置，其特征在于，所述显示重组模块还包括：

20 拖动插入模块，用于响应于对所述第二显示区域内的目标语义块的拖动操作，基于所述拖动操作拖动的目标语义块的实时位置，确定所述目标语义块的可插入位置，并在所述可插入位置显示所述目标语义块的虚拟语义块；

25 判断模块，用于检测到所述拖动操作结束时，则判断所述目标语义

块与所述虚拟语义块的重叠区域是否大于或等于预设值；

替换模块，用于若大于或等于预设值，则利用所述目标语义块替换所述虚拟语义块。

16、根据权利要求 11 至 13 任意一项所述的装置，其特征在于，所述分词模块包括：

扫描构建模块，用于利用预先设置的字典扫描所述待重组文本，基于扫描到的语义块构建所述待重组文本所有可能的有向无环图的集合；

选择模块，用于利用逆向比较语义块的概率的方式从所述集合中选择满足预设条件的有向无环图，显示选择的有向无环图中的语义块，其中，所述字典包含语义块与所述语义块的频数；所述语义块的概率基于所述语义块在所述字典中的频数得到。

17、一种终端设备，包括存储器、处理器及存储在所述存储器上且在所述处理器上运行的计算机程序，其特征在于，所述处理器执行所述计算机程序时，实现如权利要求 1 至 10 任意一项所述的文本重组方法中的各个步骤。

18、一种计算机可读存储介质，其上存储有计算机程序，其特征在于，所述计算机程序被处理器执行时，实现如权利要求 1 至 10 任意一项所述的文本重组方法的各个步骤。

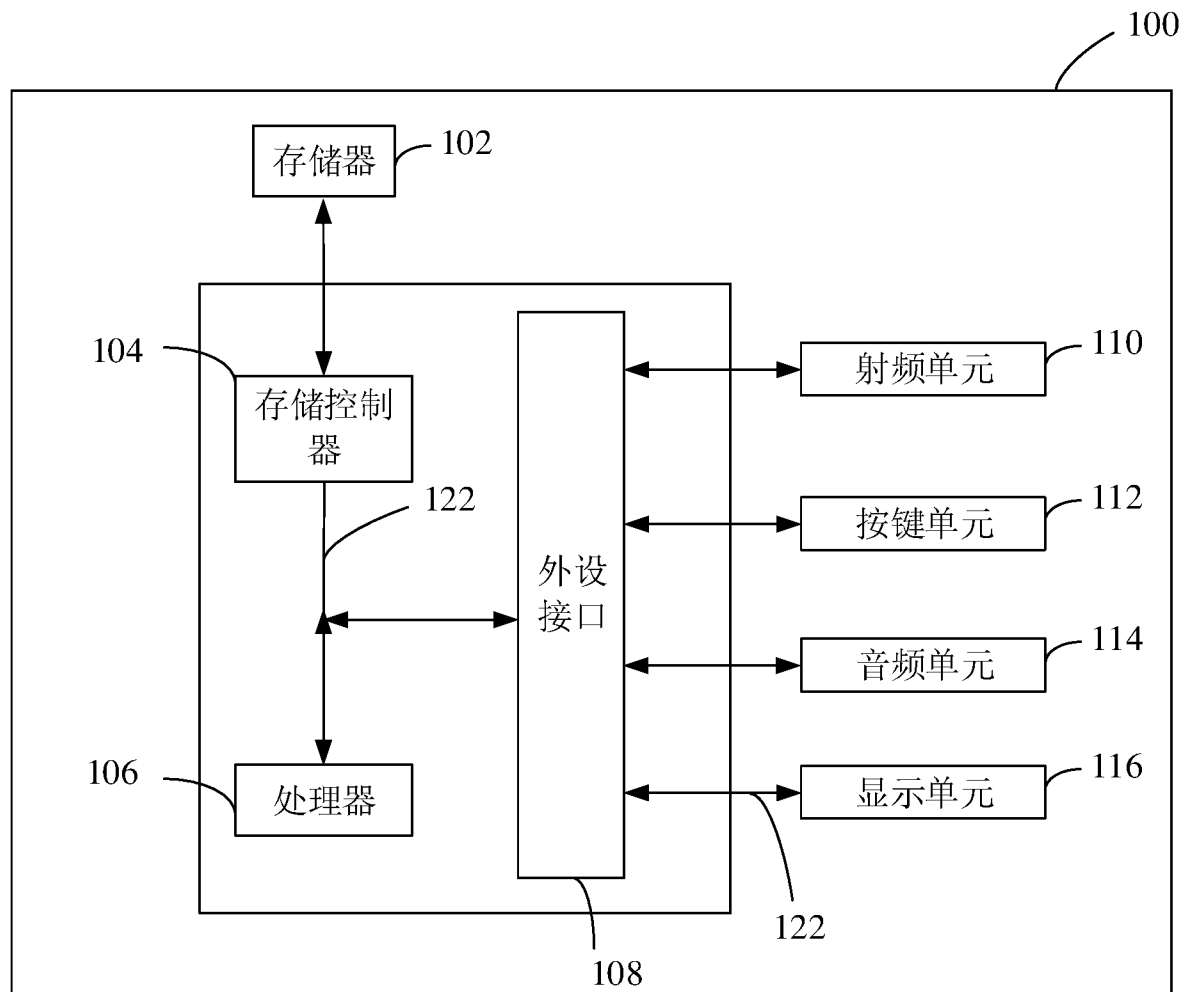


图 1

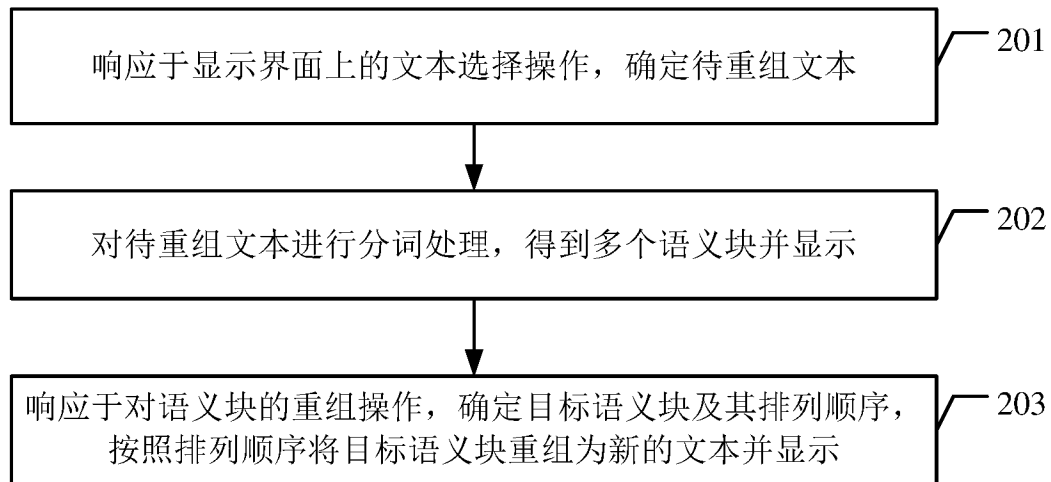


图 2

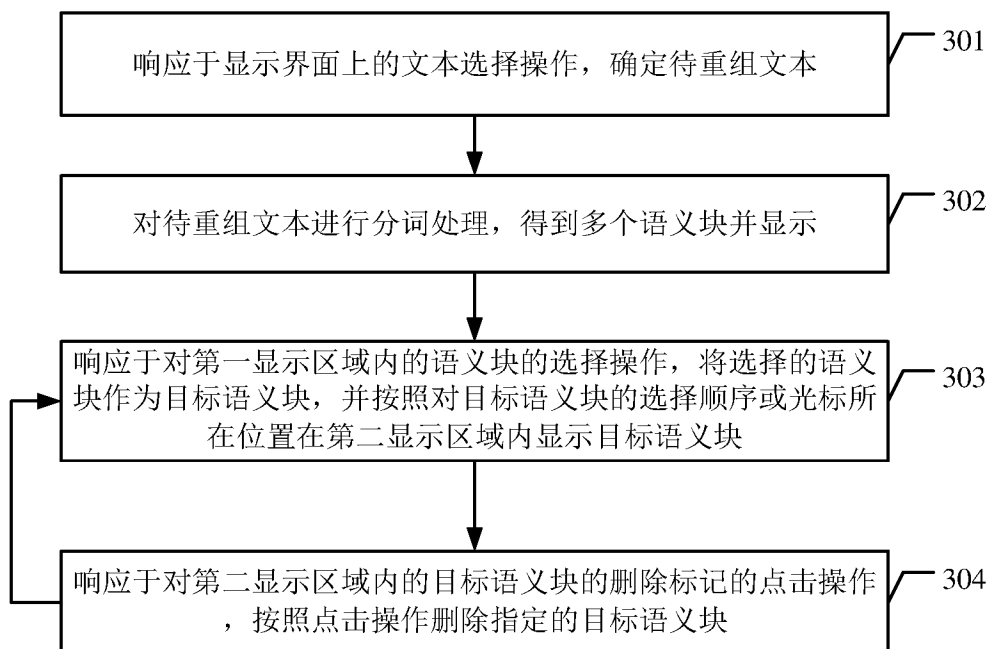


图 3

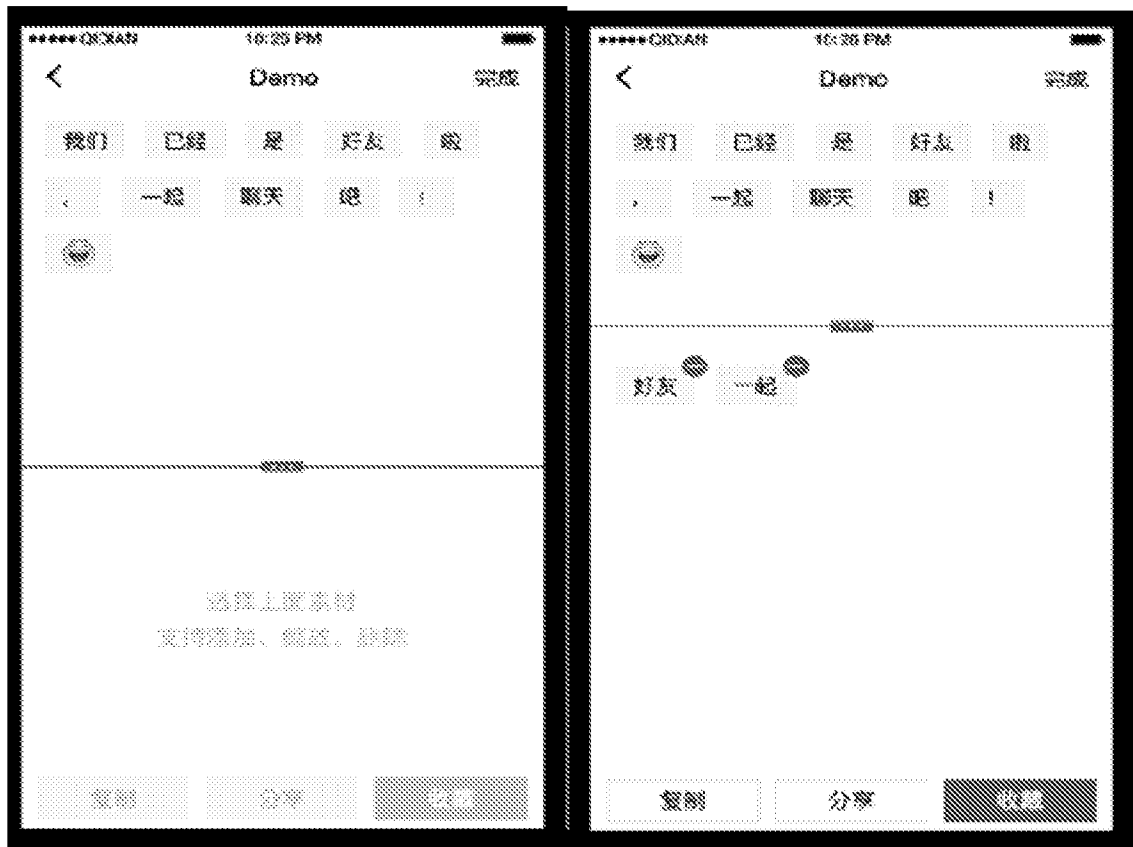


图 4a

图 4b

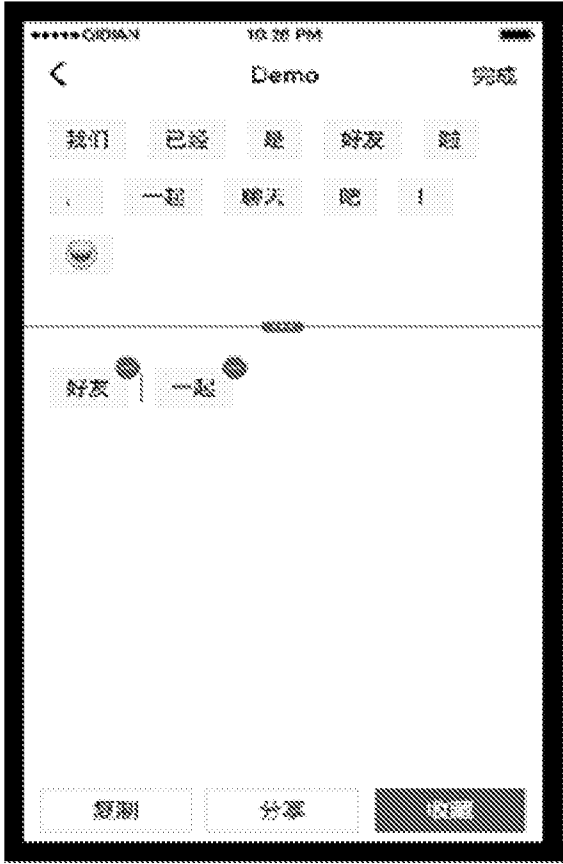


图 4c

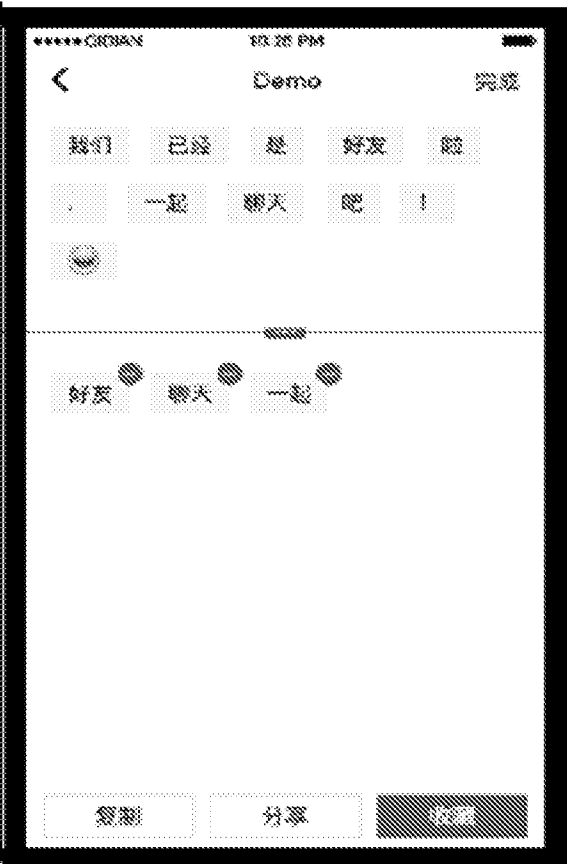


图 4d

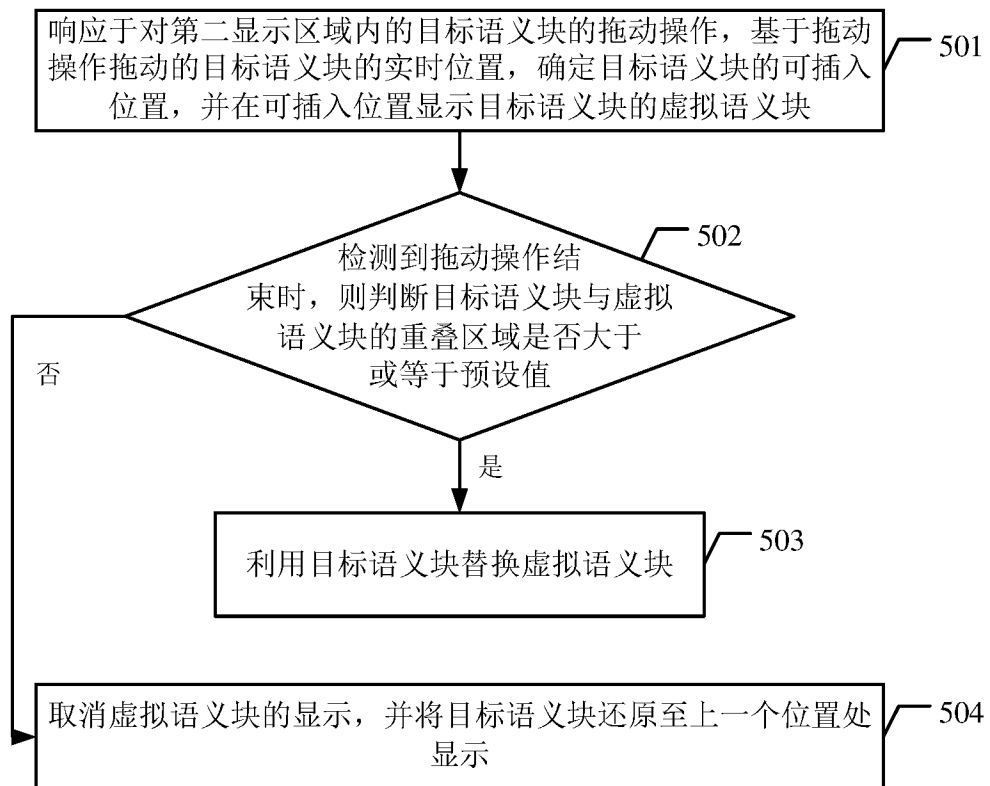


图 5

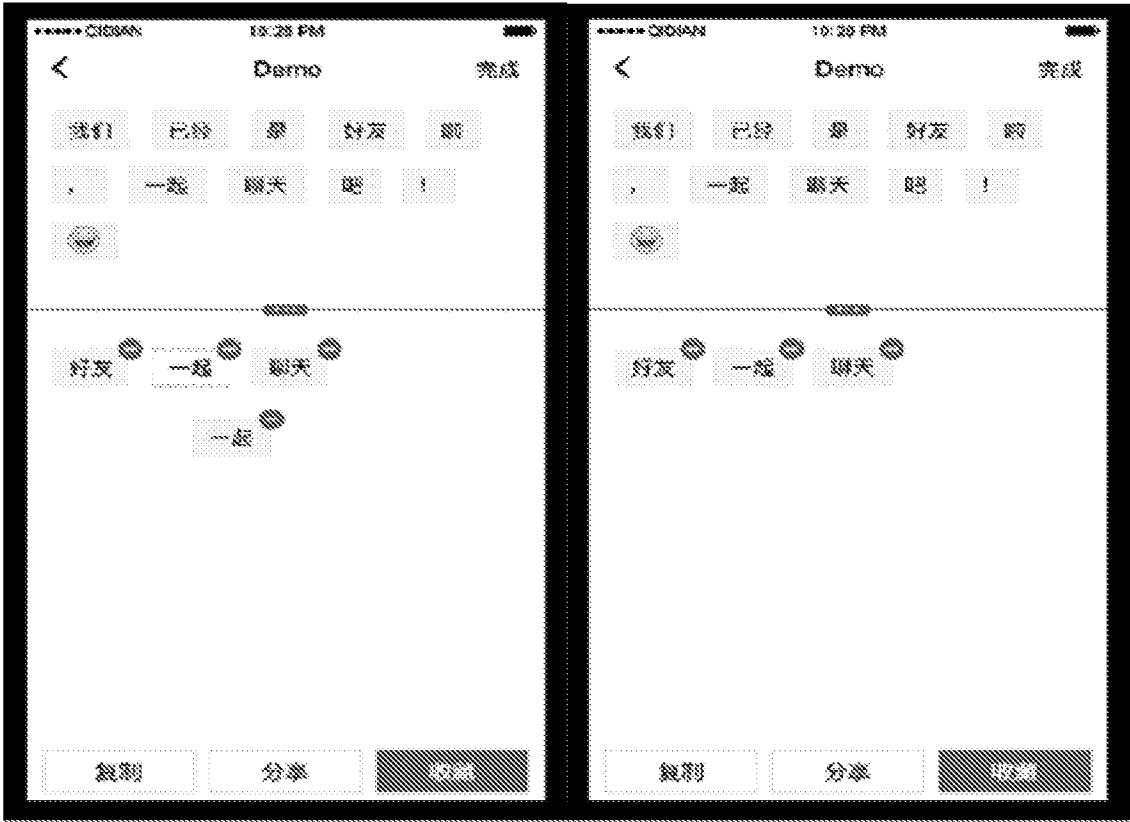


图 6a

图 6b

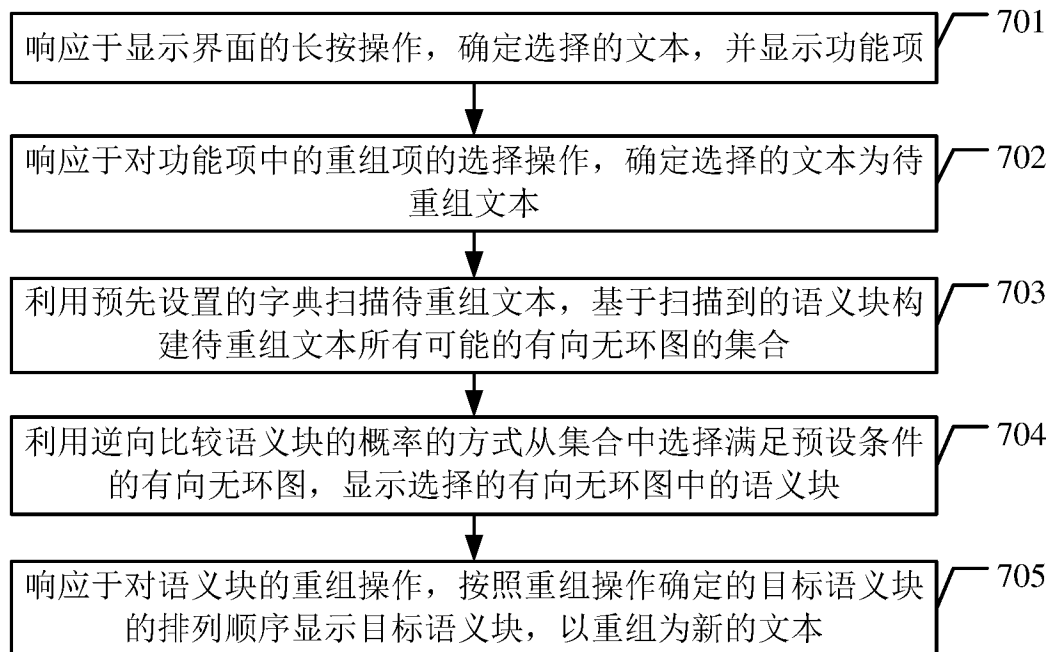


图 7



图 8

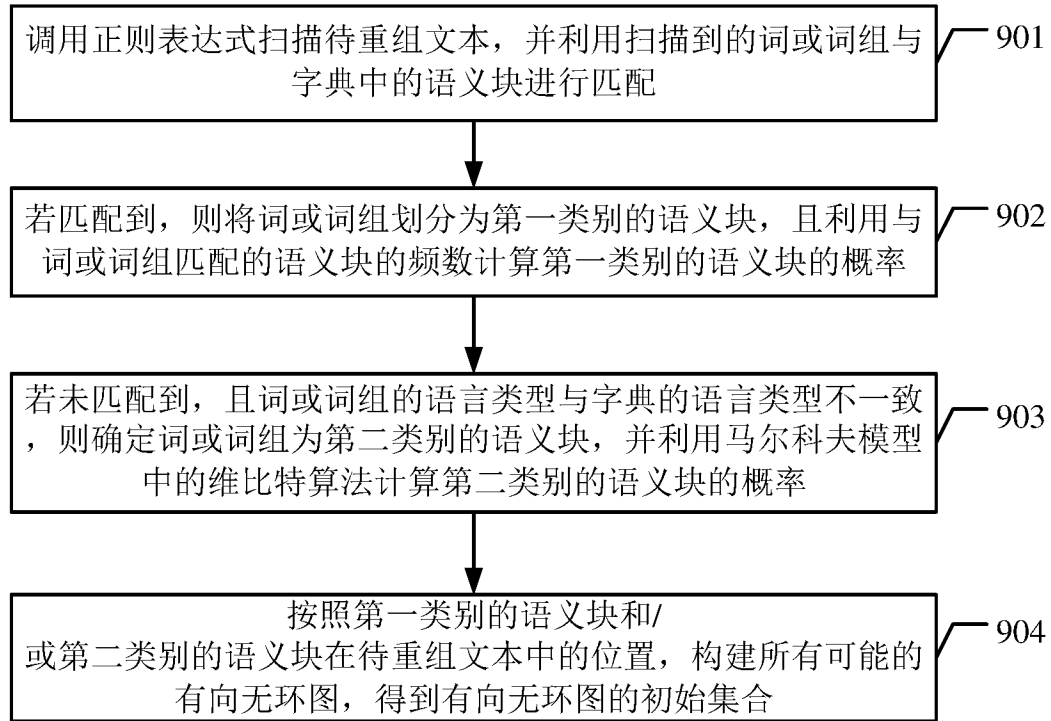


图 9

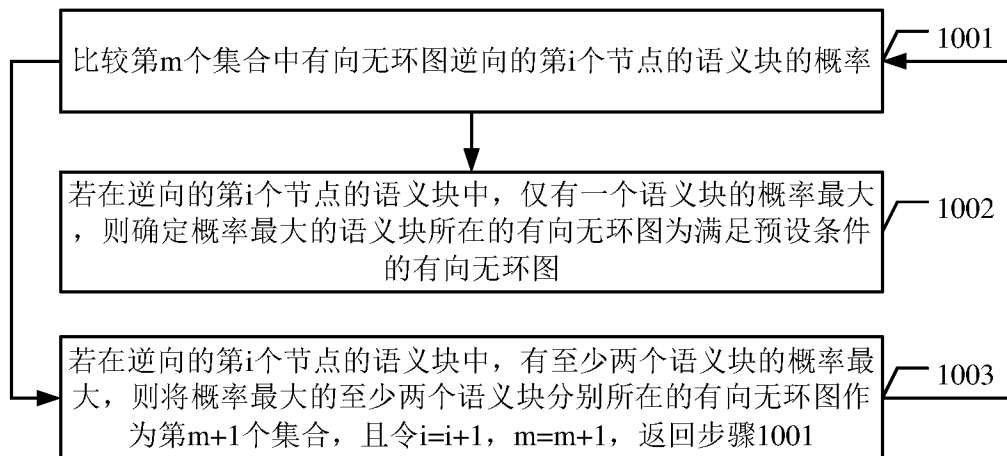


图 10

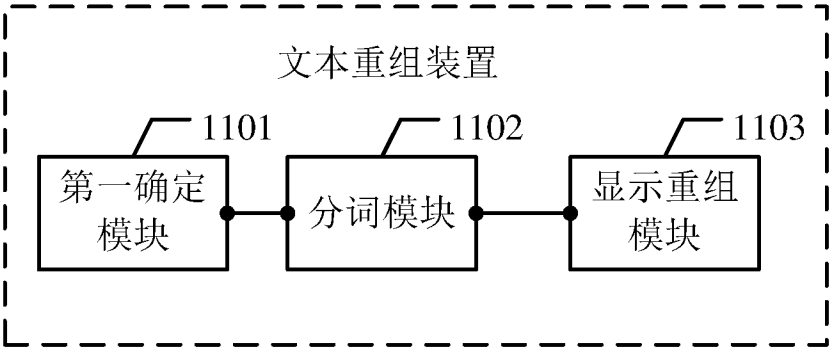


图 11

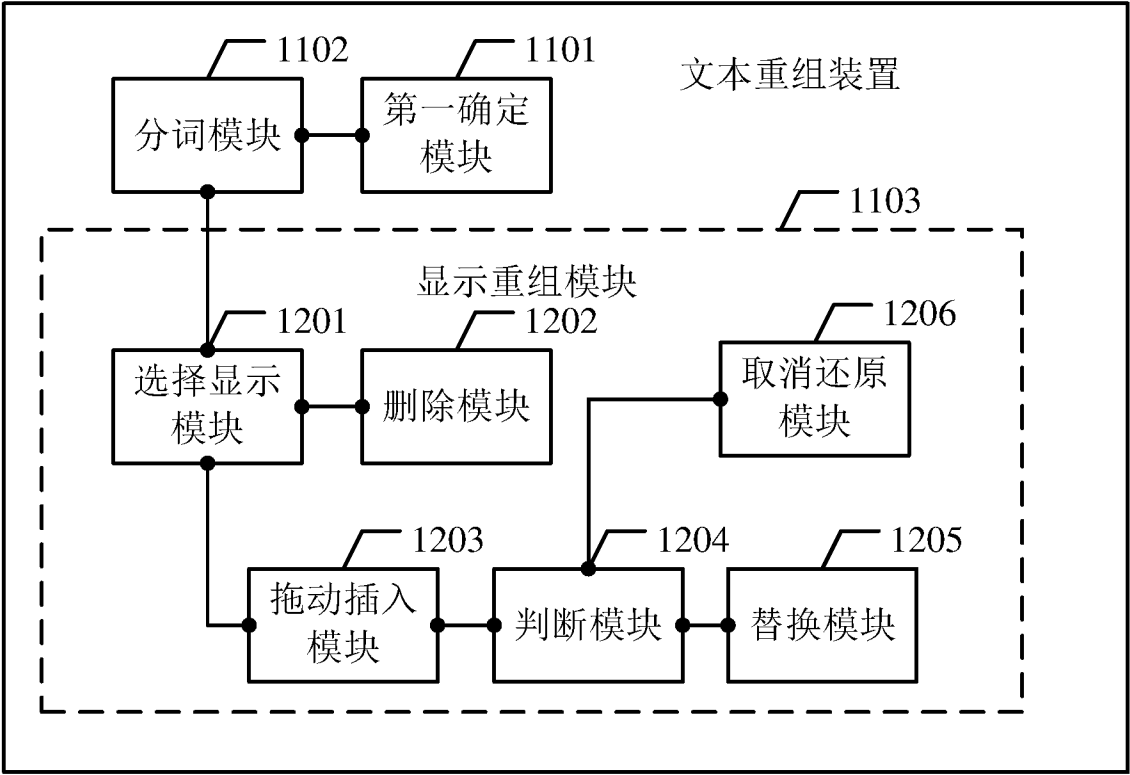


图 12

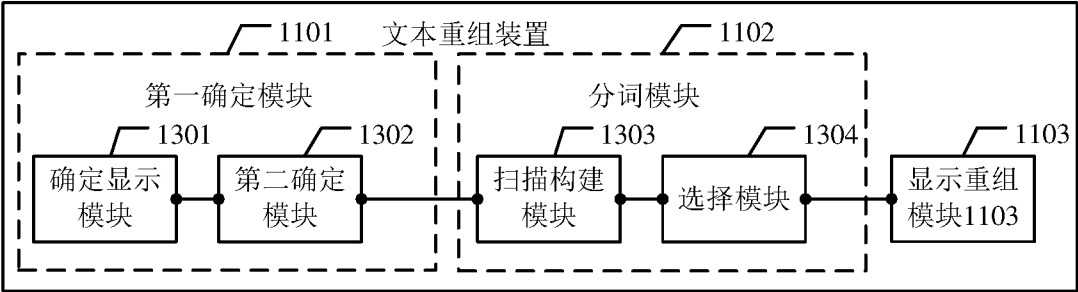


图 13

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2018/088789

A. CLASSIFICATION OF SUBJECT MATTER

G06F 3/0481(2013.01)i; G06F 3/0488(2013.01)n

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)

CNABS, CNKI, DWPI, CNTXT, CNABS: 文本, 重组, 分词, 有向无环图, text; reconstruction; word; directed acyclic graph

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 104298664 A (WANG, MEIJIN) 21 January 2015 (2015-01-21) description, paragraphs [0012]-[0034]	1-5, 11-15, 17, 18
Y	CN 104298664 A (WANG, MEIJIN) 21 January 2015 (2015-01-21) description, paragraphs [0012]-[0034]	6-10, 16-18
Y	CN 103473380 A (NANJING UNIVERSITY) 25 December 2013 (2013-12-25) description, paragraphs [0008]-[0061]	6-10, 16-18
A	CN 105468713 A (XI'AN JIAOTONG UNIVERSITY) 06 April 2016 (2016-04-06) entire document	1-18

☐ 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

09 August 2018

Date of mailing of the international search report

20 August 2018

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

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2018/088789

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	104298664	A	21 January 2015	None			
CN	103473380	A	25 December 2013	CN	103473380	B	13 April 2016
CN	105468713	A	06 April 2016	None			

国际检索报告

国际申请号

PCT/CN2018/088789

A. 主题的分类

G06F 3/0481(2013.01)i; G06F 3/0488(2013.01)n

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))

CNABS、CNKI、DWPI、CNTXT、CNABS: 文本、重组、分词、有向无环图、text; reconstruction; word; directed acyclic graph

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
X	CN 104298664 A (王美金) 2015年 1月 21日 (2015 - 01 - 21) 说明书[0012]-[0034]段	1-5、11-15、17、18
Y	CN 104298664 A (王美金) 2015年 1月 21日 (2015 - 01 - 21) 说明书[0012]-[0034]段	6-10、16-18
Y	CN 103473380 A (南京大学) 2013年 12月 25日 (2013 - 12 - 25) 说明书[0008]-[0061]段	6-10、16-18
A	CN 105468713 A (西安交通大学) 2016年 4月 6日 (2016 - 04 - 06) 全文	1-18

☐ 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

“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

2018年 8月 9日

国际检索报告邮寄日期

2018年 8月 20日

ISA/CN的名称和邮寄地址

中华人民共和国国家知识产权局(ISA/CN)
中国北京市海淀区蓟门桥西土城路6号 100088

受权官员

刘浩然

传真号 (86-10)62019451

电话号码 86-(10)-62411703

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2018/088789

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	104298664	A	2015年 1月 21日	无			
CN	103473380	A	2013年 12月 25日	CN	103473380	B	2016年 4月 13日
CN	105468713	A	2016年 4月 6日	无			