

(12) 发明专利申请

(10) 申请公布号 CN 102567300 A

(43) 申请公布日 2012. 07. 11

(21) 申请号 201110451081. 3

(22) 申请日 2011. 12. 29

(71) 申请人 方正国际软件有限公司

地址 215123 江苏省苏州市苏州工业园区星湖街 328 号创意产业园方正国际大厦

申请人 方正国际软件(北京)有限公司

(72) 发明人 胡希驰

(74) 专利代理机构 北京康信知识产权代理有限责任公司 11240

代理人 吴贵明 余刚

(51) Int. Cl.

G06F 17/24 (2006. 01)

G06T 11/60 (2006. 01)

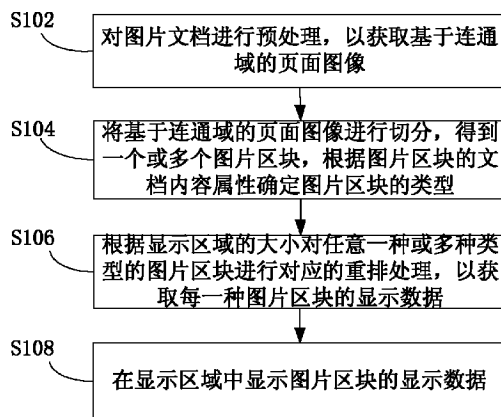
权利要求书 4 页 说明书 11 页 附图 15 页

(54) 发明名称

图片文档的处理方法及装置

(57) 摘要

本发明公开了一种图片文档的处理方法及装置。其中,该方法包括:对图片文档进行预处理,以获取基于连通域的页面图像;将基于连通域的页面图像进行切分,得到一个或多个图片区块,根据图片区块的文档内容属性确定图片区块的类型;根据显示区域的大小对任意一种或多种类型的图片区块进行对应的重排处理,以获取每一种图片区块的显示数据;在显示区域中显示图片区块的显示数据。通过本发明,能够实现直接在图片文档的图像层面上将版面重排,无需使用阅读工具,提高了阅读效率,避免了阅读工具转换过程中存在的转换错误,同时也降低了开发成本。



1. 一种图片文档的处理方法,其特征在于,包括:  
对图片文档进行预处理,以获取基于连通域的页面图像;  
将所述基于连通域的页面图像进行切分,得到一个或多个图片区块,根据所述图片区块的文档内容属性确定所述图片区块的类型;  
根据显示区域的大小对任意一种或多种类型的图片区块进行对应的重排处理,以获取每一种图片区块的显示数据;  
在所述显示区域中显示所述图片区块的显示数据。
2. 根据权利要求 1 所述的方法,其特征在于,所述图片区块的类型包括以下一种或多种类型:文字区块、图像区块、表格区块,其中,根据所述图片区块的文档内容属性确定所述图片区块的类型包括:  
检测所述图片区块的文档内容属性,其中,  
当检测到所述图片区块中各个合并连通域的矩形大小之差在预定范围之内时,确定所述图片区块为文字区块;  
当检测到所述图片区块中各个合并连通域的矩形大小之差大在预定范围之外时,确定所述图片区块为图像区块;  
当检测到所述图片区块中包括一条或多条表格线时,确定所述图片区块为表格区块。
3. 根据权利要求 2 所述的方法,其特征在于,在所述图片区块为文字区块的情况下,根据显示区域的大小对任意一种或多种类型的图片区块进行对应的重排处理,以获取每一种图片区块的显示数据的步骤包括:  
根据需求设置对应所述显示区域的字符显示特征,所述字符显示特征包括:字符大小、字符间距和字符行距;  
根据所述字符显示特征计算得到对应所述显示区域的字符行数和每行中的字符数目;  
依次读取所述文字区块中所有的字符,并按照所述显示区域的字符行数和每行中的字符数目将所述字符缩放后进行依次排序,得到所述文字区块对应所述显示区域的显示数据。
4. 根据权利要求 3 所述的方法,其特征在于,在依次读取所述文字区块中所有的字符之前,所述方法还包括:  
读取所述文字区块中的所有字符连通域;  
计算得到字符连通域的高度参考值,根据所述高度参考值来遍历所有字符连通域以对所述文字区块中的字块分行;  
根据字符的结构特征,对每行中的字符块进行单字切分和处理,以得到所述文字区块中所有的字符,其中,当所述字符为汉字的情况下,对每行中的字符块进行单字切分包括:将纵向坐标中上下关联的连通域合并为一个字符块,并将横向坐标中左右相邻距离小于等于预定值的连通域合并为一个字符块。
5. 根据权利要求 2 所述的方法,其特征在于,在所述图片区块为表格区块的情况下,根据显示区域的大小对任意一种或多种类型的图片区块进行对应的重排处理,以获取每一种图片区块的显示数据的步骤包括:  
提取所述表格区块中的表格线,并根据所述表格线对表格进行划分,得到一个或多个

具有行列坐标的单元格；

根据需求设置对应所述显示区域的单元格显示特征，所述单元格显示特征包括：单元格大小、单元格间距和单元格行距；

根据所述单元格显示特征计算得到对应所述显示区域的单元格行数和每行中的单元格数目；

依次读取所述表格区块中所有的单元格，并按照所述显示区域的单元格行数和每行中的单元格数目将所述单元格缩放后进行依次排序，得到所述表格区块对应所述显示区域的显示数据。

6. 根据权利要求 5 所述的方法，其特征在于，依次读取所述表格区块中所有的单元格，并按照所述显示区域的单元格行数和每行中的单元格数目将所述单元格缩放后进行依次排序，得到所述表格区块对应所述显示区域的显示数据包括：

提取所述表格区块中的所有表头单元格；

按照所述显示区域的单元格行数和每行中的单元格数目，确定每一个表头单元格在所述显示区域中的表头坐标位置；

将各个表头单元格缩放后复制到所述显示区域中已经确定的表头坐标位置；

读取所述表格区块中的字符单元格；

根据已经确定的表头坐标位置和所述显示区域的单元格行数和每行中的单元格数目，确定每一个字符单元格的字符坐标位置；

将各个表头单元格缩放后复制到所述显示区域中已经确定的字符坐标位置；

其中，在各个所述表头单元格的表头坐标位置确定之后，在每一个显示区域中的相同坐标位置复制相同的表头单元格。

7. 根据权利要求 2 所述的方法，其特征在于，在所述图片区块为图像区块的情况下，根据显示区域的大小对任意一种或多种类型的图片区块进行对应的重排处理，以获取每一种图片区块的显示数据的步骤包括：

根据需求设置对应所述显示区域的图像显示特征，所述图像显示特征包括：图像大小、图像间距和图像行距；

根据所述图像显示特征计算得到对应所述显示区域的图像行数和每行中的图像数目；

依次提取所述图像区块中的一个或多个子图像，并按照所述显示区域的图像行数和每行中的图像数目将所述子图像缩放后进行依次排序，得到所述图像区块对应所述显示区域的显示数据。

8. 根据权利要求 7 所述的方法，其特征在于，在提取所述图像区块中的一个或多个子图像之后，所述方法还包括：通过直方图均衡算法对各个子图像进行处理，以获取对比度超过预定值的图形。

9. 一种图片文档的处理装置，其特征在于，包括：

预处理模块，用于对图片文档进行预处理，以获取基于连通域的页面图像；

切分模块，用于将所述基于连通域的页面图像进行切分，得到一个或多个图片区块，根据所述图片区块的文档内容属性确定所述图片区块的类型；

重排模块，用于根据显示区域的大小对任意一种或多种类型的图片区块进行对应的重

排处理,以获取每一种图片区块的显示数据;

显示模块,用于在所述显示区域中显示所述图片区块的显示数据。

10. 根据权利要求 9 所述的装置,其特征在于,所述图片区块的类型包括以下一种或多种类型:文字区块、图像区块、表格区块,其中,所述切分模块包括:

检测模块,用于检测所述图片区块的文档内容属性;

第一获取模块,用于当检测到所述图片区块中各个合并连通域的矩形大小之差在预定范围之内时,确定所述图片区块为文字区块;

第二获取模块,用于当检测到所述图片区块中各个合并连通域的矩形大小之差大在预定范围之外时,确定所述图片区块为图像区块;

第二获取模块,用于当检测到所述图片区块中包括一条或多条表格线时,确定所述图片区块为表格区块。

11. 根据权利要求 10 所述的装置,其特征在于,在所述图片区块为文字区块的情况下,所述重排模块包括:

设置模块,用于根据需求设置对应所述显示区域的字符显示特征,所述字符显示特征包括:字符大小、字符间距和字符行距;

计算模块,用于根据所述字符显示特征计算得到对应所述显示区域的字符行数和每行中的字符数目;

排序模块,用于依次读取所述文字区块中所有的字符,并按照所述显示区域的字符行数和每行中的字符数目将所述字符缩放后进行依次排序,得到所述文字区块对应所述显示区域的显示数据。

12. 根据权利要求 10 所述的装置,其特征在于,在所述图片区块为表格区块的情况下,所述重排模块包括:

处理模块,用于提取所述表格区块中的表格线,并根据所述表格线对表格进行划分,得到一个或多个具有行列坐标的单元格;

设置模块,用于根据需求设置对应所述显示区域的单元格显示特征,所述单元格显示特征包括:单元格大小、单元格间距和单元格行距;

计算模块,用于根据所述单元格显示特征计算得到对应所述显示区域的单元格行数和每行中的单元格数目;

排序模块,用于依次读取所述表格区块中所有的单元格,并按照所述显示区域的单元格行数和每行中的单元格数目将所述单元格缩放后进行依次排序,得到所述表格区块对应所述显示区域的显示数据。

13. 根据权利要求 10 所述的装置,其特征在于,在所述图片区块为图像区块的情况下,所述重排模块包括:

设置模块,用于根据需求设置对应所述显示区域的图像显示特征,所述图像显示特征包括:图像大小、图像间距和图像行距;

计算模块,用于根据所述图像显示特征计算得到对应所述显示区域的图像行数和每行中的图像数目;

排序模块,用于依次提取所述图像区块中的一个或多个子图像,并按照所述显示区域的图像行数和每行中的图像数目将所述子图像缩放后进行依次排序,得到所述图像区块对

应所述显示区域的显示数据。

## 图片文档的处理方法及装置

### 技术领域

[0001] 本发明涉及图片处理领域,具体而言,涉及一种图片文档的处理方法及装置。

### 背景技术

[0002] 现有技术的用于支持版面重排的阅读工具主要针对版式文档,如 PDF, CEBX, EPUB 等。这类文件本身包含了基于内容的信息,如文字的编码、文字的位置、文字的字号字体、插图位置、图形的表达式等等。这些都为根据不同的分辨率重新排列显示形式提供了便利。但针对经过扫描后的图片格式文档,使用上述现有技术进行重排之前,需要先通过 OCR 识别等技术进行识别,而 OCR 识别技术本身还存在着错误率、兼容性等问题。而且对于漫画图像或者扫描版的 PDF 等,由于没有相关的页面与 OCR 信息,所以无法直接进行重排。为了解决该问题,可以采用借助版式文档的重排工具,但必须先将扫描图像后的图片格式文件转换成相应的版式文档,该方式需要大量的处理时间,而且转换后的内容在识别过程中会出现许多错误,影响重排结果,另外,由于阅读工具必须支持多种文件格式,增加了开发成本,不具有通用性。

[0003] 针对扫描后的图片文件,如 BMP, JPEG 格式文件,或者无版式信息的扫描版 PDF 文件,目前可以采用如下处理方式为用户提供阅读:通过对图片文件做切白边的处理,得到只显示图片中间的有效内容,可有效利用显示面积;或者按照阅读顺序切换显示焦点,如从上到下,从左到右,这种方式仅进行了局部显示,即对图片格式文件的局部内容放大后显示。上述方式存在如下问题:使用切白边对于大文档,如 A4,在显示屏幕小的设备(如手机)上,显示仍然很小,无法直接阅读。而使用焦点转移的方式阅读还是很不方便,不符合人的阅读习惯。

[0004] 目前针对相关技术的在阅读图片文档的过程中,现有的阅读工具存在阅读效率低、易出错,且开发成本高的问题,目前尚未提出有效的解决方案。

### 发明内容

[0005] 针对相关技术的在阅读图片文档的过程中,现有的阅读工具存在阅读效率低、易出错,且开发成本高的问题,目前尚未提出有效的问题而提出本发明,为此,本发明的主要目的在于提供一种图片文档的处理方法及装置,以解决上述问题。

[0006] 为了实现上述目的,根据本发明的一个方面,提供了一种图片文档的处理方法,该方法包括:对图片文档进行预处理,以获取基于连通域的页面图像;将基于连通域的页面图像进行切分,得到一个或多个图片区块,根据图片区块的文档内容属性确定图片区块的类型;根据显示区域的大小对任意一种或多种类型的图片区块进行对应的重排处理,以获取每一种图片区块的显示数据;在显示区域中显示图片区块的显示数据。

[0007] 进一步地,图片区块的类型包括以下一种或多种类型:文字区块、图像区块、表格区块,其中,根据图片区块的文档内容属性确定图片区块的类型包括:检测图片区块的文档内容属性,其中,当检测到图片区块中各个合并连通域的矩形大小之差在预定范围之内时,

确定图片区块为文字区块；当检测到图片区块中各个合并连通域的矩形大小之差大在预定范围之外时，确定图片区块为图像区块；当检测到图片区块中包括一条或多条表格线时，确定图片区块为表格区块。

[0008] 进一步地，在图片区块为文字区块的情况下，根据显示区域的大小对任意一种或多种类型的图片区块进行对应的重排处理，以获取每一种图片区块的显示数据的步骤包括：根据需求设置对应显示区域的字符显示特征，字符显示特征包括：字符大小、字符间距和字符行距；根据字符显示特征计算得到对应显示区域的字符行数和每行中的字符数目；依次读取文字区块中所有的字符，并按照显示区域的字符行数和每行中的字符数目将字符缩放后进行依次排序，得到文字区块对应显示区域的显示数据。

[0009] 进一步地，在依次读取文字区块中所有的字符之前，方法还包括：读取文字区块中的所有字符连通域；计算得到字符连通域的高度参考值，根据高度参考值来遍历所有字符连通域以对文字区块中的字块分行；根据字符的结构特征，对每行中的字符块进行单字切分和处理，以得到文字区块中所有的字符，其中，当字符为汉字的情况下，对每行中的字符块进行单字切分包括：将纵向坐标中上下关联的连通域合并为一个字符块，并将横向坐标中左右相邻距离小于等于预定值的连通域合并为一个字符块。

[0010] 进一步地，在图片区块为表格区块的情况下，根据显示区域的大小对任意一种或多种类型的图片区块进行对应的重排处理，以获取每一种图片区块的显示数据的步骤包括：提取表格区块中的表格线，并根据表格线对表格进行划分，得到一个或多个具有行列坐标的单元格；根据需求设置对应显示区域的单元格显示特征，单元格显示特征包括：单元格大小、单元格间距和单元格行距；根据单元格显示特征计算得到对应显示区域的单元格行数和每行中的单元格数目；依次读取表格区块中所有的单元格，并按照显示区域的单元格行数和每行中的单元格数目将单元格缩放后进行依次排序，得到表格区块对应显示区域的显示数据。

[0011] 进一步地，依次读取表格区块中所有的单元格，并按照显示区域的单元格行数和每行中的单元格数目将单元格缩放后进行依次排序，得到表格区块对应显示区域的显示数据包括：提取表格区块中的所有表头单元格；按照显示区域的单元格行数和每行中的单元格数目，确定每一个表头单元格在显示区域中的表头坐标位置；将各个表头单元格缩放后复制到显示区域中已经确定的表头坐标位置；读取表格区块中的字符单元格；根据已经确定的表头坐标位置和显示区域的单元格行数和每行中的单元格数目，确定每一个字符单元格的字符坐标位置；将各个表头单元格缩放后复制到显示区域中已经确定的字符坐标位置；其中，在各个表头单元格的表头坐标位置确定之后，在每一个显示区域中的相同坐标位置复制相同的表头单元格。

[0012] 进一步地，在图片区块为图像区块的情况下，根据显示区域的大小对任意一种或多种类型的图片区块进行对应的重排处理，以获取每一种图片区块的显示数据的步骤包括：根据需求设置对应显示区域的图像显示特征，图像显示特征包括：图像大小、图像间距和图像行距；根据图像显示特征计算得到对应显示区域的图像行数和每行中的图像数目；依次提取图像区块中的一个或多个子图像，并按照显示区域的图像行数和每行中的图像数目将子图像缩放后进行依次排序，得到图像区块对应显示区域的显示数据。

[0013] 进一步地，在提取图像区块中的一个或多个子图像之后，方法还包括：通过直方图

均衡算法对各个子图像进行处理,以获取对比度超过预定值的子图像。

[0014] 为了实现上述目的,根据本发明的另一方面,提供了一种图片文档的处理装置,该装置包括:预处理模块,用于对图片文档进行预处理,以获取基于连通域的页面图像;切分模块,用于将基于连通域的页面图像进行切分,得到一个或多个图片区块,根据图片区块的文档内容属性确定图片区块的类型;重排模块,用于根据显示区域的大小对任意一种或多种类型的图片区块进行对应的重排处理,以获取每一种图片区块的显示数据;显示模块,用于在显示区域中显示图片区块的显示数据。

[0015] 进一步地,图片区块的类型包括以下一种或多种类型:文字区块、图像区块、表格区块,其中,切分模块包括:检测模块,用于检测图片区块的文档内容属性;第一获取模块,用于当检测到图片区块中各个合并连通域的矩形大小之差在预定范围之内时,确定图片区块为文字区块;第二获取模块,用于当检测到图片区块中各个合并连通域的矩形大小之差在预定范围之外时,确定图片区块为图像区块;第三获取模块,用于当检测到图片区块中包括一条或多条表格线时,确定图片区块为表格区块。

[0016] 进一步地,在图片区块为文字区块的情况下,重排模块包括:设置模块,用于根据需求设置对应显示区域的字符显示特征,字符显示特征包括:字符大小、字符间距和字符行距;计算模块,用于根据字符显示特征计算得到对应显示区域的字符行数和每行中的字符数目;排序模块,用于依次读取文字区块中所有的字符,并按照显示区域的字符行数和每行中的字符数目将字符缩放后进行依次排序,得到文字区块对应显示区域的显示数据。

[0017] 进一步地,在图片区块为表格区块的情况下,重排模块包括:处理模块,用于提取表格区块中的表格线,并根据表格线对表格进行划分,得到一个或多个具有行列坐标的单元格;设置模块,用于根据需求设置对应显示区域的单元格显示特征,单元格显示特征包括:单元格大小、单元格间距和单元格行距;计算模块,用于根据单元格显示特征计算得到对应显示区域的单元格行数和每行中的单元格数目;排序模块,用于依次读取表格区块中所有的单元格,并按照显示区域的单元格行数和每行中的单元格数目将单元格缩放后进行依次排序,得到表格区块对应显示区域的显示数据。

[0018] 进一步地,在图片区块为图像区块的情况下,重排模块包括:设置模块,用于根据需求设置对应显示区域的图像显示特征,图像显示特征包括:图像大小、图像间距和图像行距;计算模块,用于根据图像显示特征计算得到对应显示区域的图像行数和每行中的图像数目;排序模块,用于依次提取图像区块中的一个或多个子图像,并按照显示区域的图像行数和每行中的图像数目将子图像缩放后进行依次排序,得到图像区块对应显示区域的显示数据。

[0019] 通过本发明,采用对图片文档进行预处理,以获取基于连通域的页面图像;将基于连通域的页面图像进行切分,得到一个或多个图片区块,根据图片区块的文档内容属性确定图片区块的类型;根据显示区域的大小对任意一种或多种类型的图片区块进行对应的重排处理,以获取每一种图片区块的显示数据;在显示区域中显示图片区块的显示数据,解决了相关现有技术的在阅读图片文档的过程中,现有的阅读工具存在阅读效率低、易出错,且开发成本高的问题,进而实现直接在图片文档的图像层面上将版面重排,无需使用阅读工具,提高了阅读效率,避免了阅读工具转换过程中存在的转换错误,同时也降低了开发成本的效果。



## 附图说明

[0020] 此处所说明的附图用来提供对本发明的进一步理解,构成本申请的一部分,本发明的示意性实施例及其说明用于解释本发明,并不构成对本发明的不当限定。在附图中:

- [0021] 图 1 是根据本发明实施例的图片文档的处理装置的结构示意图;
- [0022] 图 2a-2e 是根据图 1 所示实施例的对图片文档进行预处理的结果示意图;
- [0023] 图 3 是根据图 1 所示实施例的对图片文档进行区块分割的结果示意图;
- [0024] 图 4 是根据图 3 所示实施例的对文字区块进行字块分行处理的结果示意图;
- [0025] 图 5 是根据图 4 所示实施例的对文字区块进行单字切分处理的结果示意图;
- [0026] 图 6 是根据图 5 所示实施例的对文字区块进行重排处理的结果示意图;
- [0027] 图 7a-7c 是根据图 3 所示实施例的对表格区块进行重排处理的结果示意图;
- [0028] 图 8a-8b 是根据图 3 所示实施例的对图像区块进行重排处理的结果示意图;
- [0029] 图 9 是根据本发明实施例的图片文档的处理方法的流程图;
- [0030] 图 10 是根据图 9 所示实施例的图片文档的处理方法的详细流程图;
- [0031] 图 11a-11b 是根据图 9 所示实施例的图块的切分方法流程图;
- [0032] 图 12 是根据图 9 所示实施例的文字区块的处理方法流程图;
- [0033] 图 13 是根据图 9 所示实施例的表格区块的处理方法流程图;
- [0034] 图 14 是根据图 9 所示实施例的阅读顺序的分析流程图。

## 具体实施方式

[0035] 需要说明的是,在不冲突的情况下,本申请中的实施例及实施例中的特征可以相互组合。下面将参考附图并结合实施例来详细说明本发明。

[0036] 图 1 是根据本发明实施例的图片文档的处理装置的结构示意图;图 2a-2e 是根据图 1 所示实施例的对图片文档进行预处理的结果示意图;图 3 是根据图 1 所示实施例的对图片文档进行区块分割的结果示意图;图 4 是根据图 3 所示实施例的对文字区块进行字块分行处理的结果示意图;图 5 是根据图 4 所示实施例的对文字区块进行单字切分处理的结果示意图;图 6 是根据图 5 所示实施例的对文字区块进行重排处理的结果示意图;图 7a-7c 是根据图 3 所示实施例的对表格区块进行重排处理的结果示意图;图 8a-8b 是根据图 3 所示实施例的对图像区块进行重排处理的结果示意图。

[0037] 如图 1 所示,该图片文档的处理装置包括:预处理模块 10,用于对图片文档进行预处理,以获取基于连通域的页面图像;切分模块 30,用于将基于连通域的页面图像进行切分,得到一个或多个图片区块,根据图片区块的文档内容属性确定图片区块的类型;重排模块 50,用于根据显示区域的大小对任意一种或多种类型的图片区块进行对应的重排处理,以获取每一种图片区块的显示数据;显示模块 70,用于在显示区域中显示图片区块的显示数据。

[0038] 本申请上述实施例通过对进行预处理后的图片文档进行切分,并将各种切分后图像字块缩放后按新的显示要求贴图到显示区域的指定位置上。由于该实施例中直接利用图像处理技术对图片文档进行了预处理和分析,因此无需采用 OCR 技术进行阅读,提高了阅读效率,避免了使用阅读工具转换图片文件过程中存在的转换错误,同时也降低了开发成

本的效果。

[0039] 该技术尤其适合目前的手持设备,如智能手机、电子书、平板电脑。当前的这些设备中,使得对于扫描版的图片文档(例如 BMP 图片、JPEG 图片、扫描版 PDF 或漫画)的处理不仅仅是切除白边和按注意区域转移显示,可以进一步满足用户的阅读需求,提供更好的用户体验。

[0040] 具体的,如图 2a-2e 所示,上述实施利中对如图 2a 所示的图片文档(原始灰度图)进行预处理,可以实现根据图像质量和类型进行包括以下的一种或几种处理:降噪,灰度校正,几何校正,倾斜矫正,去黑边,二值化,连通域生成与合并等。例如,首先对图 2a 进行二值化处理得到图 2b,可以采用阈值分割算法 OTSU 将原始灰度图像转换成二值图像;然后,在如图 2b 所示的二值图像的基础上进行连通域分析得到图 2c,例如采用查找代表文字的黑像素的方式得到初始连通域,可以通过以一个黑像素点为开始,搜寻其周围 8 邻域的像素点,如果邻域上的像素点也为黑像素点则认为它们是一个连通域内的像素点,接着依次再计算邻域上黑像素点的邻域,最终找出一片相连的黑像素点区域,这个就是一个连通域。查找图像中其他没有计算过的位置,重复上述步骤,即可找出所有的连通域。对于每个连通域,其中每个像素点的  $x, y$  坐标,在一个连通域中所有的像素点计算最小和最大的  $x, y$ ,即可得到该连通域的上下左右的边界坐标,即计算出了最小的外接矩形四个顶点,坐标分别为  $(x_{min}, y_{min})$ 、 $(x_{min}, y_{max})$ 、 $(x_{max}, y_{min})$ 、 $(x_{max}, y_{max})$ ;在得到图片文档的初始连通域图 2c 之后,对图 2c 进行连通域合并得到图 2d 和 2e,例如,例如图 2e 中,由于汉字字符的笔画及偏旁,需要对初始连通域中的包含和相交的矩形进行合并,以提高后续的处理准确率。

[0041] 本申请上述实施例中的图片区块的类型可以包括以下一种或多种类型:文字区块、图像区块、表格区块,其中,切分模块 30 包括:检测模块,用于检测图片区块的文档内容属性;第一获取模块,用于当检测到图片区块中各个合并连通域的矩形大小之差在预定范围之内时,确定图片区块为文字区块;第二获取模块,用于当检测到图片区块中各个合并连通域的矩形大小之差大在预定范围之外时,确定图片区块为图像区块;第三获取模块,用于当检测到图片区块中包括一条或多条表格线时,确定图片区块为表格区块。该实施例提供了将整个图片文档中不同属性的区块区分开,以便于使用不同的方式进行重排处理。

[0042] 上述实施例中针对区块的切分模块 30 具体可以实现,将图片文档版面中的元素按内容的属性分成各类区块。具体的,可以利用空白间隙查找的方法将连通域分割成许多大块;或者直接计算图像中各个像素点的邻域特征,利用不同的特征数值将版面分割成一些区块。例如,如果确定在图片文档中分隔出多幅图像漫画,则可以利用子图间的间隙以及子图内的连通域,将整图切分成几个小图。

[0043] 具体的如图 3 所示,在以连通域为基础的图 2e 中,可以利用自底向上的合并算法或自上向下的空白区域分隔算法将文档图像分成很多区块。在分割成很多区块后,可以按照区块中的属性特征判断区块的具体类型,以便后续进一步处理,例如,需要判断每个区块是文字还是插图。可以利用图像的属性,比如文字区块中连通域的矩形大小一般比较均匀;而插图中可能大小不一;表格中会有各种相交的表格线。在切分得到多个区块之后,区块类型包括:文字区块、插图图像区块、插图图形区块(线条图)、表格区块、公式区块等等。所能利用的文档内容属性的特征包括但不限于一下特征:连通域的尺寸、连通域的空间分布周期性、尺寸的不均衡性、黑像素密度、黑游程长度及其统计特征、灰度分布特征、游程统计

特征、频域特征、直方图分布特征、梯度分布特征、分型特征、各种纹理特征等；而判断方法可以采用根据各种特征设定阈值，然后决策树判断，也可使用样本集训练的方式，如神经网络、支撑向量机等。具体的，可以对各种文档内容属性的特征设定阈值，由决策树进行判断，例如采用连通域的长宽的统计分布作为特征，则文字区域长宽较为均一，即方差较小；图像区域的连通域长宽的方差较小。根据阈值的大小即可区分；也可使用样本集训练的方式，如神经网络、支撑向量机等。

[0044] 本申请上述实施例中，在图片区块为文字区块的情况下，重排模块 50 可以包括：设置模块 501，用于根据需求设置对应显示区域的字符显示特征，字符显示特征包括：字符大小、字符间距和字符行距；计算模块 502，用于根据字符显示特征计算得到对应显示区域的字符行数和每行中的字符数目；排序模块 503，用于依次读取文字区块中所有的字符，并按照显示区域的字符行数和每行中的字符数目将字符缩放后进行依次排序，得到文字区块对应显示区域的显示数据。

[0045] 上述实施利通过对文字区块的处理为文字区块的重排操作做预备，具体的，可以对文字区块中的字符进行如下处理：成行（列），单字切分，字符分类（标点不能出现在行头，英文单词、拼音、数字不能在行尾出现断续），公式区域判断（当成图像直接抠图），文字属性分析（大小、粗细（参考 dpi））。在得到所有字符进行处理之后，可以根据设定的字体大小、字间距（可计算并保留原始值）、行间距（可计算并保留原始值）、原始 dpi 和目标显示分辨率，计算单字块、大区块的映射位置，同时在对每个字符进行缩放之后，将各个字符块拷贝到目标显示区域。

[0046] 具体的，首先，需要根据目标屏幕的大小，通过用户设定的在目标显示区域中的期望字符大小、字间距、行距，计算每个屏幕上显示区域的字行数和每行中的字数，并将字符的矩形区域图像贴到目标区域上的相应位置即可。

[0047] 在对文字区块的处理过程中还需要考虑字符类型和排版习惯，如标点不能出现在行头，英文单词、拼音、数字不能在行尾出现断续。具体的，可以判断每个字符的属性是否是标点，在，版面重排的时候，由于阅读习惯中，标点是不能放在一行的最前的，正常对于一行的宽度和要放置的字符宽度、间隔，需要计算出这一行能放多少个字符。如果检测到下一行的开始是一个标点，那么在在上一行可以细微调整字间距，将标点放置在这一行末。

[0048] 优选地，在依次读取文字区块中所有的字符之前，可以读取文字区块中的所有字符连通域；计算得到字符连通域的高度参考值，根据高度参考值来遍历所有字符连通域以对文字区块中的字块分行；根据字符的结构特征，对每行中的字符块进行单字切分和处理，以得到文字区块中所有的字符，其中，当字符为汉字的情况下，对每行中的字符块进行单字切分包括：将纵向坐标中上下关联的连通域合并为一个字符块，并将横向坐标中左右相邻距离小于等于预定值的连通域合并为一个字符块。同时，可以对合并后的字块进行判断，只有在合并后的字符宽高满足预定范围时，则对连通域进行合并。

[0049] 具体的，如图 4 所示，上述实施利具体实现方式如下：

[0050] 首先对文字区块中的字符进行字块分行处理，在字块的处理中，将字符连通域进行成行处理，有助于字块分析、单字切分。这也是版面分析中的一个常用步骤；另外，还可使用如下方式：首先统计字块中的所有连通域的高度，计算概率最大的高度值，以此作为行高的参考值。通过上述处理方式遍历所有的连通域，如果该连通域不属于任何行，则新建一个

行,以当前连通域外接矩形的中心上下半个行高做两条水平线(横排版),凡是中心点位于这两条线中间的连通域均属于这个新行,直至处理所有的连通域。

[0051] 然后,在字块分行处理完成后,如图 5 所示,对文字区块字块进行分行处理使得页面成行之后,由于汉字有上下结构,对字块进行单字切分处理,即合并行内上下关系的连通域为一个字符。同时汉字是方块字,挑选出不接近正方形的外接矩形,如果这些连通域有左右很近的,合并后的字符宽高是否符合大多数字符的宽高特征,如果符合则合并,如果不符合则保持分离。

[0052] 最后,以图 5 中所示的文字区块为例,在目标显示区域为每个字长宽均为 50 个像素,屏幕宽 500 像素,高 600 像素,字间距 10,行间距 20,如图 6 所示,每页只能排布 8 行,每行 8 个字符。因为  $50*8+9*10 = 490 < 500$ ,  $50*8+9*20 = 580 < 600$ 。图 6 为第一页显示区域,图 5 中的文字以上述方式依次以图 6 所示的布局进行显示。

[0053] 本申请上述实施例中,在图片区块为表格区块的情况下,重排模块 50 根据显示区域的大小对任意一种或多种类型的图片区块进行对应的重排处理,以获取每一种图片区块的显示数据的步骤包括:处理模块,用于提取表格区块中的表格线,并根据表格线对表格进行划分,得到一个或多个具有行列坐标的单元格;设置模块 501,用于根据需求设置对应显示区域的单元格显示特征,单元格显示特征包括:单元格大小、单元格间距和单元格的行距;计算模块 502,用于根据单元格显示特征计算得到对应显示区域的单元格行数和每行中的单元格数目;排序模块 503,用于依次读取表格区块中所有的单元格,并按照显示区域的单元格行数和每行中的单元格数目将单元格缩放后进行依次排序,得到表格区块对应显示区域的显示数据。表格区块处理模块

[0054] 上述实施例通过可将整个表格区块当做图像显示,首先通过提取表格区块中的表格线来将表格区块切分成多个单元格,然后对单元格进行排列分析,同时提取字符块,通过计算行、列数确定每个单元格在显示页面中的具体位置以及缩放的大小。在经过上述针对单元格的分析之后,可实现设置按多行显示或多列显示,或定位行列区域显示。

[0055] 具体的,如图 7a-7c 所示,利用表格线,以及文字成行的方法,可以将图 7a 所示的表格分成具有行列坐标的单元格。同上文字区块中的文字排布方式,可以根据目标屏幕大小及单元格大小,将每个单元格缩放后贴到显示区域的相应位置。为了方便阅读,可以在每页均复制贴上表头(及第一行)信息。

[0056] 优选地,上述实施例中,依次读取表格区块中所有的单元格,并按照显示区域的单元格行数和每行中的单元格数目将单元格缩放后进行依次排序,得到表格区块对应显示区域的显示数据的步骤可以包括:提取表格区块中的所有表头单元格;按照显示区域的单元格行数和每行中的单元格数目,确定每一个表头单元格在显示区域中的表头坐标位置;将各个表头单元格缩放后复制到显示区域中已经确定的表头坐标位置;读取表格区块中的字符单元格;根据已经确定的表头坐标位置和显示区域的单元格行数和每行中的单元格数目,确定每一个字符单元格的字符坐标位置;将各个表头单元格缩放后复制到显示区域中已经确定的字符坐标位置;其中,在各个表头单元格的表头坐标位置确定之后,在每一个显示区域中的相同坐标位置复制相同的表头单元格。

[0057] 本申请上述实施例中,在图片区块为图像区块的情况下,重排模块 50 包括:设置模块 501,用于根据需求设置对应显示区域的图像显示特征,图像显示特征包括:图像大

小、图像间距和图像行距；计算模块 502，用于根据图像显示特征计算得到对应显示区域的图像行数和每行中的图像数目；排序模块 503，用于依次提取图像区块中的一个或多个子图像，并按照显示区域的图像行数和每行中的图像数目将子图像缩放后进行依次排序，得到图像区块对应显示区域的显示数据。本申请上述实施例通过对图像区块进行处理，例如进行灰度调整，从而增强对比度或亮度；以及对图像区块进行二值化处理，使得显示更清晰，并将处理后的图像按照目标显示区域的大小进行放缩显示。

[0058] 具体的，如图 8a-8b 所示，将图 8a 所示的图像区块进行直方图均衡处理得到图 8b。例如，对于对比度不高的图像可以进行对比度增强，这里使用图像处理算法中常用的直方图均衡。对于文字区块，可以使用灰度图，也可以使用二值图。如果是二值图，则不需调整。该处理改善了视觉效果，提高了用户体验。

[0059] 由上对各个区块的版面重排操作，使得各类区块在目标显示区域得到预定的显示效果。在版面重排后，可实现如下的调整：设置按多行显示或多列显示，或定位行列区域显示；对于漫画文档可按照设定顺序显示，如从上到下从左到右；可通过缩放各个单字块或大的图像、表格区块，以及调整文字笔画粗细或浓淡程度重排效果进行调整；通过对字体的二值化分割和区域标定，利用填充算法，调整字符与背景的颜色。

[0060] 本申请上述实施例实现了在不利用 OCR 技术的情况下，对图片文档的页面图像进行切分。判断页面中区块的属性。如果是图像，可直接将区域抠出，显示时使用缩放技术；如果是文字块，进行行切分和字切分，在重排时按字块图像，回帖到合适位置。且利用基本的排版特征，如缩进、分栏等，可以获得段落和阅读顺序；如果是表格，利用线段检测和单元格分析，可以按列或按行或按块重新组织显示，也可将整个表格块作为插图处理。对于多格漫画，可利用其边框和插图联通情况，将原本一页的分多页显示。该技术尤其适合目前的手持设备，如智能手机、电子书、平板电脑。。

[0061] 图 9 是根据本发明实施例的图片文档的处理方法的流程图；图 10 是根据图 9 所示实施例的图片文档的处理方法的详细流程图；图 11a-11b 是根据图 9 所示实施例的图块的切分方法流程图；图 12 是根据图 9 所示实施例的文字区块的处理方法流程图；图 13 是根据图 9 所示实施例的表格区块的处理方法流程图；图 14 是根据图 9 所示实施例的阅读顺序的分析流程图。

[0062] 如图 9 所示该方法包括如下步骤：

[0063] 步骤 S102，通过图 1 中的预处理模块 10 对图片文档进行预处理，以获取基于连通域的页面图像。

[0064] 步骤 S104，通过图 1 中的切分模块 30 执行将基于连通域的页面图像进行切分，得到一个或多个图片区块，根据图片区块的文档内容属性确定图片区块的类型。

[0065] 步骤 S106，通过图 1 中的重排模块 50 来实现根据显示区域的大小对任意一种或多种类型的图片区块进行对应的重排处理，以获取每一种图片区块的显示数据。

[0066] 步骤 S108，通过图 1 中的显示模块 70 在显示区域中显示图片区块的显示数据。

[0067] 本申请上述实施例通过对进行预处理后的图片文档进行切分，并将各种切分后图像字块缩放后按新的显示要求贴图到显示区域的指定位置上。由于该实施例中直接利用图像处理技术对图片文档进行了预处理和分析，因此无需采用 OCR 技术进行阅读，提高了阅读效率，避免了使用阅读工具转换图片文件过程中存在的转换错误，同时也降低了开发成

本的效果。

[0068] 本申请上述实施例中,图片区块的类型包括以下一种或多种类型:文字区块、图像区块、表格区块,其中,根据图片区块的文档内容属性确定图片区块的类型可以包括:检测图片区块的文档内容属性,其中,当检测到图片区块中各个合并连通域的矩形大小之差在预定范围之内时,确定图片区块为文字区块;当检测到图片区块中各个合并连通域的矩形大小之差大在预定范围之外时,确定图片区块为图像区块;当检测到图片区块中包括一条或多条表格线时,确定图片区块为表格区块。该实施例提供了将整个图片文档中不同属性的区块区分开,以便于使用不同的方式进行重排处理。

[0069] 上述实施例中针对区块的切分模块 30 具体可以实现,将图片文档版面中的元素按内容的属性分成各类区块。具体的,如图 11a 和 11b 所示,可以利用空白间隙查找的方法将连通域分割成许多大块;或者直接计算图像中各个像素点的邻域特征,利用不同的特征数值将版面分割成一些区块。如图例如,如果确定在图片文档中分隔出多幅图像漫画,则可以利用子图间的间隙以及子图内的连通域,将整图切分成几个小图。

[0070] 而且,如图 10 所示,在切分得到多个区块之后,可以通过区块属性进行判断,可以按照区块中的特征判断区块的具体类型,以便后续进一步处理。区块类型包括:文字区块、插图图像区块、插图图形区块(线条图)、表格区块、公式区块等等。所能利用的文档内容属性的特征包括但不限于一下特征:连通域的尺寸、连通域的空间分布周期性、尺寸的不均衡性、黑像素密度、游程统计特征、频域特征、直方图分布特征、梯度分布特征、分型特征、各种纹理特征等;而判断方法可以采用根据各种特征设定阈值,然后决策树判断,也可使用样本集训练的方式,如神经网络、支撑向量机等。具体的,在对每种区块中的内容基于目标显示区域的标准进行处理后,可以进行阅读顺序的分析,并在显示区域进行对应的重排并根据用户体验进行效果调整。

[0071] 本申请上述实施例中,在图片区块为文字区块的情况下,根据显示区域的大小对任意一种或多种类型的图片区块进行对应的重排处理,以获取每一种图片区块的显示数据的步骤包括:根据需求设置对应显示区域的字符显示特征,字符显示特征包括:字符大小、字符间距和字符行距;根据字符显示特征计算得到对应显示区域的字符行数和每行中的字符数目;依次读取文字区块中所有的字符,并按照显示区域的字符行数和每行中的字符数目将字符缩放后进行依次排序,得到文字区块对应显示区域的显示数据。该实施例中,在执行重排操作之前,需要根据目标屏幕的大小,通过用户设定的在目标显示区域中的期望字符大小、字间距、行距,计算每个屏幕上显示区域的字行数和每行中的字数,并将字符的矩形区域图像贴到目标区域上的相应位置即可。

[0072] 具体的,上述实施利通过对文字区块的处理为文字区块的重排操作做预备,具体的,可以对对文字区块中的字符进行如下处理:成行(列),单字切分,字符分类(标点不能出现在行头,英文单词、拼音、数字不能在行尾出现断续),公式区域判断(当成图像直接抠图),文字属性分析(大小、粗细(参考 dpi))。在得到所有字符进行处理之后,可以根据设定的字体大小、字间距(可计算并保留原始值)、行间距(可计算并保留原始值)、原始 dpi 和目标显示分辨率,计算单字块、大区块的映射位置,同时在对每个字符进行缩放之后,将各个字符块拷贝到目标显示区域。考虑字符类型和排版习惯,如标点不能出现在行头,英文单词、拼音、数字不能在行尾出现断续。

[0073] 优选地,在依次读取文字区块中所有的字符之前,方法还可以包括:读取文字区块中的所有字符连通域;计算得到字符连通域的高度参考值,根据高度参考值来遍历所有字符连通域以对文字区块中的字块分行;根据字符的结构特征,对每行中的字符块进行单字切分和处理,以得到文字区块中所有的字符,其中,当字符为汉字的情况下,对每行中的字符块进行单字切分包括:将纵向坐标中上下关联的连通域合并为一个字符块,并将横向坐标中左右相邻距离小于等于预定值的连通域合并为一个字符块。上述实施例如图 12 所示,在对文字区块中的各个字符进行一系列处理之后得到字符块,便于后续字符重排的操作。

[0074] 由上分析可知,本申请中对于文字区块的处理首先对文字区块中的字符进行字块分行处理,在遍历所有的连通域得到分行处理后的文字区块;然后,在字块分行处理完成后,对文字区块字块进行分行处理使得页面成行之后,由于汉字有上下结构,对字块进行单字切分处理;最后,以图 5 中所示的文字区块为例,在目标显示区域为每个字长宽均为 50 个像素,屏幕宽 500 像素,高 600 像素,字间距 10,行间距 20,如图 6 所示,每页只能排布 8 行,每行 8 个字符。因为  $50*8+9*10 = 490 < 500$ ,  $50*8+9*20 = 580 < 600$ 。图 6 为第一页显示区域,图 5 中的文字以上述方式依次以图 6 所示的布局进行显示。

[0075] 本申请上述实施例中,在图片区块为表格区块的情况下,根据显示区域的大小对任意一种或多种类型的图片区块进行对应的重排处理,以获取每一种图片区块的显示数据的步骤可以包括:提取表格区块中的表格线,并根据表格线对表格进行划分,得到一个或多个具有行列坐标的单元格;根据需求设置对应显示区域的单元格显示特征,单元格显示特征包括:单元格大小、单元格间距和单元格行距;根据单元格显示特征计算得到对应显示区域的单元格行数和每行中的单元格数目;依次读取表格区块中所有的单元格,并按照显示区域的单元格行数和每行中的单元格数目将单元格缩放后进行依次排序,得到表格区块对应显示区域的显示数据。

[0076] 上述实施例通过可将整个表格区块当做图像显示,具体的,如图 13 所示,首先通过提取表格区块中的表格线来将表格区块切分成多个单元格,然后对单元格进行排列分析,同时提取字符块,通过计算行、列数确定每个单元格在显示页面中的具体位置以及缩放的大小。在经过上述针对单元格的分析之后,可实现设置按多行显示或多列显示,或定位行列区域显示。如果是漫画文档,按照设定顺序显示,如从上到下从左到右。

[0077] 优选地,依次读取表格区块中所有的单元格,并按照显示区域的单元格行数和每行中的单元格数目将单元格缩放后进行依次排序,得到表格区块对应显示区域的显示数据的步骤可以包括:提取表格区块中的所有表头单元格;按照显示区域的单元格行数和每行中的单元格数目,确定每一个表头单元格在显示区域中的表头坐标位置;将各个表头单元格缩放后复制到显示区域中已经确定的表头坐标位置;读取表格区块中的字符单元格;根据已经确定的表头坐标位置和显示区域的单元格行数和每行中的单元格数目,确定每一个字符单元格的字符坐标位置;将各个表头单元格缩放后复制到显示区域中已经确定的字符坐标位置;其中,在各个表头单元格的表头坐标位置确定之后,在每一个显示区域中的相同坐标位置复制相同的表头单元格。

[0078] 本申请上述实施例中,在图片区块为图像区块的情况下,根据显示区域的大小对任意一种或多种类型的图片区块进行对应的重排处理,以获取每一种图片区块的显示数据的步骤可以包括:根据需求设置对应显示区域的图像显示特征,图像显示特征包括:图像

大小、图像间距和图像行距；根据图像显示特征计算得到对应显示区域的图像行数和每行中的图像数目；依次提取图像区块中的一个或多个子图像，并按照显示区域的图像行数和每行中的图像数目将子图像缩放后进行依次排序，得到图像区块对应显示区域的显示数据。优选地，在提取图像区块中的一个或多个子图像之后，方法还包括：通过直方图均衡算法对各个子图像进行处理，以获取对比度超过预定值的子图像。本申请上述实施例通过对图像区块进行处理，例如进行灰度调整，从而增强对比度或亮度；以及对图像区块进行二值化处理，使得显示更清晰。并将处理后的图像按照目标显示区域的大小进行防缩显示。

[0079] 本申请上述实施例实现了在不利用 OCR 技术的情况下，对图片文档的页面图像进行切分。判断页面中区块的属性。如果是图像，可直接将区域抠出，显示时使用缩放技术；如果是文字块，进行行切分和字切分，在重排时按字块图像，回帖到合适位置。且利用基本的排版特征，如缩进、分栏等，可以获得段落和阅读顺序；如果是表格，利用线段检测和单元格分析，可按列或按行或按块重新组织显示，也可将整个表格块作为插图处理。对于多格漫画，可利用其边框和插图联通情况，将原本一页的分多页显示。该技术尤其适合目前的手持设备，如智能手机、电子书、平板电脑。

[0080] 需要说明的是，在附图的流程图示出的步骤可以在诸如一组计算机可执行指令的计算机系统中执行，并且，虽然在流程图中示出了逻辑顺序，但是在某些情况下，可以以不同于此处的顺序执行所示出或描述的步骤。

[0081] 本申请上述实施例为了优化用户的阅读习惯，如图 14 所示，在重排过程中还可以采用阅读顺序分析模块对排版类型自动分析（或手工输入），利用版面基础先验知识（段落缩进，段后空白，标题、章节位置，分栏情况，）判断阅读顺序为重排提供依据。同时，也可以采用显示效果调整模块缩放各个单字块或大的图像、表格区块。调整文字笔画粗细或浓淡程度以达到最佳阅读效果。另外，通过对字体的二值化分割和区域标定，利用填充算法，还可实现设置字符与背景颜色的功能。手工输入即指在操作界面上提供一个设置工具，比如采用鼠标点击单选框，选中要处理的页面是“横排版”还是“竖排版”。自动处理就是指算法自动根据文字行、列方向排布方式、间隔、周期等计算出是“横排版”还是“竖排版”。

[0082] 从以上的描述中，可以看出，本发明实现了如下技术效果：直接利用图像处理技术进行分析，无需 OCR 技术进行预先识别，将各种切分后图像字块缩放后按新的显示要求贴图到指定位置。该技术尤其适合目前的手持设备，如智能手机、电子书、平板电脑。利用上述技术的各种设备，针对扫描版的 PDF 或漫画的处理不仅仅处理是切除白边和按注意区域转移显示，满足了用户更多的阅读需求。

[0083] 显然，本领域的技术人员应该明白，上述的本发明的各模块或各步骤可以用通用的计算装置来实现，它们可以集中在单个的计算装置上，或者分布在多个计算装置所组成的网络上，可选地，它们可以用计算装置可执行的程序代码来实现，从而，可以将它们存储在存储装置中由计算装置来执行，或者将它们分别制作成各个集成电路模块，或者将它们中的多个模块或步骤制作成单个集成电路模块来实现。这样，本发明不局限于任何特定的硬件和软件结合。以上所述仅为本发明的优选实施例而已，并不用于限制本发明，对于本领域的技术人员来说，本发明可以有各种更改和变化。凡在本发明的精神和原则之内，所作的任何修改、等同替换、改进等，均应包含在本发明的保护范围之内。



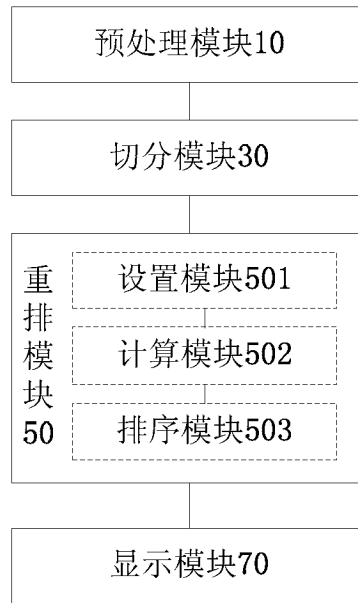


图 1



图 2a

海因里希·希特勒, 在一名党卫队军官的帮助下, 从一辆“虎”式坦克的炮塔探出身来, 神情愉悦。



### “令人畏惧的名望”带来的新光彩

9月15日, 哈尔科夫战役结束, 德国的反击遏制了苏联的冬季攻势, 并使战线稳定下来。霍塞尔率领的勇猛的党卫队装甲军团获得了“雄狮”勋章, 但他的成功受到上峰某种程度上的贬抑: 元首对霍塞尔自由机杆的行动耿耿于怀。在延宕了几周之后, 他才把霍塞尔纳入到因在哈尔科夫战役中的英勇战功而受勋者之中。但是, 无论如何, 党卫队第二装甲军团完成了希特勒对他们寄予的期望。

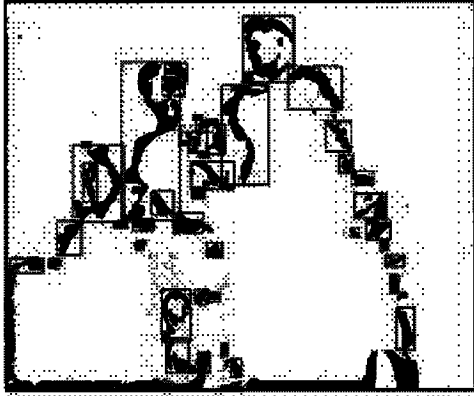
希特勒下令他钟爱的这支部队整装待发, 开赴库尔斯克附近, 投入到大规模的反击攻势当中, 这场攻势后来成为二次大战中规模最大的坦克集群战斗。为了激励士气, 党卫队帝国领袖海因里希·希姆莱受命来到哈尔科夫。他对他的部队发表了洋溢着雄心壮志的雄辩演说: “我们永远不会让我们卓越的部队, 我们在哈尔科夫战役中赢得的, 指引着我们的、令人敬畏的荣誉受损, 相反, 我们要一往无前并把它发扬光大。”



18 · 苏维埃战车

图 2b

**“奇人”的“奇事”**



**“奇人”的“奇事”**

4月8日，新华社发表了一篇题为“奇人”的奇事”的文章，引起了社会的广泛注意。文章中说，在北京，有一位奇人，他的名字叫做“奇人”，他有一项惊人的本领，就是能用他的“奇术”来治病。他治病的方法很简单，就是把病人的名字写在纸上，然后放在他的“奇术”下，病人就能马上好转。他治病的方法在国内外已经流传开来，而且他的“奇术”在国内外已经流传开来，而且他的“奇术”在国内外已经流传开来。

据新华社北京4月8日电：北京，有一位奇人，他的名字叫做“奇人”，他有一项惊人的本领，就是能用他的“奇术”来治病。他治病的方法很简单，就是把病人的名字写在纸上，然后放在他的“奇术”下，病人就能马上好转。他治病的方法在国内外已经流传开来，而且他的“奇术”在国内外已经流传开来，而且他的“奇术”在国内外已经流传开来。

新华社北京4月8日电

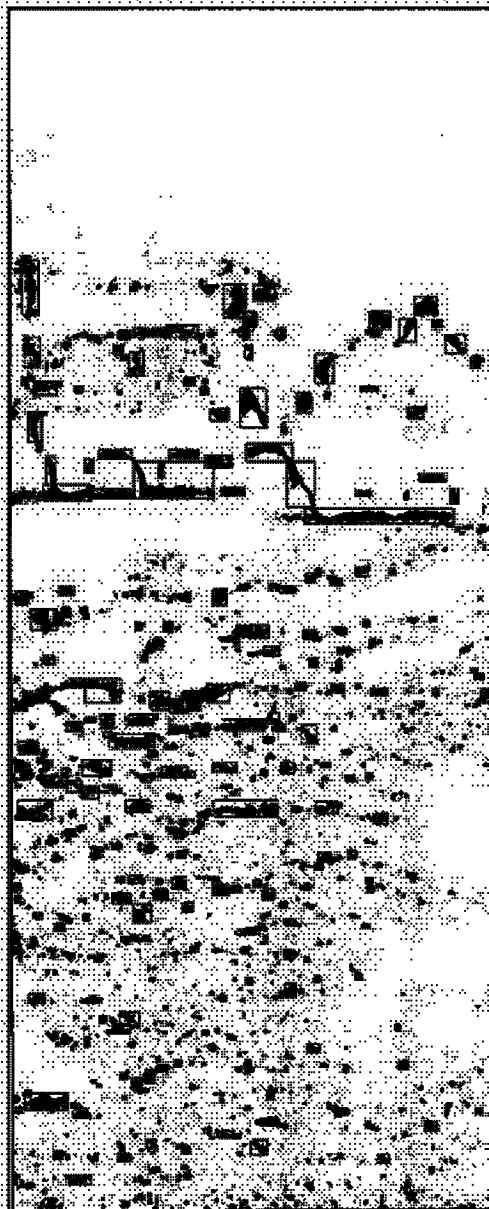


图 2c

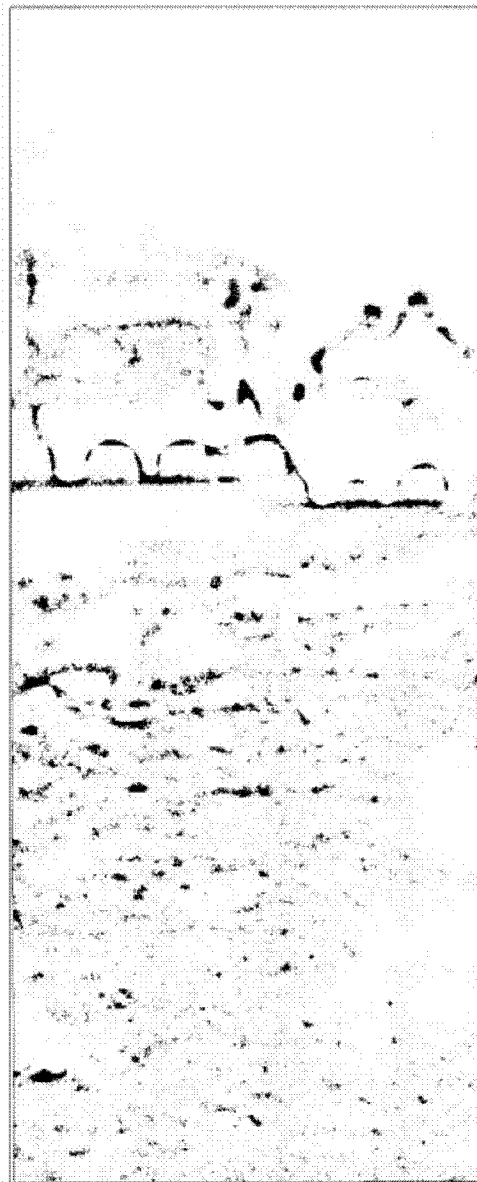
德军总司令海因里希·戈林元帅在柏林向希特勒汇报战况。希特勒在柏林的地下掩体中，通过无线电收听戈林的汇报。



### “令人生畏的名望”带来的新光彩

3月24日，隆美尔去德国汇报，德国的反击遏制了苏联的冬季攻势，并使战线稳定下来。隆美尔率领的勇猛的党卫队装甲团获得了“骷髅”勋章，但他的成功受到上峰某种程度上的嫉妒：元首对隆美尔自出机杼的行动耿耿于怀。隆美尔在汇报之后，他把他家眷亦列入因在哈尔夫战役中的英勇战功而受勋者之中。但是，无论如何，党卫队第二装甲团完成了希特勒对他们寄予的问题。

隆美尔下令他接受的任务部队整装待发，开赴库尔斯克附近，投入到大规模的反击攻势当中。这场攻势后来成为二次大战中最惨烈的坦克集群战斗。为了激励士气，党卫队帝国领袖海因里希·希特勒亲自来到哈尔夫去。他对他的部队发表了充满激情的讲话说：“我们永远不会让我们英雄的部队。我们在哈尔夫战役中获得的，指引着我们的，令人敬畏的荣誉受损。相反，我们要一往无前把它发扬光大。”



18 · 苏维埃战车

图 2d

## “令人敬畏的名望”带来的新光彩

3月15日，哈尔科夫战役结束，德国的反击遏制了苏联的冬季攻势，并使战线稳定下来。豪赛尔率领的勇猛的党卫队装甲军团获得了“雄狮”勋章，但他的成功受到上峰某种程度上的贬抑：元首对豪赛尔自出机杼的行动耿耿于怀，在延宕了几周之后，他才把豪赛尔纳入到因在哈尔科夫战役中的英勇战功而受勋者之中。但是，无论如何，党卫队第二装甲军团完成了希特勒对他们寄予的期望。

图 2e

“因为这是“斯大林”在一九四三年十月的时候，从“第一”号坦克师的成员手中，被“第一”号坦克师的成员所救下，被俘解送。



### “令人生畏的名望”带来的新光彩

5月15日，哈尔科夫战役结束，德国的反击遏制了苏联的冬季攻势，并使战线稳定下来。德军元帅曼施坦因的勇敢的党卫队装甲师获得了“维京”称号，但他的成功要受到一定程度上的贬抑：元首对塞瓦斯托波尔尔的行动耿耿于怀，在延宕了几周之后，他才把塞瓦斯托波尔尔的战斗中表现英勇的部队受勋者之中。但是，无论如何，党卫队第二装甲师完成了希特勒对他们寄予的期望。

希特勒下令他钟爱的党卫队部署在塞瓦斯托波尔尔附近，投入到大规模的反击战斗中，这场战役后来成为二次大战中规模最大的坦克集群战斗。为了激励士气，党卫队指挥官海因里希·希姆莱亲自来哈尔科夫。他对他的部队发表了洋溢着雄心壮志的演讲演说：“我们永远不会让我们卓越的部队，我们在哈尔科夫战役中赢得的，指引着我们的、令人敬畏的荣誉受损；相反，我们要一往无前，把它发扬光大。”



图 2 苏德坦克战

图 3

## “令人生畏的名望”带来的新光彩

3月15日,哈尔科夫战役结束,德国的反击遏制了苏联的冬季攻势,并使战线稳定下来。豪赛尔率领的勇猛的党卫队装甲军团获得了“雄狮”勋章,但他的成功受到上峰某种程度上的贬抑:元首对豪赛尔自出机杼的行动耿耿于怀,在延宕了几周之后,他才把豪赛尔纳入到因在哈尔科夫战役中的英勇战功而受勋者之中。但是,无论如何,党卫队第二装甲军团完成了希特勒对他们寄予的期望。

希特勒下令他钟爱的这支部队整装待发,开赴库尔斯克附近,投入到大规模的反击攻势当中,这场攻势后来成为二次大战中规模最大的坦克集群战斗。为了激励士气,党卫队帝国领袖海因里希·希姆莱受命来到哈尔科夫。他对他的部队发表了洋溢着雄心壮志的雄辩演说:“我们永远不会让我们卓越的部队,我们在哈尔科夫战役中赢得的、指引着我们的、令人敬畏的荣誉受损,相反,我们要一往无前并把它发扬光大。”

图 4



## “令人敬畏的名望”带来的新光彩

1941年12月，哈尔科夫战役陷入僵局，德国的反击遏制了苏联的冬季攻势，并使战线稳定下来。蒙塞尔多率领的勇敢的党卫队装甲团获得了“铁拳”勋章，但他的战功受到上级某种程度的怀疑：元首对蒙塞尔多自出机锋进行封锁封锁，在战争进行了几周之后，他才把蒙塞尔多列入因在哈尔科夫战役中的英勇战功而受勋者之中。但是，无论怎样，党卫队第二装甲团完成了希特勒对他们寄予的期望。

希特勒下令把蒙塞尔多支队队整编待援，在库尔斯克附近，投入到大规模的反击攻势当中，这场攻势后来成为二战中规模最大的坦克集群战役。因为激励士气，党卫队帝国领袖海因里希·希特勒亲自来到哈尔科夫，他对他的部队发表了洋溢着决心和意志的慷慨演说：“我们永远不会让我们卓越的精锐，我们在哈尔科夫战役中赢得的，指引着我们的，令人敬畏的荣誉受损，相反，我们要一往无前地把它发扬光大。”

图 5

**“令人生畏的名望”带来的新光彩**

3月15日,哈尔科夫战役结束,德国的反击遏制了苏联的冬季攻势,并使战

图 6

地区	2005	2006	2006年比2005年	
			绝对值	%
全国总计	1216.09	1363.17	147.08	12.1
内蒙古	268.57	301.56	32.99	12.3
河北	196.61	224.49	27.88	14.2
黑龙江	110.20	135.55	25.35	23.0
山西	20.41	22.02	1.61	7.9
新疆	214.88	222.92	8.04	3.7
陕西	31.22	37.54	6.32	20.2
辽宁	22.09	27.29	5.20	23.5
广西	22.09	23.03	1.94	8.8
宁夏	22.09	23.02	1.93	8.7
吉林	22.09	23.02	1.93	8.7
安徽	22.09	23.02	1.93	8.7
湖南	22.09	23.02	1.93	8.7
湖北	22.09	23.02	1.93	8.7
江西	22.09	23.02	1.93	8.7
浙江	22.09	23.02	1.93	8.7
福建	22.09	23.02	1.93	8.7
广东	22.09	23.02	1.93	8.7
海南	22.09	23.02	1.93	8.7
四川	22.09	23.02	1.93	8.7
重庆	22.09	23.02	1.93	8.7
贵州	22.09	23.02	1.93	8.7
云南	22.09	23.02	1.93	8.7
青海	22.09	23.02	1.93	8.7
甘肃	22.09	23.02	1.93	8.7
宁夏	22.09	23.02	1.93	8.7
新疆	22.09	23.02	1.93	8.7
西藏	22.09	23.02	1.93	8.7
台湾	22.09	23.02	1.93	8.7
香港	22.09	23.02	1.93	8.7
澳门	22.09	23.02	1.93	8.7

图 7a

地区	2005	2006
全国总计	1 216.09	1 363.17
内蒙古	268.57	301.56
河北	196.61	224.49
黑龙江	110.20	135.55

2006年比2005年增加	
绝对数	%
147.08	12.1
32.99	12.3
27.88	14.2
25.35	23.0

图 7b

地区	2005	2006
山东	70.41	92.50
新疆	214.88	222.92
河南	31.22	37.54
陕西	46.03	52.23

2006年比2005年增加	
绝对数	%
22.09	31.4
8.04	3.7
6.32	20.2
6.20	13.5

图 7c



图 8a



图 8b

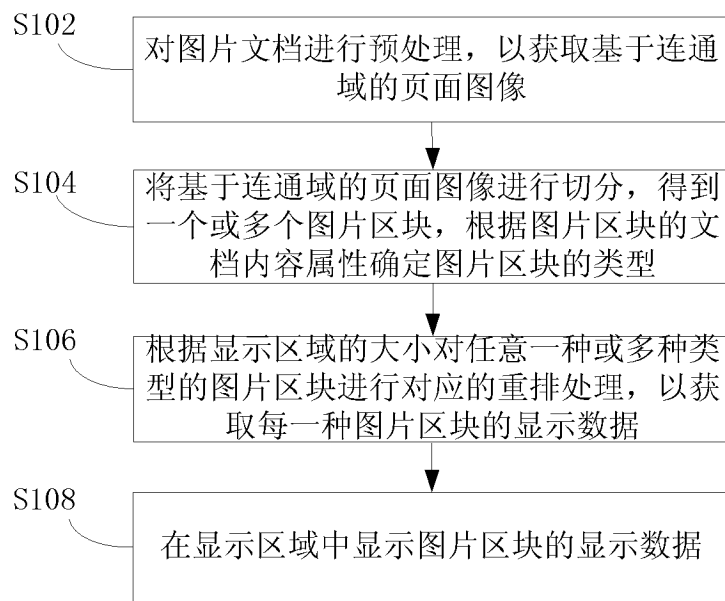


图 9

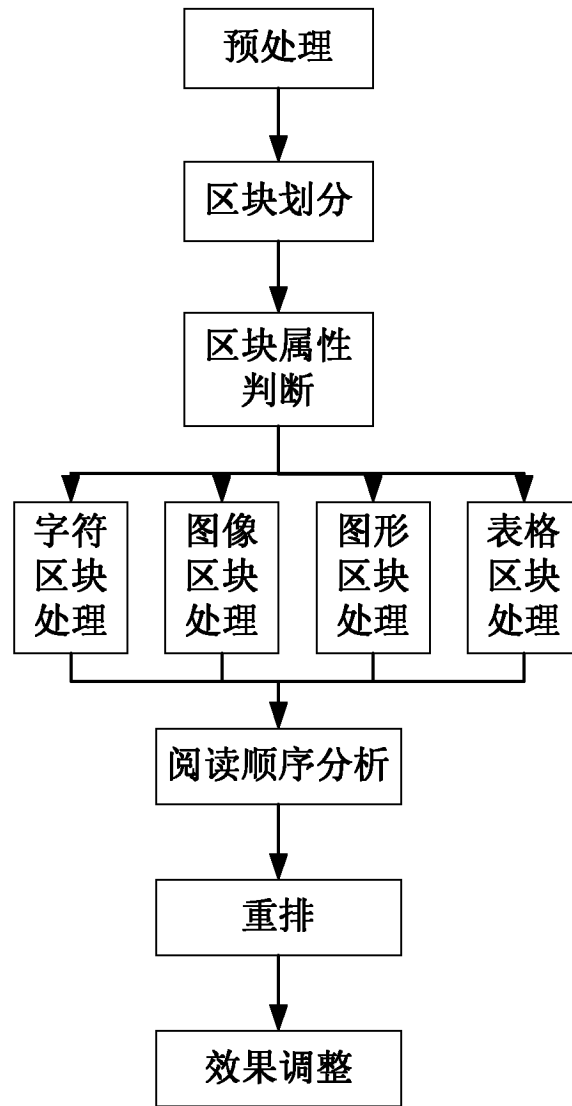


图 10

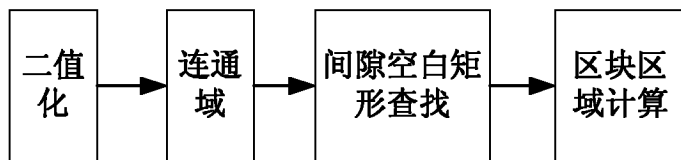


图 11a

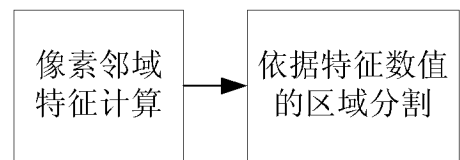


图 11b

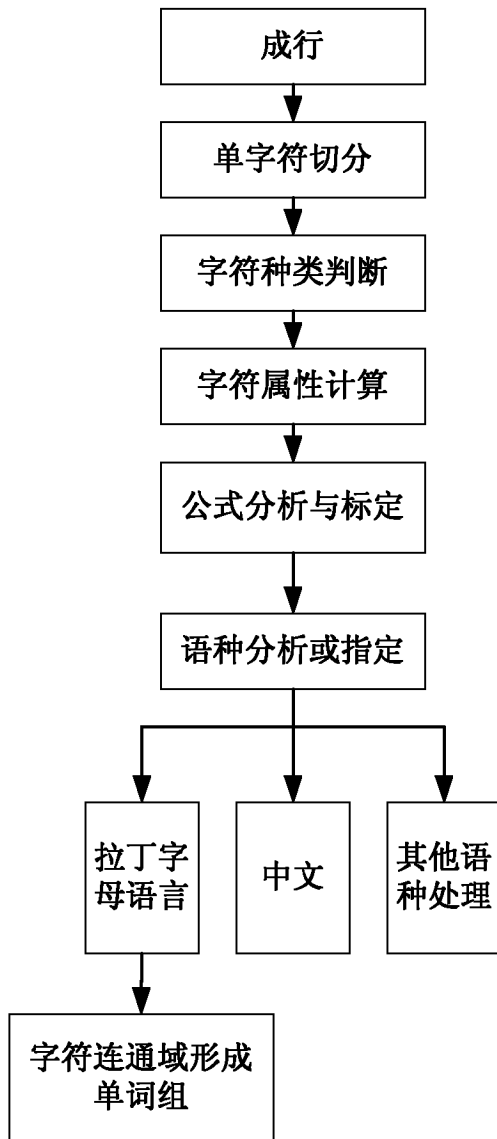


图 12

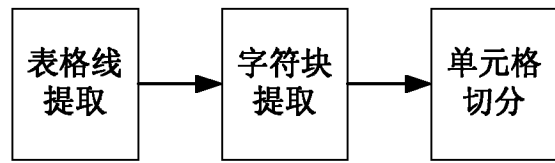


图 13

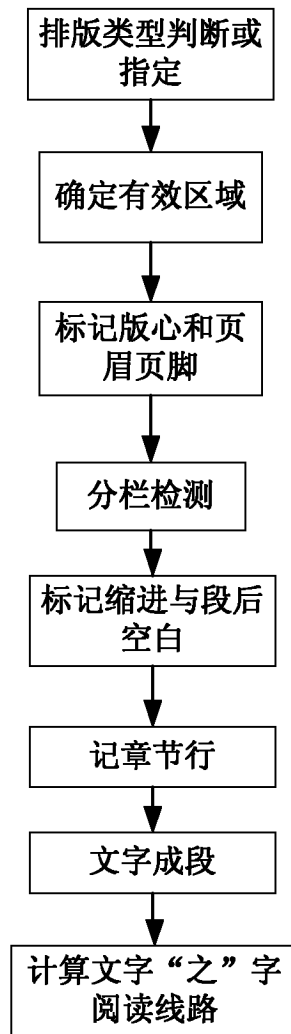


图 14