



(12)发明专利

(10)授权公告号 CN 106295486 B

(45)授权公告日 2019.11.05

(21)申请号 201510325512.X

(22)申请日 2015.06.12

(65)同一申请的已公布的文献号  
申请公布号 CN 106295486 A

(43)申请公布日 2017.01.04

(73)专利权人 富士通株式会社  
地址 日本神奈川县

(72)发明人 孙俊

(74)专利代理机构 北京集佳知识产权代理有限公司 11227  
代理人 康建峰 陈炜

(51)Int.Cl.  
G06K 9/00(2006.01)  
G06K 9/20(2006.01)  
G06K 9/32(2006.01)

(56)对比文件

JP 2004038530 A,2004.02.05,  
JP 2004038530 A,2004.02.05,  
CN 102375988 A,2012.03.14,  
CN 1928891 A,2007.03.14,  
JP 3409992 B2,2003.05.26,  
US 2011199627 A1,2011.08.18,

审查员 黄文琪

权利要求书2页 说明书9页 附图7页

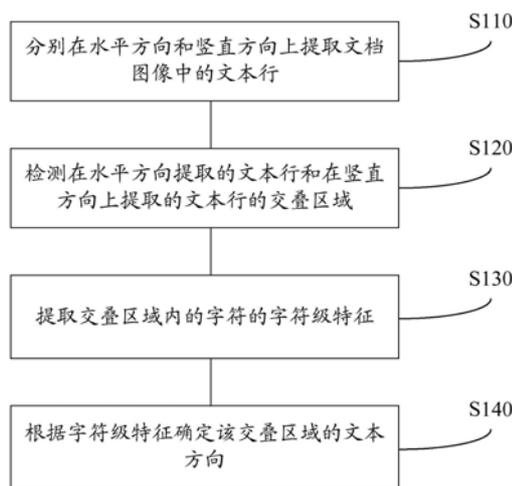
(54)发明名称

文档图像的处理方法和处理装置

(57)摘要

本发明公开了一种文档图像的处理方法和处理装置。该处理方法包括：分别在水平方向和竖直方向上提取所述文档图像的文本行；检测在水平方向提取的文本行和在竖直方向上提取的文本行的交叠区域；提取所述交叠区域内的字符的字符级特征；以及根据所述字符级特征确定所述交叠区域的文本方向。该方法利用了文档图像被提取为水平文本行和竖直文本行的交叠区域中字符的字符级特征来确定交叠区域的文本方向，从而提高了对文本行方向判断的准确性。

100



1. 一种文档图像的处理方法,包括:

分别在水平方向和竖直方向上提取所述文档图像的文本行;

检测在水平方向提取的文本行和在竖直方向上提取的文本行的交叠区域;

提取所述交叠区域内的字符的字符级特征;以及

根据所述字符级特征确定所述交叠区域的文本方向,

其中根据所述字符级特征确定所述交叠区域的文本方向包括:

根据所述交叠区域内至少一部分字符的文本方向以投票的方式来确定所述交叠区域的文本方向,其中每个字符的文本方向根据该字符的一个或多个字符级特征以投票的方式来确定;或者

根据所述交叠区域内至少一部分字符的一个或多个字符级特征的平均值以投票的方式来确定所述交叠区域的文本方向。

2. 如权利要求1所述的方法,其中字符的字符级特征包括以下特征中的一个或多个:

字符间距与字符高度的比值;

字符的高宽比;

字符间距的方差;

字符的顶部基线对齐方差;以及

字符的底部基线对齐方差。

3. 如权利要求1所述的方法,还包括:

在所述交叠区域提取文本行级特征;以及

根据所述文本行级特征来判断所述交叠区域的文本方向。

4. 如权利要求3所述的方法,还包括:

利用预设的阈值范围过滤根据所述文本行级特征的判断结果,其中当所述交叠区域作为水平文本行的文本行级特征与所述交叠区域作为竖直文本行的文本行级特征的比值处于所述预设的阈值范围内时,过滤掉根据所述文本行级特征的判断结果。

5. 如权利要求3或4所述的方法,其中所述文本行级特征包括文本行中的笔划总数和/或像素点总数;

当所述交叠区域作为水平文本行的文本行级特征与所述交叠区域作为竖直文本行的文本行级特征的比值大于1时,则判断所述交叠区域为水平方向;并且

当所述交叠区域作为水平文本行的文本行级特征与所述交叠区域作为竖直文本行的文本行级特征的比值小于1时,则判断所述交叠区域为竖直方向。

6. 如权利要求1所述的方法,其中检测在水平方向提取的文本行和在竖直方向上提取的文本行的交叠区域包括:

构建交叠矩阵,其中所述交叠矩阵中的每个元素表示在水平方向提取的一个文本行和在竖直方向提取的一个文本行的交叠情况;

分别在水平和竖直方向对所述交叠矩阵中的元素进行融合;以及

根据元素已经过融合后的交叠矩阵来确定所述交叠区域。

7. 如权利要求6所述的方法,其中在分别在水平和竖直方向对所述交叠矩阵中的元素进行融合时,对所述交叠矩阵中的元素在水平和竖直方向上交替地进行迭代融合计算。

8. 一种文档图像的处理装置,包括:

文本行提取单元,分别在水平方向和竖直方向上提取所述文档图像的文本行;

检测单元,检测所述文本行提取单元在水平方向提取的文本行和在竖直方向上提取的文本行的交叠区域;

字符级特征提取单元,提取所述检测单元检测出的交叠区域内的字符的字符级特征;  
以及

方向确定单元,根据所述字符级特征提取单元所提取的字符级特征确定所述交叠区域的文本方向,

其中根据所述字符级特征确定所述交叠区域的文本方向包括:

根据所述交叠区域内至少一部分字符的文本方向以投票的方式来确定所述交叠区域的文本方向,其中每个字符的文本方向根据该字符的一个或多个字符级特征以投票的方式来确定;或者

根据所述交叠区域内至少一部分字符的一个或多个字符级特征的平均值以投票的方式来确定所述交叠区域的文本方向。

## 文档图像的处理方法和处理装置

### 技术领域

[0001] 本发明属于图像处理领域,具体涉及文档图像的处理方法和处理装置。

### 背景技术

[0002] 随着信息技术的高速发展,涌现了包含文字信息的大量文档,这其中既有可编辑的文档,又有不可编辑的文档。对于不可编辑的文档,例如文档图像,需要提取其中的文本行,才能够获得可编辑的文字。

[0003] 现有技术中,已提出了一些提取文本行的方法,例如,E.Kim,et al,“Scene text extraction using focus of mobile camera”.Proceedings of the 10th International conference on Document Analysis and Recognition,p166~170,2009.7.26~29,Barcelona,以及N.Ezaki,et al,“Text detection from natural scene images:towards a system for visually impaired persons”.International Conference on Pattern Recognition.P683-686,2004。通过引用上述文献,其全部内容结合于此。

[0004] 然而,在版面较复杂的文档图像中,文字的方向可以是水平的或竖直的,或者其中一部分文字是水平方向的,另一部分文字是竖直方向的。对于此类文档图像,在提取文本行的过程中,需要确定文本的方向。

[0005] 为此,本申请人在申请号为201010257650.6的在先中国专利申请中,提出了一种文件图像的处理方法,其利用了文字的文本行级特征来确定文本的方向。通过引用该专利文献,其全部内容结合于此。

[0006] 然而,该方法仍有其局限性,对于某些文档图像,无法得到较好的处理效果。

### 发明内容

[0007] 有鉴于此,本发明提出了一种文档图像的处理方法和处理装置。该方法和装置利用了文档图像被提取为水平文本行和竖直文本行的交叠区域中字符的字符级特征来确定交叠区域的文本方向。

[0008] 根据本发明的一个方面,提供了一种文档图像的处理方法,包括:分别在水平方向和竖直方向上提取所述文档图像的文本行;检测在水平方向提取的文本行和在竖直方向上提取的文本行的交叠区域;提取所述交叠区域内的字符的字符级特征;以及根据所述字符级特征确定所述交叠区域的文本方向。

[0009] 根据本发明的另一方面,提供了一种文档图像的处理装置,包括:文本行提取单元,分别在水平方向和竖直方向上提取所述文档图像的文本行;检测单元,检测所述文本行提取单元在水平方向提取的文本行和在竖直方向上提取的文本行的交叠区域;字符级特征提取单元,提取所述检测单元检测出的交叠区域内的字符的字符级特征;以及方向确定单元,根据所述字符级特征提取单元所提取的字符级特征确定所述交叠区域的文本方向。

[0010] 根据本发明所提供的技术方案,利用了字符级特征来确定文档图像中的水平文本

行和垂直文本行的交叠区域的文本方向,从而提高了对文本行方向判断的准确性。

### 附图说明

[0011] 参照附图来阅读本发明的各实施方式,将更容易理解本发明的其它特征和优点,在此描述的附图只是为了对本发明的实施方式进行示意性说明的目的,而非全部可能的实施,并且不旨在限制本发明的范围。在附图中:

[0012] 图1示出了根据本发明一个实施方式文档图像的处理方法的流程图;

[0013] 图2示出了一个示例性的文档图像;

[0014] 图3示出了根据本发明另一个实施方式文档图像的处理方法的流程图;

[0015] 图4示出了根据本发明又一个实施方式文档图像的处理方法的流程图;

[0016] 图5示出了根据本发明一个实施方式检测在水平方向提取的文本行和在垂直方向上提取的文本行的交叠区域的流程图;

[0017] 图6示出了根据本发明一个实施方式文档图像的处理装置的框图;

[0018] 图7示出了根据本发明另一实施方式文档图像的处理装置的框图;

[0019] 图8示出了根据本发明另一实施方式文档图像的处理装置的框图;

[0020] 图9示出了根据本发明一个实施方式的检测单元的框图;以及

[0021] 图10示出了可用于实施根据本发明实施例的方法和系统的计算机的示意性框图。

### 具体实施方式

[0022] 现参照附图对本发明的实施方式进行详细描述。应注意,以下描述仅仅是示例性的,而并不旨在限制本发明。此外,在以下描述中,将采用相同的附图标号表示不同附图中的相同或相似的部件。在以下描述的不同实施方式中的不同特征,可彼此结合,以形成本发明范围内的其他实施方式。

[0023] 图1示出了根据本发明一个实施方式文档图像的处理方法的流程图。如图1所示,该方法100包括步骤S110至S140。在步骤S110中,分别在水平方向和垂直方向上提取文档图像中的文本行。具体提取的方法,可采用本领域已知的任何适当的方法来进行提取。在水平和垂直方向上分别提取文本行后,可能会出现一些错误。例如,对于实际上是水平的文本行,在水平方向上进行提取时,已将其提取,但在垂直方向上进行提取时,又将其误提取为垂直文本行。对此,在步骤S120中,检测在水平方向提取的文本行和在垂直方向上提取的文本行的交叠区域。在交叠区域中的字符既被提取为水平文本行,又被提取为垂直文本行。随后,在步骤S130中,提取交叠区域内的字符的字符级特征。在步骤S140中,根据字符级特征确定该交叠区域的文本方向。

[0024] 由此,利用了交叠区域中字符的字符级特征来确定交叠区域的文本方向,其准确性更高。

[0025] 根据本发明的一个实施方式,字符的字符级特征可以包括:字符间距与字符高度的比值、字符的高宽比、字符间距的方差、字符的顶部基线对齐方差和/或字符的底部基线对齐方差。下面以图2为例对这些字符级特征进行详细解释。

[0026] 图2示出了一个示例性的文档图像。如图2所示,其中的区域A中的字符为水平方向排列,区域B中的字符为垂直方向排列。对于字符间距与字符高度的比值而言,如图2中区域

A的第二行Ar2所示,其中字符的间距较小,而字符高度较大(大于字符宽度),因此,其作为水平方向的文本行,字符间距与字符高度的比值较小。而如果区域A中的字符也被提取为竖直文本行,即区域A为交叠区域,如图2中区域A的第三列Ac3所示,其中字符的间距较大,而字符高度较小(此时将字符在水平方向上的尺寸认定为字符高度)。因此,当区域A被误提取为竖直方向的文本行时,字符间距与字符高度的比值较大。由此,可通过利用字符的字符级特征的该性质来判断字符的正确文本方向。

[0027] 对于字符的高宽比而言,如图2中的区域A所示,当区域A被正确地提取为水平方向的文本行时,其中字符的高宽比应大于1,也就是说,字符的高度大于宽度。而当区域A被错误地提取为竖直方向的文本行时,其中字符的高宽比应小于1,也就是说,字符的高度小于宽度。由此,可通过利用字符的字符级特征的该性质来判断字符的正确文本方向。

[0028] 对于字符间距的方差而言,如图2中的区域A所示,当区域A被正确地提取为水平方向的文本行时,其中字符与字符之间的间距都比较接近,也就是说,字符间距的方差较小。而当区域A被错误地提取为竖直方向的文本行时,由于有些字符较高(如b、d、g),而有些字符较矮(如c、e),所以区域A的一列字符中字符与字符之间的间距差别较大,也就是说,字符间距的方差较大。由此,可通过利用字符的字符级特征的该性质来判断字符的正确文本方向。

[0029] 对于字符的顶部基线对齐方差而言,如图2中的区域A所示,当区域A被正确地提取为水平方向的文本行时,由于每一行中字符的顶部均未对齐,所以字符的顶部基线对齐方差较大。而当区域A被错误地提取为竖直方向的文本行时,由于每一列中字符的左侧或右侧(即此时认为的字符顶部)对齐程度较高,所以字符的顶部基线对齐方差较小。由此,可通过利用字符的字符级特征的该性质来判断字符的正确文本方向。与此类似,也可利用字符的底部基线对齐方差来判断字符的正确文本方向。

[0030] 类似地,对于图2所示的区域B而言,也可以利用上述字符级特征的性质来判断字符的正确文本方向。

[0031] 根据本发明一个实施方式,在上述步骤S140中确定交叠区域的文本方向时,可根据每个字符的一个或多个字符级特征以投票的方式来确定该字符的文本方向。具体地,对于任一个字符而言,均具有文本方向。如果仅考虑一个字符级特征,则根据该字符级特征就可确定该字符的文本方向。如果考虑多个字符级特征,由于每个字符级特征均会对该字符的文本方向做出一个判断,因此,可按该多个字符级特征以投票的方式来确定该字符的文本方向。

[0032] 随后,在各字符的文本方向确定后,可根据交叠区域内至少一部分字符的文本方向以投票的方式来确定该交叠区域的文本方向。具体地,在交叠区域中通常包含多个字符,由上述过程,可确定各字符的文本方向。此后,可利用交叠区域内的至少一部分字符,根据它们各自确定的文本方向,以投票的方式来确定该交叠区域的文本方向。

[0033] 根据本发明另一个实施方式,与上述方式不同,在上述步骤S140中确定交叠区域的文本方向时,可对交叠区域内的至少一部分字符的同一种字符级特征取平均值。如此,根据该字符级特征的平均值可确定交叠区域的文本方向。如果考虑多种字符级特征,可根据多种字符级特征各自的平均值,以投票的方式来确定该交叠区域的文本方向。

[0034] 图3示出了根据本发明另一个实施方式文档图像的处理方法的流程图。为了简要

起见,以下将仅描述图3所示的实施方式与图1的不同之处,并将略去其相同之处的详细描述。如图3所示,该方法100'包括步骤S110至S160。在步骤S150中,在交叠区域内提取文本行级特征。随后,在步骤S160中,根据文本行级特征来判断交叠区域的文本方向。

[0035] 根据该实施方式,既考虑了交叠区域内字符的字符级特征,又考虑了文本行级特征,来确定交叠区域的文本方向,从而能够进一步提高操作的准确性。可以理解,步骤S130和S140与步骤S150和S160之间可以并行操作,也可以彼此调换顺序。

[0036] 图4示出了根据本发明又一个实施方式文档图像的处理方法的流程图。为了简要起见,以下将仅描述图4所示的实施方式与图3的不同之处,并将略去其相同之处的详细描述。如图4所示,该方法100''包括步骤S110至S170。在步骤S170中,利用预设的阈值范围过滤根据文本行级特征的判断结果。在实际操作中,可能会遇到字符级特征与文本行级特征的判断结果相矛盾的情况。由于字符级特征的判断准确性通常高于文本行级特征,因此,可为文本行级特征的判断结果设定阈值范围,当文本行级特征的判断结果处于该范围内时,则可过滤掉根据文本行级特征的判断结果,即不考虑根据文本行级特征的判断结果。具体地,当交叠区域作为水平文本行的文本行级特征与作为竖直文本行的文本行级特征的比值处于该预设的阈值范围内时,则过滤掉根据文本行级特征的判断结果,从而可避免根据文本行级特征可能出现的错误结果所带来的干扰。

[0037] 根据本发明的另一实施方式,文本行级特征包括文本行中的笔划总数和/或像素点总数。由于在错误的方向上提取的文本行有可能在一定程度上漏掉一部分字符(即,漏掉了一些笔划或像素点)。因此,当交叠区域作为水平文本行的文本行级特征与作为竖直文本行的文本行级特征的比值大于1时(即作为水平文本行的笔划总数或像素点总数多于作为竖直文本行的笔划总数或像素点总数时),则判断交叠区域为水平方向。反之,当交叠区域作为水平文本行的文本行级特征与作为竖直文本行的文本行级特征的比值小于1时(即作为水平文本行的笔划总数或像素点总数少于作为竖直文本行的笔划总数或像素点总数时),则判断交叠区域为竖直方向。

[0038] 图5示出了根据本发明一个实施方式检测在水平方向提取的文本行和在竖直方向上提取的文本行的交叠区域的流程图。如图5所示,上述步骤S120可包括子步骤S121至S123。在子步骤S121中,构建交叠矩阵。所构建的交叠矩阵中的每个元素表示在水平方向提取的一个文本行和在竖直方向提取的一个文本行的交叠情况。交叠矩阵具体的构建方式可采用上述申请号为201010257650.6的在先中国专利申请中所披露的方式或任何其他已知的适当方式。

[0039] 随后,在子步骤S122中,分别在水平和竖直方向对所构建的交叠矩阵中的元素进行融合。交叠矩阵中的元素具体的融合方式也可采用上述申请号为201010257650.6的在先中国专利申请中所披露的方式或任何其他已知的适当方式。在对其中的元素进行融合后,在子步骤S123中,可根据交叠矩阵来确定交叠区域。

[0040] 根据本发明的一个实施方式,在分别在水平和竖直方向对交叠矩阵中的元素进行融合时,可对交叠矩阵中的元素在水平和竖直方向上交替地进行迭代融合计算。由此,通过迭代的方式,可有效地对交叠矩阵中的元素进行融合计算。并且,根据元素已经过迭代融合的交叠矩阵所确定的交叠区域更加准确。

[0041] 图6示出了根据本发明一个实施方式文档图像的处理装置的框图。如图6所示,该

装置600可包括文本行提取单元610、检测单元620、字符级特征提取单元630和方向确定单元640。文本行提取单元610可分别在水平方向和竖直方向上提取文档图像的文本行。检测单元620可检测文本行提取单元610在水平方向提取的文本行和在竖直方向上提取的文本行的交叠区域。字符级特征提取单元630可提取检测单元620检测出的交叠区域内的字符的字符级特征。方向确定单元640可根据字符级特征提取单元630所提取的字符级特征确定交叠区域的文本方向。

[0042] 根据本发明一个实施方式,方向确定单元640可包括投票子单元。该投票子单元可根据每个字符的一个或多个字符级特征以投票的方式来确定该字符的文本方向,并可根据交叠区域内至少一部分字符的文本方向以投票的方式来确定交叠区域的文本方向。可替换地,该投票子单元可根据交叠区域内至少一部分字符的一个或多个字符级特征的平均值以投票的方式来确定交叠区域的文本方向。

[0043] 根据本发明一个实施方式,字符的字符级特征可以包括:字符间距与字符高度的比值、字符的高宽比、字符间距的方差、字符的顶部基线对齐方差和/或字符的底部基线对齐方差。

[0044] 图7示出了根据本发明另一实施方式文档图像的处理装置的框图。如图7所示,除了文本行提取单元610、检测单元620、字符级特征提取单元630和方向确定单元640,装置600'还可包括文本行级特征提取单元650。文本行级特征提取单元650可在交叠区域提取文本行级特征,方向确定单元640除了可根据字符级特征提取单元630所提取的字符级特征来判断交叠区域的文本方向外,还可根据文本行级特征提取单元650所提取的文本行级特征来判断交叠区域的文本方向。

[0045] 图8示出了根据本发明另一实施方式文档图像的处理装置的框图。如图8所示,除了文本行提取单元610、检测单元620、字符级特征提取单元630、方向确定单元640和文本行级特征提取单元650,装置600"还可包括结果筛选单元660。结果筛选单元660可利用预设的阈值范围过滤方向确定单元640根据文本行级特征的判断结果。具体地,当交叠区域作为水平文本行的文本行级特征与作为竖直文本行的文本行级特征的比值处于预设的阈值范围内时,结果筛选单元660可过滤掉方向确定单元640根据文本行级特征的判断结果。

[0046] 根据本发明一个实施方式,文本行级特征包括文本行中的笔划总数和/或像素点总数。当交叠区域作为水平文本行的文本行级特征与作为竖直文本行的文本行级特征的比值大于1时,则判断交叠区域为水平方向。反之,当交叠区域作为水平文本行的文本行级特征与作为竖直文本行的文本行级特征的比值小于1时,则判断交叠区域为竖直方向。

[0047] 图9示出了根据本发明一个实施方式的检测单元的框图。如图9所示,检测单元620可包括矩阵构建子单元621、融合子单元622和区域确定子单元623。矩阵构建子单元621可构建交叠矩阵,该交叠矩阵中的每个元素表示在水平方向提取的一个文本行和在竖直方向提取的一个文本行的交叠情况。融合子单元622可分别在水平和竖直方向对矩阵构建子单元621所构建的交叠矩阵中的元素进行融合。区域确定子单元623可根据元素已经过融合子单元622融合的交叠矩阵来确定交叠区域。

[0048] 根据本发明一个实施方式,在分别在水平和竖直方向对交叠矩阵中的元素进行融合时,融合子单元622可对交叠矩阵中的元素在水平和竖直方向上交替地进行迭代融合计算。

[0049] 另外,这里尚需指出的是,上述系统中各个组成部件可以通过软件、固件、硬件或其组合的方式进行配置。配置可使用的具体手段或方式为本领域技术人员所熟知,在此不再赘述。在通过软件或固件实现的情况下,从存储介质或网络向具有专用硬件结构的计算机(例如图10所示的通用计算机1000)安装构成该软件的程序,该计算机在安装各种程序时,能够执行各种功能等。

[0050] 图10示出了可用于实施根据本发明实施例的方法和系统的计算机的示意性框图。

[0051] 在图10中,中央处理单元(CPU)1001根据只读存储器(ROM)1002中存储的程序或从存储部分1008加载到随机存取存储器(RAM)1003的程序执行各种处理。在RAM 1003中,还根据需要存储当CPU 1001执行各种处理等等时所需的数据。CPU 1001、ROM 1002和RAM 1003经由总线1004彼此连接。输入/输出接口1005也连接到总线1004。

[0052] 下述部件连接到输入/输出接口1005:输入部分1006(包括键盘、鼠标等等)、输出部分1007(包括显示器,比如阴极射线管(CRT)、液晶显示器(LCD)等,和扬声器等)、存储部分1008(包括硬盘等)、通信部分1009(包括网络接口卡比如LAN卡、调制解调器等)。通信部分1009经由网络比如因特网执行通信处理。根据需要,驱动器1010也可连接到输入/输出接口1005。可拆卸介质1011比如磁盘、光盘、磁光盘、半导体存储器等等可以根据需要被安装在驱动器1010上,使得从中读出的计算机程序根据需要被安装到存储部分1008中。

[0053] 在通过软件实现上述系列处理的情况下,从网络比如因特网或存储介质比如可拆卸介质1011安装构成软件的程序。

[0054] 本领域的技术人员应当理解,这种存储介质不局限于图10所示的其中存储有程序、与设备相分离地分发以向用户提供程序的可拆卸介质1011。可拆卸介质1011的例子包含磁盘(包含软盘(注册商标)、光盘(包含光盘只读存储器(CD-ROM)和数字通用盘(DVD))、磁光盘(包含迷你盘(MD)(注册商标))和半导体存储器。或者,存储介质可以是ROM 1002、存储部分1008中包含的硬盘等等,其中存有程序,并且与包含它们的设备一起被分发给用户。

[0055] 本发明还提出一种存储有机器可读的指令代码的程序产品。所述指令代码由机器读取并执行时,可执行上述根据本发明实施方式的方法。

[0056] 相应地,用于承载上述存储有机器可读的指令代码的程序产品的存储介质也包括在本发明的范围内。所述存储介质包括但不限于软盘、光盘、磁光盘、存储卡、存储棒等等。

[0057] 应当注意,本发明的方法不限于按照说明书中描述的时间顺序来执行,也可以按照其他的次序顺序地、并行地或独立地执行。因此,本说明书中描述的方法的执行顺序不对本发明的技术范围构成限制。

[0058] 以上对本发明各实施方式的描述是为了更好地理解本发明,其仅仅是示例性的,而非旨在对本发明进行限制。应注意,在以上描述中,针对一种实施方式描述和/或示出的特征可以以相同或类似的方式在一个或更多个其它实施方式中使用,与其它实施方式中的特征相组合,或替代其它实施方式中的特征。本领域技术人员可以理解,在不脱离本发明的发明构思的情况下,针对以上所描述的实施方式进行的各种变化和修改,均属于本发明的范围内。

[0059] 综上,在根据本发明的实施例中,本发明提供了如下技术方案。

[0060] 方案1、一种文档图像的处理方法,包括:

- [0061] 分别在水平方向和竖直方向上提取所述文档图像的文本行；
- [0062] 检测在水平方向提取的文本行和在竖直方向上提取的文本行的交叠区域；
- [0063] 提取所述交叠区域内的字符的字符级特征；以及
- [0064] 根据所述字符级特征确定所述交叠区域的文本方向。
- [0065] 方案2、如方案1所述的方法，其中根据所述字符级特征确定所述交叠区域的文本方向包括：
- [0066] 根据所述交叠区域内至少一部分字符的文本方向以投票的方式来确定所述交叠区域的文本方向，其中每个字符的文本方向根据该字符的一个或多个字符级特征以投票的方式来确定。
- [0067] 方案3、如方案1所述的方法，其中根据所述字符级特征确定所述交叠区域的文本方向包括：
- [0068] 根据所述交叠区域内至少一部分字符的一个或多个字符级特征的平均值以投票的方式来确定所述交叠区域的文本方向。
- [0069] 方案4、如方案1至3中任一项所述的方法，其中字符的字符级特征包括以下特征中的一个或多个：
- [0070] 字符间距与字符高度的比值；
- [0071] 字符的高宽比；
- [0072] 字符间距的方差；
- [0073] 字符的顶部基线对齐方差；以及
- [0074] 字符的底部基线对齐方差。
- [0075] 方案5、如方案1至4中任一项所述的方法，还包括：
- [0076] 在所述交叠区域提取文本行级特征；以及
- [0077] 根据所述文本行级特征来判断所述交叠区域的文本方向。
- [0078] 方案6、如方案5所述的方法，还包括：
- [0079] 利用预设的阈值范围过滤根据所述文本行级特征的判断结果，其中当所述交叠区域作为水平文本行的文本行级特征与所述交叠区域作为竖直文本行的文本行级特征的比值处于所述预设的阈值范围内时，过滤掉根据所述文本行级特征的判断结果。
- [0080] 方案7、如方案5或6所述的方法，其中所述文本行级特征包括文本行中的笔划总数和/或像素点总数；
- [0081] 当所述交叠区域作为水平文本行的文本行级特征与所述交叠区域作为竖直文本行的文本行级特征的比值大于1时，则判断所述交叠区域为水平方向；并且
- [0082] 当所述交叠区域作为水平文本行的文本行级特征与所述交叠区域作为竖直文本行的文本行级特征的比值小于1时，则判断所述交叠区域为竖直方向。
- [0083] 方案8、如方案1至7中任一项所述的方法，其中检测在水平方向提取的文本行和在竖直方向上提取的文本行的交叠区域包括：
- [0084] 构建交叠矩阵，其中所述交叠矩阵中的每个元素表示在水平方向提取的一个文本行和在竖直方向提取的一个文本行的交叠情况；
- [0085] 分别在水平和竖直方向对所述交叠矩阵中的元素进行融合；以及
- [0086] 根据元素已经过融合交叠矩阵来确定所述交叠区域。

[0087] 方案9、如方案8所述的方法,其中在分别在水平和垂直方向对所述交叠矩阵中的元素进行融合时,对所述交叠矩阵中的元素在水平和垂直方向上交替地进行迭代融合计算。

[0088] 方案10、一种文档图像的处理装置,包括:

[0089] 文本行提取单元,分别在水平方向和垂直方向上提取所述文档图像的文本行;

[0090] 检测单元,检测所述文本行提取单元在水平方向提取的文本行和在垂直方向上提取的文本行的交叠区域;

[0091] 字符级特征提取单元,提取所述检测单元检测出的交叠区域内的字符的字符级特征;以及

[0092] 方向确定单元,根据所述字符级特征提取单元所提取的字符级特征确定所述交叠区域的文本方向。

[0093] 方案11、如方案10所述的装置,其中所述方向确定单元包括:

[0094] 投票子单元,根据每个字符的一个或多个字符级特征以投票的方式来确定该字符的文本方向,并根据所述交叠区域内至少一部分字符的文本方向以投票的方式来确定所述交叠区域的文本方向。

[0095] 方案12、如方案10所述的装置,其中所述方向确定单元包括:

[0096] 投票子单元,根据所述交叠区域内至少一部分字符的一个或多个字符级特征的平均值以投票的方式来确定所述交叠区域的文本方向。

[0097] 方案13、如方案10至12中任一项所述的装置,其中字符的字符级特征包括以下特征中的一个或多个:

[0098] 字符间距与字符高度的比值;

[0099] 字符的高宽比;

[0100] 字符间距的方差;

[0101] 字符的顶部基线对齐方差;以及

[0102] 字符的底部基线对齐方差。

[0103] 方案14、如方案10至13中任一项所述的装置,还包括:

[0104] 文本行级特征提取单元,在所述交叠区域提取文本行级特征,其中所述方向确定单元还根据所述文本行级特征提取单元所提取的文本行级特征来判断所述交叠区域的文本方向。

[0105] 方案15、如方案14所述的装置,还包括:

[0106] 结果筛选单元,利用预设的阈值范围过滤所述方向确定单元根据所述文本行级特征判断的结果,其中当所述交叠区域作为水平文本行的文本行级特征与所述交叠区域作为垂直文本行的文本行级特征的比值处于所述预设的阈值范围内时,所述结果筛选单元过滤掉所述方向确定单元根据所述文本行级特征判断的结果。

[0107] 方案16、如方案14或15所述的装置,其中所述文本行级特征包括文本行中的笔划总数和/或像素点总数;

[0108] 当所述交叠区域作为水平文本行的文本行级特征与所述交叠区域作为垂直文本行的文本行级特征的比值大于1时,则判断所述交叠区域为水平方向;并且

[0109] 当所述交叠区域作为水平文本行的文本行级特征与所述交叠区域作为垂直文本

行的文本行级特征的比值小于1时,则判断所述交叠区域为竖直方向。

[0110] 方案17、如方案10至16中任一项所述的装置,其中所述检测单元包括:

[0111] 矩阵构建子单元,构建交叠矩阵,其中所述交叠矩阵中的每个元素表示在水平方向提取的一个文本行和在竖直方向提取的一个文本行的交叠情况;

[0112] 融合子单元,分别在水平和竖直方向对所述矩阵构建子单元构建的交叠矩阵中的元素进行融合;以及

[0113] 区域确定子单元,根据元素已经过所述融合子单元融合的交叠矩阵来确定所述交叠区域。

[0114] 方案18、如方案17所述的装置,其中在分别在水平和竖直方向对所述交叠矩阵中的元素进行融合时,所述融合子单元对所述交叠矩阵中的元素在水平和竖直方向上交替地进行迭代融合计算。

100

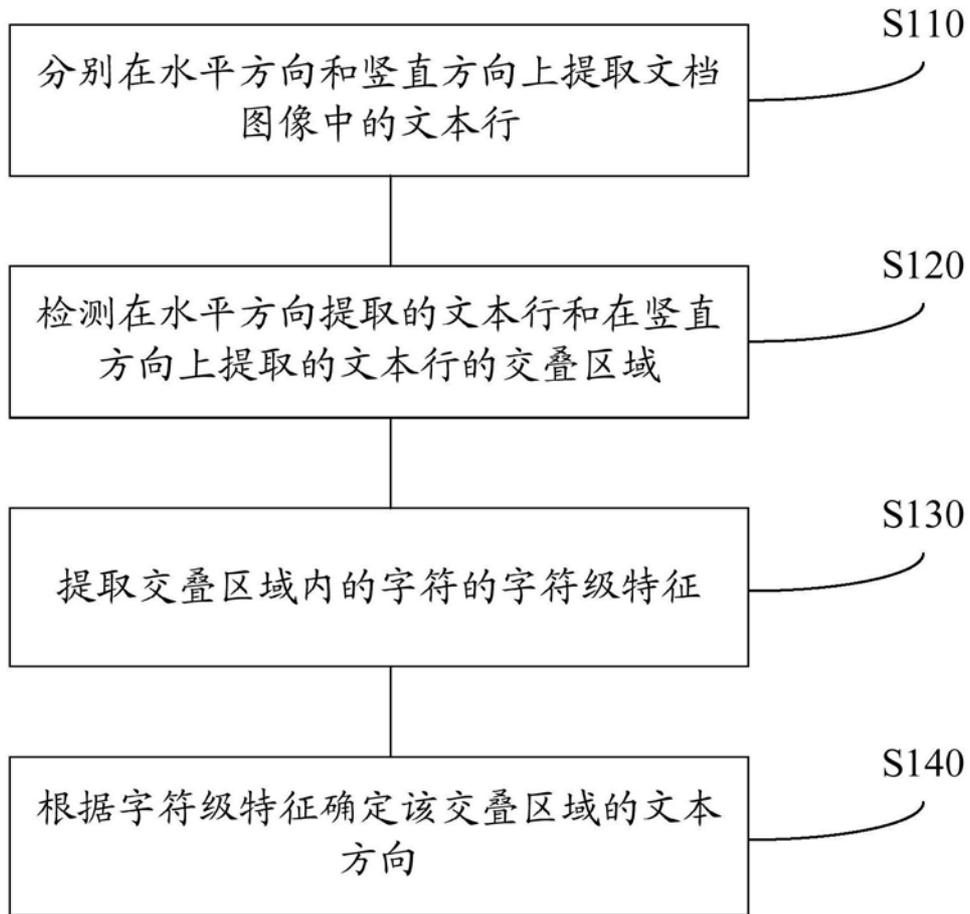


图1

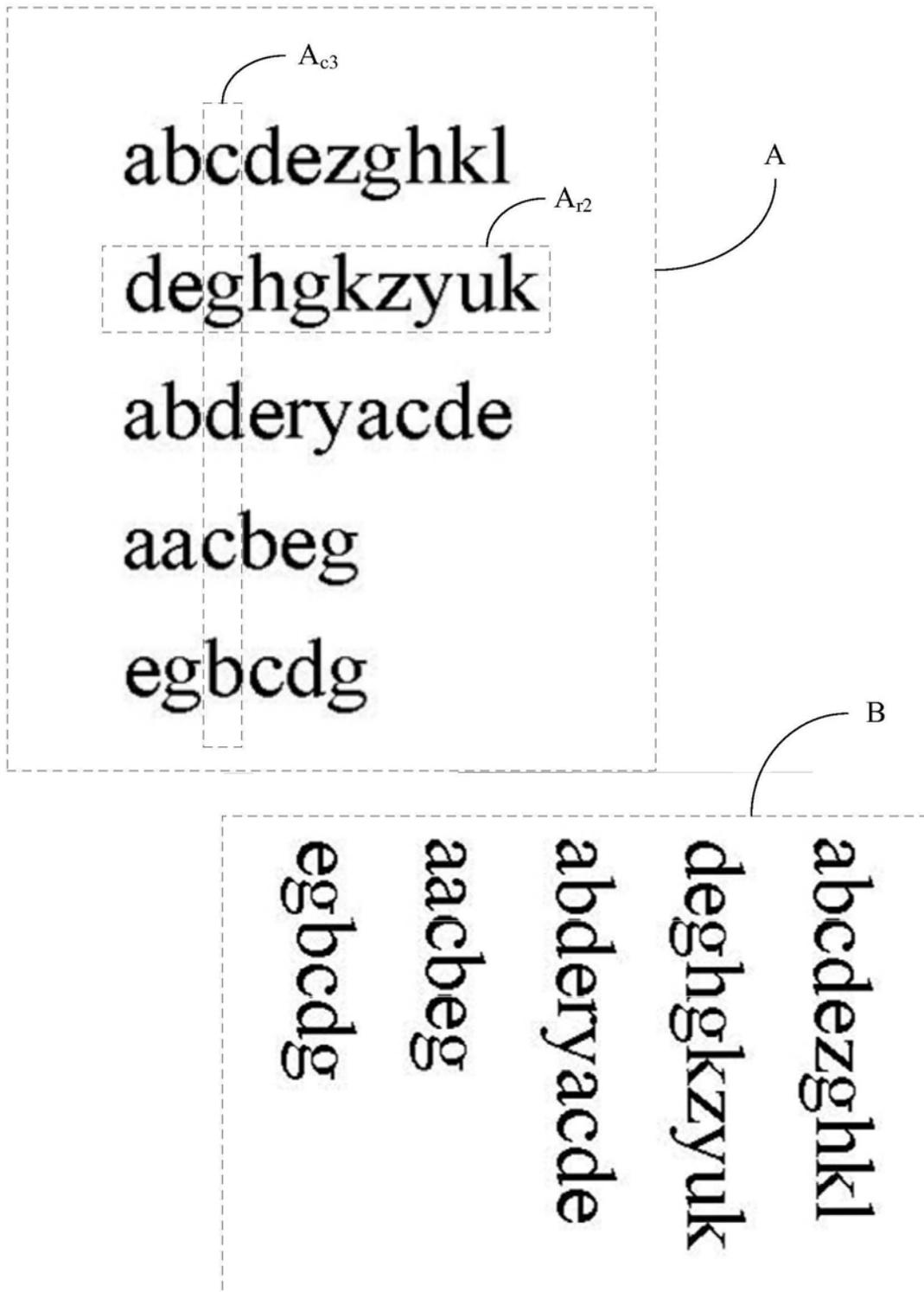


图2

100'

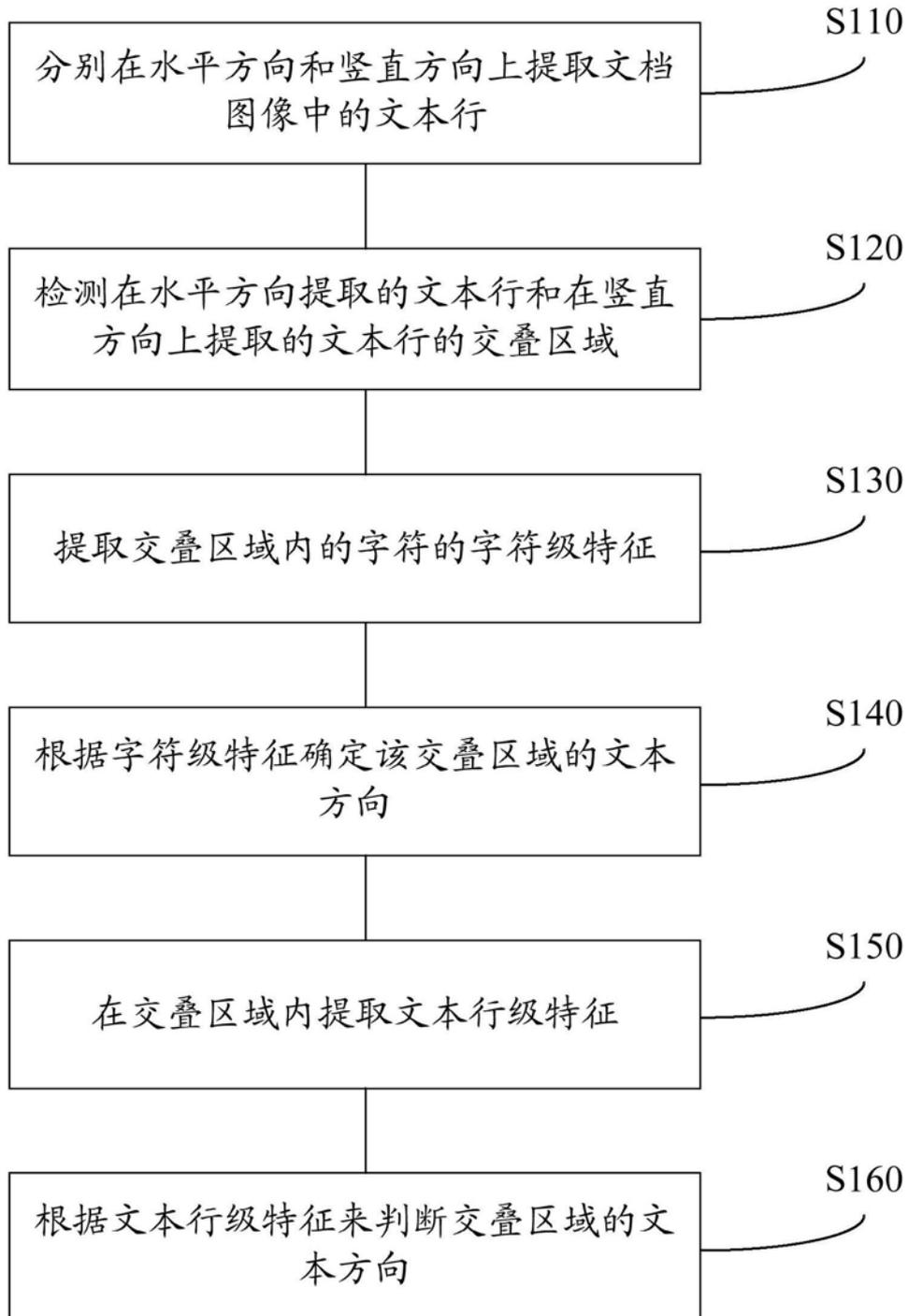


图3

100"

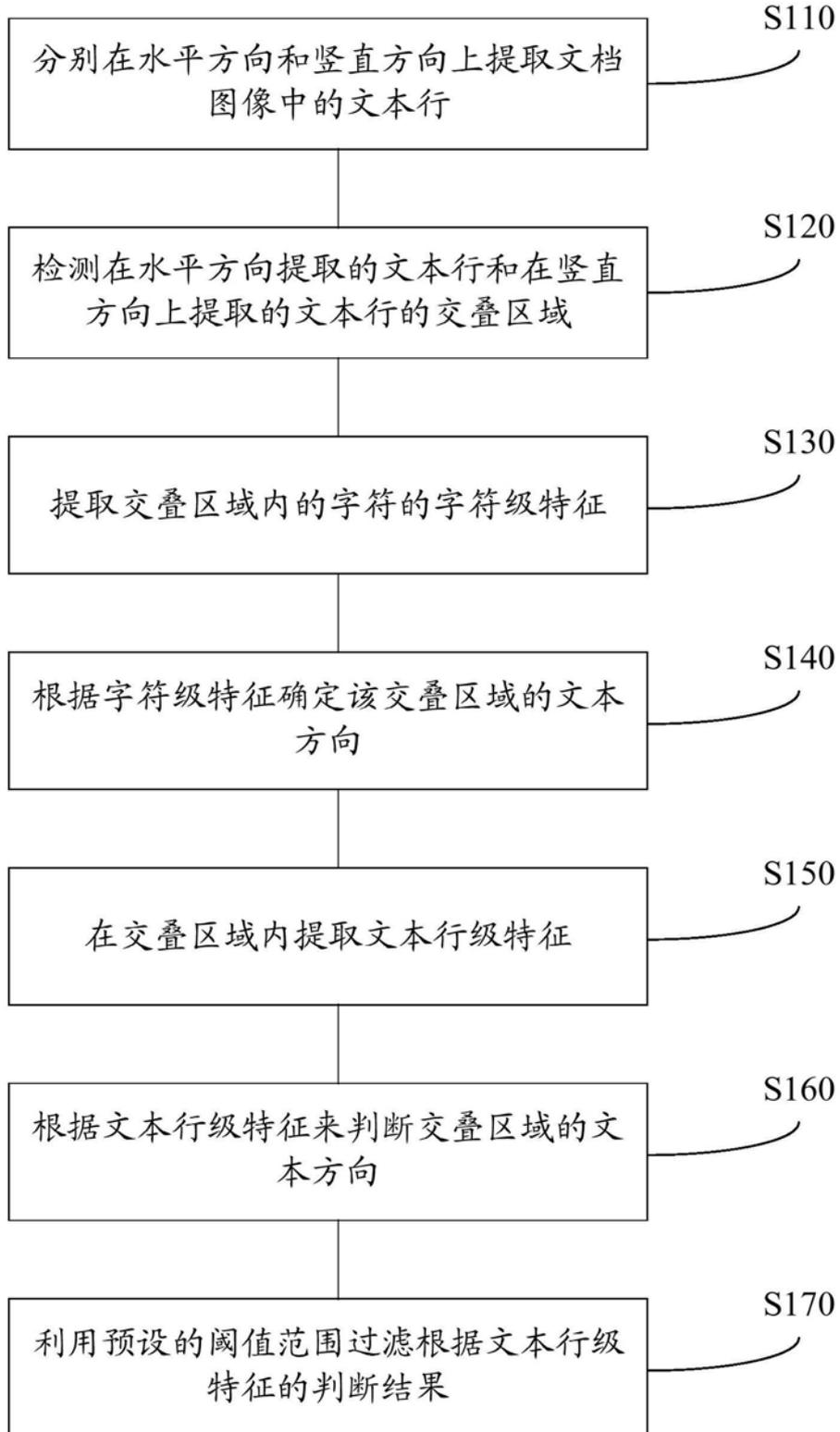


图4

S120

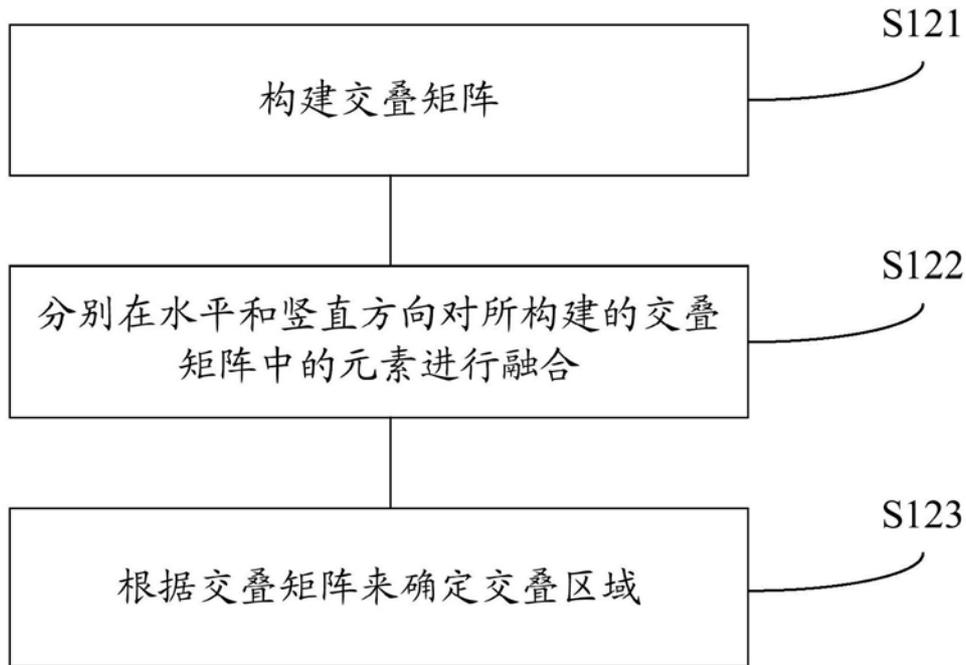


图5

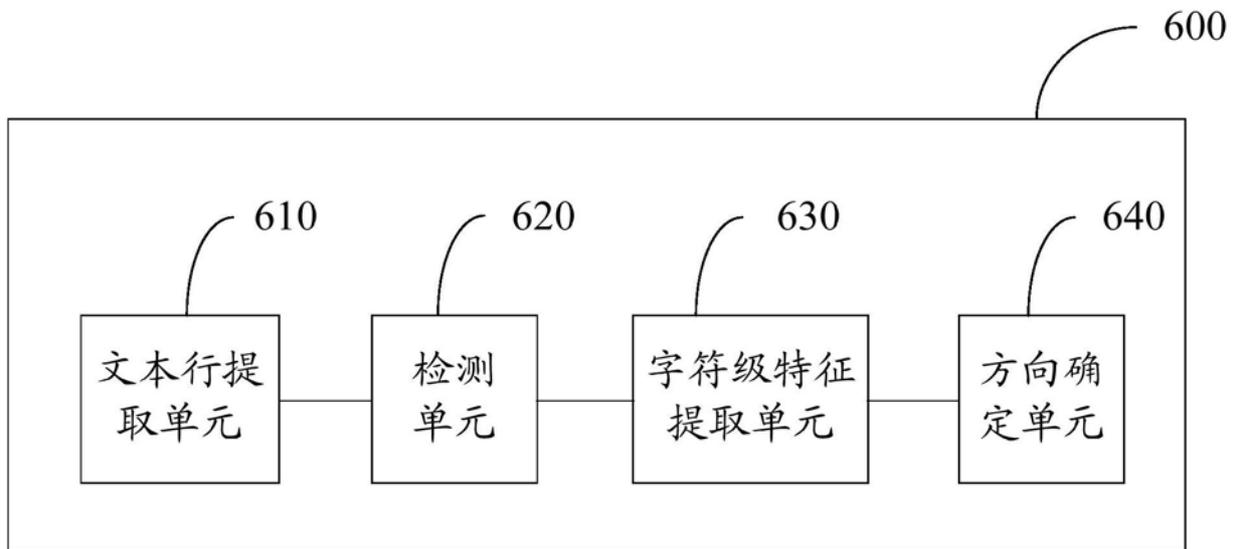


图6

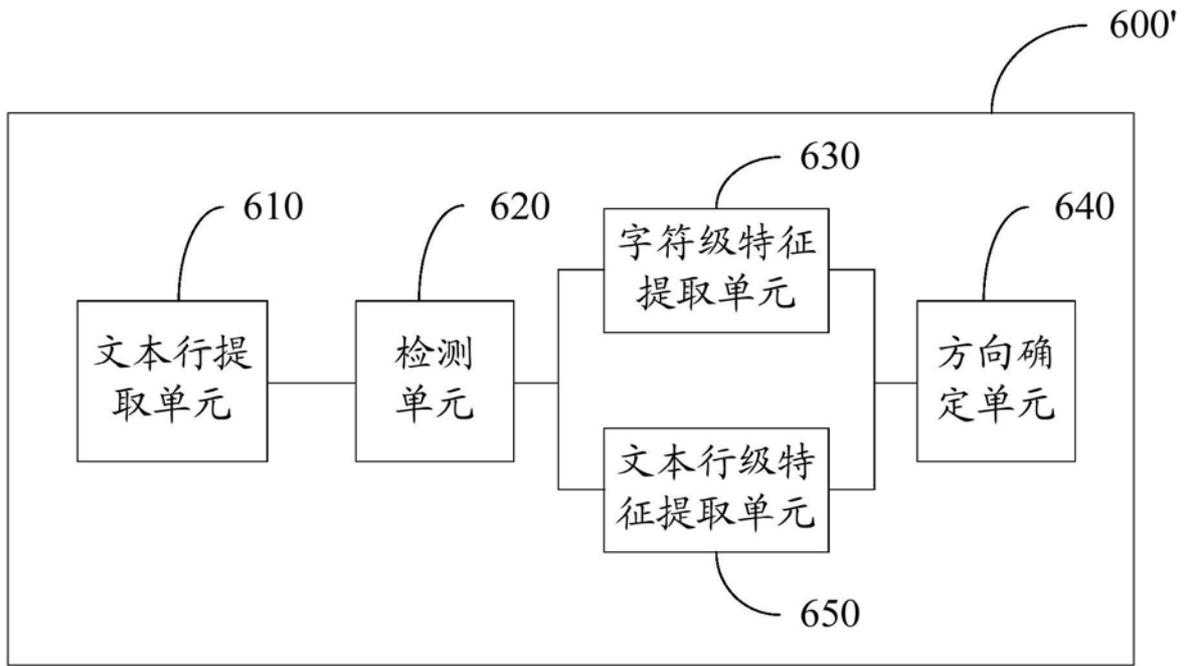


图7

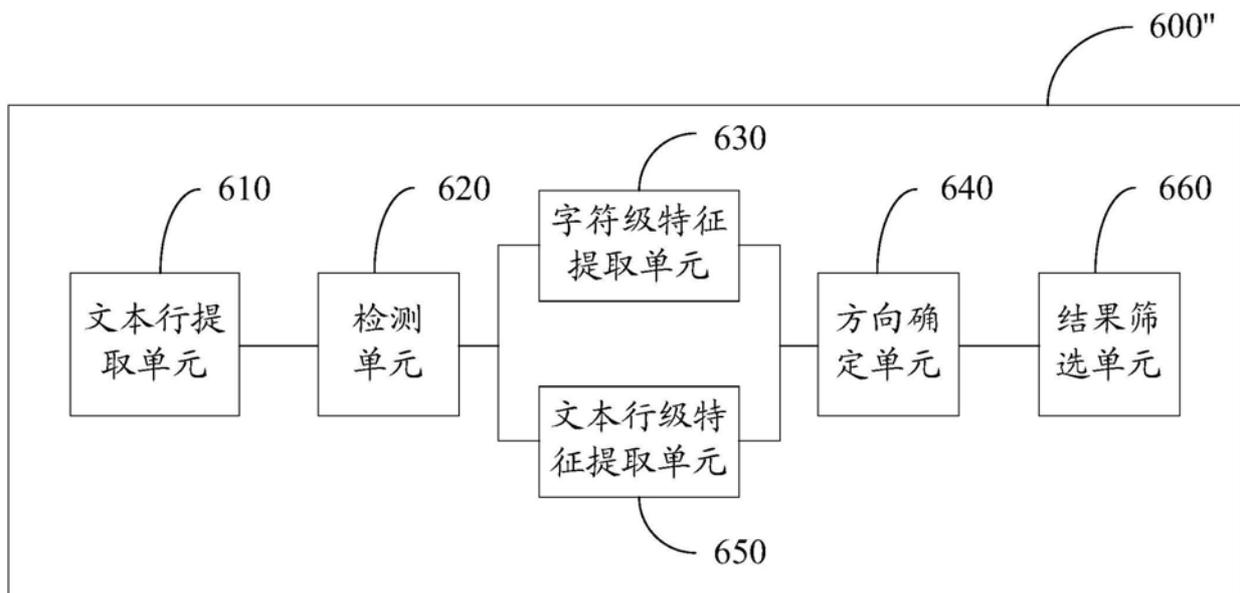


图8

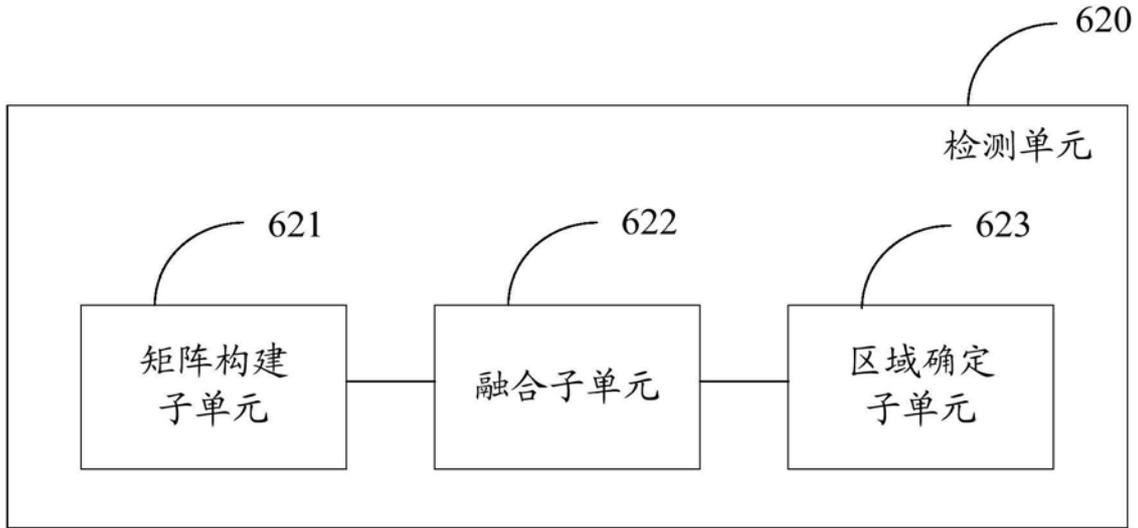


图9

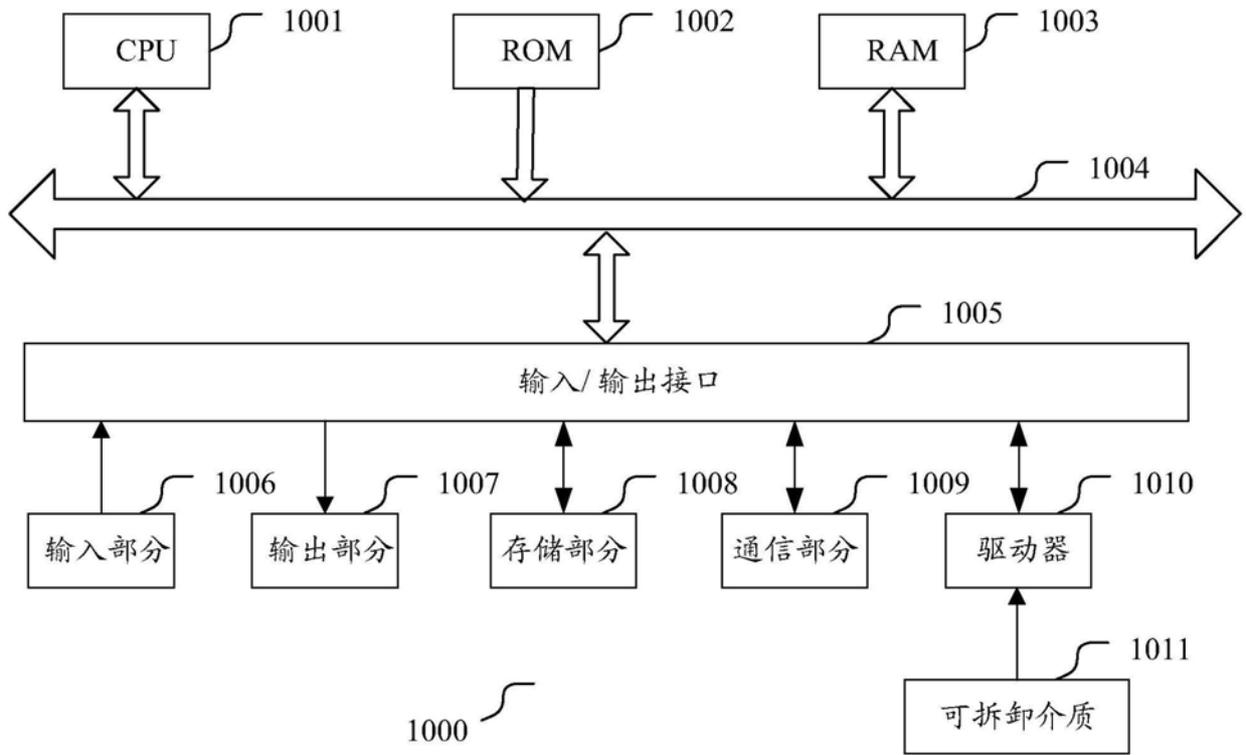


图10