



(12) 发明专利申请

(10) 申请公布号 CN 113191345 A

(43) 申请公布日 2021.07.30

(21) 申请号 202110468072.9

(22) 申请日 2021.04.28

(71) 申请人 北京有竹居网络技术有限公司
地址 101299 北京市平谷区林荫北街13号
信息大厦802

(72) 发明人 石威 林丽 黄灿 王长虎

(74) 专利代理机构 北京信远达知识产权代理有限公司 11304
代理人 姚莹丽

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

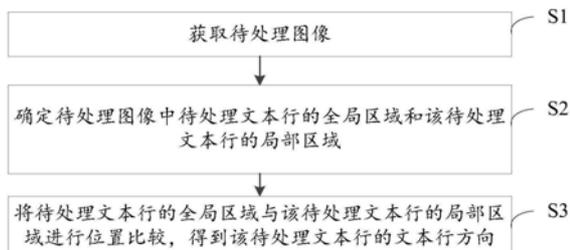
权利要求书3页 说明书24页 附图6页

(54) 发明名称

一种文本行方向确定方法及其相关设备

(57) 摘要

本申请公开了一种文本行方向确定方法及其相关设备,在获取到待处理图像之后,先确定该待处理图像中待处理文本行的全局区域和该待处理文本行的局部区域,以使该待处理文本行的全局区域表示该待处理文本行在该待处理图像内所占区域,以及该待处理文本行的局部区域表示该待处理文本行的预设部分在该待处理图像内所占区域;再将该待处理文本行的全局区域与该待处理文本行的局部区域进行位置比较,得到该待处理文本行的文本行方向。如此能够实现准确地确定出图像中文本行的文本行方向,从而有利于提高图像中文本行的字符识别准确性。



1. 一种文本行方向确定方法,其特征在于,所述方法包括:

获取待处理图像;

确定所述待处理图像中待处理文本行的全局区域和所述待处理文本行的局部区域;其中,所述待处理文本行的全局区域表示所述待处理文本行在所述待处理图像内所占区域;所述待处理文本行的局部区域表示所述待处理文本行的预设部分在所述待处理图像内所占区域;

将所述待处理文本行的全局区域与所述待处理文本行的局部区域进行位置比较,得到所述待处理文本行的文本行方向。

2. 根据权利要求1所述的方法,其特征在于,所述待处理文本行的全局区域和所述待处理文本行的局部区域的确定过程,包括:

将所述待处理图像输入预先构建的掩码图生成模型,得到所述掩码图生成模型输出的所述待处理图像对应的预测文本行全局掩码图和所述待处理图像对应的预测文本行局部掩码图;其中,所述掩码图生成模型是根据样本图像、所述样本图像对应的实际文本行全局掩码图和所述样本图像对应的实际文本行局部掩码图构建的;所述待处理图像对应的预测文本行全局掩码图包括所述待处理文本行对应的全局掩码区域;所述待处理图像对应的预测文本行局部掩码图包括所述待处理文本行对应的局部掩码区域;

将所述待处理图像对应的预测文本行全局掩码图中所述待处理文本行对应的全局掩码区域,确定为所述待处理文本行的全局区域;

将所述待处理图像对应的预测文本行局部掩码图中所述待处理文本行对应的局部掩码区域,确定为所述待处理文本行的局部区域。

3. 根据权利要求1所述的方法,其特征在于,所述待处理文本行的局部区域包括所述待处理文本行的首字符区域。

4. 根据权利要求3所述的方法,其特征在于,所述将所述待处理文本行的全局区域与所述待处理文本行的局部区域进行位置比较,得到所述待处理文本行的文本行方向,包括:

根据所述待处理文本行的全局区域中心和所述待处理文本行的首字符区域中心之间的相对位置信息,确定所述待处理文本行的文本行方向;其中,所述待处理文本行的全局区域中心表示所述待处理文本行的全局区域的中心位置;所述待处理文本行的首字符区域中心表示所述待处理文本行的首字符区域的中心位置。

5. 根据权利要求4所述的方法,其特征在于,所述文本行方向包括文本行旋转方向和/或文本行旋转角度;

所述待处理文本行的文本行旋转方向的确定过程,包括:

根据所述待处理文本行的全局区域中心和所述待处理文本行的首字符区域中心之间在第一预设方向上的相对关系,确定所述待处理文本行的第一旋转方向;根据所述待处理文本行的第一旋转方向,确定所述待处理文本行的文本行旋转方向;

所述待处理文本行的文本行旋转角度的确定过程,包括:

对所述待处理文本行的全局区域中心和所述待处理文本行的首字符区域中心进行有向连接,得到第一向量;将所述第一向量与第一预设向量之间的夹角,确定为所述待处理文本行的第一旋转角度;根据所述待处理文本行的第一旋转角度,确定所述待处理文本行的文本行旋转角度。

6. 根据权利要求1所述的方法,其特征在于,所述待处理文本行的局部区域包括所述待处理文本行的首字符区域和所述待处理文本行的字符串局部区域;其中,所述待处理文本行的字符串局部区域是由所述待处理文本行中各个字符的预设局部区域进行连接得到的。

7. 根据权利要求6所述的方法,其特征在于,所述将所述待处理文本行的全局区域与所述待处理文本行的局部区域进行位置比较,得到所述待处理文本行的文本行方向,包括:

根据所述待处理文本行的全局区域中心、所述待处理文本行的首字符区域中心和所述待处理文本行的字符串局部区域中心中至少两个之间的相对位置信息,确定所述待处理文本行的文本行方向;其中,所述待处理文本行的全局区域中心表示所述待处理文本行的全局区域的中心位置;所述待处理文本行的首字符区域中心表示所述待处理文本行的首字符区域的中心位置;所述待处理文本行的字符串局部区域中心表示所述待处理文本行的字符串局部区域的中心位置。

8. 根据权利要求7所述的方法,其特征在于,所述文本行方向包括文本行旋转方式和/或文字排列方式;

所述待处理文本行的文本行旋转方式的确定过程,包括:

根据所述待处理文本行的首字符区域中心和所述待处理文本行的字符串局部区域中心之间的相对位置信息,确定所述待处理文本行的文本行旋转方式;

所述待处理文本行的文字排列方式的确定过程,包括:

根据所述待处理文本行的全局区域中心和所述待处理文本行的字符串局部区域中心之间的相对位置信息,确定所述待处理文本行的文字排列方式。

9. 根据权利要求8所述的方法,其特征在于,所述文本行旋转方式包括文本行旋转方向和/或文本行旋转角度;

所述待处理文本行的文本行旋转方向的确定过程,包括:

根据所述待处理文本行的首字符区域中心和所述待处理文本行的字符串局部区域中心之间在第三预设方向上的相对关系以及位置差值,确定所述待处理文本行的第三旋转方向;根据所述待处理文本行的第三旋转方向,确定所述待处理文本行的文本行旋转方向;

所述待处理文本行的文本行旋转角度的确定过程,包括:

对所述待处理文本行的首字符区域中心和所述待处理文本行的字符串局部区域中心进行有向连接,得到第三向量;将所述第三向量与第三预设向量之间的夹角,确定为所述待处理文本行的第三旋转角度;根据所述待处理文本行的第三旋转角度,确定所述待处理文本行的文本行旋转角度。

10. 根据权利要求7所述的方法,其特征在于,若所述待处理文本行的文字排列方式为横排,则所述待处理文本行的字符串局部区域是由所述待处理文本行中各个字符的底部边界区域进行连接得到的;

若所述待处理文本行的文字排列方式为竖排,则所述待处理文本行的字符串局部区域是由所述待处理文本行中各个字符的中心区域进行连接得到的。

11. 根据权利要求10所述的方法,其特征在于,所述待处理文本行的文字排列方式的确定过程,包括:

若所述待处理文本行的全局区域中心与所述待处理文本行的字符串局部区域中心重合,则确定所述待处理文本行的文字排列方式为竖排;若所述待处理文本行的全局区域中

心与所述待处理文本行的字符串局部区域中心不重合,则确定所述待处理文本行的文字排列方式为横排。

12. 根据权利要求7所述的方法,其特征在于,所述文本行方向包括文本行旋转方向、文本行旋转角度、或文字排列方式中的至少一个;

所述待处理文本行的文本行方向的确定过程,包括:

根据所述待处理文本行的全局区域中心、所述待处理文本行的首字符区域中心和所述待处理文本行的字符串局部区域中心中至少两个之间的相对位置信息,确定所述待处理文本行的文本行旋转角度;

根据所述待处理文本行的全局区域中心、所述待处理文本行的首字符区域中心和所述待处理文本行的字符串局部区域中心中至少两个之间的相对位置信息,确定所述待处理文本行的文本行旋转方向;

根据所述待处理文本行的全局区域中心和所述待处理文本行的字符串局部区域中心之间的相对位置信息,确定所述待处理文本行的文字排列方式。

13. 一种文本行方向确定装置,其特征在于,所述装置包括:

图像获取单元,用于获取待处理图像;

区域确定单元,用于确定所述待处理图像中待处理文本行的全局区域和所述待处理文本行的局部区域;其中,所述待处理文本行的全局区域表示所述待处理文本行在所述待处理图像内所占区域;所述待处理文本行的局部区域表示所述待处理文本行的预设部分在所述待处理图像内所占区域;

方向确定单元,用于将所述待处理文本行的全局区域与所述待处理文本行的局部区域进行位置比较,得到所述待处理文本行的文本行方向。

14. 一种设备,其特征在于,所述设备包括处理器以及存储器:

所述存储器用于存储计算机程序;

所述处理器用于根据所述计算机程序执行权利要求1-12中任一项所述的方法。

15. 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用于存储计算机程序,所述计算机程序用于执行权利要求1-12中任一项所述的方法。

16. 一种计算机程序产品,其特征在于,所述计算机程序产品在终端设备上运行时,使得所述终端设备执行权利要求1-12中任一项所述的方法。

一种文本行方向确定方法及其相关设备

技术领域

[0001] 本申请涉及图像处理技术领域,尤其涉及一种文本行方向确定方法及其相关设备。

背景技术

[0002] 字符识别(如,光学字符识别)是指对图像中字符进行识别处理,以获取该图像中携带的文字信息。

[0003] 实际上,图像中文本行可以按照任一方向(例如,顺时针旋转 90° 等)进行呈现,故为了提高字符识别的准确性,可以在对图像中文本行(尤其是,具有一定旋转角度的文本行)进行字符识别时参考该文本行的文本行方向。其中,文本行方向是指一个文本行在图像中呈现的旋转信息。

[0004] 然而,如何确定文本行方向成为一项亟待解决的技术问题。

发明内容

[0005] 为了解决现有技术中存在的以上技术问题,本申请提供一种文本行方向确定方法及其相关设备,能够准确地确定出图像中文本行的文本行方向,如此有利于提高图像中文本行的字符识别准确性。

[0006] 为了实现上述目的,本申请实施例提供的技术方案如下:

[0007] 本申请实施例提供一种文本行方向确定方法,所述方法包括:获取待处理图像;确定所述待处理图像中待处理文本行的全局区域和所述待处理文本行的局部区域;其中,所述待处理文本行的全局区域表示所述待处理文本行在所述待处理图像内所占区域;所述待处理文本行的局部区域表示所述待处理文本行的预设部分在所述待处理图像内所占区域;将所述待处理文本行的全局区域与所述待处理文本行的局部区域进行位置比较,得到所述待处理文本行的文本行方向。

[0008] 本申请实施例还提供了一种文本行方向确定装置,所述装置包括:

[0009] 图像获取单元,用于获取待处理图像;

[0010] 区域确定单元,用于确定所述待处理图像中待处理文本行的全局区域和所述待处理文本行的局部区域;其中,所述待处理文本行的全局区域表示所述待处理文本行在所述待处理图像内所占区域;所述待处理文本行的局部区域表示所述待处理文本行的预设部分在所述待处理图像内所占区域;

[0011] 方向确定单元,用于将所述待处理文本行的全局区域与所述待处理文本行的局部区域进行位置比较,得到所述待处理文本行的文本行方向。

[0012] 本申请实施例还提供了一种设备,所述设备包括处理器以及存储器:

[0013] 所述存储器用于存储计算机程序;

[0014] 所述处理器用于根据所述计算机程序执行本申请实施例提供的文本行方向确定方法的任一实施方式。

[0015] 本申请实施例还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用于存储计算机程序,所述计算机程序用于执行本申请实施例提供的文本行方向确定方法的任一实施方式。

[0016] 本申请实施例还提供了一种计算机程序产品,所述计算机程序产品在终端设备上运行时,使得所述终端设备执行本申请实施例提供的文本行方向确定方法的任一实施方式。

[0017] 与现有技术相比,本申请实施例至少具有以下优点:

[0018] 本申请实施例提供的文本行方向确定方法及其相关设备中,在获取到待处理图像之后,先确定该待处理图像中待处理文本行的全局区域和该待处理文本行的局部区域,以使该待处理文本行的全局区域表示该待处理文本行在该待处理图像内所占区域,以及该待处理文本行的局部区域表示该待处理文本行的预设部分在该待处理图像内所占区域;再将该待处理文本行的全局区域与该待处理文本行的局部区域进行位置比较,得到该待处理文本行的文本行方向。

[0019] 其中,因待处理文本行的全局区域以及局部区域分别能够表示该待处理文本行在待处理图像内所占区域以及该待处理文本行的预设部分在待处理图像内所占区域,使得该待处理文本行的全局区域以及局部区域能够综合表示出该待处理文本行在待处理图像中的呈现方式,从而使得基于待处理文本行的全局区域以及局部区域确定出的该待处理文本行的文本行方向能够准确地描述出该待处理文本行在待处理图像中的呈现方式,如此能够实现准确地确定出图像中文本行的文本行方向,从而有利于提高图像中文本行的字符识别准确性。

附图说明

[0020] 为了更清楚地说明本申请实施例或现有技术中的技术方案,下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本申请中记载的一些实施例,对于本领域普通技术人员来讲,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。

[0021] 图1为本申请实施例提供了一种图像、其对应的全局区域以及其对应的局部区域的示意图;

[0022] 图2为本申请实施例提供了一种原图、其对应的全局掩码图以及局部掩码图的示意图;

[0023] 图3为本申请实施例提供了一种文本行方向确定方法的流程图;

[0024] 图4为本申请实施例提供了一种有向连接的示意图;

[0025] 图5为本申请实施例提供了一种横向文本行的旋转方式的特征示意图;

[0026] 图6为本申请实施例提供了一种竖直文本行的旋转方式的特征示意图;

[0027] 图7为本申请实施例提供的另一种横向文本行的旋转方式的特征示意图;

[0028] 图8为本申请实施例提供的另一种竖直文本行的旋转方式的特征示意图;

[0029] 图9为本申请实施例提供了一种文本行方向确定装置的结构示意图。

具体实施方式

[0030] 发明人在针对文本行方向的研究中发现,对于图像中的一个文本行来说,在该文本行在该图像中按照不同的文本行方向进行呈现时,该文本行在该图像内所占区域(如图1所示的全局区域)和该文本行的预设部分(如,首字符或者各个字符的预设局部区域)在该图像内所占区域(如图1所示的首字符区域或者字符串局部区域)将按照不同的相对关系进行呈现(如图2所示)。

[0031] 需要说明的是,在图2中,“文本行-1”是按照正方向(也就是旋转角度为 0°)进行呈现的横向文本行;“文本行-2”是按照逆时针旋转第一角度进行呈现的横向文本行,且该第一角度介于 0° 到 90° 之间;“文本行-3”是按照顺时针旋转 180° 或者按照逆时针旋转 180° 进行呈现的横向文本行;“文本行-4”是按照顺时针旋转 90° 进行呈现的横向文本行;“文本行-5”是按照逆时针旋转 90° 进行呈现的横向文本行;“文本行-6”是按照正方向(也就是旋转角度为 0°)进行呈现的竖直文本行;“文本行-7”是按照顺时针旋转 180° 或者按照逆时针旋转 180° 进行呈现的竖直文本行。其中,横向文本行的文字排列方式为横排;且竖直文本行的文字排列方式为竖排。

[0032] 基于上述发现,本申请实施例提供了一种文本行方向确定方法,该方法可以包括:获取待处理图像;确定该待处理图像中待处理文本行的全局区域和该待处理文本行的局部区域;将该待处理文本行的全局区域与该待处理文本行的局部区域进行位置比较,得到该待处理文本行的文本行方向。其中,因待处理文本行的全局区域以及局部区域分别能够表示该待处理文本行在该待处理图像内所占区域以及该待处理文本行的预设部分在该待处理图像内所占区域,使得该待处理文本行的全局区域以及局部区域能够综合表示出该待处理文本行在待处理图像中的呈现方式,从而使得基于待处理文本行的全局区域以及局部区域确定出的该待处理文本行的文本行方向能够准确地描述出该待处理文本行在待处理图像中的呈现方式,如此能够实现准确地确定出图像中文本行的文本行方向,从而有利于提高图像中文本行的字符识别准确性。

[0033] 另外,本申请实施例不限定文本行方向确定方法的执行主体,例如,本申请实施例提供的文本行方向确定方法可以应用于终端设备或服务器等数据处理设备。其中,终端设备可以为智能手机、计算机、个人数字助理(Personal Digital Assistant, PDA)或平板电脑等。服务器可以为独立服务器、集群服务器或云服务器。

[0034] 为了使本技术领域的人员更好地理解本申请方案,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,所描述的实施例仅是本申请一部分实施例,而不是全部的实施例。基于本申请中的实施例,本领域普通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。

[0035] 方法实施例

[0036] 参见图3,该图为本申请实施例提供了一种文本行方向确定方法的流程图。

[0037] 本申请实施例提供的文本行方向确定方法,包括S1-S3:

[0038] S1:获取待处理图像。

[0039] 待处理图像是指需要进行文本行方向确定处理的图像。其中,文本行方向用于描述一个文本行在图像中的呈现方式;而且,本申请实施例不限定文本行方向,例如,文本行方向可以包括文字排列方式和/或文本行旋转方式。

[0040] 文字排列方式用于描述一个文本行中所有字符的排列方式;而且本申请实施例不限定文字排列方式,例如,该文字排列方式可以是横排(如图2中“文本行-1”所示的文字排列方式),也可以是竖排(如图2中“文本行-6”所示的文字排列方式)。其中,横排用于表示一个文本行中所有字符的按照水平方向进行排列。竖排用于表示一个文本行中所有字符的按照垂直方向进行排列。

[0041] 文本行旋转方式用于描述一个文本行在图像中的旋转信息;而且,本申请实施例不限定文本行旋转方式,例如,文本行旋转方式可以包括旋转方向和/或旋转角度。

[0042] 另外,本申请实施例不限定待处理图像中文本行的文字排列方式,例如,对于待处理图像中第 n 个文本行来说,该第 n 个文本行的文字排列方式可以是横排,以使该第 n 个文本行属于横向文本行;或者,该第 n 个文本行的文字排列方式可以是竖排,以使该第 n 个文本行属于竖直文本行。此外,本申请实施例也不限定待处理图像中文本行的文本行旋转方式,例如,对于待处理图像中第 n 个文本行来说,该第 n 个文本行的文本行旋转方式可以是正方向(也就是旋转 0°);或者,该第 n 个文本行的文本行旋转方式可以是按照顺时针旋转第二角度,且该第二角度 $\in (0^\circ, 180^\circ)$;或者,该第 n 个文本行的文本行旋转方式可以是按照逆时针旋转第三角度,且该第三角度 $\in (0^\circ, 180^\circ)$;或者,该第 n 个文本行的文本行旋转方式可以是旋转 180° (也就是按照逆时针或者顺时针旋转 180°)。其中, n 为正整数, $n \leq N$, N 表示待处理图像中文本行的个数。

[0043] S2:确定待处理图像中待处理文本行的全局区域和该待处理文本行的局部区域。

[0044] 其中,待处理文本行用于表示待处理图像中任一文本行。例如,若待处理图像中存在 N 个文本行,则可以将待处理图像中第1个文本行、第2个文本行、……、第 N 个文本行分别确定为待处理文本行。

[0045] 待处理文本行的全局区域表示该待处理文本行在待处理图像内所占区域。

[0046] 待处理文本行的局部区域表示该待处理文本行的预设部分在待处理图像内所占区域。

[0047] 其中,预设部分可以根据应用场景预先设定。例如,预设部分可以包括首字符部分(如图2中第三掩码图内各个文本行的掩码区域对应的部分)。又如,横向文本行的预设部分可以包括该横向文本行的底部边界部分(如图2中第二掩码图内“文本行-1”的掩码区域对应的部分);而且,竖直文本行的预设部分可以包括该竖直文本行的中心部分(如图2中第二掩码图中“文本行-6”的掩码区域对应的部分)。

[0048] 基于上述“预设部分”的相关内容可知,局部区域可以包括首字符区域和/或字符串局部区域。其中,首字符区域表示图像中一个文本行的首字符在该图像内所占区域。字符串局部区域用于表示图像内一个文本行中各个字符的预设局部区域在该图像内所占区域。

[0049] 另外,预设局部区域可以预先设定,尤其可以根据一个文本行的文字排列方式进行设定。例如,若一个文本行的文字排列方式为横排,则该文本行中各个字符的预设局部区域可以是各个字符的底部边界区域;若一个文本行的文字排列方式为竖排,则该文本行中各个字符的预设局部区域可以是各个字符的中心区域(尤其是能够竖直贯穿字符的中心区域)。

[0050] 基于上述“局部区域”的相关内容可知,待处理文本行的局部区域可以包括该待处理文本行的首字符区域和/或该待处理文本行的字符串局部区域。其中,待处理文本行的首

字符区域是指该待处理文本行中首字符在待处理图像中所占区域。待处理文本行的字符串局部区域是由该待处理文本行中各个字符的预设局部区域进行连接得到的,且其具体可以包括:若待处理文本行的文字排列方式为横排,则该待处理文本行的字符串局部区域是由该待处理文本行中各个字符的底部边界区域进行连接得到的;若该待处理文本行的文字排列方式为竖排,则该待处理文本行的字符串局部区域是由该待处理文本行中各个字符的中心区域进行连接得到的。

[0051] 另外,本申请实施例不限定S2的实施方式,例如,在一种可能的实施方式下,S2具体可以包括S21-S23:

[0052] S21:将待处理图像输入预先构建的掩码图生成模型,得到该掩码图生成模型输出的该待处理图像对应的预测文本行全局掩码图和该待处理图像对应的预测文本行局部掩码图。其中,待处理图像对应的预测文本行全局掩码图包括待处理文本行对应的全局掩码区域;该待处理图像对应的预测文本行局部掩码图包括待处理文本行对应的局部掩码区域。

[0053] 掩码图生成模型用于对模型输入图像进行全局掩码图以及局部掩码图的预测。其中,“模型输入图像”是指输入至掩码图生成模型的图像(如图2所示的原图)。

[0054] 全局掩码图是针对模型输入图像中文本行在该模型输入图像内所占区域进行掩码处理后得到的,以使该全局掩码图能够表示模型输入图像中文本行在该模型输入图像内所占区域。

[0055] 局部掩码图是针对模型输入图像中文本行的预设部分在该模型输入图像内所占区域进行掩码处理后得到的,以使该局部掩码图能够表示模型输入图像中文本行的预设部分在该模型输入图像内所占区域。需要说明的是,“预设部分”的相关内容请参见上文。

[0056] 另外,本申请实施例不限定局部掩码图,例如,局部掩码图可以包括首字符掩码子图(如图2所示的第三掩码图)和/或字符串局部掩码子图(如图2所示的第二掩码图)。其中,首字符掩码子图是针对模型输入图像中文本行的首字符在该模型输入图像内所占区域进行掩码处理后得到的,以使该首字符掩码子图能够表示模型输入图像中文本行的首字符在该模型输入图像内所占区域。字符串局部掩码子图是针对模型输入图像中文本行的字符串局部部分在该模型输入图像内所占区域进行掩码处理后得到的,以使该字符串局部掩码子图表示模型输入图像中文本行的字符串局部部分在该模型输入图像内所占区域。

[0057] 需要说明的是,字符串局部部分可以预先设定,尤其可以根据一个文本行的文字排列方式进行设定。例如,横向文本行的字符串局部部分可以是该横向文本行的底部边界部分。又如,竖直文本行的字符串局部部分可以是该竖直文本行的中心区域(尤其是竖直贯穿该竖直文本行的中心区域)。

[0058] 另外,本申请实施例不限定掩码图生成模型的模型结构,例如,该掩码图生成模型可以采用任一种深度学习模型(如,基于深度学习的语义分割模型)的模型结构进行实施。

[0059] 此外,掩码图生成模型可以根据样本图像、该样本图像对应的实际文本行全局掩码图和该样本图像对应的实际文本行局部掩码图进行构建。其中,样本图像是指构建掩码图生成模型所需使用的图像;该样本图像对应的实际文本行全局掩码图用于表示该样本图像中各个文本行在该样本图像内实际所占区域;该样本图像对应的实际文本行局部掩码图用于表示该样本图像中各个文本行的预设部分在该样本图像内实际所占区域。

[0060] 另外,本申请实施例不限定掩码图生成模型的构建过程,例如,在一种可能的实施方式下,掩码图生成模型的构建过程具体可以包括步骤11-步骤15:

[0061] 步骤11:获取样本图像、该样本图像对应的实际文本行全局掩码图和该样本图像对应的实际文本行局部掩码图。

[0062] 步骤12:将样本图像输入待训练模型,得到该待训练模型输出的该样本图像对应的预测文本行全局掩码图和该样本图像对应的预测文本行局部掩码图。

[0063] 其中,待训练模型是指构建掩码图生成模型时需要进行训练的模型;而且,本申请实施例不限定待训练模型,例如,待训练模型可以是一种深度学习模型。

[0064] 样本图像对应的预测文本行全局掩码图用于表示该样本图像中至少一个文本行在该样本图像内预测所占区域。

[0065] 样本图像对应的预测文本行局部掩码图用于表示该样本图像中至少一个文本行的预设部分在该样本图像内预测所占区域。

[0066] 步骤13:判断是否达到预设停止条件,若是,则执行步骤15;若否,则执行步骤14。

[0067] 其中,预设停止条件可以预先设定;而且,本申请实施例不限定预设停止条件,例如,该预设停止条件可以是待训练模型的损失值低于第一阈值,也可以是待训练模型的损失值的变化率低于第二阈值,还可以是待训练模型的更新次数达到第三阈值。其中,第一阈值、第二阈值以及第三阈值均是预先设定的。

[0068] 待训练模型的损失值用于衡量该待训练模型的预测性能;而且该待训练模型的损失值可以根据预测数据(也就是,样本图像对应的预测文本行全局掩码图以及预测文本行局部掩码图)与标签数据(也就是,样本图像对应的实际文本行全局掩码图以及实际文本行局部掩码图)之间的差距进行确定。另外,本申请实施例不限定待训练模型的损失值的计算方式,例如,可以采用Dice系数、Dice Loss、或者Laplace smoothing进行实施。

[0069] 基于上述步骤13的相关内容可知,对于当前轮的待训练模型来说,可以判断当前轮的待训练模型是否达到预设停止条件,若达到预设停止条件,则表示当前轮的待训练模型具有较好的预测性能,故可以直接依据当前轮的待训练模型,构建掩码图生成模型,以使该掩码图生成模型也具有较好的预测性能;若未达到预设停止条件,则表示当前轮的待训练模型的预测性能依旧比较差,故可以依据样本图像对应的预测文本行全局掩码图、实际文本行全局掩码图、预测文本行局部掩码图以及实际文本行局部掩码图,对当前轮的待训练模型进行更新,以使更新后的待训练模型具有更好的预测性能。

[0070] 步骤14:根据样本图像对应的预测文本行全局掩码图、该样本图像对应的实际文本行全局掩码图、该样本图像对应的预测文本行局部掩码图和该样本图像对应的实际文本行局部掩码图,更新待训练模型,并返回执行步骤12。

[0071] 需要说明的是,本申请实施例不限定待训练模型的更新方式,可以采用现有的或者未来出现的任一种模型更新方式(例如,可以依据待训练模型的损失值进行模型更新)进行实施。

[0072] 步骤15:根据待训练模型,构建掩码图生成模型。

[0073] 本申请实施例中,在确定当前轮的待训练模型达到预设停止条件之后,可以根据当前轮的待训练模型构建掩码图生成模型(例如,将当前轮的待训练模型直接确定为掩码图生成模型;或者,根据当前轮的待训练模型的模型结构以及模型参数,确定掩码图生成模

型的模型结构以及模型参数,以使该掩码图生成模型与当前轮的待训练模型具有相同的模型结构以及模型参数),如此使得该掩码图生成模型与当前轮的待训练模型具有相同的预测性能,从而使得该掩码图生成模型也具有较好的预测性能。

[0074] 基于上述步骤11至步骤15的相关内容可知,在获取到样本图像、该样本图像对应的实际文本行全局掩码图以及实际文本行局部掩码图之后,可以将该样本图像作为构建掩码图生成模型时所需使用的模型输入数据,将该样本图像对应的实际文本行全局掩码图以及实际文本行局部掩码图作为构建掩码图生成模型时所需使用的模型标签数据,并基于该模型输入数据和该模型标签数据进行掩码图生成模型的构建,如此使得构建好的掩码图生成模型具有较好的预测性能,以便后续能够利用构建好的掩码图进行全局掩码图以及局部掩码图的预测。

[0075] 待处理图像对应的预测文本行全局掩码图是由构建好的掩码图生成模型针对待处理图像进行全局掩码图预测得到的,以使该预测文本行全局掩码图用于表示该待处理图像中至少一个文本行在该待处理图像内所占区域。可见,待处理图像对应的预测文本行全局掩码图可以包括该待处理图像中全部或部分文本行对应的全局掩码区域。其中,全局掩码区域用于表示一个文本行在该待处理图像内所占区域。

[0076] 待处理图像对应的预测文本行局部掩码图是由构建好的掩码图生成模型针对待处理图像进行局部掩码图预测得到的,以使该预测文本行局部掩码图用于表示该待处理图像中至少一个文本行的预设部分在该待处理图像内所占区域。可见,待处理图像对应的预测文本行局部掩码图可以包括该待处理图像中全部或部分文本行对应的局部掩码区域。其中,局部掩码区域用于表示一个文本行的预设部分在该待处理图像内所占区域。

[0077] 另外,本申请实施例不限定待处理图像对应的预测文本行局部掩码图,例如,其可以包括该待处理图像对应的首字符掩码子图和/或该待处理图像对应的字符串局部掩码子图。其中,待处理图像对应的首字符掩码子图用于表示该待处理图像中至少一个文本行的首字符在该模型输入图像内所占区域。待处理图像对应的字符串局部掩码子图用于表示该待处理图像中至少一个文本行的字符串局部部分在该模型输入图像内所占区域。

[0078] 基于上述S21的相关内容可知,在获取到待处理图像之后,可以将该待处理图像输入至预先构建好的掩码图生成模型,以使该掩码图生成模型能够针对该待处理图像进行全局掩码图以及局部掩码图的预测,得到并输出该待处理图像对应的预测文本行全局掩码图以及预测文本行局部掩码图,以使该预测文本行全局掩码图以及预测文本行局部掩码图能够准确地表示出该待处理图像中至少一个文本行及其预设部分在该待处理图像内所占区域。

[0079] S22:将待处理图像对应的预测文本行全局掩码图中待处理文本行对应的全局掩码区域,确定为该待处理文本行的全局区域。

[0080] 本申请实施例中,在获取到待处理图像对应的预测文本行全局掩码图之后,可以将该预测文本行全局掩码图中记录的该待处理文本行对应的全局掩码区域,确定为该待处理文本行的全局区域,以便后续能够基于该待处理文本行的全局区域进行文本行方向的确定。

[0081] S23:将待处理图像对应的预测文本行局部掩码图中待处理文本行对应的局部掩码区域,确定为该待处理文本行的局部区域。

[0082] 本申请实施例不限定待处理文本行的局部区域的确定过程(也就是,S23的实施方式),为了便于理解,下面结合三个示例进行说明。

[0083] 示例1,若待处理图像对应的预测文本行局部掩码图包括该待处理图像对应的首字符掩码子图,则S23具体可以包括步骤21-步骤22:

[0084] 步骤21:从待处理图像对应的首字符掩码子图中确定该待处理文本行的首字符区域。

[0085] 其中,待处理图像对应的首字符掩码子图用于表示该待处理图像中至少一个文本行的首字符在该待处理图像内所占区域。

[0086] 待处理文本行的首字符区域用于表示该待处理文本行的首字符在待处理图像内所占区域。

[0087] 步骤22:根据待处理文本行的首字符区域,确定该待处理文本行的局部区域,以使该待处理文本行的局部区域包括该待处理文本行的首字符区域。

[0088] 基于上述步骤21至步骤22的相关内容可知,对于待处理文本行来说,在获取到待处理图像对应的首字符掩码子图之后,可以先从该首字符掩码子图中确定出该待处理文本行的首字符区域,以使该首字符区域能够准确地表示出待处理文本行的首字符在该待处理图像内所占区域;再依据该首字符区域,确定该待处理文本行的局部区域,以使该局部区域也能够准确地表示出待处理文本行的首字符在该待处理图像内所占区域。

[0089] 示例2,若待处理图像对应的预测文本行局部掩码图包括该待处理图像对应的字符串局部掩码子图,则S23具体可以包括步骤31-步骤32:

[0090] 步骤31:从待处理图像对应的字符串局部掩码子图中确定待处理文本行的字符串局部区域。

[0091] 其中,待处理图像对应的字符串局部掩码子图用于表示该待处理图像中至少一个文本行的字符串局部部分在该待处理图像内所占区域。

[0092] 待处理文本行的字符串局部区域用于表示该待处理文本行的字符串局部部分在待处理图像内所占区域。

[0093] 步骤32:根据待处理文本行的字符串局部区域,确定待处理文本行的局部区域,以使待处理文本行的局部区域包括待处理文本行的局部区域。

[0094] 基于上述步骤31至步骤32的相关内容可知,对于待处理文本行来说,在获取到待处理图像对应的字符串局部掩码子图之后,可以先从该字符串局部掩码子图中确定出该待处理文本行的字符串局部区域,以使该字符串局部区域能够准确地表示出该待处理文本行的字符串局部部分在待处理图像内所占区域;再依据该字符串局部区域,确定该待处理文本行的局部区域,以使该局部区域也能够准确地表示出该待处理文本行的字符串局部部分在待处理图像内所占区域。

[0095] 示例3,若待处理图像对应的预测文本行局部掩码图包括该待处理图像对应的首字符掩码子图和该待处理图像对应的字符串局部掩码子图,则S23具体可以包括步骤41-步骤43:

[0096] 步骤41:从待处理图像对应的首字符掩码子图中确定该待处理文本行的首字符区域。

[0097] 步骤42:从待处理图像对应的字符串局部掩码子图中确定该待处理文本行的字符

串局部区域。

[0098] 需要说明的是,步骤41和步骤42的相关内容请分别参见上文步骤21和步骤31。

[0099] 步骤43:根据待处理文本行的首字符区域和该待处理文本行的字符串局部区域,确定该待处理文本行的局部区域,以使该待处理文本行的局部区域包括该待处理文本行的首字符区域和该待处理文本行的字符串局部区域。

[0100] 基于上述步骤41至步骤43的相关内容可知,对于待处理文本行来说,在获取到待处理图像对应的首字符掩码子图以及字符串局部掩码子图之后,先分别从该首字符掩码子图以及字符串局部掩码子图中分别确定出该待处理文本行的首字符区域以及字符串局部区域,以使该首字符区域以及字符串局部区域能够准确地表示出该待处理文本行及其字符串局部部分分别在待处理图像内所占区域;再依据该首字符区域以及字符串局部区域,确定该待处理文本行的局部区域,以使该局部区域也能够准确地表示出该待处理文本行及其字符串局部部分分别在待处理图像内所占区域。

[0101] 基于上述S2的相关内容可知,在获取到待处理图像之后,可以从该待处理图像中确定出各个待处理文本行的全局区域及其局部区域,以便后续能够依据该各个待处理文本行的全局区域及其局部区域,分别确定出各个待处理文本行的文本行方向。

[0102] S3:将待处理文本行的全局区域与该待处理文本行的局部区域进行位置比较,得到该待处理文本行的文本行方向。

[0103] 其中,待处理文本行的文本行方向用于表示该待处理文本行在待处理图像中的呈现方式。

[0104] 另外,本申请实施例不限定待处理文本行的文本行方向,例如,该待处理文本行的文本行方向可以包括该待处理文本行的文本行旋转方式和/或该待处理文本行的文字排列方式;而且该待处理文本行的文本行旋转方式可以包括该待处理文本行的文本行旋转方向和/或该待处理文本行的文本行旋转角度。

[0105] 本申请实施例也不限定上述“位置比较”的实施方式,例如,可以将待处理文本行的全局区域与该待处理文本行的局部区域之间的相对信息与至少一个候选标准相对信息进行匹配,并将匹配成功的候选标准相对信息对应的文本行方向,确定为该待处理文本行的文本行方向。其中,上文“相对信息”可以包括不同区域之间的相对参数(例如,在位置上的相对关系以及相对差值,在区域范围上的相对关系以及相对差值等)。另外,各个候选标准相对信息可以预先设定,而且本申请实施例不限定候选标准相对信息的个数。

[0106] 另外,为了进一步提高文本行方向的准确性,本申请实施例还提供了S3的一种可能的具体实施方式,其具体可以包括S31-S33:

[0107] S31:确定该待处理文本行的全局区域中心,以使该待处理文本行的全局区域中心用于表示该待处理文本行的全局区域的中心位置。

[0108] 本申请实施例中,在获取到该待处理文本行的全局区域之后,可以根据该全局区域的中心位置,确定该待处理文本行的全局区域中心,以使该待处理文本行的全局区域中心能够准确地表示出该待处理文本行在待处理图像内所占区域的中心位置。

[0109] S32:确定该待处理文本行的局部区域中心,以使该待处理文本行的局部区域中心用于表示该待处理文本行的局部区域的中心位置。

[0110] 本申请实施例不限定S32的实施方式,为了便于理解,下面结合三个示例进行说

明。

[0111] 示例一,若待处理文本行的局部区域包括该待处理文本行的首字符区域,则S32具体可以包括步骤51-步骤52:

[0112] 步骤51:从待处理文本行的首字符区域中确定该待处理文本行的首字符区域中心,以使该待处理文本行的首字符区域中心用于表示该待处理文本行的首字符区域的中心位置。

[0113] 本申请实施例中,在获取到待处理文本行的首字符区域之后,可以将该首字符区域的中心位置确定为该待处理文本行的首字符区域中心,以使该首字符区域中心能够准确地表示出该待处理文本行的首字符在待处理图像内所占区域的中心位置。

[0114] 步骤52:根据待处理文本行的首字符区域中心,确定该待处理文本行的局部区域中心,以使该待处理文本行的局部区域中心包括待处理文本行的首字符区域中心。

[0115] 基于上述步骤51至步骤52的相关内容可知,在获取到待处理文本行的首字符区域之后,可以先从该首字符区域中确定出该待处理文本行的首字符区域中心,以使该首字符区域中心能够准确地表示出该待处理文本行的首字符在待处理图像内所占区域的中心位置;再根据该首字符区域中心,确定待处理文本行的局部区域中心,以使该局部区域中心也能够准确地表示出该待处理文本行的首字符在待处理图像内所占区域的中心位置。

[0116] 示例二,若待处理文本行的局部区域包括该待处理文本行的字符串局部区域,则S32具体可以包括步骤61-步骤62:

[0117] 步骤61:从待处理文本行的字符串局部区域中确定该待处理文本行的字符串局部区域中心,以使该待处理文本行的字符串局部区域中心用于表示该待处理文本行的字符串局部区域的中心位置。

[0118] 本申请实施例中,在获取到待处理文本行的字符串局部区域之后,可以将该字符串局部区域的中心位置确定为该待处理文本行的字符串局部区域中心,以使该字符串局部区域中心能够准确地表示出该待处理文本行的字符串局部部分在待处理图像内所占区域的中心位置。

[0119] 步骤62:根据待处理文本行的字符串局部区域中心,确定该待处理文本行的局部区域中心,以使该待处理文本行的局部区域中心包括该待处理文本行的字符串局部区域中心。

[0120] 基于上述步骤61至步骤62的相关内容可知,在获取到待处理文本行的字符串局部区域之后,可以先从该字符串局部区域中确定出待处理文本行的字符串局部区域中心,以使该字符串局部区域中心能够准确地表示出该待处理文本行的字符串局部部分在待处理图像内所占区域的中心位置;再根据该字符串局部区域中心,确定该待处理文本行的局部区域中心,以使该局部区域中心也能够准确地表示出该待处理文本行的字符串局部部分在待处理图像内所占区域的中心位置。

[0121] 示例三,若待处理文本行的局部区域包括该待处理文本行的首字符区域和该待处理文本行的字符串局部区域,则S32具体可以包括步骤71-步骤73:

[0122] 步骤71:从待处理文本行的首字符区域中确定该待处理文本行的首字符区域中心,以使该待处理文本行的首字符区域中心用于表示该待处理文本行的首字符区域的中心位置。

[0123] 步骤72:从待处理文本行的字符串局部区域中确定该待处理文本行的字符串局部区域中心,以使该待处理文本行的字符串局部区域中心用于表示该待处理文本行的字符串局部区域的中心位置。

[0124] 需要说明的是,步骤71和步骤72的相关内容请分别参见上文步骤51和步骤61。

[0125] 步骤73:根据待处理文本行的首字符区域中心和该待处理文本行的字符串局部区域中心,确定该待处理文本行的局部区域中心,以使该待处理文本行的局部区域中心包括该待处理文本行的首字符区域中心和该待处理文本行的字符串局部区域中心。

[0126] 基于上述步骤71至步骤73的相关内容可知,在获取到待处理文本行的首字符区域以及字符串局部区域之后,可以分别从该首字符区域以及字符串局部区域中确定出该待处理文本行的首字符区域中心以及字符串局部区域中心;再将该首字符区域中心以及字符串局部区域中心的集合,确定为待处理文本行的局部区域中心,以使该局部区域中心能够准确地表示出待处理文本行的首字符以及字符串局部部分在待处理图像内所占区域的中心位置。

[0127] 基于上述S32的相关内容可知,在获取到待处理文本行的局部区域之后,可以从该局部区域中确定出该待处理文本行的局部区域中心,以使该局部区域中心能够准确地表示出该待处理文本行的预设部分(如首字符和/或字符串局部部分)在待处理图像内所占区域的中心位置。

[0128] S33:根据待处理文本行的全局区域中心与该待处理文本行的局部区域中心之间的相对位置信息,确定该待处理文本行的文本行方向。

[0129] 其中,相对位置信息用于描述不同对象(例如,全局区域中心与局部区域中心)在位置(例如,水平方向、垂直方向等)上的相对关系以及相对差值。

[0130] 另外,本申请实施例不限定S33的实施方式,为了便于理解S33,下面结合两个示例进行说明。

[0131] 在S33的第一个示例下,若待处理文本行的文本行方向包括该待处理文本行的文本行旋转方式,且该待处理文本行的局部区域中心包括该待处理文本行的首字符区域中心,则S331具体可以包括:根据待处理文本行的全局区域中心和该待处理文本行的首字符区域中心之间的相对位置信息,确定该待处理文本行的文本行方向。

[0132] 另外,本申请实施例不限定上述确定过程的实施方式,例如,在一种可能的实施方式中,若文本行方向包括文本行旋转方向和/或文本行旋转角度,则该待处理文本行的文本行方向的确定过程可以包括步骤81-步骤84中的全部或者部分步骤:

[0133] 步骤81:对待处理文本行的全局区域中心和该待处理文本行的首字符区域中心进行有向连接,得到第一向量。

[0134] 需要说明的是,本申请实施例不限定步骤81中“有向连接”的连接方式,例如,可以将待处理文本行的全局区域中心作为起点,并将该待处理文本行的首字符区域中心作为终点进行有向连接,以使第一向量成为一个以待处理文本行的全局区域中心为起点且以该待处理文本行的首字符区域中心为终点的有向线段。又如,可以将待处理文本行的首字符区域中心作为起点,并将该待处理文本行的全局区域中心作为终点进行有向连接,以使第一向量成为一个以待处理文本行的首字符区域中心为起点且以该待处理文本行的全局区域中心为终点的有向线段。

[0135] 步骤82:将第一向量与第一预设向量之间的夹角,确定为待处理文本行的第一旋转角度。

[0136] 其中,第一预设向量是指预先设定的用于表示全局区域中心和首字符区域中心之间有向连接的参考向量。

[0137] 实际上,具有不同文字排列方式的文本行对应于不同的参考向量。基于此,本申请实施例还提供了确定第一预设向量的一种可能的实施方式,其具体可以包括:根据待处理文本行的文字排列方式和第一映射关系,确定第一预设向量。其中,第一映射关系包括待处理文本行的文字排列方式与第一预设向量之间的对应关系。

[0138] 其中,第一映射关系用于记录不同文字排列方式对应的用于表示全局区域中心与首字符区域中心之间有向连接的参考向量;而且本申请实施例不限定第一映射关系,例如,第一映射关系可以包括横排与第一标准向量之间的对应关系、以及竖排与第二标准向量之间的对应关系。

[0139] 第一标准向量是指用于表示横向文本行的全局区域中心与首字符区域中心之间有向连接的标准向量;而且本申请实施例不限定第一标准向量,例如,第一标准向量可以根据标准横向文本行的全局区域中心和该标准横向文本行的首字符区域中心之间的有向连接确定(例如,如图4所示,可以将标准横向文本行的全局区域中心和该标准横向文本行的首字符区域中心之间的有向连接确定为第一标准向量。又如,可以将标准横向文本行的全局区域中心和该标准横向文本行的首字符区域中心之间的有向连接的任意一个平行向量,确定为第一标准向量)。其中,标准横向文本行的文字方向为水平方向,且该标准横向文本行的文本行方向为正方向(也就是,旋转 0°)。

[0140] 第二标准向量是指用于表示竖直文本行的全局区域中心与首字符区域中心之间有向连接的标准向量;而且本申请实施例不限定第二标准向量,例如,第二标准向量可以根据标准竖直文本行的全局区域中心和该标准竖直文本行的首字符区域中心之间的有向连接确定(例如,如图4所示,可以将标准竖直文本行的全局区域中心和该标准竖直文本行的首字符区域中心之间的有向连接确定为第二标准向量。又如,可以将标准竖直文本行的全局区域中心和该标准竖直文本行的首字符区域中心之间的有向连接的任意一个平行向量,确定为第二标准向量)。其中,标准竖直文本行的文字方向为垂直方向,且该标准竖直文本行的文本行方向为正方向(也就是,旋转 0°)。

[0141] 基于上述“第一预设向量”的相关内容可知,对于待处理文本来说,若确定该待处理文本行的文字排列方式为横排,则可以将第一标准向量确定为第一预设向量;若确定该待处理文本行的文字排列方式为竖排,则可以将第二标准向量确定为第一预设向量。

[0142] 另外,第一旋转角度可以用于表示待处理文本行的文本行旋转角度。

[0143] 基于上述步骤82的相关内容可知,对于待处理文本来说,在获取到第一向量之后,可以将第一向量与与第一预设向量之间的夹角,确定为该待处理文本行的第一旋转角度,以使该第一旋转角度能够准确地表示该待处理文本行的文本行旋转角度。

[0144] 步骤83:根据待处理文本行的全局区域中心和该待处理文本行的首字符区域中心之间在第一预设方向上的相对关系,确定该待处理文本行的第一旋转方向。

[0145] 其中,第一预设方向是指预先设定的参考方向。

[0146] 实际上,具有不同文字排列方式的文本行对应于不同的参考方向。基于此,本申请

实施例还提供了确定第一预设方向的一种可能的实施方式,其具体可以包括:在获取到待处理文本行的文字排列方式之后,根据该待处理文本行的文字排列方式和第二映射关系,确定第一预设方向。其中,第二映射关系包括待处理文本行的文字排列方式与第一预设方向之间的对应关系。

[0147] 第二映射关系用于记录不同文字排列方式对应的与全局区域中心与首字符区域中心相关的参考方向;而且本申请实施例不限定第二映射关系,例如,第二映射关系可以包括横排与垂直方向之间的对应关系、以及竖排与水平方向之间的对应关系。

[0148] 基于上述“第一预设方向”的相关内容可知,对于待处理文本行来说,若确定该待处理文本行的文字排列方式为横排,则可以将垂直方向确定为第一预设方向;若确定该待处理文本行的文字排列方式为竖排,则可以将水平方向确定为第一预设方向。

[0149] 另外,第一旋转方向可以用于描述待处理文本行的文本行旋转方向。

[0150] 此外,本申请实施例不限定步骤83的实施方式,为了便于理解,下面结合两种场景进行说明。

[0151] 场景一:当待处理文本行的文字排列方式为横排,且第一预设方向为垂直方向时,步骤83具体可以包括:若该待处理文本行的全局区域中心在第一预设方向上的投影位置高于该待处理文本行的首字符区域中心在第一预设方向上的投影位置,则确定待处理文本行的第一旋转方向为逆时针方向;若待处理文本行的全局区域中心在第一预设方向上的投影位置低于该待处理文本行的首字符区域中心在第一预设方向上的投影位置,则确定待处理文本行的第一旋转方向为顺时针方向。

[0152] 可见,如图5所示,若待处理文本行的文字排列方式为横排(也就是,该待处理文本行属于横向文本行),则在获取到该待处理文本行的全局区域中心(如图5中“01”)与该待处理文本行的首字符区域中心(如图5中“03”)之间在垂直方向的相对关系之后,如果该相关关系表示该待处理文本行的全局区域中心在垂直方向上的投影位置高于该待处理文本行的首字符区域中心在垂直方向上的投影位置,则可以确定该待处理文本行的第一旋转方向为逆时针方向;如果该相关关系表示该待处理文本行的全局区域中心在垂直方向上的投影位置低于该待处理文本行的首字符区域中心在垂直方向上的投影位置,则可以确定该待处理文本行的第一旋转方向为顺时针方向。

[0153] 需要说明的是,若待处理文本行的文字排列方式为横排,且该待处理文本行的全局区域中心在垂直方向上的投影位置与该待处理文本行的首字符区域中心在垂直方向上的投影位置重合,则可以确定该待处理文本行处于正方向或者该待处理文本行旋转 180° 。

[0154] 场景二:当待处理文本行的文字排列方式为竖排,且第一预设方向为水平方向时,步骤83具体可以包括:若待处理文本行的全局区域中心在第一预设方向上的投影位置比该待处理文本行的首字符区域中心在第一预设方向上的投影位置偏右,则确定该待处理文本行的第一旋转方向为逆时针方向;若待处理文本行的全局区域中心在第一预设方向上的投影位置比该待处理文本行的首字符区域中心在第一预设方向上的投影位置偏左,则确定该待处理文本行的第一旋转方向为顺时针方向。

[0155] 可见,如图6所示,若待处理文本行的文字排列方式为竖排(也就是,该待处理文本行属于竖直文本行),则在获取到该待处理文本行的全局区域中心(如图6中“01”)与该待处理文本行的首字符区域中心(如图6中“03”)之间在水平方向的相对关系之后,如果该相关

关系表示该待处理文本行的全局区域中心在水平方向上的投影位置比该待处理文本行的首字符区域中心在水平方向上的投影位置偏右,则可以确定该待处理文本行的第一旋转方向为逆时针方向;如果该相关关系表示待处理文本行的全局区域中心在水平方向上的投影位置比该待处理文本行的首字符区域中心在水平方向上的投影位置偏左,则可以确定该待处理文本行的第一旋转方向为顺时针方向。

[0156] 需要说明的是,若待处理文本行的文字排列方式为竖排,且待处理文本行的全局区域中心在水平方向上的投影位置与该待处理文本行的首字符区域中心在水平方向上的投影位置重合,则可以确定该待处理文本行处于正方向或者该待处理文本行旋转 180° 。

[0157] 基于上述步骤83的相关内容可知,在获取到待处理文本行的全局区域中心和该待处理文本行的首字符区域中心之后,可以先获取待处理文本行的全局区域中心和该待处理文本行的首字符区域中心之间在第一预设方向上的相对关系;再根据该相关关系,确定该待处理文本行的第一旋转方向,以使该第一旋转方向能够准确地表示出该待处理文本的文本行旋转方向。

[0158] 步骤84:根据待处理文本行的第一旋转角度和/或该待处理文本行的第一旋转方向,生成该待处理文本行的文本行方向。

[0159] 作为示例,若文本行方向包括文本行旋转方向和/或文本行旋转角度,则步骤84可以包括步骤841和/或步骤842:

[0160] 步骤841:根据待处理文本行的第一旋转方向,确定该待处理文本行的文本行旋转方向(如,可以直接将待处理文本行的第一旋转方向,确定为该待处理文本行的文本行旋转方向)。

[0161] 步骤842:根据待处理文本行的第一旋转角度,确定该待处理文本行的文本行旋转角度(如,可以直接将待处理文本行的第一旋转角度,确定为该待处理文本行的文本行旋转角度)。

[0162] 基于上述步骤84的相关内容可知,在获取到该待处理文本行的第一旋转角度和/或第一旋转方向之后,可以依据该第一旋转角度和/或第一旋转方向,生成该待处理文本行的文本行方向(尤其是,文本行旋转方式),以使该文本行方向能够准确地表示出该待处理文本行的文本行旋转方式,从而使得该文本行方向能够准确地表示出该待处理文本行在待处理图像中的呈现方式。

[0163] 基于上述S33的第一个示例的相关内容可知,对于待处理文本行来说,在获取到该待处理文本行的首字符区域中心之后,可以依据该待处理文本行的全局区域中心以及首字符区域中心之间的相对位置信息,确定该待处理文本行的文本行方向,以使该文本行方向能够准确地表示出该待处理文本行在待处理图像中的呈现方式。

[0164] 在S33的第二个示例下,若待处理文本行的局部区域中心包括该待处理文本行的首字符区域中心和该待处理文本行的字符串局部区域中心,则S33具体可以包括:根据待处理文本行的全局区域中心、该待处理文本行的首字符区域中心和该待处理文本行的字符串局部区域中心中至少两个之间的相对位置信息,确定该待处理文本行的文本行方向。

[0165] 本申请实施例不限定上述确定过程,例如,在一种可能的实施方式下,若文本行方向包括文本行旋转方式和/或文字排列方式,则S33具体可以包括S331和/或S332:

[0166] S331:根据待处理文本行的全局区域中心和该待处理文本行的字符串局部区域中

心之间的相对位置信息,确定该待处理文本行的文字排列方式。

[0167] 本申请实施例中,在获取到待处理文本行的全局区域中心以及字符串局部区域中心之后,可以依据该待处理文本行的全局区域中心以及字符串局部区域中心之间的相对关系(尤其是位置上的相对关系),确定该待处理文本行的文字排列方式;而且该确定过程具体可以包括:若待处理文本行的全局区域中心与该待处理文本行的字符串局部区域中心重合,则确定该待处理文本行的文字排列方式为竖排;若该待处理文本行的全局区域中心与该待处理文本行的字符串局部区域中心不重合,则确定该待处理文本行的文字排列方式为横排。

[0168] S332:根据待处理文本行的全局区域中心、该待处理文本行的首字符区域中心和该待处理文本行的字符串局部区域中心中至少两个之间的相对位置信息,确定该待处理文本行的文本行旋转方式。

[0169] 本申请实施例不限定S332,例如,S332可以采用以下三种方式进行实施。

[0170] 方式一:根据待处理文本行的全局区域中心和该待处理文本行的首字符区域中心之间的相对位置信息,确定该待处理文本行的文本行旋转方式。

[0171] 需要说明的是,方式一的相关内容可以参见上文“S33的第一个示例”的相关内容。

[0172] 方式二:根据待处理文本行的首字符区域中心和该待处理文本行的字符串局部区域中心之间的相对位置信息,确定该待处理文本行的文本行旋转方式。

[0173] 本申请实施例不限定上述确定过程,例如,在一种可能的实施方式下,待处理文本行的文本行旋转方式的确定过程可以包括步骤91-步骤94:

[0174] 步骤91:对待处理文本行的首字符区域中心和该待处理文本行的字符串局部区域中心进行有向连接,得到第三向量。

[0175] 需要说明的是,本申请实施例不限定步骤91中“有向连接”的连接方式,例如,可以将待处理文本行的首字符区域中心作为起点,并将该待处理文本行的字符串局部区域中心作为终点进行有向连接,以使第三向量成为一个以待处理文本行的首字符区域中心为起点且以该待处理文本行的字符串局部区域中心为终点的有向线段。又如,可以将待处理文本行的字符串局部区域中心作为起点,并将该待处理文本行的首字符区域中心作为终点进行有向连接,以使第三向量成为一个以待处理文本行的字符串局部区域中心为起点且以该待处理文本行的首字符区域中心为终点的有向线段。

[0176] 步骤92:将第三向量与第三预设向量之间的夹角,确定为待处理文本行的第三旋转角度。

[0177] 其中,第三预设向量是指预先设定的用于表示首字符区域中心和字符串局部区域中心之间有向连接的参考向量。

[0178] 实际上,具有不同文字排列方式的文本行对应于不同的参考向量。基于此,本申请实施例还提供了确定第三预设向量的一种可能的实施方式,其具体可以包括:根据待处理文本行的文字排列方式和第三映射关系,确定第三预设向量。其中,第三映射关系包括待处理文本行的文字排列方式与第三预设向量之间的对应关系。

[0179] 第三映射关系用于记录不同文字排列方式对应的用于表示首字符区域中心和字符串局部区域中心之间有向连接的参考向量;而且本申请实施例不限定第三映射关系,例如,第三映射关系可以包括横排与第四标准向量之间的对应关系、以及竖排与第五标准向

量之间的对应关系。

[0180] 第四标准向量是指用于表示横向文本行的首字符区域中心与字符串局部区域中心之间有向连接的标准向量；而且本申请实施例不限定第四标准向量，例如，第四标准向量可以根据标准横向文本行的首字符区域中心和该标准横向文本行的字符串局部区域中心之间的有向连接确定（例如，如图4所示，可以将标准横向文本行的首字符区域中心和该标准横向文本行的字符串局部区域中心之间的有向连接确定为第四标准向量。又如，可以将标准横向文本行的首字符区域中心和该标准横向文本行的字符串局部区域中心之间的有向连接的任意一个平行向量，确定为第四标准向量）。

[0181] 第五标准向量是指用于表示竖直文本行的首字符区域中心与字符串局部区域中心之间有向连接的标准向量；而且本申请实施例不限定第五标准向量，例如，第五标准向量可以根据标准竖直文本行的首字符区域中心和该标准竖直文本行的字符串局部区域中心之间的有向连接确定（例如，如图4所示，可以将标准竖直文本行的首字符区域中心和该标准竖直文本行的字符串局部区域中心之间的有向连接确定为第五标准向量。又如，可以将标准竖直文本行的首字符区域中心和该标准竖直文本行的字符串局部区域中心之间的有向连接的任意一个平行向量，确定为第五标准向量）。

[0182] 基于上述“第三预设向量”的相关内容可知，对于待处理文本来说，若确定该待处理文本行的文字排列方式为横排，则可以将第四标准向量确定为第三预设向量；若确定该待处理文本行的文字排列方式为竖排，则可以将第五标准向量确定为第三预设向量。

[0183] 另外，第三旋转角度可以用于表示待处理文本行的文本行旋转角度。

[0184] 基于上述步骤92的相关内容可知，对于待处理文本来说，在获取到第三向量之后，可以将第三向量与与第三预设向量之间的夹角，确定为该待处理文本行的第三旋转角度，以使该第三旋转角度能够准确地表示该待处理文本行的文本行旋转角度。

[0185] 步骤93：根据待处理文本行的首字符区域中心和该待处理文本行的字符串局部区域中心之间在第三预设方向上的相对关系以及位置差值，确定该待处理文本行的第三旋转方向。

[0186] 其中，第三预设方向是指预先设定的参考方向。

[0187] 实际上，具有不同文字排列方式的文本行对应于不同的参考方向。基于此，本申请实施例还提供了确定第三预设方向的一种可能的实施方式，其具体可以包括：根据该待处理文本行的文字排列方式和第四映射关系，确定第三预设方向；其中，第四映射关系包括待处理文本行的文字排列方式与第三预设方向之间的对应关系。

[0188] 第四映射关系用于记录不同文字排列方式对应的与全局区域中心与首字符区域中心相关的参考方向；而且本申请实施例不限定第四映射关系，例如，第四映射关系可以包括横排与垂直方向之间的对应关系、以及竖排与水平方向之间的对应关系。

[0189] 基于上述“第三预设方向”的相关内容可知，对于待处理文本行来说，若确定该待处理文本行的文字排列方式为横排，则可以将垂直方向确定为第三预设方向；若确定该待处理文本行的文字排列方式为竖排，则可以将水平方向确定为第三预设方向。

[0190] 另外，第三旋转方向可以用于描述待处理文本行的文本行旋转方向。

[0191] 此外，本申请实施例不限定步骤93的实施方式，为了便于理解，下面结合两种场景进行说明。

[0192] 场景一:当待处理文本行的文字排列方式为横排,且第三预设方向为垂直方向时,步骤93具体可以包括:

[0193] 若待处理文本行的首字符区域中心在第三预设方向上的投影位置高于该待处理文本行的字符串局部区域中心在第三预设方向上的投影位置,且该待处理文本行的首字符区域中心与该待处理文本行的字符串局部区域中心之间在第三预设方向上的投影位置差值高于预设位置差值,则确定该待处理文本行的第三旋转方向为顺时针方向;

[0194] 若待处理文本行的首字符区域中心在第三预设方向上的投影位置高于该待处理文本行的字符串局部区域中心在第三预设方向上的投影位置,且该待处理文本行的首字符区域中心与该待处理文本行的字符串局部区域中心之间在第三预设方向上的投影位置差值低于预设位置差值,则确定该待处理文本行的第三旋转方向为逆时针方向;

[0195] 若待处理文本行的首字符区域中心在第三预设方向上的投影位置不高于该待处理文本行的字符串局部区域中心在第三预设方向上的投影位置,则确定该待处理文本行的第三旋转方向为逆时针方向。

[0196] 其中,预设位置差值可以预先设定,也可以根据待处理文本行对应的正方向文本行的首字符区域中心和该正方向文本行的字符串局部区域中心之间在垂直方向的高度差值确定(例如,将待处理文本行对应的正方向文本行的首字符区域中心和该正方向文本行的字符串局部区域中心之间在垂直方向的高度差值,确定为预设位置差值)。

[0197] 可见,如图7所示,当待处理文本行的文字排列方式为横排(也就是,该待处理文本行属于横向文本行)时,在获取到待处理文本行的首字符区域中心(如图7中“03”)和该待处理文本行的字符串局部区域中心(如图7中“02”)之间在垂直方向的相对关系以及位置差值之后,若该相对关系以及位置差值表示待处理文本行的首字符区域中心在垂直方向上的投影位置比该待处理文本行的字符串局部区域中心在垂直方向上的投影位置高出的距离大于预设位置差值(如图7中“预设位置差值”),则确定待处理文本行的第三旋转方向为顺时针方向;若该相对关系以及位置差值表示待处理文本行的首字符区域中心在垂直方向上的投影位置比该待处理文本行的字符串局部区域中心在垂直方向上的投影位置高出的距离小于预设位置差值、或者该相对关系以及位置差值表示待处理文本行的首字符区域中心在垂直方向上的投影位置不高于(例如,低于或者相等)该待处理文本行的字符串局部区域中心在垂直方向上的投影位置,则确定待处理文本行的第三旋转方向为逆时针方向。

[0198] 需要说明的是,若待处理文本行的文字排列方式为横排,该待处理文本行的首字符区域中心在垂直方向上的投影位置高于该待处理文本行的字符串局部区域中心在垂直方向上的投影位置,且该待处理文本行的首字符区域中心在垂直方向上的投影位置与该待处理文本行的字符串局部区域中心在垂直方向上的投影位置之间的位置差值等于预设位置差值,则可以确定该待处理文本行处于正方向或者该待处理文本行旋转 180° 。

[0199] 场景二:当待处理文本行的文字排列方式为竖排,且第三预设方向为水平方向时,步骤93具体可以包括:若待处理文本行的首字符区域中心在第三预设方向上的投影位置比该待处理文本行的字符串局部区域中心在第三预设方向上的投影位置偏左,则确定该待处理文本行的第三旋转方向为逆时针方向;若待处理文本行的首字符区域中心在第三预设方向上的投影位置比待处理文本行的字符串局部区域中心偏右,则确定待处理文本行的第三旋转方向为顺时针方向。

[0200] 其中,因竖直文本行的字符串局部区域中心与该竖直文本行的全局区域中心重合,使得“根据竖直文本行的首字符区域中心以及字符串局部区域中心确定该竖直文本行的文本行旋转方式的过程”类似于上文“根据竖直文本行的全局区域中心以及首字符区域中心确定该竖直文本行的文本行旋转方式的过程”。

[0201] 可见,如图8所示,若待处理文本行的文字排列方式为竖排(也就是,该待处理文本行属于竖直文本行),则在获取到待处理文本行的首字符区域中心(如图8中“03”)以及字符串局部区域中心(如图8中“02”)在水平方向的相对关系之后,如果该相关关系表示待处理文本行的首字符区域中心在水平方向上的投影位置比该待处理文本行的字符串局部区域中心在水平方向上的投影位置偏左,则确定该待处理文本行的第三旋转方向为逆时针方向;如果该相关关系表示待处理文本行的首字符区域中心在水平方向上的投影位置比待处理文本行的字符串局部区域中心在水平方向上的投影位置偏右,则确定待处理文本行的第三旋转方向为顺时针方向。

[0202] 需要说明的是,若待处理文本行的文字排列方式为竖排,且该待处理文本行的首字符区域中心在水平方向上的投影位置和该待处理文本行的字符串局部区域中心在水平方向上的投影位置重合,则可以确定该待处理文本行处于正方向或者该待处理文本行旋转 180° 。

[0203] 基于上述步骤93的相关内容可知,在获取到待处理文本行的首字符区域中心和该待处理文本行的字符串局部区域中心之后,可以先获取待处理文本行的首字符区域中心和该待处理文本行的字符串局部区域中心之间在第一预设方向上的相对关系;再根据该相关关系,确定该待处理文本行的第三旋转方向,以使该第三旋转方向能够准确地表示出该待处理文本的文本行旋转方向。

[0204] 步骤94:根据待处理文本行的第三旋转角度和/或待处理文本行的第三旋转方向,生成待处理文本行的文本行旋转方式。

[0205] 作为示例,若文本行旋转方式包括文本行旋转方向和/或文本行旋转角度,则步骤94可以包括步骤941和/或步骤942:

[0206] 步骤941:根据待处理文本行的第三旋转方向,确定该待处理文本行的文本行旋转方向(如,可以直接将待处理文本行的第三旋转方向,确定为该待处理文本行的文本行旋转方向)。

[0207] 步骤942:根据待处理文本行的第三旋转角度,确定该待处理文本行的文本行旋转角度(如,可以直接将待处理文本行的第三旋转角度,确定为该待处理文本行的文本行旋转角度)。

[0208] 基于上述步骤94的相关内容可知,本申请实施例中,在获取到待处理文本行的第三旋转角度以及第三旋转方向之后,可以先依据该第三旋转角度以及第三旋转方向,分别确定该待处理文本行的文本行旋转方向和该待处理文本行的文本行旋转角度;在依据该待处理文本行的文本行旋转方向和/或该待处理文本行的文本行旋转角度,确定该待处理文本行的文本行旋转方式,以使该文本行旋转方式能够准确地表示出该待处理文本行的文本行旋转方向和/或旋转角度,从而使得该文本行旋转方式能够准确地表示出该待处理文本行在待处理图像中的呈现方式。

[0209] 基于上述方式二的相关内容可知,在获取到该待处理文本的首字符区域中心以及

字符串局部区域中心之后,可以依据该待处理文本的首字符区域中心与字符串局部区域中心之间的相对位置信息,确定待处理文本行的文本行旋转方式,以使该文本行旋转方式能够准确地表示出该待处理文本行在待处理图像中的文本行旋转信息,从而使得基于该文本行旋转方式能够准确地表示出该待处理文本行在待处理图像中的呈现方式。

[0210] 方式三:根据待处理文本行的全局区域中心、该待处理文本行的首字符区域中心和该待处理文本行的字符串局部区域中心之间的相对位置信息,确定该待处理文本行的文本行旋转方式。

[0211] 本申请实施例不限定上述确定过程,例如,在一种可能的实施方式下,待处理文本行的文本行旋转方式的确定过程可以包括步骤101-步骤103:

[0212] 步骤101:根据待处理文本行的全局区域中心、该待处理文本行的首字符区域中心和该待处理文本行的字符串局部区域中心中至少两个之间的相对位置信息,确定该待处理文本行的文本行旋转角度。

[0213] 也就是,对于待处理文本行来说,在获取到该待处理文本行的全局区域中心、首字符区域中心以及字符串局部区域中心之后,可以采用以下三种方式确定该待处理文本行的文本行旋转角度;而且该三种方式具体可以为:

[0214] 第一种方式:根据待处理文本行的全局区域中心和该待处理文本行的首字符区域中心之间的相对位置关系,确定该待处理文本行的文本行旋转角度。

[0215] 需要说明的是,第一种方式可以采用上述步骤81至步骤82进行实施,只需将上述步骤81至步骤82中“待处理文本行的第一旋转角度”替换为“待处理文本行的文本行旋转角度”即可。

[0216] 第二种方式:根据待处理文本行的首字符区域中心和该待处理文本行的字符串局部区域中心之间的相对位置信息,确定待处理文本行的文本行旋转角度。

[0217] 需要说明的是,第二种方式可以采用上述步骤91至步骤92进行实施,只需将上述步骤91至步骤92中“待处理文本行的第三旋转角度”替换为“待处理文本行的文本行旋转角度”即可。

[0218] 第三种方式:根据待处理文本行的全局区域中心、该待处理文本行的首字符区域中心和该待处理文本行的字符串局部区域中心之间的相对位置信息,确定该待处理文本行的文本行旋转角度。

[0219] 本申请实施例不限定第三种方式的实施方式,例如,其具体可以包括步骤111-步骤113:

[0220] 步骤111:根据待处理文本行的全局区域中心和该待处理文本行的首字符区域中心之间的相对位置关系,确定该待处理文本行的第一旋转角度。

[0221] 需要说明的是,步骤111可以采用上述步骤81至步骤82进行实施。

[0222] 步骤112:根据待处理文本行的首字符区域中心和该待处理文本行的字符串局部区域中心之间的相对位置信息,确定该待处理文本行的第三旋转角度。

[0223] 需要说明的是,步骤113可以采用上述步骤91至步骤92进行实施。

[0224] 步骤113:根据待处理文本行的第一旋转角度、以及该待处理文本行的第三旋转角度,确定该待处理文本行的文本行旋转角度。

[0225] 本申请实施例不限定步骤113的实施方式,例如,步骤113具体可以为:将待处理文

本行的第一旋转角度以及该待处理文本行的第三旋转角度中最大值(或者中位值,或者最小值,或者众数),确定为该待处理文本行的文本行旋转角度。又如,步骤113具体可以为:将待处理文本行的第一旋转角度以及该待处理文本行的第三旋转角度之间的平均值(或者加权平均值),确定为该待处理文本行的文本行旋转角度。

[0226] 基于上述步骤101的相关内容可知,在获取到待处理文本行的全局区域中心、首字符区域中心以及字符串局部区域中心之后,可以根据待处理文本行的全局区域中心、首字符区域中心以及字符串局部区域中心中至少两个之间的相对位置信息,确定该待处理文本行的文本行旋转角度,以便后续能够基于该文本行旋转角度,确定该待处理文本行的文本行旋转方式。

[0227] 步骤102:根据待处理文本行的全局区域中心、该待处理文本行的首字符区域中心和该待处理文本行的字符串局部区域中心中至少两个之间的相对位置信息,确定该待处理文本行的文本行旋转方向。

[0228] 也就是,对于待处理文本行来说,在获取到待处理文本行的全局区域中心、首字符区域中心以及字符串局部区域中心之后,可以采用以下四种实施方式确定待处理文本行的文本行旋转方向;而且该三种实施方式具体可以为:

[0229] 第一种实施方式:根据待处理文本行的全局区域中心和该待处理文本行的首字符区域中心之间的相对位置关系,确定该待处理文本行的文本行旋转方向。

[0230] 需要说明的是,第一种实施方式可以采用上述步骤83进行实施,只需将上述步骤83中“待处理文本行的第一旋转方向”替换为“待处理文本行的文本行旋转方向”即可。

[0231] 第二种实施方式:根据待处理文本行的首字符区域中心和该待处理文本行的字符串局部区域中心之间的相对位置信息,确定该待处理文本行的文本行旋转方向。

[0232] 需要说明的是,第二种实施方式可以采用上述步骤93进行实施,只需将上述步骤93中“待处理文本行的第三旋转方向”替换为“待处理文本行的文本行旋转方向”即可。

[0233] 第三种实施方式:根据待处理文本行的全局区域中心、该待处理文本行的首字符区域中心和该待处理文本行的字符串局部区域中心之间的相对位置信息,确定该待处理文本行的文本行旋转方向。

[0234] 本申请实施例不限定第四种实施方式的实施方式,例如,其具体可以包括步骤121-步骤123:

[0235] 步骤121:根据待处理文本行的全局区域中心和该待处理文本行的首字符区域中心之间的相对位置关系,确定该待处理文本行的第一旋转方向。

[0236] 需要说明的是,步骤121可以采用上述步骤83进行实施。

[0237] 步骤122:根据待处理文本行的首字符区域中心和该待处理文本行的字符串局部区域中心之间的相对位置信息,确定该待处理文本行的第三旋转方向。

[0238] 需要说明的是,步骤123可以采用上述步骤93进行实施。

[0239] 步骤123:根据待处理文本行的第一旋转方向以及该待处理文本行的第三旋转方向,确定该待处理文本行的文本行旋转方向。

[0240] 本申请实施例不限定步骤123的实施方式,例如,步骤123具体可以为:将待处理文本行的第一旋转方向以及该待处理文本行的第三旋转方向中最大值(或者中位值,或者最小值,或者众数),确定为该待处理文本行的文本行旋转方向。又如,步骤123具体可以为:将

待处理文本行的第一旋转方向以及该待处理文本行的第三旋转方向之间的平均值(或者加权平均值),确定为该待处理文本行的文本行旋转方向。

[0241] 基于上述步骤102的相关内容可知,在获取到待处理文本行的全局区域中心、首字符区域中心以及字符串局部区域中心之后,可以根据待处理文本行的全局区域中心、首字符区域中心以及字符串局部区域中心中至少两个之间的相对位置信息,确定该待处理文本行的文本行旋转方向,以便后续能够基于该文本行旋转方向,确定该待处理文本行的文本行旋转方式。

[0242] 步骤103:根据待处理文本行的文本行旋转角度和待处理文本行的文本行旋转方向,生成待处理文本行的文本行旋转方式。

[0243] 本申请实施例中,在获取到待处理文本行的文本行旋转角度以及旋转方向之后,可以依据该待处理文本行的文本行旋转角度以及旋转方向,生成该待处理文本行的文本行旋转方式,以使该文本行旋转方式能够准确地表示出该待处理文本行的文本行旋转方向以及文本行旋转角度,从而该文本行旋转方式能够准确地表示出该待处理文本行在待处理图像中的文本行旋转信息,进而使得基于该文本行旋转方式能够准确地表示出该待处理文本行在待处理图像中的呈现方式。

[0244] 基于上述方式三的相关内容可知,在获取到该待处理文本的全局区域中心、首字符区域中心以及字符串局部区域中心之后,可以依据这三个中心之间的相对位置信息,确定待处理文本行的文本行旋转方式,以使该文本行旋转方式能够准确地表示出该待处理文本行在待处理图像中的文本行旋转信息,从而使得基于该文本行旋转方式能够准确地表示出该待处理文本行在待处理图像中的呈现方式。

[0245] 基于上述S1至S3的相关内容可知,在获取到待处理图像之后,先确定该待处理图像中待处理文本行的全局区域和该待处理文本行的局部区域,以使该待处理文本行的全局区域表示该待处理文本行在该待处理图像内所占区域,以及该待处理文本行的局部区域表示该待处理文本行的预设部分在该待处理图像内所占区域;再将该待处理文本行的全局区域与该待处理文本行的局部区域进行位置比较,得到该待处理文本行的文本行方向。可见,因待处理文本行的全局区域以及局部区域分别能够表示该待处理文本行在待处理图像内所占区域以及该待处理文本行的预设部分在待处理图像内所占区域,使得该待处理文本行的全局区域以及局部区域能够综合表示出该待处理文本行在待处理图像中的呈现方式,从而使得基于待处理文本行的全局区域以及局部区域确定出的该待处理文本行的文本行方向能够准确地描述出该待处理文本行在待处理图像中的呈现方式,如此能够实现准确地确定出图像中文本行的文本行方向,从而有利于提高图像中文本行的字符识别准确性。

[0246] 基于上述方法实施例提供的文本行方向确定方法,本申请实施例还提供了一种文本行方向确定装置,下面结合附图进行解释和说明。

[0247] 装置实施例

[0248] 装置实施例提供的文本行方向确定装置的技术详情,请参照上述方法实施例。

[0249] 参见图9,该图为本申请实施例提供的一种文本行方向确定装置的结构示意图。

[0250] 本申请实施例提供的文本行方向确定装置900,包括:

[0251] 图像获取单元901,用于获取待处理图像;

[0252] 区域确定单元902,用于确定所述待处理图像中待处理文本行的全局区域和所述

待处理文本行的局部区域;其中,所述待处理文本行的全局区域表示所述待处理文本行在所述待处理图像内所占区域;所述待处理文本行的局部区域表示所述待处理文本行的预设部分在所述待处理图像内所占区域;

[0253] 方向确定单元903,用于将所述待处理文本行的全局区域与所述待处理文本行的局部区域进行位置比较,得到所述待处理文本行的文本行方向。

[0254] 在一种可能的实施方式中,所述区域确定单元902,具体用于:将所述待处理图像输入预先构建的掩码图生成模型,得到所述掩码图生成模型输出的所述待处理图像对应的预测文本行全局掩码图和所述待处理图像对应的预测文本行局部掩码图;其中,所述掩码图生成模型是根据样本图像、所述样本图像对应的实际文本行全局掩码图和所述样本图像对应的实际文本行局部掩码图构建的;所述待处理图像对应的预测文本行全局掩码图包括所述待处理文本行对应的全局掩码区域;所述待处理图像对应的预测文本行局部掩码图包括所述待处理文本行对应的局部掩码区域;将所述待处理图像对应的预测文本行全局掩码图中所述待处理文本行对应的全局掩码区域,确定为所述待处理文本行的全局区域;将所述待处理图像对应的预测文本行局部掩码图中所述待处理文本行对应的局部掩码区域,确定为所述待处理文本行的局部区域。

[0255] 在一种可能的实施方式中,所述待处理文本行的局部区域包括所述待处理文本行的首字符区域。

[0256] 在一种可能的实施方式中,所述方向确定单元903,具体用于:

[0257] 根据所述待处理文本行的全局区域中心和所述待处理文本行的首字符区域中心之间的相对位置信息,确定所述待处理文本行的文本行方向;其中,所述待处理文本行的全局区域中心表示所述待处理文本行的全局区域的中心位置;所述待处理文本行的首字符区域中心表示所述待处理文本行的首字符区域的中心位置。

[0258] 在一种可能的实施方式中,所述文本行方向包括文本行旋转方向和/或文本行旋转角度;

[0259] 所述待处理文本行的文本行旋转方向的确定过程,包括:根据所述待处理文本行的全局区域中心和所述待处理文本行的首字符区域中心之间在第一预设方向上的相对关系,确定所述待处理文本行的第一旋转方向;根据所述待处理文本行的第一旋转方向,确定所述待处理文本行的文本行旋转方向;

[0260] 所述待处理文本行的文本行旋转角度的确定过程,包括:对所述待处理文本行的全局区域中心和所述待处理文本行的首字符区域中心进行有向连接,得到第一向量;将所述第一向量与第一预设向量之间的夹角,确定为所述待处理文本行的第一旋转角度;根据所述待处理文本行的第一旋转角度,确定所述待处理文本行的文本行旋转角度。

[0261] 在一种可能的实施方式中,所述待处理文本行的局部区域包括所述待处理文本行的首字符区域和所述待处理文本行的字符串局部区域;其中,所述待处理文本行的字符串局部区域是由所述待处理文本行中各个字符的预设局部区域进行连接得到的。

[0262] 在一种可能的实施方式中,所述方向确定单元903,具体用于:根据所述待处理文本行的全局区域中心、所述待处理文本行的首字符区域中心和所述待处理文本行的字符串局部区域中心中至少两个之间的相对位置信息,确定所述待处理文本行的文本行方向;其中,所述待处理文本行的全局区域中心表示所述待处理文本行的全局区域的中心位置;所

述待处理文本行的首字符区域中心表示所述待处理文本行的首字符区域的中心位置；所述待处理文本行的字符串局部区域中心表示所述待处理文本行的字符串局部区域的中心位置。

[0263] 在一种可能的实施方式中，所述文本行方向包括文本行旋转方式和/或文字排列方式；

[0264] 所述待处理文本行的文本行旋转方式的确定过程，包括：根据所述待处理文本行的首字符区域中心和所述待处理文本行的字符串局部区域中心之间的相对位置信息，确定所述待处理文本行的文本行旋转方式；

[0265] 所述待处理文本行的文字排列方式的确定过程，包括：根据所述待处理文本行的全局区域中心和所述待处理文本行的字符串局部区域中心之间的相对位置信息，确定所述待处理文本行的文字排列方式。

[0266] 在一种可能的实施方式中，所述文本行旋转方式包括文本行旋转方向和/或文本行旋转角度；

[0267] 所述待处理文本行的文本行旋转方向的确定过程，包括：根据所述待处理文本行的首字符区域中心和所述待处理文本行的字符串局部区域中心之间在第三预设方向上的相对关系以及位置差值，确定所述待处理文本行的第三旋转方向；根据所述待处理文本行的第三旋转方向，确定所述待处理文本行的文本行旋转方向；

[0268] 所述待处理文本行的文本行旋转角度的确定过程，包括：对所述待处理文本行的首字符区域中心和所述待处理文本行的字符串局部区域中心进行有向连接，得到第三向量；将所述第三向量与第三预设向量之间的夹角，确定为所述待处理文本行的第三旋转角度；根据所述待处理文本行的第三旋转角度，确定所述待处理文本行的文本行旋转角度。

[0269] 在一种可能的实施方式中，若所述待处理文本行的文字排列方式为横排，则所述待处理文本行的字符串局部区域是由所述待处理文本行中各个字符的底部边界区域进行连接得到的；若所述待处理文本行的文字排列方式为竖排，则所述待处理文本行的字符串局部区域是由所述待处理文本行中各个字符的中心区域进行连接得到的。

[0270] 在一种可能的实施方式中，所述待处理文本行的文字排列方式的确定过程，包括：若所述待处理文本行的全局区域中心与所述待处理文本行的字符串局部区域中心重合，则确定所述待处理文本行的文字排列方式为竖排；若所述待处理文本行的全局区域中心与所述待处理文本行的字符串局部区域中心不重合，则确定所述待处理文本行的文字排列方式为横排。

[0271] 在一种可能的实施方式中，所述文本行方向包括文本行旋转方向、文本行旋转角度、或文字排列方式中的至少一个；

[0272] 所述待处理文本行的文本行方向的确定过程，包括：根据所述待处理文本行的全局区域中心、所述待处理文本行的首字符区域中心和所述待处理文本行的字符串局部区域中心中至少两个之间的相对位置信息，确定所述待处理文本行的文本行旋转角度；根据所述待处理文本行的全局区域中心、所述待处理文本行的首字符区域中心和所述待处理文本行的字符串局部区域中心中至少两个之间的相对位置信息，确定所述待处理文本行的文本行旋转方向；根据所述待处理文本行的全局区域中心和所述待处理文本行的字符串局部区域中心之间的相对位置信息，确定所述待处理文本行的文字排列方式。

[0273] 进一步地,本申请实施例还提供了一种设备,所述设备包括处理器以及存储器:

[0274] 所述存储器用于存储计算机程序;

[0275] 所述处理器用于根据所述计算机程序执行本申请实施例提供的文本行方向确定方法的任一实施方式。

[0276] 进一步地,本申请实施例还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用于存储计算机程序,所述计算机程序用于执行本申请实施例提供的文本行方向确定方法的任一实施方式。

[0277] 进一步地,本申请实施例还提供了一种计算机程序产品,所述计算机程序产品在终端设备上运行时,使得所述终端设备执行本申请实施例提供的文本行方向确定方法的任一实施方式。

[0278] 应当理解,在本申请中,“至少一个(项)”是指一个或者多个,“多个”是指两个或两个以上。“和/或”,用于描述关联对象的关联关系,表示可以存在三种关系,例如,“A和/或B”可以表示:只存在A,只存在B以及同时存在A和B三种情况,其中A,B可以是单数或者复数。字符“/”一般表示前后关联对象是一种“或”的关系。“以下至少一项(个)”或其类似表达,是指这些项中的任意组合,包括单项(个)或复数项(个)的任意组合。例如,a,b或c中的至少一项(个),可以表示:a,b,c,“a和b”,“a和c”,“b和c”,或“a和b和c”,其中a,b,c可以是单个,也可以是多个。

[0279] 以上所述,仅是本发明的较佳实施例而已,并非对本发明作任何形式上的限制。虽然本发明已以较佳实施例揭露如上,然而并非用以限定本发明。任何熟悉本领域的技术人员,在不脱离本发明技术方案范围情况下,都可利用上述揭示的方法和技术内容对本发明技术方案做出许多可能的变动和修饰,或修改为等同变化的等效实施例。因此,凡是未脱离本发明技术方案的内容,依据本发明的技术实质对以上实施例所做的任何简单修改、等同变化及修饰,均仍属于本发明技术方案保护的范围内。

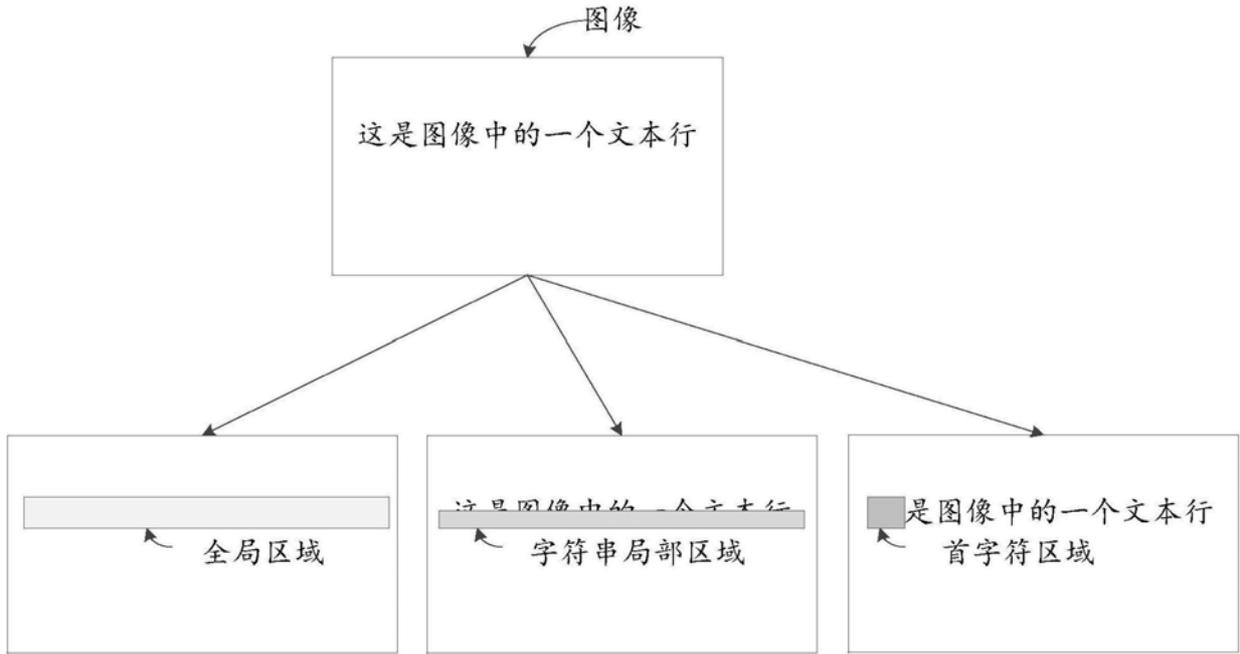


图1

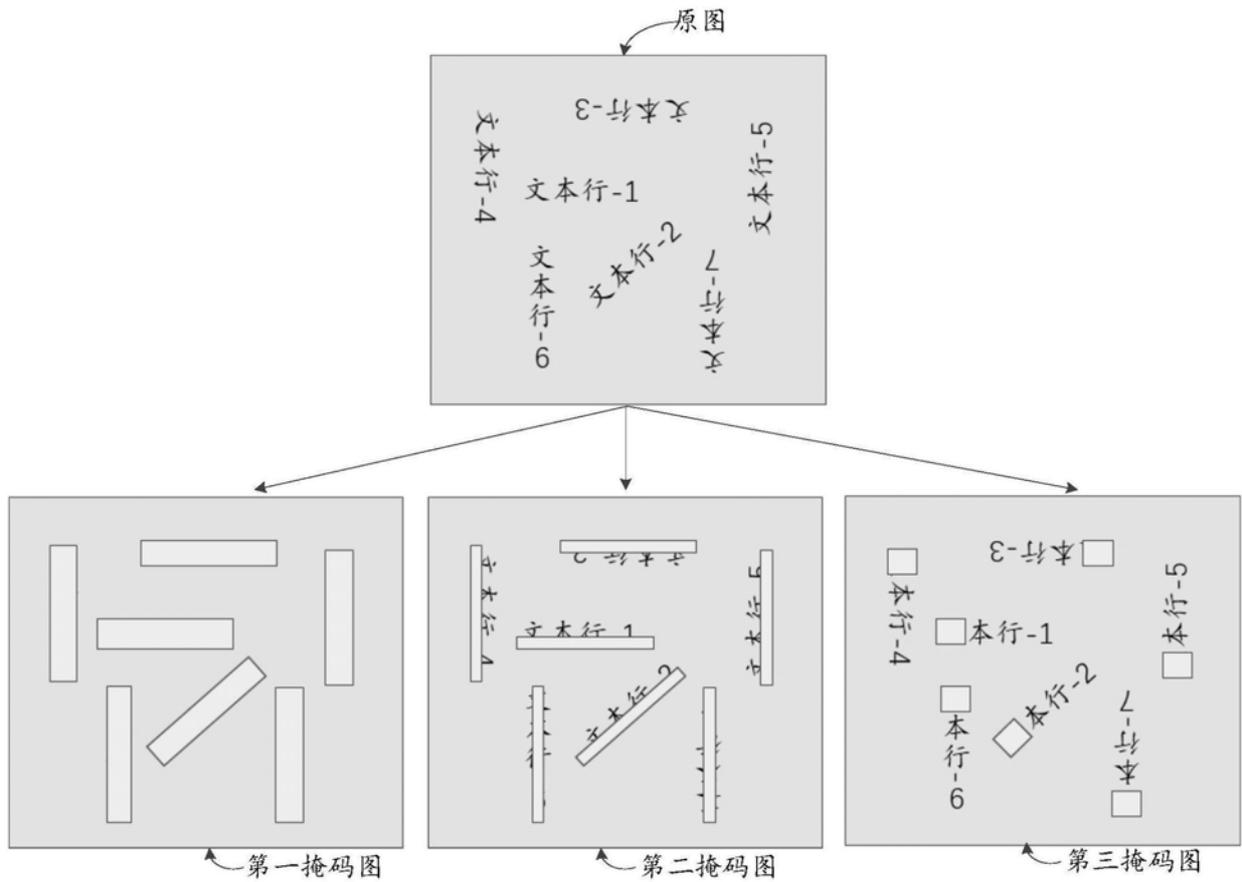


图2

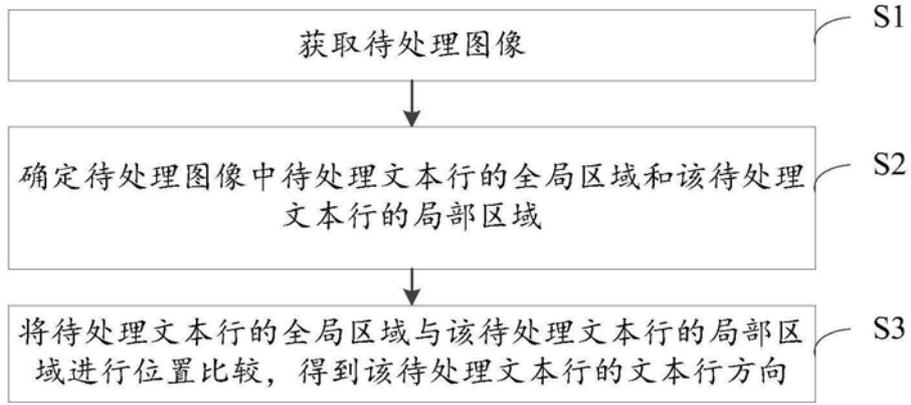


图3

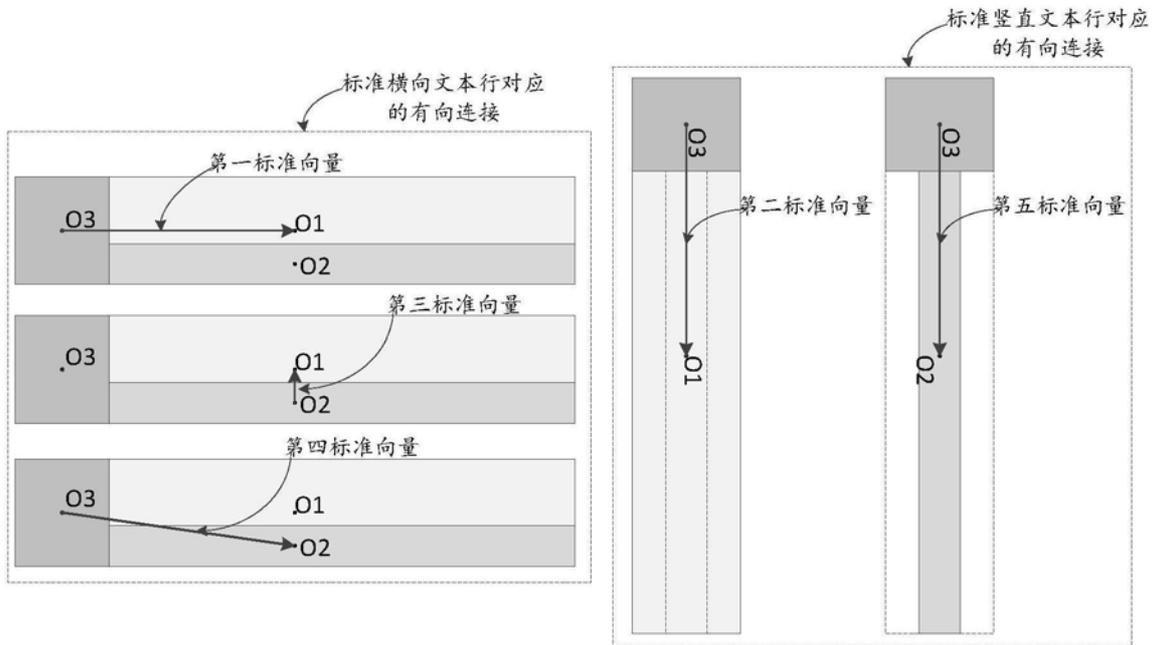


图4

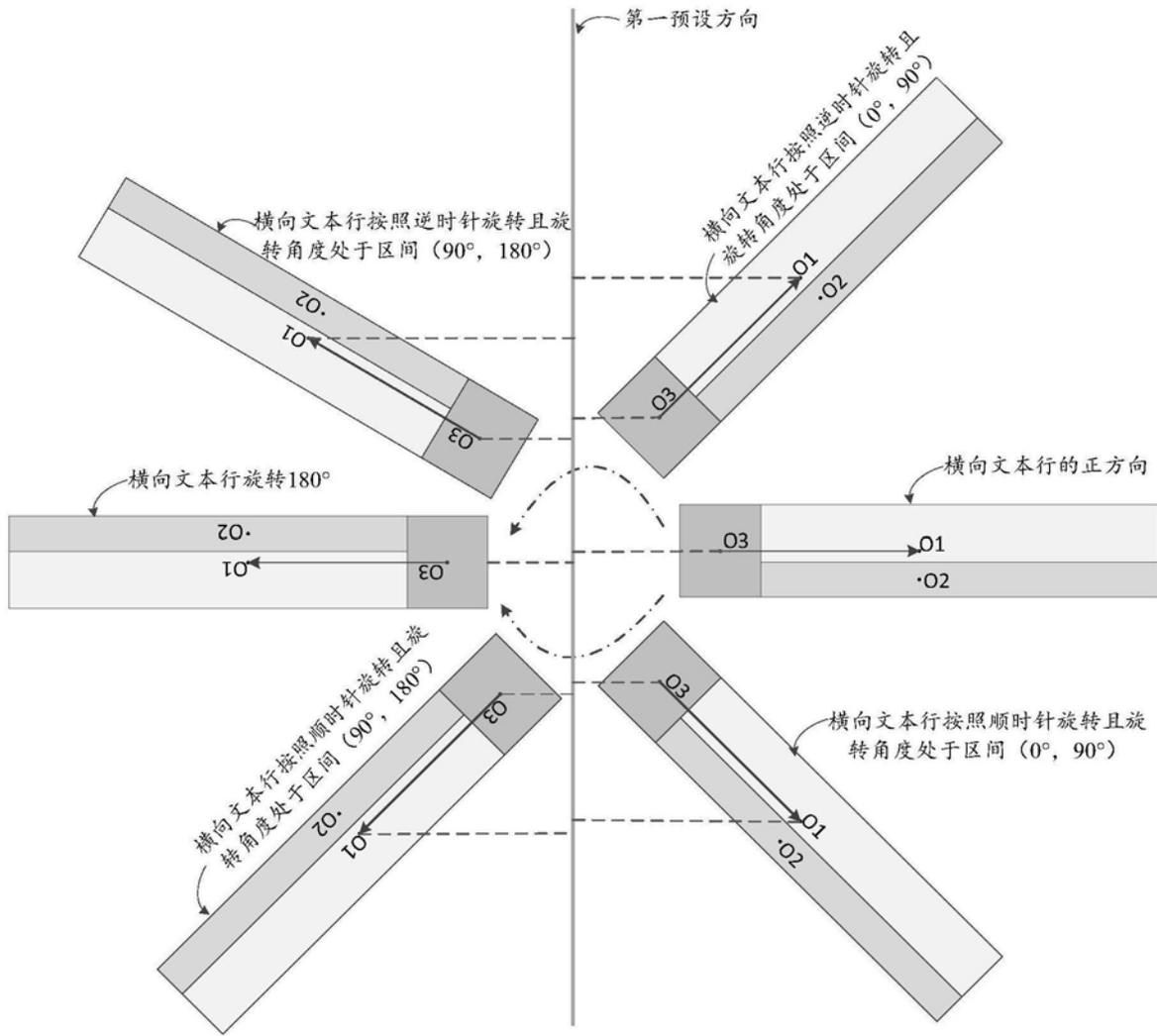


图5

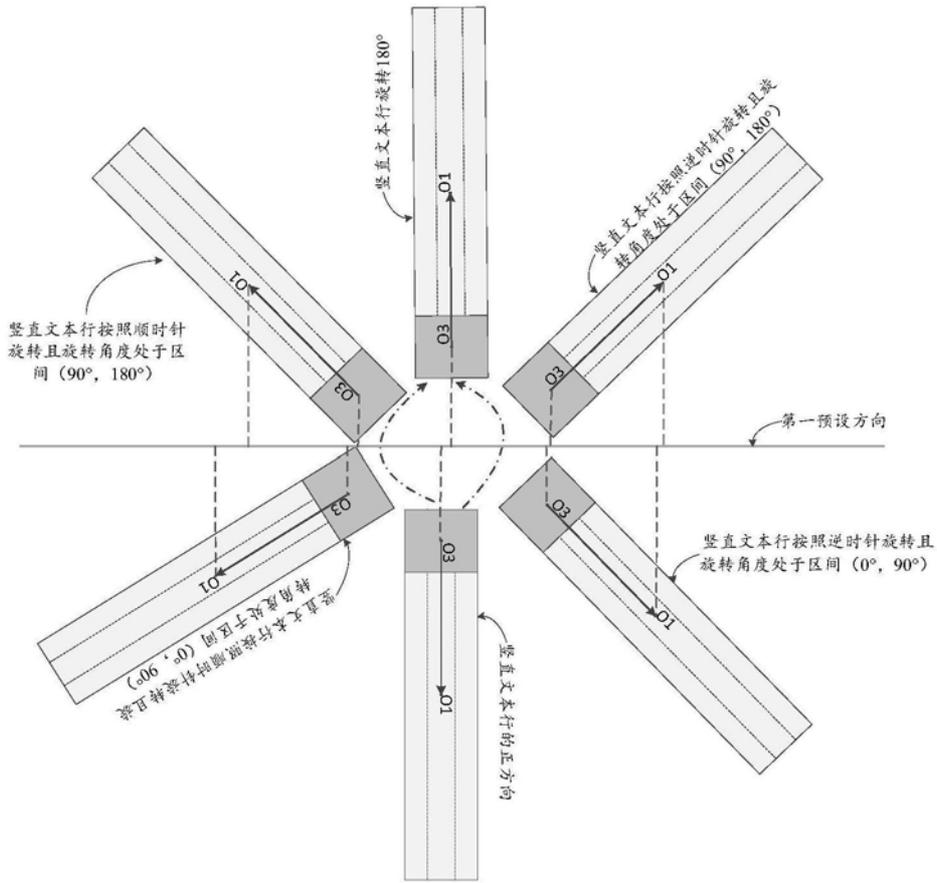


图6

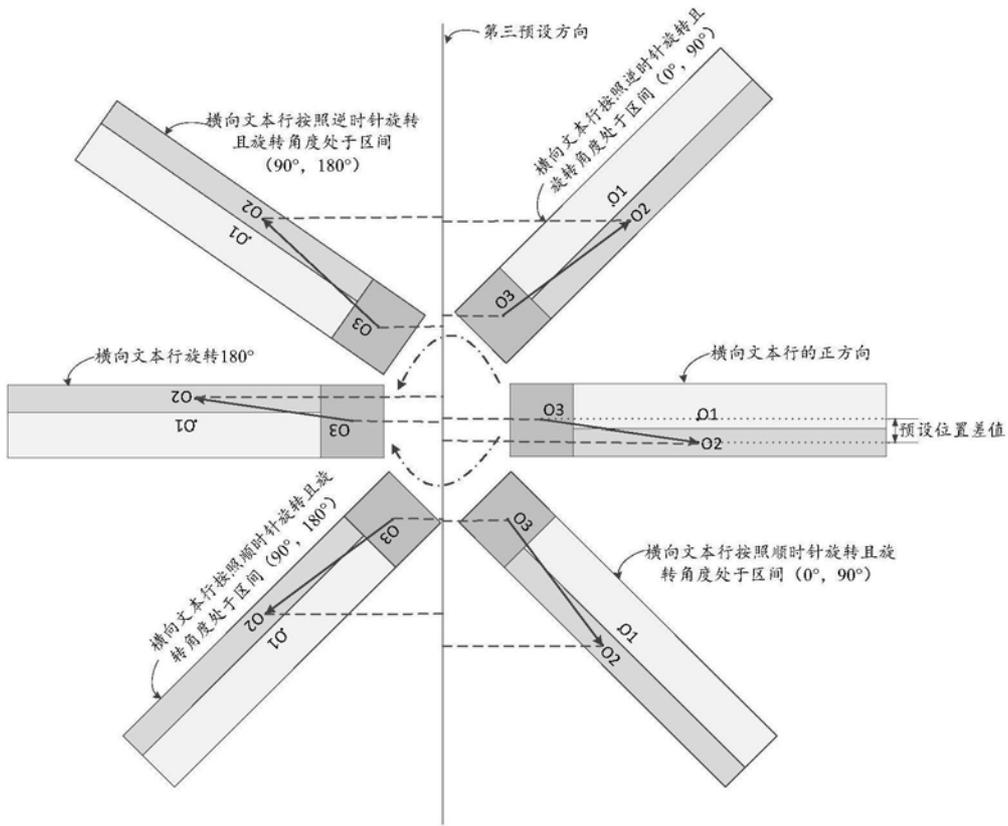


图7

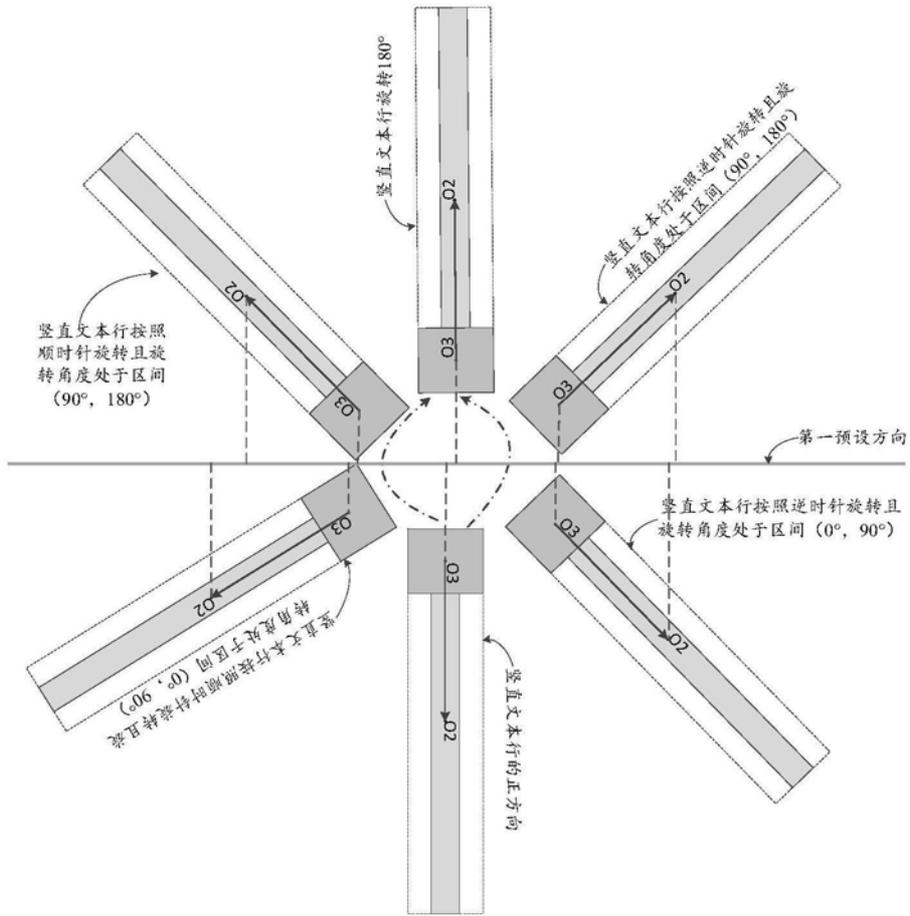


图8

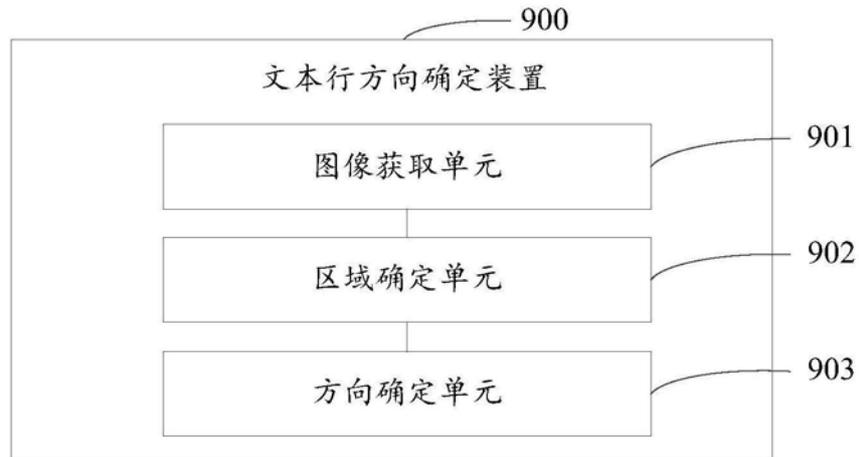


图9