

(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 9 月 20 日 (20.09.2018)



(10) 国际公布号
WO 2018/166276 A1

(51) 国际专利分类号:
G06K 9/20 (2006.01)

(21) 国际申请号: PCT/CN2017/119692

(22) 国际申请日: 2017 年 12 月 29 日 (29.12.2017)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201710152728.X 2017年3月15日 (15.03.2017) CN

(71) 申请人: 北京京东尚科信息技术有限公司 (BEIJING JINGDONG SHANGKE INFORMATION TECHNOLOGY CO., LTD.) [CN/CN]; 中国北京市海淀区杏石口路65号西杉创意园四区11号楼东段1-4层西段1-4层, Beijing 100195 (CN)。

北京京东世纪贸易有限公司 (BEIJING JINGDONG CENTURY TRADING CO., LTD.) [CN/CN]; 中国北京市北京经济技术开发区科创十一街18号C座2层201室, Beijing 100176 (CN)。

(72) 发明人: 王永亮 (WANG, Yongliang); 中国北京市海淀区杏石口路65号西杉创意园四区11号楼东段1-4层西段1-4层, Beijing 100195 (CN)。王青泽 (WANG, Qingze); 中国北京市海淀区杏石口路65号西杉创意园四区11号楼东段1-4层西段1-4层, Beijing 100195 (CN)。陈标龙 (CHEN, Biaolong); 中国北京市海淀区杏石口路65号西杉创意园四区11号楼东段1-4层西段1-4层, Beijing 100195 (CN)。

(74) 代理人: 中国国际贸易促进委员会专利商标事务所 (CCPIT PATENT AND TRADEMARK LAW)

(54) Title: TEXT REGION POSITIONING METHOD AND DEVICE, AND COMPUTER READABLE STORAGE MEDIUM

(54) 发明名称: 文字区域定位方法和装置、计算机可读存储介质

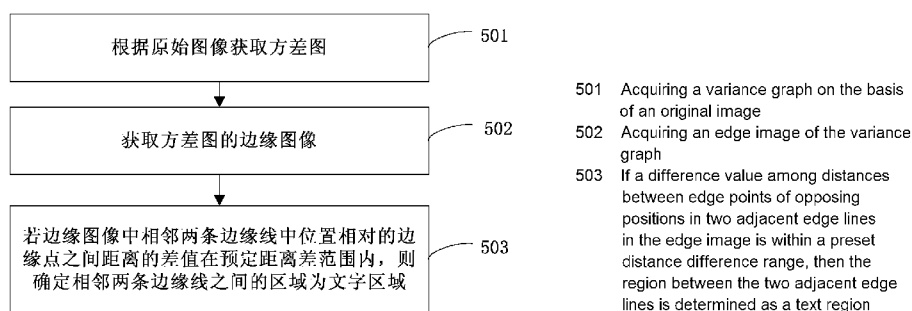


图 5

(57) Abstract: A text region positioning method and device, and a computer readable storage medium, which relate to the field of image processing. The text region positioning method comprises: acquiring a variance graph on the basis of an original image (501); acquiring an edge image of the variance graph (502); if a difference value among distances between edge points of opposing positions in two adjacent edge lines in the edge image is within a preset distance difference range, then the region between the two adjacent edge lines is determined as a text region (503). Such a method makes use of the feature wherein texts in a text region are similar in height, determines a text region on the basis of the distance between edge lines in an edge image, and is not affected by the variation in text stroke thickness. Said method is suitable for use with various types of fonts, and may prevent the complex pixel variation in an image from affecting positioning, thus improving the accuracy of text region positioning.

(57) 摘要: 一种文字区域定位方法和装置、计算机可读存储介质, 涉及图像处理领域。该文字区域定位方法包括: 根据原始图像获取方差图 (501); 获取方差图的边缘图像 (502); 若边缘图像中相邻两条边缘线中位置相对的边缘点之间距离的差值在预定距离差范围时, 则确定相邻两条边缘线之间的区域为文字区域 (503)。通过这样的方法, 能够利用文字区域中文字高度相似的特点, 根据边缘图像中边缘线之间的距离确定文字区域, 不会由于文字笔画粗细变化造成影响, 适用于各种不同的字体, 且能够避免图像复杂的像素变化情况对定位造成影响, 提高文字区域定位的准确性。



OFFICE); 中国北京市西城区阜成门外大街2号
万通新世界广场8层, Beijing 100037 (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))。

文字区域定位方法和装置、计算机可读存储介质

相关申请的交叉引用

本申请是以 CN 申请号为 201710152728.X，申请日为 2017 年 3 月 15 日的申请为
5 基础，并主张其优先权，该 CN 申请的公开内容在此作为整体引入本申请中。

技术领域

本公开涉及图像处理领域，特别是一种文字区域定位方法和装置、计算机可读存储
10 介质。

背景技术

在文字识别过程中，需要先利用某些算法在图片中定位出文字区域，再在文字区
域中利用相关的图像算法提取文字内容，因此文字定位是文字识别引擎的一个重要的
模块。目前常用的方法有笔画宽度定位法（Stroke Width Transform）和稳定极值区
15 域检测法（Maximally Stable Extremal Regions）。

笔画宽度定位法利用打印体文字的笔画宽度不变这一特性，在一个图片中寻找一
对平行线，并将平行线判定为笔画，再把相互靠近的笔画聚集成文字区域。

稳定极值区域检测法是利用图像中的文字区域会跟背景图像形成鲜明的对比这
一特点进行文字区域定位。

发明内容

发明人发现上述相关技术存在各自不同的缺陷。微软正黑体的文字笔画的宽度是
一致的，能够采用笔画宽度定位法进行定位；但是，宋体文字其笔画宽度并不相同，
因此不适用于采用笔画宽度定位法进行定位。而采用稳定极值区域检测法需要文字区
25 域的像素对比度高，但在应用过程中，对比度高的区域不一定是文字，因此该算法很
容易引入额外的噪声。另外，这两种方法都只能先定位文字区域，还需要额外的算法
将单个文字串成行，较为繁琐，降低了运算效率。

为了解决上述问题中的至少一个，本公开提出一种文字区域定位方案，能够提高
对不同字体的适应能力，提高文字区域定位的准确性。

30 根据本公开的一些实施例，提出一种文字区域定位方法，包括：根据原始图像获

取方差图；获取方差图的边缘图像；当边缘图像中相邻两条边缘线中位置相对的边缘点之间距离的差值在预定距离差范围时，确定相邻两条边缘线之间的区域为文字区域。

5 可选地，确定相邻两条边缘点之间的区域为文字区域包括：确定第一边缘点和位于相邻边缘线上的第二边缘点；根据第一边缘点和第二边缘点之间的距离确定行高；连接行高的差值在预定距离差范围内的相邻的第一边缘点，以确定第一边缘线，连接行高的差值在预定距离差范围内的相邻的第二边缘点，以确定第二边缘线，第一边缘线与第二边缘线之间的区域为文字区域。

10 可选地，确定第一边缘点和位于相邻边缘线上的第二边缘点包括：在边缘图像中取一点作为第一边缘点；从第一边缘点沿像素梯度的方向发射射线，直至下一边缘点；当第一边缘点与下一边缘点法向量的夹角小于预定角度阈值时，确定下一边缘点为第二边缘点。

15 可选地，根据原始图像获取方差图包括：获取原始图像中目标像素位置的像素值，及目标像素位置的邻近像素点的像素值，其中，邻近像素点为目标像素位置在不同方向连续预定数量的像素点；将原始图像中目标像素位置和邻近像素点的像素值取方差值，确定方差图中目标像素位置的像素值。

可选地，获取方差图的边缘图像包括：基于方差图利用 Canny 算子计算图像边缘，获取边缘图像。

20 可选地，文字区域包括横向文字区域、纵向文字区域、倾斜文字区域和/扇形文字区域中的至少一种。

通过这样的方法，能够利用文字区域中文字高度相似的特点，根据边缘图像中边缘线之间的距离确定文字区域，不会由于文字笔画粗度变化造成影响，适用于各种不同的字体，且能够避免图像复杂的像素变化情况对定位造成影响，提高文字区域定位的准确性。

25 根据本公开的另一些实施例，提出一种文字区域定位装置，包括：方差图确定模块，用于根据原始图像获取方差图；边缘图像获取模块，用于获取方差图的边缘图像；文字区域定位模块，用于当边缘图像中相邻两条边缘线中位置相对的边缘点之间距离的差值在预定距离差范围内时，确定相邻两条边缘线之间的区域为文字区域。

30 可选地，文字区域定位模块包括：边缘点确定单元，用于确定第一边缘点和位于相邻边缘线上的第二边缘点；行高确定单元，用于根据第一边缘点和第二边缘点之间

的距离确定行高；边缘线连接单元，用于连接行高的差值在预定距离差范围内的相邻的第一边缘点，以确定第一边缘线，连接行高的差值在预定距离差范围内的相邻的第二边缘点，以确定第二边缘线，第一边缘线与第二边缘线之间的区域为文字区域。

5 可选地，边缘点确定单元包括：第一边缘点指定子单元，用于在边缘图像中取一点作为第一边缘点；下一边缘点获取子单元，用于从第一边缘点沿像素梯度的方向发射射线，直至下一边缘点；第二边缘点确定子单元，用于当第一边缘点与下一边缘点法向量的夹角小于预定角度阈值时，确定下一边缘点为第二边缘点。

10 可选地，方差图确定模块用于：获取原始图像中目标像素位置的像素值，及目标像素位置的邻近像素点的像素值，其中，邻近像素点为目标像素位置在不同方向连续预定数量的像素点；将原始图像中目标像素位置和邻近像素点的像素值取方差值，确定方差图中目标像素位置的像素值。

可选地，边缘图像获取模块用于：基于方差图利用 Canny 算子计算图像边缘，获取边缘图像。

15 可选地，文字区域包括横向文字区域、纵向文字区域、倾斜文字区域和扇形文字区域中的至少一种。

这样的装置能够利用文字区域中文字高度相似的特点，根据边缘图像中边缘线之间的距离确定文字区域，不会由于文字笔画粗度变化造成影响，适用于各种不同的字体，且能够避免图像复杂的像素变化情况对定位造成影响，提高文字区域定位的准确性。

20 根据本公开的又一些实施例，提出一种文字区域定位装置，包括：存储器；以及耦接至存储器的处理器，处理器被配置为基于存储在存储器的指令执行上文中提到的任意一种文字区域定位方法。

25 这样的装置能够利用文字区域中文字高度相似的特点，根据边缘图像中边缘线之间的距离确定文字区域，不会由于文字笔画粗度变化造成影响，适用于各种不同的字体，且能够避免图像复杂的像素变化情况对定位造成影响，提高文字区域定位的准确性。

根据本公开的再一些实施例，提出一种计算机可读存储介质，其上存储有计算机程序指令，该指令被处理器执行时实现上文中提到的任意一种文字区域定位方法的步骤。

30 这样的计算机存储介质能够在文字区域定位装置运行过程中利用文字区域中文

字高度相似的特点，根据边缘图像中边缘线之间的距离确定文字区域，适用于各种不同的字体，提高文字区域定位的准确性。通过以下参照附图对本公开的示例性实施例的详细描述，本公开的其它特征及其优点将会变得清楚。

5 附图说明

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

图 1A~图 1C 为相关技术中笔画宽度定位法的示意图，其中，图 1A 为笔画放大效果图，图 1B 为轮廓示意图，图 1C 为笔画宽度计算示意图。

10 图 2 为相关技术中稳定极值区域检测法的示意图。

图 3A 为文字笔画宽度一致的字体示意图。

图 3B 为文字笔画宽度不一致的字体示意图。

图 4 为不适用于稳定极值区域检测法图片的示意图。

图 5 为本公开的文字区域定位方法的一些实施例的流程图。

15 图 6A 为采用本公开的文字区域定位方法的一些实施例的原始图像。

图 6B 为对图 6A 采用本公开的文字区域定位方法时确定的方差图。

图 6C 为对图 6B 采用本公开的文字区域定位方法时确定的边缘图像。

图 6D 为对图 6C 采用本公开的文字区域定位方法时确定的文字区域示意图。

20 图 7 为本公开的文字区域定位方法中在边缘图像中定位文字区域的一些实施例的流程图。

图 8 为本公开的文字区域定位方法中确定边缘点的一些实施例的流程图。

图 9 为本公开的文字区域定位装置的一些实施例的示意图。

图 10 为本公开的文字区域定位装置中文字区域定位模块的一些实施例的示意图。

图 11 为本公开的文字区域定位装置中边缘点确定单元的一些实施例的示意图。

25 图 12 为本公开的文字区域定位装置的另一些实施例的示意图。

图 13 为本公开的文字区域定位装置的又一些实施例的示意图。

具体实施方式

30 现在将参照附图来详细描述本公开的各种示例性实施例。应注意到：除非另外具体说明，否则在这些实施例中阐述的部件和步骤的相对布置、数字表达式和数值不限

制本公开的范围。

同时，应当明白，为了便于描述，附图中所示出的各个部分的尺寸并不是按照实际的比例关系绘制的。

以下对至少一个示例性实施例的描述实际上仅仅是说明性的，决不作为对本公开
5 及其应用或使用的任何限制。

对于相关领域普通技术人员已知的技术、方法和设备可能不作详细讨论，但在适当情况下，所述技术、方法和设备应当被视为授权说明书的一部分。

在这里示出和讨论的所有示例中，任何具体值应被解释为仅仅是示例性的，而不是作为限制。因此，示例性实施例的其它示例可以具有不同的值。

10 应注意到：相似的标号和字母在下面的附图中表示类似项，因此，一旦某一项在一个附图中被定义，则在随后的附图中不需要对其进行进一步讨论。

图 1A~图 1C 为利用笔画宽度定位法进行文字定位的示意图。

图 1A 的灰色区域是文字的一个笔画放大数倍之后的效果，其中灰色的小格子是笔画上的像素，白色的小格子是图像背景。图 1B 中，利用 Canny 算子勾勒出了笔画
15 的两条轮廓，大致可以看出这两条轮廓相互平行，p 和 q 为轮廓两侧相对的点，两点之间的直线距离为 W。图 1C 是在图 1B 的基础上计算轮廓上的像素到与其平行的轮廓上的像素点最小的距离，这个距离就是笔画宽度。

图 2 为相关技术中稳定极值区域检测法的示意图。

如图 2 所示，图像中的字体跟背景的颜色对比明显，通过不断提高二值化阈值，
20 可以获得右边的两张图片，在右边的图片中可以清晰的看到文字信息。

但是，上述相关技术存在各自不同的缺陷。

图 3A 为文字笔画宽度一致的字体示意图。图 3B 为文字笔画宽度不一致的字体示意图。如图 3A 中所示，微软正黑体的文字笔画的宽度是一致的，能够采用笔画宽度
25 定位法进行定位，但是，如图 3B 中所示的宋体文字其笔画宽度并不相同，如笔画撇（丿）就是上半部分宽度比较粗，下边比较细，因此不适用于采用笔画宽度定位法进行定位。而采用稳定极值区域检测法需要文字区域的像素对比度高，但在应用过程中，对比度高的区域不一定是文字，因此该算法很容易引入额外的噪声。

图 4 为不适用于稳定极值区域检测法图片的示意图。

如图 4 所示，矩形框所选中的位置都是稳定极值区域，但是只有不到一半的区域
30 是文字区域。另外，这两种方法都只能先定位文字区域，还需要额外的算法将单个文

字串成行，较为繁琐，降低了运算效率。

基于此，本公开提出一种文字区域定位方案，能够提高对不同字体的适应能力，提高文字区域定位的准确性。

图 5 示出本公开的文字区域定位方法的一些实施例的流程图。如图 5 所示，文字区域定位方法包括步骤 501-503。

在步骤 501 中，根据原始图像获取方差图。在一些实施例中，可以根据原始图像中每个点的像素值获取图像中每个像素点与周围几个像素点的像素值方差，如根据水平连续几个点的像素值取方差，确定其中一个点的方差图像素。通过计算每个点与周围几个点像素值的方差，可以确定方差图。

在步骤 502 中，获取方差图的边缘图像。在一些实施例中，可以采用相关技术中任何一种边缘检测算法计算边缘图像。

在步骤 503 中，当边缘图像中相邻两条边缘线中位置相对的边缘点之间距离的差值在预定距离差范围内时，确定相邻两条边缘线之间的区域为文字区域。在一些实施例中，可以在边缘图像中得到两条近似平行的边缘线，该边缘线可以为直线、曲线，中间可以有断点等。若两条边缘线之间的距离相对稳定，距离变化范围在预定距离差范围内，也可以认为这两条边缘线之间的区域为文字区域。

通过这样的方法，能够利用文字区域中文字高度相似的特点，根据边缘图像中边缘线之间的距离确定文字区域。这样就不会由于文字笔画粗度变化造成影响，能够适用于各种不同的字体，且能够避免图像复杂的像素变化情况对定位造成影响，提高文字区域定位的准确性。另外，由于无需逐字的确定文字区域再进行拼接，而是直接在复杂版面的打印体图片中直接快速的定位文字行，提高了文字区域的确定效率。

在一些实施例中，可以获取原始图像中目标像素位置的像素值、及目标像素位置的邻近像素点的像素值，将原始图像中目标像素位置和邻近像素点的像素值取方差值来确定方差图中目标像素位置的像素值。邻近像素点可以为目标像素位置在不同方向（例如，水平、竖直方向）连续预定数量的像素点。预定数量可以根据经验或者实际需求设定。例如，假定原始图像为 G ，原始图像中坐标位置为 (x, y) 的像素点的像素值为 $G(x, y)$ ，如 $G(0,0)$ 代表图像左上角的像素值。假定方差图为 I ，方差图中坐标位置为 (x, y) 的像素点的像素值为 $I(x, y)$ 。以水平方差图为例， $G(x, y)$ 的邻近像素点包括 $G(x-t,y)$ 、 $G(x-t+1,y)$ 、 $\cdots G(x-1,y)$ 、 $G(x+1,y)$ 、 $\cdots G(x+t,y)$ ，根据公式：

$$I(x,y) = \text{Var}(G(x-t,y), G(x-t+1,y), \dots, G(x,y), G(x+1,y), \dots, G(x+t,y))$$

计算方差图中 (x, y) 点的像素值 $I(x,y)$ 。在公式中, t 的数值可以根据需要或效果设定, 如设置为 5。

对于位于两端的像素点, 如 $G(0,0)$, 可以只根据 $G(0,0)$ 、 $G(1,0)$ …… $G(t,0)$ 确定 $I(0,0)$ 。

对于特殊的应用场合, 还可以计算竖直方差图, 即以竖直方向续预定数量的像素点的像素值确定方差值。还可以设定上下左右预定范围内的像素点为邻近像素点。

通过这样的方法, 能够在原始图像的基础上计算其方差图。方差图能够反应图像的变化情况, 从而获取图像发生剧烈变化的位置, 便于将文字区域与其他图像区域区分。如图 6A 所示为原始图像, 6B 为图 6A 的方差图。从方差图中能够看出, 文字区域呈现明显的长条状, 具有突出的特点。

在一些实施例中, 在方差图的基础上, 可以进一步提取方差图的边缘轮廓, 得到边缘图像。可以利用相关技术中任意一种边缘图像提取算法进行实现, 如采用 Canny 算子计算图像边缘, 得到边缘图像。

通过这样的方法, 能够在方差图的基础上进一步得到方差图的边缘轮廓, 从而便于在边缘图像的基础上进行运算, 得到位于边缘点之间的文字区域。如图 6C 所示, 在图像 6B 的基础上进行边缘轮廓提取, 能够得到 6C 中的边缘图像。图 6C 中的边缘图像线条清楚, 能够便于进行边缘点提取和距离计算, 得到如图 6D 所示的文字区域示意图。

图 7 示出本公开的文字区域定位方法中在边缘图像中定位文字区域的一些实施例的流程图。如图 7 所示, 文字区域定位方法包括步骤 701-703。

在步骤 701 中, 确定第一边缘点和位于相邻边缘线上的第二边缘点。在一些实施例中, 可以遍历边缘图像, 每次取一个边缘点作为第一边缘点, 直至完成整幅图像或整条边缘线上每个边缘点与相对的第二边缘点间的关联确认。在第一边缘点的基础上可以取与第一边缘点所处的边缘线相邻的边缘线上与第一边缘点位置相对的像素点作为第二边缘点。如两条水平的边缘线上下平行, 第一边缘点坐标为 (x, y) , 则第二边缘点坐标为 $(x, y+n)$, 其中, n 为第一、第二边缘点之间的距离。

在步骤 702 中, 根据第一边缘点和第二边缘点之间的距离确定行高。在一些实施例中, 可以遍历整幅图, 得到每个第一边缘点和对应的第二边缘点之间的行高。

在步骤 703 中, 连接行高的差值在预定距离差范围内的相邻的第一边缘点, 确定

第一边缘线，连接行高的差值在预定距离差范围内的相邻的第二边缘点，确定第二边缘线，第一边缘线与第二边缘线之间的区域为文字区域。在一些实施例中，若至少两个第一边缘点相邻、且第一边缘点对应的第二边缘点也相邻，且这些相邻的边缘点中，第一边缘点与第二边缘点之间的距离的差值在预定距离差范围内，则可以认为第一边缘点和第二边缘点分别为文字的上边线点、下边线点（在竖直状态下是左、右边线点）。从而可以将相邻边缘点连接，得到文字的上边线、文字的下边线（在竖直状态下是左、右边线），边线之间的区域即为文字区域。

通过这样的方法，能够在边缘图像的基础上得到文字的边线，从而得到文字区域，由于无需进行单个文字的判断，降低了运算量，且不受笔画粗度不同、像素值有较大差异的不规则区域的影响，提高了文字区域定位的效率和准确度。

图 8 示出本公开的文字区域定位方法中确定边缘点的一些实施例的流程图。如图 8 所示，文字区域定位方法包括步骤 801-803。

在步骤 801 中，在边缘图像中取一点作为第一边缘点。在一些实施例中，可以遍历边缘图像，每次取一个边缘点作为第一边缘点，直至确定整幅图像中每个边缘点对应的第二边缘点，或完成确定整条边缘线上每个边缘点的第二边缘点。

在步骤 802 中，从第一边缘点沿像素梯度的方向发射射线，直至下一边缘点，从而能够找到与第一边缘点所处的边缘线相邻的边缘线上与第一边缘点位置相对的点。

在步骤 803 中，若第一边缘点与下一边缘点法向量的夹角小于预定角度阈值，则确定下一边缘点为第二边缘点。在一些实施例中，预定角度阈值可以为 30 度。像素的法向量即像素的梯度或像素的导数，数字图像作为离散的点值谱，可以称为二维离散函数，可以通过对该二维离散函数求导确定其法向量。

通过这样的方法，能够在边缘图像的基础上确定第一边缘点和与其位置相对的第二边缘点，为计算第一边缘点与第二边缘点的距离提供了数据基础。由于采用沿像素梯度的方向发射射线的方式确定第二边缘点，得到的第二边缘点与第一边缘点的相对位置不确定，根据像素变化情况可能是上下、左右或其他位置关系，从而能够确定横向文字区域、纵向文字区域、倾斜文字区域，甚至是扇形文字区域，防止不规则排版造成的漏定位，提高文字定位的准确度。

图 9 示出本公开的文字区域定位装置的一些实施例示意图。如图 9 所示，文字区域定位装置包括方差图确定模块 901、边缘图像获取模块 902 和文字区域定位模块 903。

方差图确定模块 **901** 能够根据原始图像获取方差图。在一些实施例中，可以根据原始图像中每个点的像素值获取图像中每个像素点与周围几个像素点的像素值方差，如根据水平连续几个点的像素值取方差，确定其中一个点的方差图像素。通过计算每个点与周围几个点像素值的方差，确定方差图。

5 在一些实施例中，方差图确定模块 **901** 能够获取原始图像中目标像素位置的像素值，及目标像素位置的邻近像素点的像素值，将原始图像中目标像素位置和邻近像素点的像素值取方差值来确定方差图中目标像素位置的像素值。邻近像素点可以为目标像素位置水平、竖直方向连续预定数量，或上下左右预定范围内的像素点。预定数量可以根据经验或者实际需求设定。

10 通过这样的方法，能够在原始图像的基础上计算其方差图，方差图能够反应图像的变化情况，从而获取图像发生剧烈变化的位置，便于将文字区域与其他图像区域区分。

边缘图像获取模块 **902** 能够获取方差图的边缘图像。在一些实施例中，可以采用相关技术中任意一种边缘检测算法计算边缘图像。

15 在一些实施例中，边缘图像获取模块 **902** 能够在方差图的基础上进一步提取方差图的边缘轮廓，得到边缘图像。可以利用相关技术中任意一种边缘图像提取算法进行实现，如采用 Canny 算子计算图像边缘，得到边缘图像。

通过这样的方法，能够在方差图的基础上进一步得到方差图的边缘轮廓，从而便于在边缘图像的基础上进行运算，得到位于边缘点之间的文字区域。

20 文字区域定位模块 **903** 能够在边缘图像中相邻两条边缘线中位置相对的边缘点之间距离的差值在预定距离差范围内时，确定相邻两条边缘线之间的区域为文字区域。在一些实施例中，可以在边缘图像中得到两条近似平行的边缘线，该边缘线可以为直线、曲线，中间可以有断点等，若两条边缘线之间的距离相对稳定，距离变化范围在预定距离差范围内，则可以认为这两条边缘线之间的区域为文字区域。

25 这样的装置能够利用文字区域中文字高度相似的特点，根据边缘图像中边缘线之间的距离确定文字区域，不会由于文字笔画粗度变化造成影响，适用于各种不同的字体，且能够避免图像复杂的像素变化情况对定位造成影响，提高文字区域定位的准确性。

图 10 进一步示出本公开的文字区域定位装置中文字区域定位模块的一些实施例的示意图。如图 10 所示，文字区域定位模块边缘点确定单元 **1001**、行高确定单元 **1002**

30

和边缘线连接单元 1003。

边缘点确定单元 1001 能够确定第一边缘点和位于相邻边缘线上的第二边缘点。在一些实施例中，可以遍历边缘图像，每次取一个边缘点作为第一边缘点，直至完成整幅图像或整条边缘线上每个边缘点与相对的第二边缘点间关联关系的确认。在第一边缘点的基础上可以取与第一边缘点所处的边缘线相邻的边缘线上与第一边缘点位置相对的像素点作为第二边缘点。如两条水平的边缘线上下平行，第一边缘点坐标为 (x, y) ，则第二边缘点坐标为 $(x, y+n)$ ，其中， n 为第一、第二边缘点之间的距离。

行高确定单元 1002 能够根据第一边缘点和第二边缘点之间的距离确定行高。在一些实施例中，可以遍历整幅图，得到每个第一边缘点和对应的第二边缘点之间的行高。

边缘线连接单元 1003 连接行高的差值在预定距离差范围内的相邻的第一边缘点，确定第一边缘线，连接行高的差值在预定距离差范围内的相邻的第二边缘点，确定第二边缘线，第一边缘线与第二边缘线之间的区域为文字区域。

在一些实施例中，若至少两个第一边缘点相邻、且第一边缘点对应的第二边缘点也相邻，且这些相邻的边缘点中，第一边缘点与第二边缘点之间的距离的差值在预定距离差范围内，则可以认为第一边缘点和第二边缘点分别为文字的上边线点、下边线点（在竖直状态下是左、右边线点）。从而可以将相邻边缘点连接，得到文字的上边线、文字的下边线（在竖直状态下是左、右边线），边线之间的区域即为文字区域。

这样的装置能够在边缘图像的基础上得到文字的边线，从而得到文字区域，由于无需进行单个文字的判断，降低了运算量，且不受笔画粗度不同、像素值有较大差异的不规则区域的影响，提高了文字区域定位的效率和准确度。

图 11 进一步示出本公开的文字区域定位装置中边缘点确定单元的一些实施例的示意图。如图 11 所示，边缘点确定单元包括第一边缘点指定子单元 1101、下一边缘点获取子单元 1102 和第二边缘点确定子单元 1103。

第一边缘点指定子单元 1101 能够在边缘图像中取一点作为第一边缘点。在一些实施例中，可以遍历边缘图像，每次取一个边缘点作为第一边缘点，直至确定整幅图像中每个边缘点对应的第二边缘点，或完成确定整条边缘线上每个边缘点的第二边缘点。

下一边缘点获取子单元 1102 能够从第一边缘点沿像素梯度的方向发射射线，直

至下一边缘点，从而能够找到与第一边缘点所处的边缘线相邻的边缘线上与第一边缘点位置相对的点。

第二边缘点确定子单元 **1103** 能够当第一边缘点与下一边缘点法向量的夹角小于预定角度阈值时，确定下一边缘点为第二边缘点。在一些实施例中，预定角度阈值可以

以为 **30** 度。

这样的装置能够在边缘图像的基础上确定第一边缘点和与其位置相对的第二边缘点，为计算第一边缘点与第二边缘点的距离提供了数据基础。由于采用沿像素梯度的方向发射射线的方式确定第二边缘点，得到的第二边缘点与第一边缘点的相对位置不确定，根据像素变化情况可能是上下、左右或其他位置关系。这样能够确定横向文字区域、纵向文字区域、倾斜文字区域，甚至是扇形文字区域，防止不规则排版造成的漏定位，提高文字定位的准确度。

图 **12** 示出本公开文字区域定位装置的另一些实施例的结构示意图。如图 **12** 所示，文字区域定位装置包括存储器 **1210** 和处理器 **1220**。

存储器 **1210** 可以是磁盘、闪存或其它任何非易失性存储介质。存储器用于存储文字区域定位方法的对应实施例中的指令，包括仿真平台侧指令，也可以包括管理系统侧指令。

处理器 **1220** 耦接至存储器 **1210**，可以作为一个或多个集成电路来实施，例如微处理器或微控制器。该处理器 **1220** 用于执行存储器中存储的指令，能够实现文字区域的定位。

在一些实施例中，还可以如图 **13** 所示，文字区域定位装置 **1300** 包括存储器 **1310** 和处理器 **1320**。处理器 **1320** 通过 **BUS** 总线 **1330** 耦合至存储器 **1310**。该文字区域定位装置 **1300** 还可以通过存储接口 **1040** 连接至外部存储装置 **1350** 以便调用外部数据，还可以通过网络接口 **1360** 连接至网络或者另外一台计算机系统（未标出）。此处不再进行详细介绍。

在该实施例中，通过存储器存储数据指令，再通过处理器处理上述指令，能够实现文字区域定位装置的运转。

在另一些实施例中，一种计算机可读存储介质，其上存储有计算机程序指令，该指令被处理器执行时实现文字区域定位方法对应实施例中的方法的步骤。本领域内的技术人员应明白，本公开的实施例可提供为方法、装置、或计算机程序产品。因此，本公开可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的

形式。而且，本公开可采用在一个或多个其中包含有计算机可用程序代码的计算机可用非瞬时性存储介质（包括但不限于磁盘存储器、**CD-ROM**、光学存储器等）上实施的计算机程序产品的形式。

本公开是参照根据本公开实施例的方法、设备（系统）和计算机程序产品的流程图和 / 或方框图来描述的。应理解可由计算机程序指令实现流程图和 / 或方框图中的每一流程和 / 或方框以及流程图和 / 或方框图中的流程和 / 或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器以产生一个机器，使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和 / 或方框图一个方框或多个方框中指定的功能的装置。

这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中，使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品，该指令装置实现在流程图一个流程或多个流程和 / 或方框图一个方框或多个方框中指定的功能。

这些计算机程序指令也可装载到计算机或其他可编程数据处理设备上，使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理，从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和 / 或方框图一个方框或多个方框中指定的功能的步骤。

至此，已经详细描述了本公开。为了避免遮蔽本公开的构思，没有描述本领域所公知的一些细节。本领域技术人员根据上面的描述，完全可以明白如何实施这里公开的技术方案。

可能以许多方式来实现本公开的方法以及装置。例如，可通过软件、硬件、固件或者软件、硬件、固件的任何组合来实现本公开的方法以及装置。用于所述方法的步骤的上述顺序仅是为了进行说明，本公开的方法的步骤不限于以上具体描述的顺序，除非以其它方式特别说明。此外，在一些实施例中，还可将本公开实施为记录在记录介质中的程序，这些程序包括用于实现根据本公开的方法的机器可读指令。因而，本公开还覆盖存储用于执行根据本公开的方法的程序的记录介质。

最后应当说明的是：以上实施例仅用以说明本公开的技术方案而非对其限制；尽管参照较佳实施例对本公开进行了详细的说明，所属领域的普通技术人员应当理解：依然可以对本公开的具体实施方式进行修改或者对部分技术特征进行等同替换；而不

脱离本公开技术方案的精神，其均应涵盖在本公开请求保护的技术方案范围当中。

权 利 要 求

1. 一种文字区域定位方法，包括：

根据原始图像获取方差图；

获取所述方差图的边缘图像；

当所述边缘图像中相邻两条边缘线中位置相对的边缘点之间距离的差值在预定距离差范围内时，确定所述相邻两条边缘线之间的区域为文字区域。

2. 根据权利要求 1 所述的文字区域定位方法，其中，所述确定所述相邻两条边缘点之间的区域为文字区域包括：

确定第一边缘点和位于相邻边缘线上的第二边缘点；

根据所述第一边缘点和所述第二边缘点之间的距离确定行高；

连接行高的差值在预定距离差范围内的相邻的所述第一边缘点，以确定第一边缘线，连接行高的差值在预定距离差范围内的相邻的所述第二边缘点，以确定第二边缘线，所述第一边缘线与所述第二边缘线之间的区域为所述文字区域。

3. 根据权利要求 2 所述的文字区域定位方法，所述确定第一边缘点和位于相邻边缘线上的第二边缘点包括：

在所述边缘图像中取一点作为所述第一边缘点；

从所述第一边缘点沿像素梯度的方向发射射线，直至下一边缘点；

当所述第一边缘点与所述下一边缘点法向量的夹角小于预定角度阈值时，确定所述下一边缘点为所述第二边缘点。

4. 根据权利要求 1 所述的文字区域定位方法，所述根据原始图像获取方差图包括：

获取原始图像中目标像素位置的像素值、及所述目标像素位置的邻近像素点的像素值，其中，所述邻近像素点为所述目标像素位置在不同方向连续预定数量的像素点；

将所述原始图像中所述目标像素位置和所述邻近像素点的像素值取方差值，确定所述方差图中所述目标像素位置的像素值。

5. 根据权利要求 1 所述的文字区域定位方法, 其中, 所述获取所述方差图的边缘图像包括: 基于所述方差图利用 Canny 算子计算图像边缘, 获取所述边缘图像。

6. 根据权利要求 1 所述的文字区域定位方法, 其中, 所述文字区域包括横向文字区域、纵向文字区域、倾斜文字区域和扇形文字区域中的至少一种。

7. 一种文字区域定位装置, 包括:

方差图确定模块, 用于根据原始图像获取方差图;

边缘图像获取模块, 用于获取所述方差图的边缘图像;

文字区域定位模块, 用于当所述边缘图像中相邻两条边缘线中位置相对的边缘点之间距离的差值在预定距离差范围内时, 确定所述相邻两条边缘线之间的区域为文字区域。

8. 根据权利要求 7 所述的文字区域定位装置, 其中, 所述文字区域定位模块包括:

边缘点确定单元, 用于确定第一边缘点和位于相邻边缘线上的第二边缘点;

行高确定单元, 用于根据所述第一边缘点和所述第二边缘点之间的距离确定行高;

边缘线连接单元, 用于连接行高的差值在预定距离差范围内的相邻的所述第一边缘点, 以确定第一边缘线, 连接行高的差值在预定距离差范围内的相邻的所述第二边缘点, 以确定第二边缘线, 所述第一边缘线与所述第二边缘线之间的区域为所述文字区域。

9. 根据权利要求 8 所述的文字区域定位装置, 其中, 所述边缘点确定单元包括:

第一边缘点指定子单元, 用于在所述边缘图像中取一点作为所述第一边缘点;

下一边缘点获取子单元, 用于从所述第一边缘点沿像素梯度的方向发射射线, 直至下一边缘点;

第二边缘点确定子单元, 用于当所述第一边缘点与所述下一边缘点法向量的夹角小于预定角度阈值时, 确定所述下一边缘点为所述第二边缘点。

10. 根据权利要求 7 所述的文字区域定位装置, 其中, 所述方差图确定模块用于:

获取原始图像中目标像素位置的像素值，及所述目标像素位置的邻近像素点的像素值，其中，所述邻近像素点为所述目标像素位置在不同方向连续预定数量的像素点；

将所述原始图像中所述目标像素位置和所述邻近像素点的像素值取方差值，确定所述方差图中所述目标像素位置的像素值。

11. 根据权利要求 7 所述的文字区域定位装置，其中，所述边缘图像获取模块用于基于所述方差图利用 Canny 算子计算图像边缘，获取所述边缘图像。

12. 根据权利要求 7 所述的文字区域定位装置，其中，所述文字区域包括横向文字区域、纵向文字区域、倾斜文字区域和扇形文字区域中的至少一种。

13. 一种文字区域定位装置，包括：

存储器；以及

耦接至所述存储器的处理器，所述处理器被配置为基于存储在所述存储器的指令执行如权利要求 1 至 6 任一项所述的文字区域定位方法。

14. 一种计算机可读存储介质，其上存储有计算机程序指令，该指令被处理器执行时实现权利要求 1 至 6 任意一项所述的文字区域定位方法。

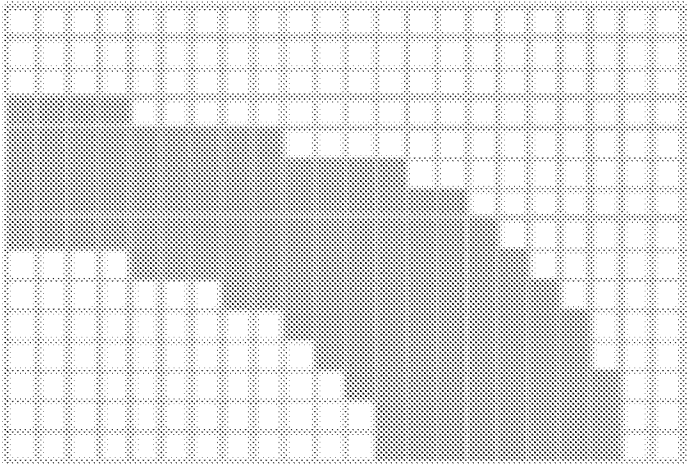


图 1A

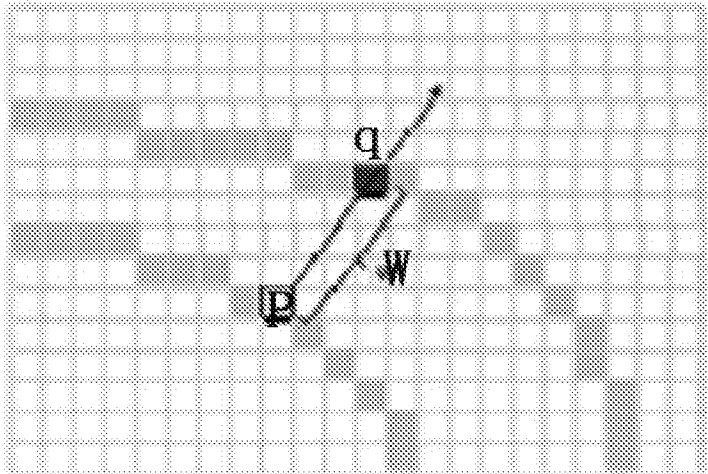


图 1B

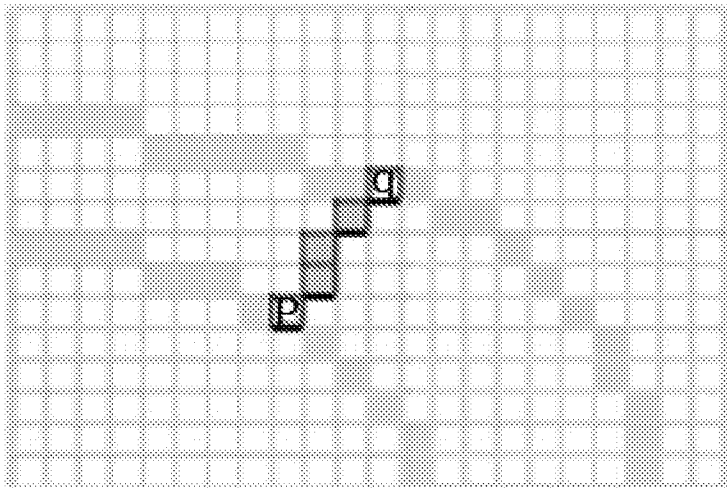


图 1C



图 2

微軟正黑體
Microsoft JhengHei

图 3A

小标宋体: 嘉善语言文字网

华文中宋: 嘉善语言文字网

图 3B



图 4

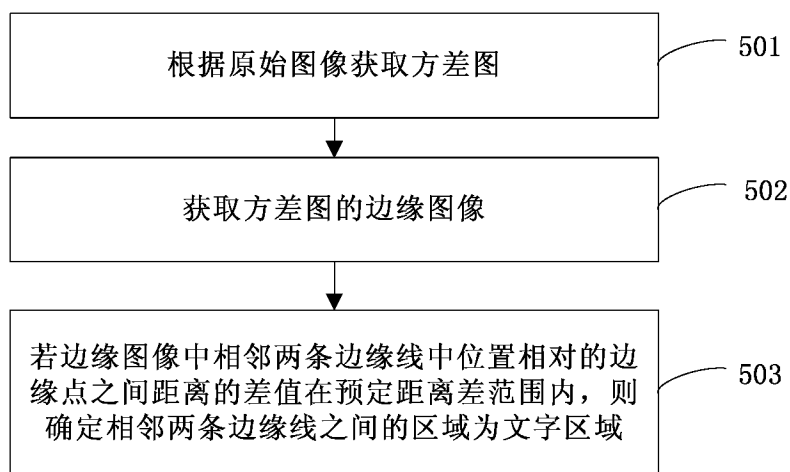


图 5



图 6A

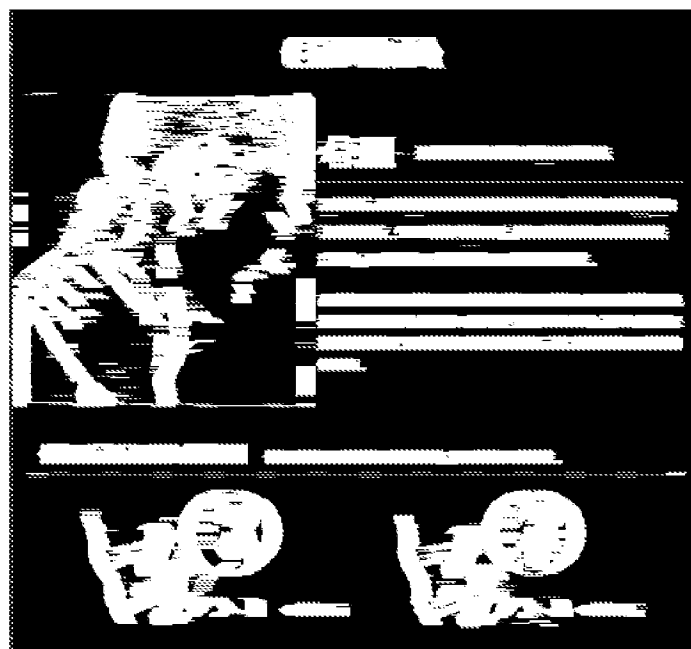


图 6B

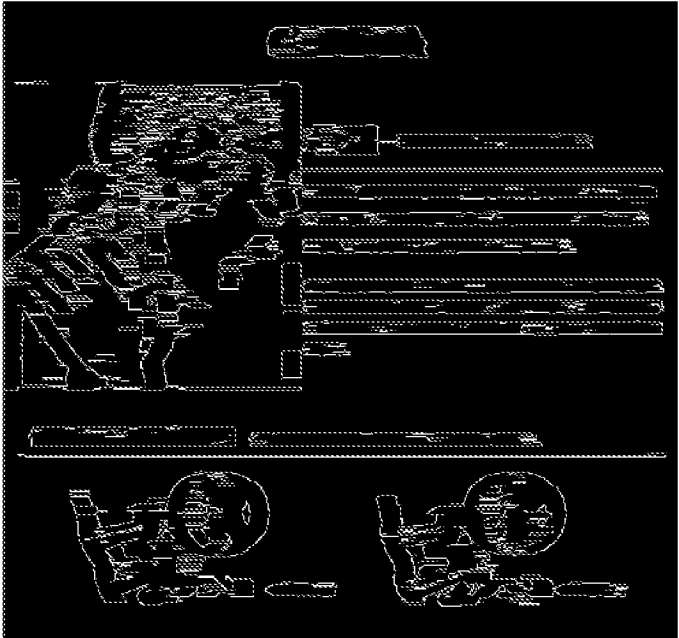


图 6C

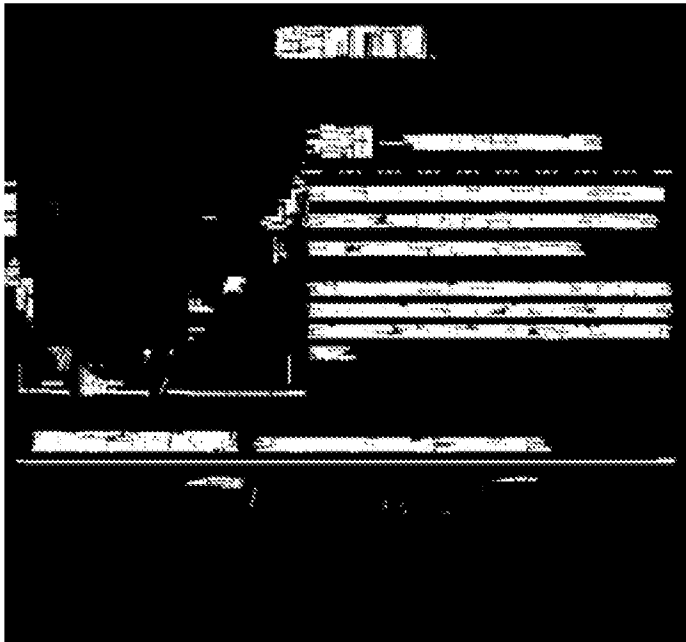


图 6D

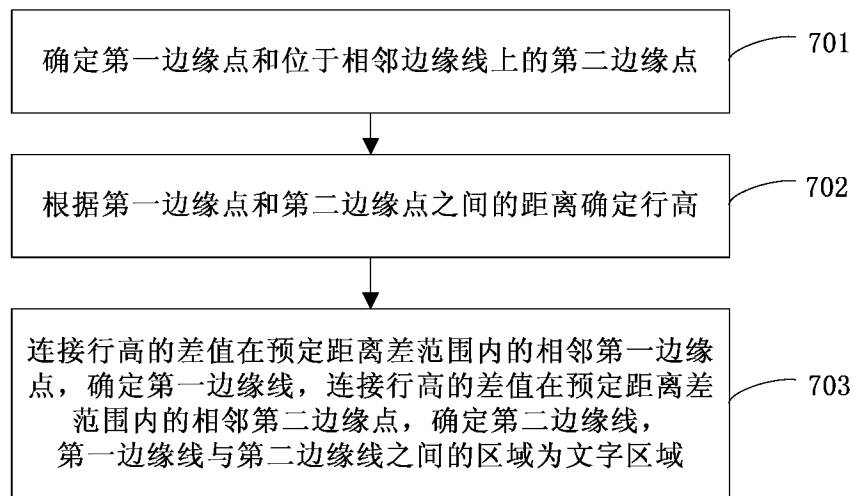


图 7

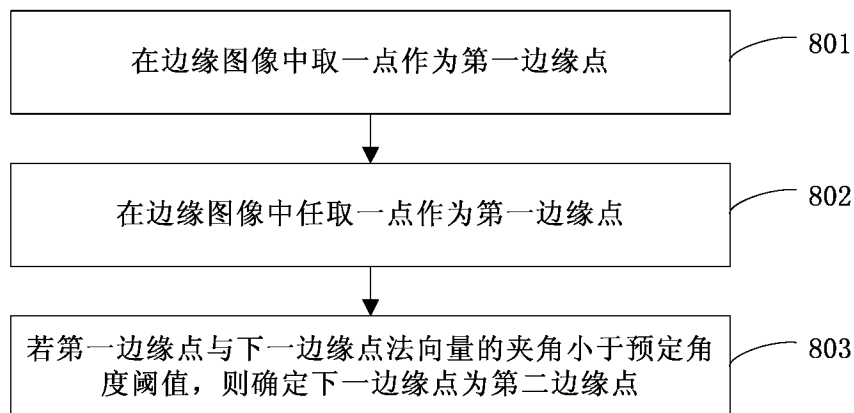


图 8



图 9



图 10

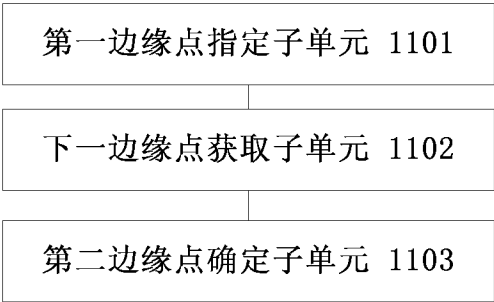


图 11

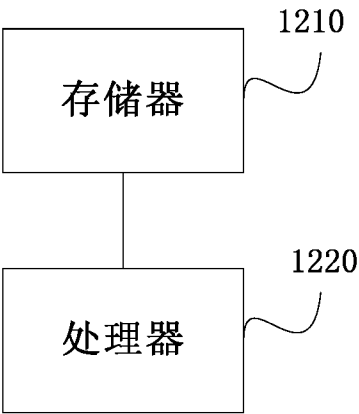


图 12

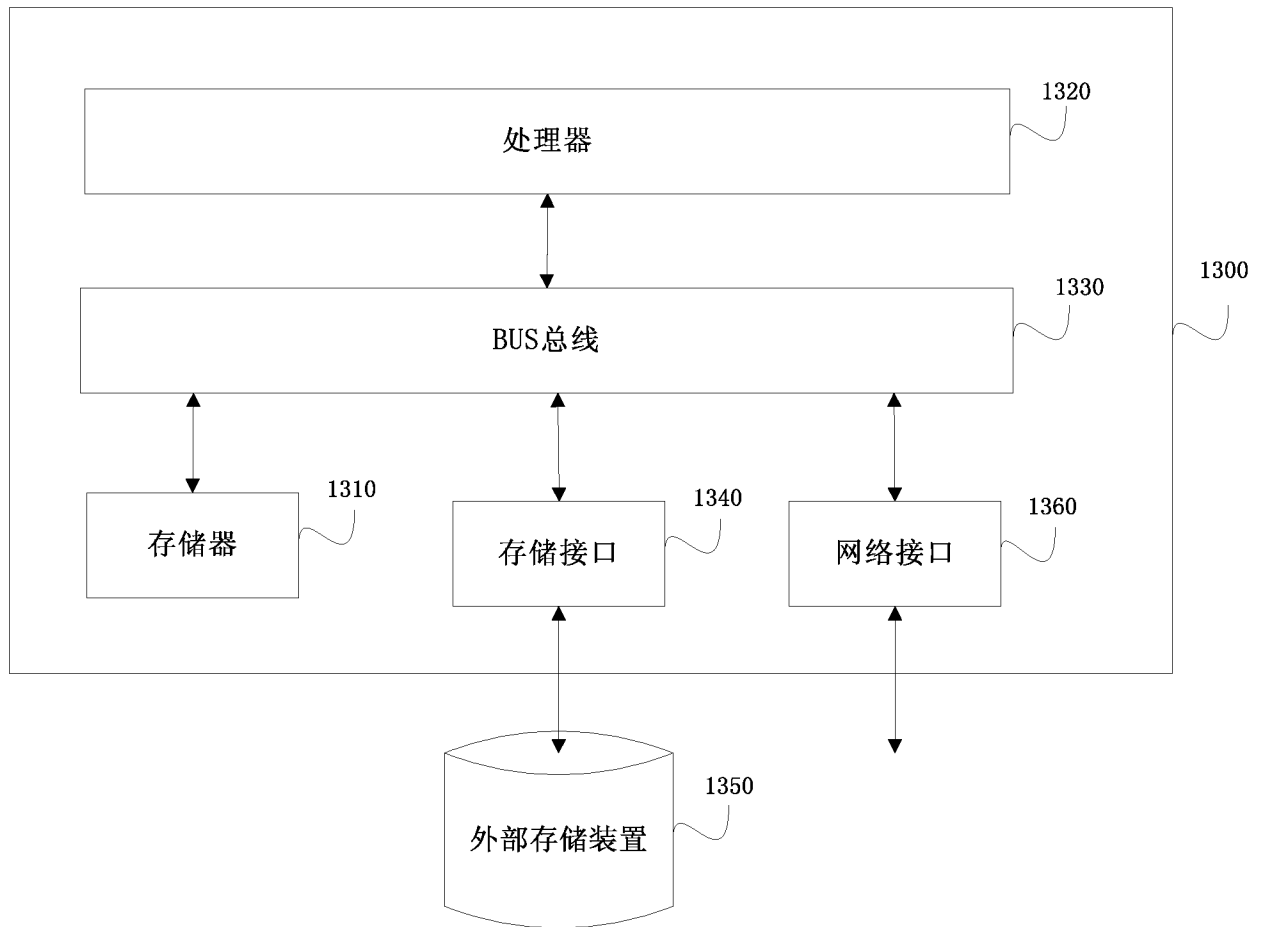


图 13

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2017/119692

A. CLASSIFICATION OF SUBJECT MATTER

G06K 9/20 (2006.01) i

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)

G06K

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)

WPI, EPODOC, CNKI, CNABS, IEEE: 文字, 区域, 文字宽度, 边缘点, 边缘线, 方差图, 图像, 梯度, character, region, width, variance, image, gradient

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 104794479 A (PEKING UNIVERSITY) 22 July 2015 (22.07.2015), description, paragraphs [0004]-[0068]	1-14
A	CN 105718926 A (SHARP K.K.) 29 June 2016 (29.06.2016), entire document	1-14
A	CN 104751142 A (UNIVERSITY OF ELECTRONIC SCIENCE AND TECHNOLOGY OF CHINA) 01 July 2015 (01.07.2015), entire document	1-14
A	CN 103324927 A (PFU CORP.) 25 September 2013 (25.09.2013), entire document	1-14
A	梁添才等. 基于笔画生长的自然场景艺术文字检测. 计算机仿真. 31 August 2015 (31.08.2015), 32(08), entire document. (LIANG, Tiancai et al. Wordart Detection in Natural Scene Based on Stroke Growing. Computer Simulation.)	1-14

☐ Further documents are listed in the continuation of Box C. ☒ See patent family annex.

* Special categories of cited documents:	“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
“A” document defining the general state of the art which is not considered to be of particular relevance	“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
“E” earlier application or patent but published on or after the international filing date	“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
“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)	“&” document member of the same patent family
“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	

Date of the actual completion of the international search 23 March 2018	Date of mailing of the international search report 04 April 2018
Name and mailing address of the ISA 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 ZHANG, Mingxia Telephone No. (86-10) 53961517

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2017/119692

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 104794479 A	22 July 2015	None	
CN 105718926 A	29 June 2016	WO 2016086877 A1	09 June 2016
		JP 2017535891 A	30 November 2017
CN 104751142 A	01 July 2015	None	
CN 103324927 A	25 September 2013	JP 2013196369 A	30 September 2013
		US 2013243321 A1	19 September 2013

国际检索报告

国际申请号

PCT/CN2017/119692

A. 主题的分类

G06K 9/20(2006.01)i

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)

G06K

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))

WPI, EPODOC, CNKI, CNABS, IEEE: 文字, 区域, 文字宽度, 边缘点, 边缘线, 方差图, 图像, 梯度, character, region, width, variance, image, gradient

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 104794479 A (北京大学) 2015年 7月 22日 (2015 - 07 - 22) 说明书第0004-0068段	1-14
A	CN 105718926 A (夏普株式会社) 2016年 6月 29日 (2016 - 06 - 29) 全文	1-14
A	CN 104751142 A (电子科技大学) 2015年 7月 1日 (2015 - 07 - 01) 全文	1-14
A	CN 103324927 A (株式会社PFU) 2013年 9月 25日 (2013 - 09 - 25) 全文	1-14
A	梁添才 等, . “基于笔画生长的自然场景艺术文字检测, ” 计算机仿真, , 第32卷, 第08期, 2015年 8月 31日 (2015 - 08 - 31), 全文	1-14

☐ 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 3月 23日

国际检索报告邮寄日期

2018年 4月 4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

张明霞

传真号 (86-10)62019451

电话号码 (86-10)53961517

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2017/119692

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	104794479	A	2015年 7月 22日	无			
CN	105718926	A	2016年 6月 29日	WO	2016086877	A1	2016年 6月 9日
				JP	2017535891	A	2017年 11月 30日
CN	104751142	A	2015年 7月 1日	无			
CN	103324927	A	2013年 9月 25日	JP	2013196369	A	2013年 9月 30日
				US	2013243321	A1	2013年 9月 19日