

[19] 中华人民共和国国家知识产权局



[12] 发明专利申请公布说明书

[21] 申请号 200810058451.5

[51] Int. Cl.

G06K 9/00 (2006.01)

G06K 9/46 (2006.01)

[43] 公开日 2008 年 12 月 3 日

[11] 公开号 CN 101315664A

[22] 申请日 2008.5.27

[21] 申请号 200810058451.5

[71] 申请人 昆明理工大学

地址 650093 云南省昆明市五华区学府路 253 号

[72] 发明人 邵玉斌

[74] 专利代理机构 昆明正原专利代理有限责任公司
代理人 金耀生

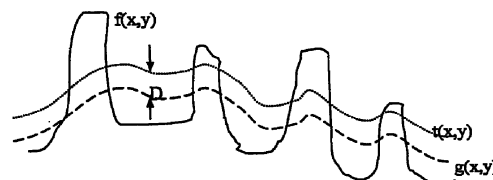
权利要求书 2 页 说明书 5 页 附图 1 页

[54] 发明名称

用于文字识别的文本图像预处理方法

[57] 摘要

本发明是用于文字识别的文本图像预处理方法。包含对文本图像的几何校正以及动态域值二值化两个步骤。所述的几何校正方法对相机摄取的文字图片的文字区域几何失真进行区域校正，得到矩形区域的校正结果；几何校正方法包含自动区域识别和手动区域指定两种。所述的动态域值二值化方法将相机摄取的文字图片中的文字前景和亮度不均的背景自适应地分离出来。本方法的特征是根据所识别的文字区域进行矩形化校正，并可结合任意的图像模糊算法，以源图像作为参考图像通过对比计算，从而将亮度不均的背景中的文字前景分离出来。本方法所得出的结果图像可用于计算机文字识别、机器视觉和机器理解等领域。本发明具有对环境适应性强、算法稳定可靠的优良效果。



1. 一种用于文字识别的文本图像预处理方法，其特征在于：对照相机摄取的文字图片的文字区域几何失真进行区域校正，得到矩形区域的校正结果；并将照相机摄取的文字图片中的文字前景和亮度不均的背景自适应地分离出来。

2. 根据权利要求 1 所述的用于文字识别的文本图像预处理方法，其特征在于：对照相机摄取的文字图片的文字区域几何失真进行区域校正的方法是：首先对照相机摄取的具有几何失真的文本图像进行区域自动识别或指定，得出文本区域的边界位置，再利用文本图像文字排列的固有矩形区域特性，对所识别或指定的文本区域进行几何校正，还原为矩形形状的文本区域；对于文本图像中的任何一个像素，根据它与所指定的文本区域的边界位置的关系，求取其在设定矩形区域中的对应位置，从而得到该像素在校正后图像中的本源位置。校正后图像的所有像素的色彩值或亮度值均以其所对应的源文字图片中的像素本源位置作为参照，通过任意一种图像插值算法确定；具体步骤如下：

1) 对原本为矩形文字区域的图像进行文字区域识别，得出文本的边界；

2) 以区域识别结果为基准，计算源图像相对于该基准下的位置关系；

3) 通过任意一种图像插值算法确定校正后图像对应像素的色彩值或亮度值，从而得出几何校正的结果；

4) 对步骤 3 所得出的几何校正的结果还可重复步骤 1、步骤 2 和步骤 3，通过迭代得到更好的几何校正的结果。

3. 根据权利要求 1 所述的用于文字识别的文本图像预处理方法，其特征在于：对照相机摄取的文字图片或由权利要求 1 所述的文字区域几何失真校正方法所得出的结果图像进行文字前景的提取，将文字前景和亮度不均的背景自适应地分离出来的方法是图像动态域值二值化预处理方法：通过图像中任意像素及其邻域像素之间的运算，可使用但不限于使用任意一种图像模糊方法，求取局部背景亮度；通过利用局部背景亮度的加权值作为该区域的二值化门限，对该区域文字前景进行分离；根据分离结果，将源图像中的文字前景部分扣除，再次使用但不限于使用任意一种图像模糊方法计算局部背景亮度，然后利用其加权值作为新的动态门限对源图像进行二值化，得出更精确的分离结果；这一过程可以进行多次迭代，具体步骤如下：

1) 采用固定域值方法对文本前景进行粗分离；

2) 利用任意一种图像模糊方法，对粗分离出来的背景区域求取其局部背景亮度；

-
- 3) 对局部背景亮度进行加权计算, 以之作为动态域值对原图像进行文献前景分离;
 - 4) 利用步骤 3 所得出的结果, 重复迭代步骤 2、3 得出更精确的分离结果。
4. 根据权利要求 1 所述的用于文字识别的文本图像预处理方法, 其特征在于几何校正方法包含自动区域识别和手动区域指定两种。

用于文字识别的文本图像预处理方法

技术领域

本发明涉及图像处理技术领域，具体地说是用于文字识别的文本图像预处理方法。

背景技术

在文字识别、计算机视觉和机器理解等领域的应用中，对复杂环境图像中的文字内容进行分离和识别是技术难点之一。机器理解文字是基于文字识别结果的，而用于文字识别的源图像的质量越好，识别率就越高。文本图像预处理方法的目的是为文字识别算法提供一种几何失真小、前景文本分离完善的二值化图像。用于文字识别的图像预处理方法一般是对图像中的文字内容区域进行识别，然后采用图像平移、旋转和伸缩等方法对文字内容区域进行校正，然后对校正结果进行全局固定域值的二值化处理。

目前，针对文本内容的图像几何校正方法主要集中于对文本图像的倾斜校正上，根据文字走向和行间空白将文本校正为水平或垂直方向的。然而，这种倾斜校正方法仅对于倾斜的矩形区域文本有效。对于实际摄像得到的图像，由于透视关系和镜头非线性失真，原为矩形的文字区域将发生失真变为不规则四边形或曲四边形。目前没有较为便捷的针对文字图片失真为不规则四边形或曲四边形的几何校正方法。针对亮度背景变化的文本图像的二值化问题，采用全局固定域值方法不足之处是不能取得好的效果，而采用分块图像局部二值化域值方法会导致分块边界效应。在文字识别预处理领域，目前尚无一种实用软件能够对照相机摄影的文字图片进行有效的几何校正和文字前景的提取。

发明内容

本发明的目的在于为文字识别和机器理解提供一种方便实用的用于文字识别的文本图像预处理方法文字图片预处理方法，利用此方法，可根据照相机拍摄的文字图片，进行文字区域的识别、并将识别区域还原为矩形区域，通过自适应域值的二值化方法实现文字前景与亮度变化背景之间分离，为文字识别方法提供良好的图像源。

本发明解决的主要技术问题是采用以下技术方案来实现的：

对照相机摄取的文字图片的文字区域几何失真进行区域校正，得到矩形区域的校正结果；并将照相机摄取的文字图片中的文字前景和亮度不均的背景自适应地分离出来。

对照相机摄取的文字图片的文字区域几何失真进行区域校正的方法是：首先对照相机

摄取的具有几何失真的文本图像进行区域自动识别或指定，得出文本区域的边界位置，再利用文本图像文字排列的固有矩形区域特性，对所识别或指定的文本区域进行几何校正，还原为矩形形状的文本区域；对于文本图像中的任何一个像素，根据它与所指定的文本区域的边界位置的关系，求取其在设定矩形区域中的对应位置，从而得到该像素在校正后图像中的本源位置。校正后图像的所有像素的色彩值或亮度值均以其所对应的源文字图片中的像素本源位置作为参照，通过任意一种图像插值算法确定；具体步骤如下：

- 1) 对原本为矩形文字区域的图像进行文字区域识别，得出文本的边界；
- 2) 以区域识别结果为基准，计算源图像相对于该基准下的位置关系；
- 3) 通过任意一种图像插值算法确定校正后图像对应像素的色彩值或亮度值，从而得出几何校正的结果；
- 4) 对步骤3所得出的几何校正的结果还可重复步骤1、步骤2和步骤3，通过迭代得到更好的几何校正的结果。

对照相机摄取的文字图片或由上述的文字区域几何失真校正方法所得出的结果图像进行文字前景的提取，将文字前景和亮度不均的背景自适应地分离出来的方法是图像动态域值二值化预处理方法：通过图像中任意像素及其邻域像素之间的运算，可使用但不限于使用任意一种图像模糊方法，求取局部背景亮度；通过利用局部背景亮度的加权值作为该区域的二值化门限，对该区域文字前景进行分离；根据分离结果，将源图像中的文字前景部分扣除，再次使用但不限于使用任意一种图像模糊方法计算局部背景亮度，然后利用其加权值作为新的动态门限对源图像进行二值化，得出更精确的分离结果；这一过程可以进行多次迭代，具体步骤如下：

- 1) 采用固定域值方法对文本前景进行粗分离；
- 2) 利用任意一种图像模糊方法，对粗分离出来的背景区域求取其局部背景亮度；
- 3) 对局部背景亮度进行加权计算，以之作为动态域值对原图像进行文献前景分离；
- 4) 利用步骤3所得出的结果，重复迭代步骤2、3得出更精确的分离结果。

本发明的方法与现有技术比较具有的优点是：由于本发明采取了上述的技术措施，与现有技术方法相比，具有对摄影环境适应性强、算法简单、稳定可靠的优良效果，此外，还设计了迭代方法以改进结果图像的质量，应用范围更广。实践中，利用本方法的文本图像预处理效果明显优于现有的文字识别预处理方法的结果。本发明可以采用软件实现，形成为计算机文字识别软件进行图像预处理的软件部分或模块，也可以通过硬件或数字信号

处理芯片实现，成为数码照相机、摄像机、机器人视觉系统等嵌入式系统的一个功能。

本发明的具体实施方式由以下实施例及其附图详细给出。

附图说明

图 1 是线性失真的被校正的文字图像的文字区域示意图（该区域对应的不失真文字区域是一个矩形）。

其中： P_{11} , P_{1n} , P_{m1} , P_{mn} 分别为该不规则四边形区域的四个顶点， P_{1k} , P_{21} , P_{in} , P_{ij} 为图像区域中的一些不同位置上的像素。

图 2 是对图 1 文字图像的校正结果区域示意图。

其中： P_{11}' , P_{1n}' , P_{m1}' , P_{mn}' 分别为对应图 1 中 P_{11} , P_{1n} , P_{m1} , P_{mn} 四顶点的校正位置结果； P_{1k}' , P_{21}' , P_{in}' , P_{ij}' 对应图 1 中的 P_{1k} , P_{21} , P_{in} , P_{ij} 点。

图 3 是非线性失真情况下的被校正的文字图像的文字区域示意图（该区域对应的不失真文字区域是一个矩形）。

其中， P_{ij} 表示该失真区域中的一个像素点。

图 4 是经过垂直方向校正之后的文字区域示意图。

其中像素点 P_{ij}' 对应图 3 中的像素点 P_{ij} 。

图 5 是图 4 区域进一步经过水平校正之后的文字区域示意图。

其中像素点 P_{ij}'' 对应图 4 中的像素点 P_{ij}' ，也即对应图 3 中的像素点 P_{ij} 。

图 6 是动态域值二值化计算方法示意图的一维表示。

其中，曲线 f 表示二维图像 $f(x,y)$ ；曲线 g 表示经过邻近像素之间运算得出的模糊图像 $g(x,y)$ ，曲线 t 表示模糊图像 $g(x,y)$ 经过平移加权得出的图像 $t(x,y)$ ， D 为平移加权值。

具体实施方式

以下结合附图实施例，对本发明的文本图像预处理方法作进一步的详述。

实施例

1. 对于线性几何失真的不规则的四边形的校正

如图 1 所示，被校正图像中的文字区域发生了线性失真，原本为矩形区域的文字部分失真为一个不规则四边形区域。由于文字区域部分与文字区域边沿有明显区别，可采用自动识别方法自动辨别出该文本所在区域，也可通过人工观察来手工设定该区域。

为了将该失真区域校正为图 2 所示的无失真矩形区域，应用本发明的方法是：

将不规则的四边形边沿按照横向和纵向划分为若干像素，例如，统计线段 $P_{11}P_{1n}$ 所

历经的像素点数目，设为 N_1 ，统计线段 $P_{m1}P_{mn}$ 所历经的像素点数目，设为 N_2 ，则横向划分像素点数 n 为

$$n=(N_1+N_2)/2$$

采用类似方法对线段 $P_{11}P_{m1}$ 和线段 $P_{1n}P_{mn}$ 进行纵向像素点划分，设纵向划分像素点数为 m 。

求出线段 $P_{11}P_{1n}$ 上均匀划分的 n 个像素点 ($P_{11}, P_{12}, \dots, P_{1n}$) 的所在坐标，并根据任何一种插值方法，求出这些像素点的取值。

然后将这些点映射到图 2 中的线段 $P_{11}'P_{1n}'$ 上的 n 个点 ($P_{11}', P_{12}', \dots, P_{1n}'$) 位置上。

以相同的方法对下一线段 $P_{21}P_{2n}$ 上的 n 个像素点进行操作。

这样，当进行第 i 次线段划分时，将线段 $P_{i1}P_{in}$ 上的 n 个像素点 ($P_{ij}, j=1, \dots, n$) 值对应到图 2 中的线段 $P_{i1}'P_{in}'$ 上的像素点 ($P_{ij}', j=1, \dots, n$) 位置上。

这样的过程一直进行到线段 $P_{m1}P_{mn}$ 为止。

可重复以上过程形成迭代。

2. 对于非线性几何失真的不规则的四边形的校正

如图 3 所示，被校正图像中的文字区域发生了非线性失真，原本为矩形区域的文字部分失真为一个由四段曲线围成的不规则区域。由于文字区域部分与文字区域边沿有明显区别，可采用自动识别方法自动辨别出该文本所在区域，也可通过人工观察来手工设定该区域。由于曲线的光滑性质，在区域辨别或设定时只需要确定曲线上的若干点，再通过样条插值等插值方法计算出近似曲线。

为了将该曲线围成的失真区域校正为图 6 所示的无失真矩形区域，应用本发明的方法是：

将一次迭代校正过程分为两个阶段：第一阶段是进行垂直校正，其结果为将区域竖直方向的围线校正为垂直等长的两条边，而水平方向上仍然存在失真，如图 4 所示。第二阶段是在图 4 所示结果的基础上再进行水平方向的校正，得到如图 5 所示的矩形区域校正结果。

这两个阶段所使用的校正算法相同。以第一阶段垂直校正为例，设定垂直方向的计算分辨率 m 后，将竖直方向的两曲边分割为 m 段，并将两曲边上对应的分割点用直线连接起来，再根据设定的水平方向分辨率 n ，将该连线分割为 n 段，于是得出这些网格交点共

$m \times n$ 个，图 3 中所示的点 P_{ij} 即表示这些网格交点中任意一个。

根据点 P_{ij} 的序号 (i,j) 可直接将之映射到图 4 所示的区域中，得到点 P_{ij}' 。这样的过程一直进行到 $i=m, j=n$ 为止。

用同样的方法完成第二阶段的水平方向校正。得到如图 5 所示的校正结果。最终将被校正图像上的点 P_{ij} 映射到图 5 所示的点 P_{ij}'' 上。

可重复以上过程形成迭代。

3. 动态域值的文字部分二值化分离方法

如图 6 所示，图中以一维曲线来示例二维图像的亮度变化。设曲线取值大的部分（凸起部分）为图像中的文字部分，而曲线取值较小部分（凹下部分）为图像背景部分。注意到背景部分的取值是变化的，表示图像背景随区域的不同而发生亮度变化。 $f(x,y)$ 表示源图像，动态域值二值化分离方法的思想是构造一种随图像局部背景亮度变化而变化的量化域值，从而将文字凸起部分和背景部分分离。为此，本发明的方法是：

对源图像 $f(x,y)$ 进行任意形式的模糊运算，也即图像的低通滤波，得出的模糊图像 $g(x,y)$ 代表了源图像局部区域的亮度。本发明利用模糊图像 $g(x,y)$ 的加权值

$$t(x,y)=g(x,y)+D$$

作为动态域值，对源图像 $f(x,y)$ 进行二值化分离，即

$$H(x,y)=255 \quad \text{当 } f(x,y)>t(x,y)$$

$$H(x,y)=0 \quad \text{当 } f(x,y)<t(x,y)$$

还可进一步采用迭代方法来改善文字部分的分离效果，如下：

根据分离结果，将源图像 $f(x,y)$ 中文字部分扣除 ($H(x,y)=255$ 部分)，即用相邻背景亮度替代相应的文字部分，得出背景部分 $f'(x,y)$ ，对之采用上述的图像模糊方法可得到新的模糊图像 $g'(x,y)$ ，加权后可作为新的动态域值对源图像实施二值化。

4. 迭代停止条件

以上方法（几何校正方法和动态域值方法）中，当前后两次迭代产生的结果之差值的均方值小于设定门限时，或迭代次数大于指定最大迭代次数时，停止迭代。

以上所述，仅是本发明的较佳的实施例，不构成对本发明的任何形式上的限制，凡是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、等效变化与修饰，均仍属于本发明技术方案范围内。

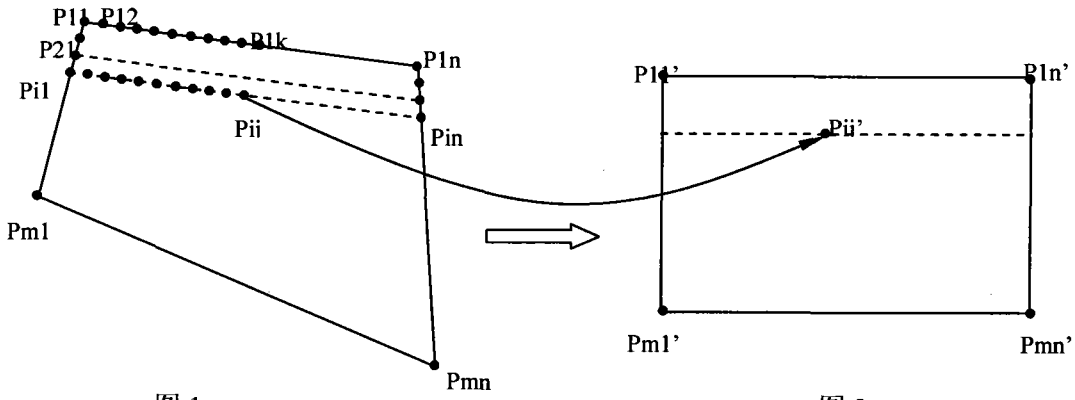


图 1

图 2

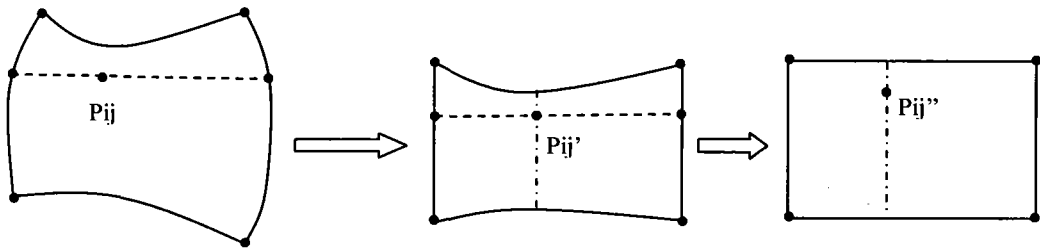


图 3

图 4

图 5

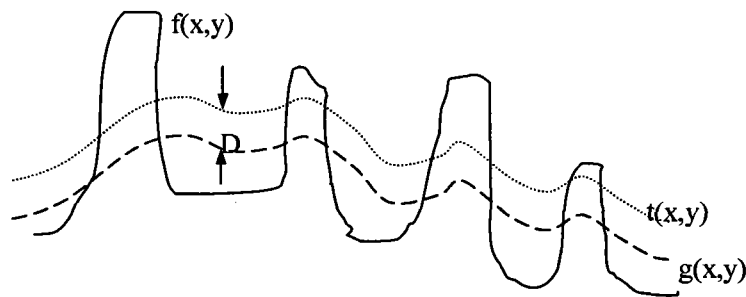


图 6