



(12)发明专利申请

(10)申请公布号 CN 113744172 A

(43)申请公布日 2021.12.03

(21)申请号 202010479951.7

(22)申请日 2020.05.29

(71)申请人 北京小米松果电子有限公司  
地址 100085 北京市海淀区清河朱房路临  
66号C栋C单元

(72)发明人 彭鑫 刘坚强 吴鹏杰

(74)专利代理机构 北京博思佳知识产权代理有  
限公司 11415

代理人 陈蕾

(51)Int.Cl.

G06T 5/50(2006.01)

G06T 7/11(2017.01)

G06T 7/136(2017.01)

G06T 5/30(2006.01)

G06T 11/40(2006.01)

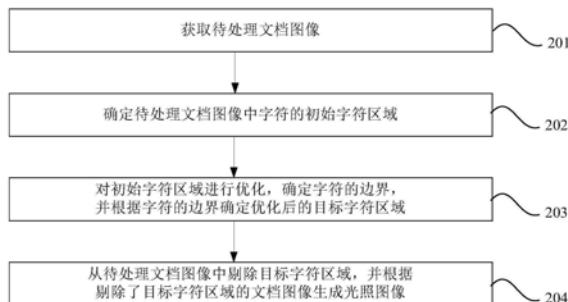
权利要求书2页 说明书13页 附图7页

(54)发明名称

文档图像的处理方法及装置、训练样本的生成方法及装置

(57)摘要

本公开提供一种文档图像的处理方法及装置、训练样本的生成方法及装置,所述方法包括:所述处理方法包括:确定待处理文档图像中字符所在的初始字符区域;对所述初始字符区域进行优化,确定所述字符的边界,并根据所述字符的边界确定优化后的目标字符区域;从所述待处理文档图像中剔除所述目标字符区域,并根据剔除了所述目标字符区域的待处理文档图像生成光照图像,所述光照图像用于反映环境的光照信息。从而能够生成包括真实光照信息的光照图像,作为训练样本生成的基础,解决了目前光照信息难以获取的难题。



1. 一种文档图像的处理方法,其特征在于,所述处理方法包括:  
确定待处理文档图像中字符所在的初始字符区域;  
对所述初始字符区域进行优化,确定所述字符的边界,并根据所述字符的边界确定优化后的目标字符区域;  
从所述待处理文档图像中剔除所述目标字符区域,并根据剔除了所述目标字符区域的待处理文档图像生成光照图像,所述光照图像用于反映环境的光照信息。
2. 如权利要求1所述的文档图像的处理方法,其特征在于,确定所述待处理文档图像中的初始字符区域,包括:  
对所述待处理文档图像进行二值化处理,得到二值图像,并从所述二值图像中识别出所述初始字符区域;  
或,将所述待处理文档图像中所述字符所在的区域进行二值化处理,并将经过二值化处理的区域确定为所述初始字符区域。
3. 如权利要求1或2所述的文档图像的处理方法,其特征在于,对所述初始字符区域进行优化,包括:  
基于形态学算法对所述初始字符区域进行优化。
4. 如权利要求1所述的文档图像的处理方法,其特征在于,根据剔除了所述目标字符区域的待处理文档图像生成光照图像,包括:  
对于空白区域中的每个像素点,基于插值算法拟合距离所述像素点预设范围内的相邻像素点的像素值,并将所述像素值填充于所述像素点处,所述空白区域为所述待处理文档图像中剔除了所述目标字符区域的区域;  
将经过像素值填充的待处理文档图像确定为所述光照图像。
5. 如权利要求1所述的文档图像的处理方法,其特征在于,所述待处理文档图像中背景区域的像素值与所述待处理文档图像中字符区域的像素值之差大于像素阈值。
6. 一种训练样本的生成方法,其特征在于,所述生成方法包括:  
获取文档图像样本;  
将光照图像叠加于所述文档图像样本上,得到叠加图像,所述光照图像由权利要求1-5中任一项所述的文档图像的处理方法获得;  
将所述叠加图像和所述光照图像作为模型训练的训练样本对,所述模型用于消除输入的文档图像中的光照信息。
7. 如权利要求6所述的训练样本的生成方法,其特征在于,所述文档图像样本包括以下至少之一:  
由电子文档转换得到的第一文档图像;  
对所述第一文档图像进行空间几何变换得到的第二文档图像。
8. 一种文档图像的处理装置,其特征在于,所述处理装置包括:  
确定模块,用于确定待处理文档图像中字符所在的初始字符区域;  
优化模块,用于对所述初始字符区域进行优化,确定所述字符的边界,并根据所述字符的边界确定优化后的目标字符区域;  
生成模块,用于从所述待处理文档图像中剔除所述目标字符区域,并根据剔除了所述目标字符区域的待处理文档图像生成光照图像,所述光照图像用于反映环境的光照信息。

9. 如权利要求8所述的文档图像的处理装置,其特征在于,在对所述初始字符区域进行优化时,所述优化模块用于:

对所述待处理文档图像进行二值化处理,得到二值图像,并从所述二值图像中识别出所述初始字符区域;

或,将所述待处理文档图像中所述字符所在的区域进行二值化处理,并将经过二值化处理的区域确定为所述初始字符区域。

10. 如权利要求8所述的文档图像的处理装置,其特征在于,在根据剔除了所述目标字符区域的待处理文档图像生成光照图像时,所述生成模块用于:

对于空白区域中的每个像素点,基于插值算法拟合距离所述像素点预设范围内的相邻像素点的像素值,并将所述像素值填充于所述像素点处,所述空白区域为所述待处理文档图像中剔除了所述目标字符区域的区域;

将经过像素值填充的待处理文档图像确定为所述光照图像。

11. 一种训练样本的生成装置,其特征在于,所述生成装置包括:

获取模块,用于获取文档图像样本;

叠加模块,用于将光照图像叠加于所述文档图像样本上,得到叠加图像,所述光照图像由权利要求8-10中任一项所述的文档图像的处理装置获得;

样本生成模块,用于将所述叠加图像和所述光照图像作为模型训练的训练样本对,所述模型用于消除输入的文档图像中的光照信息。

12. 如权利要求11所述的训练样本的生成装置,其特征在于,所述文档图像样本包括一种至少之一:

由电子文档转换得到的第一文档图像;

对所述第一文档图像进行空间几何变换得到的第二文档图像。

13. 一种文档图像的处理设备,其特征在于,包括:

处理器;

用于存储处理器可执行指令的存储器;

其中,所述处理器被配置为执行所述执行指令时实现权利要求1至5中任一项所述的文档图像的处理方法的步骤。

14. 一种训练样本的生成设备,其特征在于,包括:

处理器;

用于存储处理器可执行指令的存储器;

其中,所述处理器被配置为执行所述执行指令时实现权利要求6或7所述的训练样本的生成方法的步骤。

## 文档图像的处理方法及装置、训练样本的生成方法及装置

### 技术领域

[0001] 本公开涉及图像处理技术领域,尤其涉及文档图像的处理方法及装置、训练样本的生成方法及装置。

### 背景技术

[0002] 随着深度学习算法在计算机视觉领域的广泛应用,并且在很多方向都取得了非常不错的效果。利用深度学习算法来美化文档图片的光照并且消除拍摄带来的阴影是一种新颖的、可行的方案。但基于深度学习建立的模型依赖样本数据,具有非常强的数据驱动的特点,用于训练模型的样本数据的数目、质量好坏、多样性直接影响模型的准确度。

[0003] 目前,文档图像数据往往高度依赖人工采集,采集者需要拍摄大量不同用于模型训练的文档图像。同时为了获取对应的标签,还需在无光照条件下拍摄同样场景下的文档图像,并对光照条件下的文档图像与无光照条件下的文档图像进行图像配准和信息标注,工作量非常大,样本数据的获取成为深度学习在文档图片类的增强任务上的应用的一个极大瓶颈。

### 发明内容

[0004] 为克服相关技术中存在的问题,本公开提供了文档图像的处理方法及装置、训练样本的生成方法及装置。

[0005] 根据本公开实施例的第一方面,提供一种文档图像的处理方法,所述处理方法包括:

[0006] 确定待处理文档图像中字符所在的初始字符区域;

[0007] 对所述初始字符区域进行优化,确定所述字符的边界,并根据所述字符的边界确定优化后的目标字符区域;

[0008] 从所述待处理文档图像中剔除所述目标字符区域,并根据剔除了所述目标字符区域的待处理文档图像生成光照图像,所述光照图像用于反映环境的光照信息。

[0009] 可选地,确定所述待处理文档图像中的初始字符区域,包括:

[0010] 对所述待处理文档图像进行二值化处理,得到二值图像,并从所述二值图像中识别出所述初始字符区域;

[0011] 或,将所述待处理文档图像中所述字符所在的区域进行二值化处理,并将经过二值化处理的区域确定为所述初始字符区域。

[0012] 可选地,对所述初始字符区域进行优化,包括:

[0013] 基于形态学算法对所述初始字符区域进行优化。

[0014] 可选地,根据剔除了所述目标字符区域的待处理文档图像生成光照图像,包括:

[0015] 对于空白区域中的每个像素点,基于插值算法拟合距离所述像素点预设范围内的相邻像素点的像素值,并将所述像素值填充于所述像素点处,所述空白区域为所述待处理文档图像中剔除了所述目标字符区域的区域;

- [0016] 将经过像素值填充的待处理文档图像确定为所述光照图像。
- [0017] 可选地,所述待处理文档图像中背景区域的像素值与所述待处理文档图像中字符区域的像素值之差大于像素阈值。
- [0018] 根据本公开实施例的第二方面,提供一种训练样本的生成方法,所述生成方法包括:
- [0019] 获取文档图像样本;
- [0020] 将光照图像叠加于所述文档图像样本上,得到叠加图像,所述光照图像由上述任一项所述的文档图像的处理方法获得;
- [0021] 将所述叠加图像和所述光照图像作为模型训练的训练样本对,所述模型用于消除输入的文档图像中的光照信息。
- [0022] 根据本公开实施例的第三方面,提供一种文档图像的处理装置,所述处理装置包括:
- [0023] 确定模块,用于确定待处理文档图像中字符所在的初始字符区域;
- [0024] 优化模块,用于对所述初始字符区域进行优化,确定所述字符的边界,并根据所述字符的边界确定优化后的目标字符区域;
- [0025] 生成模块,用于从所述待处理文档图像中剔除所述目标字符区域,并根据剔除了所述目标字符区域的待处理文档图像生成光照图像,所述光照图像用于反映环境的光照信息。
- [0026] 可选地,在对所述初始字符区域进行优化时,所述优化模块用于:
- [0027] 对所述待处理文档图像进行二值化处理,得到二值图像,并从所述二值图像中识别出所述初始字符区域;
- [0028] 或,将所述待处理文档图像中所述字符所在的区域进行二值化处理,并将经过二值化处理的区域确定为所述初始字符区域。
- [0029] 可选地,在根据剔除了所述目标字符区域的待处理文档图像生成光照图像时,所述生成模块用于:
- [0030] 对于空白区域中的每个像素点,基于插值算法拟合距离所述像素点预设范围内的相邻像素点的像素值,并将所述像素值填充于所述像素点处,所述空白区域为所述待处理文档图像中剔除了所述目标字符区域的区域;
- [0031] 将经过像素值填充的待处理文档图像确定为所述光照图像。
- [0032] 根据本公开实施例的第四方面,提供一种训练样本的生成装置,所述生成装置包括:
- [0033] 获取模块,用于获取文档图像样本;
- [0034] 叠加模块,用于将光照图像叠加于所述文档图像样本上,得到叠加图像,所述光照图像由上述任一项所述的文档图像的处理装置获得;
- [0035] 样本生成模块,用于将所述叠加图像和所述光照图像作为模型训练的训练样本对,所述模型用于消除输入的文档图像中的光照信息。
- [0036] 可选地,所述文档图像样本包括一种至少之一:
- [0037] 由电子文档转换得到的第一文档图像;
- [0038] 对所述第一文档图像进行空间几何变换得到的第二文档图像。

- [0039] 根据本公开实施例的第五方面,提供一种文档图像的处理设备,包括:
- [0040] 处理器;
- [0041] 用于存储处理器可执行指令的存储器;
- [0042] 其中,所述处理器被配置为执行所述执行指令时实现上述任一项所述的文档图像的处理方法的步骤。
- [0043] 根据本公开实施例的第六方面,提供一种训练样本的生成设备,包括:
- [0044] 处理器;
- [0045] 用于存储处理器可执行指令的存储器;
- [0046] 其中,所述处理器被配置为执行所述执行指令时实现上述任一项所述的训练样本的生成方法的步骤。
- [0047] 本公开的实施例提供的技术方案可以包括以下有益效果:
- [0048] 本公开实施例中,通过对文档图像进行二值化处理和形态学处理,能够精确定位文档图像中字符区域的边界,并根据沿边界剔除字符区域后的文档图像生成包括光照信息的光照图像,作为训练样本生成的基础,从而解决了目前光照信息难以获取的难题。
- [0049] 应当理解的是,以上的一般描述和后文的细节描述仅是示例性和解释性的,并不能限制本公开。

#### 附图说明

- [0050] 此处的附图被并入说明书中并构成本公开的一部分,示出了符合本公开的实施例,并与说明书一起用于解释本公开的原理。
- [0051] 图1是本公开根据一示例性实施例示出的一种训练样本的生成方法的流程图。
- [0052] 图2是本公开根据一示例性实施例示出的一种文档图像的处理方法。
- [0053] 图3a是本公开根据一示例性实施例示出的一幅对纸质文档进行拍摄得到的文档图像的示意图。
- [0054] 图3b是本公开根据一示例性实施例示出的初始字符区域的示意图。
- [0055] 图3c是本公开根据一示例性实施例示出的对图3b进行腐蚀操作后的结果示意图。
- [0056] 图3d是本公开根据一示例性实施例示出的对图3c进行膨胀操作后的结果示意图。
- [0057] 图3e是本公开根据一示例性实施例示出的剔除字符区域之后的文档图像的示意图。
- [0058] 图4是本公开根据一示例性实施例示出的另一种文档图像的处理方法。
- [0059] 图5是本公开根据一示例性实施例示出的一种文档图像的处理装置的模块示意图。
- [0060] 图6是本公开根据一示例性实施例示出的另一种文档图像的处理装置的模块示意图。
- [0061] 图7是本公开根据一示例性实施例示出的一种训练样本的生成装置的模块示意图。
- [0062] 图8是本公开根据一示例性实施例示出的一种用于文档图像处理的设备的框图。
- [0063] 图9是根据一示例性实施例示出的另一种用于文档图像处理的设备的框图。

## 具体实施方式

[0064] 这里将详细地对示例性实施例进行说明,其示例表示在附图中。下面的描述涉及附图时,除非另有表示,不同附图中的相同数字表示相同或相似的要素。以下示例性实施中所描述的实施例并不代表与本公开相一致的所有实施例。相反,它们仅是与如所附权利要求书中所详述的、本公开的一些方面相一致的装置和方法的例子。

[0065] 在本公开使用的术语是仅仅出于描述特定实施例的目的,而非旨在限制本公开。在本公开和所附权利要求书中所使用的单数形式的“一种”、“所述”和“该”也旨在包括多数形式,除非上下文清楚地表示其他含义。还应当理解,本文中使用的术语“和/或”是指并包括一个或多个相关联的列出项目的任何或所有可能组合。

[0066] 应当理解,尽管在本公开可能采用术语第一、第二、第三等来描述各种信息,但这些信息不应限于这些术语。这些术语仅用来将同一类型的信息彼此区分开。例如,在不脱离本公开范围的情况下,第一信息也可以被称为第二信息,类似地,第二信息也可以被称为第一信息。取决于语境,如在此所使用的词语“如果”可以被解释成为“在……时”或“当……时”或“响应于确定”。

[0067] 随着便携式摄影设备如智能手机等技术的飞速发展,人们可以更加便捷地拍摄得到大量高质量的照片,除了用于记录日常生活中的精彩瞬间,人们还希望可以随时随地拍摄文档,并且得到规范的文档图像,用于重要信息的记录及分享。相比于传统的扫描仪,通过手机拍摄生成数字化文档易于使用且操作简单。但由于拍摄环境复杂多变,在光照信息的影响下,通常拍摄获得的文档图像的质量并不高,这使得高效鲁棒的文档图像增强成为亟待解决的问题。

[0068] 随着深度学习算法在计算机视觉领域的广泛应用,并且在很多方向都取得了非常不错的效果。利用深度学习算法对文档图像进行增强处理,以美化文档图片的光照并且消除拍摄带来的阴影是一种新颖的、可行的方案。但由于深度学习算法依赖样本数据,具有非常强的数据驱动的特点,用于训练模型的样本数据的数量、质量好坏、多样性会直接影响训练模型的准确度。

[0069] 目前,对于图像增强模型,其训练样本数据往往高度依赖人工采集,采集者需要拍摄大量不同光照条件、不同内容、不同角度等满足用于模型训练的文档图像。同时为了获取对应的训练标签,还需在无光照的条件下拍摄同样场景的文档图像,因为两次拍摄的距离、角度会存在差异,因此对光照条件下拍摄的文档图像(含光照信息的文档图像)与无光照条件下拍摄的文档图像(不含光照信息的文档图像)还需要进行图像配准对齐、光照信息标注,之后才能将经过图形配准和信息标注的含光照信息的文档图像以及不含光照信息的文档图像作为训练模型的图像样本对。

[0070] 对于上述样本数据的获取过程,一方面对拍摄环境要求非常严格,在实际拍摄含光照信息与不含光照信息的图像时,很难保证两次拍摄的拍摄环境条件完全一致;另一方面需要对拍摄得到的图像进行图像配准、光照信息标注,计算量大,样本数据的制作非常繁琐,无法短时间得到批量数据。

[0071] 基于上述问题,本公开实施例提供一种训练样本的生成方法,将从图像中提取的真实的光照图像叠加于不含光照信息的文档图像上,作为模型训练的训练样本,以减少人工采集和信息标注的工作量。

[0072] 接下来对本公开实施例进行详细说明。

[0073] 如图1所示,图1是本公开根据一示例性实施例示出的一种训练样本的生成方法的流程图,包括以下步骤:

[0074] 步骤101、获取文档图像样本。

[0075] 模型训练需要大量训练样本,文档图像作为生成训练样本的基础,可以但不限于通过对大量电子文档进行格式转换得到,例如,将word/pdf/txt等格式的电子文档转换为图像格式,得到文档图像样本,该文档图像样本不含光照信息。其中,为了满足模型训练对训练样本多样性的要求,电子文档可以采用各类电子杂志、电子期刊、电子书籍等。

[0076] 步骤102、将光照图像叠加于文档图像样本上,得到叠加图像。

[0077] 步骤102中,对于步骤101获得的每份文档图像样本,均需要叠加光照图像。

[0078] 其中,光照图像能够反映真实的环境光照信息,光照图像根据含光照信息的文档图像生成,具体实现过程参见下文叙述。

[0079] 可以理解地,为了扩大训练样本的数量,对于每幅不含光照信息的文档图像样本,可以叠加多幅不同的光照图像,从而对于每幅文档图像样本可以生成多幅不同的叠加图像。举例来说,若有1000幅文档图像样本和50幅光照图像,进行图像叠加之后,则能生成50000幅叠加图像,用于模型训练,大大扩大了模型训练样本的数量。

[0080] 在另一个实施例中,若光照图像与文档图像样本的尺寸大小不一致,可以先对光照图像和/或文档图像样本进行图像变换,将两幅图像的尺寸大小调整为相同。

[0081] 步骤103、将叠加图像和光照图像作为模型训练的训练样本对。

[0082] 步骤103获得的训练样本对可以用于模型训练,训练完成的模型能够消除输入的文档图像中的光照信息,输出不含光照信息的文档图像,实现文档图像的图像增强。

[0083] 本实施例中,通过组合叠加不同的光照图像与不含光照信息的文档图像样本可以得到大量的叠加图像作为训练样本,由于叠加图像是根据光照图像和文档图像样本叠加得到的,叠加图像的图像信息,包括尺寸大小、光照位置等,均可通过光照图像和文档图像样本确定,从而无需人工拍摄采集图像,无需进行图像配准和光照信息标注,自动化程度高,计算量小,且极少需要人工标注参与,能够节约大量人力成本。

[0084] 本实施例中的光照图像是通过算法处理从含有真实光照信息的文档图像中分离出来的,而非通过模拟得到,以此作为训练样本生成的基础,训练样本的真实性、准确性非常高,能够提高模型训练的准确度。

[0085] 在另一个实施例中,为了进一步扩大训练样本的数量,还可以对由电子文档转换得到的文档图像进行几何空间变化,例如,将电子文档转换得到的文档图像分割成多份子图像,将所属于不同文档图像的子图像拼接成新的文档图像,作为生成训练样本的基础;或将原文档图像的子图形进行旋转变换后拼接成新的文档图像,作为生成训练样本的基础。从而,将文档图像经过简单的空间几何变换,然后与不同的光照图像进行组合叠加,即可以得到大量可用于训练的样本数据。

[0086] 以下介绍光照图像的生成过程。

[0087] 图2是本公开根据一示例性实施例示出的一种文档图像的处理方法,包括以下:

[0088] 步骤201、获取待处理文档图像。

[0089] 其中,待处理文档图像可以是对包含字符的桌面、布匹、广告牌、纸质文档等目标



对象拍摄得的图像,待处理文档图像也可以是从网络上获取的他人拍摄的图像。

[0090] 目标对象是生成待处理文档图像的基础,为了能够从待处理文档图像中准确提取真实的光照信息,最好采用字符颜色与背景颜色相差较大的目标对象,例如,采用白纸黑字的纸质文档,这样对纸质文档进行拍摄得到的待处理文档图像中背景区域的颜色几乎为白色,字符区域的颜色为黑色,参见图3a示出的对纸质文档拍摄得到的待处理文档图像为例,图中黑色字体及标点符号所在的区域为字符区域,白色部分为背景区域,两者的像素值相差较大,有利于后续生成光照图像。

[0091] 可以理解地,待处理文档图像是在光照环境中对目标对象拍摄得到的,待处理文档图像不仅包括目标对象的原始信息,还含有光照信息。提供光照环境的光源,可以但不限于是日光、白炽灯、彩色LED(发光二极管)等。从一幅待处理文档图像中可以提取一幅能够反映光照信息的光照图像,为了增加训练样本的数量和多样性,可以对多种类型的目标对象进行拍摄,并将目标对象放在不同的光照环境下,分别从不同角度、不同距离对其进行拍摄,获得多份待处理文档图片。

[0092] 步骤202、确定待处理文档图像中字符的初始字符区域。

[0093] 在一个实施例中,对于初始字符区域的确定,可以先对待处理文档图像进行二值化处理,得到二值图像,该二值图像中像素点的像素值要么为0,要么为255,根据像素值即可识别出字符的初始字符区域。二值图像呈现出明显的黑白效果,有利于后续生成光照图像。

[0094] 在另一个实施例中,对于初始字符区域的确定,可以先确定待处理文档图像中字符所在的区域,可以但不限于采用文字检测算法,再对字符所在的区域进行二值化处理,并将经过二值化处理的区域确定为初始字符区域。

[0095] 参见图3a,图中的虚线框中的区域为初始字符区域,虚线框中的区域不仅包括字符还包括一部分的背景区域,需要执行以下步骤对初始字符区域的边界进行细化,以优化初始字符区域。

[0096] 步骤203、对初始字符区域进行优化,确定字符的边界,并根据字符的边界确定优化后的目标字符区域。

[0097] 其中,可以但不限于采用形态学算法对二值图像中的初始字符区域进行优化,以确定字符的边界,进而准确确定图像中的字符区域(目标字符区域)。

[0098] 对初始字符区域进行形态学处理,可以采用形态学开运算或形态学闭运算。形态学开运算是先对图像进行腐蚀,再进行膨胀的运算,可以用来消除小物体、在纤细点处分离物体、平滑较大物体的边界的同时并不明显改变其面积。形态学闭运算是先对图像进行膨胀,再进行腐蚀的运算,闭运算能够排除小型黑洞(黑色区域)。用户可根据文档图像中字符的特点选择使用形态学开运算或形态学闭运算。

[0099] 下面以采用形态学开运算为例,对初始字符区域进行形态学处理的具体实现过程进行说明:

[0100] 按数学方面来说,膨胀或者腐蚀操作就是将初始字符区域与卷积核进行卷积,以图3b示出的初始字符区域为例,图中每个小正方形表示一个像素点,斜线填充的正方形表示字符区域,白色的正方形表示背景区域,3\*3个虚线正方形区域表示卷积核。卷积核可以是任何的形状和大小,不限于图中示出的以3\*3个正方形组成。卷积核中有一个单独定义出

来的参考点,参见图中“☆”所在的区域位置。

[0101] 腐蚀是求局部最小值的操作:使用卷积核在初始字符区域(图3b)中进行移动,每移动一次计算一次卷积核覆盖的区域中像素点的最小值,并把这个最小值赋值给参考点所在的像素点。参见图3c,为对图3b进行腐蚀处理后的一个结果示意图,对比图3b和图3c可知,经过腐蚀操作后,斜线填充区域被缩小了。

[0102] 膨胀是求局部最大值的操作:使用卷积核在图3c中进行移动,每移动一次计算一次卷积核覆盖的区域中像素点的最大值,并把这个最大值赋值给参考点所在的像素点,这样就会使图像中的斜线填充区域逐渐增长。参见图3d,为对图3c进行膨胀处理后的一个结果示意图,图中的斜线填充区域为优化后的字符区域,其边界也即字符区域的边界。

[0103] 步骤204、从待处理文档图像中剔除目标字符区域,并根据剔除了目标字符区域的待处理文档图像生成光照图像。

[0104] 其中,光照图像能够反映环境的光照信息。

[0105] 图3e是本公开一示例性实施例示出的剔除字符区域之后的待处理文档图像,该待处理文档图像的文字区域被剔除后,只剩下光照信息。图中的点填充区域表征含有光照信息的区域,图中的白色区域表征剔除了目标字符区域后的空白区域。由于剔除图像中的字符区域时,该区域的光照信息也会被剔除,此时的图像有部分光照信息存在缺失,不能完全反映真实的光照信息,需要对空白区域的像素点进行光照信息填充,可以但不限于采用以下方式:

[0106] 对于空白区域中的每个像素点,基于插值算法拟合距离该像素点预设范围内的相邻像素点的像素值,并将像素值填充于该像素点处。其中,像素点的预设范围可以跟实际需求自行设置。

[0107] 以图3e示出的图像为例,假设预设范围为1个相邻像素点,对像素点a进行填充时,像素点a的相邻像素点为像素点b1~b8。

[0108] 在一个实施例中,可以对相邻像素点也即像素点b1~b8的像素值求均值,并将求得的均值赋值给像素点a,参见图3e,由于像素点b4、b5和b6为空白像素点,可将其像素值赋值为0后再求均值,或者仅根据像素点b1、b2、b3、b7和b8的像素值求均值。

[0109] 在另一个实施例中,可以采用最近邻插值法,还是以图3e为例,像素点a的相邻像素点中,像素点b1、b2、b3、b7和b8具有光照信息,可从这4个像素点中选择一个像素点,并将该选择的像素点的像素值赋值给像素点a。采用该算法计算量较小。

[0110] 对于空白区域中的每个像素点,均采用上述方式进行像素值填充,可以使图像中的每个像素点均包括光照信息,从而得到视觉观感较好的、包括真实光照信息的光照图像。

[0111] 本公开实施例中,能够精确定位文档图像中字符区域的边界,并根据沿边界剔除字符区域后的文档图像生成含有光照信息的光照图像,作为训练样本生成的基础,从而解决了目前光照信息难以获取的难题。

[0112] 图4是本公开根据一示例性实施例示出的另一种文档图像的处理方法,本实施例中,以对纸质文档进行拍摄得到的待处理文档图像为例,对文档图像的处理方法的具体实现过程进行说明,参见图4,方法包括以下步骤:

[0113] 步骤401、获取对纸质文档拍摄得到的待处理文档图像。

[0114] 其中,纸质文档包括字符区域及背景区域。该步骤中纸质文档是生成文档图像的

基础,为了能够从待处理文档图像中准确提取真实的光照信息,最好采用字符颜色与背景颜色相差较大的纸质文档,优选地,采用白纸黑字的纸质文档,这样对纸质文档进行拍摄得到的待处理文档图像中背景区域的颜色几乎为白色,字符区域的颜色为黑色,参见图3a示出的对纸质文档拍摄得到的文档图像为例,图中黑色字体及标点符号所在的区域为字符区域,白色部分为背景区域,两者的像素值相差较大,有利于后续生成光照图像。

[0115] 可以理解地,待处理文档图像是在光照环境中对纸质文档拍摄得到的,待处理文档图像不仅包括纸质文档的原始信息,还包括光照信息。提供光照环境的光源,可以但不限于是日光、白炽灯、彩色LED(发光二极管)等。从一幅待处理文档图像中可以提取一幅包括光照信息的光照图像,为了增加训练样本的数量和多样性,可以获取多份纸质文档,并将每份纸质文档放在不同的光照环境下,分别从不同角度、不同距离对纸质文档进行拍摄,获得多份文档图片,即获得多份待处理文档图像。

[0116] 步骤402、确定待处理文档图像中的初始字符区域。

[0117] 其中,可以但不限于采用文字检测算法,确定图像中的初始字符区域。文字检测算法只能识别出图像中字符的一个初始字符区域,参见图3a,图中虚线框出的区域为初始字符区域,该区域不仅包括字符还包括一部分的背景区域,需要执行以下步骤,以识别出准确的字符区域的边界,对字符区域进行优化。

[0118] 步骤403、基于形态学算法对初始字符区域进行优化,确定字符的边界,并根据字符的边界确定优化后的目标字符区域。

[0119] 为了便于形态学处理,进行形态学处理之前,需要先对初始字符区域进行二值化处理,将字符区域中的像素值赋值为0或255,使初始字符区域呈现出明显的黑白效果。

[0120] 在另一个实施例中,可以先对待处理文档图像进行二值化处理,得到二值图像,然后从二值图像中识别出初始字符区域,该初始字符区域中的像素点的像素值只有0、255两个值。

[0121] 对初始字符区域进行形态学处理,可以采用形态学开运算或形态学闭运算。形态学开运算是先对图像进行腐蚀,再进行膨胀的运算,可以用来消除小物体、在纤细点处分离物体、平滑较大物体的边界的同时并不明显改变其面积的特定。形态学闭运算是先对图像进行膨胀,再进行腐蚀的运算,闭运算能够排除小型黑洞(黑色区域)。用户可根据文档图像中字符的特点选择使用形态学开运算或形态学闭运算。

[0122] 下面以采用形态学开运算为例,对初始字符区域进行形态学处理的具体实现过程进行说明:

[0123] 按数学方面来说,膨胀或者腐蚀操作就是将初始字符区域与卷积核进行卷积,以图3b示出的初始字符区域为例,图中每个小正方形表示一个像素点,斜线填充的正方形表示字符区域,白色的正方形表示背景区域,3\*3个虚线正方形区域表示卷积核。卷积核可以是任何的形状和大小,不限于图中示出的以3\*3个正方形组成。卷积核中有一个单独定义出来的参考点,参见图中“☆”所在的区域位置。

[0124] 腐蚀是求局部最小值的操作:使用卷积核在初始字符区域(图3b)中进行移动,每移动一次计算一次卷积核覆盖的区域中像素点的最小值,并把这个最小值赋值给参考点所在的像素点。参见图3c,为对图3b进行腐蚀处理后的一个结果示意图,对比图3b和图3c可知,经过腐蚀操作后,斜线填充区域被缩小了。

[0125] 膨胀是求局部最大值的操作：使用卷积核在图3c中进行移动，每移动一次计算一次卷积核覆盖的区域中像素点的最大值，并把这个最大值赋值给参考点所在的像素点，这样就会使图像中的斜线填充区域逐渐增长。参见图3d，为对图3c进行膨胀处理后的一个结果示意图，图中的斜线填充区域为优化后的字符区域，其边界也即字符区域的边界。

[0126] 步骤404、沿字符的边界剔除待处理文档图像中的目标字符区域，并根据剔除了目标字符区域的待处理文档图像生成光照图像。

[0127] 图3e是本公开一示例性实施例示出的剔除目标字符区域之后的待处理文档图像，该待处理文档图像的文字区域被剔除后，只剩下光照信息。图中的点填充区域表征包括光照信息的区域，图中的白色区域表征剔除了字符区域后的空白区域。由于剔除图像中的字符区域时，该区域的光照信息也会被剔除，此时的图像有部分光照信息存在缺失，不能完全反映真实的光照信息，需要对空白区域的像素点进行光照信息填充，可以但不限于采用以下方式：

[0128] 对于空白区域中的每个像素点，基于插值算法拟合距离该像素点预设范围内的相邻像素点的像素值，并将像素值填充于该像素点处。其中，像素点的预设范围可以跟实际需求自行设置。

[0129] 以图3e的文档图像为例，假设预设范围为1个相邻像素点，对像素点a进行填充时，像素点a的相邻像素点为像素点b1~b8。

[0130] 在一个实施例中，可以对相邻像素点也即像素点b1~b8的像素值求均值，并将求得的均值赋值给像素点a，参见图3e，由于像素点b4、b5和b6为空白像素点，可将其像素值赋值为0后再求均值，或者仅根据像素点b1、b2、b3、b7和b8的像素值求均值。

[0131] 在另一个实施例中，可以采用最近邻插值法，还是以图3e为例，像素点a的相邻像素点中，像素点b1、b2、b3、b7和b8具有光照信息，可从这4个像素点中选择一个像素点，并将该选择的像素点的像素值赋值给像素点a。采用该算法计算量较小。

[0132] 对于空白区域中的每个像素点，均采用上述方式进行像素值填充，可以使文档图像中的每个像素点均包括光照信息，从而得到视觉观感较好的、包括真实光照信息的光照图像。

[0133] 本公开实施例中，能够精确定位文档图像中字符区域的边界，并根据沿边界剔除字符区域后的文档图像生成含有真实光照信息的光照图像，作为训练样本生成的基础，从而解决了目前光照信息难以获取的难题。

[0134] 与前述方法的实施例相对应，本公开还提供了装置及其所应用的设备的实施例。

[0135] 图5是本公开根据一示例性实施例示出的一种文档图像的处理装置的模块示意图，该处理装置包括：确定模块51、优化模块52和生成模块53。

[0136] 确定模块51，用于确定待处理文档图像中字符所在的初始字符区域；

[0137] 优化模块52，用于对所述初始字符区域进行优化，确定所述字符的边界，并根据所述字符的边界确定优化后的目标字符区域；

[0138] 生成模块53，用于从所述待处理文档图像中剔除所述目标字符区域，并根据剔除了所述目标字符区域的待处理文档图像生成光照图像，所述光照图像用于反映环境的光照信息。

[0139] 可选地，在对所述初始字符区域进行优化时，所述优化模块用于：

[0140] 对所述待处理文档图像进行二值化处理,得到二值图像,并从所述二值图像中识别出所述初始字符区域;

[0141] 或,将所述待处理文档图像中所述字符所在的区域进行二值化处理,并将经过二值化处理的区域确定为所述初始字符区域。

[0142] 可选地,在根据剔除了所述目标字符区域的待处理文档图像生成光照图像时,所述生成模块用于:

[0143] 对于空白区域中的每个像素点,基于插值算法拟合距离所述像素点预设范围内的相邻像素点的像素值,并将所述像素值填充于所述像素点处,所述空白区域为所述待处理文档图像中剔除了所述目标字符区域的区域;

[0144] 将经过像素值填充的待处理文档图像确定为所述光照图像。

[0145] 图6是本公开根据一示例性实施例示出的另一种文档图像的处理装置的模块示意图,该处理装置包括:获取模块61、确定模块62、优化模块63和生成模块64。

[0146] 获取模块61,用于获取待处理文档图像。

[0147] 其中,获取模块可以从本设备获取待处理文档图像,所述待处理文档图像由本设备对包括字符的纸质文档进行拍摄得到;获取模块也可以获取其他设备拍摄得到的待处理文档图像,或者从网络获取待处理文档图像。

[0148] 确定模块62,用于对所述待处理文档图像进行二值化处理,得到二值图像,并从所述二值图像中确定所述文档图像中的初始字符区域;

[0149] 优化模块63,用于对所述初始字符区域进行形态学处理,确定所述字符的边界,并根据所述字符的边界确定优化后的目标字符区域;

[0150] 生成模块64,用于所述待处理文档图像中剔除目标字符区域,并根据剔除了所述目标字符区域的待处理文档图像生成光照图像,所述光照图像用于反映环境的光照信息。

[0151] 可选地,在对所述初始字符区域进行形态学处理时,所述优化模块用于:

[0152] 对所述待处理文档图像进行二值化处理,得到二值图像,并从所述二值图像中识别出所述初始字符区域;

[0153] 或,将所述待处理文档图像中所述字符所在的区域进行二值化处理,并将经过二值化处理的区域确定为所述初始字符区域。

[0154] 可选地,在根据剔除了所述目标字符区域的待处理文档图像生成光照图像时,所述生成模块用于:

[0155] 对于空白区域中的每个像素点,基于插值算法拟合距离所述像素点预设范围内的相邻像素点的像素值,并将所述像素值填充于所述像素点处,所述空白区域为所述待处理文档图像中剔除了所述优化后的字符区域的区域;

[0156] 将经过像素值填充的文档图像确定为所述光照图像。

[0157] 图7是本公开根据一示例性实施例示出的一种训练样本的生成装置,所述生成装置包括:获取模块71、叠加模块72和样本生成模块73。

[0158] 获取模块71,用于获取文档图像样本;

[0159] 叠加模块72,用于将光照图像叠加于所述文档图像样本上,得到叠加图像,所述光照图像由上述任一项所述的文档图像的处理装置获得;

[0160] 样本生成模块73,用于将所述叠加图像和所述光照图像作为模型训练的训练样本

对,所述模型用于消除输入的文档图像中的光照信息。

[0161] 可选地,所述不包括光照信息的文档图像包括:

[0162] 由电子文档转换得到的第一文档图像;

[0163] 和/或,对所述第一文档图像进行空间几何变换得到的第二文档图像。

[0164] 关于上述实施例中的装置,其中各个模块执行操作的具体方式已经在相关方法的实施例中进行了详细描述,此处将不做详细阐述说明。

[0165] 对于装置实施例而言,由于其基本对应于方法实施例,所以相关之处参见方法实施例的部分说明即可。以上所描述的装置实施例仅仅是示意性的,其中所述作为分离部件说明的模块可以是或者也可以不是物理上分开的,作为模块显示的部件可以是或者也可以不是物理模块,即可以位于一个地方,或者也可以分布到多个网络模块上。可以根据实际的需要选择其中的部分或者全部模块来实现本实施例方案的目的。本领域普通技术人员在不付出创造性劳动的情况下,即可以理解并实施。

[0166] 图8是根据本公开的实施例示出的一种用于文档图像处理的设备的框图。该设备可以是移动电话,计算机,数字广播终端,消息收发设备,游戏控制台,平板设备,医疗设备,健身设备,个人数字助理等。

[0167] 如图8所示,设备800可以包括以下一个或多个组件:处理组件802,存储器804,电源组件806,多媒体组件808,音频组件810,输入/输出(I/O)的接口812,传感器组件814,以及通信组件816。设备还包括天线模组(例如可以连接于通信组件816),该天线模组包括:辐射体,接地点,馈电端,以及谐振电路。其中,辐射体包括开放端,接地点确定在辐射体上。馈电端电连接至辐射体上第一连接点。谐振电路的第一端电连接至第一连接点,谐振电路的第二端接地,谐振电路包括可调单元。第一连接点到开放端的距离小于第一连接点到接地点的距离。

[0168] 处理组件802通常控制装置800的整体操作,诸如与显示,电话呼叫,数据通信,相机操作和记录操作相关联的操作。处理组件802可以包括一个或多个处理器820来执行指令,以完成上述的方法的全部或部分步骤。此外,处理组件802可以包括一个或多个模块,便于处理组件802和其他组件之间的交互。例如,处理组件802可以包括多媒体模块,以方便多媒体组件808和处理组件802之间的交互。

[0169] 存储器804被配置为存储各种类型的数据以支持在装置800的操作。这些数据的示例包括用于在设备800上操作的任何应用程序或方法的指令,联系人数据,电话簿数据,消息,图片,视频等。存储器804可以由任何类型的易失性或非易失性存储设备或者它们的组合实现,如静态随机存取存储器(SRAM),电可擦除可编程只读存储器(EEPROM),可擦除可编程只读存储器(EPROM),可编程只读存储器(PROM),只读存储器(ROM),磁存储器,快闪存储器,磁盘或光盘。

[0170] 电源组件806为设备800的各种组件提供电力。电源组件806可以包括电源管理系统,一个或多个电源,及其他与为设备800生成、管理和分配电力相关联的组件。

[0171] 多媒体组件808包括在所述设备800和用户之间的提供一个输出接口的屏幕。在一些实施例中,屏幕可以包括液晶显示器(LCD)和触摸面板(TP)。如果屏幕包括触摸面板,屏幕可以被实现为触摸屏,以接收来自用户的输入信号。触摸面板包括一个或多个触摸传感器以感测触摸、滑动和触摸面板上的手势。所述触摸传感器可以不仅感测触摸或滑动动作

的边界,而且还检测与所述触摸或滑动操作相关的持续时间和压力。在一些实施例中,多媒体组件808包括一个前置摄像头和/或后置摄像头。当设备800处于操作模式,如拍摄模式或视频模式时,前置摄像头和/或后置摄像头可以接收外部的多媒体数据。每个前置摄像头和后置摄像头可以是一个固定的光学透镜系统或具有焦距和光学变焦能力。

[0172] 音频组件810被配置为输出和/或输入音频信号。例如,音频组件810包括一个麦克风(MIC),当设备800处于操作模式,如呼叫模式、记录模式和语音识别模式时,麦克风被配置为接收外部音频信号。所接收的音频信号可以被进一步存储在存储器804或经由通信组件816发送。在一些实施例中,音频组件810还包括一个扬声器,用于输出音频信号。

[0173] I/O接口812为处理组件802和外围接口模块之间提供接口,上述外围接口模块可以是键盘,点击轮,按钮等。这些按钮可包括但不限于:主页按钮、音量按钮、启动按钮和锁定按钮。

[0174] 传感器组件814包括一个或多个传感器,用于为设备800提供各个方面的状态评估。例如,传感器组件814可以检测到设备800的打开/关闭状态,组件的相对定位,例如所述组件为设备800的显示器和小键盘,传感器组件814还可以检测设备800或设备800一个组件的位置改变,用户与设备800接触的存在或不存在,设备800方位或加速/减速和设备800的温度变化。传感器组件814可以包括接近传感器,被配置用来在没有任何的物理接触时检测附近物体的存在。传感器组件814还可以包括光传感器,如CMOS或CCD图像传感器,用于在成像应用中使用。在一些实施例中,该传感器组件814还可以包括加速度传感器,陀螺仪传感器,磁传感器,压力传感器或温度传感器。

[0175] 通信组件816被配置为便于设备800和其他设备之间有线或无线方式的通信。设备800可以接入基于通信标准的无线网络,如WiFi,2G或3G,或它们的组合。在一个示例性实施例中,通信组件816经由广播信道接收来自外部广播管理系统的广播信号或广播相关信息。在一个示例性实施例中,所述通信组件816还包括近场通信(NFC)模块,以促进短程通信。例如,在NFC模块可基于射频识别(RFID)技术,红外数据协会(IrDA)技术,超宽带(UWB)技术,蓝牙(BT)技术和其他技术来实现。

[0176] 在示例性实施例中,设备800可以被一个或多个应用专用集成电路(ASIC)、数字信号处理器(DSP)、数字信号处理设备(DSPD)、可编程逻辑器件(PLD)、现场可编程门阵列(FPGA)、控制器、微控制器、微处理器或其他电子元件实现,用于执行上述任一实施例所述的方法。

[0177] 在示例性实施例中,还提供了一种包括指令的非临时性计算机可读存储介质,例如包括指令的存储器804,上述指令可由设备800的处理器820执行以完成上述方法。例如,所述非临时性计算机可读存储介质可以是ROM、随机存取存储器(RAM)、CD-ROM、磁带、软盘和光数据存储设备等。

[0178] 图9是根据一示例性实施例示出的另一种文档图像的处理设备900的框图。例如,设备900可以被提供为一服务器。参照图9,设备900包括处理组件922,其进一步包括一个或多个处理器,以及由存储器932所代表的存储器资源,用于存储可由处理部件922的执行的指令,例如应用程序。存储器932中存储的应用程序可以包括一个或一个以上的每一个对应于一组指令的模块。此外,处理组件922被配置为执行指令,以执行上述任一实施例所述的方法。

[0179] 设备900还可以包括一个电源组件926被配置为执行设备900的电源管理,一个有线或无线网络接口950被配置为将设备900连接到网络,和一个输入输出(I/O)接口958。设备900可以操作基于存储在存储器932的操作系统,例如Windows Server™,Mac OS X™,Unix™,Linux™,FreeBSD™或类似。

[0180] 本公开实施例还提供一种训练样本的生成设备,包括:处理器;用于存储处理器可执行指令的存储器;其中,所述处理器被配置为执行所述执行指令时实现上述任一实施例所述的训练样本的生成方法的步骤。

[0181] 本领域技术人员在考虑说明书及实践这里公开的公开后,将容易想到本公开的其它实施方案。本公开旨在涵盖本公开的任何变型、用途或者适应性变化,这些变型、用途或者适应性变化遵循本公开的一般性原理并包括本公开未公开的本技术领域中的公知常识或惯用技术手段。说明书和实施例仅被视为示例性的,本公开的真正范围和精神由下面的权利要求指出。

[0182] 应当理解的是,本公开并不局限于上面已经描述并在附图中示出的精确结构,并且可以在不脱离其范围进行各种修改和改变。本公开的范围仅由所附的权利要求来限制。



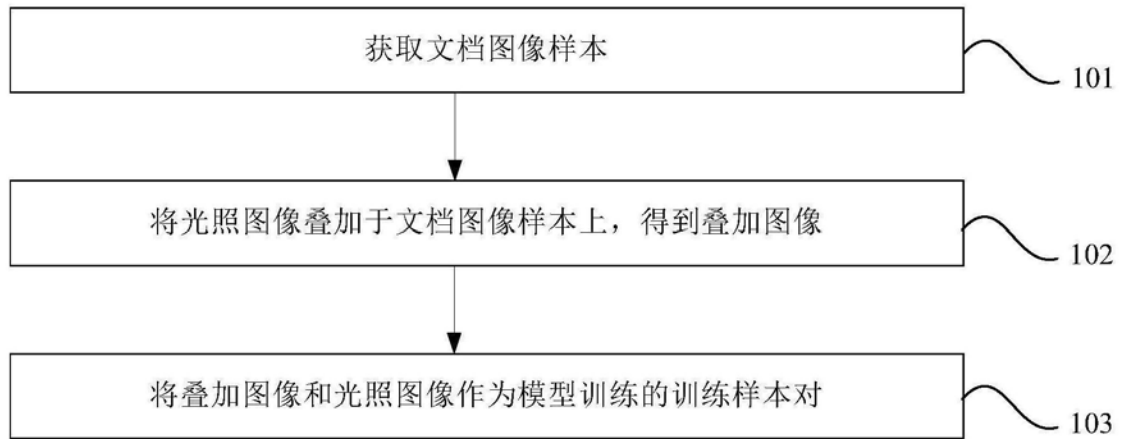


图1

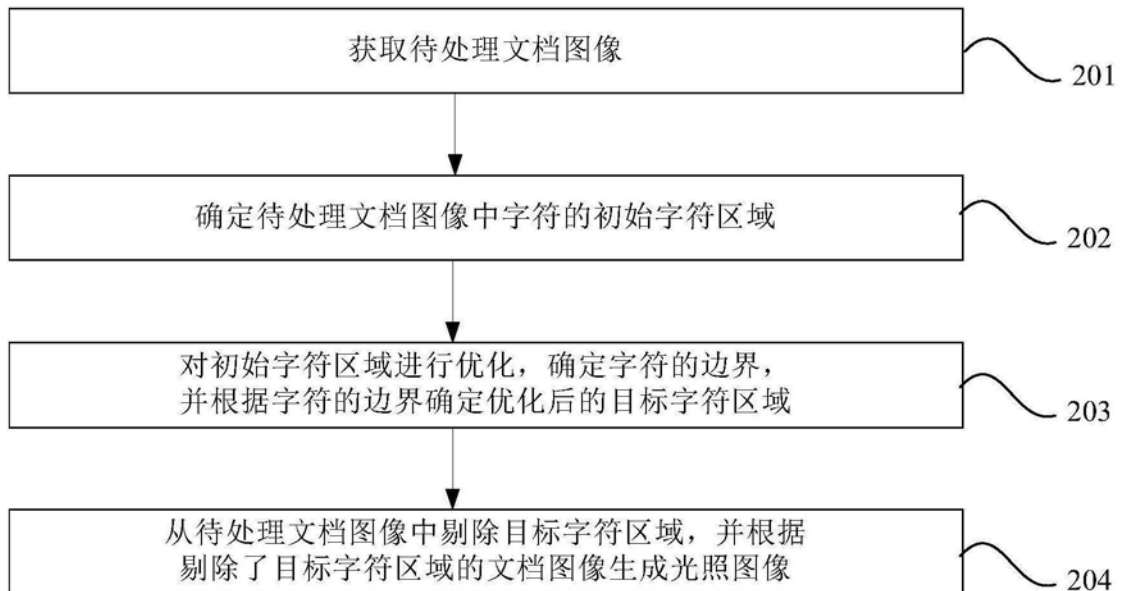


图2

# 静夜思

初始字符区域 唐 李白

床前明月光，疑是地上霜。  
举头望明月，低头思故乡。

图3a

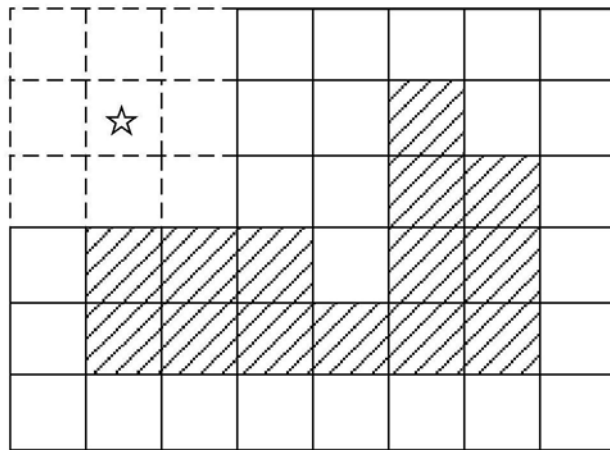


图3b

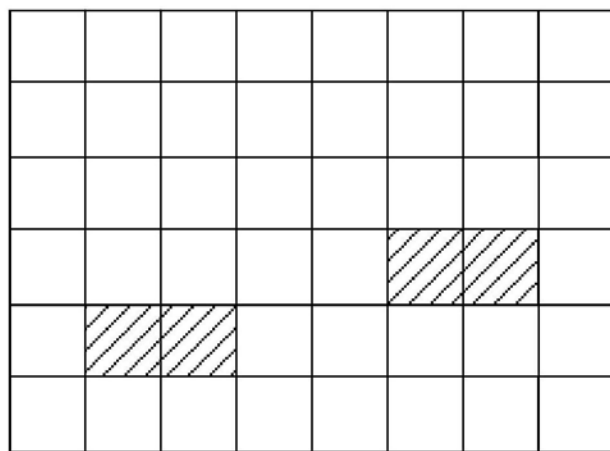


图3c

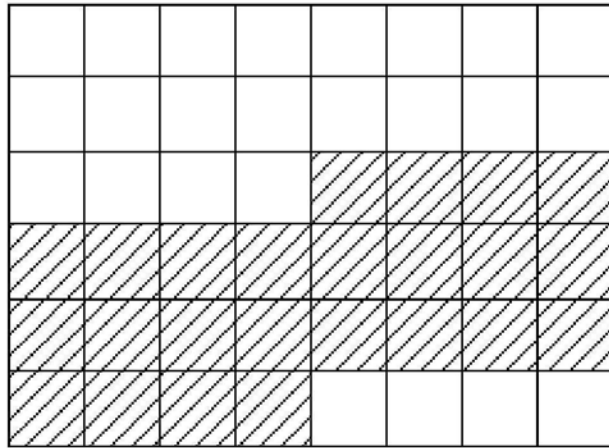


图3d

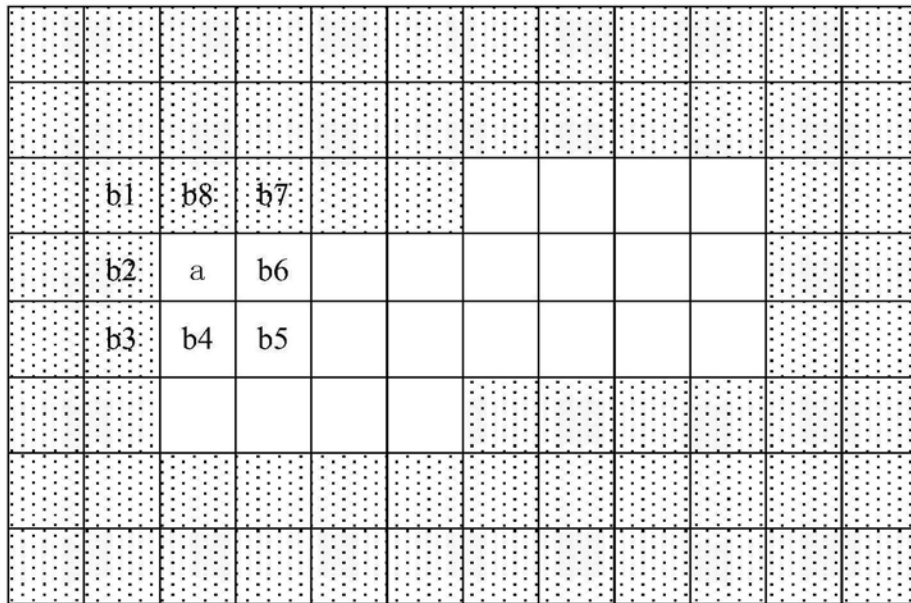


图3e

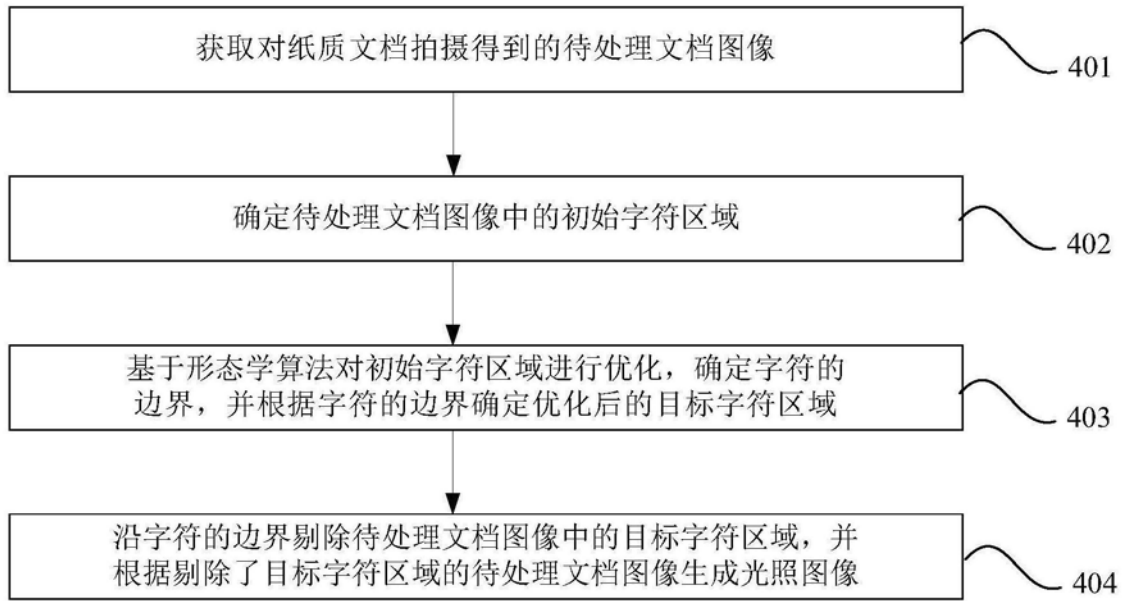


图4

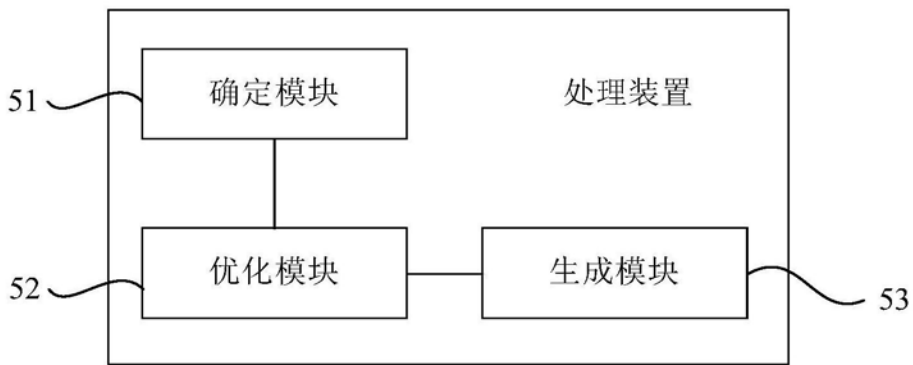


图5

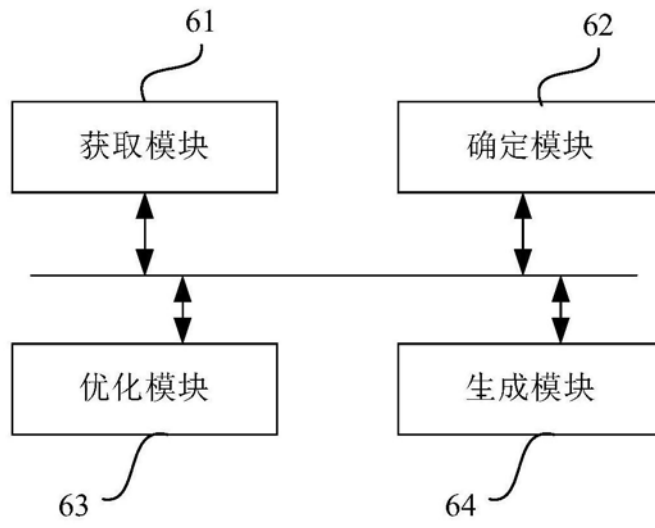


图6

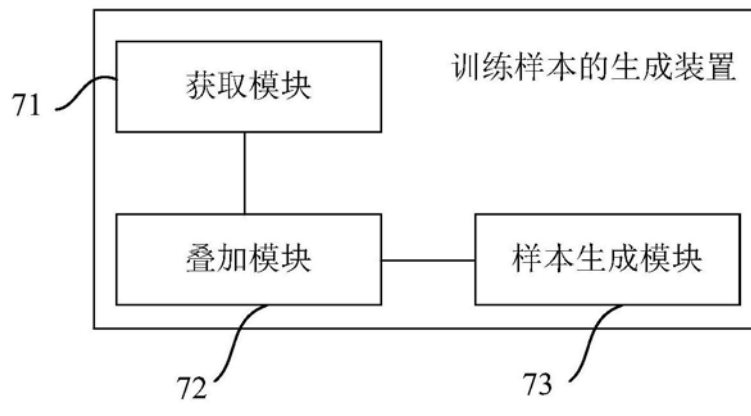


图7

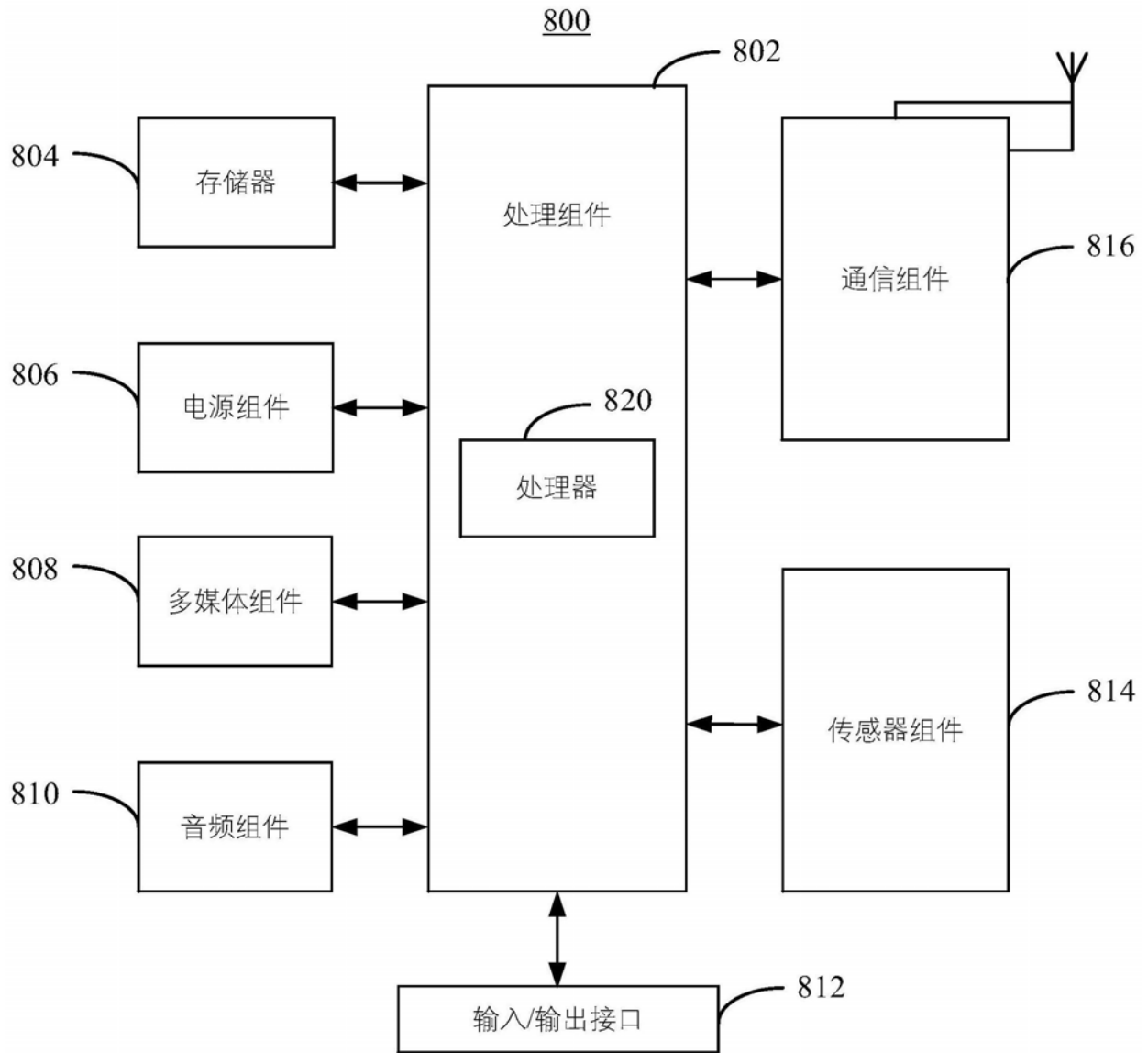


图8

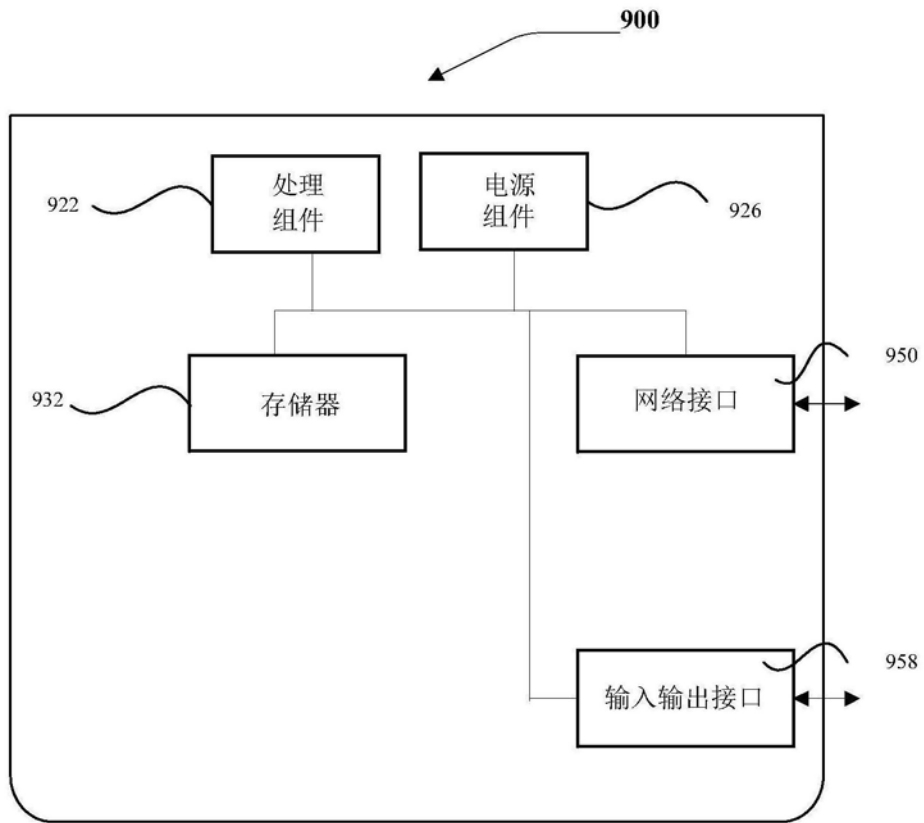


图9