



(12) 发明专利

(10) 授权公告号 CN 111461070 B

(45) 授权公告日 2023.12.08

(21) 申请号 202010356790.2

G06T 3/40 (2006.01)

(22) 申请日 2020.04.29

G06N 3/084 (2023.01)

(65) 同一申请的已公布的文献号

申请公布号 CN 111461070 A

(56) 对比文件

CN 110321913 A, 2019.10.11

CN 109948625 A, 2019.06.28

(43) 申请公布日 2020.07.28

CN 110175980 A, 2019.08.27

(73) 专利权人 OPPO广东移动通信有限公司

CN 104966269 A, 2015.10.07

地址 523860 广东省东莞市长安镇乌沙海
滨路18号

CN 110633755 A, 2019.12.31

审查员 马静

(72) 发明人 尚太章

(74) 专利代理机构 深圳市智圈知识产权代理事

务所(普通合伙) 44351

专利代理师 吕静

(51) Int. Cl.

G06V 30/413 (2022.01)

G06V 30/19 (2022.01)

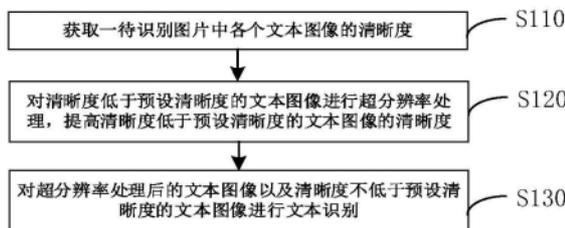
权利要求书2页 说明书11页 附图4页

(54) 发明名称

文本识别方法、装置、电子设备及存储介质

(57) 摘要

本申请公开了一种文本识别方法、装置、电子设备及存储介质,涉及图像处理技术领域。其中,该方法包括:获取一待识别图片中各个文本图像的清晰度,其中,不同文本图像为所述待识别图片中不同文本分别对应的图像;对清晰度低于预设清晰度的文本图像进行超分辨率处理,提高所述清晰度低于预设清晰度的文本图像的清晰度;对超分辨率处理后的文本图像以及清晰度不低于预设清晰度的文本图像进行文本识别。该方案提高了文本识别成功的概率。



1. 一种文本识别方法,其特征在于,所述方法包括:

获取一待识别图片中各个文本图像的清晰度,其中,不同文本图像为所述待识别图片中不同文本分别对应的图像,每个所述文本图像为单行文本对应的文本图像;

所述获取一待识别图片中各个文本图像的清晰度,包括:

将所述待识别图片输入清晰度识别模型,所述清晰度识别模型为多任务学习的神经网络模型,所述清晰度识别模型用于识别图片中的文本图像以及文本图像的清晰度,所述清晰度识别模型还用于识别各个文本图像的位置;

获取所述清晰度识别模型输出的各个文本图像的位置以及各个位置的文本图像对应的清晰度;

基于所述各个文本图像的位置从所述待识别图片中分割出所述各个文本图像;

对清晰度低于预设清晰度的文本图像进行超分辨率处理,提高所述清晰度低于预设清晰度的文本图像的清晰度,所述预设清晰度为文本识别算法识别成功概率大于预设概率的最低清晰度;

对超分辨率处理后的文本图像以及清晰度不低于预设清晰度的文本图像基于所述文本识别算法进行文本识别。

2. 根据权利要求1所述的方法,其特征在于,所述方法还包括:

获取训练样本集,所述训练样本集中包括多个图片,各个图片中文本位置、文本排列以及文本清晰度不同,每个图片中不同文本图像的位置以及不同文本图像的清晰度已经标记;

将所述训练样本集输入清晰度识别模型进行训练,根据清晰度识别模型识别出的文本图像的位置、各个位置的文本图像的清晰度、图片中标记的文本图像的位置以及图片中标记的各个文本图像的清晰度,优化所述清晰度识别模型的参数。

3. 根据权利要求1所述的方法,其特征在于,获取一待识别图片中各个文本图像的清晰度,包括:

通过文本位置检测算法检测待识别图片中各个文本的位置;

根据清晰度识别算法获取各个位置的文本对应的文本图像的清晰度。

4. 根据权利要求1至3任一项所述的方法,其特征在于,所述方法还包括:

对应所述待识别图片显示文字识别结果。

5. 一种文本识别装置,其特征在于,所述装置包括:

清晰度获取模块,用于获取一待识别图片中各个文本图像的清晰度,其中,不同文本图像为所述待识别图片中不同文本分别对应的图像,每个所述文本图像为单行文本对应的文本图像;所述获取一待识别图片中各个文本图像的清晰度,包括:将所述待识别图片输入清晰度识别模型,所述清晰度识别模型为多任务学习的神经网络模型,所述清晰度识别模型用于识别图片中的文本图像以及文本图像的清晰度,所述清晰度识别模型还用于识别各个文本图像的位置;获取所述清晰度识别模型输出的各个文本图像的位置以及各个位置的文本图像对应的清晰度;基于所述各个文本图像的位置从所述待识别图片中分割出所述各个文本图像;

超分辨率处理模块,用于对清晰度低于预设清晰度的文本图像进行超分辨率处理,提高所述清晰度低于预设清晰度的文本图像的清晰度,所述预设清晰度为文本识别算法识别

成功概率大于预设概率的最低清晰度；

文本识别模块,用于对超分辨率处理后的文本图像以及清晰度不低于预设清晰度的文本图像基于所述文本识别算法进行文本识别。

6.一种电子设备,其特征在于,包括:

一个或多个处理器;

存储器;

一个或多个程序,其中所述一个或多个程序被存储在所述存储器中并被配置为由所述一个或多个处理器执行,所述一个或多个程序被所述处理器执行用于执行如权利要求1-4任一项所述的方法。

7.一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中存储有程序代码,所述程序代码可被处理器调用执行如权利要求1-4任一项所述的方法。

文本识别方法、装置、电子设备及存储介质

技术领域

[0001] 本申请涉及图像处理技术领域,更具体地,涉及一种文本识别方法、装置、电子设备及存储介质。

背景技术

[0002] 在图像中可以存在有文本,即文本由图像中的像素点组成。对于图像中的文本,可以进行识别,获得图像中文本的信息,即确定图像中包括的文本为由哪些文字组成。但是,在对图像中文本进行识别时,即使能确定图像中包括文本,有时也会出现识别失败的现象。

发明内容

[0003] 鉴于上述问题,本申请提出了一种文本识别方法、装置、电子设备及存储介质,以改善上述问题。

[0004] 第一方面,本申请实施例提供了一种文本识别方法,所述方法包括:

[0005] 获取一待识别图片中各个文本图像的清晰度,其中,不同文本图像为所述待识别图片中不同文本分别对应的图像;对清晰度低于预设清晰度的文本图像进行超分辨率处理,提高所述清晰度低于预设清晰度的文本图像的清晰度;对超分辨率处理后的文本图像以及清晰度不低于预设清晰度的文本图像进行文本识别。

[0006] 第二方面,本申请实施例提供了一种文本识别装置,所述装置包括:清晰度获取模块,用于获取一待识别图片中各个文本图像的清晰度,其中,不同文本图像为所述待识别图片中不同文本分别对应的图像;超分辨率处理模块,用于对清晰度低于预设清晰度的文本图像进行超分辨率处理,提高所述清晰度低于预设清晰度的文本图像的清晰度;文本识别模块,用于对超分辨率处理后的文本图像以及清晰度不低于预设清晰度的文本图像进行文本识别。

[0007] 第三方面,本申请实施例提供了一种电子设备,包括:一个或多个处理器;存储器;一个或多个程序,其中所述一个或多个程序被存储在所述存储器中并被配置为由所述一个或多个处理器执行,所述一个或多个程序被所述处理器执行用于执行上述的方法。

[0008] 第四方面,本申请实施例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中存储有程序代码,所述程序代码可被处理器调用执行上述的方法。

[0009] 本申请实施例提供的文本识别方法、装置、电子设备及存储介质,对于待识别图片中的文本识别,可以获取待识别图片中不同文本分别对应的文本图像,并且,在文本图像的清晰度低于预设清晰度时,可以进行超分辨率处理,提高清晰度低于预设清晰度的文本图像的清晰度。进行文本识别时,对清晰度不低于预设清晰度的文本图像以及超分辨率处理后的文本图像进行文本识别,从而进行文本识别的文本图像清晰度较高,提高文本识别成功的概率。

附图说明

[0010] 为了更清楚地说明本申请实施例中的技术方案,下面将对实施例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的一些实施例,对于本领域技术人员来讲,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他的附图。

[0011] 图1示出了本申请一实施例提供的文本识别方法的流程图。

[0012] 图2示出了本申请另一实施例提供的文本识别方法的流程图。

[0013] 图3示出了本申请实施例提供的一待识别图片的示意图。

[0014] 图4A和图4B分别示出了从图3的待识别图片中提取的文本图像的示意图。

[0015] 图5示出了本申请另一实施例提供的文本识别过程的数据流向示意图。

[0016] 图6示出了本申请实施例提供的识别结果的一显示示意图。

[0017] 图7示出了本申请实施例提供的识别结果的另一显示示意图。

[0018] 图8示出了本申请又一实施例提供的文本识别方法的流程图。

[0019] 图9示出了本申请实施例提供的文本识别装置的功能模块图。

[0020] 图10示出了本申请实施例提供的电子设备的结构框图。

[0021] 图11是本申请实施例的用于保存或者携带实现根据本申请实施例的文本识别方法的程序代码的存储单元。

具体实施方式

[0022] 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。

[0023] 对于图片中的文字,由图片中的像素点组成,为图片的一部分,本身是图像而非文字,无法作为文字独立查看,无法进行直接编辑、复制以及剪切等文字处理操作。因此,通常需要对其进行识别,确定图片中像素点组成的文字为哪些文字,转换成一种计算机能够接受,人又能够理解的格式,可以对图片中的文字以文字的形式而非图像的形式进行重新呈现,如在可编辑文本框中显示,从而可以对获得的文字进行编辑、复制以及剪切等文字处理操作。

[0024] 另外,对于图片中的文字,由于本身是图像,可能因为分辨率过低,过于模糊,清晰度不够,用户无法有效辨识图片中的文字。因此,也可以对图片中的文字进行识别,将识别的到的文字重新以文字的形式而非图像的形式清晰呈现,或者将识别到的文字重新以清晰的图像形式进行呈现,使用户能清晰地查看到图片中包括的文字信息。

[0025] 对图像中由像素点组成的文字进行识别,确定图像中像素点组成的文字为哪些文字的过程可以定义为文本识别。在对文本识别过程中,即使能确定图像中包括文字,文本识别算法本身也是可用的,但是有时候会识别失败,无法确定图像中的文字为哪些文字,或者所确定的文字与图像中本来表现的文字不一致。发明人经过研究发现,在图片中文字的清晰度过低时,通常无法有效确定图片中的文字为哪些文字,容易识别失败。

[0026] 因此,本申请实施例提供了一种文本识别方法、装置、电子设备及存储介质,对待识别图片中不同文本对应的文本图像进行清晰度获取,并且对其中清晰度低于预设清晰度的文本图像进行超分辨率处理以提高其清晰度。对超分辨率处理后的文本图像以及待识别

图片中包括的清晰度不低于预设清晰度文本图像进行文本识别,由于所有进行识别的文本图像清晰度较高或者被提高,从而可以提高识别成功的概率。下面将通过具体实施例对本申请实施例提供的文本识别方法、装置、电子设备及存储介质进行详细说明。

[0027] 请参阅图1,示出了本申请实施例提供的文本识别方法。具体的,该方法包括:

[0028] 步骤S110:获取一待识别图片中各个文本图像的清晰度,其中,不同文本图像为所述待识别图片中不同文本分别对应的图像。

[0029] 在本申请实施例中,为便于描述,定义要进行文本识别的一个图片为一个待识别图片,对待识别图片中对应文字有文字的图像部分进行识别,确定待识别图片中显示的文字。

[0030] 在待识别图片中,不同的文本对应图像,定义每个文本对应的图像为文本图像,每个文本图像为待识别图片的一部分,或者说每个文本图像是待识别图片的部分图像区域。

[0031] 在文本图像中包括对应的文本,且仅包括对应的文本,不包括其他文本。可以理解的,文本图像包括的文本为由图像像素点形成的文本。

[0032] 在本申请实施例中,每个文本包括一个或多个文字。每个文本包括的字数在本申请实施例中并不限定。例如,每个文本可以仅包括预设单位的文字。该预设单位可以是行,即每个文本仅包括一行文字,每个文本图像仅包括一行文字;该预设单位可以是单个字,即每个文本仅包括一个文字,每个文本图像中仅包括一个文字;该预设单位可以是预设数量的字,即每个文本仅包括预设数量的文字,每个文本图像中仅包括该预设数量的文字。本申请实施例主要以行为例进行说明,即每个文本图像为单行文本对应的文本图像,其中仅包括单行文本;待识别图片中的每一行文本对应一个文本图像。

[0033] 在待识别图片中,文本中文字的大小不同,可能使文字的清晰度不同,从而文本的清晰度不同。因此,可以获取各个文本图像的清晰度。

[0034] 在本申请实施例中,定义对图像清晰度进行表示的值为清晰度值。其中,清晰度的评价方式并不限定,对于同一图像,每种图像清晰度评价方式可以对应计算得到相应的清晰度值,表示在该清晰度评价方式下表示的清晰度。例如,可以通过Brenner梯度函数、Tenengrad梯度函数、Laplacian梯度函数、SMD(灰度方差)函数、方差函数、能量梯度函数、Vollath函数、熵函数、EAV点锐度算法函数、Reblur二次模糊等清晰度评价方式中的一种或多种对图像的清晰度进行评价,根据每种清晰度评价方式对图像计算得到的清晰度值,表示在该评价方式下获得的图像的清晰度。

[0035] 另外,可选的,由于待识别图片中文字大小不同,清晰度也可能不同,也可以通过图片中文本的尺寸表示文本图像的清晰度。图片中文本的尺寸越小,清晰度越低。

[0036] 步骤S120:对清晰度低于预设清晰度的文本图像进行超分辨率处理,提高所述清晰度低于预设清晰度的文本图像的清晰度。

[0037] 对于清晰度过低的文本图像,若直接进行文本识别,可能由于清晰度过低而识别失败,因此,可以对其进行处理以提高其清晰度。

[0038] 在本申请实施例中,以清晰度低于预设清晰度的文本图像为清晰度过低的文本图像,对清晰度过低的文本图像进行超分辨率处理,以提高文本图像的清晰度。

[0039] 在本申请实施例中,表示预设清晰度的值定义为预设清晰度值,可以对应清晰度评价方式进行设置,不同评价清晰度的方式,可以对应设置不同的预设清晰度值。可以理解

的,若清晰度的评价方式中,计算获得的清晰度值越小,表示清晰度越高,则当文本图像的清晰度值高于预设清晰度值时,确定文本图像的清晰度低于预设清晰度;当文本图像的清晰度值低于预设清晰度值时,确定文本图像的清晰度高于预设清晰度。若清晰度的评价方式中,计算获得的清晰度值越小,表示清晰度越低,则当文本图像的清晰度值低于预设清晰度值时,确定文本图像的清晰度低于预设清晰度;当文本图像的清晰度值高于预设清晰度值时,确定文本图像的清晰度高于预设清晰度。

[0040] 步骤S130:对超分辨率处理后的文本图像以及清晰度不低于预设清晰度的文本图像进行文本识别。

[0041] 在进行文本识别时,对于清晰度过低的文本图像,在其超分辨率处理提高清晰度后进行识别;而对于待识别图片中包括的本身清晰度不低于预设清晰度的文本图像,可以直接进行文本识别。因此,可以实现对待识别图片中包括的不同文本的识别,确定待识别图片中包括哪些文字。

[0042] 在本申请实施例中,对待识别图片中的不同文本对应的文本图像进行清晰度获取,从而可以确定到待识别图片中清晰度过低的文本图像。对于清晰度本身不低的文本图像,可以直接进行文本识别;而对于清晰度过低的文本图像,可以进行超分辨率处理提高清晰度后,进行文本识别,从而,进行文本识别的文本图像均为清晰度较高的图像,提高了文本识别成功的概率。

[0043] 本申请还提供了一种实施例,描述了获取文本图像及其清晰度的方式。请参见图2,该实施例提供的方法包括:

[0044] 步骤S210:将待识别图片输入清晰度识别模型,所述清晰度识别模型用于识别图片中的文本图像以及文本图像的清晰度。

[0045] 步骤S220:获取所述清晰度识别模型输出的各个文本图像对应的清晰度。

[0046] 在本申请实施例中,待识别图片可以是执行本申请实施例方法的电子设备通过携带的摄像头拍摄获得,也可以是从其他设备获得,也可以是电子设备中存储的图片。

[0047] 对于不同的待识别图片,其拍摄场景不同、拍摄主体距离摄像头的距离不同、图片中文字大小不同、拍摄镜头的参数不同等原因,待识别图片中文字的清晰度可能不同。在本申请实施例中,可以通过清晰度识别模型获取待识别图片中各个文本图像的清晰度。

[0048] 具体的,用于对待识别图片中文本图像进行清晰度识别的清晰度识别模型为预先训练好的神经网络模型,可以识别图片中的文本图像以及文本图像的清晰度。将待识别图片输入清晰度识别模型,清晰度识别模型可以输出待识别图片中各个文本图像的清晰度。其中,识别图片中的文本图像,即从图片中提取文本图像,该文本图像中包括的文本仅为对应的文本。例如图3所示,可以识别到图3中待识别图片110中包括虚线框框出的文本图像111以及虚线框框出的文本图像112,并且也可以识别到文本图像111以及文本图像112的清晰度,图3中的虚线框仅用于说明文本图像,并非待识别图片的一部分。

[0049] 在本申请实施例中,可以通过训练样本集对清晰度识别模型进行训练。其中,训练样本集中可以包括大量独立的图片,每个图片中包括文本,训练样本集中的各个图片可以有各种尺寸,各个图片包括的文本可以涵盖了各种清晰度、各种大小、各种排列方式等不同文本,使训练样本集中的图片具有多样性,以提高训练完成的清晰度识别模型的清晰度识别能力。另外,在训练样本集的每个图片中,对应文本图像标记有清晰度,该标记的清晰度

作为文本图像清晰度的真值,以便于根据识别到的清晰度以及标记的清晰度计算损失函数。

[0050] 在训练时,将训练样本集中的图片输入待训练的清晰度识别模型。对于输入清晰度识别模型的图片,清晰度识别模型可以识别输入的图片中文本图像的清晰度,并根据识别出的清晰度以及文本图像标记的清晰度,优化清晰度识别模型的参数。如通过识别出的清晰度以及文本图像标记的清晰度确定清晰度识别模型的损失函数的值,在损失函数的值过大时,如大于预设的值,可以通过反向传播优化清晰度识别模型的参数,根据优化后的参数再次识别出清晰度,根据识别出的清晰度以及文本图像标记的清晰度再次计算损失函数的值,并判断损失函数的值是否足够小,如小于或等于预设值。在损失函数小于或等于预设的值的值的情况下,可以确定当前的清晰度识别模型识别出的清晰度接近于文本图像的实际清晰度;在损失函数大于预设的值时,表示该清晰度识别模型识别出的清晰度不够准确,继续优化清晰度识别模型的参数,直到损失函数的值足够小,小于或等于预设的值。

[0051] 在本申请实施例中,清晰度识别模型在输出各个文本图像的清晰度时,可以是清晰度与文本图像对应输出。其中,为准确定位到图片中的文本图像,文本图像可以通过文本图像的位置进行表示,即可以输出文本图像的位置与清晰度的对应关系。文本图像的位置为文本图像在待识别图片中的位置,该位置可以通过待识别图片的像素坐标进行表示。

[0052] 由于文本图像具有一定的形状,可以选取文本图像在待识别图片中的一个或多个坐标点对文本图像的位置进行表示。如下以文本图像为矩形图像进行说明。

[0053] 在一种实施方式中,在待识别图片中,两个对角的坐标点即可定位一个矩形区域,因此,可以通过文本图像的两个对角的坐标点表示文本图像的位置。如通过左上角和右下角的坐标点表示文本图像的位置,则文本图像在待识别图片中的位置为 (x_1, y_1) , (x_2, y_2) 时,表示该文本图像左上角的坐标点为待识别图片中的 (x_1, y_1) ,右下角的坐标点为待识别图片中的 (x_2, y_2) ,该文本图像为待识别图片中 (x_1, y_1) , (x_1, y_2) , (x_2, y_2) , (x_2, y_1) 四个坐标点围成的矩形图像区域。

[0054] 在一种实施方式中,也可以通过文本图像的四个角点或者三个角点在待识别图片中的坐标点,表示文本图像的位置。

[0055] 在一种实施方式中,也可以通过文本图像中的坐标点以及其他条件表示文本图像的位置。如通过文本图像的一个角点以及另一个角点与该角点之间形成的向量对文本图像进行表示。

[0056] 在本申请实施例中,训练好的清晰度识别模型还可以用于识别各个文本图像的位置,也就是说,清晰度识别模型可以识别到文本图像、文本图像的位置以及文本图像的清晰度。

[0057] 清晰度识别模型在输出文本图像的清晰度时,同时输出文本图像的位置,或者说,对应文本图像的位置输出清晰度,表示各个位置的文本图像的清晰度。因此获取清晰度识别模型输出的各个文本图像对应的清晰度时,可以获取清晰度识别模型输出的各个文本图像的位置以及各个位置的文本图像对应的清晰度。

[0058] 对应的,该清晰度识别模型可以是多任务学习的神经网络模型,对清晰度识别模型进行训练时,使清晰度识别模型同时学习文本图像的识别任务以及文本图像的清晰度识别任务。由于识别到文本图像,则对应获得文本图像的位置,对该清晰度识别模型的训练即

为训练清晰度识别模型输出文本图像的位置以及文本图像的清晰度。

[0059] 具体的,对清晰度识别模型进行训练的训练样本集中可以包括多个图片,各个图片中文本位置、文本排列以及文本清晰度不同,以能涵盖到各种情况下的文本位置、文本排列以及文本清晰度为宜,使训练好的清晰度识别模型能识别各种位置的文本图像的位置、各种清晰度的文本图像的清晰度。在每个图片中,不同文本图像的位置以及不同文本图像的清晰度已经标记,即训练集中的每个图片标记有文本图像的位置以及文本图像的清晰度,训练集中每个图片中各个文本图像的位置以及各个文本图像的清晰度已知。

[0060] 在训练时,将所述训练样本集输入清晰度识别模型进行训练,根据清晰度识别模型识别出的文本图像的位置、各个位置的文本图像的清晰度、图片中标记的文本图像的位置以及图片中标记的各个文本图像的清晰度,优化所述清晰度识别模型的参数。

[0061] 其中,清晰度识别模型同时进行位置识别的学习以及清晰度识别的学习,清晰度识别模型的损失函数可以包括位置损失函数以及清晰度损失函数的总和,可以表示为 $loss = L_1 + \lambda L_c$,其中, $loss$ 表示该清晰度识别模型的总的损失函数; L_1 表示位置损失函数,用于衡量文本图像在图片中的实际位置以及清晰度识别模型识别到的文本图像的位置之间的差距; L_c 表示清晰度损失函数,用于衡量文本图像的实际清晰度以及清晰度识别模型识别到的文本图像的清晰度之间的差距。文本图像的实际位置已知,即为预先进行标记的位置;文本图像的清晰度已知,也是预先标记。 λ 为调和系数,表示清晰度损失函数占据 $loss$ 函数的比重,具体的值可以根据经验设置,或者根据多次试验后选择一个最利于清晰度识别模型收敛的值,在本申请实施例中并无限定。

[0062] 在训练时,将训练样本集中的各个图片输入待训练的清晰度识别模型,根据图片中标记的各文本图像的实际位置以及识别出的文本图像的位置,可以确定位置损失函数,根据标记的各文本图像的实际清晰度以及识别出的文本图像的清晰度,可以确定清晰度损失函数,从而可以确定总的损失函数。根据总的函数是否足够小确定是否对清晰度识别模型进行反向传播以优化参数。若总的损失函数足够小,如小于或等于预设的值,可以确定当前的清晰度识别模型识别出的清晰度接近甚至等于文本图像的实际清晰度,识别出的文本图像的位置接近甚至等于文本图像的实际位置;若总的损失函数不是足够小,如大于预设的值,则可以继续反向传播以优化清晰度识别模型的参数,根据优化后的清晰度识别模型再次计算总的损失函数,直至损失函数足够小,如小于预设的值。

[0063] 在本申请实施例中,为了便于后续的文本识别以及超分辨率处理,还可以从图片中提取各个文本图像,使各个文本图像作为独立的图像。例如图4A所示的文本图像111以及图4B所示的文本图像112为根据图3所示的图片输出的独立的文本图像。在图4A以及图4B中,为了清楚进行示例说明,对文本图像111以及文本图像112进行了尺寸放大处理,在实际处理过程中,不一定会做该放大处理操作,可以对提取到的文本图像本身进行超分辨率处理或文本识别;当然,也可以对提取到的文本图像进行尺寸放大后再进行超分辨率处理或文本识别,在本申请实施例中并无限定。

[0064] 在一种实施方式中,清晰度识别模型在输出文本图像对应的清晰度时,可以对应输出该文本图像,即输出独立于图片的文本图像,或者说输出从图片中分割出的文本图像。

[0065] 在一种实施方式中,在清晰度识别模型输出文本图像的位置以及对应的清晰度后,可以根据图像分割算法以及清晰度识别模型输出位置,从图片中分割出各个文本图像。

具体的图像分割算法在本申请实施例中并不限定,如可以直接根据文本图像的位置,从图片中提取出各个文本图像,作为独立的图像。

[0066] 步骤S230:对清晰度低于预设清晰度的文本图像进行超分辨率处理,提高所述清晰度低于预设清晰度的文本图像的清晰度。

[0067] 若文本图像的清晰度低于预设清晰度,表示该文本图像的清晰度过低,可以对其进行提高清晰度的处理。其中,预设清晰度的设置在本申请实施例中并不限定,例如,可以是以文本识别算法能识别成功的概率大于预设概率的最低清晰度,作为预设清晰度。例如如图3所示的待识别图片110中识别到的文本图像111的清晰度低于预设清晰度,则对该文本图像111进行超分辨率处理提高其分辨率。

[0068] 对图像进行超分辨率处理,即对图像进行超分辨率重建,可以提高图像的分辨率。在更高分辨率的图像中,像素密度更高,每英寸图像内像素点数量越多,图像中存储的信息量越大,提供的图像细节越多,图像越清晰。

[0069] 在本申请实施例中,对清晰度低于预设清晰度的文本图像的超分辨率处理,可以是使处理后的文本图像的清晰度高于或等于预设清晰度,以使文本图像文本识别成功的概率更高。

[0070] 在本申请实施例中,具体的超分辨率算法并不限定。例如可以是SRCNN算算法,SRGAN算法以及ESRGAN算法等。

[0071] 步骤S240:对超分辨率处理后的文本图像以及清晰度不低于预设清晰度的文本图像进行文本识别。

[0072] 在本申请实施例中,文本识别的文本图像为超分辨率处理后的文本图像,以及图片中提取出的本身的清晰度高于或等于预设清晰度的文本图像。例如如图3所示的待识别图片110中识别到的文本图像111的清晰度低于预设清晰度,则对该文本图像111进行超分辨率处理后进行文本识别;文本图像112的清晰度高于预设清晰度,可以直接对该文本图像112进行文本识别。

[0073] 其中,具体可以通过文本识别算法进行识别,具体的文本识别算法在本申请实施例中并不限定,例如可以是CRNN算法、OCR算法等。

[0074] 具体的,本申请实施例可以是,如图5所示,超分辨率处理模型为通过超分辨率算法对图像进行超分辨率处理的算法模型,可以通过软件或硬件实现;文本识别模型为通过文本识别算法对图像进行文本识别的算法模型,可以通过软件或硬件实现;清晰度识别模型也可以通过软件或硬件实现。将待识别图片输入清晰度识别模型后,清晰度识别模型输出的清晰度高于或等于预设清晰度的文本图像再输入文本识别模型进行文本识别;清晰度识别模型输出的清晰度低于预设清晰度的文本图像输入超分辨率处理模型,对文本图像进行超分辨率处理,超分辨率处理模型输出的文本图像再输入文本识别模型进行文本识别,文本识别模型可以输出根据文本图像识别到的文本,其中,超分辨率处理模型输出的为超分辨率处理后的文本图像。因此,可以实现对待识别图片中不同分辨率的文本图像的有效识别,提高识别成功概率。

[0075] 另外,在本申请实施例中,还可以显示识别到的文本。其中,对应待识别图片显示文字识别结果,在显示待识别图片的同时显示识别到的文本。

[0076] 在一种实施方式中,可以将每个识别到的文本显示于待识别图片中对应的文本图

像的位置。在该实施方式中,识别到的文本显示于文本图像的位置时,可以将待识别图片中对应的文本覆盖,以避免识别到的文本与文本图像中的文本混合难以看清。

[0077] 在一种实施方式中,可以将识别到的文本与待识别图片中对应的文本图像排列显示,以便于用户查看。如对图6所示的待识别图片,将根据文本图像111识别到的文本“不清晰”与文本图像111上下排列显示;根据文本图像112识别到的文本“清晰”与文本图像112上下排列显示。

[0078] 在一种实施方式中,识别到的文本可以在与待识别图片不同的显示区域进行显示,即识别到的文本的显示与待识别图片并不重合。如图7所示,在显示界面100的不同区域显示待识别图片110以及识别到的文本“清晰”和“不清晰”,识别到的文本可以显示于文本框120中。

[0079] 可选的,在该实施方式中,文本显示时,其中文字的显示顺序以及行列数可以与文字在待识别图片中的排列顺序一致,如图7所示。

[0080] 另外,在本申请实施例中,识别到的文字也不与待识别图片在同一个显示界面进行显示,即显示识别到的文字时,可以不显示待识别到的图片。另外,可选的,文字的显示顺序以及行列数也可以与文字在待识别图片中的排列顺序一致。

[0081] 在本申请实施例中,通过清晰度识别模型可快速确定待识别图片中文本对应的文本图像的位置,同时识别文本图像的清晰度。对于清晰度不低于预设清晰度的文本图像,可以直接通过文本识别算法对其进行识别;对于清晰度低于预设清晰度的文本图像,通过超分辨率处理提高其清晰度,可以使其清晰度高于或等于预设清晰度,再进行文本识别。从而对于待识别图片中各种清晰度的文本,都可以具有更高的识别成功概率。

[0082] 本申请实施例提供的文本识别方法中,获取待识别图片中各个文本图像的清晰度还可以是检测出待识别图片中文本的位置后,根据检测到的文本位置获取文本图像的清晰度。具体的,请参见图8,本申请实施例提供的文本识别方法包括:

[0083] 步骤S310:通过文本位置检测算法检测待识别图片中各个文本的位置。

[0084] 文本检测算法可以检测待识别图片中各个文本的位置。文本检测算法具体为何种算法在本申请实施例中并不限定,例如可以是east算法, pixel-link算法以及Seglink算法等。

[0085] 其中,每个文本所包括的文字可以根据文本检测算法确定,例如,文本检测算法为检测图片中的每一行文字,获取到的是每一行文字的位置,则一个文本为一行文字,图片中有多少行文字,则可以检测到多少个文本,获得多少个文本的位置。

[0086] 在本申请实施例中,文本检测算法检测到的各个文本的位置即为文本图像的位置,例如,文本的位置为通过左上角 (x_1, y_1) 和右下角 (x_2, y_2) 的两个坐标点表示的矩形区域,则文本图像的位置也为左上角坐标点为 (x_1, y_1) 和右下角坐标点为 (x_2, y_2) 的矩形区域。从而,可以确定各个位置的文本分别对应的文本图像,从而可以对各个文本图像进行清晰度的确定等操作。

[0087] 可选的,在本申请实施例中,还可以包括文本图像的分割。即可以根据文本的位置,从待识别图片中分割出各个文本图像。具体的,可以以文本的位置作为文本图像的位置,通过图像分割算法,从待识别图片中分割出各个文本图像,作为独立的图像,并对各个独立的文本图像进行清晰度的获取。

[0088] 可选的,在本申请实施例中,文本检测算法检测到文本的位置后,也可以直接根据文本的位置输出各文本对应的文本图像,用于进行清晰度确定。

[0089] 步骤S320:根据清晰度识别算法获取各个位置的文本对应的文本图像的清晰度。

[0090] 各个位置的文本对应的文本图像,即为各个文本分别对应的文本图像。可以对各个文本图像进行清晰度识别,确定各个文本图像的清晰度。

[0091] 在本申请实施例中,具体的清晰度识别算法并不限定,可以是前述实施例中任意一种清晰度评价方式对应的算法或函数。

[0092] 步骤S330:对清晰度低于预设清晰度的文本图像进行超分辨率处理,提高所述清晰度低于预设清晰度的文本图像的清晰度。

[0093] 步骤S340:对超分辨率处理后的文本图像以及清晰度不低于预设清晰度的文本图像进行文本识别。

[0094] 步骤S330和步骤S340的描述可以参见前述实施例,在此不再赘述。

[0095] 在本申请实施例中,在确定待识别图片中文本的位置的情况下,可以对应记录各个文本的位置或者各个文本图像的位置,从而在显示时,可以根据文本的位置或者文本图像的位置,确定识别结果中的文本的显示位置。具体显示位置可以参见前述实施例,在此不再赘述。

[0096] 在本申请实施例中,可以对待识别图片进行文本位置的检测,并对检测到的各个位置的文本分别对应的文本图像获取其清晰度。对清晰度低于预设清晰度的文本图像进行超分辨率处理后进行文本识别,对待识别图片中提取的文本图像中,本身清晰度不低于预设清晰度的文本图像直接进行文本识别,从而提高了待识别图片的文本识别成功概率。

[0097] 本申请实施例还提供了一种文本识别装置400。如图9所示,该文本识别装置400可以包括:清晰度获取模块410,用于获取一待识别图片中各个文本图像的清晰度,其中,不同文本图像为所述待识别图片中不同文本分别对应的图像;超分辨率处理模块420,用于对清晰度低于预设清晰度的文本图像进行超分辨率处理,提高所述清晰度低于预设清晰度的文本图像的清晰度;以及文本识别模块430,用于对超分辨率处理后的文本图像以及清晰度不低于预设清晰度的文本图像进行文本识别。

[0098] 可选的,清晰度获取模块410可以用于将所述待识别图片输入清晰度识别模型,所述清晰度识别模型用于识别图片中的文本图像以及文本图像的清晰度;获取所述清晰度识别模型输出的各个文本图像对应的清晰度。

[0099] 可选的,清晰度识别模型还可以用于识别各个文本图像的位置,清晰度获取模块410可以用于获取所述清晰度识别模型输出的各个文本图像的位置以及各个位置的文本图像对应的清晰度。

[0100] 可选的,该装置还可以包括训练模块,用于获取训练样本集,所述训练样本集中包括多个图片,各个图片中文本位置、文本排列以及文本清晰度不同,每个图片中不同文本图像的位置以及不同文本图像的清晰度已经标记;将所述训练样本集输入清晰度识别模型进行训练,根据清晰度识别模型识别出的文本图像的位置、各个位置的文本图像的清晰度、图片中标记的文本图像的位置以及图片中标记的各个文本图像的清晰度,优化所述清晰度识别模型的参数。

[0101] 可选的,清晰度获取模块410可以用于通过文本位置检测算法检测待识别图片中

各个文本的位置;根据清晰度识别算法获取各个位置的文本对应的文本图像的清晰度。

[0102] 可选的,该装置还可以包括图像分割模块,用于从所述待识别图片中分割出各个文本图像。

[0103] 可选的,该装置还可以包括显示模块,用于从所述待识别图片中分割出各个文本图像。

[0104] 本申请实施例提供的文本识别方法及装置可以用于对各种图片进行文本识别。例如,拍摄包括文字但是距离镜头较远的主体时,拍摄获得的图片中文本的清晰度通常可能较低,识别困难。通过本申请实施例的文本识别方法及装置,可以有效提高清晰度过低的文本对应的文本图像的清晰度,同时提高了文本图像中清晰度较低的文本清晰度,提高了识别成功的概率,改善了用户对图片中文本识别的使用体验。并且,可以仅对清晰度低于预设清晰度的文本图像进行超分辨率处理,而非对所有清晰度的文本图像进行超分辨率处理,减小了数据处理量,减少运行的次数,降低时间复杂度,提高了处理速度。

[0105] 所属领域的技术人员可以清楚地了解到,为描述的方便和简洁,上述的各个方法实施例之间可以相互参照;上述描述装置和模块的具体工作过程,可以参考前述方法实施例中的对应过程,在此不再赘述。

[0106] 在本申请所提供的几个实施例中,模块相互之间的耦合可以是电性,机械或其它形式的耦合。

[0107] 另外,在本申请各个实施例中的各功能模块可以集成在一个处理模块中,也可以是各个模块单独物理存在,也可以两个或两个以上模块集成在一个模块中。上述集成的模块既可以采用硬件的形式实现,也可以采用软件功能模块的形式实现。各个模块可以配置在不同的电子设备中,也可以配置在相同的电子设备中,本申请实施例并不限定。

[0108] 请参考图10,其示出了本申请实施例提供的一种电子设备500的结构框图。该电子设备可以是手机、计算机、平板电脑、智能穿戴设备、摄像机等各种智能设备。该电子设备可以包括一个或多个处理器510(图中仅示出一个),存储器520以及一个或多个程序。其中,所述一个或多个程序被存储在所述存储器520中,并被配置为由所述一个或多个处理器510执行。所述一个或多个程序被处理器执行用于执行前述实施例所描述的方法。

[0109] 处理器510可以包括一个或者多个处理核。处理器510利用各种接口和线路连接整个电子设备500内的各个部分,通过运行或执行存储在存储器520内的指令、程序、代码集或指令集,以及调用存储在存储器520内的数据,执行电子设备500的各种功能和处理数据。可选地,处理器510可以采用数字信号处理(Digital Signal Processing,DSP)、现场可编程门阵列(Field-Programmable Gate Array,FPGA)、可编程逻辑阵列(Programmable Logic Array,PLA)中的至少一种硬件形式来实现。处理器510可集成中央处理器(Central Processing Unit,CPU)、图像处理器(Graphics Processing Unit,GPU)和调制解调器等中的一种或几种的组合。其中,CPU主要处理操作系统、用户界面和应用程序等;GPU用于负责显示内容的渲染和绘制;调制解调器用于处理无线通信。可以理解的是,上述调制解调器也可以不集成到处理器510中,单独通过一块通信芯片进行实现。

[0110] 存储器520可以包括随机存储器(Random Access Memory,RAM),也可以包括只读存储器(Read-Only Memory)。存储器520可用于存储指令、程序、代码、代码集或指令集。存储器520可包括存储程序区和存储数据区,其中,存储程序区可存储用于实现操作系统的指

令、用于实现至少一个功能的指令、用于实现上述各个方法实施例的指令等。存储数据区还可以电子设备在使用中所创建的数据等。

[0111] 另外,该电子设备还可以包括显示屏,用于对待识别图片进行显示,也用于对文本识别结果进行显示。

[0112] 请参考图11,其示出了本申请实施例提供的一种计算机可读存储介质的结构框图。该计算机可读存储介质700中存储有程序代码,所述程序代码可被处理器调用执行上述方法实施例中所描述的方法。

[0113] 计算机可读存储介质700可以是诸如闪存、EEPROM(电可擦除可编程只读存储器)、EPROM、硬盘或者ROM之类的电子存储器。可选地,计算机可读存储介质700包括非易失性计算机可读介质(non-transitory computer-readable storage medium)。计算机可读存储介质700具有执行上述方法中的任何方法步骤的程序代码710的存储空间。这些程序代码可以从一个或者多个计算机程序产品中读出或者写入到这一个或者多个计算机程序产品中。程序代码710可以例如以适当形式进行压缩。

[0114] 最后应说明的是:以上实施例仅用以说明本申请的技术方案,而非对其限制;尽管参照前述实施例对本申请进行了详细的说明,本领域的普通技术人员当理解:其依然可以对前述各实施例所记载的技术方案进行修改,或者对其中部分技术特征进行等同替换;而这些修改或者替换,并不驱使相应技术方案的本质脱离本申请各实施例技术方案的精神和范围。

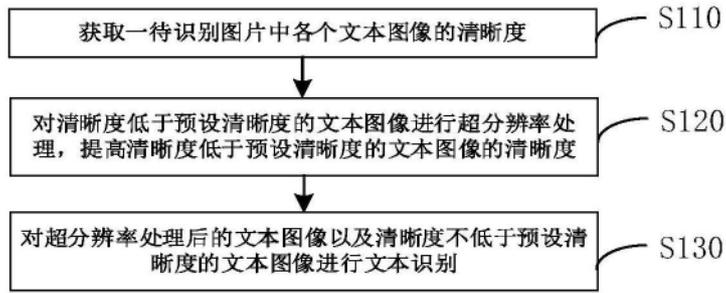


图1

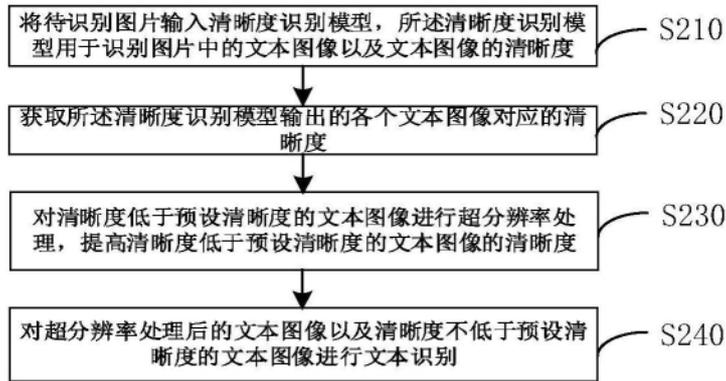


图2

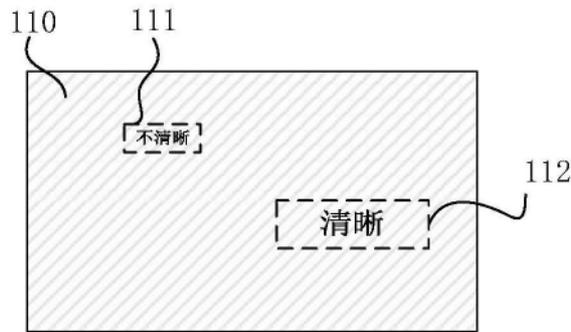


图3

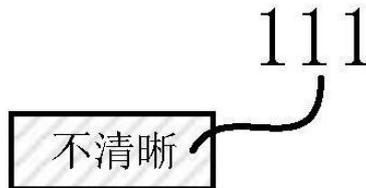


图4A



图4B

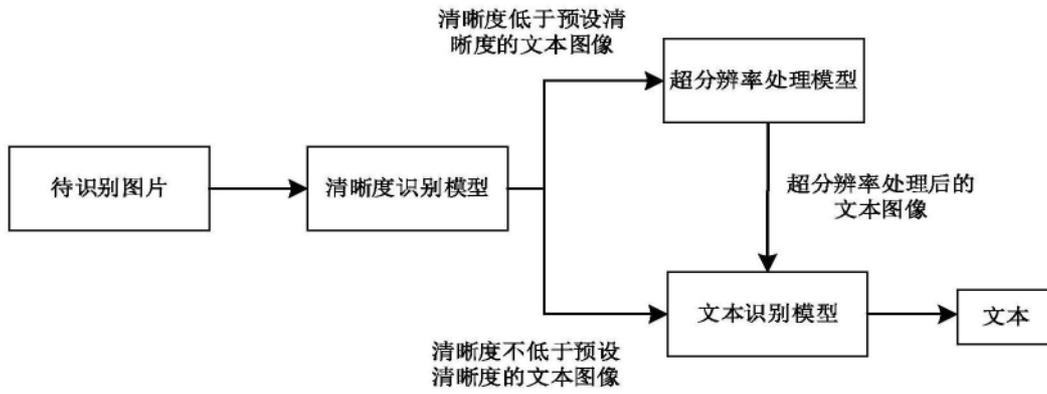


图5

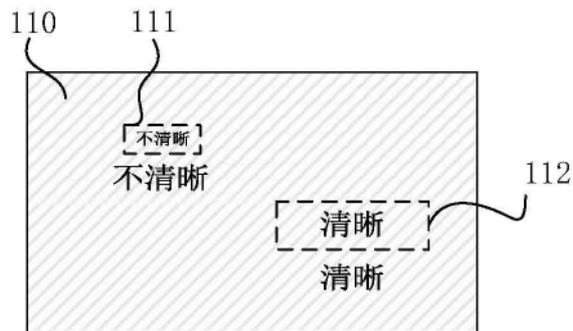


图6

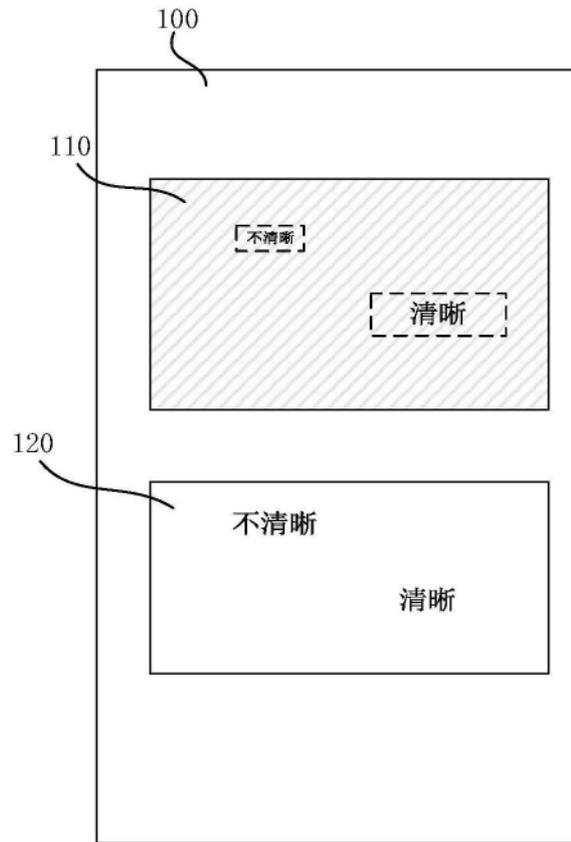


图7

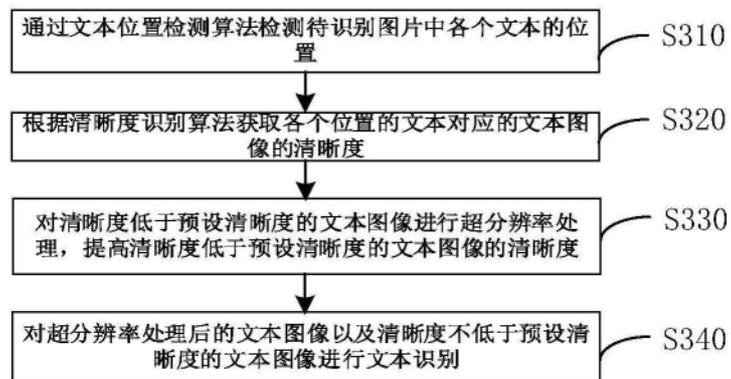


图8



图9



图10

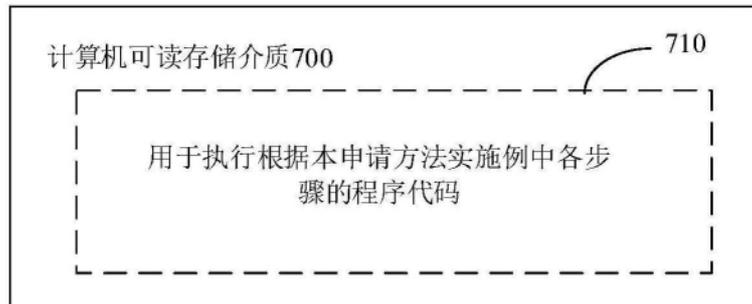


图11