



(12) 发明专利申请

(10) 申请公布号 CN 112200107 A

(43) 申请公布日 2021.01.08

(21) 申请号 202011109208.9

(22) 申请日 2020.10.16

(71) 申请人 深圳市华付信息技术有限公司

地址 518000 广东省深圳市前海深港合作
区前湾一路1号A栋201室(入驻深圳市
前海商务秘书有限公司)

(72) 发明人 文戈 陈兴委 刘磊 周先得
黄军文 汤红

(74) 专利代理机构 深圳市中兴达专利代理有限
公司 44637

代理人 黄美玲

(51) Int. Cl.

G06K 9/00 (2006.01)

G06K 9/20 (2006.01)

G06K 9/62 (2006.01)

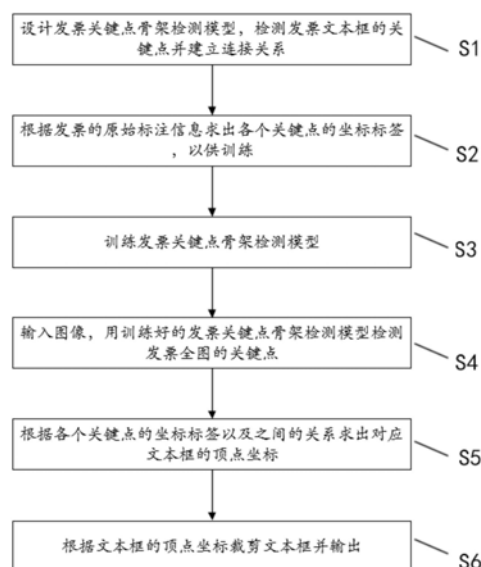
权利要求书1页 说明书4页 附图2页

(54) 发明名称

一种发票文本检测方法

(57) 摘要

本发明属于文本识别技术领域,尤其涉及一种发票文本检测方法,包括以下步骤:S1、设计发票关键点骨架检测模型,检测发票文本框的关键点并建立连接关系;S2、根据发票的原始标注信息求出各个关键点的坐标标签,以供训练;S3、训练发票关键点骨架检测模型;S4、输入图像,用训练好的发票关键点骨架检测模型检测发票全图的关键点;S5、根据各个关键点的坐标标签以及之间的关系求出对应文本框的顶点坐标;S6、根据文本框的顶点坐标裁剪文本框并输出,其中关键点的位置覆盖了所有可能检测到的字符位置,能够直接推断出文本框的位置参数,处理速度快;并且无论发票的图像在哪个方向发生形变,关键点的数值也会根据图像形变而自动调整,鲁棒性高。



1. 一种发票文本检测方法,其特征在于,包括以下步骤:

S1、设计发票关键点骨架检测模型,检测发票文本框的关键点并建立连接关系;

S2、根据发票的原始标注信息求出各个关键点的坐标标签,以供训练;

S3、训练发票关键点骨架检测模型;

S4、输入图像,用训练好的发票关键点骨架检测模型检测发票全图的关键点;

S5、根据各个关键点的坐标标签以及之间的关系求出对应文本框的顶点坐标;

S6、根据文本框的顶点坐标裁剪文本框并输出。

2. 根据权利要求1所述一种发票文本检测方法,其特征在于,S1中,发票文本框包括发票代码、发票号码、发票日期、发票金额、发票税额、发票总额和校验码;

所述发票代码包含两个关键点,分别在左右两端,并将两个关键点相连;

所述发票号码包含两个关键点,分别在左右两端,并将两个关键点相连;

所述发票代码的右端关键点与所述发票号码的左端关键点相连;

所述发票日期包含两个关键点,分别在左右两端,并将两个关键点相连;

所述发票号码和所述发票日期的左右两个端点分别相连;

所述发票金额包含一个关键点,在其中心位置;

所述发票税额包含一个关键点,在其中心位置;

所述发票日期左右两个关键点分别和所述发票金额、所述发票税额的关键点相连;

所述发票总额包含一个关键点,在其中心位置;

所述发票总额的关键点分别和所述发票金额、所述发票税额的关键点相连;

所述发票校验码包含三个关键点,两个在左右两端,一个在其中心位置,左右两端的关键点和中间的关键点相连;

当所述校验码在所述发票代码下方时,所述校验码和所述发票代码的左右两个关键点分别相连;

当所述校验码在所述发票总额下方时,所述校验码的左右两个关键点分别和所述发票总额的关键点相连。

3. 根据权利要求2所述一种发票文本检测方法,其特征在于,S2中,发票的原始标注信息为文本框的中心坐标或两端坐标。

4. 根据权利要求3所述一种发票文本检测方法,其特征在于,S3中,在制作训练样本时,当校验码在其中一个位置出现时,样本另一个位置的关键点设为不可见状态。

5. 根据权利要求2-4任一所述一种发票文本检测方法,其特征在于,S5中,对具有一个关键点的文本框和具有两个以上关键点的文本框的顶点坐标分开处理。

一种发票文本检测方法

技术领域

[0001] 本发明属于文本识别技术领域,尤其涉及一种发票文本检测方法。

背景技术

[0002] 发票的文本检测是发票识别任务中的关键环境,如何快速准备地将票据面上所需文本行检测出来是票据识别技术的关键。

[0003] 现有方法之一是采用传统机器学习方法,先用图像处理技术对输入票据图像做去噪、灰度化、二值化等处理,再用形态学、轮廓提取等方法对票据位置进行大致确定或者给出候选区域,最后使用局部特征结合分类器的方法对候选区域进行筛选确定最终的文本区域。但这些方法的准确率低,鲁棒性差,不适合商用。

[0004] 现有方法之二是使用深度学习技术对发票文本框进行检测,如使用通用检测模型来检测票据的文本,得益于卷积神经网络强大的特征提取能力,该方法的效果比传统方法有较大提高,但通用检测模型一般只能检测水平的矩形框,由于票据文本框常常存在倾斜的情况,这样输出的检测框往往会留下比较大的间隙,影响后续的字符分类效果。如果在检测模块后再加入方向矫正模块又会增加系统的冗余度,影响运行速度。

[0005] 现有方法之三是先对发票进行关键点检测以获取发票中特征明显的两个位置,再以这两个位置为基础对发票进行方向矫正与无关区域裁剪,获得裁剪对齐后的发票图片,再在此图片上进行文本框检测,一般采用两步法,先通过关键点检测及相关处理操作矫正发票位置,再通过对矫正后图像进行文本检测输出所有文本框,该方法具有鲁棒性好,准确率好的特点,但其整体过程显的冗余,在流程及速度上还有优化空间。

发明内容

[0006] 为了克服以上的技术问题,本发明提出一种发票文本检测方法,包括以下步骤:

[0007] 一种发票文本检测方法,包括以下步骤:

[0008] S1、设计发票关键点骨架检测模型,检测发票文本框的关键点并建立连接关系;

[0009] S2、根据发票的原始标注信息求出各个关键点的坐标标签,以供训练;

[0010] S3、训练发票关键点骨架检测模型;

[0011] S4、输入图像,用训练好的发票关键点骨架检测模型检测发票全图的关键点;

[0012] S5、根据各个关键点的坐标标签以及之间的关系求出对应文本框的顶点坐标;

[0013] S6、根据文本框的顶点坐标裁剪文本框并输出。

[0014] 进一步地,S1中需要检测的发票文本框包括发票代码、发票号码、发票日期、发票金额、发票税额、发票总额和校验码;

[0015] 所述发票代码包含两个关键点,分别在左右两端,并将两个关键点相连;

[0016] 所述发票号码包含两个关键点,分别在左右两端,并将两个关键点相连;

[0017] 所述发票代码的右端关键点与所述发票号码的左端关键点相连;

[0018] 所述发票日期包含两个关键点,分别在左右两端,并将两个关键点相连;

- [0019] 所述发票号码和所述发票日期的左右两个端点分别相连；
- [0020] 所述发票金额包含一个关键点，在其中心位置；
- [0021] 所述发票税额包含一个关键点，在其中心位置；
- [0022] 所述发票日期左右两个关键点分别和所述发票金额、所述发票税额的关键点相连；
- [0023] 所述发票总额包含一个关键点，在其中心位置；
- [0024] 所述发票总额的关键点分别和所述发票金额、所述发票税额的关键点相连；
- [0025] 所述发票检验码包含三个关键点，两个在左右两端，一个在其中心位置，左右两端的关键点和中间的关键点相连；
- [0026] 当所述校验码在所述发票代码下方时，所述校验码和所述发票代码的左右两个关键点分别相连；
- [0027] 当所述校验码在所述发票总额下方时，所述校验码的左右两个关键点分别和所述发票总额的关键点相连。
- [0028] 进一步地，S2中发票的原始标注信息为文本框的中心坐标或两端坐标。
- [0029] 该发明一种发票文本检测方法的有益效果：关键点的位置覆盖了所有可能检测到的字符位置，能够直接推断出文本框的位置参数，处理速度高；并且无论发票的图像在哪个方向发生形变，只要关键点检测准确，用于推断文本框的关键点向量的数值也会根据图像形变的情况而自动调整，鲁棒性高。

附图说明

- [0030] 图1为本发明实施例中发票文本检测方法的步骤示意图；
- [0031] 图2、图3为本发明实施例中求文本框顶点坐标的示意图；

具体实施方式

- [0032] 以下结合实施例对本发明作进一步的阐述，所述的实施例仅为本发明一部分的实施例，这些实施例仅用于解释本发明，对本发明的范围并不构成任何限制。
- [0033] 如说明书附图1所示，一种发票文本检测方法，包括以下步骤：
- [0034] S1、设计发票关键点骨架检测模型，检测发票文本框的关键点并建立连接关系
- [0035] 借鉴人体关键点骨架检测的方法，设计发票关键点骨架检测模型，其中以普通发票为例，需要检测的文本框为七个，分别为发票代码，发票号码，发票日期，发票金额，发票税额，发票总额，校验码。由于不同的文本框位置和大小不一，本方法根据实际情况设计关键点的个数及其连接，亦可根据自己的检测需求设计要检测的关键点和连接关系。本实施例中，对普通发票的关键点及其连接如下，
- [0036] 发票代码包含两个关键点，分别在左右两端，左右两个关键点相连；
- [0037] 发票号码包含两个关键点，分别在左右两端，左右两个关键点相连；
- [0038] 发票代码右侧关键点与发票号码左侧关键点相连接；
- [0039] 发票日期包含两个关键点，分别为左右两端，左右两个关键点相连；
- [0040] 发票号码和发票日期的左右两个关键点分别相连；
- [0041] 发票金额包含一个关键点，在其中心位置；

- [0042] 发票税额包含一个关键点,在其中心位置;
- [0043] 发票日期左右两个关键点分别和发票金额、发票税额的关键点分别相连;
- [0044] 发票总额包含一个关键点,在其中心位置;
- [0045] 发票金额和发票税额的关键点相连;
- [0046] 发票总额与发票金额、发票税额同时相连;
- [0047] 发票校验码包含三个关键点,两个在两端位置,一个在中心位置,左右两端的关键点都和中央关键点相连;
- [0048] 当校验码在发票代码下方时,校验码左右两个关键点和发票代码左右两个关键点分别相连;
- [0049] 当校验码在发票总额下方时,校验码左右关键点都和发票总额相连。
- [0050] S2、根据发票的原始标注信息求出各个关键点的坐标标签,以供训练
- [0051] 设计好关键点的具体骨架后,需要根据发票的原始标注信息求出于所设计关键点的标注坐标,原始标注信息一般为文本框的四个顶点坐标,而此设计关键点一般处在文本框的中心或者两端位置,两端的关键点坐标可以通过求两边顶点坐标的均值得来,中心关键点坐标可以通过求四个顶点坐标的均值得来。
- [0052] 在该实施例中,采用的网络架构是light-openpose,在训练过程中,关键点的连接关系会作为监督信息促进关键点位置预测的准确性,在网络训练好后,输入普通发票图片,网络应能输出各个关键点的位置信息及其属性。
- [0053] S3、训练发票关键点骨架检测模型
- [0054] 由于校验码会在两个位置随机出现(校验码在发票代码下方,或者校验码在发票总额下方),本方法在这两个位置上都设置了关键点,在制作训练样本时,当校验码在其中一个位置出现时,样本另一个位置的关键点设为不可见状态。网络训练好后预测时,不可见的关键点在逻辑上存在,但模型并不会输出不可见的关键点,这样可以保证模型能对这两种情况有合理的输出。则按照本方法的设计,一张普通发票上共有15个关键点,当校验码处在发票总金额下方时,设计在发票代码下方的校验码三个关键点不可见,当校验码处于发票代码下方时,设计在发票总金额下方的校验码三个关键点不可见。
- [0055] S4、输入图像,用训练好的发票关键点骨架检测模型检测发票全图的关键点
- [0056] S5、根据各个关键点的坐标标签以及之间的关系求出对应文本框的顶点坐标
- [0057] 从设计的关键点结构可以看到,发票代码,发票号码,发票日期,校验码这几个文本两端都有关键点,而发票金额、发票税额、发票总额三个金额文本各自只有一个关键点并且在处在文本中央,对这两种文本进行分开处理。
- [0058] 首先我们根据发票金额、发票税额、发票总额这三处的关键点(分别表示为PointA,PointB,PointC)来确定三个金额文本框顶点的坐标,如说明书附图3所示。检测到的关键点处于文本的中央位置,确定文本框的顶点坐标还需要文本框的高度、宽度以及旋转角度信息,具体来说只要求得两个正交向量Vec_hBias,Vec_wBias即可,这两个向量可以由三个关键点的坐标计算得来。
- [0059]
$$\text{Vec_wBias} = a * (\text{PointA} - \text{PointB})$$
- [0060]
$$\text{Vec_hBias} = b * \text{Vec_Height}$$
- [0061] Vec_Height为PointC引到(PointA-PointB)上的向量,它与(PointA-PointB)垂

直,长度为PointC到PointA,PointB所在直线的距离。

[0062] a,b为可调参数,具体可根据实际情况而设定,这里发票金额、发票税额、发票总额文本框都使用同样的Vec__hBias,Vec__wBias,也可以根据实际情况加上参数对尺度进行微调。

[0063] 求得Vec__hBias,Vec__wBias后用它们与关键点坐标进行加减运算可以求得文本框的顶点坐标,从而确定文本框的位置。其中Vec__Height求出后,还会在其他文本框的运算中用到。

[0064] 对于两端都有关键点的文本(发票代码,发票号码,发票日期,校验码)来说,分别由两端的关键点(分别表示为keyPointLeft,keyPointRight)再结合Vec__hBias,Vec__wBias求得文本框顶点坐标,如说明书附图3所示,这里的Vec__hBias,Vec__wBias求法如下:

[0065] $Vec_hBias = c * Vec_Height$

[0066] $Vec_wBias = d * (keyPointLeft - keyPointRight)$

[0067] $PointA = keyPointLeft + Vec_hBias + Vec_wBias$

[0068] $PointB = keyPointRight + Vec_hBias - Vec_wBias$

[0069] $PointC = keyPointRight - Vec_hBias - Vec_wBias$

[0070] 其中c,d为可调参数,对于不同属性文本框取值可以进行调整。校验码有三个关键点,在此只取两端关键点进行运算即可。

[0071] S6、根据文本框的顶点坐标裁剪文本框并输出

[0072] 通过网络输出的关键点坐标求出所有需要文本框的顶点坐标,由于关键点的属性已由网络输出,文本框的属性也可得到,整个发票文本检测过程完成。

[0073] 该方法仅需一个轻量级的网络即可得到发票关键点,再辅以简单的后处理即可确定文本框的位置,在处理速度上较多步法优势明显;并且得益于发票关键点及其连接设计的合理性,这些关键点在发票上的相对位置与结构都比较稳定,实际预测时也可以达到较高的准确率;另外关键点的检测稳定性相对于检测框来说要更强,在旋转、尺度变换等情况仍能保持较佳的准确度,由于关键点的稳定,也使得最后得到的文本框精度也有较高的鲁棒性。

[0074] 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,任何熟悉本专业的技术人员,在不脱离本发明技术方案范围内,当可利用上述揭示的技术内容作出些许更动或修饰为等同变化的等效实施例,但凡是未脱离本发明技术方案内容,依据本发明的技术实质对以上实施例所作的任何简单修改、等同变化与修饰,均仍属于本发明技术方案的范围。

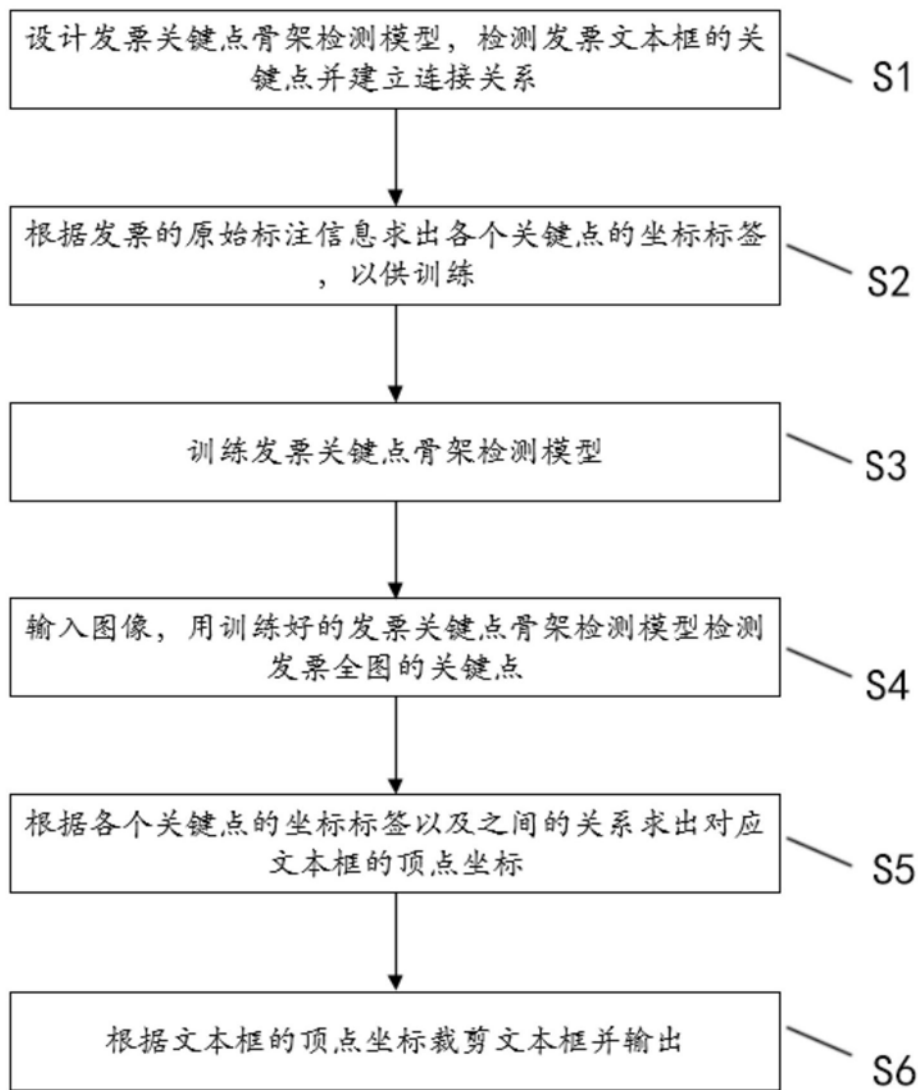


图1

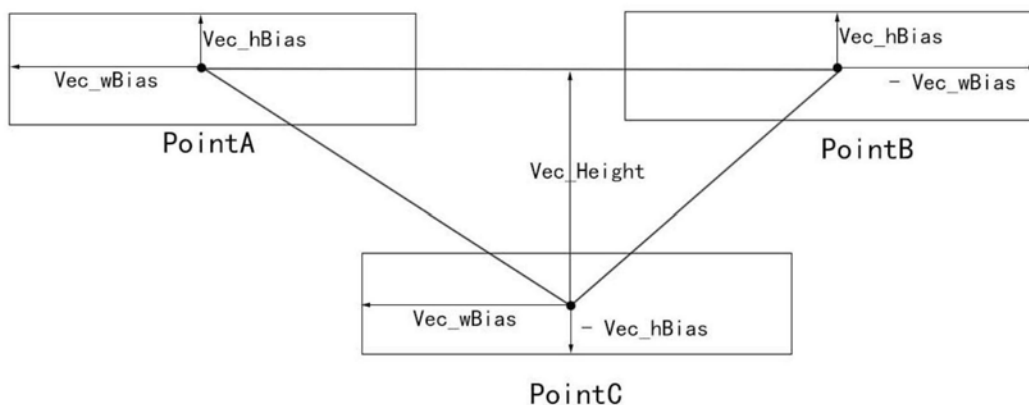


图2

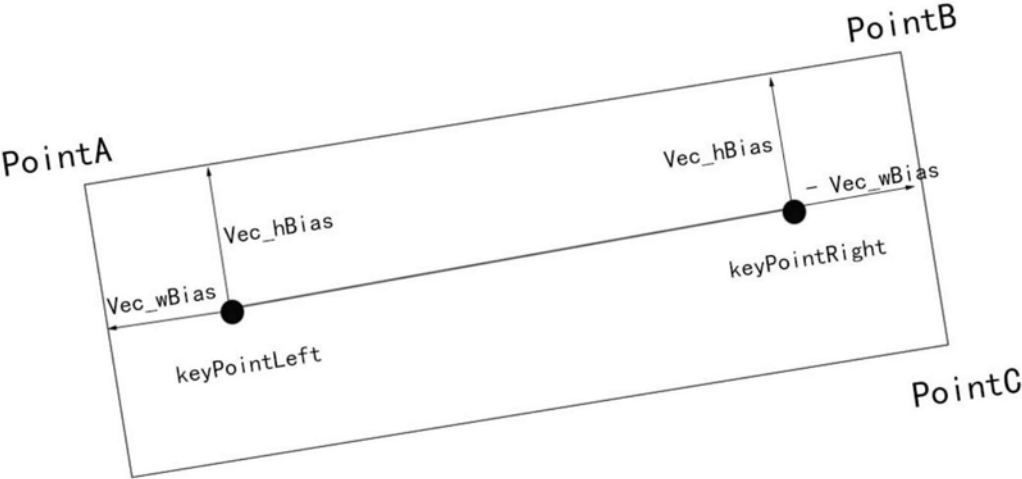


图3