



(12)发明专利申请

(10)申请公布号 CN 107025452 A

(43)申请公布日 2017. 08. 08

(21)申请号 201610065491.7

(22)申请日 2016.01.29

(71)申请人 富士通株式会社

地址 日本神奈川县

(72)发明人 许亮 范伟 孙俊 直井聪

(74)专利代理机构 北京集佳知识产权代理有限公司 11227

代理人 朱胜 李春晖

(51)Int. Cl.

G06K 9/20(2006.01)

G06K 9/62(2006.01)

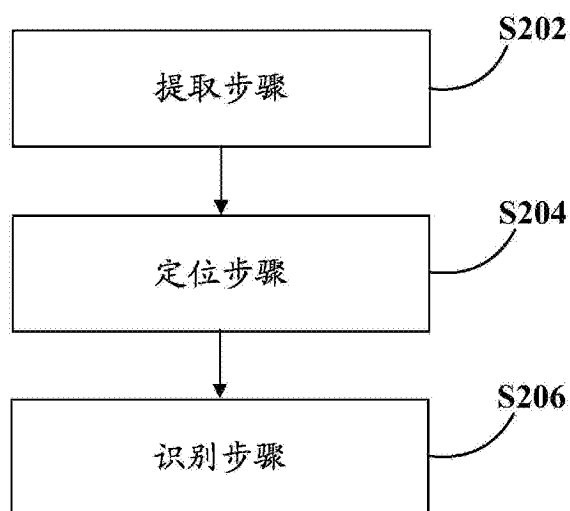
权利要求书2页 说明书12页 附图5页

(54)发明名称

图像识别方法和图像识别设备

(57)摘要

公开了一种图像识别方法和图像识别设备。一种用于对具有表格结构的输入图像进行识别的图像识别方法,其中输入图像包含有一个或多个预打印字符串以及一个或多个后打印字符串,该图像识别方法包括:提取步骤,用于从输入图像中提取出表格结构;定位步骤,用于基于表格结构的先验知识,获得一个或多个后打印字符串中的至少一个后打印字符串的位置信息;以及识别步骤,对于至少一个后打印字符串中的每个后打印字符,采用基于二值化图像的第一字符分类器和基于灰度图像的第二字符分类器来对该后打印字符进行识别。根据本公开的实施例,能够结合简单快速的字符分类器和复杂高精度的字符分类器对具有表格结构的图像进行识别。



1.一种用于对具有表格结构的输入图像进行识别的图像识别方法,其中所述输入图像包含有一个或多个预打印字符串以及一个或多个后打印字符串,所述图像识别方法包括:

提取步骤,用于从所述输入图像中提取出所述表格结构;

定位步骤,用于基于所述表格结构的先验知识,获得所述一个或多个后打印字符串中的至少一个后打印字符串的位置信息;以及

识别步骤,对于所述至少一个后打印字符串中的每个后打印字符,采用基于二值化图像的第一字符分类器和基于灰度图像的第二字符分类器来对该后打印字符进行识别。

2.根据权利要求1所述的图像识别方法,其中,在所述识别步骤中,在对输入到所述第二字符分类器的所述灰度图像进行归一化时,基于所述灰度图像的背景像素的灰度值来填充该灰度图像,以使得该灰度图像具有预定的长宽比。

3.根据权利要求2所述的图像识别方法,其中,基于所述背景像素的灰度值的均值或中值来填充该灰度图像。

4.根据权利要求1所述的图像识别方法,其中,在所述识别步骤中,所述第一字符分类器首先对所述二值化图像中的最大连通块子图像进行识别,若该识别的结果的置信度低,则所述第一字符分类器再对整个所述二值化图像进行识别,并且选择这两个识别中置信度高的识别结果作为最终识别结果。

5.根据权利要求1所述的图像识别方法,还包括分离步骤,用于基于颜色信息来从所述输入图像中分离出所述至少一个后打印字符串的图像。

6.根据权利要求5所述的图像识别方法,其中,在所述识别步骤中,对于所述至少一个后打印字符串中的每个后打印字符,采用基于所述输入图像而生成的二值化图像和灰度图像分别作为所述第一字符分类器和所述第二字符分类器的输入来对该后打印字符进行识别从而得到第一识别结果,还采用基于所述至少一个后打印字符串的图像而生成的二值化图像和灰度图像分别作为所述第一字符分类器和所述第二字符分类器的输入来对该后打印字符进行识别从而得到第二识别结果,并且选择所述第一识别结果和所述第二识别结果中置信度高的识别结果作为最终识别结果。

7.根据权利要求5所述的图像识别方法,其中,在所述识别步骤中,基于所述至少一个后打印字符串的图像来进行字符切分。

8.根据权利要求1所述的图像识别方法,其中,在所述识别步骤中,若采用所述第一字符分类器得到的识别结果的置信度低或者该识别结果的第一候选和第二候选的概率之差小,则采用所述第二字符分类器来进行识别,并且选择这两个识别中置信度高的识别结果作为最终识别结果。

9.根据权利要求1所述的图像识别方法,其中,在所述定位步骤中,基于所述表格结构中的各区域的相对位置,选择如下区域作为所述至少一个后打印字符串中的每个后打印字符串的区域:该区域包含所述表格结构中的与该后打印字符串对应的区域且比所述表格结构中的与该后打印字符串对应的区域更大。

10.一种用于对具有表格结构的输入图像进行识别的图像识别设备,其中所述输入图像包含有一个或多个预打印字符串以及一个或多个后打印字符串,所述图像识别设备包括:

提取单元,被配置成从所述输入图像中提取出所述表格结构;

定位单元,被配置成基于所述表格结构的先验知识,获得所述一个或多个后打印字符串中的至少一个后打印字符串的位置信息;以及

识别单元,被配置成对于所述至少一个后打印字符串中的每个后打印字符,采用基于二值化图像的第一字符分类器和基于灰度图像的第二字符分类器来对该后打印字符进行识别。

图像识别方法和图像识别设备

技术领域

[0001] 本公开涉及数据处理技术领域,更具体地,涉及一种能够结合简单快速的字符分类器和复杂高精度的字符分类器对具有表格结构的图像进行识别的图像识别方法和图像识别设备。

背景技术

[0002] 增值税发票(抵扣联)识别是企业的财务共享中心的一项重要内容。通过扫描增值税发票(抵扣联),得到增值税发票的数字图像,然后通过增值税发票识别系统,产生电子化的增值税发票的表格数据。

[0003] 传统的发票识别方法一般基于严格定义的发票结构模板,找到对应的条目位置,然后通过二值化提取文字,再进行字符分割和字符识别。银行流水单的识别方法与上述发票识别方法类似。传统的方法存在如下的不足:

[0004] (1)不能正确处理后打印文字和预打印文字交叠的情形。发票上的条目内容包含预打印文字和后打印文字,预打印文字的位置一般是固定的,但后打印文字的位置是不固定的,可能偏离结构模板的指定区域。图1是示出发票上后打印文字与预打印文字和表格线交叠的示例的图。如图1所示,后打印文字和预打印文字(在图1中,预打印文字为“密码区”)交叠,并且后打印文字和表格线交叠。

[0005] (2)不能正确识别后打印文字带有噪声的情形。发票上除了打印的文字,还可能会有加盖的红章、灰尘污垢等,二值化后会给后打印文字带来噪声,从而导致不能正确识别。

发明内容

[0006] 在下文中给出了关于本公开的简要概述,以便提供关于本公开的某些方面的基本理解。但是,应当理解,这个概述并不是关于本公开的穷举性概述。它并不是意图用来确定本公开的关键性部分或重要部分,也不是意图用来限定本公开的范围。其目的仅仅是以简化的形式给出关于本公开的某些概念,以此作为稍后给出的更详细描述的前序。

[0007] 鉴于以上问题,本公开的目的是提供一种图像识别方法和图像识别设备,其能够结合简单快速的字符分类器和复杂高精度的字符分类器对具有表格结构的图像进行识别,从而实现高精度的字符识别,同时识别速度较快。

[0008] 根据本公开的一方面,提供了一种用于对具有表格结构的输入图像进行识别的图像识别方法,其中输入图像包含有一个或多个预打印字符串以及一个或多个后打印字符串,该图像识别方法可以包括:提取步骤,可以用于从输入图像中提取出表格结构;定位步骤,可以用于基于表格结构的先验知识,获得一个或多个后打印字符串中的至少一个后打印字符串的位置信息;以及识别步骤,对于至少一个后打印字符串中的每个后打印字符,可以采用基于二值化图像的第一字符分类器和基于灰度图像的第二字符分类器来对该后打印字符进行识别。

[0009] 根据本公开的另一方面,还提供了一种用于对具有表格结构的输入图像进行识别

的图像识别设备,其中输入图像包含有一个或多个预打印字符串以及一个或多个后打印字符串,该图像识别设备可以包括:提取单元,可以被配置成从输入图像中提取出表格结构;定位单元,可以被配置成基于表格结构的先验知识,获得一个或多个后打印字符串中的至少一个后打印字符串的位置信息;以及识别单元,可以被配置成对于至少一个后打印字符串中的每个后打印字符,采用基于二值化图像的第一字符分类器和基于灰度图像的第二字符分类器来对该后打印字符进行识别。

[0010] 根据本公开的其它方面,还提供了用于实现上述根据本公开的方法的计算机程序代码和计算机程序产品以及其上记录有该用于实现上述根据本公开的方法的计算机程序代码的计算机可读存储介质。

[0011] 在下面的说明书部分中给出本公开实施例的其它方面,其中,详细说明用于充分地公开本公开实施例的优选实施例,而不对其施加限定。

附图说明

[0012] 本公开可以通过参考下文中结合附图所给出的详细描述而得到更好的理解,其中在所有附图中使用了相同或相似的附图标记来表示相同或者相似的部件。所述附图连同下面的详细说明一起包含在本说明书中并形成说明书的一部分,用来进一步举例说明本公开的优选实施例和解释本公开的原理和优点。其中:

[0013] 图1是示出发票上后打印文字与预打印文字和表格线交叠的示例的图;

[0014] 图2是示出根据本公开的实施例的图像识别方法的流程示例的流程图;

[0015] 图3是示出具有表格结构的输入图像的示例的图;

[0016] 图4是示出从输入图像中提取出的表格结构的示例的图;

[0017] 图5是示出从输入图像中定位出后打印字符串的示例的图;

[0018] 图6是示出包含噪声的字符串图像的示例的图;

[0019] 图7是示出包含破裂字符的字符串图像的示例的图;

[0020] 图8是示出残缺字符的示例的图;

[0021] 图9是示出CNN分类器对字符进行识别的流程的示例的图;

[0022] 图10是示出在CNN分类器中未对灰度图像进行填充和对灰度图像进行填充从而进行识别的示例的图;

[0023] 图11是示出根据本公开的实施例的图像识别设备的功能配置示例的框图;以及

[0024] 图12是示出作为本公开的实施例中可采用的信息处理设备的个人计算机的示例结构的框图。

具体实施方式

[0025] 在下文中将结合附图对本公开的示范性实施例进行描述。为了清楚和简明起见,在说明书中并未描述实际实施方式的所有特征。然而,应该了解,在开发任何这种实际实施例的过程中必须做出很多特定于实施方式的决定,以便实现开发人员的具体目标,例如,符合与系统及业务相关的那些限制条件,并且这些限制条件可能会随着实施方式的不同而有所改变。此外,还应该了解,虽然开发工作有可能是非常复杂和费时的,但对得益于本公开内容的本领域技术人员来说,这种开发工作仅仅是例行的任务。

[0026] 在此,还需要说明的一点是,为了避免因不必要的细节而模糊了本公开,在附图中仅仅示出了与根据本公开的方案密切相关的设备结构和/或处理步骤,而省略了与本公开关系不大的其它细节。

[0027] 下面结合附图详细说明根据本公开的实施例。

[0028] 首先,将参照图2描述根据本公开的实施例的图像识别方法的流程示例。图2是示出根据本公开的实施例的图像识别方法的流程示例的流程图。如图2所示,根据本公开的实施例的图像识别方法可包括提取步骤S202、定位步骤S204以及识别步骤S206。

[0029] 根据本公开实施例的图像识别方法用于对具有表格结构的输入图像进行识别,其中输入图像包含有一个或多个预打印字符串以及一个或多个后打印字符串。

[0030] 图3是示出具有表格结构的输入图像的示例的图。如图3所示,该输入图像是发票图像,该发票图像包括表格结构、印刷发票时所印刷的一个或多个预打印字符串以及在使用发票时所输入的一个或多个后打印字符串。发票图像仅是具有表格结构的图像的一个具体示例,具有表格结构的图像还可以是银行流水单等等。

[0031] 以下将结合图3中示出的图像来分别详细描述图2所示的根据本公开实施例的图像识别方法中各个步骤中的处理。

[0032] 首先,在提取步骤S202中,可以从输入图像中提取出表格结构。假设已经将发票图像进行了倾斜旋转校正,这是业界公知的技术。

[0033] 如果后打印文字和预打印文字具有不同颜色且已知,那么可以通过不同颜色投影的方法,得到包含后打印文字的灰度图像A和包含预打印文字的灰度图像B。对于灰度图像B,进行二值化以及连通部件提取,再选择其中最大的连通部件为表格。

[0034] 如果不知道颜色信息,那么可以将输入图像的灰度图像进行二值化,然后再在二值图像上寻找最长的水平游程和竖直游程,最长的水平游程和竖直游程包围的区域即是表格区域。

[0035] 图4是示出从输入图像中提取出的表格结构的示例的图。如图4所示,从图3所示的发票图像提取出其表格结构。

[0036] 在定位步骤S204中,可以基于表格结构的先验知识,获得一个或多个后打印字符串中的至少一个后打印字符串的位置信息。

[0037] 优选地,在定位步骤S204中,可以基于表格结构中的各区域的相对位置,选择如下区域作为至少一个后打印字符串中的每个后打印字符串的区域:该区域包含表格结构中的与该后打印字符串对应的区域且比表格结构中的与该后打印字符串对应的区域更大。

[0038] 在定位步骤S204中,各关键区域的相对位置可以通过关于表格结构的先验知识获得。

[0039] 具体地,基于该表格,根据各条目的弱结构模板(相对位置)信息,得到后打印文字的初步区域。这里的弱结构模板指的是目标区域不是严格地限定在表格的指定区域,而是包含比表格指定区域的更大的区域。也就是说,选择如下区域作为后打印字符串的区域:该区域包含表格结构中的与该后打印字符串对应的区域且比表格结构中的与该后打印字符串对应的区域更大。

[0040] 优选地,在定位步骤S204中,除了表格结构的先验知识之外,还可以基于颜色信息来获得至少一个后打印字符串的位置信息。

[0041] 结合后打印文字的灰度图像A,可以在条目初步区域内精细定位条目字符串。当后打印文字和预打印文字有交叠时,交叠部分的文字的颜色是介于后打印文字颜色和预打印文字颜色两者之间,因此采用该颜色信息,可以将交叠部分的后打印文字定位出来。

[0042] 图5是示出从输入图像中定位出后打印字符串的示例的图。如图5所示,用矩形所包围的区域为从图3所示的发票图像中定位出的感兴趣的后打印字符串。

[0043] 在识别步骤S206中,对于至少一个后打印字符串中的每个后打印字符,可以采用基于二值化图像的第一字符分类器和基于灰度图像的第二字符分类器来对该后打印字符进行识别。

[0044] 在识别步骤S206中,结合简单快速的第一字符分类器和复杂高精度的第二字符分类器对后打印字符进行识别。第一字符分类器可以是基于二值化图像的修正二次判别函数(MQDF)分类器。第二字符分类器可以是基于灰度图像的卷积神经网络(CNN)分类器。上述第一字符分类器和第二字符分类器仅是示例而非限制,也即第一字符分类器和第二字符分类器还可以是除了MQDF分类器和CNN分类器之外的其他字符分类器。在下文中,为了清楚,以第一字符分类器是MQDF分类器和第二字符分类器是CNN分类器来进行描述。

[0045] 具体地,本申请在传统的基于二值化图像的MQDF分类器的基础上,采用一种结合机制,还利用了深度学习的CNN分类器,该CNN分类器直接基于灰度图像提取特征来进行分类识别。这种结合既能够保证较高的字符识别正确率,又能够保证较快速的处理速度。

[0046] 下面简单介绍简单快速的第一字符分类器:二次判别函数(QDF)分类器。

[0047] 对于特征向量 x ,多类别二次函数分类器的输出如下:

[0048]
$$f_i(x) = (x - \mu_i)^T \Sigma_i^{-1} (x - \mu_i),$$

[0049] 其中, μ_i 、 Σ_i 分别是第 i ($i=1,2,\dots,N$)类别的均值向量和协方差矩阵。对于每个类别,在高斯假设的前提下,从训练样本估计均值和协方差。对于输入样本,将标签分配给具有最小距离的类别。然后,我们将分类器输出通过Sigmoid变换转换为置信度概率。

[0050] 具体地,MQDF是修正的QDF。

[0051] 优选地,在识别步骤S206中,第一字符分类器可以首先对二值化图像中的最大连通块子图像进行识别,若该识别的结果的置信度低,则第一字符分类器可以再对整个二值化图像进行识别,并且选择这两个识别中置信度高的识别结果作为最终识别结果。

[0052] 具体地,在识别步骤S206中,对于要识别的字符的位置上对应的输入图像的灰度图像部分进行二值化,为了忽略噪声的影响,先提取出二值化图像中最大的连通块子图像来采用MQDF分类器进行字符识别。图6是示出包含噪声的字符串图像的示例的图。如图6所示,其中的小数点“.”上面存在噪声,在对该小数点“.”识别时,可以先提取出其二值化图像中最大的连通块子图像来进行字符识别。如果该识别的结果的置信度较低,则采用MQDF分类器对整个二值化图像进行字符识别,最后选择这两个识别中置信度高的识别结果作为最终识别结果。

[0053] 图7是示出包含破裂字符的字符串图像的示例的图。如图7所示,其中的数字“9”、“0”、“5”和“7”是破裂字符。图8是示出残缺字符的示例的图。图8中第一行的图像是原始图像。图8中第二行的图像是去除上述原始图像中的盖章后的灰度图像。图8中第三行的图像是所提取的字符串灰度图像。图8中第四行的图像是二值化后的字符串图像,其中“6”和“4”

是残缺字符。

[0054] 然而,对于如图6所示的带噪声字符、如图7所示的破裂字符以及如图8所示的残缺字符,MQDF分类器的识别精度较低。

[0055] 由以上描述可知,第一字符分类器的优点是:快速,对于噪声较少的情形,识别精度较高。其缺点是:一般处理二值图像,图像特征是手工设计提取的;容易受噪声的影响,对于带噪声字符、破裂字符以及残缺字符,识别精度较低。

[0056] 下面简单介绍复杂高精度的第二字符分类器:CNN分类器。

[0057] 图9是示出CNN分类器对字符进行识别的流程的示例的图。

[0058] 如图9所示,首先将灰度图像输入到CNN分类器。然后,对所输入的灰度图像进行缩放处理,即对所输入的灰度图像的长度和宽度均按一定比例进行缩放。接下来,对经缩放的图像进行填充,以使得该灰度图像具有CNN分类器所要求的归一化的长宽比。最后进行特征提取和分类,从而得到识别出的字符。

[0059] 针对CNN分类器的输入样本需要保持长宽比值的归一化,本申请提出了一种自适应地估计输入样本的背景像素值来得到更加符合实际情形的字符图像的方法。

[0060] 优选地,在识别步骤S206中,在对输入到第二字符分类器的灰度图像进行归一化时,可以基于灰度图像的背景像素的灰度值来填充该灰度图像,以使得该灰度图像具有预定的长宽比。具体地,在对输入到CNN分类器的灰度图像进行归一化时,可以对该灰度图像进行二值化,得到前景像素区域,余下的为背景像素。若经缩放的灰度图像的长宽比不是CNN分类器所要求的预定的长宽比,则基于背景像素的灰度值来填充经缩放的灰度图像,以使得灰度图像具有上述预定的长宽比。

[0061] 优选地,可以基于背景像素的灰度值的均值或中值来填充该灰度图像。具体地,可以使用背景像素的灰度值的均值或者中值作为填充像素值。

[0062] 图10是示出在CNN分类器中未对灰度图像进行填充和对灰度图像进行填充从而进行识别的示例的图。具体地,如图10的第一行所示,对于所输入的灰度图像“2”,若未对缩放后的灰度图像进行填充,则会将其识别为数字“1”。而如图10的第二行所示,对于所输入的灰度图像“2”,若基于灰度图像的背景像素的灰度值对缩放后的灰度图像进行填充,则会将其识别为数字“2”。

[0063] 优选地,CNN分类器所要求的预定的长宽比可以根据实验确定,也可以根据本领域中公知的其他方法来确定。

[0064] 根据以上示例可以看出,由以上基于灰度图像的背景像素的灰度值来填充灰度图像的方法得到的字符图像更加符合实际情形,从而提升了CNN分类器的识别性能。

[0065] 第二字符分类器的优点是:直接基于灰度图像进行识别,图像特征是通过训练样本自动学习得到的;对于带噪声字符、破裂字符以及残缺字符,识别精度依然较高。其缺点是处理速度较慢。

[0066] 优选地,在识别步骤S206中,若采用第一字符分类器得到的识别结果的置信度低或者该识别结果的第一候选和第二候选的概率之差小,则可以采用第二字符分类器来进行识别,并且可以选择这两个识别中置信度高的识别结果作为最终识别结果。具体地,在识别步骤S206中,当采用第一字符分类器得到的识别结果的置信度较低或者该识别结果的第一候选和第二候选的概率之差较小时,表示该第一分类器的输出结果不太可靠。此时,进一步

采用基于灰度字符图像的复杂高精度的第二字符分类器来进行识别。最后,选择这两个识别中置信度高的识别结果作为最终识别结果。举例来说,对于无噪声的字符图像,可以直接采用基于二值化图像的MQDF分类器的识别结果。而对于包含噪声的字符图像,先得到二值化图像的MQDF识别结果,但是MQDF分类器的识别结果的置信度可能较低。那么进行灰度图像的CNN识别,然后比较两者的识别置信度,最后采用较高置信度对应的识别结果。

[0067] 优选地,根据本公开的实施例的图像识别方法还包括可以分离步骤,其可以基于颜色信息来从输入图像中分离出至少一个后打印字符串的图像。具体地,由于后打印文字和预打印文字具有不同颜色,因此可以基于颜色信息来从输入图像中分离出后打印字符串的图像。

[0068] 优选地,在识别步骤S206中,对于至少一个后打印字符串中的每个后打印字符,可以采用基于输入图像而生成的二值化图像和灰度图像分别作为第一字符分类器和第二字符分类器的输入来对该后打印字符进行识别从而得到第一识别结果,还可以采用基于至少一个后打印字符串的图像而生成的二值化图像和灰度图像分别作为第一字符分类器和第二字符分类器的输入来对该后打印字符进行识别从而得到第二识别结果,并且可以选择第一识别结果和第二识别结果中置信度高的识别结果作为最终识别结果。

[0069] 具体地,由于在利用颜色信息在分离步骤中分离后打印字符串的图像的过程中,去除了诸如红色盖章等的噪声从而使得字符更加完整,因此相比于原始输入图像而言,在分离步骤中分离出的后打印字符串的图像可能噪声更少并且可能字符更完整。从而,在识别步骤S206中,为了得到更准确的识别结果,除了将基于原始输入图像而生成的二值化图像和灰度图像分别作为第一字符分类器和第二字符分类器的输入来对后打印字符进行识别外,还可以将基于分离步骤中分离出的后打印字符串的图像而生成的二值化图像和灰度图像分别作为第一字符分类器和第二字符分类器的输入来对后打印字符进行识别,最后,识别结果中置信度高的识别结果作为最终识别结果。

[0070] 优选地,在识别步骤S206中,可以基于至少一个后打印字符串的图像来进行字符切分。为了避免预打印字符的交叠影响,本申请在分离步骤中分离出的后打印字符串的图像的灰度图像上,进行字符切分,这样可以使得字符切分的精确度更高。

[0071] 根据以上描述可知,根据本公开的实施例的图像识别方法结合简单快速的字符分类器和复杂高精度的字符分类器对具有表格结构的图像进行识别,因此可以实现高精度的字符识别,同时识别速度较快。

[0072] 与上述方法实施例相对应地,本公开还提供了以下设备实施例。

[0073] 图11是示出根据本公开的实施例的图像识别设备1100的功能配置示例的框图。

[0074] 如图11所示,根据本公开的实施例的图像识别设备1100可以包括提取单元1102、定位单元1104以及识别单元1106。接下来将描述各个单元的功能配置示例。

[0075] 根据本公开实施例的图像识别设备用于对具有表格结构的输入图像进行识别,其中输入图像包含有一个或多个预打印字符串以及一个或多个后打印字符串。

[0076] 具有表格结构的输入图像可参见以上方法实施例中相应位置的描述,在此不再重复。

[0077] 提取单元1102可以被配置成从输入图像中提取出表格结构。

[0078] 从输入图像中提取出表格结构的具体方法可参见以上方法实施例中相应位置的

描述,在此不再重复。

[0079] 定位单元1104可以被配置成基于表格结构的先验知识,获得一个或多个后打印字符串中的至少一个后打印字符串的位置信息。

[0080] 优选地,在定位单元1104中,可以基于表格结构中的各区域的相对位置,选择如下区域作为至少一个后打印字符串中的每个后打印字符串的区域:该区域包含表格结构中的与该后打印字符串对应的区域且比表格结构中的与该后打印字符串对应的区域更大。

[0081] 在定位单元1104中,各关键区域的相对位置可以通过关于表格结构的先验知识获得。

[0082] 具体地,基于该表格,根据各条目的弱结构模板(相对位置)信息,得到后打印文字的初步区域。这里的弱结构模板指的是目标区域不是严格地限定在表格的指定区域,而是包含比表格指定区域的更大的区域。也就是说,选择如下区域作为后打印字符串的区域:该区域包含表格结构中的与该后打印字符串对应的区域且比表格结构中的与该后打印字符串对应的区域更大。

[0083] 优选地,在定位单元1104中,除了表格结构的先验知识之外,还可以基于颜色信息来获得至少一个后打印字符串的位置信息。

[0084] 由于后打印文字和预打印文字具有不同颜色,那么可以通过不同颜色投影的方法,得到包含后打印文字的灰度图像A和包含预打印文字的灰度图像B。结合后打印文字的灰度图像A,在条目初步区域内精细定位条目字符串。当后打印文字和预打印文字有交叠时,交叠部分的文字的颜色是介于后打印文字颜色和预打印文字颜色两者之间,因此采用该颜色信息,可以将交叠部分的后打印文字定位出来。

[0085] 识别单元1106可以被配置成对于至少一个后打印字符串中的每个后打印字符,采用基于二值化图像的第一字符分类器和基于灰度图像的第二字符分类器来对该后打印字符进行识别。

[0086] 在识别单元1106中,结合简单快速的第一字符分类器和复杂高精度的第二字符分类器对后打印字符进行识别。第一字符分类器可以是基于二值化图像的MQDF分类器。第二字符分类器可以是基于灰度图像的CNN分类器。

[0087] 关于第一分类器的详细描述可参见以上方法实施例中相应位置的描述,在此不再重复。

[0088] 优选地,在识别单元1106中,第一字符分类器可以首先对二值化图像中的最大连通块子图像进行识别,若该识别的结果的置信度低,则第一字符分类器可以再对整个二值化图像进行识别,并且选择这两个识别中置信度高的识别结果作为最终识别结果。

[0089] 利用MQDF分类器对字符进行识别的具体示例可参见以上方法实施例中相应位置的描述,在此不再重复。

[0090] 关于CNN分类器的详细描述可参见以上方法实施例中相应位置的描述,在此不再重复。

[0091] 针对CNN分类器的输入样本需要保持长宽比值的归一化,提出了一种自适应地估计输入样本的背景像素值来得到更加符合实际情形的字符图像的方法。

[0092] 优选地,在识别单元1106中,在对输入到第二字符分类器的灰度图像进行归一化时,可以基于灰度图像的背景像素的灰度值来填充该灰度图像,以使得该灰度图像具有预

定的长宽比。

[0093] 优选地,可以基于背景像素的灰度值的均值或中值来填充该灰度图像。

[0094] 优选地,CNN分类器所要求的预定的长宽比可以根据实验确定,也可以根据本领域中公知的其他方法来确定。

[0095] 在CNN分类器中基于灰度图像的背景像素的灰度值来填充灰度图像的具体示例可参见以上方法实施例中相应位置的描述,在此不再重复。

[0096] 优选地,在识别单元1106中,若采用第一字符分类器得到的识别结果的置信度低或者该识别结果的第一候选和第二候选的概率之差小,则可以采用第二字符分类器来进行识别,并且可以选择这两个识别中置信度高的识别结果作为最终识别结果。具体地,在识别单元110中,当采用第一字符分类器得到的识别结果的置信度较低或者该识别结果的第一候选和第二候选的概率之差较小时,表示该第一分类器的输出结果不太可靠。此时,进一步采用基于灰度字符图像的复杂高精度的第二字符分类器来进行识别。最后,选择这两个识别中置信度高的识别结果作为最终识别结果。

[0097] 优选地,根据本公开的实施例的图像识别设备还包括可以分离单元,其可以被配置成基于颜色信息来从输入图像中分离出至少一个后打印字符串的图像。具体地,由于后打印文字和预打印文字具有不同颜色,因此可以基于颜色信息来从输入图像中分离出后打印字符串的图像。

[0098] 优选地,在识别单元1106中,对于至少一个后打印字符串中的每个后打印字符,可以采用基于输入图像而生成的二值化图像和灰度图像分别作为第一字符分类器和第二字符分类器的输入来对该后打印字符进行识别从而得到第一识别结果,还可以采用基于至少一个后打印字符串的图像而生成的二值化图像和灰度图像分别作为第一字符分类器和第二字符分类器的输入来对该后打印字符进行识别从而得到第二识别结果,并且可以选择第一识别结果和第二识别结果中置信度高的识别结果作为最终识别结果。

[0099] 具体地,由于在利用颜色信息在分离单元中分离后打印字符串的图像的过程中,去除了诸如红色盖章等的噪声从而使得字符更加完整,因此相比于原始输入图像而言,在分离单元中分离出的后打印字符串的图像可能噪声更少并且可能字符更完整。从而,在识别单元1106中,为了得到更准确的识别结果,除了将基于输入图像而生成的二值化图像和灰度图像分别作为第一字符分类器和第二字符分类器的输入来对后打印字符进行识别外,还可以将基于分离单元中分离出的后打印字符串的图像而生成的二值化图像和灰度图像分别作为第一字符分类器和第二字符分类器的输入来对后打印字符进行识别,最后,识别结果中置信度高的识别结果作为最终识别结果。

[0100] 优选地,在识别单元1106中,可以基于至少一个后打印字符串的图像来进行字符切分。为了避免预打印字符的交叠影响,本申请在分离单元中分离出的后打印字符串的图像的灰度图像上,进行字符切分,这样可以使得字符切分的精确度更高。

[0101] 根据以上描述可知,根据本公开的实施例的图像识别设备结合简单快速的字符分类器和复杂高精度的字符分类器对具有表格结构的图像进行识别,因此可以实现高精度的字符识别,同时识别速度较快。

[0102] 应指出,尽管以上描述了根据本公开的实施例的图像识别设备的功能配置,但是这仅是示例而非限制,并且本领域技术人员可根据本公开的原理对以上实施例进行修改,

例如可对各个实施例中的功能模块进行添加、删除或者组合等,并且这样的修改均落入本公开的范围。

[0103] 此外,还应指出,这里的装置实施例是与上述方法实施例相对应的,因此在装置实施例中未详细描述的内容可参见方法实施例中相应位置的描述,在此不再重复描述。

[0104] 应理解,根据本公开的实施例的存储介质和程序产品中的机器可执行的指令还可以被配置成执行上述图像识别方法,因此在此未详细描述的内容可参考先前相应位置的描述,在此不再重复进行描述。

[0105] 相应地,用于承载上述包括机器可执行的指令的程序产品的存储介质也包括在本发明的公开中。该存储介质包括但不限于软盘、光盘、磁光盘、存储卡、存储棒等等。

[0106] 另外,还应该指出的是,上述系列处理和装置也可以通过软件和/或固件实现。在通过软件和/或固件实现的情况下,从存储介质或网络向具有专用硬件结构的计算机,例如图12所示的通用个人计算机1200安装构成该软件的程序,该计算机在安装各种程序时,能够执行各种功能等等。

[0107] 在图12中,中央处理单元(CPU)1201根据只读存储器(ROM)1202中存储的程序或从存储部分1208加载到随机存取存储器(RAM)1203的程序执行各种处理。在RAM 1203中,也根据需要存储当CPU 1201执行各种处理等时所需的数据。

[0108] CPU 1201、ROM 1202和RAM 1203经由总线1204彼此连接。输入/输出接口1205也连接到总线1204。

[0109] 下述部件连接到输入/输出接口1205:输入部分1206,包括键盘、鼠标等;输出部分1207,包括显示器,比如阴极射线管(CRT)、液晶显示器(LCD)等,和扬声器等;存储部分1208,包括硬盘等;和通信部分1209,包括网络接口卡比如LAN卡、调制解调器等。通信部分1209经由网络比如因特网执行通信处理。

[0110] 根据需要,驱动器1210也连接到输入/输出接口1205。可拆卸介质1211比如磁盘、光盘、磁光盘、半导体存储器等等根据需要进行安装在驱动器1210上,使得从中读出的计算机程序根据需要进行安装到存储部分1208中。

[0111] 在通过软件实现上述系列处理的情况下,从网络比如因特网或存储介质比如可拆卸介质1211安装构成软件的程序。

[0112] 本领域的技术人员应当理解,这种存储介质不局限于图12所示的其中存储有程序、与设备相分离地分发以向用户提供程序的可拆卸介质1211。可拆卸介质1211的例子包含磁盘(包含软盘(注册商标))、光盘(包含光盘只读存储器(CD-ROM)和数字通用盘(DVD))、磁光盘(包含迷你盘(MD)(注册商标))和半导体存储器。或者,存储介质可以是ROM 1202、存储部分1208中包含的硬盘等等,其中存有程序,并且与包含它们的设备一起被分发给用户。

[0113] 以上参照附图描述了本公开的优选实施例,但是本公开当然不限于以上示例。本领域技术人员可在所附权利要求的范围内得到各种变更和修改,并且应理解这些变更和修改自然将落入本公开的技术范围内。

[0114] 例如,在以上实施例中包括在一个单元中的多个功能可以由分开的装置来实现。备选地,在以上实施例中由多个单元实现的多个功能可分别由分开的装置来实现。另外,以上功能之一可由多个单元来实现。无需说,这样的配置包括在本公开的技术范围内。

[0115] 在该说明书中,流程图中所描述的步骤不仅包括以所述顺序按时间序列执行的处

理,而且包括并行地或单独地而不是必须按时间序列执行的处理。此外,甚至在按时间序列处理的步骤中,无需说,也可以适当地改变该顺序。

[0116] 另外,根据本公开的技术还可以如下进行配置。

[0117] 附记1.一种用于对具有表格结构的输入图像进行识别的图像识别方法,其中所述输入图像包含有一个或多个预打印字符串以及一个或多个后打印字符串,所述图像识别方法包括:

[0118] 提取步骤,用于从所述输入图像中提取出所述表格结构;

[0119] 定位步骤,用于基于所述表格结构的先验知识,获得所述一个或多个后打印字符串中的至少一个后打印字符串的位置信息;以及

[0120] 识别步骤,对于所述至少一个后打印字符串中的每个后打印字符,采用基于二值化图像的第一字符分类器和基于灰度图像的第二字符分类器来对该后打印字符进行识别。

[0121] 附记2.根据附记1所述的图像识别方法,其中,在所述识别步骤中,在对输入到所述第二字符分类器的所述灰度图像进行归一化时,基于所述灰度图像的背景像素的灰度值来填充该灰度图像,以使得该灰度图像具有预定的长宽比。

[0122] 附记3.根据附记2所述的图像识别方法,其中,基于所述背景像素的灰度值的均值或中值来填充该灰度图像。

[0123] 附记4.根据附记1所述的图像识别方法,其中,在所述识别步骤中,所述第一字符分类器首先对所述二值化图像中的最大连通块子图像进行识别,若该识别的结果的置信度低,则所述第一字符分类器再对整个所述二值化图像进行识别,并且选择这两个识别中置信度高的识别结果作为最终识别结果。

[0124] 附记5.根据附记1所述的图像识别方法,还包括分离步骤,用于基于颜色信息来从所述输入图像中分离出所述至少一个后打印字符串的图像。

[0125] 附记6.根据附记5所述的图像识别方法,其中,在所述识别步骤中,对于所述至少一个后打印字符串中的每个后打印字符,采用基于所述输入图像而生成的二值化图像和灰度图像分别作为所述第一字符分类器和所述第二字符分类器的输入来对该后打印字符进行识别从而得到第一识别结果,还采用基于所述至少一个后打印字符串的图像而生成的二值化图像和灰度图像分别作为所述第一字符分类器和所述第二字符分类器的输入来对该后打印字符进行识别从而得到第二识别结果,并且选择所述第一识别结果和所述第二识别结果中置信度高的识别结果作为最终识别结果。

[0126] 附记7.根据附记5所述的图像识别方法,其中,在所述识别步骤中,基于所述至少一个后打印字符串的图像来进行字符切分。

[0127] 附记8.根据附记1所述的图像识别方法,其中,在所述识别步骤中,若采用所述第一字符分类器得到的识别结果的置信度低或者该识别结果的第一候选和第二候选的概率之差小,则采用所述第二字符分类器来进行识别,并且选择这两个识别中置信度高的识别结果作为最终识别结果。

[0128] 附记9.根据附记1所述的图像识别方法,其中,在所述定位步骤中,基于所述表格结构中的各区域的相对位置,选择如下区域作为所述至少一个后打印字符串中的每个后打印字符串的区域:该区域包含所述表格结构中的与该后打印字符串对应的区域且比所述表格结构中的与该后打印字符串对应的区域更大。

[0129] 附记10.根据附记1所述的图像识别方法,其中,在所述定位步骤中,除了所述表格结构的先验知识之外,还基于颜色信息来获得所述至少一个后打印字符串的位置信息。

[0130] 附记11.一种用于对具有表格结构的输入图像进行识别的图像识别设备,其中所述输入图像包含有一个或多个预打印字符串以及一个或多个后打印字符串,所述图像识别设备包括:

[0131] 提取单元,被配置成从所述输入图像中提取出所述表格结构;

[0132] 定位单元,被配置成基于所述表格结构的先验知识,获得所述一个或多个后打印字符串中的至少一个后打印字符串的位置信息;以及

[0133] 识别单元,被配置成对于所述至少一个后打印字符串中的每个后打印字符,采用基于二值化图像的第一字符分类器和基于灰度图像的第二字符分类器来对该后打印字符进行识别。

[0134] 附记12.根据附记11所述的图像识别设备,其中,在所述识别单元中,在对输入到所述第二字符分类器的所述灰度图像进行归一化时,基于所述灰度图像的背景像素的灰度值来填充该灰度图像,以使得该灰度图像具有预定的长宽比。

[0135] 附记13.根据附记12所述的图像识别设备,其中,基于所述背景像素的灰度值的均值或中值来填充该灰度图像。

[0136] 附记14.根据附记11所述的图像识别设备,其中,在所述识别单元中,所述第一字符分类器首先对所述二值化图像中的最大连通块子图像进行识别,若该识别的结果的置信度低,则所述第一字符分类器再对整个所述二值化图像进行识别,并且选择这两个识别中置信度高的识别结果作为最终识别结果。

[0137] 附记15.根据附记11所述的图像识别设备,还包括分离单元,被配置成基于颜色信息来从所述输入图像中分离出所述至少一个后打印字符串的图像。

[0138] 附记16.根据附记15所述的图像识别设备,其中,在所述识别单元中,对于所述至少一个后打印字符串中的每个后打印字符,采用基于所述输入图像而生成的二值化图像和灰度图像分别作为所述第一字符分类器和所述第二字符分类器的输入来对该后打印字符进行识别从而得到第一识别结果,还采用基于所述至少一个后打印字符串的图像而生成的二值化图像和灰度图像分别作为所述第一字符分类器和所述第二字符分类器的输入来对该后打印字符进行识别从而得到第二识别结果,并且选择所述第一识别结果和所述第二识别结果中置信度高的识别结果作为最终识别结果。

[0139] 附记17.根据附记15所述的图像识别设备,其中,在所述识别单元中,基于所述至少一个后打印字符串的图像来进行字符切分。

[0140] 附记18.根据附记11所述的图像识别设备,其中,在所述识别单元中,若采用所述第一字符分类器得到的识别结果的置信度低或者该识别结果的第一候选和第二候选的概率之差小,则采用所述第二字符分类器来进行识别,并且选择这两个识别中置信度高的识别结果作为最终识别结果。

[0141] 附记19.根据附记11所述的图像识别设备,其中,在所述定位单元中,基于所述表格结构中的各区域的相对位置,选择如下区域作为所述至少一个后打印字符串中的每个后打印字符串的区域:该区域包含所述表格结构中的与该后打印字符串对应的区域且比所述表格结构中的与该后打印字符串对应的区域更大。

[0142] 附记20.根据附记11所述的图像识别设备,其中,在所述定位单元中,除了所述表格结构的先验知识之外,还基于颜色信息来获得所述至少一个后打印字符串的位置信息。

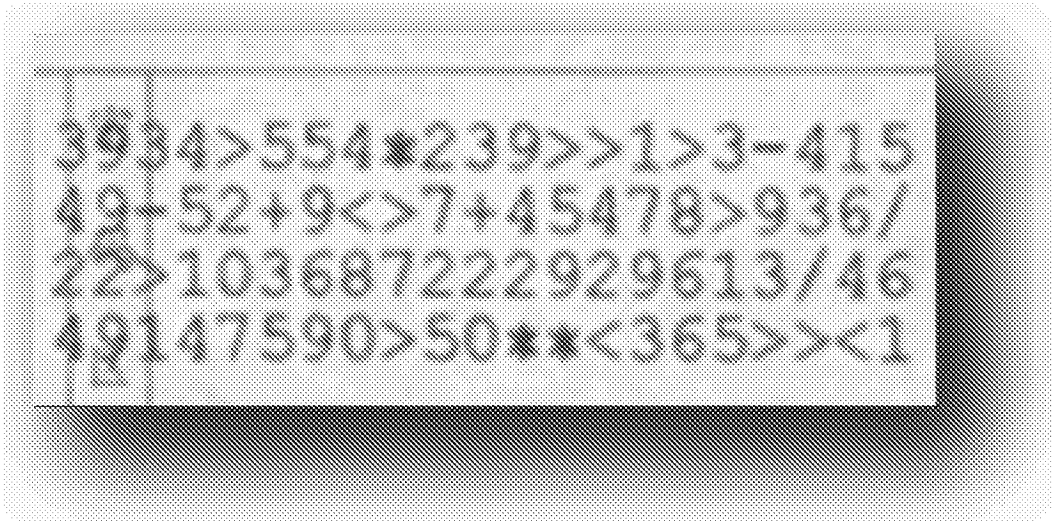


图1

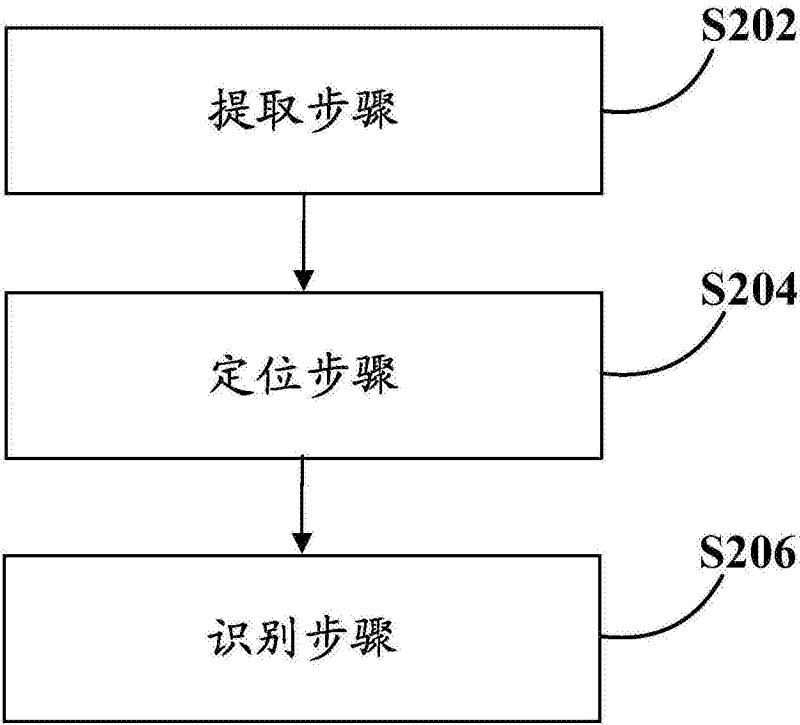


图2



图3

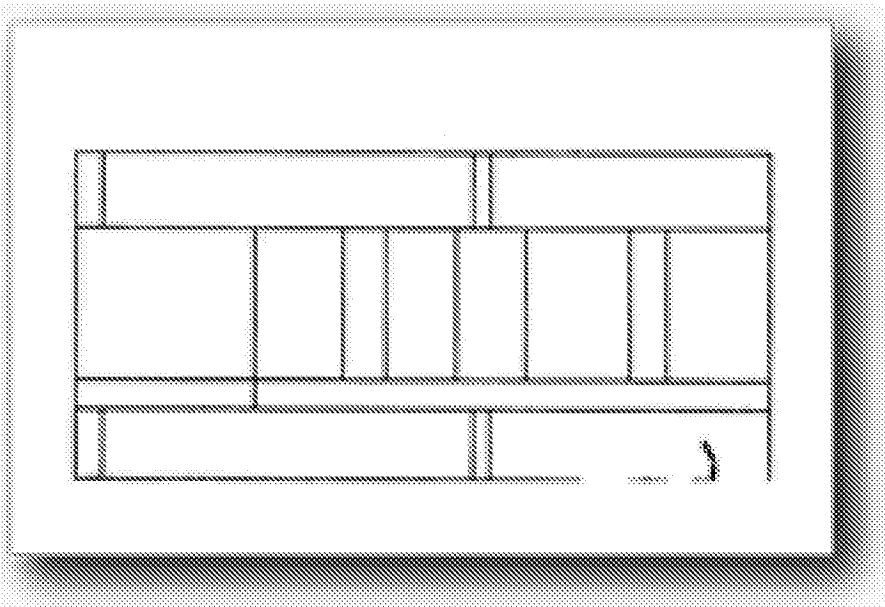


图4

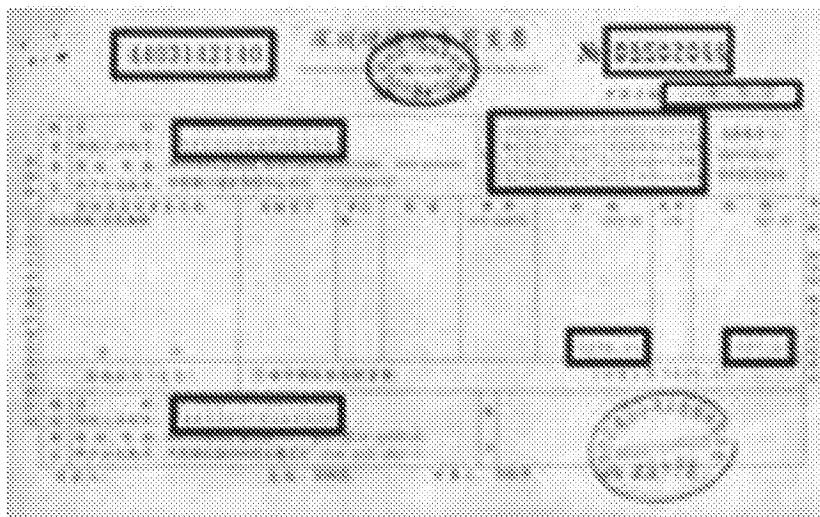


图5



图6



图7

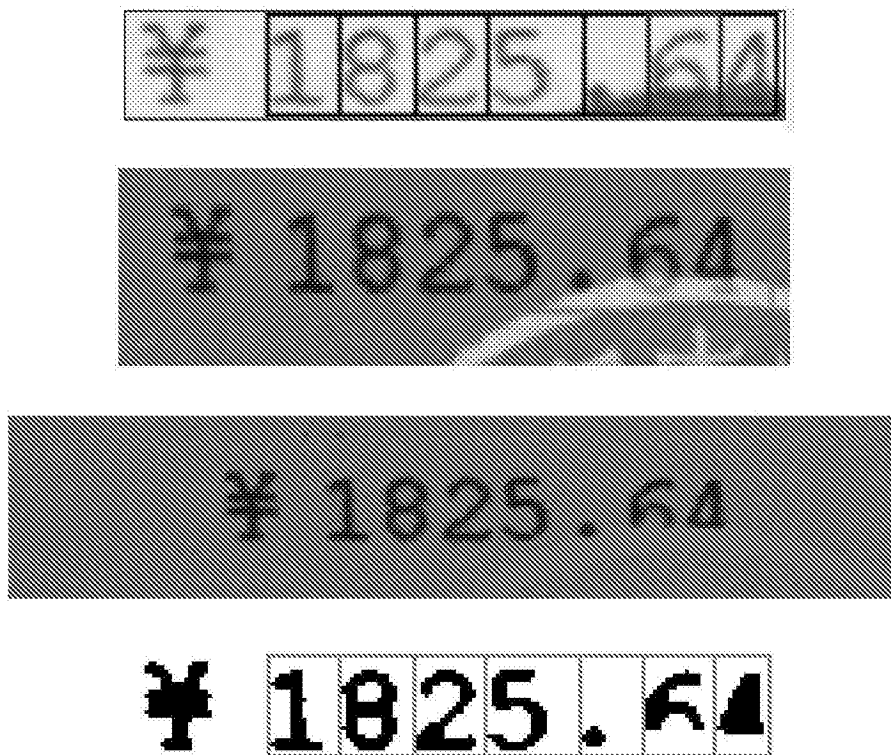


图8



图9

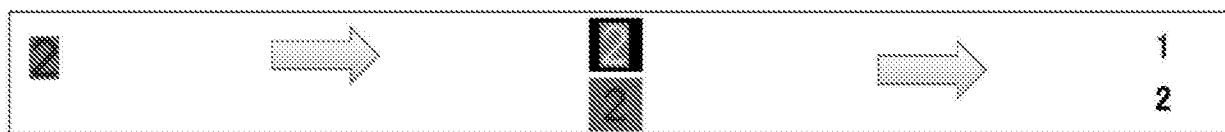


图10

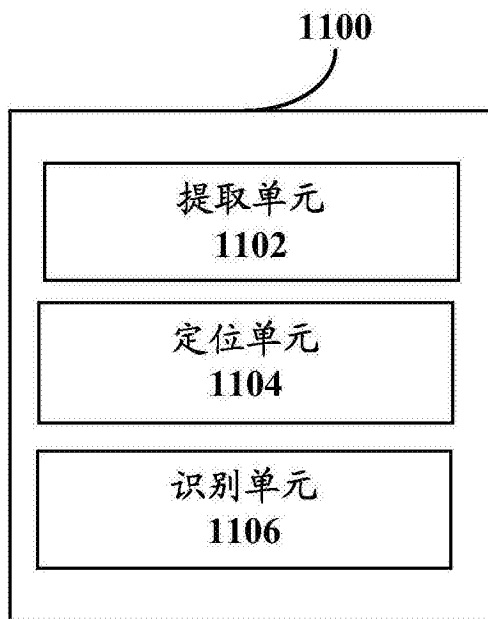


图11

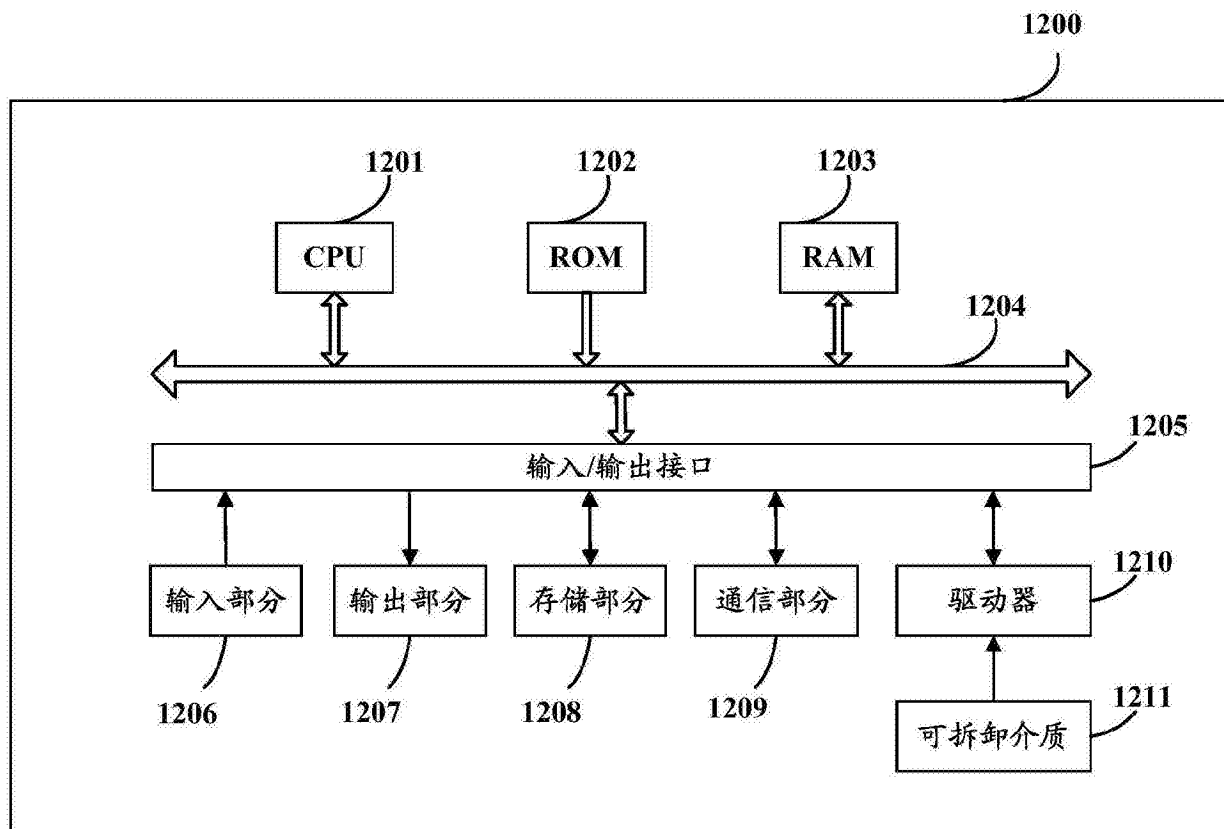


图12