



(12)发明专利申请

(10)申请公布号 CN 108885699 A

(43)申请公布日 2018. 11. 23

(21)申请号 201880001125.2

(51)Int.Cl.

(22)申请日 2018.07.11

G06K 9/32(2006.01)

G06K 9/62(2006.01)

(85)PCT国际申请进入国家阶段日
2018.08.24

(86)PCT国际申请的申请数据

PCT/CN2018/095295 2018.07.11

(71)申请人 深圳前海达闼云端智能科技有限公司

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 梁昊 南一冰 廉士国

(74)专利代理机构 北京英创嘉友知识产权代理
事务所(普通合伙) 11447

代理人 魏嘉熹 南毅宁

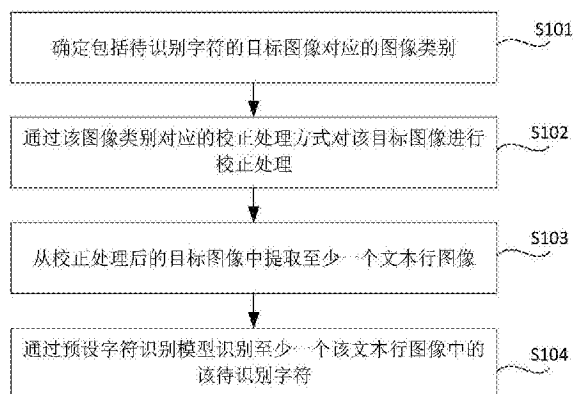
权利要求书3页 说明书9页 附图3页

(54)发明名称

字符识别方法、装置、存储介质及电子设备

(57)摘要

本公开涉及一种字符识别方法、装置、存储介质及电子设备,所述方法包括:首先,可以确定包括待识别字符的目标图像对应的图像类别;接着,通过所述图像类别对应的校正处理方式对所述目标图像进行校正处理;然后,从校正处理后的目标图像中提取至少一个文本行图像;最后,通过预设字符识别模型识别至少一个所述文本行图像中的所述待识别字符。由于不同的图像类别对应不同的校正处理方式,这样,可以将不同图像类别的图像按照对应的校正处理方式进行校正处理,并对校正处理后的图像进行字符识别,本公开能够满足对文本图像和场景图像进行字符识别,从而避免了现有技术中字符识别算法的通用性较差的问题。



1. 一种字符识别方法,其特征在于,所述方法包括:
 - 确定包括待识别字符的目标图像对应的图像类别;其中,不同的图像类别对应不同的校正处理方式;
 - 通过所述图像类别对应的校正处理方式对所述目标图像进行校正处理;
 - 从校正处理后的目标图像中提取至少一个文本行图像;
 - 通过预设字符识别模型识别至少一个所述文本行图像中的所述待识别字符。
2. 根据权利要求1所述的方法,其特征在于,所述图像类别包括文档图像和场景图像。
3. 根据权利要求1或2所述的方法,其特征在于,所述确定包括待识别字符的目标图像对应的图像类别包括:
 - 获取已确定图像类别的图像样本;
 - 根据所述图像样本确定所述目标图像对应的图像类别。
4. 根据权利要求2所述的方法,其特征在于,在所述图像类别为文档图像时,所述校正处理方式包括方向校正处理和/或畸变校正处理;在所述校正处理方式包括所述方向校正处理和所述畸变校正处理时,所述通过所述图像类别对应的校正处理方式对所述目标图像进行校正处理包括:
 - 获取所述文档图像中的所述待识别字符与水平轴之间的第一倾斜角度;
 - 在所述第一倾斜角度大于或者等于预设角度时,对所述文档图像进行方向校正处理;
 - 确定所述文档图像中的所述待识别字符是否存在畸变;
 - 在所述文档图像中的所述待识别字符存在畸变时,对所述文档图像进行畸变校正处理。
5. 根据权利要求2所述的方法,其特征在于,在所述图像类别为场景图像时,所述校正处理方式包括方向校正处理;所述通过所述图像类别对应的校正处理方式对所述目标图像进行校正处理包括:
 - 对所述场景图像进行文字区域检测得到至少一个文字区域;
 - 依次获取至少一个所述文字区域中的所述待识别字符与水平轴之间的第二倾斜角度;
 - 在至少一个所述文字区域中的所述第二倾斜角度大于或者等于预设角度时,对至少一个所述文字区域进行方向校正处理。
6. 根据权利要求1或2所述的方法,其特征在于,在所述通过预设字符识别模型识别至少一个所述文本行图像中的所述待识别字符前,还包括:
 - 获取至少一个所述文本行图像的位置信息;
 - 所述通过预设字符识别模型识别至少一个所述文本行中的所述待识别字符包括:
 - 通过所述预设字符识别模型和所述位置信息识别至少一个所述文本行图像中的所述待识别字符。
7. 根据权利要求6所述的方法,其特征在于,所述预设字符识别模型包括深度学习层、循环网络层以及编码层,所述通过所述预设字符识别模型和所述位置信息识别至少一个所述文本行图像中的所述待识别字符包括:
 - 根据所述深度学习层对至少一个所述文本行图像进行字符特征提取;
 - 将提取的字符特征输入至所述循环网络层得到至少一个所述文本行图像对应的特征向量;

将所述特征向量输入到所述编码层得到至少一个所述文本行图像的编码结果,并根据所述编码结果得到至少一个所述文本行图像的文本信息;

根据所述位置信息对至少一个所述文本行图像的文本信息进行有序排列得到所述目标图像的目标识别结果。

8. 一种字符识别装置,其特征在于,所述装置包括:

确定模块,用于确定包括待识别字符的目标图像对应的图像类别;其中,不同的图像类别对应不同的校正处理方式;

校正模块,用于通过所述图像类别对应的校正处理方式对所述目标图像进行校正处理;

提取模块,用于从校正处理后的目标图像中提取至少一个文本行图像;

识别模块,用于通过预设字符识别模型识别至少一个所述文本行图像中的所述待识别字符。

9. 根据权利要求8所述的装置,其特征在于,所述图像类别包括文档图像和场景图像。

10. 根据权利要求8或9所述的装置,其特征在于,所述确定模块包括:

第一获取子模块,用于获取已确定图像类别的图像样本;

第一确定子模块,用于根据所述图像样本确定所述目标图像对应的图像类别。

11. 根据权利要求9所述的装置,其特征在于,在所述图像类别为文档图像时,所述校正处理方式包括方向校正处理和/或畸变校正处理;在所述校正处理方式包括所述方向校正处理和所述畸变校正处理时,所述校正模块包括:

第二获取子模块,用于获取所述文档图像中的所述待识别字符与水平轴之间的第一倾斜角度;

第一校正子模块,用于在所述第一倾斜角度大于或者等于预设角度时,对所述文档图像进行方向校正处理;

第二确定子模块,用于确定所述文档图像中的所述待识别字符是否存在畸变;

第二校正子模块,用于在所述文档图像中的所述待识别字符存在畸变时,对所述文档图像进行畸变校正处理。

12. 根据权利要求9所述的装置,其特征在于,在所述图像类别为场景图像时,所述校正处理方式包括方向校正处理;所述校正模块包括:

检测子模块,用于对所述场景图像进行文字区域检测得到至少一个文字区域;

第三获取子模块,用于依次获取至少一个所述文字区域中的所述待识别字符与水平轴之间的第二倾斜角度;

第三校正子模块,用于在至少一个所述文字区域中的所述第二倾斜角度大于或者等于预设角度时,对至少一个所述文字区域进行方向校正处理。

13. 根据权利要求8或9所述的装置,其特征在于,还包括:

获取模块,用于在通过预设字符识别模型识别至少一个所述文本行图像中的所述待识别字符前,获取至少一个所述文本行图像的位置信息;

所述识别模块,用于通过所述预设字符识别模型和所述位置信息识别至少一个所述文本行图像中的所述待识别字符。

14. 根据权利要求13所述的装置,其特征在于,所述预设字符识别模型包括深度学习

层、循环网络层以及编码层,所述识别模块包括:

提取子模块,用于根据所述深度学习层对至少一个所述文本行图像进行字符特征提取;

第四获取子模块,用于将提取的字符特征输入至所述循环网络层得到至少一个所述文本行图像对应的特征向量;

第五获取子模块,用于将所述特征向量输入到所述编码层得到至少一个所述文本行图像的编码结果,并根据所述编码结果得到至少一个所述文本行图像的文本信息;

第六获取子模块,用于根据所述位置信息对至少一个所述文本行图像的文本信息进行有序排列得到所述目标图像的目标识别结果。

15. 一种计算机可读存储介质,其上存储有计算机程序,其特征在于,该程序被处理器执行时实现权利要求1-7中任一项所述方法的步骤。

16. 一种电子设备,其特征在于,包括:

存储器,其上存储有计算机程序;

处理器,用于执行所述存储器中的所述计算机程序,以实现权利要求1-7中任一项所述方法的步骤。

字符识别方法、装置、存储介质及电子设备

技术领域

[0001] 本公开涉及图像处理领域,具体地,涉及一种字符识别方法、装置、存储介质及电子设备。

背景技术

[0002] 随着计算机技术和多媒体的快速发展,越来越多的信息以图像形式传播,并且图像中的信息可以是描述性的文本,目前,文本图像可以划分为文档图像和场景图像,其中,文档图像通常包括的字符数量较多,字符分布规律,图像背景单一;与文档图像不同,场景图像通常包括的字符数量较少,字符类型丰富,字符分布随意,图像背景复杂。

[0003] 考虑到文档图像和场景图像具备上述不同的图像特征,并且目前的字符识别算法是针对特定的文本图像,使得文档图像和场景图像需要通过不同的字符识别算法分别进行字符识别,从而造成了字符识别算法的通用性较差。

发明内容

[0004] 为了解决上述问题,本公开提供一种字符识别方法、装置、存储介质及电子设备。

[0005] 根据本公开的第一方面,提供一种字符识别方法,所述方法包括:

[0006] 确定包括待识别字符的目标图像对应的图像类别;其中,不同的图像类别对应不同的校正处理方式;

[0007] 通过所述图像类别对应的校正处理方式对所述目标图像进行校正处理;

[0008] 从校正处理后的目标图像中提取至少一个文本行图像;

[0009] 通过预设字符识别模型识别至少一个所述文本行图像中的所述待识别字符。

[0010] 根据本公开的第二方面,提供一种字符识别装置,所述装置包括:

[0011] 确定模块,用于确定包括待识别字符的目标图像对应的图像类别;其中,不同的图像类别对应不同的校正处理方式;

[0012] 校正模块,用于通过所述图像类别对应的校正处理方式对所述目标图像进行校正处理;

[0013] 提取模块,用于从校正处理后的目标图像中提取至少一个文本行图像;

[0014] 识别模块,用于通过预设字符识别模型识别至少一个所述文本行图像中的所述待识别字符。

[0015] 根据本公开的第三方面,提供一种计算机可读存储介质,其上存储有计算机程序,该程序被处理器执行时实现上述第一方面所述方法的步骤。

[0016] 根据本公开的第四方面,提供一种电子设备,包括:

[0017] 存储器,其上存储有计算机程序;

[0018] 处理器,用于执行所述存储器中的所述计算机程序,以实现上述第一方面所述方法的步骤。

[0019] 在上述技术方案中,首先,可以确定包括待识别字符的目标图像对应的图像类别;

接着,通过所述图像类别对应的校正处理方式对所述目标图像进行校正处理;然后,从校正处理后的目标图像中提取至少一个文本行图像;最后,通过预设字符识别模型识别至少一个所述文本行图像中的所述待识别字符。由于不同的图像类别对应不同的校正处理方式,这样,可以将不同图像类别的图像按照对应的校正处理方式对所述目标图像进行校正处理,并对校正处理后的图像进行字符识别,本公开能够满足对文本图像和场景图像进行字符识别,从而避免了现有技术中字符识别算法的通用性较差的问题。

[0020] 本公开的其他特征和优点将在随后的具体实施方式部分予以详细说明。

附图说明

[0021] 附图是用来提供对本公开的进一步理解,并且构成说明书的一部分,与下面的具体实施方式一起用于解释本公开,但并不构成对本公开的限制。在附图中:

[0022] 图1是根据一示例性实施例示出的一种字符识别方法的流程示意图;

[0023] 图2是根据一示例性实施例示出的第一种字符识别装置的框图;

[0024] 图3是根据一示例性实施例示出的第二种字符识别装置的框图;

[0025] 图4是根据一示例性实施例示出的第三种字符识别装置的框图;

[0026] 图5是根据一示例性实施例示出的第四种字符识别装置的框图;

[0027] 图6是根据一示例性实施例示出的第五种字符识别装置的框图;

[0028] 图7是根据一示例性实施例示出的第六种字符识别装置的框图;

[0029] 图8是根据一示例性实施例示出的一种电子设备的框图。

具体实施方式

[0030] 首先,对本公开的应用场景进行说明,本公开可以应用于字符识别的场景,在该场景下,字符识别算法主要包括字符检测和字符识别两个步骤。目前,字符检测可分为单个字符检测和文本行提取两种方式,其中,单个字符检测即为直接对目标图像中的单个字符进行检测,文本行提取主要是提取出成行分布的字符区域。针对上述两种方式,单个字符检测极易发生漏检的情况,即目标图像中的一个或多个字符未被检测到,从而影响字符识别的准确率;文本行提取是将成行分布的字符当作整体,不易发生漏检,但检测出文本行后需要对文本行中的各个字符进行分割,从而对分割的准确率有较高的要求。对于上述不同的字符检测方式,字符识别方式也不同:在采用单个字符检测时,可以直接对提取的单个字符分别进行识别,并根据单个字符的字符位置信息对全部单个字符进行排列组合,从而生成最终识别结果;在采用文本行提取时,需要先对每个文本行中的字符进行分割,再对分割后的字符进行识别,并根据每个文本行的位置信息对每个文本行的字符识别结果进行排列组合,以生成最终识别结果。

[0031] 由于目前文本图像可以划分为文档图像和场景图像,其中,文档图像通常包括的字符数量较多,字符分布规律,图像背景单一;与文档图像不同,场景图像通常包括的字符数量较少,字符类型丰富,字符分布随意,图像背景复杂。对于文档图像和场景图像,由于具备上述不同的图像特征,使得当前的字符识别算法无法同时对文档图像和场景图像进行字符识别,而需要通过不同的字符识别算法分别进行字符识别,从而造成了字符识别算法的通用性较差。

[0032] 为了解决上述问题,本公开提出了一种字符识别方法、装置、存储介质及电子设备,首先,可以确定目标图像的图像类别,接着,根据图像类别确定该目标图像对应的校正处理方式,然后,根据该目标图像对应的校正处理方式对该目标图像进行校正处理,其次,可以从校正处理后的目标图像中提取至少一个文本行图像,最后,根据字符识别模型识别至少一个文本行图像中的待识别字符。由于不同的图像类别对应不同的校正处理方式,这样,可以将不同图像类别的图像按照对应的校正处理方式进行校正处理,并对校正处理后的图像进行字符识别,本公开能够满足对文本图像和场景图像进行字符识别,从而避免了现有技术中字符识别算法的通用性较差的问题。

[0033] 下面结合具体的实施例对本公开进行详细说明。

[0034] 图1是根据一示例性实施例示出的一种字符识别方法的流程示意图。如图1所示,所述方法包括:

[0035] S101、确定包括待识别字符的目标图像对应的图像类别。

[0036] 在本步骤中,该图像类别可以包括文档图像和场景图像,其中,文档图像通常包括的字符数量较多,字符分布规律,图像背景单一;与文档图像不同,场景图像通常包括的字符数量较少,字符类型丰富,字符分布随意,图像背景复杂,考虑到文档图像和场景图像之间具备上述不同的图像特征,因此,不同图像类别对应不同的校正处理方式,上述图像类别只是举例说明,本公开对此不作限定。

[0037] 在一种可能的实现方式中,可以获取已确定图像类别的图像样本,并根据该图像样本确定该目标图像对应的图像类别,进一步地,该图像样本可以包括文档图像样本和场景图像样本,并且该文档图像样本的数量和该场景图像样本的数量之间的差值小于或者等于预设阈值,这样,可以基于深度学习的方法,通过文档图像样本和场景图像样本训练预设分类器得到目标分类器,从而在该目标图像输入至该目标分类器中时,该目标分类器可以输出该目标图像对应的图像类别。

[0038] S102、通过该图像类别对应的校正处理方式对该目标图像进行校正处理。

[0039] 在该图像类别为文档图像时,由于文档图像中待识别字符通常处于密集分布,这样,若文档图像中的待识别字符存在倾斜和/或畸变时,则可能会影响到字符识别的准确率,为了避免该问题,本公开可以对该文档图像进行校正处理,该校正处理方式包括方向校正处理和/或畸变校正处理,此时,通过该图像类别对应的校正处理方式对该目标图像进行校正处理可以包括以下步骤:

[0040] S11、获取该文档图像中的该待识别字符与水平轴之间的第一倾斜角度。

[0041] 在一种可能的实现方式中,可以通过投影分析法或Hough变换法等得到该第一倾斜角度,当然,还可以对该文档图像进行阈值分割得到二值文档图像,并根据二值文档图像中的待识别字符的像素点信息获取该第一倾斜角度,具体过程可以参考现有技术,不再赘述。

[0042] S12、确定该第一倾斜角度是否大于或者等于预设角度。

[0043] 在该第一倾斜角度大于或者等于该预设角度时,执行步骤S13和S14;

[0044] 在该第一倾斜角度小于该预设角度时,执行步骤S14。

[0045] S13、对该文档图像进行方向校正处理。

[0046] 其中,方向校正处理可以是不不断旋转该目标图像,直至该文本图像中的待识别字

符与水平轴之间的第一倾斜角度小于该预设角度。

[0047] S14、确定该文档图像中的该待识别字符是否存在畸变。

[0048] 在采用扫描仪或者相机采集文本图像时,若文本本身倾斜和弯曲,或者拍摄视角倾斜等,则会导致文本图像存在畸变,这样,使得本来水平或者垂直的文本行变得弯曲,从而造成文本图像中的文本行之间存在干扰,影响待识别字符的最终识别结果。

[0049] 在该文档图像中的待识别字符存在畸变时,执行步骤S15;

[0050] 在该文档图像中的待识别字符不存在畸变时,确定完成校正处理。

[0051] S15、对该文档图像进行畸变校正处理。

[0052] 其中,畸变校正处理可以通过利用文本行之间的空白位置进行校正,使得文本行恢复成水平分布或者垂直分布,具体过程可以参考现有技术,不再赘述。

[0053] 需要说明的是,对于上述方法实施例,为了简单描述,故将其都表述为一系列的动作组合,但是本领域技术人员应该知悉,本公开并不受所描述的动作顺序的限制,因为依据本公开,某些步骤可以采用其他顺序或者同时进行,例如,步骤S14和S15可以在步骤S11之前执行,此时,可以先畸变校正处理,再进行方向校正处理;其次,本领域技术人员也应该知悉,说明书中所描述的实施例均属于优选实施例,所涉及的动作和模块并不一定是本公开所必须的。

[0054] 综上,基于文本图像的图像特征,步骤S11至S15可以将文本图像中待识别字符的第一倾斜角度和畸变进行校正,从而提高了后续步骤中的字符识别的准确率。

[0055] 在该图像类别为场景图像时,由于场景图像中待识别字符通常处于稀疏分布,并且往往存在随意分布的少量文本行,这样,场景图像中的文本行之间影响较小,从而无需进行畸变校正处理,因此,对于场景图像,对应的校正处理方式为方向校正处理,具体地,通过该图像类别对应的校正处理方式对该目标图像进行校正处理包括以下步骤:

[0056] S21、对该场景图像进行文字区域检测得到至少一个文字区域。

[0057] 其中,文字区域检测可以包括基于边缘检测、基于区域检测、基于纹理检测或者基于学习检测中的任一种,当然,还可以是上述四种检测方法中的两种、三种、或者四种的结合,上述示例只是举例说明,本公开对此不作限定。

[0058] S22、依次获取至少一个该文字区域中的该待识别字符与水平轴之间的第二倾斜角度。

[0059] 同样地,可以通过投影分析法或Hough变换法等得到该第二倾斜角度,当然,还可以对该场景图像进行阈值分割得到二值场景图像,并根据二值场景图像中的待识别字符的像素点信息获取该第二倾斜角度,具体过程可以参考现有技术,不再赘述。

[0060] 在该第二倾斜角度大于或者等于该预设角度时,执行步骤S23;

[0061] 在该第二倾斜角度小于该预设角度时,确定完成校正处理。

[0062] S23、对至少一个该文字区域进行方向校正处理。

[0063] 其中,方向校正处理可以是不断旋转该文字区域,直至该文本区域中的待识别字符与水平轴之间的第二倾斜角度小于该预设角度。

[0064] 综上,基于场景图像的图像特征,步骤S21至S23可以将场景图像中待识别字符的第二倾斜角度进行校正,从而提高了后续步骤中的字符识别的准确率。

[0065] S103、从校正处理后的目标图像中提取至少一个文本行图像。

[0066] 在本步骤中,可以基于深度学习的方法提取至少一个文本行图像,具体地,可以包括以下步骤:

[0067] S31、通过文本行检测模型中的多层卷积层提取目标图像的空间特征。

[0068] 其中,该空间特征可以是该目标图像中像素之间的相关关系。

[0069] S32、将该空间特征输入至文本行检测模型中的循环神经网络层得到该目标图像的序列特征。

[0070] 在本步骤中,该循环神经网络层可以是LSTM(长短时记忆网络;Long Short Term Memory Network)、BLSTM(双向长短时记忆网络;Bi-directional Long Short Term Memory Network)或者GRU(Gated Recurrent Unit,LSTM变体)等,上述示例只是举例说明,本公开对此不作限定。

[0071] S33、按照预设规则获取该目标图像中的候选文本框,并基于该序列特征对该候选文本框进行分类。

[0072] 在一种可能的实现方式中,可以采用预设大小和比例的滑动窗在该目标图像中滑动,以截取该候选文本框,具体过程参考现有技术,本公开不再赘述。

[0073] 其中,该分类过程可以通过该文本行检测模型中的分类层完成,示例地,该分类层可以是softmax层,并且该softmax层的输入与输出的维度一致,在该softmax层的输入与输出的维度不一致时,需要在softmax层前增加全连接层,从而达到softmax层的输入与输出的维度一致。

[0074] S34、使用文本行检测模型中的回归卷积层获取候选文本框的文本框位置信息。

[0075] S35、使用NMS(非最大值抑制;non maximum suppression)方法,根据该文本框位置信息和分类结果对候选文本框进行筛选得到文本行图像。

[0076] S104、通过预设字符识别模型识别至少一个该文本行图像中的该待识别字符。

[0077] 通常字符识别步骤以字符为单位处理,然后采用字符分类器进行字符预测,但是,在文本行图像复杂时,字符分割比较困难,可能会破坏字符结构,由于字符分割的精度直接影响到字符的最终识别结果,为了避免字符分割造成的识别准确率低的问题,本公开可以将文本行图像作为一个整体,不对该文本行图像中的待识别字符进行切割,直接识别文本行图像中的全部待识别字符,从而可以充分利用字符上下文关系进行识别。

[0078] 需要说明的是,在本步骤前,还包括:获取至少一个该文本行图像的位置信息,其中,由于在步骤S103中确定文本行图像后,可以根据文本框位置信息确定该文本行图像对应的位置信息,此时,通过该预设字符识别模型和该位置信息识别至少一个该文本行图像中的该待识别字符,该预设字符识别模型包括深度学习层、循环网络层以及编码层,具体地,字符识别过程可以包括以下步骤:

[0079] S41、根据该深度学习层对至少一个该文本行图像进行字符特征提取。

[0080] 其中,该深度学习层可以是CNN(卷积神经网络;Convolutional Neural Networks),这样,可以通过CNN将至少一个该文本行图像沿着水平方向形成多个切片,每个切片对应了一个字符特征,由于该相邻切片之间可能存在重叠,从而使得该字符特征包含一定的上下文关系。

[0081] S42、将提取的字符特征输入至该循环网络层得到至少一个该文本行图像对应的特征向量。

[0082] 其中,该循环神经网络层可以是LSTM、BLSTM或者GRU等,这样,通过该神经网络层可以进一步地学习该字符特征,从而得到切片对应的特征向量,上述示例只是举例说明,本公开对此不作限定。

[0083] S43、将该特征向量输入到该编码层得到至少一个该文本行图像的编码结果,并根据该编码结果得到至少一个该文本行图像的文本信息。

[0084] 在本步骤中,该编码层可以是CTC(时序分类算法;Connectionist Temporal Classification)层,这样,可以根据CTC层得到编码结果,由于该文本行图像中可以包括多个待识别字符,因此,该编码结果中可以包括多个编码,这样,将该编码结果中的每个编码与预设编码对应关系进行匹配得到每个编码对应的字符,根据该多个编码的编码顺序将每个编码对应的字符进行有序排列得到该文本行图像的文本信息,其中,该预设编码对应关系为编码样本与字符样本之间的对应关系,上述示例只是举例说明,本公开对此不作限定。

[0085] S44、根据该位置信息对至少一个该文本行图像的文本信息进行有序排列得到该目标图像的目标识别结果。

[0086] 在本步骤中,可以根据该位置信息得到该文本行图像中的至少一个文本行图像之间的先后顺序,从而将至少一个文本行图像的文本信息按照先后顺序进行排序得到目标识别结果。

[0087] 需要说明的是,本公开是以目标图像中的待识别字符为水平排列为例进行说明的,在该待识别字符是垂直排列时,可以提取该目标图像中的至少一个文本列图像,并通过预设字符识别模型识别至少一个该文本列图像中的该待识别字符,具体过程可以参考上述文本行图像的叙述,不再赘述。

[0088] 采用上述方法,首先,可以确定目标图像的图像类别,接着,根据图像类别确定该目标图像对应的校正处理方式,然后,根据该目标图像对应的校正处理方式对该目标图像进行校正处理,其次,可以从校正处理后的目标图像中提取至少一个文本行图像,最后,根据字符识别模型识别至少一个文本行图像中的待识别字符。由于不同的图像类别对应不同的校正处理方式,这样,可以将不同图像类别的图像按照对应的校正处理方式进行校正处理,并对校正处理后的图像进行字符识别,本公开能够满足对文本图像和场景图像进行字符识别,从而避免了现有技术中字符识别算法的通用性较差的问题。

[0089] 图2是根据一示例性实施例示出的字符识别装置20的框图,如图2所示,包括:

[0090] 确定模块201,用于确定包括待识别字符的目标图像对应的图像类别;其中,不同的图像类别对应不同的校正处理方式;

[0091] 校正模块202,用于通过该图像类别对应的校正处理方式对该目标图像进行校正处理;

[0092] 提取模块203,用于从校正处理后的目标图像中提取至少一个文本行图像;

[0093] 识别模块204,用于通过预设字符识别模型识别至少一个该文本行图像中的该待识别字符。

[0094] 可选地,该图像类别包括文档图像和场景图像。

[0095] 图3是根据一示例性实施例示出的确定模块201的框图,如图3所示,该确定模块201包括:

[0096] 第一获取子模块2011,用于获取已确定图像类别的图像样本;

[0097] 第一确定子模块2012,用于根据该图像样本确定该目标图像对应的图像类别。

[0098] 图4是根据一示例性实施例示出的校正模块202的框图,如图4所示,在该图像类别为文档图像时,该校正处理方式包括方向校正处理和/或畸变校正处理;在该校正处理方式包括该方向校正处理和该畸变校正处理时,该校正模块202包括:

[0099] 第二获取子模块2021,用于获取该文本图像中的该待识别字符与水平轴之间的第一倾斜角度;

[0100] 第一校正子模块2022,用于在该第一倾斜角度大于或者等于预设角度时,对该文本图像进行方向校正处理;

[0101] 第二确定子模块2023,用于确定该文本图像中的该待识别字符是否存在畸变;

[0102] 第二校正子模块2024,用于在该文本图像中的该待识别字符存在畸变时,对该文本图像进行畸变校正处理。

[0103] 图5是根据一示例性实施例示出的校正模块202的框图,如图5所示,在该图像类别为场景图像时,该校正处理方式包括方向校正处理;该校正模块202包括:

[0104] 检测子模块2025,用于对该场景图像进行文字区域检测得到至少一个文字区域;

[0105] 第三获取子模块2026,用于依次获取至少一个该文字区域中的该待识别字符与水平轴之间的第二倾斜角度;

[0106] 第三校正子模块2027,用于在至少一个该文字区域中的该第二倾斜角度大于或者等于预设角度时,对至少一个该文字区域进行方向校正处理。

[0107] 图6是根据一示例性实施例示出的字符识别装置20的框图,如图6所示,还包括:

[0108] 获取模块305,用于在通过预设字符识别模型识别至少一个该文本行图像中的该待识别字符前,获取至少一个该文本行图像的位置信息;

[0109] 该识别模块304,用于通过该预设字符识别模型和该位置信息识别至少一个该文本行图像中的该待识别字符。

[0110] 图7是根据一示例性实施例示出的识别模块304的框图,如图7所示,该预设字符识别模型包括深度学习层、循环网络层以及编码层,该识别模块304包括:

[0111] 提取子模块3041,用于根据该深度学习层对至少一个该文本行图像进行字符特征提取;

[0112] 第四获取子模块3042,用于将提取的字符特征输入至该循环网络层得到至少一个该文本行图像对应的特征向量;

[0113] 第五获取子模块3043,用于将该特征向量输入到该编码层得到至少一个该文本行图像的编码结果,并根据该编码结果得到至少一个该文本行图像的文本信息;

[0114] 第六获取子模块3044,用于根据该位置信息对至少一个该文本行图像的文本信息进行有序排列得到该目标图像的目标识别结果。

[0115] 采用上述装置,首先,可以确定目标图像的图像类别,接着,根据图像类别确定该目标图像对应的校正处理方式,然后,根据该目标图像对应的校正处理方式对该目标图像进行校正处理,其次,可以从校正处理后的目标图像中提取至少一个文本行图像,最后,根据字符识别模型识别至少一个文本行图像中的待识别字符。由于不同的图像类别对应不同的校正处理方式,这样,可以将不同图像类别的图像按照对应的校正处理方式进行校正处理,并对校正处理后的图像进行字符识别,本公开能够满足对文本图像和场景图像进行字

符识别,从而避免了现有技术中字符识别算法的通用性较差的问题。

[0116] 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有关该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。

[0117] 图8是根据一示例性实施例示出的一种电子设备800的框图。如图8所示,该电子设备800可以包括:处理器801,存储器802。该电子设备800还可以包括多媒体组件803,输入/输出(I/O)接口804,以及通信组件805中的一者或多者。

[0118] 其中,处理器801用于控制该电子设备800的整体操作,以完成上述的字符识别方法中的全部或部分步骤。存储器802用于存储各种类型的数据以支持在该电子设备800的操作,这些数据例如可以包括用于在该电子设备800上操作的任何应用程序或方法的指令,以及应用程序相关的数据,例如联系人数据、收发的消息、图片、音频、视频等等。该存储器802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,例如静态随机存取存储器(Static Random Access Memory,简称SRAM),电可擦除可编程只读存储器(Electrically Erasable Programmable Read-Only Memory,简称EEPROM),可擦除可编程只读存储器(Erasable Programmable Read-Only Memory,简称EPROM),可编程只读存储器(Programmable Read-Only Memory,简称PROM),只读存储器(Read-Only Memory,简称ROM),磁存储器,快闪存储器,磁盘或光盘。多媒体组件803可以包括屏幕和音频组件。其中屏幕例如可以是触摸屏,音频组件用于输出和/或输入音频信号。例如,音频组件可以包括一个麦克风,麦克风用于接收外部音频信号。所接收的音频信号可以被进一步存储在存储器802或通过通信组件805发送。音频组件还包括至少一个扬声器,用于输出音频信号。I/O接口804为处理器801和其他接口模块之间提供接口,上述其他接口模块可以是键盘,鼠标,按钮等。这些按钮可以是虚拟按钮或者实体按钮。通信组件805用于该电子设备800与其他设备之间进行有线或无线通信。无线通信,例如Wi-Fi,蓝牙,近场通信(Near Field Communication,简称NFC),2G、3G或4G,或它们中的一种或几种的组合,因此相应的该通信组件805可以包括:Wi-Fi模块,蓝牙模块,NFC模块。

[0119] 在一示例性实施例中,电子设备800可以被一个或多个应用专用集成电路(Application Specific Integrated Circuit,简称ASIC)、数字信号处理器(Digital Signal Processor,简称DSP)、数字信号处理设备(Digital Signal Processing Device,简称DSPD)、可编程逻辑器件(Programmable Logic Device,简称PLD)、现场可编程门阵列(Field Programmable Gate Array,简称FPGA)、控制器、微控制器、微处理器或其他电子元件实现,用于执行上述的字符识别方法。

[0120] 在另一示例性实施例中,还提供了一种包括程序指令的计算机可读存储介质,该程序指令被处理器执行时实现上述的字符识别方法的步骤。例如,该计算机可读存储介质可以为上述包括程序指令的存储器802,上述程序指令可由电子设备800的处理器801执行以完成上述的字符识别方法。

[0121] 以上结合附图详细描述了本公开的优选实施方式,但是,本公开并不限于上述实施方式中的具体细节,在本公开的技术构思范围内,可以对本公开的技术方案进行多种简单变型,这些简单变型均属于本公开的保护范围。

[0122] 另外需要说明的是,在上述具体实施方式中所描述的各个具体技术特征,在不矛盾的情况下,可以通过任何合适的方式进行组合。为了避免不必要的重复,本公开对各种可

能的组合方式不再另行说明。

[0123] 此外,本公开的各种不同的实施方式之间也可以进行任意组合,只要其不违背本公开的思想,其同样应当视为本公开所公开的内容。

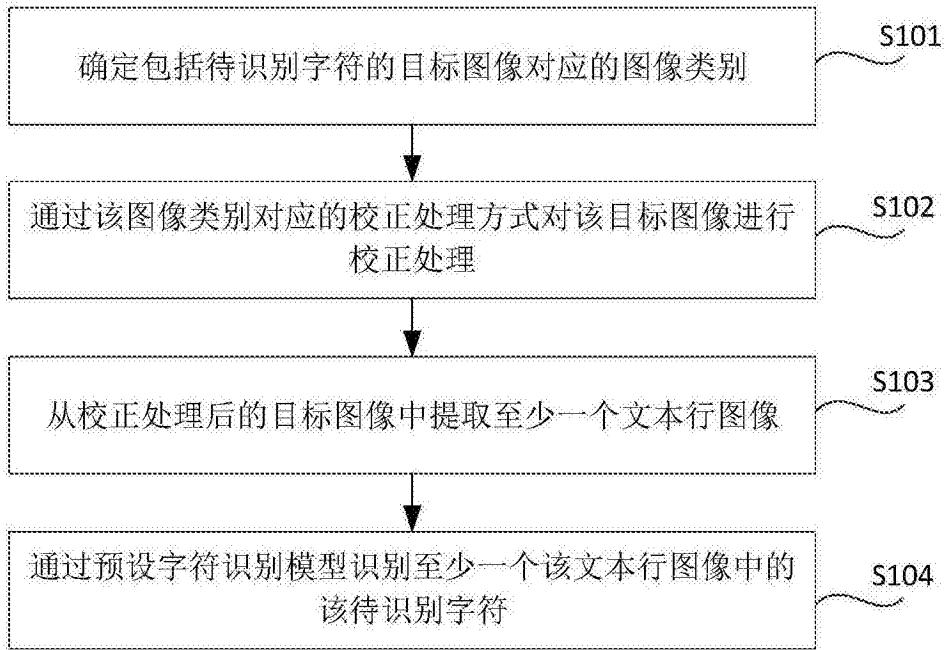


图1

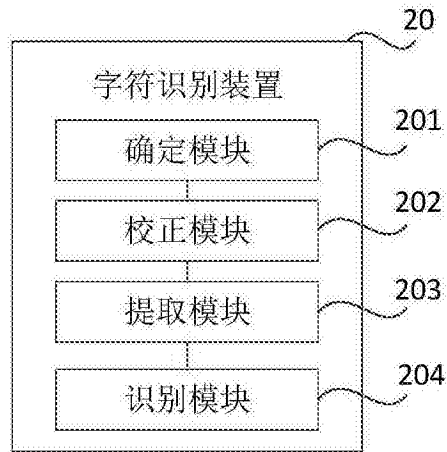


图2

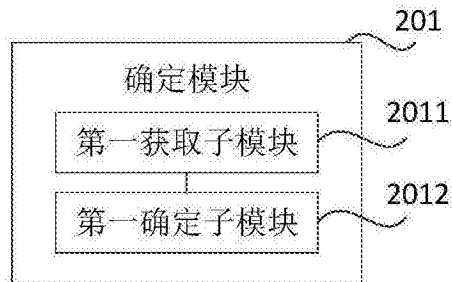


图3

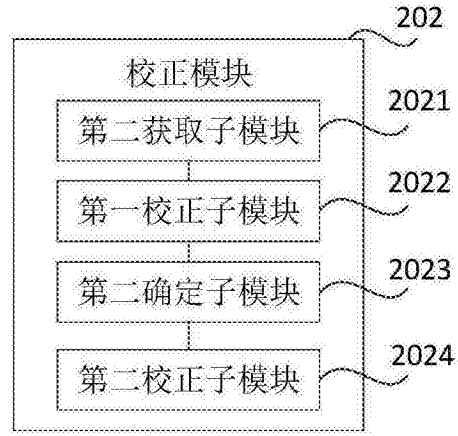


图4

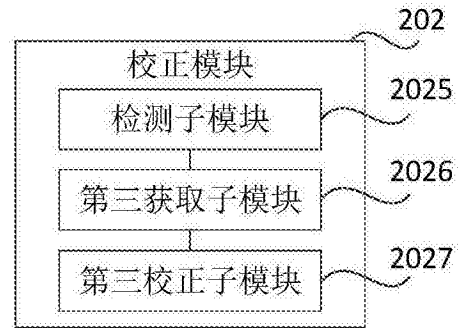


图5

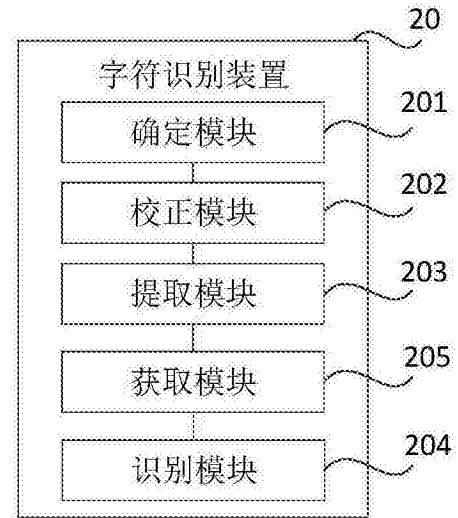


图6

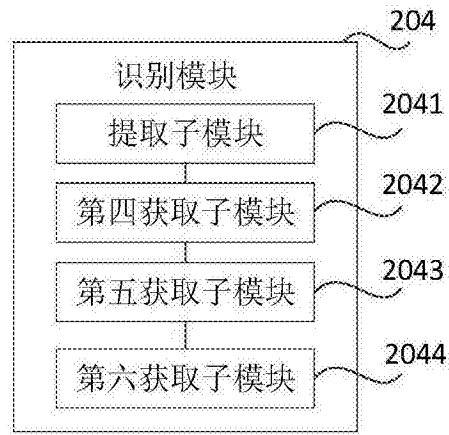


图7

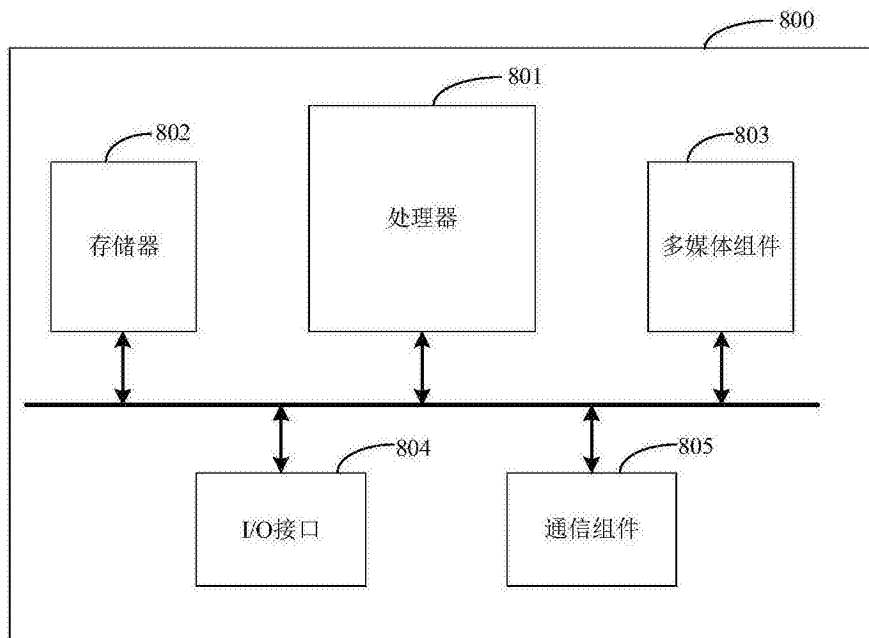


图8