



(12) 发明专利

(10) 授权公告号 CN 114511853 B

(45) 授权公告日 2022.07.12

(21) 申请号 202210417388.X

G06V 30/19 (2022.01)

(22) 申请日 2022.04.21

G06T 7/00 (2017.01)

G06K 9/62 (2022.01)

(65) 同一申请的已公布的文献号

申请公布号 CN 114511853 A

(56) 对比文件

(43) 申请公布日 2022.05.17

CN 114240804 A, 2022.03.25

JP 2011216070 A, 2011.10.27

(73) 专利权人 华南理工大学

地址 510000 广东省广州市天河区五山路
381号

Hung Tuan Nguyen et al. Online

trajectory recovery from offline
handwritten Japanese kanji characters of
multiple strokes.《2020 25th International
Conference on Pattern Recognition
(ICPR)》.2021,1-10.

专利权人 人工智能与数字经济广东省实验
室(广州)

(72) 发明人 黄双萍 陈洲楠 杨代辉 梁景麟

彭政华

詹琦梁等.一种结合多种图像分割算法的实
例分割方案.《小型微型计算机系统》.2020,第41
卷(第4期),

(74) 专利代理机构 东莞卓诚专利代理事务所

(普通合伙) 44754

审查员 邱立英

专利代理师 朱鹏

(51) Int. Cl.

G06V 30/32 (2022.01)

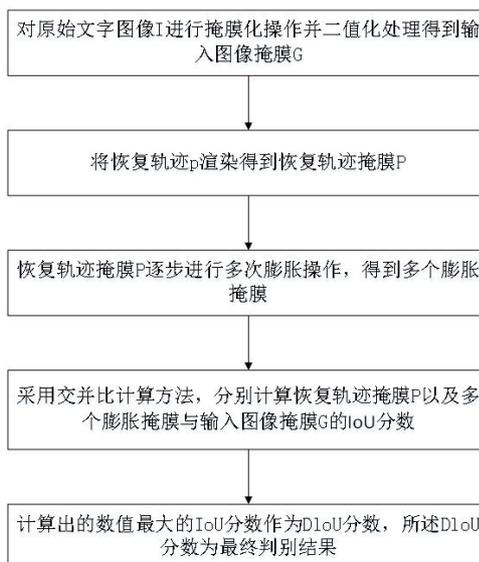
权利要求书1页 说明书4页 附图1页

(54) 发明名称

一种文字图像书写轨迹恢复效果判别方法

(57) 摘要

本发明公开了一种文字图像书写轨迹恢复效果判别方法,包括:对原始文字图像I进行掩膜化操作并二值化处理得到输入图像掩膜G;将恢复轨迹渲染得到恢复轨迹掩膜P;恢复轨迹掩膜P逐步进行多次膨胀操作,得到多个膨胀掩膜;采用交并比计算方法,分别计算恢复轨迹掩膜P以及多个膨胀掩膜与输入图像掩膜G的IoU分数;计算出的数值最大的IoU分数作为D1oU分数,所述D1oU分数为最终判别结果。本发明方法设计交并比计算方法,计算输入图像掩膜和恢复轨迹掩膜的相似度,设计动态膨胀机制,将恢复轨迹掩膜和膨胀掩膜分别与输入图像掩膜计算交并比分数,使用动态交并比分数消除输入文字图像笔画宽度对字形保真程度评价的影响。



1. 一种文字图像书写轨迹恢复效果判别方法,其特征在于,包括以下步骤:

步骤1,对原始文字图像 I 进行掩膜化操作并二值化处理得到输入图像掩膜 G ;

步骤2,将恢复轨迹 $p = (p_1, \dots, p_M)$ 渲染得到恢复轨迹掩膜 P ,其中, M 为恢复轨迹点数, p_1 为第1个恢复轨迹点的位置信息, p_M 为第 M 个恢复轨迹点的位置信息;

步骤3,恢复轨迹掩膜 P 逐步进行多次膨胀操作,得到多个膨胀掩膜;

步骤4,采用交并比计算方法,分别计算恢复轨迹掩膜 P 以及多个膨胀掩膜与输入图像掩膜 G 的IoU分数;

步骤5,计算出的数值最大的IoU分数作为D1oU分数,所述D1oU分数为最终判别结果。

2. 根据权利要求1所述的一种文字图像书写轨迹恢复效果判别方法,其特征在于,所述的交并比计算方法计算字形相似度IoU分数的公式为:

$$\text{IoU}(G, P) = \frac{|G \cap P|}{|G \cup P|}$$

其中, \cap 和 \cup 分别为“交”和“与”运算, $|\cdot|$ 计算掩膜中的前景像素个数。

3. 根据权利要求1所述的一种文字图像书写轨迹恢复效果判别方法,其特征在于,所述的原始文字图像 I 、输入图像掩膜 G 和恢复轨迹掩膜 P 的尺寸大小相同。

4. 根据权利要求2或3所述的一种文字图像书写轨迹恢复效果判别方法,其特征在于,所述的原始文字图像 I 、输入图像掩膜 G 和恢复轨迹掩膜 P 的尺寸大小为 64×64 ,在图像掩膜 G 是用0表示空白背景,用1表示文字前景的二进制位图,步骤2中渲染的过程中采用1为笔画宽度。

5. 根据权利要求1所述的一种文字图像书写轨迹恢复效果判别方法,其特征在于,步骤3中所述的膨胀操作的过程中采用 3×3 的膨胀核,膨胀次数小于10。

一种文字图像书写轨迹恢复效果判别方法

技术领域

[0001] 本发明涉及文字图像识别技术领域,尤其涉及一种文字图像书写轨迹恢复效果判别方法。

背景技术

[0002] 文字数据可以依据模态类型大致划分为图像模态数据和书写轨迹模态数据两大类,文字生成技术主要围绕这两种模态形式展开。文字图像往往是由扫描仪或是摄像头等图像采集设备得到,以点阵图像的形式保存,这种数据能直观地展示文字的形状,常用于文字的显示和阅读。文字书写轨迹由可以记录轨迹的数字笔,手写板或是触控屏等交互设备采集得到,通常以笔尖坐标点轨迹序列的方式存储,可能还会记录书写过程中的笔尖压力和速度等辅助信息。文字图像的书写轨迹恢复,是一种跨模态的文字生成技术,其目的是从不包含轨迹信息的文字图像中恢复得到其书写运动轨迹信息,常作为文字识别和数据增广的重要技术手段,也在司法笔迹鉴定、写字机器人、字体生成和文字特效生成等领域有巨大应用潜力。

[0003] 文字图像的书写轨迹恢复算法不仅要学习轨迹点在空间上的位置以及不同轨迹点之间的先后顺序,还应该保证字形保真,即恢复轨迹的字形应该在视觉上和输入图像相似。当前,轨迹恢复效果的判别方法仅仅评价轨迹点的位置和先后顺序,而忽视了轨迹的字形保真。现有的判别评价方法中,一类是依赖于主观感受,通过人对恢复轨迹和输入图像进行视觉比较,直观感受出文字轨迹的恢复效果。这种评价方法具有主观性,非定量性,易错性,一类是通过手写识别模型的准确性间接评价轨迹的恢复质量。这种评价方法在一定程度上反映轨迹的重建质量,但它依赖于识别模型,其评价效果容易受到手写识别模型效果的影响。它无法对轨迹点的位置,先后顺序和轨迹的字形保真等方面做出评价,还有一类是用于直接评价恢复质量。但它们绝大多数都只关注了轨迹点位置和先后顺序的评价,例如,RMSE(root mean squared error,根均方误差)直接计算恢复和真实轨迹之间两两轨迹点的距离,DTW(dynamic time warping,动态时间规整)先计算得到恢复和真实轨迹之间的最佳轨迹点匹配关系,随后根据该匹配关系计算两个轨迹之间的距离,这类方法可以评价恢复轨迹的轨迹点位置和先后顺序的误差,却没有对字形保真进行评价。

发明内容

[0004] 有鉴于此,本发明的目的在于提供一种文字图像书写轨迹恢复效果判别方法,以解决现有技术在进行文字图像书写轨迹恢复判别时所存在的受图像背景干扰,无法准确反映字型保真程度,且现有技术对文字笔画不具鲁棒性,导致判别结果出现偏差的问题。

[0005] 本发明公开了一种文字图像书写轨迹恢复效果判别方法,包括以下步骤:

[0006] 步骤1,对原始文字图像 I 进行掩膜化操作并二值化处理得到输入图像掩膜 G ;

[0007] 步骤2,将恢复轨迹 $p = (p_1, \dots, p_M)$ 渲染得到恢复轨迹掩膜 P , p_1 为第1

个恢复轨迹点的位置信息, p_M 为第 M 个恢复轨迹点的位置信息;

[0008] 步骤3, 恢复轨迹掩膜 P 逐步进行多次膨胀操作, 得到多个膨胀掩膜;

[0009] 步骤4, 采用交并比计算方法, 分别计算恢复轨迹掩膜 P 以及多个膨胀掩膜与输入图像掩膜 G 的IoU分数;

[0010] 步骤5, 步骤4中计算出的数值最大的IoU分数作为DIoU分数, 所述DIoU分数为最终判别结果;

[0011] 具体地, 所述的交并比计算方法计算字形相似度IoU分数的公式为:

$$[0012] \quad \text{IoU}(G, P) = \frac{|G \cap P|}{|G \cup P|}$$

[0013] 其中, \cap 和 \cup 分别为“交”和“与”运算, $|\cdot|$ 计算掩膜中的前景像素个数。

[0014] 优选地, 所述的原始文字图像 I 、输入图像掩膜 G 和恢复轨迹掩膜 P 的尺寸大小相同, 在图像掩膜 G 是用0表示空白背景, 用1表示文字前景的二进制位图。

[0015] 优选地, 所述的原始文字图像 I 、输入图像掩膜 G 和恢复轨迹掩膜 P 的尺寸大小为 64×64 , 步骤2中渲染的过程中采用1为笔画宽度。

[0016] 优选地, 步骤3中所述的膨胀操作的过程中采用 3×3 的膨胀核, 膨胀次数小于10。

[0017] 本发明方法相对于现有技术而言, 本发明方法提出动态交并比 (Dynamic Intersection on Union, DIoU) 判别方法, 用于判别恢复轨迹对输入文字图像的字形保真程度。首先, 设计交并比计算方法, 计算输入图像掩膜和恢复轨迹掩膜的相似度。随后, 设计动态膨胀机制, 逐步膨胀恢复轨迹掩膜, 得到多个膨胀掩膜, 将恢复轨迹掩膜和膨胀掩膜分别与输入图像掩膜计算交并比分数, 使用数值最大的交并比分数为动态交并比分数, 以此消除输入文字图像笔画宽度对字形保真程度评价的影响。

附图说明

[0018] 图1示出了本发明实施方法的流程示意图;

[0019] 图2示出了本发明实施例的判别流程示意图。

具体实施方式

[0020] 为了使本发明的目的、技术方案和优点更加清楚, 下面将结合附图对本发明作进一步地详细描述, 显然, 所描述的实施例仅仅是本发明一部份实施例, 而不是全部的实施例。基于本发明中的实施例, 本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其它实施例, 都属于本发明保护的范围。应当理解, 此处所描述的具体实施例仅仅用以解释本发明, 并不用于限定本发明。

[0021] 图1示出了本发明实施例的流程示意图。一种文字图像书写轨迹恢复效果判别方

法,包括以下步骤:

[0022] 步骤1,对原始文字图像 I 进行掩膜化操作并二值化处理得到输入图像掩膜 G ;

[0023] 步骤2,将恢复轨迹 $p = (p_1, \dots, p_M)$ 渲染得到恢复轨迹掩膜 P , p_1 为第1个恢复轨迹点的位置信息, p_M 为第 M 个恢复轨迹点的位置信息;

[0024] 步骤3,恢复轨迹掩膜 P 逐步进行多次膨胀操作,得到多个膨胀掩膜;

[0025] 步骤4,采用交并比计算方法,分别计算恢复轨迹掩膜 P 以及多个膨胀掩膜与输入图像掩膜 G 的IoU分数;

[0026] 步骤5,步骤4中计算出的数值最大的IoU分数作为DIoU分数,所述DIoU分数为最终判别结果。

[0027] 本实施例的具体操作步骤如下:

[0028] 如图2所示,对于输入文字图像 I ,文字图像轨迹恢复算法得到恢复轨迹 $p = (p_1, \dots, p_M)$,其中 M 为轨迹点数。设计动态交并比 (Dynamic Intersection on Union, DIoU) 判别方法,计算DIoU分数,用于判定恢复轨迹 p 对输入文字图像 I 的字形保真程度:

[0029] (1)对输入文字图像 I 进行掩膜化操作:二值化处理得到输入图像掩膜 G ,即用0表示空白背景,用1表示文字前景的二进制位图。以1为笔画宽度,将恢复轨迹 p 渲染得到恢复轨迹掩膜 P ,其尺寸与 G 相同。本实例中, I , G 和 P 的尺寸都为 64×64 。

[0030] (2)设计交并比计算方法 (Intersection on Union, IoU),用于计算掩膜间的字形相似度:

$$[0031] \quad \text{IoU}(G, P) = \frac{|G \cap P|}{|G \cup P|}$$

[0032] 其中, \cap 和 \cup 分别为“交”和“与”运算, $|\cdot|$ 计算掩膜中的前景像素个数。

[0033] (3)设计动态膨胀机制:对恢复轨迹掩膜 P 逐步进行多次膨胀操作,得到多个膨胀掩膜。将恢复轨迹掩膜 P 和膨胀掩膜分别与输入图像掩膜 G 计算IoU分数,使用数值最大的IoU分数为DIoU分数。计算伪代码如下表所示。本实例中,使用 3×3 的膨胀核,膨胀次数设置为7。动态交并比计算的计算伪代码如下表1所示。

[0034] 表1动态交并比计算步骤

[0035]

动态交并比计算步骤	
输入:	p, I
1:	将 I 掩膜化得到 G , 将 p 渲染得到 P
2:	$n \leftarrow 0, result \leftarrow 0, P_0 \leftarrow P$
3:	重复
4:	计算 $IoU(G, P_n)$
5:	如果 $IoU(G, P_n) > result$:
6:	$result \leftarrow IoU(G, P_n)$
7:	对 P_n 进行膨胀核为 3×3 的膨胀操作, 得到 P_{n+1}
8:	$n \leftarrow n + 1$
9:	直到 $n = 7$
10:	$DIoU \leftarrow result$
输出:	$DIoU$

[0036] 与现有技术相比,本发明方法提出的DIoU只对文字前景进行相似度计算,并设置了动态膨胀机制以消除输入文字图像笔画宽度对字形保真程度评价的影响。对文字图像轨迹恢复的字形保真效果进行评价,有助于更加科学地去评价轨迹恢复方法的性能,也对轨迹恢复方法的改进有很大的指导意义。

[0037] 本领域普通技术人员可以意识到,结合本文中所公开的实施例描述的各示例的单元及算法步骤,能够以电子硬件、或者计算机软件和电子硬件的结合来实现。这些功能究竟以硬件还是软件方式来执行,取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能,但是这种实现不应认为超出本发明的范围。

[0038] 以上实施例的各技术特征可以进行任意的组合,为使描述简洁,未对上述实施例中的各个技术特征所有可能的组合都进行描述,然而,只要这些技术特征的组合不存在矛盾,都应当认为是本说明书记载的范围。

[0039] 本领域技术人员应明白,本申请的实施例可提供为方法、系统或计算机程序产品。因此,本申请可采用完全硬件实施例、完全软件实施例或结合软件和硬件方面的实施例的形式。而且,本申请可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质(包括但不限于磁盘存储器、CD-ROM、光学存储器等)上实施的计算机程序产品的形式。

[0040] 以上所述仅是本发明的优选实施方式,应当指出,对于本技术领域的普通技术人员来说,在不脱离本发明原理的前提下,还可以做出若干改进和润饰,这些改进和润饰也应视为本发明的保护范围。

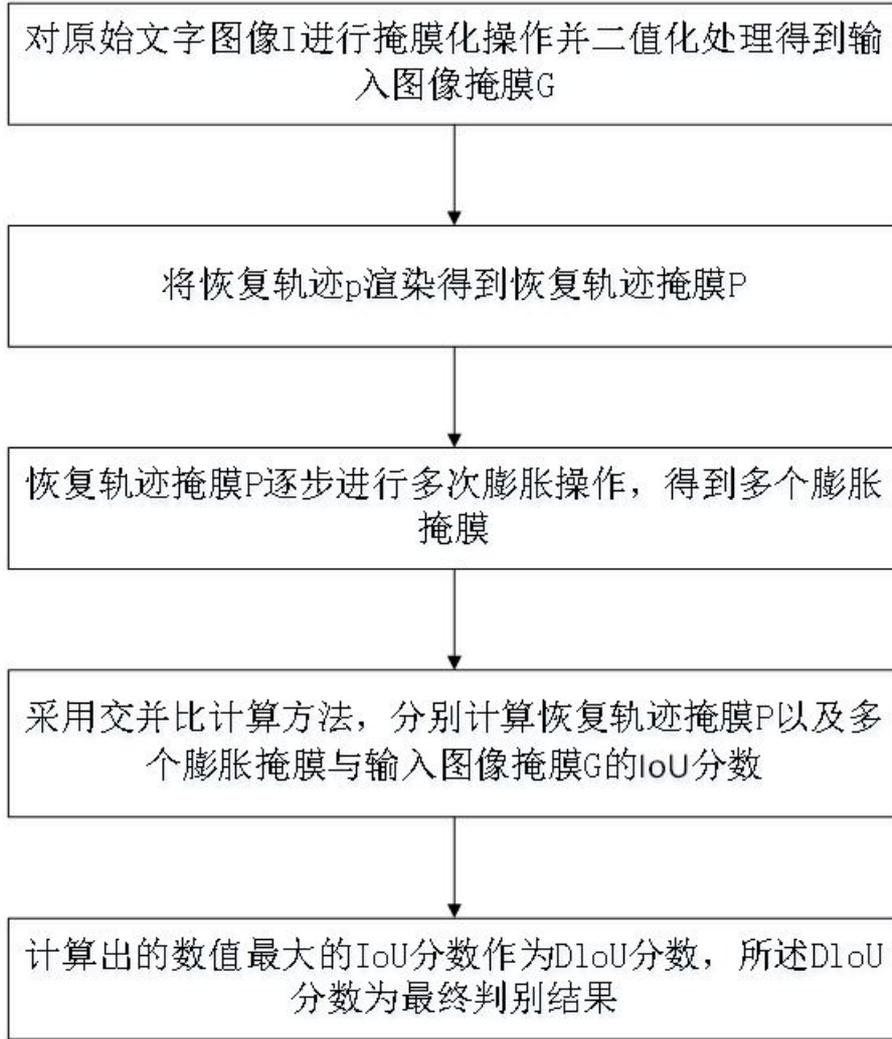


图1

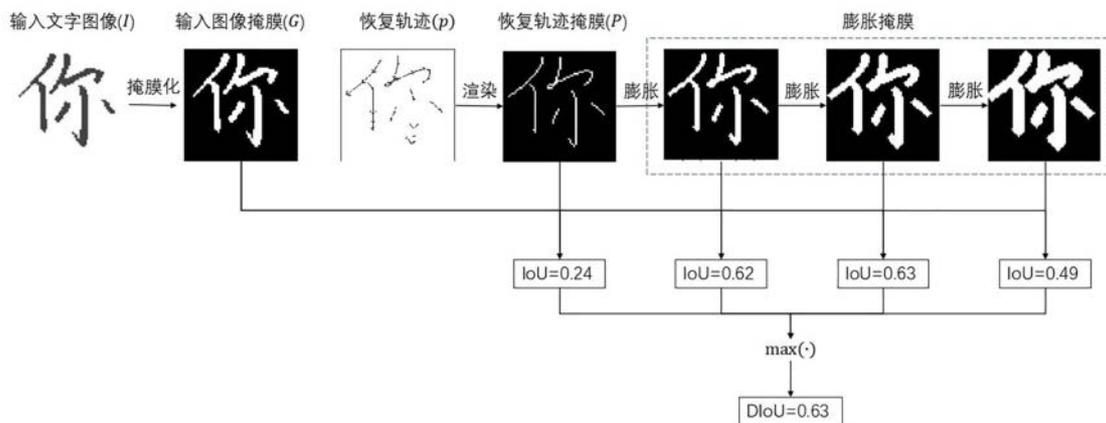


图2