(19)国家知识产权局

(12)发明专利

(10)授权公告号 CN 114463760 B
(45)授权公告日 2022.06.28

(21)申请号 202210363354.7

(22)申请日 2022.04.08

(65)同一申请的已公布的文献号
申请公布号 CN 114463760 A

(43)申请公布日 2022.05.10

(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
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

(74)专利代理机构 东莞卓诚专利代理事务所(普通合伙) 44754
专利代理师 朱鹏

(51)Int.Cl.
G06V 30/32(2022.01)
G06V 20/62(2022.01)
G06V 30/10(2022.01)
G06V 10/28(2022.01)
G06V 10/44(2022.01)
G06V 10/764(2022.01)
G06V 10/82(2022.01)
G06K 9/62(2022.01)
G06N 3/04(2006.01)
G06N 3/08(2006.01)

(56)对比文件
WO 2021136144 A1,2021.07.08
CN 110188669 A,2019.08.30
CN 109410242 A,2019.03.01
CN 114428866 A,2022.05.03
US 2021027470 A1,2021.01.28
shuangping huang.《OBC306: A Large-Scale Oracle Bone Character Recognition Dataset》.《2019 International Conference on Document Analysis and Recognition 》.2020,第681-688页.

审查员 郑明月

权利要求书3页 说明书7页 附图2页

(54)发明名称

一种基于双流编码的文字图像书写轨迹恢复方法

(57)摘要

本发明公开了一种基于双流编码的文字图像书写轨迹恢复方法,包括以下步骤:将文字图像调整为预设尺寸并进行二值化处理;构建双流编码网络,所述的双流编码网络输入为文字图像,输出为双流融合编码特征Z;构建解码网络,所述的解码网络的输入为双流融合编码特征Z,输出为预测文字书写轨迹序列;联合训练双流编码网络和解码网络,获得文字图像书写轨迹恢复网络模型;利用训练完成的文字图像书写轨迹恢复网络模型进行书写轨迹恢复。本发明方法在编码过程中,分别提取文字在竖直和水平方向上的特征,实现对特征进行降采样、减少参数量的同时,保留必要的文字字形信息,帮助后续解码精确地反映文字的字形,有效提升文字图像书写轨迹的恢复性能。

CN 114463760 B

1.一种基于双流编码的文字图像书写轨迹恢复方法,其特征在于,包括以下步骤:

步骤1,将文字图像调整为预设尺寸并进行二值化处理;

步骤2,构建双流编码网络,所述的双流编码网络输入为文字图像,输出为双流融合编码特征Z;

步骤3,构建解码网络,所述的解码网络的输入为双流融合编码特征Z,输出为预测文字书写轨迹序列;

步骤4,联合训练双流编码网络和解码网络,获得文字图像书写轨迹恢复网络模型;

步骤5,利用训练完成的文字图像书写轨迹恢复网络模型进行书写轨迹恢复;

所述的双流编码网络包括竖直卷积循环神经网络、水平卷积循环神经网络和注意力模块;

并行连接的竖直卷积循环神经网络和水平卷积循环神经网络均包含CNN编码器和BiLSTM编码器,竖直卷积循环神经网络中的CNN编码器利用竖直方向上的降采样操作来进行竖直方向上的降采样,然后配合卷积操作,对输入文字图像进行编码,得到文字在水平方向上的一维方向特征C_x,一维方向特征C_x在方向维度上拆分后得到以方向为时序的特征序列,竖直卷积循环神经网络中的BiLSTM编码器对时序的特征序列进行编码,得到双流编码特征V_x;水平卷积循环神经网络中的CNN编码器利用水平方向上的降采样操作来进行水平方向上的降采样,然后配合卷积操作,对输入文字图像进行编码,得到文字在竖直方向上的一维方向特征C_y,一维方向特征C_y在方向维度上拆分后得到以方向为时序的特征序列,水平卷积循环神经网络中的BiLSTM编码器对时序的特征序列进行编码,得到双流编码特征V_y;

在注意力模块中对双流编码特征V_x和V_y进行融合,得到双流融合编码特征Z:

$$Z=\sum_{i=1}^{|V|} v_i * w_i ,$$

其中,$w_i=\frac{e^{s_i}}{\sum_{j=1}^{|V|} e^{s_j}}$, $s_i=f(v_i)=Uv_i$,$s_i=f(v_i)=Uv_i$,通过合并特征V_x和V_y得到V,v_i和v_j为V的第i个分量和第j个分量,s_i表示v_i的注意力权重,s_j表示v_j的注意力权重,$f(\cdot)$表示一个全连接层的函数,$|V|$为V的长度,U为一个全连接层的可学习参数。

2.根据权利要求1所述的一种基于双流编码的文字图像书写轨迹恢复方法,其特征在于,所述的降采样操作为非对称池化操作、非对称卷积操作或全连接层网络操作降采样。

3.根据权利要求1所述的一种基于双流编码的文字图像书写轨迹恢复方法,其特征在于,所述的解码网络为LSTM解码器,LSTM解码器以双流融合编码特征Z为输入,依次预测轨

迹点;LSTM解码器根据$t-1$时刻的预测值p_{t-1}和隐藏层向量h_{t-1},预测t时刻的轨迹点信息p_t,$p_t=\{x_t,y_t,s_t^1,s_t^2,s_t^3\}$,其中,$x_t$和$y_t$表示$t$时刻的位置坐标,$s_t^1,s_t^2,s_t^3$表示$t$时刻笔尖的状态,3种状态对应的意义为:“笔尖正与纸面接触”,“当前笔画书写完毕,临时抬笔”和“所有笔画书写完毕”,最终,$P=\{p_1,p_2,\dots,p_N\}$为预测的文字书写轨迹序列。

4.根据权利要求3所述的一种基于双流编码的文字图像书写轨迹恢复方法,其特征在于,所述的联合训练双流编码网络和解码网络的过程中,编解码网络损失函数为:

$$L=\lambda_1 L_{l2}+\lambda_2 L_{ce}+\lambda_3 L_{dtw}$$

λ_1, λ_2, λ_3为平衡各个损失权重的预设常数,L_{l2}为L2损失,计算公式为:

$$L_{l2}=\frac{1}{\mathrm{N}}\sum_{i=1}^{\mathrm{N}}\sqrt[2]{(x_i-\hat{x}_i)^2+(y_i-\hat{y}_i)^2}$$

其中,x_i和y_i分别为所述解码网络对位置的X坐标和Y坐标的预测值,$\hat{x}_i$和$\hat{y}_i$分别为位置的X坐标和Y坐标的标签值,N为轨迹点的数目;

L_{ce}为交叉熵损失,计算公式为:

$$L_{ce}=\sum_{i=1}^{N}\sum_{k=1}^{3}\hat{s}_i^k\log\left(p(s_i^k)\right)$$

其中,$p(s_i^k)$为所述解码网络对笔尖状态s_i^k的概率预测值,$\hat{s}_i^k$为笔尖状态的标签值;

L_{dtw}为动态时间规整损失,使用动态时间规整算法寻找预测和标签轨迹序列之间的最优对齐路径,计算在最优对齐路径下的序列距离作为预测序列的全局损失:

给定预测轨迹序列$P=\{p_1,p_2,\dots,p_N\}$和标签轨迹序列$\hat{P}=\{\hat{p}_1,\hat{p}_2,\dots,\hat{p}_M\}$,序列长度分别为$N$和$M$,设置欧式距离函数$\delta(p_i,\hat{p}_j)$用于表征轨迹点$p_i$和$\hat{p}_j$的距离,定义对齐路径$\phi=\{\phi(1),\dots,\phi(T)\}$,其中,$T\le M+N$,$T$为对齐路径的长度,对齐路径的每一项定义了$p_{\phi_p(k)}$和$\hat{p}_{\phi_{\hat{p}}(k)}$的对应关系:

$$\phi(k)=\left(\phi_p(k),\phi_{\hat{p}}(k)\right)$$

$$\phi_p(k)\in\{1,\dots,N\}$$

$$\phi_{\hat{p}}(k) \in \{1, \ldots, M\}$$

其中，$p_{\phi_p(k)}$表示P的第$\phi_p(k)$个轨迹点，$\hat{p}_{\phi_{\hat{p}}(k)}$表示$\hat{P}$的第$\phi_{\hat{p}}(k)$个轨迹点，使用动态时间规整DTW算法寻找使得序列距离最小的对齐路径，作为最优对齐路径，对应的序列距离作为预测序列的全局损失：

$$L_{dtw} = \mathrm{DTW}(P, \hat{P}) = \min_{\phi} \left\{ \sum_{k=1}^{T} \delta \left(p_{\phi_p(k)}, \hat{p}_{\phi_{\hat{p}}(k)} \right) \right\}。$$

5.根据权利要求1所述的一种基于双流编码的文字图像书写轨迹恢复方法，其特征在于，使用双流编码网络中的BiLSTM编码器的隐藏层状态，作为LSTM解码器的隐藏层初始状态h_0。

6.根据权利要求4所述的一种基于双流编码的文字图像书写轨迹恢复方法，其特征在于，λ_1取值0.5，λ_2取值1.0，λ_3取值1/6000。

一种基于双流编码的文字图像书写轨迹恢复方法

技术领域

[0001] 本发明涉及文字图像模式识别领域，尤其涉及一种基于双流编码的文字图像书写轨迹恢复方法。

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] 书写轨迹恢复算法的挑战首先来源于字形结构的复杂性。以汉字为例，国标GB18030 收纳的汉字数目多达7万余个，其中不乏结构复杂或是容易类间混淆的字形，恢复模型稍有错误就可能会导致文字模糊，类别错乱，或是生成无意义的字形。其次恢复算法不仅要克服字形结构的复杂性，同时还要能学习笔尖在空间上的位置分布以及不同笔迹点之间的先后顺序(汉字的笔顺)。所以一般而言，文字书写轨迹生成任务比普通的文字图像生成更加困难。此外，由于书写轨迹恢复任务跨越了文字的图像和轨迹序列模态，要综合考虑这两种模态的特性以及他们之间复杂的映射关系，这使得轨迹恢复算法的设计存在巨大挑战。

发明内容

[0004] 有鉴于此，本发明的目的在于提供一种基于双流编码的文字图像书写轨迹恢复方法，以解决现有技术在进行文字图像书写轨迹恢复时所存在的特征表征能力差、泛化性能较弱、书写轨迹恢复精度低的问题。

[0005] 本发明公开了一种基于双流编码的文字图像书写轨迹恢复方法，包括以下步骤：

[0006] 步骤1，将文字图像调整为预设尺寸并进行二值化处理；

[0007] 步骤2，构建双流编码网络，所述的双流编码网络输入为文字图像，输出为双流融合编码特征Z；

[0008] 步骤3，构建解码网络，所述的解码网络的输入为双流融合编码特征Z，输出为预测文字书写轨迹序列；

[0009] 步骤4，联合训练双流编码网络和解码网络，获得文字图像书写轨迹恢复网络模型；

[0010] 步骤5，利用训练完成的文字图像书写轨迹恢复网络模型进行书写轨迹恢复；

[0011] 具体地，所述的双流编码网络包括竖直卷积循环神经网络、水平卷积循环神经网络和注意力模块；

[0012] 并行连接的竖直卷积循环神经网络和水平卷积循环神经网络均包含CNN编码器和BiLSTM编码器，竖直卷积循环神经网络中的CNN编码器利用竖直方向上的降采样操作来进行竖直方向上的降采样，然后配合卷积操作，对输入文字图像进行编码，得到文字在水平方向上的一维方向特征C_x，一维方向特征C_x在方向维度上拆分后得到以方向为时序的特征序列，竖直卷积循环神经网络中的BiLSTM编码器对时序的特征序列进行编码，得到双流编码特征V_x；水平卷积循环神经网络中的CNN编码器利用水平方向上的降采样操作来进行水平方向上的降采样，然后配合卷积操作，对输入文字图像进行编码，得到文字在竖直方向上的一维方向特征C_y，一维方向特征C_y在方向维度上拆分后得到以方向为时序的特征序列，水平卷积循环神经网络中的BiLSTM编码器对时序的特征序列进行编码，得到双流编码特征V_y；

[0013] 在注意力模块中对双流编码特征V_x和V_y进行融合，得到双流融合编码特征Z：

[0014] $$Z = \sum_{i=1}^{|V|} v_i * w_i,$$

[0015] 其中，$w_i = \frac{e^{s_i}}{\sum_{j=1}^{|V|} e^{s_j}}$，$s_i = f(v_i) = Uv_i$，通过合并特征$V_x$和$V_y$得到$V$，$v_i$为$V$的分量，$|V|$为$V$的长度，$U$为一个全连接层的可学习参数。

[0016] 可选地，所述的降采样操作为非对称池化操作、非对称卷积操作或全连接层网络操作降采样；

[0017] 可选地，所述的解码网络为LSTM解码器，LSTM解码器以双流融合编码特征Z为输入，依次预测轨迹点；LSTM解码器根据$t-1$时刻的预测值p_{t-1}和隐藏层向量h_{t-1}，预测t时刻的轨迹点信息p_t，$p_t = \{x_t, y_t, s_t^1, s_t^2, s_t^3\}$，其中，$x_t$和$y_t$表示$t$时刻的位置坐标，$s_t^1, s_t^2, s_t^3$表示$t$时刻笔尖的状态，3种状态对应的意义为：“笔尖正与纸面接触”，“当前笔画书写完毕，临时抬笔”和“所有笔画书写完毕”，最终，$P = \{p_1, p_2, \ldots, p_N\}$为预测的文字书写轨迹序列。

[0018] 具体地，所述的联合训练双流编码网络和解码网络的过程中，编解码网络损失函数为：

[0019] $$L = \lambda_1 L_{l2} + \lambda_2 L_{ce} + \lambda_3 L_{dtw}$$

[0020] λ_1，λ_2，λ_3为平衡各个损失权重的预设常数，L_{l2}为L2损失，计算公式

为：

[0021] $$L_{l2}=\frac{1}{\mathrm{N}}\sum_{i=1}^{\mathrm{N}}\sqrt[2]{(x_i-\hat{x}_i)^2+(y_i-\hat{y}_i)^2}$$

[0022] 其中，x_i和y_i分别为所述解码网络对位置的X坐标和Y坐标的预测值，$\hat{x}_i$和$\hat{y}_i$分别为位置的X坐标和Y坐标的标签值，N为轨迹点的数目；

[0023] L_{ce}为交叉熵损失，计算公式为：

[0024] $$L_{ce}=\sum_{i=1}^{N}\sum_{k=1}^{3}\hat{s}_i^k\log\left(p(s_i^k)\right)$$

[0025] 其中，$p\left(s_i^k\right)$为所述解码网络对笔尖状态s_i^k的概率预测值，$\hat{s}_i^k$为笔尖状态的标签值；

[0026] L_{dtw}为动态时间规整损失，使用动态时间规整算法寻找预测和标签轨迹序列之间的最优对齐路径，计算在最优对齐路径下的序列距离作为预测序列的全局损失：

[0027] 给定预测轨迹序列$P=\{p_1,p_2,\ldots,p_N\}$和标签轨迹序列$\hat{P}=\{\hat{p}_1,\hat{p}_2,\ldots,\hat{p}_M\}$，序列长度分别为$N$和$M$，设置欧式距离函数$\delta\left(p_i,\hat{p}_j\right)$用于表征轨迹点$p_i$和$\hat{p}_j$的距离，定义对齐路径$\phi=\{\phi(1),\ldots,\phi(T)\}$，其中，$T\leq M+N$，$T$为对齐路径的长度，对齐路径的每一项定义了$p_{\phi_p(k)}$和$\hat{p}_{\phi_{\hat{p}}(k)}$的对应关系：

[0028] $$\phi(k)=\left(\phi_p(k),\phi_{\hat{p}}(k)\right)$$

[0029] $$\phi_p(k)\in\{1,\ldots,N\}$$

[0030] $$\phi_{\hat{p}}(k)\in\{1,\ldots,M\}$$

[0031] 其中，$p_{\phi_p(k)}$表示P的第$\phi_p(k)$个轨迹点，$\hat{p}_{\phi_{\hat{p}}(k)}$表示$\hat{P}$的第$\phi_{\hat{p}}(k)$个轨迹点，使用动态时间规整DTW算法寻找使得序列距离最小的对齐路径，作为最优对齐路径，其对应的序列距离作为预测序列的全局损失：

[0032] $$L_{dtw} = \mathrm{DTW}(P,\hat{P}) = \min_{\phi}\left\{\sum_{k=1}^{T}\delta\left(p_{\phi_p(k)},\hat{p}_{\phi_{\hat{p}}(k)}\right)\right\}$$

[0033] 优选地，使用双流编码网络中的BiLSTM编码器的隐藏层状态，作为LSTM解码器的隐藏层初始状态h_0。

[0034] 优选地，λ_1取值0.5，λ_2取值1.0，λ_3取值1/6000。

[0035] 本发明方法相对于现有技术而言，编码过程中，分别提取文字在竖直和水平方向上的特征，实现对特征进行降采样、减少参数量的同时，保留必要的文字字形信息，帮助后续的解码精确地反映文字的字形，最终能有效提升文字图像书写轨迹的恢复性能。

附图说明

[0036] 图1示出了本发明实施方法的流程示意图；

[0037] 图2示出了本发明实施例中双流编码网络的结构示意图；

[0038] 图3示出了本发明实施例中解码网络的结构示意图。

具体实施方式

[0039] 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。

[0040] 为了引用和清楚起见，下文中使用的技术名词、简写或缩写总结解释如下：

[0041] CNN:Convolutional Neural Network, 卷积神经网络；

[0042] RNN:Recurrent Neural Network,循环神经网络；

[0043] CRNN:Convolutional Recurrent Neural Network,卷积循环神经网络；

[0044] BiLSTM:Bi-directional Long Short-Term Memory,双向长短时记忆模型；

[0045] DTW:Dynamic Time Warping,动态时间规整。

[0046] 图1示出了本发明实施例的流程示意图。一种基于双流编码的文字图像书写轨迹恢复方法，包括以下步骤：

[0047] 步骤1,将文字图像调整为预设尺寸并进行二值化处理；

[0048] 步骤2,构建双流编码网络，所述的双流编码网络输入为文字图像，输出为双流融合编码特征Z；

[0049] 步骤3,构建解码网络，所述的解码网络的输入为双流融合编码特征Z，输出为预测文字书写轨迹序列；

[0050] 步骤4,联合训练双流编码网络和解码网络，获得文字图像书写轨迹恢复网络模型；

[0051] 步骤5,利用训练完成的文字图像书写轨迹恢复网络模型进行书写轨迹恢复。

[0052] 本实施例的具体操作步骤如下：

[0053] (1)对输入文字图像进行预处理操作:保持宽高比的同时将尺寸调整为$64X64$;进行二值化处理。

[0054] (2)构建双流编码网络。

[0055] 1)如图2所示,构建两个卷积循环神经网络(Convolutional Recurrent Neural Network,CRNN)分支$CRNN_X$和$CRNN_Y$。它们分别包含CNN编码器和BiLSTM编码器。两个CNN编码器分别利用竖直或水平方向上的非对称池化操作来进行竖直或水平方向上的降采样,配合卷积操作,分别对输入文字图像进行编码,得到文字在竖直或水平方向上的一维方向特征C_x和C_y。一维方向特征C_x和C_y在方向维度上拆分后可得到以方向为时序的特征序列,BiLSTM编码器对该时序特征序列进行编码,得到双流编码特征V_x和V_y。

[0056] 2)基于注意力机制对两种特征V_x和V_y进行融合,得到双流融合编码特征Z：

[0057] $$s_i = f(v_i) = Uv_i,$$

[0058] $$w_i = \frac{e^{s_i}}{\sum_{j=1}^{|V|} e^{s_j}},$$

[0059] $$Z = \sum_{i=1}^{|V|} v_i * w_i,$$

[0060] 其中,合并特征V_x和V_y得到V,v_i和v_j为V的第i个分量和第j个分量,s_i表示v_i的注意力权重,s_j表示v_j的注意力权重,$f(\cdot)$表示一个全连接层的函数,$|V|$为V的长度,U为一个全连接层的可学习参数。

[0061] (3)构建解码网络,进行特征解码,输出预测文字书写轨迹序列。

[0062] 1)构建LSTM解码器,LSTM解码器以双流融合编码特征Z为输入,依次预测轨迹点。如图3所示,LSTM解码器根据$t-1$时刻的预测值p_{t-1}和隐藏层向量h_{t-1},预测t时刻的轨迹点信息p_t。最终,$P = \{p_1, p_2, \ldots, p_N\}$为预测的文字书写轨迹序列。使用双流编码网络中的BiLSTM编码器的隐藏层状态,作为LSTM解码器的隐藏层初始状态h_0。

[0063] 2)对于t时刻的轨迹点信息,设置$p_t = \{x_t, y_t, s_t^1, s_t^2, s_t^3\}$。其中,$x_t$和$y_t$表示该时刻的位置坐标,$s_t^1, s_t^2, s_t^3$表示该时刻笔尖的状态,构成热码$\{1,0,0\}$,$\{0,1,0\}$和$\{0,0,1\}$,分别表示书写过程中的3种状态:“笔尖正与纸面接触”,“当前笔画书写完毕,临时抬笔”和“所有笔画书写完毕”。特别的,设置初始输入轨迹点

为$p_0 = \{0,0,1,0,0\}$。

[0064] (4)构建编解码网络损失函数,端到端(end-to-end)地训练双流编码网络和解码网络构成的模型。设置的编解码网络损失函数包括L2 损失,交叉熵损失和动态时间规整损失。

[0065] L2损失:

[0066] $$L_{l2} = \frac{1}{\mathrm{N}}\sum_{i=1}^{\mathrm{N}} \sqrt[2]{(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2}$$

[0067] 其中,x_i和y_i为网络的预测值,$\hat{x}_i$和$\hat{y}_i$为标签值,N为轨迹点的数目。

[0068] 交叉熵损失(CrossEntropy Loss):

[0069] $$L_{ce} = \sum_{i=1}^{N}\sum_{k=1}^{3} \hat{s}_i^k \log\left(p\left(s_i^k\right)\right)$$

[0070] 其中,s_i^k为网络的预测值,$\hat{s}_i^k$为标签值。

[0071] 动态时间规整损失(DynamicTimeWarpingLoss):使用动态时间规整算法寻找预测和标签轨迹序列之间的最优对齐路径,计算在最优对齐路径下的序列距离作为预测序列的全局损失,从而实现轨迹序列的全局优化。

[0072] 给定预测轨迹序列$P = \{p_1, p_2, \ldots, p_N\}$和标签轨迹序列$\hat{P} = \{\hat{p}_1, \hat{p}_2, \ldots, \hat{p}_M\}$,序列长度分别为$N$和$M$,设置欧式距离函数$\delta\left(p_i, \hat{p}_j\right)$用于表征轨迹点$p_i$和$\hat{p}_j$的距离,定义对齐路径$\phi = \{\phi(1), \ldots, \phi(T)\}$(其中$T \leq M + N$为对齐路径的长度),对齐路径的每一项定义了$p_{\phi_p(k)}$和$\hat{p}_{\phi_{\hat{p}}(k)}$的对应关系:

[0073] $$\phi(k) = \left(\phi_p(k), \phi_{\hat{p}}(k)\right)$$

[0074] $$\phi_p(k) \in \{1, \ldots, N\}$$

[0075] $$\phi_{\hat{p}}(k) \in \{1, \ldots, M\}$$

[0076] 使用动态时间规整(DTW)算法寻找使得序列距离最小的对齐路径,作为最优对齐路径,其对应的序列距离作为预测序列的全局损失:

[0077] $$L_{dtw} = \mathrm{DTW}(P, \hat{P}) = \min_{\phi} \left\{ \sum_{k=1}^{T} \delta \left(p_{\phi_p(k)}, \hat{p}_{\phi_{\hat{p}}(k)} \right) \right\}$$

[0078] 编解码网络损失函数：

[0079] $$L = \lambda_1 L_{l2} + \lambda_2 L_{ce} + \lambda_3 L_{dtw}$$

[0080] λ_1，λ_2，λ_3为平衡各个损失权重的常数。在实施中，我们分别设置λ_1，λ_2，λ_3为0.5,1.0和1/6000。

[0081] 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。

[0082] 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。

[0083] 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。

[0084] 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。

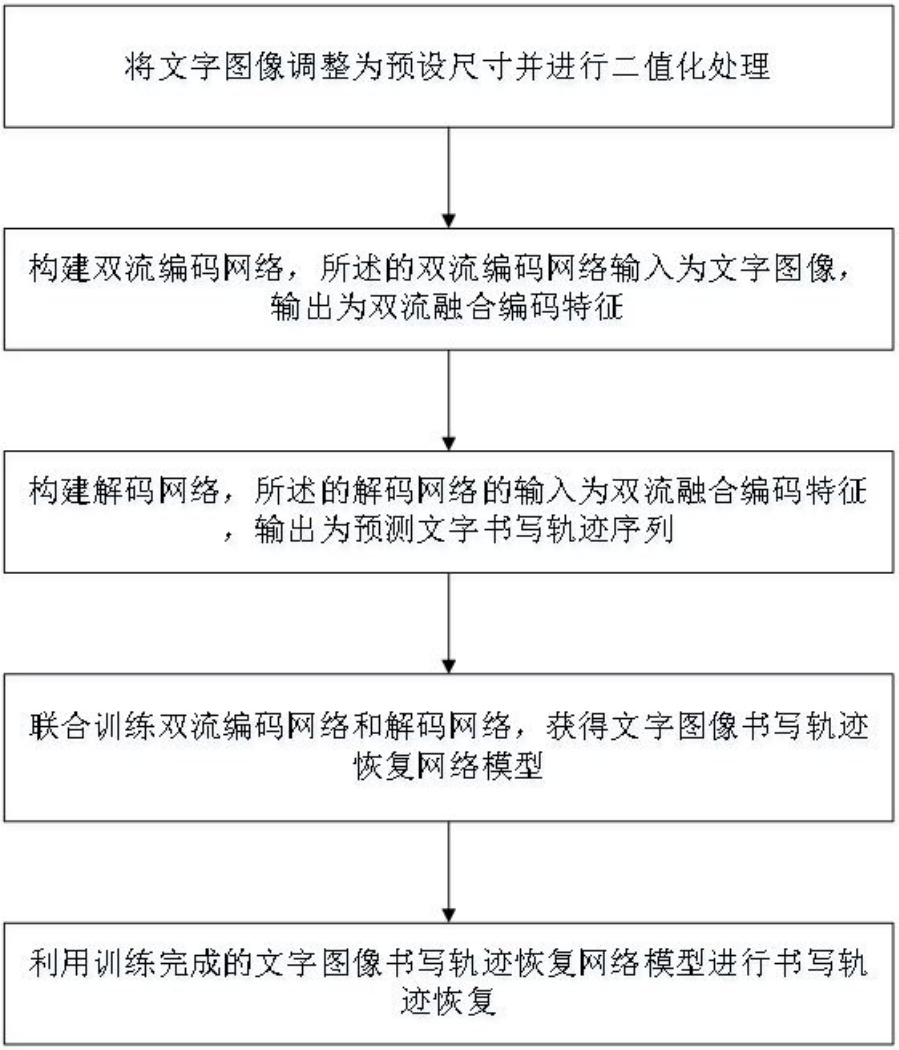
将文字图像调整为预设尺寸并进行二值化处理
构建双流编码网络，所述的双流编码网络输入为文字图像，输出为双流融合编码特征
构建解码网络，所述的解码网络的输入为双流融合编码特征，输出为预测文字书写轨迹序列
联合训练双流编码网络和解码网络，获得文字图像书写轨迹恢复网络模型
利用训练完成的文字图像书写轨迹恢复网络模型进行书写轨迹恢复

图1

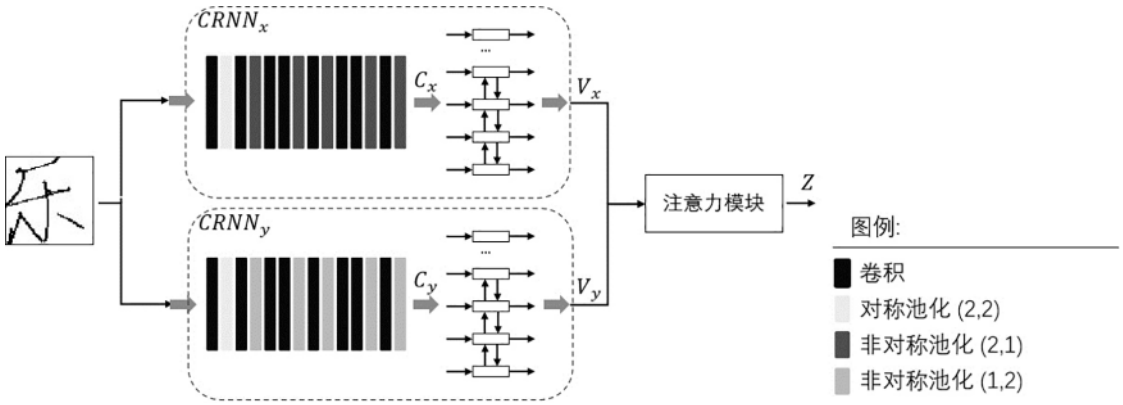
$CRNN_x$
C_x
V_x
$CRNN_y$
C_y
V_y
注意力模块
Z
图例:
卷积
对称池化 (2,2)
非对称池化 (2,1)
非对称池化 (1,2)

图2

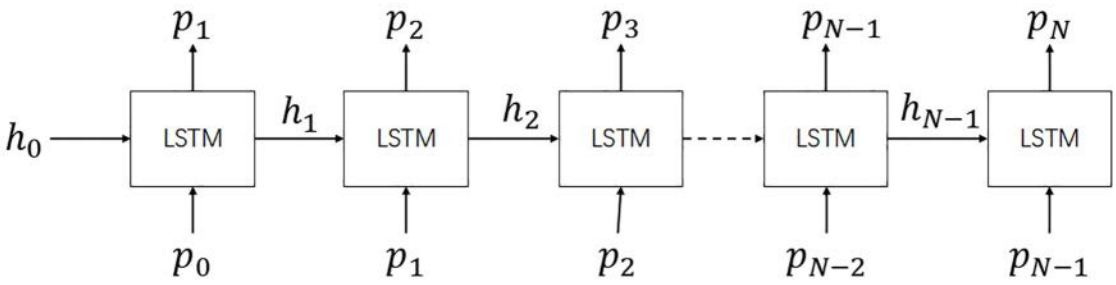
p_1
p_2
p_3
p_{N-1}
p_N
h_0
LSTM
h_1
LSTM
h_2
LSTM
LSTM
h_{N-1}
LSTM
p_0
p_1
p_2
p_{N-2}
p_{N-1}

图3