

基于多特征提取的手写汉字识别方法

崔智 崔盼盼

商丘工学院

[摘要] 手写汉字以为每个人字体不同, 风格多变而识别困难, 为解决该问题提出一种基于多特征提取的手写汉字识别方法。该方法首先通过计算Hu矩进行粗分类, 然后采用霍夫空间变换、Harris角点检测多种特征提取进行识别的算法。

[关键词] 手写汉字; 霍夫变换; Harris角点检测; Hu矩检测

[DOI] 10.12252/j.issn.2096-6261.2021.11.766

一、概述

日常生活中存在海量的手写汉字信息, 为减轻人们工作, 手写汉字识别在各个领域中得到应用。但是手写汉字风格迥异, 汉字数量庞大的特点提高了识别技术的难度。市场上手写输入法软件识别率低于90%^[1]。刘庆^[2]等人提出一种基于非监督预训练的结构优化CNN用于手写汉字识别。王来兵^[3]提出了基于贝叶斯与生成对抗网络的手写汉字识别算法。但是目前手写汉字数据集较少, 为取得良好的训练效果需要自建大样本数据集。而手写汉字的识别关键在于特征提取, 姚娜^[4]等人提出一种改进的Harris算法将汉字图像角点作为特征进行提取, 为确定线段位置和计算线段长度提供基础。叶峰^[5]等人提出一种部分级联特征识别离线手写体汉字。文生平等^[6]提出了一种改进Hu矩检测算法进行标识符识别算法。XuYao Zhang^[7]等人提出一种讲传统归一化协同方向分解特征图与深度卷积神经网络相结合的方式进行在线与离线手写字识别算法。Benatallah^[8]等人研究表明汉字可分解为“横”“竖”“捺”“撇”, 本文提出通过Harris角点检测, 然后采用霍夫空间变换统计笔画所占比例, 通过角点检测结果对其进行加权而得到融合的特征, 进而识别手写体汉字。

二、预处理

为增强汉字与背景反差, 首先通过Otsu算法进行二值化, 使汉字深浅一致, 然后采用Canny算子进行边缘检测, 后续特征均是根据边缘开展, 与手写汉字笔画的粗细无关。

三、Hu矩特征提取

Hu矩^[9]具有图像旋转、平移、伸缩不变性, 所以Hu是识别图像的重要特征之一。设灰度图像的分布为 $I(x, y)$, 其 $(p+q)$ 阶矩定义如式(1)所示:

$$m_{pq} = \sum_{x=1}^M \sum_{y=1}^N x^p y^q I(x, y) \quad (1)$$

其中, $p, q = 0, 1, 2, \dots$ 。 $(p+q)$ 阶中心矩定义如式(2)所示:

$$\mu_{pq} = \sum_{x=1}^M \sum_{y=1}^N (x-x_0)^p (y-y_0)^q f(x, y) \quad (2)$$

其中, 归一化中心距 (x_0, y_0) 为:

$$y_{pq} = \frac{\mu_{pq}}{\mu_{00}^r} \quad (3)$$

其中, $r = \frac{p+q+2}{2}, p+q = 2, 3, \dots$

Hu矩利用图像二阶矩和三阶中心矩构造的七个不变矩,

具体定义如下所示:

$$I_1 = y_{20} + y_{02} \quad (4)$$

$$I_2 = (y_{20} + y_{02})^2 + 4y_{11}^2 \quad (5)$$

$$I_3 = (y_{30} + 3y_{03})^2 + (3y_{21} - y_{03})^2 \quad (6)$$

$$I_4 = (y_{30} + y_{12})^2 + (y_{21} + y_{03})^2 \quad (7)$$

$$I_5 = (y_{30} - y_{12})(y_{30} + y_{12})[(y_{30} + y_{12})^2 - 3(y_{21} + y_{03})^2] + (3y_{21} - y_{03})(y_{21} + y_{03})[3(y_{30} + y_{12})^2 - (y_{21} + y_{03})^2] \quad (8)$$

$$I_6 = (y_{20} - y_{02})[(y_{30} + y_{12})^2 - (y_{21} + y_{03})^2] + 4y_{11}(y_{30} + y_{12})(y_{21} + y_{03}) \quad (9)$$

$$I_7 = (3y_{21} + y_{03})(y_{30} + y_{12})[(y_{30} + y_{12})^2 - 3(y_{21} + y_{03})^2] + (y_{30} - 3y_{12})(y_{21} + y_{03})[3(y_{30} + y_{12})^2 - (y_{21} + y_{03})^2] \quad (10)$$

通过上式计算得到手写汉字Hu矩特征属性, 通过计算多幅图像的Hu矩计算均值, 作为模板Hu矩, 测试图像的Hu矩特征将与其进行比较。

四、Harris角点检测

Harris角点检测的核心是设置滑窗在图像上移动, 判断局部灰度是否发生较大变化。假设图像I为灰度图, $I(x, y)$ 表示坐标 (x, y) 点的灰度值, 取大小为 $u \times v$ 的窗口滑动, 设窗口函数为高斯分布, 如果窗口中心点是角点, 那么窗口移动后中心点像素值会产生较大波动。则窗口滑动产生的灰度值变换的平方误差累计加权和 $E(u, v)$ 为:

$$E(u, v) = \sum_{x,y} w(x, y) [I(x+u, y+v) - I(x, y)]^2 \quad (11)$$

其中, $w(x, y) = \exp(-\frac{u^2+v^2}{2\sigma^2})$, 为窗口函数, $I(x+u, y+v)$ 为移动后图像灰度, 可通过泰勒级数展开进行近似, 表达式如下:

$$I(x+u, y+v) \approx I(x, y) + uI_x + vI_y \quad (12)$$

其中, I_x 和 I_y 为 I 的一阶偏导数。则式(1)可以化简为:

$$E(u, v) \approx \sum_{x,y} w(x, y) (uI_x + vI_y)^2 \quad (13)$$

将式(3)改写成矩阵形式:

$$E(u, v) \approx (u, v)A \begin{bmatrix} u \\ v \end{bmatrix} \quad (14)$$

其中, A 是自相关矩阵:

$$A = \sum_{(x,y)} w(x, y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \quad (15)$$

将A对角化后可得A特征值, 通过分析特征值可以总结出当前点是否为角点。通过Harris算法检测手写汉字角点漏检率低, 但是伪角点较多, 因此Harris角点检测不能作为单一特征去识别汉字。

五、霍夫变换

霍夫变换是1959年由Paul Hough提出的一种几何特征检测算法^[10]。直线在图像空间域可表示为 $y=kx+b$, 其中 k 为斜

率, b 为截距。霍夫变换基于对偶性将, 在 k - b 参数空间中, 使用 (k, b) 一个点表示空间中的一条直线。但是当 k 无限大时, k - b 参数空间难以表达, 所以采用 r - θ 参数空间表示, 如图1所示:

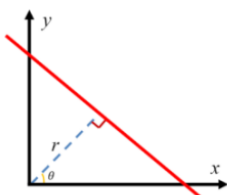


图1 霍夫变换示意图

$$y = \left(-\frac{\cos\theta}{\sin\theta}\right)x + \frac{r}{\sin\theta} \quad (16)$$

式(6)变形为:

$$r = x\cos\theta + y\sin\theta \quad (17)$$

这样通过霍夫变换可以将空间域中的直线映射到参数空间中的点 (r, θ) , 其中 $\theta \in [-\pi, \pi]$ 。本文对汉字中线段进行检测, 保留线段与水平轴夹角 α 、线段的起点与终点坐标, 与角点特征融合, 通过计算找到与角点距离最近的线段起点与终点, 舍弃距离大于阈值的线段, 然后通过 α 将线段分为“横”、“竖”、“撇”、“捺”, 并统计各种笔画所占比例。

六、实验结果及分析

本实验基于CASIA-HWDB1.1数据库^[11], 该数据库由300人手写, 包含1176000张图像。与“阿”字为例, 对“阿”字图像提取Hu矩特征, 计算结果如下:

$$I_1=0.6181 \quad I_2=0.0196 \quad I_3=0.0014 \quad I_4=0.003 \quad I_5=0 \\ I_6=0.0004 \quad I_7=0$$

对其他测试图片计算Hu矩特征, 与模板Hu矩比较, 计算距离, 进行粗分类。

然后通过Harris算法进行角点检测, 角点检测结果如图2所示:

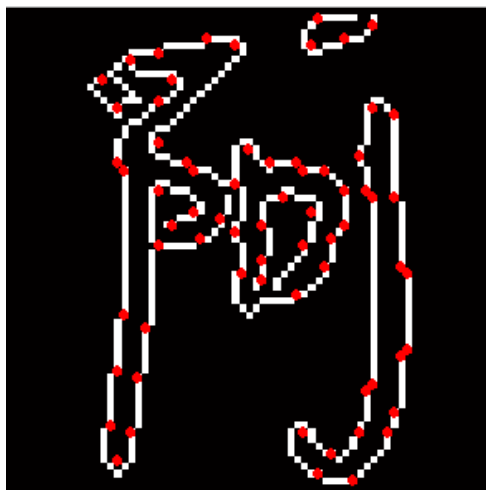


图2 角点检测结果

最后通过霍夫变换线段检测后与角点特征融合, 检测笔画结果如图3所示:

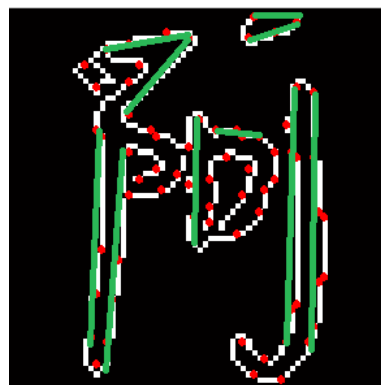


图3 笔画检测结果

结语

本文提出一种基于角点与线段融合特征的识别方式, 主要是为了统计出汉字中“横”“竖”“撇”“捺”所占比例。相比较与传统梯度模板算法, 准确率更高。

参考文献:

- [1] Qiu L, Jin L, Dai R, et al. An Open Source Testing Tool for Evaluating Handwriting Input Methods[J]. IEEE Computer Society, 2015.
- [2] 刘庆, 唐贤伦, 张娜. 基于非监督预训练的结构优化卷积神经网络[J]. 工程科学与技术, 2017, 49(S2): 210-215.
- [3] 王来兵. 基于贝叶斯与生成式对抗网络的手写字文本识别算法[J]. 黑龙江工业学院学报(综合版), 2019, 19(08): 31-35.
- [4] 姚娜, 吕海芳, 白铁成. 基于改进Harris算法的汉字图像角点检测[J]. 计算机与现代化, 2014(05): 99-103+117.
- [5] 叶锋, 邓衍晨, 汪敏, 廖茜, 郑子华, 林晖. 部分级联特征的离线手写体汉字识别方法[J]. 计算机系统应用, 2017, 26(08): 134-140.
- [6] 文生平, 冯泽锋, 洪培烽, 张施华. 基于改进Hu矩算法的AGV字符识别研究[J]. 计算机测量与控制, 2020, 28(05): 229-232+251.
- [7] Zhang X Y, Bengio Y, Liu C L. Online and Offline Handwritten Chinese Character Recognition: A Comprehensive Study and New Benchmark[J]. 2016.
- [8] Benattallah B, Dumas M, Fauvet M C, et al. Overview of some patterns for architecting and managing composite web services[J]. Acm Sigecom Exchanges, 2002.
- [9] 张伟, 何金国. Hu不变矩的构造与推广[J]. 计算机应用, 2010, 30(09): 2449-2452.
- [10] Mochizuki Y, Torii A, Imiya A. N-Point Hough transform for line detection[M]. Academic Press, Inc. 2009.
- [11] Wang D H, Liu C L, Yu J L, et al. CASIA-OLHWDB1: A Database of Online Handwritten Chinese Characters[C]// 2009: 1206-1210.