

Improved FCM Handwritten Digit Recognition Based on Zernike Moments Feature Extraction

Chunyan Miao, Yaoquan Yang, Shuo Zhang, Shenghui Han

School of Control Science and Engineering, North China Electric Power University, Baoding
Email: miaowenwen1988@163.com, yyq2201@163.com, shuoshuo316@126.com, 15831272018@163.com

Received: Mar. 27th, 2013; revised: Apr. 21st, 2013; accepted: May 2nd, 2013

Copyright © 2013 Chunyan Miao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract: This paper proposes a method of Zernike moments feature extraction, which is through the improved FCM algorithm for fuzzy clustering of handwritten numerals. Zernike moments of the image have rotational invariance, so the feature space can well reflect the characteristics of the image. Fuzzy C-Means Clustering uses a degree of membership of each data point to determine the degree of belonging to a cluster, and this paper takes weighted fuzzy C-means clustering algorithm to classify the membership. Clustering lies in feature extraction. In the image pre-processing of the premise, this paper reduced the dimension of Zernike moments, and effectively improved the recognition speed without decreasing the recognition rate.

Keywords: Pattern Recognition; Feature Extraction; Zernike Moments; Handwritten Digit Recognition; Improved FCM

基于 Zernike 矩特征提取的改进 FCM 手写体数字识别

苗春艳, 杨耀权, 张 硕, 韩升晖

华北电力大学控制科学与工程学院, 保定
Email: miaowenwen1988@163.com, yyq2201@163.com, shuoshuo316@126.com, 15831272018@163.com

收稿日期: 2013 年 3 月 27 日; 修回日期: 2013 年 4 月 21 日; 录用日期: 2013 年 5 月 2 日

摘 要: 文中提出了采用 Zernike 矩进行特征提取, 并通过改进 FCM 算法对手写体数字进行模糊聚类的一种方法。图像的 Zernike 矩具有旋转不变性, 因此用它构造的特征空间能很好的反映图像的特性。模糊 C 均值聚类是用隶属度确定每个数据点属于某个聚类的程度的一种聚类算法, 本文采取加权的模糊 C 均值聚类算法来进行分类。聚类的关键在于特征提取, 在图像预处理的前提下, 本文将 Zernike 矩进行降维处理, 在不降低识别率的基础上, 有效提高了识别速度。

关键词: 模式识别; 特征提取; Zernike 矩; 手写体数字识别; 改进模糊 C 均值聚类

1. 引言

数字识别是光学字符识别(OCR)的一个重要分支^[1], 而手写体数字识别更是多年来的研究热点。当前运用较好的主流算法还是以统计、神经网络和聚类分析的识别算法为主, 如 Adaboost 算法^[2]、BP 神经网络^[3]、支持向量机(SVM)^[4]、以及径向基函数(RBF)^[5]

等等。

在书写字符的过程中, 因人而异的书写方式, 造成手写体数字字符的各种变形, 尤其是对于无约束手写数字来说, 由于书写习惯以及因书写工具不同造成笔画粗细差异较大等问题, 导致数字的特征变得不稳定。为了提高识别的准确率, 本文提出了一种基于

Zernike 矩特征筛选的改进模糊聚类手写体数字识别方法。

2. Zernike 矩的定义及特征选取

2.1. Zernike 矩的定义

Zernike 在 1934 年引入了一组定义在单位圆 $x^2 + y^2 = 1$ 上的复值函数集 $\{V_{nm}(x, y)\}, \{V_{nm}^*(x, y)\}$ 具有完备性和正交性, 使得它可以定义在单位圆内的任何平方可积函数。而 Zernike 矩的概念首先由 Teague 于 1980 年引入^[6], n 阶 Zernike 矩的定义为:

$$Z_{nm} = \frac{n+1}{\pi} \sum_x \sum_y f(x, y) V_{nm}^*(\rho, \theta), x^2 + y^2 \leq 1 \quad (1-1)$$

式中, $f(x, y)$ 为图像的密度函数, *表示取共轭, Zernike 多项式由下式给出:

$$V_{nm}(x, y) = V_{nm}(\rho, \theta) = R_{nm}(\rho) \exp(jm\theta) \quad (1-2)$$

式中, n 为非负数, $|m| \leq n$ 并满足 $n - |m|$ 为偶数, ρ 为从原点到点 (x, y) 的向量的长度, θ 为 ρ 与 x 轴在逆时针方向的夹角, $R_{nm}(\rho)$ 为径向多项式, 其定义为:

$$R_{nm}(\rho) = \sum_{s=0}^{(n-|m|)/2} \frac{(-1)^s (n-s)!}{s! \left[\frac{(n+|m|)}{2} \right]! \left[\frac{(n-|m|)}{2} - s \right]!} \rho^{n-2s} \quad (1-3)$$

可以证明^[7], 一幅图像旋转 Φ 角后, 它的 Zernike 矩 Z'_{nm} 与初始图像的 Zernike 矩 Z_{nm} 有如下关系:

$$Z'_{nm} = Z_{nm} e^{-jm\phi} \quad (1-4)$$

由上式可得, 旋转后的图像 $f'(x, y)$ 的 Zernike 矩 Z'_{nm} 与原图像 $f(x, y)$ 的 Zernike 矩 Z_{nm} 只有相位上的差异, 不存在幅值上的差别, 即 Zernike 矩具有旋转不变性。

2.2. 特征选取

在模式识别中, 特征选取是一个关键问题, 其基本任务是从许多特征中找出那些最有效的特征。理想的特征应对字符的平移、旋转与缩放不敏感, 平移和缩放问题可以在字符规范化时解决, 在规范化时消除旋转造成的影响比较复杂, 因此提取对字符旋转角度不敏感的特征是一种更可靠与简便的方法。Zernike 矩是一种正交矩, 具有旋转不变性, 因此本文运用 Zernike 矩特征进行筛选, 建立具有平移、尺寸及旋转

不变性的样本特征库。

由式(1-1)可知, 根据 Z_{nm} 中 m 和 n 的变化可以计算出图像不同阶的 Zernike 矩。目前 Zernike 矩有很多个, 如果把所有 Zernike 矩做为识别因素, 则信息量过大。通过分析发现这些 Zernike 矩中含有一定的冗余信息, 这些冗余信息不会明显提高数字的识别率, 相反还降低了识别速度。因此如何把高维特征空间压缩到低维特征空间, 在不降低识别率的前提下, 提高识别速度, 是特征提取的关键步骤。

在统计学经典理论中, 方差分析是从质量因子的角度探讨因素不同水平对实验指标影响的差异。本文采用如下 3 个步骤对 Zernike 矩进行筛选, 即先采用特征选择的方法, 去掉那些明显没有分类信息的特征, 然后在采用特征提取的方法进行相关度分析以降低维数。

1) 确保类内特征离散度较小。文献[8]提出用 43 个矩来进行特征提取。本文对每个数字的 100 个样本的前 43 个 Zernike 特征进行方差分析, 求出各项 Zernike 特征的均值与方差, 以均值的 1.2 倍为阈值, 方差小的 Zernike 特征保留, 方差大的 Zernike 特征去除。

2) 确保类间特征离散度较大, 保留分类信息多的特征。将(1)筛选后的十个数字的 Zernike 特征的均值做为研究对象, 求出每个 Zernike 特征均值的方差, 同样以均值的 1.2 倍为阈值, 方差大的 Zernike 特征保留, 方差小的 Zernike 特征去除。

3) 特征降维处理, 以提高识别速度。将步骤(2)筛选后的十个数字的 Zernike 特征的均值做为十个数字的标准 Zernike 特征值, 相对于十个数字, 对每个 Zernike 特征之间进行相关度分析, 将相关度大的 Zernike 特征进行相加合并, 减少特征的冗余信息。

3. 基于特征加权的模糊 C 均值聚类算法

模糊 C 均值聚类算法^[9]的步骤还是比较简单的, 是用隶属度确定每个数据点属于某个聚类的程度的一种聚类算法。1973 年, Bezdek 提出了该算法, 作为早期硬 C 均值聚类(HCM)方法的一种改进。

FCM 把 n 个向量 $x_i (i = 1, 2, \dots, n)$ 分为 c 个模糊组, 并求每组的聚类中心, 使得非相似性指标的价值函数达到最小。FCM 与 HCM 的主要区别在于 FCM 用模糊划分, 使得每个给定数据点用值在 0, 1 间的隶属

度来确定其属于各个组的程度。与引入模糊划分相适应，隶属矩阵 U 允许有取值在 $0, 1$ 间的元素。不过，加上归一化规定，一个数据集的隶属度的和总等于 1 ：

$$\sum_i u_{ik} = 1, \forall k = 1, \dots, n \quad (2-1)$$

那么，FCM 的价值函数(或目标函数)就是：

$$J(U, c) = \sum_{k=1}^n \sum_{i=1}^c u_{ik}^m d_{ik}^2 \quad (2-2)$$

这里 u_{ik} 介于 $0, 1$ 间； c_i 为模糊组 i 的聚类中心， d_{ik} 表示第 i 类中的样本 x_k 与第 i 类的典型样本 c_i 之间的失真度，经常用两个矢量间的距离来度量。且 $m \in [1, \infty)$ 是一个加权指数，控制着算法的模糊程度。聚类的准则为取 $J(U, c)$ 的极小值 $\min\{J(U, c)\}$ 。

在传统的 FCM 算法中，每维特征对分类的贡献是均匀的。而在实际应用中，由于构成样本特征矢量的各维特征来自不同的传感器，存在量纲差异、精度及可靠性的不同，另一方面，所选择的特征集未必适合于模式的分类。因此，传统的 FCM 算法在实际应用中有一定的局限性，我们在聚类分析中必须考虑各维特征的不同影响。

为此，本文采取加权的模糊 C 均值聚类算法来进行分类。加权模糊 C 均值聚类可以表示为如下的规划问题^[10]：

$$\min J(U, c) = \sum_{k=1}^n \sum_{i=1}^c (u_{ik})^m (\sqrt{\omega_j} d_{ik})^2 \quad (2-3)$$

式中，

$$0 \leq u_{ik} \leq 1, 1 \leq i \leq c, 1 \leq k \leq n;$$

$$\sum_{i=1}^c u_{ik} = 1, 1 \leq k \leq n$$

$$0 < \sum_{k=1}^n u_{ik} < n, 1 \leq i \leq c;$$

$$1 \leq m < \infty$$

ω_j 通过以下方法计算得到：

- 1) 将原始数据归一化，得到无量纲矩阵 Y ；
- 2) 计算矩阵 Y 的相关系数矩阵 R ；
- 3) 计算相关系数矩阵 R 的特征值 λ_j ；
- 4) 将特征值归一化得到权向量

$$\omega = (\omega_1, \omega_2, \dots, \omega_k), \omega_j = \frac{\lambda_j}{\sum_{j=1}^k \lambda_j}。$$

4. 手写体数字识别实例

4.1. 手写数字预处理

预处理的主要目的是对一个给定的含有字符的图像，突出图像中与字符有关的某些信息，削弱或去除某些不需要的信息使它的结果对后面的识别来说比原始的字符图像更适合^[11-13]。预处理一般包括二值化、平滑、去噪声、规范化和细化等过程。由于 Zernike 矩具有旋转不变性，在字符规范化时，只需进行平移和缩放处理。如图 1 所示为手写数字图像预处理结果，主要处理过程包括位置归一化、细化和去毛刺。

4.2. 基于 Zernike 矩特征的手写体数字识别改进 FCM 聚类

从加权的改进模糊 C 均值的目标函数的角度来说，一般采用的是加权欧式距离^[14]，而该距离要求各分量相互独立。Zernike 矩的正交性正好满足了其要求。式(1-2)提出了一组正交多项式 $\{V_{nm}(x, y)\}$ ，图像的 Zernike 矩就是该图像在正交多项式上的投影。为了计算一幅给定图像的 Zernike 矩，必须将图像的中心移到坐标原点，将图像像素点映射到单位圆内。

本文运用 Matlab 进行仿真，以下为计算 Zernike 矩的几个关键点：

- 1) `circle = x2 + y2; inside = find(circle ≤ 1);` % 提取符合 `circle ≤ 1` 的数，即将图像像素映射到单位圆上。
- 2) `V_nm = zpoly(n1, m(r), p, theta);` % V_{nm} 为 n 阶的 Zernike 多项式，定义为在极坐标系中 p, θ 的函数 `zprod = cimg.*conj(V_nm);` % `conj` 是求复数的共轭 $A_{nm}(c) = (n+1) * \text{sum}(\text{sum}(zprod)) / \pi$; % $(n+1) / \pi * \sum \sum (zprod)$; 对于图像而言求和代替了求积分。

本文对手写数字样本进行了识别实验，分别用了文献[15]提出的粗网格特征提取方法和本文提到的降维后的 Zernike 矩特征提取方法进行模糊聚类。



Figure 1. The pretreatment of handwritten digital image
图 1. 手写体数字图像预处理

Table 1. Result of identification
表 1. 字符识别结果

字符	0	1	2	3	4	5	6	7	8	9
总数	100	100	100	100	100	100	100	100	100	100
本文识别数	98	97	99	97	96	97	95	98	99	95
本文识别率	98%	97%	99%	97%	96%	97%	95%	98%	99%	95%
文献[15]识别	95	96	97	93	95	94	95	97	97	96
文献[15]识别率	95%	96%	97%	93%	95%	94%	95%	98%	97%	96%

5. 结论

本文了采用 Zernike 矩进行特征提取，并通过改进 FCM 算法对手写体数字进行模糊聚类。由表 1 可知，这种方法的识别率均在 95%以上，平均识别率也为 97.1%，识别结果较好。图像的 Zernike 矩具有旋转不变性，因此用他构造的特征空间能很好的反映图像的特性。同时，本文将 Zernike 矩进行降维处理，仅选取了 25 个矩，不仅对识别率没有多大影响，还使识别速度得到了提高。当然，本文用到的方法也存在一定缺陷，比如“6”和“9”的误判率较高，究其原因在于旋转后具有很大的相似性。可以在此方法的基础上与别的识别方法相结合，以得到满足各项指标要求的创新方法。

参考文献 (References)

[1] 王亚坤, 曾德良, 李向菊. 一种新颖的数字识别算法[J]. 电力科学与工程, 2009, 25(1): 76-78.
 [2] 赵万鹏, 古乐野. 基于 Adaboost 的手写体数字识别[J]. 计算

机应用, 2005, 25(10): 2413-2414, 2417.
 [3] 刘荣华, 徐学洲. 基于模糊模式 BP 算法的手写体数字识别[J]. 电子科技, 2006, 3: 60-63.
 [4] C. J. C. Burges. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 1998, 2(2): 96-112.
 [5] 胡永东, 叶青. 基于小波和 RBF 神经网络的手写体数字识别[J]. 微处理机, 2005, 4: 24-25, 28.
 [6] M. R. Teague. Image analysis via the general theory of moments. *Journal of the Optical Society of America*, 1980, 70(8): 920-930.
 [7] 曾山. 基于模糊聚类的手写体数字识别[D]. 武汉工业学院, 2009.
 [8] 黎婷婷. 车牌识别图像处理算法的研究与实现[D]. 武汉理工大学, 2011.
 [9] 高新波. 模糊聚类分析及其应用[D]. 西安电子科技大学出版社, 2004.
 [10] 商立群, 杜亚娟. Hu 矩和 Zernike 矩在图像识别中的应用[J]. 西安科技学院学报, 2000, 20(1): 53-56.
 [11] 吴小艳, 王维庆, 杨春祥等. 基于模板匹配的数字图像识别算法[J]. 兵工自动化, 2005, 24(6): 98.
 [12] 杨宾峰, 宁欣. 主成份分析在数字图像特征提取中的应用[J]. 微计算机信息, 2007, 23(6): 282-283, 296.
 [13] 许海峰. 基于 FPGA 的数字图像处理算法研究[D]. 复旦大学, 2010.
 [14] 岑枫, 薛占熬, 卫利萍. 一种基于用户需求的加权模糊聚类分析算法[J]. 计算机技术与发展, 2008, 18(10): 82-88.
 [15] 张魁. 基于遗传-BP 神经网络的手写数字的识别方法[D]. 西安科技大学, 2012.