

# 基于改进模拟退火算法的手写体数字识别

洪沛霖, 张佑生, 邢燕

(合肥工业大学 计算机与信息学院, 安徽 合肥 230009)

**摘 要:**对模糊C-均值聚类作了简要分析和评论,在此基础上将模拟退火机制引入其中,以克服模糊C-均值聚类的局部性和对初始聚类中心的敏感性;采用了基于贴近度和择近原则的模糊识别方法;详细设计了算法。仿真结果说明,该方法在识别速度和准确率方面都达到了令人满意的效果,为手写体数字的识别提供了一种新方法,也拓展了模糊理论的应用范围。

**关键词:**模拟退火算法;模糊模式;数字识别

**中图分类号:**TP18

**文献标识码:**A

**文章编号:**1673-629X(2007)09-0015-03

## Handwritten Digit Recognition Based on Improved Simulated Annealing Algorithm

HONG Pei-lin, ZHANG You-sheng, XING Yan

(School of Computer and Information, Hefei University of Technology, Hefei 230009, China)

**Abstract:** Fuzzy C-means clustering is introduced and remarked. On the basis of systematic analysis of current algorithms, simulated annealing mechanism is inducted into fuzzy clustering to solve the locality and the sensitiveness of the initial condition of fuzzy C-means clustering. Then, proposes the fuzzy discern method based on approximation value and the principle of selecting the near. Finally, the algorithm is designed in detail. Simulation results show that the new method gets a satisfied keywords result both application on speed and the correct rate. Thus a new handwritten digit recognition method is presented, and it widens the application of fuzzy theory.

**Key words:** simulated annealing algorithm; fuzzy pattern; digit recognition

### 0 引言

采用模糊C-均值聚类(FCM)算法进行手写体数字识别,虽然能较准确地对样本进行分类。但实验证明,在数字样本集不理想、初始聚类中心设置不合理的情况下,会导致不好的聚类结果,影响后续的样本识别<sup>[1,2]</sup>。模拟退火(SA)算法是一种启发式的“健壮”的搜索算法,它的最终解不十分依赖于初始解的选取,只要降温过程足够慢,任取一初始解都会得到性能较好的最终解。文中将SA引入模糊聚类,提出一种基于SA的模糊模式识别方法,以期克服FCM搜索的局限性和对初值的敏感性<sup>[3,4]</sup>。

手写体数字模糊模式识别系统由两大部分组成:训练样本部分和识别部分。

收稿日期:2006-11-11

基金项目:教育部高等学校博士学科点专项科研基金(20050359012)

作者简介:洪沛霖(1977-),男,安徽合肥人,硕士研究生,主要研究领域为图像处理及模式识别;张佑生,教授,博导,主要研究领域为图形图像处理及人工智能。

### 1 训练样本部分

所谓训练样本,即是建立手写体数字标准分类库的过程。通过对学习样本作某种监测,得到标准模糊模式,以便后续待识别数字与之比较,达到对输入未知数字进行分类的目的。文中采用无监督的FCM学习。设 $n$ 个数据样本为 $X = \{X_1, X_2, \dots, X_n\}$ ;  $C(2 \leq C \leq n)$ 是要将数据样本分成的类型的数目;  $\{A_1, A_2, \dots, A_C\}$ 表示相应的 $C$ 个类别,  $U$ 是其相似分类矩阵,各类别的聚类中心为 $\{v_1, v_2, \dots, v_C\}$ ;  $\mu_{ik}$ 是样本 $X_i$ 对于类 $A_k$ 的隶属度,则目标函数或称聚类损失函数可以用式(1)表示:

$$J_b(U, v) = \sum_{i=1}^n \sum_{k=1}^C (\mu_{ik})^b (d_{ik})^2 \quad (1)$$

式中:  $d_{ik} = d(X_i - v_i) = [\sum_{j=1}^m (X_{ij} - v_{ij})^2]^{\frac{1}{2}}$

$d_{ik}$ 是欧几里得距离,用来度量第 $k$ 个样本 $X_k$ 与第 $i$ 类中心 $v_i$ 之间的距离。 $m$ 是样本的特征数。 $b$ 是加权参数,取值范围 $1 \leq b \leq \infty$ 。

模糊C-均值聚类方法就是寻找一种最佳的分类 $U$ ,以使该分类能产生最小的函数 $J_b$ 。它要求一个样

本对于各个聚类的隶属度值和为1,即:

$$\sum_{j=1}^c \mu_j(X_i) = 1, i=1, 2, \dots, n \quad (2)$$

式(3)和式(4)分别用于计算  $X_i$  对于类  $A_k$  的隶属度和  $c$  个聚类中心  $\{v_j\}$ :

$$\mu_{ik} = \left[ \sum_{j=1}^c \left( \frac{d_{ik}^2}{d_{jk}^2} \right)^{\frac{2}{2-\alpha}} \right]^{-1} \quad (3)$$

$$v_{kj} = \frac{\sum_{i=1}^n (\mu_{ik})^{\frac{2}{2-\alpha}} X_{ij}}{\sum_{i=1}^n (\mu_{ik})^{\frac{2}{2-\alpha}}} \quad j=1, 2, \dots, c \quad (4)$$

用式(3)和式(4)反复修改数据隶属度、聚类中心并进行分类,当算法收敛时,理论上就得到了各类的聚类中心以及各个样本对于各模式类的隶属度,从而完成了模糊聚类划分。尽管FCM有很高的搜索速度,但FCM是一种局部搜索算法,且对聚类中心的初值十分敏感,如果初值选择不当,它会收敛到局部极小点。

## 2 引入模拟退火的FCM

模拟退火是模拟固体物质的退火过程,将求解优化问题的最优点转化为求一系列随温度变化的物理系统的自由能函数的极小,由退火温度  $T$  控制着搜索过程向最优化方向进行,同时算法以一定的概率接受使目标函数变差的解,这使得该算法具有跳出局部极优解的能力。理论上已经证明SA是全局最优算法,只要退火控制参数选择合适,降温过程足够缓慢,算法以概率1接近全局最优解。

### 2.1 退火控制参数的设定

模拟退火控制参数的选取方法如下。

#### 2.1.1 初始温度的确定

一般情况下,为了使最初产生的新解都被接受,在算法开始时就应达到准平衡。初始温度  $T_0$  通常要由若干次的实验来确定。这里采用式(5):

$$T_0 = \frac{\Delta f}{\ln P_0^{-1}} \quad (5)$$

即先经过计算若干次随机扰动中的目标函数增量的平均值  $\Delta f$ ,其中  $P_0$  取0.8。

#### 2.1.2 产生新解的扰动

SA中新解的产生是对当前解进行扰动得到的。文中采用依赖于温度的似Cauchy分布产生新解<sup>[3]</sup>,目的是加快SA的收敛速度。采用式(6):

$$m'_i = m_i + x_i(B_i - A_i) \quad (6)$$

$$x_i = T \cdot \text{sgn}(u - 0.5) \cdot [(1 + 1/T)^{2u-1} - 1]$$

式中: $m_i$ 为当前解中的第 $i$ 个变量,即第 $i$ 个特征值; $u$ 为 $[0, 1]$ 均匀分布的随机数; $[A_i, B_i]$ 为 $m_i$ 的取值范围,且扰动后的 $m'_i \in [A_i, B_i]$ ;sgn为符号函数。

采用上述扰动的优点是:在高温区进行大范围的

搜索,在低温区搜索仅在当前解附近进行,而且似Cauchy分布有一平坦的“尾巴”,使其易于迅速跳出局部极优。

#### 2.1.3 降温方式

$$\text{用如下的降温方式: } T_k = T_0 \cdot a^k \quad (7)$$

其中, $T_0$ 为初始温度; $a$ 为可调参数,控制温度下降的快慢,通常取  $0.7 \leq a < 1.0$ 。

### 2.2 基于模拟退火的FCM算法描述

S1:初始化SA控制参数  $T_0, a$ , 内循环次数  $S$ , 终止温度  $T_{\text{end}}$ , 初始化FCM的聚类中心数  $C$  以及聚类中心  $V^{(0)}$ ;

S2:用式(1)~(4)计算各数据隶属度  $u^{(k)}$ ,  $C$  个聚类中心  $V^{(k)}$ , 及目标函数  $J_b^{(k)}$ ;

S3:设循环计数变量  $\text{Loop} = 0$ ;

S4:对  $V^{(k)}$  用式(6)进行扰动,产生新解  $V^{(k+1)}$ , 计算新解的目标函数  $J_b^{(k+1)}$ 。若  $J_b^{(k+1)} < J_b^{(k)}$ , 接受新解,否则,以概率  $\exp((J_b^{(k+1)} - J_b^{(k)})/T_k)$  接受新解;

S5:若  $\text{Loop} < S$ , 则  $\text{Loop} = \text{Loop} + 1$ , 转S4, 否则转S6;

S6:若  $T_k < T_{\text{end}}$ , 则算法终止,返回最优解,否则按式(7)进行降温,转S3。

上述聚类学习过程结束后,标准模式库中便已存有已知样本的聚类中心,该聚类中心将是对后续待识别样本进行分类的依据。

## 3 识别部分

文中研究的手写体数字模糊模式识别,要考察一个模糊集与哪一种已知模糊集最“贴近”,可以使用的数量指标有:模糊距离和贴近度。前者表示两者模糊集之间的差异程度,而后者则反映了两个模糊集接近的程度。

### 3.1 贴近度

贴近度<sup>[5]</sup>是用于刻画两个模糊集相似程度的,它是一个数量指标,有很多具体的表示形式。文中采用式(8)所示的格贴近度:

$$\rho(A, B) = \frac{2 \sum_{k=1}^n (A(x_k) \wedge B(x_k))}{\sum_{k=1}^n A(x_k) + \sum_{k=1}^n B(x_k)} \quad (8)$$

### 3.2 择近原则

择近原则<sup>[6]</sup>描述如下:设论域为  $U, A_1, A_2, \dots, A_n$  为  $C$  个标准模糊类型,  $x$  为待识别对象,  $\rho$  为  $U$  上的贴近度,若

$$\rho(A_j, x) = \max\{\rho(A_j, x) | 1 \leq j \leq n\} \quad (9)$$

则认为  $x$  与  $A_i$  最接近,可判定  $x$  属于  $A_i$  类型。

### 3.3 阈值原则

在实际模糊模式识别时,还要考虑确定一个阈值  $V$ 。选择阈值  $V$  时可以根据经验及实际需要来确定,通常根据下列两原则<sup>[7,8]</sup>:

① 当  $\mu_A(x) \geq V$  时,被测对象可以进行确切的分类;

② 当  $\mu_A(x) < V$  时,说明对象的某些特征远离所允许的最大偏差,因而不能归入任一类,应查找原因另作分析。

所以确定阈值的目的在于去掉隶属度很低、显然不合理的对象。

## 4 算法实现

文中使用的图像数据是 CENPAMI (Centre for Pattern Recognition and Machine Intelligence of Concordia University in Canada) 中的手写体数字图像数据库。从该数据库中对待每类数字各取 100 个样本作为训练样本,另各取 100 个样本作为测试样本。图 1 和图 2 分别给出了部分用于训练和测试的样本数字图像。

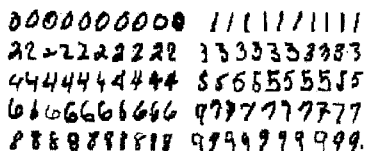


图 1 部分测试样本

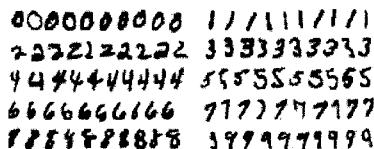


图 2 部分训练样本

### 4.1 算法描述

基于模拟退火算法的模糊模式识别,其具体算法描述如下:

S1: 输入手写体数字的  $n$  个样本;

S2: 对每个样本提取  $m$  个特征,并对特征向量归一化,将其转换为模糊向量;

S3: 用上述引入 SA 的模糊 C-均值聚类方法进行聚类,得出  $c$  个模式的聚类中心;

S4: 输入待识别手写体数字样本  $x$ ;

S5: 对样本  $x$  提取  $m$  个特征,并对特征向量归一化,将其转换为模糊向量;

S6: 用式(8)计算待识别样本  $x$  与各模式的类中心的贴近度  $\rho(A_i, x) (1 \leq i \leq c)$ ;

S7: 求出  $c$  个贴近度中的最大贴近度  $\rho_{\max}$ ;

S8: 让最大贴近度  $\rho_{\max}$  与阈值  $V$  相比,若  $\rho_{\max} < V$ ,则打印出拒识标志,转 S9;否则,根据择近原则,输出样本  $x$  的分类结果;

S9: 继续识别其它样本或结束。

## 4.2 实验结果

在 P4、256M 内存的微机上针对相同的训练样本集和测试样本集用 Visual C++ 6.0 分别实现了 FCM 聚类学习算法与引入模拟退火机制的 FCM 聚类学习算法。实验结果见表 1。

表 1 手写体数字识别实验结果比较(时间单位为 s)

	训练样本	训练耗时	测试样本	识别耗时	识别率	误识率	拒识率
FCM 聚类	1000	106.2	1000	22.5	89.4%	7.3%	3.3%
FCM+SA 聚类	1000	83.6	1000	19.6	90.7%	6.7%	2.6%

从实验结果可知,与 FCM 聚类算法相比,SA 结合 FCM 聚类算法取得了更好的识别效果,训练时间也有明显缩短。导致近 6.7% 的误识率的原因是多方面的:一是由于某些手写体数字样本的分辨率不够,导致其特征不能稳定提取;二是由于手写体样本数量有限,一方面对提取的特征值的分类能力难以用统计的方法进行度量,另一方面算法中的某些参数需要统计后确定,因此难以取得稳定的参数。

## 5 结论与展望

针对手写体数字结构识别的特点,设计了一种模糊模式识别方法。在样本学习阶段,将模拟退火机制引入模糊聚类,利用 SA 的全局搜索性,在一定程度上克服了模糊 C-均值聚类的局部性和对聚类中心初值的敏感性。需要指出的是,尽管 SA 是目前较流行的寻优搜索方法,但方法还有许多亟待完善之处,目前国内外已有人在这方面进行探索并取得了一定的成果,如在 SA 中引入自适应、遗传算法等,以提高 SA 的搜索性能。采用不同的算法与模糊模式识别联系起来,以达到最好的识别效果,是值得研究的一种方法。

### 参考文献:

- [1] 张曾科. 模糊数学在自动化技术中的应用[M]. 北京: 清华大学出版社, 1997.
- [2] 杨光正. 模式识别[M]. 合肥: 中国科学技术大学出版社, 2003.
- [3] Ingber L. Very fast simulated annealing[J]. Math Comput Modeling, 1999, 12: 967-973.

(下转第 20 页)

## 2.2 根据时间相似度聚类

根据公式(2)时间相似度函数,实现了时间相似度函数在  $k$ -中心方法下的聚类。

算法:将闭合序列模式聚类

输入:

闭合序列模式集  $P = \{\beta_1, \beta_2, \dots, \beta_m\}, \beta_i = \{a_1, a_2, \dots, a_k\}$

聚类簇数  $N$  输出:需要连接的序列位置列表。

输出:  $n$  个闭合序列模式集,  $C_1, C_2, \dots, C_n$

(1) load pattern  $a_1, a_2, \dots, a_n$

(2) initialize  $C_1, C_2, \dots, C_n$

/\* 计算两个序列模式之间的时间相似度 \*/

for each  $a_i$

for each  $a_j, j < i$

$ST_{ij} = s(a_i, a_j) / *$  对模式进行聚类 \*/

(3) repeat

(4) DistributeSamples(); /\* 指派每个剩余的對象  
给离它最近的中心点所代表的簇 \*/

(5) CalcNewClustCenters(); /\* 重新计算该簇的  
中心点 \*/

(6) until 每个簇的中心都不发生变化;

/\* 除掉每个聚类中的冗余项 \*/

for each Cluster  $C_i (i = 1, \dots, n)$

for each Pattern  $a_k$  in Cluster  $C_i$

if(IsCovered( $a_k$ ) = True

remove  $a_k$  from  $C_i$

(7) return  $C_i (i = 1, \dots, K)$

## 3 结论

在用户点击流序列的基础上,利用距离函数聚类算法和序列时间相似度函数聚类算法,可以研究序列的相关性和用户对网页内容所表现出来的兴趣的相似性。这对于开发一些目的性强的网站具有较大帮助(如电子商务网站),可以提高 Web 用户的服务质量,使用户享用到满意的个性化服务。下一步工作是在这

种算法系统基础上,减少时间复杂度和空间复杂度。

## 参考文献:

- [1] Agrawal R, Imielinski T, Swami A. Mining association rules between sets of items in large databases[C]//In: Proc of ACM SIGMOD International Conference on Management of Data. Washington D.C.: [s.n.], 1993:207-216.
- [2] Brin S, Motwani R, Silverstein C. Beyond market basket: Generalizing association rules to correlations[C]//In: Proc of 1997 ACM SIGMOD Int'l Conf on Management of Data. Tucson, Arizona, UAS: ACM Press, 1997:265-276.
- [3] Silverstein C, Brin S, Motwani R, et al. Scalable techniques for mining causal structures[C]//In: Proc of 1998 ACM SIGMOD Int'l Conf. on Management of Data. Seattle, Washington, USA: [s.n.], 1998:343-353.
- [4] Agrawal R, Srikant R. Mining sequential patterns[C]//In: Proc International Conference on Data Engineering. Taipei, Taiwan: [s.n.], 1995:3-14.
- [5] Han J, Dong G, Yin Y. Efficient mining of partial periodic patterns in time series database[C]//Int Conf Data Engineering (ICDE 99). Sydney: IEEE Press, 1999:106-115.
- [6] Han J, Pei J, Yin Y. Mining frequent patterns without candidate generation[C]//In: Proc of 2000 ACM SIGMOD international conference on management of data. Dallas, Texas, United States: [s.n.], 2000:1-12.
- [7] Ayres J, Flannick J, Gehrke J, et al. Sequential Pattern mining using a bitmap representation[C]//Proceedings of the eighth ACM SIGKDD international conference on knowledge discovery and data mining. Edmonton, Alberta, Canada: [s.n.], 2002:429-435.
- [8] Xin D, Han J, Yan X, et al. Mining compressed frequent-pattern sets[C]//Proceedings of the 31st international conference on very large databases. Trondheim, Norway: [s.n.], 2005:709-720.
- [9] Strehl A, Ghosh J, Mooney R. Impact of similarity measures on Web-page clustering[C]//In Proc. 7th natl Conf on Artificial Intelligence: Workshop of Artificial Intelligence for Web Search(AAAI 2000). Austin, TX, USA: [s.n.], 2000: 58-64.

(上接第 17 页)

- [4] Murino V. Structured neural networks for pattern recognition [J]. IEEE Transactions on System, Man, and Cybernetics part B: Cybernetics, 1998, 28(4): 53-56.
- [5] Han Hong, Yang Jing-yu. Combination of neural network classifiers[J]. Journal of Computer Research & Development, 2000, 37(12): 1488-1492.
- [6] Flandrin P, Rilling G, Gonalves P. Empirical Mode Decomposition as a Filter Bank [J]. IEEE Signal Processing Letters,

2004, 11(2): 112-114.

- [7] Slavik P, Govindaraju V. Equivalence of Different Methods of Slant and Skew Corrections in World Recognition Applications [J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 2001, 23(3): 323-326.
- [8] Pandya A S, Macy R B. Pattern Recognition with Neural Networks in C++ [M]. [s.l.]: IEEE Press, 1999: 156-172.