对手抑制式模糊 C-均值算法

魏立梅[1],谢维信[2]
(1. 西安电子科技大学电子工程学院 202 室,西安 710071; 2. 深圳大学校长办公室,深圳 518060)

摘 要: 从竞争学习的角度提出模糊 C-均值算法中隶属度的新解释,并基于隶属度的新解释提出对手抑制式模糊 C-均值算法. 理论分析和实验表明:对手抑制式模糊 C-均值算法提高了模糊 C-均值算法的收敛速度.

关键词: 模糊 C-均值算法; 隶属度; 竞争学习
中图分类号: TP183 **文献标识码**: A **文章编号**: 0372-2112 (2000) 07-0063-04

Rival Checked Fuzzy C-Means Algorithm

WEI Li-mei[1], XIE Wei-xin[2]
(1. *Lab 202, School of Electronic Engineering, Xidian University, Xi'an 710071, China*
2. *President Office, Shenzhen University, Shenzhen 518060, China*)

Abstract: A new interpretation of the membership degree of the fuzzy C-means algorithm is presented on the basis of competitive learning and then a new algorithm called rival checked fuzzy C-means algorithm is developed. The theoretical analysis and experimental results show that the new algorithm improves the convergence speed of the fuzzy C-means algorithm.

Key words: fuzzy C-means algorithm; membership degree; competitive learning

1 引言

模糊 C-均值算法(Fuzzy C-Means Algorithm,简记为 FCM)是一种有效的聚类算法,已经广泛应用于模式识别、图像处理和计算机视觉等领域[1]. 但是,FCM 在大数据量时收敛速度慢. 为克服 FCM 的这一不足,本文提出对手抑制式模糊 C-均值算法(Rival Checked Fuzzy C-Means Algorithm,简记为 RCFCM). 下面首先提出 FCM 中隶属度的新解释,然后,基于隶属度的新解释提出 RCFCM. 并通过实验证明 RCFCM 提高了 FCM 的收敛速度.

2 隶属度的新解释

2.1 FCM 简介[1]

设 FCM 中数据集 $X=\{x_1, x_2, \cdots, x_j, \cdots, x_N\}$, N 是 X 中元素总数. X 中的第 j 个元素是一个 p 维矢量, $x_j=(x_{j1}, x_{j2}, \cdots, x_{jp})$. X 中含有 C 个类别,第 i 个类的中心 $v_i=(v_{i1}, v_{i2}, \cdots, v_{ip})$. X 中第 j 个元素对第 i 个类中心的隶属度为 u_{ij}. $U=\{u_{ij}\}$、$V=\{v_i\}$. U 满足下式:

$$u_{ij}\in[0,1],\ \forall i,\ \forall j \qquad (1)$$

$$\sum_{i=1}^{C}u_{ij}=1,\ \forall j \qquad (2)$$

$$0<\sum_{j=1}^{N}u_{ij}<N,\ \forall i \qquad (3)$$

FCM 的目标函数为:

$$J(U,V)=\sum_{i=1}^{C}\sum_{j=1}^{N}u_{ij}^{m}\|x_j-v_i\|^2=\sum_{i=1}^{C}\sum_{j=1}^{N}u_{ij}^{m}d_{ij}^2 \qquad (4)$$

m 称为模糊系数,是常量. FCM 就是求在满足式(1)～(3)的情况下,式(4)中目标函数 J 的最小值. 由拉格朗日乘数法求得 J 的条件极值 $(U、V)$ 满足式(5)～(9).

对 $\forall x_j\in X$,定义集合:

$$I_j=\{i\mid 1\le i\le C, d_{ij}=0\},\ \overline{I}_j=\{1,2,\cdots,C\}-I_j \qquad (5)$$

对于给定的元素 x_j, I_j 和 $\overline{I}_j$ 是确定的; I_j 和 $\overline{I}_j$ 随 x_j 的变化而变化. 式(2)可以进一步表示为:

$$\sum_{i\in(\overline{I}_j)}u_{ij}=1,\ \forall j \qquad (6)$$

对于元素 $x_j\in X$ 和 $\forall v_i\in V$,若 $d_{ij}\neq 0$,即对于元素 $x_j\in X$,若 $I_j=\phi$,则 x_j 对类中心 $v_i\in V$ 的隶属度为:

$$u_{ij}=1/\left(\sum_{k=1}^{C}(d_{ij}^2/d_{kj}^2)^{1/(m-1)}\right) \qquad (7)$$

对于元素 $x_j\in X$,若 $\exists v_i\in V$,使 $d_{ij}=0$,即对于元素 $x_j\in X$,若 $I_j\neq\phi$,则 x_j 对类中心 $v_i\in V$ 的隶属度满足下式:

$$\forall i\in\overline{I}_j,\ u_{ij}=0 \quad \text{且} \quad \sum_{i\in\overline{I}_j}u_{ij}=1 \qquad (8)$$

$$v_i=\sum_{j=1}^{N}u_{ij}^{m}x_j\Big/\sum_{j=1}^{N}u_{ij}^{m} \qquad (9)$$

式(5～9)表示的 (U,V) 中一定包含着最小值点. FCM 的任务就是通过迭代寻找这个最小值点. FCM 的迭代方式有两种:

收稿日期:1999-01-18;修订日期:2000-01-17
基金项目:国家自然科学基金(No. 69472046)资助课题

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

(1) 给定初始聚类中心 V,按式(5)～(8)计算 U,再由 U 按式(9)计算下一次迭代的 V.

(2) 按式(1)～(3)初始化 U,由式(9)计算 V,再由 V 按式(5)～(8)计算下一次迭代的 U.

2.2 隶属度的新解释

J. C. Bezdek 曾经从物理学的角度对隶属度作出解释(见文献[1] pp. 71～73). 1998 年,范九伦和谢维信又从几何学的角度给出隶属度的解释[2]. 现在本文从模糊竞争学习的角度给出隶属度的新解释.

1996 年,钱云涛和谢维信研究了基于 FCM 的神经网络,给出了网络的构造方法和相应的模糊竞争学习算法[3]. 基于 FCM 的神经网络实质上是 FCM 的并行实现,网络在稳态时的聚类结果与 FCM 的结果等价. 只是在聚类过程中,神经网络采用模糊竞争学习算法调整各个类中心. FCM 对应的模糊竞争学习算法中各类中心的调整公式如下[3]:

$$\Delta v_i = u_{ij}^m (x_j - v_i) \tag{10}$$

从式(10)可以看到:在 FCM 对应的模糊竞争学习算法中,对任意一个输入数据 x_j,类中心 v_i 的学习率就是 u_{ij}^m.

从作用力来看,类中心 v_i 在输入数据 x_j 作用下的调整过程就是 v_i 在 x_j 吸引力作用下发生正向位移,更靠近 x_j 的过程. 位移 Δv_i 不仅与 x_j 和 v_i 之间的相对位置 $(x_j - v_i)$ 有关,而且与 x_j 对 v_i 的隶属度有关. 因此,从作用力的角度看,数据对某个类中心的隶属度是数据对该类中心作用力强度的表征.

对上述结论进行引伸,可以给出 FCM 隶属度的新解释:

在 FCM 中,每个数据都对类中心的下一次迭代值产生影响(见式(9)),数据对类中心的这种影响可以理解为数据对类中心施加的吸引力,隶属度是数据对类中心作用力强度的表征. 隶属度越大,数据对类中心的吸引力就越大,类中心的下一次迭代值受它影响也就越大.

3 新算法

3.1 FCM 的分析

对于式(4)表示的聚类问题,如果找到各类的类中心,那么由式(5)～(8)计算得到的 U 必然使 J 取得最小值,而且 U、V 同时满足式(5)～(9),因此当找到正确的类中心时,也就找到了式(4)的最优解. 所以,由 V 按式(5)～(8)计算出 U 后,接下来的任务就是如何确定下一次迭代时的 V,使 V 更靠近各类的类中心,使 FCM 尽快收敛. 从这一点考虑,FCM 中由 U 计算 V 的过程,实际上是从优化的角度求得下一次迭代的 V. 既然逐步最优并不能确保全局最优,如果可以找到一个新的方法,由它确定的 V 比用式(9)确定的 V 更合理,更靠近各类的类中心,就没有必要按照式(9)计算下一步迭代的 V.

3.2 新算法的思想基础

现在分析由式(9)确定下一次迭代的 V 并不合理,延缓了 FCM 的收敛速度.

从式(9)可以看到,X 中的每个元素都对第 i 个类中心 v_i 的下一次迭代值有影响. 根据隶属度的新解释,可以将这种影响理解为每个元素对第 i 个类中心都有吸引力,而第 j 个元素对 v_i 的隶属度 u_{ij} 则是第 j 个元素对 v_i 的吸引力强度的表征,u_{ij} 的值大时,表示元素 x_j 对 v_i 的吸引力大;u_{ij} 的值小时,表示元素 x_j 对 v_i 的吸引力小. u_{ij} 是与距离有关的,x_j 对近的类中心的隶属度大,对远的类中心隶属度值小. 所以,x_j 对近的类中心吸引力大,对远的类中心吸引力小. 具体地讲,x_j 对 v_i 的吸引力是通过 $u_{ij}^m x_j$ 项施加给 v_i 的.

x_j 对每个类中心都有吸引力的好处是:按式(9),即使是远离数据区的类中心也能在聚类过程中得到调整,C 个类中心中不会存在死节点. 这也正是 FCM 优于硬 C-均值算法(Hard C-Means Algorithm,简记为 HCM)的地方. 在 HCM 中,数据只隶属于离它最近的类中心,而对其它类中心的隶属度为 0. 因此,在 HCM 中数据只对离它最近的类中心产生吸引力,对其它类中心没有作用力. 这样,在 HCM 中,当某个类中心距离所有的数据点都比较远时,所有的数据点都对它没有吸引力. 该类中心在聚类过程中就得不到调整,它的位置就会在聚类过程中保持不变,成为"死节点". 死节点的存在,会使 HCM 算法陷入局部极值,得到错误的聚类结果.

FCM 中元素对所有的类中心都有吸引力,虽然使 FCM 不存在死节点,提高了 FCM 的寻优概率,但是 FCM 也因此付出了代价:FCM 的收敛速度慢于 HCM. FCM 中各类别对应不同的类中心. 既然每个元素对所有的类中心都有吸引力,那么 X 中的每个数据类从宏观上不仅产生一个很强的吸引力(距离最近,隶属度最大,吸引力也最大),使离它最近的类中心在下一次迭代时更靠近它,而且同时对其它的类中心也有不同程度的吸引(距离远,隶属度小,吸引力也较小). 这样,一个类中心在受到离它最近的类的元素的吸引靠近该类时,也受到来自其它类的元素的吸引,使它远离该类. 这势必延缓 FCM 的收敛速度.

如图 1 所示. 类 1 的元素距离类中心 1 最近,对类中心 1 的隶属度最大,吸引力也最大;类 1 的元素距离类中心 2 次最近,对类中心 2 的隶属度次最大,吸引力也次最大;类 1 的元素距离类中心 3 最远,对类中心 3 的隶属度最小,吸引力也最小. 类 2 则对类中心 2 的吸引力最大,对类中心 1 的吸引力次最大,对类中心 3 的吸引力最弱. 类 3 的情况类似.

类中心 1 在受到类 1 元素很强的吸引力向类 1 收敛的同时,也受到类 2 和类 3 元素的吸引力,使它远离类 1. 因此,类中心 1 向类 1 的收敛因为类 2 和类 3 元素的吸引力而被延缓了. 类 2 中元素对次最近的类中心 1 的吸引力仅次于它对类中心 2 的吸引力也很大,所以,类 2 元素对类中心 1 的吸引力对类中心 1 收敛的延缓作用大于类 3 元素的延缓作用. 对于类中心 2 和类中心 3 的收敛情况有类似的分析.

从图 1 中还可以得到:当类 2 和类 3 的元素很多时,类 2 和类 3 对类中心 1 的作用力增强,由于类 2 和类 3 造成的类中心 1 的收敛速度的延缓也就越明显. 因此,在大数据量和类数较多情况下,FCM 的收敛速度也就越

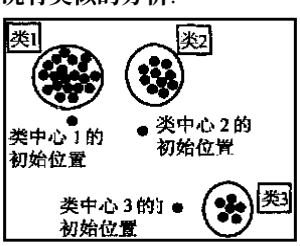


图 1 FCM 作用力示意图

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

慢.

从图 1 中还可以得到:当类 1 的元素个数远比类 2 和类 3 的元素多时,类中心 1 在类 1 元素作用下的收敛速度会大大加快. 同时,类中心 2 在受到来自类 1 大量数据的很强的吸引力下,收敛速度会大大延缓,而且类 1 元素对类中心 2 的次最大的吸引力会使类中心 2 在收敛时偏离类 2 的中心,甚至远离类 2 的中心. 所以,FCM 在聚类过程中,对数据量很少的类别难以正确聚类. 如图 2 所示,FCM 对图 2 中的数据集难以正确聚类. 图 2(b)中的类 2 的类中心在类 1 数据的吸引力下,偏离类 2 的中心,造成 FCM 的错误分类.

综上所述,FCM 按式(9)确定 V 的下一次迭代值延缓了 FCM 的收敛速度. 而且,数据对离它次最近的类中心的吸引力对 FCM 速度的延缓作用最大.

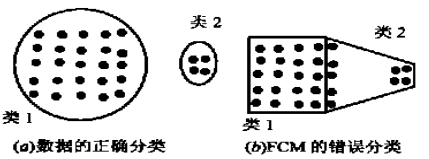


图 2　FCM 的错误聚类结果

数据对离它次最近的类中心的吸引力也是不应该的. 数据对距离它最近的类中心隶属度最大,吸引力也最大,是理所当然的. 只有这样,类中心才能在离它最近的类的吸引力作用下,收敛于该类. 数据对较远类中心的微弱吸引力能避免死节点的出现,提高 FCM 的寻优概率,同时对收敛的延缓作用不大. 可是,数据对离它次最近的类中心的吸引力是很大的,仅次于对最近类中心的吸引力,对 FCM 收敛的延缓作用也最大. 而且,一个数据没有任何理由在以最大的力将离它最近的类中心吸引过来的同时,又以较大的力将离它次最近的类中心也吸引过来. 从宏观上看,一个类别更没有任何理由将离它最近的类中心吸引过来的同时,又以较大的力将离它次最近的类中心吸引过来. 一个类别只能对应一个类中心.

所以本文认为:数据应该对离它最近的类中心施加最大的吸引力,这个力越大,类中心收敛速度也越快;数据也可以对较远的类中心有微弱的吸引力,避免聚类过程出现死节点;但不应该对离它次最近的类中心有很强的吸引力,这个力对 FCM 的收敛速度延缓作用最大. 据此,本文提出新算法.

3.3　新算法

在竞争学习算法中,距离数据最近的类中心被称为赢者,距离数据次最近的类中心被称为赢者的对手. 因此称新算法为:对手抑制式模糊 C-均值算法(RCFCM).

新算法实现步骤如下:

第一步　算法的初始化

初始化类中心为 $V(0)$. 给定模糊系数 m, $m \in [1, +\infty]$. 置吸引力抑制因子 α, $\alpha \in [0,1]$. 迭代次数为 $l = 0$.

第二步　按照式(5)～(8)计算 $U(l)$.

第三步　修正隶属度矩阵 $U(l)$.

对元素 x_j,设它对第 p 类的隶属度最大,值为 u_{pj};对第 s 类的隶属度次最大,值为 u_{sj}. 修正后,数据 x_j 对第 p 类的隶属度为:

$$u_{pj} = u_{pj} + (1 - \alpha) u_{sj} \tag{11}$$

对类 s 类的隶属度为:

$$u_{sj} = \alpha u_{sj} \tag{12}$$

对其余各类的隶属度不变.

新算法对赢者对手吸引力的抑制作用由抑制因子 α 控制. 当 $\alpha = 0$ 时,式(11)变为:$u_{pj} = u_{pj} + u_{sj}$、式(12)变为:$u_{sj} = 0$. RCFCM 算法完全抑制了数据点对赢者对手的吸引力,增大了数据点对赢者的吸引力. 当 $\alpha = 1$ 时,在隶属度修正前后,u_{pj} 和 u_{sj} 的值不变. RCFCM 算法退化为 FCM 算法. 当 $0 < \alpha < 1$ 时,RCFCM 算法对赢者对手的抑制作用随 α 增大而减小,对赢者的吸引力的增强作用随 α 增大而减弱.

因此,式(11)和式(12)体现了 RCFCM 算法抑制赢者对手的思想. 并且,从式(11)和式(12)可以看出:在隶属度修正前后,$u_{pj} + u_{sj}$ 的值保持不变. RCFCM 算法满足式(1)～(3).

第四步　用修正后的隶属度矩阵计算下一次的迭代中心 $V(l+1)$**. 计算公式为式(9).**

第五步　判断是否终止迭代.

如果不满足终止条件,更新参数,返回第二步,继续迭代. 参数更新公式:$l = l + 1$.

3.4　新算法的收敛性

新算法中,隶属度矩阵的修正不会改变 FCM 的收敛性. 容易证明新算法是收敛的,收敛性的证明过程与 FCM 的收敛性证明过程类似.

4　实验结果

给出二维平面上 5 类高斯分布的数据,数据总数为 500. 实验中取 $\varepsilon = 0.01$, $\alpha = 0$,采用 $\sum_{i=1}^{C}\sum_{j=1}^{N} \| u_{ij}(l+1) - u_{ij}(l) \| < \varepsilon$ 作为迭代结束条件. 从式(11)和式(12),可以得到:实验中 RCFCM 算法完全抑制了数据点对赢者对手的吸引力,增大了数据点对赢者的吸引力. 随机初始化聚类中心,对每组聚类中心,记录 FCM 和 RCFCM 这两种算法在模糊系数依次为 $m = 2$、$m = 10$ 时的实验结果. 对于每组实验,得到的结论是一样的. 现选 10 组实验结果列成表 1.

表 1　FCM 和 RCFCM 的性能比较

	$m = 10$				$m = 2$			
	FCM		RCFCM		FCM		RCFCM	
	聚类结果	迭代次数	聚类结果	迭代次数	聚类结果	迭代次数	聚类结果	迭代次数
1	极值	82	最优	44	最优	16	最优	12
2	最优	121	最优	42	最优	17	最优	12
3	最优	38	最优	24	最优	14	最优	10
4	最优	83	最优	43	最优	15	最优	10
5	极值	74	最优	89	最优	17	最优	13
6	最优	36	最优	23	最优	12	最优	9
7	最优	76	最优	42	最优	14	最优	10
8	极值	34	最优	39	极值	12	最优	8
9	极值	33	最优	44	最优	15	最优	11
10	最优	85	最优	44	最优	13	最优	10

从表 1 可以看出,FCM 在模糊系数取得比较合适时($m = 2$),对各组初始化类中心,FCM 都收敛于最优解(组 8 除外). 随着模糊系数增大,FCM 的性能降低,$m = 10$ 时,FCM 在一些实验中陷入局部极值. 而 RCFCM 在实验中则表现出优良的性能,无论在模糊系数取值合适时,还是在模糊系数取值很大时,都得到最优解.

在同样得到最优解时,RCFCM 在迭代过程中聚类中心向

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

正确的类中心迅速靠近,收敛速度大大快于 FCM,这一点从表 1 中的数字对比中可以一目了然.

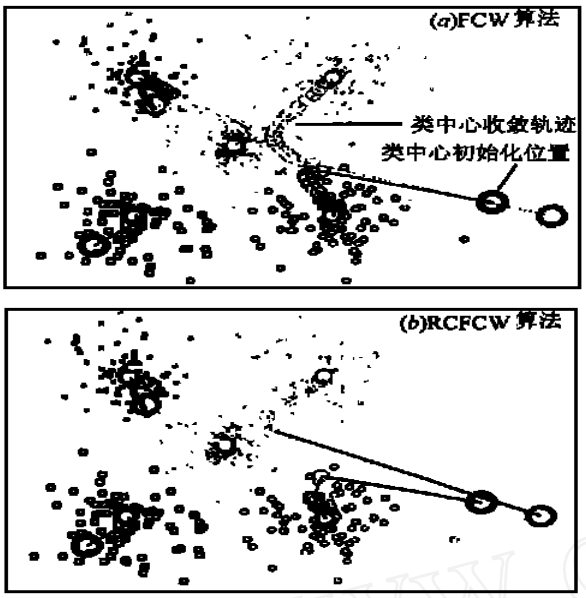


图 3 组 6 中 FCM 和 RCFCM 在 $m = 10$ 时的实验结果和类中心收敛轨迹

图 3 中给出了组 $m = 10$ 时,FCM 和 RCFCM 聚类结果和类中心收敛轨迹. 图 3 中,不同类别的数据用不同符号表示. 图中的大圆环的中心是类中心的初始位置. 通过对比图 3(*a*)和图 3(*b*)中的聚类结果和类中心收敛轨迹,可以得到结论:FCM 和 RCFCM 都得到正确的分类结果,但从五个类中心的收敛轨迹上看,在 RCFCM 中类中心很快地向各个数据类的中心处收敛,而 FCM 中类中心很缓慢地向各数据类的中心靠近.

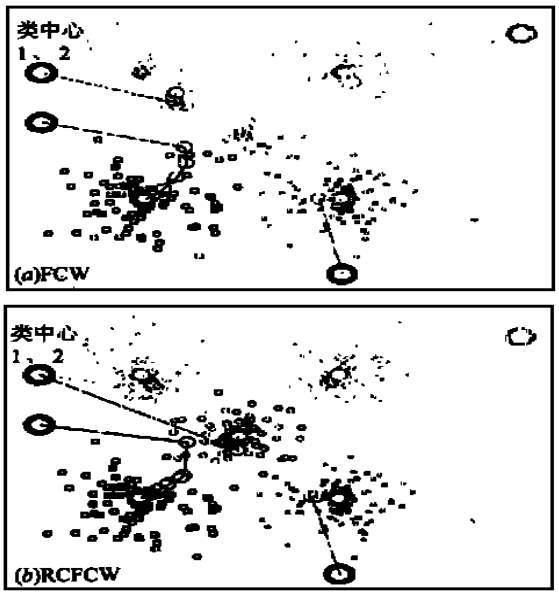


图 4 组 8 中 FCM 和 RCFCM 聚类结果的比较(*a*)FCM 算法在 $m = 2$ 时的实验结果和类中心收敛轨迹;(*b*)RCFCM 算法在 $m = 10$ 时的实验结果和类中心收敛轨迹

图 4 给出了组 8 中 FCM 在 $m = 2$ 时陷入局部极值,而 RCFCM 在 $m = 10$ 时仍得到最优解的例子. 在组 8 中,类中心 1 和类中心 2 恰巧初始化为同一个点,无论模糊系数取值如何,FCM 中这两个类中心的迭代轨迹都完全一样,无法得到正确结果(见图 4(*a*)). 而 RCFCM 却解决了这个问题,因为 RCFCM 中元素不会将两个类中心同时吸引向同一类(见图 4(*b*)).

实验有力证明,RCFCM 在同样初始聚类中心的情况下提高了收敛速度.

从实验还得到:在模糊系数增大时,RCFCM 性能明显优于 FCM. 按照式(5～8)模糊系数的增大,使 FCM 中元素对各类中心的隶属度呈现出等分配的趋势,元素对最近类中心和其它类中心的隶属度差别减小,吸引力相差不大. 此时,数据对次赢者的吸引力对 FCM 收敛速度的延缓作用将更明显. 因此,FCM 在模糊系数增大时,收敛速度慢,性能降低很大. 虽然模糊系数的增大使 RCFCM 的性能也降低了,但 RCFCM 去掉了数据对次最近类中心的吸引力,使算法的收敛性能相对于 FCM 得到很大提高.

大数据量和类数较多同样会使 FCM 的收敛速度延缓. 与模糊系数增大相比,起因不同,但对 FCM 收敛过程的影响作用却是一样的. 所以,在大数据量和类数较多时,RCFCM 对 FCM 收敛速度的提高将很明显.

5 结束语

本文从模糊竞争学习的角度提出了隶属度的新解释,并基于隶属度的新解释提出对手抑制式 FCM. 理论分析和实验表明:新算法提高了 FCM 的收敛速度.

参考文献

[1] J. C. Bezdek. Pattern recognition with fuzzy objective function algorithms. Plenum Press. New York,1981

[2] 范九伦. 模糊聚类新算法和聚类有效性问题的研究. 博士学位论文. 西安电子科技大学,1998

[3] 钱云涛. 智能信息处理中的模糊神经网络集成方法研究. 博士学位论文. 西安电子科技大学,1996



魏立梅 1970 年出生,工学博士. 1992 年 7 月获得西安电子科技大学检测技术与仪器专业学士学位. 1995 年 3 月获得该校电路与系统学科硕士学位. 1998 年 9 月获得该校信号与信息处理学科博士学位. 研究方向包括:模糊理论、图像处理、神经网络、通信理论.



谢维信 1941 年出生,深圳大学校长、博士生导师、教授. 研究方向包括:模糊理论、信号处理、神经网络、图像处理、计算机视觉等领域.

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net