# 用遗传算法实现模糊测度赋值的一种多分类器融合算法

刘汝杰[1,2],袁保宗[1],唐晓芳[1]
(1. 北方交通大学信息科学研究所,北京 100044;2. 富士通研究所开发中心有限公司,北京 100044)

**摘 要**: 由于模糊理论可以很好的表达和处理不确定性问题,因而得到了广泛的应用,并成为信息融合领域中的有效方法. 只要选取合适的模糊测度值,基于模糊积分的多分类器融合方法就可以达到比最优的单分类器更好的分类效果. 用遗传算法来计算模糊测度的值时,可以得到解空间中的最优解,从而实现比非优化方法更好的融合效果. 模拟实验结果也证实了这一点.

**关键词**: 信息融合; 模糊积分; 遗传算法
**中图分类号**: TP391 **文献标识码**: A **文章编号**: 0372-2112 (2002) 01-0145-03

# Multiple Classifiers Fusion Algorithm with the Fuzzy Measures Determined by Genetic Algorithm

LIU Ru-jie[1,2], YUAN Bao-zong[1], TANG Xiao-fang[1]
(1. *The Institute of Information Science, Northern Jiaotong University, Beijing* 100044, *China*; 2. *Fujitisu R&D Center CO., LTD. Beijing* 100016, *China*)

Abstract: Fuzzy set methods have recently achieved a high degree of popularity due to its ability of representing and managing uncertainty, and become an effective method in information fusion fields. If the values of fuzzy measure, which can be got through optimization method, are appropriate, multiple classifier fusion method based on fuzzy integral performs better than the best classifier. In this paper, the Genetic Algorithm is adopted to search the optimized values of fuzzy measure, thus a satisfied fusion result better than that without GA can be obtained, as was verified in the experiment.

Key words: information fusion; fuzzy integral; genetic algorithm

## 1 引言

近些年来,信息融合技术越来越多地引起了人们的关注,并取得了极大的进展. 它能整合来自多信息源的信息,降低单信息源中存在的不确定性,从而提高系统的整体性能[1,2]. 信息融合技术可以在数据级、特征级、决策级三个级别进行. 而对于一个模式识别问题,如果能同时设计多个不同的分类器,并将这些分类器的结果融合在一起,就可以达到单分类器所不能达到的效果. 本文将主要研究多分类器的融合问题.

由于模糊集理论可以很好的表达、处理不确定性问题,因而在控制、模式识别等领域得到了广泛的应用[3~5],并成为在决策级上进行信息融合的一种有效方法. 基于模糊积分的多分类器融合方法实质上是一种扩展的加权平均法,它可以根据分类器输出和模糊测度值"动态"地为各分类器赋予权值. 在模糊积分过程中,模糊测度(权系数)起着至关重要的作用,选取不同的模糊测度值将得到不同的融合结果. 这样,模糊积分过程就演变为权系数空间中的优化问题,在本文中,采用遗传算法来搜索最优的模糊测度值. 遗传算法[6,7]是模仿生物进化模型的一个优化搜索过程,它主要包括三种操作: 复制(Reproduction)、交叉(Crossover)和变异(Mutation). 在进化过程中,适应性强的个体以较大的概率被保存下来并产生后代,而适应性弱的个体产生后代的概率较小,这样,群体将随着进化过程逐渐向优化的方向发展. 在用该方法搜索模糊测度值时,以系统的实际输出和期望输出之间的均方误差作为适应度准则. 模拟结果表明,用基于优化权系数的模糊积分方法,可以实现比非优化方法更好的融合效果.

## 2 模糊测度和模糊积分

设 $S$ 为任意集合, $P(S)$ 表示 $S$ 的幂集. 如果集合函数 $g$ 满足:

i $g(\phi)=0$, $g(S)=1$.

ii $g(A)\le g(B)$, if $A\subset B$ 且 $A,B\in P(S)$. (1)

iii $\lim_{i\to\infty} g(A_i)=g(\lim_{i\to\infty} A_i)$ 如果 $A_i\in P(S)$, 且 $\{A_i\}$ 是单调的.

称 $g$ 是 $P(S)$ 上的一个模糊测度.

如果模糊测度还满足下面的附加条件的话,就称为 Sugeno 测度,记为 $g_\lambda$:

$$g(A\cup B)=g(A)+g(B)+\lambda g(A)g(B),\quad (\lambda>-1),A,B\in P(S)\text{ 且 }A\cap B=\phi \qquad (2)$$

假设 $S$ 是由有限个信息源组成的集合, $S=\{s_1,s_2,\cdots,s_n\}$,并记 $g_i=g_\lambda(\{s_i\})$, $g_1,g_2,\cdots,g_n$ 的值被称为模糊密度.

收稿日期:2000-07-28;修回日期:2001-04-12
基金项目:国家自然科学基金(No. 69789301)

则 $g_\lambda(S)$ 可以表示为[3]:

$$g_\lambda(S) = \sum_{i=1}^{n} g_i + \lambda \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} g_i g_j + \cdots + \lambda^{n-1} g_1 g_2 \cdots g_n \quad (3)$$

当 $\lambda \neq 0$ 时,根据模糊测度的性质 ⅰ,$\lambda$ 值可由下式确定:

$$\lambda + 1 = \prod_{i=1}^{n} (1 + \lambda g_i) \quad (4)$$

设 $f: S \to [0,1]$ 为定义在集合 $S$ 上的函数,则 $f$ 在 $A \subset S$ 上的关于模糊测度 $g$ 的 Sugeno 模糊积分定义为:

$$\int f(s) \circ g(\cdot) = \sup_{\alpha \in [0,1]} [\min(\alpha, g(f_\alpha))] \quad (5)$$

其中,$f_\alpha$ 为 $f$ 的 $\alpha$ 截集,即 $f_\alpha = \{s : f(s) \geq \alpha\}$.

Murofushi 和 Sugeno 发展了 Sugeno 积分的理论,提出了所谓的 Choquet 积分. $f$ 关于模糊测度 $g$ 的 Choquet 积分定义为[8]:

$$\int f(s)\, dg(\cdot) = \sum_{i=1}^{n} [f(s_i) - f(s_{i-1})]\, g(A_i), \quad \text{其中 } f(\delta_0) = 0 \quad (6)$$

其中,$A_i = \{s_i, s_{i+1}, \cdots, s_n\}$. 在进行多分类器融合时,$S$ 为所有信息源组成的集合,令 $f^k(s)$ 表示各信息源对第 $k$ 个命题的证据支持度,$g^k(\cdot)$ 表示相应于第 $k$ 个命题的模糊测度,利用 Choquet 积分进行融合后的系统对第 $k$ 个命题的证据支持度为:

$$w^k(S) = \int f^k(s)\, dg^k(\cdot) = \sum_{i=1}^{n} [f^k(s_i) - f^k(s_{i-1})]\, g^k(A_i) \quad (7)$$

在用模糊积分进行多分类器融合时,不同的模糊测度将形成不同的融合函数. 然而,只要选择合适的模糊测度值,就能使融合后的分类性能比最优的单个分类器的性能要好.

**定义** 如果由融合函数组成的空间 $H$ 中包含下列函数:

$$h^i(f(s_1), f(s_2), \cdots, f(s_n)) = f(s_i),\ i = 1, 2, \cdots, n$$

$f(s_i)$ 表示第 $i$ 个分类器的输出结果,则称 $H$ 具有可分解性[2]. 对分类器 $s_i$,当实际输出和理想输出分别为 $f(s_i)$ 和 $X$ 时,其均方误差为:

$$I_s(s_i) = \int X - f(s_i)]^2\, dp_{f(s_i),X} \quad (8)$$

而对融合函数 $h$,当各分类器输出是 $f(s_i)$ 时,融合后的结果 $f_h(S) = h(f(s_1), \cdots, f(s_n))$,则其均方误差为:

$$I_H(h) = \int X - f_h(S)]^2\, dp_{f_h(S),X} \quad (9)$$

**定理** 设基于 Choquet 模糊积分的融合函数空间为 $H$,则可以找到 $h^* \in H$,使得

$$I_H(h^*) = \min_{h \in H} I_H(h) = \min_{i=1}^{n} I_S(s_i) - \Delta, \Delta \in [0, \infty) \quad (10)$$

**证明** $H$ 满足可分解性. 下面以一个典型的模糊测度进行说明

设 $\dot{g}_m = 1$, $\dot{g}_l = 0$, $(\forall l \neq m)$. 则由 $g_\lambda$ 模糊测度的性质,有:

$$\dot{g}(A_k) = \dot{g}_k + \dot{g}(A_{k+1}) + \lambda \dot{g}_k \dot{g}(A_{k+1}) = \sum_{i=k}^{n-1} \dot{g}_i + \lambda \sum_{i=k}^{n-1} \sum_{j=i+1}^{n} \dot{g}_i \dot{g}_j + \cdots + \lambda^{n-k} \dot{g}_k \cdots \dot{g}_n = \begin{cases} 1, & k \leq m \\ 0, & k > m \end{cases}$$

$A_k$ 的定义同式(6). 设模糊测度 $\dot{g}$ 对应的融合函数为 $\dot{h}$,融合后的结果为:

$$\dot{h}(f(S)) = \dot{h}(f(s_1), \cdots, f(s_n)) = \sum_{i=1}^{n} [f(s_i) - f(s_{i-1})]\, \dot{g}(A_i) = \sum_{i=1}^{m} [f(s_i) - f(s_{i-1})] = f(s_m)$$

如果变化 $m$ 的值生成不同的模糊测度,就可以使融合后的结果和各单分类器的结果相同,即 $H$ 满足可分解性. 因此,

$$I_H(h^*) = \min_{h \in H} \int X - f_h(S)]^2\, dp_{f_h(S),X} \leq \int X - f(s_i)]^2\, dp_{f(s_i),X} = I_S(s_i)$$

可见,当 $H$ 满足可分解性时,总能从 $H$ 中选取一个融合函数,使得融合后的效果从均方误差意义上说比任何单分类器的分类性能要好.

## 3 基于遗传算法的模糊测度赋值

模糊测度的取值在模糊积分中起着决定性的作用,不同的模糊测度值将导致不同的分类结果. 在对模糊测度赋值时,常采用基于混淆矩阵的方法[4]. 设分类器(信息源) $s_i$ 的混淆矩阵为 $P_i = (p_i^{uv})$,与该分类器对应的关于类别(命题) $k$ 的模糊密度可由下式计算:

$$g_i^k = \left( \frac{1}{m-1} \sum_{j \neq k} (1 - p_i^{jk}) \right) p_i^{kk},\ m \text{ 为类别数} \quad (11)$$

模糊密度的不同取值将形成一个解空间,而基于混淆矩阵的模糊密度值很难与该空间中的最优值吻合. 在本文中,采用了基于遗传算法的优化过程来确定模糊密度的值.

用遗传算法确定模糊密度时,所有的模糊密度变量被编入同一条染色体中,并在开始时随机生成 $N$ 条染色体作为初始群体. 而各染色体对应的适应度函数由下式决定:

$$fit = 1 - \sqrt{\sum_{i}^{l} \sum_{j}^{m} (desired_{ij} - actual_{ij})^2 / l \cdot m} \quad (12)$$

$m$ 为类别数; $l$ 为样本数, $desired_{ij}$ 和 $actual_{ij}$ 分别表示对应于样本 $i$ 的期望输出和融合后的实际输出. 而实际输出是根据该染色体对应的模糊密度对分类器输出结果进行模糊积分得到的.

复制过程采用类似于赌轮盘的方法. 对每一染色体,可以计算一间隔区域 $r_i$[6]:

$$r_i = \left[ \frac{\sum_{i=0}^{i-1} fit_i}{\sum_{j=0}^{N-1} fit_j}, \frac{\sum_{i=0}^{i} fit_i}{\sum_{j=0}^{N-1} fit_j} \right) \quad (13)$$

$fit_j$ 为群体中与第 $j$ 个染色体对应的适应度. 当各染色体的间隔区域确定后,随机生成 $N$ 个实数 $\alpha_v$, $(0 < \alpha_v < 1, v = 0, 1, \cdots, N-1)$,则 $\alpha_v$ 的值将落在某个间隔 $r_w$ 中,而与该间隔对应的染色体 $c_w$ 将被作为种子保存在交配池中.

从交配池中选取两个染色体 $c_l$ 和 $c_{N-1-l}$ $(0 \leq l \leq (N-1)/2)$,新的染色体 $c_l'$ 和 $c_{N-1-l}'$ 可由交叉过程产生,如下所示:

$$c_l' = (c_l \cap M_x) \cup (c_{N-1-l} \cap \overline{M}_x), \quad c_{N-1-l}' = (c_l \cap \overline{M}_x) \cup (c_{N-1-l} \cap M_x) \quad (14)$$

其中,$M_x$ 为随机生成的掩码,$\overline{M}_x$ 为 $M_x$ 的补码.

交叉操作用来从解空间中生成新的解. 而在交叉之前,要对染色体进行变异操作. 变异操作是这样实现的:通常先设定

一个较小的变异概率 $P_m$(如介于 0.01 和 0.1 之间的数),对染色体上的每一个基因生成一个随机数,如果这个数小于 $P_m$ 的话,对该基因进行变异操作;否则保持该基因的值不便.

通过这些进化过程,初始群体将逐渐演变为一个适应度较好的群体. 当进化迭代次数大于预先设定的次数或群体中最好的样本对应的适应度大于预定值时,结束迭代过程,此时,群体中适应度最高的样本即为最终的结果.

## 4　模拟实验

设系统中有 3 个分类器,样本的类别数也为 3. 而每个类别中有 700 个样本可供使用,其中 400 个作为训练样本,300 个作为测试样本. 当给定一个样本时,各分类器的输出均为一个三维矢量,代表该分类器输出的样本属于各类别的后验概率的估计值(如 BP 神经网络的输出). 在本实验中,我们模拟生成分类器的输出结果,并假设各分类器关于 3 个类别的证据支持度分别服从某种分布,这些分布的参数如表 1～3 中所示.

表 1　分类器 1 对样本证据支持度的分布参数(高斯分布)

| 数据来源 \ 分类结果 | 类别 1 | | 类别 2 | | 类别 3 | |
|---|---|---|---|---|---|---|
| | 均值 | 方差 | 均值 | 方差 | 均值 | 方差 |
| 类别 1 | 0.7 | 0.1 | 0.47 | 0.15 | 0.38 | 0.2 |
| 类别 2 | 0.4 | 0.14 | 0.76 | 0.12 | 0.35 | 0.2 |
| 类别 3 | 0.41 | 0.1 | 0.25 | 0.15 | 0.7 | 0.14 |

表 2　分类器 2 对样本证据支持度的分布参数
(高斯分布 + 均匀分布)

| 数据来源 \ 分类结果 | 类别 1 | | | 类别 2 | | | 类别 3 | | |
|---|---|---|---|---|---|---|---|---|---|
| | 高斯参数 | | 均匀分 | 高斯参数 | | 均匀分 | 高斯参数 | | 均匀分 |
| | 均值 | 方差 | 布范围 | 均值 | 方差 | 布范围 | 均值 | 方差 | 布范围 |
| 类别 1 | 0 | 0.05 | (0.7 0.95) | 0 | 0.1 | (0.3 0.86) | 0 | 0.18 | (0.1 0.6) |
| 类别 2 | 0 | 0.16 | (0.05 0.63) | 0 | 0.06 | (0.61 0.95) | 0 | 0.1 | (0.01 0.7) |
| 类别 3 | 0 | 0.15 | (0.05 0.47) | 0 | 0.1 | (0.0 0.74) | 0 | 0.05 | (0.75 0.99) |

表 3　分类器 3 对样本证据支持度的分布参数(高斯分布)

| 数据来源 \ 分类结果 | 类别 1 | | 类别 2 | | 类别 3 | |
|---|---|---|---|---|---|---|
| | 均值 | 方差 | 均值 | 方差 | 均值 | 方差 |
| 类别 1 | 0.81 | 0.14 | 0.3 | 0.2 | 0.44 | 0.17 |
| 类别 2 | 0.37 | 0.19 | 0.83 | 0.12 | 0.41 | 0.23 |
| 类别 3 | 0.4 | 0.17 | 0.31 | 0.2 | 0.87 | 0.15 |

本实验比较了基于混淆矩阵和遗传算法的模糊密度赋值方法的融合结果. 由于有 3 个分类器,样本类别数也为 3,因此,模糊密度中参量个数为 9. 在遗传算法中,采用了所谓的 elitist 策略,即上一代中最优秀的染色体将被保存到下一代的群体中.

各分类器以及融合后的识别率-拒识率曲线如图 1 所示.

从图中可以看出,在用式(11)对模糊密度赋值时,融合结果比单分类器的效果要好. 然而,经过遗传算法优化后,能继续提高模糊积分的融合性能. 在用遗传算法优化模糊密度值时,由于要对群体中的所有染色体计算其适应度,而适应度的计算是以样本的分类结果为基础的,因此,训练过程要花费较长的时间. 本实验中,每个参量用一个 8 位二进制码来表示,群体大小和进化次数分别设为 28 和 40,训练时间大约为 21 分钟.

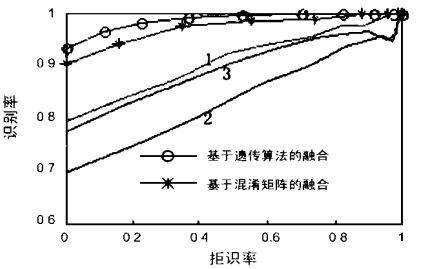


图 1　各分类器以及融合后的识别-拒识曲线. (图中标有 1、2、3 的三条曲线分别表示三个分类器的识别-拒识曲线)

## 5　结论

基于模糊积分的融合方法能达到比最优的单分类器更好的性能,然而,这并不意味着任意的模糊密度参数都可以提高系统的分类能力. 模糊密度值对系统的整体性能起着至关重要的作用. 用基于遗传算法的优化过程来确定模糊密度的值,可以得到最优解或接近于最优解的解. 从模拟实验中也可以看出,用遗传算法来计算模糊密度时,能得到更好的分类效果.

**参考文献**:

[1] D L Hall, J Llinas. An introduction to multi-sensor data fusion [J]. Proc. of the IEEE, 1997, 85(1): 6 - 23.

[2] N S V Rao. A fusion method that performs better than best sensor [C]. First Inter. Conf. on Multisource-Multisensor Information Fusion, Las Vegas, 1998: 19 - 26.

[3] T Pham, M Wagner. Similarity normalization for speaker verification by fuzzy fusion [J]. Pattern Recognition, 2000, 33: 309 - 315.

[4] J M Keller, P Gader, et al. Advances in fuzzy integration for pattern recognition [J]. Fuzzy Sets and Systems, 1994, 65(3): 273 - 283.

[5] M Grabisch, J M Nicolas. Classification by fuzzy integral: performance and tests [J]. Fuzzy Sets and Systems, 1994, 65(3): 255 - 271.

[6] C H Lin, J L Wu. Automatic facial feature extraction by genetic algorithms [J]. IEEE Trans. On Image Processing, 1999, 8(6): 834 - 845.

[7] K S Tang, K F Man, S Kwong, Q He. Genetic algorithms and their applications [J]. IEEE Signal Processing Magazine, November 1996: 22 - 37.

[8] T Murofushi, M Sugeno. An interpretation of fuzzy measure and the Choquet integral as an integral with respect to a fuzzy measure [J]. Fuzzy Sets and Systems, 1989, 29: 201 - 227.

**作者简介**:

**刘汝杰**　男. 1973 年生于山东寿光. 现为北方交通大学信息所博士研究生,主要研究方向为:信息融合,模式识别,图象处理等.