

一种自适应模糊密度赋值的决策层融合目标识别算法

付耀文, 黎湘, 庄钊文
(国防科技大学 ATR 重点实验室, 湖南长沙 410073)

摘要: 模糊积分是一种有效的决策层融合目标识别方法. 应用模糊积分进行融合的核心问题就是确定反映各传感器重要性的模糊密度. 目前确定模糊密度的方法主要有利用混淆矩阵或者采用遗传算法对模糊密度赋值. 这些方法在对模糊密度赋值时, 利用的都是训练样本的先验静态信息. 本文提出了一种利用训练样本先验静态信息结合各传感器判决包含的动态信息对模糊密度进行自适应动态赋值的方法. 仿真实验结果表明该方法能得到更好的融合效果.

关键词: 决策层融合; 模糊积分; 目标识别

中图分类号: TP391 **文献标识码:** A **文章编号:** 0372-2112 (2004) 09-1433-03

A Decision Level Fusion Algorithm for Target Recognition with Fuzzy Densities Determined Adaptively

FU Yao-wen, LI Xiang, ZHUANG Zhao-wen
(ATR Lab of National University of Defence Technology, Changsha, Hunan 410073, China)

Abstract: Fuzzy integral is an effective decision level fusion method for target recognition. The core problem in using fuzzy integral is to determine the fuzzy densities which can reflect the importance of the sensors. Presently, the main methods of determining fuzzy densities include using confusion matrix or genetic algorithm. These methods only use the a priori static information of the training samples to determine the fuzzy densities. A new method of determining fuzzy densities adaptively is presented in this paper, which uses the a priori static information of the training samples and the dynamic information contained in each sensor's decision. The simulation experiment shows that better fusion result can be obtained by using this method.

Key words: decision level fusion; fuzzy integral; target recognition

1 引言

决策层融合目标识别就是各传感器先根据各自所获得的信息分别进行目标识别, 建立对所观测目标的初步分类结论, 然后融合中心对各传感器处理的结果进行整合得到最后的判决结果. 由于信息本身的不完全性, 获得信息的不可靠性, 表达信息的不严密性, 运用信息的不成熟性, 以及各种信息的矛盾性, 各传感器根据所获取的有关事件或目标的信息给出的识别结果不可避免地具有很大的不确定性. 因此, 决策层融合目标识别实质上是一个不确定性处理问题^[1]. 由模糊集理论发展而来的模糊信息处理技术为不确定性处理和模拟人类识别机理提供了一种有效的手段, 在控制、信息融合、模式识别等领域已得到广泛应用^[2,3].

目前已有多种模糊方法应用于决策层融合目标识别, 其中最常采用的是模糊积分^[4~6]. 模糊积分在组合多源信息时不仅考虑各传感器给出的客观信息, 同时还考虑了各传感器子集的重要性. 模糊积分应用于决策层融合目标识别的核心问题是确定反映各传感器重要性的模糊密度. 目前确定各传感器模糊密度的方法主要有利用训练样本求出各传感器的混

淆矩阵来确定模糊密度或者利用训练样本采用遗传算法对模糊密度赋值. 这些方法在对模糊密度赋值时, 利用的都是训练样本的先验静态信息. 本文提出了一种利用训练样本先验静态信息结合各传感器判决包含的动态信息对模糊密度进行自适应动态赋值的方法. 仿真实验结果表明和非自适应模糊密度赋值方法相比该方法能得到更好的融合效果.

2 模糊测度与模糊积分

定义 1 设 S 为任意集合, $P(S)$ 表示 S 的幂集. 如果集合函数 $g: P(S) \rightarrow [0, 1]$, 满足:

- (1) $g(\emptyset) = 0, g(S) = 1$;
- (2) 如果 $A, B \in P(S)$, 且 $A \subseteq B \Rightarrow g(A) \leq g(B)$;
- (3) 如果 $\{A_i\}_{i=1}^n$ 是

递增可测集序列 $\Rightarrow g(\lim_{i \rightarrow \infty} A_i) = \lim_{i \rightarrow \infty} g(A_i)$.

则称 g 是 $P(S)$ 上的一个模糊测度.

两个不相交集的模糊测度一般不等于这两个子集模糊测度的和. 为此, Sugeno 定义了一类所谓的 g_A 模糊测度, 满足性质: $\forall A, B \in P(S)$, 且 $A \cap B = \emptyset$

$$g(A \cup B) = g(A) + g(B) + g(A)g(B), \quad g(A) > -1.$$

设集合 S 为有限集 $\{s_1, s_2, \dots, s_N\}$. 记 $g^i = g(\{s_i\})$, 映射 $s_i \rightarrow g^i$ 称为模糊密度函数.

定义 2 设 $h: S \rightarrow [0, 1]$ 是定义在集合 S 上的函数, 则函数 h 在集合 S 上关于模糊测度 g 的 Sugeno 模糊积分定义为:

$$h(s) \circ g(\cdot) = \int_{[0,1]} (h(s), g(F)) \quad (1)$$

其中 $F = \{s : h(s) \geq \cdot\}$.

当集合 $S = \{s_1, s_2, \dots, s_N\}$ 为一有限集, 不妨设 $1 \geq h(s_1) \geq h(s_2) \geq \dots \geq h(s_N) \geq 0$ (如果函数 h 不满足此关系, 只要将集合 S 的元素重新排列使得函数 h 满足此关系), 则函数 h 在集合 S 上关于模糊测度 g 的 Sugeno 模糊积分为:

$$h(s) \circ g(\cdot) = \bigwedge_{i=1}^N (h(s_i), g(A_i)) \quad (2)$$

其中 $A_i = \{s_1, s_2, \dots, s_i\}$.

如果 g 是 g 模糊测度, $g^i = g(\{s_i\})$, 则 $g(A_i)$ 可以由式 (3)、(4) 迭代计算得到^[4]:

$$g(A_1) = g(\{s_1\}) = g^1 \quad (3)$$

$$g(A_i) = g^i + g(A_{i-1}) - g^i g(A_{i-1}), \quad 1 < i \leq N \quad (4)$$

其中, 可通过求解方程 (5) 得到:

$$+1 = \prod_{i=1}^N (1 + g^i) \quad (5)$$

设总共有 N 个传感器 $\{s_1, s_2, \dots, s_N\}$, M 类目标 $\{1, 2, \dots, M\}$. 对一待识别目标 x , 传感器 s_i 输出的识别结果 $D_i(x) = [d_{i1}(x), d_{i2}(x), \dots, d_{iM}(x)]$ 是目标 x 属于各类目标的信度. 利用 Sugeno 模糊积分进行融合时, 集合 S 代表这 N 个传感器, 令 $h_j(s_i) = d_{ij}(x)$, $g_j(A_i)$ 表示传感器集 A_i 识别第 j 类目标的重要性, 则利用 Sugeno 模糊积分融合后目标 x 属于第 j 类目标的信度为:

$$d_j(x) = \bigwedge_{i=1}^N (h_j(s_i), g_j(A_i)) \quad (6)$$

如果系统要求输出一个目标类别, 则将最大的 $d_j(x)$ 对应的目标类别输出.

3 自适应模糊密度赋值

在利用模糊积分进行决策层融合目标识别时, 模糊密度起着决定性的作用. 不同的模糊密度赋值将形成不同的融合函数, 从而导致不同的融合识别结果. 目前一种常用的确定各传感器模糊密度的方法是基于混淆矩阵的方法. 对于 M 类目标, 传感器 s_i 的归一化混淆矩阵是一 $M \times M$ 矩阵 $P^i = (p_{kl}^i)_{M \times M}$, 其中 p_{kl}^i 表示传感器 s_i 将训练样本中第 k 类目标识别为第 l 类的比率. 由于模糊密度 $g_j^i = g_j(\{s_j\})$ 表示在判决目标属于类别 j 时传感器 s_i 的重要性, 模糊密度 g_j^i 可用式 (7) 估计^[3]:

$$g_j^i = \frac{1}{M-1} \sum_k (1 - p_{kj}^i) p_{ij}^i \quad (7)$$

另外一种常用的方法是采用遗传算法对模糊密度赋值^[6].

利用混淆矩阵或者遗传算法对模糊密度赋值时, 利用的都是训练样本的先验静态信息, 没有利用各传感器识别结果包含的动态信息. 训练结束后, 各传感器对应的模糊密度固定

不变, 这显然不太合理. 尤其是当训练条件与识别条件不一致即由训练样本所反映的各传感器的重要性与实际识别时各传感器的重要性不一致时, 融合识别效果可能会显著降低. 即使训练条件与识别条件一致, 各传感器在识别不同目标时的重要性也是不同的, 即其对应的模糊密度应该是变化的. 例如, 在各个传感器的特征空间, 来自不同类别的特征点形成各自区域, 这些不同类别对应的区域有可能发生重叠. 当待识别目标处于某传感器特征空间的重叠区域时, 反映在该传感器输出的识别结果上就是此待识别目标属于各类的信度相差不大, 该传感器很容易将其识别错. 因此, 在对此目标进行融合识别时, 该传感器的重要性应当比较低. 相反, 当待识别目标在某传感器的特征空间远离重叠区域时, 反映在该传感器输出的识别结果上就是此待识别目标属于某一类的信度很大而属于其他类的信度很小, 该传感器将其识别错的概率很小. 在对此目标进行融合识别时, 该传感器的重要性应当比较高. 因此, 在利用训练样本得到各传感器的模糊密度后, 应当根据各传感器对待识别样本的识别结果对其进行修正. 为此, 我们根据各传感器识别结果对每个传感器定义一个系数 $\alpha_i(x)$ ^[7]:

$$\alpha_i(x) = d_{ii}(x) - \frac{1}{M-1} \sum_{j=1}^M d_{ij}(x), \quad i=1, 2, \dots, N \quad (8)$$

其中 $d_{ii}(x) = \max_{j=1}^M (d_{ij}(x))$.

由式 (8) 可以看出, $\alpha_i(x)$ 表示了传感器 s_i 对目标 x 的识别结果的置信度, 也反映了目标 x 在传感器 s_i 的特征空间中处于非重叠区域的可能性. $\alpha_i(x)$ 越大, $d_{ii}(x)$ 就越大, $d_{ij}(x)$, $j=1, 2, \dots, M$ 且 $j \neq i$ 就越小, 那么传感器 s_i 对目标 x 的识别结果的置信度就越大, 目标 x 在传感器 s_i 的特征空间中处于非重叠区域的可能性也就越大. 根据 $\alpha_i(x)$ 对模糊密度自适应赋值过程如下:

(1) 首先根据训练样本采用式 (7) 或文献 [6] 介绍的遗传算法对各传感器的模糊密度确定初始值, 设为 g_j^i , $i=1, 2, \dots, N$, $j=1, 2, \dots, M$.

(2) 对待识别目标 x , 根据各传感器识别结果利用式 (8) 计算各传感器的系数 $\alpha_i(x)$.

(3) 将各传感器模糊密度的初始值乘于各自对应的系数 $\alpha_i(x)$ 作为对目标 x 进行融合识别时各传感器的模糊密度 g_j^{*i} , 即 $g_j^{*i} = g_j^i \times \alpha_i(x)$, 表示的含义是传感器对目标 x 的识别结果的置信度越低, 对目标 x 进行融合识别时, 该传感器的重要性越低.

4 仿真实验

设系统中有 3 个传感器, 目标的类别数也为 3. 我们模拟生成各传感器提取的特征向量. 设各传感器提取的目标特征向量维数为 2, 并且它们满足某种分布 (高斯分布或高斯分布与均匀分布的和), 其分布参数如表 1、2、3 所示. 依据这些分布, 对每类目标模拟生成 500 个样本的特征向量. 每类目标的前 200 个样本作为训练样本, 剩下的 300 个样本作为测试样本. 每个传感器均采用贝叶斯分类器进行分类, 当给定一个样

本时,各传感器的输出均为一个三维矢量,表示该传感器输出的样本属于各个类别的后验概率估计值。

本实验比较基于混淆矩阵、遗传算法的模糊密度赋值方法与自适应模糊密度赋值方法的融合识别结果。由于有 3 个传感器,3 类目标,因此总共有 9 个模糊密度参量。在基于遗传算法的模糊密度赋值方法中,我们对每个模糊密度参量用一个 8 位二进制码表示,9 个模糊密度参量编入一条染色体,每条染色体的适应度我们定义为根据该染色体对应的模糊密度对训练样本进行模糊积分融合识别的识别率,交换率设定为 0.7,变异率设定为 0.03,种群大小为 39。

表 1 传感器 1 提取的特征向量的分布参数

目标类别	高斯分布				均匀分布	
	m_x	m_y	σ_x	σ_y	x 范围	y 范围
1	0.8	1.3	0.5	0.4	[- 0.1 0.1]	[- 0.2 0.2]
2	0.3	1	0.4	0.5	—	—
3	1.1	0.7	0.4	0.5	—	—

表 2 传感器 2 提取的特征向量的分布参数

目标类别	高斯分布				均匀分布	
	m_x	m_y	σ_x	σ_y	x 范围	y 范围
1	0.5	1.3	0.4	0.3	—	—
2	0	0.9	0.4	0.5	[- 0.1 0.1]	[- 0.2 0.3]
3	0.6	0.8	0.4	0.4	—	—

表 3 传感器 3 提取的特征向量的分布参数

目标类别	高斯分布				均匀分布	
	m_x	m_y	σ_x	σ_y	x 范围	y 范围
1	0.6	2	0.8	1	—	—
2	1.2	1.5	0.7	0.5	—	—
3	1.7	0.8	0.4	0.9	[- 0.2 0.2]	[- 0.2 0.3]

各传感器识别率和基于混淆矩阵、遗传算法的模糊密度赋值方法及自适应模糊密度赋值方法的融合识别率如表 4 所示。表 4 中“混淆矩阵 + 自适应”表示自适应模糊密度赋值方法中的模糊密度初始值是利用混淆矩阵确定的,“遗传算法 + 自适应”表示自适应模糊密度赋值方法中的模糊密度初始值是利用遗传算法确定的。

表 4 各传感器识别率及各种模糊密度赋值方法的融合识别率

	目标 1	目标 2	目标 3	平均
传感器 1	0.622	0.684	0.712	0.673
传感器 2	0.696	0.632	0.602	0.643
传感器 3	0.646	0.630	0.662	0.646
混淆矩阵	0.750	0.708	0.868	0.7750
遗传算法	0.760	0.788	0.878	0.8087
混淆矩阵 + 自适应	0.816	0.786	0.880	0.8273
遗传算法 + 自适应	0.806	0.784	0.892	0.8273

由表 4 可以看出基于模糊积分的融合识别率比各传感器

均有较大程度提高。各种模糊密度赋值方法之间相比,基于遗传算法的模糊密度赋值方法比基于混淆矩阵的模糊密度赋值方法的融合平均识别率高 3.37%,但是利用遗传算法对模糊密度赋值需要较长的训练时间。利用混淆矩阵确定模糊密度初始值的自适应模糊密度赋值方法的融合平均识别率比基于混淆矩阵的模糊密度赋值方法提高 5.23%。利用遗传算法确定模糊密度初始值的自适应模糊密度赋值方法的融合平均识别率比基于遗传算法的模糊密度赋值方法提高 1.96%。可见自适应模糊密度赋值方法不仅利用训练样本的先验信息,而且利用各传感器识别结果包含的动态信息,能进一步提高融合识别效果。

5 结论

模糊密度的取值对基于模糊积分的融合识别性能起着重要的作用,利用混淆矩阵或遗传算法对模糊密度赋值可以得到比最优传感器更好的识别性能,根据各传感器识别结果包含的动态信息对由混淆矩阵或遗传算法确定的模糊密度进行修正后能进一步提高融合识别效果。

参考文献:

- [1] 雍少为. 信息融合的基本理论及其在自动目标识别中的应用[D]. 长沙:国防科技大学,1997.
- [2] 郭桂蓉,庄钊文. 信息处理中的模糊技术[M]. 长沙:国防科技大学出版社,1993.
- [3] J M Keller, P Gader, H Tahani, J H Chiang, M Mohamed. Advances in fuzzy integration for pattern recognition[J]. Fuzzy Sets and Systems, 1994(65):273 - 283.
- [4] H Tahani, J M Keller. Information fusion in computer vision using the fuzzy integral[J]. IEEE Trans on SMC,1990,20(3):733 - 741.
- [5] S B Cho, J H Kim. Combining multiple neural networks by fuzzy integral for robust classification[J]. IEEE Trans on SMC,1995,25(2):380 - 844.
- [6] 刘汝杰,袁保宗,唐晓芳. 用遗传算法实现模糊测度赋值的一种多分类器融合算法[J]. 电子学报,2002,30(1):145 - 147.
- [7] C C Chibelushi, F Deravi, J S D Mason. Adaptive classifier integration for robust pattern recognition[J]. IEEE Trans on SMC-Part B: Cybernetics, 1999,29(6):902 - 907.

作者简介:

付耀文 男,1976 年 5 月生于江西进贤,国防科技大学电子科学与工程学院博士研究生,感兴趣的领域包括信息融合、雷达目标识别。
 黎湘 男,1967 年 9 月生于湖南长沙,博士,副教授,主要研究领域包括精确制导、信息融合、非线性信号处理等。
 庄钊文 男,1958 年 6 月生于福建南安,教授,博士生导师,主要研究领域包括信号处理、自动目标识别、模糊技术等。