

Choquet 模糊积分的粗糙性及信息融合

管 涛, 冯博琴

(西安交通大学电子与信息工程学院, 710049, 西安)

摘要: 在分析现有模糊积分理论的基础上,提出了 Choquet 上、下模糊积分的概念和粗糙区间数的积分形式,给出了 Choquet 模糊积分的粗糙特性,建立了基于粗糙 Choquet 模糊积分形式和信息融合方法.该方法在存在各种外界干扰因素的多传感器融合环境中,可以较好地实现区间数传感器的选择和融合.所给 Choquet 上、下模糊积分是经典 Choquet 模糊积分在粗糙意义下的扩展形式,能处理粗糙区间数数据源和信息系统的对象.模拟实例验证了融合方法的有效性.

关键词: 模糊积分;粗糙集;信息融合;区间数

中图分类号: TP18 **文献标识码:** A **文章编号:** 0253 - 987X(2004)12 - 1251 - 05

Roughness of Choquet Fuzzy Integral and Information Fusion

Guan Tao, Feng Boqin

(School of Electronics and Information Engineering, Xi an Jiaotong University, Xi an 710049, China)

Abstract: Based on the analysis of existed fuzzy integral theories, the notions of the Choquet upper and lower fuzzy integral and rough interval numbers' integral form were proposed; the rough properties of Choquet fuzzy integral and an information fusion method based on rough Choquet fuzzy integral were given. The method can properly carry out the selection and fusion of interval number sensors in multi-sensor fusion environment in which various disturbance factors exist. The Choquet upper and lower fuzzy integral is the extended form of the classical Choquet one in a sense of roughness, which can deal with the data sources of rough intervals and objects of information systems. The simulated experiments have validated the effectiveness of the method.

Key words: fuzzy integral; rough sets; information fusion; interval number

近些年来,建立在不确定信息度量之上的积分在自动模式识别、图像处理、信息融合中得到了广泛的发展和应用^[1-4].根据不同测度,不确定性积分有模糊积分^[5,6]和粗糙积分^[7]两种形式.模糊积分主要用于决策支持^[8,9]、自动控制^[10]等. Sugeno 的模糊积分是定义在模糊测度上的非线性函数,特点是直接排除了次要因素的影响,与加权平均相比,强化了主要因素的作用,但却完全忽视次要因素的影响;Choquet 模糊积分^[8]考虑了各种影响因素,以避免 Sugeno 模糊积分的缺陷,而广义 Choquet 模糊积分及其在信息融合中的应用近来得到了重视和应

用^[4].粗糙积分是建立在粗糙测度之上的积分形式^[11].

模糊积分、粗糙积分模型建立在实数特征值基础上^[12],但在应用中有许多特征值并不能精确测定,只能给出其上、下界估计,而粗糙集则为这类问题提供了新的解决思路.在粗糙集理论中,对象之间的不可分辨性通过对象空间划分的粗糙性来衡量,而模糊性则度量了集合边界之间的不确定性,这 2 种不确定性从不同角度刻画了系统的特性.本文通过分析 Choquet 模糊积分的性质,建立了区间数多传感器的融合方法,为区间数传感器提供了一种选

收稿日期: 2004 - 02 - 23. 作者简介: 管 涛(1974 ~),男,博士生;冯博琴(联系人),男,教授,博士生导师. 基金项目: 国家高技术研究发展计划资助项目(2003AA1Z2610).

择策略.

1 模糊测度与 Choquet 模糊积分

给定论域 $X = \{x_1, x_2, \dots, x_n\}$, 由它得到的博雷尔集记为 \mathbf{B} , 在 \mathbf{B} 上的一个模糊测度定义^[5]为 $g: \mathbf{B} \rightarrow [0, 1]$, 使得 $g(X) = 1, g(\emptyset) = 0$, 而若 $A, B \in \mathbf{B}, A \subset B$, 则 $g(A) \leq g(B)$, 且若 $A_1 \subset A_2 \subset \dots \subset A_n \subset \dots$, 则 $\lim(g(A_i)) = g(\bigcup A_i)$.

一类常用的模糊测度为 g, μ, ν, \dots , 对于 $A, B \in \mathbf{B}$, 它满足 $g(A \cup B) = g(A) + g(B) - g(A \cap B)$. 已知 $g^i = g(\{x_i\})$ 为模糊密度, 对于 $A_i \in \mathbf{B}, g(A_i)$ 可以通过递推公式^[6]

$$g^i + 1 = \sum_{j=1}^i (1 + g^j), \quad g^i > -1$$

$g(A_i) = g(\{x_1\}) = g^1$
 $g(A_i) = g^i + g(A_{i-1}) - g^i g(A_{i-1}), \quad 1 \leq i \leq n$
得到. Choquet 模糊积分具有如下形式^[8]

$$hdg = \sum_{i=1}^n g(X_i) [h(x_i) - h(x_{i+1})]$$

其中, $h(x_{n+1}) = 0, g(x_0) = 0, h(x_1) \leq h(x_2) \leq \dots \leq h(x_n), X_i = \{x_1, \dots, x_i\}$. 该积分的另一形式为

$$hdg = \sum_{i=1}^n h(X_i) [g(x_i) - g(x_{i+1})]$$

2 区间数及其运算

根据文献 [13], 区间数来源于这样的事实: 根据近似数 A 以及它的严格边界 B 来估计未知数 X , 使得 $|X - A| \leq B$.

性质 1^[13] 区间数 $[a, b], [c, d], (a, b, c, d) \in \mathbf{R}$, 并满足

$$[a, b] + [c, d] = [a + c, b + d]$$

$$[a, b] - [c, d] = [a - d, b - c]$$

$$[a, b][c, d] = [\min(ac, ad, bc, bd), \max(ac, ad, bc, bd)]$$

$$[a, b] / [c, d] = [a, b] [1/d, 1/c], \quad 0 \notin [c, d]$$

在许多情况下, 根据已有的知识仅了解未知模糊集合 μ 的粗糙近似 $[\underline{R}\mu, \overline{R}\mu]$, 而在每个等价类中的对象具有确定的上、下界.

3 离散粗糙 Choquet 模糊积分

下面给出离散粗糙 Choquet 模糊积分的概念和描述形式.

定义 1 假设 R 是模糊测度空间 (X, \mathbf{B}, g) 上的等

价关系, 模糊集合 μ 是模糊可积的, 它的粗糙上、下 Choquet 模糊积分为

$$R\mu dg = \sum_{i=1}^r (M_i - M_{i-1}) g(\overline{H_i}) \quad (1)$$

$$\underline{R}\mu dg = \sum_{i=1}^r (m_i - m_{i-1}) g(\underline{H_i}) \quad (2)$$

而 $\frac{X}{R} = \{C_i: i=1, \dots, r\} = \{C_i: i=1, \dots, r\}, m_i = \inf\{\mu(x) : x \in C_i\}, M_i = \sup\{\mu(x) : x \in C_i\}, \overline{H_i} = \{C_k: i=1, \dots, k \leq r\}, \underline{H_i} = \{C_k: i=1, \dots, k \leq r\}$, 并且 $0 = m_0 = M_0, \dots, m_r = 0 = M_r$. 由于 Choquet 模糊积分具有单调性, 因此 $[\underline{R}\mu dg, \overline{R}\mu dg]$ 构成了 Choquet 模糊积分的一个粗糙近似.

同时, $[\underline{R}\mu dg, \overline{R}\mu dg]$ 被称之为 μ 的粗糙区间数.

性质 2 离散粗糙 Choquet 模糊积分具有如下性质.

(1) 若 $\mu = c$ (常数), 则 $\underline{R}\mu dg = \overline{R}\mu dg = c$.

(2) 若 $\mu(x_i) = c_i, x \in C_i, i=1, \dots, r$, 则 $\underline{R}\mu dg = \overline{R}\mu dg$.

(3) 由 $R\overline{R}\mu = \underline{R}R\mu$ 和 $R\underline{R}\mu = \overline{R}R\mu$ 知

$$R\overline{R}\mu dg = \underline{R}R\mu dg = \underline{R}\mu dg$$

$$\underline{R}R\mu dg = \overline{R}\mu dg = \overline{R}\mu dg$$

(4) 由 $\underline{R}(1 - \mu) = R\mu$ 和 $\overline{R}(1 - \mu) = \underline{R}\mu$ 知

$$\underline{R}(1 - \mu) dg = R\mu dg$$

$$\overline{R}(1 - \mu) dg = \underline{R}\mu dg$$

(5) 由 $\underline{R}\mu \leq R\mu$ 知 $\underline{R}\mu dg \leq R\mu dg$

(6) 若 $\mu \leq \nu$, 则 $\underline{R}\mu dg \leq \underline{R}\nu dg, R\mu dg \leq R\nu dg$.

(7) 若 $A \subset B$, 则 $\underline{R}_A \mu dg \leq \underline{R}_B \mu dg, \overline{R}_A \mu dg \leq \overline{R}_B \mu dg$.

$\underline{R}_B \mu dg$.

(8) 若 $A, B \subset X$, 则

$$\underline{R}_{A \cup B} \mu dg = \underline{R}_A \mu dg + \underline{R}_B \mu dg$$

$$\overline{R}_{A \cup B} \mu dg = \overline{R}_A \mu dg + \overline{R}_B \mu dg$$

(9) 若 $A, B < X$, 则

$$\begin{aligned} & \int_A \int_B R\mu dg & \int_A R\mu dg \\ & \int_A \int_B R\mu dg & \int_B R\mu dg \end{aligned}$$

证明 由 Choquet 模糊积分具有单调性^[6],同时根据粗糙逻辑的相关概念^[14],可证得离散粗糙 Choquet 模糊积分的上述性质成立.

4 多源传感器选择方法

图 1 给出了一种具有多个传感器的信号采集系统模型. 由于存在时间尺度差异和噪声,从传感器得到的值不一定是实数,而是数值区间,此时经典的积分形式就不再适用. 使用粗糙 Choquet 模糊积分,可以较好地实现区间数多传感器的选择问题. 实际上,选择多源传感器也是选择对象的特征.

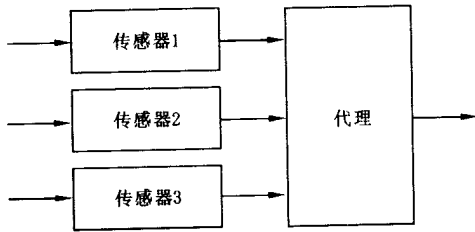


图 1 具有多个传感器输入的代理示意图

定义 2 假设指定积分值区间 $D = [a, b]$, 则基于粗糙 Choquet 模糊积分的传感器选择式为

$$= |D \cap I_A| / |I_A| \quad (3)$$

式中: $D \cap I_A$ 为传感器的积分数值区间, 显然 $[0, 1]$.

利用式(3)可以选择目标范围内的传感器, 若 $I_A \cap D = 1$, 则目标传感器被选中, 若 $|D \cap I_A| \ll |I_A|$, 则目标传感器被选中的概率很小, 这种选择可以排除位于指定区间之外的那些传感器. 当传感器生成的数据是实数时(见图 2), 则上、下粗糙 Choquet 积分与 Choquet 模糊积分相等.

使用高斯函数标准化生成的 3 种模拟传感器的曲线如图 2 所示, 图中取均值为 0.165, 方差分别为 0.120、0.115 和 0.140. 假定 $[x, x + 0.11]$, $x \in \{i/10 : 0 \leq i < 10, i \in N\}$ 构成了等价分割, 则在离散点处得到的积分值为

$$h_1 dg = 0.1778, \quad h_2 dg = 0.1490, \quad h_3 dg = 0.1375$$

当指定数值区间为 $[0.145, 0.150]$ 时, 可知传感器 2 满足条件. 若输入数据是不确定的, 仅可得到精确

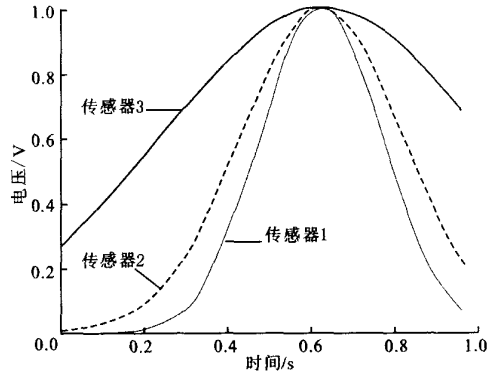


图 2 3 种传感器在某区段内的状态曲线

值的近似估计, 则离散 Choquet 模糊积分就不再适用.

5 信息融合模型

对象的特征信息, 如图像亮度、纹理, 可通过传感器的参数采样来获得, 由此得到的特征值并不能保证是数值, 它们可能是一些置信区间.

下面以粗糙 Choquet 模糊积分形式来构造信息融合模型, 并融合粗糙区间数信息. 该模型首先利用被选中的特征 f_i 产生特征信息 $[x_i, x_i]$, 然后利用模糊集合 μ 在这些特征上的取值得到置信区间 $[\underline{R}\mu(x_i), \overline{R}\mu(x_i)]$, 特别是 $\mu = 1$. 利用得到的模糊密度 g^i 可以给出对象的融合结果, 如图 3 所示, 通过与目标的重叠结果可以决定选中的概率. 该模型扩展了现有模型的适用领域.

6 模拟数值实验

实验分为两部分, 首先进行多传感器的选择, 然后对选中的传感器进行信息融合. 假定每个传感器数据是根据时间轴上等价划分得到的区间数. 给定等价划分与模糊密度, 模糊积分融合方法通过单趟扫描逐个处理对象, 实现数值对象分类. 除了模糊积分本身的复杂性之外, 该方法的时间复杂度为 $O(n)$. 图 4 给出了 3 个传感器在 $[0, 20]$ 时间区段内的波动区域. 表 1 给出了它们的值和积分计算结果, 其中假定 $D = (0.1119, 0.1130)$.

从表 1 的值可见, 传感器 1 被选中的可能性大, 传感器 3 被选中的可能性小. 与人工神经网络相比, 该方法需要先验信息, 但有两种方式可得到模糊密度值: 一是根据专家经验估计, 它们反映了专家或者经验数据对于对象特征的重视程度, 如图像融合中的“色泽”比“亮度”重要, 则“色泽”的密度值就高;

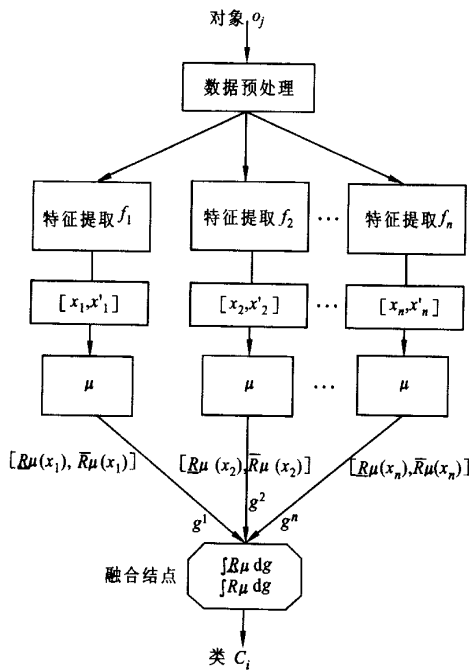


图3 粗糙区间数融合模型

表1 3种模拟传感器的各时刻数据和计算结果

等价类	传感器1 电压/V		传感器2 电压/V		传感器3 电压/V	
	最大	最小	最大	最小	最大	最小
1	01128	01119	01132	01121	01140	01117
2	01126	01120	01135	01118	01136	01120
3	01128	01118	01129	01121	01134	01119
4	01127	01118	01132	01120	01137	01117
5	01127	01120	01131	01119	01138	01119
6	01128	01119	01134	01118	01144	01125
7	01128	01121	01132	01120	01141	01120
8	01128	01119	01132	01120	01137	01122
9	01128	01119	01129	01116	01138	01119
10	01127	01118	01132	01118	01139	01118
11	01126	01119	01135	01117	01137	01120
12	01127	01118	01132	01119	01138	01118
13	01128	01120	01130	01120	01136	01118
14	01128	01121	01133	01122	01136	01119
15	01128	01119	01130	01121	01140	01118
16	01127	01120	01129	01121	01136	01119
17	01128	01119	01130	01119	01139	01118
18	01127	01120	01130	01120	01134	01121
19	01126	01120	01133	01121	01139	01119
20	01126	01118	01132	01120	01135	01117
	01972		01845		01563	

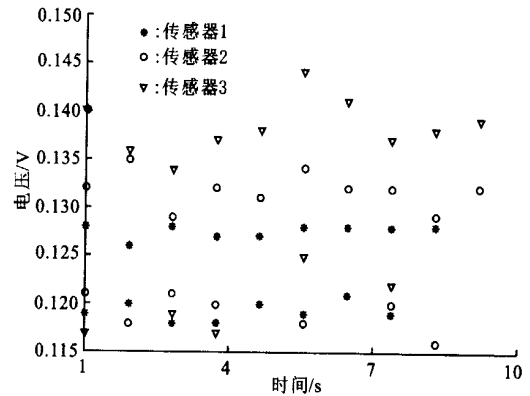


图4 3种传感器在前10个等价分割情况下的取值

二是从已有数据中训练学习得到密度值,包括一些常用的优化技术^[10]。为了简便起见,在融合模型中考虑由3对高斯函数生成模拟信号曲线,分别使用1倍和1125倍的原始数据在[0,1]区间进行标准化,它们的模糊密度如表2所示,表3给出了它们的均值与方差,融合结果如图5所示。

表2 3种传感器的模糊密度值

g^1	g^2	g^3
017	011	011

表3 3种模拟传感器的参数比较

对象	均值	方差
传感器1	0127	01012
传感器2	0175	01014
传感器3	0194	01017

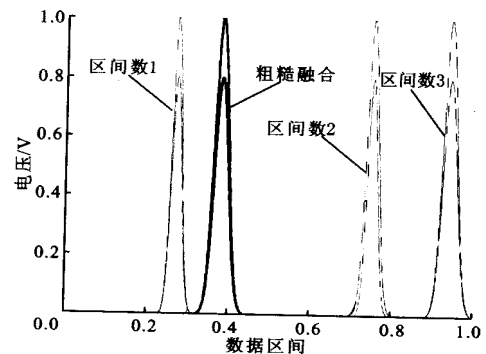


图5 粗糙区间数的融合

7 结论

与证据理论、隐马尔科夫过程(HMM)相比,在手写体识别上,模糊积分可有效提高识别率^[11]。与权重平均相比,在图像融合评价方面它更接近真实的结果^[3]。

参考文献:

- [1] Cader P D, Mohamed M A, Keller J M. Fusion of hand written word classifiers [J]. Pattern Recognition Letters, 1996, 17(6): 577 - 584.
- [2] Tahani H, Keller J M. Information fusion in computer vision using the fuzzy integral [J]. IEEE Trans on Systems, Man, and Cybernetics, 1990, 20(3): 733 - 741.
- [3] 杨 焯, 裴继红, 杨万海. 基于模糊积分的融合图像评价方法 [J]. 计算机学报, 2000, 24(8): 5 - 8.
- [4] Auephanwiriyaikul S, Keller J M, Cader P D. Generalized Choquet fuzzy integral fusion [J]. Information Fusion, 2002, 3(1): 69 - 85.
- [5] Sugeno M. Fuzzy measures and fuzzy integrals: a survey [A]. Fuzzy Automata and Decision Process [C]. New York: North-Holland, 1977. 89 - 102
- [6] Murofushi T, Sugeno M. A theory of fuzzy measures: representations, the Choquet integral, and null sets [J]. J Math Anal Appl, 1991, 159(2): 532 - 549.
- [7] Pawlak Z, Peters J F, Skowron A, et al. Rough measures: theory and applications [J]. Bulletin of International Rough Set Society, 5(1 - 2): 177 - 183.
- [8] Chiang J H. Aggregating membership values by a Choquet fuzzy integral based operator [J]. Fuzzy Sets and Systems, 2000, 114(3): 367 - 375.
- [9] Grabisch M, Sugeno M. Multiattribute classification using fuzzy integral [A]. The First IEEE Conference on Fuzzy Systems, San Diego, USA, 1992.
- [10] Keller J M, Osborn J. Training the fuzzy integral [J]. International Journal of Approximate Reasoning, 1996, 15(1): 1 - 24.
- [11] Borkowski M, Peters J F. Approximating sensor signals: a rough set approach [A]. Canadian Conference on Electrical and Computer Engineering, Winnipeg, Canada, 2002.
- [12] Grabisch M, Dispot E. A comparison of some methods of fuzzy classification on real data [A]. Second International Conference on Fuzzy Logic and Neural Networks, Lizaoka, Japan, 1992.
- [13] Moore R E. Interval analysis [M]. New York: Prentice Hall, 1966.
- [14] Lin T Y, Liu Q. First order rough logic: approximate reasoning via rough sets [J]. Fundam Inform, 1996, 27(2 - 3): 137 - 153. (编辑 苗 凌)

(上接第 1231 页)**参考文献:**

- [1] Ammann P, Wijesekera D, Kaushik S. Scalable, graph based network vulnerability analysis [A]. 9th ACM Conference on Computer and Communications Security, Washington DC, 2002.
- [2] Deraison R. Nessus [EB/OL]. <http://www.nessus.org>, 2003 - 02 - 15.
- [3] Stardust. 计算机网络系统安全漏洞分类研究 [EB/OL]. <http://www.xfocus.net/articles/200103/126.html>, 2003 - 03 - 03/2003 - 11 - 13.
- [4] Bishop M, Bailey D. A critical analysis of vulnerability taxonomies [R]. Technical Report 96 - 11. Davis, USA: Department of Computer Science, University of California, 1996.
- [5] Pawlak Z. Rough sets [J]. International Journal of Computer and Information Science, 1982 (11): 341 - 356.
- [6] Mell P. ICAT metabase: your CVE vulnerability search engine [EB/OL]. <http://icat.nist.gov/icat.cfm>, 2003 - 08 - 16. (编辑 苗 凌)

(上接第 1250 页)

- [3] Cabrera J B D, Lewis L, Qin Xinzhou, et al. Proactive detection of distributed denial of service attacks using MIB traffic variables: a feasibility study [A]. Proceedings of 2001 International Symposium on Integrated Network Management [C]. Piscataway, USA: IEEE Press, 2001. 6092622.
- [4] Kulkarni A B, Bush S F, Evans S C. Detecting distributed denial of service attacks using Kolmogorov complexity metrics [R]. Technical Report 1CRD176. New York: Research and Development Center, General Electric Company, 2001.
- [5] Tao Peng, Leckie C, Ramamohanarao K. Protection from distributed denial of service attacks using history based IP filtering [A]. IEEE International Conference on Communications, Anchorage, USA, 2003.
- [6] Cohen L. Time-frequency analysis [M]. Englewood Cliffs, USA: Prentice Hall, 1995.
- [7] Duda R O, Hart P E, Stork D G. Pattern classification [M]. 2nd ed. New York: Wiley, 2001.
- [8] MIT Lincoln Laboratory. 2000 DARPA intrusion detection scenario specific data sets [EB/OL]. <http://www.ll.mit.edu/IST/>, 2003 - 10 - 21. (编辑 苗 凌)