第16卷 第5期
2006年5月
计算机技术与发展
COMPUTER TECHNOLOGY AND DEVELOPMENT
Vol.16 No.5
May 2006

领域信息在信息融合进程中的优化处理

方宏彬[1,2],张 铃[2]
(1.安徽大学 数学系,安徽 合肥 230039;
2.安徽大学 教育部智能计算与信号处理重点试验室,安徽 合肥 230039)

摘 要:提出一种新的局部领域信息最优化组合函数算法,以解决融合过程中局部领域最优信息合理利用问题。引入最小鉴别信息熵的概念后,得到局部领域最小鉴别信息熵的概念与表示方法,应用于信息融合,得到局部领域信息最优化组合函数算法。该算法为避免其他融合方法主要考虑整体的结果而忽略具体过程提供了一种新的方法和思路。

关键词:多传感器系统;Kullback熵;信息融合
中图分类号:TP18 **文献标识码**:A **文章编号**:1673-629X(2006)05-0031-02

Optimization of Region Information in Process of Information Fusion

FANG Hong-bin[1,2], ZHANG Ling[2]
(1.Mathematics Department of Anhui Univ., Hefei 230039, China;
2.Education Ministry Key Lab. of Intelligence Computing & Signal Processing of Anhui Univ., Hefei 230039, China)

Abstract: Introduce the conception of the least discrimination information entropy and put forward the least discrimination information entropy of local region as well as its representation. Apply the method to information fusion and obtain a combinatory function algorithm about it. The algorithm is not only simple but also shuns the complicated relation of evidence theory as well as other fusion methods considering whole results.

Key words: multi-sensor system; Kullback entropy; information fusion

0 引 言

信息融合是一门综合性、交叉性很强的学科,其广阔的理论背景与应用价值引起各国学者的高度关注。以往的信息融合方法一般注重全局的结果的度量,而对融合进程中的信息优化处理未加重视。目前,信息融合方法用的较多的证据理论及其改良方法[1~3]的缺点是从整体上进行考虑,而对证据本身的可靠程度未加考虑;熵模型方法[4,5]同样是从整体概率分布角度考虑融合前后熵的变化,其他方法也很少涉及局部信息优化问题。从广义的哲学层面上来说,结果与过程之间是一种互动的关系,二者辩证发展。就信息融合过程来说,其过程中的局部领域信息优化对最后结果的影响有没有,有多大,以及如何互动作用?文中就此问题在多传感器信息融合系统下进行分析,提出了一种局部领域信息优化改进信息融合结果的方法。

收稿日期:2005-09-07
基金项目:安徽省青年科研基金(2004jq103)
作者简介:方宏彬(1972-),男,安徽池州人,博士研究生,研究方向为智能计算、信息融合;张 铃,教授,博士生导师,主要从事人工智能等领域的研究。

1 Kullback熵与最小鉴别信息函数理论简介

(1)设随机变量 X 在假设 H_1 与假设 H_2 下分别有概率密度函数 $p_1(x)$ 与 $p_2(x)$,定义Kullback熵如下[6]:

$$d(p_1,p_2)=\begin{cases}\int_x p_1(x)\log\left[\dfrac{p_1(x)}{p_2(x)}\right]\mathrm{d}x,\text{连续时}\\ \sum_x p_1(x)\log\left[\dfrac{p_1(x)}{p_2(x)}\right],\text{离散时}\end{cases}\quad(1)$$

(2)最小鉴别信息函数[6]。

利用Kullback熵,在一概率密度函数类中,对于已知概率密度函数 $p(x)$ 来说,最接近它的概率密度函数按以下方式确定:

$$d(p(x),f^*(x))=\underset{\forall f(x)\in \Gamma}{\arg\min}\, d(p,f)\quad(2)$$

由于鉴别信息函数 $d(p,f)$ 当 $p(x)$ 不变时关于 $f(x)$ 是下凸函数,从理论来说具有全局最优解。

2 局部领域信息最优化的表示

从(1)、(2)两式的表达方式可知,这是在整体上以此距离来度量寻找最优解。对于已知概率密度函数 $p(x)$,多传感器系统分别获得的概率密度函数为 $p_1(x),p_2(x),\cdots,p_n(x)$,由于受各种因素的干扰既无法确定相互独立性以及无法确定信源的重要程度,又不能说明传感器本身的

优劣。如从(1)、(2)式,可得出在整体上的最优解,这是从不同结果中找到最优结果。这种做法有其合理性,但也有不合理处:在不确定因素条件下,多传感器系统分别获得的概率密度函数 $p_1(x),p_2(x),\cdots,p_n(x)$ 各有千秋,不同时期以及不同范围区域接近 $p(x)$ 的程度不同,因此如何利用部分区域上最优接近信息从而使不同概率密度函数 $p_1(x),p_2(x),\cdots,p_n(x)$ 对 $p(x)$ 的接近程度有所贡献是合理的(见图 1)。

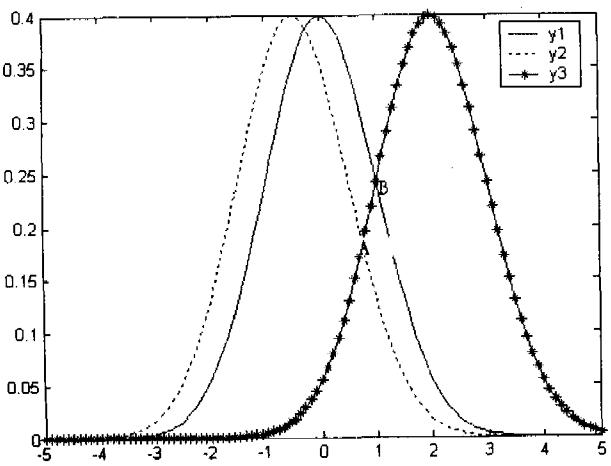

图 1 概率密度函数曲线图

图 1 中两传感器获得的概率密度函数分别为:y_2,y_3;真实概率密度函数为 y_1,函数 y_2 从整体上按推理网络或证据理论是函数 y_1 的最优接近。但从局部来看,函数曲线 y_3 上的 AB 段比函数 y_2 的相应部分与函数 y_1 的相应部分要接近,因此也应该合理利用此局部领域信息。

基于此,提出局部领域信息最优化组合函数算法。

以下是局部领域信息最优化组合函数算法:

(1) 局部最小鉴别信息熵。

对于辨别框架(X,σ_x,p),其中 X 为考察对象的论域,$p(x)$ 为其上真实概率分布,取 $X=\bigcup_{i=1}^{k}A_i$, $A_i\cap A_j=\varnothing$, $i\neq j$, $i,j=1,2,\cdots,k$。 (3)

定义:局部最小鉴别信息熵 $D_{A_i}(p_i{}^*//p)=\min\limits_{j=1,2,\cdots,n}\int_{A_i}p_j(x)\log\frac{p_j(x)}{p(x)}\mathrm{d}x$ (4)

则按(3)式对论域 U 的划分在多传感器系统分别获得的概率密度函数为 $p_1(x),p_2(x),\cdots,p_n(x)$,由(4)式融合后的表达式为:

$p^*(x)=\sum_{i=1}^{k}p_i^*(x)I(A_i)$, 其中 $I(A_i)=\begin{cases}1,x\in A_i\\0,x\notin A_i\end{cases}$,归一化为:取 $M=\int_x p^*(x)\mathrm{d}x$,

$$p^{\oplus}(x)=\frac{1}{M}p^*(x) \tag{5}$$

注:对于离散概率分布,方法完全一样,即积分号换成求和号。

(2) 局部领域信息最优化组合函数算法表示。

a. 对论域 X 按(3)式进行划分,整数 k 适当选取。

b. 对多传感器系统分别获得的概率密度函数 $p_1(x),p_2(x),\cdots,p_n(x)$ 按(4)式分别在 $A_i,i=1,2,\cdots,k$ 上计算求解 $p_i^*(x)$,再按(5)式合成 $p^{\oplus}(x)$。

c. 每一 $A_i,i=1,2,\cdots,k$ 的细度(即其连续分布情形下的区间长度或离散分布情形下的元素个数)相同,且小于给定阈值 $\varepsilon(\varepsilon>0)$ 时,k 不变(当然细度越小,有用信息的利用越有效,因此精度越高,但不可能无限逼近,因为截获的概率密度函数误差总存在而且传感器数毕竟有限),计算 $p^{\oplus}(x)$ 第 j 次,记为 $p_j^{\oplus}(x)$,这样循环整数 N 次终止,最后输出 $p_k^{\oplus}(x)=\frac{1}{N}\sum_{j=1}^{N}p_j^{\oplus}(x)$。

d. 整数 $k\leftarrow k+1$,再执行以上操作,直到输出的 $p_k^{\oplus}(x)$ 平稳为止,即:给定阈值 $\delta,\delta>0$,得到$\int_x|p_k^{\oplus}(x)-p_{k+1}^{\oplus}(x)|\mathrm{d}x\leqslant\delta$。

3 先验信息未知情形下局部领域信息最优化表示

对于不同传感器,其对待识别框架$(U,2^U,P)$为有限情形下截获数据所形成的经验分布函数分别为:$F_1(x),F_2(x),\cdots,F_n(x)$。由最大熵原理,取 $P(x)=\frac{1}{|U|}$,$\forall x\in U$,其中 $|U|$ 表示 U 中元素个数,其对应分布函数为 $F(x)$。取鉴别信息函数为:$D_{A_j}(F_i(x)//F(x))=\int_{A_j}F_i(x)\log\frac{F_i(x)}{F(x)}\mathrm{d}x$,融合步骤同上文。

4 结论

针对多传感器观察同一状态的系统,提出一种新的局部领域信息最优化组合函数算法,以解决融合过程中局部领域最优信息合理利用问题,避免证据理论中的证据之间的复杂关系与证据的可靠性以及其他融合方法主要考虑整体的结果而忽略具体过程,提供了一种新的方法和思路。

参考文献:

[1] Shafer G. A Mathematical Theory of Evidence[M]. Princeton, NJ: Princeton University Press, 1976. 133 - 185.

[2] Josang A. The consensus operator for combining beliefs[J]. Artificial Intelligence, 2002, 141: 157 - 170.

[3] 孙怀江,胡钟山,杨静宇. 基于证据理论的多分类器融合方法研究[J]. 计算机学报, 2001, 24(3): 231 - 235.

[4] Pomorski D. Entropy based optimization for binary detection networks[A]. In: proceedings of 2000 International conference on Information Fusion[C]. Paris, France: [s. n.], 2000. 1194 - 1201.

[5] Fassinut - Mombot B, Jean - Bernard C. A new probabilistic and entropy fusion approach for management of information sources[J]. Information Fusion, 2004, 5: 35 - 47.

[6] Kullback S. Information Theory and Statistics[M]. New York: Willey, 1959.