

doi: 10.3969/j.issn.1006-1576.2010.11.013

## P 距离确定信息量的证据合成方法

项浩, 姜礼平, 苏思

(海军工程大学 理学院, 湖北 武汉 430033)

**摘要:** 针对 D-S 证据理论在对高度冲突的证据进行融合时可能出现与直觉相悖的情况, 在对一些经典改进方法分析的基础上, 提出一种基于 P 距离确定信息量的证据合成方法。在保证各证据之间冲突整体最小的情况下, 建立证据权重分配模型, 利用遗传算法求出权值, 根据权值确定各证据的可取信息量, 对基本概率分配函数进行转化。实例分析表明, 改进的 D-S 算法具有抗干扰能力强和决策风险低等特点, 能提高证据组合结果的可信性。

**关键词:** 证据合成; P 距离; 可取信息量; 遗传算法

**中图分类号:** TP274<sup>+</sup>.1 **文献标识码:** A

## Evidence Combination Method Based on Determined Information by P Distance

Xiang Hao, Jiang Liping, Su Si

(College of Science, Naval University of Engineering, Wuhan 430033, China)

**Abstract:** On the basis of analyzing some classical improved methods, an evidence combination method based on information determined by P distance is presented for dealing with highly conflicting evidences. In the condition of ensuring the minimum of whole evidence conflict, the distribution model of evidence weights is set up, the evidence weights are acquired by means of genetic algorithm, then the obtainable information of evidence is determined, the BPA function are modifies according to the evidence weights. An example indicates that the improved D-S method can give a more reasonable combination results and have the capable of anti-jamming and the lower risk of decision-making.

**Keywords:** evidence combination; P distance; obtainable information; genetic algorithm

### 0 引言

经典 D-S 证据理论<sup>[1]</sup>源于 20 世纪 60 年代, 首先由 Dempster 提出, 而后他的学生 Shafer<sup>[2]</sup>对其进行了系统化完善。相比传统的概率论, 这种不确定性推理方法能更好地把握问题的未知性和不确定性, 具有较强的处理不确定信息能力, 被广泛应用于多传感器系统目标识别领域。但 D-S 证据理论也存在一些不足之处: 当证据之间出现比较严重冲突时, 合成结果会出现与直觉相悖的情况<sup>[3]</sup>。为解决这一问题, 国内外许多学者提出了很多改进策略<sup>[4-10]</sup>: 坚持 D-S 组合规则的乘性策略<sup>[4-6]</sup>、加性策略的组合规则<sup>[7-9]</sup>以及乘性策略和加性策略相结合的方法<sup>[10-11]</sup>, 均各有优缺点。考虑到参与合成各证据的可取信息量, 从证据的可取信息量角度来修正 BPA 函数, 提出一种基于 P 距离确定信息量的证据合成方法。

### 1 D-S 证据合成理论及各种改进方法概述

在证据理论中, 假设  $\Theta = \{A_1, A_2, \dots, A_w\}$  为一个各元素互不相容的完备集合, 称  $\Theta$  为辨识框架。设  $m_1$  和  $m_2$  是  $\Theta$  上的 2 个基本概率分配函数 (BPA

函数), 其正交和  $m = m_1 \oplus m_2$  定义为<sup>[1]</sup>:

$$m(A) = c^{-1} \sum_{A_i \cap A_j = A} m_1(A_i)m_2(A_j), \quad m(\Phi) = 0 \quad (1)$$

其中

$$c = 1 - \sum_{A_i \cap A_j = \Phi} m_1(A_i)m_2(A_j) = \sum_{A_i \cap A_j \neq \Phi} m_1(A_i)m_2(A_j) \quad (2)$$

式中,  $A_i \subset A, A_j \subset A, A \neq \Phi$ 。上式即为证据理论的核心, 它满足交换律、结合律以及单调性。证据合成规则认为: 若  $A = A_i \cap A_j$ , 且  $A \neq \Phi$ , 则  $m_1(A_i)m_2(A_j)$  是确切分配到  $A$  上的一部分信质, 式中,  $c$  是所有分配到  $A$  上的信质总和。但按这种理解, 当  $A = \Phi$  时, 将会产生有一部分信质  $\sum_{A_i \cap A_j \neq \Phi} m_1(A_i)m_2(A_j)$  被分配到空集上的不合理现象, 于是组合规则丢弃这部分信质, 但丢弃后辨识框架上所有焦元总信质  $\sum_{E \subset \Theta} m(E)$  就可能小于 1, 为此,

又在每一信质上乘以系数  $c^{-1}$  以满足总信质为 1 的要求。由以上分析可以看出: 在 D-S 证据理论中, 为保持 BPA 函数的归一性, 使 2 个证据的每个公共

收稿日期: 2010-06-03; 修回日期: 2010-07-01

基金项目: 国防科技重点实验室基金项目资助 (9140C2603060605)

作者简介: 项浩 (1984-), 男, 湖北人, 硕士, 从事信息融合、模式识别研究。

焦元的 BPA 函数值都变为原来的  $1/c$  倍, 同时遗弃独立焦元的每个冲突, 使 D-S 证据理论更适合于处理高置信度、低冲突证据的合成, 对低置信度、高冲突证据组合则会产生与常理相悖的不合理现象。尽管经典 D-S 证据理论存在上述问题, 但是也有其固有的优点: 不确定性随着证据的增加而减少并且满足交换律和结合律。

Yager<sup>[5]</sup>提出解决证据冲突的方法是把冲突的证据分配给辨识框架下的完备命题集合, 这种分配方法过于保守, 存在一票否决问题, 不满足结合律, 而且当证据多于 2 时, 无法得到合理的结果。孙全<sup>[6]</sup>在 Yager 合成公式的基础上做出了改进, 认为即使证据间存在着冲突, 也是部分可用的, 应该将总冲突按照一定的比例分配给所有证据的焦元集合, 但是改进后的组合公式同样不满足结合律, 而且当参与合成的证据数量很大时, 融合结果收敛到一致信息的速度比较慢。Murphy<sup>[7]</sup>的方法是对所有证据中所有命题的焦元函数进行直接平均, 再利用 D-S 证据理论进行组合, 该改进方法在有冲突的情况下可以得到较好的结果, 但是该方法只是将多源信息进行简单的平均, 没考虑各个证据之间的相互关联。

## 2 基于 P 距离确定信息量的证据合成方法

从各证据之间的全局冲突出发, 运用全局优化思想, 使各证据之间整体冲突程度最小的情况下, 利用加权证据间 P 距离建立权重分配模型, 得到各证据权值, 根据权值决定各证据的可取信息量, 对 BPA 函数进行修正, 最后采用 D-S 证据合成规则对修正后的 BPA 函数进行合成。

设目标识别框架为  $\{A_1, A_2, \dots, A_w\}$ ,  $m_i$  ( $i=1, 2, \dots, n$ ) 为 BPA 函数, 笔者基于 P 距离确定信息量的证据合成方法步骤如下:

1) 进行冲突检验。定义证据  $m_i$ 、 $m_j$  ( $i, j=1, 2, \dots, n$ ) 之间的相关系数为:

$$r_{ij} = \frac{\sum_{k=1}^W |m_i(A_k) - \bar{m}_i| |m_j(A_k) - \bar{m}_j|}{\sqrt{\sum_{k=1}^W (m_i(A_k) - \bar{m}_i)^2} \sqrt{\sum_{k=1}^W (m_j(A_k) - \bar{m}_j)^2}} \quad (3)$$

其中  $\bar{m}_i = \frac{1}{W} \sum_{k=1}^W m_i(A_k)$ 、 $\bar{m}_j = \frac{1}{W} \sum_{k=1}^W m_j(A_k)$ 。设  $\alpha$  为冲突检验阈值, 当  $r_{ij} \geq \alpha$  时, 认为证据间没有冲突或者冲突很小, 可以直接利用 D-S 证据理论合成; 当  $r_{ij} < \alpha$  时, 认为证据之间存在严重的冲突, 转到

步骤 (2)。

2) 定义各个加权证据  $\omega_i m_i$ 、 $\omega_j m_j$  之间的 P 距离:

$$d_{ij} = \left\{ \sum_{k=1}^W |\omega_i m_i(A_k) - \omega_j m_j(A_k)|^P \right\}^{1/P}, \quad i, j=1, 2, \dots, n \quad P \geq 1 \quad (4)$$

式中,  $\omega_i$  ( $i=1, 2, \dots, n$ ) 为证据  $m_i$  的权值, 这里,  $\omega_i$  反映了证据  $m_i$  的相对重要性程度;  $\omega_i$  越大, 证据  $m_i$  的可信度就越高, 可取的信息量就越多;  $\omega_i$  越小, 证据  $m_i$  的可信度就越低, 可取的信息量就越少。

设  $S$  为各个加权证据之间的 P 距离总和, 这里,  $S$  反映着各加权证据之间的整体冲突程度。考虑到各个证据的有效性, 从全局的观点出发, 运用全局优化思想, 为使修正后各证据之间更加接近, 从而在整体上使各修正后的证据间冲突达到最小, 使决策者能够更加合理、客观的做出决策, 将  $S$  作为目标函数建立优化模型如下:

$$\begin{aligned} \min S &= \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n d_{ij} \\ \text{s.t.} &\begin{cases} \omega_1 + \omega_2 + \dots + \omega_n = 1 \\ 0 \leq \omega_i \leq 1, \quad i=1, 2, \dots, n \end{cases} \end{aligned} \quad (5)$$

利用遗传算法求解上述目标规划问题得到证据权值  $\omega_i$ ,  $i=1, 2, \dots, n$ 。定义各证据的信息量为:

$$\gamma_i = \frac{\omega_i}{\max_{1 \leq i \leq n} \{\omega_i\}}, \quad i=1, 2, \dots, n \quad (6)$$

根据各个证据的可取信息量, 对 BPA 函数进行转化如下:

$$\tilde{m}(\Phi) = 0 \quad (7)$$

$$\tilde{m}_i(A_k) = \gamma_i m_i(A_k), \quad i=1, 2, \dots, n, \quad k=1, 2, \dots, W \quad (8)$$

$$\tilde{m}_i(\Theta) = \gamma_i m_i(\Theta) + (1 - \gamma_i), \quad i=1, 2, \dots, n \quad (9)$$

采用 D-S 证据合成规则对  $\{\tilde{m}_1, \tilde{m}_2, \dots, \tilde{m}_n\}$  进行合成。

3) 对合成结果进行判别。设  $\varepsilon$  为检验门限值,

$$\tilde{m}(A_p) = \max_{1 \leq i \leq W} \{\tilde{m}(A_i)\} \quad (10)$$

$$\tilde{m}(A_q) = \max \{\tilde{m}(A_1), \tilde{m}(A_2), \dots, \tilde{m}(A_{p-1}), \tilde{m}(A_{p+1}), \dots, \tilde{m}(A_w)\} \quad (11)$$

如果  $\tilde{m}(A_p) - \tilde{m}(A_q) > \varepsilon$ , 则  $A_p$  为判决结果; 否则认为识别结果不明确。证据合成流程如图 1。

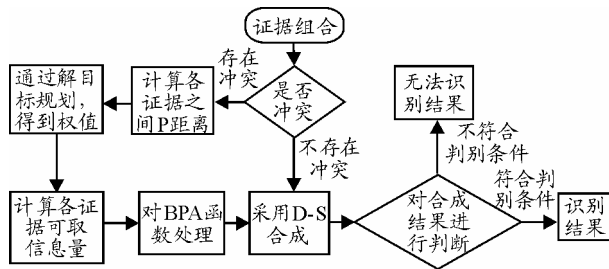


图 1 证据合成流程图

### 3 实例应用

假设识别框架为  $\{A, B, C\}$ , 识别系统中有 5 个传感器,  $m_i (i=1, 2, \dots, 5)$  为其 BPA 函数, 获得的目标基本可信度分配函数如表 1。

表 1 5 个传感器获得的关于目标的 BPA 函数

传感器	$m(A)$	$m(B)$	$m(C)$
$m_1$	0.50	0.20	0.30
$m_2$	0.00	0.90	0.10
$m_3$	0.55	0.10	0.35
$m_4$	0.55	0.10	0.35
$m_5$	0.55	0.10	0.35

将 5 个传感器获得关于目标的 BPA 函数进行冲突检验, 可知传感器 2 获得的证据与其它传感器获得的证据存在严重冲突。本算例中设检验门限值  $\varepsilon = 0.5$ , 当  $P$  分别取 1、2、3、4、5 时, 利用遗传算法得到证据权值与信息量如表 2。

表 2 证据权值与信息量

$P$	$\omega_1$	$\gamma_1$	$\omega_2$	$\gamma_2$	$\omega_3$	$\gamma_3$	$\omega_4$	$\gamma_4$	$\omega_5$	$\gamma_5$
1	0.292 1	1.000 0	0.114 6	0.392 3	0.118 5	0.405 7	0.205 5	0.703 5	0.269 5	0.922 6
2	0.238 8	1.000 0	0.083 2	0.348 4	0.225 7	0.945 4	0.225 7	0.945 4	0.225 7	0.945 4
3	0.208 2	0.972 4	0.150 8	0.704 3	0.213 0	0.994 9	0.213 8	0.998 6	0.214 1	1.000 0
4	0.221 1	0.994 2	0.110 8	0.498 2	0.222 4	1.000 0	0.222 3	0.999 6	0.222 4	1.000 0
5	0.214 8	1.000 0	0.141 5	0.658 6	0.214 4	0.998 1	0.214 7	0.999 5	0.214 4	0.998 1

利用该改进方法, 当  $P$  取 1、2、3、4、5 时,  $n(n=3, 4, 5)$  条证据合成的合成结果如图 2。

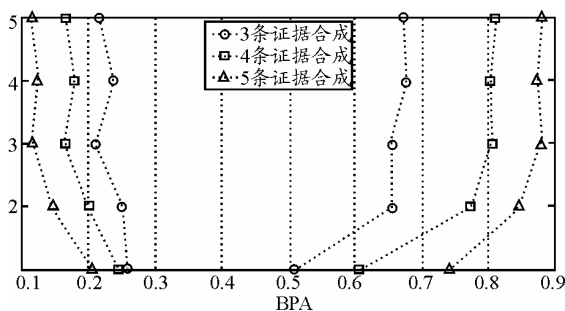


图 2 本文方法合成目标 A 与 C 的 BPA 函数

图 2 中纵坐标为  $P$  的取值, 横坐标为合成结果中 BPA 函数值最大 (A) 和次之 (C) 的识别目标

BPA 函数, 左边 3 条曲线对应于识别框架中的焦元 C, 右边 3 条曲线对应于识别框架中的焦元 A。本算例中取  $P=5$ , 由判别定义可知识别目标为 A。

在本例中取  $P=5$  时, 在分别利用 3、4、5 条证据情况下, 该改进方法与其他 4 种方法的合成结果如表 3。

表 3 各种方法的合成结果

证据组合	合成规则	$m(A)$	$m(B)$	$m(C)$	$m(\Theta)$	识别结果
$m_1 \oplus m_2 \oplus m_3$	Dempster	0	0.631 6	0.368 4	0	B
	Yager	0	0.180 0	0.010 5	0.971 5	$\Theta$
	孙全	0.160 0	0.201 0	0.125 0	0.486 0	$\Theta$
	Murphy	0.350 0	0.522 4	0.127 6	0	B
	本文	0.671 1	0.113 7	0.215 2	0	A
$m_1 \oplus m_2 \oplus m_3 \oplus m_4$	Dempster	0	0.328 8	0.671 2	0	C
	Yager	0	0.001 8	0.003 7	0.994 5	$\Theta$
	孙全	0.194 0	0.160 0	0.137 0	0.509 0	$\Theta$
	Murphy	0.602 7	0.262 7	0.134 6	0	A
	本文	0.809 7	0.025 0	0.165 3	0	A
$m_1 \oplus m_2 \oplus m_3 \oplus m_4 \oplus m_5$	Dempster	0	0.122 8	0.877 2	0	C
	Yager	0	0.000 2	0.001 3	0.998 5	$\Theta$
	孙全	0.211 0	0.138 0	0.144 0	0.507 0	$\Theta$
	Murphy	0.795 8	0.093 2	0.111 0	0	A
	本文	0.880 4	0.005 0	0.114 6	0	A

由表 3 分析可以得出: Dempster 规则不能对有冲突的证据进行有效的处理, 一旦有一证据对某命题的支持度为 0, 无论再出现对该命题多大的支持证据, 组合结果总为 0, 这显然是不合情理的; Yager 方法过于保守, 将证据冲突的那部分概率全部分配给  $\Theta$ , 不论以后收集多少支持 A 的证据, 未知项  $\Theta$  的数值始终在增加, 不利于决策者做决策; 孙全的改进方法, 可以部分克服 Yager 方法的缺点, 由表 3 可以看出随着支持命题 A 的证据越来越多, 数值有所增加, 但是分配精度不够, 未知项的数值没有明显的降低, 仍然无法做出决策。以上 3 种对证据合成公式的修改, 效果都不理想。Murphy 方法由于只是对证据进行简单平均, 没有考虑到各证据的重要程度, 使得重要程度低的证据影响了整个的合成结果。

该改进方法从降低各证据之间整体冲突的角度出发, 通过引入描述证据在合成过程中的可取信息量, 对概率分配函数重新分配, 有效地减少了不利证据 2 对组合结果的影响。从结果可见, 该改进方法最后合成结果中命题 A 的 BPA 值最大, 达到了 0.880 4, 而且在收集了 3 个证据时就能得到正确的结果。综合 5 个证据可以看出, 证据 2 可能是干扰证据, 与系统的其他证据存在较大差异。

16. If (input1 is mf4) and (input2 is mf4) then (output1 is mf4) (1)

最后, 清晰化方法采用面积重心法, 得到智能化防汽车追尾控制器的曲面图, 如图 4。从图 4 中 input1 轴可以看出, 本车与前车在 5 m 范围内就开始制动, 随着与前车距离越来越近, 其制动力度越来越大, 直至停车。进而对本车车速进行限制, 并且距离越近, 本车速度越小。当前车突然停车时, 对于单片机来讲, 其距离差(相对速度)有一个由小变大的过程, 根据图 4 中 input2 轴可以看出, 随着距离差的增大, 控制力度也在增大, 此时两车距离也越来越近, 本车速度将会在单片机实时控制下逐步减小, 直至停车, 这样可以避免突然急刹车, 避免车内人员的人身伤害。

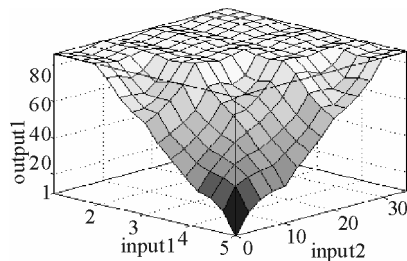


图 4 防汽车追尾控制器输出曲面图

具体在单片机中实现该程序时, 可以生成 100 行(本车在 4 cm 范围内调速) 100 列模糊控制表, 然后在单片机中直接编程进行查表控制。经过小车实地模拟测试, 该方法能取得很好的效果。

\*\*\*\*\*  
(上接第 47 页)

改进的 D-S 算法降低了干扰证据的影响, 具有抗干扰能力强和决策风险低的特点, 提高了证据组合结果的可信性。从结果中还可以看出, 随着对 A 命题支持的证据数目的增加, 目标 A 的 BPA 值稳定提高, 很好地反映了实际情况下的信息融合过程。

#### 4 结束语

该改进方法能有效处理含有严重冲突证据的组合, 不仅解决了 D-S 证据理论组合结果中可能存在与直觉相悖的情况, 同时还保留了经典 D-S 证据理论的交换律和结合律。相对于以往几种改进方法而言, 该改进方法的识别速度更快, 能更有效地降低合成结果中的不确定性, 得到更合理的结果。

#### 参考文献:

- [1] DEMPSTER A P. Upper and lower probabilities induced by a multi-valued mapping[J]. The Annals of Mathematical Statistics, 1967, 38(4): 325-339.
- [2] SHAFER G A. Mathematical theory of evidence[M].

#### 3 结束语

相比传统方法, 该方法更符合实际的刹车情况。当遇到紧急情况时, 该方法能够尽量让汽车“缓慢”停车, 能够尽量保护车内人员的安全。该方法的难点在于隶属度函数的选择。从图 4 中可以看出, 该曲面图整体上满足控制要求, 但局部曲面并不光滑。下一步, 将在生成的模糊控制表中对少数数据进行修改, 使其满足控制规律。

#### 参考文献:

- [1] 李晓霞, 李百川, 侯德藻. 汽车追尾碰撞预警系统开发研究[J]. 中国公路学报, 2001, 14(3): 93-95.
- [2] 徐杰, 杜文, 孙宏. 跟随车安全距离分析[J]. 交通运输工程学报, 2002, 2(2): 101-104.
- [3] 李晓霞, 李百川, 侯德藻. 车辆追尾碰撞避免技术[J]. 西安公路交通大学学报: 自然科学版, 2001, 21(2): 94-97.
- [4] 连晋毅, 华小洋. 汽车防追尾碰撞数学模型研究[J]. 中国公路学报, 2005, 18(3): 123-126.
- [5] 孔金生, 郭非, 王希萍. 基于安全距离模型的汽车防追尾避撞方法[J]. 微计算机信息, 2008, 24: 251-252.
- [6] 刘刚, 侯德藻, 李克强, 等. 汽车主动避撞系统安全报警算法[J]. 清华大学学报: 自然科学版, 2004, 44(5): 697-700.
- [7] 蔡自兴. 智能控制原理与应用[M]. 清华大学出版社, 2007, 11: 103-123.
- [8] Fecit. the theory of neural network and matlab7 application [M]. Beijing: Publishing House of Electronics Industry, 2005, 12.
- [9] 梅丽凤. 单片机原理及接口技术[M]. 北京: 电子工业出版社, 2003.
- [10] Princeton N J: Princeton University Press, 1976: 19-63.
- [3] Zadeh L A. Review of Shafer's a mathematical theory of evidence[J]. Artificial Intelligent Magazine, 1984, 5(3): 81-83.
- [4] Lefevre E, Colot O, V annoerenberghe P. Belief functions combination and conflict anagement[J]. Information Fusion, 2002, 3(3): 149-162.
- [5] YAGER R R. On the Dempster-Shafer framework and new combination rules[J]. Information Sciences, 1987, 41(2): 93-137.
- [6] 孙全, 叶秀清, 顾伟康. 一种新的基于证据理论的合成公式[J]. 电子学报, 2000, 28(8): 117-119.
- [7] MURPHY CK. Combining belief functions when evidence conflicts[J]. Decision Support Systems, 2000, 3(4): 1-9.
- [8] 许丽佳. D-S 理论在信息融合中的改进[J]. 系统工程与电子技术, 2004, 26(6): 717-720.
- [9] GUAN Xin, YI Xiao, HE You. An improved Dempster-Shafer algorithm for resolving the onflicting evidences[J]. Int J Info Tech, 2005, 11(12): 68-75.
- [10] 何兵, 胡红丽. 一种修正的 D-S 证据融合策略[J]. 航空学报, 2003, 24(6): 559-562.
- [11] 何兵, 毛士艺, 张有为, 等. 基于证据分类的 D-S 证据合成及判决方法[J]. 电子与信息学报, 2002, 24(7): 894-899.