

一种基于证据理论和模糊集合的信息融合方法^①

钱菲菲, 秦宁宁, 邵金涛, 徐保国

(轻工过程先进控制教育部重点实验室, 江南大学物联网工程学院, 无锡 214122)

摘要: 针对证据理论应用中基本概率分配函数难以确定和多传感器之间相互支持程度计算绝对化的问题, 提出了一种基于证据理论和模糊集合(FSB-DS)的信息融合方法。该方法首先利用相关性函数定义不确定信息的模糊支持区间和模糊支持概率, 然后由隶属函数得到各个传感器提供信息的可信度, 再将支持度和可信度转化为基本概率分配函数, 最后进行 D-S 证据合成。仿真实验表明, 该方法获得的结果具有更高的精度和可信度。

关键词: 证据理论; 模糊集合; mass 函数; 数据融合; 无线传感器网络

Information Fusion Method Based on Proof Theory and Fuzzy Set

QIAN Fei-Fei, QIN Ning-Ning, SHAO Jin-Tao, XU Bao-Guo

(Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education), School of Internet of Things Engineering, Jiangnan University, Wuxi 214122, China)

Abstract: Focused on the problem that it is difficult to set up the basic probability assignment function in the evidence theory and the calculation of degree of mutual support is absolute, an information fusion method based on evidence theory and fuzzy set(FSB-DS) is proposed. The fuzzy support range and the fuzzy support probability is obtained from the correlation function., Then by using the membership function, the reliability of information provide by each sensor is gained. and then transform support and credibility into basic probability assignment function, and finally the DS evidence combination. The simulation results shows that the fusion results have higer precision and reliability compared with other methods.

Key words: evidence theory; fuzzy set; mass function; data fusion; wireless sensor networks

1 引言

多传感器数据融合能够综合利用来自各个传感器的数据, 与单一传感器相比, 克服了被测数据的不确定性, 提高了整个传感器系统的有效性能, 能够更加准确地描述被测对象^[1]。由于不确定性的产生主要不在于信息载体的本身, 而是因为信息的利用者无法综合判断信息所反映事物的真实状态, 从而导致主观认识上的偏差, 如何利用先进的数据融合方法对不确定信息进行综合处理, 是一个值得深究的课题。Dempster-Shafer 证据理论是 Shafer 在 Dempster 理论的基础上提出的一种融合算法, 其基本方法是利用证据集合多个不相关的部分, 分别对辨识框架独立进行判断, 然后用 D-S 准则将判断结果组合起来, 提升综合判断能力, 因此能很好处理未知和不确定性信息。近年来, 证据理论已经在信息融合的目标识别方面得到

了逐步应用^[2]。

基于证据理论的信息融合方法是以 mass 函数为基础, 相互支持程度为依托, 而 mass 函数和相互支持度表示人们对目标可信程度的一种推理, 由于以往构造的 mass 函数都限定在一个具体的应用中而相互支持度计算偏于绝对化, 常常导致信息融合结果不稳定, 精度低的问题。

一种基于证据理论和模糊集合(FSB-DS)的信息融合方法以模糊置信距离测度的大小来表示传感器之间的相互支持程度, 摒弃以往支持概率选取的绝对化, 在概率融合意义下定义模糊置信距离测度和一种新的支持概率, 然后由模糊隶属函数得到各传感器提供信息的可信度, 再利用上述结果构造证据理论中的 mass 函数, 最后利用 D-S 准则对多传感器信息进行融合, 从而使融合结果更加精确、稳定。

① 收稿时间:2011-07-18;收到修改稿时间:2011-08-30

2 信息的支持程度

2.1 支持度参数

假设在一块区域内分撒了 n 个传感器节点, 第 i 个节点和第 j 个节点测到的数据分别为 X_i 和 X_j , X_i 和 X_j 都服从高斯分布, 以它们的高斯曲线作为传感器的特性函数, 记为 $p_i(x)$, $p_j(x)$ 。 x_i , x_j 分别为 X_i , X_j 的一次观测值。为了反映 x_i 和 x_j 之间的偏差大小, 采取置信距离测度, 设

$$d_{ij} = 2 \int_{x_i}^{x_j} p_i(x|x_i) dx \quad (1)$$

$$d_{ji} = 2 \int_{x_j}^{x_i} p_j(x|x_j) dx \quad (2)$$

式中,

$$p_i(x|x_i) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left\{-\frac{1}{2}\left(\frac{x-x_i}{\sigma_i}\right)^2\right\} \quad (3)$$

$$p_j(x|x_j) = \frac{1}{\sqrt{2\pi}\sigma_j} \exp\left\{-\frac{1}{2}\left(\frac{x-x_j}{\sigma_j}\right)^2\right\} \quad (4)$$

d_{ij} 的值称为第 i 个传感器与第 j 个传感器读数的置信距离测度, 借助误差函数 $erf(\theta)$ [3], 可以得出:

$$d_{ij} = erf\left(\frac{x_j - x_i}{\sqrt{2}\sigma_i}\right), \quad d_{ji} = erf\left(\frac{x_i - x_j}{\sqrt{2}\sigma_j}\right) \quad (5)$$

假设 r_{ij} 表示 x_i 被 x_j 支持的程度, 则 r_{ij} 应该为置信距离测度 d_{ij} 的函数, 并且 d_{ij} 愈大, 则 r_{ij} 愈小。根据人们认识事物的一般规律, 对于差别极大或者差别极小的一般容易区分, 而对于介于两者之间的情况一般比较难区分, 因此, 针对某一实际问题, 根据问题的具体背景, 一般可以确定置信门限值 δ_1 、 δ_2 , 其中 $0 \leq \delta_1 < \delta_2$ 。

r_{ij} 的取值可定义如下:

$$\begin{cases} r_{ij} = 1, & d_{ij} \leq \delta_1 \\ r_{ij} = 0, & d_{ij} \geq \delta_2 \\ r_{ij} = f(d_{ij}), & \delta_1 < d_{ij} < \delta_2 \end{cases} \quad (6)$$

其中, $f(x)$ 是一个严格单调下降的连续函数, 当 $x \in [\delta_1, \delta_2]$ 时, $0 \leq f(x) \leq 1$, $f(\delta_1) = 1$, $f(\delta_2) = 0$ 。论文设计如下方法确定 $f(x)$ 表达式, 确定 $\delta_3 \in (\delta_1, \delta_2)$, 确保 $f(\delta_3) = 0.5$ 成立, 则:

$$f(x) = \begin{cases} 0.5 - \frac{x - \delta_3}{2(\delta_3 - \delta_1)}, & x \in (\delta_1, \delta_3) \\ 0.5 - \frac{x - \delta_3}{2(\delta_2 - \delta_3)}, & x \in (\delta_3, \delta_2) \end{cases} \quad (7)$$

2.2 支持矩阵

如果存在 n 个传感器测量同一参数, 则存在相互

支持矩阵 $R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ r_{n1} & r_{n2} & \cdots & r_{nn} \end{bmatrix}$, 其中, r_{ij} 为 x_i 被 x_j 的

支持程度, $r_{ii} = 1, i=1,2,\dots,n$ 。

由于 x_i 的真实支持程度并不能由单一的 r_{ij} 决定, 应该是 $r_{i1}, r_{i2}, \dots, r_{in}$ 的综合支持程度, 设 x_i 的综合支持程度为 r_i , 则 r_i 越大, x_i 的真实性越高, 说明 x_i 的支持度就高, 则 r_i 应满足下面两个条件:

$$\textcircled{1} 0 \leq r_i \leq 1 \text{ 且 } \sum_{i=1}^n r_i = 1;$$

$\textcircled{2}$ 由概率源合并理论, 即由一组非负数 $\alpha_1, \alpha_2, \dots, \alpha_n$, 使得 $r_i = \alpha_1 r_{i1} + \alpha_2 r_{i2} + \dots + \alpha_n r_{in}$, 其中 $i=1,2,\dots,n$ 。

将 r_i 的关系式转化为矩阵形式 $r = R\alpha$, 其中 r 是以 r_1, r_2, \dots, r_n 为元素的列向量, α 是以 $\alpha_1, \alpha_2, \dots, \alpha_n$ 为元素的列向量。因为 R 是一个非负矩阵, 根据非负矩阵的性质可知, 存在一个最大模特征值 λ , 其对应正特征向量 α , 即 $\lambda\alpha = R\alpha$ 。这里可以将 $\lambda\alpha$ 作为传感器节点的综合支持程度的度量, 由两个条件可令 $r_i = \alpha_i / (\alpha_1 + \alpha_2 + \dots + \alpha_n)$, 则 r_i 为第 i 传感器被其他传感器支持程度。

3 证据理论和模糊隶属度

3.1 基本内涵

证据理论是建立在“识别框架 Θ ”上的理论, 它定义了一个集函数 $m: 2^\Theta \rightarrow [0,1]$, 满足: $\sum_{A \subseteq \Theta} m(A) = 1$ 且 $m(\Phi) = 0$, 称 m 为辨识框架 Θ 上的基本可信度分配, 如果 A 属于识别框架 Θ , 则 $m(A)$ 称为 A 的概率分配函数, 即 *mass* 函数, 反映了证据支持命题 A 发生的程度, 根据证据理论定义可信度函数:

$$Bel(A) = \sum_{B \subseteq A} m(B) \quad (8)$$

仅仅用可信度函数来描述对一个命题的信任程度是不够的, 须引入一个怀疑 A 的程度的量, 即:

$$Dou(A) = Bel(\bar{A}) \quad (9)$$

$$Pl(A) = 1 - Bel(\bar{A}) \quad (10)$$

Dou 称为 A 的怀疑函数, 表示 A 的怀疑度, Pl 称为似真度函数, 表示 A 的似真度。 $[Bel(A), Pl(A)]$ 表示证据的不确定性区间; $[0, Bel(A)]$ 表示命题 A 的完全可信区间; $[0, Pl(A)]$ 表示对命题 A 为真的不怀疑区间。文献 [5] 已经证明了 $m(A)$, $Bel(A)$ 和 $Pl(A)$ 之间是可以相互确定的, 因此, 本文只用 $m(A)$ 来确定命题 A 的可信度。

3.2 合成规则

D-S理论的组合规则称为正交和规则, 用 \oplus 表示, 假设 m_1 和 m_2 是 2^Θ 上的两个相互独立的基本概率分配值, Bel_1 、 Bel_2 是同一识别框架上的两个信任函数, m_1 和 m_2 为其所对应的基本概率赋值, Bel_1 的焦点是 A_1, A_2, \dots, A_j , Bel_2 的焦点是 B_1, B_2, \dots, B_K , 显然:

$$K = \sum_{A_i \cap B_j = \Phi} m_1(A_i)m_2(B_j) < 1 \quad (11)$$

扩大化重新表达:

$$m(C) = \begin{cases} 0 & C = \Phi \\ \frac{\sum_{A \cap B = C} m_1(A)m_2(B)}{1 - K} & \forall C \subset \Theta, C \neq \Phi \end{cases} \quad (12)$$

若 $K=1$, 则认为 m_1 和 m_2 矛盾, 不能对基本概率赋值进行组合; 若 $K \neq 1$, 则 $m(A)$ 确定一个基本概率分配值, 称为 m_1 和 m_2 的综合概率分配值, 对于多个证据的组合, 可采用此组合规则对证据进行两两融合。

3.3 模糊隶属度

一般, 对论域 U 的 n 个模式 A_1, A_2, \dots, A_n , 有 m 个传感器 $\mu_1, \mu_2, \dots, \mu_m$, 则对于任意识别对象 x , m 个传感器分别给出它属于各个模式的隶属度^[4]为:

$$\begin{bmatrix} \mu_{1.A_1}(x) & \mu_{1.A_2}(x) & \dots & \mu_{1.A_n}(x) \\ \mu_{2.A_1}(x) & \mu_{2.A_2}(x) & \dots & \mu_{2.A_n}(x) \\ \vdots & \vdots & & \vdots \\ \mu_{m.A_1}(x) & \mu_{m.A_2}(x) & \dots & \mu_{m.A_n}(x) \end{bmatrix}_{m \times n}$$

对象 x 属于各个模式的基本概率分配值为:

$$m_i(A_j) = \frac{r_i \mu_{i.A_j}(x)}{\sum_{j=1}^n r_i \mu_{i.A_j}(x)} \quad (13)$$

其中, $i=1, 2, \dots, m$, $j=1, 2, \dots, n$ 。 $m_i(A_j)$ 表示第 i 个传感器属于模式 A_j 的基本概率分配值, r_i 为第 i 个传感器被其他传感器支持的程度。

4 FSB-DS信息融合算法

4.1 算法步骤

通过对证据理论和模糊隶属度的分析, 论文算法

模型如图 1, 步骤可描述如下:

- 1) 将各个传感器的测量值通过隶属函数转化为模糊值, 以得到各传感器的局部决策。
- 2) 计算各传感组两两之间的距离 $\{d_{ij}\}$, 建立距离矩阵。
- 3) 根据距离矩阵计算出多传感器的相互支持程度矩阵 R 。
- 4) 利用获得的测量数据计算出每个传感器所得数据的模糊隶属度 u , 并用隶属度来表示各传感器提供数据信息的可信度。
- 5) 再将各传感器的支持度和可信度转化为基本概率赋值, 即 mass 函数。
- 6) 最后利用 D-S 合成规则对信息进行融合。计算出新组决策值, 并用新组替代被融合的两传感器, 返回到步骤 (2), 直至 $r_{ij}=0$, 或得到最大传感器连接组。

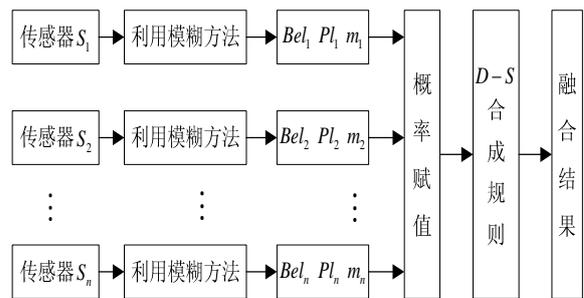


图 1 FSB-DS 融合算法模型

4.2 算法分析

通过将本文的 FSB-DS 算法与文献 [5] 提出的基于模糊理论 (FTB) 的一致性数据融合方法进行比较, 从数据测试的角度看, FSB-DS 算法的精度明显优于 FTB 算法 (将在实验部分有所体现), 将两者的方法进行对比, 会发现 FSB-DS 算法具有如下特点:

- 1) FSB-DS 算法中有关参数的确定相对于 FTB 算法中绝对阈值的设定而言, 其操作性和可接受性更高。因为, 通过将不同数据之间的相互支持程度模糊化, 避免两个数据之间相互支持的绝对化, 从而不会因为其中某个数据的微小变化而彻底改变两个数据之间的相互支持关系, 进而保证在信息融合时具有较好的稳定性和抗干扰性。

- 2) 论文采用 FSB-DS 信息融合方法, 根据测量得到的数据来计算每个传感器所得数据的模糊隶属度, 并用隶属度来表示各传感器提供信息的可信度, 然后利用隶属函数构造 mass 函数, 再根据证据理论合成公式进行融

合, 有效地实现了证据理论向模糊集的推广, 最大限度地保留了传感器给出的各个观察结果的不确定性, 最终的融合结果消除了对不明属性的判别, 可以更加准确地识别出目标的正确归属类别(具体结果实验中会提到), 使得实际问题的真实性更高, 便于在实际中具体实施。

5 实验分析

5.1 实验参数

为保证比较的公平性, 将论文算法应用于文献[5]中目标识别任务的仿真实验, 各传感器测量物体 A 和 B 的参数和测量值 (见表 1)。

表 1 测量参数及测量值

传感器	物体 A		物体 B		测量值
	μ	σ	μ	σ	
1	243	35	270	60	257
2	0.5	0.1	0.4	0.1	0.55
3	30	4	40	9	28
4	27	1.2	21	3.6	29

该任务是辨别出分别标有 A 和 B 的两个物体, 并且系统还应该能分辨出不明物体, 因此, 辨识框架为 $\{A, B, U\}$ (U 代表不明物体), 存在以下基本假设:

假设 H_0 : 物体为 A, 支持程度为 $u(A)$;

假设 H_1 : 物体为 B, 支持程度为 $u(B)$;

假设 H_2 : 物体为不明物 U, 支持程度为 $u(U)$ 。

为得到 $u(U)$, 利用模糊逻辑中的模糊关系, 可得:

$$u(U) = NOT(u(A) \text{ or } u(B))$$

$$1 - \max(u(A), u(B)) \quad (14)$$

实验采用 4 种不同类型的的传感器进行探测, 为了得到每个传感器的可信度, 采用如下所示的隶属函数:

$$u(z) = \begin{cases} 1 - \frac{|z - \mu|}{2\sigma}, & |z - \mu| < 2\sigma \\ 0 & |z - \mu| \geq 2\sigma \end{cases} \quad (15)$$

其中: z 是测量值; $\{\mu, \sigma\}$ 为均值和偏差。分别取 $\delta_1 = 0, \delta_3 = 0.150, \delta_2 = 0.300$ 。

5.2 实验过程

首先计算出置信距离测度 d_{ij} ($i, j = 1, 2, 3, 4$), 得到置信距离矩阵 D 为:

$$D = \begin{bmatrix} 0.0000 & 0.5050 & 0.4370 & 0.6990 \\ 0.3210 & 0.0000 & 0.0420 & 0.1250 \\ 0.1860 & 0.0280 & 0.0000 & 0.1110 \\ 0.8920 & 0.2500 & 0.3330 & 0.0000 \end{bmatrix}$$

得出相互支持矩阵 R :

$$R = \begin{bmatrix} 1.000 & 0.000 & 0.000 & 0.000 \\ 0.000 & 1.000 & 0.860 & 0.583 \\ 0.380 & 0.907 & 1.000 & 0.630 \\ 0.000 & 0.467 & 0.000 & 1.000 \end{bmatrix}$$

计算得最大模特征值 $\lambda = 2.1297$, 其对应的正的特征向量 $\alpha = [0.000 \ 0.668 \ 0.691 \ 0.276]^T$, 从而得到各传感器的综合支持程度 $r = [0.000 \ 0.409 \ 0.422 \ 0.169]^T$, 首先, 将测量值 z 和参数 μ, σ 代入式(15), 得到各传感器的信息可信度(见表2)。综合各传感器间的相互关系, 从而获得各传感器对目标属性的支持程度(见表3), 由式(13)求出各传感器相应的基本概率分配值(见表4)。

表 2 各传感器测量提供的信息可信度

传感器	决策 μ	
	物体 A	物体 B
1	0.800	0.892
2	0.750	0.250
3	0.750	0.333
4	0.167	0.000

表 3 各传感器对目标的支持程度

传感器	物体 A	物体 B	不明物体 U
1	0.000	0.000	0.000
2	0.307	0.102	0.693
3	0.317	0.141	0.683
4	0.028	0.000	0.972

表 4 各传感器相应的基本概率分配值

传感器	物体 A	物体 B	不明物体 U
1	0.000	0.000	0.000
2	0.279	0.092	0.629
3	0.278	0.123	0.599
4	0.028	0.000	0.972

由计算得出的综合支持程度可知, 各传感器被其他传感器支持程度依次为 3,2,4,1。因此先融合传感器 3 和 2, 融合结果为 $m_3 \oplus m_2(A) = 0.446, m_3 \oplus m_2(B) = 0.144, m_3 \oplus m_2(U) = 0.410$, 得到新的决策值 (见表 5)。

表 5 合并传感器 2 和 3 后的决策

传感器	决策 μ	
	物体 A	物体 B
1	0.800	0.892
3,2	0.446	0.144
4	0.167	0.000

与第一次融合类似,以表5为基础数据,计算出各传感器的综合支持程度 $r=[0.000\ 0.529\ 0.471]^T$,相应地,计算出合并传感器3和2后的目标支持度和基本概率赋值,再与传感器4融合,得出 $m_3 \oplus m_2 \oplus m_4(B)=0.066$,由于传感器1的支持程度为0,传感器组2,3,4的决策值0.066便为该次测量支持假设 H_1 的程度,即 $u(B)=0.066$ 。

与计算支持假设 H_1 类似,可计算出该次测量对假设 H_0 的支持程度为 $u(A)=0.913$,对假设 H_2 的支持程度为 $u(U)=1-\max(u(A),u(B))=0.087$ (见表6)表6

表6 两种方法融合结果比较

假设	支持程度(FSB-DB)	支持程度(FTB)
物体为A	0.913	0.806
物体为B	0.066	0.115
物体为U	0.087	0.194

通过融合结果可以看出:对A的概率分配函数值最大,最终判定目标是物体A。FTB融合算法对目标的支持程度从大到小分别为物体A,不明物U和物体B,结果与本文一致,不过FTB方法最终得到的融合数据为0.806,精度明显低于0.913,由此可见,FSB-DS算法的准确率比FTB融合算法要高出13.3%,由FSB-DB算法得出的目标物体的支持程度是非目标物体的十几倍,相较于FTB融合算法,FSB-DB算法具有明显的优越性(印证了前文的算法分析)。由于FTB融合算法没有考虑多传感器的相互关系,计算传感器

(上接第243页)

4 结语

扫描仪驱动程序的编译与发布非常繁琐,它首先需要编译几种类型的工程文件,然后要为各个brand将数量庞大的文件拷贝至一起,最后再通过编译InstallShield工程打包成一个可执行文件。本文将Python引入到扫描仪驱动程序的编译与发布的流程中,提出了一种可以大大简化发布人员工作量的解决方案。用户只需从服务上下载有关文件,然后运行主脚本,输入本次发布的版本号,脚本便会一直运行,直到生成所有的发布包。实践证明,该方案能够极大地提高发布效率,大大减少了发布时间,同时还避免了人工失误。此外,该方案只需稍作修改便可应用于

的支持程度时绝对化和主观化,才会导致结果误差偏大。

6 结论

多传感器对某一参数进行测量时,因受传感器自身因素和环境的影响,会有不同的测量结果,在对测量数据融合时,确定各传感器的可靠程度及数据融合方法至关重要^[6]。本文提出的FSB-DS信息融合方法简单易行,客观反映了各传感器之间的相互支持程度,能有效避免计算过程中的绝对化,使得融合结果更加精确。

参考文献

- 1 Keller J. Joining sensors through data fusion. Military and Aerospace Electronics, 2008,19(11):18-23.
- 2 Li XL, Cao JN. Coordinated workload scheduling in hierarchical sensor networks for data fusion applications. Journal of Computer Science and Technology, 2008,23(3):355-364.
- 3 Francis BA. Convergence in the boundary layer for singular perturbed equations. Automatica, 1982,18(2):57-62.
- 4 韩峰,杨万海,袁晓光.基于模糊集合的证据理论信息融合方法.控制与决策,2010,25(3):449-452.
- 5 王婷杰,施惠昌.一种基于模糊理论的一致性数据融合方法.传感器技术,1999,18(6):50-53.
- 6 刘晓光,胡学钢.D-S证据理论在决策支持系统中的应用.计算机系统应用,2010,19(10):112-116.

其他应用程序的编译与发布,特别是当发布包非常庞大时,该方案的效果更加明显。

参考文献

- 1 Chun WJ. Core Python Programming. 2nd Ed. Prentice Hall, 2006-09.
- 2 蒋崇武,刘斌,王轶辰,胡璇.基于Python的实时嵌入式软件测试脚本.计算机工程,2009,35(15):64-66.
- 3 The_winreg module. <http://effbot.org/librarybook/winreg.htm>.
- 4 Python for Windows extensions. <http://python.net/crew/mhammond/win32>
- 5 Devenv Command Line Switches.