结合SVM与DS证据理论的信息融合分类方法

雷 蕾,王晓丹
LEI Lei, WANG Xiaodan

空军工程大学 导弹学院,陕西 三原 713800
Missile Institute, Air Force Engineering University, Sanyuan, Shaanxi 713800, China

LEI Lei, WANG Xiaodan. Approach of information fusion and classification by SVM and DS evidence theory. Computer Engineering and Applications, 2013, 49(11):114-117.

Abstract: Based on the difficulty of obtaining the Basic Probability Assignment(BPA) of DS evidence theory in the practical application, an improved method of information fusion combing SVM and DS evidence theory is proposed. It uses the specific classification situation based on SVM and classifiers' reliabilities from confusion matrix to construct the basic probability assignment, which achieves the combination of SVM and the evidence theory in the information fusion. The method also presents a multi-sensor information fusion model. In the process of decision and fusion, it takes the sensors' local reliabilities into consideration and regards them as weights to integrate into BPA. The time complexity is also analyzed. The simulation results based on UCI data set and synthetic data set show that the fusion error rate can be decreased through the method proposed in this paper and the fusion reliabilities are increased.

Key words: information fusion; Support Vector Machine(SVM); evidence theory; confusion matrix

摘 要: 针对多传感器数据融合分类中,DS证据理论基本概率赋值难以解决的问题,提出了一种结合SVM与DS证据理论的信息融合改进方法。根据SVM对输入数据分类的实际情况和基于混淆矩阵得到的分类器局部识别可信度来构造基本概率赋值函数,实现了两者的有效结合,建立了SVM与DS证据相结合的多传感器信息融合模型。在决策融合过程中,重视和考虑了分类器局部识别可信度信息,并对算法进行了复杂度分析。基于UCI数据集和人工数据集的仿真结果表明该方法能够有效地降低融合识别的误差率,提高识别的可信度。

关键词: 信息融合;支持向量机;证据理论;混淆矩阵
文献标志码: A **中图分类号:** TP391 doi:10.3778/j.issn.1002-8331.1110-0377

1 引言

数据信息融合是当前信息处理领域的必然手段,怎样从不确定的信息中提取准确的信息是融合决策的关键。DS证据理论具有很强的处理不确定信息的能力。近年来成为信息融合的重要手段。然而,如何构造DS证据理论中的基本概率赋值函数(BPA),是融合中必须解决的一个重要课题,也是不易确定的问题。

许多研究者都尝试利用学习算法来获得BPA。如王毛路等利用神经网络方法通过对样本的学习,把各类条件概率作为待融合的证据[1-2],Lingmei Ai等针对医学诊断中三种不同颤动病理的分类问题,通过人工神经网络的输出结果来构造BPA[3]。考虑到神经网络在测试样本与训练样本的相差加大的情况下,可能导致结果完全错误。周皓等将支持向量机与证据理论在信息融合中结合,利用SVM的学习功能通过Platt的概率模型来确定BPA[4]。同时,结合SVM与DS证据理论的方法也在实际中得到广泛应用。张金泽等将SVM与证据理论集成方法应用于故障诊断检测[5];姜万录等利用"一对一"多类SVM分配了BPA,引入基于矩阵分析的融合算法,解决了证据理论存在的计算瓶颈问题[6]。

而在实践中,各分类器对不同类别目标的识别能力通常是不同的,因此,应估计到分类器对各个目标类别的识别可靠性。所以本文采用混淆矩阵来估计分类器局部识别可信度,提出了一种结合SVM与DS证据理论的决策融合方法:根据基分类器对输入数据分类的实际分类情况,包括分类标签、后验概率和混淆矩阵等信息来构造基本概率赋值函数,实现了SVM与DS证据理论的有效结合;同时

基金项目:国家自然科学基金(No.60975026)。
作者简介:雷蕾(1988—),女,硕士研究生,从事模式识别和智能信息处理等研究;王晓丹(1966—),女,教授,博士生导师,从事智能信息处理和机器学习等研究。
收稿日期:2011-10-19 修回日期:2011-12-07 文章编号:1002-8331(2013)11-0114-04
CNKI出版日期:2012-03-21 http://www.cnki.net/kcms/detail/11.2127.TP.20120321.1734.019.html

给出了一种结合SVM与DS证据理论的多传感器信息融合模型。

2 DS证据理论

证据理论由Dempster提出，后由Shafer进行了完善，故又称Dempster-Shafer理论，简称DS理论[7]。

在证据理论中，一个样本空间称为一个辨识框架，常用 Θ 表示，它是关于命题的彼此独立的可能答案或假设的一个有限集合，Θ 是完备的且其中的元素互不相容。Θ 的幂集记为 2^{Θ}。证据理论的基本问题就是在已知辨识框架 Θ 的条件下判明 Θ 中的一个先验的未定元素属于 Θ 中某一个子集的程度。

定义1 设 Θ 为辨识框架，Θ 的幂集构成了命题集合，如果集函数 $m:2^{\Theta}\to[0,1]$ 满足：

(1) $m(\Phi)=0$；

(2) $\sum_{A\subseteq\Theta} m(A)=1$。

则称 m 为 Θ 上的基本概率赋值(Basic Probability Assignment，BPA)函数或称mass函数。$\forall A\subseteq\Theta$，$m(A)$ 称为 A 的基本可信度[7]。

定义1包含两重含义，条件(1)表明对于空集(空命题)不产生任何信度，条件(2)反映了虽然决策者可以给一个命题赋于任意大小的信度值，但是要求决策者赋给所有命题的信度之和等于1，即总信度为1。

对于 Θ 中的任一子集 A，$m(A)$ 反映了对 A 本身的信度大小，而不去管它的任何真子集与前因后果。如果满足 $m(A)>0$，则称 A 为焦点元素，简称焦元。

定义2 设 Θ 为辨识框架，集函数 $m:2^{\Theta}\to[0,1]$ 为 Θ 上的BPA，称函数 $Bel:2^{\Theta}\to[0,1]$ 为信度函数[7]，如果其满足：

$$Bel(A)=\sum_{B\subseteq A} m(B),\ \forall A\subseteq 2^{\Theta} \tag{1}$$

$Bel(A)$ 表示对 A 的总信任度。由定义可知，$Bel(\Phi)=0$，$Bel(\Theta)=1$。

定义3 设 Θ 为辨识框架，集函数 $m:2^{\Theta}\to[0,1]$ 为 Θ 上的BPA，当其满足：

$$Pls(A)=\sum_{A\cap B\neq\Phi} m(B),\forall A\subseteq 2^{\Theta} \tag{2}$$

则称函数 $Pls:2^{\Theta}\to[0,1]$ 为似然函数(或似真函数)[3]。$Pls(A)$ 表示不否定 A 的程度，包含了所有与 A 相容的那些集合的基本可信度。

似真函数与信度函数有如下关系：

$$Pls(A)=1-Bel(\neg A) \tag{3}$$

似真函数 $Pls(A)$ 可以解释为主体在给定证据下 A 的最大可能信任程度，Pls 是一种比 Bel 更宽松的估计，对于任意的 A，显然有 $Pls(A)\geqslant Bel(A)$。区间 $[Bel(A),Pls(A)]$ 表示对命题 A 的不确定性区间，也称为 A 的信任区间。信度函数 $Bel(A)$ 和似真函数 $Pls(A)$ 分别又称为 A 的下限概率和上限概率，因此信任区间也就是 A 的概率变化范围。

需要指出的是，基本可信度 $m(A)$、信度函数 $Bel(A)$ 和似真函数 $Pls(A)$ 都是彼此唯一确定的，它们是同一证据的不同表示。

3 结合SVM与DS证据理论的信息融合方法

SVM是建立在统计学习理论的VC维理论和结构风险最小化原理基础上的学习机。标准SVM输出的是测试样本的类别标签，这就意味着在进行多个SVM基分类器信息融合时主要采用投票法。而基于数据的信息融合需要给出SVM的后验概率输出，融合前需要先把SVM输出映射为后验概率输出，即软输出。

模式识别领域中的混淆矩阵描绘了样本数据的真实类别属性与识别结果类型之间的关系，是评价分类器性能的一种常用方法。本文将混淆矩阵提供的识别率作为衡量各分类器识别能力的先验信息，对分类器的局部可信度进行描述，在构造分类器的BPA时进行加权融合。

基于以上分析，本文结合SVM与DS证据理论进行融合决策的基本思想为：首先根据SVM的硬判决输出得到其对应的软输出；其次利用混淆矩阵得到分类器针对不同目标类别的局部识别可信度估计(简称局部可信度)；最后根据SVM的软输出和分类器识别可信度估计进行基本可信度分配，而后进行DS融合，完成决策融合。

3.1 SVM的后验概率输出

对于两类SVM的后验概率输出，目前普遍接受并采用的方法是Platt提出的以Sigmod函数作为连接函数把SVM的输出 $f(x)$ 映射到 $[0,1]$ 的模型[8]：

$$P(y=1|x)=\frac{1}{1+\exp(Af(x)+B)} \tag{4}$$

其中，f 为标准的SVM输出结果，$P(y=1|f)$ 表示在输出值 f 的条件下分类正确的概率，A 和 B 是参数值，可通过求解参数集的最小负对数似然值来求得：

$$F(z)=\min_{z=(A,B)}\left(-\sum_{i=1}^{N} t_i\lg(P_i)+(1-t_i)\lg(1-P_i)\right) \tag{5}$$

其中 P_i 表示 $p(y_i=1|x_i)$。

$$t_i=\begin{cases}\dfrac{N_+ +1}{N_+ +2}, y_i=1\\ \dfrac{1}{N_- +2}, y_i=-1\end{cases}$$

N_+ 是 $y_i=1$ 的样本数量，N_- 是 $y_i=-1$ 的样本数量。

对于多类分类问题，可以结合ECOC编码等方法[9-10]获得SVM分类的后验概率输出。

3.2 基于混淆矩阵的可信度估计

假设有一个 k 类模式的分类任务，待识别数据集 X 中共有 N 个样本，每类模式中分别含有 N_i 个样本 ($i=1,2,\cdots,k$)。对数据集 X 进行分类后的混淆矩阵 $\boldsymbol{C}$ 可以表示为：

$$\boldsymbol{C}=\begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1k}\\ c_{21} & c_{22} & \cdots & c_{2k}\\ \vdots & \vdots & c_{ij} & \vdots\\ c_{k1} & c_{k2} & \cdots & c_{kk}\end{bmatrix}$$

其中 c_{ij} 表示 ω_i 类模式被分类器判断成 ω_j 类模式的数据占第 ω_i 类模式样本总数的百分比。混淆矩阵中元素的行下标对应目标的真实属性，列下标对应分类器产生的识别属性。对角线元素表示各模式能够被分类器正确识别的百分比，而非对角线元素则表示发生错误判断的百分比。

通过混淆矩阵，可以获得分类器的正确识别率和错误识别率：

各模式正确识别率：

$$R_i = c_{ii}, i = 1, 2, \cdots, k \tag{6}$$

平均正确识别率：

$$R_{\text{avg}} = \sum_{i=1}^{k}(c_{ii} \times N_i) \Big/ N \tag{7}$$

混淆矩阵行向量 $c_i(i=1,2,\cdots,k)$ 代表了模式 ω_i 的对象在进行分类时对各模式的倾向性[11]。针对当前识别问题，从输出推断输入，则由混淆矩阵可知，当分类器 L 输出类别 ω_j 时，当前样本 x 的真实类别是 ω_i 的概率[12]为：

$$PC_l(\omega_i|\omega_j) = c_{ij} \Big/ \sum_{j=1}^{k} c_{\cdot j} \tag{8}$$

将 $PC_l(\omega_i|\omega_j)$ 记作 $PC_l(\omega_i)$。则 $PC_l(\omega_i)$ 可以看作当前目标 x 属于 ω_i 类的一种支持度，即对分类器局部可信度的一种度量。

由此，将 $PC_l(\omega_i)$ 定义为分类器关于类别 i 的局部可信度，当分类器对待识别样本 x 输出一个真实类别的判决结果 ω_i 时，这个判决结果的可靠性因子就是 $PC_l(\omega_i)$。在这一思想的指导下，由分类器输出当前样本 x 的后验概率就能够根据不同类别上的可靠程度进行处理。具体来说，当分类器 L_j 对待识别目标给出 SVM 硬判决 $f(x)$ 时，将该 $f(x)$ 通过后验概率公式转化称后验概率输出；将根据混淆矩阵获得的该分类器 i 个类别的局部可信度加权融合到后验概率输出中去。这一过程可以用数学形式表示如下：

$$m_j(\omega_i) = P_i \times PC(\omega_i) \tag{9}$$

其中，$m_j(\omega_i)$ 为分类器 L_j 给出的 x 属于 ω_i 类的基本概率赋值，P_i 为SVM输出的属于 ω_i 类的后验概率，$PC(\omega_i)$ 为由混淆矩阵提供的局部可信度信息。

对每个分类器 L_j 经式(9)加权融合后的得到的BPA可由Dempster组合规则进行融合并得到最终的融合识别结果。

3.3 结合SVM与DS证据理论的多传感器信息融合模型

本节将SVM与DS证据理论用于信息融合。假设该系统中有 L_j 个传感器。首先，各局部传感器根据各自获得的信息分别进行预处理，对分类器 j 进行SVM训练确定各SVM的参数，应用时，对于传感器 L_j 的观测经SVM j 得到 P_j 和 $PC_j(\omega_i)$，再利用式(9)得到各自的BPA j，从而进行DS融合，最后给出决策融合结果。

3.4 算法复杂度分析

本节对前文提出的信息融合算法进行复杂度分析。首先，假设支持向量机的学习算法的计算复杂度为 $O(l^a)$，其中，a 对于不同的算法一般取为 $1<a<3$[13]。本文算法在规模为 l 的样本集上训练 p 个基分类器，因此，它的计算复杂度大约为 $p \cdot O(l^a)$。可见，本文提出的学习融合分类算法并未增加传统SVM的计算复杂度，着力关心解决SVM与DS的融合问题，以求获得更好的融合分类决策。

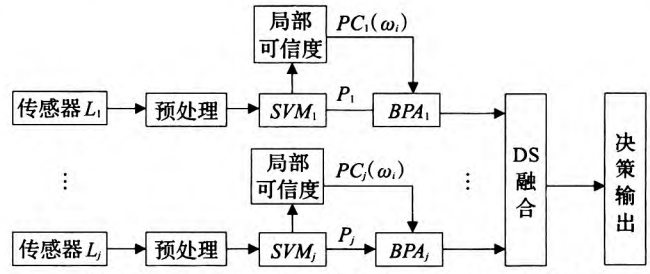

图1 结合SVM与DS证据理论的多传感器信息融合模型图

4 实验结果及分析

4.1 实验数据

实验所用的第一类数据为人工数据：产生500个以(0,0)、(2,2)为中心点，1、2为方差的两类二维正态数据，分别加以0均值高斯噪声生成正类和负类数据，如图2所示。从图中可以看出，该数据集的可分性较好。

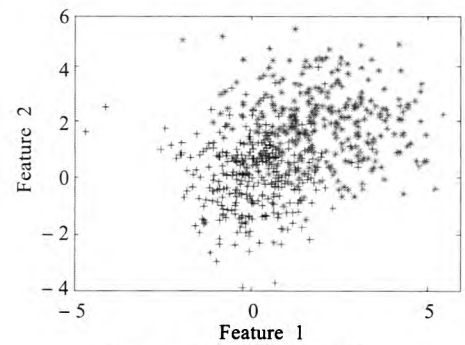

图2 正负类数据分布图

（红色为正类数据点，蓝色为负类数据点）

实验所用的第二类数据来自UCI标准数据集如表1。

表1 实验数据特征

数据集	样本数	维数
Sonar	208	60
Ionosphere	351	34
Diabetes	768	8
Breast-w	699	9
Heart-statlog	270	13

4.2 实验设计

为了验证本文方法的有效性。实验将模拟对来自5个传感器的目标数据分类。在本文信息融合过程中，采用5个基SVM分类器 $L_i(i=1,2,\cdots,5)$，均采用高斯核函数：$\sigma^2=1, C=10$。考虑到实际中各传感器性能的不同，对测试数据分别加以不同的噪声，均值均为0，方差分别为1，1.2，1.5，1.8，2。利用本文方法对基SVM分类器的输出进行DS融合，将其结果与独立的SVM分类结果进行比较，两个独立的SVM的高斯核参数分别为 $\sigma^2=10, C=50$ 和 $\sigma^2=5, C=10$。

在估计分类错误率时采用十重交叉验证来进行，并利

用双边估计 t 检验法来计算置信水平为0.95的分类错误率置信区间作为最终结果，计算公式如下：

$$\frac{|\bar{x}-\mu|}{\sigma/\sqrt{n}} \geqslant t_{0.025}(n-1) \tag{10}$$

μ,σ 分别表示十重交叉验证的均值和标准差，$t_{0.025}(9)=2.262\,2$。实验中所用基分类器均来自PRTool(http://www.prtools.org)工具箱，实验机器配置为1 GB内存，2.30 GHz CPU，算法基于Matlab7.0(R2010a)实现。

4.3 实验结果和分析

4.3.1 人工数据集

(1)实验得到5个基分类器的后验概率参数 A、B 如表2。

表2 5个基分类器后验概率参数

参数	L_1	L_2	L_3	L_4	L_5
A	-4.409 9	-3.871 7	-3.658 1	-2.972 0	-3.987 7
B	2.742 8	2.234 2	2.174 7	1.569 3	1.727 1

(2)5个基分类器得到的正负类模式的混淆矩阵：

$$L_1:\begin{pmatrix}0.969\,2 & 0.030\,8\\ 0.196\,4 & 0.803\,6\end{pmatrix}\quad L_2:\begin{pmatrix}0.938\,5 & 0.061\,5\\ 0.147\,4 & 0.877\,7\end{pmatrix}$$

$$L_3:\begin{pmatrix}0.909\,9 & 0.161\,5\\ 0.236\,4 & 0.763\,6\end{pmatrix}\quad L_4:\begin{pmatrix}0.875\,0 & 0.125\,0\\ 0.236\,4 & 0.763\,6\end{pmatrix}$$

$$L_5:\begin{pmatrix}0.949\,2 & 0.050\,8\\ 0.090\,9 & 0.909\,1\end{pmatrix}$$

(3)本文方法与独立分类器分类误差(均值±方差)(%)比较如表3。

表3 分类误差比较 (%)

分类器	SVM1	SVM2	本文方法
10重交叉	10.00±2.49	10.33±2.88	9.50±2.14

4.3.2 UCI数据集

表4为基于UCI数据集，本文方法与不同独立分类器分类结果比较。

表4 分类误差比较(均值±方差) (%)

数据集	均值±方差		
	SVM1	SVM2	本文方法
Sonar	13.03±5.17	12.48±4.43	10.92±3.00
Ionosphere	9.12±4.36	7.92±3.58	6.59±2.78
Diabetes	24.23±3.50	24.14±2.71	22.74±1.69
Breast-w	7.88±1.33	6.44±0.73	4.88±2.39
Heart-statlog	28.52±6.73	25.41±3.54	23.78±7.31

表5为不同数据集在十折交叉训练之后的时间复杂度。

表5 时间复杂度 ms

数据集	Synthetic Data Set	Sonar	Ionosphere	Diabetes	Breast-w	Heart-statlog
时间	28.593 8	97.781 3	62.868 8	45.643 1	40.984 4	35.534 3

通过实验可以得出以下结论：

(1)使用本文方法进行融合分类的分类性能优于使用单个分类器的分类器性能，证明了信息融合的优势。本文提出的信息融合方法综合考虑和利用了SVM的所有输出信息，将跟识别问题有关的信息都进行了融合，所以其分类精度大于独立分类器。

(2)本文提出的方法简单、实用、有效。提供传感器局部信息的混淆矩阵和后验概率很容易从分类结果中得到，对实际数据的适用性很强，其信息融合达到了预期的结果。且在小样本情况下，时间复杂度不是很高。

(3)因为测试样本的确定性，精度提高不是很明显，混淆矩阵提供的分类器局部可信度信息并未发挥很大的作用。考虑到实际情况的复杂性和信息的不确定性，这种结合后验概率和混淆矩阵的DS信息融合将更加真实和准确。

5 结论

本文提出了一种结合SVM与DS证据理论的信息融合改进方法。该方法根据SVM分类的实际结果，从中获取分类标签、后验概率和混淆矩阵等信息来构造待融合的证据。根据数据集本身特点及分类器性能构造BPA使获得的基本概率赋值函数更加可靠和符合实际，从而很好地解决了证据理论应用中的主要问题。从实验结果可以看出结合两种方法的信息融合的分类器的识别误差降低，达到了信息融合的目的。如何在提高融合分类正确率的前提下优化SVM与DS证据理论结合的算法复杂性是下一步的研究方向。

参考文献：

[1] 王毛路，李少洪，毛士艺.证据理论和神经网络结合的目标识别方法[J].北京航空航天大学学报，2002，28(5)：536-539.

[2] 杨露菁，郝威.多传感器目标识别的神经网络与证据理论结合方法[J].探测与控制学报，2006，28(1)：40-43.

[3] Ai Lingmei，Wang Jue，Wang Xuelian.Multi-features fusion diagnosis of tremor based on artificial neural network and D-S evidence theory[J].Signal Processing，2008，88：2927-2935.

[4] 周皓，李少洪.支持向量机与证据理论在信息融合中的结合[J].传感技术学报，2008，21(9)：1566-1570.

[5] 张金泽，单甘霖.SVM与证据理论集成的信息融合故障诊断技术研究[J].电光与控制，2007，14(4)：187-190.

[6] 姜万录，吴胜强.基于SVM和证据理论的多数据融合故障诊断方法[J].仪器仪表学报，2010，31(8)：1738-1743.

[7] Shafer G A.Mathematical theory of evidence[M].Princeton：Princeton University Press，1976.

[8] Platt J.Probabilistic outputs for support vector machines and comparison to regularized likelihood method[M]//Advance in large margin classifier.[S.l.]：MIT Press，2000：61-74.

[9] Zhou Jindeng，Wang Xiaodan，Song Heng.Research on the unbiased probability estimation of error-correcting output coding[J].Pattern Recognition，2011，44：1552-1565.

[10] Wu T F，Lin C J，Weng R C.Probability estimates for multi-class classification by pair wise coupling[J].Journal of Machine Learning Research，2004，5：975-1005.

[11] 张静.基于混淆矩阵和Fisher准则构造层次化分类器[J].软件学报，2005，16(9)：1560-1567.

[12] 贾宇平.基于信任函数理论的融合目标识别研究[D].长沙：国防科学技术大学研究生院，2009.

[13] 王磊.支持向量机学习算法的若干问题研究[D].成都：电子科技大学，2007.