

DOI: 10.7652/xjtuxb201804019

# 引入拒识的最小风险弹道目标识别

赵振冲, 王晓丹

(空军工程大学防空反导学院, 710051, 西安)

**摘要:** 针对弹道目标识别过程中误分类代价的不对称性,提出了一种引入拒识的最小风险弹道目标识别方法。该方法借助受试者工作特征(ROC)曲线,从概率分布的角度分析了真阳性率、虚警率与决策阈值之间的关系;定义了代价敏感模型及引入拒识代价的模型的代价损失函数,推导了最小化分类代价所需满足的条件,结合 ROC 曲线的切线斜率,确定了最优拒识阈值,并利用两类目标的高分辨距离像(HRRP)数据验证了方法的有效性。实验结果表明:该方法能够唯一确定使代价最小的拒识阈值,克服了传统方法由于依赖于类先验概率及训练结果而造成的随机性和不确定性的缺点,提高了可靠性和稳定性,具有理论和实际应用价值。

**关键词:** 代价敏感;拒识;弹道目标识别;受试者工作特征曲线

**中图分类号:** TP391 **文献标志码:** A **文章编号:** 0253-987X(2018)04-0132-07

## A Minimum Risk Recognition Method of Ballistic Targets with Rejection Options

ZHAO Zhenchong, WANG Xiaodan

(School of Air and Missile Defense, Air Force Engineering University, Xi'an 710051, China)

**Abstract:** A new method to minimize the recognition risk of ballistic targets with rejection options is proposed to focus on the unsymmetry problem of different misclassification costs in ballistic targets recognition. The receiver operating characteristic (ROC) curve is applied to analyze the relationship between the true positive rate and false alarm rate and the decision threshold from the perspective of probability distribution. Loss functions of a cost-sensitive model and a cost-sensitive model with rejection options are defined, and the condition to be met for minimizing the loss function is derived. The optimal rejection thresholds are obtained via the tangent of ROC curve. The proposed method is validated on high resolution range profile (HRRP) datasets of two kinds of targets. The results show that the proposed method obtains the unique optimal rejection threshold, and minimizes the classification risk as well. The weaknesses of traditional methods are overcome and the reliability and stability are improved. It shows that the method has good values in theoretical analysis and realistic applications.

**Keywords:** cost-sensitive; reject recognition; ballistic target recognition; receiver operating characteristic curve

弹道目标识别问题中,把弹头误判为诱饵比把 诱饵误判为弹头所付出的代价要大得多。代价敏感

收稿日期: 2017-09-04。 作者简介: 赵振冲(1990—),男,博士生;王晓丹(通信作者),女,教授,博士生导师。 基金项目: 国家自然科学基金资助项目(60975026,61273275)。

网络出版时间: 2018-01-14

网络出版地址: <http://kns.cnki.net/kcms/detail/61.1069.T.20180114.1111.010.html>

方法就是为解决此类误分类代价不平衡问题而提出的。传统的分类方法进行了代价敏感拓展,例如代价敏感支持向量机<sup>[1-2]</sup>,代价敏感贝叶斯<sup>[3-4]</sup>,代价敏感神经网络<sup>[5]</sup>,代价敏感决策树<sup>[6-7]</sup>,代价敏感集成<sup>[8-9]</sup>等。Chatelain 利用多目标优化算法构建受试者工作特征(ROC)前沿,然后依据最小贝叶斯风险确定给定代价矩阵下最优分类器<sup>[10]</sup>;Bernard 利用两两分解算法将 Chatelain 的方法推广到多类问题<sup>[11]</sup>;Tapkana 用代价矩阵修改近邻搜索方法和适应度函数,提出了代价敏感蜂群算法<sup>[12]</sup>;Ju 在粗糙集中引入测试代价和误分类代价,提出了一个通用的代价敏感粗糙集模型<sup>[13]</sup>。

弹道目标识别是一个时间上连续的过程,因此可以对当前类别属性不明确的目标进行拒识,在下一阶段再对其进行连续跟踪识别,以提高识别正确率并降低误判风险。构造可拒识分类器的关键在于拒识阈值的确定。刘传武等采用遍历的方法来确定使总体分类代价最小的拒识阈值<sup>[14-15]</sup>;Giorgi 通过对不同样本设置不同拒识阈值的方法来提高各类的识别正确率<sup>[16-17]</sup>;郑恩辉等在风险最小的前提下采用最优化的方法来确定各类别的拒识阈值<sup>[18]</sup>;Francesco 在构建了最小化风险函数后借助于 ROC 曲线来确定两类目标的拒识阈值<sup>[19]</sup>。尽管以上方法均取得了不错的效果,但同样存在一些不足。前 2 种方法未考虑不同误分类代价的差异性,后 2 种方法的结果由于受训练样本的影响而出现较大随机性,即使在同一训练数据上,计算的结果也会出现较大差异。

基于以上分析,本文提出了一种新的引入拒识的最小风险弹道目标识别(ORC)方法,分析了 ROC 曲线及其特性,从类先验概率的角度推导其切线的斜率方程,构造了包含拒识代价的损失代价函数,推导了最小化代价风险时 2 类阈值所需满足的条件,并使其与 ROC 曲线切线的斜率相结合,确定了最小化损失代价的两类拒识阈值。

## 1 ROC 曲线及分析

弹道目标识别可简化成只包含弹头和诱饵的两类分类问题,因此本文主要研究的是两类分类模型,而 ROC 曲线是一种用来评定两类分类器性能好坏的有效方法。

设  $I = \{(x_i, y_i) | x_i \in \mathbf{R}^d, i = 1, 2, \dots, N\}$  为样本集合,  $y_i$  为样本类别,  $y_i = \{\omega_1, \omega_2\}$  为样本类别。称第 1 类  $\omega_1$  为弹头(正类,阳性);第 2 类  $\omega_2$  为诱饵

(负类,阴性)。令  $r_{tp}$  为正类被正确分类的概率,即真阳性率(检测率); $r_{tn}$  为负类被正确分类的概率,即真阴性率; $r_{fp}$  为负类被错误分类的概率,即假阳性率(虚警率), $r_{fn}$  为正类被错分的概率,即假阴性率(漏警率)。ROC 曲线是一个以  $r_{fp}$  为横坐标,以  $r_{tp}$  为纵坐标的二维图形,如图 1 所示,描述了两类分类器的检测率  $r_{tp}$  和虚警率  $r_{fp}$  之间的相互关系。曲线上每一个点表示在特定识别阈值下分类器的性能。越靠近左上角的点准确率越高,分类器性能越好。图 1 中虚线部分描绘了基于 ROC 曲线拟合得到的 ROC 凸包,它是对分类器在特定识别阈值时可能达到的最好性能的一种估计,可以用来对分类器进行局部寻优。

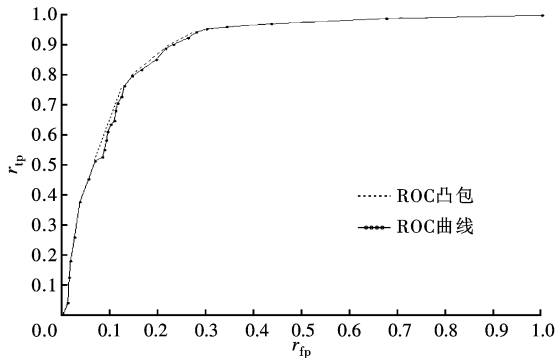


图 1 检测率随虚警率变化的 ROC 曲线

设任意目标  $\forall x_i \in I$ , 属于弹头和诱饵的后验概率分别为  $p(\omega_1 | x_i)$ ,  $p(\omega_2 | x_i)$ , 且  $p(\omega_1 | x_i) + p(\omega_2 | x_i) = 1$ 。设  $t \in [0, 1]$  为对应于  $\omega_1$  的决策阈值, 则最小错误率决策规则可表示为

$$\left. \begin{aligned} y_i &= \omega_1, & p(\omega_1 | x_i) &> t \\ y_i &= \omega_2, & p(\omega_1 | x_i) &\leq t \end{aligned} \right\} \quad (1)$$

假设  $p(\omega_1 | x_i)$  是  $x_i$  的单调增函数,  $x_i$  的维度为 1, 两类问题的后验概率如图 2 所示。对  $\forall t \in [0, 1]$ ,  $\exists x_t$ , 使得对  $\forall x_i \geq x_t$ ,  $p(\omega_1 | x_i) \geq p(\omega_1 | x_t) = t$ 。由此可知  $x_t$  是决策阈值  $t$  的函数, 设为

$$x_t = g(t) \quad (2)$$

图 3 为一维情况下两类目标的条件概率密度函数  $p(x_i | \omega_1)$  和  $p(x_i | \omega_2)$ , 其正类的检测率和虚警率

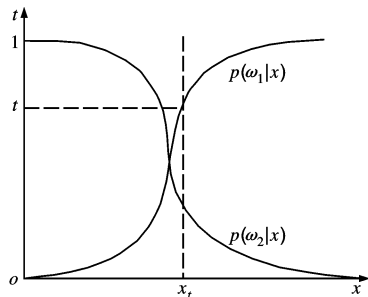


图 2 两类问题的后验概率示意图

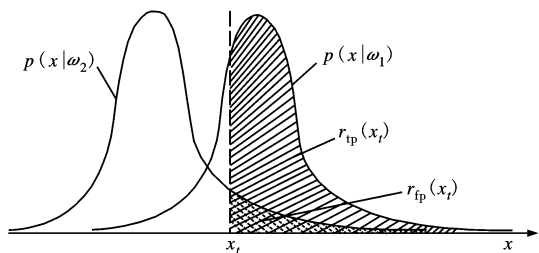


图3 一维情况下类条件概率密度示意图

可表示为

$$\left. \begin{aligned} r_{tp}(x_t) &= \int_{x_t}^{+\infty} p(x | \omega_1) dx \\ r_{fp}(x_t) &= \int_{x_t}^{+\infty} p(x | \omega_2) dx \end{aligned} \right\} \quad (3)$$

由式(3)可知,ROC曲线由 $r_{tp}(x_t)$ 、 $r_{fp}(x_t)$ 确定,则 $r_{tp}$ 对 $r_{fp}$ 求导可得

$$\frac{dr_{tp}}{dr_{fp}} = \frac{dr_{tp}(x_t)/dx_t}{dr_{fp}(x_t)/dx_t} = \frac{(\int_{x_t}^{+\infty} p(x | \omega_1) dx)'}{(\int_{x_t}^{+\infty} p(x | \omega_2) dx)'} = \frac{p(x_t | \omega_1)}{p(x_t | \omega_2)} = K(x_t) \quad (4)$$

式中: $K(x_t)$ 为ROC曲线上任一点的切线的斜率方程。综合式(1)、(4)可知, $K$ 是 $t$ 的函数,设 $K=K(g(t))$ 。

## 2 最优化拒识分类器

弹道目标识别是典型的非对称识别问题,利用代价敏感分类方法能够降低识别的整体风险。对应于式(1)的决策规则的代价矩阵可定义为

$$\mathbf{C} = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} \quad (5)$$

其中: $C_{ij}$ 为将第 $i$ 类识别为第 $j$ 类的代价,由于把弹头误识为诱饵要比把诱饵误识为弹头的代价大得多,因此令 $C_{12} > C_{21} > C_{11} = C_{22} \geq 0$ 。由式(3)可得

$$\left. \begin{aligned} r_{tn}(x_t) &= \int_{-\infty}^{x_t} p(x | \omega_2) dx = 1 - r_{fp}(x_t) \\ r_{fn}(x_t) &= \int_{-\infty}^{x_t} p(x | \omega_1) dx = 1 - r_{tp}(x_t) \end{aligned} \right\} \quad (6)$$

由式(3)、(6)可得,正、负类正确分类的概率分别为 $r_{tp}(x_t)p(\omega_1)$ 和 $r_{tn}(x_t)p(\omega_2)$ ,正、负类错误分类的概率分别为 $r_{fn}(x_t)p(\omega_1)$ 和 $r_{fp}(x_t)p(\omega_2)$ 。根据式(5)任一决策阈值 $t$ 下基于式(1)的决策规则的平均损失代价函数可表示为

$$F_{\text{cost}}(x_t) = p(\omega_1)(C_{11}r_{tp}(x_t) + C_{12}r_{fn}(x_t)) + p(\omega_2)(C_{21}r_{fp}(x_t) + C_{22}r_{tn}(x_t)) \quad (7)$$

式中: $p(\omega_1)$ 、 $p(\omega_2)$ 分别是弹头和诱饵的类先验概率,且 $p(\omega_1) + p(\omega_2) = 1$ 。将式(2)带入式(7)得

$$F_{\text{cost}} = F_{\text{cost}}(g(t)) \quad (8)$$

最小风险分类器就是要找到最优化决策阈值 $t_0$ ,使得 $F_{\text{cost}}(t_0) = \arg \min_t F_{\text{cost}}(g(t))$ 。

代价敏感分类可以形象地表示为把分类边界向误分类代价小的诱饵一类移动。由此带来的问题是对于诱饵会出现较高的误识率,考虑到弹道防御资源的有限性,这同样是不合适的。由于弹道目标识别是一个时间上连续的过程,需要在助推段、中段和再入段对其进行连续跟踪识别,因此在现阶段对风险大容易产生误分类的目标拒识,留给下一个时间点再对其进行跟踪识别更具实际意义。

设 $(t_1, t_2)$ 为决策阈值对,且 $t_1 < t_2$ 时,引入拒识的决策规划如图4所示。引入拒识的分类决策规则可表示为

$$\left. \begin{aligned} y_i &\in \omega_2, & p(\omega_1 | x_i) &< t_1 \\ y_i &\in \omega_1, & p(\omega_1 | x_i) &> t_2 \\ \text{拒识,} & & t_1 &\geq p(\omega_1 | x_i) \geq t_2 \end{aligned} \right\} \quad (9)$$

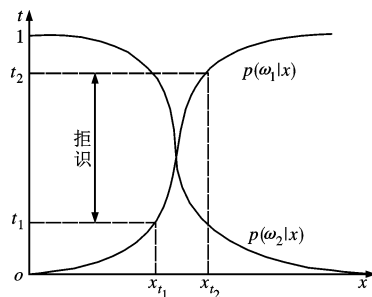


图4 引入拒识的决策规则

此时式(3)和式(6)可变换为如下形式

$$\left. \begin{aligned} r_{tp}(x_{t_2}) &= \int_{x_{t_2}}^{+\infty} p(x | \omega_1) dx \\ r_{fp}(x_{t_2}) &= \int_{x_{t_2}}^{+\infty} p(x | \omega_2) dx \end{aligned} \right\} \quad (10)$$

$$\left. \begin{aligned} r_{tn}(x_{t_1}) &= \int_{-\infty}^{x_{t_1}} p(x | \omega_2) dx \\ r_{fn}(x_{t_1}) &= \int_{-\infty}^{x_{t_1}} p(x | \omega_1) dx \end{aligned} \right\} \quad (11)$$

式中: $x_{t_1} = g(t_1)$ ,  $x_{t_2} = g(t_2)$ 。令 $r_{rn}(x_{t_1}, x_{t_2})$ 为诱饵被拒识的概率, $r_{rp}(x_{t_1}, x_{t_2})$ 为弹头被拒识的概率,其表达式如下

$$\left. \begin{aligned} r_{rp}(x_{t_1}, x_{t_2}) &= \int_{x_{t_1}}^{x_{t_2}} p(x | \omega_1) dx = 1 - r_{tp}(x_{t_2}) - r_{fn}(x_{t_1}) \\ r_{rn}(x_{t_1}, x_{t_2}) &= \int_{x_{t_1}}^{x_{t_2}} p(x | \omega_2) dx = 1 - r_{tn}(x_{t_1}) - r_{fp}(x_{t_2}) \end{aligned} \right\} \quad (12)$$

对于拒识的样本,引入拒识代价  $C_r$ ,此时式(5)中代价矩阵变换为如下形式

$$C_r = \begin{bmatrix} C_{11} & C_{12} & C_r \\ C_{21} & C_{22} & C_r \end{bmatrix} \quad (13)$$

在弹道目标识别中,当前时刻拒识类别不确定的目标是为了利用下个或多个识别间隔的综合信息进行更为准确的识别,因此拒识的风险比每一种误分类代价都小才有意义,即  $C_{12} > C_{21} > C_r > C_{11} = C_{22} \geq 0$ 。依据式(10)~(12),在阈值( $t_1, t_2$ )下,正、负类正确分类的概率分别为  $r_{ip}(x_{t_2})P(\omega_1)$  和  $r_{in}(x_{t_1})P(\omega_2)$ ,错误分类的概率分别为  $r_{fn}(x_{t_1})P(\omega_1)$  和  $r_{fp}(x_{t_1})P(\omega_2)$ ,正、负类拒分类的概率分别为  $r_{rp}(x_{t_1}, x_{t_2})P(\omega_1)$  和  $r_{rn}(x_{t_1}, x_{t_2})P(\omega_2)$ ,则根据式(13),此时的平均损失代价函数可表示为

$$F_{\text{cost}}(x_{t_1}, x_{t_2}) = p(\omega_1)(C_{11}r_{ip}(x_{t_2}) + C_{12}r_{fn}(x_{t_1}) + C_r r_{rp}(x_{t_1}, x_{t_2})) + p(\omega_2)(C_{21}r_{fp}(x_{t_2}) + C_{22}r_{rn}(x_{t_1}) + C_r r_{rn}(x_{t_1}, x_{t_2})) \quad (14)$$

将式(6)、(12)代入式(14)可得

$$F_{\text{cost}}(x_{t_1}, x_{t_2}) = p(\omega_1)(C_{11} - C_r)r_{ip}(x_{t_2}) + p(\omega_2)(C_{21} - C_r)r_{fp}(x_{t_2}) + p(\omega_1)(C_{12} - C_r)r_{fn}(x_{t_1}) + p(\omega_2)(C_{22} - C_r)r_{rn}(x_{t_1}) + C_r = p(\omega_1)(C_{11} - C_r)r_{ip}(x_{t_2}) + p(\omega_2)(C_{21} - C_r)r_{fp}(x_{t_2}) + \Delta - [p(\omega_1)(C_{12} - C_r)r_{ip}(x_{t_1}) + p(\omega_2)(C_{22} - C_r)r_{fp}(x_{t_1})] \quad (15)$$

式中:  $\Delta = p(\omega_1)C_{12} + p(\omega_2)C_{22}$ 。考虑式(4),将式(15)分别对  $r_{ip}(x_{t_1})$  和  $r_{fp}(x_{t_2})$  求偏导可得

$$\left. \begin{aligned} \frac{\partial F_{\text{cost}}(x_{t_1}, x_{t_2})}{\partial r_{ip}(x_{t_1})} &= -p(\omega_1)(C_{12} - C_r) \frac{\partial r_{ip}(x_{t_1})}{\partial r_{ip}(x_{t_1})} - p(\omega_2)(C_{22} - C_r) \\ \frac{\partial F_{\text{cost}}(x_{t_1}, x_{t_2})}{\partial r_{fp}(x_{t_2})} &= p(\omega_1)(C_{11} - C_r) \frac{\partial r_{ip}(x_{t_2})}{\partial r_{fp}(x_{t_2})} + p(\omega_2)(C_{21} - C_r) \end{aligned} \right\} \quad (16)$$

令式(16)等于0,可得

$$\left. \begin{aligned} \frac{\partial r_{ip}(x_{t_1})}{\partial r_{ip}(x_{t_1})} &= \frac{-p(\omega_2)(C_{22} - C_r)}{p(\omega_1)(C_{12} - C_r)} \\ \frac{\partial r_{ip}(x_{t_2})}{\partial r_{fp}(x_{t_2})} &= \frac{-p(\omega_2)(C_{21} - C_r)}{p(\omega_1)(C_{11} - C_r)} \end{aligned} \right\} \quad (17)$$

要取得损失代价函数  $F_{\text{cost}}(x_{t_1}, x_{t_2})$  的最小值,应使决策阈值对( $t_1, t_2$ )对应 ROC 曲线上2点的切线的斜率满足式(17)。

设任一决策阈值  $p(\omega_1 | x_t) = t$ ,则有

$$\frac{p(\omega_1 | x_t)}{p(\omega_2 | x_t)} = \frac{p(x_t | \omega_1)p(\omega_1)/p(x_t)}{p(x_t | \omega_2)p(\omega_2)/p(x_t)} = \frac{t}{1-t} \quad (18)$$

式中:  $p(x_t)$  为目标概率密度函数。式(18)变换后得

$$\frac{p(x_t | \omega_1)}{p(x_t | \omega_2)} = \frac{p(\omega_2)t}{p(\omega_1)(1-t)} \quad (19)$$

考虑式(4)可知,ROC 曲线上任一点的切线的斜率是决策阈值  $t$  的函数,即

$$K(t) = \frac{p(\omega_2)t}{p(\omega_1)(1-t)} \quad (20)$$

将式(17)与式(19)联立,可得到2个最优拒识阈值

$$\left. \begin{aligned} t_1 &= \frac{C_r - C_{22}}{C_{12} - C_{22}} \\ t_2 &= \frac{C_{21} - C_r}{C_{21} - C_{11}} \end{aligned} \right\} \quad (21)$$

由式(21)可以得出,最小风险决策阈值对能够由代价矩阵(13)唯一确定。

根据代价矩阵的含义,设正确分类的代价为0,即  $C_{11} = C_{22} = 0$ ,则式(21)可以化简为

$$\left. \begin{aligned} t_1 &= \frac{C_r}{C_{12}} \\ t_2 &= 1 - \frac{C_r}{C_{21}} \end{aligned} \right\} \quad (22)$$

根据最小风险决策阈值( $t_1, t_2$ )的含义有  $t_1 < 0.5, t_2 > 0.5$ ,则  $C_r < 0.5C_{12}, C_r < 0.5C_{21}$ ,即

$$C_r < 0.5\min(C_{12}, C_{21}) \quad (23)$$

拒识代价应小于任一种误分类代价的一半,拒识才有意义。

### 3 算法的实验验证与性能分析

本节利用两类目标的 HRRP 仿真数据验证所提方法的有效性。

#### 3.1 实验数据

实验数据来源于利用仿真软件计算得到的2类目标的动态一维距离像(HRRP)数据。2类目标分别为平底锥柱(弹头)和平底锥(诱饵),进动周期均为2s,进动角为  $10^\circ$ ,雷达发射频率为  $8.75 \sim 10.75$  GHz,采用步进雷达体制,频率步长为  $15.748$  MHz,极化方式为水平极化,雷达采样频率为  $10$  Hz。提取 HRRP 的能量聚集区长度特征、功率谱特征、幅度谱差分特征和频谱特征,4类特征组合构成维数为384的特征向量。

#### 3.2 实验设计

分类器采用支持向量机(SVM),核函数为径向

基核函数。对于 SVM 输出利用 Sigmoid 函数转换成后验概率。用主分量分析法(PCA)对特征维数进行约减,并采用累积贡献率( $A_{cr}$ )的方法来确定约减后保留的主分量的维数, $A_{cr}$ 变化区间为 70%~94%,步长为 0.02。采用 10 轮 5 倍交叉验证来计算各类的分类错误率和决策损失代价函数的大小。除本文方法外,利用循环搜索的方法来提取每一次实验中实际达到的最小分类代价(RLC)及其对应的决策阈值对,并以 Francesco 的方法<sup>[19]</sup>作为对比实验。RLC 在搜索最优化拒识阈值时搜索步长设为 0.1。Francesco 用到的类先验概率采用样本比率进行近似估计。代价矩阵各参数设为

$$C_r = \begin{bmatrix} 0 & 6 & 1 \\ 3 & 0 & 1 \end{bmatrix} \tag{24}$$

3.3 实验结果及分析

3.3.1 不同分类方法代价  $F_{cost}$  的对比 不同累积贡献率下 3 种方法的识别代价  $F_{cost}$  如图 5 所示。为了进一步对比 3 种方法的稳定性,图 6 给出了 3 种方法在未经约减的数据上采用 10 轮 5 倍交叉验证重复实验 20 次得到的分类代价。

由图 5 可以看出,不同累积贡献率下,ORC 方法的识别代价与 RLC 方法保持了高度的一致性,而 RLC 方法是可能达到的最好结果,说明 ORC 方法能够获得一个近似最优的结果,而 Francesco 方法

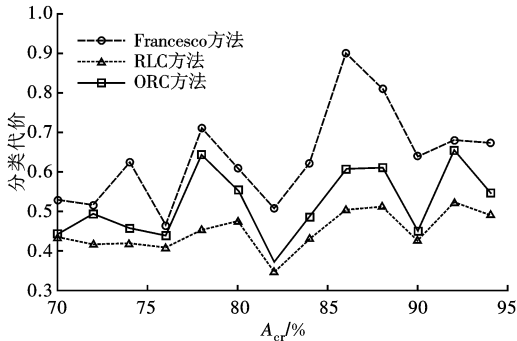


图 5 不同累积贡献率下 3 种方法的识别代价对比

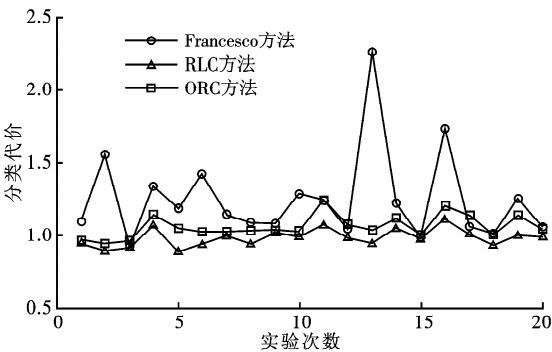


图 6 3 种方法在 20 次独立重复实验中的分类代价

产生的分类代价比 ORC 方法和 RLC 方法要大得多。由图 6 可以看出,不同的独立实验中 ORC 方法与 RLC 方法的识别代价相对稳定,Francesco 方法的识别代价则波动较大,出现了较大的不稳定性。

3.3.2 不同方法拒识阈值对的比较 为了分析 3 种方法计算得出的阈值对之间的变化关系,图 7 给出了图 6 中每个点对应的决策阈值对。

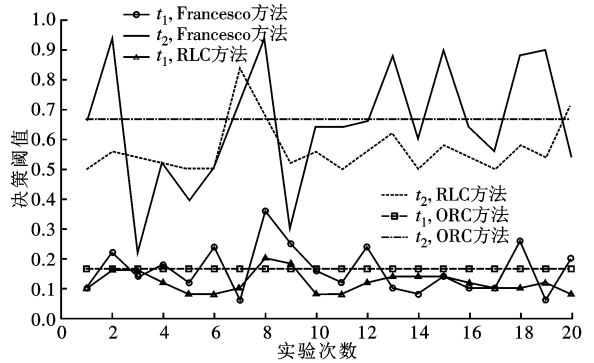


图 7 3 种方法的决策阈值对变化趋势

由图 7 可以看出,Francesco 方法的决策代价和决策阈值对出现无规律的波动,且有时在其确定的决策阈值对下分类器是无意义的。利用 RLC 方法可以确定最低的决策损失,但是其对应的决策阈值对同样出现无规律的变化。图 7 所示 Francesco 和 RLC 方法的阈值对的方差分别为 0.08、0.21 和 0.0013、0.007 5。ORC 方法不仅使分类造成的损失代价近似等于全局最小值,而且其决策阈值对由式(21)唯一确定,使得其具有很高的稳定性。

3.3.3 不同方法识别率、漏警率和可靠性比较 设正确分类样本数为  $t_s$ ,正类被正确识别的个数为  $t_p$ ,错误分类样本数为  $f_s$ ,正类识别为负类的个数为  $f_n$ ,总样本数为  $N$ 。定义识别率为  $a_{cc}$ 、可靠性为  $r_e$ 、漏警率为  $r_{in}$ ,其表达式如下

$$\left. \begin{aligned} a_{cc} &= \frac{t_s}{N} \\ r_e &= \frac{t_s}{t_s + f_s} \\ r_{in} &= \frac{t_p}{t_p + f_n} \end{aligned} \right\} \tag{25}$$

图 8~10 为 3 种方法在未经约减的数据上分别采用 10 轮 5 倍交叉验证重复实验 20 次得到的结果。由图 8~10 可以看出,ORC 方法的性能接近于 RLC 方法的性能,在所有的实验次数中,识别结果没有出现大幅度的波动,说明 ORC 方法具有较好的识别性能和稳定性,证明了本文方法的有效性。Francesco 方法识别结果的识别率、漏警率和可靠性均有明显

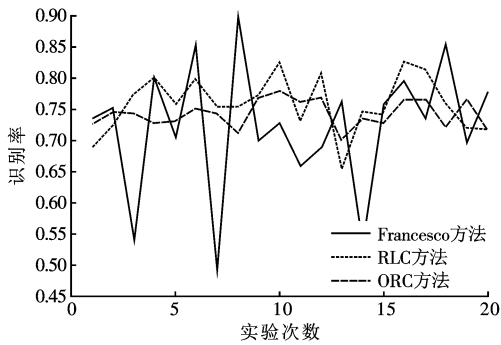


图 8 3 种方法的识别率对比结果

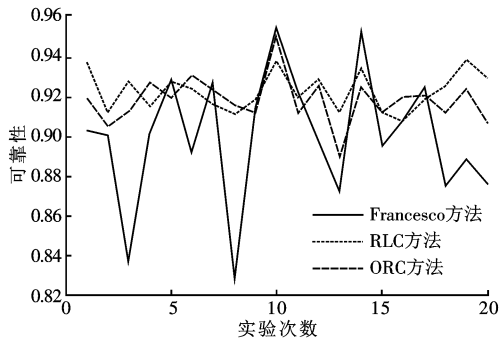


图 9 3 种方法的可靠性对比结果

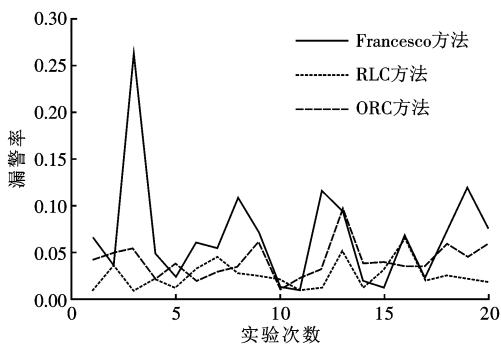


图 10 3 种方法的漏警率对比结果

的震荡,这是由于 Francesco 方法在训练过程中根据训练样本的不同以及得到 ROC 曲线的不完整性在确定阈值时有较大随机性导致的。

综上可知,ORC 方法不仅能使分类风险接近于实际能达到的最小值,而且其决策阈值由代价矩阵决定,比较稳定,表现出了较好的性能。

4 结 语

本文提出了一种引入拒识代价的最小风险弹道目标识别方法。从类条件概率密度的角度推导了真阳性率、虚警率与决策阈值之间的关系,并进一步推导了 ROC 曲线上任意一点的斜率与决策阈值之间的关系;分析了代价敏感分类器及引入拒识代价的分类器,确定了代价损失函数,推导了代价最小化所

需满足的条件,并将其与 ROC 曲线任意一点的切线斜率相结合,推导了最优化拒识阈值的确定方法。仿真实验表明,本文所提方法能够在保证代价最小的同时准确确定拒识阈值对,与传统方法相比,提高了稳定性和可靠性。

参考文献:

[1] 胡小生, 钟勇. 基于加权聚类质心的 SVM 不平衡分类方法 [J]. 智能系统学报, 2013, 8(3): 261-265.  
HU Xiaosheng, ZHONG Yong. Support vector machine imbalanced data classification based on weighted clustering centroid [J]. CAAI Transactions on Intelligent Systems, 2013, 8(3): 261-265.

[2] DATTA S, DAS S. Near-Bayesian support vector machines for imbalanced data classification with equal or unequal misclassification costs [J]. Neural Networks, 2015, 70(12): 39-52.

[3] JIANG L, LI C, WANG S S. Cost-sensitive Bayesian network classifiers [J]. Pattern Recognition Letters, 2014, 45(2): 211-216.

[4] IBANEZ A, BIELZA C, LARRANAGA P. Cost-sensitive selective naïve Bayes classifiers for predicting the increase of the h-index for scientific journals [J]. Neurocomputing, 2014, 135(4): 42-52.

[5] 吕淑静. 基于代价敏感学习的手写邮编和地址识别 [D]. 上海: 华东师范大学, 2014: 23-45.

[6] LI X, ZHAO H, ZHU W. A cost sensitive decision tree algorithm with two adaptive mechanisms [J]. Knowledge-Based Systems, 2015, 88(4): 24-33.

[7] 熊冰妍, 王国胤, 邓维斌. 分级式代价敏感决策树及其在手机换机预测中的应用 [J]. 山东大学学报, 2015, 45(5): 36-42.

XIONG Bingyan, WANG Guoyin, DENG Weibin. Hierarchical cost sensitive decision tree and its application in the prediction of the mobile phone replacement [J]. Journal of Shandong University, 2015, 45(5): 36-42.

[8] SUN Y M, KARNEL M S, WONG A, et al. Cost-sensitive boosting for classification of imbalanced data [J]. Pattern Recognition, 2007, 40(12): 3358-3378.

[9] 薛一哲, 王拓. 基于代价敏感 Adaboost 目标跟踪 [J]. 中国图像图形学报, 2016, 21(5): 544-555.  
XUE Yizhe, WANG Tuo. Object-tracking method based on improved cost-sensitive Adaboost [J]. Journal of Image and Graphics, 2016, 21(5): 544-555.

[10] CHATELAIN C, ADAM S. A multi-model selection framework for unknown and/or evolutive misclassification cost problems [J]. Pattern Recognition, 2010, 43(3): 815-823.

- [11] BERNARD S, CHATELAIN C. The multiclass ROC front method for cost-sensitive classification [J]. *Pattern Recognition*, 2016, 52(4): 46-60.
- [12] TAPKANA P, ÖZBAKIRA L, KULLUKA S, et al. A cost-sensitive classification algorithm: BEE-Miner [J]. *Knowledge-Based Systems*, 2016, 95(2): 99-113.
- [13] JU H, YANG X, YU H, et al. Cost-sensitive rough set approach [J]. *Information Sciences*, 2016, 355/356(4): 282-298.
- [14] 刘传武, 张智军, 毕笃彦. 雷达目标自动识别系统的拒识新算法 [J]. *系统工程与电子技术*, 2009, 31(8): 1846-1850.  
LIU Chuanwu, ZHANG Zhijun, BI Duyan. Novel refuse-recognition algorithm of radar automatic target recognition system [J]. *Systems Engineering and Electronics*, 2009, 31(8): 1846-1850.
- [15] 黄焱, 刘思颂, 孔瑞. 基于支持向量机的嵌入拒识代价的手写字符识别研究 [J]. *电子质量*, 2011(4): 5-7.  
HUANG Yao, LIU Sisong, KONG Rui. SVM-based hand-written character recognition with reject option [J]. *Electronics Quality*, 2011(4): 5-7.
- [16] FUMERA G, ROLI F, GIACINTO G. Multiple reject thresholds for improving classification reliability [C]// *Proceedings of the Joint International Workshops on Advances in Pattern Recognition*. Berlin, Germany: Springer-Verlag, 2000: 863-871.
- [17] FUMERA G, ROLI F, GIACINTO G. Reject option with multiple thresholds [J]. *Pattern Recognition*, 2000, 33(12): 2099-2101.
- [18] 郑恩辉, 徐欢. 嵌入非对称拒识代价的二元分类算法 [J]. *控制与决策*, 2013, 28(6): 855-860.  
ZHENG Enhui, XU Huan. Binary classification algorithm with class-dependent reject cost [J]. *Control and Decision*, 2013, 28(6): 855-860.
- [19] TORTORELLA F. An optimal reject rule for binary classifiers [C]// *Proceedings of the Joint International Workshops on Advances in Pattern Recognition*. Berlin, Germany: Springer-Verlag, 2000: 611-620.

(编辑 武红江)

(上接第 110 页)

- [11] 周雪倩, 吴晓富, 余训健. 融入 LDPC 纠错机制的认知无线电物理层认证分析 [J]. *计算机科学*, 2017, 44(7): 89-93.  
ZHOU Xueqian, WU Xiaofu, YU Xunjian. LDPC coded primary transmitter authentication schemes in cognitive radio networks [J]. *Computer Science*, 2017, 44(7): 89-93.
- [12] DAI C P, YANG J X, QIN Y N, et al. Physical layer authentication algorithm based on SVM [C]// *Proceedings of 2016 2nd IEEE International Conference on Computer and Communications*. Piscataway, NJ, USA: IEEE, 2016: 1597-1601.
- [13] WANG N, JIANG T, LV S C, et al. Physical-layer authentication based on extreme learning machine [J]. *IEEE Communications Letters*, 2017, 21(7): 1557-1560.
- [14] 杨建喜, 戴楚屏, 姜亭亭, 等. 基于支持向量机的 4G 室内物理层认证算法 [J]. *计算机应用*, 2016, 36(11): 3103-3107.  
YANG Jianxi, DAI Chuping, JIANG Tingting, et al. 4G indoor physical layer authentication algorithm based on support vector machine [J]. *Journal of Computer Applications*, 2016, 36(11): 3103-3107.
- [15] LIU R H, WADE T. Securing wireless communications at the physical layer [M]. Berlin, Germany: Springer, 2010: 23-38.
- [16] JAKES W C. *Microwave mobile communications* [M]. Piscataway, NJ, USA: IEEE Press, 1994: 11-50.
- [17] COVER T M, THOMAS J A. *Elements of information theory* [M]. Beijing: Tsinghua University Press, 1991: 155-183.
- [18] CSISZÁR I, KÖRNER J. *Information theory: coding theorems for discrete memoryless systems* [M]. Cambridge, UK: Cambridge University Press, 1981: 83-107.
- [19] CHISAR I. Almost independence and secrecy Channel [J]. *Problemy Peredachi Informatsii*, 1996, 32(1): 48-57.

(编辑 刘杨)