

一种分析噪声背景下目标特征有效性的方法^{*}

王慧频 孙仲康

(国防科技大学电子技术系 长沙 410073)

摘 要 本文根据目标特征和噪声分布,在一定假设条件下推出了最小错误率(Bayes 错误率)的表达式,从而把特征分布、噪声大小直接与分类的最小错误率联系起来,给出了一种分析噪声背景下目标特征的有效性的方法。文中利用这种方法对目标和诱饵的红外辐射特征的有效性进行了分析。

关键词 目标特征,特征有效性,Bayes 错误率,红外辐射

分类号 TN911.23

A Method Used to Analyse the Effectiveness of Target Features in Noisy Background

Wang Huipin Sun Zhongkang

(Department of electronic technology, NUDT, Changsha, 410073)

Abstract According to the distributions of target features and noise, this paper presents the formula of the minimum error probability (Bayes error probability) in some conditions. This formula relates directly the minimum error probability to the distributions of features and noise and can be used to analyse the effectiveness of target features in noisy background. As an example, the effectiveness of IR. radiation features of target and decoy is analysed in the paper.

Key words target features, the effectiveness of features, Bayes error probability, IR. radiation

在一个目标识别系统中,目标特征的提取是至关重要的。目标特征是否有效,直接影响着分类结果的好坏。实际上,特征的分布与噪声的大小决定了分类系统的最小错误

^{*} 国家自然科学基金资助项目
1996 年 10 月 3 日收稿

率，这个错误率是固定的，跟识别算法无关。本文根据目标特征及噪声分布，在一定假设条件下推出了最小错误率（即 Bayes 错误率）的表达式，把特征分布、噪声大小直接与分类的最小错误率联系起来，从而给出了一种分析噪声背景下目标特征有效性的方法。

1 基本原理

1.1 Bayes 错误率的定义

错误率是表征一个分类系统性能的重要指标。对于特征矢量 X 的分类错误率实际上是平均错误率，它的定义如下：

$$P(e) = \int P(e, X) dX = \int P(e/X) p(X) dX \quad (1)$$

对于两类问题来说，若特征矢量 X 不属于 ω_1 类，则必属于 ω_2 类，即 $X \in \omega_1$ 或 $X \in \omega_2$ 。假设各类的先验概率 $P(\omega_1)$ 、 $P(\omega_2)$ 已知，那么错误率为

$$P(e) = \int_{R_1} P(\omega_2/X) p(X) dX + \int_{R_2} P(\omega_1/X) p(X) dX \quad (2)$$

根据 Bayes 公式

$$P(\omega/X) = \frac{p(X/\omega)P(\omega)}{p(X)} \quad (3)$$

错误率还可以写成：

$$P(e) = \int_{R_1} P(\omega_2) p(X/\omega_2) dX + \int_{R_2} P(\omega_1) p(X/\omega_1) dX \quad (4)$$

式 (2) 和式 (4) 中， R_i 表示把 X 分到 ω_i 的区域。等式右边第一项表示 $X \in \omega_1$ 的样式落入 R_1 从而错分到 ω_2 的概率，第二项表示 $X \in \omega_2$ 的样本落入 R_2 从而错分到 ω_1 的概率。显然， R_i 的划分是由分类算法决定的，在图 1 中，分界线在 t_2 处的错误率比在 t_1 处的错误率要小。实际上，分界线在 t_2 处的错误率是最小的，它就是图中阴影部分的面积，它是根据下面的 Bayes 分类规则得到的：

$$\text{若 } P(\omega_1/X) = \max_{i=1,2} \{P(\omega_i/X)\}, \text{ 则 } X \in \omega_1 \quad (5)$$

在本文中把由式 (4) 定义的错误率称为 Bayes 错误率。

Bayes 错误率是由特征的分布所决定的，是衡量特征有效性的最有效的指标。但是计算 Bayes 错误率并不是一件容易的事情，如果假设概率密度函数服从正态分布，错误率的计算就容易得多。下面就讨论概率密度函数服从正态分布情况下的 Bayes 错误率。

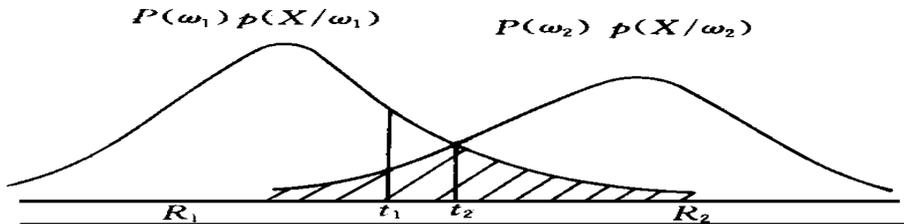


图 1 Bayes 错误率

1.2 一维情况下的错误率

假设各类的条件概率密度函数服从如下的正态分布，即

$$\begin{aligned} p(h|\omega) &= \frac{1}{2\pi\sigma} \exp\left\{-\frac{(h+\eta)^2}{2\sigma^2}\right\} \\ p(h|\omega) &= \frac{1}{2\pi\sigma} \exp\left\{-\frac{(h-\eta)^2}{2\sigma^2}\right\} \end{aligned} \quad (6)$$

那么

$$P_1(e) = \int_{t+\frac{\eta}{\sigma}} p(h|\omega) dh = \frac{1}{2\pi} \exp\left[-\frac{\xi^2}{2}\right] d\xi = \Phi\left[-\frac{t+\eta}{\sigma}\right] \quad (7)$$

$$P_2(e) = \Phi\left[\frac{t-\eta}{\sigma}\right]$$

Bayes 错误率为: $P(e) = P(\omega) P_1(e) + P(\omega) P_2(e)$ (8)

式中 $t = \ln\left[\frac{P(\omega_1)}{P(\omega_2)}\right]$, $\Phi(\cdot)$ 可查正态分布的累积分布函数表得到。

1.3 多维情况下的错误率

假设各类服从正态分布。采用负对数似然比的 Bayes 分类规则为

若 $h(X) = -\ln(l(X)) = -\ln p(X|\omega_1) + \ln p(X|\omega_2) = \ln\left[\frac{P(\omega_1)}{P(\omega_2)}\right]$, 则 $X|\omega_1$;

反之, 若 $h(X) = \ln\left[\frac{P(\omega_2)}{P(\omega_1)}\right]$, 则 $X|\omega_2$. (9)

$h(X)$ 是 X 的函数, X 是随机向量, 因此 h 是随机向量, 并且是一维的随机向量, 因此可利用前面的结果来计算错误率。

现在求 h 的均值和方差。由于

$$\begin{aligned} h(X) &= -\ln[P(X|\omega_1)] + \ln[P(X|\omega_2)] \\ &= \frac{1}{2}(X - U_1)^T \Sigma_1^{-1}(X - U_1) - \frac{1}{2}(X - U_2)^T \Sigma_2^{-1}(X - U_2) + \frac{1}{2} \ln\left[\frac{\Sigma_1}{\Sigma_2}\right] \end{aligned} \quad (10)$$

显然, $h(X)$ 是 X 的二次函数。进一步, 如果假设 $\Sigma_1 = \Sigma_2 = \Sigma$, $h(X)$ 就成为 X 的线性函数。这时, 由于 X 的各分量都服从一维正态分布, 其线性组合 h 也服从一维正态分布。对于 $P(h|\omega)$, 可以算出其一维正态分布的均值 η_1 和方差 σ_1 :

$$\eta_1 = E[h(X)|\omega_1] = -\frac{1}{2}(U_1 - U_2)^T \Sigma^{-1}(U_1 - U_2) \quad (11)$$

$$\sigma_1 = E[(h(X) - \eta_1)^2|\omega_1] = (U_1 - U_2)^T \Sigma^{-1}(U_1 - U_2)$$

令 $\eta = \frac{1}{2}(U_1 - U_2)^T \Sigma^{-1}(U_1 - U_2)$ (12)

则 $\eta_1 = -\eta$, $\sigma_1 = 2\eta$ (13)

同样可以得到 $P(h|\omega_2)$ 的均值 η_2 和方差 σ_2 :

$$\eta_2 = \eta, \quad \sigma_2 = 2\eta \quad (14)$$

利用前面的结果得到 Bayes 错误率:

$$P(e) = P(\omega)P_1(e) + P(\omega_2)P_2(e) \quad (15)$$

其中

$$\begin{aligned} P_1(e) &= \Phi\left[-\frac{t+\eta}{\sigma}\right] = \Phi\left[-\frac{t+\eta}{2\eta}\right] \\ P_2(e) &= \Phi\left[\frac{t-\eta}{\sigma}\right] = \Phi\left[\frac{t-\eta}{2\eta}\right] \end{aligned} \quad (16)$$

可以证明: $P(e)$ 是随 η 递减的。当 $\eta \rightarrow 0$ 时, $P(e) \rightarrow 1$; 当 $\eta \rightarrow \infty$ 时, $P(e) \rightarrow 0$ 。因此 η 越大, 错误率越小。由此可见, η 是分类系统的一个极为重要的指标。通常, 将 $2\eta = (U_1 - U_2)^T \Sigma^{-1} (U_1 - U_2) = J_M$ 称为 Mahalanobis 距离。

进一步, 若 X 的各分量相互独立, 即 Σ 为对角阵

$$\Sigma = \begin{bmatrix} \sigma_1^2 & 0 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & 0 & \dots & 0 \\ 0 & 0 & \dots & & \\ \dots & & & \dots & \dots \\ 0 & 0 & \dots & \dots & \sigma_d^2 \end{bmatrix}$$

并设 $U_1 = (\mu_{11}, \mu_{12}, \dots, \mu_{1d})^T$, $U_2 = (\mu_{21}, \mu_{22}, \dots, \mu_{2d})^T$

$$\text{那么 } \eta = \frac{1}{2} \sum_{i=1}^d \frac{(\mu_{1i} - \mu_{2i})^2}{\sigma_i^2} \quad (17)$$

从上式可以看出: 当特征维数 d 越大时, η 越大, 而 η 越大, 错误率越小。这就是说, 对于正态分布的情况, 当各特征相互独立时, 特征维数越大, Bayes 错误率越小。

1.4 噪声背景下的 η

在噪声背景下, 特征矢量 $X = (x_1, x_2, \dots, x_d)^T$ 由两部分组成:

$$X = S + N \quad (18)$$

其中, $S = (s_1, s_2, \dots, s_d)^T$ 为无噪声情况下的特征矢量, $N = (n_1, n_2, \dots, n_d)^T$ 为系统噪声信号。各特征分量 x_j 是由无噪声分量 s_j 和噪声分量 n_j 相加得到的, 即

$$x_j = s_j + n_j \quad j = 1, 2, \dots, d \quad (19)$$

假设 s_j 和 n_j 相互独立, 那么 x_j 的密度函数是 s_j 和 n_j 的密度函数的卷积, 即:

$$p_{x_j} = p_{s_j} * p_{n_j} \quad (20)$$

在 s_j 和 n_j 均服从正态分布的情况下, x_j 也服从正态分布, 且

$$\begin{aligned} \mu_{x_j} &= \mu_{s_j} + \mu_{n_j} \\ \sigma_{x_j}^2 &= \sigma_{s_j}^2 + \sigma_{n_j}^2 \end{aligned} \quad (21)$$

式中 μ_{x_j} 、 μ_{s_j} 、 μ_{n_j} 和 σ_{x_j} 、 σ_{s_j} 、 σ_{n_j} 分别是 x_j 、 s_j 、 n_j 的均值和方差。

本文中均假设噪声服从均值为零及方差为 σ_n^2 的正态分布, 但无噪声分量并不一定服从正态分布。为了便于分析, 在这里假设它们是近似于正态分布的。

因此

$$\mu_{x_j} = \mu_{s_j} + \mu_{n_j} = \mu_{s_j} \quad (22)$$

$$\eta = \frac{1}{2} \sum_{i=1}^d \frac{(\mu_{1si} - \mu_{2si})^2}{\sigma_{si}^2 + \sigma_{ij}^2} \quad (23)$$

无噪声信号的均值和方差可采用最大似然估计量来估计:

$$\begin{aligned} \mu_{1sj} &= \frac{1}{M} \sum_{k=1}^M x_{1j}^{(k)}, \sigma_{1sj}^2 = \frac{1}{M} \sum_{k=1}^M \left(x_{1j}^{(k)} - \mu_{1sj} \right)^2 \\ \mu_{2sj} &= \frac{1}{M} \sum_{k=1}^M x_{2j}^{(k)}, \sigma_{2sj}^2 = \frac{1}{M} \sum_{k=1}^M \left(x_{2j}^{(k)} - \mu_{2sj} \right)^2 \end{aligned} \quad (24)$$

如果噪声方差 σ_{ij} 利用信噪比 SNR 和信号的平均值 S_{av} 来得到:

$$\sigma_{ij} = \frac{S_{av}}{\text{SNR}} \quad (25)$$

这样, 式 (23) 可进一步写成:

$$\eta = 2 \sum_{i=1}^d \frac{\mu_{ij}^2}{\sigma_{si}^2 + \left(\frac{S_{av}}{\text{SNR}} \right)^2} \quad (26)$$

这样, 对于两类在理想情况下得到的特征矢量, 在已知噪声方差 (或信噪比) 的情况下, 采用式 (24) 和式 (23) 可得到 η , 进而可得到 Bayes 错误率。 η 不仅表征了两类特征矢量的分离程度, 而且还蕴含了特征对噪声的敏感程度, 并且与错误率有直接的关系。

若假设 $P(\omega) = P(\bar{\omega}) = 1/2$, 那么 $t = 0$, 因此

$$P_e = \Phi \left[- \frac{\eta}{2} \right] = \Phi \left[- \frac{J_M}{2} \right] \quad (27)$$

2 举例

本文针对大气层外目标与其伴飞诱饵的识别, 选择了三个地面观测站的两个红外谱段 ($3 \sim 5 \mu\text{m}$ 、 $8 \sim 12 \mu\text{m}$) 的辐射通量及辐射通量的时间变化量作为识别特征。下面就利用前面推出的方法分析这些红外辐射特征的有效性。

首先利用文献 [1] 中的方法计算出在一组基本参数下 (这些参数包括太阳经度、纬度、目标诱饵的形状尺寸参数、热物理参数等) 的目标诱饵的光谱辐射通量, 然后提取红外特征 (即三个地面观测站的两个红外谱段 $3 \sim 5 \mu\text{m}$ 、 $8 \sim 12 \mu\text{m}$ 的辐射通量及辐射通量的时间变化量), 设定一个信噪比, 采用式 (24)、(25)、(26) 计算出 η , 进一步得到 Bayes 错误率 P_e , 逐步减少信噪比, 可得到一组计算结果 (见表 1)。从计算结果可以看出: 目标诱饵的红外特征在无噪声情况下可以完全分开, 但它们对噪声比较敏感, 随着噪声的增加, 错误率迅速上升。这说明两类的红外特征处于两个互不交叠但距离较近的区域中。

表 1 目标诱饵的一个红外特征序列的 η 值和 Bayes 错误率

SNR		10	5	3
$\eta/2$	> 100	8.6	2.5	0.94
P_e (%)	0	0.2	5.7	16.6

变化目标诱饵的计算参数得到共 20 个红外特征序列, 用同样的方法计算其 Bayes 错

误率 (见表 2)。此时的结果显然要比表 1 的结果要差。这说明, 对于本文选择的红外特征, 随着目标诱饵的类型的增加, 其有效性将随之降低。

表 2 各种不同参数下 型目标诱饵对的 η 值和 Bayes 错误率

SNR		10	5	3
$\eta/2$	2.4	1.05	0.6	0.35
$Pe (%)$	6.3	15.6	21.8	27.8

把 12 个特征分量分成两部分: 一部分是辐射量, 另一部分是辐射量的时间变化量。分别计算各个部分的 η 值和 Bayes 错误率, 从而比较它们对于区分目标诱饵的有效性。表 3 是目标诱饵的一个特征序列 (与表 1 的序列相同) 的两个部分的 η 值和 Bayes 错误率。

表 3 特征矢量两个部分的 η 值和 Bayes 错误率

	辐射量部分	辐射量的时间变化量部分
$\eta/2$ (SNR=)	> 100	> 100
$Pe (%)$ (SNR=)	0	0
$\eta/2$ (SNR= 5)	2.5	0.0072
$Pe (%)$ (SNR= 5)	5.7	46

从表 3 可以看出: 在无噪声情况下, 两个部分的特征对于区分目标和诱饵都是有效的。但在噪声情况下, 辐射能量的时间变化量作为特征的有效性将大大降低。

利用上面的方法还可以根据需要分析各种情况下特征的有效性。

3 结束语

虽然本文的方法是在样本服从正态分布的情况下推出的, 但不管样本服从何种分布, 式 (24) 计算出的 μ 是各种样本的中心, σ 表征了样本的分散程度, 因而式 (23) 计算出的 η 值同样可以表征两类样本的分离程度, 可以用来衡量特征的有效性, 并且可用它来估算错误率。这就是说, 式 (24)、(23) 构成了一种衡量特征有效性的方法, 它不仅可以用以表征两类特征的分离程度, 而且还蕴含了特征对噪声的敏感程度, 并且可用以估算 Bayes 错误率。这是一种简单、直观而又实用的分析噪声背景下特征有效性的方法。

参 考 文 献

- 1 王慧频. 大气层外目标和诱饵的红外辐射特性及其识别方法研究: [博士论文]. 长沙: 国防科技大学电子技术系, 1996
- 2 边肇祺. 模式识别. 北京: 清华大学出版社, 1986

(责任编辑 潘 生)