

文章编号:1001-9081(2012)04-1025-05

doi:10.3724/SP.J.1087.2012.01025

基于散度比例准则的高分辨距离像特征提取

刘 敏^{1*}, 赵 峰², 刘 逸³

(1. 西安邮电学院 电子工程学院, 西安 710121; 2. 山东工商学院 计算机科学与技术学院, 山东 烟台 264005;

3. 西安电子科技大学 电子工程学院, 西安 710071)

(* 通信作者电子邮箱 zyhalj1975@163.com)

摘要:针对传统线性判别分析(LDA)的子空间倾向于保留大类间距离类对的可分性,而丢弃小类间距离类对的可分性的问题,基于子空间应均衡保留各类对可分性的思想,提出一种新的准则——散度比例(PD)准则。PD准则为各对子空间散度与原空间散度之比的均值,并推导出最大化PD准则的线性判别分析(PD-LDA)的求解过程。采用PD-LDA对高分辨距离像(HRRP)的幅度谱进行特征提取,基于外场实测数据,分别训练了最小欧氏距离分类器和支持向量机(SVM)分类器,两种分类器的识别结果均表明,PD-LDA相比LDA,可显著降低数据维数并有效提高识别率。

关键词:雷达自动目标识别;散度比例;线性判别分析;特征提取;高分辨距离像

中图分类号: TP391.4; TN957.52 文献标志码:A

HRRP feature extraction based on proportion of divergence criterion

LIU Jing^{1*}, ZHAO Feng², LIU Yi³

(1. School of Electronic Engineering, Xi'an University of Posts and Telecommunications, Xi'an Shaanxi 710121, China;

2. School of Computer Science and Technology, Shandong Institute of Business and Technology, Yantai Shandong 264005, China;

3. School of Electronic Engineering, Xidian University, Xi'an Shaanxi 710071, China)

Abstract: Traditional Linear Discriminant Analysis (LDA) faces the problem of tending to keep the separability of the class pairs having large within-class distances, while discarding the separability of those having small within-class distances. Based on the viewpoint that the feature subspace should uniformly keep the separability of each class pair, a new criterion, i.e., the Proportion of Divergence (PD), was presented. PD criterion was the mean of the proportion of the subspace divergence to original space divergence of each class pair. The solution of the Linear Discriminant Analysis (LDA) maximizing PD criterion (PD-LDA) was also presented. PD-LDA was used to perform feature extraction in the amplitude spectrum space of High Resolution Range Profile (HRRP). Shortest Euclidean distance classifier and Support Vector Machine (SVM) classifier were designed to evaluate the recognition performance. The experimental results for measured data show that, compared with traditional LDA, PD-LDA reduces data dimension remarkably and improves recognition rate effectively.

Key words: Radar Automatic Target Recognition (RATR); Proportion of Divergence (PD); Linear Discriminant Analysis (LDA); feature extraction; High Resolution Range Profile (HRRP)

0 引言

传统线性判别分析(Linear Discriminant Analysis, LDA)最大化 Fisher 准则^[1-2],其解是 $S_w^{-1}S_b$ 前 d 个最大特征值对应的特征向量^[3], S_w 和 S_b 分别为原始数据的类内散布矩阵和类间散布矩阵。多类情况下,类间距离大的类对将支配特征值分解,使得传统 LDA 子空间倾向于保留大类间距离类对的可分性,而丢弃小类间距离类对的可分性,导致低的整体识别率。目前,最大化 Fisher 准则的 LDA,多通过在计算类间散布矩阵时为各类对加权来解决这一问题。文献[4]通过加权使得 Fisher 准则近似等于子空间中各类对的平均贝叶斯正确率。但是,最大化子空间中各类对的平均贝叶斯正确率,并不能保证各类对的贝叶斯正确率均较大,且通过加权,Fisher 准则只能近似逼近平均贝叶斯正确率,在理论上尚有欠缺。文献[5]通过加权抑制可分性大的类对对类间散布矩阵的贡献,增加可分性小的类对的贡献,其权值根据实验结果确定,

该方法的权值确定过于经验化,缺少理论支持。文献[6-7]根据子空间中各类对的理论贝叶斯正确率和实际给定分类器的经验正确率确定权值,该方法将特征提取与分类识别合并成一个过程,不具有普遍性。总之,目前对传统 LDA 的改进工作,仍集中在 Fisher 准则的框架之下。

LDA 要求,在降低数据维数的同时,各类对的可分性在子空间中均尽可能地保留。显然,最大化 Fisher 准则不能满足这一要求。针对 Fisher 准则的不足,本文基于子空间应均衡保留各类对可分性的思想,提出一种新的准则——散度比例(Proportion of Divergence, PD)准则,即各类对子空间散度与原空间散度之比的均值,并推导出最大化 PD 准则的线性判别分析(LDA maximizing PD criterion, PD-LDA)的求解过程。散度反映概率距离大小,散度越大,两类可分性越大,且贝叶斯正确率与散度是单调增长关系^[8-9]。采用散度作为可分性的度量,PD 准则是子空间中各类对可分性保留比例的均值,而 Fisher 准则是子空间中各类对可分性的均值,PD 准则

收稿日期:2011-09-15;修回日期:2011-11-22。基金项目:国家自然科学基金资助项目(61003199);中央高校基本科研业务费专项(K50510020015);陕西省教育厅自然科学专项(2010JK821);西安邮电学院博士启动基金资助项目(000-1271)。

作者简介:刘敏(1975-),女,安徽萧县人,讲师,博士,主要研究方向:智能信息处理、雷达自动目标识别;赵峰(1974-),男,山东梁山人,副教授,博士,主要研究方向:机器学习、雷达自动目标识别;刘逸(1976-),男,安徽萧县人,讲师,博士研究生,主要研究方向:计算智能。

比 Fisher 准则更准确地描述了线性变换后, 子空间中各类对可分性的保留情况。PD-LDA 最大化 PD 准则, 相比传统 LDA, PD-LDA 子空间均衡地保留类对间可分性。不同于采用加权 Fisher 准则的 LDA 方法, PD-LDA 不受 Fisher 准则的框架限制。

高分辨率距离像(High Resolution Range Profile, HRRP)是目标的重要结构特征, 是目前雷达自动目标识别(Radar Automatic Target Recognition, RATR)领域的研究热点^[9-12]。HRRP 维数较高, 在分类识别前应对其进行降维。采用 PD-LDA 对 HRRP 进行特征提取, 基于外场实测数据, 最小欧氏距离分类器和支持向量机(Support Vector Machine, SVM)分类器的识别结果均表明, 相比传统 LDA, PD-LDA 可提取出更有效的分类特征, 更好地对数据进行降维, 显著提高识别效率。

1 Fisher 准则和传统 LDA

1.1 传统 LDA

设 n 维数据空间中有 c 类样本, 传统 LDA 寻找线性变换矩阵 $\mathbf{W} \in \mathbb{R}^{n \times d}$, 使得采用 $\mathbf{y} = \mathbf{W}^T \mathbf{x}$ 将原始数据 \mathbf{x} 映射到 d 维 ($d < n$) 子空间后, Fisher 准则

$$J_F(\mathbf{W}) = \text{tr}[(\mathbf{W}^T \mathbf{S}_w \mathbf{W})^{-1} (\mathbf{W}^T \mathbf{S}_b \mathbf{W})] \quad (1)$$

最大 \mathbf{S}_w 和 \mathbf{S}_b 分别为原始数据的类内散布矩阵和类间散布矩阵。 $\mathbf{S}_w = \sum_{i=1}^c P_i \boldsymbol{\Sigma}_i$, P_i 、 $\boldsymbol{\Sigma}_i$ 分别为第 i 类的先验概率和协方差矩阵; $\mathbf{S}_b = \sum_{i=1}^c P_i (\mathbf{m}_i - \mathbf{m})(\mathbf{m}_i - \mathbf{m})^T$, \mathbf{m}_i 为第 i 类的均值, $\mathbf{m} = \sum_{i=1}^c P_i \mathbf{m}_i$ 为整体均值。传统 LDA 的解是 $\mathbf{S}_w^{-1} \mathbf{S}_b$ 前 d 个最大特征值对应的特征向量。

1.2 散度

根据基于贝叶斯决策理论的最大似然比判决准则, i, j 类对的对数似然比定义为:

$$I_{ij}(\mathbf{x}) = \ln \left[\frac{p(\mathbf{x} | \omega_i)}{p(\mathbf{x} | \omega_j)} \right]$$

其中 $p(\mathbf{x} | \omega_i)$ 和 $p(\mathbf{x} | \omega_j)$ 分别为类 i 与类 j 的类条件概率密度函数。令 $I_{ij}(\mathbf{x})$ 和 $I_{ji}(\mathbf{x})$ 分别为对数似然比给类 i 与类 j 提供的平均可分性信息, 散度 J_D 定义为区分类 i 与类 j 的总平均信息^{[8]393}, 有:

$$\begin{aligned} J_D &= I_{ij}(\mathbf{x}) + I_{ji}(\mathbf{x}) = \int [p(\mathbf{x} | \omega_i) - p(\mathbf{x} | \omega_j)] \\ &\quad \ln \left[\frac{p(\mathbf{x} | \omega_i)}{p(\mathbf{x} | \omega_j)} \right] d\mathbf{x} \end{aligned}$$

当类 i 与类 j 均服从正态分布时:

$$\begin{aligned} J_D &= \frac{1}{2} \text{tr}(\boldsymbol{\Sigma}_i \boldsymbol{\Sigma}_j^{-1} + \boldsymbol{\Sigma}_j \boldsymbol{\Sigma}_i^{-1} - 2\mathbf{I}) + \\ &\quad \frac{1}{2} (\mathbf{m}_i - \mathbf{m}_j)^T (\boldsymbol{\Sigma}_i^{-1} + \boldsymbol{\Sigma}_j^{-1}) (\mathbf{m}_i - \mathbf{m}_j) \end{aligned}$$

若 i, j 两类协方差矩阵满足 $\boldsymbol{\Sigma}_i = \boldsymbol{\Sigma}_j = \boldsymbol{\Sigma}$, 则有:

$$J_D = (\mathbf{m}_i - \mathbf{m}_j)^T \boldsymbol{\Sigma}^{-1} (\mathbf{m}_i - \mathbf{m}_j) \quad (2)$$

此时, i, j 两类之间的贝叶斯错误率 $P(e)$ 与散度之间满足 $P(e) = \int_{-\infty}^{\infty} \left(\frac{1}{\sqrt{2\pi}} \right) \exp(-y^2/2) dy$, 显然贝叶斯错误率与散度是单调下降关系。

由于散度具有如下性质: 1) 散度是非负的。2) 散度的大小反映了两类的类概率密度与先验概率之积的重叠程度, 即:

重叠越大, 散度越小, 贝叶斯错误率越高, 两类越不容易区分; 重叠越小, 散度越大, 贝叶斯错误率越低, 两类越容易区分。3) 当两类的类概率密度与先验概率之积完全重叠时, 散度等于 0; 完全不重叠时, 散度取最大值。因此, 散度可作为两类间可分性的度量, 且这种度量基于类概率密度分布, 与分类错误率有直接联系。

1.3 Fisher 准则的不足

Fisher 准则^{[4]763} 可重写为:

$$J_F(\mathbf{W}) = \sum_{i=1}^{c-1} \sum_{j=i+1}^c P_i P_j \text{tr}((\mathbf{W}^T \mathbf{S}_w \mathbf{W})^{-1} (\mathbf{W}^T \mathbf{S}_{ij} \mathbf{W})) \quad (3)$$

其中: $\mathbf{S}_{ij} = \mathbf{m}_{ij} \mathbf{m}_{ij}^T$, $\mathbf{m}_{ij} = \mathbf{m}_i - \mathbf{m}_j$ 。假设 \mathbf{S}_w 为单位矩阵, \mathbf{W} 为正交归一化矩阵, 这两个假设不影响下面的结论, 则根据式(3), Fisher 准则可表达为:

$$J_F(\mathbf{W}) = \sum_{i=1}^{c-1} \sum_{j=i+1}^c P_i P_j (\mathbf{W}^T \mathbf{m}_{ij})^T (\mathbf{W}^T \mathbf{m}_{ij}) \quad (4)$$

由式(2), 此时 i, j 类对的散度为:

$$J_D = (\mathbf{m}_i - \mathbf{m}_j)^T (\mathbf{m}_i - \mathbf{m}_j) = \mathbf{m}_{ij}^T \mathbf{m}_{ij}$$

由式(4)可见, Fisher 准则是子空间中各类对的平均散度。将散度作为两类可分性的度量, 传统 LDA 最大化 Fisher 准则, 使得子空间中各类对的平均可分性最大, 但这并不能保障各类对的可分性均被很好地保留。

当各类先验概率相同时, Fisher 准则等于子空间中各类对散度之和乘以一个常数, 最大化 Fisher 准则, 即是最大化子空间中各类对散度之和。因为在多个散度相加的过程中, 只要有一个值很大, 就会使 Fisher 准则很大, 从而掩盖了散度较小的类对, 使得原空间中散度大的类对支配了 $\mathbf{S}_w^{-1} \mathbf{S}_b$ 的特征值分解, 由 $\mathbf{S}_w^{-1} \mathbf{S}_b$ 前 d 个最大特征值对应的特征向量所支撑的子空间, 倾向于保留原空间散度大的类对的散度, 而丢弃散度小的类对的散度。即传统 LDA 的子空间倾向于保留原本可分性大的类对的可分性, 而丢弃可分性小的类对的可分性, 导致低的整体识别率。

图 1 所示为 3 类, 2 维, 每类 300 个样本的情况下, 传统 LDA 和 PD-LDA 的特征提取结果对比。如图 1(a)所示, 将 3 类数据向传统 LDA 第 1 特征方向上投影, 两条点划线均垂直于第 1 特征方向, 第 1 类数据将位于第 1 特征方向上的两条点划线之间的范围。可见, 受第 3 类(野类)影响, 第 1、2 类在传统 LDA 的第 1 特征方向上发生严重重叠, 几乎不可分。

2 PD 准则和 PD-LDA

针对 Fisher 准则的不足, 基于子空间应均衡保留各类对可分性的思想, 本文提出一种新的准则——散度比例(PD)准则, 并推导出最大化 PD 准则的线性判别分析(PD-LDA)的求解过程。

2.1 PD 准则和 PD-LDA 的原理

本文提出的 PD 准则定义为:

$$J_{PD}(\mathbf{W}) = \sum_{i=1}^{c-1} \sum_{j=i+1}^c P_i P_j \frac{(\mathbf{W}^T \mathbf{m}_{ij})^T (\mathbf{W}^T \mathbf{S}_w \mathbf{W})^{-1} (\mathbf{W}^T \mathbf{m}_{ij})}{\mathbf{m}_{ij}^T \mathbf{S}_w^{-1} \mathbf{m}_{ij}} \quad (5)$$

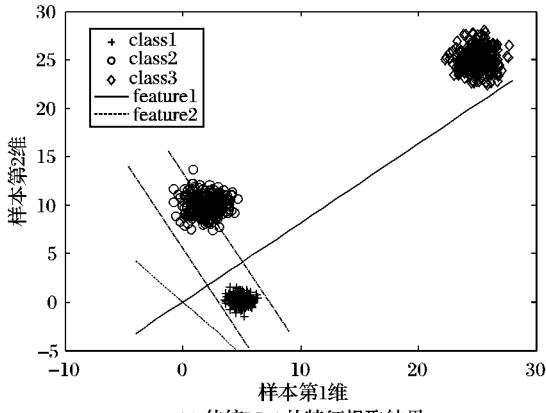
其中: 分子为子空间中 i, j 类对的散度, 分母为原空间中 i, j 类对的散度。PD 准则为各类对子空间散度与原空间散度之比的均值, 即子空间中各类对可分性保留比例的均值。

下面推导最大化 PD 准则的线性判别分析——PD-LDA 的求解过程。先白化 \mathbf{S}_w : 若 \mathbf{S}_w 的特征向量矩阵和特征值矩阵

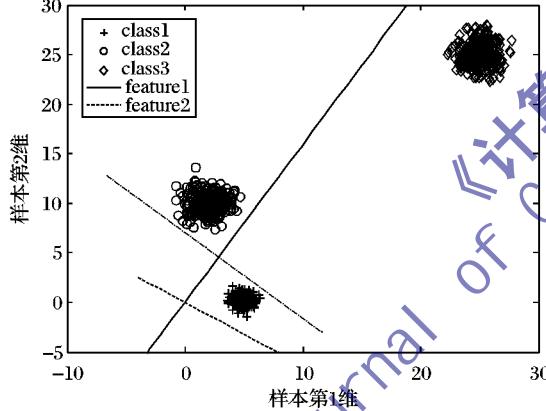
分别为 U_w 与 A_w , 则有 $W_1 = U_w A_w^{-1/2}$ 满足 $W_1^T S_w W_1 = I_n$, I_n 为 n 阶单位矩阵。设由正交归一化矩阵 W_2 将白化 S_w 后的数据映射到子空间, 因此整体映射矩阵 $W = W_1 W_2$, 代入式(5) 有:

$$\begin{aligned} J_{PD}(W) &= \\ &\sum_{i=1}^{c-1} \sum_{j=i+1}^c P_i P_j \frac{[(W_1 W_2)^T m_{ij}]^T [(W_1 W_2)^T m_{ij}]}{m_{ij}^T S_w^{-1} m_{ij}} = \\ &\sum_{i=1}^{c-1} \sum_{j=i+1}^c P_i P_j \frac{(W^T m_{ij})^T (W^T m_{ij})}{m_{ij}^T S_w^{-1} m_{ij}} \end{aligned} \quad (6)$$

在式(6) 中, 令 $m'_{ij} = W^T m_{ij}$, 则 m'_{ij} 为 PD-LDA 子空间中 i, j 类对的类心差向量, 这样的类心差向量共有 $c(c-1)/2$ 个, 它们的内积表示了 $J_{PD}(W)$ 。



(a) 传统LDA的特征提取结果



(b) PD-LDA的特征提取结果

图1 传统LDA 和 PD-LDA 特征提取结果对比

将 m'_{ij} 和 m_{ij} 的下标“ ij ”用一个下标代替, 令:

$$\begin{aligned} d'_1 &= m'_{12}, d'_2 = m'_{13}, \dots, d'_{[c(c-1)/2]} = m'_{[c-1][c]} \\ d_1 &= m_{12}, d_2 = m_{13}, \dots, d_{[c(c-1)/2]} = m_{[c-1][c]} \end{aligned}$$

内积 $r_{gh} = (d'_g)^T d'_h$ 表示了 $J_{PD}(W)$, 其中包括 $g \leq h$ 所有可能组合, 由于 $(d'_g)^T d'_h = (d'_h)^T d'_g$, 因此可只用 $g \leq h$ 情况下的 $(d'_g)^T d'_h$ 将 $J_{PD}(W)$ 表达为:

$$J_{PD}(W) = f(r_{11}, r_{12}, \dots, r_{[c(c-1)/2][c(c-1)/2]}) \quad (7)$$

其中 $r_{gh} = (d'_g)^T d'_h (g \leq h)$ 。

使式(7) 最大, 可求解:

$$\delta f = \sum_{g=1}^{c(c-1)/2} \sum_{h=g}^{c(c-1)/2} \frac{\partial f}{\partial r_{gh}} \Big|_{r_{gh}=r_{gh}^*} \delta r_{gh} = 0$$

其中 r_{gh}^* 为 r_{gh} 的最优值, 则最大化 PD 准则 $J_{PD}(W)$ 等价于最大化另一准则 J , 即:

$$J_{PD}(W) \cong J = \sum_{g=1}^{c(c-1)/2} \sum_{h=g}^{c(c-1)/2} \beta_{gh} r_{gh}$$

其中符号“ \cong ”表示等价于, $\beta_{gh} = \frac{\partial f}{\partial r_{gh}} \Big|_{r_{gh}=r_{gh}^*}$ 。将准则 J 写成

矩阵形式:

$$\begin{aligned} J_{PD}(W) &\cong J = \text{tr}(BR) = \text{tr}[B(D')^T D'] = \\ &\text{tr}[B(W^T D_0)^T (W^T D_0)] = \\ &\text{tr}[W^T D_0 B D_0^T W] = \\ &\text{tr}[W^T \Psi_b W] = \\ &\text{tr}[(W^T S_w W)^{-1} (W^T \Psi_b W)] \end{aligned} \quad (8)$$

其中:

$$B = \begin{bmatrix} \beta_{11} & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & \beta_{[c(c-1)/2][c(c-1)/2]} \end{bmatrix} =$$

$$\begin{bmatrix} P_1 P_2 & \cdots & 0 \\ \vdots & & \vdots \\ \frac{P_1 P_2}{m_{12}^T S_w^{-1} m_{12}} & \cdots & 0 \\ 0 & \cdots & \frac{P_{c-1} P_c}{m_{[c-1][c]}^T S_w^{-1} m_{[c-1][c]}} \end{bmatrix}$$

$$R = \begin{bmatrix} r_{11} & \cdots & r_{[1][c(c-1)/2]} \\ \vdots & & \vdots \\ r_{[1][c(c-1)/2]} & \cdots & r_{[c(c-1)/2][c(c-1)/2]} \end{bmatrix}$$

$$D' = [d'_1, \dots, d'_{[c(c-1)/2]}]$$

$$D_0 = [d_1, \dots, d_{[c(c-1)/2]}]$$

$$\Psi_b = D_0 B D_0^T$$

比较式(8)与式(1)可见, 相比 Fisher 准则, PD 准则的等价形式用广义类间散布矩阵 Ψ_b 代替类间散布矩阵 S_b 。因此, PD-LDA 的解是 $S_w^{-1} \Psi_b$ 前 d 个最大特征值对应的特征向量。将 PD 准则和 PD-LDA 总结如下。

PD-LDA 寻找线性变换矩阵 $W \in \mathbb{R}^{n \times d}$, 使得采用 $y = W^T x$ 将 n 维原始空间的数据 x 映射到 d 维 ($d < n$) 特征子空间后, 式(5) 定义的 PD 准则 $J_{PD}(W)$ 最大。PD-LDA 的解是 $S_w^{-1} \Psi_b$ 前 d 个最大特征值对应的特征向量。

2.2 PD 准则与 Fisher 准则的比较

将 PD 准则与 Fisher 准则相比, 可见:

1) PD 准则采用类对的子空间可分性与原空间可分性之比, 描述各类对可分性的保留; Fisher 准则采用子空间中类对可分性的绝对大小, 描述各类对可分性的保留, 前者比后者更合理。

2) 广义类间散布矩阵 Ψ_b 抑制大类间距离类对对类间散布矩阵的贡献, 增强小类间距离类对的贡献, 使得 $S_w^{-1} \Psi_b$ 的特征值分解不再受大类间距离类对的支配。因此相比传统 LDA, PD-LDA 子空间均衡地保留类对间的可分性。

如图 1(b) 所示, 将 3 类数据向 PD-LDA 的第 1 特征方向上投影, 点划线垂直于第 1 特征方向, 第 1、2 类数据将分别位于第 1 特征方向上的点划线的两侧, 可见, 第 1、2 类在 PD-LDA 的第 1 特征方向上没有重叠。因此, 与传统 LDA 相比, 所有类对在 PD-LDA 的第 1 特征方向上均能完全分开。

2.3 PD-LDA 的实现步骤

步骤 1 计算类内散布矩阵 S_w 。

步骤 2 白化 S_w 。设 S_w 的特征向量矩阵和特征值矩阵分别为 U_w 和 A_w , 即 $U_w^T S_w U_w = A_w$, 则有 $W_1 = U_w A_w^{-1/2}$ 满足 $W_1^T S_w W_1 = I_n$ 。

步骤 3 采用式(9) 计算广义类间散布矩阵 Ψ_b 。

步骤 4 对角化 $W_1^T \Psi_b W_1$, 设 W_2 的列向量为 $W_1^T \Psi_b W_1$ 前

d 个最大特征值对应的特征向量。

步骤 5 线性变换矩阵 $\mathbf{W} = \mathbf{W}_1 \mathbf{W}_2$, 原始 n 维样本 \mathbf{x} 经
 $\mathbf{y} = \mathbf{W}^T \mathbf{x}$ 映射到 d 维 PD-LDA 特征子空间。

3 基于 PD-LDA 的 HRRP 特征提取

3.1 HRRP

雷达采用宽频带信号后, 从目标接收到的是沿距离分布开的一维距离像。HRRP 有很强的方位、平移和幅度敏感性。

本文按照在视角变化 $\Delta\varphi \leq \frac{\Delta R}{L_x}$ 的角域范围内划分观测区间来解决方位敏感性^[13], 将观测区间内的 HRRP 序列定义为一帧距离像, 代表相应的一个角域。以 HRRP 的幅度谱作为距离像的平移不变特征以解决平移敏感性。对距离像的幅度谱进行归一化以解决幅度敏感性。

本文采用 PD-LDA 对 HRRP 的幅度谱进行特征提取, 并在特征子空间中设计了最小欧氏距离分类器和 SVM 多类分类器进行分类识别。

3.2 特征提取步骤

1) HRRP 预处理。

训练数据: ①划分观测角域, 建立各帧距离像。②对距离像均进行快速傅里叶变换(Fast Fourier Transformation, FFT), 并归一化各距离像幅度谱。设训练数据共有 T 个目标, 每目标有 F 帧。测试数据: 对任一测试距离像进行 FFT, 并归一化其幅度谱, 记为 \mathbf{x} 。

2) 特征提取。以一帧为一类进行特征提取, 即特征提取时数据类别数等于帧总数 TF , 采用 2.3 节介绍的 PD-LDA 实现步骤进行距离像幅度谱特征提取, 得线性变换矩阵 \mathbf{W} 。

3) 建立子空间模板库。将所有经预处理的训练数据采用式(10)映射到子空间; 再以各帧均值为模板建立模板库 $\mathbf{V}_{TF} = \{\mathbf{v}_i = (\boldsymbol{\mu}_i, b_i), i = 1, 2, \dots, TF\}$, 其中 $\boldsymbol{\mu}_i$ 为第 i 帧 HRRP 的归一化幅度谱在 PD-LDA 子空间中的均值; b_i 为该模板对应的目标类别号, 且 $b_i \in \{1, 2, \dots, T\}$ 。

4) 将测试数据映射到子空间: 采用式(10), 将任意测试数据 \mathbf{x} 映射到 PD-LDA 子空间后记为 \mathbf{y} 。

4 基于外场实测数据的实验结果

4.1 实测数据介绍

采用某研究所的 ISAR 实验雷达实测飞机数据, 雷达和飞机的参数如表 1 所示, 其中雷达参数的中心频率为 5520 MHz; 信号带宽为 400 MHz。三类飞机的飞行轨迹在地平面上的投影如图 2 所示, 图中给出了雷达的相对位置, 可大致估计飞机的方位角。其中, 雅克-42 第 2、5 段, 安-26 第 5、6 段, 奖状第 6、7 段数据作为训练数据, 其他各段作为测试数据。训练数据段基本包含了测试数据段内各种方位角的情况, 但俯仰角有所不同。训练数据共有 $T = 3$ 类目标, 每类目标有 $F = 50$ 个角域。距离像的幅度谱维数为 256。

表 1 ISAR 实验飞机和测试参数

飞机类型	机长/m	机宽/m	机高/m
雅克-42	36.38	34.88	9.83
安-26	23.80	29.20	9.83
奖状	14.40	15.90	4.57

4.2 实验结果

按照 3.2 节中的步骤进行 HRRP 特征提取, 然后在 PD-LDA 子空间中进行分类识别, 并与传统 LDA 子空间和幅度谱原空间的识别结果进行比较。

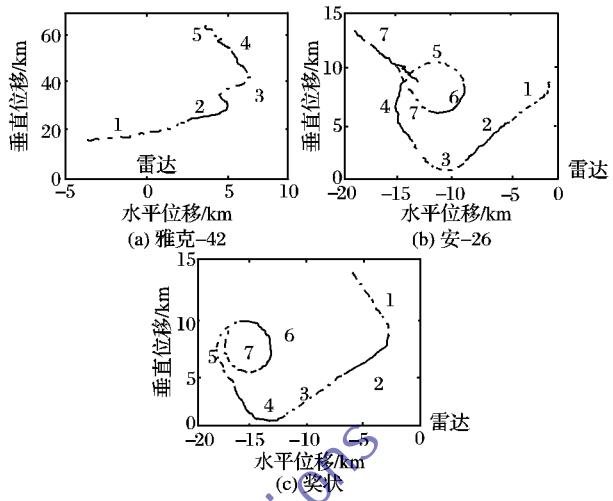


图 2 三类飞机的飞行轨迹

图 3 为在 PD-LDA 与传统 LDA 的不同维数的距离像幅度谱子空间, 采用最小欧氏距离分类器的平均识别率对比。由图 3 可见, PD-LDA 的 5 维子空间中的平均识别率为 87.25%, 明显大于传统 LDA 的 149 维子空间的平均识别率 85.25%; 随着子空间维数的增大, PD-LDA 与传统 LDA 的识别率趋于相同。以下, 本文将 PD-LDA 所提取的 5 维最优特征子空间中的识别结果, 与 LDA 的 149 维最优特征子空间以及幅度谱原空间的识别结果进行比较。

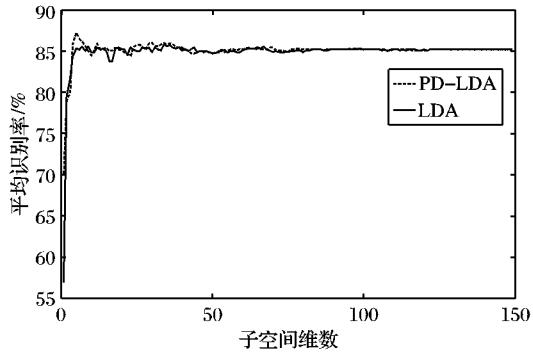


图 3 PD-LDA 与 LDA 不同维数子空间的平均识别率对比

表 2 为在 PD-LDA 的 5 维幅度谱子空间和传统 LDA 的 149 维幅度谱子空间, 采用最小欧氏距离分类器的识别结果的混淆矩阵。PD-LDA 相比 LDA 平均识别率高出 2%, 其中第 2 类目标安-26 的平均识别率提高了 4.25%。最小欧氏距离分类器是线性分类器, 说明 PD-LDA 可有效提取数据的线性可分性特征, 且 PD-LDA 在子空间维数远小于 $c - 1$ 时也可获得比 LDA 的 $c - 1$ 维子空间明显高的分类识别率, 这里类别数 $c = TF = 150$ 。采用最小欧氏距离分类器, PD-LDA 法识别过程的存储量为 $5 \times 150 + 256 \times 5$, LDA 法的存储量为 $149 \times 150 + 256 \times 149$ 。可见, 相比 LDA, PD-LDA 可大大降低数据维数并显著提高 RATR 的效率。

表 3 为在 PD-LDA 的 5 维幅度谱子空间和传统 LDA 的 149 维幅度谱子空间, 采用 SVM 多类分类器的识别结果的混淆矩阵, 前者比后者平均识别率提高了 2.75%。说明 PD-

LDA 所提取的特征,对于非线性分类器也是有效的。

表 4 为在幅度谱原空间,采用 SVM 分类器和最小欧氏距离分类器的识别结果的混叠矩阵。将表 2 和表 3 与表 4 相比可见:PD-LDA 子空间采用最小欧氏距离分类器和 SVM 分类器平均识别率比幅度谱原空间分别提高了 7.25% 和 6.42%,同时数据维数从 256 维降低到 5 维;LDA 子空间采用最小欧氏距离分类器和 SVM 分类器平均识别率比幅度谱原空间分别提高了 5.25% 和 3.67%,同时数据维数从 256 维降低到 149 维。说明 PD-LDA 在数据维数降低和识别率提高两方面均优于传统 LDA。

表 2 PD-LDA 与传统 LDA 采用最小距离分类器的识别结果

子空间	混叠矩阵			平均识别率/%
	雅克-42	安-26	奖状	
PD-LDA (5 维特征子空间)	98.75	1.25	0.00	
	18.75	71.25	10.00	87.25
	1.50	6.75	91.75	
LDA (149 维特征子空间)	96.75	3.00	0.25	
	16.25	67.00	16.75	85.25
	1.25	6.75	92.00	

表 3 PD-LDA 与传统 LDA 采用 SVM 分类器的识别结果

子空间	混叠矩阵			平均识别率/%
	雅克-42	安-26	奖状	
PD-LDA (5 维特征子空间)	97.75	2.25	0.00	
	10.75	81.75	7.50	91.8333
	1.50	2.50	96.00	
LDA (149 维特征子空间)	94.75	5.00	0.25	
	8.25	90.00	1.75	89.0833
	1.50	16.00	82.50	

表 4 256 维幅度谱原空间识别结果

分类器	混叠矩阵			平均识别率/%
	雅克-42	安-26	奖状	
SVM 多类分类器	92.25	7.5	0.25	
	5.25	75.5	19.25	85.4167
	0	11.5	88.50	
最小欧氏距离分类器	98.25	1.75	0	
	23.25	61.25	15.50	80.0000
	2.75	16.75	80.50	

5 结语

PD 准则从均衡保留类对间可分性的角度,为 LDA 有效提取数据的多类分类特征提供了思路。不同于 Fisher 准则,PD 准则用类对的子空间可分性在原空间可分性的占比,来描述各类对可分性的保留。由 PD 准则的等价形式可见,最大

化 PD 准则无需复杂的优化迭代算法。采用外场实测雷达数据的识别结果表明,相比传统 LDA,PD-LDA 可提取出更有效的多类分类特征,更好地对数据进行降维,并显著提高识别效率。今后,可通过对 PD 准则加权,使得在子空间中,权值大的类对可分性保留比例大,而权值小的类对可分性保留比例小,如何合理地确定这些权值是值得研究的问题。

参考文献:

- [1] FISHER R A. The statistical utilization of multiple measurements [J]. Annals of Human Genetics, 1938, 8(4): 376–386.
- [2] RAO C R. The utilization of multiple measurements in problems of biological classification [J]. Journal of the Royal Statistical Society: Series B (Methodological), 1948, 10(2): 159–203.
- [3] FUKUNAGA K. Introduction to statistical pattern recognition [M]. 2nd ed. Boston: Academic Press, 1990: 441–466.
- [4] LOOG M, DUIN R P W, HAEB-UMBACH R. Multiclass linear dimension reduction by weighted pairwise Fisher criteria [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001, 23(7): 762–766.
- [5] LIANG Y, LI C, GONG W, et al. Uncorrelated linear discriminant analysis based on weighted pairwise Fisher criterion [J]. Pattern Recognition, 2007, 40(12): 3606–3615.
- [6] LEE H S, CHEN B. Improved linear discriminant analysis considering empirical pairwise classification error rates [C]// Proceedings of the 6th International Symposium on Chinese Spoken Language Processing. Kunming: IEEE, 2008: 1–4.
- [7] LEE H S, CHEN B. Empirical error rate minimization based linear discriminant analysis [C]// Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing. Taipei: IEEE, 2009: 1801–1804.
- [8] 李金宗. 模式识别导论 [M]. 北京: 高等教育出版社, 1994: 392–397.
- [9] DU L, LIU H, BAO Z, et al. A two-distribution compounded statistical model for radar HRRP target recognition [J]. IEEE Transactions on Signal Processing, 2006, 54(6): 2226–2238.
- [10] DU L, WANG P, LIU H, et al. Bayesian spatiotemporal multitask learning for radar HRRP target recognition [J]. IEEE Transactions on Signal Processing, 2011, 59(7): 3182–3196.
- [11] DU L, LIU H, BAO Z, et al. Radar automatic target recognition using complex high-resolution range profiles [J]. IET Radar, Sonar & Navigation, 2007, 1(1): 18–26.
- [12] LIU J, ZHANG J, ZHAO F. Feature for distinguishing propeller-driven airplanes from turbine-driven airplanes [J]. IEEE Transactions on Aerospace and Electronic Systems, 2010, 46(1): 222–229.
- [13] 杜兰. 雷达高分辨率距离像目标识别方法研究 [D]. 西安: 西安电子科技大学, 2007.

(上接第 1024 页)

- [5] 方伟,孙俊,须文波.一种多样性控制的粒子群优化算法[J].控制与决策,2008,23(8):863–868.
- [6] 刘志雄,严新平,赵润军.置换流水车间调度粒子群算法与参数设置分析[J].武汉理工大学学报:交通科学与工程版,2010,34(6):1129–1132.
- [7] SUN KAI, YANG GEN-KE. An effective hybrid optimization algorithm for the flow shop scheduling problem[C]// Proceedings of the 2006 IEEE International Conference on Information Acquisition. [S. l.]: IEEE, 2006: 1234–1238.
- [8] 朱海梅,吴永萍.一种高速收敛粒子群优化算法[J].控制与决

- 策,2010,25(1):20–24.
- [9] WANG LING, ZHENG DA-ZHONG. A modified evolutionary programming for flow shop scheduling[J]. International Journal of Advanced Manufacturing Technology, 2003, 22: 522–527.
- [10] 乔佩利,马丽丽,郑林.基于改进粒子群算法的车间作业调度问题研究[J].哈尔滨理工大学学报,2011,16(2):35–39.
- [11] CARLIER J. Ordonnancements à contraintes disjonctives [J]. Operations Research, 1978, 12: 333–351.
- [12] REEVES C R. A genetic algorithms for flowshop sequencing[J]. Computers & Operation Research, 1995, 22(1): 5–13.