

基于非相关判别邻域保持投影的毫米波雷达目标识别

罗 磊 李跃华

(南京理工大学电子工程与光电技术学院 南京 210094)

摘 要: 该文基于流形学习的思想, 综合线性判别分析(LDA)及邻域保持投影(NPP)算法的优势, 提出了一种新的流形学习算法, 即非相关判别邻域保持投影(Uncorrelated Discriminant Neighborhood Preserving Projections, UDNPP)。该算法在保持了邻域几何结构的同时最大化了类间距离, 并通过引入一非相关约束条件使得提取的特征向量具有非相关性, 减少了冗余信息的干扰。毫米波雷达目标识别实验结果表明, 非相关判别邻域保持投影算法相对其它流行的学习算法有更好的性能。

关键词: 目标识别; 毫米波雷达; 特征提取; 流形学习; 类间距离; 非相关约束

中图分类号: TN957.51

文献标识码: A

文章编号: 1009-5896(2010)11-2751-04

DOI: 10.3724/SP.J.1146.2009.01534

Uncorrelated Discriminant Neighborhood Preserving Projections for Millimeter Wave Radar Target Recognition

Luo Lei Li Yue-hua

(School of Electronic Engineering and Optoelectronic Technology,
Nanjing University of Science and Technology, Nanjing 210094, China)

Abstract: A new algorithm named Uncorrelated Discriminant Neighborhood Preserving Projections (UDNPP) is proposed based on manifold learning. And UDNPP algorithm includes the advantages of Linear Discriminant Analysis (LDA) and Neighborhood Preserving Projections (NPP). Actually, UDNPP attempts to preserve the geometry of neighborhoods, while maximizing the between-class distance. Moreover, the features extracted are statistically uncorrelated by introducing an uncorrelated constraint. Thus the interference from redundant information are reduced. The experimental results from millimeter wave radar target recognition show that UDNPP algorithm can give competitive results in comparison with current popular algorithms.

Key words: Target recognition; Millimeter wave radar; Feature extraction; Manifold learning; Between-class distance; Uncorrelated constraint

1 引言

毫米波以其独特的优势已广泛应用于军事、航空、航天、民用等领域。在军事应用方面, 毫米波精确制导技术已发展到实用阶段, 毫米波高分辨率雷达已经应用于导弹、灵巧炸弹等武器当中。毫米波高分辨率雷达可得到目标的 1 维距离像, 包含了目标的形状及结构信息, 且不受雷达与目标之间的距离等因素的限制, 应用广泛。目标识别系统是毫米波高分辨率雷达的重要组成部分, 直接影响武器的精确度。流形学习算法以其独特优越性已成为近年机器学习领域的研究热点, 为毫米波高分辨率雷达目标识别提供了一种新的思路。

传统的特征提取方法中最著名的就是主分量分

析(PCA)和线性判别分析(LDA), 它们都是线性算法, 旨在保持全局的欧氏结构, 其中 LDA 还考虑了类间的判别信息。不同于它们, 流形学习算法旨在发现高维观测数据空间中隐藏的低维流形及低维流形的维数。典型的流形学习算法主要有等规度映射法(ISOMAP)^[1], 局部线性嵌入法(LLE)^[2], Laplacian 特征映射法(Laplacian eigenmaps)^[3]等。然而这些算法只能处理训练样本集, 不能处理新增样本。于是含有明确映射矩阵的流形学习算法被提了出来, 主要有局部保持映射(LPP)^[4], 邻域保持嵌入(NPE)^[5], 邻域保持投影(NPP)^[6]等。而且一些改进算法也被提了出来, 比如文献[7]提出了通用正交局部保持映射(GOLPP), 文献[8]提出了正交邻域保持投影(ONPP)。但是这些方法只是为了在嵌入空间更好地保持数据的拓扑结构, 而不具有更好的判别性能。随后一些有监督的方法被提了出来, 例如文献[9]提出了有监督核局部保持映射(SKLPP), 文献[10]提出

2009-12-01 收到, 2010-03-15 改回

国家部委基金资助课题

通信作者: 李跃华 hmb4507@mail.njust.edu.cn

了有监督核邻域保持投影(SKNPP)等。而且一些专门针对分类问题的算法也被提了出来,例如半监督子流形判别分析^[11]、邻域保持判别嵌入(NPDE)^[12]、局部保持判别映射(LPDP)^[13]、正交邻域保持判别分析(ONPDA)^[14]等。但是上述所有方法中的基向量都是相关的,这样就会造成特征抽取时含有冗余信息,降低识别效果。

实际上, LPP 和 NPP 分别为 Laplacian eigenmaps 和 LLE 的线性近似,不但保留了二者在保持数据局部几何结构上的优势,又解决了处理新样本的问题。但 LPP 的权重矩阵要么是简单的 0 和 1 构成,不能保持局部的几何结构,要么是基于高斯函数,则受输入参数影响较大。NPP 的权重矩阵和 LLE 相同,是经过优化的,很好地反映了近邻几何结构。本文在 NPP 的基础上提出了非相关判别邻域保持投影(Uncorrelated Discriminant Neighborhood Preserving Projections, UDNPP),致力于流形学习和分类问题。算法在保持类内几何结构的同时,最大化了类间距离,且通过引入一非相关约束使得提取的特征向量是非相关的。将算法应用到毫米波高分辨率雷达目标识别当中,实验结果证明了算法的优越性。

2 邻域保持投影

设采样得到的数据集 $\mathbf{X} = \{x_1, x_2, \dots, x_N\}$ 位于 D 维空间,特征提取或者说维数约简就是把这些数据点映射到维数是 $d(d \ll D)$ 的低维空间, \mathbf{X} 映射到 d 维空间后为 $\mathbf{Y} = \{y_1, y_2, \dots, y_N\}$ 。线性方法中, \mathbf{Y} 通过变换矩阵 \mathbf{A} 来得到,即 $y_i = \mathbf{A}^T x_i$ 。NPP 算法实际上就是为了寻找一个优化的变换矩阵 \mathbf{A} 使得变换后的数据能够保持邻域的结构,通过最小化代价函数

$$\begin{aligned} J_1(\mathbf{A}) &= \sum_{i=1}^N \left\| y_i - \sum_{j=1}^k y_j W_{ij} \right\|^2 = \|\mathbf{Y}(\mathbf{I} - \mathbf{W})\|^2 \\ &= \text{trace}\{\mathbf{Y}(\mathbf{I} - \mathbf{W})(\mathbf{I} - \mathbf{W})^T \mathbf{Y}^T\} \\ &= \text{trace}\{\mathbf{Y} \mathbf{M} \mathbf{Y}^T\} = \text{trace}\{(\mathbf{A}^T \mathbf{X}) \mathbf{M} (\mathbf{A}^T \mathbf{X})^T\} \\ &= \text{trace}\{\mathbf{A}^T \mathbf{X} \mathbf{M} \mathbf{X}^T \mathbf{A}\} \end{aligned} \quad (1)$$

来实现,且约束条件为

$$\frac{1}{N} \mathbf{Y} \mathbf{Y}^T = \frac{1}{N} \mathbf{A}^T \mathbf{X} \mathbf{X}^T \mathbf{A} = \mathbf{I} \quad (2)$$

式中 $\mathbf{M} = (\mathbf{I} - \mathbf{W})(\mathbf{I} - \mathbf{W})^T$, \mathbf{I} 是单位矩阵, \mathbf{W} 是权重矩阵,可以通过最小化下面代价函数来求得

$$\varepsilon(\mathbf{W}) = \sum_i \left| x_i - \sum_{j=1}^k W_{ij} x_j \right|^2 \quad (3)$$

实际上,这和 LLE 算法的第二步是一样的。

最终变换矩阵 $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_d]$ 的求解可以转化

为下面的广义特征值问题:

$$\mathbf{X} \mathbf{M} \mathbf{X}^T \mathbf{a} = \lambda \mathbf{X} \mathbf{X}^T \mathbf{a} \quad (4)$$

向量 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_d$ 可以根据它们对应的特征值 $0 < \lambda_1 < \lambda_2 < \dots < \lambda_d$ 重新排序得到。

实际上当约束条件为: $\mathbf{Y} \mathbf{Y}^T / (1 - N) = \mathbf{I}$ 时,算法为普通的邻域保持投影,当约束条件为: $\mathbf{A} \mathbf{A}^T = \mathbf{I}$ 时,算法则为正交邻域保持投影。

3 非相关判别邻域保持投影

根据文献[4]可知,数据的总体散布矩阵 \mathbf{S}_t 和类内散布矩阵 \mathbf{S}_w 分别为

$$\begin{aligned} \mathbf{S}_t &= (1/N) \sum_{i=1}^N (x_i - r)(x_i - r)^T \\ &= (1/N) \mathbf{X}(\mathbf{I} - \mathbf{e} \mathbf{e}^T / N) \mathbf{X}^T = \mathbf{X} \mathbf{G} \mathbf{X}^T \end{aligned} \quad (5)$$

$$\begin{aligned} \mathbf{S}_w &= \sum_{c=1}^C \sum_{x \in \omega_c} (x - q_c)(x - q_c)^T \\ &= \mathbf{X}(\mathbf{I} - \mathbf{E}) \mathbf{X}^T = \mathbf{X} \mathbf{D} \mathbf{X}^T \end{aligned} \quad (6)$$

其中 $r = (1/N) \sum_{i=1}^N x$, $q_i = (1/n_i) \sum_{x \in \omega_i} x$, $\mathbf{e} = (1, \dots, 1)^T$, $\mathbf{G} = \mathbf{I} - (1/N) \mathbf{e} \mathbf{e}^T$, $\mathbf{D} = \mathbf{I} - \mathbf{E}$ 。如果 x_i 和 x_j 同属于 ω_c , 则 $E_{ij} = 1/n_c$, 否则 $E_{ij} = 0$ 。

而 UDNPP 是针对于分类问题的,我们必须考虑其判别性能,这样就需要最大化类间距离

$$\begin{aligned} J_2(\mathbf{A}) &= \sum_{i,j=1}^C \| p_i - p_j \|^2 = \sum_{i,j=1}^C \left\| (1/n_i) \sum_{x \in \omega_i} \mathbf{A}^T x \right. \\ &\quad \left. - (1/n_j) \sum_{x \in \omega_j} \mathbf{A}^T x \right\|^2 = \sum_{i,j=1}^C \left\| \mathbf{A}^T (q_i - q_j) \right\|^2 \\ &= \text{trace} \left\{ \sum_{i,j=1}^C \mathbf{A}^T (q_i - q_j)(q_i - q_j)^T \mathbf{A} \right\} \\ &= \text{trace}\{\mathbf{A}^T \mathbf{S}_b \mathbf{A}\} \end{aligned} \quad (7)$$

其中 $p_i = (1/n_i) \sum_{y \in \omega_i} y$, \mathbf{S}_b 为类间散布矩阵,可以由下式计算,且 $\mathbf{L} = \mathbf{G} - \mathbf{D}$ 。

$$\mathbf{S}_b = \mathbf{S}_t - \mathbf{S}_w = \mathbf{X}(\mathbf{G} - \mathbf{D}) \mathbf{X}^T = \mathbf{X} \mathbf{L} \mathbf{X}^T \quad (8)$$

这样式(8)就可以写成

$$J_2(\mathbf{A}) = \text{trace}\{\mathbf{A}^T \mathbf{X} \mathbf{L} \mathbf{X}^T \mathbf{A}\} \quad (9)$$

那么就需要在最小化 J_1 的同时最大化 J_2 , 联合式(1)和式(9),最小化下面目标函数即可实现

$$J(\mathbf{A}) = \frac{\text{trace}\{\mathbf{A}^T \mathbf{X} \mathbf{M} \mathbf{X}^T \mathbf{A}\}}{\text{trace}\{\mathbf{A}^T \mathbf{X} \mathbf{L} \mathbf{X}^T \mathbf{A}\}} \quad (10)$$

接下来的问题就是引入非相关约束条件。如果低维空间中任意两个元素 y_i 和 y_j 是非相关的,根据文献[15],那么就会有

$$\mathbf{E}[(y_i - \mathbf{E}(y_i))(y_j - \mathbf{E}(y_j))] = \mathbf{a}_i^T \mathbf{S}_t \mathbf{a}_j = 0 \quad (11)$$

其中 \mathbf{a}_i 和 \mathbf{a}_j 为矩阵 \mathbf{A} 的不同的两列。再进行归一

化为

$$\mathbf{a}_i^T \mathbf{S}_i \mathbf{a}_i = 1 \tag{12}$$

综合式(12)和式(13)就得到

$$\mathbf{A}^T \mathbf{S}_i \mathbf{A} = \mathbf{I} \tag{13}$$

这样 UDNPP 就可以转化为下面的约束最小化问题：

$$\begin{aligned} \min_{\mathbf{A}^T \mathbf{S}_i \mathbf{A} = \mathbf{I}} \{J(\mathbf{A})\} &= \min_{\mathbf{A}^T \mathbf{S}_i \mathbf{A} = \mathbf{I}} \left\{ \frac{\text{trace}\{\mathbf{A}^T \mathbf{X} \mathbf{M} \mathbf{X}^T \mathbf{A}\}}{\text{trace}\{\mathbf{A}^T \mathbf{X} \mathbf{L} \mathbf{X}^T \mathbf{A}\}} \right\} \\ &= \min_{\mathbf{A}^T \mathbf{X} \mathbf{G} \mathbf{X}^T \mathbf{A} = \mathbf{I}} \left\{ \frac{\text{trace}\{\mathbf{A}^T \mathbf{X} \mathbf{M} \mathbf{X}^T \mathbf{A}\}}{\text{trace}\{\mathbf{I} - \mathbf{A}^T \mathbf{X} \mathbf{D} \mathbf{X}^T \mathbf{A}\}} \right\} \\ &= \min_{\mathbf{A}^T \mathbf{X} \mathbf{G} \mathbf{X}^T \mathbf{A} = \mathbf{I}} \{\text{trace}\{\mathbf{A}^T \mathbf{X} (\mathbf{M} + \mathbf{D}) \mathbf{X}^T \mathbf{A}\}\} \end{aligned} \tag{14}$$

这个约束最小化问题最终可以转化为广义特征值问题：

$$\mathbf{X}(\mathbf{M} + \mathbf{D})\mathbf{X}^T \mathbf{a}_i = \lambda \mathbf{X} \mathbf{G} \mathbf{X}^T \mathbf{a}_i, \quad i=1,2,\dots,d \tag{15}$$

设向量 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_d$ 为式(15)的解，且根据它们对应的特征值 $0 < \lambda_1 < \lambda_2 < \dots < \lambda_d$ 重新排序，那么变换矩阵即为 $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_d]$ ，显然，降维后的数据或者说提取的特征向量是非相关的。

由上述推导可得出非相关判别邻域保持投影算法的具体步骤，算法中的前两步与第 4 步和 NPP 相同，本文算法的改进主要在第 3 步。算法步骤如下：

步骤1 寻找与每个 x_i 的欧氏距离最小的 k 个点为邻域。

步骤2 通过最小化代价函数式(3)来计算权值矩阵 \mathbf{W} ，且权值 W_{ij} 满足约束条件 $\sum_j W_{ij} = 1$ ，且当 x_i, x_j 不是近邻点时 $W_{ij} = 0$ 。

步骤3 通过解决式(15)的广义特征值问题计算线性变换矩阵 $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_d]$ 。

步骤4 降维后的数据可以通过 $\mathbf{Y} = \mathbf{A}^T \mathbf{X}$ 计算。

UDNPP 算法是根据 NPP 算法改进的一种算法，兼有 NPP 算法在保持邻域几何结构上的优势，同时也有明晰的映射矩阵，可直接处理新样本。UDNPP 的改进包括两个方面，一是融合了 NPP 和 LDA 的算法思想，将 LDA 算法中的类间散布矩阵引入到 NPP 算法当中，且在计算类间散布矩阵时，利用了数据的类别信息，算法成了有监督方法，使算法更适应分类问题；二是在两种算法的基础上，提出了一种新的目标函数(见式(10))，在优化目标函数时能够保持两种算法的优势，且在进行目标函数优化时引入了非相关约束条件，使得提取的特征向量组具有线性无关性，减少特征向量个数的同时也减少了冗余信息的干扰，对降低后续算法的计算量和提高分类识别性能都有重要的促进作用。

4 仿真结果及分析

仿真数据来至坦克、装甲车、自行火炮、大卡车、小轿车 5 类目标。探测仰角为 0° ，方位角从 65.5° 到 85° ，每隔 0.5° 采样一次，采样点数为 256，这样每类目标就有 40 幅 1 维距离像。首先对实验数据进行归一化预处理。然后将数据平均随机分成 4 组，将其中的两组作为训练集，另外两组作为测试集，再重新分组，重复多次取平均值，其中分类算法采用最近邻分类算法。图 1 显示了不同特征提取方法所得到的平均识别率随约简后的数据维数的变化，其中低维空间维数从 10 到 90 变化，步进为 10。表 1 给出了不同特征提取方法下的最高识别率，及所对应的低维空间维数，其中 baseline 方法是指采用原始数据直接进行目标识别。

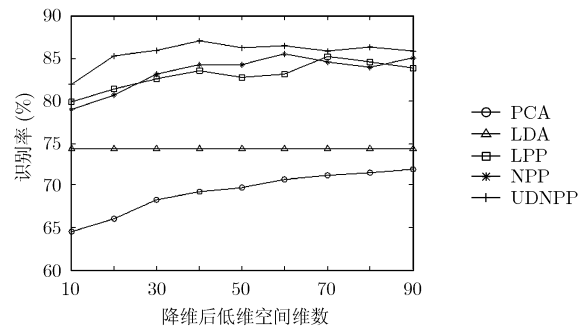


图 1 不同低维空间维数下各方法的识别率

由图 1 可知，PCA 随着低维空间维数的增加识别率不断提高，但实际上当低维空间的维数达到 112 之后，识别率不再增加，反而有下降的趋势；LDA 则由于算法本身的问题，在低维空间维数达到 5 以后就不再变化；LPP 和 NPP 算法十分接近，只是 NPP 在最高识别率时维度较低。UDNPP 在较低的维数下就达到了最高识别率，高于 NPP 和 LPP 算法，然后识别率小幅度波动，总体上好于其它算法。由式(8)可知，UDNPP 在最大化类间距离的时候，其实是利用了类别信息，属于有监督学习算法，考虑了算法的判别性能，故在进行分类时，识别效果好于其它算法，而且这也得益于 NPP 算法中优化的权值矩阵。从表 1 可知，就最高识别率而言 PCA 和 LDA 算法相对较差，LPP 和 NPP 十分接近，且都低于直接利用原始数据分类，LPP 和 NPP 的结果与原始数据差别极其细微，却降低了数据维数，大大减少了分类时计算复杂度。UDNPP 则相对较好，识别率相对采用原始数据而言，虽然仅有很小的提高，但在低维空间维数为 40 的时就达到了最大识别率，更大程度地降低了分类时的计算复杂度，比 LPP

表 1 不同方法得到的平均识别率

算法	PCA	LDA	LPP	NPP	UDNPP	Baseline
低维空间维数	90	10	70	60	40	256
识别率(%)	71.44	74.37	85.61	86.17	87.78	86.69

和 NPP 更具优势。

从 UDNPP 算法的推导及理论分析可以看出, UDNPP 是线性的有监督流形学习算法, 且更适应分类问题。而 UDNPP 算法融合了 LDA 和 NPP 算法的优点, 可应用于 LDA 及 NPP 适用的各种数据类型及机器学习问题, 例如, 图像处理、人脸识别、手写字体识别、指纹识别等领域。其它领域的应用不再讨论。

5 结论

本文提出了一种新的流形学习算法, 非相关判别邻域保持投影。算法试图保持近邻几何结构的同时, 又考虑了类间距离, 引入了判别信息, 属于有监督的线性算法; 并且通过引入非相关约束, 使得提取的特征是非相关的, 避免了信息冗余, 有利于提高识别效果, 降低分类时计算复杂度。毫米波高分辨率雷达目标识别仿真实验结果表明, 相对其它线性算法更具优势, 且算法在低维空间维数较低时仍有很好的效果。

参 考 文 献

- [1] Tenenbaum J B, Silva V, and Langford J C. A global geometric framework for nonlinear dimensionality reduction. *Science*, 2000, 290(5500): 2319-2323.
- [2] Roweis S T and Saul L K. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 2000, 290(5500): 2323-2326.
- [3] Belkin M and Niyogi P. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Computation*, 2003, 15(6): 1373-1396.
- [4] He X F, Yan S C, Hu Y X, Niyogi P, and Zhang H J. Face recognition using Laplacian faces. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 2005, 27(3): 328-340.
- [5] He X F, Cai D, Yan S C, and Zhang H J. Neighborhood preserving embedding. Proc. 10th IEEE. Conf. Computer Vision, Beijing, China, 2005, 2: 1208-1213.
- [6] Pang Y W, Zheng L, and Liu Z K, et al. Neighborhood preserving projections (NPP): a novel linear dimensionality reduction method. ICIC 2005, Part I, Lecture Notes in Computer Science, Springer, Berlin, 2005, Vol. 3644: 117-125.
- [7] Shao J D, Rong G, and Lee J M. Generalized orthogonal locality preserving projections for nonlinear fault detection and diagnosis. *Chemometrics and Intelligent Laboratory Systems*, 2009, 96(1): 75-83.
- [8] Kokiopoulou E and Saad Y. Orthogonal neighborhood preserving projections: A projection-based dimensionality reduction technique. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 2007, 29(12): 2143-2156.
- [9] Cheng J, Liu Q S, Lu H Q, and Chen Y W. Supervised kernel locality preserving projections for face recognition. *Neurocomputing*, 2006, 67(1-4): 443-449.
- [10] Yu X L, Wang X G, and Liu B Y. Supervised kernel neighborhood preserving projections for radar target recognition. *Signal Processing*, 2008, 88(9): 2335-2339.
- [11] Song Y Q, Nie F P, and Zhang C S. Semi-supervised sub-manifold discriminant analysis. *Pattern Recognition Letters*, 2008, 29(13): 1806-1813.
- [12] Han P Y, Jin A T B, and Abas F S. Neighbourhood preserving discriminant embedding in face recognition. *Journal of Vision Communication and Image Representation*, 2009, 20(8): 532-542.
- [13] Gui J, Wang C, and Zhu L. Locality preserving discriminant projections. ICIC 2009, Springer, Berlin, 2009, LNAI, Vol. 5755: 566-572.
- [14] Hu H F. Orthogonal neighborhood preserving discriminant analysis for face recognition. *Pattern Recognition*, 2008, 41(6): 2045-2054.
- [15] Jin Z, Yang J, Hu Z, and Lou Z. Face recognition based on the uncorrelated discriminant transformation. *Pattern Recognition*, 2001, 34(7): 1405-1416.

罗 磊: 男, 1982 年生, 博士生, 研究方向为信号处理、目标探测与识别。

李跃华: 男, 1959 年生, 教授, 博士生导师, 总装备部某专家组专家, 中国兵工学会引信专业委员会委员, 研究方向为探测与目标识别、信号处理与智能化技术。