

# 一种极化 SAR 影像分类中的半监督降维方法

谢欣芳, 徐新, 董浩, 吴晗, 李珞茹

武汉大学电子信息学院, 湖北 武汉 430072

**摘要** 针对极化合成孔径雷达(SAR)应用中存在的特征冗余问题,结合线性判别分析(LDA)和局部线性嵌入(LLE)的思想,提出一种半监督降维算法:半监督局部判别分析(SLDA)。该算法首先基于LLE的局部保持特性建立正则项,以避免学习中的过拟合问题。然后,在标记样本集上进行正则化的判别分析,以增强算法的推广能力,同时保持所有样本点在原始空间的局部几何结构。利用RADARSAT-2和AIRSAR卫星获得的Flevoland地区的全极化SAR数据进行降维实验,结果表明SLDA提取的低维特征具有“类内凝聚,类间分离”的特性;进一步的分类实验结果表明,SLDA只需1%~2%的标记样本就能使分类精度达到90%左右,分类性能优于其他对比方法。

**关键词** 遥感; 降维; 半监督局部判别分析; 极化合成孔径雷达; 分类

中图分类号 TP391.4

文献标识码 A

doi: 10.3788/AOS201838.0428001

## A Semi-Supervised Dimension Reduction Method for Polarimetric SAR Image Classification

Xie Xinfang, Xu Xin, Dong Hao, Wu Han, Li Luoru

School of Electronic Information, Wuhan University, Wuhan, Hubei 430072, China

**Abstract** Aiming at the problem of feature redundancy in polarimetric synthetic aperture radar (SAR) application, a semi-supervised dimension reduction algorithm: semi-supervised local discriminant analysis (SLDA) is proposed by combining the thoughts of linear discriminant analysis (LDA) and locally linear embedding (LLE). Firstly, the regularization term is established based on local preserving property of LLE to avoid overfitting problem during learning. Then, discriminant analysis with regularization is performed on labeled data set in order to improve the generalization ability and preserve the local geometric structure in original space for the whole data. Dimension reduction experiments are performed on all polarimetric SAR data from Flevoland regions obtained by RADARSAT-2 and AIRSAR satellites. The results show that the low dimensional features extracted by SLDA has the characteristics of “intra compactness and inter separation”. Further classification experiment results show that SLDA can make the classification accuracy reach about 90% only with 1%-2% labeled samples, and the classification performance of SLDA is superior to other comparison algorithms.

**Key words** remote sensing; dimension reduction; semi-supervised local discriminant analysis; polarimetric synthetic aperture radar; classification

**OCIS codes** 280.6730; 100.3008; 110.2960

## 1 引 言

极化合成孔径雷达(SAR)通过观测不同通道下的不同收发组合的极化回波信号,可以为地物分析提供更加丰富的信息,是雷达遥感领域的重要研究

方向,具有广阔的应用前景。通过变换和分解两种方法,可以从极化 SAR 中获取复杂多样的极化特征。文献[1-2]的研究表明全面利用这些极化特征可有效提高分类精度。但不同的极化特征之间存在信息冗余,比如 H/alpha 参数与 Freeman-Durden

收稿日期: 2017-10-19; 收到修改稿日期: 2017-11-16

基金项目: 高分辨率对地观测重大专项技术与开发项目(03-Y20A10-9001-15/16)

作者简介: 谢欣芳(1994—),女,硕士研究生,主要从事合成孔径雷达图像解译方面的研究。

E-mail: xiexinfang@whu.edu.cn

导师简介: 徐新(1967—),男,博士,教授,博士生导师,主要从事合成孔径雷达图像解译方面的研究。

E-mail: xinxu@whu.edu.cn(通信联系人)

分解特征都可以描述混沌体积特性。如果直接将提取的众多特征组合共同描述目标散射特性,则会使得极化 SAR 的分类面临“维度灾难”问题。因此降维或者特征选择在精确地物分类方面发挥重要作用,且日益引起广泛关注。

根据是否利用先验知识,降维方法可分为非监督、监督和半监督三种类型。在非监督方法中,主成分分析(PCA)最早被应用于极化 SAR 领域,但非监督方法不能利用标记样本的判别信息,因而取得的分类精度总体上难以满足极化 SAR 的实际应用需求。而监督方法在极化 SAR 领域的应用中虽然可取得较高的分类精度(如文献[3]提出的监督图嵌入算法以及文献[4]提出的线性判别拉普拉斯特征映射算法等),但监督方法所需的标记样本的获取比较困难而且代价昂贵,故较难将其在极化 SAR 领域推广开来。

因此,只需少量标记样本的半监督学习方法逐渐成为降维领域的主流。代表方法有半监督判别分析<sup>[5]</sup>和基于核方法的半监督判别分析<sup>[6]</sup>。这些方法能同时利用标记样本和未标记样本进行模型的学习和训练,既降低了人工标注样本的成本又能取得较好的结果,故在人脸识别和高光谱影像分类领域中被广泛应用<sup>[7-13]</sup>。近两年来,半监督学习因其强大的优势在极化 SAR 领域逐渐引起关注<sup>[11]</sup>,比如文献[7]提出的基于空间信息的半监督线性判别拉普拉斯特征映射方法和文献[8]提出的基于近邻约束的半监督特征提取方法。这两种方法在极化 SAR 影像的分类应用中都取得了不错的成果,在 RADARSAT-2 Flevoland 场景中两种方法的分类精度分别为 83.21%和 88.38%,在 AIRSAR Flevoland 场景 1 中两种方法的分类精度分别为 79.93%和 84.64%。

本文基于极化 SAR 领域的实际应用需求,考虑到线性判别分析(LDA)能够从原始特征中提取出具有判别力的新特征,但却因缺少标记样本而产生小样本问题而不能在极化 SAR 领域中被广泛应用。而局部线性嵌入(LLE)能够保持极化 SAR 数据的高维空间结构,但不能充分利用样本的判别信息,因此难以满足实际应用需求,且没有显式的映射将其特性泛化。故本文结合这两种方法的优势,提出一个具有局部保留特性的半监督判别分析算法:半监督局部判别分析(SLDA),并对提出的方法进行了理论证明和实验验证,结果表明所提算法提取的特征有利于提高极化 SAR 影像中的地物分类精度。

## 2 极化特征描述

一方面,由文献[9]可知,能够有效区分水域与其他地物类型的特征有: $|T_{22}|$ 、 $|T_{23}|$ 、 $|T_{33}|$ 、 $C_{1mHA}$ 、 $C_{SERD}$ ;能够有效区分耕地与其他地物类型的特征有: $\alpha$ 、 $C_{1mH1mA}$ 、 $\rho_{HH-HV}$  的相位;能够有效区分林地、建筑与其它地物类型的特征有: $A$ 、 $R_{\rho RR-LL}$ (圆极化相关系数  $\rho_{RR-LL}$  实部)、 $|\rho_{RR-LL}|$ (圆极化相关系数  $\rho_{RR-LL}$  幅值)、 $C_{RV1}$ 、 $C_{H1mA}$ 。另一方面,从散射机理的角度来看,水域主要表现为表面散射,建筑主要表现为偶次散射,林地主要表现为偶次散射和体散射,耕地主要表现为漫散射。

综上,为了提取与地物类别相关性更强的特征,提高特征集本身区分地物的能力,本文选取了常用的极化数据变换和极化目标分解特征: $T$ 矩阵幅值: $|T_{ij}|$ ( $0 < i \leq j \leq 3$ );Span 功率: $|P_{Span}|$ ;  $T$ 矩阵分解特征值及伪概率: $\lambda_i$ ( $i = 1, 2, 3$ )和  $P_1$ ;后向散射系数: $\sigma_{HH}^0$ 、 $\sigma_{HV}^0$ 和  $\sigma_{VV}^0$ ;极化相关系数: $\rho_{HH-VV}$ 、 $\rho_{HH-HV}$ 、 $\rho_{VV-HV}$ 和  $\rho_{RR-LL}$ ;Krogager 分解特征: $K_s$ 、 $K_d$ 和  $K_h$ ;Freeman 分解特征: $P_s$ 、 $P_d$ 和  $P_v$ ;扩展的四分量分解特征: $P_s$ 、 $P_d$ 、 $P_v$ 和  $P_f$ ;H/alpha/A 分解特征: $H$ 、 $\alpha$ 和  $A$ ;扩展 H/alpha/A 分解特征: $C_{1mHA}$ 、 $C_{H1mA}$ 、 $C_{1mH1mA}$ 、 $C_{SE1}$ 、 $C_{RV1}$ 、 $C_{SERD}$ 、 $C_{DERD}$ 、 $C_{PA}$ 和  $C_{PH}$ ;Huynen 分解特征: $A_0$ 、 $B_0 + B$ 、 $B_0 - B$ 、 $C$ 、 $D$ 、 $E$ 、 $F$ 、 $G$ 和  $H$ ;Van Zyl 分解特征: $A_1$ 、 $A_2$ 和  $A_3$ ,共 53 维。

由上述极化特征张成的极化观测空间中,由于不同维度上的特征对目标的描述角度既有不同也有相同,因此极化观测空间并非一个各维两两相互正交的特征空间,而是存在着大量的相关性,这就使得极化观测空间有了进行维数约减的必要。为了更好地揭示不同特征之间相互干扰的现象,选取来自 RADARSAT-2 卫星的 Flevoland 地区有代表性的极化特征三维一组进行显示,如图 1 所示。

由特征空间散点分布图可看出,在原始的极化观测空间中,不同类别的地物目标大多混杂在一起,即使不同类别目标点相对较为分离的观测空间,如  $H$ 、 $\bar{\alpha}$ 、 $A$  观测空间,也只是类间分离,同类目标点类内并不紧聚。若将这些观测空间合并成一个高维观测空间,则会因不同特征之间的相互影响使得整个特征空间中的数据点的分布结构更加混杂,难以区分。

## 3 算法描述

给定一景极化 SAR 图像  $X(W \times H = L_2)$  构建

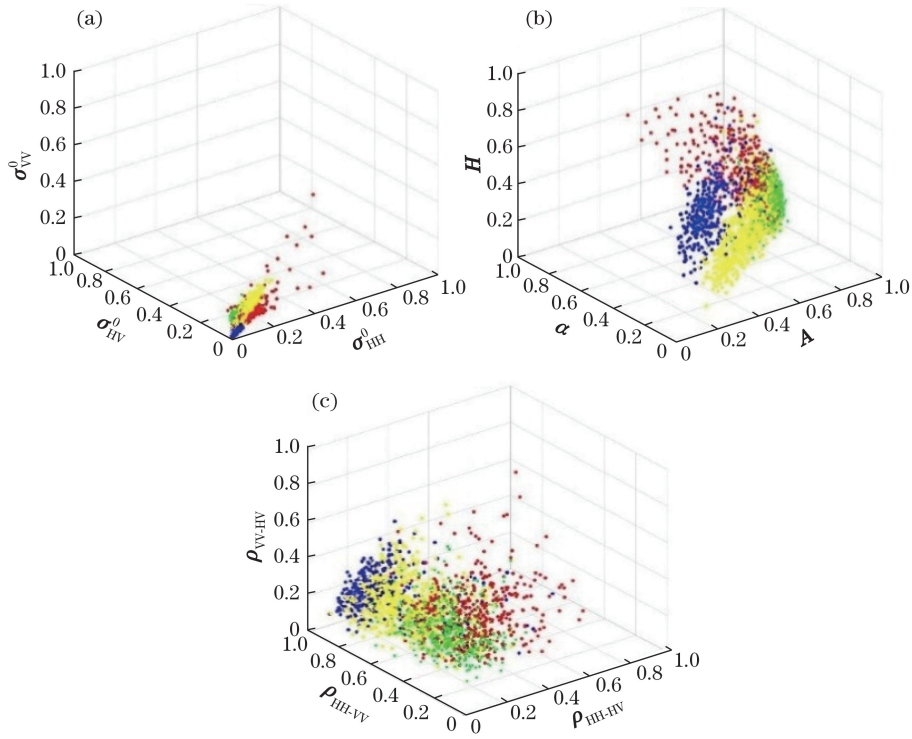


图 1 Flevoland 地区的极化特征空间散点分布图。(a)  $\sigma_{HH}^0, \sigma_{HV}^0, \sigma_{VV}^0$ ; (b)  $H, \alpha, A$ ; (c)  $\rho_{HH-VV}, \rho_{HH-HV}, \rho_{VV-HV}$

Fig. 1 Spatial scatter distribution of polarimetric features of Flevoland. (a)  $\sigma_{HH}^0, \sigma_{HV}^0, \sigma_{VV}^0$ ; (b)  $H, \alpha, A$ ; (c)  $\rho_{HH-VV}, \rho_{HH-HV}, \rho_{VV-HV}$

原始高维观测空间  $\mathbb{R}^D$  中的数据集  $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_{L_1}, \mathbf{x}_{L_1+1}, \dots, \mathbf{x}_{L_2}]$ , 其中,  $L_1$  表示有标记的样本数目,  $L_2$  表示样本总数目。数据集对应的类别标签集为  $\mathbf{C} = \{c_j, j=1, \dots, J\}$ , 类别总数为  $J$ 。记第  $j$  类数据样本构成的集合为  $c_j$ , 第  $j$  类数据样本的个数为  $n_j$ ,

且有  $\sum_{j=1}^J n_j = L_1$ 。定义  $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k]$ , 降维就是利用最优的  $\mathbf{U}$  将原始高维数据  $\mathbf{X}$  映射为低维数据  $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_k]$ , 即  $\mathbf{Y} = \mathbf{U}^T \mathbf{X}$ , 具体流程参考图 2。

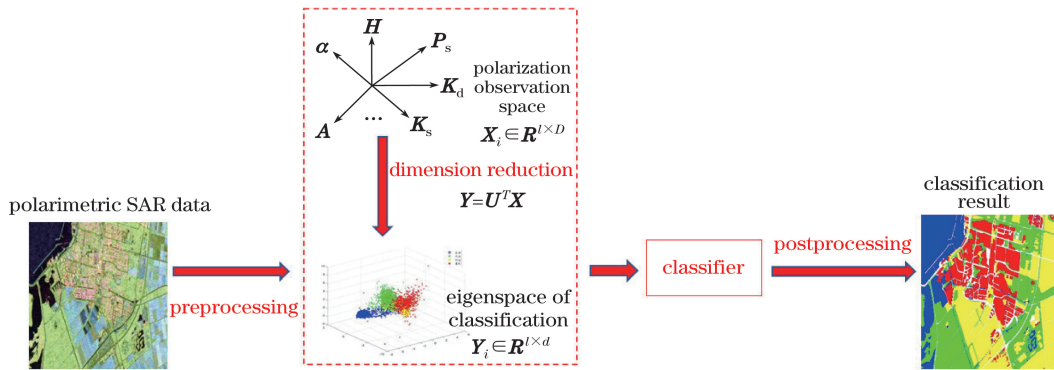


图 2 基于 SLDA 的极化 SAR 影像分类流程图

Fig. 2 Flowchart of polarimetric SAR image classification based on SLDA

### 1) 构建判别项

对于监督降维, 本研究希望映射空间能够保持原始高维样本的类间判别信息。即对于样本点  $\mathbf{x}_i$  和  $\mathbf{x}_j$ , 如果二者类别不同, 则在映射空间它们之间的距离应被最大化, 反之, 它们之间的距离应被最小化。

每一类数据样本的均值可表示为:

$$\mathbf{M}_j = \frac{1}{n_j} \sum_{\mathbf{x} \in c_j} \mathbf{x}. \quad (1)$$

所有样本的均值可表示为:

$$\mathbf{M} = \frac{1}{L_1} \sum_{\forall \mathbf{x}} \mathbf{x} = \frac{1}{L_1} \sum_{\mathbf{x} \in c_j} n_j \mathbf{M}_j. \quad (2)$$

每一类数据样本投影后的均值可表示为:

$$\bar{\mathbf{M}}_j = \frac{1}{n_j} \sum_{\mathbf{y} \in c_j} \mathbf{y} = \frac{1}{n_j} \sum_{\mathbf{x} \in c_j} \mathbf{u}^T \mathbf{x} = \mathbf{U}^T \mathbf{M}_j. \quad (3)$$

可以看出投影后的均值也就是样本中心点的投影。本研究的目标是投影后同类别的点之间相互靠近,



则构建类内判别项为：

$$\operatorname{argmin} \sum_{i,j=1}^{L_1} \|Y_i - Y_j\|^2 = \operatorname{argmin}(U^T S_w U), \quad (4)$$

式中  $S_w = \sum_{j=1}^J \sum_{x \in c_j} (x - M_j)(x - M_j)^T$  为类内散射矩阵。同理希望投影后不同类别的点尽可能远离，则构建类间判别项为：

$$\operatorname{argmax} \sum_{i,j=1}^{L_1} \|M_i - M_j\|^2 = \operatorname{argmax}(U^T S_B U), \quad (5)$$

式中  $S_B = \sum_{j=1}^J n_j (M_j - M)(M_j - M)^T$  为类间散射矩阵。

### 2) 构建保留项

受 LLE 算法<sup>[10]</sup> 的启发，利用大量无标记的样本构建保留项，保持数据流形的局部线性组合结构。本研究使用欧氏距离度量样本点间的距离。

在原始高维观测空间中，每个数据点  $x_i$  的  $k$  邻域为  $J_i$ ，以此构建邻接图  $G$ ，求取线性重构权系数  $w_{ij}$  满足以下关系式：

$$\begin{aligned} \operatorname{argmin}_Y \|Y - WY\|^2 &= \operatorname{argmin}_Y \sum_{i=1}^{L_w} \left\| y_i - \sum_{j \in J_i} w_{ij} y_j \right\|^2 = \\ \operatorname{argmin}_Y \operatorname{tr}[Y^T (I - W)^T (I - W) Y] &= \\ \operatorname{argmin}_Y \operatorname{tr}(Y^T M Y) &= \operatorname{argmin}(U^T M_p U), \end{aligned} \quad (6)$$

式中  $Y^T Y = I, M_p = X M X^T$ 。

### 3) 构建目标函数及求解

对于 SLDA，本研究的目标是映射后同类数据相互靠近，不同类数据相互分离，且能够保持高维空间数据的邻域结构，将保留项以正则项的形式加入到判别项中，得到的目标函数为：

$$\operatorname{argmax} \frac{\sum_{i,j=1}^{L_1} \|M_i - M_j\|^2}{\sum_{i,j=1}^{L_1} \|Y_i - Y_j\|^2 + \mu \sum_{i,j=1}^{L_2} \|Y - WY\|^2}. \quad (7)$$

式中  $\mu$  表示平衡因子，由(4)~(6)式可将(7)式化简为：

$$\operatorname{argmax} \frac{U^T S_B U}{U^T (S_w + \mu M_p) U}, \quad (8)$$

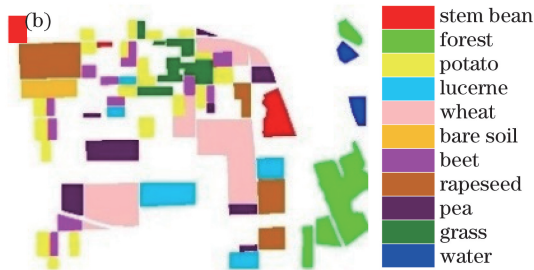


图 4 AIRSAR Flevoland 1 的(a)伪彩图；(b)标注图

Fig. 4 (a) Pseudocolor image; (b) labeled truth in classification of AIRSAR Flevoland 1

则目标函数的求解最终可转化为特征值求解问题：

$$S_B u = \lambda (S_w + \mu M_p) u, \quad (9)$$

得到前  $d$  个最大特征值对应的特征向量构成投影矩阵  $U^T = [u_1, \dots, u_d]$ ，将原始数据集  $X$  投影至由  $U^T$  张成的线性子空间上。

## 4 实验结果与分析

### 4.1 实验数据描述

第一个实验数据是 2008 年 RADARSAT-2 卫星获取的 Flevoland 地区(下称:RADARSAT-2 Flevoland)的 C 波段  $10 \text{ m} \times 5 \text{ m}$  单视全极化 SAR 数据，像素尺寸为  $1400 \text{ pixel} \times 1200 \text{ pixel}$ ，地物覆盖类型为水体、林地、耕地和建筑。其伪彩图和标注图如图 3 所示。

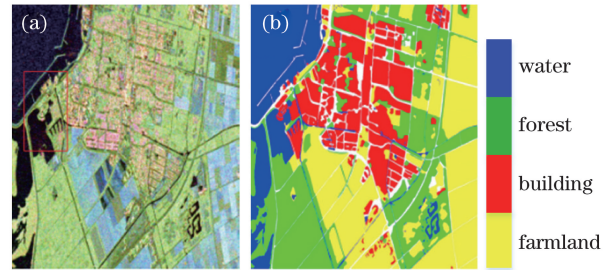


图 3 RADARSAT-2 Flevoland 的(a)伪彩图；(b)标注图

Fig. 3 (a) Pseudocolor image; (b) labeled truth in classification of RADARSAT-2 Flevoland

第二个实验数据是 1991 年 AIRSAR 卫星获取的 Flevoland 地区(下称:AIRSAR Flevoland 1)的 C 波段  $12 \text{ m} \times 6 \text{ m}$  四视全极化 SAR 数据，像素尺寸为  $750 \text{ pixel} \times 1024 \text{ pixel}$ ，地物覆盖类型为 11 种耕地类型。其伪彩图和标注图如图 4 所示。

第三个实验数据是 1991 年 AIRSAR 卫星获取的 Flevoland 地区(下称:AIRSAR Flevoland 2)的 L 波段全极化 SAR 数据，像素尺寸为  $1000 \text{ pixel} \times 1000 \text{ pixel}$ ，地物覆盖类型为 13 种耕地类型。其伪彩图和标注图如图 5 所示。

### 4.2 参数选取

需要选择的参数包括：训练样本数目、标记样本



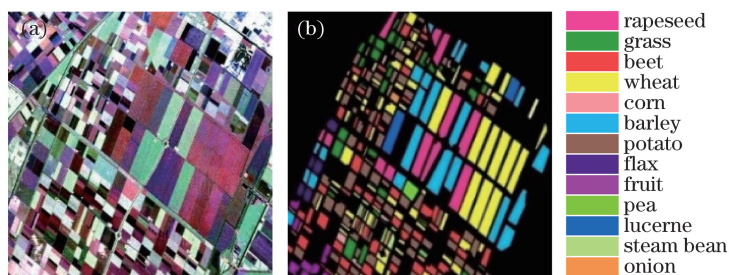


图 5 AIRSAR Flevoland 2 的(a)伪彩图;(b)标注图

Fig. 5 (a) Pseudocolor image; (b) labeled truth in classification of AIRSAR Flevoland 2

占总训练样本的比例、判别项与保留项之间平衡因子、降维维数以及参与降维的极化特征数目。图 6~图 9 分别给出不同参数选取的结果图,需要说明的

是 LDA 只能降到  $c-1$  维( $c$  为类别总数),为了方便比较将  $c-1$  维对应的结果直接平移到后面维度上, LPP 表示局部保留投影算法。

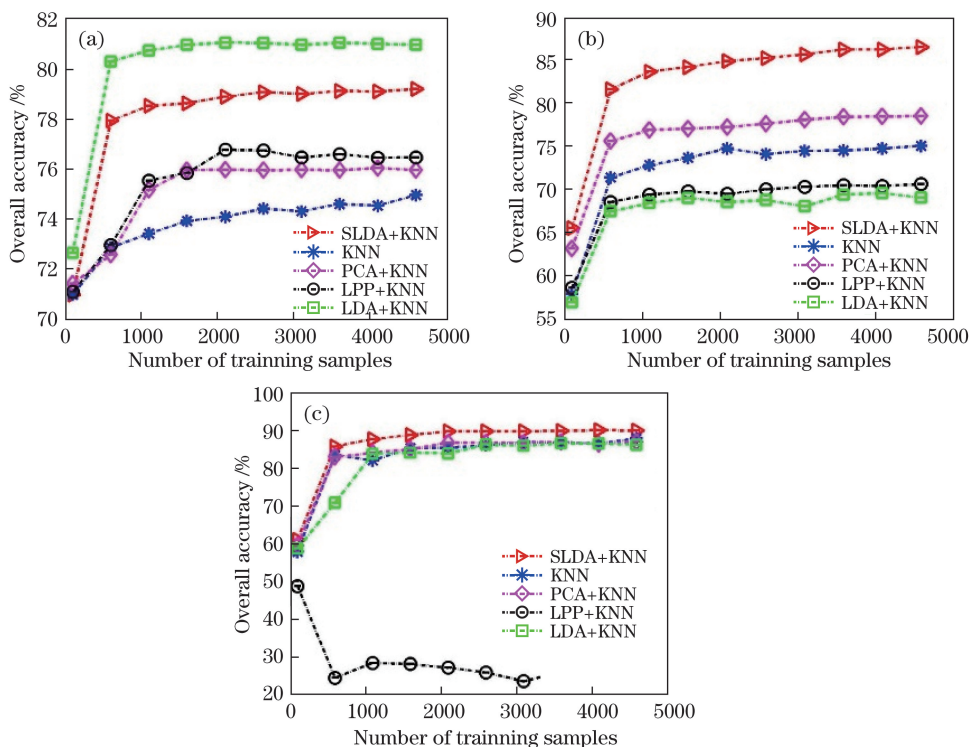


图 6 训练样本数目与分类精度关系图。(a) RADARSAT-2 Flevoland;  
(b) AIRSAR Flevoland 1; (c) AIRSAR Flevoland 2

Fig. 6 Relationship between training data number and classification accuracy. (a) RADARSAT-2 Flevoland;  
(b) AIRSAR Flevoland 1; (c) AIRSAR Flevoland 2

由图 6 可得,对于任何场景,当训练样本数目增加时,所有方法取得的分类精度都是先快速上升后趋于稳定,稳定时对应的训练样本数量在 2000 附近,故最优的训练样本数目为 2000,图中 KNN 为 k 近邻分类。

由图 7 可得,对于任何场景,随着维数的增加,分类精度都是先快速上升后趋于稳定,稳定时对应的维数可作为最优维数。RADARSAT-2 Flevoland 场景的最优维数为 4, AIRSAR Flevoland 场景的最优维数为 10。

由图 8 可得,对于任何场景,当标记比例上升时,分类精度先快速上升后趋于稳定,稳定时对应的样本比例在 10% 附近,故最优标记样本比例为 10%。10% 是相对于训练样本而言,在整个测试样本中所占比例为 1 占比 2%,可看出所提的算法具有很强的推广能力。

由图 9 可得,当平衡因子变化时,分类精度上下波动,相差甚微,则从分类精度、计算量和时间复杂度三方面折中考虑,平衡因子取 0.2。

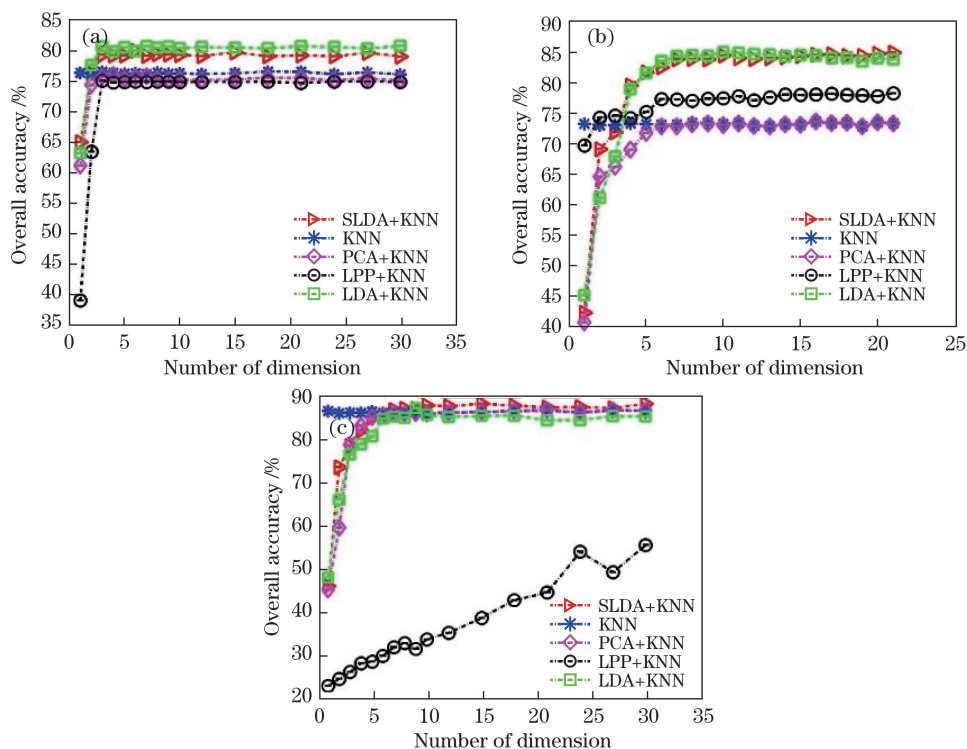


图 7 维度与分类精度关系图。(a) RADARSAT-2 Flevoland; (b) AIRSAR Flevoland 1; (c) AIRSAR Flevoland 2  
 Fig. 7 Relationship between dimension number and classification accuracy. (a) RADARSAT-2 Flevoland;  
 (b) AIRSAR Flevoland 1; (c) AIRSAR Flevoland 2

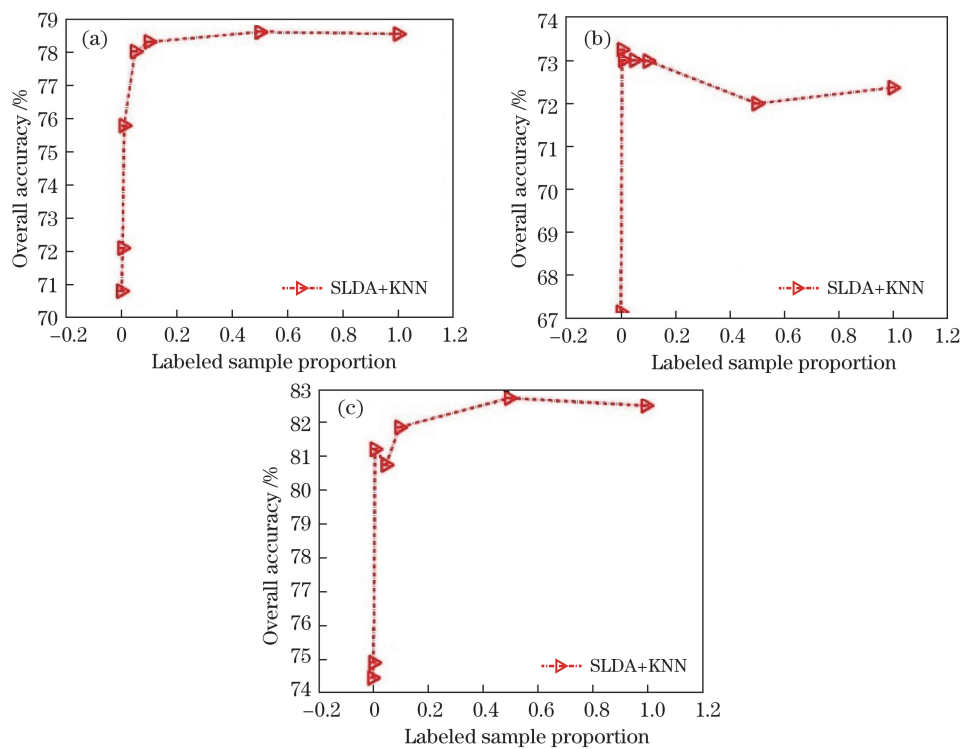


图 8 标记样本比例与分类精度关系图。(a) RADARSAT-2 Flevoland; (b) AIRSAR Flevoland 1;  
 (c) AIRSAR Flevoland 2

Fig. 8 Relationship between labeled sample proportion and classification accuracy. (a) RADARSAT-2 Flevoland;  
 (b) AIRSAR Flevoland 1; (c) AIRSAR Flevoland 2

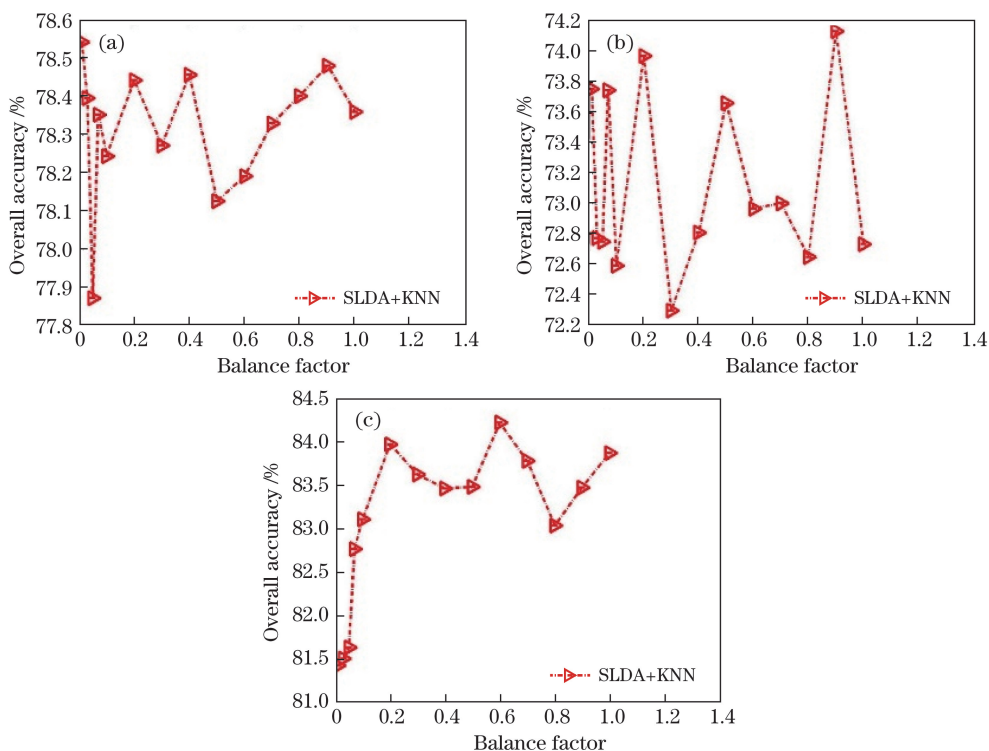


图 9 平衡因子与分类精度关系图。(a) RADARSAT-2 Flevoland; (b) AIRSAR Flevoland 1; (c) AIRSAR Flevoland 2  
 Fig. 9 Relationship between balance factor and classification accuracy. (a) RADARSAT-2 Flevoland;  
 (b) AIRSAR Flevoland 1; (c) AIRSAR Flevoland 2

在 RADARSAT-2 Flevoland 场景上使用上述最优参数,进行极化特征数目的选取,其中前 6 个特征是相干矩阵元素,之后依次叠加其他特征。进行降维和分类后的结果如图 10 所示。由图 10 可得,在相同参数条件下,随着参与降维的极化特征数目的增加,分类精度增大,这说明极化特征之间不是 100%冗余的,全面利用这些极化特征可有效提高分类精度。

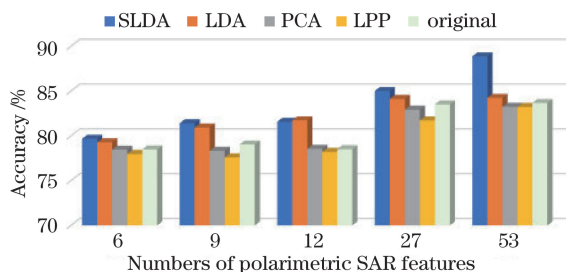


图 10 特征数目与分类精度关系图  
 Fig. 10 Relationship between feature number and classification accuracy

### 4.3 实验结果与分析

实验结果评价包括基于专业知识的定性评价和通过计算相应客观评价指标的定量评价。其中定性评价针对散点图和分类图,定量评价针对分类精度表。下面分别给出不同场景实验结果及分析。

#### 4.3.1 RADARSAT-2 Flevoland 实验结果与分析

对 RADARSAT-2 Flevoland 的分类结果如图 11 所示,其中图 11(f)是局部细节图,从左到右依次是:原始伪彩图、标注图、SLDA + KNN、KNN、PCA + KNN、LPP + KNN、LDA + KNN。由图 11 可得,SLDA 与 KNN 结合的分类效果最好,即便部分区域地物类型复杂,SLDA 也能取得比其他方法好的效果,这要归因于 SLDA 同时利用标记和未标记样本,既能保持类间和类内的判别信息,又能保持高维数据间的局部几何信息,降维后得到的特征可分性较高。

不同地物类型的分类精度如表 1 所示(OA 表示总体精度),由表 1 可得出以下结论:

- 1) 水体:因水体的散射机理较为简单,易于区分,故水体整体的分类精度很高,降维对其分类精度的提升不明显;
- 2) 林地:降维可提高其分类精度,与不降维相比 SLDA 降维可提高 5%左右,提升程度高于其他方法;
- 3) 耕地:降维可提高其分类精度,与不降维相比 SLDA 降维可提高 3%左右,提升程度高于其他方法,包括文献[7]和[8]所提出的算法(分类精度分别为 73.76%和 87.12%),这说明 SLDA 更适合用于耕地类地物的分类;



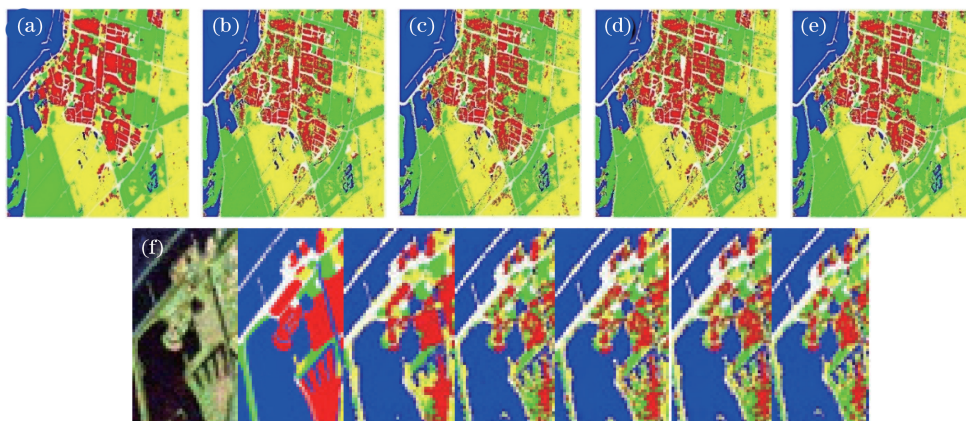


图 11 RADARSAT-2 Flevoland 分类图。(a) SLDA+KNN; (b) KNN; (c) PCA+KNN; (d) LPP+KNN; (e) LDA+KNN; (f)局部图

Fig. 11 Classification maps of RADARSAT-2 Flevoland. (a) SLDA+KNN; (b) KNN; (c) PCA+KNN; (d) LPP+KNN; (e) LDA+KNN; (f) local detail maps

表 1 RADARSAT-2 Flevoland 分类精度表

Table 1 Classification accuracy of RADARSAT-2 Flevoland %

Category	SLDA	PCA	LPP	KNN	LDA
Water	93.07	93.77	93.57	93.81	94.18
Farmland	91.94	88.15	86.50	88.94	91.13
Forest	88.39	81.59	81.68	83.01	81.05
Building	81.02	68.65	71.33	67.42	69.10
OA	88.88	83.24	83.23	83.67	84.25

4) 建筑:降维可显著提高其分类精度,与不降维相比 SLDA 降维可提高 14%左右,提升程度明显

高于其他方法;

总之,对于 RADARSAT-2 Flevoland,降维是有效的,与不降维相比选择合适的降维方法可以提高分类精度。SLDA 取得的总体分类精度是以上方法里面最高的,同时高于文献[7]和[8]所提出的方法。

#### 4.3.2 AIRSAR Flevoland 1 实验结果与分析

AIRSAR Flevoland 1 的分类结果如图 12 所示。由图 12 可发现,SLDA 与 KNN 组合的分类效果最好,即便部分区域地物类型较复杂,SLDA 也能取得较好的分类效果,这与上一场景的结论一致。



图 12 AIRSAR Flevoland 1 分类图。(a) SLDA+KNN; (b) KNN; (c) PCA+KNN; (d) LPP+KNN; (e) LDA+KNN

Fig. 12 Classification maps of AIRSAR Flevoland 1. (a) SLDA+KNN; (b) KNN; (c) PCA+KNN; (d) LPP+KNN; (e) LDA+KNN

不同地物类型的分类精度如表 2 所示,由表 2 可得出以下结论:

- 1) SLDA 的总体分类精度是以上方法里面最高的,这与上一场景的实验结论基本一致;
- 2) SLDA 在林地、土豆、油菜籽、小麦、甜菜和草地的分类任务中取得的分类精度高于其他方法。

文献[7]所提算法对以上 6 种地物的分类精度分别为 87.32%、90.05%、64.26%、90.00%、64.38%和 46.34%,文献[8]所提算法对以上 6 种地物的分类精度分别为 90.44%、85.53%、69.19%、89.33%、96.66%和 50.54%;

- 3) 对于物理特性较为复杂且与其他地物物理

表 2 AIRSAR Flevoland 1 分类精度表

Table 2 Classification accuracy of AIRSAR Flevoland 1 %

Category	SLDA	PCA	LPP	KNN	LDA
Steam bean	96.88	94.99	9.70	94.30	98.37
Forest	98.59	95.03	47.01	95.08	95.66
Potato	89.77	84.90	23.03	85.23	88.04
Lucerne	92.74	90.70	7.34	90.82	92.99
Wheat	95.02	88.48	92.31	88.97	93.54
Bare soil	89.26	92.07	5.95	92.08	93.61
Beet	88.08	81.51	1.12	82.13	80.51
Rapeseed	92.79	92.07	14.37	92.28	85.13
Pea	66.91	72.32	1.86	73.13	76.83
Grass	85.46	73.71	1.42	73.71	75.36
Water	98.55	93.87	48.20	95.52	99.46
OA	91.01	87.44	34.70	87.74	89.11

特性相似度较多的草地和干豆,SLDA 仍可取得比其他方法好的分类效果。

### 4.3.3 AIRSAR Flevoland 2 实验结果与分析

使用不同的降维方法对 AIRSAR Flevoland 2 选取的 2000 个训练样本进行降维,得到的三维散点分布图如图 13 所示。

由图 13 可以发现,与 LDA 和 LPP 相比,SLDA 降维后样本点出现明显的“类内紧聚,类间分离”现象,这要归因于 SLDA 构建的类间和类内判别项;与 PCA 相比,SLDA 的类内聚集度更大,这要归因于 SLDA 构建的局部保留项。

AIRSAR Flevoland 2 的分类结果如图 14 所示,由图 14 可得出与前两个场景一致的结论。

各地物类型的分类精度如表 3 所示,由表 3 可得出以下结论:

1) 该场景中的小麦、油菜籽、甜菜、苜蓿和土豆的分类精度和上一场景一致都高于其他方法,但同时豌豆的分类精度也都较低,这可能是因为豌豆的物理特性更加复杂,存在较多的杂散波干扰。

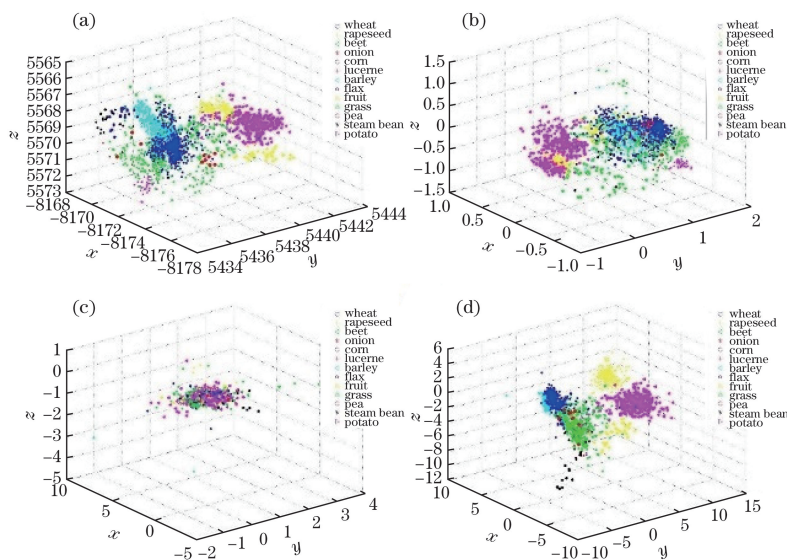


图 13 AIRSAR Flevoland 2 的三维散点分布图。(a) SLDA; (b) PCA; (c) LPP; (d) LDA

Fig. 13 3D scatter distribution diagram of AIRSAR Flevoland 2. (a) SLDA; (b) PCA; (c) LPP; (d) LDA

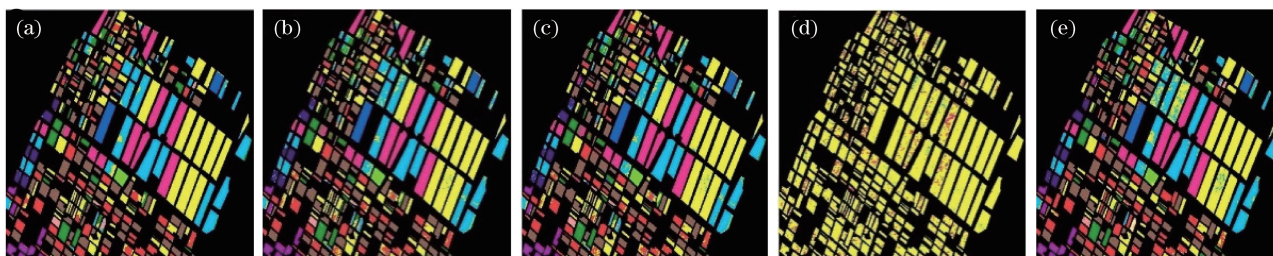


图 14 AIRSAR Flevoland 2 分类图。(a) SLDA+KNN; (b) KNN; (c) PCA+KNN; (d) LPP+KNN; (e) LDA+KNN

Fig. 14 Classification maps of AIRSAR Flevoland 2. (a) SLDA+KNN; (b) KNN;

(c) PCA+KNN; (d) LPP+KNN; (e) LDA+KNN

表 3 AIRSAR Flevoland 2 分类精度表

Table 3 Classification accuracy of  
AIRSAR Flevoland 2 %

Category	SLDA	PCA	LPP	KNN	LDA
Wheat	97.93	97.26	95.50	97.35	97.44
Rapeseed	98.81	98.44	52.58	98.51	99.52
Beet	87.74	83.38	19.35	83.68	73.61
Onion	53.64	50.79	0.93	51.29	33.75
Corn	47.38	44.12	4.78	44.06	43.76
Lucerne	97.74	96.57	5.54	96.60	91.34
Barley	92.57	89.47	36.24	89.44	73.62
Flax	97.32	96.41	1.78	96.51	92.26
Fruit	96.43	92.20	10.45	92.17	93.03
Grass	70.16	65.21	15.22	65.30	53.29
Pea	68.68	73.10	0	73.01	53.74
Steam bean	41.04	39.92	14.50	39.56	20.75
Potato	97.94	97.92	36.35	98.03	97.59
OA	92.53	90.81	47.84	90.89	85.59

2) 该场景新增了洋葱和大麦这两类地物,洋葱与干豆的物理特性接近,导致在上一场景中分类精度较高的干豆在这一场景中分类精度较低;大麦与草地的物理特性相似程度较高,导致草地的分类精度与上一场景相比出现一定程度的下降;虽然分类精度有所下降但 SLDA 仍可取得比其他方法好的分类效果。

3) SLDA 取得的总体分类精度是以上方法里面最高的,这与前两个场景的结论一致。

综合三个场景的实验结果,可得出结论: SLDA 降维得到的低维特征有利于提高极化 SAR 影像的地物分类精度。

## 5 结 论

结合极化 SAR 的实际应用需求,提出一种半监督降维方法: SLDA,并搭建了基于 SLDA 的极化 SAR 影像地物分类框架。结合近年来国内外先进的学术成果,选取了 53 维与地物类别相关性较强的极化特征子集,并调节 SLDA 的参数使其更加适用于极化特征的降维,同时发现具有普适性的参数选择规律。在真实极化 SAR 影像的特征降维以及多地物分类任务中,SLDA 的分类性能优于其他对比方法。

但本研究中直接将极化特征拉扯成二维向量可能会损失部分空间信息,这些信息可能会有利于提高特征的可分性,故在未来的研究中,如何更好地应

用空间信息应该引起重视。

## 参 考 文 献

- [1] Uhlmann S, Kiranyaz S. Integrating color features in polarimetric SAR image classification[J]. IEEE Transactions on Geoscience & Remote Sensing, 2014, 52(4): 2197-2216.
- [2] Zou T, Yang W, Dai D, *et al.* Polarimetric SAR image classification using multifeatures combination and extremely randomized clustering forests[J]. EURASIP Journal on Advances in Signal Processing, 2009, 2010: 465612.
- [3] Shi L, Zhang L, Yang J, *et al.* Supervised graph embedding for polarimetric SAR image classification[J]. IEEE Geoscience & Remote Sensing Letters, 2013, 10(2): 216-220.
- [4] Shi L, Zhang L, Zhao L, *et al.* The potential of linear discriminative Laplacian eigenmaps dimensionality reduction in polarimetric SAR classification for agricultural areas[J]. ISPRS Journal of Photogrammetry & Remote Sensing, 2013, 86(12): 124-135.
- [5] Cai D, He X, Han J. Semi-supervised discriminant analysis[C]. IEEE International Conference on Computer Vision, 2007: 1-7.
- [6] Zhang P Q, Tan X, Xu X C, *et al.* Hyperspectral imagery feature extraction based on kernel semi-supervised discriminant analysis[J]. Journal of Geomatics Science and Technology, 2016, 33(3): 258-262.  
张鹏强, 谭熊, 余旭初, 等. 基于核半监督判别分析的高光谱影像特征提取[J]. 测绘科学技术学报, 2016, 33(3): 258-262.
- [7] Sun X, Huang P P, Tu S T, *et al.* Polarimetric SAR image classification using multiple-feature fusion and ensemble learning[J]. Journal of Radars, 2016, 5(6): 692-700.  
孙勋, 黄平平, 涂尚坦, 等. 利用多特征融合和集成学习的极化 SAR 图像分类[J]. 雷达学报, 2016, 5(6): 692-700.
- [8] Zhu D X. Semi-supervised dimensionality reduction methods for polarimetric SAR classification[D]. Xi'an: Xidian University, 2015: 20-50.  
朱德祥. 极化 SAR 半监督降维方法[D]. 西安: 西安电子科技大学, 2015: 20-50.
- [9] Song C, Xu X, Gui R, *et al.* Polarimetric synthetic aperture radar feature analysis and classification based on multi-layer support vector machine classifier[J]. Journal of Computer Application, 2017, 37(1): 224-250.  
宋超, 徐新, 桂容, 等. 基于多层支持向量机的极化合成孔径雷达特征分析与分类[J]. 计算机应用, 2017, 37(1): 244-250.



- [10] Roweis S T, Saul L K. Nonlinear dimensionality reduction by locally linear embedding[J]. Science, 2000, 290(5500): 2323-2326.
- [11] Liu H, Zhu D, Yang S, *et al.* Semisupervised feature extraction with neighborhood constraints for polarimetric SAR classification[J]. IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing, 2016, 9(7): 3001-3015.
- [12] Ye Z, Bai L, Nian Y J. Hyperspectral image classification algorithm based on Gabor feature and locality-preserving dimensionality reduction[J]. Acta Optica Sinica, 2016, 36(10): 1028003.  
叶珍, 白麟, 粘永健. 基于 Gabor 特征与局部保护降维的高光谱图像分类算法 [J]. 光学学报, 2016, 36(10): 1028003.
- [13] Zhang J J, Zhou X Y, Liu Q. Improved dimensionality reduction algorithm of large-scale hyperspectral scenes using manifold[J]. Acta Optica Sinica, 2013, 33(11): 1128001.  
张晶晶, 周晓勇, 刘奇. 一种改进的大尺度高光谱流形降维算法 [J]. 光学学报, 2013, 33(11): 1128001.
- [14] Wang Y C, Guo J B, Zhou L Y. Image Hash algorithm based on data dimension reduction and symmetric binary pattern[J]. Laser & Optoelectronics Progress, 2017, 54(2): 021004.  
王彦超, 郭静博, 周丽宴. 基于数据降维与对称二值模式的图像 Hash 算法 [J]. 激光与光电子学进展, 2017, 54(2): 021004.