

基于空间约束多特征联合稀疏编码的遥感图像标注方法研究

郑歆慰^{*①②③} 胡岩峰^① 孙 显^{①②} 王宏琦^{①②}

^①(中国科学院电子学研究所 北京 100190)

^②(中国科学院空间信息处理与应用系统技术重点实验室 北京 100190)

^③(中国科学院大学 北京 100190)

摘 要:针对稀疏表示分类器不能较好地适应多特征框架的问题,本文提出一种空间约束多特征联合稀疏编码模型,并以此实现遥感影像的自动标注。该方法利用 $l_{1,2}$ 混合范数正则化多特征编码系数,约束编码系数共享相同的稀疏模式,在保持多特征关联的同时,又不添加过于严格的约束。同时,将字典学习技术扩展到多特征框架中,通过约束字典更新的变换矩阵,解决了字典学习过程丢失多特征关联的问题。另外,针对遥感影像中的空间关系常常被忽略或者利用不充分的不足,还提出了将空间一致性与多特征联合稀疏编码相结合的分类准则,提高了标注性能。在遥感公开数据集与大尺寸卫星影像上的实验证明了该方法的有效性。

关键词: 遥感图像标注; 多特征联合稀疏编码; 多特征字典学习; 空间信息

中图分类号: TP751

文献标识码: A

文章编号: 1009-5896(2014)08-1891-08

DOI: 10.3724/SP.J.1146.2013.01433

Annotation of Remote Sensing Images Using Spatial Constrained Multi-feature Joint Sparse Coding

Zheng Xin-wei^{①②③} Hu Yan-feng^① Sun Xian^{①②} Wang Hong-qi^{①②}

^①(Institute of Electronic, Chinese Academy of Sciences, Beijing 100190, China)

^②(Key Laboratory of Spatial Information Processing and Application System Technology, Chinese Academy of Sciences, Beijing 100190, China)

^③(University of Chinese Academy of Sciences, Beijing 100190, China)

Abstract: In this paper, a novel framework for remote sensing image annotation is proposed based on spatial constrained multi-feature joint sparse coding to extend the sparse representation-based classifier to multi-feature framework. The proposed framework imposed an $l_{1,2}$ mixed-norm regularization on encode coefficients of multiple features. The regularization encourages the coefficients to share a common sparsity pattern, which preserves the cross-feature information. Inspired by the success of dictionary learning, a novel dictionary learning model is proposed to promote the performance of multi-feature joint sparse coding, while the cross-feature association is preserved by consistent transformation constraint. In addition, spatial dependencies between patches of remote sensing images are useful for annotation task but usually ignored of insufficiently exploited. In this paper, a spatial relation constrained classifier is designed to incorporate spatial coherence into multi-feature sparse coding model to annotate images more precisely. Experiments on public dataset and large satellite images show the discriminative power and effectiveness of the proposed framework.

Key words: Remote sensing image annotation; Multi-feature joint sparse coding; Multi-feature dictionary learning; Spatial information

1 引言

随着遥感技术的发展,影像分辨率越来越高,遥感数据的获取数量也越来越多,面对如此海量的数据,人们急需建立一个智能化的遥感数据管理系

统来实现对数据的快速访问。其中,遥感影像的自动标注就是一个非常重要的组成部分。

自动标注就是根据图像的内容为其打上相应的标签。长久以来,这都是遥感领域的研究热点之一^[1],研究人员提出了许多方法,例如:基于 SVM 的遥感图像分类方法^[2,3],基于 LDA 的语义标注方法^[4],基于神经网络的图像分类方法^[5],等等。近年来,由于压缩感知技术的发展^[6]以及稀疏表示在计算机视

2013-09-18 收到, 2014-03-05 改回

高分对地观测领域学术交流项目(GFZX04060103)资助课题

*通信作者: 郑歆慰 zzw_1020@163.com

觉领域的成功应用^[7-9]，基于稀疏表示的分类方法(Sparse Representation-based Classification, SRC)得到了广泛的关注^[10-14]，在许多遥感应用中取得了不俗的效果^[15-17]。一方面，SRC不需要为每个类型学习复杂的模型，因而更适用于内容复杂、类型繁多的遥感影像；另一方面，由于遥感图像的特点，单一特征表示图像的能力较弱，而SRC却不能很好地使用多种特征。鉴于其实现原理，SRC要么分别使用单一特征，综合后得到分类结果，要么将多种特征串接为一个长特征。前一种方法在稀疏表示时忽略了多种特征之间的关联，削弱了多特征的代表能力；而后一种方法则相当于要求不同特征具有相同的表示系数，由于特征提取的非线性性，这样的约束就显得过于严格。

另一方面，图像的空间上下文信息也是标注过程中不可忽视的信息。一般而言，在一幅大尺寸遥感影像中，相邻的图像块会有较大的概率是属于同一类型。如何将空间信息应用到标注过程中也是研究人员关注的问题之一。Bruzzone等人^[3]提出了一种基于SVM分类器的方法，在训练过程中加入空间信息以获得更好的标注精度。Liéno等人^[4]则使用了重叠划分技术并对块与块之间的重叠部分采用多数投票策略，以此实现更精细的标注。但是，在这过程中，空间信息并没有得到充分的利用，因为不同块之间并不存在空间关系约束。

针对以上不足，本文提出了一种基于空间约束多特征联合稀疏编码的遥感图像自动标注方法。方法首先将SRC模型扩展为多特征联合稀疏编码(Multi-Feature Joint Sparse Coding, MFJSC)模型，通过对编码系数添加 $l_{1,2}$ 混合范数约束^[18,19]要求不同特征共享相同的稀疏模式，即不同特征在稀疏编码时，从字典中选择相同的元素编码，而编码系数则不尽相同，以此实现多特征关联的合理利用。随着稀疏模型的发展，字典学习技术都被广泛用于扩展SRC模型^[11-13,20,21]，不仅提高了编码的效率，而且通过加入判别信息提高了分类准确度。受此启发，本文将字典学习技术扩展到多特征框架中，并且对不同特征对应的字典元素的更新矩阵进行约束，保持了多种特征之间的关联。此外，基于MFJSC的输出，本文方法设计了新的分类准则实现空间关系约束分类器(Spatial Relationship Constraint Classifier, SRCC)，将图像块之间的空间依赖添加到模型中，影像的所有图像块的标签将通过最小化一个目标函数而同时获得。由于空间信息的融入，本文方法提高了遥感影像的标注精度。

本文第2节详细介绍了研究方法的各个环节；

第3节给出了实验结果及分析；第4节则对全文进行总结。

2 研究方法

图1给出了本文方法的主要流程。其中，训练过程包括多特征提取、多特征字典学习，测试过程则包括多特征提取、多特征联合稀疏编码以及空间关系约束分类器。

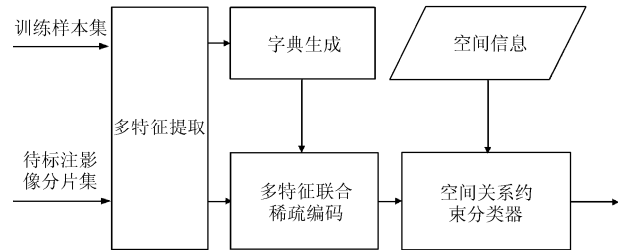


图1 自动标注方法流程框图

2.1 多特征提取

遥感图像具有内容丰富、目标背景复杂等特点，单一特征往往不能很好地表示图像，因此本文方法采用多种特征同时表示图像。记 y^k 为从图像块中提取的第 k 种特征，其中 $k = \{1, 2, \dots, K\}$ ， K 为特征类型数(在本文的实验中， $K = 4$)。特征的选取和参数设置是实验设置的重要组成部分，因此关于多特征的详细讨论将在3.1节中进行。

2.2 多特征联合稀疏编码

记 Y 为图像块提取的 K 种特征所排列成的矩阵， $Y = [y^1 y^2 \dots y^K]$ 。假设每种特征都有一个相应的字典 D^k (在这里先假设 D^k 为所有训练样本第 k 种特征按列排布组成的矩阵，更有效的字典学习算法将在2.3中讨论)。如果字典的容量够大，那么 y^k 可以用 D^k 中少量元素的线性组合很好地近似。记 w^k 为 y^k 的编码系数，其中 $w^k \in R^N$ ， N 为字典容量。记 $W = [w^1 w^2 \dots w^K]$ ， w_i 为 W 的第 i 行。定义 W 的 $l_{1,2}$ 混合范数为 W 逐行求 l_2 范数后求 l_1 范数。

$$\|W\|_{1,2} = \sum_{i=1}^N \sqrt{\sum_{k=1}^K w_{ik}^2} \quad (1)$$

压缩感知理论^[6]证明 l_1 范数是 l_0 范数很好的松弛，即 l_1 范数将诱导向量的稀疏性，因此 W 的 $l_{1,2}$ 范数将使得 W 有行稀疏的特性，即 $\|W\|_{1,2}$ 将诱导多特征间的相似稀疏模式。由此MFJSC模型可以转化为式(2)所示的最优化问题：

$$\min_W \frac{1}{2} \sum_{k=1}^K \|y^k - D^k w^k\|_2^2 + \lambda \|W\|_{1,2} \quad (2)$$

其中， λ 为控制着 W 稀疏性的非负参数， λ 值越大，

$$Z^* = \arg \min_Z \sum_{i \in V} R_i(z_i) + \gamma \sum_{i \in V} \sum_{j \in N_i} \delta(z_i, z_j) \cdot R_i(z_i) R_j(z_j) \quad (8)$$

其中 N_i 为第 i 个图像块的邻接图像块, γ 为调节参数, 而

$$\delta(z_i, z_j) = \begin{cases} 1, & z_i = z_j \\ 0, & \text{其他} \end{cases} \quad (9)$$

式(8)中的第 1 项最小化了编码误差之和, 而第 2 项则降低了相邻块之间的不相似性, 它遵循如下设计准则: 图像块的编码误差越小, 其对邻域的影响也越大。另外, γ 控制了总编码误差与空间不相似性的权衡, 若其被设置为 0, 式(8)则退化为与 SRC 相同的分类准则。

式(8)的全局最优解难于获得, 为此本文采用坐标下降方法获取局部最优解。幸运的是, 总是可以先使用 SRC 的分类准则获得一个很好的初始解, 保证算法可以得到一个较为理想的结果。

3 实验结果及分析

3.1 多特征提取

本文实验采用了如下 4 种特征: DAISY^[26], 几何模糊(Geometric Blur, GB)^[27], SIFT^[28]以及自相似特征(Self-SIMilarity, SSIM)^[29]。所有的特征都首先在一个 10×10 的网格上计算图像的局部描述子; 之后, 挑选部分局部描述子用 K-means 方法聚类形成词典(词典大小设置为 300), 最后统计局部描述子出现的频率作为图像的特征向量。

3.2 分类性能

本文采用一个公开的遥感数据集^[30]验证提出的 MFJSC 模型和多特征字典学习方法的有效性。数据集包含 21 种类型, 每种类型 100 幅图像, 图 2 给出了部分的图像样本。对于每一类型的样本, 随机划分为两部分, 50%用于训练, 50%用于测试。参与比较的方法有: SVM, LDA, SRC_L(所有特征串接为长特征), SRC_I(所有特征独立使用), MFJSC(不使用字典学习技术), MFJSC_DL。关于参数设置, SVM 方法采用了径向基核并且用交叉验证方法获取其参数, LDA 方法则采用了与文献^[31]类似的参数设置方法, SRC_L, SRC_I, MFJSC 及 MFJSC_DL 中的参数也同样地采用交叉验证方法获取。表 1 列出上述几种方法的分类精度。

相比于将所有特征串接为长向量的 SRC_L 方法以及所有特征独立使用的 SRC_I 方法, MFJSC 方法和 MFJSC_DL 方法得到了更好的实验效果, 证明了多特征联合稀疏编码对特征关联的利用更为合理。

直接使用训练样本特征作为字典组成的 MFJSC 方法, 分类效果与 LDA 和 SVM 相当, 而且它具有这样的优点: (1)无需学习, 可以避免过拟合; (2)可以很容易地扩展类型和添加新样本, 特别适用于增量学习框架。另外, 制约 MFJSC 性能的一个主要因素是字典中不包含类型间的判别信息, 为此本文将字典学习方法与 MFJSC 相结合, 并由此得到了较好的实验结果。

表 1 不同方法的分类精度(%)

方法类型	LDA	SRC_L	SRC_I	SVM	MFJSC	MFJSC_DL
分类精度	75.62	71.14	71.90	78.57	77.33	80.95

3.3 算法效率

在本小节中, 本文将比较 LDA, SRC_I, SRC_L, SVM, MFJSC 和 MFJSC_DL 的算法效率, 实验采用与 3.2 节相同的设置。表 2 给出了具

表 2 不同方法的时间效率

方法类型	训练时间(h)	测试时间(h)	每样本测试时间(s)
LDA	2.3769	1.533	5.30
SRC_L	< 0.0001	0.884	3.00
SRC_I	< 0.0001	4.361	15.00
SVM	0.0094	0.015	0.05
MFJSC	< 0.0001	0.788	2.70
MFJSC_DL	0.0890	0.717	2.50

体的训练时间和测试时间(机器配置为 Core™2 Q8400 (2.66 GHz×4) CPU 和 2 GB 内存)。

由表 2 可见, LDA 需要最多的训练时间, 而 SRC_I 因为多个特征分别稀疏表示而需要耗费最多的测试时间。SRC_L, SRC_I, MFJSC 直接使用训练样本作为字典, 因此只耗费少量的训练时间。MFJSC 在生成字典时, 需要计算式(4), 尽管矩阵的奇异值计算比较耗时, 但公式中只涉及最大奇异值, 现有软件包¹⁾可以很好地加速这一过程。MFJSC_DL 使用的字典由 2.3 节方法生成, 在字典生成过程中, 要求样本能由字典较好地联合稀疏编码, 因此在测试过程中, 该字典也能加速收敛过程。

¹⁾例如 PROPACK, <http://soi.stanford.edu/~rmunk/PROPACK/>。

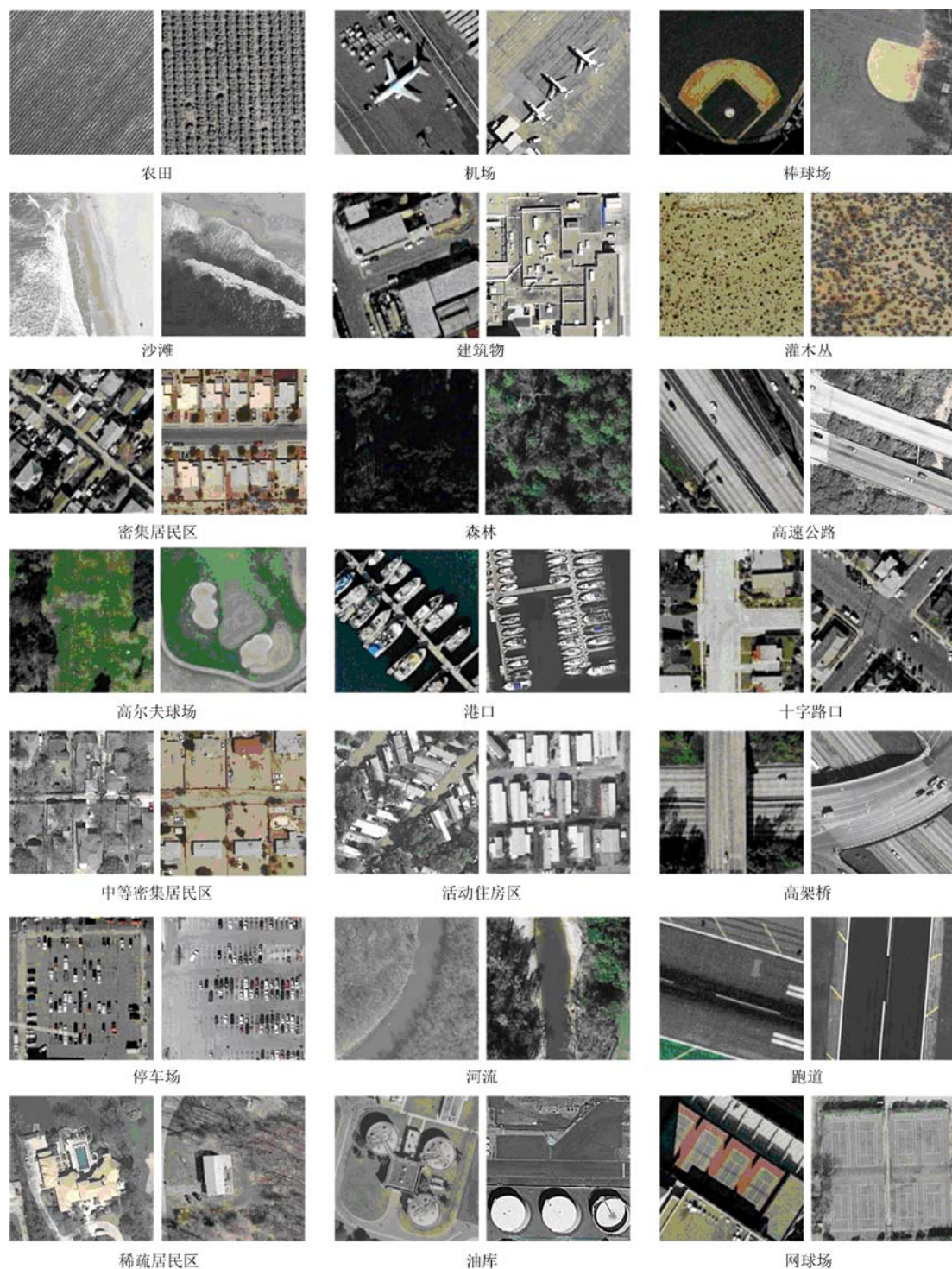


图 2 分类数据集中的部分样本图像

3.4 标注性能

在本小节中，本文用 5 幅较大的遥感影像测试所提标注方法的性能。这 5 幅影像都是 0.6 m 分辨率的全色影像，分别来自：Charleston, Claremore, Jacksonville, New Bern 和 Wilmington(见图 3(a))，影像的大小为 6000×6000 。本文方法首先将影像划分为 100×100 小块，然后自动为所有的图像块打上

标签，标签的类型为：居民区(Residential Area, RA)，公共区(Public Area, PA)，水域(Water Area, WA)和绿地(Green Area, GA)。对于每一个类型，从其他遥感影像中收集 200 个图像块作为训练样本。为了更好地评价标注方法的性能，本文参照 Google 地图和图像内容人工标注测试影像作为真值。

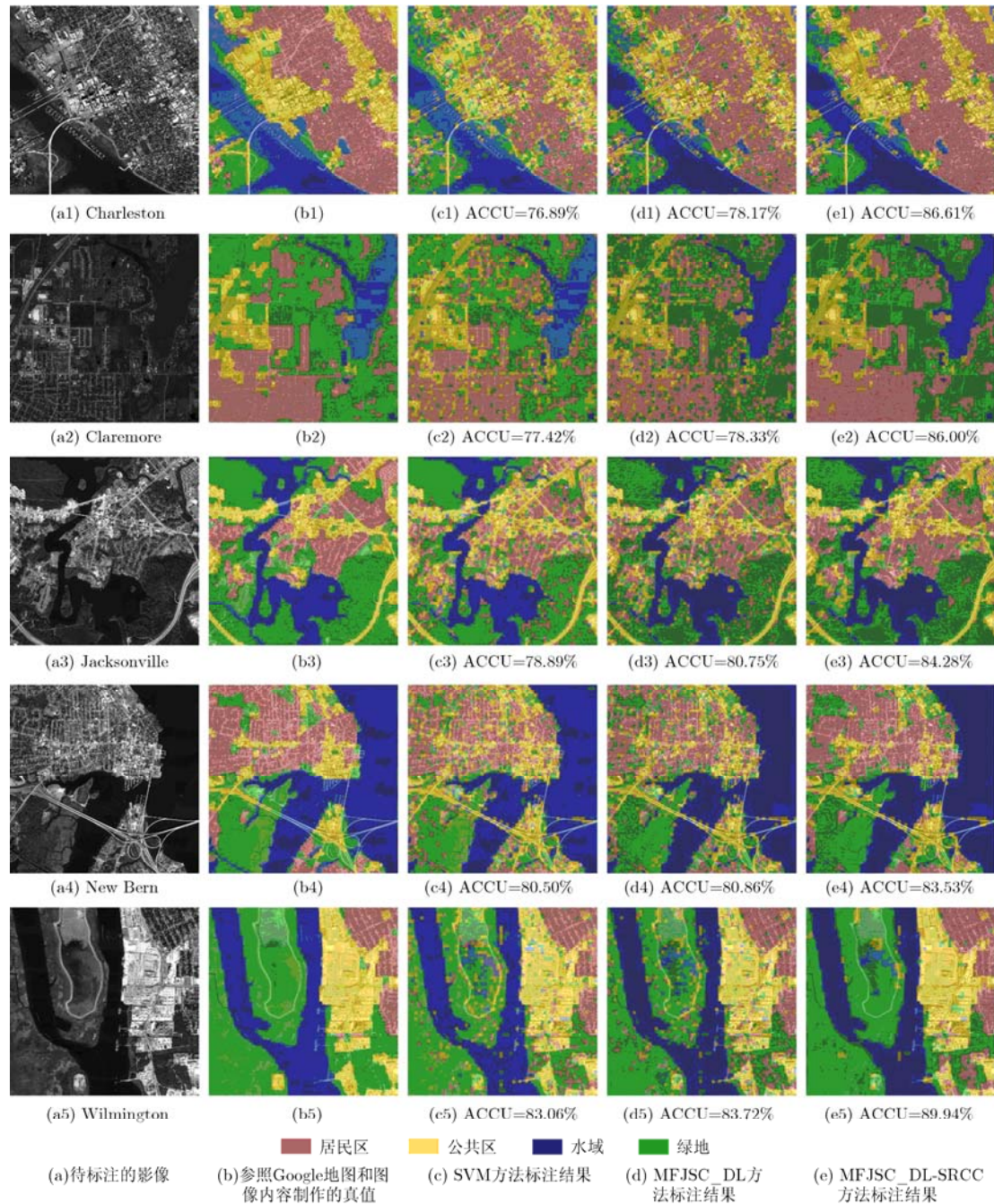


图 3 大幅遥感影像的标注结果

在本文标注方法中, MFJSC_DL 模型首先用来获取每个图像块的编码误差, 然后利用 SRCC 得到整个影像的标注结果。图 3 显示了 SVM, MFJSC_DL, MFJSC_DL-SRCC 等方法的可视化标注效果。由图 3 可见, SVM, MFJSC_DL 方法对于大部分的区域都能得到正确的结果, 而且错误标签大部分是孤立的, 类似于椒盐噪声。通过加入空间依赖, MFJSC_DL-SRCC 方法得到了更精确、更平滑的结果。也就是说, 当图像块的标签并不十分确定时, 空间连续性可以帮助标注系统决策, 从而得到更好的标注结果。

4 结束语

本文提出了空间约束多特征联合稀疏编码模型, 结合字典学习技术来更有效地标注遥感影像。该模型首先将稀疏表示分类器以及相应的字典学习方法扩展到多特征框架中, 然后设计了新的分类准则来融入空间关系约束。在遥感数据集上的实验证明了提出的多特征联合稀疏编码方法的有效性以及空间关系约束分类器对标注性能的提升。

参考文献

- [1] Wilkinson G. Results and implications of a study of fifteen

- years of satellite image classification experiments[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2005, 43(3): 433-440.
- [2] Bruzzone L and Carlin L. A multilevel context-based system for classification of very high spatial resolution images[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2006, 44(9): 2587-2600.
- [3] Bruzzone L, Marconcini M, and Persello C. Fusion of spectral and spatial information by a novel SVM classification technique[C]. Proceedings of International Geoscience and Remote Sensing Symposium, Barcelona, 2007: 4838-4841.
- [4] Liéno M, Maître H, and Datcu M. Semantic annotation of satellite images using latent dirichlet allocation[J]. *IEEE Geoscience and Remote Sensing Letters*, 2010, 7(1): 28-32.
- [5] Camps-Valls G and Rodrigo-Gonzalez A. Classification of satellite images with regularized adaboosting of RBF neural networks[C]. Proceedings of Speech, Audio, Image Biomedicine Signal Processing Using Neural Networks, Berlin, 2008: 307-326.
- [6] Donoho L. Compressed sensing[J]. *IEEE Transactions on Information Theory*, 2006, 52(4): 1289-1306.
- [7] Wright J, Ma Y, Mairal J, et al. Sparse representation for computer vision and pattern recognition[J]. *Proceedings of the IEEE, Special Issue on Applications of Compressive Sensing & Sparse Representation*, 2010, 98(6): 1031-1044.
- [8] 何艳敏, 甘涛, 陈武凡. 基于稀疏表示的两级图像去噪[J]. 电子与信息学报, 2012, 34(9): 2268-2272.
He Yan-min, Gan Tao, and Chen Wu-fan. Two-stage image denoising based on sparse representations[J]. *Journal of Electronics & Information Technology*, 2012, 34(9): 2268-2272.
- [9] Elad M. Sparse and Redundant Representations: from Theory to Applications in Signal and Image Processing[M]. New York: Springer, 2010: 227-237.
- [10] Wright J, Yang A, Ganesh A, et al. Robust face recognition via sparse representation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009, 31(2): 210-227.
- [11] Jiang Z, Lin Z, and Davis L. Learning a discriminative dictionary for sparse coding via label consistent K-SVD[C]. Proceedings of International Conference on Computer Vision and Pattern Recognition (CVPR), Colorado Springs, 2011: 1697-1704.
- [12] Yang M, Zhang L, Feng X, et al. Fisher discrimination dictionary learning for sparse representation[C]. Proceedings of IEEE International Conference on Computer Vision (ICCV), Barcelona, 2011: 543-550.
- [13] Ma L, Wang C, Xiao B, et al. Sparse representation for face recognition based on discriminative low-rank dictionary learning[C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Providence, 2012: 2586-2593.
- [14] 宋相法, 焦李成. 基于稀疏编码和集成学习的多示例多标记图像分类方法[J]. 电子与信息学报, 2013, 35(3): 622-626.
Song Xiang-fa and Jiao Li-cheng. A multi-instance multi-label image classification method based on sparse coding and ensemble learning[J]. *Journal of Electronics & Information Technology*, 2013, 35(3): 622-626.
- [15] Dai D and Yang W. Satellite image classification via two-layer sparse coding with biased image representation[J]. *IEEE Geoscience and Remote Sensing Letters*, 2011, 8(1): 173-176.
- [16] Jiang C, Zhang B, Zhang Z, et al. Experimental results and analysis of sparse microwave imaging from spaceborne radar raw data[J]. *SCIENCE CHINA Information Sciences*, 2012, 55(8): 1801-1815.
- [17] 宋相法, 焦李成. 基于稀疏表示及光谱信息的高光谱遥感图像分类[J]. 电子与信息学报, 2012, 34(2): 268-272.
Song Xiang-fa and Jiao Li-cheng. Classification of hyperspectral remote sensing image based on sparse representation and spectral information[J]. *Journal of Electronics & Information Technology*, 2012, 34(2): 268-272.
- [18] Bengio S, Pereira F, Singer Y, et al. Group sparse coding[C]. Proceedings of Advances in Neural Information Processing Systems, Vancouver, 2009: 82-89.
- [19] Gong P, Ye J, and Zhang C. Robust multi-task feature learning[C]. Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Beijing, 2012: 895-903.
- [20] Chen C, Wei C, and Wang Y. Low-rank matrix recovery with structural incoherence for robust face recognition[C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Providence, 2012: 2618-2625.
- [21] Feng Z, Yang M, Zhang L, et al. Joint discriminative dimensionality reduction and dictionary learning for face recognition[J]. *Pattern Recognition*, 2013, 46(8): 2134-2143.
- [22] Beck A and Teboulle M. A fast iterative shrinkage-thresholding algorithm for linear inverse problems[J]. *SIAM Journal on Imaging Sciences*, 2009, 2(1): 183-202.
- [23] Chen X, Lin Q, Kim S, et al. Smoothing proximal gradient method for general structured sparse regression[J]. *The Annals of Applied Statistics*, 2012, 6(2): 719-752.
- [24] Ramirez I, Sprechmann P, and Sapiro G. Classification and clustering via dictionary learning with structured incoherence and shared features[C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), San Francisco, 2010: 3501-3508.
- [25] Lafferty J, McCallum A, and Pereira F. Conditional random fields: probabilistic models for segmenting and labeling

- sequence data[C]. Proceedings of International Conference on Machine Learning, Williamstown, 2001: 282-289.
- [26] Tola E, Lepetit V, and Fua P. A fast local descriptor for dense matching[C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Anchorage, 2008: 1-8.
- [27] Berg A and Malik J. Geometric blur for template matching[C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Kauai, 2001: 607-614.
- [28] Lowe D. Distinctive image features from scale-invariant keypoints[J]. *International Journal of Computer Vision*, 2004, 60(2): 91-110.
- [29] Shechtman E and Irani M. Matching local self-similarities across images and videos[C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Minneapolis, 2007: 1-8.
- [30] Yang Y and Newsam S. Spatial pyramid co-occurrence for image classification[C]. Proceedings of IEEE International Conference on Computer Vision (ICCV), Barcelona, 2011: 1465-1472.
- [31] Li F and Perona P. A bayesian hierarchical model for learning natural scene categories[C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), San Diego, 2005: 524-531.
- 郑歆慰: 男, 1987 年生, 博士生, 研究方向为遥感图像处理、计算机视觉。
- 胡岩峰: 男, 1976 年生, 研究员, 研究方向为信号与信息处理、图像处理。
- 孙 显: 男, 1981 年生, 助理研究员, 研究方向为遥感图像处理、人工智能。

《雷达学报》“太赫兹雷达技术”专题征文通知

太赫兹波是国际科技界公认的一个非常重要的前沿技术领域, 是极具研究和开发价值的新频率资源, 它具有波束宽度窄、系统带宽大和多普勒效应显著等独特优点。太赫兹雷达探测系统融合了微波和红外可见光的优点, 在物体成像、环境监测、医疗诊断、射电天文、安全检查、反恐探测等领域具有重大的科学价值和广阔的应用前景, 特别对于高分辨雷达成像有着巨大的发展潜力。

尽管太赫兹雷达探测系统还处于理论与验证阶段, 但是其重要的科学价值和广泛的应用前景已经引起了学术界的普遍关注和极大兴趣。为及时总结太赫兹雷达研究与应用的最新进展, 《雷达学报》拟在 2015 年上半年出版“太赫兹雷达技术”专题, 专题主编为电子科技大学工程学院皮亦鸣教授。

现公开征集专题论文, 诚挚欢迎国内外专家、学者和科研人员积极投稿, 具体征文事项如下。

一. 征文方向

1. 太赫兹雷达探测系统理论
2. 太赫兹雷达目标探测技术
3. 太赫兹雷达高分辨成像技术
4. 太赫兹雷达系统应用

二. 征文要求

1. 观点明确, 论据充分, 引证准确; 尚未公开发表; 中英文皆可。
2. 来稿必须进行脱密处理, 发表前应提交保密审查表(盖章)。
3. 题目、作者姓名及单位、摘要、关键词应给出中英文; 单位和物理量使用符合国家标准。
4. 稿件请采用 Office Word 97~2003 版本通栏排版。
5. 投稿请在《雷达学报》网站(<http://radars.ie.ac.cn>)“作者投/查稿”处注册、上传, 在作者留言中注明“太赫兹雷达技术”专题论文。
6. 投稿截止日期: 2014 年 11 月 1 日。
7. 联系人: 贾守新、高华, 电话: 010-58887062, 邮箱: radars@mail.ie.ac.cn。