

利用特征选择的遥感图像场景分类

徐侃, 陈丽君, 杨文, 孙洪

(武汉大学 电子信息学院, 430079 武汉, jcarloswhu@msn.com)

摘要: 为了提高遥感图像场景分类精度, 提出了一种基于增广 LDA (Latent Dirichlet Allocation) 模型的特征选择算法. 首先对图像进行尺度不变特征变换、颜色直方图、几何模糊特征、局域二值模式和 Gabor 纹理特征提取, 然后引入一种改进的自动选择特征算法, 通过交叉验证选出最具针对性的特征组合, 再利用 LDA 将高维特征组合进行降维, 最后使用正则化逻辑回归分类器完成场景分类. 实验结果表明, 与其他特征组合相比, 经自动选择后的特征组合可以有效提高遥感图像场景分类的精度.

关键词: 场景分类; 主题模型; 特征选择; 逻辑回归

中图分类号: P237.4

文献标志码: A

文章编号: 0367-6234(2011)09-0117-05

Scene categorization of satellite images based on feature selection

XU Kan, CHEN Li-jun, YANG Wen, SUN Hong

(School of Electronic Information, Wuhan University, 430079 Wuhan, China, jcarloswhu@msn.com)

Abstract: To improve the accuracy in scene categorization of satellite images, this paper presents an algorithm of feature selection based on augmented LDA (Latent Dirichlet Allocation) model, and the algorithm is improved, which can automatically selects features from the features-pool. This method firstly extracts five kinds of features (SIFT, Geometric Blur, LBP, Gabor and Color histogram) from each image, and during the cross-validation, the combined features, which have the best performance over the dataset are got. Next, the dimensionality of the combined features is reduced by using LDA. Finally the regularized logistic regression classifier are employed to achieve the classification. Compared with other feature combination, the experimental results demonstrate that, the combination of the automatically selected features can improve the accuracy of scene categorization of satellite images effectively.

Key words: scene categorization; latent dirichlet allocation; feature selection; logistic regression

场景分类是遥感图像解译的一个重要环节, 也是遥感研究领域的热点. 如今随着图像分辨率的大幅提高, 其所包含的地物目标细节更加明显, 形状、纹理、结构等信息也更加突出. 面对丰富的特征信息, 选择有针对性的特征是提高分类性能的一个关键环节. 文献[1]的实验结果表明, 利用形状和颜色特征可以有效提取出高分辨率遥感图像中的建筑物目标. 文献[2]将遥感图像的纹理及颜色等特征组合后与 pLSA (Probabilistic Latent Semantic Analysis) 模型相结合, 利用两层分类器

来提高分类精度.

近年来, 利用语义模型来进行图像分类的方法受到了广泛关注. 与常用的词袋 (Bag of Words, BOW) 方法相比, 语义模型在对特征降维的同时, 也对潜在语义信息进行挖掘. 文献[3]将遥感图像的语义标注与 LDA 模型^[4]相结合, 取得了不错的效果. 文献[5]则将本体技术语义与 BOW 相结合, 提出了基于概率的复杂目标图像语义推理模型, 借此得到图像更深层次的语义推理结果. 语义模型根据图像中视觉词汇出现的总体情况进行分析, 不仅适用于单特征, 对于特征组合也同样有效. 文献[6]的实验结果表明将若干种特征进行组合, 分类效果要优于单特征, 但并非使

收稿日期: 2011-03-06.

基金项目: 国家自然科学基金资助项目(40801183, 60872131).

作者简介: 徐侃(1983—), 男, 博士研究生.

用特征的种类越多分类效果就越好.

对于选出有针对性的特征,文献[6]提出利用一种增广的 LDA 模型(augmented LDA, aLDA)来完成对特征的最优选择,最终将其应用在材料识别上.为了提高计算效率,本文对其算法加以改进.对 18 类遥感图像的场景分类实验表明,改进的算法不再需要进行大量迭代计算,经自动选择后组合的特征与单特征、所有特征组合以及文献[2]中的实验结果相比具有更高的分类精度.在对法国格勒诺布尔市(Grenoble)高分辨率遥感图像上所进行的大场景分类实验中,上述方法也取得了较好的效果.

1 特征提取

目前用于图像分类的特征主要包括:纹理、颜色、形状、空间位置等.它们的表达方式时各有所侧重,在分类应用中各有所长.本文在实验中运用 5 种有代表性的特征:

1) 尺度不变特征变换(SIFT)^[7].它对平移、旋转、尺度缩放、亮度变化、遮挡和噪声等具有良好的不变性,对观察视角的变化、仿射变换也具有一定的稳定性.由于其信息量丰富,容易提取,计算速度相对较快,在图像分类中得到了广泛应用.

2) 颜色直方图(colorhist).它描述的是不同颜色在整幅图像中所占的比例,反映了图像颜色的统计分布和基本色调.

3) 几何模糊特征(Geometric Blur, GB)^[8],它在模板匹配、形状匹配等方面具有良好性能,通过对图像的几何模糊来增强图像点与点之间的相关性.在实际计算中,它由图像与核函数卷积得到.

4) 局域二值模式(Local Binary Pattern, LBP).它是一种基于局部图像像素灰度值比较而提出的纹理特征,由此可定义出具体的纹理模式^[9].其核心思想是将二值矩阵看作一个二值纹理模式(纹理基元),用它来刻画邻域内像素点灰度值相对中心点的变化情况.该方法注重像素灰度的变化,符合人类视觉对图像纹理的感知.

5) Gabor 纹理特征.它作为纹理特征的一种,是对图像使用 Gabor 滤波时所得,输入图像通过二维 Gabor 函数来计算相应特征,适用于检测和描述图像纹理特性.

综上所述,SIFT 与 GB 特征主要关注图像的局部结构特征,colorhist 则是对图像颜色信息的描述,而 LBP 与 Gabor 特征侧重于图像的纹理.本实验对这些特征进行优化选择,之后将其组合应

用于遥感图像场景分类.

2 语义模型及分类器

基于语义模型的图像场景分类工作,一般通过分析图像中是否包含潜语义来完成.目前被广泛使用的 2 种语义模型 pLSA 和 LDA 都属于生成模型.它们都认为文档是由词汇构成的集合,忽略其中任何语法及出现顺序关系.当训练样本数量较小时,LDA 的优势较明显.

2.1 LDA 模型

在 pLSA 中,无法获知潜语义的分布,待估参数的数量会随文档数量增加而线性增长,相应会导致过拟合.而 LDA 模型将主题混合权重视为 k 维参数的潜在随机变量,克服了 pLSA 的不足.由模型可得到:

$$p(\theta, z, w | \alpha, \beta) = p(\theta | \alpha) \prod_{n=1}^N p(z_n | \theta) p(w_n | z_n, \beta). \quad (1)$$

将式(1)计算边缘概率得

$$p(D | \alpha, \beta) = \prod_{d=1}^M \int p(\theta_d | \alpha) \left(\prod_{n=1}^{N_d} \sum_{z_{dn}} p(z_{dn} | \theta_d) p(w_{dn} | z_{dn}, \beta) \right) d\theta_d \quad (2)$$

式中: D 为语料库; M 为语料库中文档总数; N 为文档长度; θ 为主题发生概率; z 为潜语义; α, β 分别为超参数.

在估计模型参数时可采用变分推理(variational inference)、马尔科夫链蒙特卡罗采样法(Markov Chain Monte Carlo, MCMC)等方法.

2.2 增广的 LDA 模型(aLDA)

文献[6]所提出的增广 LDA 模型(aLDA),利用贪婪算法完成对特征的最优选择.其核心思想为:在交叉验证阶段,每次从特征集中选出一种使分类正确率达到最大,即一种最好的特征,将其与之前已选出的特征进行组合,直到分类正确率不再上升为止.分类阶段,利用模型自身的参数,由最大后验原则得出类别标号,即

$$C^* = \arg \max_c (L(\alpha_c, \eta) + \lambda_c). \quad (3)$$

式中: $\lambda_c = \log \pi_c$; C 为类别标号,服从以 π 为参数的多项分布; $L(\alpha_c, \eta)$ 为模型参数估计中变分推理的最大化下界.

特征进行组合时,若有 m 种特征可供选择,则对应应有 m 个视觉词汇表,词汇表的大小记为 $|D_i| = V_i$.第 i 种特征所对应的视觉词汇为 $\{w_1^{(i)}, \dots, w_{N_i}^{(i)}\}$, $w_j^{(i)} \in \{1, 2, \dots, V_i\}$. N_i 为视觉

词汇个数. 由于各类特征独立进行聚类运算,将不同特征的视觉词汇 $\{w_1^{(1)}, \dots, w_{N_1}^{(1)}\}, \{w_1^{(2)}, \dots, w_{N_2}^{(2)}\}, \dots, \{w_1^{(m)}, \dots, w_{N_m}^{(m)}\}$ 组合起来,可表示为

$$\{W_1^{(1)}, \dots, W_{N_1}^{(1)}\}, \{W_1^{(2)}, \dots, W_{N_2}^{(2)}\} + V_1, \dots, w_1^{(m)} + \sum_{i=1}^{m-1} V_i, \dots, w_{N_m}^{(m)} + \sum_{i=1}^{m-1} V_i\}. \quad (4)$$

2.3 模型改进

在文献[6]中的交叉验证及测试阶段,均直接使用模型自身的参数得到类别标号(参见式(3)). 在确定测试样本类别标号时,其算法需要不断进行迭代计算直至模型中 λ_c 的值不再变化为止,计算时间相对较长. 除此之外,文献[11]指出在仅利用主题模型自身参数进行分类时,并不能取得令人满意的效果,若将其与判别式分类器相结合可明显提高分类精度. 为了提高分类效率,在算法中的交叉验证阶段将 LDA 模型中的参数向量 θ 输入正则化逻辑回归分类器,从而取代原算法中的式(3). 于是,在交叉验证阶段不再需要进行大量迭代运算. 另外,文献[6]中由于直接使用模型自身参数进行分类,在交叉验证及测试阶段均须基于语义模型对每一类样本单独进行训练,以获得各类别所对应的参数值,如此一来,该阶段的计算时间会随着样本类别数的增大而线性增加. 而在引入正则化逻辑回归分类器后,在交叉验证及测试阶段只需将所有样本对语义模型进行一次训练即可. 由此作出如图 1 所示的改进.

输入:视觉词汇表 $\{D_1, \dots, D_m\}$
初始化:
视觉词汇表 $D = \phi$, 平均分类准确率 $r = 0$, done = 0, iter = 0
while ~done iter = iter + 1
 利用式(4)扩充视觉词汇表 $D' = D \cup \{D_{iter}\}$
 利用 LDA 获得交叉验证数据的 θ
 将所得到的 θ 输入正则化逻辑回归分类器
 计算平均分类准确率 r_{iter}
 if $r_{iter} > r$
 $p = \arg \max_{iter} r_{iter}, D = D \cup \{D_j\}, r = r_j$
 else
 done = 1
 end
end
输出: D, r

图 1 交叉验证阶段特征选择算法

2.4 分类器

目前用于遥感图像场景分类的算法很多,例如 KNN 分类器、决策树分类器、神经网络分类器、SVM 分类器、Boosting 分类器等. 文献[10]中指出在样本有限及特征维度较高时使用正则化逻辑回

归分类器,分类效果要略优于线性 SVM 分类器,而且在速度上具有明显优势. 本实验选用 Liblinear 工具包中的正则化逻辑回归分类器.

3 实验结果

3.1 实验数据集

实验所用数据来自于 Google Earth[®]. 本文在文献[2]中数据集的基础上,新增6类场景,共有18类场景:飞机场、桥梁、商业区、沙漠、农田、足球场、森林、工业区、牧场、山脉、公园、停车场、池塘、港口、火车站、住宅区、河流以及高架桥,其中各类场景均含有 50 幅图像,大小为 600 像素 × 600 像素. 用于大场景分类实验的数据集包括6类场景:工厂、树木、草地、停车场、居民区及水域,其中每类均含有 30 幅图像,大小为 150 像素 × 150 像素.

3.2 特征计算

本文所述 5 种特征,均通过网格密集采样提取. 表 1 给出了各类特征的维数. 实验中用 K-Means 来对各类特征进行聚类运算(聚类中心的数量均为 300),从而建立视觉词汇表. 潜语义的数量设为 25.

表 1 各种特征维数

特征	SIFT	colorhist	GB	LBP	Gabor
维数	128	120	204	59	48

这样,若按式(4)将所有特征进行组合,构成的特征向量将达 1 500 维. 若使用 LDA 模型对其进行降维,便将之转化到 25 维的潜语义空间之上. 这也是将语义模型应用到遥感图像场景分类中的一个重要原因.

3.3 实验流程

实验中把视觉词汇的统计直方图作为输入语义模型的条件概率. 对于各类图像,选取 20 幅作为训练数据,10 幅作为交叉验证数据,剩余的则作为测试数据. 实验在随机抽取训练与测试图像集上进行,计算 18 个类别的平均分类准确率.

3.4 分类结果

3.4.1 场景分类实验结果

表 2 中给出了运用 BOW 方法,各类特征在 KNN 分类器下所得实验结果. 其中取 $K = 1$,即最近邻分类方法(在本文实验中 $K = 1$ 时效果较好). 表 3 给出了各类特征在经过 LDA 降维后再使用正则化逻辑回归分类器时的实验结果. 可以看出,在仅使用单特征条件下,SIFT 所对应的分类效果最佳,而 Gabor 的分类效果则相对较低.

表 4 则将不同特征组合在相同分类器下所得

结果进行了比较. 不难发现, 在特征进行组合时, 分类精度较之前仅使用单特征有了明显的提高, 但并非与使用特征种类数量成正比.

表 2 使用 BOW 单特征在 KNN ($K=1$) 分类器下所得结果					
特征	SIFT	colorhist	GB	LBP	Gabor
分类精度	58.2	55.9	52.6	51.2	50.5

表 3 使用 LDA 模型各类单特征在逻辑回归分类器下所得结果					
特征	SIFT	colorhist	GB	LBP	Gabor
分类精度	65.8	63.1	62.2	61.9	60.3

表 4 不同特征组合所得结果		%
特征组合	分类精度	
所有特征	72.5	
SIFT + colorhist + Gabor	73.6	
SIFT + colorhist	76.9	

根据经验性选择, 分别将在结构、纹理及颜色方面各取一种特征进行组合 (SIFT + colorhist + Gabor), 效果要优于使用全部特征. 而根据本实验中算法所得到的特征组合 (SIFT + colorhist), 其分类精度与前两者比相都要高. 这说明 SIFT 和 colorhist 具有一定程度的互补性, 对于该数据集, 这 2 种特征具有更强的针对性.

图 2 给出了在使用特征优化组合情况下所得到的混淆矩阵. 混淆矩阵的横轴与纵轴分别代表场景类别, 第 i 行 j 列的值表示第 i 类图像被分为第 j 类图像的比例, 其对角线上元素的值代表每类场景的分类准确率.

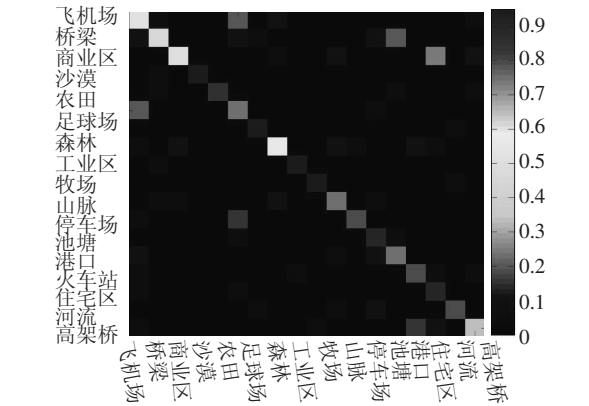


图 2 特征组合所得到的分类混淆矩阵

从图 3 中不难发现, 在所有场景中, 桥梁是最容易分错的类别, 其次则是商业区与工业区. 尤其是商业区, 有很大一部分被错分为住宅区. 由于该类中含有较多房屋与树木, 而在住宅区中同样含

有较多相似内容, 因此成为其错分的重要原因. 而对于桥梁与港口这 2 类图像, 水域在其中占有较高比例, 故桥梁中错分的图像主要被分为港口. 对于飞机场与足球场而言, 飞机场的跑道与足球场周围的公路在结构上具有一定的相似性, 是二者相互错分的主要因素.

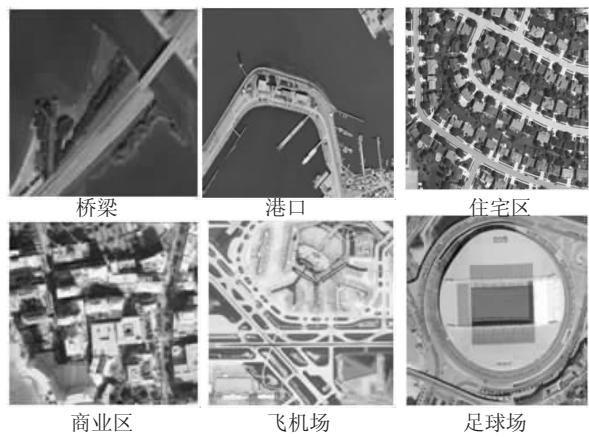


图 3 错分图像

3.4.2 大场景分类实验结果

在大场景分类实验中, 选取法国格勒诺布尔市 (Grenoble) 的城区场景作为测试图像, 大小为 6 000 像素 $\times$ 6 000 像素. 实验中把测试图像分为 1 600 个大小为 150 像素 $\times$ 150 像素的子块, 根据已建立的训练数据集选出最佳特征组合 (仍为 SIFT + colorhist). 表 5 给出了使用 2 种语义模型所得分类准确率及相应的 kappa 系数. 图 4 给出了基于 2 种模型的区域分类结果, 其中类别与颜色的对应关系为: 工厂 (粉红色), 树木 (绿色), 草地 (洋红色), 停车场 (蓝色), 居民区 (黄色) 以及水域 (灰色). 在手工标注图 (GROUND TRUTH) 中由于公路等地物不属于上述任何一类, 被标注为空类 (橙色).

表 5 大场景分类实验所得结果		
语义模型	分类精度/%	Kappa 系数
pLSA	75.7	0.61
LDA	79.0	0.64

结果显示, 除居民区之外, 其余 5 类场景都有较大一部分被错分为居民区. 这是因为在该类别中含有较多在其余类别中也会出现的地物目标, 例如, 房屋、树木、绿地、道路及位于道路上的汽车、喷水池等. 特别是草地与停车场这 2 类场景, 它们所包含的地物类型在居民区中同样也有较高出现频率, 因此其分类精度明显要低于其他类别.

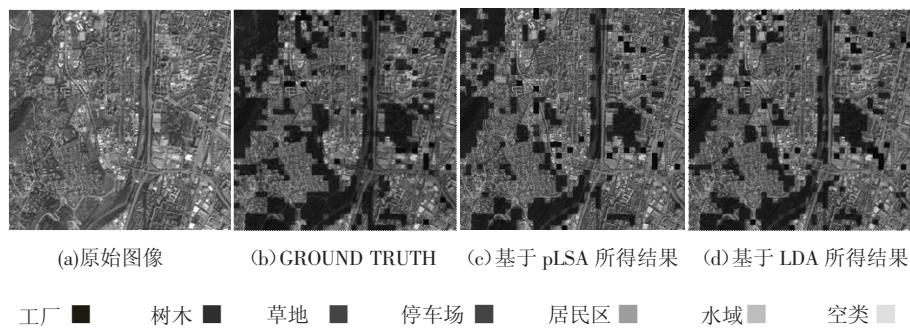


图 4 大场景分类实验结果

4 结 论

1)通过一种改进的特征自动选择算法,在交叉验证阶段选择出对于数据集最具针对性的特征,再将与语义模型相结合,完成遥感图像的场景分类。

2)实验评估结果表明,经有效选择的特征组合获得的分类结果优于使用所有特征及经验性特征组合。

参考文献:

[1] 孙 显,王宏琦,张 正. 基于对象的 Boosting 方法自动提取高分辨率遥感图像中建筑物目标[J]. 电子与信息学报,2009, 31(1):177 – 181.

[2] 徐 侃,杨 文,陈丽君,等. 基于主题模型的遥感图像场景分类[J]. 武汉大学学报(信息科学版),2011, 36(5):540 – 543.

[3] LIÉNOU M, MAÎTRE H, DATCU M. Semantic annotation of satellite images using latent dirichlet allocation [J]. IEEE, Geoscience and Remote Sensing Letters, 2009, 7(1):28 – 32.

[4] BLEI D M, NG A Y, JORDAN M I. Latent dirichlet allocation [J]. Journal of Machine Learning Research, 2003, 3: 993 – 1022.

[5] 黄宇,付琨,吴一戎,等. 基于概率的复杂目标图像语义推理模型 [J]. 武汉大学学报(信息科学版), 2009, 34(9):1043 – 1046.

[6] LIU C, SHARAN L, ADELSON E H, *et al.* Exploring features in a Bayesian framework for material recognition [C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Washington, DC: IEEE Computer Society, 2010: 239 – 246.

[7] LOWE D G. Distinctive image features from scale-invariant key-points [J]. International Journal of Computer Vision, 2004, 60(2): 91 – 110.

[8] BERG A C, MALIK J. Geometric blur for template matching [C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Washington, DC: IEEE Computer Society, 2001: 607 – 614.

[9] OJALA T, PIETIKÄINEN M, MÄENPÄÄ T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(7): 971 – 987.

[10] FAN Rong-En, CHANG Kai-Wei, HSIEH Cho-Jui, *et al.* LIBLINEAR: A library for large linear classification [J]. The Journal of Machine Learning Research, 2008, 9:1871 – 1874.

[11] ZHU J, AHMED A, XING E P. MedLDA: Maximum margin supervised topic models for regression and classification [C]//Proceedings of the 26th International Conference on Machine Learning (ICML). New York, NY: ACM, 2009: 1257 – 1264.

(编辑 张 红)