

基于词包模型的高分辨率 SAR 图像特征提取

鲁萌萌^{1,2} 赵凤军^{1,2} 李 宁^{1,2}

(1. 中国科学院电子学研究所航天微波遥感系统部 北京 100190; 2. 中国科学院大学 北京 100039)

摘 要: 特征提取在图像处理中是重要的一环,传统的特征提取算法已无法满足高分辨率图像的要求。研究运用高分辨率 SAR 图像的词包模型特征提取算法,旨在进一步优化对高分图像的解析。首先通过 SIFT 算法提取图像关键点,再对关键点进行特征向量提取。在词包模型的特征向量提取中,将边缘算子和 WLD 描述子作为新的特征向量加入词包模型中,以此提高特征分析对于边缘检测能力以及减少光照差带来的影响。通过对什邡城区 SAR 图像实测数据的特征提取和分类实验,证明新的词包模型算法具有更好的稳定性和有效性。

关键词: 词包模型;特征提取;韦伯局部描述子;高分辨率 SAR 图像;图像分类

中图分类号: TP753 **文献标识码:** A **国家标准学科分类代码:** 510.810

Feature extraction of high-resolution SAR images based on bag-of-word model

Lu Mengmeng^{1,2} Zhao Fengjun^{1,2} Li Ning^{1,2}

(1. Department of Space Microwave Remote Sensing System, Institute of Electronics, Chinese Academy of Sciences, Beijing 100190, China; 2. University of Chinese Academy of Sciences, Beijing 100039, China)

Abstract: Feature extraction is one of the most essential parts in image processing. For traditional algorithms of feature extraction cannot satisfy the high-resolution (HR) images, this paper applies bag-of-word (BoW) model algorithm to optimize the analysis of HR images feature extraction. First, key points are found by using SIFT algorithm. Second, feature vectors are extracted from key points. In terms of feature extraction of BoW model, this paper propose the mean ratio and Weber Local Descriptor (WLD) as new feature vector to improve the performance of ratio detection and decrease the illumination effect. In the feature extraction experiment using the database of Shifang SAR image, the result indicates that the new BoW model has better robustness and efficiency.

Keywords: bag-of-word (BoW) feature extraction; weber local descriptor (WLD); high-resolution SAR image; image classification

1 引 言

随着 SAR 图像处理技术的不断发展,对于图像分析的要求也在不断提高。以往传统的图像特征分析方法,例如应用广泛的灰度共生矩阵(GLCM)^[1]、Gabor 滤波器^[2]和马尔科夫随机场(GMRF)^[3]等,在以往低、中分辨率 SAR 图像中起着很大作用,但它们对于高分辨率 SAR 图像的特征分析并不能获得满意的结果。

为了满足高分辨率图像的特征分析需求,研究者们找到了很多新的特征提取分析方法。其中,在参考文献[4]中介绍的多层次局部模式直方图(MLPH)获

得了较好的效果。另外,词包模型(BoW)是近年来图像特征分析最常用的方法之一,用于图像注释^[5]、物体分类^[6]、目标检测^[7]、地物分类^[8]等,并且在 SAR 图像领域中,词包模型的应用也越来越多。在词包模型中,复杂的图像特征被描述为规范化、有序化的视觉单词表,经实验检验在图像分类中也有很好的表现。因为词包模型优越的性能,将采用词包模型对图像进行特征分析。在参考文献[9]中,使用了比边缘特征提取法提高图像处理的边缘检测能力,并获得了很好的效果。

由于词包模型容易忽略图像的整体信息,会影响

收稿日期:2015-02

其分类准确率,所以需要找到一个更好的特征提取方法来提高词包模型的准确率和鲁棒性。

2 BoW-MVR 模型

BoW 模型最初应用于文档分类领域并因其简单而有效的优点得到了广泛的应用,计算机视觉领域的研究者尝试将同样的思想应用到图像处理中。而在遥感图像处理领域,随着卫星获得遥感图像信息的技术越来越发达,SAR 图像的清晰度不断提高,词包模型也进入遥感图像处理研究者的视野。

从总体上来讲,图像分类技术可以归纳为 2 种模型:一种是基于物体的颜色、特征区域、纹理等外在信息来进行判断;另一种则是根据物体的外在信息和内在空间关系来进行判断。BoW 模型则属于前一种模型的代表。该模型最初应用于文档处理领域,将文档表示成顺序无关的关键词的组合,通过统计文档中关键词出现的频率来进行匹配。近几年来,计算机视觉领域的研究者们成功地将该模型的思想移植到图像处理领域,通过对图像进行特征提取和描述,得到大量特征进行处理,从而得到用来表示图像的关键词,并在此基础上构建视觉词典,然后对待分类图像采用相同的处理方法,将结果代入到训练的分类器中进行分类。

该模型的具体流程如图 1 所示。

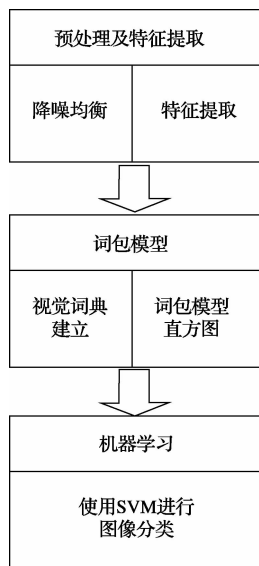


图 1 词包模型的流程

2.1 特征提取

利用 SIFT(scale-invariant feature transform,尺度不变特征转换)算法对图像进行特征点提取^[10]。SIFT 是一种电脑视觉的算法,用来侦测与描述影像中的局部性特征,它在空间尺度中寻找极值点,并提取出其位置、尺度、旋转不变量。本文研究使用的是 Lowe 提出的经典 SIFT 算法。

对获得的图像特征点进行特征分析,获得特征向量描述子,这一步骤是整个词包模型中非常重要的一环,词包模型的分类精确度主要由特征分析算法决定,因此主要研究都在特征分析算法上。最基本的特征分析向量成分为均方比,对一个特征点,取以它为中心的小图像(选择如 3×3 、 5×5 等大小)。计算获得的所有像素点的均值 m 和方差 v ,计算出均方比 L ,式(1):

$$L = \frac{m^2}{v} \quad (1)$$

均方比主要描述分辨出均匀纹理子图像,而对于类似建筑等的边缘检测没有很好的效果,并且分辨率受散斑噪声影响较大,因此引入第二个特征向量描述子:比边缘检测器。比边缘检测器对于乘性斑点噪声有很好的抑制作用,同时可以较好的进行边缘检测。

比边缘定义为特征点两侧的相邻像素区域均值的比。为了检测到所有可能出现的边缘,比边缘应该检测所有可能的方向,并找到方向中比值最大的,即为边缘方向。在本次研究中,取像素点周围的 4 个方向,即 0° 、 45° 、 90° 和 135° 两侧的像素区域将分别计算其均值 (μ_1^i 和 μ_2^i),其中 $i=0, \dots, 3$, 并计算比值。方向示意图如图 2 所示。对以上描述的 4 个方向的比边缘计算定义为:

$$r^i = 1 - \min\left(\frac{\mu_1^i}{\mu_2^i}, \frac{\mu_2^i}{\mu_1^i}\right), \quad i = 0, 1, 2, 3 \quad (2)$$

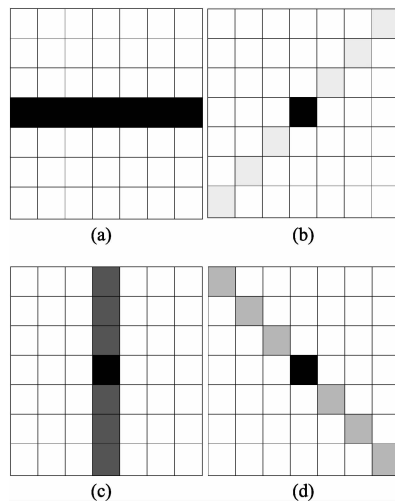


图 2 比边缘检测器的 4 个检测方向 0° 、 45° 、 90° 、 135°

最后得到的比边缘是这 4 个方向结果的最大一个,即:

$$R = \max(r^i) \quad (3)$$

利用比边缘可以较好的识别出图像的边缘方向。在该文的特征分析算法研究中,从实验图像中截取 6 种地形,分别为树林、农田、草地、城市建筑、马路和农舍,如图 3 所示。

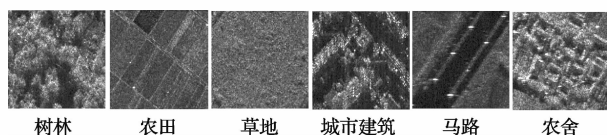


图3 6种待实验的地形

2.2 建立视觉词典

所有特征点在计算好特征向量后按照特征向量进行分类,这里分类使用的算法为 k-means 聚类算法^[11]。K-means 算法是硬聚类算法,计算数据点到原型的某种距离作为优化的目标函数,利用函数求极值的方法得到迭代运算的调整规则。在本次研究中使用欧式距离作为相似度测度,即采用误差平方和作为聚类准则函数。

由于聚类算法中的畸变函数是非凸函数,意味着不能保证取得的最小值是全局最小值,所以 K-means 对于质心初始值的选取比较敏感,即聚类准确率会随着质心初始值的变化而变化,不过一般情况下 K-means 聚类算法基本可以达到要求。在本次研究中,采取多次试验,随机选取初始值,在得到很好效果的情况下记录初始值位置,并在下次的试验中指定初始值。

根据 K-means 聚类算法,将所有特征点进行分类,编号依次为 1~100,做成直方图即为图像的词包模型。而每个类的质心的特征向量即为词包模型中每一类的特征,这 100 个特征向量即为词包模型的字典。

2.3 性能

为了测试词包模型的性能,使用高分辨率 SAR 图像的不同地形的子图形进行试验。使用的 SAR 图像为什邡城区图,98 304 × 163 840 像素大小多极化单视图,分辨率为 0.5 m × 0.5 m,入射角 45°,方位向和距离向视数分别为 4 和 1。选取了 6 个不同地区来分别分析它们的结构特征,如图 3 所示。

使用 BoW-MVR 模型的特征提取结果如图 4 显示,从图中可以看出不同地形的直方图虽然有区别,但不是很明显。另外使用了 GLCM 和 Gabor 滤波器算法进行了特征分析,结果如图 5 和图 6 所示,6 种地形分辨率极低,与词包模型进行对比之后还是可以看出传统的特征提取已经无法描述高分辨率图像的特征。

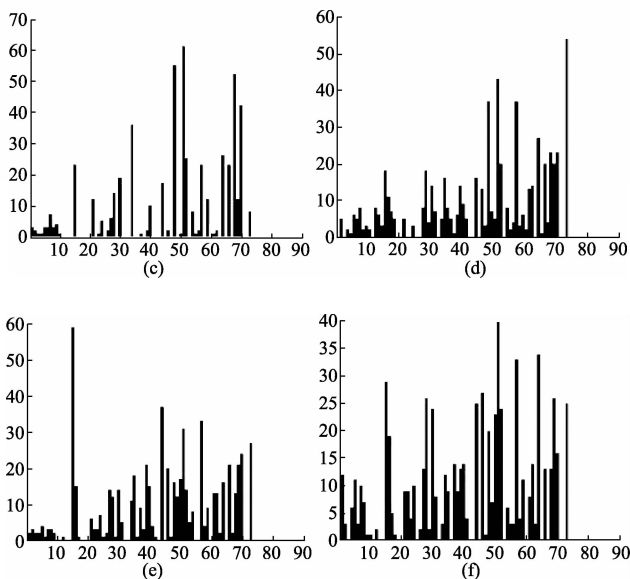
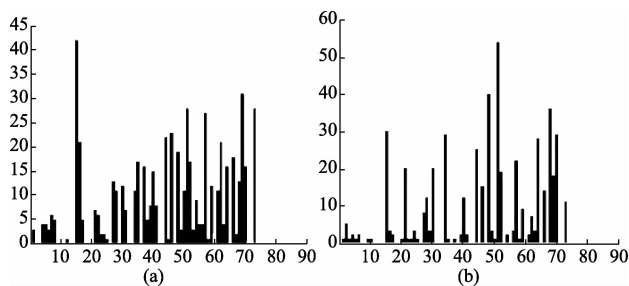


图4 6种地形的 BoW-MVR 模型特征

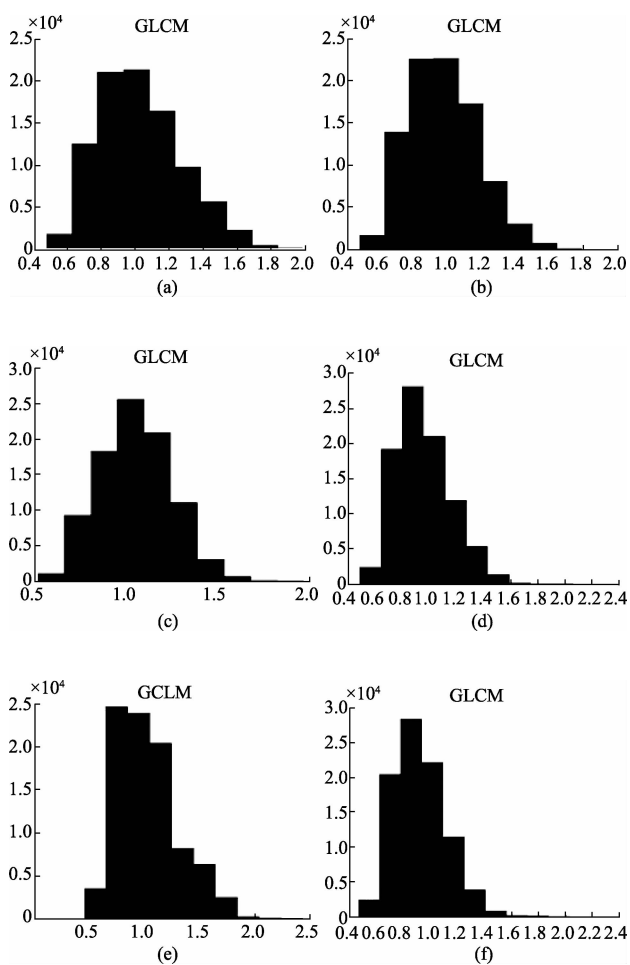


图5 6种不同地形的 GLCM 特征

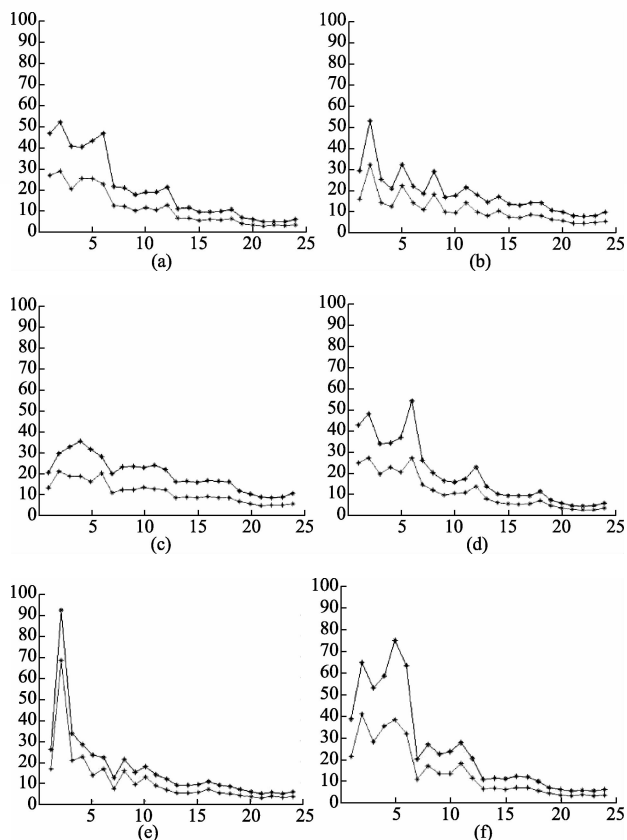


图6 6种地形的Gabor特征

3 韦伯局部描述子(WLD)

WLD描述子是根据韦伯定律提出的算法,最初用于处理光学图像。韦伯定律是指感觉的差别阈限随原来刺激量的变化而变化,并且表现为一定的规律性。引起最小可觉差的刺激变化量和原刺激强度的比值是一个常数,当刺激变化量小于这个常数时,将无法被检测出来。基于这个理论,WLD可以用于纹理特征检测。

WLD在参考文献[12]中被提出作为一个稳定的局部描述算子,它由2部分组成:差分激励 ξ 和方向 θ 。差分激励是2个部分的比值:目标像素和它周围的邻域像素的相对差分强度、以及目标像素本身的强度。方向为目标像素的方向梯度。

差分激励定义如下:

$$\xi((x_c)) = \arctan\left[\frac{v_s^{00}}{v_s^{01}}\right] = \arctan\left[\sum_{i=0}^{p-1} \left(\frac{x_i - x_c}{x_c}\right)\right] \quad (4)$$

式中: p 为邻域像素,如图7所示。在差分激励里使用了正切函数计算差分量,可以限制增长或下降过快的情况。

X_8	X_1	X_2
X_7	X_c	X_3
X_6	X_5	X_4

图7 边缘检测中定义的8个邻接像素

方向 θ 为梯度方向,定义如下:

$$\theta(x_c) = \arctan\left(\frac{v_s^{11}}{v_s^{10}}\right) \quad (5)$$

这里使用了映射:

$$f:\theta \rightarrow \theta': \theta' = \arctan2(v_s^{11}, v_s^{10}) + \pi, \text{且:}$$

$$\arctan2(v_s^{11}, v_s^{10}) = \begin{cases} \theta, v_s^{11} > 0 \text{ and } v_s^{10} > 0 \\ \pi + \theta, v_s^{11} > 0 \text{ and } v_s^{10} < 0 \\ \theta - \pi, v_s^{11} < 0 \text{ and } v_s^{10} < 0 \\ \theta, v_s^{11} < 0 \text{ and } v_s^{10} > 0 \end{cases} \quad (6)$$

尽管韦伯定律是针对光学图像而不是SAR图像邻域,但该原则对于SAR图像的特征提取依然适用。在SAR图像中有一点需要注意,与光学图像不同的是,SAR图像存在乘性散斑噪声,这将大大的减少图像分类准确率。与之前的比边缘类似,为了提高WLD描述的准确性,加入了垂直和水平方向的均差比RMD(ratio of mean differences)。RMD定义如下:

$$R_{SAR} = \frac{D_v}{D_h} = \frac{\mu_1^1 - \mu_2^1}{\mu_1^3 - \mu_2^3} \quad (7)$$

式中: $D_v = \mu_1^1 - \mu_2^1$, $D_h = \mu_1^3 - \mu_2^3$ 分别表示垂直和水平方向上的邻域像素的区域平均值。可以看出RMD和比边缘定义类似,算子大小可设定为 3×3 、 5×5 、 7×7 等。

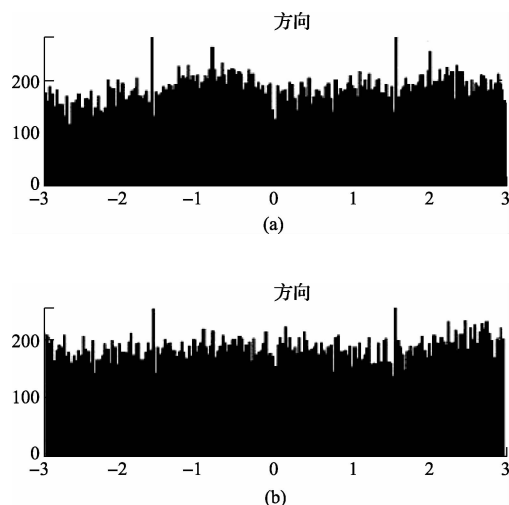
将RMD加入WLD算子中,得到改进的WLD算子,其中方向公式改进如下:

$$\theta(x_c)_{SAR} = \arctan2(R_{SAR}) = \arctan2\left(\frac{D_v}{D_h}\right) \quad (8)$$

差分激励公式改进如下:

$$\xi(x_c)_{SAR} = \arctan\left[\sum_{i=0}^3 \sum_{j=1}^2 \left(\frac{\mu_i^j - x_c}{x_c}\right)\right] \quad (9)$$

改进后的WLD对散斑噪声不敏感,并且可以很好的描述图像结构纹理。利用WLD描述子以及改进后的WLD描述子计算之前提到的6种地形子图像的差分激励和方向直方图如图8所示。



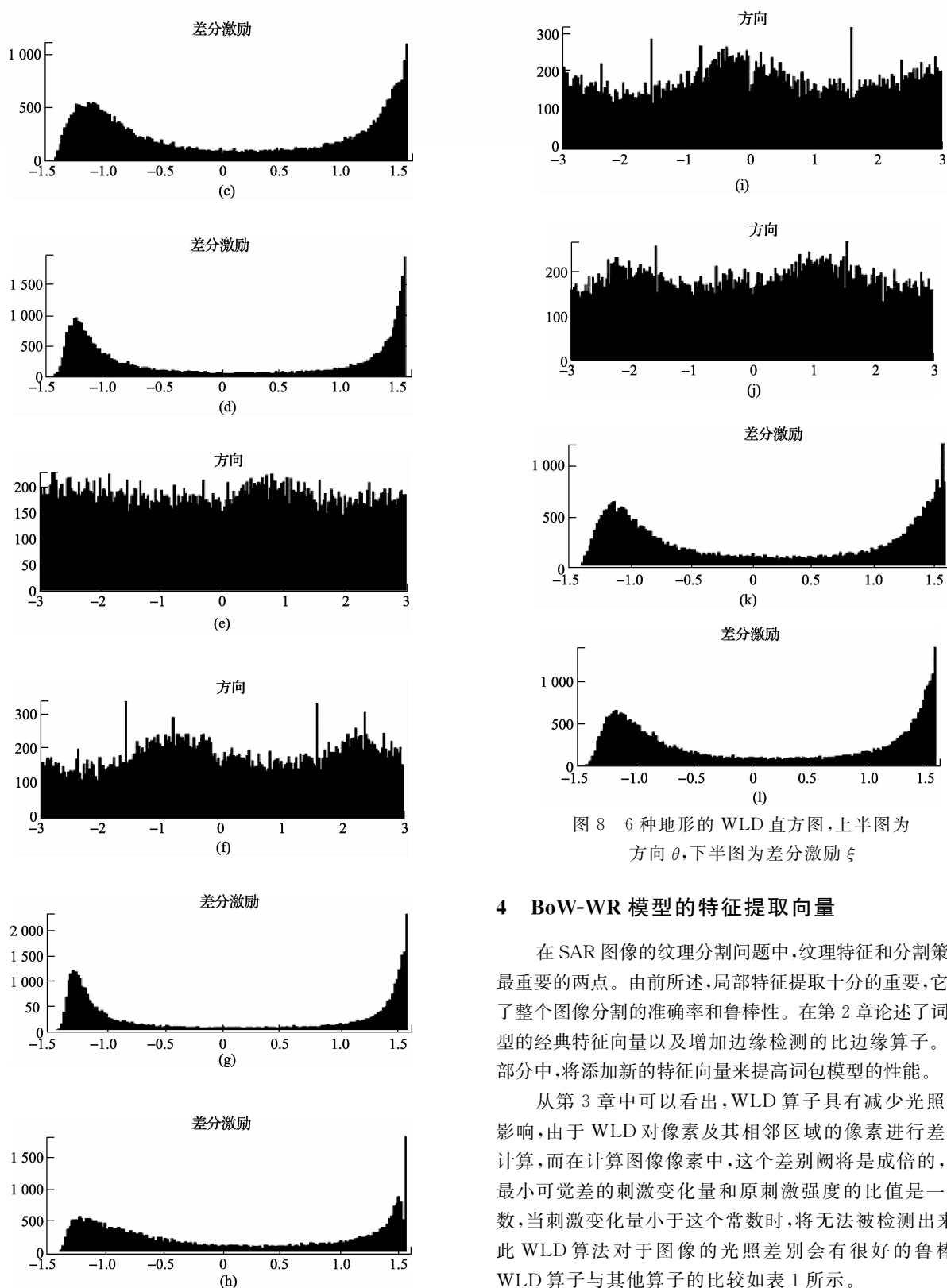


图8 6种地形的WLD直方图,上半图为方向 θ ,下半图为差分激励 ξ

4 BoW-WR模型的特征提取向量

在SAR图像的纹理分割问题中,纹理特征和分割策略是最重要的两点。由前所述,局部特征提取十分重要,它决定了整个图像分割的准确率和鲁棒性。在第2章论述了词包模型的经典特征向量以及增加边缘检测的比边缘算子。这一部分中,将添加新的特征向量来提高词包模型的性能。

从第3章中可以看出,WLD算子具有减少光照差动影响,由于WLD对像素及其相邻区域的像素进行差别阈计算,而在计算图像像素中,这个差别阈将是成倍的,引起最小可觉差的刺激变化量和原刺激强度的比值是一个常数,当刺激变化量小于这个常数时,将无法被检测出来,因此WLD算法对于图像的光照差别会有很好的鲁棒性。WLD算子与其他算子的比较如表1所示。

表1 WLD与其他算子比较

	Filtering	Labeling	statistics
LBP	Intensity difference	Thresholding at zero	Histogram over the binary strings
SIFT	Intensity	Orientation quantization	Histogram over the weighted gradient on dominant orientation
WLD	Weber ratio	Orientation quantization	Histogram over the differential excitation on dominant orientation

如上所述,将 WLD 描述子加入到词包模型的特征向量量提取中,作为新的特征向量。

5 实验结果与讨论

在之前的词包模型特征提取实验过程中,由于使用的实测数据有限,没有一个很好的学习经验,影响到机器学习能力,所以只专门对建筑和草地 2 种地形进行特征提取对比实验。提取什邡城区 SAR 图像中的一部分城区地图进行分类实验。

5.1 局部特征提取

5.1.1 BoW-MVR 模型

在特征提取过程中,在简化情况下,词包模型字典容量为 25,即横坐标为 1~25,子图像大小为 300×300 像素。首先使用 BoW-MVR 模型对两种地形进行特征提取,获得的词包模型直方图如图 9 所示。

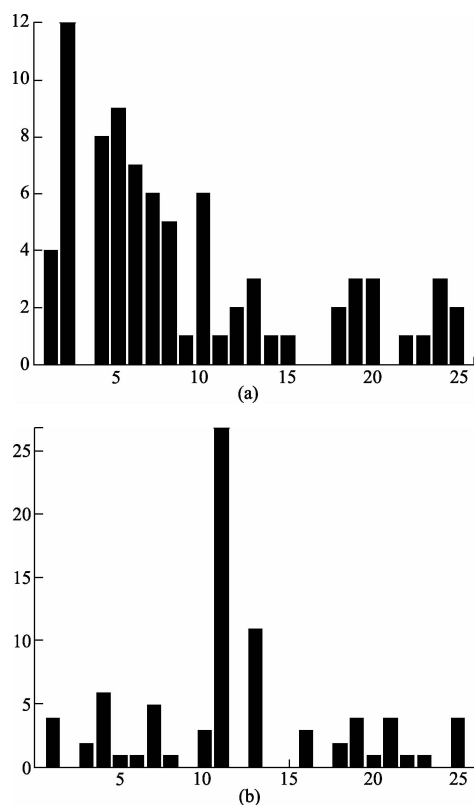


图9 两种地形的 BoW-MVR 模型直方图

5.1.2 WLD 直方图

之后先使用 WLD 描述子对两种地形进行特征向量量提取,获得 WLD 描述子直方图如图 10 所示。

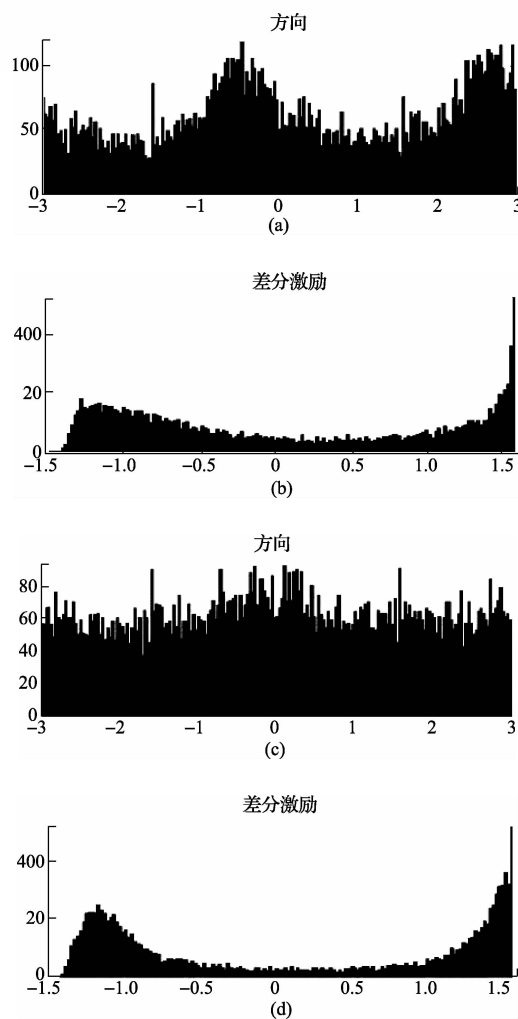
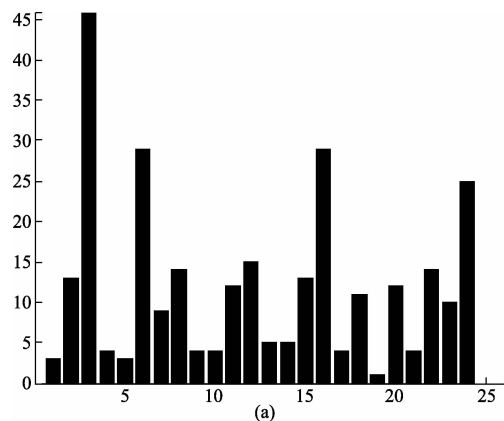


图10 两种地形的 WLD 直方图

5.1.3 使用 WLD 作为特征向量的词包模型

将 WLD 描述向量加入词包模型的特征向量中,获得 BoW-WR 模型的特征向量,BoW-WR 模型直方图如图 11 所示。



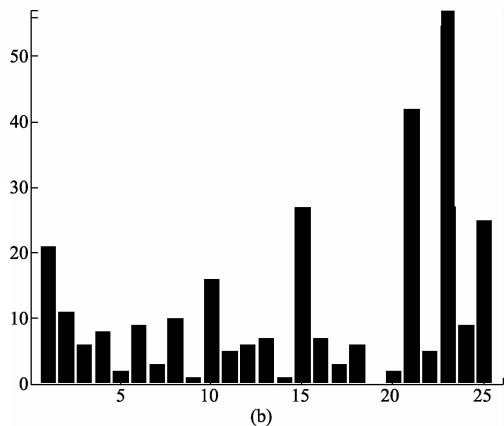


图 11 两种地形的 BoW-WR 模型直方图

5.2 实验结果

5.2.1 图像分类

传统的统计模式识别方法只有在样本趋向无穷大时,其性能才有理论保证,根据统计学习理论,学习机器的实际风险由经验风险值和置信范围值两部分组成,而基于经验风险最小化准则的学习方法只强调了训练样本的经验风险最小误差,没有最小化置信范围值,因此其推广能力较差。Vapnik 于 1995 年提出的支持向量机(support vector machine,SVM)以训练误差作为优化问题的约束条件,以置信范围值最小化作为优化目标,即 SVM 是一种基于结构风险最小化准则的学习方法,其推广能力明显优于一些传统的学习方法。

在本文图像分类就使用了支持向量机 SVM^[13]。SVM 是一种可训练的机器学习方法,主要思想可以概括为 2 点:

1)针对线性可分情况下进行分析,对于线性不可分的情况,通过使用非线性映射算法将低维输入空间线性不可分的样本转化为高维特征空间使其线性可分,从而使得高维特征空间采用线性算法对样本的非线性特征进行线性分析成为可能;

2)基于结构风险最小化理论之上在特征空间中构建最优分割超平面,使得学习器得到全局最优化,并且整个样本空间的期望风险以某个概率满足一定上界。

在本次研究中,使用了 LIBSVM 方法。LIBSVM 是台湾大学林智仁教授等开发设计的一个简单、易于使用和快速有效的 SVM 模式识别与回归的软件包。选择了 C-SVM 向量机对图像进行识别分类,核函数选择径向基核函数,参数 C 设定为 1 000。

5.2.2 实验数据

实验结果计算了分类的准确率,这可以反映出特征分析的性能高低^[14]。准确率的计算公式如下:

准确率 = $\frac{\text{正确分类}}{\text{总图像}}$ (10)

结果如表 2 所示。

表 2 3 种算法性能对比

	WLD	BoW-MVR	BoW-WR
准确率(%)	78.67	69.78	71.56
运行时间/s	35.624 467	11.162 282	10.815 475

如图 12 所示,什邡城区图中有多种复杂的地形,而本次实验为了简化只分析并分类了 2 种地形,因此准确率无法很高(有更多其他的地形被计算了进去)。图 13 显示了根据 BoW-WR 模型提取的特征进行分类的结果。因为 WLD 描述子是通过直方图来表现的,所以将这张 1 800×1 800 像素大小的图像分成 120×120 像素大小的子图像,因此共有 225 张子图像进行特征提取,每个子图像将依次计算均方比、比边缘和 WLD 计算特征向量。使用 SVM 进行分类,视觉词典之前已经建立好。图中的子图像深灰色区域代表 SVM 分类出的建筑地区,浅灰色区域代表草地地区,把其他地区的分类结果除去了,因此图中有的子图像并没有显示分类结果。



图 12 部分什邡城区图,大小为 1 800×1 800 像素

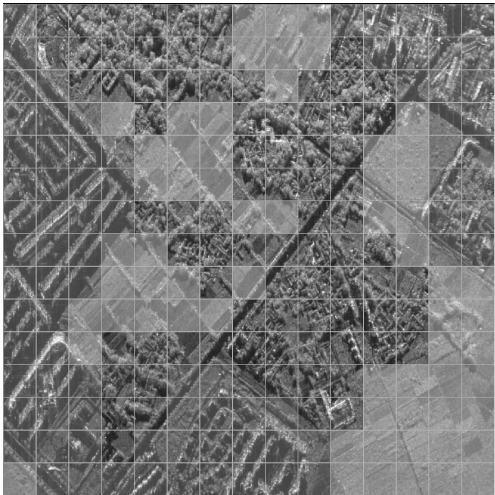


图 13 BoW-WR 模型图像分类结果

5.2.3 实验分析

分类准确率并没有很高的一个重要原因是,没有很多的实测数据(高分辨率的多种地区的 SAR 图像)可以使用,因此会很大的影响到词典建立的准确度。所以只针对 2 种地形进行特征提取分析,但即使如此,还是可以通过结果来比较这几种特征提取算法的性能高低。将 WLD 描述子放入词包模型的特征向量中,来减少光照变化带来的影响,从结果可以看出,新加了 WLD 的词包模型准确率比没有加 WLD 的词包模型有了些许的提高。之所以差距并不是很大是因为使用的 SAR 图像并没有明显的光照差,因此这 2 个词包模型都有较好的特征分析能力。WLD 直方图性能很好,但相比词包模型处理的时间要明显长很多。

6 结 论

为了满足分辨率不断提高的 SAR 图像的特征分析需求,很多新的特征提取算法取代了传统的特征分析方法,而词包模型和 WLD 描述子在这几年被引入了 SAR 图像领域,并且获得了很好的特征分析效果。基于这 2 种算法进行进一步的研究,重点关注词包模型的特征向量,在原有的均方差特征向量的基础上使用比边缘提高边缘检测能力,并把 WLD 描述子引入词包模型的特征向量,减少光照差带来的影响。通过实验可以证明新的词包模型有较好的性能,图像分类准确率比原有模型要好。在将来的工作中,会把更多的实测数据放入机器学习中,提高图像分类的准确率。

参 考 文 献

- [1] 陈强,田杰,黄海宁,等. 基于统计和纹理特征的 SAS 图像 SVM 分割研究[J]. 仪器仪表学报, 2013, 34(6):1413-1420.
- [2] 周云鹏,朱青,王耀南,等. 面部多特征融合的驾驶员疲劳检测方法[J]. 电子测量与仪器学报, 2014, 28(10):1140-1148.
- [3] 姚昆,杨学志,唐益明,等. SAR 海冰的三维区域 MRF 图像分割[J]. 仪器仪表学报, 2013, 34(11): 2551-2557.
- [4] DAI D X, YANG W, SUN H. Multilevel local pattern histogram for SAR image classification[J]. Geoscience and Remote Sensing Letters, 2011, 8(2): 225-229.
- [5] LIENOU M, MAITRE H, DATCU M. Semantic annotation of satellite images using latent dirichlet allocation[J]. Geoscience and Remote Sensing Letters, 2010, 7(1): 28-32.
- [6] XU SHENG, FANG T, LI D, et al. Object classification of aerial images with bag-of-visual words[J]. Geoscience and Remote Sensing Letters, 2010, 7(2): 366-370.
- [7] SUN H, SUN X, WANG H, et al. Automatic target detection in high-resolution remote sensing images using spatial sparse coding bag-of-words model[J]. Geoscience and Remote Sensing Letters, 2012, 9(1): 109-113.
- [8] YANG Y, NEWSAM S. Bag-of-visual-words and spatial extensions for land-use classification [C]. Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems, 2010:270-279.
- [9] CUI S, DUMITRU C O, DATCU M. Ratio-Detector-Based Feature Extraction for Very High Resolution SAR Image Patch Indexing[J]. Geoscience and Remote Sensing Letters, 2013, 10(5): 1175-1179.
- [10] 邸男,李桂菊,陈春宁,等. 结合归一化差分高斯特征的图像匹配技术研究[J]. 电子测量与仪器学报, 2014, 28(6):585-590.
- [11] FENG J, JIAO L C, ZHANG X, et al. Bag-of-visual-words based on clonal selection algorithm for SAR image classification[J]. Geoscience and Remote Sensing Letters, 2011, 8(4): 691-695.
- [12] CHEN J, SHAN S, HE C, et al. WLD: A robust local image descriptor[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 1705-1720.
- [13] 刘松松,张辉,毛征,等. 基于 HRM 特征提取和 SVM 的目标检测方法[J]. 国外电子测量技术, 2014, 33(10):38-41.
- [14] ZHANG H, LI B, WANG Y, et al. Bag-of-Words model based image classification and evaluation of image sample quality in remote sensing[C]. TENCON IEEE Region 10 Conference (31194), 2013:1-4.

作 者 简 介

鲁萌萌,1990 年出生,工学硕士。主要研究方向为 SAR 图像处理方向以及图像配准、图像特征提取分析等。
E-mail: lumeng0309@163.com