

分 类 号: TP75
研究生学号: 2012522065

单位代码: 10183
密 级: 公 开



吉 林 大 学

硕 士 学 位 论 文

(学 术 学 位)

利用空间信息的卫星图像场景分类

Satellite Image Scene Classification Using Spatial Information

作者姓名: 宋伟伟

专 业: 信号与信息处理

研究方向: 图像处理与模式识别

指导教师: 王 珂 教授

培养单位: 通信工程学院

2015 年 6 月

未经本论文作者的书面授权，依法收存和保管本论文书面版本、电子版本的任何单位和个人，均不得对本论文的全部或部分内容进行任何形式的复制、修改、发行、出租、改编等有碍作者著作权的商业性使用（但纯学术性使用不在此限）。否则，应承担侵权的法律责任。

吉林大学博士(或硕士)学位论文原创性声明

本人郑重声明：所呈交学位论文，是本人在指导教师的指导下，独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本论文不包含任何其它个人或集体已经发表或撰写过的作品成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。本人完全意识到本声明的法律结果由本人承担。

学位论文作者签名：宋伟伟

日期：2015年6月7日

利用空间信息的卫星图像场景分类

Satellite Image Scene Classification Using Spatial
Information

作者姓名：宋伟伟

专业名称：信号与信息处理

指导教师：王珂 教授

文敦伟 教授

学位类别：工学硕士

答辩日期： 2015 年 6 月 7 日

摘要

利用空间信息的卫星图像场景分类

在高分辨率卫星图像场景分类研究中，高分辨率卫星图像主要呈现两大特点：第一，随着空间分辨率的提高，图像中地物目标包含的细节更加丰富，例如结构、纹理、颜色等。第二，在高分辨率的卫星图像中，同一场景类别包含目标的尺寸和方向可能不同，不同场景类别中可能包含同样的目标。同时，在不同天气情况下，图像的亮度受光照影响较大。

以上两个特点给场景分类时图像表示带来了很大的障碍。图像的表示在场景分类中是一个关键环节，在低分辨率卫星图像分类中，纹理和强度特征就能有效地描述图像中的地物目标，由于高分辨率的卫星图像具有上述特点，利用这些简单的特征不能充分表达图像的各目标信息。为了改善卫星图像场景的表示能力和分类准确率，需要结合特征间的一些特性或者信息尽可能全面地描述各类场景。

主要的研究工作有以下三个方面：

第一，高分辨率的卫星图像的特征选择。特征的选择是卫星图像场景分类的一个关键环节。合适的特征能够更好地描述例如路面，天空，河流等一些突出的区域。SIFT (scale-invariant feature transform) 描述符具有仿射、旋转、亮度、尺度不变性的优点，考虑到上文提到高分辨率卫星图像所具有的一些特点，所以文中采用该特征描述符对图像进行描述。

第二，特征选定后，构造视觉词典。在构造视觉词典的两种方案中，本文采用了泛化能力比较好的一种方案，该方案不仅适用于在原样本库中随机挑选样本构造视觉词典，而且也同样适用于在新样本库中对视觉词典的构造。使得实验结果更加可靠真实。

第三，利用特征间的空间信息进一步加强特征的表示能力。在文中引入了空间金字塔匹配模型，通过空间金字塔建模，在图像的表示和匹配过程中，将局部特征的空间信息引入到图像的分类中，进一步改善分类性能。

实验基于 MATLAB 平台，LIBSVM 工具箱作为分类器，利用特征的空间信息的实验平均分类准确率为 82.6%。地物目标区别较大的场景类别分类准确

率较优越，通过多次实验的结果标准差比较，各组实验分类准确率受随机抽取训练和测试数据集的影响较小，鲁棒性较强。结果表明，通过空间金字塔编码，利用局部特征空间的空间位置信息，能够增强局部特征对图像的描述能力，改善卫星图像的分类性能。

关键词：

卫星图像，空间信息，空间金字塔，分类，SIFT，LIBSVM

ABSTRACT

Satellite Image Scene Classification Using Spatial Information

In the study of high resolution satellite image classification, the images show two characteristics: First, with the improvement of spatial resolution, the objects show more details in a satellite scene, such as structure, texture and color. Second, objects may appear at different orientations and scales in the same category scene of HR satellite images and different scenes may contain the same object. Meanwhile, the brightness of the same scene is influenced by lighting under different weather conditions.

The above two characteristics put big obstacles to represent images for scene classification. Image representation plays a key role in scene classification. Unlike low-resolution satellite images, which can be effectively described through texture and intensity features, these simple features cannot fully express the objects information for each HR satellite image owing to the characteristics. In order to enhance the local feature representation capacity and improve the classification performance of high resolution satellite images, we should employ some properties or relations for describing every image scene.

This research includes next three main tasks. The first is the feature selection for high resolution satellite images, which plays a key role in scene classification. A suitable feature could be good at describing some prominent regions, such as sky, road and river. Scale-invariant feature transform (SIFT) features have some invariant properties, such as affine, rotation, brightness and scale. Considering the above mentioned characteristics of high resolution satellite images, therefore, SIFT is necessary in our application.

The second is constructing visual dictionary after feature selection. There are two schemes that can be implemented to extract features for constructing visual dictionary. This research adopts one of scheme for better generalization ability. This scheme is not only suitable for the randomly selected samples for constructing the visual dictionary in original sample library, but also suitable for constructing visual dictionary in the new sample library. It makes the experimental result more convinced and reliable.

The third is employing spatial information of features to enhance the ability of

representing images. The spatial pyramid matching model (SPMM) is introduced for spatial coding and matching for HR satellite images. To improve the classification performance, through spatial pyramid modeling, the spatial information of local features is used for image classification.

The experiment is based on MATLAB platform. The LIBSVM toolbox is used to train an SVM classifier on the training set. The average classification accuracy is 82.6%. The greater the distinctiveness of different scene is, the higher the classification accuracy is. In the comparison of standard deviations for several groups experimental results, it is demonstrated that the results of each experiment are relatively stable and are less affected by choosing training set and testing set randomly, thus with a robustness. We conclude that, by utilizing the spatial information of local features through SPMM encoding, our approach can enhance the ability of describing HR satellite images and improve the classification performance.

Keywords:

Satellite Image, Spatial Information, Spatial Pyramid, Classification, SIFT, LIBSVM

目 录

第 1 章 绪论.....	1
1.1 研究背景和研究意义.....	1
1.2 国内外研究现状.....	2
1.3 论文的主要研究内容和结构安排.....	4
第 2 章 图像特征选择及视觉词典构造	7
2.1 高分辨率卫星场景图像的特点.....	7
2.2 图像局部特征的选择.....	8
2.2.1 局部二值模式 (LBP) 特征.....	8
2.2.2 尺度不变特征转换 (SIFT) 特征.....	10
2.3 视觉词典的构造.....	15
2.3.1 视觉词包模型.....	15
2.3.2 K-means 算法.....	17
2.3.3 构造视觉词典.....	18
2.4 本章小结.....	18
第 3 章 卫星图像的空间信息和分类器选择	21
3.1 空间信息的引入.....	21
3.2 空间信息与匹配.....	22
3.2.1 金字塔匹配算法.....	22
3.2.2 图像的空间金字塔匹配算法.....	24
3.3 分类器的选择.....	25
3.3.1 SVM 分类算法.....	26
3.3.2 SVM 几种常见核函数.....	28
3.4 本章小结.....	29
第 4 章 分类过程及结果分析	31

4.1 实验基本流程	31
4.2 数据集和参数设置	32
4.3 分类结果分析	35
第 5 章 全文总结	41
参考文献.....	43
作者简介及科研成果	49
致 谢.....	51

第 1 章 绪论

1.1 研究背景和研究意义

近年来, 卫星图像被广泛应用于土地测绘、农业研究、环境研究、城市区域划分等领域, 在分析和处理自然资源与人类活动关系的研究中也有许多应用。每天都有大量的卫星图像通过卫星传感器拍摄传输到地面, 供相关研究人员进行各领域的研究, 但是, 这些图像的处理给研究人员带来巨大的挑战。例如, 在图像的分析、检索与分类研究中, 仅仅通过人工操作不仅需要耗费巨大的人力和物力, 而且很难完成每天的分类任务。因此, 通过机器学习和模式识别的方法自动处理大量的卫星图像已经成为当前研究的热点。

随着拍摄设备性能的改善和发展, 所获取卫星图像的空间分辨率由过去的几十米提高到近几年的 1~5 米^[1], 例如: IKONOS、World-View-1、GeoEye-1 等, 这使得研究人员能够对所拍摄的图像进行准确的观测和测量, 甚至能够测量出地表大小是 0.5 米的物体。因此, 高分辨率卫星图像的场景识别研究将成为 21 世纪图像研究领域的一大趋势^[2]。

卫星图像分辨率的提高主要表现在为拍摄到的研究对象提供了丰富的结构信息、纹理信息和空间信息, 这些信息都有助于研究人员对研究对象特征的掌握以及场景特性的认知。例如, 图像中观测到目标对象的纹理、形状以及所处的层次属性等^[3]。然而, 高分辨卫星图像的研究模式识别中也存在很多挑战, 例如: 在同类场景图像中的同一目标会呈现不同的位置和尺度, 这也是不同类别的场景图像相互分错的主要原因。众所周知, 不同的特征在描述图像的过程中强调的重点有所不同, 纹理特征描述的纹理特性, 颜色特征描述的是图像的颜色特性。因此, 在一个大的卫星图像数据集中, 有效的以及区分力较强的特征信息在大数据集的卫星图像场景分类中扮演着一个重要的角色。

目前, 有很多方法应用于场景分类的研究, 但是对于场景类别较多的卫星图像分类研究中仍存在一些问题, 如: 光照强度、位置、尺度等因素的变化使得相同的目标对象在同类场景的不同图像中呈现的特性有所差别; 不同类别的图像中又可能包含同一目标。这就要求我们在图像中表示和分类的过程中能够尽可能有效地描述图像, 有效地区分同类别图像场景之间存在的相似性和不同

类别图像场景之间的差异性。同时，在多分类中，卫星图像类别的多少也是分类过程中要考虑的一个因素，针对以上问题，国内外很多的学者正致力于分类方法的研究。

1.2 国内外研究现状

交叉学科之间的研究已经成为时下国际研究的热点。最初人们利用机器学习和模式识别的方法处理文本，后期把该学科用于对自然图像的处理。最近，在自然图像上的研究已经开始转向对高分辨率的卫星图像场景分类的研究上。当前针对卫星图像分类的方法大多集中在传统处理文本分类的方法上，基本思路就是提取图像中某一局部特征、特征向量量化、图像表示、分类器分类等。

主题模型在文本处理中和在自然图像处理中表现出优越的性能。主题模型的基本思想就是把每个图像集视为一个文档集，一篇文档所阐述的内容一定表现在文档的主题上，而每个主题又是通过若干个相关的词语表现出来，所以主题与单词之间便存在某种潜在的映射关系，主题模型就是用来识别各图像中隐藏的主题，实现对图像的处理。因此，主题模型在高分辨率卫星图像领域中的研究开始有所发展。例如：

Lienou M^[4]等人利用图像中的语义概念对大场景的卫星图像进行标注。其主要思想是将大场景图像中各子图像块先进行有监督的分类和各图像子块之间的空间信息关系相结合。通过 LDA(Latent Dirichlet Allocation)主题模型^[5]进行建模，对图像数据集进行训练、学习和分类，场景图像中具有相同语义概念的图像子块对应于同一个主题，进而完成了对大场景卫星图像的语义标注。

Xu 等人^[6]于 2011 年提出了一种利用 pLSA(probabilistic latent semantic analysis)^[7]主题与特征组合相结合的方法对卫星场景图像进行分类。该方法的基本思路是提取图像中需要研究的不同特征，尺度不变特征对图像的结构进行描述，颜色直方图特征对图像的颜色进行描述和几何模糊特征对图像的边缘进行描述。然后通过 pLSA 主题模型对各特征向量进行降维，映射到语义空间上，实现了对各特征进行潜在主题挖掘。最后将各特征的概率主题特征组合，利用 SVM(Support Vector Machine)^[8]分类器进行训练和学习分类，实现了一种高分辨卫星场景图像在特征组合下的分类方法。

同年，Xu 等人^[9]在 LDA 主题模型的基础上对模型加以改进，提出了增广

的 LDA 模型的特征选择算法。该算法在图像的多个特征提取处理后,能够自动选择特征对卫星图像场景进行分类,最后再对具有针对性的特征进行组合,完成场景的分类。

2012 年,陶超等人^[10]提出了基于主题模型的有监督分类方法,该方法通过采用均值漂移的分割方法对卫星图像进行分割,分割后每幅图像被分成若干个图像子块,把这些图像子块组成一个集合,对该集合中的各个子块提取 Gabor 纹理特征,然后使用聚类算法构造视觉词典,最后利用主题模型进行主题分析,预测出每个图像块潜在的概率主题,进而完成识别和分类。由于高分辨率卫星图像中存在同类场景的同一目标有不同之处,不同场景的不同目标又有相似或相近之处的现象,其方法主要解决该现象对图像识别和分类过程中造成的干扰。

当然,除了利用主题模型对高分辨率卫星图像进行分类研究之外,还有一些其它的研究方法用于卫星图像场景识别的领域,以下几种方法在处理卫星图像分类时主要是从选择、提取和处理特征的角度来实现的。

Xia 等人^[11]于 2010 年提出了一种依赖于地形图^[12]和基于形状的图像检索方案^[13]的卫星图像检索方法。该方法结合了卫星图像的纹理特征和结构化特征的信息,同时该方法对于尺度变化、方向变化和对比度变化拥有较好的鲁棒性。实验结果表明该方法能平衡特征的结构信息和纹理信息,为高分辨率卫星图像的分类领域做出了一定的贡献。

Cheng 等人^[14]在 2011 年提出利用多种特征在不同分类器中对高分辨率卫星图像场景进行分类研究的方法。该方法在处理图像时提取了包括机构特征和纹理特征和颜色特征在内的 13 种特征,分别比较了每一种特征在不同分类器中获得的分类性能,然后选择分类结果比较优异的三种特征特征组合对卫星图像进行再一次分类,获得了比使用单特征分类时更优越的性能。

2012 年,Yang 等人^[15]在基于“词袋”模型的基础上,利用图像的局部不变特征对高分辨率卫星图像进行分类。作者通过调整“词袋”模型中的不同参数验证这些参数对图像分类结果的影响,包括基于网格提取局部特征时网格的大小和数量,视觉词典的尺寸大小以及在构造时采用不同的聚类算法等。

这些研究都以各自方式实现卫星图像场景分类,在分类研究中主要依赖于两个方面:一个是在利用模型方面,对图像进行建模时主要对图像特征的维数

进行降维，但是维数的降低必然会带来特征信息的丢失。同时选择模型时主要还是依赖卫星图像数据集训练样本的数目。例如，当训练样本数目比较少，LDA 主题模型的建模效果比较好，当训练样本数目比较大，pLSA 主题模型在分类时表现的综合性能比较优异。这些特点说明了在特定的卫星图像数据集上只能针对性地选择主题模型，这样就使得选择的模型在应用时泛化能力有所限制。

另一个是在特征选择方面。如果利用多特征组合来表示各场景图像，实验中特征的维数将会变得非常大，会给实验带来很大的计算复杂度和计算量。以上方法大多只是注重在特征层面上对卫星场景图像进行特征的提取以及在分类过程中图像的匹配，而忽略了在匹配时特征之间存在的一种空间位置信息，这一信息能够进一步加强特征对各图像场景的描述能力。然而在场景图像匹配时，忽略了局部特征间的这一信息，这也限制了局部特征在图像匹配时对场景图像的描述能力，使得局部特征不能充分发挥其在分类中的作用。

1.3 论文的主要研究内容和结构安排

本文依据高分辨率的卫星图像所具有的特点，在分类中大多数方法注重在特征层面上对图像场景进行分类，因此，在研究中，选择有效的特征以及如何在选定特征的基础上尽可能地增强选定局部特征对卫星场景图像进行有效的表示，最终的目的就是改善各图像场景的分类准确率。

主要研究内容如下：

(1) 数据集的获取和选择卫星图像场景中的特征

在实际的研究中，权威数据集的获取至关重要，真实权威的卫星图像场景数据集能够充分地体现现实中卫星图像的特点，研究中能够证明运用研究方法的有效性和真实性，真实数据集上的各种分类方法的实验结果比对，能够证明所采用方法较其它方法具有相对优越的性能。

以上相关的文献已经阐述了分辨率较高的卫星场景图像主要呈现两大特点，第一，随着空间分辨率的改善，图像中地物目标包含的细节更加丰富，例如结构、纹理、颜色等。第二，在高分辨率的卫星场景图像中，同一场景类别包含同一目标的尺寸和方向可能有所不同，不同场景类别中可能包含同样的目标，不同天气，图像的亮度受光照影响较大。在图像中盲目选择多特征来表示

各场景图像，可能会带来特征的维数灾难，会使分类器在训练和测试阶段耗费大量的计算量和时间。因此，针对卫星图像场景的特点，有针对性地选择某个特征或者某些特征来表示图像，是卫星图像场景分类研究中要重点考虑的一个问题。

（2）视觉词典构造过程中方案的选择

在特征选定后，如何构造有利于场景图像分类的视觉词典。下文分析了构造视觉词典的两种方案，在构造视觉词典的这两种方案中，本文采用了泛化能力比较好的一种方案，该方案不仅适用于在原样本库中随机挑选样本构造视觉词典，而且也同样适用于在新样本库中对视觉词典的构造。

（3）局部特征的空间信息

单纯地在特征层面上对场景图像进行分类，丢弃了局部特征间的空间位置信息，这也限制了所提取的特征对场景图像的描述能力。空间金字塔匹配模型对局部特征描述符进行编码建模，将局部特征的空间位置信息融入到图像的描述当中。空间金字塔编码主要思想是按照不同的分辨率对图像进行分层，然后对每层进行规则网格划分，空间几何关系较近的局部特征分到同一网格的概率较大，统计不同分辨率上各子块的局部特征直方图，从而将局部特征的空间位置关系的信息融入到对图像的描述当中。

全文的结构及主要内容安排如下，结构如图 1.1 所示。

第1章阐述高分辨卫星图像场景分类的一些主流方法和研究背景及意义，总结了图像所具有的一些主要特点。分析了在当前研究中主要存在的问题，分析了当前方法存在的一些优缺点，就这些问题本文重点研究解决其中一些问题的方法和思路。

第2章主要讨论高分辨率卫星图像的特征选择。在特征提取过程中，讨论具体采用提取的方式以及提取算法中个参数的设定。在选定的特征提取后，采用相关的聚类算法对已提取的特征进行聚类，每一个聚类中心视为一个视觉单词，构造视觉词典，在构造视觉词典时，本文选择了一种泛化能力较好的视觉词典构造方案。

第3章是图像表示和分类器选择部分。主要介绍了在图像特征提取、构造视觉词典后，在图像表示时，利用特征的空间位置信息，选择了空间金字塔匹配模型对图像进行表示，同时简单阐述了空间金字塔匹配模型的相关理论和匹配

核算法。最后阐述 SVM 分类器的相关算法以及相关的核函数。

第4章是实验对比和实验分析部分。重点阐述了实验的过程中采用的平台，实验过程存在的问题，对比调整不同参数所得到的实验结果，把实验结果同参考文献中提出的两种实验方法在相同数据集上获得的结果进行了比较和分析。

第5章全文总结

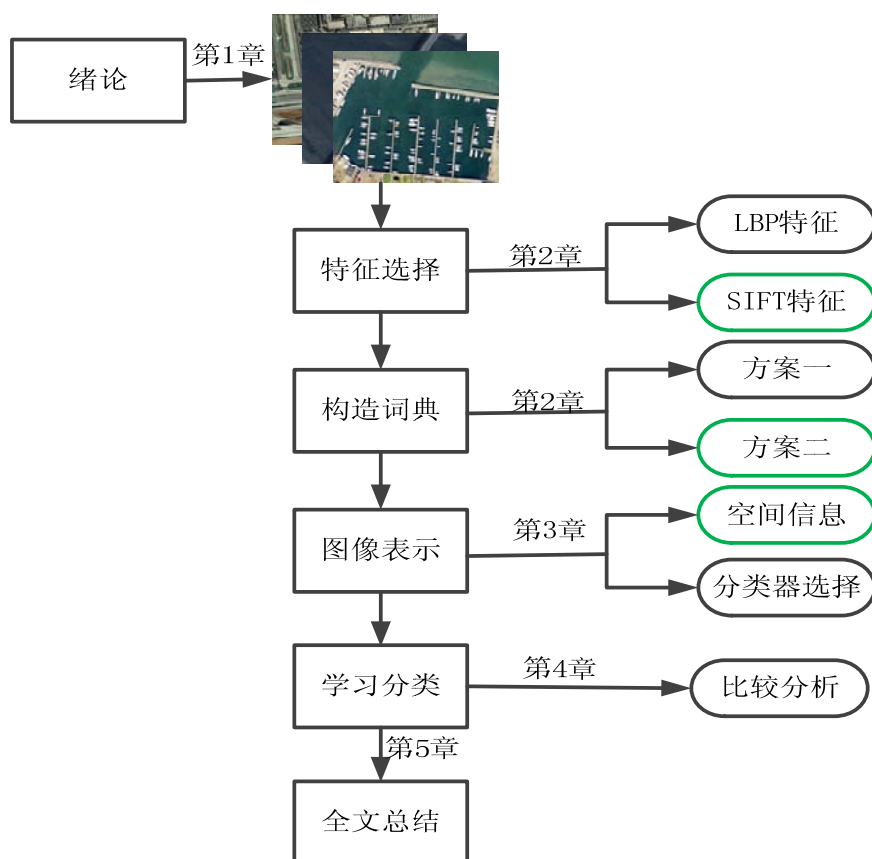


图 1.1 论文主要章节安排

第2章 图像特征选择及视觉词典构造

卫星图像场景分类过程与以往的图像场景目标识别有所不同，场景分类研究不需要对图像中的每一个具体目标进行完全的识别认知，而是对场景中整体区域各部分分布关系的语义表达。图像的局部语义概念是实现图像场景分类研究的一个重要模型，其优点是能够将高维的图像特征映射到低维的视觉词汇上，并对图像进行表示，有利于机器学习分类。其主要步骤包括：局部特征的选择，局部特征的提取，视觉词典的构造，图像的语义模型建模和图像特征向量的表达，机器学习方法进行学习和分类。

针对于上文提到的高分辨率卫星场景图像所呈现的特点，我们在选择特征对场景图像进行描述时，结合场景图像的特点有针对性地选择对图像进行描述，不能盲目的选择特征对图像场景进行描述。因此，特征的选择和提取是本章讨论和研究的第一个问题。

同时，在构造视觉词典时，有两种方案可供选择，如何选择应用能力比较好的构造词典的方案，直接影响最后的分类性能。因此，构造视觉词典时方案的设计和选择是本章要讨论和研究的第二个问题。

2.1 高分辨率卫星场景图像的特点

在场景分类中，高分辨率卫星图像主要呈现两大特点^[4]。第一，随着卫星图像分辨率的改善，场景图像中包含的地物目标所呈现的细节特点更加具体，例如结构、纹理、颜色等。第二，在图像中，同一场景类别包含同一对象所呈现的尺寸大小和方向位置可能不同，不同场景类别中可能包含同样的目标，不同天气，图像的亮度受光照影响较大。在图 2.1（a）表示的是停车场场景，在场景中，我们可以清晰的看到停车场中各个汽车停放的位置有所不同，各个车所呈现的颜色、尺度及停放的方向有所不同；在图 2.1（b）居民区和图 2.1（c）工业区两个类别场景中，建筑物是两个场景中共有的部分，同时也在两个场景类别中占主要地物目标部分。在多类别高分辨率卫星图像场景分类过程中，这些特点给图像的表达带来了很大的障碍。图像的表达是场景分类的一个关键环节，在低分辨率卫星图像分类中，纹理和强度特征就能有效地描述图像中的地物目标^[16, 17]。而由于高分辨率的卫星图像具有上述特点，单纯地利用这些

简单的特征不能充分表达图像的每个目标信息。因此，分类时需要结合特征的一些不变特性，例如尺度不变性，方向不变性，对比度不变性，特征的上下文语义信息，以及多特征间的组合等尽可能全面地描述各类场景。



图 2.1 场景样例

2.2 图像局部特征的选择

鉴于高分辨率卫星场景图像具有以上特性，所以我们在特征选择时，既要考虑特征能否就以上特点很好地描述场景图像，又要考虑特征在应用时是否有利于图像的分类。特征的选择是卫星图像场景分类的一个关键，具有代表性的特征能够更好地描述例如路面，天空，河流等一些突出的区域，本节就以下两种纹理特征做出讨论和对比，决定选择在实验中应用的特征。

2.2.1 局部二值模式（LBP）特征

局部二值模式（Local Binary Pattern）最早是由 T. Ojala 等人^[18, 19]基于局部

图像像素灰度值比较，用来描述图像局部纹理特征的算子。旋转不变性和灰度不变性是该算子所表现的两大主要特点。通过多年的研究与实验，已经验证了该特征在图像处理中可用于纹理的分割和分类^[20]，而且在人脸识别^[21, 22, 23, 24]等研究领域扮演重要的角色。求解 LBP 值的主要方法是首先固定图像中的某一个像素点，然后按照固定距离与其周围的其它像素点的灰度值进行比较，最后把比较后的数值按照规定的法则加权求和。例如，随机选择一个像素点为中心点，按照上述思想选择其周围固定距离的若干像素点，以中心点像素的灰度值为阈值，分别与其它相邻的像素点比较，比中心像素点灰度值大的点记为 1，小的点记为 0，把周围相邻各像素点的二进制值按照一定规律组合起来再化为十进制数，就是这个中心像素的 LBP 值。其中，LBP 算法可以定义为：

$$LBP_{R,N}(x,y) = \sum_{i=0}^{N-1} s(g_i - g_c) \times 2^i, \quad s(t) = \begin{cases} 1, & t \geq 0 \\ 0, & t < 0 \end{cases} \dots\dots\dots(2.1)$$

其中 $LBP_{R,N}(x,y)$ 代表坐标为 (x,y) 中心点 g_c 的 LBP 值，公式中 g_c 代表中心点的灰度值， g_i 代表距离中心点一定距离的相邻各点的灰度值， $s(t)$ 代表符号函数，当相邻各像素点的 g_i 值大于中心点 g_c 值时 $s(t)$ 置 1，反之为 0。权重 2^i 根据邻域内各点的位置，分配给各个点，然后转化为十进制求和，最后的十进制数即为该中心点的二值模式值。

$$\begin{array}{cccccccc} g_0 & g_1 & g_2 & 63 & 41 & 25 & 1 & 0 & 0 \\ g_7 & g_c & g_3 = 15 & 50 & 100 \Rightarrow 0 & & 1 \Rightarrow (00111001)_2 = LBP = 57 \dots\dots\dots(2.2) \\ g_6 & g_5 & g_4 & 28 & 53 & 74 & 0 & 1 & 1 \end{array}$$

最初计算 LBP 值的方法是在图像内选取窗口大小为 3×3 像素点，按照以上思路分别比较相邻的 8 个点与中心点的灰度值，从左上角的像素点开始，大于中心点灰度值的点标记为 1，小于中心点灰度值的点记为 0，依次顺时针进行，然后按照顺序将这 8 个二进制数进行组合转化为十进制数，这个十进制数值即为该中心点的 LBP 值，如公式(2.2)所示。需要指出的是在计算图像的局部 LBP 特征值时，各相邻的像素点之间存在较高的相关性，每一个像素点都要作为中心点，按照以上规则，重复以上步骤得到另一个点的 LBP 值。以此类推，可得到窗口内所有像素点的 LBP 值。由于 LBP 值的变化范围为 0~255，因此在统计窗口内各点 LBP 值的特征直方图时可以得到一个长度为 256 维的特征向量。

从上文可以看出，我们在计算各像素点的 LBP 值时，总是把中心点像素的

灰度值作为阈值，按照一定距离划分的其它相邻点的像素的灰度值与之比较，当图像的中局部区域的灰度值呈单调变化时，LBP 值的鲁棒性较好。但是对于复杂场景或者大场景分类时，该特征的维度过大，同时该特征不能充分地描述各图像场景特征，导致分类效果欠佳，因此，本文在进行卫星图像场景分类时没有采用该特征。

2.2.2 尺度不变特征转换（SIFT）特征

尺度不变特征变换^[25]（scale-invariant feature transform,SIFT）是一种特征描述符，主要用于描述图像局部纹理特征，在图像匹配中扮演重要角色，最早由 DAVID G. Lowe 提出。该算法通过求一幅图像中局部特征点以及该点的尺度和方向的描述子而得到特征，然后进行图像局部特征点的匹配。由于该特征具有尺度、方向及旋转度不变性，因此它具有独特性好、鲁棒性强的优点。目前该特征已广泛应用于图像场景识别、目标检测与识别等机器视觉领域^[26, 27]。SIFT 特征描述符提取的具体过程如图 2.1 所示：

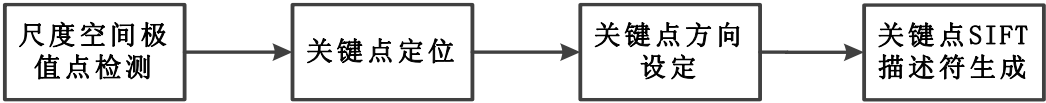


图 2.1 SIFT 特征描述符生成步骤

1. 尺度空间极值点检测

理论研究证明，高斯卷积核函数是一种独特的核函数，是实现图像尺度变换过程中的唯一一种线性核函数^[28, 29]。因此，变换图像的尺度可通过该核函数与图像进行卷积得到，其具体的表示式如公式(2.3)所示：

$$\begin{cases} L(x,y,\sigma)=G(x,y,\sigma)*I(x,y) \\ G(x,y,\sigma)=\frac{1}{2\pi\sigma^2}e^{-(x^2+y^2)/2\sigma^2} \end{cases} \dots\dots\dots(2.3)$$

其中， $L(x,y,\sigma)$ ——变换后的图像；
 $I(x,y)$ ——原始图像的矩阵， (x,y) 为像素点在图像中所处位置的坐标；
 $G(x,y,\sigma)$ ——尺度可变的二维高斯矩阵， σ 为服从高斯正态分布的方差，也是变换中的控制尺度变换的因子，尺度变换的大小直接影响图像平滑程度大小， σ 较小时，表示在变换中图像被平滑的程度也较小。

Lowe 使用了 $G(x,y,\sigma)$ 与图像进行卷积，建立差分高斯尺度空间（DoG

Scale-space)，用于关键点的稳定性优劣在该空间内进行验证。由于图像的平滑度，即图像特征的细节程度受尺度因子 σ 影响，因此可以通过 σ 乘以一个常数因子 k 来控制多个尺度，差分高斯尺度空间的各尺度的获取可以通过相邻的两个尺度之间相减得到：

$$\begin{aligned} D(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \\ &= L(x, y, k\sigma) - L(x, y, \sigma) \quad \dots\dots\dots(2.4) \end{aligned}$$

上式表明差分高斯函数 $D(x, y, \sigma)$ 可以通过相邻两个空间的图像进行不同尺度平滑，然后作差获取，这是下一步检测关键点的基础。

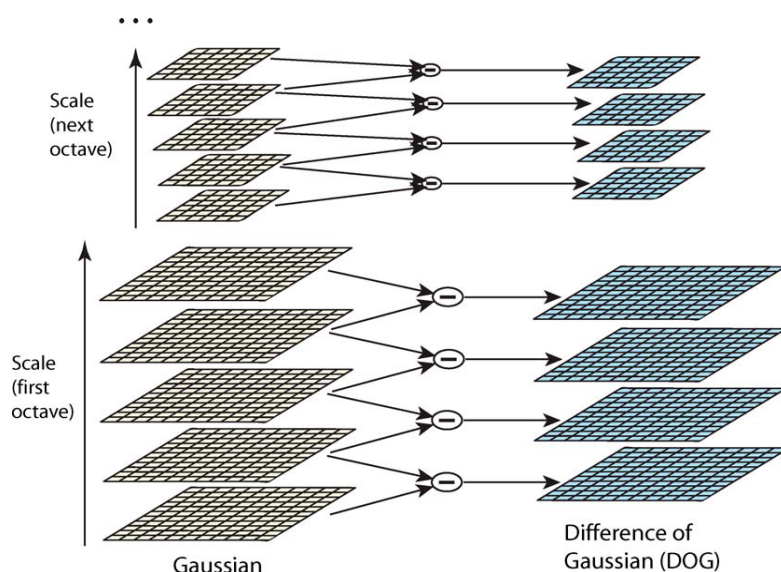


图 2.2 差分尺度空间的建立^[25]

差分高斯尺度空间的主要建立过程如图 2.2，对于每一级别的尺度空间的建立如左侧所示，每一级都是重复把原始图像与 $G(x, y, \sigma)$ 按照不同的尺度进行卷积，通过左侧相邻的两个卷积后的图像进行作差可得到右侧我们所需要的差分高斯图像，对于下一层低尺度级别的图像是上一层高尺度级别的高斯图像降 2 采样。

在确定关键点时，我们需要对各尺度空间的像素点分别采样进行比较，找出其存在的局部极值点。具体过程是选定某一个尺度内的差分图像，在该图像内随机选定某一个像素点，把该点与其周围同一尺度内相邻的 8 个点进行比较，与此同时，该点还需要与该点所处尺度相邻上下两个尺度上的 9×2 个像素点的值进行比较。如图 2.3 所示^[25]。当采样点的值大于或者小于所有比较像素点的值时，这个采样点就被确定。这样才能使得该采样点不管是在二维还是在

一维尺度空间都是局部极值点，即关键点。

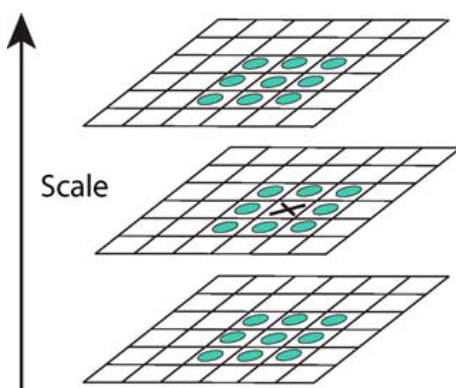


图 2.3 DoG 尺度空间局部关键点的确定

在关键点的选取时，关键点不仅要与同尺度内的相邻的其它点比较，还需要与该尺度空间相邻上下两个尺度空间内的对应的像素点比较。在图 2.2 中，有两组尺度变换后的图像，每组包含 5 层变换图像，把这 5 层通过相邻作差便得到了右侧所示的 4 层 DoG 图像，由于每组最上层和最下层的 DoG 图像不能进行极值点检测，所以只能在中间两个尺度进行。如果要检测 S 个尺度的极值点，每组的差分高斯尺度空间图像的层数为 $S+2$ ，左侧的原图像与高斯算子卷积得到的尺度空间图像的层数为 $S+3$ 。

2.关键点的定位

以上步骤是在离散空间取得的极值点，DoG 算子与图像卷积时会产生较强的边缘效应。为了消弱这种边缘效应带来的影响，更加准确检测到关键点所在的位置和尺度空间层，去除那些对比度相对较低和边缘响应不稳定的极值点，以此来增强图像匹配时特征的稳定性和特征在描述图像时的抗噪声能力。因此，我们需要用拟合三维二次函数进行处理^[30]，该函数的主要是对函数 $D(x,y,\sigma)$ 进行泰勒二次展开：

$$D(\mathbf{x}) = D + \frac{\partial D^T}{\partial \mathbf{x}} + \frac{1}{2} \mathbf{x}^T \frac{\partial^2 D}{\partial \mathbf{x}^2} \mathbf{x} \quad \dots\dots\dots (2.5)$$

其中 D 表示的是对所采样点的估计值， $\mathbf{x} = (x, y, \sigma)^T$ 表示确定精确关键点与原检测的极值点之间的偏移量。现对公式(2.5)求导，并令求导后的等式为零得到精确关键点的偏移量为：

$$\hat{\mathbf{x}} = \left(-\frac{\partial^2 D}{\partial \mathbf{x}^2} \right)^{-1} \frac{\partial D}{\partial \mathbf{x}} \quad \dots\dots\dots (2.6)$$

在关键点的位置和尺度精确定位后，由于高斯算子有很强的边缘响应，需

要剔除一些边缘响应不稳定的点和对比度相对较低的点，提高这些关键点的稳定性和抗噪声能力。对于一个定义欠佳的高斯算子的极值存在两个缺点，一个缺点是极值的主曲率在横跨边缘的地方比较大，另一个是极值的主曲率在垂直方向上较小。因此，在计算图像边缘关键点的主曲率时，在特征点处我们可以利用维数为 2×2 的 Hessian 矩阵 \mathbf{H} 来计算得到。

$$\mathbf{H} = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{yx} & D_{yy} \end{bmatrix} \dots\dots\dots (2.7)$$

从上式可以看出，矩阵中的各元素分别为差分高斯算子的偏导数，定义矩阵 \mathbf{H} 的两个特征值分别为 α 和 β ， α 表示高斯算子 x 方向的梯度和 β 代表算子，其中 α 的值大于 β 的值，同时令 $\alpha = r\beta$ ，则 Hessian 矩阵的迹的表达式和行列式值的表达式可通过两个特征值表示：

$$\begin{aligned} \text{Tr}(\mathbf{H}) &= D_{xx} + D_{yy} = \alpha + \beta \\ \text{Det}(\mathbf{H}) &= D_{xx}D_{yy} - (D_{xy})^2 = \alpha\beta \end{aligned} \dots\dots\dots (2.8)$$

由于高斯算子极值的主曲率值和 \mathbf{H} 矩阵两个特征值成正比关系，可以通过公式(2.9)来进行判定边缘的关键点是否是边缘响应强的关键点：

$$\frac{\text{Tr}(\mathbf{H})^2}{\text{Det}(\mathbf{H})} = \frac{(D_{xx} + D_{yy})^2}{D_{xx}D_{yy} - (D_{xy})^2} = \frac{(\alpha + \beta)^2}{\alpha\beta} = \frac{(r\beta + \beta)^2}{r\beta^2} = \frac{(1+r)^2}{r} \dots\dots\dots (2.9)$$

通过调节 r 的值控制主曲率是否小于 $(1+r)^2/r$ ，判定检测的关键点是否是边缘响应强的点。两特征值之间的比值越大， r 也越大，则反映了该特征点在某一方向上的梯度很大，在其它方向上的梯度较小，这正是边缘响应的特点。因此，可以定义 r 为某特定阈值，若该特征点的主曲率在该阈值之下，则这一特征点保留，否则，去除这一特征点。

3.关键点方向的设定

通过以上两步检测和筛选出需要的稳定的关键点，确定这些点的位置和尺度之后，需要对这些点的方向进行设定。我们可以采用图像采样特征点相邻的各像素点的梯度分布特性来确定方向，梯度的模值 $m(x, y)$ 和方向 $\theta(x, y)$ 。

$$\begin{aligned} m(x, y) &= \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \\ \theta(x, y) &= \arctan((L(x, y+1) - L(x, y-1)) / (L(x+1, y) - L(x-1, y))) \end{aligned} \dots\dots\dots (2.10)$$

上式中 L 为各采样像素点的尺度。然后统计与关键点相邻的各采用像素点

的梯度方向，各梯度方向可以通过直方图表示。直方图取值概括了采样点方向 360 度变化，把 10 度定义成一个等份，然后赋一个值，共计有 36 个取值。在统计的直方图中，频次出现最多的这个值所处的度数等份就是这个关键点的主要方向。若其它点的值大于峰值 80% 以上，则其它点所对应的方向即为辅方向。如此，一个关键点可能包括很多的方向，但是只有一个方向是主要的，其它的为次要的，提高了图像特征匹配的稳定性。

4. 关键点 SIFT 描述符生成

通过以上 3 个步骤，特征关键点的所在的位置、尺度和梯度方向都已确定，接下来是需要提取图像中各关键点的特征描述符。第一步是把坐标轴的方向调整为关键点所处的主方向，保证提取的特征描述符具有旋转不变性的特点。确定关键点所处的位置，在该点周围等距离取一个 16×16 像素的窗口，第二步在窗口内划分 16 个 4×4 像素的子窗口，接下来统计各个子窗口像素在 8 个方向上的梯度方向直方图，把这些梯度方向直方图统计起来组成一个 $16 \times 8 = 128$ 维的特征向量。第三步对该向量进行归一化处理，可以消除光照强弱对所提取特征的影响。提取特征时采用的这种方法能够改善该算法的抗噪声能力，可以改善稍有位置误差的局部特征点在匹配时的容错性能。如图 2.4 所示^[25]，左侧为一个 8×8 像素的窗口，右侧为每个 4×4 窗口内八个梯度方向的累加值。图 2.5 是对高分辨率卫星图像中的河流场景(a)和居民区场景(b)提取 SIFT 特征描述符的样例图。

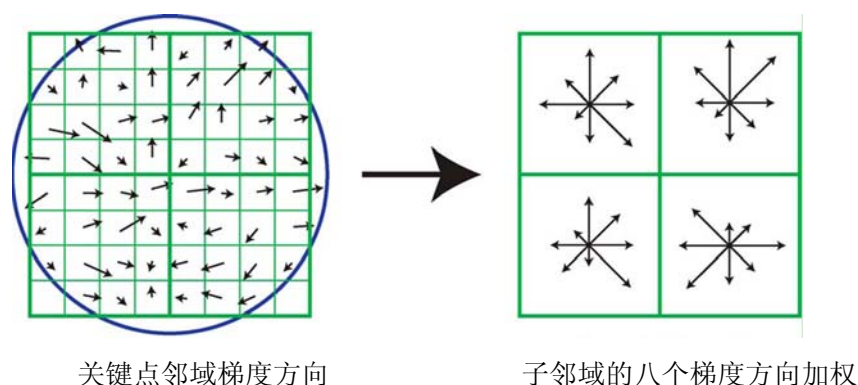


图 2.4 计算关键点描述符

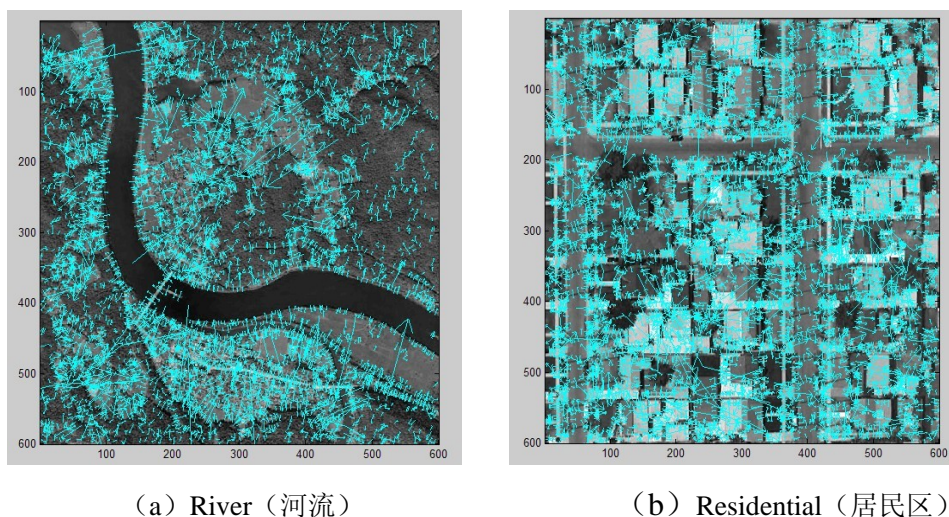


图 2.5 提取 SIFT 描述符样例图

综合以上所述，在对图像提取 **SIFT** 描述符时，以上算法的最大的优点就是具有不变性，同时对于图像受到视角变化、仿射变换、噪声干扰等具有较好的鲁棒性。该特征独特性好，信息量丰富，适合在大数据库中实现快速准确的匹配；多量性，即使少数几个物体也能提取丰富的描述符；可扩展性，能够与别的局部特征组合使用来描述表达图像。鉴于提取的 **SIFT** 特征描述符具有以上优点，因此本文对高分辨率的卫星图像分类研究中采用该特征描述符进行描述。

2.3 视觉词典的构造

视觉词典是图像视觉词包模型内的基本元素，本节在介绍视觉词典的构造之前要介绍视觉词包模型。视觉词包模型是最初从处理文档分类研究中的词袋模型发展演化而来的。

2.3.1 视觉词包模型

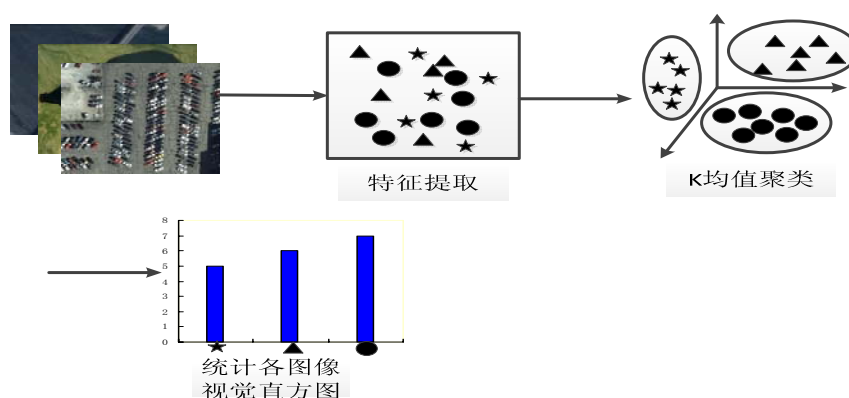


图 2.6 图像的视觉词包表示

2003 年, Sivic 等人^[31]最先将文本分析研究中的词包模型引入到视频处理研究中, 后来随着视觉词包模型在机器视觉领域的推广和应用逐渐衍生为视觉词包模型。在视觉词包模型的基础上, 研究人员利用该模型来实现目标的检测、识别及多场景的分类等研究^[32-43]。与文本处理中词包模型相似, 在图像场景分类研究中视觉词包模型可以看作是视觉单词的无序组合, 视觉单词相当于在图像中提取的局部特征通过聚类算法产生的一个聚类中心, 一幅图像可以看作是局部特征组成的一个集合。如图 2.6 所示, 视觉词包模型在图像场景分类中的应用主要包括提取图像的局部特征、利用聚类算法构造视觉词典、统计图像中各个视觉单词出现的频次组成一个向量, 该向量即为图像的视觉词包表示形式。

(1) 提取图像的局部视觉特征 局部视觉特征提取是描述各场景类别信息的基础, 也是利用词包模型对图像表示的首要工作。本文根据高分辨率的卫星场景图像的特点, 有针对性地选择了 SIFT 特征对图像进行描述。在文献^[44]中采用网格稠密采样比兴趣点采样在场景分类中的性能要好, 因此, 我们对训练和测试的卫星图像分别进行网格稠密采样, 进行 SIFT 特征的提取。

(2) 构造视觉词典 视觉词典的构造包括对局部特征进行量化、通过聚类算法把局部特征描述符聚类产生视觉单词, 所有无重复的视觉单词的集合即为我们所说的词典。局部特征提取后, 我们采用 K-means 聚类算法对提取的局部特征进行聚类, 其中每一个聚类中心视为一个视觉单词, 其中 K-means 算法将在下一小节中作详细讨论。每一个视觉单词都能够大致反映相似局部特征之间的共同属性。聚类中心的数目即为视觉词典的尺寸的大小。

(3) 图像的表达 视觉词典构造完成之后, 下一步工作是图像的表达, 在对局部特征构造视觉单词的过程中, 局部特征的提取来自于各图像。因此, 在描述各场景图像时, 可分别统计各视觉单词在词典中出现的频次, 依据频次统计每幅图像的直方图矢量。该矢量便是场景图像的视觉词包表示形式。

采用视觉词包模型对场景图像加以表示, 通过构造视觉词典, 图像的局部特征被量化为具体的视觉单词。该模型将文档分类的方法应用于图像场景的分类之中, 同时也很大程度上降低了视觉特征表示的维数, 将图像的局部特征很好地转化为中层语义对图像加以描述。但是, 该方法也存在一定的缺点, 视觉词典尺寸的大小受选择图像数据集样本数量多少的影响, 同时在对具体类别的图像描述时, 描述能力比较弱, 尤其是对场景类别比较相似的图像描述时独

特性比较弱。针对以上视觉词包模型的缺点，我们分析了描述和匹配场景图像时应该考虑的局部特征间的空间位置信息，本文的第二个主要工作是在高分辨率卫星场景图像分类中应用局部特征间的空间信息对图像加以描述。

2.3.2 K-means 算法

在构造视觉词典时，需要把提取的局部特征通过聚类算法将局部特征聚类产生视觉词汇，当前主要应用的聚类算法主要包括 K 均值聚类^[45, 46]、聚集聚类^[47]和 mean-shift^[48]等算法。由于 K 均值聚类算法的目的是各个样本与所在类的均值的误差平方和达到最小，能够对局部视觉特征集合进行合理的划分聚类，所以是构建视觉词典时应用较广泛的聚类算法^[49, 50]，本节将重点讨论高分辨率卫星图像场景分类研究过程中构造视觉词典的 K 均值算法。

该聚类算法是将 n 个待聚类的数据点通过不断迭代，将各数据点按照离聚类中心点距离最小的规则划分到 K 个聚类中心点上^[51]。称数据集合 $X=\{x_1, x_2, \dots, x_{n-1}, x_n\}$ 为聚类分析全体，集合中的元素 x_m ($m=1, 2, \dots, n$) 为一个样本，用一些参数的值来刻画某些样本所共有的某个属性。按照集合 X 中各样本所共有的属性特征进行 k 划分，样本集合应满足以下条件：

$$\bigcup_{i=1}^m x_i = X, x_i \cap x_j = \Phi, i \neq j \quad \dots\dots\dots(2.11)$$

从上式可以看出，各样本的集合是 X ，且各样本之间的交集为空。K 均值算法是将参数 K 作为输入，每一个聚类中心为一个类别，划分时类别内的各样本的相似度要满足类别间样本的相似性。该算法的基本思想是：第一步是在待聚类的样本对象中随机选择 k 个对象为待聚类的中心点，第二步计算其它剩下的数据点到已经选定的中心点的距离，离某中心点类别距离最小的样本对象归类到这个类别，第三步是在各类别中计算新的中心点，比较相邻两次的聚类中心有无变化，若有变化，重复这一过程，直到聚类准则函数达到收敛为止。主要算法描述如下：

(1)对于数据样本集合，随机选择 K 个样本对象点 $C_1(I), C_2(I), \dots, C_k(I)$ 作为 K 个聚类中心点，若对于各样本对象点的 k 次迭代中满足以下条件：

$$\|X - C_i(k)\| < \|X - C_j(k)\|, i, j = 1, 2 \dots k, i \neq j \quad \dots\dots\dots(2.12)$$

则 $X \in F_i(k)$ ，其中 $F_i(k)$ 表示聚类中心为 $C_i(k)$ 所在的类别。

(2)在各类别中获取新的聚类中心, 设 $F_i(k)$ 所在类别新的类别中心点为 $C_i(k+1)$, 计算被分在该类别中各个新的样本对象点的平均值为:

$$C_i(k+1)=\frac{1}{n_i} \sum_{X \in F_i(k)} X \quad \cdots \cdots \cdots (2.13)$$

其中 n_i 为类别 $F_i(k)$ 中具有样本的个数。

(3)若 $C_i(k)=C_i(k+1)$, 表明两次聚类中心的位置没有发生变化, 迭代终止。否则需要继续重复以上两个步骤。

2.3.3 构造视觉词典

在构造视觉词典的过程中有两种方案可供选择, 方案 1: 把所有数据集的图像进行分块, 从图像块中随机选择一部分子集进行特征提取, 构造视觉词典。方案 2: 只对训练集中的图像提取特征, 构造视觉词典。考虑到在实际的分类实验中, 待分类的图像的类别是未知的, 并且待分类的图像并未参与视觉词典的构造, 为了使实验结果更具有泛化能力, 构造视觉词典时我们选择方案 2 对图像数据集进行特征提取聚类构造视觉词典。

2.4 本章小结

本章讨论了三个方面的内容。第一是总结了在高分辨率卫星图像场景中, 不同的目标物体可能出现在不同的场景中, 同一场景类别中出现的目标物体可能又存在尺度、方向上的不同, 同时在拍摄卫星图像场景时受到气候的影响, 又使得各图像间的明亮程度有差异等等, 这些特点的存在都给场景分类中带来的困难。

第二是在图像中表示中, 由于高分辨率的卫星图像存在的特点, 特征的选择直接影响对场景图像描述的程度, 进而影响最终的分类结果。重点阐述了 LBP 特征和 SIFT 特征描述符生的相关算法, 同时也详细地讨论和比较了两种局部特征在描述图像场景时所具有的优缺点, LBP 特征的基本思想就是把图像中的某一中心点像素与其周围等距离内的像素点的灰度值进行比较, 各像素点之间的比较结果用二进制数 0 和 1 表示, 然后把这些二进制按照一定规律组合到一起, 最后的 LBP 值通过组合到一起的二进制数加权求和转化为十进制数表示, 该特征具有灰度不变性和旋转不变性的优点, 对图像中局部区域的灰度

变化特征值得鲁棒性较好，但是利用该特征对场景图像进行描述时，特征的维数过大，并不适合大场景图像的分类研究。而 **SIFT** 特征描述符具有独特性好，信息量大，对方向、旋转和尺度变化具有较强的不变性，鲁棒性好的优点，本文选择了该特征对场景图像进行表示。

第三是阐述了视觉词包模型在表示图像时的原理及思想，介绍在构建视觉词典时应用的 **K** 均值聚类算法。

第3章 卫星图像的空间信息和分类器选择

在高分辨率卫星场景图像分类中主要讨论分析了两个方面，第一是局部特征描述图像场景时空信息的引入；第二是在图像场景分类中训练和学习阶段分类器的选择。

由于高分辨率卫星图像所具有的特点，以及当前的一些研究方法在处理这些图像时未充分考虑这些特点，所以我们在上一章讨论研究有针对性地选择提取特征之后，本章将讨论在特征对图像的描述和匹配过程中，引入局部特征的空间位置信息来改善局部特征对场景图像的描述性能。在引入局部特征的空间信息过程中，讨论分析了空间金字塔模型的相关理论和算法。

本文选择 SVM（支持向量机）分类器主要是由于该分类器有两方面的优点，第一是该分类器在处理线性可分、不可分的问题时，在一定程度上能够避免“过学习”和“灾难维数”等问题。第二个优点是 SVM 的核函数种类较多，而且可以添加核函数工具箱，这为我们在实验中空间金字塔匹配核的应用提供了可能。

3.1 空间信息的引入

论文的主要工作是在卫星图像场景分类中引入空间信息对各图像场景进行表示。我们把空间信息引入到图像描述的过程中，主要出发点有：（1）因为空间信息对于地理数据的组成和分析具有重要的作用，早在 1970 年，Walter Tobler^[52] 在他地理学第一法则中就陈述了一切事物都有相关性，邻近事物的相关性要比距离远的相关性大。受这一法则启发，对图像场景进行描述时考虑了特征视觉词典的空间分布关系。为了利用局部特征间的空间位置信息，我们引入空间金字塔算法。（2）为了加强局部特征对图像的描述能力，改善卫星图像场景分类性能，引入局部空间位置信息对图像进行空间编码。

视觉词包被视为图像局部特征构造的视觉单词的无序组合，在图像表示时，无需考虑词包中各视觉单词的顺序和位置关系，但是在多场景和大场景的分类应用时，分类结果仍存在不尽如人意的地方。应用该模型中的视觉单词对图像描述时，只是单纯地停留在局部特征的层面上，没有考虑到局部特征之间存在的空间位置信息，使得局部特征在描述各场景图像时，不能充分描述各场

景图像的特点，最终导致对于大场景和多类别场景的分类结果欠佳。许多研究人员也相继提出了基于不同局部特征的生成模型应用于场景分类领域^[53, 54, 55]，2005 年，Lazebnik 等人^[56]和 Berg 等人^[57]提出了基于特征的几何对应的图像场景分类方法，同年，Grauman 等人在图像场景分类研究中提出了金字塔匹配核^[58]方法。该方法具有较优越的分类、识别性能，但是不适用于维度较高的特征集合。因此，2006 年，受金字塔匹配核思想的启发，同时为了在分类研究中利用局部特征的空间位置信息，Lazebnik 等人^[59]提出了空间金字塔匹配核的概念，在自然场景图像的分类研究中展示了良好的匹配性能。

在综合考虑了高分辨率卫星图像场景具有的特点以及局部特征对图像描述的局限性，在场景分类研究中，采用局部特征对图像进行描述表示时引入空间位置信息。

3.2 空间信息与匹配

在场景分类研究中应用局部特征的空间位置信息时，首先需要通过空间金字塔匹配模型对所提取的局部特征进行建模，空间金字塔建模的基本思路是局部视觉特征量化，空间金字塔划分，分别统计图像各子区域的视觉直方图，计算加权后各图像子区域的视觉直方图交叉。在分析空间金字塔匹配模型之前先讨论分析金字塔匹配模型。

3.2.1 金字塔匹配算法

金字塔匹配算法的方法原理就是在两个 d 维的特征空间点集 X 和 Y 中，按照不同分辨率，通过逐渐精细的网格划分，分别统计两个特征集合中各特征点划分到各网格中的直方图，然后计算两个特征集在不同分辨率中各网格中的直方图交叉，即同一网格中两个特征集合中重合的特征点数，最后对各个分辨率上网格内的局部特征的匹配点数进行加权求和来判别图像场景之间的相似度。

如图 3.1 所示^[58]， y 和 z 代表两个一维的特征空间集合，(a)表示两个特征空间点集按照不同的分辨率划分网格，特征点按照纵向分布，每行的宽度表示该行所在的尺度。(b)代表在不同尺度的划分下，两个特征空间在各网格中的特征点数，然后统计特征直方图， $H_0(y)$ 、 $H_1(y)$ 、 $H_2(y)$ 表示 y 特征空间点集在各尺度下统计的直方图， $H_0(z)$ 、 $H_1(z)$ 、 $H_2(z)$ 表示 z 特征空间点集在各尺度下统计的

直方图。(c)图表示两个特征点集在各分辨率下统计的直方图交叉，即特征点在各尺度下匹配时重合的点数， $\min(H(y)H(z))$ 表示在各尺度下，两个特征空间点集匹配的特征点数。最后两个特征空间点集的相似度通过各尺寸下的匹配点数的加权得到。

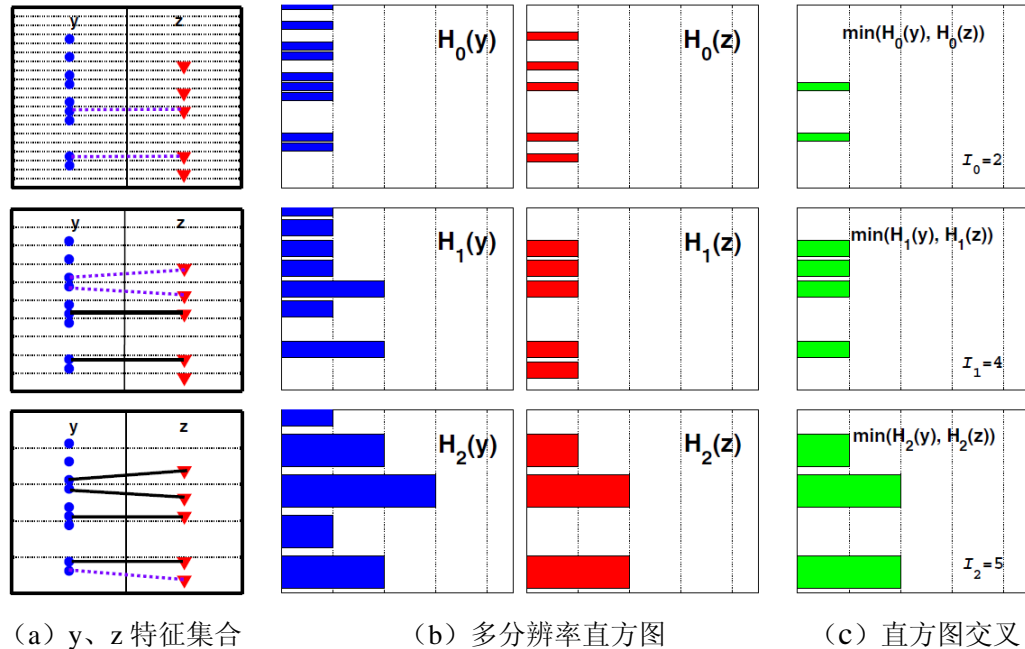


图 3.1 金字塔匹配

Grauman 和 Darrell 提出了金字塔匹配算法，目的就是为找到两个特征点集近似对应。在金字塔模型中，假设 X 和 Y 表示两个 d 维的特征空间点集，图像的特征空间按照不同的分辨率划分层次，然后在每个层次的图像进行逐渐精细的规则网格划分，比如在第 $l(l=0,1, \dots, L)$ 层,每一维特征有 2^l 个规则网格，那么 d 维特征空间总共有 $D=2^{dl}$ 个网格，然后统计每个网格上的特征直方图，把这些直方图连接起来就是图像金字塔的特征表达。 X,Y 两特征空间点集的相似度通过各层的匹配值加权得到，在特征匹配时，第 l 层的匹配值可以通过直方图交叉核函数^[60]计算：

$$I(H_X^l, H_Y^l) = \sum_{i=1}^D \min(H_X^l(i), H_Y^l(i)) \quad \dots\dots\dots (3.1)$$

其中 $I(H_X^l, H_Y^l)$ 表示直方图相交函数，下文简记为 I^l ， H_X^l 和 H_Y^l 表示特征空间 X,Y 在 l 层划分到第 i 个网格内的特征空间点数， D 表示各层加权后网格数。

由于在 l 层中存在的匹配点在其较细划分的 $l+1$ 层中也同时存在，因此在 l

层新得到的匹配点数可以通过 $I^l - I^{l+1}$ ($l=0, \dots, L-1$) 给出。特征空间划分越细时特征间的相似程度越高，所以需要为各层设置一个权重，权重设置为 $1/2^{L-l}$ ，权值的大小与各层中各网格的宽度成反比，图像的相似度通过金字塔匹配核函数（见公式（3.2,3.3））把各层匹配点数加权求和得到：

$$k^L(X, Y) = I^L + \sum_{l=0}^{L-1} \frac{1}{2^{L-l}} (I^l - I^{l+1}) \quad \dots\dots\dots(3.2)$$

$$= \frac{1}{2^L} I^0 + \sum_{l=1}^L \frac{1}{2^{L-l+1}} I^l \quad \dots\dots\dots(3.3)$$

3.2.2 图像的空间金字塔匹配算法

词包模型表示图像时，词包内的局部视觉特征之间是无序的，在表示图像时不用考虑特征间的位置顺序。金字塔匹配算法是特征空间之间匹配的一种算法。图像的空间金字塔匹配算法主要借鉴了以上两种算法的思想，该算法是词包模型的一种拓展。为了利用局部特征描述在各场景图像时能够考虑到局部特征间匹配时的空间位置信息，加强局部特征的描述能力，Lazebnik 等人将金字塔匹配模型应用于图像的二维空间，在特征空间仍采用传统词包模型对图像场景表示。

空间金字塔匹配算法主要包括局部特征的聚类，特征向量量化，同类视觉特征的匹配。在匹配的过程中只有类型相同，并且位置对应的特征之间才能匹配，这样局部特征间的空间信息被引入到图像的表达和匹配中。

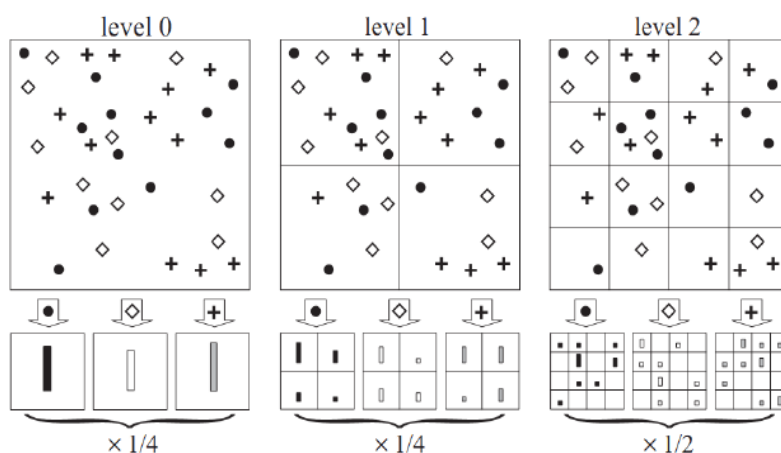


图 3.2 三层空间金字塔匹配

图 3.2 所示^[59]为图像的三层空间金字塔匹配模型，首先对图像空间按照不

同的分辨率尺度下划分层次 $l(l=0, \dots, L-1)$ ，在每个层次上对图像划分规则网格，网格数量为 $2^l \times 2^l$ ，统计各视觉单词在各尺度下各网格内的视觉直方图。

假设构造的视觉词典的尺寸为 M ，二维矢量集合 X_m 和 Y_m 表示在各自场景图像中视觉单词 m 的坐标位置，结合以上(3.1)、(3.2)、(3.3)三式，最终的空间金字塔匹配核函数为：

$$K^L(X, Y) = \sum_{m=1}^M k^L(X_m, Y_m) \quad \dots\dots\dots(3.4)$$

在实际的实验过程中，除最终实验结果的准确率是我们考虑的一个重要因素之外，计算的复杂度也是需要我们考虑的一个重要因素。核函数的计算复杂度同特征空间的维数成线性关系^[58]。通过以上算法的原理分析可知，影响特征空间维数多少的关键因素主要有两个：(1)金字塔的层数；(2)特征聚类视觉词典时词典的尺寸。最终特征空间的维数总和可定义如下：

$$M \sum_{l=0}^L 4^l = M \frac{1}{3} (4^{L+1} - 1) \quad \dots\dots\dots(3.5)$$

假设 $M=300, L=3$ 时，正交直方图的维数为 25500， $L=2$ 时，维数降为 8200。因此在实验中要兼顾两个因素对实验过程的影响。

通过分析研究得出通过空间金字塔匹配模型的建模，场景图像匹配的过程中利用局部特征对各图像描述，即利用了局部特征自身的表达信息又利用了局部特征间的空间位置信息。通过分层划分和网格划分，匹配时只有同类特征间才能匹配，将特征的空间位置信息融入到图像的表达之中。在最后的分类之中，还要结合 SVM 分类器对训练的学习来实现最终的分类实验。

3.3 分类器的选择

分类过程其实就是机器学习的整个过程，通过输入训练样本数据，学习数据的特点并掌握规律，进而在测试数据集上能够通过学习过程，在测试时对输出结果能够尽可能作出预判。因此，对于分类器的选择需要结合前文中实验的一些特点。首先，在学习和分类过程中，高分辨率卫星场景图像匹配时局部特征的维数受视觉词典的尺寸、空间金字塔的层数以及实验数据集大小等因素的影响，所以局部特征的维数不固定，有可能很大或者比较小。所以在学习和测试的过程中，分类器分类性能受样本集、特征的维数影响不大。其次，由于空

间金字塔匹配模型的应用，在分类中需要利用空间金字塔匹配核函数对各分辨率上的匹配点进行加权、学习和分类。所以在分类器上可以灵活添加自定义的核函数工具箱。

SVM（支持向量机）算法最早由 Vapnik^[61]提出的。SVM 算法最主要的优点是能够对小样本数据、非线性数据和高维特征空间进行训练统计和预测，在一定程度上解决了学习过程中的“灾难维数”和“过学习”等问题，同时具有多种分类学习核函数供选择，而且可以自定义添加核函数工具箱。因此本文选择了 SVM 分类器。

3.3.1 SVM 分类算法

目前 SVM 分类算法广泛应用于文本分类、图像场景分类、语音识别等领域。最初，SVM 分类算法主要针对数据集线性可分问题提出的，在处理线性不可分的数据集时，其主要思想是通过非线性映射方法，将线性不可分数据映射到高维的特征空间，使其在高维特征空间内线性可分，最后使用线性可分的方法对其实现分类。

如图 3.3 所示，图中的实心圆和空心圆表示两类训练数据样本，可以看出空心圆分布在平面 H_1 的附近，实心圆分布在平面 H_2 的附近， H 是把实心圆和空心圆分开的最优分类平面， H_1 、 H_2 平行于 H 为平行分类面， H_1 、 H_2 之间的距离为分类间隔。分类中，为了能把两类训练样本点尽可能准确地分开，一定范围内，还要求分割间隔越大越好，这就要求在求分类面的解唯一，从而达到在测试样本上的泛化能力最优。主要通过两类样本数据之间实现最小经验误差的同时最大化几何边缘来实现样本的正确分类^[62, 63]。

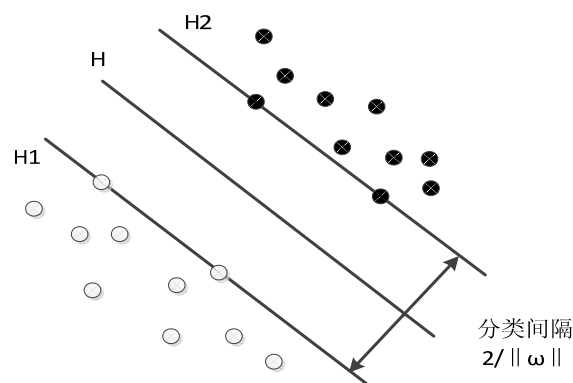


图 3.3 SVM 分类示意图

给定训练样本集 $S=((x_i, y_i)), i=1, 2, \dots, l, x_i \in R^n, y_i \in \{-1, 1\}$, 定义分类面为 $(\mathbf{w} \cdot \mathbf{x}) + b = 0$, 为了使分类面对所有样本数据点能够准确分类且具有最大的分类间隔, 因此, 样本数据点需要满足以下条件:

$$y_i = \begin{cases} 1, \mathbf{w} \cdot \mathbf{x} + b \geq 1 \\ -1, \mathbf{w} \cdot \mathbf{x} + b \leq -1 \end{cases} \Leftrightarrow y_i [(\mathbf{w} \cdot \mathbf{x}_i) + b] \geq 1 \quad \dots\dots\dots(3.6)$$

可以看出分类面一侧的点到最优分类面最大距离为 $1/\|\mathbf{w}\|$, 最大的分类间隔为 $2/\|\mathbf{w}\|$, 因此要构造最优分类面就要满足以下条件:

$$\min \Phi(\mathbf{w}) = \frac{1}{2} \|\mathbf{w}\|^2 \quad \dots\dots\dots(3.7)$$

$$\text{Subject to } y_i [(\mathbf{w} \cdot \mathbf{x}_i) + b] \geq 1 \quad \dots\dots\dots(3.8)$$

求解(3.7)式中最小值其实与求解 $1/\|\mathbf{w}\|$ 的最大值等价, 不难看出当 $\|\mathbf{w}\|$ 达到最小时, (3.7)式也达到最小值。之所以如此构造, 是为了以下求解方便。由于目标函数是严格上凹的二次型函数, 约束条件与 \mathbf{w} 成线性关系, 因此这是一个凸二次规划问题, 根据最优化理论, 可以把该问题转化为 Wolfe 对偶形式^[64]来求解, 需要构造拉格朗日函数:

$$L(\mathbf{w}, b, a) = \frac{1}{2} \|\mathbf{w}\|^2 - \sum_{i=1}^l a_i y_i [(\mathbf{w} \cdot \mathbf{x}_i) + b] + \sum_{i=1}^l a_i, a_i \geq 0, i=1, 2, \dots, l. \quad \dots\dots(3.9)$$

其中 a_i 是拉格朗日乘子, 将 (3.9) 式分别对 \mathbf{w} 和 b 求偏导数并令其等于零得:

$$\mathbf{w} = \sum_{i=1}^l a_i y_i \mathbf{x}_i \quad \dots\dots\dots(3.10)$$

$$\sum_{i=1}^l a_i y_i = 0 \quad \dots\dots\dots(3.11)$$

(3.9) 式中 y_i 其实就是分类问题样本所标记的类别标签, a_i 是拉格朗日乘子, 实际在这些乘子中只有很少一部分算子不等于 0, 这部分不等于 0 的算子乘以各样本点的类别标签, 就是这部分样本点落在了 H_1, H_2 上, 这些样本点也确定了分类函数的形式。这部分样本点就被称为支持向量。把(3.10)与(3.11)式带入到(3.9)式中, 消去 \mathbf{w} 和 b , 求解最优解的问题就转化为相应的对偶形式:

$$W(a) = L(\mathbf{w}, b, a) = \sum_{i=1}^l a_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l a_i a_j y_i y_j (x_i \cdot x_j) \quad \dots\dots\dots(3.12)$$

$$\text{Subject to } \sum_{i=1}^l a_i y_i = 0, a_i \geq 0, i=1, 2, \dots, l. \quad \dots\dots\dots(3.13)$$

Wolfe 对偶形式问题的解便是我们在处理原求解最优化问题所要求的解，解出拉格朗日乘子 \mathbf{a} 后，带入到(3.10)式得到权重 \mathbf{w} ，最后得出最优分类平面的方程为：

$$\sum_{i=1}^l a_i y_i (x_i \cdot x) + b = 0 \quad \dots\dots\dots(3.14)$$

3.3.2 SVM 几种常见核函数

在处理线性不可分的问题时，我们不能直接应用解决线性可分问题的方法来处理，而是通过非线性映射的方法把线性不可分的数据样本点转化为在高维特征空间线性可分的数据样本点。这种映射需要借助满足一定条件的核函数来实现。首先我们要找到一个高维特征空间，找出一个函数 K ，在低维空间向高维空间映射时使得 $K(x_i, x_j)$ 等价于样本点 x_i, x_j 在高维空间的内积，这样就可以用 $K(x_i, x_j)$ 代替在 Wolfe 对偶问题求解时 x_i, x_j 的内积形式，如公式(3.15, 3.16)所示。优点是大大减少了分类过程中的计算量^[65]。这类核函数称之为卷积核函数^[66]。

$$\text{Max } W(\mathbf{a}) = L(\mathbf{w}, b, \mathbf{a}) = \sum_{i=1}^l a_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l a_i a_j y_i y_j K(x_i \cdot x_j) \quad \dots\dots\dots(3.15)$$

$$\text{Subject to } \sum_{i=1}^l a_i y_i = 0, a_i \geq 0, i = 1, 2, \dots, l. \quad \dots\dots\dots(3.16)$$

只要是满足了 Mercer 条件的函数，都可以作为核函数。这类核函数的作用就是能够计算出低维空间里的线性不可分的样本数据点向量映射到高维特征空间后的样本数据点向量的内积值。几个比较常用的核函数如下：

(1) 线性核函数：

$$k(x_i, x_j) = (x_i \cdot x_j) \quad \dots\dots\dots(3.17)$$

(2) 多项式核函数：

$$k(x_i, x_j) = [(x_i \cdot x_j) + b]^p, \quad b \geq 0, \quad p \text{ 是自然数} \quad \dots\dots\dots(3.18)$$

此时把该核函数带入上式，便得到了称之为 p 阶多项式分类器，当 $b=0, p=1$ 时，该多项式核函数便转化成为线性核函数。

(3) 径向基 (RBF) 核函数：

$$k(x_i, x_j) = \exp\left(-\gamma \|x_i - x_j\|^2\right) \quad \dots\dots\dots(3.19)$$

这是一个径向基函数的支持向量机，RBF 核函数可以将一个低维的样本数据映射到更高维的空间，可以处理当类标签和特征之间的关系，是非线性分类的样例。

(4) Sigmoid 核函数:

$$k(x_i, x_j) = \tanh[k(x_i \cdot x_j) + c] \quad \dots\dots\dots(3.20)$$

支持向量机算法采用边缘最大化原则，在利用最优分类面来构造判决函数时，避免了在训练样本数据集上学习时产生过拟合现象，这是该算法的优点之一，也是其应用比较广泛的主要原因。在处理非线性问题时，可以通过采用核函数映射的方法把低维不可分的样本数据点映射到高维特征空间，使其在高维特征空间线性可分，再应用处理线性可分的思想，这也是其泛化能力较强的一个主要因素。

3.4 本章小结

本章分析了两个方面的内容。第一是引入局部特征的空间位置信息对高分辨率卫星场景图像进行描述。结合了当前的一些研究方法和图像存在的特点，为引入局部特征间的空间位置信息提供了理论铺垫，分析了空间金字塔匹配模型的相关算法和原理。

第二是分类器的选择。在选择分类器时分析了高分辨率卫星图像场景分类研究中存在的两个问题：（1）图像匹配中局部特征的维数大小依赖于视觉词典的尺寸、空间金字塔的层次和数据集数量大小等因素，特征维数可能很大或者比较小。（2）实验中引入了一种新的核函数，该分类器应具有自定义添加核函数工具箱的特点。由于 SVM 的决策函数的判定只是由若干样本数据向量中少数的样本数据向量决定，而不取决于样本空间大小。因此，在处理高维数据集时，避免了计算过程中的“维数灾难”。同时 SVM 分类器的核函数种类较多，而且易于自定义添加满足一定条件的新的核函数工具箱。因此，本文选择 SVM 分类器用于分类研究。

第 4 章 分类过程及结果分析

本章讨论卫星场景图像分类的实验过程和实验结果分析。主要包括数据集的获取，实验组成部分，实验过程中的参数设置与运行平台等。

以上步骤主要目的是通过具体的实验来验证本文做的三个方面工作的有效性。（1）有针对性地提取场景图像的局部特征；（2）构造视觉词典时选择方案的泛化能力；（3）通过引入局部特征的空间位置信息改善对场景图像的描述。

在不同参数设置下的实验获得相对优越的分类结果，并且同其它方法所获得的分类实验结果进行比较和分析得出了本文提出方法的优缺点。实验结果证明了以上主要工作的有效性。

4.1 实验基本流程

高分辨率卫星场景图像分类实验的基本组成如图 4.1 所示。首先我们把图像数据集分成两个部分，分别为训练集和测试集，在训练集和测试集上提取 SIFT 特征描述符，利用训练集提取的特征描述符通过 K-means 聚类算法构造视觉词典。然后分别对训练集和测试集进行空间金字塔建模（SPM 编码），分别统计各层各网格内视觉词典的直方图，把训练集图像的特征向量送入分类器学习，最后把已训练好的分类器对测试集图像进行分类。

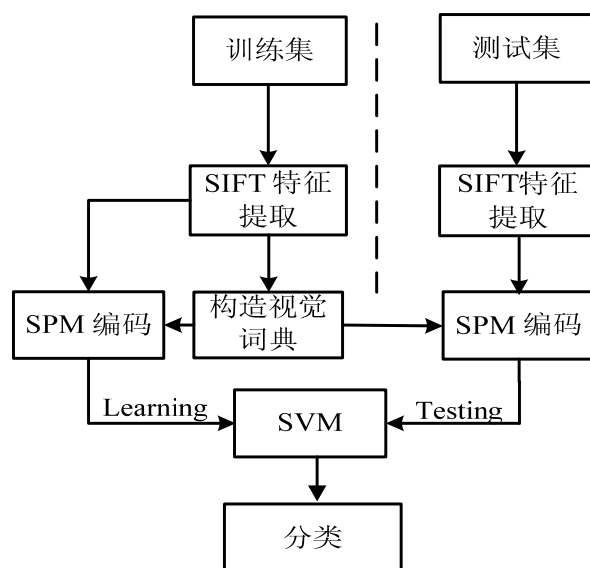


图 4.1 实验流程图

4.2 数据集和参数设置

实验中采用的高分辨率卫星图像数据集来自于文献^[6,9](原始数据集的获取均来自 Google Earth)。该数据集包括 12 个场景类别,分别为飞机场、桥梁、商业街、森林、工业区、牧场、停车场、池塘、港口、住宅区、河流和高架桥。所有的这些场景类别都是从大场景图中提取出来的。一些典型的场景类别图像如图 4.2 所示。每一类场景包括 50 幅图像,图像大小统一为 600 像素×600 像素。由于每类图像样本的拍摄受天气、气候、地域的影响,因此同类别场景的卫星图像在纹理,颜色,亮度上会有一些差别。在实际的实验分类过程中,这些特点给分类带来很大的困难,也是影响分类结果的因素。

实验过程中对于每一类别的场景图像,随机选择 5,10,15,20,25 幅图像作为训练数据集,其余的作为测试图像。实验运行的平台是 R2012a MATLAB 平台,分类器采用的是 LIBSVM 工具箱^[8],每组实验重复进行 5 次,然后统计每个类别的分类准确率,计算每组每次实验中 12 个类别的平均分类准确率。每组每次实验中各个类别的分类准确率计算公式是:

$$\text{准确率 } (R_{ac}) = \frac{\text{正确分类样本数}}{\text{测试样本数}} \times 100\% \quad \dots\dots\dots(4.1)$$

$$\text{平均分类准确率 } (R_{av}) = \frac{1}{n} \sum_{i=1}^n y_i \times R_{ac}^i \quad \dots\dots\dots(4.2)$$

其中 R_{ac} 表示每次实验中每个类别的分类准确率, R_{av} 表示每次实验后该组实验的平均分类准确率, n 表示类别数, y_i 表示第 i 类的类别标签。

Li 等人^[44]证明了采用网格稠密采样比兴趣点采样在场景分类中的性能要好,因此,我们对训练和测试的卫星图像分别进行网格稠密采样,进行 SIFT 特征的提取。对每一个采样点提取周围的 128 维 SIFT 特征。图 4.3 所示为提取的 128 维 SIFT 特征描述符通过 K-means 聚类算法聚类后所获得的视觉词典矩阵,视觉词典的尺寸为 300。图 4.4 是引入空间信息后,空间金字塔建模后图像存储的矩阵,矩阵的维数是 6300×600,金字塔的层数为 3 层,对每层进行规则网格划分,加权各层的网格数共 21 个,由于视觉词典的尺寸为 300,图像数据集总得图像数量为 600 幅,所以空间金字塔建模后的图像表示矩阵维数为 6300×600。



图 4.2 各图像场景的样例，每个场景类别 3 张

	1	2	3	4	5		124	125	126	127	128
1	0.0428	0.0456	0.0435	0.0394	0.0383		0.0667	0.0239	0.0172	0.0204	0.0307
2	0.0317	0.0320	0.0498	0.0669	0.0714		0.0157	0.0205	0.0226	0.0318	0.0815
3	0.0256	0.0382	0.0794	0.0589	0.0438		0.0196	0.0318	0.0679	0.1438	0.0796
4	0.0508	0.0420	0.0362	0.0417	0.0423		0.0292	0.0361	0.0327	0.0269	0.0256
5	0.0812	0.0525	0.0344	0.0293	0.0361		0.0625	0.0935	0.0537	0.0295	0.0297
6	0.0074	0.0119	0.0728	0.2434	0.0682		0.2401	0.0650	0.0098	0.0069	0.0112
7	0.0477	0.0481	0.0396	0.0305	0.0296		0.0170	0.0210	0.0227	0.0391	0.1139
8	0.0246	0.0303	0.0419	0.0195	0.0179		0.0200	0.0196	0.0234	0.0582	0.0263
9	0.0983	0.0477	0.0267	0.0219	0.0208		0.0496	0.0657	0.0655	0.0569	0.0500
10	0.0902	0.0310	0.0181	0.0152	0.0217	0.0326	0.0321	0.0307	0.0322	0.0321
.....										
291	0.0401	0.0439	0.0340	0.0382	0.0824		0.0486	0.0373	0.0294	0.0253	0.0325
292	0.0392	0.1091	0.1878	0.0993	0.0306		0.0432	0.0402	0.0335	0.0340	0.0367
293	0.0447	0.0439	0.0238	0.0204	0.0286		0.0292	0.0843	0.0816	0.0259	0.0238
294	0.0325	0.0213	0.0224	0.0330	0.0338		0.0346	0.0249	0.0228	0.0297	0.0575
295	0.0352	0.0292	0.0356	0.0787	0.1296		0.0277	0.0353	0.0380	0.0378	0.0305
296	0.0185	0.0119	0.0185	0.0096	0.0176		0.0117	0.0196	0.0139	0.0189	0.0661
297	0.0523	0.0275	0.0216	0.0251	0.0430		0.0391	0.0464	0.0405	0.0344	0.0311
298	0.0182	0.0576	0.1515	0.0663	0.0294		0.0651	0.0243	0.0204	0.0260	0.0135
299	0.0491	0.0357	0.0322	0.0264	0.0344		0.0387	0.0450	0.0251	0.0243	0.0302
300	0.0274	0.0558	0.0453	0.0217	0.0238	0.0299	0.0190	0.0381	0.0287	0.0131

图 4.3 视觉词典矩阵

	1	2	3	4	5		596	597	598	599	600
1	0	3.6523e-04	9.1308e-05	1.8262e-04	0		1.8262e-04	3.6523e-04	3.6523e-04	9.1308e-05	1.8262e-04
2	0	9.1308e-05	9.1308e-05	9.1308e-05	9.1308e-05		9.1308e-05	1.8262e-04	3.6523e-04	1.8262e-04	1.8262e-04
3	9.1308e-05	9.1308e-05	0	1.8262e-04	1.8262e-04		1.8262e-04	2.7392e-04	9.1308e-05	1.8262e-04	2.7392e-04
4	0	0	1.8262e-04	0	0		9.1308e-05	9.1308e-05	0	1.8262e-04	1.8262e-04
5	9.1308e-05	9.1308e-05	0	0	0		0	0	0	9.1308e-05	0
6	0	9.1308e-05	1.8262e-04	1.8262e-04	0		9.1308e-05	9.1308e-05	9.1308e-05	0	0
7	0	0	9.1308e-05	9.1308e-05	9.1308e-05		0	1.8262e-04	0	9.1308e-05	0
8	0	2.7392e-04	0	9.1308e-05	9.1308e-05		0	9.1308e-05	0	9.1308e-05	0
9	0	0	2.7392e-04	9.1308e-05	0		0	9.1308e-05	0	9.1308e-05	9.1308e-05
10	9.1308e-05	9.1308e-05	0	9.1308e-05	0	9.1308e-05	0	0	0	9.1308e-05
.....										
6290	6.3915e-04	0	4.1088e-04	4.5654e-04	3.1958e-04		3.1958e-04	9.1308e-05	2.2827e-04	2.7392e-04	1.3696e-04
6291	0.0061	0.0024	0.0030	0	0.0032		4.1088e-04	0.0013	0.0012	0.0011	6.8481e-04
6292	3.1958e-04	2.7392e-04	0.0016	5.4785e-04	4.1088e-04		8.6742e-04	0.0019	3.6523e-04	0.0012	0.0012
6293	3.1958e-04	2.7392e-04	5.9350e-04	4.5654e-04	4.1088e-04		6.3915e-04	5.9350e-04	5.0219e-04	9.5873e-04	1.8262e-04
6294	6.3915e-04	4.1088e-04	2.7392e-04	3.6523e-04	2.7392e-04		6.8481e-04	4.1088e-04	4.5654e-05	1.8262e-04	2.2827e-04
6295	4.1088e-04	5.0219e-04	2.7392e-04	4.5654e-05	0.0011		2.7392e-04	4.5654e-05	2.2827e-04	0.0012	5.4785e-04
6296	6.3915e-04	3.6523e-04	3.1958e-04	4.1088e-04	2.7392e-04		0.0016	0.0019	6.3915e-04	0.0017	0.0011
6297	8.2177e-04	6.8481e-04	0.0011	3.1958e-04	0.0018		2.2827e-04	0.0018	3.6523e-04	9.5873e-04	0.0021
6298	5.4785e-04	0.0011	0.0012	2.7392e-04	5.0219e-04		0.0016	2.2827e-04	0.0012	7.3046e-04	9.1308e-05
6299	2.7392e-04	2.7392e-04	7.3046e-04	0.0013	5.4785e-04		5.4785e-04	0.0021	7.7611e-04	0.0013	0.0018
6300	4.1088e-04	3.6523e-04	8.2177e-04	2.7392e-04	3.1958e-04					

图 4.4 空间金字塔建模后图像存储矩阵

4.3 分类结果分析

图 4.5 所示的是训练和测试集图像数量分别为 25 时，随机一组实验所获得的混淆矩阵，混淆矩阵的横轴和纵轴分别表示各个样本的类别，第 m 行第 n 列的值表示第 m 类样本被错分第 n 类样本的概率值，也可表示错分样本数。从图中我们可以清楚地看到每个测试样本类别在分类测试时，正确分类样本的个数和其它类别错分到该类别样本的数目。例如在飞机场这个测试样本中，被正确分类的样本数为 22，被错分为工业区的数量为 1，错分为停车场的样本数量为 2。对于商业区这一场景，正确分类的样本数为 21，错分为工业区的样本数为 3，错分为居民区的样本数为 1。

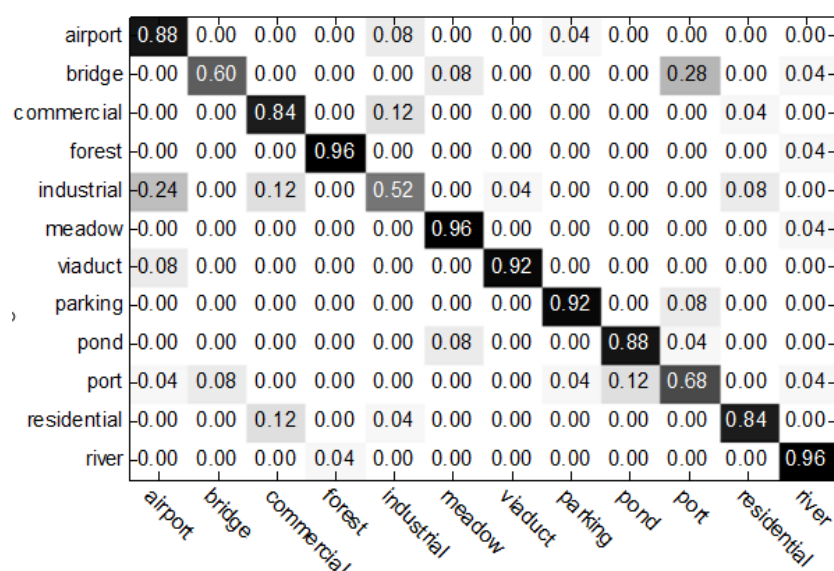


图 4.5 分类混淆矩阵

在利用 SVM 做学习和分类时，不同的核函数选择也会影响实验结果，对于有些问题，选择某些或者某个核函数分类效果会很好，选择其它核函数效果可能就欠佳。但是，一般情况下径向基核函数在学习分类过程中不会有太大偏差，因此，本文对于径向基核函数与空间金字塔匹配核函数所获得的实验结果做了对比。

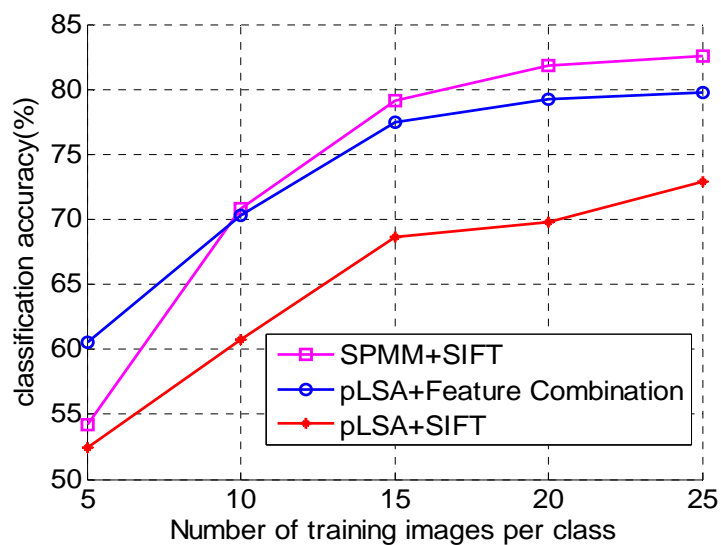
表 4.1 为两核函数所获得的分类准确率对比，可以看出通过空间金字塔建模，采样空间金字塔匹配核（直方图交叉核）作为分类的核函数所得到的分类结果更加优越。可以看出当训练图像数分别为 5,10,15,20 和 25 时，直方图交叉核函数的分类结果均优于径向基核函数。

表 4.1 两种核函数分类结果比较

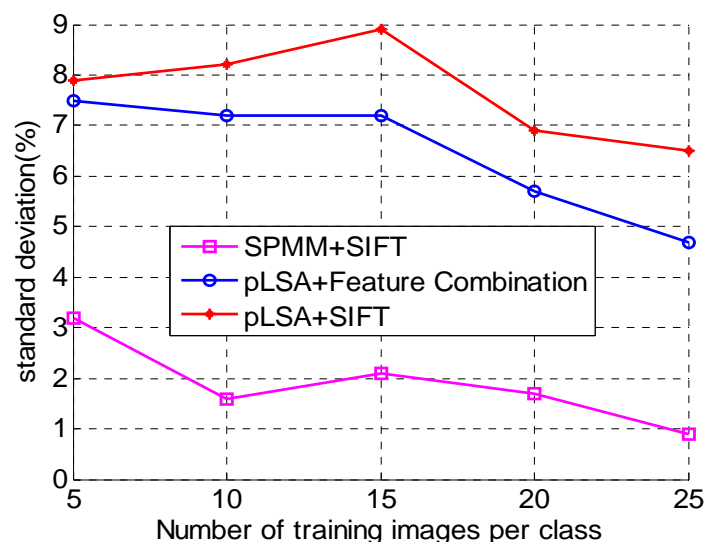
核函数 \ 图像数	5	10	15	20	25
直方图交叉核	54.2%	70.8%	79.1%	81.8%	82.6%
径向基核函数	32.4%	54.2%	62.8%	67.8%	69.7%

对卫星图像运用单特征的空间金字塔编码, 通过 SVM 学习分类, 分类实验结果如图 4.6 (a) 所示。同文献^[12]中运用单特征 (pLSA+SIFT) 及多特征组合(pLSA+multiple features combination)分别与 pLSA 主题模型相结合的分类实验结果相比, 当训练集数量较少时, 例如训练集的数量为 5, 所获得的分类性能较低, 主要原因是同一类别场景的卫星图像差别较大, 其它场景类别的图像又存在共同的目标物, 训练图像较少且使用的单特征编码, 对整个类别全部场景图像的描述能力较低, 不足以显现空间金字塔编码的优点。而 Xu *et al*^[6]采用的是多特征组合, 在图像描述层面上相对于单特征具有更强的描述能力, 并且分类时采用双层分类器, 因此, 训练集数量较少时, 多特征组合与主题模型相结合的方法对场景图像具有更好分类性能。随着训练集图像数量的增加, 分类准确率有所提高。主要原因是提取的特征描述符概括性在同类场景图像中较广, 对场景图像的描述能力增强, 同时空间金字塔编码的优点得以展示。当每类场景训练集数量为 25 时, 测试集的平均分类准确率达到 82.6%, 较文献^[6]提高近 2.8 个百分点。从以上数据分析可知在高分辨率卫星图像分类中, 基于单特征选择的基础上, 在图像表示和匹配的过程中, 空间信息的引入明显改善了分类性能。

图 4.6 (b) 给出了各实验方法分类准确率标准差的比较, 可以看出三种方法在分类研究中, 标准差的值随着训练图像数量的增加大方向上是逐渐减小的, 表明训练图像对所在类别的概括能力随着图像数量的增加逐渐增强的, 所以每组实验的结果也逐渐趋于稳定。从图中我们的方法所获得的标准差较其它两种方法有大幅度的降低, 在每组实验中, 各组实验按照一定数量随机选择训练图像数重复 5 次实验所获得实验结果之间差距不大, 结果相对稳定。可以看出, 空间信息的引入, 使得各组实验结果受随机抽取训练和测试数据集的影响较小, 鲁棒性较强。



(a) 分类准确率比较



(b) 标准差比较

图 4.6 分类准确率和标准差性能比较

图 4.7 给出了训练数据集图像的数量为 25 时，在重复 5 次该组实验后，各个场景类别分类的平均准确率的混淆矩阵，从中可以看到目标信息相差较大的不同场景类别所获得的分类准确率较高，如飞机场场景类别的平均分类准确率为 92%，森林场景类别的平均准确率为 92%，牧场场景类别为 96% 等，这是由于这些场景类别中，图像中所包含的目标信息与其它场景类别的目标信息差别较大，因此，在分类时不容易被分到其它类别。但是地物目标信息相近的不同场景容易相互分错，例如商业区场景类别平均分类准确率只有 72%，工业区场景的平均分类准确率为 72%，居民区场景的平均分类准确率为 80%，这三个场

景类别容易相互分错，是由于这 3 个场景类别中，建筑物是它们共有的主要目标物体。同时还以看到高架桥场景类别容易被分为港口场景类别，港口场景类别也容易被分为河流场景、池塘场景、高架桥场景，这是由于水域部分是这几类场景类别共有的目标信息。

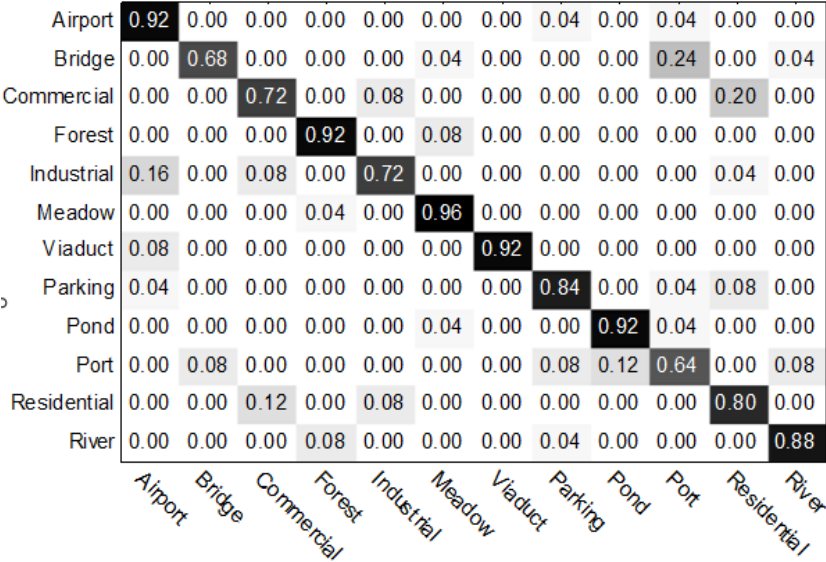


图 4.7 分类准确率的混淆矩阵

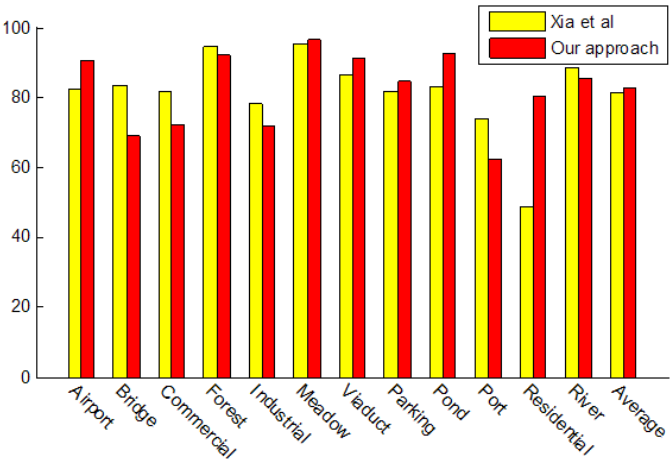


图 4.8 每一类别准确率对比

图 4.8 给出的是我们的方法同 Xia.et al^[11]方法取得分类结果的比较，Xia 在分类方法中主要利用了场景中目标的结构信息和纹理信息，利用双重信息对场景图像的描述。数据集中各类图像一半用于训练一半用于测试，差别较大的类别场景能获得良好的准确率，如飞机场、森林，牧场等场景类别，差别较小的所获得的准确率相对较低，如商业区、工业区、居民区等，特别是居民区这一类别，准确率为 48.87%，总体平均分类准确率为 81.6%。

而我们利用 **SIFT** 特征描述场景结构信息的同时，又结合了特征的空间位置信息。图中显示，各类分类准确率同 **Xia** 相比，差别较小的场景分类准确率有所下降，例如桥和港口这两类场景，这两类中水域占很大部分，商业区和工业区这两类场景中，建筑物占主要部分。差别较大场景准确率有所提高，如飞机场，牧场和高架桥等，同时，居民区这一类别所获得的准确率为 80.2%，条形图的最后显示我们在总体的平均准确率相对提高近 1 个百分点。以上各分类实验结果表明，在运用 **SIFT** 单特征的基础上，通过空间金字塔编码，引入局部特征的空间信息对各类场景图像进行描述，在图像的和匹配过程中增强了图像的描述能力，改善了分类性能。

第5章 全文总结

本文首先讨论了高分辨卫星场景图像的特点,包括(1)同一场景类别中的目标信息存在不同的方向、尺度和亮度等;(2)不同场景类别中又可能包含其它场景类别相似的目标信息,由于有的场景图像中目标在拍摄时受天气、气候、地域的影响,使得图像的亮度不同或者外观颜色差别较大。

其次,在分析高分辨率的卫星图像场景特点给分类研究带来困难的基础上,简述了当前研究卫星图像场景分类的几种主流方法,分析了这几种方法的优点以及所共有的弊端。接着,从弊端入手,提出本文研究工作的方法和理论依据。

研究工作主要从以下三个方面展开:(1)在图像描述的过程中,特征选择对于场景图像分类中图像描述的重要性;(2)构造视觉词典两种方案中,泛化能力较好方案的选择;(3)空间信息的引入对图像分类性能的改善等。论文中分析了两种局部特征算法,阐述和讨论了两种特征在实际分类研究中的优缺点;论文还分析了在构造视觉词典时运用的 **K-means** 聚类算法。论文的核心部分是在高分辨率卫星场景图像的表示和匹配过程中引入了局部特征间的空间信息,通过特征提取、视觉词典构造、空间金字塔建模、图像的表示以及最后分类器的训练和测试等过程,逐步把特征的空间信息引入到分类研究中。最后对论文中方法的分类结果和其它实验方法的分类结果做了对比分析。

本文提出了利用特征空间信息的高分辨率卫星图像场景分类的方法。采用空间金字塔模型对图像的 **SIFT** 特征进行空间编码,匹配时,通过逐层进行相似点的匹配和加权求和,获得最终的匹配值,利用 **SVM** 分类器对各场景类别的图像进行学习分类,获得了 82.6% 平均分类准确率。表明该方法能够显著增强局部特征对各场景图像的描述能力,打破了在高分辨率卫星图像分类研究中单纯依赖在局部特征层面上对各场景描述,实现场景分类这一局面,提供了高分辨率卫星图像分类研究的方法。

在场景分类中,虽然总体分类性能较文献^[6,11]有所改善,但是场景类别相近的各场景容易相互分错,而且地物目标特征相似的场景类别在分类时相互分错的可能性相对于差异较大的其它场景类别依然较大。对相似场景之间的识别能力有所欠佳,为了改善这一状况,下一步可以尝试采用多特征组合的空间金字

塔编码，通过结合其它特征的空间位置信息来描述图像场景，将会在一定程度上改善和加强局部特征对图像的描述能力。或者可以尝试利用主题模型对场景图像进行主题建模，提取图像场景中一些有用的语义信息，结合特征的空间位置信息，对目标相似的场景类别进行描述和分类的研究。

参考文献

- [1] 宫鹏, 黎夏, 徐冰. 高分辨率图像解译理论与应用方法中的一些研究问题[J]. 卫星学报, 2006, 10 (1): 1-5.
- [2] 李德仁. 论 21 世纪卫星与 GIS 的发展[J]. 武汉大学学报, 信息科学版, 2003, 28(2): 127-131.
- [3] 杜凤兰, 田庆久, 夏学齐. 面向对象的地物分类法分析与评价[J]. 卫星技术与应用, 2004, 19(1).
- [4] Lienou M, Maitre H, Datcu M. Semantic annotation of satellite images using latent dirichlet allocation [J]. Geoscience and Remote Sensing Letters, IEEE, 2010, 7(1):28-32.
- [5] Blei D M, Ng A Y, Jordan M I. Latent Dirichlet Allocation [J]. Journal of Machine Learning Research, 2003, 3: 993- 1022.
- [6] Xu K, Yang W, Chen L, et al. Satellite image scene categorization based on topic models [J]. Geomatics and Information Science of Wuhan University, 2011, 36(5):540-543.
- [7] Hofmann T. Probabilistic latent semantic indexing [C]//Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval. ACM, 1999: 50-57.
- [8] Chang C C, Lin C J. LIBSVM: a library for support vector machines [J]. ACM Transactions on Intelligent Systems and Technology (TIST), 2011, 2(3): 27.
- [9] 徐侃, 陈丽君, 杨文, 等. 利用特征选择的遥感图像场景分类[J]. 哈尔滨工业大学学报, 2011, 43(9): 117-121.
- [10] 陶超, 谭毅华, 彭碧发, 等. 一种基于概率潜在语义模型的高分辨率遥感影像分类方法[J]. 测绘学报, 2011, 40(2): 156-162.
- [11] Xia G S, Yang W, Delon J, et al. Structural high-resolution satellite image indexing[C]//ISPRS TC VII Symposium-100 Years ISPRS. 2010, 38: 298-303.
- [12] Caselles V, Coll B, Morel J M. Topographic maps and local contrast changes in natural images[J]. International Journal of Computer Vision, 1999, 33(1): 5-27.
- [13] Xia G S, Delon J, Gousseau Y. Shape-based invariant texture indexing[J]. International journal of computer vision, 2010, 88(3): 382-403.

- [14]Chen L, Yang W, Xu K, et al. Evaluation of local features for scene classification using vhr satellite images[C]//Urban Remote Sensing Event (JURSE), 2011 Joint. IEEE, 2011: 385-388.
- [15]Yang Y, Newsam S. Geographic image retrieval using local invariant features [J]. Geoscience and Remote Sensing, IEEE Transactions on, 2013, 51(2): 818-832.
- [16]Ruiz L A, Fdez-Sarría A, Recio J A. Texture feature extraction for classification of remote sensing data using wavelet decomposition: a comparative study[C]//20th ISPRS Congress. 2004.
- [17]Li C S, Castelli V. Deriving texture feature set for content-based retrieval of satellite image database[C]//Image Processing, 1997. Proceedings. , International Conference on.IEEE, 1997, 1: 576-579.
- [18]Ojala T, Pietikäinen M, Harwood D. A comparative study of texture measures with classification based on featured distributions [J]. Pattern recognition, 1996, 29(1): 51-59.
- [19]Ojala T, Pietikainen M, Maenpaa T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns [J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2002, 24(7): 971-987.
- [20]Qing X, Jie Y, Siyi D. Texture segmentation using LBP embedded region competition [J]. Electronic Letters on Computer Vision and Image Analysis, 2005, 5(1): 41-47.
- [21]Ma L, Zhu L. Integration of the optimal Gabor filter design and local binary patterns for texture segmentation[C]//Integration Technology, 2007. ICIT'07. IEEE International Conference on. IEEE, 2007: 408-413.
- [22]Jin H, Liu Q, Lu H, et al. Face detection using improved LBP under bayesian framework[C]//Multi-Agent Security and Survivability, 2004 IEEE First Symposium on. IEEE, 2004: 306-309.
- [23]Ahonen T, Hadid A, Pietikäinen M. Face recognition with local binary patterns [M]//Computer vision-eccv 2004. Springer Berlin Heidelberg, 2004: 469-481.
- [24]Ahonen T, Hadid A, Pietikainen M. Face description with local binary patterns: Application to face recognition [J]. Pattern Analysis and Machine Intelligence,

- IEEE Transactions on, 2006, 28(12): 2037-2041.
- [25]Lowe D G. Distinctive image features from scale-invariant keypoints [J]. International journal of computer vision, 2004, 60(2): 91-110.
- [26]Fergus R, Fei-Fei L, Perona P, et al. Learning object categories from Google's image search[C]//Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on. IEEE, 2005, 2: 1816-1823.
- [27]Rasiwasia N, Vasconcelos N. Latent Dirichlet Allocation Models for Image Classification [J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2013, 35(11): 2665-2679.
- [28]Koenderink J J. The structure of images [J]. Biological cybernetics, 1984, 50(5): 363-370.
- [29]Lindeberg T. Scale-space theory: A basic tool for analysing structures at different scales[C]//Journal of applied statistics. 1994.
- [30]Brown M, Lowe D G. Invariant Features from Interest Point Groups[C]//BMVC. 2002 (s 1).
- [31]Sivic J, Zisserman A. Video Google: A text retrieval approach to object matching in videos[C]//Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on. IEEE, 2003: 1470-1477.
- [32]Sivic J, Russell B C, Efros A A, et al. Discovering object categories in image collections [J]. 2005.
- [33]Karpathy A, Miller S, Fei-Fei L. Object discovery in 3D scenes via shape analysis[C]//Robotics and Automation (ICRA), 2013 IEEE International Conference on. IEEE, 2013: 2088-2095.
- [34]Deng J, Krause J, Fei-Fei L. Fine-grained crowdsourcing for fine-grained recognition[C]//Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on. IEEE, 2013: 580-587.
- [35]Yao B, Fei-Fei L. Action recognition with exemplar based 2.5 d graph matching [M]//Computer Vision–ECCV 2012. Springer Berlin Heidelberg, 2012: 173-186.
- [36]Baldassano C, Iordan M C, Beck D M, et al. Voxel-level functional connectivity using spatial regularization[J]. NeuroImage, 2012, 63(3): 1099-1106.

- [37]Savva M, Kong N, Chhajta A, et al. Revision: Automated classification, analysis and redesign of chart images[C]//Proceedings of the 24th annual ACM symposium on User interface software and technology. ACM, 2011: 393-402.
- [38]Yao B, Khosla A, Fei-Fei L. Combining randomization and discrimination for fine-grained image categorization[C]//Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on. IEEE, 2011: 1577-1584.
- [39]Zhao B, Li F, Xing E P. Large-scale category structure aware image categorization[C]//Advances in Neural Information Processing Systems. 2011: 1251-1259.
- [40]Russakovsky O, Lin Y, Yu K, et al. Object-centric spatial pooling for image classification [M]//Computer Vision–ECCV 2012. Springer Berlin Heidelberg, 2012: 1-15.
- [41]Yao B, Bradski G, Fei-Fei L. A codebook-free and annotation-free approach for fine-grained image categorization[C]//Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on. IEEE, 2012: 3466-3473.
- [42]Yao B, Fei-Fei L. Recognizing human-object interactions in still images by modeling the mutual context of objects and human poses [J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2012, 34(9): 1691-1703.
- [43]Torrallbo A, Walther D B, Chai B, et al. Good exemplars of natural scene categories elicit clearer patterns than bad exemplars but not greater BOLD activity [J]. PloS one, 2013, 8(3): e58594.
- [44]Fei-Fei L, Perona P. A bayesian hierarchical model for learning natural scene categories[C]//Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on. IEEE, 2005, 2: 524-531.
- [45]Huang Z. Extensions to the k-means algorithm for clustering large data sets with categorical values [J]. Data mining and knowledge discovery, 1998, 2(3): 283-304.
- [46]Kanungo T, Mount D M, Netanyahu N S, et al. An efficient k-means clustering algorithm: Analysis and implementation [J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2002, 24(7): 881-892.

- [47]Jurie F, Triggs B. Creating efficient codebooks for visual recognition[C]//Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on. IEEE, 2005, 1: 604-610.
- [48]Leibe B, Leonardis A, Schiele B. Robust object detection with interleaved categorization and segmentation [J]. International journal of computer vision, 2008, 77(1-3): 259-289.
- [49]Jégou H, Douze M, Schmid C. Packing bag-of-features[C]//Computer Vision, 2009 IEEE 12th International Conference on. IEEE, 2009: 2357-2364.
- [50]Cao Y, Wang C H, Li Z W, et al. Spatial-Bag-of-Features [A]. Proceeding of IEEE Conference on Computer Vision and Pattern Recognition[C], San Francisco, USA,2010: 3352-335.
- [51]MacKay D J C. Information theory, inference, and learning algorithms [M]. Cambridge: Cambridge university press, 2003.
- [52]W. Tobler. A computer movie simulating urban growth in the Detroit region. Economic Geography, 46(2):234–240, 1970.
- [53]Zhang, Jianguo, et al. "Local features and kernels for classification of texture and object categories: A comprehensive study." International journal of computer vision 73.2 (2007): 213-238.
- [54]Joly A. New local descriptors based on dissociated dipoles[C]//Proceedings of the 6th ACM international conference on Image and video retrieval. ACM, 2007: 573-580.
- [55]Fergus R, Perona P, Zisserman A. Object class recognition by unsupervised scale-invariant learning[C]//Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on. IEEE, 2003, 2: II-264-II-271 vol. 2.
- [56]Lazebnik S, Schmid C, Ponce J. A maximum entropy framework for part-based texture and object recognition[C]//Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on. IEEE, 2005, 1: 832-838.
- [57]Berg A C, Berg T L, Malik J. Shape matching and object recognition using low distortion correspondences[C]//Computer Vision and Pattern Recognition, 2005.

- CVPR 2005. IEEE Computer Society Conference on. IEEE, 2005, 1: 26-33.
- [58]Grauman K, Darrell T. The pyramid match kernel: Discriminative classification with sets of image features[C]//Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on. IEEE, 2005, 2: 1458-1465.
- [59]Lazebnik S, Schmid C, Ponce J. Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories[C]//Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on. IEEE, 2006, 2: 2169-2178.
- [60]M. Swain and D. Ballard. Color indexing. IJCV, 7(1):11–32, 1991.
- [61]Cortes C, Vapnik V. Support-vector networks [J]. Machine learning, 1995, 20(3): 273-297.
- [62]王国胜. 支持向量机的理论与算法研究[D]. 北京: 北京邮电大学, 2007.
- [63]丁世飞, 齐丙娟, 谭红艳. 支持向量机理论与算法研究综述[J]. 电子科技大学学报, 2011, 40(1): 2-10.
- [64]张学工. 关于统计学习理论与支持向量机[J]. 自动化学报, 2000, 26(1): 32-42.
- [65]Luo J, Ming D, Shen Z, et al. Multi - scale information extraction from high resolution remote sensing imagery and region partition methods based on GMRF-SVM [J]. International Journal of Remote Sensing, 2007, 28(15): 3395-3412.
- [66]Mercier G, Lennon M. Support vector machines for hyperspectral image classification with spectral-based kernels[C]//Geoscience and Remote Sensing Symposium, 2003. IGARSS'03. Proceedings. 2003 IEEE International. IEEE, 2003, 1: 288-290.

作者简介

宋伟伟（1988-），男，汉族，安徽省蚌埠市人，吉林大学通信工程学院，信号与信息处理专业硕士研究生，主要研究方向：图像处理和模式识别。

硕士期间发表的学术论文：

1.**Song W**, Wen D, Wang K, et al. Satellite image scene classification using spatial information[C]//Sixth International Conference on Graphic and Image Processing (ICGIP 2014). International Society for Optics and Photonics, 2015: 94431K-94431K-5.

2.Zang, M., Wen, D., Wang, K., Liu, T., & **Song, W.** (2015). A novel topic feature for image scene classification. Neurocomputing, 148, 467-476.

致 谢

在吉林大学的研究生生活即将结束，短短的三年时间，一晃而过，我相信这段时光将永远刻在我的记忆中，不可磨灭。在这短暂的三年当中，自己成长了很多，不只是在做学问方面，还懂得了如何去生活，也懂得了做人的很多道理，在此，要对不辞辛苦指引我前行的老师、一路陪伴我的同学和朋友、背后默默付出不求回报的父母致以最真诚的感谢。

首先，我要感谢我的指导老师。研究生期间，王老师为了我学业的完成倾注了太多的心血，从硕士入学的那一天开始，老师就一直告诫我们要正直做人，认真做事。在学习上，老师对我们前进的每一步都严格把关，教导我们要把需要做的事尽量往前赶，这样才会过的比较从容。在老师的督促和鼓励下我完成了硕士期间的第一篇论文，我知道我迈开的这一步老师为了我付出了太多太多，在我遇到困难时，老师总是鼓励我不要着急，争取每天有一点进步就行。时至今日，我真的体会到老师的话语对我是多么的受益。在生活上，老师总是问候我生活得怎么样，有什么困难没有，有困难就和他说，可以说老师对我们的照顾无微不至。在做人上，老师告诫我们不管做人做事不要弄虚作假，要认真踏实，那样将来才能站住脚。我也将铭记老师告诫我们做人的道理。千言万语也道不完我对老师的感激之情，我只能说谢谢王老师，遇到王老师真是我的幸运。

其次，我要感谢我的另一位论文指导老师--文教授，文教授在每次我们网上开会从不缺席，严谨的治学态度深深的打动了我们。每次开会，不管我们这个团队中谁遇到了什么技术困难，老师总是能耐心认真地指导解答，老师是我研究生阶段的路标，老师引导我们看国际前沿的论文，了解国际的前沿技术，避免了我们的闭门造车。在老师的帮助下，指导我们更好地去做研究，撰写论文以及修改论文。在此，我要真诚的感谢老师。

然后，我要感谢实验室这个团队的所有成员，在我成长的路上，你们给了我太多的帮助和鼓励。研究生阶段，大家学术上互相讨论，使我能够取长补短，更好地完善自己。遇到困难时大家积极出谋划策，让我在有限的时间内能做更多的事情。

最后，我要特别感谢我的父母，是你们不求回报默默的付出支撑了我的一

片天，岁月的沧桑无情在你们脸上留下了皱纹，你们就是我的精神支柱。同时我还有感谢我的家庭所有成员，是你们在背后默默的支持和鼓励，给我带来欢乐，让我享受到了大家庭的温暖和幸福。回报父母回报长辈将是我未来前进的动力和目标。