

一种改进包模型的遥感图像检索方法

杨 进¹ 刘建波¹ 戴 芹¹

1 中国科学院遥感与数字地球研究所,北京,100094

摘 要:提出了一种改进的包模型图像检索方法,使用本底图像构建视觉词汇字典。在建库过程中,通过近似 KD-Tree 搜索本底图像局部不变特征与入库图像特征间的对应匹配关系,并记录该关系;图像检索时,搜索检索图像与本底图像特征的对应关系,进而计算与入库图像的关系。该方法仅保存本底图像的局部不变特征,可实现感兴趣区域的图像检索。实验结果表明,该方法在提高图像检索准确性的同时,可以减少所需存储局部不变特征的数量。

关键词:遥感图像检索;本底图像;包模式;视觉词汇;感兴趣区

中图法分类号:P237.4; TP753 **文献标志码:**A

如何在海量遥感影像中快速检索感兴趣区域与目标成为目前急需解决的问题。传统的根据影像元数据信息的影像检索方式已不能很好地实现感兴趣区域检索^[1]。视觉词汇包(bag-of-word, BoW)方案是目前基于局部不变特征检索图像的最主要框架,在图像检索领域有多个成功的应用案例^[2-5]。但是,BoW 检索方案在图像存在较多相似纹理时,检索结果的准确性较低。BoW 方案中,视觉词汇通过 K-Means 等方法生成,这等同于全局最优约束下的分类。视觉词汇的生成不受用户指导与控制,局部特征分配到视觉词汇时采用最近邻原则。这种视觉词汇生成与分配方式存在稳定性低与可靠性差的缺点。根据在 Oxford 数据集上的测试,局部特征分配到视觉词汇时,最近邻词汇与次近邻词汇的距离比平均值大于 0.95。这意味细微的图像变化,如光照、尺度、旋转等,会使相同目标的局部特征分配到不同的视觉词汇中。这无疑会降低图像搜索的准确性。

本文提出了一种适合遥感影像的改进视觉词汇包模型图像检索方案。该方案利用遥感影像库中的图像间存在大量重叠区域的特点,建立本底特征库,用特征库中的特征作为视觉词汇表示遥感影像。该方案在将特征分配到视觉词汇的过程中可以进行空间校验,这增强了特征分配的稳定性与可靠性。实验结果表明,在遥感图像检索中,本文提出的方法同 BoW 方案相比,在提高目标检索性能的同时减少了所需的存储空间。

1 改进的 BoW 图像检索方案

BoW 方案首先在文档检索中使用,文献[6]将其扩展到计算机视觉领域,并提出了一种基于 BoW 的图像检索框架。为了提高 BoW 框架的图像检索准确性,文献[7]将文档查询中广泛应用的词频-逆文档频率(term-frequency inverse-document-frequency, TF-IDF)方法引入视觉词汇中。结合 K-NN 投票算法减少高频视觉词汇的权重,同时提高低频视觉词汇的权重。文献[8]从文本查询方法中引入一种查询词汇扩展图像检索方法。该方法利用图像串将初次查询结果中的高频视觉词汇作为查询条件进行附加查询。查询扩展在初次查询有较高准确度时能进一步提高系统整体检索性能。

上述对 BoW 基本框架的改进虽然在一定程度上提高了系统检索图像的准确性,但没有解决局部不变特征分配到视觉词汇时不稳定的现象。

在遥感影像库中,包含大量相同目标不同尺度、不同光照条件、不同成像时间、不同卫星、不同传感器拍摄的影像。也就是说,遥感影像库中有大量同名地物的多幅影像。以 Landsat5 卫星数据为例,中国区域的覆盖次数超过了 400 次。

本文从相同地区的多时相数据中选取若干本底影像作为模板,用于间接比较影像之间的相似度。

收稿日期:2014-04-18

第一作者:杨进,博士,副教授,主要从事遥感图像检索理论与方法研究。E-mail:yangjin@radi.ac.cn

利用本底影像改进的 BoW 遥感影像检索 (remote sensing image retrieval, RSIR) 框架如图 1 所示。

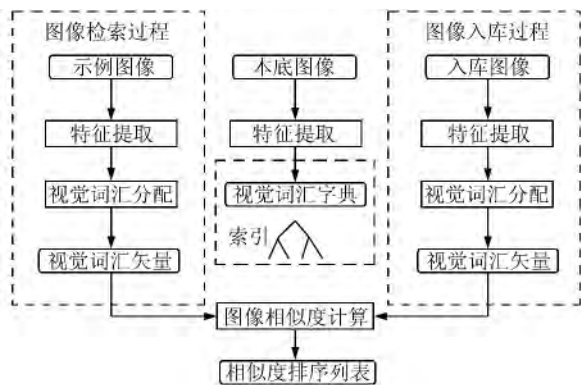


图 1 改进的 BoW 遥感影像检索框架

Fig. 1 Improved Remote Sensing Image Retrieval Framework

RSIR 框架包含特征提取、视觉词汇分配与图像相似度计算三个模块。本底特征库与入库影像特征检测使用相同的不变特征提取方法。将从本底图像中提取的特征作为视觉词汇，形成视觉词汇字典。入库图像通过视觉词汇分配模块表示为视觉词汇矢量。图像检索时，同样将检索图像转换为视觉词汇矢量，再由相似度比较模块计算最相似的图像。

1.1 局部不变特征检测

特征检测器的选择根据遥感影像的图像变化特点与性质确定。遥感影像一般不存在大侧视角情况，图像之间的仿射变形较小。图像之间的差异主要表现在尺度、光照与成像光谱上，因此采用高斯差分 (difference of Gaussian, DoG) 特征检测器^[9-10]。特征点检测后，使用 128 维尺度不变特征变换 (SIFT) 特征描述子描述该特征。

遥感影像包含多个波段，特征提取前，应将多个波段合成为灰度图像后再进行特征检测，从合成后的图像中能检测出更多的局部不变特征。以 Landsat-5 数据为例进行说明。选取 5、4、3 波段合成灰度图像，该波段组合包含了更多的地物信息。5、4、3 波段组合成一个假彩色图像，假彩色图像转换到灰度图像使用颜色空间亮度转换方法。具体方法为：

$$G_{\text{scale}} = 0.3\text{band5} + 0.59\text{band4} + 0.11\text{bands} \quad (1)$$

其中，band5、band4、band3 分别是这几个波段的灰度值； G_{scale} 是灰度值。

1.2 本底图像的选择方法

本文使用本底影像中提取的特征作为视觉词

汇，本底图像的选择决定检索的准确性。根据遥感图像的特点，设计了两条本底图像准备原则：① 图像质量原则。本底图像应是无云、无缺行数据。② 图像时相原则。每一个区域上有多个时相的遥感影像，这些遥感影像组成一个图像的按时间排列的序列串。选取本底数据时，应选择在时间维度上处于序列串中部的数据。这是因为遥感影像会随环境与人类活动发生变化。时间间隔越长，图像差异性越大。选取位于时间中部的影像不会造成本底图像与目标影像之间差异过大。

1.3 使用视觉词汇表示影像

使用视觉词汇表示图像分两个步骤进行：

① 将图像的局部不变特征映射到视觉词汇，即本底特征库中的特征；② 将映射后的词汇组织到便于比较的数据结构中。

1.3.1 特征映射

首先，通过最近邻与次近邻比值方法在本底图像特征库中筛选出初步匹配的视觉词汇。计算公式为：

$$D_{A-B} / D_{A-C} < t \quad (2)$$

其中， D_{A-B} 是图像局部特征矢量 A 到最近邻视觉词汇矢量 B 的欧氏距离； D_{A-C} 是到次近邻词汇矢量 C 的欧氏距离； t 表示阈值。根据文献^[9]， t 小于 0.6 时，可以过滤掉 80% 误匹配特征，因此，本文 t 设为 0.6。

然后，使用 RANSAC 方法^[11] 验证匹配对之间的空间关系。遥感影像之间的空间变换近似为相似性变换，匹配特征点之间的关系符合相似变换。将相似变换分解为平移、旋转、尺度等初等图像变换。相似变换的 4 个变换参数可以用 4 组对应的坐标点通过解方程组求解。

$$\begin{cases} x = a_1x' + b_1y' + c_1x'y' + d_1 \\ x = a_2x' + b_2y' + c_2x'y' + d_2 \end{cases} \quad (3)$$

式中， (x, y) 、 (x', y') 分别是校正图像与目标图像中特征点坐标；其他变量是待求解参数。

通过 RANSAC 确定匹配特征点的具体步骤为：① 在初步匹配对中随机选取 4 组数据，建立方程组求解相似变换系数；② 在剩余的数据集中验证计算出的系数，将误差在一定范围的数据归类为内点，反之为外点；③ 步骤①和②反复迭代。选取使最多数据为内点的一组参数作为最终的模型参数，并以此计算内点；④ 内点作为最终正确匹配点。

1.3.2 特征组织

将图像的局部不变特征分配到视觉词汇后，使用倒排文件 (invert file, IF) 数据存储结构表示

图像。首先为视觉词汇字典中的每一个视觉词汇建立一个存储链表,链表中记录的是图像的 ID。当入库图像的特征分配到视觉词汇后,在该视觉词汇对应的链表中记录入库图像的 ID。

1.4 图像相似度计算

本文计算图像间相似度的方法如图 2 所示。两幅图像,通过本底图像(模板图像)比较它们之间的相似性。当两幅图像都与本底影像相近时,它们也相近,反之差异较大。图像与本底图像相近时,表明图像在本底特征库中的相应位置分配有较多的视觉词汇。当两幅图像有较多相同视觉词汇,即本底特征库中重叠区较大时,图像相似。重叠区大小计算方法如图 2(a)所示:

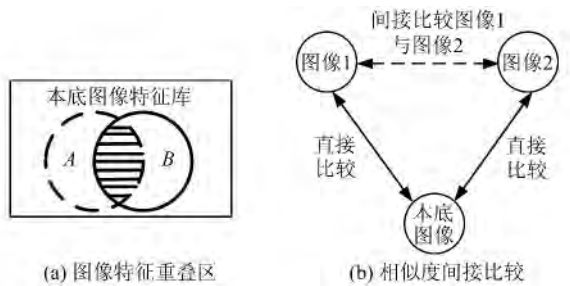


图 2 遥感图像相似度计算方法

Fig. 2 Similarity Measure for Remote Sensing Image

图 2(a)中, A 是检索图像, B 是图像库中图像,图中的矩形区域代表本底特征库,即视觉词汇字典。圆形 A、B 区域分别表示图像 A、B 在特征库中分配的视觉词汇。A 与 B 相交的部分为两幅图像在本底库中共有的视觉词汇。图像相似度数值计算公式为:

$$\text{rank}_{AB} = \frac{N_{AB}}{T_A + T_B} \quad (4)$$

其中, N_{AB} 表示图像 A 与 B 共有的视觉词汇数量; T_B 与 T_A 分别是图像 A、B 在本底特征库中分配的视觉词汇数量。根据式(4)计算出所有 rank 值大于 0 的图像,并对 rank 值排序,得出与检索图像相近的图像序列。

2 实验与分析

2.1 实验数据

测试数据集包含 20 个区域的 8 223 幅 Landsat5 快视图像,每一幅影像的大小为 512×512 (见图 3)。20 个区域遥感影像拍摄日期时间跨度为 1985~2011 年。数据集中数据的云量、成像的光照条件以及图像的质量相差较大,覆盖多种图像变形情况。此外,选取同覆盖区域的 HJ-1A 卫

星数据作为检索数据,测试在不同卫星数据间进行检索的性能。

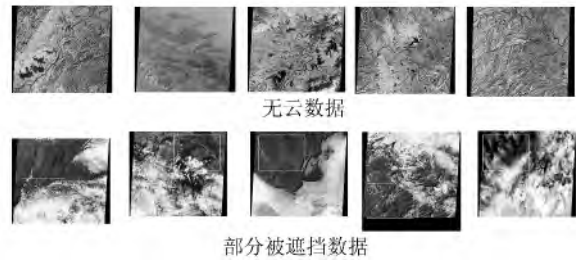


图 3 数据集中随机抽取的样例图

Fig. 3 Some Examples of Remote Sensing Image in the Dataset for Experiment

根据 § 1.2 节的数据选择方法从数据集中选出本底数据。本底数据的数量对检索准确性有影响。本文测试了一个标准分幅区域中包含 1~4 景本底数据 (base image, BI) 4 种情景下的检索性能。

BoW 框架所需的视觉词汇来自 INRIA Flickr60K^[12]。测试时,使用 10 k、50 k、200 k 和 500 k 的视觉词汇字典。

2.2 检索准确性评价指标

使用平均精度 (average precision, AP) 方法测试图像检索准确度。AP 定义为精度-召回率曲线下的面积。AP 的理论值为 0 到 1 之间的浮点数,AP 的数值越大,图像检索的性能就越高。

2.3 测试结果

2.3.1 特征占用空间

图像库占用的存储空间包括两部分,一部分用于存储视觉词汇典,另一部分用于存储每一幅图像的视觉词汇直方图矢量。第一部分占据的存储空间与所用字典大小有关。第二部分所占空间随着图像库中图像数量的增多,比重逐渐增加。表 1 显示了使用全部特征、BoW 方案与本文 RSIR 方案时所需管理的特征数量以及图像库占用空间。在使用 4 本底数据时,本文方案的视觉词汇字典中的词汇数量虽然比 BoW 方案的多,但是由于本文只记录映射正确的特征,因此,图像

表 1 不同方案特征占用空间比较

Tab. 1 Comparison of Occupied Storage Space Between BoW Framework and Our Approach

方案	图像数量	描述子数量	词汇数量	占用空间
全部特征				1.5 G
BoW	8 223	12.3 M	10 k	34 M
			50 k	38 M
			500 k	91 M
RSIR-(4BI)			92 k	16 M

特征库整体所需的存储空间少于 BoW 方案。

2.3.2 图像检索性能

表 2 给出了 AP 检索性能。本文 RSIR 方法的 AP 指标明显高于 BoW 方法。RSIR 方法中,随着本底数据的增加,检索性能有所提升,但是使用 3BI 后,性能提升有限。BoW 方案由于存在视觉词汇错误分配的情况,存在错误检索结果。

表 2 本文与 BoW 图像检索性能比较

Tab. 2 Performance Comparison Between BoW Framework and Our Approach

检索方案	检索图像		好数据	部分遮挡数据
	Landsat-5	HJ-1A		
BoW(10 k)	0.37	0.36	0.38	0.12
BoW(50 k)	0.39	0.38	0.39	0.14
BoW(200 k)	0.43	0.40	0.43	0.16
BoW(500 k)	0.43	0.39	0.42	0.16
RSIR-(1BI)	0.64	0.51	0.66	0.36
RSIR-(2BI)	0.71	0.62	0.72	0.42
RSIR-(3BI)	0.74	0.62	0.74	0.45
RSIR-(4BI)	0.76	0.63	0.76	0.47

图 4 的精度-召回率关系中,本文 RSIR 方法在大部分召回率值下都保持在 1 附近。BoW 的准确性随召回率增加衰减很快。BoW 方法通过视觉词汇分布概率矢量计算图像间的距离,它可以将检索图像与数据库中所有图像的距离进行排序,因此召回率几乎可以到 1,但此时的准确性很低。

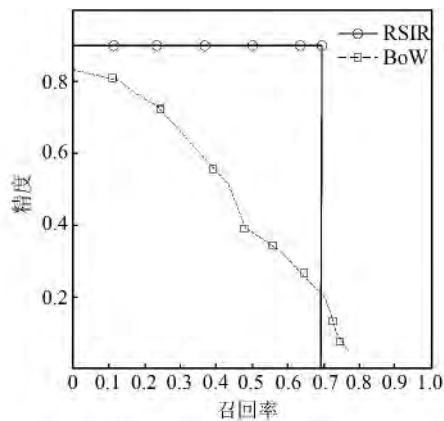


图 4 本文方法与 BoW 方法的精度-召回率关系比较

Fig. 4 Precision Recall Graph Comparison Between BoW Framework and RSIR Approach

3 结语

遥感数据具有对同一地物反复成像的时空分布特点。针对这一特点,本文提出了一种基于本底图像局部特征间接比较图像的遥感影像检索模型。该模型使用本底图像的局部特征作为视觉词

汇,在特征分配到视觉词汇过程中,应用最近邻与次近邻比例法与空间约束方法大幅度提高匹配的稳定性与可靠性,在整体上提高遥感影像检索的准确性。实验结果表明,相较于 BoW 方法,本文方法能提高遥感影像的检索准确性,减少特征存储空间,为建立大型图遥感影像库检索系统提供了一个可行的解决方案。

本文提出的检索方案具有按示例图检索数据的功能,适用于感兴趣区域的数据检索。由于该方案不依赖元数据信息,因此,可以将 Web 上无元数据信息图像作为示例图像进行检索。在建立实际检索系统时,建议本底数据可从 Google 地球的全球缓存数据中获取。

虽然本文工作主要针对遥感数据检索,但研究结果还可用于其他领域,比如遥感影像自动镶嵌时的数据自动选择等。

参 考 文 献

- [1] Mei Lixia, Tian Yumin. Region-Based Image Database Retrieval Method[J]. *Computer Engineering*, 2003, 29(5):127-129(梅丽霞, 田玉敏. 基于区域的图像数据库检索方法[J], 计算机工程, 2003, 29(5):127-129)
- [2] Molinier M, Laaksonen J, Hame T. Detecting Man-made Structures and Changes in Satellite Imagery with a Content-Based Information Retrieval System Built on Self-organizing Maps[J]. *IEEE Transactions on Geosciences and Remote Sensing*, 2007, 45(4): 861-874
- [3] Samal A, Bhatia S, Vadlamani P, et al. Searching Satellite Imagery with Integrated Measures [J]. *Pattern Recognition*, 2009, 42(11): 2 502-2 513
- [4] Shyu C R, Klaric M, Scott G J, et al. GeoIRIS: Geospatial Information Retrieval and Indexing System-content Mining, Semantics Modeling, and Complex Queries[J]. *IEEE Transactions on Geosciences and Remote Sensing*, 2007, 45(4): 839-852
- [5] Dateu M, Daschiel H, Pelizzari A, et al. Information Mining in Remote Sensing Image Archives: System Concepts[J]. *IEEE Transactions on Geosciences and Remote Sensing*, 2003, 41(12): 2 923-2 936
- [6] Sivic J, Zisserman A. Video Google: A Text Retrieval Approach to Object Matching in Videos[C]. The Ninth IEEE International Conference on Computer Vision, Washington D C, 2003
- [7] Nister D, Stewenius H. Scalable Recognition with a Vocabulary Tree[C]. IEEE Computer Society Conference on Computer Vision and Pattern Recogni-

- tion, Washington D C, 2006
- [8] Chum O, Philbin J, Sivic J, et al. Total Recall: Automatic Query Expansion with a Generative Feature Model for Object Retrieval[C]. IEEE 11th International Conference on Computer Vision, Washington D C, 2007
- [9] Lowe D. Distinctive Image Features from Scale-invariant Keypoints [J]. *International Journal of Computer Vision*, 2004, 60(2): 91-110
- [10] Wan Xue, Zhang Zuxun, Ke Tao. An Improved SIFT Operator Based on the Theory of Zero-crossing on Feature Extraction[J]. *Geomatics and Information Science of Wuhan University*, 2013, 38(3): 270-273(万雪,张祖勋,柯涛. 一种利用零交叉点理论的改进 SIFT 特征提取算法[J]. 武汉大学学报·信息科学版, 2013, 38(3): 270-273)
- [11] Zhang Zuxun, Su Guozhong, Zhang Jianqing. Airplane Pose Measurement from Image Sequences[J]. *Geomatics and Information Science of Wuhan University*, 2004, 29(4): 287-291(张祖勋, 苏国中, 张剑清. 基于序列影像的飞机姿态跟踪测量方法研究[J]. 武汉大学学报·信息科学版, 2004, 29(4): 287-291)
- [12] Jegou H, Douze M, Schmid C. Hamming Embedding and Weak Geometric Consistency for Large Scale Image Search[C]. European Conference on Computer Vision, France, 2008

An Improved Remote Sensing Image Retrieval Method Based on Bag of Word Framework

YANG Jin¹ LIU Jianbo¹ DAI Qin¹

¹ Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100094, China

Abstract: The Bag-of-Words (BoW) approach has lower retrieval accuracy performance for remote sensing image retrieval. The rate of correct results is very low especially when target textures are similar. To overcome this shortcoming, an indirect comparison between a query and target image based on local invariant feature of the background image is proposed. The solution has two phases: the process of building a database and a retrieval process. After building a database, a search of the nearest neighbor feature in a query image in the feature space of the image dataset through an approximate KD-Tree is executed, then their relationship is recorded. The relationship can be used in calculating similarity between a query and target image during a retrieval procedure. Experiments show that the proposed method has a better recognition performance than that of the BoW and, at the same time, needs less storage space for saving local invariant features.

Key words: remote sensing image retrieval; background image; bag-of-words; visual word

First author: YANG Jin, PhD, associate professor, specializes in remote sensing image retrieval. E-mail: yangjin@radi.ac.cn