

基于 ScSPM-Reranking 的高分辨率遥感影像的检索^①

弓永利^{②*} 朱盼盼^{**} 王跃宾^{③**}

(* 内蒙古建筑职业技术学院市政与路桥工程学院 呼和浩特 010070)

(** 北京师范大学地理学与遥感科学学院 北京 100875)

摘 要 为了从高分辨率遥感影像中获取详细的地表地物信息,为城市规划、环境监测以及灾情分析提供可靠的数据,进行了高分辨率遥感影像的检索研究,包括对图像的特征提取和图像之间相似度的描述。为了提高图像检索精度,运用了采用稀疏编码(Sc)的空间塔式匹配(ScSPM)技术和重排序(Reranking)技术,提出了基于 ScSPM 结合 Reranking(ScSPM-Reranking)的遥感高分辨率影像的检索方法。该方法首先使用 ScSPM 提取空间场景的特征,然后结合这些特征使用 cityblock 距离进行初步检索,最后对初步检索的结果进行 Reranking 排序,获得高精度的检索结果。同其他检索方法进行了对比实验,实验结果证明,该方法具有较高的检索精度。

关键词 高分辨率遥感影像, 图像特征描述, 图像检索, Reranking, 稀疏编码(Sc), 空间塔式匹配(SPM)

0 引 言

目前,通过卫星传感器获得的遥感影像的空间分辨率可以达到几十厘米,因此这些传感器可以为我们提供大量十分清晰的地物信息^[1,2]。但局限于地物的多样性以及地物空间结构的复杂性,给自动识别和认知这些地物带来了困难。高分辨率遥感影像的检索是对地物认知和了解的一个重要的内容^[3]。高分辨率的遥感影像的检索涉及到图像的特征提取以及对图像之间相似度的较好的描述^[4,5]。图像特征描述子的好坏,直接关系着图像识别与检索的精确度^[6,7]。所以图像的特征描述子应该具有较强的鲁棒性,较少受图像物体旋转、平移、尺度变换的影响。常用的图像特征有颜色特征、纹理特征、形状特征、空间关系等特征^[8-11]。颜色特征是一种全局特征,描述了图像上人们感兴趣目标所对应的景物的表面性质。纹理特征描述了图像或

图像区域所对应景物的表面性质。由于纹理只是一种物体表面的特性,并不能完全反映出物体的本质属性,所以仅仅利用纹理特征无法获得高层次图像内容。形状特征可以比较有效地利用图像中感兴趣的目标来进行识别^[11],但它也有一些缺陷,目前基于形状的分类方法还缺乏比较完善的数学模型,如果目标有变形时分类结果往往不太可靠。所谓空间关系,是指图像中的多个目标之间相互的空间位置或相对方向关系,这些关系也可分为连接/邻接关系、交叠/重叠关系和包含/包容关系等。空间关系特征的使用可加强对图像内容的描述区分能力,但空间关系特征常对图像或目标的旋转、反转、尺度变化等比较敏感^[12]。另外,实际应用中,仅仅利用空间信息往往是不够的,不能有效准确地表达场景信息。

比较常见的图像特征提取方法有尺度不变特征转换(scale invariant feature transform, SIFT)方法^[13]

① 国家自然科学基金(41371324)资助项目。
② 女,1976 年生,副教授,高级工程师;研究方向:工程测量,遥感,地理信息应用;E-mail: 2240924122@qq.com
③ 通讯作者, E-mail: xxgcdxwyb@163.com
(收稿日期:2017-01-10)

和梯度直方图 (histogram of oriented gradients, HOG) 方法^[14]。SIFT 是用于图像处理领域的一种描述子。这种描述具有尺度不变性,可在图像中检测出关键点,是一种局部特征描述子。后来文献 [15] 在文献 [13] 的基础上结合词袋 (bag of words, BoW) 方法,提出了塔式关键词直方图 (pyramid histogram of words, PHOW) 方法,词袋借鉴文本分类技术,将图像表示成一系列视觉关键词 (visual words) 的统计直方图。所谓视觉关键词就是由训练图像局部区域特征 (特征点) 经过聚类形成的聚类中心,视觉关键词的集合形成所谓词袋。传统的词袋模型忽略了图像的空间位置特征,并且采用稀疏采样方式,不利于图像结构特征的提取。HOG 是描述图像形状特征的有效方法,HOG 特征通过提取局部区域的边缘或梯度的分布,可以很好地表征局部区域内目标的边缘或梯度结构,进而表征目标的形状。HOG 特征没有考虑到图像的不同空间尺度对特征提取的影响,所以,文献 [16] 在文献 [15] 的基础上提出了特征提取的分层梯度方向直方图 (pyramid histogram of orientated gradients, PHOG) 方法,使用空间四叉树对图像分解,形成图像的多分辨率表示,通过合并从低分辨率到高分辨率的多级梯度方向直方图来描述图像。

上述描述遥感图像特征的方法都存在着一定的不足,因为图像本身和提取出来的图像特征都存在着大量的冗余,这些特征可以通过稀疏编码 (sparse coding, Sc) 将原始图像特征分解为一组图像基的线性组合,而且只用较少的几组基就可以将原始图像的特征恢复出来,进而减少冗余,提高图像识别的精度。因此,本文在稀疏编码 (Sc) 的基础上,基于文献 [17] 的采用稀疏编码的空间塔式匹配 (spatial pyramid matching (SPM) using Sc, ScSPM) 方法进行遥感图像特征的提取,并进一步结合重排序 (Reranking) 检索技术^[18],提出了基于 ScSPM-Reranking 的高分辨率遥感影像的检索方法,高精度地描述了图像之间的相似度。

1 基于 ScSPM 的高分辨率遥感影像特征的提取

对于遥感图像的特征,本研究首先提取图像的

dense SIFT 特征,然后结合稀疏编码技术和空间塔对原始的特征作进一步的处理。

稀疏编码是一种非监督的特征学习算法,它像主成分分析 (principal components analysis, PCA) 一样,用一组“超完备”基向量高效地表示样本数据 X 。它们能更有效地找出隐含在输入数据内部的结构与模式,如下式所示:

$$X = \sum_{i=1}^n U_i V_i \tag{1}$$

式中 U_i 表示基向量 V_i 的系数,它通常是稀疏的。查询图像和数据库图像首先提取 SIFT 特征,dense SIFT 特征从原始图像中提取的规格为 16×16 像素,步长为 8 像素,接下来对原始特征进行稀疏编码处理,稀疏编码通过一组过完备的基向量来表示样本数据,进而发现图像数据的纹理结构和模式信息。

对于一幅图像或图像特征 $X = [x_1, \dots, x_n] \in R^{m \times n}$,它可以使用矩阵因式分解的方法分解为两个矩阵: $U \in R^{m \times k}$ 和 $V \in R^{k \times n}$,使得 $X = UV$,我们可以使用 Frobenius norm 等范数来实现逼近,所以,对于一幅图像的矩阵分解可以通过下式来实现:

$$\min_{U,V} \|X - UV\|_F^2 + \|V\|_1 \tag{2}$$

式(2)可以用一个迭代算法求解,首先初始化 U ,求解 V ,关于一幅图像 x_i 的稀疏表达 v_i 可以用下式计算:

$$v_i = v_i^* (x_i, U) \equiv \operatorname{argmin}_{v_i} \|x_i - Uv_i\|_2^2 + \|v_i\|_1 \tag{3}$$

这是一个 L1 优化问题,可用很多 L1 优化方法求解,比如 Feature-sign 算法^[19]和 GPSR 算法^[20]。

其次固定 V ,求解 U ,可以用文献 [21] 描述的多变量岭回归模型 (multivariate ridge regression model),利用二次损失函数和 L2 正则项求解下式:

$$U = \operatorname{argmin}_U \|X - UV\|_F^2 + \lambda_1 \|U\|_2^2 \tag{4}$$

则 U 的解为

$$U = XV^T (VV^T + \lambda_1 I)^{-1} \tag{5}$$

重复这个过程,直到满足收敛条件。

V 为使用式(2)计算稀疏编码的结果,它是一组稀疏系数,也是对应的图像特征的表达, U 是对应的学习到的字典。

通过稀疏编码学习到的特征应该能够获取到图

像本身的特点,但是少了对图像多层次稀疏特征的表达,为了更好地提取图像的内在特征,这里使用了文献[17]的 ScSPM 的方法获取图像的特征。ScSPM 是基于稀疏编码(Sc)的空间塔式匹配(SPM)的改进。

本文通过式(6)、(7)来计算深层次的图像特征:

$$\mathbf{z} = F(\mathbf{V}) \quad (6)$$

$$z_i = \max\{|v_{1i}|, |v_{2i}|, \dots, |v_{Mi}|\} \quad (7)$$

在式(6)和(7)中, z_i 是 $\mathbf{z}$ 的第 i 个元素, v_{ij} 是 $\mathbf{V}$ 对应的 i 行 j 列的元素, F 是基于 $\mathbf{V}$ 定义的池化函数。

在实验中,首先基于图像提取 dense SIFT^[22] 特征,然后在这些特征的基础上进行稀疏编码计算,提取 $\mathbf{V}$,最后结合式(6)和(7)计算图像特征的多尺度的表达,基本流程如图 1 所示。

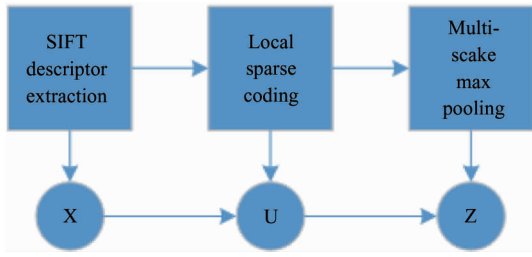


图 1 ScSPM 特征提取流程图

2 基于 Reranking 的遥感影像检索

在本文第 1 节,通过 ScSPM 提取了遥感影像的特征,进一步可以借助于 cityblock 距离或欧式距离,来衡量图像之间的相似度。在本文中,使用 city-block 距离来进行遥感影像之间的初步检索,然后在此基础上,对初步检索的结果进行 Reranking,得到高精度的检索结果。

本文基于文献[18]对初步检索的结果进行 Reranking。通过计算 cityblock 距离可以得到与查询图像目标 Q 对应的前 k 个匹配结果。为了获取高精度的匹配图像,需要对检索结果进行 Reranking。

在图 2 中,利用 cityblock 距离搜索查询图像 Q 对应的前 k 个匹配图像,存放在列表, $list_Q$ 中所检索

到的图像 $N_1, N_2, \dots, N_k$ 分别以其自身为检索图像继续检索,最后形成 $k+1$ 个列表: $list_Q, list_{N_1}, list_{N_2}, \dots, list_{N_k}$ 。由于数据库图像数量是固定的,在以上检索的结果中,存在着大量重复的彼此交叉的检索结果(数据库图像),如何更好地表达在这些检索中的图像与查询图像 Q 之间的相似度,是一个关键的问题。



图 2 Reranking 排序

通过式(8)来计算在以上检索列表中某一幅图像与查询图像之间的相似度:

$$S(Q, N_{ij}) = S(Q, N_i) \cdot S(N_i, N_{ij}) \quad (8)$$

式中, $S(Q, N_i)$ 表示查询图像 Q 与数据库图像 N_i 的关系, $S(N_i, N_{ij})$ 表示图像 N_i 与数据库图像 N_{ij} 的关系。如果 Q 与数据库图像 N_i 的关系比较近,需要图像 N_i 作为查询图像时, Q 在检索结果中的位置也是比较靠前的。所以有

$$S(Q, N_i) = \frac{1}{i + R(N_i, Q) + 1} \quad (9)$$

$$S(N_i, N_{ij}) = \frac{1}{R(N_i, N_{ij})} \quad (10)$$

式(9)中, i 表示图像 N_i 在 $list_Q$ 出现的位置, $R(N_i, Q)$ 表示图像 Q 在 $list_{N_i}$ 出现的位置。式(10)中, $R(N_i, ij)$ 表示图像 N_{ij} 在 $list_{N_i}$ 出现的位置。

最后,统计所有重复的图像 N_{ij} 与查询图像 Q 的关系:

$$S(Q, N_{ij}) = \sum_{i=0}^k \frac{1}{i + R(N_i, Q) + 1} \cdot \frac{1}{R(N_i, N_{ij})} \quad (11)$$

总体的算法流程见算法 1。

算法 1

输入:图像特征矩阵 X , 查询图像 Q , 参数 λ_1 , 最多迭代次数 t 。

输出:与查询图像 Q 相似的高精度的图像检索结果。

1 初始化 U , 利用 Feature sign 算法计算稀疏系数矩阵 V , 如下:

$$v_i = v_i^*(x_i, U) \equiv \operatorname{argmin}_{v_i} \|x_i - Uv_i\|_2^2 + \|v_i\|_1$$

2 固定 V , 利用岭回归模型求解 U :

$$U = XV^T(VV^T + \lambda_1 I)^{-1}$$

3 重复步骤 1 和 2, 直到满足收敛条件;

4 池化过程:

$$z = F(V)$$

$$z_i = \max\{|v_{1i}|, |v_{2i}|, \dots, |v_{Mi}|\}$$

5 利用 cityblock 距离搜索查询图像 Q 对应的前 k 个匹配图像, 存放在列表, $list_Q$ 中所检索到的图像 $N_1, N_2, \dots, N_k$, 分别以其自身为检索图像继续检索, 最后形成 $k+1$ 个列表: $list_Q, list_{N_1}, list_{N_2}, \dots, list_{N_k}$;

6 Reranking 过程, 计算在以上 $k+1$ 检索列表中任意一幅图像与查询图像之间的相似度:

$$S(Q, N_{ij}) = \sum_{i=0}^k \frac{1}{i + R(N_i, Q) + 1} \cdot \frac{1}{R(N_i, N_{ij})}$$

其中 $R(N_i, Q)$ 表示图像 Q 在 $list_{N_i}$ 出现的位置。式(7)中, $R(N_i, N_{ij})$ 表示图像 N_{ij} 在 $list_{N_i}$ 出现的位置。

7 根据 $S(Q, N_{ij})$, 对初步检索出来的图像库中的图像进行排序, 获取最终的检索结果。

3 应用实验

本文采用 LULC^[23,24] 的数据库对所提出的算法进行了验证。LULC 数据库包括了 21 种土地利用和土地覆盖数据: Agricultural、Airplane、Baseball diamond、Beach、Buildings、Chaparral、Dense residential、Forest、Freeway、Golf course、Harbor、Intersection、Medium residential、Mobile home park、Overpass、Parking lot、River、Runway、Sparse residential、Storage tanks、Tennis court。它来源于 U. S. Geological Survey (USGS) National Map, 各种地物影像的分辨率达 30cm, 每类地物包含 100 幅 256×256 像素的遥感影像, 在图 3 中有示例图。这些图像的空间分辨率是大约一个像素代表一英尺, 每幅图像大概是 256×256 像素, 这些数据高度地重叠, 例如密集居民地、中等居民地、稀疏的居民地, 它们主要区别在于对象空间结构的密度。

为了验证本文提出的算法的精度, 本研究做了大量的实验并且提供了详尽的分析。在实验中, 首先基于 ScSPM 提取遥感影像的特征, 然后使用 Reranking 的方法, 进行图像的检索。在实验中, 用平均检索监督 (mean average precision, MAP) 来比较

检索结果。接下来, 进行不同方法检索精度的对比。这几种方法是: (1) 使用 ScSPM 进行特征提取, 然后在此基础上使用 cityblock 距离进行图像检索; (2) 使用 ScSPM 进行特征提取, 采用 Reranking 的方法进行检索; (3) 采用 PHOG 方法对高分遥感影像进行特征提取, 然后使用 chi-square 衡量图像之间的相似度; (4) 使用 PHOW 进行特征提取, 结合 chi-square^[25] 衡量图像之间的相似度。然后, 分别设置 Reranking 中的 k 为 10, 20, 30, 进行精度评估, 评估结果如图 4、5、6 所示。

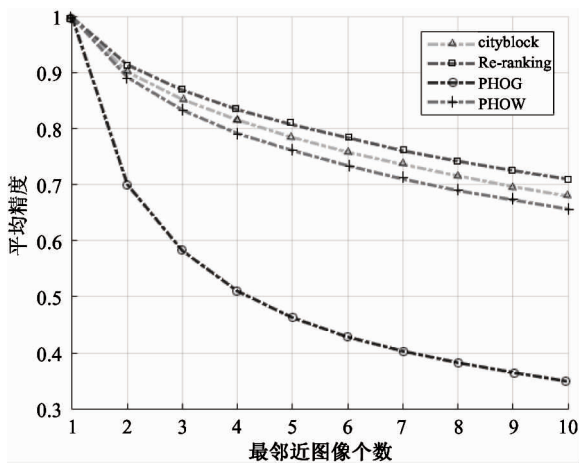
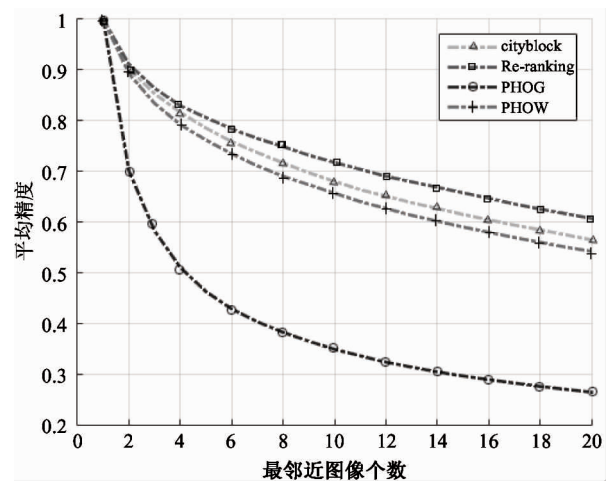
在图 4、5、6 中, 横轴表示在检索结果中前 k 个元素, 竖轴表示前 k 幅图像对应的平均检索精度。在这些检索结果中, 可以发现本文提出算法的检索精度是最高的。

首先, 固定同种相似度比较模型 chi-square 距离, 比较提取的不同特征: cityblock、PHOG、PHOW, 我们提取的遥感影像的 ScSPM 特征的精度明显好于 PHOG 和 PHOW。

其次, 固定下同一种特征 ScSPM, 来比较 Reranking 之前和之后的精度, 在图 4、5、6 中发现 Reranking 之后的精度明显好于 Reranking 之前图像检索的精度。



图3 LULC 数据集

图4 检索精度比较 ($k=10$)图5 检索精度比较 ($k=20$)

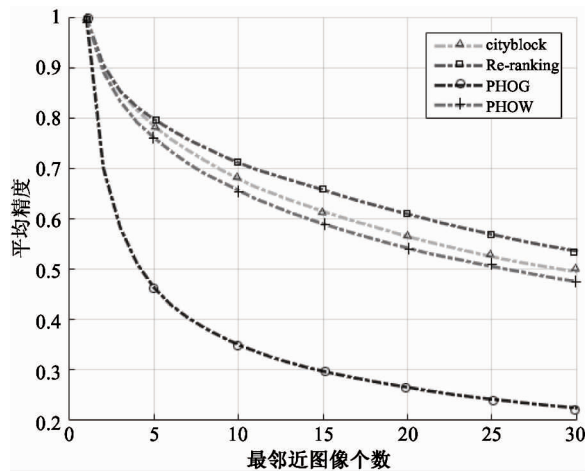


图 6 检索精度比较 ($k=30$)

4 结 论

本文提出了基于 ScSPM-Reranking 的遥感高分影像的检索方法。首先提取遥感影像的 SIFT 特征,然后基于 SIFT 特征,深层次挖掘图像的稀疏表达的特征,结合多层次的稀疏特征,最后获得遥感图像的 ScSPM 特征。然后,基于这些特征,首先使用 cityblock 进行初步检索,再进一步结合 Reranking 的方法获得高精度的检索结果。与其他相关方法对比,本文的方法检索的精度有了明显的提高。

参考文献

- [1] 戴芹,刘建波,刘士彬. 综合多特征遥感图像智能检索方法的概念设计. 地球信息科学学报,2011, 13(3): 401-408
- [2] Baraldi A, Bruzzone L. Classification of high spatial resolution images by means of a Gabor wavelet decomposition and a support vector machine. *Proceedings of the SPIE*, 2004, 5573: 19-29
- [3] Datcu M, Seidel K, Walessa M. Spatial information retrieval from remote-sensing images. Part, I: Information theoretical perspective. *IEEE Transactions on Geoscience and Remote Sensing*, 1998, 36(5): 1431-1445
- [4] 朱光良,陆丽珍. 遥感图像检索系统的设计与实现. 地球信息科学,2005,7(4): 15-19
- [5] 朱佳丽,李士进,万定生等. 基于特征选择和半监督学习的遥感图像检索. 中国图象图形学报,2011, 16(8): 1474-1482

- [6] 王利,杨征,李洋. 特征聚合的遥感图像数据库检索技术. 激光杂志,2011,37(6):78-81
- [7] 陈杰,邵权斌,邓敏等. 地物空间关系匹配的高分辨率遥感影像检索. 遥感学报,2016,20(3): 397-408
- [8] 陆丽珍,刘仁义,刘南. 一种融合颜色和纹理特征的遥感图像检索方法. 中国图象图形学报,2004, 9(3): 328-332
- [9] Aptoula E. Remote sensing image retrieval with global morphological texture descriptors. *IEEE transactions on geoscience and remote sensing*, 2014, 52(5): 3023-3034
- [10] Cheriadat A M. Unsupervised feature learning for aerial scene classification. *IEEE Transactions on Geoscience and Remote Sensing*, 2014, 52(1): 439-451
- [11] Prasad B G, Biswas K K, Gupta S K. Region-based image retrieval using integrated color, shape, and location index. *Computer Vision and Image Understanding*, 2004, 94(1): 193-233
- [12] Ko B C, Byun H. Integrated region-based image retrieval using region's spatial relationships. In: *Proceeding of the 16th International Conference on Pattern Recognition*, Quebec City, Canada, 2002, 1: 196-199
- [13] Lowe D G. Object recognition from local scale-invariant features. In: *Proceedings of the 7th IEEE International Conference on Computer Vision*, Corfu, Greece, 1999, 2: 1150-1157
- [14] Dalal N, Triggs B. Histograms of oriented gradients for human detection. In: *Proceeding of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Diego, USA, 2005, 1: 886-893
- [15] Lazebnik S, Schmid C, Ponce J. Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In: *Proceeding of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, New York, USA, 2006, 2: 2169-2178
- [16] Boiman O, Shechtman E, Irani M. In defense of nearest-neighbor based image classification. In: *Proceeding of the 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Anchorage, USA, 2008. 1-8
- [17] Yang J, Yu K, Gong Y, et al. Linear spatial pyramid matching using sparse coding for image classification. In: *Proceeding of the 2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Mi-

- ami, Florida, USA, 2009. 1794-1801
- [18] Shen X, Lin Z, Brandt J, et al. Object retrieval and localization with spatially-constrained similarity measure and k-nn re-ranking. In: Proceeding of the 2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Providence, USA, 2012. 3013-3020
- [19] Lee H, Battle A, Raina R, et al. Efficient sparse coding algorithms. In: Proceeding of the 20th Annual Conference on Neural Information Processing Systems, Vancouver, Canada, 2006. 801-808
- [20] Figueiredo M A, Nowak R, Wright S J, et al. Gradient projection for sparse reconstruction: application to compressed sensing and other inverse problems. *IEEE Journal of Selected Topics in Signal Processing*, 2007, 1(4): 586-597
- [21] Golub G H, Hansen P C, Oleary D P, et al. Tikhonov regularization and total least squares. *SIAM Journal on Matrix Analysis and Applications*, 1999, 21(1): 185-194
- [22] Van De Sande K E A, Gevers T, Snoek C G M. Evaluating color descriptors for object and scene recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, 32(9): 1582-1596
- [23] Yang Y, Newsam S. Bag-of-visual-words and spatial extensions for land-use classification. In: Proceeding of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems, San Jose, USA, 2010. 270-279
- [24] Yang Y, Newsam S. Spatial pyramid co-occurrence for image classification. In: Proceeding of the 13th International Conference on Computer Vision, Barcelona, Spain, 2011. 1465-1472
- [25] Van Gemert J C, Geusebroek J M, Veenman C J, et al. Kernel codebooks for scene categorization. In: Proceeding of the 10th European Conference on Computer Vision, Marseille, France, 2008. 696-709

Retrieval of high resolution remote sensing images based on ScSPM-Reranking

Gong Yongli*, Zhu Panpan**, Wang Yuebin**

(*Municipal and Road & Bridge Engineering College, Inner Mongolia Technical
College of Construction, Hohhot 010070)

(**School of Geography, Beijing Normal University, Beijing 100875)

Abstract

The study of remote sensing image retrieval, including image feature extraction and description of image similarity, was conducted to obtain the detailed information of the earth surface to provide reliable data for urban planning, environmental monitoring and disaster analysis. To improve the accuracy of image retrieval, the techniques of spatial pyramid matching (SPM) using sparse coding (Sc), called the ScSPM for short, as well as Reranking, were applied to the study, and a new method for retrieval of high resolution remote sensing images based on ScSPM and Reranking was proposed. The method uses ScSPM to extract the features of spatial scenes, then, uses the city-block distance to perform the preliminary retrieval of database images, and lastly uses the Reranking method to boost the accuracy of image retrieval. The proposed method was compared with other retrieval methods by experiment and the method using ScSPM-Reranking achieved the higher retrieval accuracy.

Key words: high resolution remote sensing images, image feature description, image retrieval, Reranking, sparse coding (Sc), spatial pyramid matching (SPM)