

利用 SVM 相关反馈和语义挖掘的遥感影像检索

刘婷婷^{1,2} 林 琿² 张良培³ 代汉青^{3,4}

(1 武汉大学中国南极测绘研究中心,武汉市珞喻路 129 号,430079)
(2 香港中文大学太空与地球信息科学研究所,香港新界沙田)
(3 武汉大学测绘遥感信息工程国家重点实验室,武汉市珞喻路 129 号,430079)
(4 黄冈市城乡规划局,黄冈市黄州区西湖三路 23 号,438000)

摘 要:针对语义鸿沟问题,将基于语义特征挖掘模型与支持向量机相关反馈方法相结合,建立了基于支持向量机相关反馈的人机交互遥感影像语义检索系统。实验结果表明,该方法利用高层语义特征及人机交互反馈信息缩小了语义鸿沟,提高了影像检索的精度。
关键词:影像检索;语义挖掘模型;相关反馈;SVM
中图法分类号:P237.3

随着对地观测技术的发展,遥感影像数据量日益庞大,影像的高效管理和快捷应用问题日益突出^[1, 2]。基于内容的影像检索技术的产生使解决这个问题成为可能。然而,用户视觉感知这类高层语义与光谱、纹理等底层影像特征之间存在着“语义鸿沟^[1]”,这在一定程度上影响了基于内容的影像检索技术的检索精度。近年来,基于内容影像检索的研究都围绕着如何缩小语义鸿沟展开^[3]。文献[1-5]通过利用概率统计等方法缩小了底层影像特征和高层语义特征之间的鸿沟。相关反馈通过人机交互的方式获得用户的反馈信息,并利用它进行再检索,这种方法能够在一定程度上缩小语义鸿沟^[6]。支持向量机(support vector machine, SVM)能够在有限训练样本的情况下,更好地控制机器学习的推广能力。

1 遥感影像检索系统

为了缩小语义鸿沟,使检索结果更贴近用户的要求,本研究在基于语义挖掘的遥感影像检索模型^[7]的基础上,建立了一个基于相关反馈的遥感影像语义检索系统。该系统结构如图 1 所示。
用户在利用该系统进行检索时,首先预览数

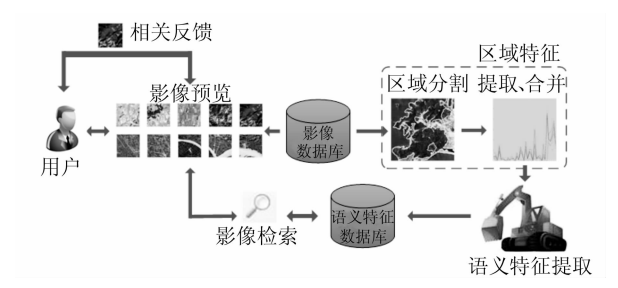


图 1 基于相关反馈的遥感影像语义检索系统
Fig. 1 SVM-based Remote Sensing Image Retrieval System

据库中影像并选取自己感兴趣的示例影像,通过语义特征挖掘等过程从数据库匹配出相似结果,再通过人机交互从初始检索结果中选择正、负样本进行相关反馈训练,获得最终的检索结果。
该语义挖掘模型可从底层影像特征中提取出隐含的高层语义特征,其过程可分为区域特征提取^[8]和语义特征提取两个阶段。

2 相关反馈

2.1 SVM 算法原理

SVM 方法通过在原空间或经投影后在高维空间中构造最优分类面,将给定的属于两个类别

的训练样本分开,构造超平面的依据是两类样本离超平面的距离最大化。

设线性可分样本集 (x_i, y_i) , 其中 $1 \leq i \leq N$, $x_i \in R^d, y_i \in \{-1, 1\}$ 为类别标号, 线性判别函数的一般形式为:

$$g(x) = wx + b \tag{1}$$

将 $g(x)$ 进行归一化, 使得距离分类面最近的样本点到分类面的距离为 1, 即必须满足:

$$y_i (wx_i + b) \geq 1 \tag{2}$$

那么, 离分类面最近的样本 $y_i (wx_i + b) = 1$, 则分类间隔为 $2 / \|w\|$ 。最优分类面是在满足式(2)约束条件下使分类间隔最大的超平面, 即要最小化 $\|w\|$, 目标函数为:

$$\min \varphi(w) = \frac{1}{2} \|w\|^2 = \frac{1}{2} (w w) \tag{3}$$

$w = \sum_i a_i y_i x_i, x_i$ 是位于分类间隔面上的样本, 这些样本被称为支撑样本, 其分类函数为:

$$f(x) = \text{sign}(wx + b) = \text{sign}\left(\sum_i a_i y_i x_i x + b\right) \tag{4}$$

对于线性不可分的情况, 通过非线性变换将输入控件变换为高维空间, 在新空间中求解最优分类面, 并引入核函数 K , 将线性可分情况下的点积运算变为 $K(x_i, x_j) = \varphi(x_i) \varphi(x_j)$, 从而实现了支持向量机在非线性分类中的应用。核函数的选择将直接影响到 SVM 分类器的泛化能力^[6]。高斯核函数具有较好的学习能力及实际应用性能^[9], 因此, 本研究选择高斯函数作为核函数。

2.2 基于 SVM 相关反馈的遥感影像语义检索

影像检索的相关反馈过程可以看作模式识别中的二分类问题, 因此, 可将 SVM 算法用于相关反馈的学习和检索过程。基于 SVM 相关反馈的遥感影像语义检索详细算法如下。

- 1) 用户从影像数据库中选出示例影像;
- 2) 根据语义特征, 利用欧氏距离的方法, 进行影像与示例影像间的相似性计算, 然后对欧氏距离计算结果排序, 并返回首次检索结果;
- 3) 用户从结果影像中选取正样本和负样本, 获得正样本集 I^+ 和负样本集 I^- ;
- 4) 准备 SVM 的训练样本集 (x_i, y_i) :
$$x_i \in I^+ \cup I^-, y_i = \begin{cases} +1, & x_i \in I^+ \\ -1, & x_i \in I^- \end{cases}$$
- 5) 利用 SVM 对训练样本进行学习并构造分类器 $f(x) = \text{sign}(\sum_i a_i y_i K(x_i, x) + b)$;
- 6) 利用分类器 $f(x)$ 对所有影像进行分类,

- 被分为正类的影像即为待选结果影像;
- 7) 计算每一幅待选结果影像与分类面的距离 $d(x) = \sum_i a_i y_i K(x_i, x) + b$;
- 8) 对 $d(x)$ 排序, 并再次返回检索结果。

3 实验与结果分析

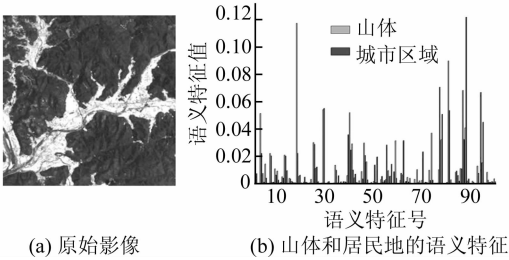
3.1 实验数据及环境

本研究利用数千幅大小 256 像素×256 像素的 Landsat TM 影像作为实验数据, 这些影像中包含海洋、植被、河流、居民地等地物。

3.2 语义特征提取实验

图 2(a) 为原始影像, 图 2(b) 展示了原始影像中山体和居民地的语义特征, 分别用黑色和灰色表示。其中, x 轴的语义特征号表示语义特征的维度, y 轴的语义特征值表示每一维特征的数值。

原始影像中山体和居民地表现出完全不同的光谱, 这在图 2(b) 中也被明显地反映。

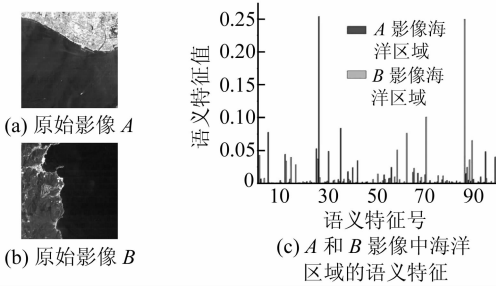


(a) 原始影像 (b) 山体和居民地的语义特征

图 2 不同区域及其语义特征

Fig. 2 Different Regions and Corresponding Semantic Features

图 3(a) 和 3(b) 为原始影像, 它们都包含海洋区域, 图 3(c) 展示了这两幅影像中海洋区域的语义特征, 分别用灰色和黑色表示。图 3(a) 和 3(b) 两幅影像中的海洋区域虽然呈现出相似的光谱特征, 但是其语义内涵却截然不同。图 3(a) 中海洋区域邻接着城市区域, 而图 3(b) 中海洋区域邻接着山体, 这种不同在图 3(c) 中被完全展现出来。



(a) 原始影像 A (b) 原始影像 B (c) A 和 B 影像中海洋区域的语义特征

图 3 相同区域及其语义特征

Fig. 3 Similar Regions and Corresponding Semantic Features

由以上两组实验可知,利用该语义挖掘模型提取出的语义特征不仅能够很好地代替光谱特征展示出不同种地物在光谱上的差异,而且能够展示出地物在不同环境中的语义差异。

3.3 不同反馈样本数检索性能实验

实验中,为 SVM 提供正、负训练样本,对其反馈学习性能进行比较。选取正、负样本总数分别为 6(正 3,负 3)、7(正 4,负 3)、8(正 4,负 4)、9(正 5,负 4)、10(正 5,负 5)、11(正 6,负 5)和 12(正 6,负 6),相应的检索结果查准率如图 4 所示。

由图 4 可以看出,检索结果查准率与样本数量成正比;由曲线的上升趋势也可以发现,当样本数为 11 时得到较高的检索精度并趋于稳定。由于在人机交互过程中,选择的样本数量越多,交互的复杂度越高,因此本研究将样本总量定为 11。

3.4 反馈前与反馈后对比检索实验

在该实验中,以图 5 中的影像为示例影像进行检索,该图中包含大片的海洋和小片的居民地及少量的植被。在语义特征的基础上,第一次检索结果如图 6 所示。

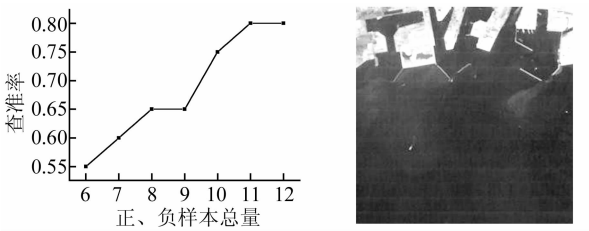


图 4 不同反馈样本数及检索查准率
Fig. 4 Different Training Samples and Retrieval Precision



图 5 示例影像
Fig. 5 Example of Image

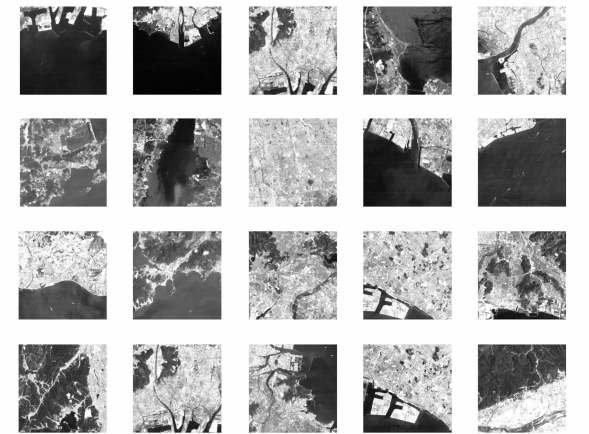


图 6 第一次检索结果影像(前 20 幅影像)
Fig. 6 First Round Image Retrieval Results

由图 6 可以看出,5 幅影像包含大片的海洋区域和小片的居民地及少量的植被,与示例影像

完全相同;7 幅影像包含海洋、居民地和少量植被,但居民地过大,与示例影像相似;4 幅包含大片的植被区域;4 幅只包含居民地和植被,未包含海洋。

根据第一次检索的结果,选择了 5 幅正样本(图 7)和 6 幅负样本(图 8)。其中,正样本包含大片的海洋区域、小片的居民地及少量的植被,负样本包含了大片的植被和居民地。



图 7 正样本
Fig. 7 Positive Samples

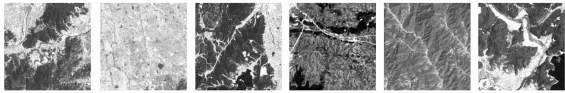


图 8 负样本
Fig. 8 Negative Samples

利用 SVM 方法对正样本和负样本进行学习并再检索,反馈结果如图 9 所示。

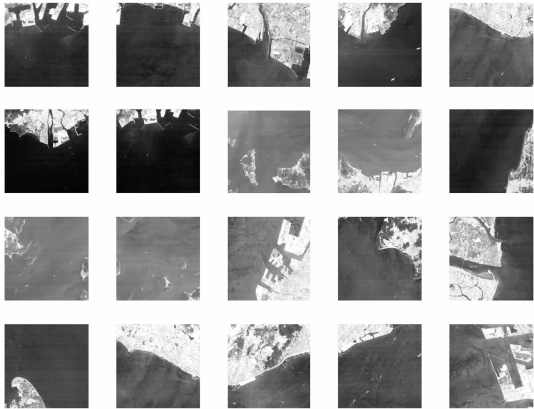


图 9 SVM 相关反馈结果影像(前 20 幅)
Fig. 9 Second Round Image Retrieval Results

从 SVM 方法进行相关反馈结果影像(图 9)可以看出,有 14 幅影像包含大片的海洋、小片的居民地和少量的植被,与示例影像完全相同;有 4 幅影像包含大片的海洋、少量的居民地和少量的植被,与示例影像相似;有 1 幅影像包含大片的海洋、极少量的居民地和植被。

最后,从影像数据库中选取 10 幅示例影像分别进行检索及相关反馈实验。10 次检索结果的查准率和查全率的均值如图 10 所示,通过对比发现,基于反馈的检索结果获得了较高的查准率和查全率,而没有反馈的检索结果则查准率及查全率均较低。因此,在用户选择的正样本和负样本的基础上,利用基于 SVM 的相关反馈,检索精度

大幅提高,检索的结果也更接近用户的需求。

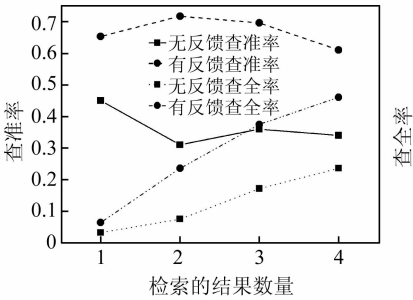


图 10 检索结果的平均查准率及查全率

Fig. 10 Average Precision and Recall of Retrieval Results

参 考 文 献

[1] Aksoy S, Koperski K, Tusk C, et al. Learning Bayesian Classifiers for Scene Classification with a Visual Grammar[J]. IEEE Transactions on Geoscience and Remote Sensing, 2005, 43:581-589

[2] 杜培军, 唐宏, 方涛. 基于内容的遥感影像检索若干问题的研究[J]. 中国矿业大学学报, 2005, 34(3): 270-278

[3] DatcuM, Seidel K. Human-Centered Concepts for Exploration and Understanding of Earth Observation Images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2005, 43(3):601- 609

[4] Datcu M, Daschiel H, Pelizzari A, et al. Information Mining in Remote Sensing Image Archives; System Concepts[J]. IEEE Transactions on Geoscience and Remote Sensing, 2003, 41(12):2 923-2 936

[5] Zhang R F, Zhang Z F. Effective Image Retrieval Based on Hidden Concept Discovery in Image Database[J]. IEEE Transactions on Image Processing, 2007, 16(2):562-572

[6] 张磊, 林福宗, 张钹. 基于支持向量机的相关反馈图像检索算法[J]. 清华大学学报(自然科学版), 2002, 429(1):80-83

[7] 刘婷婷, 李平湘, 张良培. 一个基于语义挖掘的遥感影像检索模型[J]. 武汉大学学报·信息科学版, 2009, 34(6):684-687

[8] Liu Tingting, Zhang Liangpei, Li Pingxiang, et al. Region Segmentation of Multi-spectral Remote Sensing Images Using Improved JSEG Algorithm [J]. Journal of Remote Sensing, 2009, 13(1):30-34

[9] 刘为. 基于内容图像检索关键技术的研究 [D]. 长春:吉林大学, 2010

第一作者简介:刘婷婷,博士,现从事数字图像处理和基于内容的影像检索等方面的研究。
E-mail:ttlui23@whu.edu.cn

SVM-relevance-feedback and Semantic-extraction-based RS Image Retrieval

LIU Tingting^{1,2} LIN Hui² ZHANG Liangpei³ DAI Hanqing^{3,4}

- (1 Chinese Antarctic Center of Surveying and Mapping, Wuhan University, 129 Luoyu Road, Wuhan 430079, China)
- (2 Institute of Space and Earth Information Science, the Chinese University of Hong Kong, Shatin, Hong Kong)
- (3 State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, 129 Luoyu Road, Wuhan, 430079, China)
- (4 Huanggang City Planning Bureau, 23 West Lake Road, Huanggang 438000, China)

Abstract: The semantic gap between high-level human perception and low-level image features becomes the bottleneck in content-based remotely sensed image retrieval technology. To solve this problem, in this research, a human machine interaction (HMI) remotely sensed image retrieval system is built that combines semantic mining model and SVM-based relevance feedback method. The experiments indicate that this method can well narrow semantic gap and improve retrieval precision and recall.

Key words: image retrieval; semantic mining model; relevance feedback; SVM

About the first author: LIU Tingting, Ph.D, she is engaged in image processing and content-based remote sensing image retrieval.
E-mail: ttlui23@whu.edu.cn