

基于多示例学习的遥感图像检索

葛永, 吴秀清, 洪日昌

(中国科学技术大学电子工程与信息科学技术系, 安徽合肥 230027)

摘要: 将基于多示例学习的相关反馈技术应用到遥感图像检索中. 为了避免局部最小值和减少冗余计算量, 对传统的多示例学习算法——多样密度算法进行了改进. 改进的算法利用用户标注的样本学习得到的查询概念, 指导下一轮检索. 为了提高查准率, 综合学习得到的查询概念, 提出了查询概念集来取代查询概念. 实验结果表明, 该算法比传统方法具有更好的检索性能.

关键词: 基于内容的遥感图像检索; 相关反馈; 多示例学习; 多样密度算法

中图分类号: TN911.73 **文献标识码:** A

Remote sensing image retrieval based on multiple instance learning

GE Yong, WU Xiu-qing, HONG Ri-chang

(Department of Electronic Engineering and Information Science, USTC, Hefei 230027, China)

Abstract: Relevance feedback (RF) based on multiple instance learning (MIL) was applied to the remote sensing image retrieval. In order to avoid the local minimum and reduce redundant computation, the traditional MIL algorithm—diverse density (DD) was improved. The improved DD algorithm learnt the query concept by employing the samples labeled by users, thereby facilitating the following query. In order to enhance the precision through synthesizing the query concepts learnt before the query concept set was proposed to substitute the query concepts. The experimental results show that this algorithm yields better performance than traditional approaches.

Key words: content-based remote sensing image retrieval; relevance feedback; multi-instance learning; diverse density algorithm

0 引言

随着影像传输卫星数量和可利用波段的增加以及遥感的商业应用发展, 遥感图像数据量不断增长. 为了快速有效地管理和查询这些遥感图像, 基于内容的遥感图像检索成为一个非常迫切的问题.

目前基于内容的图像检索(content-based image retrieval, CBIR) 在多媒体领域已经取得了大量的研究成果, 并且已经产生了一些成型的系统, 但是针

对遥感图像的 CBIR 研究相对较少. 通常的遥感图像检索方法^[1~4] 是通过提取纹理、颜色、频域谱特征等低层特征, 然后根据不同的相似度度量准则在特征层面上做相似度匹配, 最后输出检索结果. 由于当前图像理解技术的局限性, 提取的图像低层特征并不一定能很好地表达用户的查询概念, 所以检索的结果往往不能满足用户的查询需求. 这也是 CBIR 系统的低层特征和高层语义之间的鸿沟问题^[5]. 为了克服这个问题, 在多媒体领域, 有学者已经提出了

收稿日期: 2007-04-27; 修回日期: 2007-06-08

基金项目: 国家高新技术研究发展计划(2004AA783052)资助.

作者简介: 葛永, 男, 1982年生, 硕士生. 研究方向: 图像检索, 机器学习, 计算机视觉. E-mail: ygesi@mail.ustc.edu.cn

通讯作者: 吴秀清, 教授. E-mail: wuxq@ustc.edu.cn

一些基于高层语义的检索方法^[6]. 在遥感图像检索领域,陆丽珍^[7]提出了基于GIS语义的遥感图像检索算法. Sun等^[8]通过运用ontology和grid知识,提出了一种基于语义的遥感图像检索方法. 由于这两种算法提取的高层语义都有一定针对性,算法通用性不够. 同样为了克服低层特征和高层语义之间的鸿沟,基于机器学习算法的相关反馈技术^[9,10]被集成到检索系统中. 针对遥感图像检索,相关反馈技术的研究非常少. 文献^[11]提出了基于径向基函数神经网络的相关反馈技术. 算法将每次反馈用户标注的图像作为训练样本,不断地学习神经网络的系数,再利用学习得到的神经网络来对图像进行分类.

为了克服低层特征和高层语义之间的鸿沟,本文采用基于多示例学习(multi-instance learning, MIL)算法^[12,13]的相关反馈技术来逐步地学习用户查询概念,以达到提高检索精度的目的. 本文改进了多示例学习方法——多样密度学习算法(diverse density, DD),并把它应用到相关反馈中,避免求解时获得局部最小值的同时也降低了计算的复杂度. 此外,本文提出了查询概念集和相应的相似度量准则,并将其用于下一次的相似匹配,进一步提高了查准率. 实验结果表明,该算法在提高检索精度的同时,也降低了相关反馈的时间. 系统总体性能优于传统的遥感图像检索方法.

1 遥感图像检索系统

一般来讲,用户提供的查询图像 Q 只是包含用户感兴趣的区域,尺寸很小. 而图像库中图像 I 尺寸很大,而且图像内容具有歧义性,即一幅图像可能包含多个用户感兴趣的区域. 多示例学习的方法可以较好地处理这种歧义性. 如图1所示,首先对库中图像分割,提取区域特征,每个区域特征对应一个示例,整幅图像的所有区域特征对应一个包. 检索阶段,提取用户提供的查询图像 Q 的特征,做初次检索,返回检

索结果;用户根据检索结果进行人工标注,标注分为正、负两类,然后反馈给系统;系统从用户标注中获得正、负训练样本,利用改进的多样密度算法学习得到查询概念(特征点);系统根据特征点重新检索,然后返回检索结果. 从图1可以看出,整个检索是一个迭代过程,直到用户对返回的检索结果满意为止.

在初次检索阶段,由于没有训练样本,学习算法不可用. 本文用一种简单的基于距离的相似度量方法去度量 Q 和 I 的相似性. 假设 I 包含 n 个区域,或者说 I 对应的包包含 n 个示例,记作 $I = \{I_1 I_2 \dots I_n\}$. Q 和 I 之间的距离定义为

$$D(Q, I) = \min_{1 \leq i \leq n} \{ \|Q - I_i\| \} \quad (1)$$

式中, $\|Q - I_i\|$ 为 Q 对应的特征向量和 I_i 对应的特征向量之间的欧式距离. 在每次学习得到特征点后重新检索时,也用式(1)去度量特征点和图像库中图像的相似度.

2 多示例学习和包的生成

使用多示例学习时,首先要构造合适的包,这样才能保证多示例学习算法的有效性. 针对遥感图像的自身特点,本文提出了一种简单而有效的包构成方法. 遥感图像尺寸一般很大,而用户感兴趣的只是其中若干局部区域,直接提取整幅图像的全局特征根本无法反映感兴趣的目标概念,故有必要对整幅图像进行分解或者分割. 针对遥感图像,现有的基于块的分割主要有四叉树和九叉树算法^[16],这两种方法虽然具有分解子图像树木较少、检索效率快的特点,但是都有可能过分割感兴趣区域. 而重叠分割算法可以避免对感兴趣区域的过分割,保证其完整性. 在相同分辨率下,同类用户感兴趣区域大小一般差别不大,可以把整幅图分成一系列固定大小的区域,每个区域的尺寸为感兴趣区域尺寸的2倍,而且相邻区域有1倍感兴趣区域尺寸的重复. 经过以上处理后,每幅图像就分解为若干区域. 再通过提取每个区域的有效特征就构造了对应的示例,整幅图像的所有区域对应的示例在一起就构成了一个包.

在CBIR检索系统中,主要有四类有效特征:颜色、纹理、形状、空间位置信息. 选择合适的图像视觉特征是有效检索图像的关键. 在以上四类特征中,纹理具有不依赖颜色和亮度、反映图像中结构组织排列的特点,能很好地表达遥感图像的特征. Gabor^[14]小波变换能够最大程度地减少空间和频率的不确定性,可以方便地用来考察图像纹理的尺度和频率特

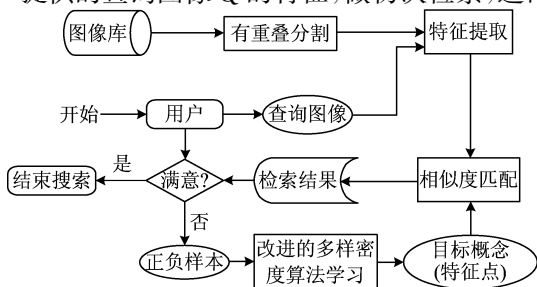


图1 图像检索系统框架

Fig. 1 Framework of the image retrieval system

性. 本文采用 Gabor 小波变换提取各个区域的纹理特征, Gabor 小波变换的基小波是 Gabor 函数, 其母小波的冲击响应 $f(x, y)$ 为

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)\right] \cos(2\pi\mu_0 x) \quad (2)$$

对母小波 $f(x, y)$ 进行适当的尺度变换和旋转变换 $f'_{m,n}(x, y) = k^{-m} f(k^{-m}x', k^{-m}y')$, 其中

$$\left. \begin{aligned} x' &= x\cos\theta + y\sin\theta \\ y' &= -x\sin\theta + y\cos\theta, \theta = n\pi/N \end{aligned} \right\} \quad (3)$$

式中, $m = 0, 1, \dots, M-1, n = 0, 1, \dots, N-1, M$ 和 N 分别为尺度数和方向数. 这样可以得到一组自相似的滤波器, 即为 Gabor 小波. 本文的 M 取 3, N 取 4, 共使用 12 组滤波器.

确定滤波器之后, 在计算各滤波器与区域卷积能量值的基础上, 计算各个区域能量值的均值和方差, 并将其作为区域的纹理特征向量. 记为

$$\mathbf{F}^{\text{Gabor}} = \{u_{(0,0)}, v_{(0,0)}, u_{(0,1)}, v_{(0,1)}, \dots, u_{(m,n)}, v_{(m,n)}, \dots, u_{(M-1,N-1)}, v_{(M-1,N-1)}\} \quad (4)$$

$u_{(m,n)}$ 为滤波后的区域能量均值, $v_{(m,n)}$ 为方差. 公式表示为

$$\left. \begin{aligned} u_{(m,n)} &= \frac{\sum_x \sum_y E_{(m,n)}(x, y)}{k \times k} \\ v_{(m,n)} &= \sqrt{\frac{\sum_x \sum_y (E_{(m,n)}(x, y) - u_{(m,n)})^2}{k \times k}} \end{aligned} \right\} \quad (5)$$

式中, $E_{(m,n)}(x, y)$ 为第 (m, n) 个滤波器在 (x, y) 处的滤波能量值, $k \times k$ 为区域大小. 这样每个区域就可以用 24 维纹理特征向量表示, 将特征向量作为示例, 整幅图有多少区域就有多少个示例, 所有示例放在一起构成一个包.

3 多样密度算法及改进

3.1 多样密度算法

假设训练样本中共有 m 个正包和 n 个反包, 正包和反包分别用 B_i^+ 和 B_i^- 表示, 正包和反包中的第 j 个示例用 B_{ij}^+ 和 B_{ij}^- 表示. 假设要寻找的特征空间中的正确目标点是个单点 t , 那么它应该满足

$$\arg \max_t \Pr(t / B_1^+ B_2^+ \cdots B_m^+, B_1^- B_2^- \cdots B_n^-) \quad (6)$$

假设特征空间中的所有点具有相同的先验概率, 使用贝叶斯法则, 同时假设给定目标概念 t 时各个包是条件独立的. 式(6)等价于

$$\arg \max_t \prod_i \Pr(t / B_i^+) \prod_i \Pr(t / B_i^-) \quad (7)$$

式(7)就是最大化多样密度的定义. Maron 和 Lozano 使用 noisy-or 模型对式(7)进行了具体化计算. 其中

$$\left. \begin{aligned} \Pr(t / B_i^+) &= 1 - \prod_j (1 - \Pr(B_{ij}^+ = t)) \\ \Pr(t / B_i^-) &= \prod_j (1 - \Pr(B_{ij}^- = t)) \end{aligned} \right\} \quad (8)$$

并假设

$$\Pr(B_{ij} = t) = \exp(-\|B_{ij} - t\|^2) \quad (9)$$

式中, $\|B_{ij} - t\|^2$ 是示例 B_{ij} 和目标点 t 之间的欧式距离.

从最大化多样密度的定义可以看出, 当特征空间特征向量很多时, 无法通过遍历整个特征空间来搜索多样密度最大的点. 为此, 多样密度算法使用了多起点的梯度下降法进行搜索. 每轮搜索时, 从特征空间的一点开始, 沿使得多样密度增长最快的方向调整目标概念, 直到达到具有多样密度极值的点. 这种梯度下降方法无法避免陷入局部极小值, 但是由于正包中至少存在一个相当靠近多样密度最大的点, 因此, 文献[13]分别从正包的每个示例开始, 使用梯度下降法寻找极值. 这种方法一定程度上减少了陷入局部极值的概率, 但是仍然会因为从每个包中选择的示例不合适而得到局部极小值. 因此只是随机避免陷入局部极小值. 另外, 在每轮搜索阶段, 式(8)的计算都用到了每个包的所有示例, 当包中含有的示例很多时, 计算量将很大. 而且, 当示例 B_{ij} 和目标点 t 之间的欧式距离很大时, $\Pr(B_{ij} = t)$ 的值就会趋近于 0, 这样的话, 在计算式(8)时, 对结果就影响很小, 冗余计算很大. 针对上面的问题, 本文对文献[12]中的多样密度算法进行了改进.

3.2 改进的多样密度算法

从用户提供的查询图像中, 可以方便地得到一个接近最大多样密度的特征点, 把这个特征点称作假设目标点, 记作 h . 首先利用这个假设目标点从每个包中挑出 k 个与 h 相似度最大的示例, 然后用这 k 个示例代表整个包. h 和每个示例的相似度用 h 与每个示例的欧式距离的倒数衡量, 即 $S(h, B_{ij}) = 1 / \|h - B_{ij}\|^2$. 这样, 式(8)就变为

$$\left. \begin{aligned} \Pr(t / B_i^+) &= 1 - \prod_{1 \leq j \leq k} (1 - \Pr(B_{ij}^+ = t)) \\ \Pr(t / B_i^-) &= \prod_{1 \leq j \leq k} (1 - \Pr(B_{ij}^- = t)) \end{aligned} \right\} \quad (10)$$

本文 k 的取值为 5. 通过上面的改进可以看出: (I) 由于每个包中除了 k 个示例保留外其他都被丢弃, 这样就大大减小了每轮搜索的计算复杂度. 同时由上面的分析可知, 仅用每个包的 k 个示例, 对式

(10)的结果影响也很小;(II)因为每个正包中保留的 k 个示例都和 h 很相似,当仍然像文献[13]那样从每个正包的每个示例开始进行搜索时,就必然在很大程度上避免了陷入局部极小值. 每轮学习得到的查询概念可以作为下一轮学习的假设目标点(h).

3.3 特征点集

受查询点移动方法(query-point movement method)^[15]的启发,为了进一步提高检索的精度,利用前面学习得到的查询概念(特征点)组成特征点集,下次计算相似度时用查询概念集取代学习得到的查询概念,进而不断地优化检索结果. 假设用户共进行了 K 次反馈,每次反馈学习得到的特征点用 t_l ($l=1,2,\dots,K$)表示. 那么在第 j 次学习之后就共有 j 个特征点,特征点集就是 $T_j = \{t_1, t_2 \dots t_j\}$. 在下次相似度匹配时,度量图像库中图像 $I(I = \{I_1 I_2 \dots I_n\})$ 和特征点集的相似性如下

$$D(T_j, I) = \min_{1 \leq l \leq j} \{ \min_{1 \leq n \leq n} (\|t_l - I_n\|) \} \quad (11)$$

4 实验结果

实验使用的图像为 5~10 m 分辨率的卫星遥感图像. 本文共选择了 4 类图像,包括城镇、林地、水域、山地,总共 200 幅图像作为待检索的图像库. 另

外每类选择了 5 幅共 20 幅图像,作为查询图像,这些图像并非直接从图像库中获得. 其中图像库中图像尺寸较大,从 1 500×1 500 像素到 3 000×3 000 像素都有. 查询图像尺寸较小,一般为 200×200 像素. 对于每幅查询图像,用户共进行 5 次反馈,包括初次查询系统总共进行了 6 次查询. 每次查询系统返回前面最相似的图像. 图 2 给出了一个查询的示例(为了节省版面,只缩小显示了第一次反馈后得到的前 10 幅图像). 查询图像为用户感兴趣区域,图 2(b)为系统一次反馈后检索得到的前 10 幅图像,相似性自左向右、自上向下依次递减. 通过相关反馈技术,系统学习用户查询意图,实现了更高的查准率. 但是由于特征的单一性,检索结果还是存在不相关图像.

本文采用传统的查准率^[13]来评价系统的性能,即用系统返回的检索结果图像中相关图像所占的百分比表示

$$\text{查准率} = \frac{\text{返回的相关图像个数}}{\text{返回的图像个数}} \times 100\% \quad (12)$$

对所有查询图像的查准率取平均得到平均查准率,本文实验中的查准率就是指平均查准率.

为了说明改进的多样密度算法检索遥感图像的有效性,将其与传统的多样密度算法进行了比较,如图 3 所示. 由图 3 可以看出,由于改进的算法避免了

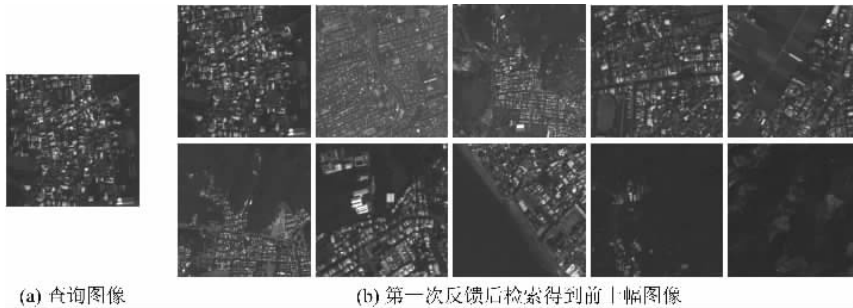


图 2 查询示例

Fig 2 An example of query

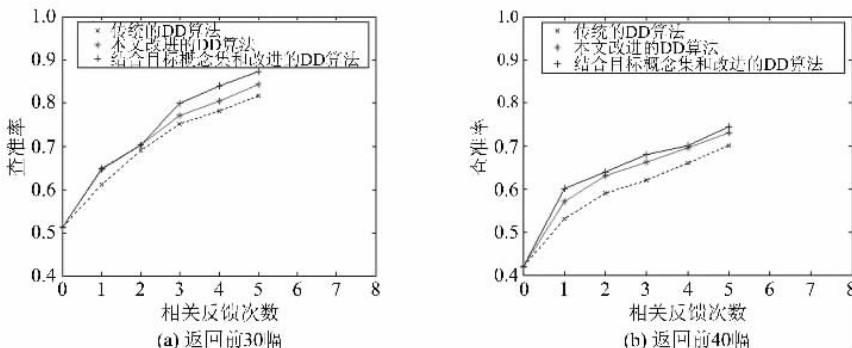


图 3 不同多示例学习算法的查准率对比

Fig. 3 Accuracy comparison of different MIL algorithms

局部最小值的出现,从而学习得到的特征点更准确,所以检索的精度有所提高.通过本文提出的特征点集,综合学习得到的特征点后,检索精度有了进一步的提高.从 3.2 节的分析可以看出,每次学习特征点的计算复杂度较传统算法也有很大的降低,从而减少了每次相关反馈的时间.表 1 列出了在 Celeron P4 2.4 G,512 M 内存,Matlab7.0 环境下改进算法和传统算法平均计算时间的比较.

表 1 平均计算时间对比

Tab. 1 Average computing time comparison

方法	一次学习耗时/s	一次相关反馈耗时/s
传统 DD	0.472 825	1.499 113
改进 DD	0.044 954	1.071 242

基于多示例学习算法的相关反馈技术有效地提高了遥感图像检索系统的查准率,本文与文献[3,4]的方法进行了比较,查准率如图 4 所示.由图 4 可以看出,只要一次反馈,本文系统的检索性能就比文献[3,4]的有很大的提升.

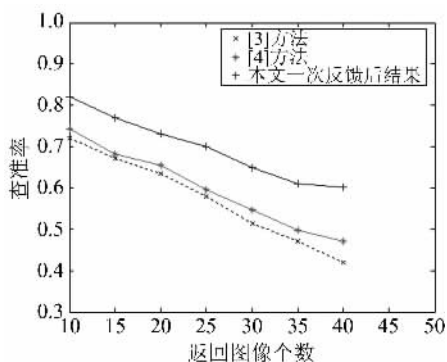


图 4 不同检索系统的查准率对比

Fig. 4 Accuracy comparison of different retrieval systems

5 结论

针对遥感图像检索,本文提出了一种基于改进的多示例学习算法的相关反馈技术,并对传统多示例学习算法进行了改进.在反馈阶段提出了目标概念集和相应的相似度量准则.实验证明,算法在大大提高了查准率的同时,对系统时间性能也有很大提升.

参考文献(References)

[1] Cheng Q M, Yang C J, Shao Z F, et al. A prototype system of content-based retrieval of remote sensing images [C]// Geoscience and Remote Sensing Symposium, Toulouse, France: IEEE Press, 2003, 6: 3 700-3 702.
[2] 牛蕾,倪林, Miao Yuan. 基于 ROI 的压缩域多谱段遥感

图像的检索[J]. 中国图象图形学报, 2005, 10(10): 1 212-1 217.

[3] Zhu B, Ramsey M, Chen H. Creating a large-scale content-based airphoto image digital library[J]. IEEE Transactions on Image Processing, 2000, 9(1):163-167.
[4] 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[J]. 中国图象图形学报, 2004, 9(3): 328-333.
[5] Smeulders A M W, Worring M, Santini S, et al. Content-based image retrieval at the end of the early years[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(12): 1 349-1 380.
[6] Jiang W, Chan K L, Li M J, et al. Mapping low-level features to high-level semantic concepts in region-based image retrieval [C]//Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington: IEEE Computer Society, 2005, 2: 244-249.
[7] 陆丽珍. 基于 GIS 语义的遥感图像检索[J]. 中国图象图形学报, 2005, 10(10): 1 207-1 211.
[8] Sun H, Li S X, Li W J, et al. Semantic-based retrieval of remote sensing images in a grid environment[J]. IEEE Geoscience and Remote Sensing Letters, 2005,2(4):440-444.
[9] Jing F, Li M J, Zhang H J, et al. Relevance feedback in region-based image retrieval[J]. IEEE Transactions on Circuits and Systems For Video Technology, 2004, 14(5): 672-681.
[10] 吴洪,卢汉清,马颂德. 基于内容图像检索中相关反馈技术的回顾[J]. 计算机学报, 2005, 28(12): 1 969-1 979.
[11] Schröder M, Rehrauer H, Seidel K, et al. Interactive learning and probabilistic retrieval in remote sensing image archives[J]. IEEE Transactions on Geoscience and Remote Sensing, 2000, 38(5): 2 288-2 298.
[12] Maron O, Lozano-Pérez T. A Framework for Multiple-Instance Learning[M]. Cambridge, MA: MIT Press, 1998.
[13] Zhang Q, Goldman S A. EM-DD: an improved multiple-instance learning technique [C]// Neural Information Processing Systems. Cambridge, MA: MIT Press, 2002: 1 073-1 080.
[14] Manjunath B S, Ma W Y. Texture feature for browsing and retrieval of image data [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1996, 18(8): 837-842.
[15] Rui Y, Huang T S. A novel relevance feedback technique in image retrieval [C]//Proceedings of the 7th ACM International Conference on Multimedia (part2). New York: ACM Press,1999: 67-70.
[16] Sheikholeslami G, Zhang A D. A multi-resolution content-based retrieval approach for geographic images [J]. Geoinformatica, 1999, 3(2):109-139.