

结合空间信息的 PTSVM 的遥感图像变化检测

高桂荣 严威 夏晨阳 吴国宝

(河海大学计算机与信息学院 南京 210096)

摘要: 基于半监督支持向量机算法的遥感图像变化检测中, PTSVM 算法采用成对标注的法则对无标签样本进行标注, 即一次只标注一对无标签样本, 取得了很好的检测结果, 一定程度上提高了检测精度。但是 PTSVM 算法只利用了样本的光谱信息, 而样本的特征信息不仅包括光谱信息, 还有空间信息, 空间信息作为样本的基本特征之一, 对样本同样重要, 所以本文在 PTSVM 的基础上, 提出了结合样本空间信息和光谱信息的 PTSVM 算法, 并用于遥感图像的变化检测中。通过实验结果分析, 该方法在遥感图像的变化检测中取得了很好的效果, 进一步提高了变化检测的准确率。

关键词: PTSVM; 成对标注; 光谱信息; 空间信息; 变化检测

中图分类号: TN79 **文献标识码:** A **国家标准学科分类代码:** 510.1050

PTSVM combined with spatial information for remote sensing image change detection

Gao Guirong Yan Wei Xia Chenyang Wu Guobao

(School of Computer and Information, Hohai University, Nanjing 210096, China)

Abstract: In the algorithms of change detection of remote sensing image based on semi-supervised SVM, PTSVM algorithm uses paired labeling rule that once marks only one pair of unlabeled samples to label unlabeled samples, which achieves good results in the change detection of remote sensing image and to some extent, improves the detection accuracy. But PTSVM algorithm uses only spectral information of samples. Sample characteristics information includes not only the spectral information, but also spatial information. Spatial information as one of the basic characteristics of the sample is equally important, so this paper proposes the improved PTSVM algorithm combined spatial information of samples with spectral information of samples in change detection for remote sensing images. Through the experimental results analysis, the method has achieved good results in change detection of remote sensing image, and further improves the change detection accuracy.

Keywords: PTSVM; paired label; spectral information; spatial information; change detection

1 引言

通过卫星等获得的遥感图像来获得地物变化信息的过程称之为变化检测^[1]。变化检测技术, 在我国的经济过程中越来越重要^[2], 研究者通过变化检测技术, 有效地检测出在一段时间内地表的变化, 然后通过研究结果, 对某一个地区的经济发展做出规划。在变化检测理论中, 基于遥感图像的变化检测技术已经成为各个领域研究中的热点, 在经济的各个方面得到了很好的发展和应用。

自从卫星技术得到发展并运用于航空事业之后, 我国就利用卫星技术来获得地表不同时期的图片, 然后研究者

们利用变化检测技术, 比对两个不同时期的图片, 对地表变化做出尽可能正确的推测。变化检测技术发展至今, 已经涉及生活的各个领域, 检测技术也多种多样, 主要的方法有非监督变化检测、监督变化检测和半监督的变化检测。非监督变化检测是指人们不需要任何先验知识, 只是根据获取的图像的光谱特征的分布规律, 来确定各类别的属性^[3]。监督变化检测是通过具有标记的训练样本进行学习, 并对训练样本中的每个样本提供类别标记^[4], 以尽可能正确地对训练样本集之外的示例进行预测^[5]。支持向量机就是一种基于监督学习的新型机器学习方法^[6]。该方法具有较强的泛化能力, 将它用于遥感图像的变化检测中, 取得了很好的效果^[7]。但是在监督学习的过程中为了获得

足够的训练样本,需要手工对大量的样本进行标记,成本高,耗时长。因此,研究人员为了满足实际需要,提出了半监督学习方法^[8],并用于支持向量机的训练过程中,即半监督支持向量机。在此基础上,最终研究出了利用渐进直推式学习算法进行学习训练样本的支持向量机分类器。但是 PTSVM 算法只利用了样本的光谱信息^[9],而样本的空间信息在遥感图像的变化检测中同样重要,所以本文提出了结合样本空间信息和光谱信息的 PTSVM 算法并用于遥感图像的变化检测中。通过实验结果分析,该方法是可行的和有效的,进一步提高了变化检测的精度和准确率。

2 理论基础

2.1 直推式支持向量机(TSVM)

由于在获取样本的过程中,实际存在着大量的未标定样本集 X,而小样本训练集的获取又比较困难。为了解决这一问题,提出了直推式支持向量机(TSVM)。直推式支持向量机是在 SVM 算法的基础上发展起来的^[10],它既保留了支持向量机的优点,又弥补了 SVM 的不足。TSVM 算法在迭代过程中,将伪标签赋予未标定的无标签样本,从而使得未标定样本变为伪标定数据,然后将初始的有标签样本和伪标定数据混合,通过对混合训练集进行迭代运算,从而得到更可信的分类超平面将数据分开,此时二次优化问题可以描述为:

$$s. t. \left\{ \begin{array}{l} \forall_{g=1}^n : y_g [\omega^T \cdot \varphi(x_g) + b] \geq 1 - \xi_g, \xi_g > 0 \\ \forall_{h=1}^d : y_h^* [\tau \omega^T \cdot \varphi(x_h^*) + b] \geq 1 - \xi_h^*, \xi_h^* > 0 \end{array} \right\}$$

$$\varphi(\omega(y_1^*, \dots, y_d^*)) = \min_{\omega, \xi, \xi^*} \left\{ \begin{array}{l} \frac{1}{2} \omega^T \cdot \omega \\ + C \sum_{g=1}^n \xi_g \\ + C^* \sum_{h=1}^d \xi_h^* \end{array} \right\} \quad (1)$$

对应的半标定样本的参数为 ξ_h^* 和 $2C^*$, 二次优化问题转化为:

$$L(\alpha, \alpha^*) = \sum_{g=1}^n \alpha_g + \sum_{h=1}^d \alpha_h^* - \frac{1}{2} \left(\sum_{g=1}^n \sum_{i=1}^n k(x_g, x_i) y_g y_i \alpha_g \alpha_i + 2 \sum_{g=1}^n \sum_{h=1}^d k(x_g, x_h^*) y_g y_h^* \alpha_g \alpha_h^* + \sum_{h=1}^d \sum_{j=1}^d k(x_h^*, x_j^*) y_h^* y_j^* \alpha_h^* \alpha_j^* \right) \quad (2)$$

$$s. t. \left\{ \begin{array}{l} 0 \leq \alpha_g \leq C, 1 \leq g \leq n \\ 0 \leq \alpha_h^* \leq C, 1 \leq h \leq d \\ \sum_{g=1}^n y_g \alpha_g + \sum_{h=1}^d y_h^* \alpha_h^* = 0 \end{array} \right.$$

式中:半标定数据的拉格朗日算子为 α^* 。经过迭代,拉格朗日算子会被确定,最终直推式支持向量机的决策函数变为:

$$y(x) = \text{sgn} \left[\sum_{g=1}^n \alpha_g y_g k(x, x_g) + \sum_{h=1}^d \alpha_h^* y_h^* k(x, x_h^*) + b \right]$$

2.2 渐进直推式支持向量机

尽管直推式支持向量机得到了广泛的运用,但是在运用于遥感图像的变化检测中,检测速度慢,变化检测的准确率低^[11]。为了进一步改善变化检测的准确率,在 TSVM 的基础上,陈毅松等人在文献^[7]提出了渐进直推式支持向量机。他的理论弥补了 TSVM 的不足,同时也提出了新的观点,他的理论在支持向量机的发展中做出了很大的贡献,该算法的提出是变化检测技术的又一次进步^[12]。在该算法中,有标签样本的分布信息不作为估计整个样本中正负样本比例的依据,而是采用渐进的方法确定无标签样本的类别。

PTSVM 算法在每一次训练完成之后,按照下面的式(4)标注一个新的正标签,同时按照式(5)标注一个新的负标签。

$$\text{Max} \|f(x_j^*)\|, s. t. 0 < f(x_j^*) < 1. \quad (4)$$

$$\text{Min} \|f(x_j^*)\|, s. t. -1 < f(x_j^*) < 0. \quad (5)$$

如果不满足式(4)和(5),则不添加正标签和负标签,算法结束,输出结果。

3 本文所提算法

考虑到如果样本点的区域发生变化,该样本点发生变化的可能性就比较大,而样本点周围的邻域对其是否发生变化会有影响的。例如一个区域单个样本点周围邻域全部都是确定变化的,那么该样本点发生变化的概率较大。基于这种思想,在 PTSVM 算法的基础上,结合样本的空间信息和光谱信息对遥感图像进行变化检测。即在 PTSVM 算法进行变化检测的过程中,通过在目标函数中加入样本的空间信息,得到最可靠的分类面,选择最有把握的无标签样本,从而提高变化检测的精度和准确率:

样本的空间信息如下式:

$$d_{ij} = \sum_{i=1}^{S_n} K(p_j, p_i) \quad (6)$$

式中: $K(\cdot)$ 是核函数,在本文中采用高斯核函数,则:

$$K(p_j, p_i) = \exp(-\|p_j - p_i\|^2 / \lambda^2) \quad (7)$$

由上述可知核函数没有被用在特征领域而是空间领域内,参数与核函数的宽度有关。

在目标函数中加入样本的空间信息,如式(8)所示。

$$\varphi(\omega) = \min_{\omega, \xi, \xi^*} \left\{ \begin{array}{l} \frac{1}{2} \omega^T \cdot \omega + C \sum_{g=1}^n \xi_g + C^* \sum_{h=1}^d \xi_h^* \\ + \sum_{i=1}^c \sum_j^m h_{ij}^m d_{ij}^2 \\ + \frac{\alpha}{N_R} \sum_{i=1}^c \sum_j^m h_{ij}^m \sum_{r \in N_i} (1 - h_r)^m \end{array} \right\} \quad (8)$$

根据选取的像素点的邻域大小选择合适的加权项,如果像素点同他周围的像素点属于相同种类,引入拉格朗日

乘法得到式(8)的极小值,对其进行求导:

$$g_{ij} = \frac{\{d_{ij}^2 + \frac{\alpha}{N_R} \sum_{r \in N_i} (1 - u_{jr})^m\}^{\frac{1}{m}}}{\sum_{j=1}^c \{d_{ij}^2 + \frac{\alpha}{N_R} \sum_{r \in N_i} (1 - g_{jr})^m\}^{\frac{1}{m}}} \quad (9)$$

由此可知,邻域内像素点属于相同种类的隶属度是最大值,这时改进的目标函数的值是最小的,因此样本点周围的样本点同属于相同的类别的概率变大。

光谱信息和空间信息作为无标签和有标签样本的基本

特征信息,很好地利用这些样本信息来检测样本的变化。将它们用在选择无标签样本时,能够唯一标识某一个无标签样本,将它们加入到样本的训练过程中,能够更有效地选择最有把握的无标签样本进行标记,然后添加到与有标签样本进行混合训练,这一过程持续下去,直到某次训练后,在当前最优分割平面的边界区域中,所有的无标签样本都不出现。

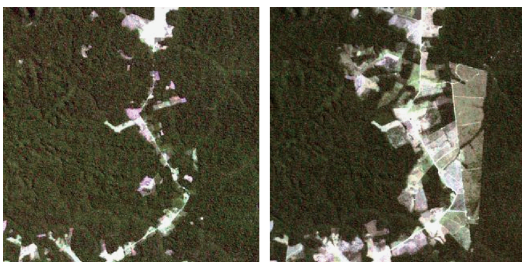
结合空间信息的 PTSVM 算法在遥感图像变化检测中的实现步骤如表 1 所示。

表 1 结合空间信息的 PTSVM 算法

算法:	结合空间信息的 PTSVM 算法的遥感图像变化检测
输入:	两幅遥感影像;
第一步:	指定初始有标签样本和无标签样本的影响因子 C 和 C ₀ ,在最初的有标签样本集上,训练得到 SVM 分类器;
第二步:	对位于当前分类超平面一定边界区域内的无标签样本,利用成对标注的法则进行标注;
第三步:	在目标函数中引入样本的空间信息和光谱信息,并对混合样本集进行训练,然后用当前得到的判别函数计算每一个无标签样本的判别函数输出;
第四步:	若无满足条件的无标签样本,则对剩下的所有无标签样本全部分类,算法结束,并输出变化检测结果;
输出:	变化检测精度。

4 实验结果与分析

本次实验采用南美洲流经巴西地区的亚马逊流域的两幅 Landsat TM(thematic mapper)遥感影像做为实验对象,该区域主要由河流、森林及砍伐后的裸地 3 部分组成,本实验对其河流区域的森林砍伐面积和地表沙漠化面积进行变化检测,两幅遥感影像的获取时间分别为 1997 年 7 月和 2003 年 7 月(如图 1),实验区大小为 400 像元×400 像元,如图 1 所示。



(1) 2000年7月

(2) 2006年7月

图 1 Landsat TM 数据

首先对遥感图像进行预处理,即对遥感图像进行辐射校正和配准。然后,为了比较新算法的优劣,选取 TSVM 和 PTSVM 两种变化检测算法与本文所提的新算法(改进的 PTSVM 算法)做进一步的对比分析。如图 2 所示。

由上图结果可以看出,结合空间的 PTSVM 算法与 TSVM 和 PTSVM 算法相比,大大提高了变化检测的准确

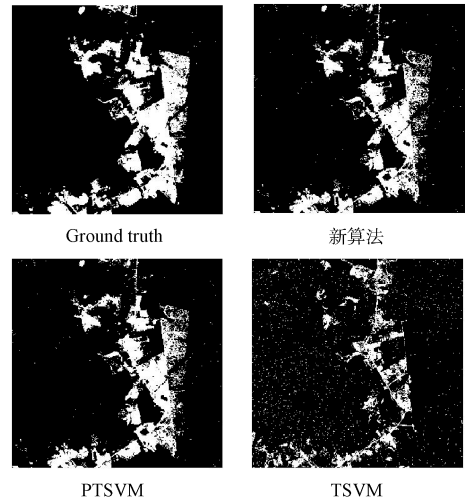


图 2 变化检测结果图以及标准参照图

率,使遥感图像变化检测的精度进一步提高,因此本文所提算法进一步改善了遥感图像变化检测的效果。

各变化检测算法随无标签样本数的变化而变化的变化结果图,如图 3 所示。

由图 3 可以看出各变化检测算法随无标签样本的增加,变化检测的误检率降低,准确率提高,最终误检率和准确率趋于稳定。且通过上图可以看出,本文所提出的结合空间信息的 PTSVM 算法,改善了变化检测的准确率。

为了进一步分析各个算法的优劣,以无标签样本数为

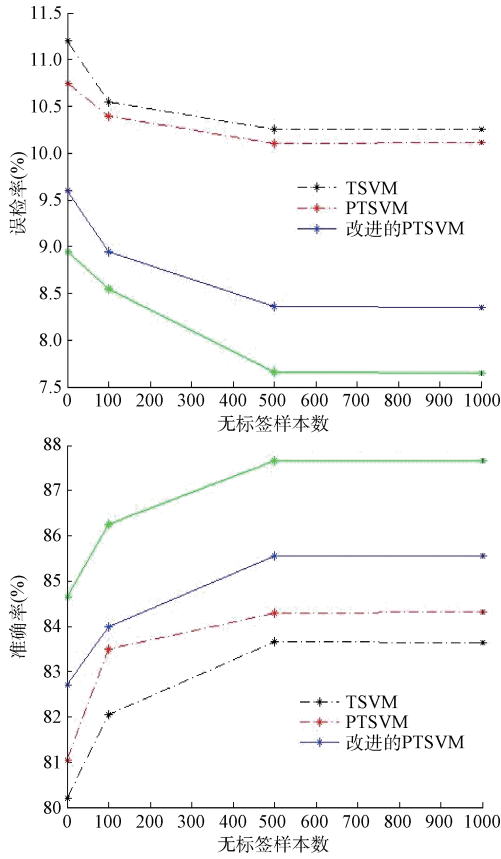


图 3 各变化检测算法随无标签样本数的变化图

指标,检测各算法随无标签样本数的变化性能。各算法变化检测的性能如表 2 所示。

表 2 变化检测性能评价

算法	无标签样本数	漏检率 (%)	误检率 (%)	准确率 (%)
T SVM	5	13.40	10.60	76.00
PTSVM	5	12.30	10.50	77.20
新算法	5	11.80	10.25	77.95
T SVM	10	10.50	10.00	79.50
PTSVM	10	10.00	9.50	80.50
新算法	10	9.50	9.10	81.40
T SVM	50	9.00	8.90	82.10
PTSVM	50	8.5	8.50	83.00
新算法	50	8.00	7.90	84.10
T SVM	100	7.45	7.35	85.30
PTSVM	100	7.00	7.01	85.99
新算法	100	6.90	6.80	86.3
T SVM	500	6.50	6.00	87.5
PTSVM	500	6.00	5.00	89.00
新算法	500	5.50	4.30	90.20

由表 2 可知,结合空间的 PTSVM 降低了变化检测的漏检率和误检率,大大提高了变化检测的准确率,使得遥感图像的变化检测性能得到很大改善。不仅如此,结合空间的 PTSVM 使得遥感图像变化检测算法发展到一个新的领域。

5 结 论

空间信息作为遥感图像的基本信息,在遥感图像的变化检测中也有非常重要的作用,因此在 PTSVM 算法的基础上结合空间信息并用于遥感图像的变化检测中。通过仿真结果分析和性能评测可知,将空间信息与光谱信息结合,改善了 PTSVM 在遥感图像变化检测中的效果,提高了变化检测的准确率。

参考文献

- [1] 钟家强. 基于多时相遥感图像的变化检测[D]. 长沙: 国防科技大学, 2009, 8(4): 407-413.
- [2] RADKE R J, ANDRA S, AL-KOFAHI O, et al. Image change detection algorithms: a systematic survey[J]. IEEE Transactions on Image Processing, 2005, 14(3): 294-307.
- [3] VAPNIK V N. The Nature of Statistical Learning Theory [M]. New York Springer, 2009, 33(23): 88-98.
- [4] GUO G, LI S Z. Content-based audio classification and retrieval by support vector machines[J]. IEEE Transactions on Neural Networks, 2003, 14(1): 209-215.
- [5] SHI L, KHUSHABA R, KODAGODA S, et al. Application of CRF and SVM based semi-supervised learning for semantic labeling of environments[C]// 2012 12th International Conference on Control Automation Robotics & Vision (ICARCV). IEEE, 2012: 835-840.
- [6] CHEN Y, WANG G, DONG S. Learning with progressive transductive support vector machine[J]. Pattern Recognition Letters, 2003, 24(12): 1845-1855.
- [7] 陈毅松, 汪国平. 基于支持向量机的渐进直推式分类学习算法[J]. 软件学报, 2012, 14(3): 451-460.
- [8] GAMMERMAN A, VOVK V, VAPNIK V. Learning by transduction[C]//Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence. Morgan Kaufmann Publishers Inc., 1998: 148-155.
- [9] 赵庆平, 陈得宝, 姜恩华, 等. 一种改进权重的非局部均值图像去噪算法[J]. 电子测量与仪器学报, 2014, 28(3): 334-339.

(下转第 52 页)