

DOI:10.3876/j.issn.1000-1980.2013.02.015

基于 MDL 准则和 EM 算法的多光谱遥感影像变化检测

石爱业¹, 吴国宝¹, 谭德宝², 申邵洪², 黄凤辰¹, 马贞立¹, 王超¹

(1. 河海大学计算机与信息学院, 江苏 南京 210098; 2. 长江科学院空间信息技术应用研究所, 湖北 武汉 430010)

摘要:针对由实际遥感地物类型难以确定导致的多光谱遥感影像变化检测精度较低的问题,提出一种结合最小描述长度(MDL)准则的 EM 变化检测方法。首先,采用主成分变换与相关系数融合法相结合的方式构造差异影像;其次,利用分支数为 k 的高斯分布混合模型对差异影像进行建模,并利用 MDL-EM 算法自适应估计模型各参数;最后,利用基于统计最小错误率的 Bayes 判别准则确定变化检测的阈值来实现多光谱遥感影像变化检测。实际遥感数据验证结果表明,所提检测方法应用于多光谱遥感影像变化检测中是可行、有效的。

关键词: 遥感影像处理; 变化检测; EM 算法; MDL 准则

中图分类号: TP751 文献标志码: A 文章编号: 1000-1980(2013)02-0171-06

Multispectral remotely sensed image change detection based on EM algorithm and MDL rule

SHI Aiye¹, WU Guobao¹, TAN Debao², SHEN Shaohong²,
HUANG Fengchen¹, MA Zhenli¹, WANG Chao¹

(1. College of Computer and Information Engineering, Hohai University, Nanjing 210098, China;
2. Spatial Information Technology Application Research Institute, Yangtze River Scientific Research Institute, Wuhan 430010, China)

Abstract: In order to overcome the disadvantages of multispectral remotely sensed image change detection due to the difficulty in determining the actual ground object type, a change detection algorithm is proposed through combination of the minimum description length (MDL) rule and the EM algorithm. First, the difference images were constructed by the principal component analysis (PCA) and the correlation coefficient fusion method. Then, a distribution model for the difference images was built with the Gaussian mixture model (GMM), and its related parameters were estimated adaptively by the minimum description length-expectation maximization (MDL-EM) algorithm. Finally, the change detection was implemented by the threshold determined by the Bayes rule. Experimental results on the multispectral images show that the proposed algorithm is feasible and effective.

Key words: remotely sensed image processing; change detection; EM algorithm; MDL rule

基于遥感影像的变化检测是根据同一区域不同时相的遥感影像辨识、观测对象或者现象的状态变化过程^[1-2]。它已广泛应用于资源管理与规划、环境保护等诸多领域^[2-4],并为相关部门提供科学决策的依据。当前,遥感影像的变化检测已成为一个非常活跃的研究方向^[2-3, 5-8]。遥感影像变化检测中的代数法是当前使用最广泛的方法之一,但其存在一些不足:构造差异影像的方法单一,不能很好地反映复杂地物的真实变化情况^[9-10];对于传统的 EM(expectation-maximization)算法^[11],当差异影像含有多种地物变化类型、光谱信息较复杂时,其变换检测精度较低^[12]。

针对上述问题,笔者首先采用主成分(principle component analysis, PCA)变换结合相关系数融合法^[9]来构造差异影像;其次,利用分支数为 k 的高斯分布混合模型(gaussian mixture model, GMM)对差异影像进行建

收稿日期: 2012-09-29
基金项目: 国家自然科学基金(61271386, 60901003)
作者简介: 石爱业(1969—),男,江苏泗洪人,副教授,博士,主要从事信号与信息处理及遥感图像处理研究。E-mail: ayshi@hhu.edu.cn

模,并引入 MDL(minmum description length)-EM 算法来估计 GMM 的各参数;最后,利用基于统计最小错误率的 Bayes 判别准则来确定变化检测的阈值,得到变化检测结果。通过实验数据比较分析,验证所提算法的可行性和有效性。

1 原理与步骤

1.1 基于 PCA 变换结合相关系数融合法的差异影像

多光谱遥感影像各波段上包含的信息存在一定程度的相关和冗余,因为 PCA 变换能很好地把原影像中多个波段的有用信息压缩到数目尽可能少的新成分影像中,并能使这些成分之间互不相关^[12],因此笔者首先对遥感影像数据进行 PCA 变换提取第一主成分,然后采用文献[9]所提的相关系数融合法来构造不同时相间的差异影像,以实现最大限度地消除各光谱波段间的相关性,从而有效地集中和突出不同时相间的变化信息,并最大限度地保留变化的细节信息,减少伪变化信息及背景噪声的干扰,为下一步的变化检测创造有利条件^[10]。

1.2 结合 MDL 准则的 EM 算法

直方图拟合是直方图分析法的一种^[13],利用模型对目标影像直方图进行拟合,这与进行遥感影像变化检测时将差异影像多类地物的光谱信息视作 k 类高斯分布的叠加相吻合。基于此,提出采用 GMM 来拟合差异影像直方图。设 L 为差异影像灰度值集合(一般为 $0,1,\cdots,255$), $h(f_i)$ 为直方图,其中 f_i 为其第 i 个像元灰度值,利用 EM 算法估计 GMM 各参数^[11]的迭代公式为

$$p^{t+1}(j) = \frac{1}{f_i h(f_i)} \sum_{f_i \in L} \frac{h(f_i) p^t(f_i | j) p^t(j)}{p^t(f_i)} \tag{1}$$

$$\mu_j^{t+1} = \frac{\sum_{f_i \in L} \frac{h(f_i) p^t(f_i | j) p^t(j)}{p^t(f_i)} f_i}{\sum_{f_i \in L} \frac{h(f_i) p^t(f_i | j) p^t(j)}{p^t(f_i)}} \tag{2}$$

$$(\sigma_j^2)^{t+1} = \frac{\sum_{f_i \in L} \frac{h(f_i) p^t(f_i | j) p^t(j)}{p^t(f_i)} (f_i - \mu_j^t)^2}{\sum_{f_i \in L} \frac{h(f_i) p^t(f_i | j) p^t(j)}{p^t(f_i)}} \tag{3}$$

式中: $p(j), \mu_j, \sigma_j^2$ ——GMM 中分支 j 的权重、均值和方差; t ——迭代次数。GMM 最优模型的选择由作为适应度函数的 MDL 准则来确定,其表达式为

$$R(k, \Theta) = -\ln P(f | k, \Theta) + \frac{T \ln NM}{2} \tag{4}$$

其中 $\Theta = (p(1), p(2), \cdots, p(k); \mu_1, \mu_2, \cdots, \mu_k; \sigma_1^2, \sigma_2^2, \cdots, \sigma_k^2)$ $T = k \left[1 + M + \frac{M(M+1)}{2} \right] - 1$

式中: k ——GMM 的分支数; N ——GMM 的像元总数; M ——GMM 中像元的维数; $\ln P(f | k, \Theta)$ ——完全数据集的对数似然函数; $\frac{T \ln NM}{2}$ ——引入的针对 GMM 分支数 k 的惩罚函数。

在选定初始最大分支数 $k_{\max}$ 的情况下,通过引入合并最小距离分支^[14]的办法自适应地依次减少分支数至最小分支数 $k_{\min}$,并依据式(5)确定最优的分支数 k^* 和参数 Θ^* 。

$$\{k^*, \Theta^*\} = \operatorname{argmin} \{R(k, \Theta), k_{\min} \leq k \leq k_{\max}\} \tag{5}$$

利用上述 MDL-EM 算法能自适应地得到 GMM 的 k^* 和 Θ^* ,进而得到对差异影像直方图的最佳拟合,再利用基于最小错误率的 Bayes 判别准则确定变化检测的阈值^[11]。

1.3 新算法实现步骤

步骤1 对选取的 2 幅遥感影像 TM_1, TM_2 进行几何校正及辐射归一化等预处理。

步骤2 分别对经过预处理的 TM_1, TM_2 进行 PCA 变换,并提取第一主成分 X_{1pc1}, X_{2pc1} 。

步骤3 求 X_{1pc1}, X_{2pc1} 的差值、比值影像,并通过相关系数融合法构造差异影像 F 。

步骤 4 设定初始参数 $k_{\max}, k_{\min}$ 及 Θ^0 , 运用 MDL-EM 算法估计 k^* 及 Θ^* 。

步骤 5 使用 GMM 对差异影像 F 的直方图进行拟合。

步骤 6 采用基于最小错误率的 Bayes 判别准则确定变化检测的阈值, 得到变化检测结果。

2 实验与分析

2.1 实验数据 1

采用巴西朗多尼亚州亚马逊流域的 2 幅 Landsat TM (Thematic Mapper) 遥感影像为实验对象, 对其森林砍伐面积进行变化检测。2 幅遥感影像的获取时间分别为 2000 年 7 月和 2006 年 7 月, 实验区大小为 512 像元 $\times$ 512 像元, 如图 1 所示。该区域主要由河流、森林及砍伐后的裸地 3 部分组成。

在几何校正及辐射归一化的基础上, 采用基于 PCA 变换结合相关系数融合法对实验数据进行构造, 得到的差异影像如图 2 所示, 其像元可视为由分支数为 k 的高斯混合分布组成, 对应的直方图如图 3 所示。

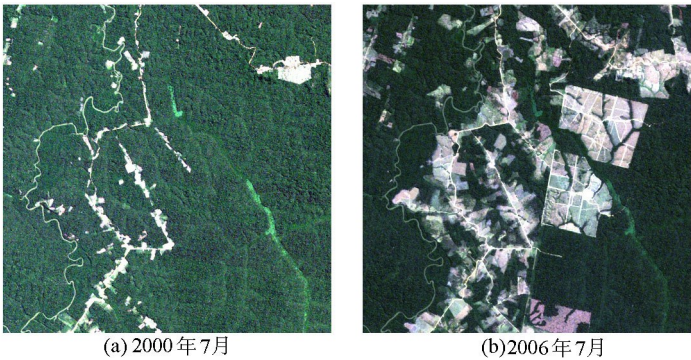


图 1 Landsat TM 数据

Fig. 1 Landsat TM data (true color)

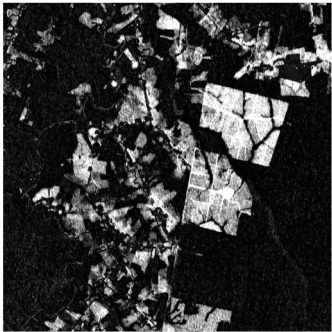


图 2 差异影像

Fig. 2 Difference image

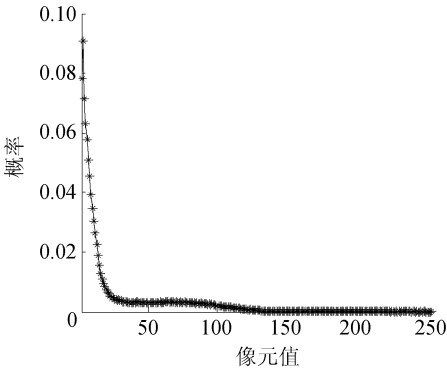


图 3 差异影像直方图

Fig. 3 Histogram of difference image

在 MDL-EM 算法中, 设置 GMM 的初始值 $k_{\max}=30, k_{\min}=2$ 。经过计算, 各分支数 k 对应的 R 如图 4 所示。由图 4 可以看到 $k=9$ 时对应的 R 最小, 因此由 MDL 准则可知, 当 $k=9$ 时的模型是最佳的拟合模型, 此时对差异影像直方图拟合的效果最好, 对应的 GMM 各分支参数如表 1 所示。由于本实验的主要目的是检测森林砍伐区变化(只检测森林是否被砍伐), 暂不考虑其他的变化类型, 所以将砍伐区统一归为“变化类”, 未砍伐区统一归为“非变化类”。

根据表 1, 利用 GMM 对图 3 所示的差异影像直方图进行拟合。由图 3 可知, 当灰度值大于 100 时, 其像

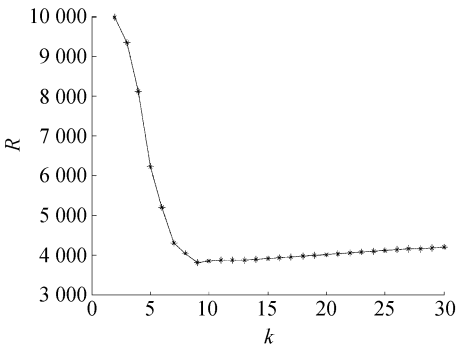


图 4 不同 k 对应的 R

Fig. 4 R with different k values

表 1 $k^*=9$ 时 GMM 各分支参数

Table 1 GMM branch parameters by $k^*=9$

类别	$p(j)$	μ_j	σ_j^2
GMM1	0.137	0.478	0.287
GMM2	0.124	2.119	0.843
GMM3	0.149	4.305	2.218
GMM4	0.130	7.385	4.831
GMM5	0.118	11.095	11.990
GMM6	0.067	22.731	18.906
GMM7	0.070	37.848	158.678
GMM8	0.196	76.531	618.989
GMM9	0.008	126.659	1948.260

素数很少,直方图与横坐标几乎重合,此时进行拟合结果比较意义不大。为了便于对拟合结果进行较清晰的比较,仅将像元灰度值在 0~100 之间的拟合部分进行放大显示, $k=9$ 时的直方图拟合结果如图 5(a) 所示。为了比较 MDL-EM 算法求得的不同分支数对差异影像直方图的拟合效果,选取具有代表性的 $k=2$ 时的直方图拟合结果作为比较,如图 5(b) 所示。由图 5 可以看出,在整体的拟合效果上, $k=9$ 的拟合效果要好于 $k=2$ 的拟合效果。依据 MDL-EM 算法参数估计的结果和最小错误率的 Bayes 判别准则,确定变化检测的阈值,进而对图 2 进行二值化处理,得到最终变化检测结果如图 6(a) 所示。

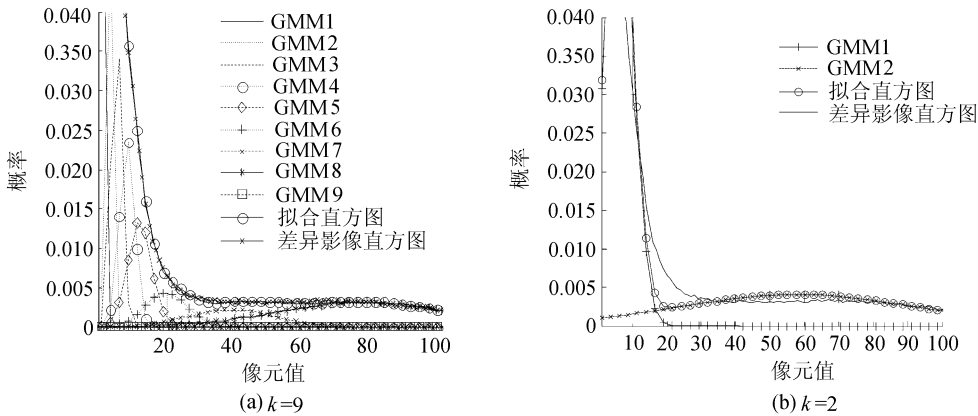


图5 差异影像直方图拟合结果

Fig. 5 Fitted histogram of difference image

为了比较新算法的优劣,选取两类变化检测算法做进一步的对比分析:(a) Bruzzone 所提的 CVA-EM 算法^[11]; (b) PCA 融合-EM 算法。所有对比实验的 EM 算法均设定 GMM 的分支数 $k=2$, EM 算法初始值设置方法和文献[11]相同。2 种算法的检测结果分别如图 6(b)(c) 所示;目视解译得到的理想变化检测结果如图 6(d) 所示。

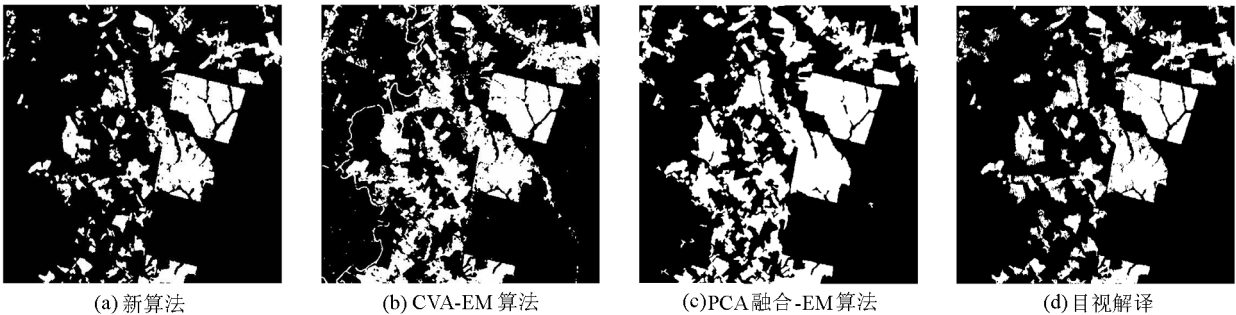


图6 实验数据1的3种算法及目视解译变化检测结果

Fig. 6 Change detection map by three algorithms and visual interpretation in experimental group 1

采用误检率、漏检率、总体精度、Kappa 系数和计算时间 5 个指标来评价新算法及其他 2 种变化检测算法的性能,结果如表 2 所示。由表 2 可知,较于其他算法,新算法除计算代价稍大外,在总体检测性能上是最好的,因此新算法提高了多光谱遥感影像的变化检测精度。

表2 实验数据1的各变化检测算法性能评价

Table 2 Performance evaluation of three algorithms in experimental group 1					
算法	误检率/%	漏检率/%	总体精度/%	Kappa 系数	计算时间/s
CVA-EM 算法	25.90	11.65	80.54	0.6140	146.81
PCA 融合-EM 算法	19.80	7.15	85.92	0.7200	135.79
新算法	8.92	8.19	91.41	0.8270	207.32

2.2 实验数据2

采用美国加利福尼亚州索尔顿湖卡利帕特里亚地区的 2 幅遥感影像,获取时间分别为 2008 年 8 月和 2009 年 8 月,实验区大小为 400 像元×400 像元,如图 7 所示。该区域地物类型较为丰富,包含有水域、山地、城市、农用耕地、裸地等,同时通过目视解译发现该区域地物变化信息也较为复杂,如水域变成裸地、山地变

成农用耕地、农用耕地变成裸地等。

利用 MDL-EM 算法对 GMM 的各参数进行估计,得到 $k^*=10$, 最终的变化检测结果如图 8(a) 所示。采用与第 1 个实验相同的比较算法,且各算法的初始参数设置亦相同,得到的变化检测结果如图 8(b)(c) 所示;目视解译得到的理想变化检测结果如图 8(d) 所示。各算法的性能评价见表 3。

通过比较图 8 的变化检测结果及由表 3 的各变化检测算法性能评价可知,相比于其他的算法,本文所提新算法在总体检测性能上是最好的,这也表明所提新算法用于多光谱遥影像的变化检测是可行的。

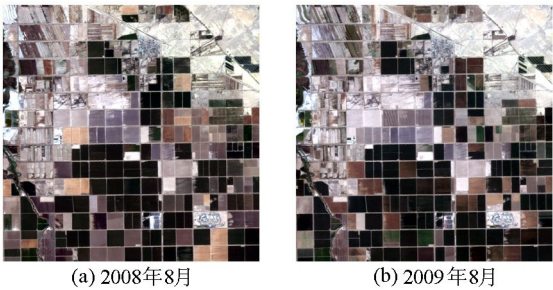


图 7 Landsat TM 数据
Fig. 7 Landsat TM data (true color)

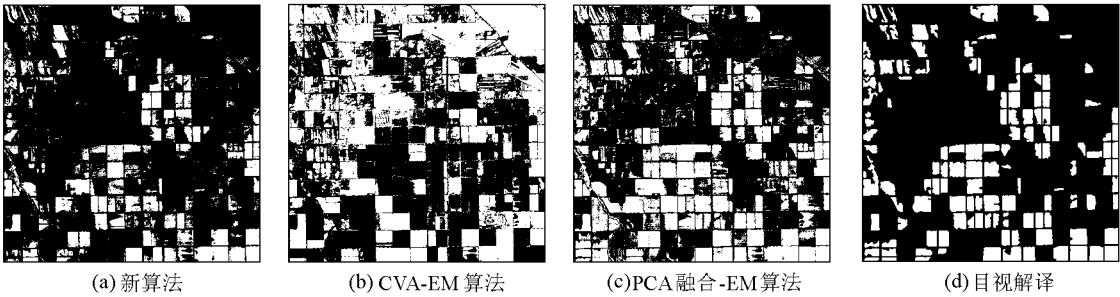


图 8 实验数据 2 的 3 种算法及目视解译的变化检测结果
Fig. 8 Change detection map by three algorithms and visual interpretation in experimental group 2

表 3 实验数据 2 的各变化检测算法性能评价

算法	误检率/%	漏检率/%	总体精度/%	Kappa 系数	计算时间/s
CVA-EM	26.05	11.60	82.23	0.6376	56.58
PCA 融合-EM	17.69	5.71	88.89	0.7734	113.68
新算法	7.29	9.10	91.70	0.8333	149.44

3 结 语

通过引入 PCA 变换并结合相关系数融合法来构造差异影像,将 MDL 准则引入 EM 算法中估计高斯混合模型的分支数及各分支参数。实验分析表明,所提新的变化检测算法是有效的。新算法中,GMM 的初始分支数通过试错法来选择。另外,新算法是在假定差异影像像元间相互独立的基础上进行的,并未考虑像元的空间关系,如何将这种空间关系嵌入到变化检测算法中以提高检测性能也是今后需要进一步研究的问题。

参考文献:

[1] SING A. Change detection techniques using remotely sensed data[J]. International Journal of Remote Sensing, 1989, 10(6): 989-1003.

[2] LU D, MAUSEL P, MORAN E. Change detection techniques[J]. International Journal of Remote Sensing, 2004, 25(12): 2365-2407.

[3] BOVOLO F, BRUZZONE L. A split-based approach to unsupervised change detection in large-size multi-temporal images: application to tsunami-damage assessment[J]. IEEE Transactions on Geoscience and Remote Sensing, 2007, 45(6): 1658-1670.

[4] CHINI M, PACIFICI F, EMERY W. Comparing statistical and neural network methods applied to very high resolution satellite images showing changes in man-made structures at rocky flats[J]. IEEE Transactions on Geoscience and Remote Sensing, 2008, 46(6): 1812-1821.

[5] GABRIELE M, ELENA A, SEBASTIANO B S. Multiscale unsupervised change detection on optical images by Markov random fields and wavelets[J]. IEEE Geoscience and Remote Sensing Letters, 2011, 8(4): 725-729.

[6] MERCIER G, SERPICO S. Conditional copulas for change detection in heterogeneous remote sensing images[J]. IEEE

Transactions on Geoscience and Remote Sensing, 2008, 46(5): 1428-1441.

[7] LUO W, LI H L. PCA based unsupervised change detection for color satellite images under the quaternion model [C] // International Symposium on Intelligent Signal Processing and Communication Systems. Chengdu: Journal of University of Electronic Science and Technology of China, 2010:1-4.

[8] GABRIELE M, ELENA A, SEBASTIANO B S. A contextual multiscale unsupervised method for change detection with multitemporal remote-sensing images [C] // 2009 Ninth International Conference on Intelligent Systems Design and Application. Pisa: Journal of University of Pisa, 2009: 572-577.

[9] 魏立飞, 钟燕飞, 张良培, 等. 遥感影像融合的自适应变化检测 [J]. 遥感学报, 2010, 14(6): 1204-1211. (WEI Lifei, ZHONG Yanfei, ZHANG Liangpei, et al. Adaptive change method of remote sensing image fusion [J]. Journal of Remote Sensing, 2010, 14(6): 1204-1211. (in Chinese))

[10] 柳思聪, 杜培军, 陈绍杰. 决策级融合的多分辨率遥感影像变化检测 [J]. 遥感学报, 2011, 15(4): 846-862. (LIU Sicong, DU Peijun, CHEN Shaojie. A novel change detection method of multi-resolution remotely sensed images based on the decision level fusion [J]. Journal of Remote Sensing, 2011, 15(4): 846-862. (in Chinese))

[11] BRUZZONE L, PRIETO D F. Automatic analysis of difference image for unsupervised change detection [J]. IEEE Transactions on Geoscience and Remote Sensing, 2000, 38(3): 1171-1182.

[12] 吴柯, 牛瑞卿, 王毅, 等. 基于 PCA 与 EM 算法的多光谱遥感影像变化检测研究 [J]. 计算机科学, 2010, 37(3): 282-284. (WU Ke, NIU Ruiqing, WANG Yi, et al. Change detection of multispectral remote sensed images based on PCA and EM algorithm [J]. Computer Science, 2010, 37(3): 282-284. (in Chinese))

[13] 邹丹平, 冯涛, 李咸伟, 等. 基于 EM 的直方图逼近及其应用 [J]. 中国图像图形学报, 2005, 10(11): 1458-1462. (ZOU Danping, FENG Tao, LI Xianwei, et al. Histogram approximation based on expectation maximization algorithm and its application [J]. Journal of Image and Graphics, 2005, 10(11): 1458-1462. (in Chinese))

[14] XIE W B, WANG X H. EM algorithm based MDL application to estimate the mixture model clustering parameters [C] // The 9th International Conference on Signal Processing. Beijing: International Journal of Scientific & Technology Research, 2008: 1637-1640.

• 简讯 •

河海大学进入全国发明专利 50 强高校

教育部科技发展中心 2013-02-23 公布了“2012 年高校获发明专利授权量前 50 名”名单。河海大学 2012 年获国家发明专利授权 199 件，位居全国高校第 47 位。

(本刊编辑部供稿)