维普资讯 http://www.cqvip.com

多尺度 Markov 模型的可适应图像分割方法

郭小卫[1],田 铮[1,2,3],林 伟[1,2],熊 毅[2]
(1. 西北工业大学计算机科学与工程系,陕西西安 710072;2. 西北工业大学应用数学系,陕西西安 710072;
3. 中国科学院自动化研究所模式识别国家重点实验室,北京 100080)

摘 要: 本文在图像分割的 TSMAP(trainable sequential maximum a posterior)方法基础上,提出基于多尺度 Markov 模型的可适应 ATSMAP(adaptive TSMAP)图像分割方法.在给定训练图像及其基本真实分割(ground truth segmentation, GTS)的基础上,通过直接对原始图像的 GTS 进行小波变换产生粗尺度上的 GTS,进而估计出图像数据的分布参数和 Markov 四叉树模型参数;上下文模型参数根据上下文的低维特征(类别数量特征)而非上下文本身来估计.该方法具有上下文模型参数估计计算量小,Markov 四叉树模型参数可针对特定的待分割图像重新优化等优点(模型适应过程),解决了 TSMAP 方法易导致过学习的问题,在待分割图像与训练图像的统计特性不匹配的情况下,仍能给出较好的分割结果.对合成图像与 SAR 图像的实验结果表明,这种方法的分割精度高于 TSMAP 和其它几种基于多尺度 Markov 模型的图像分割方法.

关键词: TSMAP (trainable sequential maximum a posterior); 多尺度 Markov 模型; ATSMAP (adaptive; trainable sequential maximum a posterior); 图像分割; SAR 图像

中图分类号: TP391.4 **文献标识码:** A **文章编号:** 0372-2112 (2005) 07-1279-05

A Method of Adaptive Image Segmentation Based on Multiscale Markov Models

GUO Xiao-wei[1], TIAN Zheng[1,2,3], LIN Wei[1,2], XIONG Yi[2]
(1. *Department of Computer Science and Technology, Northwestern Polytechnic University, Xi' an, Shaanxi* 710072, *China*;
2. *Department of Applied Mathematics, Northwestern Polytechnic University, Xi' an, Shaanxi* 710072, *China*;
3. *National Key Laboratory for Pattern Recognition, Institute of Automation Chinese Academy of Sciences, Beijing* 100080, *China*)

Abstract: In order to remedy some of the disadvantages of trainable sequential maximum a posterior (TSMAP), we introduce an adaptive trainable sequential maximum a posterior (ATSMAP) approach for image segmentation based on multiscale Markov model. This method obtains the ground truth segmentations at all scales by wavelet decomposition of the accurate segmentation of the original image, and then estimates the parameters of multiscale quadtree models. Also, the parameters of the context model are estimated by a simple way that estimates the parameters with the low dimension feature instead of the context itself. Compared with TSMAP, the advantages of ATSMAP are as follows: 1). The parameters of quadtree can be retrained for the specific image being segmented. 2) The estimation of parameters for context models is very computationally efficient. The experimental results of the segmentation for both synthetic images and SAR images indicate that the approach fairly improves the segmentation accuracy over TSMAP and some other multiscale Markov based methods.

Key words: TSMAP(trainable sequential maximum a posterior); multiscale Markov model; ATSMAP(adaptive TSMAP); image segmentation; SAR image

1 引言

近年,多尺度 Markov 模型(multiscale Markov model 简称 MRM)的理论和方法是信号与图像统计处理的研究热点.MRM 能够很好地解决非因果多尺度 Markov 场(Multiscale Markov Field 简称 MRF)模型的参数估计和分割算法需要迭代等诸多问题,因而被广泛地应用于图像分割与去噪等问题.文献[1]在 MPM (maximum posterior marginal)的基础上,提出了一种 SMAP (sequential maximum a posterior)估计的图像分割方法,这种方法较 MAP(maximum a posterior)与 MPM 估计更为合理,因为它考虑了估计误差在多尺度结构上的位置信息.文献[2]在 SMAP 估计的基础上,提出了 TSMAP(trainable SMAP)图像分割方法,其基本思想是利用一幅或多幅训练图像以及相应的基本真实分割(ground truth segmentation 简称 GTS)来训练多尺

收稿日期:2003-12-18;修回日期:2005-01-05
基金项目:国家自然科学基金(No.60375003);航空科学基金(No.03I53059)

维普资讯 http://www.cqvip.com

度 Markov 模型,并将这种训练图像的模型参数用于待分割图像的分割.虽然该方法参数估计的精度较高,但它利用"类别概率树"(class probability trees,简称 CPT)来估计上下文模型参数,计算量比较大;另外,TSMAP 的训练算法容易导致过学习,在训练图像与待分割图像的统计特性不匹配的情况下,分割精度低.

针对上述问题,本文提出了基于多尺度 Markov 模型的可适应图像分割方法 ATSMAP(adaptive TSMAP),该方法利用简单的低维特征估计上下文模型参数,有效地降低了训练算法的计算量;并且对特定的待分割图像可以重新优化 Markov 四叉树模型参数,解决了 TSMAP 的过学习问题,从而提高了分割精度.

2 TSMAP 图像分割方法

2.1 TSMAP 的多尺度图像分割模型

TSMAP 采用如图 1 所示的多尺度 Markov 模型来对图像建模,其中,$Y^{(n)}$表示尺度 n 上的图像数据,$X^{(n)}$表示其对应的类别.这里,$Y^{(n)}$和 $X^{(n)}$是随机向量,对一特定的图像,图像数据 $y^{(n)}$及对应的类别 $x^{(n)}$被看作 $Y^{(n)}$和 $X^{(n)}$的实现.TSMAP 是在 SMAP 估计基础上的一种有监督图像分割方法,它采用的仍然是 SMAP 分类器,根据文献[1],SMAP 估计可表示为如下形式:

$$\hat{x}_s = \arg\max_{1\leqslant k\leqslant M}\{\log p(y_{t(s)}\mid x_s=k)+\log p(x_s=k\mid \hat{x}_{s^+})\} \quad (1)$$

其中 $t(s)$表示以 s 为根节点的子树,$s^+\subset S^{(n+1)}$表示尺度 $n+1$ 上以 s 的父节点为中心的大小为 $d\times d$ 的窗内的节点集.邻域大小一般取 1×1,3×3,5×5 甚至更大,邻域越大,分割结果一般越光滑,但计算量也会随之增加,这种上下文模型也被称为分层图模型或金字塔图模型,其结构如图 2 所示.图 3 给出了 Markov 四叉树模型的结构.

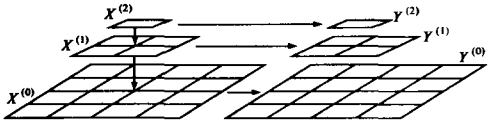

图 1 多尺度图像分割模型.$Y^{(n)}$表示尺度 n 上的图像数据,$X^{(n)}$表示尺度 n 上的像素类别

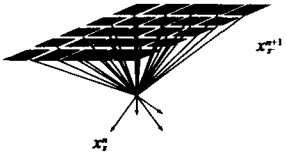

图 2 二维金字塔图像模型,邻域大小为 5×5

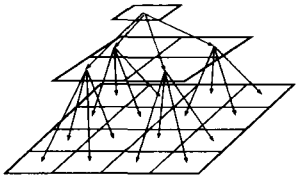

图 3 Markov 四叉树模型

2.2 TSMAP 的训练算法

SMAP 估计依赖于两类参数,一类是用于似然项计算的参数,包括条件分布 $p(y_s\mid x_s)$的参数和四叉树模型参数 $p(x_s\mid x_{s^-})$,其中 s^-表示 s 的父节点;另一类是上下文模型参数 $p(x_s\mid x_{s^+})$.TSMAP 在 SMAP 估计的基础上采用如下步骤来估计这些参数:

(1)在给定训练图像的 GTS 的条件下用 EM 算法估计 Markov 四叉树模型的参数;

(2)利用 Markov 四叉树模型参数与训练图像的 GTS,估计出所有尺度的 GTS;

(3)在每一尺度上,利用 Gauss 混合模型对每类数据的分布建模,并估计出模型参数;

(4)估计上下文模型参数.

其中,第 4 步上下文模型的参数估计较困难,例如 x_{s^+}的大小为 5×5,类别数为 3 时,x_{s^+}可能的值有个 $3^{25}\approx10^{12}$.因此,近似而有效的上下文模型参数估计方法对模型训练的计算效率和数据存储是很重要的.

3 基于 TSMAP 的可适应图像分割(ATSMAP)方法

TSMAP 方法的缺点在于:(1)容易导致过学习.估计的 Markov 四叉树模型参数和上下文模型参数过于理想,当训练图像与待分割图像的统计特性不能很好匹配的情况下,分割精度不高.(2)计算量大.利用 EM 算法估计 Markov 四叉树模型参数和利用 CPT 来估计上下文模型参数时需要迭代,计算量较大.本节将在 TSMAP 的基础上,给出一种改进方法 ATSMAP,该方法的基本步骤如下:

(1)通过对训练图像的 GTS 做小波变换直接生成粗尺度上 GTS;

(2)利用多尺度图像数据与所有尺度的 GTS 估计出条件分布 $p(y_s\mid x_s)$的参数;

(3)利用相邻尺度上各节点的分类情况来统计计算 Markov 四叉树模型参数;

(4)估计上下文模型参数.

3.1 粗尺度上 Ground Truth 的生成

利用最细尺度上的 GTS 生成所有尺度的 GTS,简单的方法如四抽一、多数表决等会造成粗尺度上细节的丢失和边界的不连续,文献[2]给出的基于极大似然估计的方法得到的效果很好,但该方法计算量较大.

本文利用小波变换的多尺度特性,对尺度 n 上的 GTS 做小波分解后将得到的尺度系数作为尺度 $n+1$

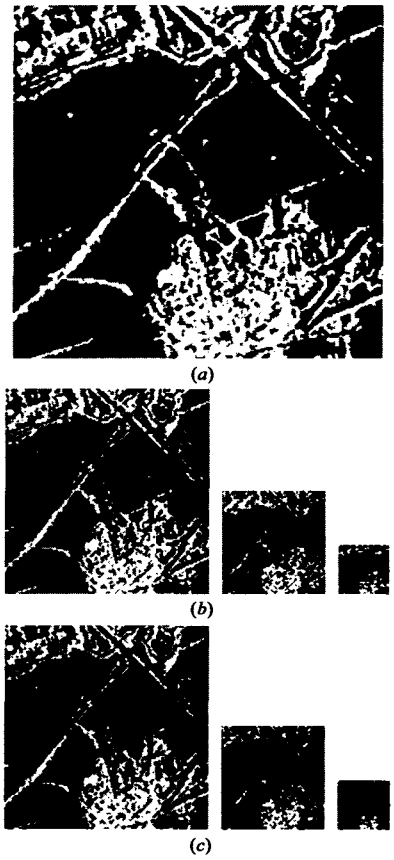

图 4 (a)尺度 0 上的 Ground Truth,(b)利用多数表决方法生成的尺度 1、2、3 上的 Ground Truth,(c)利用小波变换方法生成的尺度 1、2、3 上的 Ground Truth

维普资讯 http://www.cqvip.com

上的 GTS,这需要先对类别标记进行数值化处理(编码).通常如定义"森林"类为 1,"草地"类为 2,"阴影"类为 3 的编码方式是不合适的,因为细尺度上的 GTS 经小波变换后的尺度系数本质上是细尺度的 GTS 的某种加权平均并下抽样的结果,因而经过一次小波变换后,在"森林"类和"阴影"类的边界处可能会出现"草地"类.为此,可采用单位向量来对类别编码,如"森林"类为$[1,0,0]^t$,"草地"类为$[0,1,0]^t$,"阴影"类为$[0,0,1]^t$,这样每一尺度 GTS 都被表示为一个三维点阵,第一维和第二维对应像素的坐标,第三维对应类别,而类别维上第 m 个分量的集合就构成了第 m 类的 GTS.其中属于第 m 类的像素标记为 1,不属于第 m 类的像素标记为 0.通过对尺度 n 上每一类别的 GTS 做小波变换后保留尺度系数,就得到了尺度 $n+1$ 上向量形式表示的 GTS(如果是非单位向量需进行单位化).本文以 Haar 小波作为基函数对训练图像的 GTS 按以上编码做多级小波变换,产生粗尺度上的 GTS.图 4 给出了这种方法与多数表决法生成的 GTS 的比较,由图示结果可见,这种方法对边界的连续性和细节保留的很好.

3.2 条件分布 $p(y_s|x_s)$的参数估计

由于 $p(y_s|x_s)$的分布形式未知,而任一光滑密度可用混合元数目足够多的 Gauss 混合密度任意逼近,因此,本文用 Gauss 混合模型来对 $p(y_s|x_s)$建模,即:

$$p(Y_s=y|X_s=m)=\sum_{j=1}^{J_{m,n}}\left\{\gamma_{j,m,n}\frac{1}{\sqrt{2\pi\sigma_{j,m,n}^2}}\exp\left[-\frac{(y-\mu_{j,m,n})^2}{2\sigma_{j,m,n}^2}\right]\right\} \tag{2}$$

其中,$J_{m,n}$为尺度 n 上第 m 类的 Gauss 混合模型的阶数,$\gamma_{j,m,n}$、$\mu_{j,m,n}$和$\sigma_{j,m,n}^2$分别为尺度 n 上第 m 类的 Gauss 混合模型的第 j 个混合元的权.EM 算法是估计混合模型参数的标准算法(见文献[3]),可利用 EM 算法来估计这些参数.在所有尺度的 GTS 已知的情况下,很容易得到各尺度上各类的训练数据.

3.3 四叉树模型参数的估计

在所有尺度的 GTS 已知的情况下,尺度 n 上的四叉树模型参数可由下式来估计:

$$p(X_s=k|X_s^-=m)=\frac{N(X_s=k,X_s^-=m)}{\sum\limits_{1\leqslant k\leqslant M}N(X_s=k,X_s^-=m)},\quad s\in S^{(n)} \tag{3}$$

其中,$N(A)$表示事件 A 发生的频数.实验表明,由这种方法得到的参数估计与 TSMAP 中用 EM 算法得到参数估计差别不大.

3.4 上下文模型参数的估计

由于在金字塔图模型上 x_s^+是高维的,上下文模型参数 $p(x_s|x_s^+)$的精确估计较困难.可考虑用 x_s^+的某种低维特征 f_s 代替 x_s^+来近似估计之.考虑到在 SAR 图像与大多数自然图像中,x_s 更主要的是与 x_s^+中各类别的像素数量有关,而与每类像素的具体位置关系不大,可选择 x_s^+中各类别的像素数量作为特征向量.设 x_s^+中属于第 m 类的像素数量为 $N_{s,m}$,则 x_s 的上下文特征向量为$f_s=(N_{s,1},\cdots,N_{s,M})$.在训练图像中,可以通过统计各个尺度上节点的上下文特征向量 f_s 来近似估计上下文模型参数:

$$p(X_s=x_s|X_s^+=x_s^+)\approx p(X_s=x_s|F_s=f_s)=\frac{N_{s,x_s}}{\sum\limits_{i=1}^{M}N_{s,i}} \tag{4}$$

需要注意的是,上下文模型参数 $p(x_s/x_s^+)$反映的是尺度 $n+1$ 上的 SMAP 分割与尺度 n 上的所期望的分割之间的概率转移关系.因而,上下文模型参数估计算法如下:

(1)由 MAP 估计得到 $\hat{x}^{(L)}$;

(2)由式(4),利用尺度 $n+1$ 上的 SMAP 估计 $\hat{x}^{(n+1)}$与尺度 n 上的 GTS 来估计尺度 n 上的上下文模型参数;

(3)令 $n=n-1$,若 $n=0$,则停止,否则转到(2).

与 CPT 方法相比,该上下文模型参数估计方法虽然精度不高,但算法简单,计算量小,对存储容量的要求也很低.实验表明,这种方法对于 SAR 图像及大多数自然图像的分割是很有效的.

4 模型适应

实验表明,在待分割图像与训练图像统计特性差异较大时,分割结果往往不够精确,这是由于训练图像的四叉树模型参数与待分割图像不匹配造成的.这时就需要针对待分割图像重新优化模型参数,本文将这一过程称为模型适应过程.模型适应过程可用如下 EM 算法(详见文献[4])来解决:

(1)初始化　设定根节点的先验概率 $p(x_r=m)$,$m=1,2,\cdots,M$ 和每一尺度上的类别转移概率 $p(x_s=m|x_s^+=k)$,$m,k=1,2,\cdots,M$ 的初值;

(2)计算出两个后验概率 $p(x_s|y)$和 $p(x_s,x_{\bar{s}}|y)$　这两个概率可由自上至下的下行算法得到:

$$\begin{aligned}p(x_s|y)&=\sum_{x_s^-}p(x_s|x_s^-,y)p(x_s^-|y)\\&=\sum_{x_s^-}p(x_s|x_s^-,y_{t(s)})p(x_s^-|y)\\&=\sum_{x_s^-}\frac{p(x_s,x_s^-|y_{t(s)})}{\sum\limits_{x_s}p(x_s,x_s^-|y_{t(s)})}p(x_s^-|y)\end{aligned}$$

$$P(x_s,x_s^-|y)=\frac{p(x_s,x_s^-|y_{t(s)})}{\sum\limits_{x_s}p(x_s,x_s^-|y_{t(s)})}p(x_s^-|y)$$

(3)参数更新

$$p(x_s=m|x_{\bar{s}}=k)=\frac{\sum\limits_{s\in S^n}p(x_s=m,x_{\bar{s}}=k|y)}{\sum\limits_{s\in S^n}p(x_s=m|y)},$$

$$p(x_r=m)=p(x_r=m|y).$$

(4)判断是否收敛　若收敛,停止,否则,转(2).

初值的选取对 EM 算法的收敛速率影响较大,为减小计算量,可将原模型中的四叉树模型参数作为初值.尽管经优化后四叉树参数可能会与原有的上下文模型参数不匹配,但由于本文的上下文模型参数采用 $\hat{x}_s^+$的低维的各类别像素数量特征 f_s 代替 $\hat{x}_s^+$来近似估计,这使得上下文模型参数的估计对待分割图像与训练图像的统计特性的变化不敏感(即待分割图像与训练图像的上下文参数可认为基本相同,若这一点也不能保证,那么所选择的训练图像就是不合适的).经过模

维普资讯 http://www.cqvip.com

型适应后,模型参数针对待分割图像被优化,分割精度一般会有不同程度的提高.

5 实验结果

我们将 ATSMAP 与 TSMAP 和完全基于 Markov 四叉树的分割方法 MPM[5] 方法进行比较实验.此处给出两个实验结果,一个是对合成图像的分割结果,另一个是对 SAR 图像的分割结果.实验中所使用的多尺度图像数据由训练图像和待分割图像经多级小波变换后的尺度系数构成.小波变换的基函数为 Haar 小波,各图像的大小均为 256×256,估计上下文模型参数时邻域大小为 5×5.

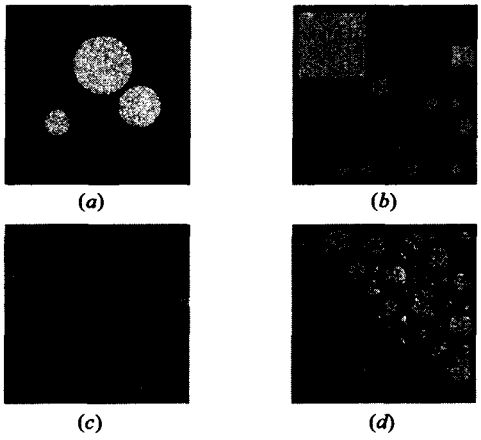

图 5 256×256 四幅合成图像

图 5 给出了四幅合成图像,(a)为训练图像,其中有三类区域,每类都被加入了服从二阶 Gauss 混合分布的白噪声,(b)~(d)为待分割图像,其中的白噪声的分布参数与(a)中的略有差异.表 1 比较了不同方法对图 5(b)~(d)的分割结果的平均错误率,其中,ATSMAP 方法后括号内的"未适应"或"适应后"表示该 ATSMAP 方法是否经过模型适应过程.图 6 给出不同方法对图 5(b)的分割结果.表 2 给出 ATSMAP 方法与 TSMAP 方法在标准 PC(PⅢ-500)和 Matlab6.1 平台上的执行时间对比.

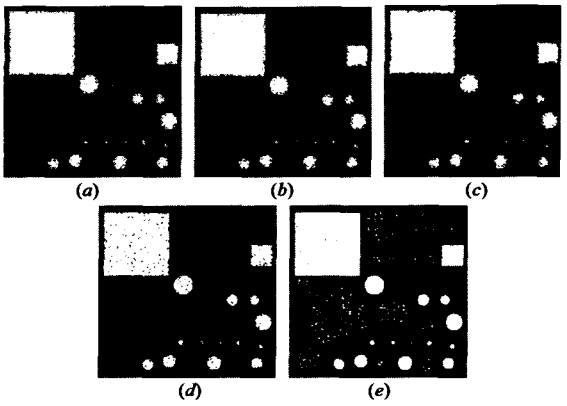

图 6 (a) ATSMAP(未适应), (b) ATSMAP(适应后), (c)TSMAP, (d)MPM(有监督), (e) MPM(无监督)

表 1 不同方法对合成图像分割的平均错误率比较

图像	图 5(b)	图 5(c)	图 5(d)
ATSMAP(未适应)	18.48%	7.48%	10.28%
ATSMAP(适应后)	9.45%	5.82%	6.65%
TSMAP	14.76%	6.93%	12.51%
MPM(有监督)	17.2%	13.34%	18.42%
MPM(无监督)	21.81%	19.11%	35.67%

表 2 ATSMAP 与 TSMAP 训练算法的执行时间(s)

方法	TSMAP	ATSMAP
粗尺度 GTS 生成	9.5×10	1.2
分布参数估计	3.7×20	3.7×20
四叉树参数估计	14.2×20	2.9
上下文参数估计	33.7×20	6.1
适应过程	0	11.3×5
总计	999.2	140.7

从该实验可以看出,未经模型适应的 ATSMAP 的分割结

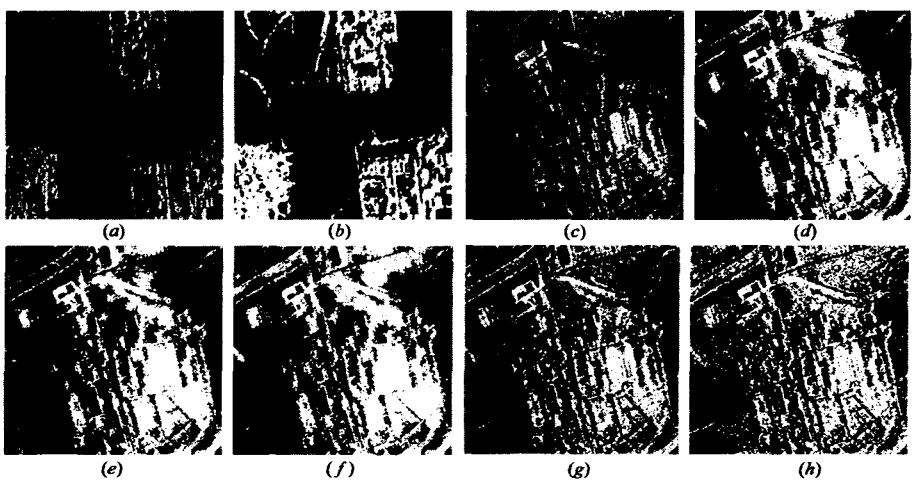

图 7 (a)训练图像, (b)训练图像的 GTS, (c)待分割图像, (d)ATSMAP(未适应), (e)ATSMAP(适应后), (f)TSMAP, (g)MPM(有监督), (h) MPM(无监督)

维普资讯 http://www.cqvip.com

果与 TSMAP 基本相当,在经过模型适应后,其结果优于 TSMAP.这两种方法都明显优于有监督的 MPM 方法,而无监督的 MPM 方法分割结果最差,这是因为它采用 Gauss 分布来对每类数据建模,模型与实验图像不匹配.

图 7 给出了 SAR 图像的实验结果.SAR 图像数据具有典型的非 Gauss 分布特征,适合用本文中的 Gauss 混合分布建模.从图中可以看出,在植被区与阴影区之间的错分率上,ATSMAP 要明显小于 TSMAP,而基于 Gauss 分布建模的无监督 MPM 分割结果最差.

6 结束语

基于多尺度 Markov 模型和 SMAP 估计的 TSMAP 图像分割方法是一种有监督方法,其主要缺点在于待分割图像与训练图像的统计特性不匹配的情况,分割精度低,另外上下文模型的参数估计需要迭代,计算量较大.本文提出的 ATSMAP 方法对此做了改进,首先,Markov 四叉树模型的参数可以针对待分割图像重新优化;其次,采用了更简单的上下文模型参数估计方法,该方法无需迭代,计算量很小.对合成图像与 SAR 图像的分割实验表明,该方法在经过模型适应后的分割精度要高于 TSMAP 和其它几种基于多尺度 Markov 模型的图像分割方法.

参考文献:

[1] C A Bouman, M Shapiro. A multiscale random field model for Bayesian image segmentation[J]. IEEE Trans Image Processing, 1994, 3(3): 162 - 177.

[2] H Cheng, C A Bouman. Multiscale Bayesian segmentation using a trainable context model[J]. IEEE Trans Image Processing, 2001, 10(4): 511 - 525.

[3] M Aitin, D B Rubin, Estimation and hypothesis testing in finite mixture models[J]. J R Statist Soc B, 1985, 47(1): 67 - 75.

[4] M S Crouse, R D Nowak, R G Baraniuk. Wavelet-based statistical signal processing using hidden markov models[J]. IEEE Trans Signal Processing, 1998, 46(4): 886 - 902.

[5] J M Laferté, P Pérez, F Heitz. Discrete Markov image modeling and inference on the quadtree[J]. IEEE Trans Image Processing, 2000, 9(3): 390 - 404.

作者简介:

郭小卫　男,1971 年 9 月生于陕西礼泉,西北工业大学计算机应用技术专业博士生,主要从事统计信号与图像处理、模式识别和多尺度统计建模等方面的研究.E-mail: gxw423@msn.com.

田　铮　女,1948 年 11 月生于辽宁沈阳,西北工业大学教授,应用数学专业博士生导师,计算机应用技术专业博士生导师,主要从事非线性时间序列分析、非参数统计的研究和多尺度非线性随机模型与图像统计处理的研究.