

文章编号: 0258-7025(2009)Supplement 2-0350-05

基于全贝叶斯神经网络的图像小波先验模型

龙兴明¹ 周 静²

(¹ 重庆师范大学物理系, 重庆 400047
² 重庆大学输配电装备及系统安全与新技术国家重点实验室, 重庆 400044)

摘要 图像小波系数先验模型在图像处理中得到广泛的应用。已有小波系数的建模方法在模型选择、模型参数估计和非高斯噪声图像恢复等方面存在一定限制。利用全贝叶斯神经网络(FBNN)模型对图像小波系数的统计特性进行建模,利用现代粒子采样技术进行估计获得该模型的参数。对单尺度和父子尺度小波系数先验模型的仿真实验表明,基于全贝叶斯神经网络的小波先验模型建模准确,较好地描述了小波系数统计特性,把由此方法获得的单尺度和父子尺度小波系数先验粒子应用于图像去噪处理,仿真结果证实去噪处理后的图像质量在客观指标和主观视觉上都有显著的提高。

关键词 图像处理;小波系数先验模型;全贝叶斯神经网络;粒子采样

中图分类号 TP751 文献标识码 A doi: 10.3788/CJL200936s2.0350

Full Bayesian Neural Network Prior Statistical Modeling for Image Wavelet Coefficients

Long Xingming¹ Zhou Jing²

(¹ Physics Department, Chongqing Normal University, Chongqing 400047, China
² State Key Laboratory of Power Transmission Equipment and System Security and New Technology, Chongqing University, Chongqing 400044, China)

Abstract Wavelet coefficients prior statistical models of image have been studied widely in the Bayesian-based image processing scopes. In this paper, we derive a precise prior statistical model based on full Bayesian neural network (FBNN). The parameters of the model can be estimated empirically from a sample image set by modern particle samplers (Montel Carlo) methods. The simulated results based on the prior models of single scale and parent-children scale show the model makes it possible to exploit the dependency between the scales. Furthermore, a novel image denoising method based on scale prior particles sampled from the fitted the single scale and parent-children prior models produces the high quality visual effects and peak signal-to-noise ratio (PSNR).

Key words image processing; wavelet coefficients prior model; full Bayesian neural network; particle sampling

1 引 言

小波变换是一种强有力的数学分析工具,利用小波多分辨率分析特性进行图像处理一直是近年来的一个研究热点^[1,2]。图像小波系数先验模型在基于贝叶斯推断的去噪、压缩和运动跟踪等图像检测中得到广泛的应用^[3]。这类方法的关键是获得准确的图像小波系数的先验信息。研究发现,自然图像小波细节系数统计具有明显的零中心对称性、零点

最大性和拖尾性。Mallat^[4]首次提出对单尺度先验小波系数进行指数功率分布建模,获得了经典加性高斯噪声模型的图像恢复问题完备解析表达;Moulin 等^[5]把单尺度先验小波系数模型推广到广义高斯模型。为了进一步准确描述先验信息,Canditiis 等^[6~9]学者对小波系数的建模进行了多方面改进,取得了重要进展。

尽管针对经典的高斯噪声自然图像恢复问题,

基金项目: 重庆市教委科学技术(KJ090829)和重庆师范大学青年基金(08XLS13)资助课题。

作者简介: 龙兴明(1976—),男,副教授,硕士,主要从事半导体光电器件检测和信号处理技术等方面的研究。

E-mail: lennydragon@163.com

上述模型能够获得简单的、完备的解,但是在先验模型选择、模型的参数估计和非高斯噪声图像恢复等方面存在一定的限制^[3,6]。为了得到更一般的图像先验模型,本文提出利用全贝叶斯神经网络对图像的先验信息进行建模。仿真实验表明,采用全贝叶斯参数估计算法获得的先验模型具有很高的稳健性和精度,该模型具有较大的实用价值。

2 全贝叶斯神经网络先验模型

研究表明,自然图像有别于随机噪声图像的统计特性,自然图像的小波系数具有明显的统计特性:类高斯性、长拖尾性和噪声干扰性^[3]。为了获得更一般的图像先验模型,针对以上小波系数的统计密度分布特点,本文提出利用全贝叶斯神经网络对图像的先验信息进行建模,选择神经网络先验密度模型为

$$y = f(x; \theta) + e, \quad (1)$$

式中 e 为余量并假定概率分布为 $p_e(e)$, 近似函数 $f(x; \theta) = b^2 + \omega^2 \tanh(b^1 + \omega^1 x)$, $\theta = \{b^2, \omega^2, b^1, \omega^1\}$ 表示所有输入/输出隐含层的权重和基参数。

结合 Andrieu 的稳健全贝叶斯神经网络学习思想^[10], 推导建议先验模型(1)的参数确定过程。

2.1 先验模型参数的贝叶斯估计

从贝叶斯学派的观点看,在已知模型 M [如(1)

式]和训练数据 $D = \{x, y\}$ 的条件下,模型未知参数 θ 包含在后验分布 $p(\theta | D, M)$ 的信息中,并且在最小均方误差准则下未知参数的最优估计为

$$\hat{\theta} = E[\theta | D, M] = \int \theta p(\theta | D, M) d\theta.$$

根据贝叶斯公式,后验分布 $p(\theta | D, M)$ 为 $p(\theta | D, M) = p(D | \theta, M) p(\theta | M) / p(D | M) \propto p(D | \theta, M) p(\theta | M)$, (2)

(2)式中 $p(\theta | M) = p(b^2, \omega^2, b^1, \omega^1 | M)$, $p(D | M)$ 为模型条件下的数据概率,它是一个已知常数, $p(D | \theta, M) = p(x, y | \theta, M) = p_e(e | \theta, M) p(x | \theta, M)$ 。

因此,当对后验分布 $p(\theta | D, M)$ 进行采样获得粒子序列 $\{\theta^{(i)}\}_{i=1}^N$ 时,先验模型(1)未知参数 θ 的高维积分估计问题可以转化为简单均值运算: $\hat{\theta} = \sum_{i=1}^N \theta^{(i)} / N$ 。

2.2 模型参数的采样

对于多变量未知模型参数 θ 的采样,利用 Gibbs 采样可以转化为单变量采样^[11]。设初始参数

$$\theta^{(0)} = \{b^2, \omega^2, b^1, \omega^1\} = \{0, 0, 0, 0\},$$

$\theta_j^{(i)}$ 表示第 i 次采样的第 j 个参数分量, $\theta_j^{(i)}$ 表示 $\theta_j^{(i)}$ 的全条件分量。根据(2)式,在已知 $\theta_j^{(i)}$ 条件下 $\theta_j^{(i+1)}$ 的 Gibbs 采样目标函数为

$$\begin{aligned} \theta_j^{(i+1)} \sim p(\theta_j | \theta_j^{(i)}, D, M) &= \{p_e[y - f(x; \theta) | \theta_j, \theta_j^{(i)}, M] p(x | \theta_j, \theta_j^{(i)}, M)\} p(\theta_j, \theta_j^{(i)} | M) = \\ & \{p_e[e^{(i)}(\theta_j)] p(x | \theta_j, \theta_j^{(i)}, M)\} p(\theta_j, \theta_j^{(i)} | M) = p_e[e^{(i)}(\theta_j)] p(\theta_j | \theta_j^{(i)}, M) p(\theta_j^{(i)} | M) \propto \\ & p_e[e^{(i)}(\theta_j)] p(\theta_j | \theta_j^{(i)}, M), \end{aligned} \quad (3)$$

式中 $e^{(i)}(\theta_j) = y - f(x; \theta) | \theta_j, \theta_j^{(i)}, M$ 表示当前参数条件下的训练误差,它是当前采样参数的函数。

由(3)式知,粒子序列 $\{\theta^{(i)}\}_{i=1}^N$ 取决于误差分布以及参数的先验分布。对于模型(1)的基、权重以及冗余噪声的先验分布,采用自动相关判决(ARD)先验分布思想^[8],例如,当具有相同输入 x_k 连接的权重 w_{kj} 取相同 H 参数 $H = \{\alpha_k^2, \alpha_{ave}^2, v_a, \alpha_0^2, v_{a,ave}\}$ 时,

权重 w_{kj} 的 ARD 先验模型为: $w_{kj} \sim N(0, \alpha_k^2)$, 并且满足 $\alpha_k^2 \sim \text{Inv_gamma}(\alpha_{ave}^2, v_a)$, $\alpha_{ave}^2 \sim \text{Inv_gamma}(\alpha_0^2, v_{a,ave})$, 这里 $\text{Inv_gamma}(\cdot)$ 表示逆 Gamma 函数。

因此,由(3)式和权重 w_{kj} 的 ARD 先验模型, w_{kj} 的 HMC 采样能量函数 $E(w_{kj})$ 和梯度矢量 $\Delta(w_{kj})$ 为

$$\begin{cases} E(w_{kj}) \propto [e^{(i)}(w_{kj})]^T [e^{(i)}(w_{kj})] / 2\alpha_k^2 + w_{kj}^T w_{kj} / 2\alpha_k^2, \\ \Delta(w_{kj}) = \partial E(w_{kj}) / \partial w_{kj}. \end{cases} \quad (4)$$

由于采用了 ARD 先验模型,故应对 H 参数进行相应采样。若已知权重的 H 参数为 $H^{(i)} = \{\alpha_k^2, \alpha_{ave}^2, v_a, \alpha_0^2, v_{a,ave}\}$, 那么根据参数的 Gibbs 采样原理(3)式有

$$\begin{aligned} H_j^{(i+1)} \sim p(H_j | H_j^{(i)}, \theta^{(i)}, D, M) &\propto p(y | \theta^{(i)}, x, H^{(i)}, M) p(\theta^{(i)} | x, H^{(i)}, M) p(x | H^{(i)}, M) p(H^{(i)} | M) = \\ & p_e[e^{(i)}(\theta^{(i)})] p(\theta^{(i)} | x, H^{(i)}, M) p(H_j | H_j^{(i)}, M) \propto p(H_j | H_j^{(i)}, M). \end{aligned} \quad (5)$$

2.3 先验模型确定步骤

根据以上分析,先验模型参数估计的步骤为:

1) 对待分析的图像 $I(s) \in L^2(R^2)$ 进行二维离散正交小波变换^[4],其中图像的细节小波系数单尺度子带记为 $\{LH, HL, HH\}$,记 HH_{12} 表示父子小波分量 HH_1 和 HH_2 的父子关系。

2) 对各单尺度或者父子尺度子带进行统计。

3) 根据统计结果对模型(1)参数进行采样。未知模型参数 θ 采样具体步骤为: I)对未知的模型参数 θ 和 H 参数置初值; II)在其他参数不变条件下利用(4)式对权重进行 HMC 采样; III)利用 Gibbs 采样(3)式对基矢量进行更新; IV)根据(5)式对噪声、基和权重的 H 参数进行更新; V)进入新一轮采样。

4) 根据采样获得粒子序列 $\{\theta^{(i)}\}_{i=1}^N$,利用简单均值运算: $\hat{\theta} = \sum_{i=1}^N \theta^{(i)} / N$ 获得所需的模型参数。

3 实际图像的小波先验模型

根据上述建议方法,在 MATLAB 7.0 中,以标准“woman”图像为例,考察该图像经二阶正交小波变换后的单尺度和父子尺度联合分布特性,在此基础上,根据获得的模型进行粒子重采样,获得先验粒子,并对真实统计、拟合统计和重采样的统计性能进行比较。

3.1 单尺度先验模型

对父系数分量进行建模。图 1 中的 3 条曲线分别表示了 HH_1 分量中大小为 $-40 \sim 40$ 范围的小波系数真实统计结果(org)、建议模型的拟合结果(fit-FBNN)和基于建议模型重采样粒子的统计结果(rj-hist)。其中建议模型(1)采用单输入/单输出及 10 隐含层的神经网络结构,该模型的 31 个参数估计是进行 500 次迭代采样,再去除初始的 50 个数据,并

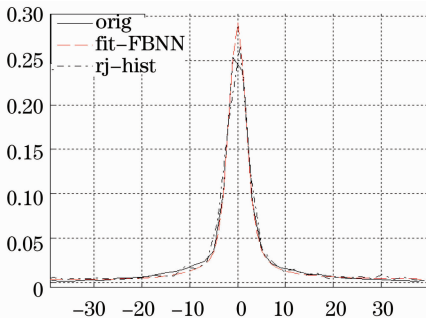


图 1 父系数分量 HH_1 统计结果

Fig. 1 Probability distribution of parent coefficients

HH_1

对余下数据隔十选一求均值所得。为了实际应用,以方差为 20 的零均值高斯函数作为比较函数(其中拒绝采样算法中的参数 $M=13$)进行拒绝采样^[11]。把经过 20000 迭代采样并剔除初始 1000 点的粒子作为单尺度先验粒子。由拟合结果可见,该模型达到了较高的拟合精度并且由此生成的先验粒子具有准确的代表性。类似地,可以得到二阶离散小波变换其余各子带的原始统计、拟合统计和先验粒子统计结果。

3.2 父子联合先验模型

为了描述小波系数之间的相关性,对二阶分解条件下的 3 个父子系数进行联合统计: HH_{12}, LH_{12} 和 HL_{12} 。图 2 表示了这 3 个父子小波系数联合统计结果。每个联合统计分布分块大小为 25,对该两输入(父和子系数)单输出(联合概率)的统计结果利用两输入单输出及 10 隐含层的先验模型进行拟合。其中模型中的 41 个参数估计结果是通过 300 次迭代采样去除初始 50 个数据后,对余下数据隔十选一求均值而获得。

类似于单尺度分析,对父子系数的拟合结果进行 MH 随机行走重采样^[11]。其中随机行走方差为 12,迭代次数为 5000。父子先验粒子为剔除初始 500 点后的 MH 随机行走重采样结果。

由图 2 可见,一方面,建议模型有效地抑制了噪声,提高了真实三维统计和二维投影结果;另一方面,根据建议模型生成的先验粒子,能够有效地代表先验信息。因此,由该方法获得的先验粒子能够有效地描述图像的先验信息,从而为推广基于粒子采样的图像非参数贝叶斯推断有着重要的作用。

4 先验模型在图像去噪中的应用

利用以上方法获得的先验模型以及由此生成的先验粒子,对噪声图像恢复问题进行研究。设一幅原始图像 x 受到噪声 n 干扰后的观测图像为 $g: g = x + n$;那么,在正交小波域中可以表示为: $y = w + e$ 。由贝叶斯理论知,最小均方误差准则下的恢复图像小波系数最优估计值为

$$\hat{w}(y) = E_w[p(w/y)] = E_w[p(w)p(y/w)] = E_w[p(w)p_e(y-w)]. \quad (6)$$

由 Montel Carlo 积分知:若获得的先验粒子为 $\{w^{(i)}\}_{i=1}^N$,则(6)式变为

$$\hat{w}(y) = \sum_i p_e(y - w^{(i)}) / N. \quad (7)$$

(7)式表明,对于任意噪声图像的恢复问题,当

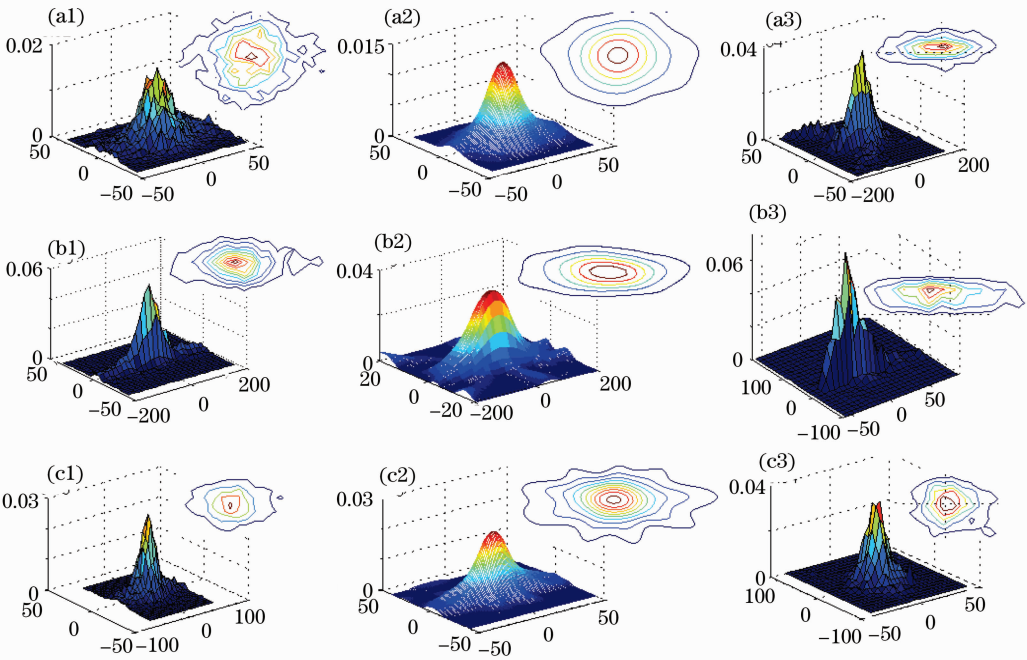


图 2 父子系数分量 $HH_{12}, LH_{12}, HL_{12}$ 的统计结果: (a1~a3)分别为父子系数分量 HH_{12} 的原始统计结果与投影、模型拟合结果与投影及重采样统计结果与投影, (b1~b3)为父子系数分量 LH_{12} 的原始统计结果与投影、模型拟合结果与投影及重采样统计结果与投影, (c1~c3)为父子系数分量 HL_{12} 的原始统计结果与投影、模型拟合结果与投影及重采样统计结果与投影

Fig. 2 Probability distribution of parent-children coefficients $HH_{12}, LH_{12}, HL_{12}$. (a1~a3): probability distributions of coefeciets HH_{12} , original, piror model and prior particles; (b1~b3): probability distributions of coefeciets LH_{12} , original, piror model and prior particles; (c1~c3): probability distributions of coefeciets HL_{12} , original, piror model and prior particles.

噪声的小波变换密度 $p_e(\cdot)$ 已知时,可以得到恢复图像小波系数的最优估计结果。对于受到独立同分布高斯白噪声干扰的经典图像去噪问题,由于经正交小波变换后的噪声密度仍然为高斯函数,即 $p_e(\cdot) =$

$N(e; 0, \sigma^2)$, 因此,经典恢复图像的收缩算子为

$$\hat{w}(y) = \sum_i N(y - w^{(i)}; 0, \sigma^2) / N.$$

选用图 3(a)图所示的标准“lenna.tif”图像,对受到独立同分布高斯白噪声干扰后的噪声图像

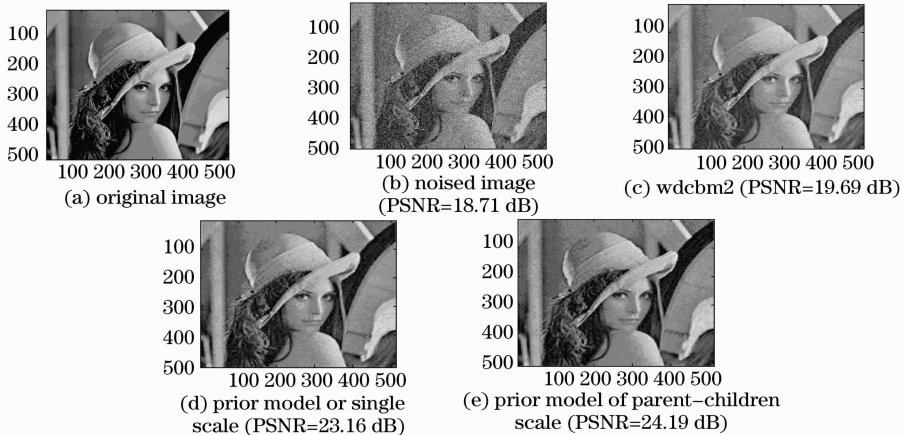


图 3 原始“lenna.tif”图像(a)、噪声图像(b)以及 3 种恢复图像结果:MATLAB7.0 标准去噪函数 wdcbm2(c)、单尺度先验粒子(d)和父子联合粒子(e)

Fig. 3 Original and erestored images "lenna.tif". (a) original; (b) noised images (PSNR=18.71 dB); (c) restored image by wdcbm2 of MATLAB7.0 (PSNR=19.69 dB); (d) restored image by the prior models of single scale (PSNR=23.16 dB); (e) restored image by the prior models of parent-children scale (PSNR=24.19 dB)

(PSNR 为 18.71 dB)利用(7)式进行去噪。在二阶正交小波分解条件下,基于单尺度先验粒子的恢复图像(PSNR 为 23.16 dB)和基于父子联合先验粒子的恢复图像(PSNR 为 24.19 dB)的仿真结果,同 MATLAB7.0 标准函数 wdcbm2 恢复的图像(PSNR 为 19.69 dB)相比,无论是客观指标还是视觉效果都有显著提高[如图 3(d),(e)所示]。

5 结 论

提出了基于全贝叶斯神经网络的图像小波系数统计特性建模方法,该方法不仅能够实现对小波系数先验模型的精确描述,而且是更一般的图像先验模型(例如,生物医学图像或宇宙图像往往不服从典型的先验模型)。其次,把粒子采样理论与小波系数先验模型进行了有效结合,通过仿真实验分析了图像小波系数单尺度和父子尺度的统计特性,由此模型生成的单尺度先验粒子和父子尺度粒子应用于基于小波贝叶斯推断的噪声图像恢复问题,获得了满足更一般噪声的粒子收缩解;与经典的图像去噪方法相比,基于该算法的恢复图像在客观指标和主观视觉都有显著的提高。

参 考 文 献

1 Zhang Jingjing, Fang Yonghua. Novel denoising method for remote sensing image based on Contourlet transform[J]. *Acta*

Optica Sinica, 2008, **28**(3): 462~466

张晶晶,方勇华.基于 Contourlet 变换的遥感图像去噪新算法[J]. *光学学报*, 2008, **28**(3): 462~466

2 Zhang Yudong, Dai Yun, Shi Guohua *et al.*. Application of one-dimensional wavelet transformation in time domain optical coherence tomography imaging[J]. *Chinese J. Lasers*, 2008, **35**(7): 1013~1016

张雨东,戴云,史国华等.一维小波变换在时域光学相干层析成像中的应用[J]. *中国激光*, 2008, **35**(7): 1013~1016

3 Peter Muller, Brani Vidakovic. Bayesian Inference in Wavelet Based Models (Lecture Notes in Statistics) [M]. New York: Springer-Verlag, 1999, **141**: 1~18

4 S. Mallat. A theory for multiresolution signal decomposition; the wavelet representation[J]. *IEEE Trans. Pattern Anal. Machine Intell.*, 1989, **11**(7): 674~693

5 P. Moulin, J. Liu. Analysis of multiresolution image denoising schemes using generalized Gaussian and complexity priors[J]. *IEEE Trans. Information Theory*, 1999, **45**(3): 909~919

6 De Canditiis, B. Vidakovic. Wavelet Bayesian block shrinkage via mixtures of normal-inverse gamma priors [J]. *J. Computational and Graphical Statistics*, 2004, **13**: 383~398

7 L. Boubchir, J. M. Fadili. Bayesian denoising based on the MAP estimation in wavelet-domain using Bessel K form prior[C]. *IEEE International Conference on Image Processing*, 2005, **1**: 113~116

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

9 L. Sendur, I. W. Selesnick. Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency[J]. *IEEE Trans. Signal Process*, 2002, **50**(11): 2744~2756

10 Christophe Andrieu, Nando de Freitas, Arnaud Doucet. Robust full Bayesian learning for radial basis networks [J]. *Neural Computation*, 2001, **13**(10): 2359~2407

11 J. C. Spall. Estimation via markov chain Monte Carlo[J]. *IEEE Control Systems Magazine*, 2003, **23**(2): 34~45