

基于贝叶斯神经网络先验模型的图像去噪研究*

龙 兴 明

(重庆师范大学 物理与信息技术学院, 重庆 400047)

摘要 基于小波系数先验模型的图像处理方法是现代图像处理技术的重要理论框架之一。针对传统的高斯型或拉普拉斯型先验模型的统计描述精度缺陷,本文提出利用贝叶斯(Bayesian)神经网络模型对图像小波系数的统计特性进行精确建模,结合现代粒子(或 Montel Carlo MC)采样技术——Gibbs 采样进行模型参数的估计,并重点考察了各尺度下的正交小波系数先验信息的建模过程,最后利用先验模型图像处理框架,实现图像的噪声抑制。仿真模拟结果表明:一方面,基于贝叶斯神经网络的小波先验模型建模准确,较好地描述了各尺度小波系数的先验信息;另一方面,从图像去噪性能来看,基于建议先验模型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 图像去噪 小波先验模型;贝叶斯神经网络 粒子采样

中图分类号:TB535

文献标识码:A

文章编号:1672-6693(2009)03-0065-04

基于图像小波系数先验模型的图像处理技术是图像去噪、压缩和运动跟踪等图像处理中的重要方法之一^[1]。这类方法的关键是获得准确的图像小波系数的先验信息。Mallat 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^[2];Moulin 等人把单尺度先验小波系数模型推广到广义高斯模型^[3]。为了进一步准确描述先验信息,Candès 扩展了 George 的标准高斯混合分布,并且假定未知高斯分布的方差先验为混合高斯-逆 Gamma 分布^[4];Boubchir 针对 Boesov 空间的图像提出 BKF (Bessel K Form) 先验分布^[5]。除了在单尺度建模外,考虑到自然图像的小波系数易于在尺度方向进行传播的特性。Crouse 采用 HMM(隐含马尔科夫模型)对先验小波系数的传播关系进行了建模^[6-7],Sendur 提出二元广义指数模型描述相邻尺度的联合统计分布^[8]。

尽管针对经典的高斯噪声自然图像恢复问题,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^[1,4]。为了得到更一般的图像先验模型,本文提出基于贝叶斯神经网络模型的先验信息建模方法,并实现基于建议先验模型的图像去噪。

1 基于先验模型的图像去噪

经典的基于小波系数先验模型的图像去噪框架描述如下,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: $y = w + e$,其中 y, w, e 分别表示观测图像 g , 原始图像 x 和噪声 n 的小波系数。

1.1 基本估计理论

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

$$\hat{w}(y) = E_w(p(w/y)) = E_w(p(w)p(y/w)) = E_w(p(w)p_e(y-w)) \quad (1)$$

其中, $p_e(\cdot)$ 表示噪声概率密度分布, $p(w)$ 表示先验小波系数的统计概率。

1.2 基于 Montel Carlo 积分估计理论

基于这一基本的估计理论框架,利用观察图像进行图像恢复时,除了高斯先验模型和高斯噪声等经典的概率分布外,对更一般的表达模型,最优估计小波系数值往往不具有完备的解析表达。

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

$$\hat{w}(y) = \sum_i p_e(y - w^{(i)})/N \quad (2)$$

(2)式表明,对于任意噪声图像的恢复问题,当噪声

* 收稿日期 2008-09-12 修回日期 2009-03-20

资助项目 重庆市教委科学技术项目(No. KJ090829) 重庆师范大学青年基金项目(No. 08XLS13)

作者简介 龙兴明,男,副教授,博士研究生,研究方向为信号检测与处理。

的小波变换密度 $p_e(\cdot)$ 已知时,可以得到恢复。

特别地,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 $p_e(\cdot) = N(e, \sigma^2)$,因此,经典恢复图像的收缩算子为

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

接下来重点讨论更一般的先验模型。

2 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^[1]。为了获得更一般的图像先验模型,本文根据小波系数的统计密度分布特点,提出利用贝叶斯神经网络模型对图像的先验模型进行建模,选择神经网络先验密度模型如下。

$$y = f(x; \theta) + e \quad (3)$$

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

结合 C. Andrieu 的鲁棒全贝叶斯神经网络学习思想^[9], 推导建议先验模型(3)的参数确定过程。

2.1 贝叶斯估计

从贝叶斯学派的观点看,在已知 M (模型(3))和 $D = \{x, y\}$ (训练数据)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

$$p(\theta | D, M) = p(D | \theta, M) p(\theta | M) / p(D | M) \propto p(D | \theta, M) p(\theta | M) \quad (4)$$

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

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(3)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:

$$\sum_{i=1}^N \theta^{(i)} / N。$$

2.2 Gibbs 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^[10]。设初始参数 $\theta^{(0)} = \{b^2, w^2, b^1, w^1\} = \{0, 0, 0, 0\}$, $\theta_j^{(i)}$ 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(4)式,在已知 $\theta_j^{(i)}$ 条件下 $\theta_j^{(i+1)}$ 的 Gibbs 采样目标函数为

$$\begin{aligned} \theta_j^{(i+1)} &\sim p(\theta_j | \theta_{-j}^{(i)}, D, M) = \\ &[p_e(y - f(x; \theta) | \theta_j, \theta_{-j}^{(i)}, M) p(x | \theta_j, \theta_{-j}^{(i)}, M)] \cdot \\ &p(\theta_j, \theta_{-j}^{(i)} | M) = [p_e(e^{(i)}(\theta_j)) p(x | \theta_j, \theta_{-j}^{(i)}, M)] \cdot \\ &p(\theta_j, \theta_{-j}^{(i)} | M) = p_e(e^{(i)}(\theta_j)) p(\theta_j | \theta_{-j}^{(i)}, M) \cdot \\ &p(\theta_{-j}^{(i)} | M) \propto p_e(e^{(i)}(\theta_j)) p(\theta_j | \theta_{-j}^{(i)}, M) \quad (5) \end{aligned}$$

其中, $e^{(i)}(\theta_j) = y - f(x; \theta) | (\theta_j, \theta_{-j}^{(i)}, M)$ 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(5)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(3)的基、权重以及冗余噪声的先验分布,笔者采用 ARD 先验分布思想^[9],例如,当具有相同输入 x_k 连接的权重 w_{kj} 取相同 H 参数, $H = \{\alpha_k^2, \alpha_{ave}^2, v_\alpha, \alpha_0^2, v_{\alpha,ave}\}$ 时,权重 w_{kj} 的 ARD 先验模型为 $w_{kj} \sim N(0, \alpha_k^2)$, 并且满足 $\alpha_k^2 \sim \text{Inv_gamma}(\alpha_{ave}^2, v_\alpha)$, $\alpha_{ave}^2 \sim \text{Inv_gamma}(\alpha_0^2, v_{\alpha,ave})$, 这里 $\text{Inv_gamma}(\cdot)$ 表示逆 Gamma 函数。

因此,由(5)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

$$\begin{cases} E(w_{kj}) \propto [e^{(i)}(w_{kj})]^T [e^{(i)}(w_{kj})] / 2\alpha_k^2 + w_{kj}^T w_{kj} / 2\alpha_k^2 \\ \Delta(w_{kj}) = \partial E(w_{kj}) / \partial w_{kj} \end{cases} \quad (6)$$

由于采用了 ARD 先验,故应对 H 参数进行相应采样。若已知权重的 H 参数为 $H^{(i)} = \{\alpha_k^2, \alpha_{ave}^2, v_\alpha, \alpha_0^2, v_{\alpha,ave}\}$, 那么根据参数的 Gibbs 采样原理, (5) 式有

$$\begin{aligned} H_j^{(i+1)} &\sim p(H_j | H_{-j}^{(i)}, \theta^{(i)}, D, M) \propto \\ &p(y | \theta^{(i)}, x, H^{(i)}, M) p(\theta^{(i)} | x, H^{(i)}, M) p(x | H^{(i)}, M) \cdot \\ &p(H^{(i)} | M) = p_e(e^{(i)}(\theta^{(i)})) p(\theta^{(i)} | x, H^{(i)}, M) \cdot \\ &p(H_j | H_{-j}^{(i)}, M) \propto p(H_j | H_{-j}^{(i)}, M) \quad (7) \end{aligned}$$

2.3 先验模型确定步骤

第 1 步 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^[2], 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。

第 2 步 对各单尺度或者父子尺度子带进行统计。

第 3 步 根据统计结果对模型(3)参数进行采样。未知模型参数 θ 采样具体步骤为: 1) 对未知的模型参数 θ 和 H 参数置初值; 2) 在其它参数不变的条件下利用(6)式对权重进行 HMC 采样; 3) 利用 Gibbs 采样(5)式对基矢量进行更新; 4) 根据(7)式

对噪声、基和权重的 H 参数进行更新 5) 最后进入新的一轮采样。

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

3 仿真实验

在 MATLAB 7.0 中,以标准 "woman" 和 "lenna.tif" 等图像为例,根据 (1) 式,进行建议小波系数先验模型的图像去噪模拟研究。实验主要包括:先验小波系数的统计概率 $p(w)$ 的性能分析以及同传统去噪方法的对比研究。

3.1 先验模型

考察标准 "woman" 图像经二阶正交小波变换后的各尺度分布特性。图 1 中(其中 x 轴表示小波系数的大小, y 轴表示所占的比重)的 3 条曲线分别表示了 HH_1 分量中小波系数大小为 $-40 \sim 40$ 的真实统计结果 (org)、建议模型的拟合结果 ($fit-FBNN$) 和基于建议模型粒子重采样的统计结果 ($rj-hist$)。其中,建议模型 (1) 采用单输入 / 单输出的 10 隐含层结构,该模型的 31 个参数估计是进行 500 次迭代采样,再去掉初始的 50 个数据后,并对余下数据隔 10 选 1 求均值所得。为了实际应用,对拟合结果以方差为 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^[9]。把经过 20 000 迭代采样并剔除初始 1 000 点的粒子作为单尺度先验粒子。

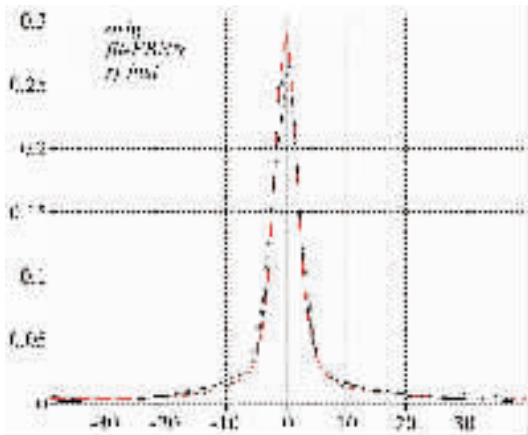


图 1 父系数分量 HH_1 统计结果

类似地,笔者给出了二阶离散小波变换其余 5 个子带的原始统计、拟合统计结果和先验粒子统计结果,如图 2 中所示。

由拟合结果可见,一方面,建议模型有效地抑制了噪声,达到了较高的拟合精度;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。

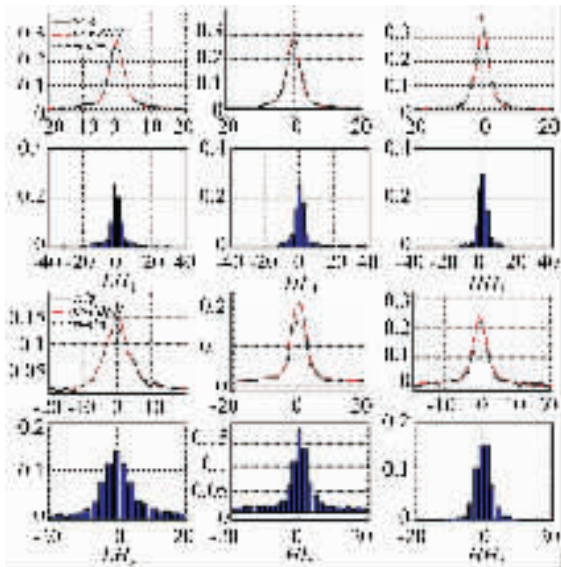


图 2 各尺度下的子带统计和相应的重采样

3.2 图像去噪

利用前部分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。选用标准 "lenna.tif" 图像(图 3 所示的原始图像),对受到独立同分布高斯白噪声干扰后的噪声图像(图 3 所示的观察图像)利用 (2) 式进行图像去噪处理。

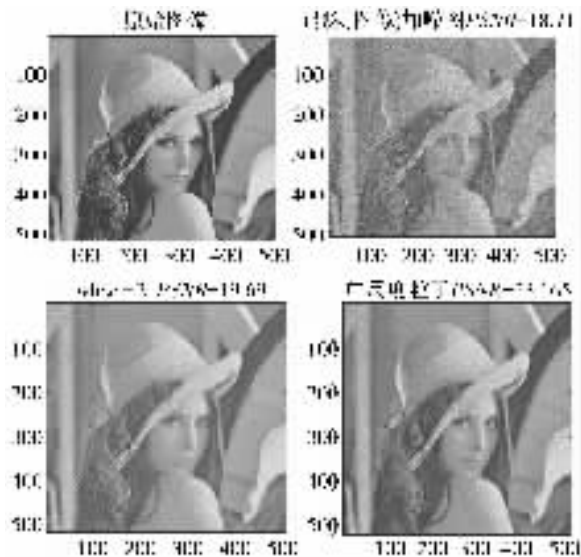


图 3 图像去噪性能对比

在二阶正交小波分解条件下,基于各单尺度先验粒子的恢复图像的仿真结果如图 3 中的“单尺度粒子”图像,其峰值信噪比 (PSNR) (matlab7.0 自带

函数为 $PSNR = 10 \times \log_{10} (\max(\max(x), \max(y))^2 / \|x - y\|^2)$ 从原来噪声图像的 $PSNR = 18.71$ dB (分贝) 提高到了 23.16 dB ;另外建议的算法其 $PSNR$ 指标也远高于 MATLAB7.0 的标准函数 wdcbm2 恢复图像 ($PSNR = 19.69$ dB) 结果。

综上所述,建议的图像去噪方法无论是客观指标还是视觉效果都有显著提高。

4 结论

本文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,通过实际仿真分析了图像小波系数在各单尺度下的统计特性,由此模型生成的先验粒子应用于基于小波贝叶斯推断的图像处理中。在噪声图像恢复问题,笔者获得了满足更一般噪声的粒子收缩解;通过与经典的图像去噪方法比较发现,基于该算法的恢复图像在客观指标和主观视觉都有显著提高。

参考文献:

[1] 龙兴明,郭世刚. 三维测量中基于脊波变换的图像预处理及拐点检测[J]. 重庆师范大学学报(自然科学版), 2007,24(2) 46-49.

- [2] Mallat S. A theory for multiresolution signal decomposition: the wavelet representation [J]. IEEE Trans Pattern Anal Machine Intell, 1989, 11(7) 774-793.
- [3] Moulin P, Liu J. Analysis of multiresolution image denoising schemes using generalized gaussian and complexity priors [J]. IEEE Trans Information Theory, 1999 45(3) 909-919.
- [4] De Canditiis D, Vidakovic B. Wavelet Bayesian block shrinkage via mixtures of normal-Inverse gamma priors [J]. Journal of Computational and Graphical Statistics, 2004, 13 383-398.
- [5] Boubchir L, Fadili J M. Bayesian denoising based on the MAP estimation in wavelet-domain using Bessel K form prior [J]. IEEE, 2005 1-116.
- [6] Crouse M S, Nowak R D. Wavelet-based signal processing using hidden Markov models [J]. IEEE Trans Signal Process, 1998, 46(4) 886-902.
- [7] 刘亮亮,敖军,高世泽. 基于灰色马尔可夫链模型的中国能源消费预测研究[J]. 重庆师范大学学报(自然科学版), 2008, 25(4) 47-49.
- [8] Sendur L, Selesnick I W. Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency [J]. IEEE Trans Signal Process, 2002, 50(11) 2744-2756.
- [9] Andrieu C, Freitas N D, Doucet A, et al. Robust full Bayesian learning for neural networks [C]. Drft of CUED/F-INFENG/TR, 1999.
- [10] Spall J C. Estimation via Markov chain monte carlo [J]. IEEE Control Systems Magazine, 2003 34-45.

Image Denoising Based Bayesian Neural Network Prior Statistical Modeling

LONG Xing-ming

(College of Physics and Information Technology, Chongqing Normal University, Chongqing 400047, China)

Abstract: Image processing based wavelet coefficients prior statistical models plays one of great important roles in modern image processing techniques. Owing to the defaults of fitting of Gaussian or Laplace functions, a Bayesian model of neural network (BMNN) to study the statistical dependency of wavelet coefficients is firstly presented. Secondly, its parameters are estimated by modern particle samplers (Monte Carlo) methods—Gibbs algorithm according to the characteristics of the suggested BMNN model. Then the relationship of wavelet coefficients is discussed in detail. Finally, a practical application of denoising image by using the BMNN model is demonstrated and the result shows that, on one hand the suggested method can express wavelet coefficients dependency efficiently, on the other, high quality visual effects and peak signal-to-noise ratio (PSNR) are achieved.

Key words: image denoising; wavelet coefficients prior models; BNN; particle samplers

(责任编辑 游中胜)