

邻域小波系数自适应的图像降噪

周登文 申晓留

(华北电力大学计算机科学与技术系, 北京 102206)

摘要 如何去除自然图像中的高斯白噪声是图像处理中的一个经典问题。基于小波收缩的 NeighShrink 降噪方法取得了很好的降噪效果,但是 NeighShrink 在所有小波子带上均使用了次优的 universal 阈值以及固定的邻域窗口尺寸,导致了较大的偏差,而且使得算法不健壮。为此,运用 Stein 的无偏风险估计改进了 NeighShrink 方法。该方法能够为每个小波子带确定最优的阈值和邻域窗口尺寸。实验结果显示,该方法取得了比 NeighShrink 更低的均方误差,也优于当前尖端的图像降噪算法—FeatShrink,其平均 MSE 大约低 6%。

关键词 自适应 图像降噪 邻域 小波

中图分类号:TP751.1 文献标识码:A 文章编号:1006-8961(2008)11-2112-05

Adaptive Image Denoising Using Neighboring Wavelet Coefficients

ZHOU Deng-wen, SHEN Xiao-liu

(Department of Computer Science and Technology, North China Electric Power University, Beijing 102206)

Abstract The denoising of a natural image corrupted by Gaussian noise is a classical problem in image processing. Chen *et al.* have proposed an effective wavelet thresholding scheme using incorporating neighboring coefficients, namely NeighShrink. NeighShrink's disadvantages are to use a suboptimal universal threshold and fixed neighboring window size for every subband. In this paper, we improve NeighShrink using Stein's unbiased risk estimation (SURE). Our method can determine an optimal threshold and neighboring window size for every subband. Experimental results show that our proposed method is significantly better than NeighShrink in all test examples. It also outperforms Balster *et al.*'s FeatShrink which is a recent sophisticated image denoising algorithm.

Keywords adaptive method, image denoising, neighborhood, wavelet thresholding

1 引言

在许多图像处理应用中,噪声去除是最普通、最重要的步骤。Donoho 和 Johnstone 创立的基于小波变换系数取阈值或收缩的方法,取得了很大的成功^[1,2]。该方法是为去除 1 维信号高斯白噪声而开发的,但是,能够推广到 2 维图像降噪。图像降噪问题可以描述为给定观察数据 $Y_{i,j}$ 作为信号 $X_{i,j}$ 的有噪实现:

$$Y_{i,j} = X_{i,j} + \varepsilon_{i,j}$$

$$i = 1, \dots, I; j = 1, \dots, J \quad (1)$$

其中, $\{\varepsilon_{i,j}\}$ 独立于信号,且是独立恒等分布 (iid) 的高斯随机变量,即 $\varepsilon_{i,j} \sim N(0, \sigma^2)$, σ^2 是噪声的方差。我们的目标是要基于 $Y_{i,j}$ 找到 $X_{i,j}$ 的最优估计,测度最优估计的客观准则通常是最小均方误差 (MSE)。

假定 \mathbf{Y}, \mathbf{X} 和 $\boldsymbol{\varepsilon}$ 分别表示式 (1) 观察数据矩阵、无噪数据矩阵和噪声矩阵, \mathbf{W} 表示正交小波变换矩阵,且 $\mathbf{w} = \mathbf{W}\mathbf{Y}$, $\boldsymbol{\theta} = \mathbf{W}\mathbf{X}$, $\mathbf{z} = \mathbf{W}\boldsymbol{\varepsilon}$,降噪过程分为以下 3 步:

(1) 计算小波系数矩阵 \mathbf{w} 。

(2)对细节系数 w 取阈值,即获得 X 的小波系数估计 \hat{w} 。

(3)对取阈值后的系数进行逆变换,即获得降噪后的估计 $\hat{X} = W^{-1}\hat{w}$ 。

在上述 3 个步骤中,关键是第 2 步,它涉及到两个问题,即阈值如何选取和对小波系数如何收缩,也就是选择怎样的收缩函数。有两种常用的收缩函数,软收缩函数 $\delta_\lambda^s(w) = \text{sgn}(w) \times \max(|w| - \lambda, 0)$ 将绝对值比阈值大的系数向零收缩,硬收缩函数 $\delta_\lambda^h(w) = w \cdot 1\{|w| > \lambda\}$ 保留绝对值比阈值大的系数,而将绝对值比阈值小的系数设置为零。由于硬收缩函数不连续,具有更大的方差,并且不够稳定,即对数据变化敏感,软收缩函数是优选的^[1]。

VisuShrink^[1]是最简单的降噪方法,它使用 universal 阈值,软或硬收缩函数。阈值 $\lambda = \hat{\sigma} \sqrt{2\log(M)}$, M 是式(1)中 I 和 J 中的较大者。VisuShrink 将这个阈值应用到小波各个分解尺度的细节子带,能够压抑噪声,但是也导致图像的过度光滑。SureShrink^[2]取得了更好的降噪效果,其阈值是数据驱动的。在各个分解子带上,使用软收缩函数,最小化 Stein 的无偏风险估计导出阈值。在 Donoho 和 Johnstone 开创性的工作之后,小波收缩降噪一直是一个活跃的研究领域。它主要往两个方向发展:一是针对某种收缩函数,优化阈值的选择^[3,4];二是寻找更优的收缩函数^[5,6]。这往往需要根据小波系数的统计性质,建立合适的小波系数模型。

Coifman 和 Donoho 指出:正交小波变换收缩降噪存在着视觉伪像^[7],它是由信号断点处 Gibbs 振动引起的。他们建议使用循环移动平均的方法来抑制这种缺陷,它等价于平移不变或稳态的小波变换。但是,这种方法增加了计算复杂度,需要更长的计算时间和更多的内存。

当前最好的图像降噪算法大都采用平移不变的小波变换,收缩小波系数时,充分利用了小波系数尺度内和尺度间的依赖关系^[8,9]。

改进了 Chen 等人中提出 NeighShrink 方法^[6]。在所有的测试实例中,该方法不仅一致胜出了 NeighShrink 方法,而且在大多数情况下,也胜出了 Balster 等人最近提出的一个尖端的图像降噪算法^[9],可称它为 FeatShrink 法。

2 自适应的邻域小波系数收缩

小波变换具有能量集中和系数簇聚的特点。也就是说,一个较大的小波系数的邻域系数也较大的概率很高。利用小波变换的这个性质,针对 1 维信号,Cai 和 Silverman 提出了邻域系数收缩估计方法^[10]。

假定 $w_{j,k}$ 是 1 维有噪信号的小波系数, $S_{j,k}^2 = w_{j,k-1}^2 + w_{j,k}^2 + w_{j,k+1}^2$,小波系数的收缩估计式为

$$\hat{w}_{j,k} = \left(1 - \frac{\lambda^2}{S_{j,k}^2}\right) w_{j,k} \quad (2)$$

λ 是 universal 阈值。计算 $S_{j,k}^2$ 时,如果 $w_{j,k}$ 超出了小波系数范围,则简单地忽略该项系数的计算。

Chen 等人将 Cai 和 Silverman 的估计方法用于图像降噪,并称之为 NeighShrink^[6]法。对每一个子带的每一个系数 $w_{j,k}$,考虑一个邻域窗口 $B_{j,k}$,窗口尺寸可以是 $3 \times 3, 5 \times 5, 7 \times 7$ 等。图 1 给出了一个邻域窗口实例,窗口中心是待取阈值的小波系数。

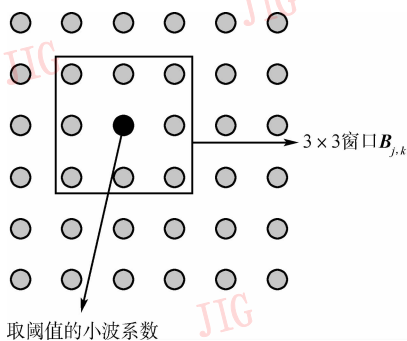


图 1 邻域窗口实例,窗口中心是取阈值的系数

Fig. 1 An illustration of the neighboring window centered at the wavelet coefficient to be shrunk

对于 2 维图像,每个子带小波系数的估计仍然使用式(2),但是, $S_{j,k}^2$ 的计算公式修改为

$$S_{j,k}^2 = \sum_{i,l \in B_{j,k}} w_{i,l}^2$$

NeighShrink 方法胜出了经典的 SureShrink 方法,存在的问题是:(1)所有子带均使用 universal 阈值,而 universal 阈值不是最优的;(2)邻域窗口的大小对降噪效果影响较大,而 NeighShrink 法只能根据经验选择邻域窗口尺寸,这往往产生很大的偏差;(3)NeighShrink 法对图像噪声级的适应性较差,不如 SureShrink 方法健壮。

为了弥补 NeighShrink 法的缺陷,建议使用自适

应的小波邻域系数估计方法。也就是对每一个子带,通过最小化 Stein 的无偏风险估计(SURE),计算最佳的阈值和最佳的邻域窗口尺寸。

为了简化记号,将子带 s 上的 N_s 个小波系数 $\mathbf{w}_s = \{w_{i,j}, i, j \in \text{位于子带 } s\}$ 重组为一个 1 维向量 $\mathbf{w}_s = \{w_n: n = 1, \dots, N_s\}$ 。同样,与 \mathbf{w}_s 对应的无噪小波系数 $\boldsymbol{\theta}_s = \{\theta_{i,j}, i, j \in \text{位于子带 } s\}$ 重组为 1 维向量 $\boldsymbol{\theta}_s = \{\theta_n: n = 1, \dots, N_s\}$ 。Stein^[11]证明:基于数据 \mathbf{w}_s ,对 $\boldsymbol{\theta}_s$ 的估计为 $\hat{\boldsymbol{\theta}}_s$,期望损失或风险(也即 MSE) $E\{\|\hat{\boldsymbol{\theta}}_s - \boldsymbol{\theta}_s\|^2\}$ 可被无偏估计。通常情况下,假定噪声标准差 $\sigma = 1$,则:

$$E\{\|\hat{\boldsymbol{\theta}}_s - \boldsymbol{\theta}_s\|_2^2\} = N_s + E\{\|g(\mathbf{w}_s)\|_2^2\} + 2\nabla \cdot g(\mathbf{w}_s) \quad (3)$$

其中, $g(\mathbf{w}_s) = \{g_n\}_{n=1}^{N_s} = \hat{\boldsymbol{\theta}}_s - \mathbf{w}_s$, $\nabla \cdot g \equiv \sum_n \frac{\partial g_n}{\partial w_n}$ 。

根据式(2),可以得到:

$$g_n(w_n) = \hat{\theta}_n - w_n = \begin{cases} -\frac{\lambda^2}{S_n^2} w_n & \text{当 } \lambda^2 < S_n^2 \text{ 时} \\ -w_n & \text{其他} \end{cases} \quad (4)$$

$$\|g_n(w_n)\|_2^2 = \begin{cases} \frac{\lambda^4}{S_n^4} w_n^2 & \text{当 } \lambda^2 < S_n^2 \text{ 时} \\ w_n^2 & \text{其他} \end{cases} \quad (5)$$

$$\frac{\partial g_n}{\partial w_n} = \begin{cases} -\lambda^2 \frac{S_n^2 - 2w_n^2}{S_n^4} & \text{当 } \lambda^2 < S_n^2 \text{ 时} \\ -1 & \text{其他} \end{cases} \quad (6)$$

假定邻域窗口的尺寸为 L (L 限定为大于 1 的奇数,如 3, 5, 7 等),令

$$SURE(\mathbf{w}_s, \lambda, L) = N_s + \sum_n \|g_n(w_n)\|_2^2 + 2 \sum_n \frac{\partial g_n}{\partial w_n} \quad (7)$$

则 $E\{\|\hat{\boldsymbol{\theta}}_s - \boldsymbol{\theta}_s\|_2^2\} = E\{SURE(\mathbf{w}_s, \lambda, L)\}$ 是期望风险的无偏估计。可通过选择给定子带 s 的邻域尺寸 L^s 和阈值 λ^s ,使 $SURE(\mathbf{w}_s, \lambda, L)$ 最小化,即

$$(\lambda^s, L^s) = \underset{\lambda, L}{\operatorname{argmin}} SURE(\mathbf{w}_s, \lambda, L) \quad (8)$$

上面的阈值和邻域尺寸是假定 $\sigma = 1$ 时导出的,对于非单位方差的数据,在使用式(8)计算阈值和邻域尺寸之前,应先估计噪声标准差 $\hat{\sigma}$,并将小波系数的噪声方差标准化为 1。一个好的噪声方差估计方法是用最精细分解层的对角子带系数绝对值的中值来估计^[1,2]:

$$\hat{\sigma} = \frac{\operatorname{median}(|\mathbf{w}_s|)}{0.6745} \quad (9)$$

($\mathbf{w}_s \in$ 最精细分解层对角子带)

3 实验结果

为了使实验结果尽可能准确、客观,选用了 4 个 512×512 标准的测试图像 Lena, Goldhill, Boats 和 Barbara,它们污染的高斯噪声有一个较宽方差范围,标准差分别为 5, 10, 20, 30 和 60。本文实现了 4 个降噪算法,包括经典的 SureShrink 方法,当前尖端的 FeatShrink 方法^[9],以及 NeighShrink 方法和本文中建议的自适应的邻域系数估计方法。4 个原测试图像如图 2 所示。



图 2 原图像

Fig. 2 Original images

所有小波变换使用 4 个分解层, Daubechies symmlet 8 个零矩小波。除 FeatShrink 方法使用平移不变的小波变换外,其他 3 种方法均使用快速正交小波变换。NeighShrink 方法采用了原作者推荐的 universal 阈值和缺省的 3×3 的邻域窗口尺寸;本文建议方法的阈值和窗口尺寸按式(8)计算。

均方误差(MSE)和信噪比(SNR)按下式计算:

$$MSE = \frac{1}{IJ} \sum_{i=1}^I \sum_{j=1}^J (X(i, j) - \hat{X}(i, j))^2 \quad (10)$$

$$SNR = 10\log_{10} \frac{\sum_{i=1}^I \sum_{j=1}^J X(i,j)^2}{\sum_{i=1}^I \sum_{j=1}^J (X(i,j) - \hat{X}(i,j))^2} \text{ (dB)} \quad (11)$$

噪声标准差均按式(9)估算。

4 个测试图像和 4 种方法在不同噪声级下计算出的均方误差(MSE)如表 1 所示,最小的 MSE 被加粗表示。

从表 1 可以看出,除 $\sigma = 60$ 情形以外, NeighShrink 的 MSE 均低于 SureShrink 的 MSE。正如所料,在所有的测试实例中,本文建议方法的 MSE 均一致低于 NeighShrink 的 MSE。当噪声级增大时, NeighShrink 的结果变坏,说明 NeighShrink 方法不够健壮,而本文建议的方法与经典的 SureShrink 方法一样健壮。

在这些测试实例中,本文建议方法的 MSE 大多数低于 FeatShrink 的 MSE,而且都是在较小噪声级时。通过统计表 1 中 MSE 的平均值,还发现,本文建议方法的 MSE 比 NeighShrink 方法大约低 19%,也比 FeatShrink 方法的 MSE 大约低 6%。FeatShrink 是当前一个尖端的图像降噪算法,采用了复杂的平移不变的小波变换以及充分利用了小波系数尺度内和尺度间的依赖关系。

一个有趣的事实是:对于多纹理的 Barbara,本文建议方法的 MSE 一致低于 FeatShrink 的 MSE, FeatShrink 方法甚至劣于 NeighShrink 方法。这个事实也说明:本文建议的方法比 FeatShrink 方法更为健壮。

通常, MSE 越低, SNR 越高, 图像降噪效果越好, 产生的图像质量越高。众所周知, MSE 和 SNR 并不总是同人的视觉一致, 还需要配合视觉观察来判断降噪效果。下面, 以污染噪声标准差为 30 的 Barbara 为例来观察由 SureShrink、FeatShrink 和本文建议方法的降噪效果。由于完整图像的降噪细节很难分辨, 故截取一个局部放大进行观察(如图 3 所示)。SNR 按式(11)计算, 3 种方法的信噪比分别为 19.47dB, 20.01dB 和 20.96dB。本文建议方法具有最低的 MSE 和最高的 SNR, 从图 3 也可以看出, 它降噪得到的图像的视觉质量也确实是最高, 清晰地保留了图像纹理细节。SureShrink 方法得到的图像更为模糊, 也残存了较多的伪像, FeatShrink 方法得到的图像虽然较为清晰, 但是, 图像被过度平

滑, 丢失了较多的纹理细节。

表 1 各种降噪方法 MSE 比较

Tab.1 The comparison of the denoising results(MSE)

		SureShrink	NeighShrink	FeatShrink	本文建议方法
Lena	$\sigma = 5$	13.61	11.67	11.96	11.48
	$\sigma = 10$	28.44	23.72	23.35	23.13
	$\sigma = 20$	59.26	52.75	45.85	48.24
	$\sigma = 30$	88.94	86.54	69.87	73.24
	$\sigma = 60$	168.19	212.34	138.84	147.21
Goldhill	$\sigma = 5$	18.82	18.23	18.71	16.69
	$\sigma = 10$	42.42	38.63	40.45	36.94
	$\sigma = 20$	87.54	81.25	79.03	75.75
	$\sigma = 30$	125.93	124.47	112.49	111.02
	$\sigma = 60$	215.06	262.25	191.94	197.02
Boats	$\sigma = 5$	20.40	18.69	18.73	18.41
	$\sigma = 10$	41.50	36.47	36.80	35.93
	$\sigma = 20$	88.80	80.03	76.16	75.63
	$\sigma = 30$	136.40	128.27	116.13	116.24
	$\sigma = 60$	251.97	285.37	220.86	227.24
Barbara	$\sigma = 5$	18.59	14.98	14.95	14.41
	$\sigma = 10$	47.81	35.81	37.04	34.70
	$\sigma = 20$	118.23	88.93	95.39	83.51
	$\sigma = 30$	189.30	149.22	167.20	134.50
	$\sigma = 60$	354.49	343.16	337.22	274.21

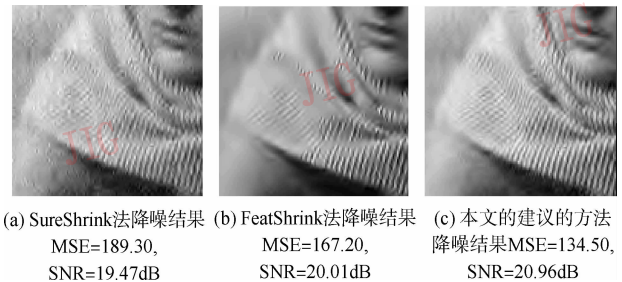


图 3 $\sigma = 30$ 时 Barbara 有噪图像降噪

Fig.3 Denoising results for Barbara, $\sigma = 30$

4 结 论

Chen 等人提出的 NeighShrink 降噪方法优于经典的 SureShrink 降噪方法。本文中, 改进了 NeighShrink 降噪方法。对小波系数的每一个分解子带, 通过最小化 Stein 的无偏风险估计(SURE), 计算最佳的阈值和最佳的邻域窗口尺寸。该方法取得了更好的效果。在所有测试实例中, 不仅均方误

差一致更低 (MSE 的平均值大约低 19%), 也克服了 NeighShrink 不够健壮的优点。

Balster 等人最近提出的 FeatShrink 方法是当前一个尖端的图像降噪算法, 它采用了平移不变的小波变换, 并充分利用了小波系数尺度内和尺度间的依赖关系。本文方法, 在大多数情况下, 也胜出了 FeatShrink, 平均 MSE 大约低 6%, 也比 FeatShrink 更健壮和取得了更好的视觉质量。

根据实验结果, 如表 2 所示给出了 4 种方法的综合性能对比。

表 2 4 种降噪方法综合性能对比

Tab. 2 The performance comparison of four denoising algorithms

	SureShrink	NeighShrink	FeatShrink	本文建议方法
平均 MSE	最大	较大	较小	最小
健壮性	健壮	不健壮	较健壮	健壮
视觉质量	差	较差	较优	优

参考文献 (References)

- 1 Donoho D L, Johnstone I M. Ideal spatial adaptation via wavelet shrinkage[J]. *Biometrika*, 1994, **81**(3): 425 ~ 455.
- 2 Donoho D L, Johnstone I M. Adapting to unknown smoothness via wavelet shrinkage [J]. *Journal of the American Statistical*

Association, 1995, **90**(432): 1200 ~ 1224.

- 3 Chang S G, Yu B, Vetterli M. Adaptive wavelet thresholding for image denoising and Compression[J]. *IEEE Transactions on Image Processing*, 2000, **9**(9): 1532 ~ 1546.
- 4 Xie J C, Zhang D L, Xu W L. Spatially adaptive wavelet denoising using the minimum description length principle [J]. *IEEE Transactions on Image Processing*, 2004, **13**(2): 179 ~ 187.
- 5 Hansen M, Yu B. Wavelet thresholding via mdl for natural images [J]. *IEEE Transactions on Information Theory*, 2000, **46**(5): 1778 ~ 1788.
- 6 Chen G Y, Bui T D, Krzyzak A. Image denoising with neighbour dependency and customized wavelet and threshold [J]. *Pattern Recognition*, 2005, **38**(1): 115 ~ 124.
- 7 Coifman R R, Donoho D L. Translation invariant denoising[A]. In *Wavelets and Statistics*, Springer Lecture Notes in Statistics 103[C], San Diego, CA, USA, 1995: 125 ~ 150.
- 8 Portilla J, Strela V, Wainwright M, *et al.* Image denoising using scale mixtures of gaussians in the wavelet domain [J]. *IEEE Transactions on Image Processing*, 2003, **12**(11): 1338 ~ 1351.
- 9 Balster E J, Zheng Y F, Ewing R L. Feature-based wavelet shrinkage algorithm for image denoising [J]. *IEEE Transactions on Image Processing*, 2005, **14**(12): 2024 ~ 2039.
- 10 Cai T T, Silverman B W. Incorporating information on neighbouring coefficients into wavelet estimation [J]. *Sankhya: The Indian Journal of Statistics, Series B*, 2001, **63**(2): 127 ~ 148.
- 11 Stein C. Estimation of the mean of a multivariate normal distribution [J]. *Annals of Statistics*, 1981, **9**(6): 1135 ~ 1151.