Image Denoising via Adaptive Soft-Thresholding Based on Non-Local Samples

Hangfan Liu, Ruiqin Xiong, Jian Zhang and Wen Gao
Institute of Digital Media, Peking University, Beijing 100871, China.

Regularization of denoising problems can generally be divided into two categories: local regularization and nonlocal regularization. Local regularization exploits the feature of image signals that they can be sparsely represented in transform domain, while nonlocal methods utilize the self-repeating patterns in images. The impressive success of nonlocal means (NLM) denoising [1] triggered a flurry of research works to utilize the nonlocal similarity of natural images. Different nonlocal denoising schemes utilize nonlocal correlation in various ways. For instance, NLM generates the estimated pixels by linear weighted average of nonlocal pixels in spatial domain, BM3D [2] enhances the sparsity of coefficients by an additional transform along the third dimension of three dimensional (3D) patch groups, LPG-PCA [4] treats nonlocal similar patches as samples, while CSR [3] gathers similar patches into a cluster and approximate them to the centroid.

This paper proposes a new image denoising approach using adaptive signal modeling and adaptive soft-thresholding. It improves the image quality by regularizing all the patches in image based on distribution modeling in transform domain. Instead of using a global model for all patches, it employ content adaptive models to address the non-stationarity of image signals. The distribution model of each patch is estimated individually and can vary for different transform bands and for different patch locations. In particular, we allow the distribution model for each individual patch to have non-zero expectation. To estimate the expectation and variance parameters for the transform bands of a particular patch, we exploit the non-local correlation of image and collect a set of similar patches as data samples to form the distribution. Irrelevant patches are excluded so that this non-local based modeling is more accurate than global modeling. Adaptive soft-thresholding is employed since we observed that the distribution of non-local samples can be approximated by Laplacian distribution.

Specifically speaking, the input image y is divided into overlapping patches of size $S \times S$. For the current patch $\mathbf{y}_i$, we search non-locally for its similar patches. The dissimilarity between two patches $\mathbf{y}_i$ and $\mathbf{y}_j$ is measured by Euclidean distance: $d(i,j) = \left\|\mathbf{y}_i - \mathbf{y}_j\right\|_2^2 / S^2$. Then the patches with distances smaller than a certain threshold τ are considered to be "similar" and stacked into group Y_i: $Y_i = \{\mathbf{y}_j | d(i,j) \leq \tau\}$, and the corresponding locations are recorded in set $\mathbb{L}_i$: $\mathbb{L}_i = \{j | d(i,j) \leq \tau\}$. Besides, the similar patches of $\mathbf{y}_i$ are used to train the PCA transform Φ_i. Apply the learned transform Φ_i to all the patches in Y_i. Let $\overrightarrow{\mu_i}$ and $\overrightarrow{\sigma_i}$ be the expectation vector and standard deviation vector of the $\mathbf{x}_i$, where the k-th entry of $\overrightarrow{\mu_i}$ and $\overrightarrow{\sigma_i}$ are the expectation and standard deviation of coefficients in band k respectively. Then the proposed objective function of this paper writes:

$$\tilde{\mathbf{x}} = \arg\min_{\mathbf{x}} \frac{c}{\sigma_n^2} \|\mathbf{y} - \mathbf{x}\|_2^2 + \sum_i \left\| \frac{\Phi_i \mathbf{x}_i - \overrightarrow{\mu_i}}{\overrightarrow{\sigma_i}} \right\|_1 . \quad (1)$$

In practice, the expectation and standard deviation of clean data is not available. The expectation $\overrightarrow{\mu_i}(k)$ is estimated as the median value of the coefficients in the k-th band because the coefficients are assumed to conform to Laplacian distribution. The advantage of such estimation is to exclude influences of outliers. Furthermore, since the noise is assumed to be i.i.d. Gaussian distributed, the standard deviation of coefficients in the k-th band $\overrightarrow{\sigma_i}(k)$ can be estimated by

$$\overrightarrow{\sigma_i}(k)^2 = \max\left(\overrightarrow{\sigma_{yi}}(k)^2 - \sigma_n^2, 0\right), \quad (2)$$

where $\overrightarrow{\sigma_{yi}}$ is the standard deviation vector calculated by coefficients of observed noisy image $\mathbf{y}$. Let $\overrightarrow{\beta_j} = \Phi_i \mathbf{y}_j$, $j \in \mathbb{L}_i$, then $\overrightarrow{\sigma_{yi}}$ is calculated as

$$\overrightarrow{\sigma_{yi}}^2 = \frac{1}{M} \sum_{j \in \mathbb{L}_i} \left(\overrightarrow{\beta_j} - \overrightarrow{\mu_i}\right)^2, \quad (3)$$

with M being the number of similar patches in the group.

Experimental results show that the proposed scheme outperforms the competing methods in most cases in terms of PSNR. When standard deviation of noise fluctuates from 10 to 90, the average PSNR of the proposed method is **0.46**~**1.14**dB higher than LPG-PCA, **0.26**~**0.39**dB higher than BM3D and **0.11**~**0.28**dB higher than CSR. As to perceptual quality, the images denoised by the proposed scheme exhibit much less noise and artifacts and better preserves details and textures (e.g. the tentacles of monarch in Fig. 1). Comparing with the three anchor schemes, the output of the proposed scheme is obviously more clean, sharp and visually pleasant.

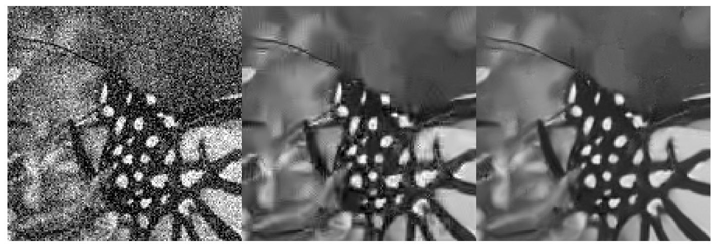

Figure 1: Perceptual comparison for *Monarch*. (a) Noisy image (σ_n = 40); (b) BM3D (PSNR = 25.77dB); (c) Proposed (PSNR = 27.17dB).

Figure 2: Perceptual comparison for *House*. (a) Noisy image (σ_n = 50); (b) BM3D (PSNR = 29.31dB); (c) Proposed (PSNR = 30.20dB).

Figure 3: Perceptual comparison for *Cameraman*. (a) Noisy image (σ_n = 30); (b) BM3D (PSNR = 28.27dB); (c)Proposed (PSNR = 28.63dB).

[1] A. Buades, B. Coll, and J.M. Morel. A non-local algorithm for image denoising. In *IEEE Conference on Computer Vision and Pattern Recognition*, June 2005.

[2] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Image denoising by sparse 3-d transform-domain collaborative filtering. *IEEE Transactions on Image Processing*, 16(8):2080–2095, Aug 2007.

[3] W. Dong, X. Li, D. Zhang, and G. Shi. Sparsity-based image denoising via dictionary learning and structural clustering. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 457–464, 2011.

[4] L. Zhang, W. Dong, D. Zhang, and G. Shi. Two-stage image denoising by principal component analysis with local pixel grouping. *Pattern Recognition*, 43(4):1531–1549, 2010.

This is an extended abstract. The full paper is available at the Computer Vision Foundation webpage.