

Nonlocal Spectral Prior Model for Low-level Vision

Shenlong Wang, Lei Zhang, Yan Liang

Northwestern Polytechnical University, The Hong Kong Polytechnic University

Abstract. Image nonlocal self-similarity has been widely adopted as natural image prior in various low-level vision tasks such as image restoration, while the low-rank matrix recovery theory has been drawing much attention to describe and utilize the image nonlocal self-similarities. However, whether the low-rank prior models exist to characterize the nonlocal self-similarity for a wide range of natural images is not clear yet. In this paper we investigate this issue by evaluating the heavy-tailed distributions of singular values of the matrices of nonlocal similar patches collected from natural images. A novel image prior model, namely nonlocal spectral prior (NSP) model, is then proposed to characterize the singular values of nonlocal similar patches. We consequently apply the NSP model to typical image restoration tasks, including denoising, super-resolution and deblurring, and the experimental results demonstrated the highly competitive performance of NSP in solving these low-level vision problems.

1 Introduction

Many image restoration problems such as denoising, super-resolution and deblurring are inherently ill-posed inverse problems. Solving these low-level vision tasks often needs regularization to yield high-quality results. Therefore, natural image prior models, which describe the ‘true’ statistics of natural image, play an important role in image restoration. The past decade has witnessed the rapid development on image prior modeling [1–14], and these prior models can be categorized into several categories: gradient (derivative, edge) based [2, 7, 8, 12, 14], filter-bank based [4, 9–11, 13], transform based [5, 15–17], etc.

Gradient-based image prior modeling is based on the fact that natural images usually contain only a small part of edge/texture regions, resulting in heavy-tailed distributions of high order statistics. Many models have been proposed to characterize these heavy-tailed distributions. For instance, Tappen et al. adopted Laplacian distribution to model gradients. Levin et al. [8] proposed a gradient sparsity-regularized approach to image deconvolution. Krishnan and Fergus [12] used hyper-Laplacian priors by minimizing the nonconvex l_q -norm ($q < 1$).

Filter-bank based prior models are also powerful for image restoration. Roth et al. [4] extended the Markov random field (MRF) framework by modeling marginal distribution of well-learned filter response. It is actually an extension of gradient based methods by formulating the responses of a series of derivative

filters as heavy-tailed distributed. Based on their work, Roth et al. [9] proposed a sparse steerable prior and Schmit et al. [18] extended MRF from a generative perspective for efficient modeling of natural statistics. Another similar approach is sparse representation based image restoration [5, 15–17]. With a well learnt dictionary over which natural image patches can be sparsely represented, restoration can be effectively conducted.

Nonlocal self-similarity (NSS) has been successfully used for image restoration problems. In their pioneer work, Buades et al. [3] regularized each image patch as the weighted average of its nonlocal neighboring patches. Dabov et al. [19] made use of NSS to construct 3D cubes of similar patches and conduct collaborative filtering on them to remove random noise. The so-called BM3D algorithm has been one benchmark for image denoising. In [16], Mairal et al. exploited NSS by using an $l_{p,q}$ -norm simultaneous sparse coding model. Dong et al. [17] proposed a centralized sparse representation model to exploit NSS in sparse domain. Zontak and Irani [20] proposed an ‘internal parametric prior’ to evaluate the patch recurrence of images for super-resolution.

Recently, Ji et al. [21] grouped the similar patches across spatial-temporal domain to form a low-rank matrix, and then presented a powerful nonlocal-based video denoising algorithm by using the recently developed low-rank matrix recovery (LRMR) technique [22]. Schaeffer et al. [23] implemented cartoon-texture separation by interpreting texture in low-rank patches. Taking advantage of low-rank interpretation, their method can effectively separated noise from texture.

The joint use of NSS and LRMR for image restoration is natural and very effective. By reforming the collected nonlocal similar patches into a 2D matrix, image restoration becomes essentially an LRMR problem. However, whether the low-rank prior models exist to characterize the NSS for a wide range of natural images has been rarely discussed. In this paper, we investigate this problem in detail. By theoretical and empirical analysis, we find that the distributions of singular values of the matrices formed by nonlocal similar patches are heavy-tailed and can be parameterized as generalized Gaussian distribution (GGD). Based on this observation, a novel natural image prior model, namely nonlocal spectral prior (NSP) model, is proposed by learning the parameters of GGD from natural images. Different from the widely used gradient prior models which exploit pixel-level high-order derivative statistics, the NSP model exploits structure self-similarities, and its parameter estimation is more robust to image degradation.

Due to the large variations of image content, using a single prior model cannot well describe natural images. Rather than learning a uniform prior, we cluster the training samples and learn a prior model for each cluster. In this way, multiple NSP models can be learnt to fit different image contents adaptively, and hence content-aware prior models can be efficiently estimated for image restoration. Moreover, MRF techniques can be adopted for robust parameter estimation along the nonlocal graph. Our experiments on image denoising, deblurring and super-resolution demonstrated the effectiveness of the proposed NSP model.

In summary, the contribution of our paper is twofold:

- We propose a novel way to use nonlocal similarity as image prior with content-awareness.
- We propose a flexible regularization of low-rank.

2 Nonlocal Spectrum Image Prior

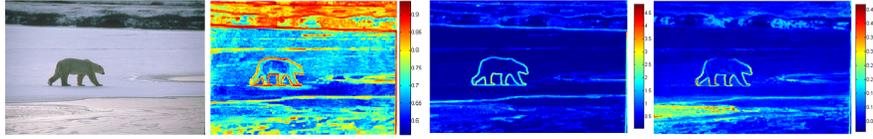


Fig. 1. Examples of singular value maps of nonlocal similar patches (patch size: 5×5). From left to right: original image, energy ratio of the first ten singular values, the largest and the 10th largest singular value maps.

2.1 Nonlocal Spectral Prior by Generalized Gaussian Distribution

The low-rank matrix recovery (LRMR) problem [22] aims at recovering the low-rank matrix \mathbf{X} from its noisy observation matrix \mathbf{Y} . Cai et al. proposed to solve this nonconvex problem by convex relaxation with the nuclear norm:

$$\min_{\mathbf{X}} \|\mathbf{X}\|_*, \text{ s.t. } \|\mathbf{X} - \mathbf{Y}\|_F^2 \leq \eta^2 \quad (1)$$

where $\|\cdot\|_*$ means the nuclear norm and η is the noise standard derivation. Eq. 1 is a convex relaxation of low-rank minimization. In [21] and some similar works [23], it is assumed that the matrix composed by similar patches (spatial-temporal or texture) exhibits low-rank characteristics. However, such a low-rank assumption in natural images lacks a solid probabilistic interpretation.

Fig. 1 illustrates the patch based singular value maps of an image. First, for each local patch (size: 5×5) we collect 49 nonlocal similar patches to it, forming a 25×50 matrix. Then the singular values of the matrix formed by these patches are calculated by SVD. From Fig. 1 we can see that in some local regions the matrix of nonlocal similar patches cannot be considered as low-rank because even the first 10 largest singular values together occupy no more than 70% of the whole energy, which means that in those areas the low-rank approximation cannot well describe the fine structures of natural images. Therefore, it is necessary to explore the probabilistic distribution of singular values of matrices formed by nonlocal similar patches (we call them *nonlocal matrices* in the following development).

Let's define function $\sigma_k(\mathbf{X}_i)$ as an operator to get the k^{th} largest singular value of matrix \mathbf{X}_i , where $\mathbf{X}_i = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_j, \dots, \mathbf{x}_{|N(i)|}]_{j \in N(i)}$ is constructed by the nonlocal similar patches \mathbf{x}_j to the given patch \mathbf{x}_i . $N(i)$ is the nonlocal neighborhood of \mathbf{x}_i by a modified nonlocal block matching scheme. In our block

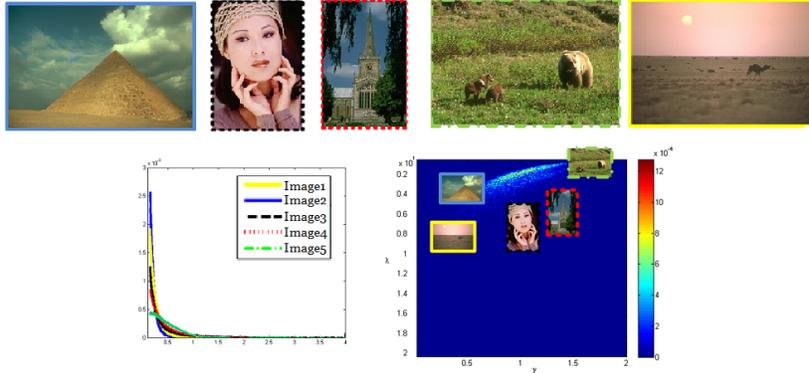


Fig. 2. Top: 5 natural images. Bottom left: heavy-tailed empirical distributions of the singular values of nonlocal matrices in the 5 images. Bottom right: distribution of estimated λ and γ in the GGD model from 20,000 image pieces of size 256×256 .

matching method, once a nonlocal patch is selected in a certain block, it will be deleted from the patch pool so that it cannot appear twice in our nonlocal blocks. We adopt this modification based on two considerations. Firstly, this will dramatically improve the computational efficiency. More importantly, in this manner, the independence between nonlocal blocks can be assured. It is a critical characteristics for constructing our jointly distribution via nonlocal graph in which nonlocal blocks are the maximal cliques. Our proposed nonlocal spectral prior aims to build a statistical model of the nonlocal redundancy. Obviously, for one patch which has many very similar patches within the searching region, $\sigma_k(\mathbf{X}_i)$ tends to be small. In Fig. 1, we have seen that for most local areas of a natural image, σ_k are relatively small since it is very likely to find many similar patches to them.

In fact, the nonlocal self-similarity (NSS) has been successfully exploited in image restoration [3]. Despite the wide use of NSS, there lacks an in-depth analysis of the low-rank characteristics of nonlocal similar patches. In Fig. 2, we plot the empirical distributions of the nonlocal singular values in five natural images. From Fig. 1 and Fig. 2, one can easily find that the NSS is *highly content dependent*, *spatially variant*, and the NSS induced nonlocal singular values are *distributed with heavy-tails*. Based on these observations, in this paper we propose a novel natural image prior, namely nonlocal spectral prior (NSP), and apply it to image restoration tasks. In particular, we parameterize the heavy-tailed distribution of nonlocal singular values by generalized Gaussian distribution (GGD)¹:

$$p(\mathbf{x}_i) = \frac{2\gamma_i \lambda_i^{1/\gamma_i}}{\Gamma(1/\gamma_i)} \exp(-\lambda_i \|\sigma(\mathbf{X}_i)\|^{\gamma_i}) \quad (2)$$

¹ Considering that all singular values are positive, the PDF that we adopted is the rectified GGD, i.e., a truncated GGD with interval $[0, \infty)$

where \mathbf{x}_i is the i^{th} patch of image \mathbf{x} and γ_i, λ_i are shape parameters.

We choose the GGD for two reasons. First, it is flexible to approximate various distributions of NSS induced singular values. Second, its parameters can be well estimated by some optimization approach. With GGD as the prior distribution of nonlocal singular values, our proposed NSP can regularize image restoration by measuring its NSS. Empirically, we have found that NSP works effectively and stably across various natural images. In Fig. 2 (bottom right) we plot the distribution of the estimated parameters λ and γ from 20,000 different image pieces (size: 256×256) collected from the Internet and BSDS 500 database [25]. The patch size is set as 5×5 and the 49 most similar patches to a given patch are collected to form a 25×50 nonlocal matrix. We can see that the estimated λ and γ are mostly located within $[2, 8], [0.6, 1.4]$, respectively.

On the other hand, let's test the sensitivity of the proposed NSP model for some typical image degradation. In Fig. 3, we add random noise to, blur, and down-sample an image, respectively, and plot the histograms of nonlocal singular values. We can find that the distributions for the degraded images differ dramatically to that of the original image. By fitting these empirical distributions with GGD in the parameter spaces $\gamma \in [0.6, 1.4]$ and $\lambda \in [2, 8]$, the KL-divergence of the fitted GGD with empirical distribution is 0.006, 3.794, 13.937, 0.333, 0.263 for the clean image, 10% noised image, 39% noised image, blurred image and down-sampled image, respectively. From this experiment, we can see that the degraded images cannot be well fitted by GGD with parameters for natural images, implying that the NSP can be used for image restoration tasks.

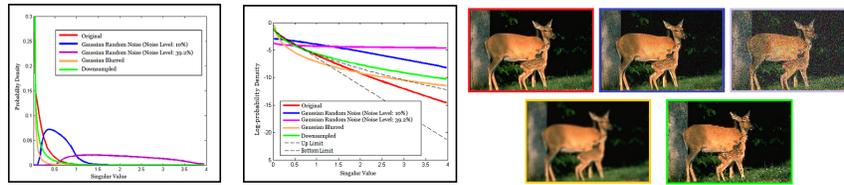


Fig. 3. Empirical analysis on degraded images. From left to right: Empirical PDF of nonlocal singular values, log-PDF of matched NSP model and ideal NSP model, illustration of degraded images

2.2 Content-awareness of Spectral Prior

From Fig. 2 we can find that although the proposed NSP model is stable in a small parametric space, the GGD fitting is still dependent on local image content. Clearly, for image patches which have more complex textural or structural patterns, it will be more difficult to find similar patches to them. The content-aware information plays a critical role in image restoration. For instance, in image denoising we would shrink the nonlocal singular value since the variance

between patches will be increased by random noise. However, for a patch which is unique in the searching range, its nonlocal singular values are relatively high and shrinkage may cause losing structural information. Therefore, in order for robust image restoration, it is crucial to model the prior adaptively to image content². For our proposed NSP model, one key issue is how to estimate λ and γ adaptively according to image content.

Several approaches have been proposed for adaptive parameter estimation in image restoration. Cho et al. [14] sampled small regions around the objective location and estimated GGD parameters from high-order statistics by support vector regression. However, in our problem it will be time consuming to calculate the high-order statistics of each patch by SVD. Moreover, without knowing clean image in prior, robust parameter estimation is necessary for practical image restoration. Since gradient-based statistics are not robust to noise, it is difficult to evaluate robustly the parameter from noisy images. Instead, by exploiting NSS in the proposed NSP model we could propose a fast parameter estimation approach which is robust to various kinds of image degradation.

3 Learning Nonlocal Spectral Prior for Image Restoration

3.1 Nonlocal Spectral Prior Learning

As discussed before, due to the variation of image content, learning a uniform NSP model for all nonlocal matrices is not accurate and robust. Therefore, multiple NSP models should be learnt and then applied adaptively based on image content. In this paper, we adopt vector quantization techniques to conduct multiple NSP model learning.

In the training stage, we extract patches from the sample images and partition the n training patches into clusters by a standard Gaussian-mixture clustering model which can be effectively solved by expectation-maximization algorithm. In each cluster, image patches share similar content, and we can assume that their nonlocal spectrums have similar GGD distribution. However, between different clusters, the nonlocal singular values have distinct distributions. Fig. 1 has actually illustrated the content-awareness of NSP.

After clustering, parameter estimation is conducted for each cluster. Since parameter estimation of GGD has no closed-form solution by methods of moments, numerical methods have to be used. Cho et al. [14] estimate γ and λ by learning a support vector regression model from second and forth moments. In this paper, we propose to estimate the GGD parameters by minimizing the KL-divergence. Since the KL-divergence of image patches is non-convex to γ and λ and the dimension of parameter space is only 2, we simply adopt a line searching strategy to learn the parameters of GGD in each cluster by minimizing

² Strictly speaking, the content-aware image prior is no longer an image prior model, because it makes use of posterior information. Similar discussions have been made by Cho et al. [14].

the KL-divergence between empirical distribution and parametric GGD:

$$[\hat{\gamma}, \hat{\lambda}] = \arg \min_{\gamma, \lambda} \text{KL}(p_E(\sigma) | p_D(\sigma)) = \arg \min_{\gamma, \lambda} \int_{\sigma} p_E(\sigma) \ln \left(\frac{p_E(\sigma)}{p_D(\sigma)} \right) d\sigma \quad (3)$$

In this way, we can compute k pairs of parameters γ and λ for the NSP model of each cluster. From Fig. 4 we can see that for different clusters, the parameters learnt are significantly different, which validates that NSP models differ with different image contents. For each cluster, due to the similarity of patch content, its learnt NSP model can better model the nonlocal singular value distribution than using a globally trained model from natural images. In the image restoration stage, we only need to estimate the model parameters for each given patch approximately by clustering. For each nonlocal matrix to the given patch centered at location i , we estimate parameters (λ_i, γ_i) by weighted sum via vector quantization.

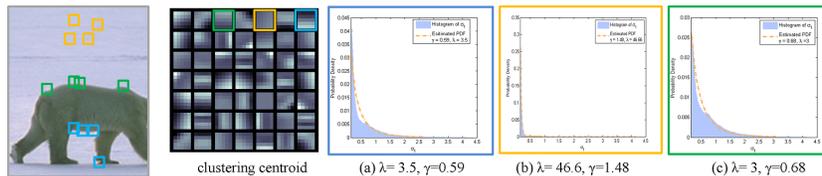


Fig. 4. Learning NSP models of different clusters. From left to right: an example training image, clustering centroid, empirical distributions and learnt GGD prior models for three clusters labeled by blue, yellow and green, respectively. Obviously, the models of smooth patches and edge patches are very different.

3.2 Robust Parameter Estimation by MRF

In addition, in order to improve the robustness of the estimation of (λ_i, γ_i) , we propose to use context information for each given patch. From Fig. 1 we can see that the nonlocal spectrum is highly correlated along edges. Therefore, we refine the parameter estimation by Markov Random Field (MRF) modeling on the graph constructed by nonlocal similar patches. Specifically, we perform the following energy minimization for accurate parameter estimation:

$$\{\gamma_i^*\} = \arg \min_{\{\gamma_i\}} \sum_{i \in \mathcal{V}} \left[(\gamma_i - \hat{\gamma}_i)^2 + \mu \sum_{j \in N(i)} w_{i,j} (\gamma_i - \gamma_j)^2 \right] \quad (4)$$

³ where μ is a regularization parameter to control the smoothness of MRF, and $w_{i,j}$ is the similarity weight defined by nonlocal similarities. As in [3], we set $w_{i,j} = \exp(-\frac{1}{\rho^2} \|\mathbf{x}_i - \mathbf{x}_j\|_2^2)$ where ρ is a constant to control the decay rate w.r.t. the distance between \mathbf{x}_i and \mathbf{x}_j . Eq. (4) is a standard MRF-MAP framework and many methods [26] can be used to effectively solve it.

³ The same for λ_i

3.3 Image Restoration with Nonlocal Spectral Prior

The image degradation process can be generally modeled as:

$$\mathbf{y} = \mathbf{h}(\mathbf{x} \otimes \mathbf{k}) + \mathbf{n} \quad (5)$$

where \mathbf{x} is the unknown clean image, \mathbf{h} is the downsampling operator, \mathbf{k} is the blurring kernel, \mathbf{n} is additive Gaussian white noise and \mathbf{y} is the degraded observation.

Image restoration aims to recover \mathbf{x} from the degraded image \mathbf{y} , given kernel \mathbf{k} and the distribution of random noise \mathbf{n} . In case \mathbf{k} is unknown, it will be a blind image restoration problem and we could estimate \mathbf{k} before estimating \mathbf{x} , or estimate them alternately. In this work, we assume that \mathbf{k} is known. The most popular approach for image restoration is to conduct maximum-a-posterior (MAP) estimation of \mathbf{x} :

$$p(\mathbf{x}|\mathbf{y}, \mathbf{k}, \mathbf{h}, \eta) \propto p(\mathbf{y}|\mathbf{x}, \mathbf{k}, \eta)p(\mathbf{x}) \quad (6)$$

where η is the standard deviation of noise \mathbf{n} and the likelihood function is standard Gaussian distribution:

$$p(\mathbf{y}|\mathbf{x}, \mathbf{k}, \mathbf{h}, \eta) \propto \exp(-\|\mathbf{y} - \mathbf{h}(\mathbf{x} \otimes \mathbf{k})\|_2^2/\eta^2) \quad (7)$$

Therefore, in order to estimate the unknown image with high quality, we can apply the NSP model learned from natural images to the estimation of \mathbf{x} :

$$p(\mathbf{x}) = \frac{1}{Z} \prod_{i=1}^n p_i(\mathbf{x}_i) = \frac{1}{Z} \prod_{i=1}^n \frac{\gamma_i \lambda_i^{\frac{1}{\gamma_i}}}{\Gamma(\frac{1}{\gamma_i})} \exp(-\lambda_i \|\sigma(\mathbf{x}_i)\|^{\gamma_i}) \quad (8)$$

where Z is a partition function to normalize density. Considering the dependence between nonlocal patches and the content-aware parameter estimation, our proposed NSP model can be technically seen as a conditional random field. Based on the specific block matching approach we adopt, the strictly equation can be ensured according to the independence between different nonlocal blocks (cliques). Therefore, the estimated image \mathbf{x} can be obtained by minimizing the log-posterior as follows:

$$\begin{aligned} \hat{\mathbf{x}} &= \arg \min_{\mathbf{x}} -\log p(\mathbf{x}|\mathbf{y}, \mathbf{k}, \mathbf{h}, \eta) = \arg \min_{\mathbf{x}} -\log p(\mathbf{y}|\mathbf{x}, \mathbf{k}, \mathbf{h}, \eta)p(\mathbf{x}) \\ &= \arg \min_{\mathbf{x}} \left\{ \frac{1}{\eta^2} \|\mathbf{y} - \mathbf{h}(\mathbf{x} \otimes \mathbf{k})\|_2^2 + \sum_{i=1}^n \lambda_i \|\sigma(\mathbf{x}_i)\|^{\gamma_i} \right\} \end{aligned} \quad (9)$$

3.4 Optimization

Similar to the patch-based likelihood proposed by Zoran et al. [13], we solve Eq. (9) by using alternating optimization as follows:

1. Solve auxiliary variables $\{\mathbf{X}_i\}$ by:

$$\hat{\mathbf{X}}_i = \arg \min_{\mathbf{X}_i} \left\{ \frac{1}{\eta^2} \|\mathbf{P}_i \mathbf{x} - \mathbf{X}_i\|_2^2 + \sum_{i=1}^n \lambda_i \|\sigma(\mathbf{X}_i)\|^{\gamma_i} \right\} \quad (10)$$

2. Reconstruct \mathbf{x} by $\{\hat{\mathbf{X}}_i\}$ and perform gradient descent: $\mathbf{x} = \mathbf{x} - \zeta \nabla_x (\|\mathbf{h}(\mathbf{x} \otimes \mathbf{k}) - \mathbf{y}\|_F^2)$

where \mathbf{P}_i is the linear operator to extract nonlocal matrix at location i from image \mathbf{x} . When $\gamma \geq 1$, it is a convex optimization problem which can be effectively solved by gradient-based optimization techniques in matrix trace function. When $\gamma < 1$, Eq. (10) is a non-convex optimization problem. In this case, convex optimization approach cannot ensure to find the global optimum. However, iterative reweighted approaches [27] can be adopted to tackle this problem by solving a series of weighted l_1 -minimization problem. Motivated by those works, we propose the following iteratively reweighted singular vector thresholding algorithm. Considering the following optimization problem:

$$(q - \text{Nuclear}) \min_{\mathbf{X}} \tau \|\sigma(\mathbf{X})\|_q + \frac{1}{2} \|\mathbf{X} - \mathbf{Y}\|_F^2 \quad (11)$$

where $q < 1$. The 1st order Taylor expansion of $\|\sigma(\mathbf{X})\|_q$ in terms of σ is:

$$\|\sigma(\mathbf{X})\|_q = \|\sigma_0\|_q + (\sigma(\mathbf{X}) - \sigma_0)/\sigma_0^{1-q} \quad (12)$$

Since each entry of $\sigma(\mathbf{X})$ is nonnegative, we can use the following updating strategy to solve Eq. (11):

$$\mathbf{X}^{(t+1)} = \arg \min_{\mathbf{X}} \sum_k \tau \sigma_k(\mathbf{X}) / \sigma_k(\mathbf{X}^{(t)})^{1-q} + \|\mathbf{X} - \mathbf{Y}\|_F^2 / 2 \quad (13)$$

Eq. (13) is a weighted nuclear norm minimization problem:

$$(\text{Weighted} - \text{Nuclear}) : \min_{\mathbf{X}} \tau \sum_k w_k \sigma_k(\mathbf{X}) + \frac{1}{2} \|\mathbf{X} - \mathbf{Y}\|_F^2 \quad (14)$$

Cai et al. have proved that the optimal solution of the standard nuclear norm minimization can be simply achieved by the singular shrinkage operator: $D_\tau(\mathbf{Y}) = \mathbf{U} \text{diag}(\{\max(\sigma_k - \tau, 0)\}) \mathbf{V}$ [22]. Similar to the Theorem 2.1 in that paper, we can easily derive the closed-form solution of the weighted-nuclear problem:

$$D_{\tau, w}(\mathbf{Y}) = \mathbf{U} \text{diag}(\{\max(\sigma_k - \tau w_k, 0)\}) \mathbf{V} \quad (15)$$

Clearly Eq. (15) is the proximal operator of convex function $\sum_j^k w_j \sigma_j(\mathbf{X})$ and thus we can adopt an iterative optimization approach to solving Eq. (9). Please find the detailed restoration algorithm in our supplementary materials.

4 Experimental Results

We evaluate the performance of the proposed NSP model for various image restoration tasks, including denoising, deblurring and super-resolution. For each task, we compare our method with state-of-the-art algorithms designed for that application. Due to the limit of page length, we only list the PSNR results and show one image for visual comparison in each task. More results and source codes are provided in our website: <http://www4.comp.polyu.edu.hk/~cslzhang/NSP.htm>.

4.1 Prior Learning

For all image restoration tasks, we used the same NSP model trained from BSDS500 dataset [25], which includes 500 pieces of 482×321 natural images. In the training stage, we first conduct block matching for each image and build the nonlocal graph. The patch size is set as 5×5 and for each patch 49 most similar patches to it are collected. SVD is used to get the nuclear norm of all nonlocal matrices. We then cluster the central patch of each nonlocal matrix by using Gaussian mixture model with 50 components. For each cluster, the histogram of its singular values is computed and the parameters λ and γ of the GGD fitting models are estimated. For our training set which includes 1 million nonlocal matrices, the whole training stage takes almost 4 hours. Block matching and SVD are the two most time-consuming processes, which however can be accelerated by adopting partial-SVD and nonlocal filtering techniques.

4.2 Image Denoising

The denoising performance of the proposed NSP algorithm is verified on the Kodak PhotoCD dataset (<http://r0k.us/graphics/kodak>), which contains 24 images of size 512×768 . Gaussian white noise of 5 different standard deviations (10, 15, 20, 25, 50) are added to the original images to simulate the noisy images. In our NSP based denoising algorithm with the trained NSP model, the balancing parameter τ is the only parameter to set. In our experiment, we set τ as 5 for all noise level. We compare our algorithm with K-SVD [5] and the benchmark BM3D [19] as well as three representative algorithms with image prior model: FoE [4], NLR-MRF [24] and EPPL [13]. Among them BM3D is the state-of-the-art in terms of accuracy. In order to prove the effectiveness of the proposed GGD model, we have also conducted a comparison experiments with the fixed nuclear norm regularization (i.e. to set γ with a fixed value of 1), which we denote as NL-LR(nonlocal low-rank regularization). The average PSNR of the 24 images



Fig. 5. Example of denoising results. From left to right: original, noisy, EPPL [13], BM3D [19], our proposed NSP method.

are reported in Table 1, from which we can find that our NSP method is slightly worse than the state-of-the-art BM3D method while being slightly better than

Table 1. Denoising Results (PSNR) on the Kodak 24 Image Datasets.

| Noise Levels | 10 | 15 | 25 | 25 | 50 |
|--------------|--------------|--------------|--------------|--------------|--------------|
| FoE[4] | 33.46 | 31.35 | 29.92 | 27.88 | 16.16 |
| K-SVD[5] | 33.72 | 31.62 | 30.21 | 29.14 | 21.90 |
| BM3D[19] | 34.39 | 32.30 | 30.92 | 29.91 | 26.98 |
| NL-MRF[24] | 34.21 | 32.11 | 30.69 | 29.67 | 26.48 |
| EPPL[13] | 34.28 | 32.21 | 30.81 | 29.77 | 26.74 |
| NL-LR | 34.39 | 32.16 | 30.60 | 29.61 | 26.30 |
| NSP | 34.46 | 32.31 | 30.84 | 29.77 | 26.63 |

the prior model based methods NLR-MRF and EPPL. All of them are much better than K-SVD and FoE. Moreover, with the flexible mixture formulation our model outperforms the fixed low-rank regularization (i.e. $\gamma = 1$). Visual quality comparison of the best three competing methods, BM3D, EPPL and NSP, is illustrated in Fig. 5. We can see that EPPL generates some artifacts; BM3D over-smooths the fine textures and brings blurring effects; our NSP method preserves texture information well without generating much blurring effect, leading to very pleasing visual quality. More results can be found in the supplementary file.

4.3 Image Super-resolution

For single image super-resolution, the low-resolution (LR) image is obtained by downsampling the blurred high-resolution (HR) image. We compare our proposed NSP based method with some well-known super-resolution methods, including Softcut [28], TV-based method [29], Sparsity based method [15] and CSR(centralized sparse representation) [17]. In our experiments, 8 commonly used images in literature are selected as the testing image. The LR images are



Fig. 6. Example of super-resolution results. From left to right: LR, HR, Sparsity [15], TV, CSR [17], our proposed NSP method.

Table 2. Super-resolution Results (PSNR).

| Image Name | Raccoon | Plants | Butterfly | Parrots | Hat | Girl | Flower | Bike |
|---------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| TV [29] | 27.54 | 31.34 | 26.56 | 27.85 | 29.20 | 31.24 | 27.51 | 23.66 |
| Softcut [28] | 27.82 | 31.19 | 24.74 | 27.99 | 29.50 | 31.82 | 27.31 | 23.15 |
| Sparsity [15] | 28.51 | 31.55 | 24.70 | 28.70 | 29.63 | 32.87 | 27.87 | 23.23 |
| CSR [17] | 29.29 | 34.00 | 28.19 | 30.68 | 31.33 | 33.67 | 29.54 | 24.72 |
| NSP | 29.23 | 34.22 | 27.84 | 30.30 | 31.34 | 33.68 | 29.70 | 24.41 |

simulated by first blurring the original HR image with a 7×7 Gaussian kernel (standard deviation: 1.6) and then downsampling with a scaling factor of 3. In this experiment, two parameters of NSP need to be set: the balancing parameter τ and gradient descent step ζ . We choose $\tau = 6$ and $\zeta = 1.0$ for all the images. The PSNR results are reported in Table 2. We can see that our method has very competitive performances with CSR and significantly outperforms all the other competing methods. The visual quality of several competing algorithms is compared in Fig. 6. From the zoom-in image, one can find that our NSP method can preserve sharp edges as well as complex texture regions, although it is not an edge-preserving image prior.

4.4 Image Deblurring

In this section, we first conduct experiments with simulated Gaussian blur kernel and uniform blur kernel, respectively, and then conduct experiments with real motion blur kernels. In the first experiment, we simulated the blur images by Gaussian and uniform kernels, respectively, and then added Gaussian noise to them (noise level: $\sqrt{2}$ and 2). We choose τ as 6 and gradient descent step length ζ as 0.6 in the deblurring experiment. We compare our method with several state-of-the-art deblurring methods, including FISTA [29], SA-DCT [30], BM3D [19] and CSR [17]. Note that the recently developed CSR has shown very strong deblurring capability. Three images are used in this experiment, and the PSNR results are listed in Table 3. We can see that our method outperforms all the competing algorithms in average. For examples of visual quality comparison and more details, please refer to the attached supplementary file.

To validate more comprehensively the effectiveness of NSP, we conduct deblurring with real motion blur kernels. The dataset we adopted is a standard testing dataset for motion deblurring from Levin et al. [31]. In this dataset, there are 4 images in total with 8 real-world motion blur kernels. For each blurred image \mathbf{y} , its corresponding original image \mathbf{x} and blur kernel \mathbf{k} are provided. We use two state-of-the-art methods for motion deblurring as our competing algorithms: the hyper-Laplacian proposed by Krishnan et al. [12] and the iterative reweighted l_1 -minimization proposed by Levin et al. [8]. Both the two competing approaches are based on gradient prior model. For each test image we draw a bar plot of the PSNR results in Fig. 7. A deblurring example is also shown in Fig. 7. We can see that our algorithm results in very competitive performance with Levin et al. [8], although it is not specially designed for motion blur deconvolution.

Table 3. Deblurring Results (PSNR) for Simulated Blur.

| Noise Image | Uniform, $\sigma = \sqrt{2}$ | | | Uniform, $\sigma = 2$ | | | Gaussian, $\sigma = \sqrt{2}$ | | | Gaussian, $\sigma = 2$ | | |
|----------------|------------------------------|--------------|--------------|-----------------------|--------------|--------------|-------------------------------|--------------|--------------|------------------------|--------------|--------------|
| | Butt. | Parr. | Leav. | Butt. | Parr. | Leav. | Butt. | Parr. | Leav. | Butt. | Parr. | Leav. |
| FISTA [29] | 28.37 | 29.11 | 26.49 | 27.73 | 28.88 | 26.03 | 30.36 | 31.23 | 29.33 | 29.67 | 30.74 | 28.62 |
| SA-DCT [30] | 27.50 | 30.07 | 27.04 | 26.46 | 29.10 | 25.86 | 29.85 | 32.46 | 29.70 | 29.42 | 31.79 | 29.16 |
| BM3D [19] | 27.21 | 30.50 | 27.45 | 26.56 | 29.75 | 26.60 | 29.01 | 32.22 | 29.67 | 28.56 | 31.72 | 29.10 |
| CSR [17] | 29.75 | 32.09 | 29.97 | 28.66 | 30.57 | 28.64 | 30.75 | 33.44 | 31.44 | 30.14 | 32.60 | 30.56 |
| NSP | 30.10 | 32.27 | 29.54 | 29.14 | 31.20 | 29.15 | 30.82 | 33.22 | 31.42 | 30.46 | 32.93 | 30.93 |

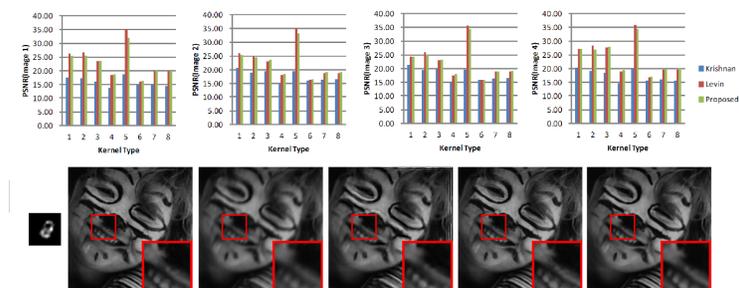


Fig. 7. Motion blurring results on dataset of Levin et al. [31]. Top: PSNR results. Bottom: An example of reconstructed images. From left to right: blur kernel, original image, blurred image, Hyper-Laplacian [12], IRLS-L1 [8], our proposed NSP.

5 Conclusions

In this paper we proposed a novel image prior, namely nonlocal spectral prior (NSP), by analyzing the heavy-tailed distribution of singular values of matrices constructed by nonlocal similar patches. The NSP builds a bridge between spectral analysis and image prior learning.

References

1. Freeman, W., Jones, T., Pasztor, E.: Example-based super-resolution. *Computer Graphics and Applications*, IEEE **22** (2002) 56–65
2. Tappen, M., Russell, B., Freeman, W.: Exploiting the sparse derivative prior for super-resolution and image demosaicing. In: *IEEE Workshop on Statistical and Computational Theories of Vision*. (2003)
3. Buades, A., Coll, B., Morel, J.: A non-local algorithm for image denoising. In: *CVPR*. Volume 2., IEEE (2005) 60–65
4. Roth, S., Black, M.: Fields of experts: A framework for learning image priors. In: *CVPR*. Volume 2., IEEE (2005) 860–867
5. Elad, M., Aharon, M.: Image denoising via sparse and redundant representations over learned dictionaries. *IEEE TIP* **15** (2006) 3736–3745
6. Fattal, R.: Image upsampling via imposed edge statistics. In: *ACM SIGGRAPH 2007*, ACM (2007) 95–es
7. Levin, A.: Blind motion deblurring using image statistics. *NIPS* **19** (2007) 841
8. Levin, A., Fergus, R., Durand, F., Freeman, W.: Image and depth from a conventional camera with a coded aperture. In: *ACM Transactions on Graphics (TOG)*. Volume 26., ACM (2007) 70
9. Roth, S., Black, M.: Steerable random fields. In: *ICCV*, IEEE (2007) 1–8
10. Weiss, Y., Freeman, W.: What makes a good model of natural images? In: *CVPR*, IEEE (2007) 1–8

11. Sun, J., Xu, Z., Shum, H., Xi'an, P.: Image super-resolution using gradient profile prior. In: CVPR, IEEE (2008) 1–8
12. Krishnan, D., Fergus, R.: Fast image deconvolution using hyper-laplacian priors. NIPS (2009)
13. Zoran, D., Weiss, Y.: From learning models of natural image patches to whole image restoration. In: ICCV, IEEE (2011) 479–486
14. Cho, T., Joshi, N., Zitnick, C., Kang, S., Szeliski, R., Freeman, W.: A content-aware image prior. In: CVPR, IEEE (2010) 169–176
15. Yang, J., Wright, J., Huang, T., Ma, Y.: Image super-resolution as sparse representation of raw image patches. In: CVPR, IEEE (2008) 1–8
16. Mairal, J., Bach, F., Ponce, J., Sapiro, G., Zisserman, A.: Non-local sparse models for image restoration. In: ICCV, IEEE (2009) 2272–2279
17. Dong, W., Zhang, L., Shi, G.: Centralized sparse representation for image restoration. In: ICCV, IEEE (2011) 1–8
18. Schmidt, U., Gao, Q., Roth, S.: A generative perspective on mrfs in low-level vision. In: CVPR, IEEE (2010) 1751–1758
19. Dabov, K., Foi, A., Katkovnik, V., Egiazarian, K.: Image denoising by sparse 3-d transform-domain collaborative filtering. IEEE TIP **16** (2007) 2080–2095
20. Zontak, M., Irani, M.: Internal statistics of a single natural image. In: CVPR, IEEE (2011) 977–984
21. Ji, H., Liu, C., Shen, Z., Xu, Y.: Robust video denoising using low rank matrix completion. In: CVPR, IEEE (2010) 1791–1798
22. Cai, J., Candès, E., Shen, Z.: A singular value thresholding algorithm for matrix completion. Siam Journal of Optimization **20** (2010) 1956–1982
23. Schaeffer, H., Osher, S.: A low patch-rank interpretation of texture. CAM Report (2011)
24. Sun, J., Tappen, M.: Learning non-local range markov random field for image restoration. In: CVPR, IEEE (2011) 2745–2752
25. Maire, M., Arbeláez, P., Fowlkes, C., Malik, J.: Using contours to detect and localize junctions in natural images. (2008)
26. Boykov, Y., Kolmogorov, V.: An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. IEEE TIP **26** (2004) 1124–1137
27. Mohan, K., Fazel, M.: Reweighted nuclear norm minimization with application to system identification. In: ACC, 2010, IEEE (2010) 2953–2959
28. Dai, S., Han, M., Xu, W., Wu, Y., Gong, Y.: Soft edge smoothness prior for alpha channel super resolution. In: CVPR, IEEE (2007) 1–8
29. Beck, A., Teboulle, M.: A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM Journal on Imaging Sciences **2** (2009) 183–202
30. Foi, A., Katkovnik, V., Egiazarian, K.: Pointwise shape-adaptive dct for high-quality denoising and deblocking of grayscale and color images. IEEE TIP **16** (2007) 1395–1411
31. Levin, A., Weiss, Y., Durand, F., Freeman, W.: Understanding and evaluating blind deconvolution algorithms. In: CVPR, IEEE (2009) 1964–1971