A Structure-Preserving Image Restoration Method With High-level Ensemble Constraints

Rui Chen, Huizhu Jia*, Xiaodong Xie, Gao Wen
National Engineering Laboratory for Video Technology, Peking University, Beijing, China

***Abstract*—In this paper, we present a new image restoration framework based on two high-level regularizations that can predict and preserve the better informative structures in the image. The sparse representation of a blurred image is first obtained to globally encode the salient structures by applying a group of coupled framelet filters. Then a physical meaning regularizer is derived to estimate the point spread function based on the frequency response characteristics of the image. Moreover, based on the operator of structure tensor, a novel nonlocal total variation as the regularizer is established to measure the image variation and non-local self-similarity. Finally, these two high-level regularizers are integrated into an objective function to constrain the ill-posedness. Compared with the state-of-the-art restoration methods, our algorithm can not only suppress strong noises effectively but also recover the sharp structures from the severe and complex blurred images.**

***Index Terms*—Image restoration, structure tensor, framelet filter, point spread function, nonlocal regularization**

I. INTRODUCTION

In many optical imaging systems, the observed images often suffer blurring due to various degradations such as defocusing, relative motion and atmospheric turbulence. Therefore, the restoration techniques are required in many image processing application domains such as microscopy, remote sensing and astronomy. Generally, the objective of image restoration is to obtain the best estimation of original clear image with known degraded versions when the point spread function (PSF) is impossible to be accurately measured. This ill-posed inverse problem becomes difficult due to less knowledge about response of the imaging system. Hence, to produce unique and stable solution, the prior information is used to constrain the latent image and PSF estimation.

Many image restoration algorithms have been proposed to solve this ill-posed inverse problem. In these schemes, the adopted regularizations are usually expressed by exploiting the smoothness, continuity and sparsity properties of natural images. Specifically, the popular regularization strategies for the restoration tasks include sparsity priors [1], total variation (TV) and its variations [2], probabilistic distributions on the image features [3-4], the wavelet tight-frame transforms [5]. Recent advancements on restoration performance have been achieved by exploring the better regularizers. Liu *et al.* [6] derived a convolution operator as the effective regularizer for PSF estimation based on the changes of frequency spectrum. Zhou *et al.* [7] used a Dirichlet approximation to estimate the blur kernel and remove the noise. However, these current methods may fail to successfully restore input image existing in the severe and complex degradations due to limited prior information used in estimation process.

In this paper, we have addressed image restoration problem by developing two powerful regularizers to recover both sharp edges and fine texture details. Natural image exhibits diverse structural content in different frequency bands that are tightly coupled. Hence, the global structural features are extracted and represented by using a group of framelet filters. All types of blur will attenuate mid- and high-frequency components in the source image. Then we devise a global frequency-reponse (GFR) regularizer to constrain the PSF solution by exploring global structural prior in the multi-scale domain. Moreover, a nonlocal TV based on structure tensor (NLTV-ST) is extended to measure both local variations and nonlocal similarity in the image. These two ensemble regularizers are integrated into an objective function. This convex optimization problem can be solved by using an alternating direction method of multipliers (ADMM) and the well-posed solutions can be found.

The rest of the organization of this paper is as follows. Section Ⅱ describes the formulation of the proposed high-level regularizers. Section Ⅲ introduces restoration procedure in details. The experimental results are given in Section Ⅳ. Section Ⅴ concludes this paper.

II. HIGH-LEVEL REGULARIZERS

Assuming that the imaging system is linear and spatial-invariant, the degraded formation process of a blurred image $\mathbf{B}$ can be usually modeled as follows:

$$\mathbf{B}=\mathbf{I}*\mathbf{K}+\mathbf{N} \tag{1}$$

where the symbol * denotes the convolution operator. $\mathbf{I}$ is the clear image, $\mathbf{K}$ is the PSF, and $\mathbf{N}$ is the white Gaussian noise. For the joint estimation of $\mathbf{I}$ and $\mathbf{K}$, we can formulate this problem by minimizing the following objective function:

$$\mathrm{E}(\mathbf{K},\mathbf{I})=\left\|\mathbf{B}-\mathbf{I}*\mathbf{K}\right\|_2^2+\alpha\,\mathrm{E}_{GFR}(\mathbf{K})+\beta\,\mathrm{E}_{NLTVST}(\mathbf{I}) \tag{2}$$

where $\mathrm{E}_{GFR}(\mathbf{K})$ and $\mathrm{E}_{NLTVST}(\mathbf{I})$ are the regularizers related with the different prior information. Two weight parameters $\alpha \geq 0$ and $\beta \geq 0$ are to control the constrained strength, respectively.

A. Framelet-based Filters

To characterize the blurring process and overcome the sensitivity of image noises, the geometric tight filters with nice properties are used for image decomposition. A set of filters $\{h_0,\ldots,h_3\}$ are first obtained from the tensor product of tight framelet filters. Then four

*The corresponding author. E-mail: hzjia@pku.edu.cn .

new filters are designed as following forms:

$$h_0 = \frac{1}{16}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \qquad h_1 = \frac{\sqrt{2}}{16}\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & -1 \end{bmatrix}$$

$$h_2 = \frac{\sqrt{2}}{16}\begin{bmatrix} 0 & 1 & 1 \\ -1 & 0 & 1 \\ -1 & -1 & 0 \end{bmatrix} \qquad h_3 = \frac{1}{16}\begin{bmatrix} -1 & 2 & -1 \\ -2 & 4 & -2 \\ -1 & 2 & -1 \end{bmatrix}$$

where h_0 and h_3 act as the weighted averaging operators. h_1 and h_2 are Sobel operators in $\pi/4$ and $-\pi/4$ directions, respectively.

This group of framelet filters can be further explored to handle complex noises and preserve the directional structural information. The matrix representation of the filter h_k is denoted by $\mathbf{W}_k$. Then a matrix $\mathbf{W}_H$ is formed by stacking the filters together

$$\mathbf{W}_H = [\mathbf{W}_1^T, \mathbf{W}_2^T, \mathbf{W}_3^T]^T \tag{3}$$

With these notations, the sparse approximation S_Γ of an image f can be obtained by

$$S_\Gamma(f) = \mathbf{W}_0^T \mathbf{W}_0 f + \mathbf{W}_H^T \mathbf{T}(\mathbf{W}_H f, \Gamma) \tag{4}$$

where the matrix $\mathbf{T}$ is the thresholding operator. Γ is a diagonal matrix determined by the orthogonal transform [5]. The constructed operator S_Γ can extract global structure features of the image.

B. The GFR Regularizer

To correct the errors of PSF estimation, the priors can accurately be modeled based on the real properties. Accordingly, the established regularizer containing the physical information can guarantee the meaningful solution. For describing the blurring process, the eigenvalue σ_i and the corresponding eigenvector κ_i of $S_\Gamma(\mathbf{B})$ is defined by the following convolution process:

$$\left\| S_\Gamma(\mathbf{B}) * \kappa_i(S_\Gamma(\mathbf{B})) \right\|_F = \sigma_i(S_\Gamma(\mathbf{B})), \quad \forall i = 1, \ldots, \mathrm{r} \tag{5}$$

where the values of κ_i and σ_i can be computed from the given matrix by Singluar Value Decomposition (SVD).

These convolution eigenvalues are linked to the multi-level structural components of the image. Based on the fact that blurring can decrease dramatically the power spectrum of an image, the eigenvalues essentially correspond to the image frequencies and will significantly decrease. After extracting the global structural features, all the convolution eigenvalues of $S_\Gamma(\mathbf{B})$ are relatively smaller than those of original image $\mathbf{I}$. This property is given by

$$\sigma_i(S_\Gamma(\mathbf{B})) \le \sigma_i(\mathbf{I}), \quad \forall i = 1, \ldots, \mathrm{r} \tag{6}$$

According to this discriminative blurred property, a convex regularizer can be derived. The regularizer based on frequency response is designed to reach the minimal value at the desired $\mathbf{K}$. Hence, the eigenvalues of $S_\Gamma(\mathbf{B})$ are smaller than those of image $\mathbf{B}$. According to this matrix analysis, the mathematical relationship between $S_\Gamma(\mathbf{B})$ and $\mathbf{I}$ can be reliably deduced. Because the frequency characteristics of a sharp image $\mathbf{I}$ is high-pass, the maximal eigenvalue $\sigma_{\max}(\mathbf{K} * S_\Gamma(\mathbf{B}))$ can effectively reflect the real property of PSF. The desired PSF can be recovered by utilizing this eigenvalue. Then $\mathrm{E}_{GFR}(\mathbf{K})$ is defined by

$$\mathrm{E}_{GFR}(\mathbf{K}) = \left| \sigma_{\max}(\mathbf{K} * S_\Gamma(\mathbf{B})) \right| \tag{7}$$

According to the fact that global features can achieve the smaller estimate errors, this regularizer puts a strong prior to guide to the accurate PSF. Although this constraint will lead to a sharp image from its noisy and blurred version, the restoration result is easy to contain the various unnatural artifacts. For performing reliable deconvolution, an additional regularized term need jointly be used.

C. The NLTV-ST Regularizer

To decrease the undesirable artifacts and further improve the restoration quality, the non-local structural self-similarity property is exploited as the regularizer [8]. Since the gradient magnitude is employed to penalize the image vairation in completely localized way, it is not very informative of the geometric image structures. To deal with these limitations of TV, the structure tensor operator is used to measure the image variation in a local neighborhood of every point. It summarizes the dominant direction of the gradient in a specified neighborhood centered at the point $\mathbf{q}$ of the image f and its formulation is defined by

$$\mathbf{T}_\mathbf{q} = \begin{bmatrix} f_x^2 & f_x f_y \\ f_x f_y & f_y^2 \end{bmatrix} \tag{8}$$

where f_x and f_y denote image derivatives along the x and y directions, respectively. The eigenvectors $v_{1,2}$ of structure tensor are computed in the directions of the lowest and highest local gradient, respectively. Correspondingly, the eigenvalues $\lambda_{1,2}$ provide the strength of the estimated edges in two orthogonal directions. With respect to $\lambda_2 > \lambda_1$, three cases are as follow:

i) If $\lambda_1 \approx \lambda_2 \approx 0$, the region is homogenous;

ii) If $\lambda_1 \approx 0$, $\lambda_2 \gg 0$, it has strong local orientation and the point $\mathbf{q}$ is located close to the edge;

iii) If $\lambda_1 \gg 0$, $\lambda_2 \gg 0$, the point $\mathbf{q}$ is close to the corner;

Based on the above, it appears that the eigenvalues of the structure tensor are more informative of the geometric structure of the image than the local gradient magnitude. To more accurately estimate the structure tensor, nonlocal mean is introduced to assign each pixel a weight. The non-negative weights are computed not only based on their relative distance of two points bus also the similarity of their values. The weighting function is defined by

$$\mathrm{W}(\mathbf{p}, \mathbf{q}) = \gamma \exp\left(-\frac{\left\| u(\mathbf{p}) - u(\mathbf{q}) \right\|_2^2}{h^2}\right) \tag{9}$$

where h is a scaling factor, and u is a patch centered at one pixel, whose size is restricted to $d \times d$ in the image $\mathbf{I}$. The parameter γ

can control the weighted strength, which is selected according to the preserving level of local structures.

Motivated by the nonlocal measure of image variation, we can construct the nonlocal version of the structure tensor. The edge content of the image is better encoded in the eigenvalues of nonlocal structure tensor and then this can lead to derive the NLTV-ST regularizer. Compared to TV regularization that preserves the edges but transforms the piecewise-smooth regions into piecewise-constant regions, the second-order regularizer can restore the local finer details and retain convexity property [9]. Here, the proposed NLTV-ST regularizer can be considered as second-order extensions of TV [10] which potentially provide better measurement of the image variation. Finally, $\mathrm{E}_{NLTVST}(\mathbf{I})$ is expressed as

$$\mathrm{E}_{NLTVST}(\mathbf{I}) = \sum_{i \in I} \sqrt{|\lambda_i^+| + |\lambda_i^-|} \tag{10}$$

where λ_i^+ and λ_i^- are the largest and smallest eigenvalues of matrix $\mathbf{W}u(i)$ at each position i, respectively. $\mathbf{W}u(i)$ is the weighted matrix of the structure tensor.

III. Restoration Method

To reliably find the global minimum solutions, the overall energy function $\mathrm{E}(\mathbf{K},\mathbf{I})$ in equation (2) can be solved via the alternating minimization scheme. In iteration processing, the sharp image and PSF are jointly estimated to ensure the meaningful solutions. Since the proposed regularizers are convex, the initial solution can be simply chosen the observed blurry image $\mathbf{B}$ as the initial condition for $\mathbf{I}$. Then the blind restoration is carried out by iterating two procedures until convergence.

While fixing the variable $\mathbf{I}$, the PSF $\mathbf{K}$ is updated by solving the following function:

$$\min_{\mathbf{K}} \left\| v(\mathbf{B}) - c(\mathbf{I})v(\mathbf{K}) \right\|_2^2 + \alpha \left| \sigma_{\max}(\mathbf{K} * S_\Gamma(\mathbf{B})) \right| \tag{11}$$

where the matrix $\mathbf{H}$ is determined by the eigenvectors of $S_\Gamma(\mathbf{B})$ and can be obtained by the algorithm of computing the Hessian matrix introduced in literature [11]. The operator $c(\mathbf{I})$ is used to convert the convolution operator into matrix multiplication. The vectorizaiton of a matrix is denoted by v.

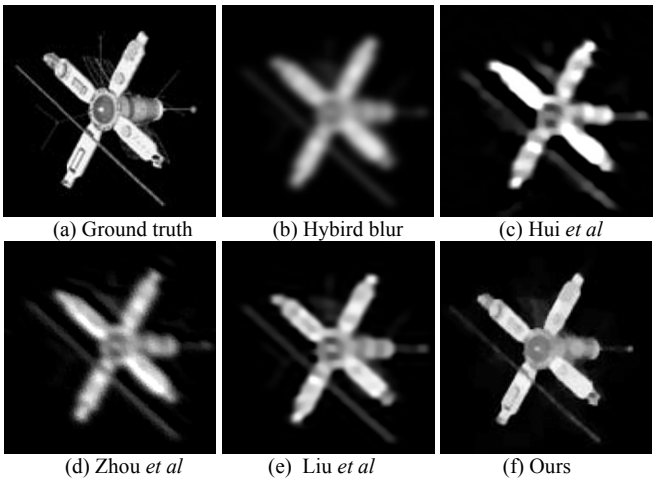
While fixing the PSF $\mathbf{K}$, the estimation of the original sharp image $\mathbf{I}$ can be solved by the proposed non-blind deconvolution method in [6] and is updated by

$$\min_{\mathbf{I}} \left\| v(\mathbf{B}) - c(\mathbf{I})v(\mathbf{K}) \right\|_2^2 + \beta \sum_{i=1}^{N} \sqrt{|\lambda_i^+| + |\lambda_i^-|} \tag{12}$$

where the index N represents the total number of pixels in matrix $\mathbf{I}$. The weight β is chosen by balancing image quality and stability.

IV. Experimental Results

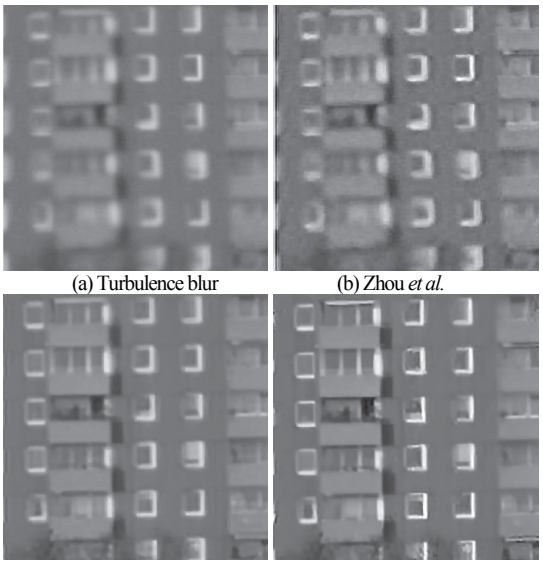
The performance of our method has been evaluated on both synthetic and real degraded images. Meanwhile, our results are compared against the ones obtained by Hui *et al.* [5], Zhou *et al.* [7] and Liu *et al.'s* methods [6]. In testing experiments, the parameters in our blind restoration framework need be first chosen. The initial kernel size of discrete PSF is set to 3×3 and the final size is usually chosen no more than 40×40. The parameter α can essentially control the sharpen degree of global structures in the original image. So this parameter is set to an appropriate value 280. The parameter β plays the role of reducing possible artifacts and suppressing the noises. This parameter is usually chosen from the range of 0.02 to 0.05. The sizes of patch are set to 5×5. γ is selected as 2.



(a) Ground truth (b) Hybird blur (c) Hui *et al*

(d) Zhou *et al* (e) Liu *et al* (f) Ours

Fig. 1. Restoration results of a degraded satellite image



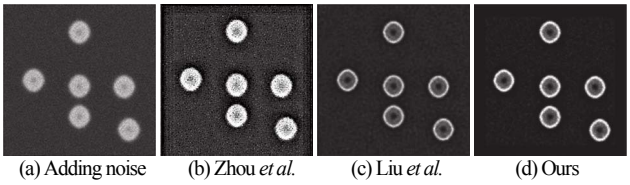
(a) Turbulence blur (b) Zhou *et al.*

(c) Liu *et al.* (d) Ours

Fig. 2. Restoration results of a building image

The first kind of degraded images are made to test the recovery capability of handling both motion and defocus blurs simultaneously. A satellite image shown in Fig. 1(a) is taken as the reference and Fig. 1(b) is blurred by a delta function with length 10 at horizontal direction and a disc function with radius 6. The second kind of degraded images are used to estimate the adaption of the Gaussian-type blur in the long-distance imaging system. A building image shown in Fig. 2(a) exhibits the distortion affected by significant atmospheric turbulence and is employed in this trial. The third type of degraded images are synthesized to observe the robustness of reconstructing the clear version from the severely blurred and

noisy input image. A blurred cell image shown in Fig. 3(a) is mainly corrupted by adding the Gaussian noise.



(a) Adding noise (b) Zhou *et al.* (c) Liu *et al.* (d) Ours

Fig. 3. Restoration results of a cell image

From the image shown in the Fig. 1, the proposed method can provide the best restoration results and faithfully recover more details distorted by complex blurs. Comparing with ground-truth, the results restored by other three methods have lost most useful structural information and resulted in the smooth regions. This is because that our regularizers contain strong and accurate structural priors and the priors adopted by other methods are difficult to handle the hybrid blurs or severe cases that human eyes might be unable to recognize the contents. The subjective results on turbulence distortion show that our result in Fig. 2(d) is most similar to the real scenes and appears more sharper structures. Moreover, the geometric deformations have not been effectively removed and the artifacts are visible in other methods. The superiority of our regularizer can be attributed to its ability to preserve the linear structures. For the high noise level added to the blurred image in Fig. 3, it is observed that our method can achieve more accurate reconstructions and suppress Gaussian noise simultaneously. These results also reflect that the NLTV-ST is stable with respect to the modeling errors. Due to its non-local nature, NLTV-ST proves more efficient in removing the noise.

TABLE I
PSNR(DB) RESULTS OF DIFFERENT METHODS

Sample	Hui	Zhou	Liu	Our
Satellite	29.85	30.66	30.54	**31.62**
Building	28.75	33.91	34.02	**36.55**
Cell	26.74	31.50	33.89	**34.56**

To quantitatively measure the improvement in the restored image quality, all the peak signal-to-noise ratio (PSNR) values of four methods are computed and compared in Table Ⅰ. According to the PSNR results, Our method is robust to restore corrupted image quite well with higher PSNR values which is consistent to the visual improvement.

In order to evaluate the converge of different methods versus to the iteration number, the normalized mean square error (NMSE) is used [4]. The best restored satellite image is selected with the lowest NMSE when the parameters of the regularization terms change for four restoration methods. The stopping criterion for different methods is set to reaching a relative norm difference of 10^{-7}. As presented in Fig. 4, the propose method has the fastest convergence and can reach the highest level of precision. Note also that our method practically converges to the vicinity of the exact solution after only 60 iterations and the solution will reach the maximum precision within 200 iterations.

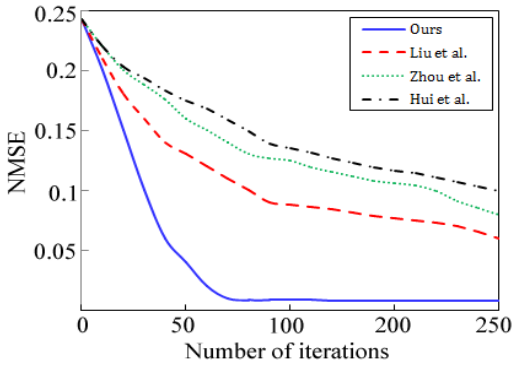


Fig. 4. Comparisons of the convergence rates

V. CONCLUSIONS

In this paper, we demonstrated a new image restoration approach by using the GFR and NLTV-ST constrains while preserving the multi-scale structures. By fully utilizing the spectrum characteristics after various blurs, we have derived the GFR regularizer to estimate the various types of PSFs. The NLTV-ST regularizer is employed to reconstruct image. Then two ensemble regularizations can render the sharpness and sparsity constraints to deal with the hybrid and severe blurs. Both the theoretical analysis and experimental results have shown the effectiveness and validity of the proposed regularizers. Moreover, our image blind restoration method can not only achieve high-quality deblurred results while preserving the informative structures but also remove the unpredictable noises and ringing artifacts.

REFERENCES

[1] R. Fergus, B. Singh, A. Hertzmann, S.T. Roweis, and W.T. Freeman, "Removing camera shake from a single photograph," *ACM Trans. Graphics*, vol. 25, no. 3, pp. 787–794, Jul. 2006.

[2] L. Xu, S.C. Zheng, and J.Y. Jia, "Unnatural L-0 sparse representation for natural image deblurring," in *IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2013, pp. 1107–1114.

[3] Z.B. Lu, H.Q. Li, and W.P. Li, "A Bayesian adaptive weighted total generalized variation model for image restoration," in *IEEE International Conference on Image Processing*. IEEE, 2015, pp. 492–496.

[4] V. Papyan, and M. Elad, "Multi-scale patch-based image restoration," *IEEE Trans. Image Process.*, vol. 25, no. 1, pp. 249–261, 2016.

[5] J. Hui, and K. Wang, "Robust image deblurring with an inaccurate blur kernel," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1624–1634, Apr. 2012.

[6] G.C. Liu, S.Y. Chang, and Y. Ma, "Blind image deblurring using spectral properties of convolution operators," *IEEE Trans. Image Process.*, vol. 23, no. 12, pp. 5047–5056, Dec. 2014.

[7] X. Zhou, J. Mateos, F.G. Zhou, R. Molina, and A.K. Katsaggelos, "Variational dirichlet blur kernel estimation," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 5127–5139, Dec. 2015.

[8] S. Lefkimmiatis, and S. Osher, "Nonlocal structure tensor functionals for image regularization," *IEEE Trans. Image Process.*, vol. 1, no. 1, pp. 16–29, Mar. 2015.

[9] X.J. Guo, and Y. Ma, "Generalized tensor total variation minimization for visual data recovery," in *IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2015, pp. 3603–3611.

[10] Y. Hu, G. Ongie, S. Ramani, and M. Jacob, "Generalized higher degree total variation regularization," *IEEE Trans. Image Process.*, vol. 23, no. 6, pp. 2423–2435, Jun. 2014.

[11] R. Chen, H. Z. Jia, X. D. Xie, and W. Gao, "Robust image/video super-resolution display," in *ACM Symposium on Virtual Reality Software and Technology*. ACM, 2015, pp. 1–4.