Research Article

Bayesian Image Restoration Using a Large-Scale Total Patch Variation Prior

Yang Chen,[1, 2, 3] Weimin Yu,[1] Yinsheng Li,[1] Zhou Yang,[1] Limin Luo,[1, 2, 3] and Wufan Chen[4]

[1] *Laboratory of Image Science and Technology, Southeast University, 211514 Nanjing, China*
[2] *Centre de Recherche en Information Biomedicale Sino-Francais (LIA CRIBs), 35000 Rennes, France*
[3] *Laboratoire Traitement du Signal et de lImage (LTSI) INSERM U642, Universite de Rennes I, Campus de Beaulieu, 263 avenue du General Leclerc, CS 74205, 35042 Rennes Cedex, France*
[4] *School of Biomedical Engineering, Southern Medical University, 510515 Guangzhou, China*

Correspondence should be addressed to
Limin Luo, luo.list@seu.edu.cn

Edge-preserving Bayesian restorations using nonquadratic priors are often inefficient in restoring continuous variations and tend to produce block artifacts around edges in ill-posed inverse image restorations. To overcome this, we have proposed a spatial adaptive (SA) prior with improved performance. However, this SA prior restoration suffers from high computational cost and the unguaranteed convergence problem. Concerning these issues, this paper proposes a Large-scale Total Patch Variation (LS-TPV) Prior model for Bayesian image restoration. In this model, the prior for each pixel is defined as a singleton conditional probability, which is in a mixture prior form of one patch similarity prior and one weight entropy prior. A joint MAP estimation is thus built to ensure the iteration monotonicity. The intensive calculation of patch distances is greatly alleviated by the parallelization of Compute Unified Device Architecture(CUDA). Experiments with both simulated and real data validate the good performance of the proposed restoration.

1. Introduction

Image restorations have wide applications in remote sensing, radar imaging, tomographic imaging, microscopic imaging, astronomic imaging, digital photography, and so forth [1–4]. For linear and shift invariant imaging systems, the transformation from f to g is well described by the following additive linear degradation model [3, 4]:

$$g = A * f + \varepsilon, \tag{1.1}$$

where g, f, and ε represent the degraded observed image, the original true image, and the corrupting white Gaussian noise with variance σ^2, respectively. A is the PSF (Point Spread

function) of the imaging system and $*$ is the linear convolution operator. Throughout this paper, we assume that the PSF A is known or could be numerically determined or estimated.

Bayesian or maximum a posteriori (MAP) approach, within Markov Random Fields (MRFs) framework, can provide a stable solution of the ill-posed inverse image restorations through incorporating a priori information about the geometrical properties of an image [3–8]. Via modelling the unknown parameters in the prior probability density functions, prior information can also be interpreted as the regularization metrics, which measures the extent to which the contextual constraint assumption is violated. With the incorporated prior information, Bayesian approach is able to distinguish solutions from less desirable ones by transforming the original ill-posed problem into a well-posed one. Generally, we can build the following posterior probability $P(f \mid g)$ for image restorations,

$$P(f \mid g) \propto P(g \mid f)P(f), \tag{1.2}$$

$$P(g \mid f) = \exp(L(g, f)) = \exp\left(-\frac{1}{2}\|g - A * f\|^2\right), \tag{1.3}$$

$$P(f) = Z^{-1} \times \exp(-\beta U(f)) = Z^{-1} \times \exp\left(-\beta \sum_j U_j(f)\right). \tag{1.4}$$

Here, $P(g \mid f)$ and $P(f)$ denote the likelihood distribution and the prior distribution, respectively. $L(g, f)$ and $U(f)$ denote the corresponding likelihood energy and the prior energy function. $U_j(f)$ is the notation of the energy function U evaluated at pixel index j. The partition function Z is a normalizing constant. β is the global parameter controlling the balance between the likelihood energy and the prior energy. We can build the posterior energy function in the logarithm form

$$\psi_\beta(f) = \log P(f \mid g) = L(g, f) - \beta U(f) = -\frac{1}{2}\|g - A * f\|^2 - \beta \sum_j U_j(f). \tag{1.5}$$

The restored or deconvolved image f can be obtained through the maximization of function $\psi_\beta(f)$.

The widely used quadratic membrane (QM) prior or the Tikhonov L2 regularization tend to smooth both noise and edge details and often lead to unfavorable oversmoothing in restorations [5]. On the other side, some edge-preserving Bayesian restoration methods have been proposed, which can be classified into three main categories: wavelet relevant regularization restorations, Bayesian restorations with nonquadratic prior energies, and Bayesian restorations with total variation (TV) prior. Wavelet relevant regularization restorations, with multiscale stochastic prior models, have been proposed in the research of edge-preserving restorations [8–11]. In [10], priors based on wavelet decompositions and heavy-tailed pdfs were incorporated into the EM restoration algorithm. In [11], Neelamani et al. proposed a Fourier-Wavelet hybrid restoration, which exploits the potentials of the sparse representation of noise for Fourier transform and the sparse representation of the coefficients in wavelet transform. All these wavelet regularization methods are based on the distribution modeling and thresholding of the decomposed wavelet coefficients [11]. Edge-preserving priors with nonquadratic energies were also proposed to preserve the edge details by using nonlinear inverse-proportional functions between edge existences and intensity

differences [5, 12, 13]. The weighting matrices in nonquadratic prior Bayesian restorations preserve the edges by turning off or suppressing smoothing at appropriate locations. Another impressive research direction in this area is the total variation (TV)-based image restorations [7, 14, 15]. This kind of approach can also be viewed as TV priors with *the half-quadratic regularization* prior energies. With no enforced global image continuity, TV prior restorations often demonstrate good property of edge preserving.

The nonquadratic priors preserve the edges by determining the pixel-wise regularizations based on the intensity-difference information within local fixed neighborhoods. The regularizations on TV priors also come from the information of the local intensity gradients. For all these nonquadratic and TV priors, the local information of intensity difference can by no means provide effective regularizations to discriminate noise from image textures in different scales. So these edge-preserving approaches often fail to cope well with some complicated restorations with high noise contaminations, and tend to produce block artifacts around the continuous edges in those situations [16–21].

Objects in the images are reasonably assumed to be a composition of edges and backgrounds with intensities changing coherently over different scales. Thus, for the traditional prior model, the intensity differences between individual pixels within local neighborhoods are often insufficient to characterize objects [19, 20]. This limit elicited our previous work in [21] which introduced a spatial adaptive (SA) prior restoration approach. This SA prior works by adaptively including the relevant neighboring pixels and excluding the negative irrelevant ones within a large neighborhood. Though showing expressive results, the iterative restoration in [21] updates the prior weight using the latest image estimate in a one step late (OSL) way, which leads to inconsistent posterior energy function and nonconvergent iterations. In addition, the approach in [21] is greatly limited by the high computational cost in calculating all the patch distances for all the pixel pairs in each prior neighborhood. In this paper, to overcome the first problem, we applied a large-scale total patch variation (LS-TPV) prior model, which is built through the specification of the conditional probability for each pixel [22–25]. The joint maximization estimation of this constrained entropy function will lead to convergent image/weight restorations. To overcome the second problem, we introduce the parallel computation structure of CUDA to calculate the patch distances [26–28]. Furthermore, another benefit from the CUDA-accelerated parallelization is that it allows further performance enhancement for the LS-TPV prior restoration by using larger search neighborhoods.

In Section 2, a review of some previous prior models is illustrated, and after that we introduce the proposed LS-TPV prior model. In Section 3, we give a joint estimation algorithm for the proposed restoration. In Section 4, we perform comparative experiments with both simulated and real data. Relevant visual/quantitative results are also presented. Conclusion and discussion are given in Section 5.

2. Prior Model

Conventionally, the value of $U_j(f)$ is commonly computed through a weighted sum of potential functions v of the differences between the pixels in the neighborhood N_j:

$$U_j(f) = \sum_{b \in N_j} w_{bj} v(f_b - f_j).$$

(2.1)

Generally, different choices of the potential function v lead to different priors. For the simple space-invariant QM prior, the potential function has the form $v(t) = t^2/2$. Some

Edge-preserving nonquadratic priors could be chosen by adopting a nonquadratic potential function v, such as the Huber potential function

$$v(t) = \begin{cases} \dfrac{t^2}{2}, & |t| \leq \gamma, \\ \gamma|t| - \dfrac{\gamma^2}{2}, & |t| > \gamma, \end{cases} \tag{2.2}$$

where γ is the threshold parameter [5, 13]. Such edge-preserving nonquadratic priors preserve structure information by choosing nonquadratic potential functions that increase less as the differences between the adjacent pixels become bigger. Weight w_{bj} is a positive value that denotes the interaction degree between pixel b and pixel j. In traditional prior models, it is simply considered to be inversely proportional to the distance between pixels b and j. So on a square lattice of image f, in which the 2D positions of pixels b and j ($b \neq j$) are, respectively, (b_x, b_y) and (j_x, j_y), w_{bj} is usually calculated by the geometric distance $1/\sqrt{(b_x - j_x)^2 + (b_y - j_y)^2}$.

Restorations using total variation (TV) prior take the prior energy as the following expression:

$$U_{\text{TV}}(f) = \sum_j \sqrt{\left(\Delta_j^h f\right)^2 + \left(\Delta_j^v f\right)^2}, \tag{2.3}$$

where Δ_j^h and Δ_j^v are linear operators corresponding to the first-order horizontal and vertical intensity differences at pixel j, respectively [7, 14, 15]. Since both the smooth and sharp edges tend to have similar prior energies, this TV prior energy given by (2.3) does not penalize weak discontinuities and often favors images with bounded variations.

Traditional quadratic priors, nonquadratic priors and TV prior, only provide local intensity information for Bayesian restorations. And the edge-preserving nonquadratic and TV priors, though being able to preserve edge information, tend to produce unfavorable block artifacts or false edges in high-noise situations.

To improve restoration quality, in [21] we proposed a spatial adaptive (SA) prior model, in which a spatial adaptive prior neighborhood is implemented by setting the weight w_{bj} to an 1-0 binary function to classify the pixels in neighborhood N into the pixels relevant to the center pixel j and those not. The building of the SA prior can be formalized as follows:

$$U_{\text{SA}}(f) = \sum_j U_j(f) = \sum_j \sum_{b \in N_j} \left(\frac{w_{bj}(f_b - f_j)^2}{|Nr_j|} \right), \tag{2.4}$$

$$w_{bj} = \begin{cases} 1, & |n_b(f) - n_j(f)|^2 \leq \delta \\ 0, & |n_b(f) - n_j(f)|^2 > \delta \end{cases}, \tag{2.5}$$

$$|n_b(f) - n_j(f)|^2 = \sum_l \left(f_{l \in n_b} - f_{l \in n_j} \right)^2. \tag{2.6}$$

Here, U_{SA} is the energy function for the SA prior. w_{bj}, which represents the classification of the neighboring pixels in the search neighborhood N_j, can be computed via (2.5) and (2.6).

$|Nr_j|$, the number of the neighboring pixels with nonzero w_{bj} in the neighborhood N_j, is the normalization factor for the different sizes of non-1 neighboring pixels in each N_j. Parameter δ in (2.5) is the threshold parameter. The distance $|n_b(f) - n_j(f)|^2$ is determined by the distance measure (2.6) between the two translated comparing patches n_b and n_j, whose centers are, respectively, located at pixel b and pixel j.

Though proved to be effective in suppressing both noise and block artifacts, all the weights w in SA prior need to be heuristically inferred using current image estimate in the way of one step late (OSL), which often leads to inconsistent posterior energy function with no guaranteed convergence [8]. To overcome this problem, in this paper, we proposed an LS-TPV prior model, in which regularization takes effect through penalizing the weighted distances/variations between the patches surrounding each neighboring pixel pair. The weights w in each N are considered variables with a nonlinear dependence on image f. For the central pixel j and its neighboring pixel b in each prior neighborhood N_j, w_{bj} is considered the similarity probability between the two patches n_b and n_j surrounding pixels b and j, respectively [24]. An entropy prior of all the weights w_{bj} for each N_j is introduced into our model in the form of $-\sum_j \sum_{b \in N_j} w_{bj} \ln w_{bj}$. Here as the "Frame" model in [25], the constraint is $\sum_{b \in N_j} w_{bj} = 1$. Then, based on joint MAP theorem [12], we can obtain

$$P(f, w, \eta \mid g) \propto P(g \mid f)P(f \mid w, \eta)P(w, \eta). \tag{2.7}$$

Here, η is the introduced pixel-wise Lagrange-multiplier parameter. For each pixel j, with $\hat{f}$ denoting the current image estimate, the prior energy $U_{\text{LS-TPV}}(f_j, w, \eta \mid \hat{f})$ of the LS-TPV prior is in the combined form of a weighted patch similarity energy, an entropy energy $-\sum_j \sum_{b \in N_j} w_{bj} \ln w_{bj}$ and a normalization constraint $\sum_{b \in N_j} w_{bj} = 1$. We can thus obtain the following posterior energy function:

$$
\begin{aligned}
\psi_\beta &(f, w, \eta \mid g) \\
&= \log P(g \mid f) + \log P(f \mid w, \eta) + \log P(w, \eta) \\
&= L(g, f) - \beta \left(\sum_j U_{\text{LS-TPV}}\left(f_j, w, \eta \mid \hat{f}\right) \right) \\
&= -\frac{1}{2}\|g - A * f\|^2 - \beta \sum_j \left(\left(\sum_{b \in N_j} w_{bj} d_{bj} + h \sum_{b \in N_j} w_{bj} \ln w_{bj} \right) + \eta_j \left(\sum_{b \in N_j} w_{bj} - 1 \right) \right).
\end{aligned}
\tag{2.8}
$$

Here, as (1.4), β is the global parameter controlling the balance between the likelihood energy and the prior energy. The patch distance d_{bj} is calculated as $\|G_\alpha * n_b - G_\alpha * n_j\|^2$, in which G_α denotes the Gaussian convolution kernel with standard deviation α. η_j is the value of Lagrange-multiplier parameter at pixel j. The parameter h controls the maximum entropy constraint on w, which is routinely modulated with respect to the noise levels (large h is routinely set to suppress more noise for high-noise situations).

3. Restoration Algorithm

To obtain tractable maximization of the posterior probability using the proposed LS-TPV prior energy $U_{\text{LS-TPV}}$ in (2.8), we use the algorithm of iterative coordinate descent (ICD),

in which each element is sequentially updated by maximizing the log-style local conditional probability function [29]. In this way, the objective image f and the weight w are jointly estimated through alternative maximization of the $\psi_\beta(f, w, \eta \mid g)$ with respect to f and w.

(1) $\arg\max_w - U_{LS\text{-}TPV}(w, \eta \mid \widehat{f})$.

$\widehat{f}$ denotes the fixed current image estimation, and each weight w_{bj} for each pixel j is updated by solving the following singleton optimization problem:

$$\arg\max_{w_{bj}} \left(-\sum_{b \in N_j} w_{bj} \|G_\alpha * n_b - G_\alpha * n_j\|^2 - h \sum_{b \in N_j} w_{bj} \ln w_{bj} - \eta_j \left(\sum_{b \in N_j} w_{bj} - 1 \right) \right). \tag{3.1}$$

With Lagrange-multiplier η_j introduced, by joint solving (3.1) with respect to each w_{bj} and η_j, we can then obtain w_{bj}:

$$w_{bj} = \frac{\exp\left(-\|G_\alpha * n_b - G_\alpha * n_j\|^2 / h\right)}{\sum_{b' \in N_j} \exp\left(-\|G_\alpha * n_{b'} - G_\alpha * n_j\|^2 / h\right)}, \tag{3.2}$$

where $\|G_\alpha * n_b - G_\alpha * n_j\|^2$ can be easily calculated using currently available $\widehat{f}$. Each w_{bj} is determined not only by the similarity metrics between patches n_b and n_j, but also by the sum of all the similarity metrics between patches with the central pixels located in N_j. So w_{bj} is different from w_{jb}.

(2) $\arg\max_f \psi_\beta(f, \widehat{w}, \widehat{\eta} \mid g)$.

Here $\widehat{w}$ denotes the fixed current weight estimate. Omitting the terms with no f, the $\psi_\beta(f, \widehat{w} \mid g)$ becomes

$$\psi_\beta(f, \widehat{w} \mid g) = L(g, f) - \beta \sum_j \left(\sum_{b \in N_j} \widehat{w}_{bj} \|G_\alpha * n_b - G_\alpha * n_j\|^2 \right). \tag{3.3}$$

To make a tractable maximization of (2.8), based on the theories of pseudolikelihood and iterated conditional mode (ICM) in [29], we can restore each f_j by solving the following factorized problem:

$$\arg\max_{f_j} \psi_\beta\left(f_j, \widehat{w}_{bj} \mid g, \widehat{f}\right) \Longrightarrow \arg\max_{f_j} L(f, g) - \beta \left(\sum_{b \in N_j} \widehat{w}_{bj} \|G_\alpha * n_b - G_\alpha * n_j\|^2 \right). \tag{3.4}$$

Considering the convexity for $\|g - A * f\|^2$ and $\|G_\alpha * n_b - G_\alpha * n_j\|^2$, we conclude that, given the fixed weight $\widehat{w}$, the Hessian matrix of each $\psi_\beta(f_j, \widehat{w}_{bj} \mid g, \widehat{f})$ in (3.4) is a negative definite. In this step of image updating, we applied the restoration algorithm in [4] to solve (3.4)

to obtain the image estimate of each pixel f_j. The $\psi_\beta(f, w, \eta \mid g)$ in (2.8) is separately convex with respect to each f and w. Denoting $\widehat{f}^n, \widehat{w}^n$ and $\widehat{f}^{n+1}, \widehat{w}^{n+1}$ as the nth and $(n+1)$th iterated estimates, we can obtain following relation: $\psi_\beta(\widehat{f}^{n+1}, \widehat{w}^{n+1} \mid g) \geq \psi_\beta(\widehat{f}^n, \widehat{w}^{n+1} \mid g) \geq \psi_\beta(\widehat{f}^n, \widehat{w}^n \mid g)$, which implies convergence to one local maximum and can be obtained for this joint updating strategy [30].

In each iteration in the proposed restoration, it is very expensive to perform the pixel-wise calculations of all the d_{bj} and w_{bj} for each pair of translated patches n in each N over the whole image region. In the experiments, to save computation cost, we do not extend the sizes of N over the whole image region, and set them to some appropriate sizes. We achieve significant acceleration by using parallel Compute Unified Device Architecture (CUDA) [26–28]. For each iteration in the restoration using the proposed LS-TPV prior model, all threads in this block grid structure execute simultaneously to perform all the involved pixel-wise operations. Furthermore, considering the symmetry property that distance d_{bj} equals to distance d_{jb}, we can further save one half computation cost by only calculating one of the two distances d_{bj} and d_{jb}.

4. Experiments

Both simulated and real data are used in the experiments. To provide comparative results, we also performed Wiener deconvolution, ForWaRD restoration, Bayesian restoration using TV prior, Huber prior, and the SA prior in [21]. The Wiener deconvolution was performed by using the degraded image to estimate the power spectrum in Fourier domain. The TWIST algorithm in [7] and Newton-Raphson algorithm in [12] are used in the restorations using TV prior and Huber prior, respectively. Restored images from above Wiener filter method are taken as the initial images for all the Bayesian restorations. In experiment, the parameters needed to be preset include the β for all Bayesian restorations, the threshold γ for Huber prior, the wavelet threshold, and decomposition level in the ForWaRD method, the δ for the SA prior, the h and the n and N settings for the proposed LS-TPV prior. In the simulation experiment, all the parameters are set based on SNR maximization criteria with SNR calculated as

$$\text{SNR} = 10 \log_{10} \left(\frac{\sum_j^M \left(F_j - \overline{F} \right)^2}{\sum_j^M \left(F_j - \widehat{f}_j \right)^2} \right). \tag{4.1}$$

Here M, $\widehat{f}$, F, and $\overline{F}$ denote the total pixel number in image, the restored image, the original true image, and the mean of the original true image, respectively. As to the real data experiment, which has no available true image for SNR calculations, all the parameters are set based on the tradeoff between edge preservation and noise suppression. N and n of sizes (11×11 N and 7×7 n) is used in our work. A full study of the effects of the sizes of N and n on the resulting restorations is performed in Section 4.5.

Generally, the solution to the restoration is considered optimized when the iteration goes stable. So for the Bayesian restorations using TV prior, Huber prior, and SA prior in which the energy functions are maximized with respect only to image, we stop the iterations when the current image estimates satisfy the condition $\psi_\beta(\widehat{f}^{n+1} \mid g) - \psi_\beta(\widehat{f}^n \mid g) < 0.995(\psi_\beta(\widehat{f}^n \mid g) - \psi_\beta(\widehat{f}^{n-1} \mid g))$. For the restorations using the proposed LS-TPV prior in which the energy functions are maximized with respect to both image and weights, we stop the iterations when the current image and weight estimates satisfy the condition

Table 1: Parameter settings in simulated data experiment.

TV prior	Huber prior	ForWaRD	SA prior	LS-TPV prior
$\beta = 0.0038$	$\beta = 0.026$, $\gamma = 0.35$	$Threshold = 3.0$, $Levels = 3$	$\beta = 0.015$, $\delta = 900$, $7 \times 7\ n$, $11 \times 11\ N$	$\beta = 0.075$, $\alpha = 0.5$, $h = 150$, $7 \times 7\ n$, $11 \times 11\ N$

$\psi_\beta(\widehat{f}^{n+1}, \widehat{w}^{n+1} \mid g) - \psi_\beta(\widehat{f}^n, \widehat{w}^n \mid g) < 0.995(\psi_\beta(\widehat{f}^n, \widehat{w}^n \mid g) - \psi_\beta(\widehat{f}^{n-1}, \widehat{w}^{n-1} \mid g))$. In practical experiments, to reach the above iteration conditions, it is found that 411, 433, 475, and 486 iterations are required for the Bayesian restorations using TV prior, Huber prior, SA prior, and the proposed LS-TPV prior, respectively.

4.1. Simulation Experiment

In this section, simulated experiments with 256×256 "Lena" image (intensity range $[-98\ 117]$) are performed. In the simulation, a uniform 9×9 PSF is assumed, and Gaussian noise with variance = 0.1663 is added. The PSF used in this experiment is normalized to 1. Figure 1(b) is the degraded image simulated by (1.1) using the PSF with the white Gaussian noise. The parameters for different restorations are listed in Table 1. All the restored images are illustrated in Figures 1(c)–1(h), respectively.

The restored images from TV prior restoration and Huber prior restoration are, respectively, shown in Figures 1(d) and 1(e). Figure 1(f) shows the restored image using the ForWaRD method. The result of Bayesian restoration using the SA prior is illustrated in Figure 1(g). For the Bayesian restoration using the proposed LS-TPV prior, the algorithm in Section 3 is used, and the corresponding restored image is shown in Figure 1(h).

Through the results, we find the ForWaRD method and all the Bayesian restorations (Figures 1(d)–1(h)) can overcome the ring effects in Wiener deconvolution (Figure 1(c)). The block artifacts (the left arrows), which are observed in the ForWaRD approaches and the restorations using Huber prior and TV prior, are effectively suppressed using the restorations using the SA prior and the proposed LS-TPV prior. The LS-TPV prior (Figure 1(h)) presents a further visual enhancement over the SA prior (Figure 1(g)) in block suppressing (the left arrows) and edge preserving (the right arrows). Figure 2 plots the profiles along one specified track (the red line in the hair region in "Lena" image in Figure 2) for different restorations, and we can see that the profile from the proposed restoration (the blue profile) has the closest match with the profile of the true "Lena" image (the red profile) of those from other restorations. Table 2 shows that, compared to other methods, the proposed LS-TPV prior restoration can lead to restored images with a higher SNR.

4.2. Real Data Experiment

In this section, a real blurred astronomical 256×256 "moon" image with intensity range $[-77\ 170]$ is obtained from "http://www.iceinspace.com.au/". Here we use a 5×5 normalized Gaussian PSF, which was tested to be effective. The parameters for different restoration approaches are listed in Table 3.

Figure 3(b) is the restored result from Wiener deconvolution. The restored images from TV prior restoration and Huber prior restoration are, respectively, illustrated in Figures 3(c) and 3(d). Figures 3(e) and 3(f) show the restored images from the ForWaRD method

(a) Original image

(b) Degraded image

(c) Wiener restoration

(d) TV prior restoration

(e) Huber prior restoration

(f) ForWaRD restoration [11]

(g) SA prior restoration [21]

(h) LS-TPV prior restoration

Figure 1: (a) Original "Lena" image. (b) Degraded "Lena" image (SNR = 6.65) with 9×9 uniform blur and additive white Gaussian noise (variance = 0.1663). (c) Wiener deconvolution. (d) TV prior Bayesian restoration. (e) Nonquadratic Huber prior Bayesian restoration. (f) The ForWaRD restoration in [11]. (g) The SA prior Bayesian restoration proposed in [21]. (h) The proposed LS-TPV prior Bayesian restoration (The left arrows: block-suppressing illustration, the right arrows: edge-preserving illustration).
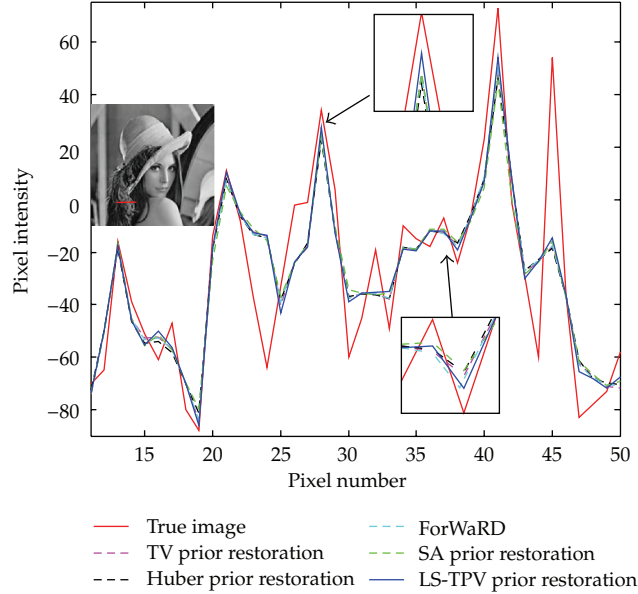
Figure 2: Profile illustrations of different restorations.

Table 2: Signal to noise ratio (SNR) of the observed degraded images and the restored images in the simulated experiment with "Lena" image.

Degraded	Wiener	TV prior	Huber Prior	ForWaRD	SA prior	LS-TPV prior
6.65	15.06	15.54	15.59	15.50	15.84	16.02

Table 3: Parameter settings in real data experiment.

TV prior	Huber Prior	ForWaRD	SA prior	LS-TPV prior
$\beta = 0.059$	$\beta = 0.045$, $\gamma = 0.45$	*Threshold* = 3.5, *Levels* = 3	$\beta = 0.032$, $\delta = 700$, $7 \times 7\ n$, $11 \times 11\ N$	$\beta = 0.095$, $\alpha = 0.5$, $h = 325$, $7 \times 7\ n$, $11 \times 11\ N$

and the Bayesian restoration using SA prior. As to the Bayesian restoration using the LS-TPV prior, the restored image is shown in **Figure** 3(g). Also, we can note that the ForWaRD method and all the Bayesian restorations (Figures 3(c)–3(g)) are free of the ring effects in Wiener deconvolution (**Figure** 3(b)). Also, the block artifacts observed in the restorations using Huber prior and TV prior are effectively suppressed in the restorations using SA prior and the LS-TPV prior (the lower red arrows in Figures 3(f) and 3(g)). The oversmoothing in the ForWaRD restoration can not be observed in the restorations using the SA prior and the proposed LS-TPV prior (the upper arrows in Figures 3(e), 3(f), and 3(g)). Compared to the restoration using the SA prior, the restoration using the LS-TPV prior also shows a better visual performance in preserving edges (the upper arrows in Figures 3(f) and 3(g)).

Table 4: CPU times (seconds) needed for different restorations for real data experiment.

Wiener filtering: *simu*: 58.16, *real*: 64.48	TV prior res: *simu*: 214.96, *real*: 228.92
Huber prior res: *simu*: 270.52, *real*: 292.36	ForWaRD res: *simu*: 2128.27, *real*: 2411.15
SA prior res: *simu*: 2128.27, *real*: 2382.43	LS-TPV prior res**1**: *simu*: 2885.61, *real*: 3058.07
LS-TPV prior res**2**: *simu*: 910.48, *real*: 954.48	LS-TPV prior res**3**: *simu*: 58.61, *real*: 64.45

4.3. Computation Cost Comparisons

Table 4 lists the recorded CPU times (in seconds) of different restorations for the experiments with simulated data and real data, respectively. Please note that the notations *simu* and *real* in Table 4 correspond to experiments with simulated data and real data, respectively. We can see that the restorations using SA prior and the LS-TPV prior are much computationally intensive than other methods. We define LS-TPV prior res**1**, LS-TPV prior res**2,** and LS-TPV prior res**3** as the original serial restorations, the restorations after optimization, and the restorations after both optimization and parallelization. Table 4 shows that, with respect to the original serial version, the proposed restorations after optimization are about 3 times faster, and the restorations after both optimization and parallelization are about 50 times faster. But we can not claim that the LS-TPV prior restoration needs less computation costs than other methods since that we did not design the parallelized versions for them.
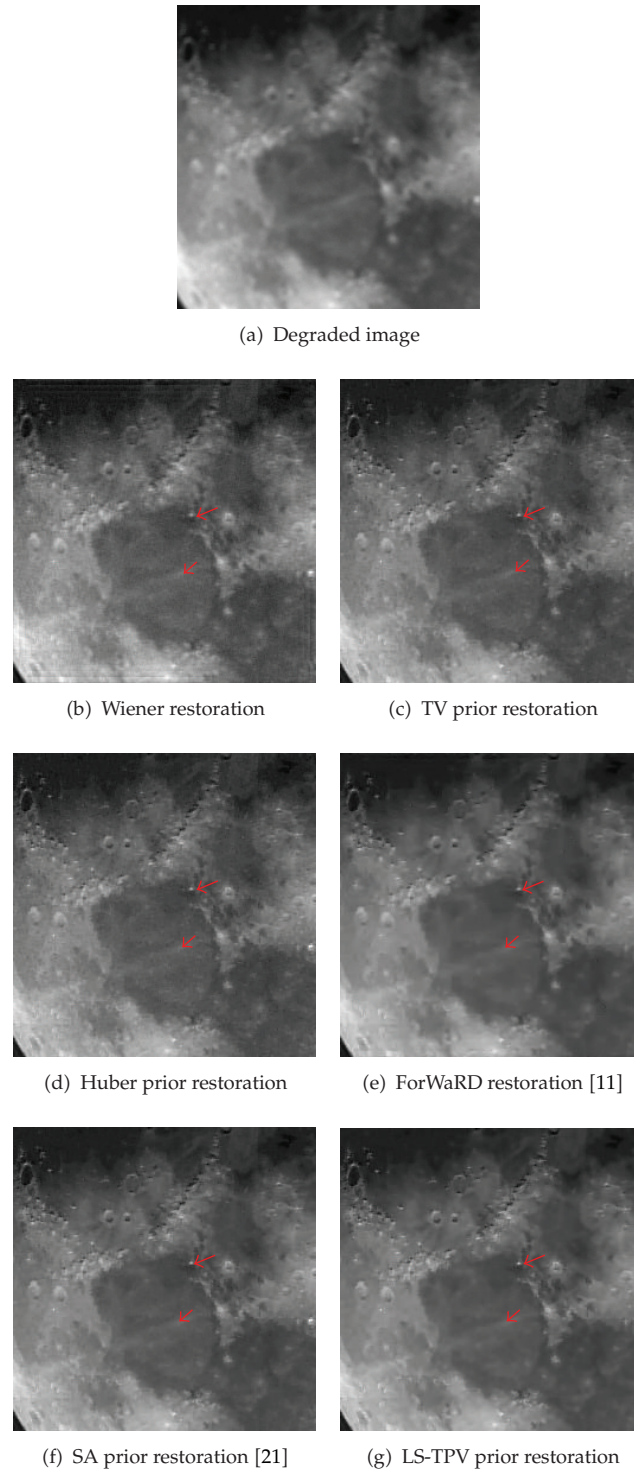
4.4. Algorithm Monotonicity

Here, to analyze the monotonicity of the joint image/weight algorithm proposed in Section 3, we calculated the posterior function energy $\psi_\beta(f, w, \eta \mid g)$ and SNR with respect to the iteration numbers for the proposed restorations. The total iteration number is set to 500. As illustrated in Figure 4, we can see that both the calculated posterior function energy and the SNR increase monotonically during the whole iteration.

4.5. Study on Neighborhood N and n of Different Sizes

To study the roles of the sizes of N and n in restorations, we perform restorations using the proposed prior with N and n of different sizes in above simulation experiment with "Lena" image. To characterize local image structures, we do not set n larger than 7×7. Considering the fact that the total pixel numbers are different for different comparing patches n (49 for 7×7 n, 9 for 3×3 n, and 1 for 1×1 n), parameter h should also change with the restorations using different patches n. The hyperparameter β and parameter h for the proposed prior are chosen by hand to produce the best stable images in terms of SNR maximization. Combinations of N and n with different sizes and the corresponding computed SNR and the recorded CPU time cost (with CUDA parallelization) are all listed in Table 5.

We can see that higher SNR can be obtained when enlarging N and n from 1×1 to 11×11 and 7×7. We also note in Table 5 that more CPU time is needed for the restorations using the proposed prior with larger N and n. We should set the sizes of N and n based on the tradeoff between performance and computation cost. We find in Table 5 that no significantly quantitative SNR differences are made between 11×11 N and larger 13×13 N. In fact, as the iteration proceeds, more global information beyond the 11×11 N will be incorporated into the regularization of the pixels.

(a) Degraded image

(b) Wiener restoration

(c) TV prior restoration

(d) Huber prior restoration

(e) ForWaRD restoration [11]

(f) SA prior restoration [21]

(g) LS-TPV prior restoration

Figure 3: (a) Original degraded "moon" image. (b) Wiener deconvolution. (c) TV prior Bayesian restoration. (d) Nonquadratic Huber prior Bayesian restoration. (e) The ForWaRD restoration in [11]. (f) The SA prior Bayesian restoration proposed in [21]. (g) The proposed LS-TPV prior Bayesian restoration (the lower arrows: block-suppressing illustration, the upper arrows: edge-preserving illustration).
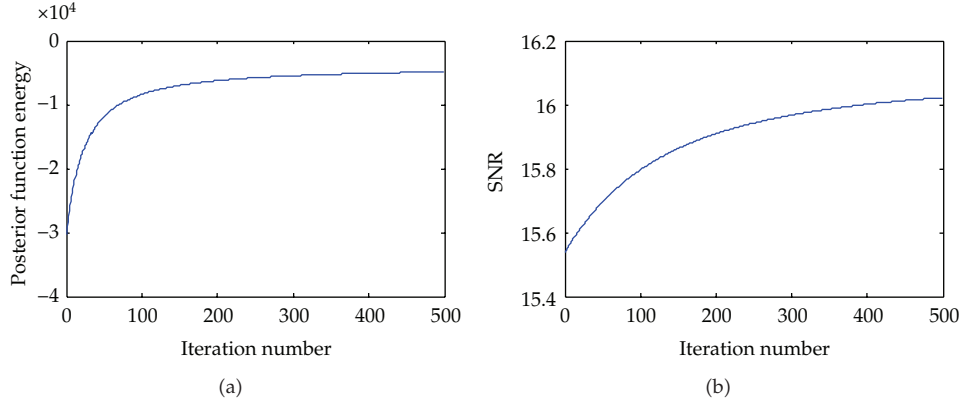
Figure 4: For the proposed restoration approach, the calculated SNR (a) and the posterior function energy (b) with respect to iteration number (1–500 iterations).

Table 5: SNR and CPU time cost (seconds) for the simulation restorations using the proposed method.

N and n	SNR	CPU times	N and n	SNR	CPU times	N and n	SNR	CPU times
(a) $N_{3\times3}n_{1\times1}$	15.48	17.98 sec	(b) $N_{3\times3}n_{3\times3}$	15.65	28.59 sec	(c) $N_{3\times3}n_{7\times7}$	15.79	53.69 sec
(d) $N_{7\times7}n_{1\times1}$	15.54	21.16 sec	(e) $N_{7\times7}n_{3\times3}$	15.70	31.37 sec	(f) $N_{7\times7}n_{7\times7}$	15.88	50.91 sec
(g) $N_{11\times11}n_{1\times1}$	15.55	24.59 sec	(h) $N_{11\times11}n_{3\times3}$	15.72	38.42 sec	(i) $N_{11\times11}n_{7\times7}$	16.02	58.61 sec
(j) $N_{13\times13}n_{1\times1}$	15.55	29.54 sec	(k) $N_{13\times13}n_{3\times3}$	15.75	77.67 sec	(l) $N_{13\times13}n_{7\times7}$	16.07	95.31 sec

5. Conclusions and Future Work Plan

In this paper on image restoration, we proposed an LS-TPV prior which penalizes the total distances between neighboring patches through a constrained mixture prior mode. A convergent joint image/weight updating algorithm, which estimates image and weights sequentially, is proposed to overcome the heuristical OSL weight determination for the SA prior restoration in [21]. We can see that, in addition to providing effective regularization, the proposed approach can lead to stable iteration with convergence guaranteed.

The application of the proposed prior model needs a pixel-wise computation of the distances between neighboring patches over a large region, which implies high computation cost of restorations. In this paper, remarkable advance in shortening computation time is achieved by optimizing patch distance computation and replacing the original pixel-wise calculation of patch distances by CUDA framework.

Further work includes applying the proposed approach in blind image deconvolutions, further justifying the proposed approach by using more image-quality measures based on human visual system, and further accelerating the computation by parallelize the PSF convolution in restorations.

Acknowledgments

References

[1] A. Srivastava, A. B. Lee, E. P. Simoncelli, and S. C. Zhu, "On advances in statistical modeling of natural images," *Journal of Mathematical Imaging and Vision*, vol. 18, no. 1, pp. 17–33, 2003.

[2] S. Tan and L. Jiao, "A unified iterative denoising algorithm based on natural image statistical models: derivation and examples," *Optics Express*, vol. 16, no. 2, pp. 975–992, 2008.

[3] R. Molina, "On the hierarchical Bayesian approach to image restoration: applications to astronomical images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, no. 11, pp. 1122–1128, 1994.

[4] G. Archer and D. M. Titterington, "On some Bayesian/regularization methods for image restoration," *IEEE Transactions on Image Processing*, vol. 4, no. 7, pp. 989–995, 1995.

[5] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, no. 6, pp. 721–741, 1984.

[6] N. P. Galatsanos and A. K. Katsaggelos, "Methods for choosing the regularization parameter and estimating the noise variance in image restoration and their relation," *IEEE Transactions on Image Processing*, vol. 1, no. 3, pp. 322–336, 1992.

[7] J. M. Bioucas-Dias and M. A. T. Figueiredo, "A new TwIST: two-step iterative shrinkage/thresholding algorithms for image restoration," *IEEE Transactions on Image Processing*, vol. 16, no. 12, pp. 2992–3004, 2007.

[8] Y. Wan and R. D. Nowak, "Bayesian multiscale approach to joint image restoration and edge detection," in *Proceedings of the International Society for Optical Engineering (SPIE '99)*, vol. 3813, pp. 73–84, 1999.

[9] M. A. T. Figueiredo and R. D. Nowak, "An EM algorithm for wavelet-based image restoration," *IEEE Transactions on Image Processing*, vol. 12, no. 8, pp. 906–916, 2003.

[10] J. M. Bioucas-Dias, "Bayesian wavelet-based image deconvolution: a GEM algorithm exploiting a class of heavy-tailed priors," *IEEE Transactions on Image Processing*, vol. 15, no. 4, pp. 937–951, 2006.

[11] R. Neelamani, H. Choi, and R. Baraniuk, "Forward: Fourier-wavelet regularized deconvolution for ill-conditioned systems," *IEEE Transactions on Signal Processing*, vol. 52, no. 2, pp. 418–433, 2004.

[12] C. A. Bouman and K. Sauer, "A generalized Gaussian image model for edge-preserving MAP estimation," *IEEE Transactions on Image Processing*, vol. 2, no. 3, pp. 296–310, 1993.

[13] D. F. Yu and J. A. Fessler, "Edge-preserving tomographic reconstruction with nonlocal regularization," *IEEE Transactions on Medical Imaging*, vol. 21, no. 2, pp. 159–173, 2002.

[14] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D*, vol. 60, no. 1–4, pp. 259–268, 1992.

[15] A. Chambolle, "An algorithm for total variation minimization and applications," *Journal of Mathematical Imaging and Vision*, vol. 20, no. 1–2, pp. 89–97, 2004.

[16] M. Mignotte, "A segmentation-based regularization term for image deconvolution," *IEEE Transactions on Image Processing*, vol. 15, no. 7, pp. 1973–1984, 2006.

[17] G. K. Chantas, N. P. Galatsanos, and A. C. Likas, "Bayesian restoration using a new nonstationary edge-preserving image prior," *IEEE Transactions on Image Processing*, vol. 15, no. 10, pp. 2987–2997, 2006.

[18] G. Chantas, N. Galatsanos, and A. Likas, "Bayesian image restoration based on variatonal inference and a product of student-t priors," in *Proceedings of the IEEE International Machine Learning for Signal Processing Workshop (MLSP '07)*, pp. 109–112, Thessaloniki, Greece, 2007.

[19] Y. Chen, J. Ma, Q. Feng, L. Luo, P. Shi, and W. Chen, "Nonlocal prior Bayesian tomographic reconstruction," *Journal of Mathematical Imaging and Vision*, vol. 30, no. 2, pp. 133–146, 2008.

[20] A. Buades, B. Coll, and J. M. Morel, "A nonlocal algorithm for image denoising," in *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. 60–65, San Diego, Calif, USA, 2005.

[21] W. Chen, Y. Chen, Y. Li, Y. Dong, L. Hao, and L. Luo, "Effective image testorations using a novel spatial adaptive prior," *EURASIP Journal on Advances in Signal Processing*, vol. 2010, Article ID 508089, 2010.

[22] S. Lyu, "An implicit Markov random field model for the multi-scale oriented representations of natural images," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1919–1925, 2009.

[23] G. Winkler, *Image Analysis, Random Fields and Markov Chain Monte Carlo Methods*, vol. 27 of *Applications of Mathematics*, Springer, Berlin, Germany, 2nd edition, 2003.

[24] G. Peyré, S. Bougleux, and L. Cohen, "Non-local regularization of inverse problems," in *Proceedings of the European Conference on Computer Vision*, pp. 57–68, 2008.

[25] S. C. Zhu, Y. Wu, and D. Mumford, "FRAME: filters, random field and maximum erntropy : towards a unified theory for texture modeling," *International Journal of Computer Vision*, vol. 27, no. 2, pp. 1–20, 1998.

[26] NVIDIA CUDA Programming Guide(Version 3.0), http://developer.download.nvidia.com/compute/cuda/3_2/toolkit/docs/CUDA_C_Programming_Guide.pdf.

[27] M. Fatica and W. Jeong, "Accelerating matlab with CUDA," in *Proceedings of the High Performance Embedded Computing*, vol. 2007, Lexington, Massachusetts, USA, September 2007.

[28] Y. Chen, Z. Zhou, W. F. Chen, X. D. Yin et al., "Improving Low-dose X-ray CT Images by Weighted Intensity Averaging over Large-scale Neighborhood," *European Journal of Radiology*, 2010. In press.

[29] C. A. Bouman and K. Sauer, "A unified approach to statistical tomography using coordinate descent optimization," *IEEE Transactions on Image Processing*, vol. 5, no. 3, pp. 480–492, 1996.

[30] W. Chen, M. Chen, and J. Zhou, "Adaptively regularized constrained total least-squares image restoration," *IEEE Transactions on Image Processing*, vol. 9, no. 4, pp. 588–596, 2000.

Advances in
Operations Research

Advances in
Decision Sciences

Mathematical Problems
in Engineering

Algebra

Journal of
Probability and Statistics

The Scientific
World Journal

International Journal of
Differential Equations

International Journal of
Combinatorics

Hindawi

Submit your manuscripts at
http://www.hindawi.com

Advances in
Mathematical Physics

Journal of
Complex Analysis

Journal of
Mathematics

International Journal of
Stochastic Analysis

Abstract and
Applied Analysis

Discrete Dynamics in
Nature and Society

International
Journal of
Mathematics and
Mathematical
Sciences

Journal of
Discrete Mathematics

Journal of
Function Spaces

Journal of
Applied Mathematics

Journal of
Optimization