Simultaneous Denoising and Illumination Correction via Local Data-Fidelity and Nonlocal Regularization

Jun Liu[1], Xue-cheng Tai[2], Haiyang Huang[1], and Zhongdan Huan[1]

[1] School of Mathematical Sciences
Laboratory of Mathematics and Complex Systems,
Beijing Normal University, Beijing 100875, P.R. China.
[2] School of Physical and Mathematical Sciences,
Nanyang Technological University, Singapore and
Department of Mathematics, University of Bergen, Norway,

Abstract. In this paper, we provide a new model for simultaneous denoising and illumination correction. A variational framework based on local maximum likelihood estimation (MLE) and a nonlocal regularization is proposed and studied. The proposed minimization problem can be efficiently solved by the augmented Lagrangian method coupled with a maximum expectation step. Experimental results show that our model can provide more homogeneous denoisng results compared to some earlier variational method. In addition, the new method also produces good results under both Gaussian and non-Gaussian noise such as Gaussian mixture, impulse noise and their mixtures.

1 Introduction

Image denoising is a fundamental technique of image processing. A large number of denoising methods have been proposed. It is common to assume that the noise is additive, i.e.

$$f(x) = u(x) + n(x),$$

where $f, u, n : \Omega \subseteq \mathbb{R}^2 \mapsto \mathbb{R}$ are the observed noisy image, true image and noise, respectively. Image denoising is to recover u for any given f and a priori knowledge of n. Variational method is one of the most efficient methods. It has now grown as a popular and widely used tool in image processing. Since the ROF model was proposed in [1], many variants based on total variation (TV) had been designed for different denoising tasks due to its good edges-preserving properties. Extending ROF, the authors in [2–5] have used L_1 norm or its linear combinations as the fidelity term to removing impulse noise. In order to better preserve some small structures such as textures, an efficient method called nonlocal mean was discussed in [6]. Motivated by the nonlocal mean and the graph theory, the nonlocal TV variational framework base on nonlocal operators was proposed in [7]. In [8], it was extended to nonlocal Mumford-Shah regularizers for image restoration. However, all these methods do not consider the varying illumination in the images. Moreover, it is also hard to treat cases that the intensity values are inhomogeneous.

Illumination correction or bias field correction is very important for real images. The artifacts caused by smooth, spatially varying illumination, although not usually a problem for visual inspection, can dramatically impede automated processing of the images. A widely accepted bias model, such as in MRI data, is the multiplicative bias field, which assumes that the observed signal f is equal to an uncorrupted signal u scaled by some bias β, i.e. $f = \beta u$. Then the application of a logarithmic transformation to the intensities allows the artifact to be modeled as an additive bias field $\ln f = \ln u + \ln \beta$. There are some works based on this logarithmic additive model for image segmentation such as [9–11] *etc.*.

Motivated by modeling the illumination bias with a multiplicative field in segmentation problem, in this paper, we propose an unified model for denoising and correcting illumination simultaneously with different types of noise include Gaussian noise, impulse noise and their mixtures. Our model is built on MLE and nonlocal regularization. To be different from the traditional regularized MLE, we construct a novel block-based adaptive data-fidelity term to handle inhomogeneous illumination and the noise. Besides, our approach do not need any additional constraints such as regularization on the bias function β to keep it smooth. Anther superiority of this model is that it can work well under different types noise like Gaussian, impulse noise, Gaussian noise plus impulse noise. The new model can be efficiently optimized by an extended augmented Lagrangian method (ALM) for nonlocal regularization according to the recently proposed ALM framework [13, 14] together with a maximum expectation process. These algorithms extend the Split-Bregman method of [17].

The rest of the paper is organized as follows: Section 2 gives our proposed model. Section 3 contains the optimization algorithms, while numerical experiments are presented in Section 4

2 The Proposed Model

2.1 Some Model Assumptions

In this paper, we consider the noise model with illumination bias

$$f(x) = \beta(x)(u(x) + n(x)), \tag{1}$$

where f is an observed noisy image, u stands for the ground truth image, n represents noise and β is a illumination bias function. In order to get a suitable denoising cost functional, we have the following assumptions:

- A1: the noise $n(x)$ at each location x is a realization of a random variable ξ with Gaussian mixture probability density function (PDF) $\sum_{k=1}^{K} \gamma_k p_k(z; c_k, \sigma_k^2)$. Here $p_k(z; c_k, \sigma_k^2)$ is the 1-D Gaussian PDF parameterized by mean c_k and variance σ_k^2, and γ_k is the mixture ratio which satisfies $\sum_{k=1}^{K} \gamma_k = 1$. This is an extended Gaussian noise model.

– A2: the bias function $\beta(x) > 0$, and β is smoothly varying. Motivated by [12], we use the following method to describe the smoothness of β: In a small neighborhood O_x centered at x, β satisfies $\beta(y) \approx \beta(x)$ when $y \in O_x$.

Now, we suppose the intensity value of the observed pixels at location x, namely $f(x)$, is a realization of a random variable η, then according to assumption A1 and model (1), we have

Proposition 1 *The PDF of η has the expression*

$$p_\eta(z) = \sum_{k=1}^{K} \frac{\gamma_k}{\sqrt{2\pi}\sigma_k \beta(x)} \exp\left(-\frac{[z - c_k\beta(x) - u(x)\beta(x)]^2}{2\sigma_k^2\beta^2(x)}\right).$$

In the next, we shall use this PDF to construct a local data-fidelity in terms of MLE and some model assumptions.

2.2 The Local Fidelity Term

Let us construct a new local data term according to the pixel density function in the section. Let $\Theta = \{\gamma_1, \cdots, \gamma_K, c_1, \cdots, c_K, \sigma_1^2, \cdots, \sigma_K^2, \beta\}$ is a parameter set. By independence assumption of $f(x)$, the PDF expression in the proposition 1 and a likelihood process, one can get the continuous functional in a neighborhood O_y centered at y

$$L_y(u, \Theta; f) = \int_{O_y} \ln \sum_{k=1}^{K} \frac{\gamma_k}{\sqrt{2\pi}\sigma_k \beta(x)} \exp\left(-\frac{[f(x) - c_k\beta(x) - u(x)\beta(x)]^2}{2\sigma_k^2\beta^2(x)}\right) \mathrm{d}x.$$

Note that $\beta(x)$ can be replaced by $\beta(y)$ when $x \in O_y$ in terms of assumption A2, thus L_y becomes

$$L_y(u, \Theta; f) = \int_{O_y} \ln \sum_{k=1}^{K} \frac{\gamma_k}{\sqrt{2\pi}\sigma_k \beta(y)} \exp\left(-\frac{[f(x) - c_k\beta(y) - u(x)\beta(y)]^2}{2\sigma_k^2\beta^2(y)}\right) \mathrm{d}x.$$

At this time, we get a local data fidelity term $D_y(u, \Theta) = -L_y(u, \Theta; f(x))$ in O_y. If we consider the different contributions to the fidelity D_y in terms of the distance from the neighborhood center, then we can assign some weights to different pixels. A common choice for this is the so-called Gaussian smoothness, and thus we get a cost functional

$$D_y(u, \Theta) = -\int_{O_y} G_\sigma(y-x) \ln \sum_{k=1}^{K} \frac{\gamma_k}{\sqrt{2\pi}\sigma_k \beta(y)} \exp\left(-\frac{[f(x) - c_k\beta(y) - u(x)\beta(y)]^2}{2\sigma_k^2\beta^2(y)}\right) \mathrm{d}x.$$

Here G_σ is a Gaussian kernel with a given standard deviation σ. Our objective is to recover all the degraded pixels. Thus we need to minimize all the local data fidelity. We shall use the cost functional

$$D(u, \Theta) = \int_{\Omega} D_y(u, \Theta)\mathrm{d}y = -\int_{\Omega}\int_{O_y} G_\sigma(y-x) \ln \sum_{k=1}^{K} \frac{\gamma_k}{\sqrt{2\pi}\sigma_k \beta(y)} \exp\left(-\frac{[\frac{f(x)}{\beta(y)} - c_k - u(x)]^2}{2\sigma_k^2}\right) \mathrm{d}x\mathrm{d}y.$$

Using properties of the Gaussian kernel, $G_\sigma(y-x) \approx 0$ when $x \notin O_y$ by choosing an appropriate σ, the neighborhood O_y in the second integration can be dropped. However, this new data-fidelity term is not easy to minimize due to the log-sum function. We use the conclusion of the following proposition [15,16] to overcome this difficult.

Proposition 2 *For all $\alpha_k(x) > 0$, let $\Delta = \{\phi(x) = (\phi_1(x), \cdots, \phi_k(x)) : \sum_{k=1}^K \phi_k(x) = 1, \phi_k(x) > 0\}$, then*

$$-\ln \sum_{k=1}^K \alpha_k(x) \exp(-\psi_k(x)) = \min_{\phi(x)\in\Delta} \left\{ \sum_{k=1}^K (\psi_k(x) - \ln \alpha_k(x))\phi_k(x) + \sum_{k=1}^K \phi_k(x) \log \phi_k(x) \right\}.$$

By applying Proposition 2 with $\alpha_k(x) = \dfrac{\gamma_k}{\sqrt{2\pi}\sigma_k \beta(y)}$, $\psi_k(x) = \dfrac{[\frac{f(x)}{\beta(y)} - c_k - u(x)]^2}{2\sigma_k^2}$, $D(u, \Theta)$ becomes

$$D(u,\Theta) = \int_\Omega \int_\Omega G_\sigma(y-x) \min_{\phi(x)\in\Delta} \{\textstyle\sum_{k=1}^K [\frac{(\frac{f(x)}{\beta(y)} - c_k - u(x))^2}{2\sigma_k^2} - \ln \gamma_k \\ + \ln(\sqrt{2\pi}\sigma_k \beta(y)) + \ln \phi_k(x)]\phi_k(x)\} \mathrm{d}x\mathrm{d}y.$$

Unlike the common methods to choose the negative log-likelihood as the data-fidelity term, we introduce a functional $E(u, \Theta, \phi)$ with an additional variable ϕ:

$$E(u,\Theta,\phi) = \\ \int_\Omega \int_\Omega G_\sigma(y-x) \sum_{k=1}^K \left[\frac{(\frac{f(x)}{\beta(y)} - c_k - u(x))^2}{2\sigma_k^2} - \ln \gamma_k + \ln(\sqrt{2\pi}\sigma_k \beta(y)) + \ln \phi_k(x) \right] \phi_k(x) \mathrm{d}x\mathrm{d}y,$$

and consider the the minimization problem

$$(u^*, \Theta^*, \quad \phi^*) = \operatorname*{arg\,min}_{u,\Theta,\phi\in\Delta} \ E(u,\Theta,\phi) \tag{2}$$

to be solved by the following alternative minimization procedure:

$$\begin{cases} \phi^{\nu+1} = \operatorname*{arg\,min}_{\phi\in\Delta} \ E(u^\nu, \Theta^\nu, \phi), \\ (u^{\nu+1}, \Theta^{\nu+1}) = \operatorname*{arg\,min}_{u,\Theta} \ E(u, \Theta, \phi^{\nu+1}). \end{cases} \tag{3}$$

Actually, the above iteration scheme can be interpreted as the well-known expectation-maximization (EM) algorithm. The updating of ϕ and Θ corresponding to the E-step and M-step, respectively. One can also prove that

Proposition 3 *The sequence u^ν, Θ^ν produced by iteration scheme (3) satisfies $D(u^{\nu+1}, \Theta^{\nu+1}) \leqslant D(u^\nu, \Theta^\nu)$.*

Thus we can take $E(u, \Theta, \phi)$ as the data-fidelity term. Compared to the model that directly uses $D(u, \Theta)$, we get some close-form solutions for the sub-problems when optimizing $E(u, \Theta, \phi)$.

2.3 Nonlocal TV

Nonlocal regularization could preserve repeated structures and textures and at the same time remove noise. The nonlocal denoising method was first proposed by Buades *etc.* [6]. In [7], Gilboa and Osher defined a variational framework based nonlocal operators. Let us review some definitions and notations on nonlocal TV regularization. Let $\Omega \subset \mathbb{R}^2$, $H_1 = L^2(\Omega)$, $H_2 = L^2(\Omega \times \Omega)$ and $\omega(x, y) \in H_2$ be a nonnegative symmetric weight function. The nonlocal gradient operator $\nabla_\omega : H_1 \mapsto H_2$ is defined as the vector of all partial derivatives at x such that:

$$(\nabla_\omega \circ u)(x) \mapsto \nabla_\omega u(x, y) \triangleq (u(y) - u(x))\sqrt{\omega(x, y)}.$$

The inner product in H_1 and H_2 is defined as

$$< u, v >_{H_1} = \int_\Omega u(x)v(x)\mathrm{d}x, \qquad < p, q >_{H_2} = \int_\Omega \int_\Omega p(x, y)q(x, y)\mathrm{d}y\mathrm{d}x.$$

Naturally, the isotropic L_1 and L_2 norms in H_2 is

$$||p||_1 = \int_\Omega \sqrt{\int_\Omega p(x, y)^2 \mathrm{d}y}\mathrm{d}x, \qquad ||p||_2 = \sqrt{\int_\Omega \int_\Omega p(x, y)^2 \mathrm{d}y\mathrm{d}x}.$$

The nonlocal divergence operator $\mathrm{div}_\omega : H_2 \mapsto H_1$ is given by the standard adjoint relation

$$< \nabla_\omega u, p >_{H_2} = - < \mathrm{div}_\omega p, u >_{H_1},$$

which leads to

$$\mathrm{div}_\omega p(x) = \int_\Omega (p(x, y) - p(y, x))\sqrt{\omega(x, y)}\mathrm{d}y.$$

Thus the nonlocal Laplacian operator $\Delta_\omega : H_1 \mapsto H_1$ is given by

$$\Delta_\omega u(x) = \mathrm{div}_\omega \nabla_\omega u(x) = 2\int_\Omega (u(y) - u(x))\omega(x, y)\mathrm{d}y.$$

With these notations, the nonlocal TV functional

$$R_\omega(u) = ||\nabla_\omega u||_1 = \int_\Omega \sqrt{\int_\Omega (u(x) - u(y))^2 \omega(x, y)\mathrm{d}y}\mathrm{d}x.$$

In this paper, we shall use the following weighting function [6]:

$$\omega^f(x, y) = \exp\{-\frac{\int_\Omega G_a(z)(f(x + z) - f(y + z))^2 \mathrm{d}z}{2h^2}\}. \tag{4}$$

2.4 The Proposed Cost Functional

The data-fidelity term $E(u, \Theta, \phi)$ together with the nonlocal TV norm yield the following new cost functional for simultaneous denoising and illumination correction :

$$J(u, \Theta, \phi) = E(u, \Theta, \phi) + \mu ||\nabla_\omega u||_1,$$

where $\mu > 0$ is a regularization parameter.

We need to impose some constraint condition on the parameters

$$\Theta = \{\gamma_1, \cdots, \gamma_K, c_1, \cdots, c_K, \sigma_1^2, \cdots, \sigma_K^2, \beta\}$$

and ϕ. For γ_k, we require $\sum_{k=1}^{K} \gamma_k = 1$ since it represents the mixture ratio. The $\phi_k(x)$ is actually a probability distribution of the pixel $f(x)$ contaminated by the noise comes from the k-th Gaussian distribution with mean c_k and variance σ_k. Thus, the constraint $\phi \in \Delta$ can guarantee this.

3 Algorithm: Augmented Lagrangian Method and EM

Operator splitting is an efficient method to solve L_1 minimization. In recent years, many efficient algorithms based on operator splitting have appeared, such as split Bregman method [17], augmented Lagrangian method (ALM) [13, 14], Douglas-Rachford splitting [18] and so on. These methods are all equivalent under certain conditions. In [13, 14], the authors only considered the local L_1 regularization, here we extend the split Bregman method [17] following the framework of Tai and Wu [13, 14]. The nonlocal TV in our model can be efficiently optimized with ALM.

In order to apply augmented Lagrangian method, the original minimization problem

$$(u^*, \Theta^*, \phi^*) = \operatorname*{arg\,min}_{u, \Theta, \phi \in \Delta} \ J(u, \Theta, \phi)$$

is reformulated as a constraint optimization minimization problem:

$$(u^*, d^*, \Theta^*, \phi^*) = \operatorname*{arg\,min}_{u, d, \Theta, \phi \in \Delta} \ E(u, \Theta, \phi) + \mu ||d||_1 \ \text{ s.t. } \ d = \nabla_\omega u. \tag{5}$$

The augmented Lagrangian functional for this constrained minimization problem is:

$$\mathcal{L}(u, d, \Theta, \phi, \lambda) = E(u, \Theta, \phi) + \mu ||d||_1 + < \lambda, (d - \nabla_\omega u) >_{H_2} + \frac{r}{2} ||d - \nabla_\omega u||_2^2,$$

where the Lagrangian multiplier $\lambda(x, y) \in H_2$, and $r > 0$ is a penalty parameter. It can be shown that one of the saddle points $(\hat{u}, \hat{d}, \hat{\Theta}, \hat{\phi}, \hat{\lambda})$ of $\mathcal{L}(u, d, \Theta, \phi, \lambda)$ is a solution of (5). We can search a saddle point by the following alternative algorithm:

$$\begin{cases} (u^{\nu+1}, d^{\nu+1}, \Theta^{\nu+1}, \phi^{\nu+1}) = \operatorname*{arg\,min}_{u, d, \Theta, \phi \in \Delta} \ \mathcal{L}(u, d, \Theta, \phi, \lambda^\nu), \\ \lambda^{\nu+1} = \lambda^\nu + r(d^{\nu+1} - \nabla_\omega u^{\nu+1}). \end{cases} \tag{6}$$

First, let us derive the updating formulations for $u^{\nu+1}$ and $d^{\nu+1}$. Denote

$$H(u) = \int_\Omega \int_\Omega G_\sigma(y-x) \sum_{k=1}^{K} \frac{[\frac{f(x)}{\beta(y)^\nu} - c_k^\nu - u(x)]^2}{2(\sigma_k^2)^\nu} \phi_k^\nu(x)\mathrm{d}x\mathrm{d}y.$$

Ignoring the constant terms, the minimization problem for u and d can be rewritten as

$$(u^{\nu+1}, d^{\nu+1}) = \arg\min_{u,d} \ H(u) + \mu|d|_1 + \frac{r}{2}||d - \nabla_\omega u + \frac{\lambda^\nu}{r}||_2^2.$$

We define $b = -\frac{\lambda}{r}$, together with the second updating formula in (6), one get the following iterative scheme:

$$\begin{cases} (u^{\nu+1}, d^{\nu+1}) = \arg\min_{u,d} \ H(u) + \mu|d|_1 + \frac{r}{2}||d - \nabla_\omega u - b^\nu||_2^2, \\ \qquad\qquad b^{\nu+1} = b^\nu + \nabla_\omega u^{\nu+1} - d^{\nu+1}. \end{cases} \tag{7}$$

Note that (7) actually is the split Bregman iteration and b is the Bregman vector [19]. We shall use an alternative minimization for u and d. The Euler-Lagrange equation for u is:

$$(\sum_{k=1}^{K} \frac{\phi_k^\nu}{(\sigma_k^2)^\nu} - r\triangle_\omega)u = \sum_{k=1}^{K} \frac{\phi_k^\nu}{(\sigma_k^2)^\nu} \left[f(x) \int_\Omega G_\sigma(y-x) \frac{1}{\beta(y)^\nu} \mathrm{d}y - c_k^\nu \right] + r\mathrm{div}_\omega(b^\nu - d^\nu), \tag{8}$$

The above equation is linear. Its approximate solution $u^{\nu+1}$ can be easily computed by a Gauss-Seidel process.

Once $u^{\nu+1}$ and b^ν is known, the minimizer $d^{\nu+1}$ is given by the following shrinkage operation:

$$d^{\nu+1} = \mathrm{shrink}(\nabla_\omega u^{\nu+1} + b^\nu, \frac{\mu}{r}) = \frac{\nabla_\omega u^{\nu+1} + b^\nu}{|\nabla_\omega u^{\nu+1} + b^\nu|} \max\{|\nabla_\omega u^{\nu+1} + b^\nu| - \frac{\mu}{r}, 0\} \tag{9}$$

For $\phi_k^{\nu+1}$ and $\Theta^{\nu+1}$, both of them have explicit solutions. To simplify the notations, we define

$$q_k^\nu(x) \triangleq \frac{\gamma_k^\nu}{\sqrt{(\sigma_k^2)^\nu}} \exp\left(-\frac{1}{2(\sigma_k^2)^\nu} \int_\Omega G_\sigma(y-x) \left(\frac{f(x)}{\beta^\nu(y)} - c_k^\nu - u^{\nu+1}(x) \right)^2 \mathrm{d}y \right),$$

$$s^{\nu+1}(y) \triangleq \sum_{l=1}^{K} \int_\Omega G_\sigma(y-x) \frac{c_l^{\nu+1} + u^{\nu+1}(x)}{(\sigma_l^2)^{\nu+1}} f(x)\phi_l^{\nu+1}(x)\mathrm{d}x,$$

$$t^{\nu+1}(y) \triangleq \sum_{l=1}^{K} \frac{1}{(\sigma_l^2)^{\nu+1}} \int_\Omega G_\sigma(y-x) f^2(x)\phi_l^{\nu+1}(x)\mathrm{d}x.$$

Then the solutions for the E-step and M-step are given by:

$$\begin{cases} \phi_k^{\nu+1}(x) = \dfrac{q_k^\nu(x)}{\sum\limits_{l=1}^{K} q_l^\nu(x)}, \\ \gamma_k^{\nu+1} = \dfrac{\int_\Omega \phi_k^{\nu+1}(x)\mathrm{d}x}{\int_\Omega 1\mathrm{d}x}, \\ c_k^{\nu+1} = \dfrac{\int_\Omega \phi_k^{\nu+1}(x)[f(x)\int_\Omega G_\sigma(y-x)\frac{1}{\beta^\nu(y)}\mathrm{d}y - u^{\nu+1}(x)]\mathrm{d}x}{\int_\Omega \phi_k^{\nu+1}(x)\mathrm{d}x}, \\ (\sigma_k^2)^{\nu+1} = \dfrac{\int_\Omega \phi_k^{\nu+1}(x)\int_\Omega G_\sigma(y-x)\left(\frac{f(x)}{\beta^\nu(y)} - c_k^{\nu+1} - u^{\nu+1}(x)\right)^2 \mathrm{d}y\mathrm{d}x}{\int_\Omega \phi_k^{\nu+1}(x)\mathrm{d}x}, \\ \beta^{\nu+1}(y) = \dfrac{-s^{\nu+1}(y) + \sqrt{(s^{\nu+1}(y))^2 + 4t^{\nu+1}(y)}}{2}. \end{cases} \tag{10}$$

Our algorithm with weight ω updating can be summarized as in the following:

Algorithm 1 (ALM-EM algorithm) *Given K, Choosing $\Theta^0, u^0, \phi^0, b^0, d^0$, and the parameters μ, r. Let $\nu = 0$ and calculate the initial weight ω^f. Do:*
1. ALM step: updating $u^{\nu+1}, d^{\nu+1}$ and $b^{\nu+1}$ according to (8), (9) and the second equation in (7), respectively.
2. If $||u^{\nu+1} - u^\nu||_2^2 < 10^{-5}||u^\nu||_2^2$, end the algorithm; else go to the next step.
3. E-step: updating $\phi^{\nu+1}$ using the first equation of (10).
4. M-step: updating the parameter set $\Theta^{\nu+1}$ using in (10).
5. Updating weight: if mod$(\nu+1, 5) == 1$, compute $\omega^{u^{\nu+1}/\beta^{\nu+1}}$ using (4). Set $\nu = \nu + 1$, and go to the ALM step.

4 Numerical Experiments

4.1 Parameters and Initial Values Selection

In this section, we give some guidelines and criterions on selection of the parameters and initial values. Here we suppose the observed image $f(x) \in [0, 1]$.

The parameter K is the number of the Gaussian PDF and it usually set to 2 or 3. Larger K can better models the true distribution of noise in some real applications, but the algorithm would be more time-consuming. In this paper, we set $K = 2$ for all the experiments.

The σ in the Gaussian kernel G_σ controls the smoothness of bias function β. Generally speaking, we need to choose a large value to keep the β smooth due to the fact that the illumination or intensity inhomogeneity in an images is often slow-varying. In our tests, we choose $\sigma = 10$.

The regularization parameter μ depends on the noise level. We find that μ in our model is not so sensitive to the noise level as in the nonlocal ROF model. This might be related to the fact that the introduction of the noise variance parameter σ_k^2, and σ_k^2 can adaptively balance the data-fidelity and the nonlocal TV terms together with μ. Experimental results show that $\mu \in [1, 15]$ can yield good results for different noise levels. In the experiments, unless otherwise specified, we set $\mu = 5$. In addition, we set penalty parameter $r = 200$.

The initial value $b^0 = d^0 = 0, \gamma_1^0 = \gamma_2^0 = \frac{1}{2}, c_1^0 = c_2^0 = 0, \sigma_1^2 = 0.1, \sigma_2^2 = 0.01, \phi_1 = 1, \phi_2 = 0$ are used. We can assume the desirable β to be around 1, and thus we set $\beta^0 = (G_\sigma * f + 1.5)/2$. Finally, we let $u^0 = G_\sigma * \frac{f}{\beta}$.

4.2 Experimental results

We first mention that the proposed model will reduced to the nonlocal ROF model by setting control parameter $\beta = 1$ and others in Θ to be equal, i.e. $\gamma_1 = \gamma_2, \sigma_1^2 = \sigma_2^2$, and so on. Thus if the illumination of an image is very homogeneous and the noise obeys a single Gaussian distribution, our method produces similar results as the nonlocal ROF model.

The superiority of our model is that it can work well under inhomogeneous illumination even with noise mixing. We shall tests these out.

Fig. 1 shows the results of the nonlocal ROF model [7] and our model under Gaussian mixture noise. The original image is displayed in Fig.1(i), one can find that the illumination of the original image itself is not homogeneous and the intensity on the left side is slightly lighter than the one on the right. We add noise and get the observed image f as shown in Fig.1(a). Here, the image f is corrupted by two additive white Gaussian noise with standard deviation $\frac{75}{255}$ and $\frac{20}{255}$, respectively. The mixture ratio is about $1 : 3$. As can be seen from the Fig.1(b) and Fig.1(c), the denoising result provided by the proposed method is better than the nonlocal ROF model. Firstly, the intensity of the reconstructed image in Fig.1(c) is more homogeneous than the one in Fig.1(b). This is caused by the use of β in our model. It can correct the inhomogeneous illumination. Secondly, our method can better preserve details in the texture areas and simultaneously clean the noise in the flat areas by adaptively adjusting the data term and nonlocal TV term through the control parameters σ_k^2 and ϕ_k. We use $\text{PSNR} = 10 \log_{10} \frac{1}{\text{var}(f - \hat{f})}$ to evaluate the quality of the denoising images, where $f, \hat{f}$ are observed and reconstructed images, respectively. For the proposed model, obviously, we need to define $\hat{f} = \beta u$ and then calculate the PSNR to make comparisons with other methods. The PSNR values for nonlocal ROF and the proposed are 23.94 and 27.34, respectively. Some estimated functions and parameters in the proposed approach are illustrated in Fig.1(d)-1(g). For visualization, we normalized β in [0,1] in Fig.1(d). The corrected noisy image can be found in Fig.1(e). We also calculate the variances $\sigma_f, \sigma_{\frac{f}{\beta}}$ of noisy image f and the corrected image $\frac{f}{\beta}$ respectively. We get $\sigma_f = 0.0576, \sigma_{\frac{f}{\beta}} = 0.0486$, which indicates the intensity in the latter image is more uniform. As mentioned earlier, our model can group the pixels into several clusters using different variances of the noise. In Fig. 1(f) and 1(g), the finally estimated partitions are displayed.

A denoising result with the proposed model under impulse noise plus Gaussian noise are given in Fig. 2. In this experiment, the image is contaminated by 25% salt-and-pepper noise together with Gaussian noise with standard deviation $\frac{15}{255}$. Here we take the common used adaptive median filter (AMF) [20] for comparison. It can be seen the AMF can clean impulse noise efficiently, but it fails in removing Gaussian noise and retaining the textures. Compared with the AMF, our method can give much better results. The denoised images and the estimated bias function provided by our method are displayed in the last two figures.

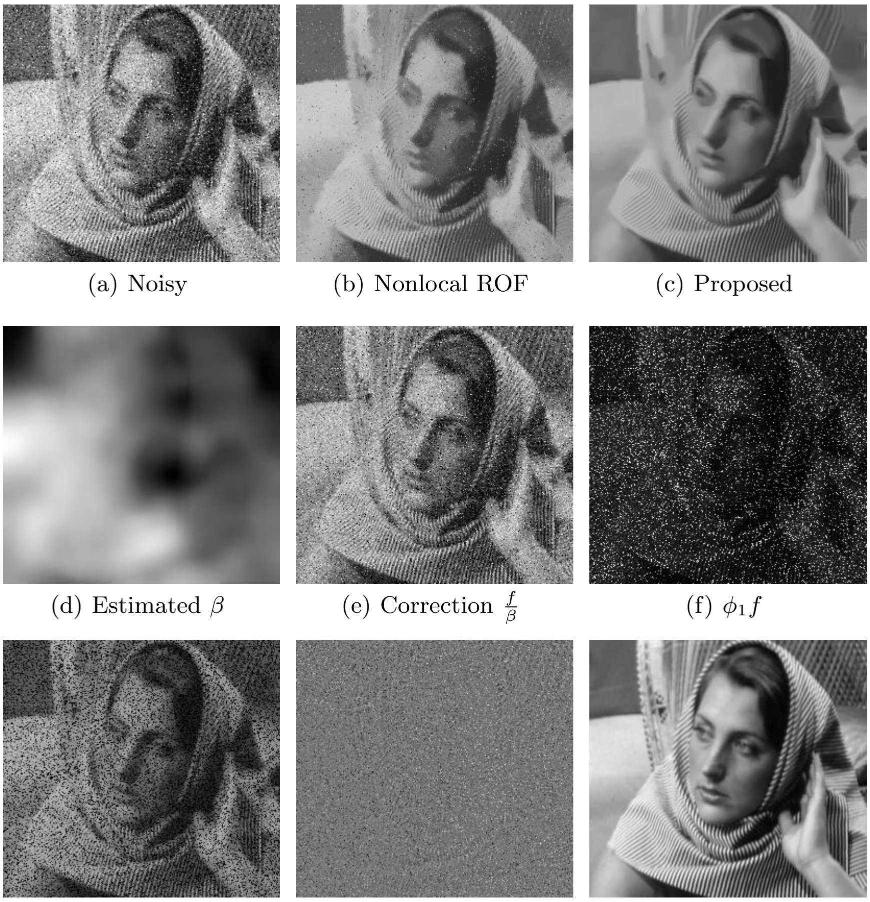
Fig. 3 shows result of applying the algorithm to MR images. In this experiment, we need to tune the regularization parameter $\mu = 15$ to get a smoothed image since the level of noise in the images is low. The denoised, corrected images and the estimated bias function β with the proposed algorithm are all illustrated in the last three figures. A benefit of the intensity correction is that the corrected images can be segmented easily with some center-based clustering methods such as Chan-Vese model, but it is very difficult to obtain a desirable segmentation result from the original data f.

5 Conclusion

We have presented an approach for simultaneous illumination correction and denoising. Numerical experiments demonstrated the method is very superior for mixed noise (e.g. impulse noise, Gaussian noise plus impulse noise etc.) compared to some earlier proposed nonlocal variational PDE based models. In addition, the non-uniform illumination function in the original data can be estimated and corrected by using the bias function. Our method can be extended to image segmentation, registration and some other computer vision problems.

References

1. Rudin, L., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. Physica D 60, 259-268 (1992)
2. Nikolova, M.: A variational approach to remove outliers and impulse noise. Journal of Mathematical Imaging and Vision 20, 99-120 (2004)
3. Bar, L., Sochen, N., Kiryati, N.: Image deblurring in the presence of impulsive noise. International Journal of Computer Vision 70, 279-298 (2006)
4. Cai, J., Chan, R., Nikolova, M.: Two-phase approach for deblurring images corrupted by impulse plus Gaussian noise. Inverse Problems and Imaging 2, 187-204 (2008)
5. Liu, J., Huang, H., Huan, Z., Zhang, H.: Adaptive variational method for restoring color images with high density impulse noise. International Journal of Computer Vision 90, 131-149 (2010)



(a) Noisy (b) Nonlocal ROF (c) Proposed

(d) Estimated β (e) Correction $\frac{f}{\beta}$ (f) $\phi_1 f$

(g) $\phi_2 f$ (h) The removed noise (i) Original

Fig. 1. A Comparison of nonlocal ROF [7] and the proposed model. (a)noisy image with PSNR= 17.01; (b) result with nonlocal ROF, PSNR= 23.94; (c) u, result with the proposed model, PSNR= 27.34 ; (d) the estimated bias function β; (e) the corrected image $\frac{f}{\beta}$; (f),(g) the estimated pixels with high level noise and low level noise, respectively; (h) the removed noise by the proposed method, i.e. $\frac{f}{\beta} - u$; (i) the original image.

(a) Impulse & Gaussian noise (b) AMF [20]

(c) Proposed (d) Estimated β

Fig. 2. Results under impulse noise plus Gaussian noise. (a)noisy image; (b) denoising with adaptive median filter (AMF) in [20], PSNR=23.47; (c) denoising with the proposed model, PSNR= 27.16 ; (d) the estimated bias function β;

(a) Observed f (b) Estimated u (c) Estimated β (d) $\frac{f}{\beta}$

Fig. 3. Applying to MR image.

6. Buades, A., Coll, B., Morel., J.M.: A review of image denoising algorithms, with a new one. Multiscale Modeling & Simulation 4(2), 490-530 (2005)
7. Gilboa, G., Osher, S.: Nonlocal operators with applications to image processing. Multiscale Modeling & Simulation 7(3), 1005-1028 (2008)
8. Jung, M., Vese, L.A.: Image restoration via nonlocal Mumford-Shah regularizers. UCLA C.A.M. Report 08-35 (2008)
9. Wells, W.M., Grimson, E.L., Kikinis, R., Jolesz, F.A.: Adaptive segmentation of MRI data. IEEE Transactions on Medical Imaging 15, 429-442 (1996)
10. Ahmed, M.N., Yamany, S.M., Mohamed, N., Farag A.A., Moriarty, T.: A modified fuzzy c-means algorithm for bias field estimation and segmentation of MRI data. IEEE Transactions on Medical Imaging 21, 193-198 (2002)
11. Ashburner, J., Friston, K.J.: Unified segmentation. Neuroimage 26(3), 839-851 (2005)
12. Li, C., Huang, R., Ding, Z., Gatenby, C., Metaxas, D.N., Gore, J.C. A variational level set approach to segmentation and bias correction of images with intensity inhomogeneity. In: MICCAI, 1083-1091 (2008)
13. Tai, X.C., Wu, C.L.: Augmented Lagrangian method, dual methods and split Bregman iteration for ROF model. UCLA C.A.M. Report 09-05 (2009)
14. Wu, C.L., Tai, X.C.: Augmented Lagrangian method, dual methods, and split bregman iteration for ROF, vectorial TV, and high order models. UCLA C.A.M. Report 09-76 (2009)
15. Rockafellar, R.T.: Convex analysis. Princeton University Press, Princeton (1970)
16. Bae, E., Yuan, J., Tai, X.C.: Global minimization for continuous multiphase partitioning problems using a dual approach. UCLA C.A.M. Report 09-75 (2009)
17. Goldstein, T., Osher, S.: The split Bregman method for L1 regularized problems. UCLA C.A.M. Report 08-29 (2008)
18. Setzer, M.: Split Bregman algorithm, Douglas-Rachford splitting and frame shrinkage. In: SSVM 2009 , 464-476 (2009)
19. Zhang, X., Burgery, M., Bresson, X., Osher, S.: Bregmanized nonlocal regularization for deconvolution and sparse reconstruction. UCLA C.A.M. Report 09-03 (2009)
20. Wang, H.H., Haddad, R.A.: Adaptive median filters: new algorithms and results. IEEE Transactions on Image Processing 4, 499-502 (1995)