

An Adaptive Method for Recovering Image from Mixed Noisy Data

Jun Liu Zhongdan Huan Haiyang Huang Haili Zhang

School of Mathematics science, Beijing Normal University

Laboratory of Mathematics and Complex Systems, Ministry of Education,

Beijing 100875, The People's Republic of China

Abstract

In this paper, we present a new version of the famous Rudin-Osher-Fatemi (ROF) model to restore image. The key point of the model is that it could reconstruct images with blur and nonuniform distributed noise. We develop this approach by adding several statistical control parameters to the cost functional, and these parameters could be adaptively determined by the given observed image. In this way, we could adaptively balance the performance of the fit-to-data term and the regularization term. The Numerical experiments have demonstrated the significant effectiveness and robustness of our model in restoring blurred images with mixed Gaussian noise or salt-and-pepper noise.

Keywords: Image deblurring, Image denoising, EM algorithm, Nonuniform distributed noise.

1 Introduction

In many applications, the images we obtain are contaminated by additive noise and blur. This process is often modeled by

$$g(x) = (k * f)(x) + n(x), \quad (1)$$

where $f(x)$ is the original clean image, $g(x)$ is the observed noisy blurred image, k is the point spread function (PSF) and also called the blur kernel, $n(x)$ is the additive noise and $*$ refers to the usual convolution. The problem of image reconstruction is to recover $f(x)$ from the degraded image $g(x)$. Traditional image recovery approaches are mainly based on variational techniques [2, 3, 4, 6, 8, 9, 10, 11, 13, 17], of which the most famous one is the ROF model, proposed by Rudin, Osher and E.Fatemi [3, 17]. In that model a regularized solution is obtained by minimizing the energy functional

$$T(f) = \frac{1}{2} \|k * f - g\|_{L^2}^2 + \lambda J_\beta(f), \quad (2)$$

where

$$J_\beta(f) = \int_{\Omega} \sqrt{|\nabla f|^2 + \beta} \, dx, \quad (3)$$

k is a known blur kernel, $\beta > 0$ is referred to as the stabilizing parameter, and $\lambda > 0$ is the regularization parameter. Numerous of experimental results (ref.[3, 4, 10, 12, 17]) have shown the effectiveness of these methods in removing Gaussian and uniform distributed white additive noise. However, in reality, images are usually degraded by mixed noise with different means, variances and even distributions. In this case, the traditional methods(e.g., ROF model) might not work well.

As is well known, it is very important to select the regularization parameter λ properly for the ROF model. Figure 1 shows that the ability of the ROF model in reconstructing images with blur and mixed Gaussian noise. We can see that the results are not satisfactory for these two particular parameters (λ s). In fact, experimental experience tells us that the results could not be further improved no matter what λ is selected. This is not accidental since model (2) employs a fixed global regularization term, therefore the blurring effect could not be sufficiently removed in areas with low level of noise, whereas the ringing effect may appear in which with high level of noise.

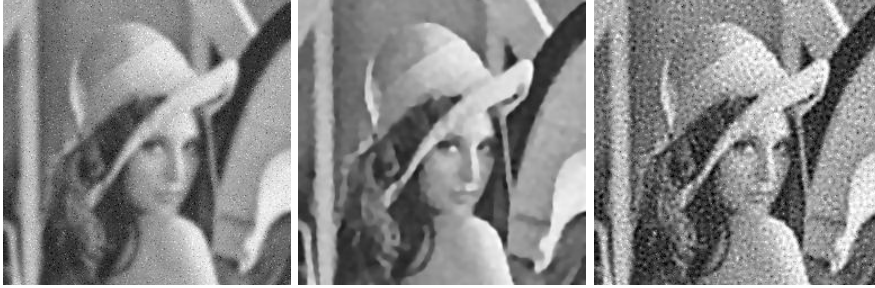


Figure 1: From left: the blurred and mixed Gaussian noisy image, where the variances of noise are $\sigma_1^2 = 1.0 \times 10^{-2}$, $\sigma_2^2 = 1.0 \times 10^{-4}$, the means are $\mu_1 = 0, \mu_2 = 0$ and the mixed ratio is 1:3; the reconstructed image based on the ROF model with $\lambda = 1.0 \times 10^{-2}$; the reconstructed image based on the ROF model with $\lambda = 1.0 \times 10^{-5}$

The above experiment tells us that the ROF model cannot work well when the blurred images are further degraded by mixed Gaussian noise. Thus to improve the quality of the reconstructed images, more information about such particular noise should be employed.

In this paper, we propose a new approach which includes some statistical information of noise. By adaptively updating the statistical control parameters of noise, we could balance the denoising and deblurring effects and thus get a better reconstruction. Meanwhile, we offer a method of how one can adaptively determine the statistical parameters of noise for the image restoration.

The paper is organized as follows: In Section 2, we give a statistical interpretation of the ROF model and propose a Gauss-Total Variation model (G-TV model). We interpret the ROF model statistically and some statistical control parameters of noise emerge automatically, then one can see that these parameters depend on the noise may play a similar role of the regularization parameter; Section 3 contains a Gaussian Mixture-Total Variation model (GM-TV model), which is used to deal with mixed noise; Experimental results regarding the proposed models are shown in Section 4; Finally, we conclude this paper in Section 5.

2 G-TV model

In this section, we develop a new interpretation of the ROF model based on statistical approaches. Suppose the image is defined in a bounded, smooth and open domain $\Omega \subset \mathbb{R}^2$. In the following, we assume at each point $x \in \Omega$, the intensity of noise $n(x)$ or $(k * f)(x) - g(x)$ is a random variable and all these random variables $\{n(x) | x \in \Omega\}$ are independent and identically-distributed (i.i.d.) as a Gaussian distribution $N(0, \sigma^2)$, i.e.,

$$p((k * f)(x) - g(x) | \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{|(k * f)(x) - g(x)|^2}{2\sigma^2} \right\}. \quad (4)$$

We want to maximize the likelihood functional

$$L(f, \sigma^2) = \prod_{x \in \Omega} \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{|(k * f)(x) - g(x)|^2}{2\sigma^2} \right\}, \quad (5)$$

or equivalently, to minimize the negative log-likelihood functional

$$-\ln(L(f, \sigma^2)) = \frac{1}{2} \int_{\Omega} \frac{|(k * f)(x) - g(x)|^2}{\sigma^2} dx + \frac{1}{2} \int_{\Omega} \ln(\sigma^2) dx + \frac{1}{2} \int_{\Omega} \ln(2\pi) dx. \quad (6)$$

Note that the last term in the expression is constant, so we can remove it and minimize the following functional,

$$E_1(f, \sigma^2) = \frac{1}{2} \int_{\Omega} \frac{|(k * f)(x) - g(x)|^2}{\sigma^2} dx + \frac{1}{2} \int_{\Omega} \ln(\sigma^2) dx, \quad (7)$$

where σ^2 is an unknown constant.

If the σ^2 is a fixed constant, minimizing (7) is equal to minimizing the residual

$$\frac{1}{2} \|k * f - g\|_{L^2}^2. \quad (8)$$

The above minimization problem is ill-posed, so we incorporate a regularization term and obtain the following cost functional

$$E(f, \sigma^2) = E_1(f, \sigma^2) + \lambda J(f). \quad (9)$$

Here we also choose the TV regularization term

$$J(f) = J_{\beta}(f) = \int_{\Omega} \sqrt{|\nabla f|^2 + \beta} dx. \quad (10)$$

By substituting (7) and (10) into (9), the cost functional becomes

$$E(f, \sigma^2) = \frac{1}{2} \int_{\Omega} \frac{|(k * f)(x) - g(x)|^2}{\sigma^2} dx + \frac{1}{2} \int_{\Omega} \ln \sigma^2 dx + \lambda \int_{\Omega} \sqrt{|\nabla f|^2 + \beta} dx. \quad (11)$$

For simplicity, we refer this as G-TV model.

The objective functional (11) is convex, continuous and lower bounded with respect to f for fixed σ^2 . Moreover, it is also continuous, convex and lower bounded with respect to σ^2 if the following condition holds,

$$\varepsilon \leq \sigma^2 \leq \frac{2}{A} \int_{\Omega} |(k * f)(x) - g(x)|^2 dx, \quad (12)$$

where ε is a very small positive number and

$$A = \int_{\Omega} 1 \, dx. \quad (13)$$

Therefore, following [7], the alternate minimization (AM) approach can be applied: in each step of the iterative procedure, we minimize with respect to one variable and keep the other one fixed. This leads to the following iterative algorithm:

Algorithm 1: Choose initial values for f^0 and $(\sigma^2)^0$. For $n = 1, 2, \dots$, do

1. Find f^{n+1} , such that

$$f^{n+1} = \arg \min_f E(f, (\sigma^2)^n); \quad (14)$$

2. Find $(\sigma^2)^{n+1}$, such that

$$(\sigma^2)^{n+1} = \arg \min_{\sigma^2} E(f^{n+1}, \sigma^2). \quad (15)$$

3. Check the convergence, if converged, stop; else goto 1.

Here and after, we will use $\arg \min E$ to denote the minimizer of E . We use a gradient based method to find the minimizers for (14), which is actually a solution of the following Euler-Lagrange equation

$$-\lambda \nabla \cdot \left(\frac{\nabla f}{\sqrt{|\nabla f|^2 + \beta}} \right) + \frac{\hat{k} * (k * f - g)}{(\sigma^2)^n} = 0, \quad (16)$$

where $\hat{k}$ is the conjugated function of k .

There are some existing numerical methods for solving the above nonlinear type partial differential equation(PDE), for instance, time marching, lagged diffusivity fixed point(FP) schemes and primal-dual methods. Considering the robustness and simplicity of the implementation of time marching algorithm, we will apply it to solve (16), that is, the solution of (16) is to be found by solving the following PDE to the steady state,

$$\begin{cases} \frac{\partial f}{\partial t} = \lambda \nabla \cdot \left(\frac{\nabla f}{\sqrt{|\nabla f|^2 + \beta}} \right) - \frac{\hat{k} * (k * f - g)}{(\sigma^2)^n}, \\ f|_{t=0} = g, \\ \frac{\partial f}{\partial \vec{n}} = 0, \end{cases} \quad (17)$$

where $\vec{n}$ is outward unit normal vector field.

The solution of step 2 is given by

$$(\sigma^2)^{n+1} = \frac{1}{A} \int_{\Omega} |(k * f^{n+1})(x) - g(x)|^2 \, dx. \quad (18)$$

According to (12), the stopping condition is

$$\int_{\Omega} |(k * f)(x) - g(x)|^2 \, dx < \varepsilon_1, \quad (19)$$

where ε_1 is a tolerated error.

The step 1 of the algorithm 1 is exactly the same as what is needed to solve the original ROF model, whereas in step 2 by updating σ^2 , we can adjust the influence of the smooth term. However, this algorithm is quite time-consuming due to the large cost to repeatedly solve the minimization problem (14). To overcome this difficulty, we modify the iterative scheme as following:

$$\frac{f^{n+1} - f^n}{\Delta t} = \lambda \nabla \cdot \left(\frac{\nabla f^n}{\sqrt{|\nabla f^n|^2 + \beta}} \right) - \frac{\hat{k} * (k * f^n - g)}{(\sigma^2)^n}, \quad (20)$$

$$(\sigma^2)^{n+1} = \frac{\int_{\Omega} |(k * f^{n+1})(x) - g(x)|^2 dx}{A}. \quad (21)$$

Consequently, we only compute the formulas (20), (21) repeatedly, and the cost of the modified algorithm is almost the same as the original ROF model. Studying this process carefully, you can find that the energy E would be descending in the iteration due to the formula (20) is actually a gradient descent step. Experimental results show that the equations (20),(21) have similar ability in reconstructing blurred and noisy images to the algorithm 1, but it is much faster.

Compared with the ROF model, the G-TV model utilizes a statistical parameter σ^2 , the variance of noise. By updating σ^2 in each iteration, we could adaptively balance the deblurring and denoising. The numerical experiments have shown that a satisfactory reconstruction result could be obtained within only a few iterations. Moreover, due to the introduction of σ^2 , the choice of λ is less sensitive than that of the ROF model. We will mention how to select λ later.

Remark: The algorithm works well when the image is contaminated by blur and various levels of uniform distributed noise. Meanwhile, it is still efficient even if there is no noise (see experiment II).

3 GM-TV model

The above G-TV model is quite effective in reconstructing images with blur and uniform distributed noise without changing the regularization parameter λ directly. However, it still could not work well when the image is contaminated with blur and mixed noise. So in this section we propose a new model to address this issue.

Assume at each point $x \in \Omega$, the intensity of noise $n(x)$ or $(k * f)(x) - g(x)$ is a random variable and all the random variables $\{n(x) | x \in \Omega\}$ are independent and identically-distributed with the following probability density function:

$$p(n(x) | \Theta) = \sum_{l=1}^M \alpha_l p_l(n(x) | \mu_l, \sigma_l^2), \quad (22)$$

where each p_l is a Gaussian density function with mean μ_l and variance σ_l^2 , and the parameter set $\Theta = \{\alpha_1, \dots, \alpha_M, \mu_1, \dots, \mu_M, \sigma_1^2, \dots, \sigma_M^2\}$ is chosen such that $\sum_{l=1}^M \alpha_l = 1$. In other words, the probability density function (PDF) is a mixture of M individual Gaussian components with different ratios.

By the same procedures as Section 2, we derive the cost functional

$$E(f, \Theta) = E_1(f, \Theta) + \lambda J(f), \quad (23)$$

where

$$E_1(f, \Theta) = \int_{\Omega} -\ln \left(\sum_{l=1}^M \frac{\alpha_l}{\sigma_l} \exp \left\{ -\frac{|(k * f)(x) - g(x) - \mu_l|^2}{2\sigma_l^2} \right\} \right) dx, \quad (24)$$

and

$$J(f) = J_{\beta}(f) = \int_{\Omega} \sqrt{|\nabla f|^2 + \beta} dx. \quad (25)$$

Consequently, the task is to minimize the above cost functional under the constraint $\sum_{l=1}^M \alpha_l = 1$.

We still use the AM algorithm to solve this problem:

Step 1. Find f^{n+1} , such that

$$f^{n+1} = \arg \min_f E(f, \Theta^n); \quad (26)$$

Step 2. Find Θ^{n+1} , such that

$$\Theta^{n+1} = \arg \min_{\Theta} E(f^{n+1}, \Theta) \quad \text{subject to} \quad \sum_{l=1}^M \alpha_l = 1. \quad (27)$$

In step 2, we need to estimate the parameters for a mixed PDF. The usual variational techniques are not suitable since it would result in an extremely complex system of nonlinear equations. To address this difficulty, we resort to the Expectation-Maximization (EM) algorithm [1, 5], an efficient algorithm for parameter estimation. The detailed information regarding this algorithm could be found in [1, 14]. The EM algorithm should be to iterate updating of α_l and updating of μ_l and σ_l^2 until convergence is achieved, with f fixed. Therefore, an expensive computational cost is required for the two steps. In fact, the process of minimizing the cost functional defined in (27) with EM algorithm is equivalent to minimizing another energy function (please see ref.[1, 5]), and we can utilize the other energy functional with the same decreasing behavior as the minimization in the step 2, so, we modify the energy $E_1(f, \Theta)$ as $\tilde{E}_1(f, \Theta)$

$$\begin{aligned} \tilde{E}_1(f, \Theta) = & \frac{1}{2} \sum_{l=1}^M \int_{\Omega} \frac{|(k * f)(x) - g(x) - \mu_l|^2}{\sigma_l^2} \omega_l^n(x) dx \\ & + \frac{1}{2} \sum_{l=1}^M \int_{\Omega} \ln(\sigma_l^2) \omega_l^n(x) dx \\ & - \sum_{l=1}^M \int_{\Omega} \ln(\alpha_l) \omega_l^n(x) dx + \sum_{l=1}^M \int_{\Omega} \alpha_l dx, \end{aligned} \quad (28)$$

where

$$\begin{aligned} \omega_l^n(x) &= \omega_l(x | (k * f^n)(x) - g(x), \Theta^n) \\ &= \frac{\alpha_l^n p_l((k * f^n)(x) - g(x) | \mu_l^n, (\sigma_l^2)^n)}{\sum_{v=1}^M \alpha_v^n p_v((k * f^n)(x) - g(x) | \mu_v^n, (\sigma_v^2)^n)}. \end{aligned} \quad (29)$$

$\tilde{E}_1(f, \Theta)$ has the same decreasing behavior as $E_1(f, \Theta)$, so we replace $E_1(f, \Theta)$

with $\tilde{E}_1(f, \Theta)$ and get the following cost functional

$$\begin{aligned}
\tilde{E}(f, \Theta) &= \tilde{E}_1(f, \Theta) + \lambda J(f) \\
&= \frac{1}{2} \sum_{l=1}^M \int_{\Omega} \frac{|(k * f)(x) - g(x) - \mu_l|^2}{\sigma_l^2} \omega_l^n(x) dx \\
&\quad + \frac{1}{2} \sum_{l=1}^M \int_{\Omega} \ln(\sigma_l^2) \omega_l^n(x) dx \\
&\quad - \sum_{l=1}^M \int_{\Omega} \ln(\alpha_l) \omega_l^n(x) dx + \sum_{l=1}^M \int_{\Omega} \alpha_l dx \\
&\quad + \lambda \int_{\Omega} \sqrt{|\nabla f|^2 + \beta} dx.
\end{aligned} \tag{30}$$

Now, we can use the variational techniques to minimize the cost functional (30). The detailed procedures can be found in Algorithm 2.

Algorithm 2: Choose initial values for f^0 , Θ^0 and calculate $\omega_l^0(x)$ by (29). For $n = 1, 2, \dots$, do

1. Find f^{n+1} , such that

$$\frac{f^{n+1} - f^n}{\Delta t} = \lambda \nabla \cdot \left(\frac{\nabla f^n}{\sqrt{|\nabla f^n|^2 + \beta}} \right) - \sum_{l=1}^M \hat{k} * \frac{(k * f^n - g - \mu_l^n) \omega_l^n}{(\sigma_l^2)^n}, \tag{31}$$

2. Find Θ^{n+1} and $\omega_l^{n+1}(x)$, such that

$$\begin{aligned}
\alpha_l^{n+1} &= \frac{1}{A} \int_{\Omega} \omega_l^n(x) dx, \quad A = \int_{\Omega} 1 dx, \\
\mu_l^{n+1} &= \frac{\int_{\Omega} ((k * f^{n+1})(x) - g(x)) \omega_l^n(x) dx}{\int_{\Omega} \omega_l^n(x) dx}, \\
(\sigma_l^2)^{n+1} &= \frac{\int_{\Omega} ((k * f^{n+1})(x) - g(x) - \mu_l^{n+1})^2 \omega_l^n(x) dx}{\int_{\Omega} \omega_l^n(x) dx}, \\
\omega_l^{n+1}(x) &= \frac{\alpha_l^{n+1} p_l((k * f^{n+1})(x) - g(x) | \mu_l^{n+1}, (\sigma_l^2)^{n+1})}{\sum_{v=1}^M \alpha_v^{n+1} p_v((k * f^{n+1})(x) - g(x) | \mu_v^{n+1}, (\sigma_v^2)^{n+1})}, \\
l &= 1, 2, \dots, M,
\end{aligned} \tag{32}$$

here

$$p_l((k * f^{n+1})(x) - g(x) | \mu_l^{n+1}, (\sigma_l^2)^{n+1}) = \frac{1}{\sqrt{2\pi(\sigma_l^2)^{n+1}}} \exp\left\{-\frac{|(k * f^{n+1})(x) - g(x) - \mu_l^{n+1}|^2}{2(\sigma_l^2)^{n+1}}\right\}. \tag{33}$$

The above approach is called GM-TV model.

The key point of this model is the introduction of $\omega_l^n(x)$, which could be determined by formulas in (32). For every point $x \in \Omega$, we have $0 \leq \omega_l^n(x) \leq 1$ and $\sum_{l=1}^M \omega_l^n(x) = 1$. So $\omega_l^n(x)$ can be viewed as a weight about the M individual deblurring components (please see the last term of (31)). If $M = 1$, we get $\omega_l^n(x) = 1$. Therefore for fixed constants $\alpha_l, \sigma_l^2, \mu_l$, the minimization problem (30) would become the ROF model.

Moreover, assume that the means $(\mu_l, l = 1, 2, \dots, M)$ of noise are all zeros, and the first term of (30) is written in the algebraic form, this fit-to-data term would become:

$$\frac{1}{2} \|\mathbf{K}\vec{f} - \vec{g}\|_{\mathbf{R}}^2 = \frac{1}{2} [(\mathbf{K}\vec{f} - \vec{g})^T \mathbf{R} (\mathbf{K}\vec{f} - \vec{g})],$$

where $\mathbf{R}$ is a diagonal matrix containing weight coefficients that are determined by σ_l^2 and $\omega_l^n(x)$. This is mentioned in reference [18], and we here give a new method which how to adaptively choose $\mathbf{R}$ according to the local variances of noise.

4 Numerical Methods and Experimental Results.

In this section, we develop the discrete scheme of the GM-TV Model based on the finite difference methods. The numerical scheme corresponding to the G-TV model in Section 2 could be deduced in the same manner. Suppose the image size is $N_1 \times N_2$ and let $g_{i,j}$ be the discrete form of $g(x)$ ($i = 1, \dots, N_1, j = 1, \dots, N_2$).

By applying the standard five-point finite difference scheme [15], the discrete equation for (31) is

$$\frac{f_{ij}^{n+1} - f_{ij}^n}{\Delta t} = \lambda s_{ij}^n(f) - v_{ij}^n(f), \quad (34)$$

where

$$s_{ij}^n(f) := \frac{T_{ij}^{xx}(f)(T_{ij}^{yy}(f)^2 + \beta) - 2T_{ij}^{xy}(f)T_{ij}^x(f)T_{ij}^y(f) + T_{ij}^{yy}(f)(T_{ij}^x(f)^2 + \beta)}{(T_{ij}^x(f)^2 + T_{ij}^y(f)^2 + \beta)^{\frac{3}{2}}}, \quad (35)$$

and

$$v_{ij}^n(f) := \left(\sum_{l=1}^M \hat{k} * \frac{(k * f^n - g - \mu_l^n) \omega_l^n}{(\sigma_l^2)^n} \right)_{ij}. \quad (36)$$

Some notations in the above equations are computed as following:

$$\begin{aligned} \omega_l^n &= \frac{\alpha_l^n p_l((k * f^n - g) | \mu_l^n, (\sigma_l^2)^n)}{\sum_{v=1}^M \alpha_v^n p_v((k * f^n - g) | \mu_v^n, (\sigma_v^2)^n)}, \\ T_{ij}^x(f) &= \frac{f_{i+1,j}^n - f_{i-1,j}^n}{2}, \\ T_{ij}^y(f) &= \frac{f_{i,j+1}^n - f_{i,j-1}^n}{2}, \\ T_{ij}^{xx}(f) &= f_{i+1,j}^n - 2f_{i,j}^n + f_{i-1,j}^n, \\ T_{ij}^{yy}(f) &= f_{i,j+1}^n - 2f_{i,j}^n + f_{i,j-1}^n, \\ T_{ij}^{xy}(f) &= \frac{f_{i+1,j+1}^n - f_{i-1,j+1}^n - f_{i+1,j-1}^n + f_{i-1,j-1}^n}{4}. \end{aligned} \quad (37)$$

The Neumann boundary condition for $f(x)$ is discretized as

$$f_{0,j} = f_{1,j}, \quad f_{N_1+1,j} = f_{N_1,j}, \quad f_{i,0} = f_{i,1}, \quad f_{i,N_2+1} = f_{i,N_2}. \quad (38)$$

Therefore, the iterative formula for solving f_{ij}^{n+1} can be summarized as

$$f_{ij}^{n+1} = f_{ij}^n + \Delta t (\lambda s_{ij}^n(f) - v_{ij}^n(f)). \quad (39)$$

The discrete scheme for (32) is

$$\begin{aligned}
\alpha_l^{n+1} &= \frac{1}{A} \sum_{i,j} (\omega_l^n)_{ij}, \quad A = N_1 \times N_2, \\
\mu_l^{n+1} &= \frac{\sum_{i,j} ((k * f^{n+1})_{ij} - g_{ij})(\omega_l^n)_{ij}}{\sum_{i,j} (\omega_l^n)_{ij}}, \\
(\sigma_l^2)^{n+1} &= \frac{\sum_{i,j} ((k * f^{n+1})_{ij} - g_{ij} - \mu_l^{n+1})^2 (\omega_l^n)_{ij}}{\sum_{i,j} (\omega_l^n)_{ij}}, \\
(\omega_l^{n+1})_{ij} &= \frac{\alpha_l^{n+1} p_l((k * f^{n+1})_{ij} - g_{ij} | \mu_l^{n+1}, (\sigma_l^2)^{n+1})}{\sum_{v=1}^M \alpha_v^{n+1} p_v((k * f^{n+1})_{ij} - g_{ij} | \mu_v^{n+1}, (\sigma_v^2)^{n+1})}, \\
l &= 1, 2, \dots, M.
\end{aligned} \tag{40}$$

In summary, the algorithm 2 can be implemented by the following procedures:

1. Choose $M, \Delta t, \beta$, initial values for $f^0 = g, \Theta^0 = \{\alpha_1^0, \dots, \alpha_M^0, \mu_1^0, \dots, \mu_M^0, (\sigma_1^2)^0, \dots, (\sigma_M^2)^0\}$ and calculate $\omega_l^0, l = 1, 2, \dots, M$ by (40);
2. Find f^{n+1} by computing (39);
3. (The EM algorithm step) Compute (40) to get Θ^{n+1} and $\omega_l^{n+1}, l = 1, 2, \dots, M$;
4. If $\sum_{i,j} ((k * f^{n+1})_{ij} - g_{ij})^2 \geq \varepsilon$, go to 2, otherwise, stop.

In this paper, all the convolutions are computed using the discrete cosine transformation (DCT) [10, 12].

As in [16], we introduce the following indexes

$$\begin{aligned}
SNR &= 10 \cdot \log_{10} \left(\frac{\sum_{i,j} (f_{ij}^2)}{\sum_{i,j} n_{ij}^2} \right), \\
BSNR &= 10 \cdot \log_{10} \left(\frac{\sum_{i,j} ((k * f)_{ij} - \bar{f})^2}{\sum_{i,j} n_{ij}^2} \right), \\
ISNR &= 10 \cdot \log_{10} \left(\frac{\sum_{i,j} (f_{ij} - g_{ij})^2}{\sum_{i,j} (f_{ij} - f_{ij}^{new})^2} \right),
\end{aligned} \tag{41}$$

where $\bar{f} = \frac{1}{N_1 N_2} \sum_{i,j} f_{ij}$, f^{new} is the reconstructed image and as in Section 1, f, g and n denote the original clear image, contaminated image and noise data respectively. Formulas in (41) are used to measure the level of noise, the ratio of blur and noise and the improvement of signal quality. Obviously, the larger the $ISNR$, the better the reconstructed image.

As we mentioned earlier, the performances of our models are less sensitive to the parameter λ than the ROF model's. But this does not mean that we can randomly choose λ as it could affect the speed of image recovery. Experiments show that thousands of iterations are needed to reconstruct images with heavy noise if we set λ to be 1, but much fewer iterations are enough if we use a bigger λ . Another thing we want to point out is that we observed that the visual effect looks similar if the λ is chosen from the interval $[5, 100]$. So, except for experiment I, we set λ to be 5 in all the following experiments.

Other parameters in our models are selected as following: the time step $\Delta t = 1.0 \times 10^{-4}$, the stabilizing parameter $\beta = 1.0 \times 10^{-6}$, $(\sigma^2)^0 = 0.1$ for the G-TV model and $M = 3$ for the GM-TV model. Initial parameters are set to be $\alpha_l^0 = \frac{1}{3}$, $\mu_l^0 = 0$, $(\sigma_l^2)^0 = 1.0 \times 10^{-l}$ for $l = 1, 2, 3$ in the GM-TV model. All images mentioned here range in intensity from 0 to 1 and the means of additive noise are all set to be zeros.

Next we present the computational results of the proposed models. The original images (denoted by f) are shown in Figure 2: the left one is a clean 256×256 Lenna image and the right one represents a piecewise constant 256×256 image with simple objects.



Figure 2: the original images for our experiments.

We begin with experiment I to test the influence of the parameter λ . Figure 3 shows the reconstructed results of the GM-TV model with different λ s. The middle one is the result for $\lambda = 1$, which takes 1000 iterations. And the right is the one for $\lambda = 200$, with 200 iterations. It can be seen that the right one is only a little smoother than the middle one, and required much less iterations.



Figure 3: From left: the noisy and blurred image, restored by the GM-TV model with $\lambda = 1$, $\lambda = 200$.

The purpose of experiment II is to show that algorithm 1 is robust with respect to

noise. The upper row in Figure 4 contains the contaminated images, with the levels of noise increasing from left to right (the left one is a blurred image with no noise). Corresponding reconstructed images are shown in the lower row. From these results, we can conclude that the G-TV model works well with different levels of noise without changing any parameters.



Figure 4: Top: the blurred images with different levels of noise; bottom: restored by the G-TV model.

The purpose of the third experiment (experiment III) is to demonstrate that the three models (ROF, G-TV, GM-TV) have similar ability in reconstructing images with blur and Gaussian uniform distributed noise. In Figure 5, the top left is the contaminated image with Gaussian blurred and Gaussian uniform distributed white noise. The standard deviation in the Gaussian blur kernel and the variance of noise are 3.0, 1.0×10^{-3} respectively. And the corresponding measures are $SNR = 25.112$, $BSNR = 15.2396$. The top right, bottom left and bottom right are the the restored images using the ROF model, G-TV model and GM-TV model. The regularization parameter for the ROF model is $\lambda = 8.0 \times 10^{-3}$. Corresponding $ISNRs$ (improvement of signal quality) are summarized in Table 1.

	ROF	G-TV	GM-TV
$ISNR$	2.1950	2.3920	2.3923

Table 1: the $ISNRs$ of the reconstructed images using the previous three models.

In this particular case, the reconstructed results of these three models have similar visual effects and the proposed GM-TV model appears to have no superiority. However, we need to select a proper λ for the ROF model to deal with the different levels of noise, whereas λ can be quite arbitrary for the proposed two models. Regardless of the selection for λ , the ROF model has a comparable reconstructing ability as the G-TV model. So in the following experiments, more attention is focused on the G-TV and GM-TV models.



Figure 5: Top left: the blurred image with uniform Gaussian noise; top right: the result of the ROF model; bottom left: the result of the G-TV model; bottom right: the result of the GM-TV model

The purpose of the fourth experiment(experiment IV) is to show the superiority of the GM-TV model in reconstructing images with mixed noise and different types of blur. In Figure 6, the top left is a Gaussian blurred ($\sigma = 2.0$) Lenna image of two mixed Gaussian white noise ($\sigma_1^2 = 1.5 \times 10^{-2}$, $\sigma_2^2 = 2.5 \times 10^{-4}$) with ratio 1 : 3 ($SNR = 19.2884$, $BSNR = 9.8200$); the bottom left is a blurred (a motion blur kernel of length=25, oriented at an angle $\theta = 0^\circ$ with respect to the horizontal line) image of three mixed Gaussian noise ($\sigma_1^2 = 1.25 \times 10^{-2}$, $\sigma_2^2 = 2.5 \times 10^{-4}$, $\sigma_3^2 = 5.0 \times 10^{-6}$) with the ratio is 1 : 1 : 2 ($SNR = 23.9096$, $BSNR = 11.6780$). The reconstructed results using the G-TV and GM-TV model are shown in the second and third column. Obviously, results in the last column are much clearer than those in the second column.

Experiment V illustrates that the proposed GM-TV model could effectively recover objects from images with blur and salt-and-pepper noise. The reason is that the salt-and-pepper noise could be approximately viewed as a mixture of two kinds of noise with different means and variances. The test images are presented in the first column with $SNR = 13.5556, 7.5284, 3.570$ and $BSNR = 3.6831, -2.3441, -6.3025$ respectively. The experiment results (Figure 7) reveal that GM-TV model could get a more satisfactory reconstructed results. The $ISNRs$ of the previous experiments are summarized in Table 2.

Until now, it is clear that the reconstructed results of the GM-TV model are much better not only visually but also numerically. Therefore, we can conclude that the GM-TV model is superior than other models in reconstructing objects from blurred and mixed noisy images.

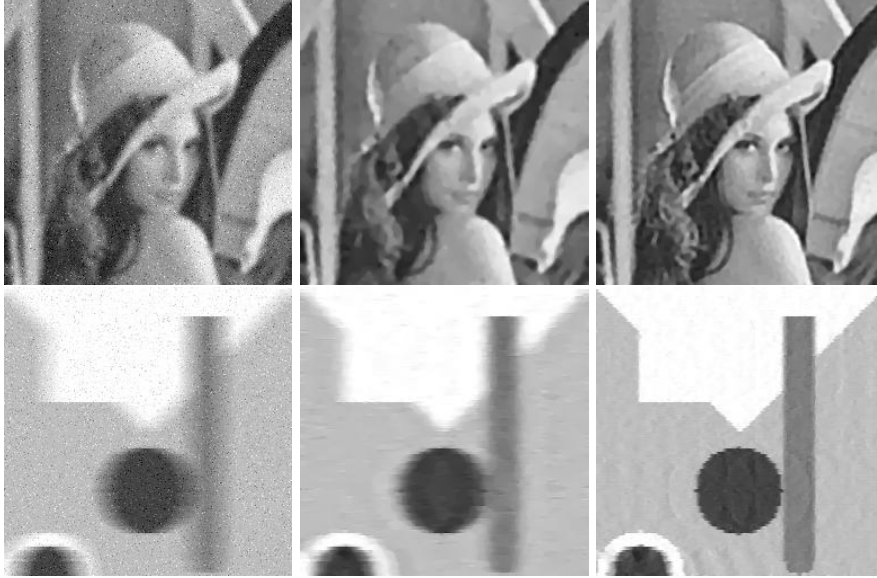


Figure 6: First Row: the Gaussian blurred and mixed noisy Lenna image, reconstructions obtained with the G-TV model and GM-TV model; Second Row: the blurred and mixed noisy image, reconstructions obtained with the G-TV model and GM-TV model.

Experiment	G-TV	GM-TV
IV	3.6378	5.1154
	4.0162	7.6336
V	5.4283	8.5004
	8.8304	13.5407
	9.6090	17.4579

Table 2: Comparison of *ISNRs* for previous experiments.

5 Conclusion and Discussion

In this paper, we present a new approach to adaptively reconstruct images from blurred and mixed noisy data. From the above experiments, we can see that this model could be used to remove Gaussian uniform distributed white noise as well as nonuniform distributed noise. We want to mention that the performance of the GM-TV model is not sensitive to the parameter M . In real applications we usually set M to be 2 or 3, which is enough to achieve good reconstructed results. Larger M would result in a more satisfactory estimation of mixed noise while more computational cost is required at the same time.

The algorithm is fast, robust and stable. Computation time for 256×256 images is about 50 seconds, using MATLAB on a Pentium(R)4 3.0GHz PC. Whereas for the highly contaminated case (the third one in experiment V), the computation time tends to be longer (about 3 minutes).

Recall that the G-TV model and GM-TV model are deduced from probability density function (PDF), so we could get an estimated PDF of noise for the G-TV model and GM-TV model. Theoretically, these estimators should be close to the true ones to guarantee satisfactory reconstructed results. Figure 8 contains some estimators



Figure 7: First Column: images with Gaussian blurred($\sigma = 3.0$) and salt-and-pepper noise(0.05, 0.2, 0.5); Second Column: the reconstructions of the first column using the G-TV model; Third Column: the reconstructions of the first column using the GM-TV model.

of the results in the experiments. Clearly, the dash line (estimated by the GM-TV) is much closer to the solid (true) than the dots line (estimated by the G-TV). Therefore, it is reasonable that the GM-TV model outperforms the G-TV model in recovering objects from mixed noisy data.

Finally, we have not given the theoretical results of the proposed GM-TV model, such as the existence and uniqueness of the regularized solution. These aspects are left for the future research.

6 Acknowledgments

We thank the anonymous reviewers for their constructive comments leading to improvements of the manuscript. The research has been supported by National Science Foundation of China (NSFC) No.10531040.

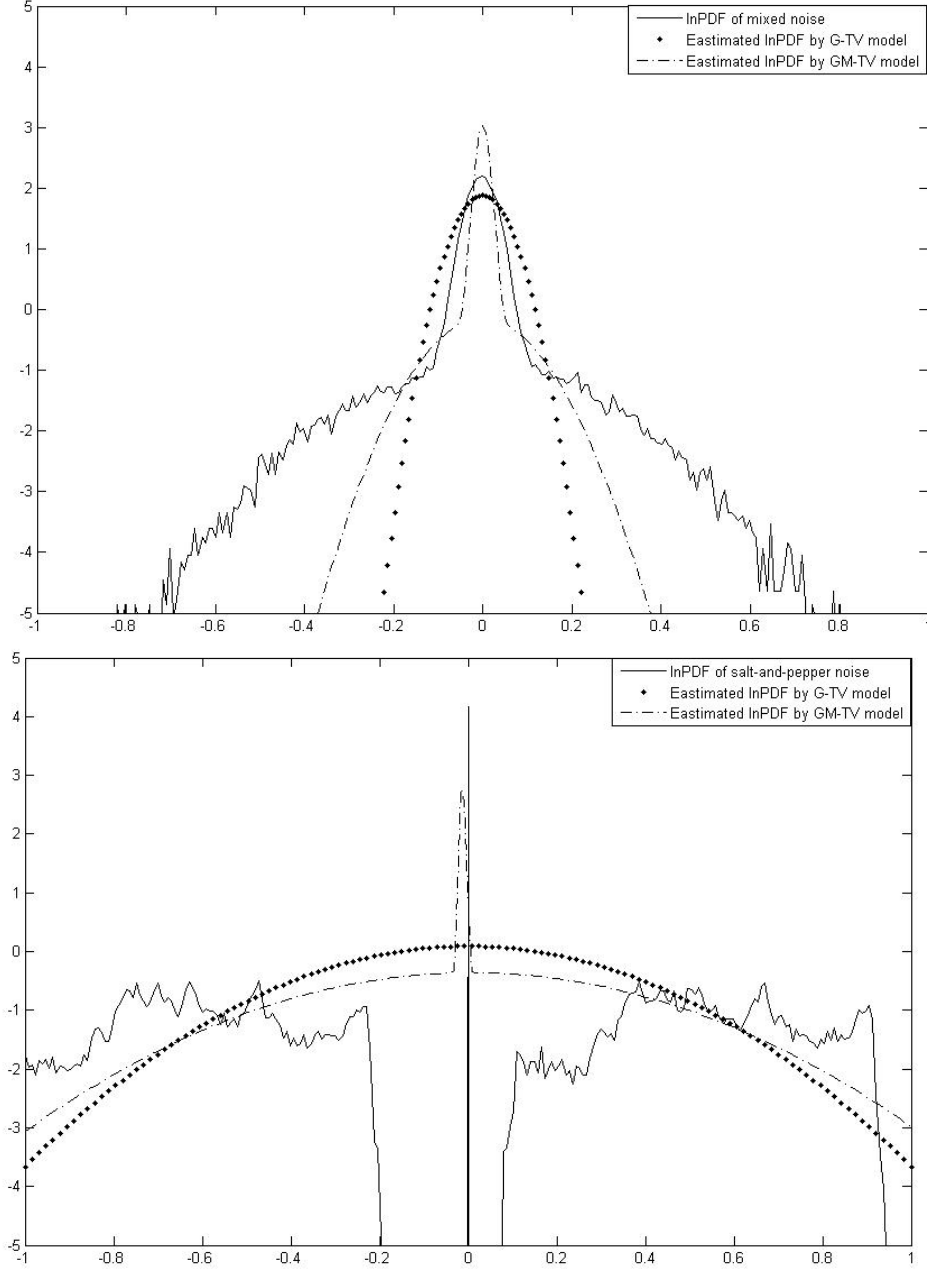


Figure 8: Top: The natural logarithm of probability density function(lnPDF) of noise in experiment IV for Lenna image, the solid line represents the true lnPDF of noise, the dots line indicates the estimated lnPDF by the G-TV model and the one by the GM-TV model is indicated by the dash line; Bottom: The same as top for experiment V, the level of salt-and-pepper noise is 0.5.

References

- [1] Bilmes, J. (1998). A Gentle Tutorial of the EM Algorithm and its Application to Parameter Estimation for Gaussian Mixture and Hidden Markov Models. Available at <http://citeseer.ist.psu.edu/bilmes98gentle.html>.
- [2] Acar, R., & Vogel, C. (1994). Analysis of Total Variation Penalty Methods. *Inverse Problems*, 10, 1217-1229.
- [3] Rudin, L., & Osher, S. (1994). Total Variation Based Image Restoration with Free Local Constraints. *Proc. IEEE ICIP*, 1, 31-35. Austin TX, USA.
- [4] Vogel, C., & Oman, M. (1998). Fast, Robust Total Variation-based Reconstruction of Noisy, Blurred Images. *IEEE Transactions on Image Processing*, 7, 813-824.
- [5] Redner, R., & Walker, H. (1984). Mixture Densities Maximum Likelihood and the EM Algorithm. *SIAM Review*, 26(2), 195-239.
- [6] Chan, R., Ho, C., & Nikolova, M. (2005). Salt-and-Pepper Noise Removal by Median-type Noise Detectors and Detail-preserving Regularization. *IEEE Transactions on Image Processing*, 14(10), 1479-1485.
- [7] Chan, T., & Wong, C. (1998). Total Variation Blind Deconvolution. *IEEE Trans. Image Processing*, 7, 370-375.
- [8] Nikolova, M. (2004). A Variational Approach to Remove Outliers and Impulse Noise. *Journal of Mathematical Imaging and Vision*, 20, 99-120.
- [9] Bar, L., Kiryati, N., & Sochen, N. (2006). Image Deblurring in the Presence of Impulsive Noise. *International Journal of Computer Vision*, 70, 279-298.
- [10] Shi, Y., & Chang, Q. (2007). Acceleration methods for image restoration problem with different boundary conditions. *Applied Numerical Mathematics*, 58(5) 602-614.
- [11] Bect, J., Blanc-Féraud, L., Aubert, J., & Chambolle, A. (2004). A l^1 -Unified Variational Framework for Image Restoration. *Proc. ECCV'2004, Prague, Czech Republic, Part IV: LNCS 3024*, 1-13
- [12] Michael, K., Chan, H., & Tang, W. (1999). A fast algorithm for deblurring models with neumann boundary conditions. *SIAM J.SCI.Comput*, 21(3), 851-866.
- [13] Vogel, R. (2002). Computational Methods for Inverse Problems. *SIAM*.
- [14] McLachlan, G., & Krishnan, H. (1997). The EM Algorithm and Extensions, *JOHN WILEY & SONS, INC, New York*.
- [15] He, L., Marquina, A., & Osher, J. (2005). Blind Deconvolution Using TV Regularization and Bregman Iteration. *Wiley Periodicals, Inc. Int J Imaging Syst Technol*, 15, 74-83.
- [16] Shi, Y., & Chang, Q. (2006). New time dependent model for image restoration. *Applied Mathematics and Computation*, 179 (1) 121-134.
- [17] Rudin, L., Osher, S., & Fatemi, E. (1992). Nonlinear total variation based noise removal algorithms. *Phys. D*, 60, 259-268.
- [18] Lagendijk, R., & Biemond, J. (1988). Regularized iterative image restoration with ringing reduction. *IEEE Transaction on Acoustics, Speech, and Signal Processing*, 36(12), 1874-1888.