Spatially adapted total variation model to remove multiplicative noise

Dai-Qiang Chen, Li-Zhi Cheng

Abstract—Multiplicative noise removal based on total variation (TV) regularization has been widely researched in image science. This paper proposes a new variational model which combines the TV regularizer with local constraints. It is also related to a TV model with spatially adapted regularization parameters. The automated selection of the regularization parameters is based on the local statistical characteristics of some random variable. The corresponding subproblem can be solved by the split Bregman method efficiently. Numerical examples demonstrate that the proposed algorithm is able to preserve small image details while the noise in the homogeneous regions is removed sufficiently. As a consequence, our method yields better denoised results than those of the recent state-of-the-art methods with respect to the SNR values.

Index Terms—total variation, Gamma noise, spatially adapted regularization, split Bregman method.

I. INTRODUCTION

Image denoising is one of the widely researched problems in image processing. For many special imaging systems such as synthetic aperture radar (SAR), laser or ultrasound imaging, single particle emission computed tomography (SPECT) and positron emission tomography (PET), image acquisition processes are different from the usual optical imaging technology and the standard additive Gaussian noise model is not suited in these situations. Instead, the multiplicative noise model provides an appropriate description of these special imaging systems. For more details about the generation of multiplicative noise in above mentioned imaging acquisition processes refer to [1, 2].

In this paper, we are interested in the problem of multiplicative noise removal. Assume $\Omega \subset \mathbb{R}^2$ is a rectangular domain. Let $u(x) : \Omega \to \mathbb{R}$ be the original image and $f(x) \in L^2(\Omega)$ be the observed image corrupted with multiplicative noise; the degradation model can be expressed as

$$f(x) = u(x)\eta(x),$$

where the noise $\eta$ is assumed to follow some distribution for a special application, e.g., in synthetic aperture radar, $\eta$ follows a Gamma distribution [2]; in ultrasound imaging, we may also confront with Rayleigh distributed [3] or Rician distributed noise [4]. Moreover, multiplicative Poisson noise appears in various applications such as electronic microscopy, positron emission tomography and single photon emission computerized tomography [5]. In this paper, we focus on Gamma distributed noise and then $\eta$ denotes Gamma noise with density function:

$$p_\eta(z) = \frac{L}{\Gamma(L)} z^{L-1} e^{-Lz} 1_{\{z \geq 0\}}.$$  (2)

The mean of $\eta$ is 1 and its standard deviation is $\frac{1}{\sqrt{L}}$.

Recently, various variational models and filter methods for removing multiplicative noise were proposed. The first total variation (TV) approach to solve the multiplicative model was presented as Rudin-Lions-Osher model [6], which used a constrained optimization approach with two Lagrange multipliers. However, their fitting term is not convex, which leads to difficulties in using the iterative regularization or the inverse space scale method. Later, Shi and Osher [7] use logarithmic transformation on both side of (1) and convert the multiplicative problem into the additive one. They then extend the relaxed inverse space scale (RISS) flows to the transformed additive problem. Numerical experiments show a good denoising effect and a significant improvement over earlier multiplicative models.

Derived from a MAP (maximum a posteriori) estimator, Aubert and Aujol [8] proposed a new total variation model to remove multiplicative Gamma noise of images

$$\min_{u \in BV(\Omega)} \int_\Omega |\nabla u| dx + \mu \int_\Omega \left( \log u + \frac{f}{u} \right) dx,$$  (3)

where $\mu > 0$ is a regularization parameter to control the trade-off between the goodness-of-fit of $f$ and a smoothness requirement due to the TV regularization. The TV regularizer $TV(u) = \int_\Omega |\nabla u|$ was first proposed by Rudin, Osher, and Fatemi [9] for Gaussian noise removal. In the space of functions of bounded variation (BV), the BV-seminorm is defined by

$$\int_\Omega |\nabla u| = \sup \left\{ \int_\Omega u \text{div}\bar{v} dx : \bar{v} \in (C_0^\infty(\Omega))^2, \|\bar{v}\|_{C_0^\infty(\Omega)} \leq 1 \right\}.$$  

and $u \in BV(\Omega)$ iff

$$\|u\|_{BV(\Omega)} = \int_\Omega |\nabla u| + \|u\|_{L^1(\Omega)} < \infty.$$  

The objective function in (3) is nonconvex, and hence the computed solution is sensitive to the initial value and is not necessary to be a global optimal solution. Recently, in many literatures [7, 10, 11], $\int_\Omega |\nabla \log u|$ is used to replace the regularizer $\int_\Omega |\nabla u|$ in the problem (3). Note that

$$\int_\Omega |\nabla \log u| = TV(\log u) = \int_\Omega (|\nabla u|/u)$$

people's Republic of China e-mail: chener050@sina.com.
is the well-known Weberized TV regularization term [12]. The essential idea of it is the use of Weber’s law, which was first described in 1834 by German physiologist E.H. Weber [13]. The law reveals the universal influence of the background stimulus $u > 0$ on human sensitivity to the intensity increment $\delta u$. Further, taking the rescaling $w = \log u$, this results in a convex minimization problem

$$\min_{w \in BV(\Omega)} \int_{\Omega} |\nabla w| dx + \mu \int_{\Omega} (w + f e^{-w}) dx,$$

which overcomes the drawbacks of the AA model (3). Experiments show that this model outperforms the AA model [10, 11].

Very recently, the equivalence of the problem (4) and I-divergence model

$$\min_{w \in BV(\Omega)} \int_{\Omega} |\nabla w| dx + \mu \int_{\Omega} (u - f \log u) dx,$$

has been researched [14, 15]. Note that the problem (5) is the classical model for Poisson noise removal. From [14, 15] we infer that (5) is appropriate for multiplicative noise as well.

The TV regularizer appears in the above models as well-known for preserving sharp discontinuities, and hence become one of the standard techniques for image restoration. However, it may not generate good enough results for images with many small structures and textures. The underlying cause of the problem is that only local features of the image are considered in the TV-based models. Recently, non-local (NL) filters [16-18] and variational models [19, 20] are developed for deblurring and denoising of images. These methods are appropriate for capturing the fine structures within images and have been shown to be very efficient. However, the computational complexity is too high and they need much more time for execution.

In order to balance the quality and efficiency of image restoration, TV models with spatially adapted regularization parameters were widely researched [21, 22]. We observe that, in the models (3)-(5), the value of $\mu$ controls the smoothness of the denoised image produced by the TV regularizer. Roughly speaking, small $\mu$ leads to over smoothing for small features in images so that fine details are removed; while large $\mu$ leads to little smoothing and thus noise will remain almost unchanged in the homogeneous regions. Therefore, spatially adapted regularization parameters are more suitable for the denoising problem. In [22], spatially varying constraints based on local variance measures were proposed for Gaussian noise removal. The similar idea is also extended to deal with multiplicative noise [23], where the proposed algorithm is to evolve the negative gradient flow based on the AA model and update the regularization parameters after each iteration. Since the proposed algorithm is an instance of the gradient descent method, it converges slowly and needs many iteration steps for good enough results. Moreover, since the AA model is non-convex and the regularization parameter is varying along the iterative steps, there is no convergence results for this algorithm.

On the other hand, a new class of methods for automated selection of the spatially adapted regularization parameters depending on the local statistical characteristics of the noise were developed for image denoising and deblurring [24-27]. This results in a convex minimization problem which can be solved efficiently by a superlinearly convergent algorithm based on Fenchel-duality and inexact semismooth Newton techniques. Moreover, the convergence properties of these methods are well established.

This paper discuss the local expected value estimator of a random variable of the form

$$\frac{f}{u} - \log \frac{f}{u},$$

and deriving from this conclusion we propose a new TV model with local constraints for multiplicative noise removal. In this model, we use the Weberized TV regularizer instead of the TV regularizer, and by adopting the logarithm transform we obtain further a convex sub-minimization problem which can be solved efficiently.

The rest of this paper is organized as follows. In section 2 we introduce the new locally constrained TV model for Gamma noise removal. In section 3 the existence and uniqueness of a solution to the proposed model are investigated, and further we describe the relation between the constrained problem and the corresponding TV model with spatially adapted regularization parameters. In section 4 we provide a new spatially adapted algorithm which combines an automated selection of the regularization parameters with the split Bregman method for solving the corresponding sub-minimization problem. In section 4 we report the numerical results comparing the proposed algorithm with those of the recent state-of-the-art methods.

II. CONSTRAINED TV MODEL FOR MULTIPLICATIVE NOISE REMOVAL

In this section, we first describe the statistical characteristic of some random variable with respect to Gamma noise $\eta$. It plays a central role in developing our method in this section.

**Proposition 1:** Let $\eta$ be a Gamma random variable (r.v) with mean 1 and standard deviation $\frac{1}{\sqrt{L}}$. Consider the following function of $\eta$:

$$I(\eta) = \eta - \log \eta.$$

Then the following estimate of the expected value of $I(\eta)$ holds true for large $L$:

$$E\{I(\eta)\} = 1 + \frac{1}{2L} + \frac{1}{12L^2} - \frac{5}{2L^3} + O\left(\frac{1}{L^3}\right).$$

**Proof:** We provide the proof in Appendix A.

Next we study the property of the formula defined by (6). Assume that a noise-free image is decomposed as follows:

$$u = u_c + u_t,$$

where $u_c > 0$ and $u_t > 0$ denote the cartoon part and the texture part of the image, respectively. Thus $f = (u_c + u_t) \cdot \eta$. Numerical Experiments demonstrate that TV-based model (3) and (4) can perform very well for cartoon images (piecewise constant regions). Thus these models using appropriate regularization parameter $\mu$ yield denoising results such that $\hat{u} = u_c$. Under this assumption we have

$$\frac{f}{\hat{u}} - \log \frac{f}{\hat{u}} = \eta - \log \eta + \left\{ \frac{u_t}{u_c} \eta - \log \left(1 + \frac{u_t}{u_c}\right) \right\}.$$
For the cartoon-like regions $\Omega_c \subset \Omega$, we have $u_t \approx 0$, and then according to (9) we obtain
\[
\int_{\Omega_c} \left( \frac{f}{u} - \log \frac{f}{u} \right) dx \approx \int_{\Omega_c} (\eta - \log \eta) dx \approx E[I(\eta)] \approx 1 + \epsilon
\]
where $\epsilon = \frac{1}{2\pi} + \frac{1}{\pi z^2}$. For the image regions $\Omega_1 \subset \Omega$ that contain rich texture information, we have $u_t \gg 0$ and then by (9) we obtain
\[
\int_{\Omega_1} \left( \frac{f}{u} - \log \frac{f}{u} \right) dx > \int_{\Omega_1} (\eta - \log \eta) dx \approx 1 + \epsilon. \tag{11}
\]
Motivated by the above arguments, we can use the local expected value estimator of $\frac{f}{u} - \log \frac{f}{u}$ as a constrained condition for the denoised image. Specifically, we define a local window centered at pixel $x$ as follows:
\[
\Omega^*_x = \left\{ y : \| y - x \|_\infty \leq \frac{r}{2} \right\},
\]
and assume that $w(x, y)$ is the mean filter, i.e.
\[
w(x, y) = \begin{cases} \frac{1}{\Omega^*_x}, & \text{if } \| y - x \|_\infty \leq \frac{r}{2}, \\ 0, & \text{else}. \end{cases}
\]
Then the local expected value estimator can be defined as follows:
\[
F(u)(x) = \int_{\Omega} w(x, y) \left( \frac{f}{u} - \log \frac{f}{u} \right) (y) dy. \tag{12}
\]
Based on the formula (12) we obtain the following Weberized TV minimization problem with local constraints
\[
\min_{u \in BV(\Omega)} \int_{\Omega} |\nabla \log u| \text{ over } u \in BV(\Omega) \tag{13}
\]
where 'a.e.' stands for 'almost everywhere'. Using the logarithm transform of the term $z = \log u$, we get the equivalent TV minimization problem
\[
\min_{z \in BV(\Omega)} \int_{\Omega} |\nabla z| \text{ over } z \in BV(\Omega) \tag{14}
\]
\[\text{s.t. } S(z) \leq 1 + \epsilon \text{ a.e. in } \Omega,
\]
where
\[
S(z)(x) = F(e^s)(x) = \int_{\Omega} w(x, y) (fe^{-s} + z - \log f)(y) dy. \tag{15}
\]
We observe that the function $g(s) = s - \log g \geq 1$ for any $s > 0$. Thus it is obvious that $S(z)(x) \geq 1$ for any $x \in \Omega$ by choosing $s = fe^{-s}$.

For later use we define the feasible set of (14) as follows
\[
C = \{ z \in BV(\Omega) : S(z) \leq 1 + \epsilon \text{ a.e. in } \Omega \}. \tag{16}
\]
Note that $\Omega$ is bounded and $f \in L^2(\Omega), z \in BV(\Omega)$. It is simple to show that $fe^{-s} + z - \log f \in L^1(\Omega)$ and $w \in L^\infty(\Omega \times \Omega)$, Thus we have $S(z) \in L^\infty(\Omega)$. Moreover, $S(\cdot)$ is continuous as a mapping from $L^2(\Omega)$ to $L^\infty(\Omega)$. Therefore, it is straightforward to show that $C$ is closed and convex.

III. SOME PROPERTIES OF THE PROPOSED TV MODEL

In this section, we address the problems of existence and uniqueness of a solution of the proposed model (14). Further, we observe that the constrained minimization problem (14) is related to a TV model with a spatially adapted regularization parameter $\lambda(x) \in L^2(\Omega)$ as follows:
\[
\min_{z \in BV(\Omega)} \int_{\Omega} |\nabla z| dx + \int_{\Omega} \lambda (z + fe^{-s}) dx. \tag{17}
\]
Based on this relationship, we propose a spatially adapted TV algorithm for multiplicative noise removal, which will be described in details next section.

In order to prove the existence of a solution to the problem (14), a technique similar to previous works [28] is adopted here. We first establish a BV-coercivity result for the function
\[
\mathcal{L}(z) = J(z) + \int_{\Omega} S(z)(x) dx
\]
with $J(z) = \int_{\Omega} |\nabla z| dx$. Then the existence is a direct result.

**Theorem 1**: Assume $f_{min}$ is a positive constant and $f \geq f_{min}$. Then $\|z\|_{BV} \to +\infty$ implies $\mathcal{L}(z) \to +\infty$. Further, the problem (14) admits a solution.

**Proof**: A sketch of the proof is given in Appendix B. ■

Due to the convexity (only) of the problem, there is no uniqueness result. However, if $w(x, y)$ in (15) is replaced by a modified mean filter
\[
\tilde{w}(x, y) = \begin{cases} \frac{1}{\epsilon_0}, & \text{if } \| y - x \|_\infty \leq \frac{r}{2}, \\ 0, & \text{else}, \end{cases}
\]
where $0 < \epsilon_0 \ll \min(1, \frac{1}{M})$ and $\tilde{w}$ satisfy $\int_{\Omega} \int_{\Omega} \tilde{w}(x, y) dx dy = 1$. Then a uniqueness result can be established. For the proof refer to Appendix C.

**Theorem 2**: Let the assumptions of Theorem 1 hold true. In addition, we suppose that for any constant $c$, $c\chi_{\Omega} \notin \mathcal{C}$, where $\chi_{\Omega}(x) = 1$ for $x \in \Omega$. Then, the solution of the problem (14) is unique.

It is obvious that the condition $c\chi_{\Omega} \notin \mathcal{C}$ is almost surely satisfied if $f$ is the product of some regular (non-constant) image and Gamma noise. Next we study the relation between the problem (14) and (17). A technique based on the first-order optimality conditions is used in some previous works such as the spatially adapted TV $- L^\tau (\tau = 1, 2)$ models for impulse or Gaussian noise removal. This method can be extended to the more general case and it is also suited for our problem here. First, the following penalty problem is considered:
\[
\min_{\mu \geq 0} \mathcal{L}_\mu(z) = J(z) + \mu \int_{\Omega} \left\{ \max(S(z) - (1 + \epsilon), 0)^2 \right\} dx
\]
\[\text{s.t. } z \in BV(\Omega), \tag{18}\]
where $\mu > 0$ denotes a penalty parameter. Then we have the following result.

**Theorem 3**: Let the assumptions of Theorem 1 hold true. Then problem (18) admits a solution $\tilde{z}_\mu \in BV(\Omega)$ for every $\mu > 0$. Furthermore, for $\mu \to +\infty$, $\{ \tilde{z}_\mu \}$ converges weakly
along a subsequence in $L^2(\Omega)$ to a solution of (14). Moreover, the formula

$$\|\max(S(z_\mu) - 1 - \epsilon, 0)\|_{L^2(\Omega)} = o\left(\frac{1}{\sqrt{\mu}}\right)$$

holds true, where $\lim_{\mu \to 0} o(x)/x = 0$ for any $x \in \mathbb{R}$.

**Proof:** Since similar proof appears in several papers [26, 27], we omit it here.

Finally, we state the first-order optimality characterization of a solution $\tilde{z}$ of the problem (14), and this implies that $\tilde{z}$ is also a solution of the spatially adapted TV model (17) with certain $\lambda$.

For subsequent results, we define

$$\lambda_\mu = \mu \max(S(z_\mu) - 1 - \epsilon, 0),$$

(20)

and $\lambda = \int_{\Omega} w(x, y) \lambda(x) dx$.

Then we have

$$\mu(\max(S(z_\mu) - 1 - \epsilon, 0)) \leq \int_{\Omega} \lambda_\mu dx - C_\mu = \int_{\Omega} \lambda_\mu q(z_\mu) dx - C_\mu,$$

(22)

where $q_f(s) = e^{-s} + s - \log f$ and $C_\mu = (1 + \epsilon) \int_{\Omega} \lambda_\mu dx$.

Inspired by the formulas (18), (22) and the results in Theorem 3, we have the following result for the relation of (14) and (17):

**Theorem 4:** Let the assumptions of Theorem 1 hold true, and let $\tilde{z}$ denote a weak limit point of $\{z_{\mu_n}\}$ as $\mu_n \to +\infty$. Moreover, we assume that $\|z_{\mu_n}\|_{L^2(\Omega)} \to \|\tilde{z}\|_{L^2(\Omega)}$ as $\mu_n \to +\infty$, and that there exists $C > 0$ such that $\|\lambda_{\mu_n}\|_{L^1(\Omega)} \leq C$ for any $n \in \mathbb{N}$. Then there exist $\tilde{\lambda} \in L^\infty(\Omega)$, a bounded Borel measure $\tilde{\mu}$ and a subsequence $\{\mu_{n_k}\}$ such that the following conclusions hold true:

(i) $\lambda_{\mu_{n_k}}$ converges weakly to $\tilde{\lambda}$ in $L^\infty(\Omega)$, and $\tilde{\lambda} \geq 0$ a.e. in $\Omega$.

(ii) There exists $j(\tilde{z}) \in \partial J(\tilde{z})$ such that

$$j(\tilde{z}), q \int_{\Omega} \tilde{\lambda}(1 - fe^{-s}) q dx = 0, \text{ for all } q \in BV(\Omega).$$

(iii) $\int_{\Omega} \psi \lambda_{\mu_{n_k}} \to \int_{\Omega} \psi d\tilde{\lambda}$ for all $\psi \in C(\Omega)$, and $\int_{\Omega} \lambda_{\mu_{n_k}}(S(z_{\mu_n}) - 1 - \epsilon) dx \to 0$.

**Proof:** The proof is similar to theorem 6 of [26], we omit it here.

Following Theorem 3 and Theorem 4(ii) we observe that one solution of the constrained problem satisfies the first order optimality condition of (17) with $\lambda = \tilde{\lambda}$. Further, assume that (19) still holds with $o(1/\sqrt{\mu})$ replaced by $o(1/\mu)$. Then from (20) we conclude that $\{\lambda_{\mu_{n_k}}\}$ is bounded in $L^2(\Omega)$ and then $\lambda^0$ is the weak limit of a subsequence $\{\lambda_{\mu_{n_k}}\}$. If the last relation in Theorem 4(iii) holds as $\int_{\Omega} \lambda^0(S(z) - 1 - \epsilon) dx = 0$, then we may equivalently write

$$\lambda^0 = \lambda^0 + \delta \max(S(z) - 1 - \epsilon, 0),$$

(23)

where $\delta > 0$ is a fixed constant.

Based on these results above an alternation iteration algorithm is proposed in the next section.

IV. SPATIALLY ADAPTED ALGORITHM FOR MULTIPlicative NOISE REMOVAL

In this section, we focus on reconstructing an image such that the local constrained conditions in (14) hold in both the detail regions and the homogeneous parts. From the discussion in section 3 we infer that it can be achieved by a suitable selection of the regularization parameter $\lambda$ in the problem (17).

Inspired by [25-27], we adopt an adjustment strategy similar to the Lagrangian multiplier update for the parameter $\lambda$. Initially, we choose $\lambda$ to be a small positive value, and then obtain an over-smoothed restored image which keeps most details in the residual. From section 2 we conclude that $S(z) > 1 + \epsilon$ in the texture-rich image regions, which implies that $\lambda$ needs to be increased there. Therefore, we propose the update rule of $\lambda$ as follows: Assume $\lambda^k$ denotes the current estimate of $\lambda^0$ in Theorem 4. Then we set the update by

$$\lambda^{k+1} = \lambda^k + \delta \max(S(z_k) - 1 - \epsilon, 0),$$

(24)

$$\lambda^{k+1} = \int_{\Omega} w(x, y) \lambda^{k+1}(x) dx,$$

(25)

where $\delta > 0$, and $z_k$ is the current estimate of the original logarithmic image $z$. Note that (4.1)-(4.2) are motivated by (20)-(21) and (23).

Based on the update formulas in (24) and (25), we obtain the spatial adaptive TV algorithm for multiplicative noise removal shown as Algorithm 1. For the convenience of the discussion below, we present it in a discrete version.

**Algorithm 1** Spatial adaptive TV algorithm for multiplicative noise removal

**Choose:** noisy image $f$; parameters $\lambda^0$ and $\delta$; local window size $r$;

**Initialization:** $k = 0, \lambda^k = \lambda^0$;

**Iteration:**

(1) Solve the discrete version of the problem

$$z^k = \arg \min_{z \in BV(\Omega)} \left\{ \int_{\Omega} \nabla z dx + \int_{\Omega} \lambda^k(z + fe^{-z}) dx \right\}$$

by the method proposed in subsection 4.1.

(2) Based on $z^k$, update $\lambda^k$ as follows:

$$\lambda^{k+1}(i, j) = (\lambda^k(i, j) + \delta \max(S(z_k)(i, j) - 1 - \epsilon, 0),$$

$$\lambda^{k+1}(i, j) = 1 - \frac{1}{\sigma^2} \sum_{(s, t) \in (i, j)} (\lambda^{k+1}(s, t)).$$

(3) Stop, or set $k = k + 1$.

In the following, we make several remarks on the proposed algorithm: (1)The initial value $\lambda^0$: the regularization parameter $\lambda$ in (17) controls the smoothness of the restored image. If $\lambda$ is a scalar, it should satisfy the following relation:

$$\lambda \propto 1/\sigma^2 = L,$$

where $\sigma^2$ is the variance of Gamma noise. Numerical examples in section 5 demonstrate that the proportional relation above is also suitable for the setting of the initial value $\lambda^0$. 

The update step \( \delta \): it controls the change rate of \( \lambda^k \) and \( \lambda^{k+1} \). Too small \( \delta \) leads to a rather slow adjustment of \( \lambda \); while too large \( \delta \) makes the algorithm unstable. We further study the influence of \( \delta \) on the denoised results in the experiments next section, and observe that a proportional relation can also be established between \( \delta \) and the variance of Gamma noise, i.e. \( \delta \propto \mu \). Moreover, for an appropriate value of \( \delta, k = 3 \) or 4 is enough for Algorithm 1.

The window size \( r \): too small window size \( r \) should yield obvious deviation between the local expected value estimator defined by (12) and the theoretical value \( 1 + \epsilon \), which is a consequence of the small sample sizes; whereas too large \( r \) makes the regularization parameter choice becomes rather global than local which compromises image details. In section 5 we research the influence of different window sizes on the restoration quality.

(4) The algorithm for solving the spatially adapted TV model defined by (17); the augmented Lagrangian method [11, 29] can be used to solve (17) directly, but it will generate a sub-minimization problem which needs to be solved by the Newton iteration method. Fortunately, based on the previous works [14, 15] we can solve the I-divergence model (4) with \( \mu = \lambda(z) \in L^2(\Omega) \) instead of it. For more details we refer to section 4.1.

A. The augmented Lagrangian method for the spatially adapted TV model

First, we explain the relation of the spatially adapted TV model (17) and the I-divergence model

\[
\min_{u \in BV(\Omega)} \int_{\Omega} |\nabla u| dx + \int_{\Omega} \lambda (u - f \log u) dx,
\]

which is given by the following result.

**Theorem 5:** Assume \( 0 < \lambda^0 \leq \lambda(x) \leq \lambda \) for any \( x \in \Omega \), where \( \lambda \) is a positive constant. Let \( \lambda_u \) and \( u_\lambda \) denote the solutions of (17) and (26) respectively. Then the following relation holds:

\[
\lambda_u \lambda \equiv e^{\epsilon u}.
\]

**Proof:** Let \( \phi(x,s) \equiv \lambda(x)(s + f(x)e^{-s}) \), \( \psi(x,s) \equiv \lambda(x) \left( s + f(x) \log \frac{f(x)}{s} + f(x) \right) \) and \( g(s) = e^s \). Since \( 0 < \lambda^0 \leq \lambda(x) \leq \lambda \), it is obvious that the assumptions (B1)-(B2) [15] hold, and thus according to Theorem 8 of [15] we conclude that \( \lambda_u \lambda \equiv e^{\epsilon u} \).

Next we describe the details of solving the I-divergence model (26) with the augmented Lagrangian method. The discrete version of the problem (26) is represented as

\[
\min_{u \in BV(\Omega)} TV(u) + \sum_{i,j=1}^{n} \lambda_{ij} (u_i - f \log u_i)_j.
\]

The isotropic discrete TV regularizer is defined by

\[
TV(u) = \sum_{i,j=1}^{n} \sqrt{(\Delta^h_{ij} u)^2 + (\Delta^v_{ij} u)^2},
\]

where \( \Delta^h_{ij} u \) and \( \Delta^v_{ij} u \) denote the horizontal and vertical first order differences at pixel \((i,j)\); while the anisotropic discrete TV regularizer is defined by the equation

\[
TV(u) = \sum_{i,j=1}^{n} |\Delta^h_{ij} u| + |\Delta^v_{ij} u|.
\]

The proposed algorithm is based on the augmented Lagrangian method. By introducing a new variable \( w \) we change the problem (27) into a constrained minimization problem as follows

\[
\min_{u,w} \left\{ TV(u) + \sum_{i,j=1}^{n} \lambda_{ij} (w - f \log w)_i,j \mid u = w \right\}.
\]

The augmented Lagrangian function for (30) is defined by

\[
\mathcal{L}^u_s = TV(u) + \sum_{i,j=1}^{n} \lambda_{ij} (w - f \log w)_i,j + \lambda (s - u - w)^2 + \frac{\eta}{2} \|u - w\|^2,
\]

where \( \eta \) is a positive penalty parameter and \( s \) is the Lagrangian multiplier. We apply the so-called alternating direction method of multipliers (ADMM) [29] to solve \( \min \mathcal{L}^u_s \). Since the details are similar to related works such as those in [11, 30], we omit them and give the results shown as Algorithm 2 directly.

**Algorithm 2** Augmented Lagrangian method for the spatially adapted TV model

**Choose:** image \( f \); regularization parameter \( \lambda \); penalty parameter \( \eta > 0 \); maximum iterative number \( K \); tolerance error \( \epsilon > 0 \).

**Initialization:** \( k = 0, u^k = w^k = f, b^k = 0 \).

**Iteration:**

\[
\begin{align*}
\min_{u,w} \left\{ ||p||_1 + \sum_{i,j=1}^{n} \lambda_{ij} (w - f \log w)_i,j \mid u = w, p = \nabla u \right\}.
\end{align*}
\]

where \( ||p||_1 = \sum_{i,j=1}^{n} \sqrt{|p^h_{ij}|^2 + |p^v_{ij}|^2} \) for isotropic total variation and \( ||p||_1 = \sum_{i,j=1}^{n} (|p^h_{ij}| + |p^v_{ij}|) \) for anisotropic total variation. However, a reasonable amount of iterations for the ROF denoising problem in Algorithm 2 is more efficient for obtaining a moderate accurate solution.

In the following experiments, we adopt the strategy of [11, section VI-A] for the running of Chambolle algorithm, i.e. in each iteration of the proposed algorithm, the internal variables
of Chambolle’s algorithm are initialized with those obtained in the previous iteration. Thus a small number of iterations such as \( K = 10 \) is needed for each call of Chambolle’s algorithm.

V. APPLICATION AND SIMULATED RESULTS

In this section, we make various experiments to evaluate the performance of the proposed algorithm. First, the setting of the initial regularization parameter \( \lambda^0 \), window size \( r \), and the step parameter \( \eta \) in Algorithm 1 is researched; second, we compare the performance of the proposed algorithm with those of the recent state-of-the-art methods introduced in [10, 11, 14], which solve the TV-based models with a scalar regularization parameter; In the last part of experiments we compare our approach with another recent state-of-the-art algorithm, proposed in [23], for which spatially varying regularization parameters are adopted and the solution is obtained by evolving the negative gradient flow for the corresponding Euler-Lagrange equation.

The code of Algorithm 1 are written entirely in Matlab, and all these algorithms are implemented under Windows XP and MATLAB 7.0 running on a Lenovo laptop with a Dual Intel Pentium CPU 1.8G and 1 GB of memory. The four original images are shown in Fig. 1.

A. Parameters selection

In this subsection we focus on the study of the parameters selection in Algorithm 1. Two images called ‘Lena’ and ‘Cameraman’ (see Figure 1) are used for our experiments. For each image, a noisy observation is generated by multiplying the original image by a realization of Gamma noise according to the formulas in (1)-(2) with \( L \in \{8,33\} \). The performance of the proposed algorithm is measured quantitatively by means of the signal-to-noise ratio (SNR), which is expressed as

\[
\text{SNR}(u, g) = 10 \log \left( \frac{\|u - \bar{u}\|_2}{\|g - \bar{u}\|_2} \right),
\]

where \( u \) and \( g \) denote the original image and the restored image respectively, and \( \bar{u} \) represents the mean of the original image. In each iteration of Algorithm 1, the sub-minimization problem (26) is solved by calling Algorithm 2. We choose \( K = 15, \epsilon = 10^{-3}, \text{and } \eta = 0.15 \) for the running of Algorithm 2.

In the following experiments, we study the influence of various parameters in Algorithm 1 on the quality of the denoised images.

**Experiment 1** (the selection of the update step \( \delta \)): For Algorithm 1, we observe that \( \delta \) controls the change speed of \( \lambda^k \) and hence has an important impact on the performance of the algorithm. Specifically, large \( \delta \) will lead to significant artifacts in image regions containing edges and details, especially for serious noise; while small \( \delta \) makes the adjustment of \( \lambda^k \) rather slow, which leads to an unacceptably large number of iterations.

In this example, we fix \( \lambda^0 = 0.1L \) and \( r = 17 \) for Algorithm 1. Under these conditions, we plot the SNR values for the denoised images with \( \delta \) varying from 2.5 to 40 in Fig. 2. From the plots we observe that the values of SNR are rather stable with respect to \( \delta \), unless \( \delta \) is too large or too small. We also find that the optimum value of \( \delta \) is inversely proportional to the variance of Gamma noise, i.e.

\[
\delta_{opt} \approx \mu_1 L,
\]

where \( \mu_1 \) is a positive constant. Moreover, we observe that the iteration number \( N = 3 \) or 4 is enough for the denoised images. Considering computational efficiency, we adopt \( N = 3 \) in the experiments below.

**Experiment 2** (the selection of the initial \( \lambda^0 \)): The regularization parameter \( \lambda^0 \) controls the trade-off between the goodness-of-fit and the regularizer defined by the TV-norm. In
other words, it reflects the smoothness of the denoised images. Therefore, a small $\lambda^0$ is suitable for images with serious noise; while a large $\lambda^0$ is needed for images with low noise.

In the following experiments, we choose $\delta = 5, 20$ for the noisy versions of the Cameraman image, and $\delta = 10, 20$ for the noisy versions of the Lena image respectively. Besides, we set $\tau = 17$ for Algorithm 1. Let $\lambda^0 = \mu_2 L$, the SNR values of the denoised images with $\mu_2$ varying from 0.02 to 0.4 are plotted in Fig. 3. From the plots we observe that the SNR values are rather stable with $\mu_2 \in [0.05, 0.2]$. In order to illustrate the influence of $\lambda^0$ on the restored images clearly, we show the results of the Lena image with $L = 33$ with different $\mu_2$ in Fig. 4. From the figure we observe that: too small $\mu_2$ will yield significant artifacts in some regions of the images (see Fig. 4(c)); while too large $\mu_2$ makes noise unchanged in the homogeneous regions (see Fig. 4(d)).

**Experiment 3** (the selection of the window size $\tau$): In this example, we test our method with different values of $\tau$. We adopt the same parameter settings as those in Experiment 2 for the update step $\delta$, and choose $\lambda^0 = 0.1L$. Fig. 5 shows the plots of the SNR values of the restored images with $\tau$ varying from 5 to 25. We observe that there is no considerable difference between the results with different values of $\tau$.

The results of the Cameraman image are shown in Fig. 6. From them we observe that some noise still persists in the denoised image with too small $\tau$, which is a consequence of the small sample sizes. However, with too large $\tau$, the regularization parameter choice becomes rather global than local which compromises image details. Therefore, a medium value for $\tau$ should be chosen for our experiments.

From the above examples, we obtain some rules for the selection of optimum parameters in Algorithm 1, and hence a fully automatic parameter choice strategy could be used for the test of the proposed approach in the following experiments.

**B. Comparison with other noise removal methods with a scalar regularization parameter**

In this subsection, we report the experimental results comparing the proposed algorithm with other TV-based models [10, 11, 14] for multiplicative noise removal. We use the stopping criterion for their iterative schemes as follows

$$\frac{\|u^{k+1} - u^k\|_2}{\|u^k\|_2} < \text{tol},$$  \hfill (34)
where \( u^k \) denotes the iterate of the scheme. We set \( \text{tol} = 0.001 \) for the following experiments.

In order to quantify the denoising performance, we list the SNR values of different denoised images by different methods in Table I. In this table, "HNW", "MIDAL", "I-divergence" represent the Huang-Ng-Wen method in [10], multiplicative image denoising by the augmented Lagrangian method in [11], and the I-divergence-TV model in [14], respectively. Note that these methods are all manual parameter models. Among them, MIDAL and I-divergence require one regularization parameter; while HNW involves two input parameters: one parameter is for the regularization, and the other is for the closeness between the two denoised images. As in [10], we fix the closeness parameter to be 19, and then we deal with only one parameter in the HNW method. In all these experiments, we adjust the regularization parameters of the three methods to be optimal, in the sense that after many trials they yield the highest SNR denoising results. Meanwhile, we choose \( \delta \approx 0.625L \), \( \lambda^0 = 0.1L \), and \( r = 17 \) for Algorithm 1.

From the table we observe that the proposed algorithm can give better denoising results in terms of SNRs than other methods. Quantitatively, we note that on average, 0.3~0.6dB improvement in SNR is obtained by our method.

<table>
<thead>
<tr>
<th>Image</th>
<th>L</th>
<th>Noisy</th>
<th>HNW</th>
<th>MIDAL</th>
<th>I-divergence</th>
<th>Our method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cameraman</td>
<td>5</td>
<td>11.33</td>
<td>11.95</td>
<td>12.32</td>
<td></td>
<td>12.57</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>13.50</td>
<td>13.55</td>
<td>13.58</td>
<td>14.43</td>
<td>14.86</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>15.79</td>
<td>16.01</td>
<td>16.39</td>
<td></td>
<td>16.59</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>11.74</td>
<td>11.74</td>
<td>11.74</td>
<td>12.31</td>
<td>12.43</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>13.80</td>
<td>13.80</td>
<td>13.83</td>
<td>14.28</td>
<td>14.54</td>
</tr>
<tr>
<td>Bars</td>
<td>5</td>
<td>8.67</td>
<td>8.51</td>
<td>8.74</td>
<td></td>
<td>8.88</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>10.41</td>
<td>10.39</td>
<td>10.49</td>
<td>10.96</td>
<td>11.19</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>12.51</td>
<td>12.51</td>
<td>12.51</td>
<td>13.03</td>
<td>13.31</td>
</tr>
</tbody>
</table>

In Fig. 9, the denoised results of the Barbara image corrupted by Gamma noise with \( L = 25 \) are shown. The result obtained by our method can be found in Fig. 9(e) and the results obtained by the other methods in Fig. 7(b)-(d) respectively. Note that in 7(e) small features are better recovered as compared to 7(b)-(d). Meanwhile homogeneous regions appear smoother in 7(e) rather than in 7(b)-(d). For example, we observe that the camera and the tripod are sharper in 7(e) than in 7(b)-(d), and noise spots in the sky appear in 7(b)-(d) and not in 7(e). The improvement of the denoised images is due to the spatially adapted regularization parameter \( \lambda \). In Fig. 7(f) the values of the regularization parameter \( \lambda \) are presented in a gray scale. Light gray regions refer to large values of \( \lambda \), whereas dark gray belongs to zones where \( \lambda \) is small. From the figure we find that \( \lambda \) is large in the regions of the camera and the tripod. For a clearer comparison, we give the zoomed versions of certain parts of the denoised images in Fig. 8.

In Fig. 9, the denoised results of the Barbara image corrupted by Gamma noise with \( L = 25 \) are shown. Fig. 9(e) and Fig. 9(b)-(d) are the results obtained by our method and the other methods respectively. Similar effects as those obtained in the above example are found here. We observe that the details in the scarf are sharper in 9(e) than in 9(b)-(d), and noise

![Image](image_url)
spots in the background and the woman’s face are removed sufficiently in 9(e), and not in 9(b)-(d). The values of $\lambda$ are shown in Fig. 9(f). From the figure we observe that $\lambda$ is large in the regions of the scarf. In order to make the comparison clearer, we zoom into certain region of these results in Fig. 10.

In Fig. 11, we present the denoised results of the Bars image corrupted by Gamma noise with $L = 5$. From these results we observe that more noise spots exist in Fig. 11(b)-(d) than in 11(e). If we choose large values for the scalar regularization parameters, more details will be smoothed and the SNR values decrease. Our method avoids the problem and is able to automatically adjust the values of the regularization parameter in the homogeneous regions and detail regions.

C. Comparison with the noise removal method with spatially varying regularization parameters

In this subsection, we compare our method with the gradient descent method proposed in [23]. As a summary, it is to evolve the negative gradient flow

$$\frac{\partial u}{\partial t} = \text{div} \left( \frac{\nabla u}{|\nabla u|} \right) + (K + \bar{\lambda}) \left( f - \frac{u}{u^2} \right).$$  (35)

Fig. 9. (a) The noisy image with $L=25$, (b) the image denoised by the HNW method, (c) the image denoised by MIDAL, (d) the image denoised by $I$-divergence, (e) the image denoised by Algorithm 1, (f) final values of $\lambda$ in Algorithm 1.

Fig. 10. The zoomed version of the denoised images in Fig. 9. (a) the HNW method, (b) MIDAL, (c) $I$-divergence, (d) Algorithm 1.

Fig. 11. (a) The noisy image with $L=5$, (b) the image denoised by the HNW method, (c) the image denoised by MIDAL, (d) the image denoised by $I$-divergence, (e) the image denoised by Algorithm 1, (f) final values of $\lambda$ in Algorithm 1.
where $K$ is a two-dimensional Gaussian kernel, and $\tilde{\lambda}$ is updated by
\[
\tilde{\lambda}(x) = \frac{D(x)}{V(x)}, \quad x \in \Omega.
\] (36)

In the formula (36), $D(x) = \text{div} \left( \frac{\nabla u}{|\nabla u|} \right) (u - f)$, and $V(x) \approx \frac{\sigma^2}{K_r(\bar{r} - r)}$, where $r = \frac{u}{f}$, $\bar{r}$ is the local mean of $r$, and $\hat{u}$ is an approximation of the original image. This algorithm is inefficient for the noise removal especially when the noise is serious. In this case, we must choose a small iteration step to ensure convergence of the algorithm, and hence thousands of iteration steps are needed for good enough results. Therefore, we only consider the case of low noise here. For the gradient descent method, we use the stopping criterion that the mean value of $|u^k - u^{k-1}|$ should be less than 0.01, where $u^k$ denotes the iterate of the scheme. Besides, the window size is set to be 17, and the iteration step is set to be 0.2.

Two images, Barbara and Lena (see Fig. 1) are chosen for our experiments. The SNR values of different denoised images by different methods are shown in Table 2. From the table we find quantitatively that our method gives better denoised results with respect to the SNR values. In Fig. 12 the denoised results of the Lena image corrupted by Gamma noise with $L = 50$ are shown. We observe that small features such as the woman’s hair are sharper in Fig. 12(b) than in 12(a). The values of corresponding $\lambda$ (for the gradient descent method, $\lambda$ corresponds to $K * \tilde{\lambda}$ in (35)) are presented in Fig. 12(c) and 12(d). It is obvious that $\lambda$ obtained by our method detect the texture regions more exactly than that obtained by the gradient descent method. The denoised results of the Barbara image polluted by Gamma noise with $L = 33$ are shown in Fig. 13. We also find the similar effect as the above example. For instance, the textures on the scarf are better recovered by our method.

<table>
<thead>
<tr>
<th>Image</th>
<th>L</th>
<th>Noisy SNR (dB)</th>
<th>Gradient descent method</th>
<th>Our method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lena</td>
<td>25</td>
<td>5.39</td>
<td>14.11</td>
<td>14.57</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>8.44</td>
<td>13.31</td>
<td>16.24</td>
</tr>
<tr>
<td>Barbara</td>
<td>33</td>
<td>6.67</td>
<td>12.31</td>
<td>12.86</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>8.46</td>
<td>13.32</td>
<td>14.02</td>
</tr>
</tbody>
</table>

We have compared both algorithms in the visual quality of the denoised images. Next the corresponding implementation efficiency is reported. We display in Fig. 14 the curves of SNRs with respect to the CPU times of both algorithms. By comparison of these curves, we observe that the SNR values of the gradient descent method increase much more slowly than those obtained by our method. From the above analysis, we conclude that our proposed algorithm clearly outperforms the gradient descent method.

VI. Conclusion

In this paper, based on the statistical characteristics of some random variable, a total variation model with local constraints is proposed for Gamma noise removal. On further research we find that this model is equivalent to a TV model with
spatially adapted regularization parameters. Deriving from this relationship, we propose a spatially adapted TV algorithm which provides an accurate instrument for the parameter adjustment. The resulting algorithm is able to preserve the image details and smooth the homogeneous regions simultaneously. Besides, the proposed algorithm is completely automatized, i.e., a process of tuning regularization parameters is not needed. Numerical experiments indicate that our method outperforms those of the recent state-of-the-art methods.

**APPENDIX A**

**Proof of Proposition 1**

We provide a proof of Proposition 1 here.

**Proof:** From the fourth-order Taylor formula for \( \log(1 + \zeta) \), we obtain

\[
\log(1 + \zeta) = \zeta - \frac{1}{2} \zeta^2 + \frac{1}{3} \zeta^3 - \frac{1}{4} \zeta^4 + \frac{1}{5} \zeta^5 - e_5(\zeta),
\]

with

\[
e_5(\zeta) = \int_0^\zeta (\zeta - t)^5 dt.
\]

Replacing \( \zeta \) with \( \eta - 1 \) we get

\[
\log \eta = (\eta - 1) - \frac{1}{2} (\eta - 1)^2 + \frac{1}{3} (\eta - 1)^3 - \frac{1}{4} (\eta - 1)^4 + \frac{1}{5} (\eta - 1)^5 - e_5(\eta - 1).
\]

Thus the expected value of \( \log \eta \) is given by

\[
E(\log \eta) = \nu_1 - \frac{1}{2} \nu_2 + \frac{1}{3} \nu_3 - \frac{1}{4} \nu_4 + \frac{1}{5} \nu_5 - E \{ e_5(\eta - 1) \},
\]

where \( \nu_k = E \{ (\eta - 1)^k \} \). Since \( E(\eta) = 1, \nu_2 = \frac{1}{2} \), by a direct computation we obtain

\[
\nu_1 = 0, \quad \nu_3 = \frac{2}{L^3}, \quad \nu_4 = \frac{3}{L^4}, \quad \nu_5 = \frac{20}{L^5} + \frac{24}{L^3}.
\]

Thus

\[
E \{ I(\eta) \} = 1 + \frac{1}{2L} + \frac{1}{12L^2} - \frac{5}{2L^3} - \frac{24}{5L^4} + E \{ e_5(\eta - 1) \}.
\]

We observe that

\[
e_5(\zeta) \leq \zeta^5 \int_0^\infty \frac{1}{(1+t)^{1.5}} dt = \frac{5}{12},
\]

Thus

\[
|E \{ e_5(\eta - 1) \}| \leq \frac{1}{5} E \{ (\eta - 1)^5 \} = \frac{4}{3} \frac{24}{5L^4} = O \left( \frac{1}{L^5} \right).
\]

Thus the proposition is proved.

**APPENDIX B**

**Proof of Theorem 1**

We provide a proof of Theorem 1 here.

**Proof:** First we verify the first part. If \( \| z \|_{BV} \to +\infty \), then we have \( J(z) \to +\infty \) or \( \| z \|_{L^1(\Omega)} \to +\infty \) according to the definition of BV-norm.

Case 1: assume \( J(z) \to +\infty \). Since \( S_\varepsilon(x) \geq 1 \) for any \( x \in \Omega \), we have \( L(z) \to +\infty \).

Case 2: assume \( J(z) \to +\infty \) and \( \| z \|_{L^1(\Omega)} \to +\infty \). Define \( \Omega^+ = \{ x \in \Omega : z(x) \geq 0 \}, \Omega^- = \Omega/\Omega^+ \) and

\[
A = \int_{\Omega^+} (f e^{-z} + z - \log f) (x) dx,
\]

\[
B = \int_{\Omega^-} (f e^{-z} + z - \log f) (x) dx.
\]

From section 2 we conclude that \( A \geq 0 \) and \( B \geq 0 \). Since \( \| z \|_{L^1(\Omega)} \to +\infty \), we have

\[
\int_{\Omega^+} z \to +\infty \quad \text{or} \quad \int_{\Omega^-} (-z) \to +\infty.
\]

(i) \( \int_{\Omega^+} z \to +\infty \): it is straightforward to show that \( A \to +\infty \). Thus we have

\[
\int_{\Omega^+} (f e^{-z} + z - \log f) (x) dx \to +\infty,
\]

and hence \( \int_{\Omega^-} S(z(x)) dx \to +\infty \). Define \( \Omega^-_M = \{ x \in \Omega : z(x) \leq -M \} \). By (37) we obtain that \( (f e^{-z} + z) (x) \geq -\alpha z(x) \), for any \( z \in \Omega^-_M \).

Then we have

\[
\int_{\Omega^-_M} (f e^{-z} + z) dx \geq -\alpha \int_{\Omega^-_M} z dx.
\]

Since \( \int_{\Omega^-_M} (-z) \to +\infty \), thus \( \int_{\Omega^-_M} z dx \to +\infty \). By (39) we immediately get \( B \to +\infty \). Thus we have

\[
\int_{\Omega^+} (f e^{-z} + z - \log f) (x) dx \to +\infty,
\]

and hence \( \int_{\Omega^-} S(z(x)) dx \to +\infty \). By combining (i) and (ii) we prove \( L(z) \to +\infty \) in case 2. Therefore, the first part of the theorem holds.

Next the second part is argued. We choose a minimizing sequence \( \{ z_n \} \subset C \). Since \( \{ z_n \} \) must satisfy the constraints in (14), we conclude that \( L(u) \) is bounded. Due to the first part of the proof we conclude that \( \{ z_n \} \) is bounded in \( BV(\Omega) \). Hence, there exists a subsequence \( \{ z_{n_k} \} \) which converges weakly in \( L^2(\Omega) \) to some \( \tilde{z} \in L^2(\Omega) \) [28, theorem 2.6], and \( \{ D z_{n_k} \} \) converges weakly as a measure to \( D \tilde{u} \) [28, lemma 2.5].

Since the function \( J(z) \) is weakly lower semicontinuous with respect to the \( L^2(\Omega) \) topology [28, theorem 2.3], we have

\[
J(\tilde{z}) \leq \liminf_{k \to \infty} J(z_{n_k}) = \inf_{z \in C} J(z).
\]

Since \( C \) is closed and convex, we have \( \tilde{z} \in C \). Therefore, \( \tilde{z} \) is a solution to the problem (14).
APPENDIX C
PROOF OF THEOREM 2

We provide a proof of Theorem 2 here.

Proof: Let \( q_f(s) = f e^{-s} + s - \log f \). Since \( f > 0 \), it is obvious that \( q_f(s) \) is strictly convex. Let \( z_1, z_2 \in BV(\Omega) \) denote two solutions of (14) with \( z_1 \neq z_2 \). Define \( \bar{z} = \frac{1}{2}(z_1 + z_2) \). By the convexity of \( q_f(s) \) we have

\[
q_f(\bar{z}) \leq \frac{1}{2} (q_f(z_1) + q_f(z_2)).
\]

(41)

If the inequality holds as an equality a.e. in \( \Omega \), then by the strictly convexity of \( q_f(s) \) we obtain that \( z_1 = z_2 \) a.e. in \( \Omega \); otherwise there exist \( \delta > 0 \) and \( \Omega_\delta \subseteq \Omega \) with \( |\Omega_\delta| > 0 \) such that

\[
q_f(\bar{z}) \leq \frac{1}{2} (q_f(z_1) + q_f(z_2)) - \delta \; \text{a.e. in } \Omega_\delta.
\]

(42)

Define \( \epsilon_\delta = \epsilon_0|\Omega_\delta| \). According to the formula (15) with \( w(x,y) = w(x,y) \) we get

\[
S(\bar{z}) \leq \frac{1}{2} (S(z_1) + S(z_2)) - \epsilon_\delta \leq 1 + \epsilon \text{ a.e. in } \Omega.
\]

(43)

Define \( z_\theta = \bar{z} \theta \) for \( \theta \in [0,1] \). Since \( S(\cdot) \) is continuous, then for \( \theta \) close to 1, we have \( z_\theta \in C \). Moreover, \( J(z_\theta) = \theta J(\bar{z}) < J(\bar{z}) \) for any \( \theta \in [0,1] \), except \( J(\bar{z}) = 0 \). This implies \( \bar{z} \not\in c \chi_\Omega \) for some \( c \in R \). However, this is impossible since \( c \chi_\Omega \not\in C \). Hence, \( z_1 = z_2 \) a.e. in \( \Omega \).

REFERENCES