A CONVEX AND EXACT APPROACH TO DISCRETE CONSTRAINED TV-L1 IMAGE APPROXIMATION

JING YUAN,*, JUAN SHI,†, AND XUE-CHENG TAI‡

Abstract. We study the TV-L1 image approximation model under the primal and dual perspective, based on the proposed equivalent convex formulations. More specifically, we apply a convex TV-L1 based approach to globally solve the discrete constrained optimization problem of image approximation, where the unknown image function $u(x) \in \{f_1, \ldots, f_n\}$, $\forall x \in \Omega$. We show that the TV-L1 formulation does provide an exact convex relaxation model to the considered nonconvex optimization problem. This result greatly extends the recent results for the CEN model for TV-L1 image processing [10]. It extends the simplest binary constrained case to a general gray-value constrained case, through some rounding off scheme. It also applies to discrete constrained image inpainting. In addition, we construct a fast multiplier-based algorithm based on the proposed primal-dual model, which properly avoids nonsmoothness of the TV-L1 energy functional. Numerical experiments validate the theoretical results and show that the proposed algorithm is both reliable and efficient.

Key words. convex optimization, primal-dual approach, total-variation regularization, image processing

AMS subject classifications. xxx

1. Introduction.

Many tasks of image processing can be formulated and solved successfully by convex optimization models, e.g. image denoising [27, 24], image segmentation [6], image labeling [25, 3] etc. The reduced convex formulations can be studied in a mathematically sound way and usually tackled by a tractable scheme in numerics. Minimizing the total-variation function for such convex image processing formulations is of great importance [27, 30, 22, 6, 23, 7, 20, 19], which preserves edges and sharp features.

In the pioneer work [10], the CEN (Chan-Esedoglu-Nikolova) TV-L1 regularized image approximation model was proposed and it takes the form:

$$\min_u \; \left\{ P(u) := \alpha \int_\Omega |f - u| \, dx + \int_\Omega |\nabla u(x)| \, dx \right\}, \tag{1.1}$$

This model was first introduced and studied by Alliney [2, 1] for discrete one-dimensional signals' denoising. Chan et al. showed an interesting property of the TV-L1 model (1.1): for the input binary image $f(x) \in \{0, 1\}$, there exists at least one global optimum $u(x) \in \{0, 1\}$. It follows that the convex TV-L1 formulation (1.1) actually solves the nonconvex optimization problem:

$$\min_{u(x) \in \{0,1\}} \; \alpha \int_\Omega |f - u| \, dx + \int_\Omega |\nabla u(x)| \, dx \,, \tag{1.2}$$

globally and exactly! Hence (1.1) provides an exact convex relaxation of the binary constrained optimization problem (1.2). The authors further proved that rounding the computed result of (1.1) by any value $\gamma \in [0, 1]$ may give a series of global optimums of the binary constrained optimization model (1.2).

*Jing Yuan, Computer Science Department, Middlesex College, University of Western Ontario, London, Ontario, N6A 5B7 (cn.yuanjing@gmail.com)

†Juan Shi, Division of Mathematical Sciences, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore. (shij0004@e.ntu.edu.sg)

‡Xue-Cheng Tai, Division of Mathematical Sciences, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore and Department of Mathematics, University of Bergen, Norway. (tai@mi.uib.no)

Previous Works and Motivations. With the help of coarea formula, Chan et al. [9, 10] proved that the energy functional $P(u)$ of (1.1) can be equivalently represented in terms of the upper level-set sequence of the image functions $u(x)$ and $f(x)$, i.e.

$$P(u) = \int_{-\infty}^{+\infty} \left\{ |\partial U^\gamma| + \alpha \, |U^\gamma \triangle F^\gamma| \right\} d\gamma \,, \tag{1.3}$$

where for each γ, U^γ and F^γ give the $\gamma-$upper level set of the unknown $u(x)$ and the input $f(x)$ respectively, such that

$$U^\gamma(x) = \begin{cases} 1 \,, & \text{when } u(x) > \gamma \\ 0 \,, & \text{when } u(x) \leq \gamma \end{cases} , \quad x \in \Omega \,, \quad i = 1, \ldots, n \,; \tag{1.4}$$

$|\partial U^\gamma|$ gives the perimeter of U^γ and $|U^\gamma \triangle F^\gamma|$ denotes the area of the symmetric difference of the two level sets.

Yin et al. [34] pointed out that minimizing such a layer-wise energy function (1.3) actually amounts to properly stacking the optimal U^γs, which corresponds to solving (1.2) for each given binary indicator function of F^γ. In other words, solving (1.1) can be reduced to optimizing a sequence of binary constrained problems as (1.2). Since $U^{\gamma_1} \subset U^{\gamma_2}$ when $\gamma_1 \geq \gamma_2$, the process recovers the optimum $u^*(x)$ of (1.1) by properly arranging all the associated level sets U^γ, $\gamma \in (-\infty, +\infty)$. The same result was also discovered by Darbon et al. [11, 12] in an image graph setting where the anisotropic total-variation term is considered and a fast graph-cut based algorithm was introduced.

However, as stated in [34], such approach means both bad and good news for processing the gray-scale image in practice: on the one hand, the total number of gray values is finite, i.e. $u(x) \in \{0, \ldots, 255\}$, hence only a finite number of optimization problems (1.2) should be considered; on the other hand, solving (1.2) for each layer F^γ is not trivial; and in order to globally tackle (1.1), one has to examine a large number of obtained level-sets to restrict its search legally. This makes Yin's results [34] impratical to a real image processing task.

In addition, the PDE-descent method is often used to numerically approximate the global optimum of (1.1) [10, 9, 34, 13], which smoothes the total-variation term by $\sqrt{\partial_x u^2 + \partial_y u^2 + \epsilon^2}$. Actually, even if ϵ takes a small enough value, the coarea formula is no longer satisfied. As a matter of fact, new gray levels appear and the indicator functions are blurred.

Motivated by the above observations, we introduce the primal and dual perspective of the TV-L1 model (1.1) and study the exactness of (1.1) as the convex relaxation of the discrete constrained optimization problem:

$$\min_{u(x) \in \{f_1, \ldots, f_n\}} \alpha \int_\Omega |f - u| \; dx + \int_\Omega |\nabla u(x)| \; dx \,, \tag{1.5}$$

given $f(x) \in \{f_1, \ldots, f_n\}$. In this paper, we assume the gray-scale values f_i, $i = 1, \ldots, n$, are ascent ordered by $f_1 < \ldots < f_n$. Clearly, integers $0, \ldots, 255$ for 8-bits images is one of the common options in most cases. Compared with the original CEN model [10], which can only handle binary images, this extension will greatly increase the application scope of proposed model.

Contributions. Our main contributions can be generalized as follows:

1. We propose equivalent formulations in terms of primal and dual, and build up a new analytical framework which results in a new variational perspective of (1.1).
2. By the proposed equivalent formulations, we show that the TV-L1 formulation (1.1) can be used as the convex relaxed model for its discrete constrained non-convex image processing model. This extends the CEN model [10] to more general cases. To the best of the authors' knowledge, this is new. The same theoretical results can be naturally extended to image inpainting.
3. In term of numerical computation, we show that the discrete constrained optimization problem (1.5) can be exactly optimized by the convex programming problem (1.1) and a proposed rounding scheme. Besides its simplicity, it largely reduces the computational and memory cost. For graph-cut based approaches e.g. [11, 21, 5, 4, 31], the reduction is especially remarkable when the total number of gray-scales is large.
4. We introduce an elegant multiplier-based algorithm which explores the equivalent primal-dual formualtion through two simple projection substeps, instead of tackling the highly nonsmooth TV-L1 energy functional directly. Its reliability and efficience are verified by standard optimization theories and different experiments.

In parallel to our multiplier-based method, several other dual formulations and algorithmic schemes were proposed recently in the literature, see [16, 33, 32, 35, 28, 29, 13]. In contrast to [16, 33, 32, 35], we apply the proposed equivalent primal-dual and dual formulations as a complete approach, including both variational analyses and algorithms to study topics concerning (1.1), not just derive the algorithmic scheme. In addition, the primal-dual algorithm we have proposed here is different from [16, 33, 32, 35]. In our algorithm, the solution u is treated as the multiplier. This seems to be a new idea.

2. Equivalent Models. We call the TV-L1 image approximation (1.1) *primal model* in this paper, as comparison to the equivalent models introduced in this section.

2.1. Equivalent Primal-Dual Model. With the help of conjugates [26], the data term of (1.1) can be equally expressed by

$$\alpha \int_\Omega |f-u| \,=\, \max_{q\in S_\alpha} \langle q, f-u\rangle \,, \quad S_\alpha \,:=\, \{q \mid |q(x)| \le \alpha \,,\ \forall x\in\Omega\,\} \,. \tag{2.1}$$

Moreover, it is well known that the total-variation term of (1.1) can also be formulated [18] as follows

$$\int_\Omega |\nabla u|\; dx \,=\, \max_{p\in C_1} \langle \operatorname{div} p, u\rangle \,, \quad C_1 \,:=\, \{p \,|\, p\in C_c^1(\Omega,\mathbb{R}^2)\,,\ |p(x)| \le 1\,,\ \forall x\in\Omega\,\} \,. \tag{2.2}$$

In view of (2.1) and (2.2), after some rearrangements, the TV-L1 approximation formulation (1.1) can be rewritten as

$$\max_{q\in S_\alpha}\max_{p\in C_1}\min_u \,\big\{E(u;q,p) \,:=\, \langle q, f\rangle + \langle \operatorname{div} p - q, u\rangle\,\big\} \,, \tag{2.3}$$

which is called the equivalent *primal-dual model* to the primal model (1.1).

2.2. Equivalent Dual Model. Observe that u is unconstrained, minimizing (2.3) over u, therefore, leads to the linear equality

$$\operatorname{div} p \,=\, q \,,$$

and the constrained maximization problem

$$\max_{q \in S_\alpha} \max_{p \in C_1} \; \big\{ D(q,p) \, := \, \langle q, f \rangle \, \big\} \,, \quad \text{s.t.} \quad \operatorname{div} p = q \,. \tag{2.4}$$

Likewise, we call (2.4) the equivalent *dual model* to (1.1).

2.3. Optimization Facts. For the primal-dual formulation (2.3), the conditions of the minimax theorem (see e.g., [15, 17]) are all satisfied. That is: the constraints of dual variables p and q are convex and the energy function is linear to both u and (p, q), hence convex l.s.c. for fixed u and concave u.s.c. for fixed p and q. This follows that there exists at least one saddle point, see [15, 17]. As a consequence, the min and max operators of the primal-dual model (2.3) can be interchanged, i.e.

$$\max_{q \in S_\alpha} \max_{p \in C_1} \big\{ \min_u \; E(u; q, p) \big\} \; = \; \min_u \big\{ \max_{q \in S_\alpha} \max_{p \in C_1} \; E(u; q, p) \big\} \,. \tag{2.5}$$

It is easy to see that the optimization of the primal-dual model (2.3) over the dual variables q and p react on the primal formulation (1.1) of TV-L1 image approximation, i.e. the right hand side of (2.5):

$$P(u) \; = \; E(u; q^*, p^*) \; = \; \max_{q \in S_\alpha} \max_{p \in C_1} \; E(u; q, p) \,.$$

Likewise, the dual model (2.4) can be achieved by optimizing the image function $u(x)$ in (2.3), i.e. the left hand side of (2.5):

$$D(q, p) \; = \; E(u^*, q, p) \; = \; \min_u \; E(u; q, p) \,. \tag{2.6}$$

2.4. Image Inpainting. In this paper, we consider image inpainting in a similar manner. We formulate its convex primal model as

$$\min_u \; \big\{ P(u) \, := \, \alpha \int_\Pi |f - u| \; dx \, + \, \int_\Omega |\nabla u(x)| \; dx \big\} \,, \tag{2.7}$$

where $\Pi \subset \Omega$ denotes the unmasked area.

The concerned equivalent primal-dual and dual models to (2.7) can be developed in a similar way as (2.3) and (2.4). Here we only list its equivalent primal-dual model:

$$\max_{q \in S_\alpha} \max_{p \in C_1} \min_u \; \big\{ E(u; q, p) \, := \, \langle q, f \rangle_\Pi + \langle \operatorname{div} p - q \sigma_\Pi, u \rangle_\Omega \, \big\} \,, \tag{2.8}$$

where σ_Π is the indicator function of the unmasked domain Π.

3. Global and Exact Optimums. In this section, we study the nonconvex optimization problem (1.5) and show that the TV-L1 formulation (1.1), which is a convex relaxed model of (1.5), solves the nonconvex minimization problems (1.5) globally and exactly through a rounding scheme. We state our results and proof in several propositions.

PROPOSITION 3.1 (Extremum Principle). *Given the image function $f(x) \in \{f_1, \ldots, f_n\}$, $\forall x \in \Omega$, along with ordering $f_1 < \ldots < f_n$, each optimum $u^*(x)$ of* (1.1) *suffices $f_1 \leq u^*(x) \leq f_n$ almost everywhere.*

The same results which state any optimum $u^*(x)$ should suffice $u^*(x) \in [f_{\min}, f_{\max}]$, i.e. $u^*(x) \in [f_1, f_n]$ considering the ascent ordering $f_1 < \ldots < f_n$ in this work, can also be found in other works, e.g. [9] where $f(x) \in \{0, 1\}$ or [14] where $f(x) \in [f_{\min}, f_{\max}]$. We also provide the proof as follows to ease readability.

Proof. Let u^* be the minimum of (1.1). Due to the convexity of (1.1), u^* is simply accepted as the global minimum. We first prove that $u^*(x) \leq f_n$ for $\forall x \in \Omega$.

If $u^*(x) > f_n$ at some area $\tilde{\Omega} \subset \Omega$, then we define the function u' which just threshholds the value $u^*(x)$ to be not larger than f_n, i.e.

$$u'(x) \,=\, \begin{cases} f_n & \text{at } x \in \tilde{\Omega} \\ u^*(x) & \text{at } x \in \Omega \backslash \tilde{\Omega} \end{cases} .$$

Obviously, in view of $f(x) \leq f_n$ and $u^*(x) > f_n$ for $\forall x \in \tilde{\Omega}$, we have

$$\begin{aligned} \int_\Omega |u^* - f| \; dx \,&=\, \int_{\Omega \backslash \tilde{\Omega}} |u^* - f| \; dx \,+\, \Big\{ \int_{\tilde{\Omega}} |f_n - f| \; dx \,+\, \int_{\tilde{\Omega}} |u^* - f_n| \; dx \Big\} \\ &=\, \int_\Omega |u' - f| \; dx \,+\, \int_{\tilde{\Omega}} |f_n - f| \; dx \,. \end{aligned}$$

It follows that

$$\int_\Omega |f - u'| \; dx \,<\, \int_\Omega |f - u^*| \; dx \,. \tag{3.1}$$

By the coarea formula of the total variation term:

$$\mathrm{TV}(u) \,=\, \int_{-\infty}^{+\infty} L_\gamma(u) \, d\gamma \,,$$

where $L_\gamma(u)$ is the length of the $\gamma-$upper level set of u, it follows that

$$\mathrm{TV}(u') \,<\, \mathrm{TV}(u^*) \,, \tag{3.2}$$

because the f_n-upper level set of u' is threshholded to vanish.

Observe (3.1) and (3.2), we must have

$$\int_\Omega |f - u'| \; dx \,+\, \alpha \mathrm{TV}(u') \,<\, \int_\Omega |f - u^*| \; dx \,+\, \alpha \mathrm{TV}(u^*) \,.$$

This is in contradiction to the fact that u^* is the global minimum of (1.1).

Likewise, we can also prove $u^*(x) \geq f_1$ $x \in \Omega$ in the same way. In consequence, we prove that each minimum $u^*(x)$ of (1.1) suffices $u^*(x) \in [f_1, f_n]$. $\square$

PROPOSITION 3.2. *Given a bounded scalar function $f_1 \leq u(x) \leq f_n$ $\forall x \in \Omega$, if an optimal vector field p^* maximizes the integral $\int_\Omega u \operatorname{div} p \, dx$ over the convex set C_1, i.e.*

$$\int_\Omega |\nabla u| \, dx \,=\, \int_\Omega u \operatorname{div} p^* \, dx \,,$$

then in view of (1.4)*, for every $\gamma-$upper level set $U^\gamma(x)$ of $u(x)$ with $\gamma \in [f_1, f_m)$, p^* also maximizes the integral $\int_\Omega U^\gamma \operatorname{div} p \, dx$ over the convex set C_1 and*

$$\int_\Omega U^\gamma \operatorname{div} p^* \, dx \,=\, |\partial U^\gamma| \,,$$

which is the perimeter of the level set $U^\gamma(x)$.

Proof. Denote the interval $\Gamma = [f_1, f_n]$. The coarea formula gives

$$\int_\Omega |\nabla u|\, dx = \int_\Gamma \int_\Omega |\nabla U^\gamma|\, dx\, d\gamma. \tag{3.3}$$

By applying this formula we can deduce

$$\int_\Omega u \,\operatorname{div} p^*\, dx = \int_\Omega |\nabla u|\, dx = \int_\Gamma \int_\Omega |\nabla U^\gamma|\, dx\, d\gamma = \int_\Gamma \left(\max_{p\in C_1} \int_\Omega U^\gamma \operatorname{div} p\, dx \right) d\gamma. \tag{3.4}$$

By the fact that $u(x) = \int_{f_1}^{u(x)} \, d\gamma = \int_\Gamma U^\gamma(x) d\gamma$ for any $x \in \Omega$, we have

$$\int_\Omega u \,\operatorname{div} p^*\, dx = \int_\Omega \left(\int_\Gamma U^\gamma(x) d\gamma \right) \operatorname{div} p^*(x)\, dx = \int_\Gamma \int_\Omega U^\gamma \operatorname{div} p^*\, dx\, d\gamma. \tag{3.5}$$

Therefore, combining (3.4) and (3.5):

$$\int_\Gamma \int_\Omega U^\gamma \operatorname{div} p^*\, dx d\gamma \;=\; \int_\Gamma \left(\max_{p\in C_1} \int_\Omega U^\gamma \operatorname{div} p\, dx \right) d\gamma. \tag{3.6}$$

This equality (3.6) together with the fact that for any $\gamma \in [f_1, f_n)$

$$\int_\Omega U^\gamma \operatorname{div} p^*\, dx \le \max_{p\in C_1} \int_\Omega U^\gamma \operatorname{div} p\, dx\,. \tag{3.7}$$

Then it follows that

$$\int_\Omega U^\gamma \operatorname{div} p^*\, dx = \max_{p\in C_1} \int_\Omega U^\gamma \operatorname{div} p\, dx$$

for almost every $\gamma \in [f_1, f_n)$. Clearly, the perimeter of the level set U^γ is given by

$$|\partial U^\gamma| = \int_\Omega |\nabla U^\gamma|\, dx = \max_{p\in C_1} \int_\Omega U^\gamma \operatorname{div} p\, dx\,.$$

$\square$

COROLLARY 3.3. *Given a bounded scalar function $f_1 \le u(x) \le f_n$ $\forall x \in \Omega$ and $n-1$ different values γ_i, $i = 1, \dots, n-1$, such that $f_1 \le \gamma_1 < \dots < \gamma_{n-1} \le f_n$, if an optimal vector field p^* maximizes the integral $\int_\Omega u \operatorname{div} p\, dx$ over the convex set C_1, then for the image function*

$$u^\gamma(x) \;=\; \sum_{i=1}^{n-1} (f_{i+1} - f_i)\, U^{\gamma_i}(x)\,,$$

p^* also maximizes the integral $\int_\Omega u^\gamma \operatorname{div} p\, dx$ over the convex set C_1, i.e. we have

$$\int_\Omega |\nabla u^\gamma|\, dx \;=\; \int_\Omega u^\gamma \operatorname{div} p^*\, dx\,.$$

Proof. By virtue of Prop. 3.2, p^* also maximize the integral

$$\int_\Omega U^{\gamma_i} \operatorname{div} p\, dx,$$

over the convex set C_1 for each γ_i, $i = 1, \dots, n-1$.

Then it follows that for the piecewise constant image function

$$u^\gamma(x) \;=\; \sum_{i=1}^{n-1} (f_{i+1} - f_i)\, U^{\gamma_i}(x)\,,$$

p^* also maximizes the integral

$$\int_\Omega u^\gamma \operatorname{div} p\, dx \;=\; \sum_{i=1}^{n-1} \big\{ (f_{i+1} - f_i) \int_\Omega U^{\gamma_i} \operatorname{div} p\, dx \big\},$$

over the convex set $p \in C_1$, because $f_1 < \ldots < f_n$ is ordered such that

$$f_{i+1} - f_i \;>\; 0\,, \quad i = 1, \dots, n-1\,.$$

Therefore, we have

$$\int_\Omega |\nabla u^\gamma|\, dx \;=\; \int_\Omega u^\gamma \operatorname{div} p^*\, dx\,.$$

□

With helps of the above facts, we can prove the following proposition:

Proposition 3.4. *Given the image function $f(x) \in \{f_1, \dots, f_n\}$, where $f_1 < \ldots < f_n$ and the boundary of each concerning upper level set $F^{f_i}(x)$, $i = 1, \dots, n$, is regular, then for any given $n-1$ values γ_i, $i = 1, \dots, n-1$ such that*

$$f_1 < \gamma_1 < f_2 < \ldots < \gamma_{n-1} < f_n\,, \tag{3.8}$$

we define the image function $u^\gamma(x)$ by the $n-1$ upper level sets (1.4) *of the computed optimum $u^*(x)$ of* (1.1)*:*

$$u^\gamma(x) \;=\; f_1 + \sum_{i=1}^{n-1} (f_{i+1} - f_i)\, U^{\gamma_i}(x)\,. \tag{3.9}$$

Then $u^\gamma(x) \in \{f_1, \dots, f_n\}$ and $u^\gamma(x)$ gives an exact global optimum of (1.5).

Proof. Let (u^*, q^*, p^*) be the optimal primal-dual pair of (2.3). Hence the optimal dual variables q^* and p^* suffice that q^* maximizes the integral $\int_\Omega q(f-u)\, dx$ over the convex set S_α and p^* maximizes the integral $\int_\Omega u \operatorname{div} p\, dx$ over the convex set C_1.

$u^\gamma(x) \in \{f_1, \dots, f_n\}$ as (3.9) can be rearranged as

$$u^\gamma(x) \;=\; f_1\, (1 - U^{\gamma_1}(x)) + \sum_{i=2}^{n-1} f_i\, (U^{\gamma_{i-1}}(x) - U^{\gamma_i}(x)) + f_n\, U^{\gamma_{n-1}}(x)\,.$$

u^γ is also a global optimum of (1.1), which can be obtained by considering the following facts: By Coro. 3.3, p^* also maximizes the integral $\int_\Omega u^\gamma \operatorname{div} p\, dx$ over the convex set C_1 and

$$\int_\Omega |\nabla u^\gamma|\, dx \;=\; \langle u^\gamma, \operatorname{div} p^* \rangle\,. \tag{3.10}$$

At the next step, we can prove

$$\alpha \int_\Omega |f - u^\gamma|\, dx \;=\; \langle q^*, f - u^\gamma \rangle\,. \tag{3.11}$$

The optimal dual variable $q^*(x)$ actually gives the sign of $f(x) - u^*(x)$ at each $x \in \Omega$, when $f(x) \neq u^*(x)$; when $f(x) = u^*(x)$, $q^*(x)$ can take any value in $[-\alpha, \alpha]$. Now we assume $u^*(x) \in [f_k, f_{k+1}]$ for the position $x \in \Omega$, then in view of (1.4) and (3.9), we have

$$u^*(x) \in [f_k, \gamma_k] \Longrightarrow u^\gamma(x) = f_k \,,$$

and

$$u^*(x) \in (\gamma_k, f_{k+1}] \Longrightarrow u^\gamma(x) = f_{k+1} \,.$$

Since $f(x) \in \{f_1, \ldots, f_n\}$, we can analyze $q^*(x)$ in two cases: $f(x) \leq f_k$ and $f(x) \geq f_{k+1}$.

- When $f(x) \leq f_k$, in view of $u^*(x) \geq f_k$, we have $q^*(x) = -\alpha$ for $u^*(x) > f_k$ or $q^*(x) \geq -\alpha$ for $u^*(x) = f(x)$ in order to maximize $q(x) \cdot (f(x) - u^*(x))$ over $q(x) \in [-\alpha, \alpha]$. Then in both cases, $q^*(x)$ also maximizes the product $q(x) \cdot (f(x) - f_k)$, or $q(x) \cdot (f(x) - f_{k+1})$, over $q(x) \in [-\alpha, \alpha]$. Hence $q^*(x)$ maximizes $q(x) \cdot (f(x) - u^\gamma(x))$ over $q(x) \in [-\alpha, \alpha]$.
- When $f(x) \geq f_{k+1}$, in view of $u^*(x) \leq f_{k+1}$, we have $q^*(x) = \alpha$ for $u^*(x) < f_{k+1}$ or $q^*(x) \leq \alpha$ for $u^*(x) = f(x)$ in order to maximize $q(x) \cdot (f(x) - u^*(x))$ $q(x) \in [-\alpha, \alpha]$. In both cases, $q^*(x)$ also maximizes the product $q(x) \cdot (f(x) - f_k)$ or $q(x) \cdot (f(x) - f_{k+1})$, over $q(x) \in [-\alpha, \alpha]$. Hence $q^*(x)$ maximizes $q(x) \cdot (f(x) - u^\gamma(x))$ over $q(x) \in [-\alpha, \alpha]$.

Therefore, we have q^* maximizes the integral $\langle q, f - u^\gamma \rangle$ over the convex set S_α. Then the fact (3.11) is proved.

By virtue of (3.10), (3.11) and the dual model (2.4), we have

$$P(u^\gamma) \,=\, E(u^\gamma, p^*, q^*) \,=\, \langle q^*, f \rangle + \langle u^\gamma, \operatorname{div} p^* - q^* \rangle \,=\, \langle q^*, f \rangle \,=\, P(u^*)\,.$$

Then it follows that u^γ is also a global minimum of (1.1) as u^* is global minimum of (1.1). Since (1.1) is just the relaxed version of (1.5), $u^\gamma(x) \in \{f_1, \ldots, f_n\}$ solves (1.5) exactly and globally. □

The proposed rounding scheme (3.9) actually gives

$$u^\gamma(x) \,=\, \begin{cases} f_1 \,, & \text{when } u^*(x) < \gamma_1 \\ f_i \,, & \text{when } \gamma_{i-1} \leq u^*(x) < \gamma_i, \; i = 2, \ldots, n-1 \\ f_n \,, & \text{when } u^*(x) \geq \gamma_{n-1} \end{cases} .$$

In the experiments of this paper, we adopt the above scheme to obtain rounding results.

3.1. Global and Exact Optimums of Image Inpainting. For the problem of image inpainting, we can prove the same result as Prop. 3.4, such that the image function $u^\gamma(x)$ given by the $n-1$ upper level sets of the optimum $u^*(x)$ of the convex image inpainting model (2.7), i.e.

$$u^\gamma(x) \,=\, f_1 + \sum_{i=1}^{n-1} (f_{i+1} - f_i)\, U^{\gamma_i}(x)\,, \tag{3.12}$$

solves the discrete constrained image inpainting problem

$$\min_{u(x) \in \{f_1, \ldots, f_n\}} \; \alpha \int_\Pi |f - u| \, dx \,+\, \int_\Omega |\nabla u(x)| \, dx\,, \tag{3.13}$$

globally and exactly.

PROPOSITION 3.5. *Given the image function $f(x) \in \{f_1, \dots, f_n\}$ at the unmasked area Π, where $f_1 < \dots < f_n$ and the boundary of each concerning upper level set $F^{f_i}(x)$, $i = 1, \dots, n$, is regular, then for any given $n-1$ values γ_i, $i = 1, \dots, n-1$ such that*

$$f_1 < \gamma_1 < f_2 < \ldots < \gamma_{n-1} < f_n \,, \tag{3.14}$$

we define the image function $u^\gamma(x)$, as follows, by the $n-1$ upper level sets (1.4) *of the computed optimum $u^*(x)$ of* (2.7)*:*

$$u^\gamma(x) \;=\; f_1 + \sum_{i=1}^{n-1} (f_{i+1} - f_i)\, U^{\gamma_i}(x) \,. \tag{3.15}$$

We have $u^\gamma(x) \in \{f_1, \dots, f_n\}$ and $u^\gamma(x)$ gives an exact global optimum of (1.5)*.* The proof just repeats Prop. 3.1, Prop. 3.2, Coro. 3.3 and Prop. 3.4 and is omitted here.

4. Multiplier-Based Algorithm. In this paper, we build up the algorithm upon the equivalent primal-dual model (2.3). Clearly, the primal variable u works as the multiplier in (2.3) for the linear equality $\operatorname{div} p - q = 0$. The energy function of (2.3) just gives the corresponding Lagrangian function to the dual formulation (2.4). By these observations, we define its augmented Lagrangian function as

$$L_c(q, p, u) \;=\; \langle q, f \rangle + \langle \operatorname{div} p - q, u \rangle - \frac{c}{2} \left\| \operatorname{div} p - q \right\|^2$$

where $c > 0$.

Thereafter, the classical augmented Lagrangian algorithm can be applied, which gives a splitting optimization framework over each dual variables q and p respectively, by exploring projections to corresponding convex sets. To this end, we call Alg. 1 the *multiplier-based algorithm.* It explores two simple projection sub-steps: (4.1) and (4.2) at each iteration, which properly avoids tackling the nonsmooth terms in (1.1) in a direct way. The projection in (4.1) is easy and cheap to compute. For projection (4.2), we can use one or a few steps of the iterative algorithm in [8]. We can also use one or a few steps of the following projected gradient decent to approximately solve (4.2) with the time step τ properly chosen:

$$p^{k+1} \;=\; \mathrm{Proj}_{\|p\|_\infty \le \lambda}(p^k + \tau \mathrm{grad}(\operatorname{div} p^k - (q^{k+1} + u^k/c)).$$

For image inpainting, its concerned augmented lagrangian function can be formulated upon (2.8). A similar algorithm as Alg. 1 can be constructed.

5. Experiments. The experiments in this work are designed in two parts: we evaluate both the theoretical results and efficiency of the proposed algorithm in terms of iterations in the first part; experiments of practical impulsive denoising and image inpainting are performed in the second part. For the implementation of algorithms, All the codes are developed on Matlab.

In the experiments, convergence is evaluated by the primal-dual gap:

$$\mathrm{err} \;=\; c \left\| \operatorname{div} p - q \right\| / \left\| u \right\| \,,$$

which gives the ratio of the primal-dual gap to the norm of image approximation $u(x)$, see (4.3).

Algorithm 1 Multiplier-Based Algorithm

- Set the starting values: q^0, p^0 and u^0, and let $k = 1$;
- Start the $k-$th iteration which includes two successive sub-steps:
  1. Optimize q^{k+1} by fixing p^k and u^k:

$$q^{k+1} := \arg \max_{\|q\|_\infty \le 1} L_c(q, p^k, u^k) = \arg \max_{\|q\|_\infty \le 1} \langle q, f \rangle - \frac{c}{2} \left\| q - (\operatorname{div} p^k - u^k/c) \right\|^2 ,$$

which is approximated by the projection

$$q^{k+1} = \operatorname{Proj}_{\|q\|_\infty \le 1}(f/c + (\operatorname{div} p^k - u^k/c)) ; \tag{4.1}$$

  2. Optimize p^{k+1} by fixing q^{k+1} and u^k:

$$p^{k+1} := \arg \min_{p \in C^\lambda} \frac{1}{2} \left\| \operatorname{div} p - (q^{k+1} + u^k/c) \right\|^2 , \tag{4.2}$$

which is the projection of $(q^{k+1} + u^k/c)$ to the convex set $\operatorname{div} C_\lambda$.

- Update u^{k+1} by

$$u^{k+1} = u^k + c\,(q^{k+1} - \operatorname{div} p^{k+1}) ; \tag{4.3}$$

and let $k = k + 1$, repeat untill convergence.

To evaluate the performance of rounded results in the following experiments, we take the energy difference associated to the computated optimum u^* and the rounded result u^γ which is measured by the ratio:

$$\text{ratio} = |P(u^*) - P(u^\gamma)| \,/ P(u^*) .$$

For the comparisons to other state of art methods, the Peak Signal to Noise Ratio (PSNR) between the ground truth and the outputs, i.e.

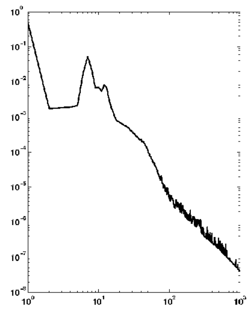
$$\text{PSNR}(u, v) = 10\log_{10} \frac{255^2}{\frac{1}{MN} \sum_{i,j} (u_{i,j} - v_{i,j})^2}$$

is measured, where $u_{i,j}$ and $v_{i,j}$ denote the pixel values of initial ground truth images and denoised images respectively.

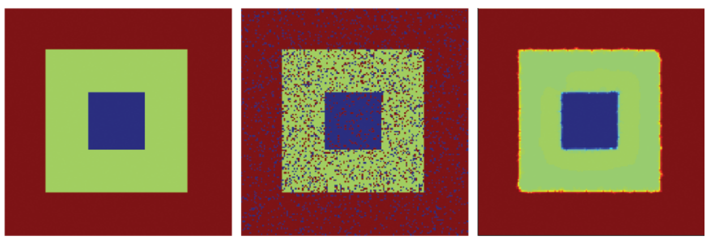
5.1. Validation and Convergence.

5.1.1. Synthetic Image. A synthetic image $f(x) \in \{0, 0.5, 1\}$ (see figure (c) of Fig. 5.1), which is colorized by red: 0, green: 0.5, blue: 1, is taken for the validation of Prop. 3.4. We set the penalty parameter $\alpha = 1$ and the augmented parameter $c = 6$. In fact, the Chambolle-projection step (4.2) is only explored by $3 - 5$ times at each iteration. The experiment still shows a fast convergence rate (see figure (a)): the algorithm runs for 1000 iterations and converges at err $\le 3 \times 10^{-7}$.

In this experiment, two rounding schemes are taken: $\{\gamma_1 = 0.25, \gamma_2 = 0.75\}$; $\{\gamma_1 = 0.35, \gamma_2 = 0.65\}$. For the computed result u^*, it gives the energy $P(u^*) = 2938.7$. The two corresponding rounded results produce the energy $P(u^\gamma) = 2937.7$, 2937.7, i.e. both rounding schemes give the same energy as the convex relaxed energy $P(u^*)$! Both energy ratios are zero!



(a)



(b) (c) (d)

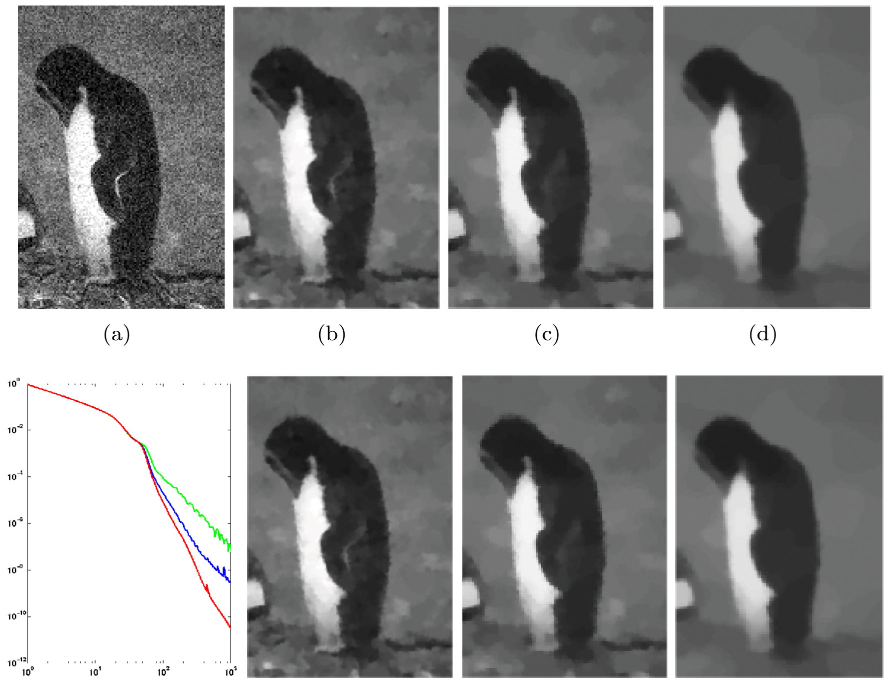
(e) (f) (g)

FIG. 5.1. *(a) convergence rate (600 iterations); (b) the ground truth image colorized by red:* 0, *green:* 0.5, *blue:* 1; *figure (c) the input image $f(x)$; (d) the computed image $u^*(x)$ where $\alpha = 1$; (e) the image u^γ rounded by $\{\gamma_1 = 0.25,\ \gamma_2 = 0.75\}$; (f) the image u^γ rounded by $\{\gamma_1 = 0.35,\ \gamma_2 = 0.65\}$; (g) the difference between two rounded results.*

5.1.2. Gray Value Images. For the given gray-value images $f(x)$ of the experiments, 256 gray-scale levels are naturally encoded, i.e. $f(x) \in \{0, \ldots, 255\}$.

The experiment results given in Fig. 5.2 show the denoising of the penguin image (see figure (a) of Fig. 5.2), which is downloaded from the middleburry data set: http://vision.middlebury.edu/MRF. The rounding scheme is simply taken by $\gamma = \{0.5, 1.5, \ldots, 255\}$, i.e. it just gives the nearest integer. For the following experiments where $\alpha = 1.3,\ 1,\ 0.5$, the Chambolle-projection step (4.2) is explored by 3 times at each iteration. Within 1000 iterations, the Algorithm 1 converges to an error below (see figure (e) of Fig. 5.2): 4×10^{-11} (red line, for $\alpha = 1.3$), 5×10^{-9} (blue line, for $\alpha = 1$), 7×10^{-8} (green line, for $\alpha = 0.5$).

The energy differences associated to the computated optimum u^* and the rounded result u^γ for the three experiments are nearly zero in numerics.



(a) (b) (c) (d)

(e) (f) (g) (h)

FIG. 5.2. *(a) the input image $f(x)$; figure (b) - (d) show the computation results when $\alpha = 1.3$, 1, 0.5 respectively; (b) plot of convergences (1000 iterations): red line: $\alpha = 1.3$, blue line: $\alpha = 1$ and green line: $\alpha = 0.5$; figure (f) - (h) show the rounding results when $\alpha = 1.3$, 1, 0.5 respectively.*

The images processed in the experiments, shown in Fig. 5.3, are downloaded from the Berkeley segmentation dataset and benchmark. For all the experiments, we set $\alpha = 0.5$ and the experiment results show the ratios of energy differences are nearly zero!

5.2. Applications and Comparisons. In this section, we apply the propsoed algorithm to some real applications: impulsive image denoising and image inpainting. In addition, we will also show comparisons to the method proposed recently by [33].

5.2.1. Impulsive Denoising. For restoration of real images corrupted by impulsive (Salt and Pepper) noises, we first make the experiment shown by Fig. 5.4, where a Dragonfly image (which has thin and elongated details) is taken for image denoising: see the image without noise (figure (b)) and the noisy image (figure (a)) where the Salt and Pepper noise with level 5% has been added. For different choice of α which trades off the balance of keeping image details and extracting small-scale structures, e.g. noises. We achieve restoration images with slight differences as shown in figure (c)-(e) of Fig. 5.4. Visually, the best result is computed by setting $\alpha = 2$. The difference between the input image $f(x)$ and the restored image, given by figure (f) of Fig. 5.4, also demostrates that detail losts of the image is very small.

In addition, we verify the performance of our method by several experiments with comparisons to the scheme proposed in [33]. Both schemes involve a substep where Chambolle-projections need to be performed iteratively. We compare our restoration results with [33], with impulsive noise levels ranging from 10% to 50%. As the com-

FIG. 5.3. *Four input images are shown in* **the first row***; the computed images $u^*(x)$ are given in* **the 2nd row** *respectively; the rounded images $u^*(x)$ are shown in* **the 3rd row** *respectively. In all experiments, we set $\alpha = 0.5$.*

parison results given by Tab. 5.1, the restored images computed by our proposed algorithm are better than [33] for experiments with low noise level; for the cases of high noise level, our method still keeps higher PSNR values, i.e. more image details.

Noise Level	Our Approach	ALM of [33]
10%	$39.61dB$	$35.04dB$
20%	$37.56dB$	$34.71dB$
30%	$35.65dB$	$33.90dB$
40%	$34.26dB$	$33.35dB$
50%	$33.58dB$	$32.83dB$

TABLE 5.1
Comparison results by PSNR for the experiment (Fig. 5.5)

For the experiment of Fig. 5.6, we try high noise levels ranging from 50% to 80%. Results show that our approach still get reasonable resutls, as shown in Fig. 5.6.

5.2.2. Image Inpainting. For the experiments of image inpainting, the masked areas of the input image are marked by red (see images in the first row of Fig. 5.7). Totally discrete optimums with 256 gray-scales can be found by the proposed approach, i.e. masked areas are properly recovered such that $u^*(x) \in \{0, 1, ..., 255\}$.

6. Conclusions. This work studies the discrete constrained TV-L1 based image approximation, with applications to image denoising and inpainting. We prove that

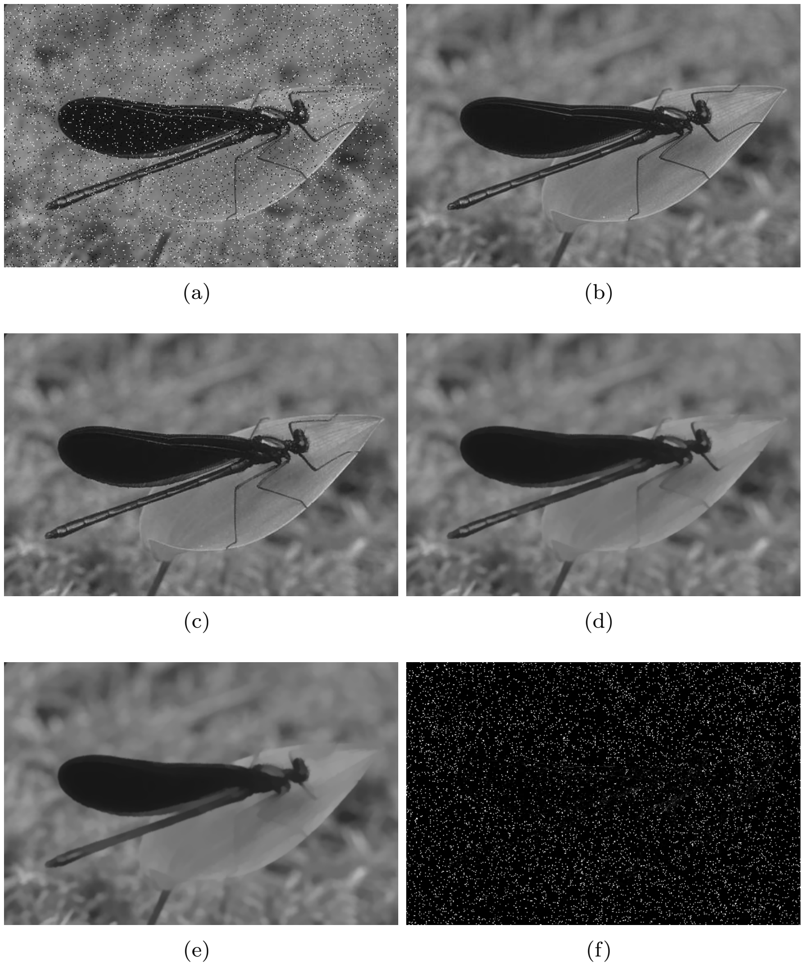


Fig. 5.4. *(a)noisy image noise level 5%,(b)ground truth,(c)restorated image with $\alpha = 2$,(d)restored image with $\alpha = 1.0$,(e)restored image with $\alpha = 0.5$, (f)image difference between (a) and (c)*

the convex TV-L1 approximation model (1.1) can be applied to solve such nonconvex optimization problem (1.5) exactly and globally, in the spatially continuous context. This greatly extends recent studies of Chan et al. [9, 10], from the simplest binary case to the general gray-scale case. In numerics, we propose a fast multiplier-based algorithm upon the constructed equivalent convex formulations, which properly avoids nonsmoothness of the considered TV-L1 energy function. Its numerical reliability and efficiency have been verified by experiments and comparisons to the state of art method, e.g. [33]. In contrast to the graph-cut based approach [11], the proposed approach also avoids heavy memory and computation load especially when the total number of discrete values is large.

Acknowledgement. The research has been supported by MOE (Ministry of Education) Tier II project T207N2202 and IDM project NRF2007IDM-IDM002-010.

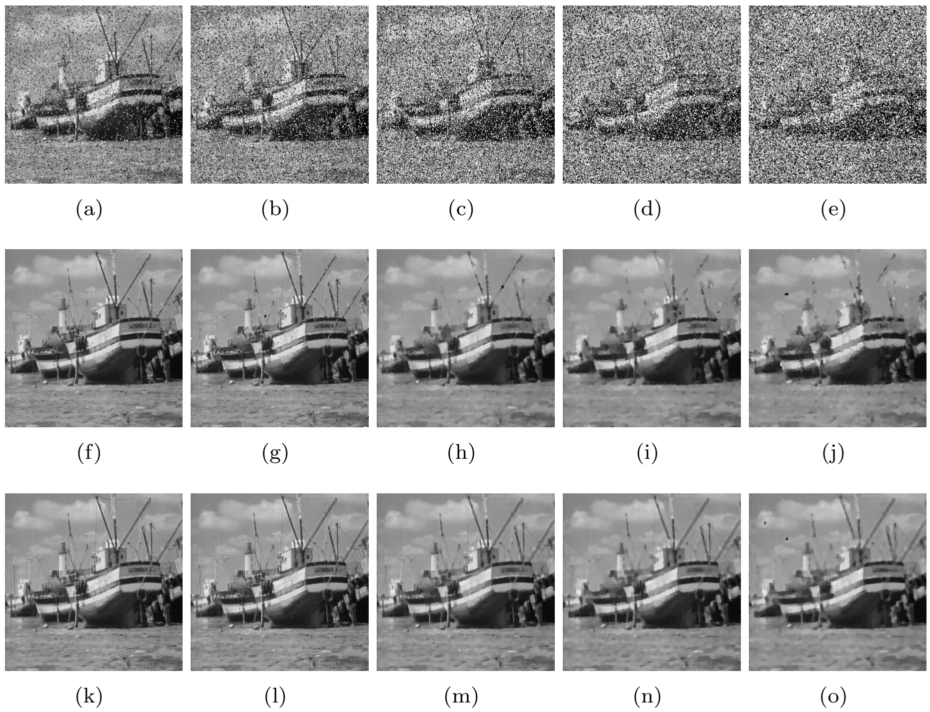


FIG. 5.5. *Boat denoising result with noise level from* 10% *to* 50%, (256 × 256). *(a)-(e)noisy image with noise level from* 10% *to* 50% *respectively,(f)-(j)denoising results by our algorithm with* $\lambda = 0.7$, *(k)-(o)denoising results by [33].*

REFERENCES

[1] S. Alliney. Digital filters as absolute norm regularizers. *Signal Processing, IEEE Transactions on*, 40(6):1548 –1562, June 1992.

[2] S. Alliney. Recursive median filters of increasing order: a variational approach. *Signal Processing, IEEE Transactions on*, 44(6):1346 –1354, June 1996.

[3] E. Bae, J. Yuan, and X.C. Tai. Global minimization for continuous multiphase partitioning problems using a dual approach. Technical report CAM09-75, UCLA, CAM, September 2009.

[4] Egil Bae and Xue-Cheng Tai. Efficient Global Minimization for the Multiphase Chan-Vese Model of Image Segmentation. In *Energy Minimization Methods in Computer Vision and Pattern Recognition: 7th International Conference, EMMCVPR 2009, Bonn, Germany, August 24-27, 2009, Proceedings*, page 28. Springer-Verlag New York Inc, 2009.

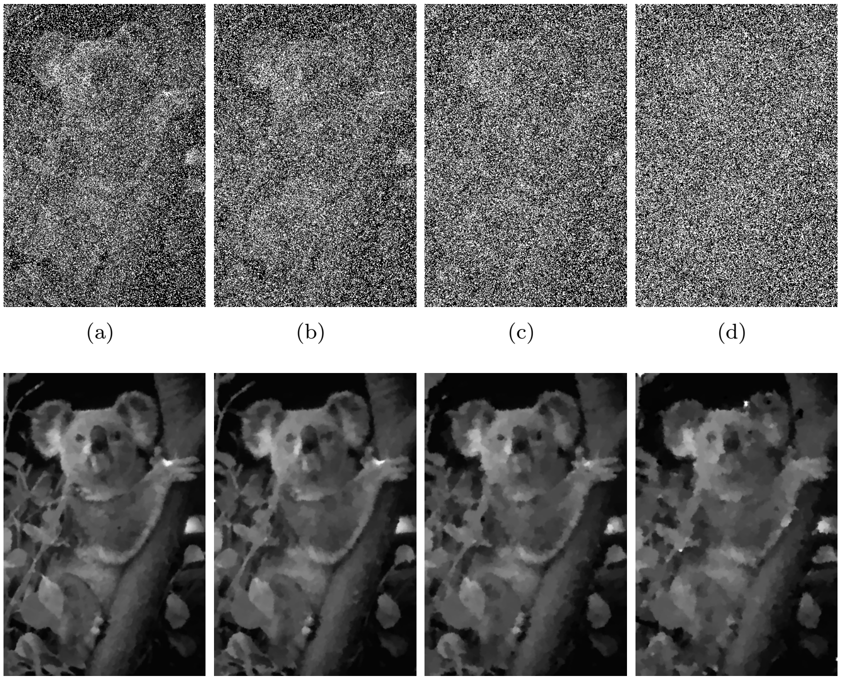
[5] Egil Bae and Xue-Cheng Tai. Graph cut optimization for the piecewise constant level set method applied to multiphase image segmentation. *Scale Space and Variational Methods in Computer Vision*, pages 1–13, 2009.

[6] X. Bresson, S. Esedoglu, P. Vandergheynst, J.P. Thiran, and S. Osher. Fast global minimization of the active contour/snake model. *Journal of Mathematical Imaging and Vision*, 28(2):151–167, 2007.

[7] M. Burger, S. Osher, J. Xu, and G. Gilboa. Nonlinear inverse scale space methods for image restoration. *Lecture Notes in Computer Science*, 3752:25, 2005.

[8] A. Chambolle. An algorithm for total variation minimization and applications. *Journal of Mathematical Imaging and Vision*, 20(1-2):89–97, Jan 2004.

[9] Tony F. Chan and Selim Esedoḡlu. Aspects of total variation regularized L^1 function approximation. *SIAM J. Appl. Math.*, 65(5):1817–1837 (electronic), 2005.



(a) (b) (c) (d)

(e) (f) (g) (h)

FIG. 5.6. *Image denoising with noise levels from* 50% *to* 80%*: (a) noisy image with noise level* 50%*,(b)noisy image with noise level* 60%*,(c)noisy image with noise level* 70%*, (d)noisy image with noise level* 80%*, (e)restored image for* 50% *noise level, (f)restored image for* 60% *noise level, (g) restored image for* 70% *noise level, (h) restored image for* 80% *noise level. All experiments are computed by* $\alpha = 1.1$.

FIG. 5.7. *Pictures of* **1st row** *show the input images where red areas denote the masked areas. Pictures of* **2nd row** *give the recovered images by the proposed approach to* (2.7).

[10] Tony F. Chan, Selim Esedoglu, and Mila Nikolova. Algorithms for finding global minimizers of image segmentation and denoising models. *SIAM J. Appl. Math.*, 66(5):1632–1648 (electronic), 2006.

[11] Jérôme Darbon and Marc Sigelle. Image restoration with discrete constrained total variation part i: Fast and exact optimization. *Journal of Mathematical Imaging and Vision*, 26(3):261–276, 2006.

[12] Jérôme Darbon and Marc Sigelle. Image restoration with discrete constrained total variation part ii: Levelable functions, convex priors and non-convex cases. *Journal of Mathematical Imaging and Vision*, 26(3):277–291, 2006.

[13] Yiqiu Dong, Michael Hintermuller, and Marrick Neri. An Efficient Primal-Dual Method for L1 TV Image Restoration. *SIAM Journal on Imaging Sciences*, 2(4):1168–1189, 2009.

[14] Vincent Duval, Jean F. Aujol, and Yann Gousseau. The tvl1 model: A geometric point of view. *Multiscale Modeling & Simulation*, 8(1):154–189, 2009.

[15] Ivar Ekeland and Roger Téman. *Convex analysis and variational problems*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1999.

[16] E. Esser. Applications of lagrangian-based alternating direction methods and connections to split bregman. Technical report CAM-09-31, UCLA, CAM, 2009.

[17] Ky Fan. Minimax theorems. *Proc. Nat. Acad. Sci. U. S. A.*, 39:42–47, 1953.

[18] Enrico. Giusti. *Minimal surfaces and functions of bounded variation*. Australian National University, Canberra, 1977.

[19] M. Hintermuller, K. Ito, and K. Kunisch. The primal-dual active set strategy as a semismooth Newton method. *SIAM Journal on Optimization*, 13:865, 2002.

[20] Y. Huang, M.K. Ng, and Y.W. Wen. A fast total variation minimization method for image restoration. *Multiscale Modeling and Simulation*, 7:774, 2008.

[21] Hiroshi Ishikawa. Exact optimization for markov random fields with convex priors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25:1333–1336, 2003.

[22] Yves Meyer. *Oscillating patterns in image processing and nonlinear evolution equations*, volume 22 of *University Lecture Series*. American Mathematical Society, Providence, RI, 2001. The fifteenth Dean Jacqueline B. Lewis memorial lectures.

[23] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin. An iterative regularization method for total variation-based image restoration. *Multiscale Modeling and Simulation*, 4(2):460–489, 2006.

[24] Nikos Paragios, Yunmei Chen, and Olivier Faugeras. *Handbook of Mathematical Models in Computer Vision*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2005.

[25] Thomas Pock, Thomas Schoenemann, Gottfried Graber, Horst Bischof, and Daniel Cremers. A convex formulation of continuous multi-label problems. In *ECCV '08: Proceedings of the 10th European Conference on Computer Vision*, pages 792–805, Berlin, Heidelberg, 2008. Springer-Verlag.

[26] R. Tyrrell Rockafellar. *Convex analysis*. Princeton Mathematical Series, No. 28. Princeton University Press, Princeton, N.J., 1970.

[27] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60(1-4):259–268, 1992.

[28] S. Setzer. *Split Bregman algorithm, Douglas-Rachford splitting and frame shrinkage*, pages 464–476. Springer, LNCS vol. 5567, 2009.

[29] S. Setzer, G. Steidl, and T. Teuber. Deblurring Poissonian images by split Bregman techniques. *J. Visual Commun. Image Repres.*, 2009. accepted.

[30] Gabriele Steidl, Joachim Weickert, Thomas Brox, Pavel Mrzek, Martin Welk, and Pavel Mr zek. On the equivalence of soft wavelet shrinkage, total variation diffusion, total variation regularization, and sides. *SIAM J. Numer. Anal*, 42:686–713, 2004.

[31] Xue-Cheng Tai and Yuping Duan. Domain decomposition methods with Graph cuts algotithms for image segmentation. *UCLA CAM Report*, pages 9–54, 2009 (to appear in Int. J. Numer. Anal. Model.).

[32] Xue-Cheng Tai and Chunlin Wu. *Augmented Lagrangian Method, Dual Methods, and Split Bregman Iteration for ROF Model*, pages 502–513. Springer, LNCS vol. 5567, 2009.

[33] Chunlin Wu, Juyong Zhang, and Xue-Cheng Tai. Augmented lagrangian method for total variation restoration with non-quadratic fidelity. Technical report CAM-09-82, UCLA, CAM, 2009.

[34] Wotao Yin, Donald Goldfarb, and Stanley Osher. The total variation regularized l^1 model for multiscale decomposition. *Multiscale Modeling and Simulation*, 6(1):190–211, 2007.

[35] Xiaoqun Zhang, Martin Burger, and Stanley Osher. A unified primal-dual algorithm framework based on bregman iteration. Technical report CAM-09-99, UCLA, CAM, 2009.