East Asian Journal on Applied Mathematics
doi: 10.4208/eajam.220310.181110a

Vol. **1**, No. 2, pp. 172-186
May 2011

A Convex and Exact Approach to Discrete Constrained TV-L1 Image Approximation

Jing Yuan[*,1], Juan Shi[2] and Xue-Cheng Tai[2,3]

[1] *Computer Science Department, Middlesex College, University of Western Ontario, London, Ontario, N6A 5B7, UK.*
[2] *Division of Mathematical Sciences, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore.*
[3] *Department of Mathematics, University of Bergen, Norway.*

Received 22 March 2010; Accepted (in revised version) 18 November 2010

Available online 7 April 2011

Abstract. We study the TV-L1 image approximation model from primal and dual perspective, based on a 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 non-convex optimization problem considered. This result greatly extends recent studies of Chan et al., from the simplest binary constrained case to the general gray-value constrained case, through the proposed rounding scheme. In addition, we construct a fast multiplier-based algorithm based on the proposed primal-dual model, which properly avoids variability of the concerning TV-L1 energy function. Numerical experiments validate the theoretical results and show that the proposed algorithm is reliable and effective.

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

1. Introduction

Many tasks of image processing can be formulated and solved successfully by convex optimization models – e.g. image denoising [21, 24], image segmentation [5], image labeling [4, 22] etc. The reduced convex formulations can be studied in a mathematically sound way and usually tackled by a tractable numerical scheme. Minimizing the total-variation function for such convex image processing formulations is of great importance [5, 6, 17–20, 24, 27], as it preserves edges and sharp features.

[*]Corresponding author. *Email addresses:* cn.yuanjing@gmail.com (J. Yuan), shij0004@e.ntu.edu.sg (J. Shi), tai@mi.uib.no (X.-C. Tai)

In their pioneer works [8,9], Chan et al. proposed the TV-L1 regularized image approximation model

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

which was first introduced and studied by Alliney [1,2] for discrete one-dimensional signals' denoising. Chan et al. [8,9] demonstrated an interesting property of the TV-L1 model (1.1) – viz. that 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

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

globally and exactly! Hence (1.1) provides an exact convex relaxation of the binary constrained optimization problem (1.2). Chan et al. [8, 9] also proved that rounding the computed result of (1.1) may give a series of global optima of the binary constrained optimization model (1.2).

Previous work and motivation

With the help of co-area formula, Chan et al. [8,9] proved that the energy functional $P(u)$ of (1.1) can be represented in terms of the upper level-set sequence of the image functions $u(x)$ and $f(x)$ – i.e.

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

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

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

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

Yin et al. [30] 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 $\gamma_1 \ge \gamma_2$, 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. [10,11] in an image graph setting, where the anisotropic total-variation term was considered and a fast graph-cut based algorithm introduced. Goldfarb and Yin also developed an efficient pre-flow based graph-cut approach to such L1 image approximation regularized by discretized total-variation.

However, in the spatial continuous setting [30], such an approach has advantages and disadvantages for processing a gray-scale image in practice. On the one hand, the total number of gray values is finite – i.e. $u(x) \in \{0, \ldots, 255\}$ – and 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. In order to globally tackle (1.1), one must be at least examine a large number of obtained level-sets – so computation directly addressing multiply layered level-sets is impractical, in a real image processing task with a large number of different gray values. A similar and interesting work in spatially continuous image labeling along linearly ordered labels (i.e. layered level sets) was recently addressed [3].

In addition, the PDE-descent method is often taken to numerically approximate the global optimum of (1.1) [8, 9, 12, 30], which smooths the total-variation term by $(\partial_x u^2 + \partial_y u^2 + \epsilon^2)^{1/2}$. Actually, even if ϵ takes a small enough value, the co-area formula is no longer satisfied. Indeed, 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

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

given $f(x) \in \{f_1, \ldots, f_n\}$. In this paper, we assume the gray values f_i, $i = 1, \ldots, n$, are ordered by $f_1 < \ldots < f_n$. Clearly, integers $0, \ldots, 255$ may be taken as the option in most cases. We greatly extend the result for the *binary constrained* image denoising (1.2), obtained by Chan et al. [8, 9], to the more general *gray-value constrained* case.

Our main contributions can be summarized 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 of its relevant discrete constrained image processing task. This greatly extends the results achieved by Chan et al. [8, 9] to more general cases, which the authors believe to be new.

3. 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 efficiency are verified by optimization theories and experiments.

In parallel to our multiplier-based method, several other dual formulations and algorithmic schemes were proposed recently in the literature, [12, 14, 25, 26, 28, 29, 31]. In contrast to [14, 28, 29, 31], we apply the proposed equivalent primal-dual and dual formulations as a complete approach to (1.1), including both variational analyses and algorithms and not just a derivation of the algorithmic scheme. In addition, the primal-dual algorithm we proposed differs from [14, 28, 29, 31] for 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, for comparison with the equivalent models introduced in this section.

2.1. Equivalent primal-dual model

With the help of conjugates [23], the data term of (1.1) can be re-expressed as

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

Moreover, it is well known that the total-variation term of (1.1) can also be formulated as [16]

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

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

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

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

2.2. Equivalent dual model

On observing that u is unconstrained and minimizing (2.3) over u, we have the linear equality

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

and hence the constrained maximization problem

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

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 (e.g. see [13,15]) are all satisfied – i.e. the constraints on the 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. It follows that there exists at least one saddle point [13, 15], so 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 reacts 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):

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

3. Global and Exact Optima

In this section, we study the nonconvex optimization problem (1.5) and show the TV-L1 formulation (1.1) gives the convex relaxed model of (1.5) – i.e. its optimum solves (1.5) globally and exactly through the proposed rounding scheme. We prove several propositions in stating our results.

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 $f_1 \le u^*(x) \le f_n$ almost everywhere.*

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 $u^*(x) \le f_n \ \forall x \in \Omega$.

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 $f(x) \le f_n$ and $u^*(x) > f_n \ \forall x \in \tilde{\Omega}$, we have

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

It follows that

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

By the co-area 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.

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

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

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 $u^*(x) \in [f_1, f_n]$. □

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 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. Consider the interval $\Gamma = [f_1, f_n]$ co-area 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

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

Since $u(x) = \int_{f_1}^{u(x)} d\gamma = \int_\Gamma U^\gamma(x) d\gamma$ 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 $\gamma \in [f_1, f_n)$

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

Then it implies 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

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

□

Corollary 3.1. *Given a bounded scalar function $f_1 \leq u(x) \leq f_n \ \forall x \in \Omega$ and $n-1$ different values γ_i, $i = 1, \ldots, n-1$ such that $f_1 \leq \gamma_1 < \ldots < \gamma_{n-1} \leq f_n$, 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, $i = 1, \ldots, n-1$.

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

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

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

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

Therefore we have

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

□

Given the above, we can prove the following proposition:

Proposition 3.3. *Given the image function $f(x) \in \{f_1,\ldots,f_n\}$, where $f_1 < \ldots < f_n$ and the boundary of each concerning upper level set $F^{f_i}(x)$, $i = 1,\ldots,n$, is regular, then for any given $n-1$ values γ_i, $i = 1,\ldots,n-1$ 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 $u^\gamma(x) \in \{f_1,\ldots,f_n\}$ 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.

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

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

Further, u^γ is also a global optimum of (1.1), since by Corollary 3.1, p^* also maximizes the integral $\int_\Omega u^\gamma \operatorname{div} p\,dx$ over the convex set C_1 and

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

At the next step, we can prove

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

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 $f(x) \le f_k$, in view of $u^*(x) \ge f_k$ we have $q^*(x) = -\alpha$ for $u^*(x) > f_k$ or $q^*(x) \ge -\alpha$ 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 $f(x) \ge f_{k+1}$, in view of $u^*(x) \le f_{k+1}$ we have $q^*(x) = \alpha$ for $u^*(x) < f_{k+1}$ or $q^*(x) \le \alpha$ 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 $q(x)\in[-\alpha,\alpha]$.

Thus we have that q^* maximizes the integral $\langle q, f-u^\gamma\rangle$ over the convex set S_α, and hence (3.11) is proven.

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 a global minimum of (1.1). Since (1.1) is just the relaxed version of (1.5), $u^\gamma(x) \in \{f_1,\dots,f_n\}$ solves (1.5) exactly and globally. □

The proposed rounding scheme (3.9) actually gives

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

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

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) gives the corresponding Lagrangian function. Thus 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 $c > 0$ – so the classical augmented Lagrangian algorithm can be applied in a splitting optimization framework over each dual variables q and p, by exploring projections to corresponding convex sets under the following *multiplier-based algorithm*.

Algorithm 4.1. 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:

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

     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 $\operatorname{div} C_\lambda$ and can be implemented by Chambolle's algorithm [7];

- Update u^{k+1} by

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

  and set $k = k + 1$, and repeat until convergence is achieved.

Clearly, Algorithm 4.1 explores two simple projection sub-steps at each iteration, which avoids tackling the less smooth terms in (1.1) directly.

5. Numerical Experiments

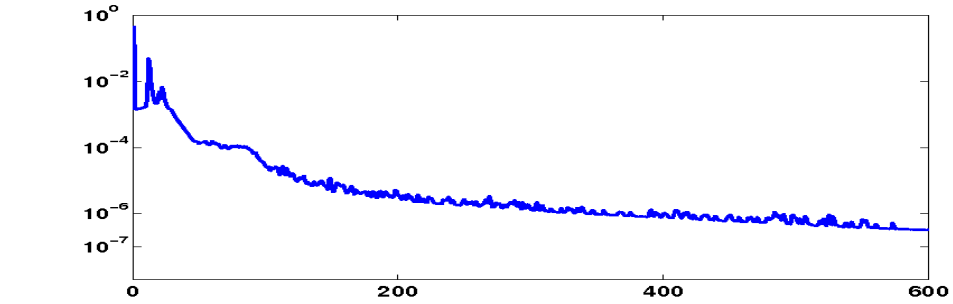
In all experiments, convergence was achieved by evaluating the following error:

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

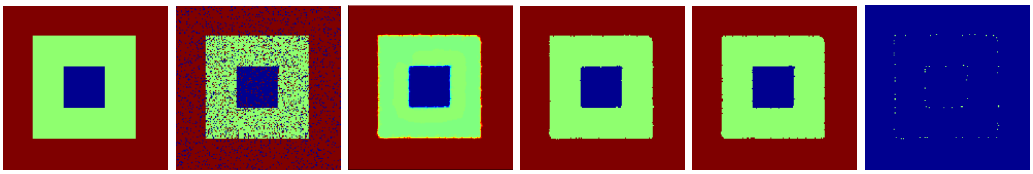
which is the ratio of the primal-dual gap to the image approximation $u(x)$, see (4.3).

To evaluate the performance of rounded results in the following experiments, we took the energy difference associated with the computed optimum u^* and the rounded result u^γ which is evaluated by

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



(a)



(b) (c) (d) (e) (f) (g)

Figure 1: (a) convergence rate (600 iterations); (b) the ground truth image color-coded as red: 0, green: 0.5, blue: 1; (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 the two rounded results.

5.1. Synthetic image

A synthetic image $f(x) \in \{0, 0.5, 1\}$ taken for this experiment is shown in Fig. 1(c), color-coded as red: 0, green: 0.5, blue: 1. We set the penalty parameter $\alpha = 1$ and the augmented parameter $c = 6$, when the algorithm ran for 600 iterations and converged with error err $= 3.23 \times 10^{-7}$. Fig. 1(a) shows the plot of convergence rate.

In this experiment, two rounding schemes were taken: $\{\gamma_1 = 0.25, \gamma_2 = 0.75\}$; $\{\gamma_1 = 0.35, \gamma_2 = 0.65\}$. For the computed result u^*, it gave the energy $P(u^*) = 2938.7$. The two corresponding rounded results produced the energy $P(u^\gamma) = 2937.3, 2937.3$, i.e. both rounding schemes gave the same energy! The energy ratios are 4.76×10^{-4}.

5.2. Gray value images

For the given gray-value images $f(x)$ of the experiments, 256 gray levels are naturally encoded by $f(x) \in \{0, \ldots, 255\}$.

The experiment results given in Fig. 2 show the denoising of the penguin image in Fig. 2(a), as downloaded from the middleburry data set: http://vision.middlebury.edu/MRF. The rounding scheme is given by $\gamma = \{0.5, 1.5, \ldots, 255\}$, i.e. it just gives the nearest integer. In all experiments where $\alpha = 1.3, 1, 0.5$, the algorithm 4.1 converged to an error below 2×10^{-6} within 300 iterations – cf. Fig. 2(b). For the $\alpha = 1.3, 1, 0.5$, the ratios of the energy differences are all nearly zero!

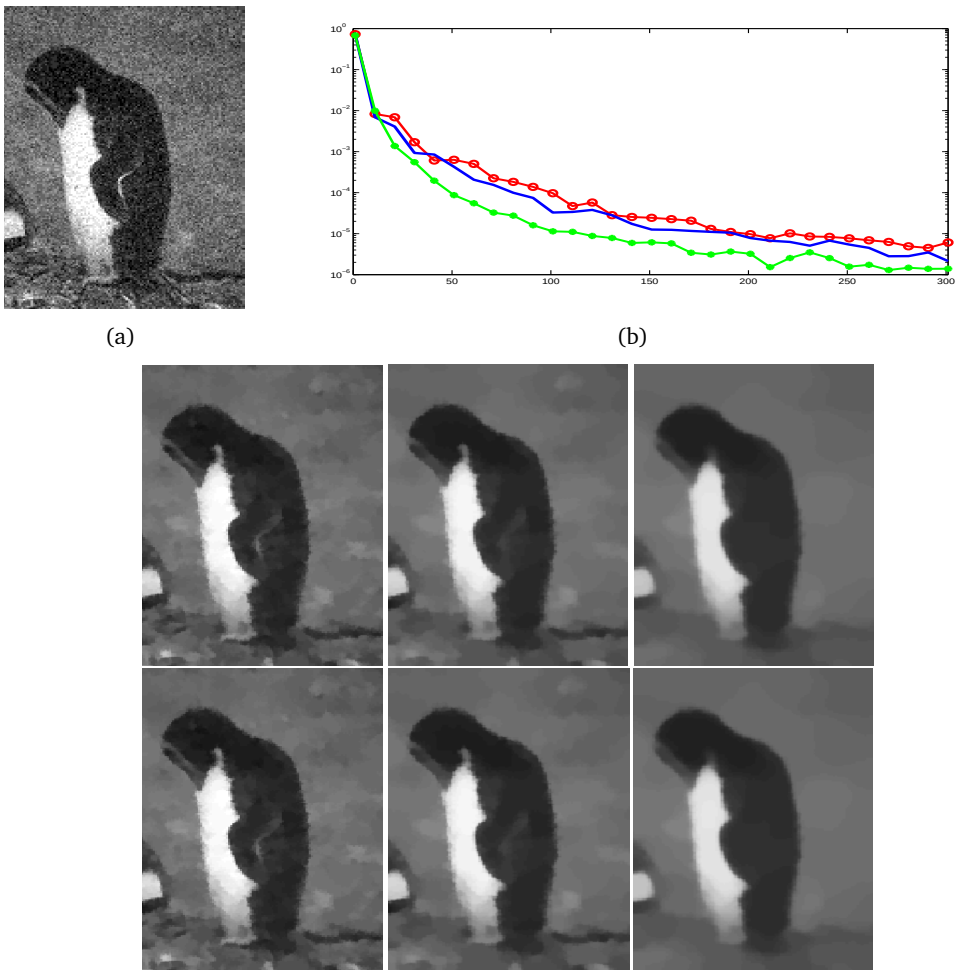


Figure 2: (a) the input image $f(x)$; (b) plot of convergences (300 iterations): red circle line : $\alpha = 1.3$, blue solid line: $\alpha = 1$ and green star line: $\alpha = 0.5$. Figures of 2nd row show the computation results when $\alpha = 1.3, 1, 0.5$, respectively. Figures of 3rd row show the rounding results when $\alpha = 1.3, 1, 0.5$, respectively.

The images processed in the experiments, shown in Fig. 3, are downloaded from the Berkeley segmentation dataset and benchmark. In all experiments, we set $\alpha = 0.5$. All experiment results show the ratios of energy differences are again all nearly zeros.

6. Conclusions

This work studies the discrete constrained image approximation, based on the corresponding TV-L1 energy function. We prove that 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

Figure 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 $\alpha = 0.5$.

al. [8, 9], from the simplest binary case to the general gray-scale case. In the numerics, we build up the multiplier-based algorithm based upon the proposed equivalent convex formulation, which avoids any variability in the considered TV-L1 energy considered. Its numerical reliability and efficiency have been verified by experiments.

Acknowledgments

The research has been supported by Ministry of Education Tier II project T207N2202 and IDM project NRF2007IDM-IDM002-010. Support from SUG 20/07 is also gratefully acknowledged.

References

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

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

[3] E. Bae, J. Yuan, X.-C. Tai, and Y. Boykov. A study on continuous max-flow and min-cut approaches part ii: Multiple linearly ordered labels. Technical report CAM-10-62, UCLA, CAM, 2010.

[4] 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.

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

[6] M. Burger, S. Osher, J. Xu, and G. Gilboa. Nonlinear inverse scale space methods for image restoration. *Lect. Notes Comput. Sci.*, 3752:25, 2005.

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

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

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

[10] J. Darbon and M. Sigelle. Image restoration with discrete constrained total variation part i: Fast and exact optimization. *J. Math. Imaging Vision*, 26(3):261–276, 2006.

[11] J. Darbon and M. Sigelle. Image restoration with discrete constrained total variation part ii: Levelable functions, convex priors and non-convex cases. *J. Math. Imaging Vision*, 26(3):277–291, 2006.

[12] Y. Dong, M. Hintermuller, and M. Neri. An Efficient Primal-Dual Method for L1 TV Image Restoration. *SIAM J. Imaging Sci.*, 2(4):1168–1189, 2009.

[13] I. Ekeland and R. Téman. *Convex analysis and variational problems*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1999.

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

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

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

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

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

[19] Y. 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.

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

[21] N. Paragios, Y. Chen, and O. Faugeras. *Handbook of Mathematical Models in Computer Vision*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2005.

[22] T. Pock, T. Schoenemann, G. Graber, H. Bischof, and D. 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.

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

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

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

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

[27] G. Steidl, J. Weickert, T. Brox, P. Mrázek and M. Welk. On the equivalence of soft wavelet shrinkage, total variation diffusion, total variation regularization, and sides. *SIAM J. Numer. Anal*, 42:686–713, 2004.

[28] X.C. Tai and C. Wu. *Augmented Lagrangian Method, Dual Methods, and Split Bregman Iteration for ROF Model*, pages 502–513. Springer, LNCS vol. 5567, 2009.

[29] C. Wu, J. Zhang, and X.C. Tai. Augmented lagrangian method for total variation restoration with non-quadratic fidelity. Technical report CAM-09-82, UCLA, CAM, 2009.

[30] W. Yin, D. Goldfarb, and S. Osher. The total variation regularized l^1 model for multiscale decomposition. *Multiscale Model. Simul.*, 6(1):190–211, 2007.

[31] X. Zhang, M. Burger, and S. Osher. A unified primal-dual algorithm framework based on bregman iteration. Technical report CAM-09-99, UCLA, CAM, 2009.