

Augmented Lagrangian Methods for p -Harmonic Flows with the Generalized Penalization Terms and Application to Image Processing

Huibin Chang¹ and Xue-Cheng Tai^{2,*}

¹ School of Mathematical Sciences, Tianjin Normal University, Tianjin 300387;
Department of Mathematics, East China Normal University, Shanghai 200241,
China.

² Department of Mathematics, University of Bergen, Johanness Brunsgate 12,
Bergen 5007, Norway.

Received 6 December 2011; Accepted (in revised version) 1 July 2012

Available online 11 January 2013

Abstract. In this paper, we propose a generalized penalization technique and a convex constraint minimization approach for the p -harmonic flow problem following the ideas in [Kang & March, IEEE T. Image Process., 16 (2007), 2251–2261]. We use fast algorithms to solve the subproblems, such as the dual projection methods, primal-dual methods and augmented Lagrangian methods. With a special penalization term, some special algorithms are presented. Numerical experiments are given to demonstrate the performance of the proposed methods. We successfully show that our algorithms are effective and efficient due to two reasons: the solver for subproblem is fast in essence and there is no need to solve the subproblem accurately (even 2 inner iterations of the subproblem are enough). It is also observed that better PSNR values are produced using the new algorithms.

AMS subject classifications: 68U10, 65N21, 74S20

Key words: p -harmonic flows, denoising, generalized penalization terms, saddle-point problem, image processing, augmented Lagrangian methods.

1. Introduction

First we present the p -harmonic flow problem in [17, 28] as

$$\min_{U \in W^{1,p}(\Omega, S^{N-1})} E(U) = \int_{\Omega} |\nabla U(\mathbf{x})|_F^p dx, \quad (1.1)$$

where $1 \leq p < \infty$. Some notations in (1.1) are defined as follows:

*Corresponding author. Email addresses: changhuibin@gmail.com (H. B. Chang), xuechengtai@gmail.com (X.-C. Tai)

- Ω : an open subset of \mathcal{R}^M .
- ∇ : differential operator, i.e.,

$$\nabla \mathbf{U} = \left(\frac{\partial U_i}{\partial x_j} \right)_{N \times M}, \quad \nabla U_i = \left(\frac{\partial U_i}{\partial x_1}, \dots, \frac{\partial U_i}{\partial x_M} \right), \quad \forall \mathbf{U} = (U_1, \dots, U_N)^T \in \mathcal{R}^N.$$

- $|\cdot|$: Euclidean norm, and $|\cdot|_F$: Frobenius norm, i.e.,

$$|\mathbf{B}|_F = \sqrt{\sum_{i,j} B_{i,j}^2}, \quad \forall \mathbf{B} = (B_{i,j})_{N \times M}.$$

- $\mathbf{W}^{1,p}(\Omega, S^{N-1}) := \mathbf{W}^{1,p}(\Omega, \mathcal{R}^N) \cap S^{N-1}$,

$$S^{N-1} := \{\mathbf{U} \in \mathcal{R}^N : |\mathbf{U}| = 1, \text{ a.e.}\}, \quad M \geq 1, \quad N \geq 2.$$

- $(\cdot)^T$ denotes the transpose of the matrix.

The minimization of (1.1) is associated with the Dirichlet boundary condition: $\mathbf{U}|_{\Omega} = \mathbf{n}_0 \in S^{N-1}$ or Neumann boundary condition: $\partial \mathbf{U} / \partial \mathbf{n} = 0$ where \mathbf{n} is the exterior unit normal to $\partial \Omega$.

The difficulties of solving (1.1) lie in three aspects, i.e., the non-convexity due to constraints of S^{N-1} , the non-regularity and non-uniqueness. Several kinds of approaches are used to solve (1.1) in literature. The authors [14, 15] dealt with the Euler-Lagrange equations for problem (1.1) using the iteration which updated the solution by normalizing $\mathbf{U} = \mathbf{V}/|\mathbf{V}|$. Analysis on the similar algorithms were done in [3–5] and constraints preserving finite element methods were proposed in [6, 7]. In [19], the authors adopted the saddle-point approach and established the proper finite element discretization in the case of two dimensional space. The second kind of approach was proposed by adding a penalization to eliminate the non-convex constraint of S^{N-1} [8, 9, 24]. Such technique is also adopted to solve the Ginzburg-Landau functional, i.e.,

$$E_\epsilon(\mathbf{U}) := E(\mathbf{U}) + \frac{1}{\epsilon} \int_{\Omega} (|\mathbf{U}^2| - 1)^2 \, dx. \quad (1.2)$$

The third kind of approach is to reformulate (1.1) to become a constrained optimization problem as follows

$$\min_{\mathbf{V} \in \mathbf{W}^{1,p}(\Omega, \mathcal{R}^N)} E(\mathbf{V}), \quad \text{s.t. } \mathbf{V} = \frac{\mathbf{U}}{|\mathbf{U}|}. \quad (1.3)$$

Such constraint was used to preserve gradient descent for solving (1.3) in [10, 28]. Further improvements based (1.3) were done in [17, 29] in which the authors proposed an innovative curvilinear search method with the global convergence property as long as satisfying Armijo-Wolfe conditions.

In this paper, by combining the second and the third approach via the relaxation and penalization, a general model is established with penalization terms following the idea in [20]. We derive the saddle-point problem for (1.1) based on the augmented Lagrangian

methods. To solve the saddle-point problem fast and efficiently, we adopt the operator splitting method or alternating optimization method, which are the classical techniques for solving the augmented Lagrangian functional. Motivated by the development of image processing, all subproblems after the operator splitting can be efficiently solved by the primal-dual method, the dual method, or fast Fourier transform (FFT).

As we know, the p -harmonic flow problem has many applications in extensive fields such as liquid crystal theory, directional diffusion, color image denoising, and etc. To demonstrate the efficiency of the proposed algorithm, we present numerical examples of the above problem. On one hand, our experiments show that the algorithms are effective and efficient. There are two reasons: one is that the solver for subproblem is fast and efficient and the other is that we do not need to solve the subproblem accurately (even 2 times of inner iterations of the subproblem works). On the other hand, the examples of cases with new penalization terms achieve better PSNR (Peak Signal to Noise Ratio) values than the penalization in [20].

The paper is organized as follows. In Section 2 we present the model with general penalization terms. In Section 3 we present the saddle-point problem solved by the augmented Lagrangian methods. Algorithms for solving the saddle-point problem with details on the subproblem solutions are then presented. Numerical examples are given in Section 4. In the last section we conclude our paper with the future work.

2. Model

We propose the model based on the ROF [26] with a known function U_0 which may be contaminated by random noise

$$\min_{U \in W^{1,p}(\Omega, S^{N-1})} \hat{F}(U) = E(U) + \frac{\eta}{2} \int_{\Omega} |U - U_0|^2 dx. \quad (2.1)$$

As we know, the constraint set $W^{1,p}(\Omega, S^{N-1})$ is nonconvex. Our aim is to eliminate the nonconvex constraint. Combining the second and the third approach in Section 1, two additional terms are added to (2.1). First we introduce a convex set K to (2.1) as follows $K = \{U \in \mathcal{R}^N : |U| \leq 1, \text{ a.e.}\}$ that would be added to (2.1) as convex relaxation. Then following the idea in [20], we propose the generalized penalization term as $\frac{1}{s_1 s_2 \epsilon} \int_{\Omega} (1 - |U|^{s_1})^{s_2}$ where ϵ is the positive constant and s_1, s_2 are the positive integers. Our main idea is to add these two conditions to (2.1) with convex relaxation and generalized penalization terms. Thus we modify the model (2.1) as

$$\begin{aligned} \min_{U \in W^{1,p}(\Omega, \mathcal{R}^N)} F_{\epsilon}(U) = & \int_{\Omega} |\nabla U|_F^p dx + \frac{\eta}{2} \int_{\Omega} |U - U_0|^2 dx \\ & + \frac{1}{s_1 s_2 \epsilon} \int_{\Omega} (1 - |U|^{s_1})^{s_2} dx + \chi_K(U), \end{aligned} \quad (2.2)$$

where

$$\chi_K(\mathbf{U}) = \begin{cases} 0, & \mathbf{U} \in K, \\ +\infty, & \mathbf{U} \notin K. \end{cases}$$

Here one readily knows that different values of s_1 and s_2 generate different models. In this paper, we concentrate on the choices of s_1 and s_2 which lead to easily computed, and effective models. Specifically speaking, we concern the following three models.

- Case I: $s_1 = 2, s_2 = 1$.
Penalization term is $\frac{1}{2\epsilon} \int_{\Omega} (1 - |\mathbf{U}|^2)$. Due to the constraint set K , this term has positive lower bound. Thus following the idea in [31], the special choice of $s_1 = 2, s_2 = 1$ is suitable. By using the augmented Lagrangian methods, the subproblems are easy to be computed. Furthermore, we can restore the data with better PSNR values than that of Case II.
- Case II: $s_1 = 1, s_2 = 2$.
Here we must emphasize the article [20]. They added this penalization term $\frac{1}{2\epsilon} \int_{\Omega} (1 - |\mathbf{U}|)^2$. We still study this case and give the new algorithms, which is faster than the simple gradient projection or time marching methods.
- Case III: $s_1 = 2, s_2 = 2$.
Penalization term is $\frac{1}{4\epsilon} \int_{\Omega} (1 - |\mathbf{U}|^2)^2$. It is the same as (1.2) which is used to solve the Ginzburg-Landau functional. However, we do not compute this case at all. Because it leads to higher order algebraic equations which do not have explicit solution. Our general penalization terms include this case as well.

In this paper we only consider the above three special cases. From our analysis in Section 3, the model with the higher order penalization term is difficult to solve i.e., $s_1, s_2 \geq 2$. Thus we choose the cases of I and II to realize the model, and develop the fast algorithms whose sub-problem has the explicit solution proposed by the ideas in the image processing areas.

The penalization term is non-convex so that the problem may have not unique minimization. Following [20], we get the theorem which reveals the equivalence of (2.1) and (2.2) when ϵ tends to zero. In our paper we concentrate on the case of $p = 1$. The existence of the minimization of (2.2) shall be considered in the bounded variation space

$$BV(\Omega, \mathcal{R}^N) = \left\{ \mathbf{U} \in L^1(\Omega, \mathcal{R}^N) : \int_{\Omega} |D\mathbf{U}| \leq +\infty \right\},$$

where $\int_{\Omega} |D\mathbf{U}|$ is defined as

$$\int_{\Omega} |D\mathbf{U}| := \sup_{\phi \in \Phi} \sum_{i=1}^N \int_{\Omega} U_i \operatorname{div} \phi_i,$$

where

$$\begin{aligned} \Phi := \left\{ \phi = (\phi_1^T, \dots, \phi_N^T)^T \in C_0^1(\Omega, \mathcal{R}^{N \times M}) : \right. \\ \left. |\phi|_F \leq 1, \text{ a.e. in } \Omega, \phi_i = (\phi_{i,1}, \dots, \phi_{i,M}) \in \mathcal{R}^M \right\}, \\ \operatorname{div} \phi_i = \sum_{j=1}^M \frac{\partial \phi_{i,j}}{\partial x_j}, \quad \mathbf{U} = (U_1, \dots, U_N)^T \in \mathcal{R}^N. \end{aligned}$$

The following theorem holds for (2.2) with the generalized term in the bounded variation space, while the norm term $\int_{\Omega} |\nabla \mathbf{U}|_F \, d\mathbf{x}$ is replaced by $\int_{\Omega} |D\mathbf{U}|$.

Theorem 2.1. (*Existence and Convergence*)

1. There exists $\mathbf{U}_{\epsilon} \in BV(\Omega, \mathcal{R}^N)$ which minimizes $F_{\epsilon}(\mathbf{U})$ with fixed ϵ .
2. Supposing that $A = \{\mathbf{U}_{\epsilon}\}$ is a sequence satisfying $\mathbf{U}_{\epsilon} = \arg \min_{\mathbf{U} \in BV(\Omega, \mathcal{R}^N)} F_{\epsilon}(\mathbf{U})$, there exists a subsequence $\{\mathbf{U}_{\epsilon_k}\} \subset A$ and \mathbf{U}^* , s.t. $\mathbf{U}_{\epsilon_k} \rightarrow \mathbf{U}^*$ in $L^q(\Omega, \mathcal{R}^N)$ as $k \rightarrow \infty$, $\forall 1 \leq q < \infty$, where $\mathbf{U}^* = \arg \min_{\mathbf{U} \in BV(\Omega, S^{N-1})} \hat{F}(\mathbf{U})$.

Proof. The proof is the same as [20] so we omit the details. \square

3. Augmented Lagrangian methods for (2.2)

In this section, we will use augmented Lagrangian methods in [16, 30] to solve (2.2). There are many efficient algorithms for the image restoration problems recently, e.g., Graph-cuts method [1, 21, 25], Bregman iteration [18, 23, 27, 32], primal-dual method [12, 13, 33, 34], dual projection method [11] and augmented Lagrangian method [16, 30]. Motivated by the above algorithms, especially the augmented Lagrangian method which is suitable for (2.2), we formulate a saddle-point model. The Uzawa algorithm for the saddle-point model benefits advantages: the variables can be separated into subproblems by decomposition techniques, and fast algorithms, e.g., FFT, dual projection methods, or primal-dual method can be adopted to solve the related subproblems. Thus we modify the model (2.2) as follows:

$$\begin{aligned} \max_{\Lambda} \min_{\mathbf{U}, \mathbf{V}} L^{s_1, s_2}(\mathbf{U}, \mathbf{V}; \Lambda) = \int_{\Omega} |\nabla \mathbf{U}|_F \, d\mathbf{x} + \frac{\eta}{2} \int_{\Omega} |\mathbf{U} - \mathbf{U}_0|^2 \, d\mathbf{x} + \frac{1}{s_1 s_2 \epsilon} \int_{\Omega} (1 - |\mathbf{V}|^{s_1})^{s_2} \, d\mathbf{x} + \chi_K(\mathbf{V}) \\ + \int_{\Omega} \Lambda \cdot (\mathbf{U} - \mathbf{V}) \, d\mathbf{x} + \frac{r}{2} \int_{\Omega} |\mathbf{U} - \mathbf{V}|^2 \, d\mathbf{x}. \end{aligned} \quad (3.1)$$

We consider the above model under Case I and Case II.

$$\begin{aligned} \max_{\Lambda} \min_{\mathbf{U}, \mathbf{V}} L^{2,1}(\mathbf{U}, \mathbf{V}; \Lambda) = \int_{\Omega} |\nabla \mathbf{U}|_F \, d\mathbf{x} + \frac{\eta}{2} \int_{\Omega} |\mathbf{U} - \mathbf{U}_0|^2 \, d\mathbf{x} + \frac{1}{2\epsilon} \int_{\Omega} (1 - |\mathbf{V}|^2) \, d\mathbf{x} + \chi_K(\mathbf{V}) \\ + \int_{\Omega} \Lambda \cdot (\mathbf{U} - \mathbf{V}) \, d\mathbf{x} + \frac{r}{2} \int_{\Omega} |\mathbf{U} - \mathbf{V}|^2 \, d\mathbf{x}, \end{aligned} \quad (3.2)$$

where r is a positive constant and Λ is the Lagrange multiplier.

On the other hand, for Case II, we can present the model as follows:

$$\begin{aligned} \max_{\Lambda} \min_{U, V} L^{1,2}(U, V; \Lambda) &= \int_{\Omega} |\nabla U|_F dx + \frac{\eta}{2} \int_{\Omega} |U - U_0|^2 dx + \frac{1}{2\epsilon} \int_{\Omega} (1 - |V|)^2 dx + \chi_K(V) \\ &+ \int_{\Omega} \Lambda \cdot (U - V) dx + \frac{r}{2} \int_{\Omega} |U - V|^2 dx. \end{aligned} \quad (3.3)$$

In the saddle-point problem (3.2), the parameter r is used to control the parameter ϵ . First of all we need to keep $r \neq 1/\epsilon$. The sufficient condition is $r > 1/\epsilon$, to guarantee the convexity of $L(U, V; \lambda)$ over V . If $r < 1/\epsilon$, the V subproblem can also be solved. However, we cannot guarantee the convergence of the algorithm. From the right one in Fig. 5.3 one readily knows the proposed algorithm diverges if $r < 1/\epsilon$.

In the following we give detailed augmented Lagrangian methods for (3.2) and (3.3). See Algorithm 1.

Algorithm 1. 1. Initialization: $\Lambda^0 = 0$, $U^0 = U_0$, $V_0 = U_0$, parameters η , ϵ , r .

2. For $k = 1, 2, \dots$, update U^k , V^k as follows:

$$U^k = \arg \min_U L^{s_1, s_2}(U, V^{k-1}; \Lambda^{k-1}), \quad (3.4a)$$

$$V^k = \arg \min_V L^{s_1, s_2}(U^k, V; \Lambda^{k-1}). \quad (3.4b)$$

3. Update the multiplier:

$$\Lambda^k = \Lambda^{k-1} + r(U^k - V^k).$$

First we show how to solve V -subproblem (3.4b). The different penalization terms lead to different algorithms for V -subproblem while they have the same U -subproblem.

3.1. V -subproblem

In the following parts we consider different cases with different s_1 , s_2 . The solutions with closed forms are obtained easily. The subproblems own almost the same structure, i.e., minimizing the quadratic functionals over a convex set. Thus the solutions to the subproblems have similar forms. Details are present as follows.

3.1.1. $s_1 = 2, s_2 = 1$

We rewrite the subproblem (3.4b) as

$$V^k = \arg \min_V \mathcal{S}_1(V) := c_{2,1} \int_{\Omega} |V - \mathbf{g}_{2,1}|^2 dx + \chi_K(V), \quad (3.5)$$

where $c_{2,1} := r - 1/\epsilon \neq 0$, $\mathbf{g}_{2,1} := (rU^k + \Lambda^{k-1})/c_{2,1}$. One readily observes that the functional is quadratic over the convex set. Thus it shall have global minimizer. If $c_{2,1} > 0$,

(3.5) has the explicit solution $\mathbf{V}^k = \mathbf{g}_{2,1} \min\{1, 1/|\mathbf{g}_{2,1}|\}$. If $c_{2,1} < 0$, the \mathbf{V} -subproblem is non-convex. The minimization can be deduced as

$$\mathbf{V}^k = \begin{cases} -\frac{\mathbf{g}_{2,1}}{|\mathbf{g}_{2,1}|}, & \mathbf{x} \in \{\mathbf{g}_{2,1}(\mathbf{x}) \neq 0\}, \\ \mathbf{b}, \quad \forall |\mathbf{b}| = 1, & \mathbf{x} \in \{\mathbf{g}_{2,1}(\mathbf{x}) = 0\}. \end{cases}$$

3.1.2. $s_1 = 1, s_2 = 2$

We rewrite the subproblem (3.4b) as

$$\mathbf{V}^k = \arg \min_{\mathbf{V}} \hat{\mathcal{S}}_1(\mathbf{V}) := \int_{\Omega} (|\mathbf{V}|^2 - 2c_{1,2}|\mathbf{V}| - 2\mathbf{g}_{1,2}\mathbf{V})d\mathbf{x} + \chi_K(\mathbf{V}),$$

where $c_{1,2} := 1/(1 + r\epsilon)$, and $\mathbf{g}_{1,2} := \epsilon c_{1,2}(r\mathbf{U}^k + \mathbf{\Lambda}^{k-1})$. The objective functional is non-convex. But we can still compute the minimizer after some simple computation. The explicit solution to the above minimization is

$$\mathbf{V}^k(\mathbf{x}) = \begin{cases} \frac{\mathbf{g}_{1,2}}{|\mathbf{g}_{1,2}|} \min\{c_{1,2} + |\mathbf{g}_{1,2}|, 1\}, & \mathbf{x} \in \{\mathbf{g}_{1,2}(\mathbf{x}) \neq 0\}, \\ c_{1,2}\mathbf{b}, \quad \forall |\mathbf{b}| = 1, & \mathbf{x} \in \{\mathbf{g}_{1,2}(\mathbf{x}) = 0\}. \end{cases} \quad (3.6)$$

From the previous part, we give the explicit solution for the subproblems. The computation complexity of these two cases are almost the same. We will compare these two cases in the numerical examples. Then we will continue to consider the \mathbf{U} -subproblem.

3.2. \mathbf{U} -subproblem

Next we consider \mathbf{U} -subproblem (3.4a). Rewrite the subproblem (3.4a) as

$$\mathbf{U}^k = \arg \min_{\mathbf{U}} \mathcal{S}_2(\mathbf{U}) := \int_{\Omega} |\nabla \mathbf{U}|_F d\mathbf{x} + \frac{c_2}{2} \int_{\Omega} |\mathbf{U} - \mathbf{f}|^2 d\mathbf{x}, \quad (3.7)$$

where $c_2 := \eta + r$, $\mathbf{f} := (\eta\mathbf{U}_0 + r\mathbf{V}^{k-1} - \mathbf{\Lambda}^{k-1})/c_2$.

Following the idea in [2, 11, 12, 30], we propose the dual projection method, primal dual method and augmented Lagrangian method in the following three subsections. Then (3.7) is considered in bounded variation space as

$$\mathbf{U}^k = \arg \min_{\mathbf{U}} \mathcal{S}_2(\mathbf{U}) := \int_{\Omega} |D\mathbf{U}| d\mathbf{x} + \frac{c_2}{2} \int_{\Omega} |\mathbf{U} - \mathbf{f}|^2 d\mathbf{x}. \quad (3.8)$$

3.2.1. Dual projection method for (3.8)

Some notations:

$$\operatorname{div} \mathbf{Y} = (\operatorname{div} Y_1, \dots, \operatorname{div} Y_N)^T, \quad \forall \mathbf{Y} = (Y_1^T, \dots, Y_N^T)^T \in \mathcal{R}^{N \times M}, \quad Y_i = (Y_{i,1}, \dots, Y_{i,M}) \in \mathcal{R}^M.$$

We can consider the dual form for the subproblem. Following this direction as [2, 11], the dual projection method is constructed as

Algorithm 2. Dual Projection Method

1. Initialization: τ and \mathbf{Y}^0 .
2. Iteration:

$$l = 0, \dots, \quad \mathbf{Y}^{l+1} = \frac{\mathbf{Y}^l + \tau \mathbf{H}}{1 + \tau |\mathbf{H}|_F}, \quad \mathbf{H} := \nabla(\operatorname{div} \mathbf{Y}^l - c_2 \mathbf{f}).$$

3. $\mathbf{U}^k = \mathbf{f} - \frac{1}{c_2} \operatorname{div} \mathbf{Y}^\infty$.

3.2.2. Primal dual method for (3.8)

Define

$$L = \|\nabla\| := \sup_{U \neq 0} \frac{\int_{\Omega} |\nabla U|_F}{\int_{\Omega} |U|},$$

where U satisfies some boundary conditions. In the discrete form, $L \approx \sqrt{8}$ in [12]. The primal dual method is listed as follows:

Algorithm 3. Arrow-Hurwicz Primal Dual Method

1. Initialization: τ_0, σ_0 , with $\tau_0 \sigma_0 L^2 \leq 1$, and $\mathbf{P}^0, \hat{\mathbf{U}}^0$.
2. Iteration: $l = 0, \dots$,

$$\begin{aligned} 1) \quad \mathbf{P}^{l+1} &= \frac{\mathbf{H}_1}{\max\{1, |\mathbf{H}_1|_F\}}, & \mathbf{H}_1 &:= \mathbf{P}^l + \sigma_l \nabla \hat{\mathbf{U}}^l; \\ 2) \quad \mathbf{U}^{l+1} &= \frac{\mathbf{H}_2 + \tau_l c_2 \mathbf{f}}{1 + \tau_l c_2}, & \mathbf{H}_2 &:= \mathbf{U}^l + \tau_l \operatorname{div} \mathbf{P}^{l+1}; \\ 3) \quad \tau_{l+1} &= \theta_l \tau_l, \quad \sigma_{l+1} = \frac{\sigma_l}{\theta_l}, & \theta_l &= \frac{1}{\sqrt{1 + 2\gamma \tau_l}}. \end{aligned}$$

3. $\mathbf{U}^k := \hat{\mathbf{U}}^\infty$.

The algorithm with variant steps can convergence fast with the order of $\mathcal{O}(n^{-2})$ as they in [12] have proved, where n is the iteration number.

3.2.3. Augmented Lagrangian method for (3.7)

By introducing the variable \mathbf{Q} , the Lagrangian multiplier Λ_s , and the penalization parameter r_1 , the augmented Lagrangian of (3.7) is

$$L_s(\mathbf{U}, \mathbf{Q}; \Lambda_s) = \int_{\Omega} |\mathbf{Q}|_F dx + \frac{c_2}{2} \int_{\Omega} |U - f|^2 dx \\ + \int_{\Omega} \Lambda_s : (\mathbf{Q} - \nabla \mathbf{U}) dx + \frac{r_1}{2} \int_{\Omega} |\mathbf{Q} - \nabla \mathbf{U}|_F^2 dx,$$

where $\mathbf{P} : \mathbf{Q} = \sum_{i,j} P_{i,j} Q_{i,j}$, $\mathbf{P} = (P_{i,j})$, $\mathbf{Q} = (Q_{i,j})$, which is Frobenius inner product.

Then following the idea in [30], we give the procedures by using the operator splitting method for solving $\max_{\Lambda_s} \min_{\mathbf{U}, \mathbf{Q}} L_s(\mathbf{U}, \mathbf{Q}; \Lambda_s)$ as follows:

Algorithm 4. Augmented Lagrangian Methods

1. Initialization: $\hat{\mathbf{U}}^0, \mathbf{Q}^0, \Lambda_s^0$.
2. Iteration: For $l = 0, \dots$,
 - solve $\hat{\mathbf{U}}^{l+1} = \arg \min_{\mathbf{U}} L_s(\mathbf{U}, \mathbf{Q}^l; \Lambda_s^l)$, that is equivalent to solving the following equation

$$-r_1 \Delta \hat{\mathbf{U}}^{l+1} + c_2 \hat{\mathbf{U}}^{l+1} = c_2 \mathbf{f} - \operatorname{div} \Lambda_s^l - r_1 \operatorname{div} \mathbf{Q}^l;$$

- solve $\mathbf{Q}^{l+1} = \arg \min_{\mathbf{Q}} L_s(\hat{\mathbf{U}}^{l+1}, \mathbf{Q}, \Lambda_s^l)$, that has the closed form

$$\mathbf{Q}^{l+1} = \frac{1}{r_1} \max \left\{ 1 - \frac{1}{|\mathbf{W}|_F}, 0 \right\} \mathbf{W}, \quad \text{where } \mathbf{W} := r_1 \nabla \hat{\mathbf{U}}^{l+1} - \Lambda_s^l;$$

- update Λ_s^{l+1} as follows

$$\Lambda_s^{l+1} = \Lambda_s^l + r_1 (\mathbf{Q}^{l+1} - \nabla \hat{\mathbf{U}}^{l+1}).$$

3. $\mathbf{U}^k := \hat{\mathbf{U}}^\infty$.

In the following section, we give some numerical example for Algorithm 1 with $s_1 = 2$, $s_2 = 1$ (Simplified as Algorithm 1^{2,1}) and with $s_1 = 1$, $s_2 = 2$ (Simplified as Algorithm 1^{1,2}). We point out the main difference between the two algorithms as follows. The U -subproblem and V -subproblem is strictly convex in Algorithm 1^{2,1} assuming the penalization parameter ϵ and the multiplier r satisfying $r > 1/\epsilon$. In Algorithm 1^{1,2}, the V -subproblem is nonconvex.

4. Difference schemes and numerical examples

We use the classical difference schemes for PDE-based image processing problems. For the simplicity, we set $\Omega \subset \mathbb{R}^2$, i.e., $M = 2$, which is a rectangle as well. We denote the domain as

$$\Omega_h := \{(x_i, y_j) \mid x_i = ih_x, y_j = jh_y, 0 \leq i \leq m, 0 \leq j \leq n\}$$

with grid size of h_x, h_y . We define the gradient and divergence of each $u_{i,j}$,

$$(\nabla u)_{i,j} = \left((\nabla u)_{i,j}^1, (\nabla u)_{i,j}^2 \right),$$

with

$$(\nabla u)_{i,j}^1 = \begin{cases} \frac{1}{h_x}(u_{i+1,j} - u_{i,j}), & i < m, \\ 0, & i = m, \end{cases} \quad (\nabla u)_{i,j}^2 = \begin{cases} \frac{1}{h_y}(u_{i,j+1} - u_{i,j}), & j < n, \\ 0, & j = n. \end{cases}$$

And the divergence of $\mathbf{p} = (p^1, p^2) \in \mathcal{R}^2$ satisfying $\text{div} = -\nabla^*$ in discrete form is defined as

$$(\text{div} \mathbf{p})_{i,j} = \begin{cases} \frac{1}{h_x}(p_{i,j}^1 - p_{i-1,j}^1), & 0 < i < m, \\ \frac{1}{h_x}p_{i,j}^1, & i = 0, \\ -\frac{1}{h_x}p_{i-1,j}^1, & i = m, \end{cases} + \begin{cases} \frac{1}{h_y}(p_{i,j}^2 - p_{i,j-1}^2), & 0 < j < n, \\ \frac{1}{h_y}p_{i,j}^2, & j = 0, \\ -\frac{1}{h_y}p_{i,j-1}^2, & j = n, \end{cases}$$

where

$$\begin{aligned} (\nabla \mathbf{U})_{i,j} &= ((\nabla U_1)_{i,j}^T, \dots, (\nabla U_N)_{i,j}^T)^T, \quad \forall \mathbf{U} = (U_1, \dots, U_N)^T, \quad U_l \in \mathcal{R}^2, \\ (\text{div} \mathbf{Y})_{i,j} &= ((\text{div} \mathbf{Y}_1)_{i,j}, \dots, (\text{div} \mathbf{Y}_N)_{i,j})^T, \quad \forall \mathbf{Y} = (\mathbf{Y}_1^T, \dots, \mathbf{Y}_N^T)^T \in \mathcal{R}^{N \times 2}, \quad \mathbf{Y}_l = (Y_{l,1}, Y_{l,2}) \in \mathcal{R}^2. \end{aligned}$$

Based on the difference schemes we present numerical examples in the cases that $M = 2$, and $N = 2, 3$. These examples consist of three types models: liquid crystals, directional denoising and chromaticity denoising for the color image.

4.1. Liquid crystals

Here we test one example for liquid crystals where $N = 2$. Consider the case similar to the examples in [28]. Assume that initial condition $\mathbf{U}_0 = (u, v)^T$ is defined as

$$\begin{aligned} u(x_1, x_2) &= \frac{x_1 - 0.5}{|\mathbf{x} - \mathbf{x}_0|} + 0.1(1 + x_1^2 - x_2^2) - 0.5\zeta_1, \\ v(x_1, x_2) &= \frac{x_2 - 0.5}{|\mathbf{x} - \mathbf{x}_0|} + 0.1(x_1 - 2x_2) + 0.5\zeta_2, \end{aligned}$$

with $\mathbf{x} = (x_1, x_2) \in \Omega = (0, 1)^2$, $\zeta_i \sim \text{Normal}(0, 1)$ and $\mathbf{x}_0 = (0.5, 0.5)$. We test the example by Algorithm 1^{2,1} and Algorithm 1^{1,2}, where the \mathbf{U} -subproblem is solved by the dual projection methods as in Subsection 3.2.1. We set $r = 5$, $\eta = 0.5$, $\tau = 0.05$, $\epsilon = 1$ for this example. From the results displayed in Fig. 1, we observe that both Algorithm 1^{2,1} and Algorithm 1^{1,2} perform well and stable. As we know, the map $\mathbf{x} \mapsto (\mathbf{x} - \mathbf{x}_0)/|\mathbf{x} - \mathbf{x}_0|$ is an exact solution to (1.1) and the restored liquid crystals are well reordered as the map indicates in Fig. 1.

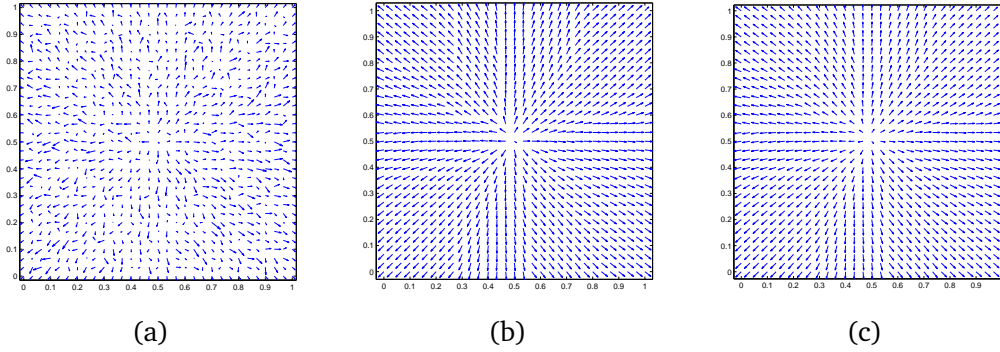


Figure 1: Liquid crystal U_0 in (a); U by Algorithm $1^{2,1}$ in (b); U by Algorithm $1^{1,2}$ in (c).

4.2. Directional denoising

4.2.1. Example 1

We define U_0 as follows: for $|x_1| + |x_2| \neq 0$

$$U_0^T = \begin{cases} (\cos(3\pi/2 + \alpha\zeta), \sin(3\pi/2 + \alpha\zeta)), & -1 < x_1 \leq 0, \quad 0 \leq x_2 < 1, \\ (\cos(\pi/2 + \alpha\zeta), \sin(\pi/2 + \alpha\zeta)), & 0 \leq x_1 < 1, \quad -1 < x_2 \leq 0, \\ (\cos(\pi + \alpha\zeta), \sin(\pi + \alpha\zeta)), & 0 < x_1 < 1, \quad 0 < x_2 < 1, \\ (\cos(\alpha\zeta), \sin(\alpha\zeta)), & -1 < x_1 < 0, \quad -1 < x_2 < 0, \end{cases}$$

with $\Omega = (-1, 1)^2$. To test the directional denoising of Algorithm $1^{2,1}$ and Algorithm $1^{1,2}$, we set $\alpha = 0.8$ and we have the following conclusions from the experiments shown in Figs. 2 and 3, respectively. For the first group tests shown in Fig. 2, set $r = 10$, $\eta = 0.2$, $\tau = 0.05$, $\epsilon = 0.2$. We let $r > 1/\epsilon$. Both Algorithm $1^{2,1}$ and Algorithm $1^{1,2}$ work well as the examples of liquid crystal reveals. For the second group tests in Fig. 3, set $r = 4$, $\eta = 0.2$, $\tau = 0.05$, $\epsilon = 0.2$ or $r = 0.2$, $\eta = 0.2$, $\tau = 0.05$, $\epsilon = 0.2$. One readily knows that $r < 1/\epsilon$. The Algorithm $1^{2,1}$ is effective in the left one in Fig. 3 while it fails to work for the right one. Thus we assume that $r > 1/\epsilon$ in the following examples.

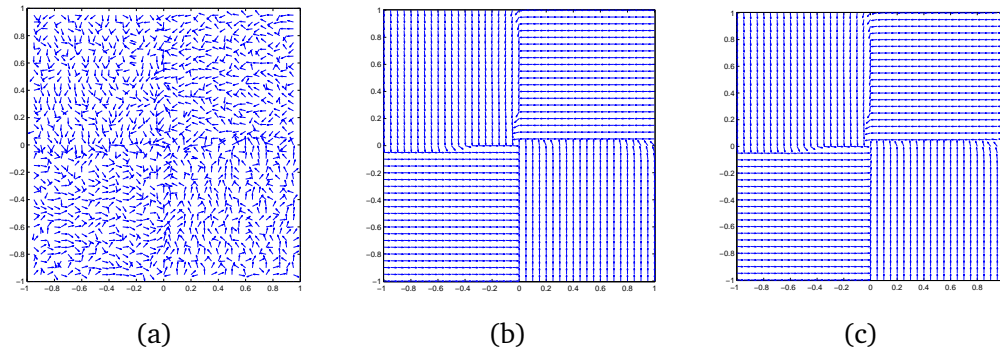


Figure 2: Directional denoising U_0 in (a); U by Algorithm $1^{2,1}$ in (b); U by Algorithm $1^{1,2}$ in (c).

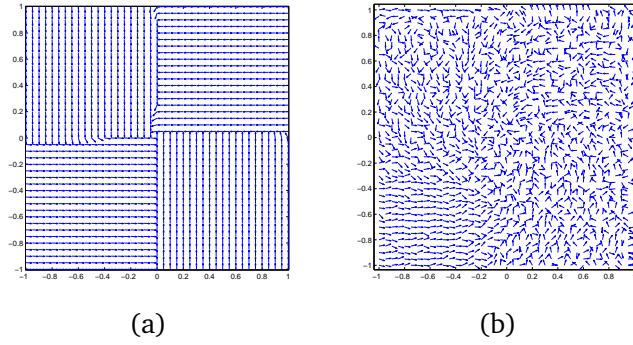


Figure 3: Directional denoising: (a) $r = 4$, $\epsilon = 0.2$; (b) $r = \epsilon = 0.2$.

4.2.2. Example 2: directional denoising represented by color image

In Fig. 4, a map from $\mathcal{R}^2 \rightarrow S^2$ is presented by three channels of RGB colors instead of vector in [28] and the noised image is also given in Fig. 4, which is corrupted by noise of level 0.5. We display the denoised images in Fig. 5. It is shown from the results in Fig. 5 that the edge of the color image are successfully restored by the proposed algorithms.

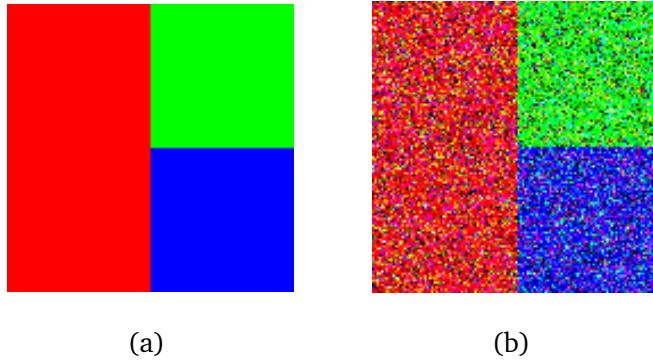


Figure 4: Directional denoising. (a) Original image; (b) noised image.

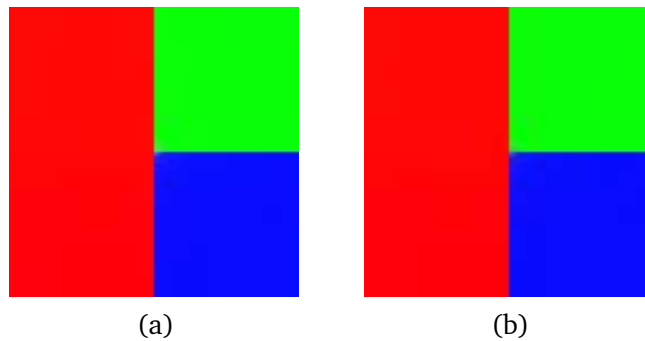


Figure 5: Directional denoising. Denoising image by (a) Algorithm $1^{2,1}$; (b) Algorithm $1^{1,2}$. $r = 10$, $\eta = 0.2$, $\tau = 0.1$, $\epsilon = 0.2$.

4.3. Chromaticity denoising

We choose a color image $I = (I_R, I_G, I_B)^T \in \mathcal{R}^3$ and define the brightness of the image as

$$B = |I| = \sqrt{I_R^2 + I_G^2 + I_B^2}$$

and also the chromaticity as

$$C = \frac{I}{|I|} = \left(\frac{I_R}{|I|}, \frac{I_G}{|I|}, \frac{I_B}{|I|} \right) \in S^2.$$

We add the noise to chromaticity $U^0 = C_n/|C_n|$, $C_n = C + \alpha\zeta$ with $\zeta = (\zeta_1, \zeta_2, \zeta_3)^T$, $\zeta_i \sim \text{Normal}(0, 1)$, where α is the noise level while keeping the brightness unchanged. After denoising the chromaticity, we reassemble the new image as $I_{new} = U|I|$ using the unchanged brightness $|I|$ and the denoised chromaticity U . We define the peak signal-to-noise ratio (PSNR) as

$$\text{PSNR} = 10 \log_{10} \frac{255^2 mn}{\sum_{i,j} |U_p - U|_{i,j}^2},$$

where U_p is the original image, U is the restored image with $m \times n$ pixels. The original images and noised images are in Fig. 6, where noise level is $\alpha = 0.6$. We also define the numerical energy E and the relative error e as follows:

$$E(U^k) := \sum_{i,j} (|\nabla U^k|_F)_{i,j} + \frac{\eta}{2} \sum_{i,j} (|U^k - U_0|^2)_{i,j} + \frac{1}{s_1 s_2 \epsilon} \sum_{i,j} ((1 - |U^k|^{s_1})^{s_2})_{i,j},$$

$$e(U^k) := \frac{\sum_{i,j} (|U^k - U^{k-1}|)_{i,j}}{\sum_{i,j} (|U^k|)_{i,j}}.$$

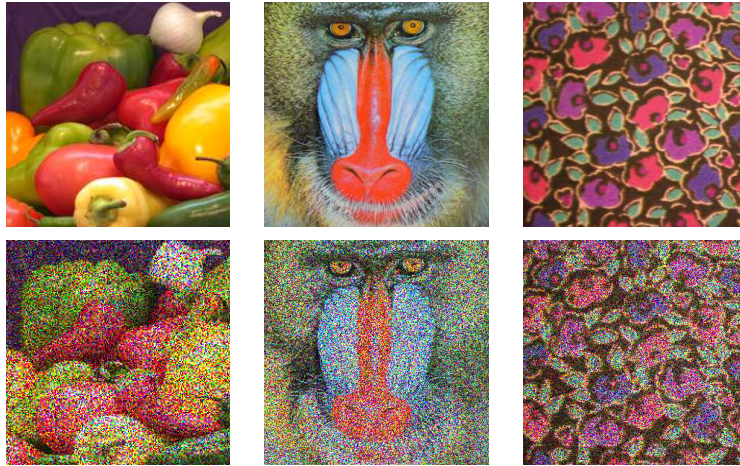


Figure 6: In the first row, Original images from left to right: peppers (resolution : 215×212), animal (resolution: 312×312), and fabric (resolution: 302×309); In the second row, different noised images.

In the following subsection, we will discuss the algorithms which are different in solving the V -subproblem. Furthermore, the different algorithms for the U -subproblem are discussed as well.

4.3.1. Algorithms $1^{2,1}$ and $1^{1,2}$ if U -subproblem is solved by dual projection methods

For Figs. 7-9, the qualities of the restored images from different algorithms are measured by the PSNR, energy E and relative error e . One readily observes that both algorithms work well as the PSNR values of the restored images by the proposed algorithm under Case I are

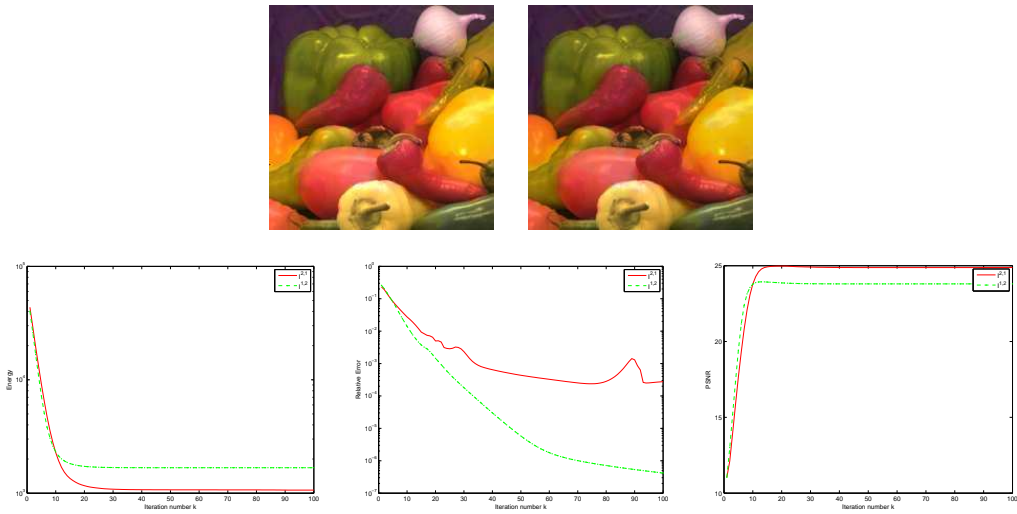


Figure 7: Peppers.

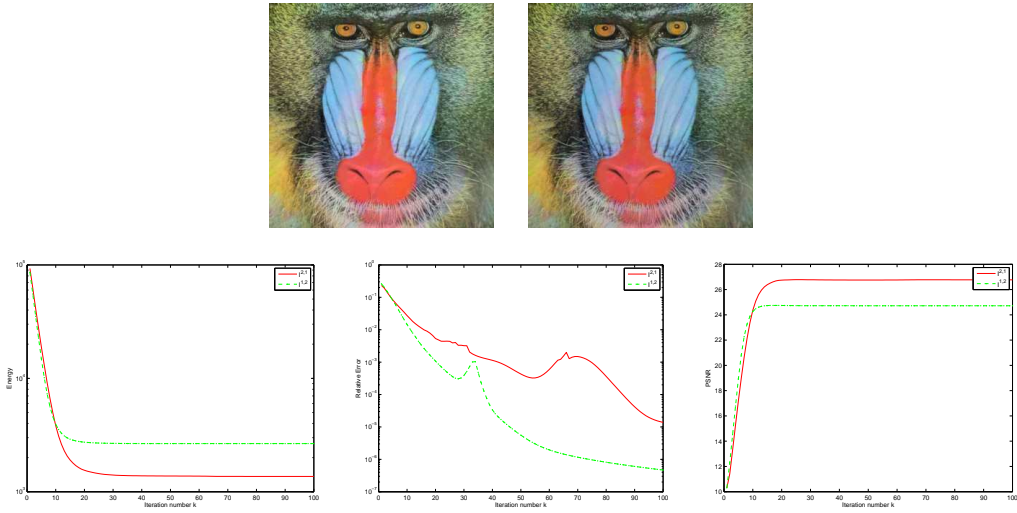


Figure 8: Animal.

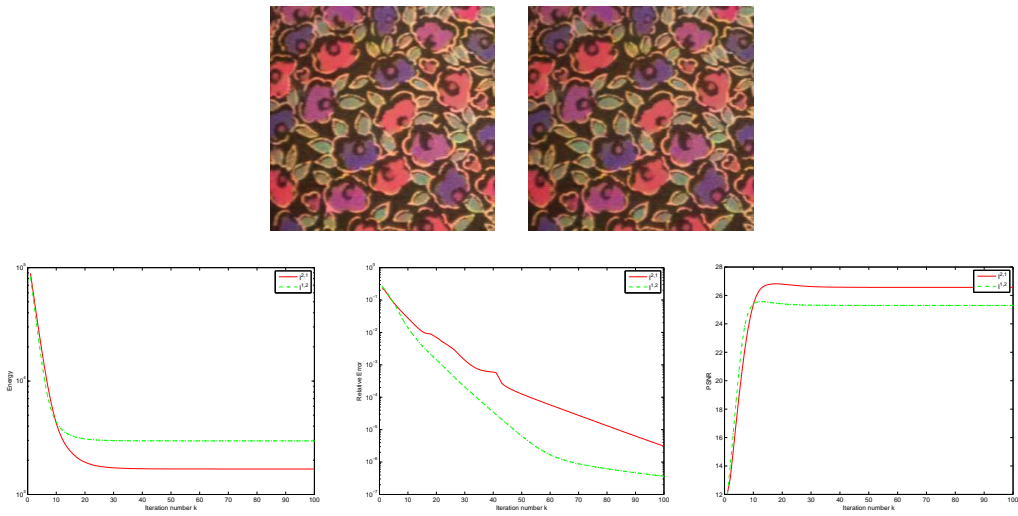


Figure 9: Fabrics.

better than Case II which was proposed in [20]. At the same time, the minimum energy from Algorithm 1^{2,1} is smaller than the energy obtained from Algorithm 1^{1,2}.

4.3.2. Algorithm 1^{2,1} with different solver for U -subproblem

For the examples in the above subsection, we just concern the Algorithm 1^{2,1} for better PSNR values of the restored images. In this paper we have proposed three different methods for solving the U -subproblem. For the simplicity, we use DP to denote the dual

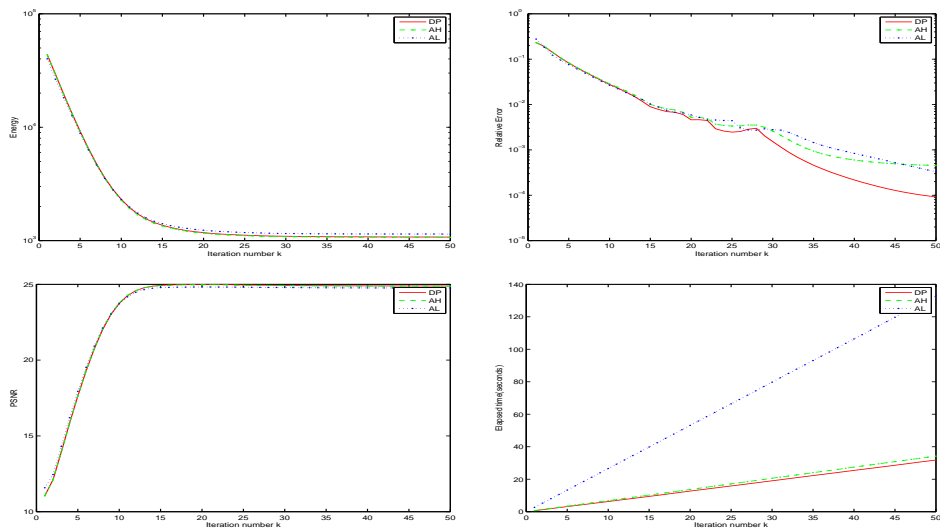


Figure 10: Data of peppers.

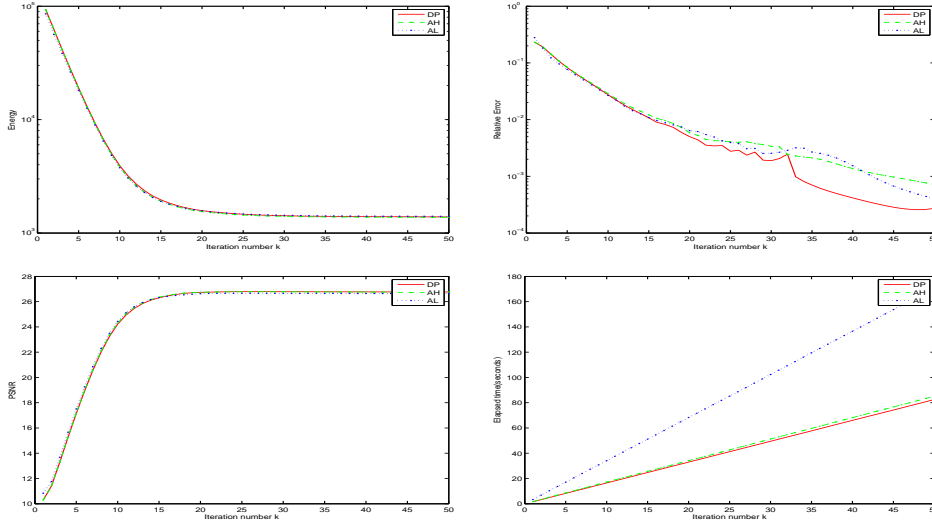


Figure 11: Data of animal.

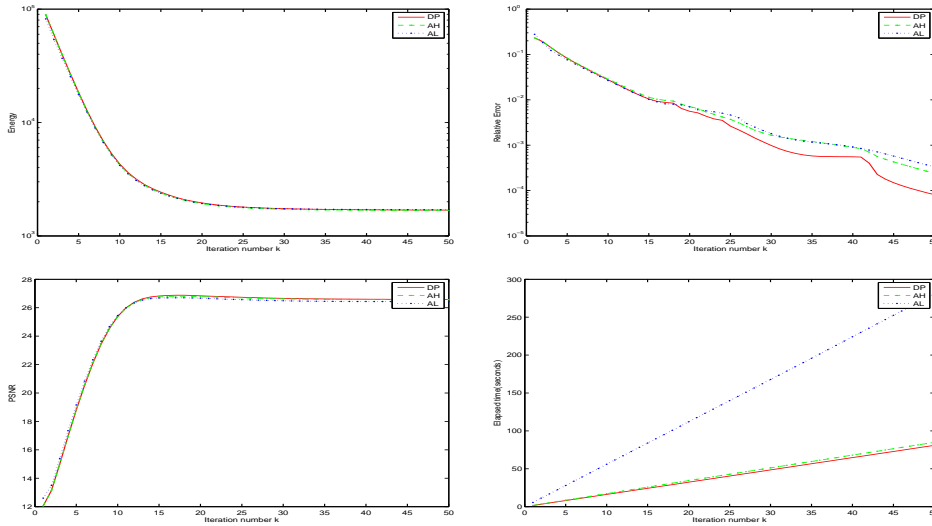


Figure 12: Data of fabric.

projection method, AH to denote the Arrow-Hurwicz primal dual method and AL to denote the augmented Lagrangian method. The numerical energy, relative error and PSNR value by these three algorithms are plotted in Figs. 10-12. In the plots, we set the inner loop for the U -subproblem to be 10. Besides, plots with 2 inner loops are shown in Figs. 13-15. These figures tell that we can always restore the images with few inner iterations because there is no need to solve the solution exactly. Not only the algorithms for the subproblem are fast, but also few iterations are needed in the inner loop for solving the subproblem. It is also shown the dual projection method is the best way to solve the U -subproblem

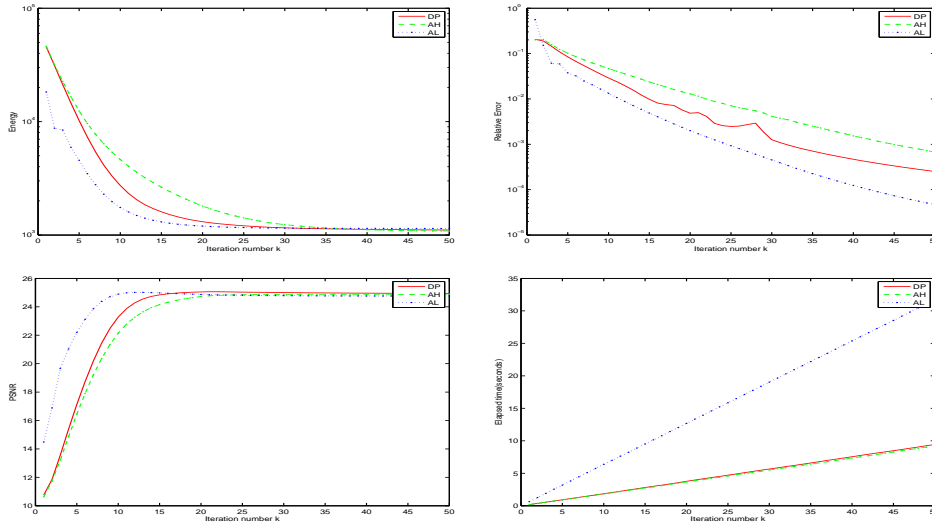


Figure 13: Data of peppers.

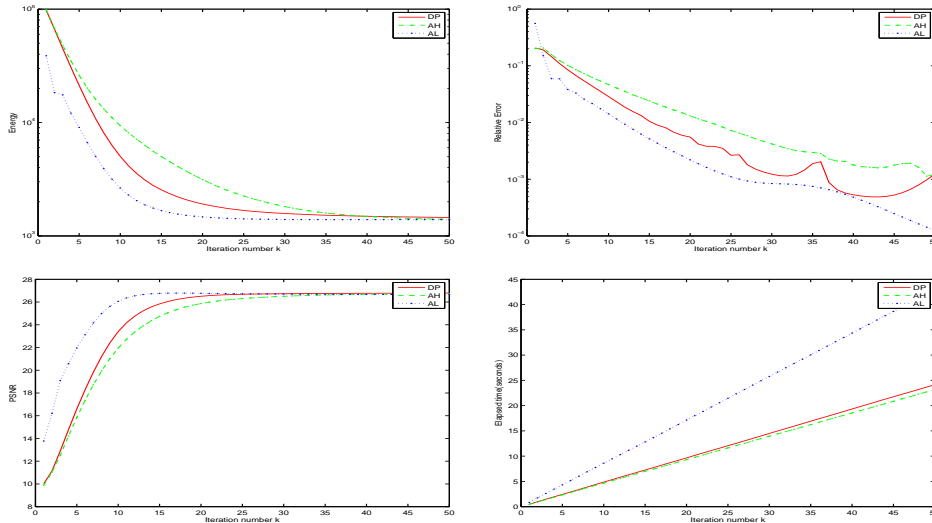


Figure 14: Data of animal.

compared to the other two approaches. This is because the dual projection is the fastest method in finishing one iteration and can recover the images with only few iterations.

5. Conclusions

The proposed algorithms can be easily applied to the high dimension manifolds without special difference schemes. That also can be extended to other models in image restoration areas including high order model proposed in [22]. If we add the pulse noise, the

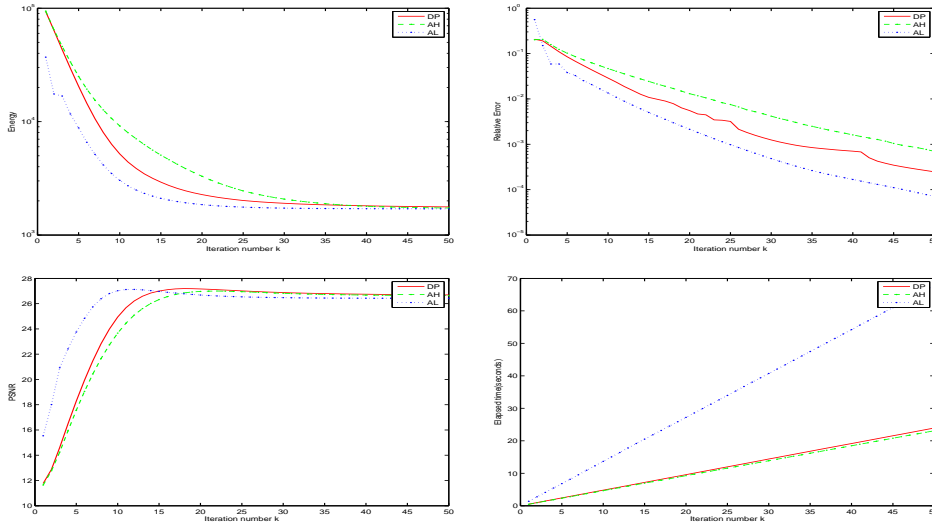


Figure 15: Data of fabric.

TV- L^1 model can be considered. It is observed that p -harmonic flow problems by adding the penalization $\frac{1}{\epsilon} \int_{\Omega} (1 - |U|^{s_1})^{s_2}$ in cases of $s_1 \geq 2$, $s_2 \geq 2$ are not easily solved. The V -subproblem leads to the algebra equation of higher order, which is hard to get an explicit solution. Thus, how to implement the algorithm fast and efficiently for the general penalizations will be considered in the future.

Acknowledgments The authors' research was supported by MOE IDM project NRF2007IDM-IDM002-010, Singapore. The first author was partially supported by PHD Program Scholarship Fund of ECNU with Grant No. 2010026 and Overseas Research Fund of East China Normal University, China. Discussions with Dr. Zhifeng Pang, Dr. Haixia Liang and Dr. Yuping Duan are helpful.

References

- [1] E. BAE AND X.-C. TAI, *Graph cut optimization for the piecewise constant level set method applied to multiphase image segmentation*, Scale Space and Variational Methods in Computer Vision, SSVN 2009, pp. 1–13.
- [2] X. BRESSON AND T. F. CHAN, *Fast dual minimization of the vectorial total variation norm and applications to color image processing*, Inverse Probl. Image., 2(4) (2008), pp. 455–484.
- [3] J. W. BARRETT, S. BARTELS, X. FENG AND A. PROHL, *A convergent and constraint-preserving finite element method for the p -harmonic flow into spheres*, SIAM J. Numer. Anal., 45(3) (2007), pp. 905–927.
- [4] S. BARTELS, *Stability and convergence of finite-element approximation schemes for harmonic maps*, SIAM J. Numer. Anal., 43(1) (2005), pp. 220–238.
- [5] S. BARTELS, *Numerical analysis of a finite element scheme for the approximation of harmonic maps into surfaces*, Math. Comput., 79(271) (2010), pp. 1263–1301.

- [6] S. BARTELS AND A. PROHL, *Convergence of an implicit finite element method for the Landau-Lifshitz-Gilbert equation*, SIAM J. Numer. Anal., 44(4) (2006), pp. 1405–1419.
- [7] S. BARTELS AND A. PROHL, *Convergence of an implicit, constraint preserving finite element discretization of p -harmonic heat flow into spheres*, Numer. Math., 109(4) (2008), pp. 489–507.
- [8] F. BETHUEL, H. BREZIS AND F. HÉLEIN, *Asymptotics for the minimization of a Ginzburg-Landau functional*, Calc. Var. Partial Differential Equations, 1(2) (1993), pp. 123–148.
- [9] F. BETHUEL, H. BREZIS AND F. HÉLEIN, *Singular limit for the minimization of Ginzburg-Landau functionals*, C. R. Acad. Sci. Paris Sér. I Math., 314(12) (1992), pp. 891–895.
- [10] T. CECIL, S. OSHER AND L. VESE, *Numerical methods for minimization problems constrained to S^1 and S^2* , J. Comput. Phys., 198(2) (2004), pp. 567–579.
- [11] A. CHAMBOLLE, *An algorithm for total variation minimization and applications*, J. Math. Imaging Vis., 20(1-2) (2004), pp. 89–97.
- [12] A. CHAMBOLLE AND T. POCK, *A first-order primal-dual algorithm for convex problems with applications to imaging*, J. Math. Imaging Vis., 40(1) (2010), pp. 120–145.
- [13] T. F. CHAN, G. H. GOLUB AND P. MULET, *A nonlinear primal-dual method for total variation-based image restoration*, SIAM J. Sci. Comput., 20(6) (1999), pp. 1964–1977.
- [14] R. COHEN, R. HARDT, D. KINDERLEHRER, S. Y. LIN AND M. LUSKIN, *Minimum energy configurations for liquid crystals: Computational results*, in *Theory and Applications of Liquid Crystals*, IMA Math. Appl., 5 (1987), pp. 99–121.
- [15] R. COHEN, S. Y. LIN AND M. LUSKIN, *Relaxation and gradient methods for molecular orientation in liquid crystals*, Comput. Phys. Commun., 53(1-3) (1989), pp. 455–465.
- [16] R. GLOWINSKI AND P. L. TALLEC, *Augmented Lagrangian and Operator-Splitting Methods in Nonlinear Mechanics*, SIAM, Philadelphia, 1989.
- [17] D. GOLDFARB, Z. W. WEN AND W. T. YIN, *A curvilinear search method for p -harmonic flows on spheres*, SIAM J. Imaging Sci., 2(1) (2009), pp. 84–109.
- [18] T. GOLDSTEIN AND S. OSHER, *The split Bregman method for $L1$ -regularized problems*, SIAM J. Imaging Sci., 2(2) (2009), pp. 323–343.
- [19] Q. HU, X.-C. TAI AND R. WINTHER, *A saddle point approach to the computation of harmonic maps*, SIAM J. Numer. Anal., 47(2) (2009), pp. 1500–1523.
- [20] SUNG HA KANG AND RICCARDO MARCH, *Variational models for image colorization via chromaticity and brightness decomposition*, IEEE T. Image Process., 16(9) (2007), pp. 2251–2261.
- [21] V. KWATRA, A. SCHODL, I. ESSA, G. TURK AND A. BOBICK, *Graphcut textures: image and video synthesis using graph cuts*, ACM Trans. Graphics, 22(3) (2003), pp. 277–286.
- [22] M. LYSAKER, S. OSHER AND X.-C. TAI, *Noise removal using smoothed normals and surface fitting*, IEEE T. Image Process., 13(10) (2004), pp. 1345–1357.
- [23] A. MARQUINA AND S. OSHER, *Image super-resolution by TV-regularization and Bregman iteration*, J. Sci. Comput., 37(3) (2008), pp. 367–382.
- [24] M. MISAWA, *Approximation of p -harmonic maps by the penalized equation*, Nonlinear Anal., 47(2) (2001), pp. 1069–1080.
- [25] C. ROTHER, V. KOLMOGOROV AND A. BLAKE, *"GrabCut": interactive foreground extraction using iterated graph cuts*, ACM Trans. Graphics, 23(3) (2004), pp. 309–314.
- [26] L. I. RUDIN, S. OSHER AND E. FATEMI, *Nonlinear total variation based noise removal algorithms*, Phys. D, 60 (1992), pp. 259–268.
- [27] S. SETZER, *Splitting Bregman algorithm, Douglas-Rachford splitting and frame shrinkage*, in *SSVM 09: Proceedings of the Second International Conference on Scale Space and Variational Methods in Computer Vision*, Springer-Verlag, Berlin, Heidelberg, pp. 464–476, 2009.
- [28] L. A. VESE AND S. J. OSHER, *Numerical methods for p -harmonic flows and applications to image processing*, SIAM J. Numer. Anal., 40(6) (2002), pp. 2085–2104.

- [29] Z. WEN AND W. YIN, *A feasible method for optimization with orthogonality constraints*, CAM Report 10-77, UCLA, Los Angeles, CA, 2010.
- [30] C. L. WU AND X.-C. TAI, *Augmented Lagrangian method, Dual methods and Split-Bregman Iterations for ROF, vectorial TV and higher order models*, SIAM J. Imaging Sci., 3(3) (2010), pp. 300–339.
- [31] X.-C. TAI, J. Y. HAHN AND G. J. CHUNG, *A fast algorithm for Euler's elastica model using augmented lagrangian*, SIAM J. Imaging Sci., 4(1) (2011), pp. 313–344.
- [32] W. YIN, S. OSHER, D. GOLDFARB AND F. DARBON, *Bregman iterative algorithms for l_1 -minimization with applications to compressed sensing*, SIAM J. Imaging Sci., 1(1) (2008), pp. 143–168.
- [33] X. ZHANG, M. BURGER AND S. OSHER, *A Unified Primal-Dual Algorithm Framework Based on Bregman Iteration*, UCLA CAM Report 09-99, UCLA, Los Angeles, CA, 2009.
- [34] M. ZHU AND T. CHAN, *An Efficient Primal-Dual Hybrid Gradient Algorithm for Total Variation Image Restoration*, UCLA CAM Report 08-34, UCLA, Los Angeles, CA, 2008.