Constrained Nonconvex Hybrid Variational Model for Edge-Preserving Image Restoration

Wen Liu

Work with Di Wu, Chuan-Sheng Wu, Tian Xu and Naixue Xiong

School of Navigation, Wuhan University of Technology

wenliu@whut.edu.cn

October 11, 2015
Overview

1 Introduction

2 Constrained Nonconvex Hybrid Variational Model

3 Numerical Optimization Algorithm
   - Alternating Direction Method of Multipliers (ADMM)
   - ADMM-Based Iteratively Reweighted Algorithm

4 Numerical Results and Discussion
   - Experimental Settings
   - Quantitative Image Quality Assessment
   - Visual Image Quality Assessment
   - Algorithm Convergence

5 Conclusions
Real-World Problem

The degraded observation of a latent image $f$ can be expressed as a linear shift-invariant system as follows

$$g = Hf + \epsilon,$$  \hspace{1cm} (1)

where $g$ is the degraded image, $\epsilon$ denotes the Gaussian white noise, and $H$ represents the convolution operator.

Mathematical Modelling

The degraded observation of a latent image $f$ can be expressed as a linear shift-invariant system as follows

$$g = Hf + \epsilon,$$  \hspace{1cm} (1)

where $g$ is the degraded image, $\epsilon$ denotes the Gaussian white noise, and $H$ represents the convolution operator.
Introduction

- **Total Variation (TV)**
  - **Advantages**
    1. Edge-preserving reconstruction
    2. Effective noise suppression
  - **Disadvantages**
    1. Staircase-like artifacts in homogeneous regions
    2. Unsuitable for images with fine structures and textures

- **Second-Order TV**
  - **Advantages**: Suppressing staircase-like artifacts
  - **Disadvantages**: Generating blurred edges and boundaries

- **Combined First- and Second-Order TV**

\[
\min_{f} \left\{ \frac{\lambda}{2} \| H f - g \|^2_2 + \omega_1 \| \nabla f \|_2 + \omega_2 \| \nabla^2 f \|_2 \right\},
\]

where \( \lambda, \omega_1 \) and \( \omega_2 \) are regularization parameters, \( \nabla \) and \( \nabla^2 \) denote the discrete differential operators of the first- and second-order, respectively.
Introduction

- **Combined First- and Second-Order TV**
  - **Advantages**: Maintaining a balance between artifacts suppression and edges preservation
  - **Disadvantages**: Mismatching the distribution of image gradients

- **Motivation**
  - Both the first- and second-order gradients follow the hyper-Laplacian distribution, \( p(x) \propto e^{-\tau|x|^\pi} \) with \( 0 < \pi < 1 \).
  
  \[
  \omega_1 \| \nabla f \|_2^2 + \omega_2 \| \nabla^2 f \|_2^2 \Rightarrow \omega_1 \| \nabla f \|_2^{\pi_1} + \omega_2 \| \nabla^2 f \|_2^{\pi_2}
  \]

  - Nonconvex minimization could provide better image restoration.
  - Adaptive regularization parameter \( \lambda \) plays an important role of improving image restoration performance.
  - Solutions yielded by traditional models easily move out of the given dynamic range (e.g., \([0, 255]\) for 8-bit images and \([0, 1]\) for normalized images).
Constrained Nonconvex Hybrid Variational Model

- **Nonconvex Hybrid Variational Model – NHTV**

\[
\min_f \left\{ \frac{\lambda}{2} \| Hf - g \|_2^2 + \omega \| \nabla f \|_2^{\pi_1} + (1 - \omega) \| \nabla^2 f \|_2^{\pi_2} \right\},
\]

where \( \pi_1, \pi_2 \in (0, 1) \) related to the image gradient distribution, and \( \omega \in [0, 1] \) denotes the adaptive weighting function.

- **Constrained Nonconvex Hybrid Variational Model – CNHTV**

\[
\min_{f \in \mathcal{B}} \left\{ \frac{\lambda}{2} \| Hf - g \|_2^2 + \omega \| \nabla f \|_2^{\pi_1} + (1 - \omega) \| \nabla^2 f \|_2^{\pi_2} \right\},
\]

where the box constraint \( \mathcal{B} = \{ f \in \mathcal{R} | f_{\text{min}} \leq f(o) \leq f_{\text{max}}, \forall o \in \Omega \} \) with denoting the dynamic range of \( f \) to be \([f_{\text{min}}, f_{\text{max}}]\).

achieving good results for pixels with values lying on the predefined dynamic range boundaries.
Given a constrained optimization problem

$$\min_{x,z} f(x) + g(z)$$

s.t. \(Ax + Bz = c\). \hspace{1cm} (5)$$

where \(x \in \mathbb{R}^{d_1}, z \in \mathbb{R}^{d_2}, A \in \mathbb{R}^{m \times d_1}, B \in \mathbb{R}^{m \times d_2}, c \in \mathbb{R}^m\), both \(f(\cdot)\) and \(g(\cdot)\) are assumed to be convex. The augmented Lagrangian function of (5) is given by

$$\mathcal{L}_A(x, z; \rho) = f(x) + g(z) + \frac{\beta}{2} \left\| (Ax + Bz - c) - \frac{\rho}{\beta} \right\|_2^2,$$

where \(\rho\) and \(\beta > 0\) denote the augmented Lagrange multiplier and pre-defined penalty parameter, respectively. At the \(k\)th iteration, ADMM attempts to solve problem (5) by iteratively minimizing \(\mathcal{L}_A(x, z; \rho)\) over \(x\) and \(z\), respectively, and updating \(\rho\) accordingly.
ADMM-Based Iteratively Reweighted Algorithm

- **Original nonconvex variational model**

\[
\min_{f \in B} \left\{ \frac{\lambda}{2} \|Hf - g\|_2^2 + \omega \|\nabla f\|_2^{\pi_1} + (1 - \omega) \|\nabla^2 f\|_2^{\pi_2} \right\}
\]  
(6)

- **Equivalent convex variational model**

\[
\min_{f \in B} \left\{ \frac{\lambda}{2} \|Hf - g\|_2^2 + \omega_k \psi_k^1 \|\nabla f\|_2 + (1 - \omega_k) \psi_k^2 \|\nabla^2 f\|_2 \right\}
\]  
(7)

where variables \( \psi_k^1 = \|\nabla f_k\|_2^{\pi_1 - 1} \) and \( \psi_k^2 = \|\nabla^2 f_k\|_2^{\pi_2 - 1} \). The unconstrained minimization problem (7) was transformed into the following constrained form

\[
\min_{u, v, w, k, f} \left\{ \frac{\lambda}{2} \|k - g\|_2^2 + \omega \psi^1 \|v\|_2 + (1 - \omega) \psi^2 \|w\|_2 \right\}
\]  
(8)

s.t. \( u = f, \ v = \nabla f, \ w = \nabla^2 f, \ k = Hf \).
Augmented Lagrangian Function

\[
\mathcal{L}_A (u, v, w, k, f; \tau, \mu, \xi, \varphi) = \frac{\lambda}{2} \|k - g\|^2_2 + \frac{\beta_1}{2} \|u - f\|^2_2 - \langle \tau, u - f \rangle + \frac{\beta_2}{2} \|v - \nabla f\|^2_2 \\
- \langle \mu, v - \nabla f \rangle + \omega \psi^1 \|v\|_2 + \frac{\beta_3}{2} \|w - \nabla^2 f\|^2_2 - \langle \xi, w - \nabla^2 f \rangle \\
+ (1 - \omega) \psi^2 \|w\|_2 + \frac{\beta_4}{2} \|k - Hf\|^2_2 - \langle \varphi, k - Hf \rangle
\]  

(9)

where \(\tau, \mu, \xi\) and \(\varphi\) represent the augmented Lagrangian multipliers, and \(\beta_1, \beta_2, \beta_3, \beta_4 > 0\) are penalty parameters which control the weights of penalty terms.

**subproblems** with respect to \(u, v, w, k\) and \(f\).
**u-subproblem**

The \( u \)-subproblem \( u_{k+1} \leftarrow \min_u \mathcal{L}_A (u, v_k, w_k, k, f_k, \tau_k, \mu_k, \xi_k, \varphi_k) \) can be solved by considering the following minimization problem

\[
\begin{align*}
    u_{k+1} &= \min_{u \in \mathcal{B}} \left\{ \frac{\beta_1}{2} \left\| u - \left( f_k + \frac{\tau_k}{\beta_1} \right) \right\|_2^2 \right\}, \\
    & \quad \text{where } \mathcal{B} = \mathbb{R} \setminus \{ f_{\text{min}} \}, \mathcal{B} = \mathbb{R} \setminus \{ f_{\text{max}} \}.
\end{align*}
\]

The solution of (10) can be simply obtained with a projection \( P_{\mathcal{B}} \) onto the box constraint \( \mathcal{B} \), i.e., \( u_{k+1} = P_{\mathcal{B}} (\tilde{u}_k) \) with \( \tilde{u}_k = f_k + \frac{\tau_k}{\beta_1} \).

The projection operator \( P_{\mathcal{B}} \) onto \( \mathcal{B} \) is defined as follows

\[
P_{\mathcal{B}} (\tilde{u}_k (o)) = \begin{cases} 
    f_{\text{min}}, & \text{if } \tilde{u}_k (o) < f_{\text{min}}, \\
    \tilde{u}_k (o), & \text{if } \tilde{u}_k (o) \in [f_{\text{min}}, f_{\text{max}}], \\
    f_{\text{max}}, & \text{if } \tilde{u}_k (o) > f_{\text{max}},
\end{cases}
\]

for any pixel \( o \in \Omega \).
**v-subproblem**

The $v$-subproblem $v_{k+1} \leftarrow \min_v \mathcal{L}_A (u_{k+1}, v, w_k, k, f_k; \tau_k, \mu_k, \xi_k, \phi_k)$ can be effectively solved through the **shrinkage** operation. Let $\tilde{v}_k^f = \nabla f_k + \frac{\mu_k}{\beta^2}$, minimization of the augmented Lagrangian function $\mathcal{L}_A$ with respect to $v$ is equivalent to

$$v_{k+1} = \min_v \left\{ \frac{\beta^2}{2} \left\| v - \tilde{v}_k^f \right\|_2^2 + \omega_k \psi_k^1 \left\| v \right\|_2 \right\}, \quad (12)$$

whose solution $v_{k+1}$ is given by

$$v_{k+1} = \text{shrinkage} \left( \frac{\tilde{v}_k^f}{\beta^2}, \frac{\omega_k \psi_k^1}{\beta^2} \right) \quad (13)$$

$$= \max \left( \left\| \tilde{v}_k^f \right\|_2 - \frac{\omega_k \psi_k^1}{\beta^2}, 0 \right) \circ \frac{\tilde{v}_k^f}{\left\| \tilde{v}_k^f \right\|_2},$$

where $\circ$ denotes the pointwise multiplication operator.
**w-subproblem**

The $w$-subproblem

\[ w_{k+1} \leftarrow \min_w \mathcal{L}_A (u_{k+1}, v_{k+1}, w, k, f_k; \tau_k, \mu_k, \xi_k, \varphi_k) \] in (9) can be expressed as the following formulation

\[
\begin{align*}
    w_{k+1} &= \min_w \left\{ \frac{\beta_3}{2} \left\| w - \tilde{w}_k^f \right\|^2_2 + (1 - \omega_k) \psi_k^2 \left\| w \right\|_2 \right\} \\
\end{align*}
\]

(14)

with $\tilde{w}_k^f = \nabla^2 f_k + \frac{\xi_k}{\beta_3}$. Analogous to the $v$-subproblem, solution $w_{k+1}$ can be obtained as follows

\[
\begin{align*}
    w_{k+1} &= \max \left( \left\| \tilde{w}_k^f \right\|_2, 0 \right) \circ \frac{\tilde{w}_k^f}{\left\| \tilde{w}_k^f \right\|_2}. \\
\end{align*}
\]

(15)
**k-subproblem**

Given the fixed $f_k$ and $\xi_k$, solution $k_{k+1}$ of the $k$-subproblem is equivalent to solving the following minimization problem

$$
k_{k+1} = \min_{k} \left\{ \frac{\lambda_k}{2} \| k - g \|_2^2 + \frac{\beta_4}{2} \| k - k_{k}^f \|_2^2 \right\},
$$

with $k_{k}^f = Hf_k + \frac{\xi_k}{\beta_4}$. Solution of this least-squares optimization problem (16) can be efficiently obtained by

$$
k_{k+1} = \mathcal{F}^{-1} \left( \frac{\lambda_k \mathcal{F} (g) + \beta_4 \mathcal{F} (k_{k}^f)}{\lambda_k + \beta_4} \right),
$$

where $\mathcal{F} (\circ)$ and $\mathcal{F}^{-1} (\circ)$, respectively, represent the forward Fast Fourier Transform (FFT) and inverse FFT.
ADMM-Based Iteratively Reweighted Algorithm

Adaptive Parameter $\lambda$ Selection – Discrepancy Principle

1. If $\|k^f_k - g\|^2 \leq cL\sigma^2$
   
   where $k^f_k = Hf_k + \xi_k / \beta_4$, $c$ is a noise-dependent parameter, $L$ and $\sigma^2$ respectively denote the image size and pre-estimated noise variance. It is easy to obtain that $\lambda_k = 0$ and $k_{k+1} = k^f_k$ which obviously satisfies the discrepancy principle.

2. If $\|k^f_k - g\|^2 > cL\sigma^2$
   
   The parameter $\lambda_k$ can be adaptively estimated by solving the following equation
   
   $$\|k_{k+1} - g\|^2 = cL\sigma^2,$$
   
   according to the discrepancy principle. By combining
   
   $k_{k+1} = \mathcal{F}^{-1} \left( \frac{\lambda_k \mathcal{F}(g) + \beta_4 \mathcal{F}(k^f_k)}{\lambda_k + \beta_4} \right)$
   
   and $\|k_{k+1} - g\|^2 = cL\sigma^2$, an automatic estimation of $\lambda_k$ is given by
   
   $$\lambda_k = \frac{\beta_4 \|k^f_k - g\|^2}{\sqrt{cL\sigma}} - \beta_4.$$

Wen Liu (WHUT)

Edge-Preserving Image Restoration

October 11, 2015 14 / 27
ADMM-Based Iteratively Reweighted Algorithm

- **f-subproblem**
  Solution of the f-subproblem can be obtained by considering the following least-squares optimization problem

  \[
  f_{k+1} = \min_f \left\{ \frac{\beta_1}{2} \| f - \tilde{f}_{k+1}^u \|^2 + \frac{\beta_2}{2} \| \nabla f - \tilde{f}_{k+1}^v \|^2 + \frac{\beta_3}{2} \| \nabla^2 f - \tilde{f}_{k+1}^w \|^2 + \frac{\beta_4}{2} \| Hf - \tilde{f}_{k+1}^k \|^2 \right\}
  \]

  where \( \tilde{f}_{k+1}^u = u_{k+1} - \frac{\tau_k}{\beta_1} \), \( \tilde{f}_{k+1}^v = v_{k+1} - \frac{\mu_k}{\beta_2} \), \( \tilde{f}_{k+1}^w = w_{k+1} - \frac{\xi_k}{\beta_3} \) and \( \tilde{f}_{k+1}^k = k_{k+1} - \frac{\zeta_k}{\beta_4} \). FFT-based numerical method can be exploited to solve the f-subproblem. The least-squares solution of f can be obtained as

  \[
  \left( \beta_1 I + \beta_2 \nabla^\top \nabla + \beta_3 (\nabla^2)^\top \nabla^2 + \beta_4 H^\top H \right) f = \beta_1 \tilde{f}_{k+1}^u + \beta_2 \nabla^\top \tilde{f}_{k+1}^v + \beta_3 (\nabla^2)^\top \tilde{f}_{k+1}^w + \beta_4 H^\top \tilde{f}_{k+1}^k,
  \]

  where \( I \) is an identity matrix, and superscript \( \top \) denotes the transpose (conjugate transpose) operator for real (complex) matrices or vectors.
ADMM-Based Iteratively Reweighted Algorithm

- **τ, μ, ξ and ϕ update**
  
  At each iteration, we update the augmented Lagrangian multiplies τ, μ, ξ and ϕ as follows

  \[
  \begin{align*}
  \tau_{k+1} &= \tau_k - \gamma \beta_1 (u_{k+1} - f_{k+1}), \\
  \mu_{k+1} &= \mu_k - \gamma \beta_2 (v_{k+1} - \nabla f_{k+1}), \\
  \xi_{k+1} &= \xi_k - \gamma \beta_3 (w_{k+1} - \nabla^2 f_{k+1}), \\
  \varphi_{k+1} &= \varphi_k - \gamma \beta_4 (k_{k+1} - Hf_{k+1}),
  \end{align*}
  \]

  where the steplength \( \gamma = 1.618 \) is used throughout experiments.
**Experimental Settings**

- **Grayscale Test Images**

  ![Test Images](image)

  (a) Phantom  (b) Text  (c) Lena  
  (d) Barbara  (e) Cameraman  (f) House

**Figure:** Different test images of size $256 \times 256$. From top-left to bottom-right: (a) Phantom, (b) Text, (c) Lena, (d) Barbara, (e) Cameraman and (f) House.
**Experimental Settings**

- **Comparison with Other Image Restoration Methods**

<table>
<thead>
<tr>
<th>Model</th>
<th>$\lambda$</th>
<th>$\pi_1$</th>
<th>$\pi_2$</th>
<th>$\omega_1$</th>
<th>$\omega_2$</th>
<th>$B$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CTV</td>
<td>constant</td>
<td>1.0</td>
<td>no</td>
<td>constant</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>HTV</td>
<td>constant</td>
<td>1.0</td>
<td>1.0</td>
<td>constant</td>
<td>constant</td>
<td>no</td>
</tr>
<tr>
<td>NHTV</td>
<td>constant</td>
<td>0.9</td>
<td>0.7</td>
<td>adaptive</td>
<td>adaptive</td>
<td>no</td>
</tr>
<tr>
<td>CNHTV-C</td>
<td>constant</td>
<td>0.9</td>
<td>0.7</td>
<td>adaptive</td>
<td>adaptive</td>
<td>yes</td>
</tr>
<tr>
<td>CNHTV-A</td>
<td>adaptive</td>
<td>0.9</td>
<td>0.7</td>
<td>adaptive</td>
<td>adaptive</td>
<td>yes</td>
</tr>
</tbody>
</table>


A singular value decomposition (SVD)-based method was used to estimate the noise variance $\sigma^2$ for adaptively updating $\lambda$. 
Quantitative Image Quality Assessment

- Gaussian blur kernel: `fspecial('gaussian', [9, 9], 3).
- Gaussian white noise: $\sigma \in \{1/255, 5/255, 10/255, 15/255, 20/255\}$.

Figure: Quantitative image quality evaluation (PSNR).
Visual Image Quality Assessment

- Imaging Conditions: $9 \times 9$ Gaussian blur kernel and Gaussian noise with $\sigma = 10/255$.
- Image “Lena”: Texture-like hair region and predominantly homogeneous background.

![Degraded](image1)
![CTV](image2)
![HTV](image3)

![NHTV](image4)
![CNHTV-C](image5)
![CNHTV-A](image6)
Imaging Conditions: $9 \times 9$ Gaussian blur kernel and Gaussian noise with $\sigma = 10/255$.

Image “Lena”: Texture-like hair region and predominantly homogeneous background.
Visual Image Quality Assessment

- Imaging Conditions: $9 \times 9$ Gaussian blur kernel and Gaussian noise with $\sigma = 10/255$.
- Image “Text”: Many pixels with values at the dynamic range boundaries.

![Images showing degraded and restored images](image-url)
The superior results benefit from the adaptive estimation of regularization parameter $\lambda$.

**Figure:** Regularization parameter $\lambda$ estimated adaptively for different test images under different degradation conditions.
Empirical Evidence

It is difficult to establish the theoretical convergence result since the proposed objective function is both nonsmooth and nonconvex.

Figure: Convergence property of our proposed method.
Conclusions

**Constrained Nonconvex Hybrid Variational Model**
This model is capable of suppressing the staircase artifacts and preserving the valuable edge information. The box constraint provides a visible positive effect on image restoration, especially when there are many pixels with values lying on the predefined dynamic range boundaries.

**Adaptive Regularization Parameter Selection**
The discrepancy principle was exploited to guarantee the optimal adaptive estimation of regularization parameter to further improve image quality.

**Numerical Optimization Algorithm**
A iteratively reweighted algorithm based on alternating direction method of multipliers (ADMM) was presented to efficiently solve the constrained nonconvex optimization problem.
References


Thank you for your attention!