Convex and Non-Convex Optimization in Image Recovery and Segmentation

Tieyong Zeng
Dept. of Mathematics, HKBU

29 May - 2 June, 2017
NUS, Singapore
Outline

1. Variational Models for Rician Noise Removal
2. Two-stage Segmentation
3. Dictionary and Weighted Nuclear Norm
Background

We assume that the noisy image $f$ is obtained from a perfect unknown image $u$

$$f = u + b.$$ 

- $b$: additive Gaussian noise.
- TV-ROF:

$$\min_u \frac{\mu}{2} \int_{\Omega} (u - f)^2 dx + \int_{\Omega} |Du|,$$

where the constant $\mu > 0$, $\Omega \subset \mathbb{R}^n$.

- The MAP estimation leads to the data fitting term $\int_{\Omega} (u - f)^2$.
- Smoothness from the edge-preserving total variation (TV) term, $\int_{\Omega} |Du|$.
Rician noise with Blur

The degradation process reads

\[ f = \sqrt{(Au + \eta_1)^2 + \eta_2^2}. \]

- \( A \) describes the blur operator
- The Rician noise was built from white Gaussian noise \( \eta_1, \eta_2 \sim \mathcal{N}(0, \sigma^2) \)
- The distribution has probability density

\[ p(f|u) = \frac{f}{\sigma^2} e^{-\frac{(Au)^2 + f^2}{2\sigma^2}} I_0\left(\frac{A uf}{\sigma^2}\right). \]
TV-model for removing Rician noise with blur

The MAP estimation leads to:

\[
\inf_u \frac{1}{2\sigma^2} \int_\Omega (Au)^2 \, dx - \int_\Omega \log l_0 \left( \frac{Afu}{\sigma^2} \right) \, dx + \gamma \text{TV}(u),
\]

- \( l_0 \) is the solution of the zero order modified Bessel function
  \[ xy'' + y' - xy = 0. \]
- the TV prior can recover sharp edges, the constant \( \gamma > 0 \)
- Non-convex!
Convex variational model for denoising and deblurring

[Getreuer-Tong-Vese, ISVC, 2011]

\[
\inf_u \int_\Omega G_\sigma(Au, f) + \gamma TV(u).
\] (1)

Let \( z = Au \) and \( c = 0.8246 \),

\[
G_\sigma(z, f) = \begin{cases} 
H_\sigma(z) & \text{if } z \geq c\sigma, \\
H_\sigma(c\sigma) + H'_\sigma(c\sigma)(z - c\sigma) & \text{if } z \leq c\sigma,
\end{cases}
\]

\[
H_\sigma(z) = \frac{f^2 + z^2}{2\sigma^2} - \log I_0\left(\frac{fz}{\sigma^2}\right)
\]

\[
H'_\sigma(z) = \frac{z}{\sigma^2} - \frac{f}{\sigma^2} B\left(\frac{fz}{\sigma^2}\right)
\]

\[
B(t) = \frac{l_1(t)}{l_0(t)} \approx \frac{t^3 + 0.950037t^2 + 2.38944t}{t^3 + 1.48937t^2 + 2.57541t + 4.65314}.
\]

- Complex and difficult to derive its mathematical property.
New elegant convex TV-model

We introduce a quadratic penalty term:

\[
\inf_u \frac{1}{2\sigma^2} \int_{\Omega} (Au)^2 dx - \int_{\Omega} \log l_0 \left( \frac{Auf}{\sigma^2} \right) dx + \frac{1}{\sigma} \int_{\Omega} (\sqrt{Au} - \sqrt{f})^2 dx + \gamma \text{TV}(u),
\]  

(2)
The quadratic penalty term

Why the penalty term $\int_\Omega \frac{(\sqrt{f} - \sqrt{Au})^2}{\sigma} \, dx$?

- the value of $e$ is always bounded, where $e$ is defined below.

**Proposition 1** Suppose that the variables $\eta_1$ and $\eta_2$ independently follow the Normal distribution $\mathcal{N}(0, \sigma^2)$. Set $f = \sqrt{(u + \eta_1)^2 + \eta_2^2}$ where $u$ is fixed and $u \geq 0$. Then we can get the following inequality,

$$e := \frac{\mathbb{E}((\sqrt{f} - \sqrt{u})^2)}{\sigma} \leq \sqrt{\frac{2}{\pi}} (\pi + 2).$$
Proof of the Proposition

**Lemma 1** Assume that $a, b \in \mathbb{R}$. Then, 
$|(u^2 + 2au + b^2)^{\frac{1}{4}} - u^{\frac{1}{2}}| \leq \sqrt{|a|} + \sqrt{|b|}$ is true whenever $u \geq 0$ and $|a| \leq |b|$.

- Based on the lemma, we can get

$$E((\sqrt{f} - \sqrt{u})^2) \leq 2E(|\eta_1|) + 2E((\eta_1^2 + \eta_2^2)^{\frac{1}{2}}).$$

- Set $Y := \frac{\eta_1^2 + \eta_2^2}{\sigma^2}$, where $\eta_1, \eta_2 \sim \mathcal{N}(0, \sigma^2)$. Thus, $Y \sim \chi^2(2)$.

- It can be shown that

$$E(\sqrt{Y}) = \frac{E((\eta_1^2 + \eta_2^2)^{\frac{1}{2}})}{\sigma} = \frac{\sqrt{2\pi}}{2},$$

$$E(|\eta_1|) = \sqrt{\frac{2}{\pi}} \sigma.$$
Approximation table

**Table:** The values of $e$ with various values of $\sigma$ for different original images $u$.

<table>
<thead>
<tr>
<th>Image</th>
<th>$\sigma = 5$</th>
<th>$\sigma = 10$</th>
<th>$\sigma = 15$</th>
<th>$\sigma = 20$</th>
<th>$\sigma = 25$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cameraman</td>
<td>0.0261</td>
<td>0.0418</td>
<td>0.0571</td>
<td>0.0738</td>
<td>0.0882</td>
</tr>
<tr>
<td>Bird</td>
<td>0.0131</td>
<td>0.0255</td>
<td>0.0371</td>
<td>0.0486</td>
<td>0.0590</td>
</tr>
<tr>
<td>Skull</td>
<td>0.0453</td>
<td>0.0754</td>
<td>0.1009</td>
<td>0.1234</td>
<td>0.1425</td>
</tr>
<tr>
<td>Leg joint</td>
<td>0.0356</td>
<td>0.0654</td>
<td>0.0906</td>
<td>0.1105</td>
<td>0.1263</td>
</tr>
</tbody>
</table>
The quadratic penalty term

Why the penalty term $\int_\Omega \frac{(\sqrt{f} - \sqrt{Au})^2}{\sigma} \, dx$?

- By adding this term, we obtain a strictly convex model (2) for restoring blurry images with Rician noise.
How to prove the convexity of the proposed model (2)?

- Let us define a function $g_0$ as

  $$g_0(t) := -\log l_0(t) - 2\sqrt{t}.$$  

- If $g_0(t)$ is strictly convex on $[0, +\infty)$, then the convexity of the first three terms in model (2) as a whole can be easily proved by letting $t = \frac{f(x)Au(x)}{\sigma^2}$.

- Since $\int_\Omega |\nabla u|$ is convex, the proposed model is strictly convex.
Convexity of the proposed model (2)

How to prove the strict convexity of function $g(t)$?

$g''_0(t) = -\frac{(l_0(t) + l_2(t))l_0(t) - 2l_1^2(t)}{2l_0^2(t)} + \frac{1}{2} t^{-\frac{3}{2}}.$

*Here, $I_n(t)$ is the modified Bessel functions of the first kind with order $n$.***

Based on the following lemma, we can prove that $g_0(t)$ is strictly convex.

**Lemma 2** Let $h(t) = t^{\frac{3}{2}} \frac{(l_0(t) + l_2(t))l_0(t) - 2l_1^2(t)}{l_0^2(t)}$. Then $0 \leq h(t) < 1$ on $[0, +\infty)$. 
Existence and uniqueness of a solution

**Theorem 1** Let $f$ be in $L^\infty(\Omega)$ with $\inf_{\Omega} f > 0$, then the model (2) has a unique solution $u^*$ in $BV(\Omega)$ satisfying

$$0 < \frac{\sigma^2}{(2\sup_{\Omega} f + \sigma)^2} \inf_{\Omega} f \leq u^* \leq \sup_{\Omega} f.$$ 

Set $c_1 := \frac{\sigma^2}{(2\sup_{\Omega} f + \sigma)^2} \inf_{\Omega} f$, $c_2 := \sup_{\Omega} f$, and define two functions as follows,

$$E_0(u) := \frac{1}{2\sigma^2} \int_{\Omega} u^2 dx - \int_{\Omega} \log l_0(\frac{fu}{\sigma^2}) dx + \frac{1}{\sigma} \int_{\Omega} (\sqrt{u} - \sqrt{f})^2 dx,$$

$$E_1(u) := E_0(u) + \gamma \int_{\Omega} |Du| dx,$$

where $E_1(u)$ is the objective function in model (2).
Existence and uniqueness of a solution

- According to the integral forms of the modified Bessel functions of the first kind with integral orders $n$, we have

$$I_0(x) = \frac{1}{\pi} \int_0^\pi e^{x \cos \theta} d\theta \leq e^x, \quad \forall x \geq 0, \quad (4)$$

thus, for each fixed $x \in \Omega$, $-\log I_0 \left( \frac{f(x)t}{\sigma^2} \right) \geq -\frac{f(x)t}{\sigma^2}$ with $t \geq 0$.

- $E_1(u)$ in (3) is bounded below.

$$E_1(u) \geq E_0(u) \geq \frac{1}{2\sigma^2} \int_\Omega u^2 \, dx - \int_\Omega \log I_0 \left( \frac{fu}{\sigma^2} \right) \, dx$$

$$\geq \int_\Omega \left( \frac{1}{2\sigma^2} u^2 - \frac{fu}{\sigma^2} \right) \, dx$$

$$\geq - \frac{1}{2\sigma^2} \int_\Omega f^2 \, dx.$$
Existence and uniqueness of a solution

- For each fixed $x \in \Omega$, let the real function $g$ on $\mathbb{R}^+ \cup \{0\}$ be defined as
  
  $$g(t) := \frac{1}{2\sigma^2} t^2 - \log I_0 \left( \frac{f(x)t}{\sigma^2} \right) + \frac{1}{\sigma} \left( \sqrt{t} - \sqrt{f(x)} \right)^2.$$

  $$g'(t) = \frac{1}{\sigma^2} t - \frac{f(x)}{\sigma^2} \frac{l_1 \left( \frac{f(x)t}{\sigma^2} \right)}{l_0 \left( \frac{f(x)t}{\sigma^2} \right)} + \frac{1}{\sigma} \left( 1 - \sqrt{\frac{f(x)}{t}} \right).$$

- We can prove that $g(t)$ is increasing if $t \in (f(x), +\infty)$ and decreasing if $0 \leq t < \frac{\sigma^2}{(2f(x)+\sigma)^2} f(x)$. This implies that $g(\min(t, V)) \leq g(t)$ if $V \geq f(x)$.

- With $\int_{\Omega} |D \inf(u, c_2)| \leq \int_{\Omega} |Du|$ [Kornprobst,Deriche,Aubert 1999], we have
  
  $$E_1(\inf(u, c_2)) \leq E_1(u).$$

  Similarly, we can get $E_1(\sup(u, c_1)) \leq E_1(u)$.

- The unique solution $u^*$ to the model (2) should be restricted in $[c_1, c_2]$. 
Lemma 3 The function $l_0(x)$ is strictly log-convex for all $x > 0$.

- In order to prove that the function $l_0(x)$ is strictly log-convex in $(0, +\infty)$, it suffices to show that its logarithmic second-order derivative is positive in $(0, +\infty)$

$$
\left(\log l_0(x)\right)'' = \frac{\frac{1}{2}(l_0(x) + l_2(x))l_0(x) - l_1(x)^2}{l_0^2(x)}.
$$

- Using Cauchy-Schwarz inequality, we obtain

$$
\frac{1}{2}(l_0(x) + l_2(x))l_0(x) = \frac{1}{\pi} \int_0^\pi \cos^2 \theta e^x \cos \theta \, d\theta \cdot \frac{1}{\pi} \int_0^\pi e^x \cos \theta \, d\theta \\
\geq \left(\frac{1}{\pi} \int_0^\pi \cos \theta e^x \cos \theta \, d\theta\right)^2 = (l_1(x))^2.
$$

Since $\cos \theta e^{\frac{1}{2}x} \cos \theta$ and $e^{\frac{1}{2}x} \cos \theta$ are not linear dependent when $\theta$ changes, the strict inequality in above holds.
Lemma 4 Let $g(x)$ be a strictly convex and strictly increasing function in $(0, +\infty)$. Meanwhile, let $g(x)$ be differentiable. Assume that $0 < a < b$, $0 < c < d$, then we have:

$$g(ac) + g(bd) > g(ad) + g(bc).$$

Based on Theorem 1, Lemma 3 and Lemma 4, we can further establish the following comparison principle (minimum-maximum principle).
**Proposition 3** Let $f_1$ and $f_2$ be in $L^\infty(\Omega)$ with $\inf_\Omega f_1 > 0$ and $\inf_\Omega f_2 > 0$. Suppose $u_1^*$ (resp. $u_2^*$) is a solution of model (2) with $f = f_1$ (resp. $f = f_2$). Assume that $f_1 < f_2$, then we have $u_1^* \leq u_2^*$ a.e. in $\Omega$.

- Note $u_1^* \wedge u_2^* := \inf(u_1^*, u_2^*)$, $u_1^* \vee u_2^* := \sup(u_1^*, u_2^*)$.
- $E_i^1(u)$ denotes $E_1(u)$ defined in (3) with $f = f_i$.
- Since $u_1^*$ (resp. $u_2^*$) is a solution of model (2) with $f = f_1$ (resp. $f = f_2$), we can easily get

\[
E_1^1(u_1^* \wedge u_2^*) \geq E_1^1(u_1^*),
\]

\[
E_1^2(u_1^* \vee u_2^*) \geq E_1^2(u_2^*).
\]
Minimum-maximum Principle

- Adding the two inequalities together, and using the fact that
  \[ \int_{\Omega} |D(u_1^* \wedge u_2^*)| + \int_{\Omega} |D(u_1^* \vee u_2^*)| \leq \int_{\Omega} |Du_1^*| + \int_{\Omega} |Du_2^*|, \]
  we obtain

  \[
  \int_{\Omega} \frac{1}{2\sigma^2} (u_1^* \wedge u_2^*)^2 - \log l_0 \left( \frac{f_1 (u_1^* \wedge u_2^*)}{\sigma^2} \right) \\
  + \frac{1}{\sigma} \left( \sqrt{u_1^* \wedge u_2^*} - \sqrt{f_1} \right)^2 dx + \int_{\Omega} \frac{1}{2\sigma^2} (u_1^* \vee u_2^*)^2 \\
  - \log l_0 \left( \frac{f_2 (u_1^* \vee u_2^*)}{\sigma^2} \right) + \frac{1}{\sigma} \left( \sqrt{u_1^* \vee u_2^*} - \sqrt{f_2} \right)^2 dx \\
  \geq \int_{\Omega} \frac{1}{2\sigma^2} (u_1^*)^2 - \log l_0 \left( \frac{f_1 u_1^*}{\sigma^2} \right) + \frac{1}{\sigma} \left( \sqrt{u_1^*} - \sqrt{f_1} \right)^2 dx \\
  + \int_{\Omega} \frac{1}{2\sigma^2} (u_2^*)^2 - \log l_0 \left( \frac{f_2 u_2^*}{\sigma^2} \right) + \frac{1}{\sigma} \left( \sqrt{u_2^*} - \sqrt{f_2} \right)^2 dx.
  \]
Minimum-maximum Principle

- As $\Omega$ can be written as $\Omega = \{u_1^* > u_2^*\} \cup \{u_1^* \leq u_2^*\}$, it is clear that
  \[
  \int_{\Omega} ((u_1^* \wedge u_2^*)^2 + (u_1^* \vee u_2^*)^2) \, dx = \int_{\Omega} ((u_1^*)^2 + (u_2^*)^2) \, dx.
  \]

- The inequality can be simplified as follows
  \[
  \int_{\{u_1^* > u_2^*\}} \left[ \log \frac{l_0\left(\frac{f_1 u_1^*}{\sigma^2}\right) l_0\left(\frac{f_2 u_2^*}{\sigma^2}\right)}{l_0\left(\frac{f_1 u_2^*}{\sigma^2}\right) l_0\left(\frac{f_2 u_1^*}{\sigma^2}\right)} \right.
  + \frac{1}{\sigma} (\sqrt{u_1^*} - \sqrt{u_2^*})(\sqrt{f_1} - \sqrt{f_2}) \right] \, dx \geq 0.
  \]

- Since $l_0(t)$ is exponentially increasing function, we get $\log l_0$ is strictly increasing.
Based on Lemma 3 and Lemma 4, we get that if \( f_1 < f_2 \) and \( u_1^* > u_2^* \),

\[
\log l_0\left(\frac{f_1 u_1^*}{\sigma^2}\right) + \log l_0\left(\frac{f_2 u_2^*}{\sigma^2}\right) < \log l_0\left(\frac{f_1 u_2^*}{\sigma^2}\right) + \log l_0\left(\frac{f_2 u_1^*}{\sigma^2}\right),
\]

which is equivalent to

\[
\log \frac{l_0\left(\frac{f_1 u_1^*}{\sigma^2}\right)l_0\left(\frac{f_2 u_2^*}{\sigma^2}\right)}{l_0\left(\frac{f_1 u_2^*}{\sigma^2}\right)l_0\left(\frac{f_2 u_1^*}{\sigma^2}\right)} < 0.
\]

From the assumption \( f_1 < f_2 \), in this case we conclude that \( u_1^* > u_2^* \) has zero Lebesgue measure, i.e., \( u_1^* \leq u_2^* \) a.e. in \( \Omega \).
Existence and uniqueness of a solution

**Theorem 2** Assume that $A \in \mathcal{L}(L^2(\Omega))$ is nonnegative, and it does not annihilate constant functions, i.e., $A1 \neq 0$. Let $f$ be in $L^\infty(\Omega)$ with $\inf_{\Omega} f > 0$, then the model (2) has a solution $u^*$. Moreover, if $A$ is injective, then the solution is unique.

- Define $E_A(u)$ as follows

$$E_A(u) = \frac{1}{2\sigma^2} \int_{\Omega} (Au)^2 dx - \int_{\Omega} \log l_0\left(\frac{A uf}{\sigma^2}\right) dx$$

$$+ \frac{1}{\sigma} \int_{\Omega} (\sqrt{Au} - \sqrt{f})^2 dx + \gamma \int_{\Omega} |Du| dx.$$

- Similar to the proof of Theorem 1, $E_A$ is bounded from below. Thus, we choose a minimizing sequence $\{u_n\}$ for (2), and have that $\{\int_{\Omega} |Du_n|\}$ is bounded.
Existence and uniqueness of a solution

- Applying the Poincaré inequality, we get

\[ \|u_n - m_\Omega(u_n)\|_2 \leq C \int_\Omega |D(u_n - m_\Omega(u_n))| = C \int_\Omega |Du_n|, \]

where \( m_\Omega(u_n) = \frac{1}{|\Omega|} \int_\Omega u_n \, dx \), \(|\Omega|\) denotes the measure of \( \Omega \), and \( C \) is a constant. As \( \Omega \) is bounded, \( \|u_n - m_\Omega(u_n)\|_2 \) is bounded for each \( n \).

- Since \( A \in \mathcal{L}(L^2(\Omega)) \) is continuous, \( \{A(u_n - m_\Omega(u_n))\} \) must be bounded in \( L^2(\Omega) \) and in \( L^1(\Omega) \).

- Based on the boundedness of \( E_A(u_n) \), \( \|\sqrt{A}u_n - \sqrt{f}\|^2 \) is bounded, which implies that \( Au_n \) is bounded in \( L^1(\Omega) \).

Meanwhile, we have:

\[ |m_\Omega(u_n)| \cdot \|A1\|_1 = \|A(u_n - m_\Omega(u_n)) - Au_n\|_1 \leq \|A(u_n - m_\Omega(u_n))\|_1 + \|Au_n\|_1, \]

which turns out that \( m_\Omega(u_n) \) is uniformly bounded, because of \( A1 \neq 0 \).
Existence and uniqueness of a solution

- As we know that \( \{ u_n - m_\Omega(u_n) \} \) is bounded, the boundness of \( \{ u_n \} \) in \( L^2(\Omega) \) and thus in \( L^1(\Omega) \) is obvious.
- Therefore, there exists a subsequence \( \{ u_{n_k} \} \) which converges weakly in \( L^2(\Omega) \) to some \( u^* \in L^2(\Omega) \), and \( \{ Du_{n_k} \} \) weakly-* converges as a measure to \( Du^* \).
- Since the linear operator \( A \) is continuous, we have that \( \{ Au_{n_k} \} \) converges weakly to \( Au^* \) in \( L^2(\Omega) \) as well.
- Then according to the lower semi-continuity of the total variation and Fatou’s lemma, we obtain that \( u^* \) is a solution of the model (2).
- Furthermore, if \( A \) is injective, then its minimizer has to be unique since (2) is strictly convex.
Primal-Dual Algorithm for solving the model (2)

1: Fixed $\tau$ and $\beta$. Initialize $u^0 = \bar{u}^0 = w^0 = \bar{w}^0 = f$, $v^0 = \nabla(u^0)$, $p^0 = (0, \cdots, 0)^\top \in \mathbb{R}^{2n}$, and $q^0 = (0, \cdots, 0) \in \mathbb{R}^n$.

2: Calculate $p^{k+1}$, $q^{k+1}$, $w^{k+1}$ and $u^{k+1}$ from

\[
p^{k+1} = \arg\max_p \langle \bar{v}^k - \nabla \bar{u}^k, p \rangle - \frac{1}{2\beta} \| p - p^k \|^2
= p^k + \beta(\bar{v}^k - \nabla \bar{u}^k),
\]

\[
q^{k+1} = \arg\max_q \langle \bar{w}^k - A\bar{u}^k, q \rangle - \frac{1}{2\beta} \| q - q^k \|^2
= q^k + \beta(\bar{w}^k - A\bar{u}^k),
\]

\[
u^{k+1} = \arg\min_{0 \leq u \leq 255} \langle u, \gamma \text{div} p^{k+1} - A^\top q^{k+1} \rangle + \frac{1}{2\tau} \| u - u^k \|^2
= u^k + \tau(A^\top q^{k+1} - \text{div} p^{k+1}),
\]

\[
w^{k+1} = \arg\min_{0 \leq w \leq 255} G(w) + \langle w, q^{k+1} \rangle + \frac{1}{2\tau} \| w - w^k \|^2
\]

\[
\bar{u}^{k+1} = 2u^{k+1} - u^k,
\]

\[
\bar{v}^{k+1} = 2v^{k+1} - v^k,
\]

\[
\bar{w}^{k+1} = 2w^{k+1} - w^k.
\]

3: Stop; or set $k := k + 1$ and go to step 2.
Numerical Experiments - Denoising

Figure: Results and PSNR values of different methods when removing the Rician noise with $\sigma = 20$ in natural image "Cameraman". Row 1: the recovered images with different methods. Row 2: the residual images with different methods. (a) Zoomed original "Cameraman", (b) MAP model ($\gamma = 0.05$), (c) Getreuer’s model ($\lambda = 20$), (d) Our proposed model ($\gamma = 0.05$), (e) Noisy "Cameraman" with $\sigma = 20$, (f)-(h) are residual images of MAP model, Getreuer’s model and our model, respectively.
## Denoising

**Table:** The comparisons of PSNR values, SSIM values and CPU-time in seconds by different methods for denoising case.

<table>
<thead>
<tr>
<th>Images</th>
<th>Methods</th>
<th>PSNR (σ = 20)</th>
<th>SSIM (σ = 20)</th>
<th>Time(s) (σ = 20)</th>
<th>PSNR (σ = 30)</th>
<th>SSIM (σ = 30)</th>
<th>Time(s) (σ = 30)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Noisy</td>
<td>22.09</td>
<td>0.4069</td>
<td></td>
<td>18.46</td>
<td>0.2874</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MAP</td>
<td>27.47</td>
<td>0.8153</td>
<td>106.85</td>
<td>24.25</td>
<td>0.7342</td>
<td>137.20</td>
</tr>
<tr>
<td>Camera-man</td>
<td>Getreuer's</td>
<td>27.28</td>
<td>0.7512</td>
<td>1.96</td>
<td>25.12</td>
<td>0.6567</td>
<td><strong>1.93</strong></td>
</tr>
<tr>
<td></td>
<td>Ours</td>
<td><strong>27.82</strong></td>
<td><strong>0.8221</strong></td>
<td><strong>1.79</strong></td>
<td><strong>25.84</strong></td>
<td><strong>0.7637</strong></td>
<td>2.23</td>
</tr>
<tr>
<td>Lumbar-Spine</td>
<td>Noisy</td>
<td>21.94</td>
<td>0.4113</td>
<td></td>
<td>18.20</td>
<td>0.2680</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MAP</td>
<td>27.52</td>
<td>0.7485</td>
<td>71.87</td>
<td>24.26</td>
<td>0.6259</td>
<td>65.06</td>
</tr>
<tr>
<td></td>
<td>Getreuer's</td>
<td>27.57</td>
<td>0.6746</td>
<td>1.27</td>
<td>24.31</td>
<td>0.4886</td>
<td>1.23</td>
</tr>
<tr>
<td></td>
<td>Ours</td>
<td><strong>28.18</strong></td>
<td><strong>0.7743</strong></td>
<td><strong>1.26</strong></td>
<td><strong>24.87</strong></td>
<td><strong>0.6444</strong></td>
<td><strong>1.14</strong></td>
</tr>
<tr>
<td>Brain</td>
<td>Noisy</td>
<td>21.13</td>
<td>0.4666</td>
<td></td>
<td>17.60</td>
<td>0.3384</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MAP</td>
<td>26.52</td>
<td>0.7191</td>
<td>72.42</td>
<td>22.70</td>
<td>0.6484</td>
<td>81.77</td>
</tr>
<tr>
<td></td>
<td>Getreuer's</td>
<td>28.84</td>
<td><strong>0.8893</strong></td>
<td><strong>1.50</strong></td>
<td>25.84</td>
<td><strong>0.8194</strong></td>
<td><strong>1.37</strong></td>
</tr>
<tr>
<td></td>
<td>Ours</td>
<td><strong>29.21</strong></td>
<td>0.8860</td>
<td><strong>1.11</strong></td>
<td><strong>26.56</strong></td>
<td>0.8067</td>
<td><strong>1.30</strong></td>
</tr>
<tr>
<td>Liver</td>
<td>Noisy</td>
<td>22.28</td>
<td>0.4395</td>
<td></td>
<td>18.82</td>
<td>0.3016</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MAP</td>
<td>28.62</td>
<td>0.7870</td>
<td>72.42</td>
<td>26.01</td>
<td>0.7129</td>
<td>87.97</td>
</tr>
<tr>
<td></td>
<td>Getreuer's</td>
<td>28.46</td>
<td>0.7726</td>
<td><strong>2.23</strong></td>
<td>26.29</td>
<td>0.6800</td>
<td><strong>2.19</strong></td>
</tr>
<tr>
<td></td>
<td>Ours</td>
<td><strong>28.96</strong></td>
<td><strong>0.7986</strong></td>
<td>2.38</td>
<td><strong>26.78</strong></td>
<td><strong>0.7326</strong></td>
<td>2.23</td>
</tr>
<tr>
<td>Average</td>
<td>Noisy</td>
<td>21.81</td>
<td>0.4311</td>
<td></td>
<td>18.27</td>
<td>0.2989</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MAP</td>
<td>27.53</td>
<td>0.7675</td>
<td>98.38</td>
<td>24.31</td>
<td>0.6804</td>
<td>104.03</td>
</tr>
<tr>
<td></td>
<td>Getreuer's</td>
<td>28.04</td>
<td>0.7719</td>
<td>1.74</td>
<td>25.39</td>
<td>0.6612</td>
<td><strong>1.68</strong></td>
</tr>
<tr>
<td></td>
<td>Ours</td>
<td><strong>28.52</strong></td>
<td><strong>0.8203</strong></td>
<td><strong>1.64</strong></td>
<td><strong>26.01</strong></td>
<td><strong>0.7369</strong></td>
<td><strong>1.73</strong></td>
</tr>
</tbody>
</table>
Deblurring and Denoising

Figure: Results and PSNR values of different methods when removing the Rician noise with $\sigma = 15$ and Motion blur in MR image ”Lumbar Spine”. Row 1: the recovered images with different methods. Row 2: the residual images with different methods. (a) Original ”Cameraman”, (b) MAP model ($\gamma = 0.05$), (c) Getreuer’s model ($\lambda = 20$), (d) Our proposed model ($\gamma = 0.04$), (e) Blurry ”Lumbar Spine” image with Rician noise $\sigma = 10$, (f)-(h) are residual images of MAP model, Getreuer’s model and our model, respectively.
Deblurring and Denoising

Table: The comparisons of PSNR values, SSIM values and CPU-time in seconds by different methods for deblurring with denoising.

<table>
<thead>
<tr>
<th>Images</th>
<th>Methods</th>
<th>Motion Blur</th>
<th></th>
<th></th>
<th></th>
<th>Gaussian Blur</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>PSNR</td>
<td>SSIM</td>
<td>Time(s)</td>
<td>PSNR</td>
<td>SSIM</td>
<td>Time(s)</td>
<td></td>
</tr>
<tr>
<td>Camera-Man</td>
<td>Degraded</td>
<td>21.70</td>
<td>0.3680</td>
<td></td>
<td>22.26</td>
<td>0.3834</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MAP</td>
<td>24.87</td>
<td><strong>0.7811</strong></td>
<td>81.49</td>
<td>25.35</td>
<td><strong>0.7877</strong></td>
<td>81.81</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Getreuer's</td>
<td>24.64</td>
<td>0.7697</td>
<td>2.87</td>
<td>25.29</td>
<td>0.7816</td>
<td>3.12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ours</td>
<td><strong>25.11</strong></td>
<td>0.7797</td>
<td><strong>2.24</strong></td>
<td><strong>25.74</strong></td>
<td>0.7866</td>
<td><strong>2.27</strong></td>
<td></td>
</tr>
<tr>
<td>Lumbar-Spine</td>
<td>Degraded</td>
<td>23.15</td>
<td>0.4290</td>
<td></td>
<td>23.66</td>
<td>0.4543</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MAP</td>
<td>26.73</td>
<td>0.7492</td>
<td>60.60</td>
<td>27.56</td>
<td>0.7810</td>
<td>60.68</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Getreuer's</td>
<td>25.76</td>
<td>0.6760</td>
<td>1.68</td>
<td>26.86</td>
<td>0.6953</td>
<td>1.40</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ours</td>
<td><strong>27.30</strong></td>
<td><strong>0.7645</strong></td>
<td><strong>1.19</strong></td>
<td><strong>28.10</strong></td>
<td><strong>0.7816</strong></td>
<td><strong>1.47</strong></td>
<td></td>
</tr>
<tr>
<td>Brain</td>
<td>Degraded</td>
<td>22.72</td>
<td>0.4912</td>
<td></td>
<td>22.96</td>
<td>0.5039</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MAP</td>
<td>27.61</td>
<td>0.7572</td>
<td>68.09</td>
<td>28.84</td>
<td>0.7728</td>
<td>69.05</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Getreuer's</td>
<td>27.91</td>
<td>0.8728</td>
<td>2.09</td>
<td>28.96</td>
<td><strong>0.8915</strong></td>
<td>2.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ours</td>
<td><strong>28.49</strong></td>
<td><strong>0.8737</strong></td>
<td><strong>1.87</strong></td>
<td><strong>29.44</strong></td>
<td>0.8905</td>
<td><strong>1.87</strong></td>
<td></td>
</tr>
<tr>
<td>Liver</td>
<td>Degraded</td>
<td>23.41</td>
<td>0.4452</td>
<td></td>
<td>23.76</td>
<td>0.4639</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MAP</td>
<td>28.02</td>
<td>0.7741</td>
<td>125.51</td>
<td>28.51</td>
<td>0.7826</td>
<td>126.78</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Getreuer's</td>
<td>27.87</td>
<td>0.7717</td>
<td>3.05</td>
<td>28.30</td>
<td>0.7798</td>
<td>3.03</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ours</td>
<td><strong>28.19</strong></td>
<td><strong>0.7820</strong></td>
<td><strong>2.22</strong></td>
<td><strong>28.79</strong></td>
<td><strong>0.7935</strong></td>
<td><strong>2.71</strong></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>Degraded</td>
<td>22.75</td>
<td>0.4334</td>
<td></td>
<td>23.16</td>
<td>0.4513</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MAP</td>
<td>26.81</td>
<td>0.7654</td>
<td>83.92</td>
<td>27.57</td>
<td>0.7810</td>
<td>84.58</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Getreuer's</td>
<td>26.55</td>
<td>0.7722</td>
<td>2.42</td>
<td>27.35</td>
<td>0.7871</td>
<td>2.41</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ours</td>
<td><strong>27.27</strong></td>
<td><strong>0.8000</strong></td>
<td><strong>1.88</strong></td>
<td><strong>28.02</strong></td>
<td><strong>0.8130</strong></td>
<td><strong>2.08</strong></td>
<td></td>
</tr>
</tbody>
</table>
Energy comparison

Table: The comparisons of the energy values $E(u_{\text{MAP}})$, $E(u_{\text{Getreuer's}})$, and $E(u_{\text{Ours}})$. Here, $E(u) := \frac{1}{2\sigma^2} \int_{\Omega} u^2 dx - \int_{\Omega} \log I_0\left(\frac{fu}{\sigma^2}\right) dx + \gamma \int_{\Omega} |Du| dx$, with the same $\gamma$. These three values are nearly the same.

<table>
<thead>
<tr>
<th>Method</th>
<th>$\sigma = 20$</th>
<th>$\sigma = 30$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MAP</td>
<td>Getreuer’s</td>
</tr>
<tr>
<td>Cameraman</td>
<td>-404.15</td>
<td>-404.04</td>
</tr>
<tr>
<td>Skull</td>
<td>-84.34</td>
<td>-83.86</td>
</tr>
<tr>
<td>Liver</td>
<td>-150.70</td>
<td>-150.67</td>
</tr>
<tr>
<td>Brain</td>
<td>-218.03</td>
<td>-220.11</td>
</tr>
<tr>
<td>Average</td>
<td>-214.30</td>
<td>-214.67</td>
</tr>
</tbody>
</table>

As one of our aim is to obtain a good approximation for the global solution of the original non-convex model (presented in the caption of the above table)

- we plugged the solutions obtained by the MAP model, Getreuer’s convex model and our model into the original functional to get the corresponding energy values.
- as these three energy values are quite close, Getreuer’s and our approximation approaches are quite reasonable.
- in the denoising case, our method is comparable to the MAP and Getreuer’s methods.
Comments on numerical results

Our convexified model is very competitive to other recently proposed methods.

- Higher PSNR values and better visual results.
- Less computational time.
- Unique minimizer, independent of initialization.

It seems the effort of convexification is justified, so we continue to another non-convex problem...
Other non-convex models

Many other non-convex models in image processing.

- Generalization?
- Possible....
- Not so easy
Multiplicative Gamma noise with blur

Degradation model

\[ f = (Au)\eta. \]

- \( \eta \): multiplicative noise.
- We assume \( \eta \) follows Gamma distribution, i.e.,
  \[
  p_{\eta}(x; \theta, K) = \frac{1}{\theta^K \Gamma(K)} x^{K-1} e^{-\frac{x}{\theta}}.
  \]
- Mean and variance of \( \eta \) are \( K\theta \) and \( K\theta^2 \).
- We assume mean of \( \eta \) equals 1.
TV-Multiplicative model

MAP analysis leads to:

\[
\min_u \int_\Omega \left( \log(Au) + \frac{f}{Au} \right) \, dx + \lambda \int_\Omega |Du|,
\]

- Known as the AA model.
- The edge-preserving TV term, \( \int_\Omega |Du| \), the constant \( \lambda > 0 \)
- Non-convex.
New approach: Convexified TV-Multiplicative model

We introduce a quadratic penalty term:

$$
\min_u \int_\Omega \left( \log(Au) + \frac{f}{Au} \right) dx + \alpha \int_\Omega \left( \sqrt{\frac{Au}{f}} - 1 \right)^2 dx + \lambda \int_\Omega |Du|.
$$

(14)

- $\alpha > 0$ is a penalty parameter.
- If $\alpha \geq \frac{2\sqrt{6}}{9}$, the model (14) is convex!
The quadratic penalty term

More reasons to add the penalty term $\int_{\Omega} \left( \sqrt{\frac{Au}{f}} - 1 \right)^2 dx$:

- Set $Y = \frac{1}{\sqrt{\eta}}$, where $\eta$ follows Gamma distribution with mean 1.
- It can be shown that (K is the shape parameter in Gamma distribution)
  \[ \lim_{K \to +\infty} E((Y - 1)^2) = 0. \]
- For large $K$, $Y$ can be well approximated by Gaussian distribution. (We will introduce the concept of the Kullback-Leibler (KL) divergence later to reveal the relationship.)
- The MAP estimation leads to $\int_{\Omega} (v - f)^2 dx$ as data fitting term in additive Gaussian noise removal.
Proposition Suppose that the random variable $\eta$ follows Gamma distribution with mean 1. Set $Y = \frac{1}{\sqrt{\eta}}$, with mean $\mu_K$ and variance $\sigma_k^2$. Then

$$
\lim_{K \to +\infty} D_{KL}(Y \| \mathcal{N}(\mu_K, \sigma_k^2)) = 0,
$$

where $\mathcal{N}(\mu_K, \sigma_k^2)$ is the Gaussian distribution.

- It can be shown that $D_{KL}(Y \| \mathcal{N}(\mu_K, \sigma_k^2)) = O\left(\frac{1}{K}\right)$.
Indeed, we can prove that:

(i) \( \int_{0}^{+\infty} p_Y(y) \log p_Y(y) \, dy = \log 2 - \log(\sqrt{K} \Gamma(K)) + \frac{2K+1}{2} \psi(K) - K \), where 
\( \psi(K) := \frac{d \log \Gamma(K)}{dK} \) is the digamma function;

(ii) \( \int_{0}^{+\infty} p_Y(y) \log p_{\mathcal{N}(\mu_K, \sigma^2_K)}(y) \, dy = -\frac{1}{2} \log(2\pi e \sigma^2_K) \), where 
\( p_{\mathcal{N}(\mu_K, \sigma^2_K)}(y) \) denotes the PDF of the Gaussian distribution 
\( \mathcal{N}(\mu_K, \sigma^2_K) \);

(iii) \( D_{KL}(Y \| \mathcal{N}(\mu_K, \sigma^2_K)) = O\left(\frac{1}{K}\right) \).
Approximate by Gaussian distribution (cont.)

Indeed, by calculation, we can get,

\[ \sigma_K^2 = \mathbb{E}(Y^2) - \mathbb{E}^2(Y) \]
\[ = \frac{K \Gamma(K - 1)}{\Gamma(K)} - \frac{K \Gamma^2(K - \frac{1}{2})}{\Gamma^2(K)}. \]

and

\[ D_{\text{KL}}(Y \| \mathcal{N}(\mu_K, \sigma_K^2)) \]
\[ = \log 2 - \log(\sqrt{K} \Gamma(K)) + \frac{2K + 1}{2} \psi(K) - K + \frac{1}{2} \log(2\pi e \sigma_K^2) \]
\[ = \log 2 + \frac{1}{2} \log K \sigma_K^2 + \mathcal{O}(\frac{1}{K}) \]
\[ = \log 2 + \frac{1}{2} \log \left(\frac{1}{4} + \mathcal{O}(\frac{1}{K})\right) + \mathcal{O}(\frac{1}{K}) \]
\[ = \mathcal{O}(\frac{1}{K}). \]
Furthermore, if we set $Z := \frac{Y - \mu_K}{\sigma_K}$, then

$$\lim_{K \to \infty} D_{KL}(Z|\mathcal{N}(0, 1)) = 0.$$ 

▶ A simple change of variable shows

$$D_{KL}(Z|\mathcal{N}(0, 1)) = D_{KL}(Y|\mathcal{N}(\mu_K, \sigma^2_K)).$$

▶ $Z$ tends to the standard Gaussian distribution $\mathcal{N}(0, 1)$. 
Figure: The comparisons of the PDFs of $Y$ and $\mathcal{N}(\mu, \sigma^2)$ with different $K$. (a) $K = 6$, (b) $K = 10$. 
Numerical Experiments - Denoising

Figure: Results of different methods when removing the multiplicative noise with $K = 10$. (a) Noisy "Cameraman", (b) AA method, (c) RLO method, (d) our method.
Cauchy noise with blur

Degradation model

\[ f = (Au) + \nu. \]

- \( \nu \): Cauchy noise.
- A random variable \( V \) follows the Cauchy distribution if it has density

\[ g(\nu) = \frac{1}{\pi} \frac{\gamma}{\gamma^2 + (\nu - \delta)^2}, \]

where \( \gamma > 0 \) is the scale parameter and \( \delta \) is localization parameter.
Cauchy noise

Figure: Alpha-stable noise in 1D: notice that the y-axis has different scale (scale between 30 and 120 on (a) and (b) and −100 and 400 on (c)). (a) 1D noise free signal; (b) signal degraded by an additive Gaussian noise; (c) signal degraded by an additive Cauchy noise. Cauchy noise is more impulsive than the Gaussian noise.
TV-Cauchy model

MAP analysis leads to:

\[
\inf_{u \in BV(\Omega)} TV(u) + \frac{\lambda}{2} \int_{\Omega} \log\left(\gamma^2 + (Au - f)^2\right) dx
\]

- \(\gamma > 0\) is the scale parameter and \(A\) is the blurring operator.
- Edge-preserving.
- Non-convex.
Image corrupted by Cauchy noise

Figure: Comparison of different noisy images. (a) Original image $u_0$; (b) $u$ corrupted by an additive Gaussian noise; (c) $u$ corrupted by an additive Cauchy noise; (d) $u$ corrupted by an impulse noise; (e)–(h) zoom of the top left corner of the images (a)–(d), respectively. Cauchy noise and impulse noise are more impulsive than the Gaussian noise.
New approach: Convexified TV-Cauchy model

We introduce a quadratic penalty term:

$$
\inf_{u \in BV(\Omega)} TV(u) + \frac{\lambda}{2} \left( \int_{\Omega} \log \left( \gamma^2 + (Au - f)^2 \right) dx + \mu \|Au - u_0\|_2^2 \right).
$$

- Cauchy noise has impulsive character.
- $u_0$ is the image obtained by applying the median filter to the noisy image.
- The median filter is used instead of the myriad filter for simplicity and computational time.
Numerical Experiments - Denoising

Figure: Recovered images (with PSNR(dB)) of different approaches for removing Cauchy noise from the noisy image "Peppers". (a) Wavelet shrinkage; (b) SURE-LET; (c) BM3D; (d) our model.
Denoising

Table: PSNR values for noisy images and recovered images given by different methods ($\xi = 0.02$). In the last line of the table, we compute the average of the values.

<table>
<thead>
<tr>
<th></th>
<th>Noisy</th>
<th>ROF</th>
<th>MD</th>
<th>MR</th>
<th>$L^1$-TV</th>
<th>Ours</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peppers</td>
<td>19.15</td>
<td>25.03</td>
<td>29.64</td>
<td>29.85</td>
<td>30.34</td>
<td>30.94</td>
</tr>
<tr>
<td>Parrot</td>
<td>19.13</td>
<td>23.88</td>
<td>27.05</td>
<td>27.13</td>
<td>28.02</td>
<td>28.98</td>
</tr>
<tr>
<td>Cameraman</td>
<td>19.07</td>
<td>24.00</td>
<td>26.14</td>
<td>26.57</td>
<td>27.21</td>
<td>27.91</td>
</tr>
<tr>
<td>Lena</td>
<td>19.06</td>
<td>24.58</td>
<td>28.94</td>
<td>28.98</td>
<td>29.84</td>
<td>30.36</td>
</tr>
<tr>
<td>Boat</td>
<td>19.03</td>
<td>24.21</td>
<td>27.27</td>
<td>27.49</td>
<td>28.70</td>
<td>29.20</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td>19.09</td>
<td>23.89</td>
<td>26.75</td>
<td>26.97</td>
<td>28.08</td>
<td><strong>28.74</strong></td>
</tr>
</tbody>
</table>
Deblurring and Denoising

**Figure:** The zoomed-in regions of the recovered images from blurry images with Cauchy noise. First row: details of original images; second row: details of restored images by $L^1$-TV approach; third row: details of restored images by our approach.
Table: PSNR values for noisy images and recovered images given by different methods ($\xi = 0.02$). In the last line of the table, we compute the average of the values.

<table>
<thead>
<tr>
<th></th>
<th>Noisy</th>
<th>ROF</th>
<th>MD</th>
<th>MR</th>
<th>$L^1$-TV</th>
<th>Ours</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peppers</td>
<td>18.31</td>
<td>24.21</td>
<td>25.19</td>
<td>25.01</td>
<td>26.70</td>
<td>27.46</td>
</tr>
<tr>
<td>Parrot</td>
<td>18.23</td>
<td>24.06</td>
<td>24.48</td>
<td>24.57</td>
<td>25.75</td>
<td>26.79</td>
</tr>
<tr>
<td>Cameraman</td>
<td>18.29</td>
<td>23.98</td>
<td>24.43</td>
<td>24.39</td>
<td>25.49</td>
<td>26.27</td>
</tr>
<tr>
<td>Lena</td>
<td>18.64</td>
<td>25.74</td>
<td>26.70</td>
<td>26.72</td>
<td>27.26</td>
<td>28.14</td>
</tr>
<tr>
<td>Baboon</td>
<td>17.42</td>
<td>20.84</td>
<td>21.54</td>
<td>21.49</td>
<td>21.36</td>
<td>21.81</td>
</tr>
<tr>
<td>Goldhill</td>
<td>18.47</td>
<td>24.84</td>
<td>25.88</td>
<td>25.85</td>
<td>26.17</td>
<td>26.76</td>
</tr>
<tr>
<td>Boat</td>
<td>18.48</td>
<td>24.36</td>
<td>25.42</td>
<td>25.43</td>
<td>26.18</td>
<td>26.69</td>
</tr>
<tr>
<td>Average</td>
<td>18.28</td>
<td>24.00</td>
<td>24.81</td>
<td>24.78</td>
<td>25.56</td>
<td>26.31</td>
</tr>
</tbody>
</table>
Remarks

- We introduced three convex models based on Total Variation.
- Under mild conditions, our models have unique solutions.
- Because of convexity, fast algorithms can be employed.
- Numerical experiments suggest good performance of the models.
- Generalization?
Outline

1. Variational Models for Rician Noise Removal
2. Two-stage Segmentation
3. Dictionary and Weighthed Nuclear Norm
Mumford-Shah model (1989)

Mumford-Shah model is an energy minimization problem for image segmentation:

$$\min_{g,\Gamma} \frac{\lambda}{2} \int_{\Omega} (f - g)^2 dx + \frac{\mu}{2} \int_{\Omega \setminus \Gamma} |\nabla g|^2 dx + \text{Length}(\Gamma).$$

- $f$ observed image, $g$ a piecewise smooth approximation of $f$,
- $\Gamma$ boundary of segmented region.

- $\lambda$ and $\mu$ are positive.

- Non-convex, very difficult to solve!
Chan-Vese model
When $\mu \to +\infty$, it reduces to:

$$
\min_{\Gamma, c_1, c_2} \lambda \int_{\Omega_1} (f - c_1)^2 dx + \frac{\lambda}{2} \int_{\Omega_2} (f - c_2)^2 dx + \text{Length}(\Gamma),
$$

where $\Omega$ is separated into $\Omega_1, \Omega_2$ by $\Gamma$.

- Level-set method can be applied.
- Rather complex!
- Still non-convex.
- Multi-phases?
Our approach

We propose a novel two-stage segmentation approach.

▶ First stage, solve for a convexified variant of the Mumford-Shah model.
▶ Second stage, use data clustering method to threshold the solved solution.
Stage One

- \text{Length}(\Gamma) \leftrightarrow \int_{\Omega} |\nabla u|; \text{ and equal when } u \text{ is binary.}
- \text{phases of } g \text{ can be obtained from } u \text{ by thresholding.}
Stage One

- \( \text{Length}(\Gamma) \leftrightarrow \int_{\Omega} |\nabla u| \); and equal when \( u \) is binary.
- phases of \( g \) can be obtained from \( u \) by thresholding.

**Observation:** (especially when \( g \) is nearly binary)

- jump set of \( g \) \( \approx \) jump set of \( u \).
- \( \int_{\Omega} |\nabla g| \approx \int_{\Omega} |\nabla u| \leftrightarrow \text{Length}(\Gamma) \).
Stage One: Unique Minimizer

Our convex variant of the Mumford-Shah model (Hintermüller $H^1$-norm for image restoration) is:

$$E(g) = \frac{\lambda}{2} \int_\Omega (f - Ag)^2 \, dx + \frac{\mu}{2} \int_\Omega |\nabla g|^2 \, dx + \int_\Omega |\nabla g| \, dx$$  \hspace{1cm} (15)

Its discrete version:

$$\frac{\lambda}{2} \| f - Ag \|^2_2 + \frac{\mu}{2} \| \nabla g \|^2_2 + \| \nabla g \|_1$$

**Theorem**

Let $\Omega$ be a bounded connected open subset of $\mathbb{R}^2$ with a Lipschitz boundary. Let $\text{Ker}(A) \cap \text{Ker}(\nabla) = \{0\}$ and $f \in L^2(\Omega)$, where $A$ is a bounded linear operator from $L^2(\Omega)$ to itself. Then $E(g)$ has a unique minimizer $g \in W^{1,2}(\Omega)$.
Stage Two

With a minimizer \( g \) from the first stage, we obtain a segmentation by thresholding \( g \).

- We use the K-means method to obtain proper thresholds.
- Fast multiphase segmentation.
- Number of phases \( K \) and thresholds \( \rho \) are determined after \( g \) is calculated. Little computation to change \( K \) and \( \rho \). No need to recalculate \( u \)!
- Users can try different \( K \) and \( \rho \).
Extensions to Other Noise Models

First stage: solve for

\[
\min_g \left\{ \lambda \int_\Omega (Ag - f \log Ag) dx + \frac{\mu}{2} \int_\Omega |\nabla g|^2 dx + \int_\Omega |\nabla g| dx \right\}
\]

(16)

- data fitting term good for Poisson noise from MAP analysis
- also suitable for Gamma noise (Steidl and Teuber (10)).
- objective functional is convex (solved by Chambolle-Pock)
- admits unique solution if \( \text{Ker}(\mathcal{A}) \cap \text{Ker}(\nabla) = \{0\} \).

Second stage: threshold the solution to get the phases.
Example: Poisson noise with motion blur

Original image | Noisy & blurred | Yuan et al. (10) |
--- | --- | ---
Dong et al. (10) | Sawatzky et al. (13) | Our method
Example: Gamma noise

Original image

Noisy image

Yuan et al. (10)

Li et al. (10)

Our method
Three-stage Model

We propose a SLaT (Smoothing, Lifting and Thresholding) method for multiphase segmentation of color images corrupted by different types of degradations: noise, information loss, and blur.

- Stage One: Smoothing
- Stage Two: Lifting
- Stage Three: Thresholding
Three-stage Model

Let the given degraded image be in $\mathcal{V}_1$.

- **Smoothing:** The convex variational model ((15) or (16)) for grayscale images is applied in parallel to each channel of $\mathcal{V}_1$. This yields a restored smooth image.

- **Color Dimension Lifting:** We transform the smoothed color image to a secondary color space $\mathcal{V}_2$ that provides us with complementary information. Then we combine these images as a new vector-valued image composed of all the channels from color spaces $\mathcal{V}_1$ and $\mathcal{V}_2$.

- **Thresholding:** According to the desired number of phases $K$, we apply a multichannel thresholding to the combined $\mathcal{V}_1$-$\mathcal{V}_2$ image to obtain a segmented image.
Smoothing stage

Let \( f = (f_1, \ldots, f_d) \) be a given color image with channels
\( f_i : \Omega \rightarrow \mathbb{R}, \ i = 1, \ldots, d \). Denote \( \Omega_0^i \) the set where \( f_i \) is known.

We consider the minimizing the functional \( E \) below

\[
E(g_i) = \frac{\lambda}{2} \int_{\Omega} \omega_i \cdot \Phi(f_i, g_i) dx + \frac{\mu}{2} \int_{\Omega} |\nabla g_i|^2 dx + \int_{\Omega} |\nabla g_i| dx, \quad i = 1, \ldots, d,
\]

where \( \omega_i(\cdot) \) is the characteristic function of \( \Omega_0^i \), i.e.

\[
\omega_i(x) = \begin{cases} 
1, & x \in \Omega_0^i, \\
0, & x \in \Omega \setminus \Omega_0^i.
\end{cases}
\]

For \( \Phi \) in (17) we consider the following options:

1. \( \Phi(f, g) = (f - Ag)^2 \), the usual choice;
2. \( \Phi(f, g) = Ag - f \log(Ag) \) if data are corrupted by Poisson noise.
Uniqueness and existence

The Theorem below proves the existence and the uniqueness of the minimizer of (17) where we define the linear operator $(\omega_iA)$ by

$$(\omega_iA) : u(x) \in L^2(\Omega) \mapsto \omega_i(x)(Au)(x) \in L^2(\Omega).$$

Theorem

Let $\Omega$ be a bounded connected open subset of $\mathbb{R}^2$ with a Lipschitz boundary. Let $A : L^2(\Omega) \to L^2(\Omega)$ be bounded and linear. For $i \in \{1, \ldots, d\}$, assume that $f_i \in L^2(\Omega)$ and that $\text{Ker}(\omega_iA) \cap \text{Ker}(\nabla) = \{0\}$ where $\text{Ker}$ stands for null-space. Then

$$E(g_i) = \frac{\lambda}{2} \int_{\Omega} \omega_i \cdot \Phi(f_i, g_i) dx + \frac{\mu}{2} \int_{\Omega} |\nabla g_i|^2 dx + \int_{\Omega} |\nabla g_i| dx,$$

with either $\Phi(f_i, g_i) = (f_i - Ag_i)^2$ or $\Phi(f_i, g_i) = Ag_i - f_i \log(Ag_i)$ has a unique minimizer $\bar{g}_i \in W^{1,2}(\Omega)$. 
Segmentation of Real-world Color Images

Figure: Four-phase sunflower segmentation (size: $375 \times 500$). (A): Given Gaussian noisy image with mean 0 and variance 0.1; (B): Given Gaussian noisy image with 60% information loss; (C): Given blurry image with Gaussian noise;
Observations for Two/Three-stage segmentation

- Convex segmentation model with unique solution. Can be solved easily and fast.
- No need to solve the model again when thresholds or number of phases changes.
- Easily extendable to e.g. blurry images and non-Gaussian noise.
- Link image segmentation and image restoration.
- Efficient algorithms of color image segmentation.
Outline

1. Variational Models for Rician Noise Removal
2. Two-stage Segmentation
3. Dictionary and Weighted Nuclear Norm
Deblurring under impulse noise

Degraded model

\[ g = \mathbb{N}_{imp}(Hu), \]

where \( g \) is corrupted image, \( H \) is the blur kernel.

- one-phase model:

\[
\min_u J(u) + \lambda F(Hu - g),
\]

where \( J \) is the regularization term, \( \lambda \) is a positive parameter, \( F \) is the data fidelity-term.

- two-phase methods:

\[
\min_u J(u) + \lambda \sum_s F(\Lambda_s(Hu - g)_s),
\]

where

\[ \Lambda_s = \begin{cases} 
0, & \text{if } s \in \mathcal{N}, \\
1, & \text{otherwise},
\end{cases} \]

with \( \mathcal{N} \) the noise candidates set.
Variational model

[Ma, Yu, and Z, SIIMS 2013]
In order to restore image from degraded model

\[ g = \mathbb{N}_{imp}(Hu). \]

we consider the following model:

\[
\min_{\alpha_s, u, D} \sum_{s \in \mathcal{P}} \mu_s \|\alpha_s\|_0 + \sum_{s \in \mathcal{P}} \|D\alpha_s - R_s u\|_2^2 + \eta \|\nabla u\|_1 + \lambda \|\Lambda(Hu - g)\|_1,
\]

(19)

where \(\|\nabla u\|_1\) denotes the discrete version of the isotropic total variation norm defined as:

\[
\|\nabla u\|_1 = \sum_s |(\nabla u)_s|, \quad |(\nabla u)_s| = \sqrt{|(\nabla u)_s^1|^2 + |(\nabla u)_s^2|^2}.
\]
Experimental results

PSNR (dB) values for various methods for the test images corrupted by Gaussian blur and random-valued noise with noise levels $r = 20\%, 40\%, 60\%$ respectively.

<table>
<thead>
<tr>
<th>Image/r</th>
<th>FTVd</th>
<th>Dong</th>
<th>Cai1</th>
<th>Cai2</th>
<th>Yan</th>
<th>Ours</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bar./20%</td>
<td>25.29</td>
<td>25.18</td>
<td>28.37</td>
<td>29.34</td>
<td>27.33</td>
<td><strong>31.30</strong></td>
</tr>
<tr>
<td>Bar./40%</td>
<td>24.17</td>
<td>24.19</td>
<td>25.99</td>
<td>26.68</td>
<td>26.21</td>
<td><strong>28.82</strong></td>
</tr>
<tr>
<td>Bar./60%</td>
<td>22.13</td>
<td>21.27</td>
<td>24.02</td>
<td>24.07</td>
<td>25.02</td>
<td><strong>25.28</strong></td>
</tr>
<tr>
<td>Cam./20%</td>
<td>29.56</td>
<td>28.19</td>
<td>33.25</td>
<td>37.22</td>
<td>33.32</td>
<td><strong>39.24</strong></td>
</tr>
<tr>
<td>Cam./40%</td>
<td>24.92</td>
<td>25.43</td>
<td>29.44</td>
<td>31.74</td>
<td>30.64</td>
<td><strong>35.21</strong></td>
</tr>
<tr>
<td>Cam./60%</td>
<td>19.92</td>
<td>21.00</td>
<td>23.33</td>
<td>23.92</td>
<td>24.55</td>
<td><strong>28.59</strong></td>
</tr>
<tr>
<td>Lena/20%</td>
<td>35.08</td>
<td>32.99</td>
<td>37.60</td>
<td>40.68</td>
<td>36.08</td>
<td><strong>42.42</strong></td>
</tr>
<tr>
<td>Lena/40%</td>
<td>31.26</td>
<td>30.85</td>
<td>35.74</td>
<td>36.67</td>
<td>35.20</td>
<td><strong>39.53</strong></td>
</tr>
<tr>
<td>Lena/60%</td>
<td>25.77</td>
<td>23.47</td>
<td>30.33</td>
<td>30.59</td>
<td>30.25</td>
<td><strong>34.41</strong></td>
</tr>
</tbody>
</table>
Numerical Experiments - Denoising

**Figure:** Recovered images (with PSNR(dB)) of different methods on image Cameraman corrupted by Gaussian blur and random-valued noise with noise level 60%.
SSMS deblurring

1. Applying Wiener filter to remove the blur effect;
2. SSMS denoising to remove the color noise.

The underlying clear image patch $x$ can be estimated from the noisy patch $y$ via

$$\tilde{f} = \sum_{l \in \Lambda} \langle f, \phi_{k_0,l} \rangle \phi_{k_0,l},$$

where

$$\Lambda = \{ l : |\langle f, \phi_{k_0,l} \rangle| > T \},$$

and the threshold value $T$ depends on the noise variance.

**Drawback**: Wiener filter doesn’t work well when the noise level is high.
Deblurring via total variation based SSMS


Model:

\[
\min_{u, D, \alpha_s} \frac{\lambda}{2} \| Au - g \|_2^2 + \frac{\beta}{2} \sum_{s \in \mathcal{P}} \| R_s u - D \alpha_s \|_2^2 \\
+ \sum_{s \in \mathcal{P}} \mu_s \| \alpha_s \|_1 + \| \nabla u \|_1,
\]

(20)

For each patch \( R_s u \), a best orthogonal basis \( \Phi_{k_0} \) of size \( N \times N \) is selected by SSMS.

We employ the alternating minimization method to solve the model.
The proposed algorithm is compared to five related existing image deblurring methods:

1. TV based method
2. Fourier wavelet regularized deconvolution (ForWaRD) method
3. Structured Sparse Model Selection (SSMS) deblurring
4. Framelet based image deblurring approach
5. Non-local TV based method
## Experimental results

Comparison of the PSNR (dB) of the recovered results by different methods, with respect to the noise level $\sigma = 2$.

<table>
<thead>
<tr>
<th></th>
<th>Kernel</th>
<th>TV</th>
<th>ForWaRD</th>
<th>SSMS</th>
<th>Framelet</th>
<th>NLTV</th>
<th>Ours</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Boat.</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>box</td>
<td></td>
<td>27.97</td>
<td>27.69</td>
<td>27.75</td>
<td>28.90</td>
<td>28.86</td>
<td><strong>29.39</strong></td>
</tr>
<tr>
<td>gaussian</td>
<td></td>
<td>27.91</td>
<td>27.80</td>
<td>27.73</td>
<td>28.62</td>
<td>28.64</td>
<td><strong>28.80</strong></td>
</tr>
<tr>
<td>motion</td>
<td></td>
<td>28.24</td>
<td>28.34</td>
<td>28.03</td>
<td>29.53</td>
<td>29.57</td>
<td><strong>30.33</strong></td>
</tr>
<tr>
<td><strong>Lena</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>box</td>
<td></td>
<td>28.79</td>
<td>29.72</td>
<td>29.82</td>
<td>30.65</td>
<td>29.64</td>
<td><strong>31.18</strong></td>
</tr>
<tr>
<td>gaussian</td>
<td></td>
<td>30.01</td>
<td>30.45</td>
<td>30.68</td>
<td>31.76</td>
<td>30.73</td>
<td><strong>32.28</strong></td>
</tr>
<tr>
<td>motion</td>
<td></td>
<td>29.56</td>
<td>30.60</td>
<td>30.55</td>
<td>31.82</td>
<td>30.97</td>
<td><strong>32.72</strong></td>
</tr>
<tr>
<td><strong>Cam.</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>box</td>
<td></td>
<td>25.44</td>
<td>25.66</td>
<td>25.69</td>
<td>26.43</td>
<td>26.29</td>
<td><strong>27.05</strong></td>
</tr>
<tr>
<td>gaussian</td>
<td></td>
<td>24.83</td>
<td>25.06</td>
<td>25.04</td>
<td>25.27</td>
<td>24.92</td>
<td><strong>25.62</strong></td>
</tr>
<tr>
<td>motion</td>
<td></td>
<td>26.62</td>
<td>26.89</td>
<td>26.74</td>
<td>28.11</td>
<td>28.38</td>
<td><strong>29.16</strong></td>
</tr>
</tbody>
</table>
**Experimental results**

Comparison of the PSNR (dB) of the recovered results by different methods, with respect to the noise level $\sigma = 10$.

<table>
<thead>
<tr>
<th></th>
<th>Kernel</th>
<th>TV</th>
<th>ForWaRD</th>
<th>SSMS</th>
<th>Framelet</th>
<th>NLTV</th>
<th>Ours</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Boat.</strong></td>
<td>box</td>
<td>25.06</td>
<td>25.15</td>
<td>25.26</td>
<td>25.72</td>
<td>25.71</td>
<td><strong>25.90</strong></td>
</tr>
<tr>
<td></td>
<td>motion</td>
<td>24.67</td>
<td>24.94</td>
<td>25.07</td>
<td>25.29</td>
<td>25.50</td>
<td><strong>25.94</strong></td>
</tr>
<tr>
<td><strong>Lena</strong></td>
<td>box</td>
<td>26.47</td>
<td>27.46</td>
<td>27.71</td>
<td>27.82</td>
<td>26.92</td>
<td><strong>28.45</strong></td>
</tr>
<tr>
<td></td>
<td>gaussian</td>
<td>27.82</td>
<td>28.96</td>
<td>29.05</td>
<td>29.17</td>
<td>28.43</td>
<td><strong>29.92</strong></td>
</tr>
<tr>
<td></td>
<td>motion</td>
<td>26.18</td>
<td>27.21</td>
<td>27.62</td>
<td>27.53</td>
<td>27.10</td>
<td><strong>28.64</strong></td>
</tr>
<tr>
<td><strong>Cam.</strong></td>
<td>box</td>
<td>23.09</td>
<td>22.35</td>
<td>22.34</td>
<td>23.50</td>
<td>23.38</td>
<td><strong>24.02</strong></td>
</tr>
<tr>
<td></td>
<td>gaussian</td>
<td>23.55</td>
<td>23.43</td>
<td>23.41</td>
<td>24.04</td>
<td>24.11</td>
<td><strong>24.37</strong></td>
</tr>
<tr>
<td></td>
<td>motion</td>
<td>22.89</td>
<td>22.26</td>
<td>22.27</td>
<td>23.43</td>
<td>23.52</td>
<td><strong>24.28</strong></td>
</tr>
</tbody>
</table>
Low rank

Figure: Grouping blocks

**BM3D:**
Grouping by matching; Collaborative filtering.

**Low rank:**
Grouping by matching; Low rank minimization.
Low rank minimization

\[
\min_X \| Y - X \|_F^2, \quad s.t. \quad \text{rank}(X) \leq r
\]

**Drawback:** nonconvex

**Convex relaxation:** nuclear norm minimization (NNM:

\[
\min_X \frac{1}{2} \| Y - X \|_F^2 + \lambda \| X \|_*
\]

with solution \( X^* = US_\lambda(\Sigma)V^T \).

\( Y = U\Sigma V^T \) is the singular value decomposition (SVD) of matrix \( Y \), where

\[
\Sigma = \begin{pmatrix}
\text{diag}(\sigma_1, \sigma_2, \cdots, \sigma_n) \\
0
\end{pmatrix},
\]

and the singular value shrinkage operator is defined as

\[
S_\lambda(\Sigma) = \max(0, \Sigma - \lambda),
\]

which actually is the proximal operator of the nuclear norm function (Cai, Candès and Shen).
**WNNM**

**NNM**: using a same value to penalize every singular value
However, the larger singular value is more important and should shrink less in many cases.

The weighted nuclear norm minimization (**WNMM**) problem:

\[
\min_X \frac{1}{2} \| Y - X \|_F^2 + \| X \|_{\bar{w},*},
\]  

(22)

where

\[
\| X \|_{\bar{w},*} = \sum_{i=1}^{n} w_i \sigma_i(X),
\]

(23)

with \( \sigma_1(X) \geq \sigma_2(X) \geq \cdots \geq \sigma_n(X) \geq 0 \), the weights vector
\( \bar{w} = [w_1, w_2, \cdots, w_n] \), and \( w_i \geq 0 \).
Existent mathematical properties of WNNM

**Proposition:** Assume $Y \in R^{m \times n}$, the SVD of $Y$ is $Y = U\Sigma V^T$, and $0 \leq w_1 \leq w_2 \leq \cdots \leq w_n$, the global optimal solution of the WNNM problem in (22) is

$$X^* = UDV^T, \quad (24)$$

where $D = \begin{pmatrix} \text{diag} (d_1, d_2, \cdots, d_n) \\ 0 \end{pmatrix}$ is a diagonal non-negative matrix and $d_i = \max(\sigma_i - w_i, 0)$, $i = 1, \cdots, n$.

**Proposition:** Assume $Y \in R^{m \times n}$, and the SVD of $Y$ is $Y = U\Sigma V^T$, the solution of the WNNM problem in (22) can be expressed as $X^* = UDV^T$, where $D = \begin{pmatrix} \text{diag} (d_1, d_2, \cdots, d_n) \\ 0 \end{pmatrix}$ is a diagonal non-negative matrix and $(d_1, d_2, \cdots, d_n)$ is the solution of the following convex optimization problem:

$$\min_{d_1, \cdots, d_n} \sum_{i=1}^{n} \frac{1}{2} (d_i - \sigma_i)^2 + w_id_i, \; \text{s.t.} \; d_1 \geq d_2 \geq \cdots \geq d_n \geq 0. \quad (25)$$
It is extremely important and urgent if one can find explicit solution for Problem (22) which uses the weights in an arbitrary order. This is one of our main contributions.

**Theorem 3:** Assume that \((a_1, \cdots, a_n) \in \mathbb{R}^n\) and \(n \in \mathbb{N}\). Then the following minimization problem

\[
\min_{d_1 \geq \cdots \geq d_n \geq 0} \sum_{i=1}^{n} (d_i - a_i)^2,
\]

(26)

has a unique global optimal solution in the closed form. Moreover, for any \(k \in \mathbb{N}\), we assume that \(M_k = (d_1, \cdots, d_k)\) is the solution to the \(k\)-th problem of (26) (that is (26) with \(n = k\)), in particular, we denote by \(d_k^* = d_k\) the last component of \(M_k\) and \(d_0^* = +\infty\).
Existential mathematical properties of WNNM

If

\[ s_0 := \sup \{ s \in \mathbb{N} \mid d_{s-1}^* \geq \frac{\sum_{k=s}^n a_k}{n-s+1}, \quad s = 1, \ldots, n \} \], \quad (27) \]

then the solution \( M_n = (d_1, \ldots, d_n) \) to (26) satisfies

\[(d_1, \ldots, d_{s_0-1}) = M_{s_0-1} \quad \text{and} \quad d_{s_0} = \cdots = d_n = \max \left( \frac{\sum_{k=s_0}^n a_k}{n-s_0+1}, 0 \right). \quad (28)\]

Furthermore, the global solution of (22) is given in (24) with

\[ D = \begin{pmatrix} \operatorname{diag} (d_1, d_2, \ldots, d_n) \\ 0 \end{pmatrix}. \]
The uniqueness is clear since the objective function in (26) is strictly convex. We should employ the method of induction to prove the existence result of (26).

Firstly, for $n = 1$, (26) has a unique solution $d_1 = \max\{a_1, 0\}$.

Assume that for any $k \leq n - 1$, the $k$-th problem of (26) has a unique solution $M_k = (d_1, \ldots, d_k)$. Then to solve the $n$-th problem (26), we compare $d_{n-1}^*$ and $a_n$ as follows:

- if $d_{n-1}^* \geq a_n$, then $d_n = \max\{a_n, 0\}$ and (26) is solvable. Moreover, the solution $M_n$ satisfies (28) with $s_0 = n$
- if $d_{n-1}^* < a_n$, we take $d = d_{n-1} = d_n$ and we have to solve the following reduced $(n - 1)$-th problem

$$\min_{d_1 \geq \ldots \geq d_{n-2} \geq d \geq 0} \left( \sum_{i=1}^{n-2} (d_i - a_i)^2 + 2(d - \frac{a_{n-1} + a_n}{2})^2 \right). \quad (29)$$
Mathematical properties

For the latter case, we have to solve (29). To do so, we compare $d_{n-2}^*$ and $\frac{a_{n-1}+a_n}{2}$ in the similar way as above:

- if $d_{n-2}^* \geq \frac{a_{n-1}+a_n}{2}$, then $d_{n-1} = d_n = d = \max\{\frac{a_{n-1}+a_n}{2}, 0\}$ and (26) is solvable. Moreover, the solution $M_n$ satisfies (28) with $s_0 = n - 1$;
- if $d_{n-2}^* < \frac{a_{n-1}+a_n}{2}$, we take $d = d_{n-2} = d_{n-1} = d_n$ and we have to solve the following reduced $(n-2)$-th problem

$$
\min_{d_1 \geq \cdots \geq d_{n-3} \geq d \geq 0} \left( \sum_{i=1}^{n-3} (d_i - a_i)^2 + 3(d - \frac{a_{n-2} + a_{n-1} + a_n}{3})^2 \right).
$$

(30)
We shall repeat the above arguments and stop at the following $s_0$-th problem

$$
\min_{d_1 \geq \cdots \geq d_{s_0-1} \geq d \geq 0} \left( \sum_{i=1}^{s_0-1} (d_i - a_i)^2 + (n - s_0 + 1)(d - \frac{\sum_{k=s_0}^{n} a_k}{n-s_0+1})^2 \right).
$$

(31)

where $d = d_{s_0} = \cdots = d_n$ and $s_0$ is defined in (27). Since $d_{s_0-1}^* \geq \frac{\sum_{k=s_0}^{n} a_k}{n-s_0+1}$, we have

$d_{s_0} = \cdots = d_n = d = \max(\frac{\sum_{k=s_0}^{n} a_k}{n-s_0+1}, 0)$. Then (26) is solvable and the solution $M_n$ satisfies (28). Thus, it completes the proof of the theorem.
Deblurring based on WNNM

The proposed model is

$$\min_u \sum_{j \in P} \|R_j u\|_{\tilde{w},*} + \|\nabla u\|_1 + \frac{\lambda}{2} \|A u - g\|_F^2,$$  \hspace{1cm} (32)

where $P$ denotes the set of indices where small image patches exist. The operator $R_j$ firstly collects the similar patches of the reference patch located at $j$, and then stacks those patches into a matrix which should be low rank.

Introduce an auxiliary variable $v$:

$$\min_{u,v} \sum_{j \in P} \|R_j v\|_{\tilde{w},*} + \|\nabla u\|_1 + \frac{\lambda}{2} \|A u - g\|_F^2, \text{ s.t. } v = u. \hspace{1cm} (33)$$

Then using the split Bregman method, we get

$$\begin{cases} (u^{k+1}, v^{k+1}) = \min_{u,v} \sum_{j \in P} \|R_j v\|_{\tilde{w},*} + \|\nabla u\|_1 + \frac{\lambda}{2} \|A u - g\|_F^2 + \frac{\alpha}{2} \|u - v - b^k\|_F^2, \\
 b^{k+1} = b^k + (v^{k+1} - u^{k+1}), \end{cases}$$
Minimization for each patches group

For each similar patches group $R_j \nu$, we have:

$$\| R_j \nu \|_{\tilde{w}^*, \alpha} + \frac{\alpha}{2} \left\| R_j \nu - \left( R_j u^{k+1} - R_j b^k \right) \right\|_F^2.$$ 

Then, using the WNNMG to solve this problem, we have

$$R_j \nu^{k+1} = U \begin{pmatrix} \text{diag} (d_1, d_2, \cdots, d_n) \\ 0 \end{pmatrix} V^T,$$

where

$$\left( R_j u^{k+1} - R_j b^k \right) = U \begin{pmatrix} \text{diag} (\sigma_{ub,1}, \sigma_{ub,2}, \cdots, \sigma_{ub,n}) \\ 0 \end{pmatrix} V^T,$$

and $d_i = \max \left( 0, \left( \sigma_{ub,i} - \frac{w_i^2}{\sigma_{ub,i}} \right) \right)$, $w_i = \frac{c_1 \sqrt{N_{sp}}}{\sigma_i (R_j \nu^{k+1}) + \varepsilon}$, for $i = 1, 2, \cdots, n$. 
Experimental results

Comparison of the PSNR (dB) of the recovered results by different methods, with respect to the noise level $\sigma = 5$.

<table>
<thead>
<tr>
<th>Image</th>
<th>Kernel</th>
<th>ROF</th>
<th>ForWaRD</th>
<th>Framelet</th>
<th>NLTV</th>
<th>BM3DDEB</th>
<th>Ours</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bar.</td>
<td>box</td>
<td>22.73</td>
<td>23.57</td>
<td>23.79</td>
<td>23.12</td>
<td>24.03</td>
<td>24.05</td>
</tr>
<tr>
<td></td>
<td>gaussian</td>
<td>23.11</td>
<td>23.87</td>
<td>24.02</td>
<td>23.32</td>
<td>24.09</td>
<td>24.14</td>
</tr>
<tr>
<td></td>
<td>motion</td>
<td>23.13</td>
<td>23.74</td>
<td>24.22</td>
<td>23.72</td>
<td>25.02</td>
<td>25.15</td>
</tr>
<tr>
<td>Cam.</td>
<td>box</td>
<td>24.00</td>
<td>23.83</td>
<td>24.67</td>
<td>24.73</td>
<td>24.92</td>
<td>25.84</td>
</tr>
<tr>
<td></td>
<td>gaussian</td>
<td>24.17</td>
<td>24.28</td>
<td>24.62</td>
<td>24.64</td>
<td>24.79</td>
<td>25.27</td>
</tr>
<tr>
<td></td>
<td>motion</td>
<td>24.28</td>
<td>24.15</td>
<td>25.15</td>
<td>25.32</td>
<td>25.67</td>
<td>27.02</td>
</tr>
<tr>
<td>Lena</td>
<td>box</td>
<td>27.43</td>
<td>28.52</td>
<td>28.90</td>
<td>28.23</td>
<td>29.38</td>
<td>29.96</td>
</tr>
<tr>
<td></td>
<td>gaussian</td>
<td>28.70</td>
<td>29.93</td>
<td>30.46</td>
<td>29.64</td>
<td>30.85</td>
<td>31.34</td>
</tr>
<tr>
<td></td>
<td>motion</td>
<td>27.48</td>
<td>28.55</td>
<td>29.43</td>
<td>28.78</td>
<td>30.27</td>
<td>31.10</td>
</tr>
</tbody>
</table>
Figure: Recovered results (with PSNR(dB) of different methods on image Cameraman corrupted by $9 \times 9$ uniform blur and Gaussian noise with standard deviation $\sigma = 5$.
Remarks

- Good recovered results (blur + Gaussian noise ∖ Impulse noise);
- Extent to non-Gaussian noises such as Poisson noise and multiplicative noise;
- Extent to other image applications such as image classification.
Summary

There are many non-convex image recover and segmentation models, we consider 3 ways to solve those models

1. Add an extra convex term
2. Relax the functional
3. Compute the analytic solution

Messages:

1. Extra convex term, good to overcome non-convexity
2. Relaxation technique, good for segmentation
3. Sparsity models lead to good image recovery results
References


Contributors

- Xiaohao Cai (University of Cambridge)
- Raymond Chan (The Chinese University of Hong Kong)
- Liyuan Chen (University of Texas at Dallas)
- Yiqiu Dong (Technical University of Denmark)
- Jian Yu (Beijing Jiaotong University)
- Lionel Moisan (Universit Paris Descartes)
- Zhi Li (Michigan Stage University)
- Liyan Ma (Institute of Microelectronics, CAS)
- Mila Nikolova (École Normale Supérieure de Cachan)
- Federica Sciacchitano (Technical University of Denmark)
- Li Xu (Chinese Academy of Science)
- Hongfei Yang (The Chinese University of Hong Kong)
THANK YOU!