Alternating minimization algorithm for shifted speckle reduction variational model

Sangwoon Yun\textsuperscript{1} \quad \textbf{Hyenkyun Woo}\textsuperscript{2}

\textsuperscript{1}School of Computational Science  
Korea Institute for Advanced Study

\textsuperscript{2}Institute of Mathematical Sciences  
Ewha Womans University

International Geoscience and Remote Sensing Symposium,  
July 2011
Motivation

Figure: Proposed Algorithm (1Mpixel@8sec, 2.1GHz CPU Notebook)

- We solve variational models with Total Variation (TV) prior
  \[
  \min_{u \in U} f(u) + \lambda \|\nabla u\|
  \]

- The proposed method is the \textbf{fastest} one to solve total variation (TV) based speckle reduction variational models
Outline

1. Previous Speckle Reduction Methods
   - Convex Variational Models with Total Variation
   - Augmented Lagrangian based algorithms

2. Proposed Speckle Reduction Methods
   - Shifted Variational Models with Total Variation
   - Alternating Minimization Algorithm

3. Numerical Experiments
Outline

1 Previous Speckle Reduction Methods
   - Convex Variational Models with Total Variation
   - Augmented Lagrangian based algorithms

2 Proposed Speckle Reduction Methods
   - Shifted Variational Models with Total Variation
   - Alternating Minimization Algorithm

3 Numerical Experiments
Multiplicative Speckle Noise of L-look SAR image

- $b(\text{SAR image}) = u(\text{unknown reflectivity}) \times \eta(\text{speckle})$
- $b$ is the L-look SAR intensity image
- $\eta$ is fully developed multiplicative speckle noise
- PDF of $\eta$ is the Gamma distribution

$$P(\eta) = \frac{L^L \eta^{L-1}}{\Gamma(L)} e^{-L\eta} \rightarrow E(\eta) = 1; \quad \text{Var}(\eta) = \frac{1}{L}$$
Variational Model with Total Variation regularizer

- Maximum a *Posteriori* (MAP) Criterion:

\[
\max_{u \in U} P(u|b) \propto P(b|u)P(u),
\]

where

\[
P(b|u) = \frac{L^L b^{L-1} e^{-L_b u}}{\Gamma(L)} \frac{e^{-L_u}}{u^L}, \quad P(u) = e^{-\lambda \|\nabla u\|}
\]

- Variational model with Total Variation (Aubert & Aujol 08):

\[
\min_{u \in U} F_b(u) = -\log[P(b|u)P(u)]
\]

\[
= \left\langle \log u + \frac{b}{u}, 1 \right\rangle + \tilde{\lambda} \|\nabla u\|
\]

- Assume that \( u \in U = (0, C]^n \), where \( n \) is the size of image.
Algorithm for MAP based Variational Model

- The original MAP based model:
  \[
  \min_{u \in U} F_b(u) = \left\langle \log u + \frac{b}{u}, \mathbf{1} \right\rangle + \tilde{\lambda} \| \nabla u \|
  \]

- We can find solution by Gradient Descent method
  \[
  \frac{du}{d\tau} = - \frac{\partial F_b(u)}{\partial u}
  \]
  \[
  \frac{u^{k+1} - u^k}{\delta} = \frac{b - u^k}{(u^k)^2} + \tilde{\lambda} \text{div}\left( \frac{\nabla u^k}{|\nabla u^k| + \varepsilon} \right)
  \]

- Difficulties of this approach:
  - \( F_b(u) \) is not convex
  - \( \delta \) is small and the algorithm is slow
  - TV is convex but non-differentiable: \( \partial_u \| \nabla u \| \) is not unique
The original MAP based model:

$$\min_{u \in U} F_b(u) = \left\langle \log u + \frac{b}{u}, 1 \right\rangle + \lambda \| \nabla u \|$$

(Shi & Osher 08) recommended log transformation.

**EXponential model** ($u_l = \log u$)

$$u^* = e^{u_l^*}; \quad u_l^* = \arg \min_{u_l \in \log U} \left\langle u_l + be^{-u_l}, 1 \right\rangle + \lambda \| \nabla u_l \|$$

(Steidl & Teuber 10) show that the solution of the I-divergence model equals to that of the exponential model.

**I-DIVERgence model**

$$u^* = \arg \min_{u \in U} \left\langle u - b \log u, 1 \right\rangle + \lambda \| \nabla u \|$$
Augmented Lagrangian for Exponential Model

$$\min_{u_l, d} \left\{ f(u_l) + g(d) = \langle u_l + be^{-u_l}, 1 \rangle + \lambda \| \nabla d \| \mid u_l = d \right\}$$

Augmented Lagrangian Function:

$$\mathcal{L}_\alpha(u_l, d, p) := f(u_l) + g(d) + \langle p, d - u_l \rangle + \frac{\alpha}{2} \| d - u_l \|^2_2.$$  

- (Bioucas-Dias & Figueiredo 10) proposed MIDAL$^1$:

\[
\begin{aligned}
  u^{k+1}_l &= \arg \min_{u_l \in \log U} \langle u_l + be^{-u_l}, 1 \rangle + \frac{\alpha}{2} \| d^k - u_l + p^k \|^2_2 \\
  d^{k+1} &= \arg \min_{d} \lambda \| \nabla d \| + \frac{\alpha}{2} \| d - u^k_l + p^k \|^2_2 \\
  p^{k+1} &= p^k + (d^{k+1} - u^{k+1}_l).
\end{aligned}
\]

- Use Newton iteration to obtain $u^{k+1}_l$ and solve subproblem to get $d^{k+1}$ at each iteration.

---

$^1$Multiplicative Image Denoising by Augmented Lagrangian
Augmented Lagrangian for I-divergence model

(Steidl & Teuber 10) recommend the following reformulation (ADMM-III)

\[
\begin{align*}
    u^* &= \arg\min_{u \in U} \left\{ \langle u - b \log u, 1 \rangle + \lambda \|z\| \mid \begin{pmatrix} u \\ z \end{pmatrix} = \begin{pmatrix} I \\ \nabla \end{pmatrix} d \right\} \\
\end{align*}
\]

The corresponding Augmented Lagrangian based algorithm becomes

\[
\begin{align*}
    d^{k+1} &= \arg\min_d \left\| \begin{pmatrix} p_1^k \\ p_2^k \end{pmatrix} + \begin{pmatrix} I \\ \nabla \end{pmatrix} d - \begin{pmatrix} u^k \\ z^k \end{pmatrix} \right\|_2^2, \\
    u^{k+1} &= \arg\min_{u \in U} \left\{ \langle u - b \log u, 1 \rangle + \frac{\alpha}{2} \|p_1^k + d^{k+1} - u\|_2^2 \right\}, \\
    z^{k+1} &= \arg\min_z \left\{ \lambda \|z\| + \frac{\alpha}{2} \|p_2^k + \nabla d^{k+1} - z\|_2^2 \right\}, \\
    \begin{pmatrix} p_1^{k+1} \\ p_2^{k+1} \end{pmatrix} &= \begin{pmatrix} p_1^k \\ p_2^k \end{pmatrix} + \begin{pmatrix} I \\ \nabla \end{pmatrix} d^{k+1} - \begin{pmatrix} u^{k+1} \\ z^{k+1} \end{pmatrix}. \\
\end{align*}
\]
Augmented Lagrangian for I-divergence model

We can express the algorithm in a more simplified form:

\[
\begin{align*}
d^{k+1} &= (I - \Delta)^{-1}(u^k - p_1^k - \text{div}(z^k - p_2^k)) \\
u^{k+1} &= \mathcal{P}_U \left[ \frac{1}{2} \left( p_1^k + d^{k+1} - \frac{1}{\alpha} \mathbf{1} + \sqrt{(p_1^k + d^{k+1} - \frac{1}{\alpha} \mathbf{1})^2 + 4 \frac{b}{\alpha}} \right) \right] \\
z^{k+1} &= \text{shrink}(p_2^k + \nabla d^{k+1}, \frac{\lambda}{\alpha}) \\
\begin{pmatrix}
p_1^{k+1} \\
p_2^{k+1}
\end{pmatrix} &= \begin{pmatrix}
p_1^k \\
p_2^k
\end{pmatrix} + \begin{pmatrix}
I \\
\nabla
\end{pmatrix} d^{k+1} - \begin{pmatrix}
u^{k+1} \\
z^{k+1}
\end{pmatrix},
\end{align*}
\]

where

- The discrete cosine transform (DCT) is used to inverse \( I - \Delta \), where \( \Delta \) is the Laplacian operator.
- \( \mathcal{P}_U(u) \) denotes the projection of \( u \) onto the set \( U \)
- \( \text{shrink}(a_i, c) = \max(\|a_i\|_2 - c, 0) \frac{a_i}{\|a_i\|_2} \)
Proposed Speckle Reduction Methods

Outline

1 Previous Speckle Reduction Methods
   • Convex Variational Models with Total Variation
   • Augmented Lagrangian based algorithms

2 Proposed Speckle Reduction Methods
   • Shifted Variational Models with Total Variation
   • Alternating Minimization Algorithm

3 Numerical Experiments
Property of multiplicative speckle noise model \((b = u\eta)\)

- Since \(E(\eta) = 1\), \textbf{averaging property} is invariant w.r.t. \(T\) shifting:
  \[
  E(b + T) = E(u + T), \quad \forall T \in \mathbb{R}_+. 
  \]

- In homogeneous region \(A\), despeckled image \(u^*\) is given by
  \[
  u^* = E_A(b) = E_A(b + T) - T. 
  \]

- We can find solution by the following process

  1. **Shift up**: \(B = b + T\)
  2. **Despeckling**: \(\hat{u}^* = \arg \min_u F_B(u)\)
  3. **Shift down**: \(u^* = \hat{u}^* - T\)

  The solution \(u^*\) roughly equals to the despeckled image without shifting
  \[
  u^* \approx \arg \min_u F_b(u). 
  \]

Shifted Exponential Variational Model

Let $B = b + T$ and $U_T = U + T$, where $T \geq 0$ and $U = [0, C]^n$

- **Shifted EXP model**
  
  $u^* = e^{u_i^*} - T; \quad u_i^* = \arg\min_{u_i\in\log U_T} f(u_i) + \lambda \|\nabla u_i\|$

  where

  \[
  f(u_i) = \langle u_i + B e^{-u_i}, 1 \rangle
  \]

- **The fidelity term $f(u_i)$ is strongly convex**

  \[
  \nabla^2_{u_i} f(u_i) \geq \frac{\min B}{\max U_T} I \geq \frac{T}{C + T} I.
  \]

- The strongly convex modulus of $f(u_i)$ is $\frac{T}{C+T}$. 

S. Yun, H. Woo (KIAS, Ewha Univ.)
Shifted I-divergence Variational Model

Let $B = b + T$ and $U_T = U + T$, where $T \geq 0$ and $U = [0, C]^n$

- Shifted I-DIV model

$$u^* = \tilde{u}^* - T; \quad \tilde{u}^* = \arg\min_{u \in U_T} f(u) + \lambda \| \nabla u \|$$

where

$$f(u) = \langle u - B \log u, 1 \rangle$$

- The fidelity term $f(u)$ becomes strongly convex

$$\nabla_u^2 f(u) \geq \frac{\min B}{(\max U_T)^2} I \geq \frac{T}{(C + T)^2} I.$$

- The strongly convex modulus of $f(u)$ is $\frac{T}{(C + T)^2}$. 
Alternating Minimization Algorithm (Tseng 1991)

\[
\min_{u,z} \left\{ f(u) + \lambda \|z\| \mid \nabla u = z \right\}
\]

- Lagrangian & Augmented Lagrangian Function:
  \[
  \mathcal{L}(u, z, p) := f(u) + \lambda \|z\| + \langle p, z - \nabla u \rangle \\
  \mathcal{L}_\alpha(u, z, p) := f(u) + \lambda \|z\| + \langle p, z - \nabla u \rangle + \frac{\alpha}{2} \|z - \nabla u\|^2_2.
  \]

- Alternating Minimization Algorithm becomes \(0 < \alpha < \sigma/4\)

\[
\begin{align*}
  u^{k+1} &= \arg \min_{u \in \log U} \mathcal{L}(u, z^k, p^k) \\
  z^{k+1} &= \arg \min_{z} \mathcal{L}_\alpha(u^{k+1}, z, p^k) \\
  p^{k+1} &= p^k + \alpha (z^{k+1} - \nabla u^{k+1}),
\end{align*}
\]

where \(\sigma\) is a strongly convex modulus, i.e. \(\nabla^2_u f(u) \geq \sigma I\).
Proposed Speckle Reduction Methods

Alternating Minimization Algorithm

AMA for Shifted Exponential Model

\[
\min_{u_l \in \log U_T, z} \{ f(u_l) + g(z) = \langle u_l + B e^{-u_l}, 1 \rangle + \lambda \|z\| \mid \nabla u_l = z \}
\]

• (Proposed method) AMA for Shifted exponential Model (AMASM) :

\[
\begin{cases}
    u_l^{k+1} = P_{\log U_T} \left[ \log \left( \frac{B}{1 + \text{div} p^k} \right) \right], \\
    z^{k+1} = \text{shrink} \left( \nabla u_l^k - \frac{p^k}{\alpha}, \frac{\lambda}{\alpha} \right), \\
    p^{k+1} = p^k + \alpha (z^{k+1} - \nabla u_l^{k+1}),
\end{cases}
\]

\[
u^* = e^{u_l^{k+1}} - T
\]

where \( B = b + T1, \ U_T = U + T1, \) and

\( P_{\log U_T}(u_l) \) denotes the projection of \( u_l \) onto the set \( \log U_T \)

\( \text{shrink}(a_i, c) = \max(\|a_i\|_2 - c, 0) \frac{a_i}{\|a_i\|_2} \)
Proposed Speckle Reduction Methods

Alternating Minimization Algorithm

AMA for Shifted I-divergence Model

\[
\min_{u \in U_T, z} \{ f(u) + g(z) = \langle u - B \log u, 1 \rangle + \lambda \|z\| \mid \nabla u = z \}
\]

(Proposed method)AMA for Shifted I-divergence Model (AMASM):

\[
\left\{
\begin{array}{l}
    u^{k+1} = \mathcal{P}_{U_T}\left[ \frac{B}{1 + \text{div} p^k} \right], \\
    z^{k+1} = \text{shrink}(\nabla u^k - \frac{p^k}{\alpha}, \frac{\lambda}{\alpha}), \\
    p^{k+1} = p^k + \alpha(z^{k+1} - \nabla u^{k+1}), \\
    u^* = u^{k+1} - T
\end{array}
\right.
\]

where \( B = b + T1, U_T = U + T1 \), and

\( \mathcal{P}_{U_T}(u) \) denotes the projection of \( u \) onto the set \( U_T \)

\( \text{shrink}(a_i, c) = \max(\|a_i\|_2 - c, 0) \frac{a_i}{\|a_i\|_2} \)
AMASM performance vs. $T$ shifting

\[ [\text{EXP}] \sigma_1 = \frac{\min B}{\max U_T} \geq \frac{T}{C+T}, \quad [\text{IDIV}] \sigma_2 = \frac{\min B}{(\max U_T)^2} \geq \frac{T}{(C+T)^2} \]
Outline

1. Previous Speckle Reduction Methods
   - Convex Variational Models with Total Variation
   - Augmented Lagrangian based algorithms

2. Proposed Speckle Reduction Methods
   - Shifted Variational Models with Total Variation
   - Alternating Minimization Algorithm

3. Numerical Experiments
Twelve Test Images (256 × 256, ..., 1500 × 1500).

2.1GHz CPU Notebook (64bit Linux).

All algorithms (AMASM, MIDAL, ADMM-III, PPBit(Nonlocal)) are implemented in C.

Stopping Criterion \( \Vert u^{k+1} - u^k \Vert_2 \leq 3 \times 10^{-4} \Vert u^k \Vert_2 \).
<table>
<thead>
<tr>
<th>IMAGE</th>
<th>Exponential model</th>
<th>I-Divergence model</th>
<th>Nonlocal model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PSNR/Time/Iter</td>
<td>PSNR/Time/Iter</td>
<td>PSNR/Time/Iter</td>
</tr>
<tr>
<td>Barbara</td>
<td>22.2/1.3/17</td>
<td>22.1/0.5/120</td>
<td>22.1/0.4/33</td>
</tr>
<tr>
<td>Boat</td>
<td>23.9/5.7/17</td>
<td>23.7/2.1/114</td>
<td>23.9/2.3/32</td>
</tr>
<tr>
<td>House</td>
<td>24.5/1.2/17</td>
<td>24.3/0.5/111</td>
<td>24.6/0.4/31</td>
</tr>
<tr>
<td>Lena</td>
<td>25.6/5.7/17</td>
<td>25.4/2.1/115</td>
<td>25.6/2.4/33</td>
</tr>
<tr>
<td>Remote1</td>
<td>21.1/4.5/18</td>
<td>21.0/1.9/124</td>
<td>21.1/1.9/32</td>
</tr>
<tr>
<td>Remote2</td>
<td>21.2/6.4/18</td>
<td>21.0/2.3/112</td>
<td>21.2/2.8/32</td>
</tr>
<tr>
<td>Remote3</td>
<td>20.8/1.6/18</td>
<td>20.7/0.6/118</td>
<td>20.7/0.5/32</td>
</tr>
<tr>
<td>Remote4</td>
<td>21.6/5.9/18</td>
<td>21.5/2.3/122</td>
<td>21.6/2.2/31</td>
</tr>
<tr>
<td>Remote5</td>
<td>20.7/5.8/18</td>
<td>20.5/2.4/128</td>
<td>20.7/2.5/34</td>
</tr>
<tr>
<td>Remote6</td>
<td>22.7/14.8/18</td>
<td>22.5/5.1/122</td>
<td>22.7/7.7/33</td>
</tr>
<tr>
<td>Remote7</td>
<td>23.8/26.5/18</td>
<td>23.7/8.5/122</td>
<td>23.8/14.2/33</td>
</tr>
<tr>
<td>Remote8</td>
<td>24.2/55.6/17</td>
<td>24.2/18.6/121</td>
<td>24.2/42.4/32</td>
</tr>
<tr>
<td>Average</td>
<td>22.7/11.2/18</td>
<td>22.6/3.9/119</td>
<td>22.7/6.6/32</td>
</tr>
<tr>
<td></td>
<td></td>
<td>22.6/4.2/141</td>
<td>22.8/180.0</td>
</tr>
</tbody>
</table>

- **[1]** MIDAL : Bioucas-Dias and Figueiredo 2010 IEEE TIP
- **[2]** ADMM-III: Steidl and Teuber 2010 JMIV
- **[3]** PPBit : Deledalle, Denis, and Tupin. 2009 IEEE TIP

Remote7(1000x1000), Remote8(1500x1500) : Megapixel size

Proposed method (AMASM) for the exponential model is the fastest method with reasonable image quality.
Figure: Remote8 (1500x1500) Image
Figure: Barbara (256x256) Image
Figure: Real SAR Image ($ENL = 3.5$), © Sandia NL
Figure: Real SAR Image ($ENL = 3.5$), ©Sandia NL
Conclusion

- Alternating minimization algorithm for shifted exponential model is the fastest one.

- The proposed method is highly parallelizable and so can be accelerated easily by using GPGPU (CUDA).

- We can going to improve PSNR of the proposed method by using nonlocal total variation regularization.
Thank you for your attention!