Edge-Preserving FWI via Regularization by Denoising

Amsalu Anagaw and Mauricio Sacchi

Department of Physics, University of Alberta

SAIG 2018 Annual Meeting
Calgary, AB, Canada

November 26-27, 2018

Outline

- Introduction
- Primal-Dual Total Variation
- Regularization by Denoising (RED) technique for FWI
- Numerical Examples

Introduction

- Preconditioning or regularization method often applied to a misfit function
  - to stabilize ill-posed inverse problems
  - to retrieve and preserve desired features of model parameters
- If edges of the model or if the model is blocky, often an edge-preserving regularization method is used:Total Variation (TV)
  - TV regularization method provides high-resolution images of the subsurface where edges and discontinuities are properly preserved
  - However, finding proper parameters that control nonlinearity of the model that associated with its implementations is cumbersome & is time-consuming to adapt for full waveform inversion

$$J(\mathbf{m}) = ||\mathbf{d}^{obs} - \mathbf{d}^{cal}(\mathbf{m})||^2 + \alpha R(\mathbf{m}) \tag{1}$$

$$R(\mathbf{m}) = \frac{1}{h} \sum_{i,j=1}^{nx,nz} \sqrt{(m_{i+1,j} - m_{i,j})^2 + (m_{i,j+1} - m_{i,j})^2 + \beta} \tag{2}$$

Total Variation Regularization

- Total Variation Regularization

$$J(\mathbf{u}) = \frac{1}{2}||Au - z||^2 + \alpha R(\mathbf{u}), \tag{3}$$

where u is true image, z is noisy image and A is forward modelling operator.

$$R(\mathbf{u}) = \int_\Omega |\nabla u| dxdz = \int_\Omega \sqrt{u_x^2 + u_z^2} dxdz. \tag{4}$$

- Image restoration problem

$$\min_u \int_\Omega (\alpha \sqrt{u_x^2 + u_z^2} + \frac{1}{2}||Au - z||^2) dxdz \tag{5}$$

$$-\alpha \nabla . \left(\frac{\nabla u}{|\nabla u|} \right) + A^*(Au - z) = 0 \tag{6}$$

- if $|\nabla u| = 0$, then the first-order condition cannot be satisfied, and also the second-order (in the Newton's method)

...

- We perturb the TV norm functional:- Primal problem

$$\min_u P(u) = \min_u \int_\Omega \alpha\sqrt{u_x^2 + u_z^2 + \beta}\, dxdz + \frac{1}{2}||Au - z||^2), \tag{7}$$

where β is a small positive parameter.
- Then, image restoration can be performed using a time marching scheme

$$-\alpha\nabla.\left(\frac{\nabla u}{\sqrt{|\nabla u|^2 + \beta}}\right) + A^*(Au - z) = 0 \tag{8}$$

$$u_t = \alpha\nabla.\left(\frac{\nabla u}{\sqrt{|\nabla u|^2 + \beta}}\right) + A^*(Au - z), \quad u(x, 0) = z(x) \tag{9}$$

Effect of β

- if β is too small, a solution to the image restoration with the first- or second-order condition will yield a wrong solution or it might fail
- if β is too large, it will smear the edges of u

Primal-Dual Total Variation

- The term, $\nabla . \left(\frac{\nabla u}{|\nabla u|}\right)$ is high nonlinear and is the source of numerical problems
- In the Primal-Dual TV problem, a unitary vector w is introduced to linearize the problem

$$w = \left(\frac{\nabla u}{|\nabla u|}\right) = \left(\frac{\nabla u}{\sqrt{|\nabla u|^2 + \beta}}\right). \tag{10}$$

- Then, the image restoration problem yields the following system of equations

$$-\alpha \nabla . w + A^*(Au - z) = 0 \tag{11}$$

$$w\sqrt{|\nabla u|^2 + \beta} - \nabla u = 0, \tag{12}$$

where solution to the above equations is solved with $|w|_2 \leq 1$ bound condition

$$w = \begin{cases} \frac{\nabla u}{|\nabla u|} & \text{if } \frac{\nabla u}{|\nabla u|} \neq 0 \\ 1 \text{ or not unique} & \text{if } \frac{\nabla u}{|\nabla u|} = 0 \end{cases}$$

Solution often convergence even if $\beta = 10^{-10}$ or small, and is stable

Primal-Dual Total Variation

- Then, the Primal-Dual problem is given by

$$\min_{u} \max_{|w|_2 \leq 1} \int_{\Omega} -u\nabla.w \; dxdz + \frac{1}{2}||Au - z||^2. \tag{13}$$

- For simple denoising problem if we set $A = I$, then u becomes

$$u(w) = z + \alpha\nabla.w \tag{14}$$

$$\min_{u} \max_{|w|_2 \leq 1} \int_{\Omega} -\alpha u\nabla.w \; dxdz + \frac{1}{2}||u - z||^2. \tag{15}$$

- If we eliminate u in the Primal-Dual problem, we get the Dual problem

$$\max_{|w|_2 \leq 1} D(w) = \max_{|w|_2 \leq 1} \int_{\Omega} -\alpha z(\nabla.w)^2 \; dxdz - \frac{\alpha^2}{2}||\nabla.w||^2 \tag{16}$$

- The Dual problem can be written as in a simple form as

$$\max_{|w|_2 \leq 1} D(w) = \max_{|w|_2 \leq 1} \frac{1}{2}\left[||z||^2 - ||\alpha\nabla.w + z||^2\right] \tag{17}$$

Primal-Dual Total Variation

- The Dual problem can be written as in a simple form as

$$\max_{|w|_2 \leq 1} D(w) = \max_{|w|_2 \leq 1} \frac{1}{2}\left[||z||^2 - ||\alpha \nabla . w + z||^2\right] \tag{18}$$

- The above Dual problem is equivalent to solving

$$\min_{|w|_2 \leq 1} \frac{1}{2}||\nabla . w + \frac{1}{\alpha} z||^2 \tag{19}$$

- Once the Dual problem is solved, the solution to the Primal problem is followed by updating u with

$$u(w) = z + \alpha \nabla . w \tag{20}$$

  Note that w is the dual variable and u is the primal variable. Hence the problem is viewed as a primal-dual problem.

Regularization by Denoising (RED)

- Elad et al 2016 proposed the following regularization term for image denoising problem

$$\rho(\mathbf{m}) = \frac{1}{2}\mathbf{m}^{\mathbf{T}}\left[\mathbf{m} - f(\mathbf{m})\right], \tag{21}$$

  where $f(.)$ is any arbitrary denoising engine of choice and $\mathbf{m}$ is model parameter.

- Conditions and properties $f(\mathbf{m})$
  - For small positive scaling parameter c, $f(\mathbf{m})$ is local homogeneous

$$f(c\mathbf{m}) = cf(\mathbf{m}) \tag{22}$$

  - The Jacobian $\nabla_{\mathbf{m}} f(\mathbf{m})$ of a denoising algorithm is stable and satisfies the condition

$$||f(\mathbf{m})|| \leq ||\mathbf{m}|| \tag{23}$$

- With the above conditions, the gradient of the regularization term leads to the following simple equation

$$\nabla_{\mathbf{m}}\rho(\mathbf{m}) = \mathbf{m} - f(\mathbf{m}) \tag{24}$$

Regularization by Denoising (RED) ...

- In the case of regularized FWI, we minimize $J(\mathbf{m})$ using constrained optimization

$$\begin{aligned} \underset{\mathbf{m}}{\text{minimize}} \quad & J(\mathbf{m}) \\ \text{subject to} \quad & \rho(\mathbf{m}) \leq \epsilon \end{aligned} \tag{25}$$

- Then constrained optimization leads to the following

$$J(\mathbf{m}) = \frac{1}{2}||\mathbf{d}_{obs} - \mathbf{d}(\mathbf{m})_{cal}||^2 + \frac{\lambda}{2}\mathbf{m}^\mathbf{T}\left[\mathbf{m} - f(\mathbf{m})\right] \tag{26}$$

$$J(\mathbf{m}) = J_o(\mathbf{m}) + \frac{\lambda}{2}\mathbf{m}^\mathbf{T}\left[\mathbf{m} - f(\mathbf{m})\right] \tag{27}$$

- The gradient of the misift function becomes

$$\nabla_\mathbf{m} J(\mathbf{m}) = \nabla_\mathbf{m} J_o(\mathbf{m}) + \lambda\mathbf{m} - \frac{\lambda}{2}\left[f(\mathbf{m}) + \nabla_\mathbf{m} f(\mathbf{m})\mathbf{m}\right] \tag{28}$$

- From the second condition of RED, we have

$$\nabla_\mathbf{m} f(\mathbf{m})\mathbf{m} = f(\mathbf{m}) \tag{29}$$

Regularization by Denoising (RED) ...

- Then, the gradient of misfit function of the regularized FWI becomes
$$\nabla_{\mathbf{m}} J(\mathbf{m}) = \nabla_{\mathbf{m}} J_o(\mathbf{m}) + \lambda(\mathbf{m} - f(\mathbf{m})) \tag{30}$$
- Then, model building via regularization by denoising is carried out iterative as
$$\mathbf{m}_{k+1} = \mathbf{m}_k - \alpha \nabla_{\mathbf{m}_k} J_o(\mathbf{m}_k) + \lambda(\mathbf{m}_k - f(\mathbf{m}_{k-1})) \tag{31}$$

Newton's Method!

- In the case of Newton's method or second-order optimization techniques, the model perturbation update will be
$$\delta\mathbf{m} = -\nabla^2_{\mathbf{m}} J_o(\mathbf{m}) + \lambda(\mathbf{I} - \frac{f(\mathbf{m})}{\mathbf{m}}) \tag{32}$$
- Note, by linearized the Total Variation regularization, it satisfies the conditions and properties of of the denoising engine for RED

Examples:- Image Denoising with TV & Primal-Dual TV

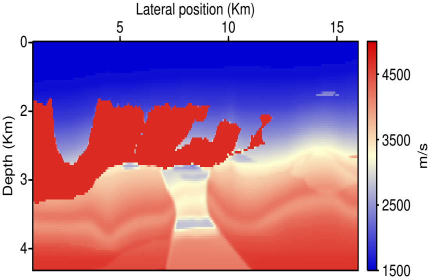


Figure 1 : BP Velocity model

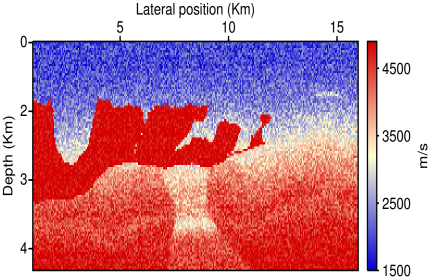


Figure 2 : Model contaminated by noise

Image Denoising with TV

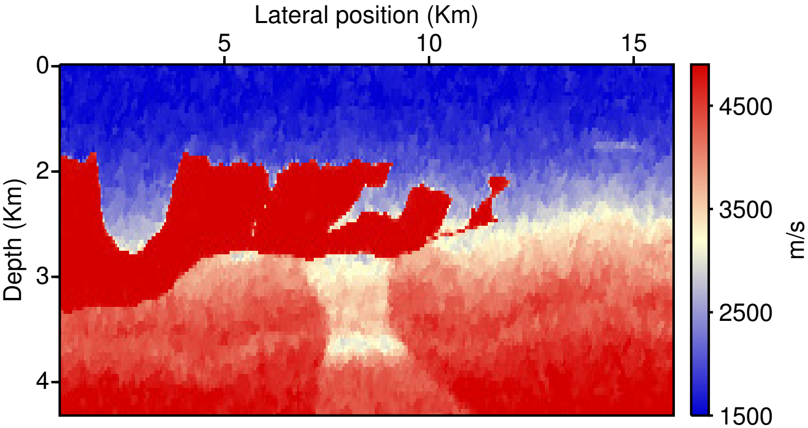
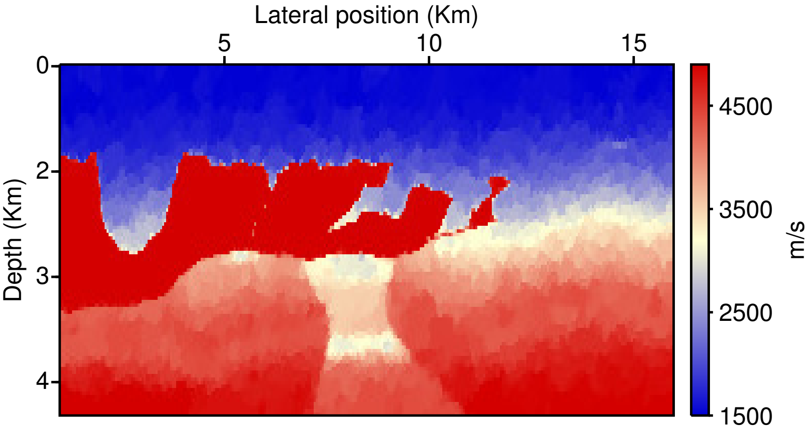
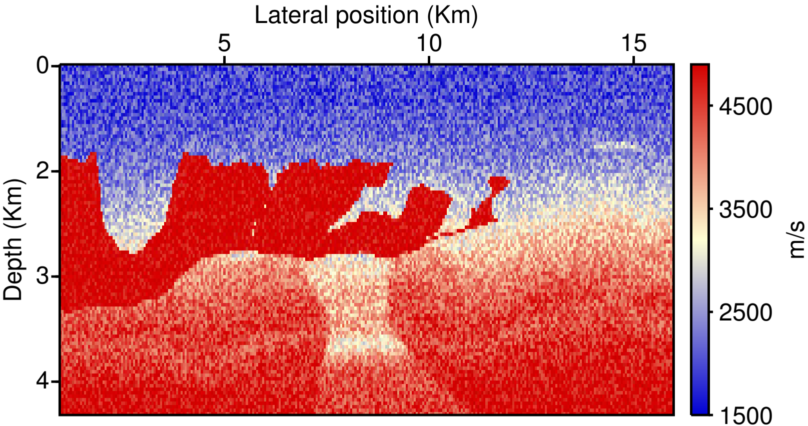


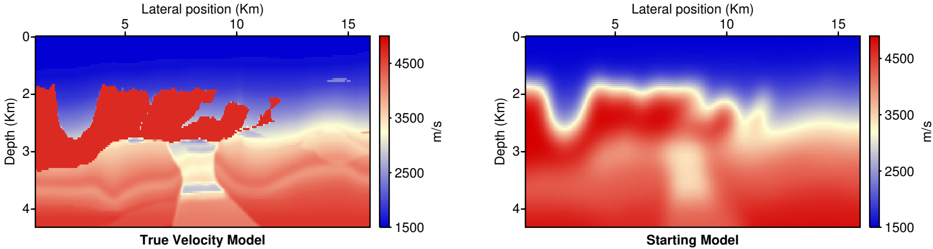
Image Denoising with Primal-Dual TV



Noisy BP Velocity Model

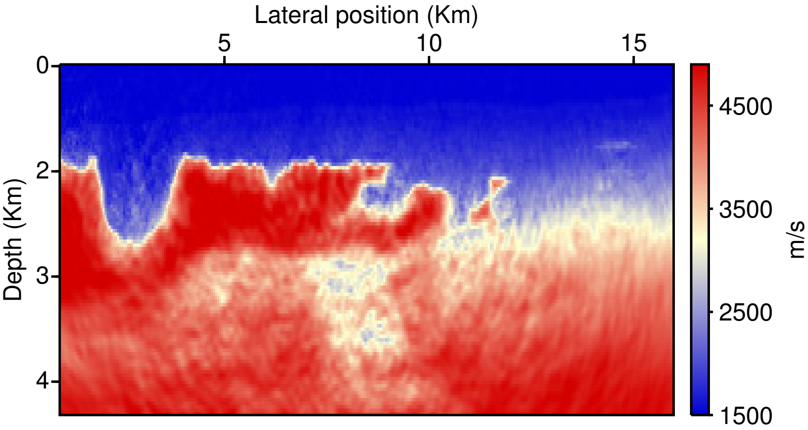


FWI via RED scheme

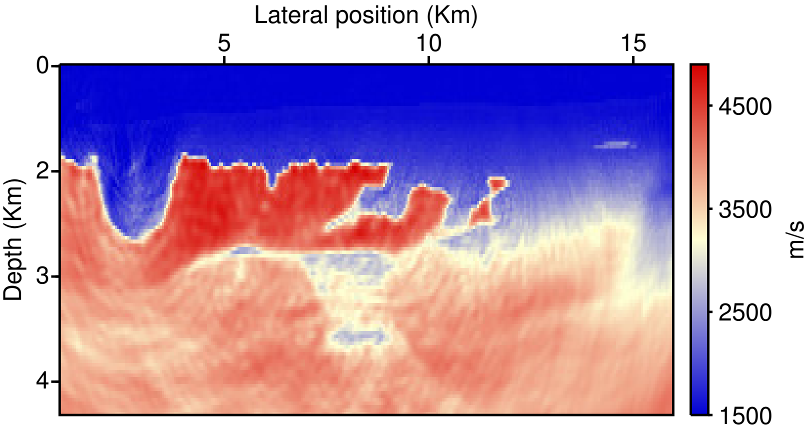


- Number of shots:- 100
- Number of receivers:- 298
- In order to introducing artifacts in the model, we used simultaneous source techniques via source-encoding techniques
  - 4 super-shots are created, where each super-shot contains 25 individual monochromatic sources
- Maximum number of iteration for each frequency group is 50

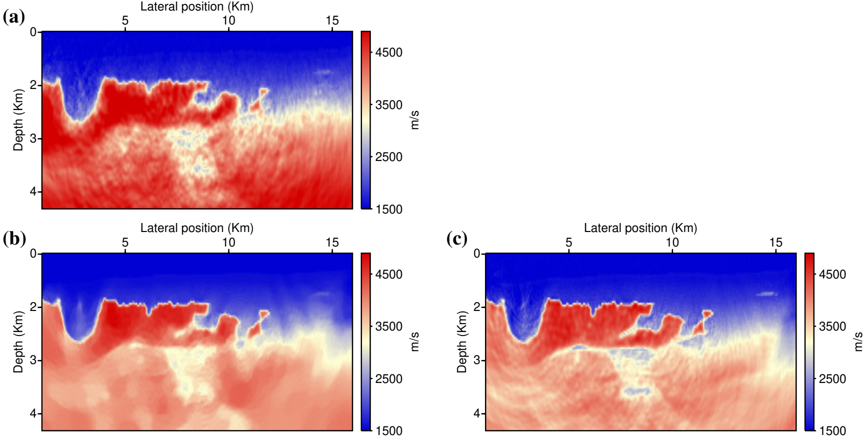
Reconstructed velocity model without regularization



Reconstructed model via RED with Prime-Dual TV

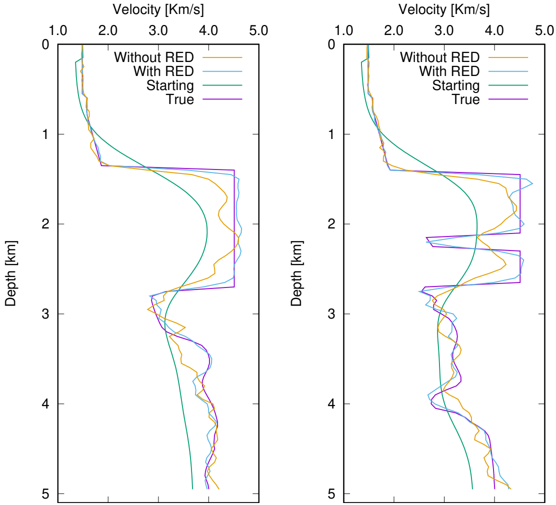


Reconstructed model via RED with Prime-Dual TV



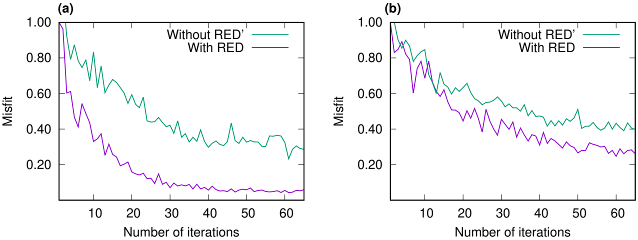
Reconstructed velocity model without RED (a), with RED $\lambda = 1.5$ (b) and $\lambda = 0.08$ (c)

Vertical velocity profiles



Comparison of vertical velocity profiles. The depth velocity profiles are extracted at (4.50 Km,0 Km) and (7.20 km,0 Km)

Misfit function



Relative data misfit reduction curve with and without applying RED with the Primal-Dual Total Variation denosing engine 6 Hz (a) and 12 Hz (b)

Summary

- The aim of this work is to feasibility and advantage of incorporating an edge-preserving denoising algorithm via regularization by denoising (RED) technique in full waveform inversion for velocity model buildings with strong velocity contrasts and sharp discontinuities.
- One advantage of the regularization by denoising algorithm in FWI is the easiness of implementation. The regularization by denoising technique only requires an image denoising engine to handle the structure of the inverse problems.
- For the regularization by denoising technique, we implement the Primal-Dual Total Variation as our denoising engine.
- The primary objective of the Primal-Dual Total Variation denoising technique is to remove some of the singularity caused by the non-differentiability of the L_1 TV norm and is performed by applying a linearization technique.
- The known edge-preserving, for example, TV-norm constraint method requires a cumbersome work in finding optimal parameters that control the nonlinearity properties of the model to impose and retrieve desired features on subsurface images.

Acknowledgements

- We want to acknowledge the sponsors of Signal Analysis and Imaging Group (SAIG) at the University of Alberta and NSERC for supporting this project.

Thank you!