

MA3111: Mathematical Image Processing

Variational Image Deblurring



Suh-Yuh Yang (楊肅煜)

Department of Mathematics, National Central University
Jhongli District, Taoyuan City 320317, Taiwan

First version: October 23, 2016/Last updated: June 30, 2025

Outline of “variational image deblurring”

In this lecture, we will give a brief introduction to the topics:

- *The blurring kernels of motion blur and Gaussian blur.*
- *The standard total variation model for variational image deblurring.*



The material of this lecture is mainly based on

- T. F. Chan and C.-K. Wong, Total variation blind deconvolution, *IEEE Transaction on Image Processing*, 7 (1998), pp. 370-375.
- Y. Wang, W. Yin, and Y. Zhang, A fast algorithm for image deblurring with total variation regularization, *CAAM Technical Report TR 07-10*, 2007, Rice University.

Blurry and noisy image restoration

- **Image restoration (影像修復):** One of the important tasks in image processing is to recover images from noisy and blurry observations.

To recover a sharp image from its blurry observation is the problem known as image deblurring (影像去模糊).

- These blurring artifacts may come from different sources, such as atmospheric turbulence, diffraction (繞射), optical defocusing, camera shaking, and more.
- The blurry and noisy observation is generally modeled as

$$f(x) = (K\bar{u})(x) + n(x), \quad x \in \overline{\Omega},$$

where \bar{u} is the clean image, n is the Gaussian noise, and K is a blurring operator. We may assume the image domain is $\overline{\Omega}$ and zero-valued for all $x \in \mathbb{R}^2 \setminus \overline{\Omega}$.

Linear and shift-invariant blurring operator K

The blurring operator K is typically assumed to be a “*linear*” and “*shift-invariant*” operator, expressed in the convolutional form:

$$(Ku)(x) = \int_{\Omega} h(x-s)u(s)ds =: (h \star u)(x), \quad x \in \overline{\Omega},$$

where \star denotes the convolution operation and h is the so-called point spread function (blurring kernel) associated with the linear blurring operator K .

Therefore, the image deblurring is also called the image deconvolution.

- K is linear: \forall images u and v , and $\forall \alpha, \beta \in \mathbb{R}$

$$\begin{aligned} (K(\alpha u + \beta v))(x) &= \int_{\Omega} h(x-s)(\alpha u(s) + \beta v(s))ds \\ &= \dots = \alpha(Ku)(x) + \beta(Kv)(x), \quad \forall x \in \overline{\Omega}. \end{aligned}$$

- K is shift-invariant (平移不變): Let $g(x) = f(x - \tau)$ for $\tau \in \mathbb{R}^2$. Then

$$\begin{aligned} (Kg)(x) &= \int_{\mathbb{R}^2} h(x-s)g(s)ds = (h \star g)(x) = (g \star h)(x) \\ &= \int_{\mathbb{R}^2} g(x-s)h(s)ds = \int_{\mathbb{R}^2} f(x - \tau - s)h(s)ds \\ &= (f \star h)(x - \tau) = (Kf)(x - \tau), \quad \forall x \in \overline{\Omega}. \end{aligned}$$

Creating a 2-D blurring filter H in Matlab

Motion blur:

```
>> H = fspecial('motion', len, theta)
```

returns a filter to approximate the linear motion of a camera by the length of `len` pixels of the motion, with an angle of `theta` degrees in a counterclockwise direction.

The default `len` is 9 pixels and the default `theta` is 0 degree.

Examples:

```
>> H = fspecial('motion', 5, 45)
```

$$H = \begin{bmatrix} 0 & 0 & 0 & 0.0501 & 0.0304 \\ 0 & 0 & 0.0519 & 0.1771 & 0.0501 \\ 0 & 0.0519 & 0.1771 & 0.0519 & 0 \\ 0.0501 & 0.1771 & 0.0519 & 0 & 0 \\ 0.0304 & 0.0501 & 0 & 0 & 0 \end{bmatrix}$$

Motion blur (cont'd)

```
>> H = fspecial('motion', 5, 30)
```

$$H = \begin{bmatrix} 0 & 0 & 0.0268 & 0.1268 & 0.1464 \\ 0 & 0.1000 & 0.2000 & 0.1000 & 0 \\ 0.1464 & 0.1268 & 0.0268 & 0 & 0 \end{bmatrix}$$

```
>> H = fspecial('motion', 5, 60)
```

$$H = \begin{bmatrix} 0 & 0 & 0.1464 \\ 0 & 0.1000 & 0.1268 \\ 0.0268 & 0.2000 & 0.0268 \\ 0.1268 & 0.1000 & 0 \\ 0.1464 & 0 & 0 \end{bmatrix}$$

A motion filter and blurred image: cameraman

Read image cameraman.png and display it:

```
>> I = imread('cameraman.png');  
>> imshow(I);
```

Create a motion filter and use it to blur the image:

```
>> H = fspecial('motion', 30, 45);  
>> motion.blur = imfilter(I, H, 'replicate');
```

Display the blurred image:

```
>> imshow(motion.blur);
```



Gaussian blur

```
>> H = fspecial('gaussian', hsize, sigma)
```

returns a rotationally symmetric Gaussian lowpass filter of size hsize with standard deviation sigma.

Example:

```
>> H = fspecial('gaussian', 5, 1)
```

$$H = \begin{bmatrix} 0.0030 & 0.0133 & 0.0219 & 0.0133 & 0.0030 \\ 0.0133 & 0.0596 & 0.0983 & 0.0596 & 0.0133 \\ 0.0219 & 0.0983 & 0.1621 & 0.0983 & 0.0219 \\ 0.0133 & 0.0596 & 0.0983 & 0.0596 & 0.0133 \\ 0.0030 & 0.0133 & 0.0219 & 0.0133 & 0.0030 \end{bmatrix}$$

Here `fspecial` creates Gaussian filters using

$$H_g(n_1, n_2) := e^{\frac{-(n_1^2 + n_2^2)}{2\sigma^2}} \quad \text{and} \quad H(n_1, n_2) := \frac{H_g(n_1, n_2)}{\sum_{n_1} \sum_{n_2} H_g}.$$

A Gaussian filter and blurred image: cameraman

Read image cameraman.png and display it:

```
>> I = imread('cameraman.png');  
>> imshow(I);
```

Create a Gaussian filter and use it to blur the image:

```
>> H = fspecial('gaussian', 30, 5);  
>> gaussian.blur = imfilter(I, H, 'replicate');
```

Display the blurred image:

```
>> imshow(gaussian.blur);
```



Blurry and noisy image restoration

The total variation (TV) regularization has become one of the standard techniques known for preserving sharp discontinuities such as edges and object boundaries.

Let $f : \overline{\Omega} \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ be a given blurry and noisy image. The standard TV model recovers an image from f by solving the TV/L2 problem:

$$\min_u \int_{\Omega} |\nabla u(x)| dx + \frac{\lambda}{2} \int_{\Omega} ((Ku)(x) - f(x))^2 dx,$$

where $\lambda > 0$ is a model parameter, K is a linear blurring operator, u is the unknown image to be restored, and

$$|\nabla u(x)| := \|\nabla u(x)\|_2 = \sqrt{(\partial u / \partial x)^2 + (\partial u / \partial y)^2}.$$

We assume that $(Ku)(x) = (h \star u)(x)$ for all $x \in \overline{\Omega}$ and the point spread function h is given.

If both the blur kernel h and the latent sharp image u are unknown, the problem is called “blind image deblurring” or “blind image deconvolution.”

The energy functional

Since the energy functional in the TV/L2 problem is convex, u is optimal if and only if it satisfies the first-order optimality condition. Define the energy functional

$$E[u] := \int_{\Omega} |\nabla u(x)| + \frac{\lambda}{2} ((Ku)(x) - f(x))^2 dx.$$

For any smooth function η with $\eta = 0$ on $\partial\Omega$, let $\Phi(\varepsilon) := E[u + \varepsilon\eta]$, then we have

$$\begin{aligned} 0 &= \Phi'(0) = \frac{d}{d\varepsilon} \Phi(\varepsilon) \Big|_{\varepsilon=0} = \lim_{\varepsilon \rightarrow 0} \frac{E[u + \varepsilon\eta] - E[u]}{\varepsilon - 0} \\ &= \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \left(\int_{\Omega} |\nabla u(x) + \varepsilon \nabla \eta(x)| + \frac{\lambda}{2} ((Ku + \varepsilon K\eta)(x) - f(x))^2 dx \right. \\ &\quad \left. - \int_{\Omega} |\nabla u(x)| + \frac{\lambda}{2} ((Ku)(x) - f(x))^2 dx \right) \\ &= \left(\int_{\Omega} \frac{\nabla u(x) + \varepsilon \nabla \eta(x)}{|\nabla u(x) + \varepsilon \nabla \eta(x)|} \Big|_{\varepsilon=0} \cdot \nabla \eta(x) dx \right) \\ &\quad + \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \frac{\lambda}{2} \left(\int_{\Omega} (\varepsilon (K\eta)(x))^2 + 2\varepsilon (K\eta)(x) ((Ku)(x) - f(x)) \right). \end{aligned}$$

The Euler-Lagrange equation

Then, by Green's formula, we obtain

$$\begin{aligned} 0 &= \int_{\Omega} -\nabla \cdot \frac{\nabla u(\mathbf{x})}{|\nabla u(\mathbf{x})|} \eta(\mathbf{x}) + \lambda(K\eta)(\mathbf{x})((Ku)(\mathbf{x}) - f(\mathbf{x})) d\mathbf{x} \\ &= \int_{\Omega} -\nabla \cdot \frac{\nabla u(\mathbf{x})}{|\nabla u(\mathbf{x})|} \eta(\mathbf{x}) + \lambda\eta(\mathbf{x})K^*((Ku)(\mathbf{x}) - f(\mathbf{x})) d\mathbf{x} \\ &= \int_{\Omega} \left(-\nabla \cdot \frac{\nabla u(\mathbf{x})}{|\nabla u(\mathbf{x})|} + \lambda K^*((Ku)(\mathbf{x}) - f(\mathbf{x})) \right) \eta(\mathbf{x}) d\mathbf{x} \end{aligned}$$

for any smooth function η with $\eta = 0$ on $\partial\Omega$, where K^* is the adjoint operator of K . Therefore, we attain the Euler-Lagrange equation,

$$-\nabla \cdot \left(\frac{\nabla u}{|\nabla u|} \right) + \lambda K^*(Ku - f) = 0 \quad \text{for } \mathbf{x} \in \Omega,$$

or equivalently,

$$\nabla \cdot \left(\frac{\nabla u}{|\nabla u|} \right) - \lambda K^*(Ku - f) = 0 \quad \text{for } \mathbf{x} \in \Omega,$$

along with the Neumann boundary condition, $\partial u(\mathbf{x}) / \partial \mathbf{n} = 0$ on $\partial\Omega$.

The adjoint operator

Let \mathcal{V} be a real (or complex) Hilbert space with inner product $\langle \cdot, \cdot \rangle$, e.g., $L^2(\Omega)$ with the inner product $\langle f, g \rangle := \int_{\Omega} fg \, d\Omega$.

- Consider a continuous (i.e., bounded) linear operator $T : \mathcal{V} \rightarrow \mathcal{V}$. Then the adjoint of T is the continuous linear operator $T^* : \mathcal{V} \rightarrow \mathcal{V}$ satisfying

$$\langle Tx, y \rangle = \langle x, T^*y \rangle, \quad \forall x, y \in \mathcal{V}.$$

- Existence and uniqueness of this operator follows from the Riesz representation theorem.*

Sketch of the proof: Fixed $y \in \mathcal{V}$, then $f(x) := \langle Tx, y \rangle$, $\forall x \in \mathcal{V}$, is a bounded linear functional. By the Riesz representation theorem, \exists unique $z_y \in \mathcal{V}$ such that $\langle Tx, y \rangle = f(x) = \langle x, z_y \rangle$. Define $T^* : \mathcal{V} \rightarrow \mathcal{V}$ by $T^*(y) = z_y$. Then we can show that T^* is a linear bounded operator.

- This can be seen as a generalization of the adjoint matrix of a square matrix, i.e., the conjugate transpose of a square matrix. For example, let $A \in \mathbb{R}^{3 \times 3}$. Then

$$\langle Ax, y \rangle = y^T A x = \langle x, A^T y \rangle, \quad \forall x, y \in \mathbb{R}^3.$$

What is the adjoint operator K^* of K ?

Suppose that the linear and shift-invariant blurring operator $K : L^2(\Omega) \rightarrow L^2(\Omega)$ is defined as

$$(Ku)(x) := (h \star u)(x) = \int_{\Omega} h(x-s)u(s)ds \quad \forall x \in \overline{\Omega},$$

where h is the given kernel function.

$$\begin{aligned} \langle Ku, v \rangle_{L^2(\Omega)} &= \int_{\Omega} \left(\int_{\Omega} h(x-s)u(s)ds \right) v(x)dx \\ &= \int_{\Omega} u(s) \left(\int_{\Omega} h(x-s)v(x)dx \right) ds. \end{aligned}$$

Let $\tilde{h}(x) = h(-x)$ for all $x \in \mathbb{R}^2$. Then for all $u, v \in L^2(\Omega)$, we have

$$\begin{aligned} \langle u, K^*v \rangle_{L^2(\Omega)} &= \langle Ku, v \rangle_{L^2(\Omega)} = \int_{\Omega} u(s) \left(\int_{\Omega} \tilde{h}(s-x)v(x)dx \right) ds \\ &= \langle u, \tilde{h} \star v \rangle_{L^2(\Omega)}. \end{aligned}$$

Therefore, $(K^*v)(x) = (\tilde{h} \star v)(x)$.

Nonlinear PDE based image restoration

Consider the E-L equation with the homogeneous BC, $\frac{\partial u}{\partial n} = 0$ on $\partial\Omega$.

$$\nabla \cdot \left(\frac{\nabla u}{|\nabla u|_\delta} \right) - \lambda K^*(Ku - f) = 0,$$

where $|\cdot|_\delta := \sqrt{|\cdot|^2 + \delta^2}$, $0 < \delta \ll 1$, to avoid division by zero.

- Rudin-Osher (1994) used the artificial time marching method:

$$u \leftarrow u + \Delta t \left\{ \nabla \cdot \left(\frac{\nabla u}{|\nabla u|_\delta} \right) - \lambda K^*(Ku - f) \right\}.$$

This method is very easy to implement but converges slowly due to the nonlinearity of the diffusion operator.

- Vogel-Oman (1996) used *a lagged diffusivity procedure* to partially overcome this difficulty by solving the following equation for $u^{(n+1)}$ iteratively:

$$\nabla \cdot \left(\frac{\nabla u^{(n+1)}}{|\nabla u^{(n)}|_\delta} \right) - \lambda K^*(Ku^{(n+1)} - f) = 0.$$

An equivalent constrained convex problem

By introducing a new variable $w(x) := \nabla u(x)$, we obtain an equivalent constrained convex minimization problem:

$$\min_{u,w} \int_{\Omega} |w(x)|dx + \frac{\lambda}{2} \int_{\Omega} ((Ku)(x) - f(x))^2 dx,$$

subject to $w(x) = \nabla u(x), x \in \Omega.$

Wang-Yin-Zhang (2007) considered the L^2 -norm-square penalty formulation to obtain the unconstrained problem:

$$\min_{u,w} \int_{\Omega} |w(x)|dx + \frac{\lambda}{2} \int_{\Omega} ((Ku)(x) - f(x))^2 dx + \frac{\beta}{2} \int_{\Omega} |w(x) - \nabla u(x)|^2 dx,$$

where $\beta > 0$ is a sufficiently large penalty parameter in order to approximate the solution of the original problem.

The discrete form of the unconstrained problem

Suppose that $f = [f_{ij}]$ is an $N \times N$ digital image. Let us consider the discrete form of the unconstrained problem:

$$\min_{u, w} \sum_{i,j=1}^N \|w_{ij}\| + \frac{\lambda}{2} \|Ku - f\|_F^2 + \frac{\beta}{2} \sum_{i,j=1}^N \|(\partial^+ u)_{ij} - w_{ij}\|^2,$$

where K is the discrete convolution operator, $\|\cdot\|$ is the Euclidean norm in \mathbb{R}^2 , i.e., $\|\cdot\| := \|\cdot\|_2$, and $\|\cdot\|_F$ is the Frobenius norm,

$$w_{ij} = \begin{pmatrix} (w_1)_{ij} \\ (w_2)_{ij} \end{pmatrix} \in \mathbb{R}^2.$$

Moreover, ∂^+ denotes the forward finite difference operator,

$$(\partial^+ u)_{ij} = \begin{pmatrix} (\partial_1^+ u)_{ij} \\ (\partial_2^+ u)_{ij} \end{pmatrix} = \begin{pmatrix} u_{i+1,j} - u_{ij} \\ u_{i,j+1} - u_{ij} \end{pmatrix} \in \mathbb{R}^2.$$

An alternating method

We will solve the discrete problem by alternately minimizing the objective function with respect to w while fixing u , and vice versa.

w -subproblem: For a fixed u , we solve

$$\min_w \sum_{ij=1}^N \left(\|w_{ij}\| + \frac{\beta}{2} \|w_{ij} - (\partial^+ u)_{ij}\|^2 \right),$$

which permits a closed-form solution

$$w_{ij} = \max \left(\|(\partial^+ u)_{ij}\| - \frac{1}{\beta}, 0 \right) \frac{(\partial^+ u)_{ij}}{\|(\partial^+ u)_{ij}\|}, \quad 1 \leq i, j \leq N,$$

where we follow the convention that $0 \cdot (0/0) := 0$. The computation complexity is of order $O(N^2)$.

An alternating method (cont'd)

u -subproblem: For a fixed $w = (w_1, w_2)^\top$, we solve the following problem with a special structure:

$$\min_u \frac{\lambda}{2} \|Ku - f\|_F^2 + \frac{\beta}{2} \|\partial_1^+ u - w_1\|_F^2 + \frac{\beta}{2} \|\partial_2^+ u - w_2\|_F^2,$$

where $Ku = H \star u$ with a given blurring filter H , $\partial_1^+ u = [(\partial_1^+ u)_{ij}]$, $w_1 = [(w_1)_{ij}]$, and so on, and all are matrices in $\mathbb{R}^{N \times N}$.

Therefore, we can solve a linear least-squares problem in the form:

$$\min_u \| \begin{bmatrix} A \\ B \\ C \end{bmatrix} u - \begin{bmatrix} f \\ w_1 \\ w_2 \end{bmatrix} \|_2^2,$$

where u , f , w_1 , and w_2 are vectorization of $[u_{ij}]$, $[f_{ij}]$, $[w_{1ij}]$, and $[w_{2ij}]$, respectively. *However, the linear least-squares solver (by solving the normal equations, or using the QR decomposition, or using the SVD) has high complexity, leading to significant costs!*

u-subproblem: an FFT-based algorithm

We can use the FFT to solve the *u*-subproblem:

- Since $K, \partial_1^+, \partial_2^+$ are all discrete convolutions, if we transform the *u*-subproblem into the Fourier domain, then these operations become element-wise products, e.g., $\mathcal{F}(H \star u) = \mathcal{F}(H) \circ \mathcal{F}(u)$.
- Since the Fourier transform preserves the Frobenius norm, we obtain an equivalent problem (set $\gamma := \beta/\lambda$):

$$\begin{aligned} \min_u & \|\mathcal{F}(H) \circ \mathcal{F}(u) - \mathcal{F}(f)\|_F^2 + \gamma \|\mathcal{F}(\partial_1^+) \circ \mathcal{F}(u) - \mathcal{F}(w_1)\|_F^2 \\ & + \gamma \|\mathcal{F}(\partial_2^+) \circ \mathcal{F}(u) - \mathcal{F}(w_2)\|_F^2. \end{aligned}$$

- After solving for $\mathcal{F}(u)$ (using first-order optimality condition), we obtain the solution to the *u*-subproblem by

$$u = \mathcal{F}^{-1} \left(\frac{\mathcal{F}(H)^* \circ \mathcal{F}(f) + \gamma (\mathcal{F}(\partial_1^+)^* \circ \mathcal{F}(w_1) + \mathcal{F}(\partial_2^+)^* \circ \mathcal{F}(w_2))}{\mathcal{F}(H)^* \circ \mathcal{F}(H) + \gamma (\mathcal{F}(\partial_1^+)^* \circ \mathcal{F}(\partial_1^+) + \mathcal{F}(\partial_2^+)^* \circ \mathcal{F}(\partial_2^+))} \right),$$

where “*” denotes complex conjugacy and the division is element-wise. *Therefore, it requires two ffts and one ifft per iteration.*

Selection of model parameters

- **Noisy level control parameter λ :** An appropriate λ should give a solution u satisfying

$$\|Ku - f\|^2 \approx \|K\bar{u} - f\|^2 = \sigma^2 = \text{Var}(n).$$

- **Constraint penalty parameter β :** Parameter β cannot be too small because it would allow $\nabla u = w$ to be violated excessively. However, β cannot be too large either because the larger the β is the less updates applied to w and u , making the algorithm take more iterations. Therefore, we should choose β in a continuation way to balance the speed and accuracy.
- **Prescribed maximum value β_{max} :** The initial value of β is relatively small (e.g., $\beta = 4$). Then β is increased (e.g., doubled) until a prescribed maximum value β_{max} is reached (e.g., 2^{20}).

Numerical experiments

% creat a blurring filter

```
>> H = fspecial('motion', 41, 135)
```

% add Gaussian white noise with mean 0 and variance 10^{-3}

```
>> f = imnoise(original, 'gaussian', 0, 1e-3)
```

Original image size = 512x512



Blurry image (SNR 5.9065)



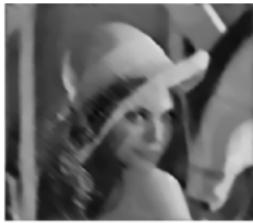
Blurry and noisy image (SNR 5.5328)



$\lambda = 10$ (SNR 7.2649)



$\lambda = 50$ (SNR 9.0242)



$\lambda = 250$ (SNR 10.9436)



(All the numerical experiments are performed by Pei-Chiang Shao)

Numerical experiments

% creat a blurring filter

```
>> H = fspecial('gaussian', 41, 10)
```

% add Gaussian white noise with mean 0 and variance 10^{-6}

```
>> f = imnoise(original, 'gaussian', 0, 1e-6)
```

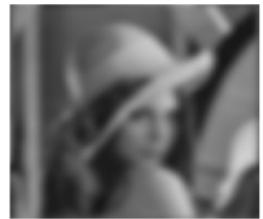
Original image size = 512x512



Blurry image (SNR 6.2287)



Blurry and noisy image (SNR 6.2282)



$\lambda = 10000$ (SNR 9.6682)



$\lambda = 50000$ (SNR 10.5387)



$\lambda = 250000$ (SNR 11.2205)



Total variation blind deconvolution

Chan-Wong (1998) formulated the blind deconvolution problem as

$$\min_{u,h} \frac{1}{2} \int_{\Omega} ((h \star u)(x) - f(x))^2 dx + \alpha_1 \int_{\Omega} |\nabla u(x)| dx + \alpha_2 \int_{\Omega} |\nabla h(x)| dx,$$

where the use of TV regularization for the blurring kernel h is due to the fact that some blurring kernels can have edges.

The first-order optimality conditions give

$$u(-x) \star ((u \star h)(x) - f(x)) - \alpha_2 \nabla \cdot \left(\frac{\nabla h(x)}{|\nabla h(x)|} \right) = 0, \quad x \in \Omega,$$

$$h(-x) \star ((h \star u)(x) - f(x)) - \alpha_1 \nabla \cdot \left(\frac{\nabla u(x)}{|\nabla u(x)|} \right) = 0, \quad x \in \Omega.$$

These are the associated Euler-Lagrange equations, which should be supplemented with the homogeneous Neumann boundary conditions: $\frac{\partial u}{\partial n} = 0$ and $\frac{\partial h}{\partial n} = 0$ on $\partial\Omega$.

A further study is needed!

References

- ① T. F. Chan and C.-K. Wong, Total variation blind deconvolution, *IEEE Transaction on Image Processing*, 7 (1998), pp. 370-375.
- ② P. Getreuer, Total variation deconvolution using split Bregman, *Image Processing On Line*, 2 (2012), pp. 158-174.
- ③ L. I. Rudin, S. Osher, and E. Fatemi, Nonlinear total variation based noise removal algorithms, *Physica D*, 60 (1992), pp. 259-268.
- ④ Y. Wang, W. Yin, and Y. Zhang, A fast algorithm for image deblurring with total variation regularization, *CAAM Technical Report TR 07-10*, 2007, Rice University.
- ⑤ Y. Wang, J. Yang, W. Yin, and Y. Zhang, A new alternating minimization algorithm for total variation image reconstruction, *SIAM Journal on Imaging Sciences*, 1 (2008), pp. 248-272.