Source separation by minimum total variation

Huo Zhuoxi

June 5, 2013

1 Minimization

Total variation (TV) of a given image $\mathbf{f}$ contains N pixels in total is defined as

$$\mathrm{TV}(\mathbf{f}) = \sum_{i}^{N} \left(\nabla^2 f\right)_i^2, \tag{1}$$

where ∇^2 is the Laplacian operator. It smoothes the image $\mathbf{f}$ while minimizing its TV. A mean squared error (MSE) of an image $\mathbf{f}$ with respect to its model $\overline{\mathbf{f}}$ and weighted by reciprocal of the model uncertainty is defined as

$$\mathrm{MSE}\left(\mathbf{f} \,\middle|\, \overline{\mathbf{f}}\right) = \sum_{i}^{N} \frac{\left(f_i - \overline{f}_i\right)^2}{2\sigma_i^2}, \tag{2}$$

where σ_i is the standard deviation of the model $\overline{\mathbf{f}}$ at the i-th pixel.

The objective function is defined as

$$J = \mathrm{TV}(\mathbf{f} - m\mathbf{g}) + \alpha \mathrm{MSE}\left(\mathbf{g} \,|\, \mathbf{p}\right), \tag{3}$$

where $\mathbf{f}$ is the original image contains a foreground point source and a background smooth object, $\mathbf{g}$ is the normalized profile of the separate foreground image, i.e., $\sum_i g_i = 1$, with m being its total flux, $\mathbf{p}$ is the normalized profile of point source, i.e., the point spread function (PSF), and α is the factor balancing weights of TV term and MSE term in the objective function.

The TV term in Eq. 3 is a quadratic function of $\mathbf{g}$

$$\mathrm{TV}(\mathbf{f} - m\mathbf{g}) = m^2 \langle \mathbf{Lg}, \mathbf{Lg} \rangle + \langle \mathbf{Lf}, \mathbf{Lf} \rangle - 2m \langle \mathbf{Lf}, \mathbf{Lg} \rangle, \tag{4}$$

where $\mathbf{L}$ is a toeplitz matrix serves as the Laplacian operator

$$\mathbf{Lf} = \nabla^2 \mathbf{f}, \tag{5}$$

and $\langle \cdot, \cdot \rangle$ denotes the inner product. The MSE term in Eq. 3 is also a quadratic function of $\mathbf{g}$

$$\mathrm{MSE}\left(\mathbf{g} \,|\, \mathbf{p}\right) = \frac{1}{2} \langle \mathbf{Wg}, \mathbf{Wg} \rangle - \langle \mathbf{Wp}, \mathbf{Wg} \rangle, \tag{6}$$

where $\mathbf{W}$ is a diagonal matrix of weights

$$\mathbf{W} = \operatorname{diag}\left(\frac{1}{\sigma_1}, \frac{1}{\sigma_2}, \cdots, \frac{1}{\sigma_N}\right) = \begin{pmatrix} \frac{1}{\sigma_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sigma_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sigma_N} \end{pmatrix}. \tag{7}$$

Therefore the objective function J is a quadratic function of $\mathbf{g}$, i.e.,

$$J(\mathbf{g}, m) = \frac{1}{2}\langle\left(2m^2\mathbf{L}^{\mathrm{T}}\mathbf{L} + \alpha\mathbf{W}^{\mathrm{T}}\mathbf{W}\right)\mathbf{g}, \mathbf{g}\rangle - \langle\left(2m\mathbf{L}^{\mathrm{T}}\mathbf{L} + \alpha\mathbf{W}^{\mathrm{T}}\mathbf{W}\right)\mathbf{f}, \mathbf{g}\rangle, \tag{8}$$

given the constant term $\langle\mathbf{Lf}, \mathbf{Lf}\rangle$ in Eq. 4 is omitted.

The foreground point source is separated from the image through finding $\mathbf{g}$ and m that minimize J. To find m that minimizes J while having $\mathbf{g}$ fixed we can solve the following equation

$$\frac{\partial J}{\partial m} = m\langle\mathbf{Lg}, \mathbf{Lg}\rangle - \langle\mathbf{Lf}, \mathbf{Lg}\rangle = 0 \tag{9}$$

directly and find

$$m = \frac{\langle\mathbf{Lf}, \mathbf{Lg}\rangle}{\langle\mathbf{Lg}, \mathbf{Lg}\rangle}. \tag{10}$$

Minimizing J with fixed m is equivalent to solving the following linear system of linear equations

$$\mathbf{Ag} = \mathbf{b}, \tag{11}$$

where

$$\mathbf{A} = 2m^2\mathbf{L}^{\mathrm{T}}\mathbf{L} + \alpha\mathbf{W}^{\mathrm{T}}\mathbf{W}, \tag{12}$$

and

$$\mathbf{b} = 2m\mathbf{L}^{\mathrm{T}}\mathbf{Lf} + \alpha\mathbf{W}^{\mathrm{T}}\mathbf{Wp}. \tag{13}$$

The following algorithm is designed to find a solution $\mathbf{g}^*$ as well as the corresponding m iteratively.

1. Let $k = 0$, initial total flux $m = 1$, initial profile of separate point source $\mathbf{g}_0 = \mathbf{p}$, the residual $\mathbf{r}_0 = \mathbf{b} - \mathbf{Ag}_0$ and $\mathbf{q}_0 = \mathbf{r}_0$.
2. If the residual $\mathbf{r}_k$ is small enough then let $\mathbf{g}^* = \mathbf{g}_k$ and exit, otherwise proceed to the following loop.
   (a) Update the profile of separate foreground point source by the following conjugate gradient iteration

$$\begin{aligned} a_k &= \frac{\mathbf{r}_k^{\mathrm{T}}\mathbf{r}_k}{\mathbf{q}_k^{\mathrm{T}}\mathbf{A}\mathbf{q}_k}, \\ \mathbf{g}_{k+1} &= \mathbf{g}_k + a_k\mathbf{q}_k, \\ \mathbf{r}_{k+1} &= \mathbf{b} - \mathbf{A}\mathbf{g}_{k+1}, \\ \beta_k &= \frac{\mathbf{r}_{k+1}^{\mathrm{T}}\mathbf{r}_{k+1}}{\mathbf{r}_k^{\mathrm{T}}\mathbf{r}_k}, \\ \mathbf{q}_{k+1} &= \mathbf{r}_{k+1} + \beta_k\mathbf{q}_k. \end{aligned} \tag{14}$$

(b) Impose non-negative constraint

$$g_{i,k+1} = \begin{cases} g_{i,k+1}, & g_{i,k+1} \geq 0 \\ 0, & \text{otherwise} \end{cases} \tag{15}$$

as well as finite-support constraint

$$g_{i,k+1} = \begin{cases} g_{i,k+1}, & i \in S \\ 0, & \text{otherwise} \end{cases} \tag{16}$$

on each pixel of $\mathbf{g}_{k+1}$, where S is the finite support of $\mathbf{g}$.

(c) Update total flux of separate foreground point source

$$m = \frac{\mathbf{f}^{\mathrm{T}}\mathbf{L}^{\mathrm{T}}\mathbf{L}\mathbf{g}_{k+1}}{\mathbf{g}_{k+1}^{\mathrm{T}}\mathbf{L}^{\mathrm{T}}\mathbf{L}\mathbf{g}_{k+1}}. \tag{17}$$

(d) Let $k = k + 1$ and continue.