Course 049064: Variational Methods in Image Processing

Notes 5: TV Denoising and Deconvolution

Guy Gilboa

1 TV Denoising

Model:

$$f = g + n,$$

where g is the clean image, n is random additive noise and f is the input image. The noise model is often white Gaussian noise with standard deviation σ.

There are two energy minimization formulations to solve this (which can be tuned to coincide). The first one is the unconstrained problem with a total energy based on a smoothness and a fidelity term:

$$E(u, f) = J_{TV}(u) + \lambda \|f - u\|^2_{L^2}, \tag{1}$$

where $J_{TV}(u) = \int_\Omega |\nabla u(x)| dx$ and the solution is u which minimizes the above energy. There is one degree of freedom - the value of λ. Usually this is selected in relation to the expected or known standard deviation of the noise. We can denote $v = f - u$ and require $\hat{\text{var}}(v) = \text{var}(n)$, where var() is the true variance, $\text{var}(n) = \sigma^2$, and $\hat{\text{var}}()$ is the empirical variance:

$$\hat{\text{var}}(q) = \frac{1}{|\Omega|} \int_\Omega (q(x) - \bar{q})^2 dx,$$

where $\bar{q} \equiv \frac{1}{|\Omega|} \int_\Omega q(x) dx$ is the mean value of q and $|\Omega|$ is the area of the domain. As $\bar{f} = \bar{u}$ we have $\bar{v} = 0$ and

$$\hat{\text{var}}(v) = \frac{1}{|\Omega|} \|v\|^2_{L^2} = \frac{1}{|\Omega|} \|f - u\|^2_{L^2}.$$

As we expect v to contain mostly noise, this estimation is quite reasonable (although not optimal, in the SNR sense, in most cases, see [5]).

One can formulate this also as a constrained minimization problem:

$$\begin{array}{ll} \min_u \{J_{TV}(u)\} & \text{subject to} \\ \int_\Omega (u(x) - f(x))^2 dx = |\Omega|\sigma^2. & \end{array} \tag{2}$$

We can use the technique of Lagrange multipliers. The functional is defined as:

$$F(u, \lambda) = J_{TV}(u) + \lambda \left(\|u - f\|_{L^2}^2 - |\Omega|\sigma^2 \right).$$

A steady-state is reached for:

$$\begin{array}{lll} (i) & \partial_u F = & -\operatorname{div}\left(\frac{\nabla u}{|\nabla u|}\right) + 2\lambda(u - f) = 0, \\ (ii) & \partial_\lambda F = & \|u - f\|_{L^2}^2 - |\Omega|\sigma^2 = 0. \end{array} \tag{3}$$

For this convex problem one can do the following iterations:

1. Pick some value λ (e.g. $\lambda = \frac{1}{2}$). Initialize $u(t = 0) = f$.

2. Following Eq. (3).(i), evolve until convergence the descent flow:

$$u_t = \operatorname{div}\left(\frac{\nabla u}{|\nabla u|}\right) + 2\lambda(f - u).$$

3. Update λ by

$$\lambda = \frac{1}{2|\Omega|\sigma^2} \int_\Omega \operatorname{div}\left(\frac{\nabla u}{|\nabla u|}\right)(u - f)dx. \tag{4}$$

4. Repeat Steps 2 and 3 until λ converges.

The update of λ in Eq. (4) is reached by multiplying (3).(i) by (u-f), integrating over Ω and using (3).(ii).

Remark: In practice, often one does not need to converge in Step 2, a few numerical iterations can suffice before updating λ.

See more detailed theoretical analysis in [7, 2]. For a spatially varying constrained formulation see [6]. In the implementation, the TV term is often replaced by its epsilon approximation.

2 Deconvolution

Model:

$$f = g * H + n,$$

where H is a *known* blurring kernel and $*$ denotes convolution.

The fidelity term changes and we minimize the following energy:

$$E(u, f) = J_{TV}(u) + \lambda \|f - u * H\|_{L^2}^2, \tag{5}$$

The corresponding constrained problem is

$$\begin{array}{ll} \min_u \{J_{TV}(u)\} & \text{subject to} \\ \int_\Omega (u(x) * H(x) - f(x))^2 dx = |\Omega|\sigma^2. & \end{array} \tag{6}$$

The Lagrange multiplier formulation is

$$F(u, \lambda) = J_{TV}(u) + \lambda \left(\|u * H - f\|_{L^2}^2 - |\Omega|\sigma^2 \right).$$

A steady-state is reached for:

$$\begin{array}{lll} (i) & \partial_u F = & -\operatorname{div}\left(\frac{\nabla u}{|\nabla u|}\right) + 2\lambda \hat{H} * (u * H - f) = 0, \\ (ii) & \partial_\lambda F = & \|u * H - f\|_{L^2}^2 - |\Omega|\sigma^2 = 0, \end{array} \tag{7}$$

where $\hat{H}(x) = H(-x)$. Similar alternating iterations as described above can be used here also to get the correct λ for a given noise variance.

3 Blind Deconvolution

Model:

$$f = g * H + n,$$

where H is an *unknown blurring kernel* that we need to estimate in addition to the deblurred image u.

We follow the paper of Chan and Wong [3]. The assumption is that also the blur kernel is smooth, in the TV sense. Therefore a joint minimization is done with respect to both the deblurred image u and the kernel H:

$$E(u, H, f) = J_{TV}(u) + \alpha J_{TV}(H) + \lambda \|f - u * H\|_{L^2}^2. \tag{8}$$

The basic principle for solving this problem is to use an alternating minimization algorithm where one assumes u is known, solves for H. Then uses the solution found for H and solves for u. These steps are alternated until convergence.

The Euler-Lagrange with respect to each variable are:

$$\begin{array}{ll} \partial_u E(u, H, f) = & -\operatorname{div}\left(\frac{\nabla u}{|\nabla u|}\right) + 2\lambda \hat{H} * (u * H - f) = 0, \\ \partial_H E(u, H, f) = & -\alpha \operatorname{div}\left(\frac{\nabla H}{|\nabla H|}\right) + 2\lambda \hat{u} * (u * H - f) = 0, \end{array} \tag{9}$$

where $\hat{u}(x) = u(-x)$.

This problem can have multiple possible solutions. To increase robustness and regularity of the solution, the following assumptions were made in addition in [3]:

$$\begin{array}{ll} (i) & \int_\Omega H(x)dx = 1, \\ (ii) & u(x) \geq 0;\ H(x) \geq 0, \\ (iii) & H(x) = H(-x). \end{array} \tag{10}$$

The algorithm for solving the problem is:

1. Start with $u^0(x) = f(x)$, $H^0(x) = \delta(x)$. Set $n = 1$.
2. To find H^n, fix $u = u^{n-1}$ and minimize Eq. (8) with respect to H.
3. Impose on H^n:
   (a) $H^n(x) = H^n(x)$ for $x \geq 0$ and 0 otherwise.
   (b) $H^n(x) = (H^n(x) + H^n(-x))/2$.
   (c) $H^n(x) = H^n(x)/\bar{H}$, where $\bar{H} = \int_\Omega H^n(x)dx$.
4. Solve for u^n by fixing $H = H^n$ and minimize Eq. (8) with respect to u.
5. Impose: $u^n(x) = u^n(x)$ for $x \geq 0$ and 0 otherwise.
6. Repeat steps 2-5 until convergence.

4 Numerics for solving TV

Recent sophisticated numerical schemes to solve TV efficiently will be presented in the next lesson. We describe below three more classical methods: the simplest explicit method, Vogel-Oman lagged diffusivity [8] and Chambolle's projection [1]. We begin by defining discrete gradient and divergence operators in 2D (using the notations of [1]). We assume a $N \times N$ pixel grid with indices $i, j = 1, .., N$ specifying the row and column.

Gradient: Let the discrete gradient for pixel i, j be defined as:

$$(\nabla u)_{i,j} := \left((\nabla u)^1_{i,j}, (\nabla u)^2_{i,j}\right), \tag{11}$$

where

$$\begin{array}{l} (\nabla u)^1_{i,j} := \begin{cases} u_{i+1,j} - u_{i,j}, & \text{if } i < N, \\ 0, & \text{if } i = N, \end{cases} \\ (\nabla u)^2_{i,j} := \begin{cases} u_{i,j+1} - u_{i,j}, & \text{if } j < N, \\ 0, & \text{if } j = N, \end{cases} \end{array} \tag{12}$$

Divergence: Let a vector p be defined as $p = (p^1, p^2)$ (where p^k is a discrete 2D image). The discrete divergence is

$$(\operatorname{div} p)_{i,j} = \begin{cases} p^1_{i,j} - p^1_{i-1,j}, & \text{if } 1 < i < N, \\ p^1_{i,j}, & \text{if } i = 1, \\ -p^1_{i-1,j}, & \text{if } i = N, \end{cases} + \begin{cases} p^2_{i,j} - p^2_{i,j-1}, & \text{if } 1 < i < N, \\ p^2_{i,j}, & \text{if } i = 1, \\ -p^2_{i,j-1}, & \text{if } i = N. \end{cases} \tag{13}$$

This divergence definition admits an adjoint relation with the gradient defined above: $\langle -\operatorname{div} p, u\rangle = \langle p, \nabla u\rangle$.

4.1 Explicit method

Approximates $J_{TV-\varepsilon}(u) + \lambda\|f-u\|^2_{L^2}$. Let

$$F(u) = \operatorname{div}\left(\frac{\nabla u}{\sqrt{|\nabla u|^2 + \varepsilon^2}}\right) + 2\lambda(f-u), \tag{14}$$

where ∇, and div are the discrete operators defined in Eqs. (11), (13), respectively. Δt is bounded by the CFL condition.

Algorithm: initialize with $u^0 = f$, set $n = 0$. Evolve until convergence:

$$u^{n+1} = u^n + \Delta t F(u^n).$$

4.2 Lagged diffusivity

Approximates $J_{TV-\varepsilon}(u) + \lambda\|f-u\|^2_{L^2}$.

We write the Euler-Lagrange equation:

$$\operatorname{div}\left(\frac{\nabla u}{\sqrt{|\nabla u|^2 + \varepsilon^2}}\right) + 2\lambda(f-u) = 0.$$

We assume to have a previous value u^n (initializing with $u_0 = f$) and compute a new u^{n+1}. We can fix the diffusivity and define the following linear operator, which operates on v:

$$L_u v = \operatorname{div}\left(\frac{\nabla}{\sqrt{|\nabla u|^2 + \varepsilon^2}}\right) v.$$

Using the E-L equation, changing order of variables, constructing L_{u^n} based on the previous iteration and dividing by 2λ we get

$$(1 - \frac{1}{2\lambda} L_{u^n}) u^{n+1} = f. \tag{15}$$

Thus we have to solve a linear system of equations. The system is very sparse and there are several efficient ways to solve this. See more details in [8, 4].

References

[1] A. Chambolle. An algorithm for total variation minimization and applications. *JMIV*, 20:89–97, 2004.

[2] A. Chambolle and P.L. Lions. Image recovery via total variation minimization and related problems. *Numerische Mathematik*, 76(3):167–188, 1997.

[3] T.F. Chan and C.-K. Wong. Total variation blind deconvolution. *Image Processing, IEEE Transactions on*, 7(3):370–375, 1998.

[4] Tony F Chan and Pep Mulet. On the convergence of the lagged diffusivity fixed point method in total variation image restoration. *SIAM journal on numerical analysis*, 36(2):354–367, 1999.

[5] G. Gilboa, N. Sochen, and Y.Y. Zeevi. Estimation of optimal PDE-based denoising in the SNR sense. *IEEE Trans. on Image Processing*, 15(8):2269–2280, 2006.

[6] G. Gilboa, N. Sochen, and Y.Y. Zeevi. Variational denoising of partly-textured images by spatially varying constraints. *IEEE Trans. on Image Processing*, 15(8):2280–2289, 2006.

[7] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.

[8] R.V. Vogel and M.E. Oman. Iterative methods for total variation denoising. *SIAM J. Scientific Computing*, 17(1):227–238, 1996.