42-505/705: Variational Image Processing
Assignment 2: Image Denoising – Theory
Due 6 October 2004

The problem of removing noise from an image without blurring sharp edges can be formulated as an infinite-dimensional minimization problem. Given a possibly noisy image $u_0(x, y)$, we would like to find the image $u(x, y)$ that is closest in the L_2 sense, i.e. we want to minimize

$$\mathcal{F}_{LS} := \int_\Omega (u - u_0)^2 \, d\boldsymbol{x},$$

while also removing noise, which is assumed to comprise very "rough" components of the image. This latter goal can be incorporated as an additional term in the objective, in the form of a penalty,

$$\mathcal{R}_{TN} := \frac{1}{2} \int_\Omega k(\boldsymbol{x}) \nabla u \cdot \nabla u \, d\boldsymbol{x},$$

where as in the previous assignment $k(\boldsymbol{x})$ is a "diffusion" coefficient that controls how strongly we impose the penalty, i.e. how much smoothing occurs. Unfortunately, if there are sharp edges in the image, this so-called *Tikhonov (TN) regularization* will blur them. Instead, in these cases we prefer the so-called *total variation (TV) regularization*,

$$\mathcal{R}_{TV} := \int_\Omega k(\boldsymbol{x}) (\nabla u \cdot \nabla u)^{\frac{1}{2}} \, d\boldsymbol{x}$$

where (we will see that) taking the square root is the key to edge preservation. Since $\mathcal{R}_{TV}$ is not differentiable when $\nabla \boldsymbol{u} = \boldsymbol{0}$, it is usually modified to include a positive parameter ε as follows:

$$\mathcal{R}^\varepsilon_{TV} := \int_\Omega k(\boldsymbol{x}) (\nabla u \cdot \nabla u + \varepsilon)^{\frac{1}{2}} \, d\boldsymbol{x}.$$

We wish to study the performance of the two denoising functionals $\mathcal{F}_{TN}$ and $\mathcal{F}^\varepsilon_{TV}$, where

$$\mathcal{F}_{TN} := \mathcal{F}_{LS} + \mathcal{R}_{TN}$$

and

$$\mathcal{F}^\varepsilon_{TV} := \mathcal{F}_{LS} + \mathcal{R}^\varepsilon_{TV}.$$

In this assignment we will study analytical aspects of the denoising problem; the next assignment will investigate computational aspects. This problem also serves as an excellent example of nonlinear minimization. We will use Newton optimization methods to minimize the two functionals. For boundary conditions, you can take the homogeneous Neumann condition $\nabla u \cdot \boldsymbol{n} = 0$, which amounts to assuming that the image intensity does not change normal to the boundary.

1. For both $\mathcal{F}_{TN}$ and $\mathcal{F}^\varepsilon_{TV}$, derive the first order necessary condition for optimality using calculus of variations, in both weak form and strong form. Use v to represent the test function.

2. Show that when ∇u is zero, $\mathcal{R}_{TV}$ is not differentiable, but $\mathcal{R}^\varepsilon_{TV}$ is.

3. For both $\mathcal{F}_{TN}$ and $\mathcal{F}^\varepsilon_{TV}$, derive the infinite-dimensional Newton step, in both weak and strong form. For consistency of notation, please use $\tilde{u}$ as the differential of u (i.e. Newton step). The strong form of the second variation of $\mathcal{F}^\varepsilon_{TV}$ will give you an anisotropic diffusion operator of the form $-\nabla \cdot (\boldsymbol{A}(u)\nabla u)$, where $\boldsymbol{A}(u)$ is an anisotropic tensor that plays the role of the diffusivity coefficient. (In contrast, you can think of the second variation of $\mathcal{F}_{TN}$ giving an *isotropic* diffusion operator, i.e. with $\boldsymbol{A} = \alpha \boldsymbol{I}$ for some α.)

4. Derive expressions for the two eigenvalues and corresponding eigenvectors of $\boldsymbol{A}$. Based on these expressions, give a explanation of why $\mathcal{F}_{TV}^{\varepsilon}$ is effective at preserving sharp edges in the image, while $\mathcal{F}_{TN}$ is not.

5. Show the equivalence between the following two approaches:

   - *Ritz method*: First make a finite element approximation of the infinite-dimensional functional $\mathcal{F}_{TV}^{\varepsilon}$, and then derive the finite-dimensional Newton step.
   - *Galerkin method*: Directly make a finite element approximation of the (weak form of the) infinite-dimensional Newton step you derived above.

6. Based on the results of Part 5 above, show that for large enough ε, $\mathcal{R}_{TV}^{\varepsilon}$ behaves like $\mathcal{R}_{TN}$, and for $\varepsilon = 0$, the discrete Hessian of $\mathcal{R}_{TV}^{\varepsilon}$ is singular. This suggests that ε should be chosen small enough that edge preservation is not lost, but not too small that ill-conditioning occurs.