Total Variation Inpainting using Split Bregman

Pascal Getreuer

Yale University (`pascal.getreuer@yale.edu`)

Abstract

Given an image where a specified region is unknown, *image inpainting* or *image completion* is
the problem of inferring the image content in this region. Traditional retouching or inpainting is
the practice of restoring aged artwork, where damaged or missing portions are repainted based
on the surrounding content to approximate the original appearance. In the context of digital
images, inpainting is used to restore regions of an image that are corrupted by noise or where
the data is missing. Inpainting is also used to solve disocclusion, to estimate the scene behind an
obscuring foreground object. A popular use of digital inpainting is object removal, for example,
to remove a trashcan that disrupts a scene of otherwise natural beauty.

Inpainting is an interpolation problem, filling the unknown region with a condition to agree
with the known image on the boundary. A classical solution for such an interpolation is to solve
Laplace's equation. However, Laplace's equation is usually unsatisfactory for images since it is
overly smooth. It cannot recover a step edge passing through the region.

Total variation (TV) regularization is an effective inpainting technique which is capable of
recovering sharp edges under some conditions (these conditions will be explained). The use of
TV regularization was originally developed for image denoising by Rudin, Osher, and Fatemi [3]
and then applied to inpainting by Chan and Shen [13]. TV-regularized inpainting does not create
texture, the method is limited to inpainting the geometric structure.

Source Code

ANSI C source code to produce the same results as the demo is accessible on the article web
page http://dx.doi.org/10.5201/ipol.2012.g-tvi. Future software releases and updates
will be posted at http://dev.ipol.im/~getreuer/code.

1 Introduction

The general problem of interpolation has a very long history and numerous works across multiple
disciplines, see for example the chronology by Meijering [15]. In the context of imaging, inpainting
is closely-related to image replacement [5], image editing [10], image zooming [11, 13], and edge
completion [4].

The first variational approach to the image inpainting problem was Nitzberg and Mumford's 2.1-D
sketch [2], based on a variant of the Mumford–Shah functional, and the second variational approach
was the work of Masnou and Morel [6], based on solving for level lines with minimal curvature.
Bertalmío, Sapiro, Caselles, and Ballester [9] introduced the term "image inpainting" in analogy to
artistic inpainting and proposed an anisotropic diffusion PDE model.

Chan and Shen [13] applied TV regularization to inpainting. Chan and Shen have also suggested an improvement, a second order curvature-driven diffusion [12] that can connect level lines across larger gaps.

In order to inpaint textures, the method of Igehy and Pereira [5] fills the unknown region with texture synthesized from a second image. Efros and Leung [8] proposed a patch copy-paste method to synthesize texture within the same image, and Criminisi, Pérez, and Toyama [17] refined the fill-order to improve structural propagation. The methods of Bertalmío, Vese, Sapiro, and Osher [16] and Elad, Starck, Querre, and Donoho [19] are attempts to mix the variational and patch-based approaches to recover both structure and texture.

2 H^1 Inpainting (Laplace's Equation)

Laplace's equation is a classical solution to the inpainting problem. Let $f : \Omega \to \mathbb{R}$ be a given grayscale image and let $D \subset \Omega$ be an open set representing the region to be inpainted. It is supposed that f is known in $\Omega \backslash D := \{x \in \Omega : x \notin D\}$ and unknown in D. The inpainting solution by Laplace's equation satisfies

$$\begin{cases} \Delta u = 0 & \text{in } D, \\ u = f & \text{in } \partial D, \end{cases} \tag{1}$$

where Δ denotes the Laplacian and $u = f$ in $\Omega \backslash D$.

The $H^1(D)$ seminorm of u is defined as the $L^2(D)$ norm of its gradient,

$$|u|^2_{H^1(D)} := \int_D |\nabla u(x)|^2 \, dx, \tag{2}$$

where $|\nabla u(x)|$ denotes the Euclidean magnitude of the gradient at x. The function space H^1 is the set of L^2 functions with finite H^1 seminorm, i.e., the Sobolev space $W^{1,2}$,

$$H^1(D) := \{u \in L^2(D) : |u|_{H^1(D)} < \infty\}. \tag{3}$$

Dirichlet's principle [7, section 2.2.5] states that if the Laplace inpainting solution is in $C^2(D)$, then it has the minimal H^1 seminorm among all C^2 functions satisfying the boundary condition. The connection is through Green's identity: if $\Delta u = 0$, then

$$0 = -\int_D (\Delta u) v \, dx = \int_D \nabla u \cdot \nabla v \, dx \quad \text{for all } v \in C^2_0(D), \tag{4}$$

and from the other direction, if u minimizes the H^1 seminorm, then

$$0 = \partial_\epsilon |u + \epsilon v|^2_{H^1}\big|_{\epsilon=0} = \int_D \nabla u \cdot \nabla v \, dx \quad \text{for all } v \in H^1_0(D). \tag{5}$$

3 TV Inpainting

Total variation (TV) inpainting replaces H^1 with the functions of bounded variation (BV) and the TV seminorm. A function u is in $BV(\Omega)$ if there exists a Radon measure Du such that

$$\int_\Omega u(x) \operatorname{div} \vec{g}(x) \, dx = -\int_\Omega \langle \vec{g}, Du(x) \rangle \quad \text{for all } \vec{g} \in C^1_c(\Omega, \mathbb{R}^2)^2, \tag{6}$$

and the *total variation seminorm* of u is

$$\|u\|_{\text{TV}(\Omega)} := \int_\Omega |Du| := \sup\left\{ \int_\Omega u \operatorname{div} \vec{g} \, dx : \vec{g} \in C^1_c(\Omega, \mathbb{R}^2)^2, \sqrt{g_1^2 + g_2^2} \le 1 \right\}. \tag{7}$$

When u is smooth, TV is the integral of its gradient magnitude,

$$\|u\|_{\mathrm{TV}(\Omega)} = \int_\Omega |\nabla u|\, dx. \tag{8}$$

The TV inpainting method is to find the BV function u that solves the minimization problem

$$\underset{u \in BV(\Omega)}{\arg\min} \|u\|_{\mathrm{TV}(\Omega)} + \frac{\lambda}{2} \int_{\Omega \backslash D} \big(f(x) - u(x)\big)^2 dx, \tag{9}$$

where λ is a positive parameter. This minimization problem is identical to the Rudin, Osher, and Fatemi denoising problem [3] except that here the second integral is over $\Omega \backslash D$ instead of Ω. If the exact solution is in BV and takes values in $[0, 1]$, then minimizers u exist but are generally not unique [14].

Inpainting may be viewed as denoising with a spatially-varying regularization strength $\lambda(x)$,

$$\underset{u}{\arg\min} \|u\|_{\mathrm{TV}(\Omega)} + \frac{1}{2} \int_\Omega \lambda(x)\big(f(x) - u(x)\big)^2 dx, \tag{10}$$

where in the case of inpainting $\lambda(x)$ is set to zero in D and $\lambda(x) > 0$ outside of D. For $x \in D$ where $\lambda(x) = 0$, the value $f(x)$ is unused and $u(x)$ is only influenced by the $\|u\|_{\mathrm{TV}}$ term. Outside of D, the model performs TV-regularized denoising and $\lambda|_{\Omega \backslash D}$ specifies the denoising strength. This denoising behavior may be desirable when it is difficult to specify the inpainting domain accurately. By setting λ to a very large value, the denoising effect is limited so that the image remains nearly unchanged outside of D.

4 Split Bregman

We describe the split Bregman algorithm of Goldstein and Osher [21] applied to inpainting as developed in tvreg [22]. Due to the similarity between TV denoising and TV inpainting, the split Bregman algorithm for denoising [23] is readily adapted to inpainting. This inpainting algorithm is identical to the denoising algorithm aside from the spatially-varying λ in the u subproblem.

Our approach to discrete derivatives is to use one-sided differences and half-sample symmetric extension at points near the boundaries. In one dimension, we define the discrete derivative ∂ of a uniformly-sampled signal $f_0, f_1, \ldots, f_{N-1}$ as its forward difference,

$$\begin{pmatrix} \partial f_0 \\ \partial f_1 \\ \vdots \\ \partial f_{N-2} \\ \partial f_{N-1} \end{pmatrix} = \begin{pmatrix} -1 & 1 & & & \\ & -1 & 1 & & \\ & & \ddots & \ddots & \\ & & & -1 & 1 \\ & & & & 0 \end{pmatrix} \begin{pmatrix} f_0 \\ f_1 \\ \vdots \\ f_{N-2} \\ f_{N-1} \end{pmatrix}. \tag{11}$$

Since the half-sample symmetric extension is $2N$-periodic, the discrete gradient may also be viewed as a cyclic convolution of the reflected signal $(f_0, \ldots, f_{N-1}, f_{N-1}, \ldots, f_0)$ with the filter $h_{-1} = 1$, $h_0 = -1$, and h equal to zero otherwise.

In two dimensions, we define the discrete gradient of an $N \times N$ image u using the discrete derivative as $\nabla u = (\partial_x u, \partial_y u)^T$, where the subscript on ∂ denotes the dimension along which the difference is applied. We define discrete divergence through the relationship $\mathrm{div} := -\nabla^* = -\partial_x^* - \partial_y^*$, where $*$ denotes adjoint, and the discrete Laplacian as $\Delta := \mathrm{div}\, \nabla$. In the image interior, the discrete divergence is $\mathrm{div}\, v = v_{i,j}^x - v_{i-1,j}^x + v_{i,j}^y - v_{i,j-1}^y$ and the discrete Laplacian is the 5-point formula $\Delta u = -4u_{i,j} + u_{i+1,j} + u_{i-1,j} + u_{i,j+1} + u_{i,j-1}$. Please refer to the section on discrete derivatives in the denoising article [23] for details.

Many different discretizations of the TV seminorm have been proposed, see the denoising article [23] for discussion. Here, TV is approximated by summing the vector magnitude $|\nabla u_{i,j}|$ over all pixels,

$$\|u\|_{\text{TV}(\Omega)} \approx \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} |\nabla u_{i,j}|, \tag{12}$$

where ∇u is the discrete gradient developed above. The split Bregman algorithm is applied to solve the constrained minimization problem

$$\begin{cases} \arg\min_{d,u} \sum_{i,j} |d_{i,j}| + \dfrac{1}{2} \sum_{i,j} \lambda_{i,j}(f_{i,j} - u_{i,j})^2 \\ \text{subject to } d = \nabla u. \end{cases} \tag{13}$$

The auxiliary variable d is a vector field which is constrained to equal ∇u. The Bregman iteration [1] method is applied to solve the constrained problem. In each iteration of the Bregman method, the following problem is solved:

$$\arg\min_{d,u} \sum_{i,j} |d_{i,j}| + \frac{1}{2} \sum_{i,j} \lambda_{i,j}(f_{i,j} - u_{i,j})^2 + \frac{\gamma}{2} \sum_{i,j} |d_{i,j} - \nabla u_{i,j} - b_{i,j}|^2 \tag{14}$$

where b is a variable related to the Bregman iteration algorithm and the parameter γ may be set to any positive constant. The joint minimization over d and u is solved by an alternating direction method, in each step minimizing either d or u while keeping the other variable fixed.

d **subproblem**   With u fixed, the d subproblem is

$$\arg\min_{d} \sum_{i,j} |d_{i,j}| + \frac{\gamma}{2} \sum_{i,j} |d_{i,j} - \nabla u_{i,j} - b_{i,j}|^2. \tag{15}$$

This problem is solved as in [23]:

$$d_{i,j} = \frac{\nabla u_{i,j} + b_{i,j}}{|\nabla u_{i,j} + b_{i,j}|} \max\big\{|\nabla u_{i,j} + b_{i,j}| - 1/\gamma, 0\big\}. \tag{16}$$

u **subproblem**   With d fixed, the u subproblem is

$$\arg\min_{u} \frac{1}{2} \sum_{i,j} \lambda_{i,j}(f_{i,j} - u_{i,j})^2 + \frac{\gamma}{2} \sum_{i,j} |d_{i,j} - \nabla u_{i,j} - b_{i,j}|^2. \tag{17}$$

The optimal u satisfies

$$\tfrac{1}{\gamma}\lambda u - \Delta u = \tfrac{1}{\gamma}\lambda f - \text{div}(d - b), \tag{18}$$

where div and Δ are the discrete divergence and Laplacian defined above. Analogous to Goldstein and Osher's approach for denoising, we approximate the solution to this equation using one sweep of Gauss–Seidel per Bregman iteration. The subproblem is solved once for each Bregman iteration, so the combined effect of the sweeps over multiple iterations solves the subproblem accurately.

The auxiliary variable b is initialized to zero and updated according to

$$b^{k+1} = b^k + \nabla u - d. \tag{19}$$

The penalty parameter γ may be selected as any positive constant. The limiting solution exactly satisfies the constraint $d = \nabla u$ for any $\gamma > 0$. As discussed in the denoising article [23], γ should be a moderate value for good convergence speed.

4

In the d subproblem, the solution d is the vector shrinkage of $(\nabla u + b)$ by $1/\gamma$. The shrinkage has the greatest effect when γ is small. In the u subproblem, however, the solution u satisfies

$$\lambda u - \gamma \Delta u = \lambda f - \gamma \operatorname{div}(d - b), \tag{20}$$

where the spatial interaction is stronger for larger γ. Therefore, γ should be neither extremely large nor small for good convergence in both subproblems.

The overall algorithm is

Initialize $u = 0$, $d = 0$, $b = 0$
while $\|u_{cur} - u_{prev}\|_2 > Tol$ **do**
    | Solve the d subproblem
    | Solve the u subproblem
    | $b = b + \nabla u - d$

Algorithm 1

The iteration continues until the L^2 difference between the current and previous iterate of u is below a threshold Tol or a maximum iteration limit is reached. The maximum iterations limit should be large enough so that convergence within Tol is usually reached, but not too large to prevent impractical runtimes.

For color images, the vectorial TV (VTV) is used in place of TV,

$$\|u\|_{\text{VTV}(\Omega)} := \int_\Omega \Big(\sum_{i \in \text{channels}} |\nabla u_i(x)|^2 \Big)^{1/2} dx. \tag{21}$$

The grayscale algorithm extends directly to VTV-regularized inpainting.

5 Examples

In the examples and online demo, we fix $\gamma = 5$, $Tol = \|f\|_2/10^5$, and the maximum number of iterations is 250. In the online demo, the inpainting region may be randomly generated or specified manually by clicking on the image. For visualization, the mask color may be set to yellow, blue, or black. The mask color has no effect on the inpainting result.

Online demo interface

5.1 Text removal

TV inpainting (and inpainting in general) is most successful when the inpainting domain is thin. Excellent results are obtained in removing text overprinted on an image ($\lambda = 10^4$).

Input	D	TV Inpainting
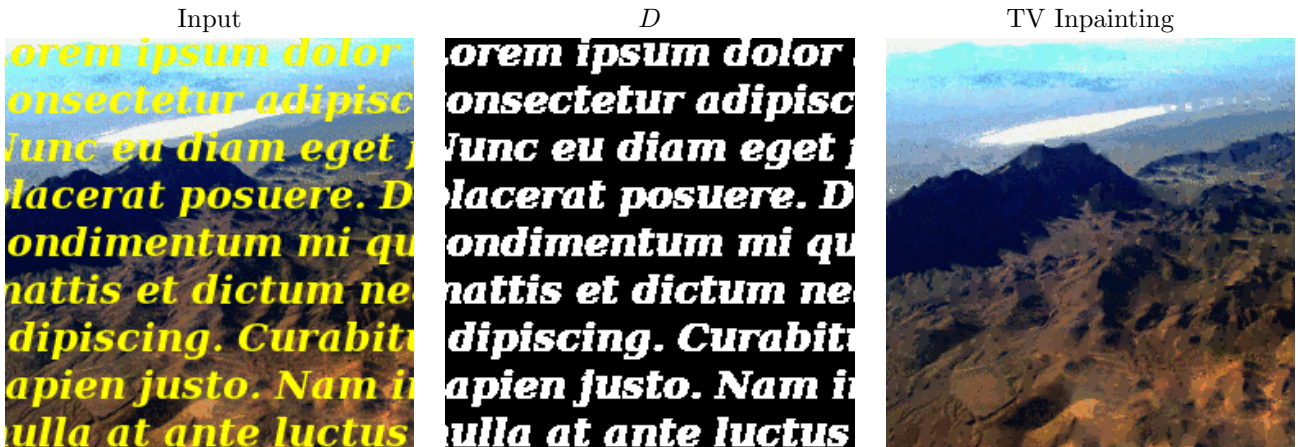
A good feature of TV inpainting is that it reconstructs edges rather than smoothing them. The following figure shows the detail where the edge of the mountain is inpainted:

Input	TV Inpainting

Some other well-suited applications having thin inpainting domains are

- removing wires, e.g., utility lines, fences, cages;

- restoring photographs with scratches or folding creases;

- removing impulsive noise, provided that the set D of corrupt pixels is already determined.
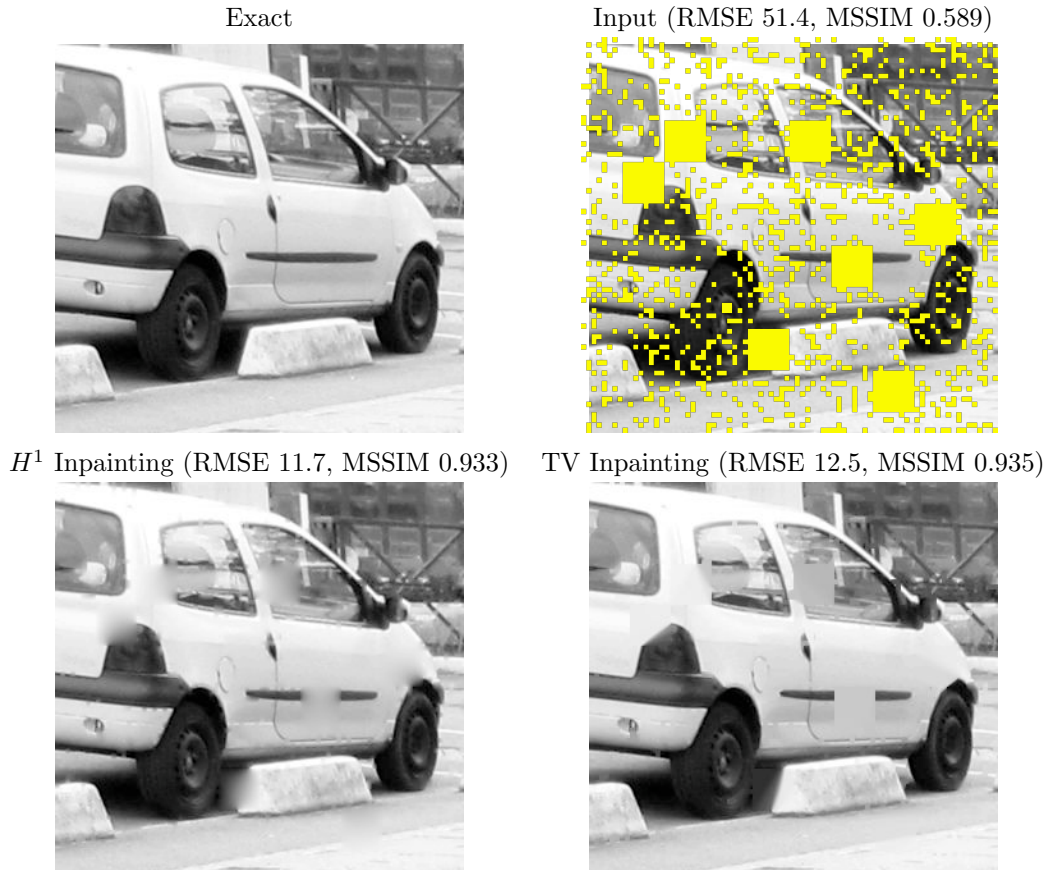
5.2 Effect of λ

Outside of the inpainting region D, the TV inpainting model denoises the image. The denoising strength is controlled by the value of λ, where a smaller value implies stronger denoising. In the figure below, the previous experiment is repeated with three different values of λ.

$\lambda = 10$	$\lambda = 40$	$\lambda = 10^4$

5.3 Comparison with H^1 Inpainting

The following example compares H^1 and TV inpainting where the inpainting domain D consists of 4×4-pixel and 32×32-pixel squares ($\lambda = 10^4$). The inpainting domain covers about 26% of the image.

Exact

Input (RMSE 51.4, MSSIM 0.589)

H^1 Inpainting (RMSE 11.7, MSSIM 0.933)

TV Inpainting (RMSE 12.5, MSSIM 0.935)

Notice that for both methods, the inpainting is usually better in the 4×4 squares than in the larger 32×32 squares. The H^1 inpainting is unable to recover sharp edges: in squares containing an edge, it produces a smooth gradient. TV inpainting can create sharp edges, so it inpaints object boundaries more convincingly.

According to the root mean squared errors (RMSE), H^1 inpainting is a little better. However, it is well known that RMSE may disagree with perception. We also compare using the mean structure similarity (MSSIM) index of Wang et al. [18], which is more sensitive to the quality of edges and yields a maximum value of 1 for perfect agreement. The MSSIM suggests that the TV inpainting result is slightly better than H^1.

Input

H^1 Inpainting
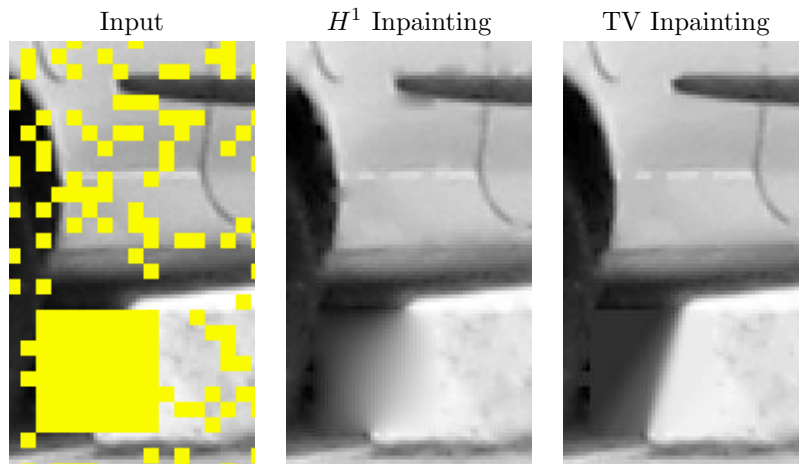
TV Inpainting

TV inpainting prefers straight contours as they have minimal TV, but this is less successful for recovering curved boundaries. The closeups below show the car's tail light. TV inpainting connects the tail light boundary with a straight edge, producing sharp corners at the inpainting boundary. The result is also poor in 32×32 square containing the corner of the window frame.

| Input | H^1 Inpainting | TV Inpainting |
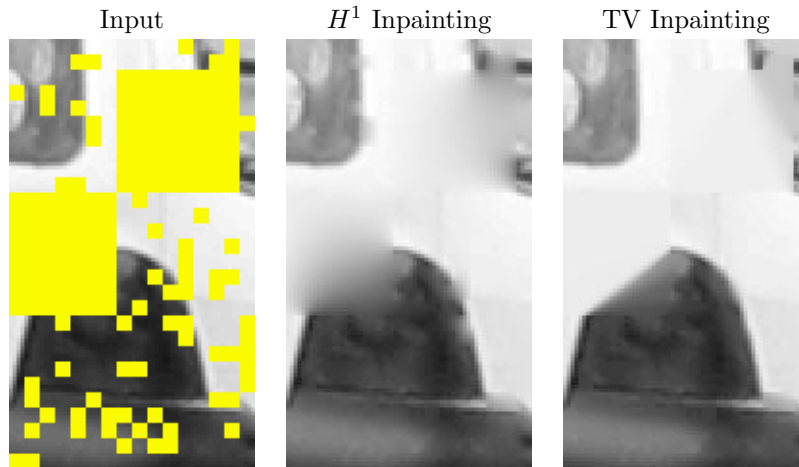
TV inpainting can reconstruct a stripe passing through the inpainting domain, but only if the length to be bridged is less than the stripe thickness. TV inpainting breaks the stripe if the length is greater. TV inpainting always prefers to interpolate the level lines using the shortest possible connection. Chan and Kang [20] investigate this effect theoretically in their error analysis of TV inpainting. Here, this effect is most prominent where a rub strip is broken by a 32×32 square and can also be seen on the bottom edge of the door.

| Input | TV Inpainting |
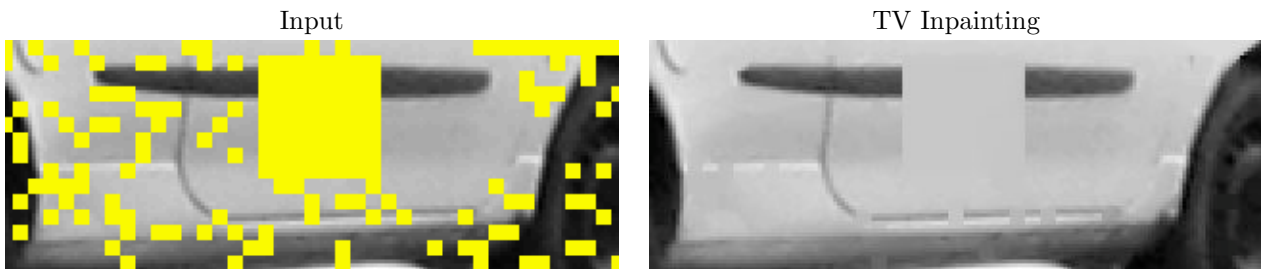
5.4   Object Removal

In this example, TV inpainting is used to attempt to remove a lamppost from an image ($\lambda = 250$).

| Input | D | TV Inpainting |
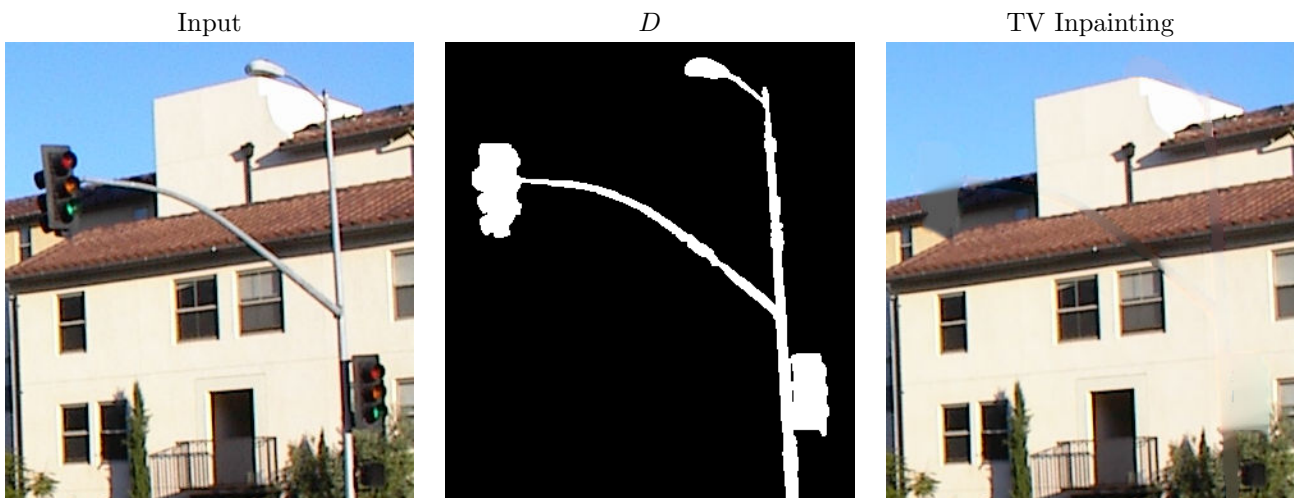
The result is reasonable over the pole where D is thin, but is poor over the signal where it is thicker.

Acknowledgments

Image Credits

References

[1] L.M. Bregman, "The relaxation method of finding the common points of convex sets and its application to the solution of problems in convex optimization," *USSR Computational Mathematics and Mathematical Physics*, vol. 7, no. 3, pp. 200–217, 1967. http://dx.doi.org/10.1016/0041-5553(67)90040-7

[2] M. Nitzberg, D. Mumford, "The 2.1-D sketch," *Proceedings of the Third International Conference on Computer Vision*, pp. 138–144, 1990. http://dx.doi.org/10.1109/ICCV.1990.139511

[3] L.I. Rudin, S. Osher, E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D*, vol. 60, pp. 259–268, 1992. http://dx.doi.org/10.1016/0167-2789(92)90242-F

[4] M. Nitzberg, D. Mumford, T. Shiota, "Filtering, Segmentation, and Depth," *Lecture Notes in Computer Science*, vol. 662, Springer-Verlag, Berlin, 1993.

[5] H. Igehy, L. Pereira, "Image replacement through texture synthesis," *Proceedings of the International Conference on Image Processing*, vol. 3, pp. 186–189, 1997. http://dx.doi.org/10.1109/ICIP.1997.632049

[6] S. Masnou, J.M. Morel, "Level lines based disocclusion," *Proceedings of the International Conference on Image Processing*, pp. 259–263, 1998. http://dx.doi.org/10.1109/ICIP.1998.999016

[7] L. Evans, *Partial Differential Equations*, Graduate studies in mathematics, vol. 19, American Mathematical Society, 1998. ISBN: 0821807722.

[8] A. Efros, T. Leung, "Texture synthesis by non-parametric sampling," *International Conference on Computer Vision*, pp. 1033–1038, 1999. http://dx.doi.org/10.1109/ICCV.1999.790383

[9] M. Bertalmío, G. Sapiro, V. Caselles, C. Ballester, "Image inpainting," *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, pp. 417–424, 2000. http://dx.doi.org/10.1145/344779.344972

[10] J.H. Elder, R.M. Goldberg, "Image editing in the contour domain," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 3, pp. 291–296, 2001. http://dx.doi.org/10.1109/34.910881

[11] M. Bertalmío, A.L. Bertozzi, G. Sapiro, "Navier-stokes, fluid dynamics, and image and video inpainting," *Proceedings of Computer Vision and Pattern Recognition*, vol. 1, pp. 355–362, 2001. http://dx.doi.org/10.1109/CVPR.2001.990497

[12] T.F. Chan, J. Shen, "Nontexture inpainting by curvature-driven diffusions," *Journal of Visual Communication and Image Representation*, vol. 12, no. 4, pp. 436–449, 2001. http://dx.doi.org/10.1006/jvci.2001.0487

[13] T.F. Chan, J. Shen, "Mathematical models of local non-texture inpaintings," *SIAM Journal on Applied Mathematics*, vol. 62, no. 3, pp. 1019–1043, 2001. http://www.jstor.org/stable/3061798

[14] T.F. Chan, S.-H. Kang, J. Shen, "Euler's Elastica and Curvature-Based Inpainting," *SIAM Journal on Applied Mathematics*, vol. 63, no. 2, pp. 564–592, 2002. http://dx.doi.org/10.1137/S0036139901390088

[15] E. Meijering, "A Chronology of Interpolation: From Ancient Astronomy to Modern Signal and Image Processing," *Proceedings of the IEEE*, 90, pp. 319–342, 2002. http://dx.doi.org/10.1109/5.993400

[16] M. Bertalmío, L. Vese, G. Sapiro, S. Osher, "Simultaneous structure and texture image inpainting," *IEEE Transactions on Image Processing*, vol. 12, no. 8, pp. 882–889, 2003. http://dx.doi.org/10.1109/TIP.2003.815261

[17] A. Criminisi, P. Pérez, K. Toyama, "Region filling and object removal by exemplar-based image inpainting," *IEEE Transactions on Image Processing*, vol. 13, no. 9, pp. 1200–1212, 2004. http://dx.doi.org/10.1109/TIP.2004.833105

[18] Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004. http://dx.doi.org/10.1109/TIP.2003.819861

[19] M. Elad, J.L. Starck, P. Querre, D.L. Donoho, "Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA)," *Applied and Computational Harmonic Analysis*, vol. 19, no. 3, pp. 340–358, 2005. http://dx.doi.org/10.1016/j.acha.2005.03.005

[20] T.F. Chan, S.H. Kang, "Error Analysis for Image Inpainting," *Journal of Mathematical Imaging and Vision*, vol. 26, no. 1–2, pp. 85–103, 2006. http://dx.doi.org/10.1007/s10851-006-6865-7

[21] T. Goldstein, S. Osher, "The Split Bregman Method for L1-Regularized Problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 2, pp. 323–343, 2009. http://dx.doi.org/10.1137/080725891

[22] P. Getreuer, "tvreg v2: Variational Imaging Methods for Denoising, Deconvolution, Inpainting, and Segmentation," 2010. http://www.mathworks.com/matlabcentral/fileexchange/29743-tvreg

[23] P. Getreuer, "Rudin-Osher-Fatemi Total Variation Denoising using Split Bregman," *Image Processing On Line*, 2012. http://dx.doi.org/10.5201/ipol.2012.g-tvd