Semi-local Total Variation for Regularization of Inverse Problems

Laurent Condat

Proc. of the 22nd European Signal Processing Conf. (EUSIPCO), Lisbon, Portugal, Sept. 2014.

SEMI-LOCAL TOTAL VARIATION FOR REGULARIZATION OF INVERSE PROBLEMS

Laurent Condat

GIPSA-lab, Dept. Images & Signals, University of Grenoble–Alpes, Grenoble, France

ABSTRACT

We propose the discrete semi-local total variation (SLTV) as a new regularization functional for inverse problems in imaging. The SLTV favors piecewise linear images; so the main drawback of the total variation (TV), its clustering effect, is avoided. Recently proposed primal-dual methods allow to solve the corresponding optimization problems as easily and efficiently as with the classical TV.

Index Terms— total variation, non-local regularization, inverse problem, convex optimization, proximal method

1. INTRODUCTION

Many inverse problems in imaging can be regularized and put under the form of convex optimization problems: given the data y and the linear observation operator A, one aims at solving problems like

$$\hat{x} = \arg\min_x \ \mathcal{J}(x) \quad \text{s.t.} \quad Ax = y, \quad \text{or} \tag{1}$$

$$\hat{x} = \arg\min_x \ \tfrac{\lambda}{2}\|Ax - y\|^2 + \mathcal{J}(x), \tag{2}$$

where $\mathcal{J}$ is a convex regularization functional and $\lambda > 0$ is the regularization parameter. The classical Tikhonov regularizer $\mathcal{J}(x) = \|\nabla x\|_2^2$ generally makes the problem easy to solve, but yields over-smoothing of the textures and edges in the recovered image $\hat{x}$. A popular and better regularizer is total variation (TV), see e.g. [1]:

$$\mathcal{J}_{\mathrm{TV}}(x) = \|\nabla x\|_{1,2} = \textstyle\sum_{\mathbf{k}\in\mathbb{Z}^2} \|\nabla x[\mathbf{k}]\|_2, \tag{3}$$

where $\|\mathbf{a}\|_2 = \sqrt{|a_1|^2 + |a_2|^2}$,

$$\nabla x[\mathbf{k}] = \big[x[k_1+1,k_2] - x[\mathbf{k}], x[k_1,k_2+1] - x[\mathbf{k}]\big]^{\mathrm{T}} \tag{4}$$

and $x[\mathbf{k}]$ is the pixel value of x at location $\mathbf{k} = [k_1, k_2]^{\mathrm{T}}$. TV regularization yields images with sharp edges but the textures are still over-smoothed, there are staircasing effects and the pixel values in smooth regions are clustered in piecewise constant areas, which gives a synthetic look to the reconstructed images.

To overcome the drawbacks of regularizers based on the local interactions of adjacent pixels solely, non-local methods have become increasingly popular. The non-local functional of Gilboa *et al.* [2, 3] can be expressed as

$$\mathcal{J}_{\mathrm{NL}}(x) = \sum_{\mathbf{k}\in\mathbb{Z}^2} \sum_{\mathbf{l}\in\mathbb{Z}^2} \phi(|x[\mathbf{k}] - x[\mathbf{l}]|) w(\mathbf{k},\mathbf{l}) \tag{5}$$

for a positive convex function ϕ. The nonnegative and symmetric weight function $w(\mathbf{k},\mathbf{l})$ accounts for the similarity between the image features at locations $\mathbf{k}$ and $\mathbf{l}$. Its choice is critical. It can be obtained based on patch distances in a first estimate of the solution obtained by solving the problem with Tikhonov or TV regularization; or it can be defined implicitly from the geometry of the image and updated iteratively at the same time as the solution [4].

In this work, we propose instead a regularizer based on gradient differences instead of pixel values differences and without the introduction of the weight function, which is difficult to determine. The functional is as follows:

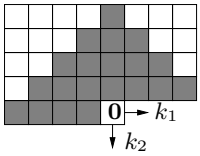
$$\mathcal{J}_{\mathrm{SLTV}}(x) = \sum_{\mathbf{k}\in\mathbb{Z}^2} \sum_{\mathbf{l}\in\mathbb{Z}^2 \,|\, \mathbf{l}-\mathbf{k}\in\Omega} \|\nabla x[\mathbf{k}] - \nabla x[\mathbf{l}]\|_2, \tag{6}$$

for some set of pixels $\Omega \subset \mathbb{Z}^2$. $\mathcal{J}_{\mathrm{SLTV}}$ is semi-local, since the gradient is compared to other gradients in its neighborhood. This functional was proposed and studied in the continuous domain by Kindermann *et al.* [5], as a semi-local extension of the total variation. Since $\mathcal{J}_{\mathrm{SLTV}}(x) = 0$ if x represents an affine image, it is expected that the minimization of $\mathcal{J}_{\mathrm{SLTV}}$ favors piecewise affine solutions over piecewise constant ones, avoiding staircasing. This has been confirmed by experiments in [5].

Note that that if the set Ω is symmetric, (6) can be rewritten as

$$\mathcal{J}_{\mathrm{SLTV}}(x) = 2 \sum_{\mathbf{k}\in\mathbb{Z}^2} \sum_{\mathbf{l}\in\mathbb{Z}^2 \,|\, \mathbf{l}-\mathbf{k}\in\Omega, \mathbf{l}<\mathbf{k}} \|\nabla x[\mathbf{k}] - \nabla x[\mathbf{l}]\|_2, \tag{7}$$

where $\mathbf{l} < \mathbf{k}$ is understood in the lexicographic order. Therefore, we can choose Ω as one half of a symmetric set. This is advantageous because the size of the memory buffers in the algorithms is proportional to the size of Ω. The following set is used:



Indeed, we found out empirically that a smaller neighborhood is not able to efficiently capture the local correlations, while there is virtually no difference when further increasing the size of Ω.

In [5], the problem (2) was solved approximately by smoothing, i.e. replacing $\|\mathbf{a}\|_2$ by $\sqrt{|a_1|^2+|a_2|^2+\varepsilon}$ for a small $\varepsilon > 0$, and using a Euler method for the steepest descent flow, a particularly slow approach. Recent advances in optimization theory have made the computational solution of the problems (1) and (2), with convex non-smooth regularizers like the TV, easy and fast [6–10]. In the next section, we detail the implementation of an efficient optimization method to regularize problems with SLTV. We demonstrate the improvement of SLTV over TV by experiments in Sect. 3. We stress that this work does not aim at giving state-of-the-art results in inverse imaging problems. Instead, we modestly show how regularization by SLTV can be harnessed easily to a variety of applications, potentially yielding better results than the popular TV.

2. A PRIMAL-DUAL ALGORITHM FOR SLTV MINIMIZATION

Numerous problems in engineering can be formulated as the minimization of a sum of convex functions, not necessarily differentiable, possibly composed with linear operators. *Proximal splitting methods* solve the problems iteratively by calling either the gradient or the proximity operator of each function [11]. The proximity operator of a convex function g is defined by

$$\operatorname{prox}_g(x) = \underset{x'}{\arg\min}\ g(x') + \frac{1}{2}\|x - x'\|^2. \tag{8}$$

The classical splitting methods, like the forward-backward or Douglas–Rachford methods [11], cannot be used to solve the problems (1) and (2), since they would require evaluations of $\operatorname{prox}_{\mathcal{J}}$, for which there is no closed form. Recent advances in the field have enabled to solve the generic problem

$$\text{Find } \hat{x} \in \underset{x\in\mathcal{X}}{\arg\min}\ f(x) + g(x) + h(Lx), \tag{9}$$

where f, g, h are convex functions, with f supposed differentiable, $L : \mathcal{X} \to \mathcal{U}$ is a linear operator, $\mathcal{X}$ and $\mathcal{U}$ are real Hilbert spaces. The primal-dual algorithms in [8, 9] allow to solve this problem, using at every iteration calls to ∇f,prox_g, prox_h, L and its adjoint L^*. In this paper, we present the algorithm of [9, 12], which is as follows:

Splitting algorithm to solve (9)

Choose the parameters $\tau > 0$, $\sigma > 0$, $\rho > 0$ and the initial estimates $x^{(0)} \in \mathcal{X}$, $u^{(0)} \in \mathcal{U}$. Then iterate, for $i = 0, 1, \ldots$

$$\begin{aligned}
\tilde{x}^{(i+1)} &:= \operatorname{prox}_{\tau g}\big(x^{(i)} - \tau\nabla f(x^{(i)}) - \tau L^* u^{(i)}\big),\\
x^{(i+1)} &:= \rho\,\tilde{x}^{(i+1)} + (1-\rho)x^{(i)},\\
\tilde{u}^{(i+1)} &:= \operatorname{prox}_{\sigma h^*}\big(u^{(i)} + \sigma L(2\tilde{x}^{(i+1)} - x^{(i)})\big),\\
u^{(i+1)} &:= \rho\,\tilde{u}^{(i+1)} + (1-\rho)u^{(i)}.
\end{aligned}$$

In the algorithm, h^* is the convex conjugate of h, about which it is generally sufficient to know that $\operatorname{prox}_{\sigma h^*}(u) = u - \sigma\operatorname{prox}_{h/\sigma}(u/\sigma)$.

There are several ways to recast the problems (1) and (2) as particular cases of (9); in all cases, the regularizer $\mathcal{J}_{\mathrm{SLTV}}$ is assigned to the term $h \circ L$:

- The problem (1) corresponds to $f = 0$ and $g(x) = \{0 \text{ if } Ax = y,\ +\infty \text{ else}\}$. We have $\operatorname{prox}_{\tau g}(x) = x + A^\dagger(y - Ax)$, where $A^\dagger$ is the Moore-Penrose pseudo-inverse of A. If AA^* is invertible, then $A^\dagger = A^*(AA^*)^{-1}$. Convergence of the algorithm is guaranteed if $\tau\sigma\|L\|^2 \leq 1$ and $\rho < 2$.
- For the problem (2), one can set $f(x) = \frac{\lambda}{2}\|Ax - y\|^2$ and $g = 0$. Then, $\nabla f(x) = \lambda A^*(Ax - y)$. Convergence of the algorithm is guaranteed if $\tau(\lambda\|A\|^2/2 + \sigma\|L\|^2) < 1$ and $\rho = 1$.
- For the problem (2), another choice (adopted for the experiments in Sect. 3) is to set $f = 0$ and $g(x) = \frac{\lambda}{2}\|Ax - y\|^2$. Then, $\operatorname{prox}_{\tau g}(x) = (\mathrm{Id} + \lambda\tau A^*A)^{-1}(x + \lambda\tau A^*y)$. Note that if $AA^* = \mu\mathrm{Id}$ for some $\mu > 0$, this simplifies to $\operatorname{prox}_{\tau g}(x) = x + \frac{\lambda\tau}{1+\lambda\tau\mu}A^*(y - Ax)$. Convergence of the algorithm is guaranteed if $\tau\sigma\|L\|^2 \leq 1$ and $\rho < 2$.

Let us now define the operator L corresponding to SLTV. Let N be the number of pixels in Ω (20 with our choice) and $\mathbf{m}_1, \ldots, \mathbf{m}_N \in \mathbb{Z}^2$ be the elements of Ω. Then, $Lx = \mathbf{u}$ with

$$\mathbf{u}[\mathbf{k}]_n = \nabla x[\mathbf{k}] - \nabla x[\mathbf{k} + \mathbf{m}_n] \in \mathbb{R}^2, \tag{10}$$

for every $\mathbf{k} \in \mathbb{Z}^2$ and $n = 1, \ldots, N$. We also have $L^*\mathbf{u} = x$ with, for every $\mathbf{k} \in \mathbb{Z}^2$,

$$\begin{aligned}
x[\mathbf{k}] = \sum_{n=1}^{N} & u[k_1 - 1, k_2]_{n,1} - u[\mathbf{k}]_{n,1} - \\
& u[k_1 - m_{n,1} - 1, k_2 - m_{n,2}]_{n,1} + \\
& u[k_1 - m_{n,1}, k_2 - m_{n,2}]_{n,1} + \\
& u[k_1, k_2 - 1]_{n,2} - u[\mathbf{k}]_{n,2} - \\
& u[k_1 - m_{n,1}, k_2 - m_{n,2} - 1]_{n,2} + \\
& u[k_1 - m_{n,1}, k_2 - m_{n,2}]_{n,2}.
\end{aligned} \tag{11}$$

We have $\|L\|^2 = \|L^*L\|$ and L^*L is a linear shift-invariant operator on images; that is, it corresponds to a convolution: $L^*L(v) = v * p$ for some filter p. Hence, $\|L^*L\| = \sup_{\boldsymbol{\omega}\in[-\pi,\pi]^2}\hat{p}(\boldsymbol{\omega})$, where $\hat{p}(\boldsymbol{\omega}) = \sum_{\mathbf{k}\in\mathbb{Z}^2} p[\mathbf{k}]e^{-j\boldsymbol{\omega}^{\mathrm{T}}\mathbf{k}}$ is the Fourier transform of p. For the set Ω depicted in the Introduction, we have $\|L\|^2 \approx 325.63$.

Now, the function h so that $\mathcal{J}_{\mathrm{SLTV}} = h \circ L$ is

$$h(\mathbf{u}) = \sum_{\mathbf{k}\in\mathbb{Z}^2}\sum_{n=1}^{N}\|\mathbf{u}[\mathbf{k}]_n\|_2. \tag{12}$$

The convex conjugate of h is $h^* : \mathbf{u} \mapsto \{0 \text{ if } \max_{\mathbf{k}\in\mathbb{Z}^2}\max_{n=1,\ldots,N}\|\mathbf{u}[\mathbf{k}]_n\|_2 \leq 1,\ +\infty \text{ else}\}$.

Proof: with this definition of h^*, we have

$$(h^*)^*(\mathbf{u}) = \sup_{\mathbf{v}} \langle \mathbf{u}, \mathbf{v} \rangle - h^*(\mathbf{v}) \quad (13)$$

$$= \sup_{0<\rho\leq 1} \sup_{\mathbf{v} \mid \max_{\mathbf{k}\in\mathbb{Z}^2} \max_{n=1,\ldots,N} \|\mathbf{v}[\mathbf{k}]_n\|_2 = \rho} \langle \mathbf{u}, \mathbf{v} \rangle \quad (14)$$

$$= \sup_{0<\rho\leq 1} \rho \sum_{\mathbf{k}\in\mathbb{Z}^2} \sum_{n=1}^{N} \|\mathbf{u}[\mathbf{k}]_n\|_2 \quad (15)$$

$$= \sum_{\mathbf{k}\in\mathbb{Z}^2} \sum_{n=1}^{N} \|\mathbf{u}[\mathbf{k}]_n\|_2 = h(\mathbf{u}). \quad \square \quad (16)$$

Hence, $\mathrm{prox}_{\sigma h^*}$ is the orthogonal projection which maps $\mathbf{u}$ to $\mathbf{v}$ with, for every $\mathbf{k} \in \mathbb{Z}^2$, $n = 1, \ldots, N$,

$$\mathbf{v}[\mathbf{k}]_n = \frac{\mathbf{u}[\mathbf{k}]_n}{\max(\|\mathbf{u}[\mathbf{k}]_n\|_2, 1)}. \quad (17)$$

In the experiments of the next section, with pixel values in the range $[0, 255]$, we chose $\tau = 0.1$, $\sigma = 1/\|L\|^2/\tau$, $\rho = 1$.

3. EXPERIMENTAL EXAMPLES

3.1. Denoising

We first consider the denoising problem, where the observation operator is the identity: $A = \mathrm{Id}$. In Fig. 2, four parts of popular test images are shown, corrupted by additive white Gaussian noise (AWGN) of standard deviation $\sigma = 20$. The second row of Fig. 2 shows the denoised images using total variation ($\mathcal{J} = \mathcal{J}_{\mathrm{TV}}$), where the value of λ has been tuned manually to maximize the PSNR, for each image. This optimal value of λ yields images where noise is still visible, while some image details have disappeared, see e.g. the stripes of the pants in image (f), and the pixel values tend to be clustered into piecewise constant regions. The third row of Fig. 2 shows the results with semi-local total variation ($\mathcal{J} = \mathcal{J}_{\mathrm{SLTV}}$). As is visible in the images (i)–(l), the strong edges are sharp, but the clustering and staircasing effects proper to total variation have disappeared. Hence, the tradeoff between noise removal and details preservation is better with SLTV than with TV.

3.2. Demosaicing

Another classical interpolation problem in imaging is demosaicing, which consists in reconstructing a color image $\mathbf{x} = [x_R, x_G, x_B]^{\mathrm{T}}$ with red (R), green (G), blue (B) channels, knowing only one of these three values at each pixel location [13]. That is, $A\mathbf{x} = y$ with $y[\mathbf{k}] = \{x_G[\mathbf{k}]$ if $k_1 + k_2$ is even, $x_R[\mathbf{k}]$ else if k_2 is even, $x_B[\mathbf{k}]$ else$\}$, $\forall \mathbf{k} \in \mathbb{Z}^2$. Note that $AA^* = \mathrm{Id}$. In [14], the author proposed an extension of the total variation to color images as follows:

$$\mathcal{J}_{\mathrm{TV}}(\mathbf{x}) = \mu\, \mathcal{J}_{\mathrm{TV}}(x_L) + \mathcal{J}_{\mathrm{TV}}(x_C), \quad (18)$$

where $x_C = x_{G/M} + j.x_{R/B}$ is the complex chrominance field and $x_L, x_{G/M}, x_{R/B}$ are the channels of $\mathbf{x}$ expressed in the luminance, green-magenta and red-blue chrominance orthonormal basis [14]. The important parameter $\mu < 1$ in (18) ensures that the reconstructed image has its chrominance channels smoother than its luminance channel, a known property of natural images. It is straightforward to extend the definition of the SLTV to color images, the same way as TV is extended in (18). For the computations in the algorithms, one can switch between the R,G,B and luminance,chrominance bases, since this operation is unitary.

Table 1. PSNR (in dB) for the demosaicing experiments over the 24 images of the classical Kodak test set.

image	TV	SLTV	13	35.75	**36.36**
1	39.20	**39.72**	14	34.95	**36.68**
2	38.24	**39.92**	15	37.99	**39.33**
3	41.32	**42.52**	16	42.31	**42.67**
4	39.19	**40.52**	17	40.08	**41.27**
5	35.95	**37.61**	18	35.88	**37.33**
6	38.72	**39.44**	19	38.74	**39.38**
7	39.78	**42.07**	20	40.09	**41.27**
8	35.00	**35.67**	21	38.82	**39.73**
9	40.80	**42.06**	22	36.85	**38.09**
10	40.41	**41.84**	23	40.80	**42.62**
11	38.00	**39.55**	24	33.26	**34.93**
12	42.38	**43.36**	mean	38.52	**39.75**

The results of solving (1) with $\mathcal{J} = \mathcal{J}_{\mathrm{TV}}$ and $\mathcal{J}_{\mathrm{SLTV}}$ are reported in Tab. 1. We used $\mu = 0.625$. The large average improvement of 1.2dB obtained with the SLTV over the TV shows that the SLTV is a better regularization for the demosaicing problem.

We also considered the joint demosaicing-denoising problem, in which the mosaicked image is corrupted by AWGN, with std. dev. 20. The image is reconstructed by solving (2). One result is illustrated in Fig. 1. The visual quality of the images reconstructed with the TV and the SLTV is comparable, but the latter is free from the piecewise constant clustering effect of the TV. Moreover, the SLTV tends to give images with more accurate colors, while with the TV, the colors are desaturated, especially on small objects. This can be seen in Fig. 1 on the blue rudder in the man's hands, which appears more blue in (c) than in (b).

4. CONCLUSION

We proposed the semi-local total variation (SLTV), as an alternative to the total variation (TV) for regularization of inverse problems in imaging. We have shown that with recent primal-dual splitting methods, there is no difficulty in adapting an algorithm from the TV to the SLTV. The computational cost with SLTV is higher, relatively to the size of the neighborhood set Ω, but SLTV yields more pleasant images, where the sharpness of edges is maintained without the typical

(a) Original Image

(b) reconstructed with TV, 26.83dB $\mu = 0.42$, $\lambda = 0.06$

(c) reconstructed with SLTV, 27.03dB $\mu = 0.32$, $\lambda = 1.3$.

Fig. 1. Joint demosaicing-denoising experiment on image 14 of the Kodak test base. The PSNR values correspond to the whole image, not to the crop selected here. The parameters μ and λ were empirically optimized to maximize the PSNR. 100 iterations of the algorithm were run.

clustering effect of TV.

REFERENCES

[1] A. Chambolle, V. Caselles, D. Cremers, M. Novaga, and T. Pock, "An introduction to total variation for image analysis," in *Theoretical Foundations and Numerical Methods for Sparse Recovery*. De Gruyter, Radon Series Comp. Appl. Math., 2010, vol. 9, pp. 263–340.

[2] G. Gilboa, J. Darbon, S. Osher, and T. F. Chan, "Nonlocal convex functionals for image regularization," UCLA CAM, Tech. Rep. Report 06-57, Oct. 2006.

[3] G. Gilboa and S. Osher, "Nonlocal operators with applications to image processing," *Multiscale Model. Simul.*, vol. 7, no. 3, pp. 1005–1028, 2008.

[4] G. Peyré, S. Bougleux, and L. Cohen, "Non-local regularization of inverse problems," in *Proc. of ECCV*, 2008, pp. 57–68.

[5] S. Kindermann, S. Osher, and P. W. Jones, "Deblurring and denoising of images by nonlocal functionals," *Multiscale Model. Simul.*, vol. 4, no. 4, pp. 1091–1115, 2006.

[6] A. Chambolle and T. Pock, "A first-order primal–dual algorithm for convex problems with applications to imaging," *Journal of Mathematical Imaging and Vision*, vol. 40, no. 1, pp. 120–145, 2011.

[7] I. Loris and C. Verhoeven, "On a generalization of the iterative soft-thresholding algorithm for the case of non-separable penalty," *Inverse Problems*, vol. 27, no. 12, 2011.

[8] P. L. Combettes and J.-C. Pesquet, "Primal–dual splitting algorithm for solving inclusions with mixtures of composite, Lipschitzian, and parallel-sum type monotone operators," *Set-Valued and Variational Analysis*, vol. 20, no. 2, 2012.

[9] L. Condat, "A primal-dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms," *J. Optimization Theory and Applications*, vol. 158, no. 2, pp. 460–479, 2013.

[10] P. Chen, J. Huang, and X. Zhang, "A primal-dual fixed point algorithm for convex separable minimization with applications to image restoration," *Inverse Problems*, vol. 29, no. 2, 2013.

[11] P. L. Combettes and J.-C. Pesquet, "Proximal splitting methods in signal processing," in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, H. H. Bauschke, R. Burachik, P. L. Combettes, V. Elser, D. R. Luke, and H. Wolkowicz, Eds. New York: Springer-Verlag, 2010.

[12] L. Condat, "A generic proximal algorithm for convex optimization — application to total variation minimization," *IEEE Signal Processing Lett.*, vol. 21, no. 8, pp. 1054–1057, Aug. 2014.

[13] B. K. Gunturk, J. Glotzbach, Y. Altunbasak, R. W. Schaffer, and R. M. Mersereau, "Demosaicking: Color filter array interpolation," *IEEE Signal Processing Mag.*, vol. 22, no. 1, pp. 44–54, Jan. 2005.

[14] L. Condat and S. Mosaddegh, "Joint demosaicking and denoising by total variation minimization," in *Proc. of IEEE ICIP*, Orlando, USA, Sept. 2012.

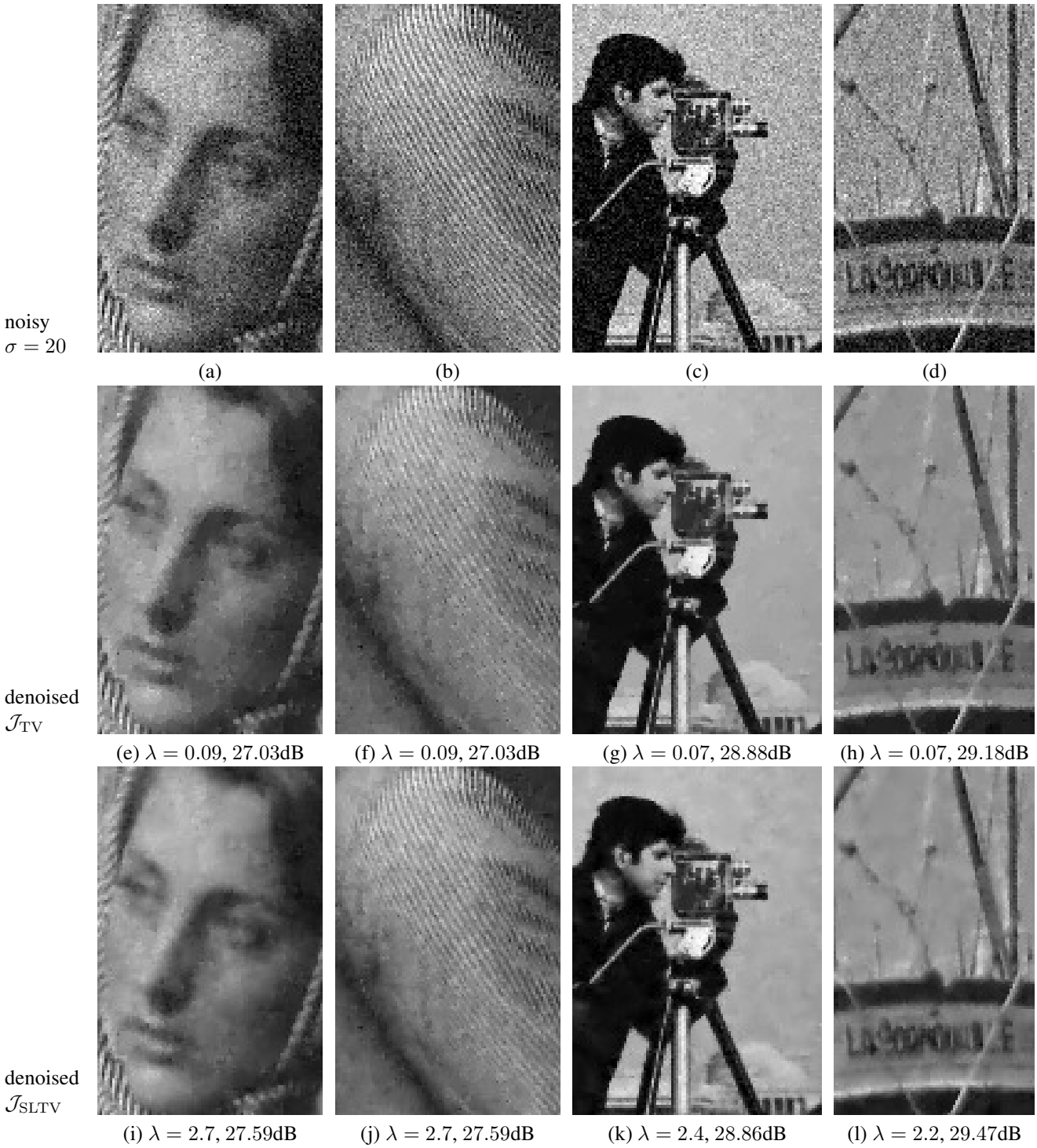


Fig. 2. Denoising experiments using regularization with total variation and proposed semi-local total variation. For each denoised image, 100 iterations of the proposed algorithm were run. The PSNR values correspond to the whole Barbara, Camera and Boat, images, not to the crops selected here.