Image Denoising

Ted (Yining) Ding
yd2007@hw.ac.uk

July 2021 (last updated June 2022)

This document is a supplement to the following GitHub repository on image denoising: `https://github.com/tedyiningding/Image-Denoising`.

Grayscale image denoising

The following two methods of denoising a grayscale image will be investigated: TV denoising and TGV denoising.

TV denoising

Consider a grayscale image of size $M \times N$, Total Variation (TV) denoising (more specifically, we mean the ROF model [1]) finds the unique solution $x^\star \in \mathbb{R}^{M\times N}$ which minimises

$$\frac{1}{2}\left\|x-y\right\|^2 + \lambda \left\|x\right\|_{\mathrm{TV}}$$

where $y \in \mathbb{R}^{M\times N}$ is the observed image, λ is a regularisation parameter that balances the two terms, and $\|\cdot\|_{\mathrm{TV}}$ is the total variation of an image which is defined by

$$\left\|x\right\|_{\mathrm{TV}} = \left\|\mathrm{D}x\right\|_{p,1} = \sum_{i=1,j=1}^{M,N} \left|(\mathrm{D}x)_{i,j}\right|_p = \sum_{i=1,j=1}^{M,N} \left((\mathrm{D}x)_{i,j,1}^p + (\mathrm{D}x)_{i,j,2}^p\right)^{1/p}$$

where $\mathrm{D} : \mathbb{R}^{M\times N} \to \mathbb{R}^{M\times N\times 2}$ is the discrete gradient operator. That is, $\|\cdot\|_{\mathrm{TV}}$ is the ℓ_1-norm of the p-norm of the pixelwise image gradients [2, pp. 168]. When $p = 1$, it is called the anisotropic TV; whereas when $p = 2$, it is called the isotropic TV. The code uses the latter.

The code solves the aforementioned problem using the over-relaxed Chambolle-Pock algorithm [3, Algorithm 3.1] after obtaining a saddle-point problem [2, Example 5.6].

TGV denoising

TV regularisation only promotes piecewise constant structures therefore the result could suffer from staircasing artefacts (as will be seen from the denoised images shown below). To combat this, Total Generalised Variation (TGV) was proposed in [4]. The second order TGV promotes not only piecewise constant structures, but also piecewise affine structures. Again, consider a grayscale image of size $M \times N$, TGV denoising finds the unique solution $u^\star \in \mathbb{R}^{M\times N}$ (and the unique optimiser of an auxiliary variable $\mathbf{v}^\star = (v_1^\star, v_2^\star) \in \mathbb{R}^{M\times N\times 2}$ which minimises

$$\frac{1}{2}\left\|u-y\right\|^2 + \lambda_0 \left\|\mathrm{J}\mathbf{v}\right\|_{2,1} + \lambda_1 \left\|\mathrm{D}u - \mathbf{v}\right\|_{2,1}$$

where the discrete gradient operator D is the same as in the TV denoising example, the operator $\mathrm{J} : \mathbb{R}^{M\times N\times 2} \to \mathbb{R}^{M\times N\times 4}$ can be decomposed into $\mathrm{J}\mathbf{v} = (\mathrm{D}v_1, \mathrm{D}v_2)$ [2, Sec. 7.2.]

As can be seen from the equation above, piecewise affine parts in an image will lead to small values to the last two terms, whereas piecewise constant parts do not.

Similar to the TV Denoising code, the code solves the problem using the over-relaxed Chambolle-Pock algorithm [3, Algorithm 3.1] after obtaining a saddle-point problem [2, Sec. 7.2.].

Results comparison

Qualitative results are shown in Figure 1. We also show some quantitative results including the reconstructed signal-to-noise ratio (RSNR) and the structural similarity index measure (SSIM) in Table 1.

(a) The original image (b) The noisy image (c) TV denoised image (d) TGV denoised image

Figure 1: Gray image denoising comparison

Images	**RSNR (dB)**	**SSIM**
Noisy	14.1727	0.1949
TV denoised	25.3879	0.8702
TGV denoised	25.3985	0.8859

Table 1: Gray image error metrics

We can see that the TGV method is only marginally higher in RSNR and SSIM but there is no obvious staircasing artefact (see the background).

Colour image denoising

Pixels in a colour image are vector-valued and therefore colour images can be considered as a concatenation of multiple scalar-valued images (i.e. multiple colour channels) in the third dimension. A naive way of denoising a colour image would be to apply techniques for scalar-valued image denoising (e.g. TV denoising) to each colour channel individually, ignoring any coupling between the colour channels. We will investigate some slightly more sophisticated methods.

Now we consider an RGB image $\mathbf{x} = (x_r, x_g, x_b) \in \mathbb{R}^{M\times N\times 3}$, where $x_r, x_g, x_b \in \mathbb{R}^{M\times N}$ represent the three colour channels. Consider a new discrete gradient operator $\mathbf{D} : \mathbb{R}^{M\times N\times 3} \to \mathbb{R}^{M\times N\times 2\times 3}$ such that $\mathbf{Dx} = (\mathrm{D}x_r, \mathrm{D}x_g, \mathrm{D}x_b)$, where D is the discrete gradient operator for scalar-valued images which we have defined previously in TV denoising. Without loss of generality, we can switch the last two dimensions of $\mathbf{Dx}$ so now after applying $\mathbf{D}$, each pixel in $\mathbf{x}$ corresponds to a matrix of size 3×2.

The Schatten p-norm of a matrix $A \in \mathbb{R}^{P\times Q}$ is defined by the ℓ_p-norm of the vector consisting of the singular values of A:

$$|A|_{\mathcal{S}_p} \coloneqq \left(\sum_{n=1}^{\mathbf{min}\{P,Q\}} \sigma_n^p(A) \right)^{1/p}$$

where $\sigma_n(A)$ is the nth singular value of A, or equivalently the square root of the nth eigenvalue of $A^\top A$ or $AA^\top$. Three special cases of the Schatten p-norm are listed below.

- $p = 1$: the nuclear norm of matrix (i.e. the sum of its singular values)
- $p = 2$: the Frobenius norm of matrix (i.e. the square root of the sum of squares of all of its entries)
- $p = \infty$: the operator or spectral norm of matrix (i.e. its largest singular value)

Furthermore, calculating the prox of any Schatten p-norm of a matrix A is quite simple:

1. Apply SVD to the matrix: $A = U\Sigma V^\top$
2. Extracting the diagonal entries of Σ to make a vector $\sigma(A)$
3. Calculate the prox of the ℓ_p-norm of $\sigma(A)$ and denote the result by $\hat{\sigma}(A)$. For example, when $p = 1$, it is the well-known soft-thresholding operator.
4. The prox of the Schatten p-norm of a matrix A is just $U\ \mathrm{diag}(\hat{\sigma}(A))\ V^\top$

Nuclear norm denoising

From the explanations above we can already see that the nuclear norm denoising (i.e. $p = 1$) forces every matrix corresponding to a pixel in a vector-valued image to be of low rank (i.e. suppress small non-zero singular values to zeros and shrinks big non-zeros ones). More concretely, it finds the unique solution $\mathbf{x}^\star \in \mathbb{R}^{M\times N\times 3}$ which minimises

$$\frac{1}{2}\|\mathbf{x}-\mathbf{y}\|^2 + \lambda\|\mathbf{D}\mathbf{x}\|_{\mathcal{S}_1,1}$$

Again, the code solves the problem using the over-relaxed Chambolle-Pock algorithm [3, Algorithm 3.1] after obtaining a saddle-point problem [2, Sec. 7.3.].

Because the convex conjugate of a norm is the indicator function of a ball based on the dual norm [5, Sec. 6.5], one will need to calculate the projection of A onto the Schatten ∞-norm ball with radius λ when solving the saddle-point problem, which requires an SVD of A associated with each pixel, or equivalently an eigenvalue decomposition of $A^\top A$ associated with each pixel. Thankfully, when $A \in \mathbb{R}^{3\times 2}$, its SVD has a closed-form solution, allowing not only efficient calculations of the projection, but also efficient evaluations of the nuclear norm of A. This can be achieved by first parameterising the orthogonal matrix $V \in \mathbb{R}^{2\times 2}$ by an angle θ:

$$\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

Please refer to this tutorial for steps of solving θ and the two singular values (σ_1 and σ_2) of A. Next, we show how to efficiently solve the projection of A onto the Schatten ∞-norm ball with radius λ.

It is easy to see that the columns of AV have norms σ_1 and σ_2 respectively:

$$AV = U\Sigma = \begin{bmatrix} | & | & | \\ u_1 & u_2 & u_3 \\ | & | & | \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} | & | \\ \sigma_1 u_1 & \sigma_2 u_2 \\ | & | \end{bmatrix}$$

Furthermore, because both U and V are orthogonal matrices, they preserve matrix norms. Therefore, we can perform the projection on the norm of the two columns. When $p = \infty$, it amounts to the following element-wise operation:

$$\hat{\sigma} = \min\{\sigma, \lambda\}$$

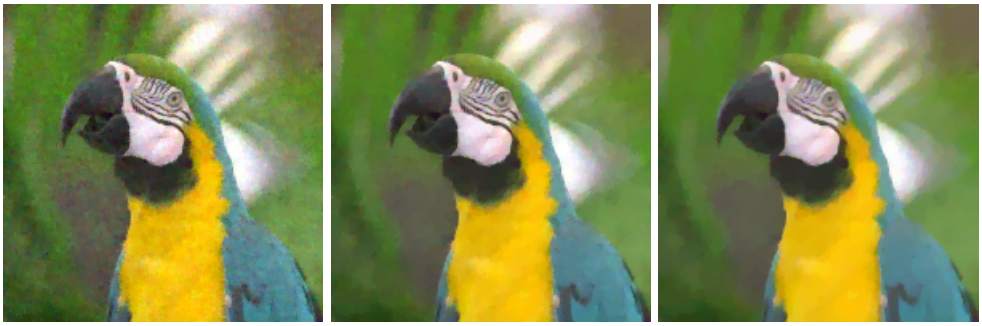
Therefore, we just need to rescale each column's norm

$$\hat{A}V = \begin{bmatrix} | & | \\ \hat{\sigma}_1 u_1 & \hat{\sigma}_2 u_2 \\ | & | \end{bmatrix}$$

And therefore

$$\hat{A} = \begin{bmatrix} | & | \\ \hat{\sigma}_1 u_1 & \hat{\sigma}_2 u_2 \\ | & | \end{bmatrix} V^\top$$

(a) The original image (b) The noisy image (c) Nuclear norm denoised



(d) Spectral norm denoised (e) Frobenius norm denoised (f) Channel-wise TV denoised

Figure 2: Colour image denoising comparison

Images	**RSNR (dB)**
Noisy	14.1294
Nuclear norm denoised	24.3723
Spectral norm denoised	22.7215
Frobenius norm denoised	24.1266
Channel-wise TV denoised	23.0542

Table 2: Colour image error metrics

Results comparison

Qualitative results are shown in Figure 2. We also show the RSNR results in Table 2.

We can see that the nuclear norm denoised has the highest RSNR, which is slightly higher than the one denoised by Frobenius norm. Channel-wise applying TV denoising without considering any coupling between the channels gives the lowest RSNR. In Figure 3 we show a close-up comparison on a patch of the image with detailed textures.

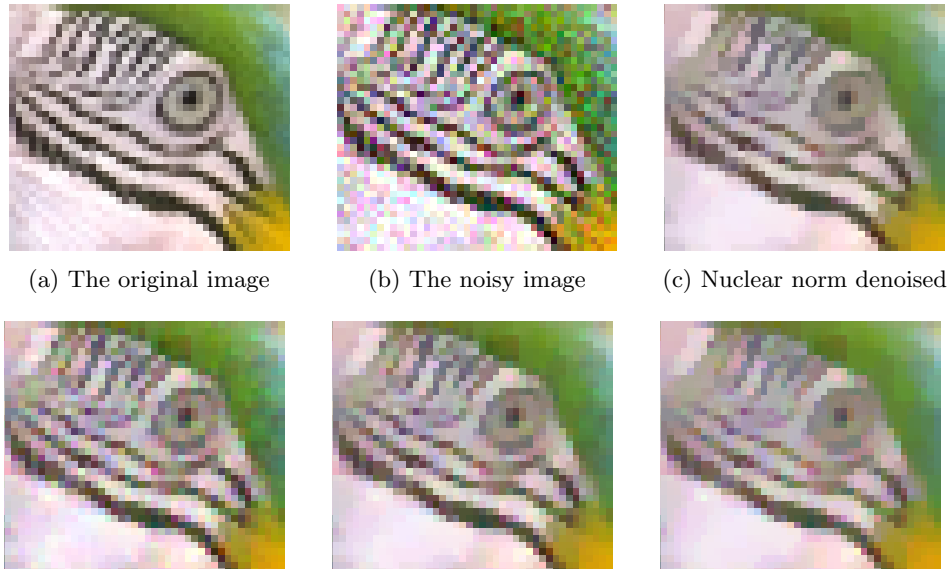
References

[1] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: nonlinear phenomena*, vol. 60, no. 1-4, pp. 259–268, 1992.

[2] A. Chambolle and T. Pock, "An introduction to continuous optimization for imaging," *Acta Numerica*, vol. 25, pp. 161–319, 2016.

[3] L. Condat, "A primal–dual splitting method for convex optimization involving lipschitzian, proximable and linear composite terms," *Journal of optimization theory and applications*, vol. 158, no. 2, pp. 460–479, 2013.

[4] K. Bredies, K. Kunisch, and T. Pock, "Total generalized variation," *SIAM Journal on Imaging Sciences*, vol. 3, no. 3, pp. 492–526, 2010.



(a) The original image (b) The noisy image (c) Nuclear norm denoised

(d) Spectral norm denoised (e) Frobenius norm denoised (f) Channel-wise TV denoised

Figure 3: Colour image denoising detailed comparison

[5] N. Parikh and S. Boyd, "Proximal algorithms," *Foundations and Trends in optimization*, vol. 1, no. 3, pp. 127–239, 2014.