CS-439: Image Denoising Using a Chambolle Scheme with Isotropic and Anisotropic Total Variations

Christopher Cherfan, Lenny Del Zio, Romain Rochepeau
Department of Computer Science, EPFL, Switzerland

***Abstract*—The goal of this paper is to use the Chambolle dual algorithm with an isotropic and anisotropic total variation regularization to perform denoising on sample images. We use some mathematical tricks to calculate the required proximal operators for the iterative algorithm. The results will then be compared according to the choices of the hyperparameters via analysis of some image quality metrics.**

I. Introduction

Image restoration is an important task in computer vision. It aims to improve the quality of a picture by removing various types of distortions. Given a noisy image $y \in \mathbb{R}^{n_1 \times n_2}$, one would like to to retrieve the original x by minimizing the following objective [1]:

$$\min_x \frac{1}{2}\|Ax - y\|_2^2 + \mathcal{R}(x) \tag{1}$$

where A is some linear operator that depends on the specific task, and $\mathcal{R}$ is a regularizer. In the case of denoising, A is just the identity [2] and the regularizer is usually chosen to be the isotropic or anisotropic total variation (TV). Some papers have proposed a weighted difference of both regularizers that slightly improved the denoised image [3] [4]. In this report, we will investigate the effect of adding a sum of isotropic and anisotropic TV norms, and minimize the corresponding objective function with a dual Chambolle scheme.

II. Theory

A. Discrete Gradient and TV

Define the discrete gradient D as the linear operator which maps an image to another image whose values are pairs of vertical and horizontal finite differences:

$$\begin{aligned}(Dx)_{n_1,n_2} &= ((Dx)_{n_1,n_2,v}, (Dx)_{n_1,n_2,h}) \\ &= (x_{n_1+1,n_2} - x_{n_1,n_2}, x_{n_1,n_2+1} - x_{n_1,n_2})\end{aligned} \tag{2}$$

We assume Neumann boundary conditions, *ie* all differences across the boundary are zero. One can similarly define its adjoint D^*. We can now write the expressions of the isotropic and anisotropic TV regularizer [5]:

$$TV_{iso}(x) = \sum_{n_1,n_2} \sqrt{(Dx)^2_{n_1,n_2,v} + (Dx)^2_{n_1,n_2,h}} \tag{3}$$

$$TV_{aniso}(x) = \sum_{n_1,n_2} |(Dx)_{n_1,n_2,v}| + |(Dx)_{n_1,n_2,h}| \tag{4}$$

Notice that they respectively correspond to the $\ell_{2,1}$ and ℓ_1 matrix norms of Dx [6]. This will be important in the dual Chambolle algorithm when calculating proximal operators.

B. Dual Chambolle Scheme

Consider the primal objective

$$\min_{x \in \mathbb{R}^n} f(x) + g(Ax) \tag{5}$$

where $A \in \mathbb{R}^{p \times n}$ is a linear operator, $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^p \to \mathbb{R}$ are convex functions. We further assume that f is β-strongly convex, which ensures that its convex conjugate f^* is differentiable and its gradient is β-Lipschitz continuous, *ie* f^* is β-smooth. The Fenchel-Rockafellar theorem states that it is equivalent to solving the dual problem [7]

$$\min_{u \in \mathbb{R}^p} f^*(-A^*u) + g^*(u) \tag{6}$$

Hence, we can find the optimal solution x_{opt} of (5) by getting u_{opt} from (6) and calculating

$$x_{opt} = \nabla f^*(-A^* u_{opt}) \tag{7}$$

Since f^* is a "nice" function, equation (6) can be solved with proximal gradient descent:

$$u_{t+1} = \text{prox}_{\gamma, g^*}(u_t + \gamma A \nabla f^*(-A^* u_t)) \tag{8}$$

where we have taken the gradient of the function $F(u) = f^*(-A^*u)$ in the second term. Convergence to the optimal is guaranteed for $0 < \gamma < \frac{2}{\beta \|L\|^2}$ [8].

C. Formulation and Algorithm

In our case, the primal objective has the form

$$\min_{x \in \mathbb{R}^{n_1 \times n_2}} \frac{1}{2}\|x - y\|_2^2 + \lambda \|Dx\|_{2,1} + \alpha \|Dx\|_1 \tag{9}$$

and the dual is

$$\min_{u \in \mathbb{R}^{n_1 \times n_2 \times 2}} \frac{1}{2}\|-D^*u + y\|_2^2 - \frac{1}{2}\|y\|_2^2 + g^*(u) \tag{10}$$

where g^* is the convex conjugate of the sum of isotropic and anisotropic regularizers, whose expression is not important because we only care about its proximal operator, as shown in (8). We thus summarize the algorithm in Algorithm 1.

Algorithm 1 Chambolle Dual Algorithm

1: Input: Initial guess $u_0 \in \mathbb{R}^{n_1 \times n_2 \times 2}$, Stepsize γ, number of iterations T
2: **for** $i \leftarrow 1$ to T **do**
3: $\quad x_t = \nabla f^*(-D^* u_t)$
4: $\quad u_{t+1} = \text{prox}_{\gamma, g^*}(u_t + \gamma D x_t)$
5: **end for**

D. Calculating Proximal Operators

In step 3 of the algorithm, it is easy to calculate the gradient of f^*:

$$\nabla f^*(u) = u + y \tag{11}$$

However, $\text{prox}_{\gamma, g^*}$ is usual not that trivial to evaluate, but it can be found exactly in our case. By Moreau's identity [9],

$$x = \text{prox}_{\gamma, g^*}(x) + \gamma \text{prox}_{g/\gamma}(x/\gamma) \tag{12}$$

Setting $g_1(x) = \lambda \|x\|_{2,1}$ and $g_2(x) = \alpha \|x\|_1$, we get that

$$\text{prox}_{\gamma, g^*}(x) = x - \gamma \text{prox}_{1/\gamma, (g_1 + g_2)}(x/\gamma) \tag{13}$$

For the chosen g_1 and g_2, it turns out that [10]

$$\text{prox}_{\gamma, (g_1 + g_2)} = \text{prox}_{\gamma, g_1} \circ \text{prox}_{\gamma, g_2} \tag{14}$$

Hence, we can calculate u_{t+1} in step 4 of the algorithm via

$$\text{prox}_{\gamma, g^*}(x) = x - \gamma \text{prox}_{1/\gamma, g_1}(\text{prox}_{1/\gamma, g_2}(x/\gamma)) \tag{15}$$

for $x = u_t + \gamma D x_t$. Additionally, the proximal operators of g_1 and g_2 can be evaluated exactly [11] [8]:

$$\text{prox}_{\gamma, g_1}(x_j) = \big(1 - \frac{\gamma \lambda}{\max(\|x_j\|_2, \lambda)}\big) x_j \tag{16}$$

$$\text{prox}_{\gamma, g_2}(x_i) = \begin{cases} x_i + \gamma\alpha & x_i \leq -\gamma\alpha \\ 0 & -\gamma\alpha \leq x_i \leq \gamma\alpha \\ x_i - \gamma\alpha & x_i \geq \gamma\alpha \end{cases} \tag{17}$$

III. Implementation

A. Noise

The image of interest is a simple black and white stair photo displayed below. We decided to corrupt it with Gaussian noise with a noise level of 45.

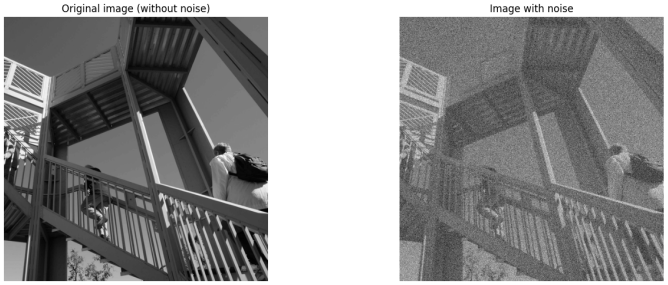

Fig. 1. Original image without noise (on the left) and its corresponding noisy version (on the right)

B. Hyper-Parameter Tuning

We first aimed at finding a combination of α and λ that would the highest peak signal-to-noise ratio (PSNR) and structural similarity index measure (SSIM) value [12]. The former is an easy to calculate logarithmic scale that measures the ratio between the maximum possible power of a signal and the power of corrupting noise, while the latter considers three parameters namely luminance, contrast and structural information between the two images. In other words, PSNR is a pixel-based measure and SSIM aligns with visual human perception. Our grid search can be seen in the following two heat maps.

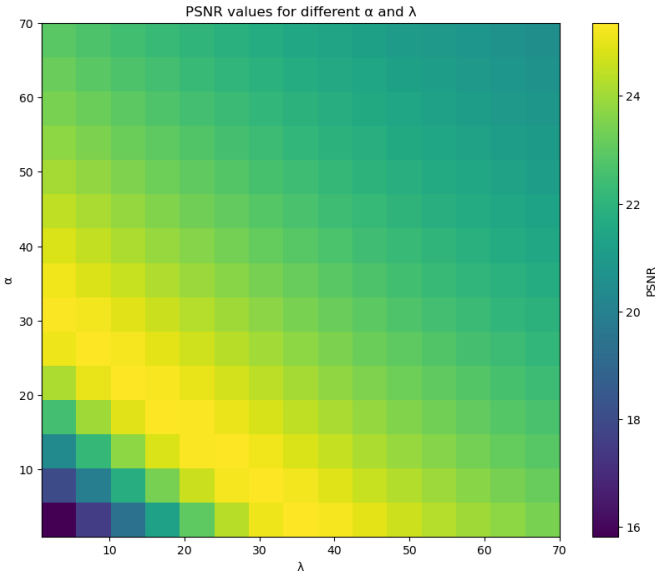

Fig. 2. Heat map showing the variation of the PSNR of the image in figure 1 as a function of the hyperparameters λ and α

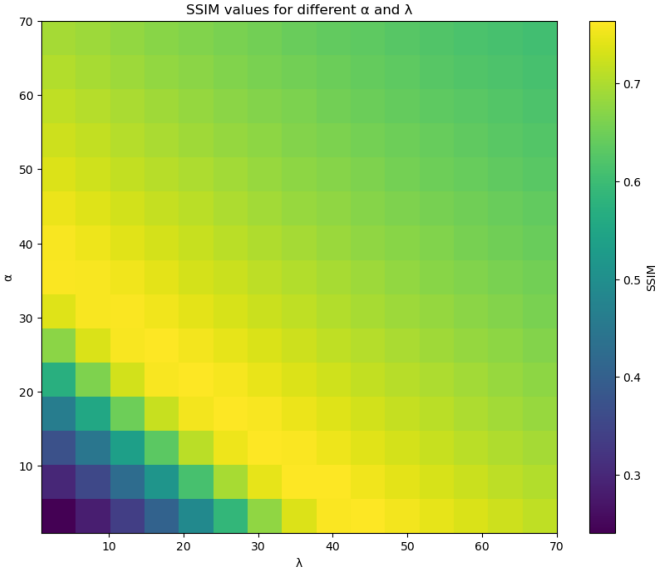

Fig. 3. Heat map showing the variation of the SSIM of the image in figure 1 as a function of the hyperparameters λ and α

Fig. 4. Denoised images with the dual Chambolle algorithm for different values of α and λ, with their corresponding PSNR and SSIM.

C. Results

Based on the heat maps, we choose combinations of λ and α that maximize the PSNR and SSIM. In particular, we compare when the regularizers' coefficients are balanced to when only one of the isotropic or anisotropic term is conserved. When removing the contribution of the isotropic term, we set λ as a very small value because putting zero leads to a numerical error in the calculation of the proximal operator.

We find that the SSIM is practically the same for all three combinations. However, the PSNR is the highest when we only consider the anisotropic TV regularizer while it is slightly lower when we consider the sum of the isotropic and anisotropic terms.

Finally, we plot the value of the primal objective (equation (5)) as a function of the number of iterations. Again, the proposed regularizer leads to a slightly higher converging value compared to the other two cases.

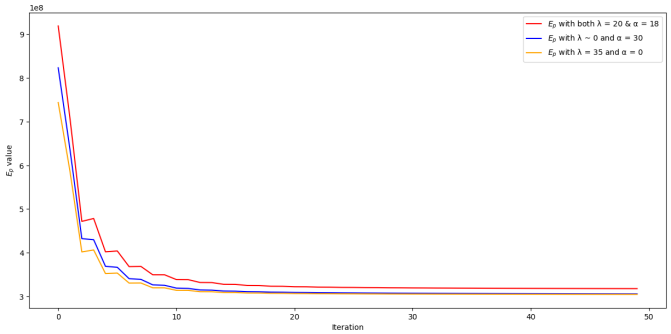

Fig. 5. Graph showing the variation of the primal objective in terms of the number of iterations of the dual Chambolle algorithm, for the three combinations of λ and α chosen.

IV. Conclusion

As demonstrated, the proposed objective is able to perform image denoising almost as efficiently as the ones that use a single TV regularizer, as they display very close PSNR and SSIM. Future work might involve adapting the algorithm to minimize a similar function with a weighted difference of the two norms, as it is done in [3] [4].

References

[1] L. Yu, M. Ng, M. Nikolova, and J. Barlow, "Efficient minimization methods of mixed l2-l1 and l1-l1 norms for image restoration," *SIAM Journal on Scientific Computing*, vol. 27, pp. 1881–1902, 01 2006.

[2] L. Fan, F. Zhang, H. Fan, and C. Zhang, "Brief review of image denoising techniques," *Visual Computing for Industry, Biomedicine, and Art*, vol. 2, 12 2019.

[3] K. Bui, F. Park, Y. Lou, and J. Xin, "A weighted difference of anisotropic and isotropic total variation for relaxed mumford-shah color and multiphase image segmentation," 2021.

[4] K. Bui, Y. Lou, F. Park, and J. Xin, "Difference of anisotropic and isotropic tv for segmentation under blur and poisson noise," 2023.

[5] L. Condat, "Discrete total variation: New definition and minimization," *SIAM Journal on Imaging Sciences*, vol. 10, no. 3, pp. 1258–1290, 2017. [Online]. Available: https://doi.org/10.1137/16M1075247

[6] Y. Yang, H. T. Shen, Z. Ma, Z. Huang, and X. Zhou, "l2,1-norm regularized discriminative feature selection for unsupervised learning," in *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence - Volume Volume Two*, ser. IJCAI'11. AAAI Press, 2011, p. 1589–1594.

[7] O. Nachum and B. Dai, "Reinforcement learning via fenchel-rockafellar duality," *CoRR*, vol. abs/2001.01866, 2020. [Online]. Available: http://arxiv.org/abs/2001.01866

[8] 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, pp. 120–145, 2011. [Online]. Available: https://api.semanticscholar.org/CorpusID:261281173

[9] *Chapter 6: The Proximal Operator*, pp. 129–177. [Online]. Available: https://epubs.siam.org/doi/abs/10.1137/1.9781611974997.ch6

[10] H.-J. M. Shi, S. Tu, Y. Xu, and W. Yin, "A primer on coordinate descent algorithms," 2017.

[11] N. Parikh and S. Boyd, "Proximal algorithms," *Found. Trends Optim.*, vol. 1, no. 3, p. 127–239, jan 2014. [Online]. Available: https://doi.org/10.1561/2400000003

[12] U. Sara, M. Akter, and M. S. Uddin, "Image quality assessment through fsim, ssim, mse and psnr—a comparative study," *Journal of Computer and Communications*, vol. 07, pp. 8–18, 01 2019.