Fixed Pattern Noise Removal For Multi-View Single-Sensor Infrared Camera Supplementary material

Arnaud Barral[1], Pablo Arias[2], Axel Davy[1]
[1]Université Paris-Saclay, CNRS, ENS Paris-Saclay, Centre Borelli, France
[2]Universitat Pompeu Fabra, Dept. of Information and Communication Technologies, Spain
{arnaud.barral,axel.davy}@ens-paris-saclay.fr pablo.arias@upf.edu

1. Derivations for the Chambolle-Pock method

Chambolle-Pock [1] is a first-order primal-dual optimization algorithm for convex problems of the form:

$$\min_{b\in\mathbb{R}^d} F(Kb) + G(b), \tag{1}$$

where $F : \mathbb{R}^c \to [0, +\infty]$, $G : \mathbb{R}^d \to [0, +\infty]$ are proper, convex and lower semi-continuous functions, and K is a linear operator from $\mathbb{R}^d$ to $\mathbb{R}^c$ (i.e. a $c \times d$ matrix). We will refer to this as the primal problem, and b as the primal variable. CP solves instead the following min-max problem:

$$\min_{b\in\mathbb{R}^d} \max_{\mathbf{q}\in\mathbb{R}^c} \langle Kb, \mathbf{q}\rangle + G(b) - F^*(\mathbf{q})$$

where F^* is the convex conjugate of F and $\mathbf{q}$ is an auxiliary *dual variable*. The convex conjugate is defined as follows:

$$F^*(\mathbf{q}) = \max_{\mathbf{v}\in\mathbb{R}^c} \langle \mathbf{q}, \mathbf{v}\rangle - F(\mathbf{v}). \tag{2}$$

This min-max problem is the *primal-dual problem*. Due to the convexity of F and G, if $(b^*, \mathbf{q}^*)$ solves the primal-dual problem, then b^* solves the primal problem.

The CP algorithm is given in Algorithm 1, where $(I + \sigma\partial F^*)^{-1}$ and $(I + \tau\partial G)^{-1}$ are the *resolvent operators* of F^* and G. The resolvent operator of a function $f : \mathbb{R}^n \to [0, +\infty]$ is an operator from $\mathbb{R}^n$ to $\mathbb{R}^n$ defined for each $\hat{z} \in \mathbb{R}^n$ as the solution of the following problem

$$(I + \tau\partial f)^{-1}(\hat{z}) = \arg\min_z \left\{ \frac{\|z - \hat{z}\|_2^2}{2\tau} + f(z) \right\}. \tag{3}$$

In order to derive the expressions for the CP algorithm in the main paper, we need to express the proposed FPN estimation energy

$$E(b) = \sum_{n=1}^{N} \mathrm{TV}(y_n - b) + \frac{\lambda_b}{2}\|b\|_2^2, \tag{4}$$

in the form of (1), thus defining F, G and K, and compute the corresponding resolvent operators.

For our energy E in (4) the primal variable is the FPN b, which we consider as a vector in $\mathbb{R}^{HW}$. G corresponds to the data attachment term

$$G(b) = \frac{\lambda_b}{2}\|b\|^2 = \frac{\lambda_b}{2}\sum_{p\in\Omega} b_p^2 \tag{5}$$

(recall that Ω denotes the $H \times W$ image domain). The linear operator $K : \mathbb{R}^{HW} \to \mathbb{R}^{2NHW}$ is given by

$$K(b) = (\nabla^+ b, ..., \nabla^+ b), \tag{6}$$

where $\nabla^+ : \mathbb{R}^{HW} \to \mathbb{R}^{2HW}$ is the discrete forward difference gradient operator with Neuman boundary conditions at the borders of the image. For an image $b \in \mathbb{R}^{HW}$ and $p = (i, j) \in \Omega$ a pixel in the image domain, the two components of the gradient are computed as follows:

$$(\nabla^+ b)_{p,1} = \begin{cases} b_{i+1,j} - b_{i,j} & \text{if } i < H, \\ 0 & \text{if } i = H, \end{cases} \tag{7}$$

$$(\nabla^+ b)_{p,2} = \begin{cases} b_{i,j+1} - b_{i,j} & \text{if } j < W, \\ 0 & \text{if } j = W. \end{cases} \tag{8}$$

Thus K concatenates N copies of the gradient of b. Its transpose, denoted by $K^\star$ in Alg. 1, is given by $K^* = \sum_{n=1}^{N} \mathrm{div}^-$, where div^- is the discrete backwards difference divergence operator, and is defined as $\mathrm{div}^- = -(\nabla^+)^\star$. For an image $v \in \mathbb{R}^{2HW}$ with 2 channels at each pixel, the div^- operator at pixel $p = (i, j) \in \mathbb{R}^\Omega$ is given by

$$(\mathrm{div}^- v)_p = \begin{cases} v_{i,j,1} - v_{i-1,j,1} & \text{if } 1 < i < H, \\ v_{i,j,1} & \text{if } i = 1, \\ -v_{i-1,j,1} & \text{if } i = H, \end{cases} + \begin{cases} v_{i,j,2} - v_{i,j-1,2} & \text{if } 1 < j < W, \\ v_{i,j,2} & \text{if } j = 1, \\ -v_{i,j-1,2} & \text{if } j = W. \end{cases} \tag{9}$$

The input of F are elements in $\mathbb{R}^{2NHW}$, i.e. N concatenated $H \times W$ images with two channels at each pixel. Thus,

if $\mathbf{v} \in \mathbb{R}^{2NHW}$, we use the following notation to access its pixel values:

$$\mathbf{v}_{n,p} = (\mathbf{v}_{n,p,1}, \mathbf{v}_{n,p,1}) \in \mathbb{R}^2,$$

for $n = 1, \ldots, N$ and $p \in \Omega$. Then $F : \mathbb{R}^{2NHW} \to \mathbb{R}$ is given by

$$F(\mathbf{v}) = \|\mathbf{v} - \mathbf{g}\|_1 = \sum_{n=1}^{N} \sum_{p \in \Omega} |\mathbf{v}_{n,p} - \mathbf{g}_{n,p}|_2 = \sum_{n=1}^{N} \sum_{p \in \Omega} \sqrt{(\mathbf{v}_{n,p,1} - \mathbf{g}_{n,p,1})^2 + (\mathbf{v}_{n,p,2} - \mathbf{g}_{n,p,2})^2} \tag{10}$$

where

$$\mathbf{g} = (\nabla^+ y_1, \nabla^+ y_2, ..., \nabla^+ y_N) \in \mathbb{R}^{2NHW}, \tag{11}$$

is the concatenation of the discrete gradients of the N observed images.

Algorithm 1: Chambolle-Pock min-max solver

Input F, G**:** proper, convex, l.s.c. functions from $\mathbb{R}^n \to [0, +\infty]$
Input K**:** $c \times d$ matrix
Output b**:** estimated FPN

1 $b^0, \bar{b}^0, \mathbf{q}^0 \coloneqq$ `zeros` `# init with zeros`
2 **for** m **from** 1 **to** M **do**
3 $\quad \mathbf{q}_{n,p}^{m+1} = (I + \sigma^m \partial F^\star)^{-1}(\mathbf{q}^m + \sigma^m K \bar{b}^m)$
4 $\quad b^{m+1} = (I + \tau^m \partial G)^{-1}(b^m - \tau^m K^\star \mathbf{q}^{m+1})$
5 $\quad \theta = \frac{1}{(1+2\gamma\tau^m)^{\frac{1}{2}}};\ \tau^{m+1} = \theta\tau^m;\ \sigma^{m+1} = \frac{\sigma^m}{\tau^m}$
6 $\quad \bar{b}^{m+1} = b^{m+1} + \theta(b^{m+1} - b^m)$

1.1. Resolvent operator of G

Plugging in our definition of G for the FPN estimation energy, the resolvent operator results from the following quadratic minimization problem

$$(I + \tau\partial G)^{-1}(\hat{b}) = \arg\min_b \left(\frac{\|b - \hat{b}\|^2}{2\tau} + \frac{\lambda_b}{2}\|b\|^2 \right). \tag{12}$$

The solution can be obtained by setting to zero the derivative of the objective function with respect to b. This yields a simple scaling of $\hat{b}$:

$$b = (I + \tau\partial G)^{-1}(\hat{b}) = \frac{1}{1 + \tau\lambda_b}\hat{b}. \tag{13}$$

Line 4 in Algorithm 1 of the main paper results from applying this scaling to $\hat{b} = b^m - \tau^m K^\star \mathbf{q}^{m+1}$ with $\tau = \tau^m$ (see line 4 in Algorithm 1) in this supplemental).

1.2. Resolvent operator of F^*

To compute the convex conjugate of F, we express F in terms of the following norm:

$$H(\mathbf{v}) = \|\mathbf{v}\|_{2NHW} = \sum_{n=1}^{N} \sum_{p \in \Omega} |\mathbf{v}_{n,p}|_2 = \sum_{n=1}^{N} \sum_{p \in \Omega} \sqrt{\mathbf{v}_{n,p,1}^2 + \mathbf{v}_{n,p,1}^2}, \tag{14}$$

such that $F(\mathbf{v}) = H(\mathbf{v} - \mathbf{g})$. Then by the properties of the convex conjugate (see for instance [4]), we have that the conjugate of F is

$$F^*(\mathbf{q}) = H^*(\mathbf{q}) + \langle \mathbf{g}, \mathbf{q} \rangle. \tag{15}$$

Since H is a norm, its convex conjugate is the indicator function of the unit ball of the corresponding dual norm [4]. We denote this unit ball P. Thus we have

$$H^*(\mathbf{q}) = \delta_P(\mathbf{q}) = \begin{cases} 0 & \text{if } \mathbf{q} \in P \\ +\infty & \text{if } \mathbf{q} \notin P \end{cases} \tag{16}$$

where it can be shown that

$$P = \{\mathbf{q} \in \mathbb{R}^{2NHW} : \forall n = 1, ..., N, p \in \Omega,\ |\mathbf{q}_{n,p}|_2 \leq 1\}. \tag{17}$$

Thus we have that the convex conjugate of F is

$$F^*(\mathbf{q}) = \begin{cases} \langle \mathbf{q}, \mathbf{g} \rangle & \text{if } \mathbf{q} \in P \\ +\infty & \text{if } \mathbf{q} \notin P \end{cases}. \tag{18}$$

The resolvent operator of F^* results from the following optimization problem

$$(I + \sigma\partial F^*)^{-1}(\hat{\mathbf{q}}) = \operatorname*{argmin}_{\mathbf{q} \in \mathbb{R}^{2NHW}} \left(\frac{\|\mathbf{q} - \hat{\mathbf{q}}\|_2^2}{2\sigma} + F^*(\mathbf{q}) \right) = \operatorname*{argmin}_{\mathbf{q} \in P} \left(\frac{\|\mathbf{q} - \hat{\mathbf{q}}\|_2^2}{2\sigma} + \langle \mathbf{g}, \mathbf{q} \rangle \right) = \operatorname*{argmin}_{\mathbf{q} \in P} \sum_{n=1}^{N} \sum_{p \in \Omega} \left(\frac{|\mathbf{q}_{n,p} - \hat{\mathbf{q}}_{n,p}|_2^2}{2\sigma} + \langle \mathbf{g}_{n,p}, \mathbf{q}_{n,p} \rangle_2 \right).$$

We expressed the indicator function over P as the constraint. Note from the definition of P that this constraint $\mathbf{q} \in P$ can be written as NWH per-pixel constraints on the Euclidean norm of the 2 dimensional vector at each pixel $|\mathbf{q}_{n,p}|_2 \leq 1$. This makes the problem separable, and allows us to compute the solution by solving per-pixel optimization 2 dimensional problems:

$$\operatorname*{argmin}_{\mathbf{q}_{n,p}, |\mathbf{q}_{n,p}|_2 \leq 1} \frac{|\mathbf{q}_{n,p} - \hat{\mathbf{q}}_{n,p}|_2^2}{2\sigma} + \langle \mathbf{g}_{n,p}, \mathbf{q}_{n,p} \rangle_2$$

The Karush-Kuhn-Tucker [4] conditions for this constrained problem are

$$\frac{\mathbf{q}_{n,p} - \hat{\mathbf{q}}_{n,p}}{\sigma} + \mathbf{g}_{n,p} - \lambda \mathbf{q}_{n,p} = 0 \tag{19}$$

$$|\mathbf{q}_{n,p}|_2^2 - 1 \leq 0 \tag{20}$$

$$\lambda \geq 0 \tag{21}$$

$$\lambda(|\mathbf{q}_{n,p}|_2^2 - 1) = 0 \tag{22}$$

where λ is the Lagrange multiplier associated to the norm constraint. Solving this set of equalities and inequalities results in

$$\mathbf{q} = (I + \sigma \partial F^*)^{-1}(\hat{\mathbf{q}})$$

$$\iff \mathbf{q}_{n,p} = \frac{\hat{\mathbf{q}}_{n,p} - \sigma \mathbf{g}_{n,p}}{\max\{|\hat{\mathbf{q}}_{n,p} - \sigma \mathbf{g}_{n,p}|_2, 1\}}. \tag{23}$$

Line 3 in Algorithm 1 of the main paper results from applying this resolvent operator with $\sigma = \sigma^m$, to $\hat{\mathbf{q}} = \mathbf{q}^m + \sigma^m K \bar{b}^m$.

2. Additional figures

In this section we discuss some plots that were left out of the main paper due to the paper length limitation.

Minimization of the offline multi-view energy. Figure 1 shows the evolution of the minimization of the offline multi-view FPN estimation energy (Section 4.1 in the main paper). The plot compares the convergence of the CP algorithm and Adam for some values of the hyperparameters. The top plot displays the PSNR between the estimated FPN and the real FPN on the y-axis, whereas the bottom plot shows the evolution of the objective energy being minimized (energy (4)). The plot corresponds to the results with $N = 16$ synthetic images from the DIV2K dataset (see figure 3 and table 1 in the main paper). We can see the impact of the step sizes on the algorithms, and also the different asymptotic values for CP and Adam.

Online single-view FPN estimation in the presence of temporally uncorrelated noise. In Section 4.3 of the main paper we presented some results obtained with the online single-view rolling window version of our energy. In one experiment we consider the more realistic case in which temporal uncorrelated noise is also present (which we model as AWGN). To denoise the video, we first use an FPN removal algorithm, and then remove the remaining temporally uncorrelated noise with FastDVDNet [5]. The rightmost plot in figure 4 of the main paper shows the PSNR obtained between the denoised video and the clean ground truth. However, it is also interesting to see how does the temporally uncorrelated noise affect the precision of the

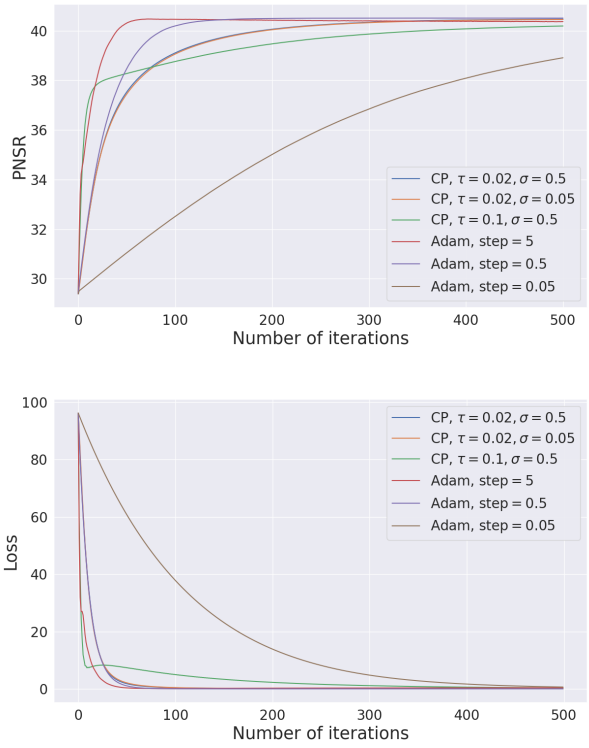

Figure 1. Minimization of the multi-view offline FPN estimation energy (4) with $N = 16$ synthetic images of the DIV2K dataset. Energy parameter λ_b is fixed at $\lambda_b = 5 \times 10^{-2}$. Evolution of the PSNR (top) and energy E (bottom) with Adam and Chambolle Pock with different step sizes.

FPN estimation. The plot in figure 2 shows the PSNR between the estimated FPN and the real FPN. This plot should be compared with the leftmost plot in figure 4 in the main paper, where we estimate the FPN for the same sequence, but this time without temporally uncorrelated noise. We can see that the temporally uncorrelated noise significantly affects the precision of the estimated FPN. The drop in PSNR is between 3dBs for our methods. It is smaller for the THPF methods, suggesting that they are producing a suboptimal result when FPN is dominant.

3. Comparison with deep learning methods

In this section, we will compare our methods with SNR-WDNN [2], a more recent deep learning method for column noise removal from a single image. We used the implementation of SNRWDNN provided by the authors to train two networks. One with only Gaussian noise vertical stripes as used by the authors with a standard deviation of $\sigma_{b_c} = 5$ and a second one with our noise model, both spatially structured and spatially independent noise with a standard deviation of $\sigma_{b_c} = \sigma_{b_u} = 5$. The networks were trained on the training

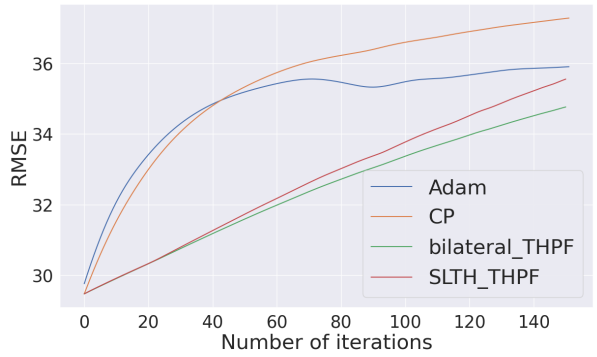

Figure 2. Graph of the evolution of the PSNR between the estimated FPN and the ground truth FPN added to the images by Chambolle Pock, by Bilateral THPF [6], by SLTH THPF [3]. Synthetic additive FPN with standard deviation of 5 and temporally uncorrelated AWGN with standard deviation of 10 have been added to the sequence. N is set to 16. For Adam the step size is equal to 0.5 and $\lambda_O = 5 \times 10^{-2}$. For CP, τ and σ are equal to 0.1 and $\lambda_O = 5 \times 10^{-2}$.

dataset of the DIV2K dataset and tested on a small subset of the validation dataset as the one used in the offline setting of our method. In the main paper we include a comparison against SNRWDNN on the use case of FPN removal from a single sequence. In this section we show a comparison in the offline multi-image setting, where our method uses 16 frames.

Table 1. Average PSNR results of 16 frames from the DIV2K validation dataset, the same as the ones used in the offline setting. For column noise, Gaussian noise vertical stripes with a standard deviation of 5 were added to the image. Full noise refers to our model, which means simulated additive FPN (no temporal independent noise), spatially structured and spatially independent with a standard deviation of $\sigma_{b_c} = \sigma_{b_u} = 5$ were added to the images. N is set to 16. For Adam the step size is equal to 0.5.

FPN	Noisy	SNRWDNN [2]	Adam
Column noise	34.5	39.4	41.7
Full	29.5	33.2	40.5

Our method is better both quantitatively and qualitatively. There is a gap of 2dBs between SNRWDNN and our method on stripes noise and 7dBs on our full noise model. Visual results show there are much more artefacts from SNRWDNN on images (e) and (m), structured noise remain on the image. This is not surprising since our method is using $N = 16$ input frames where as SNRWDNN uses a single frame.

(a) ground truth (b) noisy (c) ground truth (d) noisy

(e) Denoised Adam (f) Denoised SNRWDNN (g) Denoised Adam (h) Denoised SNRWDNN

(i) Noisy column (j) Denoised Adam (k) Noisy column (l) Denoised Adam

(m) Denoised snrwdnn (n) Denoised snrwdnn

Figure 3. Visual Comparison of some images from the DIV2K dataset. Simulated additive (no temporal independent noise), spatially structured and spatially independent with a standard deviation of $\sigma_{b_c} = \sigma_{b_u} = 5$ was added to frames (b), (c), (d), (e), (f), (g), (h) and stripe noise with standard deviation of 5 was added to the image (i), (j), (k), (l), (m), (n). (a), (b), (e), (f), (i), (j), (m), are respectively the ground truth clean image, the noisy image with simulated FPN, the image (b) denoised by our method with Adam, the image (b) denoised by SNRWDNN, the image with simulated stripes noise, the image (j) denoised by our method with Adam, the image (j) denoised by SNRWDNN. Images (c), (d), (g), (h), (k), (l) and (n) are the same for another frame from the DIV2K validation dataset.

References

[1] Antonin Chambolle and Thomas Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision*, 40, 05 2011. 1

[2] Juntao Guan, Rui Lai, and Ai Xiong. Wavelet deep neural network for stripe noise removal. *IEEE Access*, 7:44544–44554, 2019. 3, 4

[3] Weixian Qian, Qian Chen, and Guohua Gu. Space low-pass and temporal high-pass nonuniformity correction algorithm. *Optical Review*, 17:24–29, 02 2010. 4

[4] R Tyrrell Rockafellar. *Convex analysis*, volume 11. Princeton university press, 1997. 2, 3

[5] Matias Tassano, Julie Delon, and Thomas Veit. Fastdvdnet: Towards real-time deep video denoising without flow estimation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 1354–1363, 2020. 3

[6] Chao Zuo, Qian Chen, Guohua Gu, and Weixian Qian. New temporal high-pass filter nonuniformity correction based on bilateral filter. *Optical Review*, 18:197–202, 03 2011. 4