Parallel ProXimal Algorithm for data restoration

C. Chaux, J.-C. Pesquet and N. Pustelnik

Lab. IGM - Université Paris-Est

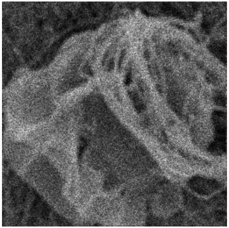
Porquerolles, October 19-21, 2009

nD data restoration

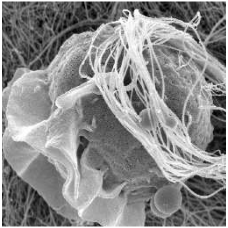
- We observe an image $z \in \mathbb{R}^M$ degraded by
  - a linear operator T (e.g. a blur)
  - a noise (e.g. Gaussian, Poisson noise)

$$\boxed{z = T\overline{y} + w}$$

- Objective: restore the unknown original image $\bar{y} \in \mathbb{R}^N$



$\Longrightarrow$



z $\bar{y}$?

- Use of convex optimization to restore the degraded image.

Outline

1. **Parallel ProXimal Algorithm (PPXA)**
2. **PPXA for image restoration** : proximity operator associated to linear degradation model and some class of discrete approximation of TV.
3. **Accelerated PPXA**
4. **Conclusion**

Convex optimization

Minimization problem

$$\text{Find} \quad \min_{\xi \in \mathcal{H}} \quad \sum_{j=1}^{J} f_j(\xi)$$

where $(f_j)_{1 \leq j \leq J}$ be functions in the class $\Gamma_0(\mathcal{H})$.
This criterion can be non differentiable.

$J = 2$: $\Longrightarrow$ a Bayesian interpretation can be formulated letting f_1 be a data fidelity term and f_2 an a priori term.

- Forward-Backward algorithm f_1 or f_2 is β-Lipschitz differentiable ($\beta \in \,]0, +\infty[$) [Figueiredo and Nowak, 2003][Bect et al., 2004][Daubechies et al., 2004][Combettes and Wajs, 2005][Chaux et al., 2007],...
- Douglas-Rachford algorithm [Lions and Mercier, 1979] [Eckstein and Bertsekas, 1992][Combettes and Pesquet, 2007]

Convex optimization

Minimization problem

$$\text{Find} \quad \min_{\xi \in \mathcal{H}} \quad \sum_{j=1}^{J} f_j(\xi)$$

where $(f_j)_{1 \leq j \leq J}$ be functions in the class $\Gamma_0(\mathcal{H})$.
This criterion can be non differentiable.

$J = 3$: the problem can be $\min_{\xi \in \mathcal{H}} f_1(\xi) + f_2(\xi) + \iota_C(\xi)$

- Nested algorithms
  β-Lipschitz differentiability of f_1 or f_2
  [Dupé et al., 2008][Chaux et al., 2009]

$J \geq 3$ and no $(f_j)_{1 \leq j \leq J}$ is β-Lipschitz differentiable

- Parallel ProXimal Algorithm (PPXA)
  [Combettes and Pesquet, 2008]

PPXA : $\min_{\xi\in\mathcal{H}} \sum_{j=1}^{J} f_j(\xi)$

Initialization

Set $\gamma \in \,]0,+\infty[$.
Set $(\omega_j)_{1\le j\le J} \in]0,1]^J$ such that $\sum_{j=1}^{J} \omega_j = 1$.
Set $(u_{j,0})_{1\le j\le J} \in (\mathcal{H})^J$ and $\xi_0 = \sum_{j=1}^{J} \omega_j u_{j,0}$.

Iterations [Combettes and Pesquet, 2008]

$$
\begin{array}{ll}
\text{For } \ell = 0,1,\ldots & \\
\quad \text{For } j = 1,\ldots,J & \text{Prox. computation} \\
\quad \lfloor \quad p_{j,\ell} = \operatorname{prox}_{\gamma f_j/\omega_j} u_{j,\ell} + a_{j,\ell} & \leftarrow \text{with possible errors} \\
\quad p_\ell = \sum_{j=1}^{J} \omega_j p_{j,\ell} & \leftarrow \text{Weighted sum} \\
\quad \text{Set } \lambda_\ell \in \,]0,2[& \\
\quad \text{For } j = 1,\ldots,J & \\
\quad \lfloor \quad u_{j,\ell+1} = u_{j,\ell} + \lambda_\ell \left(2\, p_\ell - \xi_\ell - p_{j,\ell}\right) & \leftarrow \text{Update} \\
\quad \xi_{\ell+1} = \xi_\ell + \lambda_\ell \left(p_\ell - \xi_\ell\right) & \leftarrow \text{Update}
\end{array}
$$

PPXA : convergence

The sequence $(\xi_\ell)_{\ell\geq 1}$ generated by the PPXA can be shown to converge weakly to a minimizer of $\sum_{j=1}^{J} f_j$ under the following assumption [Combettes and Pesquet, 2008].

1. $\lim_{\|\xi\|\to+\infty} f_1(\xi)+\ldots+f_J(\xi)=+\infty$.
2. $\operatorname{dom} f_1 \cap \bigcap_{j=2}^{J} \operatorname{int}\operatorname{dom} f_j \neq \varnothing$.
3. $(\forall j\in\{1,\ldots,J\})\ \sum_{\ell\in\mathbb{N}} \lambda_\ell \,\|a_{j,\ell}\| < +\infty$.
4. $\sum_{\ell\in\mathbb{N}} \lambda_\ell\,(2-\lambda_\ell) = +\infty$.

Convex optimization

Minimization problem

Find $\min_{\xi \in \mathcal{H}} \sum_{j=1}^{J} f_j(\xi)$ where $(f_j)_{1 \le j \le J}$ are in the class $\Gamma_0(\mathcal{H})$.

- f_j can be related to noise
  - $\forall y \in \mathbb{R}^N$, $f_j(y) = \frac{1}{2\sigma^2}\|Ty - z\|^2$ for Gaussian noise
  - $\forall y \in \mathbb{R}^N$, $f_j(y) = D_{KL}(Ty, z)$ for Poisson noise
- f_j can be related to a constraint
  - $\forall y \in \mathbb{R}^N$, $f_j(y) = \iota_C(y)$ where $C = [0, 255]^N$ (pixel range constraint)
- f_j can be related to some a priori on the target solution
  - $\forall y \in \mathbb{R}^N$, $f_j(y) = \|y\|_2^2$ for Tikhonov regularization
  - $\forall y \in \mathbb{R}^N$, $f_j(y) = \mathrm{tv}(y)$ for total variation regularization
  - $\forall y \in \mathbb{R}^N$, $f_j(y) = \|y\|_1$ to promote sparsity

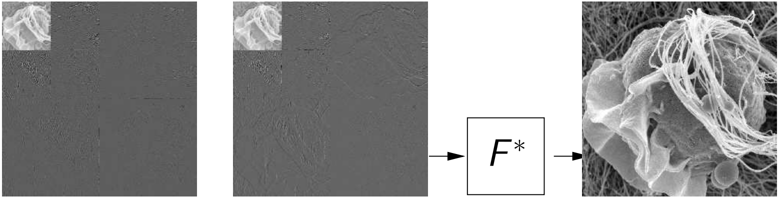
Convex optimization

Minimization problem

Find $\min_{\xi \in \mathcal{H}} \sum_{j=1}^{J} f_j(\xi)$ where $(f_j)_{1 \leq j \leq J}$ are in the class $\Gamma_0(\mathcal{H})$.

- f_j can be related to noise
  - $\forall y \in \mathbb{R}^N$, $f_j(y) = \frac{1}{2\sigma^2}\|Ty - z\|^2$ for Gaussian noise
  - $\forall y \in \mathbb{R}^N$, $f_j(y) = D_{KL}(Ty, z)$ for Poisson noise
- f_j can be related to a constraint
  - $\forall y \in \mathbb{R}^N$, $f_j(y) = \iota_C(y)$ where $C = [0, 255]^N$ (pixel range constraint)
- f_j can be related to some a priori on the target solution
  - $\forall y \in \mathbb{R}^N$, $f_j(y) = \|y\|_2^2$ for Tikhonov regularization
  - $\forall y \in \mathbb{R}^N$, $f_j(y) = \mathrm{tv}(y)$ for total variation regularization
  - $\forall y \in \mathbb{R}^N$, $\boxed{f_j(y) = \|y\|_1}$ to promote sparsity

Frame representation



Frame coefficients ($\overline{x}$) Original ($\overline{y}$)

- $\overline{x} \in \mathbb{R}^K$: **Frame coefficients** of original image $\overline{y} \in \mathbb{R}^N$
- $F^* : \mathbb{R}^K \to \mathbb{R}^N$: **Tight frame synthesis operator** such that $\exists \nu \in]0, +\infty[, \; F^* \circ F = \nu \mathrm{Id}$

$$\boxed{\bar{y} = F^* \overline{x}}$$

Motivations

Our objective is to take $J > 2$. Why ?

- it allows us to mix constraints and regularization functions which has proved to be fruitful.

  [Bect et al., 2004][Bioucas-Dias and Figueiredo, 2008][Combettes and Pesquet, 2008]

Wavelet regularization

$\Downarrow$

lack of regularity

Total variation regularization

$\Downarrow$

staircase effects

Considered minimization problem

Our objective

$\min_{x\in\mathbb{R}^K} \quad d(TF^*x, z) + \kappa\,\mathrm{tv}(F^*x) + \iota_C(F^*x) + \vartheta\, f(x)$
where $\kappa > 0$, $\vartheta > 0$.

- $d(\cdot, z) \in \Gamma_0(\mathbb{R}^M)$: data fidelity term .

- tv: total variation term

- ι_C: indicator function of a closed convex set $C = [0, 255]^N$

- $\forall x = (x^{(k)})_{1\le k\le K} \in \mathbb{R}^K$, $f(x) = \sum_{k=1}^K \phi_k(x^{(k)})$ where, for every $k \in \{1, \dots, K\}$, ϕ_k is a finite function of $\Gamma_0(\mathbb{R})$ such that $\lim_{|x^{(k)}|\to+\infty} \phi_k(x^{(k)}) = +\infty$

Considered minimization problem

Our objective

$\min_{x\in\mathbb{R}^K} \quad d(TF^*x, z) + \kappa\,\mathrm{tv}(F^*x) + \iota_C(F^*x) + \vartheta\, f(x)$
where $\kappa > 0$, $\vartheta > 0$.

- $d(\cdot, z) \in \Gamma_0(\mathbb{R}^M)$: data fidelity term .
  $\Rightarrow$ No explicit proximity operator expression except when d is quadratic.
- tv: total variation term
  $\Rightarrow$ Proximity operators proposed for several forms of tv.
- ι_C: indicator function of a closed convex set $C = [0, 255]^N$
  $\Rightarrow$ P_C.
- $\forall x = (x^{(k)})_{1\le k\le K} \in \mathbb{R}^K$, $f(x) = \sum_{k=1}^K \phi_k(x^{(k)})$ where, for every $k \in \{1, \ldots, K\}$, ϕ_k is a finite function of $\Gamma_0(\mathbb{R})$ such that $\lim_{|x^{(k)}|\to+\infty} \phi_k(x^{(k)}) = +\infty$
  $\Rightarrow$ Explicit form.

Proximity operator computation of $d(TF^*., z)$

- $\forall x \in \mathbb{R}^K$, $\boxed{d(TF^*x, z) = \Psi(TF^*x)} = \sum_{m=1}^{M} \psi_m((TF^*x)^{(m)})$
- $\forall m \in \{1, \ldots, M\}$, explicit form for prox_{ψ_m} $\Rightarrow$ Explicit form for prox_{Ψ}
- How to circumvent $\mathrm{prox}_{\Psi \circ T \circ F^*}$?

Proximity operator computation of $d(TF^*., z)$

- $\forall x \in \mathbb{R}^K$, $\boxed{d(TF^*x, z) = \Psi(TF^*x)} = \sum_{m=1}^{M} \psi_m((TF^*x)^{(m)})$
- $\forall m \in \{1, \dots, M\}$, explicit form for prox_{ψ_m} $\Rightarrow$ Explicit form for prox_{Ψ}
- How to circumvent $\mathrm{prox}_{\Psi \circ T \circ F^*}$?

Proposition

- $\mathcal{H}$ and $\mathcal{G}$: real separable Hilbert spaces.
- $(e_m)_{m \in \mathbb{K}}$ an orthonormal basis of $\mathcal{G}$ such that $(\forall u \in \mathcal{G})$, $\Phi(u) = \sum_{m \in \mathbb{K}} \varphi_m(\langle u, e_m \rangle)$ where $(\varphi_m)_{m \in \mathbb{K}} \in \Gamma_0(\mathbb{R})$.
- $L \colon \mathcal{H} \to \mathcal{G}$: bounded linear operator
- Suppose that the composition of L and L^* is an isomorphism which is diagonalized by $(e_m)_{m \in \mathbb{K}}$ (*i.e.* $(\forall m \in \mathbb{K}) \quad \underbrace{L \circ L^*}_{D} e_m = d_m e_m$)

then, $\mathrm{prox}_{\Phi \circ L} = \mathrm{Id} + L^* \circ (\mathrm{prox}_{D^{-1}\Phi(D\cdot)} - \mathrm{Id}) \circ D^{-1} \circ L$.

Proximity operator computation of $d(TF^*., z)$

- $\forall x \in \mathbb{R}^K$, $\boxed{d(TF^*x, z) = \Psi(TF^*x)} = \sum_{m=1}^{M} \psi_m((TF^*x)^{(m)})$
- $\forall m \in \{1, \ldots, M\}$, explicit form for prox_{ψ_m} $\Rightarrow$ Explicit form for prox_{Ψ}
- How to circumvent $\mathrm{prox}_{\Psi \circ T \circ F^*}$?

Proposition

- $\mathcal{H}$ and $\mathcal{G}$: real separable Hilbert spaces.
- $(e_m)_{m \in \mathbb{K}}$ an orthonormal basis of $\mathcal{G}$ such that $(\forall u \in \mathcal{G})$, $\Phi(u) = \sum_{m \in \mathbb{K}} \varphi_m(\langle u, e_m \rangle)$ where $(\varphi_m)_{m \in \mathbb{K}} \in \Gamma_0(\mathbb{R})$.
- $L \colon \mathcal{H} \to \mathcal{G}$: bounded linear operator
- Suppose that the composition of L and L^* is an isomorphism which is diagonalized by $(e_m)_{m \in \mathbb{K}}$ (*i.e.* $(\forall m \in \mathbb{K})$ $\underbrace{L \circ L^*}_{D} e_m = d_m e_m$)

then, $\mathrm{prox}_{\Phi \circ L} = \mathrm{Id} + L^* \circ (\mathrm{prox}_{D^{-1}\Phi(D\cdot)} - \mathrm{Id}) \circ D^{-1} \circ L$.

- If $L = TF^*$ with $T = \mathrm{Id}$ and $F^* \circ F = \nu \mathrm{Id} \Rightarrow$ Proposition can be used
- If $L = TF^*$ with $T \neq \mathrm{Id}$ and $F^* \circ F = \nu \mathrm{Id} \Rightarrow$ Problem: $(TF^*) \circ (TF^*)^*$ is non necessarily diagonalized in the canonical basis of $\mathbb{R}^M$

Example: 1D periodic convolution

For a kernel size equal to 3, $T \circ T^* \neq D$

$$T = \begin{bmatrix} \theta_2 & \theta_1 & \theta_0 & 0 & & & \cdots & & & 0 \\ 0 & \theta_2 & \theta_1 & \theta_0 & 0 & & & & & \vdots \\ \vdots & 0 & \theta_2 & \theta_1 & \theta_0 & 0 & & & & \\ & & 0 & \theta_2 & \theta_1 & \theta_0 & 0 & & & \\ & & & 0 & \theta_2 & \theta_1 & \theta_0 & 0 & & \\ \vdots & & & & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ & & & & & 0 & \theta_2 & \theta_1 & \theta_0 & 0 \\ 0 & & & & & & 0 & \theta_2 & \theta_1 & \theta_0 \\ \theta_0 & 0 & & & & & & 0 & \theta_2 & \theta_1 \\ \theta_1 & \theta_0 & 0 & & \cdots & & & & 0 & \theta_2 \end{bmatrix}$$

Example: 1D periodic convolution

$T_1 \circ T_1^* = \sum_{q=0}^{2} |\theta_q|^2 \mathrm{Id}$

$$T = \begin{bmatrix}
\theta_2 & \theta_1 & \theta_0 & 0 & & & \cdots & & & 0 \\
0 & \theta_2 & \theta_1 & \theta_0 & 0 & & & & & \vdots \\
\vdots & 0 & \theta_2 & \theta_1 & \theta_0 & 0 & & & & \\
 & & 0 & \theta_2 & \theta_1 & \theta_0 & 0 & & & \\
 & & & 0 & \theta_2 & \theta_1 & \theta_0 & 0 & & \\
\vdots & & & & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\
 & & & & & 0 & \theta_2 & \theta_1 & \theta_0 & 0 \\
0 & & & & & & 0 & \theta_2 & \theta_1 & \theta_0 \\
\theta_0 & 0 & & & & & & 0 & \theta_2 & \theta_1 \\
\theta_1 & \theta_0 & & 0 & & \cdots & & & 0 & \theta_2
\end{bmatrix}$$

T_1

Example: 1D periodic convolution

$$T_2 \circ T_2^* = \sum_{q=0}^{2} |\theta_q|^2 \mathrm{Id}$$

$$T = \begin{bmatrix}
\theta_2 & \theta_1 & \theta_0 & 0 & & \cdots & & & & 0 \\
0 & \theta_2 & \theta_1 & \theta_0 & 0 & & & & & \vdots \\
\vdots & 0 & \theta_2 & \theta_1 & \theta_0 & 0 & & & & \\
 & & 0 & \theta_2 & \theta_1 & \theta_0 & 0 & & & \\
 & & & 0 & \theta_2 & \theta_1 & \theta_0 & 0 & & \\
\vdots & & & & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\
 & & & & & 0 & \theta_2 & \theta_1 & \theta_0 & 0 \\
0 & & & & & & 0 & \theta_2 & \theta_1 & \theta_0 \\
\theta_0 & 0 & & & & & & 0 & \theta_2 & \theta_1 \\
\theta_1 & \theta_0 & 0 & & \cdots & & & & 0 & \theta_2
\end{bmatrix}$$

T_2

Example: 1D periodic convolution

$$T_3 \circ T_3^* = \sum_{q=0}^{2} |\theta_q|^2 \mathrm{Id}$$

$$T = \begin{array}{cc} \left[\begin{matrix}
\theta_2 & \theta_1 & \theta_0 & 0 & & & \cdots & & & 0 \\
0 & \theta_2 & \theta_1 & \theta_0 & 0 & & & & & \vdots \\
\vdots & 0 & \theta_2 & \theta_1 & \theta_0 & 0 & & & & \\
 & & 0 & \theta_2 & \theta_1 & \theta_0 & 0 & & & \\
 & & & 0 & \theta_2 & \theta_1 & \theta_0 & 0 & & \\
\vdots & & & & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\
 & & & & & 0 & \theta_2 & \theta_1 & \theta_0 & 0 \\
0 & & & & & & 0 & \theta_2 & \theta_1 & \theta_0 \\
\theta_0 & 0 & & & & & & 0 & \theta_2 & \theta_1 \\
\theta_1 & \theta_0 & 0 & & \cdots & & & & 0 & \theta_2
\end{matrix}\right] & \begin{matrix} \\ \\ T_3 \\ \\ \\ \\ \\ \\ \\ \\ \end{matrix} \end{array}$$

Proximity operator computation of $\Psi(TF^*.)$

- Notations:
  - $(\mathbb{I}_i)_{1\le i\le I}$ is a partition of $\{1,\dots,M\}$ in nonempty sets
  - $\forall i \in \{1,\dots,I\}$, M_i denotes the number of elements in $\mathbb{I}_i$
  - $\Upsilon_i : \mathbb{R}^{M_i} \to \,]0,+\infty[\, : (\eta^{(m)})_{m\in\mathbb{I}_i} \mapsto \sum_{m\in\mathbb{I}_i} \psi_m(\eta^{(m)})$
- We have then $\Psi\circ T\circ F^* = \sum_{i=1}^{I} \Upsilon_i \circ T_i \circ F^*$ where T_i is the linear operator from $\mathbb{R}^N$ to $\mathbb{R}^{M_i}$ associated with the matrix $[t_{m_1},\cdot,t_{m_{M_i}}]^\top$ with $\mathbb{I}_i = \{m_1,\dots,m_{M_i}\}$.

Particular case of Periodic Convolution:
For all $i\in\{1,\dots,I\}$, $(t_m)_{m\in\mathbb{I}_i}$ is a family of non zero orthogonal vectors such that $T_i\circ T_i^* = \sigma_i \mathrm{Id}$ where
$\sigma_i = \sum_{q_1=0}^{Q_1-1}\sum_{q_2=0}^{Q_1-1} |\theta_{q_1,q_2}|^2$. For $F^*\circ F = \nu\mathrm{Id}$ (tight-frame) and
$\forall i \in (1,\dots I)$,

$$\boxed{\mathrm{prox}_{\Upsilon_i\circ T_i\circ F^*} = \mathrm{Id} + \frac{F\circ T_i^*}{\nu\sigma_i}\circ(\mathrm{prox}_{\nu\sigma_i\Upsilon_i} - \mathrm{Id})\circ T_i\circ F^*}.$$

Proximity operator computation of $\Psi(TF^*.)$

Complexity:

$$\boxed{\mathrm{prox}_{\Upsilon_i \circ T_i} = \mathrm{Id} + T_i^* \circ (\mathrm{prox}_{D_i^{-1}\Phi(D_i \cdot)} - \mathrm{Id}) \circ D_i^{-1} \circ T_i}$$

- If we have Q proximity operators to compute:
  - Complexity of each T_i and T_i^* : $O(N)$
  - Complexity of $\mathrm{prox}_{D_i^{-1}\Upsilon_i(D_i \cdot)}$: $O(M_i)$
  - Q proximity operators $\mathrm{prox}_{\Upsilon_i \circ T_i}$:
    $O(Q(2N + M_i)) = O(N(2Q + 1))$
- If we have M proximity operators to compute (when $M \sim N$):
  - Complexity of each T_i and T_i^* : $O(Q)$
  - Complexity of $\mathrm{prox}_{D_i^{-1}\Upsilon_i(D_i \cdot)}$: $O(1)$
  - N proximity operators $\mathrm{prox}_{\Upsilon_i \circ T_i}$: $O(N(2Q + 1))$

Proximity operator computation of $\Psi(TF^*.)$

Initialization

Set $\gamma \in\ ]0, +\infty[$.
Set $(\omega_j)_{1\leq j\leq J} \in]0,1]^J$ such that $\sum_{j=1}^J \omega_j = 1$.
Set $\boxed{(u_{j,0})_{1\leq j\leq J}} \in (\mathbb{R}^K)^J$ and $\xi_0 = \sum_{j=1}^J \omega_j u_{j,0}$.

Iterations [Combettes and Pesquet, 2008]

For $\ell = 0, 1, \ldots$
- For $j = 1, \ldots, J$ — Prox. computation
  - $p_{j,\ell} = \mathrm{prox}_{\gamma f_j/\omega_j} u_{j,\ell} + a_{j,\ell}$ — $\leftarrow$ with possible errors
- $p_\ell = \sum_{j=1}^J \omega_j p_{j,\ell}$ — $\leftarrow$ Weighted sum
- Set $\lambda_\ell \in\ ]0, 2[$
- For $j = 1, \ldots, J$
  - $u_{j,\ell+1} = u_{j,\ell} + \lambda_\ell \left(2\, p_\ell - \xi_\ell - p_{j,\ell}\right)$ — $\leftarrow$ Update
- $\xi_{\ell+1} = \xi_\ell + \lambda_\ell \left(p_\ell - \xi_\ell\right)$ — $\leftarrow$ Update

Considered minimization problem

Our objective

$$\min_{x\in\mathbb{R}^K} \quad d(TF^*x, z) + \kappa\,\mathrm{tv}(F^*x) + \iota_C(F^*x) + \vartheta\, f(x)$$

where $\kappa > 0$, $\vartheta > 0$.

- $d(\cdot, z)$: data fidelity term.
  $\Rightarrow$ No explicit proximity operator expression except when d is quadratic.
- tv: total variation term
  $\Rightarrow$ Proximity operators proposed for several forms of tv.
- ι_C: indicator function of a closed convex set $C = [0, 255]^N$
  $\Rightarrow P_C$.
- $\forall x = (x^{(k)})_{1\leq k\leq K} \in \mathbb{R}^K$, $f(x) = \sum_{k=1}^{K} \phi_k(x^{(k)})$ where, for every $k \in \{1, \ldots, K\}$, ϕ_k is a finite function of $\Gamma_0(\mathbb{R})$ such that $\lim_{|x^{(k)}|\to+\infty} \phi_k(x^{(k)}) = +\infty$
  $\Rightarrow$ Explicit form.

Proximity operator associated to various TV discretization

Total variation of a digital image

$y = (y_{n_1,n_2})_{0\le n_1<N_1, 0\le n_2<N_2} \in \mathbb{R}^{N_1\times N_2}$:

$$\mathrm{tv}(y) = \sum_{n_1=0}^{N_1-P_1}\sum_{n_2=0}^{N_2-P_2} \rho_{\mathrm{tv}}\big((y_h)_{n_1,n_2},(y_v)_{n_1,n_2}\big), \tag{1}$$

- $\rho_{\mathrm{tv}} \in \Gamma_0(\mathbb{R}^2)$ (ex. $\rho_{\mathrm{tv}}(\cdot,\cdot) = \sqrt{|\cdot|^2+|\cdot|^2}$, $\rho_{\mathrm{tv}}(\cdot,\cdot) = |\cdot|+|\cdot|$)
- $H \in \mathbb{R}^{P_1\times P_2}$ and $V \in \mathbb{R}^{P_1\times P_2}$ are the filter kernel matrices
- $Y_{n_1,n_2} = (y_{n_1+p_1,n_2+p_2})_{0\le p_1<P_1, 0\le p_2<P_2}$
- $(y_h)_{n_1,n_2} = \mathrm{tr}(H^\top Y_{n_1,n_2})$ and $(y_v)_{n_1,n_2} = \mathrm{tr}(V^\top Y_{n_1,n_2})$

Proximity operator for split $\mathrm{tv}(y) = \sum_{p_1=0}^{P_1-1}\sum_{p_2=0}^{P_2-1} \mathrm{tv}_{p_1,p_2}(y)$

- Possible when $\mathrm{tr}(HV^\top) = 0$ and $\|H\|_{\mathrm{F}}^2 = \|V\|_{\mathrm{F}}^2 = \tau > 0$
- Examples: Haar, Finite difference, Prewitt, Sobel.

Results

- CPU time

Image size	128×128		256×256		512×512	
Kernel blur size	3×3	7×7	3×3	7×7	3×3	7×7
Iteration numbers	30	50	41	50	50	50
CPU time (in sec.)	117.2	633.0	411.7	1298	1458	4514

- High computational time due to the number of F and F^* to compute at each iteration.

PPXA $\quad \min_{x\in\mathbb{R}^K} \sum_{j=1}^J f_j(x)$

For $\ell = 0, 1, \ldots$

For $j = 1, \ldots, J$

$\quad p_{j,\ell} = \mathrm{prox}_{\gamma f_j/\omega_j} u_{j,\ell} + a_{j,\ell}$

$p_\ell = \sum_{j=1}^J \omega_j p_{j,\ell}$

Set $\lambda_\ell \in]0, 2[$

For $j = 1, \ldots, J$

$\quad u_{j,\ell+1} = u_{j,\ell} + \lambda_\ell \left(2\, p_\ell - x_\ell - p_{j,\ell}\right)$

$x_{\ell+1} = x_\ell + \lambda_\ell (p_\ell - x_\ell)$

PPXA $\min_{x\in\mathbb{R}^K} \sum_{j=1}^{S} g_j(F^*x) + \sum_{j=S+1}^{J} f_j(x)$

For $\ell = 0, 1, \ldots$

- For $j = 1, \ldots, S$
  - $p_{j,\ell} = \mathrm{prox}_{\nu\gamma/\omega_j g_j \circ F^*}(u_{j,\ell}) + a_{j,\ell}$
- For $j = S+1, \ldots, J$
  - $p_{j,\ell} = \mathrm{prox}_{\gamma f_j/\omega_j} u_{j,\ell} + a_{j,\ell}$
- $p_\ell = \sum_{j=1}^{J} \omega_j p_{j,\ell}$
- Set $\lambda_\ell \in]0,2[$
- For $j = 1, \ldots, J$
  - $u_{j,\ell+1} = u_{j,\ell} + \lambda_\ell \left(2\, p_\ell - x_\ell - p_{j,\ell}\right)$
- $x_{\ell+1} = x_\ell + \lambda_\ell(p_\ell - x_\ell)$

PPXA $\quad \min_{x\in\mathbb{R}^K} \sum_{j=1}^{S} g_j(F^*x) + \sum_{j=S+1}^{J} f_j(x)$

For $\ell = 0, 1, \ldots$

- For $j = 1, \ldots, S$
  - $p_{j,\ell} = u_{j,\ell} + \frac{1}{\nu} F\left(\operatorname{prox}_{\nu\gamma/\omega_j g_j}(F^* u_{j,\ell}) - F^* u_{j,\ell}\right) + a_{j,\ell}$
- For $j = S+1, \ldots, J$
  - $p_{j,\ell} = \operatorname{prox}_{\gamma f_j/\omega_j} u_{j,\ell} + a_{j,\ell}$
- $p_\ell = \sum_{j=1}^{J} \omega_j p_{j,\ell}$
- Set $\lambda_\ell \in]0, 2[$
- For $j = 1, \ldots, J$
  - $u_{j,\ell+1} = u_{j,\ell} + \lambda_\ell \left(2\, p_\ell - x_\ell - p_{j,\ell}\right)$
- $x_{\ell+1} = x_\ell + \lambda_\ell(p_\ell - x_\ell)$

- Large number of F and F^* to compute

PPXA $\min_{x\in\mathbb{R}^K} \sum_{j=1}^{S} g_j(F^*x) + \sum_{j=S+1}^{J} f_j(x)$

For $\ell = 0, 1, \ldots$

For $j = 1, \ldots, S$

$F^* p_{j,\ell} = F^* u_{j,\ell} + \frac{1}{\nu} F^* F\left(\mathrm{prox}_{\nu\gamma/\omega_j g_j}(F^* u_{j,\ell}) - F^* u_{j,\ell}\right) + F^* a_{j,\ell}$

For $j = S+1, \ldots, J$

$p_{j,\ell} = \mathrm{prox}_{\gamma f_j/\omega_j} u_{j,\ell} + a_{j,\ell}$

$p_\ell = \sum_{j=1}^{J} \omega_j \boxed{p_{j,\ell}}$

Set $\lambda_\ell \in]0, 2[$

For $j = 1, \ldots, J$

$u_{j,\ell+1} = u_{j,\ell} + \lambda_\ell \left(2\boxed{p_\ell} - x_\ell - \boxed{p_{j,\ell}}\right)$

$x_{\ell+1} = x_\ell + \lambda_\ell(\boxed{p_\ell} - x_\ell)$

- Decompose as $p_{j,\ell} = F q_{j,\ell} + p_{j,\ell}^{\perp}$

Accelerated PPXA

For $\ell = 0, 1, \dots$

$$
\left\lfloor
\begin{array}{l}
\text{For } j = 1, \dots, S \\
\lfloor \quad q_{j,\ell} = \frac{1}{\nu}\mathrm{prox}_{\nu\gamma/\omega_j g_j} v_{j,\ell} + \widetilde{a}_{j,\ell} \\
\text{For } j = S+1, \dots, J \\
\lfloor \quad p_{j,\ell} = \mathrm{prox}_{\gamma/\omega_j f_j} u_{j,\ell} + a_{j,\ell} \\
p_\ell = \sum_{j=1}^{S} \omega_j u_{j,\ell}^{\perp} + F \sum_{j=1}^{S} \omega_j q_{j,\ell} + \sum_{j=S+1}^{J} \omega_j p_{j,\ell} \\
r_\ell = 2\, p_\ell - x_\ell; \qquad \widetilde{r}_\ell = F^* r_\ell; \qquad r_\ell^{\perp} = r_\ell - \frac{1}{\nu} F \widetilde{r}_\ell \\
\text{Set } \lambda_\ell \in \left]0, 2\right[\\
\text{For } j = 1, \dots, S \\
\left\lfloor
\begin{array}{l}
u_{j,\ell+1}^{\perp} = u_{j,\ell}^{\perp} + \lambda_\ell \left(r_\ell^{\perp} - u_{j,\ell}^{\perp} \right) \\
v_{j,\ell+1} = v_{j,\ell} + \lambda_\ell \left(\widetilde{r}_\ell - \nu q_{j,\ell} \right)
\end{array}
\right. \\
\text{For } j = S+1, \dots, J \\
\lfloor \quad u_{j,\ell+1} = u_{j,\ell} + \lambda_\ell \left(r_\ell - p_{j,\ell} \right) \\
x_{\ell+1} = x_\ell + \lambda_\ell (p_\ell - x_\ell)
\end{array}
\right.
$$

Results

Image size	128×128		256×256		512×512	
Kernel blur size	3×3	7×7	3×3	7×7	3×3	7×7
Iteration numbers	30	50	41	50	50	50
CPU time (s)	117.2	633.0	411.7	1298	1458	4514
CPU time acc. version (s)	13.53	29.82	60.59	89.48	263.6	405.0
Gain	8.67	21.2	6.79	14.5	5.53	11.1

Table: Comparisons between PPXA and its accelerated version.

Conclusion

- Adaptation of PPXA for a **large class of image recovery problems**.
- Convergence of accelerated PPXA.
- Parallel implementation of PPXA / acc. PPXA on 8 core processor with OpenMP
- Future work: GPU implementation