Proximal tools for image reconstruction in dynamic Positron Emission Tomography

Nelly Pustelnik[1]

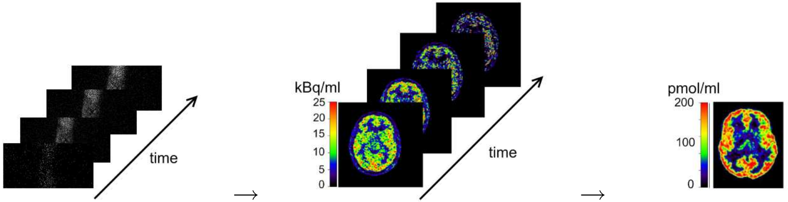
joint work with
Caroline Chaux[2], **Jean-Christophe Pesquet**[3], **and Claude Comtat**[4]

[1] Laboratoire de Physique, ENS Lyon – CNRS UMR 5672
[2] LATP, Université d'Aix-Marseille 1 –CNRS UMR 7353
[3] LIGM, Université Paris-Est – CNRS UMR 8049
[4] SHFJ, CEA

Séminaire DROITE
September, 17th 2012

Positron Emission Tomography (PET)

- **Image reconstruction for dynamic PET**



Measures → Reconstructed images → Benzodiazepine receptors density

Difficulties:

✗ Degradation: projection operator and noise.

✗ Large data size.

Outline

1. Degradation model and state-of-the-art

   → Variational approach, sparsity, frames

2. Proximal algorithms

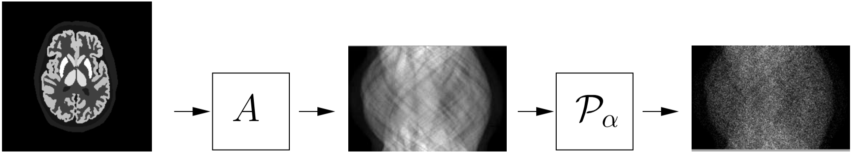
   → Proximity operators, properties, DR, PPXA, ...

3. Space+time PET reconstruction

   → PPXA, hybrid regularization, simulated and preclinical data

4. Conclusion and future works

Degradation model



Original image $\overline{y} \in \mathbb{R}^N$ — $A\overline{y}$ — Sinogram $z \in \mathbb{R}^M$

Degradation model

$$z = \mathcal{P}_\alpha(A\overline{y})$$

- $A \in \mathbb{R}^{M \times N}$: matrix associated with the projection operator.
- $\mathcal{P}_\alpha$: Poisson noise with scaling parameter $\alpha > 0$.

GOAL: Find a vector $\widehat{y}$ the closest to $\overline{y}$ from the observations z

Variational approach: State-of-the-art

- Quadratic regularization - Wiener filtering [Wiener,1949]:

$$\widehat{y} = \underset{y \in \mathbb{R}^N}{\operatorname{argmin}} \|Ay - z\|_2^2 + \lambda \|y\|_2^2 \qquad \text{where} \qquad \lambda > 0$$

Variational approach: State-of-the-art

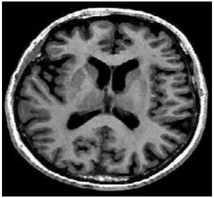
- Quadratic regularization - Wiener filtering [Wiener,1949]:

$$\widehat{y} = \underset{y \in \mathbb{R}^N}{\operatorname{argmin}} \|Ay - z\|_2^2 + \lambda \|y\|_2^2 \qquad \text{where} \qquad \lambda > 0$$

- Use of frames [Haar,1910] [Mallat,2009] - Soft-thresholding:

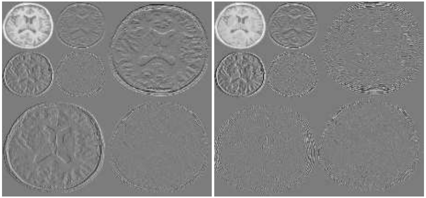
$$\widehat{y} = \underset{y \in \mathbb{R}^N}{\operatorname{argmin}} \|Ay - z\|_2^2 + \lambda \|Fy\|_1 \qquad \text{where} \qquad \lambda > 0$$

Variational approach: State-of-the-art



$\overset{F}{\longrightarrow}$

$\underset{F^\top}{\longleftarrow}$



$y \in \mathbb{R}^N$

$x = Fy \in \mathbb{R}^K$ with $K \geq N$

$F \in \mathbb{R}^{K \times N}$: matrix associated with the analysis frame operator.
$F^* \in \mathbb{R}^{N \times K}$: matrix associated with the synthesis frame operator.

Tight frame condition: $\boxed{F^*F = \mu \mathrm{Id}}$ pour $\mu > 0$.

Variational approach: State-of-the-art

- Quadratic regularization - Wiener filtering [Wiener,1949]:

$$\widehat{y} = \underset{y \in \mathbb{R}^N}{\operatorname{argmin}} \|Ay - z\|_2^2 + \lambda \|y\|_2^2 \qquad \text{where} \qquad \lambda > 0$$

- Use of frames [Haar,1910] [Mallat,2009] - Soft-thresholding:

$$\boxed{\widehat{y} = \underset{y \in \mathbb{R}^N}{\operatorname{argmin}} \|Ay - z\|_2^2 + \lambda \|Fy\|_1} \qquad \rightarrow \text{2 interpretations}$$

Variational approach: Bayesian formulation

- $u = Ay = (u^{(i)})_{1\leq i\leq N}$: realization of a random vector U.
- z: realization of a random vector Z.
- $x = Fy = (x^{(k)})_{1\leq k\leq K}$: realization of a random vector $\overline{X} = (X^{(k)})_{1\leq k\leq K}$ having independent components.

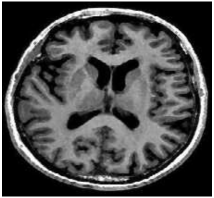
MAP estimator (Maximum A Posteriori)

$$\max_y \ \mathsf{P}(U = Ay \mid Z = z)$$

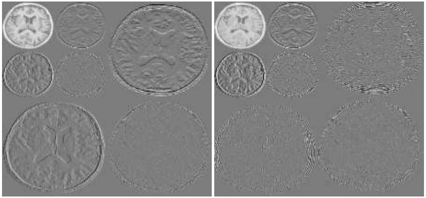
$$\max_x \ \mathsf{P}(Z = z \mid U = AF^*x) \ \cdot \ \mathsf{P}(X = x)$$

$$\min_x \ -\underbrace{\ln \mathsf{P}(Z = z \mid U = AF^*x)}_{\text{Data fidelity}} \underbrace{- \ln \prod_{k=1}^{K} \mathsf{p}_{X^{(k)}}(x^{(k)})}_{\text{A priori}}$$

Variational approach: Bayesian formulation

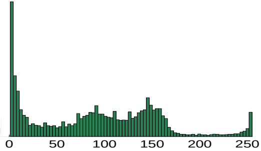


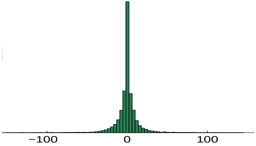
$$\xrightarrow{F} \qquad \xleftarrow[F^\top]{}$$



$y \in \mathbb{R}^N$

$x = Fy \in \mathbb{R}^K$ with $K \geq N$





Variational approach: Bayesian formulation

$$\boxed{\min_x \; -\underbrace{\ln \mathsf{P}(Z=z \mid U=AF^*x)}_{\text{Fidelity term}} - \underbrace{\ln \prod_{k=1}^{K} \mathsf{p}_{X^{(k)}}(x^{(k)})}_{\text{A priori}}}$$

where

$$\mathsf{P}(Z=z \mid U=AF^*x) = \frac{1}{(2\pi\alpha)^{N/2}} \exp\left\{-\frac{\|AF^*x - z\|_2^2}{2\alpha}\right\}$$

and

$$p_{X^{(k)}}(x^{(k)}) = \frac{1}{C_k}\exp\{-\lambda_k |x^{(k)}|\}$$

Consequently
$$\boxed{\min_x \; \underbrace{\frac{1}{2\alpha}\|AF^*x - z\|_2^2}_{\text{Fidelity term}} + \underbrace{\sum_{k=1}^{K} \lambda_k |x^{(k)}|}_{\text{A priori}}}$$

Variational approach: Compressed sensing

- Sparsity assumption $$\boxed{y_0 = \arg\min_{y \in \mathbb{R}^N} \|Fy\|_0 \ \text{ s.t. } \ \|Ay - z\|_2^2 \leq \varepsilon}$$
  where $(\forall y = (y_i)_{1 \leq i \leq N} \in \mathbb{R}^N)\ \|y\|_0 \equiv \#\{i : y_i \neq 0\}$ and $\varepsilon > 0$

- Convex relaxation : $$\boxed{\widehat{y} = \arg\min_{y \in \mathbb{R}^N} \|Fy\|_1 \ \text{ s.t. } \ \|Ay - z\|_2^2 \leq \varepsilon}$$
  where $(\forall y = (y_i)_{1 \leq i \leq N} \in \mathbb{R}^N)\ \|y\|_1 \equiv \sum_{i=1}^N |y_i|$

- Lagragian formulation : $$\boxed{\widehat{y} = \arg\min_{y \in \mathbb{R}^N} \|Ay - z\|_2^2 + \lambda \|Fy\|_1}$$
  where $\lambda > 0$

$\Rightarrow$ Compressed sensing [Donoho,2004][Candès,2008]

Variational approach: Criterion choice

Minimization problem

$$\text{Find} \qquad \widehat{y} \in \underset{y \in \mathbb{R}^N}{\operatorname{Argmin}} \sum_{j=1}^{J} f_j(y)$$

where $(f_j)_{1 \le j \le J}$ belong to the class of convex functions, l.s.c., and proper from $\mathbb{R}^N$ to $]-\infty, +\infty]$.

Variational approach: Criterion choice

Minimization problem

$$\text{Find} \qquad \widehat{y} \in \underset{y \in \mathbb{R}^N}{\operatorname{Argmin}} \sum_{j=1}^{J} f_j(y)$$

- f_j **may model the data fidelity term**
  - $\forall y \in \mathbb{R}^N,\ f_j(y) = \frac{1}{2\alpha}\|Ay - z\|^2$ for a Gaussian noise
  - $\forall y \in \mathbb{R}^N,\ f_j(y) = D_{KL}(\alpha Ay, z)$ for a Poisson noise

Variational approach: Criterion choice

Minimization problem

$$\text{Find} \qquad \widehat{y} \in \underset{y \in \mathbb{R}^N}{\text{Argmin}} \sum_{j=1}^{J} f_j(y)$$

- **f_j may model the data fidelity term**
  - $\forall y \in \mathbb{R}^N$, $f_j(y) = \frac{1}{2\alpha}\|Ay - z\|^2$ for a Gaussian noise
  - $\forall y \in \mathbb{R}^N$, $f_j(y) = D_{KL}(\alpha Ay, z)$ for a Poisson noise

Poisson distribution:

$$P(Z = z|U = Ay) = \prod_{i=1}^{N} \frac{\exp\big(-\alpha(Ay)_i\big)}{z_i!}(\alpha(Ty)_i)^{z_i}$$

Poisson anti-log likelihood:

$$-\log P(Z = z|U = Ay) = \sum_{i=1}^{N} -z_i \log(\alpha(Ay)_i) + \alpha(Ay)_i$$

Variational approach: Criterion choice

Minimization problem

$$\text{Find} \qquad \widehat{y} \in \underset{y \in \mathbb{R}^N}{\operatorname{Argmin}} \sum_{j=1}^{J} f_j(y)$$

- f_j **may model the data fidelity term**
  - $\forall y \in \mathbb{R}^N$, $f_j(y) = \frac{1}{2\alpha}\|Ay - z\|^2$ for a Gaussian noise
  - $\forall y \in \mathbb{R}^N$, $f_j(y) = D_{KL}(\alpha Ay, z)$ for a Poisson noise
- f_j **may model a prior**
  - $\forall y \in \mathbb{R}^N$, $f_j(y) = \lambda\|y\|_2^2$: Tikhonov regularization
  - $\forall y \in \mathbb{R}^N$, $f_j(y) = \lambda \mathrm{tv}(y)$: total variation
  - $\forall y \in \mathbb{R}^N$, $f_j(y) = \lambda\|Fy\|_1$: model sparsity

Variational approach: Criterion choice

Minimization problem

$$\text{Find} \qquad \widehat{y} \in \underset{y \in \mathbb{R}^N}{\operatorname{Argmin}} \sum_{j=1}^{J} f_j(y)$$

- **f_j may model the data fidelity term**
  - $\forall y \in \mathbb{R}^N,\ f_j(y) = \frac{1}{2\alpha}\|Ay - z\|^2$ for a Gaussian noise
  - $\forall y \in \mathbb{R}^N,\ f_j(y) = D_{KL}(\alpha Ay, z)$ for a Poisson noise
- **f_j may model a prior**
  - $\forall y \in \mathbb{R}^N,\ f_j(y) = \lambda\|y\|_2^2$: Tikhonov regularization
  - $\forall y \in \mathbb{R}^N,\ f_j(y) = \lambda \mathrm{tv}(y)$: total variation
  - $\forall y \in \mathbb{R}^N,\ f_j(y) = \lambda\|Fy\|_1$: model sparsity
- **f_j may model constraints**
  - $\forall y \in \mathbb{R}^N,\ f_j(y) = \iota_C(y)$ $(C = [0,255]^N)$: data range constraint

Variational approach: Algorithm choice

Minimization problem

$$\text{Find} \qquad \widehat{y} \in \underset{y \in \mathbb{R}^N}{\operatorname{Argmin}} \sum_{j=1}^{J} f_j(y)$$

- Properties of the involved functions
  - smooth functions
    $\rightarrow$ gradient-based methods (Newton, Quasi-Newton, . . .)
  - constraints
    $\rightarrow$ projection based methods (POCS, SIRT, . . .)
  - non-smooth functions
    $\rightarrow$ proximal algorithms (FB, DR, PPXA, . . .)

Proximity operator

Definition [Moreau (1965)]

Let $\varphi\colon \mathcal{H} \to \mathbb{R}$ be a convex, l.s.c., and proper function. The proximity operator of φ at point $u \in \mathcal{H}$ is the unique point denoted by $\mathrm{prox}_{\varphi} u$ such that

$$(\forall u \in \mathcal{H}), \qquad \boxed{\mathrm{prox}_{\varphi} u = \arg\min_{v \in \mathcal{H}} \varphi(v) + \frac{1}{2} \|u - v\|^2}$$

Examples:

- If $\varphi = \iota_C \Rightarrow \mathrm{prox}_{\iota_C} = P_C$ where P_C denotes the projection onto the convex set C.
- If $\varphi = \chi|\cdot|$ with $\chi > 0 \Rightarrow \mathrm{prox}_{\varphi}$ is the soft-thresholding with a fixed threshold χ.
- If $\varphi = D_{KL}(\alpha\cdot, \zeta)$ with $\zeta, \alpha > 0 \Rightarrow \mathrm{prox}_{\varphi} = \frac{1}{2}(\cdot - \alpha + \sqrt{|\cdot - \alpha|^2 + 4\zeta})$.

Proximity operator

Proposition [Combettes,Pesquet,2007]

- $\mathcal{H}_1$ and $\mathcal{H}_2$: real separable Hilbert spaces,
- φ: convex, l.s.c., and proper function from $\mathcal{H}_2$ to $\mathbb{R}$,
- $L\colon \mathcal{H}_1 \to \mathcal{H}_2$: bounded linear operator such that $L \circ L^* = \sigma \mathrm{Id}$ where $\sigma > 0$. Then,

$$\boxed{\mathrm{prox}_{\varphi \circ L} = \mathrm{Id} + \frac{L^*}{\sigma} \circ (\mathrm{prox}_{\sigma\varphi} - \mathrm{Id}) \circ L}.$$

Proximity operator

Proposition [Combettes,Pesquet,2007]

- $\mathcal{H}_1$ and $\mathcal{H}_2$: real separable Hilbert spaces,
- φ: convex, l.s.c., and proper function from $\mathcal{H}_2$ to $\mathbb{R}$,
- $L\colon \mathcal{H}_1 \to \mathcal{H}_2$: bounded linear operator such that $L \circ L^* = \sigma \mathrm{Id}$ where $\sigma > 0$. Then,

$$\boxed{\mathrm{prox}_{\varphi \circ L} = \mathrm{Id} + \frac{L^*}{\sigma} \circ (\mathrm{prox}_{\sigma\varphi} - \mathrm{Id}) \circ L}.$$

- If L denotes a tight frame, i.e. $L = F^*$ where $F^* \circ F = \mu \mathrm{Id}$

  $\Rightarrow$ Proposition can be used (example: $\mathrm{prox}_{D_{KL}(F^* \cdot, z)}$)

Proximity operator

Proposition [Combettes,Pesquet,2007]

- $\mathcal{H}_1$ and $\mathcal{H}_2$: real separable Hilbert spaces,
- φ: convex, l.s.c., and proper function from $\mathcal{H}_2$ to $\mathbb{R}$,
- $L\colon \mathcal{H}_1 \to \mathcal{H}_2$: bounded linear operator such that $L \circ L^* = \sigma \mathrm{Id}$ where $\sigma > 0$. Then,

$$\boxed{\mathrm{prox}_{\varphi \circ L} = \mathrm{Id} + \frac{L^*}{\sigma} \circ (\mathrm{prox}_{\sigma\varphi} - \mathrm{Id}) \circ L}.$$

- If L denotes a tight frame, i.e. $L = F^*$ where $F^* \circ F = \mu \mathrm{Id}$

  $\Rightarrow$ Proposition can be used (example: $\mathrm{prox}_{D_{KL}(F^* \cdot, z)}$)

- Extension to the case of a separable φ and a linear operator such that $L \circ L^* = D$ where D denotes a diagonal matrix [Pustelnik et al, 2011].

Proximal algorithm: Forward-backward

Forward-backward algorithm

Let $u_0 \in \mathcal{H}$.
The algorithm constructs a sequence $(u_\ell)_{\ell \geq 1}$ by the iterations

$$u_{\ell+1} = u_\ell + \lambda_\ell \big(\mathrm{prox}_{\gamma_\ell f_1}(u_\ell - \gamma_\ell \nabla f_2(u_\ell)) - u_\ell\big)$$

Proximal algorithm: Forward-backward

Forward-backward algorithm

Let $u_0 \in \mathcal{H}$.
The algorithm constructs a sequence $(u_\ell)_{\ell \geq 1}$ by the iterations

$$u_{\ell+1} = u_\ell + \lambda_\ell \big(\mathrm{prox}_{\gamma_\ell f_1} (u_\ell - \gamma_\ell \nabla f_2(u_\ell)) - u_\ell \big)$$

Forward-backward convergence [Combettes and Wajs, 2005]

- f_2 is β-Lipschitz differentiable on $\mathcal{H}$ with $\beta > 0$
- $\gamma_\ell \in]0, 2/\beta[$: algorithm step-size
- $\lambda_\ell \in]0, 1]$: relaxation parameter

Under these assumptions, $(u_\ell)_{\ell \in \mathbb{N}}$ converges to a solution of $\min f_1(u) + f_2(u)$.

Proximal algorithm: PPXA

Parallel ProXimal Algorithm [Combettes and Pesquet, 2008]

- $\gamma \in \,]0, +\infty[\,$, $(\omega_j)_{1 \le j \le J} \in]0,1]^J$ such that $\sum_{j=1}^{J} \omega_j = 1$,
- $(v_{j,0})_{1 \le j \le J} \in (\mathbb{R}^N)^J$ and $u_0 = \sum_{j=1}^{J} \omega_j v_{j,0}$,

The sequence $(x_\ell)_{\ell \ge 1}$ is generated by the iterations:

$$\begin{cases} \text{For} \quad j = 1, \ldots, J \\ \quad \lfloor p_{j,\ell} = \mathrm{prox}_{\gamma/\omega_j f_j} v_{j,\ell} \leftarrow \quad \text{compute every proximity operators} \\ p_\ell = \sum_{j=1}^{J} \omega_j p_{j,\ell} \\ \lambda_\ell \in \,]0, 2[\\ \text{For} \quad j = 1, \ldots, J \\ \quad \lfloor v_{j,\ell+1} = v_{j,\ell} + \lambda_\ell \left(2\, p_\ell - u_\ell - p_{j,\ell}\right) \\ u_{\ell+1} = u_\ell + \lambda_\ell (p_\ell - u_\ell) \end{cases}$$

Convergence [Combettes and Pesquet, 2008]

Under some assumptions, the sequence $(u_\ell)_{\ell \ge 1}$ converges to a solution of $\min \sum_{j=1}^{J} f_j(u)$

Proximal algorithms: $\widehat{y} \in \underset{y \in \mathbb{R}^N}{\operatorname{Argmin}} \sum_{j=1}^{J} f_j(L_j y)$

- **Forward-Backward** [Combettes-Wajs, 2005]

  $\rightarrow J = 2$, $L_i = \mathrm{Id}$ and a gradient Lipschitz function.

Proximal algorithms: $\widehat{y} \in \underset{y\in\mathbb{R}^N}{\operatorname{Argmin}} \sum_{j=1}^{J} f_j(L_j y)$

- **Forward-Backward** [Combettes-Wajs, 2005]

  $\rightarrow J = 2$, $L_i = \mathrm{Id}$ and a gradient Lipschitz function.
- **Douglas-Rachford** [Combettes-Pesquet, 2007]

  $\rightarrow J = 2$ and $L_i = \mathrm{Id}$.

Proximal algorithms: $\widehat{y} \in \underset{y \in \mathbb{R}^N}{\operatorname{Argmin}} \sum_{j=1}^{J} f_j(L_j y)$

- **Forward-Backward** [Combettes-Wajs, 2005]

  $\rightarrow J = 2$, $L_i = \mathrm{Id}$ and a gradient Lipschitz function.

- **Douglas-Rachford** [Combettes-Pesquet, 2007]

  $\rightarrow J = 2$ and $L_i = \mathrm{Id}$.

- **PPXA** [Combettes-Pesquet, 2008]

  $\rightarrow L_i = \mathrm{Id}$.

Proximal algorithms: $\widehat{y} \in \underset{y \in \mathbb{R}^N}{\operatorname{Argmin}} \sum_{j=1}^{J} f_j(L_j y)$

- **Forward-Backward** [Combettes-Wajs, 2005]

  $\rightarrow J = 2$, $L_i = \mathrm{Id}$ and a gradient Lipschitz function.

- **Douglas-Rachford** [Combettes-Pesquet, 2007]

  $\rightarrow J = 2$ and $L_i = \mathrm{Id}$.

- **PPXA** [Combettes-Pesquet, 2008]

  $\rightarrow L_i = \mathrm{Id}$.

- **PPXA +** [Pesquet-Pustelnik, 2012]

  $\rightarrow \sum_i L_i^* L_i$ invertible.

Proximal algorithms: $\widehat{y} \in \underset{y \in \mathbb{R}^N}{\operatorname{Argmin}} \sum_{j=1}^{J} f_j(L_j y)$

- **Forward-Backward** [Combettes-Wajs, 2005]
  $\rightarrow J = 2$, $L_i = \mathrm{Id}$ and a gradient Lipschitz function.
- **Douglas-Rachford** [Combettes-Pesquet, 2007]
  $\rightarrow J = 2$ and $L_i = \mathrm{Id}$.
- **PPXA** [Combettes-Pesquet, 2008]
  $\rightarrow L_i = \mathrm{Id}$.
- **PPXA +** [Pesquet-Pustelnik, 2012]
  $\rightarrow \sum_i L_i^* L_i$ invertible.
- **M+SFBF** [Combettes-Briceño, 2011]
  $\rightarrow$ no required conditions.

Proximal algorithms: $\widehat{y} \in \underset{y\in\mathbb{R}^N}{\text{Argmin}} \sum_{j=1}^{J} f_j(L_j y)$

- **Forward-Backward** [Combettes-Wajs, 2005]

  $\rightarrow J = 2$, $L_i = \text{Id}$ and a gradient Lipschitz function.
- **Douglas-Rachford** [Combettes-Pesquet, 2007]

  $\rightarrow J = 2$ and $L_i = \text{Id}$.
- **PPXA** [Combettes-Pesquet, 2008]

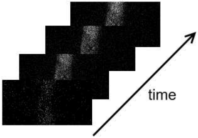
  $\rightarrow L_i = \text{Id}$.
- **PPXA +** [Pesquet-Pustelnik, 2012]

  $\rightarrow \sum_i L_i^* L_i$ invertible.
- **M+SFBF** [Combettes-Briceño, 2011]

  $\rightarrow$ no required conditions.
- **M+LFBF** [Combettes-Pesquet, 2012]

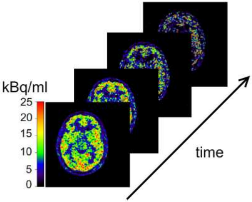
  $\rightarrow$ one function with a Lipschitz gradient.

PET reconstruction (dynamic)



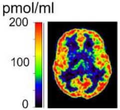
Measures

→



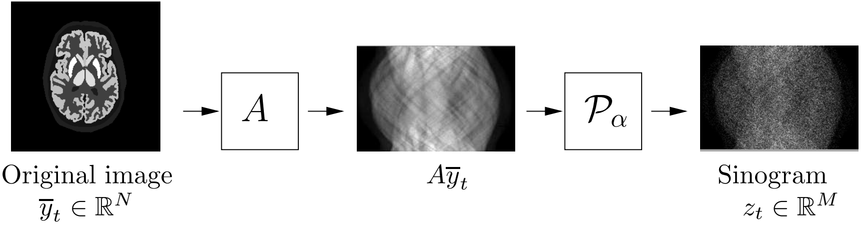
Reconstructed images

→



Benzodiazepine receptors distribution

PET reconstruction (dynamic)



Degradation model

$$\forall t \in \{1, \ldots, T\},\ z_t = \mathcal{P}_\alpha(A\bar{y}_t)$$

- $A \in \mathbb{R}^{M \times N}$: matrix associated with the projection operator.
- $\mathcal{P}_\alpha$: Poisson noise with scaling parameter $\alpha > 0$.

PET reconstruction: State-of-the-art

- Filter Back Projection: FBP
- EM-ML [Shepp, Vardi, 1982][Lange, Carson, 1984]
- Ordered Subsets EM: OSEM [Hudson, Larkin, 1994]
- Space-alternating Generalized EM: SAGE [Fessler, Hero, 1994]
- RAMLA [Browne, De Pierro, 1996]
- EM with stopping criteria [Veklerov, Llacer, 1987]
- SIEVES [Snyder, Miller, 1985]
- MAP [Herman et al., 1979], [Sauer, Bouman, 1993], [Fessler, 1995]
- Multiresolution [Turkheimer et al., 1999], [Alpert et al., 2006], [Zhou et al., 2007],
- Spatio-temporal assumptions [Nichols et al., 2002], [Kamasac et al., 2005], [Reader et al., 2006], [Verhaeghe et al., 2008]

Proposed method

Criterion

$$\hat{y} \in \operatorname*{Argmin}_{y \in \mathbb{R}^{NT}} \sum_{t=1}^{T} D_{\mathrm{KL}}(\alpha A y_t, z_t)$$

- $D_{\mathrm{KL}}(\alpha A \cdot, z)$: minus Poisson log-likelihood (data fidelity term)

Proposed method

Criterion

$$\hat{y} \in \underset{y \in \mathbb{R}^{NT}}{\operatorname{Argmin}} \sum_{t=1}^{T} D_{\mathrm{KL}}(\alpha A y_t, z_t) \qquad + \kappa \|Fy\|_1$$

- $D_{\mathrm{KL}}(\alpha A \cdot, z)$: minus Poisson log-likelihood (data fidelity term)
- $F \in \mathbb{R}^{K \times NT}$: wavelet transform such that $F^* F = F F^* = \mathrm{Id}$
- $\| \cdot \|_1$: ℓ_1-norm

Proposed method

Criterion

$$\hat{y} \in \underset{y \in \mathbb{R}^{NT}}{\operatorname{Argmin}} \sum_{t=1}^{T} \Big(D_{\mathrm{KL}}(\alpha A y_t, z_t) + \vartheta \operatorname{tv}(y_t) \Big) + \kappa \|Fy\|_1$$

- $D_{\mathrm{KL}}(\alpha A \cdot, z)$: minus Poisson log-likelihood (data fidelity term)
- $F \in \mathbb{R}^{K \times NT}$: wavelet transform such that $F^* F = F F^* = \mathrm{Id}$
- $\| \cdot \|_1$: ℓ_1-norm
- tv: total variation

Proposed method

Criterion

$$\hat{y} \in \underset{y \in \mathbb{R}^{NT}}{\operatorname{Argmin}} \sum_{t=1}^{T} \Big(D_{\mathrm{KL}}(\alpha A y_t, z_t) + \vartheta \operatorname{tv}(y_t) \Big) + \kappa \|Fy\|_1 + \iota_C(y)$$

- $D_{\mathrm{KL}}(\alpha A \cdot, z)$: minus Poisson log-likelihood (data fidelity term)
- $F \in \mathbb{R}^{K \times NT}$: wavelet transform such that $F^* F = F F^* = \mathrm{Id}$
- $\| \cdot \|_1$: ℓ_1-norm
- tv: total variation
- $C \subset \mathbb{R}^{NT}$: convex constraint

Proposed method

Criterion

$$\hat{y} \in \underset{y \in \mathbb{R}^{NT}}{\operatorname{Argmin}} \sum_{t=1}^{T} \Big(D_{\mathrm{KL}}(\alpha A y_t, z_t) + \vartheta \operatorname{tv}(y_t) \Big) + \kappa \|Fy\|_1 + \iota_C(y)$$

- $D_{\mathrm{KL}}(\alpha A \cdot, z)$: minus Poisson log-likelihood (data fidelity term)
- $F \in \mathbb{R}^{K \times NT}$: wavelet transform such that $F^* F = F F^* = \mathrm{Id}$
- $\| \cdot \|_1$: ℓ_1-norm
- tv: total variation
- $C \subset \mathbb{R}^{NT}$: convex constraint
- $\vartheta \geq 0$ and $\kappa \geq 0$

Proposed method

Criterion

$$\hat{y} \in \underset{y \in \mathbb{R}^{NT}}{\operatorname{Argmin}} \sum_{t=1}^{T} \Big(D_{\mathrm{KL}}(\alpha A y_t, z_t) + \vartheta \operatorname{tv}(y_t) \Big) + \kappa \|Fy\|_1 + \iota_C(y)$$

- $D_{\mathrm{KL}}(\alpha A \cdot, z)$: minus Poisson log-likelihood (data fidelity term)
- $F \in \mathbb{R}^{K \times NT}$: wavelet transform such that $F^* F = F F^* = \mathrm{Id}$
- $\| \cdot \|_1$: ℓ_1-norm $\leftarrow$ non differentiable
- tv: total variation
- $C \subset \mathbb{R}^{NT}$: convex constraint $\leftarrow$ non differentiable
- $\vartheta \geq 0$ and $\kappa \geq 0$

Proposed method

Criterion

$$\hat{y} \in \operatorname*{Argmin}_{y \in \mathbb{R}^{NT}} \sum_{t=1}^{T} \Big(D_{\mathrm{KL}}(\alpha A y_t, z_t) + \vartheta \mathrm{tv}(y_t) \Big) + \kappa \|Fy\|_1 + \iota_C(y)$$

- $D_{\mathrm{KL}}(\alpha A\cdot, z)$: minus Poisson log-likelihood (data fidelity term)
- $F \in \mathbb{R}^{K\times NT}$: wavelet transform such that $F^*F = FF^* = \mathrm{Id}$
- $\|\cdot\|_1$: ℓ_1-norm $\leftarrow$ non differentiable
- tv: total variation
- $C \subset \mathbb{R}^{NT}$: convex constraint $\leftarrow$ non differentiable
- $\vartheta \geq 0$ and $\kappa \geq 0$

Proposed method

Criterion

$$\hat{y} \in \operatorname*{Argmin}_{x \in \mathbb{R}^K} \sum_{t=1}^{T} \Big(D_{\mathrm{KL}}(\alpha A y_t, z_t) + \vartheta \operatorname{tv}(y_t) \Big) + \kappa \|Fy\|_1 + \iota_C(y)$$

- tv: anisotropic total variation
- $\| \cdot \|_1$: ℓ_1-norm
- $C \subset \mathbb{R}^{NT}$: convex constraint
- $D_{\mathrm{KL}}(\alpha A \cdot, z)$: Poisson anti-log likelihood

Proposed method

Criterion

$$\hat{y} \in \underset{x \in \mathbb{R}^K}{\operatorname{Argmin}} \sum_{t=1}^{T} \Big(D_{\mathrm{KL}}(\alpha A y_t, z_t) + \vartheta \mathrm{tv}(y_t) \Big) + \kappa \|Fy\|_1 + \iota_C(y)$$

- tv: anisotropic total variation $\rightarrow$ Closed form for $\mathrm{prox}_{\mathrm{tv}}$ after splitting
- $\|\cdot\|_1$: ℓ_1-norm $\rightarrow$ Closed form for $\mathrm{prox}_{\|\cdot\|_1}$
- $C \subset \mathbb{R}^{NT}$: convex constraint $\rightarrow$ Closed form for prox_{ι_C}
- $D_{\mathrm{KL}}(\alpha A \cdot, z)$: Poisson anti-log likelihood

Proposed method

Criterion

$$\hat{y} \in \underset{x \in \mathbb{R}^K}{\operatorname{Argmin}} \sum_{t=1}^{T} \Big(D_{\mathrm{KL}}(\alpha A y_t, z_t) + \vartheta \mathrm{tv}(y_t) \Big) + \kappa \|Fy\|_1 + \iota_C(y)$$

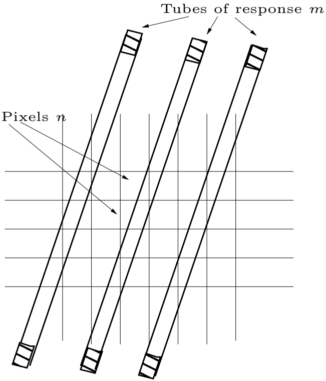
- tv: anisotropic total variation $\rightarrow$ Closed form for $\mathrm{prox}_{\mathrm{tv}}$ after splitting
- $\|\cdot\|_1$: ℓ_1-norm $\rightarrow$ Closed form for $\mathrm{prox}_{\|\cdot\|_1}$
- $C \subset \mathbb{R}^{NT}$: convex constraint $\rightarrow$ Closed form for prox_{ι_C}
- $D_{\mathrm{KL}}(\alpha A \cdot, z)$: Poisson anti-log likelihood

  $\rightarrow$ closed form for $\mathrm{prox}_{D_{\mathrm{KL}}}(\alpha \cdot, z)$,

  $\Rightarrow$ NO closed form for $\mathrm{prox}_{D_{\mathrm{KL}}(\alpha A \cdot, z)}$ because $A \circ A^* \neq D$

  (where D denote a diagonal matrix)

Proposed method

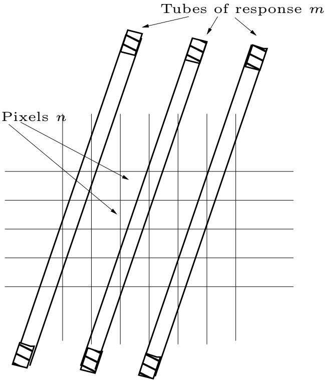


How to compute $\text{prox}_{D_{\text{KL}}(\alpha A\cdot,z)}$**?**

- For every $m \in \{1,\ldots,M\}$, $\Psi_m \in \Gamma_0(\mathbb{R})$
- We can write

$$D_{\text{KL}}(\alpha Ay, z) = \sum_{m=1}^{M} \Psi_m(\alpha A_{m,\bullet}\, y)$$

Proposed method



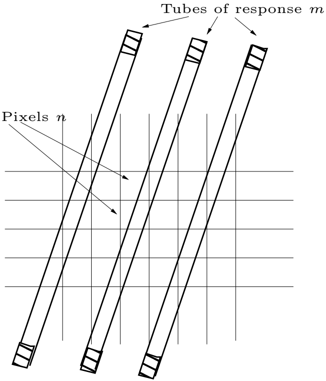
How to compute $\mathrm{prox}_{D_{\mathrm{KL}}(\alpha A\cdot,z)}$**?**

- $\mathbb{I}_r$ is a partition of $\{1,\ldots,M\}$
- For every $m \in \{1,\ldots,M\}$, $\Psi_m \in \Gamma_0(\mathbb{R})$
- We can write

$$
\begin{aligned}
D_{\mathrm{KL}}(\alpha A y, z) &= \sum_{m=1}^{M} \Psi_m(\alpha A_{m,\bullet} y) \\
&= \sum_{r=1}^{R} \sum_{m \in \mathbb{I}_r} \Psi(\alpha A_{m,\bullet} y)
\end{aligned}
$$

- Choice of $\mathbb{I}_r$

Proposed method



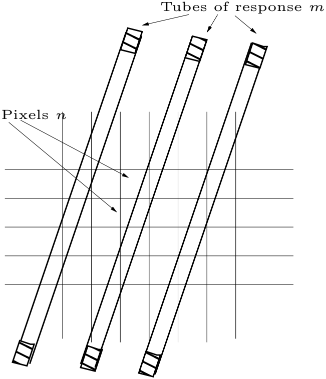
How to compute $\operatorname{prox}_{D_{\mathrm{KL}}(\alpha A\cdot, z)}$**?**

- $\mathbb{I}_r$ is a partition of $\{1,\ldots,M\}$
- For every $m \in \{1,\ldots,M\}$, $\Psi_m \in \Gamma_0(\mathbb{R})$
- We can write

$$
\begin{aligned}
D_{\mathrm{KL}}(\alpha A y, z) &= \sum_{m=1}^{M} \Psi_m(\alpha A_{m,\bullet} y) \\
&= \sum_{r=1}^{R} \sum_{m\in\mathbb{I}_r} \Psi(\alpha A_{m,\bullet} y) \\
&= \sum_{r=1}^{R} \Psi^{(r)}(\alpha A^{(r)} y)
\end{aligned}
$$

- Choice of $\mathbb{I}_r$: $A^{(r)} = (A_{m,n})_{m\in\mathbb{I}_r, 1\le n\le N}$?

Proposed method



How to compute $\mathrm{prox}_{D_{\mathrm{KL}}(\alpha A\cdot, z)}$**?**

- For every $m \in \{1, \ldots, M\}$, $\Psi_m \in \Gamma_0(\mathbb{R})$
- We can write

$$
\begin{aligned}
D_{\mathrm{KL}}(\alpha A y, z) &= \sum_{m=1}^{M} \Psi_m(\alpha A_{m,\bullet} y) \\
&= \sum_{r=1}^{R} \sum_{m \in \mathbb{I}_r} \Psi(\alpha A_{m,\bullet} y) \\
&= \sum_{r=1}^{R} \Psi^{(r)}(\alpha A^{(r)} y)
\end{aligned}
$$

- Choice of $\mathbb{I}_r$: $A^{(r)} = (A_{m,n})_{m \in \mathbb{I}_r, 1 \le n \le N}$? containing non-overlapping and thus orthogonal rows of A $\Rightarrow$ $\mathrm{prox}_{\Psi^{(r)} \circ A^{(r)}}$ takes a closed form (i.e. $A^{(r)} \circ (A^{(r)})* \neq D$.

Simulated data: Materials & methods

Simulated data - Materials

- PET 2D+time cerebral data $[^{18}F]$-FDG – HR+ scanner, based on Zubal cerebral phantom.
- 16 temporal frames, time: between 30 secondes and 5 minutes.
- Events number: between 48 and 26804.

Simulated data: Materials & methods

Simulated data - Materials

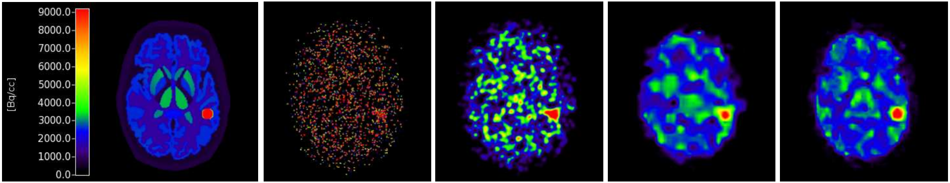
- PET 2D+time cerebral data [^{18}F]-FDG – HR+ scanner, based on Zubal cerebral phantom.
- 16 temporal frames, time: between 30 secondes and 5 minutes.
- Events number: between 48 and 26804.

Method

- PPXA: 400 iterations – 10 s each.
- System matrix split is $R = 432$ subsets.
- Wavelets: symmlet filters length 6 on 3 levels (space) and 2 levels of Daubechies-6 on interval (time).
- EM-ML with/without post-reconstruction smoothing, PPXA with/without temporal regularisation.

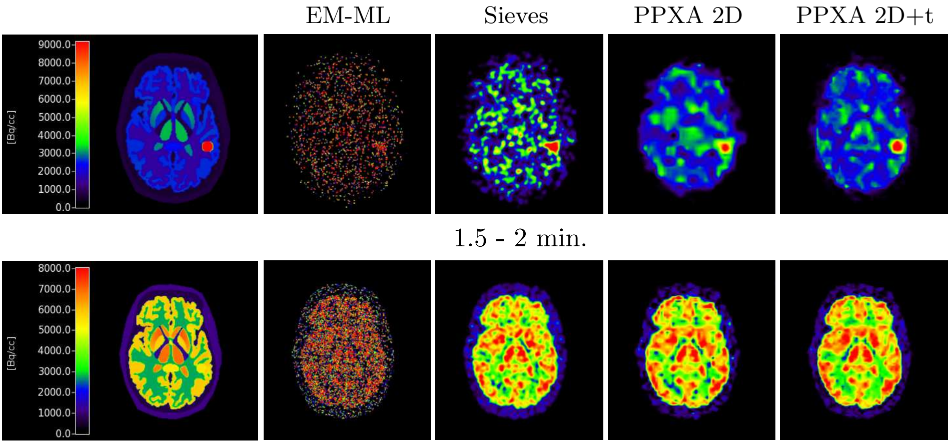
Simulated data: Results

EM-ML Sieves PPXA 2D PPXA 2D+t



1.5 - 2 min.

Simulated data: Results

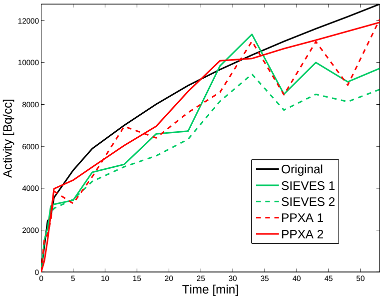


8 - 13 min.

Simulated data: MSE

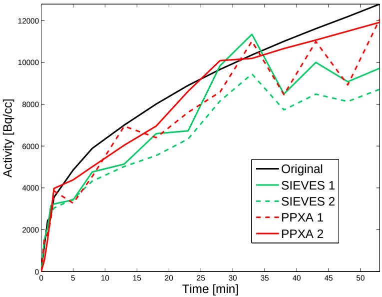
Cortex				
Frame [min]	EM-ML	SIEVES	PPXA 2D	PPXA 2D+t
1.5 - 2	22.4	0.52	0.19	**0.12**
8 - 13	3.49	0.13	0.14	**0.12**
43 - 48	2.54	**0.16**	0.18	**0.16**
Striatum				
Frame [min]	EM-ML	SIEVES	PPXA 2D	PPXA 2D+t
1.5 - 2	18.6	0.39	0.11	**0.08**
8 - 13	3.05	**0.10**	**0.10**	**0.10**
43 - 48	2.23	**0.10**	**0.10**	**0.10**

Simulated data: TAC

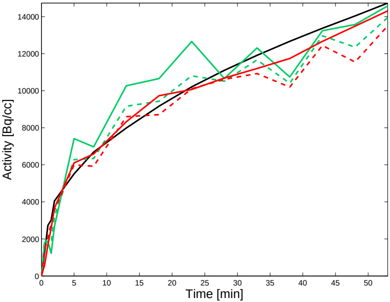


Cortex

Simulated data: TAC



Cortex



Striatum

Preclinical data: materials & methods

- Baboon 2D brain PET [^{18}F]-FDG study on the HR+ scanner
- Prompt and delayeds stored separatly
- 128 time frames, from 10 to 30 second duration

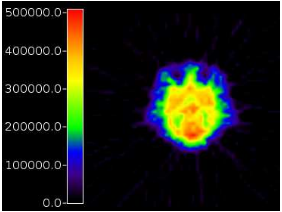
Preclinical data: materials & methods

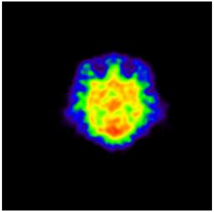
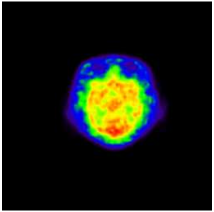
- Baboon 2D brain PET [^{18}F]-FDG study on the HR+ scanner
- Prompt and delayeds stored separatly
- 128 time frames, from 10 to 30 second duration

- Sinograms rebinned into 16 time frames, from 80 to 240 second duration
- OP-PPXA: all corrections (except scatters) included in the model

Preclinical data: results

FBP Sieves PPXA 2D+t

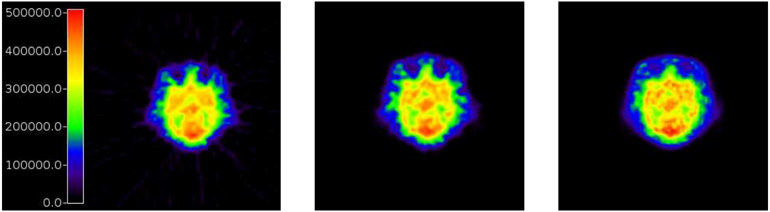


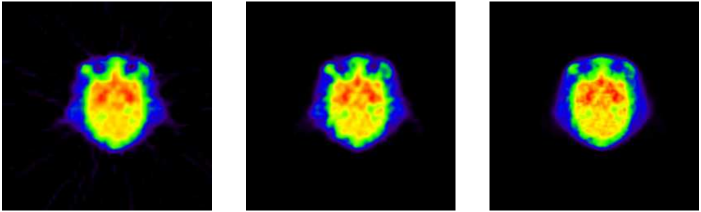
4 - 6 min.

Preclinical data: results

FBP Sieves PPXA 2D+t



4 - 6 min.



42 - 46 min.

Conclusion

- Efficient method to reconstruct dynamic PET data and obtain parametric map
- Taking into account time-frame characteristics by considering **wavelets**
- Using **Parallel** ProXimal Algorithm
- Possibility to deal with a large panel of regularization terms and constraints
- Results presented on simulated and preclinical data

Future work:

- Impact of considering 16, 32, 64,... temporal frames for parametric map estimation
- Improve the convergence rate of the algorithm

Reference

- J. Verhaegue, D. Van De Ville, I. Khalidov, Y. D'Asseler, I. Lemahieu, and M. Unser. "Dynamic PET reconstruction using wavelet regularization with adapted basis functions." IEEE TMI, 27(7):943–959, 2008.
- F. C. Sureau, J.-C. Pesquet, C. Chaux, N. Pustelnik, A. J. Reader, C. Comtat, R. Trebossen, "Temporal wavelet denoising of PET sinograms and images." IEEE NSS, 6 pp., Dresden, Germany, Oct. 19-25 2008.
- N. Pustelnik, C. Chaux, J.-C. Pesquet, F. C. Sureau, E. Dusch and C. Comtat, "Adapted Convex Optimization Algorithm for Wavelet-Based Dynamic PET Reconstruction," Fully 3D, 4 pp., Bejing, Chine, Sept 5–10 2009.
- N. Pustelnik, C. Chaux, J.-C. Pesquet, and C. Comtat, "Parallel Algorithm and Hybrid Regularization for Dynamic PET Reconstruction," IEEE MIC, 4 pp., Knoxville, Tennessee, Oct. 30 - Nov. 6 2010.
- N. Pustelnik, C. Chaux and J.-C. Pesquet, "Parallel ProXimal Algorithm for image restoration using hybrid regularization," IEEE TIP, 20(9):2450–2462, Sep. 2011.
- S. Anthoine, J.-F. Aujol, Y. Boursier, and C. Mélot, "Some proximal methods for Poisson intensity CBCT and PET", Inverse Problems and Imaging, to appear, 2012.