Parallel Algorithm and Hybrid Regularization for Dynamic PET Reconstruction
Nelly Pustelnik, Caroline Chaux, Jean-Christophe Pesquet, Claude Comtat

To cite this version:
Nelly Pustelnik, Caroline Chaux, Jean-Christophe Pesquet, Claude Comtat. Parallel Algorithm and Hybrid Regularization for Dynamic PET Reconstruction. IEEE Medical Imaging Conference, Oct 2010, Knoxville, Tennessee, United States. pp.2423–2427. hal-00733493

HAL Id: hal-00733493
https://hal-upec-upem.archives-ouvertes.fr/hal-00733493
Submitted on 18 Sep 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Parallel algorithm and hybrid regularization for dynamic PET reconstruction

N. Pustelnik, Student Member, IEEE, C. Chaux, Member, IEEE, J.-C. Pesquet, Senior Member, IEEE, and C. Comtat, Member, IEEE

Abstract—To improve the estimation at the voxel level in dynamic Positron Emission Tomography (PET) imaging, we propose to develop a convex optimization approach based on a recently proposed parallel proximal method (PPXA). This class of algorithms was successfully employed for 2D deconvolution in the presence of Poisson noise and it is extended here to (dynamic) space + time PET image reconstruction. Hybrid regularization defined as a sum of a total variation and a sparsity measure is considered in this paper. The total variation is applied to each temporal-frame and a wavelet regularization is considered for the space+time data. Total variation allows us to smooth the wavelet artifacts introduced when the wavelet regularization is used alone. The proposed algorithm was evaluated on simulated dynamic fluorescent oxygen glucose (FDG) brain data and compared with a regularized Expectation Maximization (EM) reconstruction. From the reconstructed dynamic images, parametric maps of the cerebral metabolic rate of glucose (CMRglu) were computed. Our approach shows a better reconstruction at the voxel level.

I. INTRODUCTION

The extraction of physiological parameters at the voxel level (parametric imaging) is a challenging problem in dynamic PET imaging (space+time). In [1], [2], [3] the authors demonstrated the advantage of taking into account time-frame characteristics during the reconstruction for improving the image quality and the physiological parameter estimation. More recently, the space+time information was used by considering wavelet-frame decomposition [3], [4], [5]. To efficiently deal with such representations, the criterion to minimize often involves a \( \ell_1 \)-norm (regularization term) which leads to a non-differentiable problem. New convex optimization algorithms (within the class of proximal methods) have been developed in [6], [7] aiming at minimizing criteria which are non-necessarily differentiable. The first application combining wavelet-frame representations and proximal algorithms in dynamic PET image reconstruction was suggested by Verhaeghe et al. [3] by considering Forward-Backward (FB) iterations [6] to minimize a criterion involving a Gaussian data fidelity term and a wavelet spline regularization. In [5], the authors performed Temporal Activity Curves (TAC) denoising in the sinogram space before reconstruction by using the Douglas-Rachford (DR) algorithm so as to deal with a Kullback-Leibler divergence as a data fidelity term, which is well-adapted to Poisson noise. To perform simultaneously denoising and reconstruction with wavelet regularization, we previously designed in [8] a nested iterative algorithm (FB-DR algorithm). This method allowed us to improve the reconstruction quality, but its two main drawbacks were (i) a high computational time due to the required subiteration and (ii) the presence of artifacts introduced by wavelet regularization. To tackle these problems, we consider in this paper a more recent algorithm called Parallel ProXimal Algorithm (PPXA) [9] which is well-adapted to dynamic PET due to (i) its parallel structure and (ii) the possibility of dealing with hybrid regularization such as the combination of a wavelet-frame penalty and a total variation measure [10]. This method was presented for image deblurring applications in the presence of Poisson noise [11] and its application to dynamic PET is now investigated.

This paper is organised as follows. First, we will present the considered degradation model and the criterion we aim at minimizing. Then, convex optimization tools such as the proximity operator and PPXA will be introduced. The main technical difficulty in this work consists of computing the proximity operator of the data fidelity term. A solution to overcome this difficulty will be proposed. Finally, numerical experiments will be provided by considering simulated and real dynamic PET \(^{18}\text{F}\)-FDG exams.

II. PROBLEM

We consider a finite parameterization of the original image denoted by \( \tilde{y}_t = (\tilde{y}_{i,t})_{1 \leq i \leq N} \) which corresponds to the spatial activity distribution for a time-frame \( t \). These data are observed through the linear system associated with the matrix \( A = (A_{j,i})_{1 \leq j \leq M, 1 \leq i \leq N} \) where each element \( A_{j,i} \) represents the probability for a voxel \( i \) to be detected in the tube of response (TOR) \( j \). Moreover, during the acquisition process, data are contaminated by Poisson noise. The effect of the noise is denoted by \( \mathcal{P} \). The dynamic PET data \( z_t = (z_{j,t})_{1 \leq j \leq M} \) are related to the original images by the following degradation model:

\[
(\forall t \in \{1, \ldots, T\}) \quad z_t = \mathcal{P}(A\tilde{y}_t) \tag{1}
\]

where \( T \) denotes the number of time-frame. The reconstruction process requires to find the finite parameterization \( \hat{y} = (\hat{y}_t)_{1 \leq t \leq T} \) closest to \( \tilde{y} = (\tilde{y}_t)_{1 \leq t \leq T} \) from the measured data \( z = (z_t)_{1 \leq t \leq T} \).

In order to perform this task, we consider a variational formulation where the criterion is made up of a data fidelity term related to Poisson noise, a hybrid regularization term
(total variation + sparsity term in the wavelet-frame domain) and a dynamic range constraint.

The most suited data fidelity term in the presence of Poisson noise is the generalized Kullback-Leibler divergence (related to the Poisson minus log-likelihood) denoted in the following by $D_{KL}(z_t, \cdot)$. Concerning the regularization term, the total variation (denoted by $tv$) is applied to each time-frame and the wavelet-frame regularization deals simultaneously with the whole space-time data. The total variation allows us to smooth the artifacts introduced by wavelet-frame regularization ($\ell_1$-norm). In the following, $F^* \in \mathbb{R}^{N_T \times L}$ denotes a general tight wavelet-frame synthesis operator such that $F^* \circ F = \nu \text{Id}$ for some $\nu \in [0, +\infty[$. Thus, the image $\hat{y}$ can be expressed as $\hat{y} = F^* \mathcal{T} = (F_i^* \mathcal{T})_{1 \leq t \leq T}$, where $\mathcal{T} \in \mathbb{R}^L$ denotes the wavelet-frame coefficients and $F_i^* \in \mathbb{R}^{N_T \times L}$. The convex constraint $C \subset \mathbb{R}^{NT}$ limits the data value range of the solution (for example, it may be a positivity constraint). Below, $\iota_C$ denotes the indicator function of $C$ such that, for every $y \in \mathbb{R}^{NT}$, $\iota_C(y) = 0$ if $y \in C$ and $+\infty$ otherwise. The resulting minimization problem amounts to finding the wavelet-frame coefficients such that

$$\hat{x} \in \text{Argmin}_{x \in \mathbb{R}^L} \sum_{t=1}^{T} (D_{KL}(z_t, AF_i^* x) + \vartheta tv(F_i^* x)) + \kappa \|x\|_1 + \iota_C(F^* x). \tag{2}$$

$\vartheta > 0$ and $\kappa > 0$ denote the regularization parameters. They allow us to adjust the impact of the total variation and sparsity penalties on the solution. The reconstructed image is $\hat{y} = F^* \hat{x}$. Due to the definition of each functional, this criterion is convex but non-differentiable. To solve efficiently such a minimization problem, we will consider recent convex optimization tools.

### III. Convex Optimization

One of the most famous convex convex optimization algorithm is the Projection Onto Convex Sets (POCS) algorithm [12] based on alternated projections. A generalization of projections was proposed by Moreau in [13], which is known as the proximity operator. This tool is defined for a convex, lower semi-continuous and proper function $\varphi : \mathbb{R}^X \rightarrow ]-\infty, +\infty]$, at a point $v \in \mathbb{R}^X$ and it is such that

$$\text{prox}_{\varphi} : \mathbb{R}^X \rightarrow \mathbb{R}^X : v \mapsto \arg \min_{u \in \mathbb{R}^X} \frac{1}{2} \|u - v\|^2 + \varphi(u).$$

Particularly, if $C$ is a nonempty closed convex set of $\mathbb{R}^X$, $\text{prox}_{\iota_C}$ reduces to the projection $PC \circ \iota_C$.

In the recent literature, proximal algorithms (that is based on proximity operators) have been proposed to minimize efficiently non-differentiable convex criteria. One of the most famous ones is the FB algorithm [14] also known as thresholded Landweber or iterative soft-thresholding [6] when a $\ell_1+\ell_2$ criterion is considered. More generally, the FB algorithm allows us to minimize a sum of two functions one of which is Lipschitz differentiable. This framework is however too restrictive to solve the minimization problem (2) due to the non-Lipschitz differentiability of $D_{KL}$ [7]. Another proximal algorithm known as PPXA [9] allows us to relax the Lipschitz assumption and to solve any minimization problem formulated as

$$\hat{x} \in \text{Argmin}_{x \in \mathbb{R}^L} \sum_{q=1}^{Q} f_q(x) \tag{3}$$

where $Q \in \mathbb{N}^+$ and where $(f_q)_{1 \leq q \leq Q}$ are convex, lower semi-continuous and proper functions from $\mathbb{R}^L$ to $]-\infty, +\infty]$. The iterations are recalled in Algorithm 1.

**Algorithm 1**: General form of PPXA

1. Set $\gamma \in [0, +\infty[$.
2. For every $q \in \{1, \ldots, Q\}$, set $(\omega_q)_{1 \leq q \leq Q} \in [0, 1]^Q$ such that $\sum_{q=1}^{Q} \omega_q = 1$.
3. Set $(u_{q,0})_{1 \leq q \leq Q} \in (\mathbb{R}^L)^Q$ and $x_0 = \sum_{q=1}^{Q} \omega_q u_{q,0}$.
4. For $n = 0, 1, \ldots$
   4a) For $q = 1, \ldots, Q$
       - $p_{q,n} = \text{prox}_{\omega_q \gamma f_q} u_{q,n}$
   4b) $p_n = \sum_{q=1}^{Q} \omega_q p_{q,n}$
   4c) Set $\lambda_n \in [0, 2[$
   4d) For $q = 1, \ldots, Q$
       - $u_{q,n+1} = u_{q,n} + \lambda_n (2p_n - x_n - p_{q,n})$
   4e) $x_{n+1} = x_n + \lambda_n (p_n - x_n)$

The convergence of the sequence $(x_n)_{n \in \mathbb{N}}$ generated by Algorithm 1 to a solution $\hat{x}$ of (3) is established under the following assumption.

**Assumption III.1**

(i) $\lim_{\|x\| \rightarrow +\infty} f_1(x) + \cdots + f_Q(x) = +\infty$.
(ii) $\bigcap_{q=1}^{Q} \text{rint dom } f_q \neq \emptyset$.
(iii) $\sum_{n \in \mathbb{N}} \lambda_n (2 - \lambda_n) = +\infty$.

The main difficulty of this algorithm lies in Step 4a) which requires to compute the proximity operator of each function $f_q$. Fortunately, for a large class of functions, this operator takes a closed form [14]. For example, the $\ell_1$-norm proximity operator corresponds to a soft-thresholding. The way to proceed with $tv$ is detailed in [9]: it results that $tv$ can be split in a sum of four functions, denoted by $tv_j$ when Haar filters are considered. However, a remaining difficulty in using this algorithm for dynamic PET reconstruction is to obtain an explicit form at Step 4a) for $D_{KL}(z_t, A)$. Subsequently, we assume that for every $r \in \{1, \ldots, R\}$, $I_r$ is a partition of $\{1, \ldots, M\}$ in nonempty sets. Additionally, we can write $D_{KL}(z_t, \cdot)$ as a sum of $M$ functions $\Psi_j,t : \mathbb{R} \rightarrow ]-\infty, +\infty]$ operating on each $j$-th component $z_{j,t}$ of $z_t$. By regrouping the terms in this sum such that $j \in I_r$, for every $r \in \{1, \ldots, R\}$, $\Psi_r((z_{j,t})_{z \in I_r}) = \sum_{j \in I_r} \Psi_j,t(z_{j,t})$. $D_{KL}(z_t, A)$ can be decomposed as a sum of $R$ functions involving reduced-size linear operators $A^{(r)} = (A_{j,i})_{i \in I_r, 1 \leq j \leq N}$ containing non-overlapping and thus orthogonal rows of $A$. For every $r \in \{1, \ldots, R\}$, $I_r$ denotes a set of TOR involving distinct pixels. The resulting PPXA iterations are detailed in Algorithm 2.

1. Let $\varphi : \mathbb{R}^X \rightarrow ]-\infty, +\infty]$ be a convex, l.s.c. and proper function. dom $\varphi = \{u \in \mathbb{R}^X \mid \varphi(u) < +\infty\}$ and the interior of $C \subset \mathbb{R}^X$ relative to its affine hull is the set denoted by rint $C$.
Algorithm 2 PPXA for dynamic PET reconstruction

1) Set $\gamma \in ]0, +\infty[$.
2) Set $\omega_1 = \ldots = \omega_R = \frac{1}{2\pi}$, $\omega_{R+1} = \ldots = \omega_{R+6} = \frac{1}{16}$ and $\omega_{R+5} = \omega_{R+6} = \frac{1}{9}$.
3) Set, for every $q \in \{1, \ldots, R + 6\}$, $u_{q,0} = F z$
   and $x_0 = \sum_{q=1}^{R+6} \omega_q u_{q,0}$.
4) For $n = 0, 1, \ldots$
   4a) For $t = 1, \ldots, T$
      \[
      \begin{align*}
      p_{r,t,n} &= \text{prox}_{q \omega_t} \Psi_t A_t F_r^*(u_{r,n}) \\
      p_{R+1,t,n} &= \text{prox}_{q \omega_t} \frac{1}{2} \omega_t \psi_{t+1} F_r^*(u_{r+1,n}) \\
      p_{R+2,t,n} &= \text{prox}_{q \omega_t} \frac{1}{2} \omega_t \psi_{t+2} F_r^*(u_{r+2,n}) \\
      p_{R+3,t,n} &= \text{prox}_{q \omega_t} \frac{1}{2} \omega_t \psi_{t+3} F_r^*(u_{r+3,n}) \\
      p_{R+4,t,n} &= \text{prox}_{q \omega_t} \frac{1}{2} \omega_t \psi_{t+4} F_r^*(u_{r+4,n}) \\
      p_{R+5,n} &= \text{prox}_{q \omega_t} \frac{1}{2} \omega_t \psi_{t+5} F_r^*(u_{r+5,n}) \\
      p_{R+6,n} &= \text{prox}_{q \omega_t} \frac{1}{2} \omega_t \psi_{t+6} F_r^*(u_{r+6,n})
      \end{align*}
      \]
   4b) $p_n = \sum_{q=1}^{R+6} \omega_q p_{q,n}$.
   4c) Set $\lambda_n \in ]0, 2[$
   4d) For $q = 1, \ldots, Q$
      \[
      \begin{align*}
      u_{q,n+1} &= u_{q,n} + \lambda_n \left(2 p_n - x_n - p_{q,n}\right) \\
      x_{n+1} &= x_n + \lambda_n \left(p_n - x_n\right)
      \end{align*}
      \]

IV. MATERIALS AND METHODS

A. Simulated data

The method was evaluated on simulated 2D+time data, using one slice of the Zubal brain phantom [15]. A two
compartment model with $\lambda_4 = 0$ was used to simulate a
dynamic $[18F]$-FDG brain study, with 16 time frames
with a duration varying between 30 seconds for the first
time-frames to 5 minutes for the last ones. Moreover, small local variations of $K_c$ were created in the cortex and putamen
(see Figure 2, top-left). 288 (radial) $\times$ 144 (angles) sinograms
with a radial sampling of 2.247 mm were simulated by
analytically projecting the phantom and adding Poisson noise.
Attenuation, random and scattered coincidences were not
simulated in this example. The number of events varies from 48
for the first time-frame to 26804 for the last time-frame.

The PPXA algorithm is implemented in parallel with the
OpenMP library and run over 400 iterations of 10s each.
The relaxation parameter is here equal to $\lambda_n \equiv 1$ and the step-
size is $\gamma = 6 \times 10^5$. The system matrix is split in $R = 432$
subsets and the function number is $Q = R + 6$. The chosen separable orthonormal wavelets correspond to symmlet filters
of length 6 over 3 resolution levels of spatial decomposition.
PPXA 1 refers to our approach without temporal regularization
and PPXA 2 considers a two-level decomposition with
Daubechies-6 wavelets on the interval for temporal decomposi-
tion [16]. The latter choice is motivated by the small number
of time-frames and aims at avoiding boundary effects. The
constraint on the data range is $C = [0, 10^5]^{NT}$ Bq/cc and
the chosen regularization parameters are $\kappa = 0.1$ and $\vartheta = 10^{-5}$.
A limited value of the parameter $\vartheta$ was chosen so as to
avoid “cartoon effects”. The proposed solution without/with
temporal regularization in PPXA 1/PPXA 2 is compared with
EM-ML and EM-ML with post-smoothing (Sieves). SIEVES 1
denotes a Gaussian filtering optimized to yield the minimum
MSE and SIEVES 2 uses a Gaussian kernel leading to a similar
noise level than our approach. For EM-ML and SIEVES 1/2,
the iteration number is 250. A full width at half maximum (FWHM)
of 12 mm $\times$ 12 mm (resp. 14 mm $\times$ 14 mm) is used for
SIEVES 1 (resp. SIEVES 2).

B. Pre-clinical data

We also processed real data resulting from a $[18F]$-FDG
exam on the ecat HR+ scanner (Siemens medical solutions,
Knoxville, TN, USA). 284 MBq were injected to a 4.5 kg
baboon. The injection was done under camera. A dynamic
acquisition was performed with septa extended. The prompt
and delayed coincidences were registered separately in 16
frames with a duration varying from 1.5 minutes to 4 minutes.
The global exam duration was 54 minutes. Transmission
was realized before injection by a rotative source of germanium 57.

We reconstructed prompt data by considering three different
approaches: filtered backprojection (FBP), EM-ML recon-
struction with Gaussian (12 mm $\times$12 mm FWHM) post-
filtering (Sieves) and our approach (PPXA). During the re-
construction process, attenuation effects, normalization effects,
and random coincidences have been taken into account. How-
ever, scattered coincidences have not been considered (this
choice is justified by the septa extended). The system matrix
used for SIEVES and PPXA does not model resolution (PSF)
effects.

V. RESULTS

A. Simulated data

Figures 1 and 2 display the reconstruction results for diffe-
rent temporal-frames ($t = 4$: 1.5-2 min and $t = 7$: 8-13 min).
During first frames when the noise level is particularly high,
PPXA 2 restores quite well the brain activity even if artifac-
tes can be observed (for example in the thalamus area).
Table I compares the normalized MSE for different areas
of interest such as the cortex, the thalamus and the striatum. It
can be observed that PPXA 2 leads to the smallest normalized
MSE for the first frames and equivalent normalized MSE for
the last ones.

Figures 3, 4, and 5 represent Temporal Activity Curves for
individual pixels in different area of interest. This allows us
to evaluate the impact of the spatio-temporal reconstruction
and the good results obtained with PPXA 2.

Using Pmod software (Pmod Technologies Ltd, Zürich, CH),
parametric images of CMRglu were computed from the EM-
ML, SIEVES, and PPXA reconstructions. The results are shown
in Figure 6. The differences between the CMRglu images is
not as much important as the TAC results or the reconstructed
images. This is mainly due to the Patlak model which essen-
tially uses the last frames for CMRglu computation, which are
less sensitive to noise.
Fig. 1. Reconstructed images (in Bq/cc) for $t = 4$.

Fig. 2. Reconstructed images (in Bq/cc) for $t = 7$.

Fig. 3. TAC in the thalamus.

Fig. 4. TAC in the cortex.

Fig. 5. TAC in the striatum.
B. Pre-clinical data

Figure 7 shows the reconstructed images obtained with the different reconstruction methods for temporal-frames $t = 4$ (4-6 min), $t = 7$ (14-18 min), and $t = 14$ (42-46 min). These results illustrate the good behaviour and the quantitative correctness of the proposed method on real data.

VI. CONCLUSION

We have presented a convex optimization approach to deal with spatio-temporal PET reconstruction. The minimized criterion is composed with different terms: generalized Kullback-Leibler divergence (minus Poisson log-likelihood) term, $\ell_1$-norm for wavelet-based regularization, total variation and dynamic range constraint. To minimize efficiently this criterion, we employed PPXA which is a recent algorithm belonging to the class of proximal methods. The convergence of this algorithm is guaranteed and it is implementable in parallel.

REFERENCES


