2013 IEEE 10th International Symposium on Biomedical Imaging:
From Nano to Macro
San Francisco, CA, USA, April 7-11, 2013

GLOBALLY CONVERGENT 3D DYNAMIC PET RECONSTRUCTION WITH PATCH-BASED NON-CONVEX LOW RANK REGULARIZATION

K. S. Kim [a], Y. D. Son[b], Z. H. Cho[b], J. B. Ra[c] and J. C. Ye[a,]*

[a]Bio Imaging & Signal Processing Lab., Dept. of Bio& Brain Engineering, KAIST
[b] Neuroscience Research Institute, Gachon University of Medicine and Science
[c]Image Systems Lab., Electrical Engineering, KAIST

ABSTRACT

Dynamic positron emission tomography (PET) is widely used to measure variations of radiopharmaceuticals within the organs over time. However, conventional reconstruction algorithm can produce a noisy reconstruction if there are not sufficient photon counts. Hence, the main goal of this paper is to develop a novel spatio-temporal regularization approach that exploits inherent similarities within intra- and inter-frames to overcome the limitation. One of the main contributions of this paper is to demonstrate that such correlations can be exploited using a low rank constraint of overlapping similarity blocks. The resulting optimization framework is, however, non-smooth and non Lipschitz due to the low-rank penalty terms and Poisson log-likelihood. Therefore, we propose a novel globally convergent optimization method using the concave-convex procedure (CCCP) by exploiting Legendre-Fenchel transform, which overcomes the memory and computational limitations. We confirm that the proposed algorithm can provide significantly improved image quality and extract accurate kinetic parameters.

Index Terms— Dynamic PET reconstruction, patch, low-rank, concave-convex procedure, convex conjugate functions, Legendre-Fenchel transform

1. INTRODUCTION

Quantification of spatial and temporal radiotracer distribution is one of the important topics in dynamic PET studies [1]. Typically, the time activity curves (TACs) of concentration and kinetic parameters are obtained from the reconstructed images from time frames. For the case of short acquisition time, each time frame data has low photon counts, which results in very noisy reconstruction. To enhance the image quality from such photon limited data, various dynamic reconstruction algorithms have been proposed [1]. By extending the existing researches, we are interested in developing a novel spatio-temporal regularization scheme that exploits correlation structures within each intra and inter frames. In particular, this paper exploits geometric similarities in intra and inter frames. We found that a row rank constraint, which is originated from matrix completion problem in compressed sensing [2], is very useful to exploit self-similarities since the matrix rank is less sensitive to the intensity offset and is easier to capture edges, etc. In addition, we propose an overlapping patch based non-convex low rank penalty to exploit geometric self similarities. However, there are several technical challenges. First, the self-similarity structures are unknown before the final reconstruction is obtained. To address the first issue, this paper proposes a re-fineable patch search method, which iteratively refines the similarity blocks during reconstruction. Second, non-convex patch based low rank penalty term is non-smooth and the gradient of Poisson log-likelihood is non-Lipschitz, so optimization problem is non-trivial. Direct application of conventional Poisson image deconvolution by augmented Lagrangian (PIDAL) [3] algorithm is, however, problematic due to a large memory requirements, which prohibit its parallelization. To deal with this issue, the concave-convex procedure (CCCP) [4] is employed to convexify the concave rank prior and to make a EM-type separable likelihood function. Interestingly, the resulting subprogram have a pixel-by-pixel close form expression, which allows the algorithm converge fast, and the graphic process unit (GPU) implementation can be applied easily without additional memory allocation as in PIDAL. We perform simulation studies to validate the proposed algorithm. Our results demonstrate that the proposed algorithm can provide significantly improved reconstruction quality and accurate kinetic parameters. GPU implementation shows the reconstruction time is practical for clinical environment.

This research was supported by by the Converging Research Center Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2012K001492). The authors would like to thank Prof. Yoram Bresler for many discussions.

2. PROBLEM FORMULATION

2.1. Notation, Loglikelihood

First, we define the negative loglikelihood function from Poisson statistics as:

$$L(\mathbf{X}) \quad = \quad \sum_{s=1}^{S} \langle \mathbf{1}, \mathbf{A}\mathbf{x}_s \rangle - \langle \mathbf{y}_s, \log(\mathbf{A}\mathbf{x}_s) \rangle, \qquad (1)$$

where $\mathbf{1}$ denotes a vector with elements of ones, and $\mathbf{X} = [x_{ns}]_{n,s=1}^{N,S}$, $\mathbf{Y} = [y_{ms}]_{m,s=1}^{M,S}$, $\mathbf{A} = [a_{ns}]_{m,n=1}^{M,N}$ where x_{ns} denotes the unknown image at voxel n at time s; and y_{ms} represents the m-th detector measurement at time s, and a_{mn} denotes the probability that an emission photon from n-th voxel is detected at the m-th detector position, respectively.

2.2. Spatio-Temporal Patch-based Low Rank Penalty

It is well-known that a natural image has geometric self-similarities, *i.e.* some parts of the images are similar to different parts of the images. Thus, in our dynamic PET reconstruction, we are interested in imposing a patch-based low rank penalty to exploit self similarities. More specifically, to Eq. (1), we add patch-based low rank penalty for group of patches written as following:

$$\Psi_o(\mathbf{X}, \mathbf{R}) = \sum_{p=1}^{P} \lambda_p \mathrm{Rank}(\mathbf{V}_p), \tag{2}$$

where

$$\mathbf{V}_p = \left[\begin{array}{cccc} \mathbf{R}_{p1}\mathbf{x} & \mathbf{R}_{p2}\mathbf{x} & \cdots & \mathbf{R}_{pQ_p}\mathbf{x} \end{array} \right] \in \mathbb{R}^{B \times NS}$$

where $\mathbf{R}_{pq}, q = 1, \cdots, Q_p$ denotes a size B patch extraction operator from the vectorized spatio-temporal volume of image $\mathbf{x} = \mathrm{vec}(\mathbf{X})$. We denote $\mathbf{R} = \{\mathbf{R}_{pq}\}_{p,q=1}^{P,Q_p}$ as a collection of such operator.

There exists two important issues in using the penalty in Eq. (2). First, the rank operator is not convex, we could use the nuclear norm as a convex relaxation [2]:

$$\Psi(\mathbf{V}, \mathbf{R}) = \sum_{p=1}^{P} \lambda_p \|\mathbf{V}_p\|_* \,, \tag{3}$$

where $\|\mathbf{V}_p\|_* = \sum_{k=1}^{\mathrm{Rank}(\mathbf{V}_p)} \sigma_k(\mathbf{V}_p)$ and $\sigma_k(\mathbf{V}_p)$ denotes the k-th largest singular value of $\mathbf{V}_p$. However, since it has been shown that concave penalty outperforms that convex nuclear norm [5] and the CCCP optimization framework in this paper prefers to concave prior, we use the following concave rank prior [5]:

$$\|\mathbf{V}_p\|_\nu = \sum_{k=1}^{\mathrm{Rank}(\mathbf{V}_p)} h_{\mu,\nu}\left(\sigma_k(\mathbf{V}_p)\right), \quad 0 < \nu \leq 1. \tag{4}$$

where the generalized Huber function $h_{\mu,\nu}(t)$ is defined as

$$h_{\mu,\nu}(t) = \left\{ \begin{array}{ll} |t|^2/2\mu, & \text{if } |t| < \mu^{1/(2-\nu)} \\ |t|^\nu/\nu - \delta & \text{if } |t| \geq \mu^{1/(2-\nu)} \end{array} \right. \tag{5}$$

where $\delta = (1/\nu - 1/2)\mu^{\nu/(2-\nu)}$ to make the function continuous.

Second, to construct a group of similarity patches, we need to find a similarity relationships $\{\mathbf{R}_{pq}\}_{p,q=1}^{P,Q_p}$. However, in dynamic PET, the vectorized image $\mathbf{x}$ is the very object that needs to be estimated, so $\{\mathbf{R}_{pq}\}_{p,q=1}^{P,Q_p}$ is not known *a priori*. To address this issue, we perform a re-fineable similarity searches, *i.e.* we fix a similarity mapping $\{\mathbf{R}_{pq}\}_{p,q=1}^{P,Q_p}$ using the previous estimation of $\mathbf{X}$, and find a new estimate from the updated image.

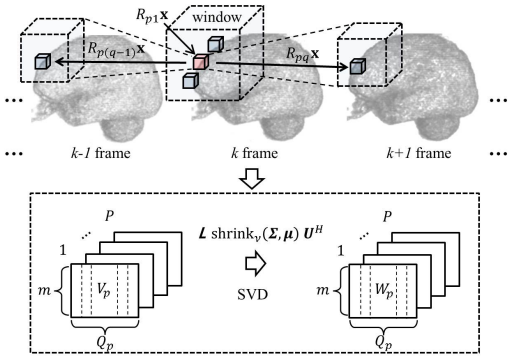


Fig. 1. Description of the patch grouping and low-rank calculation.

3. OPTIMIZATION FRAMEWORK

In this paper, we employ the concave-convex procedure (CCCP) [4] to minimize the resulting optimization problem. The CCCP is especially useful when non-convex terms exist. Consider a typical minimization problem given by

$$\min_{\mathbf{x}} L(\mathbf{A}\mathbf{x}) + \Psi(\mathbf{x}) \,.$$

Suppose $\Psi(\mathbf{x})$ is concave. As $-\Psi(\mathbf{x})$ is convex, there exist a convex function $\Psi^*(\mathbf{v})$ such that

$$\Psi(\mathbf{x}) = \min_{\mathbf{v}} \Psi^*(\mathbf{v}) - \langle \mathbf{v}, \mathbf{x} \rangle \,.$$

Therefore, the original primal problem can be converted as

$$\min_{\mathbf{x}} \min_{\mathbf{v}} L(\mathbf{A}\mathbf{x}) + \Psi^*(\mathbf{v}) - \langle \mathbf{v}, \mathbf{x} \rangle \,.$$

Even if neither L nor Ψ are concave, there exists several ways to apply CCCP. For example, let us define a nonlinear concave coordinate transform $\mathbf{z} = \mathbf{G}(\mathbf{A}\mathbf{x})$. Then, $L(\mathbf{A}\mathbf{x}) = L\left(\mathbf{G}^{-1}(\mathbf{z})\right) = \tilde{L}(\mathbf{z})$. Even though L is convex, there often exists a nonlinear transform $\mathbf{G}(\cdot)$ such that $\tilde{L}$ can be concave. In this case, we know that there exists a convex conjugate $\tilde{L}$ such that

$$\tilde{L}(\mathbf{z}) = L(\mathbf{A}\mathbf{x}) = \min_{\mathbf{v}} \tilde{L}^*(\mathbf{v}) - \langle \mathbf{v}, \mathbf{G}(\mathbf{A}\mathbf{x}) \rangle \,, \tag{6}$$

where $\tilde{L}^*$ is a conjugate for $\tilde{L}$. Therefore, we can convert the primal problem as

$$\min_{\mathbf{x}} \min_{\mathbf{v}} \tilde{L}^*(\mathbf{v}) - \langle \mathbf{v}, \mathbf{G}(\mathbf{A}\mathbf{x}) \rangle + \Psi(\mathbf{x}).$$

Going back to the original problem, now assuming that $\mathbf{R}$ is known, we need to solve the following minimization problem:

$$\min_{\mathbf{X}\in\mathbb{R}^{N\times S}} J(\mathbf{X}) \qquad \text{where } J(\mathbf{X}) = L(\mathbf{X}) + \Psi(\mathbf{X},\mathbf{R}) \,. \quad (7)$$

$L(\mathbf{X})$ denotes the negative loglikelihood and $\Psi(\mathbf{X},\mathbf{R})$ is the patch-based low rank penalty. Here, the penalty term is non-convex, and the gradient of $L(\mathbf{X})$ is non-Lipschitz, and each element of $\mathbf{X}$ should be nonnegative. Now we show how CCCP framework work for our optimization framework.

Note that the $L(\mathbf{X})$ for the Poisson noise in Eq. (1) is convex. To apply the CCCP, we use a concave coordinate transform. More specifically, we define $z^s_{mn} = \log(a_{mn}x_{ns})$. Then, $-\log\left(\sum_{n=1}^N a_{mn}x_{ns}\right) = -\log\left(\sum_{n=1}^N e^{z^s_{mn}}\right)$, which becomes concave with respect to $\{z^s_{mn}\}_{n=1}^N$. Therefore, we have

$$L(\mathbf{X}) \;=\; \min_{\mathbf{c}} L_c(\mathbf{X},\mathbf{c}),$$

where

$$L_c(\mathbf{X},\mathbf{c}) = \sum_{s=1}^S \mathbf{A}\mathbf{x}_s + \langle \mathbf{c}_s, \log \mathbf{c}_s\rangle - \langle \mathbf{c}_s, \log(\mathbf{A}\mathbf{x}_s)\rangle. \quad (8)$$

In the case of the patch-based rank penalty, we use the generalized Huber function in Eq. (5). Here, $|t|^2/\mu - h_{\mu,\nu}(t)$ is strictly convex. Therefore, the Legendre-Fenchel transform tells us that there exist $g_{\mu,\nu}$ such that

$$h_{\mu,\nu}(t) = \min_s\{|s-t|^2/\mu + g_{\mu,\nu}(s)\}. \quad (9)$$

The corresponding rank penalty for a matrix $\mathbf{V}$ is given by

$$\begin{aligned}\|\mathbf{V}\|_{h_{\mu,\nu}} &= \sum_{k=1} h_{\mu,\nu}(\sigma_k(\mathbf{V})) \\ &= \min_{\mathbf{W}}\left\{\frac{1}{\mu}\|\mathbf{V}-\mathbf{W}\|_F^2 + \|\mathbf{W}\|_{g_{\mu,\nu}}\right\} \quad (10)\end{aligned}$$

where $\|\mathbf{W}\|_{g_{\mu,\nu}} = \sum_{k=1} g_{\mu,\nu}(\sigma_k(\mathbf{V}))$. Chartrand [5] showed that $g_{\mu,\nu}(s)$ is convex when $\nu = 1$, but in general it is not convex. However, even when $g_{\mu,\nu}$ is non-convex and does not have close form expression, there exist a close form expression for the minimizer of Eq. (9) given as

$$\text{shrink}_\nu(t,\mu) = \max\{0, |t| - \mu|t|^{\nu-1}\}t/|t| \,. \quad (11)$$

Now, using Eq. (8) and Eq. (10), we have the following minimization problem:

$$\min_{\mathbf{X},\mathbf{c},\{\mathbf{W}_p\}_{p=1}^P} L_c(\mathbf{X},\mathbf{c}) + \sum_{p=1}^P \lambda_p\left\{\frac{1}{\mu}\|\mathbf{V}_p - \mathbf{W}_p\|_F^2 + \|\mathbf{W}_p\|_{g_{\mu,\nu}}\right\}.$$

One of the main advantage of the proposed CCCP framework is that each subproblem has close form solutions. Here, we describe them in detail.

1. Minimization with respect to $\mathbf{W}_p$
   First, note that the minimization problem is independent to $L_c(\mathbf{X},\mathbf{c})$. Moreover, the problem can be decomposed for each patch by patch. More specifically, we have

   $$\mathbf{W}_p^{(k+1)} = \arg\min_{\mathbf{W}}\left\{\frac{1}{\mu}\|\mathbf{V}_p^{(k)} - \mathbf{W}\|_F^2 + \|\mathbf{W}\|_{g_{\mu,\nu}}\right\}. \quad (12)$$

   Using the shrinkage relationship in Eq. (11), the close form solution for Eq. (12) is given by

   $$\mathbf{W}_p^{(k+1)} = \mathbf{L}\text{shrink}_\nu(\mathbf{\Sigma},\mu)\mathbf{U}^H,$$

   where $\text{shrink}_\nu(\mathbf{\Sigma},\mu)$ denotes an element by element singular value shrinkage operator.

2. Minimization with respect to $\mathbf{c}$
   Using the constraint $\sum_{n=1}^N c^s_{mn} = y_m$, we have the following close form solution for the constrained optimization problem:

   $$c^{s(k+1)}_{mn} = y_{ms}\frac{a_{mn}x^{(k)}_{ns}}{\sum_{n'=1}^N a_{mn'}x^{(k)}_{n's}}, \quad \forall\, m,n,s \,. \quad (13)$$

3. Minimization with respect to $\mathbf{X}$
   Finally, for given $\mathbf{c}^{(k+1)}$ and $\mathbf{W}^{(k+1)}$, we can obtain a close form solution for update of $\mathbf{X}^{(k+1)}$. We calculate a fixed point equation of the gradient of the cost function with respect to x_{ns}. Then, the close form solution is given by

   $$x^{(k+1)}_{ns} = \frac{-b_{ns} + \sqrt{(b_{ns})^2 + 4d_{ns}x^{EM}_{ns}\sum_{m=1}^M a_{mn}}}{2d_{ns}}, \quad (14)$$

   where

   $$x^{EM(k+1)}_{ns} = \frac{c^{s(k+1)}_{mn}}{\sum_{m=1}^M a_{mn}}, \quad (15)$$

   and

   $$d_{ns} := \frac{1}{\mu}\sum_{p\in I_{ns}}\lambda_p, \quad b^{(k+1)}_{ns} = \sum_m a_{mn} - \frac{1}{\mu}\sum_{p\in I_{ns}}\lambda_p w^{(k+1)}_{p,ns}.$$

Note that the solution is always non-negative, and our update equation is pixel-by-pixel update similar to OSEM algorithm, and guarantees the global convergence. Compared to the augmented Lagrangian method [6], our framework does not require any additional memory for storing Lagrangian parameter, so it is good for memory intense 3-D dynamic PET. Moreover, each subprogram is completely parallelizable, so GPU implementation is very straightforward.

4. EXPERIMENTAL RESULTS

Parametric imaging has been widely used to analyze dynamic PET images. We set our simulation model using the two compartmental model with four rate constants with respect to the

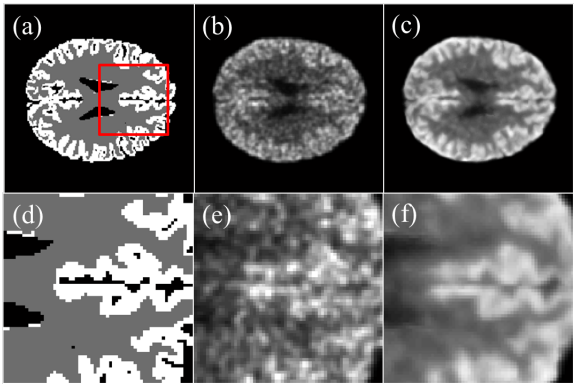


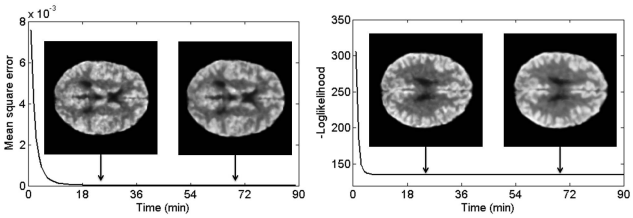
Fig. 2. Reconstruction images of (a)(d) Ground truth, (b)(e) Gaussian smoothing of conventional OSEM, and (c)(f) the proposed algorithm.

Grey matter		K_1	k_2	k_3
Ground-Truth		0.101	0.071	0.042
Gaussian smoothing	mean	0.1075	0.1177	0.0611
	std.	0.0245	0.1376	0.0813
Proposed method	mean	0.1022	0.0735	0.0418
	std.	0.0071	0.0178	0.0098

Table 1. Voxel-wise parameter analysis for the grey matter.

entry and return of tracers in tissue. The corresponding kinetic parameters of K1, k2, k3, k4($\approx$ 0) are shown in Table 1. We then generated 15 frames of low photon count emissions in grey and white matters using the Monte-Carlo simulation in a high-resolution research tomograph (HRRT) geometry. Specifically, in each frame, the size of image is $256 \times 256 \times 207$, and the size of sinogram is $256 \times 288 \times 2209$. We used an Nvidia Tesla C2070 GPU on the Intel i7-2600 CPU platform.

We compared reconstruction images using Gaussian smoothing of the conventional OSEM and the proposed algorithm. In Fig. 2, the details of shapes were difficult to be identified in Gaussian smoothing images, whereas our proposed algorithm provided significantly enhanced details of shapes and the boundaries of the grey matters are clearly visible. Parameter extraction was also conducted in Table 1. Here, our proposed algorithm provides accurate parameters with small variations. Finally, to demonstrate the convergence of the algorithm, the mean square error (MSE) and negative loglikelihood in Eq. (1) were calculated as shown in Fig. 3. We can confirm that the algorithm converges very quickly. In our GPU implementation, the OSEM takes 4.5 sec, and the patch based low-rank penalty takes 13.8 sec, which are 445 times and 78 times faster than CPU, respectively. Total execution time of the proposed algorithm with



(a) (b)

Fig. 3. (a) Mean square error by iteration, (b) Negative log-likelihood by iteration.

15 frames takes about 10 minutes, which is a reasonable time for clinical environment.

5. CONCLUSIONS

In conclusion, we proposed a dynamic PET reconstruction using non-convex low rank patch based regularization and derived a memory efficient globally convergent algorithm using CCCP procedure. In simulation experiment, our proposed algorithm provided significantly improved images. Furthermore, our GPU implementation reduced the computational cost significantly so that it can be used in real clinical environment.

6. REFERENCES

[1] A. Rahmim, J. Tang, and H. Zaidi, "Four-dimensional (4D) image reconstruction strategies in dynamic PET: beyond conventional independent frame reconstruction," *Medical Physics*, vol. 36, pp. 3654–3670, 2009.

[2] E.J. Candès and B. Recht, "Exact matrix completion via convex optimization," *Foundations of Computational Mathematics*, vol. 9, no. 6, pp. 717–772, 2009.

[3] M.A.T. Figueiredo and J.M. Bioucas-Dias, "Restoration of Poissonian images using alternating direction optimization," *IEEE Transactions on Image Processing*, vol. 19, no. 12, pp. 3133–3145, 2010.

[4] A.L. Yuille and A. Rangarajan, "The concave-convex procedure," *Neural Computation*, vol. 15, no. 4, pp. 915–936, 2003.

[5] R. Chartrand, "Nonconvex splitting for regularized low-rank sparse decomposition," *Los Alamos National Laboratory Report: LA-UR-11-11298*, 2012.

[6] S. Ramani and J.A. Fessler, "A splitting-based iterative algorithm for accelerated statistical X-ray CT reconstruction," *IEEE Transactions on Medical Imaging*, vol. 31, no. 3, pp. 677–688, 2012.