WAIT: Workshop on the Advances in Information Technology, Budapest, 2017

Direct Parametric Reconstruction for Dynamic PET with Floating Frames

László Szirmay-Kalos and Ágota Kacsó

Budapest University of Technology and Economics (e-mail: szirmay@iit.bme.hu)

Abstract

This paper proposes a scalable dynamic PET reconstruction approach for GPUs. In dynamic PET reconstruction the space-time activity function needs to be recovered from measurements. The computation is based on the maximum likelihood principle, which means that we look for the space-time function that maximizes the probability of the actual measurements. The enormous computational burden can be handled by GPUs if the algorithm is decomposed to parallel, independent threads and the storage requirements are kept under control. As in practical systems, the number of spatial basis functions is in the range of a hundred million, while the number of events can exceed billions, we should avoid storing many voxel arrays or multiple LOR arrays. Both the computational complexity and the storage space are proportional to the number of frames, i.e. how fine the temporal discretization is, which determines the ability to reconstruct high frequency phenomena. To reduce the storage requirements and the computation time, the number of frames should be minimized, but without compromising the reconstruction of quick changes. The paper addresses this problem and presents techniques for minimizing the number of frames while maintaining the temporal accuracy of the reconstruction.

1. Introduction

In dynamic Positron Emission Tomography (PET), we focus on the dynamic nature of biological processes, like accumulation and emptying drugs in certain organs. Such studies are essential in pharmaceutical research. Dynamic tomography reconstructs space-time activity density of the points of interest at time t. To represent the spatial dependence with finite data, the space-time activity is expressed in a finite function series form. In the simplest case, the domain is decomposed to N_V regions, which are assumed to be homogeneous. A region can be a larger volume of an organ or anatomic part that is believed to be homogeneous, or a small voxel that discretizes the volume of interest. In this respect there is no difference between region-based and voxel-based reconstruction approaches, only the number and definition of the regions differ. In this paper we consider both options and use the word region even if it is as small as a voxel.

We assume that the radiotracer concentration can be expressed by a common *kinetic model* $C(\mathbf{p}_V, t)$, where spatial dependent properties are encoded in a low dimensional vector of N_p kinetic parameters $\mathbf{p}_V$. Such models can be defined based on the mathematical description of the biological/chemimal processes or on compartment analysis. Then, the activity density is the product of the concentration and the exponential fall-off $\exp(-\lambda t)$ due to the decay, where λ is the decay rate of the radiotracer.

If the reconstruction and kinetic parameter estimation are separated and executed after each other, the approach is indirect. In a *direct method*, reconstruction and parameter estimation are merged into a single process. Direct reconstruction can be considered as a regularization in the time domain since we impose the requirement that the resulting temporal functions must belong to the class that can be represented by the kinetic model. The state of the art and previous work on direct estimation of kinetic parametric images for dynamic PET are surveyed in review articles [16, 7].

Using the time activity, the expected number of radioactive decays, i.e. number of positrons generated in region V and in time interval $[t_s, t_e]$ is

$$\tilde{x}(\mathbf{p}_V, t_s, t_e) = \int_{t_s}^{t_e} C(\mathbf{p}_V, t) \exp(-\lambda t) \mathrm{d}t. \qquad (1)$$

The positron emitted at a decay may annihilate with an elec-

tron, when two oppositely directed gamma-photons are born, which might be detected by the tomograph. A PET/CT system collects the *events* of simultaneous photon incidents in detector pairs. An event is a composition of the identification of the detector pair, also called *Line Of Response* or *LOR*, and its time of occurrence.

List mode reconstructions process the raw data of events. If the measurement time is discretized by interval boundaries $t_1, t_2, \ldots, t_{N_T}$ and events are binned in frames (t_T, t_{T+1}), the time complexity reduces from the number of events to the number of frames N_T, but the information of the actual time of the events within a frame is lost.

If the kinetic model is a linear function of non-negative parameters, then the classical ML-EM scheme can be extended to find all parameters [5]. Biologically plausible kinetic models often depend on their parameters non-linearly, rendering the classical ML-EM approach not applicable anymore. A first approach aiming at the reconstruction of parameters of the two-tissue-compartment model for each region rather than only for a few ROIs searched the maximum of a penalized likelihood with the parametric iterative coordinate descent algorithm [2]. The one-compartment model was considered in [18]. Wang et al. [17] proposed the application of optimization transfer principle to find the optimum for general kinetic models, which locally approximated the objective function of the optimization. This method alternates independent reconstructions of the frames and model fitting in iteration steps using, for example, the Levenberg-Marguard algorithm.

Our approach builds on that of Wang [17] and executes the maximum-likelihood dynamic reconstruction method with a two-level iteration scheme that decomposes the solution to phases where each phase can be efficiently implemented on the GPU, i.e. can be computed by parallel threads with no communication. Data exchange may happen just between phases. Additionally, the storage compexity of the method is low, we need to represent the compressed list of events, a single LOR array, and just a few voxel arrays, in additional to the final results, which is the voxel array storing one parameter vector at each voxel.

To handle the problem of the fine frames, we present a *floating frame* approach, where the advantages of list mode and binned mode are combined. The time discretization is neither uniform nor fixed during the reconstruction process, but frames are *floating* during the iterations to minimize the discretization error while working with relatively low number of frames at a time.

2. Dynamic PET reconstruction

The correspondence between positron generation and gamma photon detection is established by *system matrix* $\mathbf{A}$ [12, 11]. that expresses the probability of generating an event in LOR L given that a positron is emitted in voxel V. The expected number of events $\tilde{y}_L(t_s, t_e)$ in LOR L during frame (t_s, t_e) is the sum of the contributions of all voxels in the volume during this time.

$$\tilde{y}_L(t_s, t_e) = \sum_V \mathbf{A}_{L,V} \tilde{x}(\mathbf{p}_V, t_s, t_e). \tag{2}$$

Unknown coefficients $\mathbf{p}$ are found to maximize the probability of the actually measured data. The measured number of hits in LOR L in frame (t_s, t_e) follows a Poisson distribution of expectation $\tilde{y}_L(t_s, t_e)$. Because of the statistical independence of different LORs and different frames as we required them to be disjoint, the combined probability considering all LORs and all frames is the product of elementary probabilities. According to the concept of maximum-likelihood reconstruction, unknown parameters are found to maximize the following log-likelihood:

$$\log \mathcal{L} = \sum_L \sum_T \left(y_L(t_T, t_{T+1}) \log \tilde{y}_L(t_T, t_{T+1}) - \tilde{y}_L(t_T, t_{T+1}) \right). \tag{3}$$

The optimization problem has very high computational complexity. The number of free variables, i.e. the dimension of the search space is $N_P \times N_V$, the log-likelihood acting as the optimization target is a sum of $N_L \times N_T$ terms. The range of N_V and N_L is typically several hundred millions, N_T is typically in the order of a hundred, and N_P is less than 10 since we wish to describe a time function with a few parameters. Thus, both the dimension of the search space and the number of terms in the optimization target can be in the order of millions. This means that the search method should be carefully designed and high performance computation platforms should be exploited.

The reconstruction means the maximization of the log-likelihood of Equation 3. The likelihood has an extremum where all partial derivatives are zero. Computing the partial derivatives, we obtain

$$\sum_T \frac{\partial \tilde{x}(\mathbf{p}_V, t_T, t_{T+1})}{\partial \mathbf{p}_{V,P}} \sum_L \left(\mathbf{A}_{L,V} \frac{y_L(t_T, t_{T+1})}{\tilde{y}_L(t_T, t_{T+1})} - \mathbf{A}_{L,V} \right) = 0. \tag{4}$$

for $V = 1, 2, \ldots, N_V$ and $P = 1, \ldots, N_P$. Thus, we have $N_V \times N_P$ equations, each containing N_T terms that depend on unknown parameters of all voxels, and the computation of each equation requires the consideration of all LORs L for which $\mathbf{A}_{L,V}$ is not zero. Note that accurate reconstruction requires the computation of positron range and scattered particle paths as well, which makes system matrix $\mathbf{A}$ not sparse.

The computation of the derivatives of the log-likelihood requires a static forward projection and a back projection in each frame T. Indeed, in frame T, the expected number of radioactive decays in region V is $\tilde{x}(\mathbf{p}_V, t_T, t_{T+1})$, which is forward projected to obtain $\tilde{y}_L(t_T, t_{T+1})$ according to Equation 2. A gathering type approach is obtained if computational threads are assigned to LORs and each thread computes the expected hits for a single LOR L during frame T

An ML-EM static back projection would obtain a new estimate of the activity as

$$x_V(t_T,t_{T+1}) = \tilde{x}(\mathbf{p}_V,t_T,t_{T+1}) \cdot \frac{\sum_L \mathbf{A}_{L,V} \frac{y_L(t_T,t_{T+1})}{\tilde{y}_L(t_T,t_{T+1})}}{\sum_L \mathbf{A}_{L,V}}.$$

Note that this step is a classical ML-EM step, where a gathering algorithm assigns computational threads to regions and each thread computes a single value $x_V(t_T,t_{T+1})$.

From this equation, we can express

$$\sum_L \mathbf{A}_{L,V} \frac{y_L(t_T,t_{T+1})}{\tilde{y}_L(t_T,t_{T+1})} = \frac{x_V(t_T,t_{T+1})}{\tilde{x}(\mathbf{p}_V,t_T,t_{T+1})} \cdot \sum_L \mathbf{A}_{L,V},$$

which can be substituted into Equation 4:

$$\sum_T \frac{\partial \tilde{x}(\mathbf{p}_V,t_T,t_{T+1})}{\partial \mathbf{p}_{V,P}} \left(\sum_L \mathbf{A}_{L,V} \right) \left(\frac{x_V(t_T,t_{T+1})}{\tilde{x}(\mathbf{p}_V,t_T,t_{T+1})} - 1 \right) = 0. \tag{5}$$

Dividing both sides by the sensitivity of the voxel, i.e. by $\sum_L \mathbf{A}_{L,V}$, we get an equivalent requirement for the extremum of the likelihood

$$\sum_T \frac{\partial \tilde{x}(\mathbf{p}_V,t_T,t_{T+1})}{\partial \mathbf{p}_{V,P}} \left(\frac{x_V(t_T,t_{T+1})}{\tilde{x}(\mathbf{p}_V,t_T,t_{T+1})} - 1 \right) = 0. \tag{6}$$

In this equation $\tilde{x}(\mathbf{p}_V,t_T,t_{T+1})$ depends on the unknown parameter vector of the given voxel $\mathbf{p}_V$, while x_V involves forward and back projections and depends on the parameter vectors of all voxels. Thus, if x_V were known, then the computation could be decoupled for different voxels, where a system of equations with N_P unknowns needs to be solved. In this way, forward/backward projection can be separated from the parameter fitting, thus the complexity of the algorithm will be the sum of the complexities of the two steps and not their product.

Having fixed $x_V(t_T,t_{T+1})$, the non-linear equation is solved with the Levenberg-Marquard algorithm. This can also be imagined as a curve-fitting process.

Concerning the evaluation and the solution of this equation on massively parallel architectures, the process should be decomposed to phases where computational threads are assigned either to voxels or LORs and can be executed without inter-thread communication. Note that in forward projection we should compute many LOR values from many voxel values, in back projection many voxel values from many LORs, in the phase of voxel activity evaluation many voxel parameter values to many voxel activity values. On parallel architectures *gathering type* or *output-driven* algorithms are preferred where the parallel threads are assigned to the output values and a thread gathers the contributions of inputs that can affect this particular output. As number of LORs N_L is huge and the number of voxels N_V is also large, at most one LOR array and just a few voxel arrays can be used by the algorithm. To achieve this goal, the computation of the derivative of the log-likelihood function is decomposed to steps that can be executed by the GPU, and its iterative solution is decomposed to sub-iterations. Let us first consider the evaluation of the derivative of the log-likelihood function.

There are various options to regularize the solution, which is essential in the case of inverse problems. One option is the modification of the optimization target in Equation 3 by a regularization term that penalizes unacceptable solutions [4], where, for example, the spatial or temporal variation is too high. The *method of sieves* [9, 10, 15, 13], on the other hand, does not modify the optimization target, but filters the iterated approximation in each iteration step [1, 6]. Mathematically, this approach projects the current estimate into the subspace of acceptable solutions in each iteration [8, 3]. Filtering can also exploit anatomic information gathered by a CT or an MR [14].

Putting the discussed steps together, the pseudo-code of the reconstruction is as follows:

for $n = 1$ **to** $n_{\max}$ **do** *// main iteration*
  for $T = 1$ **to** N_T **do** *// iterate through frames*
    // evaluation
    foreach voxel V **in parallel** $\tilde{x}_V = \int_{t_T}^{t_{T+1}} C(\mathbf{p}_V,t)\exp(-\lambda t)\mathrm{d}t$
    // forward projection
    foreach LOR L **in parallel** $\tilde{y}_L = \sum_{V'} \mathbf{A}_{L,V'} \tilde{x}_{V'}$
    // back projection
    foreach voxel V **in parallel** $x_V = \tilde{x}_V \cdot \frac{\sum_L \mathbf{A}_{L,V} \frac{y_L}{\tilde{y}_L}}{\sum_L \mathbf{A}_{L,V}}$
  endfor
  // curve fitting
  foreach voxel V **in parallel** $\hat{\mathbf{p}}_V^{(n+1)} =$ Solve Equation 5
  // filtering
  foreach voxel V **in parallel** $\mathbf{p}_V^{(n+1)} = \mathrm{Filter}(\hat{\mathbf{p}}_V^{(n+1)})$
endfor

The most time consuming steps of the reconstruction are the forward projection and back projection, and these steps should also be repeated in every iteration and for every frame. Even if its time is reduced by efficient parallel computation of the GPUs, it still takes long and requires the temporary storage of a voxel array for each frame. Thus, not only the time complexity but also the storage requirement will be prohibitive if the number of frames is large. Thus, the reduction of the number of frames is imperative. However, distributing events just in a few bins may lead to information loss, and as a result, the quality of the reconstruction of the high frequency phenomena may degrade. This paper proposes a method that minimizes the computational error due to such frame reduction.

3. The method of frame definition

Suppose that the time is decomposed to a lot of fine frames, but the number of frames makes both time and storage requirements prohibitive. Thus, we should reduce the number of frames making them longer. We focus on how the frames should be merged and defined to introduce just a small error

in the reconstruction. Of course, a frame definition method is feasible only if it does not require a reconstruction with many frames since our objective is to reduce the number of frames to speed up reconstruction. Thus, sub-optimal or approximate methods are also acceptable if they meet this requirement.

We first examine how the reconstruction error is modified if two frames are merged together. Then, using the obtained error measure, different frame definition methods are proposed.

Note that during reconstruction we need to solve equations of form

$$F(\mathbf{p}_V) = \sum_T \frac{\partial \tilde{x}(\mathbf{p}_V, t_T, t_{T+1})}{\partial \mathbf{p}_{V,P}} \left(\frac{x_V(t_T, t_{T+1})}{\tilde{x}(\mathbf{p}_V, t_T, t_{T+1})} - 1 \right) = 0. \tag{7}$$

Let us now consider the change of expected activity $\tilde{x}(\mathbf{p}_{V,P}, t)$ as a result of merging frames T^* and T^*+1. We use the following shorthand notations, where index i is 1 for the first frame, i.e. when $t_s = t_{T^*}$ and $t_e = t_{T^*+1}$, index i is 2 for the second frame, i.e. when $t_s = t_{T^*+1}$ and $t_e = t_{T^*+2}$, and index i is 12 for the merged frame, i.e. when $t_s = t_{T^*}$ and $t_e = t_{T^*+2}$:

$$\tilde{x}(\mathbf{p}_V, t_s, t_e) = \tilde{x}_i, \quad \frac{\partial \tilde{x}(\mathbf{p}_V, t_s, t_e)}{\partial \mathbf{p}_{V,P}} = \tilde{x}'_i, \quad x_V(t_s, t_e) = x_i,$$

$$y_L(t_s, t_e) = y_{L,i}, \quad \tilde{y}_L(t_s, t_e) = \tilde{y}_{L,i}.$$

Variable $\tilde{x}$ giving the expected number of radioactive decays and its derivative are additive in frames since they are time integrals (Equation 1):

$$\tilde{x}_1 + \tilde{x}_2 = \tilde{x}_{12}, \quad \tilde{x}'_1 + \tilde{x}'_2 = \tilde{x}'_{12}.$$

As the expected number of hits $\tilde{y}_L(t_s, t_e)$ depends linearly on expected voxel activities $\tilde{x}_V(t_s, t_e)$ (Equation 2), it is also additive, so is the number of measured hits $y_L(t_s, t_e)$:

$$\tilde{y}_{L,1} + \tilde{y}_{L,2} = \tilde{y}_{L,12}, \quad y_{L,1} + y_{L,2} = y_{L,12}.$$

We show that the result of the back projection

$$x_V(t_T, t_{T+1}) = \tilde{x}(\mathbf{p}_V, t_T, t_{T+1}) \cdot \frac{\sum_L \mathbf{A}_{L,V} \frac{y_L(t_T, t_{T+1})}{\tilde{y}_L(t_T, t_{T+1})}}{\sum_L \mathbf{A}_{L,V}}$$

is also approximately additive if the ratios of expected decays $\tilde{x}$ in frame T^* and frame T^*+1 are similar in all voxels and mainly due to the different lengths of the frames. Note that this requirement is met if the time dependence of $C(\mathbf{p}_V, t)\exp(-\lambda t)$ is roughly constant in the time interval of the two frames since ratio

$$\frac{\tilde{x}_V(t_T, t_{T+1})}{\tilde{x}_V(t_{T+1}, t_{T+2})} = r$$

is then equal to the lengths of the two merged frames. If the ratios are the same, then

$$\frac{\tilde{x}_1}{\tilde{y}_{L,1}} \approx \frac{\tilde{x}_2}{\tilde{y}_{L,2}} \approx \frac{\tilde{x}_{12}}{\tilde{y}_{L,12}}.$$

Taking into account that the number of measured hits in LORs $y_L(t_s, t_e)$ is also additive

$$\frac{\tilde{x}_1 y_{L,1}}{\tilde{y}_{L,1}} + \frac{\tilde{x}_2 y_{L,2}}{\tilde{y}_{L,2}} \approx \frac{\tilde{x}_{12} y_{L,12}}{\tilde{y}_{L,12}}.$$

Using this, the additivity of the back projection results can be shown:

$$x_1 + x_2 = \tilde{x}_1 \cdot \frac{\sum_L \mathbf{A}_{L,V} \frac{y_{L,1}}{\tilde{y}_{L,1}}}{\sum_L \mathbf{A}_{L,V}} + \tilde{x}_2 \cdot \frac{\sum_L \mathbf{A}_{L,V} \frac{y_{L,2}}{\tilde{y}_{L,2}}}{\sum_L \mathbf{A}_{L,V}} \approx$$

$$\tilde{x}_{12} \cdot \frac{\sum_L \mathbf{A}_{L,V} \frac{y_{L,12}}{\tilde{y}_{L,12}}}{\sum_L \mathbf{A}_{L,V}} = x_{12}.$$

If frames are merged, then error ΔF is introduced into the equation, so instead of the original equation we solve

$$F(\mathbf{p}_V) + \Delta F = 0$$

where error

$$\Delta F = \tilde{x}'_{12} \frac{x_{12}}{\tilde{x}_{12}} - \tilde{x}'_1 \frac{x_1}{\tilde{x}_1} - \tilde{x}'_2 \frac{x_2}{\tilde{x}_2} \tag{8}$$

is due to frame merging. Using the additivity relations, the error can be approximated as

$$|\Delta F| = \frac{\tilde{x}_1 \tilde{x}_2}{\tilde{x}_1 + \tilde{x}_2} \left| \frac{x_1}{\tilde{x}_1} - \frac{x_2}{\tilde{x}_2} \right| \left| \frac{\tilde{x}'_1}{\tilde{x}_1} - \frac{\tilde{x}'_2}{\tilde{x}_2} \right|$$

The difference of the solved equation ΔF makes some error $\Delta \mathbf{p}$ in parameter vector $\mathbf{p}$, which in turn, modifies the complete activity density and its integrals in frames. What we are interested in is the scale of this modification, for which a tractable error measure should be found. Based on the recognition that the solution of Equation 7 is equivalent to a time curve fitting problem, the difference of $\tilde{x}$ obtained with and without merging the frames can be well represented by the difference of its values at the point of time where the two frames are merged. It means that we assume that error ΔF translates only to error $\Delta \tilde{x}$ of $\tilde{x}_{12}$. Thus, we look for the difference of $\tilde{x}_{12}$ that can alone compensate the modification of F, while assuming that the temporal function is similar in other frames:

$$F(\mathbf{p}_V + \Delta \mathbf{p}) - F(\mathbf{p}_V) = \Delta F = \tilde{x}'_{12} \frac{x_{12}}{\tilde{x}_{12} + \Delta \tilde{x}} - \tilde{x}'_{12} \frac{x_{12}}{\tilde{x}_{12}}.$$

Using first order approximation

$$\frac{x_{12}}{\tilde{x}_{12} + \Delta \tilde{x}} \approx \frac{x_{12}}{\tilde{x}_{12}} \left(1 - \frac{\Delta \tilde{x}}{\tilde{x}_{12}} \right).$$

and the result of Equation 8, we can express error $\Delta \tilde{x}$ as

$$\Delta \tilde{x} = \frac{\tilde{x}^2_{12}}{\tilde{x}'_{12} x_{12}} \left(\tilde{x}'_{12} \frac{x_{12}}{\tilde{x}_{12}} - \tilde{x}'_1 \frac{x_1}{\tilde{x}_1} - \tilde{x}'_2 \frac{x_2}{\tilde{x}_2} \right). \tag{9}$$

With the additivity relations, the absolute value of error $\Delta \tilde{x}$ can be expressed in the following form:

$$|\Delta \tilde{x}| \approx \frac{\tilde{x}_1 \tilde{x}_2 (\tilde{x}_1 + \tilde{x}_2)}{(\tilde{x}'_1 + \tilde{x}'_2)(x_1 + x_2)} \left| \frac{x_1}{\tilde{x}_1} - \frac{x_2}{\tilde{x}_2} \right| \left| \frac{\tilde{x}'_1}{\tilde{x}_1} - \frac{\tilde{x}'_2}{\tilde{x}_2} \right|. \tag{10}$$

With this error formula we can select that pair of consecutive frames that can be merged together introducing the smallest error while reducing the number of frames by one, and iterating the same approach, reducing the number of frames to an arbitrary number. The direct application of this approach requires a fine frame solution at least approximately since the error formula contains variables of the two fine frames that are merged. We call this the *bottom-up approach*, which needs efficient approximations of the reconstruction taking higher number of frames.

The error formula can also be used in a *top-down approach* that starts with a very few number of frames and refines them where necessary. However, in this case, we need heuristics to guess the data of the fine frames since only the data of the merged frame is available.

Finally, we can modify the frame boundaries without changing the number of frames. Again, heuristics is needed to approximate the data of the fine frame that is attached to a different larger frame.

3.1. Bottom-up approach

The bottom-up approach needs initial reconstruction with many fine frames, which is feasible only if the number of regions is small. Thus, at the first phase of the method, we use large voxels and execute real region based reconstruction. If the number of regions is small, then the system matrix can be stored in the computer memory, so we should compute it only once making forward and back projections fast. Having obtained a reconstruction with high temporal but low spatial resolution, the spatial resolution is increased while the temporal reconstruction is decreased controlled by Equation 10. It means that we merge those frames where the introduced error in the low spatial resolution model is minimal.

3.2. Top-down approach

In the top-down approach we start with a relatively few number of frames and split a frame into two if necessary. To decide which frame should be split, we should use equation 10 in reverse direction and ask how the error can be maximally decreased by decomposing a frame into two frames. Before subdivision, we have the data of the combined frame only, so decision should be based only on this. Factor

$$\left|\frac{x_1}{\tilde{x}_1}-\frac{x_2}{\tilde{x}_2}\right|\left|\frac{\tilde{x}'_1}{\tilde{x}_1}-\frac{\tilde{x}'_2}{\tilde{x}_2}\right|$$

depends on the difference of the relative errors of refined frames, so this factor is ignored during the decision. When a frame is broken into two, we decompose the time interval into two equal parts, so $\tilde{x}_1$ and $\tilde{x}_2$ are approximately equal, thus their product can be well approximated as

$$\tilde{x}_1\tilde{x}_2 \approx \left(\frac{\tilde{x}_1+\tilde{x}_2}{2}\right)^2$$

according to the relation of the geometric and arithmetic means. Using this approximation we decompose that frame where

$$\frac{(\tilde{x}_1+\tilde{x}_2)^3}{(\tilde{x}'_1+\tilde{x}'_2)(x_1+x_2)} \approx \frac{\tilde{x}^3_{12}}{\tilde{x}'_{12}x_{12}}$$

is maximal.

4. Results

To examine the proposed method, we use a 2D model [12] where $N_L = 2115$ and $N_V = 1024$ (Figure 1) and reconstruct the activity curves of a 2D brain model. For this model, the system matrix can be precisely computed.

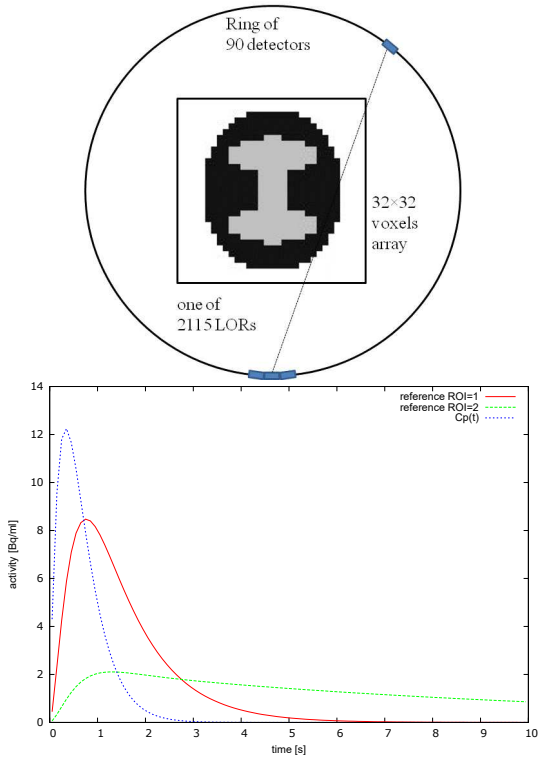


Figure 1: 2D tomograph model: The detector ring contains 90 detector crystals and each of them is of size 2.2 in voxel units and participates in 47 LORs connecting this crystal to crystals being in the opposite half circle, thus the total number of LORs is $90 \times 47/2 = 2115$. The voxel array to be reconstructed is in the middle of the ring and has 32×32 resolution, i.e. 1024 voxels. The lower image shows the blood input function and the simulated time activity curves in the gray matter (ROI 2) and white matter (ROI 1) of the brain.

For defining the frame boundaries we use the bottom–up approach. We split $20s$ the measurement time into 100 frames and merged the two consecutive frames introducing the smallest error. we repeat this prcess until the number of

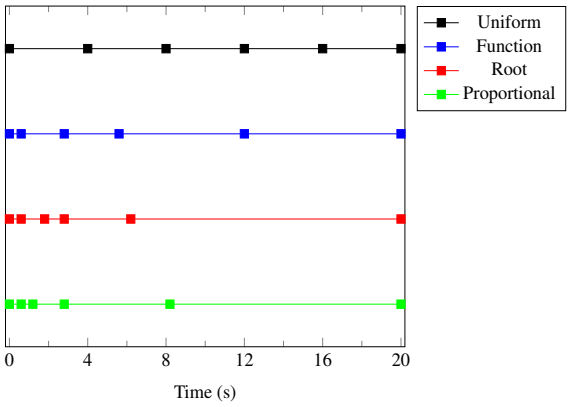


Figure 2: The resulting 5 frames using the discussed methods and a simple uniform split of the measurement time.

frames is sufficiently low, in this particular case at the end we have 5 time frames. The introduced error can be expressed by the Equation 10 or 9. An other option is to merge those two frames where the sum of activities is minimal, thus resulting in frames proportional with the number of hits.

Figure 2 shows the results of these frame merging methods, while on Figure 3 we see the different stages of the frame merging process, and the normalized merging errors for the consecutive time frames.

For the reconstruction with high spatial resolution we use the obtained 5 frames. Figure 4 shows the activity error during the iterations, and the resulting averaged time activity functions obtained with the different method used for frame definition. Note that the results are much better if the measurement time is not just simply split into 5 equal length frames but we use the one of the proposed methods.

5. Conclusions

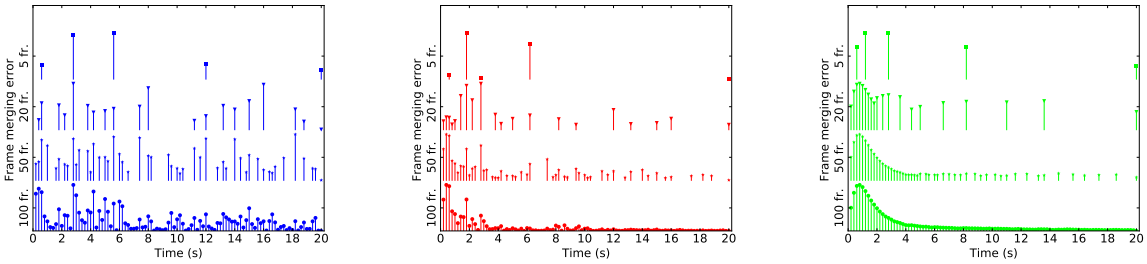
In this paper we investigated the problem of dynamic PET reconstruction when the total activity in a region of interest needs to be reconstructed as a function of time. We have shown that binning in time but still using non-constant basis functions is possible and can greatly reduce the computation time while maintaining the accuracy of the time consuming list mode reconstruction. We have also shown that the proposed methods for defining a low number of time frames provides more accurate results than the simple uniform division of the measurement time.

Acknowledgement

This work has been supported by OTKA K–104476 and VKSZ-14 PET/MRI 7T.

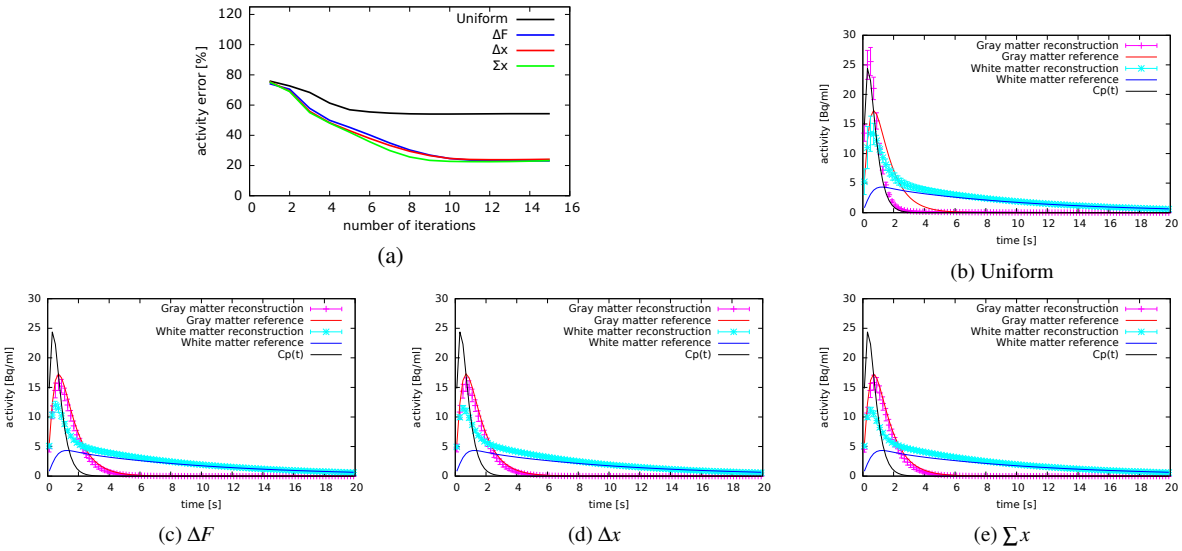
References

1. B. Csébfalvi and G. Rácz. Retailoring box splines to lattices for highly isotropic volume representations. *Computer Graphics Forum*, 35(3):411–420, 2016.
2. M.E. Kamasak, C.A. Bouman, E.D. Morris, and K. Sauer. Direct reconstruction of kinetic parameter images from dynamic pet data. *IEEE Trans Med Imaging*, 24(5):636–650, 2005.
3. M. Magdics, L. Szirmay-Kalos, B. Tóth, and T. Umenhoffer. Filtered sampling for PET. In *IEEE Nuclear Science Symposium Conference Record (NSS/MIC)*, 2012.
4. Milán Magdics and László Szirmay-Kalos. *Optimization in Computer Engineering - Theory and Applications*, chapter Total Variation Regularization in Maximum Likelihood Estimation. Scientific Research Publishing, 2011.
5. J. Matthews, D. Bailey, P. Price, and V. Cunningham. The direct calculation of parametric images from dynamic PET data using maximum-likelihood iterative reconstruction. *Physics in Medicine and Biology*, 42:1155–1173, June 1997.
6. G. Rácz and B. Csébfalvi. Tomographic reconstruction on the body-centered cubic lattice. In *Spring Conference on Computer Graphics (SCCG)*, 2015.
7. Andrew J Reader and Jeroen Verhaeghe. 4d image reconstruction for emission tomography. *Physics in Medicine and Biology*, 59(22):R371, 2014.
8. Eddy T.P. Slijpen and Reek J. Beekman. Comparison of post-filtering and filtering between iterations for SPECT reconstruction. *IEEE Trans. Nuc. Sci.*, 46(6):2233–2238, 1999.
9. Donald L. Snyder and M.I. Miller. The use of sieves to stabilize images produced with the em algorithm for emission tomography. *IEEE Transactions on Nuclear Science*, 32(5):3864–3872, Oct 1985.
10. Donald L. Snyder, M.I. Miller, Lewis J. Thomas, and D.G. Politte. Noise and edge artifacts in maximum-likelihood reconstructions for emission tomography. *IEEE Transactions on Medical Imaging*, 6(3):228–238, Sept 1987.
11. L. Szirmay-Kalos, M. Magdics, and B. Tóth. Multiple importance sampling for PET. *IEEE Trans Med Imaging*, 33(4):970–978, 2014.
12. L. Szirmay-Kalos, M. Magdics, B. Tóth, and T. Bükki. Averaging and Metropolis iterations for positron emission tomography. *IEEE Trans Med Imaging*, 32(3):589–600, 2013.
13. László Szirmay-Kalos and Ágota Kacsó. Regularizing direct parametric reconstruction for dynamic pet with the method of sieves. In *Molecular Imaging Conference*, MIC '16, pages M16D–1, 2016.
14. László Szirmay-Kalos, Milán Magdics, and Balázs Tóth. Volume enhancement with externally controlled anisotropic diffusion. *The Visual Computer*, pages 1–12, 2016.
15. Eugene Veklerov and J. Llacer. The feasibility of images reconstructed with the methods of sieves. *IEEE Transactions on Nuclear Science*, 37(2):835–841, Apr 1990.
16. G. Wang and J. Qi. Direct estimation of kinetic parametric images for dynamic pet. *Theranostics*, 3(10):802–815, 2013.



(a) ΔF (b) Δx (c) $\sum x$

Figure 3: The results of the frame merging process at different stages, using the bottom–up approach: starting from 100 frames and merging the two consecutive frames with the smallest (a) ΔF value, (b) Δxs value and (c) $\sum x$ value.



(a) (b) Uniform

(c) ΔF (d) Δx (e) $\sum x$

Figure 4: (a) Error of the time activity functions during the iterations. (b-e) The reconstructed parametric time activity curves using the obtained 5 time frames from the different frame definition methods. Averages and the standard deviations are depicted.

17. Guobao Wang and Jinyi Qi. An optimization transfer algorithm for nonlinear parametric image reconstruction from dynamic pet data. *IEEE Trans Med Imaging*, 31(10):1977–1988, 2012.

18. Jianhua Yan, Beata Planeta-Wilson, and R.E. Carson. Direct 4d list mode parametric reconstruction for pet with a novel em algorithm. In *Nuclear Science Symposium Conference Record, 2008. NSS '08. IEEE*, pages 3625–3628, Oct 2008.