Fully 4D List-Mode Reconstruction with Temporal Basis Function Estimation

Andrew J. Reader, Florent C. Sureau, Claude Comtat, Régine Trébossen and Irène Buvat

***Abstract*– A fully 4D image reconstruction algorithm for emission tomography is proposed, which regards the temporal sequence of 3D images as a single 4D image composed of 4D basis functions. Each 4D basis function is factorised into a 3D spatial basis function and a 1D temporal basis function. The reconstruction algorithm estimates both a set of temporal basis functions as well as the corresponding set of weights for each of these functions for each spatial position in the field of view. The method is designed and evaluated within the framework of a direct 4D reconstruction of time-dependent list-mode positron emission tomography (PET) data. The advantages of the proposed method include: a) a notable reduction in the number of parameters to be estimated in the reconstruction, and b) every single image in the time series can potentially benefit from the acquired data over all points in time (depending on the temporal extent of the estimated time-basis functions). Evaluation of the method with simulated 4D Monte Carlo list-mode data indicates that this method has markedly lower spatial and temporal noise compared to the conventional independent time-frame reconstruction approach.**

I. Introduction

THE reconstruction of 4D PET data normally involves the use of *a priori* temporal information. For example, the conventional approach to reconstructing a time series of images in emission tomography is to consider each time frame independently from the others. In such a case the assumed prior information is that the activity does not change within the time duration of each frame (or, alternatively viewed, the reconstruction is regarded as giving the time-averaged radioactivity concentration for a specified time window). This conventional independent frame-by-frame approach suffers from a substantial reduction in image statistics for each time frame, related to the duration of the time frame. Recently, more advanced methods have been proposed e.g. [1-3], demonstrating the considerable improvement that can be obtained through removal of the assumption of time-frame independence. The approach proposed here is a more general formulation for direct 4D reconstruction, whereby temporal basis functions, possibly extending across all time points, are estimated as part of the reconstruction process – along with the amplitudes/weights of the basis functions for the final 4D reconstructed image. This approach is in contrast to all previous dynamic reconstruction methods which adopt *a priori* temporal basis functions. The method described here can of course be simplified to include fixed temporal basis functions, either selected *a priori* (such as a limited set of Fourier basis functions), or *a posteriori* (such as time activity curves from factor analysis [4], spectral analysis [5] or even a dictionary of predetermined basis functions [6]). Hence the proposed algorithm implicitly includes, as a special case, the conventional frame by frame reconstruction method (choice of top-hat functions as the time bases), as well as the model of Nichols *et al*. [1] (choice of spline basis functions).

Manuscript received November 11, 2005.

A.J. Reader is with the School of Chemical Engineering and Analytical Science (SCEAS) at the University of Manchester, Manchester, PO BOX 88, M60 1QD, United Kingdom (e-mail: A.J.Reader@manchester.ac.uk).

F. C. Sureau, C. Comtat and R. Trébossen are with the Service Hospitalier Frédéric Joliot, CEA/DSV/DRM, Orsay, France

F. C. Sureau is also with Siemens Medical Solutions Saint-Denis, France.

I. Buvat is with UMR 678 INSERM - UPMC, CHU Pitié-Salpêtrière, Paris, France

II. Theory

A. Model of the 4D Image

This work models the space (**r**) - time (t) (4D) distribution of the radiotracer as a superposition of 4D basis functions $g(\mathbf{r},t)$, with each basis function factorised according to:

$$g(\mathbf{r},t) = s(\mathbf{r})\tau(t) \tag{1}$$

where $s(\mathbf{r})$ is a 3D spatial basis function (such as a voxel) and $\tau(t)$ is a 1D temporal basis function. Note that equation (1) is relatively general, and this model incorporates a number of other techniques as special cases, simply through the appropriate choice of $\tau(t)$. Based on equation (1), the complete space-time (4D) radiotracer distribution is represented by

$$f(\mathbf{r},t) = \sum_{j=1}^{J}\sum_{c=1}^{C} w_{jc} s(\mathbf{r}-\mathbf{r}_j)\tau_c(t) \tag{2}$$

where the j=1…J spatial basis functions $s(\mathbf{r})$ located at $\mathbf{r}_1 \ldots \mathbf{r}_J$ (covering the field of view (FOV)) are shift-invariant, and for the location j of each one of these spatial basis functions there is a collection of c=1…C different temporal basis functions $\tau_c(t)$. Hence, the reconstruction task becomes one of estimating the set of weights $\{w_{jc}\}$ in equation (2), i.e. finding the amplitude of each temporal basis function τ_c for each spatial location j. If voxels are used, along with a set of sampled temporal basis functions (of T samples each), equation (2) can be written as:

$$\mathbf{f} = \mathbf{B}\mathbf{w} \tag{3}$$

where $\mathbf{f}$ is a JT–D vector containing the time variation (at time points t=1…T) of activity for each voxel j=1…J, $\mathbf{B}$ is a $JT \times JC$ matrix containing the temporal basis functions (repeated J times) and $\mathbf{w}$ is a JC–D vector consisting of the C amplitudes for each temporal basis function for each of the J spatial locations. The precise form of $\mathbf{B}$ is as follows:

$$\mathbf{B} = \begin{bmatrix} \boldsymbol{\Gamma} & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Gamma} & \mathbf{0} & \ldots \\ \ldots & \mathbf{0} & \ldots & \ldots \\ \mathbf{0} & \ldots & \ldots & \boldsymbol{\Gamma} \end{bmatrix} \quad (4)$$

where $\boldsymbol{\Gamma}$ is a $T \times C$ matrix containing $c=1 \ldots C$ columns, each column being a T-D vector holding a temporal basis function:

$$\boldsymbol{\Gamma} = \begin{bmatrix} \boldsymbol{\tau}_1 & \boldsymbol{\tau}_2 & \ldots & \boldsymbol{\tau}_C \end{bmatrix}, \quad (5)$$

and each $\mathbf{0}$ in (4) is a $T \times C$ matrix of zero values. Hence $\mathbf{B} = \{b_{jc}\}_{JT \times JC}$.

B. Estimation of Weights for 4D Image Reconstruction

An expected time-dependent projection data vector of IT dimensions (I lines of response (LORs) for T time points) $\mathbf{q}$ can thus be obtained by

$$\mathbf{q} = \mathbf{ABw} \quad (6)$$

where $\mathbf{A}$ is an $IT \times JT$ matrix used to model the measurement process of the scanner (the probability that an emission from voxel j is detected in LOR i) for each time point t (normally the elements of $\mathbf{A}$ are regarded as time-invariant). Modeling the measured data vector $\mathbf{p}$ (projection data or list-mode data) as Poisson distributed, the maximum likelihood – expectation maximization (ML-EM) algorithm [7] can be applied to find the weights vector $\mathbf{w}$ (i.e. to estimate the amplitudes $\{w_{jc}\}$ for each temporal basis function τ_c at each spatial location j):

$$\mathbf{w}^{k+1} = \frac{\mathbf{w}^k}{\mathbf{B}^T\mathbf{A}^T\mathbf{1}} \mathbf{B}^T\mathbf{A}^T \left\{ \frac{\mathbf{p}}{\mathbf{ABw}^k} \right\} \quad (7)$$

C. Estimation of Basis Functions

In algorithm (7) the matrix $\mathbf{B}$ containing repeat copies of the matrix of temporal basis functions, $\boldsymbol{\Gamma}$, was held constant and assumed to be known. In fact this is a key issue with algorithm (7), as in general the time bases are not known *a priori* (although they could be obtained a posteriori from initial estimates from an independent reconstruction, or from a factor analysis for example). To obviate this issue entirely, an alternating updating scheme is proposed. The following algorithm estimates the time basis functions themselves whilst holding the amplitudes/weights vector $\mathbf{w}$ constant. The matrix-vector version of equation (2) for this case will be written as

$$\mathbf{f} = \mathbf{Mt} \quad (8)$$

where $\mathbf{t}$ is a TC-D vector, holding the C temporal basis functions (each of T samples), and so $\mathbf{t}$ is simply obtained by converting the $T \times C$ matrix $\boldsymbol{\Gamma}$ into a single vector. $\mathbf{M}$ is the $JT \times TC$ "mixing" matrix:

$$\mathbf{M} = \begin{bmatrix} diag\{w_{(1,1)}\} & diag\{w_{(1,2)}\} & \ldots & diag\{w_{(1,C)}\} \\ diag\{w_{(2,1)}\} & diag\{w_{(2,2)}\} & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ diag\{w_{(J,1)}\} & \ldots & \ldots & diag\{w_{(J,C)}\} \end{bmatrix} \quad (9)$$

where $\text{diag}\{w_{(j,c)}\}$ is defined as a square $T \times T$ diagonal matrix with the central diagonal elements all set equal to the scalar value w_{jc}, and all other elements zero. Expected 4D data can hence be obtained by

$$\mathbf{q} = \mathbf{AMt} \quad (10)$$

The EM algorithm for updating the estimates of the C time basis functions is given by

$$\mathbf{t}^{k+1} = \frac{\mathbf{t}^k}{\mathbf{M}^T\mathbf{A}^T\mathbf{1}} \mathbf{M}^T\mathbf{A}^T \left\{ \frac{\mathbf{p}}{\mathbf{AMt}^k} \right\} \quad (11)$$

Equation (11) uses *all* of the measured 4D data to update just the $T \times C$ parameters contained in the C different basis functions in $\mathbf{t}$. Initial values for $\mathbf{t}^0$ were chosen to correspond to a set of equally spaced broad Gaussian functions.

Algorithms (7) and (11) can be applied in an interleaving fashion, whereby a set number of iterations of algorithm (7) are first applied, then a set number of iterations of algorithm (11) (to form one 'cycle' of the algortihm) and so on, in an alternating manner. The time basis functions found from equation (11) can be regularized using an inter-update filter:

$$\mathbf{t}_{REG}^{k+1} = (1-\beta)\mathbf{t}^{k+1} + \beta\mathbf{t}_{SMOOTH}^{k+1} \quad (12)$$

where $0 \leq \beta \leq 1$ controls the level of regularization, and $\mathbf{t}^{k+1}_{SMOOTH}$ is obtained by applying a 3 point box-car filter to each time basis function in $\mathbf{t}^{k+1}$.

III. Methods

The new approach to 4D reconstruction was assessed using simulated phantom data, and compared with the conventional independent time-frame reconstruction method. Four versions of the direct 4D reconstruction method were considered in this evaluation. First, the direct reconstruction using the exact temporal basis functions (matching perfectly those used in the generation of the simulated data) was assessed. For this case only the weights were estimated in reconstruction (algorithm (7) only). Second, the direct method which estimates both the temporal basis functions themselves as well as the weights, and thirdly, the direct method with estimation of an "excessive" number of basis functions (in this case 8 were estimated when in fact 4 would suffice). This case was considered as in practice the number of temporal basis functions is unknown, and so it is more appropriate to overestimate the number rather than risk a misrepresentation of the 4D image. Finally, this last case was also tested with the regularization of equation (12).

A. Simulation Phantom

A simple Monte Carlo simulation for the High Resolution Research Tomograph (HRRT) [8] was created to generate list-mode data, incorporating the scanner geometry (with detector gaps), but excluding effects such as attenuation, scatter and randoms. This permitted an initial assessment of the direct method and its performance relative to the independent method, without the further complication of data corrections. The phantom used for evaluation of the algorithms was a cylinder of length 50 mm, diameter 62 mm, with an additional outer rim of 5 mm thickness (total cylinder diameter 72 mm). The simulated radioactivity in the main cylinder body followed a time activity curve corresponding to brain white matter, and the outer rim followed a time activity curve corresponding to gray matter. The cylinder contained 6 spheres of radius 2.5 mm (3 of which were cold spheres (representing ventricles), and 3 of which followed a time

activity curve corresponding to an arterial time-activity curve) and 6 spheres of radius 5 mm (again, 3 were cold, and 3 followed an arterial time-activity curve). Finally, the phantom included two line sources of length 25 mm (one parallel to the scanner axis, displaced by 22.5 mm, and one perpendicular to the axis, displaced by 8 mm from the central transverse plane) and 49 point sources located within a transverse plane (displaced by 8 mm from the central plane). The activity of the line and point sources was held constant in time. Example slices from the phantom can be seen in Fig. 1. Ten million events were generated over a time interval of 5 minutes (simulating a mean count rate of 33 kcps). This short time scale and phantom were selected to allow a more computationally practical means of assessing the performance of the 5 methods. Images were reconstructed into 64×64×50 matrices (1.21875 mm voxel side length) for 50 time frames (interval 6 seconds). For the independent reconstruction method this gave a single 4D image of 64×64×50×50, whereas for the direct method this gave a set of C parametric images, each of size 64×64×50, where C corresponds to the chosen number of temporal basis functions. Since 4 different types of temporal behaviour were simulated (arterial, gray matter, white matter and constant time-activity curves), a minimum of 4 temporal basis functions were used in the direct reconstruction ($C \geq 4$).

B. *Evaluation and Figures of Merit (FOMs)*

The spatial properties of the reconstruction methods were evaluated for the mid-point time frame 25 (2.5 minutes). The image for this point in time was used to assess background noise level, using a simple spatial standard deviation FOM, which calculated the standard deviation of the voxel values belonging to the background region (white matter) of the cylinder, divided by the mean value. The resolution of the two line sources was also assessed, by summing all the profiles for all the length of each line source. This gave a full width at half maximum (FWHM) resolution measure for the line source. These spatial noise and spatial resolution measures were performed on a number of images, obtained for two iteration numbers, and a range of post-smoothing levels applied post-reconstruction. The temporal properties of the images were assessed using a bias FOM, evaluated for each region of interest (ROI) (white matter (background), gray matter (rim), artery (sphere)). This FOM simply summed the differences between the measured time activity curve (TAC) and the corresponding true TAC, normalizing for the number of samples. The noise present in the measured TAC for each image region was assessed visually, plotting the TAC obtained for each method for each region.

IV. Results

Fig. 1 shows example transverse slices (from 3D images) from the true, the independent reconstruction and the direct reconstructions. The direct methods show visually reduced image noise. The similarity between the time-frame summed slices and the single time frame slices for the direct methods arises from the use of parametric images which give the weights for time basis functions extending through all time frames. Fig. 2 shows example profiles, indicating the noise reduction achieved by the direct method. Fig. 3 gives the resolution-noise trade off for all methods, for different iteration numbers and post-smoothing levels. It is clear that the direct reconstruction method consistently gives markedly lower spatial noise levels for a given spatial resolution when compared to the independent time frame reconstruction. This advantage is lost to some extent when 8 time basis functions are estimated (whilst only 4 are needed), but the advantage is recovered through incorporation of regularization. Fig. 4 shows the TACs for all methods for all regions – the direct method clearly outperforms the independent approach. Fig. 5 shows the bias levels for the various methods, indicating that the direct approaches give lower bias for all regions of interest.

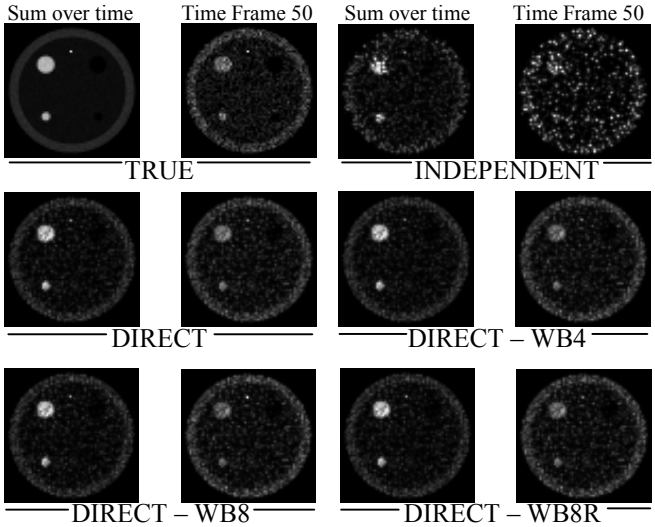

Fig. 1. Transverse slices of the 3D true (sites of positron emission), and for five different reconstruction methods. The first the third columns show slice 25 summed over all 50 time frames, the second and fourth columns give slice 25 for time frame 50 only. "DIRECT" indicates that only the weights were estimated, using the exact TACs from the simulation as the basis functions. "WB4" indicates estimation of both the weights for each voxel, as well as the 4 basis functions associated with these weights. "WB8" indicates estimation of weights and 8 basis functions, and "R" signifies that regularization has been included (β=0.25 in this case).

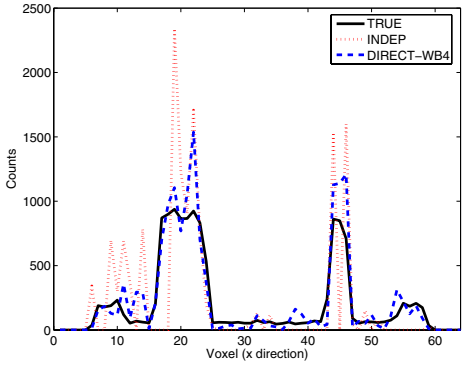

Fig. 2. Profile through the true, the independent and the DIRECT-WB4 reconstruction. Note that the background is almost a cold region, resulting in bias in this region for the noisy independent reconstruction when reconstructed with the positivity constraint of the EM algorithm.

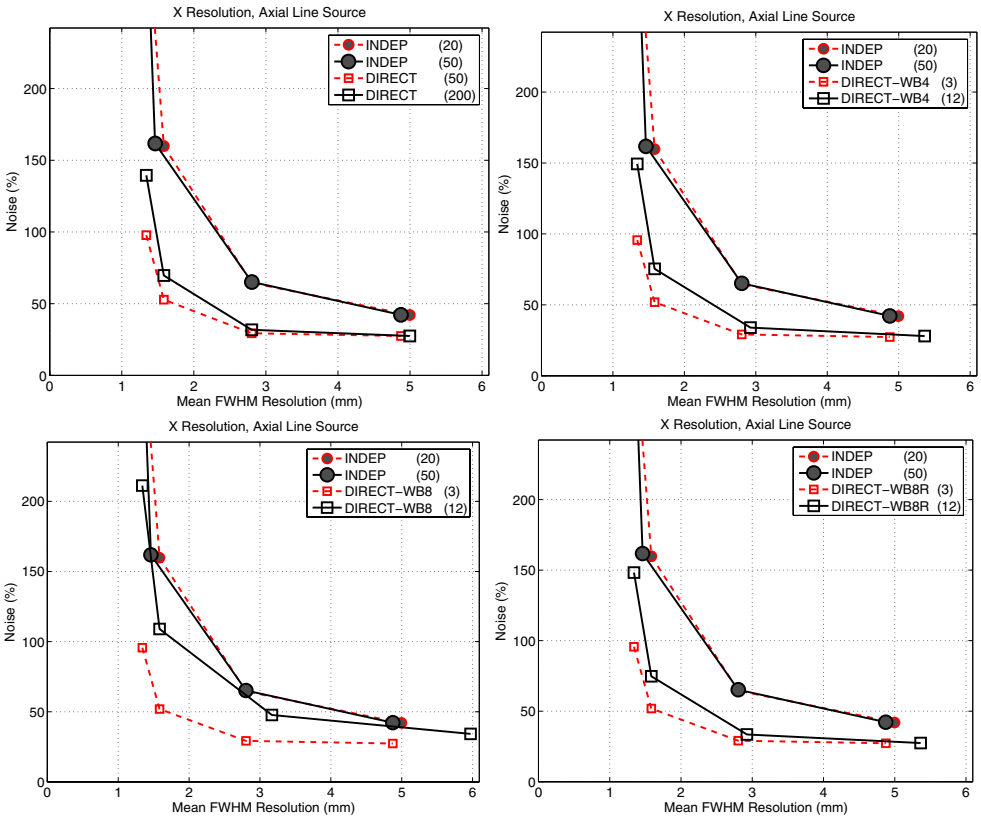

Fig. 3. Noise-resolution trade off for the independent time frame reconstruction method (INDEP, same in all graphs, shown for iterations 20 and 50, for 4 different levels of post-reconstruction smoothing (3D Gaussian)), and for the direct reconstruction methods (DIRECT – estimated only the weights, using perfect time bases, as used by the simulation data, DIRECT-WB4 estimated the weights as well as 4 time bases, DIRECT-WB8 estimated 8 time bases and their weights, DIRECT-WB8R included regularization of the time basis functions). For the DIRECT-WB methods, the number in brackets indicates the number of repeats of the weights-then-time-basis estimation cycle (each cycle consisting of 16 iterations for the weights, and 16 iterations for the time bases).

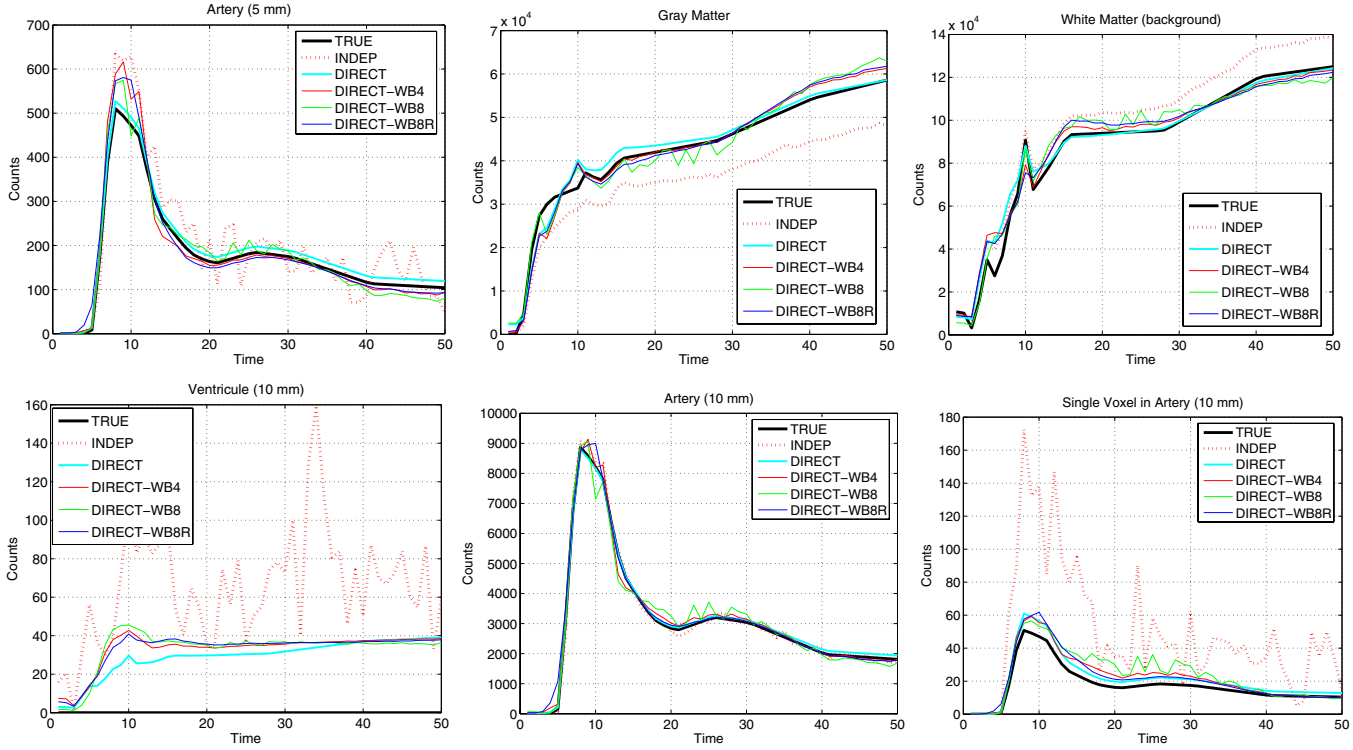

Fig. 4. TACs for all of the ROIs considered. The lower-right graph shows the result for a single voxel ROI located within one of the spheres, showing the substantial noise reduction of the direct method in comparison to the independent method. The TACs shown correspond to those from 50 iterations of the independent method, 200 iterations of the direct method, and 12 cycles (each of 16 iterations for both the weights and the bases) for the other direct methods.

Figs. 6 and 7 show the parametric images and corresponding estimated temporal basis functions for the direct method when using 8 basis functions and including regularization (DIRECT-WB8R, with β=0.25). The key features of the phantom (the line source, the spheres, the rim) are identified (i.e. separated) by the method. The eighth basis function is effectively not used by the algorithm in this particular case: both the estimated parametric image and basis function are of negligible amplitude.

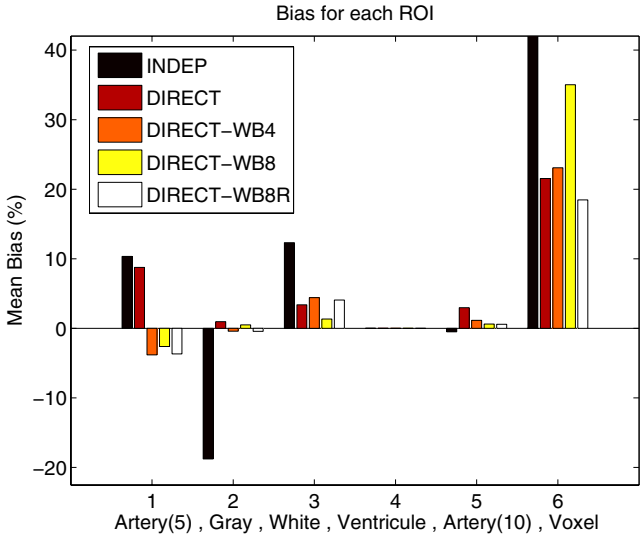

Fig. 5. Mean Bias. For the voxel ROI, the independent reconstruction method has a mean bias of 173.13%. Note that for ROI number 4, the ventricle, a mean bias % is not available as the true is zero valued.

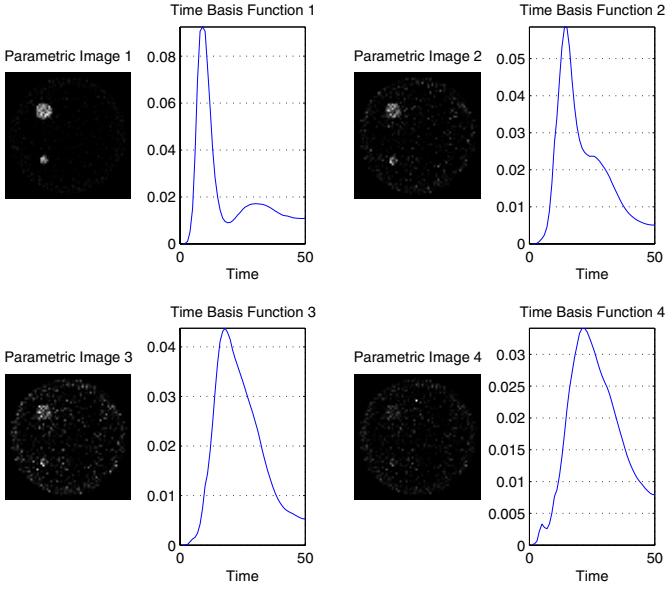

Fig. 6. First four parametric images and corresponding estimated time basis functions for the DIRECT-WB8R method after 12 cycles of weights/bases estimation (8 time bases estimated, regularized with β=0.25).

V. Conclusion

A new approach to 4D reconstruction is presented, which models the 4D image as a set of 4D basis functions, where both the amplitude/weight of the temporal component of each basis function is estimated, as well as all the basis functions themselves. The method permits each time frame to benefit from substantially more of the measured data than the conventional independent time frame approach, leading to notable image quality advantages in both the space and time dimensions. This benefit is drawn from the assumption that a limited number of time basis functions is sufficient to describe the temporal behaviour of the radioactivity in each voxel.

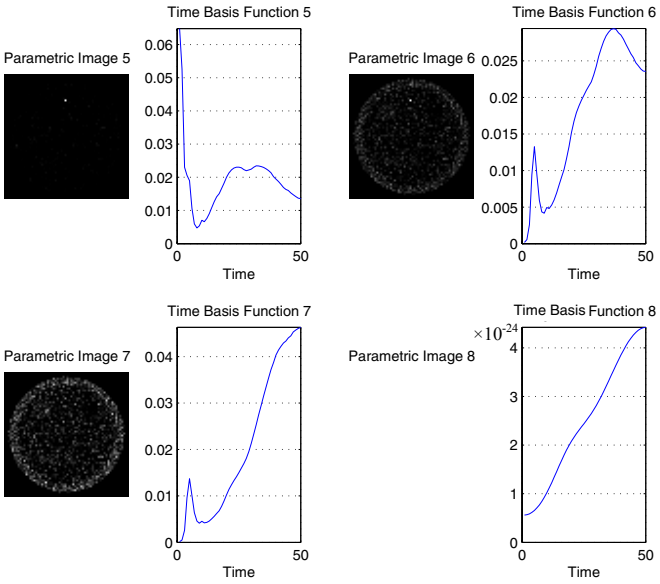

Fig. 7. Last four parametric images and corresponding estimated time basis functions for the DIRECT-WB8R method after 12 cycles of weights/bases estimation (8 time bases estimated in total, regularization included, β=0.25). Note that the 8th image and basis function are effectively not used (the scale for basis function 8 is 10^{-24}).

References

[1] T. E. Nichols, J. Qi, E. Asma, and R. M. Leahy "Spatiotemporal Reconstruction of List-mode PET Data" *IEEE Trans. Med. Im.,* vol. 21, pp.396-404, 2002.

[2] R.J. Walledge, R. Manavaki, M. Honer, and A.J. Reader "Inter-Frame Filtering for List-Mode EM Reconstruction in High Resolution 4D PET" *IEEE Trans. Nucl.. Sci.* vol. 51, pp. 705 – 711, 2004.

[3] M.E. Kamasak, C. A. Bouman, E. D. Morris, K. Sauer "Direct reconstruction of kinetic parameter images from dynamic PET data" *IEEE Trans. Med. Im.* vol. 24, pp. 636-50, 2005.

[4] K. S. Nijran and D. C. Barber "Towards automatic analysis of dynamic radionuclide studies using principal principal components factor analysis" *Phys. Med. Biol*. vol. 30, pp. 1315–25, 1985.

[5] S. R Meikle, J. C. Matthews, V. J. Cunningham, D. L. Bailey, L. Livieratos, T. Jones, and P. Price "Parametric image reconstruction using spectral analysis of PET projection data" *Phys. Med. Biol.* vol. 43, pp. 651-666, 1998.

[6] R. N. Gunn, S. R. Gunn, F. E. Turkheimer, J. A. D. Aston, and V. J. Cunningham "Positron emission tomography compartmental models: A basis pursuit strategy for kinetic modeling" *J. Cer. Blood Flow and Metab.* vol. 22, pp. 1425-1439, 2002.

[7] L.A. Shepp and Y. Vardi "Maximum Likelihood reconstruction in positron emission tomography" *IEEE Trans. Med. Im.,* vol. 1, pp.113-122, 1982.

[8] K. Wienhard, M. Schmand, M.E. Casey, K. Baker, J. Bao, L. Eriksson, W.F. Jones, C. Knoess, M. Lenox, M. Lercher, P. Luk, C. Michel, J.H. Reed, N. Richerzhagen, J. Treffert, S. Vollmar, J.W. Young, W.D. Heiss, R. Nutt "The ECAT HRRT: Performance and First Clinical Application of the New High Resolution Research Tomograph," *IEEE Trans. Nucl. Sci.*, vol. 49, pp. 104-110, 2002.