STATISTICAL METHODS FOR RECONSTRUCTION OF PARAMETRIC IMAGES FROM DYNAMIC PET DATA

A Thesis

Submitted to the Faculty

of

Purdue University

by

Mustafa E. Kamasak

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

May 2005

To the memory of my father, İhsan Kamaşak

ACKNOWLEDGMENTS

First of all, I would like to thank my father for inspiring and encouraging me. To be a worthwhile son, I always strive to be one the best in whatever I do in life. I am grateful to my mother and sister for their love and endless support during my graduate studies. I thank for all they have done and endured for me.

I would like to thank my advisor Prof. Charles A. Bouman for his excellent guidance, mentoring, and patience. He taught me to be meticulous and never content myself with mediocrity in my work.

Before I met Prof. Evan D. Morris all I had in my research were a bunch of equations on paper and a vague idea. He extricated us from this state and bring my work to the point reported here. I am grateful to him for all his help and effort.

I would like to thank Prof. Brad Christian for providing the real data and for his extremely useful comments and suggestions.

I would like to thank Prof. Ken Sauer, Dr. Ned Rouze, and Dr. Adam Alessio for their involvement in my work and useful discussions.

I would like to thank Prof. Jan P. Allebach, Prof. Edward J. Delp, and Prof. Bradley Lucier for being in my advisory committee and dedicating their time.

I would like to thank Cristian Constantinescu, Chunzhi Wang, Dr. Karmen Yoder and Dr. Ti-Qiang Li at Indiana University School of Medicine for their help in constructing the rat phantom from MR data and their hospitality when I visit their campus.

Last but certainly not the least, I would like to thank my colleagues Dr. Murat Dundar, Dr. Du Yong Ng, Bulent Bayraktar, Osman Arslan, Guotong Feng, Maribel Figuera, Dr. Seungseok Oh and many others who are not listed here for their friendship, support, and discussions.

ABBREVIATIONS

BIG	bismuth germanate
BP	binding potential
Ci	Curie
CPU	central processing unit
CSF	cerebral spinal fluid
dpi	dots per inch
dSPECT	dynamic single photon electronic computed tomography
DSP	digital signal processing
DVR	distribution volume ratio
EM	expectation maximization
EMPIRA	expectation maximization parametric image reconstruction algorithm
FBP	filtered back projection
FDG	2-deoxy-2-[^{18}F]fluoro-D-glucose
FORE	Fourier rebinning
FOV	field of view
GMRF	Gaussian Markov random field
ICA	independent component analysis
ICD	iterative coordinate descent
MAP	maximum a posteriori
ML	maximum likelihood
ml	milliliter
MRF	Markov random field
MRI	magnetic resonance imaging

MSE	mean squared error
NRMSE	normalized root mean squared error
ODE	ordinary differential equation
OSEM	ordered subset expectation maximization
PA	paper advance
PCA	principal component analysis
PET	positron emission tomography
PICD	parametric iterative coordinate descent
pmol	picomole
PWLS	pixelwise weighted least squares
PWLSR	pixelwise weighted least squares with regularization
PWLSZ	pixelwise weighted least squares with Zhou regularization
PSF	point spread function
PTAC	plasma time activity curve
RMSE	root mean squared error
ROI	region of interest
SA	specific activity
SIME	simultaneous estimation of kinetic parameters and plasma
SNR	signal to noise ratio
SPECT	single photon electronic computed tomography
TAC	time activity curve
TTAC	tissue time activity curve
VD	volume distribution
WLS	weighted least squares

TABLE OF CONTENTS

Page

Page

LIST OF TABLES

LIST OF FIGURES

Figure Page

ABSTRACT

Kamasak, Mustafa E. Ph.D., Purdue University, May, 2005. Statistical Methods for Reconstruction of Parametric Images from Dynamic PET Data. Major Professor: Charles A. Bouman.

In this thesis, we demonstrate and evaluate a new and efficient method for direct voxel-wise reconstruction of kinetic parameter images using all frames of the PET data. The direct parametric image reconstruction is formulated in a Bayesian framework, and uses the parametric iterative coordinate descent (PICD) algorithm to solve the resulting optimization problem. The PICD algorithm is computationally efficient and is implemented with spatial regularization in the domain of the physiologically relevant parameters. Our experimental simulations of a rat head imaged in a working small animal scanner indicate that direct parametric reconstruction can substantially reduce estimation error in kinetic parameters, as compared to indirect methods, without appreciably increasing computation.

The PICD algorithm is also used to reconstruct the parametric images of a monkey brain directly from the EXACT HR+ sinograms. PICD reconstructions are compared with the estimates using image domain methods. The results have shown that the proposed direct reconstruction algorithm can produce higher resolution parametric reconstructions.

We simultaneously estimate both the kinetic parameters at each voxel and the model-based plasma input function directly from the sinogram data. The plasma model parameters are initialized with an image domain method to avoid local minima, and multiresolution optimization is used to perform the required reconstruction. Good initial guesses for the plasma parameters are required for the algorithm to converge to the correct answer. This method can estimate some of the kinetic

parameters (k_2, k_3, k_4, BP), but it can only estimate the others (k_1, VD) within a scale factor.

1. DIRECT RECONSTRUCTION OF KINETIC PARAMETER IMAGES FROM DYNAMIC PET DATA

Abstract

Our goal in this chapter is the estimation of kinetic model parameters for each voxel corresponding to a dense 3D PET image. Typically, the activity images are first reconstructed from PET sinogram frames at each measurement time, and then the kinetic parameters are estimated by fitting a model to the reconstructed time-activity response of each voxel. However, this "indirect" approach to kinetic parameter estimation tends to reduce signal-to-noise ratio (SNR) because of the requirement that the sinogram data be divided into individual time frames.

In 1985, Carson and Lange proposed [1], but did not implement, a method based on the expectation-maximization (EM) algorithm for direct parametric reconstruction. The approach is "direct" because it estimates the optimal kinetic parameters directly from the sinogram data, without an intermediate reconstruction step. However, direct voxel-wise parametric reconstruction remained a challenge due to the unsolved complexities of inversion and spatial regularization.

In this work, we demonstrate and evaluate a new and efficient method for direct voxel-wise reconstruction of kinetic parameter images using all frames of the PET data. The direct parametric image reconstruction is formulated in a Bayesian framework, and uses the parametric iterative coordinate descent (PICD) algorithm to solve the resulting optimization problem [2]. The PICD algorithm is computationally efficient and is implemented with spatial regularization in the domain of the physiologically relevant parameters. Our experimental simulations of a rat head imaged in a working small animal scanner indicate that direct parametric reconstruc-

tion can substantially reduce estimation error in kinetic parameters, as compared to indirect methods, without appreciably increasing computation.

1.1 Introduction

Positron Emission Tomography (PET) is a powerful molecular imaging technique with the sensitivity to detect picomolar quantities of a labelled tracer with reasonable (seconds to minutes) temporal resolution. Through the application of kinetic models, the dynamic PET data can be transformed into physiological parameters that indicate the functional state of the imaged tissue. Ideally, one would like to reconstruct parametric images from PET data (i.e., images which specify the estimated kinetic parameters for each voxel in the imaged volume). Such parametric images could serve many uses. For example, they may be particularly desirable when testing a new tracer whose sites of action are not completely known. In the brain, parametric images might be useful in identifying new brain circuits or discovering unsuspected connectivity between disparate brain regions. As new tracers continue to be developed with greater specific to nonspecific binding ratios, the impetus grows for looking at their uptake in all regions of the brain, rather than in a few pre-selected regions of interest. For example, $[^{18}F]$-fallypride, a high affinity dopamine tracer, can be used to image dopamine receptors outside the striatum because the non-specific background is low [3].

This chapter introduces a novel algorithm for directly reconstructing parametric images from PET sinogram data. We demonstrate that this method can generate parametric images with superior quality; and, perhaps surprisingly, we also show that it has computational requirements that are similar to a two-step approach of iterative reconstruction followed by kinetic parameter estimation.

Kinetic compartmental models are often used to describe the movement of a tracer between different physically or chemically distinct states or compartments [4–6]. The exchange of tracer between these compartments can be modeled by a system

of first order ordinary differential equations (ODEs) whose coefficients are the kinetic parameters. The resulting kinetic models have been validated as producing reliable quantitative indices of various clinically and scientifically important physiological processes [7–17].

In some cases, a single set of kinetic parameters can describe the tracer behavior in a homogeneous region of tissue such as the myocardium or perhaps the entire striatum in brain images. If the region of interest can be delineated using some form of segmentation, then the PET activity can be averaged over the region at each time frame and a single set of kinetic parameters can be estimated by fitting a single kinetic model to the time sequence of average activities. This case is illustrated in Fig. 1.1. The PET data are first reconstructed into K time frames, then a region of interest (ROI) is segmented from each frame, and a single set of kinetic parameters is fit to the regional-average time sequence. These ROI-based methods may be further classified into linear methods and nonlinear methods. Linear techniques [18–23] transform the data, so that the parameters of interest can be estimated by linear regression methods, while nonlinear techniques [5, 6] generally estimate the kinetic parameters by iteratively minimizing a properly weighted distance metric between the reconstructed time-activity curves and the model output.

Recently, there has been increasing interest in the formation of parametric images which model the kinetic behavior of each voxel individually. This approach is more appropriate when the volume cannot be effectively segmented into homogeneous regions that would be modeled with a single kinetic parameter set. Existing approaches to the creation of parametric images can be roughly categorized as "indirect", "semi-direct", and (our new method) "direct" [24] reconstruction.

Indirect methods work by first reconstructing the PET emission images for each of the K measurement times, and then estimating the kinetic parameters at each voxel. The primary difficulty of the indirect approach is that the low signal-to-noise ratio of the time-activity curve (TAC) for each voxel makes accurate estimation of parameters difficult. To improve estimation accuracy, O'Sullivan *et al.* [25] applied

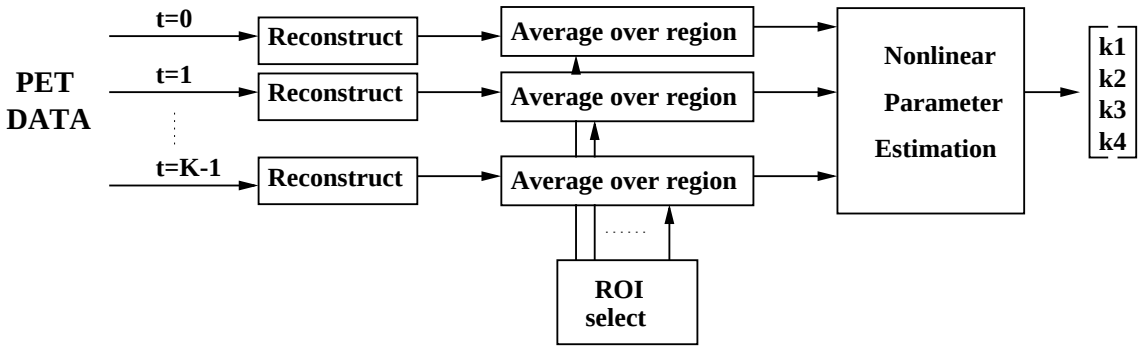


Fig. 1.1. ROI-based kinetic parameter estimation methods.

ridge regression techniques to regularize the parameters using prior knowledge of their means and variances derived from the analysis of a reasonably large patient group. Huang *et al.* [26] applied a spatial smoothing step between the iterations of a nonlinear estimation process at each voxel. Later, Zhou *et al.* [27], [28] developed a two-stage algorithm whereby the kinetic parameters were estimated first using standard nonlinear techniques. In a second step the initial results were smoothed spatially and used to constrain the final estimates. (This method is further discussed in section 1.4.2.) Kimura *et al.* [29] and Zhou *et al.* [30] have developed algorithms that cluster the images before estimation and regularize the data within the clusters.

Semi-direct algorithms, as they are sometimes named, attempt to improve signal-to-noise by constraining the possible choices of time-courses for each voxel via signal sub-spaces or splines. Kao *et al.* [31] and Narayanan *et al.* [32] used principal component analysis (PCA) to decorrelate the sinograms in time. Similarly, Wernick *et al.* [33] applied PCA decomposition of PET data followed by reconstruction of tracer concentrations in the principal-component space. Nichols *et al.* [34], and Reutter *et al.* [35, 36] proposed reconstruction methods that use a b-spline specification of the time-activity curves. Kinetic parameters must then be estimated from the b-spline representation. It is important to note that spline-based methods have certain computational advantages when processing list mode data. However, the two-step process of first computing spline coefficients and then kinetic parameters still results in

a loss of optimality, particularly if the number of spline coefficients is much larger than the number of kinetic parameters.

Ideally, one would like to estimate directly the space-domain kinetic parameters from the measured sinogram data. In fact, Carson and Lange [1] proposed direct estimation of kinetic parameters from PET data in 1985. In that paper, the authors outlined a general framework for a direct reconstruction algorithm based on expectation-maximization (EM) [37] iterations. Unfortunately, the Carson and Lange direct parametric reconstruction algorithm has never, to our knowledge, been fully implemented for nonlinear estimation of a dense set of voxels. Limber *et al.* [38] proposed an algorithm for direct parametric reconstruction using maximum likelihood (ML) estimation of kinetic parameters from PET data, but only demonstrated the algorithm for an 8×8 array of voxels. A number of authors have implemented direct nonlinear parameter estimation methods that were designed for segmented ROI's [39–44] rather than a dense set of voxels. In an alternative approach, Meikle *et al.* [45] first pre-computed the time-activity curves for a range of possible nonlinear parameters and then directly reconstructed the linear weights for each of the nonlinear "basis" curves. Similarly, Matthews *et al.* [46] obtained pre-defined time-curves from other patients' reconstructions and used the EM algorithm to compute the weights of each curve. In other work, Carson [47] proposed an ML framework to estimate the ROI values from the projections, and Farncombe *et al.* [48] estimated organ uptake parameters that were incorporated into the reconstruction algorithm for dSPECT applications.

In this chapter, we present an algorithm for direct nonlinear estimation of space-domain kinetic parameters in a dense volume of voxels. Our novel parametric reconstruction algorithm, which we call parametric iterative coordinate descent (PICD) [2, 49] is in the spirit of Carson and Lange's method. However, PICD is a completely specified and implemented algorithm (See Appendix A) which we show to be computationally efficient with robust convergence properties. In fact, the computation required for parametric reconstruction using PICD is comparable to that

required for more conventional maximum *a posteriori* (MAP) reconstruction of an image sequence from PET sinogram data. In other words, it is our claim that direct parametric reconstruction can have comparable computational requirements to indirect methods (recall that indirect methods require an initial reconstruction of all the data). The key to computational efficiency of the PICD method is the use of state variables and nested optimization to decouple the nonlinearities of the forward tomographic model, the nonlinear kinetic model, and the Bayesian prior model. Notably, PICD is designed to compute the MAP estimate of the kinetic parameters using a prior distribution defined on any well-behaved transformation of the parameter space. This allows the regularization to be applied to the parameters that are deemed to be physiologically important. Simulation results, presented below, indicate that the PICD-generated parametric reconstructions have lower mean squared error and better visual quality than the best indirect methods.

Section 1.2 reviews the 2-tissue compartment model and the set of ODE's that govern a tracer's kinetics. Section 1.3 introduces the PICD algorithm for direct parametric reconstruction and gives a detailed description of its implementation. Section 1.4 first reviews some existing methods for image domain parameter estimation, and then suggests a useful method for regularization of pixel-wise approaches. Section 1.5 compares the computational complexity of the proposed methods. Section 1.6 presents simulation results. Discussion and Conclusion follow the results.

1.2 Two-tissue Compartment Model

In this work, we used a 2-tissue compartment model to describe the kinetic processes that are represented by the signal from each voxel of a reconstructed image. This model is commonly used to describe the uptake and retention of an analog of glucose, 2-deoxy-2-[^{18}F]fluoro-D-glucose (FDG). The model can also be properly applied to receptor ligand studies provided that there is no non-specific binding and that the tracer has been prepared at sufficiently high specific activity. Figure 1.2

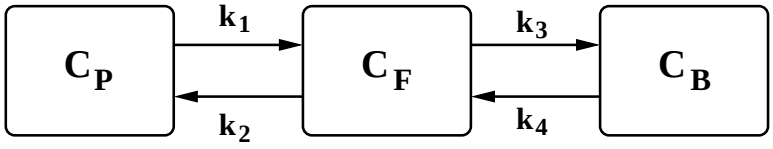


Fig. 1.2. 2 tissue compartment model with 4 kinetic parameters.

illustrates the model: C_P is the molar concentration of tracer in the plasma, C_F is the molar concentration of unbound tracer, and C_B is the molar concentration of metabolized or bound tracer. The model depends on the kinetic parameters, k_1, k_2, k_3, and k_4, which specify the tracer exchange rates between compartments in units of inverse minutes. The parameters k_1, k_2, and k_4 are first order rate constants, and k_3 is an *apparent* first order rate constant describing a process (metabolism or receptor-binding) that proceeds in proportion to the concentration of the labelled tracer only, as long as the number of sites available for binding do not become rate-limiting.

In addition to the above-stated parameters, there are two compound parameter groups that have ready physiological interpretations and practical application, particularly for receptor-ligand imaging: binding potential (BP), and total volume of distribution (VD). BP is proportional to the number of receptors and VD represents the steady state distribution of tracer between the plasma and tissue. BP and VD can be expressed in terms of the aforementioned kinetic parameters

$$BP = \frac{k_3}{k_4} \tag{1.1}$$

$$VD = \frac{k_1}{k_2}\left(1 + \frac{k_3}{k_4}\right) . \tag{1.2}$$

In applying the model in Fig. 1.2 to all voxels, we assume that the delivery of tracer is the same to all regions being imaged. In other words, the value of C_P is not a function of voxel position. However, the values of the kinetic parameters will be allowed to vary for each voxel location, s. Using these assumptions, the

time variation of the concentrations for a single voxel are governed by the following ODE's.

$$\frac{dC_F(s,t)}{dt} = k_{1s}C_P(t) - (k_{2s}+k_{3s})C_F(s,t) + k_{4s}C_B(s,t) \quad (1.3)$$

$$\frac{dC_B(s,t)}{dt} = k_{3s}C_F(s,t) - k_{4s}C_B(s,t)\ . \quad (1.4)$$

In this work, $C_P(t)$ is assumed known. In practice, it can be measured directly from arterial plasma samples during the imaging procedure [6], or it may be estimated from imaged volumes that consist primarily of blood [50–54]. The solution to the ODE's in (1.3,1.4) is given by

$$C_F(s,t) = \left\{\frac{k_{1s}}{\alpha_2-\alpha_1}[(k_{4s}-\alpha_1)e^{-\alpha_1 t} + (\alpha_2-k_{4s})e^{-\alpha_2 t}]u(t)\right\} * C_P(t) \quad (1.5)$$

$$C_B(s,t) = \left\{\frac{k_{1s}k_{3s}}{\alpha_2-\alpha_1}[e^{-\alpha_1 t} - e^{-\alpha_2 t}]u(t)\right\} * C_P(t) \quad (1.6)$$

where $*$ indicates continuous-time convolution, α_1 and α_2 are real valued constants that result from the subtraction and addition of terms in (1.7) respectively, and $u(t)$ is the unit step function.

$$\alpha_1, \alpha_2 = \frac{(k_{2s}+k_{3s}+k_{4s}) \mp \sqrt{(k_{2s}+k_{3s}+k_{4s})^2 - 4k_{2s}k_{4s}}}{2}\ . \quad (1.7)$$

Next, we transform the kinetic parameters (k_1, k_2, k_3, k_4) to form the new parameters (a, b, c, d) as shown in Table 1.1. This transformation is important because while the parameters (a, b, c, d) are well suited for optimization, (k_1, k_2, k_3, k_4) are physiologically more relevant. We use $\varphi_s = [a_s, b_s, c_s, d_s]^t$ to denote the parameter vector for each voxel s.

The total activity (e.g., in nCi/ml) for voxel s at time t is denoted by

$$\begin{aligned} f(\varphi_s, t) &\triangleq (1-V_B)\left[C_F(s,t) + C_B(s,t)\right] S_A e^{-\lambda t} + V_B C_{WB}(t) \\ &= (1-V_B)\left[(a_s e^{-c_s t} + b_s e^{-d_s t})u(t) * C_P(t)\right] S_A e^{-\lambda t} + V_B C_{WB}(t) \end{aligned} \quad (1.8)$$

where S_A is the initial specific activity of the tracer (nCi/pmol), λ is the decay rate of the isotope (min^{-1}), V_B is the volume fraction of the voxel that contains blood,

Forward Transforms	Inverse Transforms
$a_s = \frac{k_{1s}}{2\Delta}(k_{2s} - k_{3s} - k_{4s} + \Delta)$	$k_{1s} = a_s + b_s$
$b_s = \frac{k_{1s}}{2\Delta}(-k_{2s} + k_{3s} + k_{4s} + \Delta)$	$k_{2s} = \frac{a_s c_s + b_s d_s}{a_s + b_s}$
$c_s = \frac{1}{2}(k_{2s} + k_{3s} + k_{4s} + \Delta)$	$k_{3s} = \frac{a_s b_s (c_s - d_s)^2}{(a_s + b_s)(a_s c_s + b_s d_s)}$
$d_s = \frac{1}{2}(k_{2s} + k_{3s} + k_{4s} - \Delta)$	$k_{4s} = \frac{c_s d_s (a_s + b_s)}{a_s c_s + b_s d_s}$
$\Delta = \lvert\sqrt{(k_{2s} + k_{3s} + k_{4s})^2 - 4k_{2s}k_{4s}}\rvert$	

Table 1.1
Forward and inverse transformations from standard kinetic parameters $[k_{1s}, k_{2s}, k_{3s}, k_{4s}]$ for the voxel s to new parameters $[a_s, b_s, c_s, d_s]$.

and C_{WB} is the tracer concentration in whole blood (i.e., plasma plus blood cells plus other particulate matter). We can simplify the expression for $f(\varphi_s, t)$ by defining the following functions

$$\alpha(c_s, t) \triangleq \left\{C_P(t) * \left[e^{-c_s t}u(t)\right]\right\}(1 - V_B)S_A e^{-\lambda t} \tag{1.9}$$

$$\beta(d_s, t) \triangleq \left\{C_P(t) * \left[e^{-d_s t}u(t)\right]\right\}(1 - V_B)S_A e^{-\lambda t} \tag{1.10}$$

$$\gamma(t) \triangleq V_B C_{WB}(t) \tag{1.11}$$

With these definitions, $f(\varphi_s, t)$ can be written as

$$f(\varphi_s, t) = [a_s, b_s] \begin{bmatrix} \alpha(c_s, t) \\ \beta(d_s, t) \end{bmatrix} + \gamma(t) \ . \tag{1.12}$$

We next define some vector and matrix notation that will be useful in discretization of the problem. Let $t_0, \cdots, t_{K-1}$ be the K discrete times at which the tissue is imaged. Then we may construct the vectors

$$\alpha(c_s) \triangleq [\alpha(c_s, t_0), \alpha(c_s, t_1), \cdots, \alpha(c_s, t_{K-1})] \tag{1.13}$$

$$\beta(d_s) \triangleq [\beta(d_s, t_0), \beta(d_s, t_1), \cdots, \beta(d_s, t_{K-1})] \tag{1.14}$$

$$\gamma \triangleq [\gamma(t_0), \gamma(t_1), \cdots, \gamma(t_{K-1})] \ . \tag{1.15}$$

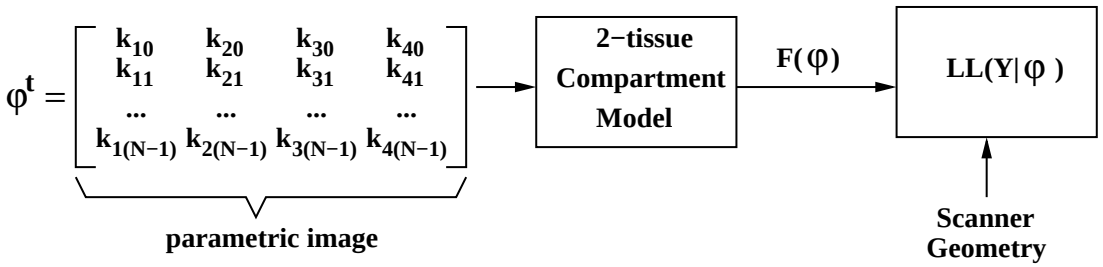


Fig. 1.3. Model used for direct parametric reconstruction of images.

Using this notation, the activity at each time for voxel s is given by the $1 \times K$ row vector

$$f(\varphi_s) = [f(\varphi_s, t_0), f(\varphi_s, t_1), \cdots, f(\varphi_s, t_{K-1})] \tag{1.16}$$

$$= [a_s, b_s] \begin{bmatrix} \alpha(c_s) \\ \beta(d_s) \end{bmatrix} + \gamma \ . \tag{1.17}$$

Let the N voxels be indexed by the values $s = 0, 1, \cdots, N-1$, and let $\varphi = [\varphi_0, \varphi_1, \cdots, \varphi_{N-1}]$ denote the $4 \times N$ matrix of parameters at all voxels. With this, we define the $N \times K$ function

$$F(\varphi) = \begin{bmatrix} f(\varphi_0) \\ \vdots \\ f(\varphi_{N-1}) \end{bmatrix}$$

which maps the parametric image, φ, to the activity of each voxel at each time. Finally, let $F(\varphi, t_k)$ denote the k^{th} column of $F(\varphi)$, so $F(\varphi, t_k)$ contains the activity for each voxel at time t_k.

1.3 Parametric Reconstruction from Sinogram Data

In this section, we describe our method for directly reconstructing the parametric image, φ, from sinogram data. We will do this by first formulating a conventional scanner model under the assumption that the sinogram measurements are Poisson random variables. We will then use the kinetic model of Section 1.2 as the input

to the scanner model as shown in Fig. 1.3. Once the complete forward model is formulated, we will present an iterative algorithm for computing the maximum a posteriori (MAP) estimate of the parametric image $\hat{\varphi}$ from the sinogram data. Once $\hat{\varphi}$ is computed, the activity images can be computed at any time t simply by evaluating $F(\varphi, t)$ using the kinetic model equations of (1.8).

1.3.1 Scanner Model

Let Y_{mk} denote the sinogram measurement for projection $0 \leq m < M$ and time frame $0 \leq k < K$, and let Y be the $M \times K$ matrix of independent Poisson random variables that form the sinogram measurements. Furthermore, let A be the forward projection matrix, with elements A_{ms} (counts-ml/nCi), and let μ be the number of accidental coincidences. Then the expected number of counts for each measurement at a given time, t_k is given by

$$E[Y_{mk}|F(\varphi, t_k)] = \sum_{s=0}^{N-1} A_{ms} f(\varphi_s, t_k) + \mu \ . \tag{1.18}$$

This relationship can be compactly expressed using matrix notation as

$$E[Y|F(\varphi)] = AF(\varphi) + \mu \ . \tag{1.19}$$

It is easily shown that under these assumptions the probability density for the sinogram matrix is given by [55]

$$p(Y|\varphi) = \prod_{k=0}^{K-1} \prod_{m=0}^{M-1} \frac{(A_{m*}F(\varphi, t_k) + \mu)^{Y_{mk}} e^{-(A_{m*}F(\varphi,t_k)+\mu)}}{Y_{mk}!} \tag{1.20}$$

where A_{m*} is the m^{th} row of the system matrix, A. The log likelihood of the sinogram matrix is then given by

$$LL(Y|\varphi) = \sum_{k=0}^{K-1} \sum_{m=0}^{M-1} Y_{mk} \log(A_{m*}F(\varphi, t_k)+\mu) - (A_{m*}F(\varphi, t_k)+\mu) - \log(Y_{mk}!) \ . \tag{1.21}$$

This is a very general formulation. For specific scanners, the form of the system matrix A may vary considerably, and accurate determination of the matrix A can be critical to obtaining accurate tomographic reconstructions [56, 57].

1.3.2 MAP Estimation Framework

We will use MAP estimation to reconstruct the parametric image. For this purpose, a cost function is formed by negating the log likelihood given in (1.21) and adding a stabilizing function.

$$C(Y|\varphi) = -LL(Y|\varphi) + S(\varphi) \tag{1.22}$$

The MAP reconstruction, $\hat{\varphi}$, will be the parametric image that minimizes this cost function.

$$\hat{\varphi} = \arg\min_{\varphi} C(Y|\varphi) \tag{1.23}$$

The stabilizing function can be obtained from an assumed prior probability distribution for the parametric image. In this work, we model the distribution of the parametric image as a Markov random field (MRF) with a Gibbs distribution of the form

$$p(\varphi) = \frac{1}{z}\exp\{-\sum_{\{s,r\}\in\mathcal{N}} g_{s-r}\|T(\varphi_s) - T(\varphi_r)\|_W^q\} \tag{1.24}$$

where z is the normalization constant, $\mathcal{N}$ is the set of all neighboring voxel pairs in φ, g_{s-r} is the coefficient linking voxels s and r, q is a constant parameter that controls the smoothness of the edges in the parametric image, $T(\cdot)$ is a transform function, and W is the diagonal weighting matrix.

In this work, we will assume $q = 2$ and that $\mathcal{N}$ is formed with voxel pairs using an 8-point neighborhood system. In this case, the probability density function corresponds to a Gaussian Markov random field, and we choose the negative logarithm of this function as our stabilizing function.

$$S(\varphi) = \sum_{\{s,r\}\in\mathcal{N}} g_{s-r}\|T(\varphi_s) - T(\varphi_r)\|_W^2 \ . \tag{1.25}$$

By choosing an appropriate transform function, $T(\cdot)$, the regularization can be done in the space of the physiologically relevant parameters. Typically, we will select $T(\cdot)$ to transform from the a, b, c, d space to the k_1, k_2, k_3, k_4 as show in Table 1.1; however, any well-behaved one-to-one transformation, $T(\cdot)$, is suitable for our algorithm.

1.3.3 Parametric Image Reconstruction using PICD

The MAP reconstruction described in equation (1.23) is computed efficiently by an algorithm which we call parametric iterative coordinate descent (PICD). This algorithm is similar to the iterative coordinate descent (ICD) algorithm used in conventional PET image reconstruction [55], but it is adapted to account for the nonlinear parameters of the compartmental model. PICD sequentially updates the parameters of each voxel thereby monotonically decreasing the cost function given in Equation (1.23). When $F(\varphi)$ is a nonlinear function, the PICD algorithm reduces computation by decoupling the dependencies between the compartment model nonlinearities and the forward tomography model.

In order to compute a PICD voxel update, we must compute

$$\varphi_s \leftarrow \arg\min_{\varphi_s} C(Y|\varphi_s) \ . \tag{1.26}$$

To do this efficiently, we use the second order Taylor expansion of the change in the cost function.

Suppose we are updating the parameters of voxel s from $\varphi_s = [a_s, b_s, c_s, d_s]^t$ to $\tilde{\varphi}_s = [\tilde{a}_s, \tilde{b}_s, \tilde{c}_s, \tilde{d}_s]^t$, and that we represent the change in the time response function of voxel s by the $1 \times K$ vector function

$$\Delta f(\tilde{\varphi}_s, \varphi_s) = f(\tilde{\varphi}_s) - f(\varphi_s) \ .$$

We next define a simplified cost functional

$$\Delta C(\tilde{\varphi}_s, \varphi_s) \ = \ -LL(Y|\tilde{\varphi}_s) + LL(Y|\varphi_s) + \sum_{r \in \partial s} g_{s-r} \|T(\tilde{\varphi}_s) - T(\varphi_r)\|^2_W \ .$$

Notice that since $\Delta C(\tilde{\varphi}_s, \varphi_s)$ is equal to the change in the cost functional $C(Y|\tilde{\varphi}_s)$ within a constant, so it may be used to compute the voxel update of (1.26). The value of $\Delta C(\tilde{\varphi}_s, \varphi_s)$ can then be locally approximated with a second order Taylor series as

$$\Delta C(\tilde{\varphi}_s, \varphi_s) \ \approx \ \Delta f(\tilde{\varphi}_s, \varphi_s)\theta_1 + \frac{1}{2} \|\Delta f(\tilde{\varphi}_s, \varphi_s)\|^2_{\theta_2} + \sum_{r \in \partial s} g_{s-r} \|T(\tilde{\varphi}_s) - T(\varphi_r)\|^2_W$$

where ∂s denotes the set of voxels that are 8-neighbors of voxel s, θ_1 is a $K \times 1$ vector, θ_2 is a $K \times K$ diagonal matrix, and $\|x\|_{\theta_2}^2 = x^t \theta_2 x$. Here the values of θ_1 and θ_2 consist of the first and second derivatives respectively of the log likelihood function evaluated at each time frame. These derivatives at time frame k can be iteratively updated using the equations of the conventional ICD algorithm [55], given in (1.27) and (1.28).

$$[\theta_1]_k \leftarrow \sum_{m=0}^{M-1} A_{ms} \left(1 - \frac{Y_{mk}}{A_{m*} F(\varphi, t_k) + \mu} \right) \tag{1.27}$$

$$[\theta_2]_{k,k} \leftarrow \sum_{m=0}^{M-1} Y_{mk} \left(\frac{A_{ms}}{A_{m*} F(\varphi, t_k) + \mu} \right)^2 \tag{1.28}$$

Using the notation defined in (1.13), (1.14), and (1.15), the PICD update can then be expressed as

$$\tilde{\varphi}_s \leftarrow \arg\min_{\tilde{\varphi}_s} \left\{ \Delta f(\tilde{\varphi}_s, \varphi_s)\theta_1 + \frac{1}{2} \|\Delta f(\tilde{\varphi}_s, \varphi_s)\|_{\theta_2}^2 + \sum_{r \in \partial s} g_{s-r} \|T(\tilde{\varphi}_s) - T(\varphi_r)\|_W^2 \right\} \tag{1.29}$$

where

$$\Delta f(\tilde{\varphi}_s, \varphi_s) = \left[\tilde{a}_s, \tilde{b}_s \right] \begin{bmatrix} \alpha(\tilde{c}_s) \\ \beta(\tilde{d}_s) \end{bmatrix} - [a_s, b_s] \begin{bmatrix} \alpha(c_s) \\ \beta(d_s) \end{bmatrix} . \tag{1.30}$$

We have found that the PICD update is best implemented using two-stage nested optimization.

$$(c_s, d_s) \leftarrow \arg \min_{\tilde{c}_s \geq \tilde{d}_s \geq 0} \left\{ \arg \min_{\tilde{a}_s, \tilde{b}_s \geq 0} \left\{ \Delta C([\tilde{a}_s, \tilde{b}_s, \tilde{c}_s, \tilde{d}_s], \varphi_s) \right\} \right\} . \tag{1.31}$$

This nested optimization strategy is very important in reducing computation and assuring robust convergence. The inner optimization over $\tilde{a}_s$ and $\tilde{b}_s$ must be performed many times since this result is required for each update of outer optimization over $\tilde{c}_s$ and $\tilde{d}_s$. Fortunately, optimization over $\tilde{a}_s$ and $\tilde{b}_s$ can be done very efficiently with a simple steepest descent algorithm because this optimization does not require updating of θ_1, θ_2, $\alpha(\tilde{c}_s)$, or $\beta(\tilde{d}_s)$. Optimization with respect to $(\tilde{c}_s, \tilde{d}_s)$ is done using iterative 1-D golden section search along the $\tilde{c}_s$ and $\tilde{c}_s + \tilde{d}_s$ directions. This method assures the convergence is to a local minimum that meets the Kuhn-Tucker conditions [58]. Appendix A contains pseudocode that specify details of the algorithm.

1.3.4 Multiresolution Initialization

It is well known that for the tomographic problem the ICD reconstruction algorithm tends to have slow convergence at low spatial frequencies [59]. Normally, this problem is solved by initializing the ICD iterations with a FBP reconstruction. In this case, most of the residual error is only at high frequencies, so the ICD iterations converge quickly. However, for parametric reconstruction there is no simple direct reconstruction algorithm, such as FBP, to use as an initialization for the PICD iterations.

To solve this problem, we use a multi-resolution reconstruction scheme, which first computes coarse resolution reconstructions and then and proceeds to finer scales. The coarsest resolution reconstruction is initialized with a single set of parameters obtained by weighted least squares curve fitting to the average emission rate of each time frame. Importantly, the average activity of each time frame can be calculated directly from the sinogram data with little computation. Finer resolution reconstructions are then initialized by interpolating the parametric reconstruction of the previous coarser resolution. This recursive process reduces computation because the computationally inexpensive reconstructions at coarse levels provide a good initialization for finer resolution reconstructions.

1.4 Image Domain Parameter Estimation Methods

For purposes of comparison, we will also consider image domain methods which estimate parameters at each voxel from reconstructed images at each time. Each of these methods requires that the sinogram at each time frame be reconstructed using conventional reconstruction methods. For these methods, let $x_s(t_k)$ denote the reconstructed activity of voxel s at time frame k collected at time t_k, and let

$$x_s = [x_s(t_0), x_s(t_1), \cdots, x_s(t_{K-1})]$$

denote the activity of voxel s at all time frames.

1.4.1 Pixel-wise Weighted Least Square (PWLS) Method

The pixel-wise weighted least squares method estimates the parameters of each voxel by iteratively minimizing the weighted square error between the reconstructed time response of the voxel and the model output.

The parameters of voxel s are estimated as

$$\hat{\varphi}_s = \arg\min_{\varphi_s} \|x_s - f(\varphi_s)\|^2_{W_s} \tag{1.32}$$

where W_s is the $K \times K$ diagonal weighting matrix for voxel s. The weight of each time frame is chosen to be inversely proportional to the variance of the voxel activity in that time frame. This variance can be approximated by the activity estimate of this voxel, normalized by the duration of the time frame. In this case, W_s is a diagonal matrix with diagonal elements given by

$$[W_s]_{k,k} = \frac{\Delta t_k}{\max\{x_{MIN}, x_s(t_k)\}} \tag{1.33}$$

where Δt_k is the duration of time frame k, and x_{MIN} controls the maximum allowable value for the weights.

The parameters are estimated using the same nested optimization strategy as specified in equation (1.31). In fact, this algorithm differs from the parametric reconstruction in only two respects. First, the data derivatives of equations (1.27) and (1.28) are replaced by

$$\theta_1 = -2W_s(x_s - f(\varphi_s)) \tag{1.34}$$

$$\theta_2 = 2W_s \; ; \tag{1.35}$$

and second, the stabilizing functional $S(\varphi)$ is set to 0.

1.4.2 Pixel-wise Weighted Least Square Method with Spatial Regularization

The spatial variation of the PWLS parameter estimates can be reduced by adding a stabilizing function to equation (1.32). The resulting estimate is given by

$$\hat{\varphi} = \arg\min_{\varphi} \sum_{s=0}^{N-1} \|x_s - f(\varphi_s)\|^2_{W_s} + S(\varphi) \tag{1.36}$$

where $S(\cdot)$ is the spatial stabilizing functional [25, 26].

In the first method, which we call the pixel-wise least squares regularized (PWLSR) method, the stabilizing function has the form specified in equation (1.25). This is the same stabilizing function as was used for direct parametric reconstruction.

For the second method, which we call the PWLSZ method, we implemented the stabilizing functional described in [26]. This method smooths the PWLS estimate and uses it in the stabilizing function. Let $H(\cdot)$ be a smoothing operator and φ^P be the PWLS parameter estimate. The constrained parametric image is then given by

$$\varphi^C = H(\varphi^P) \ .$$

Next, a weight is calculated for each voxel. For voxel s the corresponding weight is

$$w_s = \frac{\|x_s - f(\varphi^P)\|^2_{W_s}}{(\varphi^P_s - \varphi^C_s)^2} \ .$$

Using the constraint images and weights, the stabilizing function is given by

$$S(\varphi) = \sum_{s=0}^{N-1} w_s(\varphi_s - \varphi^C_s)^2 \ . \tag{1.37}$$

Notice that the stabilizing function of (1.37) penalizes the difference between the parameters and a smoothed version of the parameters. Alternatively, the more traditional stabilizing function of (1.25) penalizes the spatial derivatives of the parameters.

For both of these methods the solution to (1.36) is computed using the nested optimization strategy specified in (1.31) and the data derivatives specified in (1.34) and (1.35).

1.4.3 Linear (Logan) Method

Kinetic parameter groups can sometimes be easily estimated by properly transforming the data. The Logan plot is a popular integral transform of the model given in equations (1.3), (1.4), and (1.8). This transformation can be expressed as follows:

$$\left[\frac{\int_0^{t_k} x_s(t)dt}{x_s(t_k)}\right] = \frac{k_{1s}}{k_{2s}}\left(1+\frac{k_{3s}}{k_{4s}}\right)\left[\frac{\int_0^{t_k} C_P(t)dt}{x_s(t_k)}\right] + const\ . \tag{1.38}$$

When the transformed variables (quantities in square brackets above) are plotted against each other, the resulting line has a slope equal to the compound parameter VD_s which is defined in (1.2).

To calculate BP_s the brain is segmented into a target region and a reference region. The target region consists of voxels within the brain that contain receptors for the tracer; and the reference region consists of the voxels that do not contain receptors for the tracer (i.e. $k_3 = 0$). Let, $\mathcal{T}$ be the set of voxel indices from target region, and $\mathcal{R}$ be the set of voxel indices from reference region.

For a voxel $r \in \mathcal{R}$ (from reference region), the distribution volume is

$$VD_r = \frac{k_{1r}}{k_{2r}}, \quad r \in \mathcal{R}\ .$$

For each voxel $s \in \mathcal{T}$ (from target region), the distribution volume ratio (DVR_s) is

$$\begin{aligned} DVR_s &= \frac{VD_s}{\frac{1}{|\mathcal{R}|}\sum_{r\in\mathcal{R}} VD_r} \\ &= 1+\frac{k_{3s}}{k_{4s}}\ , \end{aligned} \tag{1.39}$$

where $|\mathcal{R}|$ denotes the number of voxels in the region $\mathcal{R}$. Hence, the binding potential for the target region can be calculated as $BP_s = DVR_s - 1$.

The assumptions that are used in the derivation of equations (1.38) and (1.40) are as follows:

- k_1/k_2 ratio is constant for every voxel in the brain (i.e., both target and reference regions)

- The tracer has high specific activity (so binding can be described as an apparent first order process)
- Blood volume fraction, V_B, is zero inside the target and the reference
- $k_3 = 0$ for all the voxels in the reference region

1.5 Computational Complexity Analysis

In order to better understand the computational requirements of parametric reconstruction, we derive expressions for the computational complexity of a number of parameter estimation algorithms.

First consider the PICD algorithm. For each voxel update, the data derivatives, θ_1 and θ_2, are computed once. The complexity of this computation is $\mathcal{O}(KM_0)$, where K is the number of time frames, and M_0 is the average number of nonzero projections per voxel. Assume the nested search described in equation (1.31) requires L_{cd} evaluations of nonlinear parameters $\tilde{c}_s$ and $\tilde{d}_s$. Furthermore, assume that each update of $\tilde{c}_s$ or $\tilde{d}_s$ requires L_{ab} evaluations of linear parameters $\tilde{a}_s$ and $\tilde{b}_s$.

Each evaluation with respect to $\tilde{c}_s$ or $\tilde{d}_s$ requires a convolution with the plasma input function and L_{ab} evaluations with respect to $\tilde{a}_s$ and $\tilde{b}_s$. The convolution has requires $\mathcal{O}(K_cK)$ operations, and the evaluation with respect to $\tilde{a}_s$ and $\tilde{b}_s$ requires $\mathcal{O}(K)$ operations; so the total complexity of a voxel update is given by $\mathcal{O}(KM_0 + L_{cd}(K_cK + L_{ab}K))$, and the total complexity of PICD per full iteration for an N voxel image is given by $\mathcal{O}(KN(L_{cd}L_{ab} + K_cL_{cd} + M_0))$. The complexity of PWLS, PWLSR, and ICD are then derived by removing the terms corresponding to operations that are not performed.

1.6 Simulations

The following section compares the accuracy and computational burden of direct parametric reconstruction and image domain estimation methods.

Algorithm	Function	Per Iteration Complexity
PICD	Direct parametric reconstruction	$KN(L_{cd}L_{ab} + K_cL_{cd} + M_0)$
PWLS	Parameter estimation from reconstruction	$KN(L_{cd}L_{ab} + K_cL_{cd})$
PWLSR/PWLSZ	Regularized parameter estimation from reconstruction	$KN(L_{cd}L_{ab} + K_cL_{cd})$
ICD	MAP image reconstruction	$KN(M_0)$

Table 1.2
Computational complexity for a single full iteration of PICD, PWLS, PWLSR, and ICD. Notation: N = number of voxels; M_0 = is average number of projections per voxel; K = is the number of time frames; K_c = number of time points in the time-convolution kernel; L_{ab} = number iterations required for each update of $(\tilde{a}, \tilde{b})$; L_{cd} = number iterations required for each update of $(\tilde{c}, \tilde{d})$. Expressions do not include the computational cost of regularization.

1.6.1 Phantom Design

Our simulation experiments are based on a phantom of a rat's head. Figure 1.4 shows a schematic representation of the rat phantom and its constituent regions. The phantom has 7 regions including the background. These regions were obtained by segmenting an MRI scan of a rat through automated and manual techniques [60]. The regions and their corresponding parameters [61] are given in Table 1.3, and their time activity curves are shown in Fig. 1.5. Time frames of emission images are generated using these parameter images and the 2-tissue compartment model equations, and the plasma function, $C_P(t)$, is generated using (2) from reference [54]. In order to achieve sufficient accuracy, the convolution is implemented with $K_c = 691$ sample points. The blood contribution to the PET activity is assumed to be zero, and the tracer is assumed to be raclopride with ^{11}C, which has a decay constant of $\lambda = 0.034$ min^{-1}. Total scan time is 60 min., divided into 18 time frames with 4×0.5 min, 4×2 min, and 10×5 min. The phantom had a resolution of 128×128 with each voxel having dimensions of (1.2 mm)^3.

The rat phantom image at each time frame is forward projected into a sinogram using a Poisson model for the detected counts with a background (accidental coincidence) level of $\mu = 0.001$. Each sinogram consists of 180 angles and 200 radial bins per angle. This results in a value of $M_0 \simeq 934$. A triangular point spread function with a 4 mm base width is used in forward projections. The blood function, $C_p(t)$ is scaled so that the total number of counts in all sinogram frames is approximately 10 million.

1.6.2 Algorithm Implementation

Direct reconstructions were computed using the PICD algorithm with three levels of multiresolution optimization corresponding to resolutions of 32×32, 64×64 and 128×128. The reconstructions used $L_{cd} \simeq 35$ and $L_{ab} = 15$. In most cases, regular-

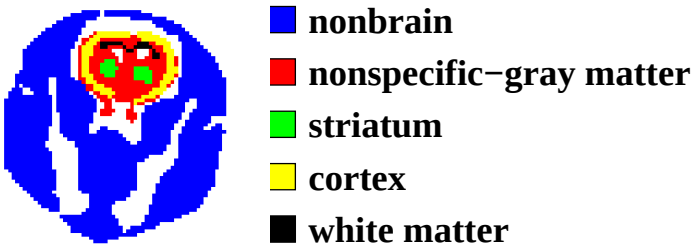


Fig. 1.4. Regions of the rat phantom derived from a segmented MR Image. Colors indicate kinetically distinct tissue regions. Red: nonspecific-gray matter tissue containing no specific binding sites for tracer but comparable blood flow values (k_1, k_2) to striatal area; Blue: non-brain; Green: striatum, containing high concentration of binding sites for tracer; Yellow: cortex, containing low concentration of binding sites; Black: White matter, contains no specific binding sites and low flow. White areas represent a mixture of background regions that do not contain any activity over time. The small white areas dorsal to (above) the striatum are ventricles that contain cerebral spinal fluid and no tracer. White areas surrounding brain correspond to skull which does not take up appreciable amounts of tracer.

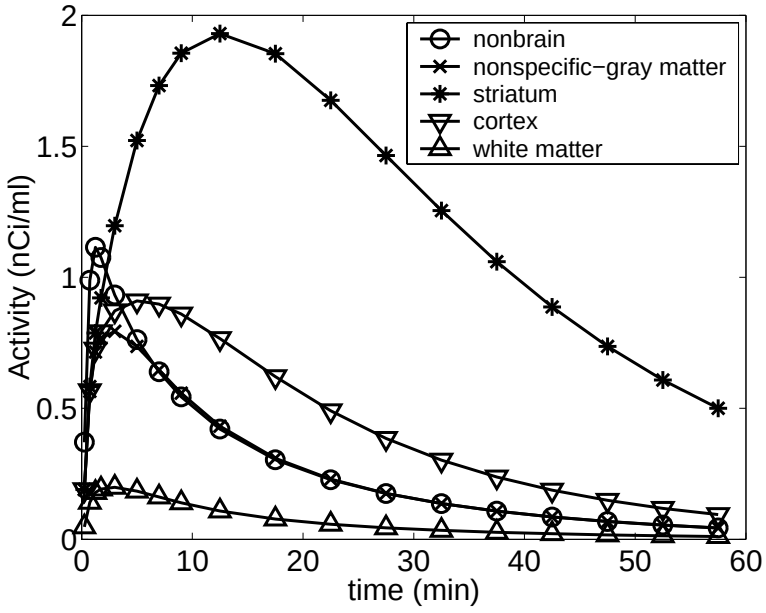


Fig. 1.5. Time-activity curves for 5 distinct tissue regions in rat brain phantom.

Region	k_1 min^{-1}	k_2 min^{-1}	k_3 min^{-1}	k_4 min^{-1}	a min^{-1}	b min^{-1}	c min^{-1}	d min^{-1}
Background	0	0	0	0	0	0	0	0
CSF	0	0	0	0	0	0	0	0
Nonbrain	.1836	.8968	0	0	.1836	0	.8968	0
Nonspecific-gray matter	.0918	.4484	0	0	.0918	0	.4484	0
Striatum	.0918	.4484	1.2408	.1363	.02164	.07016	1.7914	.0312
Cortex	.0918	.4484	.141	.1363	.0607	.0311	.628	.09725
White matter	.02295	.4484	0	0	.02295	0	.4484	0

Table 1.3
Kinetic parameters used in the simulations for distinct tissue regions of the rat head.

ization was applied directly to the k_1, k_2, k_3, and k_4 parameters; so the stabilizing functional had the form

$$S(k_1, k_2, k_3, k_4) = \sum_{i=1}^{4} \frac{1}{2\sigma_{k_i}^2} \sum_{\{s,r\}\in\mathcal{N}} g_{s-r} |k_{i,s} - k_{i,r}|^2 \ . \tag{1.40}$$

where the function g_{s-r} is inversely proportional to the distance between the voxels s and r and normalized to sum to 1, and the constants $\sigma_{k_i}^2$ control the regularization for each of the four parameters. The maximum likelihood (ML) estimate of $\sigma_{k_i}^2$ was computed for each parameter from the original parametric image using the formula [62]

$$\hat{\sigma}_{k_i}^2 = \frac{1}{N} \sum_{\{s,r\}\in\mathcal{N}} g_{s-r} |k_{i,s} - k_{i,r}|^2 \ . \tag{1.41}$$

In the original formula, N is the number of voxels in the image; however some parameter images have very few nonzero voxels, so we choose N to be the number of nonzero voxels in the image. These ML parameters are then linearly scaled all together to find a set of regularization parameters that minimize the RMSE of the estimated kinetic parameters. Some results use regularization in the k_1, k_2, BP, and VD parameters. In this case, scaling parameters are selected similarly using the appropriate parameter values.

The image domain parameter estimation methods of section 1.4 require that the image be reconstructed for each time frame. For this purpose, we used MAP image reconstruction with a quadratic prior (1.40) and a single fixed regularization parameter for all frame times. This single fixed parameter was chosen to minimize the total mean square error of the reconstructed emission image frames. The weighting matrix required for the PWLS, PWLSZ, and PWLSR algorithms was computed using (1.33) with $x_{MIN} = 0.05$ nCi/ml.

A fixed number of iterations is used for each method. The multiresolution PICD method uses 30 iterations at 32×32 resolution, 20 iterations at 64×64 resolution, and 20 iterations at 128×128 resolution. Image domain methods use 15 iterations.

In order to compute the PWLSZ reconstruction as described in section 1.4.2, we smoothed the result of PWLS reconstruction with a 3×3 equal weight filter to

calculate the constraints and weights. The weights were then scaled to minimize the MSE of the parameter estimates.

The PWLSR method was computed using a prior model on the k_1, k_2, k_3, and k_4 parameters in a manner similar to that used for parametric reconstruction. As with parametric reconstruction, the $\sigma^2_{k_i}$ constants were first selected using the ML estimation method described above, and then scaled to yield the minimum RMSE estimates of the parameters.

For the linear (Logan) method, the cortex and striatum regions are selected as target regions, and the nonspecific-gray matter was used as the reference region. Since these regions were selected precisely from simulated data, all assumptions of this method are perfectly satisfied.

1.6.3 Results

Figure 1.6 and figure 1.7 show the reconstructions of the kinetic parameters. The first row contains the original parametric images. The remaining rows are respectively the reconstructions of PWLS, PWLSZ, PWLSR, PICD reconstruction regularized on k_1, k_2, k_3, and k_4, and PICD reconstruction regularized on k_1, k_2, BP, and VD.[1] In addition, the normalized RMSE of parameters k_1, k_2, k_3, and k_4 estimated by these algorithms are listed in Fig. 1.8. The normalized RMSE of a parameter is computed as

$$\text{normalized RMSE}(k_i) = \frac{\sqrt{\frac{1}{|\mathcal{S}|}\sum_{s\in\mathcal{S}}(k_{i,s}^{true} - k_{i,s}^{estimated})^2}}{\sqrt{\frac{1}{|\mathcal{S}|}\sum_{s\in\mathcal{S}}(k_{i,s}^{true} - k_{i,s}^{PWLS})^2}} \ ,$$

where $\mathcal{S}$ is the domain where RMSE is computed, $|\mathcal{S}|$ is the number of voxels in this domain, $k_{i,s}^{true}$ is the original value, $k_{i,s}^{PWLS}$ is the PWLS estimate of the parameter, and $k_{i,s}^{estimated}$ is the estimated value of the parameter for voxel s. The RMSE of k_1 is

[1] A very small amount of regularization was also used for k_3 and k_4 (i.e. $\sigma^2_{k_3} = 1\text{min}^{-2}$ $\sigma^2_{k_4} = 0.1\text{min}^{-2}$) to suppress impulsive noise in these reconstructions.

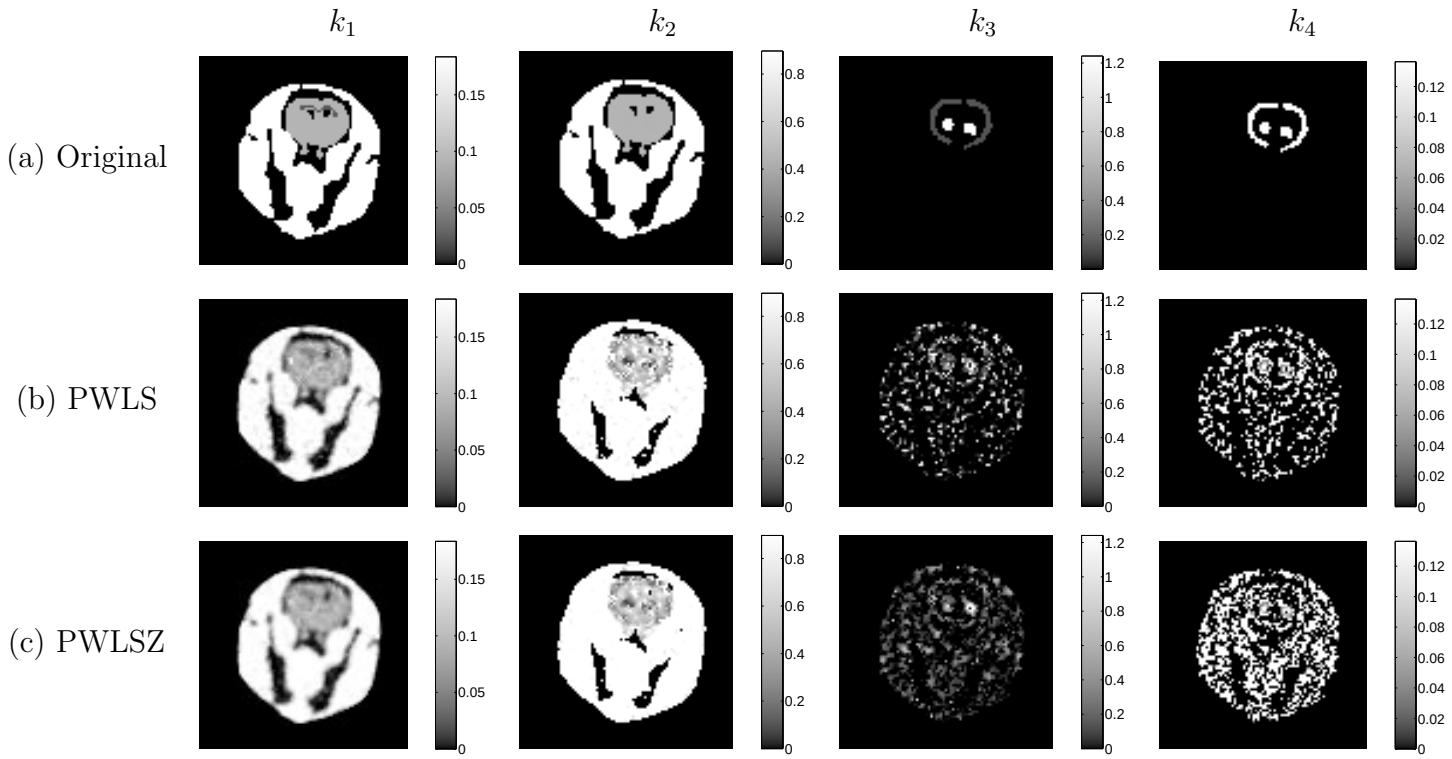


Fig. 1.6. Parametric images of k_1, k_2, k_3 and k_4 estimated by the algorithms; (a) original (b) PWLS (c) PWLSZ.

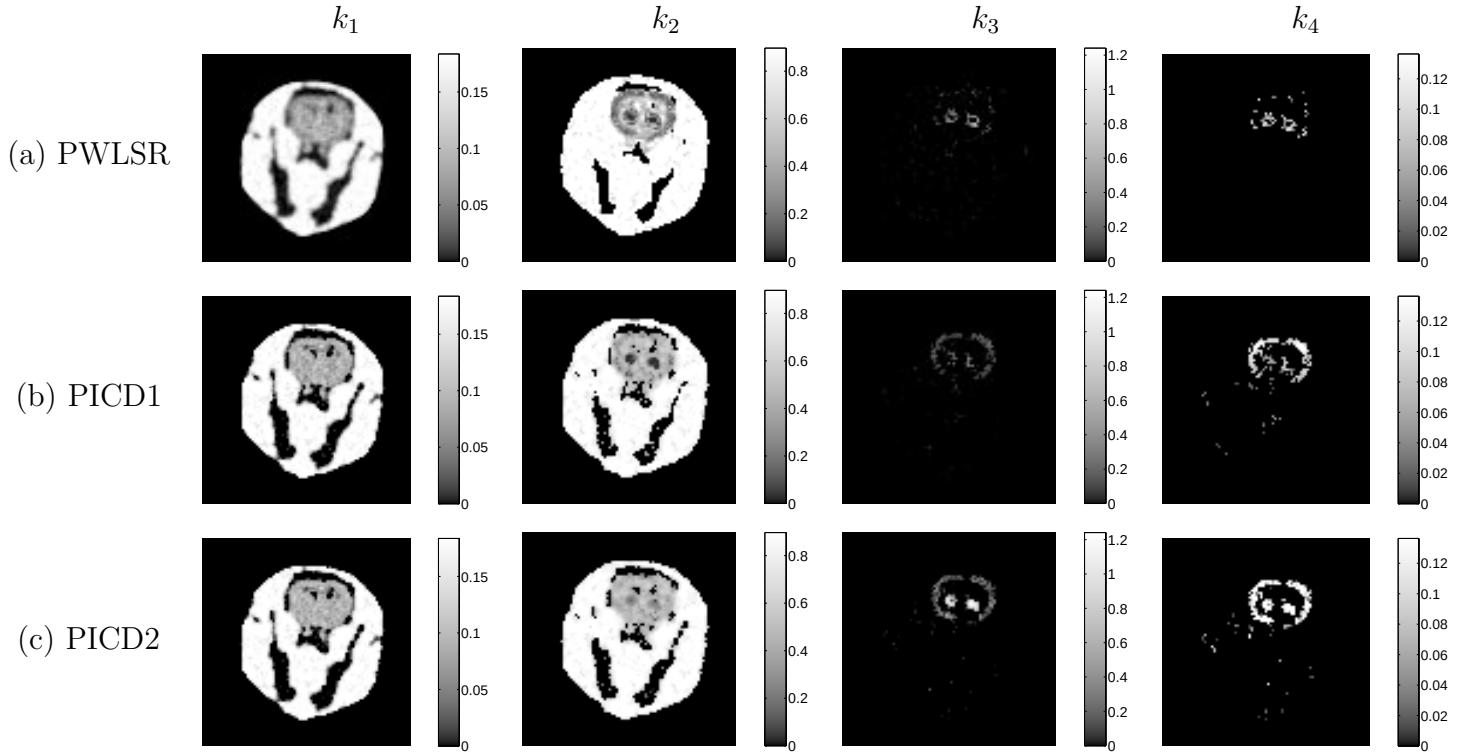


Fig. 1.7. Parametric images of k_1, k_2, k_3 and k_4 estimated by the algorithms; (a) PWLSR (b) PICD reconstruction regularized on k_1, k_2, k_3, and k_4 (c) PICD reconstruction regularized on k_1, k_2, BP, and VD.

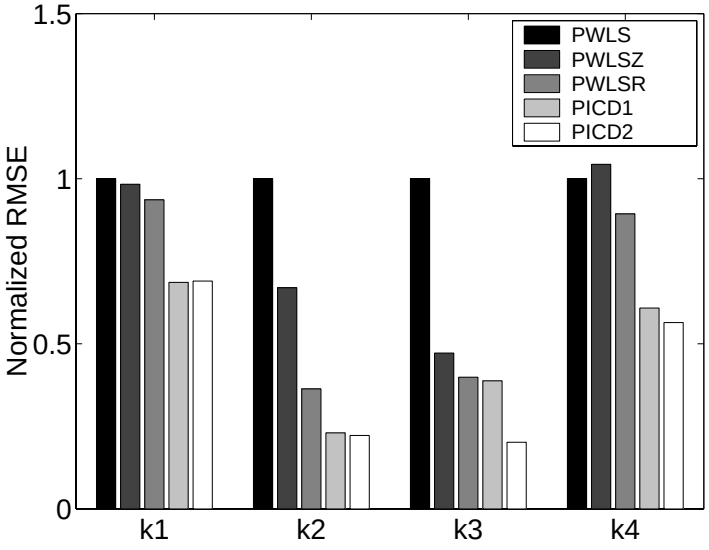


Fig. 1.8. Normalized RMSE for the reconstructed parametric images. PICD1 denotes the PICD reconstruction regularized on k_1, k_2, k_3, and k_4. PICD2 denotes the PICD reconstruction regularized on k_1, k_2, BP, and VD. Notice that PICD2 produces the lowest RMSE.

calculated over the whole image. The RMSE of parameters k_2 and k_3 are calculated over the support of k_1, and the RMSE of k_4 is calculated over the support of k_3.[2]

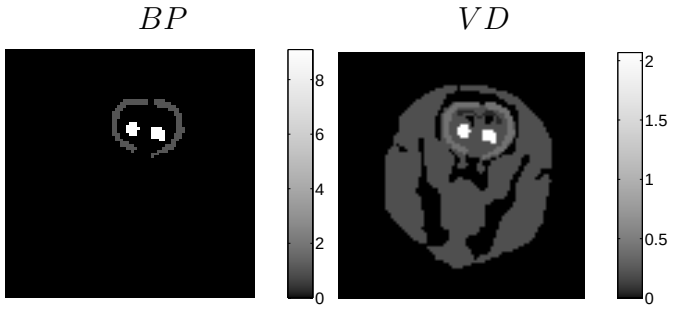
For the nonlinear parameters k_3 and k_4, the PWLS and PWLSZ methods both produced reconstructions which are very noisy, and this is reflected in the RMSE calculations. The PWLSR method with the GMRF prior produces lower RMSE reconstructions with more visually acceptable results for k_3 and k_4; however some details in these nonlinear parameters are lost. The parametric reconstruction regularized on k_1, k_2, k_3, and k_4 produces higher SNR reconstructions than any of the image domain methods, and the reconstructed images are visually similar to the original phantom. However, the parametric reconstructions with regularization on k_1, k_2, BP, and VD yield the best quality results judging from both the visual quality and the computed RMSE.

[2]When k_1 is zero, then k_2 and k_3 are not defined. Similarly, when k_3 is zero, k_4 is not defined.

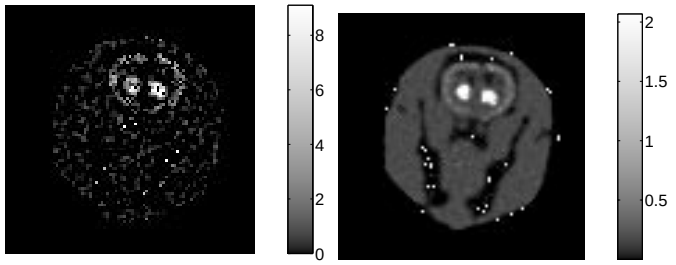
For the comparison of parameters BP and VD, spatial regularization is applied on k_1, k_2, BP, and VD. In this case, the scaling of the four regularization constants are chosen to minimize the RMSE of the BP and VD estimates alone. The results are shown in Fig. 1.6.3 and Fig. 1.6.3, and the normalized RMSE of the estimates of all methods are given in Fig. 1.11. The RMSE of BP is estimated over the support of k_3, and the RMSE of VD is estimated over the support of k_1. Again, parametric image reconstruction produces the lowest RMSE estimation for both BP and VD.

Once the parametric image is reconstructed, the forward ODE's can be solved for any particular time to reconstruct the corresponding emission image. Fig. 1.12 compares these reconstructions to the conventional reconstructions computed using FBP and MAP reconstruction for time frames 5, 10, and 15. The FBP reconstructions use a Hamming filter with cutoff at the Nyquist frequency. The RMSE of these reconstructions for each frame and for total RMSE of all frames are given in Fig. 1.13.

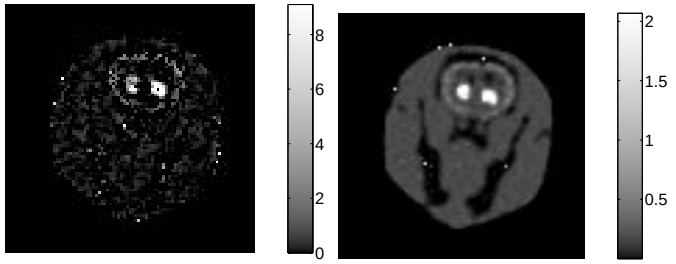
Finally, the convergence speed as a function of CPU time for all algorithms is given in Fig. 1.14. The time needed to reconstruct emission images required by image domain methods is included in this figure. As can be seen from this figure, the convergence speed of direct parametric reconstruction is comparable to the pixel-wise methods. Table 1.4 lists the CPU time required for a single iteration of each method. Notice that direct parametric reconstruction using PICD does not require substantially more computation per iteration than the image domain methods, and the image domain methods require that the images first be reconstructed. This result is consistent with the complexity listed in Table 1.2 since in this example, $(K_c L_{cd} = 24,185) >> (M_0 = 934) >> (L_{cd} L_{ab} = 525)$; so the computational complexity of the time convolution required for kinetic parameter estimation dominates the computations required for the tomographic reconstruction.



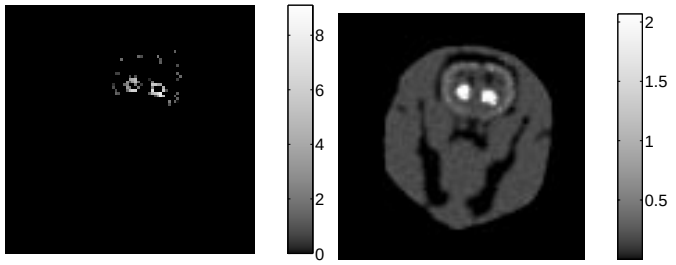
(a) Original



(b) PWLS

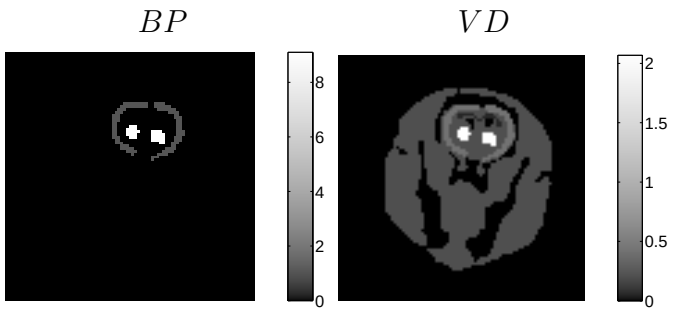


(c) PWLSZ

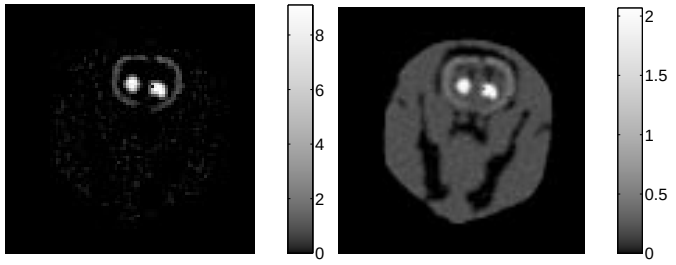


(d) PWLSR

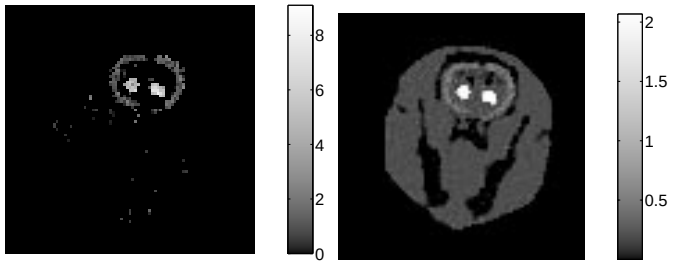
Fig. 1.9. Parametric images of BP and VD estimated by the image-domain algorithms; a) original (b) PWLS (c)PWLSZ (d) PWLSR



(a) Original



(b) Logan



(c) PICD

Fig. 1.10. Parametric images of BP and VD estimated by the Logan and PICD algorithms; a) original (b) Logan (c) PICD reconstruction (new method).

1.7 Discussion

In section 1.6.3, we demonstrated that the kinetic parameters estimated by the direct parametric image reconstruction (ie. the "direct") have lower overall error as compared to those estimated in the image domain (i.e., the "indirect"). The

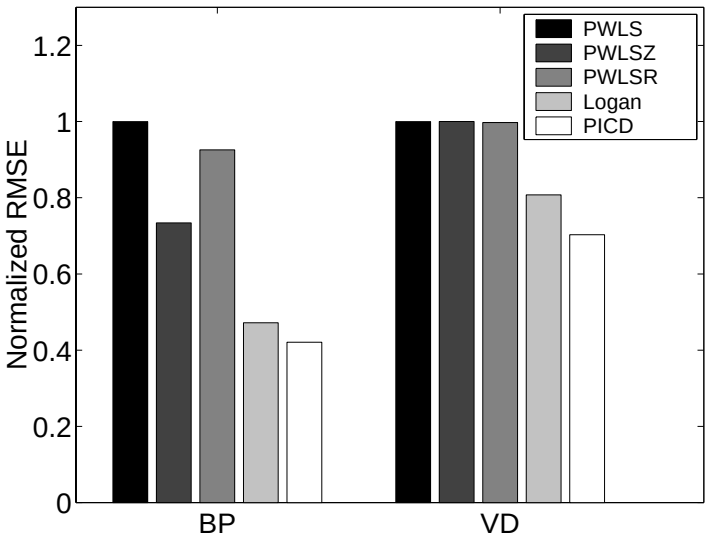


Fig. 1.11. Normalized RMSE for the reconstructed BP and VD. PICD reconstruction uses regularization on k_1, k_2, BP, and VD. Notice that PICD reconstruction gives the lowest RMSE results.

Method	time for 1 iteration (sec.)
PWLS	474
PWLSZ	487
PWLSR	526
PICD	594

Table 1.4
CPU time for a single iteration.

improvement in the visual quality and the error of the kinetic parameter estimation may be due to the following factors:

- All the available data are used simultaneously.
- Kinetic parameters are estimated directly from PET sinogram data (for which we have a very good error model).

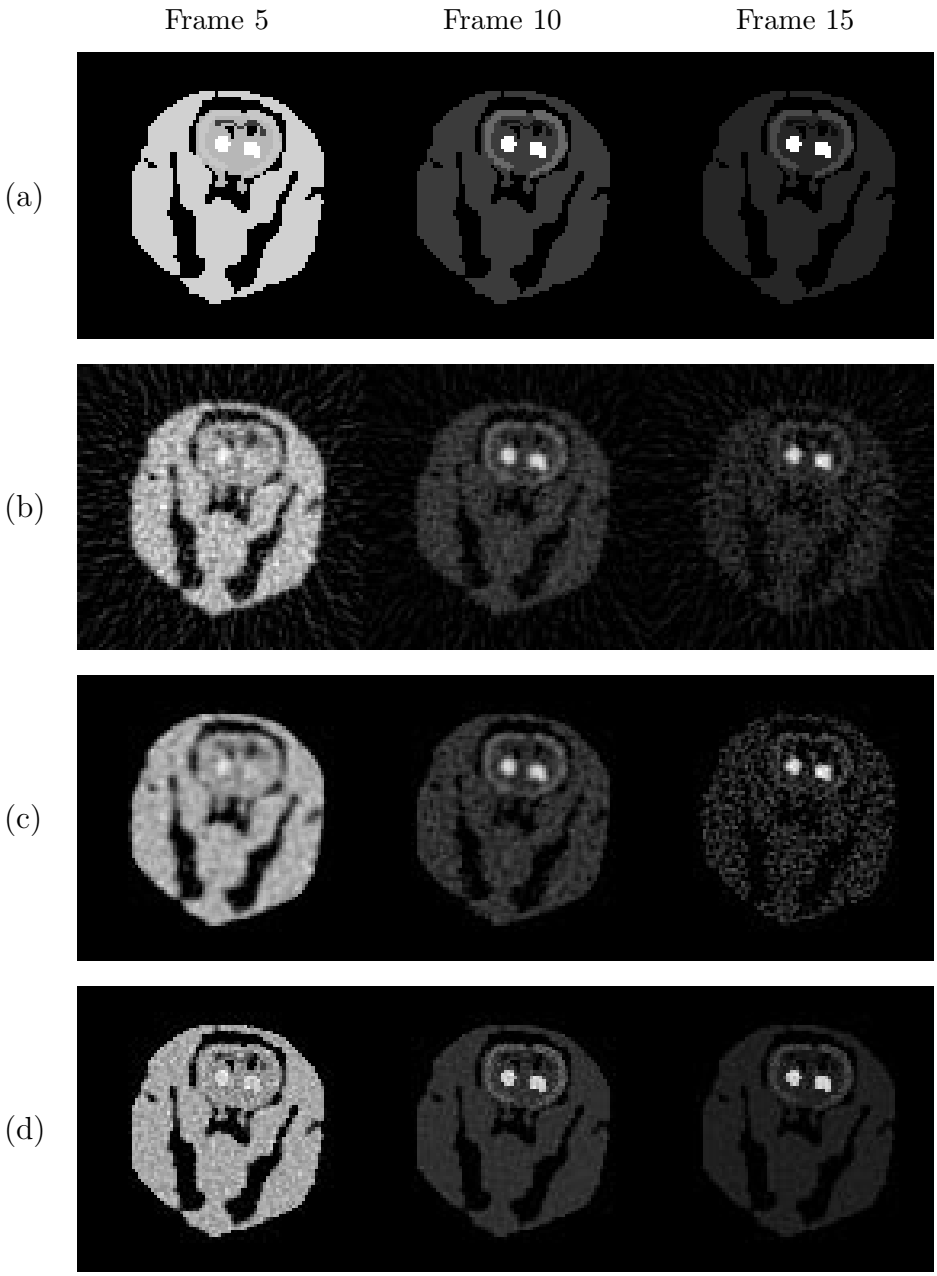


Fig. 1.12. Activity images (a) original phantom (b) FBP reconstruction (c) MAP (d) PICD reconstruction (new method) for frames 5, 10, and 15.

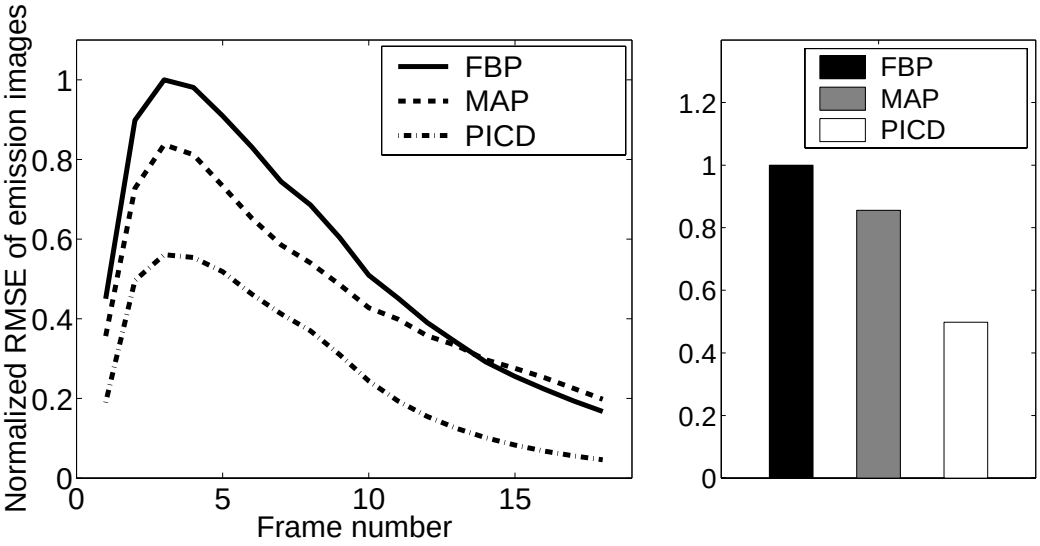


Fig. 1.13. Normalized RMSE of emission image reconstructions (a) frame by frame and (b) total. Notice that images generated using PICD reconstructed parameters have the lowest RMSE.

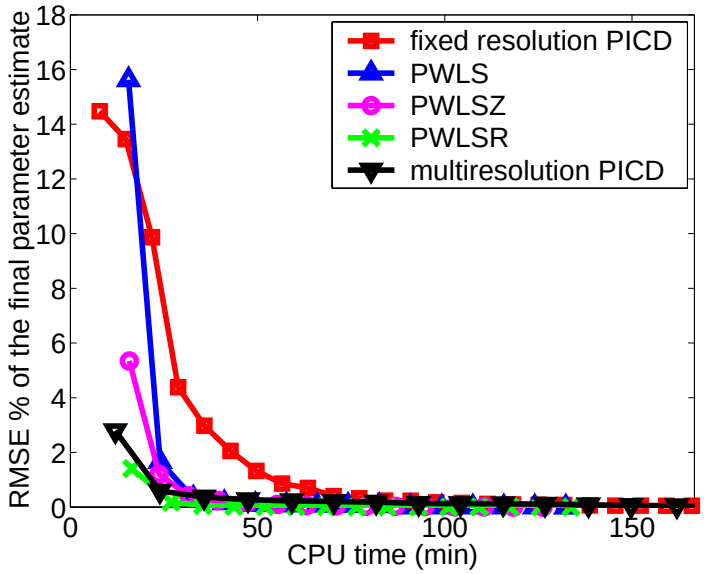


Fig. 1.14. Convergence curves for the estimation algorithms. Notice that the parametric reconstruction method with multiresolution initialization converges much faster than fixed resolution parametric reconstruction, and it is comparable in speed to the image-domain methods.

- Nonlinear estimation methods are used (so there is no need to linearize the model and introduce unwanted inaccuracy).

- Spatial regularization is done in the kinetic parameter domain (because neighboring voxels probably have similar *function*)

In contrast, the various image domain methods (described in Section 1.4) depend on the quality of the tomographic reconstructions of time-activity curves. Filtered backprojection (FBP) is still commonly used to reconstruct the dynamic PET data; unfortunately, it cannot produce the quality and the resolution achieved by the iterative reconstruction techniques. On the other hand, for the iterative methods (e.g., EM, ordered subset EM [63], or MAP [55]) to achieve their full resolution, they require that optimal reconstruction parameters be set for each reconstruction (i.e., every time frame and slice). Therefore, direct parametric reconstruction reduces the dimensionality of the estimation problem from the number of time frames to the number of kinetic parameters in the compartment model. In our simulations, the dimensionality of the estimation problem was reduced by a factor of 4.5 (from 18 time frames to 4 parameters) by the direct method.

When using an image domain approach, spatial regularization can reduce the high spatial variance in the parametric images. We have found that spatial regularization based on a Gaussian Markov model produces less estimation error for all kinetic parameters except BP, compared with a smoothing filter-based constraint.

The linear (Logan plot) method described in section 1.4.3 is a very fast estimation technique. However, this method can only estimate some of the (compound) kinetic parameters. In receptor-ligand imaging, it provides no means for estimating k_1, k_2, k_3, or k_4, individually. Furthermore, to derive BP from distribution volume ratio, there must exist a reference region in the brain devoid of receptors ($k_3 = 0$). For some tracers (e.g., muscarinic or nicotinic ligands), there is no readily apparent reference region and so the value of the Logan method is compromised. Even when an appropriate reference region exists in theory (e.g., for dopaminergic ligands) the

validity of the parameter estimates in the rest of the tissue can be biased by the placement of (or spillover of activity into) the reference ROI. In our simulations, we use the precise target (striatum and cortex) and reference (nonspecific-gray matter) regions for this method which are selected from the original image.

Another drawback to linearizations of the model is that they achieve some of their computational simplicity by unmet model assumptions (e.g., that the blood volume fraction in the reference and target tissues is zero over all time). These simplifications have been shown to introduce biases that are aggravated with decreasing SNR [64–66]. Another common assumption that is implicit in the use of Logan-plot methods is that the k_1/k_2 ratio everywhere in the brain is constant (although we satisfy this constraint in our simulated data, the direct estimation method does not require it). This ratio can, of course, be regularized spatially in the direct method. Local regularization, however, is not nearly as rigid a requirement as expecting k_1/k_2 to be constant everywhere.

Although direct parameter estimation from the PET sinograms has been proposed previously as the EMPIRA algorithm [1], this or equivalent methods have not been fully implemented. This is likely due to the computational complexity of the M-step which was not fully specified, and the slow convergence of conventional EM iterations. With the development of computationally efficient and rapidly converging methods such as have been demonstrated in this chapter, direct reconstruction to parametric images should become widely applied to dynamic PET data for which a kinetic model has been already established. It should be mentioned that there is nothing to prevent us from incorporating more complicated kinetic models into the PICD algorithm. Even though the solution to these models cannot be expressed in closed form, the power of our method, to decouple the (numerical) solution of the model from the other steps in the optimization procedure, is preserved.

We also believe that an extension of the PICD algorithm to list-mode data is possible. The function $f(\varphi_s)$, defined in equation (1.17), can be viewed as the coefficients of a zeroth order piece-wise constant spline. By using higher order splines,

the activity of a voxel at any time can be computed. In this way, event arrival times can be incorporated into the probability function and log likelihood given in equations (1.20) and (1.21) respectively. In this case, $f(\varphi_s)$ would denote the spline coefficients for voxel s, and the quantities θ_1 and θ_2 would be the first and second derivatives of the likelihood function with respect to the spline coefficients.

Our current implementation of direct reconstruction has certain limitations. For example, in image domain estimation methods it is possible to register the images for motion compensation. External measurement devices can allow us to record motion during acquisition [67] and correct the data in an automated fashion. The current implementation of our algorithm does not allow for this type of compensation. Another limitation is that our method assumes that all voxels are well modeled by the same family of model kinetics, which might not be the case in practice. However we note that single families of model kinetics have been sufficient for describing receptor ligands in different regions of the brain on an ROI-basis.

1.8 Conclusions

In this chapter, we introduce a method for the direct reconstruction of kinetic parameters at each voxel from dynamic PET sinogram data. Our algorithm, which we call parametric iterative coordinate decent (PICD), decouples the nonlinearities between the tomographic model, the kinetic model, and the regularized parameters. It also allows one to regularize with respect any desired parametrization, even if the parameters that are selected are nonlinearly related to the projections or the kinetic model parameters. Using an anatomically and physiologically realistic small animal phantom, we demonstrated that our method can reduce the mean squared error in model parameter estimates; and we show that for our example, it does not require substantially more computation than more conventional methods for computing dense parameter estimates in the image domain.

2. DIRECT RECONSTRUCTION OF PARAMETRIC IMAGES FROM EXACT HR+ DATA

2.1 Introduction

In this chapter, we directly reconstruct kinetic parameters of EXACT HR+ data using the PICD algorithm described in Chapter 1. [1]

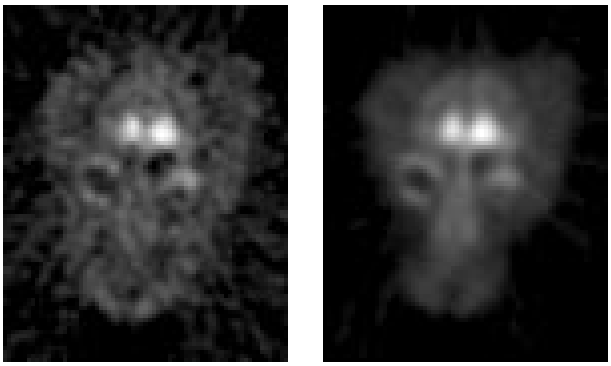
A healthy male rhesus monkey (*Macaca mulatta*) was scanned using an EXACT HR+ scanner. [2] The animal was anesthetized using ketamine (10mg/kg) and subsequently maintained on 0.5-1.5% isoflurane. The radioligand was administered through an i.v. catheter placed in the sephanous vein. A 5-min transmission scan using ^{68}Ge/^{68}Ga rod sources was acquired prior to administration of the radiopharmaceutical. The data were collected into 40 time frames consisting of 6×0.5 min., 7×1 min., 5×2 min., 4×5 min., and 18×10 min. frames. Total acquisition time was 220 min.

The raw data were corrected for randoms, deadtime, scatter, attenuation, and normalization. The data were then rebinned into 2D data sinograms using Fourier rebinning (FORE) [74]. The filtered backprojection (FBP) reconstruction of the data slice 25 at frame 20 before the corrections is shown in figure 2.1(a). The FBP reconstruction of the same slice after the sinograms were corrected and Fourier rebinned is shown in 2.1(b). For both FBP reconstructions, a Hamming filter at the Nyquist frequency was used.

Arterial blood samples were collected for the PET studies from an arterial catheter placed in the femoral artery. Coinciding with injection of the [^{18}F]fallypride, 1 ml

[1] Data provided by Dr. Brad Christian, Kettering Medical Center, Kettering, OH

[2] See [68–72] for more information on the EXACT HR+ scanner geometry and sinogram structure. See [73] for the description ECAT 7.2 data format.



(a) (b)

Fig. 2.1. FBP reconstruction of slice 25 at frame 20 (a) no corrections and no rebinning, (b) corrected for randoms, deadtime, scatter, attenuation, and normalization. FORE rebinned.

samples were initially drawn at 10-sec intervals for the first 3 min, 5-min intervals for 15 min, and 30-min intervals until the end of the study. (See [75] for further details on plasma function measurements.) The metabolite corrected plasma input function is shown in figure 2.2.

2.2 Results and Conclusion

The parametric images were reconstructed by the proposed direct reconstruction algorithms. For comparison, we used PWLS and PWLSR algorithms that are described in section 1.4, For these image domain methods, the tomographic reconstructions were performed using FBP with a Hamming filter at the Nyquist frequency.

The parametric images of k_1, k_2, k_3, and k_4 that were produced by PWLS, PWLSR, and PICD (direct reconstruction) are shown in figure 2.3(a), figure 2.3(b), and figure 2.3(c) respectively.

The parametric images of BP and VD that were produced by PWLS, PWLSR, and PICD (direct reconstruction) are shown in figure 2.4(a), figure 2.4(b), and figure 2.4(c) respectively.

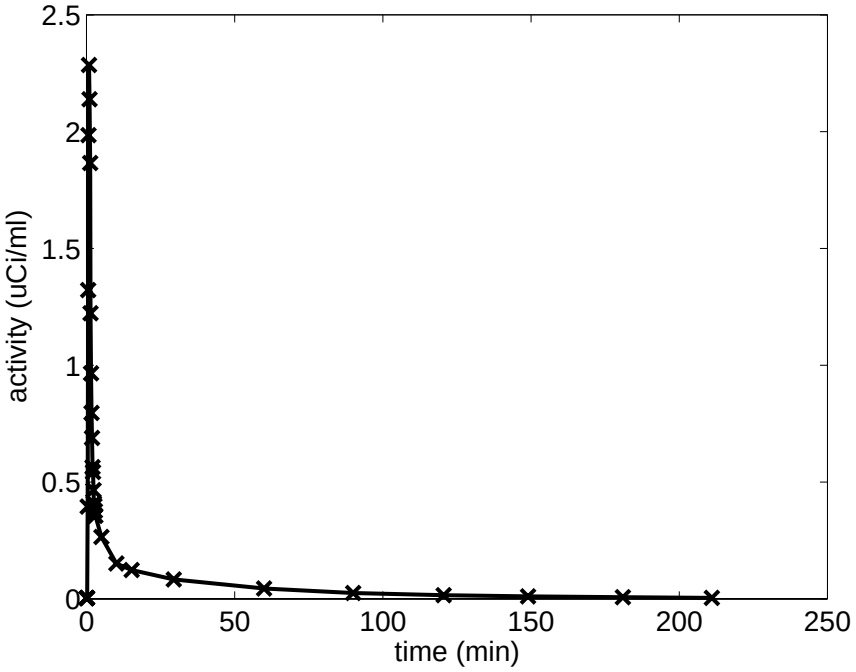


Fig. 2.2. Metabolite corrected plasma input function for the EXACT HR+ data.

	PWLSR	PICD
σ_{k1}	1	0.02
σ_{k2}	1	0.01
σ_{k3}	1	0.06
σ_{k4}	1	0.01
σ_{BP}	30	20
σ_{VD}	8	10

Table 2.1
Regularization parameters used in PWLSR and PICD reconstructions.

The regularization parameters used in the PWLSR and PICD algorithms are listed in Table 2.1. The regularization parameters for these two methods were quite different, because they used different noise models.

The reconstructed TAC, PWLS fits, and residuals for three arbitrarily selected voxels from the monkey head are shown in figures 2.5, 2.6, and 2.7. Similarly, the measured sinogram and the forward projection of the voxels' time responses are shown in figures 2.8, 2.9, and 2.10.

As it can be seen from figures 2.3 and 2.4 the PWLS estimations have high spatial variance especially in the nonlinear parameters k_2, k_3, and k_4. With the introduction of a stabilizing function, the high variance in the estimations were avoided. The PICD reconstructions have higher resolution than the image domain estimation methods.

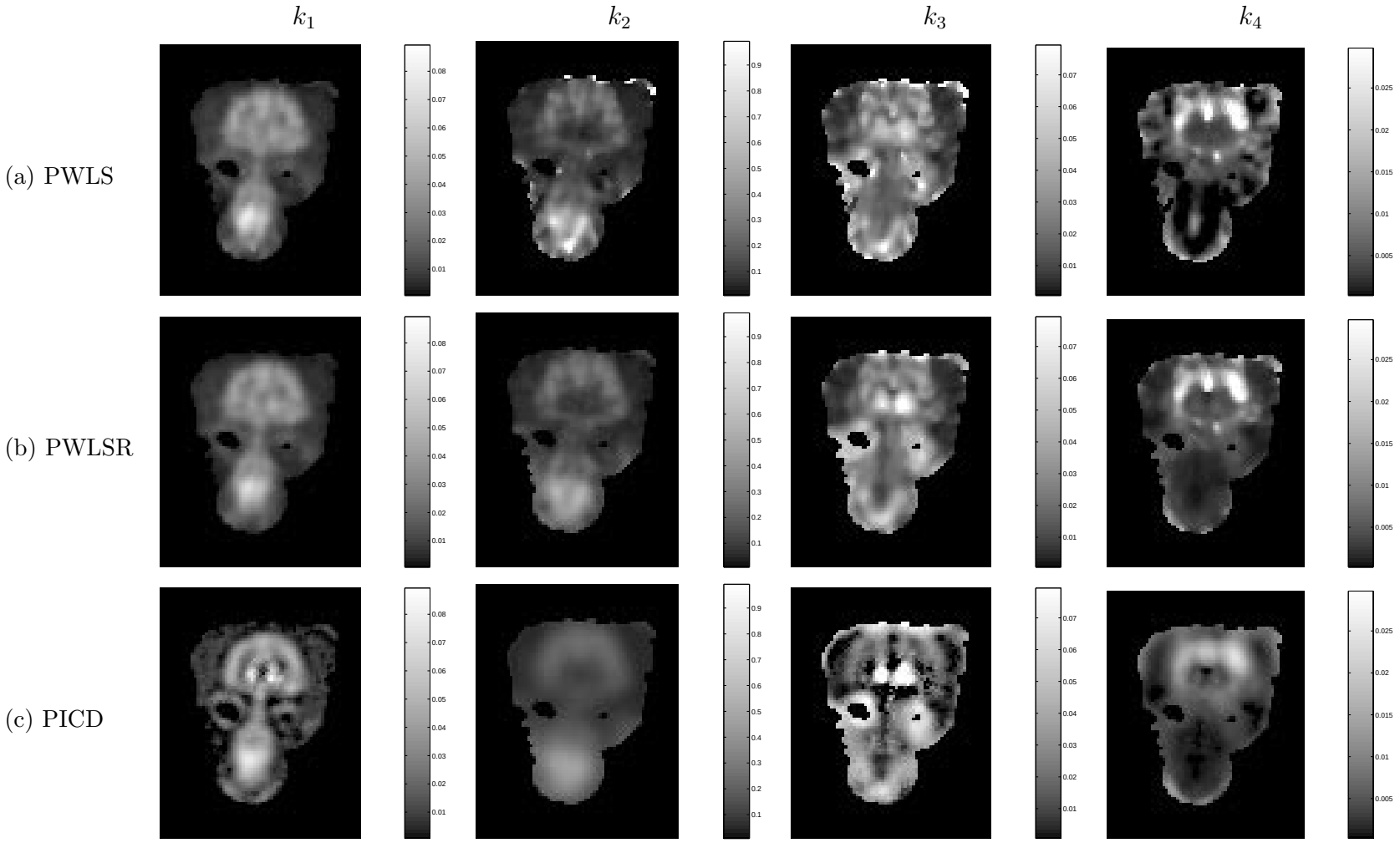


Fig. 2.3. k_1, k_2, k_3, and k_4 images of the EXACT HR+ data reconstructed by (a) PWLS, (b) PWLSR, and (c) PICD.

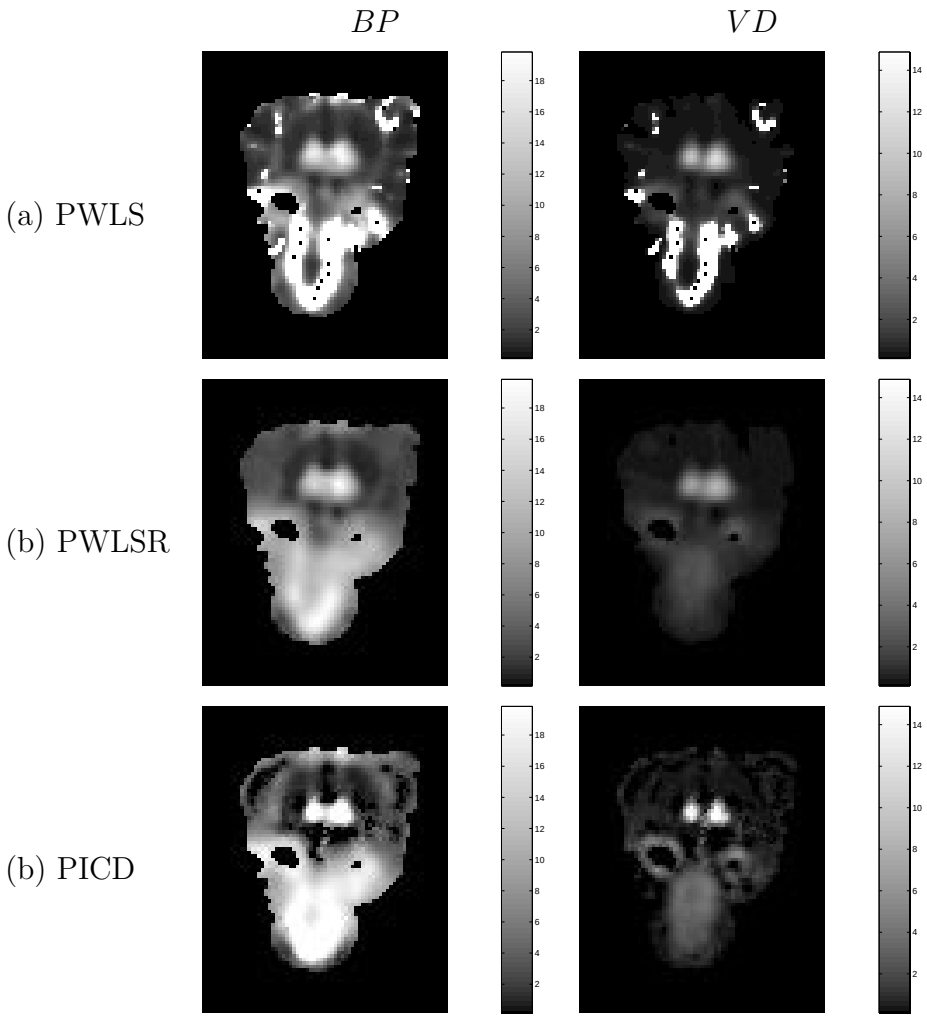


Fig. 2.4. BP and VD images of the EXACT HR+ data reconstructed by (a) PWLS, (b) PWLSR, and (c) PICD.

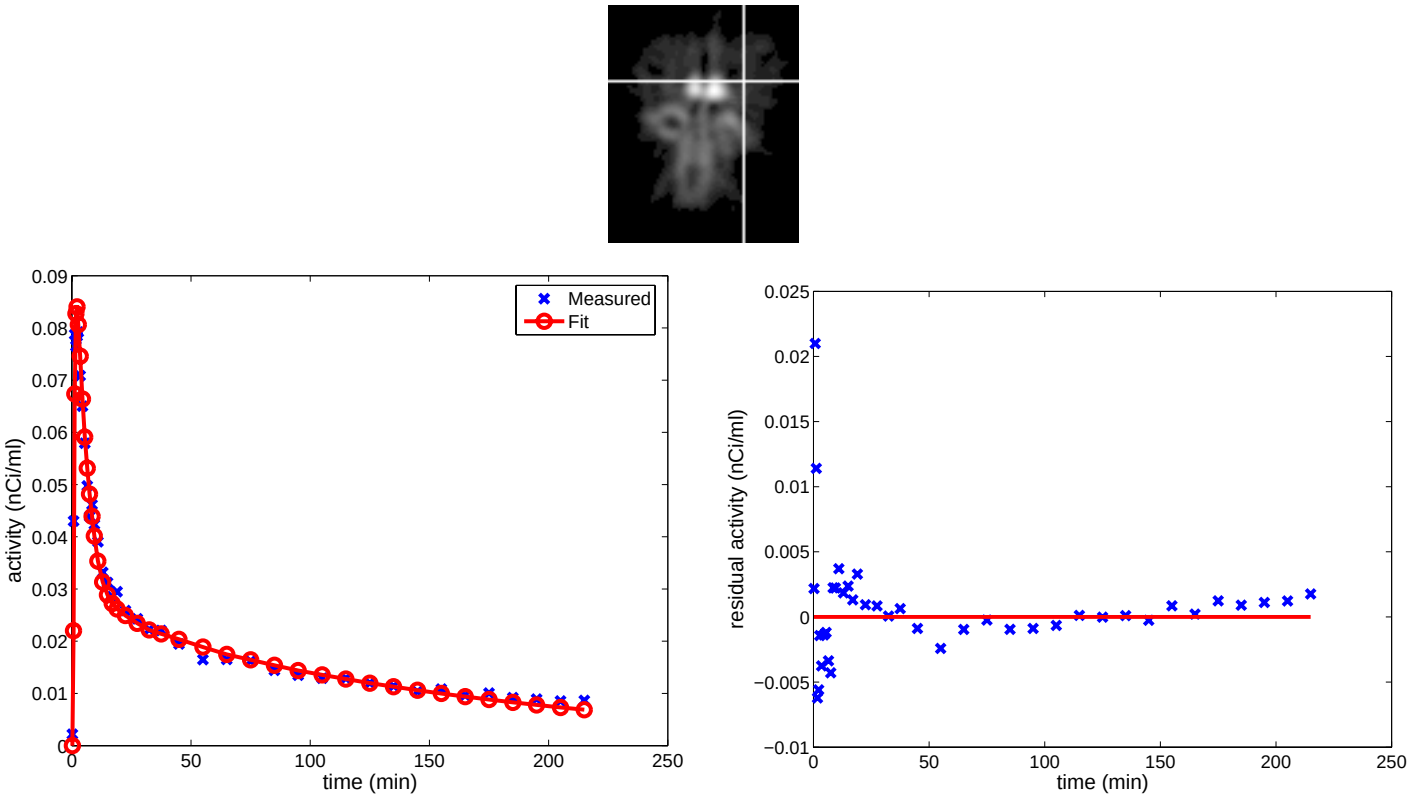


Fig. 2.5. Fit and residuals of PWLS estimations on time domain for voxel (68,70).

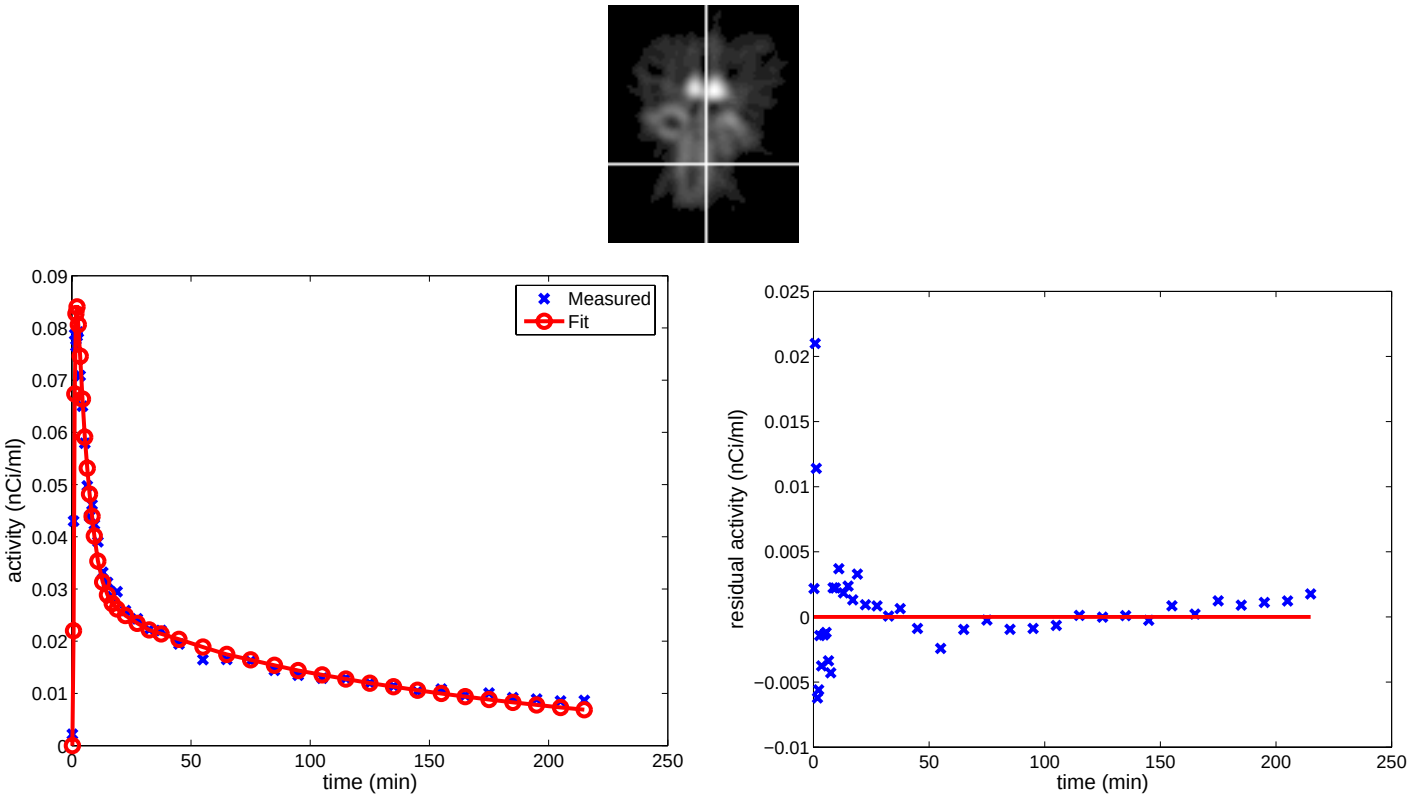


Fig. 2.6. Fit and residuals of PWLS estimations on time domain for voxel (99,56).

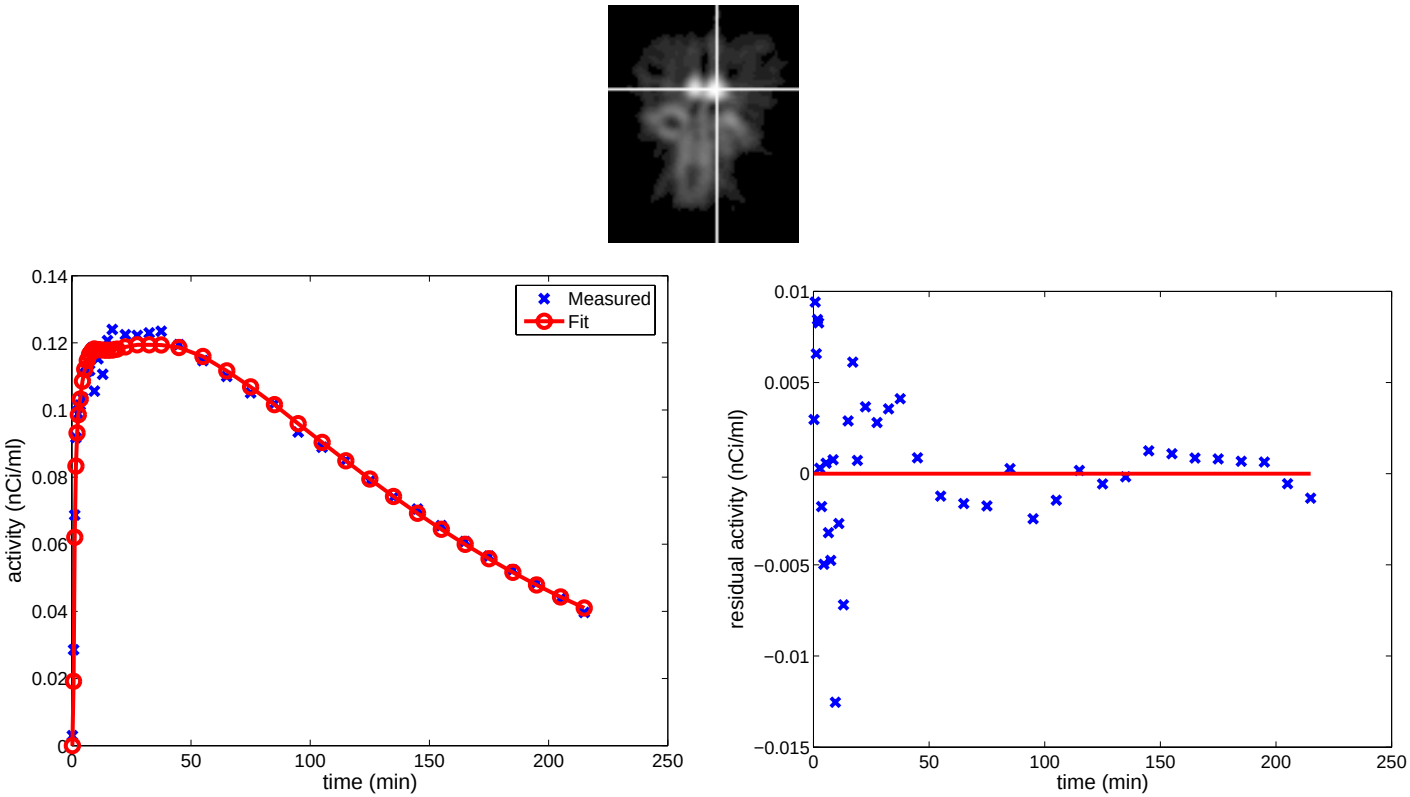


Fig. 2.7. Fit and residuals of PWLS estimations on time domain for voxel (71,60).

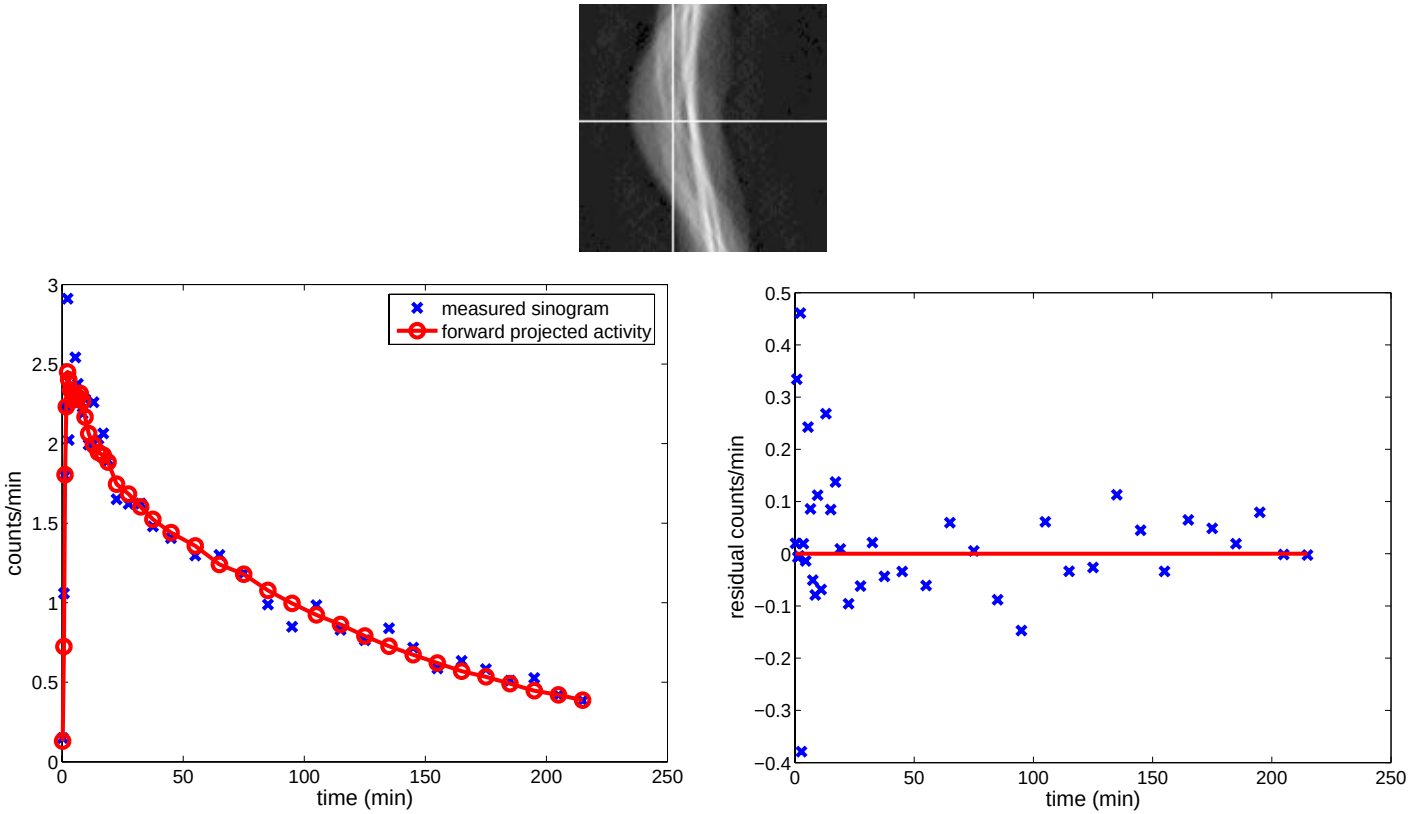


Fig. 2.8. Fit and residuals of PICD estimations on sinogram domain for sinogram element (61,49).

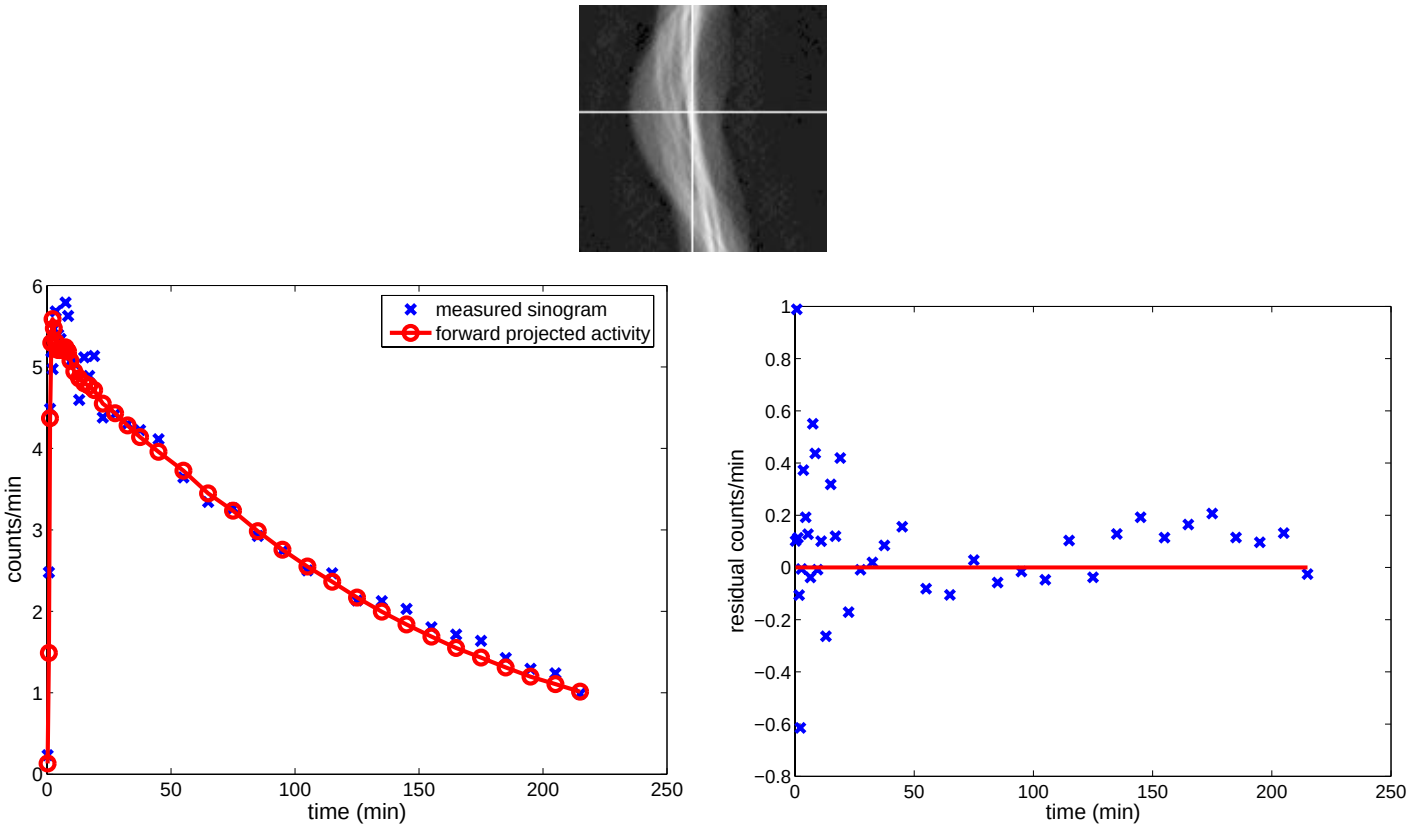


Fig. 2.9. Fit and residuals of PICD estimations on sinogram domain for sinogram element (56,59).

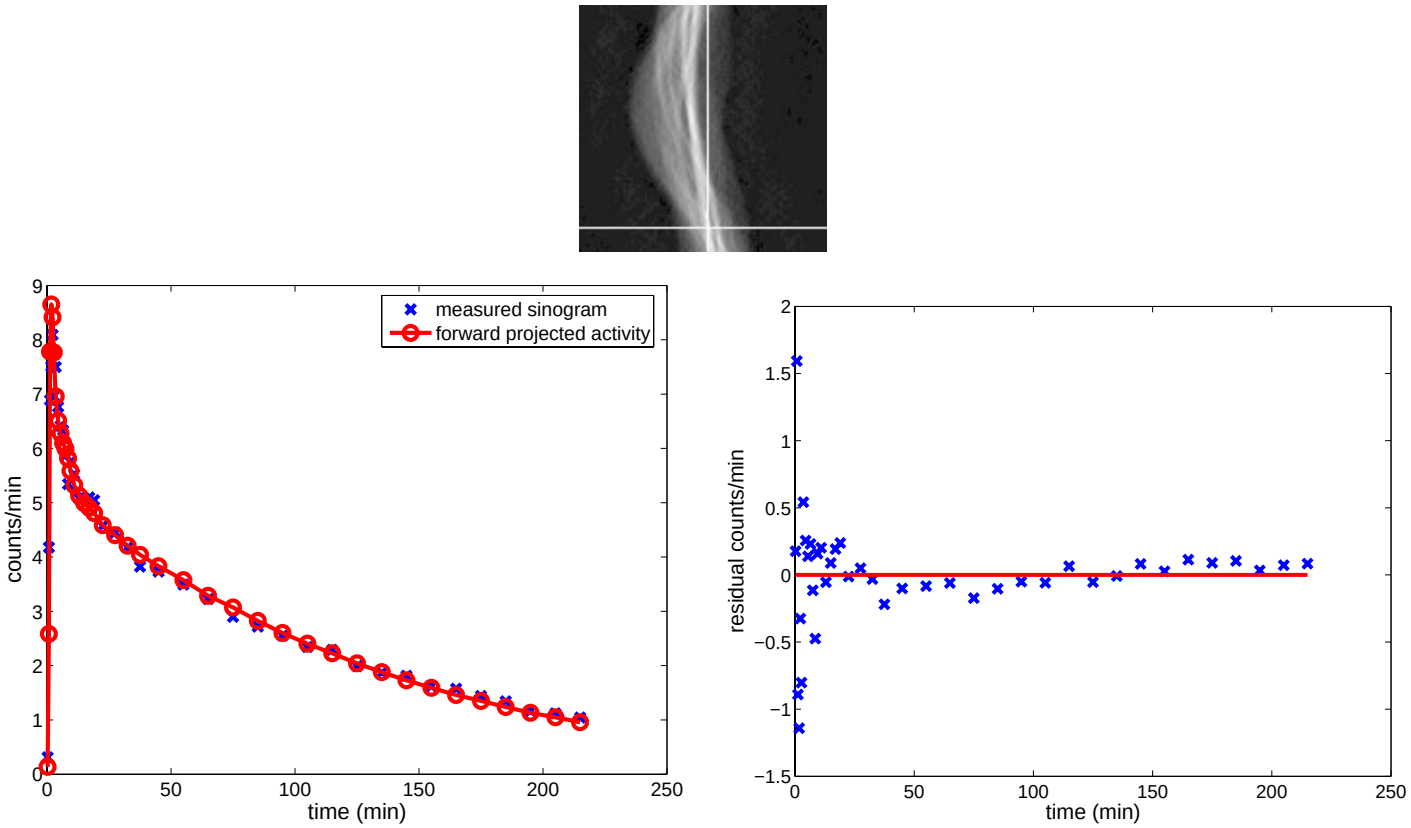


Fig. 2.10. Fit and residuals of PICD estimations on sinogram domain for sinogram element (116,67).

3. PARAMETRIC RECONSTRUCTION OF KINETIC PET DATA WITH PLASMA FUNCTION ESTIMATION

Abstract

It is often necessary to analyze the time response of a tracer. A common way of analyzing the tracer time response is to use a compartment model and estimate the model parameters. The model parameters are generally physiologically meaningful and called "kinetic parameters". In this chapter, we simultaneously estimate both the kinetic parameters at each voxel and the model-based plasma input function directly from the sinogram data. Although the plasma model parameters are not our primary interest, they are required for accurate reconstruction of kinetic parameters. The plasma model parameters are initialized with an image domain method to avoid local minima, and multiresolution optimization is used to perform the required reconstruction. Good initial guesses for the plasma parameters are required for the algorithm to converge to the correct answer. Therefore, we devised a preprocessing step involving clustering of the emission images by temporal characteristics to find a reasonable plasma curve that was consistent with the kinetics of the multiple tissue types. We compare the root mean squared error (RMSE) of the kinetic parameter estimates with the measured (true) plasma input function and with the estimated plasma input function. Tests using a realistic rat head phantom and a real plasma input function show that we can simultaneously estimate the kinetic parameters of the two-tissue compartment model and plasma input function. The RMSE of the kinetic parameters increased for some parameters and remained the same or decreased for other parameters.

3.1 Introduction

The tracer concentration in the plasma, also called the input function, is required for the estimation of physiological parameters. This plasma concentration is generally obtained by sampling blood from arteries and direct measurement of activity in these samples [76]. However, there are many risks associated with this invasive procedure such as arterial thrombosis, arterial sclerosis, and irreversible tissue ischemia. Furthermore, the arterial sampling causes discomfort to the patients and may expose the medical staff to additional radiation. As a result of these drawbacks, there has been growing interest to develop algorithms that eliminate the need for arterial blood sampling. Gunn *et al.* [16] proposed a reference region model. In this model, the brain is segmented into target (with specific binding) and reference (no specific binding) regions. A mathematical expression for the input function can be derived in terms of kinetic parameters from a reference region in the brain, and this expression can be substituted in the target region to eliminate the need for plasma concentrations. Takikawa *et al.* [77], Onishi *et al.* [78], and Eberl *et al.* [79] proposed population-based methods. The arterial sample measurements obtained from a population of former patients were used to approximate the input function. This approach has been validated for $[^{18}F]$fluorodeoxyglucose (FDG) in positron emission tomography (PET) [79], and has been validated $[^{123}I]$iomazenil in single-photon emission computed tomography (SPECT) [78]. Litton [80], Chen *et al.* [52], and Liptrot *et al.* [50] proposed an image-derived input function estimation. In these cases, the input function is estimated from imaged volumes that consist primarily of blood. Morris *et al.* [81], used a shape-based method to derive the input function by automatically identifying voxels that were most likely to represent the blood signal in perfusion MR images. Feng *et al.* [53] and Wong *et al.* [54] estimated a model-based input function and kinetic parameters simultaneously from the images.

Recently, we have proposed a method to estimate the kinetic parameters directly from PET sinograms using parametric iterative coordinate descent (PICD) [2]. We

also showed that for simulated data the PICD method can improve over image domain estimation methods. As with image domain estimation methods, the PICD algorithm requires the arterial input function to be known. In this chapter, we extend the idea of direct reconstruction from sinograms to include the estimation of a model-based plasma input function. The scale of the plasma input function and tracer uptake rate cannot be estimated individually, because a linear increase in the plasma concentration or tracer uptake rate produces the same measurements. Therefore, we fix the scale of the plasma input function and estimate the tracer uptake rate within an unknown scale factor. With this extension, the PICD algorithm can be used in cases where the input function is not known.

This chapter is organized as follows; Section 3.2 reviews the 2-tissue compartment model and the set of ODE's that govern a tracer's kinetics. Section 3.3 introduces the PICD algorithm for direct parametric reconstruction. Section 3.4 presents simulation results. Conclusion follows the results.

3.2 Two-Tissue Compartment Model

Compartment models are commonly used in modelling physiological processes [5]. A compartment model is characterized by the number of its compartments and their interactions. Each compartment in a model represents a distinct physical space or different states of the tracer. In this work, we use a 2-tissue compartment model with 4 parameters. This model is commonly used to describe the uptake and retention of an analog of glucose, 2-deoxy-2-[^{18}F]fluoro-D-glucose (FDG). The model can also be properly applied to receptor ligand studies provided that two conditions hold. First, the tracer should only bind to the specific receptor of interest. Non-specific binding will cause deviations from a simple 2-compartment model. Second, the specific activity (i.e. the molar concentration percentage of tracer which is radioactive) must be sufficiently high to insure that receptors are not saturated since this would cause the simple 2-compartment model to behave nonlinearly.

Figure 1.2 illustrates the model: C_P (pmol/ml) is the molar concentration of tracer in the plasma, C_F (pmol/ml) is the molar concentration of unbound tracer, and C_B (pmol/ml) is the molar concentration of metabolized or bound tracer. The model depends on the kinetic parameters, k_1, k_2, k_3, and k_4, which specify the tracer exchange rates between compartments in units of inverse minutes. The parameters k_1, k_2, and k_4 are first order rate constants, and k_3 is an *apparent* first order rate constant describing a process (metabolism or receptor-binding) that proceeds in proportion to the concentration of the labelled tracer only, as long as the number of sites available for binding do not become rate-limiting.

In addition to the above-stated parameters, there are two compound parameter groups known as binding potential (BP), and total volume of distribution (VD) that have physiological interpretations and practical application, particularly for receptor-ligand imaging.

$$BP = \frac{k_3}{k_4} \tag{3.1}$$

$$VD = \frac{k_1}{k_2}\left(1 + \frac{k_3}{k_4}\right) . \tag{3.2}$$

The parameter BP is proportional to the number of receptors and VD represents the steady state distribution of tracer between the plasma and tissue. In applying the model in Fig. 1.2 to all voxels, we assume that the delivery of tracer is the same to all regions being imaged. In other words, the value of C_P is not a function of voxel position. However, the values of the kinetic parameters will be allowed to vary for each voxel location, s.

In this work, the plasma concentration, C_P, is modelled using 5 parameters [82] as

$$C_P(\phi, t) = \{(A_1(t-\tau) - A_2)e^{-\lambda_1(t-\tau)} + A_2 e^{-\lambda_2(t-\tau)}\}u(t-\tau) , \tag{3.3}$$

where $\phi = [A_1, A_2, \lambda_1, \lambda_2, \tau]^t$ is the array of plasma model parameters, and $u(\cdot)$ is the unit step function. Using these assumptions, the time variation of the concentrations for a single voxel are governed by the following ordinary differential equations (ODE).

$$\frac{dC_F(s,\phi,t)}{dt} = k_{1s}C_P(\phi,t) - (k_{2s}+k_{3s})C_F(s,\phi,t) + k_{4s}C_B(s,\phi,t) \quad (3.4)$$

$$\frac{dC_B(s,\phi,t)}{dt} = k_{3s}C_F(s,\phi,t) - k_{4s}C_B(s,\phi,t)\ . \quad (3.5)$$

The solution to the ODE's in (3.4,3.5) is given by

$$C_F(s,\phi,t) = \left\{\frac{k_{1s}}{\alpha_2-\alpha_1}[(k_{4s}-\alpha_1)e^{-\alpha_1 t} + (\alpha_2-k_{4s})e^{-\alpha_2 t}]u(t)\right\} * C_P(\phi,t) \quad (3.6)$$

$$C_B(s,\phi,t) = \left\{\frac{k_{1s}k_{3s}}{\alpha_2-\alpha_1}[e^{-\alpha_1 t} - e^{-\alpha_2 t}]u(t)\right\} * C_P(\phi,t) \quad (3.7)$$

where $*$ indicates continuous-time convolution, and

$$\alpha_1, \alpha_2 = \frac{(k_{2s}+k_{3s}+k_{4s}) \mp \sqrt{(k_{2s}+k_{3s}+k_{4s})^2 - 4k_{2s}k_{4s}}}{2}\ . \quad (3.8)$$

where α_1 and α_2 are real valued constants that result from the subtraction and addition of terms in (3.8) respectively.

Next, we transform the kinetic parameters (k_1, k_2, k_3, k_4) to form the new parameters (a, b, c, d) as shown in Table 1.1. This transformation is important because while the parameters (a, b, c, d) are well suited for optimization, (k_1, k_2, k_3, k_4) are more physiologically relevant. We use $\varphi_s = [a_s, b_s, c_s, d_s]^t$ to denote the parameter vector for each voxel s.

The total activity concentration (e.g., in nCi/ml) for voxel s at time t is denoted by

$$\begin{aligned} f(\varphi_s,\phi,t) &\triangleq (1-V_B)\left[C_F(s,\phi,t) + C_B(s,\phi,t)\right] S_A e^{-\lambda t} + V_B C_{WB}(t) \\ &= (1-V_B)\left[(a_s e^{-c_s t} + b_s e^{-d_s t}) u(t) * C_P(\phi,t)\right] S_A e^{-\lambda t} + V_B C_{WB}(t) \end{aligned} \quad (3.9)$$

where S_A is the initial specific activity of the tracer (nCi/pmol), λ is the decay rate of the isotope (min^{-1}), V_B is a known constant for the volume fraction of the voxel that contains blood, and C_{WB} (nCi/ml) is the tracer activity concentration in whole blood (i.e., plasma plus blood cells plus other particulate matter). [1]

[1]Notice that both $f(\varphi_s, t)$ and $C_{WB}(t)$ in equation (3.9) include decay, either explicitly or implicitly. Therefore, the sinogram data should not be decay corrected for the implementation of this method.

We next discretize $f(\varphi_s, \phi, t)$ using $t_0, \cdots, t_{K-1}$ as the K discrete times at which the tissue is imaged. The activity at each time for voxel s is given by the $1 \times K$ row vector

$$f(\varphi_s, \phi) = [f(\varphi_s, \phi, t_0), f(\varphi_s, \phi, t_1), \cdots, f(\varphi_s, \phi, t_{K-1})] \; . \tag{3.10}$$

Let the N voxels be indexed by the values $s = 0, 1, \cdots, N-1$, and let $\varphi = [\varphi_0, \varphi_1, \cdots, \varphi_{N-1}]$ denote the $4 \times N$ matrix of parameters at all voxels. With this, we define the $N \times K$ function

$$F(\varphi, \phi) = \begin{bmatrix} f(\varphi_0, \phi) \\ \vdots \\ f(\varphi_{N-1}, \phi) \end{bmatrix}$$

which maps the parametric image, φ, to the activity of each voxel at each time. Finally, let $F(\varphi, \phi, t_k)$ denote the k^{th} column of $F(\varphi, \phi)$, so $F(\varphi, \phi, t_k)$ contains the activity for each voxel at time t_k.

3.3 Parametric Reconstruction from Sinogram Data with Plasma Function Estimation

In this section, we describe our method for simultaneously estimating the plasma function, $C_P(\phi, t)$, and the parametric image, φ directly from sinogram data. We do this by first formulating a conventional scanner model under the assumption that the sinogram measurements are Poisson random variables. We then use the kinetic model of Section 3.2 as the input to the scanner model. Once the complete forward model is formulated, we present a cost function that consists of a negative log likelihood and a prior for the kinetic parameters. Our approach is then to simultaneously minimize this cost function over the kinetic image parameters and the plasma model parameters. This algorithm can also be viewed as joint MAP estimation with a uniform prior distribution for the plasma model parameters [83]. Although the plasma model parameters are not of direct interest, they are required for the accurate reconstruction of kinetic parameters.

3.3.1 Scanner Model

Let Y_{mk} denote the sinogram measurement for projection $0 \leq m < M$ and time frame $0 \leq k < K$, and let Y be the $M \times K$ matrix of independent Poisson random variables that form the sinogram measurements. Furthermore, let A be the forward projection matrix, with elements A_{ms} (counts-ml/nCi), and let μ be the number of accidental coincidences. Then the expected number of counts for each measurement at a given time, t_k is given by

$$E[Y_{mk}|F(\varphi,\phi,t_k)] = \sum_{s=0}^{N-1} A_{ms} f(\varphi_s,\phi,t_k) + \mu \ . \tag{3.11}$$

This relationship can be compactly expressed using matrix notation as

$$E[Y|F(\varphi,\phi)] = AF(\varphi,\phi) + \mu \ . \tag{3.12}$$

It is easily shown that under these assumptions the probability density for the sinogram matrix is given by [55]

$$p(Y|\varphi,\phi) = \prod_{k=0}^{K-1}\prod_{m=0}^{M-1} \frac{(A_{m*}F(\varphi,\phi,t_k)+\mu)^{Y_{mk}} e^{-(A_{m*}F(\varphi,\phi,t_k)+\mu)}}{Y_{mk}!} \tag{3.13}$$

where A_{m*} is the m^{th} row of the system matrix, A. The log likelihood of the sinogram matrix is then given by

$$LL(Y|\varphi,\phi) = \sum_{k=0}^{K-1}\sum_{m=0}^{M-1} Y_{mk}\log(A_{m*}F(\varphi,\phi,t_k)+\mu) - (A_{m*}F(\varphi,\phi,t_k)+\mu) - \log(Y_{mk}!) \ . \tag{3.14}$$

This is a very general formulation. For specific scanners, the form of the system matrix A may vary considerably, and accurate determination of the matrix A can be critical to obtaining accurate tomographic reconstructions [56].

3.3.2 Estimation Framework

For the joint optimization of the kinetic parameters and the plasma model parameters, a cost function is formed by negating the log likelihood given in (3.14) and adding a stabilizing function.

$$C(Y|\varphi,\phi) = -LL(Y|\varphi,\phi) + S(\varphi) \tag{3.15}$$

The kinetic parameter reconstructions and the plasma model parameters can be estimated by minimizing this cost function;

$$\{\hat{\varphi},\hat{\phi}\} \leftarrow \arg\min_{\{\varphi,\phi\}} C(Y|\varphi,\phi)\ . \tag{3.16}$$

The stabilizing function can be obtained from an assumed prior probability distribution for the parametric image. In this work, we model the distribution of the parametric image as a Markov random field (MRF) with a Gibbs distribution of the form

$$p(\varphi) = \frac{1}{z}\exp\left\{-\sum_{\{s,r\}\in\mathcal{N}} g_{s-r}\|T(\varphi_s) - T(\varphi_r)\|_W^q\right\} \tag{3.17}$$

where z is the normalization constant, $\mathcal{N}$ is the set of all neighboring voxel pairs in φ, g_{s-r} is the coefficient linking voxels s and r, q is a constant parameter that controls the smoothness of the edges in the parametric image, $T(\cdot)$ is a transform function, and W is the diagonal weighting matrix.

In this work, we will assume $q = 2$ and that $\mathcal{N}$ is formed with voxel pairs using an 8-point neighborhood system. In this case, the probability density function corresponds to a Gaussian Markov random field, and we choose the negative logarithm of this function as our stabilizing function.

$$S(\varphi) = \sum_{\{s,r\}\in\mathcal{N}} g_{s-r}\|T(\varphi_s) - T(\varphi_r)\|_W^2\ . \tag{3.18}$$

By choosing an appropriate transform function, $T(\cdot)$, the regularization can be done in the space of the physiologically relevant parameters. Typically, we will select $T(\cdot)$ to transform from the a, b, c, d space to the k_1, k_2, k_3, k_4 as shown in Table 1.1; however, any well behaved one-to-one transformation, $T(\cdot)$, is suitable for our algorithm.

foreach iteration {

$$\hat{\varphi} \leftarrow \arg\min_{\varphi} C(Y|\varphi, \phi)$$

$$\hat{\phi} \leftarrow \arg\min_{\phi} C(Y|\hat{\varphi}, \phi)$$

}

Fig. 3.1. Optimization Strategy using PICD. Each iteration of optimization has two steps; first the plasma model parameters were kept fixed and parametric image is updated, then plasma model parameters are estimated for the updated parametric image.

3.3.3 Optimization Strategy using PICD

Simultaneous update of the parametric image, φ, and plasma model parameters, ϕ, is not tractable. Therefore, we chose an iterative optimization strategy. Each iteration had two steps: 1) estimate the kinetic parameters using parametric iterative coordinate descent (PICD) algorithm [2] by keeping the plasma model parameters constant, 2) update the plasma model parameters (See Fig. 3.1).

The PICD algorithm is similar to the ICD algorithm used in conventional PET image reconstruction [55], but it is adapted to account for the nonlinear parameters of the compartmental model. PICD sequentially updates the parameters of each voxel thereby monotonically decreasing the cost function given in Equation (3.16);

$$\varphi_s \leftarrow \arg\min_{\varphi_s} C(Y|\varphi_s, \varphi) \ . \tag{3.19}$$

When $F(\varphi, \phi)$ is a nonlinear function, the PICD algorithm reduces computation by decoupling the dependencies between the compartment model nonlinearities and the forward tomography model. Therefore, it is computationally efficient.

After updating the parametric image, the plasma model parameters are each sequentially updated using line search in the order listed below.

$$\begin{aligned}
\hat{A}_2 &\leftarrow \arg\min_{A_2} C(Y|\varphi, [A_1, A_2, \lambda_1, \lambda_2, \tau]) \\
\hat{\lambda}_1 &\leftarrow \arg\min_{\lambda_1} C(Y|\varphi, [A_1, \hat{A}_2, \lambda_1, \lambda_2, \tau]) \\
\hat{\lambda}_2 &\leftarrow \arg\min_{\lambda_2} C(Y|\varphi, [A_1, \hat{A}_2, \hat{\lambda}_1, \lambda_2, \tau]) \\
\hat{\tau} &\leftarrow \arg\min_{\tau} C(Y|\varphi, [A_1, \hat{A}_2, \hat{\lambda}_1, \hat{\lambda}_2, \tau]) \\
\hat{\phi} &\leftarrow [A_1, \hat{A}_2, \hat{\lambda}_1, \hat{\lambda}_2, \hat{\tau}]
\end{aligned}$$

Note that in equation (3.4) k_{1s} and $C_P(\phi, t)$ are multiplied. Therefore, we can estimate the value of the quantity $k_{1s}C_P(\phi, t)$, but we cannot individually estimate the scale of k_{1s} and $C_P(\phi, t)$. So for example, if k_{1s} is doubled, and $C_P(\phi, t)$ is halved, then the resulting observations are indistinguishable. In our optimization framework, the scale of the plasma function is determined by the parameters A_1 and A_2, and the scale of the tracer uptake rate is a function the transformed kinetic parameters a_s and b_s. Therefore, only three parameters out of A_1, A_2, a_s, and b_s can be identified from the measured sinogram data. In order to address this unidentifiability issue, A_1 can be fixed to a constant, and in this case, parameters A_2, k_1 and VD can only be estimated within a scale factor. Other parameters (k_2, k_3, k_4, and BP) are not effected. It may be possible to use some additional experimental data such as injected dosage or a single late blood sample or prior information such as population-averaged blood curve to properly scale the plasma input function (A_1 and A_2) and the kinetic parameters (k_1 and VD) [84].

3.3.4 Initialization

The joint estimation strategy described in Section 3.3.3 can converge to local minimum with an arbitrary initial plasma input function. To avoid local minima, a good initial plasma input function is required. In order to choose good initial plasma model parameters, we used an approach similar to Feng *et al.* [53] (See fig. 3.2.)

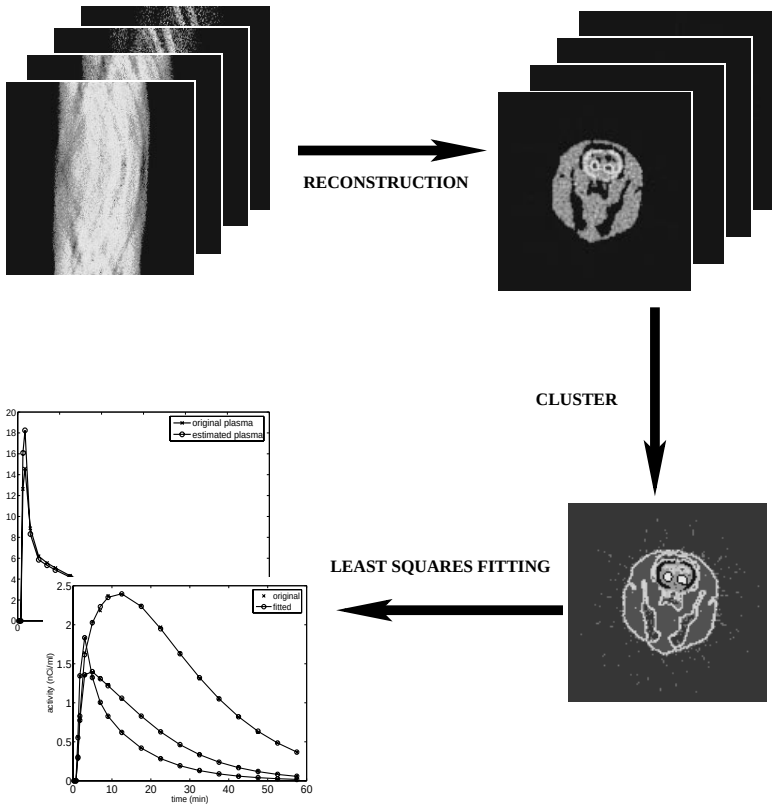


Fig. 3.2. Initialization strategy for plasma model parameters. First sinograms are reconstructed, then the tomographic reconstructions are clustered. Finally plasma model parameters and kinetic parameters are estimated using least squares curve fitting.

First, we reconstructed the sinograms using filtered back projection (FBP). The voxels were then segmented into a pre-determined number of clusters according to their reconstructed time responses [85]. Each cluster was represented by a single time response. Let L be the number of clusters, and x_l be the representative time response of cluster l. Then, initial plasma model parameters and the kinetic parameters for each cluster were chosen to minimize the weighted least squares between the time responses of the clusters and the model;

$$\{\varphi_0^{init}, \cdots, \varphi_{L-1}^{init}, \phi^{init}\} \leftarrow \arg \min_{\{\varphi_0, \cdots, \varphi_{L-1}, \phi\}} \sum_{l=0}^{L-1} \|x_l - f(\varphi_l, \phi)\|_{W_l}^2 \ . \tag{3.20}$$

In this equation, W_l denotes the $K \times K$ diagonal weighting matrix for cluster l. W_l is formed by the inverses of the time-response variances, i.e. the k^{th} diagonal element of W_l is given by

$$[W_l]_{k,k} = \frac{\Delta t_k}{x_{l,k}} \ ,$$

where Δt_k is the duration of the k^{th} time frame, and $x_{l,k}$ is the time response of cluster l at time t_k. Note that this initialization may also converge to a local minimum. However, it is a relatively fast method, and it can be executed many times starting from different points. The solution that minimizes equation (3.20) can be chosen as the initial point for our algorithm.

It is well known that for the tomographic problem the ICD reconstruction algorithm tends to have slow convergence at low spatial frequencies [59]. To solve this problem, we use a multiresolution reconstruction scheme, which first computes coarse resolution reconstructions and then proceeds to finer scales. The coarsest resolution reconstruction is initialized with ϕ^{init} and a single set of parameters obtained by weighted least squares curve fitting to the average emission rate of each time frame. Importantly, the average activity of each time frame can be calculated directly from the sinogram data with little computation. Finer resolution reconstructions are then initialized by interpolating the parametric reconstruction of the previous coarser resolution. This recursive process reduces computation because the computationally inexpensive reconstructions at coarse levels provide a good initialization for finer resolution reconstructions.

3.4 Simulation and Results

The following section compares the accuracy of kinetic parameter estimates with measured plasma input function, and with estimated plasma input function.

3.4.1 Phantom Design

Our simulation experiments are based on a phantom of a rat's head. Figure 1.4 shows a schematic representation of the rat phantom and its constituent regions. The phantom has 7 regions including the background. These regions were obtained by segmenting an MRI scan of a rat through automated and manual techniques [60]. The regions and their corresponding parameters [61] are given in Table 1.3, and their time activity curves are shown in Fig. 1.4. Time frames of emission images are generated using these parameter images and the 2-tissue compartment model equations. The plasma function, $C_P(t)$, is obtained by arterial plasma sampling of a rat scanned in IndyPET-II [86]. The blood contribution to the PET activity is assumed to be zero, and the tracer is assumed to be raclopride with ^{11}C, which has a decay constant of $\lambda = 0.034$ min^{-1}. Total scan time is 60 min., divided into 18 time frames with 4×0.5 min, 4×2 min, and 10×5 min. The phantom had a resolution of 128×128 with each voxel having dimensions of (1.2 mm)^3.

The rat phantom image at each time frame is forward projected into a sinogram using a Poisson model for the detected counts with a background (accidental coincidence) level of 0.001 nCi/ml. Each sinogram consists of 180 angles and 200 radial bins per angle. A triangular point spread function with a 4 mm base width is used in forward projections.

3.4.2 Algorithmic Implementation

The plasma model parameters were initialized as described in Section 3.3.4. Since we cannot identify A_1,A_2 and k_1 simultaneously, we can fix A_1 to any arbitrary value. However, with an arbitrary A_1 the estimated values of A_2, k_1 and VD will be off by a scale factor. Therefore we fixed A_1 to its true value in these simulations in order to simplify presentation of results and allow for the use of the same regularization parameters in all comparisons.

In the initialization, the tomographic reconstructions are clustered into 8 regions. To find initial plasma model parameters, 10 starting points in the range of $2 \leq A_2 \leq 10$, $0.5 \leq \lambda_1 \leq 5$, $0 \leq \lambda_2 \leq 0.2$, $0 \leq \tau \leq 2$ were used, and for each of these the solution that minimized equation (3.20) was chosen as the initial point.

The maximum likelihood (ML) estimate of the stabilizing function parameters were computed from the true parametric image and used in the simulations [62].

The kinetic parameters were reconstructed using PICD algorithm with three levels of multiresolution optimization corresponding to resolutions 32×32, 64×64, and 128×128. Regularization was applied directly to the k_1, k_2, BP, and VD parameters. The multiresolution PICD method was executed with a fixed number of iterations at each resolution; 40 iterations at 32×32 resolution, 20 iterations at 64×64 resolution, and 10 iterations at 128×128 resolution.

3.4.3 Results

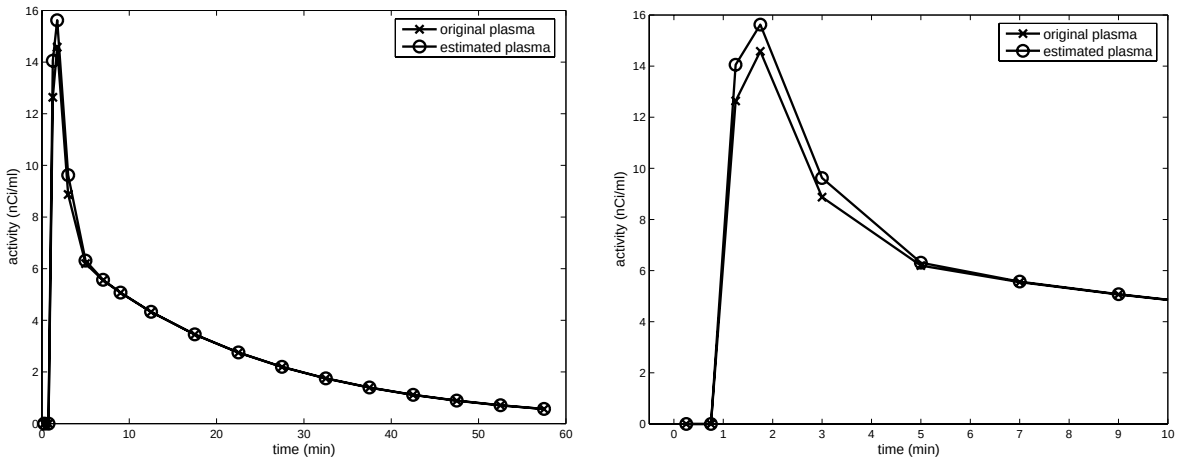
Table 3.1 shows the true and estimated plasma model parameters. Figure 3.3(a) shows the measured (true) plasma input function and estimated plasma input function, and Fig. 3.3(b) displays only the first 10 min of the measured and estimated plasma input functions.

Figure 3.4 and 3.5 display the kinetic parameter reconstructions with the measured plasma input function and with the estimated plasma input function. All kinetic parameters in these figures are estimated with regularization on k_1, k_2, BP, and VD.

Figure 3.6 shows the normalized RMSE of the kinetic parameter estimates, k_2, k_3, k_4, and BP with the measured and estimated plasma input function. Note that these are the only kinetic parameters that we can estimate without any side information. The RMSE of parameters k_2 and k_3 are calculated over the support of k_1, and the

	true	estimated
A_1	49.325	–
A_2	7.310	7.341
λ_1	1.789	1.617
λ_2	0.045	0.045
τ	0.893	0.825

Table 3.1
True and estimated plasma model parameters. Note that A_1 is not estimated.



(a) (b)

Fig. 3.3. (a) Measured and estimated plasma input functions (b) Measured and estimated plasma function in the first 10 minutes.

RMSE of k_4 is calculated over the support of k_3.[2] From this figure, the RMSE of k_2, and k_4 estimates increase with estimated plasma input function.

[2]When k_1 is zero, then k_2 and k_3 are not defined. Similarly, when k_3 is zero, k_4 is not defined.

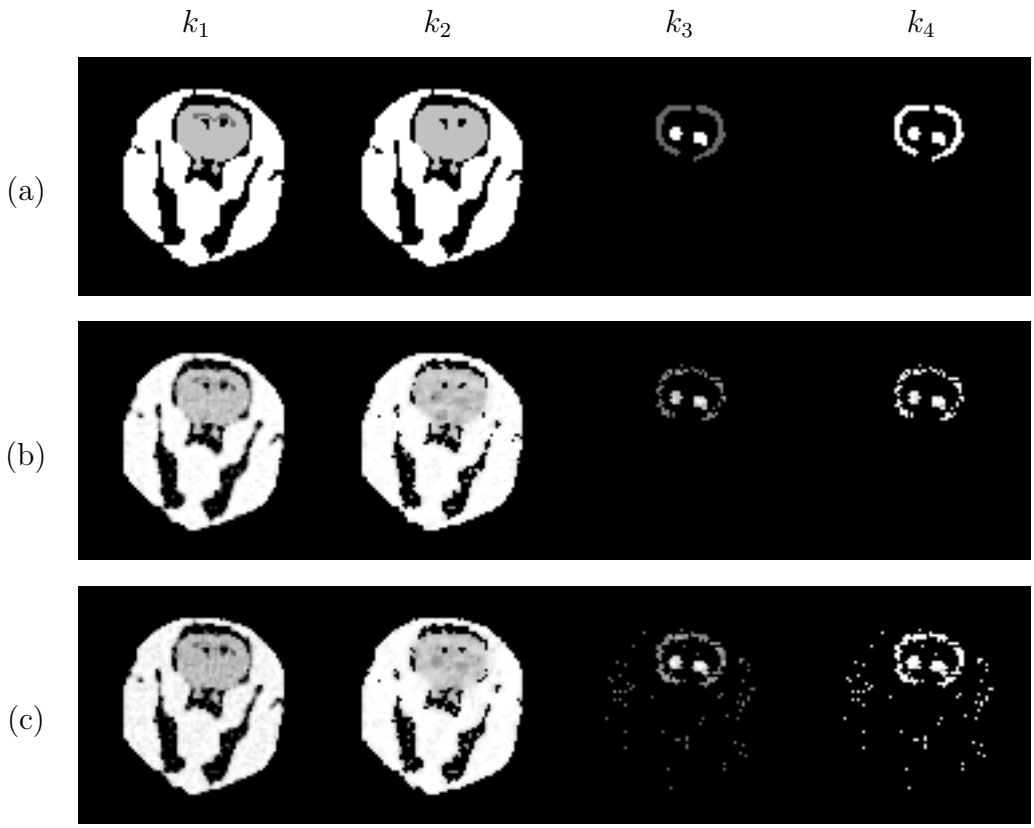


Fig. 3.4. (a) Original k_1, k_2, k_3, and k_4. Estimated k_1, k_2, k_3, and k_4 (b) estimated with measured input function, and (c) estimated with simultaneous estimation of input function.

3.5 Conclusion and Future Work

We have demonstrated that it is possible to estimate kinetic parameters k_2, k_3, k_4, and BP directly from the PET sinograms without plasma input function measurements. The tracer uptake rate, k_1, and VD can only be estimated to within a scale factor since the scale of the plasma input function is not known. A model-based plasma input is estimated jointly with kinetic parameters. In our simulation, which uses a real plasma input function and realistic phantom, the estimated plasma input function was close to the true input function. Furthermore, some of the estimated

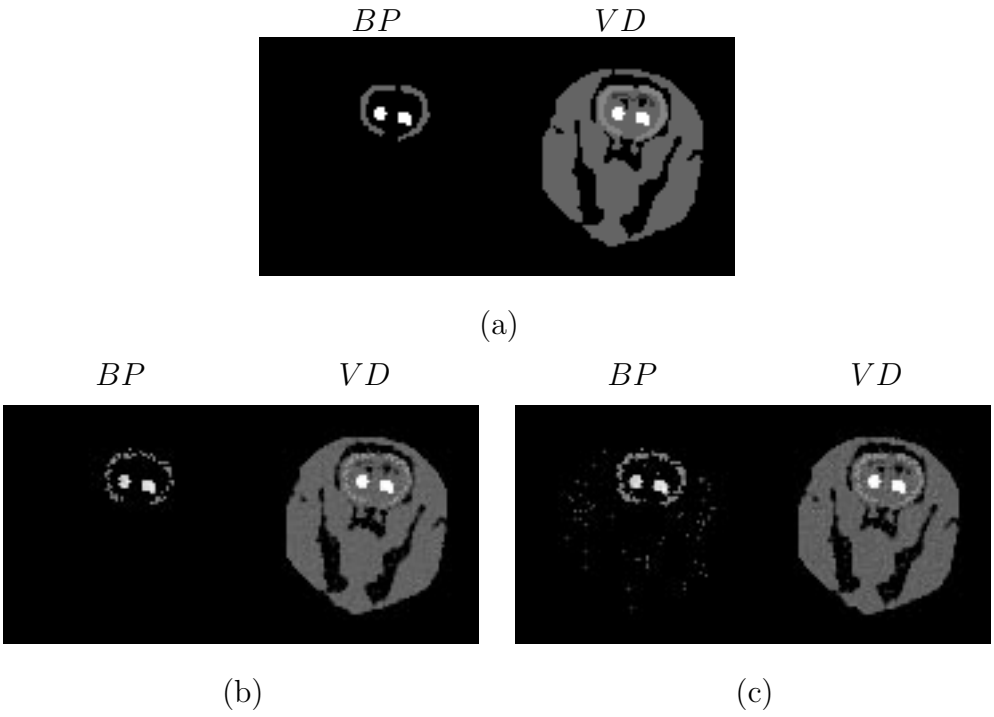


Fig. 3.5. (a) Original BP and VD. Estimated BP and VD (b) estimated with measured input function, and (c) estimated with simultaneous estimation of input function.

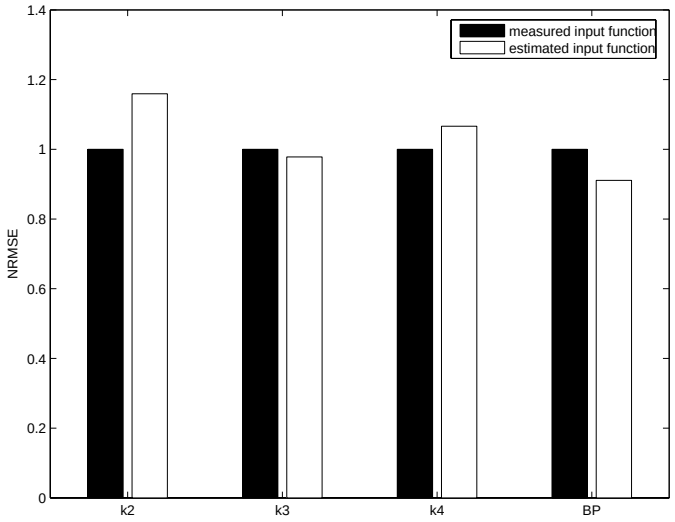


Fig. 3.6. Normalized RMSE of the kinetic parameter estimations with measured input function and with estimated input function.

kinetic parameters have higher RMSE (k_2 and k_4), some of them (k_3 and BP) have either similar or lower RMSE when the input function is estimated.

Better optimization for plasma model parameters may also be needed for avoiding local minima and for computational efficiency. Further simulations and tests with real data are needed to analyze the the RMSE, bias, and variance of the kinetic parameter estimations.

4. DYNAMIC PRINT MODE CONTROL FOR INKJET PRINTING

Abstract

We introduce a novel printing scheme that allows arbitrarily shaped regions in a document to be printed using different modes. Our implementation first segments the document into two basic modes: one pass and four pass. The one pass mode results in greater print speed but has lower quality, whereas the four pass mode results in greater quality but lower print speed. Typically, the one pass mode is more suitable for text and graphics, while the four pass mode is more suitable for continuous tone pictures.

We demonstrate our method using a Lexmark Z52 print engine controlled by a Texas Instruments C6211 DSP. The resulting dynamic print mode control system can be used to continuously trade-off print speed for print quality, or to optimize print quality at a desired print speed.

4.1 Introduction

Typically, ink jet printers use a variety of different print modes to control the quality and speed of printing. In one pass printing, each horizontal motion of the print head is used to print a full swath of the image. This print mode is fastest, but it can result in print quality defects, particularly at swath boundaries. One pass printing is also not robust to defects or nonuniformities in the print nozzles. Alternatively, in multi-pass printing, each pixel is covered by more than one pass of the print head, and the order that dots are fired is determined by a print mask

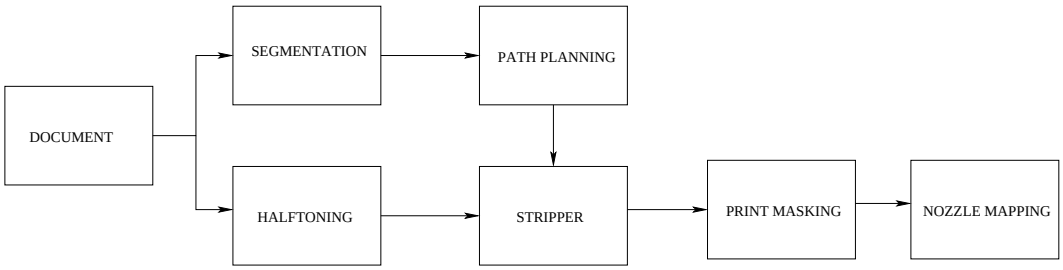


Fig. 4.1. Overview of dynamic print mode system.

[87]. Multi-pass printing is much slower, but it produces much higher print quality, particularly for pictures and continuous tone graphics.

A limitation of typical printers is that they must print the entire document using a single print mode. This means that if a small region of the document requires the quality of multi-pass printing, then the entire document must be printed using multi-pass printing; thereby substantially reducing print speed.

In this chapter, we introduce a novel printing system which we call dynamic print mode control. This system allows arbitrary regions of the document to be printed using different modes. So a small region containing a photo can be printed using four pass bi-directional printing, while the remaining portions of the document can be printed using one pass bi-directional printing. The print mode of each pixel is individually determined by first segmenting the document into one pass and four pass pixel types. For our purposes, we use a very simple segmentation method based on the point-wise gray level of each pixel.

Dynamic print mode control allows for a continuous trade-off between the print-speed and print-quality of the document. It also allows the quality to be optimized at a particular print speed by selecting the multi-pass print modes for the portions of the document that are most sensitive to printing artifacts.

4.2 The Dynamic Print Mode System

Figure 4.1 shows an overview of the architecture we used for implementing the dynamic print mode system. We used a layered architecture to both simplify implementation, and allow for portability to print mechanisms other than the Lexmark Z52 platform used in this work.

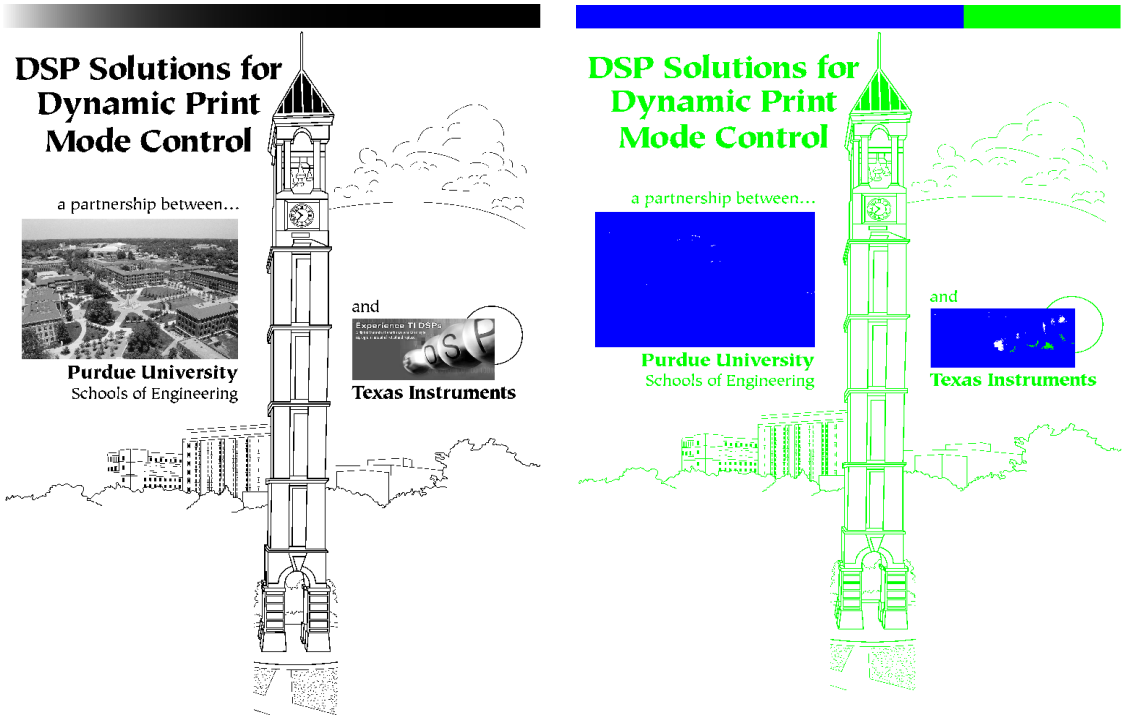
The document to be printed is assumed available as input in 8 bit raster form. For these experiments, we used Adobe Illustrator to render the original postscript document as a 600 dpi 8 bit raster with no anti-aliasing. The document is then processed along two paths. The first path performs binary halftoning. The second path segments the image into regions corresponding to different print modes, and then plans the path that the print head will take as it moves across the page. We refer to each each pass of the print head for a fixed paper position as a print swath.

The final three operations use the segmentation, path plan, and halftone as input. The "stripper" operation extracts the binary image strips corresponding to the path of the inkjet head as it moves across a swath. The print mask then determines exactly which pixels are printed and which are not on each pass of the print head. We will see that the print mask is critical in multipass printing since it can be used to minimize print defects. Finally, the nozzle mapping organizes the output data so that it is time sequenced to suit the specific nozzle format of the printer.

4.2.1 Segmentation

The function of the segmentation step is to label each pixel with its required print mode. For this work, we use three print modes.

- **No print (NP)** - These are pixels which are not printed i.e. white space.
- **1 Pass (1P)** - These are pixels that are printed in one pass print mode.
- **4 Pass (4P)** - These are pixels that are printed in four pass print mode.



(a) (b)

Fig. 4.2. a) The 600 dpi 8 bit raster test image used in the experiments. b) Segmentation of test image into no print (white), 1 pass (blue), and four pass (green) regions.

Figure 4.2 shows the test image and its segmentation into these three modes. We use a very simple segmentation algorithm computed directly from the value of each 8 bit pixel.

gray level 0 $\Leftrightarrow$ No print (NP)

gray level 255 $\Leftrightarrow$ 1 pass (1P)

other gray levels $\Leftrightarrow$ 4 pass (4P)

Note that the segmentation can have regions of arbitrary shape. So, for example, a small region of four pass pixels can exist inside a region of one pass pixels. This

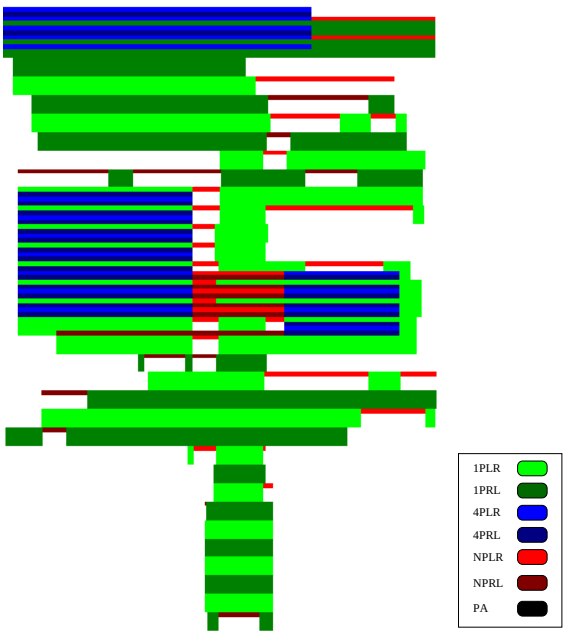


Fig. 4.3. A diagram illustrating the path chosen to allow printing of each pixel with the desired mode.

is very important since it allows the mode of each pixel to be chosen in a manner which is best, independently of other regions in the document.

4.2.2 Path Planning

Once the document is segmented, the path of the print head must be planned so that each pixel can be printed using the desired mode. More specifically, each 1P pixel must be covered by at least one pass of the print head, and each 4P pixel must be covered by exactly four passes of the print head. We present a relatively simple method for path planning, but more complex methods could be used to further minimize head motions.

The basic head motions are given in Table 4.1.

The paper advance (PA) operation moves the paper forward by $N/4$ where $N = 192$ is the height of the print head in pixels for our system. We refer to the first

Code	HeadMotion	Parameters
1PLR	1 pass left to right	Start: First 1P or 4P pixel in the swath Stop: Last 1P or 4P pixel in the swath
1PRL	1 pass right to left	Start: First 1P or 4P pixel in the swath Last: Last 1P or 4P pixel in the swath
4PLR	4 pass left to right	Start: First 4P pixel in the swath Stop: Last 4P pixel in the swath
4PRL	4 pass right to left	Start: First 4P pixel in the swath Stop: Last 4P pixel in the swath
NPLR	No print left to right	
NPRL	No print right to left	
PA	Paper advance by N/4	

Table 4.1
Basic head motions used in the system.

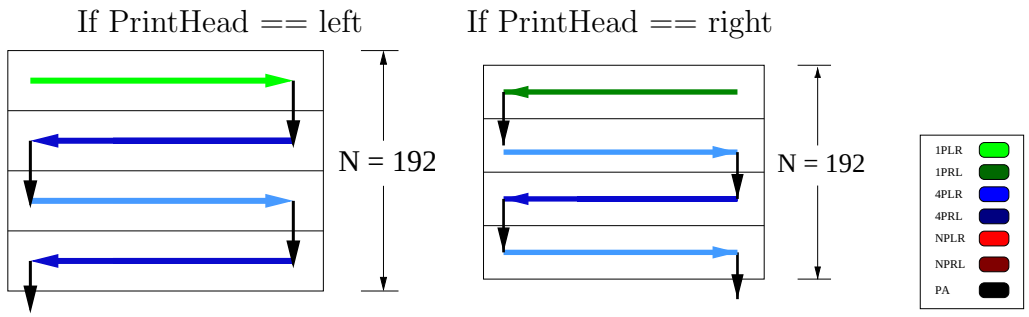
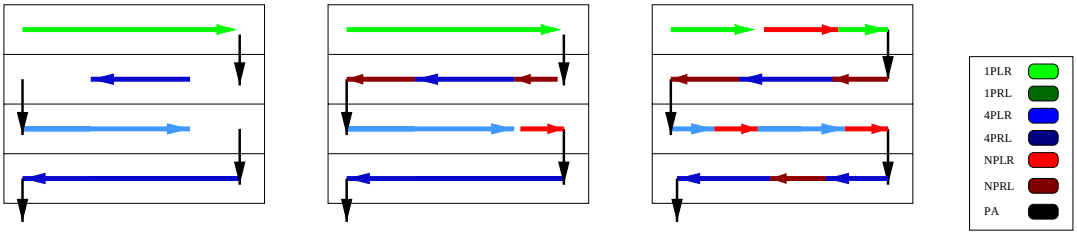


Fig. 4.4. Path plan initialization.

four operations (i.e. 1PLR, 1PRL, 4PLR, 4PRL) as printing head motions, since the remaining operations do not cause ink to be fired from the print head.



(a) (b) (c)

Fig. 4.5. Path plan optimization.

Depending on the initial position of the head, the following initial set of head motions is chosen, shown in Fig. 4.4.

```
If(HeadPosition == left){                                    (4.1)
  NewMotions =
    (1PLR, PA, 4PRL, PA, 4PLR, PA, 4PRL, PA)
}
If(HeadPosition == right){
  NewMotions =
    (1PRL, PA, 4PLR, PA, 4PRL, PA, 4PLR, PA)
}
```

Note that there are four PA operations in the sequence, so at the end of the sequence the paper has advance a full print head height. Next the 1PLR, 1PRL, 4PLR, and 4PRL head motions are processed to determine the horizontal starting and ending position of the head for each motion.

- For 1PLR or 1PRL

```
If(number of 1P or 4P pixels in swath == 0)
  empty = true; otherwise empty = false;
```

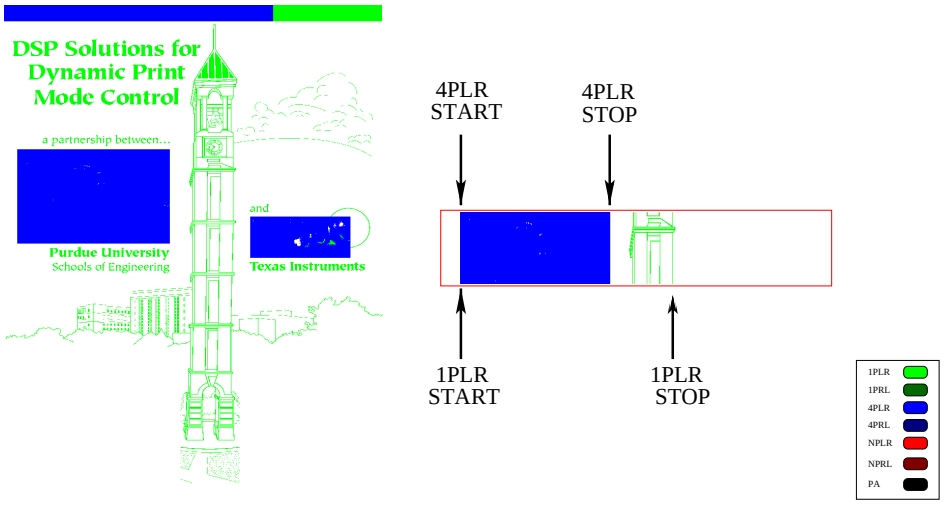


Fig. 4.6. Head motions are processed to determine the horizontal starting and ending position of the print head for each motion.

If(empty == false) {

Start = first position containing 1P or 4P pixel.

Stop = last position containing 1P or 4P pixel.

}

- For 4PLR or 4PRL

If(number of 4P pixels in swath == 0)

empty = true; otherwise empty = false;

If(empty == false) {

Start = first position containing 4P pixel.

Stop = last position containing 4P pixel.

}

Fig. 4.6 shows this process for 1PLR and 4PLR. For 1PLR the path starts from the first 1P or 4P pixel of the document. The stop position is the last 1P or 4P pixel, which is the last pixel. For 4PLR the start position is the first 4P pixel, same with

the 1PLR. However, the stop position is the last 4P pixel, which is the last pixel of the image.

Note that one pass operations consider both 1P and 4P pixels, whereas four pass operations only consider 4P pixels. Since any head motion that is empty results in no printing, it is deleted from the NewMotions sequence.

After deletions of empty head motions, the print head may finish in either the left or right position. For example, the following sequence might result after deletions of empty head motions, leaving the head in the right position.

$$(1PLR, PA, 4PRL, PA, 4PLR, PA, PA) \tag{4.2}$$

Depending on the final state of the head, we set the HeadPostion variable to either $left$ or $right$. Then the new head motions are concatenated to the existing head motions

$$\text{HeadMotions} \leftarrow \text{cat}\left(\text{HeadMotions}, \text{NewHeadMotions}\right),$$

and the process is then repeated starting with (4.1) until the end of the page is reached.

This algorithm produces a list "HeadMotion" which contains a sequence of print head movements that are not necessarily connected horizontally ,shown in Fig. 4.5a. In other words, the horizontal stop position of one head movement may not correspond to the horizontal start position of the subsequent head movement. To fix this, we insert NPLR and NPRL commands as necessary to connect the head motions. Fig. 4.5b shows the reconnected paths. We note that during the NPLR and NPRL motions the head moves substantially faster to reduce printing times.

Finally, the path plan is optimized inserting by NPLR and NPRL operations in white regions of the document. So for example, a command such as 1PLR may contain a white space in which the print head can be accelerated. In this case, the command might be replaced with the sequence 1PLR, NPLR, 1PLR, depending on the size of the white space and the dynamics of the print mechanism. Figure 4.5c shows the optimal path plan.

Figure 4.3 shows the result of applying this path planning algorithm to the test page. For the purposes of illustration, the 1 pass motions are printed as color stripes that are $N = 192$ pixels high, and 4 pass motions are printed as color stripes that are $N/4 = 48$ pixels high. However, the color corresponding to later motions overwrites earlier motions; so some green strips are less than 192 rows high.

4.2.3 Halftoning and Stripping

We use Floyd and Steinberg error diffusion to convert the 300 dpi 8 bit raster document into a 300 dpi binary raster document [88]. Once the halftoning is complete, the halftone, path plan, and segmentation are sent to the "stripper". In our implementation, the stripper extracts the halftone and segmentation data required for each printing head motion. The extracted data is stored in data structures which are individually processed by the print masking function. This approach simplifies implementation, but increases computation and memory requirements.

4.2.4 Print Masking

In four pass printing, each pixel can be printed in any one of four possible passes of the print head. Four pass printing of an image is shown in Fig. 4.7. Each frame shows the pixels of the image that are printed in that pass. The specific pass used to print the pixel is determined by the print mask. Let i be the index of the print nozzle, where $i = 0$ indicates the top row, and $i = 191$ indicates the bottom row of the nozzle. When four pass printing is used, the pass number of nozzle i is given by

$$PixelPassNumber = (191 - i)\%48 \ .$$

Note that the bottom 1/4 of the nozzles are always used to print the first pass, and the top 1/4 are always used to print the last pass.

A dot is fired when the pixel's pass number is equal to the entry in the tiled print mask. We used an 8×8 print mask, as shown in Fig. 4.8. We designed our print

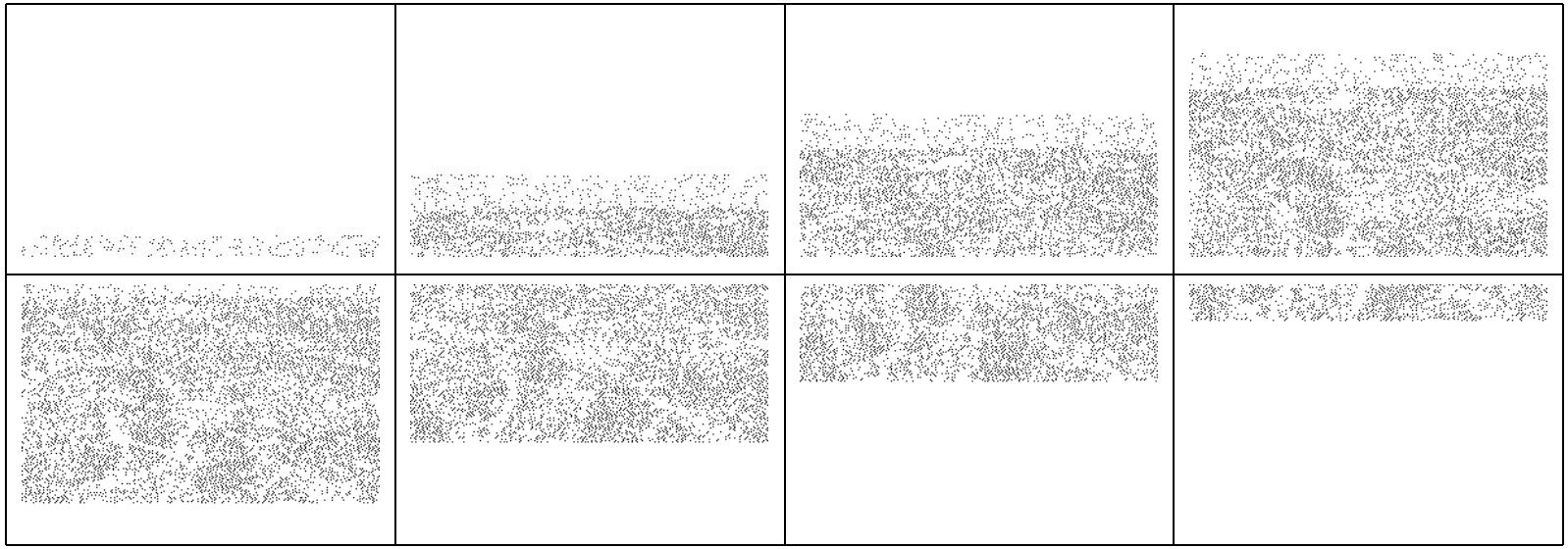


Fig. 4.7. Each pass of the print head prints only certain pixels (and skips the rest) depending on the print mask

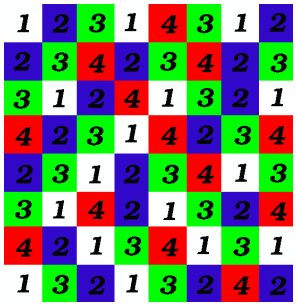


Fig. 4.8. The 8×8 print mask used in the experiments.

```
For each printed strip {
  if(Head Motion == 1PLR or 1PRL) {
    for each pixel in strip {
      test1 = (segmentation[x][y] == 1P);
      test2 = (Mask[x%8][y%8]==PixelPassNumber);
      if(test1 or test2)
        fire dot at position [x,y] on page;
      else
        do not fire dot;
    }
  if(Head Motion == 4PLR or 4PRL) {
    for each pixel in strip {
      test1 = (segmentation[x][y] == 4P);
      test2 = (Mask[x%8][y%8]==PixelPassNumber);
      if(test1 and test2)
        fire dot at position [x,y] on page;
      else
        do not fire dot;
    }
  }
}
```

Fig. 4.9. Pseudo-code illustrating the logic used to fire dots in the dynamic print mode system. Note that one pass printing motions are used to print both one pass and four pass pixels, whereas four pass printing motions are only used to print four pass pixels.

mask to minimize the number of adjacent pixels that are fired sequentially [87]. So the test for firing a 4 pass pixel is

$$Mask[x\%8][y\%8] == PixelPassNumber$$

where x and y are the absolute position of the printed pixel on the page.

Figure 4.9 shows pseudo-code specifying the logic used to fire dots in the dynamic print mode system. The decision to fire a dot is based on both the mode of the print head motion (i.e. 1PLR, 1PRL, 4PLR, or 4PRL) and the class of the pixel (i.e. 1P and 4P). In particular, one pass printing motions are used to print both one pass and four pass pixels, whereas four pass printing motions are only used to print four pass pixels. This is the key factor which allows regions with differing print modes to have arbitrary shape.

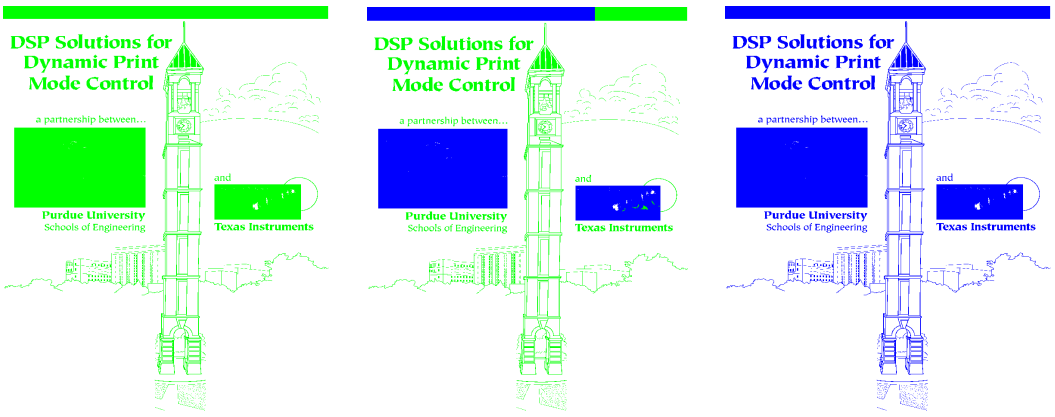
4.2.5 Nozzle Mapping

The last function is nozzle mapping. This function reorganizes the data in a form suitable for sequential input to the print head.

4.2.6 Experimental Results

This dynamic print mode control system was implemented on a Lexmark Z-52 printer mechanism. A Texas Instruments Tiger II DSP board based on the C6211 VLIW DSP was used to control carriage motion, paper motion, and nozzle firing in real-time. The system was successfully demonstrated at the Comdex 2000 exhibition. While the system was developed for full bi-directional printing, we limit our examples to left-right printing in order to eliminate artifacts due to horizontal misregistration.

Figure 4.10 shows the test page processed in three different schemes. In (a), the whole page is printed using one-pass printing mode, in (b) dynamic print mode is used and in (c) whole page is printed using four-pass printing mode. Some of the statistics about these three schemes are given in Table 4.2. Total swath length, the



(a) all one pass (b) dynamic print mode (c) all four pass

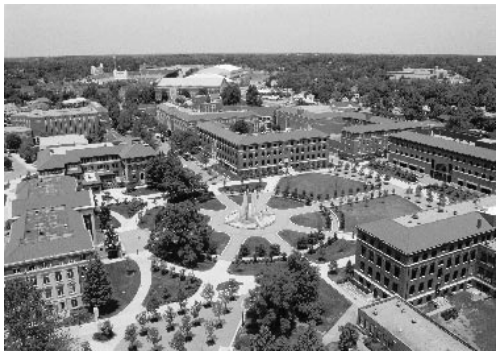
Fig. 4.10. Documents printed in different modes to compare the distance taken by print-head. (a) all one pass (b) dynamic print mode (c) all four pass.

	all one pass	dynamic print mode	all four pass
Number of path segments	99	146	383
Number of swaths	33	64	132
NP swath length	17765	28178	51622
Total swath length	92499	176625	356872

Table 4.2
Distance taken by print head in all one pass, dynamic print mode and all four pass modes.

distance taken by printhead, of the dynamic print mode is approximately half of the length needed by all four pass mode. This means, ignoring other printer dependent delays, dynamic print mode can print a page twice as fast as all four pass print mode, without a significant quality degradation.

Figures 4.11 and 4.12 contain results scanned from actual printer output. Figures 4.11 shows results of printing an image in both one pass and four pass print



(a)

(b) (c)

Fig. 4.11. Image printed using four-pass printing and one-pass printing; a) A 600 dpi original image. The same image printed using b) four-pass printing, and c) one-pass printing.

modes. Note that the image quality is substantially improved by using the four pass print mode since the print mask covers most of the banding artifacts generated by variations in nozzle performance. Figure 4.12 shows portions of the printed output page. Region a), which was printed with one pass mode, shows some defects due to nonuniformities in nozzle output. However, this region is printed with much greater speed. Region b), which was printed with four pass mode, has no visible banding artifacts because they are hidden by the print mask. However, this region is printed at approximately 1/4 the speed of the one pass regions.

(a)

(b)

Fig. 4.12. Portions of printed output page. a) A region printed with one pass mode shows some defects due to miss-firing nozzles. However, this region is printed with greater speed. b) A region printed with four pass mode shows no banding defects because defects are hidden by the print mask. However, this region is printed at approximately 1/4 the speed of the one pass mode.

4.3 Conclusions

We introduced a novel printing scheme that allows arbitrarily shaped regions in a document to be printed using different modes. The one pass mode results in greater print speed but has lower quality; whereas the four pass mode results in greater quality but lower print speed. Typically, the one pass mode is more suitable for text and graphics, and the two pass mode is more suitable for pictures. More

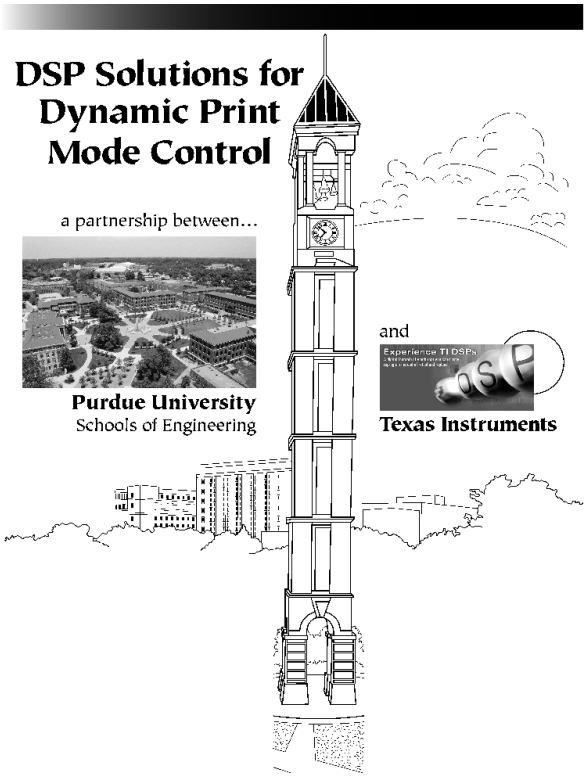


Fig. 4.13. Output page printed using dynamic print mode control and scanned at 600 dpi

sophisticated segmentation algorithms could be used to produce a more optimal trade-off between print quality and print speed.

5. CONCLUSIONS AND FUTURE WORK

In Chapter 1, we introduce a new and efficient algorithm to reconstruct the kinetic parameters of a two-tissue compartment model directly from dynamic PET data. The direct reconstruction approach reduces the dimensionality of the estimation problem, since it does not require tomographic reconstructions. We also proposed the regularization of the kinetic parameters on physiologically relevant kinetic parameters. We show that the proposed algorithm can outperform the image domain methods in terms of RMSE. Furthermore, once the kinetic parameters are reconstructed, the tomographic reconstructions can be computed for any time by solving the compartment model ODE's. In our simulations, we show that the tomographic reconstructions obtained by solving compartment model ODE's have lower RMSE than conventional image reconstruction methods such as FBP and MAP.

In neuroimaging more complicated compartment models may be required for the tracers that have non-specific binding in the brain. The analysis of this type of tracers needs an extra compartment for the non-specific binding. If the specific activity of the tracer is not high enough the model behaves nonlinearly as the receptors are limited in number. The proposed method can be extended to incorporate complicated and nonlinear models with more kinetic parameters. This extension requires the numeric solution of the compartment model ODE's. Our current implementation analytically solves the compartment model ODE's. As the number of kinetic parameters in the model increases, the optimization problem becomes more vulnerable to local minima. Therefore, more advanced optimization techniques may be required. It is possible to use multi-grid [89] optimization to solve the optimization at different resolutions of the parametric image and the sinogram data [90]. Multi-grid methods are shown to

be more resilient to the local minima and converges faster than fixed grid methods for inverse problems [91].

In Chapter 2, we use the proposed PICD algorithm to reconstruct the parametric images of a monkey brain directly from the EXACT HR+ sinograms. We compare the PICD reconstructions with the estimations of the image domain methods. The results show that the proposed direct reconstruction algorithm can produce higher resolution parametric reconstructions.

It is possible to model EXACT HR+ scanner very accurately [71]. We corrected the data for scatter, attenuation, deadtime and detector normalization using CTI software. [1] However it may be better to incorporate these corrections into the forward model. The PICD algorithm may achieve even higher resolutions with a more accurate forward model.

The proposed method can be extended to reconstruct 3D parametric images directly from 3D sinogram data. However, the computational complexity of 3D parametric image reconstruction will be very high.

In Chapter 3, we show that a model-based plasma input function can be simultaneously estimated with the kinetic parameters. However, the simultaneous estimation of the plasma input function with the kinetic parameters is an ill-posed problem. Therefore, the proposed estimation algorithm required an image domain initialization method to avoid local minima. This method can estimate some of the kinetic parameters (k_2, k_3, k_4, BP), but it can only estimate the others (k_1, VD) within a scale factor.

The image domain estimation methods have been researched throughly. Many problems have been aroused and solved. With the proposed direct kinetic parameter reconstruction method, all these problems should be addressed before the adaptation of this algorithm for clinical usage. Some of these problems and our intuitive suggestions are as follows.

[1] CTI Molecular Imaging Inc., Knoxville, TN

Estimation reliability is a measure of confidence for each of the estimated parameters. The partial derivative of the time activity curve with respect to the estimated parameters, given in (5.1) can be used to develop a measure for estimation reliability.

$$\frac{\partial}{\partial \varphi_s} f(\varphi_s, t) = \begin{bmatrix} \frac{\partial}{\partial k_1} f(\varphi_s, t) \\ \frac{\partial}{\partial k_2} f(\varphi_s, t) \\ \frac{\partial}{\partial k_3} f(\varphi_s, t) \\ \frac{\partial}{\partial k_4} f(\varphi_s, t) \end{bmatrix} \tag{5.1}$$

This derivative can be estimated for each voxel in the PICD algorithm.

Parameter identifiability requires the estimation of Hessian for each voxel. The Hessian for voxel s is

$$H_s = \left[\frac{\partial}{\partial \varphi_s} f(\varphi_s, t)\right] \left[\frac{\partial}{\partial \varphi_s} f(\varphi_s, t)\right]^T . \tag{5.2}$$

Once the Hessian is computed for a voxel, the unidentifiable parameters can be detected according to the nondiagonal elements of the Hessian.

Goodness-of-fit analysis is required for any parameter estimation problem that requires fitting. It is a measure that shows the goodness of the fit. For image domain methods the residuals between the reconstructed time-activity curve (TAC) and model output are analyzed to measure the goodness of the fit. The residuals of a good fit should not have any temporal correlation. In other words, the residuals should resemble white noise with zero mean. Temporal structures in the residuals indicates a bad fit. A bad fit may be a result of convergence problems, local minima, and model inadequacy.

The sum of weighted squared residuals given in (5.3) is a commonly used goodness-of-fit measure for the image domain estimation methods.

$$\chi_s^2 = \sum_{k=0}^{K-1} w_{sk}(x_{sk} - f(\varphi_s, t_k))^2 \tag{5.3}$$

In this equation, x_{sk} is the reconstructed time-activity, $f(\varphi_s, t_k)$ is the model output, and w_{sk} is the weight for voxel s at time frame k. The weight, w_{sk}, should be the

inverse of the variance of the measurement at voxel s at time frame k, and it is generally assumed to be $w_{sk} = 1/x_{sk}$.

Another popular measure is the "runs test". In this measure the zero crossing of the residuals are counted. The residuals with less temporal correlation will have more zero crossings.

In direct parameter reconstruction, it may be possible to analyze the residuals in the sinogram domain. The sinogram residuals (shown in figures 2.8,2.9, and 2.10), denoted by Y_{res}, are the difference between the measured sinograms and the forward projection of the compartment model output. Following the notation of Chapter 1,

$$Y_{res} = Y - AF(\varphi) \ . \tag{5.4}$$

If we back project this using FBP, we can obtain the residual of activity curves in time domain. Let $FBP(\cdot)$ denote the FBP reconstruction, then the FBP reconstruction of the residuals in sinogram domain is

$$\begin{aligned} FBP\{Y_{res}\} &= FBP\{Y - AF(\varphi)\} \\ &= FBP\{Y\} - FBP\{AF(\varphi)\} \ . \end{aligned} \tag{5.5}$$

The first term, $FBP\{Y\}$, is equivalent to the FBP reconstruction of the sinogram data, and the second term is similar to the time-activity curves of the model output. Therefore, the FBP of the residuals in sinogram domain is similar to the residuals in time domain. We may use the measures used time domain residual analysis to analyze the back projected sinogram domain residuals.

Compartment models are constructed and validated for well-known tracers. However for new tracers, the corresponding compartment model may not be known. In these cases, the model (ie. the number of compartments and model parameters) should also be selected. In image domain techniques, *model selection* is performed by using an information criteria such as Akaike information criteria (AIC) [92], Bayesian information criteria (BIC) [93], or minimum description length (MDL) [94]. A similar method can be developed for model selection with direct parameter reconstruction algorithm.

In dynamic PET, the total scan duration may quite long. During the acquisition the patient's head can move. Therefore, *interframe image registration and motion compensation* is required before the kinetic parameter estimation. To eliminate this problem a head tracking system is proposed [67]. This system can record the head movements or correct the collected data in an automated fashion. In an other study, the simultaneous estimation of an interframe motion field with the kinetic parameters is proposed [95]. Direct estimation method can adapt similar approaches.

LIST OF REFERENCES

LIST OF REFERENCES

[1] Richard E. Carson and Kenneth Lange. The EM parametric image reconstruction algorithm. *Journal of the American Statistical Association*, 80(389):20–22, 1985.

[2] M. Kamasak, C. A. Bouman, E. D. Morris, and K. Sauer. Direct reconstruction of kinetic parameter images from dynamic PET data. In *Proceedings of 37th Asilomar Conference on Signals, Systems and Computers*, pages 1919–1923, Pacific Grove, CA, November 9-12 2003.

[3] B. T. Christian, T. Narayanan, B. Shi, E. D. Morris, J. Mantil, and J. Mukherjee. Measuring the in vivo binding parameters of [18F]-Fallypride in monkeys using a PET multiple-injection protocol. *Journal of Cereb. Blood Flow and Metab.*, 24(3):309–322, 2004.

[4] S. C. Huang and M. E. Phelps. Principles of tracer kinetic modeling in positron emission tomography. In M. E. Phelps, J. Mazziotta, and H. Schelbert, editors, *Positron Emission Tomography and Autoradiography*, pages 287–346. Raven Press, New York, 1986.

[5] E. D. Morris, C. J. Endres, K. C. Schmidt, B. T. Christian, R. F. Muzic Jr., and R. E. Fisher. Kinetic modeling in PET. In M. Wernick and J. Aarsvold, editors, *Emission Tomography: The Fundamentals of PET and SPECT*, chapter 23. Academic Press, San Diego, 2004.

[6] R. E. Carson. Tracer kinetic modeling in PET. In P. E. Valk, D. L. Bailey, D. W. Townsend, and M. N. Maisey, editors, *Positron Emission Tomography, Basic Science and Clinical Practice*. Springer, London, 2002.

[7] S. S. Kety and C. F. Schmidt. The nitrous oxide method for the quantitative determination of cerebral blood flow in man: theory, procedure and normal values. *J. Clin. Invest.*, 27:476–483, 1948.

[8] L. Sokoloff, M. Reivich, C. Kennedy, M. H. Des Rosiers, C.S. Patlak, K. D. Pettigrew, O. Sakurada, and M. Shinohara. The [14c]deoxyglucose method for the measurement of local cerebral glucose utilization: theory, procedure, and normal values in the conscious and anesthetized albino rat. *Journal of Neurochemistry*, 28(5):897–916, May 1977.

[9] M. E. Phelps, S. C. Huang, E. J. Hoffman, C. Selin, L. Sokoloff, and D. E. Kuhl. Tomographic measurement of local cerebral glucose metabolic rate in humans with (f-18)2-fluoro-2-deoxy-d-glucose: validation of method. *Annals of Neurology*, 6(5):371–388, November 1979.

[10] M. A. Mintun, M. E. Raichle, M. R. Kilbourn, G. F. Wooten, and M. J. Welch. A quantitative model for the in vivo assessment of drug binding sites with positron emission tomography. *Ann. Neurol.*, 15:217–227, 1984.

[11] G. Blomqvist, S. Pauli, L. Farde, L. Ericksson, A. Persson, and C. Halldin. Dynamic models of reversible ligand binding. In C. Beckers, A. Goffiet, and A. Bol, editors, *Clinical Research and Clinical Diagnosis*. Kluwer Academic Publishers, 1989.

[12] V.J. Cunningham, S. P. Hume, G. R. Price, R. G. Ahier, J. E. Cremer, and A. K. Jones. Compartmental analysis of diprenorphine binding to opiate receptors in the rat in vivo and its comparison with equilibrium data in vitro. *Journal of Cereb. Blood Flow and Metab.*, 11(1):1–9, 1991.

[13] S. P. Hume, R. Myers, P. M. Bloomfield, J. Opacka-Juffry, J. E. Cremer, R. G. Ahier, S. K. Luthra, D. J. Brooks, and A. A. Lammertsma. Quantitation of carbon-11-labeled raclopride in rat striatum using positron emission tomography. *Synapse*, 27:4754, 1992.

[14] A. A. Lammertsma, C. J. Bench, S. P. Hume, S. Osman, K. Gunn, D. J. Brooks, and R. S. Frackowiak. Comparison of methods for analysis of clinical [11c]raclopride studies. *Journal of Cereb. Blood Flow and Metab.*, 16(1):42–52, 1996.

[15] A. A. Lammertsma and S. P. Hume. Simplified reference tissue model for PET receptor studies. *NeuroImage*, 4:153–158, 1996.

[16] R. N. Gunn, A. A. Lammertsma, S. P. Hume, and V. J. Cunningham. Parametric imaging of ligand-receptor binding in PET using a simplified reference region model. *NeuroImage*, 6(4):279–287, 1997.

[17] H. Watabe, R. E. Carson, and H. Iida. The reference tissue model: Three compartments for the reference region. *NeuroImage*, 11(6):S12, 2000.

[18] J. Logan, J. S. Fowler, N. D. Volkow, A. P. Wolf, D. J. Schlyer S. L. Dewey, R. R. MacGregor, R. Hitzemann, S. J. Gatley B. Bendriem, and D. R. Christman. Graphical analysis of reversible radioligand binding from time-activity measurements applied to [n-11c-methyl]-(-)-cocaine PET studies in human subjects. *Journal of Cereb. Blood Flow and Metab.*, 10(5):740–747, 1990.

[19] Joseph A. Thie, Gary T. Smith, and Karl F. Hubner. Linear least squares compartmental-model-independent parameter identification in PET. *IEEE Trans. on Medical Imaging*, 16(1):11–16, 1997.

[20] K. Chen, M. Lawson, E. Reiman, A. Cooper, D. Feng, S. C. Huang, D. Bandy, and D. Ho. Generalized linear least squares method for fast generation of myocardial blood flow parametric images with n-13 ammonia PET. *IEEE Trans. on Medical Imaging*, 17:236–243, 1998.

[21] R. E. Carson, S. C. Huang, and M. V. Green. Weighted integration method for local cerebral blood flow measurements with positron emission tomography. *Journal of Cereb. Blood Flow and Metab.*, 6:245–258, 1986.

[22] C. S. Patlak and R. G. Blasberg. Graphical evaluation of blood-to-brain transfer constants from multiple-time uptake data: generalizations. *Journal of Cereb. Blood Flow and Metab.*, 5:584–590, 1985.

[23] D. Fang, Z. Wang, and S. C. Huang. A study on statistically reliable and computationally efficient algorithms for generating local cerebral blood flow parametric images with positron emission tomography. *IEEE Trans. on Medical Imaging*, 12(2):182–188, June 1993.

[24] Grant T. Gullberg. Dynamic SPECT imaging: Exploring a new frontier in medical imaging. In *IEEE International Symposium on Biomedical Imaging*, pages 607–610, Arlington, VA, April 15-18 2004.

[25] F. O'Sullivan and A. Saha. Use of ridge regression for improved estimation of kinetic constants from PET data. *IEEE Trans. on Medical Imaging*, 18(2):115–125, February 1999.

[26] S. C. Huang and Y. Zhou. Spatially-coordinated regression for image-wise model fitting to dynamic PET data for generating parametric images. *IEEE Trans. on Nuclear Science*, 45(3):1194–1199, June 1998.

[27] Y. Zhou and S. C. Huang. Linear ridge regression with spatial constraint for generation of parametric images in dynamic positron emission tomography studies. *IEEE Trans. on Nuclear Science*, 48(1):125–130, 2001.

[28] Y. Zhou, C.J. Endres, J.R. Brasic, S.C. Huang, and D.F. Wong. Linear regression with spatial constraint to generate parametric images of ligand-receptor dynamic PET studies with a simplified reference tissue model. *Neuroimage*, 18:975–989, 2003.

[29] M. Senda Y. Kimura and N.M. Alpert. Fast formation of statistically reliable FDG parametric images based on clustering and principal components. *Phys. Med. Biol.*, 47:455–468, August 2002.

[30] Y. Zhou, S.C. Huang, M. Bergsneider, and D. F. Wong. Improved parametric image generation using spatial-temporal analysis of dynamic PET studies. *NeuroImage*, 15:697–707, 2002.

[31] Chien-Min Kao, Jeffrey T. Yap, Jogeshwar Mukherjee, and Miles Wernick. Image reconstruction for dynamic PET based on low-order approximation and restoration of the sinogram. *IEEE Trans. on Medical Imaging*, 16(6):738–749, 1997.

[32] Manoj V. Narayanan, Micheal A. King, Miles N. Wernick, Charles L. Byrne, Edward J. Soares, and P. Hendrik Pretorius. Improved image quality and computation reduction in 4d reconstruction of cardiac-gated SPECT images. *IEEE Trans. on Medical Imaging*, 19(5):423–433, 2000.

[33] Miles N. Wernick, E. James Infusino, and Milos Milosevic. Fast spatio-temporal image reconstruction for dynamic PET. *IEEE Trans. on Medical Imaging*, 18(3):185–195, March 1999.

[34] Thomas E. Nichols, Jinyi Qi, Evren Asma, and Richard M. Leahy. Spatiotemporal reconstruction of list-mode PET data. *IEEE Trans. on Medical Imaging*, 21(4):396–404, 2002.

[35] B. W. Reutter, G. T. Gullberg, and R. H. Huesman. Direct least-squares estimation of spatiotemporal distributions from dynamic SPECT projections using a spatial segmentation and temporal b-splines. *IEEE Trans. on Nuclear Science*, 19(5):434–450, 2000.

[36] B. W. Reutter, G. T. Gullberg, and R. H. Huesman. Accuracy and precision of compartmental model parameters obtained from directly estimated dynamic SPECT time-activity curves. *IEEE Trans. on Nuclear Science*, 51(1):170–176, 2004.

[37] L. Shepp and Y. Vardi. Maximum likelihood reconstruction for emission tomography. *IEEE Trans. on Medical Imaging*, MI-1(2):113–122, October 1982.

[38] M. A. Limber, M. N. Limber, A. Cellar, J. S. Barney, and J. M. Borwein. Direct reconstruction of functional parameters for dynamic SPECT. *IEEE Trans. on Nuclear Science*, 42(4):1249–1256, 1995.

[39] Gengsheng L. Zeng, Grant T. Gullberg, and Ronald H. Huesman. Using linear time-invariant system theory to estimate kinetic parameters directly from projection measurements. *IEEE Trans. on Nuclear Science*, 42(6):2339–2346, 1995.

[40] Andreas Robert Formiconi. Least squares algorithm for region-of-interest evaluation in emission tomography. *IEEE Transactions on Medical Imaging*, 12(1):90–100, March 1993.

[41] Elenora Vanzi, Andreas Robert Formiconi, Dino Bindi, Giuseppe La Cava, and Alberto Pupi. Kinetic parameter estimation from renal measurements with a three-headed SPECT system: A simulation study. *IEEE Trans. on Medical Imaging*, 23(3):363–373, 2004.

[42] Ronald H. Huessman, Bryan W. Reutter, G. Larry Zeng, and Grant T. Gullberg. Kinetic parameter estimation from SPECT cone-beam projection measurements. In *International Meeting on Fully 3-D Image Reconstruction Conference Record*, pages 121–125, 1997.

[43] Ping-Chun Chiao, W. Leslie Rogers, Neal H. Clinthorne, Jeffrey A. Fessler, and Alfred O. Hero. Model-based estimation for dynamic cardiac studies using ECT. *IEEE Transactions on Medical Imaging*, 13(2):217–226, June 1994.

[44] P-C. Chiao, W.L. Rogers, J.A. Fessler, N.H. Clinthorne, and A.O. Hero. Model-based estimation with boundary side information or boundary regularization. *IEEE Trans. on Medical Imaging*, 13(2):227–234, June 1994.

[45] Steven R. Meikle, Julian C. Matthews, Vincent J. Cunningham, Dale L. Bailey, Lefteris Livieratos, Terry Jones, and Pat Price. Parametric image reconstruction using spectral analysis of PET projection data. *Phys. Med. Biol.*, 43:651–666, 1998.

[46] Julian Matthews, Dale Bailey, Pat Price, and Vin Cunningham. The direct calculation of parametric images from dynamic PET data using maximum-likelihood iterative reconstruction. *Phys. Med. Biol.*, 42:1155–1173, 1997.

[47] R. E. Carson. A maximum likelihood method for region-of-interest evaluation in emission tomography. *Journal of Computer Assisted Tomography*, 10(4):654–663, 1986.

[48] T. Farncombe, A. Celler, C. Bever, D. Noll, J. Maeght, and R. Harrop. The incorporation of organ uptake into dynamic SPECT (dSPECT) image reconstruction. *IEEE Trans. on Nuclear Science*, 48(1):3–9, 2001.

[49] Mustafa E. Kamasak, Charles A. Bouman, Evan D. Morris, and Ken Sauer. Direct reconstruction of kinetic parameter images from dynamic PET data. *IEEE Trans. on Medical Imaging*, 24(5), 2005. to appear.

[50] M. Liptrot, K. H. Adams, L. Martiny, L. H. Pinborg, M. N. Lonsdale, N. V. Olsen, S. Holm, C. Svarer, and G. M. Knudsen. Cluster analysis in kinetic modelling of the brain: a noninvasive alternative to arterial sampling. *NeuroImage*, 21(2):483–493, 2004.

[51] L. M. Wahl, M. C. Asselin, and C. Nahmias. Regions of interest in the venous sinuses as input functions for quantitative PET. *Journal of Nuclear Medicine*, 40(10):1666–1675, 1999.

[52] K. Chen, D. Bandy, E. Reiman, S. C. Huang, M. Lawson, D. Feng, L. Yun, and A. Palant. Noninvasive quantification of the cerebral metabolic rate for glucose using positron emission tomography, 18F-fluoro-2deoxyglucose, the patlak method, and an image-derived input function. *Journal of Cereb. Blood Flow and Metab.*, 18(7):716–723, 1998.

[53] Dagan Feng, Koon-Pong Wong, Chi-Ming Wu, and Wan-Chi Siu. A technique for extracting physiological parameters and the required input function simultaneously from PET image measurements: Theory and simulation study. *IEEE Transactions on Information Technology in Biomedicine*, 1(4):243–254, December 1997.

[54] Koon-Pong Wong, Dagan Feng, Steven R. Meikle, and Michael J. Fulham. Simultaneous estimation of physiological parameters and the input function - in *vivo* PET data. *IEEE Transactions on Information Technology in Biomedicine*, 5(1):67–76, March 2001.

[55] C. A. Bouman and K. Sauer. A unified approach to statistical tomography using coordinate descent optimization. *IEEE Trans. on Image Processing*, 5(3):480–492, March 1996.

[56] E. Mumcuoglu, R. M. Leahy, S. Cherry, and E. Hoffman. Accurate geometric and physical response modeling for statistical image reconstruction in high resolution PET scanners. In *Proceedings of IEEE Nuclear Science Symposium and Medical Imaging Conference*, pages 1569–1573, 1996.

[57] Thomas Frese, Ned C. Rouze, Charles A. Bouman, Ken Sauer, and Gary D. Hutchins. Quantitative comparison of FBP, EM, and bayesian reconstruction algorithms, including the impact of accurate system modeling, for the IndyPET scanner. *IEEE Trans. on Medical Imaging*, 22(2):258–276, 2003.

[58] H. W. Kuhn and A. W. Tucker. Nonlinear programming. In *Proceedings of 2nd Berkeley Symposium on Mathematical Statistics and Probabilistics*, pages 481–492. University of California Press, 1951.

[59] K. Sauer and C. A. Bouman. A local update strategy for iterative reconstruction from projections. *IEEE Trans. on Signal Processing*, 41(2):534–548, February 1993.

[60] G. Paxinos and C. Watson. *The Rat Brain in Stereotaxic Coordinates*. Academic Press, 4th edition, 1998.

[61] S. Pappata, S. Dehaene, J. B. Poline, M. C. Gregoire, A. Jobert, J. Delforge, V. Frouin, M. Bottlaender, F. Dolle, L. Di Giamberardino, and A. Syrota. In vivo detection of striatal dopamine release during reward: a PET study with [^{11}C]raclopride and a single dynamic scan approach. *Neuroimage*, 16(4):1015–1027, 2002.

[62] S. S. Saquib, C. A. Bouman, and K. Sauer. ML parameter estimation for Markov random fields with applications to Bayesian tomography. *IEEE Trans. on Image Processing*, 7(7):1029–1044, July 1998.

[63] H.M. Hudson and R.S. Larkin. Accelerated image reconstruction using ordered subsets of projection data. *IEEE Trans. on Medical Imaging*, 13(4):601–609, December 1994.

[64] M. Slifstein and M. Laruelle. Effects of statistical noise on graphical analysis of PET neuroreceptor studies. *Journal of Nuclear Medicine*, 41(12):2083–2088, 2000.

[65] M. Ichise, H. Toyama, R. B. Innis, and R. E. Carson. Strategies to improve neuroreceptor parameter estimation by linear regression analysis. *Journal of Cereb. Blood Flow and Metab.*, 22:1271–1281, 2002.

[66] J. Logan, J. S. Fowler, N. D. Volkow, Y. S. Ding, G. J. Wang, and D. L. Alexoff. A strategy for removing the bias in the graphical analysis method. *Journal of Cereb. Blood Flow and Metab.*, 21(3):307–320, 2001.

[67] P. M. Bloomfield, T. J. Spinks, J. Reed, L. Schnorr, A. M. Westrip, L. Livieratos, R. Fulton, and T. Jones. The design and implementation of a motion correction scheme for neurological PET. *Phys. Med. Biol.*, 48(8):959–978, 2003.

[68] M. Defrise and P. Kinahan. Data acquisition and image reconstruction for 3D PET. In B. Bendriem and D. W. Townsend, editors, *The Theory and Practise of 3D PET*, pages 11–53. Kluwer Academic Publishers, 1998.

[69] Thomas E. Nichols. *Spatiotemporal Modeling of Positron Emission Tomography*. PhD thesis, Carnegie Mellon University, Department of Statistics, Carnegie Mellon University Pittsburgh, Pennsylvania, 15213, January 2001.

[70] Lars-Eric Adam, Joachim Zaers, Hermann Ostertag, Herbert Trojan, Matthias E. Bellermann, and Gunnar Brix. Performance evaluation of the whole-body PET scanner ECAT EXACT HR+ following the IEC standard. *IEEE Trans. on Nuclear Science*, 44(3):1172–1179, 1997.

[71] Jinyi Qi, Richard M. Leahy, Chinghan Hsu, Thomas H. Farquhar, and Simon R. Cherry. Fully 3D bayesian image reconstruction for the ECAT EXACT HR+. *IEEE Trans. on Nuclear Science*, 45(3):1096–1103, 1998.

[72] Frederic H. Fahey. Data acquisition in PET imaging. *J. Nucl. Med. Technol.*, 30(2):39–49, 2002.

[73] CTI-PET. ECAT software operating instructions, version 7.2. Manual, October 1999. (Part Number 9340029-00 Rev A).

[74] Michel Defrise, Paul Kinahan, David Townsend, Christian Michel, Merence Sibomana, and Danny Newport. Exact and approximate rebinning algorithms for 3D PET data. *IEEE Trans. on Medical Imaging*, 16(2):145–158, 1997.

[75] Bradley T. Christian, Tanjore K. Narayanan, Bingzhi Shi, and Jogeshwar Mukherjee. Quantitation of striatal and extrastriatal D-2 dopamine receptors using PET imaging of ^{18}Ffallypride in nonhuman primates. *Synapse*, 38(1):71–79, October 2000.

[76] Micheal E. Phelps, John C. Maziotta, and Heinrich R. Schelbert. *Positron Emission Tomography and Autoradiography, Principles and Applications for the Brain and Heart*, chapter 7. Raven Press, Raven Press, 1140 Avenue of the Americas, New York, NY 10036, 1986.

[77] S. Takikawa, V. Dhawan, P. Spetsieris, W. Robeson, T. Chaly, R. Dahl, D. Margouleff, and D. Eidelberg. Noninvasive quantitative fluorodeoxyglucose PET studies with and estimated input function derived from a population-based arterial blood curve. *Radiology*, 188:131–136, 1993.

[78] Y. Onishi, Y. Yonekura, S. Nishizawa, F. Tanaka, H. Okazawa, K. Ishiza, T. Fujita, J. Konishi, and T. Mukai. Noninvasive quantification of iodine-123-iomazenil SPECT. *Journal of Nuclear Medicine*, 37(2):374–378, 1996.

[79] S. Eberl, A. R. Anayat, R. R. Fulton, P. K. Hooper, and M. J. Fulham. Evaluation of two population-based input functions for quantitative neurological FDG PET studies. *Eur. J. Nucl. Med.*, 24(3):299–304, 1997.

[80] J.-E. Litton. Input function in PET brain-studies using MRI defined arteries. *J. Comp. Ass. Tom.*, 21(6):907–909, 1997.

[81] E. D. Morris, J. W. Vanmeter, T. A. Tasciyan, J. M. Maisog, and T. A. Zeffiro. Automated determination of the arterial input function for quantitative MR perfusion analysis. In *Proc. Intl. Soc. Magn. Reson. Med.*, page 740, 2000.

[82] Dagan Feng, Sung-Cheng Huang, and Xinmin Wang. Models for computer simulation studies of input functions for tracer kinetic modeling with positron emission tomography. *International Journal of Biomedical Computing*, 32(2):95–110, 1993.

[83] A. Mohammad-Djafari. Joint estimation of parameters and hyperparameters in a Bayesian approach of solving inverse problems. In *Proc. of IEEE Int'l Conf. on Image Proc.*, volume II, pages 473–476, Lausanne, Switzerland, September 16-19 1996.

[84] G. J. Hunter, L. M. Hamberg, N. M. Alpert, N. C. Choi, and A. J. Fischman. Simplified measurement of deoxyglucose utilization rate. *Journal of Nuclear Medicine*, 37(6):950–955, 1996.

[85] Koon-Pong Wong, Dagan Feng, Steven R. Meikle, and Michael J. Fulham. Segmentation of dynamic PET images using cluster analysis. *IEEE Trans. on Nuclear Science*, 49(1):200–207, 2002.

[86] N. C. Rouze, K. M. Stantz, and G. D. Hutchins. Design of IndyPET-II, a high-resolution, high-sensitivity dedicated research scanner. In *Proceedings of Nuclear Science Symposium Conference*, pages 1545–1549, 2001.

[87] J. Yen, M. Carlsson, M. Chang, J.M. Garcia, and H. Nguyen. Constraint solving for inkjet print mask design. *Journal of Electronic Imaging*, 44(5):391–397, Sept./Oct. 2000.

[88] R. W. Floyd and L. Steinberg. An adaptive algorithm for spatial greyscale. *Journal of the Society for Information Display*, 17(2):75–77, 1976.

[89] W. L. Briggs, V. E. Henson, and S. F. McCormick. *A Multigrid Tutorial, 2nd Ed.* Society for Industrial and Applied Mathematics, Philadelphia, 2000.

[90] Seungseok Oh, Charles A. Bouman, and Kevin J. Webb. Multigrid inversion algorithms for poisson noise model-based tomographic reconstruction. In *Proc. of SPIE Conf. on Electronic Imaging: Computational Imaging*, volume 5674, pages 455–467, San Jose, CA, January 16-21 2005.

[91] Seungseok Oh, Adam B. Milstein, Charles A. Bouman, and Kevin J. Webb. A general framework for nonlinear multigrid inversion. *IEEE Trans. on Image Processing*, 14(1), January 2005.

[92] H. Akaike. A new look at the statistical model identification. *IEEE Trans. Automatic Control*, AC-19(6):716–723, December 1974.

[93] G. Schwarz. Estimating the dimension of a model. *The Annals of Statistics*, 6:461–464, 1978.

[94] J. Rissanen. A universal prior for integers and estimation by minimum description length. *The Annals of Statistics*, 11(2):417–431, September 1983.

[95] David R. Gilland, Bernard A. Mair, James E. Bowsher, and Ronald J. Jaszczak. Simulatenous reconstruction and motion estimation for gated cardiac ECT. *IEEE Trans. on Nuclear Science*, 49(5):2344–2349, 2002.

APPENDIX

APPENDIX A
PSEUDOCODE FOR PICD ALGORITHM

$\varphi \leftarrow$ ParametricReconstruct(sinograms){

for each iteration {

for each voxel s {

$[\theta_1,\theta_2] \leftarrow$ ComputeDerivatives(sinograms,φ_s)

$\tilde{\varphi}_s \leftarrow \varphi_s$

$[\alpha, \beta] \leftarrow \text{ConvolveWithPlasma}(c_s, d_s, [t_0, \cdots, t_{K-1}], \lambda, V_B, C_P)$

$\tilde{c}_s \leftarrow \arg\min_{\tilde{c}_s} \left\{ \text{CostFunction}([\tilde{a}_s, \tilde{b}_s, \tilde{c}_s, \tilde{d}_s], \varphi_s, \alpha, \beta, \theta_1, \theta_2, \{\varphi_r : r \in \partial s\}, W) \right\}$

$[\tilde{\alpha}, \tilde{\beta}] \leftarrow \text{ConvolveWithPlasma}(\tilde{c}_s, \tilde{d}_s, [t_0, \cdots, t_{K-1}], \lambda, V_B, C_P)$

$[\tilde{a}_s, \tilde{b}_s] \leftarrow \text{EstimateAandB } (\tilde{\varphi}_s, \varphi_s, \tilde{\alpha}, \tilde{\beta}, \alpha, \beta, \theta_1, \theta_2, \{\varphi_r : r \in \partial s\}, W)$

$l \leftarrow \arg\min_{l} \left\{ \text{CostFunction}([\tilde{a}_s, \tilde{b}_s, \tilde{c}_s + l, \tilde{d}_s + l], \varphi_s, \alpha, \beta, \theta_1, \theta_2, \{\varphi_r : r \in \partial s\}, W) \right\}$

$\tilde{c}_s \leftarrow \tilde{c}_s + l; \quad \tilde{d}_s \leftarrow \tilde{d}_s + l$

$[\tilde{\alpha}, \tilde{\beta}] \leftarrow \text{ConvolveWithPlasma}(\tilde{c}_s, \tilde{d}_s, [t_0, \cdots, t_{K-1}], \lambda, V_B, C_P)$

$[\tilde{a}_s, \tilde{b}_s] \leftarrow \text{EstimateAandB}(\tilde{\varphi}_s, \varphi_s, \tilde{\alpha}, \tilde{\beta}, \alpha(c_s), \beta(d_s), \theta_1, \theta_2, \{\varphi_r : r \in \partial s\}, W)$

$\varphi_s \leftarrow \tilde{\varphi}_s$

}

}

}

cost $\leftarrow$ CostFunction$(\tilde{\varphi}_s, \varphi, \alpha, \beta, \theta_1, \theta_2, \{\varphi_r : r \in \partial s\}, W)$ {

$[\tilde{\alpha}, \tilde{\beta}] \leftarrow \text{ConvolveWithPlasma}(\tilde{c}_s, \tilde{d}_s, [t_0, \cdots, t_{K-1}], \lambda, V_B, C_P)$

$[\tilde{a}_s, \tilde{b}_s] \leftarrow \text{EstimateAandB}(\tilde{\varphi}_s, \varphi_s, \tilde{\alpha}, \tilde{\beta}, \alpha, \beta, \theta_1, \theta_2, \{\varphi_r : r \in \partial s\}, W)$

$\text{cost} \leftarrow \text{DeltaCost}(\tilde{\varphi}_s, \varphi_s, \tilde{\alpha}, \tilde{\beta}, \alpha, \beta, \theta_1, \theta_2, \{\varphi_r : r \in \partial s\}, W)$

}

$[\tilde{a}, \tilde{b}] \leftarrow \text{EstimateAandB}(\tilde{\varphi}_s, \varphi_s, \tilde{\alpha}, \tilde{\beta}, \alpha, \beta, \theta_1, \theta_2, \{\varphi_r : r \in \partial s\}, W)\{$

for $L_{ab}/3$ iterations {

$\Delta C \leftarrow \text{DeltaCost}([\tilde{a}, \tilde{b}, \tilde{c}_s, \tilde{d}_s], \varphi_s, \tilde{\alpha}, \tilde{\beta}, \alpha, \beta, \theta_1, \theta_2, \{\varphi_r : r \in \partial s\}, W)\}$

$\frac{dC}{da} \leftarrow \frac{-1}{\epsilon}\{\text{DeltaCost}([\tilde{a} + \epsilon, \tilde{b}, \tilde{c}_s, \tilde{d}_s], \varphi_s, \tilde{\alpha}, \tilde{\beta}, \alpha, \beta, \theta_1, \theta_2, \{\varphi_r : r \in \partial s\}, W)$ - $\Delta C\}$

$\frac{dC}{db} \leftarrow \frac{-1}{\epsilon}\{\text{DeltaCost}([\tilde{a}, \tilde{b} + \epsilon, \tilde{c}_s, \tilde{d}_s], \varphi_s, \tilde{\alpha}, \tilde{\beta}, \alpha, \beta, \theta_1, \theta_2, \{\varphi_r : r \in \partial s\}, W)$ - $\Delta C\}$

if $\tilde{a} = 0$ and $\frac{dC}{da} < 0$ then $\frac{dC}{da} \leftarrow 0$

if $\tilde{b} = 0$ and $\frac{dC}{db} < 0$ then $\frac{dC}{db} \leftarrow 0$

if $\left\{|\frac{dC}{da}| + |\frac{dC}{db}|\right\} > 0$ then {

$\left(\frac{dC}{da}, \frac{dC}{db}\right) \leftarrow \frac{\left(\frac{dC}{da}, \frac{dC}{db}\right)}{\sqrt{\frac{dC}{da}^2 + \frac{dC}{db}^2}}$

$\zeta \leftarrow \arg\min_{\zeta \in [0,1]} \text{DeltaCost}([\tilde{a} + \zeta\frac{dC}{da}, \tilde{b} + \zeta\frac{dC}{db}, \tilde{c}_s, \tilde{d}_s], \varphi_s, \tilde{\alpha}, \tilde{\beta}, \alpha, \beta, \theta_1, \theta_2, \{\varphi_r : r \in \partial s\}, W)$

$\tilde{a} \leftarrow \tilde{a} + \zeta\frac{dC}{da}; \quad \tilde{b} \leftarrow \tilde{b} + \zeta\frac{dC}{db}$

}

}

$\Delta C \leftarrow \text{DeltaCost}(\tilde{\varphi}_s, \varphi_s, \tilde{\alpha}, \tilde{\beta}, \alpha, \beta, \theta_1, \theta_2, \{\varphi_r : r \in \partial s\}, W) \ \{$

$\Delta f \leftarrow \tilde{a}_s\tilde{\alpha} + \tilde{b}_s\tilde{\beta} - a_s\alpha - b_s\beta$

$\Delta C \leftarrow \Delta f\theta_1 + \frac{1}{2}\|\Delta f\|^2_{\theta_2} + \sum_{r\in\partial s} g_{s-r}\|T(\tilde{\varphi}_s) - T(\varphi_r)\|^2_W$

}

VITA

VITA

Mustafa E. Kamasak received his B.S. and M.S. degrees in Electrical Engineering from Boğaziçi University, Istanbul, Turkey in 1997 and 1999, respectively. He is currently working towards a Ph.D. degree at School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN. His research interests include image processing, medical imaging, and pattern recognition.