Multiresolution imaging of in vivo ligand-receptor interactions

P. Thévenaz[a] and P. Millet[b]

[a]Swiss Federal Institute of Technology Lausanne
[b]Hôpitaux universitaires de Genève

ABSTRACT

The aim of this study is to obtain voxel-by-voxel images of binding parameters between [^{11}C]-flumazenil and benzodiazepine receptors using positron emission tomography (PET). We estimate five local parameters (k_1, k_2, $B'_{\max}$, $k_{\text{on}}/V_{\text{R}}$, k_{off}) by fitting a three-compartment ligand-receptor model for each voxel of a PET time series.

It proves difficult to fit the ligand-receptor model to the data. We trade noise and spatial resolution to get better results. Our strategy is based on the use of a multiresolution pyramid. It is much easier to solve the problem at coarse resolution because there are fewer data to process. To increase resolution, we expand the parameter maps to the next finer level and use them as initial solution to further optimization, which then proceeds at a fast pace and is more likely to escape false local minima.

For this approach to work optimally, the residue between data at a given pyramid level and data at the next level must be as small as possible. We satisfy this constraint by working with spline-based least-squares pyramids. To achieve speed, the optimizer must be efficient, particularly when it is nearing the solution. To that effect, we have developed a Marquardt-Levenberg algorithm that exhibits superlinear convergence properties.

Keywords: PET, benzodiazepine, parametric imaging, multiresolution pyramid, spline

1. INTRODUCTION

Receptors are sites that interact with extracellular physiological signals called neurotransmitters or *ligands*, and that convert them into intracellular effects.[1] Ligand-binding methods have been enormously important for mapping the anatomical distribution of different neurotransmitter receptors in the brain.

Several researchers attempted to obtain voxel-by-voxel images of ligand-binding parameters between [^{11}C]-flumazenil and benzodiazepine receptors using Positron Emission Tomography (PET) in human volunteers.[2–4] When a molecule is known to interact selectively with specific receptors, which happens for the flumazenil molecule and the benzodiazepine receptor, one can use the molecule to probe in vivo the regional distribution of, and affinity to, the relevant binding site. The procedure requires that one administers flumazenil molecules that are labeled with a positron emitting radionuclide, which then allows one to monitor the time course of the local tracer concentration. The corresponding data-acquisition protocol involves multiple injections of radioligand; their effect is usually recorded as an extensive time-series of PET volumes. By fitting one three-compartment ligand-receptor model per voxel, five physiologically relevant local parameters can be estimated (k_1, k_2, $B'_{\max}$, $k_{\text{on}}/V_{\text{R}}$, k_{off}).

However, so as not to harm the patient, it is common to severely limit the amount of injected radioactivity when acquiring PET scans. This typically results in noisy data. In the analysis of ligand-receptor interactions, the problem is exacerbated because the cumulated dose of radioactivity must be shared between injections. Moreover, the duration of individual volume acquisitions is short because it is necessary to achieve a sufficient time resolution when acquiring the series of PET volumes, which results in even noisier data. For these reasons, it proves difficult to fit the ligand-receptor model to the data in a way that is sound from a physiological point of view.

An additional difficulty comes from the complexity of the ligand-receptor model which requires an iterative optimization procedure to achieve the fit, with the risk of getting trapped into a non-global optimum. This model complexity also adds to the computational burden which is already large due to the amount of data to process.

Further author information:
P. Thévenaz: Biomedical Imaging Group, DMT/IOA, BM-Ecublens 4.137, CH–1015 Lausanne EPFL, Switzerland, E-mail: philippe.thevenaz@epfl.ch, URL: http://bigwww.epfl.ch/
P. Millet: Unité de neuroimagerie psychiatrique, Chemin du Petit-Bel-Air 2, CH–1225 Chêne-Bourg, Switzerland, E-mail: philippe.millet@dim.hcuge.ch

Medical Imaging 2001: Physiology and Function from Multidimensional Images, Chin-Tu Chen, Anne V. Clough, Editors, Proceedings of SPIE Vol. 4321 (2001) © 2001 SPIE · 1605-7422/01/$15.00

To get better results, a natural idea is to trade noise and spatial resolution. Traditionally, one selects a volume of interest and averages data within. In this paper instead, we propose an alternative strategy based on the use of a spatial multiresolution pyramid. Starting at its coarsest level, it is much easier to solve the problem because there are fewer data to process and because data have been denoised by the averaging effect of the pyramid; the price to pay is low spatial resolution. To achieve better resolution, we expand the parameter maps to the next finer level and use them as initial solution when applying the optimization procedure at this finer level. Since the initial conditions are already good estimates of the true solution, the optimization can proceed at a fast pace and is more likely to escape false local minima. We proceed similarly until we have reached the desired spatial resolution.

For this approach to work optimally, it is important that the residue between data at a given pyramid level and data at the next level be as small as possible. We satisfy this constraint by working with spline-based pyramids that are optimal in the sense that the residue between levels is minimized with respect to a mean-squares criterion. To achieve speed, it is crucial that the optimizer be efficient, particularly when it is nearing the solution. To that effect, we use a Marquardt-Levenberg algorithm which exhibits superlinear convergence properties near the solution. The superlinear regime is available at once, without the need to explore the surroundings of the solution in the parameter space.

We have obtained parametric images both with, and without the proposed multiresolution strategy. We are able to show that the proposed strategy is more robust since it results in images that are in better agreement with the expected physiology; meanwhile, the computation time is significantly reduced. Finally, the freedom to stop the optimization at any resolution level offers us the possibility to perform measurements in a time and at a resolution that can be clinically acceptable.

2. COMPARTMENTAL MODEL

2.1. Structure and Equations

Figure 1 shows the compartment model that is used in this study. It is a nonequilibrium, nonlinear model that includes three compartments, one for the concentration of each molecule variety (unmetabolized flumazenil in plasma, free and nonspecific ligand that has left the plasma, ligand bound to receptor sites), and five parameters.[5,6] The parameters k_1 and k_2 are associated with the exchange of molecules between the plasma and the free ligand compartment; $B'_{\max}$ represents the concentration of receptors available for binding; k_{on} and k_{off} are the association and dissociation rate constants between free and bound ligand, respectively. The goal of this study is to adjust these parameters to fit the model of Figure 1 to the experimental data. Unfortunately, the quantities k_{on} and V_{R} cannot be estimated separately, where V_{R} is a dimensionless coefficient that controls the volume of reaction—after correction for tissue inhomogeneities. Only the ratio $k_{\text{on}}/V_{\text{R}}$ can be identified.

The model of Figure 1 has often been used to study in the human brain the binding of flumazenil, which is a convenient ligand because it is an antagonist with high affinity and selectivity for central benzodiazepine receptors.[7–10] In this paper, three injections of flumazenil (FMZ) are performed. A first injection of labeled [^{11}C]FMZ at time T_1 is followed by a displacement injection at time T_2, where the dose contains only unlabeled FMZ ligand. The last injection at time T_3 contains a mixture of FMZ and [^{11}C]FMZ. The unlabeled and labeled ligand kinetics are assumed to be similar; thus, the model contains two parts with the same structure and the same parameters.

To estimate the arterial concentration of labeled and unlabeled ligand, $C^*_{\text{p}}(t)$ and $C_{\text{p}}(t)$, respectively, we draw a series of plasma samples and measure $C^*_{\text{pm}}(t)$, the concentration of labeled* ligand in the plasma. This concentration must be corrected because of metabolite buildup; we proceed as follows:

$$C^*_{\text{p}}(t) = \begin{cases} C^*_{\text{pm}}(t)\ (A_1\ \exp(-B_1\,(t-T_1)) + A_2\ \exp(-B_2\,(t-T_1))) & T_1 < t \le T_3 \\ C^*_{\text{pm}}(t)\ (A_1\ \exp(-B_1\,(t-T_3)) + A_2\ \exp(-B_2\,(t-T_3))) & T_3 < t, \end{cases} \tag{1}$$

*We can assume with confidence that our injections provide the sole source of labeled ligand. This confidence vanishes for unlabeled ligand, because some amount is naturally present at all times in the body due to physiological processes. Thus, C_{p} cannot be measured directly but must be inferred from C^*_{p}.

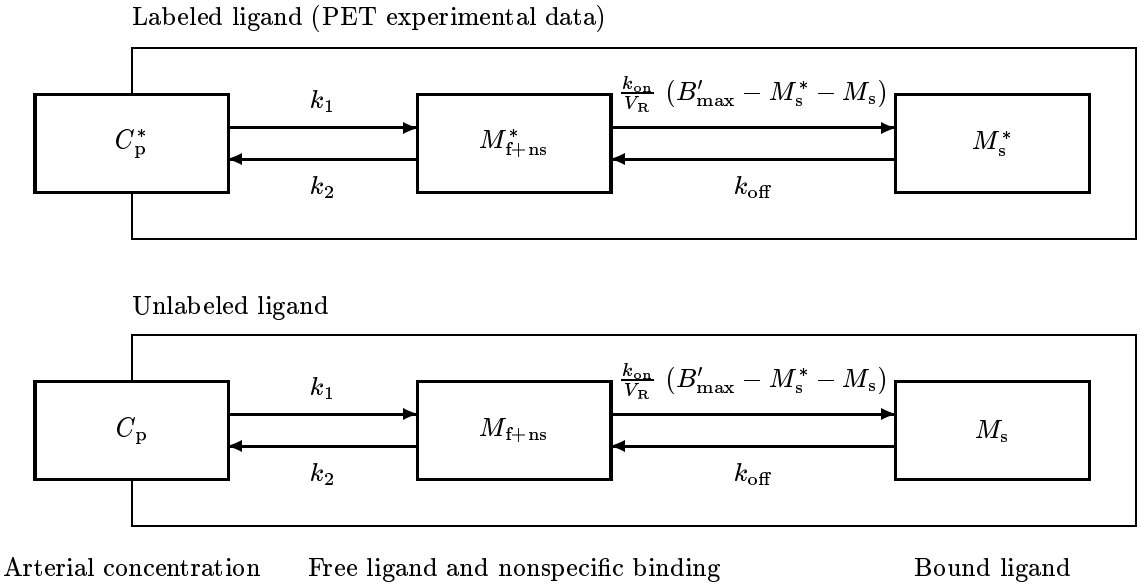

Figure 1. Model of the flow of FMZ concentration

where A_n and B_n represent the bi-exponential coefficients used for metabolite correction. The unlabeled ligand concentration related to the injection dose is then given by

$$C_{\mathrm{p}}(t) = \begin{cases} 0 & t \le T_2 \\ \frac{J_2}{J_1^*}\, C_{\mathrm{p}}^*(t - T_2) & T_2 < t \le T_3 \\ \frac{J_3}{J_3^*}\, C_{\mathrm{p}}^*(t - T_3) + C_{\mathrm{p}}(T_3)\, \frac{C_{\mathrm{p}}^*(t)}{C_{\mathrm{pm}}^*(t)} & T_3 < t, \end{cases} \tag{2}$$

where J_i^* and J_i correspond to the dose at time t_i of labeled and unlabeled ligand, respectively.

Finally, the system of differential equations that describes Figure 1 can be established as follows:

$$\frac{\mathrm{d}M_{\mathrm{f+ns}}^*}{\mathrm{d}t} = k_1\, C_{\mathrm{p}}^*(t) - k_2\, M_{\mathrm{f+ns}}^*(t) - \frac{\mathrm{d}M_{\mathrm{s}}^*}{\mathrm{d}t} \tag{3}$$

$$\frac{\mathrm{d}M_{\mathrm{s}}^*}{\mathrm{d}t} = \frac{k_{\mathrm{on}}}{V_{\mathrm{R}}}\, (B'_{\mathrm{max}} - M_{\mathrm{s}}^*(t) - M_{\mathrm{s}}(t))\ M_{\mathrm{f+ns}}^*(t) - k_{\mathrm{off}}\, M_{\mathrm{s}}^*(t) \tag{4}$$

$$\frac{\mathrm{d}M_{\mathrm{f+ns}}}{\mathrm{d}t} = k_1\, C_{\mathrm{p}}(t) - k_2\, M_{\mathrm{f+ns}}(t) - \frac{\mathrm{d}M_{\mathrm{s}}}{\mathrm{d}t} \tag{5}$$

$$\frac{\mathrm{d}M_{\mathrm{s}}}{\mathrm{d}t} = \frac{k_{\mathrm{on}}}{V_{\mathrm{R}}}\, (B'_{\mathrm{max}} - M_{\mathrm{s}}^*(t) - M_{\mathrm{s}}(t))\ M_{\mathrm{f+ns}}(t) - k_{\mathrm{off}}\, M_{\mathrm{s}}(t), \tag{6}$$

where Equations 4 and 6 include the interaction between the labeled and the unlabeled ligand.

The experimental data $f(\mathbf{x}; t_i)$ at time t_i collected by the PET scanner at location $\mathbf{x}$ measures the accumulated radioactive decay between time t_{i-1} and t_i. We model it by

$$g(\mathbf{x}; t_i) = \frac{1}{t_i - t_{i-1}} \int_{t_{i-1}}^{t_i} \left(M_{\mathrm{f+ns}}^*(\mathbf{x}; t) + M_{\mathrm{s}}^*(\mathbf{x}; t) + F_{\mathrm{V}}\, C_{\mathrm{pm}}^*(\mathbf{x}; t) \right) \mathrm{d}t,$$

where the parameter F_{V} represents the fraction of blood present in the tissue volume, and where the various concentrations all depend on the five parameters we want to measure, thanks to the model of Figure 1.

2.2. Experimental Protocol

The protocol consists of three injections of radiolabeled flumazenil and/or cold flumazenil. At the beginning ($T_1 = 0$), about 15 mCi of [^{11}C]FMZ were injected intravenously over a 1-minute period. At $T_2 = 39$ min., an intravenous injection of 0.01 mg/kg of unlabeled ligand was performed (displacement injection). At $T_3 = 69$ min., a third injection of a mixture of labeled (about 9 mCi at time of the injection) and unlabeled flumazenil in the same syringe (coinjection), was performed. Although this third injection introduced a lower amount of radioactivity than the first, the corresponding amount of labeled flumazenil was much higher because of the physical decay of [^{11}C] activity that took place between T_1 and T_3.

The whole experiment lasted about two hours, during which fifty sequential PET scans of size $128 \times 128 \times 32$ and of increasing duration (from 30 seconds to 5 minutes) were reconstructed after each injection of of radiolabeled and/or cold FMZ.

The identification of the five model parameters $\boldsymbol{\mu} = (k_1,\ k_2,\ B'_{\max},\ k_{\text{on}}/V_{\text{R}},\ k_{\text{off}})$ requires the knowledge of the blood time-activity curves $C^*_{\text{pm}}(t)$. During the study, seventy arterial blood samples were collected from the radial artery. The time interval between each sample varied from five seconds (during the two minutes following each injection of labeled FMZ) to 10 minutes (when the change in the blood radioactivity concentration slowed down). The time-activity curves were then corrected for the physical decay of [^{11}C] activity and the plasma radioactivity concentrations were translated to FMZ time-concentration curves using the corresponding specific radioactivity of [^{11}C]FMZ. The values of the coefficients for metabolite correction were $A_1 = 0.7$, $A_1 = 0.13$, $B_1 = 0.35$, and $B_2 = 0$.

Finally, the value of the fraction of blood present in the tissue volume is assumed to be $F_{\text{V}} = 0.04$.

2.3. Parameter Identification

In this study, we identify for each position $\mathbf{x}$ the model parameters $\boldsymbol{\mu}(\mathbf{x})$ by minimizing the weighted least-squares cost function defined by

$$\varepsilon^2(\mathbf{x};\boldsymbol{\mu}) = \sum_{i=1}^{m} w_i(\mathbf{x})\,(f(\mathbf{x};t_i) - g(\mathbf{x};\boldsymbol{\mu},t_i))^2$$

where $\{t_i : i \in [1 \ldots m]\}$ is the set of data sampling times, where $w_i(\mathbf{x})$ is the weight associated at time t_i with the decay-corrected measures $f(\mathbf{x};t_i)$ at position $\mathbf{x}$, and where $g(\mathbf{x};\boldsymbol{\mu},t_i)$ is the value predicted by the model with parameters $\boldsymbol{\mu}$, according to the experimental protocol. The cost function ε^2 is minimized by using a Marquardt-Levenberg algorithm.[11] The parameter values that minimize ε^2 give the best estimate and are noted by $\boldsymbol{\mu}_0$.

We have no analytical solution to the system of differential equations 3–6, especially considering the effect of the experimental boundary values $C^*_{\text{pm}}(t)$ given in Equation 1; for this reason, we proceed numerically. Similarly, the gradient $\partial\varepsilon^2/\partial\mu$ is estimated by a finite difference.

The choice of a weighting matrix is a very important step. Absolute scaling is paramount since the identification results are identical when the weights are multiplied by a constant factor, but their relative values should reflect the relative credibility of the measures. If the variance denoted by $\text{var}(f(\mathbf{x};t_i))$ is assumed to be known for each measure $f(\mathbf{x};t_i)$, then the optimal choice for weights is given by[12]

$$w_i(\mathbf{x}) = \frac{1}{\text{var}(f(\mathbf{x};t_i))}.$$

In practice however, the value of data variance is not often well known. Here, we have used the weight defined by

$$w_i(\mathbf{x}) = \frac{t_i - t_{i-1}}{f(\mathbf{x};t_i)},$$

which is a reasonable choice in experiments involving radioactivity counts in an image.[13]

Figure 2 shows typical time-concentration curves obtained with our multi-injection protocol. These curves correspond to the large regions of interest (ROI) used in most quantification studies and include from about 150 to 300 pixels, both from receptor-rich and receptor-poor structures, along with an ROI consisting of a whole slice of the brain that contains an average density of receptors. It is easy to discern the first increase in radioligand concentration after the injection at time $T_1 = 0$ min. of [^{11}C]FMZ, which reaches a maximum that depends on the receptor density, and gently decays thereafter. The concentration then plummets at time $T_2 = 39$ min., due to the displacement

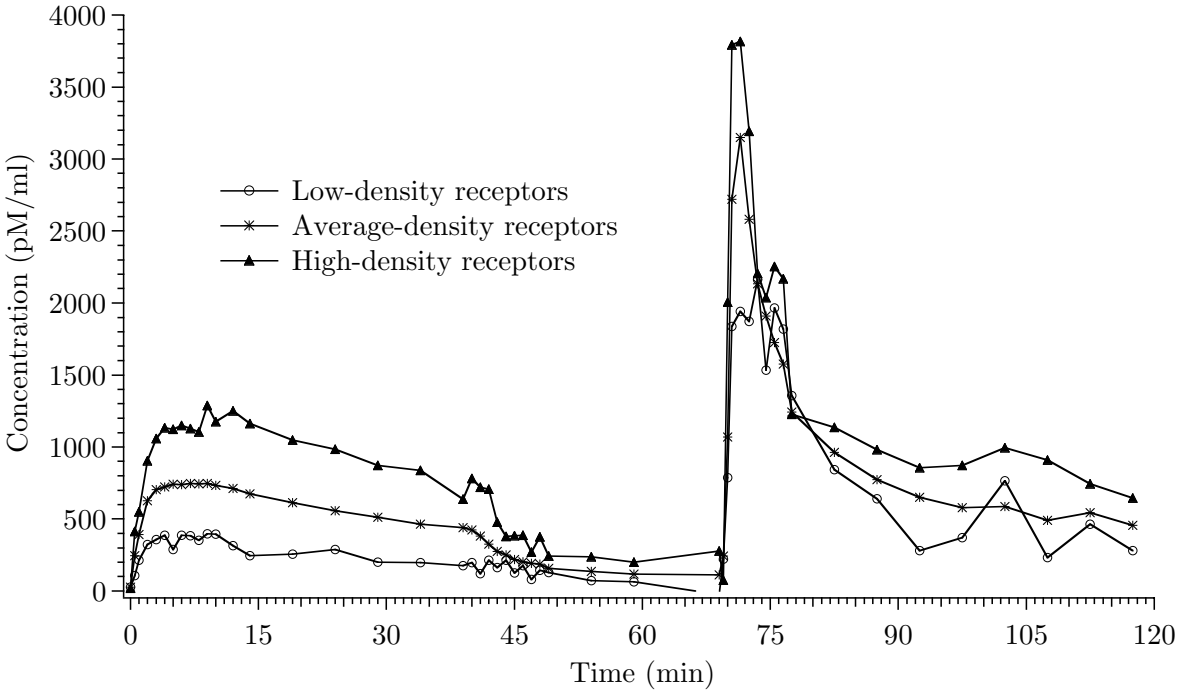

Figure 2. Several time-concentration curves $\sum_{\mathbf{x}\in\mathbb{D}} f(\mathbf{x};t)$ averaged over different regions of interest $\mathbb{D}$

injection that introduces a large amount of unlabeled ligand which quickly washes the labeled bound ligand from the receptor sites, denoting a fast dissociation rate. Finally, the coinjection of labeled and unlabeled FMZ at time $T_3 = 69$ min. produces a sharply rising peak in the time-concentration curves due to the large amount of injected FMZ, after which the decay of the radioactive ligand resumes.

2.4. Error Analysis

We estimate the parameter uncertainties by using both the information matrix and the measure of the residual difference between experimental data and model-predicted values. This methodology provides an approximate parameter covariance matrix. From this matrix, we deduce an estimation of the standard error of each parameter.

The starting points of most studies on parameter error estimates are the sensitivity functions which are defined as the derivatives $\partial g(\mathbf{x};\boldsymbol{\mu},t_i)/\partial\mu_k$ and can always be calculated using a numerical procedure. Let $s_k(\mathbf{x};\boldsymbol{\mu},t_i)$ be the sensitivity function corresponding to a parameter μ_k and let S be the so-called sensitivity matrix defined by

$$[S(\mathbf{x})]_{i,k} = s_k(\mathbf{x};\boldsymbol{\mu},t_i).$$

Moreover, let r be the number of parameters (five in the present case), and let W be the diagonal matrix composed of the elements $w_i(\mathbf{x})$ which is called the weighting matrix. The matrix $M = S^\top W\, S$ is called the information matrix. If the data variances are unknown and if the weighting matrix is correctly justified, an approximation $\tilde{C}$ of the covariance matrix C can be constructed as follows[12,14–16]:

$$\tilde{C} = s^2\, M^{-1},$$

where

$$s^2 = \frac{\varepsilon(\boldsymbol{\mu}_0)}{m-r}.$$

Finally, assuming that the sampling distribution of the estimate $\boldsymbol{\mu}_0$ is approximately Gaussian, an estimate of the standard deviation $\Delta\mu_k = \sqrt{\mathrm{var}(\mu_k)}$ of the parameter μ_k is given by the square root of the k-th diagonal component of $\tilde{C}$, denoted by $[\tilde{C}]_k$, as can be seen in[12,14]

$$\Delta\mu_k = \sqrt{[\tilde{C}]_k} = s\,\sqrt{[M^{-1}]_k}.$$

3. MULTIRESOLUTION

3.1. Motivation

Given the time-series of data $f(\mathbf{x};t)$ and measurements $C^*_{\mathrm{pm}}(t)$, our goal is the estimation of the parameter map $\boldsymbol{\mu}_0(\mathbf{x})$. If we proceed directly, as described in Section 2.3, then we have to face a difficult problem because the high level of noise inherent in the experimental data will perturb the solution of the system of differential equations 3–6. To lessen the influence of noise, we are willing to trade it with spatial resolution: noise reduction is generally achieved by any of a number of smoothing mechanisms, the simplest being the selection of a region of interest and the averaging of data within—we shall soon consider a better method.

Meanwhile, our optimization is based on a Marquardt-Levenberg algorithm. Thus, we must first obtain an initial solution before we let the optimizer refine it. Since no a priori knowledge is available, this initial map of parameters is usually selected to be constant $\forall\mathbf{x}$, which may markedly differ from the optimal solution. Thus, the possibility that the optimizer gets trapped into a non-global minimum during its search is likely. There are two complementary ways to avoid this condition. First, a reduction in the complexity of the data (i.e., of the data amount) will lead to a less convoluted criterion function $\varepsilon^2(\boldsymbol{\mu})$, which in turn will obliterate many a local minimum. This increases the odds of the optimizer to locate the global optimum. Second, a better initial solution will not only help bypass non-global optima, it will also bring large speed benefits.

One may think that a better initial condition leads to a faster optimization simply because the optimizer has less territory to explore before it finds $\boldsymbol{\mu}_0$. While this is true, the effect of this mechanism becomes irrelevant when compared to a far more potent one, namely, the superlinear property that a Marquardt-levenberg optimizer exhibits near an optimum. This superlinearity strongly boosts the speed of convergence; unfortunately, it comes at the price of a poorer performance than some optimizers (e.g., steepest gradient) when the actual parameters being refined are remote from a minimum. Thus, the availability of a good starting condition is a very critical aspect of the Marquardt-Levenberg optimizer.

A dyadic multiresolution pyramid offers an elegant solution to both the noise problem and to the initial condition problem. Let $f^{(j)}(\mathbf{x};t)$ be the representation of our data at resolution level indexed by j, where $(j+1)$ is the next finer level, up to the original (finest) resolution J of the pyramid. For all resolutions but the finest one, the noise will have been reduced by the smoothing property of a pyramid. Moreover, the data will have been made less complex by the downsampling operation. Both effects concur to the avoidance of non-global optima and lead to increased robustness.

Furthermore, suppose that we have at our disposal the map of optimal parameters $\boldsymbol{\mu}_0^{(j)}$ obtained by the optimizer at the j-th level. Then, we can expand this map to the next resolution and use it as initial condition for optimizing at level $(j+1)$, where the amount of details that tell the two resolutions apart is hopefully small because most of the energy of practically all images or biomedical volumes is concentrated toward low frequencies. This will favor the establishment of the superlinear regime of the Marquardt-Levenberg algorithm. By proceeding recursively, we can benefit from speed and robustness all the way up to the final J-th level.

Performing optimization at the coarsest pyramid level is an easy task too, even without the benefit of an informed initial solution: there is few data to process, which allows for a thorough exploration of the space of parameters $\boldsymbol{\mu}$ at virtually no computational cost. Moreover, the paucity of data tends to simplify the topology of the criterion $\varepsilon^2(\boldsymbol{\mu})$; in addition, the data has undergone the largest amount of smoothing, which favors a well-defined optimum of the criterion.

While building the pyramid, we smooth level $(j+1)$ and then downsample it to get level j. For optimal noise reduction, it is important that the amount of smoothing before downsampling be just right, so that no aliasing be present, nor oversmoothing. For best performance, it is also important that the amount of detail that distinguishes one pyramid level from the next be as small as possible. We satisfy these two conditions by using a least-squares pyramid based on splines.[17] The essential property of this pyramid is that the fidelity of any level to the next finer level is the largest possible, in a least-squares sense.

The pyramid algorithm that we cited[17] is most general. We will shortly derive a version that applies to the specific context of a dyadic pyramid; but first, we must establish the grounds for a continuous model of our data.

3.2. Interpolation Model

Let β^n be the B-spline of degree n defined as the n-fold convolution of $\beta^0(x) = \frac{1}{2}\left(\mathrm{sign}(x+\frac{1}{2}) - \mathrm{sign}(x-\frac{1}{2})\right)$. Some of its many interesting properties are a finite-support of width $(n+1)$, an $(n+1)$-th order of approximation, and multiresolution space embedding.

The cardinal spline η^n is an interpolating function that satisfies $\eta^n(k) = \delta_k$ and that shares the properties of β^n mentioned above, except that it has an infinite support. Since its approximation order is also $(n+1)$, it closely approximates the sinc function, and is very appropriate for interpolation in theory; in practice, its infinite support forbids its direct use, an inconvenience that we are about to lift. It is defined by

$$\eta^n(x) = \sum_{k_0 \in \mathbb{Z}} \left(b^n\right)^{-1}(k_0)\, \beta^n(x - k_0),$$

where the discrete sequence $\left(b^n\right)^{-1}(k)$ has the z-transform

$$\left(b^n\right)^{-1}(k) \quad \overset{z}{\longleftrightarrow} \quad \frac{1}{\beta^n(0) + \sum_{k=1}^{\lfloor \frac{n}{2} \rfloor} \beta^n(k)\left(z^k + z^{-k}\right)}. \tag{7}$$

We can thus easily express a continuous model of a function f given by a sequence of samples $f(k)$ as

$$\begin{aligned} f(x) &= \sum_{k_1 \in \mathbb{Z}} f(k_1)\, \eta^n(x - k_1) \\ &= \sum_{k_1 \in \mathbb{Z}} f(k_1) \sum_{k_0 \in \mathbb{Z}} \left(b^n\right)^{-1}(k_0)\, \beta^n(x - k_0 - k_1) = \sum_{k_2 \in \mathbb{Z}} \underbrace{\sum_{k_1 \in \mathbb{Z}} f(k_1)\, \left(b^n\right)^{-1}(k_2 - k_1)}_{c_{k_2}}\, \beta^n(x - k_2) \\ &= \sum_{k \in \mathbb{Z}} c_k\, \beta^n(x - k), \end{aligned}$$

where we have just converted the infinite support of the interpolating basis function η^n into the finite support of β^n, without approximation, but at the cost of the determination of coefficients c_k. As seen in the derivation, these coefficients can be obtained by digital filtering; an efficient procedure is described in.[18] Note that, to obtain the continuous model for volumetric instead of linear data, it is enough to consider the tensor product of splines.

3.3. Pyramid Algorithms

Multiresolution space embedding is the most relevant property of splines for pyramids. For an odd degree n, it allows one to express an expanded B-spline as a linear weighted sum of normal B-splines

$$\beta^n\left(\frac{x}{2}\right) = \sum_{k \in \mathbb{Z}} u_2^n(k)\, \beta^n(x - k),$$

where u_2^n is the binomial filter defined by

$$u_2^n(k) = \begin{cases} 2^{-n} \begin{pmatrix} n+1 \\ k + \frac{n+1}{2} \end{pmatrix} & |k| \le \frac{n+1}{2} \\ 0 & \frac{n+1}{2} < |k|. \end{cases}$$

3.3.1. Expansion

Let $f_0(x) = \sum c_0(k)\, \beta^n(x-k)$ be a known function. We want to construct a version $f_{-1}(x) = \sum c_{-1}(k)\, \beta^n(2\,x - k)$ that is expanded by a factor two and that is as close as possible to f_0, in a least-squares sense (in fact, we show that we can even reach perfection with $f_0 = f_{-1}$). This algorithm will be used to obtain $\boldsymbol{\mu}^{(j)}$ out of $\boldsymbol{\mu}_0^{(j+1)}$.

$$f_0(x) = \sum_{k \in \mathbb{Z}} c_0(k)\, \beta^n(x-k) = \sum_{k \in \mathbb{Z}} c_0(k) \sum_{l \in \mathbb{Z}} u_2^n(l)\, \beta^n(2\,x - 2\,k - l) = \sum_{m \in \mathbb{Z}} \underbrace{\sum_{k \in \mathbb{Z}} c_0(k)\, u_2^n(m - 2\,k)}_{c_{-1}(m)}\, \beta^n(2\,x - m).$$

In short, the expansion of the coefficients is given by $c_{-1}(k) = \left(u_2^n * \uparrow_2 (c_0)\right)(k)$, where the operator $\uparrow_2(\cdot)$ upsamples its argument, and where the operator $*$ represents a discrete convolution such that $(u * v)(k) = \sum_{l \in \mathbb{Z}} u(l)\, v(k-l)$.

3.3.2. Reduction

Let $\langle f(\cdot), g(\cdot) \rangle = \int_{-\infty}^{\infty} f(x)\, g(x)\, \mathrm{d}x$ represent the scalar product of two functions. Let $f_1(x) = \sum c_1(k)\, \beta^n(x/2 - k)$ be a reduced version of f_0. We want that the residual error $(f_1 - f_0)$ be orthogonal to the space of functions that can be represented by the coarser basis function $\beta^n(\cdot/2)$. This is equivalent to mean-square minimization, or to asking that $\langle f_1 - f_0, f_1 - f_0 \rangle$ be minimal. Thus, we write

$$0 = \langle f_1(\cdot) - f_0(\cdot), \beta^n(\frac{\cdot}{2} - k) \rangle \quad \forall k \in \mathbb{Z}.$$

Developing this requirement, we obtain

$$\sum_{l \in \mathbb{Z}} c_1(l)\, \langle \beta^n(\frac{\cdot}{2} - l), \beta^n(\frac{\cdot}{2} - k) \rangle = \sum_{l \in \mathbb{Z}} c_0(l)\, \langle \beta^n(\cdot - l), \beta^n(\frac{\cdot}{2} - k) \rangle.$$

Because of space embedding, and since a B-spline is defined as an n-fold convolution of a symmetric function,

$$2 \sum_{l \in \mathbb{Z}} c_1(l)\, \beta^{2\,n+1}(k - l) = \sum_{l \in \mathbb{Z}} c_0(l) \sum_{m \in \mathbb{Z}} u_2^n(m)\, \langle \beta^n(\cdot - l), \beta^n(\cdot - 2\,k - m) \rangle.$$

By rearranging the terms, using again symmetry and n-fold convolution, we finally get

$$c_1(k) = \frac{1}{2} \left(\left(b^{2\,n+1} \right)^{-1} * \downarrow_2 \left(u_2^n * b^{2\,n+1} * c_0 \right) \right)(k), \tag{8}$$

where $b(k) = \beta(x)|_{x=k}$ and where $\downarrow_2 (\cdot)$ downsamples its argument by two. Note that, should several pyramid levels be computed at once, additional computational savings are considered in.[17]

3.4. Signal Extension

Section 3.3 has shown that a very limited repertoire of operations is sufficient to perform either expansion or reduction of data by a factor two in a least-squares sense: upsampling, downsampling, and discrete convolution. Since u_2^n and b^n are FIR filters, the only apparent difficulty comes from the IIR filter b^{-n} of Equations 7 and 8. We show now how to adapt a previously published algorithm[18] to suit our present needs.

Practical signals are always finite. Thus, data extrapolation is necessary to calculate the infinite sum of a convolution. A popular extrapolation method is known as zero padding; though the simplest of all, it compromises ease of implementation for quality because an arbitrary value (that is, zero) leaks into the signal for as far as the convolution filter extends, up to infinity in the b^{-n} case. Moreover, a zero-padding method creates an artificial edge between the arbitrary value zero and the endpoint values of the data.

Another solution—signal double-mirroring around the endpoints—can be found in the literature,[18] where observed instead of arbitrary data values are used to perform the extrapolation; moreover, no artificial edge is created at the data endpoints. A serious drawback of this otherwise appealing extrapolation method is as follows: let a signal $s(x)$, $x \in [x_0, x_1]$ be extended as

$$\begin{cases} s(x + x_0) = s(x_0 - x) \\ s(x + x_1) = s(x_1 - x). \end{cases}$$

Then, it is not difficult to show that $s(x) = s(x + 2\,(x_1 - x_0))$, which implies that the signal is periodic. For discrete data s_k, $k \in [1 \ldots K]$, the endpoints proposed in[18] are $x_0 = 1$ and $x_1 = K$, in which case the resulting period $2\,K - 2$ is inadequate for a dyadic pyramid: among other reasons, if K is even, then the period of a reduced version of the data will be $K - 1 = (2\,K - 2)/2$, which is odd and cannot be further reduced.

Here, we propose instead to extend the data by the simple periodization $s(x) = s(x + K)$. No arbitrary value is introduced, but an edge may be present at the data endpoints. Along with a shift of the range of validity from $[1 \ldots K]$ to $[0 \ldots K - 1]$, this modification requires that Equations (2.6) and (2.5) in[18] be rewritten as follows:

$$\begin{cases} y^+(0) &= \frac{1}{1 - z_i^K} \left(x(0) + \sum_{l=1}^{K-1} z_i^l\, x(K - l) \right) & k = 0 \\ y^+(k) &= x(k) + z_i\, y^+(k - 1) & k \in [1 \ldots K - 1] \\ y(K - 1) &= \frac{-z_i}{1 - z_i^K} \left(y^+(K - 1) + z_i \sum_{l=0}^{K-2} z_i^l\, y^+(l) \right) & k = K - 1 \\ y(k) &= z_i\, (y(k + 1) - y^+(k)) & k \in [0 \ldots K - 2]. \end{cases}$$

These expressions are exact. Of course, since the decay of the sequence b^{-n} is exponential, it is also possible to truncate the sum early, as soon as the terms z_i^l take a negligible value. Alternatively, one can perform the convolution $y = b^{-n} * x$ in Fourier, by using the substitution $z = e^{j\,\omega}$ in Equation 7.

4. EXPERIMENTAL RESULTS

We have acquired data according to Section 2. So as to conform to common practice in the field, we have considered one slice at a time (2D processing), even though it may be more beneficial to consider the whole volume simultaneously (3D processing). Each slice has size 128×128.

The first task is the generation of the multiresolution pyramid for the time-series of PET data. For this paper, we achieved the deepest possible depth and produced an eight-level pyramid with sizes 128×128, 64×64, 32×32, 16×16, 8×8, 4×4, 2×2, 1×1. Figure 3 shows the four finest levels of the pyramid at time t_i, where the first row gives the amount of effective data, which results in varying sizes of fixed resolution, and where the second row lets the resolution vary while fixing the size. Note that these two representations are exactly equivalent in terms of intrinsic data amount. These data are to be compared to those of Figure 4.

Then, starting with the 1×1 level, we consider the data time-series and fit the compartmental model. This produces five 1×1 parametric maps, one for each component of $\boldsymbol{\mu}$. The curve showing the average density of receptors of Figure 2 corresponds to the 1×1 map for $B'_{\max}$. This map is expanded to yield the initial condition for the next level, as explained in Sections 3.1 and 3.3.1. This procedure is iterated until we reach the finest level. We show in Figure 5 the concentration $B'_{\max}$ that represents the best map resulting from optimizing at some level, and that gives at the same time the initial condition for the next level. The spatial resolution increases from left to right as $\frac{1}{8}$, $\frac{1}{4}$, $\frac{1}{2}$, and 1 (full size) in the last column.

All images have the same spatial extent; only their resolution differ. We observe that the result at half resolution is very close to the result at full resolution; from a clinical perspective, the difference in quality is immaterial, while the gain in speed is certainly significant (four times faster).

5. CONCLUSION

Given a time series of noisy positron emission tomography data of the head and given a time series of blood samples, we have proposed a solution to the problem of fitting five parameters of a compartmental model for ligand-receptor interactions. We have used a spline least-squares pyramid to trade noise for spatial resolution; at each pyramid level, we have optimized the fit thanks to a Marquardt-Levenberg algorithm. By propagating the fit from a coarse pyramid level to the next finer one, we were able to gain in robustness and speed; this scheme benefits strongly from the high quality of the pyramid and from the superlinear performance of the optimizer near the optimum. We have performed an experiment with real data that shows that an additional substantial gain in speed can be achieved by stopping the optimization at half resolution, without prejudice to the clinical applicability.

ACKNOWLEDGMENTS

This work was supported in part by the Swiss National Science Foundation (Nb 32-52894.97, Nb 31-64020.00) and by the Geneva Academic Society.

REFERENCES

1. F. Hucho, *Neurotransmitter Receptors*, Elsevier Science, 1994.
2. N. Lassen, P. Bartenstein, A. Lammertsma, M. Prevett, D. Turton, S. Luthra, S. Osman, P. Bloomfield, T. Jones, P. Patsalos, M. O'Connell, J. Duncan, and J. Andersen, "Benzodiazepine receptor quantification in vivo in humans using [(11)C]flumazenil and PET: Application of the steady-state principle," *Journal of Cerebral Blood Flow and Metabolism* **15**, pp. 152–165, January 1995.
3. J. Delforge, L. Spelle, B. Bendriem, Y. Samson, and A. Syrota, "Parametric images of benzodiazepine receptor concentration using a partial-saturation injection," *Journal of Cerebral Blood Flow and Metabolism* **17**, pp. 343–355, March 1997.
4. P. Millet, C. Graf, A. Buck, B. Walder, G. Westera, C. Broggini, M. Arigoni, D. Slosman, C. Bouras, and V. Ibanez, "Similarity and robustness of PET and SPECT binding parameters for benzodiazepine receptors," *Journal of Cerebral Blood Flow Metabolism* **20**, pp. 1587–1603, November 2000.
5. A. Syrota, G. Paillotin, J. Davy, and M. Aumont, "Kinetics of in vivo binding of antagonist to muscarinic cholinergic receptor in the human heart studied by positron emission tomography," *Life Science* **35**, pp. 937–945, August 27 1984.

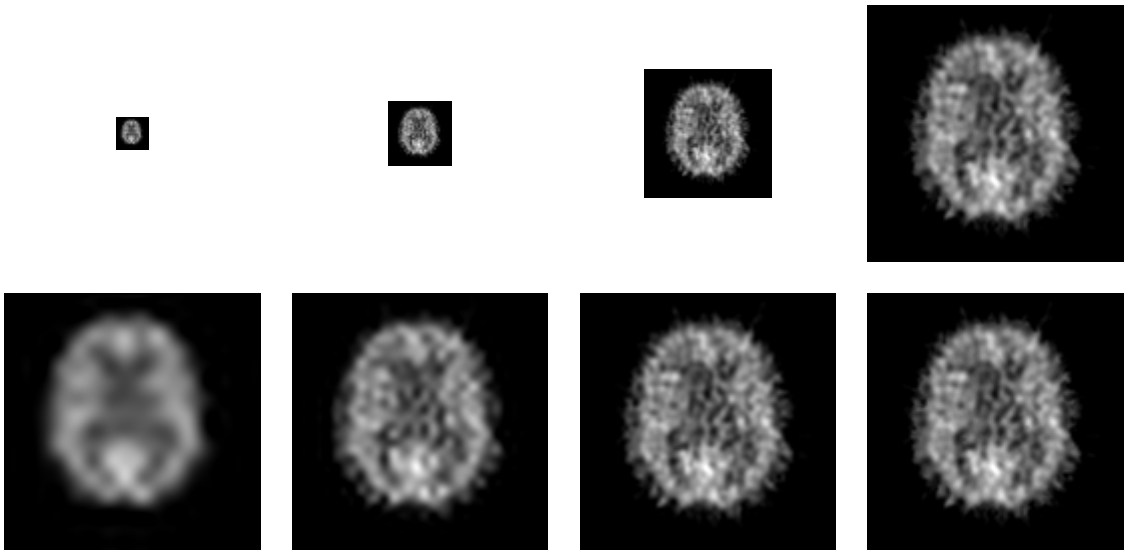

Figure 3. Multiresolution pyramid for the PET data (single slice at time t_i)

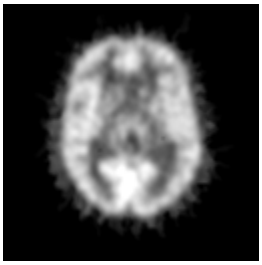

Figure 4. Pet slice averaged over all recorded t_i

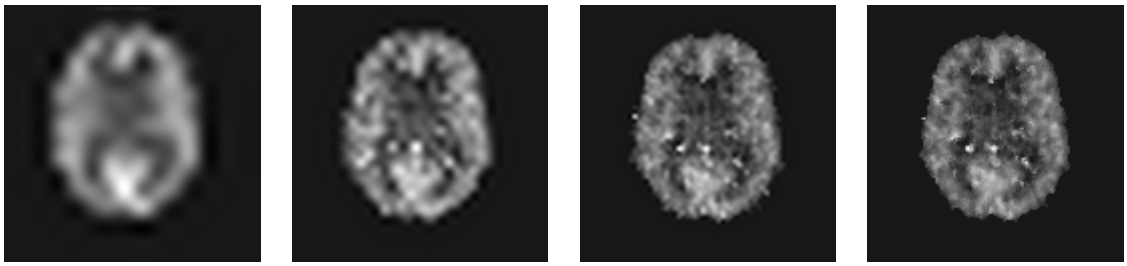

Figure 5. Final steps of the optimization process ($B'_{\max}$)

6. J. Delforge, D. Le Guludec, A. Syrota, B. Bendriem, C. Crouzel, M. Slama, and P. Merlet, "Quantification of myocardial muscarinic receptors with PET in humans," *Journal of Nuclear Medicine* **34**, pp. 981–991, June 1993.
7. D. Comar, M. Mazière, J. Godot, G. Berger, F. Soussaline, C. Menini, G. Arfel, and R. Naquet, "Visualization of 11C-flunitrazepam displacement in the brain of the live baboon," *Nature* **280**, pp. 329–331, July 1979.
8. M. Mazière, C. Prenant, J. Sastre, M. Crouzel, D. Comar, P. Hantraye, M. Kaijima, B. Guibert, and R. Naquet, "11C-RO 15-1788 et 11C-flunitrazepam, deux coordinats pour l'étude par tomographie par positrons des sites de liaison des benzodiazépines," *Comptes-rendus de l'académie des sciences de Paris* **296**, pp. 871–876, 1983.
9. A. Persson, E. Ehrin, L. Eriksson, L. Farde, C. Hedstrom, J. Litton, P. Mindus, and G. Sedvall, "Imaging of [(11)C]-labelled RO 15-1788 binding to benzodiazepine receptors in the human brain by positron emission tomography," *Journal of Psychiatric Research* **19**(4), pp. 609–622, 1985.
10. Y. Samson, P. Hantraye, J. Baron, F. Soussaline, D. Comar, and M. Mazière, "Kinetics and displacement of [(11)C]RO 15-1788, a benzodiazepine antagonist, studied in human brain in vivo by positron tomography," *European Journal of Pharmacology* **110**, pp. 247–251, April 1985.
11. D. Marquardt, "An algorithm for least-squares estimation of non-linear parameters," *Journal of the Society for Industrial and Applied Mathematics* **11**(2), pp. 431–441, 1963.
12. E. Landaw and J. r. DiStefano, "Multiexponential, multicompartmental, and noncompartmental modeling. II. Data analysis and statistical considerations," *American Journal of Physiology* **246**, pp. R665–R677, May 1984.
13. B. Mazoyer, R. Huesman, T. Budinger, and B. Knittel, "Dynamic PET data analysis," *Journal of Computer Assisted Tomography* **10**, pp. 645–653, 1986.
14. J. Beck and K. Arnold, *Parameter Estimation in Engineering and Science*, John Wiley & Sons, 1977.
15. E. Carson, *Parameter Estimation on PET Positron Emission Tomography and Autoradiography, Principle and Application for the Brain and Heart*, R. Press, 1977.
16. J. Delforge, A. Syrota, and B. Mazoyer, "Experimental design optimization: Theory and application to estimation of receptor model parameters using dynamic positron emission tomography," *Physics in Medicine and Biology* **34**, pp. 419–435, April 1989.
17. M. Unser, A. Aldroubi, and M. Eden, "The L_2 polynomial spline pyramid," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **15**, pp. 364–379, April 1993.
18. M. Unser, A. Aldroubi, and M. Eden, "B-spline signal processing: Part II—Efficient design and applications," *IEEE Transactions on Signal Processing* **41**, pp. 834–848, February 1993.