Atlas-Based Segmentation For HRRT Brain PET

Philip Novosad

Master of Science

Medical Physics Unit

McGill University

Montreal, Quebec

August 2014

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree Master of Science.

ACKNOWLEDGEMENTS

I would like to thank my supervisor Dr. Andrew Reader for his enthusiastic support during the past year. I would also like to extend my gratitude to my labmates at the BIC, and Dr. Alain Dagher, Dr. Sylvia Cox, and Kevin Larcher for generously providing [^{11}C]raclopride data, and to Dr. Jean-Paul Soucy for his assistance with image segmentations. Finally, I'd like to thank my family for always being there.

This work was partially supported by the CREATE Medical Physics Research Training Network grant of the Natural Sciences and Engineering Research Council (grant number: 432290) and also partially supported by NSERC discovery grant (NSERC RGPIN 387067-10) and the Canada Foundation for Innovation (CFI project no. 209302).

ABSTRACT

Positron emission tomography (PET) images often require segmentation for analysis. In practice, segmentation is performed manually by an expert on the PET images, or regions of interest (ROIs) are delineated on co-registered anatomical images. However, manual methods are operator dependent and time-consuming, and registration between anatomical and functional images of the same subject can be inaccurate. To overcome these limitations, we propose an automated segmentation scheme based on a PET-to-PET image registration. Image segmentation is achieved by registration to a labeled three dimensional (3D) radio-tracer specific brain template; dynamic PET data can be segmented by registering the sum image of all dynamic PET frames to the template. Crucially, the template image is a mean shape and mean intensity brain image formed from multiple subjects, making it an ideal image for registration and for labeling. Image segmentations can then be used to extract regional time-activity curves (TACs) to enable robust and fully automated regional binding potential (BP) estimation using a reference tissue model. We validate our approach with a comprehensive simulated dynamic [^{11}C]raclopride phantom study, and assess its performance on real dynamic PET [^{11}C]raclopride data from the high resolution research tomograph (HRRT).

Using simulated data, we found that our atlas-based segmentation achieved a mean Dice similarity coefficient greater than or equal to 0.881 over all regions of interest, and that binding potential estimates using the segmented regions differed

from the binding potential estimates using the true regions by a maximum of ${}^{\%}Abs.Error = 3.423$, comparing favourably with the "gold standard" conventional segmentation method of manual delineation. Applied to real data, the atlas segmentations agreed with manual delineations in terms of the linear correlation coefficient on binding potential estimates ($r = 0.979$ when considering all regions of interests). These findings suggest that atlas-based segmentation for fully automated kinetic analysis of dynamic [^{11}C]raclopride images performs at a level which at least matches, and in many cases outperforms, manual segmentation.

ABRÉGÉ

Les images de la tomographie par émission de positrons (TEP) ont souvent besoin de segmentation pour l'analyse. En pratique, la segmentation est effectuée manuellement par un expert sur les images, où des régions d'intérêt sont délimitées sur les images anatomiques co-enregistrés. Cependant, les méthodes manuelles dépendent de l'opérateur et prennent beaucoup de temps, et le recalage entre les images anatomique et fonctionnelles du même sujet peu être inexact. Pour surmonter ces limitations, nous proposons un système de segmentation automatique base sur un enregistrement TEP-à-TEP optimisé. La segmentation des images est réalisée par alignement avec un atlas spécifiques au traceur-radio du cerveau. Plus important encore, l'atlas est une d'image de moyenne forme et moyenne intensité formé à partir de plusieurs images, ce qui en fait une image optimale pour l'enregistrement et l'étiquetage manuel. Les segmentations des images peuvent être utilisé pour extraire des courbes régionales temps-activité pour permettre l'estimation automatique de BP régionale en utilisant un modèle de tissu de référence. Nous validons notre approche avec des images simules, et evaluons sa performance sur de vraies images TEP de [^{11}C]raclopride du tomographe de recherche à haute résolution.

En utilisant des images simulées, nous avons constaté que notre segmentation basée sur l'atlas était exact (moyen coefficient Dice de similarité supérieur ou égal à 0,881 pour toutes les régions d'intérêt) et que estimation BP en utilisant les régions segmentées différer de l'estimation de BP en utilisant les véritables

régions par un maximum de $^{\%}Abs.Error = 3.423$, se compare favorablement avec la référence de la méthode de segmentation classique de délimitation manuelle. Appliquée à des images réelles, la segmentation est d'accord avec les délimitations manuelles en termes de coefficient de corrélation linéaire sur des estimations de BP ($r = 0.979$ lorsque l'on considère toutes les régions d'intérêt). Ces résultats suggèrent que la segmentation d'images par atlas fonctionnelles effectue à un niveau qui correspond au moins, et dans de nombreux cas surpasse, segmentation manuelle.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS ii

ABSTRACT iii

ABRÉGÉ v

LIST OF TABLES ix

LIST OF FIGURES x

1 Positron Emission Tomography 1

1.1 Positron emission and annihilation 1
1.2 Photon-pair detection 3
1.3 Data Storage and Sinograms 6
1.4 From Sinograms to Images 7
1.4.1 Filtered Back Projection 8
1.4.2 MLEM: Derivation Using Gradient Ascent 10
1.4.3 MLEM: Derivation From General EM Framework 13
1.5 MLEM: Limitations and Solutions 15
1.6 Data Corrections 19
1.6.1 Attenutation Correction 19
1.6.2 Random Coincidences 20
1.6.3 Scatter Coincidences 21
1.6.4 Normalization 22
1.6.5 Data Corrections: Implementation 23

2 Kinetic Analysis and Conventional Segmentation 25

2.1 Compartmental Models: Two-tissue Model 27
2.2 Compartmental Models: One-tissue Model 29
2.3 Reference Tissue Models 30
2.4 Implementation of SRTM: Basis Function Method 33
2.5 Segmentation 34

2.5.1 k-means . . . 34
2.5.2 Fuzzy c-means . . . 36
2.5.3 Gaussian Mixture Model . . . 37

3 Image Registration . . . 40

3.1 Affine Registration . . . 40
3.2 Demons Algorithm . . . 42
3.3 Log-Domain Diffeomorphic Demons Algorithm . . . 46

4 Atlas-Based Segmentation of HRRT Brain PET . . . 49

4.1 Introduction . . . 49
4.2 Methods . . . 51
4.2.1 Atlas-Based Segmentation and Kinetic Analysis . . . 51
4.2.2 Simulations . . . 53
4.2.3 Real Images . . . 56
4.2.4 Atlas Construction . . . 59
4.2.5 Manual Segmentation . . . 60
4.2.6 Segmentation Quality Metrics . . . 60
4.3 Experiments . . . 62
4.3.1 Simulation Validation . . . 62
4.3.2 Real Data . . . 63
4.4 Results . . . 63
4.4.1 Simulation Validation . . . 63
4.4.2 Real Data . . . 69
4.5 Discussion . . . 72
4.6 Conclusion . . . 75

REFERENCES . . . 77

LIST OF TABLES

Table		page
1–1	Widely used radionuclides for PET imaging	3
4–1	Resolution kernel parameters	56
4–2	Simulated image characteristics	57
4–3	Simulation validation results	68
4–4	Overall simulation validation results	68

LIST OF FIGURES

Figure		page
1–1	Positron annihilation	2
1–2	Scanner geometry	4
1–3	Coincidence events	5
1–4	Sinograms	7
2–1	Dynamic Pet Scan	26
2–2	Two-tissue, three-comparment kinetic model	28
2–3	Two-tissue, two-comparment kinetic model	30
2–4	Reference tissue model	31
3–1	Affine transformations	42
3–2	Example demons transformation	45
3–3	Demons transformations	48
4–1	Simulation model	54
4–2	Simulated images	58
4–3	Atlas construction	61
4–4	Simulated atlas	63
4–5	Simulated image segmentation	64
4–6	Simulation validation results (DSC)	65
4–7	Simulation validation results (errors)	67
4–8	Real atlas	69

4–9 Real subject segmentation 70
4–10 Correlation of BP estimates. 71

CHAPTER 1
Positron Emission Tomography

1.1 Positron emission and annihilation

Positron emission describes the radioactive process by which a proton (p) of an unstable proton-rich nucleus is converted into a neutron (n) while releasing a positron (e^+) and a neutrino (ν):

$$ {}^{1}_{1}p \rightarrow {}^{1}_{0}\,n + {}^{0}_{1}\,e^{+} + \nu \tag{1.1} $$

In general, therefore, the equation for positron decay is:

$$ {}^{A}_{Z}X \rightarrow {}^{A}_{Z-1}\,Y + {}^{0}_{1}\,e^{+} + \nu + Q \tag{1.2} $$

where Q is the kinetic energy that is shared between the positron and neutrino. Positrons are thus emitted with a range of energies, from zero to a maximum energy determined by the difference in mass between the parent atom X and the daughter atom Y. The positron that is ejected following positron decay has a very short lifetime in electron-dense materials, such as biological tissue. It rapidly loses kinetic energy via inelastic interactions with atomic electrons, and once most of the energy is dissipated (typically with a range of a few mm at FWHM [1], depending on the initial kinetic energy of the positron), it combines with an electron (both being approximately at rest) and both annihilate, giving off energy in the form of two photons (the emission of a single photon is strictly prohibited

because of conservation of momentum considerations). The two photons are each of 511 keV energy and are emitted at approximately 180 degrees to each other (in rare cases, three photons may be emitted following positron-electron annihilation).

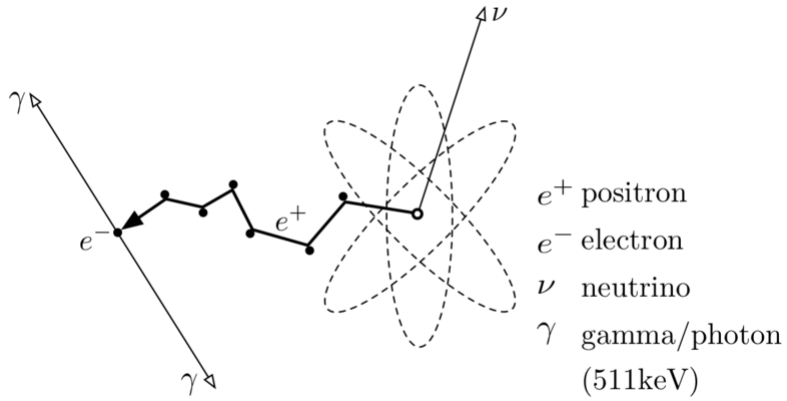


Figure 1–1: Positron annihilation. An unstable proton-rich nucleus decays into a neutron, a positron and a a neutrino. The positron loses kinetic energy rapidly via inelastic iteractions with atomic electrons and annihilates with an electron, giving off energy in the form of two approximately co-linear 511 keV photons.

Radionuclides (atoms with unstable nuclei) that undergo radioactive decay (with sufficiently short half-lives) through positron emission can be used to label molecules with biologically interesting functions. Such a labeled molecule is called a *radiotracer*, since it it is radioactive, and since it's path can be traced through the body (once injected, swallowed or inhaled) by detection of photon-pair emissions. Generally, radionuclides are incorporated either into compounds normally used by the body (i.e. glucose or glucose analogues), or into molecules that bind to receptors (i.e. [^{11}C]raclopride, a D2/D3 dopamine antagonist). In table 1–1, several widely used radionuclides are characterized.

Radionuclide	$T_{1/2}$ (mins)	E_{max} (MeV)	R (mm at FWHM)	Example Radiotracer
^{11}C	20.4	0.96	1.1	[^{11}C]raclopride
^{13}N	10.0	1.19	1.4	Nitrogen-13-ammonia
^{15}O	2.0	1.70	1.5	Oxygen-15-water
^{18}F	109.7	0.64	1.0	Fluorodeoxyglucose (FDG)

Table 1–1: Widely used radionuclides for PET imaging. $T_{1/2}$ is the halflife of the radionuclide. E_{max} is the maximum energy of the positron, R is the range of the positron in water.

1.2 Photon-pair detection

A PET scanner consists of a ring of photon detectors that surround the photon-pair-emitting subject. The goal of the physical PET system is to detect and characterize "lines of response" (LORs), or lines in space connecting pairs of photon-activated detectors. A convenient way of representing these LORs (in 2 dimensions) is to adopt the co-ordinate scheme described in figure 1–2, enabling one to characterize a given LOR with two parameters: the radial offset from the center of the scanner, s, and the angle ϕ. Event detection in PET relies on electronic collimation; a "detection event" is regarded as valid if:

- two photons are detected within a predefined time window, known as the "coincidence window"
- the subsequent line-of-response (LOR) formed between the detected photons is within a valid range
- the energy deposited by both photons is within the selected energy window

Detection events which satisfy the above criteria are known as prompt events (or "prompts"). These criteria are important in order to preferentially detect photon

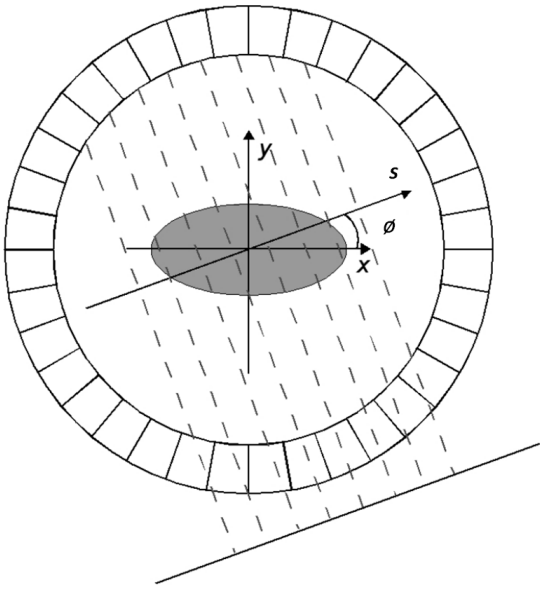


Figure 1–2: Scanner geometry. The "line of response" that connects a pair of activated photon detectors is characterized by the radial offset from the center of the scanner, s, and the angle ϕ.

pairs which originate from a *tractable point* in the subject. However, many prompt events satisfying the above criteria are, in fact, not due to an ideal emission of a photon pair originating from a single positron annihilation event. In other words, many registered prompt events are spatially uncorrelated with the spatial distribution of the tracer. The simplest reasons for this are due to the non-zero positron range and to photon-pair colinearity (photon pairs are not emitted at exactly 180 degrees with respect to each other, creating LORs which do not cross the actual point of annihilation). There are three additional categories of events which lead to unwanted prompts: random, scattered and multiple coincidences. Each of these effects have a degrading effect on the measurement, and need to be

corrected for to produce an image that resembles, as closely as possible, the true radiotracer distribution. The categories are depicted in figure 1–3.

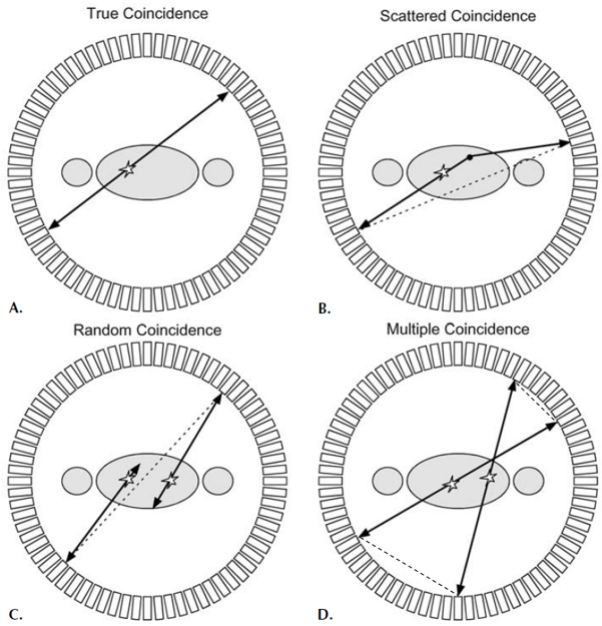


Figure 1–3: Prompt events do not always give useful information about the origin of the detected photon pair, due to scattered, random and multiple coincidence events. The (inaccurate) LOR as seen by the scanner is depicted by a dashed line.

Random coincidences are the result of two unrelated single annihilation photons being registered as a valid prompt. Because these events are produced by photons emitted from spatially unrelated annihilations, they carry no information about the tracer distribution. Scattered coincidences occur when one or both of the annihilation photons, originating from a single annihilation event, undergo inelastic Compton scattering and change direction prior to reaching the detector pair. Although it is true that scattered photons (undergoing a change in direction) have an associated energy loss, many scatter photons cannot be discriminated

due to the limited energy resolution of the ring detector. At high count rates, it is possible that three or more detectors are activated within the allowed time window. In this case, it is ambiguous where the event should be positioned. These events are called multiple coincidences, and are usually discarded.

1.3 Data Storage and Sinograms

As mentioned previously, PET scanners detect and record coincidence events. There are two formats for storing such emission data: list mode and sinograms. Storing emission data in list mode is, as the name implies, a simple list of recorded events. Such events can be characterized by time of detection and detected photon energy [2]. On the other hand, a sinogram is a vector containing indices that each represent the counts detected along a given LOR. We will focus on the case of a two dimensional sinogram for simplicity. Referring back to figure 1–2, we note that a detected event corresponds to a given LOR, which can be parametrized by s and ϕ. A sinogram $m_\phi(s)$ is simply a 2D array populated by pixels corresponding to the number of events detected along a given LOR. From figure 1–2, we can relate the spatial radioactivity distribution $n(l, s)$ to the sinogram count for a given LOR by the follow integral equation, also known as the Radon or x-ray transform:

$$m_\phi(s) = \int n(l, r)dl \tag{1.3}$$

where l is the distance along the LOR. The reason for the name "sinogram" is that if the radioactivity distribution is a simple point source emitter, then the 2D array resembles a sine wave. By way of example, consider a single radioactive point source emitting photon pairs (figures 1–4).

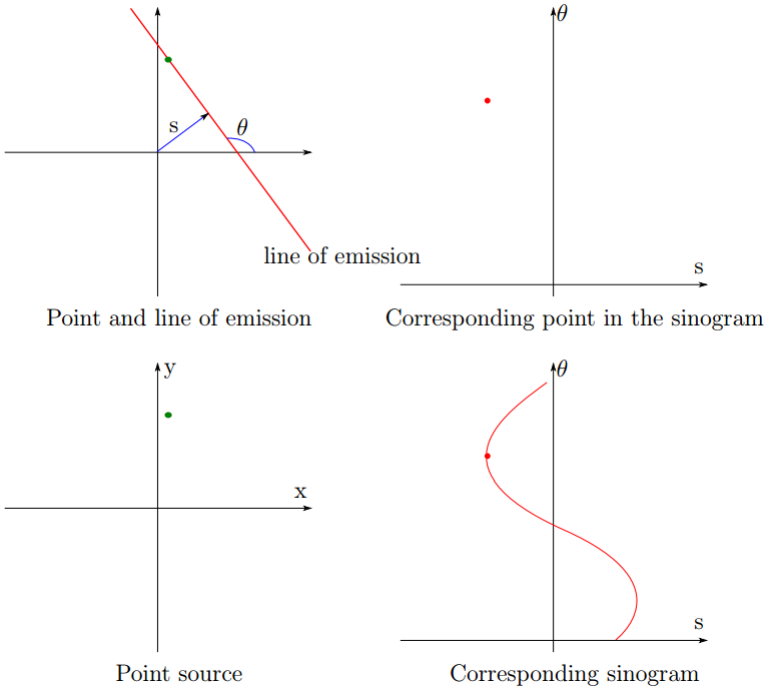


Figure 1–4: Construction of a sinogram. Top left: colinear photon-pair emitted from point source along an LOR characterized by two parameters, s and ϕ. Top right: This LOR can be represented as a single point on a sinogram (value = 1 count). Bottom row: the same point source continues emitting colinear photon pairs isotropically, which creates a sinogram resembling a sine wave.

Equation 1.3 allows us to predict what a given sinogram looks like for a given spatial radioactivity distribution. This knowledge is necessary to tackle the image reconstruction problem, which will now be discussed.

1.4 From Sinograms to Images

The essential problem in PET image reconstruction is to estimate the radioactive distribution n_j (the average number of emissions from each pixel j within the subject) given the noisy measured data m_i (the noisy sinogram data containing

values that represent the number of counts per projection bin i). Several algorithms can be used to reconstruct PET images; filtered back projection (FBP) [3], least squares techniques [4] and the Maximum Likelihood Expectation Maximization (MLEM) algorithm [5] are among the most popular choices. In this section, a brief overview of FBP and MLEM will be given, with a greater emphasis on the latter; the MLEM algorithm is used for the HRRT [6] [7] at the Brain Imaging Center (BIC).

1.4.1 Filtered Back Projection

Rewriting the Radon transform equation (equation 1.3) in terms of Cartesian coordinates x and y, we have

$$m_\phi(s) = \int\int n(x,y)\delta(xcos\phi + ysin\phi - s)dxdy \tag{1.4}$$

The image $n(x,y)$ can then be reconstructed from the sinogram data m using the inverse Radon transform, which is usually performed using the filtered back projection (FBP) algorithm [8]. Note that the inverse Radon transform assumes that there is no attenuating medium, and that detectors are perfect.

A key theorem in analytic image reconstruction is the *central slice theorem*, which says that the one dimensional (1D) Fourier transform of a projection of an image $n(x,y)$ at an angle ϕ is equivalent to a slice through the center of the two dimensional (2D) Fourier transform $N(u,v)$ at the same angle. In other words, we have

$$M_\phi(w) = N(u,v) \tag{1.5}$$

where $M_\phi(w)$ is the 1D FT of the projection $m_\phi(s)$, and $u = wcos\phi$, $v = wsin\phi$. Taking the inverse 2D Fourier transform allows for the recovery of the object:

$$n(x,y) = \int\int N(u,v)e^{j2\pi(ux+vy)}dudv \tag{1.6}$$

and, by exchanging the Cartesian coordinate system in the frequency domain for a polar coordinate system, one can show that

$$n(x,y) = \int_0^\pi \lambda_\phi(xcos\phi + ysin\phi)d\phi \tag{1.7}$$

where

$$\lambda_\phi(s) = \int M_\phi(w)|w|e^{j2\pi ws}dw \tag{1.8}$$

Explicitly, equation 1.8 is a simple ramp filtering operation applied to the 1D Fourier transform of a projection at angle ϕ. $\lambda_\phi(s)$ is the filtered projection, and equation (1.7) is the backprojection operation. The steps of the FBP algorithm are outlined below

for *each projection angle* ϕ **do**

Take the 1D Fourier transform of the parallel projection;

Apply the ramp filter in Fourier space;

Take the inverse 1D Fourier transform;

Backproject;

end

Algorithm 1: The FBP algorithm

Often, the ramp filter $|w|$ is modified to improve the signal-to-noise ratio (SNR) of the resulting image by introducing a cut-off frequency; however, this has the unwanted effect of reducing spatial resolution. Therefore, choosing a suitable filter is akin to finding an optimal trade-off between SNR and spatial resolution [9].

1.4.2 MLEM: Derivation Using Gradient Ascent

While analytic reconstruction methods (e.g. FBP) are fast, however, they are limited by their reliance on the approximation inherent in the line-integral equation (equation 1.4). The advantage of using an iterative reconstruction algorithms (such as MLEM), by contrast, is that they allow for more complex and therefore more realistic models of the acquired PET data [10]. In the following derivation[1] , n_j denotes the mean number of emissions from pixel j, m_i denotes the noisy measured sinogram data in bin i, and q_i denotes the "mean sinogram" that could theoretically be obtained by repeating a particular experiment an infinite number of times. One way to estimate q_i is to make use of equation 1.3; given a discrete image n representing the mean number of emissions per pixel, we can determine q by "ray-tracing". We can discretize the integral operation into a summation, such that:

$$q_i = \sum_{j=1}^{J} a_{ij} n_j \tag{1.9}$$

where a_{ij} can be thought of as the discretized equivalent of the Radon transform equation; a_{ij} represents the probability that an annihilation event taking place at

[1] This derivation is largely based on Dr. Andrew Reader's MDPH607 class notes, Medical Physics Unit, McGill University.

voxel j is detected at sinogram bin i. For any given scan, the actual obtained sinogram data m_i is a random vector distributed according to the Poisson distribution:

$$p(m_i|q_i) = \frac{q_i^{m_i} exp(-q_i)}{m_i!} \tag{1.10}$$

Equation (1.10) gives the probability of obtaining m_i counts in bin i given that the mean counts is q_i. We seek to maximize the Poisson likelihood $O(n;m) = p(m|n)$ of obtaining a noisy sinogram vector m given an image n. We start by writing the Poisson likelihood for the entire noisy measured sinogram:

$$O_{ML}(n;m) = \prod_{i=1}^{I} \frac{(q_i)^{m_i} exp(-q_i)}{m_i!} \tag{1.11}$$

Taking the log of equation (1.11) simplifies the derivation considerably without affecting the value of n that maximizes the likelihood:

$$L_{ML}(n;m) = \sum_{i=1}^{I} m_i ln(q_i) - \sum_{i=1}^{I} q_i = \sum_{i=1} m_i ln(\sum_{b=1}^{J} a_{ib} n_b) - \sum_{i=1}^{I} \sum_{b=1}^{J} a_{ib} n_b \tag{1.12}$$

where L_{ML} is the Poisson log-likelihood. Because the log-likelihood function is concave [5], a gradient ascent approach can be used. In other words, given that we evaluate the log-likelihood equation with some image n^k, we can move closer to the optimal value n^{ML} by independently updating the value of each image pixel n_j as follows:

$$n_j^{k+1} = n_j^k + \tau \frac{\partial L_{ML}}{\partial n_j} \tag{1.13}$$

where τ is a step size to be determined and k is the iteration number. Taking the partial derivative of equation 1.12 with respect to an image pixel n_j, and

substituting it into equation 1.13, we obtain:

$$n_j^{k+1} = n_j^k + \tau \left(\sum_{i=1}^{I} \frac{a_{ij}^k m_i}{\sum_{b=1}^{J} a_{ib} n_b^k} - \sum_{i=1}^{I} a_{ij} \right) \tag{1.14}$$

And choosing the step size[2] to be equal to the current image value (n_j) normalized by the sensitivity of the scanner at that point ($\sum_{i=1}^{I} a_{ij}$), we obtain the MLEM algorithm:

$$n_j^{k+1} = \frac{n_j^k}{\sum_{i=1}^{I} a_{ij}} \sum_{i=1}^{I} a_{ij} \frac{m_i}{\sum_{b=1}^{J} a_{ib} n_b^k} \tag{1.15}$$

or, in compact matrix notation,

$$\mathbf{n^{k+1}} = \frac{\mathbf{n^k}}{\mathbf{A^T 1}} \mathbf{A^T} \left(\frac{\mathbf{m}}{\mathbf{A n^k}} \right) \tag{1.16}$$

where $\mathbf{n}$ (size $J \times 1$) and $\mathbf{m}$ (size $I \times 1$) are column vectors containing the image values and sinogram counts, respectively, $\mathbf{A}$ is an $I \times J$ matrix with elements a_{ij}, and $\mathbf{1}$ is simply a column vector of size $I \times 1$ with each entry equal to 1. Note that the division is carried out element by element. Convergence towards an image estimate n is therefore accomplished iteratively: explicitly, a set of correction values between the measured sinogram data and the forward projected image estimate is formed. The correction values are backprojected into image space via the transpose of the system matrix, and applied multiplicatively to the image estimate, voxel-by-voxel. The corrected image estimate is normalized by the

[2] An alternative and elegant derivation of MLEM making use of hidden complete data avoids this potential ambiguity, but the derivation is more involved. It is presented in the next section.

sensitivity image (again, voxel-by-voxel), yielding the image estimate for the next iteration.

1.4.3 MLEM: Derivation From General EM Framework

Given a statistical model consisting of a set x of observed data, a vector of unknown parameters θ, and a likelihood function $L(\theta; x) = p(x|\theta)$, the maximum likelihood estimate of the unknown parameters θ can be found by likelihood of the observed data, $L(\theta; x)$. For some models, a maximum likelihood estimation of the parameters θ can be found as an explicit function of the observed data. For many other models, however, no explicit and closed-form solution is available, and optimization techniques must be used.

The goal of the MLEM algorithm is to facilitate maximum likelihood estimations by introducing a *hidden* random vector z. The intuition behind the algorithm is that we can calculate the *expected* value of the log-likelihood function with respect to the conditional distribution of z given x and under the current estimate of the parameters θ^k:

$$Q(\theta|\theta^k) = \mathbb{E}_{z|x,\theta^k}[\log\ L(\theta; x, z)] \tag{1.17}$$

Then we can maximize (more easily, often analytically) the *expected* log likelihood function with respect to the unknown parameters θ:

$$\theta^{k+1} = \arg\max_{\theta} Q(\theta|\theta^k) \tag{1.18}$$

and repeat until convergence. It can be shown that by increasing $Q(\theta|\theta_k)$, we also increase the likelihood of the observed data [11].

Applied to PET imaging, we have a set of observed sinogram data m, a set of unknown parameters n to estimate, and a statistical model relating the two (Poisson distribution). As we have seen, however, maximizing the log-likelihood function is not trivial, but can be made more tractable by introducting a complete data random hidden vector z_{ij} which represents the number of sinogram counts in bin i originating from emissions in pixel j, and we have $\bar{z}_{ij} = a_{ij}n_j$ and $\sum_{j=1}^{J} z_{ij} = m_i$ by definition. Then we can write down the Poisson log-likelihood, $\log p(z|\bar{z}) = \log L(n; z)$:

$$\log L(n; z) = \log \prod_{j=1}^{J} \prod_{i=1}^{I} \frac{exp(-a_{ij}n_j)a_{ij}n_j^{z_{ij}}}{z_{ij}!} \tag{1.19}$$

and then the *expected* value of the log-likelihood with respect to the conditional distribution of z:

$$\mathbb{E}_{z|m,n^k}[\log\ L(n; z)] = \mathbb{E}_{z,n^k} \sum_{b=1}^{J} \sum_{i=1}^{I} \left(z_{ib} ln(a_{ib}n_b) - a_{ib}n_b - ln(z_{ib}!) \right) \tag{1.20}$$

And because of the linearity of expectation, we have:

$$Q(n|n^k) = \sum_{b=1}^{J} \sum_{i=1}^{I} \left(\mathbb{E}[z_{ib}|m, n^k] ln(a_{ib}n_b) - a_{ib}n_b - ln(\mathbb{E}[z_{ib}|m, n^k]!) \right) \tag{1.21}$$

For a given sinogram bin i, the hidden complete data z_{ij} is a Poisson random vector conditioned on it's sum, m_i, and as such forms a multinomial distribution[3] ,

[3] See appendix for proof.

with expectation value:

$$\mathbb{E}[z_{ij}|m, n^k] = NP = m_i \frac{a_{ij} n_j^k}{\sum_{b=1}^{J} a_{ib} n_b^k} \tag{1.22}$$

where NP is the mean of the binomial distribution. Substituting equation (1.22) into equation (1.21), we have:

$$Q(n|n^k) = \sum_{b=1}^{J} \sum_{i=1}^{I} \left(m_i \frac{a_{ib} n_b^k}{\sum_{c=1}^{J} a_{ic} n_c^k} ln(a_{ib} n_b) - a_{ib} n_b - C \right) \tag{1.23}$$

where C is a constant and can be disregarded. Note that this equation can be easily maximized with respect to a single pixel value n_j. Taking the partial derivative with respect to n_j and setting to zero, we obtain the MLEM algorithm:

$$n_j^{k+1} = \arg\max_n Q(n|n^k) = \frac{n_j^k}{\sum_{i=1}^{I} a_{ij}} \sum_{i=1}^{I} a_{ij} \frac{m_i}{\sum_{b=1}^{J} a_{ib} n_b^k} \tag{1.24}$$

1.5 MLEM: Limitations and Solutions

Despite the advantages of the MLEM algorithm compared to FBP, the algorithm can be slow, partly due to its use the entire measured data set m at each iteration. One solution is the so-called "Ordered Subsets Expectation Maximization" (OSEM) algorithm, which addresses the problem by using only a subset of the measured data when computing an image update:

$$n_j^{k,l+1} = \frac{n_j^{k,l}}{\sum_{i \in S_l} a_{ij}} \sum_{i \in S_l} \frac{m_i}{q_i^{k,l}} \tag{1.25}$$

where the backprojection sums only over a subset S_l ($l = 1...L$) of the measured sinogram data, and the image is updated each subiteration (i.e. L subiterations for each iteration k). There are a variety of approaches for dividing the sinogram data

into subsets. Most approaches use non-overlapping subsets, commonly dividing the sinogram data into balanced subsets containing data from different azimuthal angles. With regard to subset ordering, it is recommended that substantial new information is introduced into the image estimate as fast as possible [12]; for instance, the order can be chosen such that there is a maximum separation angle between successive subsets. Because the computational cost of MLEM is proportional to the number of sinogram projections, one iteration of OSEM will have require roughly the same as one iteration of MLEM. However, a single iteration of OSEM consists of a number of subiterations, each of which converges similarly to one MLEM iteration. OSEM therefore accelerates convergence by a factor proportional to the number of chosen subsets [12]. Another approach for accelerating the MLEM algorithm is to use an over-relaxation parameter, which effectively increases the influence of the correction image applied to the image estimate at each iteration. Re-writing the MLEM algorithm (equation (1.24)) in the following form

$$n_j^{k+1} = n_j^k + (c_j^k - 1)n_j^k \tag{1.26}$$

shows that the MLEM algorithm can be interpreted as applying an *additive*, rather than multiplicative, correction to the image estimate at each iteration. The influence of the additive correction term can be amplified using a scalar parameter λ:

$$n_j^{k+1} = n_j^k + \lambda(c_j^k - 1)n_j^k \tag{1.27}$$

where a choice of $\lambda > 1$ will accelerate the iterative process.

Maximum likelihood estimators are advantageous in the sense that they offer unbiased, minimum variance image estimates. Nonetheless, due to the inherent noise in photon counting, maximum likelihood estimate images are still almost always unacceptably noisy: as the ML estimate is approached with increasing iterations, higher frequency definition is introduced in the image at the cost of a lack of low frequency definition. In practice, therefore, image reconstructions are often terminated prior to reaching the ML solution [13], resulting in an image with less variance but increased bias. Alternatively, the algorithm is permitted to run more iterations, and the resulting image is smoothed [14].

Another approach to image reconstruction regularization are the so-called "Bayesian methods." These methods seek to enforce conditions on the image estimate during the reconstruction process by the addition of *a priori* information in a Bayesian formulation. Conventionally, this is done within a maximum *a posteriori* (MAP) framework [15], where the goal of optimization is to maximize the *posterior* probability of the image estimate n given the measured data m. Using Baye's rule, we have

$$p(n|m) = \frac{p(m|n)p(n)}{p(m)} \tag{1.28}$$

where $p(n|m)$ is the posterior probability, $p(m|n)$ is the likelihood and $p(n)$ is the prior probability on the image estimate. Taking the logarithm and dropping the constants, image reconstruction becomes the following optimization problem

$$n = \arg\max_n \{\log p(m|n) + \log p(n)\} \tag{1.29}$$

The function of the prior $p(n)$ is to regurarize the image estimate. Conventionally, the prior is assumed to be in the form of a Gibb's distribution [16]:

$$p(n) = \frac{1}{Z} \exp^{-\beta U(n)} \tag{1.30}$$

where $U(n)$ is the Gibbs energy function. A number of energy functions have been proposed in the literature; most commonly, the energy function attempts to produce local smoothing while not blurring boundaries in the image [17] [16]. Recently, information-theoretic energy functions have been proposed to encourage correspondence between the reconstructed image and co-registered MR anatomical images [18]. With the Gibb's prior defined, the reconstruction problem described in equation (1.29) can be rewritten as

$$n = \arg\max_{n} \{\log p(m|n) - \beta U(n)\} \tag{1.31}$$

The parameter β controls the degree of regularisation, and we can note that as β approaches zero, optimization of equation (1.31) reduces to typical MLEM reconstruction. Because of the addition of the regulariser term, optimization of equation (1.31) is more involved, and the reader is referred to [19] for more details on this topic. Mainstream clinical acceptance of Bayesian image reconstructions is challenged by the added complexity in implementation, and due to the challenge of choosing an appropriate β parameter for controlling the influence of the regularising term.

1.6 Data Corrections

As discussed previously, reconstruction must take into account that many sinogram counts are due to randoms and scatters, and therefore are not representative of the underlying radioactivity distribution. In addition, it is important to correct for attenuation and normalization effects. The following section will offer a brief description of these correction factors, and how they are normally implemented into analytic and iterative reconstruction approaches.

1.6.1 Attenutation Correction

Referring back to figure 1–2, and assuming that $I_o(s,\phi)$ is the number of coincidence events detected along a given LOR in the absence of attenuation, then the number of events detected on that line in the presence of attenutation, $I_a(s,\phi)$, can be expressed as:

$$I_a(s,\phi) = I_o(s,\phi) \times e^{-\int \mu(l,s)dl} \tag{1.32}$$

where l is the distance along the LOR, and μ is the attenuation coefficient, and we now define $\frac{I_a(s,\phi)}{I_o(s,\phi)}$ as the "attenuation correction factor" (ACF) for a given LOR or projection bin. This equation is the basis for attenuation correction methods. In practice, attenuation correction factors are physically measured or calculated.

As equation (1.32) shows, the number of detected events along a given LOR in the presence of attenuation does not depend on the locations along the LOR where the actual annihilations occur. Exploiting this convenience, a radioactive source placed outside the body can be used to measure attenutation factors [20]. This method places a radioactive source *outside* the object, then performs a transmission scan. Secondly, the object is removed, and a black scan is performed.

The ACFs can then be calculated as the ratio of the counts of the blank scan to the transmission scan. This method is hampered by significant noise levels since it can be difficult to obtain enough counts per projection bin.

Alternatively, attenuation factors can be calculated if the shape of the object and the attenutations coefficients are known; the ACFs can be directly computed from equation (1.32). With the advent of PET/CT scanners, ACFs can be directly computed from computed tomography (CT) images since attenuation coefficients can be directly determined from the CT image. As a result, these approaches have the potential to yield more accurate ACFs [21].

1.6.2 Random Coincidences

Recall that in a random coincidence event the two deteted annihilation photons come from different decay sites, and therefore contribute no useful information for the image reconstruction. To correct for the influence of random coincidences, two main approaches are usually considered. The first is to estimate the random count rate from the 'singles' count rate using:

$$N_r = 2\tau N_1 N_2 \tag{1.33}$$

where N_r is the count rate of the random coincidences, τ is the coincidence timing window (the amount of time below which the difference in detection times between two detected photons are considered as a single coincidence event) and N_1 and N_2 are the count rates of single events in each of two detectors forming a single LOR. This method has the advantage that both N_1 and N_2 are both large in comparison to N_r, so that the statistical quality of the random count rate N_r is

good. However, the method can be biased in the case that the coincidence timing window is not know accurately for each detector.

The second method based on direct measurement is the delayed coincidence method. In this method, the logic pulse from one of two detectors forming a LOR is delayed such that the detector pair cannot produce true coincidences; therefore, any measured coindences can only be caused by random events (in practice, a parallel coincidence circuit is implemented in the electronics to enable measurement of random events during the scan). While free from systematic bias, this method is hindered by poor statistical quality.

1.6.3 Scatter Coincidences

Many schemes have been proposed for scatter correction, including convolution-subtraction [22], direct measurement methods [23], dual energy window methods [24], simulation-based algorithms [25] [26] and "Gaussian fit" techniques [27]. The latter two, being the most widely used, are briefly considered here.

This Gaussian fit approach to scatter correction is based on the insight that (i) events in the sinogram outside of the source object are due to scatter, and (ii) the scatter distribution is a low frequency function that is insensitive to the activity distribution. The Gaussian fit method estimates the scatter component of each sinogram projection by fitting the activity outside of the source object with a Gaussian distribution. While this method is fast and gives a smooth scatter distribution, it may fail on whole body scans where the scatter tails are short (since the body occupies a large portion of the field of view). In addition, the

accuracy of the scatter component estimate is highly dependent on the parameters chosen for the Gaussian function fit.

Simulation-based algorithms use a reconstructed image and attenuation map, and then simulates the scatter contribution based on the reconstructed images. In practice, the simulation is done using a simple single-scatter model or a Monte-Carlo simulation. These methods are accurate, but computationally expensive.

1.6.4 Normalization

Non-uniformities in detector efficiency, electrons and geometrical variations in different LORs can cause different sensitivies; "normalization" is the process of correcting such non-uniformities. Normalization coefficents for each LOR can be estimated most easily by direct measurement [28]. In this approach, a source of uniform radioactivity is scanned, and normalization coefficients are then proportional to the inverse of the events recorded in any given LOR. This accuracy of this method is mainly hindered by the requirement that enough counts be collected per LOR, so as to attain a sufficient statistical quality on the estimate.

An attractive alternative to the direct measurement approach is the so called "component-based" method [29], which models the normalization coefficient for each LOR as the product of an intrinsic crystal efficiency and a set of geometric factors. Because the normalization coefficients for each LOR are not independent (since a single crystal efficiency is a factor of many normaliztion coefficients) the number of unknowns is reduced from the number of LORs to the number of crystals, and so the statistical noise is largely reduced.

1.6.5 Data Corrections: Implementation

Unlike iterative reconstruction methods, analytic methods require that the raw sinogram data be corrected prior to reconstruction. Firstly, scatter and random sinograms are subtracted from the raw measured sinogram. Secondly, the attenuation correction factors are applied to the corrected sinogram. Lastly, the normalization correction factors (NCFs) are applied. The resulting sinogram (equation 1.34) is then suitable for analytic reconstruction.

$$m_{corrected} = NCF \times ACF \times (m_{raw} - r - s) \tag{1.34}$$

Iterative reconstruction algorithms (such as MLEM) have the advantage that corrective effects can be incorporated directly into the reconstruction algorithm. This avoids possible negative sinogram values, and retains the Poisson nature of the raw acquired data (which is required for the Poisson statistical model in the MLEM algorithm to be correct). The corrections can be implemented into the MLEM framework by modeling the mean sinogram, as $\mathbf{q} = \mathbf{An} + \mathbf{r} + \mathbf{s}$ where $\mathbf{r}$ are the sinogram counts due to the random events and $\mathbf{s}$ are the sinogram counts due to scatter events. Furthermore, attenuation, normalization and positron range can be modeled by the system matrix. The system matrix $\mathbf{A}$ can be modeled by $\mathbf{A} = \mathbf{NLXH}$. $\mathbf{N}$ is a matrix containing normalization factors which account for differences in detector response. $\mathbf{L}$ is a matrix containing attenuation factors, accounting for the reduced sinogram counts in bins which correspond to LORs which traverse through more attenuating tissue. $\mathbf{X}$ is the line transform, and $\mathbf{H}$ is a blurring matrix which approximates for positional uncertainties in positron range

and crystal resolution [30]. The MLEM algorithm then has the form, using matrix notation:

$$\mathbf{n^{k+1}} = \frac{\mathbf{n^k H^T X^T L^T N^T}}{\mathbf{H^T X^T L^T N^T 1}} \left(\frac{\mathbf{m}}{\mathbf{NLXHn^k + r + s}} \right) \tag{1.35}$$

where $\mathbf{n}$ (size $J \times 1$) is the image values, and $\mathbf{m}, \mathbf{r}$ and $\mathbf{s}$ (all size $I \times 1$) are column vectors containing the raw measured sinogram counts, the random counts and the scatter counts, respectively, and the $\mathbf{X}$ (size $I \times J$) is line transform matrix, $\mathbf{L}$ and $\mathbf{N}$ (each size $I \times I$) are the attenuation and normalization matrices respectively and $\mathbf{H}$ (size $J \times J$) is the blurring matrix

CHAPTER 2
Kinetic Analysis and Conventional Segmentation

PET imaging produces quantitative images representing the spatial distribution of radioactivity within the subject. A single image (static image) may be reconstructed, or the full time-course of the spatial distribution of the radioactivity can be measured by reconstructing images as a function of time of scan (figure 2–1). If an appropriate radio-tracer is used, the time dependent activity values measured in an ROI should correlate to physiologically relevant parameters. Kinetic models attempt to describe the relationship between the time-course evolution of measured activity and these physiological parameters of interest.

It is common practice to correct dynamic images for both *frame length* and *isotope decay*. For instance, a typical dynamic [^{11}C]raclopride scan usually lasts 1 hour, and the reconstruction produces a sequence of 26 images. Each image represents the spatial distribution of detected emissions *during a particular time frame*. Due to the decay of the radiotracer within the body, earlier time frames are short (30 seconds) and later time frames are longer (300 seconds), since more time is required to detect a comparable number of coincidence events during the later part of the scan. Frame length corrections normalize the effect of the varying length of the time frames, such that the activity distribution of a single time frame is representative of the activity at a single time located in the middle of the frame. On the other hand, correcting for isotope decay takes into account the fact that

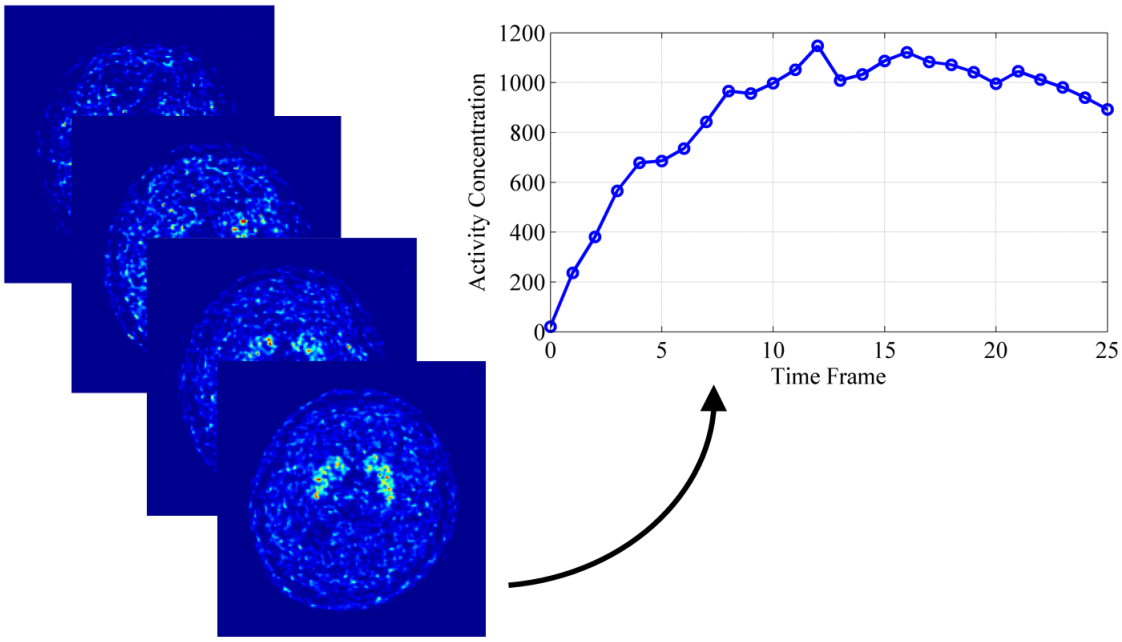


Figure 2–1: Example of a time-activity curve extracted from a real dynamic [^{11}C]raclopride image. Here, the mean activity concentration of the left putamen is calculated for every frame and plotted.

the activity of the radiotracer within the body decreases naturally with a known half-life, resulting in relatively lower counts for the later time frames.

The time-course evolution of the (corrected) measured activity depends on two factors. Firstly, it depends on tissue physiology (for example, the D2/D3 receptor density in the case of [^{11}C]raclopride). Secondly, it depends on the input function; the time-course evolution of the tracer radioactivity within the blood, which defines the actual availability of the tracer for uptake by the target organ. A kinetic model is a mathematical description of the relationship between these factors; a model can predict the time-course evolution of tissue radioactivity given

knowledge of the underlying physiology and the input function. The usefulness of such a model lies in the fact that we can fit the measured time-activity data to the model, and estimate physiologically relevant parameters. Though many types of kinetic models exist [31], we will focus on the most commonly used method for modeling tracer pharmacokinetics: compartmental models. These models assume that tracer molecules exist, at any given time, in one of many compartments. Each compartment defines a specific physical location of the tracer (i.e. blood, extracellular space, intracellular space) and the molecule's physical state (i.e. bound or unbound). Emphatically, compartmental models also specify how the tracer "moves" between compartments; for instance, a receptor-binding tracer that is free may become bound to a receptor. This "movement" is described by a so-called *rate constant* "k" which specifies the fractional rate of change of the tracer concentration in one compartment, and has units of inverse time.

2.1 Compartmental Models: Two-tissue Model

In a two-tissue model, the tracer is taken up (K_1) by the plasma (concentration C_a) into a second compartment (brain tissue; non-specifically bound receptor with concentration C_2). A fraction diffuses back (k_2) into the plasma, and another fraction (k_3) moves further into a third compartment (brain tissue, specifically bound receptor with concentration C_3). Finally, the tracer is allowed to return from the specifically bound chemical state to the non-specifically bound state (k_4). See figure 2.1[1] .

[1] Compartmental model images from Marie Bieth, used with permission.

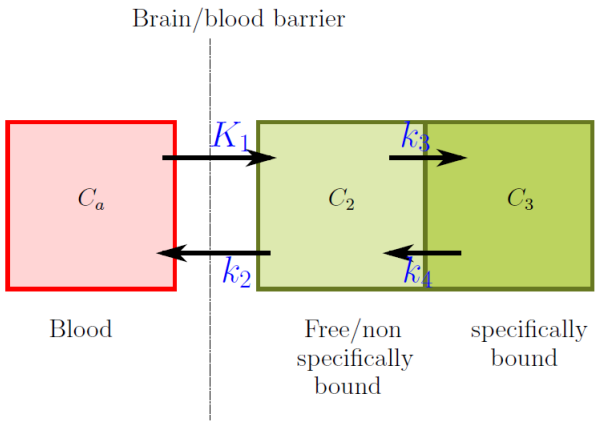


Figure 2–2: Two-tissue, three-comparment kinetic model. The ligand is free to diffuse between the blood and the free and bound states in the brain. The diffusion rates are described by the constants K_1, k_2, k_3 and k_4.

The parameter of interest is the binding potential (BP) defined as $BP = \frac{K_1 k_3}{k_2 k_4}$. The two-tissue, three compartment model is governed by the following set of differential equations:

$$\begin{aligned} \frac{dC_2}{dt} &= K_1 C_a - (k_2 + k_3) C_2 + k_4 C_3 \\ \frac{dC_3}{dt} &= k_3 C_2 - K_4 C_3 \end{aligned} \tag{2.1}$$

In practice, C_2 and C_3 cannot be differentiated on the image, and we only have knowledge of $C_T = C_2 + C_3$. Solving for (equation 2.1) yields:

$$C_T(t) = \alpha_1 C_a(t) \otimes \exp(\alpha_2 t) + \alpha_3 C_a(t) \otimes \exp(\alpha_4 t) \tag{2.2}$$

where the α_i are functions of the rate constants K_1, k_2, k_3 and k_4. To compute the parameters of interest, equation (2.2) can be fitted to a given time-activity curve (TAC), yielding estimates of the rate constants and there of BP. Note that

TACs can be extracted voxel-wise from the reconstructed images, or from regions as a whole. The two-tissue compartmental model is of limited practical use since this method requires a measurement of C_a, the plasma tracer concentration as a function of time. In practice, this is done by arterial blood sampling through the length of the scan, which is uncomfortable and subject to error due to the complicated analysis required of the blood sample. Still, the two-tissue compartmental model is well established in the literature and is therefore a suitable benchmark when developing simpler or more practical kinetic models.

2.2 Compartmental Models: One-tissue Model

Recall that in the two-tissue/three-compartment model, the tracer could exist in two compartments corresponding to two chemical states: specifically bound or non-specifically bound. When these two tissue compartments are practically indistinguishable, (i.e. when the exchange between these two compartments are sufficiently fast) the one-tissue model simplifies the two tissue model (see figure 2–3) to the following single differential equation:

$$\frac{dC_T}{dt} = K_1 C_a - k_2 C_T \tag{2.3}$$

which can be solved to obtain the following model of the time-activity curve:

$$C_T(t) = K_1 C_a(t) \otimes \exp(k_2 t) \tag{2.4}$$

which can again be fitted to a given TAC to characterize it's physiologically relevant parameters. In this case, the binding potential BP is defined as $\frac{K_1}{k_2}$. As

with the two-tissue model, the one-tissue model suffers from the need to perform arterial sampling for the length of the scan.

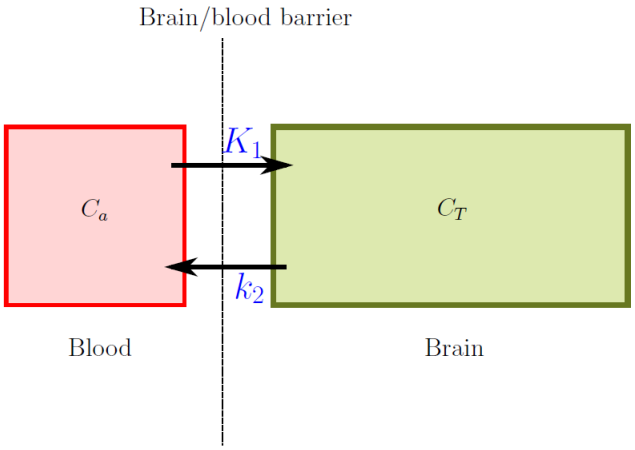


Figure 2–3: Two-tissue, two-comparment kinetic model. The ligand is free to diffuse between the blood and the brain. The diffusion rates are described by the constants K_1 and k_2.

2.3 Reference Tissue Models

Reference tissue models were proposed to overcome the need for arterial blood sampling. The strategy here is to consider compartments corresponding to functional or anatomical regions with useful physiological properties; a *reference region* is defined as a region in which no specific binding occurs. In the reference tissue model, the time course uptake of a radiotracer in a tissue of interest is modeled in terms of its uptake in the reference tissue, assuming that the level of nonspecific binding is the same in both tissues. The model is depicted in figure 2–4 (a).

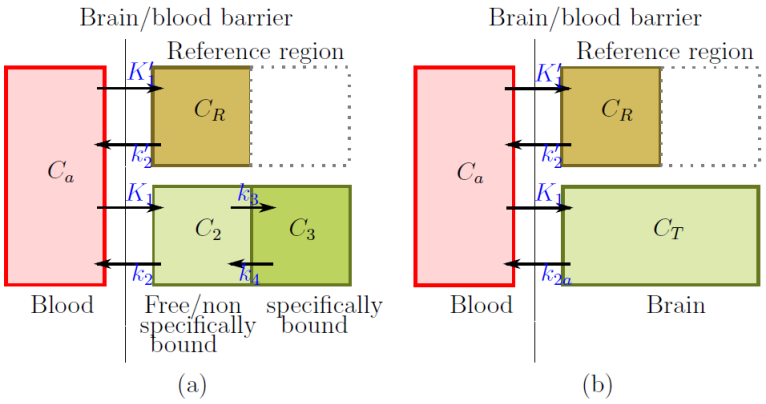


Figure 2–4: Reference tissue models. (a) The classic reference tissue model is comprised of three tissue types and four compartments. The reference region is assumed to be a region where no specific binding occurs. In contrast, both non-specific and specific binding can occur in the brain compartment. (b) The simplified reference tissue model (SRTM) is comprised of three tissue types and only three comparments. It simplifies the former model by allowing the non-specific and specific binding states to be indistinguishable.

The original reference tissue model is governed by the following differential equations:

$$\begin{aligned} \frac{dC_R}{dt} &= K_1' C_a - k_2' C_R \\ \frac{dC_2}{dt} &= K_1 C_a - (k_2 + k_3) C_2 + k_4 C_3 \\ \frac{dC_3}{dt} &= k_3 C_2 - k_4 C_3 \end{aligned} \tag{2.5}$$

After sufficient time, the net tracer flux between the arterial blood and the non-specific tissue compartment will be zero. The *volume of distribution* is defined as the ratio between the equilibrium concentrations of a given tissue and the arterial blood. By letting $R_1 = \frac{K_1}{K_1'}$ and assuming that the volume of distribution is the same in the reference region and in the non-specific compartment of the tissue of

interest:

$$\frac{K_1'}{k_2'} = \frac{K_1}{k_2} \tag{2.6}$$

equations (2.5) can be solved and an expression for C_T can be derived in terms of R_1, k_2, k_3 and k_4. Fitting a TAC to the model of C_T allows for the estimation of the *non-displaceable binding potential*:

$$BP_{ND} = \frac{k_3}{k_4} \tag{2.7}$$

However, the reference model can be significantly simplified again assuming that the free and non-specifically bound compartments are indistinguishable. This model, known as the *simplified reference tissue model* (SRTM) is widely used to model radiotracers that bind to brain receptors, such as [^{11}C]raclopride binding to D2/D3 dopamine receptors. The model is depicted in figure 2–4 (b) and is governed by the following set of differential equations:

$$\begin{aligned} \frac{dC_R}{dt} &= K_1' C_a - k_2' C_R \\ \frac{dC_2}{dt} &= K_1 C_a - k_{2a} C_T \end{aligned} \tag{2.8}$$

The key simplification is that the volume of distribution of the simplified and non-simplified tissues of interest remain unchanged:

$$\frac{K_1}{k_{2a}} = \frac{K_1}{k_2}(1 + BP_{ND}) \tag{2.9}$$

allowing a solution of equation (2.8) where C_T depends on only three paramters: R_1, k_2 and BP_{ND}:

$$C_T(t) = R_1 C_R(t) + k_2 \left(1 - \frac{R_1}{1 + BP_{ND}}\right) C_R(t) \otimes \exp\left(\frac{k_2}{1 + BP_{ND}} t\right) \tag{2.10}$$

2.4 Implementation of SRTM: Basis Function Method

The goal in kinetic analysis is to fit a relevant mathematical model to a TAC extracted from a relevant tissue of interest (normally, TACs are extracted from voxels or homogeneous regions consisting in many voxels). Re-writing equation (2.10) in a more tractable form, we have:

$$C_T(t) = \theta_1 C_R(t) + \theta_2 C_R(t) \otimes \exp(\theta_3 t) \tag{2.11}$$

where the θ_i are functions of the parameters R_1, k_2 and BP_{ND}. While equation (2.11) is non-linear in the unknowns, given a fixed θ_3, θ_1 and θ_2 can be estimated using standard least squares. Gunn *et al.* [32] proposed the "basis function method" (BFM) and defined a set of basis functions $B_i(t) = C_R(t) \otimes \exp(\theta_{3_i} t)$ in order to re-write equation (2.11) into a linear equation for each basis function:

$$C_T(t) = \theta_1 C_R(t) + \theta_2 B_i(t) \tag{2.12}$$

Equation (2.12) can then be solved using standard linear least squares for each basis function; the index i which minimizes the residual sum of squares is determined by a direct search, and the associated parameters for this index (θ_1, θ_2 and θ_3) are obtained. A fast algorithm for the basis function method is described by Gunn *et al.* which computes weighted least squares solutions by QR decomposition.

2.5 Segmentation

A natural problem associated with kinetic fitting is that of segmentation. Voxel-wise kinetic fitting using a reference tissue model requires a reference region, and delineation of the reference region from a PET image is not a straightforward task. ROI-based kinetic fitting requires, in addition to the delineation of the reference region, the delination of relevant ROIs. As discussed before, conventional segmentation techniques rely on manual drawing of the ROIs on the PET image itself or on co-registered anatomical images, or applying an automatic segmentation strategy to the co-registered anatomical image. Both options suffer from limitations; manually drawn ROIs are prone to user-dependency and poor reproducibility, and co-registration with anatomical images is costly. For these reasons, there has been an increased interest in automatic segmentation of PET images using purely functional information. In this section we will outline several popular segmentation algorithms for brain PET, and address their limitations.

2.5.1 k-means

The k-means is one of the simplest unsupervised clustering algorithms [33]. Given a set of observations $(x_1, x_2, ..., x_N)$, where each observation x_i is a finite-dimensional vector, the k-means algorithm seeks to partition the N observations into K clusters, so as to minimize the objection function:

$$J = \sum_{j=1}^{K} \sum_{i=1}^{N} \left\| x_i^{(j)} - \mu_j \right\|^2 \tag{2.13}$$

where $\left\| x_i^{(j)} - \mu_j \right\|^2$ is the distance between a data point belonging to cluster j and the cluster mean (also called the cluster "centroid"), μ_j. The objective function is therefore often referred to as the "within-cluster sum of squares." The k-means algorithm is performed as follows:

while *not converged* **do**

Initialize the cluster centroids μ_j;

Assign each data point x_i to the cluster j with the closest centroid μ_j;

Determine the new cluster centroids μ_j;

Iterate between steps 2 and 3 until the centroids no longer move;

end

Algorithm 2: k-means algorithm

Wong *et al.* [34] applied the k-means algorithm to dynamic PET and clustered voxels based on TAC similarity (i.e. x_i is a single voxel TAC). The application showed only limited success: the algorithm was found to be highly sensitive to initial cluster centroids, and the optimum number of clusters for cluster analysis is not known a priori. In addition, "false kinetics" created by partial volume effects and motion were wrongly assumed to belong to a single cluster. Lastly, no spatial contextual information is used (for instance, neighbouring voxels have correlated TACs) and, like other crude cluster analysis algorithms, it is highly sensitive to noise.

2.5.2 Fuzzy c-means

The fuzzy c-means (FCM) algorithm extends and improves the k-means method by allowing a single data point to belong to more than one cluster; associated with each data point is a set of membership levels. The membership levels indicate the strength of association between the data point and each cluster; the FCM algorithm is a means of determining these optimal membership levels. The FCM algorithm seeks to minimize the objective function:

$$J = \sum_{j=1}^{K} \sum_{i=1}^{N} u_{ij}^{m} \left\| x_i - \mu_j \right\|^2 \tag{2.14}$$

where m is a real number greater than 1, u_{ij} is a matrix which contains the degree of membership of x_i in the cluster j, and μ_j is the centroid of cluster j. The membership matrix u_{ij} has the additional properties that:

$$\begin{aligned} u_{ij} \in [0,1],\ \forall i,j \\ \sum_{j=1}^{K} u_{ij} = 1 \\ 0 \leq \sum_{i=1}^{N} u_{ij} \leq N,\ \forall j \end{aligned} \tag{2.15}$$

The objective function therefore differs from the k-means objective function by the addition of the membership matrix u_{ij} and the m term. The m term is called the "fuzzifier" and determines the level of cluster fuzziness: large values of m result in smaller memberships u_{ij} and vice versa. It can be shown [35] that the objective function can be optimized by iterative updates of the membership matrix u_{ij} and the cluster centroids c_j. The FCM algorithm has inherent advantages

over the k-means algorithm for the segmentation of PET images [36]. Due to motion and partial volume effects, single voxels often have contributions from several tissue types. Using "fuzzy"' techniques, which allow voxels to belong to several groups, enables a more realistic model and therefore delivers more meaningful clusters. However, FCM suffers from largely the same problems as the k-means algorithm. Due to a disregard of spatial contextual information, the FCM algorithm generally fails to segment images corrupted by noise. A comprehensive study of PET brain image clustering methods found that the FCM algorithm failed to satisfactorily segment dynamic [^{11}C]raclopride data [37]. In particular, the FCM algorithm failed to separate the striatum from the surrounding white matter, and could not differentiate the cerebellum (the reference region for kinetic analysis of [^{11}C]raclopride data using the SRTM) from grey matter.

2.5.3 Gaussian Mixture Model

Gaussian mixture models (GMMs) model the observed data $(x_1, x_2, ..., x_N)$ as independent realizations of a random vector distributed with a probability density function given by a weighted sum of K component probability distribution functions [38]. The relative probability of observing a value x_i with respect to the k^{th} component probability distribution function is (assuming 1-dimensional data for convenience):

$$f_k\left(x_i|\theta_k\right) = \frac{1}{\sqrt{2\pi\sigma_k^2}} \exp \frac{\left(x_i - \mu_k\right)^2}{2\sigma_k^2} \qquad \theta_k = (\mu_k, \sigma_k) \tag{2.16}$$

where μ_k and σ_k are the mean and standard deviations of the k^{th} Gaussian component probability distribution, respectively. Summing the component distributions to obtain the complete model probability distribution function of the observed data, we have:

$$f\left(x_i|\Phi\right) = \sum_{k=1}^{K} \pi_k f_k\left(x_i|\theta_k\right) \tag{2.17}$$

where $\Phi = (\pi_1, \pi_2, ..., \pi_K, \theta_1, \theta_2, ..., \theta_K)$ are the model parameters to be estimated, and the π_k's are called the "mixing proportions," satisfying the conditions:

$$\begin{aligned} 0 \leq \pi_k \leq 1 \\ \sum_{k=1}^{K} \pi_k = 1 \end{aligned} \tag{2.18}$$

GMM seeks to optimize the fit between the observed data and the complete model by seeking the model parameters Φ that minimize the likelihood function:

$$L\left(\Phi\right) = \prod_{i=1}^{N} f\left(x_i|\Phi\right) = \prod_{i=1}^{N} \frac{\pi_k}{\sqrt{2\pi\theta_k^2}} \exp \frac{\left(x_i - \mu_k\right)^2}{2\theta_k^2} \tag{2.19}$$

It can be shown [38] that minimizing the ML estimate of the parameters can be done in the Expectation-Maximization (EM) framework. Koivistoinen *et al.* [37] found that EM clustering of real dynamic [^{11}C]raclopride PET brain images based on TACs performed better than FCM, successfully segmenting the striatum, but was unable to distinguish the cerebellum and grey matter. Chen *et al.* [39] applied a modified EM segmentation scheme to dynamic PET images with the goal of automatic extracting of reference tissue TACs for kinetic modeling. They included a Markov random field model (MRF) to provide a way to incorporate

spatial contextual information between voxels, demonstrating increased robustness to the intrinsic noise of dynamic PET. However, the number of clusters and the parameters that control the spatial correlation imposed by the MRF model are often somewhat *ad hoc*.

CHAPTER 3
Image Registration

Medical image registration is the process of determining the correspondence of features between images collected at different times or using different imaging modalities [40]. These correspondences can be used to change the appearance of one image so that it better resembles another, so that the pair can be combined, compared and/or analyzed directly. In the following sections, we will review several popular intra-modality registration algorithms (to be used for inter-subject image registration), with particular emphasis on the demons algorithm, which is used in this work.

3.1 Affine Registration

An affine registration is the simplest non-linear registration technique. An affine transformation is any transformation such that parallel lines remain parallel after being transformed. Affine transformations can therefore consist of any combinations of scaling, rotation, shearing and translation. In practice, an affine registration offers a global aligment between images. This may suffice for some applications, such as monomodal intra-subject registration of different time frames. For other applications, such as monomodal intersubject registration, affine registrations are not sufficient, since they are incapable of accounting for local differences between the images. In these cases, an affine registration is normally done prior to deformable registration to ensure an adequate initialization. For

each point (x_1, x_2, x_3) in an image, and affine mapping can be defined into the coordinates of another space (x'_1, x'_2, x'_3). This can be expressed in matrix notation as $x' = \mathbf{S} \cdot x$:

$$\begin{bmatrix} x'_1 \\ x'_2 \\ x'_3 \\ 1 \end{bmatrix} = \begin{bmatrix} s_{11} & s_{12} & s_{13} & s_{14} \\ s_{21} & s_{22} & s_{23} & s_{24} \\ s_{31} & s_{32} & s_{33} & s_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ 1 \end{bmatrix} \tag{3.1}$$

The moving image $M(p)$ is transformed to the image $M'(p)$ via the matrix $\mathbf{S}$, such that

$$M'(p) = M(\mathbf{S} \cdot p) \tag{3.2}$$

The matrix parameters s_{ij} can be estimated by optimizing a cost function. A popular choice of cost function is the mean squared difference between images, defined as:

$$E(F, M, \mathbf{S}) = \sum_p b_i(\mathbf{S}) = \sum_p (M(\mathbf{S} \cdot p) - F(p))^2 \tag{3.3}$$

A simple choice for the minimization of the cost function E is the Gauss-Newton algorithm [41]: starting with an initial estimate of the matrix $\boldsymbol{s}$, it is iteratively updated:

$$\mathbf{S}^{t+1} = \mathbf{S}^t + \delta \tag{3.4}$$

with $\delta = -(\mathbf{J^T J})^{-1} \mathbf{J^T b}$, $\mathbf{b}$ being the vector of functions b_i, and $\mathbf{J}$ the Jacobian of $\mathbf{b}$ with respect to $\mathbf{S}$ and evaluated at $\mathbf{S^t}$. An example of affine transformations applied to a real time-summed [^{11}C]raclopride image is shown in figure 3–1.

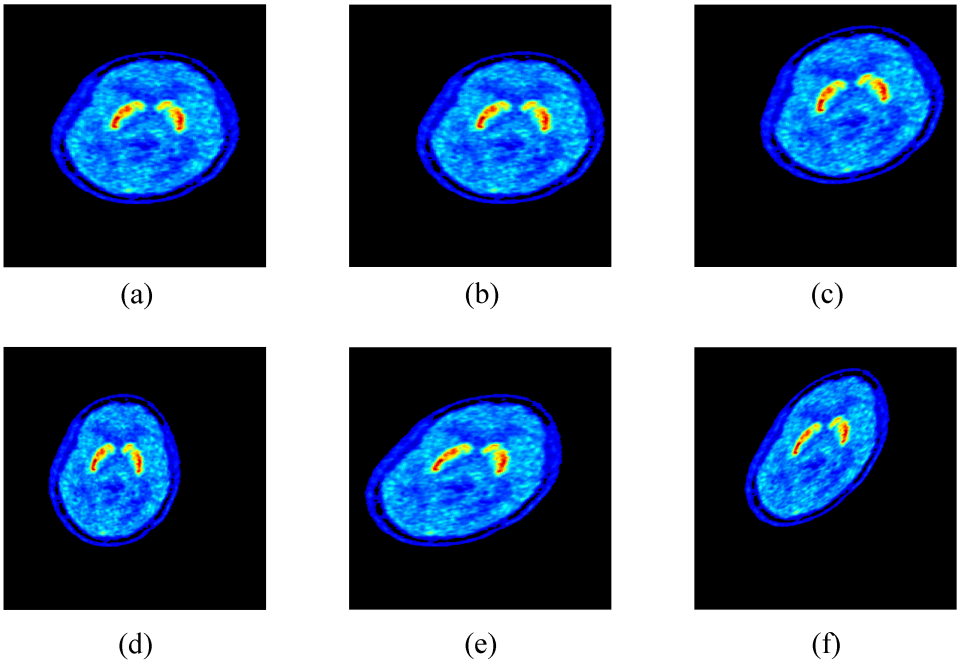


Figure 3–1: Examples of different affine transformations applied to a real time-summed [^{11}C]raclopride image. (a) Original image, (b) translation, (c) rotation, (d) scale, (e) shear, (f) combination of all previous transformations.

3.2 Demons Algorithm

The goal of the classical demons algorithm is to minimize the distance between two images $F(p)$ and $M(s \circ p)$, where F is a fixed image, M is a moving image, p is the voxel and s is a spatial transformation. The distance is defined as:

$$Sim(F, M, s) = \frac{1}{2}\|F - M \circ s\|^2 = \frac{1}{2}\sum_p \|F(p) - M(s(p))\|^2 \tag{3.5}$$

In order to register the fixed and moving images, the distance needs to be optimized over a given space of spatial transformations. However, a simple optimization of $Sim(F, M, s)$ over the space of non-parametric transformations leads to

an ill-posed problem with unstable solutions. The classical demons algorithm, re-interpreted by Vercauteren *et al.* [42], aims to minimize instead the distance $S(F, M, s)$ by minimizing the global energy E defined by:

$$E(c, s) = \frac{1}{\sigma_i^2} Sim(F, M, c) + \frac{1}{\sigma_x^2} dist(s, c)^2 + \frac{1}{\sigma_T^2} Reg(s) \tag{3.6}$$

where σ_i accounts for the uncertainty on the image intensity, σ_x accounts for the uncertainty on point correspondences c, σ_T controls the amount of regularization on the transformation, and c is a hidden variable called the 'correspondence'. Classically, we have $Sim(c) = \|F - M \circ c\|^2$, $dist(s, c) = \|c - s\|$ and $Reg(s) = \|\nabla s\|$. The inclusion of the hidden variable c allows for an alternating iterative optimization, between s and c, of the energy E: the first step solves for the correspondences by optimizing $\frac{1}{\sigma_i^2} Sim(F, M, c) + \frac{1}{\sigma_x^2} dist(s, c)^2$ with respect to c, and with s given, by making a step from $c = s$. The second steps solves for the regularization by optimizing $\frac{1}{\sigma_x^2} dist(s, c)^2 + \frac{1}{\sigma_T^2} Reg(s)$ with respect to s, and with c being given. The solution of the second step, with $Reg(s) = \|\nabla s\|^2$, is a simple convolution of the correspondence field c by a Gaussian kernel. The energy E can be modified to include a fluid regularization term whose solution is also a convolution with a Gaussian kernel, but applied directly to the update field u. The classical demons algorithm is as follows:

while *E is not stationary* **do**

Given s, compute update field u by minimizing

$E^{corr}(u) = \|F - M \circ (s+u)\|^2 + \frac{\sigma_i^2}{\sigma_x^2}\|u\|^2$ with respect to u ;

If using fluid-like registration, let $u \leftarrow K_{fluid} \star u$;

Let $c \leftarrow s + u$;

If using diffusion-like registration, let $s \leftarrow K_{diff} \star c$;

end

Algorithm 3: Demons Algorithm

To minimize $E^{corr}(u)$ we start with the approximation $F(p) - M \circ (s+u)(p) \approx F(p) - M \circ s(p) + J^p.u(p)$ where J^p is the Jacobian of $F(p) - M \circ s(p)$, evaluated at pixel p, and must be approximated. The energy E^{corr} can then be rewritten in matrix form:

$$E^{corr} \approx \frac{1}{2}\sum_p \|\begin{bmatrix} F(p)-M\circ s(p) \\ 0 \end{bmatrix} + \begin{bmatrix} J^p \\ \frac{\sigma_i(p)}{\sigma_x} \end{bmatrix}.u(p)\|^2 \tag{3.7}$$

Taking the gradient of E^{corr} and setting it to zero, we need to solve at each pixel p the following normal equations:

$$\begin{bmatrix} {J^p}^T \frac{\sigma_i(p)}{\sigma_x} I \end{bmatrix}.\begin{bmatrix} J^p \\ \frac{\sigma_i(p)}{\sigma_x} I \end{bmatrix}.u(p) = -\begin{bmatrix} {J^p}^T \frac{\sigma_i(p)}{\sigma_x} I \end{bmatrix}.\begin{bmatrix} F(p)-M\circ s(p) \\ 0 \end{bmatrix} \tag{3.8}$$

$$({J^p}^T.J^p + \frac{\sigma_i^2(p)}{\sigma_x^2} I).u(p) = -(F(p) - M \circ s(p)).{J^p}^T \tag{3.9}$$

Applying the Sherman-Morrison formula[1] to find the inverse matrix, we have:

$$u(p) = -\frac{F(p) - M \circ s(p)}{\|J^p\|^2 + \frac{\sigma_i^2(p)}{\sigma_x^2}} J^{p^T} \tag{3.10}$$

An example transformation s is shown in figure 3–2.

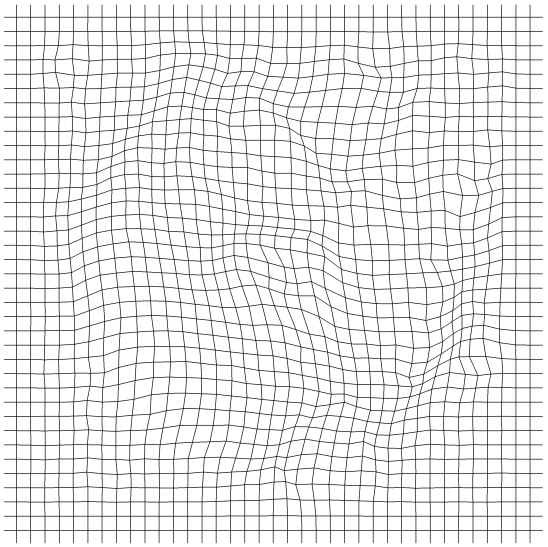


Figure 3–2: Example of a demons transformation applied to a uniform grid.

[1] The Sherman-Morrison formula computes the inverse of a matrix A perturbed by an outer product, uv^T, of vectors u ad v. If A is an inveritble square matrix then $(A + uv^T)^{-1} = A^{-1} - \frac{A^{-1}uv^TA^{-1}}{1+v^TA^{-1}u}$.

3.3 Log-Domain Diffeomorphic Demons Algorithm

The algorithm discussed in the previous section returns a smooth displacement, but it is not always diffeomorphic. However, for many applications the inverted transformation is necessary. Like most spatial transformation spaces, the space of diffeomorphisms do not form a vector space with respect to addition; i.e. adding two diffeomorphisms may result in non-invertibility. However, diffeomorphisms can be smoothly composed and inverted.

An efficient computational framework for diffeomorphisms was proposed in [43], using a Lie group structure that defines an exponential mapping from the vector space of smooth stationary velocity fields to diffeomorphisms. The main idea of the diffeomorphic log-demons is to represent the current spatial transformation s as an exponential of a smooth velocity field v, i.e. $s = exp(v)$ and to search for an update field u on the Lie algebra (the vector space of velocity fields). Then, the update step has the form $s \circ exp(u)$, i.e. $exp(v) \circ exp(u)$.
If a log-domain optimization is to be carried out, one needs to know whether for any velocity fields v and u, there exists another velocity field w such that $exp(w) = exp(v) \circ exp(u)$.

Making the assumption that the update velocity field u is small, a smooth velocity field $Z(v, \epsilon u)$ is desired such that

$$exp(Z(v, \epsilon u)) \approx exp(v) \circ exp(\epsilon u) \tag{3.11}$$

where ϵ is used to emphasize that we are looking for an approximation valid for small update fields ϵu but for arbitrary velocoty field v. By using the first terms

of the Baker-Campbell-Hausdorff (BCH) formula, the authors in [44] showed that there exists a valid approximation for small ϵu of the form

$$Z(v, \epsilon u) = v + \epsilon u + \frac{1}{2}[v, \epsilon u] + ... \tag{3.12}$$

where the lie bracket $[v, u]$ provides a velocity field defined at each pixel p by:

$$[v, u](p) = Jac(v)(p).u(p) - Jac(u)(p).v(p) \tag{3.13}$$

The BCH approximation allows the update step $s \leftarrow s \circ exp(u)$ to be re-written into a log-domain update $v \leftarrow Z(v, u)$. To keep the same alternate optimization approach as in the classical demon algorithm, the Gaussian smoothing is performed directly on the velocity fields in the log-domain, and the diffeomorphic log-demons algorithm is as follows:

while *E is not stationary* **do**

Given $s = exp(v)$, compute update field u using equation (3.10);

If using fluid-like registration, let $u \leftarrow K_{fluid} \star u$;

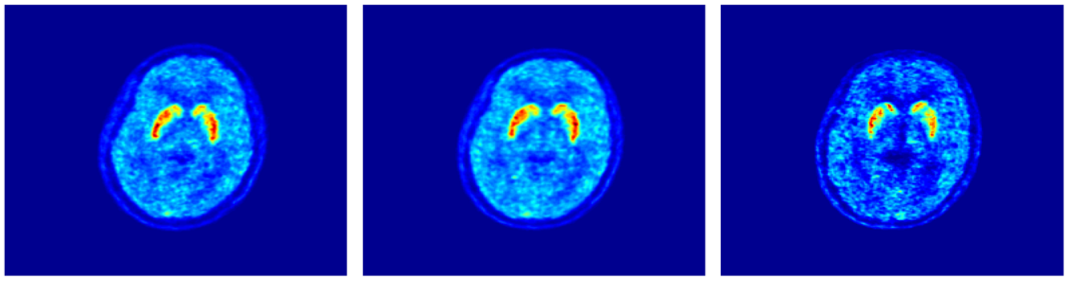
Let $v \leftarrow v + u + \frac{1}{2}[v, u]$;

If using diffusion-like registration, let $v \leftarrow K_{diff} \star v$;

end

Algorithm 4: Diffeomorphic log-demons algorithm

An example of a diffeomorphic log-domains registration of two real [^{11}C]raclopride images is shown in figure 3–3.



(a) (b) (c)

Figure 3–3: Example of log-domain diffeomorphic demons registration. An image registered via affine registration (a) to the target (c) is non-linearly registered to the target using the demons algorithm (b).

CHAPTER 4
Atlas-Based Segmentation of HRRT Brain PET

The following chapter is largely based on a manuscript entitled "Radiotracer-Specific Atlases for Automated Segmentation and Kinetic Analysis of [^{11}C]raclopride High Resolution Brain PET" authored by Philip Novosad, Paul Gravel and Andrew J. Reader, submitted to Physics, Medicine and Biology (article reference PMB-101213)

4.1 Introduction

Positron emission tomography can generate functional and quantitative dynamic images of tracer uptake in vivo. By segmenting the dynamic image into regions of interest, regional time-activity curves can be extracted to estimate kinetic parameters, which correlate to physiologically relevant parameters [45]. Despite the intrinsic noise and limited spatial resolution of PET, the accuracy of ROI delineation is essential to ensure the extraction of relevant TACs. In practice, segmentation of functional ROIs is performed manually by an expert on the PET images, or structural ROIs are obtained by co-registration to segmented anatomical images [46]. Manually defined ROIs are prone to operator error and poor reproducibility, whereas co-registration to segmented anatomical images necessitates the use of a second imaging system, which is costly and time-consuming. In addition, if the PET tracer uptake is limited to a small number of discrete regions (as in the case of [^{11}C]raclopride images), registration between PET and anatomical images

may be inaccurate [47]. As a result, there has been an increased interest in unsupervised segmentation of PET images, without anatomical knowledge, into regions with homogeneous functional behaviour. Clustering algorithms, for instance, have been applied with some success to dynamic PET images [34] [48] but still suffer from the need for a priori information such as initialization and the optimal number of clusters. In addition, clustering techniques are often insensitive to spatial information, such as the strong correlation between neighbouring voxels, making them highly sensitive to noise. Conventional clustering techniques are mostly unable to distinguish between regions with subtely different functional behaviour, such as the grey matter and cerebellum of [^{11}C]raclopride images [37], limiting their usefulness to specific applications. Recently, segmentation approaches based on radiotracer-specific PET atlases have been explored [47] [49] in which segmentation is achieved by co-registration of the atlas with a source image (the image to be segmented). These approaches have shown promising results compared to conventional methods relying on co-registered anatomical images, while overcoming the aforementioned limitations associated with clustering methods. These recent efforts construct a template by voxel-wise averaging of PET images in a common space. In contrast, the work presented here proposes the use of mean intensity and mean shape three-dimensional (3D) radiotracer-specific brain atlases, constructed using diffeomorphic (one-to-one and invertible) transformations. The term mean shape is to be understood as an image that minimizes the distance between itself and the individual images in the set, where distances are expressed in the space of diffeomorphic transformations [50]. This approach has the distinct advantage of

exploiting a highly appropriate PET-to-PET registration to the template, rather than an MR-to-PET or other inter-modality registration. Dynamic PET images can be segmented by registration of the summed image, obtained by summing the dynamic PET frames, to the radiotracer-specific atlas. The segmented regions can then be used to enable robust and fully automated regional kinetic analysis using a reference tissue kinetic model. Because PET images are characterized by functional contrasts, it is important to emphasize that our segmentation scheme works by delineating functional regions; functional and anatomical regions correspond only when the anatomical region shows homogeneous tracer kinetics. Together with the proposed method, we present its validation by performing a comprehensive simulation of a whole brain [^{11}C]raclopride study, consisting of twenty functionally and anatomically unique subject images, and assess its performance on real dynamic [^{11}C]raclopride images from the high resolution research tomograph (HRRT) [6].

4.2 Methods

4.2.1 Atlas-Based Segmentation and Kinetic Analysis

The method for automated atlas-based segmentation and kinetic analysis, given an image to be segmented (the source image) and a labeled template (the atlas), is described below. Construction of the template and corresponding atlas is described later in section 4.2.4.

Intensity Normalization As a pre-registration step, the source and template images are normalized to account for differences in injected tracer uptake. Both the affine and non-linear registration algorithms used in this work are based on a

sum of squared differences (SSD) optimization, making this step crucial for high quality registration and segmentation results. In this work, histogram matching between the source and template images is performed (256 histogram bins), with only voxels contained within the brain considered when forming the histograms. Brain voxels were masked by thresholding the image at the mean intensity level of all voxels of the image (background voxels included). Dynamic images are time summed prior to intensity normalization.

Atlas Registration The (moving) atlas is then co-registered to the (fixed) source image using an initial affine registration followed by a non-linear log-domain diffeomorphic demons (LDDD) registration (detailed in section 2.3). The affine registration is based on the optimization of a 12-parameter transform (3 translations, 3 rotations, 3 scales and 3 shears) with respect to an SSD objective function. A multiresolution strategy (down-sampling factors of 4, 2 and 1) is used for both the affine and LDDD registrations. After registration, the atlas label set is spatially transformed with the found deformation field, completing the segmentation of the source image. A nearest neighbour interpolation scheme is used to preserve the integer label values.

Kinetic Analysis For quantitative kinetic analysis of dynamic [^{11}C]raclopride, a reference tissue model has been shown to be valid and is widely used to avoid arterial sampling [51]. From the image segmentation, a set of average regional TACs are extracted by voxel-wise averaging of TACs in each segmented region of the dynamic image. Subsequently, regional binding potential estimates are calculated using the simplified reference tissue model with the basis function

method (SRTM-BFM) [32] with 100 basis functions, ($0.001\text{s}^{-1} < \theta_3 < 0.01\text{s}^{-1}$) and with the cerebellum as the reference region.

4.2.2 Simulations

To evaluate the performance of the proposed method, we simulated a 3D [^{11}C]raclopride PET study by generating a total of twenty realistic dynamic brain images. The construction of each simulated image required an anatomical model of the brain (a realistic labeled 3D brain image) and a unique set of regional TACs (each set describing the dynamic behaviour of the radiotracer in specific anatomical structures) for each brain, as shown in figure 4–1. Twenty unique sets of regional TACs were extracted from twenty segmented dynamic [^{11}C]raclopride images (image acquisition details given in the following section) by computing the mean TAC in each region; regions were defined by co-registration of the time-summed dynamic PET images to subject-specific T1-weighted magnetic resonance (MR) images. The MR images were automatically segmented via the CIVET pipeline [52] [53] [54] [55] into structural regions corresponding to grey matter, white matter, the cerebellum, and the left and right caudate and putamen.

The anatomical model is based on the numerical phantom from Rahmim *et al.* [56]. The phantom is a 256×256×207 3D brain image containing integer-valued labels representing the cerebellum, white matter, grey matter, and the left and right putamen and caudate. The twenty sets of tissue TACs were each incorporated homogeneously into the corresponding regions defined on the numerical phantom to yield twenty dynamic brain phantoms, each with twenty-six time frames.

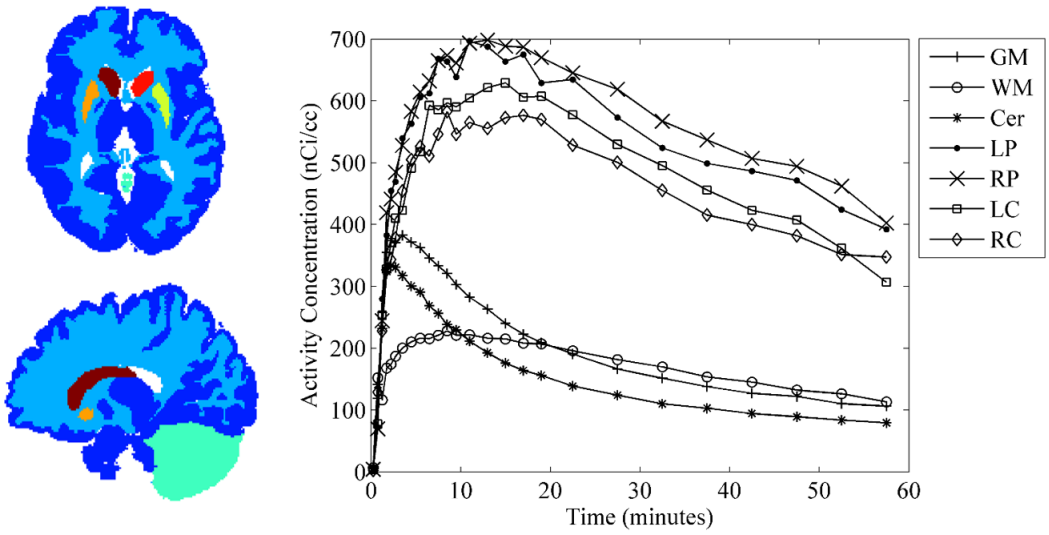


Figure 4–1: The original numerical phantom (anatomical model) and a sample set of regional TACs extracted from a real dynamic [^{11}C]raclopride image. Left: original numerical phantom with seven labeled regions of interest. Right: example of a typical set of regional TACs extracted from real dynamic [^{11}C]raclopride images from the HRRT scanner (GM: grey matter, WM: white matter, Cer: cerebellum, LP: left putamen, RP: right putamen, LC: left caudate, RC: right caudate).

Simulating Variation in Anatomy and Position To simulate realistic variability among brain scans, each of the twenty dynamic brain phantoms were time summed and co-registered with a different time-summed [^{11}C]raclopride image of a healthy subject. A global nine-parameter rigid registration (3 scales, 3 rotations and 3 translations) followed by a non-linear registration using the ANIMAL methodology [55] were used, both based on a cross-correlation objective function. The resulting transformations were then applied (i) to each frame of the dynamic phantom, and (ii) to the original numerical phantom, yielding (i) a unique deformed dynamic phantom, and (ii) the labels corresponding to that dynamic phantom (see figure 2). A nearest-neighbor interpolation scheme was used to preserve the integer

label values of the anatomical model. It is important to emphasize that we used different registration methods (and objective functions) for (i) creating the phantom images and for (ii) atlas-based segmentation (LDDD, based on optimization of an SSD objective function), in order to perform a more realistic validation study.

Dynamic Phantom Reconstruction Dynamic sequences of sinograms were generated by forward projecting each image of the dynamic sequence into a fully 3D set of 2209 sinograms (each of 256×288 bins, radial bin size of 1.21875 mm). Noise was introduced in each sinogram by drawing counts from the Poisson distribution with the mean as the float value in the sinogram bin. The sum of the mean sinogram counts was chosen to be the same as a typical one hour 370 MBq [^{11}C]raclopride scan on the HRRT (a mean total sinogram count of 450×10^6). Attenuation effects were simulated by generating attenuation factor sinograms from a forward projection of a 3D linear-attenuation coefficient distribution, assuming the brain to be equivalent to water for attenuation of 511 keV photons. Resolution modeling [57] [30] was simulated by first degrading the resolution prior to forward projecting and introducing noise. The kernel used for resolution recovery was slightly sharper than the kernel used for resolution degradation to avoid creating Gibbs' artifacts in the reconstructed images. Both kernels are defined by:

$$h(r) = N\left[exp^{\left(\frac{r^2}{2\sigma_1^2}\right)} + p \times exp^{\left(\frac{r^2}{2\sigma_2^2}\right)}\right] \tag{4.1}$$

where $r^2 = x^2 + y^2 + z^2$. The values for both kernels are shown in table 4–1.

Kernel	σ_1 (mm)	σ_2 (mm)	ρ	FWHM (mm)
Resolution Degradation	1.05	2.7	0.07	2.2
Resolution Recovery	0.9	2.5	0.05	1.7

Table 4–1: Parameters used for resolution modeling simulation. The parameters used for resolution recovery are the same as those used on the HRRT scanner.

Twenty different noisy sinogram datasets were then reconstructed using 10 iterations of the Ordinary Poisson Ordered Subset Expectation Maximization (OP-OSEM) algorithm with 16 subsets to obtain twenty simulated [^{11}C]raclopride dynamic brain images, each with 26 time frames (see figure 4–2).

Ground Truth Binding Potential Due to the protocol used to simulate and reconstruct the dynamic phantoms, regional BP estimates obtained by fitting the dynamic phantoms (figure 4–2(b)), using the ground truth regions, differ from the regional BP estimates obtained by fitting the corresponding reconstructed dynamic phantoms (figure 4–2(c)) using the same ground truth regions. To assess the performance of our proposed work independently of any particular reconstruction protocol, we define the ground truth BP as those obtained by fitting the reconstructed dynamic phantoms using the ground truth regions, with BP calculations performed as described in section 4.2.1. Some key characteristics of the twenty simulated images are detailed in table 4–2.

4.2.3 Real Images

A set of twenty real dynamic [^{11}C]raclopride images were used. The dynamic images were acquired on the HRRT scanner over 60 minutes with typical frame

BP

ROI	Min	Max	Mean	SD
Grey Matter	0.173	0.327	0.257	0.046
White Matter	0.366	0.616	0.469	0.076
Cerebellum	—	—	—	—
Left Putamen	2.627	5.001	3.940	0.576
Right Putamen	2.660	4.809	3.847	0.504
Left Caudate	1.493	4.021	3.055	0.623
Right Caudate	1.792	3.839	3.140	0.556

Volume (cc)

ROI	Min	Max	Mean	SD
Grey Matter	573.723	867.758	697.793	68.161
White Matter	366.276	538.700	443.456	41.172
Cerebellum	115.148	176.692	141.361	15.246
Left Putamen	3.213	4.661	3.883	0.413
Right Putamen	2.748	4.452	3.495	0.455
Left Caudate	4.622	7.210	5.661	0.693
Right Caudate	3.785	5.491	4.597	0.459

Table 4–2: Characteristics of the twenty simulated [^{11}C]raclopride images used for validation. BP estimates are obtained by fitting regional TACs extracted from the reconstructed PET images using the corresponding ground truth regions. Volumes are reported in cubic centimeters (cc). Min: minimum, Max: maximum, SD: standard deviation.

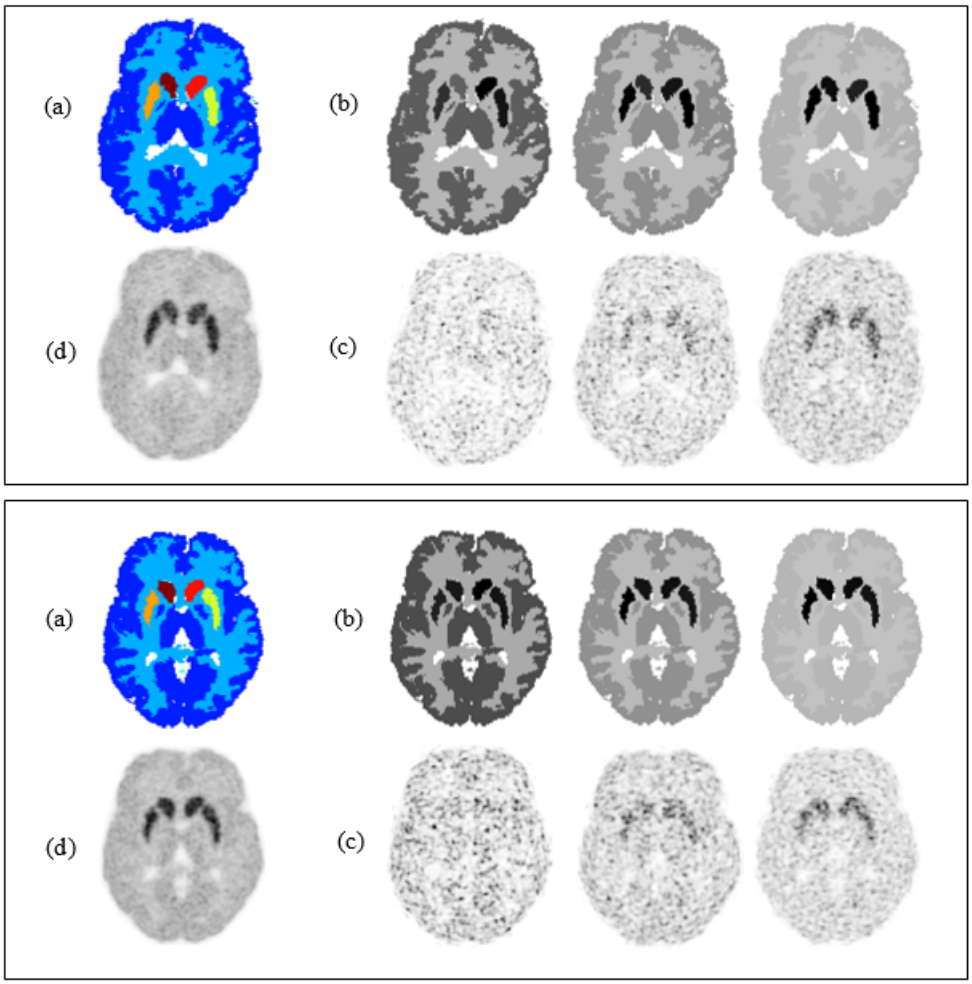


Figure 4–2: Representative sample of two simulated 3D [^{11}C]raclopride images (one transverse slice shown from each image). (a) Spatially transformed ground truth numerical phantom defining true regions. (b) Three time-frames (5, 10 and 15) of the spatially transformed dynamic phantom corresponding to ground truth image. (c) Corresponding reconstructed time-frames. (d) Timesummed reconstructed dynamic image of all twenty-six time-frames.

lengths (6×30s, 7×60s, 5×120s and 8×300s). The images were reconstructed using 10 iterations of the OP-OSEM algorithm [58] with 16 subsets.

4.2.4 Atlas Construction

Templates are constructed using the methodology described by Lombaert *et al.* [59] and based on the work of Guimond *et al.* [60], using the LDDD algorithm. In this method, an arbitrary initial reference image iteratively converges towards the mean intensity and mean shape of all the images in the dataset; a "mean shape image" is to be understood as an image that minimizes the differences between itself and the images in the set, where differences are expressed in the space of diffeomorphic transformations. For a set of 3D brain images (or time-summed dynamic images) $I_1, \ldots, I_N$ in the affine space of an arbitrary reference image, the atlas template R can be expressed as:

$$R^{(t+1)} = \left(\frac{1}{N} \sum_{i=1}^{N} I_i \circ \exp v_i \right) \circ \left(-\frac{1}{N} \sum_{i=1}^{N} v_i \right) \tag{4.2}$$

where v_i is the stationary velocity field of the warping from I_i to the current reference volume, and $\circ$ is the composition function. Explicitly, for a set of N brain images in the affine space of the reference brain image $R^{(t)}$, a non-linear LDDD registration between each (moving) image and the (fixed) reference image is performed. The registered brain images are then intensity-averaged in the shape of the reference image, and the transformation updating the reference image towards the mean shape of all the images is computed (using the negated velocity fields) and applied to yield the reference image for the next iteration, $R^{(t+1)}$ (see figure 4–3). The method produces practically unbiased (independent of the initial reference image) templates and only a few iterations are required to converge to a stable average image. It should be emphasized that all steps and iterations of

the algorithm are fully automated and require no user interaction. In practice, four iterations were sufficient to ensure that the average image converged to a stable estimate. Since the LDDD algorithm is based on the optimization of an SSD-based energy function, and applied to PET images which may have different injected dose, histogram matching between the images and the reference image (256 histogram bins) was performed prior to the estimation of each transformation. The transformations were then applied to the original images, to preserve the original intensity values. Finally, the constructed template is manually labeled (as described in section 4.2.5) into functional regions corresponding to the left and right caudate and putamen, and the cerebellum.

4.2.5 Manual Segmentation

Manual segmentations were performed using the MNI Display software (Montreal Neurological Institute, Canada, www.bic.mni.mcgill.ca/ServicesSoftware) that allows for manual painting of individual image slices from three orthogonal views. Dynamic images were segmented on time summed images, obtained by summation of all dynamic PET frames. The quality of the manual segmentations was verified by an experienced nuclear medicine physician.

4.2.6 Segmentation Quality Metrics

Letting A and B denote two image segmentations, the Dice similarity coefficient (DSC), which quantifies the similarity of the two segmentations for a particular segmented region is defined as:

$$DSC = 2 \times \frac{n(A \cap B)}{n(A) + n(B)} \tag{4.3}$$

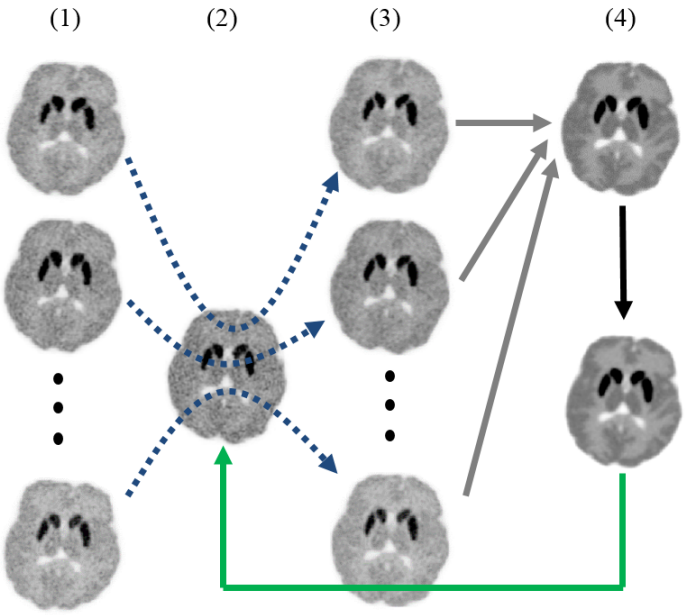


Figure 4–3: Construction of average 3D brain template (simulated data shown). A set of images (column 1) in the affine space ofan arbitrary reference image are non-linearly transformed into the shape of the reference image (column 2) using the LDDD registration algorithm (dashed blue arrows). The registered images (column 3) are intensity averaged in the shape of the reference image (grey arrows, column 4) and the transformation updating the image towards the mean of all images is computed and applied (black arrow). Finally, the resulting image is used as the reference image for the next iteration (green arrow).

where $\cap$ is the intersection of the two regions, and $n()$ returns the number of voxels in a region. The second and third metrics are the percentage error ($^{\%}Error$) and the percentage absolute error ($^{\%}Abs.Error$) on regional BP estimates using an image segmentation:

$$^{\%}Error = 100\% \times \frac{BP_s - BP_t}{BP_t} \tag{4.4}$$

$$^{\%}Abs.Error = 100\% \times \frac{|BP_s - BP_t|}{BP_t} \tag{4.5}$$

where BP_t is the true BP of a region and BP_s is the estimated BP using the image segmentation. As for the DSC metrics, the ($^{\%}Error$) and ($^{\%}Abs.Error$) metrics are defined for each region and image; each image has distinct ground truth regions and ground truth regional BP values.

4.3 Experiments

4.3.1 Simulation Validation

For the simulation study, we compared our atlas-based method to a suitable conventional method. The conventional method for segmentation is defined by manual segmentation of the summed PET images, obtained by summing the dynamic PET images over all time frames. We compared manual segmentation with our atlas-based method using the segmentation quality metrics described in section 4.2.6. The manual segmentation of each simulated image involved identifying the left and right caudate and putamen, and the cerebellum, on the time-summed image.

From the set of twenty simulated dynamic images, an atlas was generated (as described in section 4.2.4) from ten randomly selected "training subjects", using the time-summed images . The atlas (labeled template) was then co-registered to each of the remaining ten "test subjects" for segmentation and kinetic analysis. The experiment was repeated five times (each time randomly selecting the training and test subjects, and manually labeling the resulting template), resulting in five sets of ten segmented images, each image with four regional BP estimates.

4.3.2 Real Data

Without a ground truth image defined, we assessed the performance of atlas-based segmentation visually. Secondly, we assessed the correlation between (i) regional BP estimates from TACs obtained by functional ROIs using atlas-based segmentation, and (ii) regional BP estimates from TACs obtained by functional ROIs using manual segmentation. From the set of twenty real dynamic [^{11}C]raclopride images, fifteen training subjects were selected to build a template. Functional ROIs corresponding to the left and right putamen and caudate, and the cerebellum, were manually labeled on the template. The atlas was co-registered to each of the remaining five test subjects for segmentation and kinetic analysis.

4.4 Results

4.4.1 Simulation Validation

Figure 4–4 depicts a typical atlas derived from 10 simulated training subjects with manually defined labels.

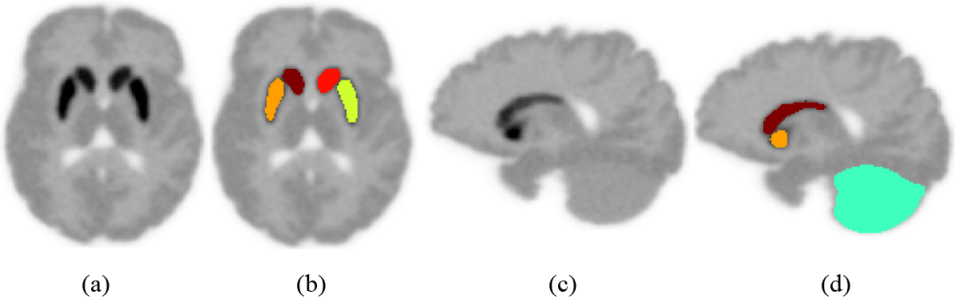


Figure 4–4: (a) Transverse slice of a [^{11}C]raclopride atlas with corresponding manually-defined labels (b), generated from simulated data. (c) Sagittal slice of the same [^{11}C]raclopride atlas with corresponding labels (d).

Figure 4–5 compares typical segmentation results on a single simulated test subject. Visually, both the atlas-based and manualy segmentations are satisfactory.

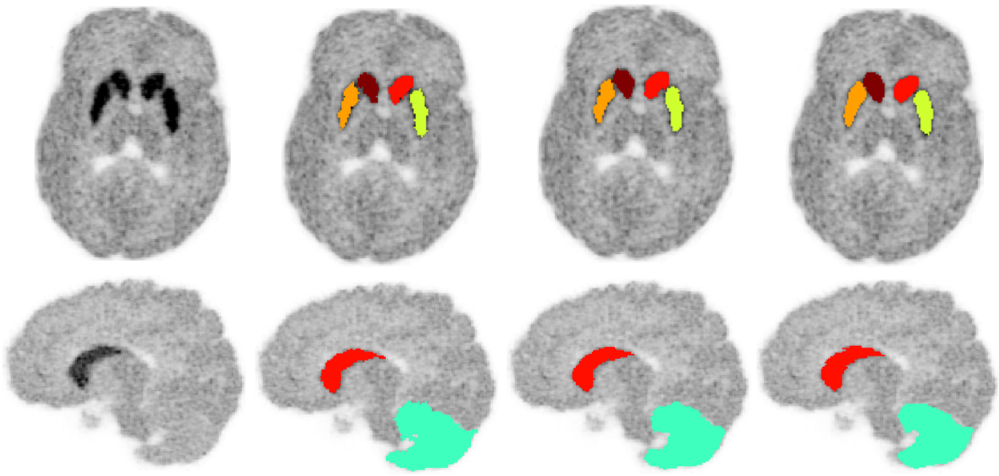


Figure 4–5: Segmentation of a typical simulated image. (a) Typical simulated image, (b) truel labels from the ground truth, (c) atlas-based segmentation, (d) manual segmentation.

The segmentation results were compared to the ground truth regions using the *DSC* metric. The distributions of *DSC*s over all ROIs (cerebellum, left and right caudate and putamen) are shown in the top panel of figure 6, for each experiment and for both segmentation methods: atlas-based segmentation and manual segmentation. To enrich the statistics for our analysis, we combined the data from the five experiments in order to assess the performance of each segmentation method at the ROI level. This step is justified because the distributions of *DSC*s vary only insignificantly between experiments. The bottom panel of figure 4–6 depicts the distributions of *DSC*s for each ROI and each segmentation method

when considering all five experiments together. The data are summarized in table 4–3.

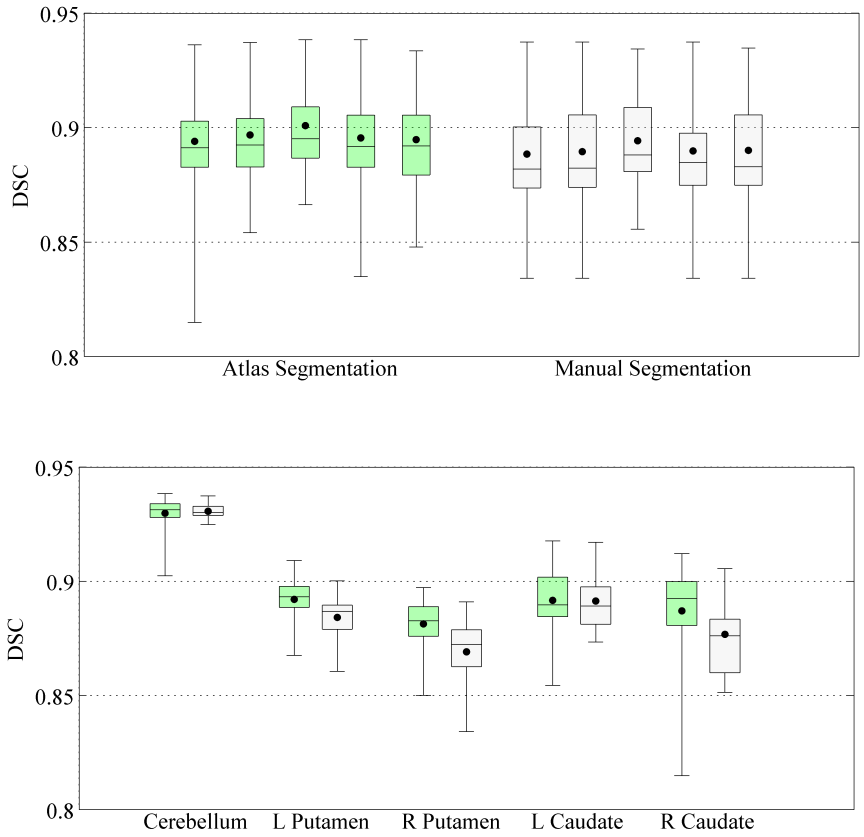


Figure 4–6: Top: distributions of $DSCs$ over all ROIs for each of the five experiments and for two image segmentation methods: atlas-based segmentation and manual segmentation. Each distribution consists of fifty $DSCs$ (ten segmentation images, each image with five ROIs). The upper and lower bounds of the boxes represent the 25th and 75th percentiles of the distribution. The whiskers extend to the most extreme data points, and the black circle represents the mean DSC of the distribution. Bottom: distributions of $DSCs$ when all five experiments are considered together for each ROI (L: left, R: right). Each distribution contains fifty $DSCs$.

For each experiment, atlas-based segmentation yielded superior mean DSCs across all ROIs when compared to manual segmentation. At the ROI level, mean DSCs from atlas-based segmentations were superior for the left and right caudate and putamen (mean $DSC \geq 0.881$ for all ROIs) compared to manual segmentation (mean $DSC \geq 0.869$ for all ROIs). Manual segmentation of the cerebellum (mean $DSC = 0.931 \pm 0.007$) performed comparably to atlas-based segmentation (mean DSC of 0.930 ± 0.003).

For each segmented image, the regional BP of the left and right caudate and putamen were calculated using TACs extracted from the image segmentations. Regional BP values were compared to the ground truth BP values using the $^{\%}Error$ and $^{\%}Abs.Error$ metrics. The top panel of figure 4–7 depicts the distributions of $^{\%}Error$ values on BP estimates over all ROIs for each experiment. To assess the $^{\%}Error$ at the ROI level, we again combined the data from the five experiments. The bottom panel of figure 4–7 depicts the distributions of $^{\%}Error$ on BP estimates for each ROI when all five experiments are considered together. The data are summarized in table 4–3.

For each experiment, atlas-based segmentation achieved both lower $^{\%}Error$ and lower$^{\%}Abs.Error$ values across all ROIs when compared to manual segmentation ($^{\%}Abs.Error \leq 3.364$ for all experiments using atlas-based segmentation, $^{\%}Abs.Error \leq 5.532$ for all experiments using manual segmentation). At the ROI level, $^{\%}Error$ values on regional BP estimates were less negatively biased when using atlas-based segmentation compared to manual segmentation. Using atlas-based segmentation for regional BP calculations also achieved reduced $^{\%}Abs.Error$ values

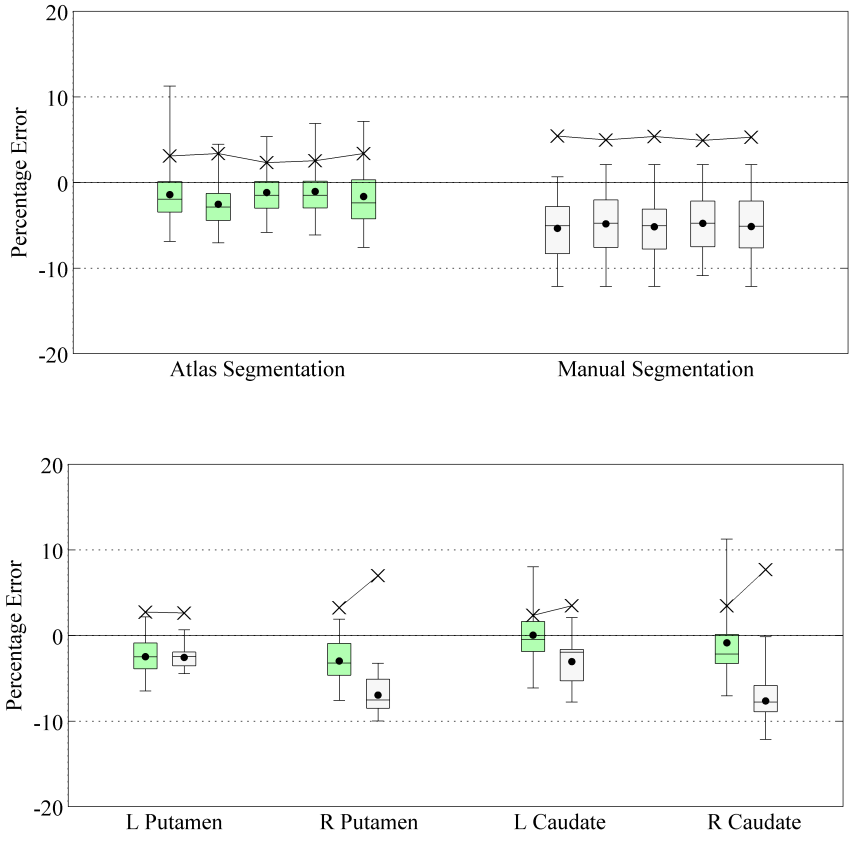


Figure 4–7: Top: distributions of $^{\%}Error$ values over all ROIs for each of the five experiments. Each distribution consists of forty $^{\%}Error$ (ten segmented images each with four regional BP estimates). The upper and lower bounds of the boxes represent the 25th and 75th percentiles of the distribution. The whiskers extend to the most extreme data points, and the black circle represents the mean $^{\%}Error$ of the distribution. The mean $^{\%}Abs.Error$ of each distribution is plotted with the black cross. Bottom: distributions of $^{\%}Error$ values over all five experiments for each ROI (L: left, R: right). Each distribution contains forty $^{\%}Error$ values.

for all ROIs with the exception of the left putamen, where manual segmentation ($^{\%}Abs.Error = 2.714 \pm 1.770$) performed comparably to atlas-based segmentation ($^{\%}Abs.Error = 2.608 \pm 1.157$). Table 4–3 summarizes the results of the ROI level

		Mean DSC	Mean $^{\%}Error$	Mean $^{\%}Abs.Error$
Cerebellum	Atlas	0.930 (0.007)	—	—
Cerebellum	Manual	0.931 (0.003)	—	—
L Putamen	Atlas	0.891 (0.009)	-2.493 (2.073)	2.714 (1.770)
L Putamen	Manual	0.884 (0.010)	-2.583 (1.214)	2.603 (1.157)
R Putamen	Atlas	0.881 (0.010)	-2.993 (3.339)	3.222 (2.006)
R Putamen	Manual	0.869 (0.014)	-6.976 (1.982)	6.976 (1.982)
L Caudate	Atlas	0.892 (0.014)	0.022 (3.018)	2.334 (1.884)
L Caudate	Manual	0.891 (0.013)	-3.069 (2.830)	3.452 (2.337)
R Caudate	Atlas	0.887 (0.020)	-0.872 (4.043)	3.423 (2.273)
R Caudate	Manual	0.877 (0.015)	-7.655 (2.411)	7.655 (2.411)

Table 4–3: Summary of results from ROI level analysis. Mean DSC, $^{\%}Error$ and $^{\%}Abs.Error$ are reported over fifty images segmentations for each image segmentation method. Standard deviation is shown in parentheses.

	Atlas Segmentation	Manual
Mean DSC	0.896 (0.022)	0.890 (0.024)
Mean $^{\%}Error$	-1.584 (3.189)	-5.071 (3.144)
Mean $^{\%}Abs.Error$	2.923 (2.022)	5.173 (2.972)

Table 4–4: Overall performance of each image segmentation method. Mean DSC, $^{\%}Error$ and $^{\%}Abs.Error$ are reported over fifty images segmentations and all ROIs for each image segmentation method. Standard deviations are shown in parentheses.

analysis: the mean DSC, $^{\%}Error$ and $^{\%}Abs.Error$ is shown for each ROI and for each image segmentation method, when considering all five experiments together.

To compare the overall performance of the three methods, we combined all the data from each method (five experiments and all ROIs) to assess the overall mean DSC, $^{\%}Error$ and $^{\%}Abs.Error$. These data are shown in table 4–4.

Overall, atlas-based segmentation performed only slightly better than manual segmentation in terms of regional DSC, but markedly better than manual

segmentation in terms of regional BP calculations using TACs extracted from the segmented regions. This apparent discrepancy can be explained by considering that manual segmentation of the striatal regions (caudate and putamen) on average slightly overestimated the volume of interest (percentage volume difference of 4.9 ± 8.2), while atlas-based segmentation of the same regions on average slightly underestimated the volume of interest (percentage volume difference of -2.7 ± 8.5). Overestimating the volume leads to the inclusion of neighbouring low-intensity voxels in the segmented regions, causing more negatively biased regional BP values.

4.4.2 Real Data

Figure 4–8 depicts the atlas derived from the training set consisting of fifteen real [^{11}C]raclopride images, with manually defined labels corresponding to the left and right putamen and caudate, and the cerebellum.

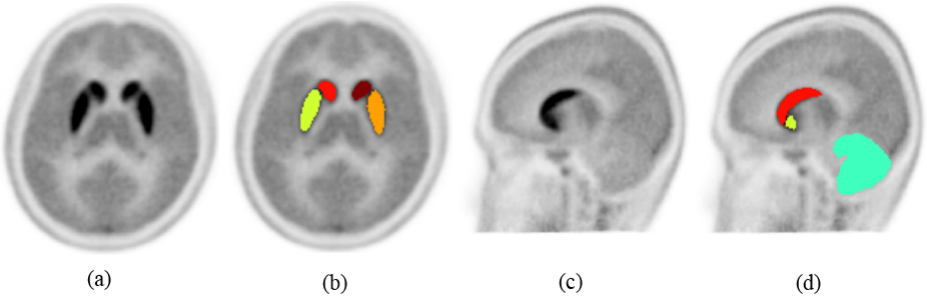


Figure 4–8: (a) Transverse slice of the [^{11}C]raclopride atlas with corresponding labels (b). (c) Sagittal slice of [^{11}C]raclopride atlas with corresponding labels (d).

The atlas was registered to five real time-summed dynamic [^{11}C]raclopride images for segmentation and subsequent regional BP calculations. Segmentation

of a real [^{11}C]raclopride image is shown in figure 4–9, for visual comparison of atlas-based and manual segmentation approaches.

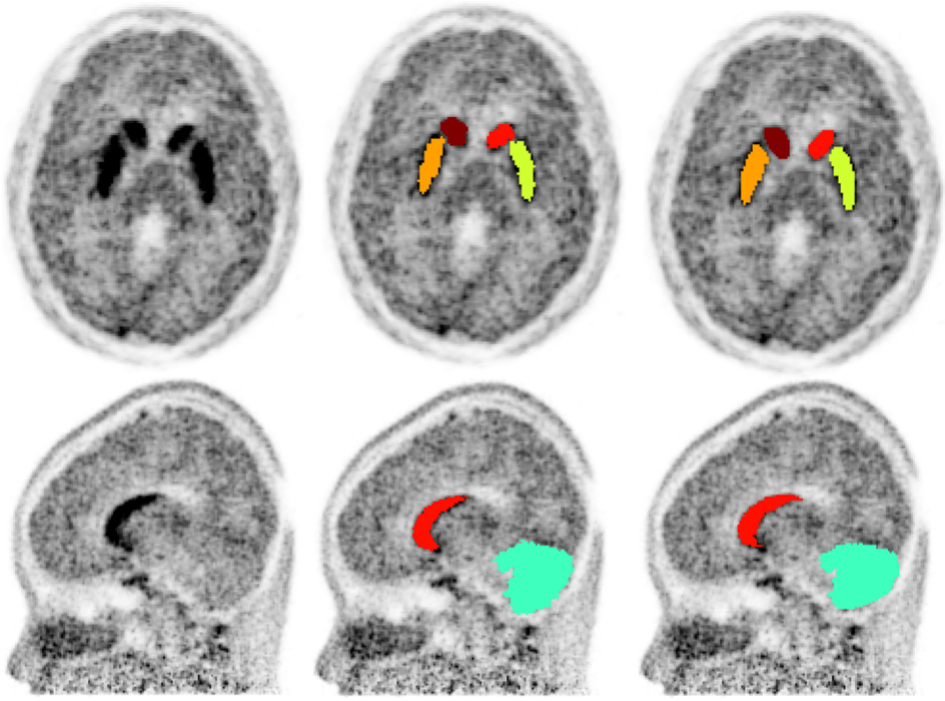


Figure 4–9: Visual comparison of atlas-based segmentation (b) with manual segmentation (c) of a typical real time-summed dynamic [^{11}C]raclopride images into (a) functional regions corresponding to the cerebellum, and left and right caudate and putamen.

The correlation between regional BP values derived from (i) atlas-based segmentations and (ii) manual segmentations is shown in figure 4–10. For reference, the correlation between regional BP estimates derived from the same two segmentation approaches are also shown for the simulated data for each of the five experiments.

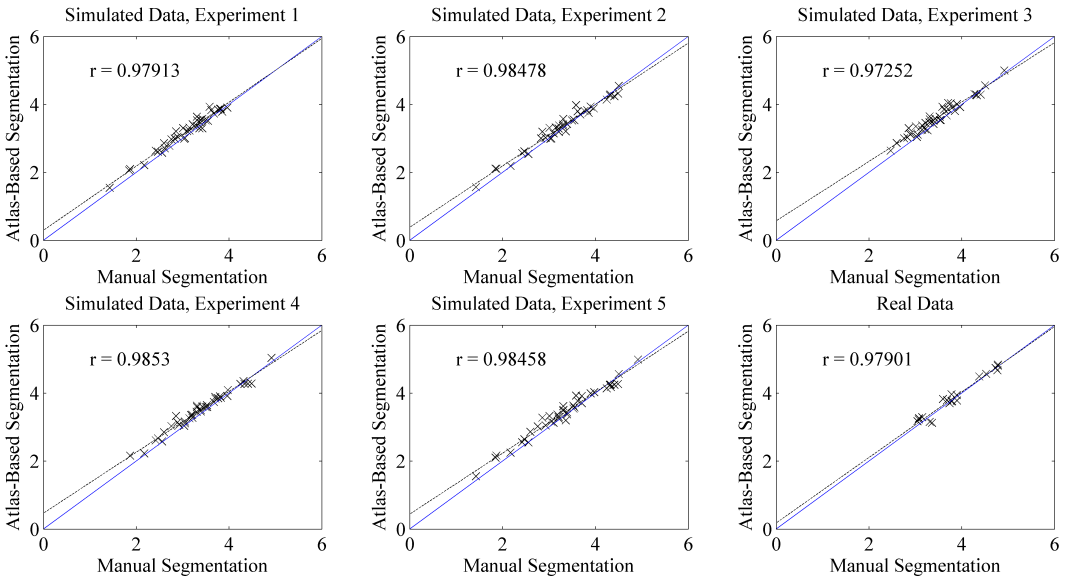


Figure 4–10: Correlation between regional BP values obtained using TACs derived from manual segmentations and atlas-based segmentations, for both simulated and real [^{11}C]raclopride data. Plots from simulated data each contain forty BP values (ten segmented images). The plot from real data contains twenty BP values (five segmented images). The solid blue line is the reference line of zero intercept and unity slope and the dashed black line is the linear line of best fit.

Correlations between regional BP values from manual segmentations and atlas-based segmentations are excellent for both simulated and real [^{11}C]raclopride data, with linear correlation coefficients r greater than or equal to 0.973 in all cases. Using real data, the linear correlation coefficient is comparable to that obtained using simulated data and shows the same trend (regional BP values from atlas-based segmentations tend to be greater than those obtained from manual segmentation), corroborating the results of the simulation study.

4.5 Discussion

This work proposes and validates the performance of a fully automated radiotracer-specific atlas-based segmentation and kinetic analysis approach using simulated and real [^{11}C]raclopride PET image data from the HRRT scanner. In this approach, segmentation is achieved by a fully automated one-step registration of a radiotracer-specific atlas with a given image, and the segmented regions are used to extract regional TACs for kinetic analysis using a reference tissue model. Importantly, the atlas template is an unbiased mean shape image, derived from diffeomorphic transformations, making it highly suitable for diffeomorphic registration to any given PET image. As an initial validation of the atlas-based image processing protocol, we applied it to a set of twenty simulated 3D dynamic [^{11}C]raclopride images, each with unique anatomy and unique dynamic functional behaviour. Considering all ROIs and all five experiments, the overall performance of atlas-based segmentation exceeded manual segmentation in terms of DSC (with respect to the ground truth regions) and $^{\%}Error$ (with respect to ground truth regional BPs) on regional BP estimates derived from the image segmentations. At the ROI level, atlas-based segmentation performed either comparably to, or out-performed manual segmentation in terms of both the DSC and $^{\%}Error$ metrics. A preliminary study on real [^{11}C]raclopride data using five dynamic images corroborated the simulation validation study, showing excellent correlation between regional BP values derived from TACs extracted from either manually defined or atlas-segmented ROIs.

These findings are consistent with our expectation that manual labeling of a high contrast average template is prone to less ambiguity than manual labeling of a single noisy image, and that the propagation of labels from the atlas to a given noisy image via co-registration is therefore a more reliable approach for image segmentation. Importantly, our proposed method overcomes two important limitations associated with manual segmentation: (i) a significant inter- and intra-operator dependency and (ii) significant time requirement. Concerning the first limitation, given a labeled template, our proposed atlas-based method yields fully deterministic segmentations and regional BP values. In addition, we expect that manual segmentation of a high contrast average template is prone to less user dependency than a noisy single image. Concerning the second limitation, manual segmentation of a typical 3D [^{11}C]raclopride image into functional regions corresponding to the cerebellum and left and right caudate and putamen required on average over an hour of time. On the other hand, our proposed method requires only a single manual segmentation of the average template, and reliable and accurate segmentation and kinetic analysis of a given image can then be performed in minutes, offering a drastic reduction in required time.

We noted a particularly accurate atlas-based segmentation of the cerebellum (mean DSC of 0.930 ± 0.007 using simulated data, compared to ground truth regions). These results indicate that atlas-segmented cerebellar regions can be used to automatically and reliably extract the reference region TAC of [^{11}C]raclopride images for a voxel-level parametric fit with a reference tissue kinetic model, which makes no assumptions about homogeneity of the functional regions. This is

particularly important since it has been shown that conventional cluster analysis methods for segmentation perform poorly when segmenting subtlety different functional regions (such as the grey matter and cerebellum for [^{11}C]raclopride images) [37]. In contrast, for cases in which the functional regions are known to be adequately homogeneous, region-based kinetic analysis offers benefits in the form of a reduced impact of noise, and atlas-based segmentation of these regions should prove beneficial.

One limitation of the current work is that validation was performed on images of healthy subjects. In the clinic, for instance, [^{11}C]raclopride is routinely used for the differentiation of Parkinsons disease from atypical parkinsonian symptoms [61]. [^{11}C]raclopride images of such patients are characterized by a pronounced increase (as in early Parkinsons disease), decrease, or lack of change in tracer uptake confined to the striatal regions [62]. We therefore expect that an intensity normalization procedure prior to registration between the radiotracer-specific atlas and a given image will suffice to ensure a reliable registration. Indeed, recent work using radiotracer-specific has shown that a radiotracer-specific template derived from healthy subjects is a suitable registration target for images derived from patients with suspected parkinsonian syndrome [47].

Finally, the proposed method can in principle be extended to work with PET images derived from other radiotracers; no assumptions about the specific distribution of the tracer was made in this work.

4.6 Conclusion

We have demonstrated that radiotracer-specific PET atlases can be used for fully automated segmentation and regional BP estimation of of high resolution dynamic [^{11}C]raclopride PET brain images, offering comparable performance to manual segmentation in terms of regional DSC, and improved performance in terms of the error on regional BP estimates. Atlas-based segmentation of PET brain images offers distinct advantages over conventional methods for image segmentation (such as cluster analysis and manual segmentation) by non-linear registration between a single PET image and a highly suitable PET atlas. Our proposed method requires no user intervention or anatomical images, making it a practical tool for automated analysis of PET images.

Appendix

Proof that a Poisson random variable conditioned on its sum is a multinomial distribution

Let $\mathbf{Z_i} = Z_{i1}, Z_{i2}, ..., Z_{iJ}$ be a hidden random vector from a Poisson distribution with parameters $\lambda_{i1}, \lambda_{i2}, ..., \lambda_{iJ}$. The joint probability mass function of the z_{ij} (a noisy realization of the random vector $\mathbf{Z_i}$) is given by (via Poisson statistics):

$$p_{\mathbf{Z_i}}(\mathbf{z_i}) = \prod_{j=1}^{J} e^{(-\lambda_{ij})} \frac{\lambda_{ij}^{z_{ij}}}{z_{ij}!} = e^{-\sum_j \lambda_{ij}} \frac{\lambda_{ij}^{\sum_j z_{ij}}}{z_{i1}! z_{i2}! \ ... \ z_{iJ}!} \tag{4.6}$$

Note that the sum m_i, being a sum of independent Poisson random variables, is itself a Poisson random variable. Let m_i be a noisy realization of Poisson random variable M_i with parameter $\bar{m}_i = \sum_j \lambda_{ij}$ such that $P(M_i = m_i) = \frac{e^{-\sum_j \lambda_{ij}} \left(\sum_j \lambda_{ij}\right)}{m_i!}$. Now we have:

$$P[(Z_{i1} = z_{i2}, ..., Z_{iJ} = z_{iJ}) \cap (M_i = m_i)] = \begin{cases} p_{\mathbf{Z_i}}(\mathbf{z_i}) & if \quad \sum_j z_{ij} = m_i \\ 0 & if \quad \sum_j z_{ij} \neq m_i \end{cases} \tag{4.7}$$

therefore the conditional distribution of the z_{ij} for a given i can be written:

$$\begin{aligned} &p_{\mathbf{Z_i}}(\mathbf{z_i}|M = m_i) = \frac{P[(Z_{i1}=z_{i2},...,Z_{iJ}=z_{iJ}) \cap (M_i=m_i)]}{P(M=m_i)} \\ &= \frac{m_i!}{z_{i1}! z_{i2}! \ ... \ z_{iJ}!} \left(\frac{\lambda_{ij}}{\sum_j \lambda_{ij}} \right)^{\sum_j \lambda_{ij}} \end{aligned} \tag{4.8}$$

which is the multinomial distribution with $N = m_i$ and $P_i = \left(\frac{\lambda_{ij}}{\sum_j \lambda_{ij}} \right)$.

REFERENCES

[1] C. S. Levin and E. J. Hoffman. Calculation of positron range and its effect on the fundamental limit of positron emission tomography system spatial resolution (vol 44, pg 781, 1999). *Physics in Medicine and Biology*, 45(2):559–559, 2000.

[2] A Rahmim, M Lenox, AJ Reader, C Michel, Z Burbar, TJ Ruth, and V Sossi. Statistical list-mode image reconstruction for the high resolution research tomograph. *Physics in medicine and biology*, 49(18):4239, 2004.

[3] James G Colsher. Fully-three-dimensional positron emission tomography. *Physics in medicine and biology*, 25(1):103, 1980.

[4] J. A. Fessler. Penalized weighted least-squares image reconstruction for positron emission tomography. *IEEE Trans Med Imaging*, 13(2):290–300, 1994.

[5] L. A. Shepp and Y. Vardi. Maximum likelihood reconstruction for emission tomography. *IEEE Trans Med Imaging*, 1(2):113–22, 1982.

[6] K. Wienhard, M. Dahlbom, L. Eriksson, C. Michel, T. Bruckbauer, U. Pietrzyk, and W. D. Heiss. The ecat exact hr: performance of a new high resolution positron scanner. *J Comput Assist Tomogr*, 18(1):110–8, 1994.

[7] F. H. van Velden, R. W. Kloet, B. N. van Berckel, F. L. Buijs, G. Luurtsema, A. A. Lammertsma, and R. Boellaard. Hrrt versus hr+ human brain pet studies: an interscanner test-retest study. *J Nucl Med*, 50(5):693–702, 2009.

[8] P. E. Kinahan and J. G. Rogers. Analytic 3d image reconstruction using all detected events. *Nuclear Science, IEEE Transactions on*, 36(1):964–968, 1989.

[9] TH Farquhar, A Chatziioannou, G Chinn, M Dahlbom, and EJ Hoffman. An investigation of filter choice for filtered back-projection reconstruction in pet. In *Nuclear Science Symposium, 1997. IEEE*, volume 2, pages 1042–1046. IEEE.

[10] Andrew J Reader and Habib Zaidi. Advances in pet image reconstruction. *PET Clinics*, 2(2):173–190, 2007.

[11] Roderick JA Little and Donald B Rubin. Statistical analysis with missing data. 2002.

[12] H Malcolm Hudson and Richard S Larkin. Accelerated image reconstruction using ordered subsets of projection data. *Medical Imaging, IEEE Transactions on*, 13(4):601–609, 1994.

[13] E. Veklerov and J. Llacer. Stopping rule for the mle algorithm based on statistical hypothesis testing. *IEEE Trans Med Imaging*, 6(4):313–9, 1987.

[14] J. Llacer, E. Veklerov, K. J. Coakley, E. J. Hoffman, and J. Nunez. Statistical analysis of maximum likelihood estimator images of human brain fdg pet studies. *IEEE Trans Med Imaging*, 12(2):215–31, 1993.

[15] T. Hebert and R. Leahy. A generalized em algorithm for 3-d bayesian reconstruction from poisson data using gibbs priors. *IEEE Trans Med Imaging*, 8(2):194–202, 1989.

[16] P. J. Green. Bayesian reconstructions from emission tomography data using a modified em algorithm. *IEEE Trans Med Imaging*, 9(1):84–93, 1990.

[17] S. Geman and D. Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Trans Pattern Anal Mach Intell*, 6(6):721–41, 1984.

[18] J. Tang and A. Rahmim. Bayesian pet image reconstruction incorporating anato-functional joint entropy. *Phys Med Biol*, 54(23):7063–75, 2009.

[19] J. Qi and R. M. Leahy. Iterative reconstruction techniques in emission computed tomography. *Phys Med Biol*, 51(15):R541–78, 2006.

[20] H. Ostertag, W. K. Kubler, J. Doll, and W. J. Lorenz. Measured attenuation correction methods. *Eur J Nucl Med*, 15(11):722–6, 1989.

[21] E. Kamel, T. F. Hany, C. Burger, V. Treyer, A. H. Lonn, G. K. von Schulthess, and A. Buck. Ct vs 68ge attenuation correction in a combined pet/ct system: evaluation of the effect of lowering the ct tube current. *Eur J Nucl Med Mol Imaging*, 29(3):346–50, 2002.

[22] M. Bentourkia, P. Msaki, J. Cadorette, and R. Lecomte. Assessment of scatter components in high-resolution pet: correction by nonstationary convolution subtraction. *J Nucl Med*, 36(1):121–30, 1995.

[23] S. R. Cherry, S. R. Meikle, and E. J. Hoffman. Correction and characterization of scattered events in 3-dimensional pet using scanners with retractable septa. *Journal of Nuclear Medicine*, 34(4):671–678, 1993.

[24] S. Grootoonk, T. J. Spinks, D. Sashin, N. M. Spyrou, and T. Jones. Correction for scatter in 3d brain pet using a dual energy window method. *Physics in Medicine and Biology*, 41(12):2757–2774, 1996.

[25] C. C. Watson, D. Newport, and M. E. Casey. A single scatter simulation technique for scatter correction in 3d pet. *Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, 4:255–268, 1996.

[26] C. S. Levin, M. Dahlbom, and E. J. Hoffman. A monte-carlo correction for the effect of compton-scattering in 3-d pet brain imaging. *Ieee Transactions on Nuclear Science*, 42(4):1181–1185, 1995.

[27] S. R. Cherry and S. C. Huang. Effects of scatter on model parameter estimates in 3d pet studies of the human brain. *Ieee Transactions on Nuclear Science*, 42(4):1174–1179, 1995.

[28] P. E. Kinahan, D. W. Townsend, D. L. Bailey, D. Sashin, F. Jadali, and M. A. Mintun. Efficiency normalization techniques for 3d pet data. *1995 Ieee Nuclear Science Symposium and Medical Imaging Conference Record, Vols 1-3*, pages 1021–1025, 1996.

[29] E. J. Hoffman, T. M. Guerrero, G. Germano, W. M. Digby, and M. Dahlbom. Pet system calibrations and corrections for quantitative and spatially accurate images. *Ieee Transactions on Nuclear Science*, 36(1):1108–1112, 1989.

[30] Andrew J Reader, Peter J Julyan, Heather Williams, David L Hastings, and Jamal Zweit. Em algorithm system modeling by image-space techniques for pet reconstruction. *Nuclear Science, IEEE Transactions on*, 50(5):1392–1397, 2003.

[31] Richard E Carson. Tracer kinetic modeling in pet. *Positron Emission Tomography: Basic Sciences*, pages 127–159, 2005.

[32] 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):279–87, 1997.

[33] J. A. Hartigan and M. A. Wong. Algorithm as 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 28(1):100–108, 1979.

[34] K. P. Wong, D. G. Feng, S. R. Meikle, and M. J. Fulham. Segmentation of dynamic pet images using cluster analysis. *Ieee Transactions on Nuclear Science*, 49(1):200–207, 2002.

[35] James C Bezdek. *Pattern recognition with fuzzy objective function algorithms*. Kluwer Academic Publishers, 1981.

[36] S. Belhassen and H. Zaidi. A novel fuzzy c-means algorithm for unsupervised heterogeneous tumor quantification in pet. *Medical Physics*, 37(3):1309–1324, 2010.

[37] Heidi Koivistoinen, Jussi Tohka, and Ulla Ruotsalainen. Comparison of pattern classification methods in segmentation of dynamic pet brain images. In *Proc. of Sixth Nordic Signal Processing Symposium, NORSIG*, pages 73–76.

[38] Douglas A Reynolds. Gaussian mixture models, 2009.

[39] Jun L Chen, Steve R Gunn, Mark S Nixon, and Roger N Gunn. Markov random field models for segmentation of pet images. In *information processing in medical imaging*, pages 468–474. Springer.

[40] W. R. Crum, T. Hartkens, and D. L. G. Hill. Non-rigid image registration: theory and practice. *British Journal of Radiology*, 77:S140–S153, 2004.

[41] Ake Bjrck. *Numerical methods for least squares problems*. Siam, 1996.

[42] T. Vercauteren, X. Pennec, A. Perchant, and N. Ayache. Diffeomorphic demons: efficient non-parametric image registration. *Neuroimage*, 45(1 Suppl):S61–72, 2009.

[43] V. Arsigny, O. Commowick, X. Pennec, and N. Ayache. A log-euclidean framework for statistics on diffeomorphisms. *Medical Image Computing and Computer-Assisted Intervention - Miccai 2006, Pt 1*, 4190:924–931, 2006.

[44] M. Hernandez, M. N. Bossa, and S. Olmos. Registration of anatomical images using geodesic paths of diffeomorphisms parameterized with stationary vector fields. *2007 Ieee 11th International Conference on Computer Vision, Vols 1-6*, pages 2322–2329, 2007.

[45] L. Farde, L. Eriksson, G. Blomquist, and C. Halldin. Kinetic-analysis of central [c-11] raclopride binding to d2-dopamine receptors studied by pet - a comparison to the equilibrium-analysis. *Journal of Cerebral Blood Flow and Metabolism*, 9(5):696–708, 1989.

[46] H. Ito, F. Kodaka, H. Takahashi, H. Takano, R. Arakawa, H. Shimada, and T. Suhara. Relation between presynaptic and postsynaptic dopaminergic functions measured by positron emission tomography: Implication of dopaminergic tone. *Journal of Neuroscience*, 31(21):7886–7890, 2011.

[47] Felix P Kuhn, Geoffrey I Warnock, Cyrill Burger, Katharina Ledermann, Chantal Martin-Soelch, and Alfred Buck. Comparison of pet template-based and mri-based image processing in the quantitative analysis of c11-raclopride pet. *EJNMMI research*, 4(1):1–7, 2014.

[48] J Ashburner, J Haslan, C Taylor, V Cunningham, and T Jones. A cluster analysis approach for the characterization of dynamic pet data. *Journal of Cerebral Blood Flow and Metabolism*, 15(1):S626, 1995.

[49] P Edison, SF Carter, JO Rinne, G Gelosa, K Herholz, A Nordberg, DJ Brooks, and R Hinz. Comparison of mri based and pet template based approaches in the quantitative analysis of amyloid imaging with pib-pet. *Neuroimage*, 70:423–433, 2013.

[50] Sarang Joshi, Brad Davis, Matthieu Jomier, and Guido Gerig. Unbiased diffeomorphic atlas construction for computational anatomy. *NeuroImage*, 23:S151–S160, 2004.

[51] A. A. Lammertsma and S. P. Hume. Simplified reference tissue model for pet receptor studies. *Neuroimage*, 4(3):153–158, 1996. 1 Wc002 Times Cited:920 Cited References Count:18.

[52] J. G. Sled, A. P. Zijdenbos, and A. C. Evans. A nonparametric method for automatic correction of intensity nonuniformity in mri data. *IEEE Trans Med Imaging*, 17(1):87–97, 1998.

[53] J. Mazziotta, A. Toga, A. Evans, P. Fox, J. Lancaster, K. Zilles, R. Woods, T. Paus, G. Simpson, B. Pike, C. Holmes, L. Collins, P. Thompson, D. MacDonald, M. Iacoboni, T. Schormann, K. Amunts, N. Palomero-Gallagher, S. Geyer, L. Parsons, K. Narr, N. Kabani, G. Le Goualher, D. Boomsma, T. Cannon, R. Kawashima, and B. Mazoyer. A probabilistic atlas and reference system for the human brain: International consortium for brain mapping (icbm). *Philos Trans R Soc Lond B Biol Sci*, 356(1412):1293–322, 2001.

[54] S. M. Smith. Fast robust automated brain extraction. *Hum Brain Mapp*, 17(3):143–55, 2002.

[55] D. L. Collins, P. Neelin, T. M. Peters, and A. C. Evans. Automatic 3d intersubject registration of mr volumetric data in standardized talairach space. *Journal of Computer Assisted Tomography*, 18(2):192–205, 1994.

[56] A. Rahmim, J. C. Cheng, K. Dinelle, M. Shilov, W. P. Segars, O. G. Rousset, B. M. Tsui, D. F. Wong, and V. Sossi. System matrix modelling of externally tracked motion. *Nucl Med Commun*, 29(6):574–81, 2008.

[57] F. C. Sureau, A. J. Reader, C. Comtat, C. Leroy, M. J. Ribeiro, I. Buvat, and R. Trebossen. Impact of image-space resolution modeling for studies with the high-resolution research tomograph. *J Nucl Med*, 49(6):1000–8, 2008.

[58] C Comtat, F Bataille, C Michel, JP Jones, M Sibomana, L Janeiro, and R Trebossen. Osem-3d reconstruction strategies for the ecat hrrt. In *Nuclear Science Symposium Conference Record, 2004 IEEE*, volume 6, pages 3492–3496. IEEE.

[59] H. Lombaert, J. M. Peyrat, P. Croisille, S. Rapacchi, L. Fanton, F. Cheriet, P. Clarysse, I. Magnin, H. Delingette, and N. Ayache. Human atlas of the cardiac fiber architecture: study on a healthy population. *IEEE Trans Med Imaging*, 31(7):1436–47, 2012.

[60] Alexandre Guimond, Jean Meunier, and Jean-Philippe Thirion. *Automatic computation of average brain models*, pages 631–640. Springer, 1998.

[61] V. Kaasinen, H. M. Ruottinen, K. Nagren, P. Lehikoinen, V. Oikonen, and J. O. Rinne. Upregulation of putaminal dopamine d2 receptors in early parkinson's disease: a comparative pet study with [11c] raclopride and [11c]n-methylspiperone. *J Nucl Med*, 41(1):65–70, 2000.

[62] K. Van Laere, K. Clerinx, E. D'Hondt, T. de Groot, and W. Vandenberghe. Combined striatal binding and cerebral influx analysis of dynamic 11c-raclopride pet improves early differentiation between multiple-system atrophy and parkinson disease. *J Nucl Med*, 51(4):588–95, 2010.