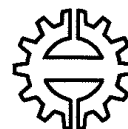


Tampereen teknillinen korkeakoulu
Julkaisu 265

Tampere University of Technology
Publications 265



Sakari Alenius

On Noise Reduction in Iterative Image Reconstruction Algorithms for Emission Tomography: Median Root Prior

Thesis for the degree of Doctor of Technology to be presented with due permission
for public examination and criticism in Auditorium HB116, Hermitech at Tampere
University of Technology, on the 9th of October 1999, at 12 o'clock noon.

Tampere 1999

ISBN 978-952-15-0279-7 (printed)
ISBN 978-952-15-4216-9 (PDF)
ISSN 0356-4940

Abstract

In this study, a simple method was developed for reducing noise during the iterative reconstruction of emission tomography images.

The emission images express the spatial distribution of a chemical compound, if possible in quantitative terms. The concentration of the tracer gives information about the active metabolism of a living tissue. As a computational medical imaging method, the emission tomography involves a significant amount of data processing. The algorithms that carry out the computation of the estimate of the transaxial image from the measured data suffer from noise due to the statistical nature of the acquisition.

The new method (MRP, median root prior) is novel and generally applicable for emission reconstruction, independent of the organ or the tracer. The principle of penalizing locally non-monotonic noise is suitable for both emission and transmission image reconstruction. Instead of explicitly producing visually pleasing images, the method is designed to compute locally accurate images with no constraints on pixel value differences.

In practice, MRP is quantitatively accurate and simple in implementation. The transmission reconstruction with this method allows for a short-time attenuation correction. As a result of this, the scanning time for the patient can be reduced.

Preface

This work has been done in the Signal Processing Laboratory of Tampere University of Technology in a close cooperation with Turku PET Centre, Tampere University Hospital, and Kuopio University Hospital.

I thank my supervisor Prof. Jaakko Astola for valuable advice. I owe my deepest gratitude especially to Ph.D. Ulla Ruotsalainen for leading me into the world of PET and for pointing out the essential point of view.

I thank the preliminary assessors of this dissertation, Prof. Johan Nuyts and Assoc. Prof. Jeffrey Fessler, for their thorough work.

I express my sincere thanks to the personnel of Turku PET Centre for providing an exiting environment for both business and pleasure. I thank Prof. Juhani Knuuti, Ph.D. Chietsugu Katoh, Ph.D. Tapio Utriainen, and M.D. Meri Kähkönen for linking my work to the medical field, where I am a mere layman. Thanks to Lic.Tech. Päivi Laarne, M.Sc. Vesa Oikonen, and Mr. Mika Klemetsö other people can use some of my programs in their work. I am thankful to the physicists M.Sc. Mika Teräs, M.Sc. Tuula Tolvanen, and B.Sc. Petri Numminen for carrying out important tasks when implementing MRP.

I am grateful to Ph.D. Matti Koskinen and Ph.D. Antti Virjo in TAUH and Lic.Sc. Tomi Kauppinen in KUH for broadening the scope of the application from PET to SPECT by working hard with a professional attitude.

I thank my fellow workers in TUT/DMI/SPAG/MFGI for an enjoyable and relaxed atmosphere I have been privileged to share. I extend my thanks also to the personnel of Tampere Image Processing Systems for bringing me into the professional life during the early years of my career.

The financial support provided by the City of Tampere, the Ulla Tuominen Foundation, the Research Foundation of Instrumentarium, Tampere Graduate School in Information Science and Engineering, and the Research Scholarship Found of Tampere (Tamperealaisen tutkimustyön tukisäätiö) is gratefully acknowledged.

I thank my closest family for support and especially my nieces Marika and Pauliina for giving delight to my life. The other Pauliina deserves my warmest thanks for moments not related to this thesis.

Tampere, August 1999
Sakari Alenius

Contents

Abstract	i
Preface	iii
List of publications	ix
Symbols and abbreviations	xi
1 Introduction	1
1.1 Objectives of emission tomography	1
1.2 Problems in the image reconstruction	3
1.3 Structure of the work	4
2 Data Acquisition	5
2.1 PET emission	5
2.2 PET transmission	6
2.2.1 Attenuation correction	7
2.3 SPECT	8
2.4 Noise and errors in data	10
3 Image Reconstruction	11
3.1 Filtered back projection	11
3.1.1 The Radon transform and its inverse	11
3.1.2 Applicability to PET	14
3.2 On iterative reconstruction algorithms	16
3.2.1 Notations	16
3.2.2 General concept	17
3.3 Maximum likelihood	18
3.3.1 MLEM algorithm	18
3.3.2 Noise and bias in iterations	21
3.4 Bayes priors and penalties	22
3.4.1 Gibbs priors	23

3.4.2	One step late algorithm	24
3.4.3	Mathematical priors	25
3.4.4	Anatomical priors	28
3.5	Implementational aspects	29
3.5.1	Acceleration techniques	29
3.5.2	Corrections	30
3.6	Performance measures of estimators	30
3.7	Summary	31
4	Aims of the work	33
5	Median Root Prior	35
5.1	General description	35
5.2	MRP Algorithm	36
5.2.1	Emission OSL algorithm	36
5.2.2	Interpretations on the prior	36
5.3	Transmission algorithm	38
5.3.1	Motivation	38
5.3.2	Data fitting algorithm	39
5.3.3	tMRP algorithm	40
5.4	On convergence properties	41
5.5	Noise and bias	41
5.6	Implementation	41
5.6.1	Neighborhood size	41
5.6.2	Parameter β	42
5.6.3	Iterations	42
6	Results and Images	45
6.1	Simulations	45
6.1.1	Error histograms	46
6.1.2	Reliability of ROI estimation	48
6.1.3	Noise sensitivity	49
6.1.4	Parameter sensitivity	49
6.1.5	Images	52
6.2	Clinical data	53
6.2.1	Emission studies	53
6.2.2	Transmission	55
6.2.3	Whole body studies	56
6.3	SPECT images	57
6.4	Summary of results and the publications	58

<i>CONTENTS</i>	vii
-----------------	-----

7 Discussion	61
---------------------	-----------

7.1 General principle	61
---------------------------------	----

7.2 Applicability of MRP	63
------------------------------------	----

7.3 Conclusion	64
--------------------------	----

Bibliography	65
---------------------	-----------

Publications	75
---------------------	-----------

List of publications

- I : S. Alenius. Bayesian Image Reconstruction in Positron Emission Tomography using Median Root Prior. *The Proceedings of the First TUT Symposium on Signal Processing '94*, pp. 113-116, 1994, Tampere, Finland.
- II : S. Alenius, U. Ruotsalainen. Bayesian image reconstruction for emission tomography based on median root prior. *European Journal of Nuclear Medicine*, vol. 24, pp. 258-265, Mar 1997.
- III : S. Alenius, U. Ruotsalainen, J. Astola. Using local median as the location of the prior distribution in iterative emission tomography image reconstruction. *IEEE Transactions on Nuclear Science*, selected papers from the 1997 Medical Imaging Conference (MIC), Albuquerque, New Mexico, November 13-15 1997, vol. 45, no. 6(2), pp. 3097-3107, Dec 1998.
- IV : S. Alenius, U. Ruotsalainen, and J. Astola. Attenuation Correction for PET Using Count-Limited Transmission Images Reconstructed with Median Root Prior. *IEEE Transactions on Nuclear Science*, selected papers from the 1998 Medical Imaging Conference (MIC), Toronto, Canada, November 11-14 1998, vol. 46, no. 3(2), pp. 646-651, June 1999.

Symbols and abbreviations

β	Bayes weight
C	constant, terms independent of the variable in question
$\delta()$	Dirac delta function
$\mathcal{F}_n\{\}$	n-dimensional Fourier transform
λ	emission image
μ	transmission image
$\mathcal{R}\{\}$	Radon transform
ACF	attenuation correction factor
ART	algebraic reconstruction technique
Bq	Bequerel, unit of activity, one disintegration per second (1/s)
CT	computerized (computed) tomography
ET	emission tomography
eV	electron volt, unit of energy
FBP	filtered back projection
FDG	fluoro-2-deoxy-D-glucose
FFT	fast Fourier transform
FIR	finite impulse response
FOV	field of view
FT	Fourier transform
LOR	line of response
MAE	mean absolute error
MAP	maximum <i>a posteriori</i>
ML	maximum likelihood
MLEM	maximum likelihood expectation maximization
MRF	Markov random field
MRI	magnetic resonance imaging
MRP	median root prior
MSE	mean squared error
OSL	one step late
PDF	probability density function
PET	positron emission tomography

PSF	point spread function
ROI	region of interest
RT	Radon transform
SPECT	single photon emission computed tomography, a.k.a. SPET
SPET	single photon emission tomography, a.k.a. SPECT
tMRP	transmission MRP

Chapter 1

Introduction

1.1 Objectives of emission tomography

Tomography (Greek *tomos* = section, + -graphy) means a method of producing an image of the internal structures of a solid object by the observation and recording of the differences in the effects on the passage of waves of energy impinging on those structures [64]. The medical applications of tomography utilize the non-invasive nature of the procedure, which makes it possible to examine living objects *in vivo*. Different tomographic imaging modalities provide different kind of information depending on the data recording means. The traditional X-ray based computerized tomography (CT) produces images of photon attenuation in the tissue, and magnetic resonance imaging (MRI) describes the proton or water density [45]. These images are generally anatomical because they reveal the physical structure of the tissue, serving as an "electronic knife". Such properties as the density of the tissue, the water density, and the membrane structure are examples of objects of interest in the anatomical imaging.

Emission tomography (ET) is intended to express functional properties of the tissue. Emission refers to the fact that the energy source is not external but it is brought into the tissue as a part of the body metabolism. This is done by a tracer which is a chemical compound labeled with a radioactive isotope. The spatial distribution of the radiating source is the object of interest as it tells the concentration of the tracer in the tissue. As the labeled compound is chemically identical to its non-labeled counterpart, an image of the spatial source distribution shows information about the metabolism relating to the compound during the data acquisition period. Thus, functional ET images are generally different from anatomical images. Primarily, ET images do not express what the tissue looks like, but how it functions. Although functional and anatomical images of the same organ have some similarities, such as boundaries between tissue types (see Fig. 1.1), the essential information in an ET image is where it differs from

the corresponding anatomical image. An ET image can be used when studying the normal behavior of the tissue, or when diagnosing abnormal cases in clinical work. The main areas of medicine that utilize ET are neurology, cardiology, and oncology.

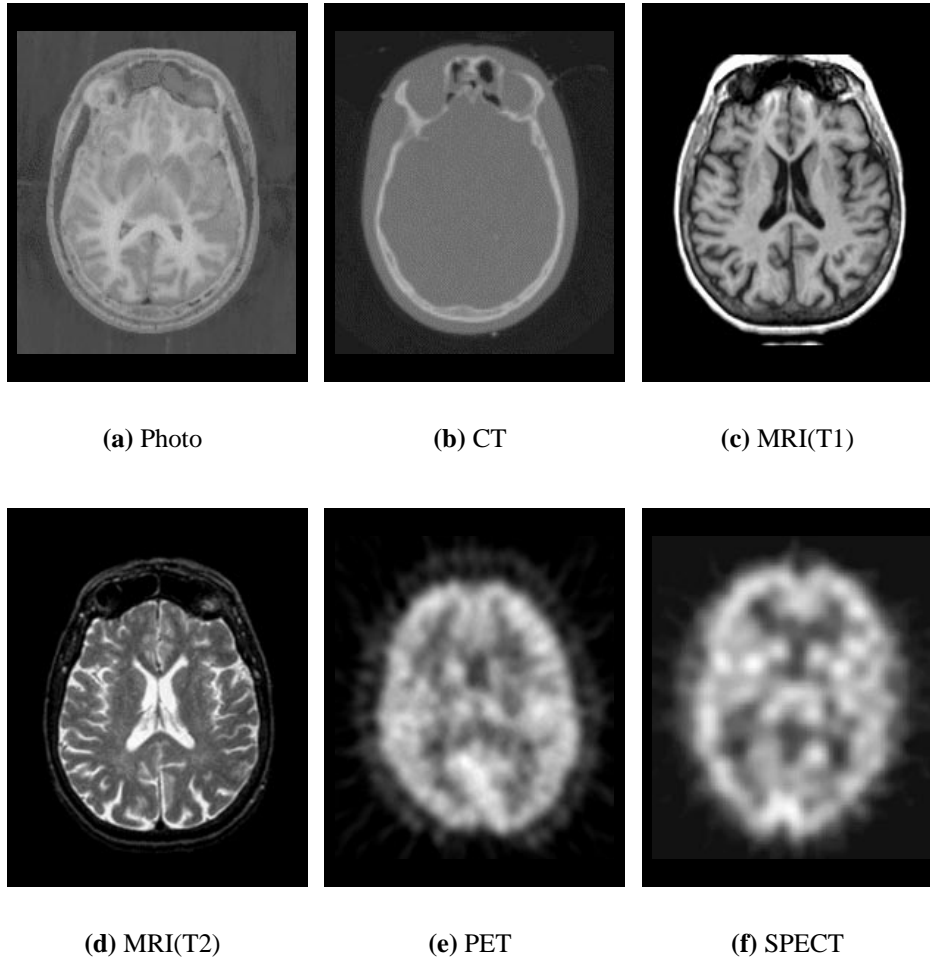


Figure 1.1: Examples of various imaging modalities. Images are not of the same person. Sources: [12], [43], Turku PET Centre, and Kuopio University Hospital.

Tomographic images as slices are originally two-dimensional, but they can be combined to a 3D image. What is more, a time sequence of ET images form a dynamic 3D or 4D image. The dimension of time is medically important because the concentration of the tracer in the tissue changes in time. A parametric image computed from a dynamic image sequence depicts values of a certain metabolic rate as an image [73].

There are two ET modalities that share some common properties. Positron

emission tomography (PET) is capable to quantitative measurements of tracer radioactivity concentration. Simpler and cheaper single photon emission computed tomography (SPECT, a.k.a. SPET) has traditionally been used for relative measurements.

1.2 Problems in the image reconstruction

The measured data in tomography are not an image of the object, but projections of it, a sinogram. The unknown image has to be estimated from the available data computationally. This task is called the image reconstruction from projections [37]. Its goal is a reasonably accurate image with a reasonably low noise level. The ET data acquisition is subject to a substantial amount of statistical noise, originating from the statistical nature of the decay of the label isotope. Fig. 1.2(a) shows a PET sinogram. The noise in the sinogram contributes to reconstruction artifacts visible in the reconstructed image of the object 1.2(b).

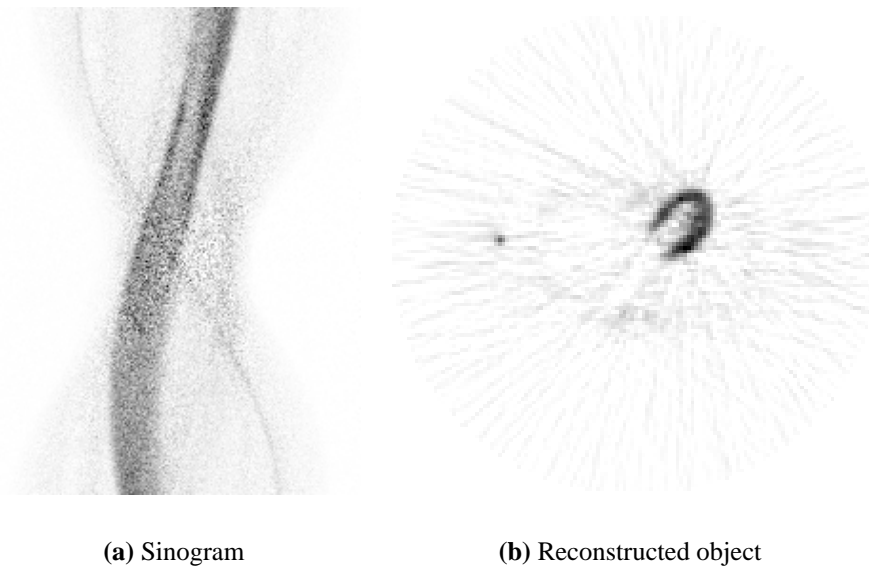


Figure 1.2: A PET sinogram and the reconstructed image. Both are shown as negative pictures.

The accuracy of the image reconstruction is the basis of ET studies. The region of interest (ROI) analysis is based on determining the activity in the tissue averaged over selected set of image pixels. Noise and reconstruction artifacts deteriorate the reliability of the ROI analysis in two ways. First, the ROI is difficult to draw on an image with reconstruction artifacts and noise. Secondly, they may cause some local bias to the quantitative values of ROI pixels. Because the

dose given to the patient should be fairly low and the acquisition time should be relatively short, the noise level of the acquired data is unavoidably quite high. Thus, there is a need for improvement of the quality of the reconstructed images.

Image enhancement and filtering methods that are convenient in general purpose image processing may not serve as reliable noise reduction means in ET. Particularly the quantitative accuracy of PET sets special requirements to image improvement operations. What is more, the image reconstruction process itself should take advantage of the special nature of image formation process. The system's sensitivity in distinguishing between two different activity levels should not be compromised by the noise reduction operation. Also, the resolution of how well two neighboring areas of different activity can be separated spatially should be maintained.

1.3 Structure of the work

The main result of this work is a practical way to remove a significant amount of noise during the iterative reconstruction process without causing substantial bias and without removing relevant information. The developed method is general and simple, it applies to various iterative reconstruction algorithms, and it can be used independently of the organ under the study. The main focus is in PET.

Chapter 2 addresses the data acquisition process. The principles of the emission detection are discussed. The basis of quantitative measurements, the attenuation correction, is shortly described. The differences between PET and SPECT and the noise contamination in the data are discussed, too.

Chapter 3 discusses two main classes of image reconstruction methods, analytical and statistical algorithms. Various approaches for noise reduction of the iterative algorithms are reviewed.

Chapter 4 presents the main goals of the work.

Chapter 5 describes the main contribution of the author, the MRP algorithm. MRP is used in conjunction with an iterative algorithm as a regularizer or a penalty.

Chapter 6 reports results and experiments of MRP reconstructions applied both on simulated and on clinical data.

Chapter 7 draws conclusions about the relevance of MRP in the field ET image reconstruction.

Chapter 2

Data Acquisition

2.1 PET emission

In PET, common labels are ^{11}C , ^{13}N , ^{15}O , and ^{18}F isotopes [80]. They can be attached to medically interesting molecules by radiopharmaceutical processes. The tracer can be introduced into the patient by injection or by inhalation. The nucleus of the label is unstable and it transforms into a stable isotope by emitting a positron. The positron and the electron are anti-particles. Their masses annihilate into two photons of 511 keV according to the equation $E = mc^2$ [7]. That energy is enough for escaping the body. Because the photons travel at almost opposing directions, the data acquisition is accomplished by detecting coincident events by a detector ring surrounding the patient (Fig. 2.1). Each event is recorded related to a line of response (LOR), a line along which the annihilation took place. The area overlapped and enclosed by the LORs is the field of view (FOV). The angles and the positions of the LORs are saved as 2D tables, sinograms. Each row of the sinogram corresponds to a projection profile at certain angle, and the bins of the row correspond to the detector pairs on that direction. A point source generates half of a sine wave into the sinogram, as visible in Fig. 1.2 on page 3. After the measurement interval, each sinogram bin contains the number of detected events (counts) for its LOR.

The reconstructed image is a temporal average of the spatial tracer distribution on the transaxial image plane over the data acquisition period. The unit of quantitative images is commonly kBq/ml. Quantitative accuracy can be achieved by elimination the effect of the body attenuation and scatter. Even though transmittance properties are not the main interest in PET, the transmission measurements are carried out in order to get quantitative images.

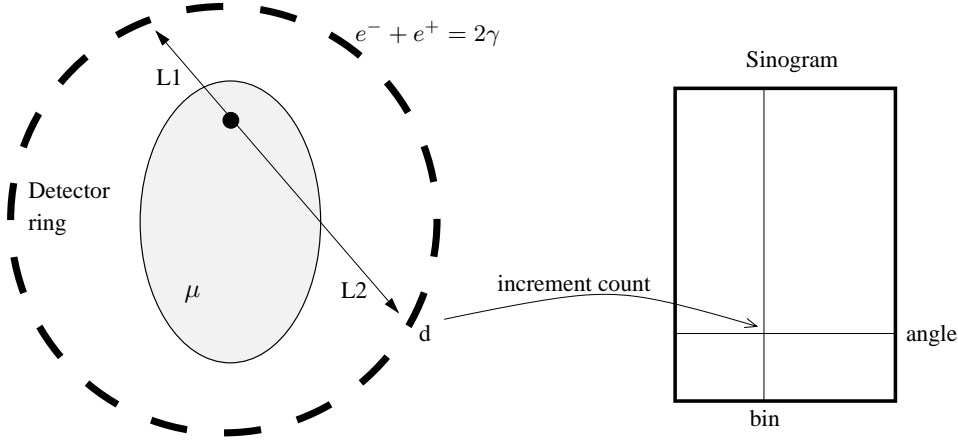


Figure 2.1: Photon acquisition in PET. The detectors work in pairs registering coincident γ -photons. The sinogram contains the registered counts. The attenuation effect is independent of the position of the annihilation.

2.2 PET transmission

The emitted photon may interact with the matter and scatter from its original fly path. If that happens, no true coincident event occurs. Thus, not all annihilations are detected as the body attenuates the photon flux. The photon attenuation has to be taken into consideration in order to find out the spatial emission source distribution quantitatively. Also, an image of the transmittance properties of the tissue may be useful in some studies [11].

The probability that both of the emitted photons travel without scattering to the detectors is [45]

$$P = P_1 P_2 = e^{-\int_{L1_d} \mu(x) dx} e^{-\int_{L2_d} \mu(x) dx} = e^{-\int_{L_d} \mu(x) dx} \triangleq e^{-a_d}, \quad (2.1)$$

where μ is the 2D function of linear mass absorption coefficients, L_d is the path corresponding to LOR for detector pair d , and $L1_d$ and $L2_d$ are the sub-paths from the place of annihilation to detectors ($L_d = L1_d + L2_d$, Fig. 2.1). Because of the coincidence detection, the attenuation effect on d is independent of the position of the annihilation along the LOR. This is why an external source can be used when measuring the attenuation of the body.

2.2.1 Attenuation correction

Measured attenuation correction

In order to determine the attenuation of the body, two additional measurements are done: the blank scan and the transmission scan [41]. The blank scan is recorded using an external source without the patient. This represents the unattenuated case. For the transmission scan the patient and the bed are placed into the scanner and the attenuated data are measured using the external source. The external source can be ring-shaped on a fixed position, or a rotating line or point source [11].

The attenuation correction factor (ACF) can be calculated as the ratio of the measured counts without and with the attenuating object

$$ACF_d = \frac{B_d e^0}{B_d e^{-a_d}} = e^{a_d}, \quad (2.2)$$

where B_d is the measured activity of the external source. The ACFs form a sinogram, whose elements express the amount of attenuation for the detector pair d . Figure 2.2 shows the two scans and the resulting ACFs. Attenuation pre-correction is carried out as the multiplication of the emission sinogram count by the corresponding ACF. The ACF in Eq. (2.2) is intended to cancel out the attenuation expressed by Eq. (2.1).

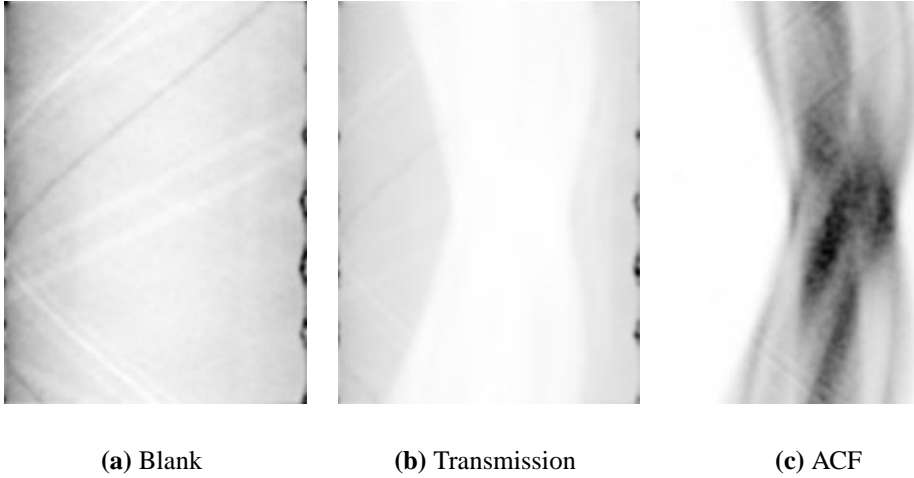


Figure 2.2: The smoothed blank scan, the smoothed transmission scan, and the attenuation correction factors according to (2.2) shown as negative pictures.

Short transmission scan times increase the noise level and worsen the quality of ACFs. Due to noise, the scan data are usually smoothed prior to the

division. Otherwise the noise in the ACFs propagates to the corrected emission sinogram. The drawback of smoothing is that the resulting blurring of ACFs propagates to the emission sinogram as well [19].

The measurement of the transmission data can be done before the emission acquisition, which increases the time spent in clinical routines. In order to reduce the patient discomfort, the transmission acquisition can be done simultaneously to the emission scan by utilizing a dedicated hardware design. Consequently, the cross contamination of the transmission data and the emission data must be taken care of [11].

Computational attenuation correction

If the transmission acquisition must be avoided, attenuation correction can be performed by computational methods [41]. The attenuating matter is approximated by an area with a uniform value of linear mass absorption coefficient μ . The ACFs can be computed by projecting the graphically defined μ -image and taking the exponential

$$ACF_d = e^{\int_{L_d} \mu(x) dx} . \quad (2.3)$$

The assumption that the μ -values are uniform and the manual work usually required restrict the usage of (2.3) [41].

Hybrid attenuation correction

The hybrid methods for attenuation correction combine the transmission measurements and the computational method [41]. First, the μ -image is reconstructed from the natural logarithm of the ACFs (2.2). Then this transmission image is segmented to assumed tissue types. The pixel values in the segmented areas are assigned with the known average μ -values and the resulting image is projected to ACFs (2.3). The segmentation step reduces the noise in the transmission from propagating to the emission scan when the emission sinogram are multiplied with the ACFs. The assumptions that there is a known number of tissue types and that they are uniform limit the applicability of the hybrid method with organs with diverse set of μ -values, as lungs [77].

2.3 SPECT

SPECT devices register only single photons usually without any coincidence detection [42]. This allows for detection of photons of various energy levels, not only those produced by the positron annihilation. Commonly used labels

are ^{99m}Tc and ^{201}Tl . In the absence of the coincidence detection, a physical collimation is required in order to restrict the possible ray path of photons entering the detector. Usually the detector heads (one, two or three) can be rotated mechanically around the patient. Each head contains a matrix of detectors such that the head records an image at its rotational position. These images can be used as planar images without the reconstruction. The counts along a transversal line of the acquisition images form the sinogram, as shown in Fig. 2.3. The reconstruction of the sinogram gives the image of the selected transaxial slice.

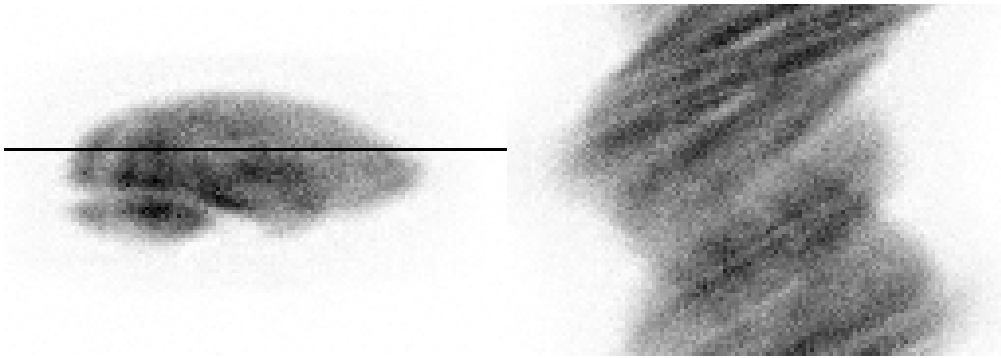


Figure 2.3: An emission SPECT study of a brain. **Left:** The acquisition image. **Right:** The sinogram with full 360° rotation. The projections 180° apart can be combined before the reconstruction.

Often there are no attenuation correction used in SPECT. Without the coincidence detection the attenuation depends on the depth of the source location in the object. Each detector collects photons originating from various depths and the total count cannot be compensated by a single factor, even if the opposing counts are combined [79]. The measured attenuation correction in SPECT can be accomplished in a similar way than in PET, but the pre-correction (2.2) is only approximate [41, 11]. For quantitative results, the sketched or reconstructed μ -image can be used for computing the attenuated projections inside the emission reconstruction algorithm [11].

The lack of the attenuation correction makes SPECT studies commonly non-quantitative, and the visual quality of the reconstructed image is more important. Sometimes pre-filtering is applied on the acquired data or post-filtering on the reconstructed images.

2.4 Noise and errors in data

The data acquisition process in ET measures the product of the radioactive decay, which is a random process. The emission of a positron is a rare event, but in a large population (high number of atoms) the number of such events is distributed according to Poisson distribution [70].

An underlying activity generates a random number of disintegrations per a time interval. The realized number of events (counts), n , is drawn from a Poisson distribution with the mean equal to the underlying activity λ [76], $Poisson(\lambda)$

$$P\{n = \lambda\} = e^{-\lambda} \frac{\lambda^n}{n!} . \quad (2.4)$$

Technically speaking, the main interest in ET is to estimate the unknown mean activity λ .

In addition to the statistical noise, the data acquisition system results in some errors in the data because of the finite energy and time resolution of the detectors. If the photon is *scattered*, it can still be accepted by a detector, which results a LOR with no real annihilation. Also, almost simultaneous single photons can be accepted as a pair, which accounts for an *accidental* coincident. The detectors have individual gains, which must be *normalized*. There is a *dead time* after each event before the detector is able to record again. Photons arriving during that time are lost. The positron travels a few millimeters in the tissue losing some energy before it can interact with an electron. The distance between the place of the annihilation and the place of the decay is the *positron range*, which contributes to the point spread function (PSF) of the device. These errors can be partly compensated by approximate statistical methods prior to or during the reconstruction. However, the Poisson nature of the noise in the data may be distorted by the operations applied on the raw data [72, 85].

Chapter 3

Image Reconstruction

The image reconstruction process computes a 2D slice of the object emitting radiation, or in the transmission case, exposed to radiation. The measured radiation profiles organized as a sinogram are the input to the reconstruction. As an inverse problem the reconstruction is to some extent similar to the image restoration, but the main difference is the transform from the projection space into the image space.

The noise and the finite amount of measured data make the reconstructed image only an approximation of the true object. The properties of the image depend on the choice of the reconstruction method. The main differences between reconstruction methods come from how the data and the acquisition are modeled.

3.1 Filtered back projection

3.1.1 The Radon transform and its inverse

The tomographic data acquisition is conventionally modeled by the Radon transform (RT) [40]

$$m(t, \theta) \triangleq \mathcal{R}\{f\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta(x \cos \theta + y \sin \theta - t) dx dy . \quad (3.1)$$

The measured data $m(t, \theta)$ are considered as a set of line integrals passing through the unknown 2D object $f(x, y)$ at angle θ ($0 \leq \theta < \pi$), see Fig. 3.1. Then the problem of reconstructing the image $f(x, y)$ from its projections is to compute the inverse RT. This is based on the

Fourier slice theorem: *The 1D Fourier transform of a projection taken at angle θ equals the central radial slice at angle θ of the 2D Fourier transform of the original object.* [40, 45]

The Fourier slice theorem can be proved in a following way. A projection profile is expressed using a rotated coordinate system (t, s)

$$\begin{bmatrix} t \\ s \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \quad \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} t \\ s \end{bmatrix}; \quad (3.2)$$

and (3.1) can be written as

$$m(t, \theta) = \int_{-\infty}^{\infty} f(t \cos \theta - s \sin \theta, t \sin \theta + s \cos \theta) ds. \quad (3.3)$$

Using (3.3), the 1D Fourier transform (FT) of $m(t, \theta)$ with respect to t is

$$\begin{aligned} \mathcal{F}_1\{m(t, \theta)\} &\triangleq M(\omega, \theta) = \int_{-\infty}^{\infty} m(t, \theta) e^{-j2\pi\omega t} dt \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(t \cos \theta - s \sin \theta, t \sin \theta + s \cos \theta) e^{-j2\pi\omega t} dt ds \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{-j2\pi\omega(x \cos \theta + y \sin \theta)} \underbrace{\begin{vmatrix} \frac{\partial t}{\partial x} & \frac{\partial s}{\partial x} \\ \frac{\partial t}{\partial y} & \frac{\partial s}{\partial y} \end{vmatrix}}_{\cos^2 \theta + \sin^2 \theta = 1} dx dy \\ &= F(\omega \cos \theta, \omega \sin \theta), \end{aligned} \quad (3.4)$$

where $F(u, v) = \mathcal{F}_2\{f(x, y)\}$, and ω is the frequency variable of the 1D FT. (3.4) gives the values of the 2D FT evaluated along the line $(u = \omega \cos \theta, v = \omega \sin \theta)$ across the (u, v) space.

The Fourier slice theorem states that if the 2D Fourier space could be filled, the inverse 2D FT would give the original object. This is illustrated in Fig. 3.1. Unfortunately, a discrete implementation of (3.4) using fast Fourier transforms (FFT) requires interpolations especially at the high frequencies, where the density of the resulting 2D Fourier space is low. This is why direct Fourier methods are not popular as inverse Radon transform algorithms [37]. The most common algorithm for the inverse RT is the filtered back projection (FBP) [45], which is described next.

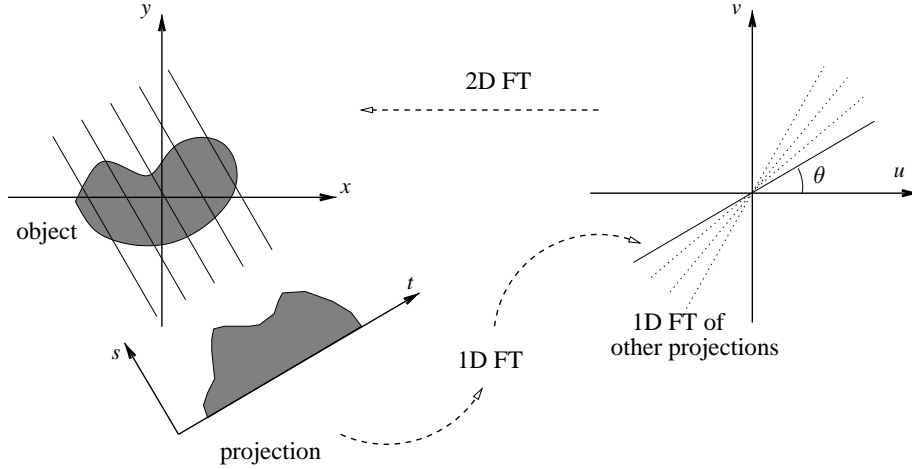


Figure 3.1: The relation of the image space, the Radon space, and the Fourier space according to the Fourier slice theorem.

The inverse 2D FT expressed using the polar coordinates ω and θ in the frequency space ($u = \omega \cos \theta$, $v = \omega \sin \theta$) is

$$\begin{aligned}
 f(x, y) &= \mathcal{F}_2^{-1}\{F(u, v)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(u, v) e^{j2\pi(xu+yv)} du dv \\
 &= \int_0^{2\pi} \int_0^{\infty} F(\omega \cos \theta, \omega \sin \theta) e^{j2\pi\omega(x \cos \theta + y \sin \theta)} \underbrace{\begin{vmatrix} \frac{\partial u}{\partial \omega} & \frac{\partial v}{\partial \omega} \\ \frac{\partial u}{\partial \theta} & \frac{\partial v}{\partial \theta} \end{vmatrix}}_{=\omega} d\omega d\theta \\
 &= \int_0^{\pi} \left[\int_{-\infty}^{\infty} M(\omega, \theta) |\omega| e^{j2\pi\omega(x \cos \theta + y \sin \theta)} d\omega \right] d\theta \\
 &= \int_0^{\pi} \hat{m}(x \cos \theta + y \sin \theta, \theta) d\theta \triangleq \mathcal{B}\{\hat{m}(t, \theta)\}, \quad (3.5)
 \end{aligned}$$

where the Fourier slice theorem (3.4) and the definition $\hat{m}(t, \theta) \triangleq \int_{-\infty}^{\infty} M(\omega, \theta) |\omega| e^{j2\pi\omega t} d\omega$ were used. The multiplication by $|\omega|$ serves as a filter applied to each projection profile in the frequency space. The filtered profile \hat{m} is summed along the ray paths to the image space. The notation $\mathcal{B}\{\hat{m}(t, \theta)\}$ is the back projection of \hat{m} over the image. Now (3.5) gives an algorithm to reconstruct the image $f(x, y)$ from its projections $m(t, \theta)$:

Filtered back projection, FBP: Set $f(x, y) = 0, \forall x, y$. For each projection profile:

- *Take FT*
- *Apply the frequency domain filter*
- *Take inverse FT*
- *Back project over the image at the given angle*

In the discrete implementation of FBP, the integrals are replaced by finite summations and FFTs can be used. The filtering can be performed also as a convolution in the spatial domain, giving the alternative algorithm called the convolution – back projection [60]. The back projection step requires interpolation, which is the most time consuming part of the algorithm. Otherwise, FBP is fast and simple in implementation.

3.1.2 Applicability to PET

FBP is capable to compute accurately the inverse Radon transform. FBP works well especially in the transmission tomography (CT), where the radiation source and the detector are on the opposing sides of the object. In the emission tomography the source is introduced into the object. The measured radiation profiles are organized as a sinogram, and FBP can be used for the emission case, too. Attenuation correction can be applied to the emission sinogram before the reconstruction. However, in ET the measured data are intrinsically noisy, and RT is not an accurate model for the measurement process. This is why the inverse RT image provided by FBP suffers from heavy noise. FBP is especially sensitive to noise because the ramp filter $|\omega|$ in (3.5) amplifies high frequencies. This effect can be weakened by applying a window function to the ramp filter (Fig. 3.2) [36].

A cutoff frequency ω_c can also be used as a truncation threshold for eliminating high frequencies. Including the window $W_{\omega_c}(\omega)$, FBP becomes

$$f(x, y) = \int_0^\pi \left[\int_{-\infty}^\infty M(\omega, \theta) |\omega| W_{\omega_c}(\omega) e^{j2\pi\omega(x \cos \theta + y \sin \theta)} d\omega \right] d\theta . \quad (3.6)$$

Because the frequency bands of the true signal and noise overlap, a trade-off between the resolution and noise rejection is unavoidable. This compromise is tuned by the filter function and the cutoff. The reconstruction artifacts of ET images reconstructed by FBP are most prominent in the background, where the correct value should be zero. The radial streaks are due to the fact that filtered noisy projection profiles do not cancel out each others in the back projection. This results in both positive and negative background pixel values. What is more, the back projection smears each profile over the entire image, and the artifacts are not necessarily canceled out in any part of the image.

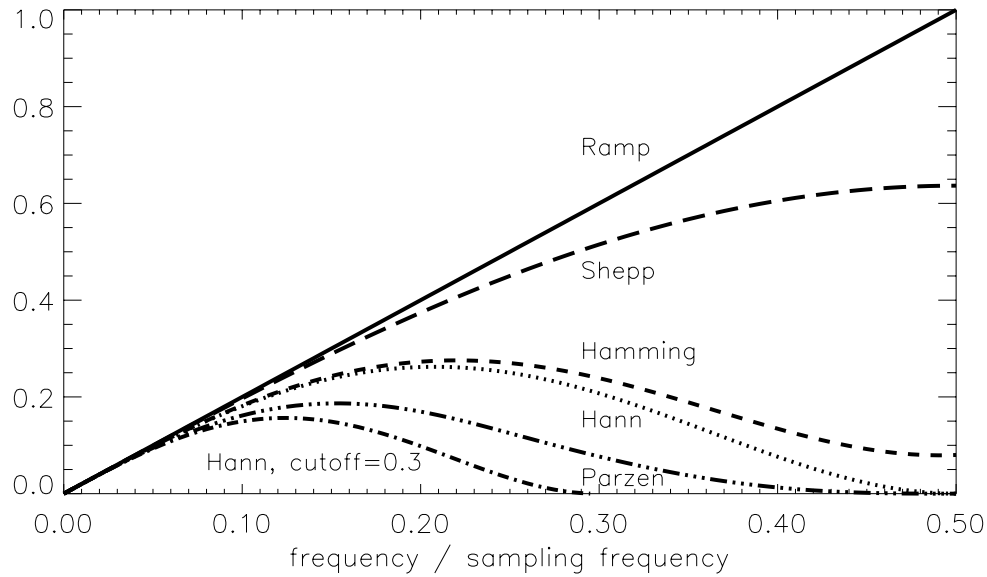


Figure 3.2: Common filter functions for FBP.

The noise in the data originating from a local radioactive decay process contributes over a large spatial area of the image reconstructed by the FBP algorithm. This makes the bias of a ROI not independent of the rest of the image. For instance, the noise and the quantitative value of a low-activity area may depend on whether there is a high activity area present or not. Fig. 3.3 shows typical FBP-artifacts. The quantitative value of a local activity may be contaminated by high activity levels in the arm and in the heart through the streak artifacts across a long spatial distance. Using the Hann window function as noise reduction applies some global filtering and is obviously suboptimal.

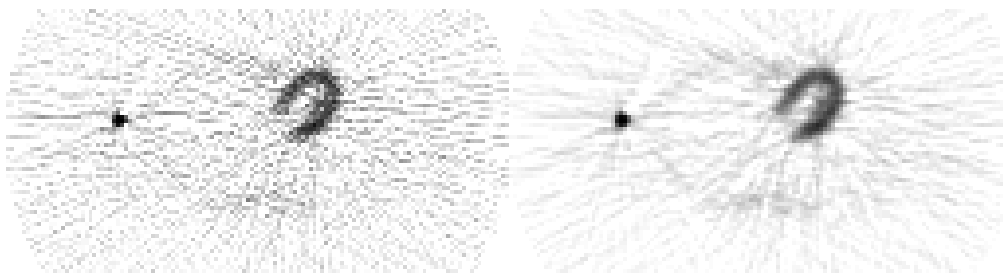


Figure 3.3: Reconstruction artifacts of PET FBP images. **Left:** Ramp. **Right:** Hann. High activity (black pixels) in the vein of the arm contributes streak patterns to the surroundings. Pixels with negative values are not shown.

A transmission image can be reconstructed from the natural logarithm of the ACFs (2.2) using FBP. The image is usually noisy or blurred due to strong filtering, but it can be used for a patient movement check [11].

3.2 On iterative reconstruction algorithms

In Chapter 3.1, the reconstruction problem was solved using transforms and continuous functions for describing the data acquisition and the object. The analytical solution was implemented in a discrete way for the computer. Most iterative algorithms discussed from now on use a statistical model for the acquisition [76, 58]. The measured sinogram data as well as the image are defined to be spatially discrete. The object of interest is the spatial distribution of radioactivity concentrations, that is, the mean number of emissions for each image pixel. Because each tomographic slice has a certain thickness, a pixel is actually a volumetric box, a voxel.

3.2.1 Notations

The notations are listed in Table 3.1. Bold symbols refer to an image or a sinogram as a column vector. Symbols with subscripts are scalars. The measured emission data are assumed to be corrected for attenuation and detector normalization. For PET, a detector refers to a detector pair.

Table 3.1: Notations

$\lambda^{(k)}$	image of k th iteration
b	pixel index, $1 \leq b < B$
d	detector (LOR) index, $1 \leq d < D$
p_{db}	projection weighting factor
x_{db}	complete data, counts from b detected at d
λ_{db}	$\lambda_b p_{db}$, mean of x_{db}
n_d	measured counts at detector d
λ_d	mean counts at detector d
λ_b	mean counts at pixel b

3.2.2 General concept

The mean count of a pixel b , λ_b , is the desired information because it is related to the tracer concentration. For each pixel, the radioactive decay process generates a random number of emissions x_b drawn from the Poisson distribution with the nonnegative mean λ_b [58].

The detected sinogram count is a sum of realized outcomes of independent random variables along the ray path

$$n_d = \sum_b x_{db} \approx \sum_b x_b p_{db} . \quad (3.7)$$

Each pixel-detector pair has a weight p_{db} that describes the contribution of pixel b to detector d (Fig. 3.4). This may include physical aspects like attenuation, photon range, and non-isotropic PSF, but it is often defined using simple geometrical rules, such as an area of the intersection or a bilinear interpolation according to the distances from the pixel to the LOR path [38].

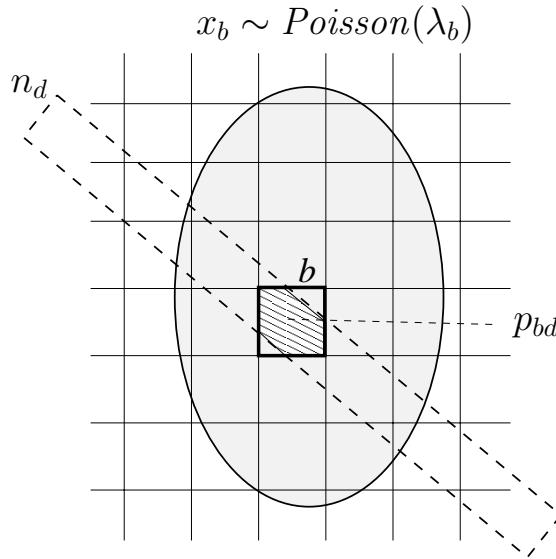


Figure 3.4: The discretization of the image and the acquisition process. n_d is the weighted sum of number of emissions in the pixels overlapping with the LOR of detector d .

There are various approaches for solving \mathbf{x} from $\mathbf{n} = \mathbf{P}\mathbf{x}$, Eq. (3.7). The $D \times B$ matrix \mathbf{P} is sparse but large. Using direct matrix algebra an exact solution does not usually exist with noisy \mathbf{n} and approximate \mathbf{P} , and the linear least-squares solution with pseudo inverses contains negative pixel values [68].

There are iterative algorithms for solving Eq. (3.7), such as algebraic reconstruction techniques (ART) and its variants, simultaneous iterative reconstruction technique (SIRT), and simultaneous algebraic reconstruction technique (SART) [45]. The idea behind them is heuristically easy to understand. At each iteration, the current guess of the image is checked against the data. The criteria is how well the re-projection of the image matches with the measurements. In its simplest form this is a subtraction. This error between the two is projected back to the image space and added to the image, thus updating it. The problem is that the image becomes very noisy as the iterations proceed. What is more, the solution of (3.7), $\hat{\mathbf{x}}$, actually tries to describe the number of occurred counts, not the mean counts. The interest is not to find out how many annihilations did happen during a particular acquisition period because they are just outcomes of random variables and as such inherently noisy. The real interest are the means of those random variables, $\boldsymbol{\lambda}$.

Similar to (3.7), the sum of the means of the random variables along the ray path equals to the mean of the measured data

$$\lambda_d = \sum_b \lambda_{db} = \sum_b \lambda_b p_{db} . \quad (3.8)$$

Eq. (3.8) contains the desired quantity $\boldsymbol{\lambda}$. However, the mean λ_d is not available. With realistic acquisition times, noise contributes heavily to \mathbf{n} ; that is, $\lambda_d \neq n_d$, and Eq. (3.8) can not be solved as Eq. (3.7). The solving of $\boldsymbol{\lambda}$ requires a statistical link between $\boldsymbol{\lambda}$ and \mathbf{n} .

3.3 Maximum likelihood

3.3.1 MLEM algorithm

The maximum likelihood (ML) estimation utilizes the classical estimation approach, where the quantity of interest is unknown but a deterministic constant [51]. This constant is a parameter to a probability density function (PDF) related to the measured random data. Adapted to the reconstruction problem in ET, the ML algorithm searches iteratively for such an estimate or a guess $\hat{\boldsymbol{\lambda}}$ of the unknown image $\boldsymbol{\lambda}$ that it maximizes the conditional probability of the data, given the image. As a function of the unknown image, it is called the likelihood function:

$$l(\boldsymbol{\lambda}) = f(\mathbf{n}|\boldsymbol{\lambda}) , \quad (3.9)$$

where $f(\cdot)$ is the PDF of the data \mathbf{n} . As with ART, at each iteration the current guess of the true image is checked against the data. Now the criteria is

statistically more rigorous: the likelihood. Intuitively that is, if the image were the true image, it should generate such simulated 'measured' data that fit well with the actual measured data. The fit is evaluated according to the likelihood criteria, and the estimate is made iteratively better than previous guesses. The reconstructed ML image is then the estimate

$$\hat{\lambda} = \arg \max_{\lambda} [l(\lambda)] . \quad (3.10)$$

The likelihood in (3.10) can be replaced by the log-likelihood, $\ln(l(\lambda)) = L(\lambda)$, in order to simplify expressions.

The unobserved complete data \mathbf{x} is exploited in the derivation of the maximum likelihood expectation maximization (MLEM) algorithm [66, 58]. The complete data are related to the observed data \mathbf{n} through the many-to-one mapping (3.7). In order to illustrate the reason for introducing \mathbf{x} , let us first build the log-likelihood without using the complete data. Each measured sinogram bin count, n_d , is distributed according to the distribution $Poisson(\lambda_d)$. Then the objective function is

$$\begin{aligned} L(\lambda) &= \ln(f(\mathbf{n} | \lambda)) = \ln \left(\prod_d e^{-\lambda_d} \frac{(\lambda_d)^{n_d}}{n_d!} \right) \\ &= \sum_d [-\lambda_d + n_d \ln(\lambda_d) - \ln(n_d!)] \\ &= \sum_d \left[-\sum_b \lambda_b p_{db} + n_d \ln \left(\sum_b \lambda_b p_{db} \right) \right] + C , \end{aligned} \quad (3.11)$$

where C contains terms independent of λ_b . The function (3.11) should be maximized with respect to λ_b , which is cumbersome due to the nested sums with logarithms.

Equation (3.11) becomes easier to deal with when the complete data \mathbf{x} are used. Each x_{db} is distributed according to $Poisson(\lambda_{db})$. Thus, the complete data log-likelihood is then

$$\begin{aligned} \ln(f(\mathbf{x} | \lambda)) &= \ln \left(\prod_{db} e^{-\lambda_{db}} \frac{(\lambda_{db})^{x_{db}}}{x_{db}!} \right) \\ &= \sum_{b,d} \left[-\lambda_{db} + x_{db} \underbrace{\ln(\lambda_{db})}_{=\lambda_b p_{db}} - \ln(x_{db}!) \right] \\ &= \sum_{b,d} [-\lambda_b p_{db} + x_{db} \ln(\lambda_b)] + C . \end{aligned} \quad (3.12)$$

As x_{db} in Eq. (3.12) is unavailable, (3.12) is replaced by its conditional expectation, given the data \mathbf{n} and the current guess $\boldsymbol{\lambda}^{(k)}$

$$Q(\boldsymbol{\lambda}, \boldsymbol{\lambda}^{(k)}) = E [\ln [f(\mathbf{x}|\boldsymbol{\lambda})] \mid \mathbf{n}, \boldsymbol{\lambda}^{(k)}] . \quad (3.13)$$

A Poisson random variable (x_{db}) conditioned on its sum (n_d) is a multinomial [25], whose expectation can be computed as

$$E [x_{db} \mid n_d, \boldsymbol{\lambda}^{(k)}] = NP = n_d \frac{\lambda_{db}^{(k)}}{\sum_{b'} \lambda_{db'}^{(k)}} = \frac{n_d \lambda_b^{(k)} p_{db}}{\sum_{b'} \lambda_{b'}^{(k)} p_{db'}} \triangleq \bar{x}_{db} , \quad (3.14)$$

where the product NP is the mean of the multinomial. (3.14) is the expectation (E) step of the MLEM algorithm. After replacing x_{db} in (3.12) by \bar{x}_{db} from (3.14), the expectation of the complete data log-likelihood (3.13) is maximized with respect to λ_b . This also maximizes the observed data likelihood $f(\mathbf{n}|\boldsymbol{\lambda})$ [58]. The M step is then

$$\begin{aligned} \frac{\partial}{\partial \lambda_b} \{Q(\boldsymbol{\lambda}, \boldsymbol{\lambda}^{(k)})\} &= \frac{\partial}{\partial \lambda_b} \left\{ \sum_{b,d} [-\lambda_b p_{db} + \bar{x}_{db} \ln (\lambda_b)] + C \right\} \\ &= - \sum_d p_{db} + \sum_d \bar{x}_{db} / \lambda_b = 0 \\ \Rightarrow \hat{\lambda}_b &= \frac{\sum_d \bar{x}_{db}}{\sum_d p_{db}} . \end{aligned} \quad (3.15)$$

MLEM updates the image $\boldsymbol{\lambda}^{(k)}$ sequentially by using (3.15) as the guess for the pixel b in the next iteration [76]. Combining (3.15) and (3.14), MLEM becomes as

$$\lambda_b^{(k+1)} = \frac{\lambda_b^{(k)}}{\sum_d p_{db}} \sum_d \frac{n_d p_{db}}{\sum_{b'} \lambda_{b'}^{(k)} p_{db'}} = \frac{\lambda_b^{(k)} c_b^{L(k)}}{\sum_d p_{db}} . \quad (3.16)$$

The MLEM algorithm (3.16) computes a new pixel value iteratively by multiplying the current pixel by the likelihood coefficient $c_b^{L(k)}$. The term $\sum_d p_{db}$ can include the attenuation of the body, scatter and detector inefficiency [58], if the pre-correction (2.2) is not used. The normalization pre-correction effectively assumes that all emitted photons are detected by some detector, $\sum_d p_{db} = 1$ [76, 66].

MLEM is usually initialized by setting the first image to be a uniform disk, such that it is enclosed by the FOV and the total sum of pixel values matches

with the sum of the sinogram counts [62]. Because the update is multiplicative, it is ensured that the background pixels remain zero and that all pixels are nonnegative. Also, MLEM preserves the total count for all iterations [76]: $\sum_b \lambda_b^{(k)} = \sum_d n_d, \forall k$.

MLEM increases the likelihood $f(\mathbf{n}|\boldsymbol{\lambda})$ in a nondecreasing way towards the fixed point, where the likelihood does not increase any more [66]. If MLEM is allowed to converge, the resulting image provides a good global fit with the measured data.

3.3.2 Noise and bias in iterations

Sometimes an image estimate with a good fit with the data may not be as good as it intuitively might be [71]. This is due to the fact that the reconstruction problem in ET is often an ill-posed problem [68]. It means that a small change in the measured data may cause a large change in the estimated image [14]. Moreover, as the data are noisy, a good fit makes the image noisy, too. This is a result from the fact that when estimating the mean count of a Poisson distribution, the best guess is indeed the number of realized counts in that pixel. In a way, the converged MLEM solution falls into the same pitfall as ART, that is, the pixel values attempt to describe the actual counts, not the mean counts.

A common problem with MLEM is that it generates noisy images when the iterations proceed. This can be seen from Fig. 3.5, which shows images from a sequence of MLEM iterations. In order to avoid this sort of over-fitting, the iterations can be stopped before the convergence [82]. This approach suffers from a noise / bias trade-off: if the convergence is reached, the image is too noisy [78]. On the other hand, if only a small number of iterations are used, the image is less noisy but the quantitative level of pixel values are biased towards the initial starting image [62]. If the initial image is a uniform disk, the local bias may be substantial [3]. If the initial image is a corresponding FBP image, the FBP-based reconstruction artifacts may still remain in the prematurely stopped MLEM image.

Even with the noise / bias trade-off, the MLEM image may still be better than the corresponding FBP image, but the compromise is unavoidable. Thus, most reconstructions called "MLEM images" should be called "aborted MLEM images" in order to emphasize the fact that some of the theoretical attractiveness of the MLEM algorithm may have lost due to the early stoppage.

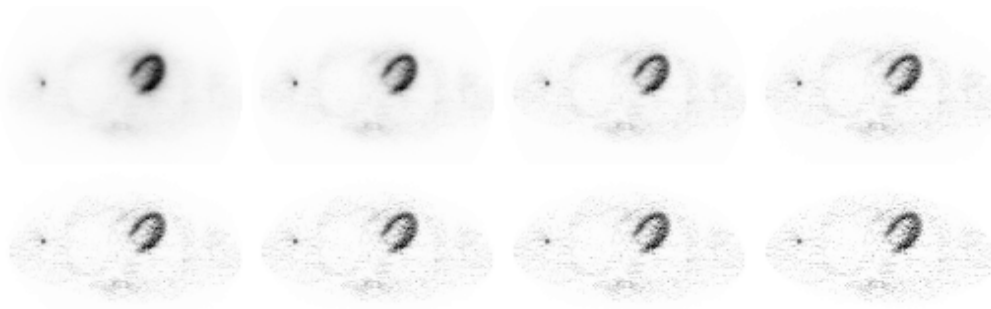


Figure 3.5: Intermediate MLEM reconstructions. **Top row:** iterations 5, 10, 20, and 30. **Bottom row:** iterations 40, 50, 60, and 100.

3.4 Bayes priors and penalties

The concept of making an ill-posed reconstruction problem to a well-posed one is based on introducing extra control on which solutions are more favorable than others. This means that the reconstructed image is required not to fit with the data as well as possible, but also be consistent with additional criteria. Those criteria are set independently of the data. Depending on the point of view, these restrictions can be considered as Tikhonov regularization [14], penalty functions [32], or as Bayesian priors [51]. They all are designed to push the solution towards a predefined assumption about the nature of the true image.

In terms of Bayesian estimation, the desired λ is not an unknown deterministic constant, but a random variable [51]. The *a priori* PDF $f_p(\lambda)$ gives extra information about the image, independently of the measured data \mathbf{n} . The objective function to be maximized is not the likelihood (3.9), but the *a posteriori* PDF

$$f(\lambda|\mathbf{n}) = \frac{f(\mathbf{n}|\lambda) f_p(\lambda)}{f(\mathbf{n})} \propto f(\mathbf{n}|\lambda) f_p(\lambda) . \quad (3.17)$$

Because $f(\mathbf{n})$ is independent of the image, it can be left out from the objective function. $f(\mathbf{n}|\lambda)$ is the likelihood function (3.9). The reconstructed image is then the estimate

$$\hat{\lambda} = \arg \max_{\lambda} [f(\mathbf{n}|\lambda) f_p(\lambda)] = \arg \max_{\lambda} [l(\lambda) f_p(\lambda)] . \quad (3.18)$$

This estimator is called maximum *a posteriori*, MAP. The difference between (3.10) and (3.18) is the introduction of the prior $f_p(\lambda)$. Taking the logarithm, (3.18) separates into a sum

$$\hat{\lambda} = \arg \max_{\lambda} [\ln(l(\lambda)f_p(\lambda))] = \arg \max_{\lambda} [L(\lambda) + P(\lambda)] . \quad (3.19)$$

Now there are two conditions to be met. The first, $L()$, is the likelihood, which requires that the image λ has to be to some extent consistent with the data. The second term in (3.19), $P()$, sets a penalty if the image violates the prior assumptions on what kind of an image is favored. The MAP estimate searches for a balance between the two terms. A noisy image is penalized by the second term, even if it fits well with the data according to the first term.

Depending on the algorithm, the term $L()$ in Eq. (3.19) can express some other data fitting criteria than likelihood, such as least squares error or cross-entropy [48, 18]. Then the estimate is not MAP, but the general problem of over-fitting can be regularized using *ad hoc* penalty functions $-P()$.

The choice of the prior is crucial. The reconstructed MAP image reflects the assumptions made when constructing the prior. The nature of the application of ET is to reveal the unknown activity concentration, with potential unexpected deviations from the normal uptake of the tracer. Too strict conditions may cause loss of relevant information, especially in an unusual case. This is why the prior should be as general as possible, as long as the ill-posedness can be dealt with it. Ideally, the true image should pass the check against the prior unpenalized.

3.4.1 Gibbs priors

A common Bayesian prior is formulated according to the Gibbs distribution, whose general form is [32]

$$f_p(\lambda) = C e^{-\beta U(\lambda)} = C e^{-\beta \sum_b U(\lambda, b)} , \quad (3.20)$$

where β is the Bayes weight of the prior, and C is a normalizing constant. The non-negative energy function $U(\lambda)$ has its minimum and the prior has its maximum when the image λ meets the prior assumptions. $U(\lambda, b)$ is the notation for the value of the energy function $U()$ evaluated on λ at pixel b .

A common choice for $U()$ in (3.20) is an energy function computed using a potential function $v()$ of the differences between pixels in the neighborhood N_b

$$\beta U(\lambda, b) = \beta \sum_{i \in N_b} w_{bi} v(\lambda_b - \lambda_i) , \quad (3.21)$$

where w_{bi} is the weight of pixel i in the neighborhood of pixel b [56]. The parameter β expresses the confidence of the prior. If β is close to 0, the prior

$f_p(\boldsymbol{\lambda})$ is close to its maximum over wide range of $\boldsymbol{\lambda}$'s and (3.20) ranks images cautiously. With a large β , the prior is more peaked and some images are significantly more favorable than others.

As the neighborhood N_b at pixel b is spatially finite, $U()$ has a local support and the Gibbs distribution defines a Markov random field (MRF) [32]. Considering the image as a MRF, its local characteristics are conveniently modeled using the Gibbs prior (3.20).

3.4.2 One step late algorithm

In order to implement the prior, it is necessary to modify the MLEM algorithm according to (3.19). The complete data formulation and the E step (3.14) are same as before. The M step now maximizes the expectation of the log-posterior probability

$$L_p(\boldsymbol{\lambda}, \boldsymbol{\lambda}^{(k)}) = E [\ln [f(\mathbf{x}|\boldsymbol{\lambda})] \mid \mathbf{n}, \boldsymbol{\lambda}^{(k)}] + \ln (f_p(\boldsymbol{\lambda})) = Q(\boldsymbol{\lambda}, \boldsymbol{\lambda}^{(k)}) - \beta U(\boldsymbol{\lambda}) \quad (3.22)$$

with respect to λ_b (ignoring the constant term C). This is maximized by solving

$$\frac{\partial}{\partial \lambda_b} L_p(\boldsymbol{\lambda}, \boldsymbol{\lambda}^{(k)}) = 0. \quad (3.23)$$

Using (3.15),

$$\begin{aligned} \frac{\partial}{\partial \lambda_b} L_p(\boldsymbol{\lambda}, \boldsymbol{\lambda}^{(k)}) &= \frac{\partial}{\partial \lambda_b} Q(\boldsymbol{\lambda}, \boldsymbol{\lambda}^{(k)}) - \beta \frac{\partial}{\partial \lambda_b} \sum_{b'} U(\boldsymbol{\lambda}, b') \\ &= - \sum_d p_{db} + \sum_d \bar{x}_{db} / \lambda_b - \beta \frac{\partial}{\partial \lambda_b} U(\boldsymbol{\lambda}, b) = 0. \end{aligned} \quad (3.24)$$

The one step late (OSL) algorithm uses the current image $\boldsymbol{\lambda}^{(k)}$ when calculating the value of the derivative of the energy function $U()$ in (3.24) [34]. This decouples λ_b from the prior term and, using (3.14), the OSL update can be solved as

$$\lambda_b^{(k+1)} = \frac{\lambda_b^{(k)} c_b^{L(k)}}{\sum_d p_{db} + \beta \frac{\partial}{\partial \lambda_b} U(\boldsymbol{\lambda}, b) |_{\boldsymbol{\lambda}=\boldsymbol{\lambda}^{(k)}}} = \lambda_b^{(k)} c_b^{L(k)} c_b^{P(k)}. \quad (3.25)$$

$c_b^{L(k)}$ is computed using MLEM (3.16). The coefficient that updates the pixel consists of two parts, $c_b^{L(k)}$ as in (3.16), and $c_b^{P(k)}$. With the simplifying normalization assumption $\sum_d p_{db} = 1$, the penalty coefficient

$$c_b^{P\langle k \rangle} = \frac{1}{1 + \beta \frac{\partial}{\partial \lambda_b} U(\boldsymbol{\lambda}, b) |_{\boldsymbol{\lambda}=\boldsymbol{\lambda}^{(k)}}} \quad (3.26)$$

is close to one if the current image meets the prior assumptions at b . Then the values of $\frac{\partial}{\partial \lambda_b} U(\boldsymbol{\lambda}, b)$ are small. Large values of the derivative imply that the image $\boldsymbol{\lambda}^{(k)}$ deviates from the prior assumption.

3.4.3 Mathematical priors

Priors based on pixel difference penalties

The potential function $v()$ in (3.21) defines the behavior of the prior. With the quadratic choice, $v(r) = r^2$, the derivative in (3.25) is $2r$, and the penalty term in (3.26) is linear with respect to the pixel difference $r = \lambda_b^{(k)} - \lambda_i^{(k)}$

$$\begin{aligned} c_b^{P\langle k \rangle} &= \frac{1}{1 + 2\beta \sum_{i \in N_b} w_{bi} (\lambda_b^{(k)} - \lambda_i^{(k)})} \\ &= \frac{1}{1 + 2\beta \left(\lambda_b^{(k)} - \sum_{i \in N_b} w_{bi} \lambda_i^{(k)} \right)} , \end{aligned} \quad (3.27)$$

assuming scaling with $\sum_{i \in N_b} w_{bi} = 1$. The penalty is set with respect to the sum $\sum_{i \in N_b} w_{bi} \lambda_i^{(k)}$, against which the current pixel $\lambda_b^{(k)}$ is compared. In effect, the penalty reference is the output of a linear finite impulse response (FIR) filter with filter coefficients w . With proposed w 's [56], the FIR is a low-pass filter [69], and non-smooth and noisy images are penalized. Images that are close to the result of the reference FIR filter are smooth and they are not much penalized. This results in blurring of sharp edges between different activity concentrations in the tissues. The prior according to (3.27) can therefore be called the smoothing prior.

The number of coefficients w in the reference FIR is limited because the neighborhood of the corresponding MRF can not be expanded spatially too wide. The correlation between pixels in the ET image tends to drop quickly as the spatial distance between them grows. Thus, a realistic prior has a narrow range in the Markovian sense. A window size of a few pixels leaves little space for a proper FIR design.

With (3.27), the MAP estimate is less noisy than the MLEM image. The prior assumption is that the true image is smooth. But an ET image is known to contain non-smooth areas, such as boundaries between different tissues.

Hyperparameter tuning

In order to avoid blurring, the quadratic $v(r)$ in (3.21) can be replaced by other functions, whose derivative is less prominent in penalizing step edges. Some of these potential functions and their derivatives are listed in Table 3.2. The chosen potential function can be further tuned by an extra parameter, T , as $v(r/T)$ [56]. The Huber prior uses T a threshold in order to use the constant penalty function instead of the linear one for large pixel differences [15]. The additional parameters of the prior are called hyperparameters. The adjustment of the Bayes weight β and especially the hyperparameter T is difficult in general [48]. The main difficulty is that the parameter T adjusts the penalty by the height of the step r [55, 8].

Table 3.2: Suggested Potential Functions

$v(r)$	$\frac{\partial}{\partial \lambda_b} v$	
r^2	$2r$	Quadratic, [56]
$\ln(\cosh(r))$	$\tanh(r)$	[34]
$r^2/(1+r^2)$	$2r/(1+r^2)^2$	[56]
$\ln(1+r^2)$	$2r/(1+r^2)$	[56]
$[r + (1 + r)^{-1} - 1] / 2$	$[\text{sign}(r) - \text{sign}(r)/(1 + r ^2)] / 2$	[56]
$\begin{cases} r^2/2 & , r \leq T \\ T r - T^2/2 & , r > T \end{cases}$	$\begin{cases} r & , r \leq T \\ T \text{sign}(r) & , r > T \end{cases}$	Huber, [15]

By convenient choice of parameters β , T , or $v()$ itself, the derivative of the potential function is tailored such that edges that are high enough are preserved. Fig. 3.6 shows various proposed potential and penalty functions from Table 3.2. The functions in the lower plot (Fig. 3.6(b)) should be designed in such a way that the relevant pixel differences were not penalized too much [35, 55].

There are two possible pitfalls in this approach. First, the true edge height is unknown, except for artificial phantoms, for which most priors work well. Secondly, low edges are the most important ones to be detected because they represent small quantitative differences between concentrations in adjacent tissue areas. If the threshold T is set properly, high edges are not blurred too much and spatial changes of low amplitude are penalized. This may have potential risks of losing important information and degrading the sensitivity in distinguishing between two different activity levels, because the power of ET should be in detecting small and unexpected, often abnormal changes in tracer

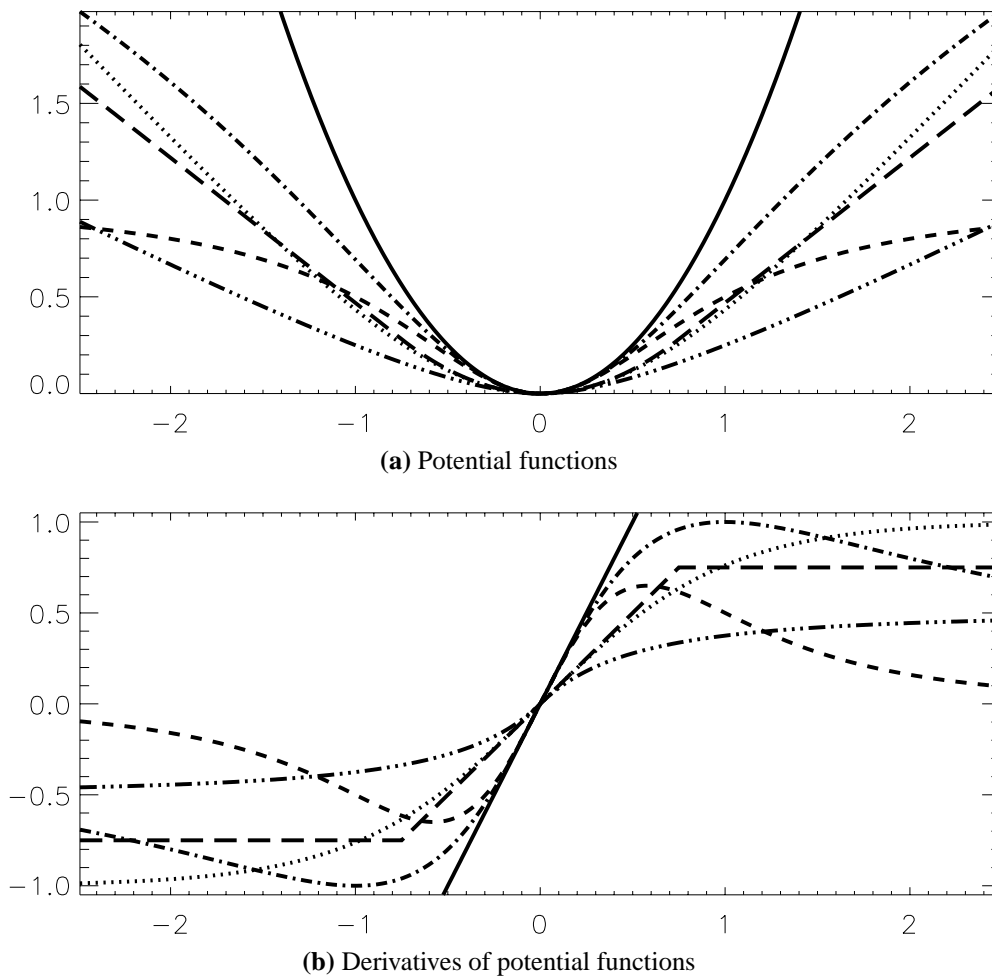


Figure 3.6: Potential and penalty functions as a function of the pixel difference r (in arbitrary units). The solid line is the quadratic function, the others are in the order of Table 3.2: dotted, dashed, dash dot, dash dot dot dot, long dashes. Parameter $T = 0.75$ for Huber prior.

activity.

The choice of parameter T has a trade-off in the amplitude domain, much in a similar way than the cutoff parameter in FBP reconstruction. Also, the proper choice of the derivative function (Fig. 3.6(b) on page 27) is similarly difficult and arbitrary as the choice of the FBP filter function (Fig. 3.2 on page 15). The interesting and harmful parts of the signal overlap, both in the frequency domain (cutoff) and in the amplitude domain (T). Setting thresholds in order to distinguish the two parts may lead to compromises in terms of resolution or sen-

sitivity. However in the transmission image reconstruction, setting thresholds is reported to be less critical [67].

The gamma prior is another suggestion for the prior PDF in (3.22) [57]. It has two pixel-wise parameters, α_b and β_b :

$$f_p(\lambda_b) = \Gamma(\alpha_b)^{-1} (\alpha_b/\beta_b)^{\alpha_b} \lambda_b^{\alpha_b-1} e^{-\alpha_b \lambda_b/\beta_b}, \quad (3.28)$$

where $\Gamma(\alpha)$ is the gamma function $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt$ [71]. The gamma prior (3.28) is a Gibbs prior with a single-pixel neighborhood. The mean of the PDF is β_b and the variance is β_b^2/α_b . The log-prior $P()$ in (3.19) is

$$P(\lambda_b) = (\alpha_b - 1) \ln(\lambda_b) - \alpha_b \lambda_b/\beta_b. \quad (3.29)$$

The motivation for the gamma prior is that it is analytically convenient to deal together with Poisson-based likelihoods. The Hessian matrix of the log-prior (3.29) is diagonal with strictly negative diagonal entries, which keeps the objective function in (3.19) concave [57]. Because the gamma prior is independent of other pixels, the objective function can be solved without the OSL-technique. Unfortunately, the parameters α_b and β_b are difficult to adjust [83].

In general, the fundamental problem with designing mathematical priors is the difficulty to define analytically such a PDF or a energy function and its hyperparameters, which would not penalize the unknown true image.

3.4.4 Anatomical priors

Anatomical images (MRI, CT) have been used as priors in order to improve the edge sharpness in the emission images [20, 8, 17, 61]. The anatomical image of the same subject brings information on the tissue borders into the emission reconstruction. From this edge map data a line site process is iteratively updated in parallel with the actual reconstruction algorithm. The line sites are set in between pixels and they effectively guide the prior not to smooth out the edges, given that the multimodal image registration can be accomplished [9, 21].

Anatomical priors are justified by the fact that there is a strong correlation between the anatomical image and the corresponding emission image (see Fig. 1.1). For instance, an anatomical boundary between bone and gray matter in the brain implies a probable boundary in the emission activity, too. The last term in (3.19) can be tuned using the anatomical image, expressing the believed similarity of the emission image and the anatomical edge map [8]. The penalty is not set if there is a match between the emission activity and the anatomical edge map. This allows for strong edges not to be penalized, but requires some sort of knowledge of the tissue types. The resulting emission image has often

sharp edges and the smoothing prior has removed most of the noise in the areas surrounded by the anatomical borders.

The critical aspect with this complicated approach is that even though the correlation between a MRI image and a PET image is high, the most important information is the differences between them, as stated in Section 1.1 on page 2. The penalty according to the differences between anatomical and emission data nicely penalizes noise, and the resulting image look almost as good as the anatomical image. The risk is that the reconstructed image is too general, kind of an atlas picture, not a precise description of the status of the tissue of the particular patient at the given time.

3.5 Implementational aspects

3.5.1 Acceleration techniques

The MLEM algorithm is slow in convergence [68]. During the first iterations the likelihood increases rapidly but the rate slows down at later iterations. The rate of the convergence depends by the design of the objective function [48, 18, 23, 28]. If the objective function is made more shallow near the global maximum, the step size is increased [29]. The speed depends on the strategy how the algorithm performs the optimization. Gradient based algorithms can be accelerated by going further to the direction indicated by the update [74]. Space-alternating generalized EM (SAGE) updates a subset of pixels at a time [31]. Coordinate ascent (or descent) algorithms update one pixel at a time [28, 16].

One of the most popular ways to speed up the rate of convergence is the ordered subsets (OS) method [39]. In OS, the projections are divided into smaller subsets of projection angles (sinogram rows). Eq. (3.16) is applied after processing one subset. The increase in speed comes from the fact that the image is updated more frequently. If one subset contains every fourth projection angle, the pixel values are assigned a new value four times during the processing of all projections once. There are two drawbacks in the heuristic nature of OS. There are no proofs of convergence and the usage of priors is ambiguous. The prior can be applied in between the sub-iterations or after a complete iteration [52]. With high OS-factors, both of the strategies may be unsatisfactory. The effect of one sub-iteration to the image is a streak-like pattern with visible gaps between the projection angles. This may be penalized too much by the prior, changing the net effect in an unexpected way. On the other hand, a complete cycle may increase the amount of noise too much. Nevertheless, several algorithms have been speeded up by applying the OS method [52, 46].

Most of the iterative reconstruction algorithms can be and need to be regularized with a prior. The effect of the prior is that more iterations can be run and the image will be more precise, given that the prior does not cause a bias of its own.

3.5.2 Corrections

The background events (accidentals, scattered) in the measured data can be estimated and collected to a separate randoms sinogram in some devices [62]. Because the pre-correction destroys the Poisson nature, the data can be left uncorrected and the randoms can be taken into account in the mean number of detected counts: $\lambda'_d = \lambda_d + r_d$ [68]. Also, the statistical model can include the effect of the pre-correction [85]. These choices may change the unregularized algorithm, but the prior remains the same independent of whether the randoms are pre-corrected or taken care of inside the algorithm. In this work, the randoms pre-correction provided by the PET manufacturer was applied.

Similarly, the pre-correction of the emission sinogram for the body attenuation has been a *de facto* method, especially in older ET devices. This is the way how attenuation correction was accomplished for the data used in this work, unless otherwise stated. The effect of the attenuation can be included in the projection weighting factors p_{db} , as mentioned on page 17. This is theoretically more sound than the pre-correction and it is reported to improve the results of an iterative algorithm [65]. However, the problem of the regularization and the definition of penalties or priors still remains.

3.6 Performance measures of estimators

The reconstruction algorithm can be considered as an estimator of the unknown quantity. There are several ways to compare the estimator performance [51]. The *unbiasedness* is a desirable property

$$bias(\hat{\lambda}) = E(\hat{\lambda}) - \lambda = 0 , \quad (3.30)$$

where $\hat{\lambda}$ is the estimate of the quantity λ .

There is often a trade-off between the bias and the variance of the estimate. The variance measures the variability of the calculated estimate. An estimator with a small bias is correct on the average, but it may have a large variance, which makes an individual estimate unreliable.

A proportional measure of two estimators A and B is the *relative efficiency*, the relation of the two variances

$$efficiency = \frac{Var(\hat{\lambda}_A)}{Var(\hat{\lambda}_B)} . \quad (3.31)$$

The *mean square error* (MSE) is a measure of both the bias and the variance

$$MSE(\hat{\lambda}) = E \left[(\hat{\lambda} - \lambda)^2 \right] = \left[bias(\hat{\lambda}) \right]^2 + Var(\hat{\lambda}) . \quad (3.32)$$

The *mean absolute error* (MAE) is another measure of the quantitative accuracy and the variability of the estimate

$$MAE(\hat{\lambda}) = E \left[|\hat{\lambda} - \lambda| \right] . \quad (3.33)$$

Because some of these measures need the knowledge of the true quantity, they are most useful in simulations.

3.7 Summary

The standard image reconstruction algorithm in ET, FBP, has been the same as the one that is successfully used in CT. The quality of such an ET image is compromised by the fact that the data acquisition process is different in the two cases. In ET, the measured data have a strong noise component due to the statistical nature of the decay of the radiating source. This is why the FBP images are often contaminated by noise and reconstruction artifacts. In order to take into account the acquisition process of ET, the iterative reconstruction methods utilize a statistical model. The main problem with these iterative algorithms is the increase of noise in the image when the number of iterations increases. This is unfortunate because the quantitative accuracy requires a reliable evaluation of the concentration value of a region of interest (ROI) of the image.

In Section 3.4, various approaches of including prior information were reviewed. Mathematically defined priors may serve as convenient regularizers if their hyperparameters can be tuned properly. But rather than expressing the true nature of the unknown image, they serve as possible techniques to penalize the noise.

There are some figures of merit that a decent image reconstruction method should comply with. A small bias, and at the same time a small variability and insensitivity to noise are important in ET. The practical usability of the algorithm depends more on the hardware. The visual quality is subjective and it depends on human experiences and practices, which makes it a very complex issue.

Regarding the various algorithms for implementing the chosen reconstruction method, the choices are many. In order to keep the solution practical, the

method should not be too complex, yet general enough. This is the problem with anatomical priors. Other modalities as well as many reconstruction parameters (Bayes weight, tuning parameters, number of iterations) would make it difficult to set up a procedure, which is simple enough to be used in practice. Alone the choice of the two parameters of FBP (window, cutoff) is more or less based on subjective issues and carried out empirically [26].

The generality is important because with few assumptions and restrictions, the solution does not contain built-in descriptions about the image properties. The object is, after all, unknown.

Chapter 4

Aims of the work

The large variety of different reconstruction methods indicates that the image reconstruction problem has been addressed using many different approaches. None of them is the best one; each method expresses the assumptions and the point of view of its own. When using FBP, the starting point is, or should be, that the measured data is close to the Radon transform of the object. Applying a smoothing prior with iterative algorithms means that the object is believed to be smooth, parameter-tuned priors express some known specific properties of the object, and so on. All these solutions have their justifications, and the reconstructed image expresses the underlying assumptions.

Our aim was to develop an effective noise reduction method for statistical ET reconstruction, without significantly compromising the quantitative accuracy. The noise reduction procedure should not cause significant bias to local pixel values. In this sense the requirements are different from general image enhancement operations [40, 10]. Especially in some PET studies, the quantitative aspect is often more important than the visual appearance.

One of the main emphases with the design of the penalty or the prior was on avoiding unnecessary complexity. This means both the practical operations required in the clinical work and the technical usability of the algorithm. The number of the parameters of an iterative algorithm should be few, if possible just one. Also, the parameter should not be too sensitive to suboptimal choice of values because it is difficult to assign the optimal value objectively.

The emission image is the main goal of the ET imaging. The transmission measurements are important because they make quantitative images possible. The transmission images themselves are quite rarely needed, but they can be used in attenuation correction. The statistical transmission reconstruction algorithms suffer from the ill-posedness just like the emission algorithms. Thus, the general principle for reducing noise during the reconstruction of ET images should apply for both emission and transmission.

Given the simplicity of the developed method, it is sufficient to compare it against the main standard algorithms, FBP and MLEM. A survey of the relative performance of other more complex priors is beyond the scope of this work.

Chapter 5

Median Root Prior

5.1 General description

The new approach starts with the general description of the unknown emission image: *the desired image is assumed to be locally monotonic*. Inside a local neighborhood, the changes of the pixel values are spatially non-decreasing or non-increasing. Using that assumption as a penalty to MLEM, the median root prior (MRP) algorithm was developed [1, 2, 4, 3]. Images that are invariant under median filtering are called as root images, and they are locally monotonic [10]. By setting the penalty of a pixel against the local median, the penalty is set only if inside a local neighborhood the image contains non-monotonic structures. No other constraints are used. This constraint is plugged in the term $P()$ of Eq. (3.19). The difference between adjacent pixel values is not penalized in MRP, which makes it possible to avoid the tuning of the derivative of the energy function, as discussed in Section 3.4.3.

The Bayesian interpretation of the penalty is that MRP's prior favors any locally monotonic image. This is what the otherwise unknown discrete version of the continuous true image is assumed to be. The density of the image grid affects on how well the assumption can be justified. With too few pixels, the assumption of local monotonicity may be not well met, but with reasonably small pixels the assumption can be considered fairly valid. In the iterative reconstruction process the image is guided towards the root by MRP; the final image is and needs not to be a root in general.

5.2 MRP Algorithm

5.2.1 Emission OSL algorithm

Using the OSL-form (3.25, 3.26), the MRP algorithm was defined for emission as [2]

$$\lambda_b^{\langle k+1 \rangle} = \frac{\lambda_b^{\text{MLEM} \langle k+1 \rangle}}{1 + \beta \frac{\lambda_b^{\langle k \rangle} - M_b}{M_b}} = \lambda_b^{\langle k \rangle} c_b^{L \langle k \rangle} c_b^{P \langle k \rangle}, \quad (5.1)$$

where $M_b = \text{Med}\{\lambda_i^{\langle k \rangle} | i \in N_b\}$ (the median of pixels in the neighborhood N_b centered at b), and β is the weight of the prior. The positivity constraint for the pixel is: $0 < \beta \leq 1$. The penalty coefficient $c_b^{P \langle k \rangle}$ sets a penalty if the old pixel value is not close to the local median. This encourages the solution towards locally monotonic images.

The penalty of MRP is set according to how much the center pixel differs from the local median. In contrast to priors of Table 3.2, individual pixel differences are not penalized. Since the median follows an edge, a simple penalty function can be used. The prior needs not be explicitly instructed to behave differently in flat and edgy image areas based on information from other modalities, nor do quantitative aspects such as the edge height or the noise amplitude cause any parameter tuning. If the image is locally monotonic, no penalty is applied. The restriction of local monotonicity is quite modest because it does not assume anything about the appearance or the shape of the true object. Only non-monotonic local changes of activity between pixels in the neighborhood are penalized. Typically, noise is locally non-monotonic and the true signal is locally monotonic in ET. This is quite different from the case of the conventional smoothing prior, for which a uniform or a flat image passes the prior without penalty.

5.2.2 Interpretations on the prior

Because of the median, an analytical analysis of the properties of MRP is not straightforward. The following is not a rigorous derivation of MRP, but it is simply intended to point out an interesting *intuitive* connection between MRP and Gaussian type of priors.

A general prior distribution in the form of (3.20) for Gaussian PDF's is

$$f_p(\boldsymbol{\lambda}) = C e^{-\frac{\beta}{2} \sum_b (\lambda_b - m_b)^2 / m_b}, \quad (5.2)$$

where the mean of the prior PDF, m_b , is a hyperparameter. The log-prior or the Gibbs energy function of (5.2) is of the simple quadratic form

$$\beta U(\boldsymbol{\lambda}) = \beta \sum_b U(\boldsymbol{\lambda}, b) = \beta \sum_b \frac{(\lambda_b - m_b)^2}{2m_b}, \quad (5.3)$$

and the derivative term in (3.26) is

$$\beta \frac{\partial}{\partial \lambda_b} U(\boldsymbol{\lambda}, b) = \frac{\beta(\lambda_b - m_b)}{m_b}. \quad (5.4)$$

By assigning (5.3) into Eq. (3.22), the objective function or the log-posterior function is

$$L_p(\boldsymbol{\lambda}, \boldsymbol{\lambda}^{(k)}) = Q(\boldsymbol{\lambda}, \boldsymbol{\lambda}^{(k)}) - \beta U(\boldsymbol{\lambda}) = Q(\boldsymbol{\lambda}, \boldsymbol{\lambda}^{(k)}) - \beta \sum_b \frac{(\lambda_b - m_b)^2}{2m_b}, \quad (5.5)$$

where $Q(\boldsymbol{\lambda}, \boldsymbol{\lambda}^{(k)})$ is the expectation of the complete data likelihood function (3.13) used in (3.15) when deriving the MLEM algorithm. In principle, the likelihood can be replaced by another data fitting criteria, such as least squares or entropy, as noted in Section 3.4.

An algorithm for this hypothetical prior can be derived by the maximization of (5.5) with respect to λ_b , using the OSL-technique that gives

$$\frac{\partial}{\partial \lambda_b} \{L_p(\boldsymbol{\lambda}, \boldsymbol{\lambda}^{(k)})\} = - \sum_d p_{db} + \frac{\sum_d \bar{x}_{db}}{\lambda_b} - \frac{\beta(\lambda_b - m_b)}{m_b} = 0 \quad (5.6)$$

$$\Rightarrow \quad \lambda_b^{(k+1)} = \frac{\lambda_b^{(k)} c_b^{L(k)}}{\sum_d p_{db} + \beta \frac{\lambda_b^{(k)} - m_b}{m_b}}. \quad (5.7)$$

In order to have any practical value for (5.7), the values of the hyperparameters m_b must be set somehow, possibly depending on the image $\boldsymbol{\lambda}$, which is again depended on the measured data. Also, m_b 's can be made spatially depended on each others by $m_b = g\{\lambda_i^{(k)} | i \in N_b\}$, where the function g and the neighborhood N_b link the means m_b together.

As discussed in Section 3.2.2, λ_b 's are the means of the realized outcomes x_b for each pixel. Then by defining a neighborhood N_b , a MRF it set up, which expresses a limited spatial dependency. The dependency is caused by computing the hyperparameters values of overlapping neighborhoods. The random variables x_b and λ_b are themselves each pixel-wise independent.

By comparing (5.1) and (5.7), an intuitive connection between MRP and the above Gaussian prior can be seen: $m_b = M_b = \text{Med}\{\lambda_i^{(k)} | i \in N_b\}$ (with normalization $\sum_d p_{db} = 1$). So it looks like the location of the Gaussian prior is

selected to be the median of the neighborhood. Unfortunately, if this definition of m_b is assigned into (5.2), the analytical derivation becomes intractable. Due to the median operation, the dependence of M_b on the image $\lambda^{(k)}$ is non-linear.

The following heuristic simplification was used in [3]: the hyperparameter M_b was treated as an unknown constant, whose value was selected later using the robust median operator. For some justification, Fig. 5.1 illustrates the dependence of M_b on λ_b . When the value of M_b equals the center pixel value, the penalty term vanishes and the prior is effectively not in use (sloped part of the plot in Fig. 5.1). Otherwise, M_b is selected to be one of the other pixels $\{\lambda_i | i \neq b, i \in N_b\}$ (constant part of the plot in Fig. 5.1), and as such treated as an independent variable when taking derivatives with respect to λ_b .

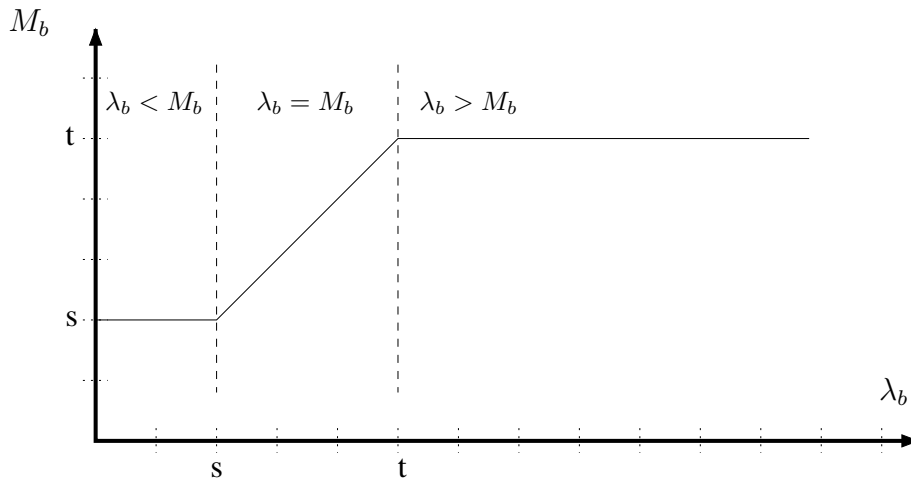


Figure 5.1: The hyperparameter M_b as a function of λ_b . $s = \max\{\lambda_i | \lambda_i \leq M_b, i \in N_b\}$, and $t = \min\{\lambda_i | \lambda_i \geq M_b, i \in N_b\}$. The denominator in (5.1) is 1 when $s \leq \lambda_b \leq t$.

Thus, the dependence of M_b on λ_b either disappears from the penalty term of MRP (5.1), or, when $M_b \neq \lambda_b$, M_b is effectively a constant. However, as the value of M_b is not strictly independent of λ_b , the connection to Gaussian priors is up to intuitive interpretations.

5.3 Transmission algorithm

5.3.1 Motivation

Quantitative PET studies require the computation of attenuation correction factors (ACF) for compensating the body attenuation effect in the emission data,

as mentioned in Section 2.2.1. When emission data are multiplied by the ACFs, the noise or blurring caused by smoothing propagate from ACFs to the emission data.

If the image of linear mass absorption coefficients μ is available, the ACFs can be generated by implementing the principle of Eq. (2.3). From the reconstructed μ -image, the ACFs are computed as

$$ACF_d = e^{R(l_d, \mu)}, \quad (5.8)$$

where $R(l_d, \mu) \triangleq \sum_b l_{db} \mu_b$ is the forward-projection of the image μ to the LOR d , and l_{db} is the weight of the contribution of pixel b to LOR d .

As mentioned in Section 3.1.2, transmission image can be reconstructed using FBP from blank and transmission measurements. However, FBP images are noisy and some bias may introduced in the transmission image [28]. Usually the image is segmented before (5.8) in order to reduce noise and artifacts. This means that some assumptions need to be made about tissue types and their typical μ -values [63, 84]. The segmentation method is rather complex to apply for a diverse set of attenuation values. In practice, some details may not properly contribute to the ACFs [77]. Thus, high quality transmission images are desired for accurate attenuation correction.

5.3.2 Data fitting algorithm

The iterative methods do not usually reconstruct the logarithm of the ACFs (2.2), but use a more precise statistical model for the scans [30]. The likelihood for the measured transmission data is Poisson

$$f(\mathbf{T} | \mu) = \prod_d e^{-B_d} e^{-R(l_d, \mu)} \frac{(B_d e^{-R(l_d, \mu)})^{T_d}}{T_d!}, \quad (5.9)$$

where \mathbf{B} is the blank sinogram, and \mathbf{T} is the transmission sinogram. μ_b is the probability of photon capture per unit length for pixel b . The mean of the Poisson can include random events r_d such as accidentals or scatter: $B_d e^{-R(l_d, \mu)} + r_d$, if available [28].

The objective function to be maximized is the log-likelihood

$$L(\mu) = \sum_d [-B_d e^{-R(l_d, \mu)} - T_d R(l_d, \mu)] + C. \quad (5.10)$$

The MLEM algorithm for (5.10) was developed in [58] using a complete data formulation, but it is too slow in practice [59]. The convex algorithm for transmission by Lange et al. [59] is commonly used instead. It uses Newton's iteration and updates the pixel as

$$\mu_b^{C\langle k+1 \rangle} = \mu_b^{\langle k \rangle} + \frac{\mu_b^{\langle k \rangle} \sum_d l_{db} [B_d e^{-R(l_d, \boldsymbol{\mu}^{\langle k \rangle})} - T_d]}{\sum_d l_{db} R(l_d, \boldsymbol{\mu}^{\langle k \rangle}) B_d e^{-R(l_d, \boldsymbol{\mu}^{\langle k \rangle})}}. \quad (5.11)$$

If \mathbf{B} and \mathbf{T} have different acquisition times, \mathbf{B} is scaled accordingly. \mathbf{B} and \mathbf{T} are both normalized for detector compensation. There are various ways to include mathematical priors of the type of Table 3.2 into (5.11) [59]. Anatomical priors are not an option because a CT or MRI image could be used in place of the ET transmission image. The OS acceleration method (Section 3.5.1) can be applied for the convex algorithm [46].

5.3.3 tMRP algorithm

The algorithm (5.11) is ill-posed and some regularization is required. Because the assumption of local monotonicity of the true image is equally valid for both emission and transmission, MRP can be used as the penalty. The aim is to make high quality transmission images with as short transmission acquisition times as possible. These images are used in computing low noise ACFs for emission data.

MRP can be applied to (5.11) analogously to the emission OSL algorithm (5.1). The transmission image is computed by the tMRP (transmission MRP) algorithm [6, 5]

$$\mu_b^{\langle k+1 \rangle} = \frac{1}{1 + \beta \frac{\mu_b^{\langle k \rangle} - M_b}{M_b}} \mu_b^{C\langle k+1 \rangle}, \quad (5.12)$$

where $\boldsymbol{\mu}^{\langle k \rangle}$ is the current transmission image, and $M_b = \text{Med}\{\mu_i^{\langle k \rangle} | i \in N_b\}$. The role of the parameter β is the same as for the emission OSL-algorithm.

(5.12) is simplest way to apply the penalty of MRP to any transmission algorithm. The convex transmission algorithm (5.11) is relatively slow due to the exponentiations. There are other transmission algorithms [28, 30], and their speed depends on the update strategy of the algorithm, as discussed in Section 3.5.1.

With tMRP, the ACFs can be generated from count-limited data with poor statistics without smoothing or segmentation, and no assumptions need to be made about tissue types and typical μ -values. tMRP can be applied to various procedures carrying out the transmission scan, such as the simultaneous transmission-emission scan mentioned in Section 2.2.1. tMRP can also be used for generation of the attenuation map for the cases where attenuation is included in the projection weighting factors p_{db} of the emission reconstruction algorithm.

5.4 On convergence properties

As an OSL, MRP shares with the convergence properties of an OSL algorithm [56]. The effect of the prior on the convergence of the updating algorithm can be examined by the properties of the log-prior part of the objective function [57]. The fact that M_b is nonlinearly dependent on λ_b , makes it not easy to rigorously derive the convergence properties of MRP. However, with the simplistic but intuitive interpretation of the nature of the hyperparameter $M_b \in \{\lambda_i \mid i \neq b\}$ (page 38 and [3]), MRP enjoys same properties that motivated the usage of the gamma prior (3.29, page 28) [57]. A rigorous derivation of convergence properties of MRP is still a challenging problem.

In the absence of a proof of the convergence of MRP, the experience of using MRP in practice has not brought up any problems of instability. This issue will be addressed in Section 6.1.

5.5 Noise and bias

MRP operates solely on the current image estimate computed by the updating algorithm (MLEM or any other one). As there are no quantitative assumptions or thresholds in the design of the prior itself, MRP is supposed to have a very small bias. The noise suppression is based on a general description of any object, which leaves the updating algorithm the liberty to proceed with as many iterations as is needed.

5.6 Implementation

5.6.1 Neighborhood size

The spatial size of the median mask defines the area, over which the criteria of monotonicity is required, or in other words, the neighborhood system of the corresponding MRF. This gives the possibility to pre-set the width of the smallest detail to be preserved. The width of the smallest horizontal or vertical detail that a root signal can contain is half of the mask size

$$dw = \frac{(mw + 1)}{2} pw = \frac{(mw + 1) fov}{2 iw zoom}, \quad (5.13)$$

where mw is the median mask width [pixels], pw is the pixel width [mm], fov is the width of FOV [mm], iw is the image width [pixels], and $zoom = fov/(iw pw)$. By setting the image dimension, the zoom factor, and the mask size accordingly (so that small neighboring objects do not both fall inside the

mask), MRP penalizes only details smaller than a known size (5.13). This requires that the number of pixels in the image is adjusted to be large enough, so that the mask size in pixels is reasonable (3×3 or perhaps preferably 5×5). The size of those small neighboring objects is easier to set beforehand than quantitative thresholds. Thus, the mask size affects the granularity rather than the resolution or the sensitivity. Unfortunately, the computation time of the data fitting part of the algorithm increases heavily along with the image dimension.

5.6.2 Parameter β

After choosing the neighborhood size, the weight of the prior must be chosen. Neither of the MRP-algorithms, emission (5.1) and transmission (5.12), have other user-set tuning parameters than the Bayes weight, β . The robustness of median makes the choice of β rather uncritical (see Section 6.1.4). Experimentally [2, 4, 3, 52, 47], $\beta \approx 0.3$ is a good choice for (5.1) when $mw = 3$. If $mw = 5$, there are more pixels inside the mask and β can be smaller: $\beta = 0.3 \times 3^2/5^2 \approx 0.11$.

Greater values of β shift the new pixel value closer to M_b setting more weight on the assumption that the image should be locally monotonic. The closer the image is to its median-filtered image, the less the prior needs to change the image. MRP does not use the median filter as a direct operation, but rather as a way to compute a reference for the penalty. As pointed out on page 25, the smoothing prior uses a FIR filter for the penalty reference in a similar way. On flat areas the difference between a low-pass FIR and the median is quite small, especially with a larger 5×5 -mask. The difference is that the median brings the edge preservation capability into the algorithm without tuned hyperparameters.

5.6.3 Iterations

The iterations of MRP start with a uniform disk as an initial image. The few first iterations are plain MLEM iterations. The penalty is taken in use after 2 – 4 iterations. This is because during the early iterations MLEM does not yet generate noise to the image, but forms a coarse approximation of the image. Small details and high spatial frequencies are produced at later iterations, and there is no need to control the monotonicity of the image right from the start. Also, the approximations made with the OSL-algorithm (page 24) are less significant at later iterations, when the changes of the pixels between the iterations are generally small. Fig. 5.2 shows images of a MRP sequence of iterations. Compared with the MLEM sequence of Fig. 3.5 on page 22, later iterations do not generate noise in the MRP images.

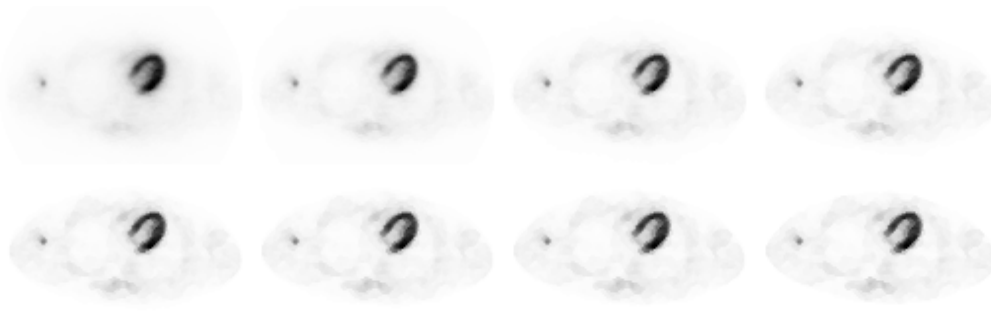


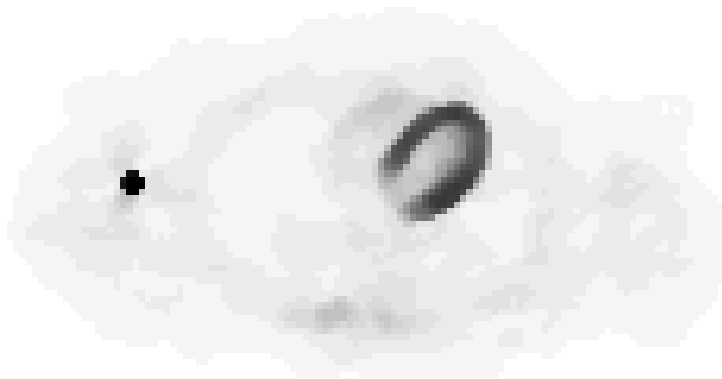
Figure 5.2: Intermediate MRP reconstructions. **Top row:** iterations 5, 10, 20, and 30. **Bottom row:** iterations 40, 50, 60, and 100.

A smoothed FBP image as the first image instead of the disk might save a few iterations in the beginning. However, in order to get rid of FBP-artifacts and develop sharp edges, the number of iterations needs to be quite high. Because MRP can tolerate a high number of iterations, the initial image is not critical to the result. Using a uniform disk as the first image is the simplest and the most neutral choice.

Based on what was mentioned in Section 3.5.1, the OS acceleration method can be used with MRP with moderate number of subsets. With OS factor no more than 4, the MRP penalty can be applied after each sub-iteration. A possible way to use temporal averaging with higher OS factor is shown in Fig. 5.3. The sinogram frame is split into few (3) frames of shorter acquisition times. They are reconstructed separately with a larger OS factor (16 or 32). The sum of the sub-images averages out some of the temporal fluctuations caused by different noise realizations. The common part of the sub-images represent the non-random signal, which is preserved in the summation. As the averaging is carried out in the temporal dimension, there seems to be hardly any spatial blurring. In contrast, FBP as a linear operation deals with the noise in the split and in the total sinograms the same way. This simple experiment suggests, even without considering the OS-factors, that spatial fluctuations due to noise and the "true" image may be distinguished in the time dimension. Noise is dynamic and the spatial activity distribution is nearly static.



(a) MRP, 144 iterations, 3 sec frame.



(b) MRP, 9 iterations with OS=16, 144 total iterations, 1+1+1 sec frame.

Figure 5.3: MRP reconstructions of a 3 second frame and the sum of three 1 second frames of the same acquisition period. The computation time is saved by the factor of $16/3 \approx 5$.

Chapter 6

Results and Images

6.1 Simulations

The effect of noise on the accuracy and the quality of the reconstructed images was examined using a simulated phantom. The phantom was based on the Shepp-Logan phantom [76]. Because a piecewise constant phantom is an unnatural object and maybe too easy for the reconstruction task, a part of the phantom image was made smooth by repetitive filtering using 5×5 , 7×7 , and 9×9 running averages. The phantom was not meant to be the most realistic one for ET, but it serves as a basis for quantitative tests. Methodologically, the reconstruction method should be able to generate any kind of an image.

The phantom was re-projected to a noiseless sinogram. Poisson noise was added to it. Altogether 100 different noise realizations were generated by using a different seed of the random number generator [71]. The deviation was increased by the factor of 4, 6, 8, 10, and 12 at multiples of ten during the 50 first realizations [72]. For the last 50 realizations, the deviation was not changed, just the seed. These sinograms were reconstructed using different algorithms and parameters (see Table 6.1). The aim is to compare MRP with the well known *de facto* standards, FBP and MLEM. Because the various other priors may require careful parameter optimization and more complex implementation techniques, they were excluded from the comparison.

In order to examine the consequences of the selection of reconstruction method and its parameters, the images were checked against the known true phantom image and the statistical properties of the images were examined. Four ROIs were used, see Table 6.1.

Table 6.1: Notations used in Section 6.1.

Abbreviation	Explanation
FBP	FBP with Hann
FBPr	FBP with ramp
ML16	MLEM with 16 iterations
ML48	MLEM with 48 iterations
ML	MLEM with 144 iterations
MRP1	MRP with $\beta = 0.1$ and 144 iterations
MRP3	MRP with $\beta = 0.3$ and 144 iterations
MRP5	MRP with $\beta = 0.5$ and 144 iterations
MRP7	MRP with $\beta = 0.7$ and 144 iterations
MRP9	MRP with $\beta = 0.9$ and 144 iterations
ROI1	329 pixels of a uniform low value
ROI2	3259 pixels of a uniform medium value
ROI3	140 pixels of a uniform high value
ROI4	1960 pixels, the blurred part

6.1.1 Error histograms

Fig. 6.1 shows the histograms of the error images (the reconstructed image – the true phantom) over all 100 images for two different ROIs. The histogram of the error was calculated over the set

$$\left\{ \frac{\lambda_{bi} - t_b}{\bar{t}_{ROI}} 100\% \mid b \in ROI, \forall i \right\}, \quad (6.1)$$

where b is the pixel index, i is the noise realization (sinogram or image) index ($1 \leq i \leq I$), t_b is the true pixel value, and \bar{t}_{ROI} is the true ROI average. A wide histogram means a large variability and weak reliability of the quantitative result of an individual image. A shift of the position of the peak of the histogram away from the vertical line of zero error in Fig. 6.1 indicates a risk of a systematic bias.

Fig. 6.1(a) shows how sensitive the local bias of MLEM is to the number of iterations and to the initial value of the image. When the number of iterations increase, the histogram shifts closer to the zero error position, but at the same time it gets much wider. Interestingly, the fully converged MLEM gives a negative mode for ROI1 and its distribution is roughly as wide as FBP's. MRP's

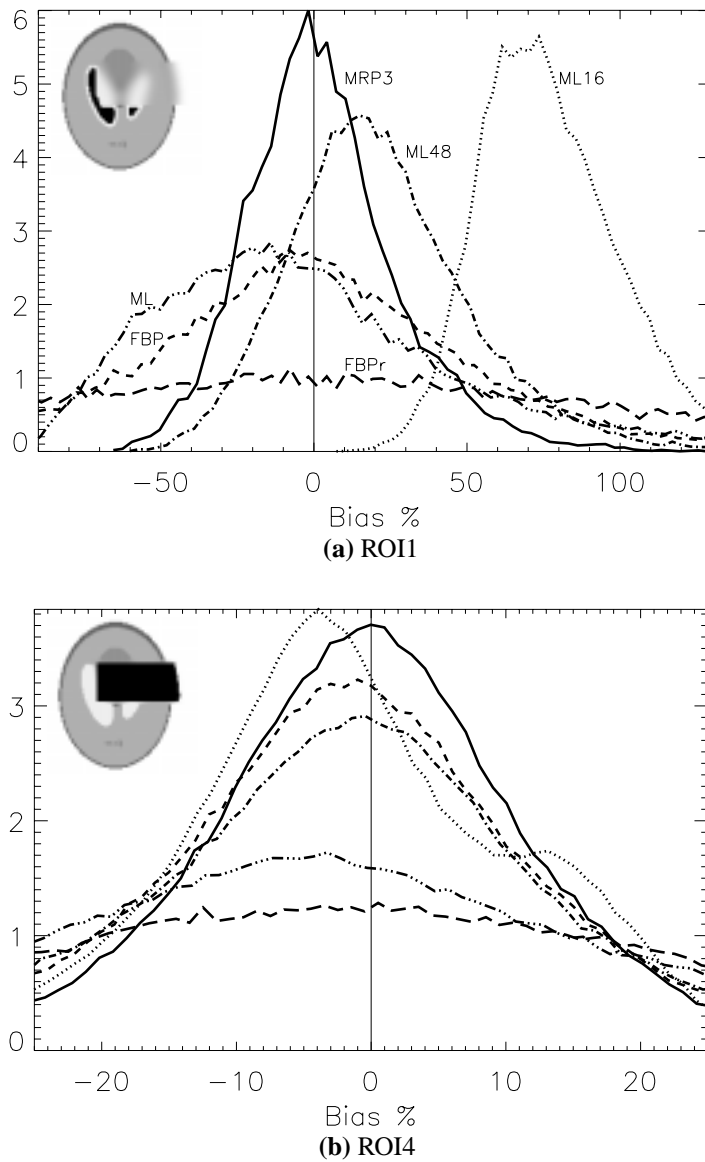


Figure 6.1: Histograms of the error of the modified Shepp-Logan phantom reconstructions. The ROI pixels are marked black on the negative phantom images. Solid line: MRP3. Dashed: FBP. Long dashes: FBP. Dotted: ML16. Dash-dotted: ML48. Dash-dot-dot-dotted: ML. The histogram cell size is 3% for ROI1 and 1% for ROI4. The y-axis shows the probability (%) of error falling in the cell.

error histogram is both narrow and at the right position.

The non-uniform ROI4 in Fig. 6.1(b) demonstrates that also a smooth part of the MRP3 image is quantitatively accurate. This is because the requirement of

local monotonicity includes smoothness. ML16 produces an asymmetric error histogram.

For larger ROIs all methods perform quite well, but the width of the error histogram of ML16 and MRP3 are the two narrowest ones. With more iterations, MLEM is close to FBP or FBPr.

The error histograms for MRP are compact and centered around zero for all ROIs. Different values of β are not shown for MRP because they had only a small effect on the histograms, with $\beta = 0.9$ slightly the most compact. On the other hand, the number of iterations used for MLEM has a strong effect both on the shape and on the location of the error histogram, similar to the difference between FBP and FBPr.

6.1.2 Reliability of ROI estimation

A desirable property of a ET reconstruction method is to produce images from which ROI estimates can be reliably analyzed. It is important to notice that the common ROI average calculation is yet another estimation task after the reconstruction task itself. A sample mean of the ROI pixels as an estimate of the true ROI value is more reliable if the values of the sample pixels form a symmetric distribution. Even if the ROI average is expected to be unbiased, heavy tails of an asymmetric sample distribution increase the possibility of outliers. The outliers affect on the sample mean of an individual realization causing potential bias, especially for a small ROI.

One of the simplest measures of the symmetry of the distribution is the skewness [71]. Fig. 6.2 shows the average skewness values of ROI2 for each reconstruction method. The average skewness was calculated as

$$\overline{skew}(\{\lambda_{bi}|b \in ROI, \forall i\}) = \frac{1}{I} \sum_{i=1}^I \left[\frac{1}{B} \sum_{b=1}^B \left(\frac{\lambda_{bi} - \bar{\lambda}_i}{\sigma_i} \right)^3 \right], \quad (6.2)$$

where $\bar{\lambda}_i$ is the ROI mean of the i th image, and σ_i is the standard deviation of the ROI of the i th image. The closer the skewness is to zero, the more symmetrical the sample distribution is. If the skewness is not close to zero, there is a risk that the ROI average of an individual image is a biased estimator for the desired quantity. ML tends to produce tails to the positive direction. In terms of skewness, MRP performs well. FBP is very good, but FBP may have some local bias (Fig. 6.4(a)).

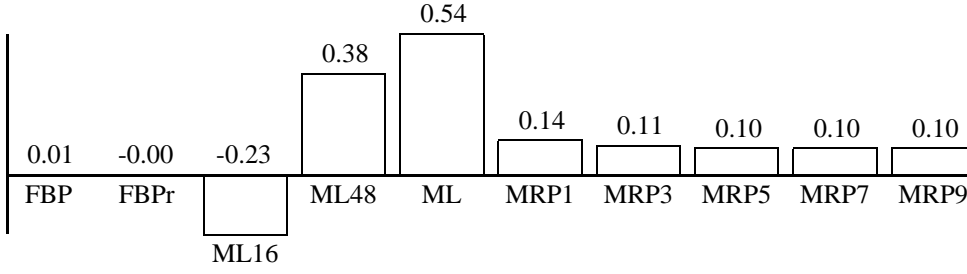


Figure 6.2: The skewness values of a ROI2 averaged over all noise realizations. Skewness values close to zero indicate symmetric distributions and reliable sample means. The approximate limit for significant value of the skewness [71] is $3\sqrt{15/B} = 0.2$.

6.1.3 Noise sensitivity

The estimators may behave in a different way when operating on noisy data. Fig. 6.3 shows MAE for each noise realization. MAE is calculated for each image as

$$MAE(i) = \frac{\sum_{b \in ROI} |\lambda_{bi} - t_b|}{B_{ROI} \bar{t}_{ROI}} 100\% , \quad (6.3)$$

where i is the noise realization index, and B_{ROI} is the number of pixels in the ROI.

In Fig. 6.3 the five different noise deviations used when generating the input data can be seen as an increase of MAE of MLEM at multiples of ten ($i < 50$). MRP3 is less sensitive to the amount of noise in the sinogram. On the other hand, what it comes to MAE, ML serves as a noise amplifier. There are no other MRP reconstructions with different β values plotted in Fig. 6.3 because they are almost indistinguishable from each others. Because the positive and negative errors do not cancel out in MAE, a small MAE value indicates that also the smooth part of the phantom is accurately reconstructed by MRP.

6.1.4 Parameter sensitivity

Fig. 6.4 shows the robustness of the behavior of the reconstruction method. The figures of merit are calculated for each method as

$$bias = \frac{\sum_{b \in ROI} (\bar{\lambda}_b - t_b)}{B_{ROI} \bar{t}_{ROI}} 100\% , \quad (6.4)$$

$$efficiency = \frac{\sigma_{FBP}^2}{\sigma^2} , \quad (6.5)$$

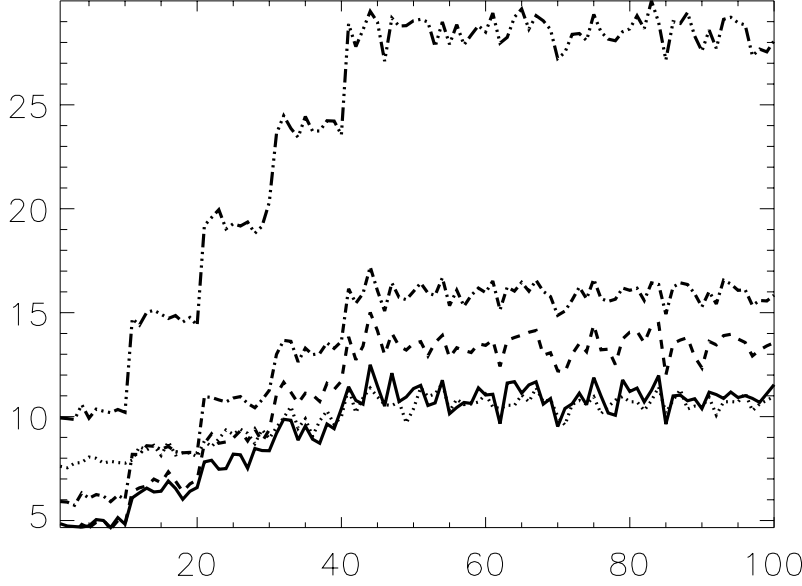


Figure 6.3: The average (%) of MAE of ROI4 for each noise realization. Solid lines: MRP3. Dashed: FBP. Dotted: ML16. Dash-dotted: ML48. Dash-dot-dot-dotted: ML.

$$MSE = bias^2 + \sigma^2, \quad (6.6)$$

and

$$MAE = \frac{\sum_{b \in ROI} |\bar{\lambda}_b - t_b|}{B_{ROI} \bar{t}_{ROI}} 100\%, \quad (6.7)$$

where $\bar{\lambda}_b = (\sum_i \lambda_{bi})/I$ is the sample mean of reconstructed pixels at b in the set of I images, and $\sigma^2 = Var(\lambda_{bi})$ is the variance of IB_{ROI} reconstructed pixels.

In Fig. 6.4(a), the bias is the smallest for MRP, and different β values of MRP make a very small difference. The amount of bias introduced by using a limited number of iterations for MLEM depended on the chosen ROI.

Fig. 6.4(b) depicts the variability of the estimate in terms of the relative efficiency (6.5) compared to FBP. The relative efficiency should be more than 1 in order to perform better than FBP in noise reduction. Only ML16 and MRP with $\beta > 0.1$ achieve this. The price paid for the excellent efficiency of ML16 is the bias.

In Fig. 6.4(c) MSE is shown. MSE takes both the bias and the variance into account, but the bias is so small that the variance is the dominating part. ML16 performs well, but with full number of iterations ML has almost as large MSE as FBPr. Again, MRP performs equally well independently of β . FBP

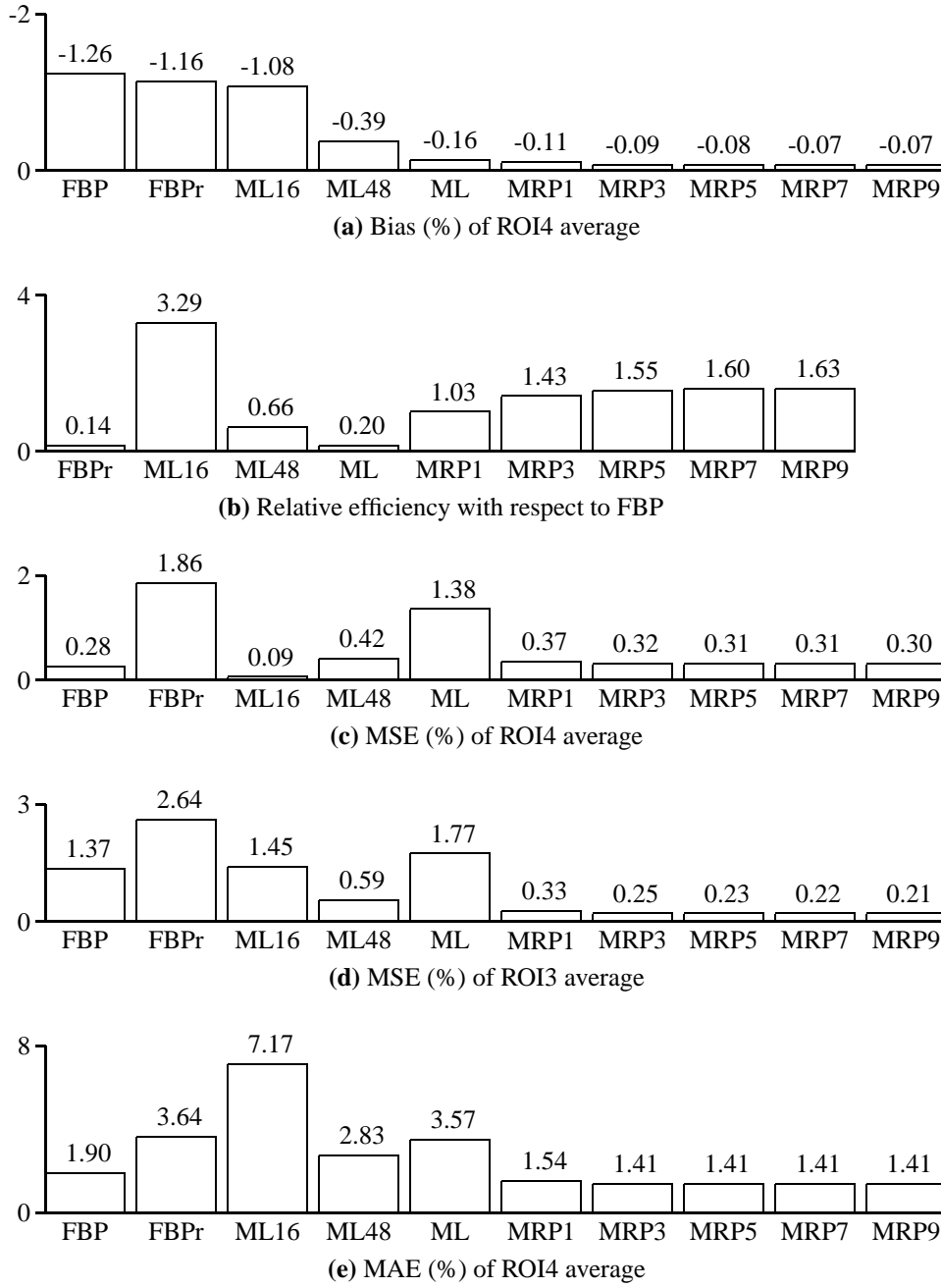


Figure 6.4: Bias, MSE, the relative efficiency, and MAE of each reconstruction method calculated for a ROI. MRP is robust with respect to β .

performs remarkably well. The change of ROI4 to ROI3 in Fig. 6.4(d) worsens FBP and ML16. Thus, the optimal number of iterations for MLEM depends on

the chosen ROI.

Fig. 6.4(e) shows how MAE ranks the methods. Although 6.4(e) and 6.4(c) are computed for the same ROI, the choice of the figure of merit (MSE or MAE) gives different results regarding the optimal number of iterations for MLEM. MRP is more consistent in this sense.

6.1.5 Images

Fig. 6.5 shows images of the phantom study. The first and the second row show the reconstructed images, with a small and a large amount of noise, respectively. ML48 is close to FBP, and ML is close to FBPr. This suggest that the number of iterations as a global regularization method suffers from the same kind of trade-off as FBP's filter functions.

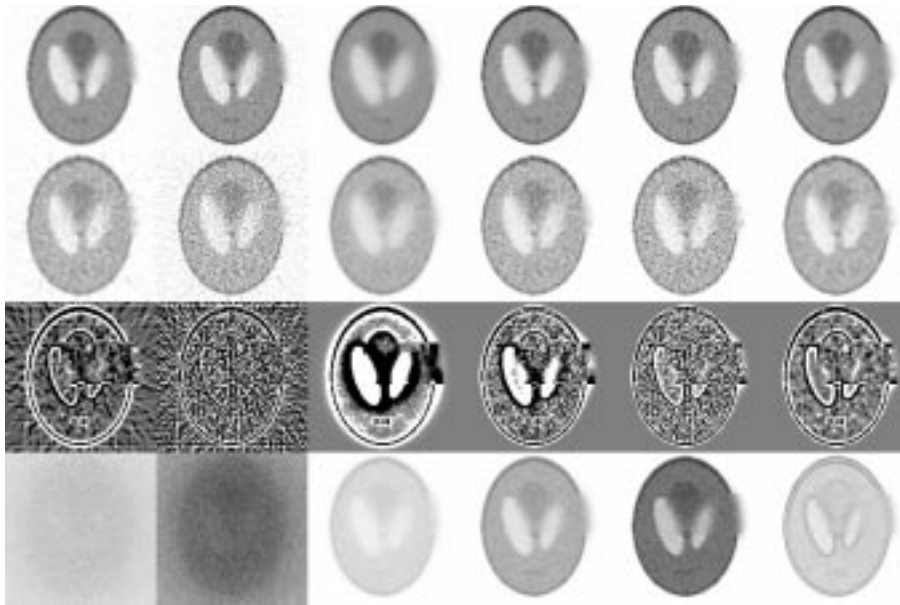


Figure 6.5: Phantom images. From left to right: FBP, FBPr, ML16, ML48, ML, and MRP3. **1st row:** reconstructed images with the smallest amount of noise in the sinogram. **2nd row:** reconstructed images with the largest amount of noise in the sinogram. **3rd row:** average bias images. Pixels with an error larger (smaller) than 10% of the low activity of ROI1 are marked white (black). **4th row:** average pixel-wise standard deviation. Images on each row are scaled for display using the same factor.

On the third row there are the average bias images computed for each pixel with respect to all 100 reconstructions of the noise realizations

$$bias(b) = \frac{\sum_i (\lambda_{bi} - t_b)}{I}. \quad (6.8)$$

Bias values greater than 10% of the activity of ROI1 are marked with white pixels, and negative bias values of the same amount are marked black. Thus, large white or black areas (ML16) are undesirable because they suggest that the pixel values of an individual image may be systematically biased. On the other hand, areas with mixed white and black pixels (ML, FBPr) may be quantitatively accurate on the average across different noise realizations, but the ROI value of an individual image may be sensitive to noise.

The fourth row of Fig. 6.5 shows the pixel-wise standard deviation images computed with respect to the noise realizations, $\sqrt{Var(\lambda_b)}$. ML seems to have a strong data-dependent variance, which makes areas with a high activity quite noisy. MRP3 expresses some edge jitter [10].

The MRP3 images on the rightmost column show some desirable properties. There are no large areas with a systematic bias, and the variability of the estimate is small in general. The MRP images itself may look unconventional, especially if the sinogram is very noisy. But the quantitative properties should overweight the pleasing of the human eye.

6.2 Clinical data

Clinical PET data acquired using ECAT931/08-12 (CTI/Siemens, Knoxville, USA) were reconstructed using MRP (5.1). More clinical images are shown in the publications II – IV.

6.2.1 Emission studies

Fig. 6.6 shows reconstructions of the same study as Fig. 3.3 on page 15. Compared to FBP, the iterative methods do not spread out strong artifacts due to the high activity in the vein. OS factor 4 was used for both MLEM and MRP. The difference between the images in Fig. 6.6 is that the early stoppage of MLEM has performed some global noise reduction, whereas MRP does it locally.

By increasing the dimensions of both the median mask and the image, the spatial size of the mask can be kept the same. Then there are more pixels in the set from which the local median is selected and the appearance of the image is less blocky. Fig. 6.7 shows an MRP image with the mask size increased from 3×3 to 5×5 mask and image dimension from 128 to $128 \times 5/3 \approx 212$. The computation time increases along with the size of the image.

The general form of MRP allows its usage for any organ and for any tracer uptake. In order to illustrate that, Figs. 6.8 and 6.9 show images of legs and

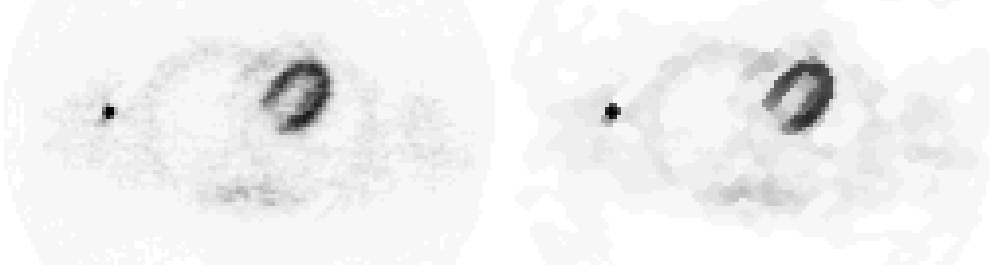


Figure 6.6: PET FDG images. Image dimension = 128. **Left:** MLEM, 5 iterations with OS=4, 20 total iterations. **Right:** MRP with $\beta = 0.3$, 3×3 mask, 36 iterations with OS = 4, 144 total iterations. High activity in the vein does not contribute to other parts of the image. The early stoppage of MLEM may have caused some noise / bias trade-off.



Figure 6.7: MRP with $\beta = 0.11$, mask size 5×5 and image dimension 212, 36 iterations with OS = 4, 144 total iterations. The mask size is spatially the same (1.4 cm) as for the MRP image in Fig. 6.6.

the brain, respectively. The difference images $A - B$ in Fig. 6.8 show excess in A as white pixels and excess in B as black pixels. The difference images show how noise artifacts in the FBP image are removed from the MRP image. MLEM with only a few iterations is not yet very noisy, but the quantitative level in the bone is too high [3].

One of the simplest edge detection operations, Sobel [40], was used in Fig. 6.9 for emphasizing that MRP effectively removes noise without blurring the image. FBP and aborted MLEM do not produce sharp edges where there should

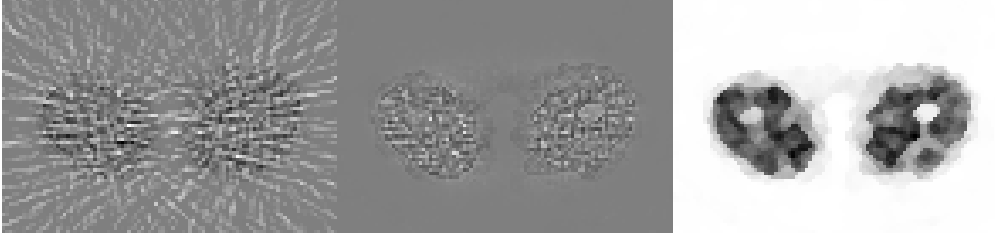


Figure 6.8: PET FDG thigh images. **Left:** Difference image FBP (Hann) – MRP. **Middle:** Difference image MLEM (5 iterations with OS=4, 20 total iterations) – MRP. **Right:** MRP with $\beta = 0.3$, 3×3 mask, 36 iterations with OS = 4, 144 total iterations. Image dimension = 128.

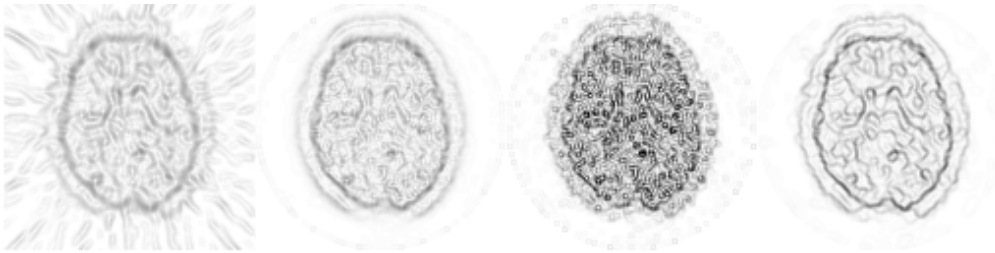


Figure 6.9: Sobel edges of PET FDG brain images. Image dimension = 128. From left to right: FBP; MLEM, 5 iterations with OS=4, 20 total iterations; MLEM, 36 iterations with OS = 4, 144 total iterations; and MRP with 3×3 mask, 36 iterations with OS = 4.

be. Converged MLEM is notoriously noisy, whereas MRP is sharp and expresses a low amount of noise. The good separability of tissue borders facilitates the drawing of ROIs. However, it is noteworthy that there are also low and smooth transitions between different concentration levels. The MRP image as such seems to be suitable for various post-reconstruction operations.

6.2.2 Transmission

The motivation for using tMRP (5.12) for transmission image reconstruction is to avoid the count-limited noise from occurring in the image and, consequently, in the ACFs. Fig. 6.10 shows ACFs computed by the conventional method with a long scan time and by using tMRP with the short time. In spite of the increased count-limited noise in the latter, the ACFs are as good as, if not better than in the conventional case. This was accomplished without any object dependent post-reconstruction operations such as the segmentation of the transmission image

and the tissue type classification [63, 13, 77, 84].

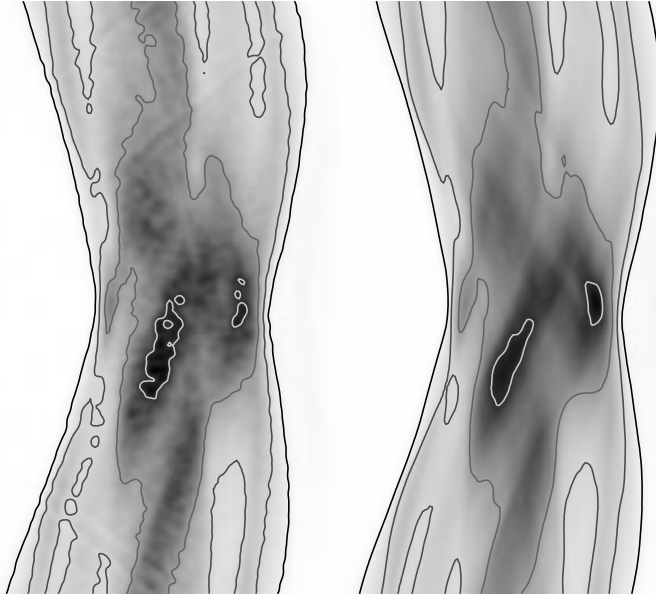


Figure 6.10: PET ACFs. **Left:** conventional, 20 min scan time. **Right:** ACFs computed using tMRP with $\beta = 0.5$, 5×5 mask, 100 iterations, 2 min scan time. The contour lines are drawn at ACF values of 2, 4, 8, and 16.

More results and tMRP transmission images are presented in publication IV [5].

6.2.3 Whole body studies

A single set of acquired sinogram planes covers only 10 – 15 cm in the axial direction. Whole body studies are commonly done without attenuation correction because multiple transmission scans at different axial positions take too much time using the conventional method. The possibility to use short transmission acquisition times and the absence of post-reconstruction operations make tMRP suitable for computing the transmission images and the ACFs for a whole body study. Fig. 6.11 shows coronal transmission images and attenuation corrected emission images. The transmission scan time for each axial position was 2 min. Although the computation time is significantly longer than using the conventional method, it is more easily expandable than the scanner time for the patient.

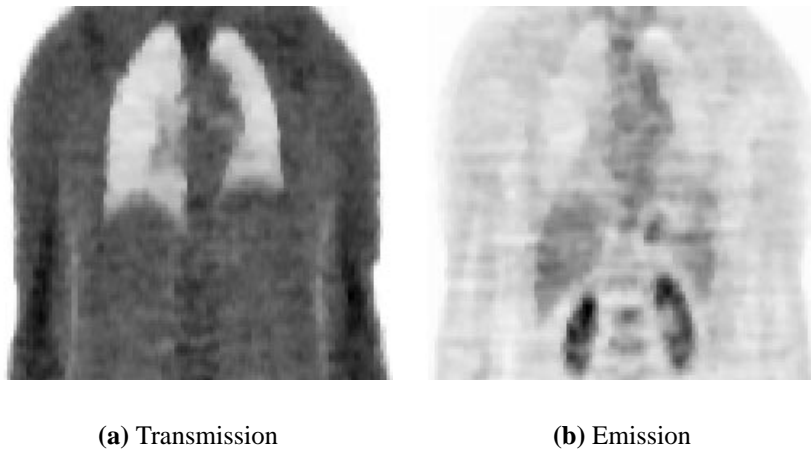


Figure 6.11: Coronal whole body PET images of a clinical study. The images are made by stacking transaxial images. The emission images were corrected for attenuation.

6.3 SPECT images

MRP has been applied for SPECT transmission and emission reconstruction [49]. Fig. 6.12 shows heart emission images acquired using a SPECT device with 64 projection angles / 360° (Siemens Orbiter, Siemens Medical Systems, Inc., Hoffman Estates, Il., USA). No attenuation correction was applied.

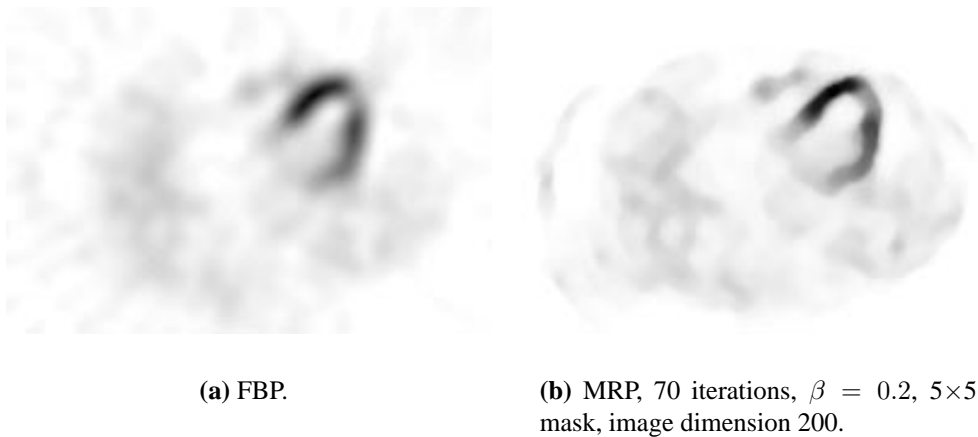
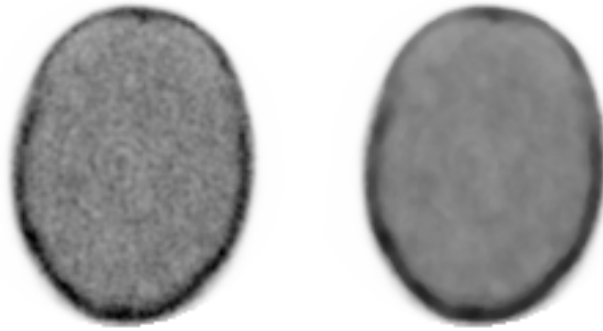


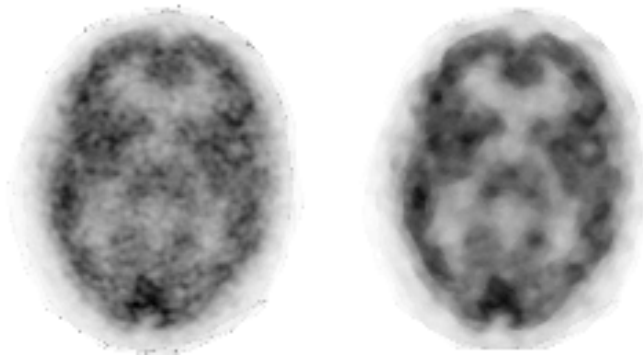
Figure 6.12: SPECT images of a clinical heart study.

Siemens Multispect3 three-head gamma camera with fan-beam collimators and 128 projection angles / 360° was used for quantitative studies with mea-

sured attenuation correction. No attenuation pre-correction was applied on the emission data. Fig. 6.13 compares aborted MLEM and MRP. As in PET, MRP reduces noise without blurring details. MRP is especially suitable for the reconstruction of the attenuation map, Fig. 6.13(a).



(a) Transmission. **Left:** MLEM, 24 iterations. **Right:** MRP, 24 iterations, $\beta = 0.3$, 5×5 mask.



(b) Emission. **Left:** MLEM, 12 iterations. **Right:** MRP, 60 iterations, $\beta = 0.15$, 5×5 mask.

Figure 6.13: SPECT images of a clinical brain study. The transmission images were used as the attenuation maps in the emission reconstruction. Larkin's OSEM package [39] version 5 was modified for MRP.

6.4 Summary of results and the publications

A large portion of the material of the enclosed publications is rewritten for this manuscript for the sake of clarity. However, the detailed results are not repeated

here.

Publication I [1] documents the first appearance of MRP in a scientific forum. The results were tentative but consistent with later examinations. Publication II [2] included more quantitative tests of the new method presented to a wider audience. A stable behavior with respect to iterations were demonstrated. Both a good resolution and a low noise level were possible to achieve at the same time. Publication III [3] further expands the theoretical inspection of the prior. The interpretation for the Gaussian PDF of the prior was formulated. Good convergence properties as well as a small noise / bias trade-off were reported. As a possible step towards visual enhancement, a FIR-median-hybrid filter was suggested to be used instead of the standard median, if more convenient. Publication IV [5] extends the application of MRP to the transmission image reconstruction for attenuation correction. The good tolerance against the count-limited noise made short time transmission scans possible.

The new tests in this section are to demonstrate the insensitivity of MRP against different noise realizations, both in terms of bias and variability of the estimate. As the simulations are free from attenuation, the effects of an imprecise statistical model due to pre-corrections do not contribute here.

The histograms of the pixel-by-pixel errors in Section 6.1.1 showed that MRP produces errors of moderate magnitude. According to the skewness test in Section 6.1.2, the errors form symmetric distributions. This is important in parametric images.

In [15], it is pointed out that a non-convex penalty function resulting from a non-linearity (e.g. threshold) may cause a sudden change in the estimate if the input data is slightly varied. The results in Section 6.1.3 anecdotally indicate that the non-linear and strictly speaking non-convex penalty function of MRP seem not to cause such sudden changes in the pixel values when the noise is varied and increased in the sinogram.

Fig. 6.4 in Section 6.1.4 indicates that MRP is robust with respect to the selected value of the parameter β . MRP performs well in terms of all four figures of merit. Furthermore, the behavior of MRP does not significantly depend on the chosen ROI, which reflects robustness with respect to the activity in other parts of the image.

The different values of β behave in a robust way for all the tests. A moderate value ($\beta \approx 0.3$) lets the image be not exactly locally monotonic, which helps in depicting local extremes in the image.

MRP performs well in light of the variability. Concerning the complexity of many other priors (parameter tuned, anatomical), the possible gain with respect to MRP, if any, may be too small for the extra effort.

Visually MRP images may lack of conventionally common features, such as blurred edges. For pure visual analysis, some post-reconstruction filters may

be used, if more convenient. SPECT might be such a case.

As the results in this section indicate, the MRP images share some quantitatively important features: the bias and the variability of the estimate are small at the same time, and the ROI averages are insensitive both to the input noise and to the neighboring image structure. The results denote that it is safe to use MRP images as a basis of quantitative pixel-by-pixel or ROI analysis.

Chapter 7

Discussion

7.1 General principle

A priori is a term used to identify a type of knowledge which is obtained independently of experience. In MRP, the prior assumption is that an ideal emission image is locally monotonic. Inside a neighborhood of a certain size, there are only monotonic changes in activity between pixels. The monotonicity includes sharp step edges, shallow ramp edges, and arbitrarily smooth edges of any height. This is also the structure of a root signal of the median filter. The root signal is the penalty reference of MRP. Thus, MRP contains implicitly the general description of the unknown tracer concentration. When the spatial size of the neighborhood is reasonably selected and it contains a sufficient number of pixels, the true image is left automatically almost completely unpenalized. No special knowledge of the appearance of the true image is required. This universal nature of the prior makes the definition of the penalty of MRP virtually independent of the organ or the tracer.

The Bayesian interpretation of the penalty of MRP is close to an independent Gaussian prior PDF for the mean activity of the pixel. The location of the heuristic PDF is the median of the neighborhood. Thus, the general description of the true image is modeled using a Markov random field, where the spatial dependence of the pixels is accomplished only by the median. This makes unknown differences between the pixel values permitted in the prior model of MRP.

The problem with mathematical priors is to describe the properties of the true image as a function of the pixel differences in the neighborhood. The quadratic energy function results in a penalty with respect to the difference between the center pixel and the weighted average. The operation is similar to a lowpass filter. As the frequency contents of noise and the true image overlap, a compromise between noise rejection and edge blurring must be done. The

assumption that the unwanted roughness of the image can be expressed using a linear function of adjacent pixel differences is a restriction rather than a generalization.

Other choices for the energy function with tuning parameters are justified in terms of the behavior of the penalty in the vicinity of an edge, but the corresponding Bayesian prior and the MRF are set up in quantitative terms: only certain kind of pixel differences are allowed in the neighborhood. This results in an almost certain penalty to the unknown true image. The resulting image may be a good representative for an *ensemble* of images similar to the true image. But in tomography the key question is "what is the unknown source that gave rise to the *particular* set of measurements we have?". An average of all plausible images may not be an answer to that question.

Priors that apply penalties based on an anatomical image of the target (brain) are more specific to the given task; they rely on the assumption that the true emission image *usually* resembles the given anatomical image in terms of local homogeneity. This may require further assumptions about the tissue types and their *normal* behavior. This, in effect, may result in another good answer to a wrong question.

Multimodal images are important as they reveal relations between various imaging modalities, but with MRP the regularization of the reconstruction process itself can be left "clean" of external data. Then the combination of the two modalities after reconstruction is most informative. From the Bayesian point of view, calling anatomical images as priors is somewhat artificial. That is, there are two sets of measured data, one type of anatomical (CT, or MRI with proton density T1 or T2 [45]) and functional (PET). Using one as a constraint when reconstructing the other might be effective technically in terms of noise reduction, but a Bayes prior should express more fundamental nature of the properties of the image.

The distribution for the mean tracer concentration in a pixel can be interpreted to resemble Gaussian for MRP. The Gaussian or normal distribution is usually considered as the most cautious one when there is no other information. What is more, the central limit theorem suggests that under the influence of many small and unrelated random effects the overall distribution is Gaussian [27, 24, 33]. Despite the fact that the connection between MRP and a Gaussian prior is more intuitive than strictly rigorous, it gives some insight into otherwise analytically intractable MRP algorithm. Taking the location of the Gaussian as the median of neighboring means brings the spatial dependency in a simple, general and robust form into the algorithm.

7.2 Applicability of MRP

MRP has been applied to many clinical studies [44, 81, 73, 49, 47, 53, 54, 50]. The applicability of MRP is good because the prior is very general and simple, applying for both PET and SPECT. MRP is also used in [52, 75].

In contrast to the non-Bayesian ways to control the noise, such as Hann window in FBP or the stoppage of iterations in MLEM, a Bayes prior deals more with specific local properties. The two other methods restrict the noise more globally, which is suboptimal in local terms. A Bayes prior reflects the spatially limited range of the underlying MRF. The reconstructed image is subject to local ROI analysis, and it is important that local characteristics are not affected by global noise reduction. MRP is especially independent of global or external assumptions and parameters. The bias towards the initial image is effectively avoided when MRP is used because the number of iterations can be high enough. The results in Section 6.1 demonstrate that MRP is able to reduce noise and maintain the quantitative accuracy at the same time.

Concerning the practical usability, it is important that the method is not too sensitive to suboptimal choices of parameter values. The parameter β of MRP is the weight of the prior, which controls the strength of monotonicity. The value of the parameter is rather uncritical, which makes MRP easy to use for any study. MRP behaves consistently independently of the chosen ROI, the tracer uptake, or the tissue type.

The quantitative sensitivity of MRP is not compromised by the applied noise reduction because differences in activity with sufficient spatial size are preserved independently of the type of the spatial transition from one activity level to other. This size is an object dependent factor, but it can be selected before the reconstruction according to Eq. (5.13).

For the transmission, the aim was to compute the ACFs directly from the count-limited transmission image reconstructed using tMRP, without any post-reconstruction steps, such as tissue segmentation. The usage of tMRP is straightforward and simple. The principle of local monotonicity is especially suitable for transmission images.

The parametric images are especially sensitive to pixel-by-pixel noise. MRP can be used in reconstructing the whole sequence. Optionally, the parametric mapping can be done before the reconstruction [73]. In that case, the statistical model used in the objective function should be carefully chosen, but the general principle of MRP serves for the regularization of the reconstruction of the parametric sinogram as well.

7.3 Conclusion

The main strength of MRP is its simplicity. A generally accepted principle in science is the Occam's razor (or Ockham), which states that entities must not be multiplied beyond what is necessary, that is, the simplest model with fewest number of assumptions is the one that should be selected [22]. In terms of quality and usability, MRP is very competitive. As a comment to [2], Ph.D. I.G. Zubal writes [86]:

I am convinced that as iterative (and Bayesian) reconstruction techniques become more popular in nuclear medicine, we will need to slowly adjust to a new "class" of images. The reconstruction here shown ... [Fig. 5A in publication II], has a different quality than we are used to seeing in filtered backprojection images. This image, however, is probably a better representation of the true distribution of radioactivity within the patient. We thank the speed of modern day computers for making such improved images possible.

As pointed out by Dr. Zubal, the professionals working with ET images are used to analyze mostly smooth or noisy images. While MRP images express the absence of both noise and blur, they may be visually unlike what is conventionally considered as pleasing to the human eye. The visual aspects can be emphasized with MRP by choosing the image grid density and the mask width carefully, or using a FIR median hybrid in place of the median [3]. However, the modifications of the basic idea of MRP open up a large variety of median-related operations [10], which might be unnecessary. Although the visual aspects are important, they may often reflect individual tastes, which makes it difficult to define a visually good image in general. The quantitative merits of MRP should not be sacrificed for the sake of subjective visual customs.

In spite of the convenient simplicity of MRP, more theoretical analysis of the algorithm still remains an open and challenging problem.

Bibliography

- [1] S. Alenius. Bayesian image reconstruction in positron emission tomography using median root prior. In *The Proceedings of the First TUT Symposium on Signal Processing*, pages 113–116, Tampere, Finland, May 20 1994.
- [2] S. Alenius and U. Ruotsalainen. Bayesian image reconstruction for emission tomography based on median root prior. *European Journal of Nuclear Medicine*, 24(3):258–265, Mar. 1997.
- [3] S. Alenius, U. Ruotsalainen, and J. Astola. Using local median as the location of the prior distribution in iterative emission tomography image reconstruction. *IEEE Transactions on Nuclear Science*, 45(6):3097–3104, Dec. 1998.
- [4] S. Alenius, U. Ruotsalainen, and J. Astola. Using local median as the location of the prior distribution in iterative emission tomography image reconstruction. In *1997 IEEE Medical Imaging Conference, Conference Record*, pages 1726–1730, 1998.
- [5] S. Alenius, U. Ruotsalainen, and J. Astola. Attenuation correction for PET using count-limited transmission images reconstructed with median root prior. *IEEE Transactions on Nuclear Science*, 46(3):646–651, June 1999.
- [6] S. Alenius, U. Ruotsalainen, and J. Astola. Transmission images and attenuation correction with short acquisitions using median root prior reconstruction in PET. In *1998 IEEE Medical Imaging Conference, Conference Record*, pages 1711–1715, 1999.
- [7] M. Alonso and E. Finn. *Fundamental University Physics*, volume III Quantum and Statistical Physics. Addison-Wesley, Reading, Massachusetts, 1983.
- [8] B. Ardekani, M. Braun, and B. Hutton. Improved quantification with the use of anatomical information in PET image reconstruction. In K. Uemura

- et al., editors, *Quantification of Brain Function*, pages 351–359. Elsevier Science Publishers, Amsterdam, 1993.
- [9] B. Ardekani, M. Braun, B. Hutton, I. Kanno, and H. Iida. A fully automatic multimodality image registration algorithm. *Journal of Computer Assisted Tomography*, 19(4):615–623, July/August 1995.
- [10] J. Astola and P. Kuosmanen. *Fundamentals of nonlinear digital filtering*. CRC Press, Boca Raton, 1997.
- [11] D. Bailey. Transmission scanning in emission tomography. *European Journal of Nuclear Medicine*, 25(7):774–787, July 1998.
- [12] T. Barker and J. Young. The QUT digital anatomy lab, Queensland university of technology, Australia. <http://www.dal.qut.edu.au/>, ©1996.
- [13] V. Bettinardi, M. Gilardi, S. Cargnel, G. Rizzo, M. Teräs, and G. Striano. A hybrid method of attenuation correction for positron emission tomography brain studies. *European Journal of Nuclear Medicine*, 21(12):1279–1284, Dec. 1994.
- [14] J. Biemond, R. Lagendijk, and R. Mersereau. Iterative methods for image deblurring. *Proceedings of the IEEE*, 78(5):856–883, May 1990.
- [15] C. Bouman and K. Sauer. A generalized Gaussian image model for edge-preserving MAP estimation. *IEEE Transactions on Image Processing*, 2(3):296–310, July 1993.
- [16] C. Bouman and K. Sauer. A unified approach to statistical tomography using coordinate descent optimization. *IEEE Transactions on Image Processing*, 5(3):480–492, Mar. 1996.
- [17] J. Bowsher, V. Johnson, T. Turkington, R. Jaszczak, C. Floyd, and R. E. Coleman. Bayesian reconstruction and use of anatomical a priori information for emission tomography. *IEEE Transactions on Medical Imaging*, 15(5), Oct. 1996.
- [18] C. Byrne. Iterative image reconstruction algorithms based on cross-entropy minimization. *IEEE Transactions on Image Processing*, 2(1):96–103, Jan. 1993.
- [19] A. Chatziioannou and M. Dahlbom. Detailed investigation of transmission and emission data smoothing protocols and their effects on emission images. *IEEE Transactions on Nuclear Science*, 43(1):290–294, Feb. 1996.

- [20] C. Chen, V. Johnson, W. Wong, X. Hu, and C. Metz. Bayesian image reconstruction in positron emission tomography. *IEEE Transactions on Nuclear Science*, 37(2):636–641, Apr. 1990.
- [21] C. Comtat, P. Kinahan, J. Fessler, T. Beyer, D. Townsend, M. Defrise, and C. Michel. Reconstruction of 3D whole-body PET data using blurred anatomical labels. In *1998 IEEE Medical Imaging Conference, Conference Record*, pages 1651–1655, 1999.
- [22] F. Copleston. *A History of Philosophy*, volume III Ockham to Suárez. Burns and Oates Ltd, London, 1968.
- [23] A. De Pierro. A modified expectation maximization algorithm for penalized likelihood estimation in emission tomography. *IEEE Transactions on Medical Imaging*, 14(1):132–137, Mar. 1995.
- [24] C. Derman, L. J. Gleser, and I. Olkin. *A Guide to Probability Theory and Application*. Holt, Rinehart and Winston, Inc., New York, 1973.
- [25] M. Dwass. *Probability and Statistics*. W.A. Benjamin, Inc., New York, 1970.
- [26] T. Farquhar, A. Chatziioannou, G. Chinn, M. Dahlbom, and E. Hoffman. An investigation of filter choice for filtered back-projection in PET. In *1997 IEEE Medical Imaging Conference, Conference Record*, pages 1042–1047, 1998.
- [27] W. Feller. *An Introduction to Probability Theory and Its Applications*, volume I. Wiley & Sons, Inc., New York, 2nd edition, 1957.
- [28] J. Fessler. Hybrid Poisson/polynomial objective functions for tomographic image reconstruction from transmission scans. *IEEE Transactions on Image Processing*, 4(10):1439–1450, Oct. 1995.
- [29] J. Fessler and N. Clinthorne. *Statistical Methods for Image Reconstruction and Medical Imaging System Design*. 1998 IEEE NSS-MIC tutorial course, Nov. 10 1998.
- [30] J. Fessler, E. Ficaro, N. Clinthorne, and K. Lange. Grouped-coordinate ascent algorithms for penalized-likelihood transmission image reconstruction. *IEEE Transactions on Medical Imaging*, 16(2):166–175, Apr. 1997.
- [31] J. Fessler and A. Hero. Space-alternating generalized expectation-maximization algorithm. *IEEE Transactions on Signal Processing*, 42(10):2664–2677, Oct. 1994.

- [32] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6):721–741, Nov. 1984.
- [33] L. Gonick and W. Smith. *The Cartoon Guide to Statistics*. HarperCollins Publishers, Inc., New York, 1993.
- [34] P. Green. Bayesian reconstructions from emission tomography data using a modified EM algorithm. *IEEE Transactions on Medical Imaging*, 9:84–93, Mar. 1990.
- [35] T. Hebert and R. Leahy. A generalized EM algorithm for 3D-Bayesian reconstruction from Poisson data using Gibbs priors. *IEEE Transactions on Medical Imaging*, 8(2):194–202, June 1989.
- [36] G. Herman. *Image Reconstruction from Projections: Implementation and Applications*. Springer-Verlag, Berlin, 1979.
- [37] G. Herman. *Image Reconstruction from Projections: The Fundamentals of Computerized Tomography*. Academic Press, New York, 1980.
- [38] G. Herman and D. Odhner. Performance evaluation of an iterative image reconstruction algorithm for positron emission tomography. *IEEE Transactions on Medical Imaging*, 10(3):336–346, Sept. 1991.
- [39] H. Hudson and R. Larkin. Accelerated image reconstruction using ordered subsets of projection data. *IEEE Transactions on Medical Imaging*, 13(4):601–609, Dec. 1994.
- [40] A. Jain. *Fundamentals of Digital Image Processing*. Prentice-Hall International, Englewood Cliffs, NJ, 1989.
- [41] R. Jaszczak and E. Hoffman. Scatter and attenuation. In H. Wagner, editor, *Principles of Nuclear Medicine*, chapter 19, section 3. W.B. Saunders Company, Philadelphia, 2nd edition, 1995.
- [42] R. Jaszczak and B. Tsui. General principles. In H. Wagner, editor, *Principles of Nuclear Medicine*, chapter 18, section 1. W.B. Saunders Company, Philadelphia, 2nd edition, 1995.
- [43] K. Johnson and J. Becker. The whole brain atlas. <http://www.med.harvard.edu/AANLIB/home.html>, ©1995-1997.

- [44] M. Kähkönen, S. Alenius, V. Oikonen, and U. Ruotsalainen. Improved quantitative images produced by sinogram Patlak and a new iterative reconstruction method. *Proceedings of the 44th Annual Meeting (San Antonio, USA, June 1-5, 1997), the Society of Nuclear Medicine, Journal of Nuclear Medicine (suppl)*, 38(5):196, 1997.
- [45] A. Kak and M. Slaney. *Principles of Computerized Tomographic Imaging*. IEEE Press, New York, 1988.
- [46] C. Kamphuis and F. Beekman. Accelerated iterative transmission CT reconstruction using an ordered subsets convex algorithm. *IEEE Transactions on Medical Imaging*, 17(6):1101–1105, Dec. 1998.
- [47] C. Katoh, U. Ruotsalainen, H. Laine, S. Alenius, H. Iida, P. Nuutila, and J. Knuuti. Iterative reconstruction based on median root prior in quantification of myocardial blood flow and oxygen metabolism. *Journal of Nuclear Medicine*, 40(5):862–7, 1999.
- [48] L. Kaufman. Maximum likelihood, least squares, and penalized least squares for PET. *IEEE Transactions on Medical Imaging*, 12(2):200–214, June 1993.
- [49] T. Kauppinen. Comparison between iteration and back-projection based on emission tomography reconstruction methods. Lic.Sc. Thesis (in Finnish), Tampere University of Technology, Tampere, Finland, 1997.
- [50] T. Kauppinen, S. Alenius, V. Turjanmaa, J. Kuikka, and M. Koskinen. Comparison between iterative median root prior algorithm and filtered back-projection in SPET. *European Journal of Nuclear Medicine*, 25(8):948, 1998. European Association of Nuclear Medicine (EANM) 1998 abstract.
- [51] S. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice-Hall International, Englewood Cliffs, NJ, 1993.
- [52] G. Kontaxakis, L. Strauss, and G. van Kaick. Optimized image reconstruction for emission tomography using ordered subsets, median root prior, successive substitutions and a web-based interface. In *1998 IEEE Medical Imaging Conference, Conference Record*, pages 1347–1352, 1999.
- [53] H. Laine, M. Knuuti, U. Ruotsalainen, T. Utriainen, V. Oikonen, M. Raitakari, M. Luotolahti, O. Kirvelä, P. Vicini, C. Cobelli, P. Nuutila, and H. Yki-Järvinen. Preserved relative dispersion but blunted stimulation of mean flow, absolute dispersion, and blood volume by insulin in skeletal

- muscle of patients with essential hypertension. *Circulation*, 97(21):2146–53, June 1998.
- [54] H. Laine, M. Luotolahti, M. Knuuti, M. Raitakari, U. Ruotsalainen, H. Iida, J. Kapanen, O. Kirvelä, M. Haaparanta, H. Yki-Järvinen, and P. Nuutila. Insulin resistance in essential hypertension is characterized by impaired insulin stimulation of blood flow in skeletal muscle. *Journal of Hypertension*, 16(2):211–219, Feb. 1998.
- [55] D. Lalush and B. Tsui. Attenuation and detector response compensations used with Gibbs priors distributions for maximum a posteriori SPECT reconstruction. *IEEE Transactions on Nuclear Science*, 39(5):1454–1459, Oct. 1992.
- [56] K. Lange. Convergence of EM image reconstruction algorithms with Gibbs smoothing. *IEEE Transactions on Medical Imaging*, 9(4):439–446, Dec. 1990. Corrections ITMI 10(2) 1991 p. 228.
- [57] K. Lange, M. Bahn, and R. Little. A theoretical study of some maximum likelihood algorithms for emission and transmission tomography. *IEEE Transactions on Medical Imaging*, 6(2):106–114, June 1987.
- [58] K. Lange and R. Carson. EM reconstruction algorithms for emission and transmission tomography. *Journal of Computer Assisted Tomography*, 8(2):306–316, Apr. 1984.
- [59] K. Lange and J. Fessler. Globally convergent algorithms for maximum a posteriori transmission tomography. *IEEE Transactions on Image Processing*, 4:1430–1438, Oct. 1995.
- [60] R. Lewitt. Reconstruction algorithms: Transform methods. *Proceedings of the IEEE*, 71(3):390–408, Mar. 1983.
- [61] B. Lipinski, H. Herzog, R. Kops, W. Oberschelp, and H. Muller-Gärtner. Expectation maximization reconstruction of positron emission tomography images using anatomical magnetic resonance information. *IEEE Transactions on Medical Imaging*, 16:129–136, Apr. 1997.
- [62] J. Llacer, E. Veklerov, K. Coakley, E. Hoffman, and J. Nunez. Statistical analysis of maximum likelihood estimator images of human brain FDG PET studies. *IEEE Transactions on Medical Imaging*, 12(2):215–231, June 1993.

- [63] S. Meikle, M. Dahlbom, and S. Cherry. Attenuation correction using count-limited transmission data in positron emission tomography. *Journal of Nuclear Medicine*, 34(1):143–150, Jan. 1993.
- [64] Merriam-Webster Incorporated. WWWebster dictionary. <http://www.m-w.com/>, ©1996.
- [65] C. Michel, M. Sibomana, A. Bol, X. Bernard, M. Lonneux, M. Defrise, C. Comtat, P. Kinahan, and D. Townsend. Preserving Poisson characteristics of PET data with weighted OSEM reconstruction. In *1998 IEEE Medical Imaging Conference, Conference Record*, pages 1323–1329, 1999.
- [66] T. Moon. The expectation maximization algorithm. *IEEE Signal Processing Magazine*, pages 47–60, Nov. 1996.
- [67] J. Nuyts, P. Dupont, S. Stroobants, A. Maes, L. Mortelmans, and P. Suetens. Evaluation of maximum-likelihood based attenuation correction in positron emission tomography. In *1998 IEEE Medical Imaging Conference, Conference Record*, pages 1836–1841, 1999.
- [68] J. Ollinger and J. Fessler. Positron-emission tomography. *IEEE Signal Processing Magazine*, pages 43–55, Jan. 1997.
- [69] A. Oppenheim and R. Schaffer. *Discrete-Time Signal Processing*. Prentice-Hall International, Englewood Cliffs, NJ, 1989.
- [70] J. Pollard. *Handbook of numerical & statistical techniques*. Cambridge University Press, Cambridge, 1977.
- [71] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery. *Numerical Recipes in C, The Art of Scientific Computing*. Cambridge University Press, Cambridge, second edition, 1992.
- [72] R. Rowe and S. Dai. A pseudo-Poisson noise model for simulation of positron emission tomographic projection data. *Medical Physics*, 19(4):1113–1119, Jul/Aug 1992.
- [73] U. Ruotsalainen. *Quantification and Data Analysis in Positron Emission Tomography: Organ Blood Flow, Graphical Analysis and Radiation Dosimetry*. PhD thesis, Tampere University of Technology, Tampere, 1997.
- [74] K. Sauer and C. Bouman. A local update strategy for iterative reconstruction from projections. *IEEE Transactions on Signal Processing*, 41(2):534–548, Feb. 1993.

- [75] A. Seret. Median root prior and ordered subsets in Bayesian image reconstruction of single-photon emission tomography. *European Journal of Nuclear Medicine*, 25(3):215–219, Mar. 1998.
- [76] L. Shepp and Y. Vardi. Maximum likelihood reconstruction for emission tomography. *IEEE Transactions on Medical Imaging*, 1(2):113–122, Oct. 1982.
- [77] R. Smith, J. Karp, F. Bénard, A. Alavi, E. Gualtieri, G. Muehllehner, and M. Geagan. A comparison of segmentation and emission subtraction for singles transmission in PET. *IEEE Transactions on Nuclear Science*, 45(3):1212–1218, June 1998.
- [78] D. Snyder, M. Miller, L. Thomas, and D. Politte. Noise and edge artifacts in maximum likelihood reconstructions for ET. *IEEE Transactions on Medical Imaging*, 6(3):228–238, Sept. 1987.
- [79] Sorenson. Chapter 9, Quantitative measurement of radioactivity in vivo by whole-body counting. In G. Hine and J. Sorenson, editors, *Instrumentation in nuclear medicine*, volume 2. Academic Press, New York, 1974.
- [80] M. Ter-Pogossian. General principles. In H. Wagner, editor, *Principles of Nuclear Medicine*, chapter 19: Positron Emission Tomography, section 1. W.B. Saunders Company, Philadelphia, 2nd edition, 1995.
- [81] T. Utriainen, P. Nuutila, T. Takala, P. Vicini, U. Ruotsalainen, T. Rönnemaa, T. Tolvanen, M. Raitakari, M. Haaparanta, O. Kirvelä, C. Cobelli, and H. Yki-Järvinen. Intact insulin stimulation of skeletal muscle blood flow, its heterogeneity and redistribution, but not of glucose uptake in non-insulin-dependent diabetes mellitus. *The Journal of Clinical Investigation*, 100(4):777–785, Aug. 1997.
- [82] E. Veklerov and J. Llacer. Stopping rule for the MLE algorithm based on statistical hypothesis testing. *IEEE Transactions on Medical Imaging*, 6:313–319, Dec. 1987.
- [83] W. Wang, C. Goldstein, and G. Gindi. Noise and resolution properties of gamma-penalized likelihood reconstruction. In *1998 IEEE Medical Imaging Conference, Conference Record*, pages 1136–1140, 1999.
- [84] M. Xu, P. Cutler, and W. Luk. Adaptive, segmented attenuation correction for whole-body PET imaging. *IEEE Transactions on Nuclear Science*, 43(1):331–336, Feb. 1996.

- [85] M. Yavuz and J. Fessler. New statistical models for randoms-precorrected PET scans. In *Information Processing in Medical Imaging*, volume 1230 of *Lecture Notes in Computer Science*, pages 190–203. Springer Verlag, Berlin, 1997.
- [86] I. Zubal. An edited abstract of [2]. In A. Gottschalk, I. Zubal, et al., editors, *1998 Year book of nuclear medicine*, chapter 13, pages 416–417. Mosby Inc., St. Louis, 1998. Original paper by Alenius and Ruotsalainen selected for summary and comments.

Publications



Publication I

S. Alenius. Bayesian Image Reconstruction in Positron Emission Tomography using Median Root Prior. *The Proceedings of the First TUT Symposium on Signal Processing '94*, pp. 113-116, 1994, Tampere, Finland.



Publication II

S. Alenius, U. Ruotsalainen. Bayesian image reconstruction for emission tomography based on median root prior. *European Journal of Nuclear Medicine*, vol. 24, pp. 258-265, Mar 1997.



Publication III

S. Alenius, U. Ruotsalainen, J. Astola. Using local median as the location of the prior distribution in iterative emission tomography image reconstruction. *IEEE Transactions on Nuclear Science*, selected papers from the 1997 Medical Imaging Conference (MIC), Albuquerque, New Mexico, November 13-15 1997, vol. 45, no. 6(2), pp. 3097-3107, Dec 1998.

Publication IV

S. Alenius, U. Ruotsalainen, and J. Astola. Attenuation Correction for PET Using Count-Limited Transmission Images Reconstructed with Median Root Prior. *IEEE Transactions on Nuclear Science*, selected papers from the 1998 Medical Imaging Conference (MIC), Toronto, Canada, November 11-14 1998, vol. 46, no. 3(2), pp. 646-651, June 1999.