UNIVERSITY OF KUOPIO, Faculty of Natural and Environmental Sciences
Physics
Medical Physics
Seppänen Aku O.: Correction of Collimator Blurring and Attenuation in Single Photon Emission Computed Tomography
Master of Science thesis, 58 pages,
Supervisors:
Ph.D. Pasi Karjalainen
Ph.D. Jari Kaipio
March 2000

Keywords: Single Photon Emission Computed Tomography (SPECT), collimator, attenuation, least squares estimation, maximum likelihood estimation, inverse problem

Single Photon Emission Computed Tomography (SPECT) is a widely used method in nuclear medicine. In SPECT a radiopharmaceutical is injected into the patient and the radiation emitted by nuclei of radioactive labeled pharmaceuticals is detected using scintillation camera. The scintillation camera is rotated around the patient to form projections at several acquisition angles. The three dimensional distribution of radiopharmaceuticals is reconstructed using these projections.

There are several methods for computing the activity distribution from projections. The traditional method is called the filtered back projection which is a method based on Fourier analysis. Another approach to SPECT problem is the least squares estimation which is a linear algebraic reconstruction method. Modern class of methods are the iterative reconstruction techniques in which the statistical nature of radiation is taken into account. Conventionally the activity distribution is reconstructed slice-by-slice, but the least squares method and the iterative methods are also suitable for fully three dimensional imaging. Mathematically the reconstruction problem of SPECT is an ill-posed inverse problem. The solution of such a problem is not unique or stable without regularization of the problem.

When using least squares estimation or iterative methods in SPECT, all the physical factors affecting to formation of planar images can be taken into account. This is done through modeling the forward problem of SPECT. When the forward problem is modeled the model is used for solving the inverse problem of SPECT

In this thesis we present the modeling of two main physical features having influence on SPECT imaging: the collimator blurring and the attenuation. We consider the collimator correction in the case of conventional slice-by-slice reconstruction and also in the case of fully three dimensional reconstruction. In addition to modeling, a comparison of reconstruction techniques is carried out. The models and the reconstruction techniques are evaluated through various simulations and one real phantom study. The SPECT images are improved when better models are used in reconstruction. We also show that taking into account the statistics of radiation can improve the images considerably.

This thesis is available on `http://venda.uku.fi/~aoseppan/gradu_index.html`

Acknowledgements

This study was carried out in the Department of Applied Physics at the University of Kuopio during 1998-1999.
I want to thank my supervisors Professor Pasi Karjalainen, PhD. and Professor Jari Kaipio, PhD. for their support and guidance. Special thanks I express to Professor Karjalainen for his interest and patience to flux of questions I put forward during this work.
I also thank the official reviewer Samuli Siltanen, PhD.
I thank Vesa Tapio and Aimo Hietanen from North-Carelian hospital for providing me the real measurement data.
I warmly thank my parents Aino and Heimo Seppänen for their love and support.
Most of all I want to thank my dear Viivi, for her love and respect. And for bringing sense to this world of insanity.

Kuopio, March 2000

Aku Seppänen

Abbreviations

ART	Algebraic reconstruction technique
CG	Conjugate gradient
CGLS	Conjugate gradient least squares
CT	Computed tomography
DTFT	Discrete time Fourier transform
EM	Expectation maximization
FBP	Filtered back projection
FFT	Finite Fourier transform *or* Fast Fourier transform
FIR	Finite impulse response
IFFT	Inverse finite Fourier transform
LS	Least squares
MAP	Maximum a posteriori
ML	Maximum likelihood
MN	Minimum norm
MRI	Magnetic resonance imaging
MS	Mean square
PET	Positron emission tomography
SIRT	Simultaneous iterative reconstruction technique
SPECT	Single photon emission computed tomography
SVD	Singular value decomposition
TSVD	Truncated singular value decomposition

Contents

Chapter I

Introduction

Single photon emission computed tomography (SPECT) is a method for nuclear medical diagnostics. When using SPECT the aim is to determine the concentration of some biologically active molecules in a certain region of human body to evaluate some physiological function [38, 48]. The procedure of SPECT is briefly as follows: First radiopharmaceutical is injected into patient. The radiopharmaceutical is specific for some physiological function and it gather to organ of interest, for example myocardium, liver or brains. The radiopharmaceuticals consist of γ-emitting isotope, e.g. Technetium-99m, which is incorporated with physiologically active agent. γ-photons emitted by the radioactive tracer are detected using a scintillation camera, which is rotated around the patient. The photon counts detected at one acquisition angle form a planar image (scintigram). The acquisitions are made at several angles; usually the number of angles varies from 64 to 120. The aim is to compute the three dimensional distribution of the radiopharmaceutical using the information of the planar images.
It is important to notice that radioisotope imaging is not anatomical imaging. Instead, SPECT is a method for functional imaging. Thus the quality of SPECT images is not comparable with anatomical CT or MRI images. Another method using radioisotopes is *positron emission tomography* (PET) [35]. In the PET positron emitters are used as labels. A consequence of positron emission is that a positron annihilates with an electron emitting a pair of 511 keV photons at opposite directions. These are detected by a couple of detectors in the detector ring of scintillation crystals around patient. In the PET less information is lost by the detection system and images provided by PET have better quality than SPECT images. The advantage of SPECT over PET is that it is much cheaper method. PET requires more expensive equipments for both the detection system and for the production of the isotopes [48].
In addition to traditional SPECT imaging there is also some other applications where same kind of measurements are done to reconstruct activity distribution: One is detection of absorbed radiation dose in radionuclide therapy [24], and an another application is detection of nuclear waste [44].

1.1 Formation of planar image

The structure of scintillation camera is presented schematically in Figure 1.1. The main elements of a scintillation camera are a lead collimator, a scintillation crystal and photomultiplier tubes. The purpose of the collimator is to allow only photons having direction parallel to direction of holes of the collimator to pass through it. Therefore collimator is a kind of lens of scintillation camera. Holes of the collimator can be parallel, diverging or converging depending on size of object to be imaged. Photons that pass through the collimator interact with the scintillation crystal. When interacting with the crystal a γ-photon is absorbed and thousands of photons with wavelength of visible light are emitted. These visible light photons are detected by photomultiplier tubes behind the crystal. The output of each photomultiplier tube is an electric signal proportional to the intensity of the light observed by the tube. Consequently the electric signal is proportional to distance between the tube and the point of the photon interaction in scintillation crystal. The signals of the photomultiplier tubes are now used to produce the x- and y-coordinates of the point of the photon interaction. Also, the so called z-pulse is produced. This pulse is proportional to the energy of the detected photon. If the z-pulse is of right magnitude the x- and y-pulses are digitized and used to increment an address location corresponding to x- and y- coordinates in

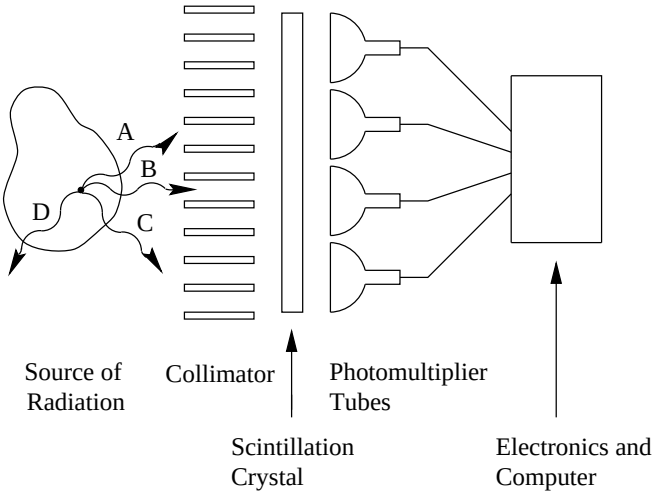


Figure 1.1: Cross section of a scintillation camera. A, B, C and D are the γ-photons emitted from one point. Only photon B has direction parallel to holes of the collimator. Thus only B passes through the collimator and interacts with the crystal causing a photon detection.

the computer memory. When a sufficient amount of photon counts is detected the planar image is ready. The planar image is a discrete map of pixels each pixel having a value that is equal to the number of the photon interactions in the crystal in location of pixel. If the holes of the collimator in scintillation camera are small, one row of pixels in planar image is contributed only by the activity in one fine slice at height of the row of pixels. Thus, approximately, one row of the planar image is a projection of activity in a slice the at same height, and the whole planar image is a projection of the three dimensional activity distribution. Then the formation of one row of the planar image can be written in the integral form

$$P(s,\theta) = \int_{I_{s,\theta}} f(x,y)\mathrm{d}I := (R_\theta f)(s), \tag{1.1}$$

where $P(s,\theta)$ is the value of the pixel in the planar image, θ is the acquisition angle, $f(x,y)$ denotes the activity distribution in the slice, and $I_{s,\theta}$ is a straight line passing through the pixel $P(s,\theta)$ and having direction parallel to the holes of the collimator. R_θ is the integral operator that projects the (two dimensional) activity distribution f to angle θ. $(R_\theta f)(s)$ is called the Radon transform of f and it will be defined in more accurate terms in Chapter 3.
Equation (1.1) models the formation of the planar image mathematically. Later we will refer to this model by *line integral model*, because using this model we assume that the planar images are formed by integrating the activity distribution over straight lines in the direction of the holes of the collimator. We will discuss about the validity of this assumption in Section 1.3, but we already empress that the line integral model is a rough approximation. There are many physical and geometrical issues that must be taken into account when modeling the formation of planar image accurately.

1.2 Single photon emission computed tomography

In the case of tomographic imaging the planar images are acquired at several angles to determine the three dimensional activity distribution. However, traditionally SPECT is not considered as a fully 3D imaging problem. Instead, two dimensional slices of activity distribution are reconstructed. Next we will present the principle of reconstruction of these cross sectional slices. The fully 3D imaging will be presented later in Section 4.4
Figure 1.2 presents the measurement process for SPECT imaging. We consider rows at particular height in every planar image. If we still assume that the holes of the collimator are tiny, these rows are projections of activity in same slice at different angles. The formation of these rows can be described (approximately)

by equation (1.1). If we combine all the pixels in these rows in one vector P, we can write

$$P = Rf(x, y), \tag{1.2}$$

where R is an integral operator obtained by combining the operators R_θ. Computation of projections P using equation (1.2) when $f(x, y)$ is known is the *forward problem* of SPECT, whereas computation of the activity distribution when P is known is the *inverse problem*. We can solve formally the activity distribution in the slice writing

$$f(x, y) = R^{-1}P, \tag{1.3}$$

where R^{-1} stands for the inverse operator of R. However, a problem arises in forming the inverse operator R^{-1}. It is obvious that the inverse operator of R does not really exist, because R is an operator which maps a continuous function $f(x, y)$ to discrete vector P. Some kind of approximate inversion of R has to be done to estimate the activity distribution. Several different approaches to this problem have been presented in literature. We will present the main methods in Chapter 3.

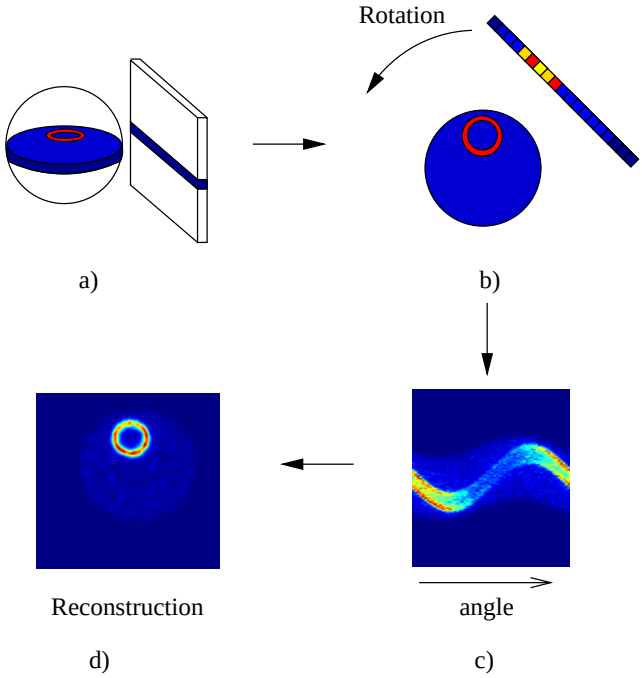


Figure 1.2: Measurement process for SPECT: a) Formation of the projection, one row of planar image is approximately a projection of activity in a slice at the same height. b) A top view of the slice. The scintillation camera is rotated to obtain projections at several angles. c) Projections of the slice at all angles are gathered together and presented in a form of sinogram. d) The reconstruction is computed using the acquired projections.

1.3 Image blurring

In the previous section we claimed that the line integral model described by equation (1.1) is not sufficient for modeling the formation of planar image. There are several physical factors that have effect on SPECT data [22]. In this section we will consider these factors that cause the blurring of planar image, and thus should be taken into account when modeling the forward problem. Later in this thesis we will present modeling of the first three of features described below: *collimator response, attenuation* and *Poisson statistics of radiation*.

1.3.1 Response of collimator

When using the line integral model we assume that the pixels of a planar image are formed integrating the activity over straight lines. This would be true if the holes of the collimator were extremely narrow. But, in the real case, this can not be arranged, because the amount of detected photons per unit time would reduce. This would cause longer acquisition times or higher doses on patient. Diameter of the holes of collimators varies approximately from 1 mm to 5 mm depending on the use of the collimator. The collimator types are named according to the dimensions of the holes: collimators with narrow holes are called high resolution collimators and ones with the holes of larger diameter high sensitivity collimators. Taking into account geometry of the collimator, the forward problem can be modeled more accurately. A pixel in the planar image is actually formed by integrating photons emitted inside a cone instead of a line. Consequently, when modeling the forward problem, the line integral in the equation (1.1) should be replaced by an integral over a cone. Clearly this leads to a genuine 3D problem: every cone includes also parts of the other slices instead of only one slice at the height of the projection bin [42, 45]. We will consider the modeling of collimator in sections 4.2 and 4.4. In section 4.2 we will neglect the effect of photons emitted from other slices. The genuine 3D modeling will be considered in section 4.4.

1.3.2 Attenuation

So far we haven't discussed about photon interaction in the body inside of which the photons are emitted. The principal interaction mechanisms are *electric absorption* and *Compton scattering* for γ-photons in the energy range of interest for medical diagnostics [23]. The interactions of the γ-photons cause attenuation of the photon flux. The intensity of the photon flux reduces exponentially

$$I(l) = I_0 \mathrm{e}^{-\int_0^l \mu(l)\mathrm{d}l}, \tag{1.4}$$

where I_0 is the original intensity, $\mu(l)$ is attenuation coefficient of medium and l is thickness of the medium. Taking into account the attenuation, the forward problem gets a form

$$P(s,\theta) = \int_V f(x)\mathrm{e}^{-\int_{x,(s,\theta)} \mu(l)\mathrm{d}l}\mathrm{d}V, \tag{1.5}$$

where V is a cone shaped volume, x is three dimensional coordinate and $\int_{x,(s,\theta)} \mu(l)\mathrm{d}l$ is the line integral of attenuation distribution over a segment of line connecting point x and projection bin $P(s,\theta)$.
The total mass attenuation coefficient of water is approximately $0.15\mathrm{cm}^2/\mathrm{g}$ for γ-photons having energy 140 keV (Tc-99m) [40]. Thus, the attenuation coefficient is approximately $0.15\mathrm{cm}^{-1}$. Consequently the intensity reduces to one half of original in about 5 cm of water. This implies that attenuation has a strong effect on formation of planar image. Activities near to the camera suffer less attenuation than the distant activities and thus contribute to image more.
The attenuation coefficient is not a constant inside the body. It varies between tissues in the body. In the computation of SPECT the attenuation coefficient is traditionally either assumed to have a constant value [8] or the map of coefficients, the μ-map, is obtained using x-ray imaging (CT) [36, 30, 6]. A new approach to SPECT is to reconstruct both the activity and the attenuation distribution simultaneously from SPECT data alone [4, 37]. This leads to a nonlinear inverse problem which is more ill-conditioned than a basic SPECT problem. In this thesis we will assume that the μ-map of patient is known. The attenuation correction for SPECT will be presented in section 4.3.

1.3.3 Poisson noise

Decay of radioactive isotopes is a statistical process. If f is the activity of the isotope in some region, the actual number of emissions from that region is a random variable A with probability density function $p(A)$. The density function of A can be approximated by Poisson distribution function, that is $A \sim \mathrm{Poisson}(f)$ or:

$$p(A) = \mathrm{e}^{-f}\frac{f^A}{A!}. \tag{1.6}$$

where the parameter f is both the mean and the variance of A. Thus, if we consider the equation (1.5) as observation model of SPECT, we should use A in the equation (1.5) instead of f. This reveals that

the data P is also a random variable. The traditional methods for SPECT imply that the activity of γ-emitting isotope is a deterministic variable. Newer *iterative methods* (see section 3.4) take into account the statistical nature of radioactive decay.
There is also some additive noise involved in nuclear imaging. This is due to the background activity. The additive noise can be considered as measurement error of the imaging system.

1.3.4 Other features

We also describe briefly some other features that cause blurring in SPECT imaging. In this thesis, however, these effects are ignored in the modeling of the forward problem.

- **Scatter:**

  As we noted in 1.3.2 Compton scatter is the other of the main effects that cause attenuation. Unfortunately some of the scattered photons have a new direction parallel to the holes of the collimator of the scintillation camera. Energy of a scattered photon depends on the scattering angle. Using energy window in scintillation camera the photons scattered at wide angle (having low energy) can be rejected. Still, the scattered photons that have larger energies are detected, which causes image blurring. The scatter correction is a subject of number of publications, for example [47, 1, 17]. In this thesis the effect of Compton scatter is ignored.

- **Electronics of scintillation camera:**

  A dead time effect causes errors on images if the count rate is high. In modern cameras the dead time is only a few micro seconds [4], and does not have a remarkable effect on image. Also some other errors may happen due to imperfect camera [48]. The scintillation cameras are tested frequently in hospitals to eliminate these equipment dependent errors.

- **Time varying activity distribution:**

  Usually it is assumed that the activity distribution does not vary during the acquisition. Sometimes this is not a valid assumption and the activity of tracer decreases during the acquisition. For example the half life of Tc-99m is about 6 hours. Compared to this a typical acquisition time, about 20, minutes) is relatively long, and the time dependence should not be ignored. Another reason for the time dependence is that concentration of a radiopharmaceutical can vary during the acquisition. In the dynamic SPECT imaging the aim is to construct a sequence of tomographic images of the time varying activity distribution. Karjalainen *et.al.* [20] have presented a recursive estimation method for dynamic SPECT imaging using the Kalman filter algorithm. In this thesis we will make an assumption of time independence, but the static models presented here can be used as the observation equation in Kalman filter approach for dynamic SPECT.

1.4 Aims of the thesis

There are three aims in this thesis. First, we want to present the estimation theory, where the different methods of SPECT imaging are based on. The second aim is to present the modeling of the two main factors that affect to SPECT imaging, the collimator blurring and the attenuation. The third aim is to compare the different estimation methods in SPECT both theoretically and by simulations.

Chapter II

Estimation and regularization theory

Estimation theory is needed when some physical quantity is tried to be calculated based on a set of measured data. In the case of the SPECT a set of planar images are measured and the task is to calculate the 3D activity distribution. The activity distribution is presented by a finite number of parameters that are to be estimated. In this section we present some issues of the estimation theory. For more details on estimation theory we refer to [19].

2.1 Introduction

We denote the measurements with a vector P and the parameters with a vector f. The task is then to estimate the parameters f based on the measurements P. The estimate of the parameters is denoted by $\hat{f} = \hat{f}(P)$. The parameters f can be interpreted either as random or non-random. In the Bayesian estimation the parameters are assumed to be random variables, whereas in least squares estimation the parameters are treated as deterministic variables.
The observations P are related to the parameters by an observation model

$$P = h(f, v), \tag{2.1}$$

where v are the measurement errors. If the error is additive, the observation model can be written in the form

$$P = h(f) + v. \tag{2.2}$$

Furthermore, if the observation model is linear the equation (2.2) is of the form

$$P = Hf + v, \tag{2.3}$$

where H is the *observation matrix*, that does not depend on the parameters f.

2.2 Bayesian estimation

Assume that the parameters f and measurements P are random vectors, and the joint density of f and P is $p(f, P)$. In the Bayesian estimation the task is to solve the posterior density $p(f|P)$ of the parameters when the observations P are known (measured). Bayesian point estimation means solving some parameters of the posterior density $p(f|P)$. Typical point estimates are the mean and the mode of $p(f|P)$.
Bayes' theorem [19] makes a connection between prior and posterior densities of f and P

$$p(f|P)p(P) = p(P|f)p(f). \tag{2.4}$$

Thus, the posterior density of f is

$$p(f|P) = \frac{p(P|f)p(f)}{p(P)} \tag{2.5}$$

$$\propto p(P|f)p(f). \tag{2.6}$$

The function $p(P|f)$ is called the likelihood of data. It is the density of data P for given parameter f, and it contains the physical model that connects the measurements P to the parameters f. The function $p(f)$ is the density of f. In the Bayesian estimation the prior assumptions of parameters f are included in $p(f)$.

2.3 Maximum a posteriori estimation

The maximum a posteriori estimate $\hat{f}_{\mathrm{MAP}}$ is defined by

$$\hat{f}_{\mathrm{MAP}} = \arg\max p(f|P). \tag{2.7}$$

That is, $\hat{f}_{\mathrm{MAP}}$ is an estimate of f maximizing the posterior density $p(f|P)$. In the other words $\hat{f}_{\mathrm{MAP}}$ is a Bayesian point estimate solving the mode of the posterior density function. Another important Bayesian point estimate is the *Mean Square estimate* $\hat{f}_{\mathrm{MS}}$, which is the solution for the mean of the posterior density. If the posterior density is symmetric and unimodal the estimates $\hat{f}_{\mathrm{MAP}}$ and $\hat{f}_{\mathrm{MS}}$ are equal.

2.4 Maximum likelihood estimation

In the maximum likelihood estimation the parameter f is assumed to be nonrandom vector and the data P is a random vector. Maximum likelihood estimate $\hat{f}_{\mathrm{ML}}$ is defined by

$$\hat{f}_{\mathrm{ML}} = \arg\max p(P|f). \tag{2.8}$$

That is, $\hat{f}_{\mathrm{ML}}$ maximizes the likelihood function $p(P|f)$ for given data P. Thus $\hat{f}_{\mathrm{ML}}$ is an estimate which makes the observed data most probable.
There is also an another interpretation for maximum likelihood estimation. Assume that f is a random vector, and the prior density $p(f)$ has a constant value. This is called an improper prior density, see: [19]. In this case the maximum likelihood estimate equals to the maximum a posteriori estimate. Thus the maximum a posteriori estimation gets the form of maximum likelihood estimation when no prior information of f is used.

2.5 Least squares estimation

Next consider the situation, where the data P is related to non-random parameters f by observation model

$$P = h(f) + v, \tag{2.9}$$

where v is a non-random measurement error. In the least squares (LS) estimation the estimate $\hat{f}_{\mathrm{LS}}$ minimizes the Euclidean norm of the residual, that is

$$\hat{f}_{\mathrm{LS}} = \arg\min ||P - h(f)||. \tag{2.10}$$

Thus, the LS estimate is the minimizer of the index

$$l_{\mathrm{LS}} = ||P - h(f)||^2 = (P - h(f))^T (P - h(f)). \tag{2.11}$$

Nonlinear model

First assume that $h(f)$ is a nonlinear function. To find the minimum of the index l_{LS}, we form the second order Taylor approximation of l_{LS} in the neighborhood of f^*

$$l(f) \approx l(f^*) + \left[\left(\frac{\partial l}{\partial f}(f^*)\right) + \frac{1}{2}(f - f^*)^T \left(\frac{\partial^2 l}{\partial f^2}(f^*)\right)\right](f - f^*) =: g(f). \tag{2.12}$$

If function g is at minimum when $f = \hat{f}$, the gradient of g must vanish at $\hat{f}$. That is

$$\frac{\partial g}{\partial f}(\hat{f}) = \left(\frac{\partial l}{\partial f}(f^*)\right) + (\hat{f} - f^*)^T \left(\frac{\partial^2 l}{\partial f^2}(f^*)\right) = 0 \tag{2.13}$$

and we obtain

$$\hat{f} = f^* - \left(\frac{\partial^2 l}{\partial f^2}(f^*)\right)^{-1} \left(\frac{\partial l}{\partial f}(f^*)\right)^T, \tag{2.14}$$

where the gradient of l is

$$\frac{\partial l}{\partial f}(f^*) = -2\left(P - h(f^*)\right)^T \frac{\partial h}{\partial f}(f^*) \tag{2.15}$$

and the Hessian of l is

$$\frac{\partial^2 l}{\partial f^2}(f^*) = -2\left(\sum_{j=1}^{M}\left(P_j - h_j(f^*)\right)\frac{\partial^2 h_j}{\partial f^2}(f^*)\right) + 2\left(\frac{\partial h}{\partial f}(f^*)\right)^T \left(\frac{\partial h}{\partial f}(f^*)\right). \tag{2.16}$$

Substitution of (2.15) and (2.16) into (2.13) gives a recursion

$$\hat{f}^{(i+1)} = \hat{f}^{(i)} + k_i \left(J_i^T J_i - \sum_{j=1}^{M}\left(P_j - h_j(\hat{f}^{(i)})\right)\frac{\partial^2 h_j}{\partial f^2}(\hat{f}^{(i)})\right)^{-1} \left(J_i^T\left(P - h(\hat{f}^{(i)})\right)\right), \tag{2.17}$$

where $J_i = \frac{\partial h}{\partial f}(\hat{f}^{(i)})$ and k_i is the step size parameter. Recursion (2.17) defines the Newton-Raphson method for solving the nonlinear LS problem. If the Hessian of l is approximated by

$$\frac{\partial^2 l}{\partial f^2}(\hat{f}^{(i)}) \approx 2J_i^T J_i, \tag{2.18}$$

the recursion (2.17) gets the form

$$\hat{f}^{(i+1)} = \hat{f}^{(i)} + k_i \left(J_i^T J_i\right)^{-1} \left(J_i^T\left(P - h(\hat{f}^{(i)})\right)\right), \tag{2.19}$$

that is called the Gauss-Newton method. Approximating the Hessian by identity matrix gives the steepest descent method

$$\hat{f}^{(i+1)} = \hat{f}^{(i)} + k_i \left(J_i^T\left(P - h(\hat{f}^{(i)})\right)\right). \tag{2.20}$$

Linear model

In the case of linear model

$$P = Hf + v \tag{2.21}$$

the gradient of l is simply

$$\frac{\partial l}{\partial f} = -2\left(P - Hf\right)^T H \tag{2.22}$$

from which we obtain the condition for LS estimate

$$(P - H\hat{f}_{\mathrm{LS}})^T H = 0 \tag{2.23}$$

or

$$H^T(P - H\hat{f}_{\mathrm{LS}}) = 0. \tag{2.24}$$

Furthermore

$$H^T H\hat{f}_{\mathrm{LS}} = H^T P \tag{2.25}$$

and we get the linear least squares solution

$$\hat{f}_{\mathrm{LS}} = \left(H^T H\right)^{-1} H^T P \tag{2.26}$$

in the case that the inverse matrix of $H^T H$ exist.

Minimum norm solution

If the matrix $H^T H$ is noninvertible the unique solution for the set of equations (2.25) does not exist. In that case one of the solutions must be chosen, for example the one whose norm is at minimum

$$\hat{f}_{\text{MN}} = H^\dagger P = \arg\min \{||f|| \mid f = \arg\min ||P - h(f)||\}, \tag{2.27}$$

where $H^\dagger$ is the *generalized (Moore-Penrose) inverse* or *pseudoinverse* of H. The estimate $\hat{f}_{MN}$ defined by the equation (2.27) is called the *minimum norm solution* of the least squares problem. If the model is underdeterministic and Rank(H) is equal to number of observations, the pseudoinverse gets a form [28]

$$H^\dagger = H^T (H H^T)^{-1}. \tag{2.28}$$

2.6 Regularization

Many inverse problems are *ill-posed.* The ill-posedness of an inverse problem can be briefly defined so that in ill-posed problems small errors in the data cause big errors to the estimates [10]. In linear case this means that the matrix H is ill-conditioned. That is, the condition number of H is large. The nonlinear inverse problem is ill-posed when the Jacobians J_i are ill-conditioned [46]. In that case the problems must be modified to get a stable solution. Regularization means that the ill-posed problem is modified so that it becomes well-posed.

One regularization method for the ill-posed linear least squares problem is the *truncated singular value decomposition* (TSVD). [10]. The idea of TSVD is as follows. In the case of a well-posed linear least squares problem the minimum norm solution (2.27) can be solved using the *singular value decomposition* (SVD). When using SVD the solution $\hat{f}_{MN}$ is written as a linear combination of the eigenvectors of the matrix $H^T H$. In the TSVD the problem is regularized so that the solution is written as a linear combination of the eigenvectors of the matrix $H^T H$ that correspond to n biggest eigenvalues of the matrix $H^T H$.

Another widely used regularization method is the *(generalized) Tikhonov regularization* [12]. In Tikhonov regularization the minimization problem (2.10) is replaced by a problem

$$\hat{f}_\alpha = \arg\min \left\{ ||P - h(f)||^2 + \alpha ||L(f - f_*)||^2 \right\}, \tag{2.29}$$

where α is a regularization parameter. The term $L(f - f_*)$ is sometimes called the *side constraint.* Obviously the Tikhonov regularized solution of LS problem is such that it tends to minimize both the residual $||P - h(f)||$ and the norm of the side constraint. The regularization parameter α controls the weight of side constraint in minimization.

In the linear case, when $h(f) = Hf$ the solution of the problem (2.29) is equal to the solution of the least squares problem [12]

$$\hat{f}_\alpha = \arg\min ||\tilde{H} f - \tilde{P}||^2, \tag{2.30}$$

where

$$\tilde{H} = \begin{pmatrix} H \\ \sqrt{\alpha} L \end{pmatrix}, \quad \tilde{P} = \begin{pmatrix} P \\ \sqrt{\alpha} L f_* \end{pmatrix}. \tag{2.31}$$

Thus, the formal solution of the regularized problem is

$$\hat{f}_\alpha = \left(\tilde{H}^T \tilde{H} \right)^{-1} \tilde{H}^T \tilde{P}. \tag{2.32}$$

The matrix L contains the prior information of parameters f. For example if the norm of the parameter vector f is known to be small the choice for the side constraint is $L = I$, the identity matrix, and $f_* = 0$. In that case the minimization problem is of the form

$$\hat{f}_\alpha = \arg\min \left\{ ||P - h(f)||^2 + \alpha ||f||^2 \right\}, \tag{2.33}$$

This is called the *standard Tikhonov regularized solution.* Sometimes we can make an assumption that the parameter vector f is smooth. That is, the difference between parameters next to each other small.

In that case the side constraint is chosen so that $L = D_1$, the first difference matrix

$$D_1 = \begin{pmatrix} 1 & -1 & 0 & \dots & 0 & 0 \\ 0 & 1 & -1 & \dots & 0 & 0 \\ \vdots & & & & & \vdots \\ 0 & 0 & \dots & \dots & 1 & -1 \end{pmatrix}. \tag{2.34}$$

Also, for example the second difference matrix D_2 is used in side constraint. The choice of difference matrix is called the *smoothness priors* [5].

2.7 Iterative methods

Solving the linear least squares problem straightforward by equation (2.26) is computationally heavy task if the size of the matrix H is large. For example in SPECT a typical size of matrix H is 15360×16384. In such a case the system of equations (2.25) must be solved iteratively. There are several iterative methods for solving the linear LS problem. Here we will present the idea of three algorithms.

Kaczmarz iteration or ART

Now consider the linear observation model, that is the system of equations (2.21). Let's denote the l^{th} row of the matrix H with H_l. If we neglect the error v, the l^{th} equation of (2.21) is

$$P_l = H_l f. \tag{2.35}$$

The equation (2.35) defines a $(N-1)$-dimensional hyperplane in $\mathbb{R}^N$ for each $l = 1, \dots, M$. If the unique solution of the system of equations exists, it is located in the point where the M hyperplanes intersect. In the overdeterministic case the point where all the hyperplanes intersect does not exist necessarily, and the LS solution of the problem is the point, which is "close to all the hyperplanes". In the underdeterministic case the number of intersection points is infinite, and the unique solution does not exist.
In Kaczmarz iteration the idea is to solve the LS problem in the following way: First make an initial guess $\hat{f}^{(0)} \in \mathbb{R}^N$. The vector $\hat{f}^{(0)}$ is projected at the hyperplane defined by the first equation in (2.21). As a result we obtain a new estimate $\hat{f}^{(1)}$. Next $\hat{f}^{(1)}$ is projected at the hyperplane defined by the second equation in (2.21) and so on. After the M^{th} projection the estimate is projected again at the first hyperplane. This procedure is continued until the residual is small enough.
The projection of $\hat{f}^{(i)}$ at the l^{th} hyperplane is [18]

$$\hat{f}^{(i+1)} = \hat{f}^{(i)} + \frac{1}{||H_l||^2} H_l^T (p_l - H_l \hat{f}^{(i)}). \tag{2.36}$$

This equation defines the Kaczmarz iteration also known as algebraic reconstruction technique (ART). In the underdeterministic case the point of convergence depends on the initial guess.
It is worth to notice that

$$\frac{1}{||H_l||^2} H_l^T (p_l - H_l \hat{f}^{(i)}) \in R(H^T) = N(H)^{\perp}. \tag{2.37}$$

That is, the correction term in the iteration belongs to the orthogonal complement of the null space of H. This reveals that the component of initial guess $\hat{f}^{(0)}$ belonging $N(H)$ remains the same during the iteration. As a consequence, with an initial guess $\hat{f}^{(0)} = (0, \dots, 0)^T$ the ART algorithm converges to minimum norm solution.

SIRT

Simultaneous iterative reconstruction technique (SIRT) [18] has basicly the same principle as ART in solving the LS problem. The exception is that in SIRT in each iteration the estimate $\hat{f}^{(i)}$ is projected at all M hyperplanes using equation (2.36) and then all the obtained correction terms added to $\hat{f}^{(i)}$ simultaneously. Clearly the SIRT iteration can be written in a form

$$\hat{f}^{(i+1)} = \hat{f}^{(i)} + \left(\text{diag}(H^T H)\right)^{-1} H^T (P - H\hat{f}^{(i)}), \tag{2.38}$$

where we have noted that

$$\mathrm{diag}(H^T H) = \begin{pmatrix} H_1^T H_1 & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & H_M^T H_M \end{pmatrix} = \begin{pmatrix} ||H_1||^2 & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & ||H_M||^2 \end{pmatrix} \tag{2.39}$$

Comparison of Gauss-Newton iteration (2.19) and SIRT (2.38) reveals an interesting point. In the case of the linear observation model the Jacobian J in equation (2.19) is $J = H$. In that case the Gauss-Newton iteration gets a form

$$\hat{f}^{(i+1)} = \hat{f}^{(i)} + k_i (H^T H)^{-1} H^T (P - H\hat{f}^{(i)}) \tag{2.40}$$

from which we see that actually SIRT approximates the Gauss-Newton algorithm with $H^T H \approx \mathrm{diag}(H^T H)$.

Conjugate gradient iteration

Conjugate gradient (CG) algorithm is also an iterative method for solving a system of equations. The implementation of the CG algorithm is presented in [12] and [21]. The CG algorithm solves a system $P = Hf$, where H is positive definite and symmetric. However, the algorithm can be modified so that it solves the least squares problem. In that case the CG algorithm is applied to normal equations $H^T Hf = H^T P$. The resulting conjugate gradient least squares (CGLS) algorithm is included in the regularization toolbox for MATLAB.
In CG algorithm the i^{th} iteration provides the following estimate $\hat{f}^{(i)}$

$$\hat{f}^{(i)} = \arg\min ||Hf - P|| \quad \text{subject to} \quad f \in \mathcal{K}_i(H^T H, H^T P), \tag{2.41}$$

where

$$\mathcal{K}_i(H^T H, H^T P) = \mathrm{span}\{H^T P, H^T H H^T P, \dots, (H^T H)^{i-1} H^T P\} \tag{2.42}$$

is the Krylov subspace. In CG algorithm the low-frequency components of the solution tend to converge faster than the high-frequency components [12]. Actually the CG algorithm approximates the truncated SVD method mentioned in the beginning of section 2.6. Thus the number of iterations serves as a regularization parameter in CG algorithm.

Chapter III

Reconstruction methods in SPECT

There are three different approaches for solving the SPECT problem. The traditional method is the *filtered back projection*. It is based on a result called the Fourier slice theorem. The main idea is to filter the projections P_i in the frequency domain separately and form the reconstruction backprojecting those filtered projections.
The second approach to SPECT problem are the *algebraic reconstruction techniques*. The formation of the projections is then modeled by a set of linear equations. The reconstruction of the activity distribution is then obtained by calculating the least squares solution for the set of equations. The LS problem is usually an ill-posed inverse problem and some requlariza tion is needed. Also, a straightforward solution for LS problem is computationally heavy and iterative methods are used. The algebraic reconstruction techniques, such as ART and SIRT are methods for solving the LS problem using different types of iterations.
In the newest class of methods the Poisson statistics of radiation is taken into account. The solution for activity distribution is obtained by Bayesian point estimation: maximum a posteriori or maximum likelihood estimation. The methods for Bayesian point estimation are iterative and they are called the *iterative methods* for SPECT.

3.1 Filtered back projection

In this section we present the filtered back projection method. First we derive the solution in the continuous case following the presentation in [2]. After that the discrete form of filtered back projection method is implemented.

3.1.1 The Radon transform

First we define the Radon transform Rf in continuous case using notations in Figure 3.1. If we project the activity distribution $f(x)$ to given direction $\theta^\perp$, which is orthogonal to θ, we get

$$(R_\theta f)(s) = \int_{-\infty}^{\infty} f(s\theta + t\theta^\perp)\mathrm{d}t. \tag{3.1}$$

The Radon transform is now obtained combining the projections to all directions, that is

$$(Rf)(s\theta) = (R_\theta f)(s). \tag{3.2}$$

3.1.2 The Fourier slice theorem

It is easy to prove [18] that the Fourier transform of the projection defined in equation (3.1) is equal to (two dimensional) Fourier transform of f in the straight line passing through the origin and having direction θ, that is

$$(\widehat{Rf})(\omega) = \hat{f}(\omega\theta). \tag{3.3}$$

This is the Fourier slice theorem. The meaning of equation (3.3) is illustrated in figure 3.2.

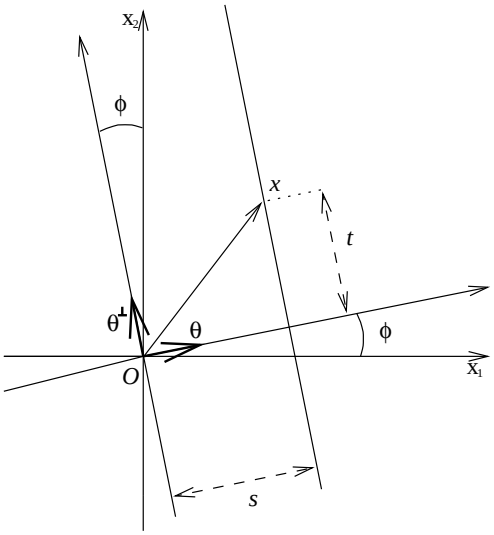


Figure 3.1: The geometry in the Radon transform

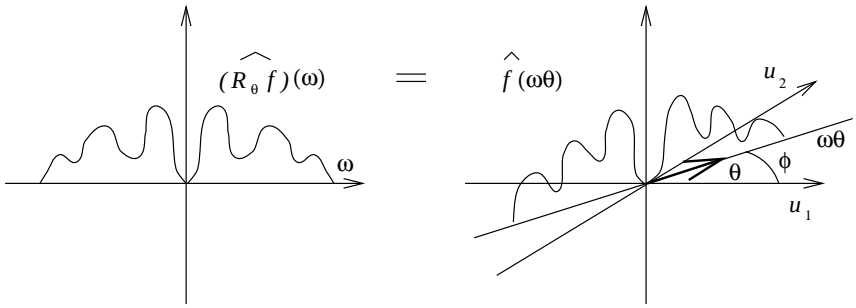


Figure 3.2: Geometrical presentation of the Fourier slice theorem

3.1.3 Implementation of the filtered back projection method

The two dimensional Fourier transform of function f is defined by equation

$$\hat{f}(u) = \int \mathrm{e}^{-iu\cdot x} f(x)\mathrm{d}x, \tag{3.4}$$

where $x = (x_1, x_2)$ and $u = (u_1, u_2)$. The inverse Fourier transform of $\hat{f}$ is

$$f(x) = \frac{1}{(2\pi)^2} \int \hat{f}(u)\mathrm{e}^{ix\cdot u}\mathrm{d}u. \tag{3.5}$$

Using polar coordinates (ω, ϕ), so that $u = \omega(\cos\phi, \sin\phi) = \omega\theta$, we get $du = \omega d\omega d\phi$ and equation (3.5) becomes

$$f(x) = \frac{1}{(2\pi)^2} \int_0^{2\pi} \left(\int_0^{\infty} \omega \hat{f}(\omega\theta)\mathrm{e}^{i\omega\theta\cdot x} d\omega \right) \mathrm{d}\phi. \tag{3.6}$$

Replacing ϕ with $\phi + \pi$ means replacing θ with $-\theta$. Thus, using the fact that $\hat{f}(\omega\theta)$ is periodic with period 2π, we can write

$$f(x) = \frac{1}{(2\pi)^2} \int_0^{2\pi} \left(\int_0^{\infty} \omega \hat{f}(-\omega\theta)\mathrm{e}^{-i\omega\theta\cdot x}\mathrm{d}\omega \right) \mathrm{d}\phi. \tag{3.7}$$

This equals with

$$f(x) = \frac{1}{(2\pi)^2} \int_0^{2\pi} \left(\int_{-\infty}^{0} |\omega| \hat{f}(\omega\theta)\mathrm{e}^{i\omega\theta\cdot x}\mathrm{d}\omega \right) \mathrm{d}\phi. \tag{3.8}$$

If we add equations (3.6) and (3.8), we obtain

$$f(x) = \frac{1}{2(2\pi)^2} \int_0^{2\pi} \left(\int_{-\infty}^{\infty} |\omega| \hat{f}(\omega\theta) \mathrm{e}^{i\omega\theta \cdot x} \mathrm{d}\omega \right) \mathrm{d}\phi. \tag{3.9}$$

Now, let's assume that we have measured the (continuous) projections of f. We denote the projections by $P(s, \theta)$. In this case P is exactly the Radon transform of f, that is, $P(s, \theta) = (Rf)(s, \theta)$. Thus we can use the Fourier slice theorem (3.3), and equation (3.9) gets a form

$$f(x) = \frac{1}{2(2\pi)^2} \int_0^{2\pi} \left(\int_{-\infty}^{\infty} |\omega| \hat{P}(\omega, \theta) \mathrm{e}^{i\omega\theta \cdot x} \mathrm{d}\omega \right) \mathrm{d}\phi. \tag{3.10}$$

Furthermore, we can write the equation (3.10) in the following form

$$f(x) = \frac{1}{4\pi} \int_0^{2\pi} \tilde{P}(s, \theta) \mathrm{d}\phi, \tag{3.11}$$

where $\tilde{P}$ is (note that $s = \theta \cdot x$):

$$\tilde{P}(s, \theta) = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\omega| \hat{P}(\omega, \theta) \mathrm{e}^{i\omega s} \mathrm{d}\omega. \tag{3.12}$$

We see immediately, that equation (3.12) presents high pass filtering of the projection P. Thus we call $\tilde{P}$ the filtered projection.
It was proved in [2], that operator R^* defined by

$$(R^* g)(x) = \int_0^{2\pi} g(x \cdot \theta, \theta) \mathrm{d}\phi \tag{3.13}$$

is the adjacent operator of the Radon transform operator defined in equation (3.2). The operator R^* is called the back projection operator, because it backprojects g to the point x over the straight lines passing through x.
Using (3.13) we can write equation (3.11) in form

$$f(x) = \frac{1}{4\pi} (R^* \tilde{P})(x), \tag{3.14}$$

which shows us that f is obtained by backprojecting the filtered projections. This gives us reason to call the method described by the equation (3.10) the filtered back projection.

3.1.4 FFT and filtering in matrix form

The next aim is to get the discrete version of the filtered back projection. For this purpose we present the filtering process in discrete case.
First we will present the finite Fourier transform (FFT) in matrix form. Let's consider a vector $g = (g_0, \ldots, g_{N-1})^T$. When using FFT, the idea is to present g as a linear combination of functions $e^{ikl2\pi/N}$, where $k, l = 0, \ldots, N-1$. That is,

$$g_l = \sum_{k=0}^{N-1} \hat{g}_k e^{ikl2\pi/N}, \ l = 0, \ldots, N-1. \tag{3.15}$$

We denote $\omega := 2\pi/N$ and write

$$\begin{pmatrix} g_0 \\ g_1 \\ \vdots \\ g_{N-1} \end{pmatrix} = \begin{pmatrix} 1 & 1 & \ldots & 1 \\ 1 & e^{i\omega} & \ldots & e^{i(N-1)\omega} \\ 1 & e^{i2\omega} & \ldots & e^{i(N-1)2\omega} \\ \vdots & & & \\ 1 & e^{i(N-1)\omega} & \ldots & e^{i(N-1)(N-1)\omega} \end{pmatrix} \begin{pmatrix} \hat{g}_0 \\ \hat{g}_1 \\ \vdots \\ \hat{g}_{N-1} \end{pmatrix}, \tag{3.16}$$

that is

$$g = \Psi \hat{g}. \tag{3.17}$$

We notice that the columns of the matrix Ψ are orthogonal, and thus the inverse matrix of Ψ exists. It is easy to see, that the inverse matrix of Ψ is

$$\Psi^{-1} = \frac{1}{N}\Psi^T = \frac{1}{N}\begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & e^{-i\omega} & e^{-i2\omega} & \dots & e^{-i(N-1)\omega} \\ \vdots & & & & \\ 1 & e^{-i(N-1)} & e^{-i(N-1)2\omega} & \dots & e^{-i(N-1)(N-1)\omega} \end{pmatrix}. \tag{3.18}$$

Consequently, we can write

$$\begin{aligned} \hat{g} &= \Psi^{-1} g \\ &= \frac{1}{N}\Psi^T g. \end{aligned} \tag{3.19}$$

This is the discrete time Fourier transform (DTFT) of g, but it is also called the finite Fourier transform (FFT) [18]. The fast Fourier transform, also denoted by FFT, is a fast method for calculating (3.19). Correspondingly, the equation (3.17) defines the inverse finite Fourier transform, IFFT.
The Filtering of a signal means that the signal g is converted in some system. The result of the conversion is a filtered signal $\tilde{g}$. Now assume that the signal g is discrete. In the case of a finite impulse response (FIR) system the filtering of g can be written in the form

$$\tilde{g}_t = \sum_{l=0}^{N-1} a_l g_{t-l}, \tag{3.20}$$

where $a = (a_1, \dots, a_{N-1})^T$ is the impulse response of the system. Equation (3.20) can be written in the matrix form

$$\tilde{g} = Wg, \tag{3.21}$$

where

$$W = \begin{pmatrix} a_0 & a_1 & a_2 & \dots & a_{N-1} \\ a_{N-1} & a_0 & a_1 & \dots & a_{N-2} \\ \vdots & & & & \\ a_1 & a_2 & a_3 & \dots & a_0 \end{pmatrix} \in \mathbb{C}^{N\times N} \tag{3.22}$$

Now assume that we want to construct a discrete system with a given frequency response $A(\omega)$. In the case of a FIR system we can approximate the frequency response of a filter as the finite Fourier transform of the impulse response of the system a. That is,

$$A_k = \sum_{l=0}^{N-1} a_l e^{-ikl\omega}, \tag{3.23}$$

or

$$A = \frac{1}{N}\Psi^T a, \tag{3.24}$$

where $A = (A_1, \dots, A_{N-1})^T$. Thus, if $A(\omega)$ is given, we can sample the values A_k to form a vector A and then compute the impulse response vector using the IFFT of A

$$a = \Psi A. \tag{3.25}$$

Next, the matrix W can be constructed using the equation (3.22) and the filtering of the vector g can be performed by using the equation (3.21).

3.1.5 The discrete version of the filtered back projection

In 3.1.3 we assumed that the projection $P(s,\theta)$ is a continuous function of s and θ. In real cases we have discrete projections at a limited number of acquisition angles. Thus the filtered back projection method must be modified for the discrete case.
In discrete case we try to calculate the activity in a set of pixels. The activity distribution can now be presented by a vector $f = (f_1, \ldots, f_N)^T$, where N is the number of pixels. We denote the discrete projection vector at the first angle by $P_1 = (p_1, \ldots, p_L)^T \in \mathbb{R}^{L\times 1}$, where L is the number of bins in one projection. Similarly we denote the vectors $P_2, \ldots, P_M$ corresponding the other angles. For further use we present the projection vector P formed by combining the projections at all angles:

$$P = \begin{pmatrix} P_1 \\ \vdots \\ P_M \end{pmatrix} = \begin{pmatrix} p_1 \\ \vdots \\ p_B \end{pmatrix} \in \mathbb{R}^{B\times 1}, \tag{3.26}$$

where $B = M \cdot L$ is the total number of projection bins.
Now we can present the discrete Radon transform

$$\tilde{H}f = \begin{pmatrix} \tilde{H}_{11} & \ldots & \tilde{H}_{1N} \\ \vdots & \ddots & \\ \tilde{H}_{B1} & \ldots & \tilde{H}_{BN} \end{pmatrix} \begin{pmatrix} f_1 \\ \vdots \\ f_N \end{pmatrix}. \tag{3.27}$$

In the equation (3.27) the matrix $\tilde{H}$ is line integral modeled observation matrix. That is, $\tilde{H}_{ij}$ is proportional to the area of j:th pixel covered by the 'fat' line passing through the i:th projection bin and having direction $\theta^{\perp}$ defined in figure 3.1.
In the continuous case the backprojection operator R^* was noted to be the adjacent operator of the Radon transform operator. The discrete Radon transform operator is the real valued matrix $\tilde{H}$ and thus the discrete backprojection operator is $\tilde{H}^T$, the transpose of the discrete Radon transform operator. Thus, we get the discrete version of the equation (3.14)

$$\hat{f}_{\mathrm{FBP}} = \tilde{H}^T \tilde{P}, \tag{3.28}$$

where $\tilde{P}$ is the filtered projection vector.
When using the filtered back projection, each projection P_i is filtered separately as we see from equation (3.12). In continuous case we obtained that the frequency response of the filter is $|\omega|$, see (3.12). But a filter normally used differs from that, because the noise is lowered filtering the highest frequencies from the projections. Figure 3.3 shows the frequency response $A(\omega)$ of a typical filter used in the filtered back projection.

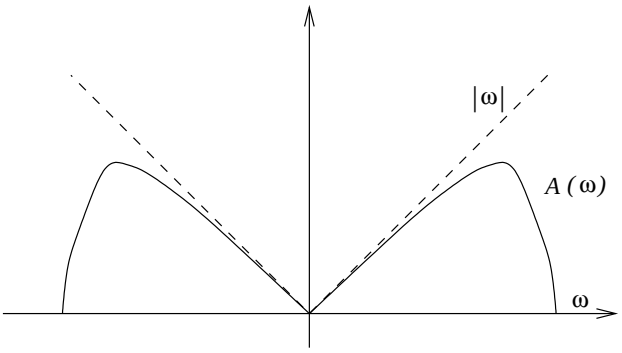


Figure 3.3: Frequency response of a typical filter used in the filtered back projection

We can write the filtering of each projection in the form

$$\tilde{P}_i = WP_i, \tag{3.29}$$

where W is obtained from equations (3.22) and (3.25). Now we get the filtered projection $\tilde{P}$ combining $\tilde{P}_i$'s

$$\tilde{P} = \begin{pmatrix} \tilde{P}_1 \\ \vdots \\ \tilde{P}_M \end{pmatrix}. \tag{3.30}$$

By combining equations (3.28)-(3.30) we get the discrete filtered back projection in matrix form

$$\hat{f}_{\mathrm{FBP}} = \tilde{H}^T \begin{pmatrix} W & & & & \\ & \ddots & & 0 & \\ & & W & & \\ & 0 & & \ddots & \\ & & & & W \end{pmatrix} \begin{pmatrix} P_1 \\ P_2 \\ \vdots \\ P_M \end{pmatrix}, \tag{3.31}$$

or

$$\hat{f}_{\mathrm{FBP}} = \tilde{H}^T \Omega P. \tag{3.32}$$

Thus, in filtered back projection the filtering process of the projection P is equal to multiplying the projection by Ω. The matrix Ω is a block matrix, in which all the diagonal blocks are equal to W and all the off-diagonal blocks are all-zero matrices. Furthermore, all the rows of the diagonal blocks W are obtained by shifting the first row of W.
It is worth to notice that when using filtered back projection, the reconstruction is not computed in the matrix form in practice. Instead, filtered back projection is calculated using the fast Fourier transform (FFT) and inverse fast Fourier transform (IFFT), see [18]. The idea of matrix presentation in this thesis is to illustrate the meaning of filtered back projection. Equation (3.31) shows clearly how the filtered back projection algorithm actually calculates the estimate: the discrete data vectors P_i are first filtered and then backprojected along straight lines having direction parallel to holes of collimator.

3.1.6 Attenuated Radon transform

If the attenuation is implemented in Radon transform (3.1), the projection gets a form

$$A(f,\mu)(s\theta) = \int_{-\infty}^{\infty} f(s\theta + t\theta^{\perp}) \mathrm{e}^{-\int_t^{\infty} \mu(s\theta+\tau\theta^{\perp})\mathrm{d}\tau} \mathrm{d}t, \tag{3.33}$$

where $A(f,\mu)$ is called the exponential Radon transform [18] or the attenuated Radon transform [4]. Similarly as in the case of Radon transform the reconstruction can be calculated by filtered back projection method based on the attenuated Radon transform. It appears [18] that in this case the reconstruction is computed so that the filtered projections are backprojected and the obtained pixel values are multiplied by a distance dependent exponential weight.

3.2 Algebraic reconstruction

Now we take a linear algebraic approach to SPECT problem. We model the projections by a set of linear equations. Now the solution of the reconstruction problem is obtained by the LS solution for the set of equations.

3.2.1 Observation model

As in the discrete filtered back projection the activity is to be reconstructed in N pixels. Again, we denote the vector of the activities in the pixels by $f = (f_1, \ldots, f_N)^T \in \mathbb{R}^{N\times 1}$. The vector that contains the values of the pixels in the planar images in the height of the slice to be reconstructed is denoted by $P = (p_1, \ldots, p_B)^T \in \mathbb{R}^{B\times 1}$, where $B = M \cdot L$. M is the number of acquisition angles and L is the number of pixels in one row of the planar image.
Assume that H_{ij} is the contribution of j^{th} pixel to i^{th} projection bin. In the other words H_{ij} is the probability that a γ-photon that emitted from area of j^{th} pixel is detected in i^{th} projection bin. Figure

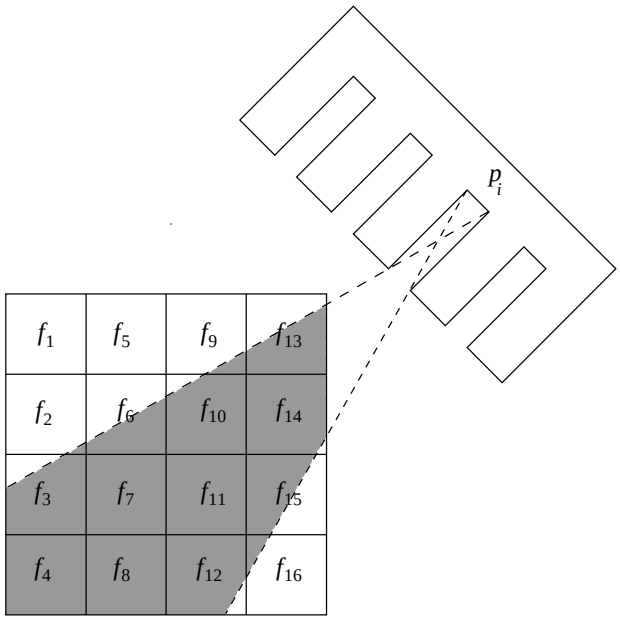


Figure 3.4: Formation of projection bin p_i: p_i is linear combination of activities f_j, that is $p_i = \sum_{j=1}^{N} H_{ij} f_j$. Projection bin p_i is contributed by pixels having parts in the colored area. Thus $H_{ij} \neq 0$, when $j \neq 1, 5$ and $H_{i1} = H_{i5} = 0$.

3.4 shows how the projection is formed in a simple case, when $N = 16$. If we use the additive model for measurement error the projection bin p_i gets the form

$$p_i = \sum_{j=1}^{N} H_{ij} f_j + v_i, \tag{3.34}$$

where v_i is the additive noise. From (3.34) we obtain the matrix presentation

$$\begin{pmatrix} p_1 \\ \vdots \\ p_B \end{pmatrix} = \begin{pmatrix} H_{11} & \cdots & H_{1N} \\ \vdots & \ddots & \vdots \\ H_{B1} & \cdots & H_{BN} \end{pmatrix} \begin{pmatrix} f_1 \\ \vdots \\ f_N \end{pmatrix} + \begin{pmatrix} v_1 \\ \vdots \\ v_B \end{pmatrix} \tag{3.35}$$

or

$$P = Hf + v, \tag{3.36}$$

where $H \in \mathbb{R}^{B \times N}$ is the observation matrix. If the matrix H is modeled correctly it includes all the physical features of the detecting system. The modeling of the observation matrix is the subject of Chapter 4. The observation matrix $\tilde{H}$ that was presented in previous section is constructed using the line integral model. This is one special case of observation matrix corresponding to the ideal collimator. Equation (3.36) presents the linear observation model of SPECT. Forming projection bins as a linear combination of activities is clearly a discrete version of integrating activities over some volume as presented in Chapter 1. When using equation (3.36) as an observation model, it is assumed that the number of emissions in each pixel is a deterministic variable.

3.2.2 The linear LS problem

If we approximate that v in equation (3.36) is a Gaussian noise vector, we can solve the activity distribution by using the linear LS-estimation. That is, the aim is to solve the estimate $\hat{f}_{LS}$ satisfying the condition

$$\hat{f}_{\mathrm{LS}} = \arg\min ||P - Hf|| \tag{3.37}$$

where P is the measured projection vector.
Assume that we have the measured data P and we have formed the matrix H. Now the task is to solve the LS problem defined by (3.37). Usually the observation model (3.36) is underdeterministic, that is, $N > B$. The LS estimate can not then be computed using equation (2.26). In that case the LS problem has not unique solution and some of the solutions must be chosen. Usually the minimum norm solution (2.27) is used. Furthermore, the inverse problem of SPECT is usually ill-posed and some regularization may be needed in solving the LS problem. In addition, the matrix H is very large and the LS solution must be computed iteratively.
Any of the iterative methods presented in Section 2.7 can be used to solve the LS problem. For example the algorithms ART (2.36) and SIRT (2.38) are often used for solving the LS problem of SPECT [18]. The approximate minimum norm solution is computed in [42] and [43] by using truncated SVD. The conjugate gradient least squares (CGLS) algorithm introduced in Section 2.7 is an attractive method for solving the ill-posed SPECT problem because of the regularization effect characteristic to conjugate gradient iterations. In this thesis we will use the CGLS algorithm for solving the LS problems. The conjugate gradient method can also be modified so that instead of applying the CG algorithm to normal equations (2.25), the observation equation (3.36) is multiplied by some matrix Z and the CGLS algorithm is applied to resulting preconditioned equations [25]. This method is called the preconditioned CGLS and it provides faster convergence, if the matrix Z is chosen well. In addition to these algorithms there are several other iterative algorithms converging to LS solution, see for example [15].

3.2.3 Regularization

Tikhonov regularization (see Section 2.6) can be used in the SPECT to get a stable solution. For the generalized Tikhonov regularization we have to determine the side constraint. In the SPECT one type of prior information is the non-negativity of the activity. Unfortunately the use of non-negativity leads to a non-linear LS problem. Thus we will deal with simpler side constraints to preserve the linearity of the LS problem.
The simplest form of the Tikhonov regularization is the standard Tikhonov regularization, in which the norm of the solution is forced to be small.
Another choice is the use of smoothness priors. For example the difference of values in pixels next to each other can be forced to be small. However the activity distribution is a two dimensional map of pixels and the parameters f_i equal to values of pixels are presented in (one dimensional) vector f. Thus, all the pixels being next to each other in activity map can not be next to each other in the parameter vector f. As a consequence, in the case of SPECT we can not use normal difference matrices in the side constraint. Instead, when forming the matrix L in side constraint we have to take into account which pixels are close to each other.
A simple way to form the matrix L is to use some two dimensional filter matrix Υ and perturbation technique: First we set value of j^{th} pixel equal to 1 and value of other pixels equal to 0. This pixel map is filtered with Υ and the impulse response is obtained. If the values of the impulse response are presented in the vector l, the vector l includes the contribution of j^{th} pixel to other pixels in filtering process. Thus the j^{th} column of L is set equal to vector l. This is done for every j to form the matrix L.
For example, if we make a choice that Υ is Laplacian filter, the side constraint Lf in (2.29) results in estimate $\hat{f}_\alpha$ which is smooth when presented in the form of a pixel map.

3.3 Comparison of the deterministic methods

In this section we will compare the deterministic methods in SPECT: the filtered back projection and the least squares methods. Table 3.1 presents the deterministic reconstruction methods for SPECT. The first four equations are the formal solutions of the SPECT problem. As noted before, the reconstructions are not actually computed in these forms. Instead, the LS solutions are obtained using iterative methods and FBP estimate is computed using FFT. Three of the iterative methods for solving the LS problem are also included in the table.
First we will consider the formal solutions. The intuitive explanation of the LS solution $\hat{f}_{\text{LS}}$ is as follows. The projection P is first back projected to image domain with H^T and then "deblurred" with matrix $(H^T H)^{-1}$. The matrix $H^T H$ is called the blurring matrix, because the j^{th} row of $H^T H$ gives the

Table 3.1: The deterministic reconstruction methods: The formal solutions and two iterative methods for solving the LS problem

	Method	Formula
Formal solutions	Least squares	$\hat{f}_{\mathrm{LS}} = (H^T H)^{-1} H^T P$
	Minimum norm least squares	$\hat{f}_{\mathrm{MN}} = H^T (H H^T)^{-1} P$
	Back projection	$\hat{f}_{\mathrm{BP}} = H^T P$
	Filtered back projection	$\hat{f}_{\mathrm{FBP}} = \tilde{H}^T \Omega P$
Iterative methods	Gauss-Newton	$\hat{f}^{(i+1)} = \hat{f}^{(i)} + k_i (H^T H)^{-1} H^T (P - H\hat{f}^{(i)})$
	ART	$\hat{f}^{(i+1)} = \hat{f}^{(i)} + \frac{1}{\|\|H_l\|\|^2} H_l^T (p_l - H_l \hat{f}^{(i)})$
	SIRT	$\hat{f}^{(i+1)} = \hat{f}^{(i)} + \left(\mathrm{diag}(H^T H)\right)^{-1} H^T (P - H\hat{f}^{(i)})$

probability density of from which pixel the detection system senses the photon being emitted when the photon is emitted from j^{th} pixel, [32, 33, 34].
Usually in the case of SPECT the observation model is underdeterministic. In that case the minimum norm solution is often used (note, that for example ART converges to $\hat{f}_{\mathrm{MN}}$). We see that in minimum norm solution the projection P is first modified multiplying with matrix $(HH^T)^{-1}$ and then backprojected.
If we consider the simple back projection $\hat{f}_{\mathrm{BP}}$, we see that it is an approximation of $\hat{f}_{\mathrm{LS}}$ and $\hat{f}_{\mathrm{MN}}$ if the blurring matrix $(H^T H) \approx I$, the identity matrix. Furthermore, the comparison of $\hat{f}_{\mathrm{BP}}$ to the equation (2.20) gives the following interpretation: $\hat{f}_{\mathrm{BP}}$ is the result of first iteration of the steepest descent algorithm with an initial guess $\hat{f}_0 = 0$.
In the filtered back projection the projection P is (formally) first multiplied with the matrix Ω presenting the filtering process, and then backprojected along straight lines. It is easy to see the analogy between the estimates $\hat{f}_{\mathrm{MN}}$ and $\hat{f}_{\mathrm{FBP}}$. If we choose $H = \tilde{H}$ in the minimum norm solution, then $\hat{f}_{\mathrm{MN}} = \tilde{H}^T (\tilde{H}\tilde{H}^T)^{-1} P$. If we want to get the FBP estimate equal to the minimum norm solution, the matrix Ω must be

$$\Omega = (\tilde{H}\tilde{H}^T)^{-1}. \tag{3.38}$$

As we noticed in 3.1.5 the matrix Ω in filtered back projection is a block matrix, whose all the diagonal blocks are equal and the blocks are obtained shifting the first row of first row of the block. Thus we can say that the filtered back projection solution is an approximation of the minimum norm solution with line integral model approximating the matrix $(\tilde{H}\tilde{H}^T)^{-1}$ with Ω described above.
In literature there is also papers about choosing the filter in FBP so that the collimator blurring is taken into account [8]. Now we try to find the matrix Ω so that $\tilde{H}^T \Omega P$ equals with $\hat{f}_{\mathrm{MN}}$, in the case that the collimator blurring is modeled in H. That is, we demand that

$$\tilde{H}^T \Omega P = H^T (H H^T)^{-1} P. \tag{3.39}$$

In that case

$$\tilde{H}^T \Omega = H^T (H H^T)^{-1}. \tag{3.40}$$

We solve this by writing

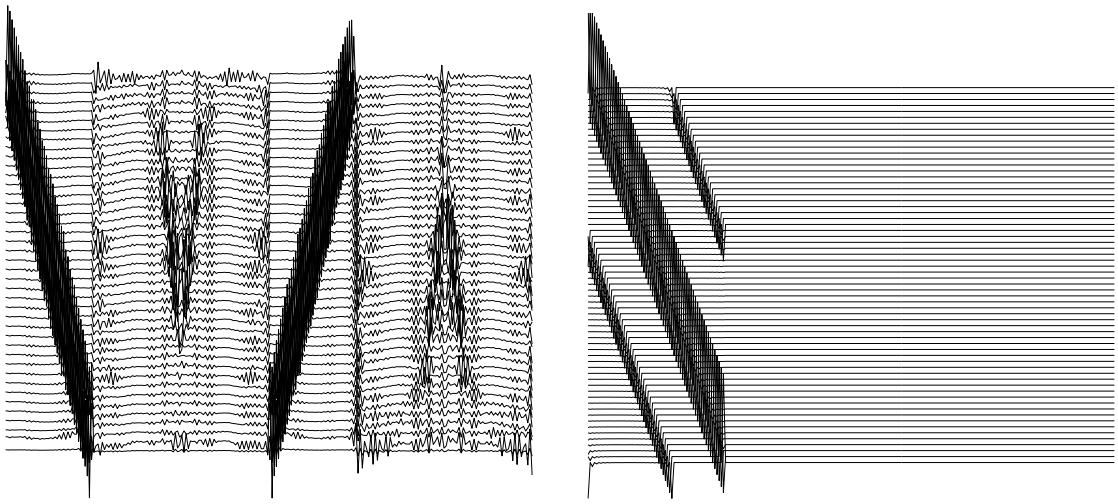
$$\tilde{H}\tilde{H}^T \Omega = \tilde{H} H^T (H H^T)^{-1}, \tag{3.41}$$

and multiplying with $(\tilde{H}\tilde{H}^T)^{-1}$:

$$\Omega = (\tilde{H}\tilde{H}^T)^{-1} \tilde{H} H^T (H H^T)^{-1}. \tag{3.42}$$

Thus we can say that the collimator corrected filtered back projection estimate $\hat{f}_{\mathrm{FBP}}$ is an approximation of the collimator corrected minimum norm estimate $\hat{f}_{\mathrm{MN}}$ approximating that $(\tilde{H}\tilde{H}^T)^{-1} \tilde{H} H^T (H H^T)^{-1}$ is a block matrix Ω as described above.

We formed the observation matrices $\tilde{H}$ and H in one specific case of very limited number of acquisition angles. The object to be imaged was divided to 41×41 pixels. We assumed that the number of acquisition angles was six: $\phi = 0^0, 60^0 \ldots, 300^0$ and each projection consisted of 41 projection bins. We computed both the matrix $(\tilde{H}\tilde{H}^T)^{-1}\tilde{H}H^T(HH^T)^{-1}$ related to the minimum norm LS solution and Ω related to the FBP solution. Figure 3.5 represents the 41 first rows of both of the matrices These rows correspond to the filtered projection at the angle $\phi_0 = 0^0$. The rows form six block matrices, the first of which is the diagonal block.



a) b)

Figure 3.5: a) The 41 uppermost rows of matrix $(\tilde{H}\tilde{H}^T)^{-1}\tilde{H}H^T(HH^T)^{-1}$ b) The 41 uppermost rows of the matrix Ω

As presented above, in the FBP method we approximate the matrix $(\tilde{H}\tilde{H}^T)^{-1}\tilde{H}H^T(HH^T)^{-1}$ by a block matrix Ω where all the off-diagonal blocks are all-zero matrices. In the case of Figure 3.5 the matrix $(\tilde{H}\tilde{H}^T)^{-1}\tilde{H}H^T(HH^T)^{-1}$ is certainly not of that form. The off-diagonal blocks are not all-zero blocks. There seems to be a high correspondence between the projections to opposite directions, especially with the opposite projection bins. (see Figure 3.5 a: the first block and the fourth block correspond to projections at angles 0^0 and 180^0). This is reasonable, because the projections at the opposite angles are approximately formed integrating the activities along the same lines. However, in the filtered back projection this correspondence is not taken into account as can be seen from the structure of the matrix Ω in Figure 3.5. As a consequence we can assume that in the case of limited number of observation angles the FBP method approximates the LS solution very roughly.
As we have presented previously there are several methods for solving the linear LS problem iteratively. Table 3.1 includes two of them: ART and SIRT. Clearly, the idea of these methods is to add the back-projection of the residual $P - H\hat{f}^{(i)}$ to current estimate until the condition $P = H\hat{f}^{(i)}$ is satisfied with certain accuracy. Table 3.1 shows that both of these methods can be considered as approximations of the Gauss-Newton method.

3.4 Statistical approach

In this section the statistical nature of the radioactive decay is taken into account. Consequently, the SPECT problem is treated as a statistical estimation problem.

3.4.1 Maximum likelihood estimation in SPECT

In the previous sections we assumed that the number of emissions in j^{th} pixel is a deterministic variable being equal to activity f_j. But as we noticed in 1.3.3 the number of emissions in pixel j is a Poisson distributed random variable. Taking this into account the equation (3.34) presenting the formation of

the i^{th} projection bin must be rewritten as

$$p_i = \sum_{j=1}^{N} H_{ij} A_{ij} + v_i, \tag{3.43}$$

where $A_{ij} \sim \text{Poisson}(f_j)$. If we neglect the measurement error v_i, the projection bin p_i is a linear combination of the Poisson distributed variables A_{ij}. Thus we get from superposition properties of Poisson distributed variables that $p_i \sim \text{Poisson}(\sum_{j=1}^{N} H_{ij} f_j)$. That is, the posterior density of p_i is

$$p(p_i|f) = \mathrm{e}^{-\sum_{j=1}^{N} H_{ij} f_j} \cdot \frac{(\sum_{j=1}^{N} H_{ij} f_j)^{p_i}}{p_i!}. \tag{3.44}$$

Moreover, the posterior density of P is

$$p(P|f) = \prod_{i=1}^{B} \mathrm{e}^{-\sum_{j=1}^{N} H_{ij} f_j} \cdot \frac{(\sum_{j=1}^{N} H_{ij} f_j)^{p_i}}{p_i!}. \tag{3.45}$$

Now we aim at forming the maximum likelihood estimate $\hat{f}_{\text{ML}}$ for the activity distribution, when the data P is known. According to the condition (2.8) the ML estimate is the value of f that maximizes the posterior density $p(P|f)$. The so-called expectation maximization (EM) algorithm is an iterative method for solving the maximum likelihood estimate.

3.4.2 EM algorithm for SPECT

The EM algorithm for the emission tomography was presented independently in [41] and [27]. Here we will follow the presentation of [27].
Because maximizing the posterior density (3.45) is difficult, we present the so-called "complete data" $z_{ij} = H_{ij} A_{ij}$. That is, z_{ij} is the number of photons which are emitted from pixel j and detected at projection bin i. The "complete data" z_{ij} is Poisson distributed: $z_{ij} \sim \text{Poisson}(H_{ij} f_j)$. If all variables z_{ij} are combined to one vector z, the posterior density of z is clearly

$$p(z|f) = \prod_{i=1}^{B} \prod_{j=1}^{N} \mathrm{e}^{-H_{ij} f_j} \cdot \frac{(H_{ij} f_j)^{z_{ij}}}{z_{ij}!}. \tag{3.46}$$

The logarithm of $p(z|f)$ is then

$$\ln p(z|f) = \sum_{i=1}^{B} \sum_{j=1}^{N} \left(-H_{ij} f_j + z_{ij} \ln(H_{ij} f_j) - \ln(z_{ij}!)\right). \tag{3.47}$$

Suppose that we have an parameter estimate $\hat{f}^{(n)}$. The conditional expectation of $\ln p(z|f)$ with respect to the data P and the current estimate $\hat{f}^{(n)}$ is

$$E\left\{\ln p(z|f) \,\middle|\, P, \hat{f}^{(n)}\right\} = \sum_{i=1}^{B} \sum_{j=1}^{N} \left(-H_{ij} f_j + N_{ij} \ln(H_{ij} f_j)\right) + R \tag{3.48}$$

where R does not depend on the new f, and N_{ij} is the conditional expectation of z_{ij}, that is

$$N_{ij} = E\left\{z_{ij} \,\middle|\, p_i, \hat{f}^{(n)}\right\} \tag{3.49}$$

The value of N_{ij} can be deduced as follows: The number of counts in the i^{th} projection bin is p_i. The probability that a count registered in the i^{th} projection bin has been emitted from the j^{th} pixel is $\frac{H_{ij} \hat{f}_j^{(n)}}{\sum_{k=1}^{N} H_{ik} \hat{f}_k^{(n)}}$. Thus, the conditional expectation of z_{ij} is

$$N_{ij} = \frac{H_{ij} \hat{f}_j^{(n)} p_i}{\sum_{k=1}^{N} H_{ik} \hat{f}_k^{(n)}}. \tag{3.50}$$

Now the task is to find the value of f that maximizes the conditional expectation of log likelihood $E\left\{\ln p(z|f)\,\middle|P,\hat{f}^{(n)}\right\}$. In the maximum point the partial derivatives of $E\left\{\ln p(z|f)\,\middle|P,\hat{f}^{(n)}\right\}$ must vanish:

$$\frac{\partial}{\partial f_j}\left(E\left\{\ln p(z|f)\,\middle|P,\hat{f}^{(n)}\right\}\right)=0 \tag{3.51}$$

$$\frac{\partial}{\partial f_j}\left(\sum_{i=1}^{B}\sum_{j=1}^{N}(-H_{ij}f_j+N_{ij}\ln(H_{ij}f_j))+R\right)=0 \tag{3.52}$$

$$-\sum_{i=1}^{B}H_{ij}+\sum_{i=1}^{B}N_{ij}\frac{1}{f_j}=0. \tag{3.53}$$

Combining equations (3.53) and (3.50) we obtain the following recursion

$$\hat{f}_j^{(n+1)}=\frac{\sum_{i=1}^{B}N_{ij}}{\sum_{i=1}^{B}H_{ij}}=\frac{\hat{f}_j^{(n)}}{\sum_{i=1}^{B}H_{ij}}\sum_{i=1}^{B}\frac{H_{ij}p_i}{\sum_{k=1}^{N}H_{ik}\hat{f}_k^{(n)}} \tag{3.54}$$

which is the EM algorithm. The intuitive interpretation of EM algorithm is as follows [27]: The sum $\sum_{i=1}^{B}N_{ij}$ is the estimated number of detected photons, which have been emitted from j^{th} pixel. In the recursion this is set equal to the expected number $\sum_{i=1}^{B}H_{ij}f_j$.
One remarkable feature of EM algorithm is the non-negativity of the estimate: if the initial guess $\hat{f}^{(0)}$ is positive, the estimate is positive after every iteration. Thus, the EM algorithm regularizes the problem with implicit non-negativity constraint.

3.4.3 Maximum a posteriori estimation in SPECT

Maximum a posteriori (MAP) estimation can be used in SPECT if some prior information of the activity distribution is known. Maximum a posteriori estimate satisfies the condition (2.7). According to the equation (2.6) the function to be maximized in MAP estimation is $p(P|f)p(f)$, where $p(P|f)$ is the posterior density of P as in ML estimation and $p(f)$ is the prior density of f containing the prior information of f. There are several publications in literature of MAP estimation. In those papers the EM algorithm (3.54) is modified to obtain the MAP estimate.
In [9, 26, 13, 14] the choice of prior density is the smoothness prior. Then $p(f)$ is a functional

$$p(f)=\mathrm{e}^{-\beta V(f)}, \tag{3.55}$$

where $V(f)$ is the so-called energy function, which is at minimum when f is smooth.
In [7, 29, 3, 31] anatomical information is used in prior density. The anatomical information is obtained from CT or MRI images. In that case the prior density $p(f)$ is chosen so that the activity distribution is forced to be smooth inside the anatomical regions but gaps are accepted in the boundaries of these regions.
Clearly the prior density $p(f)$ plays the same role in MAP estimation as the side constraint in Tikhonov regularized LS estimation. The use of prior density regularizes the reconstruction problem.

3.5 Three dimensional reconstruction

In 2D reconstruction it is assumed that all photon counts in a particular row of the planar image are due to photons emitted from the slice at same height. Actually every projection bin is also contributed by voxels located in other slices. Figure 3.6 shows, how one projection bin is contributed by voxels inside a cone. Dimensions of the cone depend on the dimensions and shape of holes of the collimator.
In order to compute the fully 3D reconstruction we must discretize the whole 3D region where the activity distribution is to be reconstructed. The 3D region is discretized by $N_x\times N_y\times N_z$ voxels. Thus the activity distribution is unknown vector $f=(f_1,\ldots,f_{N_xN_yN_z})^T\in\mathbb{R}^{N_xN_yN_z\times 1}$. We denote the i^{th} planar image at angle ϕ_i by $P_i\in\mathbb{R}^{P_yP_z\times 1}$ Figure 3.7 illustrates the dimensions of the vectors f and P_i.

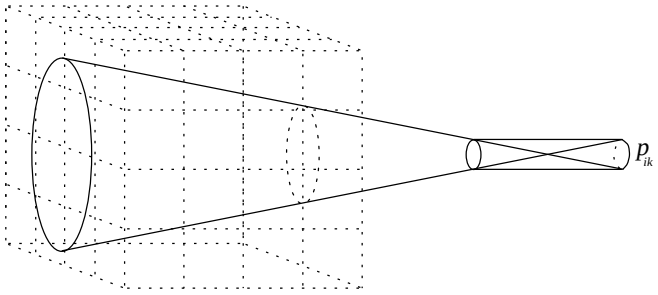


Figure 3.6: Representation of 3D collimator effect: projection bin p_{ik} is contributed by voxels inside the cone

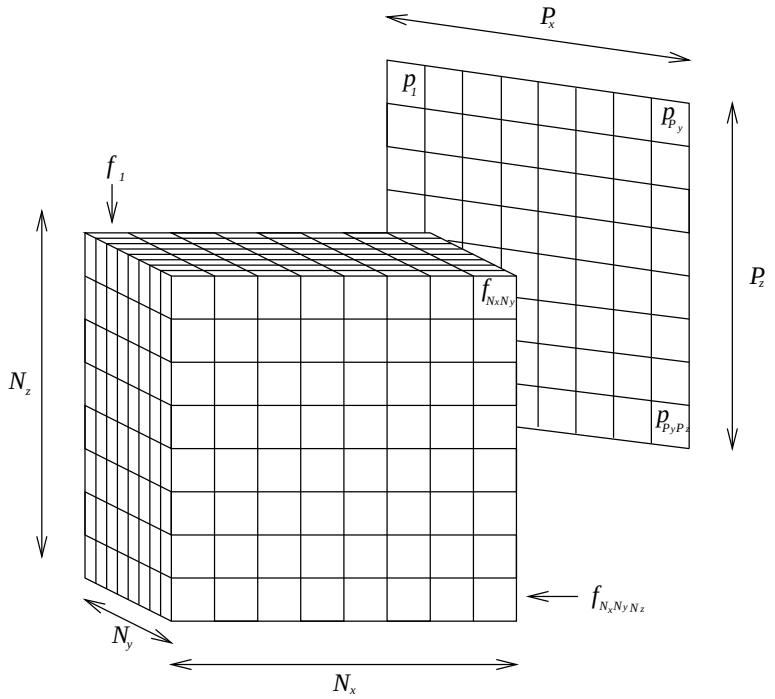


Figure 3.7: The 3D region is discretized in $N_x \times N_y \times N_z$ voxels. The planar image is a discrete map with $P_y \times P_z$ pixels.

First we present the LS estimation in the case of the fully 3D model. In the deterministic case we can now write a model for the first planar image P_1:

$$P_1 = H_1 f + V_1, \tag{3.56}$$

where the dimensions of matrix H_1 are $P_y P_z \times N_x N_y N_z$ and the measurement error vector $V_1 \in \mathbb{R}^{P_y P_z \times 1}$. The element $H_1(k, j)$ presents probability of a γ-photon emitted from voxel j to be detected at projection bin p_k.
Equation (3.56) defines the first planar image. The equation that defines many planar images can be written in the matrix form

$$\begin{pmatrix} P_1 \\ \vdots \\ P_M \end{pmatrix} = \begin{pmatrix} H_1 \\ \vdots \\ H_M \end{pmatrix} f + \begin{pmatrix} V_1 \\ \vdots \\ V_M \end{pmatrix} \tag{3.57}$$

or

$$P = Hf + v, \tag{3.58}$$

where $P_1, \ldots, P_M$ are the planar images at acquisition angles $\phi_1 \ldots \phi_M$.
Matrix H is a sparse matrix, whose dimensions are $MP_yP_z \times N_xN_yN_z$. Construction of the matrix H will be presented later in Section 4.4.
The equation (3.58) presents the 3D deterministic observation model of SPECT. Now we can compute the LS solution $\hat{f}_{\mathrm{LS}}$ for the set of equations (3.58) as in the case of the 2D model. The difference to 2D case is that now the vector $\hat{f}_{\mathrm{LS}}$ is an estimate for three dimensional activity distribution, while in the 2D case $\hat{f}_{\mathrm{LS}}$ is the estimate for the activity distribution in one slice of the three dimensional region.
Now the computation of the ML estimate in the 3D case should be trivial. If we use the notations P, f and H presented in this section, the ML estimate can be computed using EM algorithm (3.54) as in the 2D case.

Chapter IV

Modeling of the forward problem

The algebraic reconstruction techniques and the iterative methods are methods for solving the inverse problem of SPECT. To solve the inverse problem we have to form the model for the observations first. In Section 3.2 we presented the linear observation model of SPECT

$$P = Hf + v, \tag{4.1}$$

In Section 3.4 we took into account the Poisson statistics of the radioactive decay and obtained the posterior density of the data P

$$p(P|f) = \prod_{i=1}^{B} \mathrm{e}^{-\sum_{j=1}^{N} H_{ij} f_j} \cdot \frac{(\sum_{j=1}^{N} H_{ij} f_j)^{p_i}}{p_i!}. \tag{4.2}$$

In the deterministic model case the forward problem is to compute P when f is known. In the statistical sense the forward problem is to compute the posterior density $p(P|f)$ when f is known. In both cases solving the forward problem requires determining the matrix H. Matrix H is the observation matrix whose element H_{ij} gives probability for a photon emitted from pixel j to be detected in i^{th} projection bin. When H is constructed, that is, the forward problem is modeled, the inverse problem or the reconstruction problem can be solved using methods of Chapter 3.
In literature there is lots of papers considering the reconstruction algorithms, but the construction of the observation matrix H is not often described in detail. The basic model, the line integral model is explained in [18]. In [16] and [44] different kinds of models for collimator response are proposed. The modeling of attenuation is presented in [30] and [44]. In this chapter we present in detail how the observation matrix H can be constructed taking into account both the collimator blurring and attenuation.

4.1 Line integral model

The most simple model for SPECT is called called here the line integral model. When using this model we approximate the projection bins p_i as line integrals or ray-sums of the activities in pixels. More precisely, we consider a fat line passing through the i^{th} projection bin and having direction parallel to the holes of the collimator. The width of the fat line is equal to width of the projection bin p_i. The element H_{ij} is now the area of the fat line belonging to j^{th} pixel, see Figure 4.1. Approximation that the projection bins are formed integrating the activities over straight lines equals to approximation of extremely narrow holes in the collimator. Thus we also refer the line integral model as the ideal collimator model.

4.2 Modeling of the collimator

We already noted in Section 1.3 that the collimator is never ideal. A projection bin is actually sum of activities inside a cone instead of a fat line. In two dimensional modeling we take into account the activity in the slice at the height of the projection bin. Thus, in two dimensional case the i^{th} projection bin is assumed to be formed as in Figure 3.4. The modeling presented in this section is suitable for *parallel*

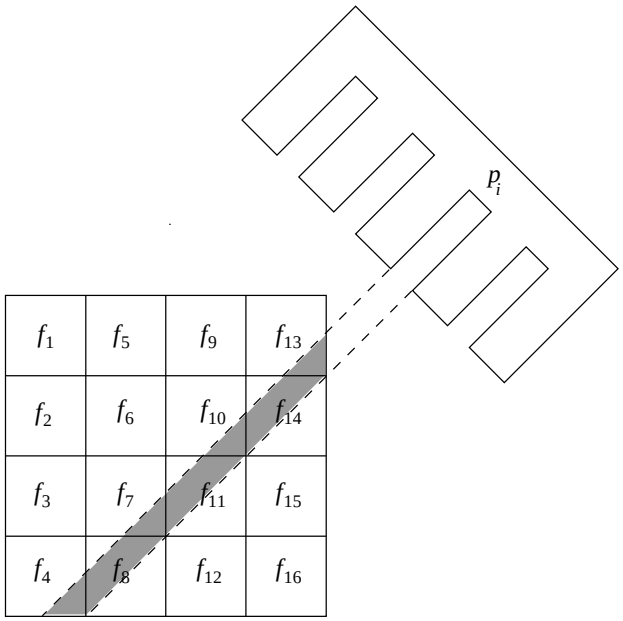


Figure 4.1: The line integral model in SPECT: The projection bin p_i is assumed to be ray-sum of the activities. The contribution of activity f_j to projection bin p_i is approximately the area of the j^{th} pixel inside the colored area.

beam collimators in which the holes are parallel. In Section 4.4 when forming the 3D collimator model we will use formalism which is suitable for other types of collimators too.
The computation of all the elements of the observation matrix H separately would be a time requiring task. To quicken the computation we use following procedure in construction of the matrix H: First we compute a zero-angle observation matrix H_0. The zero-angle observation matrix means the matrix that models the the projection at angle $\phi_0 = 0^0$. Next we form the matrix H block by block computing the matrices for other observation angles ϕ_k using the zero-angle observation matrix and bilinear interpolation.

4.2.1 Zero-angle observation matrix

Now we present the formation of the collimator modeled zero-angle observation matrix H_0, which we will use in forming the observation matrix H. As we defined above, the zero-angle observation matrix H_0 is the matrix modeling the formation of the projection at angle $\phi_0 = 0^0$. Thus, for computation of H_0 we set the coordinates as in Figure 4.2. The origin situates in the center of rotation of the Anger camera. The distance of the camera from the center of rotation is denoted by l. Knowing the dimensions of the object to be imaged and the Anger camera we can calculate the coordinates (x, y) for every pixel of the image and the y-coordinate Y for every projection bin. Here we have droped the subindexes of the coordinates for simplicity. Actually the coordinates are (x_j, y_j) and (l, Y_i).
Now, assume that the coordinates of the center point of the j^{th} pixel are (x, y). The matrix element $H_0(i, j)$ is the probability that a photon emitted from pixel j is detected by i^{th} projection bin. Thus, $H_0(i, j)$ is equal to the ratio

$$H_0(i,j) = \frac{\varphi_{ij}}{4\pi} \tag{4.3}$$

where φ_{ij} is the solid angle subtended by visible part of projection bin p_i to point (x, y) and 4π is the solid angle subtended by a sphere surrounding the point (x, y). We compute the solid angle φ_{ij} approximately using assumptions that holes of the collimator are square and parallel.
If we write (see Figure 4.3)

$$dx = l - x \ , \quad dy = |Y - y|, \tag{4.4}$$

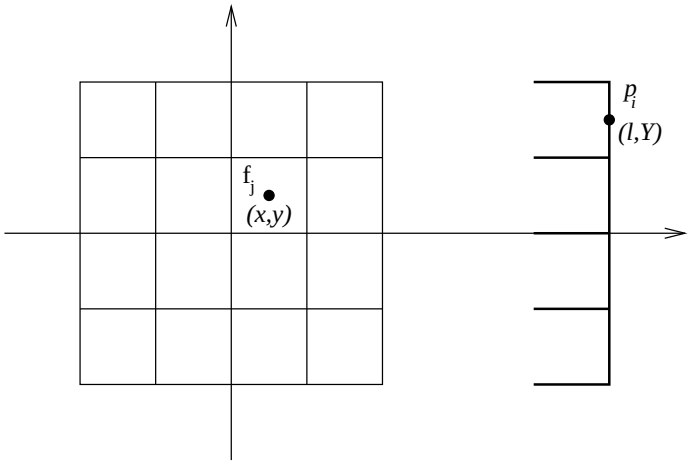


Figure 4.2: Coordinate system in construction of the zero-angle observation matrix. The origin is located in the center of rotation of the Anger camera. The coordinates of center of the j^{th} pixel are (x, y) and coordinates of center of the i^{th} projection bin are (l, Y).

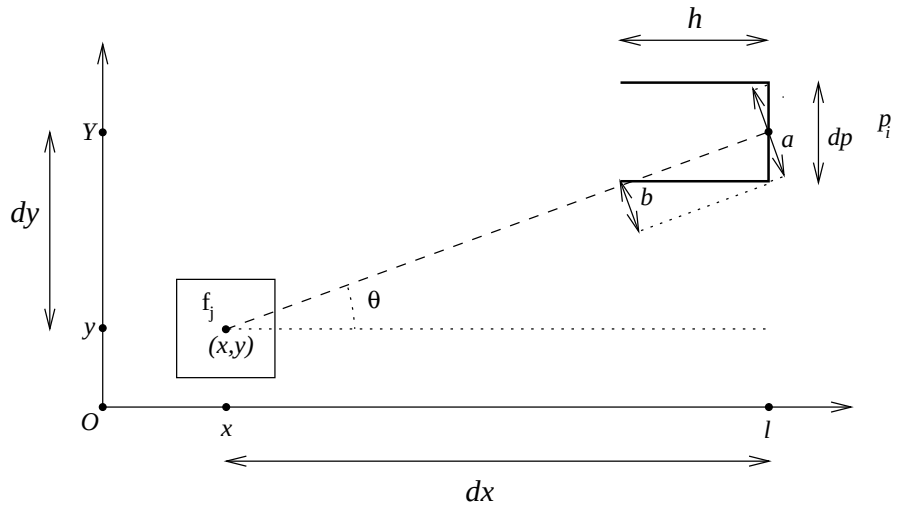


Figure 4.3: Computation of the element $H_0(i, j)$. Part of the projection bin p_i is seen by the j^{th} pixel. The contribution of the j^{th} pixel to i^{th} projection bin is proportional to solid angle subtended by the visible part of projection bin to (x, y).

the distance of j^{th} pixel from i^{th} projection bin is

$$d = \sqrt{dx^2 + dy^2}. \tag{4.5}$$

Then, we can write for the angle θ in Figure 4.3

$$\cos\theta = \frac{dx}{d}\ , \quad \sin\theta = \frac{dy}{d} \tag{4.6}$$

We denote the width of projection bin p_i by dp and the length of the collimator by h. The projection of dp at direction of pixel j is clearly

$$a = dp\cos\theta = dp\frac{dx}{d}, \tag{4.7}$$

and projection of h at direction of pixel j is

$$b = h \sin\theta = h\frac{dy}{d}. \tag{4.8}$$

If the height of the slice is dp, that is, the projection bin is a square, the solid angle subtended by the visible part of the projection bin p_i to point (x, y) is approximately

$$\varphi_{ij} = \frac{(a - \kappa b)dp}{d^2}, \tag{4.9}$$

where the factor κ is number of collimator holes in area of projection bin. In Figure 4.3 the width of the hole is equal to width of the projection bin, when $\kappa = 1$. However, usually $\kappa > 1$ for real collimators. Obviously, in the case that $\varphi_{ij} < 0$ the pixel j does not contribute to projection bin p_i, and we have to set $\varphi_{ij} = 0$.
Now we get the element $H_0(i, j)$ using the equation (4.3). Repeating the same procedure for all $j = 1, \ldots, N$ and $i = 1, \ldots, L$ we obtain the zero-angle matrix H_0.

4.2.2 Other angles

Assume that we have computed the zero-angle observation matrix H_0 as presented above. Now, we want to compute the other blocks H_k of the matrix H corresponding to the projections to angles ϕ_k.
The zero-angle observation matrix is actually an observation matrix fixed to the coordinates of the collimator. The rotation of the Anger camera around the object to be imaged equals to rotation of the object in the coordinates fixed to the collimator. If the collimator is rotated counter clockwise at angle ϕ_k, it equals to rotating the object clockwise at angle ϕ_k in the coordinates fixed to the collimator. Thus, we can construct the matrix H as follows.
Let f be a vector presenting the activity map. To compute the projection of activity distribution at angle ϕ_k we rotate the activity map to angle ϕ_k in the coordinates fixed to the collimator. Let $R_k \in \mathbb{R}^{N\times N}$ be the *interpolation matrix*, so that the rotated activity map is $\tilde{f}^k = R_k f$, see figure 4.4.

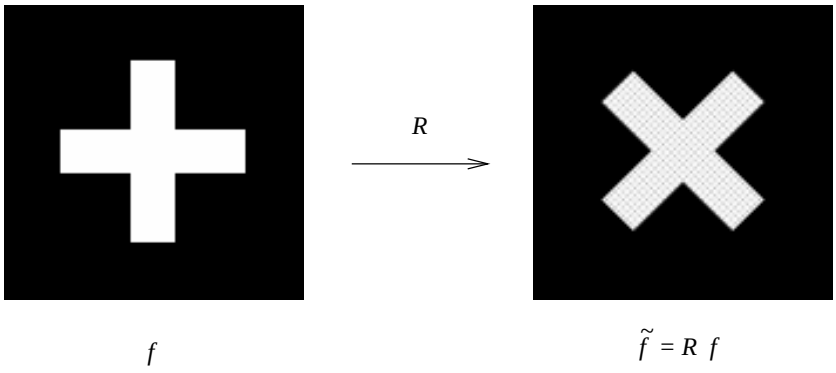


Figure 4.4: Rotation of an image. R is the interpolation matrix. Multiplying the image vector f with R gives the rotated image $\tilde{f}$.

Now, when we have interpolated the rotated image in the chosen coordinates, we can use the zero-angle observation matrix H_0 to form the projection at angle ϕ_k:

$$P_k = H_0 R_k f \tag{4.10}$$

from which we see that the k^{th} block of the matrix H is $H_k = H_0 R_k$. Now we have constructed the observation matrix

$$H = \begin{pmatrix} H_0 R_1 \\ H_0 R_2 \\ \vdots \\ H_0 R_M \end{pmatrix}, \tag{4.11}$$

where M is the number of acquisition angles. Usually we make a choice that the first acquisition angle ϕ_1 is the zero-angle, $\phi_1 = 0^0$. In that case $R_1 = I$, the identity matrix.

4.3 Attenuation corrected model

We discuss next the modeling of the attenuation in the forward problem. As presented in 1.3.2 the intensity of the photon flux reduces exponentially

$$I(l) = I_0 \mathrm{e}^{-\int_0^l \mu(l)\mathrm{d}l}, \tag{4.12}$$

Next we construct the observation matrix taking the attenuation into account. Let $\mu \in \mathbb{R}^{N\times 1}$ be the vector that includes the attenuation coefficients of the pixels.
First assume that we have already computed the collimator modeled observation matrix H. Using the law of attenuation (4.12) we get the attenuation corrected observation matrix H^{Att}

$$H^{Att} = H \odot \Gamma, \tag{4.13}$$

where $\odot$ stands for the element-by-element multiplication operator. The matrix Γ is the *attenuation matrix*, whose element Γ_{ij} is

$$\Gamma_{ij} = \mathrm{e}^{-\sum_l d_{lij}\mu_l}, \tag{4.14}$$

where μ_l is the attenuation coefficient of l^{th} pixel and d_{lij} is the distance that a segment of line connecting the j^{th} pixel and the i^{th} projection bin passes through the l^{th} pixel.
But the computation of the distances d_{lij} for every l, i, j would be computationally heavy task. Thus we use same kind of method as in forming the collimator modeled matrix H. Again, we consider the situation of Figure 4.2, where the acquisition angle is $\phi_0 = 0^0$. We form the collimator modeled zero-angle observation matrix H_0 again. In addition, we compute the distances d_{lij} for all combinations of l, i, j in the case that the acquisition angle is 0^0. Now we can compute the block Γ_k of the attenuation matrix Γ corresponding to the angle ϕ_k as follows: First we rotate the attenuation map to angle ϕ_k in the coordinates fixed to the collimator:

$$\tilde{\mu}^k = R_k \mu, \tag{4.15}$$

where R_k is the same interpolation matrix as in the Section 4.2 and $\tilde{\mu}^k$ is the rotated attenuation vector. Now that we have rotated the attenuation map in the coordinates fixed to the collimator, we can use the distances d_{lij} in constructing the matrix block Γ_k so that

$$\Gamma_k(i,j) = \mathrm{e}^{-\sum_l d_{lij}\tilde{\mu}_l^k}. \tag{4.16}$$

In addition to rotating the attenuation map we have to rotate the activity map f to form projection vector P_k at angle ϕ_k (as in Section 4.2). Thus, the vector P_k is

$$P_k = (H_0 \odot \Gamma_k) R_k f, \tag{4.17}$$

from which we see that the attenuation modeled observation matrix H^{Att} is

$$H^{Att} = \begin{pmatrix} (H_0 \odot \Gamma_1) R_1 \\ (H_0 \odot \Gamma_2) R_2 \\ \vdots \\ (H_0 \odot \Gamma_M) R_M \end{pmatrix}. \tag{4.18}$$

4.4 Three dimensional modeling

In Section 3.5 we presented the 3D observation model of SPECT. As in the two dimensional case to solve the forward problem we have to construct the observation matrix H

$$H = \begin{pmatrix} H_1 \\ \vdots \\ H_M \end{pmatrix} \in \mathbb{R}^{MP_yP_z \times N_xN_yN_z}, \tag{4.19}$$

where the block matrices H_k are the matrices corresponding to the planar images P_k at angles ϕ_k. In this section we will present the construction of the collimator modeled 3D observation matrix. The procedure of forming the matrix H in 3D case is basicly the same as in 2D collimator modeled case (Section 4.2). First we form the zero-angle observation matrix H_0 corresponding to the planar image at acquisition angle $\phi_0 = 0^0$. Then we can form the matrix H block by block using bilinear interpolation.
Similarly as in 2D collimator modeled case we get that the contribution of the j^{th} voxel to i^{th} projection bin is

$$H_0(i,j) = \frac{\varphi_{ij}}{4\pi}, \tag{4.20}$$

where φ_{ij} is a solid angle subtended by the visible part of p_i to j^{th} voxel.
The calculation of the solid angles φ_{ij} depends on the shape of the holes in the collimator. Next we assume, that the holes are circular and there is one hole per each projection bin. In this case we present the computation of the solid angles φ_{ij} so that the holes do not have to be perpendicular to the crystal plain. That is, the collimator can also be e.g a *conebeam collimator* [39], or a *multifocal collimator* [11, 49] not necessarily the parallel beam collimator. We denote the radius of the circular holes by r and length of the collimator by h.
Let $\bar{d} \in \mathbb{R}^3$ be a vector pointing from center of p_i to voxel j as seen in Figure 4.5 a). Collimator hole vector, a vector pointing from center of p_i to center of the other end of the hole is denoted by $\bar{n}$. Angle between the vectors $\bar{d}$ and $\bar{n}$ can be calculated from vector product of $\bar{n}$ and $\bar{d}$:

$$\beta = \arccos\left(\frac{\bar{n}\cdot\bar{d}}{|\bar{n}||\bar{d}|}\right), \tag{4.21}$$

where $|\bar{n}|$ and $|\bar{d}|$ are Euclidean norms of the vectors. Note that for normal parallel beam collimator $|\bar{n}| = h$ for all holes.
In order to calculate the solid angle φ_{ij} we have to calculate the area of the visible part of p_i for voxel j. Figure 4.5 b) shows how projection bin p_i is seen by voxel j. The visible part of p_i is determined by two intersecting ellipses. The ellipses are projections of both ends of the collimator hole to direction of voxel j. The ellipses are not exactly equal, but if $|\bar{d}|$ is big enough we can approximate them as equal.
Figure 4.6 represents the same situation as Figure 4.5 b) with an exception that the coordinates have been chosen so that the origin is located at the center of the ellipse corresponding to more distant end of the collimator hole, and y-axis passes the center of the ellipse corresponding to closer end of the collimator hole. Next we have to calculate area A bounded by the intersecting ellipses.

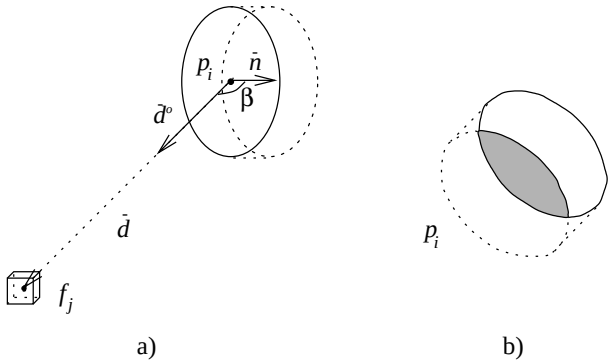


Figure 4.5: a) Voxel j and the part of the projection bin p_i that is behind a circular hole of the collimator. $\bar{d}$ is a vector pointing from p_i to j, and $\bar{d}^0$ is the corresponding unit vector. $\bar{n}$ is a vector pointing along the hole of collimator b) Projection bin p_i from the angle of view of the voxel j

First we formulate equations for both ellipses. Horizontal radius for both of the ellipses is r. Vertical radius of both of the ellipses is $b = r\cos\beta$, which is obtained by projecting radius of circle to the direction of the voxel j. Similarly by projecting the collimator hole vector $\bar{n}$ we obtain the distance between the

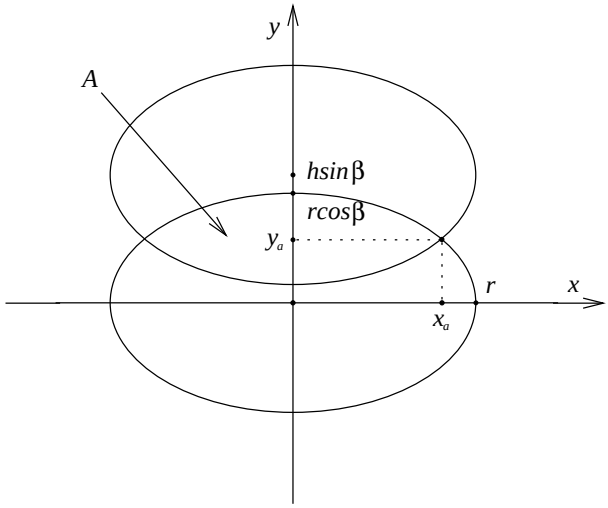


Figure 4.6: Calculation of the area A seen by voxel j. A is the area bounded by two intersecting ellipses.

centers of the ellipses $k = |\bar{n}| \sin\beta$ (For parallel beam collimator this is $k = h \sin\beta$). Thus the center of the upper ellipse is in $(0, k)$ and the equations of the ellipses are:

$$\frac{x^2}{r^2} + \frac{y^2}{b^2} = 1 \tag{4.22}$$

for the lower, and

$$\frac{x^2}{r^2} + \frac{(y-k)^2}{b^2} = 1 \tag{4.23}$$

for the upper ellipse.
Next we try to find the points of intersection of the ellipses. From symmetry we get that y-coordinate of both of the points of intersection is $y_a = k/2$. The x-coordinates are obtained by substituting y_a in (4.22):

$$x = \pm r\sqrt{1 - \frac{k^2}{4b^2}}. \tag{4.24}$$

Thus, the intersecting points are $(-x_a, y_a)$ and (x_a, y_a), where

$$x_a = r\sqrt{1 - \frac{k^2}{4b^2}}. \tag{4.25}$$

Now we can calculate A with integration. We get from symmetry:

$$\begin{aligned} A &= 2\int_{-x_a}^{x_a} (y(x) - y_a)\, dx \\ &= 4\int_{0}^{x_a} (y(x) - y_a)\, dx \end{aligned} \tag{4.26}$$

where $y(x)$ is a function defining the upper half of the lower ellipse. We get from (4.22):

$$y(x) = b\sqrt{1 - \frac{x^2}{r^2}}. \tag{4.27}$$

Then equation (4.26) becomes

$$A = 4\int_0^{x_a} \left(b\sqrt{1 - \frac{x^2}{r^2}} - \frac{k}{2} \right) dx$$

$$
\begin{aligned}
&= 4\Big/_{0}^{x_a}\left(\frac{b}{r}\left(\frac{x}{2}\sqrt{r^2-x^2}+\frac{r^2}{2}\arcsin\left(\frac{x}{r}\right)\right)-\frac{k}{2}x\right) \\
&= 4\left(\frac{b}{r}\left(\frac{x_a}{2}\sqrt{r^2-x_a^2}+\frac{r^2}{2}\arcsin\left(\frac{x_a}{r}\right)\right)-\frac{k}{2}x_a\right).
\end{aligned} \tag{4.28}
$$

Now φ_{ij}, the solid angle subtended by A to voxel j can be calculated by

$$\varphi_{ij} = \frac{A}{|\bar{d}|^2} \tag{4.29}$$

and the corresponding matrix element $H_0(i,j)$ is obtained from the equation (4.20).
When we have formed the zero-angle observation matrix we can construct the observation matrix H block-by-block as in the 2D case. The only exception is that now we can not rotate the activity map at angle ϕ_k simply multiplying the vector f with R_k. The vector f represents the three dimensional activity distribution, and the elements $f_1,\ldots,f_{N_xN_y}$ represent the uppermost slice, the elements $f_{N_xN_y+1},\ldots,f_{2N_xN_y}$ the second highest slice and so on. Thus, these blocks of the vector f must be multiplied with R_k separately presenting the rotation of all the slices at the angle ϕ_k. We can write the zero-angle matrix H_0 in the form

$$H_0 = \left(H_0^{(1)},\ldots,H_0^{(N_z)}\right), \tag{4.30}$$

where a block $H_0^{(l)}$ models the projection of l^{th} slice of activity distribution at angle $\phi_0 = 0^0$. Now the formation of the planar image at angle ϕ_k can be written as

$$P_k = \left(H_0^{(1)}R_k, H_0^{(2)}R_k,\ldots,H_0^{(N_z)}R_k\right) f, \tag{4.31}$$

and thus, the observation matrix H is

$$H = \begin{pmatrix} H_0^{(1)}R_1 & H_0^{(2)}R_1 & \ldots & H_0^{(N_z)}R_1 \\ H_0^{(1)}R_2 & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ H_0^{(1)}R_M & \ldots & \ldots & H_0^{(N_z)}R_M \end{pmatrix}. \tag{4.32}$$

Chapter V

Simulations and phantom studies

In this chapter we evaluate the models and reconstruction techniques with four simulation studies and one set of real measurement data. The first and second simulation are based very simple geometries; the purpose of these studies is to illustrate the effect of collimator blurring and attenuation in SPECT. The third simulation presents a more realistic case with simulated Poisson noise and matrices with larger dimensions. Various models and reconstruction techniques are evaluated through this simulation. In the fourth simulation we study the effect of 3D modeling. The real measurement study is based on cardiac phantom measurements.

5.1 Simulation study 1

In the first simulation we formed a two dimensional mathematical phantom with 32×32 pixels. The width of the phantom was 40 cm. We generated three point sources in the phantom by setting activity to unity in three pixels and zero in other pixels (see Figure 5.1).

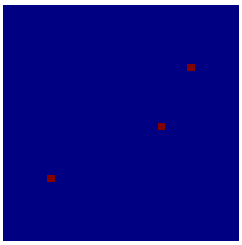


Figure 5.1: The phantom of the first simulation. Activity is unity in the three red points and zero in other parts of the phantom.

We assumed that the dimensions of the collimator were as follows: width 40 cm, thickness 5 cm, the number of holes 32 and the width of the square shaped holes 6.25 mm. The number of projection bins was 32 and the radius of rotation 46 cm.
To illustrate the collimator blurring effect we constructed the observation matrices using both the line integral model and the collimator model. In both cases we computed the projection vectors P_k at 16 angles: $\phi_k = 0^0, 22.5^0, \ldots, 337.5^0$ using the deterministic equation

$$P = Hf, \tag{5.1}$$

where the vector P includes all the vectors P_k, H is the line integral or collimator modeled observation matrix, and the vector f contains the values of the pixels in the phantom, as presented previously. Figure 5.2 presents the computed projections in the form of sinograms. The projections of the points are clearly wider in the case of the collimator modeled case. The spreading of a pixel into several pixels in planar image is due to the collimator blurring effect as explained in Section 1.3

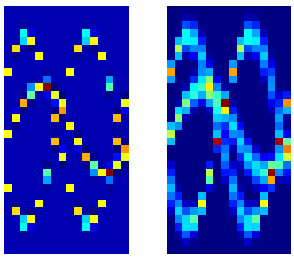


Figure 5.2: The sinograms simulated using line integral modeled observation matrix (left) collimator modeled observation matrix (right).

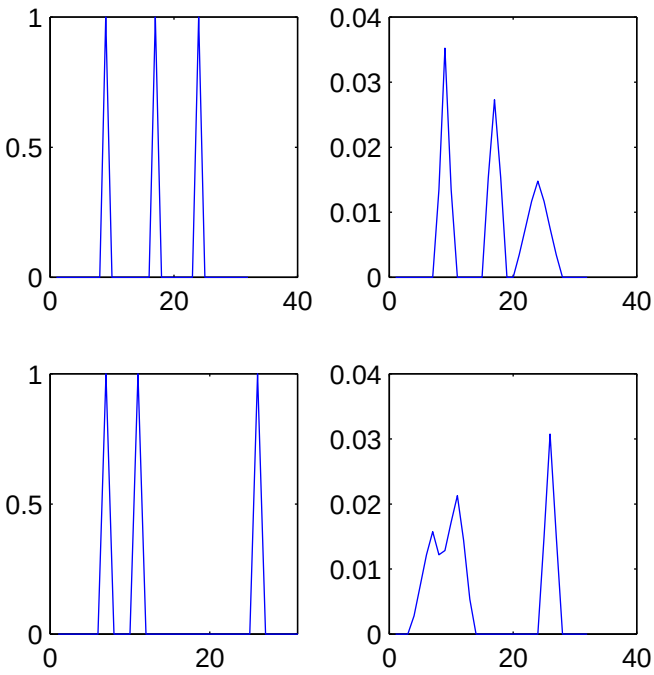


Figure 5.3: Top) The projection vectors at angle 0^0 with line integral model (left) and collimator model (right). Bottom) The projection vectors at angle 90^0 with line integral model (left) and collimator model (right).

Figure 5.3 presents the projections P_k at angles 0^0 and 90^0 computed using both models. Here we see the effect of collimator blurring better. In the case of the line integral modeled forward problem the projections of pixels do not depend on the distance of the activity from the collimator. The points at different distances cause peaks with same height in the projection. In the collimator modeled case the height of the peak depends on the distance: the further points cause peaks which are wider and lower.
Next we added Gaussian noise to projection vector P simulated using the collimator model. We used Gaussian noise with variance 10 % of mean of projection values p_i. We used these projections with Gaussian noise as data to compute the inverse problem of SPECT. We computed the LS estimate for activity distribution using the CGLS algorithm. Figure 5.4 shows the estimates obtained using both the line integral modeled and the collimator modeled observation matrix H. The estimate obtained using the collimator model seems to be smoother and more accurate. However, it is important to notice that this does not prove that the collimator modeled reconstruction would give better estimates in general. We

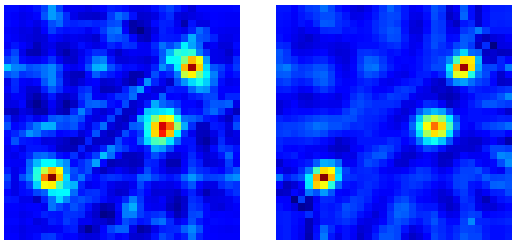


Figure 5.4: The LS estimates of the activity distribution in the first simulation. The estimates were computed using CGLS algorithm with the line integral model (left) and the collimator model (right).

used the same collimator model for computing both the forward problem and the inverse problem. If any noise was not added to observations, this would have been so-called *inverse crime*. However, we added a considerable amount of noise to projections, and thus, we have not strictly made an inverse crime here.

5.2 Simulation study 2

In the second simulation we illustrate the effect of attenuation in SPECT. In this simulation we assumed the detection system to have dimensions equal to the first simulation. Here we generated a phantom with two point sources, having activity 2 in another point and 3 in another. In this phantom we also added two attenuating objects (see Figure 5.5).

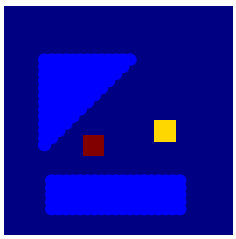


Figure 5.5: The phantom in the second simulation. Activity is 3 in the red point, 2 in the yellow point and zero elsewhere. The blue objects have a non-zero attenuation coefficient, elsewhere the attenuation coefficient is zero.

We constructed the observation matrix using the collimator model with attenuation correction in three different cases. We variated the attenuation map so that first we set the attenuation coefficient $\mu_1 = 0.9\text{cm}^{-1}$ for the quadrangle shaped object and $\mu_2 = 0.6\text{cm}^{-1}$ for the triangle shaped object, next we set $\mu_1 = 0.3\text{cm}^{-1}$ and $\mu_2 = 0.2\text{cm}^{-1}$, and finally $\mu_1 = 0.15\text{cm}^{-1}$ and $\mu_2 = 0.1\text{cm}^{-1}$. Figure 5.6 shows sinograms computed in each case. Again, the projections were computed using the deterministic model (5.1).
In each case we added Gaussian noise to projection vector P. The variance of the Gaussian noise was 3 % of mean of projection values p_i. Next we computed LS estimates for the activity distribution in the case of each data set using the collimator modeled observation matrix both with and without attenuation correction. When computing the attenuation corrected estimates we used the same observation matrices as in the forward problem. Again, we computed the LS estimates using CGLS algorithm. All the estimates are presented in Figure 5.7.
Clearly the estimates computed using the attenuation correction have better quality than those without attenuation correction. This is reasonable because the attenuation corrected model is the complete model in the case of this simulation. However, we can note that when the attenuation coefficients of

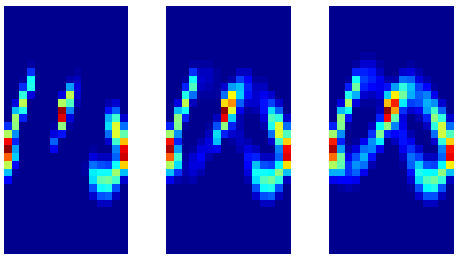


Figure 5.6: The sinograms computed using attenuation corrected observation model. The attenuation coefficient of the attenuating objects is highest in the case of the image in left side and lowest in the case of the image in right side.

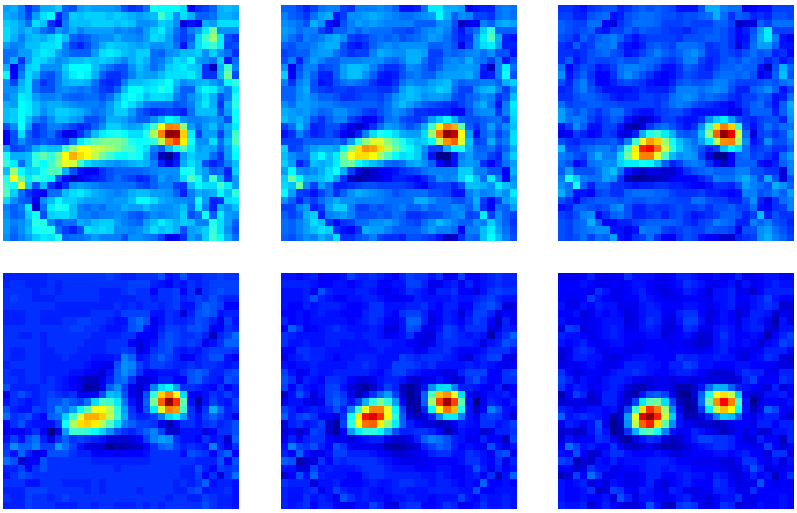


Figure 5.7: The LS estimates for the activity distribution computed using collimator model without attenuation correction (top row) and with attenuation correction (bottom row). The estimates in left side correspond to the situation of high attenuation coefficients and the estimates in right side correspond to situation of low attenuation coefficients.

the attenuating objects lowers the estimates without attenuation correction improve. Moreover, also the estimates with attenuation correction improve when the attenuation coefficients get smaller. This is due to fact that the ill-posedness of the inverse problem in SPECT gets worse, when attenuation increases.

5.3 Simulation study 3

In the third simulation we used a mathematical phantom with higher number of pixels: 64×64. Here we simulated the data from Poisson distribution. Thus, it is reasonable to use a statistic method in reconstruction and compare it to deterministic methods.
Figure 5.8 shows the phantom used in this simulation. It presents a cardiac phantom with a ring having activity 10 million s^{-1}/pixel. Elsewhere in a circular area the activity was 100000 s^{-1}/pixel and outside the circle the activity was 0. Attenuation coefficient was 0.015mm^{-1} inside the circular area and 0 in other parts of the phantom.
The dimensions of the collimator were as follows: width 40 cm, thickness 25 mm, the number of holes and projection bins 64 and the width of the square shaped holes 2 mm. The radius of rotation was 32.5

cm.

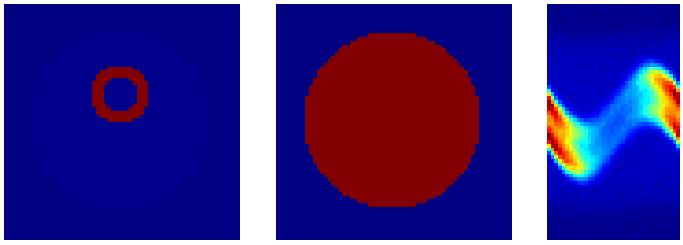


Figure 5.8: The activity distribution of the phantom (left), the attenuation map (middle) and the sinogram computed using Poisson statistics and additive Gaussian noise (right)

We computed the projection vectors P_k at 36 angles $\phi_k = 0^0, 10^0, \ldots, 350^0$ using Poisson statistics and additive Gaussian noise. That is, we first computed the expected values of P:

$$\hat{P} = Hf, \tag{5.2}$$

where H is the collimator modeled observation matrix with attenuation correction. Then we simulated the noise-free projection bins $\tilde{p}_i$ from Poisson distribution using the fact that $\tilde{p}_i \sim \text{Poisson}(\hat{p}_i)$. Finally, we added Gaussian noise to noise-free projection bins. The variance of the Gaussian noise was 3 % of mean of noise-free values $\tilde{p}_i$). The resulting projection bins were combined to a vector P presenting the data. The data P is presented in the form of sinogram in Figure 5.8.
We computed reconstructions using various methods and models. First we computed the reconstruction using filtered back projection method. No collimator or attenuation corrections were used in this reconstruction. We also computed the LS estimate for activity using the line integral modeled observation matrix and CGLS algorithm. The both reconstructions are shown in Figure 5.9.
Next we computed the LS estimate using CGLS algorithm with collimator modeled observation matrix with and without attenuation correction. In attenuation correction we used the attenuation map of the phantom. We computed also the regularized LS solution in the attenuation modeled case. We posed a smoothness prior for generalized Tikhonov regularization using a Laplacian filter as described in 3.2.3. These LS reconstructions are shown in Figure 5.10.
Finally we computed maximum likelihood estimates for the activity distribution using EM algorithm. The ML estimates were computed using all the observation matrices. The estimates are shown in figure 5.11.
We evaluated the reconstructions taking three different profiles of the images. The first profile presents the smoothness of the reconstructed activity in the ring where the top activity lies. The activity density along a circle is presented using a polar plot in Figures 5.9-5.11. In polar plot the radius at given angle gives the activity at same angle in the circle. The other profile was taken at x-direction in the height of the center of the circle and the third profile was taken at y-direction in the width of the center of the circle. For comparison, Figure 5.9 shows also the same profiles taken from the phantom.
The estimate obtained using filtered back projection is very rough. The line integral modeled LS estimate is superior when compared to FBP estimate. This can be seen both from the image and the profiles at x- and y-directions.
The non-regularized LS estimates with collimator and attenuation correction seem to be worse than the line integral modeled LS estimate. The active ring is imaged well but the smaller activity in the bigger circle is lost. Also, the both collimator modeled LS estimates have negative components with large magnitude and the images are not smooth. The regularization with smoothness prior improves the image considerably. The activity is quite smooth in the ring of high activity, and the circle with lower activity is also imaged. The explanation is that the LS problem with collimator and attenuation modeling is more ill-posed than LS problem with line integral model and regularization is needed. But also the regularized solution includes an unwanted negative gap outside the ring of high activity.

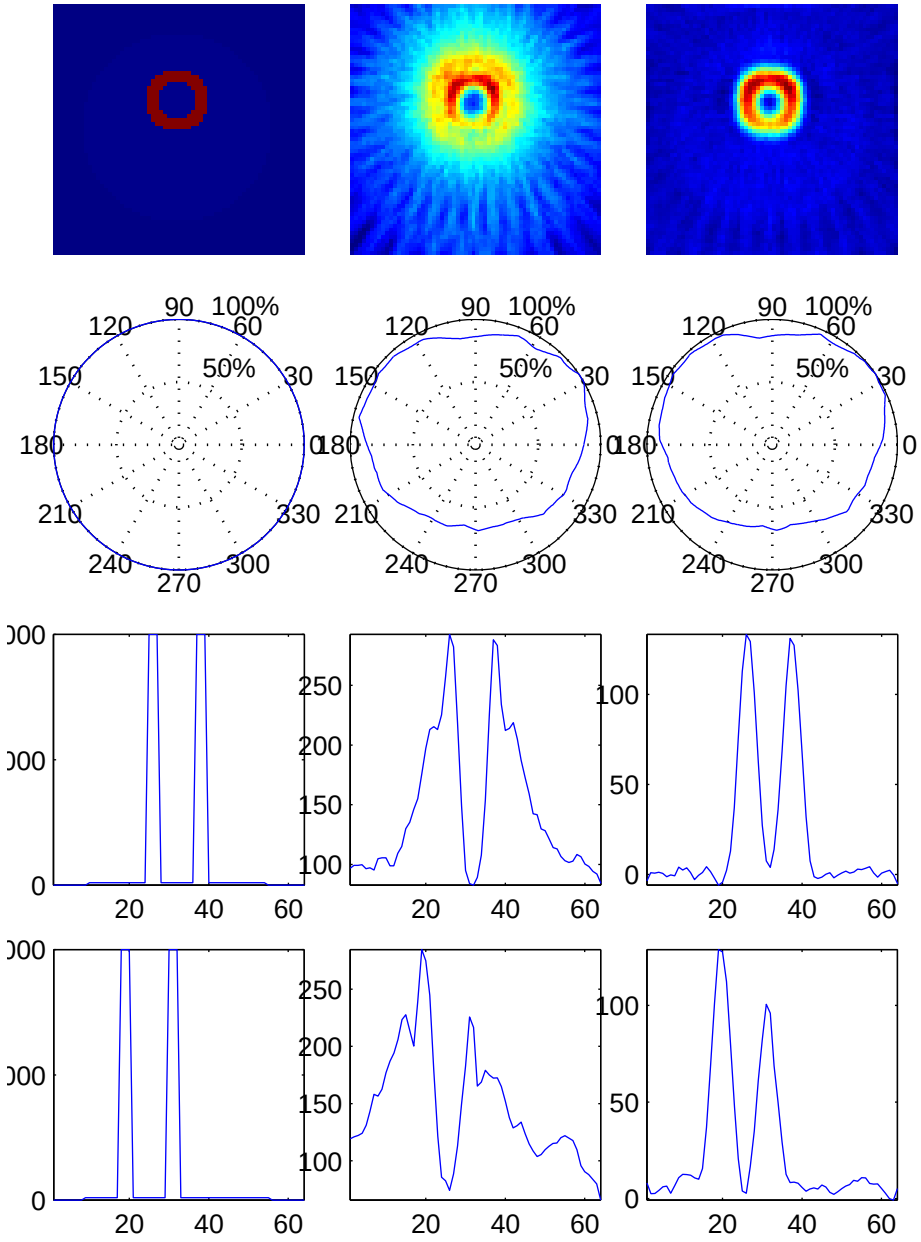


Figure 5.9: Top row - The phantom (left), filtered back projection estimate (middle) and line integral modeled LS estimate (right). Second row - The profiles of images along circles with maximum intensity. In the polar plot the radius gives the intensity as function of angle. Third row - Profiles of images along lines passing through the center of circle having x-direction. Fourth row - Profiles of images along lines passing through the center of circle having y-direction.

No doubt the ML estimates are superior over the LS estimates. All the ML estimates are quite smooth and both the ring with high activity and the circle with lower activity is imaged. The estimates without attenuation correction clearly underestimate the activity at the center of the phantom. The modeling of attenuation corrects this error and the ring is smoother in this estimate. Actually the activity is a bit overestimated at the center in this image. Also, the polar plot of this estimate is not perfectly circular. This seems to be due to effect of pixel size.

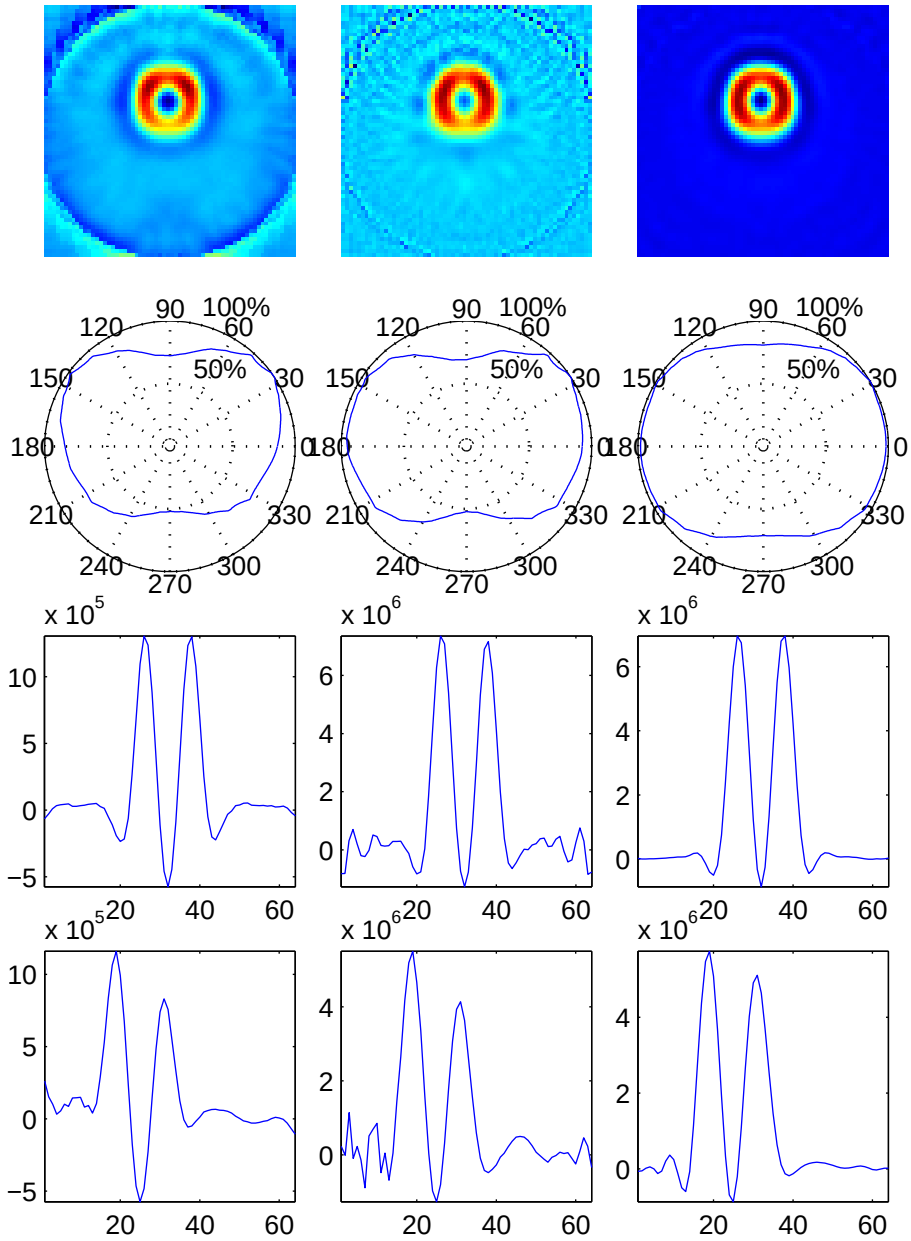


Figure 5.10: Images and profiles of the collimator modeled LS estimate (left), the collimator and attenuation modeled LS estimate (middle) and the Tikhonov regularized collimator and attenuation modeled LS estimate with smoothness prior (right).

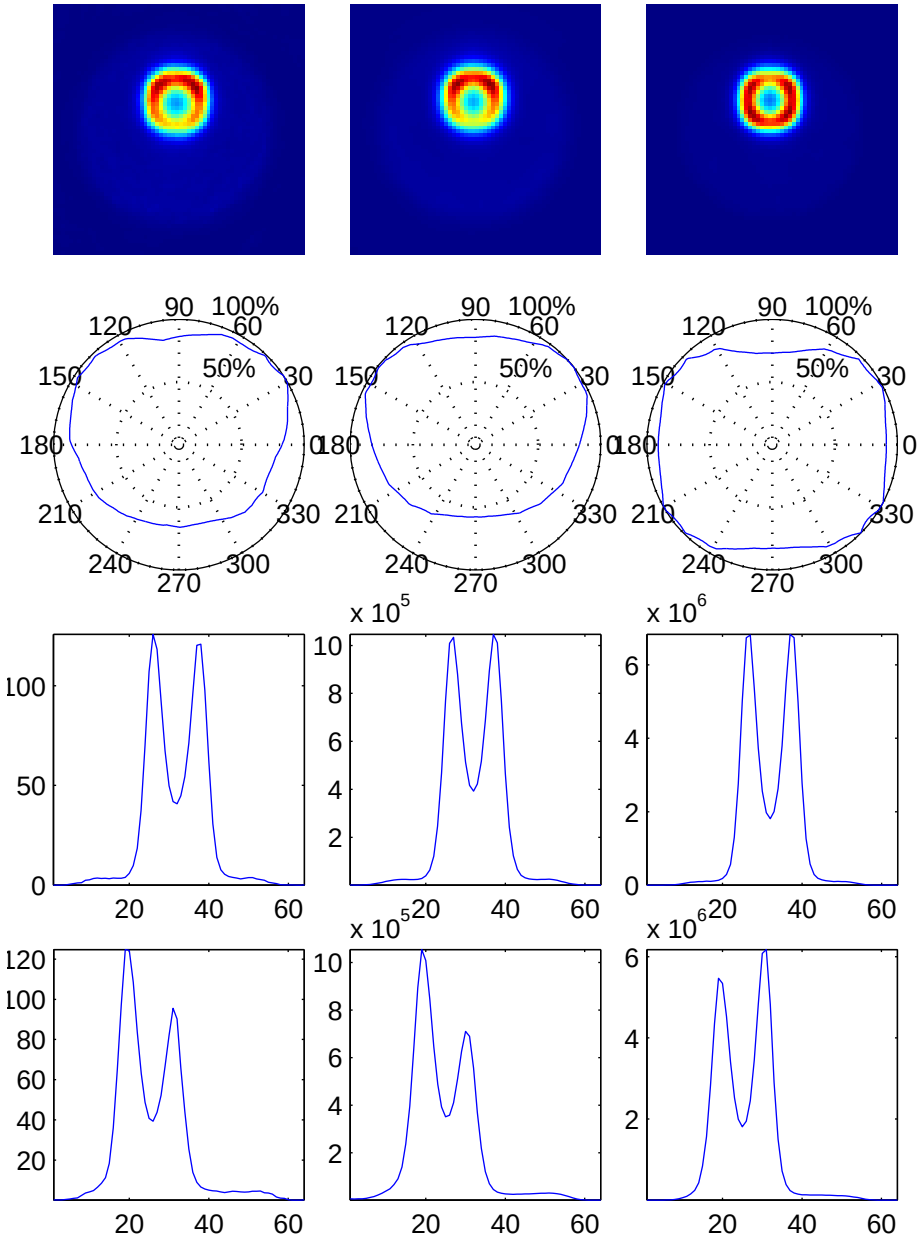


Figure 5.11: Images and profiles of the ML estimates with line integral model (left), collimator model (middle) and collimator model with attenuation correction (right).

5.4 Simulation study 4

In this section we present a 3D simulation. Here we used a three dimensional mathematical phantom with two active regions, other having activity 3s^{-1}/voxel and the other 1s^{-1}/voxel. In the other parts of the phantom the activity was zero. The phantom consisted of $16 \times 16 \times 9$ voxels. The size of the phantom was 32cm $\times$ 32cm $\times$ 18cm. The three dimensional presentation of the phantom is shown in Figure 5.12. Figure 5.13 represents the phantom slice-by-slice.

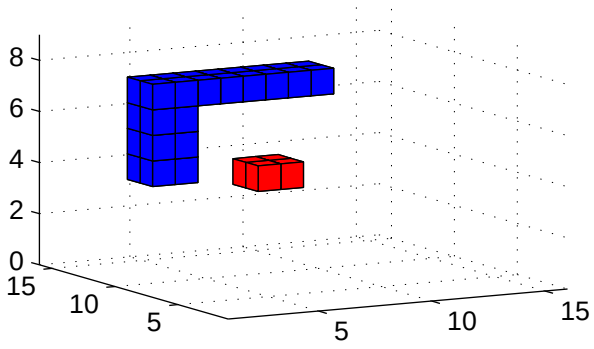


Figure 5.12: 3D presentation of phantom. Activity is 3s^{-1}/voxel in the red object and 1s^{-1}/voxel in the blue object.

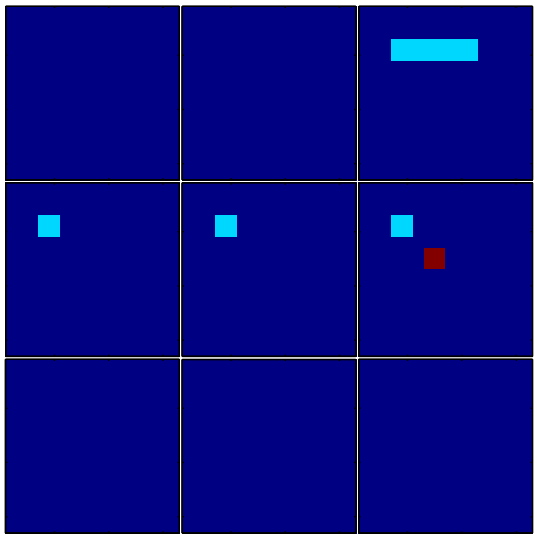


Figure 5.13: The slice-by-slice representation of the phantom. The image at top left presents the highest slice, the image at top middle is the second highest slice and so on.

The dimensions of collimator were as follows: area 32cm $\times$ 16cm, thickness 22 cm, the number of holes and the projection bins 16×9 and the diameter of the circular holes 2 cm. The radius of rotation was 38 cm. We computed the planar images at 36 angles $\phi_k = 0^0, 10^0, \ldots, 350^0$ using 3D collimator modeled observation matrix. The planar images were computed using the deterministic equation (5.1). Again we added Gaussian noise (variance 10 % of mean of projection bins p_i) to projection vector. Figure 5.4 shows the planar images at angles $0^0, 40^0, \ldots, 320^0$. We see clearly the 3D collimator effect from the planar images: projection of an object located in one slice only is spreaded also to neighboring slices.

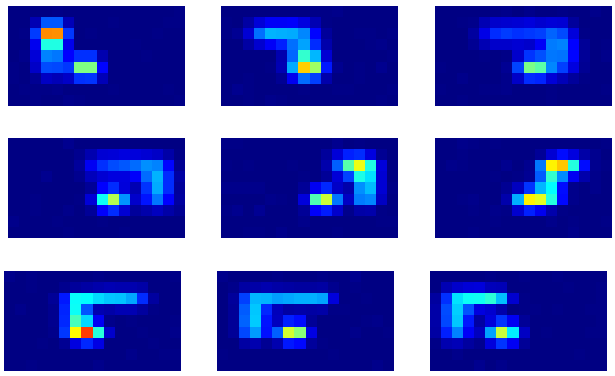


Figure 5.14: Planar images at angles $0^0, 40^0, \ldots, 320^0$

We computed the inverse problem using both 2D model and the 3D model. The simulated planar images with Gaussian noise were considered as data. The 2D slice-by-slice reconstruction was computed using line integral model in LS estimation. In 3D reconstruction we used the same 3D observation matrix as in the forward problem. Both reconstructions were computed using CGLS algorithm. The estimates are shown slice-by-slice in Figure 5.15.

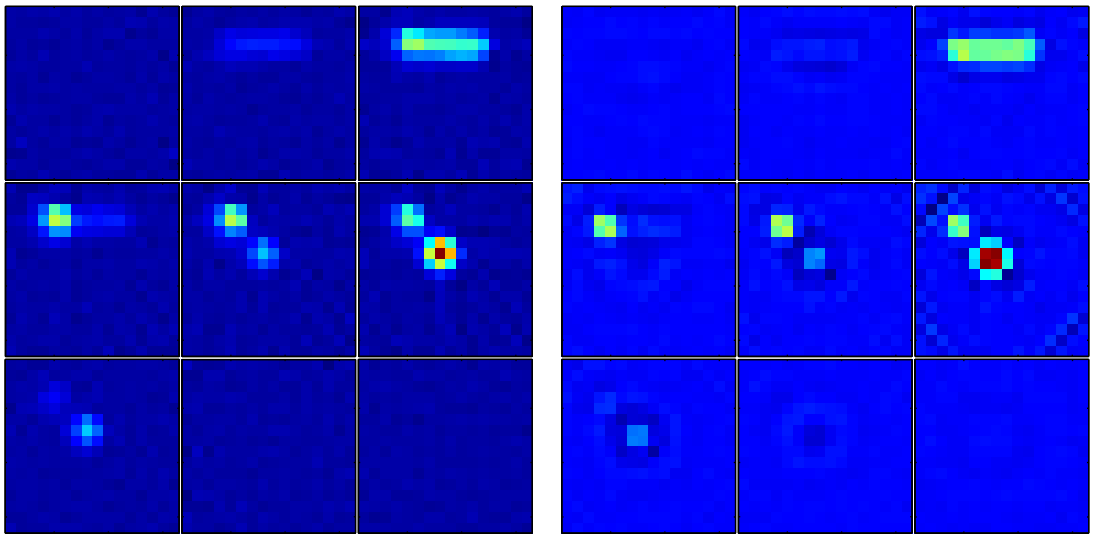


Figure 5.15: LS-estimates of the activity distribution. Left: 2D line integral modeled. Right: 3D collimator modeled.

The comparison of the 2D and 3D reconstructions shows the lack in two dimensional reconstruction. In 2D reconstruction all the counts in one row of planar image are considered as emitted from the slice at the same height. Now because the 3D collimator blur spreads the images, the 2D model causes ghosts in the slices where the activity is actually zero. There is also some ghosts in the 3D reconstruction, but they are considerably smaller. In addition, the shape of the objects with activity is well imaged. In this simulation we used the same 3D model both in the forward problem and the inverse problem. This can be considered as an inverse crime. However, the simulation illustrates the 3D collimator blurring effect.

5.5 Real phantom study

In this study we used real measurement data obtained from cardiac phantom measurements. The cardiac phantom consisted of two cylinders in a tank being one within the other. The outer cylinder was filled with water including high concentration of radioactive labels. The inner cylinder and the tank were filled with water having low radioactivity. Figure 5.16 shows the cross-sectional image of the phantom.

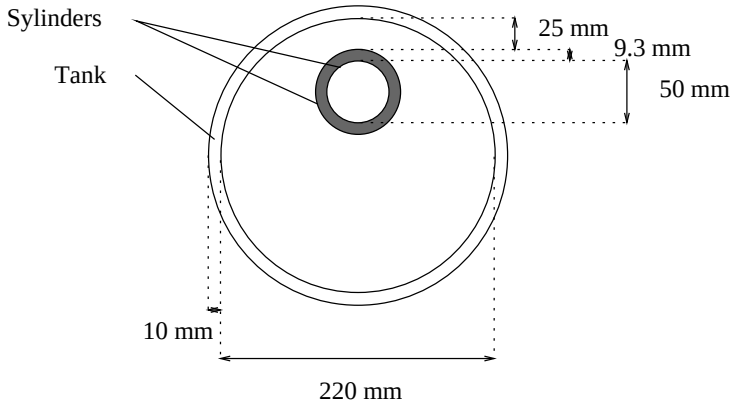


Figure 5.16: The dimensions of the cardiac phantom. Cross-sectional image of the cylindrical phantom, where the colored ring presents cross section of the compartment with higher activity.

The planar images were acquired at 64 angles: $0^0, 5.6250^0, \ldots, 354.3750^0$. The duration of the measurement of each planar image was 15s. Each planar image consists of 128×128 projection bins. Eight of the planar images are shown in Figure 5.17. The total number of counts was 484410112. The minimum value of projection bins was 0 and the maximum value was 35328. We selected one row of the planar images, the 60^{th} row. We formed a projection vector P including all the projection bins p_i in the 60^{th} row of the planar images. The vector P is presented in the form of sinogram in Figure 5.17.
Next we solved the activity distribution in the 60^{th} slice. To make the modeling of the collimator we selected an area of size 340mm $\times$ 340mm including the cross section of the tank. The selected area was divided in 128×128 pixels. We didn't have the true dimensions of the collimator, but we used the following dimensions to form the observation matrices. Width was approximated to be 614mm and thickness 24.05mm. The holes were approximated to be square shaped and having width 1.11mm. The radius of rotation was approximately 230 cm. We formed the observation matrices for line integral model and collimator model with and without attenuation correction. In attenuation correction we obtained the attenuation map so that we defined the pixels being inside the tank from the reconstruction without attenuation correction. In these pixels the attenuation coefficient was approximated to be 0.015mm^{-1} (the attenuation coefficient of water) and 0 in the other parts of the phantom.
We computed the ML estimates using EM algorithm with all the observation matrices. All the estimates are shown in Figure 5.18. As in Section 5.3 we took the three different profiles from images. Note, that the scale of x- and y- profiles differs in the case of line integral model. This is due to that discretization of the image domain is different in this case.
Again, the estimates without attenuation correction underestimate the activity at center of the phantom, whereas the estimate with attenuation correction overestimates a bit the activity at center. None of the estimates is too smooth but the one with both collimator and attenuation correction seems to be the smoothest.
Based on visual observation the estimate with both collimator and attenuation correction images best the form of the active ring. The ring seems to be too thick in the estimate obtained with line integral model. The quantitative results prove this: We defined the dimensions of the active ring from the x- and y- profiles of the estimates so that we measured the half widths of the peaks in profiles. Table 5.5 shows the results of measurements. We denote the diameter of the inner cylinder with d, and the thickness of

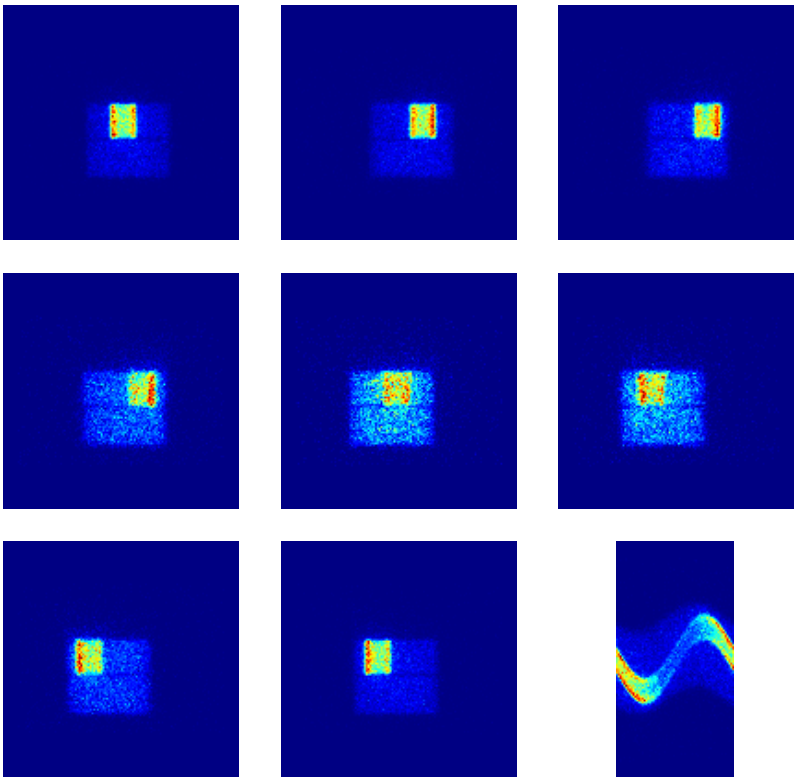


Figure 5.17: Planar images at angles $0^0, 45^0, \ldots, 315^0$ and sinogram corresponding to the 60^{th} row of planar images.

the active ring with s. The subindexes of s, 1 and 2 are for the first and the second edge in each profile. From Figure 5.16 we get that the real measures are $d = 50$mm and $s = 9.3$mm. In all the estimates the diameter d is underestimated and s is overestimated. However, the measures approach to the real values as the model is improved.

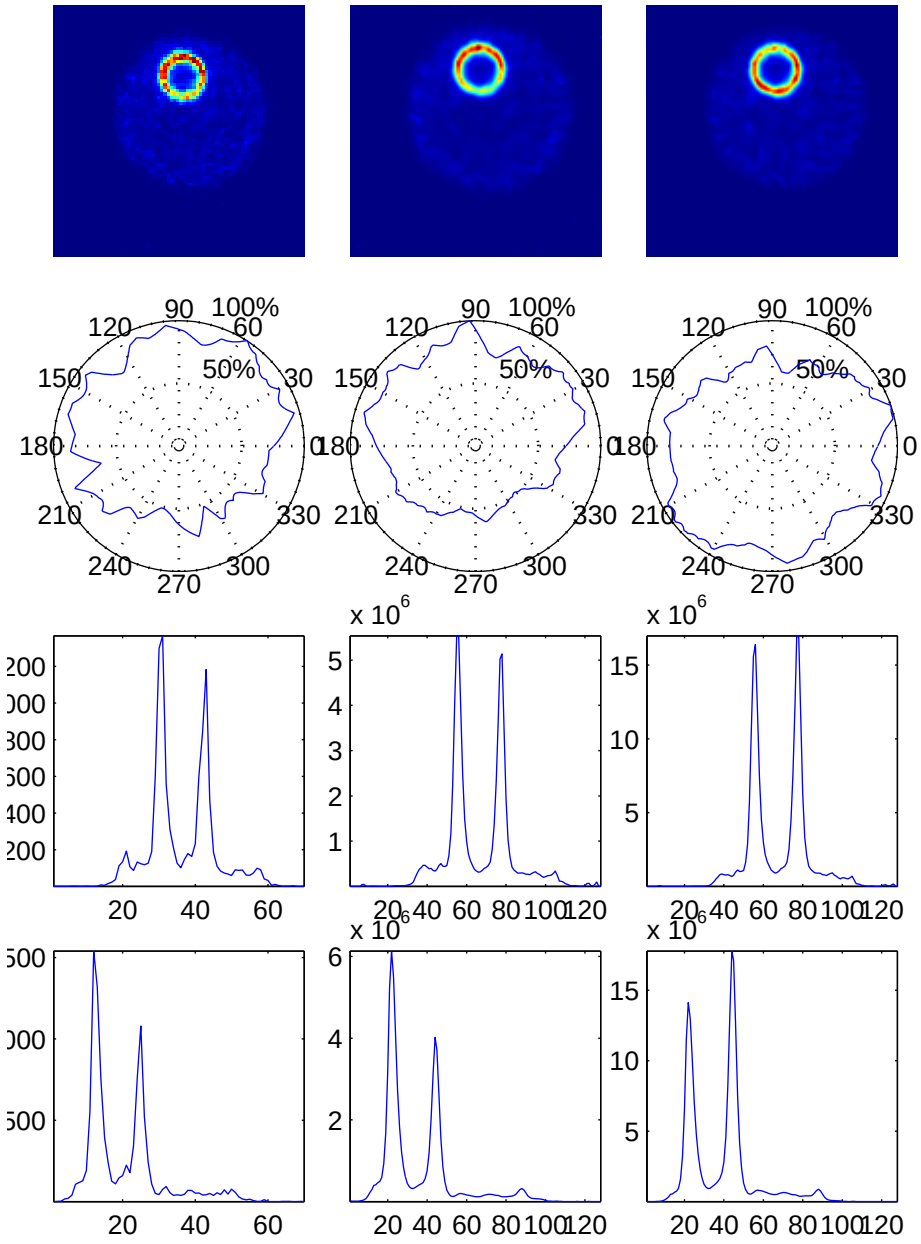


Figure 5.18: Images and profiles of the ML estimates for the activity of cardiac phantom. The estimates were computed using line integral model (left), collimator model (middle) and collimator model with attenuation correction.

Table 5.1: The measured widths. d is the measured diameter of the inner cylinder and s_i's stand for the measured thickness of the active ring with s. The real measures are $d = 50$mm and $s_1 = s_2 = 9.3$mm.

Model	Direction	s_1(mm)	d(mm)	s_2(mm)
Line integral	x	14.2	43.6	12.9
Line integral	y	14.6	43.7	12.9
Collimator	x	12.8	45.4	12.5
Collimator	y	13.7	45.2	12.7
Collimator and attenuation	x	12.2	45.4	12.3
Collimator and attenuation	y	13.5	45.0	12.7

Chapter VI

Discussion

We have considered various methods and models for SPECT. We noticed that the reconstruction methods of SPECT can be divided in three groups: the filtered back projection, the least squares estimation and the statistical estimation. Variety of different algorithms follows from the fact that the least squares solution and the statistical estimates can not be solved non-iteratively. Instead, some iterative algorithms are needed to get the solution. We are not particularly interested in the algorithms used in solution but the estimation criterions which the algorithms are used to satisfy. That is the reason why we compared the different estimation methods in Chapter 3. We obtained that filtered back projection solution can be considered as an approximation for the LS solution.

When using the LS methods or the ML estimation the key question is the modeling of the forward problem, that is, the construction of the observation matrix H. In this thesis we presented the modeling of some of the factors effecting to SPECT imaging. This was done because the modeling of the forward problem is not considered very often in literature despite of the fundamental role of it in reconstruction. We also presented the fully three dimensional modeling and reconstruction in this thesis. In 3D simulation we had to use very small dimensions, because in the fully three dimensional reconstruction the size of observation matrix tends to rise considerably when the dimensions get larger. That is why the fully 3D method is not yet applicable. But when the computers develop, the use of fully 3D modeling may improve the quality of SPECT imaging.

Simulation 3 gives us many interesting results. In the case of Simulation 3 we had a limited number of observations in comparison to number of pixels to be reconstructed. That is, the observation model was highly underdeterministic. As noticed in Section 3.3, in the case of the case of limited number of acquisition angles the FBP estimation can not provide a reconstruction with good quality. More interesting result was that the line integral modeled LS estimate was better than the LS estimates with collimator and attenuation modeling although the forward problem was modeled using the collimator model with attenuation correction. As we noted, this is due to that the inverse problem more ill-conditioned in the case of the collimator and attenuation model. Thus, the regularization is more important in the case of better modeling. One might consider the improvement of the modeling being useless as it leads to ill-posedness of the inverse problem. But we claim that instead of using a clumsy model to obtain a well-posed problem, it is preferable to use as perfect model as possible and then add all the prior information to reconstruction to get a well-posed problem as we do in the regularization.

The ML estimates show the importance of correct modeling of the forward problem. The reconstructions improve considerably when better models are used. The superiority of the ML estimates over the LS estimates can be caused by a couple of reasons. First, the statistics of decay is taken into account. We do not demand the equation $P = Hf$ to be satisfied as in LS estimation. The negative peaks in the LS estimates can be considered as some kind of "correction terms" which are needed to fulfill the deterministic equation. The other reason is the non-negativity of the solution obtained by EM algorithm. The non-negativity is actually the only true prior information that we know about the activity distribution. As noted in 3.2.3 the choice of this prior information leads to nonlinear problem and it is not used in LS estimation. But when using EM algorithm the estimate is always positive.

In this thesis we did not consider the MAP estimation in details. The use of MAP estimation would

give smooth estimates as Tikhonov regularization, but it is important to remember that the distribution is not necessarily smooth. The MAP estimation, as Tikhonov regularized estimation with smoothness priors should be tested also with non-smooth phantoms to obtain reliable results.

Bibliography

[1] F.J. Beekman and M.A. Viergever. Fast SPECT simulation including object shape dependent scatter. *IEEE Trans Med Imaging*, 14:271–282, 1995.

[2] M. Bertero. *Introduction to inverse problems in imaging*. IOP Publishing Ltd, 1998.

[3] J.E. Bowsher, V.E. Johnson, T.G. Turkington, R.J. Jaczczak, and Jr. C.E. Floyd. Bayesian reconstruction and use of anatomical *a priori* information for emission tomography. *IEEE Trans Med Imaging*, 15:673–686, 1996.

[4] V. Dicken. *Simultaneous activity and attenuation reconstruction in single photon emission computed tomography, a nonlinear ill-posed problem*. PhD thesis, Universität Potsdam, 1998.

[5] W. Gersch. Smoothness priors. In *New Directions in Time Series Analysis, Part II*, pages 113–146. Springer-Verlag, 1991.

[6] D.R Gilland, R.J. Jaszczak, H. Wang, T.G. Turkington, K.L. Greer, and R.E. Coleman. A 3D model of non-uniform attenuation and detector response for efficient iterative reconstruction in SPECT. *Phys Med Biol*, 39:547–561, 1994.

[7] G. Gindi, M. Lee, A. Rangarajan, and I.G. Zubal. Bayesian reconstruction of functional images using anatomical information as priors. *IEEE Trans Med Imaging*, 12:670–680, 1993.

[8] S.J. Glick, B.C. Penney, M.A. King, and C.L. byrne. Noniterative compensation for the distance-dependent detector response and photon attenuation in SPECT imaging. *IEEE Trans Med Imaging*, 13:363–374, 1994.

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

[10] C.W. Groetsch. *Inverse Problems in the Mathematical Sciences*. Vieweg, 1993.

[11] R. Guillemaud and P. Grangeat. A multifocal collimator with circularly distributed focal points for SPECT imaging. *IEEE Trans Nucl Sci*, 41:1473–1480, 1994.

[12] P.C. Hansen. *Regularization Tools: A Matlab package for analysis and solution of discrete ill-posed problems*. UNI-C, Technical university of Denmark, 1992.

[13] T. Hebert and R. Leahy. A generalized EM algorithm for 3-D Bayesian reconstruction from Poisson data using Gibbs priors. *IEEE Trans Med Imaging*, 8:194–202, 1989.

[14] T. J. Hebert and S. S. Gopal. The GEM MAP algorithm with 3-D SPECT system response. *IEEE Trans Med Imaging*, 11:81–90, 1992.

[15] G. T. Herman and A. Lent. Iterative reconstruction algorithms. *Comput Biol Med*, 6:273–294, 1976.

[16] YL. Hsieh, G.L. Zeng, and G.T. Gullberg. Projection space image reconstruction using strip functions to calculate pixels more 'natural' for modeling the geometric response of the SPECT collimator. *IEEE Trans Med Imaging*, 17:24–44, 1998.

[17] D.J. Kadrmas, E.C. Frey, S.S. Karimi, and B.M.W. Tsui. Fast implementations of reconstruction based scatter compensation in fully 3D SPECT image reconstruction. *Phys Med Biol*, 43:857–873, 1998.

[18] A.C. Kak and M. Slaney. *Principles of computerized tomographic imaging*. IEEE Press, 1987.

[19] P.A. Karjalainen. *Regularization and Bayesian methods for evoked potential esimation*. PhD thesis, University of Kuopio, Department of Applied Physics, 1997.

[20] P.A. Karjalainen, M. Vauhkonen, and J.P. Kaipio. Dynamic reconstruction in SPET. In *Proc 19th Int Conf IEEE Eng Med Biol Society*, pages 777–780, Chicago, October 30-November 2 1997.

[21] C.T. Kelley. *Iterative Methods for Linear and Nonlinear Equations*. SIAM, 1995.

[22] H-J. Kim, B.R. Zeeberg, F.H. Fahey, E.J. Hoffman, and R. Reba. 3-D SPECT simulations of a complex 3-D mathematical brain model: effects of 3-D geometric detector response, attenuation, scatter and statistical noise. *IEEE Trans Med Imaging*, 11:176–184, 1992.

[23] G.F. Knoll. Single-photon emission computed tomography. *Proc IEEE*, 71:320–329, 1983.

[24] J. Kulmala. *Absorbed radiation in targeted radionuclide therapy: a method based on transversal limited-projection emission tomography*. PhD thesis, 1997.

[25] V. La and P. Grangeat. Minimal residual cone-beam reconstruction with attenuation correction in SPECT. *Phys Med Biol*, 43:715–727, 1998.

[26] D.S. Lalush and B.M.W. Tsui. Simulation evaluation of Gibbs prior distributions for use in maximum *a posteriori* SPECT reconstructions. *IEEE Trans Med Imaging*, 11:267–275, 1992.

[27] K. Lange and R. Carson. EM reconstruction algorithms for emission and transmission tomography. *J Comput Assist Tomogr*, 8:306–316, 1984.

[28] C.L. Lawson and R.J. Hanson. *Solving Least Squares Problems*. SIAM, 1995.

[29] SJ. Lee, A. Rangarajan, and G. Gindi. Bayesian image reconstruction in SPECT using higher order mechanical models as priors. *IEEE Trans Med Imaging*, 14:669–680, 1995.

[30] Z. Liang, T.G. Turkington, D.R. Gilland, R.J. Jaszczak, and R.E. Coleman. Simultaneous compensation for attenuation, scatter and detector response for SPECT reconstruction in three dimensions. *Phys Med Biol*, 37:587–603, 1992.

[31] B. Lipinski, H. Herzog, E.R. Kops, W. Oberschelp, and H.W. Müller-Gärtner. Expectation maximization reconstruction of positron emission tomography images using anatomical magnetic resonance information. *IEEE Trans Med Imaging*, 16:129–136, 1997.

[32] J. Llacer. Theory of imaging with a very limited number of projections. *IEEE Trans Nucl Sci*, NS-26:596–602, 1979.

[33] J. Llacer. Tomographic image reconstruction by eigenvector decomposition: its limitations and areas of applicability. *IEEE Trans Med Imaging*, Ml-1:34–42, 1982.

[34] J. Llacer and J.D. Meng. Matrix-based image reconstruction methods for tomography. *IEEE Trans Nucl Sci*, 32:855–864, 1985.

[35] A. Mallon and P. Grangeat. Three-dimensional PET reconstruction with time-of-flight measurement. *Phys Med Biol*, 37:717–729, 1992.

[36] S.H. Manglos, F.D. Thomas, and R.B. Capone. Attenuation compensation of cone beam SPECT images using maximum likelihood reconstruction. *IEEE Trans Med Imaging*, 10:66–73, 1991.

[37] J. Nuyts, P. Dupont, S. Stroobants, R. Benninck, L. Mortelmans, and P. Suetens. Simultaneous maximum a posteriori reconstruction of attenuation and activity distributions from emission sinograms. *IEEE Trans Med Imaging*, 18:393–403, 1999.

[38] M. E. Phelps. Emission computed tomography. *Semin Nucl Med*, 7:337–365, 1977.

[39] Ph. Rizo, P. Grangeat, and R. Guillemaud. Geometric calibration method for multiple-head cone-beam SPECT system. *IEEE Trans Nucl Sci*, 41:2748–2757, 1994.

[40] F. D. Rollo. *Nuclear medicine physics, instrumentation, and agents.* The C. V. Mosby Company, 1977.

[41] L. A. Shepp and Y. Vardi. Maximum likelihood reconstruction for emission tomography. *IEEE Trans Med Imaging*, MI-1:113–122, 1982.

[42] M.F. Smith, jr C.E. Floyd, R.J. Jaszczak, and R.E. Coleman. Reconstruction of SPECT images using generalized matrix inverses. *IEEE Trans Med Imaging*, 11:165–175, 1992.

[43] M.F. Smith, jr C.E. Floyd, R.J. Jaszczak, and R.E. Coleman. Three-dimensional photon detection kernels and their application to SPECT reconstruction. *Phys Med Biol*, 37:605–622, 1992.

[44] R. Thierry, J-L. Pettier, and L. Desbat. Simultaneous compensation for attenuation, scatter and detector response for 2D-emission tomography on nuclear waste with reduced data. In *Proc. 1st world congress on industrial process tomography*, 1999.

[45] B.M.W. Tsui, E.C. Frey, X. Zhao, D.S. Lalush, R.E. Johnston, and W.H. McCartney. The importance and implementation of accurate 3D compensation methods for quantitative SPECT. *Phys Med Biol*, 39:509–530, 1994.

[46] M. Vauhkonen. *Electrical Impedance Tomography and Prior Information.* PhD thesis, University of Kuopio, Kuopio, Finland, 1997.

[47] X. Wang and K.F. Koral. A regularized deconvolution-fitting method for Compton-scatter correction in SPECT. *IEEE Trans Med Imaging*, 11:351–360, 1992.

[48] S. Webb. *The physics of medical imaging.* IOP Publishing Ltd, 1988.

[49] G.L. Zeng and G.T. Gullberg. Iterative and analytical reconstruction algorithms for varying-focal-length cone-beam projections. *Phys Med Biol*, 43:811–821, 1998.