Positron Emission Tomography with Three-dimensional reconstruction

Kjell Erlandsson

Radiation Physics Department
The Jubileum Institute
Lund University
Sweden

1996

Doctoral Dissertation 1996

Radiation Physics Department
The Jubileum Institute
Lund University

Printed in Sweden by
KF-Sigma, Lund, 1996
ISBN 91-628-2189-X

"I'm guided by the beauty
of our weapons."

Leonard Cohen

LIST OF ORIGINAL PAPERS

This thesis is based on the following papers, which will be referred to in the text by their Roman numerals:

I. Sandell A, Erlandsson K, Ohlsson T, Strand S-E, "A PET Scanner Based on Rotating Scintillation Cameras: Development and Experience", submitted to J Nucl Med, 1996.

II. Erlandsson K, Strand S-E, "A New Approach to Three-Dimensional Image Reconstruction in PET", IEEE Trans Nucl Sci, 39:1438-1443, 1992.

III. Erlandsson K, Esser PD, Strand S-E, van Heertum RL, "3D Reconstruction for a Multi-Ring PET Scanner by Single-Slice Rebinning and Axial Deconvolution", Phys Med Biol, 39:619-629, 1994.

IV. Erlandsson K, Strand S-E, "3D Reconstruction for 2D PET", submitted to Nucl Instr Meth A, 1996.

V. Sandell A, Erlandsson K, Bertenstam L, "A Scanner for Positron Emission Mammography", submitted to Phys Med Biol, 1996.

ABSTRACT

In this thesis, the developments of two different low-cost scanners for positron emission tomography (PET), based on 3D acquisition, are presented. 3D reconstruction methods for these systems were also developed. The first scanner consists of two rotating scintillation cameras, and produces quantitative images, which have shown to be clinically useful . The second one is a system with two opposed sets of detectors, based on the limited angle tomography principle, dedicated for mammographic studies. The development of low-cost PET scanners can increase the clinical impact of PET, which is an expensive modality, only available at a few centres world-wide and mainly used as a research tool.

The conventional way to image a 3D volume in PET, by acquisition of data in a set of parallel slices, which are reconstructed as independent planes, makes an inefficient use of the available photon fluence. With 3D acquisition, also photons which traverse more than one transaxial slice are detected. In this case, fully 3D reconstruction is needed. The main problems in 3D reconstruction for PET are the size of the data sets, the computation time and how to theoretically incorporate all the data.

A 3D reconstruction method was developed that utilizes all the available data. The size of the data-sets is considerably reduced, using the single-slice rebinning approximation. The 3D reconstruction is divided into 1D axial deconvolution and 2D transaxial reconstruction, which makes it relatively fast. This method was developed for the rotating scanner, but was also implemented for multi-ring scanners with and without inter-plane septa. An iterative 3D reconstruction method was developed for the limited angle scanner, based on the new concept of "mobile pixels", which reduces the finite pixel errors and leads to an improved signal to noise ratio.

CONTENTS

AIM OF STUDY

The aims of the present study can be summarized as follows:

◆ Development of a PET scanner based on two rotating scintillation cameras using 3D data acquisition and reconstruction, including correction methods for obtaining quantitative images.

◆ Development of a 3D reconstruction method for PET capable of utilizing all available data. The single-slice rebinning approximation was used in order to reduces the size of the projection data-sets. The 3D reconstruction problem was divided into 1D axial deconvolution and 2D transaxial reconstruction, making it relatively fast.

◆ Implementation of this 3D reconstruction method for a multi-ring scanner with the interplane septa removed in order to get a higher sensitivity, and for a 2D multi-ring scanner, with the aim of obtaining an improved axial resolution.

◆ Development of a simple scanner dedicated for positron emission mammography with two opposed sets of detectors, based on the limited angle tomography principle, and of an iterative 3D reconstruction method for this scanner, based on the new concept of "mobile pixels", which reduces the finite pixel errors and leads to an improved signal to noise ratio.

1. INTRODUCTION

Positron emission tomography is a non-invasive technique for imaging the biodistribution of radionuclide-labelled molecules in vivo. Using tracer amounts of a radio-labelled molecule, quantitative physiological and biochemical information can be obtained from PET images.

The radionuclides used in PET are neutron-deficient as compared to their stable isotopes. They decay by transformation of a nuclear proton into a neutron through emission of a positron and a neutrino. The positron interacts with an atomic electron, and both particles annihilate with emission of two γ-photons, which can be utilized for PET measurements.

The most commonly used radionuclides are ^{11}C, ^{13}N, ^{15}O and ^{18}F. They are produced by a beam of accelerated ions, resulting in carrier-free preparations with high specific activity. Since they represent biologically important elements, they can be used for labelling biomolecules without changing their biological properties, and thereby offer a particularly attractive method for the regional study of biochemical activity.

Many of the nuclides used have short half-lifes (^{11}C - 20 min, ^{13}N - 10 min, ^{15}O - 2 min, ^{18}F - 110 min), which means that more information is obtained for a given absorbed dose of radiation to the patient. It also gives the possibility to make repeated studies on the same patient without interference between different measurements. However, fast and simple chemical routines are needed.

PET can be applied to many areas of medicine. It is especially useful for neurological and psychiatric disorders, heart disease and cancer. The interpretation of PET images often requires physiological modelling. The biological and physiological parameters possible to measure with PET include blood flow, fatty acid and glucose utilisation, oxygen metabolism, protein synthesis and receptor density. This allows the study of diseases such as epilepsy, stroke, brain tumours and dementia. Also the mechanisms behind normal brain processes, such as language processing, speech and vision, as well as brain development can be studied. The heart applications are myocardial perfusion and myocardial viability. In oncology, PET studies are valuable for determining the extent of tumours prior to surgery, differentiating radiation necrosis from viable tumour, assessing the degree of malignancy of lesions, and

their response to therapy.

PET is an expensive modality. It generally requires an on-site cyclotron for production of radionuclides, an imaging device (scanner), in-house labelling of molecules and a team of experts to handle the different steps. Until now PET has mainly been used as a research tool, but the interest for clinical PET is increasing. There are a number of ways to increase to clinical impact of PET: Generator-produced positron-emitters (e.g. $^{62}Zn/^{62}Cu$, $^{68}Ge/^{68}Ga$, $^{82}Sr/^{82}Rb$), regional distribution centres for ^{18}F, automated synthesis units, low-cost low-energy accelerators, and low-cost PET scanners. The development of a low-cost clinical PET facility has been described by Ohlsson (1996).

2. PET SCANNERS

2.1 Annihilation coincidence detection

The radionuclides used in PET decay by emission of a β^+-particle, a positron, which is the anti-particle of the electron. After travelling a short distance and being slowed down almost to the thermal energy level, the positron interacts with an atomic electron, and the two particles are annihilated. Their mass-energy is converted to two 511 keV γ-photons, emitted in opposite directions. The fundamental principle of PET is to use the annihilation photons to get an image of the activity distribution by tomographic measurements (see e.g. Ott et al., 1988).

Due to the law of conservation of momentum, the relative angle of the annihilation photons in the electron-positron centre of mass system is exactly 180°. However, due to the small remaining kinetic energy, in the laboratory system there will be a slight deviation from 180° by approximately 0.5°. This effect together with the range of the positron before annihilation (about 1 mm, depending on the radionuclide) sets a physical limit to the spatial resolution obtainable with PET.

PET scanners are composed of at least two detectors and based on the annihilation coincidence detection (ACD) principle. This means that, if two photons simultaneously trigger two different detectors, an annihilation is assumed to have occurred somewhere along the line joining the two detection points, and the coincidence event is registered. In this case, "simultaneously" means within a time window, 10-60 ns wide.

PET scanners can be divided into two groups; those based on a small number of large area position sensitive detectors, and those based on a large number of small individual detectors (ring scanners). In the following sections a review will be made of some of the PET scanners that have been developed within these two groups over the years.

2.2 Area detector scanners

A PET scanner can be constructed with two opposed large area detectors connected in coincidence mode (figure 2.1a). A 2D coordinate from each

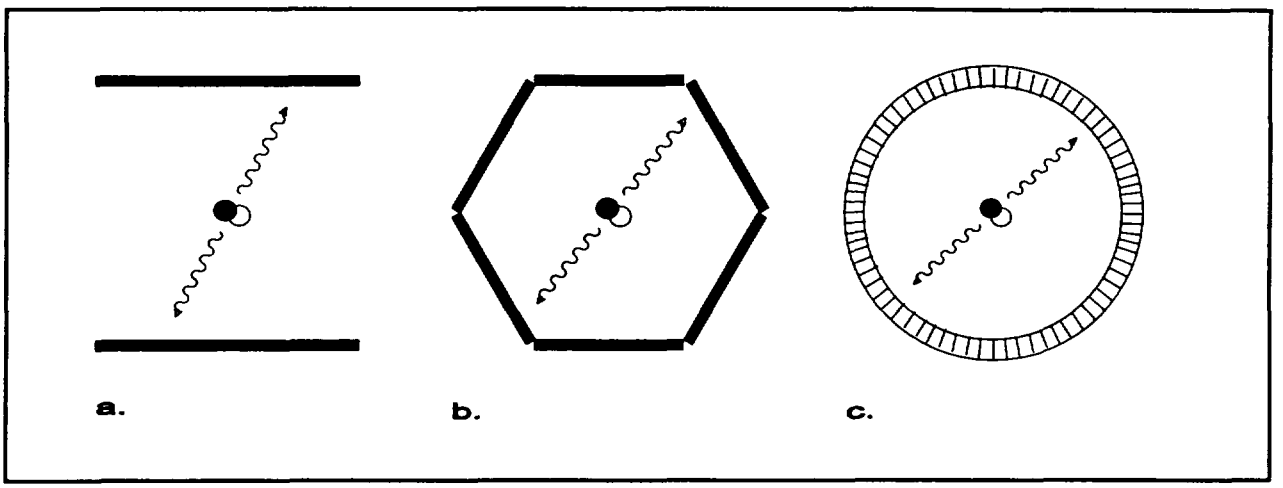

Figure 2.1: PET scanners, a) planar detectors, b)hexagonal geometry, c) detector ring.

detector is obtained for each coincidence detection event. With such a detector system, coincidence lines in a number of different angles will be registered, resulting in tomographic measurements. Two different types of area detectors have been used for PET scanners; scintillation cameras and multi-wire proportional chambers (MWPC).

The scintillation camera, developed by Anger (1967), is a position sensitive scintillation detector. A large, thin NaI(Tl)-crystal is coupled to a number of photo-multiplier (PM) tubes. The relative signals from the different PM-tubes are used to produce position coordinates. Anger (1967) suggested using two scintillation cameras for ACD measurements, using the "focal-plane" method to produce longitudinal images parallel to the detector planes. This procedure is equivalent to backprojection (see below). Muehllehner et al. (1976 and 1977) developed a positron camera based on two scintillation cameras with 2.54 cm thick crystals, which could be used in both stationary and rotational mode. In stationary mode, tomographic data is only obtained in a limited angular interval, which results in a lower resolution in the direction perpendicular to the detector planes. This type of measurement is called limited angle tomography. By rotating the detectors, a complete tomographic data set is obtained, and transaxial images can be reconstructed.

A positron camera was later developed based on six scintillation camera-type detectors in a hexagonal arrangement, as shown in figure 2.1b (Muehllehner and Karp, 1986, Muehllehner et al., 1988). This scanner gave a full tomographic data set without rotation, except for some missing data in the gaps between the detectors. Rotating dual headed scintillation cameras have for some time now been used for single photon emission computed

tomography (SPECT) studies. Recently, two such commercial systems have been modified for ACD measurements (Nellemann et al., 1996, Miyaoka et al., 1996). However, the crystal-thickness used is only 1 cm, which is optimal for SPECT with photon energies around 150 keV, but gives a low efficiency for 511 keV annihilation photons. In paper I, a PET scanner which was developed in Lund, based on two rotating scintillation cameras, is presented. A preliminary description of this system was previously made (Sandell et al., 1992).

Another type of large area detector, which has been used in PET, is the multi-wire proportional chamber (Bateman et al., 1980, Jeavons et al., 1983). These detectors have the advantage of a good spatial resolution and a low cost, but also the disadvantage of a low intrinsic efficiency. PET scanners based on MWPCs have been shown to be clinically useful (Flower et al., 1984, Townsend et al., 1987, Cherry et al., 1989). However, a significantly higher sensitivity would be desirable. A new positron camera is presently being developed based on hybrid BaF_2-MWPC detectors, which combines the high sensitivity of the BaF_2 scintillation crystal and the high resolution of the MWPC (Visvikis et al., 1995).

2.3 Ring detector scanners

Another approach for the design of a PET scanner is to use a large number of small individual detectors surrounding the patient. This line of development started when Phelps et al. (1975) constructed a scanner based on 24 NaI(Tl) detectors arranged in a hexagonal array. Each detector was connected in coincidence with one opposing detector. The detectors in the three different angular views were placed at different radial positions, so that a 360° rotation gave tomographic data with adequate radial and angular sampling density. Phelps et al. (1978) developed a new hexagonal scanner with 66 NaI(Tl) detectors. Each detector was now in coincidence with 11 opposing detectors. Adequate sampling was obtained by linear translation and rotation over 60°.

A PET scanner with a ring geometry (figure 2.1c) was built by Bohm et al. (1978) with 95 NaI(Tl) detectors, each one connected in coincidence with 40 others on the opposite side of the ring. A new technique called "wobbling" was introduced, to get an improved radial sampling. The detector system moved so that its centre described a small circle with a diameter equal to the

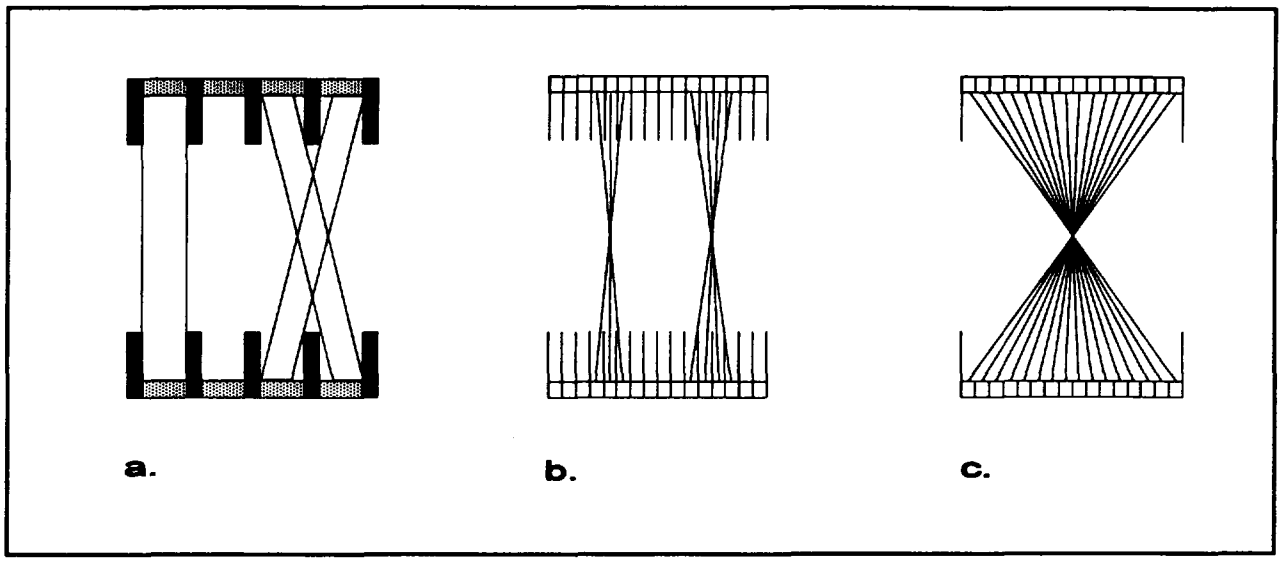

Figure 2.2: Axial ring scanners geometries, a) 4 rings with septa, b) 16 rings with septa, 2D mode, c) 16 rings without septa, 3D mode. The axial dimension has been exaggerated.

detector spacing. A multi-ring PET scanner was developed by Litton et al. (1984), consisting of 4 rings with 96 detectors each. The scintillation material bismuth germanate oxide (BGO) was used instead of NaI(Tl) due to its higher intrinsic efficiency for 511 keV photons. Each detector was in coincidence with 36 opposing detectors in the same ring and 36 in adjacent rings. With this arrangement 7 transaxial planes can be imaged at once. The data acquired from coincidences between detectors in the same ring represents 4 "direct planes". Coincidences between adjacent rings represent 3 "cross-planes", which are treated as transaxial planes parallel to and midway between the direct planes (see figure 2.2a). Between the detector rings, lead shields called "septa" were placed to reduce the number of unwanted scattered and random events (see below). Wobbling was still used to improve the radial sampling density. In paper IV we present an implementation of 3D reconstruction for this scanner.

Over the years, the trend in the development of multi-ring scanners has been to use smaller and smaller detectors, in order to improve the spatial resolution. However, with individual crystal-to-PM-tube coupling, the size and cost of the PM-tubes represents a limiting factor. This has lead to the development of block-detectors, consisting of a block of BGO, which is cut into e.g. 8x8 elements and viewed by 4 PM-tubes. Anger-logic is used for identification of the individual block elements, which in practice function as individual detectors. A system with 16 rings using block detectors was described by Spinks et al. (1992). In order to obtain sufficient sensitivity in each plane, 3 and 4 ring combinations were used to form direct and cross-planes,

respectively, as shown in figure 2.2b. With this scanner, wobbling was still available, but with newer scanners, with even smaller detectors, it is no longer necessary.

With improved spacial resolution it is important to also have an increased sensitivity. This can be obtained by removing the inter-plane septa and utilizing all possible ring combinations, as shown in figure 2.2c. The sensitivity will increase partly due to more coincidence lines, partly due to less shielding from the septa. However, there will be a variation in sensitivity within the field of view (FOV). Monte Carlo simulations were made by Dahlbom et al. (1989) to study the advantages and disadvantages of using 3D acquisition with a multi-ring scanner. Although the sensitivity increases significantly, the number of scattered and random coincidences also increases. The count rate capability of the system will remain the same, however, and the maximum count rate will be reached for a lower amount of activity in the FOV. 3D data acquisition requires 3D image reconstruction. In paper II, a 3D reconstruction method is presented, which was developed for a rotating PET scanner. In paper III, an implementation of this method for a multi-ring scanner is presented.

Individual detectors can also be used in other-than-ring configurations. Townsend et al. (1993) developed a rotating PET scanner based on opposed arrays of block detectors covering only a third part of the circumference around the patient. The principle is the same as for the rotating scanners with planar detectors described above. In paper V, a PET scanner is described, which consists of two opposed arrays of individual detectors and is based on the limited angle tomography principle.

2.4 Data corrections

There are a number of corrections which must be applied to the data obtained from a PET scanner before the tomographic reconstruction is made. The effects that must be taken into account are detector efficiency variations, electronic dead-time, random coincidences, photon scattering and photon attenuation. Some of these corrections may alternatively be performed during or after the reconstruction process. In papers I and V, the implementation of these corrections for two different PET scanners is presented.

In a scanner composed of many individual detectors, the intrinsic efficiency for

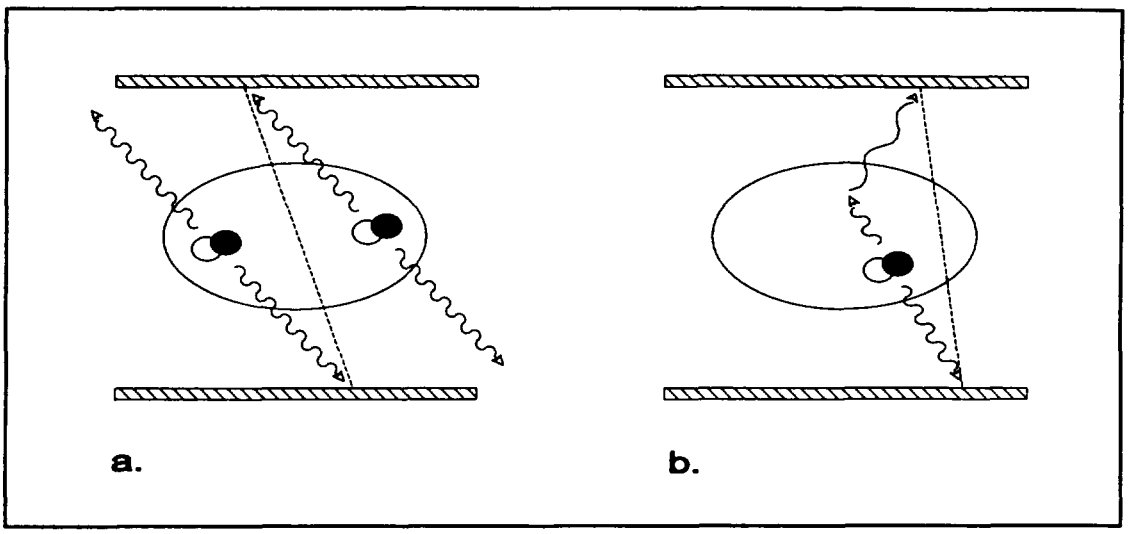

Figure 2.3: Background events, a) random coincidence, b) scattered photon.

each detector will not be the same. Also for large area detectors, the efficiency may vary over the detector plane. To correct for this effect, the acquired data is normalized using a measurement with a uniform plane or ring source. For high count rates, the response of the system will no longer be a linear function of the incoming photon fluence rate due to electronic dead-time. Using a model of the count rate performance of the system, based on the singles and coincidence rates, a correction can be made.

There is no natural background in PET measurements due to the coincidence requirement. However, there are two effects that creates background events during a PET scan: random coincidences and Compton scattering in the object. Since the coincidence detection technique is based on a finite time window of 10-60 ns, there is always a chance that two photons originating from different annihilations may cause a coincidence event, as shown in figure 2.3a. The number of these "random" or "accidental" coincidences can be calculated using equation (2.1).

$$n_R = n_1 n_2 2\tau \tag{2.1}$$

where n_R is the count rate of random coincidences between detectors 1 and 2, n_1 and n_2 are the single count rates in each detector, and 2τ is the time window (the maximum time difference between the two registrations is τ).

It is also possible to measure the random count rate using a delayed coincidence window, which means that one signal is delayed so that no true coincidences can be detected. Casey and Hoffman (1986) presented a method

for reducing the noise in random coincidence or normalization measurements, utilizing the correlation between the different coincidence lines.

The annihilation photons may interact with atomic electrons in the patient either by the photo-electric absorption or by Compton (incoherent) scattering. In the latter process (see figure 2.3b), the photon changes direction and looses some of its energy. If detected, the two photons will still be in coincidence. The photon energies before ($h\nu$) and after ($h\nu'$) scattering by a polar angle θ are related by the following equation:

$$h\nu' = h\nu\left(1 + \frac{h\nu}{m_o c^2}(1-\cos\theta)\right)^{-1} \qquad (2.2)$$

where $m_o c^2$ is the rest mass of the electron (511 keV). The energy resolution of the detectors used in PET is not sufficiently high to allow efficient discrimination of scattered photons, although NaI(Tl) is better than BGO in this sense.

Scatter correction method have been developed for both 2D and 3D PET. Bergström et al. (1983) showed that the point scatter distribution in the 1D projections from a 2D PET scanner can be described by a monoexponential function. They developed a correction method in which the scattered component was calculated by convolution of the measured data with a position dependent kernel and subsequently subtracted. Bailey & Meikle (1994) used a 2D monoexponential function for 3D scatter correction. Lercher and Wienhard (1994) described the scatter distribution in the 2D projections from a 3D PET scanner by a 2D Gaussian function. Cherry et al. (1993) developed a 3D correction method for a PET scanner with retractable septa, based on a comparison between data acquired in 2D mode with septa in place and in 3D mode with septa retracted. Ott et al. (1995) developed a method for scatter correction in 3D backprojected data in object space.

Attenuation correction is essential for an accurate reconstruction (Huang et al., 1979). Some of the annihilation photons emitted inside the patient will never reach the detectors due to photo-electric absorption or Compton scattering in the patient. For each photon there is a certain probability of this happening. For a large number of photons, the fraction passing through an object of thickness d is given by the attenuation factor $e^{-\mu d}$, where μ is the linear attenuation coefficient. The total path length for the two annihilation photons

through an object along a certain coincidence line is equal to the object thickness along that line. If the attenuation properties of the object are uniformly distributed and the contour of the object is known, the attenuation factors for each coincidence line can be calculated. Alternatively, the attenuation factors can be measured with an external source. The latter method is more accurate, but requires an extra measurement, which takes time and adds noise to the final image.

Bergström et al. (1982) developed a method for determination of the head contour in PET scans of the brain, based on the activity distribution in the projections. Huang et al. (1981) developed a hybrid method based on a short transmission scan, segmentation of the body into regions with different attenuation coefficients, and calculation of the attenuation factors. A transmission scan is preferably made before administration of the radiopharmaceutical to the patient. However, this may not be convenient for studies that require a long uptake period. In this case, the transmission scan can be made after administration if corrections are made for the emitted photons (Carson et al., 1988).

2.5 Monte Carlo simulations

Some of the results presented in papers II-IV were obtained by Monte Carlo (MC) simulation. This technique is based on the use of random numbers to simulate events with a known probability distribution. In most cases, only simple geometric simulations were made to test the 3D reconstruction methods. However, for MC simulations of basic spread functions (BSF) used in the reconstructions, photon interactions were also simulated.

For geometric simulations, a phantom was first defined, consisting of different regions with uniform activity distributions with different concentrations. The problems were usually rotationally invariant, and the simulations were then made for one parallel projection at a time. A random annihilation position was generated in the phantom. A direction of emission was simulated by generating a random polar angle. Two photons, emitted in opposite directions, were traced. The detector efficiencies were assumed to be 100%. If both photons hit a detector, a coincidence event was recorded.

For BSF simulations, Compton and photo-electric interactions in the lead septa

(if present) and in the scintillation crystals were simulated. When a photon entered a material, a random distance, d, travelled before interaction was generated from the relation $R=(1-e^{-\mu d})$, where R is a random number in the interval [0,1). An interaction type was randomly selected based on the relative probabilities. If a Compton interaction had occurred, a random scattering angle, θ, was generated using Kahn's method, based on the Klein-Nishina cross-section (see e.g. Zerby, 1963). The energy of the scattered photon was then given by equation (2.2). When a photon had deposited all or part of its energy in a detector, the energy resolution of the scanner was simulated by randomly selecting an energy from a Gaussian distribution centred at the energy deposited. If this was above the predefined energy discriminator level, a detection had occurred.

3. ANALYTICAL 2D RECONSTRUCTION

In tomographic measurements, line integrals in different directions of a 2D or 3D function are obtained. By line integral is meant the integral of the function-values along a straight line. The fundamental problem of tomographic reconstruction is to recover the function from its line integrals. In the case of emission computed tomography (PET and SPECT), the function to be recovered is the distribution of a radioactive tracer inside a patient. In the case of transmission computed tomography (TCT or CT), it is the x-ray attenuation property of the material (the "electron-density"), and in magnetic resonance imaging (MRI) it is the proton-density (to make a simplification). Tomographic methods can also be used in areas other than medical diagnosis, e.g. electron microscopy, industrial testing, geology and astronomy.

There are several ways to approach the reconstruction problem. The algorithms used can generally be sorted into two categories: analytical and algebraic methods. There are many good reviews on 2D analytical image reconstruction, see e.g. Herman (1980), Lewitt (1983), Kak & Slaney (1988). Here the basic theory for analytical reconstruction methods will be presented.

3.1 The 2D Radon transform

Different geometries can be used in tomographic measurements. In PET the parallel beam geometry is normally used. In the following we will use the two Cartesian coordinate systems (x,y) and (r,s), which are related by rotation by an angle ϕ (see figure 3.1a). This relation can be formulated as follows:

$$\begin{cases} x = r\cdot\cos\phi - s\cdot\sin\phi \\ y = r\cdot\sin\phi + s\cdot\cos\phi \end{cases} \Leftrightarrow \begin{cases} r = x\cdot\cos\phi + y\cdot\sin\phi \\ s = -x\cdot\sin\phi + y\cdot\cos\phi \end{cases} \tag{3.1}$$

A parallel projection, p, of a two-dimensional function, f, is a 1D functions consisting of a set of parallel line integrals as illustrated in figure 3.1b. This can be described by the following equation:

$$p_\phi(r) = \int_{-\infty}^{\infty} f(x,y)\, ds = \int_{-\infty}^{\infty} f(r\cdot\cos\phi - s\cdot\sin\phi,\, r\cdot\sin\phi + s\cdot\cos\phi)\, ds \tag{3.2}$$

where ϕ is the projection angle.

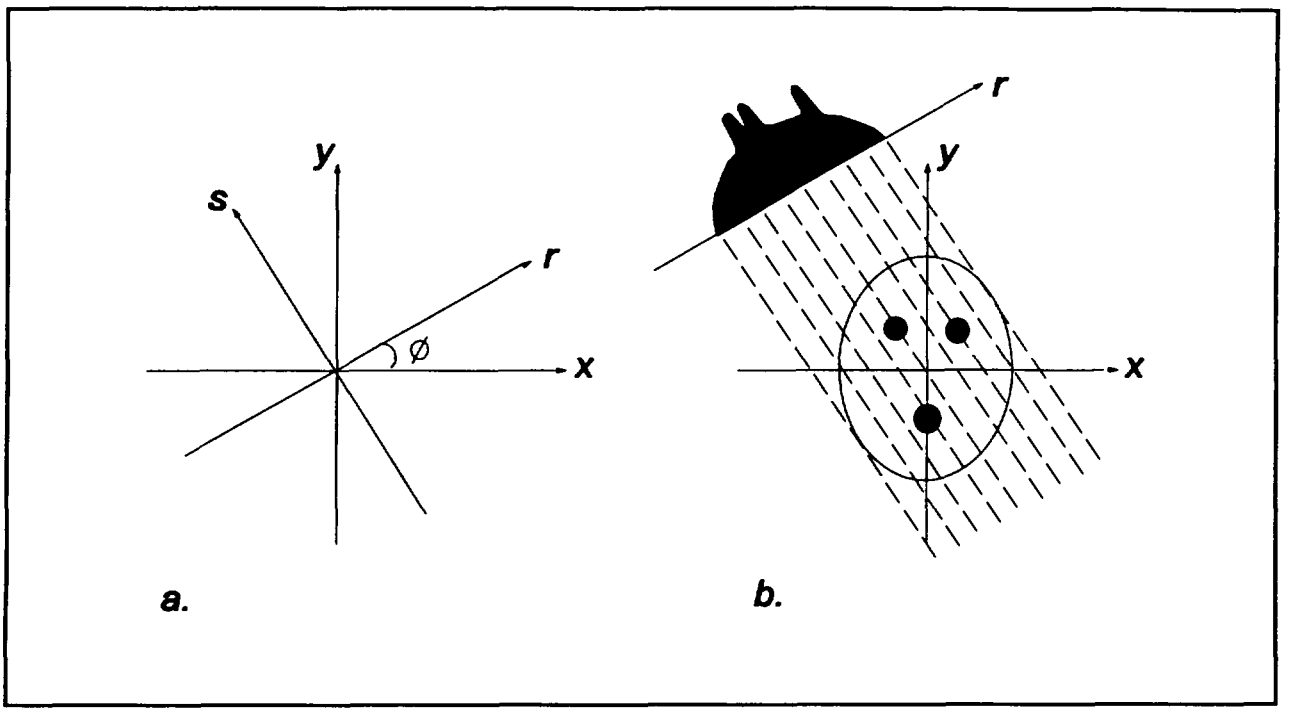

Figure 3.1: a) Coordinate systems, b) Parallel projection.

The collection of parallel projections for all possible angles, ϕ, is call the Radon transform of the function f, also known as the X-ray transform. The analytical approach to image reconstruction is to obtain an inversion formula for the continuous Radon transform, which then is discretized, since in practice data are only measured for a finite number of samples in a finite number of angles. Radon defined the transform and also formulated an inversion formula already in 1917 [Radon, 1917 and 1986], long before computed tomography became practically feasible.

3.2 Backprojection

The Radon transform is obtained by forward projection of a 2D function. The reverse operation to forward projection is called backprojection, and is defined in equation (3.3).

$$g(x,y) = \int_0^{\pi} p_{\phi}(r)\, d\phi = \int_0^{\pi} p_{\phi}(x\cdot\cos\phi + y\cdot\sin\phi)\, d\phi \qquad (3.3)$$

Backprojection is not the inverse of the Radon transform. The result obtained, g, is just a blurred version of the original function, f.

3.3 The Fourier transform

The Fourier transform (FT) is a way to express a function as an infinite sum of sine- and cosine-terms with different frequencies. Through the Fourier transform, a function in the spatial domain becomes a function in the frequency domain. The opposite is true for the inverse Fourier transform. There is an extensive literature on the FT and its applications, see e.g. Brigham (1974).

The FT has a fundamental roll in the derivation of analytical reconstruction methods, also known as transform methods. This is due to a its relations to the Radon transform, through the projection theorem. Another important properties of the FT is its relations to the convolution operator, through the convolution theorem.

The 1D FT of a function, f, and it's inverse are defined as follows:

$$F(u) = \mathcal{F}\{f(x)\} = \int_{-\infty}^{\infty} f(x)\, e^{-i2\pi xu}\, dx \tag{3.4a}$$

$$f(x) = \mathcal{F}^{-1}\{F(u)\} = \int_{-\infty}^{\infty} F(u)\, e^{i2\pi xu}\, du \tag{3.4b}$$

In two or more dimensions, the definitions are analogous using vector representation:

$$F(\mathbf{u}) = \mathcal{F}\{f(\mathbf{x})\} = \int f(\mathbf{x})\, e^{-i2\pi \mathbf{x}\cdot\mathbf{u}}\, d\mathbf{x} \tag{3.5}$$

$$F(u,v) = \mathcal{F}_2\{f(x,y)\} = \int_{-\infty}^{\infty}\int f(x,y)\, e^{-i2\pi(xu+yv)}\, dxdy =$$

$$= \int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} f(x,y)\, e^{-i2\pi xu}\, dx \right) e^{-i2\pi yv}\, dy$$

This means that the 2D FT can be calculated by 1D FT in the x-direction for each value of y, followed by 1D FT in the y-direction for each value of u. The inverse 2D FT is given by:

$$f(x,y) = \mathcal{F}_2^{-1}\{F(u,v)\} = \int_{-\infty}^{\infty}\int F(u,v)\, e^{i2\pi(xu+yv)}\, dudv \tag{3.6}$$

3.4 The convolution theorem

Convolution between two functions, f and h, yield a new function, g, defined as follows:

$$g(x) = (f \circledast h)(x) = \int_{-\infty}^{\infty} f(s)\, h(x-s)\, ds \tag{3.7}$$

where $\circledast$ denotes the convolution operator. Taking the FT of both sides gives:

$$G(u) = \mathcal{F}\{g(x)\} = \int_{-\infty}^{\infty} g(x)\, e^{-i2\pi xu}\, dx = \tag{3.8}$$

$$= \int_{-\infty}^{\infty}\int f(s)\, h(x-s)\, e^{-i2\pi xu}\, ds dx$$

Changing the order of integration we obtain:

$$G(u) = \int_{-\infty}^{\infty} f(s) \int_{-\infty}^{\infty} h(x-s)\, e^{-i2\pi xu}\, dx ds \tag{3.9}$$

Substituting t for (x-s) in the inner integral gives:

$$G(u) = \int_{-\infty}^{\infty} f(s) \int_{-\infty}^{\infty} h(t)\, e^{-i2\pi (s+t)u}\, dt ds = \tag{3.10}$$

$$= \int_{-\infty}^{\infty} f(s) \left(\int_{-\infty}^{\infty} h(t)\, e^{-i2\pi tu}\, dt \right) e^{-i2\pi su}\, ds =$$

$$= \int_{-\infty}^{\infty} f(s)\, H(u)\, e^{-i2\pi su}\, ds =$$

$$= H(u) \int_{-\infty}^{\infty} f(s)\, e^{-i2\pi su}\, ds = H(u)\, F(u)$$

This result is the convolution theorem, which in words states that convolution in the spatial domain is equivalent to multiplication in the frequency domain. This means that the relatively complex operation of convolution of two functions can be performed by taking the FT of both functions, multiplying in frequency space, and taking the inverse FT of the result. The frequency convolution theorem states that convolution in the frequency domain is equivalent to multiplication in the spatial domain, and can be shown in a similar manner.

3.5 Filtering

Convolution is often used in signal processing for applications such as smoothing, edge-enhancement or restoration. The data to be processed is convolved with a suitable function called kernel. When performed by multiplication in the frequency domain, this operation is called filtering, and the FT of the kernel is called filter. Filters for different applications can be designed directly in the frequency domain. Low pass filters are used for smoothing and high pass filters for edge enhancement. One example of a often used low pass filter is the Hamming filter, described in the following equation:

$$H(u) = A + (1-A)\cos(\pi u/u_c) \qquad (3.11)$$

where A is a constant, usually set to 0.5, and u_c is the cut-off frequency, above which the filter value is set to zero.

3.6 Deconvolution

An imaging system can be characterized by its point response function (PRF), also called point spread function, which is the response to a point source or δ-function input. The FT of the PRF is called the modulation transfer function (MTF). If the PRF is independent of position, the response, g, to an arbitrary input, f, can be described by the equation:

$$g = f \otimes h + n \qquad (3.12)$$

where h is the PRF, and n is a noise term. In the frequency domain, this equation becomes:

$$G = FH + N \qquad (3.13)$$

where upper case letters denotes the FT of the functions denoted by the corresponding lower case letters. If we want to restore the original input function, a first attempt could be to multiply G by the inverse filter 1/H. This is called deconvolution and is risky because the MTF usually tapers off and approaches zero at high frequencies, which means that little information is recovered at these frequencies, but the noise will get highly amplified. An alternative is to use a combination of the inverse filter and a low pass filter.

One example of this is the Wiener filter. A simplified Wiener filter is described by the following equation.

$$W(u) = \frac{H(u)}{H^2(u) + \gamma} \tag{3.14}$$

where γ is a constant determining the trade off between inverse and low pass filtering.

3.7 The 2D projection theorem

The most important property of the Radon transform is its relation to the Fourier transform through the projection theorem, also called the central section theorem or the Fourier slice theorem. This is the basis for all analytical reconstruction methods. In words the theorem can be expressed as follows: The 1D FT of the parallel projection, $p_\phi(r)$, at an angle ϕ of a function f(x,y) is equal to data along a line through the 2D FT of f(x,y), F(u,v), passing through the origin with an angle ϕ to the u-axis. The theorem is illustrated in figure 3.2 and can be derived as follows:

$$\begin{aligned} \mathcal{F}\{p_\phi(r)\} &= \int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} f(x,y)\,ds\right) e^{-i2\pi r\rho}\,dr = \\ &= \int_{-\infty}^{\infty}\int f(r\cdot\cos\phi - s\cdot\sin\phi,\, r\cdot\sin\phi + s\cdot\cos\phi)\,ds\, e^{-i2\pi r\rho}\,dr = \\ &= \int_{-\infty}^{\infty}\int f(r\cdot\cos\phi - s\cdot\sin\phi,\, r\cdot\sin\phi + s\cdot\cos\phi)\, e^{-i2\pi r\rho}\,dr ds \end{aligned} \tag{3.15}$$

Substituting x and y for r and s we obtain:

$$\begin{aligned} \mathcal{F}\{p_\phi(r)\} &= \int_{-\infty}^{\infty}\int f(x,y)\, e^{-i2\pi(x\cdot\cos\phi + y\cdot\sin\phi)\rho}\,dxdy = \\ &= \int_{-\infty}^{\infty}\int f(x,y)\, e^{-i2\pi(x\rho\cdot\cos\phi + y\rho\cdot\sin\phi)}\,dx\bar{dy} = \\ &= F(u,v)\,|_{u=\rho\cdot\cos\phi,\; v=\rho\cdot\sin\phi} \end{aligned} \tag{3.16}$$

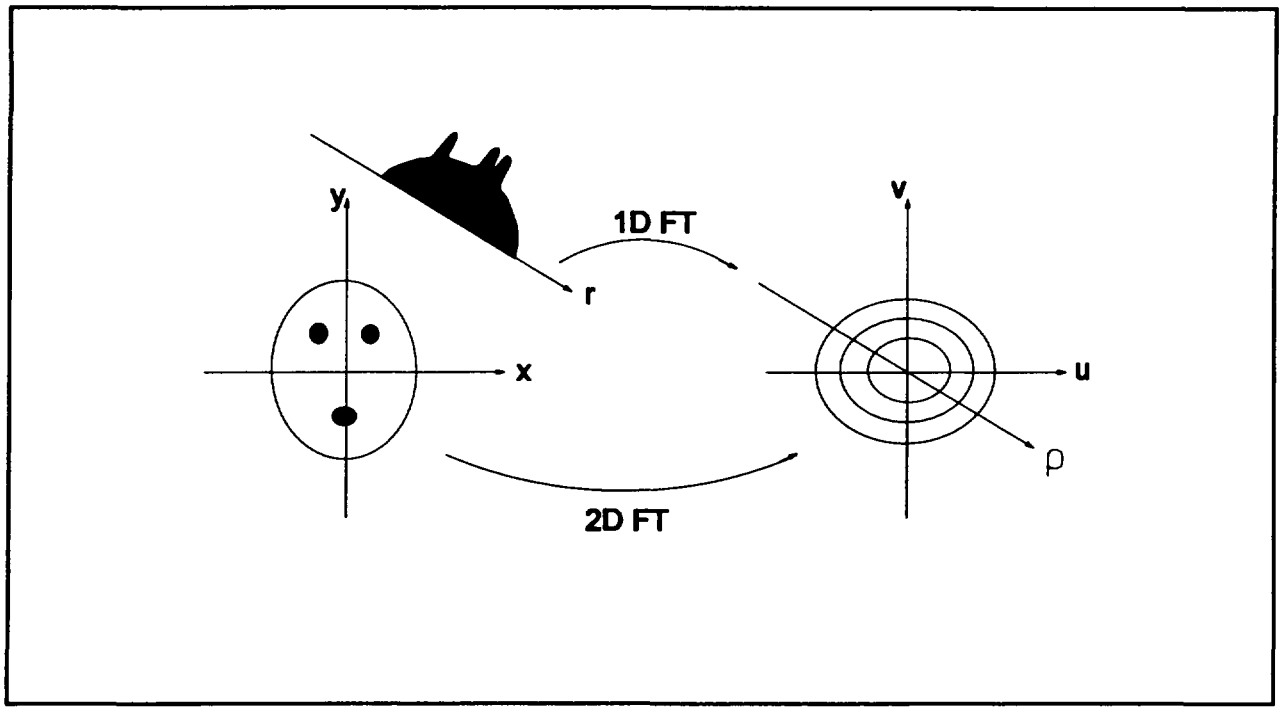

Figure 3.2: The projection theorem.

The projection theorem shows that it is theoretically possible to reconstruct an image from its projections. Taking the FT of a large number of parallel projection over the interval $[0,\pi)$ a polar map of the function is obtained in frequency space (see fig. 3.3). A 2D inverse FT will then give the original function. The most common reconstruction algorithm, based on this principle, is described in the next section.

3.8 Filtered Backprojection

The derivation of the filtered backprojection algorithm (FBP) starts from the definition of the inverse 2D Fourier transform:

$$f(x,y) = \int_{-\infty}^{\infty}\int F(u,v)\, e^{i2\pi(xu+yv)}\, du dv \qquad (3.17)$$

Changing to polar coordinates results in:

$$f(x,y) = \int_{0}^{2\pi}\int_{0}^{\infty} F_p(\rho,\phi)\, e^{i2\pi(x\rho\cdot\cos\phi+y\rho\cdot\sin\phi)}\, \rho\, d\rho d\phi \qquad (3.18)$$

where F_p is the polar representation of the function F. Since data is collected for angles in the interval $[0,\pi)$, the integration limits are changed. Also, the

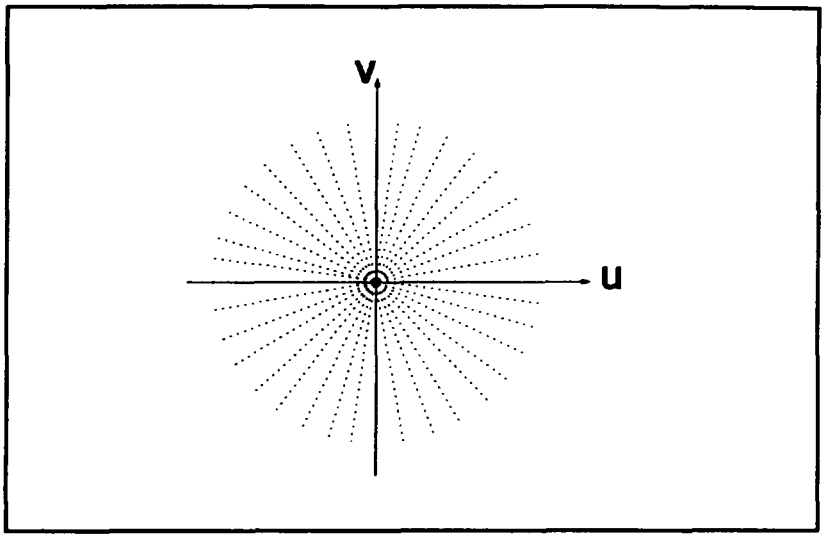

Figure 3.3: Mapping of frequency space.

variable r from equation (3.2) is introduced:

$$f(x,y) = \int_0^{\pi} \int_{-\infty}^{\infty} F_p(\rho,\phi)\,|\rho|\,e^{i2\pi r\rho}\,d\rho d\phi \tag{3.19}$$

Utilizing the projection theorem, the following equation is obtain:

$$f(x,y) = \int_0^{\pi} \left(\int_{-\infty}^{\infty} P_\phi(\rho)\,|\rho|\,e^{i2\pi r\rho}\,d\rho \right) d\phi \tag{3.20}$$

where $P_\phi(\rho)$ is the FT of the projection $p_\phi(r)$.

This equation is the basis for the filtered backprojection algorithm: The parallel projections are first individually filtered with the "ramp filter", $|\rho|$, and then backprojected, to form the reconstructed image. The high frequency part of the data is often dominated by noise, and the ramp filter is therefore usually combined with a low pass filter. The filtering step can also be performed by convolution in the spatial domain. This algorithm is called "convolution and backprojection" (CBP).

3.9 Discrete implementation

In a practical situation we cannot work with continuous data. A continuous function has to be sampled and truncated so that we get a finite number of data values. The 1D discrete Fourier transform (DFT) and its inverse are defined

in equation (3.21a-b) for a function with N samples. The definitions for two or more dimensions are analogous.

$$F(n) = \sum_{k=0}^{N-1} f(k)\, e^{-i2\pi kn/N}, \quad n = 0, \ldots N-1 \tag{3.21a}$$

$$f(k) = \frac{1}{N} \sum_{n=0}^{N-1} F(n)\, e^{i2\pi kn/N}, \quad k = 0, \ldots N-1 \tag{3.21b}$$

If the N data samples, f(k), represent one period of a periodic discrete function, it can be shown (Brigham, 1974) that the continuous FT of this function gives as result a periodic discrete function where one period is represented by the N values, F(n), given by equation (3.21a).

The convolution theorem is valid also for discrete functions, which are treated as periodic functions. This means that interference can be obtained between different periods when performing a convolution between two discrete functions by multiplication in the frequency domain. This can be avoided by zero padding, which means that a function with N values is extended with N zeros, and the FT is calculated for 2N values. An interference free convolution is obtained by zero padding the two functions and in the end discarding the N last values of the result. Another effect of zero padding is a finer sampling in frequency space.

Equation (3.21) is never used directly for calculating the DFT. It is much more efficient to use the fast Fourier transform (FFT) algorithm, resulting in a computation time reduced by a factor $2N/\log_2 N$, where N is the number of data points (Brigham, 1974).

FBP and CBP are mathematically equivalent methods. It has been shown (Lewitt, 1979) that the convolution step in CBP can be speeded up significantly with an insignificant amount of error by using reduced precision multiplications or by approximating the convolution kernel with a piecewise constant function.

3.10 The sampling theorem

The sampling interval is an important parameter for discrete functions. For band-limited functions, the optimal sampling interval can be determined from the Nyqvist sampling theorem (see e.g. Brigham, 1974). A function is band-

limited if it has no components for frequencies larger than a certain value, u_c. The sampling theorem states that a continuous band-limited function is uniquely determined by its sampled values if a sampling interval, Δx, is used such that $\Delta x \leq (1/2u_c)$. All discrete functions have a periodic FT, and if $\Delta x > 1/(2u_c)$, there will be interference between different periods in the frequency domain. This effect is called aliasing, and can be reduced by using a low-pass filter. The frequency $1/(2\Delta x)$ is the highest frequency component in a discrete function, and is known as the Nyqvist frequency.

The activity distributions studied in PET are always limited in space, which means that they cannot be band-limited. However, all PET scanners have a finite resolution, which means that the data is filtered with a low-pass filter before sampling. Huang et al. (1980) concluded that the sampling interval used should be $\leq$ W/3, where W is the full width half maximum (FWHM) of the system PRF.

When the radial sampling interval has been set, a suitable angular sampling interval should be chosen. If the number of radial samples is N_r, equal radial and tangential sampling intervals are obtained at the highest frequencies by using a number of angular samples, $N_\phi = (\pi/2)N_r$.

3.11 Alternative reconstruction methods

An alternative to FBP reconstruction is to start with the backprojection operation, and then filter the backprojected image with a 2D ramp filter. We can call this method "backprojection and filtering" (BPF), also known as the "ρ-filtered layergram" method. Zero padding must be used in the filtering step. The 2D ramp filter can of course be combined with a low pass filter.

Another alternative analytical reconstruction method, which constitutes the most obvious utilisation of the projection theorem, is the direct Fourier inversion (DFI). This method consists of taking the FT of all projections, resulting in a polar mapping of the image in frequency space, interpolating to a Cartesian grid representation, and finally obtaining the reconstructed image by 2D inverse FT. The one step that causes difficulties in a practical implementation of this method is the interpolation in frequency space. This is due to the fact that the samples are closely spaced near the origin but more sparsely at higher frequencies. An improved interpolation can be obtained, at

the cost of a longer computation time, if the projections are zero-padded before the FT to obtain a finer frequency sampling. Another way to improve the interpolation with negligible extra computation time is to shift the frequency samples radially by half a sampling interval for every other projection angle (Lewitt, 1983). This can be accomplished by introducing a phase-shift in the data before the FT.

The interpolation step can be reduced the from a 2D problem to a 1D problem if the frequency samples are collected along concentric squares in stead of concentric circles, which can be done by appropriately selecting different sampling intervals for different projection angles. This is the linogram method (see e.g. Jacobson, 1996).

4. ITERATIVE RECONSTRUCTION

While analytical reconstruction methods are based on an inversion formula for continuous functions, algebraic methods begin with a discrete model of the problem, and attempts to solve a system of linear equations. Most algebraic reconstruction methods are iterative. These methods have the advantage that a priori knowledge about the image can be incorporated into the model (e.g. non-negative activity concentration). Also the effects of attenuation, scatter and randoms can be incorporated. They are basically slower than analytical methods, but can be speeded up with different techniques. Algebraic methods are not limited by theoretical requirements of a complete set of projections.

General reviews on algebraic reconstruction techniques can be found in e.g. Colsher (1977), Herman (1980) and Kak & Slaney (1988). The theory that will be presented here is valid for both 2D and 3D iterative reconstruction.

4.1 The discrete model

An discrete image can be represented by a set of picture-elements (pixels) as shown in figure 4.1, each one containing a uniform activity distribution. In 3D, the term volume element (voxel) is used. The measured projection values can be regarded as linear combinations of the pixel values, called ray-sums. This leads to a set of linear equations, which can be expressed as follows, with N pixel values, f_j, and M projection values, p_i:

$$\sum_{j=1}^{N} f_j w_{ij} = p_i \,, \quad i = 1, \ldots M \tag{4.1}$$

where w_{ij} represents the contribution from the j:th pixel to the i:th ray-sum. In the past, the solution of this equation system by ordinary matrix inversion techniques was not possible due to large sizes of M and N, and the presence of noise in the measured data. Instead different iterative methods were developed, which are basically trial-and-error procedures, starting from an initial image, that is updated in each iteration. These methods can be based on either additive or multiplicative techniques, with either row-action or simultaneous updating schemes. Today the matrix inversion approach could possibly be an alternative. Although the calculation of the inverse takes a very long time, this only has to be done once.

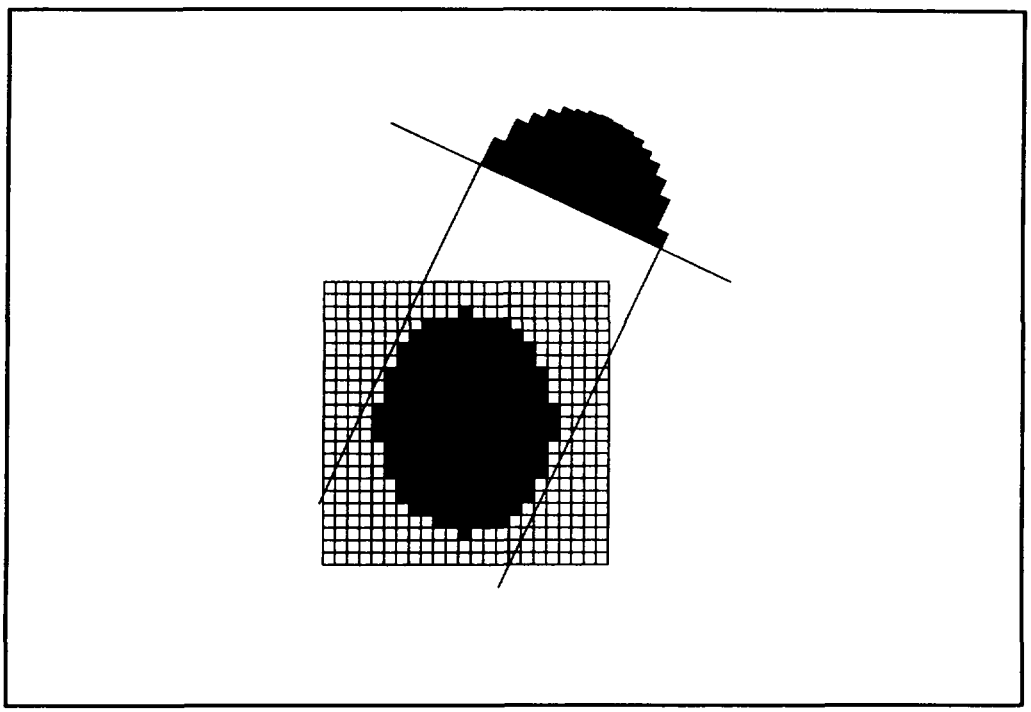

Figure 4.1: The discrete image and projection model.

4.2 Algebraic reconstruction technique

Algebraic reconstruction technique (ART) is an iterative reconstruction method that was developed by Gordon et al. (1970), based on a method for solving linear equation systems proposed by Kaczmarz (1937). The solution can be regarded as a point in an N-dimensional space. Each equation in (4.1) would then represent different hyperplanes (one for each value of i). If a unique solution exits, it will be the intersection point of all the hyperplanes. Kaczmarz's iterative process for finding this point starts with an initial guess, f^0, that is projected on the first hyperplane, giving a new point, f^1, which is then projected on the second hyperplane, et cetera. This process converges to the unique solution, and can be described as follows:

$$q_i^k = \sum_{j=1}^{N} f_j^k w_{ij} \tag{4.2}$$

$$f_j^{k+1} = f_j^k + \frac{(p_i - q_i^k)\, w_{ij}}{|w_i|^2} \quad , \quad j = 1, \ldots N \tag{4.3}$$

where k is the iteration number, i=mod(k,M) is the remainder after division of k with M, and

$$|w_i|^2 = \sum_{j=1}^{N} w_{ij}^2 \qquad (4.4)$$

In practice the ART method works as follows: The initial guess is an image with a uniform distribution (e.g. zero in all pixels). The ray-sum corresponding to one projection line is calculated according to equation (4.2). A correction term is obtained as the difference between the measured projection value and the calculated ray-sum. The correction term is backprojected, and each pixel along the projection line is corrected according to equation (4.3). This procedure is then repeated for the next projection line. A smoother image can be obtained at the cost of a slower convergence by multiplying the correction terms by a relaxation factor <1. An increased convergence speed can be achieved by processing the different projection angles, not in a cyclic order, but with a larger difference between consecutive angles.

Equation (4.3) is the additive ART algorithm. An alternative is to use multiplicative correction factors in stead of additive correction terms. Multiplicative ART can be described as follows:

$$f_j^{k+1} = f_j^k \frac{p_i}{q_i^k} w_{ij} , \quad j=1,\ldots N \qquad (4.5)$$

4.3 Simultaneous iterative reconstruction

ART is a row action method, which means that the image is updated once for each projection line. A different approach is to consider all the projections simultaneously in each iteration. Landweber (1951) proposed a simultaneous additive iterative algorithm that can be described as follows:

$$f_j^{k+1} = f_j^k + \sum_{i=1}^{M} (p_i - q_i^k) w_{ij} , \quad j = 1,\ldots N \qquad (4.6)$$

This means that in each iteration, each pixel value is corrected by a weighted mean of the correction terms for all projection angles. This gives a slower convergence, but often a better result. The initial estimate of the image can be zero. A modified Landweber method was implemented for image reconstruction (Pan & Yagle, 1991). In paper V, a 3D reconstruction method is described, based on a simultaneous additive iterative algorithm.

4.4 Maximum likelihood - Expectation maximization

The most popular iterative reconstruction method used in PET is the expectation maximization algorithm for maximum likelihood estimation (ML-EM) proposed by Shepp and Vardi (1982). This method was developed, based on a Poison model for the emission of photons. It attempts to find the solution which has the highest probability of generating the observed projection data; the maximum likelihood estimate. If w_{ij} is the probability for an emission in pixel j to be detected in coincidence line i, the expected projection value for a given image, f, is given by:

$$q_i = \sum_{j=1}^{N} f_j w_{ij} , \quad i=1,...M \tag{4.7}$$

The likelihood, L(f), that the image f has generated the projection values p is given by:

$$L(f) = \prod_{i=1}^{M} e^{-q_i} \frac{q_i^{p_i}}{p_i!} \tag{4.8}$$

To find the maximum value, we calculate the derivatives of the logarithm of L(f) with respect to each image element f_j:

$$\frac{\partial}{\partial f_j} \ln L(f) = - \sum_{i=1}^{M} w_{ij} + \sum_{i=1}^{M} \frac{p_i}{q_i} w_{ij} , \quad j=1,...N \tag{4.9}$$

The maximum likelihood solution is obtained when this expression is zero for all values of j. Shepp and Vardi (1982) showed that an EM algorithm, in which the expectation step is given by equation (4.7) and the maximization step by equation (4.10), converges to the maximum likelihood solution.

$$f_j^{k+1} = f_j^k \frac{1}{\sum_{i=1}^{M} w_{ij}} \sum_{i=1}^{M} \frac{p_i}{q_i^k} w_{ij} , \quad j=1,...N \tag{4.10}$$

ML-EM is a simultaneous multiplicative iterative algorithm. The iteration is normally initiated with a uniform distribution $f^0>0$. From equation (4.10) follows two important properties of this algorithm, namely that $f_j^{k+1}>0$ (non-negativity) and that $\sum_{j=1}^{N} f_j^{k+1} = \sum_{i=1}^{M} p_i$ (constant sum of counts). The popularity of the ML-EM algorithm is due to the high signal to noise ratio (SNR) in the reconstructed images. This reconstruction method was used for the scanner presented in paper I.

4.5 Accelerated EM

The slow convergence rate of the ML-EM algorithm is a major disadvantage. A number of methods have been proposed for accelerating the convergence. One way is to raise the correction factors to a power >1. This would require a renormalization of the image after each iteration. An alternative method was proposed by Lewitt and Muehllehner (1986). It consists of rewriting equation (4.10) in additive form, and applying an overrelaxation factor $\lambda>1$ to the correction terms:

$$f_j^{k+1} = f_j^k + \lambda^k f_j^k \left(\frac{1}{\sum_{i=1}^{M} w_{ij}} \sum_{i=1}^{M} \frac{p_i}{q_i^k} w_{ij} - 1 \right) \quad , \; j = 1, \ldots N \quad (4.11)$$

where lambda is adjusted at each iteration, so that the non-negativity is preserved. This method also preserves the constant sum of counts property.

A more sophisticated accelerated EM algorithm was suggested by Tanaka (1987), which includes high pass filtering of the correction factors in order to obtain a more uniform frequency response. A modification is also included that gives a more uniform correction over the image plane.

Another method, which has received a great deal of attention lately, is the ordered subsets EM algorithm (OS-EM) (Hudson & Larkin, 1994). With this method, only a subset of the measured projections are used in each iteration. The convergence can be accelerated by an order of magnitude (a factor equal to the number of subsets). However, it has not been shown that this algorithm converges to the maximum likelihood solution. In practice, OS-EM is somewhere between ML-EM and multiplicative ART.

4.6 A stopping rule

The main advantage of the ML-EM algorithm is that it produces images with a high SNR as compared to FBP, especially in regions of_low activity. However, it has been noticed that after a certain number of iterations, the image quality starts to deteriorate rather than improve. This can be explained as follows: The algorithm tries to find image values, f_j, so that each expected projection value q_i, is close to the corresponding measured projection value, p_i. In the first part of iteration process, the image quality improves, as the

expectation values get closer to the measured values. However, with a Poison distribution, it's statistically unlikely that all measured projection values would be close to their respective expectation values. Therefore the image quality starts to deteriorate after a certain number of iterations, when to many of the q_i-value are close to the corresponding p_i-values. Veklerov and Llacer (1987) proposed using a χ^2-test after each iteration to test the hypothesis that the M p_i-values are jointly statistically valid realizations of Poison-distributed random variables with mean values q_i. This method can be used to determine when to stop the iteration process.

4.7 Basis functions

In iterative reconstruction methods, the image is modelled as a linear combination of a number of shifted copies of a basis function (pixels or voxels). An important factor for the behaviour of these methods is how well the model chosen describes the real object. Of special importance is how the factors w_{ij} are selected. These factors represents the contribution from each pixel to each ray-sum, and can in principle be precalculated and stored in a file before the reconstruction. This would be a very large file, but that is less of a problem today than it has been in the past. An alternative is to calculate the values as they are needed during the reconstruction. The most simple way to do this is to set w_{ij}=1 or 0 depending on the distance between the centre of pixel j and projection line i. The results obtained with this method are usually quite poor. Better results are obtained with values that decreases linearly with increasing distance. In a practice, these two methods can be regarded as nearest neighbour and linear interpolation, respectively. Both methods assume that the image is composed of circularly symmetric pixels or spherically symmetric voxels. Superior results can be obtained with more sophisticated basis functions, such as the modified Kaiser-Bessel functions suggested by Lewitt (1992b).

A way to avoid the errors due to finite pixels was proposed by Buonocore et al. (1981) by introducing the "natural pixels"-model. Each natural pixel consists of the area in the image corresponding to one projection strip. A discrete image is then represented by M natural pixel values, q_j. This in fact means that the image is constructed by backprojection of the q_j-values. The relation between the M natural pixel values, q_j, and the M measured projection values, p_i, is given by equation (4.12).

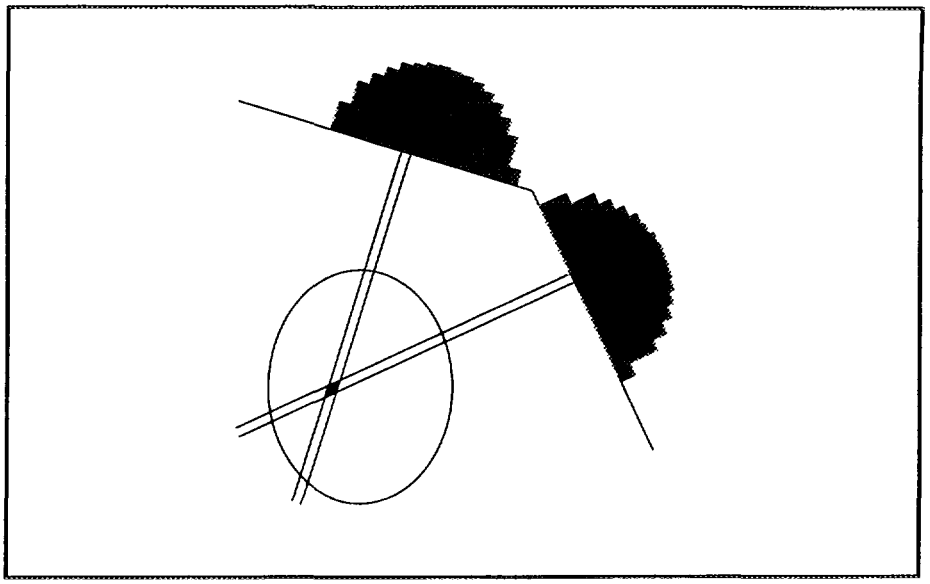

Figure 4.2: The natural pixels model.

$$\sum_{j=1}^{M} H_{ij} q_j = p_i \ , \quad i = 1, \ldots M \tag{4.12}$$

where H_{ij} is the area of intersection between natural pixels i and j, as shown in figure (4.2). Equation (4.12) can be solved by an iterative process (Buonocore et al., 1981) or by singular value decomposition (SVD) (Gullberg & Zeng, 1994), which is a method that can be used for matrix inversion (see e.g. Press et al., 1986).

In paper V, an iterative 3D reconstruction method based on the new concept of "mobile pixels" is presented. This method has a certain similarity to the natural pixel approach, in that it is intended to reduce the finite pixel errors. Each plane in a 3D volume is represented by M overlapping pixels, one for every projection line, placed on a very fine grid. The location of the pixels vary from plane to pane following the projection lines, thereby the term "mobile".

5. 3D RECONSTRUCTION

A number of 3D reconstruction methods have been developed for PET over the past two decades. Most of them are extensions to 3D of the 2D methods described above for the parallel beam geometry. A review of reconstruction methods for the cone-beam geometry has been made by Jacobsson (1996). A description of the general principles of these 3D methods will be given here, without too many mathematical details.

5.1 Data formats

The acquisition data from PET scanners can be stored in different formats. The simplest one is the list format, which means that, for each detected coincidence, the coordinates of the detection points of the two annihilation photons are saved in a long list. After the acquisition, the data can be transformed to some other format before image reconstruction. The advantage is that no extra computation is needed during the data acquisition. What could be a disadvantage is the size of the data files, which will be dependent on the number of detected events. This method has mostly been used for PET scanners based on position sensitive planar detectors. A coincidence event can be represented by a 2D planar coordinate for each detector and an angular coordinate for the rotation angle of the scanner, (r_1,t_1,r_2,t_2,ϕ).

For 2D ring-scanners, the acquisition data is saved in sinogram format. A sinogram is a 2D matrix with one dimension for the projection angle (ϕ) and one for the radial position of the coincidence line (r). There is one sinogram for each plane. When a coincidence event is detected, the matrix-entry corresponding to the angular and radial position of the coincidence line is incremented by 1. Figure 5.1 shows an example of a sinogram corresponding to two point sources. A 2D sinogram set can be represented by the 3D array $p(r,\phi,z)$, where z the axial position of the plane.

In 3D acquisition with multi-ring scanners, the sinogram format has also been used, with one sinogram for each ring combination. E.g. a 16 ring scanner produces 256 sinograms in 3D mode. For intra-ring coincidences we get normal 2D sinograms. However, the sinograms for coincidences between different rings will be inconsistent in the sense that the data in different projection angles does not represent the same plane through the activity

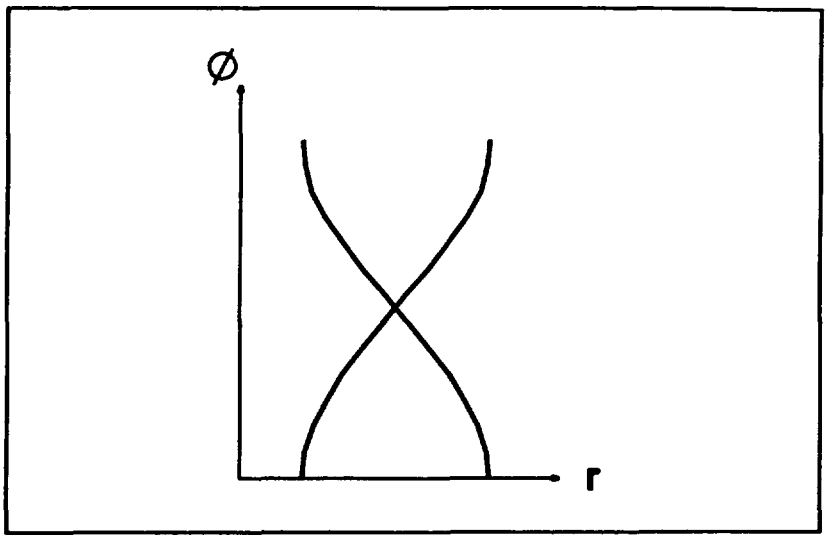

Figure 5.1: Sinogram for two point sources.

distribution. A 3D sinogram set is represented by the 4D array $p(r,\phi,z_1,z_2)$, where z_1 and z_2 are the axial positions of the two detector rings involved.

The latest generation of multi-ring scanners use 2D projection format. The data is sorted into 2D parallel projections for different azimuthal and polar angles. A 2D projection data set is represented by the 4D array $p(r,t,\phi,\theta)$, where r and t are Cartesian coordinates on the projection plane, ϕ is the azimuthal angle, and θ is the polar angle.

It may be advantageous to convert projection data from one format to another depending on the reconstruction method that will be used. This is called rebinning. E.g. 3D sinogram data may be converted to 2D projection format. List mode data can be converted to sinograms or 2D projections before reconstruction. An alternative is to perform a 3D event-by-event backprojection into a 3D image matrix. For each detected event, all voxels along the coincidence line are incremented by an amount proportional to the line length through the voxel. This method can be used if the number of collected events is not sufficient to get adequate statistical accuracy in all projection bins. A fast algorithm for 3D backprojection has been presented by Siddon (1985).

5.2 The 3D projection theorem

The projection theorem in 3D is analogous to the 2D case. It states that the 2D FT of a 2D projection of a 3D function is equal to a slice through the origin

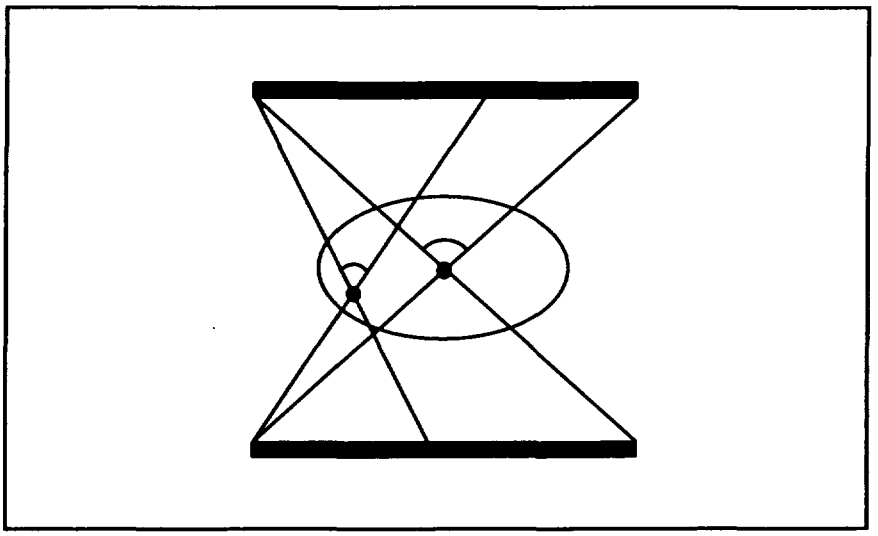

Figure 5.2: Acceptance angles for different points in the FOV.

of the 3D FT of the function. A strict proof can be developed in a way similar to the 2D case, but a more intuitive one is obtained as follows: A 3D function can be regarded as a set of parallel planes with any orientation. The 2D projection theorem is valid for each plane. 1D FT in the third direction gives the relationship stated above.

5.3 Limited angle tomography, 3D BPF

3D reconstruction was first used in PET for scanners based on two stationary planar detectors (see figure 2.1a). With this geometry, data is only acquired in a restricted angular interval around the object. This kind of measurement is therefore called limited angle tomography. A 3D reconstruction method based on Fourier techniques was developed by Chu and Tam (1977).

In the following we will use a 3D Cartesian coordinate system, with the 3D coordinate $\mathbf{r}=(x,y,z)$. The y-axis is perpendicular to the detector planes. The angle between a coincidence line and the y-axis is denoted θ. For different positions in the field-of-view (FOV), the angular range within which coincidences can be detected ("acceptance angle") varies, as illustrated in figure 5.2.

The first step in the reconstruction is a 3D event-by-event backprojection. If f is the true distribution, the image, g, obtained after the backprojection can be described by equation (5.1).

$$g(\mathbf{x}) = \int f(\mathbf{x}')\, h(\mathbf{x},\mathbf{x}')\, d\mathbf{x}' \qquad (5.1)$$

where h(**r**,**r**') is the point response function (PRF), giving the relative contribution from a point source at position **r**' to the backprojected image at position **r**. The response function varies with the position in the field-of-view (FOV), due to the varying acceptance angle. Equation (5.1) is a Fredholm equation of the first kind and has no general solution. However, if we know the extent of the object, we can define a restricted acceptance angle, ψ, equal to the smallest acceptance angle at some point within the object. For each point in the object, there is cone-shaped region in which coincidences with $\theta \leq \psi$ can be detected. This region is called the universal cone. If we backproject only those events which fall within the universal cone, the response function will be position-independent, and equation (5.1) becomes a convolution equation:

$$g(\mathbf{x}) = \int f(\mathbf{x}')\, h(\mathbf{x}-\mathbf{x}')\, d\mathbf{x}' \qquad (5.2)$$

Taking the FT of both sides, equation (5.2) becomes (due to the convolution theorem):

$$G(\rho) = F(\rho)\, H(\rho) \qquad (5.3)$$

where G, F and H are 3D FTs of g, f and h, respectively, and ρ is a 3D frequency coordinate, ρ=(u,v,w). An estimate, $\hat{F}$, of F can then be obtained as:

$$\hat{F}(\rho) = R(\rho)\, G(\rho) \qquad (5.4)$$

where $R(\rho)$ is a restoration filter, which in principle could be the inverse filter $H^{-1}(\rho)$, but in practice it is necessary to combine it with a low pass filter, due to the noise domination at high frequencies. Chu and Tam (1977) proposed the following filter (compare with equation (4.14)):

$$R(\rho) = \frac{H(\rho)}{H^2(\rho) + \gamma (2\pi)^4 |\rho|^4} \qquad (5.5)$$

where γ is a Lagrange multiplier which depends on the noise level. The optimum value of γ can be determined by an iterative process. Schorr and Townsend (1981) derived analytical expressions for H^{-1}. The practical implementation of 3D BPF was studied by Webb et al. (1984).

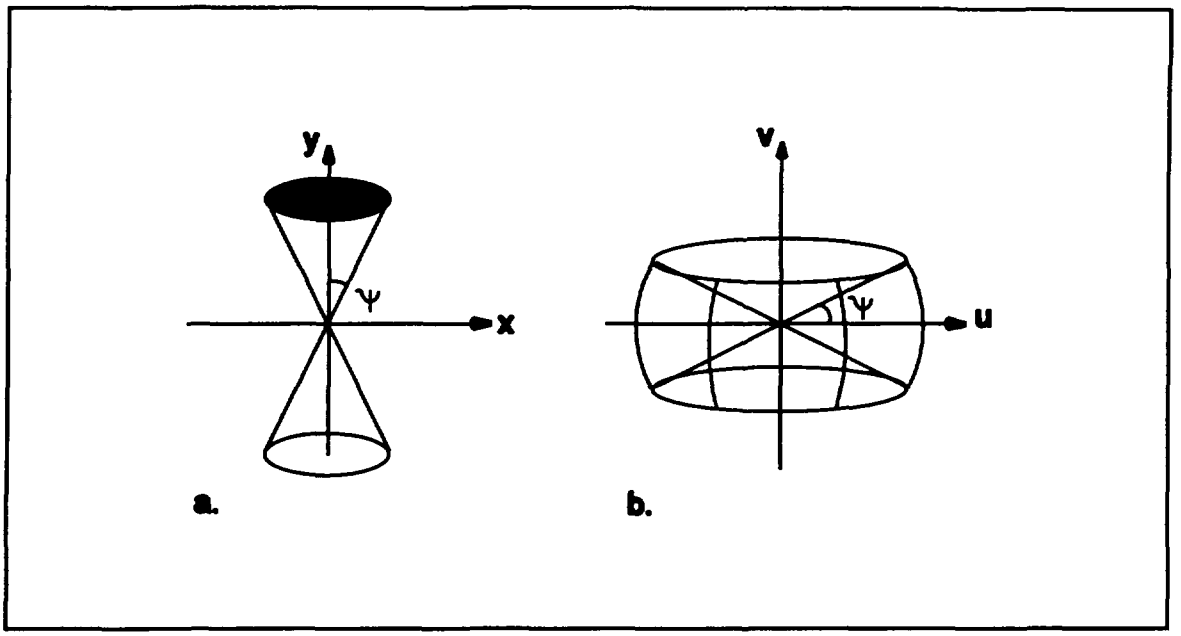

Figure 5.3: Point response function in a) the spatial domain, b) the frequency domain.

5.4 Missing data

The use of a restricted acceptance angle, ψ, in the backprojection is equivalent to using only data in parallel projections in directions $\theta \leq \psi$. According to the projection theorem, this means that frequency space is mapped only in the region outside a cone with an angle $\pi/2-\psi$ to the v-axis, as illustrated in figure 5.3. A consequence of this is that $H(\rho)=0$ inside the "missing cone", and the values of $F(\rho)$ in this region cannot be recovered using equation (5.4).

Tam et al. (1979) proposed a method for recovering the missing frequency components, utilizing a priori information about the extent of the object. The method was based on the fact that a function that is limited in the spatial domain cannot be limited in the frequency domain (compare with band-limited 1D functions). While iterating back and forth between the two domains, the function was reset to the known values in different regions. In practice, however, the missing data can never be fully recovered. To quote Harrison Barrett (San Francisco, October 1995): "Limited angle tomography is... limited!"

5.5 Iterative limited angle reconstruction

A major disadvantage of the reconstruction method described above is that a large part of the data has to be rejected due to the restricted acceptance angle,

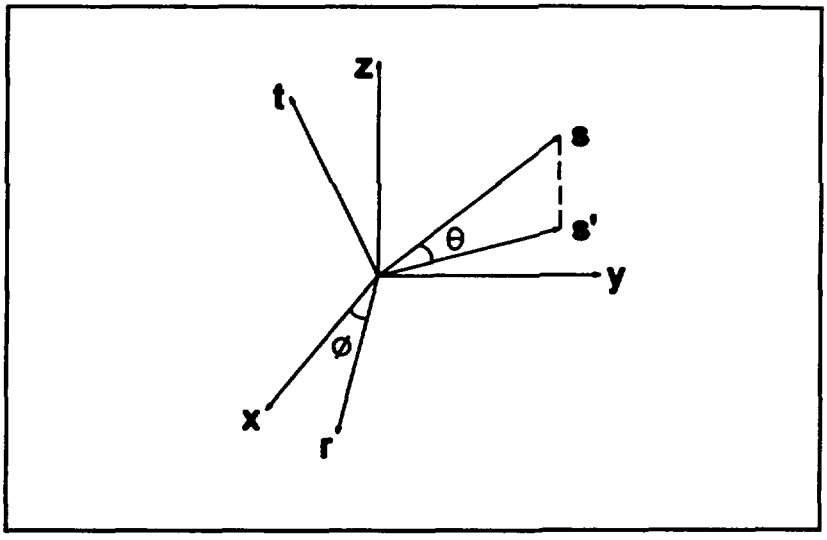

Figure 5.4: 3D coordinate systems.

and that the FOV must be limited; the larger the FOV, the smaller the acceptance angle, and the larger the amount of data that has to be rejected.

In paper V we present a small PET scanner for mammographic studies, which is based on the limited angle tomography principle. In order to use all the available data and to have a FOV as large as possible, we developed a reconstruction method based on a simultaneous additive iterative algorithm. The method utilizes a new concept called "mobile pixels", which reduces the rate of increase of the noise level during the iteration process due to finite pixel errors.

5.6 Rotating scanners, 3D BPF

The obvious way to obtain the missing data with a dual-headed PET scanner is to rotate the scanner. In the following, the coordinate systems shown in figure 5.4 will be used. The (x,y,z)-system is a stationary Cartesian coordinate system, with the z-axis on the rotation axis of the scanner. The (r,s,t)-system, where the s-axis indicates the direction of projection, is obtained by rotating the (x,y,z)-system an angle ϕ about the z-axis, giving the intermediate (r,s',z)-system, which is then rotated an angle θ about the r-axis. The relationship between the two system is given by:

$$\begin{cases} r = x\cdot\cos\phi + y\cdot\sin\phi \\ s = (-x\cdot\sin\phi + y\cdot\cos\phi)\cos\theta + z\cdot\sin\theta \\ t = (x\cdot\sin\phi - y\cdot\cos\phi)\sin\theta + z\cdot\cos\theta \end{cases} \tag{5.6}$$

If the scanner is rotated so that the azimuth angle ϕ spans over 180°, the entire frequency space will be mapped. All the necessary data for a correct tomographic reconstruction will certainly be obtained, since a subset of the data (θ=0) consists of complete 2D data sets for different transaxial planes. Still using the restricted acceptance angle ψ in the backprojection, the PRF, h, will now have it's support in a region outside the cone with an angle $\pi/2-\psi$ to the z-axis. Colsher (1980) derived an analytical expression for the 3D FT, H, of h. The reconstructed image can then be obtained by filtering the backprojected image with the inverse filter H^{-1}.

An alternative to reconstruction by 3D backprojection and filtering (BPF) is 3D filtered backprojection (FBP), which can be used if the projection data is saved in 2D projection format. Each 2D projection is filtered with a 2D filter, and the image is obtained by 3D backprojection. According to the projection theorem, the 2D filters to be used are slices through the 3D filter, H^{-1}. Colsher (1980) also presented analytical expressions for these 2D filters, which depends of the polar angle, θ, but are independent of the azimuth angle, ϕ.

Correct images can be obtained with the Colsher-filter, but the restricted acceptance angle is still used, and a large fraction of the collected data rejected. Clack et al. (1984) showed that, for rotating scanners, the angular constraint can be removed in the direction of the rotation, ϕ, and consequently also the limit for the transaxial FOV. This can be explained as follows: Even if the coincidence line for a detected event falls outside the universal cone at the actual scanner position, it may well be inside the cone for another scanner position. After removing the ϕ-constraint, the support of the response function, h, will be the same. However, the sensitivity varies for different coincidence lines depending on the time during which it can be seen by the rotating detectors. This geometric sensitivity, gs, decreases linearly with the perpendicular distance, |r|, from the coincidence line to the axis of rotation, z. Clack et al. (1984) proposed 2 methods for taking this sensitivity variation into account in 3D BPF reconstruction. The first and most obvious method is to apply a weight inversely proportional to gs(r) to each event during backprojection. The second method suggested, which includes some approximation, is to apply a position dependent scale factor to the reconstructed image after deconvolution. The geometric sensitivity, gs, is the same function as the one that was used in paper I to correct the projections from a rotating scanner before reconstruction.

The angular constraint has now be removed in ϕ but not in θ, so still some data has to be rejected and the axial FOV limited. Defrise et al. (1988) suggested a method to increase the FOV and the amount of data utilized by dividing the image into different subregions. A large region, D_1, covering the entire object is first reconstructed by 3D BPF, using a small acceptance angle, ψ_1. A low statistics estimate of the image is then obtained. This image is convolved with the response function corresponding to an acceptance angle $\psi_2 > \psi_1$. A new backprojection is made with acceptance angle ψ_2 into a region, D_2, smaller than D_1, and the old data in D_2 is replaced. The whole image can then be deconvolved with inverse filter corresponding to the acceptance angle ψ_2, after appropriate normalisation. This procedure can then be repeated with an even smaller region, in order to incorporate more data. The implementation of this method was studied by Suckling et al. (1992).

5.7 Multi-ring scanners, 3D FBP

Different reconstruction approaches have been suggested for multi-ring scanners operated in 3D mode. The one most used is the "reprojection algorithm", which is a 3D filtered backprojection (FBP) algorithm that utilizes all detected events. The use of 3D FBP was suggested by Pelc and Chesler (1979). The acquired data is sorted in 2D parallel projections. The reconstruction is done by 2D filtering and 3D backprojection. The only problem is that the measured oblique 2D projections ($|\theta|>0$) are truncated and do not cover the entire FOV. Just as in the case of 3D BPF, 3D FBP cannot be applied directly without restricting the θ-acceptance angle and the axial field of view. This problem can be solved with the reprojection algorithm (Kinahan & Rogers, 1989). First a 2D reconstruction is made of each transaxial plane using only a 2D subset of the data ($\theta=0$), giving a low statistics estimate of the image. Forward projection (reprojection) of this image is then used for calculation of the missing data in the truncated projections. Finally 3D FBP can be applied, including all the data. This method was implemented for multi-ring scanners by Townsend et al. (1989 and 1991), Defrise et al. (1990) and Cherry et al. (1992a).

Contrary to the case of 2D FBP, there is no unique filter function for 3D FBP, giving mathematically correct reconstruction. This is due to the overdetermined character of the problem, as explained by Defrise et al. (1995a). One method for obtaining the filter function is to calculate the FT of

the point response function for backprojection (Colsher, 1980).

Cherry et al. (1992b) developed a method for using 3D FBP in connection with whole body scanning without the need of 2D reconstruction and 3D reprojection to obtain the missing data. They used a "quasi-continuous" axial scanning, which consisted of moving the patient couch, and acquiring data at a large number of closely spaced axial positions. This way, complete 2D projections are obtained for a large part of the axial FOV, which can be reconstructed directly by 3D FBP. The missing data is thus obtained from subsequent axial positions. This approach is analogous to the method described above to eliminate the azimuthal angular constraint for rotating dual headed PET systems.

A completely different approach to 3D FBP was used by Ra et al. (1982). They developed a method based on the observation that 2D FBP can be used to reconstruct any oblique plane through the object with a normal that has an angle to the z-axis $\leq$ the maximum oblique acceptance angle of the scanner. The image values at each point can then be obtained as a mean value of all the reconstructed oblique planes passing through that point. A direct implementation of this method would be very inefficient, due to the large number of planes involved. But the basic principle was used for development of a 2D filter for 3D FBP.

5.8 1D/3D reconstruction methods

Feldkamp et al. (1984) developed a simple but approximate 3D FBP method for cone-beam tomography. Each cone-beam projection is treated as a number of tilted fan-beams, which are independently filtered with a transaxial 1D filter (ramp-filter). The image is then obtained by 3D backprojection of each filtered tilted fan-beam. The reconstruction is exact in the central transaxial plane, but becomes more and more approximate with increasing distance from this plane. Although this method was developed for cone-beam tomography, the basic principle of 1D/3D filtering/backprojection can also be applied to the parallel beam projection geometry used in PET. This results in a fast reconstruction, but the filtering approximation leads to some cross-talk between the different transaxial planes.

Tanaka et al. (1992) developed a more sophisticated reconstruction method

based on the 1D/3D-principle, which eliminates the axial cross-talk. The oblique projections ($|\theta|>0$) are filtered with a high-pass filter, so that they only contribute to the high frequency components of the image. The low frequency components are reconstructed using the transaxial projection data ($\theta=0$), and the final frequency response is normalized.

Another method based on the 1D/3D principle was developed by Defrise et al. (1992). A 1D filter is still used for the oblique projections, but for the transaxial projections a 2D filter is used. The modulation transfer function (MTF=FT{PRF}) for the backprojected filtered oblique projections was analysed, and a 2D filter for the transaxial projections was developed, which corrects for the approximation errors introduced. For each transaxial plane to be reconstructed, different 2D filters must be used to filter the 2D transaxial projections.

5.9 Alternative 3D reconstruction methods

A few other methods for 3D image reconstruction in PET will be mentioned here briefly. Daube-Witherspoon and Muehllehner (1986) developed an iterative 3D reconstruction method based on the ML-EM algorithm. The method starts with a 3D backprojection of the measured data, in order to reduce the number of data values. At each iteration, the comparison between measured and calculated values is done in image space. Kinahan et al. (1988) developed a method for 3D CBP as an alternative to 3D FBP, using 2D convolution instead of 2D filtering before the 3D backprojection. Stearns et al. (1990) developed a 3D direct Fourier inversion method. The missing data is obtained from the frequency domain. This is a fast method since 3D backprojection is not needed. Stazyk and Rogers (1992) developed a method for inversion of the 3D Radon transform, which consists of plane-integrals rather than line-integrals, using a fast backprojection technique. Wu et al. (1995) developed a fast reconstruction method based on direct Fourier inversion of the 3D Radon transform.

5.10 Rebinning algorithms

The methods described above can all be used to reconstruct data from 3D PET scanners, but there are certain disadvantages. 3D FBP is based on 2D parallel

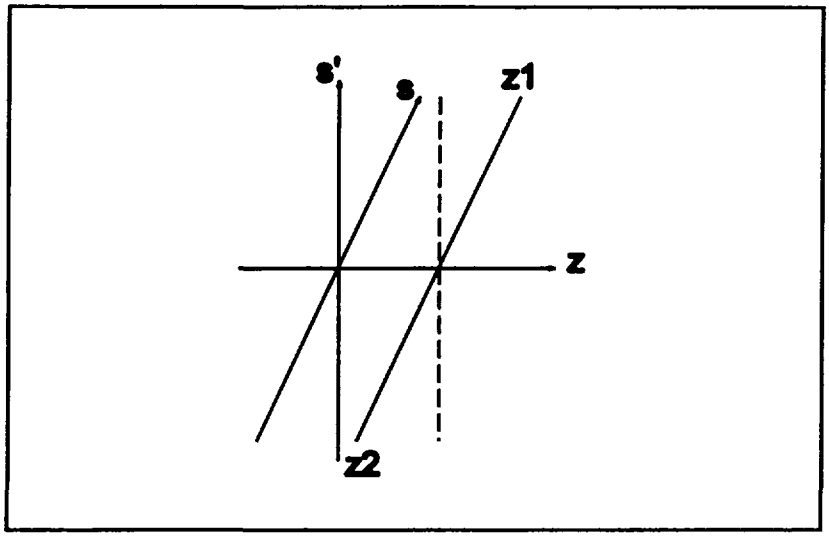

Figure 5.5: Single slice rebinning.

projections in different azimuthal and polar angles. These 4D projection data sets require large storage space on the computer, and a high number of detected events, in order to get sufficient statistical accuracy in all the projection bins. With 3D BPF, the amount of data is reduced by backprojection into a 3D image matrix. However, the corrections for randoms, scatter and attenuation are more difficult to perform in object space than in projection space, which limits the quantitative accuracy of these methods.

An alternative approach to the reconstruction problem in 3D PET is to rebin the 4D projection data to a 3D format with one 2D sinogram for each transaxial plane. This way the corrections can still be made in projection space with reduced data sets and increased statistical accuracy. Basically, the 3D reconstruction problem is divided into a 1D axial rebinning problem and 2D transaxial reconstruction. Several methods based on rebinning algorithms have been developed.

The most simple of these methods is the single slice rebinning (SSRB) algorithm proposed by Daube-Witherspoon and Muehllehner (1987). Each oblique coincidence line is included in the sinogram corresponding to the transaxial plane midway between the two detection points, as shown in figure 5.5. 2D reconstruction is then used for each plane. This method has several good qualities: Its easy to implement, the rebinning can be performed in real time during the data acquisition, and the reconstruction is fast. The approximations used are $\theta=0$ and $z=(z_1+z_2)/2$. For annihilations occurring centrally between the two detectors ($s=0$), the axial position (z) chosen will be correct but, with increasing distance from the centre, the mispositioning errors

increase. This method is basically 2D reconstruction of 3D data. Due to its simplicity it has been found to be useful in some situations (Sossi et al., 1994).

In paper II, a 3D reconstruction method based on SSRB is presented, which is also used in papers III and IV. Before the final 2D reconstruction, an axial deconvolution is made to correct for the mispositioning errors of SSRB. To calculate the axial spread functions (ASF) to be deconvolved, we have to know the activity distribution. This can be obtained by 2D reconstruction of a 2D data subset, giving a low-statistics estimate.

The axial deconvolution is done for each radial and angular positions in the sinograms. Frequency space deconvolution cannot be used because the ASFs are not position independent. Instead we have used iterative methods. In paper II we used a multiplicative simultaneous method, but in papers III and IV we opted for an additive method. We have also tested using SVD (Erlandsson and Strand, 1996), which can give similar results but is somewhat more time consuming. The axial deconvolution can be done for all data or alternatively only for the data with oblique angles ($|\theta|>0$), which is later added to the transaxial data (Erlandsson et al., 1994). This approach has the advantage of small data sets, and real time rebinning. The reconstruction algorithm can be summarized as follows: Preliminary 2D reconstruction, 3D forward projection, 1D axial deconvolution and final 2D reconstruction.

Multi-slice rebinning (MSRB) is a method developed by Lewitt et al. (1992a, 1994). With this rebinning technique, the sinograms for all transaxial planes between the two axial detection points are incremented by an amount inversely proportional to the number of planes. This operation can be called "axial backprojection". It results in more axial blurring than SSRB, but in this case the response function is fairly independent of the source distribution, and correction can be made by axial filtering without the need of a preliminary reconstruction.

The axial filtering can be made either before or after 2D reconstruction of transaxial planes. The point response function depends on the axial position, so frequency space filtering cannot be used. Lewitt et al. first used an iterative method (1992a), but then opted for SVD (1994). Since the response functions are known in advance, the decomposition step of SVD is only performed once. The noise can be controlled by combining SVD with a "low-pass filter", with an amplitude that decreases for decreasing singular values, or a Wiener filter

(Shao, 1995). One disadvantage with MSRB is that a special processor is needed for real time rebinning. This reconstruction method can be described as 1D axial filtered backprojection and 2D transaxial reconstruction.

Defrise (1995b) developed a rebinning method called Fourier rebinning (FRB), based on the frequency distance principle (FDP) introduced by Edholm et al. (1986). FDP is an approximate method which can be used to separate data in a 2D sinogram depending on the position, s, along the projection line. The 2D FT of the sinogram, $p(r,\phi)$, gives a function $P(\rho,k)$. FDP states that, for each projection angle ϕ, an image point mainly contributes to a Fourier coefficient, $P(\rho,k)$, such that $s= -k/\rho$. This means that data along a straight line passing through the origin of the P-graph all correspond to the same s-value.

The FRB method is based on acquisition of 3D sinograms, $p(r,\phi,z_1,z_2)$, with one 2D sinogram for each detector ring combination (z_1,z_2). After 2D FT of these sinograms, the Fourier coefficients, $P(\rho,k,z_1,z_2)$, are sorted into a matrix, $P(\rho,k,z)$, using FDP to obtain the axial position, z, according to:

$$z = \frac{z_1+z_2}{2} - \frac{k}{\rho}\tan\theta \qquad (5.7)$$

where θ is the oblique angle of the 3D sinogram. After normalization, inverse 2D FT will give a set of 2D sinograms, $p(r,\phi,z)$, and 2D reconstruction can be used.

A disadvantage with FRB is that a full 4D data set is required. Also, FDP is not completely reliable. E.g. if the activity distribution has rotational symmetry, the sinograms will be constant in the ϕ-direction. Since FT of a constant is a delta-function, the 2D FT of the sinogram will be zero for all points (ρ,k) except those with $k=0$. In this case FRB is reduced to SSRB. However, this situation is not very likely to occur in practice.

6. Conclusions and future prospects

PET is an expensive modality, only available at a few centres world-wide and mainly used as a research tool. The use of PET in clinical practice has been increasing during the past few years (Wagner, 1991), and will continue to do so in the future, partly due to developments of low-cost PET scanners. In this thesis, the developments of two such systems were presented. In order to fully utilize these systems, 3D data acquisition was used and 3D reconstruction methods were developed.

A rotating scanner was built with two scintillation camera heads (paper I), and quantitative images were obtained after correction for electronic dead-time, physical decay of the radionuclide, position non-linearities, intrinsic and geometric efficiency variations, random coincidences, attenuation and scatter. This scanner has shown to be clinically useful for tumour studies (Brun et al., 1996), and similar systems will be used in the future for clinical PET (Nellemann et al., 1996, Miyaoka et al., 1996, Visvikis et al., 1995).

There will also be simple PET scanners dedicated for particular examination, e.g. mammography (Thompson et al., 1995, Freifelder & Karp, 1996). In paper V, a positron emission mammography scanner with two opposed sets of detectors, based on the limited angle tomography principle, was presented, including corrections for sensitivity variations, randoms and attenuation.

A number of different methods for 3D image reconstruction for PET have been developed over the years, both for large area detector scanners and for multi-ring scanners without inter-plane septa. The method that has become most popular is the reprojection algorithm for 3D FBP (Kinahan & Rogers, 1989). Defrise et al. (1994) compared different 3D methods. They found that the reprojection algorithm gave the highest SNR, and also that no dramatic improvement in computation speed was obtained with direct Fourier reconstruction (Stearns et al., 1990) or with the FAVOR algorithm (Defrise et al., 1992). One way to increase the reconstruction speed could be to use methods based on plane-integrals and inversion of the 3D Radon transform (Stazyk and Rogers, 1992, Wu et al., 1995). Another way is to use rebinning methods. SSRB (Daube-Witherspoon and Muehllehner, 1987) has the advantage of reducing the data sets from 4D to 3D, and a direct 2D reconstruction is fast but results in a degraded spatial resolution and possible artifacts. It has, however, shown to be useful in particular situations (Sossi et

al., 1994). The adverse effects of SSRB can be removed using the SSAD method (papers II-III), which includes preliminary 2D reconstruction, forward projection and axial deconvolution before the final 2D reconstruction. This is a method which also can be applied to data from a 2D scanner (paper IV). MSRB (Lewitt et al., 1992a, 1994) is a similar 3D reconstruction method that has the advantage of more uniform axial spread functions and does not require preliminary reconstruction and reprojection before axial filtering. However, it is more complicated to performed this type of rebinning in real time during the data acquisition. The Fourier rebinning method (Defrise, 1995) has some good qualities, but requires a complete 4D data set.

Iterative reconstruction algorithms have several advantages over analytic methods. Physical effects, such as attenuation, scatter, randoms and detector efficiency, can be incorporated into the model, as well as *a priori* knowledge about the activity distribution. Furthermore, there are no theoretical requirements of a complete projection data set or an invariant PRF, and the images obtained often have a higher SNR, especially in areas of low activity concentration. The major disadvantage is usually longer computation times. With the developments on the computer front, iterative reconstruction becomes more feasible, even in 3D. Johnson et al. (1995) implemented the ML-EM algorithm for 3D PET. Kinahan et al. (1995) compared the 3D reprojection FBP algorithm with iterative 3D ART and ML-EM reconstruction. They concluded that ART can produce superior results with a similar computation time as the reprojection algorithm. They also noted that, although each iteration requires a considerably longer computation time in 3D than in 2D, far fewer iterations are needed. In paper V, iterative reconstruction was used for a limited angle scanner. In this case the computation time was not a problem due to the small size of the data sets. An iterative 3D reconstruction algorithm could be developed for data sets obtained by SSRB, using the axial spread functions from SSAD. In the future, the role of algebraic reconstruction methods will be increasing, although analytical methods will continue to be used for a long time.

7. ABBREVIATIONS

nD	= n-Dimensional (n=1,2,3,4)
ART	= Algebraic Reconstruction Technique
BPF	= BackProjection and Filtering
CBP	= Convolution and BackProjection
FBP	= Filtered BackProjection
FOV	= Field Of View
FT	= Fourier Transform
MC	= Monte Carlo
ML-EM	= Maximum Likelihood - Expectation Maximization
MSRB	= Multi-Slice ReBinning
MTF	= Modulation Transfer Function
PET	= Positron Emission Tomography
PRF	= Point Response Function
SNR	= Signal to Noise Ratio
SSAD	= Single Slice rebinning with Axial Deconvolution
SSRB	= Single Slice ReBinning
SVD	= Singular Value Decomposition

8. ACKNOWLEDGMENTS

I want to thank...

☺ my supervisor Sven-Erik Strand, who has been my cicerone through all these years.
☺ my colleagues in the PET group Anders Sandell and Tomas Ohlsson. Without their work, mine would not have been possible.
☺ Michael Ljungberg for helping me with innumerable computer problems.
☺ Freddy Ståhlberg, for introducing me to the fascinating field of tomographic image reconstruction.
☺ John Palmer for reviewing this thesis.
☺ all those at the Radiation Physics Department and the University Hospital in Lund who have helped me in one way or another.

I also want to acknowledge the generous financial support from...

○ The Mrs Bertha Kamprad Foundation.
○ The Gunnar, Arvid and Elisabeth Nilsson Foundation.
○ The Royal Physiographic Society in Lund.
○ The Swedish Cancer Foundation.
○ The Lund University Mathematical and Natural Science Faculty.
○ The Lund University Medical Faculty.
○ The Lund University Hospital Funds.
○ The Swedish Society for Radiation Physics.
○ The Swedish Society for Medical Radiology.

9. REFERENCES

Anger HO, "Radioisotope cameras", In Hine GJ (editor), *Instrumentation in nuclear medicine,* Academic press, New York, 485-552, 1967.

Bailey DL, Meikle SR, "A convolution-subtraction scatter correction method for 3D PET", Phys Med Biol, 39:411-424, 1994.

Bateman JE, Connoly JF, Stephenson R, Flesher AC, "The development of the Rutherford laboratory MWPC positron camera", Nucl Instr Meth, 176:83-88, 1980.

Bergström M, Litton J, Eriksson L, Bohm C, Blomqvist G, "Determination of object contour from projections for attenuation correction in cranial positron emission tomography", J Comp Assist Tomogr, 6:365-372, 1982.

Bergström M, Eriksson L, Bohm C, Blomqvist G, Litton J, "Corrections for scattered radiation in a ring detector positron camera by integral transformation of the projection", J Comp Assist Tomogr, 7:42-50, 1983.

Bohm C, Eriksson L, Bergström M, Litton J, Sundman R, Singh M, "A computer assisted ringdetector positron camera system for reconstruction tomography of the brain", IEEE Trans Nucl Sci, 25:624-637, 1978.

Brigham EO, "The Fast Fourier transform", Prentice-Hall Inc, Englewood Cliffs, NJ, 1974.

Brun E, Ohlsson T, Erlandsson K, Kjellén E, Sandell A, Tennvall J, Wennerberg J, Strand S-E, "Early prediction of treatment in head and neck cancer with 2-^{18}FDG PET", Submitted to J Nucl Med, 1996.

Buonocore MH, Brody WR, Macovski A, "A natural pixel decomposition for two-dimensional image reconstruction", IEEE Trans Biomed Eng, 28:69-78, 1981.

Carson RE, Daube-Witherspoon ME, Green MV, "A method for postinjection PET transmission measurements with a rotating source", J Nucl Med, 29:1558-1567, 1988.

Casey ME, Hoffman EJ, "Quantitation in positron emission computed tomography: A technique to reduce noise in accidental coincidence measurements and coincidence efficiency calibration", J Comput Assist Tomogr, 10:845-850, 1986.

Cherry SR, Marsden PK, Ott RJ, Flower MA, Webb S, Babich JW, "Image quantification with a large area multiwire proportional chamber positron camera (MUP-

PET)", Eur J Nucl Med, 15:694-700, 1989.

Cherry SR, Dahlbom M, Hoffman EJ, "Evaluation of a 3D reconstruction algorithm for multi-slice PET scanners", Phys Med Biol, 37:779-790, 1992a.

Cherry SR, Dahlbom M, Hoffman EJ, "High sensitivity, total body PET scanning using 3D data aquisition and reconstruction", IEEE Trans Nucl Sci, 39:1088-1092, 1992b.

Cherry SR, Meikle SR, Hoffman EJ, "Correction and caracterization of scattered events in three-dimensional PET using scanners with retractable septa", J Nucl Med, 34:671-678, 1993.

Chu G, Tam K-C, "Three-dimensional imaging in the positron camera using Fourier techniques", Phys Med Biol, 22:245-265, 1977.

Clack R, Townsend D, Jeavons A, "Increased sensitivity and field of view for a rotating positron camera", Phys Med Biol, 29:1421-1431, 1984.

Colsher JG, "Iterative three-dimensional reconstruction from tomographic projections", Comput Graph Im Process, 6:513-537, 1977.

Colsher JG, "Fully three-dimensional positron emission tomography", Phys Med Biol, 25:103-115, 1980.

Dahlbom M, Eriksson L, Rosenqvist G, Bohm C, "A study of the possibility of using multi-slice PET systems for 3D imaging", IEEE Trans Nucl Sci, 36:1066-1071, 1989.

Daube-Witherspoon ME, Muehllehner G, "An iterative image space reconstruction algorithm suitable for volume ECT", IEEE Trans Med Im, 5:61-66, 1986.

Daube-Witherspoon ME, Muehllehner G, "Treatment of axial data in three-dimensional PET", J Nucl Med, 28:1717-1724, 1987.

Defrise M, Kuijk S, Deconinck F, "A new three-dimensional reconstruction method for positron cameras using plane detectors", Phys Med Biol, 33:43-51, 1988.

Defrise M, Townsend DW, Geissbuhler A, "Implementation of three-dimensional image resonstruction for multi-ring positron tomographs", Phys Med Biol, 35:1361-1372, 1990.

Defrise M, Townsend DW, Clack R, "FaVoR: A fast reconstruction algorithm for volume imaging in PET", Conference Record, 1991 IEEE Nuclear Science Symposium and Medical Imaging Conference, Santa Fe, 1919-1923, 1992.

Defrise M, Geissbuhler A, Townsend DW, "A performance study of 3D reconstruction algorithms for positron emission tomography", Phys Med Biol, 39:305-320, 1994.

Defrise M, Clack R, Townsend DW, "Image reconstruction from truncated, two-dimensional, parallel projections", Inv Probl, 11:287-313, 1995a.

Defrise M, "A factorization method for the 3D x-ray transform", Inv Probl, 11: 983-994, 1995b.

Edholm PR, Lewitt RM, Lindholm B, "Novel properties of the Fourier decomposition of the sinogram", SPIE Proc. Int. Wkshp. Physics Eng. Computerized Multidimensional Image Processing, 671:8-18, 1986.

Erlandsson K, Esser PD, Strand S-E, van Heertum RL, "3D reconstruction for a multi-ring PET scanner", Conference Record of the 1993 IEEE Nuclear Science Symposium and Medical Imaging Conference, San Francisco, 1562-1566, 1994.

Erlandsson K, Strand S-E, "Improved axial resolution in 2D PET with 3D reconstruction", Conference Record of the 1995 IEEE Nuclear Science Symposium and Medical Imaging Conference, San Francisco, 1267-1271, 1996.

Feldkamp LA, Davis LC, Kress JW, "Practical cone-beam algorithm", J Opt Soc Am, 1:612-619, 1984.

Flower MA, Ott RJ, Webb S, Leach MO, Marsden P, Khan O, McCready VR, Bateman JE, Flesher AC, Sharma HL, Smith AG, "A clinical eveluation of a prototype positron camera for longitudinal emission tomography", Brit J Radiol, 57:1103-1117, 1984.

Freifelder R, Karp JS, "A dedicated PET scanner for breast cancer", Conference Record of the 1995 IEEE Nuclear Science Symposium and Medical Imaging Conference, San Francisco, 1358-1362, 1996.

Gordon R, Bender R, Herman GT, "Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and x-ray photography", J Theor Biol, 29:471-481, 1970.

Gullberg GT, Zeng GL, "A reconstruction algorithm using singular value decomposition of a discrete representation of the exponential Radon transform using natural pixels", IEEE Trans Nucl Sci, 41:2812-2819, 1994.

Herman GT, "Image reconstruction from projections. The fundamentals of computerized tomography", Academic Press, New York, 1980.

Huang S-C, Hoffman EJ, Phelps ME, Kuhl DE, "Quantitation in positron emission computed tomography: 2. Effects of inaccurate attenuation correction", J Comput Assist Tomogr, 3:804-814, 1979.

Huang S-C, Hoffman EJ, Phelps ME, Kuhl DE, "Quantitation in positron emission computed tomography: 3. Effects of sampling", J Comput Assist Tomogr, 4:819-826, 1980.

Huang S-C, Carson RE, Phelps ME, Hoffman EJ, Schelbert HR, Kuhl DE, "A boundary method for attenuation correction in positron computed tomography", J Nucl Med, 22:627-637, 1981.

Hudson HM, Larkin RS, "Accelerated image reconstruction using ordered subsets of projection data", IEEE Trans Med Im, 13:601-609, 1994.

Jacobson C, "Fourier methods in 3D-reconstruction from cone-beam data", Linköping Studies in Science and Technology, Dissertations No. 427, Linköping University, Sweden, 1996.

Jeavons A, Hood K, Herlin G, Parkman C, Townsend D, Magnanini R, Frey P, Donath A, "The high-density avalanche chamber for positron emission tomography", IEEE Trans Nucl Sci, 30:640-645, 1983.

Johnson CA, Yan Y, Carson RE, Martino RL, Daube-Witherspoon ME, "A system for 3D reconstruction of retracted-septa PET data using the EM algorithm", IEEE Trans Nucl Sci, 42:1223-1227, 1995.

Kaczmarz S, "Angenaherte auflosung von systemen linearer gleichungen", Bull Acad Pol Sci Lett A, 6-8A:355-357, 1937.

Kak AC, Slaney M, "Principles of computerized tomographic imaging", IEEE Press, New York, 1988.

Kinahan PE, Rogers JG, Harrop R, Johnson RR, "Three-dimensional image reconstruction in object space", IEEE Trans Nucl Sci, 35:635-638, 1988.

Kinahan PE, Rogers JG, "Analytic 3D image reconstruction using all detected events", IEEE Trans Nucl Sci, 36:964-968, 1989.

Kinahan PE, Matej S, Karp JS, Herman GT, Lewitt RM, "A comparison of transform and iterative reconstruction techniques for a volume-imaging PET scanner with a large axial acceptence angle", IEEE Trans. Nucl. Sci., vol. 42, pp. 2281-2287, 1995.

Landweber L, "An iterative formula for Fredholm integral equations of the first kind", Amer J Math, 73:615-624, 1951.

Lercher MJ, Wienhard K, "Scatter correction in 3-D PET", IEEE Trans Med Im, 13:649-657, 1994.

Lewitt RM, "Ultra-fast convolution approximations for computerized tomography", IEEE Trans. Nucl. Sci., 26:2678-2681, 1979.

Lewitt RM, "Reconstruction algorithms: Transform methods", Proceedings of the IEEE, 71:390-408, 1983.

Lewitt RM, Muehllehner G, "Accelerated iterative reconstruction for positron emission tomography based on the EM algorithm for maximum likelihood estimation", IEEE Trans Med Im, 5:16-22, 1986.

Lewitt RM, Muehllehner G, Karp JS, "3D image reconstruction for PET by multislice rebinning and axial filtering", Conference Record of the 1991 IEEE Nuclear Science Symposium and Medical Imaging Conference, Santa Fe, 2054-2061, 1992a.

Lewitt RM, "Alternatives to voxels for representation in iterative reconstruction algorithms", Phys Med Biol, 37:705-716, 1992b.

Lewitt RM, Muehllehner G, Karp JS, "Three-dimensional image reconstruction for PET by multi-slice rebinning and axial image filtering", Phys Med Biol, 39:321-339, 1994.

Litton J, Bergström M, Eriksson L, Bohm C, Blomqvist G, Kesselberg M, "Performance study of the PC-384 positron camera system for emission tomography of the brain", J Comput Assist Tomogr, 8:74-87, 1984.

Miyaoka RS, Lewellen TK, Kim JS, Kaplan MS, Kohlmyer SK, Costa W, Jansen F, "Performance of a dual headed SPECT system modified for coincidence detection", Conference record of the 1995 IEEE Nuclear Science Symposium and Medical Imaging Conference, San Francisco, 1348-1352, 1996.

Muehllehner G, Buchin MP, Dudek JH, "Performance parameters of a positron imaging camera", IEEE Trans Nucl Sci, 23:528-537, 1976.

Muehllehner G, Atkins F, Harper PV, "Positron camera with longitudinal and transverse tomographic capabilities", in Medical Radionuclide Imaging, Vienna, IAEA-SM-210/84, 1:291-307, 1977.

Muehllehner G, Karp JS, "A positron camera using position-sensitive detectors: PENN-

PET", J Nucl Med, 27:90-98, 1986.

Muehllehner G, Karp JS, Mankoff DA, Beerbohm D, Ordonez CE, "Design and performance of a new positron tomograph", IEEE Trans Nucl Sci, 35:670-674, 1988.

Nellemann P, Hines H, Braymer W, Muehllehner G, Geagan M, "Performance caracteristics of a dual headed SPECT scanner with PET capability", Conference record of the 1995 IEEE Nuclear Science Symposium and Medical Imaging Conference, San Francisco, 1751-1755, 1996.

Ohlsson T, "A clinical positron emission tomography facility: Development and results", PhD Thesis, Lund University, Sweden, 1996.

Ott RJ, Flower MA, Babich JW, Marsden PK, "The physics of radioisotope imaging", in Webb S (Editor), *The physics of medical imaging*, IOP Publishing, Bristol, 1988.

Ott RJ, Veugen R, Flower MA, Visvikis D, Wells K, "3D scatter characteristics for large area PET cameras", Proceedings of the 1995 International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine, Aix les Bains, France, 121-125, 1995.

Pan T-S, Yagle AE, "Numerical study of multigrid implementations of some iterative image reconstruction algorithms", IEEE Trans Med Im, 10:572-588, 1991.

Pelc NJ, Chesler DA, "Utilization of cross-plane rays for three-dimensional reconstruction by filtered back-projection", J Comput Assist Tomogr, 3:385-395, 1979.

Phelps ME, Hoffman EJ, Mullani NA, Ter-Pogossian MM, "Application of annihilation coincidence detection to transaxial reconstruction tomography", J Nucl Med, 16:210-224, 1975.

Phelps ME, Hoffman EJ, Huang SC, Kuhl DE, "ECAT: A new computerized tomographic imaging system for positron-emitting radiopharmaceuticals", J Nucl Med, 19:635-647, 1978.

Press WH, Flannery BP, Teukolsky SA, Vetterling WT, "Numerical recipes. The art of scientific computing", Cambridge University Press, Cambridge, 1986.

Ra JB, Lim CB, Cho ZH, "A true three-dimensional reconstruction algorithm for the spherical positron emission tomograph", Phys Med Biol, 27:37-50, 1982.

Radon J, "Uber die bestimmung von funktionen durch ihre integralwerte längs gewisser mannigfaltigkeiten", Berichte Sächsische Akademie der Wissenschaften, Math-Phys Kl,

69:262-267, 1917.

Radon J, "On the determination of functions from their integral values along certain manifolds", (Translation by PC Parks from the original German text), IEEE Trans Med Im, 5:170-176, 1986.

Sandell A, Ohlsson T, Erlandsson K, Hellborg R, Strand S-E, "A PET system based on 2-^{18}FDG production with a low energy electrostatic proton accelerator and a dual headed PET camera", Acta Oncologica, 31:771-776, 1992.

Schorr B, Townsend D, "Filters for three-dimensional limited-angle tomography", Phys Med Biol, 26:305-312, 1981.

Shao L, Lewitt RM, Karp JS, Muehllehner G, "Combination of Wiener filtering and singular value decomposition for volume imaging PET", IEEE Trans Nucl Sci, 42:1228-1234, 1995.

Shepp LA, Vardi Y, "Maximum likelihood reconstruction for emission tomography", IEEE Trans Med Im, 1:113-121, 1982.

Siddon RL, "Fast calculation of the exact radiological path for a three-dimensional CT array", Med Phys, 12:252-255, 1985.

Sossi V, Stazyk MW, Kinahan PE, Ruth TJ, "The performance of the single-slice rebinning technique for imaging human striatum as evaluated by phantom studies", Phys Med Biol, 39:369-380, 1994.

Spinks TJ, Jones T, Bailey DL, Townsend DW, Grootoonk S, Bloomfield PM, Gilardi M-C, Casey ME, Sipe B, Reed J, "Physical performance of a positron tomograph for brain imaging with retractable septa", Phys Med Biol, 37:1637-1655, 1992.

Stazyk MW, Rogers JG, "Analytic image reconstruction in PVI using the 3D Radon transform", IEEE Trans Nucl Sci, 39:1153-1160, 1992.

Stearns CW, Chessler DA, Brownell GL, "Accelerated image reconstruction for a cylindrical positron tomograph using Fourier domain methods", IEEE Trans Nucl Sci, 37:773-777, 1990.

Suckling J, Ott RJ, Deehan BJ, "Quantitative analysis of a reconstruction method for fully three-dimensional PET", Phys Med Biol, 37:751-766, 1992.

Tam K-C, Perez-Mendez V, Macdonald B, "3-D object reconstruction in emission and transmission tomography with limited angular input", IEEE Trans Nucl Sci, 26:2797-

2805, 1979.

Tanaka E, "A fast reconstruction algorithm for stationary positron emission tomography based on a modified EM algorithm", IEEE Trans Med Im, 6:98-105, 1987.

Tanaka E, Mori S, Shimizu K, Yosikawa E, Yamashita T, Murayama H, "Moving slice septa and pseudo three-dimensional reconstruction for multi-ring PET", Phys Med Biol, 37:661-672, 1992.

Thompson CJ, Murthy K, Picard Y, Weinberg IN, Mako R, "Positron emission mammography (PEM): A promising technique for detecting breast cancer", IEEE Trans Nucl Sci, 42:1012-1017, 1995.

Townsend D, Frey P, Jeavons A, Reich G, Tochon-Danguy HJ, Donath A, Christin A, Schaller G, "High Density Avalanche Chamber (HIDAC) Positron Camera", J Nucl Med, 28:1554-1562, 1987.

Townsend DW, Spinks TJ, Jones T, Geissbuhler A, Defrise M, Gilardi M-C, Heather J, "Three-dimensional reconstruction of PET data from a multi-ring camera", IEEE Trans Nucl Sci, 36:1056-1065, 1989.

Townsend DW, Geissbuhler A, Defrise M, Hoffman EJ, Spinks TJ, Bailey DL, Gilardi M-C, Jones T, "Fully three-dimensional reconstruction for a PET camera with retractable septa", IEEE Trans Med Im, 10:505-512, 1991.

Townsend DW, Wensveen M, Byars LG, et al, "A rotating PET scanner using BGO block detectors: Design, performance and applications", J Nucl Med, 34:1367-1376, 1993.

Veklerov E, Llacer J, "Stopping rule for the MLE algorithm based on statistical hypothesis testing", IEEE Trans Med Im, 6:313-319, 1987.

Visvikis D, Wells K, Ott R, et al, "PETRRA: Preliminary experimental results from the first full size detector and dead time simulation of the count rate performance of a unique whole body PET camera", IEEE Trans Nucl Sci, 42:1031-1037, 1995.

Wagner HN jr, "Clinical PET: Its time has come", J Nucl Med, 32:561-564, 1991.

Webb S, Ott RJ, Bateman JE, Flesher AC, Flower MA, Leach MO, Marsden P, Khan O, McCready VR, "Tumour localization in oncology using positron emitting radiopharmaceuticals and a multiwire proportional chamber positron camera; Techniques for 3D deconvolution", Nucl Instr Meth, 221:233-241, 1984.

Wu C, Ordonez CE, Chen C-T, "FIPI: Fast 3-D PET reconstruction by Fourier inversion of rebinned plane integrals", Proceedings of the 1995 International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine, Aix-les-Bains, 241-245, 1995.

Zerby CD, "A Monte Carlo calculation of the response of gamma-ray scintillation counters", in *Methods in computational physics*, Academic Press, New York, 1:89-134, 1963.