



Emission Tomography and Image Reconstruction

16

Magdy M. Khalil

Contents

16.1 Introduction	409
16.2 Image Reconstruction	419
16.3 Conclusions	437
References	437

16.1 Introduction

In the early days of nuclear medicine, measurement of radioactivity administered into a human body was simply acquired by placing a Geiger counter over the desired region of interest. Further progress was undertaken using a rectilinear scanner. The breakthrough, as mentioned in Chap. 10, came from the development of the gamma camera and the use of the scintillation crystal coupled to photomultiplier tubes (PMTs). To this end, there was no available tool to measure the spatial extent of tracer distribution in three-dimensional (3D) fashion, and all measurements were confined to two-dimensional (2D) planar imaging. The third dimension is important to fully depict radiopharmaceutical uptake, hence enabling the interpreting physician to make a confident decision. Another feature of 3D imag-

ing is the ability to quantify tracer concentrations more accurately than with 2D imaging. Tracer uptake, residence time, and clearance rates are important dynamics of tracer biodistribution in diseased and healthy tissues, in which temporal sampling is particularly useful for studying tracer or organ kinetics. Adding the time dimension to 2D planar imaging is important in some scintigraphic studies, such as renal scintigraphy and planar equilibrium radionuclide angiocardigraphy (ERNA). In the former case, kidney function is studied through a time course of about half an hour, dividing the examination time into two phases (perfusion and function) such that the first minute is assigned to depict organ perfusion while the rest of the study is used to assess renal function. In planar ERNA, the time dimension is essential to make snapshots of different phases of the heart cycle through identification of the R-R signal during heart contraction. This helps to obtain valuable information about heart motion and to assessment of functional parameters.

M. M. Khalil (✉)
Medical Biophysics, Department of Physics, Faculty
of Science, Helwan University, Cairo, Egypt

Many nuclear medicine procedures are performed by acquiring planar views of the area under investigation. However, planar images are manifested by poor image contrast and lack of quantitative accuracy. Nuclear examinations such as bone scintigraphy and thyroid, parathyroid, and lung scanning are among those studies for which planar imaging is commonly used; however, under many circumstances the 2D nature of the acquired data have shortcomings in their yield of accurate diagnostic results, especially when dense overlying structures obscure the inspection of tracer spread and accumulation. This directly influences the interpretation results and may lead to an inconclusive diagnosis.

Many nuclear medicine procedures have been revolutionized by use of 3D imaging in terms of the amount of information that can be extracted and incorporated into the decision-making process. Among those are myocardial perfusions, brain, and bone imaging, for which tomographic acquisition provides a greater opportunity to visualize organs from different angular perspectives. This allows reading physicians to thoroughly investigate pathological lesions from many directions, especially when appropriate visualization tools are available on the viewing workstation. This in turn has had a positive impact on the diagnostic accuracy of many nuclear medicine examinations. For example, a tomographic bone scan is more sensitive than planar imaging and has been reported to improve the diagnostic accuracy for detecting malignant bone involvement [1].

16.1.1 History

Emission and transmission CT rely on the fact that to obtain a 3D picture of the human body a set of multiple 2D projections is required for image reconstruction. This necessitates collection of a sufficient amount of information about the object under examination. Johann Radon (1887–1956) introduced the principles of data formation through what is called the radon transform, which describes an object in 3D space as a sum of line integrals. In 1917, Radon developed a

solution for image reconstruction utilizing projection data sets and applied his technique to non-medical applications, namely, gravitational problems. In 1956, the reconstruction technique developed by Radon found another application by Bracewell in the field of radioastronomy [2]. Allan Cormack, a few years later after Bracewell, independently and without knowledge about Radon's work, developed a method for calculating radiation absorption distributions in the human body based on transmission scanning. Kuhl and Edwards were the first to introduce the concept of emission tomography using backprojection in 1963, and about 10 years later, Godfrey Hounsfield, the inventor of CT, succeeded to practically implement the theory of image reconstruction in his first CT scanner. Shortly after the invention, Hounsfield and Cormack were recognized by sharing the Noble Prize in Medicine and Physiology in 1979. By analogy to PET imaging, CT scanning was focused on brain imaging; however, body examinations were introduced a few years later, and the first body images taken in the body prototype machine were of Hounsfield himself on December 20, 1974 [3].

One of the earlier works on tomography was to move the object while keeping the imaging system stationary. This was in the late 1960s and early 1970s, when investigators used transaxial tomography to image a patient setting on a rotatable chair placed in front of a stationary gamma camera. After the mid-1970s, a gamma camera detector was mounted on a rotated gantry to take multiple images around the patient under investigation [4].

16.1.2 SPECT and PET

There are two 3D techniques provided by nuclear imaging namely single-photon emission computed tomography (SPECT) and positron emission tomography (PET). Both are noninvasive diagnostic modalities that are able to provide valuable metabolic and physiologic information about many pathophysiologic and functional disorders. A remarkable feature of SPECT and PET is their ability to improve contrast resolution

manyfold compared to planar scintigraphic imaging. The two imaging modalities have proved useful as applications in molecular imaging research and translational medicine. In addition, attempts to derive quantitative parameters are more accurate than planar imaging. Furthermore, when the timing factor is added to the 3D imaging, the amount of information that can be obtained from analyzing the data is significantly high. Two examples are worth mentioning when SPECT or PET is used to collect tracer spatial distribution and its associated temporal component. One of these is gated myocardial perfusion tomographic imaging, in which the tracer distribution and heart function can be captured in one imaging session, providing an assessment of myocardial perfusion (or metabolic) parameters, such as defect extent and severity or tissue viability, in addition to calculations of regional and global left ventricular function and ejection fraction. Another important area of application is the study of tracer distribution during the time course of tracer uptake and clearance from biological tissues. In these acquisition protocols, dynamic frames are collected over predefined timing intervals (or reframed in case of list-mode acquisition) to record tracer flow, extraction, retention, and clearance from the tissue of interest. The recorded data are then presented to an appropriate mathematical model to obtain physiologically important parameters, such as transport rate constants and calculation of tissue metabolic activity or receptor density (see Chaps. 20 and 21).

The addition of temporal sampling to tomographic imaging has other utilities, such as recording the respiratory cycle to correct for lung motion on myocardial perfusion imaging and to correct for spatial coregistration errors arising from temporal mismatch between computed tomography (CT) and PET in lung bed positions during whole-body fluorodeoxyglucose-F18 (FDG) PET/CT examinations. The time information required for motion characterization in four-dimensional (4D) imaging can be obtained either prospectively or retrospectively using respiratory-gating or motion-tracking techniques [5].

The advances in hybrid imaging and introduction of PET/CT and SPECT/CT to the clinic have

added another dimension to the diagnostic investigations; currently, hybrid modalities provide greater opportunity to study functional as well as morphological changes that occur at different stages of disease progression or regression. The characterizing aspects that distinguish these imaging methods from other imaging modalities are the underlying physical principles, the way data are acquired, image reconstruction and correction techniques, and finally image visualization, quantitation and display.

16.1.3 Resolution and Sensitivity

SPECT and PET imaging modalities have common and different characteristics in terms of spatial resolution and sensitivity. In general, clinical PET systems have better spatial resolution than SPECT; the former can provide an intrinsic spatial resolution of about 4–6 mm, and the latter can hardly achieve 10 mm full width at half-maximum (FWHM) using the conventional NaI(Tl) designs. The resolution of PET images is determined by many factors, which differ from those that affect SPECT resolution. Detector size, positron range, photon acollinearity, and some instrumental factors contribute by different degrees to the spatial resolution of PET images, as discussed in Chap. 12. On the other hand, SPECT imaging uses multihole collimation to identify structures and to determine directionality of the emitted radiation. This type of data collection imposes constraints on the overall system sensitivity and spatial resolution. There is often a trade-off between sensitivity and spatial resolution in collimator design. For instance, collimators with high spatial resolution have reduced count efficiency and vice versa. Another aspect of this trade-off is realized in some other (divergent) collimator geometry, in which the spatial resolution is improved while keeping the sensitivity at the same level, but this comes with a reduced imaging field of view.

In PET imaging, there is also a trade-off of these performance parameters but not in a similar manner as the principles of photon detection in PET imaging obviate the need for such photon collimation, leading to increased system

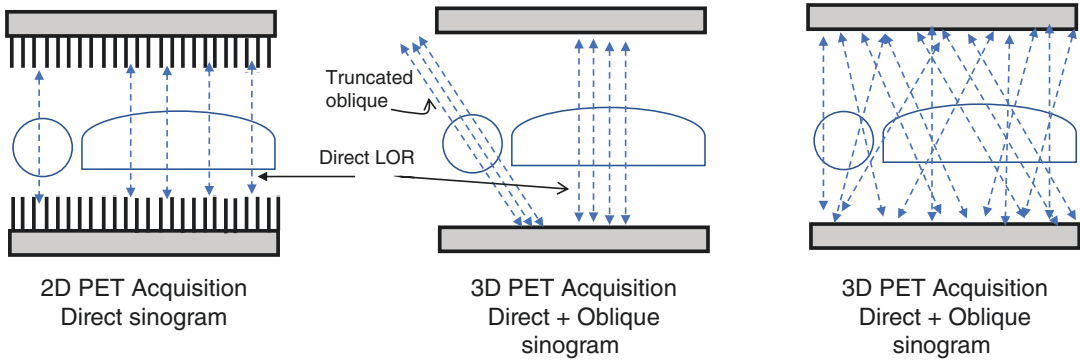


Fig. 16.1 Acquisition modes in positron emission tomographic (PET) scanning

sensitivity. However, collimation also exists in PET imaging in a form of 2D acquisition modes by using collimating septa between scanner rings. In 3D scanner configuration where the interplane septa are removed, a significant increase in detection efficiency is obtained (four- to sixfold) than that when the scanner is operated in 2D mode. In the later mode, collimator septa are placed between detector rings to confine the acquired projections to a set of 2D projection arrays. This facilitates image reconstruction so that any 2D reconstruction algorithms can be used similar to 2D SPECT data reconstruction. In this way, image reconstruction is implemented in an independent slice-by-slice manner. Image reconstruction in 2D is a straightforward procedure, while in 3D some kind of data manipulation is required to utilize the increased system sensitivity in improving image quality. In septaless or 3D acquisition mode, however, the scanner sensitivity is not uniform across the axial field of view, and approaches to reconstruct images are either to use fully 3D reconstruction techniques or to rebin the data into a 2D projection array. Figure 16.1 shows the two acquisition modes offered by PET scanning.

16.1.4 Image Acquisition

Data sampling by gamma camera detector is implemented by computer digitization for the events detected on the scintillation crystal. The computer matrix varies according to system sen-

sitivity, resolution, and data storage capacity. A lower matrix size, such as 64×64 and 128×128 , is commonly used in SPECT while being higher in PET due to the improved spatial resolution. The matrix size in X-ray CT is even higher than nuclear techniques due to the submillimeter resolution capabilities and superior photon statistics. However, in nuclear PET and SPECT imaging, the relatively lower photon flux due to radioactive decay properties, restrictions on the injected dose, lower detection efficiency, acquisition time, and the expected spatial resolution are among the factors for reducing the matrix size.

In planar imaging, the patient is positioned in front of the detection system, and adequate time is given to form an image. The resulting image is a depiction of tracer distribution in two dimensions, x and y . The third dimension cannot be realized as the collected counts over a particular point of the detector matrix are a superimposition of tracer activities that lie along the accepted beam path. This manner of data acquisition does not allow for extracting valuable information about source depth. However, to solve such a problem additional information must be provided to obtain further details about tracer spatial distribution. Moving the detector to another position can produce another image for the given tracer distribution, and another angular view would again reveal information that was not present in the previous two views. This process can be repeated several times to complete an acquisition arc of at least 180° , which is the smallest angular arc that can be applied to reliably reconstruct an

image [6]. Let us go through some basic geometric and mathematical principles of this type of data acquisition.

In SPECT imaging, data acquisition is performed by a one-, two-, or three-head camera that is adjusted to rotate around the patient over small angular intervals to acquire an adequate set of 2D projections. The increased number of detector heads serves to improve study sensitivity and reduces the acquisition time. This is dissimilar to PET scanning, in which the circular ring design circumvents the patient in a 2π fashion; thus, detector rotation is not necessary. However, in the old dual-detector coincidence gamma camera and partial-ring design, detector motion is required to satisfy angular requirements imposed by reconstruction algorithms. The detection process relies on annihilation photons ($\sim 180^\circ$ apart) and coincidence circuitry to record events in an emission path, called the tube or line of response (LOR).

Fig. 16.2 shows one projection view for SPECT and PET cameras such that the former is positioned to acquire a cardiac study while the latter was chosen to scan a brain patient. Suppose that we select one-detector row (i.e., 1D) of the detector 2D matrix, and in 2D PET this corresponds to one projection angle acquired using a single ring. The activity distribution within a patient injected by myocardial tracer or FDG is defined by $f(x,y)$, where x and y are the coordinates of the tracer uptake inside the patient boundaries. The counts collected over the elements of the projection row at an angle θ is denoted by the function $p(s,\theta)$, where θ is the angle subtended by the SPECT camera and the cartesian x and y coordinates as shown in Fig. 16.2. This is also the same angle at which the PET scanner was chosen to look at the brain study.

According to Radon, every projection bin of the 1D image is a result of count accumulation along the path traversed by the emitting radiations and falling perpendicular on the detector plane. However, in PET it is the line that connects a detector pair in coincidence. One can therefore consider the acquired counts at a given angle as a compressed version of the slice under investiga-

tion. At the end of data acquisition, we obtain a multiple number of projection angles; each is a compressed version of the object distribution viewed from different angles. In PET imaging, data format is mostly represented by rebinning the coincidence data in a sinogram. Another format of data acquisition and event storage is list mode, in which events are individually recorded for their timing, position, and possibly any other relevant attributes, such as energy.

The problem now is how to reconstruct or get a solution for the tracer concentration given the information provided by the set of projections. It is easy to understand that once we are able to reconstruct or find a solution for one transverse image, then it becomes possible to obtain the contiguous slices following the same pathway. In the given examples, the function of the reconstruction algorithm is to find the best estimate of the tracer spatial distribution within the slices taken across the myocardium or the brain tissues. Apart from considering the effect of photon scatter and detector response, one can write the measured projection data as

$$p(s,\theta) = \int f(x,y) e^{-\int \mu(x,y) dt'} dt \quad (16.1)$$

This is the attenuated radon transform, and solving the equation for $f(x,y)$ is the way to find estimates of tracer activity within the patient and hence image reconstruction. Here, t and s are elements of a coordinate system such that t is passing along the direction of the rays and perpendicular on the detector plane while s is the axis parallel to the detector. In terms of x and y directions, s and t are defined as follows:

$$s = x \cos\theta + y \sin\theta$$

$$t = -x \sin\theta + y \cos\theta$$

By neglecting the exponential term, the resulting formula will be the Radon transform equation, which states that the acquired count over a particular projection bin $p(s,\theta)$ is the integration of tracer activity along the line that passes through the object studied and, in SPECT, falling perpendicular on the detector plane, while in PET it is the line that connects a coincident detector pair.

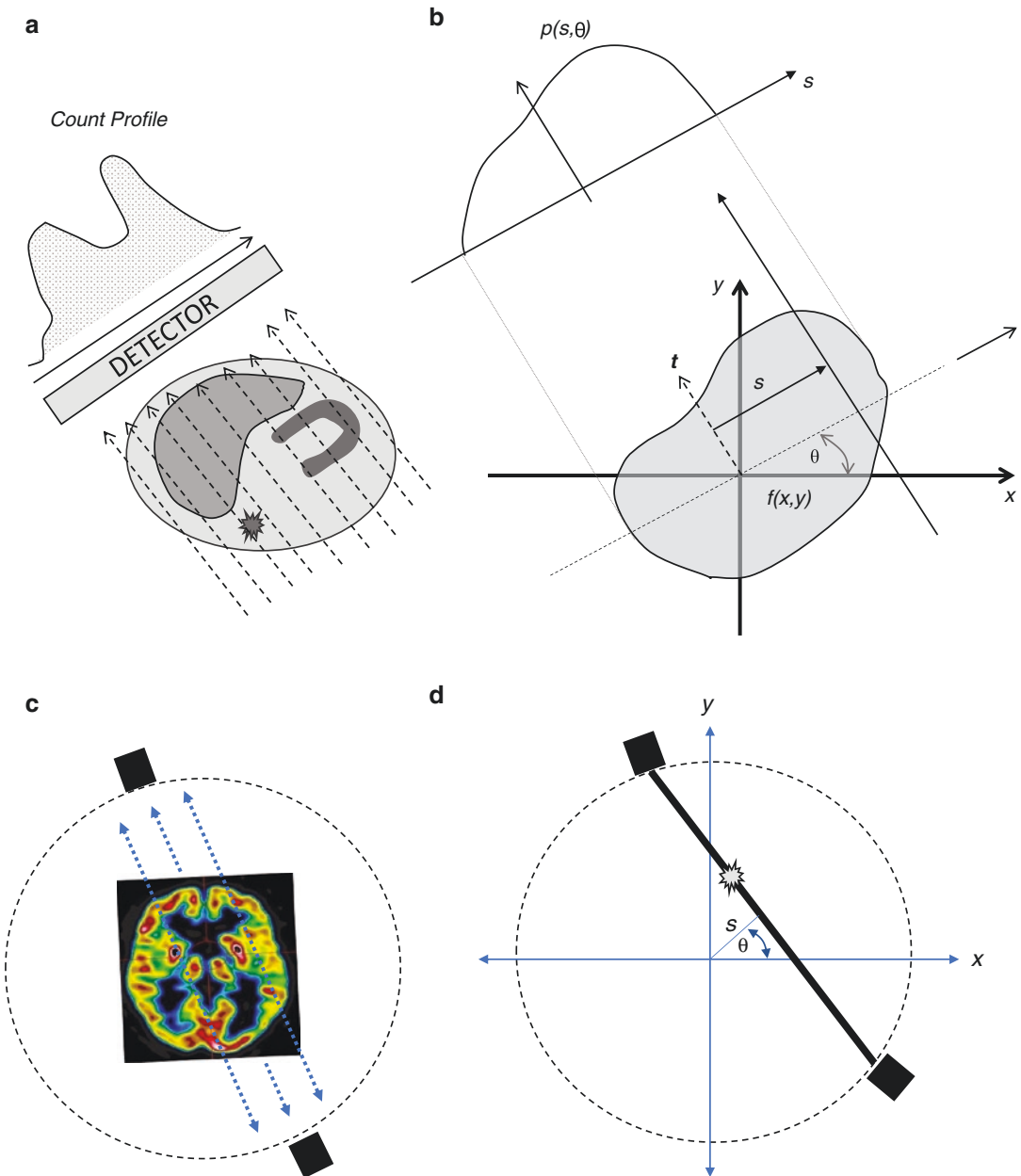


Fig. 16.2 (a) Projection profile for a one-dimensional (1D) row of the detector is displayed; a varying count intensity is evident. Any point on the profile is the line integral of all activity concentrations lying along the path of the ray. (b) The coordinate system (t, s) is defined so that s is parallel to the detector plane, while t is perpen-

dicular on it. This coordinate system is used to define the projection profile $p(s, \theta)$ in relation to the stationary coordinate system (x, y) . In (c, d) positron emission tomographic (PET) acquisition setting, similar projections are defined by sinogram variables s and θ , where both determine the location of the annihilation site on the sinogram

The process that maps the tracer activity $f(x, y)$ onto the projection image $p(s, \theta)$ is defined as the X-ray transform.

In the case of SPECT, the exponential term of the formula denotes the amount of photon attenuation that extends from the site of emission $f(x, y)$

to the detector plane, demonstrated in Fig. 16.3, whereas in PET it refers to the amount of attenuation experienced by annihilation photons while traversing the corresponding patient thickness. Figure 16.4 shows how an attenuation correction in PET imaging can be solved by calculating the

probability of detecting two coincident photons by a detector pair.

It is noted that the attenuation correction factor is a function of the patient thickness and independent of the emission site given a recorded LOR. By moving the exponential term to outside the integra-

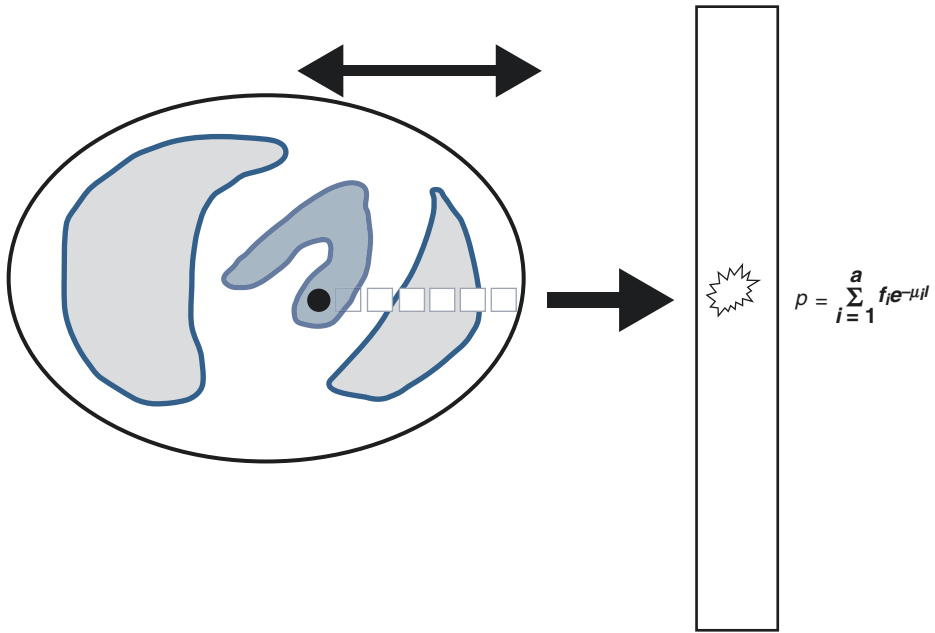
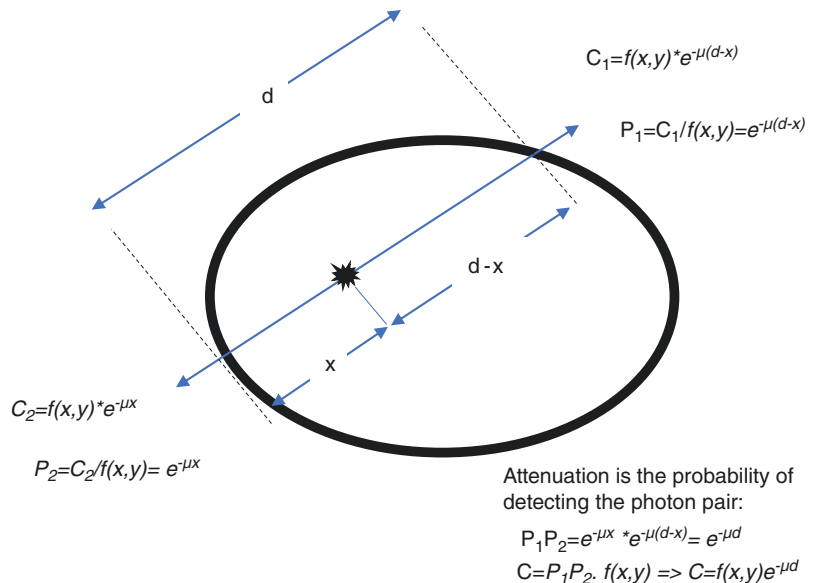


Fig. 16.3 Photon attenuation in single-photon emission computed tomography (SPECT)

Fig. 16.4 Photon attenuation in positron emission tomography (PET)



tion, denoting the measured projection as I and the integration term as I_0 , and rearranging the formula, we can obtain the attenuation correction factors (ACF) required to correct a measured LOR:

$$\text{ACF} = I_0 / I$$

This is simply achievable in practice using a transmission source where I_0 is the measurements performed while the patient is outside the field of view (i.e., blank scan), and I corresponds to the data taken when the patient is positioned inside the field of view (i.e., transmission scan).

Each LOR can be corrected for attenuation by multiplication with the corresponding correction factors, or the latter data can be reconstructed to obtain a spatial distribution of attenuation coefficients. In SPECT attenuation correction, the direct multiplication of the emission data by the correction factors is not applicable due to dependence of photon attenuation on the emission site, which is unknown. Instead, the logarithmic ratios of the initial and transmitted projections are reconstructed to obtain a spatial distribution of attenuation coefficients or what is known as an attenuation map.

The introduction of hybrid imaging such as SPECT/CT and PET/CT has allowed the use of CT images to correct the radionuclide emission data for photon attenuation. CT images provide low noise correction factors and faster scanning times, but corrected data may suffer from quantitative bias and correction artifacts. A CT scan also provides high-resolution anatomical images and with image coregistration serves to strength the confidence of lesion localization detected in radionuclide images. Radioactive sources provide more noise, less bias, and increased imaging time. Different methodologies have been devised to correct for the bias introduced by CT-based attenuation correction and methods to reduce noise propagation into radionuclide emission images when radioactive transmission scanning is used.

16.1.5 X-Ray CT

For a monoenergetic X-ray beam passing through an object of thickness L and linear attenuation

coefficient μ , the transmitted radiation can be calculated from

$$I = I_0 e^{-\mu L}$$

where I and I_0 are the transmitted and initial beam intensity, respectively. For an X-ray beam in CT, the rays traverse various body tissues of different attenuation properties due to their various compositions and effective Z number. Thus, the amount of attenuation that the beam encounters is equal to the total sum of all μ values that lie along the beam path.

Therefore, the measured transmission data for an X-ray beam of initial intensity I_0 passing through a human body can be written as

$$p(s, \theta) = I_0 e^{-\int \mu(x, y) dt}$$

Rearranging the formula and renaming the measured projection $p(s, \theta)$ as described, we obtain

$$\ln \frac{I_0}{I} = \int \mu(x, y) dt$$

where $\mu(x, y)$ is the linear attenuation coefficient for a pixel located at position (x, y) , and the integration is the line integral of attenuation coefficients along the transmission beam (see Fig. 16.5).

The reconstruction algorithm here does not try to find the activity distribution of the tracer, but it estimates the spatial distribution of attenuation coefficients using two pieces of information, the initial beam intensity I_0 and the transmitted projection data. Note that this is the same equation used to derive the correction factors for PET emission data since it accounts for the total amount of photon attenuation experienced by the initial X-ray beam I_0 while moving through the object. Actually, in real practice and data analysis of X-ray CT, determination of the distribution of attenuation coefficients is not simply performed by solving the equation stated here; several pre- and postprocessing steps are taken to correct for many variables and confounders that deviate the practical measurements from being consistent with the theoretical ideal conditions; for further details, see [7].

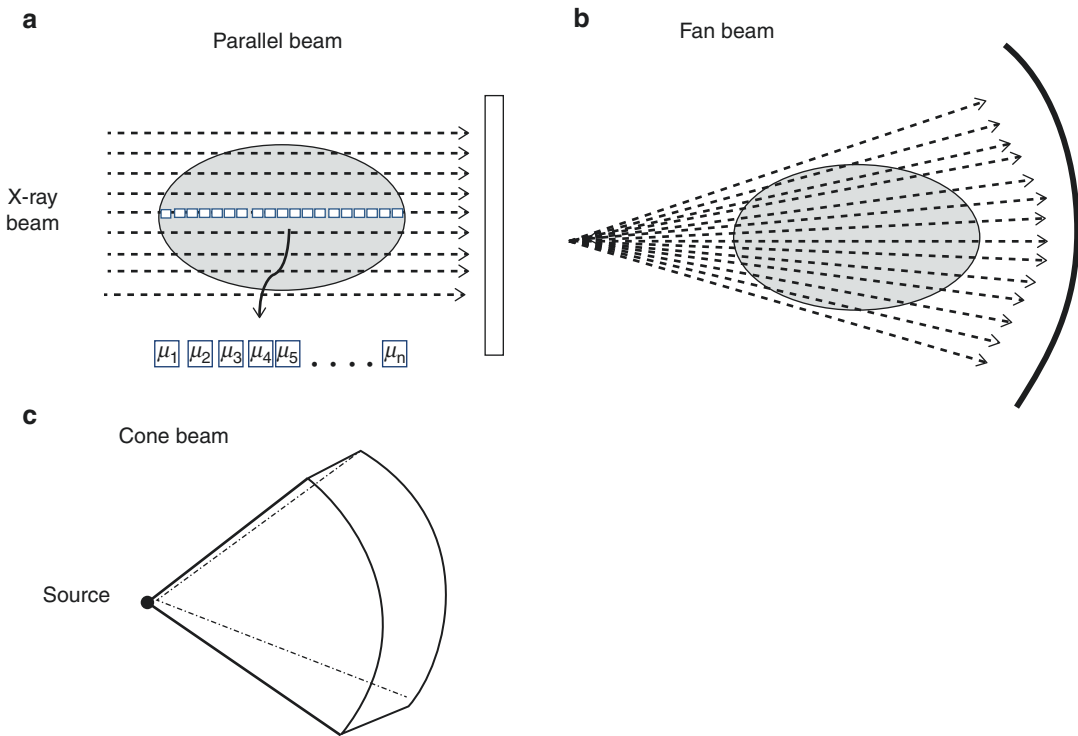


Fig. 16.5 In X-ray CT, different geometries have been used in image acquisition which in turn posed different requirements on image reconstructions. (a) Image reconstruction is straightforwardly implemented using direct filtered backprojection (FBP). The other geometry is

shown in (b) where image can be either reconstructed using rebinning or direct FBP algorithm. (c) Cone beam: another design currently used in commercial CT scanners where image reconstruction is modified to adapt and account for beam geometry

16.1.6 Sinogram

Rebinning the acquired data in a single diagram such that the projection bin represents the horizontal axis while the projection angle is placed on the vertical direction produces a sine wavelike pattern called a *sinogram*. Representing the acquired data in a sinogram has several benefits in terms of data processing, image reconstruction, and correction techniques. Also, it is useful in inspecting detector failure, in which a diagonal black line in a PET sinogram indicates an artifact in a single detector element, while a diagonal band could indicate a malfunction of a detector block [8]. It can also be used to correct for patient motion and for other correction techniques. Note that one selected pixel on the sinogram should indicate the total counts collected for a particular LOR regardless of any contamination from any

other events. In SPECT, it is the integral of counts that lie along the emission path and falling perpendicular on the detector surface (i.e., line-integral model).

For a point source located at the center of the field of view, the resulting sinogram is just a vertical line that extends from the top to the bottom of the sinogram. Further, a horizontal line passing through the sinogram indicates a particular projection angle taken for a transverse slice. Figure 16.6 shows sinograms for an object having one and two hot spots on the transverse section. The figure clearly represents location (angle and position) and intensity of “two lesions” and also real complex emission similar to clinical studies (i.e. brain) where the object consists of a large number of points taken at multiple projection angles. Figure 16.7 shows also how sinogram formation relies on number of disintegrations collected from positron emission as opposed to

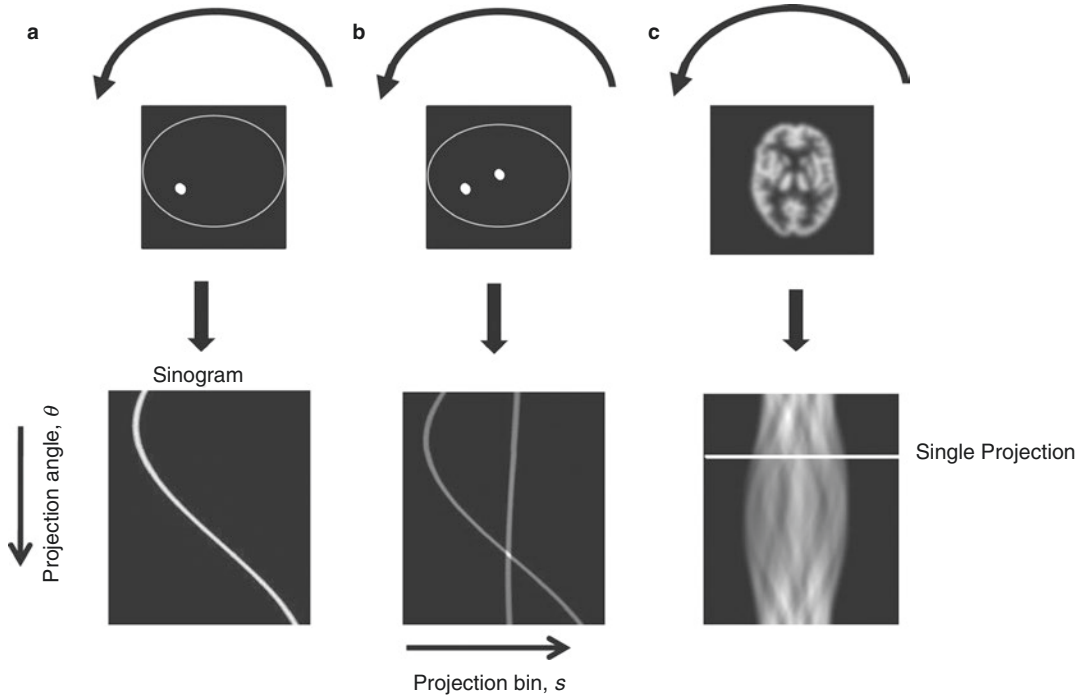


Fig. 16.6 Sinograms for different activity distributions. The sine wave pattern can be seen for a single hot lesion (a) and two hot lesions (b). The third sinogram (c) is more

complicated due to its representation for many points in the projection profile, including all the angular views

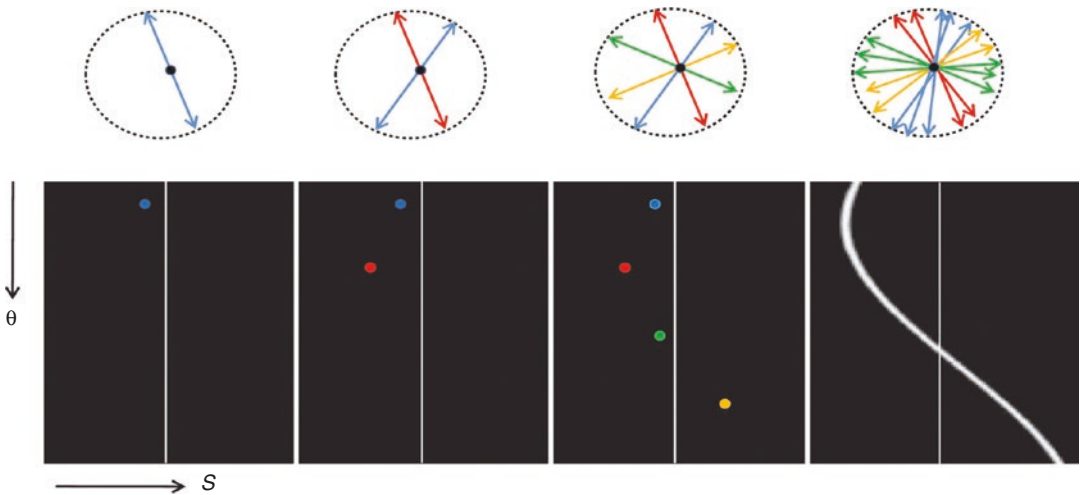


Fig. 16.7 The rebinning of the acquired events as the source decays into a sinogram that represents the projection bin on the horizontal axis, while the vertical axis denotes the projection angle. Note that each colored line

of response (LOR) refers to a particular projection view; in other words, it points to a certain set of parallel LORs taken at a given angle

conventional SPECT systems, in which detector rotation is necessary to build up a complete sinogram. This is one of the advantages provided by PET scanners based on circular design since all projections are acquired simultaneously and also possibly in 3D fashion. This characteristic is absent in most conventional designs of the gamma camera, for which detector rotation is essential to accomplish the task.

16.2 Image Reconstruction

16.2.1 Analytic Methods

16.2.1.1 Simple Backprojection

As described, a collected count from a projection element according to the Radon transform is a line integral of tracer concentration along the emission path length. The task placed on the reconstruction algorithms is to find the spatial distribution of tracer activity within the body segment in question. One way to reconstruct an image from the raw data is to redistribute the collected counts (i.e., backproject) over the contributing individual pixels that lie along the path of the rays in the reconstruction matrix. Repeating this process for each projection element and for each acquired angle, one can obtain a picture of the tracer concentration as shown in Fig. 16.8. It can be seen that this method of image reconstruction cannot reveal useful information about tracer

distribution due to the blurry appearance and substantially degraded signal-to-noise ratio.

Backprojection operation at point b can be represented as

$$f_{BP} = \int_0^{\pi} p(x \cos \theta + y \sin \theta, \theta) d\theta$$

The backprojected image f_{BP} at a particular point b is the result of summing all the corresponding projection bin values across all angular views taken during data acquisition. Here, s is the location of the projection bin on the detector. In PET geometry, the backprojection operation is performed for those LORs that connect detector pairs in coincidence. For obvious reasons, this process of count redistribution cannot determine the exact site where photon annihilation took place. Therefore, all pixels along the ray path are equally likely to get the same amount of counts. In PET systems with time of flight, calculation of the arrival of the two photons allows reduction of this LOR to a significantly smaller distance (based on system timing resolution) that, if included in image reconstruction, would result in an improvement of signal-to-noise ratio.

This is actually not the exact description of the backprojection operation since it is implemented on a grid of finite elements or computer matrix (and acquisition geometry), and thus it is possible that, for a given pixel, the backprojected ray can pass through a small part or intersect the pixel at

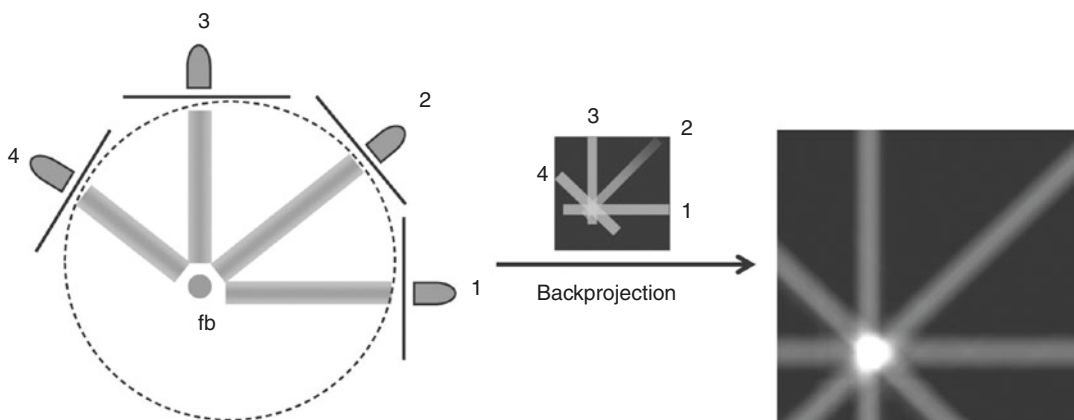


Fig. 16.8 Image reconstruction using simple backprojection

its full length. Therefore, a number of backprojection methods have been developed to deal with this point. Methods used in forward- and backprojection are pixel driven, ray driven, distance driven, distance weighted, matrix rotation, and others. Also, a combination of these methods, such as ray driven and pixel driven, can be used [9]. However, these methods differ in their computational efficiency, interpolation, and estimation accuracy. In iterative reconstruction, they should be carefully selected since several iterations may accumulate interpolation errors, introducing reconstruction artifacts.

Two clinical examples are shown in Fig. 16.9, one slice from a myocardial perfusion SPECT and another one from brain-FDG PET

study. The characteristic blurring appearance of simple backprojection is clear in both studies, with most low-frequency components overexpressed with a remarkable reduction of high frequencies. This significant artifact is attributed to the fact that the sampling criteria do not match the model assumptions; hence, the reconstructed image is far from an accurate estimate of the tracer distribution. Simple backprojection assumes that data are collected with infinite linear and angular sampling, and the data collected are free from attenuation and scattered radiation in addition to shift-invariant and perfect system response (Fig. 16.10). These assumptions are violated in practice due to image digitization and the discrete angular intervals undertaken in image

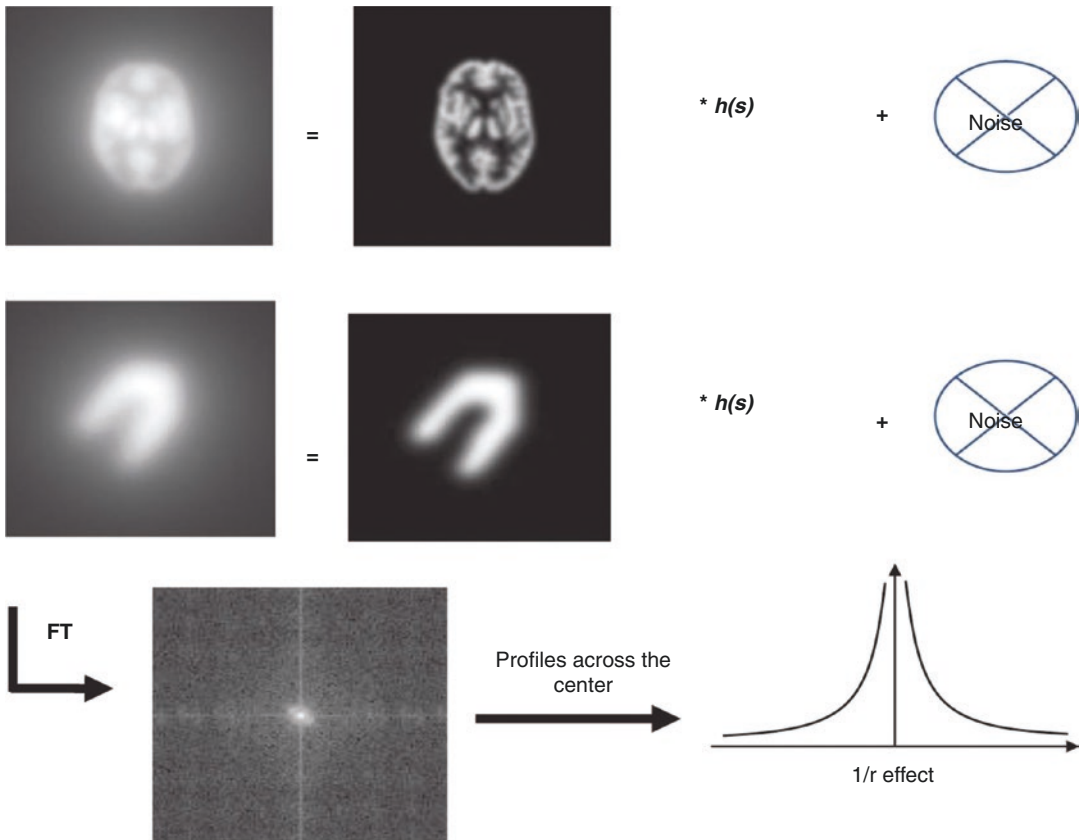
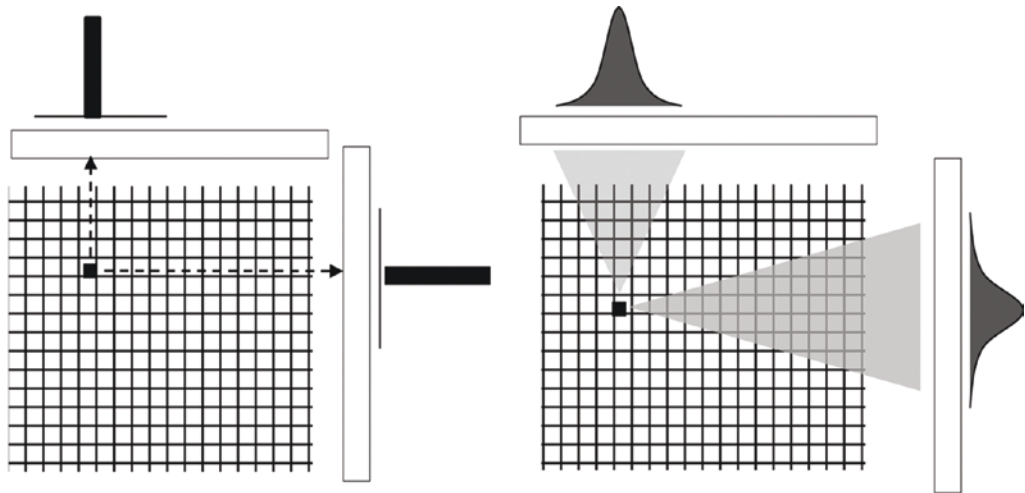


Fig. 16.9 When the backprojection images are Fourier transformed the low frequencies are overemphasized while the high frequencies are reduced showing a pattern called 1/r effect. This backprojected blurry images are therefore corresponds to a convolution of the underlying true activity distribution with the $h(s)$ function. Analytic

approaches remove this effect by deconvolving the acquired data with the blurring function, neglecting the noise component and leading to a tremendous increase in image noise. As a result, regularization using a smoothing function is required



Analytic Approach

- Perfect system response
- No attenuation
- No scatter or randoms
- No Noise Model
- Infinite Sampling

Realistic Approach

- Beam divergence
- Photon attenuation
- Photon scatter
- Noise
- Finite Sampling

Fig. 16.10 Ideal versus realistic model of analytic image reconstruction

acquisition. Furthermore, the emitted photons undergo different types of interactions, resulting in photon loss or recoiling from the original path, and hence invalidate the absence of photon attenuation and scatter assumption in data acquisition. The measured projections are noisy due to the Poisson statistics of the radioactive decay process, and ignoring the noise component serves to alter the statistical properties of the reconstructed images and degrades image quality.

A profile drawn over the Fourier transform (FT) of the backprojected image shows a damping function that extends from the center of the spectrum (low-frequency region) toward the periphery (high-frequency region). This is referred to as the $1/r$ effect, in which the reconstructed image can be described as a convolution of the underlying activity distribution and a $1/r$ blurring function (Fig. 16.9). This situation can be written in the frequency domain as

$$F_{\text{BP}} = \frac{1}{\sqrt{v_x^2 + v_y^2}} F(v_x, v_y) = \frac{1}{v} F(v_x, v_y)$$

where the backprojected image F_{BP} is equal, in theory, to the original image $F(v_x, v_y)$ multiplied by the inverse of the function $h(s)$ in the frequency space. The latter function is defined as the system output to an ideal point source object and describes the system blurring effects on image formation. It is usually called the system spread function or point spread function (PSF). It is the key to solving the problem of backprojection by removing the blurring effect shown in Fig. 16.9 by either convolving the measured projections with the function $h(s)$ or multiplication in the frequency domain as described by the equation. Similarly, in 3D image reconstruction without data truncation, the backprojected image can be convolved with an appropriate 3D filter function to get an estimate of the original object distribution; alternatively, the measured projections are convolved with the 3D filter function. Before proceeding further to use this approach in image reconstruction, an important theorem that is central to many analytic reconstruction techniques should be discussed.

16.2.1.2 Fourier Reconstruction Theorem

Fourier analysis has a wide range of applications in many disciplines of science and engineering. This includes image and signal processing, filtering, image reconstruction, and many other biomedical applications. It has also been used in radioastronomy, electron microscopy, optical holography, magnetic resonance imaging (MRI), CT, and radionuclide SPECT and PET imaging [10]. Refer to Chap. 15, in which the FT is applied to a number of useful applications in nuclear medicine. In short, projection data and reconstructed slices can be represented in two different domains: spatial and frequency. The FT for a given input function can be represented by the sum of the sine and cosine waves with different amplitudes and phases.

Image reconstruction based on the FT is different from simple backprojection, and both can be combined to yield a variety of reconstruction approaches, as will be discussed further. The concept can easily be understood if we reversely assumed that we already have a transverse section of a patient thorax in which we can see the myocardium, and the 2D FT of this section has been calculated. The reconstruction theorem based on Fourier analysis states that a profile taken at a certain angle (θ) from the 2D FT of the transaxial section is equal to the 1D FT of the projection profile computed at the same angle. This is the underlying assumption of Fourier reconstruction theorem or the *central section* theorem, which relates the acquired projection

data to the reconstructed image by the aid of Fourier transformation (Fig. 16.11).

Suppose the Fourier coefficients (intensity values in the frequency domain) are defined by the function $F(u, v)$, which is the 2D FT of the activity distribution $f(x, y)$ for a given cross-sectional slice; then, it can be proven that

$$F(v_x, v_y) = P(v, \theta)$$

where $P(v, \theta)$ is the FT of the projection $p(s, \theta)$, which is the function we have used to describe the counts collected over a 1D detector row.

Figure 16.12 summarizes the steps involved in Fourier reconstruction for a myocardial perfusion study, where the 1D FT of projection data is first calculated for all angular views, then data are collected in a 2D format and interpolated to account for gaps between views. Finally, inverse 2D FT is computed to yield a reconstructed myocardial image. Here, u and v are the spatial frequencies in the Fourier space and are defined in a square matrix; however, the polar sampling regime taken by the detector does not match the rectangular requirements, and therefore interpolation is required. Such a problem could be dealt with using standard interpolation methods or interpolation by gridding, taking into account that the accuracy of the results depends strongly on the interpolation method [11].

By analogy to 2D Fourier image reconstruction, a central plane through the FT of the 3D activity distribution is equal to the 2D FT of the 2D parallel projection data taken at the same ori-

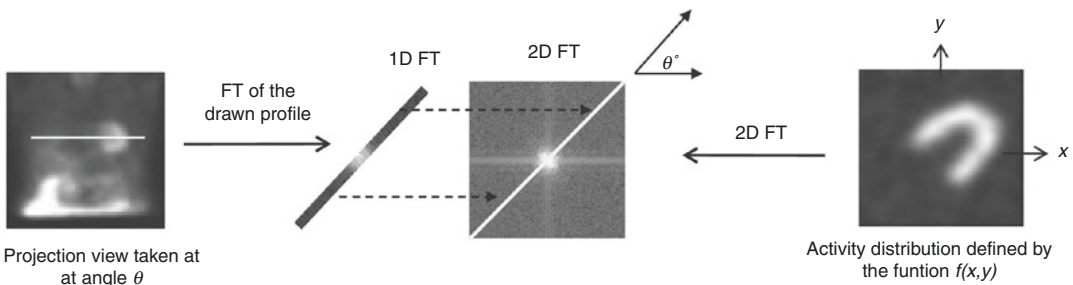


Fig. 16.11 The principle of two-dimensional (2D) Fourier reconstruction. The one-dimensional Fourier transform (1D-FT) of a horizontal profile drawn over a

projection image at angle θ° is equal to the 2D-FT of the reconstructed image taken at the same angle

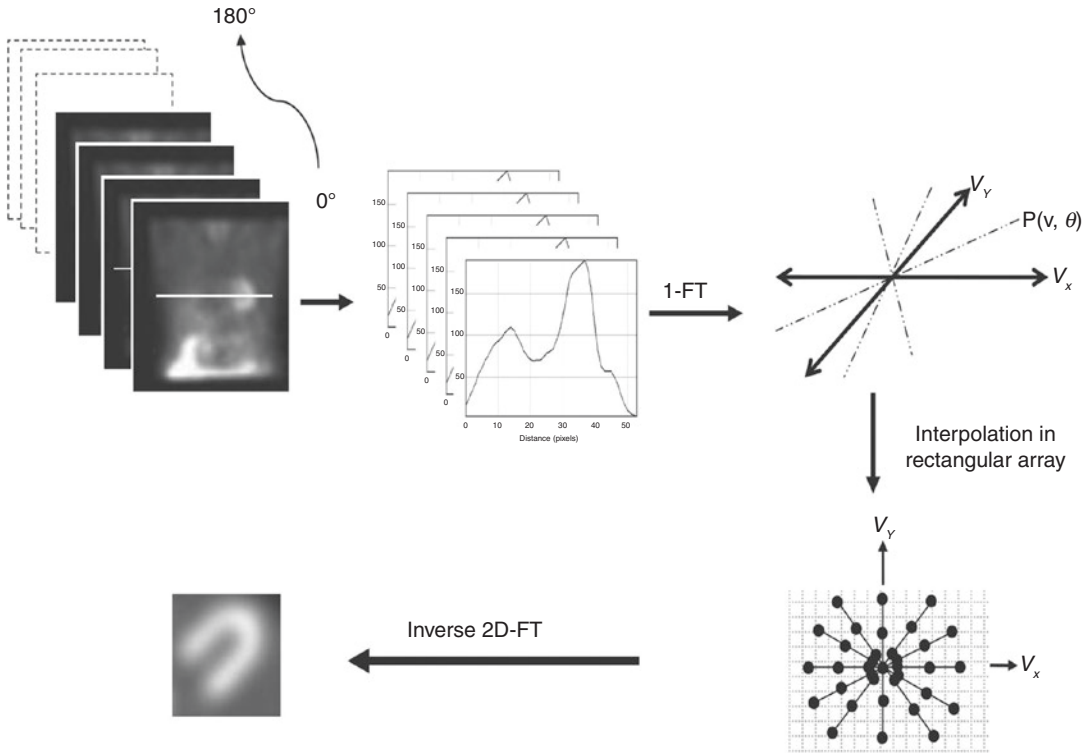


Fig. 16.12 Fourier transform reconstruction theorem states that the Fourier transform of a one-dimensional (1D) projection profile is equal to the two-dimensional (2D) Fourier transform of the corresponding activity distribution imaged at the same angle. This example shows a 1D profile taken across the patient's heart for all angles;

then, the FT was calculated and interpolated in a rectangular array to obtain a 2D data set. Finally, the inverse Fourier transform is computed to generate the corresponding activity distribution represented here by the transaxial myocardial slice

entation. However, the 3D transform of the object has different and more complex structure manifested by local sampling density when compared to 2D and thus requires special interpolation and weighting approaches [12].

The central section theorem and simple back-projection can be combined in different forms of image reconstruction utilizing the mathematical properties of FT and convolution theorem, which states that convolution in the spatial domain is equivalent to multiplication in the frequency domain. However, these methods differ in the order of reconstruction steps regarding whether convolution or backprojection is accomplished first and if convolution is implemented in the spatial or frequency domain, together with their computational efficiency.

Backprojection filtering (BPF) or filtering of the backprojection is one of these reconstruction approaches that combines Fourier reconstruction and backprojection in one procedure. BPF starts first by backprojecting the image into a reconstruction matrix, 2D FT is then computed, the result is multiplied by 2D ramp filter, and finally image reconstruction is performed by taking the inverse 2D FT. Image reconstruction can also be implemented by convolving the projection data with a convolution kernel, and then the product is simply backprojected to produce an image of the object activity distribution. However, the most computationally efficient and easy to implement is filtered backprojection, which has been extensively used in the routine practice of image reconstruction.

16.2.1.3 Filtered Backprojection

The most analytic approach that is used in SPECT and PET reconstruction is filtered backprojection. It has a historical dominance in many applications due to its speed and easy implementation in software reconstruction programs. It relies on filtering the projection data after Fourier transformation of all the acquired angular views; then, backprojection is carried out to give an estimate of the activity distribution. Data filtering is performed to eliminate the $1/r$ effect that works to blur the reconstructed images and is implemented in the Fourier space. Backprojection alone yields an image dominated by low-frequency components. By looking at the reconstructed brain and cardiac slices in Fig. 16.9, one can perceive the smoothing appearance of the images due to the prevalence of low frequencies with difficulty in identifying small details, a situation that results in a significant loss of signal-to-noise ratio. This problem can be tackled by using a ramp filter, which serves to suppress low frequencies and enhance high-frequency components of the projection data.

The ramp filter function $|v|$, as can be seen in Fig. 16.13, is a diagonal line that extends from the center in the frequency space to a sharp cutoff value. This significantly reduces the drawbacks of the backprojection step in image reconstruction. However, the sharp cutoff value has a disadvantage of producing count oscillations over

regions of sharp contrast [13]. Further, it increases the image noise due to the enhancement of the high-frequency components. To overcome this problem, an additional filter function is often used with the ramp filter to roll off this sharp cut-off value and to suppress high frequencies to a certain level.

The steps involved in reconstructing one slice using filtered backprojection (FBP) are demonstrated in Fig. 16.14 and summarized as follows:

1. 1D FT is calculated for each projection profile.
2. The Fourier transformed projections are multiplied with the ramp filter (plus a smoothing filter) in the frequency domain.
3. The inverse FT of the product is computed.
4. The filtered data are backprojected to give an estimate of the activity distribution.

These steps can be written mathematically as

$$f(x,y) = \int_0^{\pi} p^F(s,\theta) d\theta = \int_0^{\pi} p^F(x \cos \theta + y \sin \theta, \theta) d\theta$$

The reconstructed image $f(x,y)$ is obtained by filtering the projection data in the frequency space (by multiplication with the ramp-smoothing function); then, the filtered data p^F are backprojected in the spatial domain to obtain the object activity distribution. 2D FBP is used in the

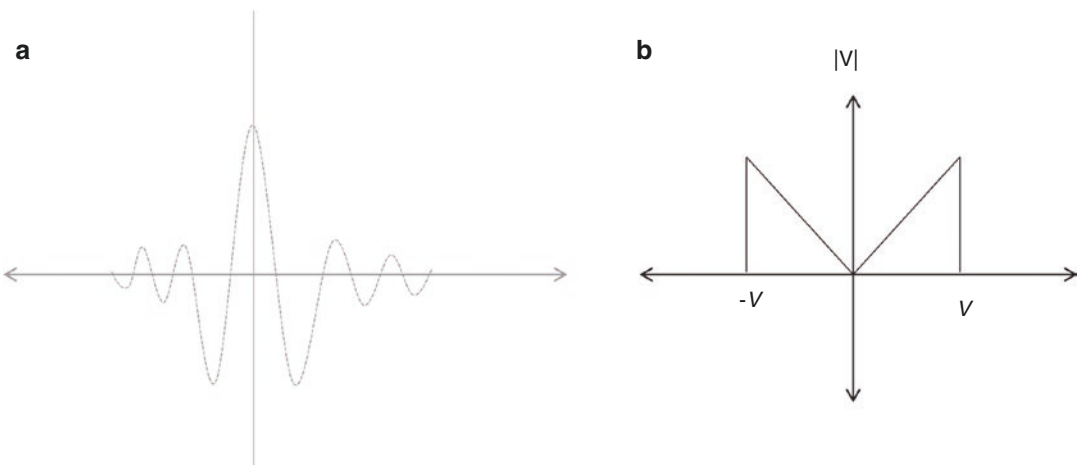


Fig. 16.13 Ramp filter in (a) spatial and (b) frequency domain

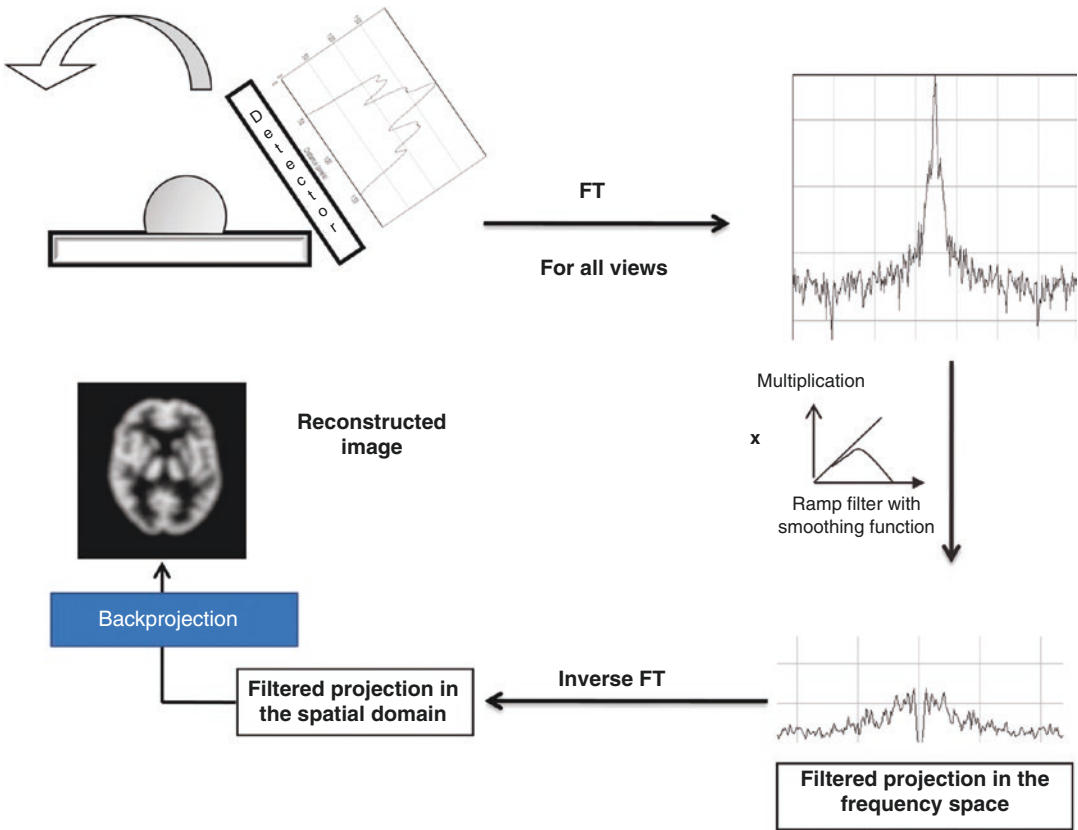


Fig. 16.14 [14] Steps involved in filtered backprojection (FBP) image reconstruction. The projection profiles are Fourier transformed and are then multiplied by the ramp function to yield filtered data in the frequency domain.

The inverse Fourier transform is then computed for the filtered data to move back to the spatial domain, and then backprojection is implemented. The filtration step can be performed prior to or after backprojection

reconstructions of the 2D PET (septa extended) and SPECT images acquired with parallel hole or fan beam collimators. An image reconstructed with FBP is demonstrated in Fig. 16.15 using different projection angles.

16.2.1.4 Filtering

As shown in Fig. 16.13, a ramp is a high-pass filter that does not permit low frequencies to appear in the image; therefore, it is used to overcome the problem of simple backprojection in image reconstruction. However, this filter has positive coefficients near the center and negative values at the periphery, as can be seen in Fig.16.13a, in which the filter is plotted in the spatial domain. These characteristics of a ramp filter can introduce artifacts at regions that lie close to areas of

high activity concentrations. This can be noted in the clinic in patients with full bladder activity undergoing bone SPECT imaging over the pelvic region. A severe cold artifact could be seen on the femoral head due to multiplying the ramp negative values with the projection counts. This could adversely affect the interpretation process and might be resolved by emptying the bladder and repeating the scan or reconstructing the image using iterative techniques [15]. Another example can be seen in patients scheduled for whole-body FDG scanning and who have full bladder activity. This negative lobe effect introduced by a ramp filter can also cause a reduction of the inferior wall counts in myocardial perfusion SPECT studies if there are increased extracardiac activity concentrations in close proximity to the heart

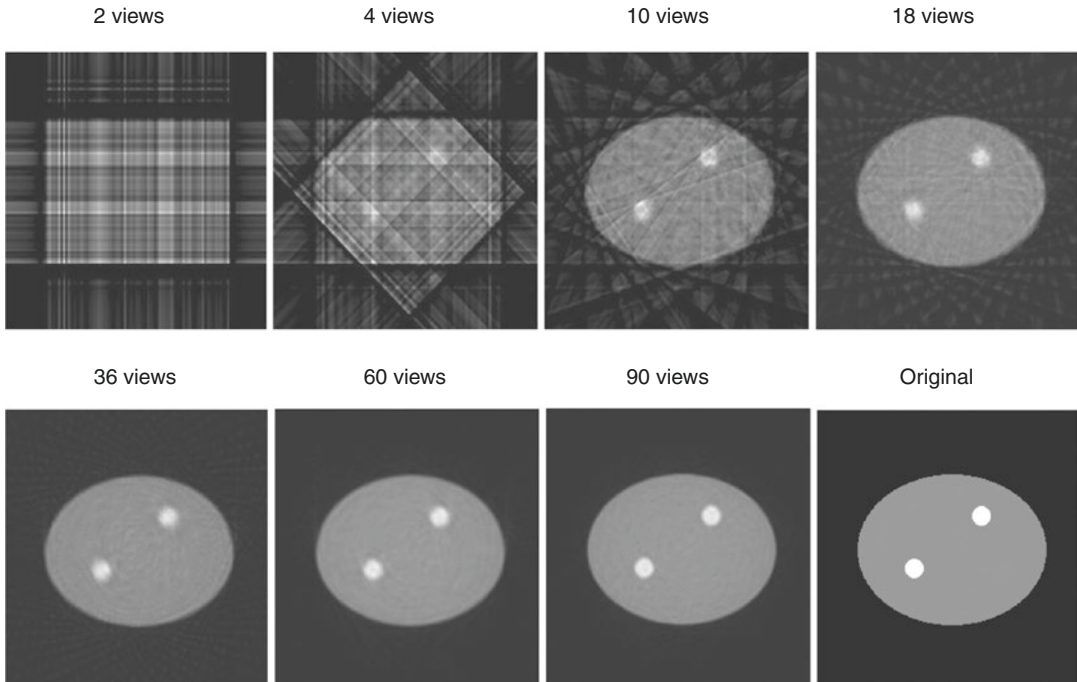


Fig. 16.15 Filtered backprojection (FBP) using different viewing angles

boundaries. This could result in an impression of diseased myocardial segments, causing false-positive results.

Another drawback of a ramp filter is its property of elevating the high-frequency components, thus increasing the noise level of the reconstructed images. An analytic solution for data acquired with noise is an ill-posed problem in which small perturbations (noise) in the input data cause a significant impact on the solution. Thus, a smoothing filter (regularization) is commonly used with a ramp filter to eliminate the noisy appearance of the ramp-filtered data and to improve image quality. Many filter functions were used with a ramp filter in several applications of nuclear medicine, such as Shep-Logan, parzen, hann, Hamming, and the commonly used Butterworth filter. Another class of filters has been proposed to correct for detector response function in image reconstruction, such as Metz and Wiener. Both filters rely on a system modulation transfer function taken at a certain depth and thus do not match the requirement of the shift-

variant response imposed by the detector system. The inclusion of the detector response function in iterative reconstruction showed superior performance over other methods of image restoration.

A low cutoff value may smooth the image to a degree that does not permit perceiving small structures in the image, leading to blurred details and resolution loss. On the other hand, higher cutoff values serve to sharpen the image, but this occurs at the expense of increasing the amount of noise in the reconstructed images. The optimum cutoff value is therefore the value at which a fair suppression of noise is achieved while maintaining the resolution properties of the image. This trade-off task of the cutoff frequency is important to properly use a given filter function and to improve the image quality as much as possible. The cutoff value depends on factors such as the detector response function, spatial frequencies of the object, and count density of the image [16]. Better isotropic resolution properties are produced with 3D smoothing, and therefore it is preferred over 1D filters applied for individual slices.

However, a 2D filter for the projection data may produce almost equal smoothing effects and is also computationally less intensive.

16.2.2 Summary of Analytic Image Reconstructions

Analytic approaches for image reconstruction in emission tomography seek to find an exact solution for tracer activity distribution. There are a number of assumptions that are invalid under the imaging conditions encountered in practice. Thus, the results provided by FBP are suboptimal to restore the true activity concentrations accumulated in target tissues. Images reconstructed with FBP need a number of corrections to improve the reconstruction results. As mentioned, effects of attenuation, scatter, and detector response are potential degrading factors that FBP does not account for in the reconstruction process. Nevertheless, this reconstruction method has the advantages of being fast and easy to implement, and nuclear physicians have long-term experience working with its outcome. Most image reconstruction in SPECT is implemented on a 2D slice-by-slice basis, so that at the end of image reconstruction one can obtain a complete set of transverse slices that, if stacked together, would represent the tracer distribution within the reconstructed volume. In PET image reconstruction, however, the same situation exists when data are acquired using 2D acquisition mode or the 3D data set are sorted into 2D projection arrays. Analytic image reconstruction can be summarized as follows:

1. Analytic reconstruction using FBP does not account for the inherent statistical variability associated with radioactive decay, and data collected are assumed to follow Radon transform, for which the object measured is approximated by line integrals. Regularization using linear filtering is necessary to control the propagation of noise into the reconstructed images. However, the noise is signal dependent, and filtering to achieve an optimal noise resolution trade-off is not an appropriate solution. Therefore, to solve the problem as accurately as possible, iterative refinement can be a better alternative.
2. Images reconstructed by FBP show streak artifacts as a result of the backprojection step along with the possibility of generating negative reconstruction values in regions of low count or poor tracer uptake. Both artifacts can be treated using iterative reconstruction techniques.
3. While many factors affect the PET LORs and serve to deviate the data to be approximated as line integrals when reconstructed by analytic image reconstruction, it remains an approximate reasonable approach in PET rather than SPECT [17]. Photon attenuation is an exact and straightforward procedure to implement in PET scanning, and the detector response function is not substantially degraded with source depth. In contrast, SPECT images suffer from photon attenuation in a more complicated way in addition to significant resolution loss as the source position increases.
4. The assumption of line integrals does not hold true for some imaging geometries, such as SPECT systems equipped with coded apertures and PET scanners based on hexagonal or octagonal detectors. In the former, analytical inversion of the acquired data is not a simple task and constitutes a considerable challenge, while for the latter the gaps between detector modules (e.g., C-PET and HRRT) need to be filled before applying the analytic approach. Methods to account for the missed data were therefore developed, such as linear and bilinear interpolation or constraint Fourier space gap filling [18, 19].
5. In 3D data acquisition, coincidence events are allowed to be recorded among all scanner rings; accordingly, the collected data result in direct as well as oblique sinograms. For a point source located in a scanner operating in 2D mode, the in-plane system sensitivity does not depend on source location when compared to 3D imaging. In the latter scenario, the solid angle subtended by the scanner detectors differs from one position to another, especially when the source moves in the axial direction.

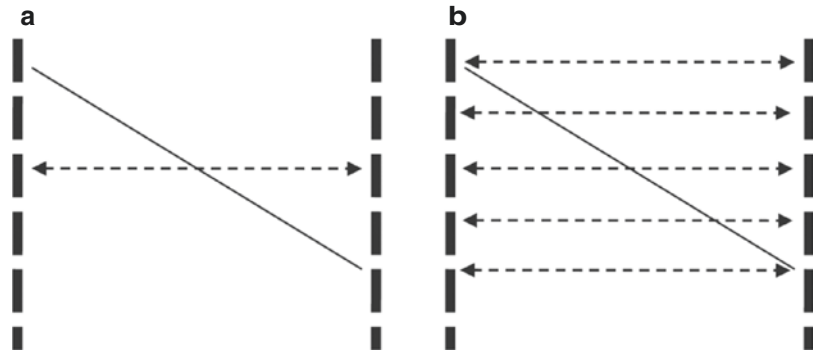
6. Another point that must be discussed is data truncation due to the fact that the axial extent of the PET scanner is limited. However, in the 3D situation, the oblique LORs are redundant in the sense that their statistical contribution to data reconstruction is unexploited. Direct 2D reconstruction uses LORs that arise from the direct planes to form an image, but this leads to compromising a lot of useful coincident events recorded as oblique LORs. The incorporation of these events into image reconstruction serves to improve the statistical quality of the scan by increasing count sensitivity. Analytic FBP with using a Colsher filter can reconstruct the oblique projection if data are not truncated [20]. In the case of data truncation, however, the missed information due to the limited axial extent of the scanner can be estimated by reconstructing the direct planes of the 2D projections (they are adequate for data reconstruction) and then reprojecting the resulting images to get an estimate of the truncated oblique projections. This method is called 3D reconstruction by reprojection (3DRP) [11]. In other words, 3DRP estimates the missed information of the oblique sinogram in the forward projection step, assuming the scanner axis is extended beyond the practical limit of data acquisition. This step is important to satisfy the requirements of (axial) data shift invariance. Image reconstruction is then carried out using 3D FBP with a 2D Colsher filter. 3DRP is computationally demanding and was extensively used as a standard analytic 3D method of choice for volumetric PET imaging.
7. The other alternative to make use of the oblique LORs is to rebin the data so that the 3D data set is reduced to a 2D problem. A number of rebinning approaches have been developed to overcome the increased reconstruction times and to utilize the count sensitivity of the scanner, yet this occurs with some drawbacks placed on spatial resolution and image noise.

16.2.3 Rebinning Methods

For many reasons, 3D PET imaging was not the acquisition mode of choice; an important one is the lack of an acceptable algorithm suited to provide clinically feasible reconstruction times. Another problem is the large amount of data that need to be processed along with extensive computational demands. An alternative way to handle this problem is to rearrange the oblique LORs into a direct array of parallel projections or a 2D data set. The latter allows for reconstruction times that are practically acceptable when compared to 3D reconstruction as the data can be reconstructed by any available 2D reconstruction algorithm. As mentioned, rebinning methods have been developed to benefit from the increased system sensitivity and to reduce computational speed requirements imposed by 3D reconstruction. Some of these rebinning approaches are summarized as follows:

1. Single-slice rebinning (SSRB) is a simple geometric approach to reduce the 3D PET data into 2D parallel sinograms [21] (Fig. 16.16a). The method is implemented by rebinning an oblique sinogram that connects a two-detector pair into a parallel sinogram that lies midway between the two detectors. Although this method can simply be applied to rearrange the 3D information into direct planes consisting of parallel sinograms, it is valid when the oblique lines are close to the center of the field of view and in systems with small aperture size.
2. The geometric simplification provided by SSRB has been refined by the multislice rebinning (MSRB) method, in which the sinograms that lie across two detectors that connect an oblique LOR are incremented as shown in Fig. 16.16b. Stated another way, for each oblique LOR, the transverse slices intersected are identified, and the corresponding sinogram is incremented. Thus, it can be viewed as a backprojection on the z -direction [22]. This process depends on the number of

Fig. 16.16 (a) Single-slice rebinning and (b) multislice rebinning



sinograms to be incremented, and the increment varies with different oblique lines. However, axial blurring and amplification of noise are the drawbacks of MSRB.

- By utilizing the properties of FT, the estimate of the FT of direct sinograms can be exactly and approximately derived in the frequency domain from the FT of the oblique sinograms using the frequency–distance relationship [23]. It is based on an acceptable equivalence between the Fourier transformed sinograms arising from direct and oblique LORs. This is called Fourier rebinning or FORE. It has significantly improved the computation time required to rearrange the 3D data sets into 2D direct sinograms with an order of magnitude gain in reconstruction times when compared to the 3DRP. FORE showed little differences compared to 3DRP, with good accuracy and stability in a noisy environment, but was less accurate in scanners with a large aperture [24, 25].
- Besides the reconstruction time gained from FORE, it can be combined with statistical iterative 2D image reconstruction techniques [26] to improve image quality when compared to FORE plus FBP or 3DRP and to exploit the incorporation of the imaging physics into the reconstruction model.
- Several studies have shown that iterative techniques have the capabilities to improve image quality and quantitative accuracy when compared to analytic techniques or hybrid approaches (rebinning + 2D reconstruction) with the drawback of increased computational

burdens. However, this has been tackled using accelerating reconstruction algorithms implemented on fast computer systems.

16.2.4 Iterative Reconstructions

The task of the reconstruction algorithm is to solve $p = Af$ to find the best estimate of f . Here, p is the measured projection data, and A is a matrix that maps the tracer activity to the projection space. The presence of image noise does not allow finding a unique solution for the problem, or the solution might not exist or might not depend continuously on the data.

The better alternative to find a solution is to perform the task in an iterative manner. In this way, an initial estimate is assumed for the reconstructed image (solution), and the image is forward projected, simulating and accounting for all possible factors that work together to form the projection data. This initial estimate or guess can be a uniform image or FBP image and can be a zero image for additive-type algorithms. Many physical factors can be handled in the projection step to produce a projection image that is a close match to the acquired projections. Then, the measured and estimated projections are compared in such a way that allows derivation of a correction term. This last step allows the algorithm to modify the reconstructed slice through what is known as image update, and the process is controlled by the cost function or the objective likelihood function, as in the maximum likelihood (ML) algorithm. It is clear that the initial estimate will be

far from the solution; thus, the process is continued by repeating the same steps to reach the best estimate of the solution: convergence. This means that the algorithm will alternate through several steps of forward- and backprojection, in contrast to direct analytic methods, for which the estimated solution is obtained through a few pre-defined steps.

Most iterative techniques share the aforementioned idea and generally differ in the objective function, the optimization algorithm, and the computation cost [17]. The combined selection of the cost function and the optimization algorithm, as underlined above, is important in optimizing the iterative reconstruction technique. Both should not be confused and are distinguished in terms of their functionality as the first denotes the governing principle or the statistical basis on which the best estimate of the solution is determined, while the latter is the “driving” tool to achieve that estimate through a number of defined steps [27].

Iterative reconstructions have the advantages of incorporating corrections for image-degrading factors in the system matrix to handle an incomplete, noisy, and dynamic data set more efficiently than analytic reconstruction techniques. An important outcome of these advantages is that the final results enjoy better qualitative features in addition to more accurate estimation of tracer concentration, improved image contrast, spatial resolution and better noise properties.

Iterative reconstruction can be statistical, such as ML or ordered subset (OS) expectation maximization (EM) algorithms, or nonstatistical, as in conventional algebraic reconstruction methods like algebraic reconstruction techniques (ARTs), steepest descent, simultaneous iterative reconstruction, and others. Another group of iterative methods based on FBP image reconstruction has also been proposed. Statistical methods can further be categorized into Gaussian or Poisson based on the noise model assumed. In Gaussian methods, the objective function can be weighted or nonweighted least square, while in Poisson-based models the objective function is the log likelihood function. The latter guarantees positivity constraint so that the pixel value is always in

the positive direction, while in the Gaussian least square model, additional requirements are needed to maintain positivity.

Another possible classification for the statistical techniques is whether they consider prior information. The inclusion of prior information in image reconstruction allows driving the reconstructed images to the desired solution using penalty terms or prior function. This can be applied when Bayes’s theorem is used in defining the objective function so that information regarding image distribution can be included in the reconstruction formula in advance. Morphological or patient anatomy, pixel smoothness, or nonnegativity constraints are different types of prior that can be used in Bayesian-based image reconstruction. The increased variance as the number of iterations increases is one of the noticeable but undesired features of statistical reconstruction techniques such as ML. Regularization using a smoothness penalty function can thus be applied to reduce image noise and to improve detectability of the reconstructed images.

16.2.4.1 System Matrix

A projection or system matrix (and also a transition matrix) is a key component in iterative techniques. It is based on the fact that the projection data are constructed by differential contributions of the object voxels being imaged. This transition from the image space to the projection space (Fig. 16.17) is the forward projection and is described in a matrix form as

$$P = Af$$

Unlike FBP, a system matrix in iterative reconstruction takes into account that each image voxel has a probability to contribute to a particular projection bin or sinogram. The system matrix A is the information reservoir that describes how the projection image is formed. It contains the coefficients a_{ij} that denote the probabilities of detecting a photon (or LOR) emitted from a particular site and detected in a particular bin.

Many physical phenomena can therefore be incorporated as far as they significantly contribute to data formation. In other words, the image space is mapped to the projection space by the

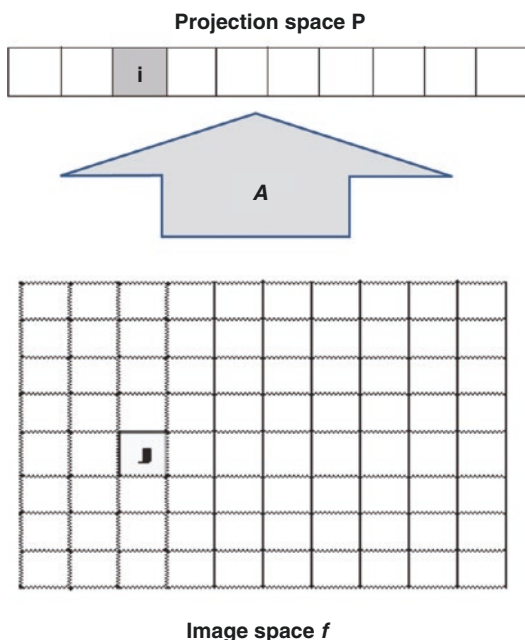


Fig. 16.17 The system matrix maps the data from the image space to the projection space

aid of the transition matrix that describes the probability of detecting a photon emitted from pixel j and measured in projection bin i such that

$$p_i = \sum_j a_{ij} f_j$$

where f is the image vector representing the activity distribution indexed by pixel j , and p is the measured projection and indexed by pixel i . A is the transition matrix of elements and is equal to $i \times j$.

However, this is not only for a one-detector row at one angle but also for all the acquired views, including all the detector elements. The situation becomes more problematic in building up a transition matrix for 3D image reconstruction when the interslice plane (3D SPECT) or oblique LOR (3D PET) is considered. Overall, the size of the system matrix is a function of the type and dimension of the data acquisition, number of detectors, number of projection angles, and size of the reconstructed image [28].

The system matrix can be structured so that it can account for the imaging physics and detector characteristics. In the context of SPECT imaging, attenuation, scatter, and detector response are

major degrading factors that can be incorporated in the iterative scheme. An accurate correction for these image-degrading elements can lead to a significant improvement in image quality and quantitative accuracy. In PET imaging, the system matrix can also be built to handle geometric components and many physical parameters of positron emission and detection. It can be decomposed into individual matrices so that each matrix can account for particular or combined physical effects [29]. The accuracy of the system matrix is essential to ensure that the sources of degrading effects are well addressed and to realize the benefits underlying the modeling procedure. Otherwise, oversimplification or inaccuracies of the system matrix would transfer the signal into noise due to inconsistencies that would arise as the estimated projection will no longer match the measured data [30, 31].

It can be calculated on the fly using efficient geometric operators, or it can be computed and stored prior to image reconstruction. Analytical derivation, Monte Carlo simulation, experimental measurements, or a combination of these techniques can be used to compute the system matrix. However, these estimation approaches vary in terms of their complexity, computational burdens, accuracy, and validity. To reduce storage capacity, the sparseness and intrinsic symmetry of the scanner is utilized to generate a compressed version of the probability matrix. Also, for efficient use of the 3D-PET matrix, it can be decomposed into individual matrices, such as geometric, attenuation, sensitivity, detector blurring, and physics of positron emission.

The inclusion of many effects that degrade image quality and contribute to image formation has expensive computational requirements. Attempts made to overcome these computational demands have been the development of accelerated image reconstruction approaches such as OSEM (ordered subset expectation maximization), the rescaled block iterative expectation maximization (RBI-EM) method [32], and the row action ML algorithm (RAMLA). Other approaches were to use an unmatched pair of projection–backprojection in the iterative scheme to accelerate the reconstruction process by not

taking into account the effect of all degrading factors in both operations [33, 34]. Efficient algorithms that include dual-matrix and variance reduction techniques have significantly reduced the processing times of Monte-Carlo-based statistical reconstructions to clinically feasible limits [34].

16.2.4.2 Maximum Likelihood Expectation Maximization

Maximum likelihood expectation maximization (MLEM) is a popular iterative reconstruction technique that gained wide acceptance in many SPECT and PET applications. The technique comprises two major steps:

1. Expectation
2. Maximization

The algorithm works to maximize the probability of the estimated slice activity given the measured projection data with the inclusion of count statistics. Stated another way, the ML algorithm seeks to find the best estimate of the reconstructed image f that with the highest likelihood can produce the acquired projection counts p . The probability function is derived from the Poisson statistics and is called the likelihood objective function:

$$L(p|f) = \text{prob}[p|f] = \prod_i e^{-q_i^k} \frac{(q_i^k)^{p_i}}{p_i!} \quad (16.2)$$

where q_i^k is the estimated forward projection data and equal to $\sum_j a_{ij} f_j^k$, while the measured projection data are represented by p_i . The ML estimate can be calculated by Eq. 16.2 but it is more convenient and easier to work with the log of the likelihood function. The selection of Poisson function is appropriate since it maintains positivity of the pixel values and agrees with the statistics of photon detection. As a result, ML reconstruction has good noise properties and is superior to FBP, especially in areas of poor count statistics. One important issue in implementing the ML algorithm is that the input data (projections/sinograms) should be matched with the

noise hypothesis of the ML model, and prior treatments or corrections for the acquired data would serve to alter the noise properties assumed by the algorithm. This can be solved by either modifying the noise model (e.g., shifted Poisson) or feeding the data directly into the iterative process without a prior correction for any of the noise-disturbing elements.

Expectation maximization is the algorithm of choice to solve the likelihood function and works to estimate the projection data from knowledge of the system matrix and the current estimate of the image. The estimated and measured projection data are then compared by taking the ratio, which in turn is used to modify the current estimate of the slice. An image update takes place by multiplying that ratio with the current estimate to get a new image estimate “update.” This process continues for several iterations until convergence is obtained and can be summarized as follows for iteration numbers k and $k + 1$:

1. The slice activity in the k th iteration is forward projected using the proposed imaging model to form a new projection image.
2. The ratio of the measured and estimated projection is calculated for each bin.
3. The result of the previous step is backprojected and normalized by dividing over the coefficients a_{ij} (see Eq. 16.3).

The new image f^{k+1} is produced by the multiplying the image in the k th iteration with the normalized backprojected data.

The equation used to define the MLEM reconstruction algorithm is [36]

$$f_j^{k+1} = \frac{f_j^k}{\sum_j a_{ij}} \left[\sum_i a_{ij} \frac{p_i}{q_i^k} \right] \quad (16.3)$$

It tells us that the $(k + 1)$ th iteration is equal to the immediate previous iteration k multiplied by a correction term. The correction term is a normalized backprojection of the ratio of the measured projection p_i and the estimated projection of the slice activity resulting from iteration k , or q_i^k .

The drawback of using the Poisson formula is that it makes the algorithm reach a solution (reconstructed image) that is statistically consistent with the proposed activity distribution of the acquired projections. The reconstructed images therefore tend to be noisy, especially at a high number of iterations. As the number of iterations increases and the algorithm approaches the solution, the log-likelihood of the function also increases but with image deterioration due to high variance estimate. This is one of the major drawbacks of the ML algorithm, which can be overcome using stopping criteria, postreconstruction smoothing filters, or regularization by Gaussian kernels: “the method of sieves” [35, 36]. This last approach is implemented by restricting the range of the optimization in least squares or ML to a subset of smooth functions on the parameter space.

Penalized likelihood and Bayesian algorithms are also applied to regularize the solution and reduce noise artifacts. In practice, however, noise reduction is accomplished mostly using postreconstruction smoothing filters. However, in analytic image reconstruction, regularization is implemented using linear filtering, compromising spatial resolution.

Convergence of the MLEM is slow, but guaranteed, and depends on the spatial frequency (object dependent) such that low-frequency regions converge faster than high-frequency regions. At a large number of iterations, however, resolution tends to be uniform across the reconstructed slice.

The second limitation of ML is the computation requirements since it converges slowly, and high-speed computer devices are needed to make it feasible in practice. However, new computer technology is continuously advancing to resolve this issue (Moore’s law). The other alternative to ML estimation is the OS algorithm, which has gained wide acceptance in many areas of research and clinical practice as it provides a significant improvement in computation time by accelerating the reconstruction process.

16.2.4.3 Ordered Subset Expectation Maximization (OSEM)

The accelerated version of the ML algorithm is the OS. This type of algorithm is also called block iterative or row action as it relies on using a single datum or subset of data at each iteration. OSEM was derived by Hudson and Larkin to speed up the iteration process [37].

The underlying concept of OSEM reconstruction is that instead of using the whole data set to obtain an update for the reconstructed image, all projection data are divided into smaller groups of projections, or subsets, and thus the image update is implemented when one subset is used; this is called *subiteration*. However, full iteration takes place when the algorithm uses all the available subsets in the image reconstruction.

The number of projections is divided equally into subsets. For example, in SPECT acquisition of 72 projections, the data set can be divided into 8 subsets, each with 9 projections. The projections in each subset are not contiguous but are spread over the whole set of angular views such that the first subset includes the projection numbers 1, 9, 18, and so on, and the second subset would have the projection numbers 2, 10, 19, and so on, and the same holds for the remaining subsets. The standard EM reconstruction of projection/back-projection is applied to each subset, one by one, so that the resulting reconstruction from subset 1 is the starting value for subset 2 and so on. In that example, a reduction of the reconstruction time by a factor of 8 can be achieved when using the OSEM technique as the rate of convergence is accelerated by a factor proportional to the number of subsets [37].

The properties of OSEM are similar to MLEM. Low-frequency regions converge faster than high-frequency regions. Thus, stopping iterations at an early stage may result in suboptimal results represented in a biased contrast; however, running a large number of iterations produces noisy images. Therefore, a trade-off between the number of iterations and detail recovery should be considered [38]. In regions of low tracer con-

centration, OSEM reconstruction might underestimate tracer activity concentration. This has been shown in a number of reports, including myocardial FDG studies and brain DatScan SPECT imaging [38, 39]. The spatially variant and object dependency convergence of iterative reconstruction is a limitation in determining the optimal number of iterations particularly with the increased noise as the iteration progresses. It is therefore of importance to optimize the reconstruction parameters, including the filtration step, given a particular detection task to exploit the full potential of the iterative technique in improving the observer performance or quantitative measurements [40].

Both 2D- and 3D-OSEM have found a number of successful applications in the reconstruction of SPECT and PET images, including corrections for many potentially degrading factors in addition to noise handling. These results have been exploited and commercialized in different software packages provided by scanner manufacturers. Attenuation-weighted OSEM reconstruction has been implemented in commercial PET scanners. Instead of precorrecting for photon attenuation before image reconstruction and presenting the data to the iterative technique in a Poisson-corrupted form, attenuation correction factors can be included in the system matrix to yield images with less noise and superior quality than data precorrected for attenuation. Not only attenuation but also other degrading factors, such as system response, has been incorporated into iterative OSEM and resulted in remarkable improvement of PET image quality and spatial resolution [41]. Also, it has become evident that including all corrections starting from random, dead time, normalization, geometric scatter, attenuation, and arc correction (the problem of unevenly spaced acquired projections) in the system matrix of iterative reconstruction allows preservation of the statistical nature of the raw data and satisfies the Poisson likelihood function of OSEM or MLEM, yielding an image with superior noise properties [42, 43].

16.2.4.4 Maximum A Posteriori

Maximum a posteriori (MAP) is a Bayesian reconstruction method that found several applications in SPECT and PET imaging [44]. It has a superior performance over analytic image reconstruction, especially when image-degrading factors are taken into account [45, 46]. However, in contrast to the ML mentioned here, MAP reconstruction uses prior knowledge to force the solution in the preferred or desired direction. According to Bayes's theorem, the probability of estimating an image provided the measured projection data is given by the posterior density function

$$\text{prob}[f|p] = \frac{\text{prob}[P|f]\text{prob}[f]}{\text{prob}[p]}$$

The first term of the nominator refers to the likelihood, while the second term denotes the distribution of the prior. The denominator is a constant (not a function of f) and can be dropped [40]. Note that in ML no preferences are placed on the reconstructed image; therefore, the objective function returns to the ML form once no information about the prior is assumed. The property given by MAP reconstruction to incorporate prior knowledge in the iterative procedure allows the associated noise elevation to be overcome as the number of iterations increases, as mentioned. This is implemented by penalizing the likelihood function by a prior term, driving the log-likelihood to the favored solution. The prior function is often selected to smooth the reconstructed images; however, this occurs with drawbacks of blurring sharp edges. Functions designed to smooth the image while being able to preserve edges have also been suggested. Another type of prior attempts to utilize morphological information provided by anatomical imaging modalities such as CT and MRI and based on the assumption that tracer uptake within a given structure or organ is uniformly distributed. However, using MAP reconstruction with anatomical priors has a number of limitations that, if

properly addressed, could significantly improve lesion detectability and image quality.

One of the commonly used is Gibbs distribution prior, which penalizes a given pixel based on differences with the neighboring pixels. It has the following mathematical representation:

$$P(x) = \frac{1}{Z} e^{-\beta U(x)}$$

$$U(x) = \frac{1}{2} \sum_{j=1}^N \sum_{k \in N_j} \psi(x_j - x_k)$$

where Z is a normalization constant, and β is a weighting parameter that determines the strength of the prior. $U(x)$ is the energy function and often contains potentials, $U(\cdot)$, defined on a pairwise cliques of neighboring pixels [46]. Prior functions based on absolute pixel differences have been devised as well as functions that use relative pixel differences. It is the selection and design of the potential function that allows penalization of the reconstructed images in favor of smoothing the images or preserving sharp edges, and this is implemented by increasing or decreasing the probability of the desired solution [47]. In the same vein, MAP-based reconstruction techniques produce an image with complex and object-dependent spatial resolution; this again can be controlled by the prior function. A nonuniform spatial resolution is obtained if a shift-invariant prior is used, whereas a uniform resolution comparable to postsmoothed ML (with a sufficient number of iterations) can be achieved with appropriate tuning of the prior [48, 49].

The availability of multimodality imaging devices such as SPECT/CT, PET/CT, and PET/MRI allows the introduction of morphological information in the iterative algorithm and thus has the potential to improve the quality of the diagnostic images. However, some problems could arise, such as image coregistration errors, identification of lesion location within the anatomical structures or segmentation errors, addition of lesion or organ boundaries or both, and selection of the penalty function and optimal prior strength [50]; ultimately, research efforts need to optimize the technique and prove an

improved diagnostic confidence over other methods that do not rely on prior information. Furthermore, an underutilized application of MAP-type reconstruction is the unexploited feature of incorporating an anatomical prior to correct for the partial volume effect. There is an interest in improving the spatial resolution of PET images using resolution recovery approaches; however, investigations could also be directed to make use of the anatomical data provided by CT or MRI images to formulate feasible correction schemes in multimodality imaging practice [50–52].

16.2.5 Time of Flight Reconstruction

Time of flight as explained in Chap. 13 has received a renewed interest after its initial inception in 1980s. Initial results presented due to TOF image reconstruction in commercial LSO scanner have shown measurable gain in signal-to-noise ratio despite the use of relatively poor timing resolution of 1.2 ns [53]. Once the timing resolution of the PET scanner is very short such that the time arrival of the individual photons of the coincidences can be precisely measured, the theory of image reconstruction of the PET data is no longer required as photons “point” of annihilation can be determined. The current clinical PET systems still have not reached that goal and hence image reconstruction is required while having TOF data as additional piece of information to the reconstruction algorithm. PET data acquired with 3D mode are four dimensional and the additional TOF information increases data sparsity (a matrix with many zero elements) [54]. A Gaussian distribution function, the kernel, is then given instead of using the whole line of response in estimating the location of positron annihilation. The Gaussian is determined primarily by the system timing resolution (denoted by full width at half-maximum, FWHM) using the time-distance relationship $\Delta x = c\Delta\tau/2$ where c is the speed of light.

List-mode data format is an efficient method for storing uncompressed PET containing TOF information. While being very slow in image

reconstruction due to handling the coincidences on an event-by-event basis, some advantages are obtained for the sake of image quality. Full utilization of the 3D PET data including TOF information and plugging into reconstruction without data reduction or rebinning could provide an improved image characteristics including uniform spatial resolution and noise-contrast trade-off [55, 56]. However, this process is computationally intensive and computer clusters are used in commercial TOF systems. A factor that may enhance the computation time and speed up the process is kernel truncation but inaccurate assignment of the TOF kernel could result in image quality deterioration [57]. In Philips GEMINI TOF, for example, the optimal kernel width was found less critical for the recovered contrast but influential on the background uniformity. Moreover, a smaller or wider kernels yielded less uniform background and reduced contrast recovery [58].

The 3D TOF data can also be transformed into different format including TOF as well as non-TOF in the 2D and 3D domains. This comes at the expense of increased image noise in the direction of going from 3D to 2D and also from TOF to non-TOF [59]. However, data rebinning via optimal weightings may have better variance and contrast recovery assessments in contrast to data without optimal weightings [60]. Furthermore, data rebinning into non-TOF sinograms retains significant signal to noise ratio over sinograms collected in absence of TOF information [61].

SSRE can also be implemented for TOF data in a similar fashion described for data acquired without TOF information [21, 62]. Fourier rebinning mapping in frequency space or native coordinate were also devised to reduce the size and dimensionality of the 3D TOF data into 2D data set [56, 63]. Fast accelerating methods employing graphics processing units (GPUs) using the compute unified device architecture (CUDA) framework was also presented to reconstruct TOF list-mode data set [54].

Earlier methods of TOF data reconstruction was the analytic methods [53, 64]. As described above, the analytic TOF reconstruction is carried out through backprojection of the sinogram data using confidence weighting function that utilize

the uncertainty of the time resolution window of the system and then an inverse filter is employed to reconstruct the activity distribution in the image space. Various TOF reconstruction filters were proposed such as the most likely position (MLP), confidence weighting (CW), transverse ramp (TR), convolved ramp and Gaussian and others. The confidence weighting was shown to have minimal noise variance when having Poisson data derived from infinite uniform source distribution [64].

While analytic methods can be utilized to reconstruct the TOF PET data providing more speed and consistent quantification, model based statistical methods such as ML (or OSEM) are the most commonly used [65]. Direct reconstruction of the list-mode data is computationally demanding and used in some clinical systems but rebinning or transverse mashing methods could serve in data reduction and time saving.

16.2.6 Machine Learning

The topic of machine learning has been discussed in more than on instance in this textbook. There are several reasons behind this interest among which is the successful implementation of machine and deep learning in several aspects of radiology and nuclear medicine including image processing, classification, segmentation, super-resolution, and denoising, and many other disciplines related to disease detection, characterization and monitoring [66–68]. However, there are continued interests in improving network performance and more applications are emerging. Image reconstruction using convolutional neural networks or deep learning has been reported in several reports. DeepPET is a convolutional network devised as end-to-end encoder-decoder PET image reconstruction technique [69]. The method reconstructs the PET image from the sinogram data with high quality and quantitative accuracy. Initial results showed better relative error, peak signal to noise ratio, structural similarity index, and faster performance than iterative and analytic image reconstruction. On the SPECT side, a specialized method called SPECTNet was developed such that it split the

deep network into two subsystems and trained them separately; thus, avoiding training difficulty [70]. The projection space was mapped and compressed to a low-dimensional space in the image domain and then the compressed image was upsampled to the original dimension. More accurate images were obtained with less sensitivity to noise. Another approach was used to utilize the convolutional neural network in deriving SPECT images with quality comparable to Monte-Carlo-based image reconstruction but in a faster rate of processing [71]. In a similar manner, it was also demonstrated that deep learning reconstruction of scattered data in Y90 studies could provide comparable performance to Monte-Carlo based scatter estimates in the context of patient dosimetry and safety [72]. The merits achieved are accelerated image reconstruction by orders of magnitude faster than Monte-Carlo approach while able to maintain high accuracy.

Image noise in tomographic PET and SPECT is one of the most annoying factors in image reconstruction. Neural network may be trained on a predetermined noise level but this prior may lose generalizability due to noise sparsity in training or testing data and introduce additional bias if not properly treated. An approach to incorporate a local linear fitting function with denoising convolutional network was reported to robust versus noise level disparities while the network was trained with a predetermined noise level. A better quantitative and qualitative results were obtained in comparison to conventional methods [73]. The future of machine and deep learning in tomographic image reconstruction looks promising and would be able to overcome many of the current limitations providing a significant improvement in image quality, quantitative accuracy, and diagnostic performance.

16.3 Conclusions

Image reconstruction is a key element in conveying the diagnostic information given an activity distribution within different tissues. Analytic approaches are simple, fast, and easy to imple-

ment in research and clinical practice. However, they have some drawbacks that can be eliminated using iterative techniques. These provide improved image quality and quantitative accuracy, with some efforts to be done on optimizing the reconstruction parameters given a particular detection task. The system matrix of iterative reconstruction can be considered an information reservoir that allows the technique to reach the most accurate solution and thus should be optimally constructed. The future of machine and deep learning in tomographic image reconstruction looks promising and would be able to overcome many of the current limitations providing a significant improvement in image quality, quantitative accuracy, and diagnostic performance.

References

1. Even-Sapir E. Imaging of malignant bone involvement by morphologic, scintigraphic, and hybrid modalities. *J Nucl Med.* 2005;46(8):1356–67.
2. Bracewell RN. Strip integration in radio astronomy. *Aust J Phys.* 1956;9:198–217.
3. Beckmann EC. CT scanning the early days. *Br J Radiol.* 2006;79(937):5–8.
4. Jaszczak RJ. The early years of single photon emission computed tomography (SPECT): an anthology of selected reminiscences. *Phys Med Biol.* 2006;51(13):R99–115.
5. Li G, Citrin D, Camphausen K, Mueller B, Burman C, Mychalczak B, Miller RW, Song Y. Advances in 4D medical imaging and 4D radiation therapy. *Technol Cancer Res Treat.* 2008;7(1):67–81.
6. Noo F, Wagner JM. Image reconstruction in 2D SPECT with 180° acquisition. *Inverse Probl.* 2001;17:1357–71.
7. Hsieh J. Computed tomography: principles, design, artifacts, and recent advances. In: SPIE. International society for optical engineering, Bellingham; 2003.
8. Zanzonico P. Routine quality control of clinical nuclear medicine instrumentation: a brief review. *J Nucl Med.* 2008;49(7):1114–31.
9. Zeng GL, Gullberg GT. Unmatched projector/back-projector pairs in an iterative reconstruction algorithm. *IEEE Trans Med Imaging.* 2000;19(5):548–55.
10. Brooks RA, Di Chiro G. Principles of computer assisted tomography (CAT) in radiographic and radioisotopic imaging. *Phys Med Biol.* 1976;21(5):689–732.
11. Kinahan PE, Defrise M, Clackdoyle R. Analytic image reconstruction methods. In: Wernick M, Aarsvold J, editors. *Emission tomography: the funda-*

- mentals of PET and SPECT. San Diego: Academic; 2004.
12. Matej S, Kazantsev IG. Fourier-based reconstruction for fully 3-D PET: optimization of interpolation parameters. *IEEE Trans Med Imaging*. 2006;25(7):845–54.
 13. Madsen MT, Park CH. Enhancement of SPECT images by Fourier filtering the projection image set. *J Nucl Med*. 1985;26(4):395–402.
 14. National Electrical Manufacturers Association. Performance measurements of positron emission tomographs (PET). Rosslyn: NEMA Standards Publication NU 2-2012; 2012.
 15. Wells RG, Farncombe T, Chang E, Nicholson RL. Reducing bladder artifacts in clinical pelvic SPECT images. *J Nucl Med*. 2004;45(8):1309–14.
 16. Gilland DR, Tsui BMW, McCartney WH, Perry JR, Berg J. Determination of the optimum filter function for SPECT imaging. *J Nucl Med*. 1988;29(5):643–50.
 17. Qi J, Leahy RM. Iterative reconstruction techniques in emission computed tomography. *Phys Med Biol*. 2006;51(15):R541–78.
 18. van Velden FH, Kloet RW, van Berckel BN, Molthoff CF, Lammertsma AA, Boellaard R. Gap filling strategies for 3-D-FBP reconstructions of high-resolution research tomography scans. *IEEE Trans Med Imaging*. 2008;27(7):934–42.
 19. Karp JSMG, Lewitt RM. Constrained Fourier space method for compensation of missing data in emission computed tomography. *IEEE Trans Med Imaging*. 1988;7(1):21–5.
 20. Colsher JG. Fully three-dimensional positron emission tomography. *Phys Med Biol*. 1980;25:103.
 21. Daube-Witherspoon ME, Muehllehner G. Treatment of axial data in three-dimensional PET. *J Nucl Med*. 1987;28(11):1717–24.
 22. Lewitt RM, Muehllehner G, Karp JS. Three-dimensional reconstruction for PET by multi-slice rebinning and axial image filtering. *Phys Med Biol*. 1994;39:321–40.
 23. Defrise M, Kinahan PE, Townsend DW, Michel C, Sibomana M, Newport DF. Exact and approximate rebinning algorithms for 3D PET data. *IEEE Trans Med Imaging*. 1997;MI-16:145–58.
 24. Matej S, Karp JS, Lewitt RM, Becher AJ. Performance of the Fourier rebinning algorithm for 3D PET with large acceptance angles. *Phys Med Biol*. 1998;43:787–97.
 25. Krzywinski M, Sossi V, Ruth TJ, Nuclear Science Symposium CRI. Comparison of FORE, OSEM and SAGE algorithms to 3DRP in 3D PET using phantom and human subject data. In: *IEEE nuclear science symposium, conference record*, 1998, vol. 3, 1998. p. 1546–51.
 26. Kinahan PE, Michel C, Defrise M, Townsend DW, Sibomana M, Lonneux M, Newport DF, Luketich JD. Fast iterative image reconstruction of 3D PET data. In: *IEEE nuclear science and medical imaging conference*, Anaheim, 1996. p. 1918–22.
 27. Lalush DS, Wernick MN. Iterative image reconstruction. In: Wernick M, Aarsvold J, editors. *Emission tomography: the fundamentals of PET and SPECT*. San Diego: Academic Press; 2004.
 28. Loudos GK. An efficient analytical calculation of probability matrix in 2D SPECT. *Comput Med Imaging Graph*. 2008;32:83.
 29. Qi J, Leahy R, Cherry S, Chatzioannou A, Farquhar T. High-resolution 3-D Bayesian image reconstruction using the microPET small-animal scanner. *Phys Med Biol*. 1998;43:1001–13.
 30. Qi J, Huesman RH. Effect of errors in the system matrix on maximum a posteriori image reconstruction. *Phys Med Biol*. 2005;50(14):3297–312.
 31. Ortuno J, Pedro Guerra-Gutierrez P, Rubio J, Kontaxakis G, Santos A. 3D-OSEM iterative image reconstruction for high-resolution PET using precalculated system matrix. *Nucl Instrum Meth Phys Res A*. 2006;569:440–4.
 32. Byrne CL. Block-iterative methods for image reconstruction from projections. *IEEE Trans Imaging Process*. 1996;5:792.
 33. Kamphuis C, Beekman F, Van Rijk PP, Viergever MA. Dual matrix ordered subsets reconstruction for accelerated 3D scatter compensation in single-photon emission tomography. *Eur J Nucl Med*. 1998;25:8–18.
 34. De Wit TC, Xiao J, Beekman FJ. Monte Carlo-based statistical SPECT reconstruction: influence of number of photon tracks. *IEEE Trans Nucl Sci*. 2005;52:1365–9.
 35. Snyder DL, Miller MI. The use of sieves to stabilize images produced with the EM algorithm for emission tomography. *IEEE Trans Nucl Sci*. 1985;NS-32:3864–72.
 36. Snyder DL, Miller MI, Thomas LJ, Polite DG. Noise and edge artifacts in maximum-likelihood reconstructions for emission tomography. *IEEE Trans Med Imaging*. 1987;6(3):228–38.
 37. Hudson H, Larkin R. Accelerated image reconstruction using ordered subsets of projection data. *IEEE Trans Med Imaging*. 1994;13:601.
 38. Seret A, Boellaard R, van der Weerd A. Number of iterations when comparing MLEM/OSEM with FBP. *J Nucl Med*. 2004;45(12):2125–6.
 39. Dickson JC, Tossici-Bolt L, Sera T, Erlnadsson K, Tatsch K, Hutton B. The impact of reconstruction method on the quantification of DaTSCAN images. *Eur J Nucl Med Mol Imaging*. 2010;37(1):23–35.
 40. Hutton B, Nuyts J, Zaidi H. Iterative image reconstruction methods. In: Zaidi H, editor. *Quantitative analysis in nuclear medicine imaging*. New York: Kluwer/Plenum; 2004.
 41. Panin VY, Kehren F, Michel C, Casey M. Fully 3-D PET reconstruction with system matrix derived from point source measurements. *IEEE Trans Med Imaging*. 2006;25:907–21.
 42. Kadmas DJ. LOR-OSEM: statistical PET reconstruction from raw line-of-response histograms. *Phys Med Biol*. 2004;49:4731.

43. Comtat C, Kinahan PE, Defrise M, Michel C, Townsend DW. Fast reconstruction of 3D PET data with accurate statistical modeling. *IEEE Trans Nucl Sci.* 1998;45:1083–9.
44. Levitan E, Herman GT. A maximum a posteriori probability expectation maximization algorithm for image reconstruction in emission tomography. *IEEE Trans Med Imaging.* 1987;6(3):185–92.
45. Frese T, Rouze NC, Bouman CA, Sauer K, Hutchins GD. Quantitative comparison of FBP, EM, and Bayesian reconstruction algorithms for the IndyPET scanner. *IEEE Trans Med Imaging.* 2003;22(2):258–76.
46. Qi J. Theoretical evaluation of the detectability of random lesions in Bayesian emission reconstruction. *Inf Process Med Imaging.* 2003;2732:354–65.
47. Qi J. Analysis of lesion detectability in Bayesian emission reconstruction with nonstationary object variability. *IEEE Trans Med Imaging.* 2004;23:321.
48. Fessler JA. Penalized weighted least squares image reconstruction for PET. *IEEE Trans Med Imaging.* 1994;13:290.
49. Fessler JA, Rogers WL. Spatial resolution properties of penalized-likelihood image reconstruction: spatialinvariant tomographs. *IEEE Trans Image Process.* 1996;5(9):1346–58.
50. Lehovich A, Gifford HC, Schneider PB, King MA. Choosing anatomical-prior strength for MAP SPECT reconstruction to maximize lesion detectability. *IEEE Nucl Sci Symp Conf Rec.* 2007;6(1):4222–5.
51. Baete KNJ, Van Laere K, Van Paesschen W, Ceyskens S, De Ceuninck L, Gheysens O, Kelles A, Van den Eynden J, Suetsens P, Dupont P. Evaluation of anatomy based reconstruction for partial volume correction in brain FDG-PET. *NeuroImage.* 2004;23(1):3.
52. Alessio AM, Kinahan PE. Improved quantitation for PET/CT image reconstruction with system modeling and anatomical priors. *Med Phys.* 2006;33(11):4095–103.
53. Conti M, Bendriem B, Casey M, Chen M, Kehren F, Michel C, et al. First experimental results of time-of-flight reconstruction on an LSO PET scanner. *Phys Med Biol.* 2005;50(19):4507–26.
54. Prax G, Surti S, Levin C. Fast list-mode reconstruction for time-of-flight PET using graphics hardware. *IEEE Trans Nucl Sci.* 2011;58(1):105–9.
55. Defrise M, Panin V, Michel C, Casey ME. Continuous and discrete data rebinning in time-of-flight PET. *IEEE Trans Med Imaging.* 2008;27(9):1310–22.
56. Vandenberghe S, Daube-Witherspoon ME, Lewitt RM, Karp JS. Fast reconstruction of 3D time-of-flight PET data by axial rebinning and transverse mashing. *Phys Med Biol.* 2006;51(6):1603–21.
57. Daube-Witherspoon ME, Surti S, Matej M, Werner S, et al. Influence of time-of-flight kernel accuracy in TOF-PET reconstruction. In: *IEEE nuclear science symposium conference record, San Diego, CA, 2006.* p. 1723–7.
58. Vandenberghe S, van Elmbt L, Guerschaf M, Clementel E, Verhaeghe J, Bol A, et al. Optimization of time-of-flight reconstruction on Philips GEMINI TF. *Eur J Nucl Med Mol Imaging.* 2009;36(12):1994–2001.
59. Bai B, Asma E. PET image reconstruction: methodology and quantitative accuracy. In: Khalil MM, editor. *Basic science of PET imaging.* Berlin: Springer; 2016. p. 259–84.
60. Ahn S, Cho S, Li Q, Lin Y, Leahy RM. Optimal rebinning of time-of-flight PET data. *IEEE Trans Med Imaging.* 2011;30(10):1808–18.
61. Cho S, Ahn S, Li Q, Leahy RM. Exact and approximate Fourier rebinning of PET data from time-of-flight to non time-of-flight. *Phys Med Biol.* 2009;54(3):467–84.
62. Mullani N, Wong W, Hartz P, et al. Sensitivity improvement of TOFPET by the utilization of the inter-slice coincidences. *IEEE Trans Nucl Sci.* 1982;29:479–83.
63. Defrise M, Casey ME, Michel C, Conti M. Fourier rebinning of time-of-flight PET data. *Phys Med Biol.* 2005;50:2749–63.
64. Tomitani T. Image reconstruction and noise evaluation in photon time-of-flight assisted positron emission tomography. *IEEE Trans Nucl Sci.* 1981;NS-28(6):4582–8.
65. Vandenberghe S, Mikhaylova E, D'Hoe E, Mollet P, Karp JS. Recent developments in time-of-flight PET. *EJNMMI Phys.* 2016;3(1):3.
66. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts H. Artificial intelligence in radiology. *Nat Rev Cancer.* 2018;18(8):500–10.
67. Choi H. Deep learning in nuclear medicine and molecular imaging: current perspectives and future directions. *Nucl Med Mol Imaging.* 2018;52(2):109–18.
68. Veit-Haibach P, Buvat I, Herrmann K. *EJNMMI supplement: bringing AI and radiomics to nuclear medicine.* *Eur J Nucl Med Mol Imaging.* 2019;46(13):2627–9.
69. Haggstrom I, Schmidlein CR, Campanella G, Fuchs TJ. DeepPET: a deep encoder-decoder network for directly solving the PET image reconstruction inverse problem. *Med Image Anal.* 2019;54:253–62.
70. Shao W, Du Y. SPECT image reconstruction by deep learning using a two-step training method. *Med Image Anal.* 2019;60(Suppl 1):1353.
71. Dietze MMA, Branderhorst W, Kunnen B, Viergever MA, de Jong H. Accelerated SPECT image reconstruction with FBP and an image enhancement convolutional neural network. *EJNMMI Phys.* 2019;6(1):14.
72. Xiang H, Lim H, Fessler JA, Dewaraja YK. A deep neural network for fast and accurate scatter estimation in quantitative SPECT/CT under challenging scatter conditions. *Eur J Nucl Med Mol Imaging.* 2020;47(13):2956–67.
73. Kim K, Wu D, Gong K, Dutta J, Kim JH, Son YD, et al. Penalized PET reconstruction using deep learning prior and local linear fitting. *IEEE Trans Med Imaging.* 2018;37(6):1478–87.