
AN OVERVIEW OF SPECT IMAGING

Junyu Chen
jchen245@jhmi.edu
Johns Hopkins University

June 8, 2021

Contents

1 Modes of Radioactive Decay	4
2 Interaction of Radiation with Matter	4
3 Fundamentals of SPECT imaging system	5
3.1 Collimator	5
3.2 Scintillation crystal	5
3.3 Light guide	6
3.4 Photomultiplier tubes	6
3.5 Pre-amplifiers and amplifiers	6
3.6 Positioning logic and pulse-height analyzer	6
4 SPECT system resolution	6
4.1 Intrinsic resolution	7
4.2 Collimator resolution	8
5 SPECT energy resolution	9
6 SPECT system sensitivity	10
7 Spatial, energy, and sensitivity corrections	11
7.1 Energy correction	11
7.2 Spatial correction	11
7.3 Sensitivity correction	11
8 Image quality for detection tasks	11
8.1 Contrast	11
8.2 Noise	12
8.2.1 Random noise	12
8.2.1.1 Coefficient of variance	12
8.2.1.2 Contrast-to-noise rate	12
8.2.2 Structured noise	13
8.3 Observer Performance Studies	13
9 Quantitative analysis in SPECT	14
9.1 Quantitative tasks	14
9.2 Figure of merit in SPECT quantification	14
9.2.1 Bias and relative error	14
9.2.2 Precision	15
9.2.3 Root mean square error	15
9.3 Factors that affect SPECT quantification	15

10 SPECT image reconstruction	16
10.1 Forward projection	16
10.2 Simple backprojection	17
10.3 Projection-slice theorem	18
10.4 The Fourier method of image reconstruction	18
10.5 Filtered backprojection	19
10.6 Image reconstruction in a matrix view	20
10.7 Iterative reconstruction methods	20
10.7.1 Maximum-likelihood Criterion	20
10.7.2 Least-Squares and Weighted-Least-Squares	21
10.7.3 The Maximum-Likelihood Expectation-Maximization algorithm	21
10.7.4 Shortcoming of Maximum-Likelihood, Least-Squares, and Weighted-Least-Squares	23
10.7.5 Bayesian methods	23
10.7.6 The Ordered-Subset Expectation-Maximization algorithm	24
10.7.7 Important remarks	24
11 Attenuation, scatter, and collimator-detector response compensation in SPECT	25
11.1 Attenuation compensation	25
11.2 Scatter estimation	25
11.3 Collimator-detector response	26

1 Modes of Radioactive Decay

There are four main modes of decay:

- **Alpha decay** - results in emission of an alpha particle.
- **Beta decay** - results in emission of a beta particle.
- **Positron decay** - results in emission of a positron. A positron is an antimatter electron. After a positron is emitted, it travels a short distance (a.k.a., *positron range*) then meets a free electron in the tissue. The mutual annihilation occurs from the combination of the positron and the electron. Two 511 KeV photons are then emitted 180° back-to-back.
- **Isometric transition** - A radionuclide may decay to a more stable nuclide that has the same atomic and mass numbers. The excess energy is released in the form of gamma rays.

PET imaging system relies on Positron decay, whereas SPECT imaging system relies on isometric transition. Radioactive decay is a random process and follows the Poisson distribution:

$$Pr[N = k] = \frac{a^k e^{-a}}{k!}, \quad (1)$$

where a denotes the mean, and for Poisson distribution the variance, $\sigma^2 \equiv a$.

2 Interaction of Radiation with Matter

Gamma photon can pass through matter (i.e., tissues, water, etc.) without loss of energy and change of direction, or it can be completely absorbed by the matter, or it can scatter results in a change of direction of travel path. Formally, there are three ways that a gamma photon can interact with matter:

- **Photoelectric effect** - it is an atomic absorption process in which an atom completely absorbs the energy of an incident gamma photon. The absorbed energy results in the ejection of a photoelectron, and later the vacancy in an orbital electron shell leads to the emission of characteristic x rays. Fig. 1a illustrated the idea of photoelectric effect.
- **Compton scattering** - it is a "collision" between a photon and a loosely bound outer-shell orbital electron of an atom. Fig. 1b illustrated the idea of Compton scattering. The gamma photon does not disappear in Compton scattering. Instead, it is deflected through a scattering angle θ with a loss of its energy:

$$E_f = \frac{E_0}{1 + \frac{E_0}{m_e c^2} (1 - \cos \theta)}. \quad (2)$$

- **Coherent scattering** - because the incident photon does not have enough energy to liberate the electron from its bound state so no energy transfer occurs. The incident photon is therefore deflected with no loss of energy (but the direction of traveling does change). Coherent scattering happens at relatively low energies ($\ll 50$ KeV).

Fig. 2 shows graphs of Photoelectric, Compton, pair-production, and total mass attenuation coefficients in water and NaI. We can observe that pair-product only occurs at high energies. A beam of photons is not degraded in energy as it passes through matter, but due to scattering and photoelectric effect, some photons from the beam are removed entirely through scattering or absorption. The total number of photons is reduced by the number that have interacted:

$$I = I_0 e^{-\mu x}, \quad (3)$$

where I_0 is the nonattenuated incident beam intensity, x is the thickness of the matter, and μ is the attenuation coefficient.

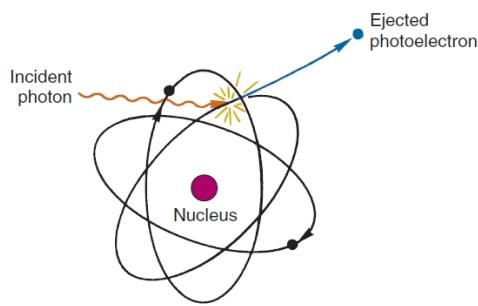


FIGURE 6-11 Schematic representation of the photoelectric effect. The incident photon transfers its energy to a photoelectron and disappears.

(a) Photoelectric effect.

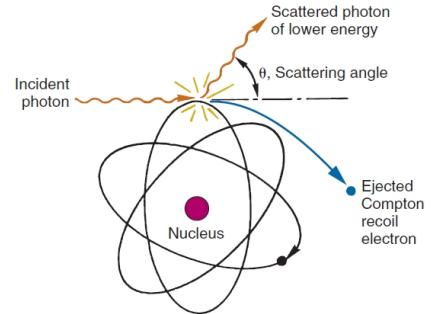


FIGURE 6-12 Schematic representation of Compton scattering. The incident photon transfers part of its energy to a Compton recoil electron and is scattered in another direction of travel (θ , scattering angle).

(b) Compton scattering.

Figure 1: Photoelectric effect and Compton scattering. Images were obtained from [1].

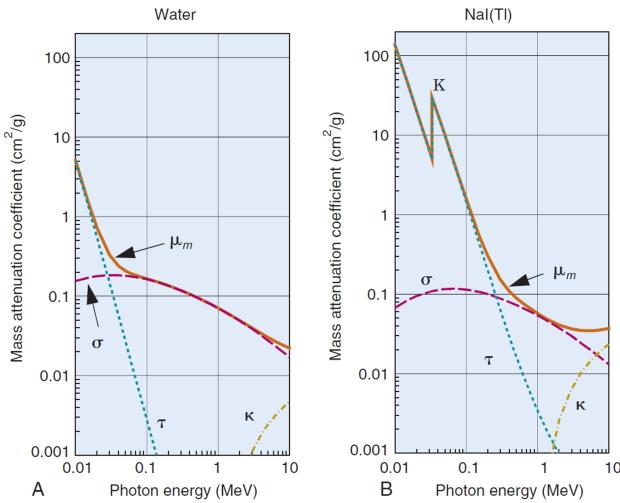


Figure 2: Photoelectric τ , Compton σ , pair-production κ , and total mass attenuation coefficients μ for water and NaI. Images were obtained from [1].

3 Fundamentals of SPECT imaging system

Fig. 3a shows the main components inside a SPECT camera (i.e., Anger camera). It consists of five major parts: 1. Collimator 2. Scintillation crystal 3. PMTs 4. Pre-amplifiers and amplifiers 5. Positional logic and pulse height analyzer.

3.1 Collimator

A basic collimator geometry is shown in Fig. 3b. A collimator is used to reject photons that are not traveling in the desired direction. For a parallel hole collimator, that direction would be 90 degrees to the face of the detector. The resolution of a SPECT system is primarily limited by the collimator resolution, which is described in more detail in section 4.2. The collimator should be made of high z (i.e., atomic number) materials so that the gamma photons cannot penetrate through the septa. Septa penetration can have an impact on image contrast, which is described in section 8.1.

3.2 Scintillation crystal

A scintillation crystal (it is typically made of NaI(Tl)) is used to convert an incident gamma photon into many scintillation photons (light photons) via a series of photoelectric effects. Some gamma photons may pass through the

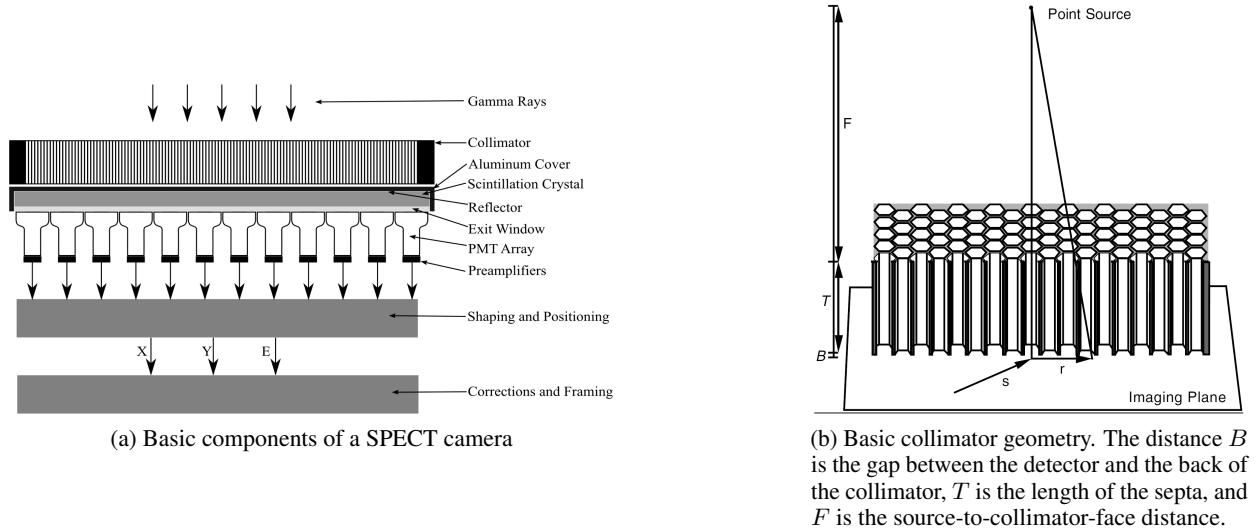


Figure 3: A typical SPECT camera (left), and a collimator (right). Collimator image was obtained from [2].

detector without interaction, some will scatter via Compton effect (see left panel of Fig. 2), and some will interact with the detector and produce scintillation photons. The crystal should have a high z to absorb the majority of the incident gamma photons efficiently. It should also be transparent enough or have a high light output so that the converted scintillation photons can easily go through the crystal with minimum quantum noise. The ideal crystal should also have a fast rise time and decay time for the light emission to avoid pulse pileup. The total light emitted is proportional to the energy deposited in the crystal. Therefore, the total light emitted is used to estimate the energy of the gamma-ray.

3.3 Light guide

Light guide guides the scintillation photons coming from the scintillator into photomultiplier tubes (PMTs).

3.4 Photomultiplier tubes

PMTs are used to convert scintillation light photons into electrical signals. Basic structures of a PMT are shown in Fig. 4a. It consists of three main parts: 1. Photocathode 2. Several dynodes 3. An anode. The photocathode converts light photons into electrons via the photoelectric effect. The dynodes are used to amplify the electrons ejected from the photocathode. Each dynode has a higher voltage than the previous one so that the emitted electrons are multiplied.

3.5 Pre-amplifiers and amplifiers

The pre-amplifier integrates the signal from PMT, so the height of output voltage is proportional to input charge (as shown in Fig. 4b). Ideally, output pulses should be narrow and flat on top to allow easy digitization of the pulse and avoid pulse pileup. However, the pre-amplifier generally has a long decay time, so it is not useful for handling high count rates. Then the amplifier is used to shape the signal so that it is easier to measure the voltage and to handle the high count rates. The sum of the voltages is proportional to the gamma camera energy.

3.6 Positioning logic and pulse-height analyzer

In the end, we have the positioning logic and the pulse-height analyzer. Positioning logic is used to determine the position of where the incident gamma photon interacts in the crystal. A pulse-height analyzer is used to determine and reject scattered photons or photons with energy different from the energy of interest.

4 SPECT system resolution

The resolution of a SPECT system is about 10mm [3]. The system resolution of SPECT has two components: 1. Intrinsic resolution 2. Collimator resolution. *The system resolution of SPECT is primarily limited by the collimator*

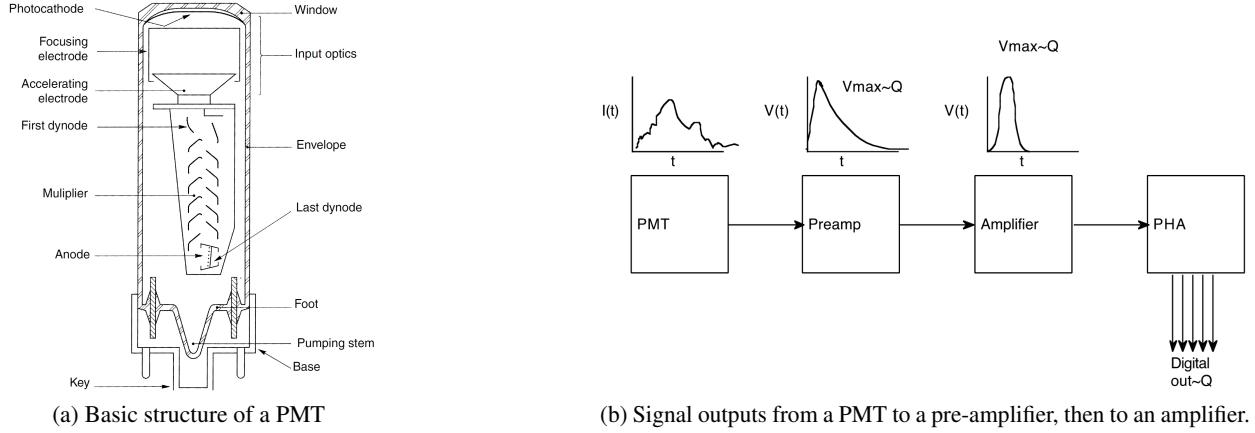


Figure 4: A typical PMT (left). Signal processing of a PMT output (right). PMT image was obtained from [2].

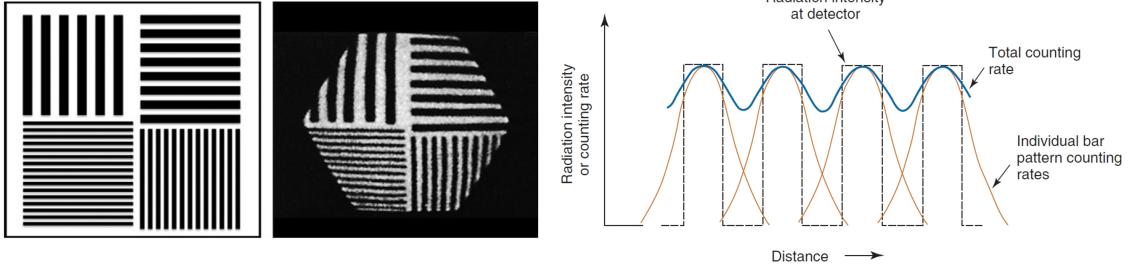


Figure 5: Left: The design and a SPECT image of a bar phantom. Right: Counting-rate profiles obtained on a bar phantom. Images were obtained from [1].

resolution. Resolution in SPECT can be measured objectively or subjectively. In the subjective way, an organ phantom is imaged, then some human observers (e.g., physicians) perform visual inspection on the resulting images. A more objective way to quantify the resolution is to use a bar phantom (as shown in Fig. 5), where one can visually tell what the minimum resolvable bar width is. The resolution can also be quantitatively measured using the full-width at half maximum (FWHM) and the modulation transfer function (MTF). The overall system resolution can be characterized by the FWHM of a point-spread function (PSF) by imaging a point source. The SPECT system's FWHM can be mathematically written as:

$$\text{FWHM} = \sqrt{\text{FWHM}_{\text{int}}^2 + \text{FWHM}_{\text{col}}^2}, \quad (4)$$

where FWHM_{int} and FWHM_{col} denote, respectively, the intrinsic FWHM and the collimator FWHM. The system resolution can also be evaluated in frequency domain via MTF:

$$\text{MTF} = \text{MTF}_{\text{int}} \times \text{MTF}_{\text{col}}. \quad (5)$$

MTF is defined as the absolute frequency response of the system over the d.c. (i.e., direct current, a term used in electrical engineering) of the response. Typically, the bars in the bar phantom will be hard to resolve if the MTF drops below 0.1. Where the low frequencies correspond to low-contrast large structures, and the high frequencies correspond to fine details, such as edges and small bright lesions.

4.1 Intrinsic resolution

The intrinsic resolution of a SPECT system is limited due to the variations in scintillation photon generation and fluctuations in PMTs. The factors are:

- The thickness of the crystal. *The thicker the crystal, the poorer the intrinsic resolution.* This is because of the variations in light distributions as a function of depth of interaction due to geometric effects (as shown in the left panel of Fig. 7). The position of the incident event depends on the centroid estimation. Therefore the variations in light distribution cause variations in the centroid estimation.

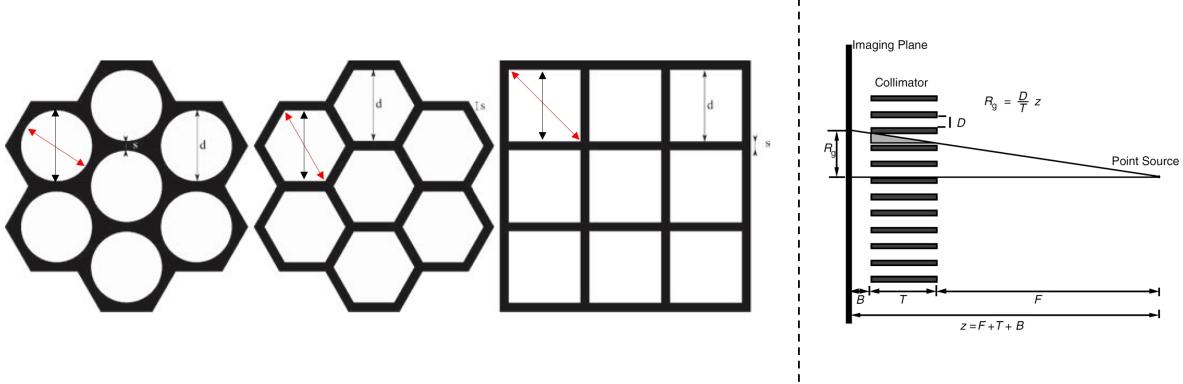


Figure 6: Left: Examples of the hole shapes of collimator (image was obtained from [4]). Right: Relationship between collimator geometries, source placement, and resolution (image was obtained from [2]).

- Scattering of the gamma photons in crystal. Some gamma photons scatter off an electron in an atom and change their direction with a photon energy loss. Then these scattered gamma photons may then be absorbed by the crystal and release scintillation photons, which can cause variations in light distribution, thus, reducing the intrinsic resolution.
- The inhomogeneity of the crystal (flaws during production, such as air bubbles or nonuniform density). The inhomogeneity causes the light distribution to vary by reflecting or absorbing scintillation photons, thus, reducing the intrinsic resolution.
- The fluctuation in the photoelectron generation in PMTs. Electron generated in the photocathode follows a Poisson random process, and the number of electrons emitted which are focused onto the first dynode is also random. Therefore, the output from each PMT will be different. These variations worsen the intrinsic resolution.
- Different gains of PMTs. The gains of PMTs are very sensitive to magnetic fields, even those as small as the earth's magnetic field. The variations can be caused by the improper magnetic shielding of individual PMT. For example, the earth's magnetic field can cause variations in the gains of PMTs with different SPECT camera position.
- The number of PMTs is much smaller than the number of pixels in images from the SPECT camera. Therefore, increase the number of PMTs used per unit area crystal can improve spatial resolution.
- **Projection binning.**

4.2 Collimator resolution

The collimator resolution is affected by three factors:

- Collimator hole diameter.
- The length of septa.
- The source-to-collimator distance.

Collimator resolution is proportional to:

$$R_{\text{col}} = \frac{D}{T} Z, \quad (6)$$

where D is the diameter of the collimator hole, T is the length of the septa, and Z is the source-to-detector distance. As illustrated in the right panel of Fig. 6, the wider the hole is the more photons with different degrees we accept. Thus, worsen the resolution. The same thing for T and Z , smaller T and larger Z lead to more accepting angles, thus decrease the resolution. Different designs of collimator hole shape are shown in the left panel of Fig. 6. Round holes have a uniform resolution in all directions because the hole diameter, d , remains the same in all directions. Hex and square holes have a nonuniform resolution because the hole diameter varies (notice the length difference between the red and black arrows in Fig. 6).

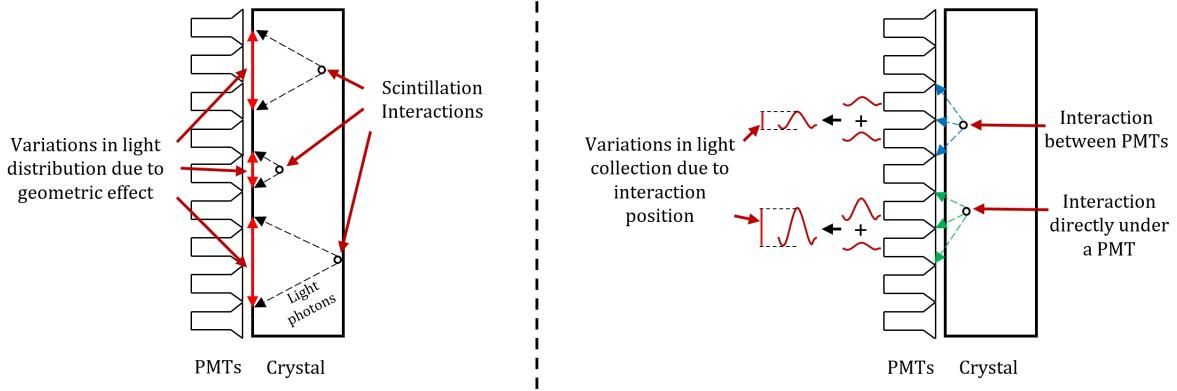


Figure 7: Left: Variations in light distribution caused by geometric effects inside a scintillation crystal. Right: Variations in light collection as a function of interaction positions.

5 SPECT energy resolution

Energy resolution in SPECT can be characterized as the FWHM of the energy of interest, for example, for Technetium-99m (^{99m}Tc) that is 140 KeV. The energy resolution is linked to the ability to distinguish scattered energy from primary photon energy. *It is affected by several factors in a similar way to the intrinsic resolution.* The energy estimation of an incident event depends on the output from the PMTs, which then depends on the light output from the scintillator. This indicates that the estimated energy is proportional to the number of scintillation light photons detected, and the number of scintillation photons emitted per gamma photon is proportional to the energy of the incident gamma photon. The conversion of the scintillation photons from gamma photons follows Poisson distribution. Therefore, to improve the energy resolution (i.e., to make the FWHM smaller), the coefficient-of-variance (COV, i.e., $\frac{\sigma}{\mu} \equiv \frac{1}{\sqrt{\mu}}$) of the number of scintillation photons detected needs to be decreased, meaning that the mean number of scintillation photons detected needs to be increased (i.e., μ needs to be increased). As a result, the energy of the incident gamma photon is inversely proportional to the FWHM. The higher the energy deposited in the crystal, the better the energy resolution. Therefore, any variations in the light transmission and electrical signal conversion will worsen the precision in the energy estimation. Here are the factors that affect energy resolution:

- The variations in light collection. For example, as shown in the right panel of Fig. 7, if the interaction is directly under a PMT, a larger fraction of the total light will be collected, resulting in a larger energy value than if the interaction is in the gap between the PMTs. Because the light collection depends on position (i.e., the light collection is poorer in between PMTs than directly under the PMT), it results in a proportional change in the estimated energy from each PMT as shown in the right panel of Fig. 8. Therefore, the overall energy resolution will decrease. When an energy window is used, the shifts in the energy spectrum reduces the system's sensitivity.
- The energy resolution can also be affected by the inhomogeneities in the scintillator crystal due to the imperfections in the production of the crystal, such as air bubbles or variations in density. The inhomogeneities cause variations in the light output, which reduces energy resolution.
- The statistical fluctuations happened in the PMTs. Different PMTs may have different gains (due to magnetic field fluctuations, as stated in section 4.1). Besides, the photoelectron conversion in the photocathode follows a Poisson random process, and the number of electrons emitted that is focused onto the first dynode is also random. Therefore, the electrical output from each PMT may vary, which limits the energy resolution.
- Scintillator crystal may respond non-linearly as a function of energy deposited. In practice, gamma photons emitted from a radio-isotope is not monoenergetic (single energy). They are, in fact, polyenergetic (various energy), and a portion of photons may scatter inside the patient, resulting in photons with various energy. If the crystal responds non-linearly to the energies deposited, then the outcome FWHM may spread (consider it as the impact of the Gamma correction to an intensity histogram), which limiting the energy resolution.

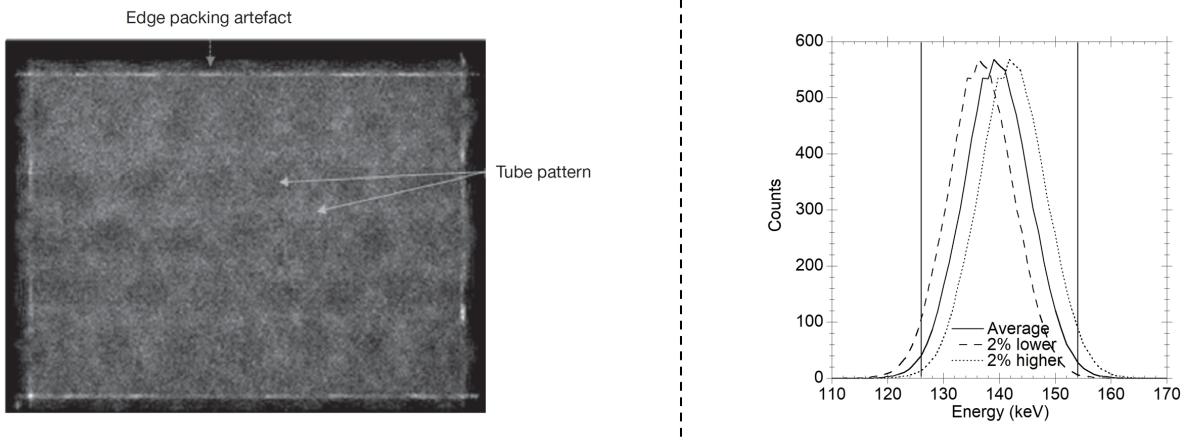


Figure 8: Left: Intrinsic flood image of gamma camera with collimator *removed* and without energy, spatial or sensitivity correction. Right: Shifts in energy spectrum. Images were obtained from [4].

6 SPECT system sensitivity

There has always been a trade-off between system sensitivity and system resolution. The sensitivity is also important for nuclear medicine imaging. Because the noise level in a SPECT or PET image is a function of the total detected counts, and more mean counts lead to reduced relative noise. The collimator sensitivity is given as:

$$S_c = \frac{a_{open}}{4\pi T^2} \frac{a_{open}}{a_{total}}, \quad (7)$$

where T is the length of the septa (as shown in the right panel of Fig. 6), a_{open} is the area of the hole itself, and a_{total} is the total area of the cell including the part of collimator septa. It is obvious that collimator sensitivity can be affected by:

- The diameter of the collimator hole. The larger the hole, the better the sensitivity because more photons are accepted.
- The length of the septa. Reducing the length of the septa can also improve sensitivity, but it decreases the resolution.
- The thickness of the septa. Reducing the thickness of the septa leads to increased sensitivity by allowing septa penetration (i.e., more counts will be detected). However, it will reduce the contrast of the resulting image (increases full width at tenth maximum) and resolution if there is catastrophic penetration.

The sensitivity of the crystal can be characterized as:

$$S_i = 1 - e^{-\mu t}, \quad (8)$$

where μ denotes the linear attenuation coefficient of the crystal, and t denotes the thickness of the crystal. Crystal sensitivity can be affected by

- The thickness of the scintillator crystal t . By increasing t , the crystal stops more gamma photons, resulting in an improved sensitivity. However, the possibility of scattering in crystal and increased quantum noise are expected, which will worsen resolution.
- The energy of the incident gamma photon. The decreasing of gamma photon energy yields improved detector sensitivity, because μ decreases with the increase of the photon energy. Lower photon energy leads to more scattering and less attenuation in crystal, therefore, more detected photons. However, the trade-off is the decreased intrinsic resolution and energy resolution.
- Fluctuation in light collection among PMTs. For example, if the interaction is directly under a PMT, a larger fraction of the total light will be collected, resulting in a larger number of detected counts than if the interaction is in the gap between the PMTs.

In SPECT, sensitivity can be improved only if the resolution is decreased. A final factor that can reduce sensitivity is that:

- The width of the energy acceptance window used. The wider the window, the more accepted photons, but the reduced contrast due to the acceptance of more scattered and septa penetrated photons.

7 Spatial, energy, and sensitivity corrections

Due to the factors that affects intrinsic resolution (section 4) and energy resolution (section 5), the energy and position estimation are non-ideal. As shown in the left panel of Fig. 8, these errors lead to non-uniform sensitivity in the camera, resulting in edge packing artifacts and tube pattern artifacts.

7.1 Energy correction

As mentioned in section 5, there are shifts in the individual energy spectrum due to the variations in light collections among different PMTs, i.e., the variations of the fraction of the scintillation light collected as function of interaction position. These variations may be caused by the geometric effects (as shown in the left panel Fig. 7), or the variations in light collection due to scintillation interaction positions (as shown in the right panel Fig. 7). A typical energy correction algorithm measures the *individual* energy spectrum for each PMT as a function of position in the image using a source with known energies. A linear correction is then made to the estimated energy.

7.2 Spatial correction

Due to the biases produced in PMT's light collecting process (illustrated graphically in the right panel of Fig. 7), the lines in the resulting image are typically bent towards space between PMTs. To perform the spatial correction, a grid of holes or lines in combination with a flood source is imaged to create an image with a series of bright spots or lines. Note that the actual position of each spots or lines are known. Then, a function is fitted to the set of actual positions as a function of the set of measured positions. The resulting fitted function is used to correct each measured position.

7.3 Sensitivity correction

The goal of this correction is to make images of a flood source as uniform as possible (see an example of a non-uniform flood source image in the left panel of Fig. 8). It is the same as spatial correction? There are two types of uniformity corrections: intrinsic, which corrects for non-uniform sensitivity of the detector systems (i.e., excluding the collimator) and extrinsic, which corrects for both detector and collimator non-uniformities. Uniformity corrections are done by imaging a high-count flood image, and then multiplying each pixel value in the acquired image by a factor that is equal to the average counts in the *active portion* of the flood image divided by the counts in the corresponding pixel in the flood image.

8 Image quality for detection tasks

8.1 Contrast

Image contrast refers to differences in intensity in parts of the image corresponding to different levels of radioactive uptake in the patient. Image contrast is computed as:

$$C_l = \frac{R_\ell - R_o}{R_o} = \frac{\Delta R_\ell}{R_o}, \quad (9)$$

where R_ℓ denotes the counting rate over a lesion, and R_o denotes the counting rate over a normal tissue. C_ℓ can be negative or positive. The factors that affect contrast are the ones that add background counting rates:

- Natural background noise
- Gamma photon scattering in patient
- Septa penetration

The new contrast after adding a background counting rate will be:

$$C'_\ell = \frac{(R_\ell + R_b) - (R_o + R_b)}{(R_o + R_b)} = \boxed{C_\ell \times \frac{1}{1 + (R_b/R_o)}}, \quad (10)$$

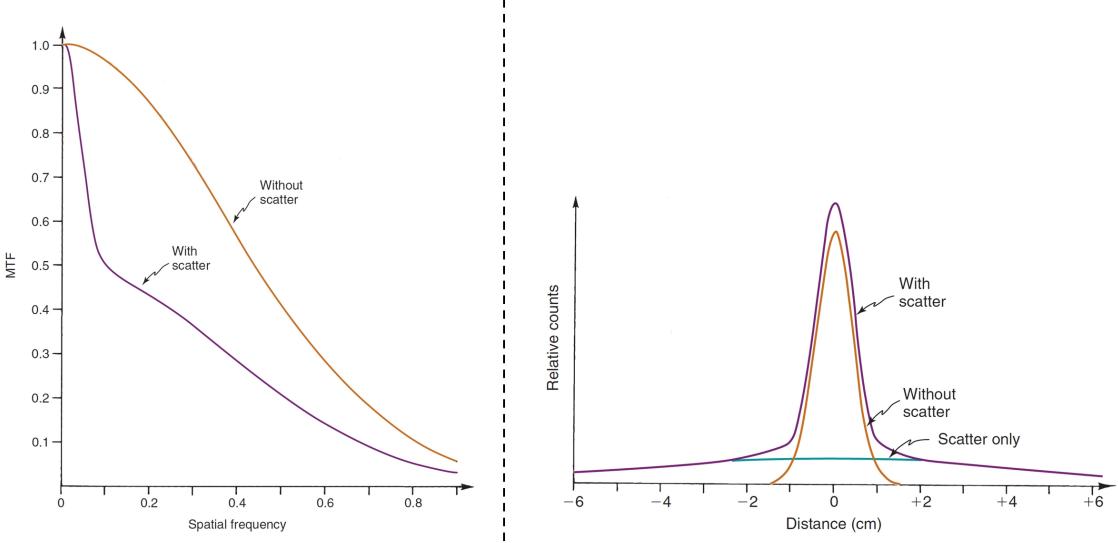


Figure 9: Effects of added background counting rate on the FWHM of PSF (left panel) and MTF (right panel) Images were obtained from [1].

where R_b denotes the background counting rate. Notice that $\frac{1}{1+(R_b/R_o)} < 1$ for any nonzero R_b . Therefore, added a background counting rate will always worsen image contrast and decrease the visibility of all structures, which shows as the suppression of low and high frequencies (as shown in the left panel of Fig. 9). However, the added background does not affect the FWHM or the system resolution by much as shown in the right panel of Fig. 9).

8.2 Noise

There are two types of noise in a SPECT image, random noise and structured noise.

8.2.1 Random noise

8.2.1.1 Coefficient of variance Random noise in NMI comes from the random nature of radioactivity decay and photon detection, following the Poisson distribution. The random noise can be characterized by the coefficient of variance (COV):

$$\boxed{COV = \frac{\sigma}{m} = \frac{1}{\sqrt{m}}}, \quad (11)$$

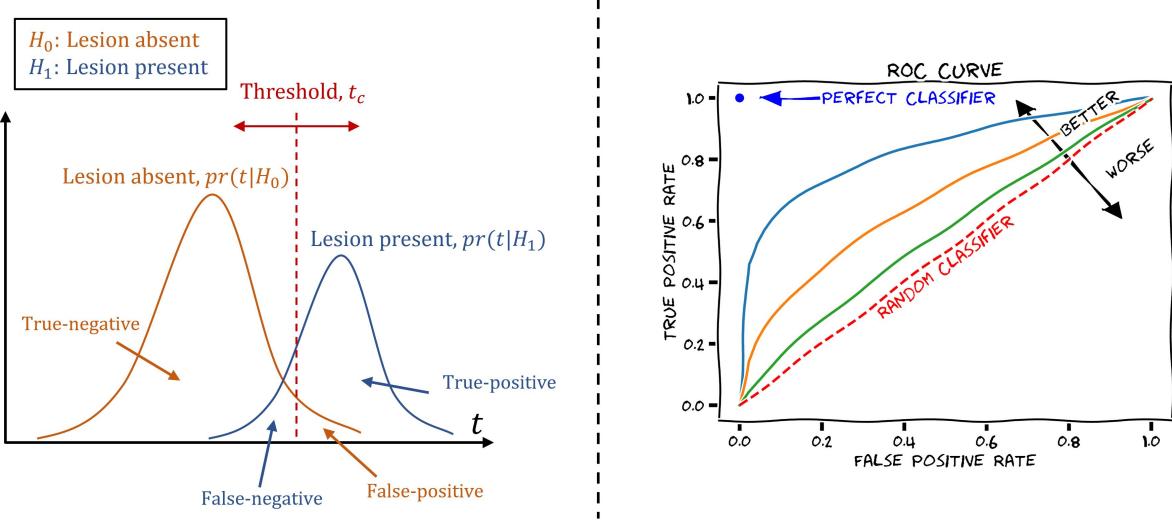
where σ and m represent, respectively, the standard deviation and the mean detected counts. Photon detection in SPECT follows the Poisson distribution, for which the variance σ^2 equals to the mean m (i.e., $\sigma^2 = m$). The lower the COV, the lower the relative noise. Therefore, to decrease COV, \sqrt{m} must be increased, i.e., the mean detected counts need to be increased.

8.2.1.2 Contrast-to-noise rate To measure the detectability of a lesion presented in a SPECT image, we can measure the contrast-to-noise ratio (CNR). CNR over a lesion is defined as:

$$CNR_\ell = \frac{|C_\ell|}{C_{noise}}, \quad (12)$$

where the measure background contrast C_{noise} can be represented as:

$$C_{noise} = \frac{\sigma_o}{N_o} = \frac{1}{\sqrt{N_o}}, \quad (13)$$

Figure 10: Left: Plots of $\text{pr}(t|H_0)$ and $\text{pr}(t|H_1)$. Right: The ROC curves.

where $N_o = R_o \times \frac{\pi}{4} d_\ell^2 \times t$ is the number of counts recorded in a background area of the same size as the lesion (characterized by the lesion diameter d_ℓ) during an acquisition time t . Therefore, CNR can be approximated by:

$$\begin{aligned} \text{CNR}_\ell &= \frac{|C_\ell|}{C_{\text{noise}}} = |C_\ell| \times \sqrt{N_o} \\ &= |C_\ell| \times \sqrt{R_o \times \frac{\pi}{4} d_\ell^2 \times t} \\ &\approx |C_\ell| \times d_\ell \times \sqrt{R_o \times t}. \end{aligned} \quad (14)$$

The quantity $R_o \times t = \text{ID}_o$ is known as the background *information density*. Generally, in order for a lesion to be detectable, CNR must be larger than or equal to 4 (i.e., $|C_\ell| \times d_\ell \times \sqrt{R_o \times t} \geq 4$). Therefore,

- Lesion contrast C_ℓ
- Lesion size (characterized by its diameter) d_ℓ
- Information density $\sqrt{R_o \times t}$

are three factors that can affect lesion detectability. It seems like increasing R_o would increase lesion detectability. However, the increased R_o would lead to the *decreased* C_ℓ as shown in Eqn. (10), thus limiting CNR. For example, if scatter is added to the image formation, the contrast would decrease by a factor of squared, because R_o only increases under the square root. In reality, C_ℓ and d_ℓ are directly linked, CNR increases as the square of spherical lesion diameter.

8.2.2 Structured noise

We described structured noise in section 7, and it is visually shown in the left panel of Fig. 8. Structured noise are non-uniformities and artifacts presented in a SPECT image. Because of the fluctuation in light collection among PMTs, resulting in non-uniformities and edge packing artifacts.

8.3 Observer Performance Studies

The measurements of image quality discussed in the previous sections are not directly related to the object detectability of human observers. Thus, the *observer performance studies* (i.e., receiver operating characteristic (i.e., ROC) studies) are often used to determine the actual detectability. In an ROC study, a set of images obtained from different imaging systems, reconstruction algorithms, or image processing methods is given to the *observer* (human or mathematical observer) to rate the detectability of some targets (e.g., lesion present or lesion absent). The images with "lesion-present" is referred to as *positive images*, and the images with "lesion-absent" is called *negative-images*. Usually the confidence levels are numbered and different levels are permitted; for example, 1 = definitely absent, 2 =

probably absent, 3 = probably present, and 4 = definitely present. Then the following results are computed for each confidence level:

True-positive fraction (TPF) = fraction of positive images *correctly* identified as positive by the observer (15)

False-positive fraction (FPF) = fraction of negative images *incorrectly* identified as positive by the observer

Some additional quantities are:

True-positive fraction (TNF) = fraction of negative images *correctly* identified as negative by the observer (16)

False-negative fraction (FNF) = fraction of positive images *incorrectly* identified as negative by the observer

where FNF = 1 - TPF, and TNF = 1 - FPF. The TPF and TNF are referred to as *sensitivity* and *specificity*, respectively. The ROC curve is generated by plotting TPF versus FPF for progressively relaxed degree of confidence. Mathematically, if we assume the TPF and FPF are normal distributed (as shown in the left panel of Fig. 10):

$$\begin{aligned} TPF &= \frac{1}{2} \left[1 - \operatorname{erf}\left(\frac{t_c - \bar{t}_1}{\sqrt{2\sigma_1^2}}\right) \right], \\ FPF &= \frac{1}{2} \left[1 - \operatorname{erf}\left(\frac{t_c - \bar{t}_0}{\sqrt{2\sigma_0^2}}\right) \right], \end{aligned} \quad (17)$$

where the *error function* is defined as:

$$\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-y^2} dy. \quad (18)$$

Another figure of merit is the area under the ROC curve (AUC), which is computed as the integration of the ROC curve. A better image quality should lead to a better lesion detectability, which would have a larger AUC.

9 Quantitative analysis in SPECT

9.1 Quantitative tasks

There are two general tasks in Nuclear Medicine Imaging (NMI): classification and quantitative tasks.

- *Classification tasks*: Placing a patient into one of several discrete classes, such as cancer diagnostic tasks in fluorodeoxyglucose (FDG) PET cancer imaging.
- *Quantitative tasks*: Extracting a numerical value or values from data obtained in a nuclear medicine procedure, such as measuring the total activity in an organ or tumor or object of interest based on the SUV in FDG imaging.

Classification tasks can, however, be performed on the basis of quantitative values obtained from images: for example, in the diagnosis of cancer based on a strict SUV threshold.

9.2 Figure of merit in SPECT quantification

The figure of merit for SPECT quantification are bias, relative error, precision, and mean squared error. The definitions of each measurement is described in next subsections.

9.2.1 Bias and relative error

Bias is the difference between the mean across measurements and the true activity. Bias and relative error are defined as:

$$b = E\{m_i - t|i\} = \bar{m} - t, \quad (19)$$

$$\text{Relative error} = \frac{b}{t}, \quad (20)$$

where m_i is the i -th measurement, and t is the true activity value. In nuclear medicine applications, it is often more relevant to take the average over a population of patients or in one patient imaged at different times where the true value (t) is expected to differ. The ensemble bias, b_{Ensemble} :

$$b_{\text{Ensemble}} = E\{m_j - t_j|j\} = \bar{m} - \bar{t}. \quad (21)$$

9.2.2 Precision

Precision is defined as the standard deviation across the measurements. The common measures of precision are the standard deviation of activities, σ , and the coefficient of variation, COV:

$$\sigma = \sqrt{E\{(m_i - \bar{m})^2|i\}}, \quad (22)$$

$$\text{COV} = \frac{\sigma}{\bar{m}}. \quad (23)$$

The ensemble standard deviation is defined as:

$$\sigma_{\text{Ensemble}} = \sqrt{E\{(m_j - \bar{m}_j^2)|j\}}. \quad (24)$$

9.2.3 Root mean square error

Root mean squared error (or mean squared error) is the bias squared and the variance combined. A single measurement if reliability is the root mean square error, and the ensemble root mean square error:

$$\begin{aligned} \text{RMSE} &= \sqrt{b^2 + \sigma^2} = \sqrt{E\{(m_i - t)^2|i\}} \\ \text{RMSE}_{\text{Ensemble}} &= \sqrt{b_{\text{Ensemble}}^2 + \sigma_{\text{Ensemble}}^2} = \sqrt{E\{(m_j - t_j)^2|j\}}. \end{aligned} \quad (25)$$

9.3 Factors that affect SPECT quantification

There are ten things that can affect activity quantification of a lesion:

1. Natural background.
2. Radioactive decay.
3. Noise due to the nature of radioactive decay and detection.
4. Photon attenuation is the largest factor.
5. Scatter inside patient.
6. Motion in the imaging.
7. Resolution and partial volume.
8. Count rate problems (dead time caused by pulse pile-up, which reduces counts.)
9. Variations in regions and volumes defined by user.
10. Image reconstruction algorithms.

Each of them is described here:

1. Noise in the background, such as natural background activity or patients in other rooms. The background noise can cause overestimation in the radioactivity, which affects bias and quantification error, but not precision as the background noise level is static, i.e., does not change over time.
2. Radioactive decay can also lead to errors in activity estimation. One should carefully mark the injection time and the imaging time to compensate for the reduced activity from radioactive decay.
3. The noise can also contribute to quantitative error, but not by much. Noise is random, and it follows Poisson distribution, so an increase in noise level will degrade precision (*but not bias*). To reduce noise, the mean detected counts should be increased. This means that we should either increase the dose injected into patient for higher activities or we can image the patient longer to increase acquisition times.
4. Photon attenuation is another factor. Without attenuation correction, the activity will be underestimated by a factor of 5-20. Photon attenuation is spatially varying because of the different components in the human body. Therefore, accurate attenuation correction is needed, which means we need to use iterative reconstruction methods rather than simple FBP.

5. Scatter in the patient body causes an overestimation in activity quantification, which worsens bias, and error, but not precision (scatter contributes to a static level of activity. Thus, the precision remains the same). This is because the error can be calculated as $b^2 + \sigma^2$. The scattering ratio increases approximately linear with the source depth inside the body, which means that in tomographic imaging (i.e., SPECT), the scatter occurs more often when closer to the *center of rotation*. If the image is attenuation compensated but without scatter correction, it will boost uptake value to an even higher overestimation because scatter occurs more often with the source depth, which is the same case for the attenuation. Once the attenuation compensation is applied, overestimated counts due to scattering will be boosted again.
6. Motion can cause errors in quantification. Because of the blurring, ROIs will be hard to define. Variations in ROI will contribute errors to quantification.
7. Because of the limited spatial resolution, which is limited by intrinsic resolution and collimator resolution, there are spill-out (known as partial volume effect) of the uptakes from a region with higher uptake to a region with lower uptake, and vice versa for the spill-in. With the fix-sized ROI, regions that suffer from PVE will result in an increase in the absolute bias. Depending on the task performed, PVE can also affect precision. If we compute precision across different lesions, PVE will contribute to a decrease in precision because PVE have different effects on different lesion sizes. However, if we measure precision on the same lesion but across different scan times, PVE will not contribute to a decrease in precision. In general, PVE worsens the quantification accuracy.
8. Count rate losses for very high count rates. This is due to the dead time of a SPECT system. Because for the PMTs and positional logic circuits to process an incident pulse requires some time, there is a certain amount of time that the system will ignore some of the incident pulses. Also, two incident pulses may get combined into one large pulse with higher estimated energy. These pulses will be rejected by the system. This is known as pulse-pile up, which may contribute to an underestimation of the activity (an increase in the absolute bias).
9. Because of the degrading factors mentioned before, such as motion, limited resolution, and partial volume effect, it is hard to define an accurate volume-of-interest to quantify the activities. The variations in the volume of the VOIs can cause variations in activity quantification, which degrades precision.
10. Image reconstruction method used can also have an impact on quantification accuracy. A SPECT reconstructed image without collimator-detector-response, scatter and attenuation compensations yields an increase in the absolute bias, in which scatter causes an overestimation and attenuation contributes to an underestimation of radioactivities (they are also spatially varying). Different reconstruction methods may produce different noise properties in the resulting image. Therefore, precision can also be affected by the reconstruction method used. The methods for compensations are described in section 11.

10 SPECT image reconstruction

10.1 Forward projection

SPECT detectors are used to collect radiation emitted from the radioactivity distribution to from projection data from different views around the object [5] (an example is shown in the left panel of Fig. 11). The right panel of Fig. 11 shows the geometrical configuration of the 2D image projection. The 2D activity distribution is represented by $f(x, y)$. We denote the rotated coordinate system as $f(x(s), y(s))$, where the rotation angle is defined as θ , and

$$x(s) = \ell \cos \theta - s \sin \theta, \quad (26)$$

$$y(s) = \ell \sin \theta + s \cos \theta. \quad (27)$$

The 1D projection, $g(\ell, \theta)$, is composed of sum of counts from the line of response (i.e., line integral):

$$g(\ell, \theta) = \int_{-\infty}^{\infty} f(x(s), y(s)) \exp \left\{ \int_s^R \mu(x(s'), y(s'); E) ds' \right\} ds, \quad (28)$$

where $R-s$ is the distance between the point $(x(s), y(s))$ and the edge of $f(x(s), y(s))$ along the direction perpendicular to the ℓ axis, μ represents the attenuation coefficient at gamma photon energy E . Here, we assume there is no attenuation effect, i.e., $\mu(x(s'), y(s'); E) = 0$, and Eqn. (28) reduces to:

$$g(\ell, \theta) = \int_{-\infty}^{\infty} f(x(s), y(s)) ds, \quad (29)$$

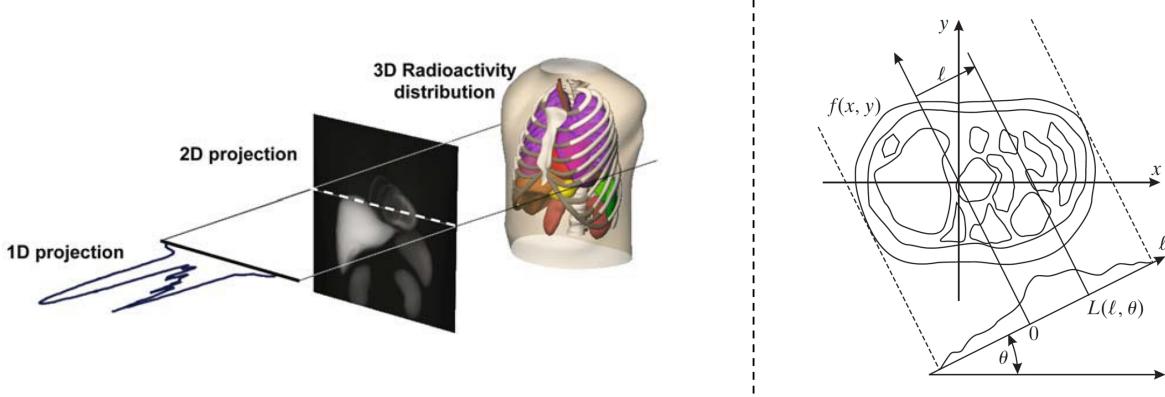


Figure 11: Left: The form of 2D and 1D projection data in SPECT (image was obtained from [5]). Right: The geometry of lines and projections (image was obtained from [6]).

which is simply the line integral of activity distribution (i.e., the radon transform):

$$g(\ell, \theta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta(x \cos \theta + y \sin \theta - \ell) dx dy, \quad (30)$$

This radon transform has an interesting property that $g(\ell, \theta) = g(-\ell, \theta + \pi)$, which is shown below:

$$\begin{aligned} g(-\ell, \theta + \pi) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta(-x \cos \theta - y \sin \theta + \ell) dx dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta(-[x \cos \theta + y \sin \theta - \ell]) dx dy, \end{aligned} \quad (31)$$

where $\delta(-t) = \delta(t)$. Thus,

$$\begin{aligned} g(-\ell, \theta + \pi) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta(-[x \cos \theta + y \sin \theta - \ell]) dx dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta(x \cos \theta + y \sin \theta - \ell) dx dy \\ &= g(\ell, \theta). \end{aligned} \quad (32)$$

10.2 Simple backprojection

The left panel of Fig. 12 illustrates the idea of forward projection and the simple backprojection. The simple backprojection method takes each of the projection data $g(\ell, \theta)$ and smears it over the image pixels in the direction of θ . The simple backprojection can be written as (with attenuation ignored):

$$\begin{aligned} \hat{f}(x, y) &= \int_0^\pi \int_{-\infty}^{\infty} g(\ell, \theta) \delta(x \cos \theta + y \sin \theta - \ell) d\ell d\theta \\ &= \boxed{\int_0^\pi g(x \cos \theta + y \sin \theta, \theta) d\theta}. \end{aligned} \quad (33)$$

As shown in the left panel of Fig. 12, the reconstructed image using the simple backprojection is blurring. It is later analytically shown that the simple backprojection does not correctly reconstruct image. However, the correct procedure still uses this simple backprojection but on filtered versions of the projections.

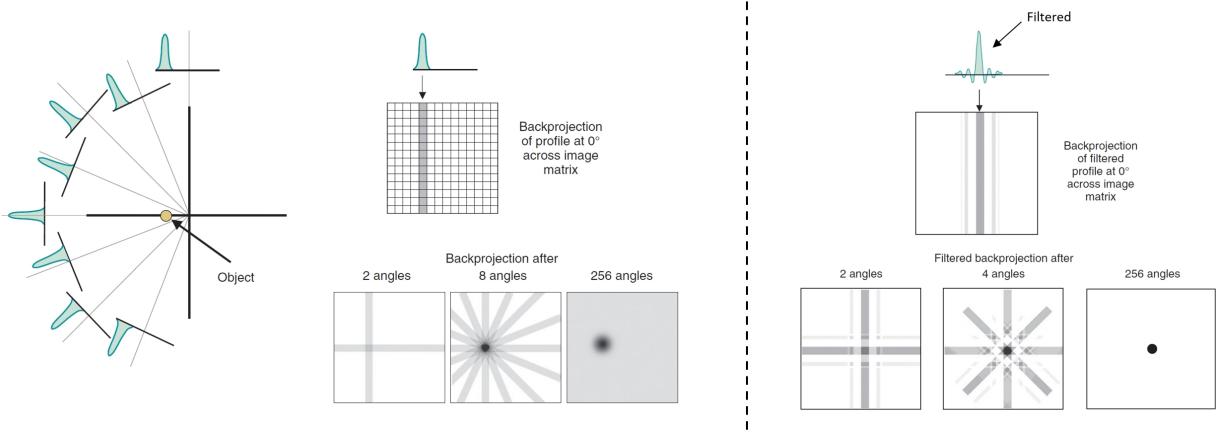


Figure 12: Left: Forward projection and the simple backprojection. Right: The filtered backprojection. Image was obtained from [1].

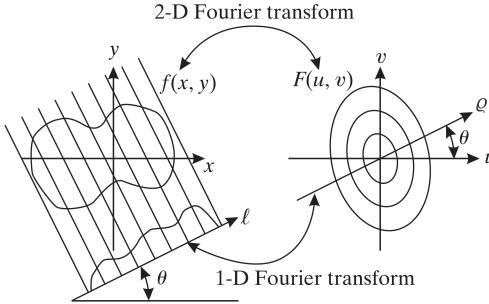


Figure 13: Graphical illustration of the projection-slice theorem. Image was obtained from [6].

10.3 Projection-slice theorem

In this section, we demonstrate a very important relationship between the 1D Fourier transform of a projection $g(\ell, \theta)$ and the 2D Fourier transform of the activity distribution. The 1D Fourier transform of $g(\ell, \theta)$ is written as:

$$\begin{aligned}
 G(\rho, \theta) &= \mathcal{F}_{1D}\{g(\ell, \theta)\} = \int_{-\infty}^{\infty} g(\ell, \theta) e^{-2\pi\rho\ell} d\ell \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta(x \cos \theta + y \sin \theta - \ell) dx dy e^{-2\pi\rho\ell} d\ell \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \int_{-\infty}^{\infty} \delta(x \cos \theta + y \sin \theta - \ell) e^{-2\pi\rho\ell} d\ell dx dy \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{-2\pi\rho(x \cos \theta + y \sin \theta)} dx dy \\
 &= \boxed{F(\rho \cos \theta, \rho \sin \theta)}.
 \end{aligned} \tag{34}$$

This equation (i.e., $G(\rho, \theta) = F(\rho \cos \theta, \rho \sin \theta)$) is known as the *projection-slice theorem*. The 1D Fourier transform of a projection $g(\ell, \theta)$ is a *slice* of the 2D Fourier transform of the activity distribution (with attenuation ignored). This idea is illustrated graphically in Fig. 13.

10.4 The Fourier method of image reconstruction

The Fourier method simply takes the 1D Fourier transform of each projection, inserts it with the corresponding θ to form a complete Fourier space, and then takes the inverse 2D Fourier transform of the result:

$$f(x, y) = \mathcal{F}_{2D}^{-1}\{G(\rho, \theta)\}, \tag{35}$$

where the 2D inverse Fourier transform is defined as:

$$\begin{aligned} f(x, y) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(u, v) e^{2\pi(ux+vy)} du dv \\ &= \int_0^{2\pi} \int_0^{\infty} F(\varrho \cos \theta, \varrho \sin \theta) e^{2\pi\varrho(x \cos \theta + y \sin \theta)} \varrho d\varrho d\theta, \end{aligned} \quad (36)$$

which is done by converting the integrals into a polar form, i.e.,

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s(x, y) dx dy \equiv \int_0^{2\pi} \int_0^{\infty} s(r \cos \theta, r \sin \theta) r dr d\theta. \quad (37)$$

Therefore, the 2D inverse transform of $G(\varrho, \theta)$ is defined as:

$$\begin{aligned} f(x, y) &= \mathcal{F}_{2D}^{-1}\{G(\varrho, \theta)\} \\ &= \int_0^{2\pi} \int_0^{\infty} F(\varrho \cos \theta, \varrho \sin \theta) e^{2\pi\varrho(x \cos \theta + y \sin \theta)} \varrho d\varrho d\theta, \end{aligned} \quad (38)$$

from which follows that

$$f(x, y) = \int_0^{\pi} \int_{-\infty}^{\infty} |\varrho| F(\varrho \cos \theta, \varrho \sin \theta) e^{2\pi\varrho(x \cos \theta + y \sin \theta)} d\varrho d\theta, \quad (39)$$

by using the fact that $g(\ell, \theta) = g(-\ell, \theta + \pi)$. Finally, we replace $F(\varrho \cos \theta, \varrho \sin \theta)$ with $G(\varrho, \theta)$, Eqn. (39) reduces to

$$f(x, y) = \int_0^{\pi} \left[\int_{-\infty}^{\infty} |\varrho| G(\varrho, \theta) e^{2\pi\varrho\ell} d\varrho \right]_{\ell=x \cos \theta + y \sin \theta} d\theta. \quad (40)$$

10.5 Filtered backprojection

Eqn. (40) can be further reduced by transforming $|\varrho|G(\varrho, \theta)$ into spatial domain via the 1D inverse Fourier transform:

$$f(x, y) = \int_0^{\pi} [\mathcal{F}_{1D}^{-1}\{|\varrho|\} * g(\ell, \theta)]_{\ell=x \cos \theta + y \sin \theta} d\theta. \quad (41)$$

This leads to the *convolution backprojection*,

$$\begin{aligned} f(x, y) &= \int_0^{\pi} [c(\ell) * g(\ell, \theta)]_{\ell=x \cos \theta + y \sin \theta} d\theta \\ &= \int_0^{\pi} \left[\int_{-\infty}^{\infty} g(\hat{\ell}, \theta) c(\ell - \hat{\ell}) d\hat{\ell} \right]_{\ell=x \cos \theta + y \sin \theta} d\theta \\ &= \boxed{\int_0^{\pi} \int_{-\infty}^{\infty} g(\ell, \theta) c(x \cos \theta + y \sin \theta - \ell) d\ell d\theta}, \end{aligned} \quad (42)$$

where we define $\mathcal{F}_{1D}^{-1}\{|\varrho|\} \equiv c(\ell)$. ϱ is known as a ramp function (as shown in the left panel of Fig. 14). Notice that $|\varrho|$ is unbounded, i.e., the magnitude of $|\varrho|$ increases with ϱ . Therefore, $c(\ell)$ does not exist, since $|\varrho|$ is not integrable. Besides, the activity distribution in SPECT is often assumed to be band-limited in practice due to the low-passing effects of the collimator-detector response and projection binning. When high frequency noise is present in the acquired projection data, applying ramp function amplifies the high frequency noise, which yields noisy reconstructed image. For above reasons, a smoothing filter (or windowing filter) is often applied in practice in conjunction with the ramp function. As shown in the right panel of Fig. 14, typical choices of the smoothing filters are:

- Rectangular window
- Hanning window
- Butterworth filter

An graphical illustration of an image reconstructed by the filtered backprojection is shown in the right panel of Fig. 12.

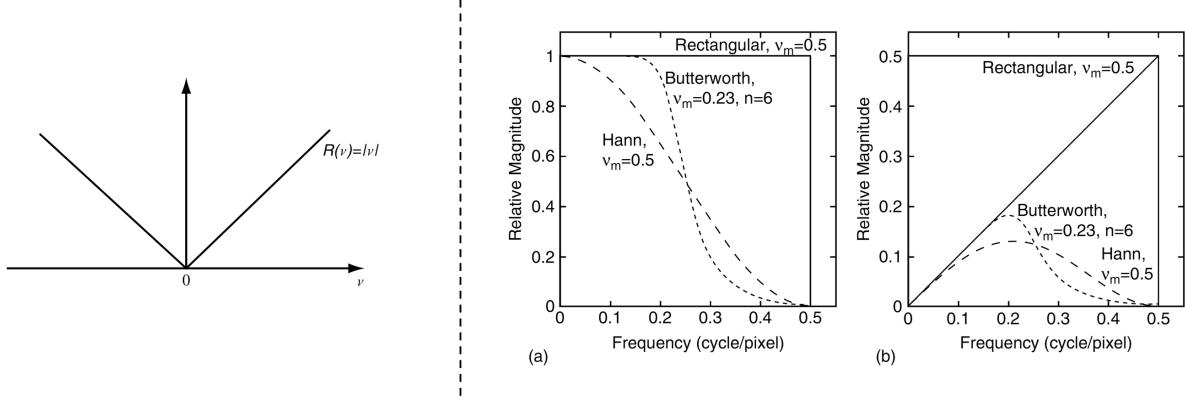


Figure 14: Graphical illustration of the projection-slice theorem. Image was obtained from [6].

10.6 Image reconstruction in a matrix view

The left panel of Fig. 15 describes a discrete version of image projection in a matrix point of view, where \mathbf{g} denotes the 1D projections stacked from all views (i.e., all camera positions, or all θ). Each bin g_i can be defined as the following system of linear equations:

$$g_i = \mathbf{h}_i^T \mathbf{f}, \quad (43)$$

where \mathbf{h}_i is the i^{th} row of \mathbf{H} (highlighted by red box), and each element of \mathbf{f} (i.e., f_k , $k = 1, \dots, N$) represents one pixel (or voxel) in the image space. The forward projection can be summarized by a single matrix equation as:

$$\mathbf{g} = \mathbf{H}\mathbf{f}. \quad (44)$$

Here, \mathbf{f} may represent either a vectorized 2D image slice or a vectorized 3D image volume, and complicated imaging systems is presented as the system matrix \mathbf{H} . In consideration of the randomness in the projection data, this matrix equation should truly be:

$$E[\mathbf{g}] = \mathbf{H}\mathbf{f}, \quad (45)$$

where the i^{th} bin in \mathbf{g} is:

$$\bar{g}_i = \sum_{k=1}^N h_{ik} f_k. \quad (46)$$

Assuming the numbers of events detected in the projection bins are independent of one another, and we write the probability for \mathbf{g} as a Poisson model:

$$p(\mathbf{g}; \mathbf{f}) = \prod_{i=1}^P \frac{\bar{g}_i^{g_i} e^{-\bar{g}_i}}{g_i!}. \quad (47)$$

It is sometimes useful to approximate $p(\mathbf{g}; \mathbf{f})$ as a Gaussian distribution, then:

$$\begin{aligned} p(\mathbf{g}; \mathbf{f}) &= k \exp \left[-\frac{1}{2} \sum_{i=1}^P \frac{(g_i - \bar{g}_i)^2}{\bar{g}_i} \right] \\ &= k \exp \left[-\frac{1}{2} (\mathbf{g} - \mathbf{H}\mathbf{f})^T \mathbf{C}^{-1} (\mathbf{g} - \mathbf{H}\mathbf{f}) \right], \end{aligned} \quad (48)$$

where k is a normalizing constant, and $\mathbf{C} = \text{diag}\{\bar{g}_1, \dots, \bar{g}_P\}$ is the covariance matrix of \mathbf{g} (because mean equals variance). With this, we proceed to the iterative reconstruction methods in the next section.

10.7 Iterative reconstruction methods

10.7.1 Maximum-likelihood Criterion

Maximum-likelihood (ML) is described as: choose the reconstructed image $\hat{\mathbf{f}}$ to be the object function \mathbf{f} for which the measured data would have had the greatest likelihood $p(\mathbf{g}; \mathbf{f})$, which can be mathematically described as:

$$\hat{\mathbf{f}} = \arg \max_{\mathbf{f}} p(\mathbf{g}; \mathbf{f}), \quad (49)$$

that is, choose the \mathbf{f} for which $p(\mathbf{g}; \mathbf{f})$ is the greatest.

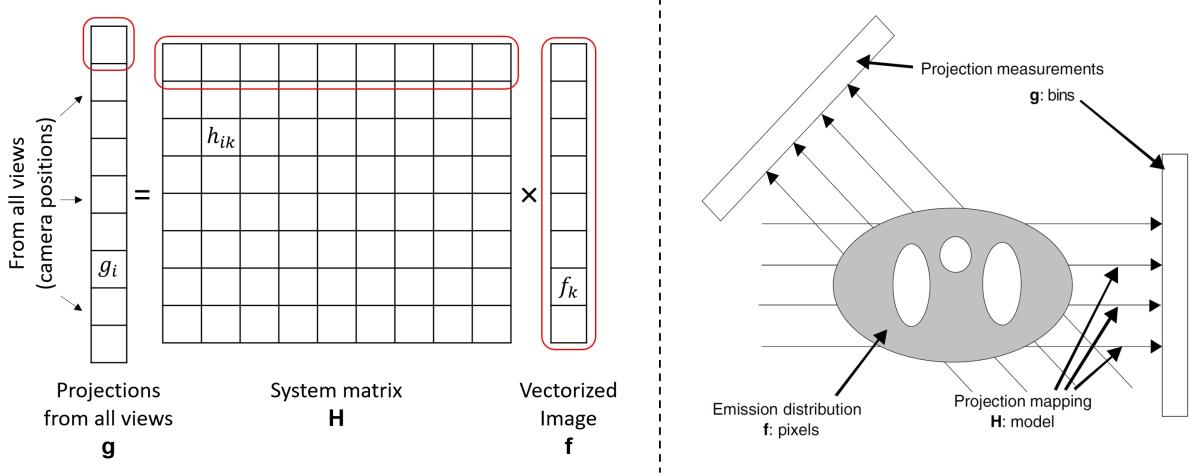


Figure 15: Left: A discrete model (or a matrix view) of the projection process. Right: A general model of tomographic projection (image was obtained from [2]).

10.7.2 Least-Squares and Weighted-Least-Squares

Least-squares (LS) is described as [2]: choose the value of \mathbf{f} that, if observed through the system matrix \mathbf{H} , would yield projections $\mathbf{H}\mathbf{f}$ that are most similar to the observer projections \mathbf{g} in terms of Euclidean distance. LS method can be written as:

$$\begin{aligned}\hat{\mathbf{f}}_{LS} &= \arg \min_{\mathbf{f}} \|\mathbf{g} - \mathbf{H}\mathbf{f}\|^2 = \arg \min_{\mathbf{f}} [(\mathbf{g} - \mathbf{H}\mathbf{f})^T(\mathbf{g} - \mathbf{H}\mathbf{f})] \\ &= \arg \min_{\mathbf{f}} \sum_{i=1}^P \left(g_i - \sum_{k=1}^N h_{ik} f_k \right)^2.\end{aligned}\quad (50)$$

To consider variance in the projection data g_i , we can use the weighted-LS (WLS) method:

$$\begin{aligned}\hat{\mathbf{f}}_{WLS} &= \arg \min_{\mathbf{f}} [(\mathbf{g} - \mathbf{H}\mathbf{f})^T \mathbf{D} (\mathbf{g} - \mathbf{H}\mathbf{f})] \\ &= \arg \min_{\mathbf{f}} \sum_{i=1}^P d_i \left(g_i - \sum_{k=1}^N h_{ik} f_k \right)^2,\end{aligned}\quad (51)$$

where $\mathbf{D} = \text{diag}\{(var[g_1])^{-1}, \dots, (var[g_p])^{-1}\}$ or $\mathbf{D} = \text{diag}\{\bar{g}_1^{-1}, \dots, \bar{g}_p^{-1}\}$ because SPECT data is Poisson-distributed, the variance equals mean.

The closed-form solution to LS and WLS problems can be derived analytically using the pseudo-inverse:

$$\hat{\mathbf{f}}_{LS} = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{g} \quad (52)$$

$$\hat{\mathbf{f}}_{WLS} = (\mathbf{H}^T \mathbf{D} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{D} \mathbf{g} \quad (53)$$

where we assume $\mathbf{H}^T \mathbf{H}$ and $\mathbf{H}^T \mathbf{D} \mathbf{H}$ are invertible. Notice that LS and WLS solutions are equivalent to ML solution if we assume the log-likelihood $\log[p(\mathbf{g}; \mathbf{f})]$ is Gaussian while LS assumes equal variance for g_i .

10.7.3 The Maximum-Likelihood Expectation-Maximization algorithm

We define s_{im} to be the number of photons emitted from within pixel m and detected in projection bin i . The observation projection data in i^{th} bin can be defined as:

$$g_i = \sum_m s_{im}, \quad (54)$$

and

$$E[s_{im}] = h_{im} f_m. \quad (55)$$

As mentioned previously, SPECT data is Poisson-distributed. The Poisson likelihood can be written as

$$p(\mathbf{s}; \mathbf{f}) = \prod_i \prod_m \frac{E[s_{im}]^{s_{im}} e^{-E[s_{im}]} }{s_{im}!}, \quad (56)$$

and the log-likelihood is

$$\boxed{\ln p(\mathbf{s}; \mathbf{f}) = \sum_i \sum_m s_{im} \ln(E[s_{im}]) - E[s_{im}] - \ln(s_{im}!)}. \quad (57)$$

We first compute the *E-step* of MLEM:

$$\begin{aligned} Q(\mathbf{f}; \hat{\mathbf{f}}^{(n)}) &= E[\ln p(\mathbf{s}; \mathbf{f}) | \mathbf{g}; \mathbf{f}^{(n)}] = \sum_i \sum_m E[s_{im} | \mathbf{g}; \mathbf{f}^{(n)}] \ln(E[s_{im}]) - E[s_{im}] - E[\ln(s_{im}!)] \\ &= \sum_i \sum_m E[s_{im} | \mathbf{g}; \mathbf{f}^{(n)}] \ln(h_{im} f_m) - h_{im} f_m - E[\ln(s_{im}!)], \end{aligned} \quad (58)$$

where

$$E[s_{im} | \mathbf{g}; \mathbf{f}^{(n)}] = \frac{h_{im} f_m}{\sum_k h_{ik} f_k} g_i \triangleq p_{im}. \quad (59)$$

Therefore,

$$Q(\mathbf{f}; \hat{\mathbf{f}}^{(n)}) = \sum_i \sum_m p_{im} \ln(h_{im} f_m) - h_{im} f_m - E[\ln(s_{im}!)]. \quad (60)$$

Next, we compute *M-step*:

$$\frac{\partial Q(\mathbf{f}; \hat{\mathbf{f}}^{(n)})}{\partial f_j} = \sum_i p_{ij} \frac{h_{ij}}{h_{ij} f_j} - h_{ij} = \sum_i \frac{p_{ij}}{f_j} - h_{ij}. \quad (61)$$

By setting $\frac{\partial Q(\mathbf{f}; \hat{\mathbf{f}}^{(n)})}{\partial f_j} = 0$, we have

$$\begin{aligned} \frac{\partial Q(\mathbf{f}; \hat{\mathbf{f}}^{(n)})}{\partial f_j} &= \sum_i \frac{p_{ij}}{f_j} - h_{ij} = 0 \\ \sum_i \frac{p_{ij}}{f_j} &= \sum_{i'} h_{i'j} \\ f_j &= \frac{\sum_i p_{ij}}{\sum_{i'} h_{i'j}} = \frac{f_j^{(n)}}{\sum_{i'} h_{i'j}} \sum_i h_{ij} \frac{g_i}{\sum_k h_{ik} f_k^{(n)}} \end{aligned} \quad (62)$$

Therefore,

$$\boxed{f_j^{(n+1)} = \frac{f_j^{(n)}}{\sum_{i'} h_{i'j}} \sum_i h_{ij} \frac{g_i}{\sum_k h_{ik} f_k^{(n)}}}. \quad (63)$$

Notice the similarities between Eqn. (53) and (63), in which $\frac{1}{\sum_k h_{ik} f_k^{(n)}} \equiv \frac{1}{g_i}$ can be think of as the i^{th} element of \mathbf{D} (i.e., $d_i = \frac{1}{g_i}$), and $\sum_i h_{ij} \frac{g_i}{\sum_k h_{ik} f_k^{(n)}}$ can be think of as $\mathbf{H}^T \mathbf{D} \mathbf{g}$. By convention, \mathbf{H} and \mathbf{H}^T (i.e., $\sum_j h_{ij} [\cdot]$ and $\sum_i h_{ij} [\cdot]$) denote, respectively, the forward- and backward- projection operators. It can be shown that if \mathbf{H} is an orthogonal or a semi-orthogonal (non-square) matrix, $\mathbf{H}^T \mathbf{g}$ is actually the simple backprojection (described in section 10.2) of the projection data \mathbf{g} [7]. In the field of image reconstruction, Eqn. 63 is often divided into four separate steps (this is graphically shown in Fig. 16):

- Forward projection of image f at n^{th} iteration: $\sum_k h_{ik} f_k^{(n)}$.
- Compare projection entries by taking the ratio: $\frac{g_i}{\sum_k h_{ik} f_k^{(n)}}$.
- Backward projection of the errors: $\sum_i h_{ij} \frac{g_i}{\sum_k h_{ik} f_k^{(n)}}$.
- Update the image: $\frac{f_j^{(n)}}{\sum_{i'} h_{i'j}} \sum_i h_{ij} \frac{g_i}{\sum_k h_{ik} f_k^{(n)}}$.

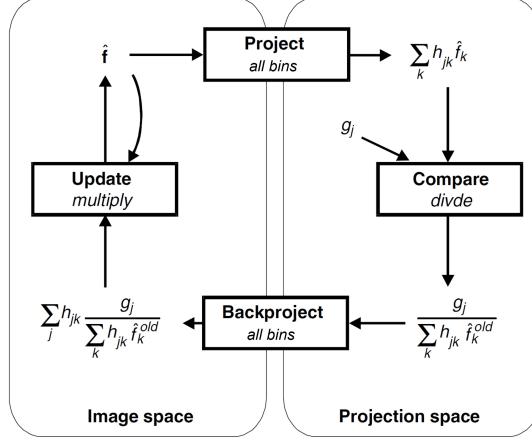


Figure 16: An overview of the MLEM algorithm. Image was obtained from [2].

Note that if there is no error in the projection, that is, $\frac{g_i}{\sum_k h_{ik} f_k^{(n)}} = 1$, then the update equation becomes:

$$f_j^{(n+1)} = \frac{f_j^{(n)}}{\sum_{i'} h_{i'j}} \sum_i h_{ij} \cdot 1 = f_j^{(n)}. \quad (64)$$

Therefore $\frac{1}{\sum_{i'} h_{i'j}}$ is a normalizing term.

10.7.4 Shortcoming of Maximum-Likelihood, Least-Squares, and Weighted-Least-Squares

In practice, the projection data \mathbf{g} is noisy due to various reasons (e.g., quantum noise, background noise, etc.). If we view noise in \mathbf{g} as an additive zero-mean contribution \mathbf{n} , i.e., $\mathbf{g} = \mathbf{H}\mathbf{f} + \mathbf{n}$, the LS solution becomes:

$$\begin{aligned} \hat{\mathbf{f}} &= \mathbf{H}^+ (\mathbf{H}\mathbf{f} + \mathbf{n}) \\ &= \mathbf{f} + \mathbf{H}^+ \mathbf{n}. \end{aligned} \quad (65)$$

In SPECT, the system matrix \mathbf{H} is often a low-pass operator due to collimator-detector response. Therefore, the pseudo-inverse \mathbf{H}^+ is a high-pass operator, which tends to amplify noise in the image. The result $\hat{\mathbf{f}}$ is generally an extremely noisy image.

10.7.5 Bayesian methods

Bayesian methods in SPECT reconstruction are often used to reduce noise in the reconstructed image by introducing prior knowledge, which leads to the maximum a posterior (MAP) criterion: choose \mathbf{f} that maximizes the posterior PDF, $p(\mathbf{f}|\mathbf{g})$:

$$\hat{\mathbf{f}} = \arg \max_{\mathbf{f}} p(\mathbf{f}|\mathbf{g}). \quad (66)$$

By using Bayes' law, we have:

$$\hat{\mathbf{f}} = \arg \max_{\mathbf{f}} p(\mathbf{f}|\mathbf{g}) = \arg \max_{\mathbf{f}} \frac{p(\mathbf{g}|\mathbf{f})p(\mathbf{f})}{p(\mathbf{g})}. \quad (67)$$

We can take the logarithm and omit $p(\mathbf{g})$ because it is not a function of \mathbf{f} . Then, we have:

$$\hat{\mathbf{f}} = \arg \max_{\mathbf{f}} [\ln p(\mathbf{g}|\mathbf{f}) + \ln p(\mathbf{f})]. \quad (68)$$

MAP estimator uses $\ln p(\mathbf{f})$ to penalize image solutions that do not meet the expected properties (e.g., piece-wise smoothness, sharp edges, etc.) but the trade-off is the increased bias (it is known as the *bias-variance trade-off*). There are many choices of $\ln p(\mathbf{f})$, such as Gibbs prior, Total-variational prior, and anatomical priors. However, defining a prior can be difficult. Here is an example described in [2], suppose that our prior model did not anticipate patients with a particular type of lesion. In this case, the reconstruction algorithm might view images containing this lesion as improbable and thus attempt to suppress the lesion in the reconstruction.

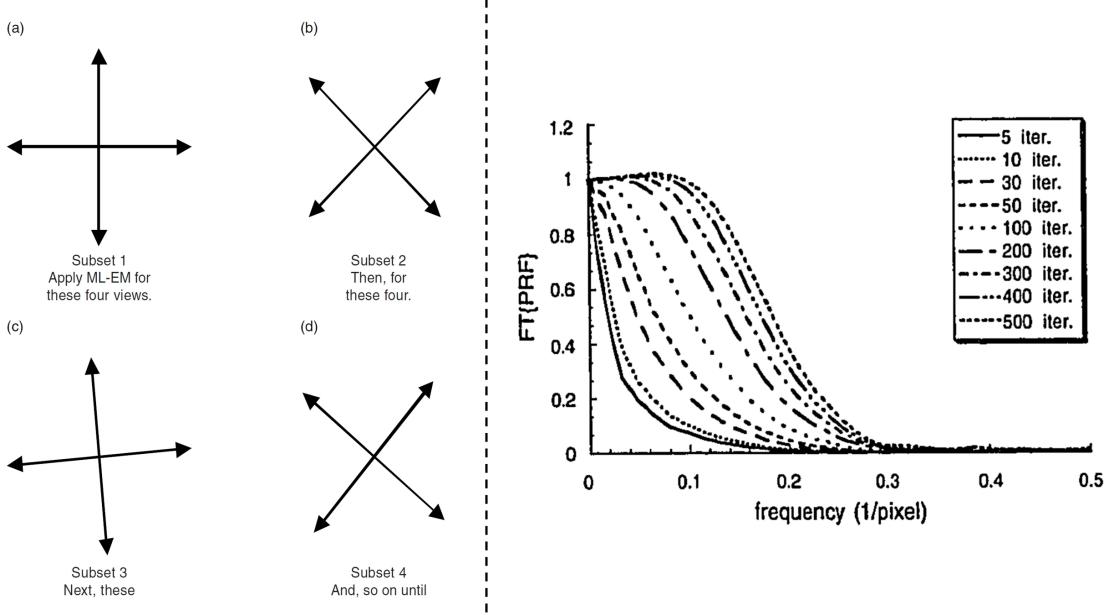


Figure 17: Left: Examples of subset selection and projection views selection in a subset (image was obtained from [2]). Right: MTF of ML-EM with different iterations (image was obtained from [8]).

10.7.6 The Ordered-Subset Expectation-Maximization algorithm

A major drawback of the ML-EM algorithm is that it suffers from slow convergence. The OS-EM algorithm was introduced to speed up the convergence rate of the ML-EM by breaking up the full set of projection data (from all views/angles/camera-positions) into a series of mutually exclusive subsets and apply ML-EM to each subset sequentially. The OS-EM is given by:

$$f_j^{(n+1)} = \frac{f_j^{(n)}}{\sum_{i' \in S_n} h_{i'j}} \sum_{i \in S_n} h_{ij} \frac{g_i}{\sum_k h_{ik} f_k^{(n)}}, \quad (69)$$

where the backprojections (i.e., $\sum_{i' \in S_n} h_{i'j}$ and $\sum_{i \in S_n} h_{ij}$) are performed only for a subset of projection bins S_n . At each update (or known as *sub-iteration*), a different subset of projection bins S_n is used. One pass through all of the subsets is referred to as one *iteration*. Note that the processing time for one iteration of OS-EM is comparable to that of one iteration of ML-EM [2].

The organization of the subsets is very important not only to the performance of the algorithm but also to the mathematical correctness of the algorithm. If any subset does not contain some contribution from every pixel in the field of view, then $\frac{f_j^{(n)}}{\sum_{i' \in S_n} h_{i'j}} \rightarrow \frac{f_j^{(n)}}{0}$, which leads to mathematical error. Usually, subsets are chosen in groups of projection bins associated with one projection view or camera position [2]. The members of a subset are selected to have maximum angular distance between them. As shown in the left panel of Fig. 17, to create 32 subsets for SPECT data with 128 projection views (i.e., camera positions) over a 360° arc, each subset would consist of projections from $\frac{128}{32} = 4$ camera positions and the angular interval between the camera positions is $\frac{360^\circ}{4} = 90^\circ$.

OS-EM at n iterations reaches roughly the same point of convergence as ML-EM at $\text{num. of subsets} \times n$. Therefore, it is capable of speeding up the convergence by a factor of num. of subsets . Similar to ML-EM, OS-EM models low spatial frequencies, with higher spatial frequencies improving with further iterations. As shown in the right panel of Fig. 17, the MTF gradually include more higher frequency components with further iterations.

10.7.7 Important remarks

The FBP method described in section 10.5 is referred to as an analytical method, i.e., FBP provides an exact solution to image reconstruction. However, it is important to emphasize that FBP ignored image degrading factors (e.g., attenuation, scatter, etc.), limiting its reconstruction performance. On the other hand, ML-EM, OS-EM, and LS methods described in section 10.7.2, section 10.7.3 and section 10.7.6 are known as iterative methods (although LS methods provide

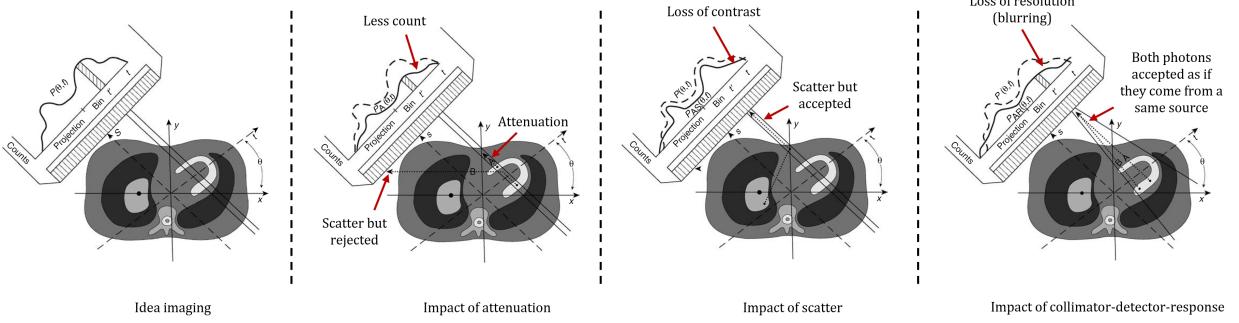


Figure 18: Geometry of ideal SPECT imaging, and impact of attenuation, scatter, and collimator-detector-response on SPECT imaging. Images were obtained from [2].

closed-form solution), in which ML-EM and OS-EM enable the exact modeling of imaging physics in their forward- and backward-projection operations. Due to this reason, OS-EM is currently the state-of-the-art method for SPECT image reconstruction. The commonly used compensation methods in OS-EM are introduced in the next section.

11 Attenuation, scatter, and collimator-detector response compensation in SPECT

As mentioned earlier, SPECT imaging is not ideal. There are several factors degrade SPECT projection data, such as photon attenuation, the inclusion of scattered photon in the projection due to limited energy resolution (see section 5), and the limited spatial resolution caused by collimator-detector-response (CDR) (see section 4). In the next subsections, we briefly introduce some methods used for compensating the image degrading factors.

11.1 Attenuation compensation

As shown in the second figure of Fig. 18, photons emitted from a source inside a patient may be either photoelectrically absorbed (see Eqn. (3)) or scattered so that they are not detected by the detector. This effect shows as a reduction in the number of photons detected varies with the thickness and the material composition of the matter. As shown in Eqn. (28), the transmitted fraction, $TS(x(s), y(s), \theta)$, which is the fraction of the photons from location $(x(s), y(s))$ that will be transmitted through an attenuator at angle θ . It is defined as:

$$TS(x(s), y(s), \theta) = \exp \left(- \int_s^\infty \mu(x(s'), y(s')) ds' \right), \quad (70)$$

where μ is the distribution of linear attenuation coefficients as a function of location, and

$$x(s) = \ell \cos \theta - s \sin \theta, \quad (71)$$

$$y(s) = \ell \sin \theta + s \cos \theta. \quad (72)$$

To compensate for attenuation, we first compute $TS(x(s), y(s), \theta)$ at each $(x(s), y(s))$ and θ . The attenuation compensation is performed by modeling attenuation in the forward projection stage of the ML-EM or OS-EM algorithm, attenuation is modeled by taking the product of $TS(x(s), y(s), \theta)$ and its corresponding h_{ik} (see Eqn. (63) and (69)), which is the system matrix at the i^{th} bin (specified by $(x(s), y(s))$ and θ) and k^{th} pixel location (specified by $(x(s), y(s))$).

11.2 Scatter estimation

In previous sections, we have discussed how photon scattering is originated (section 2) and how it affects image quality for quantitative (section 9) and detection tasks (section 8). Its impact on SPECT projection data is shown in the third figure of Fig. 18. The inclusion of scattered photons in projection data is primarily due to the limited energy resolution (as shown in Fig. 19).

A commonly used method to estimate scattered photons is the Triple-energy-window (TEW) method [9]. In this method, scatter is estimated as the area under a trapezoid formed by the heights of the counts per KeV in each of the two windows (lower and upper windows) on either side of the photopeak window, and a width of the photopeak window

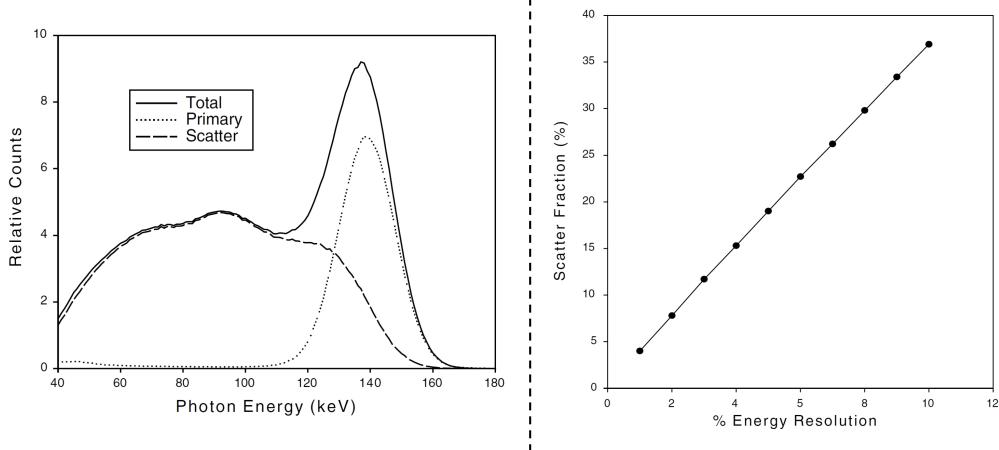


Figure 19: Left: Total, primary (140 KeV), and scatter energy spectrums generated using SIMIND simulation. Right: Percentage scatter fraction versus percentage energy resolution generated using SIMIND simulation. Images were obtained from [2]).

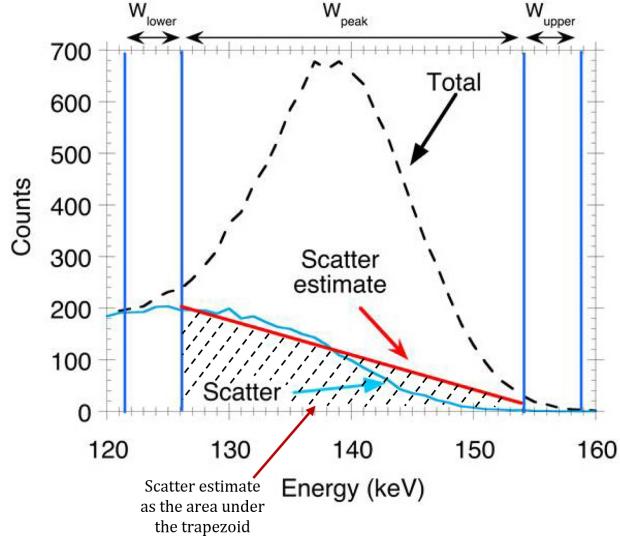


Figure 20: Triple-energy-window scatter estimation method. Image was obtained from [10]).

(as shown in Fig. 20). The scatter estimate image can be calculated as:

$$I_{scatt}^{TEW}(x, y) = \left[\frac{I_{lower}(x, y)}{W_{lower}} + \frac{I_{upper}(x, y)}{W_{upper}} \right] \cdot \frac{W_{peak}}{2}. \quad (73)$$

Note that this scatter estimate image is in projection space (i.e., it is calculated based on projection images). The scatter estimate is subtracted from the photopeak projection image.

11.3 Collimator-detector response

Collimator-detector response (CDR) comes from the finite distance-dependent spatial resolution of SPECT imaging system. In the absence of septal penetration and scatter, the PSF for parallel-hole collimators is approximated as a Gaussian function. The standard deviation ($\sigma_C(d)$) of the Gaussian function is a linear function of distance to collimator face:

$$\sigma_C(d) = \sigma_0 + \sigma_d \cdot d, \quad (74)$$

where d is distance from the face of the collimator (i.e., F in Fig. 3b), σ_0 is the standard deviation at the face of the collimator, and σ_d is the change in standard deviation per unit distance. Combine with a Gaussian function that models

the intrinsic spatial resolution, the standard deviation of the overall Gaussian function is written as:

$$\sigma_s(d) = (\sigma_C^2(d) + \sigma_I^2)^{1/2}, \quad (75)$$

where σ_I denotes the standard deviation of the Gaussian function modeling the intrinsic spatial resolution. This knowledge can be used to model CDR effects during the forward projection of ML-EM or OS-EM. The system response at some distance d_{i+1} can be expressed as the convolution of the response at distance d_i . It can be shown that the convolution of two Gaussians produce a Gaussian whose σ equals the square root of the sum of the σ s squared. Therefore, to reduce computational burden, instead of calculating σ at each distance, we calculate σ_{inc} , which is defined as:

$$\sigma_{inc} = \sqrt{\sigma_s^2(d_{i+1}) - \sigma_s^2(d_i)}. \quad (76)$$

The advantage here is that σ_{inc} is much smaller than $\sigma_s(d_{i+1})$ so that it requires less computational power. The CDR blurring is modeled as incrementally convolving projection data with σ_{inc} as a function of distance.

References

- [1] Simon R Cherry, James A Sorenson, and Michael E Phelps. *Physics in nuclear medicine e-Book*. Elsevier Health Sciences, 2012.
- [2] Miles N Wernick and John N Aarsvold. *Emission tomography: the fundamentals of PET and SPECT*. Elsevier, 2004.
- [3] Magdy M Khalil, Jordi L Tremoleda, Tamer B Bayomy, and Willy Gsell. Molecular spect imaging: an overview. *International journal of molecular imaging*, 2011, 2011.
- [4] Dale L Bailey and JL Humm. *Nuclear medicine physics: a handbook for teachers and students*. Iaea, 2014.
- [5] Habib Zaidi. *Quantitative analysis in nuclear medicine imaging*. Springer, 2006.
- [6] Jerry L Prince and Jonathan M Links. *Medical imaging signals and systems*. Pearson Prentice Hall Upper Saddle River, 2006.
- [7] Nasser M. Abbasi. Finding image forward projection and its transpose matrix. https://www.12000.org/my_notes/image_projection_matrix/index.htm. Accessed: 2021-05-10.
- [8] Donald W. Wilson. *Noise and resolution properties of FB and MLEM reconstructed SPECT images*. PhD thesis, University of North Carolina, Chapel Hill, 1994.
- [9] Takashi Ichihara, Koichi Ogawa, Nobutoku Motomura, Atushi Kubo, and Shozo Hashimoto. Compton scatter compensation using the triple-energy window method for single-and dual-isotope spect. *Journal of Nuclear Medicine*, 34(12):2216–2221, 1993.
- [10] Irene Buvat, Eric Frey, Alan Green, and Michael Ljungberg. Quantitative nuclear medicine imaging: concepts, requirements and methods. *Human Health Reports*, 9, 2014.