A Short Course in Medical Imaging and Positron Emission Tomography

Rome, 2010

This course is meant to be an introduction to medical imaging with an emphasis on positron emission tomography (PET). I will put medical imaging into the context of the formalism of linear imaging and put PET into the context of medical imaging in general. I will thus describe how PET fits into the spectrum of medical imaging techniques, its rationale, applications and utility in clinical medicine and research. I am including references to appropriate Wikipedia pages and other public-domain references and to published articles.

Outline

1. Principles of Linear Imaging
    (a) shift invariance
    (b) impulse response and transfer function
    (c) Nyquist-Shannon sampling
    (d) aliasing
2. Projection Imaging
    (a) pinhole imager
    (b) transmission imager
    (c) optical camera
    (d) xray imaging
    (e) single γ nuclear imaging
3. Computed Tomography (CT)
    (a) Radon transform
    (b) central slice theorem
    (c) Fourier reconstruction and impulse response
    (d) sampling in tomography
    (e) xray CT
    (f) MRI
    (g) single γ tomographic imaging (SPECT)
    (h) positron emission tomography (PET)
        i. superiority of PET
            A. image quality
            B. quantitative physiology

ii. PET camera architecture
iii. PET detectors
iv. PET attenuation correction
v. PET backgrounds
vi. PET image reconstruction
vii. time of flight PET
viii. PET radiochemistry
ix. clinical PET
x. PET research
xi. PET physiological modeling

4. Image Processing
   (a) imaging as a stochastic process
   (b) stationary and ergotic stochastic processes
   (c) Poisson process
   (d) Wiener-Khinchin theorem
   (e) Filtering
      i. simple filters
      ii. optimum filters
         A. Wiener filter
         B. matched filter
      iii. Kalman filter in PET processing
      iv. tomographic reconstruction as a filter

Definition of an Image: The measurements of a physical quantity $o(\mathbf{x}_o)$ on a grid in object space, normally stored and displayed as an array of values $i(\mathbf{x}_i)$ in image space. The measurements may be multidimensional such as a vector or tensor quantity. Measurements are random variables and in general are correlated. We limit ourselves to linear imaging systems. We characterize the imaging operator by a transfer function $g(\mathbf{x}_i, \mathbf{x}_o)$ such that $E[i(\mathbf{x}_i)] = \int g(\mathbf{x}_i, \mathbf{x}_o) o(\mathbf{x}_o) d\mathbf{x}_o$. We have severely limited ourselves. For example, diffractive imaging isn't included unless the object is spatially incoherent. A further limitation is shift invariance $g(\mathbf{x}_i, \mathbf{x}_o) = g(\mathbf{x}_i - \mathbf{x}_o)$, which asserts that the response to a (point) impulse is the same for every location in object space and is translated appropriately in image space. A good impulse response $g(\mathbf{x})$ is narrow and symmetrical. The point spread function (psf) is the impulse response normalized to unit area. A good psf thus approaches a δ function.

Impulse Response

The shift-invariant impulse response makes the image a convolution $E[i(\mathbf{x}_i)] = \int_{-\infty}^{\infty} g(\mathbf{x}_i - \mathbf{x}_o) o(\mathbf{x}_o) d\mathbf{x}_o = g(\mathbf{x}_i) \star o(\mathbf{x}_i)$. Taking the Fourier transform we represent the imaging operation as multiplication in $\mathbf{k}$ space, $E[I(\mathbf{k})] = G(\mathbf{k})O(\mathbf{k})$.

$G(\mathbf{k})$ is the transfer function(TF) of the system and $|G(\mathbf{k})|/G(0)$ is the modulation transfer function (MTF). A good transfer function is real and flat from $\mathbf{k} = 0$ to large $|\mathbf{k}|$. The imaging operator has eigenfunctions $e^{i\mathbf{k}\cdot\mathbf{x}}$ with eigenvalues $G(\mathbf{k}) = \int_{-\infty}^{\infty} g(\mathbf{x})e^{-i\mathbf{k}\cdot\mathbf{x}}$, the Fourier transform of $g(\mathbf{x})$. Shift-invariant linear imaging is encompassed in LTI system theory, which applies to signal processing, control theory, circuits, seismology, etc.
`http://en.wikipedia.org/wiki/LTI_system_theory`

We can generalize shift invariance to imaging with magnification such that $g(\mathbf{x}_i, \mathbf{x}_o) = g(\mathbf{x}_i - M\mathbf{x}_o)$. Then $E[I(\mathbf{k})] = G(\mathbf{k})O(M\mathbf{k})$. There are no eigenfuctions but $e^{i\mathbf{k}\cdot\mathbf{x_o}}$ is taken to $\frac{1}{M^n}G(\mathbf{k}/M)e^{i\mathbf{k}\cdot\mathbf{x_i}/M}$, where n is the dimensionality of the space.

Our intuition for imaging is based on human vision. The optics of the eye, when adjusted to the object distance, causes the irradiance ($\mathrm{W/m^2}$) on the retina to be proportional to the radiance ($\mathrm{W/m^2/ster}$) of a surface emitter. The detector (retina) translates this signal into a measurement of illuminance ($\mathrm{lm/m^2}$). If there are multiple surfaces in the field of view, the eye superposes the images, all but one necessarily out of focus, and all with different magnifications. The brain is clever enough to make sense of the resulting data but intrinsically the eye is not a very good imaging device. The optical instruments we are familiar with are variants of the eye. We would prefer an image that is intrinsically 3-dimensional, with magnification independent of coordinates, and all elements simultaneously "in focus". We can make devices that produce such images for a variety of physical quantities, and many of these methods now play a prominent role in medical imaging.
`http://en.wikipedia.org/wiki/Eye`

Projection Imaging

It is straightforward to compute the impulse response for projection imaging for simple systems.

For a "pinhole" imager with plane aperture $a(\mathbf{x})$ a distance $L1$ from the object plane and $L2$ from the image plane, in the paraxial approximation, $g(\mathbf{x}_i, \mathbf{x}_o) = a((\mathbf{x}_i - M\mathbf{x}_o)/m)/(L1 + L2)^2$, where $M = -L2/L1$ is the object magnification and $m = (L1 + L2)/L1$ is the aperture magnification. The object measure is radiance and the image measure is irradiance.

For a transmission imager with source of radiance $\mathrm{L}(\mathbf{x})$ placed a distance $L1$ from the object with transmission $t(\mathbf{x}_o), 0 \leq t \leq 1$, which is $L2$ from the detector, $g(\mathbf{x}_i, \mathbf{x}_o) = \frac{m^2}{M^2}L((\mathbf{x}_i - m\mathbf{x}_o)/M)/(L1+L2)^2$. (same definitions of M and m)

For an ideal lens camera with aperture area $\mathcal{A}$ (neglecting diffraction), $g(\mathbf{x}_i, \mathbf{x}_o) = \frac{\mathcal{A}}{L1^2}\delta(\mathbf{x}_i - M\mathbf{x}_o)$. Here $1/L1 + 1/L2 = 1/f$.

A "pinhole" camera including diffraction in the Fraunhofer approximation ($L1(2) > R^2/\lambda$ for aperture radius R) has impulse response $g(\mathbf{x}_i, \mathbf{x}_o) = \frac{1}{\lambda^2 L1^2 L2^2} |A(\frac{\mathbf{x}_o - M\mathbf{x}_i}{L2\lambda})|^2$. The expression is the same for a lens when $1/L1 + 1/L2 = 1/f$. A is the Fourier transform of $a(\mathbf{x})$. This impulse response is only meaningful for a fully incoherent object.

Sampling

The theorem of Nyquist and Shannon gives the sampling required to fully utilize the frequency content of a given impulse and describes the aliasing artifacts of a sampled image. Say we measure a 1-dimensional density $f(x)$ with impulses $g(x - nX), -N \leq n \leq N$. The sampled function is $f_s(x) = f(x) \sum_n g(x - nX)$. For simplicity take $g(x) = \delta(x)$. The Fourier transform is $F_s(u) = \frac{1}{X} \sum_n \delta(u - n/X) \star F(u) = \frac{1}{X} \sum_n F(u - n/X)$. Thus sampling causes replication of the Fourier transform at all points n/X. To recover $F(u)$ we must truncate $F_s(u)$ at $|u| = \frac{1}{2X}$, i.e. $F'_s(u) = F_s(u) rect(uX)$. If $|u_{max}| \leq \frac{1}{2X}$ we recover $F(u)$. Otherwise we lose information for $u > \frac{1}{2X}$ and $F'_s(u)$ is contaminated by the replications, causing aliasing artifacts. Thus the Nyquist-Shannon Sampling Theorem states *If a function x(t) contains no frequencies higher than B hertz, it is completely determined by giving its ordinates at a series of points spaced 1/(2B) seconds apart.* The (spatial) Nyquist frequency is $\frac{1}{2X}$. The result implies that "sinc interpolation" is optimal; $f'_s(x) = \sum_n f(nX) sinc(\frac{x - nX}{X})$.

It is worth noting that any finite object cannot be bandwidth limited so aliasing is inevitable. Sampling is chosen to reasonably suppress aliasing without generating a nearly useless excess of data. It is interesting that the diffractive inpulse response of an aperture is nominally bandwidth limited. The Fourier transform is the convolution of the aperture with itself. However 1. this is the paraxial approximation and 2. the image is finite so the bandwidth is in fact infinite.

For rectilinear sampling we apply the Shannon-Nyquist criterion in each dimension. In CT spatial sampling is radial and angular. For MRI our data-acquisition method produces a Fourier-space impulse response. We discuss sampling for these methods below.

Applications of Projection Imaging

The optical imaging camera is familiar. Film as a recoding medium has recently been replaced by pixelated semiconductor detectors. The fundamental resolution limitation in the image plane is set by the geometry of the camera and giving a maximum spatial frequency of roughly $D/(\lambda f)$ for lens diameter D and focal length f. The pixel separation should meet the S-N criterion for this spatial frequency. More pixels are useless.

X-ray projection imaging is based on small-focal-spot X-ray tubes and flat-panel detectors of several technologies, operating in the X-ray energy regime 30-100 keV. For general use the X-ray focal spot is $\sim$ 1 mm across, for mammography 0.1-0.3 mm. The combination of film with intensifying screens was the prevailing detector technology for a long time. The screen is a thin phosphor of calcium tungstate (K-edge 69.4 keV) or containing the rare earths lanthanum (38.9 keV)or gadolinium (50.2 keV). The latter have higher absorption in the usual diagnostic X-ray regime (40-50 keV) and produce more light. The film is made to match the screen emissions. The optical density of film has a "toe" at low exposure where it is relatively insensitive to exposure, producing poor contrast, followed by a regime where the optical density is proportional to the log of exposure (over an exposure dynamic range of about 10), followed by a shoulder where the film saturates. The spatial resolution of X-ray systems is given in line pairs per mm and film-screen systems are capable of resolutions as high as 20 lp/mm.
`http://www.e-radiography.net/radtech/f/film.htm`

Currently the photostimulable phosphor plate (CR) is probably the most common detector type, with spatial resolution of 5-10 lp/mm When the plate is exposed to Xrays, electrons are trapped in metastable states "color centers" in the crystal lattice. When the phosphor is then exposed to visible light from a scanning laser, visible light is emitted and detected to form an image. The plate is erased and reused. Another recording technique is xeroradiography, where a plate of charged selenium is exposed to the X-ray flux, which discharges it locally due to photoconduction. The resulting differentially charged plate attracts toner particles which are transferred to paper. Pixelated detectors for projection radiography (DR) are now made commercially with pixel pitch as small as 50 μm. GE uses CsI coupled to an amorphous Si photodiode array. Samsung has an analogous system (FPXD). There are CCD-based systems. The overall spatial resolution of projection radiography is $\sim$ 5 lp/mm for general purpose application and $\sim$ 20 lp/mm for mammography, which is still mostly film-based.

The background in projection imaging is Compton scattering from tissue. A (moving) collimator (Bucky grid) is used to reduce background. However most of the recorded photons are Compton scatters, resulting in very poor discrimination between regions of different attenuation coefficient. Pixel detectors with energy-measurement capability, e.g. medipix, may make it possible to partly suppress Compton-scattered photons. A difficulty is that at diagnostic X-ray energies (40-50 keV), Compton scatters are not greatly reduced in energy.

Nuclear Imaging

Nuclear imaging measures the spatial distribution of γ-emitting radionuclides. In nuclear medicine the radionuclide is generally conjugated to an entity with anatomical or physiological specificity, such as a bone-seeking molecule, a

particle that is trapped in the liver, spleen, bone marrow or lungs, a sugar, lipid or amino acid analog, etc. It is possible to image in the 100-300 keV regime by single γ projection and tomographic (SPECT) imaging. Positron emission tomography (PET) is accomplished using coincidence detection of the 511 keV annihilation photons. ^{99m}Tc (142 keV) and ^{123}I (159 keV) are suitable radionuclides for single γ imaging and ^{11}C and ^{18}F are the radionuclides mainly used in PET.

Nuclear projection imaging requires collimation for image formation. The pinhole collimator is used for selected studies but more often the imaging device is a multiple parallel hole collimator, which yields a nearly shift-invariant image of unit magnification. The detector is usually a large-area NaI crystal 1/4 in to 3/8 in thick, coupled to a close-packed array of photomultipliers. Position determination is obtained from the weighted averages of the photomultipler signals. The position resolution at the crystal is $\sim$ 3 mm FWHM but collimator geometry degrades the resolution to $\sim$ 1 cm FWHM. The geometrical efficiency is roughly 10^{-3}. The detection efficiency is reduced by the photofraction of NaI ($\sim$0.8 at 140 keV)).
http://en.wikibooks.org/wiki/Basic_Physics_of_Nuclear_Medicine/Nuclear_Medicine_Imaging_S

A commercial gamma camera based on CsI and silicon photodiodes has been developed (Digirad) and gamma cameras based on cadmium zinc telluride diode arrays are in development. For the latter a typical geometry is 2.5 mm square pixels that are 5mm thick. At 140 keV CZT has photopeak efficiency $\sim$ 65% that of NaI but has better energy resolution (5% vs 12%).

The background in nuclear projection imaging is mainly Compton scattered photons. These are suppressed by recording only photopeak events. At γ energies above 100 keV, energy discrimination is useful.

Isotopes and Chemistry

The important characteristics of nuclides for single-photon imaging are photon energy (100 - 200 keV), lifetime, suitability for chemistry and favorable radiation dosimetry. For dosimetric reasons and because the γ is monochromatic, the best single-photon emitters are metastable gamma emitters (notably ^{99m}Tc, 142 keV, $t_{1/2}$=6 h) and electron-capture nuclides (notably ^{123}I, 159 keV, $t_{1/2}$=13 h). Most isotopes are cyclotron produced, with the notable exception of ^{99}Mo, the precursor of ^{99m}Tc, which is made by irradiating highly enriched uranium in a reactor to yield ^{99}Mo, $t_{1/2}$=67 h, a fission product of ^{235}U. The ^{99}Mo is made into a generator system to yield chemically pure ^{99m}Tc. As produced this way, ^{99m}Tc is inexpensive and is the principal isotope for nuclear imaging. Synthetic methods are available to make a variety of tracers with suitable physiological properties for diagnostic imaging. ^{123}I is much more expensive and is used in thyroid imaging and to prepare tracers where ^{99m}Tc is unsuitable.

Tomographic Imaging

Tomography has nearly taken over medical imaging. I'll discuss X-ray CT, SPECT and PET. Other techniques include optical tomography of the breast and tomographic ultrasound.

The principle of CT is due to Radon (1917), who defined the Radon transform and provided a formula for its inverse.
`http://en.wikipedia.org/wiki/Radon_transform`
Proof of the projection-slice theorem in 2 dimensions
`http://en.wikipedia.org/wiki/Projection-slice_theorem`

The Fourier transform of the projection of an N-dimensional function f(r) onto an m-dimensional linear submanifold is equal to an m-dimensional slice of the N-dimensional Fourier transform of that function consisting of an m-dimensional linear submanifold through the origin in the Fourier space which is parallel to the projection submanifold.

For a 2D function the FT of a 1D projection equals the 2D FT along the corresponding diameter. If we determine the 1D projections of a 2D object of diameter 2R and compute the FT to maximum frequency ρ_{max}, the inverse FT will yield a 2D image with impulse response $\frac{\rho_{max}}{r} J_1(2\pi\rho_{max} r)$, which has negative lobes but becomes a δ function in the limit $\rho_{max} \to \infty$. The first zero of this impulse response is at $r = 0.61/\rho_{max}$. If ρ_{max} is equal to $\frac{1}{2X}$ in accordance with the sampling requirement, the full width of the impulse response is $2.44X$ and the FWHM approximately $1.22X$. Normally an additional filter, which attenuates high spatial frquencies, is applied to reduce the oscillation in the impulse response. There is substantial data at frequencies higher than $\frac{1}{2X}$ and a fixed-ring scanner significantly undersamples. Some earlier scanners increased the sampling by wobbling the rings during acquisition.

Sampling in CT

The sampling parameters in CT are the interval X in the projection and the angle Φ between adjacent projections. Within projections, $X \leq \frac{1}{2}\rho_{max}$ as above. We obtain the required angular sampling by considering a circular object with radius R. In Fourier space the frequency sampling interval must satisfy $|\Delta\rho| \leq \frac{1}{2R}$. The sparsest sampling is at the periphery of Fourier space where $\Delta\rho = \rho_{max}\Phi \leq \frac{1}{2R}$. Then $\Phi < \frac{1}{2\rho_{max}R}$. Overall CT is oversampled by $\times 2$.

X-ray CT
`\http://en.wikipedia.org/wiki/X-ray_computed_tomography`

Mainly because the geometry of CT greatly reduces Compton scattered photons, CT has much better absorption coefficient resolution than X-ray projection imaging. Absorption coefficient is expressed in Hounsfield units (HU), where air

is -1000 and water is 0. Bone has values between 1000 and 3000, liver 50, kidney 30, fat -70, blood 40, inflated lung -900. A typical scan measures HU with uncertainty less than 5 units, compared to no better than 100 units for X-ray projection imaging.

An X-ray CT scanner
http://en.wikipedia.org/wiki/File:Ct-internals.jpg

The scanner consists of a collimated X-ray tube and detector array that rotate helically around the patient. The tube produces a fan beam or cone beam, which together with the detectors selects the slices. Modern CT scanners simultaneously scan up to 320 1 mm slices and rotate with a period as short as 1/3 sec. Detectors are selected for high efficiency and short "afterglow". The detector array is typically an array of scintillators (BGO or CdWO4) coupled to photodiodes. Xe ionization chambers have been used. The background in CT is mainly Compton scatters in tissue. Collimation in CT reduces background to a level much less than in projection imaging. An example of a state-of-the art CT scanner is the Toshiba Aquilion.
http://www.medical.toshiba.com/products/ct/dynamic-volume/clinical-cardiac-01.php

SPECT

SPECT (Single photon emission computed tomography) is an incremental improvement over nuclear projection imaging, where one or two conventional γ cameras with their collimators rotate around the subject to acquire projections. The tomographic algorithm is used to reconstruct 3D images. The LORs are defined by the collimators and are cone-shaped. The attenuation is depth-dependent within an LOR and a direct measurement cannot be obtained. However SPECT images are frequently more informative than projection images and the method is widely applied.

Positron Emission Tomography (PET)
http://en.wikipedia.org/wiki/Positron_emission_tomography
mandelkern_ann_rev_nuc_part_sci_45:205_1995
phelps_ann_rev_nuc_part_sci_52:303_2002

PET exploits the coincident 511 keV photons from positron annihilation to define tomographic lines of response (LOR). The following positron-emitting nuclides are especially suitable for PET: ^{11}C,^{13}N, ^{15}O, ^{18}F. As isotopes of elements that are found widely in tissues they can be incorporated into molecules which then retain physiological activity. They can be made by bombardment of readily available nuclear targets with protons of energy < 15 MeV with reasonably large cross sections.They have suitable lifetimes, long enough to perform organic syntheses but not so long to produce excessive radiation doses. ^{18}F is the most useful of these because of its 110 minute halflife, low maximum positron energy of 635 keV, and small atomic size, which permits it to replace a hydrogen atom

without untoward steric effects. It is made by the ^{18}O(p,n)^{18}F reaction, which has a very large low-energy cross section. ^{11}C has a maximum positron energy of 970 keV and halflife of 10 minutes and is made by bombarding ^{14}N with protons. It is widely used in research. ^{15}O and ^{13}N have special applications as do several heavier isotopes that cannot be made in small cyclotrons, including ^{124}I, ^{62}Cu, ^{68}Ga, ^{82}Rb. The latter three nuclides are conveniently obtained the decay of longer-lifetime nuclides in generator systems.

There are two sources of intrinsic spatial resolution uncertainty. One is due to the distance travelled by the positron before annihilation. The 635 keV e^+ from ^{18}F yields a vertex distribution with FWHM 0.1 mm and FWTM 1.03 mm The 970 keV e^+ from ^{11}C gives FWHM 0.19 mm and FWTM 1.86 mm. The annihilation photons are noncollinear with a roughly Gaussian angular distribution with FWHM 0.58^o, which for a scanner ring of 80 cm diameter yields a distribution with FWHM of 4 mm for annihilations in the center of the ring, narrower for peripheral events.

Iwata et al.,Phys Rev A55:3586, 1997.
Levin Phys Med Biol 44:781, 1999.

The PET scanner consists of a cylindrical array of detectors with a system for detecting coincidences and selecting photopeak events. A human scanner is typically 80 cm in diameter and 15 cm in height. Coincidences are made with resolving time 5-10 ns. The amount of activity in the field of view is $\sim$1 mCi (3.7×10^7 sec^{-1}) or less and the geometrical efficiency for this geometry is about 8%.

The geometrical efficiency of a scanner is approximately proportional to cylinder height squared and inversely proportional to cylinder radius. However spatial resolution suffers for small cylinder radius because of geometrical parallax, which limits detector depth reducing efficiency. Depth of interaction information redresses this limitation.

The backgrounds in PET consist of accidental coincidences and coincidences where one or more of the γs Compton scatter in the subject. Background rates are crucially dependent on the detector characteristics.

Detectors

Since 511 keV is well above the K-edge for all elements and the coincidence rate is quadratic in the detector efficiency, we require much thicker detectors than for nuclear projection imaging. The scintillators with the best stopping power have $\lambda \geq 1$ cm. Detector elements are typically 4-6 mm transversely and 20-30 mm deep for human scanners and $\sim$ 2 mm transversely and 10 mm deep for animal scanners. It is important that the detectors have high linear absorption coefficient to minimize geometrical parallax, high photofraction for 511 keV to reject scatter-background, and high light output and fast timing for accidental-background suppression . The best available scintillator is LSO(Ce), with mean free path and photofraction for 511s of 1.2 cm and $\sim$ 80% respectively, fast light output of 31500 photons/MeV (roughly 80% of NaI), and

fast decay time of 37 ns . The search for better scintillators continues. We currently detect the scintillation light with photomultipliers in a multiplexed design. Fast semiconductor light detectors are potential successors, notably the silicon photomultiplier, which is extremely thin and correspondly fast, is efficient and produces very low noise.

"Parallax" and depth of interaction (DOI)

A significant limitation to spatial resolution in PET is due to the oblique orientation of the detectors for LORs away from the scanner origin, causing errors in defining the LOR of the coincidences. The use of semiconductor ohotodetectors makes it possible to instrument both ends of the detectors in order to determine the depth of the photoelectric interaction (DOI). A module consisting of CMOS silicon photomultipliers (SSPM) with a LYSO scintillator is shown to yield a depth of interaction measurememt with FWHM$<$ 3 mm.
`Dokhale et al., 2009 IEEE NSS Conference Record M05-268.`

Attenuation correction in PET

The linear attenuation coefficient of 511 keV photons in water is $(10.5\text{ cm})^{-1}$ (HVL 7.3 cm) [at 140 keV 6.7 (4.6)]. LOR passing through the center of the body are substantially attenuated. An exact correction can be obtained by recording CT LOR at 511 keV for each PET LOR. This can be done by moving a 511 keV source, e.g. Ge-68/Ga-68, in a helical path external to the body but within the detector ring, so that every LOR is sampled. Most PET scanners are now coupled to CT scanners and the CT data is energy extrapolated to obtain PET attenuation factors. Scatter is a sigificant problem for PET, both for the emission and transmission scans and how best to correct is an open problem.

Accidental background in PET

Accidental background is due to annihilation γs from different annihilations arriving within the coincidence window. Relative to the "trues" rate, accidental background is proportional to activity and to the coincidence resolving time. It is usually corrected by collecting data in a delayed coincidence window and subtracting the accidental sinogram from the trues sinogram level prior to reconstruction. A consequence is that the subtracted sinogram bin populations are no longer Poisson.

Scatter background in PET

Compton scatter in the patient is the most difficult background in PET. It is suppressed by selecting photopeak events so that high-Z scintillators that give a sharp photopeak are preferred. It is essential to model the remaining background so it can be subtracted from the image. This is done iteratively but currently is not fully successful.

It is worth noting that because of the relatively small linear attenuation of 511s in tissue, for small-animal PET (mice and rats) scatter background is small. Reasonable images can be made from events collected without an energy threshold, i.e. including detector Compton scatters.This allows the use of detectors with a small photofraction. Alberto del Guerra and Guido Zavattini developed an animal scanner based on YAP.

Reconstruction in PET

The Fourier reconstruction method implicit in the projection-slice theorem does not take into account the statistical character of the LOR measurements. For methods with extremely high counting statistics such as X-ray CT, Fourier reconstruction is pretty good and a variant of it "filtered back-projection" is generally used.

More generally the LOR measurements are significant inconsistent due to statistical variability and the maximum likelihood method is appropriate to best estimate the parameters (the voxel values) characterizing the object. This is especially a problem for nuclear methods - SPECT and PET. However in an effort to reduce radiation exposure in xray CT, statistical methods for reconstruction are now being applied. For $\mathbf{f}$ the column vector of voxel intensities and $\mathbf{g}$ the column vector of Poisson measurements such that $E[\mathbf{g}] = \mathbf{af}$, the log likelihood function $w(\mathbf{f}) \sim \sum_i (g_i ln(\sum_j a_{ij} f_j) - \sum_j (a_{ij} f_j))$, yielding $\sum_i a_{ij} = \sum_i \frac{g_i a_{ij}}{\sum_k a_{ik} f_k}$. These equations are extremely complicated and practically insoluble. The usual method of "solution" is the iterative ML-EM procedure, i.e. $\hat{f}_j^{\,n+1} = \hat{f}_j^{\,n} \sum_i \frac{g_i a_{ij}}{\sum_k a_{ik} \hat{f}_k^{\,n}} / \sum_l a_{lj}$. It was proved by Dempster et al., J Roy Stat Soc B39, 1-78, 1977, that the successive iterations lead to monotonic increases of the Likelihood Function. In this method we make successive forward projections of intermediate solutions followed by multiplicative corrections. The solution of interest is specified by the set of f_j for which the initial values are set finite positive, since any initial $f_j = 0$ remains zero after multiplication with the positive-definite correction factor. (The f_j are non-negative Poisson means.) However the approach to the ML solution is slow and this method is not suitable for the very large data arrays of medical tomography (128x128 and larger). An accelerated method named OS-EM [Hudson and Larkin, (1994) Accelerated image reconstruction using ordered subsets of projection data, IEEE Trans. Medical Imaging, 13 (4), 601-609] is much faster but doesn't quite converge. It has become the standard but isn't entirely satisfactory and has inherent biases. This is an area of active research but little progress has been made recently.

3D Reconstruction

We have implicitly been discussing tomographic reconstruction for LOR in a plane. However the vast majority of LOR in a multiring scanner do not lie in transaxial planes. To the extent that these LOR can be placed in fully

sampled oblique planes they can can be reconstructed to yield voxel activities that are independent measurements and can be statistically combined with the measurements from the transaxial planes. Howeer most of the oblique LOR lie in incompletely sampled planes that cannot be directly reconstructed. There are a number of approaches. One can derive information from the oblique LOR that can be added to the already fully sampled transaxial planes and proceed with Fourier single-plane reconstruction. This method is called Fourier rebinning. `Defrise et al., IEEE Trans Nuc Imag 16:145, 1997.` Iterative maximum likelihood reconstruction in 3 dimensions is very computer-intensive but is implemented in commercial scanners equipped with substantial parallel computing resources.

Statistics of CT

For CT the statistical errors in voxel estimates are greater than for projection imaging with the same amount of data, because each measurement contains one-dimensional rather than two-dimensional information, localized only to all n voxels along an LOR. For all reconstruction methods an $n \times n$ image has voxel uncertainty greater than a projection image with the same number of counts by about $\sqrt{n}$.

Time of Flight PET

The uncertainty in determining the annihilation point is $\delta x = c\delta\Delta t/2$, where Δt is the time difference between detector pulses. If we localize the annihilation point within Δx for a uniform positron-emitting distribution of diameter D, we decrease the voxel statistical uncertainty by a factor $\sqrt{D/\Delta X}$. Time of Flight PET is implemented using the fast scintillators LSO and LYSO, where the FWHM is $\sim$ 600 ps giving $\Delta x \sim 9$ cm, for a factor 2.1 reduction in voxel statistical unertainty for an object of 40 cm diameter. (The activity distributions of the heart and brain are roughly 10 cm in diameter.)

EM reconstruction for TOF PET is done with a straightforward generalization. The data consists of g_{ik}, where k identifies the time bin for the ith LOR. The log likelihood function becomes $w(\mathbf{f}) \sim \sum_{ik}(g_{ik} ln(\sum_j a_{ij}b_{kj}f_j) - \sum_j(a_{ij}b_{kj}f_j))$, where b_{kj} is the probability that a count from voxel j in LOR i appears in the k time bin.

Better scintillators, possibly lanthanum bromide $LaBr_3$(Ce), which has 160% the light output of NaI and $\tau_d < 20$ns, better photodetectors such as silicon photomultipliers,and better electronics, will make TOF PET better.

PET Radiochemistry

PET radioisotopes have short lifetimes and are made in local cyclotrons. Several $^{[}18]$F tracers are made commercially but PET research centers generally operate cyclotrons. There are methods for producing isotopes with high specific activity in chemical forms suitable for subsequent synthesis. The synthetic process, purification and quality control are performed within 1 or 2

half lives. Automated systems that couple to the cyclotron and synthesize radiopharmaceticals are available commercially. These are effective in delivering consistent syntheses and in avoiding high radiation doses to personnel. However the current technology is suitable for bulk synthesis, for several patients. FDG can be synthesized in quantities suitable for an entire day of scans.

Clinical PET currently uses non-specific tracers such as ^{18}F-FDG. However in the future one envisions tailor-made tracers, possibly based on the specific receptors associated with a patient's illness, and different for every case. We require methods for making short cyclotron runs to produce enough isotope for one dose, followed by a fully automated synthetic method using a fully charged and disposable microfabricated reaction system. Proof of principle has been established for the use of microfluidic technology in PET.
`Gillies et al., Appl Radiat Isot 2006 Mar;64(3):333-6`

Proton bombardment is required to make the proton-rich nuclei that decay by positron emission. Carbon-11 is made by the ^{14}N(p,α)^{11}C reaction and fluorine-18 by the ^{18}O(p,n)^{18}F reaction, both with incoming beam energy 10-20 Mev. The cross section for the ^{18}F reaction is especially large giving a saturation yield for 11 MeV protons of 180 mCi/μA. A typical target can handle at least 10 μA for Curie-quantity yields. The desired chemical form for the synthesis is often achieved during the bombardment. For example a trace of oxygen is included to make $^{11}CO_2$. Reduction to CH_4 is accomplished in the target or by passing the CO_2 with H_2 over a catalyst. The ^{18}O(p,n)^{18}F reaction in a ^{18}O-liquid-water target produces $^{18}F^-$ ion in solution, suitable for nucleophilic fluorination. Electrophilic fluorination employs ^{18}F-^{19}F gas, which can be made by ^{20}Ne(d,α)^{18}F or by ^{18}O(p,n)^{18}F in a $^{18}O_2$ gas target applying a second bombardment after removing the oxygen gas and refilling the target with $^{19}F_2$. This reagent necessarily has very low specific activity.

Specific activity (SA) quantifies the radioactive fraction of the tracer molecules, quoted as Ci/μmol. Because of chemical contamination of the environment and relatively short radionuclide lifetimes it is fractionally small. For pure ^{18}F the specific activity is about 150 Ci/μmol. We generally achieve up to 10 Ci/μmol. The SA drops by 2 for every half life, here 110 minutes. For ^{11}C the maximum possible SA is about 30 Ci/μmol and reaching 5 Ci/μmol is quite good, the SA falling by 2 every 20 min.

The clinical and research value of PET

PET uniquely has the capability of making diverse quantitative physiological measurements noninvasively in an imaging context. 1. The radiotracer method has extraordinary sensitivity to picomolar concentrations of labeled molecules. 2. PET is fully quantitative because of its accurate attenuation correction. 3. PET tracers contain radioisotopes of the light elements carbon, nitrogen, oxygen and fluorine, which allows maintenance of fairly accurate physiological activity.

Clinical applications of PET are currently semiquantitative. By far the most common use is oncology, followed by cardiology and neurology and these methods are mainly based on ^{18}FDG. This tracer illustrates an important principle, that suitable tracer molecules are almost always analogs of physiological substances with restricted physiological activity. Labeled ^{11}C-glucose can be made but it is a useless tracer because it is rapidly metabolized by tissues and the label excreted as $^{11}CO_2$, preventing localization and quantification of metabolism. DG (2-deoxy-glucose) is missing the oxygen on the 2 carbon of glucose. This modification allows transport into cells by the glucose transporter and subsequent phosphorylation to DG-6-P. At that point metabolism stops and the DG is effectively trapped in cells where it is observed by virtue of ^{18}F decay.

Clinically we work with the relative uptake of tracer in tissue, quantified as the specific uptake value (SUV), uptake per gram of tissue divided by total dose calibrated to patient size. However to quantitatively assess the transport and phosphorylation of glucose and determine the metabolic rate for gucose metabolism, grams of glucose metabolized per gram of tissue per minute, we model the metabolism of FDG and experimentally relate the FDG parameters to those of ordinary glucose. In particular the equations for DG kinetics are:

$$dC_1(t)/dt = K_1^* C_p(t) - k_2^* C_1(t) - k_3^* C_1(t) + k_4^* C_2(t) \tag{1}$$

$$dC_2(t)/dt = k_3^* C_1(t) - k_4^* C_2(t) \tag{2}$$

Here $C_1(t)$ is the tissue concentration of DG, $C_2(t)$ is the tissue concentration of DG-6-P and $C_p(t)$ is the measurable vascular concentration of DG. K_1^*, k_2^*, k_3^* and k_4^* are the parameters to be measured. The solution of these equations is straightforward and the parameters are found by nonlinear regression. From experiment it has been shown that K_1^*, k_2^*, k_3^* and k_4^* are related in a consistent way to K_1, k_2, k_3 and k_4, the corresponding parameters for glucose metabolism. The capacity of tissue for glucose metabolism is so large that this is a true tracer measurement. The parameters are unaffected by the introduction of the tracer.

A similar measurement is that of the tissue concentration of receptors. All tissues are populated by specific receptors for neurotransmitters, hormones and other molecules produced by the body for signaling and modulation. In the brain, signal transmission between neurons is modulated by the release of transmitters at synapses. The model for tracer uptake is characterized by:

$$dC_1(t)/dt = K_{in} C_p(t) - k_{out} C_1(t) - k_{on}(B'_{max} - C_2(t)/f)C_1(t) + k_{off} C_2(t) \tag{3}$$

$$dC_2(t)/dt = (B_{max} - C_2(t)/f)C_1(t) - k_{off} C_2(t) \tag{4}$$

Here $C_1(t)$ is the tissue concentration of free tracer, $C_2(t)$ is the tissue concentration of bound tracer and $C_p(t)$ is the vascular concentration of tracer. f is the fraction of molecules chemically identical to the tracer that are labeled. K_{in}, k_{out}, k_{on}, B'_{max} and k_{off} are the parameters to be measured. This system is nonlinear but if $C_2(t)/f$ is small compared to B'_{max} the nonlinear term can be dropped. B'_{max} is generally a very small concentration in the picomolar to nanomolar range so f is constrained to be large, a high specific activity. We

have effective tracers for many of the important receptor systems, including the acetylcholine, dopamine, serotonin, opioid systems. Unlike glucose metabolism, which is present in all tissues, receptors are highly tissue and function-specific. Many diseases such as cancers have unique receptors. We already diagnose certain brain diseases, such as Parkinsonism, by identifying receptor deficits. A probably future application of PET is diagnosis using patient-specific receptor tracers. We need to make our synthetic methods more efficient to make this practical.

An application that relies upon the very high sensitivity of PET is imaging transgene expression. A PET reporter gene (PRG) is attached to the therapeutic gene to be transferred. It too is incorporated into the genome of the recipient organism and is expressed. The expression signals the success of the the therapeutic gene transfer. One type of PRG causes expression of an enzyme that causes a PET tracer to be trapped in certain cells. Another leads to specific receptor synthesis that causes a PET tracer to bind to certain cells.

Magnetic resonance imaging (MRI)

MRI employs sequences of magnetic-field gradient and RF pulses to manipulate the proton magnetization in material, so that the resulting signal due to the precessing magnetization is proportional to the free proton density weighted by the longitudinal (T1) and transverse (T2) relaxation times and other properties of the magnetization. It can be applied to other magnetic nuclei, including ^{23}Na, ^{13}C, ^{31}P. The usual acqusition method uniformly measures, on a point-by-point basis, the 2-dimensional FT of planar slices on a square grid of dimension $2k_{max} \times 2k_{max}$. The impulse response is then $\text{sinc}(k_{max}x)$ and the sampling requirement is that $\Delta k \leq \frac{1}{2X}$ where the object is $2X$ in overall dimension. Aliasing artifacts appear as repeats of the image or "wrap around" , centered at spatial intervals $\frac{1}{k_{max}}$. This implies that we have no artifacts if we satisfy the sampling requirement.
`http://en.wikipedia.org/wiki/Magnetic_resonance_imaging`

MRI images contain physiological information to the degree that proton density, T1 and T2 reflect tissue chemistry. This is extremely valuable for discriminating tissue types, for example gray vs white matter in the brain. A developing area of MRI is the use of magnetonanoparticles (MNP) to label physiological compounds to produce tracers. The magnetoparticles are made of iron oxide and dextran and produce very large signals in T2-weighted sequences. There is evidence that MNPs conjugated to amino acids or sugars have physiological acivity. Since the MNPs are much larger than the molecules to which they are attached it is unlikely that quantitative measures of metabolism are possible. For most neural and hormonal receptors the concentrations in tissue are so small that it is unlikely that MRI techniques are sufficiently sensitive.

Electroencephalography (EEG) and Magnetoencephalography (MEG)

The sources of the EEG and MEG are synchronized currents flowing in the aligned apical dendrites of neurons in organized layers of the cerebral cortex. At least 10,000 aligned neurons are required to produce an observable signal.

The EEG is modeled as the potential of a curent dipole density, and is substantially attenuated by the induced charge densities in the skull and scalp. Measured EEG is in the 10-100 μV regime. In EEG we apply 32-128 electrodes to the surface of the scalp and record the electric potential. An image of the brain primary current density is obtained taking into account the conductivity of the brain, skull and scalp. EEG is effectively limited to the surface of the brain because of the relatively high tissue conductivity. It is sensitive only to radial current dipoles which appear in the brain gyri.
`http://en.wikipedia.org/wiki/Electroencephalography`

MEG is described as the magnetic field of intracellular ionic currents flowing in the aligned apical dendrites. Extracellular return currents substantially contribute to the observed field. MEG is observed with up to 300 SQUID gradiometers at the scalp surface. It is primarily sensitive to transverse current dipoles, which are present in the cortical sulci. MEG fields are 10-1000 fT and are much smaller than ambient magnetic noise.
`http://en.wikipedia.org/wiki/Magnetoencephalography`

Image processing

Image processing is used to adapt the imaging process to the known properties of the object and to the purpose of imaging. We tailor the process to specific features of the signal, reduce image "background" compared to signal, and/or detect specific features in the image. Standard image-processing methods are crucially dependent on our premise of shift-invariant imaging. Only if the imaging system is shift-invariant is the signal fully characterized by an impulse response that is a function of a single coordinate-space vector, or equivalently by a transfer function dependent only on a single frequency-space vector. Under these conditions we generally "filter" in frequency space. Simple signal processing is done by applying a low-pass, high-pass, band-pass, or bandstop filter that favors signal frequencies compared to background frequencies. More sophisticated methods optimize the filter, based on detailed knowledge of the signal and background frequency spectra. The best-known of these is the Wiener filter, which is developed in the context of the theory of stationary stochastic processes.
`http://en.wikipedia.org/wiki/Stochastic_process`
`http://en.wikipedia.org/wiki/Stationary_process`
`http://en.wikipedia.org/wiki/Ergodic_process`
`http://en.wikipedia.org/wiki/Wiener_filter`
`http://en.wikipedia.org/wiki/Matched_filter`

In a stochastic (random) process the sample is a function in some discrete

or continuous sample space, e.g. time or n-dimensional space. The set of sample functions can be finite or infinite and may be determined by one or more random variables. For example the height of the function at each point can be random. In imaging the observed image is a particular sample of a random process. The discrete voxel values are random variables, e.g. multinomial, where the moments are determined by the non-random object and imaging process. A stationary random process is one where the joint pdf, and thus all expectation values such as moments are invariant under translation in the sample space. White noise is stationary. However a signal that appears preferentially at certain places is not. A random walk is a non-stationary stochastic process. The mean is zero everywhere but the variance increases with distance from the starting point. A weaker form is wide-sense stationary, where only the first and second moments must be translation invariant. An ergotic random process is one where an expectation value is equal to the corresponding "time average", i.e. all the moments can be obtained from a single sufficiently long sample. An ergotic process is necessarily stationary.

A stationary process useful in imaging is the Poisson process, $p(t) = \sum_i^n x(t-t_i)$. Here $x(t)$ is the shape function of the pulses, n is Poisson with mean $E[n]$ and t_i are uniform on some interval. The signal is modeled with shape $s(t)$ and mean $E[n_s]$ and the background with shape $b(t)$ and mean $E[n_b]$. Filters are best described in frequency space. To obtain the frequency content of a process we use the Wiener-Khinchin theorem which states that the power spectral density of a wide-sense-stationary random process is the Fourier transform of the corresponding autocorrelation function.
`http://en.wikipedia.org/wiki/Wiener-Khinchin_theorem`

We model the overall random process $x(t)$ as the signal process $s(t)$ plus a wide-sense-stationary additive background process $b(t)$, $x(t) = s(t) + b(t)$, and construct a linear shift-invariant filter $g(t)$ such that $\hat{s}(t) = g(t) \star [s(t) + b(t)]$. If $s(t)$ is wide-sense stationary the best-known optimum filter is the Wiener filter, which minimizes the mean squared error (MSE) $E[(\hat{s}(t) - s(t))^2]$ in order to best estimate the signal. In frequency space $G(\nu) = S_{x,s}(\nu)/S_x(\nu)$, where $S_{x,s}(\nu)$ is the FT of the cross-correlation between $x(t)$ and $s(t)$ and $S_x(\nu)$ is the autocorrelation of $x(t)$. If $s(t)$ and $b(t)$ are uncorrelated (the common case), $G(\nu) = S_s(\nu)/(S_s(\nu) + S_n(\nu)), \nu > 0; G(\nu) = 1/2, \nu = 0$.
`http://en.wikipedia.org/wiki/Wiener_filter`
If $s(t)$ is non-stochastic, we have the matched filter, which maximizes $E[(g(t) \star s(t))^2]/E[(g(t) \star n(t))^2]$ in order to optimally detect the known signal.
`http://en.wikipedia.org/wiki/Matched_filter`
A further generalization is the adaptive filter, which self-adjusts based on the input data.
`http://en.wikipedia.org/wiki/Adaptive_filter`

For random processes where one can define a dynamical principle relating the "state" (statistical mean) of a process at successive times the Kalman fil-

ter is useful. In time-lapse photography, for example, the principle can be that the current state (mean image) $\mathbf{x}_k$ is the same as the previous state $\mathbf{x}_{k-1}$. The Kalman filter uses previous measurements to predict an *a priori* state estimate $\hat{\mathbf{x}}_k^-$, which is then updated by the current measurement $\mathbf{z}_k$ to yield the *a posteriori* state estimate $\hat{\mathbf{x}}_k$. The MSE $E[(\mathbf{x}_k - \hat{\mathbf{x}}_k)^2]$ is minimized. `http://rsbweb.nih.gov/ij/plugins/kalman.html`

An example of filtering as an intrinsic part of the image-creation process is tomography. We can think of FBP as simple back-projection followed by the application of a shift-invariant filter. 2-dimensional projection followed by simple back-projection is a shift-invariant operation with impulse response $1/r$. We restore the simple back-projection image by convolving with $[1/r]^{-1}$. This is simple in frequency space since the 2-D FT of $1/r$ is $1/\rho$ thus the filter is ρ, which of course must be truncated to control noise at high frequency.