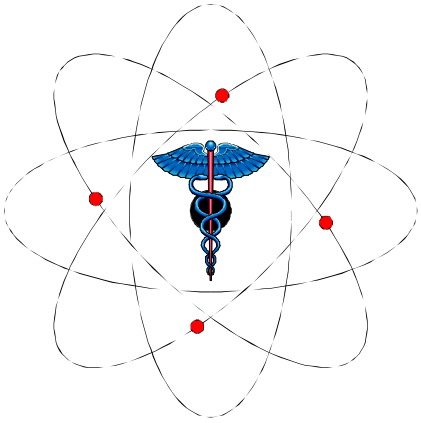
PARAPET



Determination of a Set of Existing Algorithms For Comparative Evaluation

Document: D/TECHNION/LEVKOVITZ-ZIBULEVSKY-FALIKMAN /GENEVA/LABBE-ZAIDI-MOREL/1.3/05.05.1997/1/1.1

FINAL DRAFT

Approved by:

C. Labbé
(Author)

F. Wray
(Project manager)

CONTENTS

GLOSSARY

ANALYTICAL ALGORITHMS

(x,y) : Coordinates in object space

(ν_x,ν_y) : Coordinates in frequency space

R : Radius of field of view

z : Axial coordinate

θj : Co-polar angle (i.e. aperture angle between an LOR and the transaxial plane j)

ϕ : Angle with respect to x

ω : Azimuthal angle

s : Radial distance where $s = x \cos\phi + y \sin\phi$

ν_s : Coordinate on a frequency axis

***f*(x,y)** Imaged object (i.e tracer distribution)

$F_r(\nu)$: Reconstructed Fourier transform of the image for all frequencies ν

$p_\phi(s)$: one row of the sinogram corresponding to differents values of s for a fixed value of ϕ

$P_\phi(\nu_s)$: Fourier transform of $p_\phi(s)$

$\hat{u}$: Unit direction vector specified by ϕ and a polar angle $\pi/2 - \theta$

p($\hat{u}$, s) : 2D parallel projection of f(x,y)

$p^F(\hat{u}, s)$: Filtered projection

h($\hat{u}$,s) : Convolution kernel

H($\hat{u}$,s) : 2D Fourier transform of the kernel h($\hat{u}$,s)

Ψ : Polar angle of the frequency vector $\vec{\nu}$ (i.e. $\nu_Z/|\nu| = \cos\Psi$

N_{rings} : Number of rings

δ : Ring difference between i and j

s_k : Sinogram corresponding to plane k

$S_{i,j}$: Sinogram between pairs of rings i,j

ITERATIVE ALGORITHMS :

λ_j : Local radioactivity distribution in pixel j

y_i : Numer of observed counts collected in bin i

n_j : Number of photons in pixel j

$\nu=(x,y,z)$: Unit box (=voxel)

$\lambda(\nu)$: Emission density in voxel ν

b_{ij} : Single bin containing all events detected by two detectors d_i and d_j

$p(\nu,b)$: Probability of an emitted events from voxel ϖ detected in bin b

$P(y/\lambda)$: Posteriori distribution

$\mu(b)$: Means of the Poisson variables y(b)

θ : Step size for OSEM algorithm

p_{nk} : Probability density function that relate the number of photons emitted from pixel (n) to

$R_{k(\theta)}$: Calculated ray sum within ray $k(\theta)$

μ^k : Relaxation parameter for the k^{th} iteration

$N_{k(\theta)}$: Number of pixels lying along the ray $k(\theta)$

1. INTRODUCTION

This document is the main deliverable for Task 1.3 of the PARAPET project. The objective of this task is to survey existing work and practice in order to choose a set of two- (2D) and three- (3D) dimensional reconstruction algorithms currently in use which can form the basis for a comparative assessment of the algorithms developed and enhanced during the project. The establishment of this base set of algorithms, using both analytic and iterative methods, will facilitate the assessment of the quality of reconstructions achieved by new algorithms.

At present, the most widely used method for 2D and 3D image reconstruction in PET is filtered backprojection, which is a direct analytical method. This technique is very quick for 2D reconstruction, but the images tend to be "streaky" and display interference between regions of low and high tracer concentration.

There is a general trend to reduce radiation dosages to minimise patient exposure which reduces photon counts. It is expected that this trend will continue. Low administered dose are compensated in part by a move to 3D scanning modes which greatly increases data volumes with consequent increase in reconstruction times. The general problem in nuclear medicine of low counting statistics favours the use of iterative techniques. Such techniques may require up to 10^{13} floating point operations to reconstruct a detailed PET volume. To do this in a clinically acceptable time frame requires two orders of magnitude speedup in computer performance over that available from a current powerful workstation.

2. REVIEW OF ANALYTIC ALGORITHMS

The analytic solution to the two-dimensional inversion problem of recovering a 2D image $f(x,y)$ from the set of its one-dimensional projections $p(s,\phi)$ has been known for nearly 80 years [Radon et al., 1917] and is presented below as a useful introduction to the more complex three-dimensional problem. This derivation follows that of Bracewell [Bracewell et al., 1967]. The existing analytic algorithms fall into two categories. The first category includes exact algorithms obtained by discretizing analytical inversion formulae of 3D X-ray transform (3DRP, FAVOR). The second category of algorithms includes approximate algorithms which factorise the 3D reconstruction problem into a set of independent 2D reconstruction and are therefore considerably faster than the exact algorithm (FORE, SSRB, MSRB).

2.1. "FBP": Filtered backprojection

2.1.1. Solution of the 2D problem

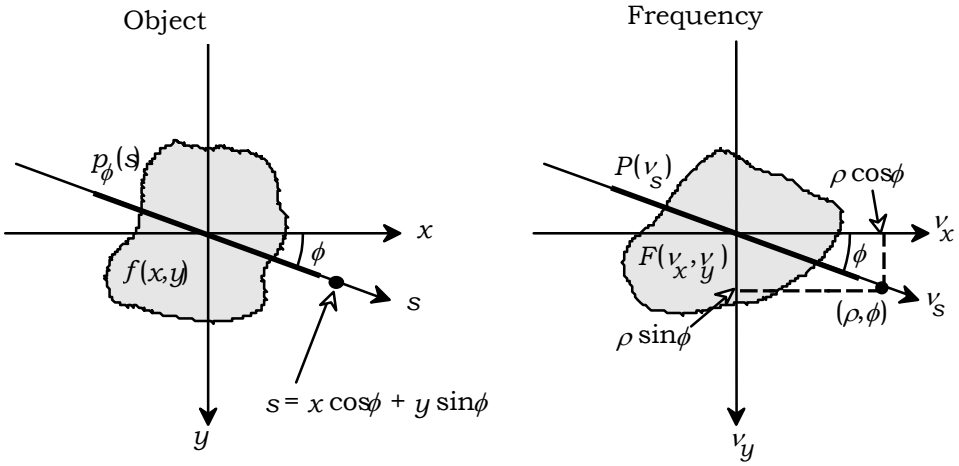
The key result that links the imaged object $f(x,y)$ to its projections $p_\phi(s)$ is the central-section theorem, which can be expressed as, in 2D,

$$P_\phi(\nu_s) = F(\nu_s \cos\phi, \nu_s \sin\phi) \qquad \text{(Eq. 1)}$$

where capital letters denoting Fourier transforms, and ν_s the coordinate on a frequency axis rotated by ϕ relative to the x-frequency axis, the ν_x-axis. A proof of this theorem is given in [Natterer et al., 1986] by the central-section theorem (Eq. 1) and Fig. 1 is an illustration of the various coordinates used in the following derivation. This yields the following final result:

$$f(x,y) = \int_0^\pi d\phi \left[\int_{-\infty}^{+\infty} d\nu_s \, |\nu_s| \, e^{2\pi i \nu_s s} \, P_\phi(\nu_s) \right], \quad s = x\cos\phi + y\sin\phi \qquad \text{(Eq. 2)}$$

From this equation, $f(x,y)$ may be recovered from $p_\phi(s)$ in two steps: (i) the projections are filtered by



the ramp filter $|\nu_s|$ in Fourier space; and (ii) the contribution to the image grid of each of the filtered projections is computed using $s = x\cos\phi + y\sin\phi$. This last step is called backprojection.

Figure 1: *Coordinates used in two-dimensional reconstruction. In image space, the axis s perpendicular to the direction of projection is rotated by ϕ with respect to x, and $s = x\cos\phi + y\sin\phi$. Similarly, in frequency space, the axis ν_s is rotated by ϕ with respect to ν_x, and polar co-ordinates (ρ,ϕ) are introduced, with $\nu_x = \rho\cos\phi$ and $\nu_y = \rho\sin\phi$. This figure also illustrates the 2D central-section theorem: a section through F at an angle ϕ is equal to the Fourier transform P_ϕ of the projection p_ϕ of f at the same angle ϕ.*

2.1.2. Solution of the 3D problem

Let us now recall the essential difference between two- and three-dimensional image reconstruction from projections. In 2D, line integrals determined by two parameters, a slope and an

intercept, are used to reconstruct an unknown tracer distribution, function of two spatial variables $f(x,y)$. The problem is analytically fully determined and the solution exists and is unique. In 3D, however, integration lines are specified by four parameters, two slopes and two intercepts in our case, but the image to be reconstructed depends on three spatial variables only, $f(x,y,z)$. The problem is greatly over-determined, since analytically the set of *direct* projections – projections perpendicular to the scanner axis – is sufficient to recover $f(\vec{x})$ in the conventional slice-by-slice fashion. The additional information on oblique projections are analytically superfluous, as long as we are not concerned with the estimation of statistical variations. As a consequence the recovery procedure is no longer unique, and an infinite number of analytically equivalent reconstruction methods may be devised, whose behaviour in real cases, i.e. in the presence of noisy measurements, may nevertheless vary substantially. Following Defrise et al. (1995), we shall now show that a recovery condition may be derived, which must be satisfied by any exact analytical algorithm.

The central section theorem is the key result that links the projections of the image to the image itself, just as in two dimensions: a section of the 3D Fourier transform of the image through the origin and perpendicular to $\hat{u}$ is equal to the 2D Fourier transform of the projection of the image along $\hat{u}$. That is,

$$F(\vec{\nu}) = P(\hat{u}, \vec{\nu}), \qquad \hat{u} \cdot \vec{\nu} = 0 \qquad \text{(Eq. 3)}$$

where capital letters denote Fourier transforms.

Therefore, every measured projection samples one plane through the origin of the image in frequency space. The object $f(\vec{x})$ is completely determined if $F(\vec{\nu})$ is known for all the frequencies $\vec{\nu}$; so recovery of $f(\vec{x})$ from $p(\hat{u},\vec{s})$ is possible if the set Ω of all the projections $\hat{u}$ is such that for every frequency $\vec{\nu}$ there exists at least one projection such that $\hat{u} \cdot \vec{\nu} = 0$. In fact since these planes generally intersect, some frequency components will be available from more than one projection, a fact that illustrates again the over-determination of the 3D problem. Since different frequencies $\vec{\nu}$ are then sampled more or less often, a filter function $H(\vec{\nu})$ is required to compensate for this imbalance by weighting every frequency component $F(\vec{\nu})$ by the number of projection planes that contain $\vec{\nu}$.

The sufficiency condition stated above can be expressed in terms that relate more directly to the scanner geometry [Orlov et al., 1975]. A stable solution to the inversion problem can be achieved if any great circle on the unit sphere S^2 has a non-empty intersection with the set Ω of available projection directions (*Orlov's sufficiency condition*). This condition is satisfied for instance for two-dimensional (ring-by-ring) acquisition, in which case the set Ω reduces to the equatorial circle on the unit sphere. It is satisfied *a fortiori* for 3D acquisition in a cylindrical scanner, since Ω in this case is not limited to, but contains the equatorial circle.

The above considerations suggest a direct method to recover $f(\vec{x})$ by placing the Fourier transformed projection planes into the Fourier image matrix and applying a 3D inverse Fourier transform to obtain the image. This Fourier reconstruction method [Stearns et al., 1987] suffers from complications however due to the need to interpolate between the available projection samples and the desired image elements. Filtered backprojection (FBP) solutions to this problem have been known for the past two decades for the particular case of full acceptance, i.e. $\Omega = S^2$ [Solmon et al., 1976], and

for the limited-angle case [Pelc et al., 1979]. As its name suggests, FBP consists of a filtering step followed by backprojection, or smearing back of the projection data into the image matrix:

Filtering step: Each 2D parallel projection is convolved with a kernel $h(\hat{u},\vec{s})$, which will be further specified below, to give the filtered projection $p^F(\hat{u},\vec{s})$:

$$p^{\mathrm{F}}(\hat{u},\vec{s}) = \iint_{\vec{s}'\cdot\hat{u}=0} \mathrm{d}^2\vec{s}'\; h(\hat{u},\vec{s}-\vec{s}')\; p(\hat{u},\vec{s}') \qquad \text{(Eq. 4)}$$

Backprojection : The 2D filtered projections $p^F(\hat{u},\vec{s})$ are backprojected by redistributing the value $p^F(\hat{u},\vec{s})$ uniformly along the line $(\hat{u},\vec{s})$ so as to form the reconstructed image:

$$f_r(\vec{x}) = \iint_{\Omega} \mathrm{d}^2\hat{u}\; p^{\mathrm{F}}(\hat{u},\vec{x}-(\vec{x}\cdot\hat{u})\hat{u}) \qquad \text{(Eq. 5)}$$

where projection element $\vec{s} = \vec{x} - (\vec{x}\cdot\hat{u})\hat{u}$ is the projection of vector $\vec{x}$ onto the projection plane perpendicular to $\hat{u}$. In practice, backprojection algorithm is carried out in reverse order with respect to that suggested by equation Eq. 5 for every $\hat{u}$, the contributions from the corresponding projection are backprojected into all image voxels, rather than repeating equation Eq. 5 for each voxel individually.

As stated earlier the convolution kernel $h(\hat{u},\vec{s})$ is no longer unique, unlike in the 2D case. Several different filters have been proposed in the literature [Colsher et al., 1980; Ra et al., 1982; Schorr et al., 1983], until Defrise derived the filter equation for $h(\hat{u},\vec{s})$ and proposed a general solution [Defrise et al., 1989]. The Dirac delta-function restricts the two-dimensional integral over all possible directions to those orthogonal to $\vec{\nu}$, i.e. to the curved line integral along the intersection of Ω with the great circle orthogonal to $\vec{\nu}$. This intersection corresponds to the set of available projections which each allow an independent estimate of $F(\vec{\nu})$, each of these projections being assigned a weight given by the filter $H(\hat{u},\vec{\nu})$. The reconstructed Fourier transform of the image $F_r(\vec{\nu})$ is expressed as the product of the Fourier transform of the original object $F(\vec{\nu})$ by $T(\vec{\nu})$ with :

$$T(\vec{\nu}) = \iint_{\Omega} \mathrm{d}^2\hat{u}\; \delta(\hat{u}\cdot\vec{\nu})\, H(\hat{u},\vec{\nu}), \forall\vec{\nu}\in\mathrm{R}^{3*} \qquad \text{(Eq. 6)}$$

This is the modulation transfer function (MTF) of the entire process consisting of the measurement of the projections, their filtering and the backprojection. The filter equation must hold for the reconstruction to be exact and should follow the following equation :

$$T(\vec{\nu}) = 1, \forall\vec{\nu}\in R^{3*} \qquad \text{(Eq. 7)}$$

A general solution can be constructed by normalising an arbitrary function $G(\hat{u},\vec{\nu})$ by its value integrated over Ω.

$$H(\hat{u},\vec{\nu}) = \frac{G(\hat{u},\vec{\nu})}{\iint_{\Omega} \mathrm{d}^2\hat{u}'\; \delta(\hat{u}'\cdot\vec{\nu})\, G(\hat{u}',\vec{\nu})}, \forall\vec{\nu}\in\mathrm{R}^{3*} \qquad \text{(Eq. 8)}$$

provided that *G* is such that the denominator is defined and does not vanish.

Alternatively, the solution can also be constructed by adding an appropriate correction factor to an arbitrary function *G*:

$$H(\hat{u}, \vec{v}) = G(\hat{u}, \vec{v}) + \frac{w(\hat{u})\left(1 - \iint_{\Omega} \mathrm{d}^2\hat{u}'\ \delta(\hat{u}' \cdot \vec{v})\, G(\hat{u}', \vec{v})\right)}{\iint_{\Omega} \mathrm{d}^2\hat{u}'\ \delta(\hat{u}' \cdot \vec{v})\, w(\hat{u}')}, \forall\, \vec{v} \in \mathrm{R}^{3*} \qquad \text{(Eq. 9)}$$

where w is any integrable function of the direction of projection $\hat{u}$ such that the denominator does not vanish. Since there is at least one such $w(\hat{u})$ (for example $w(\hat{u}) = 1$), additive normalisation is always possible, and is preferred in practice; The FAVOR filter described in section 2.6 is an example of such an additive correction factor.

The class of filters built from equation Eq. 9 with $G(\hat{u}, \vec{v}) = w(\hat{u})$ independent of the frequency are called "factorisable filters" and possess the property that filtering (Eq. 4) and backprojection (Eq. 5) with an angular weight $w(\hat{u})$ can be commuted to yield an equivalent backprojection-filtering reconstruction. Although this method is not generally used it remains the only possible approach when events are backprojected one by one rather than being histogrammed into projection data.

These filters all lead to the same exact reconstruction in the case of consistent, noise free data. However, with real data corrupted by noise they yield different results and are no longer equivalent. The next section deals with the particular filter which minimises the reconstruction error in the presence of noise.

The role of the filter is to account for the non-uniform sampling density in Fourier space as detailed in the following sections. In 2D, there is no data redundancy and the filter is a ramp. In 3D there is data redundancy and as a consequence the filter is not unique. All should satisfy the Defrise equation (Eq. 7). The Colsher filter is used in the reprojection algorithm (3DRP) of Kinahan and Rogers [Kinahan et al., 1989], also named 3DRP (3D Reprojection). Fast volume reconstruction algorithm (FAVOR) uses a filter for which there is no need to reprojected the missing sinograms and is faster than 3DRP. Furthermore, recent developments of beamwise incremental backprojectors allow to speed up significantly FBP processing [Egger, 1996]. Fourier rebinning [Defrise 1995] (FORE) is a new algorithm which rebins the 3D data into 2D direct sinogram and the image is reconstructed in 2D slice-by-slice using, for instance, the standard 2D FBP algorithm.

2.2. Colsher filter

The usual approach to over-determined inverse problems is to select the least squares solution of minimum norm. It can easily be verified that this solution is obtained using the filter defined by equation Eq. 8 with $G(\hat{u}, \vec{v}) \equiv 1$ [Defrise et al., 1995]. This filter was originally derived by Colsher [Colsher et al., 1980] for the case of a shift-invariant scanner with limited acceptance θ_o. His initial derivation was based on a relation between the activity distribution $f(\vec{x})$ and the backprojected image

$$g(\vec{x}) = \iint_{\Omega} \mathrm{d}^2\hat{u}\ p(\hat{u}, \vec{x} - (\vec{x} \cdot \hat{u})\hat{u}) \qquad \text{(Eq. 10)}$$

of the unfiltered projections, in a backproject-then-filter framework. In a scanner with a spatially invariant point spread function $q(\vec{x})$, resulting from the backprojected projections of a single object point, these quantities are linked by the relation

$$g(\vec{x}) = \int f(\vec{x}')q(\vec{x} - \vec{x}')d^3\vec{x}' \qquad \text{(Eq. 11)}$$

The expression of $q(\vec{x})$ as a shift invariant convolution integral is possible only under the condition of spatial invariance of the point spread function, as above. The practical consequences for its implementation are deferred to the next section; for now, let us simply force spatial invariance by restricting the projections to directions $\hat{u}$ such that their co-polar angle θ is less than the minimum value θ_o of the acceptance angle within the entire field of view; that is, all projections with $|\theta| < |\theta_o|$ completely sample the object $f(\vec{x})$. Simple geometrical considerations lead to a mathematical formulation of function $q(\vec{x})$. Recovery of the activity distribution $f(\vec{x})$ then requires the inversion of Eq. 11, which may be written in Fourier space as the deconvolution process

$$F(\vec{\nu}) = G(\vec{\nu})Q^{-1}(\vec{\nu}) = G(\vec{\nu})H_{Colsher}(\vec{\nu}) \qquad \text{(Eq. 12)}$$

with $H_{Colsher}(\vec{\nu}) = Q^{-1}(\vec{\nu})$. More detailed developments may be found in [Colsher et al., 1980] and we simply state the final result:

$$H_{Colsher}(\vec{\nu}) = \begin{cases} \dfrac{|\vec{\nu}|}{2\pi} & if\ cos\psi \geq \cos\theta_o \\ \dfrac{|\vec{\nu}|}{4} \dfrac{1}{\arcsin(\sin\theta_o / \sin\psi)} & if\ cos\psi < \cos\theta_o \end{cases} \qquad \text{(Eq. 13)}$$

where ψ is the polar angle of the frequency vector $\vec{\nu}$. In practice, filtering is carried out before backprojection, and the two-dimensional projections along direction $\hat{u}$ are filtered with a cross-section of this filter through the origin such that $\hat{u} \cdot \vec{\nu} = 0$.

The weighting property of the filter function *H* becomes clear if the result of the equation above is related to the description of Orlov's sphere: the value of *H* is the reciprocal of the arc length given by the intersection of the great circle perpendicular to $\vec{\nu}$ with the set Ω. The multiplication by the filter function may thus be understood explicitly as a compensation for the number of projections contributing to one given frequency component.

2.3. Shift invariance constraint

So far we have assumed that projections along a direction $\hat{u} \in \Omega$ are completely measured, i.e. that the solid angle covered by the detection channels is the same everywhere inside the field of view. This would be the case in a cylindrical scanner of infinite axial extent, or in a finite cylinder if the set Ω is restricted to directions orthogonal to the scanner axis. However, making full use of all available detection channels in a truncated cylindrical scanner results in a severely shift variant problem and truncated projection data.

Kinahan and Rogers [Kinahan et al., 1989] proposed their reprojection method to work around shift variance: a first estimate of the image is reconstructed using a sufficient data set satisfying Orlov's condition and with all projections fully measured; this is usually the set of transaxial, or direct,

projections. By calculating line integrals through this first image along the missing detection channels, the truncated parts of the projections at all remaining angles can be recovered. This is equivalent to numerically forward-projecting what would have been detected by an axially longer scanner. Having artificially restored shift invariance, the image can be reconstructed using FBP with the Colsher filter. Due to the hybrid nature of the data – one projection consists of data both measured and estimated by forward projection – noise propagation properties and smoothing characteristics of this method become involved. The validity of this approach is, however, confirmed daily in practical use since the reprojection method is the most widely used 3D reconstruction algorithm and is implemented in most commercial 3D PET scanners. Clearly, the need to perform the reprojection step represents a significant increase in both calculation and memory required to hold the data set. This effect will be felt even more with the trend going towards the development of ever longer scanners.

2.4. Limitations of simple 1D transaxial ramp filter

The filter function $H(\vec{u},\vec{v})$ in equation Eq. 8 is the two-dimensional function that requires the two-dimensional projections to be completely measured. Let us now consider the problem from the other end. Rather than devising a filter function that satisfies the recovery condition (Eq. 7), we shall now estimate the error caused by the use of a particularly simple filter function, the one-dimensional transaxial ramp filter used in 2D image reconstruction. Using this filter for doing FBP reconstruction results in a significant speed-up and it is interesting to determine to what extent the non-conformation to the filter condition deteriorates the reconstructed images.

The 1D transaxial ramp filter is proportional to the magnitude of the transaxial component v_a of $\vec{v}$, i.e. $H_{ramp}(\hat{u},\vec{v}) \propto |\vec{v}|\cos\omega|$. For reasons which will become apparent later, we shall introduce a normalisation factor and write the filter function as

$$H_{ramp}(\hat{u},\vec{v}) = \frac{1}{4\theta_o}|\vec{v}|\cos\omega| \qquad \text{(Eq. 14)}$$

where $\vec{v} = |\vec{v}|(\hat{a}\cos\omega + \hat{b}\sin\omega)$ is a vector in the projection plane in which the coordinate system is defined by two unit vectors $\hat{a}$ and $\hat{b}$ with vector $\hat{a}$ in the transaxial direction i.e $a_z = 0$ and $\hat{a}.\hat{b} = 0$ as shown in Figure 2.

The dependence of this filter function on the transaxial co-ordinate v_a only, and not on v_b, makes it attractive in the case of a typical scanner with a truncated cylindrical geometry: filtering of a particular transaxial section of the 2D projection (that is, $\vec{s}\cdot\hat{b} = constant$), where $\vec{s}$ is the position vector of a projection element within the projection plane is unaffected by the values of the projection in other transaxial sections. Filtering may therefore be carried out separately for each transaxial section, independently of the values in other transaxial sections, which may not even be measured due to the truncated geometry.

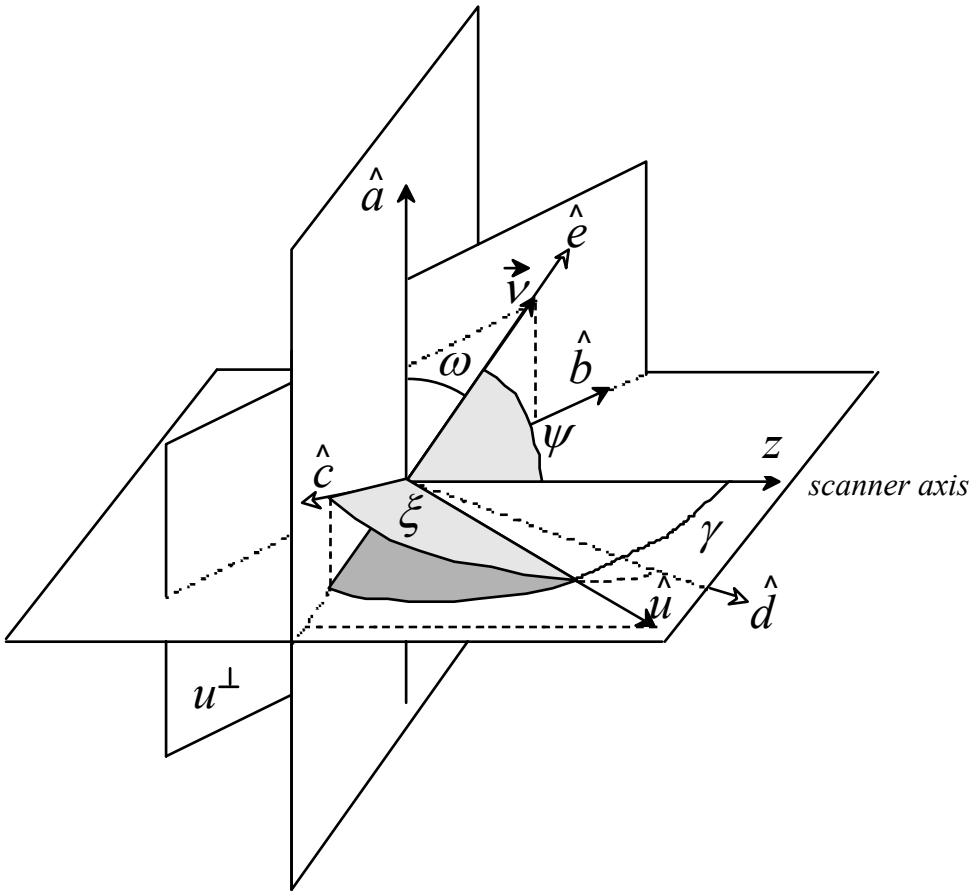


Figure.2 : *Set of co-ordinates $\hat{c}$, $\hat{d}$ and $\hat{e}$ used to compute the modulation transfer function of the transaxial ramp filter.*

In order to estimate the accuracy of the transaxial ramp filter it is necessary to compute its modulation transfer function (MTF) $T(\vec{\nu})$ (Eq. 6) which is:

$$T_{ramp}(\vec{\nu}) = \begin{cases} \psi / \theta_j & \text{if } 0 \le \psi < \theta_j \\ 1 & \text{if } \theta_j \le \psi < \pi - \theta_j \\ (\pi - \psi) / \theta_j & \text{if } \pi - \theta_j < \psi \le \pi \end{cases} \qquad \text{(Eq. 15)}$$

This result shows that all frequencies outside the cone defined by $|\nu_z / \nu| > \cos\theta_j$ are recovered exactly (MTF=1), so that the Ramp algorithm is expected to be a reasonable approximation for small aperture scanners.

2.5. "3DRP": 3D Reprojection algorithm

Three-dimensional reconstruction is complicated by the fact that all the oblique 2D projections (corresponding in a multi-ring scanner to coincidences with a ring difference larger than one) are incompletely measured due to the finite length of the scanner. This is done by forward projecting a first image estimate, reconstructed from the direct projections since these are not truncated. The algorithm combining this forward projection step with the filter of Colsher is referred to here as the 3DRP

algorithm (for 3D Reprojection). In practice, as much as 40% of the total reconstruction time is spent in estimating, and then backprojecting the unmeasured projection data.

2.6. "FAVOR": Fast volume reconstruction algorithm

In some situations, and in particular when the axial aperture of the scanner is large, the error due to the approximations of the Ramp algorithm may no longer be negligible with respect to other sources of error (measurement noise, sampling, background, etc...). The properties of the ramp algorithm above may now be used to devise an analytically exact filter – the FAVOR filter – satisfying the filter Eq. 7 by use of an additive correction, as in Eq. 9. The key point is that this will be an exact method that does not require all 2D projections to be completely measured [Comtat et al., 1993; Comtat et al., 1994; Defrise et al., 1992]. One-dimensional transaxial filtering may be applied to oblique projections, regardless of the truncation in the axial direction. The error due to the MTF not being equal to one for all frequencies (Eq. 15) may be corrected for by applying a complementary 2D filter to the fully measured direct projections.

The MTF of the correction filter H_{corr} must be $1 - T_{ramp}(\vec{\nu})$ rather than 1 to compensate for the error introduced by the Ramp algorithm. It may be based on any filter satisfying condition (Eq. 7), the simplest being Colsher's filter, so that the denominator of Eq. 9 is 1:

$$H_{corr}(\vec{\nu}) = (1 - T_{ramp}(\vec{\nu}))H_{Colsher}(\vec{\nu}) \qquad \text{(Eq. 16)}$$

Combining H_{corr} with the ramp filter we obtain the FAVOR filter:

$$H_{Favor}(\vec{\nu}) = \frac{1}{4\theta_j}|\nu|\cos\omega| + w(\cos\theta)H_{corr}(\vec{\nu}) \qquad \text{(Eq. 17)}$$

where $w(\cos\theta) = \begin{cases} 1 & \text{if } |\cos\theta| > \cos\theta_a \\ 0 & \text{if } |\cos\theta| \le \cos\theta_a \end{cases}$

The angle θ_a is the limiting angle of acceptance for direct projections, and is usually taken to correspond to coincidence channels in adjacent detector rings. The second term on the right hand side of this expression vanishes for oblique projections, and the filter reduces to the transaxial 1D ramp filter.

Comparing the FAVOR and the Colsher filters, FAVOR will exhibit different noise propagation properties: The low-frequency hump characteristic of FAVOR includes most of the useful frequency range occurring in a typical image, amplifies it to a multiple of the 1D Ramp value, and hence introduces a bias by attributing a much greater weight to the direct projections than to the oblique projections. High-frequency statistical noise is however not affected since it occurs at frequencies superior to those corrected by the FAVOR filter.

2.7. Rebinning algorithms

The large amount of calculations required by truly three-dimensional reconstruction algorithms such as those discussed in the previous sections have motivated the development of fast approximate methods for incorporating all the information measured by a volume scanner into the reconstructed image. An improvement of the 3D reconstruction time requires some form of data compression to reduce huge number of data sampling. One approach consists in undersampling the data by averaging groups of contiguous samples (*mashing*). Another approach is considered in this section: *rebinning*.

2.7.1. "SSRB" : Single slice rebinning

The single-slice rebinning (SSRB) algorithm [Daube-Witherspoon et al., 1987] is based on a very simple approximation but its accuracy is acceptable only near the axis of the scanner and for small apertures. The SSRB assigns an LOR between two detectors to the sinogram of the transaxial slice lying midway, axially between the two detectors. A sinogram s_k is synthesised for each transaxial slice k=1...2N_{rings}–1. Each slice k is then reconstructed by applying 2D FBP to the corresponding sinogram s_k. In the SSRB, s_k is the average of the sinogram $S_{i,j}$ between pairs of rings i,j located symmetrically, axially, with respect to slice k:

$$s_k = \frac{1}{k}\sum_{i=1}^{k} S_{i,k+1-i}, 1 \le k \le N_{rings} \qquad \text{(Eq. 18)}$$

and a similar expression holds for $N_{rings} \le k \le 2N_{rings} - 1$

2.7.2. "MSRB": Multi-slice rebinning

The multi-slice rebinning algorithm (MSRB) [Lewitt et al., 1994] involves a more sophisticated rebinning, where an a sinogram $S_{i,j}$ contributes to all slices k such that the LORs between diametrically opposite detectors in rings i and j intersect the slice k within a cylindrical field-of-view characterised by a transaxial radius R_{FOV}. The sinogram s_k is then built as a weighted average of all sinograms $S_{i,j}$ which satisfy that condition. After 2D reconstruction of the $2N_{rings}$–1 sinograms s_k, each column of voxels is then deconvolved using the multiplicative iterative algorithm described by Lewitt et al. (1994). MSRB is more accurate than SSRB but suffers from instabilities in the presence of noisy data.

2.7.3. "FORE": Fourier rebinning

The Fourier rebinning algorithm is a newly proposed algorithm for 3D image reconstruction in which oblique rays are binned to a transaxial slice using the frequency-distance relationship of the data in Fourier space [Defrise et al., 1995; Lewitt et al., 1994]. This allows the processing of the 3D data as 2D slices and thus speeds up the reconstruction by an order of magnitude. This algorithm re-histogram (*rebin*) the 3D data into a stack of ordinary 2D direct sinograms and the image is reconstructed in 2D

slice-by-slice, for instance using the standard 2D filtered-backprojection algorithm. The way the data from the oblique sinograms are rebinned into the smaller 2D data set is the crucial step of any rebinning algorithm and has a vital influence on the accuracy and variance of the reconstructed image.

A preliminary step necessary for this approach is the resampling of the measured data. Consider a cylindrical N ring scanner with a radius R and its axis along the z axis. Let us now derive the frequency-distance relation for the 3D X-ray transform; first let us take the continuous Fourier transform of one sinogram $p_r(s,\phi,z,\tan\theta)$ where s and ϕ are ordinary sinogram variables, $z = (z_A + z_B)/2$ the average of the axial coordinates of the two detectors in coincidence, and $\tan\theta = (z_A - z_B)/2R$, the tangent of the angle θ between a line of response (LOR) and the transaxial plane :

$$P_r(\omega,k,z,\tan\theta) = \int_0^{2\pi} \mathrm{d}\phi \int_{s_{\min}}^{s_{\max}} \mathrm{d}s\ p_r(s,\phi,z,\tan\theta)\, e^{-ik\phi - i\omega s} \qquad \text{(Eq. 19)}$$

with k being the integer index of the angular Fourier series, and ω the continuous spatial frequency corresponding to the co-ordinate s. Sinograms are usually sampled over half the azimuthal range, $\phi \in [0,\pi)$, so that complete 2π sinograms must be constructed using the relation

$$p_r(s,\phi+\pi,z,\tan\theta) = p_r(-s,\phi,z,-\tan\theta) \qquad \text{(Eq. 20)}$$

Let us now substitute Eq. 19 into Eq. 20:

$$P_r(\omega,k,z,\tan\theta) = \iint_{x^2+y^2 \le s_{\max}^2} \mathrm{d}x\,\mathrm{d}y \int_0^{2\pi} \mathrm{d}\phi\, f(x,y,z+\tan\theta\,(-x\sin\phi + y\cos\phi))\, e^{-ik\phi - i\omega(x\cos\phi + y\sin\phi)}$$

(Eq. 21)

For large values of k and ω, the phase of the exponential in the above equation varies rapidly with ϕ, so that the exponential oscillates rapidly between positive and negative values. If the function f is sufficiently smooth this will yield a negligible contribution to the inner integral over ϕ; consequently, the only important contributions to this integral come from the neighbourhood of values of ϕ where the phase is stationary:

$$\frac{\partial}{\partial\phi}(k\phi + \omega(x\cos\phi + y\sin\phi)) = k + \omega(-x\sin\phi + y\cos\phi) = 0 \qquad \text{(Eq. 22)}$$

When $|k| \le |\omega|\sqrt{x^2+y^2}$ this equation has two solutions in the interval [0, 2π], both corresponding to the same signed distance $t = (-x\sin\phi + y\cos\phi)$ along the LOR, the distance t being measured on the transaxial projection of the LOR, from the midpoint. Each frequency component (k,ω) thus receives contributions mainly from points located at a fixed distance $t = -k/\omega$ along all LORs of the original sinogram.

The stationary-phase approximation is obtained by replacing, in the third argument of f in Eq. 21, $-x\sin\phi + y\cos\phi$ by $-k/\omega$, and factoring f out of the integral:

$$P_r(\omega,k,z,\tan\theta) \cong \iint_{x^2+y^2 \le s_{\max}^2} \mathrm{d}x\,\mathrm{d}y\ f(x,y,z-\tan\theta\ k/\omega) \int_0^{2\pi} \mathrm{d}\phi\, e^{-ik\phi - i\omega(x\cos\phi + y\sin\phi)} \qquad \text{(Eq. 23)}$$

The axial shift ($-\tan\theta\, k/\omega$) is independent of the transaxial co-ordinates x and y, so that this equation may be rewritten in the simple form:

$$P_r(\omega, k, z, \tan\theta) \cong P_r(\omega, k, z - \tan\theta\, k/\omega, 0) \qquad \text{(Eq. 24)}$$

This relation forms the heart of the algorithm. It allows the rebinning of the elements of the 2D Fourier transform of an oblique sinogram with oblicity $\tan\theta$ and mid-slice z into the elements of the 2D Fourier transform of a direct sinogram ($\tan\theta = 0$) for slice $z - \tan\theta\, k/\omega$. The equivalent relation in the single-slice rebinning approximation is:

$$P_r(\omega, k, z, \tan\theta) \cong P_r(\omega, k, z, 0) \qquad \text{(Eq. 25)}$$

3. REVIEW OF ITERATIVE ALGORITHMS

Iterative methods for reconstruction of PET images have been investigated for more than a decade now. It is well known that iterative methods produce images which are visually more attractive and in some cases diagnostically better. It is not clear, however, that this approach produces more information from a medical research or diagnostic point of view in the majority of cases. Most of the reported results of computational studies apply data set which are test "phantoms" of simulated shapes. In this section, we intend to establish the relative merits of different iterative methods.

The general form of an iterative method is to optimise some function object to a large system of equational and/or non-negativity restrictions. An attractive property of iterative reconstruction methods is that it is possible to introduce weights or penalties which reflect the nature of the problem and characteristics of the scanning system. Also additional constraints and penalty functions can be included which ensure that the image has certain desirable properties. This enables algorithms to be tuned to specific clinical requirements. Iterative techniques for both 2D and 3D image reconstruction in PET include :

- Maximum likelihood-expectation maximisation (ML-EM)
- Ordered subset expectation maximum (OSEM)
- Space-alternative generalized expectation maximum (SAGE)
- Least squares minimization (LSQ)
- Algebraic reconstruction technique (ART)
- Maximum a posetriori (Bayesian) approach (MAP)
- Maximum entropy (ME)

All those algorithms will be described in the following sections.

3.1. "ML-EM": Maximum likelihood - expectation maximization

The application of the expectation maximization (EM) algorithm in emission tomography by Shepp and Vardi has led to the introduction of many related techniques [Shepp et al., 1982]. These include a number of important Bayesian (or equivalent penalized likelihood) adaptations of EM. The EM generally employs a postulated, unobservable "complete data" set, which is sampled from a larger

or richer space, to facilitate the maximization process. A conditional expectation can be formulated by defining a new likelihood function based on this postulated "complete data" set, given the measured data and a current estimate. In theory, maximizing this conditional expectation is equivalent to the basic ML task. Therefore, the EM algorithm provides a pratical means for computing ML estimates iteratively. The method of maximum likelihood (ML) is a general technique for finding estimators. ML estimators are the values of unobservable parameters of some theoretical model, based on the physical and statistical processes involved in the study, that make the measured data most likely.

The problem of PET image reconstruction can be formulated in the context of an incomplete data problem. The parameters to be estimated are local radioactivity distributions $\lambda = \{\lambda_j\}$, in different image cells, which is equivalent to the mean Poisson intensity of photons emitted from those cells. The observed data, $y = \{y_i\}$, are counts of photons collected in various detectors bins. The latent (or missing) data, $n = \{n_j\}$, are number of photons actually emitted from different image cells, which cannot be observed because of their interactions with tissues, misregistration in detectors arrays, and the fact that each detector bin collects photons emitted from all image cells along the same projection path. The task of PET image reconstruction is to estimate λ given y. In general, a likelihood function can be defined as the joint probability function of the measured data in terms of the unobservable parameters to be estimated in the task. Maximising this likelihood function with respect to the unobservable parameters yields the parameters with which the data are more consistent.

Let a body with a variable radiation density be contained in the discretized cube $B(x,y,z)$ then the emission density of each unit box (voxel) $v=(x,y,z)$ in the cube is defined as $\lambda(v)$. When the radiation emitted by the body is detected by detectors encircling the cube, all coincidence acquisition reconstruction algorithms try to determine the unknown density distribution in the cube given the list of detected events. The PET scanner is built of rows of detector rings whereby each detector is regarded as a discrete unit. All the events detected simultaneously by the same two detectors d_i and d_j are collected in the single bin b_{ij}. Thus, each bin b defines a single line of response (LOR) and every event detected in this bin is assumed to have been originated along this LOR.

Let $v=1,..,V$ represent the voxels of the field of view cube and let independent Poisson variables with unknown means $\lambda(v)$ represent the number of unobserved emission in each of the V voxels. Suppose that an emission from the voxel v is detected in bin b with probability $p(v,b)$, $b=1,..,B$ then $p(v,b)$ defines a transition matrix (likelihood matrix) assumed known. We observe the total number $y = y(b)$ of events detected by each bin b and want to estimate the unknown $\lambda = \lambda(v)$, $v=1,..,V$. For each λ, the observed data y has the conditional probability or likelihood :

$$P(y|\lambda) = \prod_{1\le b\le B} e^{-\mu(b)} \frac{\mu(b)^{y(b)}}{y(b)!} \qquad \text{(Eq. 26)}$$

where $\mu(b)$ are the means of the Poisson variables $y(b)$ that is

$$\mu(b) = \sum_{1\le v\le V} \lambda(v) p(v,b) \qquad \text{(Eq. 27)}$$

The maximum likelihood estimate is :

$$\bar{\lambda} = \arg\max_{\lambda} P(y|\lambda) \qquad \text{(Eq. 28)}$$

The optimality condition of the above equation (based on the log-likelihood) is :

$$\lambda(\nu) = \frac{1}{P(\nu)} \sum_{b=1}^{B} \frac{y(b)\lambda(\nu)p(\nu,b)}{\sum_{\nu'}^{V} \lambda(\nu')p(\nu',b)}, \quad \nu = 1,\ldots,V \qquad \text{(Eq. 29)}$$

where $P(\nu)$ is the probability to detect an event emitted from the voxel ν :

$$P(\nu) = \sum_{b=1}^{B} p(\nu,b) \qquad \text{(Eq. 30)}$$

The EM algorithm can be considered as a fixed point iterative algorithm based on Eq. 30.

$$\lambda^{new}(\nu) = \frac{1}{P(\nu)} \sum_{b=1}^{B} \frac{y(b)\lambda^{old}(\nu)p(\nu,b)}{\sum_{\nu'}^{V} \lambda^{old}(\nu')p(\nu',b)}, \quad \nu = 1,\ldots,V \qquad \text{(Eq. 31)}$$

where $p(\nu',b)$ is normalised such that the total probability of detection emitted from voxel ν is unity, i.e. $\sum_{\nu'}^{V} p(\nu',b) = 1$

The probability $p(\nu',b)$ is influenced by a number of factors such as the geometrical relation between pixel and detector pair, duration of data acquisition, rate of radioactivity decay, detector response, scatter, attenuation, etc.

It has been known for more than a decade that the ML-EM algorithm can produce good quality images in 2D, but due to various problems, the algorithm has not been exploited. Nevertheless, it is considered by many as the "gold standard" reconstruction algorithm [Vermeulen et al., 1989; Llacer et al., 1993; Fessler et al., 1995; Depierro et al., 1995; Belleman et al., 1995]. There are three major problems that prevented the use of ML-EM in commercial produced scanners, namely :

- the large memory requirements, which arises due to the size of the maximum-likelihood matrix ;
- the computational complexity of the single iteration which is similar to that of the FBP algorithm. However, in some cases, several hundred iterations are needed for a reasonable reconstruction ;
- the lack of a good stopping criterion, which is the most serious one. Since, when the relative error is small enough, the natural stopping criterion of an iterative algorithm is known to produce snowy images.

3.2. "OSEM": Ordered subsets expectation maximum

While the quality of reconstruction is good, the application of EM is computer intensive, and convergence slow, even with standard acceleration techniques. Recently, Hudson and Larkin presented an accelerated version of the EM algorithm based on an ordered sets approach [Hudson et al., 1994]. The Ordered Sets EM (OSEM) algorithm processes the data in subsets (blocks) within each iteration and show that this procedure accelerates convergence by a factor proportional to the number of subsets. Many independent tests proved that the OSEM produces images which are similar in quality to those produced by the EM algorithm in a fraction of the time [Meikle et al., 1996].

Let y_i be the number of photon emissions recorded in the i^{th} projection bin and let Y_θ be the set of parallel projections $\{y_i, y_{i+1}, \ldots\}$ that view the object at angle θ, orthogonal to the tomograph axis. The

projection data are grouped into n subsets, denoted $S_1, S_2, \ldots, S_n$. If there are P projection bins in total, the elements of each subset, in no particular order, are :

$$S_1 = \left\{ y_1, y_2, \ldots, y_{P/n} \right\} S_2 = \left\{ y_{P/n+1}, y_{P/n+2}, \cdots, y_{2P/n} \right\} \ldots, \; S_n = \left\{ y_{(n-1)P/n+1}, y_{(n-1)P/n+2}, \ldots, y_P \right\} \quad \text{(Eq. 32)}$$

The subsets normally consist of projection views separated by some fixed angle about the object. For example, each subset might consist of two sets of parallel projection views, spaced 90 degrees apart : $S_1 = \left\{ Y_0, Y_{\pi/2} \right\}, S_2 = \left\{ Y_{\pi/4}, Y_{3\pi/4} \right\}$, and so on.

In the OSEM algorithm, we increase the log-likelihood objective function for each of the subsets using the result of the previous subset as a starting point. Therefore, from the Eq. 31 by replacing λ^{old} by λ^{k} where λ^{k} is the estimated number of emissions following the introduction of the k^{th} subset of projections, the EM iterations become:

$$\lambda^{k+1}(\nu) = \frac{1}{P(\nu)} \sum_{b \in S_k} \frac{y(b) \lambda^k(\nu) p(\nu, b)}{\sum_{\nu'}^{V} \lambda^k(\nu') p(\nu', b)}, \quad \nu = 1, \ldots, V \qquad \text{(Eq. 33)}$$

The EM procedure is repeated until all n subsets have been exhausted. Such a cycle is considered a single iteration of OSEM. The cycle can be repeated iteratively until a satisfactory reconstruction is obtained.

In this framework, EM is a special case of OSEM with the number of subsets set to $n = 1$. For the case of noiseless projections (infinite statistics), it has been shown that each OSEM estimate based on a subset of projections converges as far towards a maximum likelihood solution as a full iteration of EM using all projections [Hudson et al., 1994]. In other words, if the projection data are divided into n subsets, then once all projections have been used in a single iteration of OSEM, an estimate has been produced which is similar to n iterations of EM. It is this property that gives OSEM considerable acceleration compared with EM. In general, the greater the level of subdivision of the projection data, the greater the acceleration that is achieved. It was shown, however, that after a certain number of subsets the results deteriorate. It is assumed, however, that 16 are adequate giving the OSEM a 16 fold speedup compared to the EM algorithm.

A drawback of the OSEM algorithm when applied to real, noisy, data is its tendency to cycle between several solutions and not to converge to the maximum of likelihood function. Although it seems that this deficiency has, in practice, only little influence there is a way to overcome it. Browne et al. (1996) propose a framework whereby the OSEM step size is controlled and is calculated at every outer iteration. The authors claim that this algorithm, which can be seen as a combination of the ART and the OSEM approaches, performs better than standard EM or OSEM. Their algorithm, called row action maximum likelihood algorithm (Ramla) can be interpreted as the following :

$$\lambda^{new\_Ramla} = (1-\theta)\lambda^{old} + \theta \lambda^{newosem}, \; 0 \le \theta \le 1 \qquad \text{(Eq. 34)}$$

where θ is the step size and $\lambda^{newosem}$ is the result of applying one iteration of the algorithm, achieving a significant higher likelihood than several iterations of ML-EM [Matej et al., 1995] . Note that for $\theta = 1$ we get the standard OSEM algorithm.

3.3. "SAGE": Space-alternative generalised expectation maximum

The space-alternative general expectation (SAGE) algorithm has been proposed as a fast method to compute penalised maximum likelihood estimates in PET. Sequential update methods replace the simultaneous update of all pixels dictated by the EM algorithm with a scheme that sequentially updates each voxel. The most promising of these is the SAGE algorithm proposed by Fessler et al. [Fessler et al., 1995; Hero et al., 1994; Ollinger et al., 1996]. The SAGE algorithm is a generalised EM algorithm that speeds convergence by sequentially updating voxels using a sequence of cleverly chosen hidden data spaces. The hidden data spaces are chosen to minimise the largest eigenvalue of the Fisher information matrix. Since this eigenvalue governs the rate of convergence, better choices of the hidden data space yield faster algorithms. A penalised version the algorithm by Fessler et al. known as PNL-SAGE3 is presented below [Fessler et al., 1995].

The system is modelled by a transition matrix $A=\{p_{nk}\}$ such that the observed data are given by $\{y_n\}$and are modelled by:

$$\bar{y}_n(\lambda) = \sum_k p_{nk}\lambda_k + r_n + s_n \qquad \text{(Eq. 35)}$$

where $\{\lambda_k\}$ is the image, $\{r_n\}$ the randoms and $\{s_n\}$ the scattered coincidences. The log-likelihood is generalised by adding the quadratic penalty function:

$$\frac{1}{2}\beta\sum_k \sum_{j\in N_k} \frac{1}{2} w_{kj}(\lambda_k - \lambda_j)^2 \qquad \text{(Eq. 36)}$$

where β is an adjustable parameter and the weights $\{^{w_{kj}}\}$ are chosen here to be the inverse of the distance between the center of voxel k and voxel j. The neighbourhood N_k is defined in three dimensions to be the 26 voxels surrounding a given voxel. The algorithm is given by :

$$\lambda_k^{i+1} = \left[-B_k + \sqrt{B_k^2 + \frac{e_k(\lambda)(\lambda_k^i + z_k^i)}{\beta w_k}} - z_k^i \right] \qquad \text{(Eq. 37)}$$

where $\{^{z_k^i}\}$ are the design parameters associated with the choice of hidden data space $w_k = \sum_j w_{kj}$.

The expression

$$e_k(\lambda) = \frac{\sum_n p_{nk} y_n}{\bar{y}_n(\lambda^i)} \qquad \text{(Eq. 38)}$$

is recognisable as the update factor in the EM algorithm and

$$B_k = \frac{p_{\bullet k} - \beta \sum_{j\in N_k} w_{kj}(\lambda_j^i + z_k^i)}{2\beta w_{k\bullet}} \qquad \text{(Eq. 39)}$$

The design parameters are chosen according to the rule

$$z_k^i = \min_{n:\, p_{nk}\neq 0} \left\{ \frac{\bar{y}_n(\lambda)}{p_{nk}} \right\} - \lambda_k^i \qquad \text{(Eq. 40)}$$

Note that the penalised method above which is described by Ollinger et al (1996) breaks each iteration into n sub-iterations. Each such sub-iteration computes a correction matrix from orthogonal projections. It is regularised using the one step late approximation. The algorithm is, then, given by :

$$\lambda_k^{i+1} = \lambda_k^{i,m=M}, \lambda_k^{i,0} = \lambda_k^i \text{ and } \lambda_k^{i,m+1} = \lambda_k^{i,m} \frac{e_k(\lambda^{i,m})}{p_{k\bullet} + \frac{\beta}{2} \sum_{j \in N_k} (\lambda_k^{i,m} - \lambda_j^{i,m})} \quad \text{(Eq. 41)}$$

3.4. "LSQ": Least square minimisation of the likelihood matrix equations

The EM algorithm the likelihood that an annihilation which was detected in a certain area of the filed of view originated in a certain voxel of the image. Another approach for doing this is to use the least square minimisation (LSQ). The LSQ method is used by many scientific disciplines as an optimisation tool [Kaufman et al., 1993; Osullivan et al., 1995]. However, these methods suffer from the same difficulties. Furthermore, non-negativity constraints have to be imposed artificially. In many ways, the internal step of the LSQ algorithm is similar to that of the ML-EM algorithm. The use of known least-squares methods by using accelerators and preconditionners allows to speedup convergence. This will also have a qualitative effect since the preconditionner and accelerator will also influence the quality of the reconstructions.

Let y_i be the number of coincidences registered on the fixed coordinates LOR i and let P^T be the likelihood matrix. It is well known that due to ill conditioning, the problem

$$P^T \lambda = y \quad \text{(Eq. 42)}$$

is very difficult to solve. This leads to the typical approach in this cases, that is, least square (LSQ) minimisation :

$$Min: \ r(\lambda) = \frac{1}{2} \left\| P^T \lambda - y \right\|^2 \text{ subject to } \lambda \geq 0 : \quad \text{(Eq. 43)}$$

The pure LSQ minimisation and the EM algorithm have a similar problem, namely, that the reconstruction often leads to a snowy image. This is especially evident when there is only a low statistical count. In such cases, the noise is fitted together with the signal data. It is well known that in the first few iterations, the quality of the picture is visually improving and only then it starts to deteriorate. The common-sense approach is therefore to stop the optimisation as soon as this phenomena happens. Unfortunately, a quantitative stopping criteria that will incorporate this phenomena is yet to be discovered.

To resolve this problem, many researchers [Kaufman et al., 1995; Kaufman et al., 1993] have offered to use some kind of smoothing filter that will choose, from all possible solutions to the least square problem, those which are less streaky. In other words, instead of minimising $r(\lambda)$ we would like to minimise :

$$f(\lambda) = r(\lambda) + \mu q(\lambda) \quad \text{(Eq. 44)}$$

where $q(\lambda)$ is some smoothing filter and μ the assigned weight. A typical choice of $q(\lambda)$ is the following quadratic function similar to the one used in SAGE :

$$q(\lambda) = \sum_k \frac{1}{26} \sum_{j \in N_k} (\lambda_k - \lambda_j)^2 \quad \text{(Eq. 45)}$$

Where the $j \in N_k$ means that j is one of the 26 voxel neighbours of k. This filter reflects the a-priori knowledge that many parts of the image are likely to very fairly smooth. Of course, for some parts of the image, in many cases the important parts, this is not valid. Thus, if we would like to avoid eliminating these areas, an appropriate choice of μ is crucial.

The LSQ problem or the penalised LSQ can be solved using the preconditioned conjugate gradient (PCG) [Kaufman et al., 1996; Lalush et al., 1995] or any other iterative solution scheme. To a large extent, the solution scheme will determine the choice of μ. Kaufman et al. (1995), observed that if the PCG algorithm is used, the convergence graph resembles an L shaped curve. It is evident that the proper place to stop the algorithm is when the curve starts to bend. They developed several algorithms to identify a proper spot and suggest an adaptive μ that is re-calculated for each iteration. Using this approach, they reconstructed good quality images after small number of iterations (sometimes less then 8). It is easy to show that the work load of a single iteration of the PCG algorithm is at least as high as an iteration of the EM algorithm. However, a proper choice of the pre-conditioner can considerably reduce the number of iterations. This LSQ algorithm has similar property to ML-EM, but is better understood and various accelerators and preconditionners exist.

3.5. "ART": Algebraic reconstruction technique

The interest in algebraic reconstruction technique (ART) algorithms was rekindled after researchers showed that the algorithms, on simulated data, converged in slightly more than a single iteration. Although this may be not be the case in real clinical acquisition, the possibility of achieving reasonable reconstruction with relatively small number of iterations make this algorithms attractive. The ART algorithm uses a successive over-relaxation (SOR) method known in numerical linear algebra [Golub et al., 1983]. ART attempts to find a solution to Eq. 43 by processing each row equation m separately :

$$P_m^T \lambda = y_m \qquad \text{(Eq. 46)}$$

In fact, ART consists of guessing at a value for all the picture elements $\lambda(i,j)$, and then modifying each element along each ray by a factor which compensates for the discrepancy between the measured ray sum $y_{k(\theta)}$ and the calculated ray sum $R_{k(\theta)}$.

$$\lambda^{n+1}(i,j) = \lambda^n(i,j)\frac{y_{k(\theta)}}{R_{k(\theta)}} \qquad \text{(Eq. 47)}$$

where $\lambda(i,j)$ is the value of a particular pixel (i,j) of the image to the ray-sum $y_{k(\theta)}$.

If the calculated ray sum is the same as the measured value, it implies that the guessed value is correct for a particular projection; however, for another projection there might be a large discrepancy, thus the pixel of the last views while lying in the ray for the new view will be modified according to the discrepancy between the new ray and the measured value. Thus, each ray from each projection is examined and values of $\lambda(i,j)$ falling within that ray are changed iteratively for all the projections for 5 to 10 iterations. Eq. 47 is called multiplicative ART.

Another method of correcting the discrepancy between the measured projections consists of adding the difference between the measured ray sum and the simulated ray sum. This is called the additive form of ART.

$$\lambda^{n+1}(i,j) = \max\left\{\lambda^{n}(i,j) + (y_{k(\theta)} - R_{k(\theta)}) / N_{k(\theta)}, 0\right\} \qquad \text{(Eq. 48)}$$

Here $N_{k(\theta)}$ is the number of pixels lying along the particular ray $k(\theta)$ which passes through pixel (i,j).

There are two important modifications of ART. One consists of setting to zero those values in the array that are clearly zero because they correspond to a ray sum that was observed as zero. This effective bounds the data and is an important boundary condition for any of the iterative techniques.

Let π be a permutation defined on the rows of P^T and let μ^k be the relaxation parameter defined for the k^{th} iteration then:

$$\lambda_b^{k,m} = \lambda_b^{k,m-1} + \mu^k \frac{y_{\pi(m)} - \sum_{b'=1}^{B} \lambda_{b'}^{k,m-1} p_{b',\pi(m)}}{\sum_{b''=1}^{B} \left(p_{b'',\pi(m)}\right)^2} p_{b,\pi(m)} \qquad , b=1,..,B \ , \ \lambda^{k+1,0} = \lambda^{k,m=M} \qquad \text{(Eq. 49)}$$

If a proper relaxation parameters μ^k are chosen and proper ordering permutation π of the row equations is maintained, then it is possible that the algorithm will converge very quickly. Herman et al. (1995) performed many experiments with different relaxation parameters and concluded that the best method of choosing the set of parameters (0,.., μ^{k-1}, μ^k,..) is experimental. That is, to find a typical set which represent a specific medical task (brain scan, say) and to optimise the ART algorithm by finding, at every iteration, the best μ^k parameter to continue. He also shows that if proper series of regularization parameters μ^k 's are chosen then a good quality reconstruction is obtained after 2-3 iterations of the ART algorithm.

From Eq. 49, it is evident that the computational effort of the ART algorithm is relatively low. Furthermore, it can be parallelized efficiently to vector operations on different processors. Data matrices can be several hundred of Megabytes in size. Processing times can be tens of minutes on the proper selection of control parameters and the order of processing data subset. Currently, these parameters are fixed in the experimental way and are highly dependent on the acquisition type.

3.6. "MAP": Maximum a posteriori (Bayesian) approach

An undesirable property of the ML-EM algorithm is that after a certain number of iterations, the reconstructed images become increasingly noisy. Bayesian reconstruction methods form a powerful extension of the maximum likelihood reconstruction method. Bayesian reconstruction methods allow the incorporation of prior information (such as smoothness constraints or partial specified topological information) and therefore further reduce the noise sensitivity. Hence, the development of scanners combining both anatomical and functional information is motivating the development of Bayesian maximum a posteriori (MAP) reconstruction techniques. The morphological information extracted from an X-ray computed tomography (CT) or magnetic resonance (MR) images may be used as prior information in the reconstruction of the PET images [Ardekani et al., 1996; Murase et al., 1996].

Studies in 2D showing this have been carried out successfully, but the 3D problem has not been yet addressed. Through the use of Bayes theorem, one can introduce any prior distribution to describe the statistical properties of the unknown image and thus produce *a posteriori* probability distributions from the image conditioned upon the data. Maximisation of the *a posteriori* probability over the set of possible images results in the MAP estimate. From the developer's point of view, a significant advantage of this type of approach is its modularity : the various components of the prior, such as non-negativity of the solution, pseudo-Poisson nature of statistics, local voxel correlations (local smoothness), or known existence of anatomical boundaries, may be added one by one into the estimation process, assessed individually, and used to guarantee a fast working implementation of preliminary versions of the algorithms [Lalush et al., 1995; Lange et al., 1995].

Using a Bayesian model, the incorporation of prior information derived from a registered anatomical CT/MR image by a mechanism aids the good quality reconstruction of functional SPECT/PET images. This method encourages the reconstructed image to be piecewise smooth; the contribution is the inclusion of a coupling term that influences the creation of edges in the vicinity of significant anatomical edges which is generally modelled in a Gibbs prior distribution. Also, a Gibbs prior of piecewise smoothness could be used in the Bayesian model. Gibbs prior whose potential function is determined by a set of parameters, is designed to smooth noise with minimal smoothing edge information. The Gibbs prior has three parameters: one which determines the overall weight placed on the prior in the reconstruction process, and two others which affect the relative smoothing of noise and edges in the reconstructed image estimates.

However, it is generally more difficult to find the optimum reconstruction as compared to the ML-EM algorithm. To overcome this problem, Green et al. (1990) proposed the one-step-late (OLS) algorithm where near optimums solutions of the Bayesian formulation of the problem can be obtained by modifying the general form of ML-EM (Eq. 30) by a multiplicative factor using the derivative of an energy function.

$$\lambda_i^{n+1} = \frac{\lambda_i^n}{\sum_j p_{i,j} + W(\Lambda^n)} \sum_j \frac{p_{i,j}\, y_j}{p_{k,j}\, \lambda_k^n} \qquad \text{(Eq. 50)}$$

where $\Lambda^n = \{\lambda_i^n\}$ and $W(\Lambda^n)$ may be a penalty function. This function depends on the prior distribution assumed for the image and acts to smooth the image by scaling down λ_i when its value is higher than the neighbours and scaling it up when its value is smaller than its neighbours. The smoothing is the same throughout the image regardless of the position λ_i. While this method generally yields better noise properties, the reconstructed images have lower resolutions and increased partial volume effects.

To overcome this problem, some other new penalty function could be proposed which allows adjustments of the degree of smoothing used depending upon prior anatomical information about the image, or information derived during the iterations.

The Bayesian model could utilise a Gibbs prior as first suggested by German et al. (1984), to describe the spatial correlation of neighbouring regions and takes into account the effect of limited spatial resolution as well.

In terms of the unobservable latent data n and the observed counts y, the posterior distribution $p(\lambda|y)$, according to the Bayes theorem, can be expressed as

$$p(\lambda \mid y) = \int p(\lambda \mid n, y) p(n \mid y)\, dn \qquad \text{(Eq. 51)}$$

Calculation of an estimate of λ using the formula requires knowledge of both the predictive density p($n|y$) and the conditional posterior density p($\lambda|n,y$). It can be shown that sampling from $p(n|\lambda,y)$ is equivalent to redistributing the counts y to neighbouring pixels according to the multinomial probabilities

$$q_{ij} = \frac{\lambda_i p_{ij}}{(\sum_m p_{im}\ \lambda_{mj})} \qquad \text{(Eq. 52)}$$

where $p = \{p_{ij}\}$ as defined earlier. Hence, if the conditional posterior density $p(n|\lambda,y)$ can be specified, a new sample from λ can be drawn using Eq. 51 and Eq. 52. The algorithm proceeds iteratively until some preset conditions are satisfied.

From a Bayesian point, the conditional posterior density $p(n|\lambda,y)$ can be expressed as the product of the likelihood function and a prior density function. The likelihood function based on n can be defined as the product of independent Poisson distributions. The prior density function, on the other hand, is to be specified by the *a priori* information that one intends to incorporate into the reconstruction.

As mentioned earlier, one piece of *a priori* information that can be incorporated into the Bayesian model is the local continuity; i.e. pixel intensities with homogeneous regions tend to be similar. The local continuity could be modelled by using the Gibbs priors which has a density in the form of

$$P(\lambda) = (1/Z)\ exp\{-U(\lambda)\} \qquad \text{(Eq. 53)}$$

where Z is called the partition function and is a normalisation constant independent of λ. $U(\lambda)$ is defined as the "energy function" and can be expressed as

$$U(\lambda) = \sum_C V_C(\lambda) \qquad \text{(Eq. 54)}$$

where the functions $V_C(\lambda)$ are arbitrary functions that represent characteristics of the pixel intensities in a local neighbouring C. The functions $V_C(\lambda)$ are referred to as the "potentials".

The MAP algorithm can produce good quality reconstructions due to the ability to incorporate information from other sources such as CT scan and MRI.

3.7. "ME": Maximum entropy

In MAP methods, not all images are considered equally likely in the absence of data, i.e., images have a nonuniform *a priori* probability distribution. Maximum entropy methods form a class of MAP methods where the *a priori* probability of the images depends on the entropy of the images. Since the

entropy of an image increases with the smoothness of the image, the maximum entropy methods encourage smoothness in the reconstructed images [Desmedt et al., 1991].

From the equation Eq. 51, the solution proposed by the ML-EM algorithm is to find for which $p(\lambda|y)$ is maximised, i.e the probability $p(\lambda)$ is considered independent of λ. However, after a number of iterations increasing noise can considerably deteriorate the image. The introduction of prior information can solve this problem. The entropy prior proposed by Liang et al. (1989) (with α a weight factor):

$$p(\lambda) \approx \exp\left[\alpha \Psi(\lambda)\right] = \exp\left[-\alpha \sum_{i=1}^{V} \lambda_i \ln \lambda_i\right] \qquad \text{(Eq. 55)}$$

and the following iteration scheme is proposed

$$\lambda^{k+1} = \frac{1_V}{P^T 1_T - \alpha \nabla \Psi(\lambda^k)} \left(P^T \left(\frac{y}{P\lambda^k}\right)\right) \lambda^k \qquad \text{(Eq. 56)}$$

λ and y are column matrices and 1_V, 1_T are column matrices of dimension V, respectively T with all elements equal to 1. In the case of the Liang entropy prior, the following is valid:

$$\nabla \Psi(\lambda) = 1 + \ln(\lambda) \qquad \text{(Eq. 57)}$$

where the operation $P\lambda$, is the computation of the projection of the image on the detectors for different angles, and the operation $P^T d$, the computation of the backprojection of the data on the image. Due to the storage requirement for the matrix P, the latter is replaced by a product of different matrices. Each of these matrices models one physical effect, e.g. the scanner geometry, the attenuation, the efficiency of detectors, the positron range, etc...

Many different forms of the entropy formula are in use. The most obvious form is the Shannon entropy. Recently, the entropy has been extended to include a model of the image. This new entropy estimates the distance of the image to that model. In the maximum entropy method, this distance is kept minimal while a reasonable fit for the data is obtained. Many features of the image to reconstruct can be included in this model. Here, an intermediate method is studied: a single model entropy with changing model. The influence of changing the information in the model is examined.

The latest entropy formulas have two models [Desmedt et al., 1991]. The purpose of the extra model is to account for the general features of the image. This model typically represents a low frequency estimation of the image. The low frequency estimation is continuously updated in the iterative scheme. The second problem with maximum entropy methods is the determination of the regularization parameter. The regularization parameter determines the relative importance of the *a priori* information versus the information obtained by the measurement. Two methods to automatically determine this parameter are investigated. The first method is the method of moments. This method adapts the regularization parameter to fit the first two moments of the measurements. The second method is based on a marginalization procedure. During the iterations, the probability of the regularization parameter is estimated by marginalizations. The regularization parameter that maximises this posterior probability is used in the reconstruction. Reconstructions with software phantoms are used to evaluate the performance of the different methods. These experiments show that the regularization method can adapt itself to the correctness of the model used; i.e., if the model is not a good approximation of the image to reconstruct, not much attention is paid to that model, and vice versa.

The temporary conclusions are: the maximum entropy methods can indeed improve the reconstructed images of the standard ML-EM method. When using an entropy with model, the model should contain enough information. The performance of the filtered-backprojected images as a model in these methods will be studied in the near future.

4. CONCLUSION

The principal objective of reviewing all the set of existing reconstruction algorithms is :

- for comparative investigation of iterative and analytic methods using simulated and real patient data,
- for development and qualitative and quantitative assessment of improved or new volume reconstruction methods through the development of new advanced parallel algorithms [Egger et al., 1996],
- for achieving satisfactory processing rate and high quality of reconstruction which, for example, are essential for accurate analysis of the behaviours of tracers with time in different organs of the body.

All these analytic and iterative algorithms described above have inherent parallelism. Whilst their parallelisation is itself a research topic, the consortium feels it has the necessary expertise to undertake this effectively without distracting from the main objectives of the work. There are several computational and algorithmic challenges associated with iterative and analytic reconstruction techniques. These are :

- The shift from 2D to 3D modes of scanning has resulted in an order of magnitude increase in the volume of data. This greatly affects the computational performance and memory requirements of iterative methods which involve the generation and storage of a large matrix of dimensions directly proportional to the size of the data.
- For some iterative techniques, such as ML-EM, good stopping criteria are not known and various accelerators and preconditionners need to be set up in order to speed up the image reconstruction.
- The reconstructed images must be of a sufficiently high statistical quality in order to identify the time-varying components in the emission data.

In conclusion, the PARAPET project aims to address the above computational constraints and algorithmic challenges with reference to real clinical requirements. By the end of project, it is intended to be able demonstrate such reconstruction in a clinically acceptable time frame through the use of parallelism combined with the increase in performance of individual microprocessors. For some algorithms, data acquisition and processing can proceed simultaneously. The implementation of such algorithms can have clear clinical benefits.

5. REFERENCES

Ardekani B.A., Braun M. et al. **(1996)** Minimum cross entropy reconstruction of PET images using prior anatomical information. *Phys Med Biol*. 41:11:2497-2517.

Belleman M.E., Schmidlin P. et al. **(1995)** Characteristics of a maximum-likelihood algorithm for iterative PET image reconstruction. *Radiology*. 197:223-223.

Bracewell R.N. and Riddle A.C. **(1967)** Inversion of fan-beam scans in radio astronomy. *Astrophysical Journal*. 150:2:427-434.

Browne J. and Depierro A. **(1996)** Row-action alternative to the EM algorithm for maximizing likelihood in emission tomography. *IEEE Trans Med Imag*. 15:5:687-699.

Colsher J.G. **(1980)** Fully three-dimensional positron emission tomography. *Phys Med Biol*. 25:1:103-115.

Comtat C., Egger M.L., Scheurer A.K.H., Morel C., Defrise M. and Townsend D.W. **(1994)** Use of the 3D Favor algorithm with large aperture PET scanners. *Conference Records 1993 IEEE Medical Imaging Conference, San Francisco, USA:* 1677-1681.

Comtat C., Morel C., Defrise M. and Townsend D.W. **(1993)** The Favor algorithm for 3D PET data and its implementation using a network of transputers. *Phys Med Biol*. 38:929-944.

Daube-Witherspoon M.E. and Muehlener G. **(1996)** Iterative image space reconstruction algorithm suitable for volume. *IEEE Trans Med Imag*. 5:2:61-66.

Daube-Witherspoon M.E. and Muehllehner G. **(1987)** Treatment of axial data in three-dimensional PET. *J Nucl Med*. 82:1717-1724.

Defrise M **(1995)** A factorisation method for the 3D x-ray transform, *Inverse Problems 11 pp 983-994*.

Defrise M., Kinahan P. and Townsend D. **(1995)** A new rebinning algorithm for 3D PET: Principle, implementation and performance. *Conference Records 1995 International meeting on fully three-dimensional image reconstruction in radiology and nuclear medicine*, *Aix-Les-Bains, France*. 235-239.

Defrise M., Sibomana M., Michel C. and Newport D. **(1995)** 3D PET reconstruction with the ECAT EXACT HR using Fourier rebinning. *Conference Records 1995 IEEE Medical Imaging Conference, San Francisco, USA*: 1316-1320.

Defrise M., Townsend D.W. and Clack R. **(1989)** Three-dimensional image reconstruction from complete projections. *Phys Med Biol*. 34:5:573-587.

Defrise M., Townsend D.W. and Clack R. **(1992)** Favor: a fast reconstruction algorithm for volume imaging in PET. *Conference Records 1991 IEEE Medical Imaging Conference, Santa Fe, USA*. 1919-1923.

Depierro A.R. **(1995)** Modified expectation maximization algorithm for penalized likelihood estimation in emission tomography. *IEEE Trans Med Imag*. 14:1:132-137.

Desmedt P. **(1991)** Maximum Entropy and Bayesian Image Reconstruction: Overview of Algorithms. LEM Technical report DG91-13, University of Gent.

Egger M.L., Hermann Scheurer A., Joseph C. and Morel C. **(1996)** Fast volume reconstruction in positron emission tomography: Implementation of four algorithms on a high performance scalable parallel platform. *Conference Records 1996 IEEE Medical Imaging Conference, Anaheim,CA, USA*: 1574-1578.

Egger M. **(1996)** Fast volume reconstruction in positron emission tomography. *PhD thesis, University of Lausanne, Lausanne, Switzerland*.

Fessler A. and Hero A.O. **(1994)** Space Alternating Generalized Expectation Maximization Algorithm. *IEEE Trans Sig Proc*. 42:2664-2677.

Fessler A. and Hero A.O. **(1995)** Penalized maximum-likelihood image reconstruction using space-alternating generalized EM algorithms. *IEEE Trans Imag Proc*. 4:1417-1429.

German S. and German D. **(1984)** Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Trans Patt Anal Mach Intell*. 6:721-741.

Golub G. and Van-Loan C. **(1983)** Matrix computation. Press John Hopkins University. USA.

Green P.J. **(1990)** Bayesian reconstructions from emission tomography data using a modified EM algorithm. *IEEE Trans Med Imag*. 9:84-93.

Herman G. **(1995)** Image Reconstruction From Projections. *Real Time Imaging*. 1:3-18.

Herman G.T. and Meyer L.B. **(1993)** Algebraic reconstruction techniques can be made computationally efficient. *IEEE Trans Med Imag*. 12:600-609.

Hudson H.M. and Larkin R.S. **(1994)** Accelerated image reconstruction using ordered subsets of projection data. *IEEE Trans Med Imag*. 13:4:601-609.

Johnson V.E., Wong W.H., Hu X. and Chen C.-T. **(1991)** Bayesian restoration of PET images using Gibbs priors. *Information Processing in Medical Imaging*. New York. Wiley-Liss: 15-28.

Kaufman L. **(1993)** Maximum Likelihood, Least Squares and Penalized Least Squares for PET. *IEEE Trans Med Imag*. 12:2:200-214.

Kaufman L. and Neumaier A. **(1995)** Image Reconstruction Through Regularization by Envelope Guided Conjugate Gradients. AT&T Bell Laboratories, 600 Mountain Avenue, Murray Hill, NJ 07974-0636, USA.

Kaufman L. and Neumaier A. **(1996)** PET regularization by envelope guided conjugate gradients. *IEEE Trans Med Imag*. 15:3:385-389.

Kinahan P.E. and Rogers J.G. **(1989)** Analytic 3D image reconstruction using all detected events. *IEEE Trans Nucl Sci*. 36:1:964-968.

Lalush D.S. and Tsui B.M.W. **(1995)** Fast and stable maximum a-Posteriori conjugate gradient reconstruction algorithm. *Med Phys*. 22:8:1273-1284.

Lange K. and Fessler J.A. **(1995)** Globally convergent algorithms for maximum a-posteriori transmission tomography. *IEEE Trans Imag Proc*. 4:10:1430-1438.

Lewitt R.M., Muehllehner G. and Karp J.S. **(1994)** Three-dimensional reconstruction for PET by multi-slice rebinning and axial image filtering. *Phys Med Biol*. 39:321-340.

Liang Z., Jaszczak R. and Geer K. **(1989)** On bayesian image reconstruction from projections: uniform and nonuniform a priori source information. *IEEE Trans Med Imag*. 8:227-235.

Llacer J., Veklerov E., Baxter L.R., Grafton S.T., Griffeth L.K., Hawkins R.A., Hoh C.K., J.C. Mazziotta, Hoffmann E.J. and Metz C.E. **(1993)** Results of a clinical receiver operating characteristic study comparing filtered backprojection and maximum likelihood estimation images in FDG PET studies. *J Nucl Med*. 34:7:1198-1203.

Matej S. and Browne J.A. **(1995)** Performance of a fast maximum likelihood algorithm for fully 3D PET reconstruction. *1995 International meeting on fully three-dimensional image reconstruction in radiology and nuclear medicine*. Aix-les-Bains, FRANCE. 247-251.

Meikle S.R., Hutton B.F., Bailey D.L., Hooper P.K. and Fulham M.J. **(1996)** Accelerated EM Reconstruction in Total Body PET: Potential For Improving Tumor Delectability. *Research Report, Department of Nuclear Medicine Royal Prince Alfred Hospital Missenden Road , Camperdown 2050, Australia*.

Murase K., Zhang Y. et al. **(1996)** Usefulness of a newly developed MRI constrained PET image: reconstruction method. *J Nucl Med*. 37:5 Suppl:1009.

Natterer F. **(1986)** *The Mathematics of Computerized Tomography*. Wiley.

Ollinger J.M. and Goggin A.S. **(1996)** Maximum Likelihood Reconstruction in Fully 3D PET via the SAGE algorithm. Technical Report, Washington University, St. Louis, MO 63110.

Orlov S.S. **(1975)** Theory of three-dimensional reconstruction: Conditions of a complete set of projections. *Soviet Physics Crystallography*. 20:312.

Osullivan F. **(1995)** Study of least-squares and maximum-likelihood for image reconstruction in positron emission tomography. *Annals of statistics*. 23:4:1267-1300.

Pelc N.J. and Chester D.A. **(1979)** Utilization of cross-plane rays for three-dimensional reconstruction by filtered backprojection. *J Comp Assist Tom*. 3:385-395.

Ra J.B., Lim C.B., Cho Z.H., Hilal S.K. and Correll J. **(1982)** A true three-dimensional reconstruction algorithm for the spherical positron emission tomograph. *Phys Med Biol*. 27:1:37-50.

Radon J. **(1917)** Über die Bestimmung von Funktionen durch ihre Integralwerte langs gewisser Mannigfaltigkeiten. *Berichte der Schsischen Akademie der Wissenschaften*. 69:262-277.

Schorr B., Townsend D.W. and Clack R. **(1983)** A general method of three-dimensional filter computation. *Phys Med Biol*. 28:305-312.

Shepp L.A. and Vardi Y. **(1982)** Maximum Likelihood Reconstruction for Emission Tomography. *IEEE Trans Med Imag*. 1:2:113-122.

Solmon D.C. **(1976)** The X-ray transform. *Jour Math Anal Appl* 56:61-83.

Stearns C.W., Chesler D.A. and Brownell G.L. **(1987)** Three-dimensional image reconstruction in the Fourier domain. *IEEE Trans Nucl Sci*. 36:374-378.

Vermeulen F. and Lemahieu I. **(1989)** An Evaluation of Maximum-Likelihood Reconstruction of PET Images in the Presence of Low Quality Attenuation Correction Measurements. *Proceedings of the Vth Mediterranean Conference on Medical and Biological Engineering* (*Medicon 89*), *Patras, Greece*: 336-337.

Zhou Z., Leahy R.M. and Mumcuoglu E.U. **(1994)** A comparative study of the effects of using anatomical priors in PET reconstruction. *Conference Records 1994 IEEE Medical Imaging Conference, Norfolk, USA*: 1749-1753.