Intrinsity™ FastMATH™ Application to Computed Tomography

Written by G. P. Yost, PhD

1. Introduction

This document describes the science of computed tomography and the application of the Intrinsity™ FastMATH™ processor to this problem. The FastMATH processor is designed as a high-speed parallel-processing matrix engine, programmable in C or other high-level languages, and we will show that it is well-suited to tomography problems.

We begin with a brief discussion of computerized tomography (CT, sometimes called CAT for computerized axial tomography) systems. For specificity, we concentrate on x-ray systems designed for scanning of airport baggage, but most of the same principles apply to medical x-ray tomography, positron emission tomography (PET), single-photon emission tomography (SPECT), and other systems. In its most sophisticated form, the system makes use of a collimated cone-beam x-ray source that emits pulsed x-rays at rapid intervals. A detector system configured in finite-sized elements lies in the beam on the opposite side of a conveyor belt upon which the scanned object (e.g., luggage) is being moved continuously. The source and detector co-rotate about the conveyor belt as it moves so that the luggage is scanned from multiple angles. The detector elements record the line integral of the absorptivity on each exposure. By use of the Fourier slice theorem, the x-ray absorption of each volume element can be reconstructed from a complete set of these line integrals. The space resolution will depend on a number of factors, including the number of scans per unit angle, the angular range covered, and several others. By using more than one energy band, some discrimination of chemical elements is possible.

Systems capable of up to 30 images per second are available. Modern CT systems can generate huge amounts of data. Computational requirements can be a limiting factor in achieving high spatial resolution, especially in real-time situations. Most of the image-reconstruction computation involves manipulation of arrays; therefore, a processor optimized for vector and matrix processing is what is needed.

Intrinsity, Inc. (www.intrinsity.com) has launched the FastMATH processor, designed for exactly that type of requirement: very fast vector and matrix mathematics involving large data sets. This processor is configured with a MIPS32®-compliant scalar unit operating synchronously with a 4 × 4 array processor capable of full vector and matrix mathematics and single instruction, multiple data (SIMD) mode. The cycle speed is 2 GHz, verified on the hardware. With the matrix unit it is capable of sustaining a 32 giga-operations per second (GOPS) computation rate, with a peak rate of 64 GOPS. Two bi-directional RapidIO™ ports, capable of simultaneous 1 Gbyte per second data rates in both directions, can provide data at a rate compatible with the processor. A core tomography routine has been benchmarked and impressive results have been achieved.

2. Computed Tomography

Tomography is the science of reconstructing hidden internal structure from radiation transmitted through a subject. It is therefore non-invasive and has proven extremely valuable in medical and industrial applications. The radiation may be in the form of x-rays that transmit the subject entirely, or it can be radiation emitted from a source within the subject. In this case, the subject is usually a patient who has ingested a radioactive substance. In tomography what is usually desired is three-dimensional information from inside the subject, and not just a two-dimensional projection like a standard chest x-ray provides. If the subject is luggage or cargo being scanned at an airport, for example, the object is to locate and identify suspicious objects that might be explosives or other contraband. Such objects might be

concealed within x-ray absorbent material, which could be as simple as a book inside which a small sheet of plastic explosive is placed. In a patient, the object might be to image an organ in three dimensions, sometimes in real time. For example, a beating heart, the progress of a surgical procedure such as an angioplasty, or the real-time functioning of the brain are common real-time applications. A recent review of the field is given by Kalendar [12].

Three-dimensional reconstruction will require multiple exposures from many angles. In x-ray tomography a common approach is to circle the subject with a rotating gantry to which is attached one or more sources collimated to irradiate the subject. After passing through the subject the beam strikes an array of detectors also attached to the gantry. As the gantry rotates the source emits pulses of x-rays. The detectors measure the absorption profiles, pulse by pulse. Therefore, the subject is exposed from many angles.

We draw heavily on the thesis by Turbell [2].

2.1. X-Rays

An x-ray tube will typically emit a broad angular distribution of photons over a broad energy spectrum. Definitions of what constitutes an x-ray are not consistent in the literature. Some authors choose to describe the difference between x-rays and what is known as gamma rays based on their production mechanism (atomic transitions for x-rays, nuclear transitions for gamma rays; it is not clear which definition is used for synchrotron radiation). Here, we will concentrate on the more physically interesting definition in terms of energy because that is what determines utility. Even there, different authors draw different boundaries; best is to leave the boundaries vague and overlapping. A common definition is that x-ray wavelengths range from around 0.1 micron (low energy) down to maybe 0.01 nm (high energy), a range of four orders of magnitude. Recall that the energy of a photon is inversely proportional to its wavelength. For comparison the highest energy visible photons, at the extreme visible range of blue, have a wavelength only down to about 0.4 μm. At the high-energy end the x-ray spectrum overlaps the gamma-ray spectrum; at the low energy end it overlaps with UV. Such short wavelengths for x-rays minimize diffraction so that straight-line optics provides a good approximation down to very small angles. X-rays may be scattered out of the beam (occasionally into a different detector element) or absorbed. These wavelengths do not permit coherent beams to be easily generated (without a particle accelerator), but that is not needed for tomography. An x-ray beam is typically “steered” by a collimator, which simply removes those photons going in undesired directions. Note that a collimator is not any sort of lens; it is a block of absorbent material (e.g., lead) with one or more holes in it. It is an aperture. Unused x-rays are prevented from irradiating other parts of the room (or patient).

X-ray lenses are not practical. Crystal structure can be analyzed due to diffraction off the atomic planes. This does involve changing the directions of x-rays by small amounts, but this is not practical as a focusing mechanism. Crystal grazing-incidence diffraction gratings have been successfully used in space-borne astronomical observatories to energy-select and re-direct x-rays, but this is a specialized and expensive operation. The Chandra x-ray observatory has obtained high-resolution images of point sources using Indium-coated highly nonspherical grazing incidence mirrors. See [6] for a discussion of x-ray optics.

Note that, in x-ray tomography, images of the source are not of interest because the source is outside the subject. However, other medical imaging technologies do have interest in forming an image of the source, as will be discussed.

2.2. Detectors

Numerous technologies exist for x-ray detection [1]. These will give a signal proportional to the number of photons striking the detector element. This can be a voltage or a simple photon count. Detector response is energy-dependent. Large area flat-panel detectors with very fine pixilation (as small as 50 μm has been achieved over large areas) and fast response times are becoming available. These hold the promise of lower patient dose and higher space and time resolution. All of this can translate into shorter exposure times to achieve a given resolution. If and when these are put into use for high-resolution tomography, the processing requirements will increase enormously.

Detector elements usually require a set of calibration constants. These will express the efficiency of each element for the energy spectrum directed towards it. The angular distribution of the source may also not be uniform, requiring additional calibration as a function of the angle of each detector element.

2.3. Linear Absorption Coefficient

X-ray systems do not directly measure density; they measure the absorptivity, also called the linear absorption coefficient. This is dependent on the density but also on the atomic number, *Z*, and can be strongly affected by specific elements in the subject. What we can infer from the absorption is some function of its local density and *Z* as a function of location. Tomography enables us to start from the total absorption along a set of rays through the subject and find the local absorption coefficient in small volumes. It does this from a set of exposures at different angles through the same volume element and using Fourier analysis. Roughly speaking, one is able to "average out" the contributions of intervening materials by looking from all angles.

2.4. Setup; Slices

CT scanning equipment provides for a beam to pass through the subject and strike a set of detector elements on the other side. In the simplest mode the source and detectors co-rotate around each subject element 360°, exposing the subject at many equally-spaced angles, called view angles, before the subject moves. A complete set of view angles defines one slice. The axis of rotation is defined as the z axis. It is usually configured down a longitudinal axis through the center of the subject. The normal to the slice is parallel to the z direction. After exposure at a chosen z, the subject may be moved on a platform and stopped for a new scan. The direction of motion is along z. This is called constant z-axis (CZA) scanning because an entire scan of one slice of the subject is at the same z. Most present-day medical tomography uses CZA. However, some newer systems move baggage or patients continuously for practical reasons. Therefore, the source and detectors trace a helical path around each volume element and so a given volume element may be at the periphery of the beam for some of the scanning angles. This is called constant-speed helical scanning (CSH). If the beam has an angular spread in the z-direction, and the detector elements are in a corresponding two-dimensional array, then a number of slices may be defined for each scan. Some of these will be at an angle with respect to the z-axis.

We begin with an assumption of parallel x-rays from multiple angles at a fixed z and then progress to a fan beam, to a cone beam, and finally to a helical scan with a cone beam. Each new level means increased complexity. We also discuss the advantages of dual-energy scanning for distinguishing chemical elements.

These techniques are suitable for medical scanning as well as luggage scanning. Of course, lower doses are required for medical scanning, with consequences for resolution.

3. Introduction to X-Ray Tomography

We construct a coordinate system in which the *z*-axis is the axis of rotation, assumed to pass through the scanned subject near its center line. Let the *y*-axis be vertical and construct the *x*-axis for a right-handed coordinate system. If *f(x,y,z)* is the x-ray absorptivity (linear attenuation coefficient) at a specified location, then the intensity, *I*, of transmitted x-rays is related to the initial intensity I_0 by

$$I = I_0 \exp\left\{-\int_L f(x,y,z)\,dl\right\} \quad , \qquad \textit{Eq.(1)}$$

where the line integral is taken along the ray through the sample. To achieve adequate resolution, it is desirable to use a strong enough flux that a significant number of x-ray photons pass through the object and are registered in each detector element. The constraints on this are obvious: safety to operating personnel and objects (e.g., unexposed film in luggage), or to patients. There may also be a saturation constraint from the detector or its electronics.

The object of the exercise is to measure *f* in three dimensions throughout the volume of interest.

3.1. The Physics of X-Ray Absorption

The physics of x-ray absorption tells us that *f* will depend both on the density, ρ, and the atomic number, *Z*, at each location. There are several principal mechanisms for x-ray interactions with matter, which are sometimes listed differently by different authors. We categorize them as photoelectric absorption, Compton scattering, coherent scattering, pair production, and photonuclear absorption. Photoelectric absorption refers to the absorption of the photon by

an atomic (electron energy level) transition. The photon may be thought of as having been absorbed by the atom as a whole, rather than a particular electron (although it may be just one electron that changes energy level). In some cases, the absorbed energy is subsequently re-emitted as one or more photons, but often not in the same direction or even necessarily as a single emission at the same energy. Sometimes an electron is emitted. Compton scattering refers to the scattering of a photon off an electron, often accompanied by electron ejection. A reduced-energy, deflected, photon is emitted. Coherent scattering results in a small-angle deflection of the x-ray leaving the atom neither excited nor ionized. Depending on the size of the detector elements, it may not cause loss of beam to that element and, therefore, not contribute to attenuation. Pair production contributes only at energies above twice the mass of the electron, i.e., a wavelength of about 1 fm or shorter, well into the gamma ray region and above the energies of interest for x-ray tomography. Photonuclear absorption may be accompanied by emission of a neutron or other particle; it tends to contribute at energies well above the x-ray region. Some of these higher-energy mechanisms may be of interest to tomographs that use gamma rays.

Photoelectric absorption dominates at low energies (long wavelengths), being replaced primarily by Compton scattering as the energy increases (the wavelength decreases). Strong peaks and valleys as a function of photon energy are observed for most elements, depending on atomic structure. Photon attenuation is often tabulated as the mass attenuation coefficient, which is the attenuation divided by the density. This separates the Z-dependent effects from density effects. A summary and detailed information may be found in reference [5].

3.2. Energy Dependence

The reason these mechanisms are important is that their strengths vary differently with atomic number and x-ray energy. Total attenuation at a given energy, therefore, depends on both ρ and Z. A difference in attenuation could result from a difference in either one. This could be important in a lot of contexts. For example, in a baggage-scanning system the atomic number of substances in the luggage could range up to $Z = 92$ (uranium) or 94 (plutonium).

It is possible to separate density effects from atomic number effects by using more than one energy. The photoelectric effect varies approximately as $\rho Z^4/AE^3$ and Compton scattering varies approximately as $\rho Z f_{KN}(E)/A$, where A is the atomic number (which increases through the periodic table somewhat faster than Z). The Klein-Nishina function, f_{KN}, for the energy dependence of Compton scattering is given in many physics texts and may be found in [3], one of the classic references in the field of energy-dependent tomography. It has a much slower energy fall-off than $1/E^3$. Therefore, the attenuation of a low-energy beam will vary with Z primarily as Z^4/A (therefore, roughly as Z^3) and a high-energy beam primarily as Z/A (roughly independent of Z; note that the number of electrons per unit volume, which will determine the reaction rate, is roughly proportional to density). Measuring attenuation by at least two well-separated energies allows estimation of both ρ and Z. In practice, integrals over energy intervals, with corrections for emission spectra and detector sensitivity, are usually involved. Typically, different energy intervals may be selected via exposures through different energy-selective filters (this requires different exposures, with the different filters interposed) or by use of a two-component detector sensitive to two different energy bands at the same time. It is also possible to use different sources with different emission lines to selectively provide high-energy or low-energy beams (again requiring different exposures).

Note that, even with two energies, we can only obtain two data points. These can be taken to be average density and average Z. Any further differentiation would require additional energy-dependent data points.

3.3. Background Processes

Complicating this analysis are absorption lines present in many elements (resonance absorption around specified energies) and uncertainties in the emission spectrum and detector efficiencies. Beam-hardening is another complicating effect. If intervening matter has high Z the low-energy components of the beam may be preferentially absorbed and the energy spectrum assumed for the beam may not apply behind the high Z material.

Secondary photons may enter the detectors, especially after Compton scattering. These secondary photons may not follow the same trajectory as the primaries and, therefore, are a source of background. If they happen to enter the direct-line detector element, then the attenuation of the object is less than assumed; some of the photons that were calculated to be scattered away, were not. If they enter a different detector element, they reduce the apparent attenuation there. Because Compton scattering gives an angular distribution peaked in the forward direction (whose width is

energy- and *Z*-dependent), detector elements slightly off the direct path through a region of high *Z* can see a spurious attenuation dip (extra photons). This often reconstructs, for example, as bright streaks near metallic objects. Other sources of background are cosmic radiation, natural radioactivity, statistical fluctuations, and noise effects in the detectors. Some of these can be reduced by the use of electronic timing gates. There can also be photons scattered off the collimators or other hardware.

Schemes have been proposed [4] to actually take advantage of off-axis Compton scattering, which is dependent upon electron density. One can, in principle, gain a handle on chemical composition in this way.

Blurring artifacts can arise from the finite size of the x-ray source and detector elements. Image blurring can be reduced with appropriate image-processing algorithms, such as iterative maximum likelihood deblurring. Motion artifacts can arise with subject movement during exposure.

3.4. Accuracy

Algorithms may be separated into "exact" and "non-exact" types. Reconstruction accuracy will depend on the background effects above, the number of scans per unit angle, the segmentation of the detector, and the statistical significance of the detector output (dependant upon the incident flux). An exact algorithm is capable of reconstruction that can be brought closer to the true attenuation coefficient (in the absence of background) than any specified tolerance by sufficient improvements in those parameters. A non-exact algorithm cannot be brought to arbitrary accuracy, but it is useful when accuracy requirements permit and computational convenience is desired. Given the presence of background, non-exact algorithms are sometimes perfectly adequate and may be preferred for computational or other reasons.

The measurements are fundamentally counting experiments and Poisson statistics apply. Statistical errors are determined by Poisson fluctuations in the number of counts and, thus, obey a $\sqrt{n}$ law, where n is the expectation value of the number of counts. Typical variances for tomography are given in [3].

3.5. Related Technologies

A number of imaging technologies use tomographic reconstruction. With some differences in detail, most of the reconstruction mathematics will be the same. Of course, data acquisition, formatting, and noise reduction may vary.

Magnetic resonance imaging (MRI) may be one of the best known tomographic systems (see [15] for a review). The technology makes use of the spin magnetic dipole moment of many atomic nuclei (electron dipole moments can also be used). With application of a strong magnetic field around the patient, this dipole moment causes these types of nuclei to preferentially align in the minimum-energy state for the field. Energy from a radio frequency pulse will be absorbed if it causes the nucleus to flip its spin. This can occur if the pulse frequency corresponds to a resonance of the nuclear magnetic moment. MRI provides for this absorption to be detected as a function of volume. This resonance frequency will depend to a small but significant extent on the electronic structure in which the nucleus is embedded. This is because of shielding of the nuclear moment by the electrons' own intrinsic magnetic moments and their orbital moments. Thus, the resonance lines can vary (typically by parts per million) with chemical structure, so MRI can identify molecular concentrations. Fourier transforms are a central part of MRI processing.

Also in the medical field, positron emission tomography (PET) uses back-to-back gamma rays produced in electron-positron annihilation to image the location where a radioactive source has settled within the patient. In a typical implementation, a radioactive isotope of fluorine (Z = 9) can be incorporated into glucose and injected into the patient. The glucose will concentrate in energy-using centers of the body, such as speech centers in the brain during conversation. The fluorine decay emits a low-energy positron that almost immediately annihilates on a near-by electron. This annihilation, because it takes place at low energy, produces two (almost) back-to-back gamma rays of energy about 1 MeV each. Detectors at opposite sides of the patient can register a coincidence that helps reduce background and tomography may be performed.

Single-photon emission computed tomography (SPECT) is a bit similar, except that it directly uses a gamma ray emitted by a chosen radioactive source. At least a couple of dozen different radionuclides are in common use. This has additional challenges from background but has considerable use.

Both of these have challenges measuring the direction of travel of the gammas, a problem not encountered in the x-ray CT systems discussed here. Typically, this is done with collimators. In x-ray CT, the direction of travel is deter-

mined by the known locations of the source and the detector (apart from background processes as discussed above). Another difference is that x-ray tomography measures attenuation to infer properties of the subject, whereas PET and SPECT investigate the location and activity of the source within a patient. The source is what is being imaged. For those, attenuation is an undesirable effect contributing to patient-absorbed ionizing radiation dose while reducing or corrupting the signal.

Electrical impedance tomography (EIT), imaging the distribution of electrical conductivity in three dimensions, has been developed for early detection of breast tumors [13]. Magnetic induction tomography (MIT) [14] uses an oscillating magnetic field and measures phase differences between inductor (inside the subject) and detector. The filtered backprojection (described below) takes place along lines of the magnetic field.

4. Parallel-Beam Tomography

A setup for parallel-beam tomography is shown schematically in Figure 1. The coordinate system is as described above, with the z-axis out of the page. The object being scanned is shown shaded, with two of the many angular projections also shown. At each angle the parallel beam makes an absorption projection onto the detector array (dark line). At this point we assume that the rays are collimated to lie all in a narrow plane at fixed z, so the detector array is a single row of elements. Each detector element measures the total x-ray absorption along a known ray through the object. This detector array describes an axis called the t-axis, which is fixed to the detector; its origin lies at the projection of the space coordinate system origin. A particular volume element is shown with darker shading. Note that for one projection it is to the left of the t-axis origin and for the other it is to the right. As the equipment rotates about the z axis, the projection of any off-axis volume element will move sinusoidally about the origin. For this reason the projection onto the t-axis is sometimes called a sinogram. The two rays shown passing through that volume element have no other regions of the object in common. For this reason it is possible to isolate the contribution of that volume element working from the full set of absorption profiles.

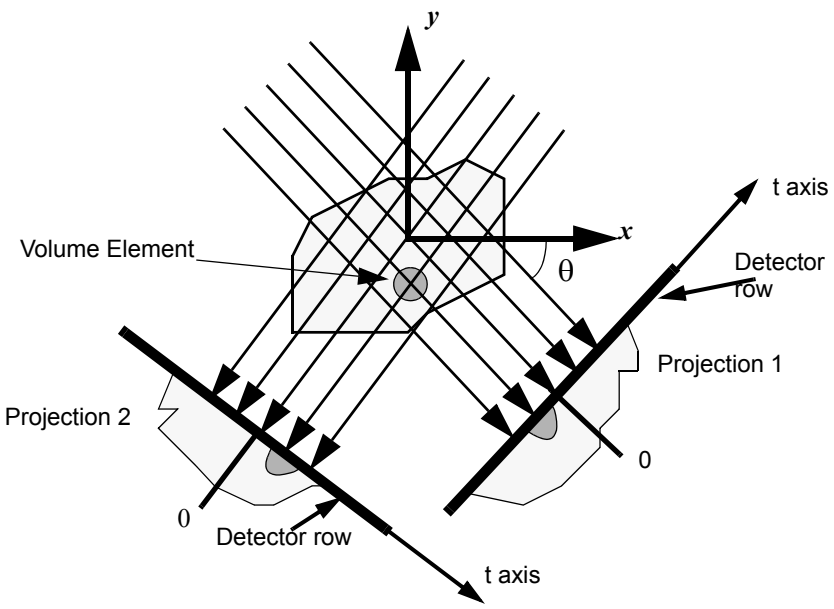
If we have the two-dimensional Fourier transform of the object's attenuation coefficient, *f(x,y,z)*, then we can reconstruct the attentuation coefficient itself by taking the inverse transform. From any one projection along an axis t at angle θ, we can calculate the one-dimensional Fourier transform. The Fourier slice theorem states that this is a one-dimensional slice of the full two-dimensional transform, and it lies along an axis in frequency space parallel to the t-axis. We will call this axis τ. We can reconstruct the two-dimensional transform with a full scan of slices covering the two-dimensional space. This is true whether these slices are arranged in a parallel array or are arranged in a rotating pattern. We can then transform back to the Cartesian coordinate system in measurement space.

With parallel slices in a Cartesian system every unit area will be covered by an equally dense set of slices. With a pattern of slices that rotate about a common origin, the area around the origin will be mathematically much more densely covered than areas farther out. This radial distribution of coverage will be corrected by a weighting filter that arises from the transformation to polar coordinates to follow.

First construct a Cartesian coordinate system in frequency space, with axes τ_x and τ_y. Then,

$$f(x,y,z) = F^{-1}{}_2F_2[f(x,y,z)] = \iint [F_2 f](\tau_x, \tau_y, \tau_z) e^{j2\pi(x\tau_x + y\tau_y)} d\tau_x d\tau_y \quad , \qquad \textit{Eq.(2)}$$

where z and τ_z are constants, denoting the scan plane normal to the z axis. F_2 and its inverse are the two-dimensional Fourier transforms. The inverse transform is explicitly shown by the double integral. Note in the integrand that the square bracket notation expresses the Fourier transform of *f*, $[F_2 f]$, in terms of the τ_x, τ_y, and τ_z axes; the parentheses specify the argument of this operator.



View of tomograph along z-axis (out of page). Parallel beam source (not shown) and detector co-rotate about z-axis; object to be scanned is shown shaded; a small volume element is shown with darker shading. Two rotated exposures shown. Projections give the line integrals of the linear absorption coefficient along the rays to each detector element.

Figure 1: Schematic of parallel beam tomography

We now note in Eq.(2) that the integrations are performed in a Cartesian coordinate system whereas the measurements are made in a rotating coordinate system. We therefore must transform the integrand into a space of rotated slices. Construct axis τ in frequency space parallel to a given t-axis. Let the propagation direction of the rays be θ, relative to the positive x-axis (Figure 1). Then θ is the angle the t-axis makes with the positive y-axis, and also the angle the τ axis makes with the positive τ_y axis. The differential volume element transforms into polar coordinates as

$$d\tau_x d\tau_y = \tau d\tau d\theta \quad , \qquad \textit{Eq.(3)}$$

and therefore the integral, Eq.(2), becomes

$$f(x,y) = \int_0^{2\pi}\int_0^{\infty} [F_2 f](-\tau\sin\theta, \tau\cos\theta) e^{j2\pi\tau(-x\sin\theta + y\cos\theta)} \tau d\tau d\theta \quad , \qquad \textit{Eq.(4)}$$

where we suppress the z-coordinate for simplicity. The factor τ corrects for the radial coverage problem mentioned earlier. Although the integral over τ has infinite range, in practice we will take the Fourier transform over a finite number of samples in t-space and therefore there will be a finite number of frequencies in τ-space.

The term $[F_2 f](-\tau \sin\theta, \tau \cos\theta)$ is the two-dimensional Fourier transform constrained to lie along the τ-axis at angle θ. That is a one-dimensional Fourier transform along that axis. If we write $p(t, \theta)$ as the t-axis projection of $f(x,y)$ at a particular θ and perform the $[F_2 f]$ operation, then we can recast Eq.(4) as

$$f(x,y) = \int_0^{2\pi}\int_0^{\infty}\int_{-\infty}^{\infty} p(t,\theta) e^{-j2\pi\tau t} dt e^{j2\pi\tau(-x\sin\theta + y\cos\theta)} \tau d\tau d\theta \quad . \qquad \textit{Eq.(5)}$$

Construct the distribution

$$g_\infty(t) = \frac{1}{2}\int_{-\infty}^{\infty} |\tau| e^{j2\pi\tau t} d\tau \quad . \qquad \textit{Eq.(6)}$$

This distribution functions as a filter, sometimes loosely called a ramp filter due to the shape of |τ|, a radial coordinate in frequency space. Note the convenient mathematical fiction of integration of dτ over (-∞,∞), replacing τ by |τ|, and dividing by 2 to compensate for doubling the range. The factor |τ| will assign higher weights to higher frequencies. This can have undesirable side effects in finite sampling problems, as all real problems are. High-frequency components from sharp edges can get amplified and form image streaks. Note that |τ| is not square integrable over an infinite range and therefore g_∞ is not a Fourier transform. Continuing,

$$f(x,y)=\int_0^{2\pi}\int_{-\infty}^{\infty} p(t,\theta)g_\infty(y\cos\theta - x\sin\theta - t)dtd\theta \qquad \textit{Eq.(7)}$$

$$= \int_0^{2\pi}[p\otimes g_\infty](\theta, y\cos\theta - x\sin\theta)d\theta \quad , \qquad \textit{Eq.(8)}$$

where the ⊗ notation signifies convolution over t, justifying the designation of g_∞ as a filter, and the two-dimensional argument following the bracket applies to the result of the convolution.

4.1. Filtered Backprojection (FBP)

Eq.(8) is known as filtered backprojection. It is filtered because of the transformation to polar coordinates, and it is backprojection because it mathematically "projects" the integrated linear attenuations back into the sample to reconstruct the attenuation coefficient as a function of volume element.

Real apparatus requires finite sampling. Let the detector length be 2 t_{max} and the number of equally-spaced elements be N_t. Let N_θ be the number of equally-spaced angles at which the exposures are made. We note at this point that the attenuation should be the same in one direction as in the reverse direction: *p(t, θ) = p(-t, π-θ)*. Therefore, it is not always necessary to scan the full rotation before moving to the next z position, and N_θ may only refer to a scan through π radians. However, any beam-hardening effects will be different from one direction than from the other, since a volume element that is behind a high-Z concentration from one direction will be in front of it from the other. We will, therefore, restrict ourselves for the present analysis to full (2π) scans.

If $\Delta t = 2\, t_{max}/N_t$, then the band-limited ramp filter becomes

$$g(t)=\int_{-1/(2\Delta t)}^{1/(2\Delta t)} |\tau| e^{j2\pi\tau t}d\tau = \frac{1}{2(\Delta t)^2}\mathrm{sinc}\left(\frac{2t}{2\Delta t}\right) - \frac{1}{4(\Delta t)^2}\mathrm{sinc}^2\left(\frac{t}{2\Delta t}\right) \quad . \qquad \textit{Eq.(9)}$$

We use the definition of the sinc function,

$$\mathrm{sinc}(x)=\frac{\sin(\pi x)}{\pi x} \quad . \qquad \textit{Eq.(10)}$$

Introduce indexes *i* to refer to the angle sample θ_i and *k* to refer to detector sample t_k, where $t_0 = -t_{max} + \Delta t$. Then, as a function of detector sample number, *k*, write

$$g[k] \equiv \Delta t \cdot g(k\Delta t) = \begin{cases} \frac{1}{4\Delta t}; \text{ k = 0} \\ 0; \text{ k even} \\ -\frac{1}{k^2\pi^2\Delta t}; \text{ k odd} \end{cases} \quad . \qquad \textit{Eq.(11)}$$

The factor Δ*t* is incorporated here for notational convenience in what follows. Following Turbell [2], we denote index arguments referencing samples with square brackets. In the z-plane defined by the scan define two-dimensional pixels (x_l, y_m), *l* and *m* indices denoting locations in a two-dimensional grid. We wish to find *f*[*l*,*m*] from Eq.(8) (see page 8):

$$f[l,m] = \frac{2\pi}{N_\theta} \sum_{i=0}^{N_\theta - 1} p[i,k] \otimes g[k] \quad , \qquad \textit{Eq.(12)}$$

where the convolution refers to a sum over a dummy index *k'*:

$$p[i,k] \otimes g[k] = \sum_{k'} p[i, k-k'] g[k'] \quad . \qquad \textit{Eq.(13)}$$

The index k must refer to the ray intersecting pixel $[l,m]$ at angle θ_i. Because only a finite number of samples over a finite range are taken, the allowed values of k' in the summation are $0 \le k - k' \le N_t$. This filtering may be performed via a fast Fourier transform (FFT) technique rather than as the direct sum, Eq.(13). Note that for most i and k the number of taps equals the number of terms in the sum, so the FFT technique is attractive. Unused elements of g for a given k are zero-padded.

Note that a ray to the center of a given detector element may not pass directly through (x_l, y_m). This is frequently addressed by interpolation between adjacent detector elements.

Collecting data over a full 2π turn means that there is redundancy between rays in opposite directions (neglecting beam-hardening). Some equipment allows for the detector to be shifted by $\Delta t/2$ during half the turn. Then rays from opposite directions are interleaved for increased resolution for the same applied dose.

This logic has followed Turbell [2]. A different discussion is given by Herman in [16], where other algorithms and references are also given. The Linogram algorithm is a fast FBP algorithm that is well-suited for certain situations. Series expansion methods assume that the absorptivity f can be expanded in terms of a series of a suitable set of basis functions. A possible choice for basis functions is the set of image pixels. These methods offer certain advantages of flexibility but Herman reports these to be computationally expensive, and the results are sensitive to the model (basis functions, a projection matrix, and estimation criteria). However, he reports that there are approaches to these difficulties that give satisfactory results without requiring complete, time-consuming convergence. The series expansion methods are iterative methods, to be discussed again below.

4.2. Fast FBP

Turbell [2] references a number of algorithms for faster computation of the filtered backprojection algorithm. These are all of the form of divide-and-conquer that has been so successful with the FFT and that is often well-suited for parallelization on parallel processing machines. Without optimization, an $N \times N$ image can be reconstructed in $O(N^3)$ computations. With optimization this can be reduced to $O(N^2 \log N)$, where the logarithm is taken to base 2.

5. Fan-Beam Tomography

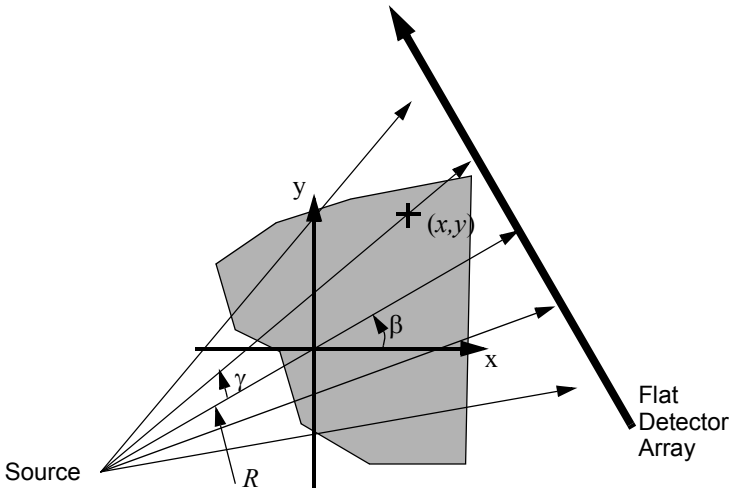
The parallel-ray apparatus has been largely replaced by the much simpler fan-beam setup, where a single source of x-rays illuminates a detector row that may be either flat (Figure 2, page 10) or curved. A few of the beams from the source are shown. The central ray is at angle β with respect to the x-axis and a specified ray through point (x,y) is at fan angle γ, where $\gamma = 0$ is the central ray. Exact fan-beam algorithms exist [2] (and references therein). An equivalent parallel projection set can be constructed from the fan-beam projections by rebinning and interpolating,

$$p(\theta, t) = p_{fan}\left(\theta - \operatorname{asin}\frac{t}{R}, \operatorname{asin}\frac{t}{R}\right) \quad , \qquad \textit{Eq.(14)}$$

where θ and t are as defined for parallel beams and R is the radius of curvature of the arc of travel of the source and detector. In this equation, p_{fan} refers to the projection of the fan beam onto a *curved* detector row. Note that

$$\beta = (\theta - \gamma); \qquad \gamma = \operatorname{asin}\frac{t}{R} \quad . \qquad \textit{Eq.(15)}$$

Redundant information is discarded and reconstruction is done using FBP as for parallel beams. With these interpolations, parallel-beam reconstruction will work. Turbell [2] reports that an advantage of rebinning to a parallel approach is increased computational simplicity.



X-ray source illuminates detector row through subject (shaded) with a fan beam. Source and detector co-rotate about origin (z-axis) at radius *R*. Location of source yields central ray at angle β. Ray through given point (*x,y*) is at angle γ with respect to central ray.

Figure 2: Fan Beam Setup

We should recall that most x-ray sources will not have a uniform angular distribution of luminance. Therefore, a correction will have to be applied, along with the other obvious calibrations of detector element response, etc.

6. Circular Cone-Beam Reconstruction

For reconstruction of a volume image a cone beam is often more efficient than the slice beam discussed above. In this case the detector array extends in two dimensions, i.e., with segmentation in the z-direction. This could be a plane or a cylinder with center of curvature at the source. A third variable is necessary to describe data from this system. This could be the z-coordinate of the detector row being examined. It is sometimes convenient to think of slices at an angle to the z-axis, defined by the plane containing the source and one entire row of detector elements at some z. Then we want to define the angle of inclination of the slice. We may define an angle, κ, that is the polar angle of the normal of the slice plane with respect to the z-axis. In the CZA slices described previously, $\kappa = 0$.

As the source/detector gantry rotates about a subject the angle κ will precess about the z axis. If the subject is on a continuously-moving platform or conveyor belt, then the scan will be helical, in which case the motion of κ is sometimes said to be nutating.

6.1. Feldkamp, Davis, and Kress (FDK) Algorithm

FBP reconstruction requires that all the slices lie in the same plane. In cone-beam systems this can be approximately achieved by interpolation between turns. That requires two complete turns so that two views at the same angle at opposite sides of the region of interest are available. Even so, FBP requires that the slice be normal to the z-axis, a condition that is satisfied only for the central portion of a cone beam. The conventional FBP approach to a cone beam with a two-dimensional detector ignores some of the extra information available in the cone beam and treats each detector row as an independent one-dimensional constant z projection, even though only the central row is actually at constant z. This sometimes leads to streaks and other image artifacts.

The FDK algorithm [7] is said to be the most commonly-used algorithm for 3-D reconstruction using cone beams. It is an approximate (non-exact) algorithm, but the approximations are reported to be small and the algorithm has the advantage that it can deal with subjects whose full longitudinal extent may not have been scanned. It attempts to deal with the true cone-beam geometry; the backprojection is in the actual cone angles. For example, in medical applications only one portion (e.g., the region of a single organ) may be scanned and not the entire patient. In this case, exact cone-beam tomography requires scan elements that may not be available.

FDK derive weighting and filtering from an assumption of a two-dimensional filtered backprojection from the source through an object at a projection angle. For other projection angles different planes are considered.

6.1.1. Fast FDK

A fast FDK algorithm is described by Turbell [2]. He reports that the situation is much more complex than for planar fast-filtered backprojection. However, he has successfully implemented an algorithm that executes in $O(N^3 \log N)$ time. Considerably more computation is required than for fast FBP with slices normal to the z-axis. Other algorithms have been proposed; a few recent approaches are given in [11].

7. Constant-Speed Helical (CSH) Cone-Beam Reconstruction

Newer systems place the patient or other object on a continuously moving platform or conveyor belt. For airport scanning, for example, it is simply not practical to start and stop the luggage every few centimeters; it must be kept continuously moving. In continuously-moving systems the gantry will trace a helix around every volume element of the object being scanned. The CT algorithm must deal with the fact that no volume element will be scanned from every angle in a plane normal to the z-axis.

7.1. Backprojection Approaches

As for CZA cone-beam reconstructions, CSH cone-beam reconstruction requires a two-dimensional detector array. The subject is irradiated by a cone beam that rotates about a common z-axis while the subject is translated longitudinally. This field is evolving rapidly and new work becomes available frequently. A number of algorithms are introduced by Wang et al. [8]. We summarize some of their conclusions.

The benefits of CSH reconstruction with a cone beam are faster data collection and more accurate reconstruction. One aspect of helical scanning is that the scan is longitudinally truncated. A condition for exact cone-beam reconstruction in traditional algorithms is that every plane that intersects the subject contains a source point. In other words, it must be possible to surround the subject completely, with at least one source point in every plane. If the opening angle of the cone is large, this might be almost satisfied, but most apparatus now in use feature small opening angles.

Newer algorithms permit exact reconstruction with longitudinally truncated data. They do require two complete turns around the helix.

Approximate algorithms are also available. Sometimes these are capable of better resolution. They require only one-dimensional data (1-D) filtering, instead of 2-D filtering. This is less sensitive to noise and background. Also, interpolation is less complex for the approximate algorithms; this has a favorable impact on image blurring. Furthermore, since less data is required, the data may be acquired in shorter time. A full two turns is not required. This enables faster reconstruction and real-time capabilities are improved. Since there is less data the computation is less problematical. Particularly for small regions of interest (ROIs), the computation is highly parallel and, therefore, amenable to parallel processor solution.

Some additional algorithms that have been proposed for three-dimensional CSH reconstruction may be found in [9]. Rodet et al. [10] recently proposed a backprojection algorithm based on a divide-and-conquer approach that may be suitable for parallelization.

7.2. Iterative Approaches

We rely on Wang et al. [8] for the following points. Full references may be found therein.

Iterative algorithms are becoming increasingly attractive because of improvements in computing power and because they are more capable of dealing with Poisson fluctuations in photon counts. Especially in medical applications, where the dose must be kept as low as possible, statistical fluctuations can be important. In this case, the reconstruction can be viewed as a parameter estimation problem. Maximum likelihood algorithms have proven successful, especially in PET and SPECT imaging. Weighted least-squares algorithms are also available.

Expectation maximization (EM) algorithms have been shown to be successful in PET and SPECT problems and have also been adapted to x-ray tomography. This approach has shown promise for noise reduction and deblurring, as well as reduction of the streaking sometimes observed near high-Z objects, for example from metal objects in the subject being studied.

Parallel processing engines have been proven successful at dealing with iterative approaches.

8. Utilization of the FastMATH Processor for CT

Intrinsity, Inc. produces a high data-rate, high clock-speed matrix engine, the FastMATH processor [17, 18]. The first product delivers 32 giga multiply-accumulate operations (GMACs) per second, sustained. The two bi-directional RapidIO interfaces together provide 4 Gbytes per second total bandwidth to bring the required large amounts of data into the 1-Mbyte L2 cache. The FastMATH engine contains a 4 × 4 array of interconnected processing elements designed for matrix and vector mathematics and for use in SIMD mode. It is programmable in high-level languages, including C. The instruction set is the industry-standard MIPS32 instruction set architecture with coprocessor 2 extensions for the matrix unit. This gives the designer the flexibility to revise and improve algorithms without changing the hardware.

The FastMATH processor has been benchmarked on a core portion of the FBP algorithm in an x-ray tomography application. In this code, a grid of 256 × 256 picture elements (space points in one CZA slice) is reconstructed from a summation over data obtained from 256 equally-spaced detector elements. These detector elements view the object over 180 degrees of rotation, with one x-ray exposure per degree of rotation. Because the spatial grid points do not in general project into the centers of the detector elements at any view angle, a linear interpolation between the two nearest-neighbor projected detector elements is done for each view for each picture element. Thus, a total of 180 × 256 × 256 interpolations (11.8 million total) is required, followed by summation over the 180 view angles for each picture element.

The FastMATH processor can efficiently compute these in its 4 × 4 array of processing elements in groups of 16 at a time, SIMD fashion, at its 2 GHz cycle rate. The entire task described above can be performed in under 25 ms. The data can be pre-fetched into the unusually large L2 cache via the RapidIO ports. Comparable systems with other technologies require multiple processors and/or expensive and inflexible hardware solutions in order to keep up with the input data rate. The FastMATH speed and programmable capability therefore allow the designer to achieve finer resolution, shorter time intervals between slices, or to include multiple functions in the same processor, such as control, data acquisition, detector calibration, or other tasks.

9. Summary

Computed tomography has expanded beyond its initial uses in slow-speed medical imaging. With advances in radionuclide availability a number of different gamma-emitting chemicals can be constructed, specialized to image different organs of the body. With advances in detector and data acquisition technology high-speed, high-resolution imaging has become almost routine. Real-time medical imaging is now possible. With advances in x-ray tomographic equipment, including gantries that do not drag wires and are, therefore, capable of full rotations, full scans can be made for high resolution. Scans of patients or other subjects can be made with reduced dose and improved scan rates. With advances in algorithms it is practical to extract large amounts of information from the data.

Algorithms must be tailored for the data and for the processing engines available. An increase in processing power can make a more complex algorithm, or higher data rates, capable of delivering higher quality results practical. Most tomography applications deal with arrays of data from the detectors. This data may be collected at high rates, espe-

cially in real-time situations. Therefore, processors optimized for vector or matrix processing may be key to achieving sufficient performance. For example, algorithms for most technologies, such as MRI, rely heavily on FFTs or other divide-and-conquer strategies. These are usually accessible to attack by single instruction, multiple data (SIMD) engines, which have proven capable of high performance on FFTs.

Algorithm development continues at a fast pace. It is unlikely that the optimum algorithm for a given technology is yet available, and processors such as FastMATH with improved capability and flexibility open new doors.

We demonstrate the Intrinsity FastMATH matrix processor with a key FBP benchmark. To do 11.8 million data interpolations and projections for a CZA slice of a 256 × 256 point x-ray tomographic image, with 180 projections, requires 25 ms on the 2-GHz FastMATH processor. We are able to take advantage of the FastMATH SIMD capability to process 16 interpolations at once and thus benefit from both the fast cycle speed and the matrix architecture.

10. Acronyms

CAT	computerized axial tomography (also known as CT)
CSH	constant speed helical scanning
CT	computerized tomography
CZA	constant z-axis scanning
EIT	electrical impedance tomography
EM	expectation maximization
FBP	filtered backprojection
FDK	Feldkamp, Davis and Kress algorithm
FFT	fast Fourier transform
GMAC	giga multiply-accumulate operations
MIT	magnetic induction tomography
MRI	magnetic resonance imaging
PET	positron emission tomography
ROI	region of interest
SIMD	single instruction, multiple data
SPECT	single-photon emission tomography
UV	ultraviolet light

11. References

[1] Some technologies are reviewed by J. A. Rowlands, “Current Advances and Future Trends in X-Ray Digital Detectors for Medical Applications,” *IEEE Trans. Instr. and Meas.*, **47**, p.1415-1418 (1998); see also J. A. Rowlands, “The Physics of Computed Radiography,” *Phys. Med. Biol.*, **47**, p.R123-R166 (2002).

[2] Henrik Turbell, *Cone-Beam Reconstruction using Filtered Backprojection*, Ph. D. thesis, Institute of Technology, Linköpings Universitet, Linköping, Sweden (2001).

[3] Robert E. Alvarez and Albert Macovski, “Energy-Selective Reconstructions in X-ray Computed Tomography,” *Phys. Med. Biol.* **21**, p.733-744 (1976).

[4] Pratondo Busono and Esam M. A. Hussein, “Algorithms for Density and Composition-Discrimination Imaging for Fourth-Generation CT Systems,” *Phys. Med. Biol.* **44**, p.1455-1477 (1999).

[5] Particle Data Group, “Review of Particle Properties,” *The European Physical Journal C***15**, p.1-878 (2000), and other (biennial) editions; considerably more detail and tabulated values are available at http://physics.nist.gov/PhysRefData.

[6] B. Aschenbach, *Rep. Prog. Phys.*, **48**, p.579-629 (1985).

[7] L. A. Feldkamp, et al., “Practical Cone-Beam Algorithm,” *J. Opt. Soc. Am.*, **1**, p.612-619 (1984).

[8] G. Wang, C. R. Crawford, and W. A. Kalender, “Guest Editorial Multirow Detector and Cone-Beam Spiral/Helical CT,” *IEEE Trans. Med. Imaging*, **19**, p.817-821 (2000) introduce a number of original papers in the same journal summarizing developments in helical cone-beam reconstruction.

[9] H. Kudo and T. Saito, "Three-Dimensional Helical-Scan Computed Tomography Using Cone-Beam Projections," *Journal of the Electronics, Information, and Communications Society*, **J74-D-II**, p. 1108-1114 (1991); D. X. Yan and R. Leahy, "Cone-Beam Tomography with Circular, Elliptical, and Spiral Orbits," *Phys. Med. Biol.*, **37**, p.493-506 (1992); S. Schaller, T. Flohr, and P. Steffen, "New Efficient Fourier Reconstruction Method for Approximate Image Reconstruction in Spiral Cone-Beam CT at Small Cone Angles," *SPIE International Symposium on Medical Imaging*, (Feb., 1997); G. Wang, T-H Lin, P. Cheng, and D. M. Shinozaki, "A General Cone Beam Algorithm," *IEEE Trans. Med. Imag.*, **12**, p.486-496 (1993).

[10] T. Rodet, P. Grangeat, and L. Desbat, "Multichannel Algorithm for Fast 3D Reconstruction," *Phys. Med. Biol.*, **47**, p.2659-2671 (2001).

[11] K. Taguchi, G. L. Zeng, and G. T. Gullberg, "Cone-Beam Image Reconstruction Using Spherical Harmonics," *Phys. Med. Biol.,* **46**, p.N127-N138 (2001).

[12] W. A. Kalendar, *Computed Tomography: Fundamentals, System Technology, Image Quality and Applications*, (Publicis MCD, Munich 2000).

[13] V. Cherepenin et al., "A 3D Electrical Impedance Tomography (EIT) System for Breast Cancer Detection," *Physiol. Meas.*, **22**, p.9-18 (2001).

[14] A. Korjenevsky, V. Cherepenin, and S. Sapetsky, "Magnetic Induction Tomography: Experimental Realization," *Physiol. Meas.*, 21, p.89-94 (2000).

[15] S. Stergiopolous, ed., *Advanced Signal Processing Handbook* (CRC Press, New York, 2000).

[16] G. T. Herman, "Algorithms for Computed Tomography," in V. K. Madisetti and D. B. Williams, eds., *The Digital Signal Processing Handbook* (CRC Press, New York 1998).

[17] V. Anantha, G. Crabtree, C. Harlé, Tim Olson, and G. Yost, "An Innovative Processor for Parallel Processing Applications: FFT Implementation Results," presented to International Signal Processing Conference and Global DSP Expo, Dallas, TX (2003).

[18] T. Olson, "Advanced Processing Techniques Using the Intrinsity FastMATH Processor," presented to Embedded Processor Forum, San Jose, CA (2002); also available on www.intrinsity.com.

12. Revision History

Revision	Date	Description
1.0	11/25/02	Initial Issue
1.1	1/17/03	Incorporate first benchmark results

Copyright © 2002-2003 Intrinsity, Inc. All Rights Reserved. Intrinsity, the Intrinsity logo, "the Faster processor company", and FastMATH are trademarks/registered trademarks Intrinsity, Inc. MIPS32 is a trademark of MIPS Technologies, Inc. RapidIO is a trademark of the RapidIO Trade Association. All other trademarks are the property of their respective owners.

INTRINSITY, INC. MAKES NO REPRESENTATIONS OR WARRANTIES, EXPRESS OR IMPLIED, AS TO THE ACCURACY OF COMPLETENESS OF INFORMATION CONTAINED IN THIS DOCUMENT. INTRINSITY RESERVES THE RIGHT TO MODIFY THE INFORMATION PROVIDED HEREIN OR THE PRODUCT DESCRIBED HEREIN OR TO HALT DEVELOPMENT OF SUCH PRODUCT WITHOUT NOTICE. RECIPIENT IS ADVISED TO CONTACT INTRINSITY, INC. REGARDING THE FINAL SPECIFICATIONS FOR THE PRODUCT DESCRIBED HEREIN BEFORE MAKING ANY EXPENDITURE IN RELIANCE ON THE INFORMATION CONTAINED IN THIS DOCUMENT.

No express or implied licenses are granted hereunder for the design, manufacture or dissemination of any information or technology described herein or the use of any trademarks used herein.

Any and all information, including technical data, computer software, documentation or other commercial materials contained in or delivered in conjunction with this document (collectively, "Technical Data") were developed exclusively at private expense, and such Technical Data is made up entirely of commercial items and/or commercial computer software. Any and all Technical Data that may be delivered to the United States Government or any governmental agency or political subdivision of the United States Government (the "Government") are delivered with restricted rights in accordance with Subpart 12.2 of the Federal Acquisition Regulation and Parts 227 and 252 of the Defense Federal Acquisition Regulation Supplement. The use of Technical Data is restricted in accordance with the terms set forth herein and the terms of any license agreement(s) and/or contract terms and conditions covering information

containing Technical Data received between Intrinsity, Inc. or any third party and the Government, and the Government is granted no rights in the Technical Data except as may be provided expressly in such documents.