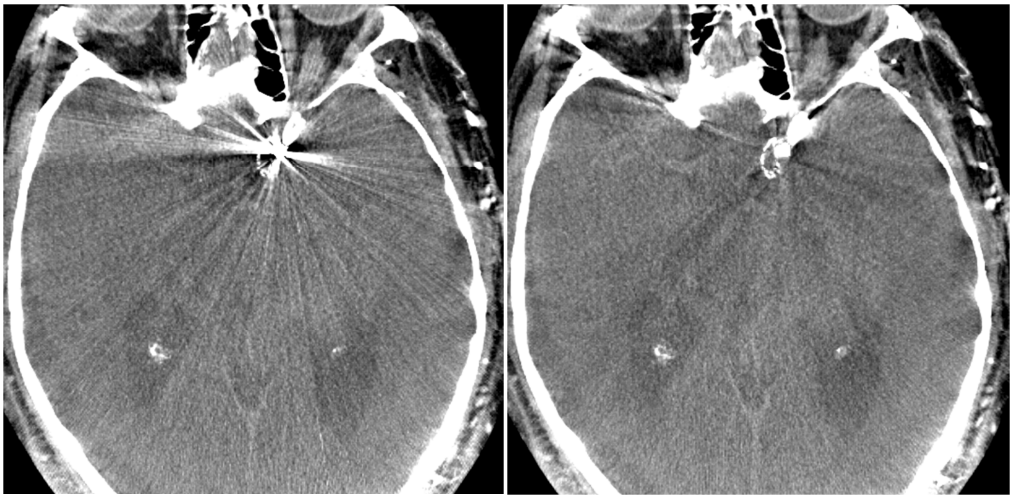
METAL ARTIFACT REMOVAL IN C-ARM CONE-BEAM CT

Group 4
Carolina Cay-Martinez, Marta Wells

Report #2
Technical Background



CT image of coil before and after MAR algorithm application
Image provided by Radvany, MD

The Johns Hopkins University
Advanced Computer Integrated Surgery

Project Advisors:
Jeffrey H. Siewerdsen, Ph.D.
Martin Radvany, MD (Interventional Radiology)
Tina Ehtiati, Ph.D. (Siemens Healthcare)

The primary purpose of medical imaging systems is to create images of the internal structure and function of the body for diagnostic purposes or interventional treatment of diseases. The ability of medical professionals to successfully accomplish these tasks strongly depends on the quality of the images, i.e. the degree to which the image successfully represents the anatomy. Degradations, distortions or artifacts introduced by the imaging system hinder the correct representation of the true objects. Fundamental knowledge of imaging principles, image quality and factors that introduce image degradation provides tools for optimization of imaging systems.

I. Medical Imaging: CT Imaging Principles

Electromagnetic Radiation: Strength, Attenuation and Dose

The simplest measure of the strength of a burst of electromagnetic radiation is simply the number of photons N in the burst. However, it is also important to measure the area over which these photons are spread; thus, a more accurate measure of the beam strength is the photon fluence Φ as the number of photons N per unit area A,

$$\Phi = N/A$$

where the area is oriented at right angle to the direction of the radiation beam propagation. It is also important to account for time, since measurements may take place over a fixed interval Δt. Accordingly, photon fluence rate is defined as

$$\varphi = \frac{N}{A\Delta t}$$

In addition, it is useful to be able to determine the energy of the burst and how much of that energy could be deposited into a material if completely absorbed. If the beam is monoenergetic, i.e. each photon in the burst carries the same energy hν, then the total energy of the burst is Nhν. The energy fluence and energy fluence rate can be defined as

$$\Psi = \frac{N\,h\nu}{A}$$

$$\psi = \frac{N\,h\nu}{A\Delta t}$$

The energy fluence rate is also called the intensity, I, of an electromagnetic radiation beam. Thus,

$$I = E\varphi$$

where E = hν. Intensity has units of energy per unit area per unit time.

Attenuation occurs when photons are absorbed within material by the photoelectric effect and when photons are deflected away from the detector by Compton events; the detector count N' of photons will be smaller than N, the original number of photons in the burst. If n photons are "lost" is each burst, we would expect that both doubling the width of the sample material, Δx, or doubling N should also double n. Thus, n is proportional to both N and Δx, or

$$n = \mu N \Delta x$$

where μ is a constant of proportionality called the linear attenuation coefficient, defined as

$$\mu = (n/N)/\Delta x$$

which is the fraction of photons that are lost per unit length.

The change in the number of photons upon interaction with the material is

$$\Delta N = N' - N$$
$$= -n$$
$$= -\mu N \Delta x$$

Letting the material become infinitely thin and treating N as a continuous quantity leads to the differential equation

$$dN/N = -\mu\, dx,$$

which can be integrated to yield

$$N = N_0 e^{-\mu \Delta x},$$

where N_0 is the number of photons at x = 0. This is called the fundamental photon attenuation law and can be used to compute the fraction of photons that will be stopped or transmitted by a noninfinitesimal layer of homogeneous material. The thickness of a given material that will attenuate half of the incident photons is known as the half value layer (HVL) is

$$N/N_0 = \tfrac{1}{2} = e^{-\mu HVL},$$

or

$$HVL = \ln 2/\mu.$$

In general, the linear attenuation coefficient is different for different materials as well as varying as a function of energy for the same material. Electromagnetic radiation experiences less attenuation at higher energies, i.e. it is more penetrating at higher energies. For a polyenergetic beam, we treat the linear attenuation coefficient of a given material as a function of E – μ(E).

A homogeneous slab of material of thickness Δx has an incident beam of spectrum $S_0(E)$. The spectrum leaving the material is attenuated in an identical way to that of the monoenergetic case, except that the linear attenuation coefficient depends on E,

$$S(E) = S_0(E) e^{-\mu(E)\Delta x}$$

For a heterogeneous slab of material, the linear attenuation coefficient is dependent on both the position along the line as well as energy. As such, the integral form of the attenuation law obeys

$$S(x; E) = S_0(E')\, e^{\int_0^{\infty} -\mu(x';E')dx}$$

To calculate the overall of the beam,

$$I = \int_0^{\infty} S_0(E')E'\, e^{-\mu(E')\Delta x}\, dE'$$

and

$$I = \int_0^{\infty} S_0(E')E'\, e^{\int_0^{\infty} -\mu(x';E')dx}\, dE'$$

However, while this equation is a valid model for heterogeneous materials, it is not actually useful in developing imaging equations. The energy dependence of μ, while useful in developing an understanding of the basic imaging properties, is intractable from a mathematical standpoint.

Radiation dosage is primarily concerned with what electromagnetic radiation does than what it is, i.e. the biological effects of the radiation. By definition, ionizing radiation ionizes hydrogen atoms and produces ions in the air. Exposure, X, refers to the number of ion pairs produced in a specific volume of air by electromagnetic radiation. The unit of measure of exposure is coulombs per kilogram of air (C/kg); however, a more useful unit in medical imaging is the roentgen (R), which equals 2.58×10^{-4} C/kg, or 1 C/kg = 3876 R.

As ionizing electromagnetic radiation passes through a material, it deposits energy into the material by both the photoelectric effect and Compton scattering. This is the fundamental concept of dose. The unit of absorbed dose, D, is the rad, which is defined as the absorption of 100 ergs per gram of material, and refers to an energy-deposition concentration rather than a total amount of energy. The SI unit for absorbed dose is the gray (Gy); 1 Gy = 1 J/kg = 100 rads. For the degree of accuracy required in biomedical dosimetry, one roentgen of exposure yields one rad of absorbed dose in soft tissue.

The linear energy transfer (LET) is a measure of the energy transferred by radiation to the material through which it is passing per unit length; higher LET radiation tends to produce greater adverse biological consequences. Specific ionization (SI) is the number of ion pairs formed per unit length. LET and SI are

related to each other by the average amount of energy required to form on ion pair (often referred to as W), which is a characteristic of the material.

Often it can be useful to measure exposure but express the results as dose to an individual in that radiation field. By the definitions of the roentgen and the rad, 1 R = 0.87 rad, and to compute the dose to a material other than air, the f-factor is used as a conversion,

$$D = fX.$$

The f-factor is defined as

$$f = 0.87 \frac{(\mu/\rho)material}{(\mu/\rho)air},$$

where μ is the linear attenuation coefficient and ρ is the mass density of the material and air. The quantity (μ/ρ) is called the mass attenuation coefficient.

Filtration, Restriction and Contrast Agents

Filtration and restriction both deal with modifications to the x-ray beam that take place before the x-rays enter the body. Filtration is the process of absorbing low-energy x-ray photons before they enter the patient, while restriction is the process of absorbing the x-rays outside a certain field of view. X-ray photons emitted from an x-ray tube have a distribution of energies and x-ray sources used in medical imaging systems are polyenergetic; in practice, it is very undesirable for low-energy photons to enter the body, as they are almost entirely absorbed within the body, contributing to patient radiation dose but not to the image. Additionally, body tissues attenuate x-rays in different amount depending on their linear attenuation coefficients and the x-ray energies. It is this differential attenuation that gives rise to contrast and the ability to differentiate tissues in an x-ray image. In some circumstances, we may want to artificially change the natural attenuation of the body prior to detecting the x-rays. Thick body parts or dense materials stop more x-rays than thinner body parts and soft tissues; imaging both locations within the same exposure is difficult due to the limited dynamic range of the x-ray detectors. A compensation filter, comprising a specially shaped aluminum or plastic object, is placed between the x-ray source and the patient, or between the patient and the detector. This filter is thicker where the body part is thinner and vice versa, so that the x-ray detector requires a smaller dynamic range.

In addition, x-ray tubes generate x-rays in all directions. By design, the x-rays that exit through the tube window form a cone that is ordinarily much larger than the desired body region to be imaged. The exiting beam must be further restricted both to avoid exposing parts of the patient that do not need to be imaged and to help reduce the adverse effects of Compton scattering. The three basic types of beam restrictors are diaphragms, cones or cylinders, and collimators.

Finally, differential attenuation in the body gives rise to contrast in the x-ray image; however, often different soft tissue structures are difficult to visualize due to insufficient intrinsic contrast. This can be improved by using contrast agents, chemical compounds that are introduced into the body in order to increase x-ray absorption within the anatomical regions into which they are introduced, increasing x-ray contrast.

CT Parameters

The main variable parameters that contribute to radiation dose are tube current (mA), peak kilovoltage (kVp), pitch, collimated detector row width, z-coverage, and gantry cycle time in seconds. The number of electrons flowing from the negative to the positive electrode in an x-ray tube is the x-ray tube current. A range of tube currents are available on clinical CT systems the specific range is variable but is typically within 50 to 300 mA. The relationship between tube current and radiation dose is linear: decreasing tube current by half will also decrease radiation dose by half. An increase in x-ray tube current increases the number of x-ray photons produced per unit time; therefore the selection of a higher x-ray tube current increases the number of photons penetrating the patient and reaching the detectors, thereby decreasing image noise. However, this will result in increased radiation exposure to the patient; thus, the lowest x-ray tube current that meets the image noise requirements of the clinical application is desired.

In contrast to tube current, the relationship between peak kilovoltage and dose in nonlinear. Tube voltage is the potential difference maintained between the anode and the cathode in an x-ray tube. It provides the energy to the electrons ejected by the x-ray filament to accelerate towards the anode. When electrons accelerating as the result of the potential difference initiated by the kilovoltage peak (kVp) hit the anode, Bremsstrahlung x-rays are produced of a range of energies. The average energy of x-rays coming out of the x-ray tube is typically one third of the peak tube voltage (kVp). In CT, the x-ray tube voltage ranges from 100 kVp to 140 kVp, with 120 kVp being the most commonly used tube voltage. Of the tube voltages greater than 100 kVp, 120 kVp typically is the chosen voltage in the majority of CT protocols because it yields x-ray energy sufficient to penetrate most parts of the anatomy. The larger the patient, the greater the tube voltage required to penetrate the thickest portion of the anatomy.

Pitch is defined as the table distance traveled in one complete rotation divided by the total collimated width of the x-ray beam. It is linearly inversely proportional to patient dose: larger pitches lower the radiation dose. A pitch of 1 indicates acquisition of overlapping data slabs while a pitch greater than 1 indicates gaps in the acquisition of successive data slabs. Lower pitch values are typically necessary to ensure continuity in z-coverage between reconstructed image sets; however, lower pitch values result in increased radiation exposure and increased total scan time.

The collimated width of one detector row determines the minimum reconstruction slice thickness. State-of-the-art CT systems have 0.5- or 0.625 mm wide detector rows at the center of rotation. Although smaller collimated slice widths offer improved z-axis spatial resolution, fewer x-ray photons contribute to the reconstructed image, resulting in increased image noise. This increased noise can be overcome by increasing the x-ray tube voltage or the tube current.

In addition to the width of the detector rows, the number of detector rows is critical for data acquisition. Increased z-coverage per rotation and decreased scan time are associated with a greater number of detector rows in a multidetector array. However, the number of detector rows does not necessarily equal the number of slices that can be acquired per rotation; some scanners sample each detector twice per rotation such that the number of slices acquired per rotation is two times the number of detector rows used. The z-coverage in state-of-the-art scanners ranges from approximately 2 to 16 cm per rotation.

Finally, decreasing gantry rotation time decreases radiation dose linearly: the faster the gantry rotation, the lower the dose. The minimum gantry rotation time depends on the make and model of the CT scanner; values for state-of-the-art scanners range from 270 to 350 ms. A faster gantry rotation time, however, decreases the number of x-ray photons produced and, subsequently, the number of photons interacting with the body, resulting in increased image noise. Therefore, an increase in radiation exposure is required with improvements in temporal resolution to maintain image noise.

CT Measurements

In a CT system, the x-ray tube makes a short burst of x-rays that propagate through a cross-section of the patient. The detectors measure the exit beam intensity integrated along a line between the x-ray source and each detector. However, the integration over energy is in a mathematically intractable form for CT image reconstruction. To correct this, we use the concept of effective energy, E, which is defined as the energy that, in a given material, will produce the same measured intensity from a monoenergetic source as is measured using the actual polyenergetic source. Thus,

$$I_d = I_0 \, e^{\int_0^d \mu(s;\, E)\, ds}$$

Given a measurement of I_d and knowledge of I_0, this equation can be rearranged to yield a basic projection measurement g_d,

$$g_d = -\ln(I_d/I_0),$$

$$= \int_0^d \mu(s;\, E)\, ds$$

Thus, the most basic measurement of a CT scanner is a line integral of the linear attenuation coefficient at the effective energy of the scanner. In an actual CT system, the reference intensity I_0 must be measured for each detector as calibration.

A CT scanner reconstructs the value of μ at each pixel within a cross-section. However, different CT scanners have different x-ray tubes, which in turn have different effective energies. Thus, the same object

may produce different numerical values of μ on different scanners. In order to compare data from different scanners, CT numbers are computed from the measured linear attenuation coefficients at each pixel. The CT number is defined as

$$\text{HU} = 1000 \cdot \frac{\mu - \mu_{water}}{\mu_{water}}$$

in Hounsfield units (HU). h = 0 HU for water and since μ = 0 in air, h = -1000 HU for air. The largest CT numbers typically found naturally in the body are in bone where h ≈ 1000 HU on average, although CT numbers can surpass h ≈ 3000 HU for metal and contrast agents. Usually, CT numbers are rounded or truncated to the nearest integer; they are typically reproducible within 2 HU between scans and across scanners.

The most basic CT measurement is the line integral of the effective linear attenuation coefficient within a cross-section; however, for imaging purposes we need a picture of μ, or equivalently the CT number h, over the entire cross-section. The parametric form of the line integral of the function f(x,y) is given by

$$g(l,\theta) = \iint_{-\infty}^{\infty} f(x,y)\delta(xcos\theta + ysin\theta - l)dxdy$$

For a fixed θ, g(l,θ) is called a projection; for all l and θ, g(l,θ) is called the 2-D Radon transform of f(x,y). Given that

$$f(x,y) = \mu(x,y;E')$$
$$g(l,\theta) = -\ln(\frac{I_d}{I_0})$$

then this mathematical equation exactly characterizes the CT measurements.

An image of g(l,θ) with l and θ as rectilinear coordinates is called a sinogram. It is a pictorial representation of the Radon transform of f(x,y) and represents the data that are necessary to reconstruct f(x,y); the Radon transform of points in the image correspond to sinusoids in the sinogram. Data represented in the sinogram is then processed to create the cross-sectional image through a procedure known as reconstruction.

Reconstruction: Backprojection, Summation, Filtering, and Convolution

The simplest method for reconstructing an image from a projection along an angle is through backprojection, in which we assign every point on the line, $L(l_0,\theta_0)$ the value $g(l_0,\theta_0)$. When this process is repeated for all l for the given θ, the resulting function is called a backprojection image. To incorporate information about the projections at other angles, we can integrate their backprojection images, yielding the backprojection summation image:

$$f_b(x,y) = \int_0^{\pi} b_\theta(x,y)d\theta$$

also known as a laminogram.

The projection-slice theorem is the basis of three important reconstruction methods: the Fourier method, filtered backprojection, and convolution backprojection. The relationship states that the 1-D Fourier transform of a projection is a slice of the 2-D Fourier transform of the object, i.e. that the 1-D Fourier transform of the projection equals a line passing through the origin of the 2-D Fourier transform of the object at that angle corresponding to the projection.

The Fourier method involves taking the 1-D Fourier transform of each projection, inserting it with the corresponding correct angular orientation into the correct slice of the 2-D Fourier transform plane, and taking the inverse 2-D Fourier transform of the result. However, the Fourier method is not commonly used in CT due to the problem of interpolating polar data onto a Cartesian grid, and the need to utilize time-consuming 2-D inverse Fourier transforms. Further analytic manipulation leads to the improved method of filtered backprojection, written as

$$f(x,y) = \int_0^{\pi} [\int_{-\infty}^{\infty} |\varsigma| G(\varsigma,\theta) e^{j2\pi\varsigma l} d\varsigma]_{l=xcos\theta+ysin\theta} d\theta$$

where ς denotes spatial frequency. The inner integral is an inverse 1-D Fourier transform so the Fourier transform of the projection g(l,θ) is multiplied by the frequency filter |ς| and inverse-transformed. After the inverse transform, the filtered projection is backprojected, which is followed by a summation of all filter

projections. The straight application of the Fourier method leads to sampling that is inversely proportional to ς; however, additional area is required to compensate for the sparser sampling at higher frequencies.

From the convolution theorem of Fourier transforms, the equation for filtered backprojection, as seen above, can be rewritten as

$$f(x,y) = \int_0^{\pi} [\mathrm{F}_{1D}^{-1}\{|\varsigma|\} * g(l,\theta)]_{l=x\cos\theta+y\sin\theta} d\theta$$

Defining $c(l) = \mathrm{F}_{1D}^{-1}\{|\varsigma|\}$ results in

$$f(x,y) = \int_0^{\pi} [c(l) * g(l,\theta)]_{l=x\cos\theta+y\sin\theta} d\theta$$

$$f(x,y) = \int_0^{\pi} \int_{-\infty}^{\infty} g(l,\theta)\, c(x\cos\theta + y\sin\theta - l) dl\, d\theta$$

where the above equation is the mathematical representation for convolution backprojection and results from substituting in the convolution integral. Both filtered and convolution backprojection reconstruct images from a sinogram using three basic steps: filtering or convolution, backprojection and summation. Convolution backprojection is generally more efficient than filtered backprojection if the impulse response is narrow, and generally most CT scanners perform a form of convolution rather than filtered backprojection.

II. Image Quality

CT imaging systems introduce some artifacts or distortion throughout the acquisition of the signal and the reconstruction of the image. These distortions may decrease the perceived quality of the final reconstructed image; therefore, quality assessment is an important issue to address. Image quality is a measure of the perceived degradation of an acquired image. Following are quantitative parameters critical for image quality assessment.

Contrast

Contrast refers to the difference between the image intensity of an object and its surrounding objects or background. The use of a periodic signal and its modulation is an effective way to quantify contrast. The *modulation* of a signal, with f_{max} maximum and f_{min} minimum values, is defined by

$$m_f = \frac{f_{max} - f_{min}}{f_{max} + f_{min}}$$

Modulation quantifies the relative amount by which the amplitude of f(x,y) stands out from the average value. If $m_f = 0$, we say it has no contrast. The *modulation transfer function* (MTF) is the ratio of the output modulation to the input modulation as a function of spatial frequency.

$$\mathrm{MTF}(u) = \frac{m_g}{m_f} = \frac{|\mathrm{H}(u,0)|}{\mathrm{H}(0,0)}$$

This shows the MTF of an imaging system is, in essence, the "frequency response" of the system, and it can be directly obtained from the Fourier transform of the response of the system to a point source (point impulse); $\mathrm{H}(u,v) = \boldsymbol{\mathcal{F}_{2D}}\{h(x,y)\}$. The MTF quantifies degradation of contrast as a function of spatial frequency, taking values from 0 to 1. If the image reproduced the object perfectly, the MTF would have a value of 1. If the image reproduced contained no information from the object, the MTF would be 0. For most medical imaging systems, the MTF becomes significantly less than unity (or even zero) at high spatial frequencies. The lowest MTF that will result in a visible CT image happens at about 0.1 MTF. An MTF curve (plotted MTF vs. spatial frequency) that extends further to the right indicates higher spatial resolution.

It is common on CT imaging modality to consider an object of interest (e.g. tumor) that is referred to as the target. If the target has a nominal intensity of f_t and the background has the nominal intensity of f_b then the difference between the intensities of the target and the background is captured by *local contrast*, defined as

$$c = \frac{f_t - f_b}{f_b}$$

Note that the definition of local contrast differs from the definition of modulation in that the intensities f_t and f_b may be selected locally (e.g. within the liver) and they don't need to be the maximum and minimum of the entire image.

Resolution

In general terms, resolution is the ability of a medical imaging system to accurately depict two distinct events in time, space or frequency as separate. Therefore, we can talk about *spatial, temporal or spectral resolution*. Resolution can also be thought as the degree of smearing or blurring a medical imaging system introduces during the acquisition and reconstruction of data.

Spatial Resolution

Spatial resolution is the ability to resolve, as separate forms, small objects that are very close together. Spatial resolution can be measured using two methods: directly or by analyzing the spread of information within the system by using the modulation transfer function.

The most common method for measuring spatial resolution is by using the MTF. Suppose that a line impulse, passing through the origin of the spatial domain, is imaged through the system and the line impulse is represented by $f(x,y) = \delta(x,y)$. Since the system is isotropic, it is sufficient to consider the response to a vertical line through the origin, i.e., $f(x,y) = \delta(x)$. The output of the system (the resulting image) $g(x,y)$, is only a function of x and it is known as the *line spread function* (LSF). The LSF can be used to calculate the 1-D Fourier transform $L(u)$, and $H(u,0) = L(u)$. Given the LSF, its resolution can be quantified using a measure called the *full width at half maximum* (FWHM). This is the (full) width of the LSF at one-half its maximum value. Recall that the MTF takes values from 0 to 1. If the image reproduced the object exactly, the MTF would have a value of 1. If the image reproduced contained no information of the object, the MTF would be 0. As the size of the object increases, the MTF also increases. An MTF curve (plotted MTF vs. spatial frequency) that extends further to the right indicates higher spatial resolution. The lowest MTF that will result in a visible CT image happens at about 0.1 MTF, referred to as the *limiting resolution.*

Spatial resolution can also be measured directly through the use of resolution tools, or bar phantoms. This type of phantom is made of acrylic and has closely spaced metal strips imbedded in it. The phantom is scanned, and the number of strips that are visible are counted. A line pair is *not* a set of two lines, but rather a line and the space between lines. This is the conventional definition because it would be impossible to differentiate neighboring strips if there was no space between the lines. With this in mind, if 20 lines can be seen, the spatial resolution can be reported as 20 line pairs per centimeter (lp/cm).

There are a variety of interrelated factors that affects the degree of spatial resolution in a CT image. These factors are matrix size, pixel size, field of view (FOV), voxel size, slice thickness, focal spot size, and blur.

Matrix size and pixel size: To create an image, the system must segment raw data into tiny sections. A matrix is a grid that is used to break the data into columns and rows of tiny squares. Each square is a *picture element,* more commonly referred to as a pixel. The pixel value represents the proportional amount of x-ray energy that passes through anatomy and strikes the detector. The information contained in each pixel is averaged so that one Hounsfield unit is assigned to each pixel. If an object is smaller than a pixel, its density will be averaged with the information in the remainder of the pixel. This phenomenon is referred to as the *partial volume effect* or *volume averaging*. It results in a less accurate image. Because no object smaller than a pixel can be accurately displayed due to volume averaging (and the matrix size influences the size of the pixel), it follows that matrix size affects spatial resolution.

Field of View: Selecting the FOV determines how much of the total raw data available will be used to create an image. Decreasing the field of view will decrease the pixel size. The pixel size increases as more patient information is crammed into each pixel. This causes the spatial resolution to decrease.

Voxel size: A voxel (volume element) represents a volume of patient data. Voxel size also plays a role in volume averaging. As stated earlier, in order to create an image, the system must break up the patient data into segments. We have seen how a matrix is used to divide the data into pixels with X and Y dimensions. This allows the system to create a two-dimensional image. However, it is important to keep in mind that a three-dimensional object is being represented. By accounting for the slice thickness, the voxel represents a volume of patient data. Decreasing the slice thickness affects the resolution in two ways. First, it reduces the amount

of tissue averaged together. Second, it will increase the image noise if technique is not increased to compensate for photon absorption from increased collimation.

Blur: In considering the resolution of a system, we must consider an aspect known as *sharpness*. Sharpness is the ability of a system to define an edge. It is measured by the amount of blur in a system.

Temporal and Spectral Resolution

Temporal resolution is the ability to distinguish two events in time as being separate. *Spectral resolution* is the ability to distinguish two different frequencies (or, equivalently, energies). Conceptually, a frequency distribution histogram of the number of observed events as a function of time or energy from a single-time or single-energy process can be created. Therefore, the actual FWHM (in time or energy) would quantify the resolution.

Noise

Noise is a generic term that refers to any type of random fluctuation in an image, and it can have dramatic impact on image quality. The source and amount of noise depend on the physics and instrumentation of the particular imaging modality and the particular medical imaging system at hand. In projection radiography, x-rays arrive at the detector in discrete packets of energy, called *quanta* or *photons*. The discrete nature of their arrival leads to random fluctuations, called *quantum mottle*, which gives and x-ray image a textured or grainy appearance.

Usually, noise in imaging systems result from a summation of a large number of independent noise sources. According to the *central limit theorem* of probability, a random variable that is the sum of a large number of independent causes tends to be Gaussian. Therefore, it is natural to model noise in medical imaging systems by means of a Gaussian random variable. The *probability density function* of a Gaussian random variable is given by,

$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2}$, where μ is the mean(expected value) and σ^2 is the variance

More specifically, noise can also be characterized by discrete random variables given that photons are discrete particles. In x-ray imaging, noise is characterized by Poisson random variables, which are used to statistically characterize the distribution of photons counted per unit area by an x-ray image intensifier (i.e. to model the number of photons that arrive at the detector). The *probability mass function* of a Poisson random variable is given by,

$P(k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}$, where λ is the $arrival\ rate$ of the photons

A useful way to quantify noise degradation is by means of the *signal to noise ratio* (SNR). The SNR describes the relative strength of a signal *f* with respect to the noise, *n*. Higher SNR values indicate low noise, while lower SNR values indicate higher noise degradation. Most frequently, SNR is expressed as the ratio of signal amplitude to noise amplitude: SNR = Amplitude(*f*)/Amplitude(*n*). In x-ray imaging systems, photon count follows a Poisson distribution. Consider the signal *f* to be the average photon count per unit area and the noise *n* to be the random variation of this count around the mean, with amplitude quantified by the standard deviation of the number of photons. Therefore, the *intrinsic SNR* of x-ray is defined as,

$$SNR = \frac{\mu_n}{\sigma_n} = \frac{\mu}{\sqrt{\mu}} = \sqrt{\mu}$$

Now, consider an object (or target) of interest placed on a background. Let f_t and f_b be the average image intensities within the target and background, respectively. A useful choice for SNR is obtained by taking the "signal" to be the difference in average image intensity values between the target and the background integrated over the area of the target (A), and by taking the "noise" to be the random fluctuation of image intensity from its mean over and area of the background. This lead to the *differential signal to noise ratio* (DSNR), given by,

$$DSNR = \frac{A(f_t - f_b)}{\sigma_b(A)}$$

In x-ray imaging, f_b can describe the average photon count per unit area in the background region around a target, in which case $f_b = \lambda_b$. λ_b is the mean of the underlying Poisson distribution governing the number of background photons counted per unit area. Therefore, $\sigma_b(A)=\sqrt{\lambda_b A}$,

$$DSNR = \frac{C A \lambda_b}{\sqrt{A \lambda_b}} = C\sqrt{A \lambda_b} = C\sqrt{N_b}, \text{where C is the contrast and } N_b \text{ is the average number of photons per burst area}$$

As seen, SNR can be defined and measured in various ways, but any useful definition of it must contain contrast and noise. Therefore, a basic SNR measurement should contain,

$$SNR = \frac{C \mu}{\sigma_\mu}$$

Distortion

Distortion is geometrical in nature and refers to the inability of any imaging system to give an accurate impression of the shape, size and/or position of objects of interest.

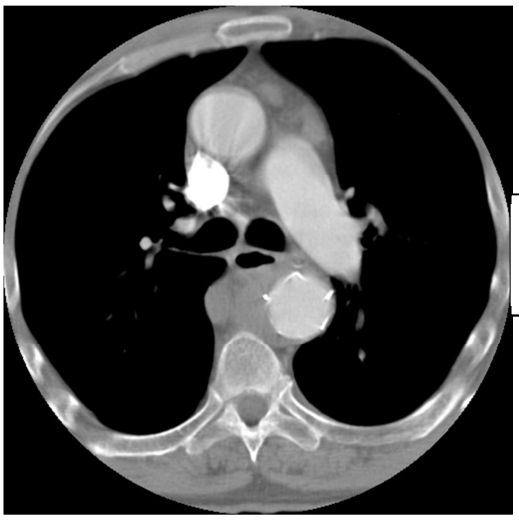
III. Artifacts

The term artifact is applied to any systematic discrepancy between the CT numbers in the reconstructed image and the true attenuation coefficients of the object. Artifacts do not represent valid anatomical objects and can obscure important targets or be falsely interpreted as valid image features. The types of artifact that can occur are as follows: **(a)** streaking, which is generally due to an inconsistency in a single measurement; **(b)** shading, which is due to a group of channels or views deviating gradually from the true measurement; **(c)** rings, which are due to errors in an individual detector calibration; and **(d)** distortion, which is due to helical reconstruction.

Physics-Based Artifacts

Physics-based artifacts result from the physical processes involved in the acquisition of CT data. They can also include helical and multisection artifacts, which are produced by the image reconstruction process. Following are various causes for physics-based artifacts.

Beam Hardening: An x-ray beam is composed of individual photons with a range of energies. As the beam passes through an object, it becomes "harder," i.e. its mean energy increases, because the lower energy photons are absorbed more rapidly than the higher-energy photons. Two types of artifact can result from this effect: so-called cupping artifacts and the appearance of dark bands or streaks between dense objects in the image. These artifacts can be corrected with the use of proper filtration, calibration correction, and an iterative correction algorithm.

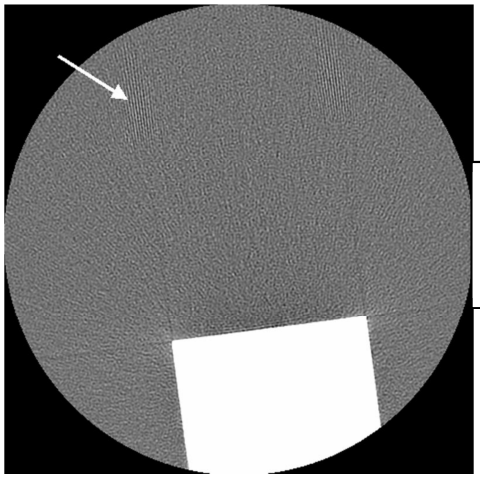


CT image shows streaking artifacts due to the beam hardening effects of contrast medium. Image provided by Barrett and Keat, Artifacts in CT: Recognition and Avoidance

Partial Volume: One type of partial volume artifact occurs when a dense object lying off-center protrudes partway into the width of the x-ray beam. The inconsistencies between the views cause shading

artifacts to appear in the image. Partial volume artifacts can best be avoided by using a thin acquisition section width.

Undersampling: The number of projections used to reconstruct a CT image is one of the determining factors in image quality. Too large an interval between projections (undersampling) can result in misregistration by the computer of information relating to sharp edges and small objects. This leads to an effect known as view aliasing, where fine stripes appear to be radiating from the edge of, but at a distance from, a dense structure. Stripes appearing close to the structure are more likely to be caused by undersampling within a projection, which is known as ray aliasing. Acquiring the largest possible number of projections per rotation can minimize view aliasing. Ray aliasing can be reduced by using specialized high-resolution techniques, such as quarter- detector shift or flying focal spot, which manufacturers employ to increase the number of samples within a projection.



CT image of a Teflon block in a water phantom shows aliasing (arrow) due to undersampling of the edge of the block. Image provided by Barrett and Keat, Artifacts in CT: Recognition and Avoidance

Photon Starvation: A potential source of serious streaking artifacts is photon starvation, which can occur in highly attenuating areas. When the x-ray beam is traveling horizontally, the attenuation is greatest and insufficient photons reach the detectors. The result is that very noisy projections are produced at these tube angulations. The reconstruction process has the effect of greatly magnifying the noise, resulting in horizontal streaks in the image. These artifacts can be corrected with automatic tube current modulation and adaptive filtration.

CT image of a shoulder phantom shows streaking artifacts caused by photon starvation. Image provided by Barrett and Keat, Artifacts in CT: Recognition and Avoidance

Scanner-Based Artifacts

Scanner-based artifacts result from imperfections in scanner function.

Ring Artifacts: If one of the detectors is out of calibration on a third-generation (rotating x-ray tube and detector assembly) scanner, the detector will give a consistently erroneous reading at each angular position, resulting in a circular artifact. A scanner with solid-state detectors, where all the detectors are separate entities, is in principle more susceptible to ring artifacts than a scanner with gas detectors, in which the detector array consists of a single xenon-filled chamber subdivided by electrodes. The presence of circular artifacts in an image is an indication that the detector gain needs recalibration or may need repair

services. Selecting the correct scan field of view may reduce the artifact by using calibration data that fit more closely to the patient anatomy.

Cone Beam Effect: As the number of sections acquired per rotation increases, a wider collimation is required and the x-ray beam becomes cone-shaped rather than fan-shaped. As the tube and detectors rotate around the patient, the data collected by each detector correspond to a volume contained between two cones, instead of the ideal flat plane. This leads to artifacts similar to those caused by partial volume around off-axis objects.

Patient-Based Artifacts

Patient-based artifacts are caused by such factors as patient movement or field of view (FOV) problems.

Patient motion: patient motion can cause misregistration artifacts, which usually appear as shading or streaking in the reconstructed image. Special care is taken to avoid voluntary motion, such as the use of positioning aids. Manufacturers minimize motion artifacts by using overscan and underscan modes, software correction, and cardiac gating.

CT image of a head shows motion artifacts. Image provided by Barrett and Keat, Artifacts in CT: Recognition and Avoidance

Incomplete projections: If any portion of the patient lies outside the scan field of view, the computer will have incomplete information relating to this portion and streaking or shading artifacts are likely to be generated. To avoid artifacts due to incomplete projections, it is essential to position the patient so that no parts lie outside the scan field.

IV. Metal Artifacts

Metal artifacts on CT are caused by the manifestation of several effects discussed in previous sections. (a) Metal artifacts represent the extremes of the beam hardening phenomenon since the complete attenuation of the beam results in gaps on the data from these shadows in the projection data. When using a *filtered-backprojection method,* these gaps in the projection data produce dark bands on the images around the metallic object. (b) The attenuation coefficient of metallic object is another factor to consider. (c) Motion of the interface between the metal and surrounding tissue can cause an accentuation artifact. Following are the parameters that relate to the mentioned generation of CT artifacts from metal hardware: X-ray kVp, tube current (mA), pitch, hardware composition, geometry, location, and the image reconstruction parameters.

Imaging Parameters

X-ray kVp: A higher X-ray kVp can increase the ability of the X-ray beam to penetrate metal.

X-ray tube current (mA): A higher tube current setting may increase the ability of the X-ray beam to penetrate metal. These settings are possible with multichannel CT because of its lower pitch settings

Pitch: With multichannel CT, a lower pitch setting allows the collection of redundant data, thereby increasing the likelihood that adequate projection data will be collected. In addition, lower pitch settings reduce splay artifacts (radially alternating bands of higher- and lower-attenuation projection across the

image) which are inherent in multichannel CT. Given that scanner setup and technical factors remain unchanged, splay artifacts are reduced as the number of detector rows increases

Metal composition: The artifacts generated by orthopedic hardware are related to the composition of the hardware. Materials with lower X-ray beam attenuation coefficients (density) produce fewer artifacts (e.g., plastic < titanium < vitallium < stainless steel < cobalt–chrome).

Geometric factors: The amount of artifacts generated at CT is also related to the cross-sectional area of the hardware. Artifacts generally predominate in the direction of the maximal thickness of the hardware; in fact, the degree of X-ray beam attenuation is proportional to the thickness of the hardware. If it is possible, the affected body part should be positioned so that the X-ray beam passes the metal at its smallest cross-sectional area, although the repositioning of body parts is generally not possible for spine imaging. X-ray beam attenuation is also greatest in the regions of greatest patient girth and bone mass. Hence, X-ray beam attenuation in the cervical spine is less than that in the thoracic spine, which in turn is less than that in the lumbar and sacral spine.

Image Reconstruction Parameters

Filter (Kernel): Selection of an appropriate reconstruction filter may play a critical role in the appearance of a metal-related artifact. The use of a standard or smooth reconstruction filter is preferred, particularly in the presence of dense metallic hardware and in patients with a large body habitus.

Multiplanar reconstruction: Reformatted images can be thicker than those initially acquired by averaging pixel values within slices. On these thicker reformats, metal-induced artifacts are reduced and signal-to-noise ratio is enhanced.

Volumetric reconstructions: Surface 3D or volumetric rendered images are occasionally useful in providing an additional view on the relationship of hardware to the bones or by providing a global view of the regional anatomy. They can display metal hardware fracture that is easily overlooked on multiplanar reconstruction (MPR). Volumetric rendering techniques provide semitransparent views of bones that tend to reduce metal artifacts. They clearly show location of metal hardware and their relationship to the adjacent bone.

Window levels: The use of wide windows to review or to film images (3,000–4,000 HU window width, 800 window level) facilitates visualization of structures adjacent to metal hardware and reduce the effects of metal artifacts. The use of extended-scale CT and dedicated workstations also will decrease the appearance of metal artifacts by allowing for large window widths.

V. Metal Artifact Removal Techniques

Various types of metal artifact reduction methods have been proposed and the can be grouped into two categories:

- *Projection completion methods*: missing data are replaced by synthetic data, obtained by interpolation, pattern recognition, or linear prediction methods
- *Iterative methods*: existing iterative methods, such as the maximum likelihood expectation maximization algorithm or the algebraic reconstruction technique are modified in order to ignore missing data.

Sinogram Inpainting Method

Sinogram inpainting methods are the most common MAR technique. They utilize interpolation or forward projections to complete the sinogram, where metal-affected values are treated as missing data. The first step to reduce the existing metal artifacts consists of a segmentation of the metal trace. Metal objects are labeled in a preliminary filtered backprojection reconstruction through the use of a simple threshold segmentation. Then, a forward projection of the metal-only image is calculated resulting in a sinogram mask. All sinogram values which are different from zero in this mask are assumed to be inconsistent. Next, the inconsistent data is replaced through a two-dimensional sinogram restoration technique based on an approach adapted from image inpainting ideas. The goal of image inpainting is to modify an image as far as possible in an undetectable way; here, it is used to replace the inconsistent projection data. The gap inside the sinogram data, the inpainting region, is filled in such a way that it is not detectable in the resulting image. After a number of iterations of inpainting, the repaired sinogram data are reconstructed. However, this method tends to introduce new streak artifacts during sinogram restoration. The origin of these new artifacts

is related to the loss of edge information of the objects by using surrogate data. This loss of edge information affects the entire image and is not restricted to edges next to metal objects.

Transmission Maximum A Posteriori Algorithm

This algorithm is centered on the hypothesis that improving the accuracy of the mathematical model employed by the reconstruction algorithm should reduce or eliminate artifacts. It uses a Markov random field smoothness prior and applies increased sampling in the reconstructed image and is based on a transmission maximum-likelihood algorithm in which an attenuation image is reconstructed by optimizing the likelihood, assuming that the measured detector read-outs have a Poisson distribution. As such, less weight is attributed to low-count detector read-outs, making the algorithm robust against other sources of artifacts that are most prominent in directions of low counts, such as beam hardening, scatter, and partial volume effects. In addition to the transmission maximum-likelihood algorithm, the a posteriori algorithm employs a Markov random field smoothness prior with a Huber potential function, as well as performs iterations using a reconstruction image at double resolution. After the last iteration, the image is re-sampled to normal resolution. This higher resolution provides a better model for sharp transitions in the image.

Normalization Method

While sinogram interpolation methods efficiently remove metal artifacts, new artifacts are often introduced in the process, as interpolation cannot completely recover the information from the metal trace. The length normalization method allows for efficient reduction of metal artifacts while adding almost no new ones. In the first step, metal is segmented in the image domain by thresholding, the same process as used in sinogram inpainting. A 3D forward projection identifies the metal trace in the original projections. Before interpolation, the projections are normalized based on a 3D forward projection of a prior image. This prior image is obtained by a multithreshold segmentation of the initial image. The original raw data are then normalized by dividing by the projection data of the prior image and, after interpolation, denormalized again. This denormalization is done by multiplying the interpolated and normalized sinogram with the intersection lengths. Reconstruction of this corrected sinogram yields the corrected image.

The steps of normalized metal artifact reduction (NMAR) are shown in the diagram below:

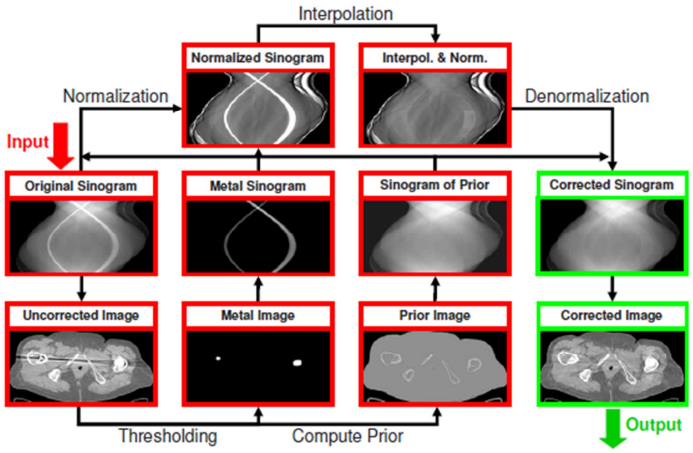


Image provided by Meyer et al.: Normalized metal artifact reduction (NMAR) in computed tomography

Frequency Split Metal Artifact Reduction

FSMAR combines a raw data inpainting-based MAR method with an image-based frequency split approach. It consists of five steps, shown in the diagram below:

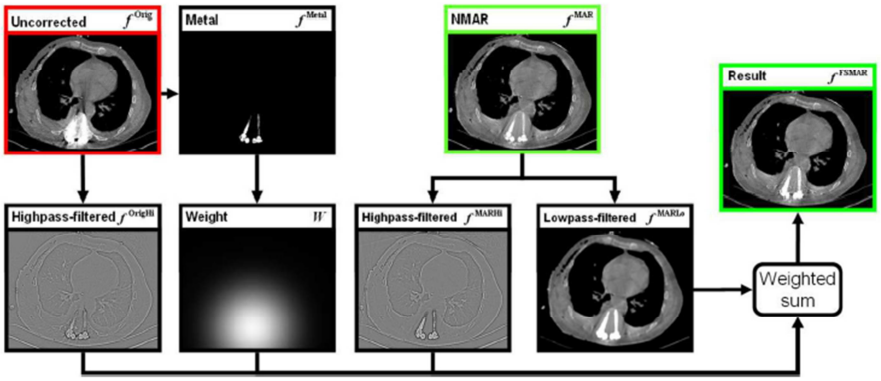


Image provided by Meyer et al.: Frequency split metal artifact reduction

1) Preprocessing: An adaptive filter is applied to the raw data; this is unnecessary in standard inpainting MAR approaches because the data from the shadow of the metal artifact is completely replaced. With FSMAR, data from the metal shadow contribute to the final image and thus, noise reduction is recommended. From the preprocessed raw data, an image is reconstructed.
2) Segmentation of Metal: From the reconstructed image, a metal image is segmented by simple thresholding. This image contains only the metal parts, and any pixel that does not belong to the metal implant is set to zero.
3) MAR by Inpainting: The metal image from the second step is then forward projected. The location of the positive entries in this projection dataset defines the metal trace in the original data, which must be replaced by an inpainting method.
4) Frequency Split: FSMAR combines parts of the high frequencies of the images reconstructed in the first and second steps. The metal image is more reliable than the original image with respect to the low frequencies as it no longer contains beam hardening and scatter artifacts. The original image contains the edges and fine anatomical structures in addition to some remaining streak artifacts. Thus a spatially varying weight is used to combine the advantages of both high frequency images.
5) Spatial Weighting: Using the complete high frequency part of the original image would unnecessarily increase the noise in the final corrected image, thus, a weight is computed for each image pixel. Pixels close to metal implants are weighted higher than pixels more distant from the metal artifacts. This weight image is obtained through binarizing the metal image, smoothing it with a strong low-pass filter, and normalizing it to a maximum value of one.

VI. Summary

A fundamental knowledge of the basic principles of CT imaging, such as strength of radiation, attenuation, and dose, as well as the parameters and measurements of the CT machine itself, are critical to being able to understand the factors responsible for CT image quality. Artifacts and their sources, in particular, metal artifacts, cause serious degradation of image quality and hinder the ability of medical professionals to diagnose and treat disease. Effective metal artifact removal techniques are essential in order to reduce the distortive results of such artifacts and to reveal the relevant anatomical structures.

References

[1] Barrett, J. F., and N. Keat. "Artifacts in CT: Recognition and Avoidance." Radiographics 24.6 (2004): 1679-691. Print.
[2] Prince, Jerry L., and Jonathan M. Links. *Medical Imaging Signals and Systems*. Upper Saddle River, NJ: Pearson Prentice Hall, 2006. Print.
[3] Müller, J.; Buzug, T. M. "Spurious structures created by interpolation-based CT metal artifact reduction." Medical Imaging 2009: Physics of Medical Imaging. Proceedings of the SPIE, Volume 7258 (2009).
[4] B. De Man1, J. Nuyts2, P. Dupont2, G. Marchal3, and P. Suetens."Reduction of metal streak artifacts in x-ray computed tomography using a transmission maximum a posteriori algorithm." IEEE Transactions on Nuclear Science, vol. 47, nr. 3, 2000, pp. 977-981.
[5] Meyer, Esther, Rainer Raupach, Michael Lell, Bernhard Schmidt, and Marc Kachelrieß. "Normalized Metal Artifact Reduction (NMAR) in Computed Tomography." Medical Physics 37.10 (2010): 5482. Print.
[6] Meyer, Esther, Rainer Raupach, Michael Lell, Bernhard Schmidt, Marc Kachelrieß. "Frequency split metal artifact reduction (FSMAR) in computed tomography." Medical Physics 39.00 (4), April 2012. Print.
[7] P. Stradiotti, A. Curti, G. Castellazzi, A. Zerbi. "Metal-related artifacts in instrumented spine. Techniques for reducing artifacts in CT and MRI: state of the art" Eur Spine J. 2009 June; 18(Suppl 1): 102–108. Published online 2009 May 13.
[8] M. Oehler, T. M. Buzug. "Sinogram Inpainting for Metal Artifact Reduction in CT Images." Proc. 4th European Congress for Medical and Biomedical Engineering, Springer IFMBE Series, 2008.