Cone Beam Computed Tomography

Oral and Maxillofacial Diagnosis and Applications

COPYRIGHTED MATERIAL

1 Technology and Principles of Cone Beam Computed Tomography

Matthew W. Jacobson

This chapter aims to convey a basic technical familiarity with compact Cone Beam Computed tomography (CBCT) systems, which have become prevalent since the late 1990s as enablers of in-office CT imaging of the head and neck. The technical level of the chapter is designed to be accessible to current or candidate end users of this technology and is organized as follows. In Section 1, a high-level overview of these systems is given, with a discussion of their basic hardware components and their emergence as an alternative to conventional, hospital CT. Section 2 gives a treatment of imaging basics, including various aspects of how a CT image is derived, manipulated, and evaluated for quality.

Section 1: Overview of compact cone beam CT systems

Computed tomography (CT) is an imaging technique in which the internal structure of a subject is deduced from the way X-rays penetrate the subject from different source positions. In the most general terms, a CT system consists of a gantry which moves an X-ray source to different positions around the subject and fires an X-ray beam of some shape through the subject, toward an array of detector cells. The detector cells measure the amount of X-ray radiation penetrating the subject along different lines of response emanating from the source. This process is called the *acquisition* of the X-ray measurements. Once the X-ray measurements are acquired, they are transferred to a computer where they are processed to obtain a CT image volume. This process is called *image reconstruction*. Once image reconstruction has been performed, the computer components of the system make the CT image volume available for display in some sort of image viewing software. The topics of image reconstruction and display will be discussed at greater length in Section 2.

Cone beam computed tomography refers to CT systems in which the beam projected by the X-ray source is in the shape of the cone wide enough to radiate either all or a significant part of the volume of interest. The shape of the beam is controlled by the use of collimators, which block X-rays from being emitted into undesired regions of the scanner field of view. Figure 1.1 depicts a CBCT system of a compact variety suitable for use in small clinics. In the particular system shown in the figure, the gantry rotates in a circular path about the subject firing a beam of X-rays that illuminates the entire desired field of view. This results in a series of

Cone Beam Computed Tomography: Oral and Maxillofacial Diagnosis and Applications, First Edition. Edited by David Sarment.
© 2014 John Wiley & Sons, Inc. Published 2014 by John Wiley & Sons, Inc.

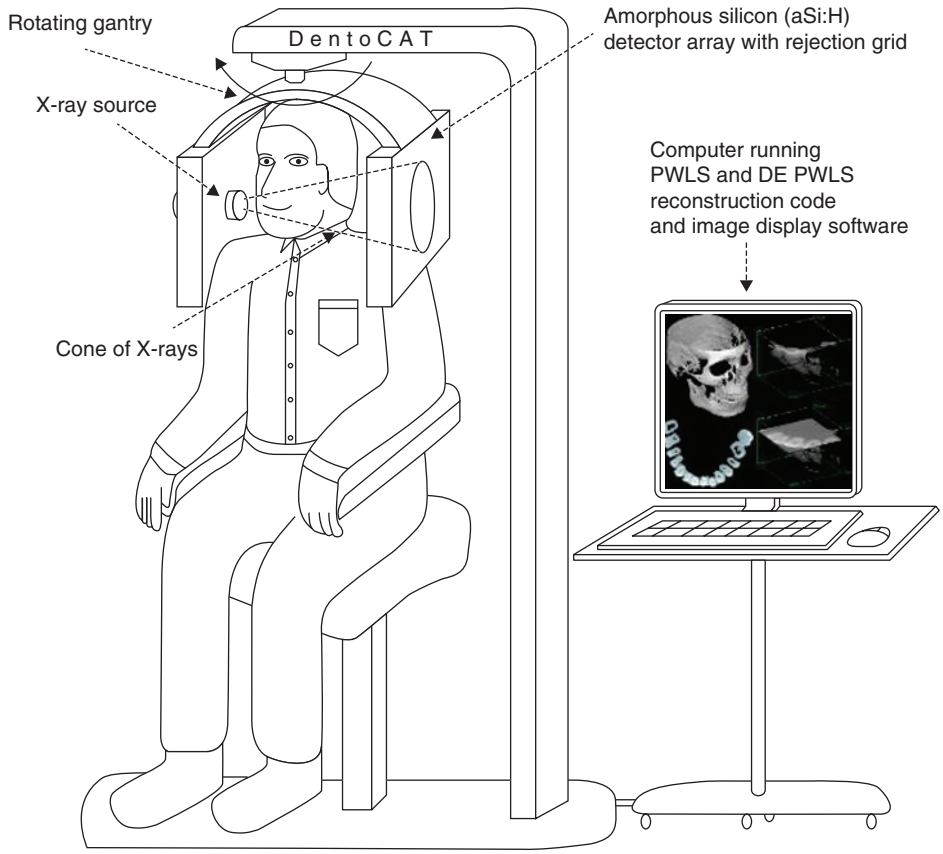

Figure 1.1 The proposed design of DentoCAT. The patient is seated comfortably in chair (the chin-rest is not shown). DentoCAT features cone beam geometry, aSi:H detector array, PWLS and DE PWLS reconstruction methods.

two-dimensional (2D) images of the X-ray shadow of the object that is recorded by a 2D array of detector cells. Cone beam CT systems with this particular scan geometry will be the focus of this book, but it is important to realize that in the broader medical imaging industry, CT devices can vary considerably both in the shape of the X-ray beam and the trajectory of the source.

Prior to the introduction of CBCT, it was common for CT systems to use so-called *fan beam* scan geometries in which collimators are used to focus the X-ray beam into a flat fan shape. In a fan beam geometry, the source must travel not only circularly around the subject but also axially along the subject's length in order to cover the entire volume of interest. A helical (spiral) source trajectory is the most traditional method used to accomplish this and is common to most hospital CT scanners. The idea of fan beam geometries is that, as the source moves along the length of the subject, the X-ray fan beam is used to scan one cross-sectional slice of the subject at a time, each of which can be reconstructed individually. There are several advantages to fan beam geometries over cone beam geometries. First, since only one cross-section is being acquired at a time, only a 1-dimensional detector array is required, which lowers the size and cost of the detector. Second, because a fan beam only irradiates a small region of the object at a given time, the occurrence of scattered X-rays is reduced. In cone beam systems, conversely, there is a much larger component of scattered radiation, which has a corrupting effect on the scan (see "Common Image Artifacts" section). Finally, in a fan beam geometry, patient movement occurring during the scan will only degrade image quality in the small region of

the subject being scanned when motion occurs. Conversely, in cone beam systems, where larger regions of anatomy are irradiated at a given time, patient movement can have a much more pervasive effect on image quality.

The disadvantage of fan beam geometries, however, is their inefficient use of X-ray output. Because collimators screen away X-ray output from the source except in the narrow fan region of the beam, much of the X-rays generated by the source go unused. Accordingly, the source must generate more X-ray output than a cone beam geometry for the same region scanned, leading to problems with source heating. Regulating the temperature of the source in such systems requires fast rotating source components, accompanied by a considerable increase in mechanical size, complexity, and expense. As the desire for greater volume coverage has grown in the CT industry, the difficulties with source heating have been found to outweigh the advantages of fan beam scanning, and the CT industry has been gradually moving to cone beam scan geometries. Cone beam geometries have other advantages as well, which have further motivated this shift. The spatial resolution produced by cone beam CT scanners, when used in conjunction with flat panel X-ray detectors, tends to be more uniform than fan beam–based systems.

Although the CT industry as a whole has been trending toward cone beam scanning, the hardware simplifications brought on by CBCT have played a particularly important role in the advent of compact in-office CT systems, of the kind shown in Figure 1.1. Conventional hospital CT scanners are bulky and expensive devices, not practical for in-office use. The reason for their large size is in part due to source cooling issues already mentioned and in part due to the fact that hospital CT systems need to be all-purpose, accommodating a comprehensive range of CT imaging tasks. To accommodate cardiac imaging, for example, hospital CT systems must be capable of very fast gantry rotation (on the order of one revolution per second) to deal with the movement of the heart. This has further exacerbated the mechanical power requirements, and hence the size and expense of the system.

The evolution of compact CT came in part from recognizing how cone beam scanning and other system customizations can mitigate these issues. As discussed, the use of a cone beam scanning geometry increases the efficiency of X-ray use, leading to smaller and cheaper X-ray sources that are easier to cool. Additionally, the imaging needs of dentomaxillofacial and otolaryngological medical offices have generally been restricted to high-contrast differentiation between bone and other tissues in nonmotion prone head and neck anatomy. CT systems customized for such settings can therefore operate both at lower X-ray exposure levels and at slower scanning speeds (on the order of 20–40 sec) than hospital systems. Not only does this further mitigate cooling needs of the X-ray source, it also leads to cheaper and smaller gantry control components.

The emergence of compact CBCT was also facilitated in part by recent progress in fast computer processor technology and in X-ray detector technology. The mathematical operations needed to reconstruct a CT image are computationally intensive and formerly achievable at clinically acceptable speeds only through expensive, special purpose electronics. With the advent of widely available fast computer processors, especially the massively parallel programming now possible with common video game cards, the necessary computer hardware is cheaply available to CT manufacturers and hence also to small medical facilities. Improvements in X-ray detector technology include the advent of flat panel X-ray detectors. Early work on compact CT systems (circa 2000) proposed using X-ray detectors based on image intensifier technology, then common to fluoroscopy and conventional radiography. However, flat panels have provided an alternative that is both cheaply available and also offers X-ray detection with less distortion, larger detector areas, and better dynamic range.

The development of compact CBCT for the clinic has made CT imaging widely and quickly accessible. Where once patients may have had to wait weeks for a scan referred out to the hospital, they may now be scanned and treated in the same office visit. The prompt availability of CT has also been cited as a benefit to the learning process of physicians, allowing them to more quickly correlate CT information with observed symptoms. Some controversy has sprung up around this technology, with questions including how best to regulate X-ray dose to patients. The financial compensation that physicians receive when prescribing a CT scan is argued to be a counterincentive to minimizing

patient X-ray dose. In spite of the controversy, CBCT has found its way into thousands of clinics over the last decade and is well on its way to becoming standard of care.

Section 2: Imaging basics for compact cone beam CT systems

This section describes the image processing software components of compact CBCT systems that go into action once X-ray measurements have been acquired. Tasks performed by these components include the derivation of a CT image volume from the X-ray measurements (called *image reconstruction*) and the subsequent display, manipulation, and analysis of this volume. In the subsections to follow, these topics will be covered in a largely qualitative manner suited to practitioners, with a minimum of mathematical detail.

Overview of image processing and display

The volume image data obtained from a CT system is a 3D map of the *attenuation* of the CT subject at different spatial locations. Attenuation, often denoted μ, is a physical quantity measuring the tendency of the anatomy at a particular location to obstruct the flight of X-ray photons. Because attenuation is proportional to tissue density, a 3D map of attenuation can be used to observe spatial variation in the tissue type of the subject anatomy (e.g., soft tissue versus bone). The attenuation applied to an X-ray photon at a certain location also depends on the photon energy. Ideally, when the X-ray source emits photons of a single energy level only, this energy dependence is of minor consequence. In practice, however, an X-ray source will emit photons of a spectrum of different energies, a fact that introduces complications to be discussed later.

Once the X-ray measurements have been acquired, the first processing step performed is to choose an imaging field of view (FOV), a region in space where the CT subject is to be imaged. For circularly orbiting cone beam CT systems, this region will typically be a cylindrical region of points in space that are all visible to the X-ray camera throughout its rotation and that cover the desired anatomy. A process of image reconstruction is then performed in which the X-ray measurements are used to evaluate the attenuation at various sample locations within the FOV. The sample locations typically are part of a 3D rectangular lattice, or reconstruction grid, enclosing the FOV cylinder (see Figure 1.2). The sample locations are thought of as lying at the

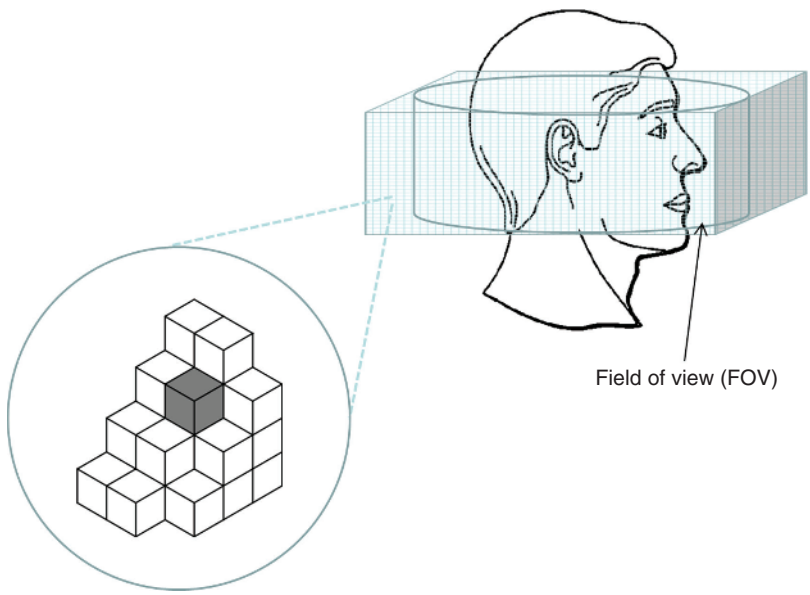

Figure 1.2 The concept of a reconstruction grid and field of view.

center of small box-shaped cells, called voxels. For image analysis and display purposes, the attenuation of the subject is approximated as being uniform over the region covered by a voxel. Thus, when the reconstruction software assigns an attenuation value to a grid sample location, it is in effect assigning it to the entire box-shaped region occupied by the voxel centered at that location.

The following section will delve into image reconstruction in more detail. For now, we simply note that the selection of an FOV and reconstruction grid brings a number of design trade-offs into play, and must be optimized to the medical task at hand. The selected FOV must first of all be large enough to cover the anatomy to be viewed. In addition, certain medical tasks will require the voxel sizes (equivalently, the spacing between sample points), to be chosen sufficiently small, to achieve a needed resolution. On the other hand, enlarging the FOV and/or increasing the sampling fineness will, in turn, increase the number of voxels in the FOV that need reconstructing. For example, simply halving the voxel size in all three dimensions while keeping the FOV size fixed translates into an eight-fold increase in the number of FOV voxels. This leads in turn to increased computational burden during reconstruction and slows reconstruction speed. Moreover, when sampling fineness in 3D space is increased, the sampling fineness of the X-ray measurements must typically be increased proportionately in order to reconstruct accurate values. This leads to similar increases in computational strain. Finally, as the FOV size is increased, there is a corresponding increase both in radiation dose to the patient, and also the presence of scattered radiation, which leads to a degrading effect on the CT image (see "Common Image Artifacts" section).

Attenuation is measured in absolute units of inverse length (mm^{-1} or cm^{-1}). However, for purposes of analysis and display, it is standard throughout the CT industry to re-express reconstructed image intensities in CT numbers, a normalized quantity which measures reconstructed attenuation relative to the reconstructed attenuation of water:

$$\text{CT No.} = \frac{\mu - \mu_{\text{water}}}{\mu_{\text{water}}} \times 1000$$

The value of μ_{water} is obtained in a system calibration step by reconstructing a calibration phantom consisting of water-equivalent material. CT numbers are measured in Hounsfield units (HU). In this scale, water always has a CT number of zero, while for air (with $\mu_{\text{air}}=0$), the CT number is –1000. Expressing image intensity in HU instead of physical attenuation units provides a more sensitive scale for measuring fine attenuation differences. Additionally, it can help to cross-compare scans of the same object from different CT devices or using different X-ray source characteristics. The effect of the different system characteristics on the contrast between tissue types is more easily observed in the normalized Hounsfield scale, in which waterlike soft tissue is always anchored at a value near zero HU.

Once the reconstructed 3D volume is converted to Hounsfield units, it is made available for display in the system's image viewing software. Typically, an image viewer will offer a number of standard capabilities, among them a multiplanar rendering (MPR) feature that allows coronal, sagittal, or axial slices of the reconstructed object to be displayed (see Figure 1.3). The slices can be displayed as reconstructed, or one can set a range of neighboring slices to be averaged together. This averaging can reduce noise and improve visibility of anatomy at some expense in resolution. Other typical display functions include the ability to rotate the volume so that MPR cross-sections at arbitrary angles can be displayed, a tool to measure physical distances between points in the image, and a tool for plotting profiles of the voxel values across one-dimensional cross-sections.

CT display systems will also provide a drawing tool allowing regions of interest (ROIs) to be designated in the display. The drawing tool will typically show the mean and standard deviation of the voxel values as well as the number of voxels within the ROI to be computed. For CT systems in the U.S. market, this feature is in fact federally required under 21 CFR 1020. Figure 1.4 illustrates a circular ROI drawn in a commercial CT viewer, with the relevant ROI statistics displayed. One function of this tool is to verify certain performance specifications that the CT manufacturer is federally required to provide in the system data sheets and user manual. These metrics will be discussed in greater detail in the "Imaging Performance" section.

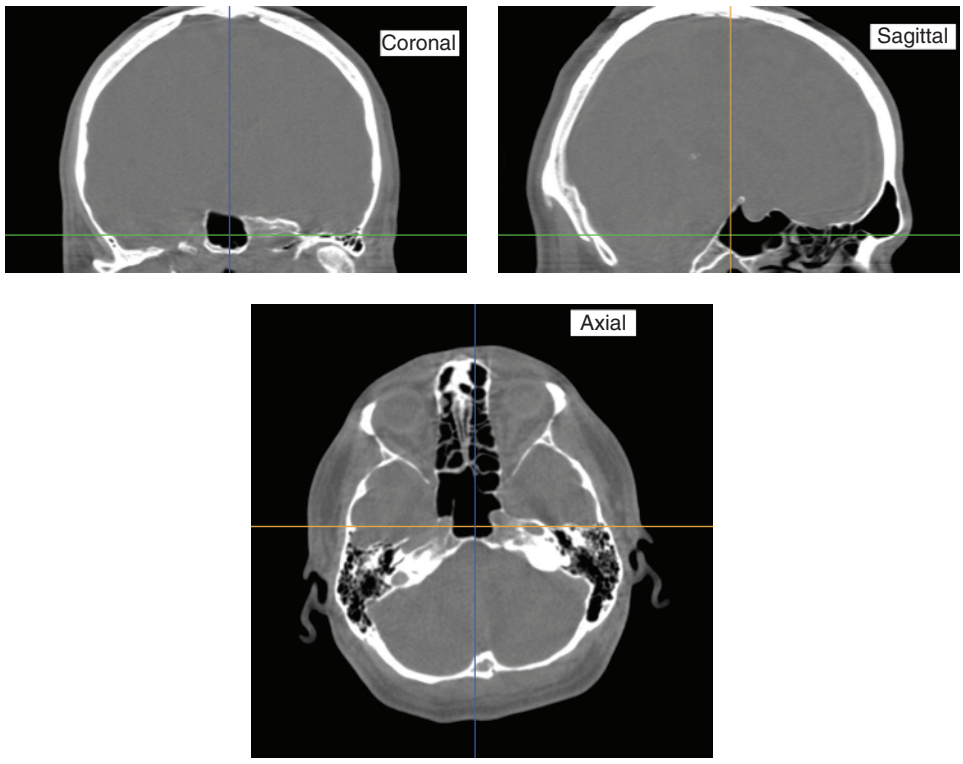

Figure 1.3 Multiplanar rendering of a CT subject.

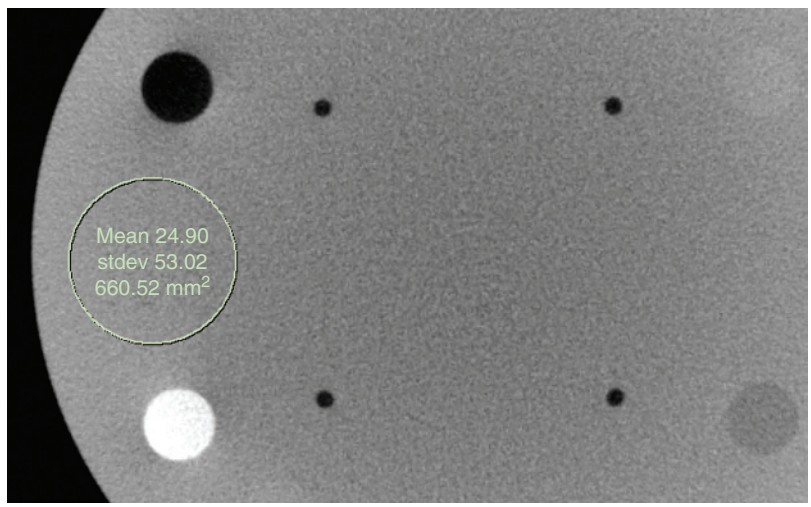

Figure 1.4 Illustration of a region of interest drawing tool in the display of a reconstructed CT phantom.

Another important display capability is the ability to adjust the viewing contrast in the image. Because there are a limited number of different brightness levels that can be assigned to a voxel for display purposes, the viewing software will divide the available brightness levels among the CT numbers in a user-selected range, or window. Image voxels whose CT numbers fall between the

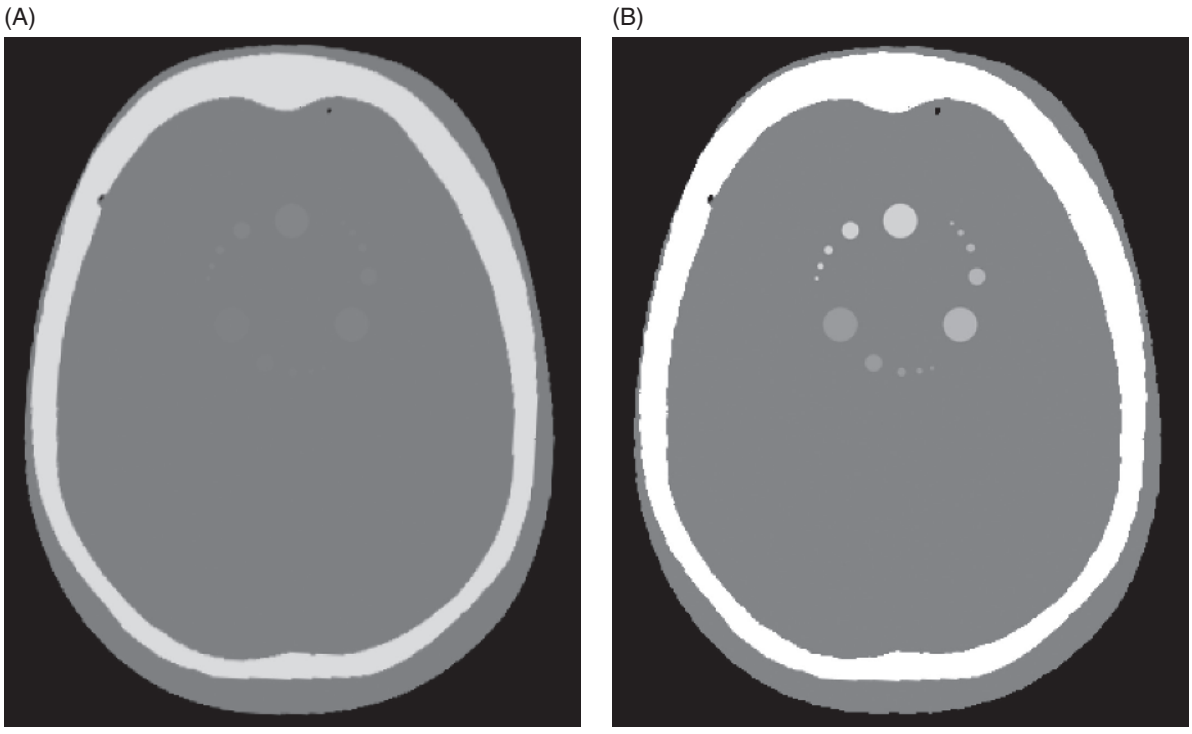

Figure 1.5 Axial slice of computer-generated phantom in (A) a high-contrast viewing window (L/W = 50/1200 HU), and (B) a low-contrast window (L/W = 30/90 HU).

minimum and maximum values set by the window are assigned a proportionate brightness level. If a voxel value falls below the minimum CT number in this range, it will be given zero brightness, whereas if it lies above the maximum CT number, it will be assigned the maximum brightness. It is common to express a window setting in terms of a level (L), meaning the CT number at the center of the range, and a window width (W), meaning the difference between the maximum and minimum CT number in the range. For example, a window ranging between 400 HU and 500 HU would be specified as L=450 HU and W=100 HU.

Narrowing the display window about a particular intensity level allows for better contrast between subtly different image intensities within the window. Figure 1.5 shows an axial slice of a computer-generated head phantom as displayed in both a wide, high-contrast window (Figure 1.5A) and a narrow, low-contrast window (Figure 1.5B). Clearly, the narrower window offers better visibility of the pattern of low-contrast discs in the interior of the slice. At the time of this writing, however, low-contrast viewing windows are more commonly employed by users of compact CBCT systems. This is because certain limitations of the cone beam geometry and of current flat panel technology, to be elaborated upon later, render image quality poor when viewed in high-contrast windows. The industry has therefore been limited to head and neck imaging where often only the coarse differentiation between bone and soft tissue are needed. For these applications, low-contrast viewing windows, such as in Figure 1.5B, tend to be sufficient. The terms *soft tissue window* and *bone window* are commonly used to distinguish between display range settings appropriate, respectively, to soft tissue differentiation and coarse bone/soft tissue differentiation tasks. Soft tissue windows will use window levels of 30–50 HU and window widths of one to several hundred HU. The bone window will use window levels of 50–500 HU and window widths of anywhere from several hundred to over a thousand Hounsfields.

The images in Figure 1.3 and Figure 1.5A are displayed at L/W=50/1200 HU, a setting representative of the bone window. Figure 1.5B is displayed at L/W=30/90 HU, a setting at the narrower end of different possible soft tissue windows.

Image reconstruction

Image reconstruction is the process by which attenuation values for each voxel in the CT image are calculated from the X-ray measurements. This process tends to be the most computationally intensive software task performed by a CBCT system. There are tens of millions of voxels in a typical reconstruction grid and each computed voxel value derives information from X-ray measurements taken typically at hundreds of different gantry positions. A complete image reconstruction task may hence require, at minimum, tens of billions of arithmetic and memory transfer operations. CT manufacturers therefore invest considerable development effort in making reconstructions achievable within compute times acceptable in a clinical environment. Because of the computational hurdles associated with image reconstruction, commercial systems have historically resorted to filtered back projection algorithms. These are among the simplest reconstruction approaches computationally but have certain limitations in the image quality they can produce. As computer processor power has increased over time, however, and especially with the recent proliferation of cheaply available parallel computing technology, the CT industry has begun to embrace more powerful, if more computationally demanding, iterative reconstruction algorithms. The next section will overview conventional filtered back projection reconstruction, which is still the most prevalent approach. The section titled "Iterative Reconstruction" will then give a short introduction to emerging iterative reconstruction methods and some rudimentary demonstrations.

Conventional filtered back projection

To understand conventional image reconstruction, one must first consider a particular line of X-ray photon flight, one that emanates from the X-ray focal spot (see Figure 1.6) to a particular pixel on the detector panel for some particular gantry position. One then considers sample attenuation values of the CT subject along this line, with sample locations spaced at a separation distance, d. If the samples are weighted by this separation distance and summed, then as the separation distance is taken smaller and smaller (making the sampling more and more dense), this weighted sum approaches a

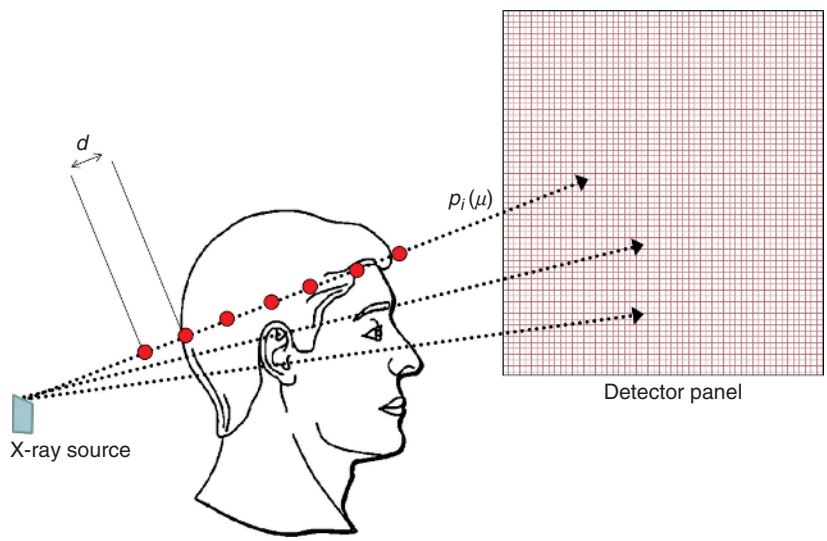

Figure 1.6 The concept of a geometric projection.

limiting value, p_i (μ), known as the *geometric projection*, or X-ray transform, of the attenuation map, μ, along the *i*-th measured X-ray path. The idea behind most conventional reconstruction techniques is to extract measurements of the geometric projections from the raw physical X-ray measurements and to then apply known mathematical formulas for inverting the X-ray transform.

The calculation of geometric projections from raw X-ray measurements requires the knowledge of certain physical properties of the source-detector X-ray camera assembly. For example, it is necessary to know the sensitivity of each detector pixel to X-rays fired in air, with no object present in the field of view. It is also necessary to know the detector offset values, which are nonzero signals measured by the detector even when no X-rays are being fired from the source. The offset signals originate from stray electrical currents in the photosensitive components of the detector. These properties are measured in a calibration step performed at the time of scanner installation, by averaging together many frames of an air scan and a blank scan (a scan with no X-rays fired). The air scan and blank scan response will drift over time due to temperature sensitivity of the X-ray detector and gradual X-ray damage, and therefore they must be refreshed periodically, typically by recalibrating the device at least daily.

Once the geometric projections have been calculated, an inverse X-ray transform formula is applied. Commonly, such formulas reduce to a filtering step, applied view-by-view to the geometric projections, followed by a so-called *back projection* step in which the filtered projection values are smeared back through the FOV. Algorithms that implement the reconstruction this way are thus called *filtered back projection* (FBP) algorithms and are used in a range of tomographic systems, both in CT and other modalities. The fine details of both the filtering step and the back projection step are somewhat dependent on the scanning geometry, that is, on the shape of the gantry orbit and the shape of the radiation beam. Generally speaking, however, the filtering step will be an operation that sharpens anatomical edges in the X-ray projections while dampening regions of slowly varying intensity. The smearing action of back projection, meanwhile, will typically be along the measured X-ray paths connecting the X-ray source to the panel, in a sense undoing the forward projecting action of the radiation source. For circular orbiting cone beam CT systems, our primary focus here, a well-known FBP algorithm is the Feldkamp Davis Kress (FDK) algorithm (Feldkamp and Davis, 1984). We will focus on the FDK algorithm for the remainder of this section.

Figure 1.7 illustrates the stages of FDK reconstruction up through filtering, including the data

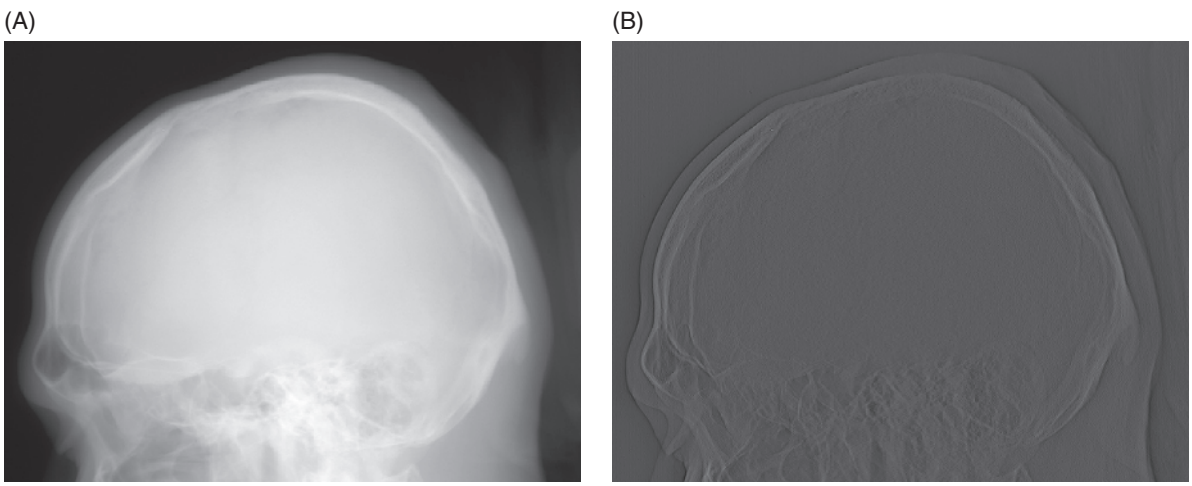

Figure 1.7 Illustration of the precorrection and filtering stages of the FDK algorithm for a CT subject. (A) One frame of precorrected geometric projection measurements. (B) The same frame after filtering.

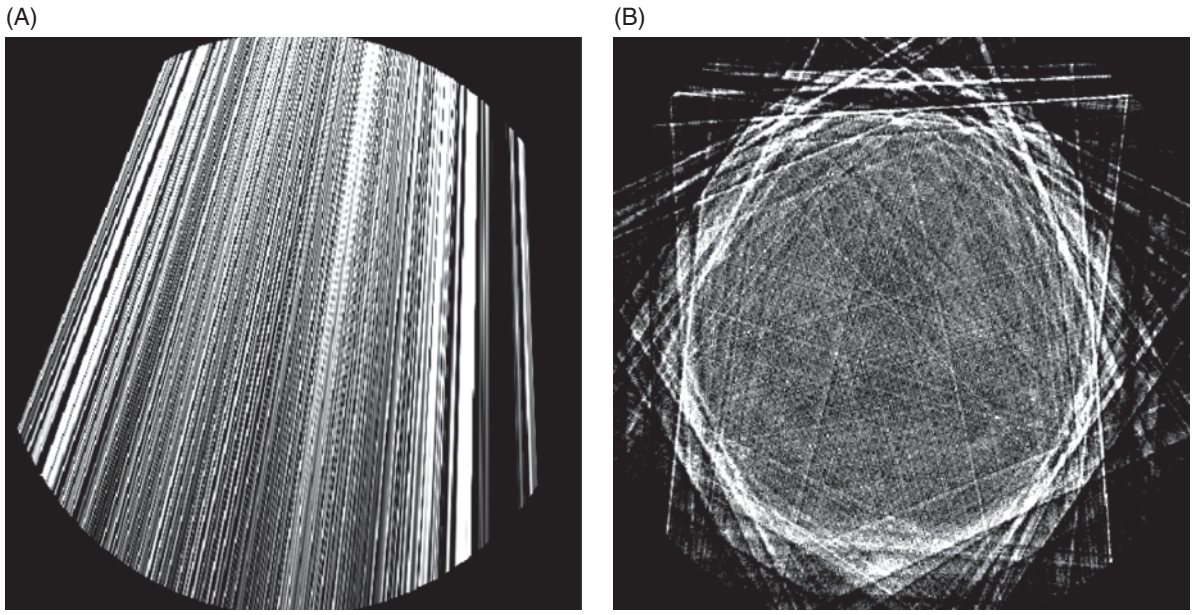

Figure 1.8 The back projection step of the FDK algorithm for progressively larger numbers of frames: (A) 1 frame. (B) 12 frames.

precorrection step, for one frame of a cone beam CT scan. The edge sharpening effect of the filter is clear in Figure 1.7B. Because the sharpening operation can also undesirably amplify sharp intensity changes due to noise, the filtering operation will also employ a user-chosen cutoff parameter. Intensity changes that are "too sharp," as determined by the cutoff, are interpreted by the filter as noise, rather than actual anatomy, and are therefore smoothed. Generally speaking, it is impossible to distinguish anatomical boundaries from noise with perfect reliability, and so applying the cutoff always leads to some sacrifice in resolution in the final image. A judgment must be made by the system design engineers as to the best tradeoff between noise suppression and resolution preservation.

Figure 1.8 shows the result of back projecting progressively larger sets of X-ray frames. In Figure 1.8A, where only a single frame is back projected, one can see how smearing the projection intensities obtained at that particular gantry position back through the FOV results in a pattern demarcating the shape of the X-ray cone beam. In Figure 1.8B, C, D, and E, as contributions of more gantry positions are added, the true form of the CT subject gradually coalesces.

As mentioned earlier, image reconstruction is computationally expensive compared to other processing steps in a CT scan. For conventional filtered back projection, most of that expense tends to be concentrated in the back projection step. For the filtering step, very efficient signal processing algorithms exist so that filtering can be accomplished in a few tens of operations per X-ray measurement. Conversely, in back projection, each X-ray measurement contributes to hundreds of voxels lying along the corresponding X-ray path and therefore results in hundreds of computations per data point. Perhaps even more troublesome is that both the voxel array and the X-ray measurement array are too large to be held in computer cache memory. When naively implemented, a back projection operation can therefore result in very time-consuming memory-access operations. Accordingly, a great deal of research over the years has been devoted to acceleration of back projection operations. For example, a method for approximating a typical back projection with greatly reduced operations was proposed by Basu and Bresler (2001). Later, the same group proposed a method that makes memory access patterns more efficient, resulting in strong acceleration over previous methods (De Man and Basu, 2004).

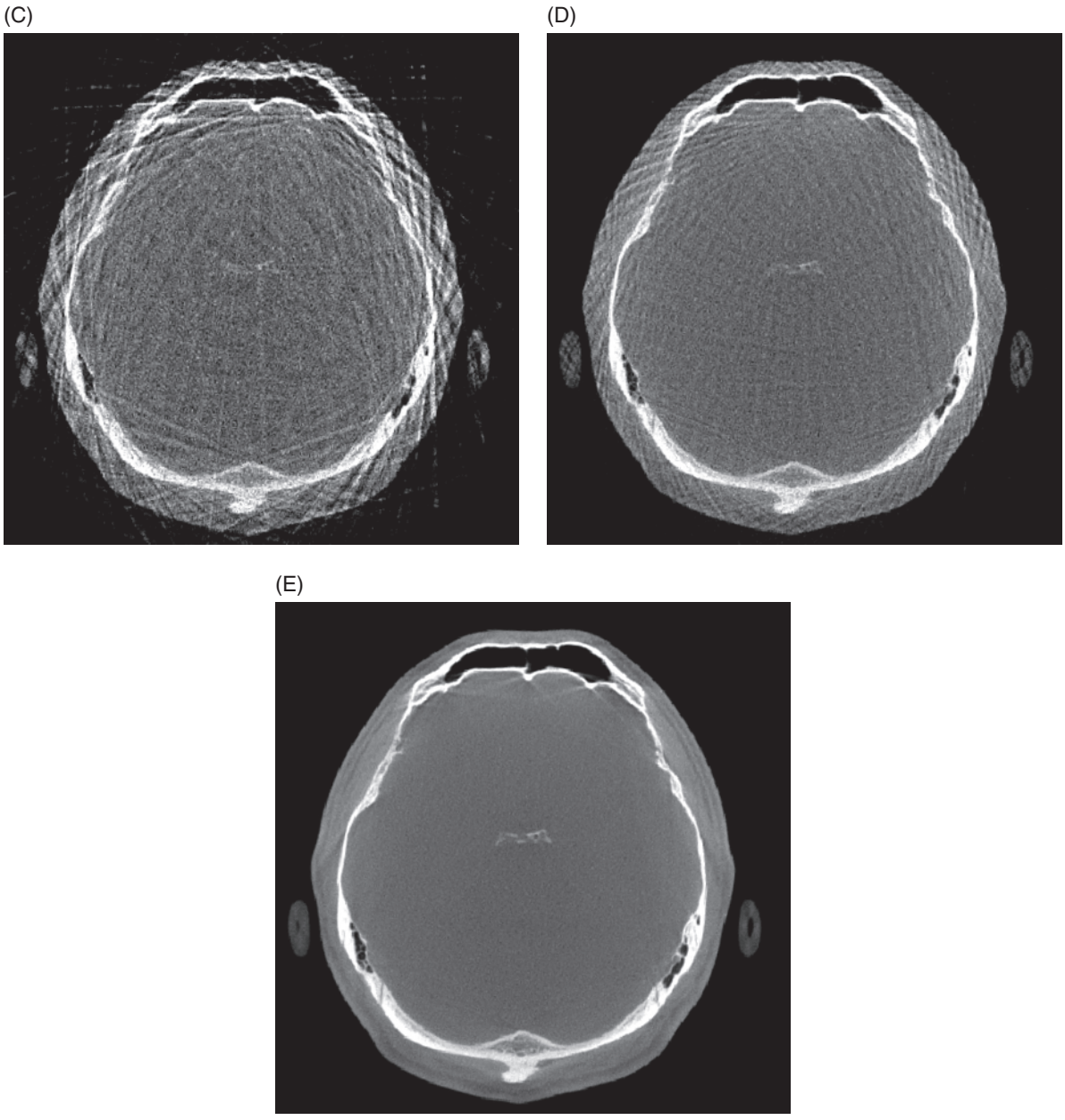

Figure 1.8 (*Continued*) (C) 40 frames. (D) 100 frames. (E) 600 frames.

Much of the acceleration of image reconstruction seen over the years has also been hardware-based. For high-end CT systems, specialized circuit chips known as application-specific integrated circuits (ASICs) have been used in place of software to implement time-consuming reconstruction operations (Wu, 1991). Since the cost of developing such specialized chips can run into millions of dollars, this route has generally been available only to large CT manufacturers. Parallel computing technology has also often been used as an approach to acceleration. Operations like back

projection often consist of tasks that are independent and can be dispatched to several processors working in parallel. For example, the contribution of each X-ray frame to the final image can be computed independently of other frames. Similarly, different collections of slices in the reconstruction grid can be reconstructed in parallel.

Although parallel computing has become increasingly available to smaller manufacturers with the emergence of multicore CPUs, it has taken a particular significant leap forward in recent years with the advent of general purpose graphics processing units (GPGPUs). Essentially, it has been found that the massive parallel computing done by common video game graphics cards can be adapted to a variety of scientific computing problems, including FDK back projection (Vaz, McLin, et al., 2007; Zhao, Hu, et al., 2009). This advance has first of all led to a dramatic speed-up in reconstruction time. Whereas five years ago a typical head CT reconstruction took on the order of several minutes, it can now be performed in approximately 10 seconds. Additionally, the use of GPGPU has greatly cut costs of both the relevant hardware and software engineering work. In terms of hardware, the only equipment required is a video card, costs for which may be as low as a few hundred dollars, thanks to the size of the video gaming industry. The necessary software engineering work has been simplified by the emergence of GPGPU programming languages, such as CUDA and OpenCL (Kirk and Hwu, 2010).

While the FDK reconstruction algorithm is the most common choice for circular-orbit cone beam CT systems, there are limitations to a circular-orbiting CT scanner that appear when the FDK algorithm is applied. Specifically, it is known that a circular-orbiting cone beam camera does not offer complete enough coverage of the object to reliably reconstruct all points in the FOV (or at least not by an algorithm relying on the projection measurements alone). Conditions for a point in 3D space to be recoverable in a given scan geometry are well studied and are given, for example, in Tuy (1983). For circular-orbiting cameras, only points in the plane of the X-ray source satisfy these conditions. Because of this, the accuracy and quality of the reconstructed image gradually deteriorate with distance from the source plane. This is illustrated in Figure 1.9, which shows sagittal views of a computer-generated head phantom and its FDK reconstruction from simulated cone beam CT measurements. Comparing Figure 1.9B to Figure 1.9A, one can clearly see an erroneous drop-off in the image intensity values with distance from the plane of the source, as well as the appearance of streaks and shading artifacts. These so-called cone beam artifacts become more pronounced where the axial cross-sections are less symmetric, for example, in the bony region of the sinuses. It is important to emphasize that artifacts such as these can arise from a number of different causes in actual CT scans, such as scatter and beam hardening (see "Common Image Artifacts"). Here, however, the simulation has not included any such corrupting effects. The artifacts we see here are therefore assuredly and entirely due to the limitations of the circular scan geometry and the FDK algorithm.

In spite of this fundamental weakness in circular cone beam scans, the circular scan geometry has nevertheless been historically favored in the compact CT device industry. This is in part because it simplifies mechanical design. It is also because a range of these artifacts are obscured when the phantom is viewed in a high-contrast bone window (as illustrated in Figure 1.9C and Figure 1.9D), and bone window imaging has been an application of predominant interest for compact CT. On the other hand, this can also be seen as one reason why circular cone beam CT has had difficulty spreading in use from bone imaging to lower contrast imaging applications. In the next section, we discuss iterative reconstruction, which among other things offers possibilities for mitigating the problem of cone beam artifacts.

Iterative reconstruction

Although filtered back projection methods have been commercially implemented for many years, the science has continued to look for improvements using iterative reconstruction methods, both in CT and in other kinds of tomography (Shepp and Vardi, 1982; Lange and Carson, 1984; Erdoğan and Fessler, 1999a). With iterative reconstruction, instead of obtaining a single attenuation map from an explicit reconstruction formula, a sequence of attenuation maps is generated that converges to a final desired reconstructed map. While iterative methods are more computationally

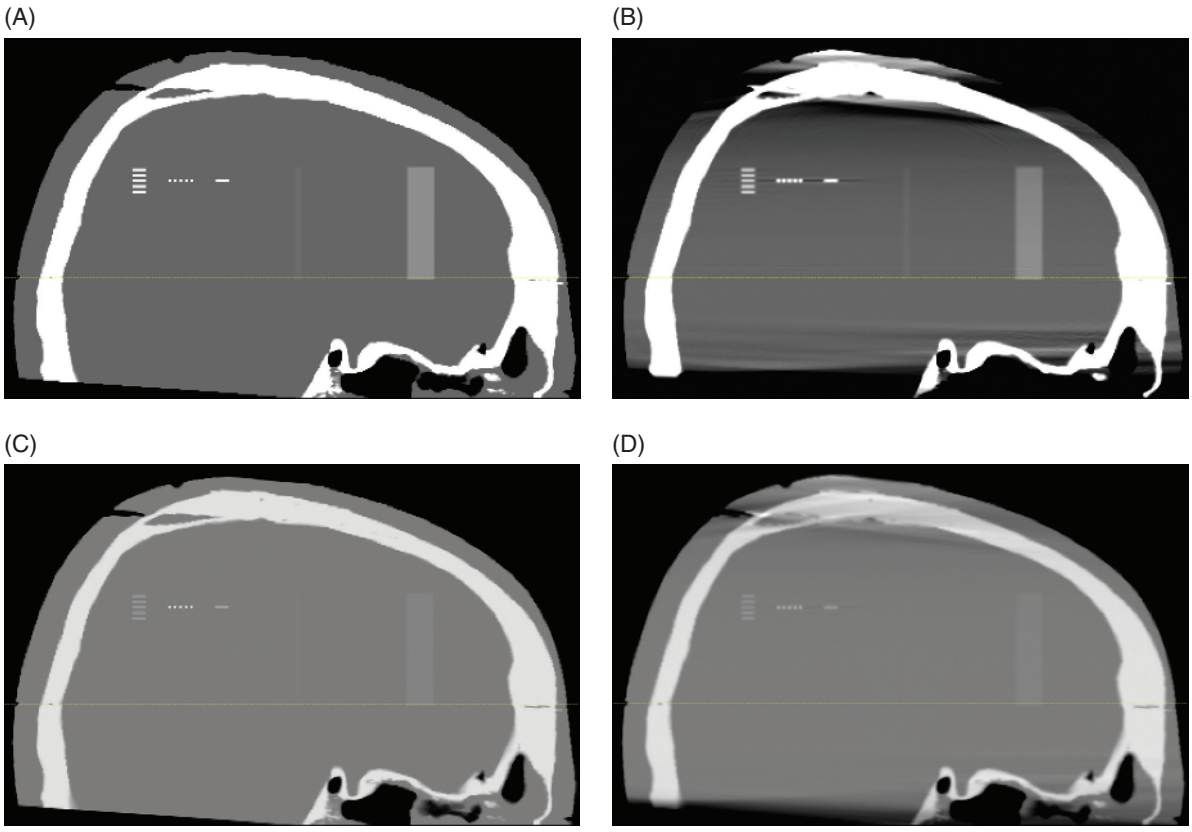

Figure 1.9 Comparison of sagittal views of a computer-generated phantom and its FDK reconstruction in low- and high-contrast viewing window. The dashed line marks the position of the plane of the x-ray source. (A) True phantom, low-contrast window (L/W = 50/200 HU). (B) FDK reconstruction, low-contrast window (L/W = 50/200 HU). (C) True phantom, high-contrast window (L/W = 50/1200 HU). (D) FDK reconstruction, high-contrast window (L/W = 50/1200 HU).

demanding than filtered back projection, they provide a flexible framework for using better models of the CT system, leading to better image quality, sometimes at reduced dose levels. At this writing, iterative methods have also begun to find their way into the commercial CT device market. Notably, the larger medical device companies have commercialized proprietary iterative methods with claims of reducing X-ray dose by several factors without compromising image quality (Freiherr, 2010). Iterative reconstruction software is also marketed by private software vendors such as InstaRecon, Inc., sample results of which are shown subsequently.

In the design of image reconstruction algorithms, there is a trade-off between the amount/accuracy of physical modeling information included in an algorithm, which affects image quality, and the computational expense of the algorithm, which affects reconstruction speed. The previous section overviewed traditional filtered back projection algorithms, which are among the simplest and fastest reconstruction methods. An explicit formula is used to obtain the reconstructed image, and only one pass over the measured X-ray data is required. However, the amount of physical modeling information used in filtered back projection is fairly limited. As an example, filtered back projection ignores statistical variation in the X-ray measurements, leading to higher noise levels in the reconstructed image (or alternatively higher radiation dose levels) than are actually necessary. FBP also ignores the fact that realistic X-ray beams consist of a multitude of X-ray photon energies,

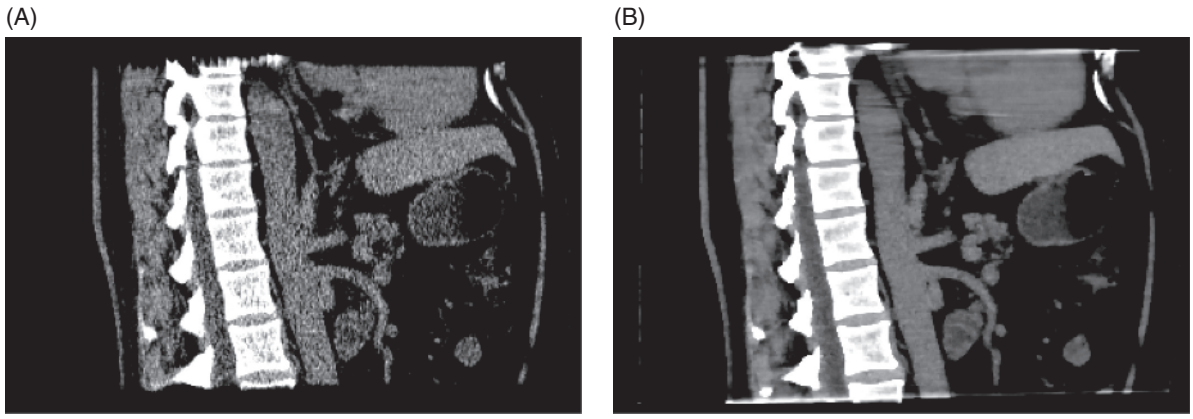

Figure 1.10 Reconstructions of a clinical helical CT scan of the abdomen using (A) filtered back projection and (B) a proprietary iterative algorithm developed by InstaRecon.

approximating the beam instead as a monoenergetic one. This leads to beam hardening artifacts, to be discussed under "Common Image Artifacts." Finally, FBP only incorporates information available in the X-ray measurements, whereas more complicated iterative algorithms can also incorporate a priori knowledge about the characteristics of the patient anatomy. This has important implications for circular-orbit CBCT systems, because for this scanning geometry (see "Conventional Filtered Back Projection" section), the X-ray measurements alone cannot provide enough information to accurately reconstruct the object at all points in the field of view. The FDK algorithm, a variation of FBP specific to circular-orbit systems, produces cone beam artifacts, as a result.

The desire to improve image quality has led many researchers over the years to propose reconstruction algorithms based on more detailed and complicated physical models of CT systems. These more complicated models lead to reconstruction equations that have no explicit solution. Instead, the solution must be obtained by iterative computation, in which a sequence of images is generated that gradually converges to the solution. Generally speaking, every iteration of an iterative reconstruction algorithm tends to have a computational cost comparable to an FBP reconstruction. This extra computation puts a significant price tag on the image quality improvements that iterative reconstruction proposes to bring, a price tag that delayed the clinical acceptability of these methods for many years. Nevertheless, the advantages of iterative reconstruction over filtered back projection are readily demonstrated. Some relevant illustrations are provided in Figure 1.10, Figure 1.11, and Figure 1.12.

Figure 1.10A and Figure 1.10B show a performance comparison of a proprietary iterative algorithm developed by InstaRecon with filtered back projection for a clinical abdominal scan. This particular scan was acquired using a conventional helical CT system, and so the filtered back projection algorithm used was not cone beam FDK. The iterative algorithm achieves reduced image noise and hence more uniform images. Furthermore, since image noise generally trades off with X-ray exposure, noise-reducing iterative algorithms such as these also allow one to scan with reduced X-ray dose, while achieving the same noise levels in the reconstructed image as conventional filtered back projection. Figure 1.11A and Figure 1.11B show a similar comparison for simulated CT measurements of a phantom commonly used to measure low-contrast imaging performance. One sees how the iterative algorithm improves the detectability of low-contrast objects as compared to filtered back projection.

Figure 1.12A and Figure 1.12B show iterative reconstructions of the same computer-generated CBCT phantom scan as in Figure 1.9. This reconstruction algorithm incorporates prior information about the piece-wise smooth structure of the patient anatomy. Reconstruction algorithms

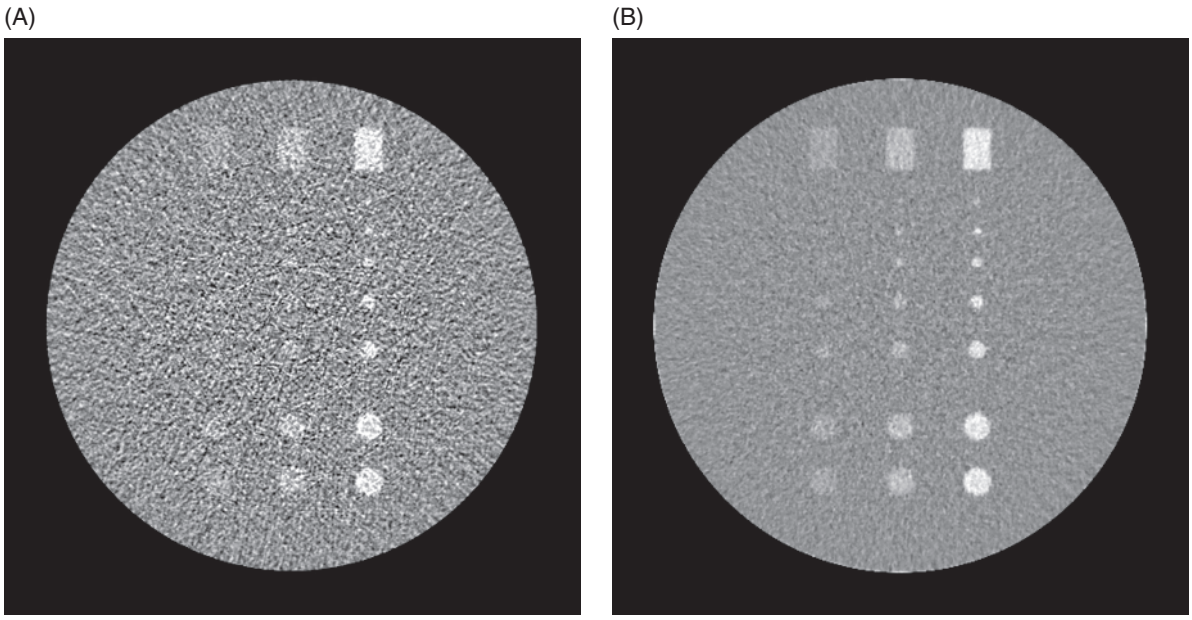

Figure 1.11 Reconstructions of a simulated CBCT scan of a CIRS061 contrast phantom using (A) filtered back projection and (B) a proprietary iterative algorithm developed by InstaRecon.

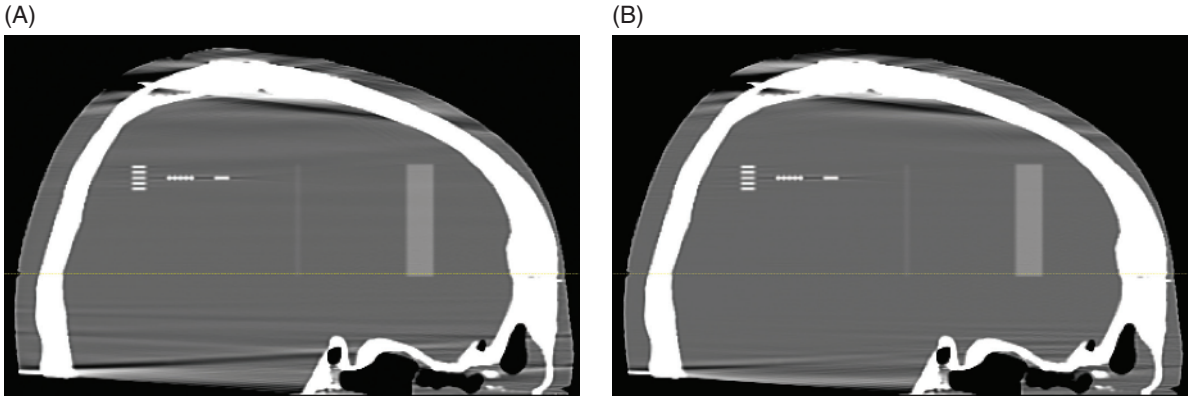

Figure 1.12 Sagittal views of a computer-generated phantom reconstructed using a rudimentary iterative algorithm in a low-contrast viewing window (L/W = 50/200 HU). (A) Result after 30 iterations. (B) Result after 300 iterations.

that incorporate such information (Sukovic and Clinthorne, 2000) are abundant in the medical imaging literature. The reconstruction algorithm used here was more rudimentary than InstaRecon's algorithm. Among other things, it has not been optimized for speed and it takes many more iterations to converge. However, it was sufficient to show how adding prior smoothness information can mitigate cone beam artifacts. Figure 1.12 shows that the intensity values in the region of the sinuses are much closer to their true value as compared to the FDK results in Figure 1.9B. This occurs because the addition of prior information about anatomical smoothness compensates for the geometric incompleteness of the circular X-ray camera orbit.

Although the image quality benefits of iterative algorithms have been known for many years, it has

only recently become possible to run at sufficient speed to make them clinically acceptable for CT imaging. Computing hardware improvements over the years, such as GPGPU discussed earlier, have contributed to reducing computation time per iteration. Additionally, much medical imaging research has been devoted to finding iterative reconstruction algorithms requiring as few as possible iterations to converge (Kamphuis and Beekman, 1998; Erdoğan and Fessler, 1999b; Ahn, Fessler, et al. 2006).

Imaging performance

This section discusses several quantitative measures of image quality that are commonly used to assess the performance of a CT device, namely noise performance, low-contrast detectability, and spatial resolution. CT manufacturers will typically report such quality measurements in the user manuals issued with their devices. Typically also, manufacturers provide customers equipment to repeat these measurements and specify in the user manual how reproducible the measurements should be. For CT manufacturers in the United States, providing this information is legally required by the Code of Federal Regulations (21 CFR 1020.33).

Image noise

The term *measurement noise* refers to random variations in CT measurements. *Image noise* refers to the ensuing effect of these variations on the reconstructed image. In a CT scan, there are several sources of measurement noise that make the measurements not precisely repeatable. When X-rays are fired through a patient along a certain straight-line path, there is randomness in the number of photons that will penetrate through the object to interact with the detector. There is also randomness in the number of photons that, after penetrating the object, will successfully interact with the X-ray detector panel to produce a signal. Finally, there are also elements of random fluctuation in the detector electronics itself, independent of the object and the X-ray source.

Measurement noise leads to sharp discontinuities among the measured values of neighboring detector pixels. When the X-ray measurements are put through the image reconstruction process, the reconstructed CT volume will exhibit correspondingly sharp discontinuities among neighboring voxel values that would otherwise be uniform or gradually varying. This is the visual manifestation of image noise. A common way to measure image noise is to compute the standard deviation of some region of voxels in a phantom of some uniform material (as in Figure 1.4, for example). As mentioned in the "Overview of Image Processing and Display," most CT image viewing software provides this capability. In manuals for a CT device, the noise standard deviation will often be reported as a fraction of the attenuation of water.

CT system engineers make design choices to control noise but must take certain trade-offs into account. Measurement noise can be reduced, for example, by increasing X-ray exposure to the patient, although health concerns place obvious limits on doing so. Certain types of detector panels have better photon detection efficiency than others, giving better resistance to noise. However, such detectors are also more expensive and lead to increased system cost. Other methods of reducing noise involve configuring the X-ray detection and image reconstruction process in a certain way, although these methods entail trade-offs in image resolution. For example, most detector panels allow one to combine neighboring detector pixels to form larger pixels. This "binning" of pixels effectively averages together the signal values that would be measured by the smaller pixels separately and reduces noise. However, projection sampling fineness, and hence resolution, are also reduced as a trade-off. Similarly, the reconstruction software can be designed to include smoothing operations. As mentioned previously, filtered back projection methods include smoothing in the filtering step, while iterative reconstruction methods can enforce image smoothness using a priori anatomical information. These smoothing methods reduce noise but can also blur anatomical tissue borders as a side effect, and so resolution is again sacrificed. Reconstruction algorithms are often compared based on how favorably noise trades off with spatial resolution.

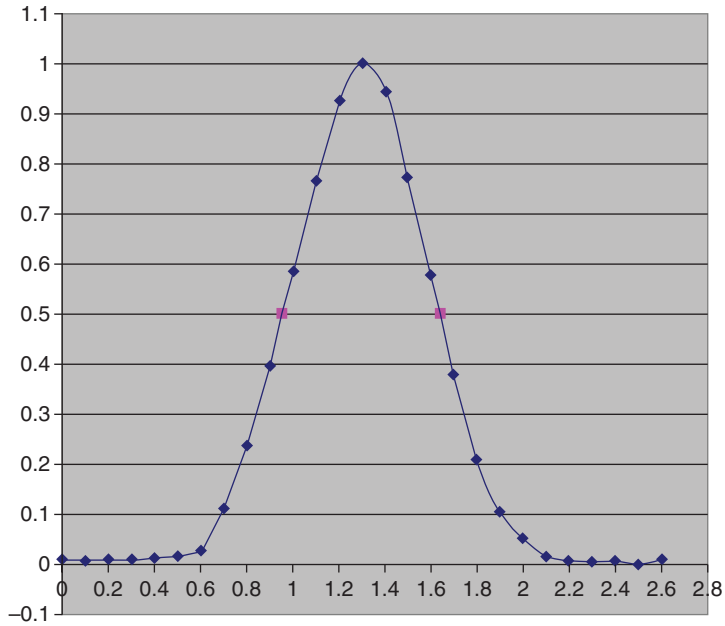

Figure 1.13 Example of a slice sensitivity profile (SSP) illustrated with data from the xCAT-ENT, a commercial mobile cone beam CT scanner for sinus imaging. The profile is plotted on a horizontal axis in units of millimeters.

Spatial resolution

Spatial resolution refers to how well small or closely spaced objects are visualized in an image. Spatial resolution in a cone beam CT system is partly limited by the size of the image voxels used for reconstruction. However, resolution is further limited by various sources of system blur. As discussed in the previous section, certain sources of blur arise as a side effect of various engineering measures taken to reduce image noise. Other sources of blur arise from the physics of the X-ray detection process. Detector glare is an effect whereby X-ray photons striking the detector induce a scattering event that causes a signal to be detected in several neighboring pixels. This leads to a blurring of the projection views and an ensuing blur in the reconstructed image. A similar effect is detector lag, in which the signal detected in one X-ray shot fails to dissipate before the next X-ray shot is taken. This has the effect of blurring together adjacent X-ray shots. Finally, imperfect modeling of the CT system geometry in the reconstruction process can also blur the image. For example, no cone beam CT system produces a perfectly cone-shaped X-ray beam because X-rays are emitted from different points on the surface of the source, rather than from a single apex point. However, this effect is commonly ignored by the reconstruction software, at the expense of spatial resolution.

In conventional helical fan beam CT systems, the amount of blur along the axis of the scanner has historically been significantly different than the blur within an axial slice. This difference has led to common practices, and in some cases regulations, for CT manufacturers to report separate measurements of axial and in-plane spatial resolution. With the advent of cone beam systems, the difference in axial versus in-plane resolution has greatly diminished, but laws designed for helical fan beam systems are so well established that they are still applied to CBCT. To measure spatial resolution axially, an object such as a wire or bead, whose cross-section along the scanner axis is narrow and pointlike, is imaged. Due to blur effects, the cross-section in the image will have a smeared, lobelike profile, such as that shown in Figure 1.12. The amount of blur is reported on a *slice sensitivity profile* such as the one in Figure 1.13. The width of

this profile at half its peak value is known as the *nominal tomographic section thickness*.

To measure in-plane spatial resolution, it is traditional to report the *modulation transfer function* (MTF). An MTF is a graph showing how the imaged contrast of densely clustered objects decreases, as a result of system blur, with the clustering density. As a result of blur effects, the intensity of small or narrow objects is diluted with background material in the image, thereby lowering their apparent contrast. Since objects must be of decreasing size to be clustered more densely, an accompanying decrease in contrast with density is typically observed. This is illustrated in Figure 1.14A, which shows a series of progressively denser line pair targets, with the density expressed in line pairs per centimeter (lp/cm). One can see how not only the separation between the more densely spaced line pairs diminishes as a result of blur, but also their percent contrast with the background medium. By measuring the percent contrast of line pair phantoms, one can plot contrast versus line pair density, which is how MTF plots are often expressed. MTFs can also be obtained more indirectly by measuring an in-plane blur profile, similar to the slice sensitivity profile (Boone, 2001). The MTF plots in Figure 1.14B were obtained in such a manner. They show the MTFs for two imaging modes of a commercial ear-nose-throat scanner. The temporal bone mode has a more slowly decreasing MTF, indicative of less blur and higher spatial resolution, than the sinus mode. This is typical, due to the higher resolution needs of temporal bone imaging tasks.

Low-contrast detectability

Low-contrast detectability is a performance parameter of CT systems that measures its overall ability to resolve small differences in intensity between objects. To test low-contrast detectability in a CT system, phantoms such as that in Figure 1.11, containing low-contrast targets of a range of sizes, are often used.

As discussed in the previous section, system blur reduces the contrast of small objects. However, there are other contrast-limiting effects in a CBCT system that can affect the visibility of large objects as well. One contrast-limiting effect in CT systems is the energy spectrum of the X-ray source. At lower average photon energies, obtained by lowering the X-ray source voltage, attenuation differences among different materials generally increase, leading to better contrast. The engineering trade-off in lowering source energy, however, is that the ability of X-ray photons to penetrate the CT subject is reduced, leading to higher noise and photon starvation artifacts. Contrast is also limited by certain features in the electronics of the X-ray detector. When detected X-rays are converted from analog to digital signals, information about tissue contrast is somewhat degraded. This degradation can be reduced by using A/D converters which digitize signals more finely, but the trade-off in doing so is an increase in the cost of the detector panel, and hence the overall system.

Common image artifacts

Image artifacts are visible patterns in an image arising from systematic errors in the reconstruction process. Common kinds of artifacts include streaks and nonuniformity trends, such as in Figure 1.15A. For circular-orbiting CT systems, ring artifacts such as in Figure 1.16A are also commonly encountered. Current use of compact CT systems is often tolerant to artifacts, since bone window viewing of CT images is still very prevalent, and many artifacts are obscured in the bone window. An understanding of artifacts and their causes can still be important, however, for several reasons. First, there are exceptions where artifacts are severe enough to appear even in the bone window viewing applications. When scanning very bony anatomy, for example in dental or skull base imaging, very strong streak artifacts can be present. Artifacts can also be a sign that a CT system is in need of maintenance. Strong ring artifacts can appear when the system is in need of recalibration, for instance. Finally, as practitioners expand their use of compact CT to low-contrast soft tissue imaging applications, the influence of artifacts becomes more noticeable in the less forgiving low-contrast viewing windows. Means of suppressing artifacts will be important to extending compact CT to these applications.

Causes of artifacts can be either advertent or inadvertent. Inadvertent causes include inaccuracies in the calibration of the CT system. When a CT system

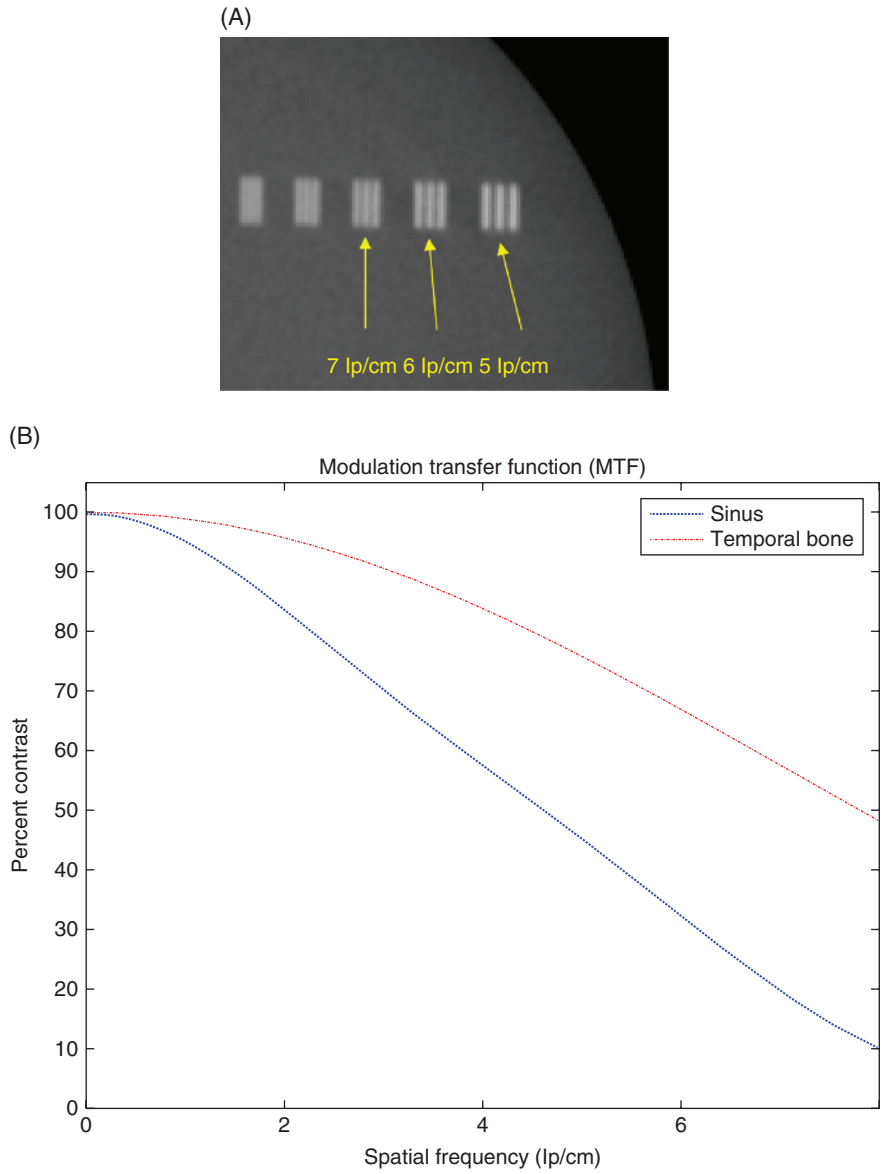

Figure 1.14 Concepts of in-plane resolution measurement illustrated with data from the MiniCAT, a commercial cone beam CT scanner for sinus and temporal bone imaging. (A) Reconstructed image of a phantom containing line pair targets of different densities. lp/cm = line pairs per centimeter. (B) Modulation transfer function for the MiniCAT's sinus and temporal bone scan protocols.

is first installed, and possibly periodically thereafter, certain physical properties of the system must be measured through a calibration procedure. The physical properties to be calibrated are ones that cannot be precisely controlled by the manufacturer, or that may drift over the lifetime of the machine in some uncontrollable way. In the section "Conventional Filtered Back Projection," for example, it was discussed how certain detector pixel parameters must be calibrated periodically using air scans and blank scans. These kinds of calibrated quantities serve as input to the image

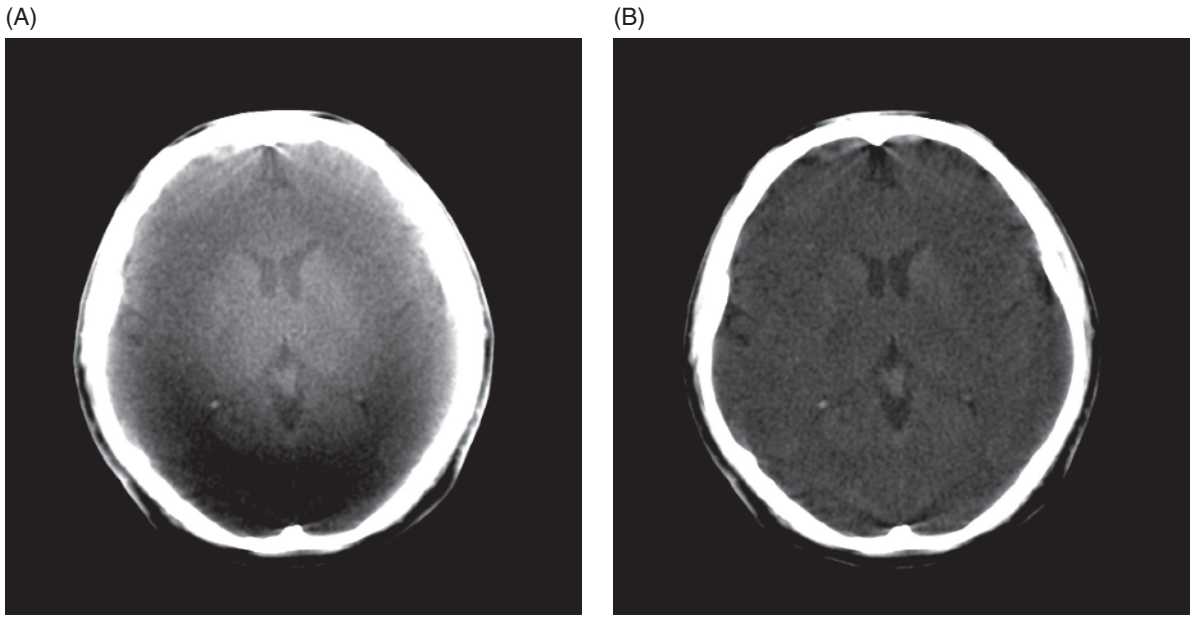

Figure 1.15 (A) Illustration of streaks and nonuniformity artifacts in an axial slice of a low-contrast CBCT scan. (B) The same slice after a postcorrection method is applied.

reconstruction process, which uses them to model system behavior. Inaccuracies in the calibration create disagreement between the true physical X-ray measurements and the mathematical model used by the reconstruction software, resulting in image artifacts. In circular-scanning CBCT systems, inaccuracies in pixel sensitivities and offsets are a typical cause of tree trunk–like ring artifacts, like those shown in Figure 1.16A. Miscalibration of a given pixel will introduce errors in how that pixel's measurement is processed in every X-ray shot. The repetition of these measurement errors throughout the circular orbit of the X-ray camera leads to circularly symmetric artifact patterns in the image, thus showing as rings.

Artifacts can also result from deliberate mathematical errors and approximations made by the reconstruction algorithm to simplify computation. As an example, in the "Conventional Filtered Back Projection" section, it was discussed how cone beam artifacts are an engineering trade-off to the mechanical simplicity of a circular-orbiting CT camera, as well as to the computational simplicity of the FDK reconstruction algorithm. Similar kinds of trade-offs have historically been made in the treatment of other corrupting physical effects such as beam hardening and scatter. Beam hardening is a physical effect whereby the average energy content of an X-ray beam gradually increases as the photons in the beam pass through an object. This occurs because lower energy X-ray photons have a lower probability than higher energy photons of passing through the object unattenuated, and are progressively sifted out of the beam. Scatter is an effect whereby some X-ray photons traveling through the CT subject are deflected from a straight-line path, due to interaction with matter, and generate signal in the wrong detector pixels. When ignored by the reconstruction process, both beam hardening and scatter can contribute to coarse nonuniformity artifacts, such those as shown in Figure 1.15A. Moreover, when scanning bony, asymmetric anatomy, beam hardening and scatter can contribute to streak artifacts, also shown in the figure. Streaks result whenever certain particular X-ray shots contain much more measurement errors than at other positions of the X-ray camera. Beam hardening and scatter effects are a common cause of such errors because their effect varies strongly with the thickness and density of tissue through which the X-ray beam passes.

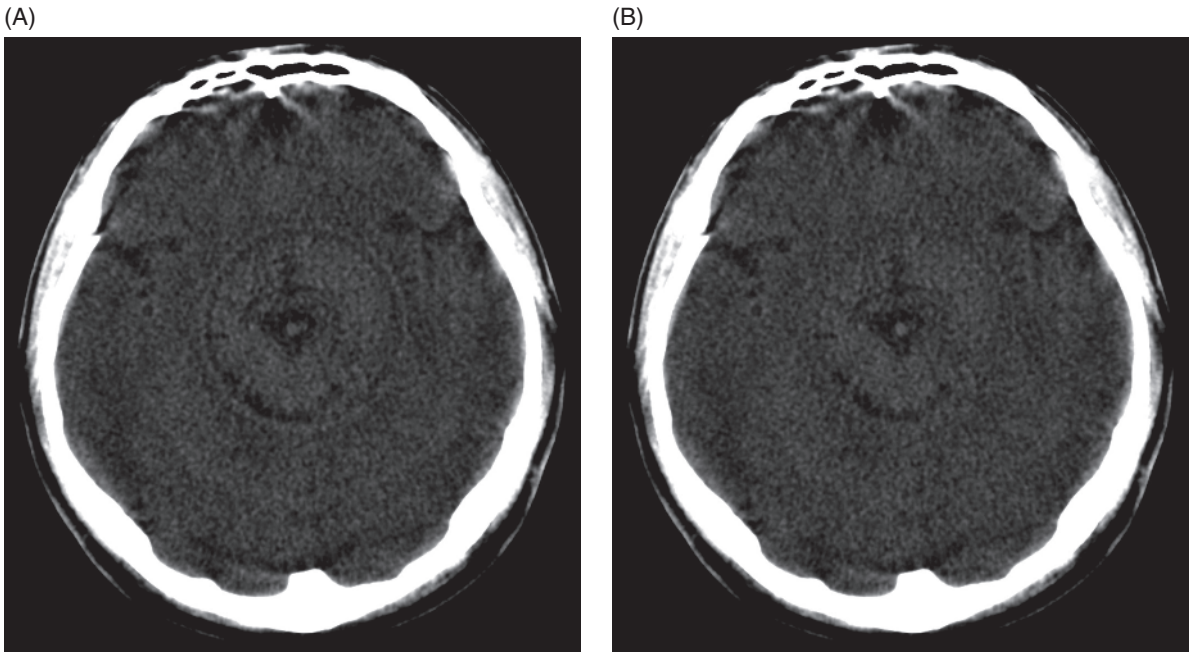

Figure 1.16 (A) Illustration of ring artifacts in an axial slice of a low-contrast CBCT scan. (B) The same slice after a ring correction method is applied.

For asymmetric patient anatomy, these in turn vary strongly with the position of the X-ray camera relative to the patient.

Beam hardening and scatter have historically been computationally expensive to handle in the image reconstruction process in a mathematically precise way, which means that in practice they are either ignored or corrected using computationally cheaper compromises. One of the more mathematically rigorous ways of dealing with beam hardening, for example, is to use an image reconstruction algorithm that models the energy variation of the beam (Elbakri and Fessler, 2002; Elbakri and Fessler, 2003). However, reconstruction algorithms with this level of modeling generally require iterative methods, and only in recent years has computing technology become fast enough to consider using such methods clinically. Similarly, scientific literature has proposed very accurate scatter modeling and correction approaches (Zbijewski and Beekman, 2006). However, achieving clinically viable computation time remains a challenge with these methods.

In situations where rigorous image reconstruction is too expensive computationally, but where the resulting artifacts cannot be tolerated, commercial systems will often remove artifacts from the reconstructed image using fast postcorrection methods. These methods are often proprietary, and therefore it is hard to comment authoritatively on how they work for different CT vendors. However, a variety of postcorrection methods have been proposed in public-domain scientific literature. It is likely that at least some methods used commercially are derived from these. The degree of mathematical or physical modeling rigor on which postcorrection methods are based can vary greatly. There is therefore much ongoing debate in scientific literature over their limitations, as compared to their more computationally expensive, mathematically rigorous alternatives. However, postcorrection methods have certainly proven effective enough to make them popular compromises. Figure 1.15B, for example, demonstrates the reduction of streak and nonuniformity artifacts using a combination of postprocessing approach (Zbijewski and Stayman, 2009; Hsieh, Molthen, et al., 2000). Figure 1.16B demonstrates the reduction of ring artifacts using a postcorrection method (Sijbers and Postnov, 2004).

References

Ahn, S., Fessler, J.A., et al. (2006). Convergent incremental optimization transfer algorithms: Application to tomography. *IEEE Transactions on Medical Imaging* 25(3): 283–96.

Basu, S., and Bresler, Y. (2001). Error analysis and performance optimization of fast hierarchical backprojection algorithms. *IEEE Trans Im Proc* 10(7): 1103–17.

Boone, J.M. (2001). Determination of the presampled MTF in computed tomography. *Med Phys* 28(3): 356–60.

De Man, B., and Basu, S. (2004). Distance-driven projection and backprojection in three dimensions. *Phys Med Biol* 49(11): 2463–75.

Elbakri, I.A., and Fessler, J.A. (2002). Statistical image reconstruction for polyenergetic X-ray computed tomography. *IEEE Transactions on Medical Imaging* 21: 89–99.

Elbakri, I.A., and Fessler, J.A. (2003). Segmentation-free statistical image reconstruction for polyenergetic X-ray computed tomography with experimental validation. *Phys Med Biol* 48(15): 2543–78.

Erdoğan, H., and Fessler, J.A. (1999a). Monotonic algorithms for transmission tomography. *IEEE Transactions on Medical Imaging* 18(9): 801–14.

Erdoğan, H., and Fessler, J.A. (1999b). Ordered subsets algorithms for transmission tomography. *Phys Med Biol* 44(11): 2835–51.

Feldkamp, L.A., and Davis, L.C. (1984). Practical cone-beam algorithm. *J Opt Soc Amer* 1: 612–19.

Freiherr, G. (2010). Iterative reconstruction cuts CT dose without harming image quality. *Diagnostic Imaging* 32(11). Available at www.diagnosticimaging.com.

Hsieh, J., Molthen, R.C., et al. (2000). An iterative approach to the beam hardening correction in cone beam CT. *Med Phys* 27(1): 23–9.

Kamphuis, C., and Beekman, F.J. (1998). Accelerated iterative transmission CT reconstruction using an ordered subsets convex algorithm. *IEEE Transactions on Medical Imaging* 17(6): 1001–5.

Kirk, D.B., and Hwu, W.W. (2010). *Programming Massively Parallel Processors: A Hands-on Approach*. Morgan Kaufman.

Lange, K., and Carson, R. (1984). EM reconstruction algorithms for emission and transmission tomography. *J Comp Assisted Tomo* 8(2): 306–16.

Shepp, L.A., and Vardi, Y. (1982). Maximum likelihood reconstruction for emission tomography. *IEEE Trans Med Imag* 1(2): 113–22.

Sijbers, J., and Postnov, A. (2004). Reduction of ring artefacts in high resolution micro-CT reconstructions. *Phys Med Biol* 49(14): N247-54.

Sukovic, P., and Clinthorne, N.H. (2000). Penalized weighted least-squares as a metal streak artifacts removal technique in computed tomography. *Proc IEEE Nuc Sci Symp Med Im Conf*.

Tuy, H.K. (1983). An inversion formula for cone-beam reconstruction. *SIAM J Appl Math* 43(3): 546–52.

Vaz, M.A., McLin, M., et al. (2007). Current and next generation GPUs for accelerating CT reconstruction: Quality, performance, and tuning. *Proc Intl Mtg on Fully 3D Image Recon in Rad and Nuc Med*.

Wu, M. A. (1991). ASIC applications in computed tomography systems. Fourth Annual IEEE International ASIC Conference and Exhibit.

Zbijewski, W., and Beekman, F.J. (2006). Efficient Monte Carlo based scatter artifact reduction in cone-beam micro-CT. *IEEE Trans Med Imag* 25(7): 817–27.

Zbijewski, W., and Stayman, J.W. (2009). Volumetric soft tissue brain imaging on xCAT: A mobile flat-panel x-ray CT system. *Proc SPIE 7258, Medical Imaging 2009: Phys Med Im*.

Zhao, X., Hu, J.J., et al. (2009). GPU-based 3D cone-beam CT image reconstruction for large data volume. *Int J Biomed Imaging* 2009: 149079.