Radiology

Glenn S. Slavin, PhD
David A. Bluemke, MD, PhD

Published online before print
10.1148/radiol.2342031990
Radiology 2005; 234:330–338

Abbreviations:
DFT = discrete Fourier transform
DTFT = discrete-time Fourier transform
FOV = field of view
TR = repetition time
VPS = views per segment

[1] From the Applied Science Laboratory, GE Medical Systems, Baltimore, Md (G.S.S.); and Russell H. Morgan Department of Radiology and Radiological Science, Johns Hopkins University School of Medicine, MRI Room 143 (Nelson Basement), 600 N Wolfe St, Baltimore, MD 21287 (G.S.S., D.A.B.). Received December 10, 2003; revision requested February 19, 2004; revision received April 12; accepted May 24. **Address correspondence to** D.A.B. (e-mail: *dbluemke@jhmi.edu*).

Spatial and Temporal Resolution in Cardiovascular MR Imaging: Review and Recommendations[1]

Because of the nature of digital imaging, the number of pixels in a reconstructed image is often unrelated to the actual spatial resolution of the image. Similarly, the number of reconstructed frames of a dynamic or cine examination can be unrelated to the acquired temporal resolution. These discrepancies can result in misinterpretations and inaccuracies when image resolution is reported in the literature. The goal of this report is to clarify the differences between acquired and displayed resolution, both spatial and temporal, in magnetic resonance imaging. The effects of imaging parameters on acquired resolution are discussed, as are the mathematic effects of the reconstruction process on the displayed resolution of the resulting image. Finally, recommendations to authors are offered to promote accurate and unambiguous reporting of spatiotemporal resolution in the literature.

With the advent of increasingly more sophisticated magnetic resonance (MR) and computed tomographic (CT) imagers, applications in cardiovascular imaging have increased rapidly. A question we are often asked is just how "good" these images are in relationship to a standard of reference (generally, conventional angiography). This is frequently translated into a question of spatial or temporal resolution, which is the ability to discriminate between two points in space and time, respectively. For example, in nearly all journal articles regarding clinical applications of MR or CT angiography, resolution is usually taken to be related to the size of a pixel (or voxel, in three dimensions).

Unfortunately, with digital imaging techniques, there may be little or no relationship between pixel size and inherent image resolution. An everyday example of this is a video camcorder that promises a digital "zoom" factor of 120 times. Clearly, this zoom factor is related only to digital pixel enlargement and has little to do with picture resolution or image quality. Similar "empty" resolution can now be achieved with three-dimensional CT and MR acquisitions by using sophisticated reconstruction algorithms; sections are reconstructed routinely at fine intervals that are smaller than the acquired section thickness. Both authors and readers of journal articles may mistakenly interpret this as improved spatial resolution.

As multi–detector row CT and MR imaging may come to replace diagnostic angiography for certain applications, it is useful to review the principles of image resolution, to describe pitfalls in stating resolution, and to suggest standard means for researchers who report resolution that will ultimately aid journal authors, reviewers, and critical readers in comparing different imaging modalities. The topics that follow will be primarily focused on MR imaging, but relevance to multi–detector row CT will also be addressed.

BACKGROUND

Case example: A resident is creating a protocol for an MR examination, for which he realizes that his attending physician requires a "high-resolution" image. Instead of the usual 256 × 256 (frequency × phase) image, he asks the technologist to boost the resolution to 512 × 512. Recognizing that doubling the phase-encoding steps from 256 to 512 will double the imaging time, the technologist adeptly suggests an alternative to the resident. The technologist can instead use the zero-filling option on the imager to increase

the image size to 512 × 512 without an increase in imaging time. At the time of image interpretation, the attending physician understands that the technologist has outwitted the resident, and the inherent resolution of the zero-filled image is still 256 × 256.

Understanding the effects of digital interpolation schemes in the above example is complex but is aided by dividing image acquisition into three general steps: examination prescription, MR signal reception and sampling, and image reconstruction. We will discuss the role that image resolution, both spatial and temporal, plays in each of these steps. Before we do this, however, we will review the relationship between the mathematic domains of the object, the image, and the MR signal (or, similarly, the raw data). These are referred to as the object domain, the image domain, and the frequency domain, respectively. The image and object domains are similar, with the discrete image domain being a sampled version of the continuous object domain. The image and object domains are related to the frequency domain by means of Fourier transformation.

To begin, the MR signal *s(t)* as a function of time (where *s* represents signal and *t* represents time) can be written as (1)

$$s(t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F_0(x, y)e^{-2\pi(k_x(t)\,x+k_y(t)\,y)}\,dxdy, \quad (1)$$

where $F_0(x,y)$ represents the object, *x* and *y* are the object-domain variables in centimeters, k_x and k_y are the frequency-domain variables (ie, the spatial frequencies in the frequency- and phase-encoding directions, respectively) in 1/centimeters. The right side of Equation (1) can be recognized as the two-dimensional Fourier transform of $F_0(x,y)$ (2). We can therefore write *s(t)* more simply as

$$s(t) = f\big(k_x(t),\, k_y(t)\big) = \mathrm{FT}[F_0(x, y)]. \quad (2)$$

From this, we can see that the received MR signal is simply the Fourier transform (FT) of the object at each point in time. For completeness, we have indicated the time dependence of k_x and k_y, which is attributable to the time-dependent imaging gradients G_x and G_y:

$$k_x(t) = \frac{\gamma}{2\pi}\int_0^t G_x(t')dt' \quad (3)$$

and

$$k_y(t) = \frac{\gamma}{2\pi}\int_0^t G_y(t')dt', \quad (4)$$

where γ is the gyromagnetic ratio.

In the expressions for *s(t),* the variables are continuous and infinite in extent. To represent the imaged object in digital form, *s(t)*—or equivalently $f\big(k_x(t),k_y(t)\big)$—must be sampled a finite number of times. We refer to the sampled form of $f\big(k_x(t),k_y(t)\big)$ as the raw data. This sampling results in tissue volumes (or voxels) in the object domain and pixels in the image domain. Later, we will encounter the unexpected result that what appears to be a one-to-one correspondence between the number of voxels and the number of pixels does not yield the expected spatial resolution on the image.

SPATIAL RESOLUTION

During examination prescription, the desired spatial resolution is specified by selecting the field of view (FOV) and the number of frequency- and phase-encoding points. These selections determine the degree (ie, the duration and frequency) of sampling of the MR signal, which in turn specifies the extent of $k_x(t) - k_y(t)$ − space (ie, k-space) that is acquired (1,3). The extent of k-space is specified by

$$k_{x\max} = \frac{1}{2\Delta x} \quad (5)$$

and

$$k_{y\max} = \frac{1}{2\Delta y}, \quad (6)$$

where Δ*x* and Δ*y* are the prescribed acquired (or nominal) spatial resolutions in the frequency- and phase-encoding directions, respectively. These are determined by the FOV and the number of frequency- and phase-encoding points (N_x and N_y) selected on the imager.

$$\Delta x = \frac{\mathrm{FOV}_x}{N_x} \quad (7)$$

and

$$\Delta y = \frac{\mathrm{FOV}_y}{N_y}. \quad (8)$$

The sampling interval between k-space points is determined by the FOV.

$$\Delta k_x = \frac{1}{\mathrm{FOV}_x} \quad (9)$$

and

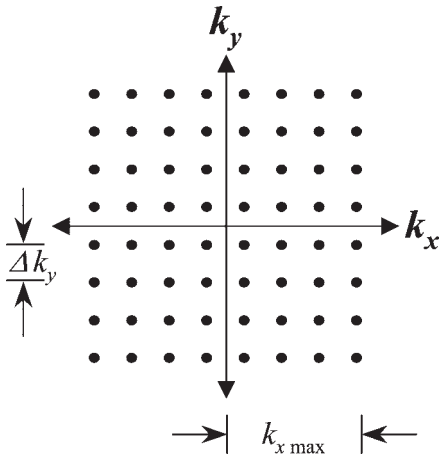

Figure 1. Schematic of the k-space representation of MR raw data.

$$\Delta k_y = \frac{1}{\mathrm{FOV}_y}. \quad (10)$$

A graphic representation of k-space is shown in Figure 1.

One way to reconstruct the raw data is use of the discrete-time Fourier transform (DTFT) (4). The one- and two-dimensional forms of the DTFT are shown in Equations (11) and (12), respectively (see Appendix).

$$F(x) = \sum_{\hat{k}_x=-\infty}^{+\infty} f[\hat{k}_x]e^{-(2\pi i/\mathrm{FOV}_x)\hat{k}_x x} \quad (11)$$

for all *x*.

$$F(x, y) = \sum_{\hat{k}_x=-\infty}^{+\infty}\sum_{\hat{k}_y=-\infty}^{+\infty} f[\hat{k}_x, \hat{k}_y]e^{-2\pi i((\hat{k}_x x/\mathrm{FOV}_x)+(\hat{k}_y y/\mathrm{FOV}_y))} \quad (12)$$

for all *x* and *y*.

Here, $f[\hat{k}_x,\hat{k}_y]$ are the sampled points of $f\big(k_x(t),k_y(t)\big)$, and $\hat{k}_x$ and $\hat{k}_y$ are the discrete k-space indices (ie, $\hat{k}_x,\hat{k}_y = 0, \pm1, \pm2, \pm3, \ldots$). The DTFT is performed for discrete sequences, which are the k-space data $f[\hat{k}_x,\hat{k}_y]$ in this case. $F(x,y)$ represents the imaged object as a continuous periodic function of position. Because $F(x,y)$ is continuous, it effectively has infinite spatial resolution, since *x* and *y* can take on any value, no matter how small.

Consequently, the DTFT of the raw data is the exact representation of the object as imaged (ie, at the acquired spatial resolution). It is not, however, an exact representation of the object. To achieve a truly exact representation, sampling would have to be performed for an infi-

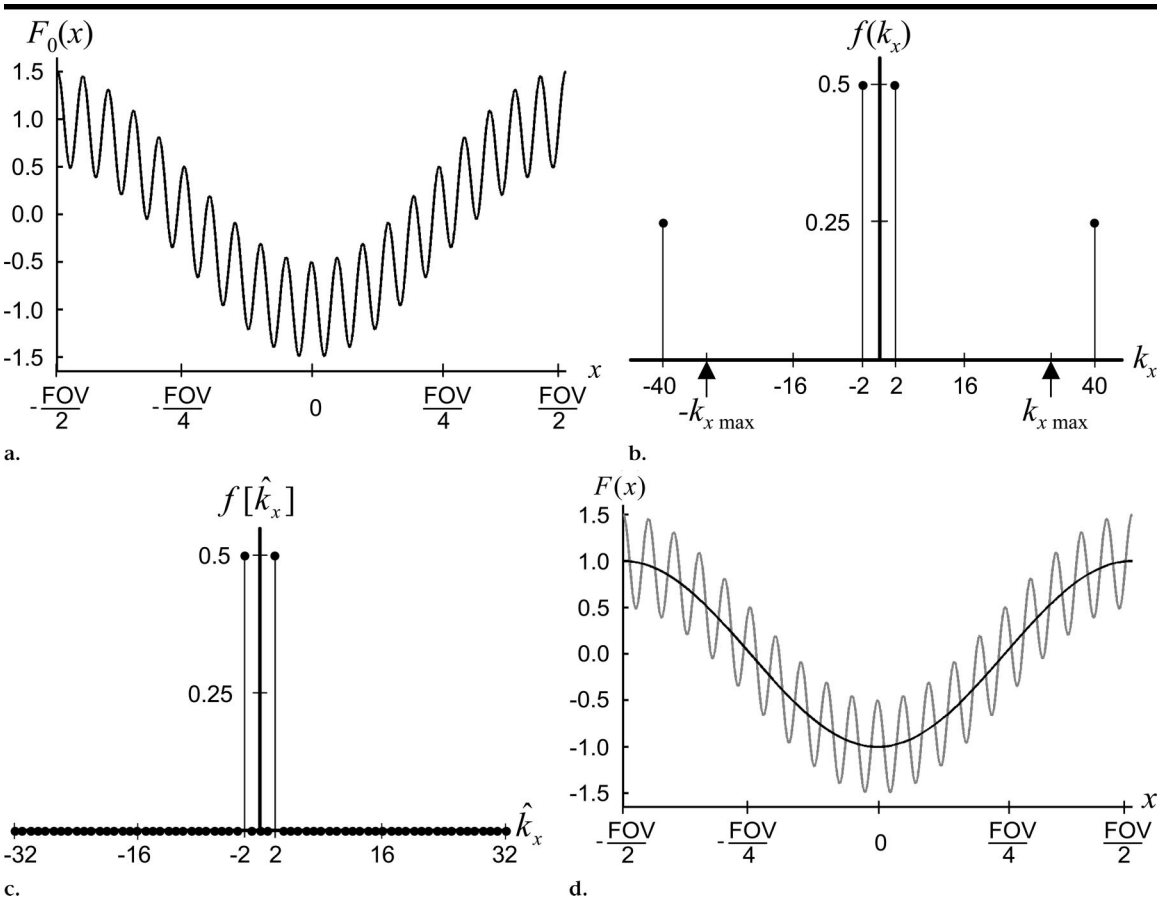

Figure 2. **(a)** Hypothetical one-dimensional object consisting of only two spatial sinusoids. **(b)** Continuous Fourier transform of the object contains nonzero values only at the frequencies of the two sinusoids. **(c)** The k-space raw data are essentially a sampled version of the Fourier transform in **b**. Imaging parameters determine the frequency and extent of the sampling. In this case, the spatial resolution was chosen such that sampling was performed only out to $k_{x\max} = \pm 32$; therefore, the high-frequency information at $k_x = \pm 40$ are not present in the raw data. **(d)** Plot of the DTFT of the data in **c** (dark curve). The DTFT function (Eq [13]) represents the sampled raw data with infinite spatial resolution; however, it does not contain the high-frequency components present in the original object (light curve and part **a**).

nite amount of time. In practice, sampling is performed long enough such that the necessary spatial frequencies (as specified by the prescribed spatial resolution in Eqq [5, 6]) are collected.

How can *F(x,y)* have infinite spatial resolution yet be limited by the acquired resolution? The basis of Fourier theory is that a given function can be decomposed into a sum of sine waves of different amplitudes and frequencies (5). A point in Fourier space (ie, k-space) is identified by two pieces of information: a value and a location. These attributes are equivalent to the amplitude and frequency, respectively, of a sinusoid; this is all the information necessary to characterize the sinusoid with infinite precision.

As an example, assume that the one-dimensional object shown in Figure 2a is to be imaged. This object is composed of only two spatial cosine waves, and its true Fourier transform (ie, its k-space [k_x, in this case] representation) is shown in Figure 2b. If, during the examination prescription, we choose a particular spatial resolution such that k_x is sampled only out to $k_{x\max} = \pm 32$, then we fail to collect any information about any higher spatial frequencies, specifically the points at $k_x = \pm 40$. The raw data we actually collect, which are shown in Figure 2c, contain only the low-frequency component of the original object. By applying the DTFT (Eq [11]) to the raw data (which consist of only two nonzero values at $\hat{k}_x = 2$ and $\hat{k}_x = -2$ with values $f[2] = ½$ and $f[-2] = ½$), we get

$$\begin{aligned} F(x) &= \tfrac{1}{2}\, e^{(-2\pi i(2)\,x/\mathrm{FOV})} + \tfrac{1}{2}\, e^{(-2\pi i(-2)\,x/\mathrm{FOV})} \\ &= \tfrac{1}{2}\left(e^{(-2\pi i 2x/\mathrm{FOV})} + e^{(2\pi i 2x/\mathrm{FOV})}\right) \\ &= \cos\left(2\pi\,\frac{2x}{\mathrm{FOV}}\right). \end{aligned} \qquad (13)$$

The last step follows from Euler's identity, $e^{ix} = \cos(x) + i\sin(x)$. Equation (13) exactly describes the imaged object for all values of x (ie, with infinite spatial resolution), even though the highest-frequency (ie, smallest-structure) information from the actual object is no longer present (Fig 2d).

It should be emphasized that the DTFT produces a function, not an image. A discrete image can be created by sampling the continuous function yielded by the DTFT. Alternatively, the image can be derived directly from the raw data by using the discrete Fourier transform (DFT) (4,6).

The one-and two-dimensional forms of the DFT are shown in Equations (14) and (15), respectively.

$$F[\hat{x}] = \sum_{\hat{k}_x=0}^{N-1} f[\hat{k}_x] e^{-(2\pi i/N)\hat{k}_x\hat{x}} \qquad (14)$$

for $\hat{x} = 0, 1, 2, \ldots, N - 1$.

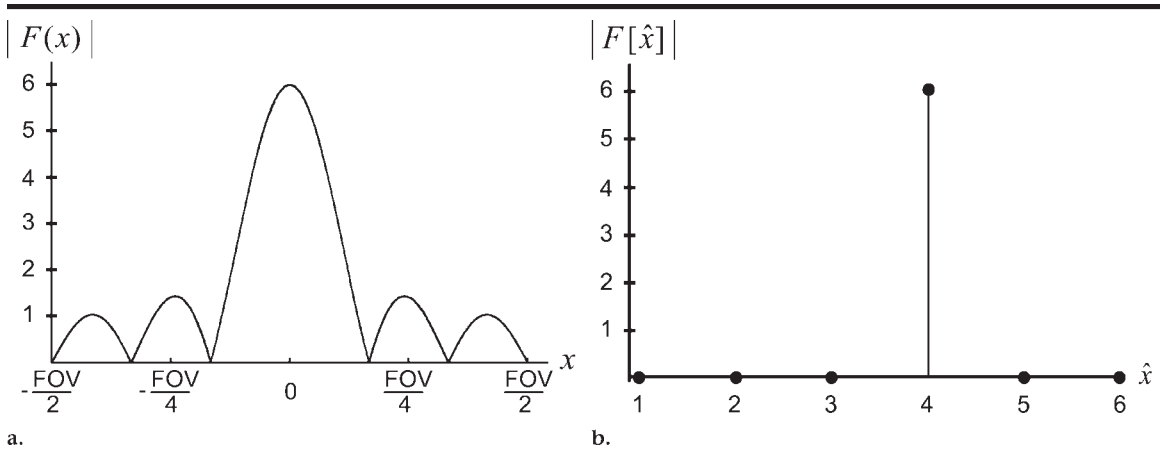

Figure 3. **(a)** The ideal reconstruction of $f[\hat{k}_x] = \{1, 1, 1, 1, 1, 1\}$ with use of the DTFT is a continuous and periodic function of x (periodicity not shown). **(b)** Six-point DFT reconstruction of the same data.

$$F[\hat{x}, \hat{y}] = \sum_{\hat{k}_x=0}^{N-1} \sum_{\hat{k}_y=0}^{M-1} f[\hat{k}_x, \hat{k}_y] e^{-2\pi i((\hat{k}_x\hat{x}/N)+(\hat{k}_y\hat{y}/M))} \quad (15)$$

for $\hat{x} = 0, 1, 2, \ldots, N - 1$ and $\hat{y} = 0, 1, 2, \ldots, M - 1$. Here, $F[\hat{x}]$ and $F[\hat{x},\hat{y}]$ represent reconstructed one- and two-dimensional images, x and y are the discrete image-domain pixel indices, and N and M are the total number of reconstructed pixels in the x and y directions, respectively. The DFT, which is implemented as the fast Fourier transform on MR imagers, is essentially a sampled version of the DTFT. Because of their similarities, it is important to recognize the difference between the expressions for the DTFT (Eqq [11, 12]) and the DFT (Eqq [14, 15]). Although both operations are performed for discrete sequences (ie, the sampled k-space data), the DFT yields a discrete sequence at a finite number of equidistant points ($\hat{x}$ and $\hat{y}$) along x and y, while the DTFT yields a continuous function of x and y.

What are the effects of reconstructing the data with the DFT? In the same way that sampling the MR signal necessarily limits acquired resolution, sampling $F(x,y)$ to obtain $F[\hat{x},\hat{y}]$ limits displayed resolution. While it may seem obvious that sampling an object N_x times in the frequency-encoding direction and N_y times in the phase-encoding direction results in raw data that have lower spatial resolution than that of the actual object, it may be less obvious that the resulting $N_x \times N_y$ image has lower resolution than that contained in the $N_x \times N_y$ raw data. These are the unavoidable results of sampling (eg, causing the resolution seen in the image to be lower than what was actually acquired) and can be demonstrated in the following example.

Let $f[\hat{k}_x] = \{1, 1, 1, 1, 1, 1\}$ be the sampled raw data of a one-dimensional ($N_x = 6$) object. By using Equation (11) for the DTFT, the ideal reconstructed object is represented by

$$F(x) = 1 + e^{-(1/3)i\pi x} + e^{-(2/3)i\pi x} + e^{-i\pi x} + e^{-(4/3)i\pi x} + e^{-(5/3)i\pi x}, \quad (16)$$

the magnitude plot of which is shown in Figure 3a. Recall that this is a continuous and periodic function, valid for all values of x. To generate an image, the DFT (Eq [14]) is used to yield

$$F[\hat{x}] = 1 + e^{-(1/3)i\pi\hat{x}} + e^{-(2/3)i\pi\hat{x}} + e^{-i\pi\hat{x}} + e^{-(4/3)i\pi\hat{x}} + e^{-(5/3)i\pi\hat{x}} \quad (17)$$

for $0 \leq \hat{x} \leq 5$. The result is the sequence of six points (the image) shown in Figure 3b. The image derived from the six-point DFT appears to bear very little resemblance to the ideal reconstructed object (Fig 3a), both of which were derived from the same raw data. In generating this hypothetical one-dimensional MR image, we have had two losses in spatial resolution: first during the acquisition of the N_x-point raw data and then again during the N_x-point DFT reconstruction.

Can this lost resolution be recovered? Yes and no. Acquired resolution cannot be improved because it is defined only by the examination prescription (ie, FOV, N_x, N_y). Nothing smaller than a tissue volume of this size can be distinguished; such information is simply not contained in the raw data. Displayed resolution, on the other hand, can be improved, but not beyond the acquired resolution. This can be achieved by means of sinc interpolation (also known as zero padding and zero filling) (6,7) and is performed routinely with clinical MR imagers. Zero padding simply involves appending zeroes to the raw data before the DFT.

The effect of zero padding is demonstrated in Figure 4, which involves the use of the same hypothetical raw data as in Figure 3. As the number of appended zeroes increases, the reconstructed image $F[\hat{x}]$ looks increasingly like the ideal reconstruction $F(x)$. With sufficiently increased size of the DFT, the true resolution inherent in the raw data can be represented accurately. An example of this is shown in Figure 5. A phantom was imaged twice with identical imaging parameters (256 × 256 acquisition matrix, 20-cm FOV, 0.78-mm nominal resolution) and reconstructed with both a 256 × 256 DFT and a 512 × 512 DFT. The spatial resolution is clearly superior with the 512 × 512 reconstruction and is much closer to the prescribed 0.78 mm than is the 256 × 256 reconstruction.

In practice, a zero-padding factor of two is used most commonly to improve the displayed resolution. Increasing the degree of zero padding would further improve the displayed resolution but with diminishing returns. A zero-padding factor of two is a reasonable compromise between desired displayed resolution and available hard disk space, considering that a 512 × 512 image requires four times the disk space of a 256 × 256 image.

Zero padding can be performed along any combination of the phase-, frequency-, and section-encoding directions during reconstruction. It is performed most often when the prescribed number of

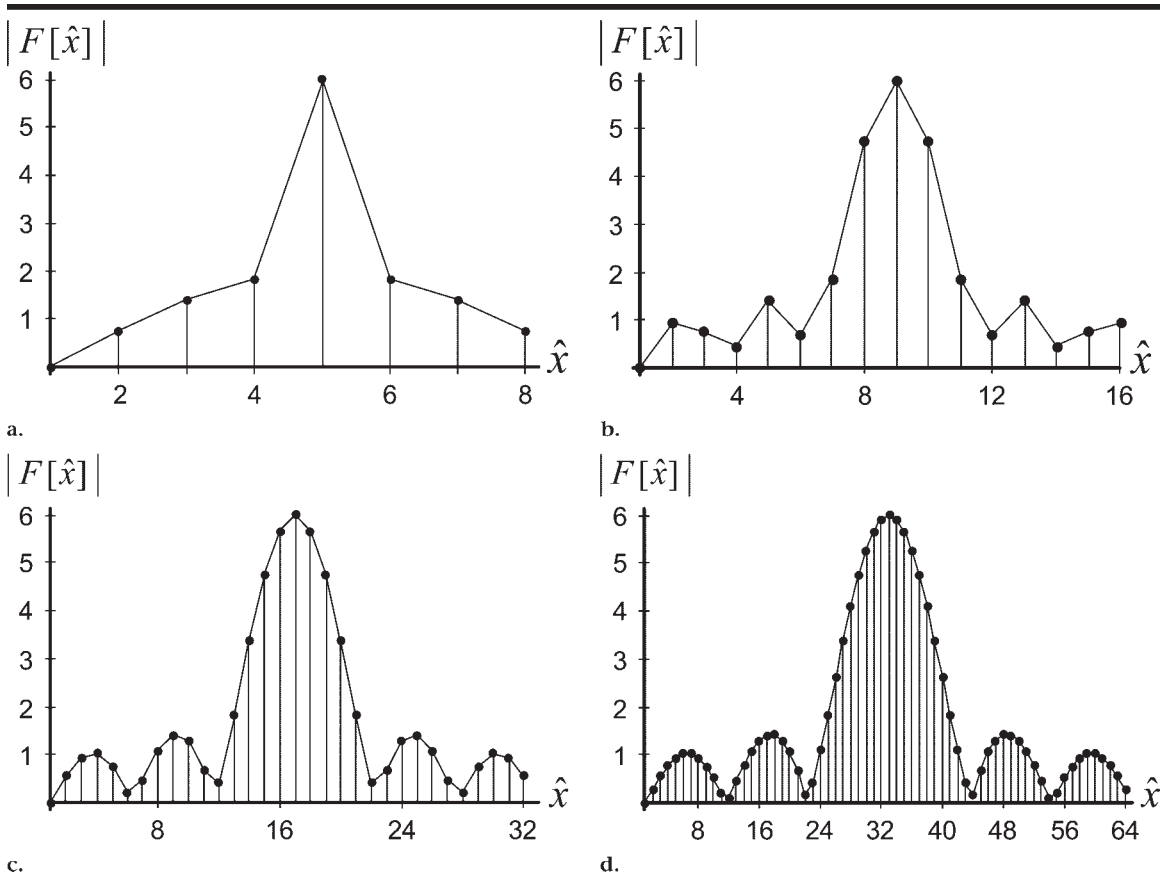

Figure 4. By using the same six-point "raw data" as in Figure 3, increasing the size of the DFT to **(a)** eight points, **(b)** 16 points, **(c)** 32 points, and **(d)** 64 points causes the reconstructed image to more closely resemble the ideal reconstruction in Figure 3a.

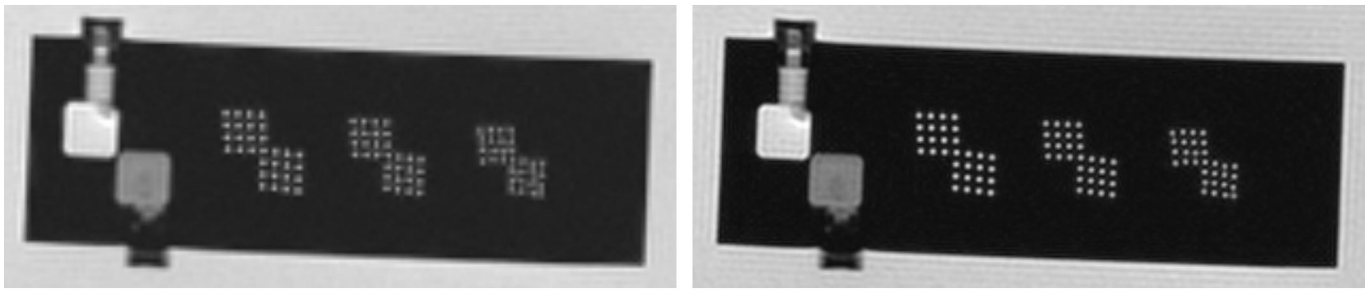

Figure 5. Phantom images demonstrate the effect of DFT size. Both images were acquired with identical imaging parameters (256 × 256 acquisition matrix, 20-cm FOV, 0.78-mm nominal resolution). **(a)** Reconstruction with a 256 × 256 DFT. **(b)** Reconstruction with a 512 × 512 DFT and zero filling. The zero-filled reconstruction is a better representation of the acquired resolution than is the 256 × 256 DFT.

phase-encoding lines is less than the number of frequency-encoding points. For example, assume that the following imaging protocol is used: 25.6 × 25.6-cm FOV, 256 (frequency) × 128 (phase) acquisition matrix, and 1.0 × 2.0-mm nominal resolution. In this case, the 256 × 128 raw data matrix is usually zero padded to 256 × 256 by appending 128 lines of zeroes to the phase-encoding direction. This yields a 256 × 256 image. Although the 25.6-cm FOV in the phase-encoding direction is now represented by 256 points (as opposed to the 128 acquired points), the nominal resolution remains 2.0 mm despite a pixel size of 1.0 mm. The displayed resolution, however, is closer to 2.0 mm in this case than if a 128-point DFT were performed in the phase-encoding direction.

These examples demonstrate how misunderstandings about resolution arise. If one assumed that the resolution in Figure 5a is actually that which was prescribed, then the logical conclusion on seeing the better resolution in Figure 5b is that the interpolation process improved the resolution beyond that which was acquired. Since we have seen that this cannot be the case, then the original assumption

must be incorrect. Reconstruction can often yield images with lower spatial resolution than what is actually present in the acquired data. Zero padding can improve the apparent resolution of an image by displaying more completely the information contained in the raw data. The acquired resolution, however, remains the same.

The preceding sample protocol can be extended to three dimensions: Assume that a 32-mm volume is to be encoded with 16 sections with a nominal section thickness of 2.0 mm. The most common method of zero padding in the section-encoding direction is to double the number of reconstructed sections. In this case, 16 "sections" of zeroes are added to the raw data, and the reconstruction yields 32 images. What is the spatial resolution in this case? Clearly, the section thickness is still 2.0 mm. Because of the imaging prescription that specified 16 partitions of the 32-mm volume, no information about structures smaller than 2.0 mm is present in the data.

Nevertheless, it is increasingly common in the literature for authors to refer to the section thickness in this case in a variety of ways: "1 mm after interpolation," "interpolated resolution of 1 mm," "reduced section spacing of 1 mm," or "overlapping 2-mm sections." In fact, image display algorithms can allow even higher degrees of interpolation; for example, 2.0-mm sections reconstructed at 0.5-mm intervals for renal or coronary imaging are achieved easily. Incorrect interpretation of spatial resolution relating to the 0.5-mm "image reconstruction interval" can be misleading and can cause confusion in relating results of one study to another.

As with the in-plane example discussed previously, while it is true that the displayed resolution in the section-encoding direction is improved with interpolation, it can never be better than the acquired resolution. For MR imaging, we frequently observe that such "interpolated resolution" is reported in the literature only with respect to the section-encoding direction, even though the identical process is performed routinely in the phase-encoding direction for most two- and three-dimensional examinations. Thus, it is likely that both readers and authors may not understand the nominal resolution of these images.

TEMPORAL RESOLUTION

Case example: A young investigator was performing an electrocardiographically gated cine phase contrast examination and was deciding how many temporal frames he wanted to reconstruct. Information he gathered from experience, recommendations, and the literature indicated that 20 reconstructed frames per R-R interval would be acceptable. One of his normal volunteer subjects was athletic and had a heart rate of 45 beats per minute. The investigator was surprised to learn subsequently that reconstruction of 20 frames for this bradycardiac athlete provided images every 67 msec, even though the pulse sequence had a nominal temporal resolution of 50 msec and actually acquired 26 frames. The investigator was in fact throwing away information during reconstruction.

Up to this point, we have discussed spatial resolution. There are interesting and direct analogies for specification of temporal resolution of cine images for both CT and MR imaging. A common example for cardiovascular MR imaging is the evaluation of left ventricular function. By imaging the contraction of the left ventricle, information about wall motion abnormalities, myocardial strain, and ventricular volumes can be obtained. The technique of choice for most of these applications involves the use of an electrocardiographically gated segmented cine acquisition (8).

Cine imaging is also widely used for phase contrast imaging of blood flow. The basic feature of cine imaging is that each cardiac cycle (R-R interval) is divided into several cardiac phases, and images acquired at these cardiac phases are played back as a movie. The duration of each cardiac phase is $\Delta t = \text{VPS} \cdot \text{TR}$, where VPS (views per segment) is a user-prescribed variable and TR is repetition time. Because the sequence TR is fixed for a given set of imaging parameters, the VPS is used to select the desired temporal resolution. But what is meant by "temporal resolution"?

If one uses the analogy for spatial resolution, then the temporal resolution can essentially be defined as the smallest increment of time over which a change in an imaged dynamic process can be observed. This definition is still somewhat vague when applied to cine MR imaging. Consider the following cardiac examination: Two cine acquisitions are performed with identical imaging parameters, except for the delay after the electrocardiographic trigger. The cardiac phases that are imaged with the two acquisitions are shown in Figure 6a. It is apparent that the duration of each cardiac phase with both acquisitions is the same (VPS · TR), but the respective temporal locations of the cardiac phases are offset by ½(VPS · TR). If we were to combine the data sets from the two acquisitions by interleaving the images, we would essentially double the number of points sampled during the cardiac cycle. If we take this one step further, we realize that because we are imaging the same process twice with the same imaging protocol, we can eliminate the second acquisition and generate the offset data from the first acquisition (Fig 6b). This concept of view sharing introduced by Foo et al (9) offers a way, without increasing the imaging time, to effectively double the number of temporal images. View sharing is now applied routinely in cardiac MR imaging and has the desirable effect of causing cine MR images to appear "smoother" as the cine loop is played.

Has view sharing actually doubled the temporal resolution, however? As Foo et al (9) were careful to point out, it has not. It is important to distinguish between the two temporal measurements in this case. First, there is the period of data acquisition for each image (Δt). Second, there is the temporal spacing between reconstructed images ($\Delta t/2$), where the reference point for each image is considered to be at the center of its data acquisition window. Temporal resolution should be defined as the period of data acquisition for a given image, which in this case is Δt. This is analogous to the definition of spatial resolution in that it is specified only by the acquisition. Any steps taken during reconstruction do not affect the temporal attributes of the acquired data; they merely alter its display. So although the temporal spacing has been halved to $\Delta t/2$, each image is still represented by data collected over a period Δt.

How then do we interpret view-shared data? On the surface, it can be seen as similar to zero padding, in that the reconstruction yields intermediate points and better visualization. However, there is an important distinction between the two techniques. Unlike view sharing, which improves only the appearance of the images, zero padding improves the appearance by actually displaying more of the information contained in the raw data. In other words, whereas view sharing cannot add new information to the images, zero padding can.

How can such a routinely applied algorithm such as view sharing be confusing to the user? One difficulty is that cardiac CT and MR images may be annotated with the starting time at which the image was acquired. For example, a "time delay" written on the first cine MR image may indicate that the acquisition started

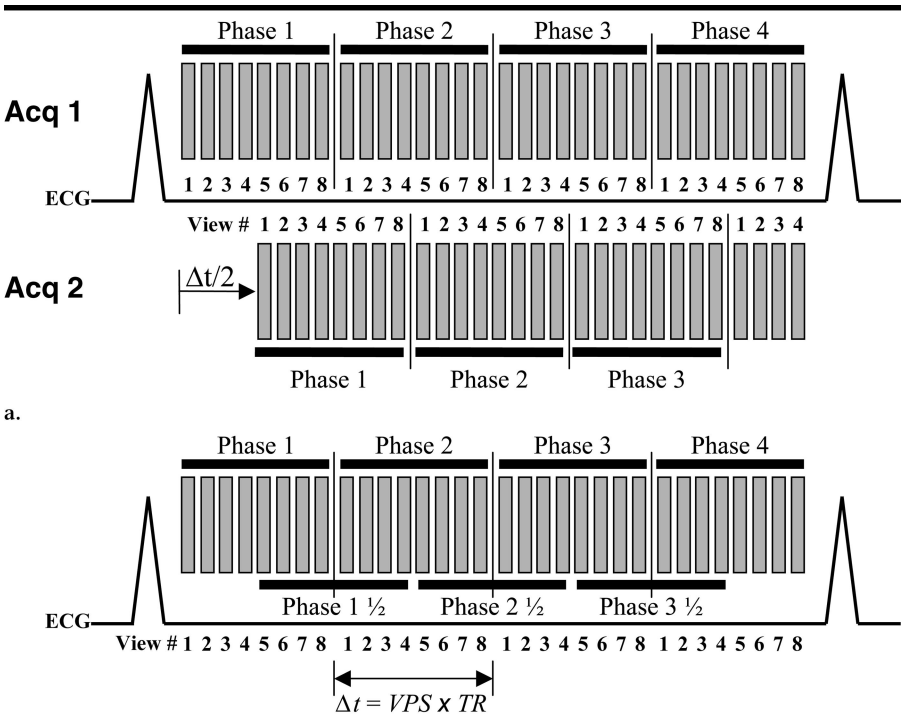

Figure 6. View sharing. **(a)** Two segmented cine acquisitions (VPS = 8) performed with identical imaging parameters. Temporal resolution Δt equals VPS · TR. The only difference is that data acquisition and hence the cardiac phase images in acquisition 2 *(Acq 2)* are delayed by $\Delta t/2$ with respect to acquisition 1 *(Acq 1)*. **(b)** Offset data from acquisition 2 in **a** can be created from data from acquisition 1. By regrouping cardiac phase data, additional data points (phases 1½, 2½, and 3½) can be reconstructed between original data points (phases 1, 2, 3, and 4). Although the effective time between adjacent reconstructed images is $\Delta t/2$, the temporal resolution for each cardiac phase image is still Δt.

25 msec after the R-wave of the cardiac cycle. The second and third images, for example, may be annotated as 50 and 75 msec, respectively. It may seem reasonable to assume that the temporal resolution of the cine sequence in this case is 25 msec (ie, the time difference between successive images). If the data were acquired with a TR of 5 msec and a VPS of 10, however, then the true temporal resolution is only 10 · 5 msec = 50 msec.

Thus, with current MR and CT platforms, it is frequently the case that the number of images that have been reconstructed in the R-R interval may have little or no relationship to the nominal temporal resolution. A relatively new example of this issue is the use of multi–detector row cine CT with retrospective electrocardiographic gating.

As pointed out by Kopp et al (10), although the temporal resolution for their coronary CT scanner was 250 msec (by using single-segment reconstruction), the authors were able to generate image reconstructions at intervals that corresponded to 10% of the cardiac cycle. If the patient's heart rate is 60 beats per minute (1 beat per second), the CT scanner could only generate 1000 msec, divided by 250 msec per cardiac cycle, thereby producing four images to display the entire cardiac cycle. However, a 10% reconstruction interval in this case would result in 1000 msec divided by 10 images—hence, 100 msec per image. Yet, each of the 10 images contains information (acquired temporal resolution) that covers 250 msec of the cardiac cycle.

The concept of "effective" temporal resolution has been used to denote the temporal spacing of reconstructed images; however, it is important to understand that this reconstruction-based spacing is entirely independent of the acquisition parameters. As we have discussed in the context of view sharing, high effective temporal resolution is valuable for improving the appearance of cine movies, but it has no acquisition-based relevance.

DISCUSSION AND RECOMMENDATIONS

We have reviewed some of the basic concepts of MR image acquisition and reconstruction as they pertain to spatial and temporal resolution. Our goal has been to clarify the distinction between the resolution acquired during imaging and that displayed after reconstruction. As the spatial and temporal resolution of MR and CT begin to challenge "standard of reference" examinations such as angiography, we would like to offer some recommendations with regard to reporting image resolution in the literature. We hope that this will be beneficial to authors, reviewers, and readers alike.

The most frequent misstatement of spatial resolution appears to be in the area of three-dimensional MR imaging or CT, and in particular, three-dimensional MR angiography. When not reported correctly, there can be confusion when readers attempt to understand or reproduce the results. Consequently, we believe that investigators should be discouraged from calculating the "interpolated resolution" that results from zero padding in the section-encoding direction. While it is certainly true that zero padding improves visualization when reformatting three-dimensional data sets (eg, surface and volume renderings, maximum intensity projections), authors should be careful when reporting its use. Ideally, reconstruction options should not be included in the description of the acquisition parameters.

To avoid misinterpretation or misunderstanding, we recommend that authors state the following: the slab thickness, the number of encodings of the slab (ie, the number of acquired sections), and the corresponding numeric resolution in millimeters, defined as the slab thickness divided by the number of encodings. If zero padding is used, the number of "reconstructed sections" or "interpolated sections" should be reported specifically as such (eg, "16 acquired sections interpolated to 32 sections"), without any quantitative measure of "interpolated resolution."

Unless there is a specific reason to do so, however, the use of zero padding and its effect on reconstructed sections may not even need to be included. (For example, zero padding is rarely reported in the phase- and frequency-encoding directions despite its extensive use for partial echoes, fractional number of signals acquired, rectangular phase FOV, or simply when $N_y < N_x$.) In general, when spatial resolution is stated in conventional MR imaging, it should be specified (in any of the three directions) as the FOV divided by the number of encodings.

Similar misunderstandings can also

arise in cine imaging. This kind of misunderstanding can be avoided if investigators become more accustomed to thinking in terms of true acquired temporal resolution rather than in terms of reconstructed frames per second. In segmented cine examinations, the temporal resolution should be reported as the minimum span of time from which data for a single reconstructed image (frame) are collected. This is usually VPS · TR. However, it should be recognized that there are novel imaging techniques in which the same data are used in multiple images or in which various parts of the data have different temporal resolutions. In some of these cases, true temporal resolution can be difficult if not impossible to define. Regardless of the acquisition technique (and analogous to the three-dimensional imaging case discussed earlier), there are visualization benefits derived from reconstructing the acquired data with view sharing or other retrospective interpolation schemes. Nevertheless, only the true temporal characteristics should be stated as part of the description of the acquisition protocol.

CONCLUSIONS

When reporting spatial or temporal resolution, it is imperative that investigators distinguish between acquisition and reconstruction parameters. To accomplish this, they must fully understand not only the prescribed imaging parameters but how their particular imager implements them. For example, investigators should know when and if zero padding is performed and whether their cine acquisitions involve the use of view sharing or some other method of retrospective interpolation. With this knowledge, terms such as *effective resolution* or *interpolated resolution* can be avoided, and misunderstandings between authors and their audience can be minimized.

Like the digital camera with a zoom factor of 120 mentioned in the Introduction, our CT and MR imagers are capable of tremendous image processing tricks that improve our perception of spatial and temporal resolution. As these noninvasive modalities approach the resolution of standards of reference, such as screen-film radiography or angiography, it is becoming critical to have an accurate understanding of these concepts.

APPENDIX

Some technical comments regarding the form of the DTFT shown in Equations (11) and (12) are warranted. In the digital signal processing literature, the one-dimensional DTFT is most commonly written as

$$F(\omega) = \sum_{k=-\infty}^{+\infty} f[k]e^{-i\omega k}, \qquad \text{(A1)}$$

where ω denotes the continuous radian frequency variable, which is periodic over the interval $-\pi$ to $+\pi$. If $f[k]$ is taken to be the sampled MR raw data, then Equation (A1) yields an expression for the object as a function of radian frequency. However, radians are not a natural unit for representing a physical object. Because the object domain is typically represented by the variable x in centimeters, we would prefer to have the reconstructed object also be a function of x—that is, $F(x)$. To do this, we need to rewrite the variables in Equation (A1).

In general, the radian frequency ω can be written as

$$\omega = \frac{2\pi f}{f_s}, \qquad \text{(A2)}$$

where f is the frequency variable, and f_s is the sampling rate. In Equation (A1), the sampling frequency f_s is set equal to 1, which has the effect of normalizing the frequency axis. Because we are interested in preserving the units on the frequency axis, we must avoid this normalization. In our case, the "frequency" variable is x, and the sampling rate of k-space is FOV (see Eqq [9, 10], where the sampling interval is the inverse of the sampling rate). We can then rewrite Equation (A2) as

$$\omega = \frac{2\pi x}{\text{FOV}}. \qquad \text{(A3)}$$

Substitution of Equation (A3) into Equation (A1) yields

$$F(x) = \sum_{k=-\infty}^{+\infty} f[k]e^{-i(2\pi x/\text{FOV})k}, \qquad \text{(A4)}$$

which is the basic form shown in Equations (11) and (12). This allows us to represent the reconstructed object as a function of x, where $-(\text{FOV}/2) \le x \le (\text{FOV}/2)$ (see Figs 2 and 3).

GLOSSARY

Terminology

Acquired spatial resolution.—The smallest tissue volume size that can be represented on an image. It is defined as the prescribed FOV in the frequency- and phase-encoding directions divided by the number of frequency- and phase-encoding points ([FOV_x/N_x] and [FOV_y/N_y]), respectively.

Acquired temporal resolution.—The duration of data acquisition (Δt) for each image in a dynamic or cine examination. For example, for a cine examination, Δt = VPS · TR. This definition can vary substantially, depending on the acquisition scheme.

DFT.—Discrete Fourier transform. A variant of the Fourier transform that is performed with discrete (ie, digital) data and yields a discrete result.

Displayed spatial resolution.—The apparent tissue volume size displayed on the reconstructed image. It can never be higher than the acquired resolution, but it can be lower.

Displayed temporal resolution.—The temporal spacing between reconstructed images in a dynamic examination.

DTFT.—Discrete-time Fourier transform. A variant of the Fourier transform that is performed with discrete (digital) data and yields a continuous result.

FOV_x, FOV_y.—FOV in the frequency- and phase-encoding directions, respectively.

FT.—Fourier transform.

M.—Number of reconstructed points (pixels) in the phase-encoding direction.

N.—Number of reconstructed points (pixels) in the frequency-encoding direction.

Nominal spatial resolution.—Same as acquired spatial resolution.

N_x.—Number of frequency-encoding points.

N_y.—Number of phase-encoding points.

Pixel size.—The size of a pixel on a reconstructed image. Its dimensions are (FOV_x/N) · (FOV_y/M). It is unrelated to and should not be confused with the acquired spatial resolution.

View sharing.—A technique to improve displayed temporal resolution by sharing data from images at two adjacent time points to create an image at an intermediate time point. See reference 9 for details.

VPS.—The number of phase-encoding lines (views) that define the acquired temporal resolution for a cine examination. See temporal resolution (acquired).

Zero padding (zero filling).—A technique to improve displayed spatial resolution by appending zeroes to the raw data before image reconstruction. It allows the displayed spatial resolution to match the acquired spatial resolution more closely.

Mathematic Notation

$F_0(x,y)$.—Object-domain representation of the object to be imaged.

$F(x,y)$.—Continuous representation of object after reconstruction with DTFT. This is a function, not an image; x and y are expressed in centimeters, for example.

$f(k_x(t),k_y(t))$.—Continuous representation of the Fourier transform of the object. k_x and k_y are expressed in 1/centimeters.

$F[\hat{x},\hat{y}]$.—Discrete representation of the object—that is, the image. It is a sampled ver-

sion of $F(x,y)$; $\hat{x}$ and $\hat{y}$ are dimensionless indices.

$f[\hat{k}_x,\hat{k}_y]$.—Discrete representation of the Fourier transform of the object—that is, the raw data. It is a sampled version of $f(k_x(t), k_y(t))$; $\hat{k}_x$ and $\hat{k}_y$ are dimensionless indices.

References

1. Haacke EM, Brown RW, Thompson MR, Venkatesan R. Magnetic resonance imaging: physical principles and sequence design. New York, NY: Wiley-Liss, 1999.
2. Bracewell RN. The Fourier transform and its applications. New York, NY: McGraw-Hill, 1978.
3. Mezrich R. A perspective on k-space. Radiology 1995; 195:297–315.
4. Oppenheim AV, Schafer RW. Discrete-time signal processing. Englewood Cliffs, NJ: Prentice Hall, 1986.
5. Brigham EO. The fast Fourier transform and its applications. Upper Saddle River, NJ: Prentice Hall, 1988.
6. Lyons RG. Understanding digital signal processing. Reading, Mass: Addison-Wesley, 1997.
7. Ludeman LC. Fundamentals of digital signal processing. New York, NY: Wiley, 1986.
8. Atkinson DJ, Edelman RR. Cineangiography of the heart in a single breath hold with a segmented TurboFLASH sequence. Radiology 1991; 178:357–360.
9. Foo TK, Bernstein MA, Aisen AM, Hernandez RJ, Collick BD, Bernstein T. Improved ejection fraction and flow velocity estimates with use of view sharing and uniform repetition time excitation with fast cardiac techniques. Radiology 1995; 195: 471–478.
10. Kopp AF, Schroeder S, Kuettner A, et al. Coronary arteries: retrospectively ECG-gated multi–detector row CT angiography with selective optimization of the image reconstruction window. Radiology 2001; 221:683–688.