Refraction corrected transmission ultrasound computed tomography for application in breast imaging

Shengying Li
Department of Computer Science, Stony Brook University, Stony Brook, New York 11794-4400

Marcel Jackowski[a]
Department of Computer Science, University of Sao Paulo, 05508-090 Sao Paulo, Brazil

Donald P. Dione[b]
Ultrasound Detection Systems LLC, Madison, Connecticut 06443

Trond Varslot[c]
Department of Applied Mathematics, Australian National University, Canberra, ACT 2601, Australia

Lawrence H. Staib[d]
Department of Biomedical Engineering and Department of Diagnostic Radiology, Yale University, New Haven, Connecticut 06520-8042

Klaus Mueller[e]
Department of Computer Science, Stony Brook University, Stony Brook, New York 11794-4400

Purpose: We present an iterative framework for CT reconstruction from transmission ultrasound data which accurately and efficiently models the strong refraction effects that occur in our target application: Imaging the female breast.

Methods: Our refractive ray tracing framework has its foundation in the fast marching method (FNMM) and it allows an accurate as well as efficient modeling of curved rays. We also describe a novel regularization scheme that yields further significant reconstruction quality improvements. A final contribution is the development of a realistic anthropomorphic digital breast phantom based on the NIH Visible Female data set.

Results: Our system is able to resolve very fine details even in the presence of significant noise, and it reconstructs both sound speed and attenuation data. Excellent correspondence with a traditional, but significantly more computationally expensive wave equation solver is achieved.

Conclusions: Apart from the accurate modeling of curved rays, decisive factors have also been our regularization scheme and the high-quality interpolation filter we have used. An added benefit of our framework is that it accelerates well on GPUs where we have shown that clinical 3D reconstruction speeds on the order of minutes are possible. © *2010 American Association of Physicists in Medicine.* [DOI: 10.1118/1.3360180]

I. INTRODUCTION

After lung cancer, breast cancer strikes more women in the world than any other type of cancer (excluding cancers of the skin). In 2007, according to the American Cancer Society (ACS), an estimated 1.2 million women worldwide were diagnosed with breast cancer, and about 465 000 women died from the disease. Early detection of breast cancer has repeatedly been shown to improve the chance of survival,[1–3] and x-ray mammography is currently the method of choice for breast cancer screening. However, false positive x-ray mammograms result in a large number of unnecessary biopsies. The ACS reports that in the U.S. approximately 80% of the million biopsies performed each year are benign. X-ray mammography also subjects patients to significant ionizing radiation, which renders short-term follow-up monitoring prohibitive. Furthermore, mammography yields 2D whole-breast projection images, which requires the breast to be compressed in order to minimize the superposition of tissues and therefore reduce structural noise. Apart from being in-

convenient to the patient, this compression also makes the localization of lesions difficult once the compression is removed. While these shortcomings are addressed or at least alleviated by tomo-synthesis and 3D cone-beam imaging,[4,5] the radiation dose still remains. The use of MRI technology can address this problem, but here the high cost is a limiting factor. Several newer technologies are also emerging (that is, acoustical holography, infrared, electrical, optical, and elasticity methods) but none have yet proven to be of use in a clinical setting.[6]

On the other hand, (low-cost) ultrasound echo imaging is routinely employed as an adjunct to x-ray mammography for the differentiation of simple cysts from solid masses. However, echo is often not effective enough at differentiating malignant and benign masses to avoid the need for biopsy. This shortcoming can be overcome by replacing (or augmenting) reflective by transmission ultrasound followed by tomographic reconstruction. Beginning with the seminal work of Greenleaf,[7] several research groups have attempted this[3,8–12] with promising results. Their methods can be dis-

tinguished by the framework employed for modeling the propagation of sound in the heterogeneous medium that makes up the female breast.

The most accurate propagation framework for transmission ultrasound image reconstruction is based on the full wave equation. Solving the full wave equation, however, is computationally very expensive. To keep the computational complexity reasonable, Born or Rytov approximations are often used. These are first-order scattering approximations and reasonable when multiple reflections can be ignored. The Born approximation is generally seen to giving a better description of backscattering, while the Rytov approximation yields better results for the forward propagation in particular in the presence of objects which are large relative to the wavelength of the propagating field. Born and Rytov approximations are frequently used in diffraction tomography[13–15] since they allow for the modeling of effects such as interference, diffraction, and refraction. Geometrical acoustics, on the other hand, describes energy propagating along rays and is a high-frequency approximation. As such, it does not account for finite-frequency phenomena such as diffraction and interference. However, it does allow for a much more efficient numerical solution while accurately modeling refraction.

For the purpose of reconstruction, straight ray approximations are often used in geometrical acoustics in order to have the simplest least computationally complex propagation method. The general limit of any of the above approximations to the full wave equation is that the maximal resolution obtainable is bounded by $\lambda/2$, where λ is the acoustic wavelength in the medium. Thus, higher frequencies yield better resolution but have an increased amount of small-scale scattering which causes faster signal decay and less penetration depth. While techniques[16,17] have been developed that can resolve detail at resolution of $\lambda/3$ and better, these are still in an experimental stage and it is not clear if they can work in practice at feasible overhead.

The decisive question when selecting the most appropriate approximation to the wave equation is what effects dominate in the target setting, here the female breast. Zhu and Steinberg[18] have shown in both *in vivo* and phantom studies that it is refraction, rather than weak scattering, that dominates the distortion of the ultrasonic wave front. The main structures in the breast are the fat layers under the skin which are highly refractive. Diffraction effects, on the other hand, are less significant. Thus it appears advantageous to select a framework rooted in geometrical acoustics, but with bent (curved) rays to properly reproduce the strong refractions at the fat-skin interface and elsewhere. As Andersen has shown,[19,20] ray tracing can resolve refractions of up to 20% which is well above the most refractive interfaces in the breast where adjacent material refractive indices vary up to 8%.

In this paper we describe a high-fidelity transmission ultrasound CT (UCT) reconstruction method based on ray tracing. Our method is iterative, which is essential to cope with noise and a goal-directed solution space traversal in the presence of inconsistent data. Iterative methods require a physi-

cally accurate projector to obtain the estimate for correction computation, and proper regularization in the correction stage is also needed to drive the estimate closer to an accurate solution. Finally, since it is well known that these measures are computationally expensive they must be cast into an efficient framework and implemented on an affordable computational platform (we use the GPU) that provides feasible computation times enabling deployment in a clinical setting without compromising reconstruction fidelity. We describe a complete framework that fulfills all of these stated design goals and tests it using realistic phantom data. A preliminary version of our system has appeared in Ref. 21.

II. BACKGROUND AND SIGNIFICANCE

Our approach is rooted in *ray theory*, popular in seismology and seismic imaging. Ray theory allows refraction effects to be modeled accurately, assuming a *high-frequency (short wavelength)* assumption is satisfied. With this assumption, the length scale of the features of interest in the medium cannot be lower than half the signal's wavelength, $\lambda/2=0.5\cdot v/\omega$, with v being the speed of sound in the medium and ω the signal frequency. We will show that for our particular application of ultrasound CT breast imaging this assumption is well satisfied.

II.A. Eikonal equation

A central mathematical driver in ray theory is the Eikonal equation which derivation we sketch here. We start by formulating the propagation of a compression wave in a heterogeneous medium,

$$\nabla^2\Phi = \frac{1}{v^2}\frac{\partial^2\Phi}{\partial t^2}. \tag{1}$$

Here, Φ is the wave's scalar potential and ∇^2 is the Laplacian operator. We then assume a harmonic solution of the form

$$\Phi = A(x)e^{-i\omega(T(x)+t)}, \tag{2}$$

where $A(\boldsymbol{x})$ is the wave's amplitude at location $\boldsymbol{x}$ and $T(\boldsymbol{x})$ is a phase function that describes a surface of constant phase. Substituting the above expression into the wave equation yields

$$|\nabla T|^2 - \frac{1}{v^2} = \frac{\nabla^2 A}{A\omega^2}. \tag{3}$$

The aforementioned high frequency approximation assumes that the normalized Laplacian of the amplitude is much smaller than ω^2, forcing the right hand side of Eq. (3) to zero. Using the high frequency approximation transforms Eq. (3) to the Eikonal equation,

$$|\nabla T|^2 = \frac{1}{v^2} = U^2, \tag{4}$$

where U is called the *slowness* (or *reciprocal sound speed)*. Using this equation we can define a wave front as those points in space for which $T=$ constant, and $T(x_0)$ defines the time required by the wave front to reach a certain x_0 (see Fig.
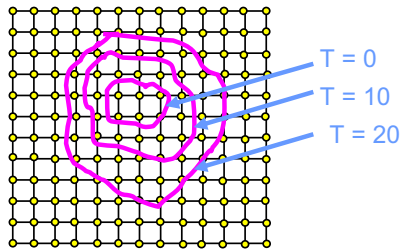
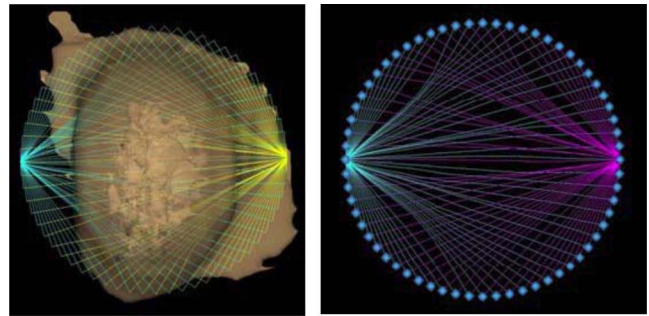FIG. 1. Propagating wave front with temporal isocontours.

FIG. 2. Shown are the ray paths from two ultrasound emitters (calculated using our projection method) through a simulated human breast phantom as viewed from above (the rays have been shaded differently to indicate their originating sources). There are 64 transducer elements, arranged equiangularly spaced in a single ring detector. On the left are the paths through a volume rendering of the breast. On the right are the same paths without the volume rendering. We observe strong refractive effects along the sides of the breast due to the subcutaneous fat layer.

1 for an illustration). This time is often called *time of flight* (*TOF*), giving rise to a *TOF-field* spanning the domain. The gradient of the wave front (the TOF-field at x) $\nabla T(x)$ then yields the instantaneous vector of a ray passing across the wave front. We will use this gradient to efficiently trace rays across a refractive medium.

II.B. Projection simulator

It is well known for iterative CT reconstruction that the fidelity of the forward projector, simulating the ray integrals of the current reconstruction estimate, has a crucial effect on reconstruction quality. The fidelity of the backprojector is less crucial in this respect.[22] In the projector, the more closely the line integrals model the relevant aspects of the physical energy propagation process, the more accurate the grid corrections computed by comparing these integrals with the acquired data will be. However, in contrast to transmission x-ray CT where rays are approximately linear, a major difficulty in transmission ultrasound CT (using ray-based approaches) has been to efficiently determine the curved, nonlinear paths of the rays across the refractive media, from source to receivers and back. Several investigators have explored the use of iterative ray tracing via ray linking, ray rebinning, and others to accomplish this.[7,20,23,24] However, none of these methods have been efficient enough to be deployed in practice, and thus the more recent literature is dominated by straight ray propagation methods. But as was shown by Denis *et al.*[25] and more recently by Li *et al.*,[26] nonlinear ray tracing can achieve substantial qualitative improvements over straight ray casting, even for moderately refracting fields. Yet that work did not offer efficient methods to accomplish this. Figure 2 shows the results of an ultrasound simulation in a female breast phantom, using our projection method. Strong refraction effects at the subcutaneous fat layer can be observed, and these prominent ray distortions suggest that time of flight (TOF) is an important and dominant feature in the transmission ultrasonic imaging of the female breast.

As motivated above, the gradient of the wave front provides a natural instrument for guiding the rays across the refractive media. All one requires is the TOF field of the current estimate, originating from a given source. Our framework computes this TOF field very efficiently using the fast marching method (FMM). The FMM was first described by Tsitsiklis[27] and has become quite popular in computer graphics and vision, and also in seismology.[28] Our experiments

indicated, however, that a projector based on the traditional FMM is not sufficiently accurate within a CT reconstruction framework. We thus employ a high-accuracy extension to the FMM, called the HAFMM.

We have conducted simulation experiments that indicated that there can also be out-of-plane refractions. Incorporating these effects into the UCT reconstruction procedure extends the ray tracing from the typical 2D slice-stack procedure into a 3D volumetric setting. However, since 3D wave tracking is computationally expensive and the out-of-plane refraction is not extensive, we devised a reconstruction procedure that operates only within a certain slice neighborhood from the given center slice.

II.C. Reconstruction algorithm

The nonlinear rays, data noise, and possibly incomplete data make the use of classical tomography reconstruction algorithms based on filtered backprojection less favorable.[29] Thus, similar to other UCT researchers, we employ an algebraic reconstruction approach, SART,[19] which attempts to iteratively solve the following system of equations,

$$p_i = \sum_{j=1}^{N} w_{ij} v_j, \quad i = 1,..,M, \tag{5}$$

where the p_i are the acquired projection (transducer) data, the v_j are the Cartesian lattice pixels (or voxels in 3D) subject to reconstruction, and the w_{ij} are the weight factors that relate the pixels to the data. In the simplest form these weights are determined by the interpolation filter but they could incorporate ray effects as well, such as attenuation. Each equation in Eq. (5) represents the ray integral of the energy arriving at receiver i. In our ultrasound CT application there are S transducers each emitting once, with the $(S-1)$ remaining transducers receiving the transmissions, thus $M = S \cdot (S-1)$. Equation (5) may also be expressed in matrix notation $P = W \cdot V$, with W being the (sparse) system matrix. SART seeks to solve $V = W \cdot P$ via the following iterative framework:
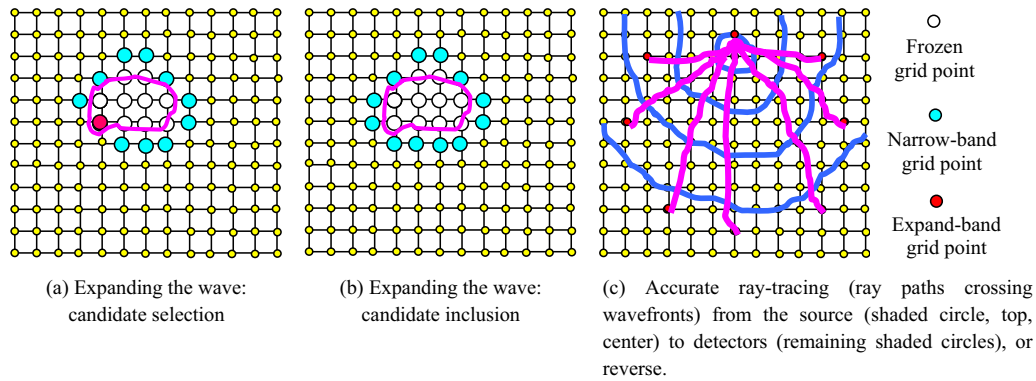
(a) Expanding the wave: candidate selection

(b) Expanding the wave: candidate inclusion

(c) Accurate ray-tracing (ray paths crossing wavefronts) from the source (shaded circle, top, center) to detectors (remaining shaded circles), or reverse.

○ Frozen grid point

● Narrow-band grid point

● Expand-band grid point

FIG. 3. Illustration of the fast marching method.

$$v_j^k = v_j^{k-1} + \alpha \frac{\sum\limits_{p_i \in P_\varphi} c_i w_{ij}}{\sum\limits_{p_i \in P_\varphi} w_{ij}}, \quad c_i = \frac{p_i - \sum\limits_{l=1}^{N} w_{il} v_l^{k-1}}{\sum\limits_{l=1}^{N} w_{il}}, \quad (6)$$

where k enumerates the iteration. The correction/update factors c_i are computed by subtracting the result of a discrete (nonlinear) ray integration [given the current object estimate reconstructed at iteration $(k-1)$] from the physical integration acquired at receiver i, normalized by the sum of weights. SART is a block-based algorithm, that is, a grid update occurs after all rays for a given source (emitter) due to a certain projection P_φ have been traced and the correction factors computed. Here, the subscript φ denotes the emitter's angular location in the circular arrangement of transducers. Finally, α is the relaxation factor. It controls the convergence speed of the reconstruction in the presence of noise. Much research has been devoted to the choice of α, typically either leaving the factor constant or decreasing it as iterations proceed. We propose an estimate-driven selection which improves reconstruction quality significantly.

A key contribution of our paper is that we combine SART with FMM to find the accurate ray directions by wave propagation. In this way we can avoid the complicated bent-ray computations that previous UCT reconstruction algorithms had to contend with, replacing them with the simple and linear computations afforded by the FMM. We thus call our iterative reconstruction scheme FMM-SART.

III. METHODS

Figure 2 illustrates the setup of our system. It consists of a cylindrical multirow transducer unit with up to 256 transducers on each row, equiangularly spaced. The transducers are considered point sources and operate at a center frequency of 5 MHz. The pulse is sinusoidal and windowed using a Gaussian to yield a −6 dB FWHM bandwidth of about 2 MHz.

During acquisition, each transducer functions once as the emitter while the others are receiving. In a practical, but experimental system that we have conceived in our laboratory (not the focus of this paper) we record as TOF the time

at which the extremum of the first half wave is detected, while the attenuation data are formed by the magnitude of the extremum of the first half wave. The results presented in this paper, however, are based on data obtained by simulation with a high-quality wave PDE solver. We choose to employ simulated instead of real data to highlight the algorithmic aspects of our work, using phantoms with realistic sound speed and attenuation properties. We also note that we use the term attenuation in a more general sense, in the form of *insertion loss*. While attenuation in a strict sense is the reduction in signal only due to energy absorption, insertion loss is the loss in energy due to anything (e.g., attenuation, scatter, reflection, diffraction) that reduces the energy that arrives at the receiver.

Our UCT reconstruction framework iteratively tracks wave fronts through space, based on the current estimate of the sound speed (SS) field, and then uses the resulting TOF field to guide the refractive rays from the sensors to the source, updating the SS field by the iterative corrections. Once the SS distribution is reconstructed we use it for ray guidance in the reconstruction of the attenuation distribution. We now detail this approach.

III.A. Transmission ultrasound simulator

We first rewrite the Eikonal equation given in Eq. (3) into a more implementation-accessible form[30] (here, to be most general, we formulate the expression in 3D),

$$(\partial t/\partial x)^2 + (\partial t/\partial y)^2 + (\partial t/\partial z)^2 = 1/\mathrm{SS}^2(x,y,z). \quad (7)$$

The (x,y,z) are discrete spatial grid coordinates, and the sound speed term $\mathrm{SS}(x,y,z)$ is a measure of the sound conductance properties at that point. As mentioned, we solve the Eikonal equation efficiently by employing the fast marching method (FMM), originally proposed by Tsitsiklis.[27] The FMM is related to Dijkstra's method,[31] which is a classical algorithm for identifying the shortest path in a network of links. The FMM is a single-pass, upwind finite difference scheme, which produces the correct viscosity solution to the Eikonal equation.[28]

The FMM computes for every voxel (x,y,z) the time $T(x,y,z)$ at which the wave from a particular emitter has traversed it (see Fig. 3 for an illustration). It starts by putting

all voxels connected to the source into a *narrow band.* Then, as the wave front proceeds across the grid, the FMM *selects* that voxel (x,y,z) in this narrow band (situated immediately upwind from the current wave boundary) which minimizes the time increment, given the speed values of the narrow band voxels. It then *freezes* the voxel, writing to it the TOF and inserts its neighbors into the narrow band (if not already present). The final result of the FMM is the time of flight (TOF) image. There is one such image for each emitter. Using the TOF image one can then compute accurate ray paths from the source to the detectors.

We use a binary heap to quickly find the voxel with the smallest postulated wave arrival time in the narrow band of the advancing wave front. This voxel's wave arrival value is then written to the TOF image, and its neighbors and their arrival times are updated in the heap. Note that a translation table with double pointers (as is further described in Refs. [28] and [32]) is used in order to quickly map the spatial domain voxels to the heap voxels.

The original FMM solves the Eikonal equation using only first-order finite differences. However, we found that this approximation leads to inaccuracies at high curvature boundaries. For a more accurate approximation of Eq. (6) we employ a second-order approximation to the partial derivatives, the higher accuracy FMM (HAFMM),[28]

$$(\partial t/\partial x) = (3t(x,y,z) - 4t(x-1,y,z) + t(x-2,y,z))/2. \quad (8)$$

We note that this formulation also requires accurate second-order estimates for the initialization around the propagation seed points (the emitter locations).

Once the wave propagation is complete we use the resulting TOF image to calculate the path of the rays from the receivers back to the emitter [the reverse of Fig. 3(c)]. The TOF image allows us to locally compute the ray direction vectors, given by the TOF image gradients. This approach ensures that a given ray will not miss the emitter (within a tolerance of 1–2 voxels if a high-quality interpolation filter is used to estimate the image gradients, see below), and thus the need for the tedious ray linking and path assembly of earlier approaches is eliminated. Figure 1 shows the acoustic ray paths from 31 receivers, which are distorted when passing through the object.

An important issue here is the interpolation of the gradients from the TOF image. Since derivatives are increasingly more sensitive to interpolation filter deficiencies than scalars, as documented in our earlier work,[33] we use a high-quality B-spline derivative filter to calculate the TOF gradients.

III.B. Iterative FMM-SART UCT reconstruction algorithm

Our framework consists of two consecutive phases. In phase 1, we iteratively reconstruct the sound speed (SS) image from the TOF data collected at the receivers. In phase 2, we use this SS image to guide the nonlinear rays for the iterative reconstruction of the sound attenuation (SA) image from the attenuation data collected at the receivers. Both phases employ SART as the iterative reconstruction engine.

```
Initialize sound speed image SS
Until convergence, loop
    Until all emitters have been chosen
        Randomly select an emitter E
        Load E's acquired TOF receiver data TOF_acquired
        Propagate wave from E using current SS estimate
        Record TOF at each voxel
        Record TOF at receivers → TOF_simulated
        For each receiver R_i
            Trace rays from R_i to E using TOF gradient
            Store ray path rp for backprojection
            Accumulate ray length L
            Compute correction factor as
                ΔTOF = ( TOF_acquired - TOF_simulated )/L
        Back project ΔTOF along rp
        Update SS estimate by ΔTOF at each rp voxel
```

FIG. 4. The FMM-SART algorithm for sound speed reconstruction.

Note that the SA image is easy to reconstruct once an accurate SS image is available for guiding the distorted rays, provided the gradients are faithfully reconstructed using good interpolation filters. We can save further time by storing (and subsequently using) the ray paths computed in the final SS-reconstruction iteration.

When computing the SA image, we need to account for the geometrical spreading of the acoustic energy as the wave front expands away from the source. This expansion scales down the energy (and subsequently scales up the estimated attenuation). In a homogeneous medium, the energy in a wave front falls off at a factor of $1/r$ (with r being the distance from the source) while in 3D this factor is $1/r^2$. In a heterogeneous medium the wave front deviates from circular or spherical (see Fig. 3) and the most accurate method would estimate the local wave front curvature and then scale the local attenuation accordingly. However, we found that a simple scaling by a distance measure was sufficient in our case. Further, we do not perform such corrections in the back-projection phase, referring to Ref. [22] who promote the use of unmatched projector/backprojector pairs.

Figures 4 and 5 list the SS and SA reconstruction algorithms, respectively. Figure 6 illustrates the grid update along a refracted ray. The grid correction delivered by a given ray is distributed into the grid voxels using the weights of a B-spline interpolation filter centered at sample points along the ray.

In the following we assume spherical voxels or *blobs.*[34]

```
Initialize sound attenuation image SA
Until convergence, loop
    Until all emitters have been chosen
        Randomly select an emitter E
        Load E's acquired attenuation receiver data SA_acquired
        Load E's rp data from final SS reconstr. iteration
        For each receiver R_i
            Trace rays from R_i to E using rp_i → SA_simulated
            Compute correction factor as
                ΔSA = ( SA_acquired - SA_simulated )/L
        Back project ΔSA along rp
        Update SA estimate by ΔSA at each rp voxel
```
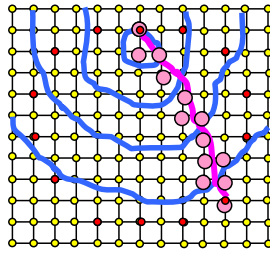
FIG. 5. The FMM-SART algorithm for sound attenuation.

FIG. 6. Updating the grid voxels along the path *rp* of a refracted ray.

Initialize sound speed volume *SS*
Until convergence, loop
    Until all slices in *SS* have been chosen
        Randomly select a slice in *SS* → central slice *CS*
        Load *SS* estimate of slices peripheral to *CS* → *PS*
        Until all emitters in *CS* have been chosen
            Update *SS* estimate of all voxels in *PS*

FIG. 7. The FMM-SART algorithm for sound speed reconstruction in a limited multirow detector scenario.

We match this representation by using spherical (radially symmetric) B-spline interpolation filters. Spherical interpolation filters result in a spherical frequency response which makes both the interpolation and the reconstruction result orientation independent.

For the SS update step, we use the relationship $SS = d/t$, where d is the diameter of a spherical voxel (we use $d=1$, that is, one grid spacing). The following voxel update equation arises:

$$SS^{(k+1)} = \frac{d}{d/SS^{(k)} + \sum_{i}^{n_{rays}} \Delta TOF_i/n_{rays}}. \quad (9)$$

Here, n_{rays} are the number of rays updating the given voxel. One SS (SA) iteration completes after all emitters have been processed once, and the iterations continue until the difference between the simulated TOF (SA) data at the receiver positions and the acquired TOF (SA) data there becomes smaller than a threshold. In our experiments, convergence typically requires three to four iterations.

While using the notion of "image," our discussion so far is valid for both 2D slice stack and fully-3D UCT reconstruction. The former acquires the data with a single transducer ring stepping along the axis of the breast, while the latter uses the cylindrical multirow transducer assembly axially covering the entire breast. Previous work has attempted to reconstruct these data in serial 2D, ignoring 3D refraction effects. To advance these works we have implemented a fully 3D version of FMM. However, in practice we find that performing the FMM in fully 3D is computationally rather expensive. Furthermore, it also requires a significant amount of memory to store the heap with the (3D) narrow band.

As an alternative, for these computational reasons, we have investigated reconstruction within a moderately thick volumetric section. The advantage of this scheme is that it allows us to reduce full 3D ray tracing and front tracking to a narrow neighborhood of the plane in which the transducers reside. We currently use a section width of 10%–20% of the full volume height. Given the limits on the variation in propagation velocities and the geometric arrangement with the volume, this section width appears sufficient. If there are large refraction effects, however, the resulting image will be less accurate.[35]

In this *locally* fully 3D scheme we acquire data within a short multirow assembly in which solely the central ring transducers act as emitters as well, and then perform reconstruction only within the covered volume slices for each such axial central ring position. In order to capture the out-of-plane refraction effects properly, we increase the resolution along the transducer stack axis, currently by a factor of 3. That is, in addition to the original slices spanned by each transducer ring, for each two consecutive such slices, we maintain two further slices located in-between, spaced equidistantly.

The corresponding reconstruction algorithm for the SS volume is outlined in Fig. 7. The reconstruction of the SA volume proceeds in a similar manner. The wave front and rays are now traced in 3D, within the short stack of slices. The algorithm updates slice voxels several times per iteration, that is, each time they fall within the range of the short multidetector stack.

III.C. Regularization

We observed in our experiments that the first sequence of corrective updates is crucial in achieving a faithful object recovery from the data. We found that the overall shape and geometry (the SS distribution) of the object is often formed quickly, but the SS values themselves are low.[36] This effect reduces resolved detail and contrast in the simulated projections, which in turn has a negative impact on the detail and contrast delivered in the corrective updates. The algorithm then often drives the solution estimate away from the true desired solution estimate, leading to significant distortions. Thus, we aim to "boost" the SS iterate prior to the forward projection, to the effect that the more pronounced reconstructed detail creates more contrast in the wave propagation, producing a more detail-sensitive TOF image and corresponding simulated projection. To achieve this effect, we explored two different strategies: (i) *fixed scaling* of the corrective updates and (ii) *state-driven scaling* of the corrective updates. Both methods use two separate SS images for projection and backprojection—we shall call them SS and SS_{bp}, respectively. SS_{bp} is obtained as usual via Eq. (8), but the result is then translated into SS_p in the following manner. For the fixed scaling we simply use $SS_p = aSS_{bp}$, where a is a constant. On the other hand, the idea behind the state-driven scaling is to boost the contrast of the reconstructed SS image to the contrast expected in the final reconstruction. Since we do not know the pixel-level contrast prior to reconstructing the object, a more accessible model available is the expected value ranges, that is, the minimum and maximum of these values. These can be obtained from tables (see Table I which lists values for both sound speed and attenuation). Once ob-

TABLE I. Acoustic properties of the female human breast. Outside the object we set the sound attenuation to nearly zero.

Ultrasound properties	Tissue	Fat	Large lesions	Small lesions	Skin
Sound Speed (in m/s)	1475	1375	1560	1530	1680
Attenuation (insertion loss, in %)	60	20	70 or 40	70	50

tained, we use the following equation to scale the reconstructed SS_{bp} values at iteration k into the SS_p values used in the subsequent projection at iteration $k+1$:

$$SS_p^{(k+1)} = SS_{exp,min} + a(SS_{bp}^{(k)} - SS_{bp,min}^{(k)}),$$

$$a = \frac{SS_{exp,max} - SS_{exp,min}}{SS_{bp,max}^{(k)} - SS_{bp,min}^{(k)}}. \tag{10}$$

Here, $SS_{exp,min}$ and $SS_{exp,max}$ are the tissue values obtained from the tables, and $SS_{bp,min}^{(k)}$ and $SS_{bp,max}^{(k)}$ are the minimum and maximum values in the present SS reconstruction. In this mechanism the scale factor a is initially large and then tends to 1.0 as the values in the reconstruction grow to the expected tissue values. We note that the same mechanism is also used in the attenuation image reconstruction, just replacing the SS with SA values in Eq. (10).

Our regularization scheme will benefit from an accurate assessment of $SS_{exp,min}$ and $SS_{exp,max}$, which can vary as a function of breast density. For this, it may be possible to gauge this density beforehand, by classifying the subject by age, prior physical exams, or information gained in prior scans (including x-ray). Alternatively, one may be able to obtain some estimates during the ongoing reconstruction, making the proper adjustments on the fly as the iterations proceed. Since we use a GPU-accelerated implementation,[37] a few extra iterations will not be overly critical to time performance. Both methods, user assisted or not, could be applied in conjunction with a known statistical distribution in min/max values of breast tissue in women to estimate the most likely setting for the case at hand. This is subject of further investigations.

IV. UCT BREAST PHANTOMS

We tested our UCT reconstruction framework with two phantoms. The first is a relatively simple traditional 2D phantom, composed of various circular primitives with different densities that mimic the values typically found in a breast. Its quantitative parameters are given in Table I. The sound speed phantom is shown in Fig. 8(a), and the corresponding attenuation phantom is shown in Fig. 8(f). Both share the same geometry and feature various lesions of diameters ranging between 2 and 8 pixels. We note that for all figures presented throughout this paper, the calibration bar is only valid for the focus object—the black embedding background is only provided for contrast.

The second phantom is aimed to be more realistic. While in x-ray mammography a number of breast phantoms have been defined,[38,39] these are typically too complex and some of the structures are not well differentiated by ultrasound or

are irrelevant for this modality. We therefore designed a specific digital anthropomorphic breast phantom for ultrasound. We used the Visible Female cryosection data set (NIH National Library of Medicine)[40] since it contains all the structures present in real breast tissue. We then augmented the breast region of this data set with cancerous lesions of various sizes, shapes, densities and contrasts. The cryosectional data of the visible female has 5189 slices of 4096×2700 RGB pixels each. To construct our anthropomorphic breast phantom, we began by extracting the breast portion from the data set, which yielded 74 slices of 280×280 pixels each (a spatial resolution of 0.33×0.33 mm). One such slice is shown in Fig. 9(a)—we observe that fat, lobules (tissue), and ducts are visually well distinguished. We then mapped the RGB values to the corresponding ultrasound property values in a three-step procedure. First, we transformed the data from RGB to HSV (Hue, saturation, value). Since we found that materials in the breast of the visible female can be differentiated very well by Hue (H), we subsequently used the
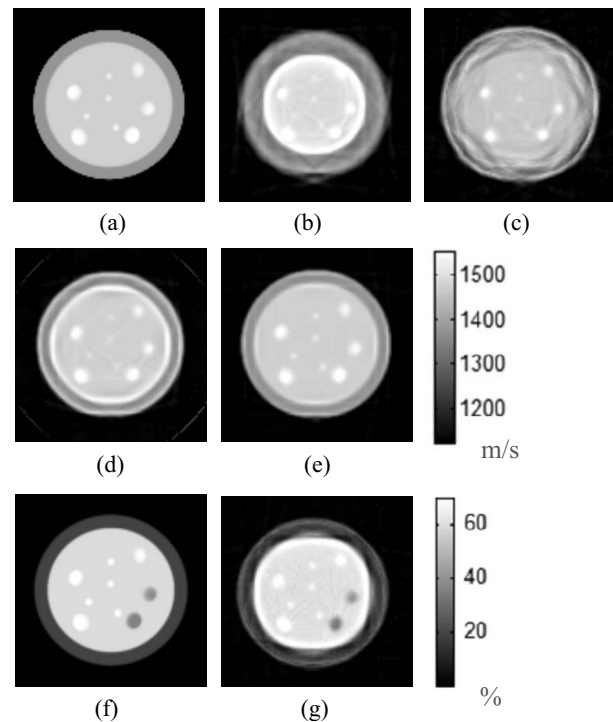

FIG. 8. Experiments with a simple phantom. (a) Original sound speed phantom; (b) SS reconstruction with straight rays, fixed scaling; (c) SS reconstruction with curved rays, HAFMM, fixed scaling; (d) SS reconstruction with curved rays, FMM, state-driven scaling; (e) SS reconstruction with curved rays, HAFMM, state-driven scaling; (f) original attenuation phantom, (g) reconstructed attenuation image with curved rays, HAFMM, and state-driven scaling.
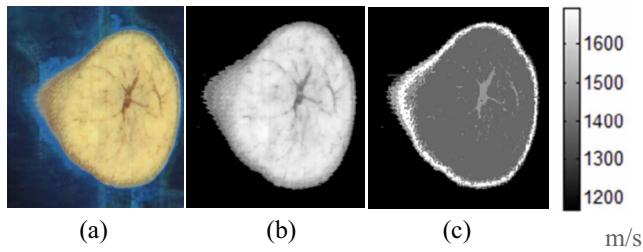
FIG. 9. A slice of the breast portion of the Visible Female data set. (a) A cryosection RGB slice; (b) the same slice with the *H* value transformed to gray scale; (c) the slice after the mapping to acoustic speed (see Table I).

H-channel only [see Fig. 9(b)]. We then found the mean and the range of *H* for the different materials by statistically sampling the corresponding regions (skin, fat, tissue, ducts). Using the ultrasound density values for these materials (Table I), we created a discrete mapping from *H* to these ultrasound densities. Here, for each material, we applied a Gaussian kernel with its mean as the density distribution center and the range as its sigma. When the Hue value fell into the tails of the Gaussian kernel functions of multiple materials, the material's kernel function with higher value was chosen in order to maintain continuity. When the Hue fell out of any range of the Gaussian kernel function of materials, it was mapped to the background value. Figure 9(c) shows a slice of the resulting continuous transformation from the cryosection data to ultrasound properties.

V. RESULTS

We studied our framework with these two phantoms. In these efforts we focused primarily on reconstruction quality, but we will show that the FMM provides a very cost-efficient means for achieving high-fidelity UCT projections and reconstructions. A companion publication[37] has explored the computational efficiency further, describing a GPU-accelerated framework that was able to achieve an 80-fold speed-up over an optimized CPU implementation, without any loss in reconstruction quality. It allowed us to reconstruct an object of size $256^2 \times 44$ with 512 transducers per ring in less 5 min, which is a time that can easily satisfy real-life clinical scenarios. The same task took over 5 h on a CPU.

V.A. Suitability of the FMM and the proposed regularization

The first series of experiments is based on the simple digital phantom (size 128×128 pixels). Our simulation assumes 256 emitter/receiver positions along a circle, as discussed before. The resolution of the reconstructed image matches the size of the phantom images. The reconstruction took about 60 s for three iterations using a 2.8 GHz Pentium 4 CPU. The data were obtained using a high-quality full wave equation solver (further described below) which operated on a significantly upsampled version of the phantom (2048×2048 pixels) in order to gain a realistic near-continuous scale simulation.

For the sound speed phantom shown in Fig. 8(a), the reconstruction results are given in Fig. 8(b)–8(e). The reconstructed image obtained with straight (linear) rays is shown in (b) and reconstructions obtained with curved (nonlinear) rays are shown in (c) and (e). For the curved ray reconstruction, we further compare the results obtained for the different regularization strategies: Fixed scaling in (c) and state-driven scaling in (e). We calculate the error as the average squared intensity difference (the RMSE) for all grid points (inside the phantom area) between the reconstructed image and the original phantom. We observe that the use of straight rays distorts the image with an error of 25%. Because refraction is ignored, the size of the phantom's regions grows or shrinks dramatically, which has been reported before. However, we also find that when correcting for refraction (using curved rays) the regularization method has a large impact on the resulting image accuracy. We observe that our state-driven scaling results in a significantly better estimate of the original image. It has an error of 3%, while the fixed speed strategy overcorrects for refraction and results in an error of 19%.

The qualitative improvement that can be obtained with the HAFMM method (over the standard FMM) is significant as well. The image computed with FMM, shown in Fig. 8(d), has poor boundary delineation (error of 10%), while the HAFMM, shown in Fig. 8(e) does much better in that regard. HAFMM also more accurately preserves detail, for example, the interior lesions are much rounder than with the FMM. The improvement in accuracy of the HAFMM is due to the inclusion of second-order curvature information when solving the Eikonal equation, which means that boundaries, at which refraction occurs, are better defined and determined.

Finally, the reconstruction results for the attenuation phantom in Fig. 8(f) are presented in Fig. 8(g). The attenuation image reconstruction is based on the curved ray paths determined in the SS phantom reconstruction shown in Fig. 8(e). We see that the reconstruction recovers both intensity and shape accurately, with an error of less than 7%.

V.B. Comparison with full wave equation solver

We compared our HAFFM ultrasound simulation with a more traditional (but computationally much more expensive) method which solves the wave equation via a PDE-solver (for the 2D case). This solver takes a pseudospectral approach,[41] where the spatial derivatives were obtained using the fast Fourier transform, while the temporal propagation is performed using a fourth-order accurate Adams–Bashforth explicit scheme. Our approach, on the other hand, advocates an alternative way, rooted in geometric acoustics, which solves the curved-ray problem by directly simulating the acoustic sound wave propagation using the HAFMM.

For our comparative study, we collected time-of-flight (sound speed) simulation data within a circle of sensors, again using the phantom of Figs. 8(a) and 8(f), for both the HAFMM-based and the PDE-based solver (the latter used the high-resolution version of the phantom as discussed above). We observed an excellent correlation between the two simulation data sets, with R^2 greater than 99%, which
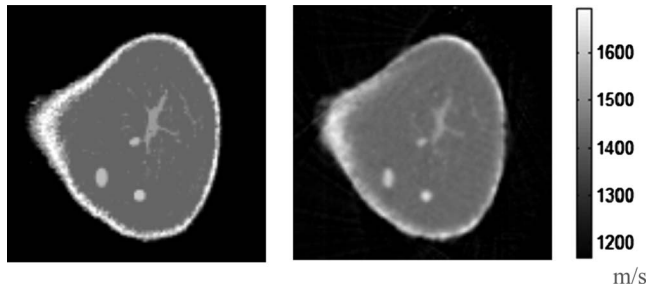
FIG. 10. Our digital anthropomorphic breast phantom augmented with three lesions. (a) Original; (b) speed of sound reconstructed from 280 transducers. The size of the lesions ranges between 6 and 9 pixels with densities from 100 to 250 (for tissue properties, see Table I).

demonstrates that our HAFMM method is in excellent linear agreement with a physically accurate PDE solver. Likewise, we also compared the two solvers in terms of the ratios of the estimated attenuation due to just water vs. the estimated attenuation with the object immersed. Here we obtained R^2 =0.95. Both of these results demonstrate that for the special case of ultrasound breast mammography, the HAFMM produces accurate estimates of both time of flight and tissue attenuation in a heterogeneous material setting. It, however, does so at a great speed advantage. For example, for the wave propagation simulation in a 512×512 image, our HAFMM-based method takes only minutes, while the PDE-solver method requires several hours to complete. These speed gains are very important in a practical iterative reconstruction framework.

V.C. Influence of algorithmic parameters

In this section, we explore the influence of the various algorithm parameters and their settings on reconstruction quality. For all studies in this section (and the ones following), we use the digital anthropomorphic breast phantom described in Sec. IV, with three lesions of different sizes and densities. A slice and its UCT reconstruction are shown in Fig. 10. The data for this phantom were obtained using a high-accuracy nonlinear ray-tracing algorithm.

V.C.1. Number of transducers

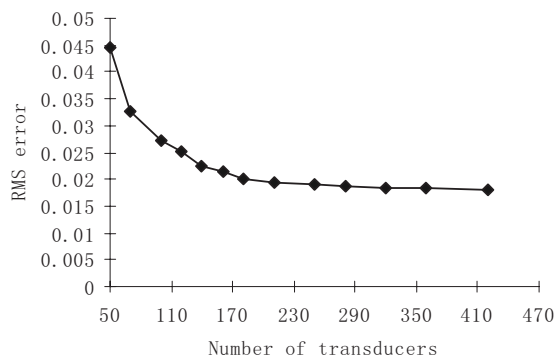Figure 11 compares the errors obtained as a function of

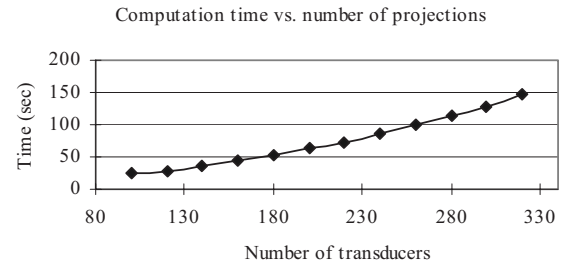FIG. 11. Number of transducers vs. reconstruction error.

FIG. 12. Number of transducers vs. computation time.

the number of transducers, N_t. We use one slice of the phantom of size 140×140 pixels. When N_t is half that size, that is, 70, aliasing artifacts are noticeable in the reconstruction. With N_t doubling to 140, the error drops sharply by 47%. As N_t continues to double (to 280), the error drops further by 18%. As N_t increases further the improvements are marginal.

Next, Fig. 12 plots the cost of computation vs. N_t. We observe that the computational cost increases with the number of transducers. A trade-off between quality and computational cost is always to be made. In this particular case, to achieve a good balance between accuracy and computation time, the optimal number of transducers is 280, which is about double the grid size. It gives a reconstruction error of 1.8%, at a computational cost of 112.5 s [see also Fig. 10(b)]. We use this configuration for all experiments presented. This configuration results in a clear reconstruction of inner structures and lesions, at a reasonable reconstruction time.

V.C.2. Additional transducer ring positions

Sometimes one may not be able to afford more transducers in a ring, but one may attempt to gather additional data by simply rotating the ring by a percentage of the equiangular distance between the individual transducers. We investigated whether such a strategy can produce improved reconstructions. For this, we added to each original data set additional sets of observations, based on one or two interspaced incremental rotational ring displacements. The results of this experiment are shown in Fig. 13. We tested four configurations using 50 (experiment ID 1–3), 70 (ID 4–6), 140 (ID 7–9), and 240 transducers (ID 10–12). The first measurement in each series is due to the original configuration, without using extra data due to ring displacement, while the other two use additional data due to one or two such displacements, respectively. We observe that such a strategy does not produce better reconstructions. It appears that adding additional ring positions to a fixed number of transducers does not add much independent information. We therefore conclude that a better correction of the SS and SA estimates requires a denser set of reconstruction rays for a single emitter position, and not multiple emitter positions.

V.C.3. Relaxation coefficient

The relaxation coefficient α is an important parameter for SART. It provides a means to reduce (that is, smooth) the influence of noise in the data, but there is a trade-off with
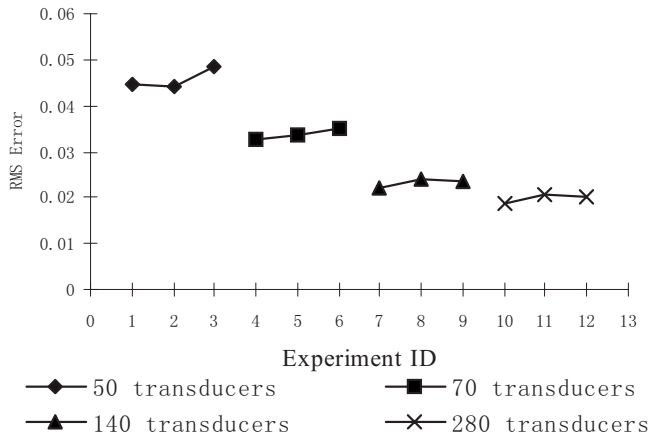
FIG. 13. The effect of acquiring additional data by incremental rotational transducer ring displacement, acquiring data at intermediate angular positions without increasing the number of transducers. Each triple has the same number of transducers—only the number of intermittent positions increases from left to right for each. Doing so does not significantly reduce reconstruction error, if at all.

speed of convergence. In x-ray CT with limited data typically a value in the neighborhood of 0.1 is chosen for α. In the case of FMM-SART, Fig. 14 shows that if α is greater than 0.1, the error of the reconstruction image can increase dramatically as iterations continue. Such growth is often caused by overcorrection. For FMM-SART, the refraction effects that give rise to curved rays and which in turn cause nonuniform grid updates require a less aggressive relaxation schedule than when linear rays are used. Therefore, smaller α are needed. We find that the best reconstruction results are achieved when α is between 0.1 and 0.05. The lowest error occurs at about three to five iterations. Of course, in practice the optimal α-setting depends on many parameters, such as number of transducers, noise in the data, image size, and the like. To that end, our study only seeks to demonstrate the trends that come with the choice of α.

V.C.4. Interpolation filter

While using an interpolation filter of good quality is always important in iterative reconstruction,[42] it is even more important when one seeks to estimate gradients from discrete data. Earlier work has shown[43] that the sensitivity to suboptimal interpolation increases with the order of the estimated derivative. This sensitivity is especially crucial in iterative reconstruction tasks, since interpolation errors committed in the forward projection get fed back into the corrective updates which can drive the reconstruction away from the desired result. Linear and cubic convolution filters are widely used in CT reconstruction. However, in previous work[33] we have shown that the B-spline filter family exhibits better performance, in particular, in the presence of refractions. Although B-spline filters are slightly more computationally expensive, their benefits are often worth this expense. The B-spline filter is a noninterpolating filter. This provides more freedom in the filter design since the kernel does not need to be zero at integer positions. First, coefficients are calculated
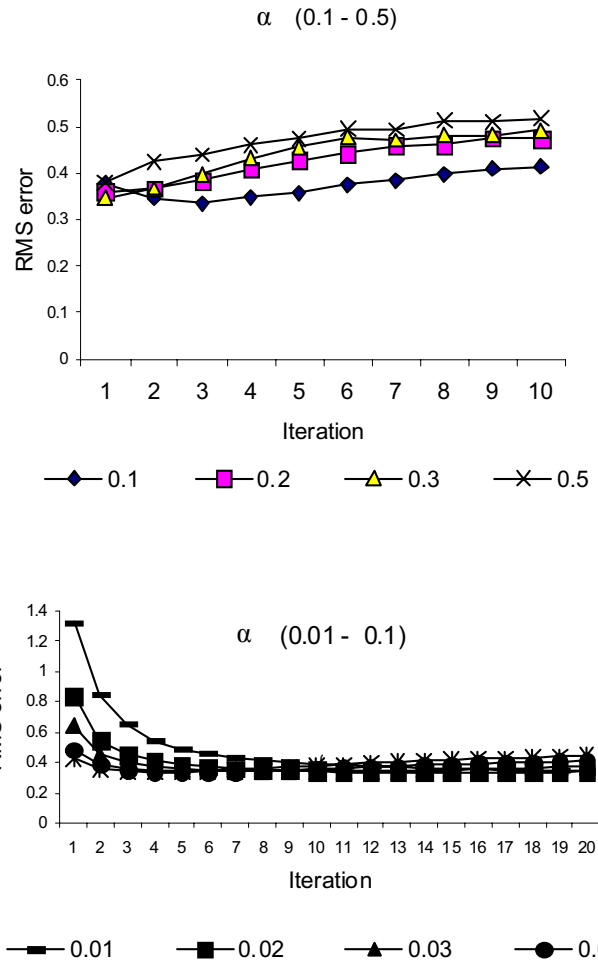

FIG. 14. The effect of the relaxation coefficient α on reconstruction quality.

from the original signal at each grid point. This coefficient field emphasizes high-frequency information in the signal, that is, it oversharpens (high passes) it. The subsequent interpolation by the B-spline kernel during the ray tracing then provides the smoothing, (low passing) that neutralizes the sharpening in the prior coefficient calculation step. This pipeline results in a very high-quality interpolation, with only very little smoothing and aliasing remaining. In Fig. 15, we plot the error vs. the number of iterations for various filters. We observe that for the linear and cubic convolution filters, the error drops initially as the number of iterations grows, but then starts to increase as iterations continue. The B-spline filters, on the other hand, exhibit a monotonic reduction in reconstruction errors, and we do not observe the divergence trends of the other filters at later iterations. This result, in fact, is an excellent demonstration of the impact of interpolation filter quality on reconstruction convergence behavior. In our particular application, that is, breast phantom reconstruction, we also find that the B-spline gradient interpolation filters achieve significantly more accurate results than the traditional linear and cubic convolution filters. Further, we find that the less expensive second-order B-spline (B-spline 2) achieves a similar quality as the more expensive third-order B-spline filter (B-spline 3). Using the B-spline 2
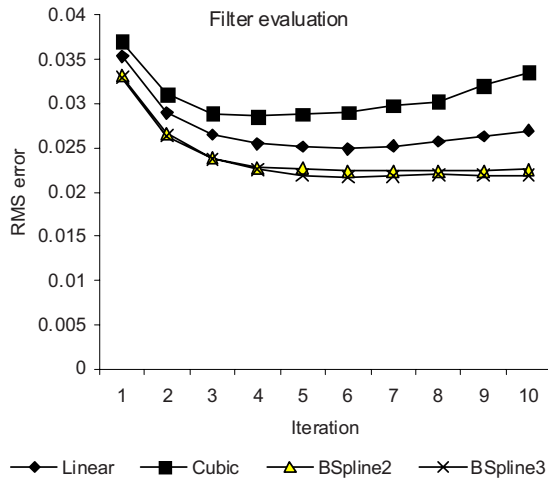
FIG. 15. The effect of the TOF gradient estimation filter on reconstruction error.

filter saves time, since the computational cost of the B-spline 2 filter is about 1/4 less than that of the B-spline 3.

V.D. Detectability, resolution, and noise tolerance

We now discuss (i) the framework's spatial coverage by the bent rays, (ii) its spatial resolution determining the size and spacing of detectable lesions, (iii) its ability to distinguish among various diagnostic shapes, and (iv) its noise behavior.

Since the rays are curved there is a chance that some regions are never covered during the corrective updates, even when the transducers are sufficiently close on the periphery. In this case we would be bounded to overlook important features. In the 3D wavefront construction method by Vinje *et al.*,[44] rays emitted from the source form the vertices of an advancing triangular mesh, which is synonymous to the wave front. Then, if the side length of one of these triangles exceeds a certain threshold the mesh is subdivided at that location and a new ray is spawned. We performed experiments to see if such an approach was needed in our tomographic reconstruction scenario, or if our current raytracing approach was sufficient. In this context, we should note that our forward projection using the FMM is in fact space cov-

ering. The FMM is an implementation of Eikonal wave tracking, which represents an alternative approach to mesh tracking in Vinje *et al.*

In order to verify the complete space coverage of the collection of curved rays we visualized the refraction effects of the various wave propagations (see Fig. 16). First, Fig. 16(a) shows the wave front that bends forward or backward at refractive interfaces. As a consequence, the rays from the emitter to the receivers deviate, as is shown in Fig. 16(b). This bending results in emptier spaces through which fewer rays pass—for this emitter. In Fig. 16(c) we show the image due to the interpolation filter weights, collected from the rays in one grid update. This image is used to normalize these grid updates [compare Eq. (5)]. The brighter the points in the weight images, the more rays have passed through their neighborhood. Finally, Fig. 16(d) shows the accumulated weight image due to all projection directions. One can easily see that all grid points receive some amount of contributions, and thus no areas exist that remain unknown. In fact, we can also observe in these weight images the structure and the shape of the phantom and its interior lesions.

Next we evaluate two important metrics: (i) lesion detectability and (ii) lesion detail preservation, quantifying spatial resolution.

V.D.1. Detectability: Lesion size

It is important to faithfully detect lesions of small sizes at an early stage and at high fidelity. Figure 9 shows that traditional straight ray methods tend to produce a great amount of distortions of small and large structures. Our framework captures the impact of refractions and effectively corrects the deformation. The images in Fig. 17(a) and 17(b) show the original phantom and its reconstruction, augmented with various lesions of different sizes. We see that FMM-SART clearly identifies lesions even at a size as small as 2 pixels in radius, at clinically relevant contrast levels.

V.D.2. Resolution: Lesion spacing and shape

Lesions can grow in the vicinities of others. For accurate diagnosis, a cluster of small lesions has to be reconstructed with each individual lesion having the correct size and shape, such that it can be differentiated from a large lobulated lesion. To test this ability, we inserted two lesions separated by

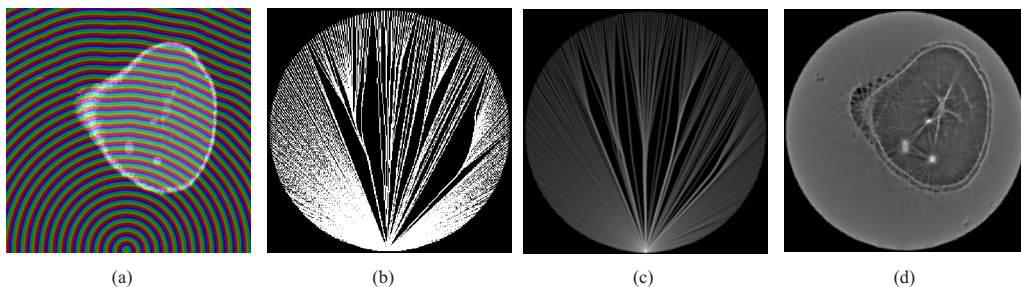(a)   (b)   (c)   (d)

FIG. 16. Wave propagation and ray tracing. (a) Wave front (TOF) image for FMM; (b) rays traced using the TOF image; (c) filter weights of corrective grid updates (backprojections) for one emitter; (d) weight image accumulated for all backprojections.
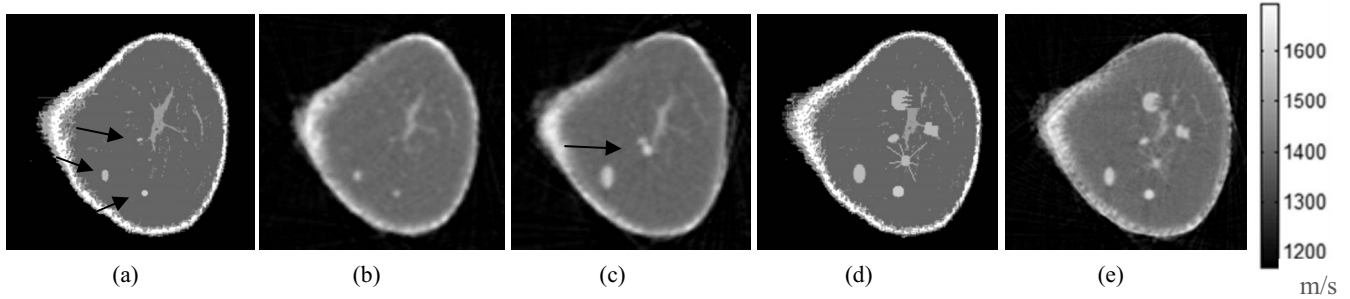
FIG. 17. Reconstructed images of lesions with different sizes. The smallest lesion is 2 pixels in radius. (a) Phantom with various small lesions, pointed to by arrows; (b) speed of sound reconstruction.

a small distance (1–2 pixels) into the phantom. The reconstructed image (Fig. 17(c)) clearly reveals the tiny gap (pointed to by the arrow) between these lesions and correctly identifies them as two lesions, instead of a single large one.

The shape of lesions is another critical factor to help diagnose the properties of lesions. Benign lesions may be sharp and circumscribed, while malignancies may have spiculated, lobulated/undulating shapes. Circumscribed margins are sharply demarcated with an abrupt transition between the lesion and surrounding tissue, with nothing to suggest infiltration. Spiculated margins, on the other hand, are suggestive of surrounding parenchymal infiltration.[45] Deformation caused by refraction effects makes it hard to reconstruct the original lesion shape faithfully. For example, for rays reaching one lesion, some of them need to traverse other lesions before reaching the lesion, while other rays enter through surrounding tissue directly. In straight ray methods, different ultrasound speeds in lesions and tissue may affect the calculation and make the reconstructed image of part of the lesion appear dilated and deform the shape. Our curved-ray method overcomes these challenges by taking into consideration the speed variation in different materials and thus correctly calculates their deflected trajectories. We simulated spiculated, lobulated, and obscured lesions in our phantom. Figure 17(d) and 17(e) compares these simulated lesions with their reconstruction using our UCT framework. We observe that our method can successfully reconstruct the different lesion shapes (the reconstruction error is about 4%).

This ability of FMM-SART to separate small lesions 1–2 pixels apart (Fig. 17(c)) and to resolve small lesion detail at similar granularities (for example, the comblike structure in the upper third of Fig. 17(d) and 17(e)) demonstrates its excellent spatial resolution. Further experiments with physical simulation data will be needed to verify these findings for use in clinical practice. This is subject of ongoing work.

V.D.3. Noise tolerance

In practice, ultrasound projections tend to be noisy—the signal-to-noise ratio (SNR) can range within 100 to 10. We tested our algorithm by adding Gaussian noise into the projection results, using SNR levels at the lower end of this range, 5 to 20, in order to gauge the robustness of our framework. The resulting reconstruction errors are plotted in Fig. 18. We see that only for SNR=5 the reconstruction strongly diverges at an early stage. The images in Figs. 19(a)–19(d) show the reconstructed results for SNR=5, 10, 15, and 20. We observe that even when the SNR is as small as 5 and the reconstructed image is quite noisy, the inner lesions and overall structure are still visible. With the SNR increasing to 10 (still a significant noise level) the inside lesion can be well distinguished from the surrounding tissue.
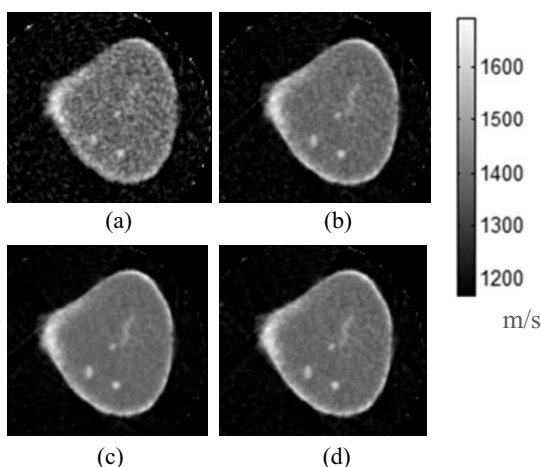

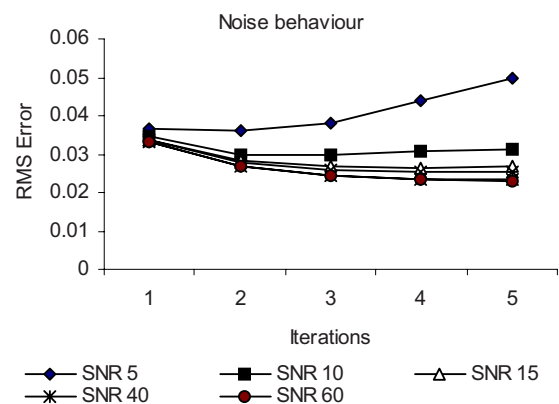FIG. 18. Impact of noise on projection on reconstruction quality.

FIG. 19. Sensitivity to noise speed of sound images reconstructed from data with (a) SNR=5, (b) SNR=10; (c) SNR=15; (d) SNR=20. We observe that even for SNR=5 the lesions can be well recovered.
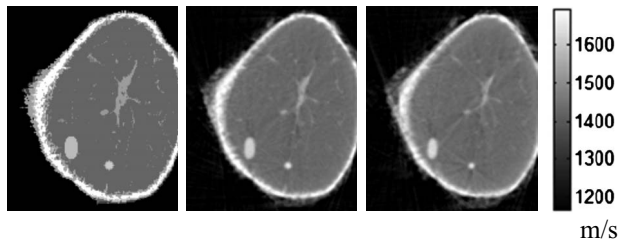
FIG. 20. Comparing 2D with short-stack 3D reconstruction. (a) A slice of the phantom; (b) 2D reconstruction; (c) the same slice in a 3D short-stack reconstruction.

V.E. Short-stack 3D reconstruction

The reconstructions presented so far have been on a slice-by-slice basis, that is, the data have been acquired (simulated) with a single row transducer ring. The remaining discussion now presents the quality gains that can be achieved by acquiring (simulating) data with a short multirow transducer assembly and performing the reconstruction as outlined at the end of Sec. III B. We currently assume a short-stack size of six transducer rings, which is 14% of the 44 slices reported in our companion paper.[37] As discussed in Sec. III B we maintain two additional slices between each such ring pair. Thus the total number of slices reconstructed within one such short stack is 16. Figure 20 presents one of these slices, reconstructed both in 2D [Fig. 20(a)] and with the short-stack 3D approach [Fig. 20(b)]. We observe some improvements in feature definition, but both reconstructions can resolve all required detail satisfactorily. The RMSE error for the 2D slice reconstruction is 3.6%, while the RMSE error for the 3D approach is 3.4%. The computation time is proportional to stack size.[37]

VI. DISCUSSION

To bring our work into the proper clinical perspective, we need to relate the wavelength used in the simulation with the reconstruction's resolution, in the context of the female human breast. A typical breast is about 110–140 mm wide at its base, and thus at our current slice resolution of 140^2 the size of a pixel is 1 mm. Figure 17 indicates that our method is able to detect features with a diameter of 2–4 pixels, which corresponds to a size of 2–4 mm. As lesions less than 5 mm are considered small, this is an encouraging result. For smaller lesion sizes, we will need to reconstruct the object on a finer grid, and for this we will need more transducers to ensure sufficient angular sampling (compare Fig. 11). A natural limit here is imposed by the frequency of the transducers. Our physics-based simulation uses a center frequency of $\omega = 5$ MHz. With a sound speed v in breast tissue of around 1500 m/s, this gives us a lower bound on the size of the features that can be faithfully reconstructed of around $\lambda/2 = 0.5 \cdot v/\omega = 0.15$ mm. However, such a resolution is beyond the reported capabilities of ray theory-based reconstructions, which is more on the scale of $2-3 \cdot \lambda$ or 0.6–0.9 mm in this case.

With this in mind, we are encouraged by our finding (Sec.

V B) that the HAFFM is capable of performing *forward simulations* (predictions) of the TOF through a complex structure with accuracy comparable to a numerical solution of the full wave equation. The PDE solution was discretized to accurately describe scattering from features down to at least about $\lambda/2$. Therefore, the HAFFM has the potential of performing accurate forward predictions through a medium with features down to about 0.15 mm. However, this does not necessarily translate into the ability to perform *inversion* at this accuracy. Since inversion is an ill-posed problem, accurately solving the forward problem does not guarantee similar accuracy also in the inversion. This is subject to further study.

VII. CONCLUSIONS

We have presented a novel iterative framework termed FMM-SART for CT reconstruction from transmission ultrasound data. Our approach is rooted in geometric acoustics and, in contrast to most other systems of its kind it accurately takes into account the strong refraction effects that occur in the female breast. We have shown that breast lesions of even small sizes and distances can be well reconstructed even in the presence of significant noise. Our method also differentiates lesion shape quite well which is important in cancer diagnostics. Decisive factors in achieving these promising results are, apart from the accurate modeling of curved rays, a novel regularization scheme and a high-quality interpolation filter. An added benefit of our refractive ray tracing framework is that its foundation in the fast marching method allows for an efficient implementation on the CPU and also accelerates well on commodity multiprocessor architectures, such as the GPU. With the latter, we have shown that clinical reconstruction speeds are possible, requiring less than 5 min to reconstruct a $256^2 \times 44$ volume. Further contributions are the development of a realistic anthropomorphic breast phantom based on the Visible Female data set.

In this paper, we aimed to demonstrate the feasibility of our framework for breast transmission ultrasound CT and make contributions with respect to its accuracy (both gradient interpolation and wave evolution) and efficiency (both algorithmic and computational). We further discussed the impact on feature resolution and tolerance to noise. We find that given physics-based simulated data, our method is able to detect features at clinically relevant contrasts with good preservation of detail. While it is too premature to claim (at this stage) that this would also be the case *in vivo*, we do believe that our results are quite encouraging in that respect.

In our setup, we calculate attenuation (insertion loss) for the central frequency and the parameters we reconstruct are with respect to this frequency. Additional information regarding tissue properties may be obtainable by measuring at multiple frequencies and we intend to investigate this in the future. As is typically done, we assume constant density; variation in tissue density is more difficult to measure.

Finally, our simulations and the subsequent comparative study with the full wave equation solver have only been in 2D. While we do not expect the outcome to change dramati-

cally in 3D, since the TOF-simulations should still be valid, the reconstruction of attenuation may possibly degrade a bit due to the fact that interference effects could be more important. We aim to study these 3D effects in the future.

ACKNOWLEDGMENTS

[a]Electronic mail: mjack@ime.usp.br
[b]Electronic mail: dione@quixnet.net
[c]Electronic mail: trond.varslot@anu.edu.au
[d]Electronic mail: lawrence.staib@yale.edu
[e]Author to whom correspondence should be addressed. Telephone: 631-632-1524; Fax: 631-632-8445; Electronic mail: mueller@cs.sunysb.edu

[1]I. Christoyianni, E. Dermatas, and G. Kokkinakis, "Fast detection of masses in computer-aided mammography," IEEE Signal Process. Mag. **17**, 54–64 (2000).

[2]B. Drukker, "Breast disease: A primer on diagnosis and management," Int. J. Fertil. **42**, 278–287 (1997).

[3]N. Duric, P. Littrup, A. Babkin, D. Chambers, and S. Azevedo, "Development of ultrasound tomography for breast imaging: Technical assessment," Med. Phys. **32**, 1375–1386 (2005).

[4]B. Chen and R. Ning, "Cone-beam volume CT breast imaging. Feasibility study," Med. Phys. **29**, 755–70 (2002).

[5]T. Wu, R. Moore, E. Rafferty, and D. Kopans, "A comparison of reconstruction algorithms for breast tomosynthesis," Med. Phys. **31**, 2636–2647 (2004).

[6]S. Feig, "Role and evaluation of mammography and other imaging methods for breast cancer detection, diagnosis, and staging," Semin. Nucl. Med. **29**, 3–15 (1999).

[7]J. Greenleaf and R. Bahn, "Clinical imaging with transmissive ultrasonic computerized tomography," IEEE Trans. Biomed. Eng. **BME-28**, 177–185 (1981).

[8]H. Azhari and D. Sazbon, "Volumetric imaging with ultrasonic spiral CT," Radiology **212**, 270–275 (1999).

[9]N. Duric, P. Littrup, L. Poulo, A. Babkin, R. Pevzner, E. Holsapple, O. Rama, and C. Glide, "Detection of breast cancer with ultrasound tomography: First results with the Computed Ultrasound Risk Evaluation (CURE) prototype," Med. Phys. **34**, 773–785 (2007).

[10]J. Jeong, T. Kim, D. Shin, S. Do, M. Singh, and V. Marmarelis, "Soft tissue differentiation using multi-band signatures of high resolution ultrasonic transmission tomography," IEEE Trans. Med. Imaging **24**, 399–408 (2005).

[11]C. Pintavirooj, A. Romputtal, A. Ngamlamiad, W. Withayachumnankul, and K. Hamamoto, "Ultrasonic refractive index tomography," J. Int. Conf. Central Europe on Computer Graphics, Visualization and Computer Vision **12**, 133–139 (2004).

[12]Y. Quan and L. Huang, "Sound-speed tomography using first-arrival transmission ultrasound for a ring array," Proc. SPIE 6513 (2007).

[13]M. Andre, H. Janee, P. Martin, G. Otto, B. Spivey, and D. Palmer, "High-speed data acquisition in a diffraction tomography system employing large-scale toroidal arrays," Int. J. Imaging Syst. Technol. **8**, 137–147 (1997).

[14]K. Ladas and A. Devaney, "Application of an ART in an experimental study of ultrasonic diffraction tomography," Ultrason. Imaging **15**, 48–58 (1993).

[15]X. Pan, "Unified Reconstruction theory for Diffraction Tomography, with Consideration of Noise Control," J. Opt. Soc. Am. A Opt. Image Sci. Vis. **15**, 2312–2326 (1998).

[16]H. Pratt, L. Huang, N. Duric, and P. Littrup, "Sound-speed and attenuation imaging of breast tissue using waveform tomography of transmission ultrasound data," Proc. SPIE 6510 (2007).

[17]F. Simonetti, L. Huang, N. Duric, and O. Rama, "Imaging beyond the Born approximation: An experimental investigation with an ultrasonic ring array," Phys. Rev. E **76**, 036601 (2007).

[18]Q. Zhu and B. Steinberg, "Wavefront amplitude distribution in the female breast," J. Acoust. Soc. Am. **96**, 1–9 (1994).

[19]A. Andersen and A. Kak, "Simultaneous algebraic reconstruction technique (SART)," Ultrason. Imaging **6**, 81–94 (1984).

[20]A. Andersen, "A ray tracing approach to restoration and resolution enhancement in experimental ultrasound tomography," Ultrason. Imaging **12**, 268–291 (1990).

[21]S. Li, K. Mueller, M. Jackowski, D. Dione, and L. Staib, "Fast Marching Method to correct for refraction in ultrasound computed tomography," in *IEEE International Symposium in Biomedical Imaging (ISBI)*, pp. 896–899 (2006).

[22]G. Zeng and G. Gullberg, "Unmatched projector/backprojector pairs in an iterative reconstruction algorithm," IEEE Trans. Med. Imaging **19**, 548–555 (2000).

[23]A. Andersen, "Ray linking for computed tomography by rebinning of projection data," J. Acoust. Soc. Am. **81**, 1190–1192 (1987).

[24]S. Norton, "Computing ray trajectories between two points: A solution to the ray-linking problem," J. Opt. Soc. Am. **4**, 1919–1922 (1987).

[25]F. Denis, O. Basset, and G. Gimenez, "Ultrasonic transmission tomography in refracting media: Reduction of refraction artifacts by curved-ray techniques," IEEE Trans. Med. Imaging **14**, 173–88 (1995).

[26]C. Li, N. Duric, P. Littrup, and L. Huang, "In vivo breast sound-speed imaging with ultrasound tomography," Ultrasound Med. Biol. **35**, 1615–1628 (2009).

[27]J. Tsitsiklis, "Efficient algorithms for globally optimal trajectories," IEEE Trans. Autom. Control **40**, 1528–1538 (1995).

[28]J. Sethian, *Cambridge Monographs on Applied and Computational Mathematics*, 2nd ed. (Cambridge University Press, Cambridge, U.K., 1999).

[29]A. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging* (IEEE, New York, 1988).

[30]V. Pereyra, "Ray tracing methods for inverse problems," Inverse Probl. **16**, R1–R35 (2000).

[31]E. Dijkstra, "A note on two problems in connection with graphs," Numer. Math. **1**, 269–271 (1959).

[32]J. A. Barentzen, "On the implementation of Fast Marching Methods for 3D lattices," Technical University of Denmark Technical Report No. MM-REP-2001-13, 2001.

[33]S. Li, K. Mueller, "Spline-based gradient filters for high-quality refraction computations in discrete datasets," Eurographics/IEEE VGTC Symposium on Visualization, June 2005, pp. 217–222.

[34]S. Matej and R. Lewitt, "Practical considerations for 3-D image reconstruction using spherically symmetric volume elements," IEEE Trans. Med. Imaging **15**, 68–78 (1996).

[35]G. Trahey, P. Freiburger, L. Nock, and D. Sullivan, "In vivo measurement of ultrasonic beam distortion in the breast," Ultrason. Imaging **13**, 71–90 (1991).

[36]C. Meyer, T. Chenevert, and P. Carson, "A method for reducing multipath artifacts in ultrasonic computed tomography," J. Acoust. Soc. Am. **72**, 820–823 (1982).

[37]S. Li, K. Mueller, M. Jackowski, D. Dione, and L. Staib, *MICCAI*, Springer Lecture Notes in Computer Science Vol. 5242 (Springer, New York, 2008).

[38]K. Bliznakova *et al.*, "A three-dimensional breast software phantom for mammography simulation," Phys. Med. Biol. **48**, 3699–3719 (2003).

[39]P. Taylor, R. Owens, and D. Ingram, "Simulated mammography using synthetic 3D breast digital mammography," Proceedings of the Fourth International Workshop on Digital Mammography, Nijnegen, 1998, pp. 283–290.

[40]NIH Visible Human, http://www.nlm.nih.gov/research/visible (accessed 9/24/08).

[41]B. Fornberg, *A Practical Guide to Pseudospectral Methods* (Cambridge University Press, Cambridge, 1995).

[42]K. Mueller, R. Yagel, and J. Wheller, "Anti-aliased 3D cone-beam reconstruction of low-contrast objects with algebraic methods," IEEE Trans. Med. Imaging **18**, 519–537 (1999).

[43]T. Möller, R. Machiraju, K. Mueller, and R. Yagel, "Evaluation and design of filters using a Taylor Series expansion," IEEE Trans. Vis. Comput. Graph. **3**, 184–199 (1997).

[44]V. Vinje, E. Iversen, K. Astebøl, and H. Gjøystdal, "Estimation of multivalued arrivals in 3D models using wavefront construction—part I," Geophys. Prospect. **44**, 819–842 (1996).

[45]Elements Of Breast Imaging Basics, GE Online Course (accessed 9/24/08) http://www.gehealthcare.com/usen/ultrasound/education/products/cme_breast.html.