where the first term enforces the fit of the solution to the measured fluence, the second enforces a constraint based on an a priori assumption about how a reasonable solution should behave, and λ is called the regularization parameter and governs the tradeoff between the two terms. Typically R is a norm or the norm of a derivative so as to penalize solutions that are too large or too rough. Choosing the value of λ is often critical and sensitive, and there is a vast literature on this topic (see [66], for instance, for a discussion). Regularization-type formulations can be posed in a variety of optimization frameworks and thus solved by a corresponding variety of optimization methods. Statistical modeling is important to accurately deal with the differing types of noise sources and can also provide an alternative formulation for regularization. Parametric modeling, in particular of the geometric structure of the reconstruction, can serve to reduce or eliminate under-determinedness and ill-posedness as well as to provide attractive new formulations for constraints for regularization. Our discussion must necessarily be brief and omits many contributions. We emphasize recent work and organize the presentation into three general themes: new approaches to reconstruction with linear models, new formulations of forward models for inverse solutions, and two novel approaches based on statistical regularization and parametric geometric representations, respectively.

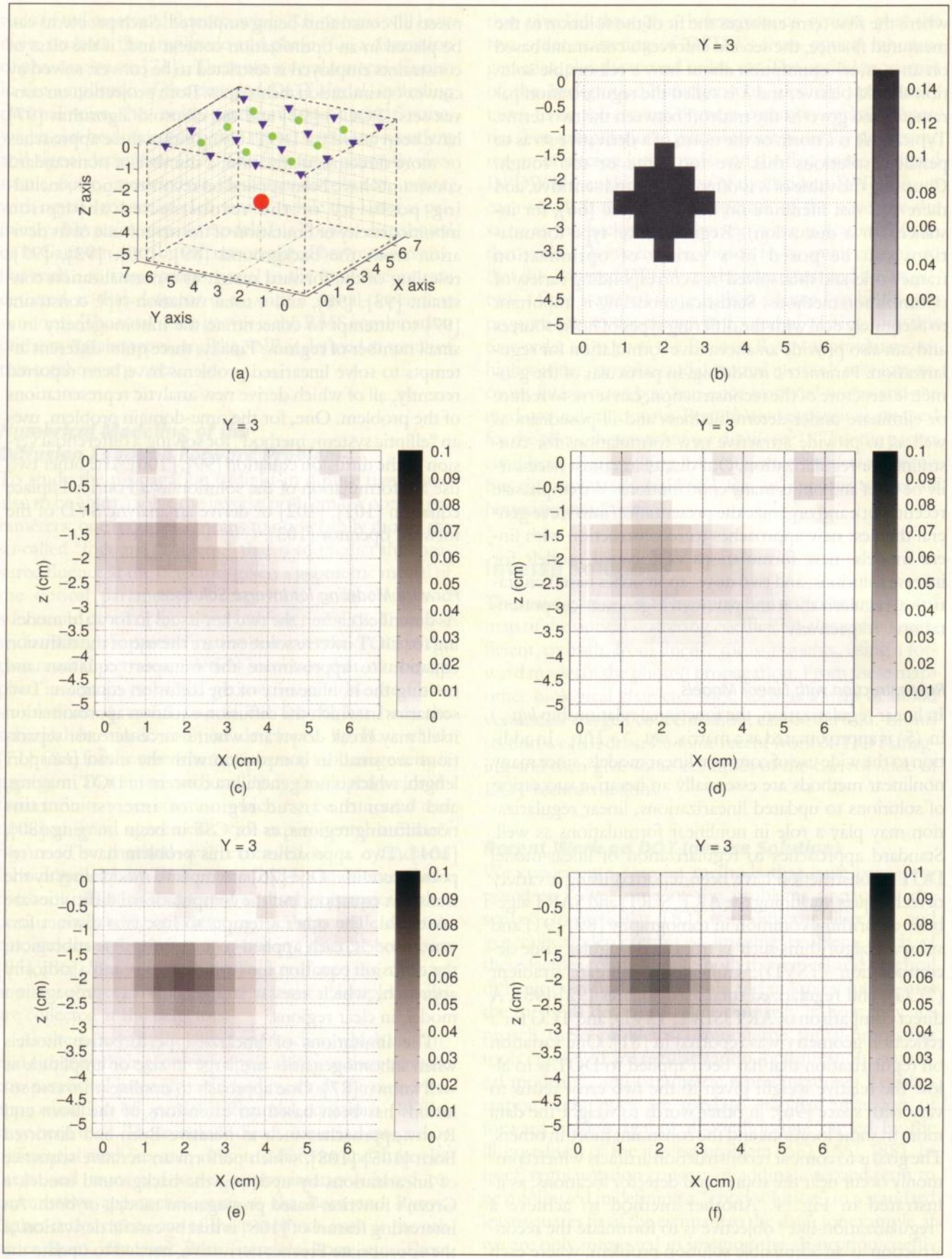
Reconstruction with Linear Models

In linear regularization, the functional relationship $h(\mu_a)$ in (5) is approximated as a matrix, $h(\mu_a) \approx H\mu_a$. In addition to the wide use of common linear models, since many nonlinear methods are essentially an iterative succession of solutions to updated linearizations, linear regularization may play a role in nonlinear formulations as well. Standard approaches to regularization of linear-model DOT reconstruction have been reported using a variety of techniques, including the ART, SIRT, and SART algebraic algorithms common in tomography [89]-[91] and subspace algorithms such as truncated singular value decomposition (TSVD), truncated conjugate gradient (TCG), and regularized total least squares [92]-[95]. A direct comparison of ART, SIRT, TSVD, and TCG for a reflection geometry was reported in [81]. One variation on regularization that has been applied to DOT is to allow the relative weight given to the two error terms to vary with space [96], in other words to weight the data more in some locations and the constraint more in others. The goal is to combat reconstruction artifacts which commonly occur near the source and detector locations, as illustrated in Fig. 4. Another method to achieve a "regularization-like" objective is to formulate the reconstruction as an admissible solution problem, where each constraint (including one on the difference between forward-projected "data estimate" and the measured fluence) is seen as describing a constraint set in the solution space. An admissible solution is then one which meets all constraints being employed. Such problems can be placed in an optimization context and, if the class of constraints employed is restricted to be convex, solved by convex optimization techniques. Both projection on convex sets (POCS) [92] and the ellipsoid algorithm [97] have been tested for DOT. Using either these approaches or more direct regularization, a number of nonstandard constraints have been applied on reconstructions, including positivity of the reconstructed absorption inhomogeneity or limitation of the magnitude of its deviation from the background [89], [90], [93], [97], rescaling of the forward matrix as a normalization constraint [93], [98], and a total variation-type constraint [97] to attempt to concentrate the inhomogeneity in a small number of regions. Finally, three quite different attempts to solve linearized problems have been reported recently, all of which derive new analytic representations of the problem. One, for the time-domain problem, uses an "elliptic systems method" for solving a differential version of the diffusion equation [99], [100]. The other two use a reformulation of the solution as a Fourier-Laplace equation [101], [102] or derive an analytic SVD of the forward operator [103].

Forward Modeling for Inverse Solutions

As described earlier, the two key issues in forward modeling for DOT inverse solutions are the use of the diffusion equation to approximate the transport equation and treating the nonlinearity of the diffusion equation. Two scenarios in which the diffusion equation approximation itself may break down are when source-detector separations are small in comparison with the mean transport length, which is not generally a concern in DOT imaging, and when the tissue region of interest contains nondiffusing regions, as for CSF in brain imaging [80], [104]. Two approaches to this problem have been reported recently. One [76] attempts to model directly the transport equation, but the computational difficulties are nontrivial. The other attempts to fuse two distinct forward models, each applied to an appropriate subregion, the diffusion equation for turbid regions and a radiosity approach, which uses a straight-line ray propagation model, in clear regions.

The limitations of linearized perturbative models when inhomogeneities are large in size or amplitude is well known [87]. One approach to nonlinear inverse solutions has been based on extensions of the Born and Rytov approaches such as iterative Born and distorted Born [105]-[108], which perform an iterative sequence of linearizations by updating the background model, a Green's function-based propagation model, or both. An interesting feature of [108] is that by careful derivation of the appropriate Frechet derivative, needed to update the Green's function, the authors found a particular weighting of the current estimate of the unknowns used to update the model. Recent work by Boas et al. [109] reports on expansion of the forward model to include as



▲ *4. Reconstruction examples for four linear reconstruction techniques at 20 dB SNR. Panel (a) shows computational volume with location of inhomogeneity, and (b) shows a vertical slice through the center of the true image. Each subsequent image shows the same vertical slice through a reconstruction using the following algorithms: (c) the ART algorithm, (d) the SIRT algorithm result, (e) the TSVD algorithm, and (f) the TCG algorithm.*

unknowns the "coupling coefficients" that model the efficiency of transfer of light from source to tissue and from tissue to detector. These coefficients are multiplicatively related to the detected fluence, so that even a linear model becomes nonlinear, but by use of the Rytov approximation (which employs a logarithmic transformation of an exponential model of the perturbation relationship) they become linear and can be incorporated within a linear reconstruction method. Another variation is to include fluence from more than one illumination wavelength (since at least two are required to determine chromophore concentrations) into a joint inverse problem and estimate the results simultaneously, taking advantage for instance of similarities in spatial structure [110].

Statistical and Parametric Inverse Models

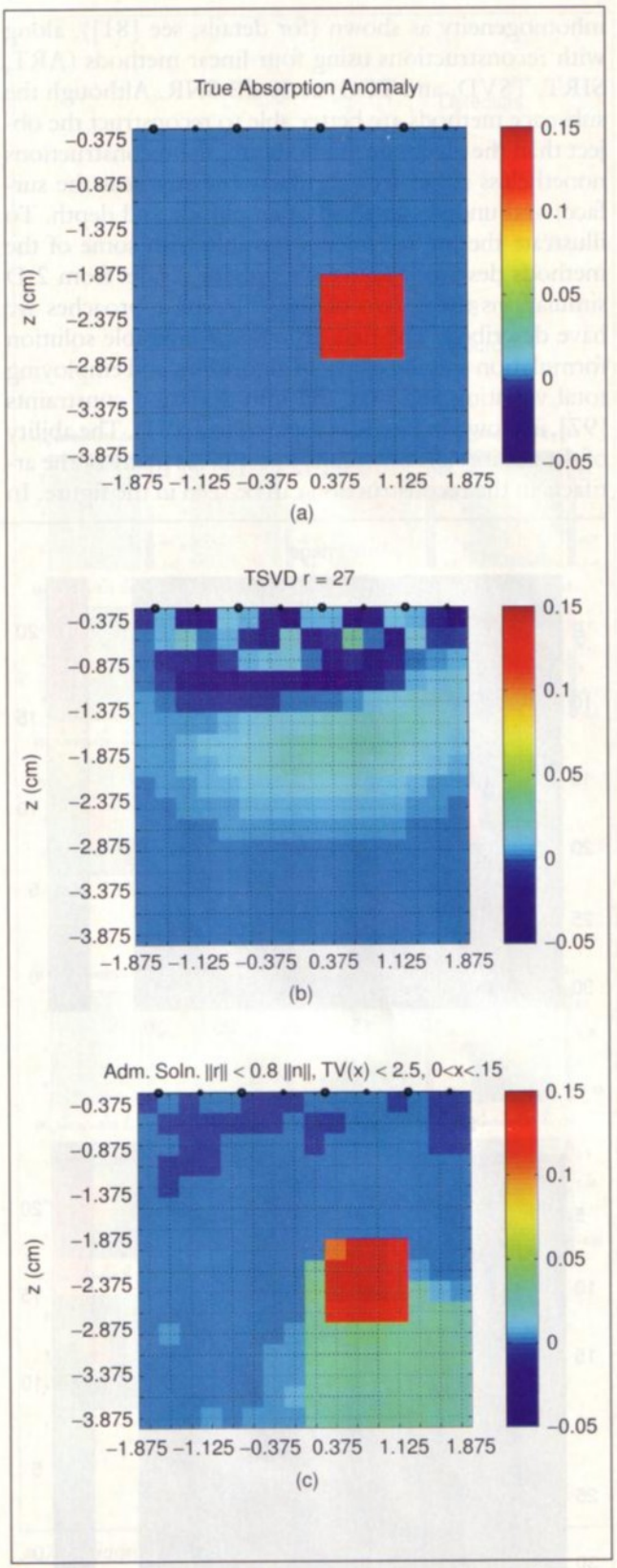
When a statistical prior model is assumed for the unknown parameters, a Bayesian maximum a posteriori (MAP) solution has formal similarity to a regularization solution, but offers extended possibilities for both algorithms and modeling. In [61], Eppstein et al. use this approach to pose the reconstruction as an extended Kalman filter, a scheme which they originally developed for geohydrology. They impose upper and lower limits on the prior probabilities via a transformation from a true beta density to a more tractable Gaussian approximation and reduce the number of unknowns at each iteration by grouping like voxels together into a large voxel with the same properties, thus ameliorating the under-determinedness.

The latter goal, reducing the number of unknowns, has also been the goal of recent work based on a sparse parameterization of the reconstruction domain. Two groups have published articles in this area, with similar but distinct approaches [111]-[115]. These methods divide the tissue region into background regions and inhomogeneities, assume that both the background and inhomogeneities can be modeled by some low-order variation (i.e., constant, linear, etc.), and assume that the boundaries of the regions are either known or can themselves be expanded as a small set of basis functions (trigonometric or B-spline polynomials). Thus the unknowns are reduced from the parameters of interest at every voxel to the coefficients of the background and inhomogeneity models and the location of the boundaries.

Some Examples of DOT Reconstructions

Below we present some reconstruction results drawn from our work, illustrating aspects of the discussion above. These are not intended to be a comprehensive presentation of the state-of-the-art, but rather to give the reader a visual impression of the abilities, limitations, and possibilities of some current approaches.

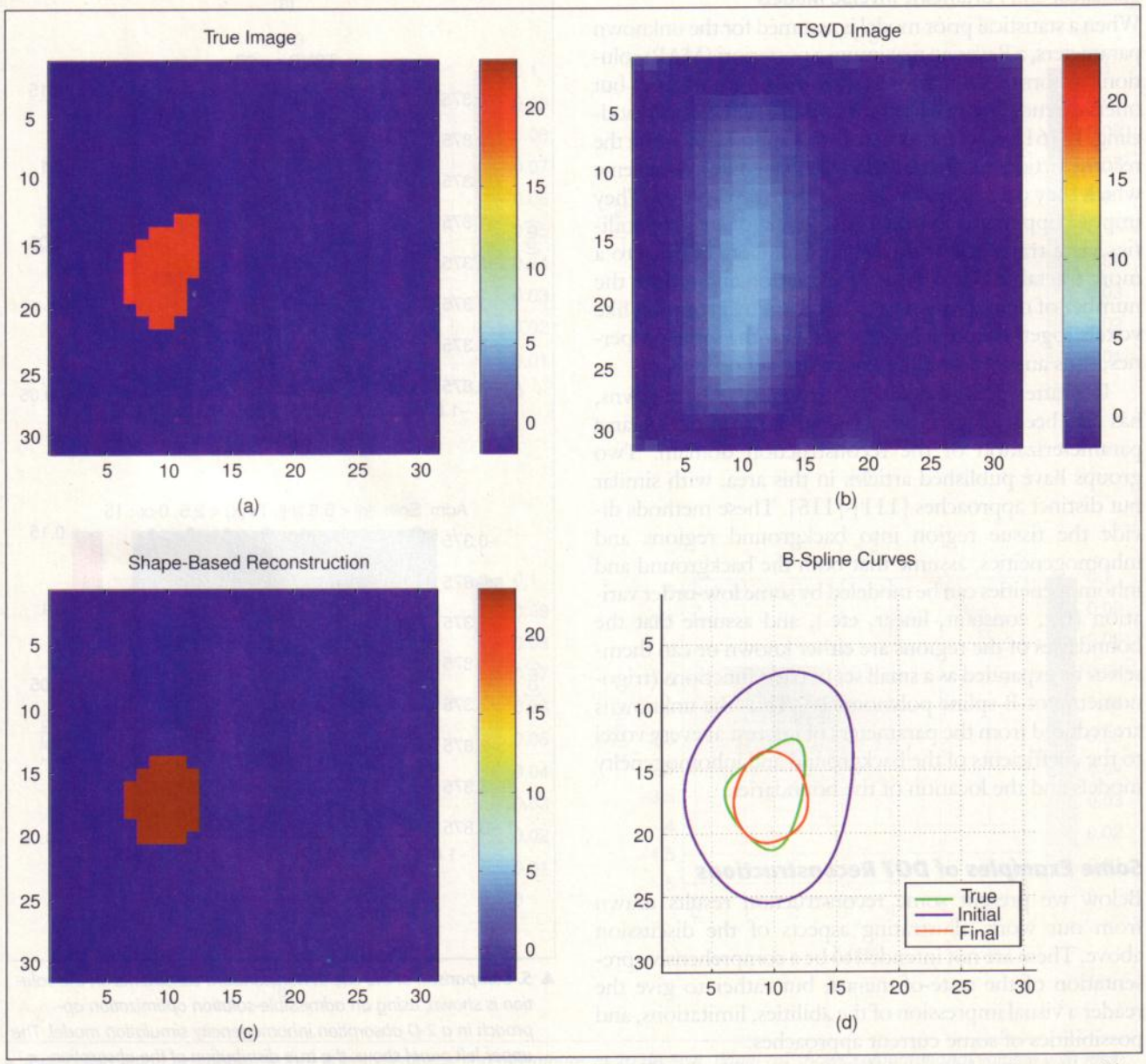
To illustrate the possibilities and problems of standard regularized DOT reconstructions, in Fig. 4 we show a particular simulated computational volume in a reflection imaging scenario with an absorption



▲ *5. Comparison of the use of nonstandard constraints on the solution is shown, using an admissible-solution optimization approach in a 2-D absorption inhomogeneity simulation model. The upper left panel shows the true distribution of the absorption coefficient, the upper right shows a TSVD reconstruction, and the lower left shows an admissible solution reconstruction using constraints on the residual of the scattered fluence and the total variation and max and min deviations of the absorption.*

inhomogeneity as shown (for details, see [81]), along with reconstructions using four linear methods (ART, SIRT, TSVD, and TCG) at 20 dB SNR. Although the subspace methods are better able to reconstruct the object than the algebraic methods are, the reconstructions nonetheless suffer from artifacts (mostly near the surface) and underestimation of amplitude and depth. To illustrate the improvements possible with some of the methods described above we present results from 2-D simulations using two of the different approaches we have described. The first, using an admissible solution formulation and the ellipsoid algorithm, and employing total variation and max and min deviation constraints [97], is shown in Fig. 5, compared to TSVD. The ability of the additional constraints to suppress many of the artifacts in the reconstruction can be seen in the figure. In Fig. 6 we present a 2-D example of a parametric basis function approach, in which the region is modeled as consisting of a constant background with constant inhomogeneity, and the inhomogeneous region is bounded by a B-spline curve. The values of the absorption in each region and the shape of the boundaries are found through iterative updates which are locally optimal. The improvement over TSVD is illustrated in the figure, and objective measures such as mean square error confirm the result. Details can be found in [115].

We also present one example of reconstruction from measured data using a phantom built at MGH and the reconstruction scheme outlined above for recovering both coupling and absorption coefficients. The phantom consisted of a highly scattering solution of Intralipid [116] mixed with India ink to produce optical properties of

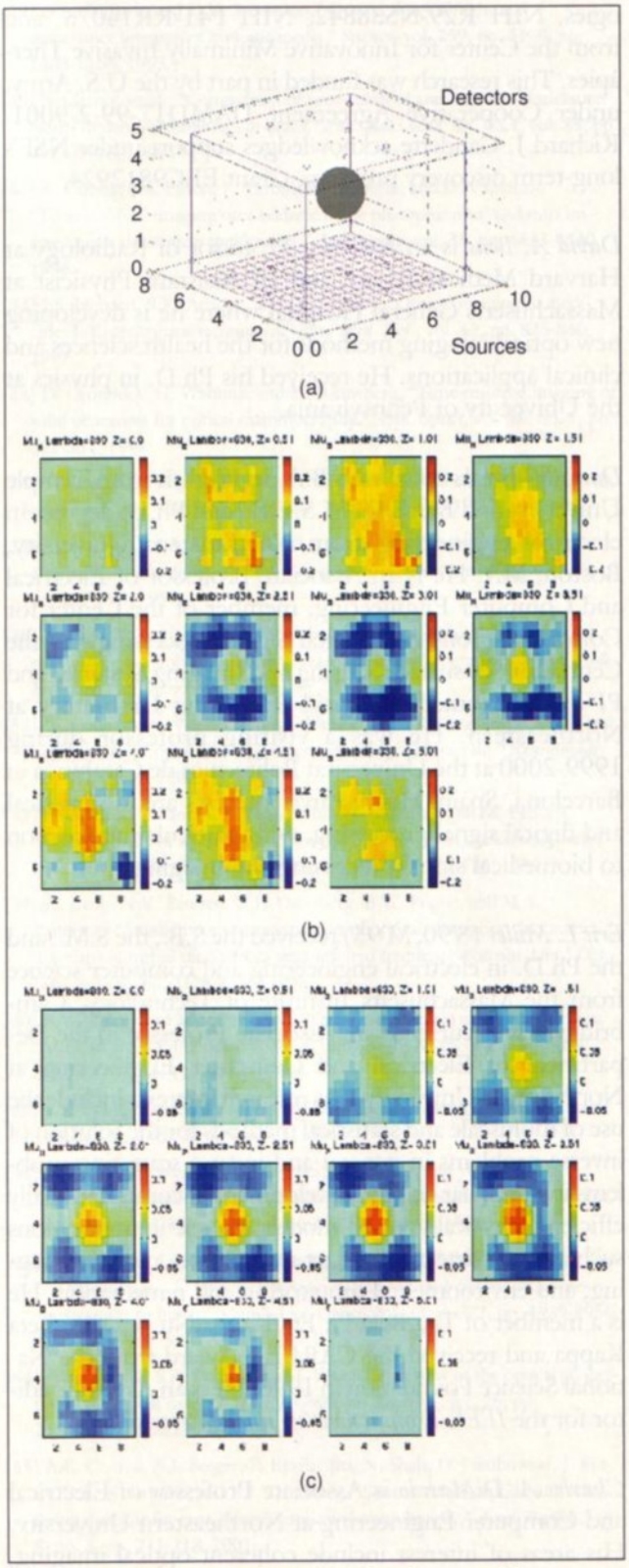


▲ *6. A 2-D example of a shape-based reconstruction. The upper left panel shows the true model of the inhomogeneity. The upper right panel shows a truncated SVD reconstruction, while the lower left panel shows a shape-based reconstruction. The lower right panel shows the outline of the true object along with the initial guess at the outline and the final outline reconstructed by the shape-based method. In all four panels the x and y axes are marked in centimeters.*

$\mu_s' = 5$ cm^{-1} and $\mu_a = 0.025$ cm^{-1}. Thirty sources were placed on one XY boundary of the phantom at $z = 0$, while nine detectors were placed on the other XY boundary at $z = 5.1$ cm. The sources and detectors spanned 8×10 cm in the XY planes. A spherical absorber ($\mu_a > 0.2$ cm^{-1}) with a diameter of 2 cm with the same scattering as the background medium was centered in the phantom. The arrangement of the source and detectors relative to the absorbing sphere are shown in Fig. 7(a). Note that this imager is different from the one pictured earlier. Measurements were made with an RF system, 830 nm illumination wavelength and 70 MHz intensity modulation, and both amplitude and phase of the detected signals were measured. Three-dimensional images were reconstructed from the 270 independent measurements. Fig. 7(b) shows the image reconstructed from the actual measured fluence with an approximate calibration of the source and detector amplitudes made prior to the phantom experiment. The absorbing object can be discerned within the center of the image, but it is overshadowed by large amplitude fluctuations near the source and detector planes. These artifacts near the surfaces result from errors in the calibration. In Fig. 7(c) we show a reconstruction from the same data after including the coupling coefficients of each source and detector in the inverse problem.

Conclusions

As both imaging technology and our understanding of the physical modeling of propagation of scattered light have developed over the last decade, DOT has become increasingly able to take advantage of sophisticated signal processing for both acquisition and reconstruction. The current challenge is twofold: in the near-term the challenge is to provide compelling evidence of its potential on clearly relevant applications such as detection of breast tumors and functional imaging of the brain. In parallel with this effort, in the longer-term the challenge is to develop better imaging devices, physical models, inverse reconstructions, and associated efficient algorithms, to extract the information which the multiply-scattered light is now known to possess. In particular, we believe there may be a role to play for many sophisticated image reconstruction and signal modeling techniques developed in the signal processing community for other purposes. Careful attention must be paid, however, to the integration of such techniques with appropriate models of light propagation to achieve useful and reliable results.



▲ *7. A diagram of the configuration of a phantom with 30 sources, nine detectors, and the absorbing sphere is shown in (a). All reconstructions in panels (b) and (c) are done using data measured by this experimental setup. The sequence in panel (b) shows an absorption image reconstructed without source/detector amplitude fitting. XY slices are shown at different z levels. Axes are labeled in centimeters. The sequence in (c) shows the same image reconstructed simultaneously with the source/detector amplitude factors.*

Acknowledgments

This work was supported in part by CenSSIS, the Center for Subsurface Sensing and Imaging Systems, under the Engineering Research Centers Program of the National Science Foundation (NSF) (Award Number EEC-9986821). David A. Boas and Quan Zhan acknowledge financial support from Advanced Research Technol-