

Image analysis methods for diffuse optical tomography

Brian W. Pogue

Scott C. Davis

Xiaomei Song

Ben A. Brooksby

Hamid Dehghani

Keith D. Paulsen

Dartmouth College

Thayer School of Engineering

Hanover, New Hampshire 03755

Abstract. Three major analytical tools in imaging science are summarized and demonstrated relative to optical imaging *in vivo*. Standard resolution testing is optimal when infinite contrast is used and hardware evaluation is the goal. However, deep tissue imaging of absorption or fluorescent contrast agents *in vivo* often presents a different problem, which requires contrast-detail analysis. This analysis shows that the minimum detectable sizes are in the range of 1/10 the outer diameter, whereas minimum detectable contrast values are in the range of 10 to 20% relative to the continuous background values. This is estimated for objects being in the center of the domain being imaged, and as the heterogeneous region becomes closer to the surface, the lower limit on size and contrast can become arbitrarily low and more dictated by hardware specifications. Finally, if human observer detection of abnormalities in the images is the goal, as is standard in most radiological practice, receiver operating characteristic (ROC) curve and location receiver operating characteristic curve (LROC) are used. Each of these three major areas of image interpretation and analysis are reviewed in the context of medical imaging as well as how they are used to quantify the performance of diffuse optical imaging of tissue. © 2006 Society of Photo-Optical Instrumentation Engineers. [DOI: 10.1117/1.2209908]

Paper 05331R received Nov. 5, 2005; revised manuscript received Jan. 6, 2006; accepted for publication Jan. 10, 2006; published online Jun. 20, 2006.

1 Introduction

Imaging with light through tissue has been a major area of research for the past two decades, and while several different generations and geometries of imaging systems exist, the tools for analyzing and comparing these systems are still rudimentary. This work provides an overview of the three major methods utilized for image analysis in the imaging science and medical physics communities, and illustrates how they can be applied to imaging through tissue with near-infrared (NIR) light. The pertinent categories lie in the areas of 1. spatial resolution, 2. contrast-detail analysis, and 3. human perception of images. This overview is organized along these three categories, and provides an introduction to the best approach in each area for diffuse optical imaging. Specific illustrative examples are taken from the medical community where relevant, and from diffuse optical imaging research.

1.1 Light Transport in Tissue

Light propagation in tissue is dominated by multiple scattering and in the near-infrared (NIR) part of the spectrum, with an especially low absorption window between the wavelengths of 610 to 940 nm. Here, the absorption is orders of magnitude lower than that observed in the visible (blue-green), ultraviolet, or infrared regions. Light transport in all wavelength regions is thought to be accurately modeled by radiation transport methods (both deterministic and stochastic models), which can predict the optical fluence patterns based

on knowledge of the microscopic interaction coefficients. These coefficients are the absorption coefficient μ_a , representing the probability of photon absorption per unit pathlength, the scattering coefficient μ_s , representing the probability per unit pathlength of a scattering event, and the microscopic phase function $P(\theta)$, which describes the probability distribution for the angular direction of outgoing scattered photons. Radiation transport models can be simplified if the absorption is low and the fluence is being detected in the far field [i.e., distance $\gg 1/(\mu_a + \mu_s)$]. In this case, the fluence can be modeled well by an assumption of isotropic scattering using a modified scattering coefficient, called the transport or reduced scattering coefficient, defined as $\mu_s = \mu_s(1-g)$, where g is the average cosine of the scattering angle:

$$g = \frac{\int_0^{2\pi} \cos(\theta)P(\theta)d\theta}{\int_0^{2\pi} P(\theta)d\theta}.$$

In this isotropic scattering regime, the field becomes diffuse, and it can be shown that the Boltzmann transport equation, used to analytically predict the irradiance, simplifies considerably to the diffusion approximation,¹⁻⁴ given by:

Address all correspondence to Brian Pogue, Thayer School of Engineering, 8000 Cummings Hall - Dartmouth College, Hanover, NH 03755. Tel: 603 646-3861; Fax: 603 646-3856; E-mail: pogue@Dartmouth.EDU

$$-\nabla \cdot D \nabla \Phi(r) + \mu_a \Phi(r) = S(r),$$

where $D=(1/3\mu_s)$ is the diffusion coefficient, $\Phi(r)$ is the optical fluence, and $S(r)$ is the isotropic source of photons. Modeling the light fluence within tissue as “diffuse” has led to great advances in the ability to reconstruct images based on diffuse projection tomography and to predict interior distributions of absorption, fluorescence, and luminescence concentrations within tissue.⁵ Several more exact radiation transport modeling studies have been reported, but application to clinical or preclinical imaging has not succeeded yet; however, as computational power grows, it is likely that this will be achieved soon.

In addition to analytic light transport modeling, stochastic methods to predict light fluence have been developed and formalized for many years,^{6–9} and have led to an improved understanding of light transport in tissue, especially over smaller distances where the light is not a true diffusive field. These forward models provide the basis for understanding the physical limitations of imaging with light through tissue,^{10,11} and are utilized in model-based imaging of tissue with light. Application of this forward modeling approach to clinical or preclinical imaging methods has not been successful yet. Thus, in this review, diffusion-model-based imaging is analyzed and characterized in terms of resolution, contrast, and detectability of objects.

1.2 Imaging with Light: Planar Techniques

The mechanisms to form functional and physiologically relevant optical images of tissue with NIR light fall largely into two broad categories, namely planar imaging and tomographic image reconstruction. In planar imaging, similar to single photon emission corrugated tomography (SPECT) nuclear medicine imaging, the remitted light from tissue is imaged with a planar device, such as a charge-coupled device (CCD) when using optical photons, which inherently captures the image pixelated by the available detector resolution. This type of imaging is predominant in clinical procedures such as endoscopy, colposcopy, and ophthalmology,^{12–14} and experimental studies in reflectance, fluorescence, and bioluminescence imaging. The image resolution and performance of these procedures is highly dependent on the geometry and the specific optical system used to acquire data. In fluorescence imaging, the issue of background suppression is most often the dominating factor affecting image quality, since rejecting the excitation light from a diffuse body is a challenging problem. In bioluminescence imaging of tissue, though there is little background signal, the intensity of the emitted light is orders of magnitude lower than that seen in intrinsic optical tomography,^{15–17} and thus low signal to noise typically dominates imaging performance, but the noise is from the readout or dark noise of the camera itself. In cases with ample signal, the shot noise will dominate, but this is often not the case in practical imaging where the light signal is low. In all of these applications, imaging performance is still further affected by the external shape of the tissue and the system’s ability to compensate for irregular boundaries, as well as the light propagation through overlying tissues. Indeed, accurate knowledge of these parameters can improve the resulting images, and throughout NIR imaging of tissue, the problems associated with knowledge of the tissue boundaries and opti-

cal properties have been a key problem. As is discussed in Sec. 2, even though the hardware may have good resolution in the absence of scattering, ultimately as the depth increases the resolution of localized emitting tissues below the surface will be governed more by the light transport than by the specifics of the hardware.

In absorption-based and fluorescence imaging, there is a unique feature of having to introduce a light field into the tissue, and then having to remove this light field from the tissue prior to detection of the true remitted signal. The most significant problem of all is specular reflectance from the surface, which can be orders of magnitude larger than signals remitted from the interior of the tissue. To reduce specular reflection, using a crossed polarizer is standard, yet this does not completely remove the reflection, and it still does not compensate for differences in remittance due to the curvature of the tissue. Interpreting all of these issues in the presence of light that has traveled an unknown indirect path between the time it went into and was remitted out of the tissue is problematic. While radiation transport models can be used, they become most challenging for prediction at complex surfaces, and the entire interpretation of background, reflections, tissue coupling, and radiation transport can significantly degrade image quality if not carefully designed. Here, the focus is on using image analysis tools to interpret image quality, which may or may not correctly account for or compensate for all of these nontrivial issues. Planar imaging in particular as compared to tomographic imaging is prone to errors due to curvature of the tissue boundary. Corrective methods have been successfully proposed, such as using a phase signal to correct for tissue curvature,¹⁸ focusing on the second derivative of the overall remission image,¹⁹ or utilizing scanner systems to outline the surface volume.²⁰ Thus the problem is not insurmountable, yet in many commercial systems this issue is largely ignored. Again, as in bioluminescence imaging, the deeper that the objects being imaged are located, the more they are affected by blurring from the light transport process, and inherently the lower the resolution will be. A more comprehensive analysis of this is provided in Secs. 2 and 3.

1.3 Imaging with Light: Tomographic Techniques

In tomographic imaging, the remitted signals are acquired, and an analytical or numerical algorithm is used to calculate images of the interaction coefficients or the distribution of tissue constituents that would give rise to the measured data. *In-vivo* absorption tomography, scatter tomography,^{21–25} and fluorescence tomography^{26–29} are now all experimentally demonstrated, and tools for quantitative analysis of these systems are required. The issues involved in interpreting tomographic images are distinctly different from planar imaging, since the assumption of a nonscattered field for resolution assessment is rarely possible. Indeed, resolution assessment itself implies that the contrast is high and scattering is insignificant, such that the dominant factor is the physical limitations of resolution. Computed tomographic imaging is almost always done in a manner that sacrifices spatial resolution for improved contrast resolution; an unavoidable tradeoff due to the need to choose a finite number of projections to keep the computational burden within what is feasible. In comparison, projection tomography without computational recovery can

often use an arbitrarily large number of projects to maximize detectability, as there is little computational burden. However, in both cases, detectability of the object is more important than resolution assessment. Thus, characterizing contrast resolution becomes more important than high resolution assessment of imaging performance. This assessment is readily achieved through contrast-detail analysis of the imaging system discussed in detail within Sec. 3.

In many cases, the recovery of images is done iteratively using a Newton method, requiring inversion of a highly ill-posed and ill-conditioned matrix. After acquisition of a set of measurements of the outgoing optical flux at the boundary Φ_m , these data are simulated in a calculation numerically or analytically with a light transport model Φ_c . This calculation is based on an initial estimate of the property distribution, and then the difference between the Φ_m and Φ_c values is minimized by iteratively updating the set of spatial absorption, scattering, or emission coefficients μ . This is done by solving a matrix equation by precalculating a linear sensitivity matrix, also known as the weight matrix or Jacobian matrix \mathbf{J} , which is used to map a change in μ to a change in Φ , providing the update equation:

$$\Delta\Phi = \mathbf{J}\Delta\mu.$$

The methods to formulate and solve this equation vary considerably between research groups, but in general, \mathbf{J} being a nonsquare and ill-conditioned matrix cannot be directly inverted, requiring the solution to be of the form:

$$\Delta\mu = [\mathbf{J}^T\mathbf{J} + \lambda\mathbf{I}]^{-1}\mathbf{J}^T\Delta\Phi,$$

for the overdetermined problem. Or for the underdetermined problem, it is:

$$\Delta\mu = \mathbf{J}^T[\mathbf{J}\mathbf{J}^T + \lambda\mathbf{I}]^{-1}\Delta\Phi,$$

where λ is a regularization parameter and \mathbf{I} is the identity matrix. The images are formed by starting with an initial homogeneous or heterogeneous guess of μ and iteratively updating this estimate with the matrix equation solution by calculating $\mu_{i+1} = \mu_i + \Delta\mu_i$. Detailed discussions of these methods for diffuse tomography have been published previously.^{4,30-33} The particular reconstruction algorithm used in these studies has been tailored for the circular tomography geometry and optimized for regularization and inclusion of *a priori* information.³⁴⁻³⁷

2 Spatial Resolution

When image performance is analyzed in the classical ray optics sense, the contrast between effectively black and white regions is used to assess the minimum detectable size or separation between any two points. This approach is routinely used to test the resolution of telescopic imaging systems or microscopic imaging system hardware performance.^{38,39} It should be noted early on in this discussion that this approach is rarely useful for analyzing routine medical imaging of soft tissues, especially *in-vivo* optical imaging, as it assumes that infinite contrast is available. Thus, the tools discussed in Sec. 3 are generally more appropriate for most analyses of optical tissue imaging systems; however, this section on resolution

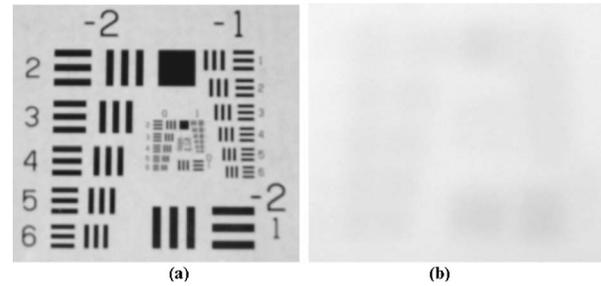


Fig. 1 Images of the USAF test field, imaged (a) through air and (b) through 2 mm of 1% intralipid solution. The ability to use this type of resolution tool in a diffusing medium is very limited.

assessment is included for completeness. Resolution analysis is always useful for characterizing the limiting hardware and software performance of an imaging system, and may be useful for imaging thin tissues or tissues with minimal scattering, where extremely high contrast is possible.

2.1 Line Pair Imaging Analysis

Line pair analysis is the standard in most imaging systems, where the assumption of extremely high contrast makes sense, and where the ultimate spatial resolution of the system needs to be tested. Phantoms or test fields that do not cause significant background scatter, and have effectively infinite attenuation contrast, are standard tools to assess the limiting spatial resolution. The United States Air Force (USAF) resolution test chart, shown in Fig. 1(a), is classically used as a universal test field containing a range of line pairs per millimeter. It is the ideal test field for mesoscopic or telescopic imaging fields, where the field size ranges from the order of millimeters to near $5\ \mu\text{m}$. Below the micron range though, this test field does not provide sufficient resolution for useful imaging analysis, and precalibrated microscopic line pair test fields can be purchased. Resolution limits for a given imaging system are measured by visually assessing the number of line pairs per millimeter that are discernable in an image. More quantitative measures can also be applied, such as determining the line pairs per millimeter grouping that shows a discernable decrease between the two dark lines. In the limit of high resolution coherent light, this could be determined with the Rayleigh criterion, where the peak of one object appears to overlap the first Airy disk trough of the other.³⁸

This method of line pair imaging analysis has been used in several biomedical papers mainly focused on thin tissue imaging. Examples include investigations into the use of polarization filtering,⁴⁰ ultra-fast time-resolved detection,⁴¹ or coherent gating methods,^{42,43} all of which focus on rejecting multiple scattered light from transmittance or reflectance signals.^{44,45} The use of line pair resolution testing at depths beyond the transport scattering length ($\ell' = 1/\mu'_s$) is not very informative, since the anticipated resolution is larger than the maximum width between line pairs in the USAF test field. Line pair analysis is usually restricted to imaging in high resolution situations, since imaging multiple lines requires a large field of view that has high resolution and can thus image many lines within a field. For imaging systems with lower to moderate spatial resolution, or systems that image in a scatter dominated field, it is more useful to discuss the resolution as

assessed by the point spread, line spread, or edge spread function, or as interpreted by the Fourier transform of these, the modulation transfer function. This is discussed in the next section.

2.2 Point/Line/Edge Spread Functions

Assessment of the point spread function (PSF), line spread function (LSF), or edge spread function (ESF) can provide nearly equivalent information in imaging systems that have a linear and spatially homogeneous response. Unfortunately, this latter criteria is rarely true in most useful imaging systems, and so a comprehensive assessment of these systems would involve assessing the PSF at multiple locations in the imaging field⁴⁶ to assess the spatial response. For example, chromatic aberration effects lead to imaging field responses that vary in space, and so assessment of the PSF at different areas in the field is often required to characterize a system. Similarly, in diffuse imaging, the point spread function can easily be distorted by the boundaries of tissue, as the diffuse response function is altered by orders of magnitude when near a boundary.

Point spread function imaging in tomography has been a standard practice to assess imaging system and algorithm quality.⁴⁷ However, many studies have blurred the line between point imaging and circular region imaging when defining spatial resolution. When infinitely high contrast is used, the size of the object being imaged should not affect the size of the resulting image, but rather as the object size decreases, the resulting image response decreases in magnitude. PSF determination can be achieved using black spherical inclusions in the medium to simulate the small infinite absorbers present. These types of studies are tedious, but when done exhaustively can provide fundamentally important information about the imaging response and the field.^{10,48,49}

Measurements of the line spread function have been completed in many early studies of diffuse tomography, and illustrate the classic banana-shaped sensitivity function observed between the source and detector. As a line is translated across the path between source and detector and oriented perpendicular to the path of the light travel, the change in intensity obtained in this profile study is exactly proportional to the intensity of the photon path, or the sensitivity function of each source-detector pair.^{10,50-55} These basic measurements can be obtained and compared to computational predictions to illustrate the accuracy of the forward model in predicting the Jacobian matrix.⁵⁴

One useful empirical observation applicable to imaging through diffusely transmitting slabs is that the lowest spatial resolution is most likely limited by the broadest point of the photon path, which has been estimated at 20% of the slab thickness.⁵⁶ While this estimation is empirical and only true in diffuse regimes, it provides a “rule of thumb” to estimate the resolution at the center of a slab for a point source and point detector. There are analytic predictions of the photon paths for both reflectance and transmittance imaging.^{54,57} These predicted paths are effectively weights representing the ensemble of statistical paths taken by photons, and are analogous mathematically to the adjoint Jacobian matrix, as developed for image reconstruction.^{4,33} These mathematical expressions and algorithms provide a useful way to estimate the lower resolu-

tion limit for given source-detector and boundary geometries.

The use of an edge spread function provides a strong contrast that is readily imaged, and analytic derivations of the transmission past an opaque edge in an otherwise scattering medium have been shown⁵⁸⁻⁶⁰ for infinite medium geometries. Measurements in bounded tomography regions show considerably less predictable response,⁶¹ but still provide strong insight into the nonlinear response across the imaging field. The disadvantage of this approach is that most diffuse reconstruction programs are inherently based on the concept of the Newton method, where the background medium is assumed to be homogeneous and the reconstruction algorithm function is to recover the embedded regions within this. This scenario is not true when a large and distributed field such as the edge of a large absorber is present, and so the response of most diffuse image reconstruction algorithms to a large flat object can be substantially different than the response to smaller round-shaped objects. Therefore, results of an ESF assessment in diffuse imaging should probably be interpreted alongside similar measures of the PSF or LSF.⁶¹

2.3 Modulation Transfer Function

In most medical imaging settings, the data of PSF/LSF/ESF are interpreted in the frequency domain by Fourier transforming the dataset to provide a modulation transfer function (MTF). While the information content is the same, representing the spatial frequencies provides a direct linear method to spatially filter or modify the response function. In many systems, the standard approach has been to use a measurement of the LSF and transform this to the MTF.⁶² However, in certain systems, generating a line that is thin enough to test the system satisfactorily may be problematic due to constraints on the setup, and so it is often simpler to generate a sharp edge for measurement of the ESF. The MTF for each of the PSF/LSF/ESF curves in Fig. 2 are shown in Fig. 3. It can be seen that similar information is provided in both the spatial and frequency domains for this simple case. In this case, the response at the center of the phantom is seen to have lower spatial frequencies in content, corresponding to wider values of the PSF/LSF/ESF functions. This difference is understood to be caused by the diffusive path between source and detector, which is widest when most distant from a source or detector. Narrowing the source detector distance is the only physical way to decrease PSF/LSF/ESF values, and points nearest either source or detector will always have the lowest values. Decreasing scatter or increasing absorption will reduce the PSF/LSF/ESF values and increase the spatial frequency bandwidth as well, throughout the entire imaging field.

2.4 Application of Resolution Testing in the Field of Diffuse Tomography

In diffuse light imaging, it has long been recognized that light follows a statistical path in which the predominant path between source and detector is a line surrounded by a banana-shaped distribution.⁶³ This spreading of the photon paths is induced by the inherent multiple scattering present, and decreases in width in a medium with lower scattering or increased absorption. The effect of increased absorption is somewhat counterintuitive, but generally leads to a loss of

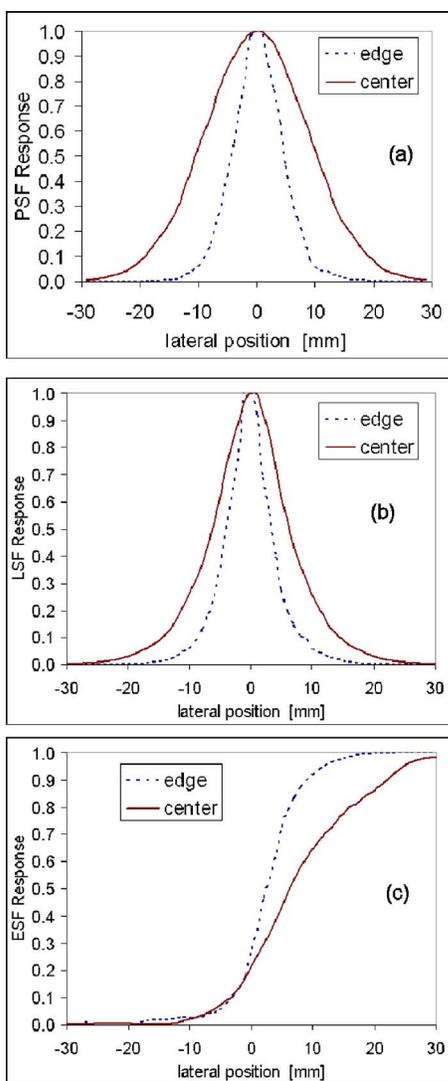


Fig. 2 Graphs of the (a) point spread function (PSF), (b) line spread function (LSF), and (c) edge spread function (ESF) for “pencil-beam” transmission through a 60-mm slab, having diffuse interaction coefficients of $\mu_a=0.01 \text{ mm}^{-1}$ $\mu'_s=1.0 \text{ mm}^{-1}$. In all three cases, two locations were analyzed using a target near the edge (2 mm inside the surface) and then at the center of the slab.

photons that have traveled farther in tissue that subsequently narrows the average path of travel. These distributions have been studied by many investigators, and specifically quantified by papers in the early 1990's.^{30,51,52,56,64,65} Imaging of edges has not proven all that useful, as the wide spread of photons really limits the ability to visualize the edge of objects clearly, and the spatial variation in the resolution ultimately complicates the analysis.

In diffuse tomography imaging, it is easier to resolve a smooth circular heterogeneity embedded in a field than step changes.⁶⁶ This is because of the fact that heterogeneities appear as symmetrically Gaussian filtered objects in the image. Almost all papers in the field of diffuse optical imaging have focused on assessing resolution by placing point objects or line objects in the field to assess spatial resolution.^{4,30,64,67-73} This focus has emerged from a fundamental limitation in the

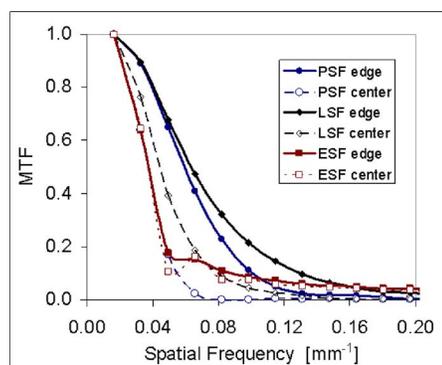


Fig. 3 MTF profiles of the graphs shown in Fig. 2, illustrating the information contained as a function of spatial frequency, with LSF having highest resolution, PSF having next highest, and ESF having the lowest. Resolution is always considerably worse in the interior of the diffusing medium than at the edge near a source or detector.

field of diffuse tomography stemming from the fact that all currently used reconstruction algorithms are derived at some level from perturbation theory. The Born, Rytov, and Newton methods for minimizing an objective function are all based on perturbing an initial field to find the solution. This approach is inherently optimized for imaging point objects, and the ill-posed nature of the problem, combined with significant regularization, leads to a solution that is significantly smoother than the original test field.

Once a system or algorithm is established in its ability to recover point objects, extension to multiple objects has been a common theme; however, this step is both important and problematic. The most significant problem is the nonlinear response of the measured field to multiple or extended inhomogeneities, requiring an infinite number of heterogeneity configurations to fully analyze system performance.

Perhaps the only reasonable approach to characterizing the imaging field response to multiple heterogeneities is to simulate the expected distributions of values possible *in vivo* and use this as the limited calibration of the system and corresponding algorithm. Even with these measures taken, it is critical to evaluate these distributions with the full range of object sizes and contrasts expected *in vivo*, as is discussed in Sec. 3 on contrast-detail analysis.

2.5 Analysis of Luminescence and Fluorescence Diffuse Imaging Resolution

Imaging of the minimum spatial resolution is only reasonable when an effectively infinite contrast is expected. Fluorescence protein imaging or bioluminescence imaging are two of the few situations in optical imaging *in vivo* where it may be reasonable to expect nearly infinite contrast, when the background emission issues might be neglected or corrected.⁷⁴⁻⁷⁷ When cells are specifically transfected or modified to express an optical signal, such as a specific organ or a tumor, the background emission in the neighboring organs should be effectively zero. In green or red fluorescent protein (GFP or RFP) imaging, the background and leakage of excitation light through the filters does provide the most significant background signal; however, this can be significantly reduced when wavelength-dependent fitting or wavelength-based

background subtraction is used. In bioluminescence, little real background is present in most cases, and background is often simply the dark noise in the camera or light leakage into the enclosure from the room. Thus, the spatial resolution of bioluminescence or fluorescence protein imaging *in vivo* can be assessed by point spread function or line spread function imaging, yet little study of this has been reported. One comprehensive paper on this issue by Troy et al.⁷⁷ showed effective point spread functions as measured in phantoms and *in vivo*, using small numbers of cells to assess the minimum detectable number of photons and cells. This analysis illustrated that bioluminescence is a more sensitive imaging technique in the remission geometry, by a considerable margin, due to the decrease of fluorescent protein imaging sensitivity caused by background autofluorescence. However, recent reports of fluorescence imaging in the transmission geometry will likely be more sensitive. In most applications of fluorescence or bioluminescence, the actual resolution was not the most important parameter in distinguishing the two systems, but actually the sensitivity. Resolution of bioluminescence and fluorescence appeared to be similar, because the photon spread within a spectral window was effectively equivalent. This study and other similar studies focus on minimum contrast or signal detectable, because the issue of resolution is not governed by system constraints, but rather by the physical constraints of the light transport in tissue. Resolution limits in this regime have less to do with system design than with the depth of the object to be resolved in the tissue. While the resolution of objects at the surface of a tissue can, in principle, be as high as the diffraction limit of light (i.e., near 250 nm) given sufficient contrast, the presence of tissue motion and the quality of the imaging system typically contribute to the real imaging resolution being lowered to typically near 1 to 2 μm when imaging at the surface of the tissue. Again, resolution clearly degrades by orders of magnitude in just a few millimeters of depth into the tissue, due to the overwhelming presence of scattering.

3 Contrast-Detail Analysis

Resolution of an imaging system is a common term that is often used inappropriately in medical imaging. While resolution refers to the lowest resolvable size in the field of view, it does not have any appreciation of the contrast of that object being imaged. In most cases where there is a finite contrast between the region to be detected and the background, the pertinent measure is whether the object is “detectable” with a given size at a given contrast. This measure is more subjective than resolution, as it implies some human imposed decision of what is “detectable.” The process of contrast-detail analysis was developed precisely for this purpose.

3.1 Contrast-Detail Curves

Contrast-detail (C-D) analysis is commonly used to determine the performance of medical imaging systems and is an effective method for assessing the imaging capabilities of prototype systems. This technique was introduced to medical imaging in the 1970’s and has since been used extensively for CT^{78–81} magnetic resonance imaging,⁸² ultrasound,⁸³ mammography,^{84,85} fluoroscopy,⁸⁶ whole body x-ray systems,⁸⁷ as well as imaging displays.^{88,89} This technique is

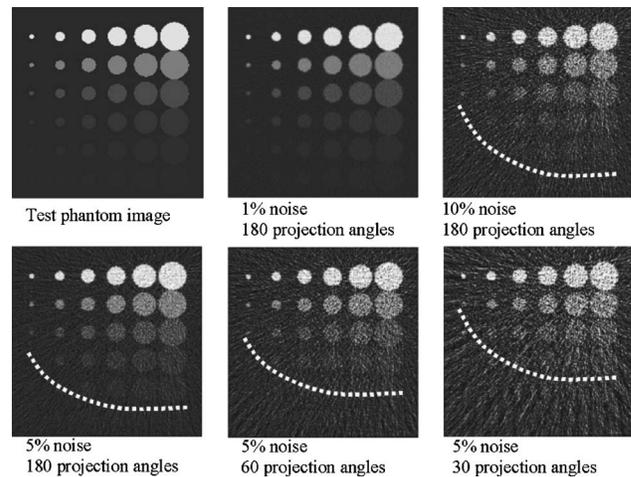


Fig. 4 An example of contrast-detail analysis is shown for a simulated computed tomography test phantom having six different sized objects at six different contrast values from the background. This type of phantom is used routinely to test imaging system performance, by imaging and determining the minimum contrast detectable for each object size. As seen in (b) and (c), the reconstructed images show a degradation of the image quality due to noise and sampling error, and in (c) the line of minimum detection is shown. In (d), (e), and (f), a 5% noise level was used, and the number of sampling (or projection) angles was systematically decreased to show the degradation of the image and the resulting increase in the position of the contrast-detail curve.

used to quantify the combined performance of the imaging system and the image reader in detecting objects representing a clinically relevant range of sizes and contrasts within a domain, focusing on assessing the lower limits of each possible range.^{90,91} A C-D graph of minimum detectable contrast level for all sizes of objects provides limiting data on two major regimes of system operation, namely 1. the spatial resolution for high contrast objects (high contrast, small object size), and 2. the lower level of contrast detectable for larger-sized objects.⁸¹

Typical contrast-detail study test fields contain a series of objects representing a range of contrasts and diameters often in a regularly spaced pattern. The phantom can have either discretely sized objects or continuously varying sizes, and each object size is repeated with varying contrasts relative to the background. A theoretical representation of a C-D phantom is shown in Fig. 4(a), with object size increasing from left to right and contrast increasing vertically. The phantom is imaged tomographically by simulating x-ray computed tomography transmission data with the addition of varying amounts of Gaussian distributed noise in each detector. The noise and geometry of the system contribute to the resulting image quality, as represented by the recovered images in Figs. 4(b) through 4(f).

In assessing each image, the lower limit of detection for a given size is determined for each column, so the minimum “detectable” contrast level is determined for each size. This is repeated for all sizes, and a curve (as shown in the images of Fig. 4) is displayed representing the minimum detectable range of contrasts for all sizes. In this analysis, “contrast” represents the real object contrast prior to image reconstruction, relevant to the imaging modality, rather than the contrast

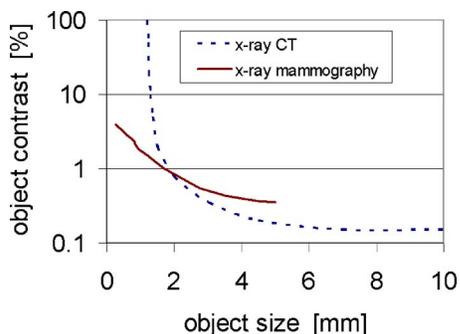


Fig. 5 Contrast-detail curves showing the values for a mammography system and for a CT system, illustrating the inherent strengths and weaknesses of each system. The mammography system is strongest for imaging smaller objects, below 2 mm diam, as lower contrasts can be detected. In comparison, CT cannot image objects smaller than 1 mm, but can image subtle changes in contrast better when the objects are larger than 2 mm.

within the generated image. In Fig. 4, it can be seen that an increase in noise results in an increased contrast required to detect an object of a given size. Furthermore, a decreased number of projection angle samples results in a higher C-D curve, and therefore lower *contrast resolution*. In tomography systems, the ability to resolve subtle contrast changes is typically most important, and a lower C-D threshold curve indicates the better system contrast resolution and therefore a higher level of imaging performance.⁹⁰

3.2 Contrast-Detail Analysis in Medical Imaging

Detection thresholds usually are determined by human readers,⁹² but automated assays of the image can also be based on contrast-to-noise ratio (CNR) and have been used in prototype systems.^{93–96} In the clinical setting, human readers are the gold standard through which comparisons are made, since all radiological images are read by radiologists for diagnosis. There has been significant research to develop automated algorithms that systematically “detect” objects by acting as ideal observers or appropriately mimicking human observers.⁹²

Contrast-detail phantoms are commercially available for all conventional clinical imaging systems, and the American College of Radiology has established guidelines for phantoms, which should be used for calibration and validation of specific imaging modalities. Common applications of these studies include periodically assuring clinical system image quality, optimizing developing technologies, and comparing intersystem performance. Figure 5 is an example of contrast-detail curves illustrating the typical performance of a mammography system⁸⁵ and an x-ray CT system.^{97–99} Objects with sizes and contrasts in the range above and to the right of the curves are considered detectable, while those below and to the left are too small or have too little contrast to be detected in the image. Therefore, imaging systems with curves closer to the x and y axes indicate a lower contrast detection threshold for all objects and are considered to have better imaging performance for the phantom being imaged.

As a performance measurement tool for standard radiography and mammography, contrast-detail analysis is used for scheduled quality assurance, optimizing system settings,^{98,99}

assessing and comparing digital and film-screen detection methods,^{87,100–104} and determining minimum requirements for image viewing and digital file storage protocols.^{88,105–107}

Contrast-detail analysis is an efficient means to complete these studies, since a given C-D plot often requires only a single image and minimal time commitments by professional readers. Some C-D studies in radiography seek to optimize the tradeoff between image contrast and x-ray dose to the patient by extending the analysis to include total dose levels. This is particularly important for computed tomography systems, where patient dose is a major concern. A series of studies in the late 1970's and early 1980's by Cohen et al. were the first to directly apply contrast-detail analysis to the assessment of CT scanners.^{78–80,90,108} Faulkner et al. published a fairly comprehensive set of contrast-detail results to explore the effect on CT imaging performance using different filtering techniques, reconstruction algorithms, and CT scanners.⁸¹ Contrast-detail analysis has also been used in assessing ultrasound,^{83,109–112} fluoroscopy,^{86,113–115} and magnetic resonance (MR) systems.^{82,116} The key in this process is to utilize a contrast-detail phantom, which is representative of the size and contrast scale of the tissues that the system will be used to image routinely. The shape of the C-D curve then allows comparison between systems with the same settings and analysis of which applications would be suitable for the system.

3.3 Contrast-Detail Applications in Planar Versus Tomographic Imaging

Contrast-detail studies for planar imaging systems with relatively large and flat response fields can be completed using a phantom containing the full set of object contrasts and diameters. This allows the full C-D analysis to be performed with a single, or limited number of, image(s). Applying contrast detail to NIR tomography presents a more difficult problem, since the presence of multiple optical contrast objects in the test field has a substantial effect on image quality¹¹⁷ due to the inherent soft field nature of the problem. However, this has been addressed by using a series of images, each with a single object within the test field.⁹⁴ The object size and contrast are varied between images, and detection thresholds are extracted and compiled to produce the C-D curve. Additionally, simulated studies are easily completed to quantify expected contrast-detail system performance for best case and a variety of other simulated conditions.^{95,96}

Further imaging performance assessments for emerging technologies should be reported following current medical imaging protocols. Contrast-detail analysis can provide a reasonably comprehensive method for quantitatively and systematically assessing system capabilities. Though other techniques, such as receiver operating characteristic (ROC) analysis, may be more rigorous, these tests are not always practical due to professional reader time constraints.^{91,109,112,118} The relative ease of producing contrast-detail plots makes this type of analysis attractive as a preliminary or primary method for ranking and optimizing medical imaging systems.

The key factor in the observation of higher spatial resolution in planar imaging is that tomographic image reconstruction spatial resolution is limited by the number of projects that can be obtained, due to the image reconstruction problem be-

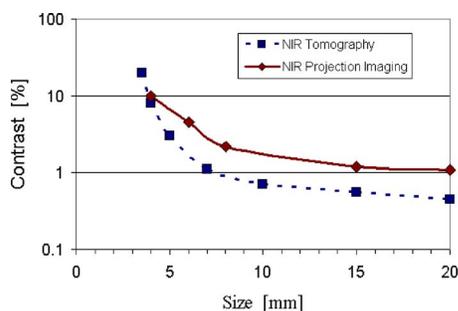


Fig. 6 Contrast-detail curves are shown for diffuse optical imaging simulations, comparing transmission imaging with projection data through a 6-cm-diam slab, to NIR tomography of an 8-cm cylinder. These geometries are chosen to mimic that of breast imaging, and the test object is placed at the center of the imaging field to simulate the most difficult lesion to detect. The minimum detectable contrast for each size object is shown as a square, illustrating that for reconstruction tomography, the minimum detectable contrast is lower than that for projection imaging. Objects smaller than 4 mm are not easily simulated in these calculations, but are presumably more readily detected with projection imaging than with tomographic imaging, as is observed in Fig. 5.

coming overwhelming computationally at higher numbers of projections. This computational limit on the image reconstruction often limits the resolution that can be achieved; however, in projection imaging there is typically no major computation required, and so larger numbers of projections can be taken, and hence the resulting resolution can be maximized for objects near the surface of the tissue. This is certainly true in x-ray imaging throughout tissue volumes, but the scattering process in diffuse tomography complicates this a little further. It is generally believed that diffuse image reconstruction for interior objects (distant from any surfaces) could be recovered with higher resolution in tomography, than with projection imaging.

3.4 Applications in Endogenous Versus Exogenous Contrast Imaging

With the recent advent of absorption, scattering, and fluorescence tomography methods,^{26,27,119,120} and potentially bioluminescence tomography,^{121,122} it is important to fully understand the capabilities and limitations of each approach for characterizing tissues, in terms of the minimum detectable contrast, or cell number *in vivo*. As stated before, it should be anticipated that in most cases tomography is not the best way to improve spatial resolution, but rather is the optimal way to recover low contrast information (i.e., optimal contrast resolution) when a limited number of projects are able to be obtained due to computational limits. An illustration of this is shown in Fig. 6, where the contrast-detail curve of absorption-based imaging is compared to tomographic imaging. Recent breakthroughs in enzymatically activatable fluorophores have created a significant interest in the possibilities for molecular imaging.²⁶ Systematic characterization of fluorescence tomography is important for the implementation of this modality to imaging tumors, where specific markers of cellular or vascular expression may be localized.

The contrast-detail response of fluorescence tomography is expected to be similar to that of near-infrared absorption to-

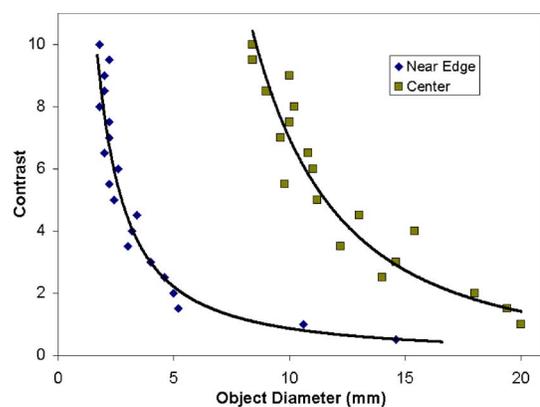


Fig. 7 Contrast-detail analysis of diffuse imaging demonstrating the difference in sensitivity for objects located at the center and near the edge of the imaging field. These simulations are based on fluorescence tomographic imaging, and are similar quantitatively to absorption-based imaging. The simulation field size was 86 mm diam, using scattering and absorption parameters typical of soft tissue imaging in the near-infrared ($\mu'_s=1.0 \text{ mm}^{-1}$, $\mu_a=0.01 \text{ mm}^{-1}$). The points present the contrast at which a minimum contrast-to-noise ratio of 3 is recovered in the resulting tomographic images.

mography, as the path of light propagation is similarly diffuse in nature. Computational studies of this type have recently been completed⁹⁶ and the results, shown in Fig. 7, illustrate a best-case C-D curve for imaging through 86 mm of tissue.

Future studies with experimental systems can demonstrate how the experimental apparatus and extension to real tissue affects the contrast-detail performance reported in Fig. 7. Ultimately, this will help determine the potential role of the system in a research or clinical setting.

Though most current studies in fluorescence optical tomography focus on small animal studies, the high tissue absorption and small size of these test subjects results in a less diffuse light field, which departs from the diffusion approximation of photon propagation and thus can complicate the image reconstruction process. This problem has been well studied, yet no clear solution exists other than attempting to model the light propagation with radiation transport theory or Monte Carlo models.^{10,123} An emerging approach is to incorporate high absorption coefficients into a diffusion-based model. This technique has been reported with visible wavelengths and has been shown to work considerably well in small animal imaging.¹²⁴ Further work in this area is ongoing, and a clearer analysis of contrast-detail characterization of these systems would be a significant benefit.

3.5 Contrast-Detail Application in Assessment of Hybrid Imaging Systems

In recent years, there has been a significant interest in developing near-infrared imaging systems that are coupled to standard clinical systems, or systems that can contribute structural prior information to the image reconstruction process. This has been shown with MRI^{21,125–128} and ultrasound.^{129–131} The hypothesis driving the development of these hybrid techniques is that the accuracy of the image or the image property values will be somehow improved. While this is implied, it has not clearly been proven, and must be demonstrated for

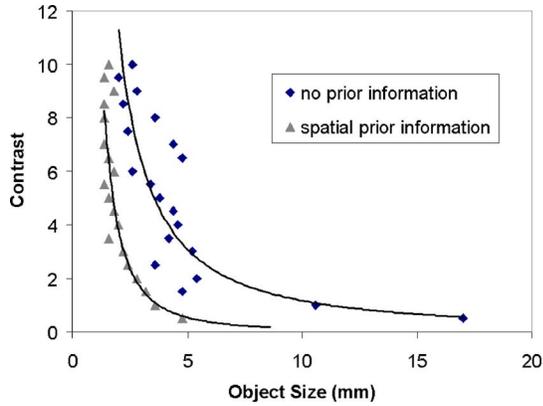


Fig. 8 Contrast-detail analysis of diffuse imaging with and without a *priori* information about the fat and fibroglandular layer of the breast, estimating the detectable levels of contrast for given object sizes. With the inclusion of a *priori* fat/glandular tissue layers, the minimum contrast required for each size is decreased to achieve the same CNR value.

each system. Indeed, there is thought to be a delicate balance between forcing the solution to converge to a solution imposed by the *a priori* constraints, versus allowing the constraints to minimally coerce the solution toward the most accurate image. In recent studies, Brooksby et al.^{21,125–128} have demonstrated *in-vivo* imaging with MRI-coupled NIR tomography, and used the MRI information to segment the adipose from the glandular tissue, providing input spatial information about the two tissue regions. This approach provides important *a priori* information, which then improves the image reconstruction. Analysis of this improvement can be accomplished through contrast-detail analysis, thereby quantifying the improvement in contrast that can be expected for each size object, given a decision criterion. This is demonstrated in Fig. 8 with contrast-detail curves for images generated with and without structural *a priori* information. These curves were determined using CNR=3.0 as the threshold for assessing object detection; however, this study can also be completed using human observers or a computer-generated “ideal observer.”

The shift of the curve to the left in Fig. 8 with the inclusion of *a priori* information from MRI is a quantitative indication that the CNR of the imaging system is superior to the imaging capability without *a priori* information. The experimental extension of this work is ongoing, and use of contrast-detail analysis in this application will demonstrate improved contrast resolution with increasing levels of spatial constraints that can be implemented. As hybrid imaging systems become more established, as are PET/CT systems, this type of analysis will become increasingly important to quantitatively evaluate the system capabilities.

4 Analysis of Human Observer Interpretation of Images

4.1 Sensitivity, Specificity, and Receiver Operating Characteristic Analysis

The receiver operating characteristic (ROC) methodology has been widely used to address the clinical efficacy of medical

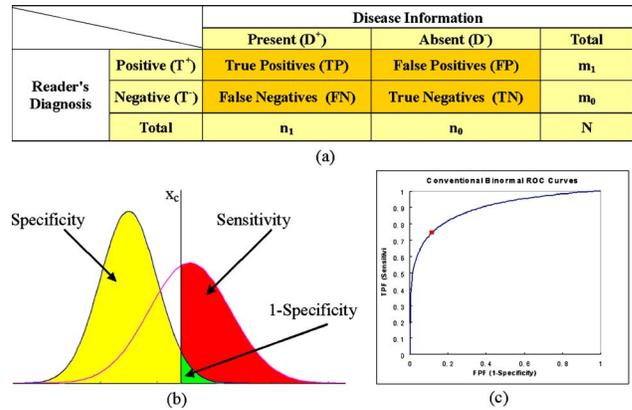


Fig. 9 Statistical analysis of data given two distributions. (a) The 2×2 diagnosis table. The observer gives m_1 positive and m_0 negative diagnosis to N patient images of which n_1 are diseased and n_0 are nondiseased. (b) Example of probability density distributions of an observer's confidence in a diagnostic test, which are analyzed by translating the threshold criteria through all possible values, and then plotting the sensitivity versus specificity for the test. (c) The resulting graph from (b) represents the ROC curve, and imaging tests that maximize the area under this curve are considered more beneficial for accurate clinical classification of lesions. The square point on the curve corresponds to the observer's confidence threshold line shown in (b).

imaging systems.^{132–135} In an ROC study, the readers view a cohort of normal and abnormal radiology images and assign numeric ratings (typically four to six) to each image as an indication of their confidence level that the image shows a clinical abnormality. The resulting rating data are then analyzed, summarized, and plotted on an ROC curve. This graph reveals the relationship between the true-positive fraction (TPF) and the false-positive fraction (FPF) as the reader's confidence level varies. Summary measures of the curve are typically used as an objective measure of the ability of the reader to detect objects in the images, representing the quality of the medical imaging modality when applied in a human diagnosis task. These summary measures include the area under the entire curve and the partial area under the curve in a particular region of interest.

ROC analysis is widely employed to evaluate observer diagnostic performance in a possible situation in which two alternatives exist. In this case, the classification of the stimulus (patients' real condition) and response (radiologists' diagnosis) have only one of two possible choices, normal (nondiseased) or abnormal (diseased). For a diagnostic test of N patients, of which n_1 patients are abnormal and n_0 patients are normal (here $N=n_0+n_1$), there is a 2×2 diagnostic table that completely describes the observer's diagnostic performance, as shown in Fig. 9(a). The table is a listing of all the possible combinations for a pair of binary variables, and the data represent the number of occurrences of combinations of the two variables. In Fig. 9(a), the two variables are the patient's real condition and the observer's diagnosis. The observer gives m_1 positive and m_0 negative diagnosis readings (where also $N=m_0+m_1$) for patients of which n_1 are diseased and n_0 are nondiseased. Within this, the true positives (TP) and false positives (FP) are the number of diseased and nondiseased patients who are diagnosed as diseased (i.e., positive), respec-

tively, while the true negatives (TN) and false negatives (FN) are the number of nondiseased and diseased patients who are diagnosed as not diseased (i.e., negative), respectively.

From the 2×2 diagnostic table, the sensitivity and specificity can be calculated and are commonly used as indications of discriminatory accuracy of the diagnostic study. The sensitivity is the conditional probability of a positive diagnosis ($T+$), given that the patient is in fact diseased or abnormal ($D+$), i.e.:

$$\text{sensitivity} = P(T+ | D+) = \frac{TP}{n_1}.$$

Sensitivity represents the proportion of truly diseased persons in a screened population who are identified as being diseased by the test, and is a measure of the probability of correctly diagnosing a condition. The specificity is the conditional probability of negative diagnosis ($T-$), given that the patient is, in fact, normal ($D-$):

$$\text{specificity} = P(T- | D-) = \frac{TN}{n_0}.$$

Specificity is the proportion of truly nondiseased persons who are identified by the screening test. It is a measure of the probability of correctly identifying a nondiseased person. More clearly, in some medical literature, the sensitivity and specificity are also called “true positive rate” and “true negative rate.”

In practice, the diagnostic discrimination capacity of an observer in a specific diagnostic test is usually not perfect, because diagnoses are made from various states of symptom or evidence. In other words, the diagnosis depends on the confidence level of diagnostic evidence, i.e., the confidence threshold. Thus, it is more informative and meaningful to design the diagnostic tests on a confidence rating scale, either on a fixed number of discrete response categories or a continuous test variable, and then calculate different sensitivity and specificity pairs, which are used to generate a receiver operating characteristic (ROC) curve.

The probability density distribution function of a radiologist’s confidence in a positive diagnosis for a particular diagnostic task is shown schematically in Fig. 9(b). The degree of overlap of the diseased and nondiseased distribution functions completely determines the ability of the test to distinguish diseased patients from nondiseased. As shown in Fig. 9(b), for a specific decision or confidence threshold value x_c , the sensitivity and specificity values can be calculated. As x_c increases, the specificity increases at the expense of sensitivity. To graphically present the relationship of sensitivity and specificity, the ROC curve is generated [Fig. 9(c)], which plots sensitivity versus false positive rate (FPR), defined as (1-specificity). This graph provides the sensitivity and specificity for a given imaging system/reader combination for a range of reader confidence thresholds.

4.2 Interpretation of the Receiver Operating Characteristic Curve

In practice, human-generated ROC curves for a full range of confidence thresholds require a substantial time commitment from the professional image readers involved. This is com-

monly addressed by reducing the number of confidence thresholds required to be assessed by the reader and completing the curve by using various techniques to model the probability density distributions shown in Fig. 9(b). This provides a more continuous representation of the sparse study data. There are several algorithms that have been published to construct ROC curves based on discrete or continuous test data. These algorithms can be divided into two basic categories, nonparametric or parametric, depending on whether the implementation of the algorithm assumes a parametric model. The empirical, nonparametric approach is used to calculate the ROC curve using empirically determined histogram distributions, in which there is no need for structural assumptions and parameters for modeling or fitting. Though the empirical, nonparametric method is easy to implement and robust in general cases, it does not provide a smooth fitted curve and there are no standard statistical measurements, such as confidence levels, available for evaluation. A nonparametric kernel smoothing technique can significantly improve the empirical nonparametric method,^{134,136} but must be done with a careful interpretation of how the smoothing kernel affects the data. In this method, a kernel function estimating the densities of the distribution functions in the diseased and nondiseased populations and a bandwidth is applied and optimized to numerically represent the distribution functions. This analysis results in the generation of a smooth and optimal ROC curve. Having stated this, parametric modeling of the data is most often chosen and provides a strong proven approach to generating smooth ROC curves with statistically useful estimates of the confidence intervals. The standard normal distribution is most commonly used, and provides a mean and standard deviation value, which are used in generation of the confidence interval lines.

One of the great practical challenges in ROC analysis in a typical observer performance study is how to deal with the variation of skill levels among different observers. Although the basic concept of ROC analysis has been understood since the early 1980’s,¹³⁷ the analytical techniques or research tools available at the time had limited practical applicability, until the introduction of the so-called multiple-reader multiple-case (MRMC) ROC paradigm¹³⁸ in the early 1990’s. The MRMC ROC paradigm uses a theoretical model and proposes applying procedures such as the jackknife method (jackknife readers or jackknife cases) or the bootstrap method, to the area under the ROC curve obtained for each reader. These approaches have allowed inclusion of multiple readers into ROC analysis.

4.3 Location Receiver Operating Characteristic Analysis

In standard ROC methods, the major focus is assessing the diagnostic utility of the medical images, where the complexity of the target object location is often eliminated by clearly specifying the region of interest (ROI) in the images. Recent developments in localization-response ROC (LROC) analysis offers more understanding of medical imaging methodology, in terms of measuring the ability to detect and correctly localize the actual target within the reconstructed image. These developments include simultaneous ROC/LROC fitting¹³⁹ and alternative free-response ROC (AFROC) analysis.¹⁴⁰ The

LROC measures the probability of successfully detecting and locating objects within images, versus the probability of falsely detecting objects in normal images, as a function of the detection criteria.

Despite the essential simplicity of the fundamental concept of ROC/LROC analysis, the reproducibility and repeatability of human observers depends on many nonimage-related conditions such as experience, physiological or ambient conditions, number of tests, etc.¹⁴¹ ROC/LROC analysis therefore must be carried out with careful attention to these possibly confounding conditions.

4.4 Receiver Operating Characteristic/Location Receiver Operating Characteristic Analysis in Diffuse Optical Imaging

In diffuse optical imaging, little attention has been paid to assessing the detectability of objects, as assessed by human observers. Yet ROC analysis can be readily applied here, and has been shown useful in assessing diffuse images of arthritic finger joints.¹⁴² It is also likely that assessing how humans perceive nonlinearly reconstructed images may help interpret how image reconstruction algorithms should be tailored. As part of this assessment, the ROC analysis of simulated NIR tomography images were completed and are shown in Song et al.¹⁴³ The purpose of the study was to evaluate diffuse optical imaging by determining the limitation of human readers' ability to detect objects within the image by ROC and LROC analysis. Given a fixed system noise level, four key parameters determine the quality of a reconstructed image: 1. the size of the heterogeneities, 2. the absorption and scattering coefficients' contrast between the heterogeneities and the homogeneous tissue, 3. the number of reconstruction iterations used in the image formation, and 4. the location of the region of interest (ROI) in heterogeneous images. Human observer performance is reported in terms of the related area under the curve (AUC) value with error correction of ROC and LROC curves.

Typical reconstructed images are shown in Fig. 10. Each image contains an inclusion with a different size and contrast level located to the right of the image field. These images are shown to illustrate the visible quality of the tomographic images, and how decreasing size or contrast diminishes detectability of the object within the reconstructed image. The human observer analysis was completed on a large number of images, similar to the ones in Fig. 10. The resulting ROC data of area under the curve (AUC) is shown in Fig. 11, using 600 images for each object size, and four observers, as discussed in Song et al.¹⁴³ The data are separated into studies, one in which the contrast was held constant while the size was varied [Fig. 11(a)], and the other in which the effect of varying contrast was considered for a constant anomaly size [Fig. 11(b)]. For each of these studies, the location of the object was randomly moved around within the imaged domain, and control images with no objects were also used. Not surprisingly, as the object size or contrast decreases, the ability of humans to detect the objects decreases. One interesting observation is that the LROC AUC values drop to almost zero for the smallest anomaly sizes and lowest anomaly contrast values, as seen in Figs. 11(a) and 11(b), while the ROC AUC values remain high in these regimes. These numbers indicate that humans

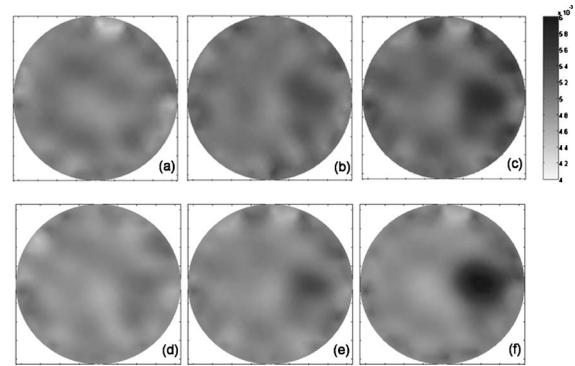


Fig. 10 Diffuse optical tomography images reconstructed from simulated data showing the effect of ROI size and contrast on image quality. The ROI location was held constant in each test field, as was the reconstruction iteration number, while the size and absorption contrast was varied. In the top row of images, the size was fixed at 12 mm diam, and the contrast values were varied in the original data with (a) 1.1, (b) 1.4, and (c) 2.0. In the bottom row of images, the contrast was fixed at 2.0 and the size was varied with (d) 4 mm, (e) 10 mm, and (f) 16 mm diam. The background optical properties were $\mu_a = 0.004 \text{ mm}^{-1}$ and $\mu_s' = 1.0 \text{ mm}^{-1}$.

are still able to detect the presence of local abnormalities in the images, even when their ability to determine the location of the object is diminished or near zero.

In Fig. 11(c), ROC and LROC AUC analysis was used to examine the effect of the number of iterations in the nonlinear reconstruction algorithm by presenting the observers with image sets at different levels of iterations. The object size was maintained at 10 mm, and the contrast and location of the object were varied randomly. The AUC values for both ROC and LROC are nearly constant, but with a subtle but significant rise in LROC at lower iteration numbers. This indicates that observers are better able to localize the region at fewer numbers of iterations. This may be somewhat counterintuitive; however, it is explained by understanding that even though the modeled and measured data more closely match after more iterations, these reconstructed images may also contain a higher degree of spatial noise.

In Fig. 11(d), the effect of the location of the object was assessed by similar LROC and ROC analyses. It is well known that the imaging field response is highly spatially variant in diffuse tomography, and this is demonstrated directly from the images using LROC curves. The ability to localize objects increases (i.e., the AUC of the LROC increases) as the object moves toward the edge of the imaging field. This reflects the fact that the diffuse imaging provides a better reconstruction of objects closer to the edge of the field.

Further use of ROC analysis in detection of objects will allow improved assessment of the factors that limit the use of diffuse tomography methods in detection tasks.

4.5 Receiver Operating Characteristic Analysis of Mammography with Adjuvant Imaging

Near-infrared tomography is now in clinical trials at several centers for breast cancer tumor imaging. The use of NIR as a screening tool has been the subject of research for several years. A clinical study of multispectral NIR tomography in breast cancer imaging was recently summarized by Poplack et

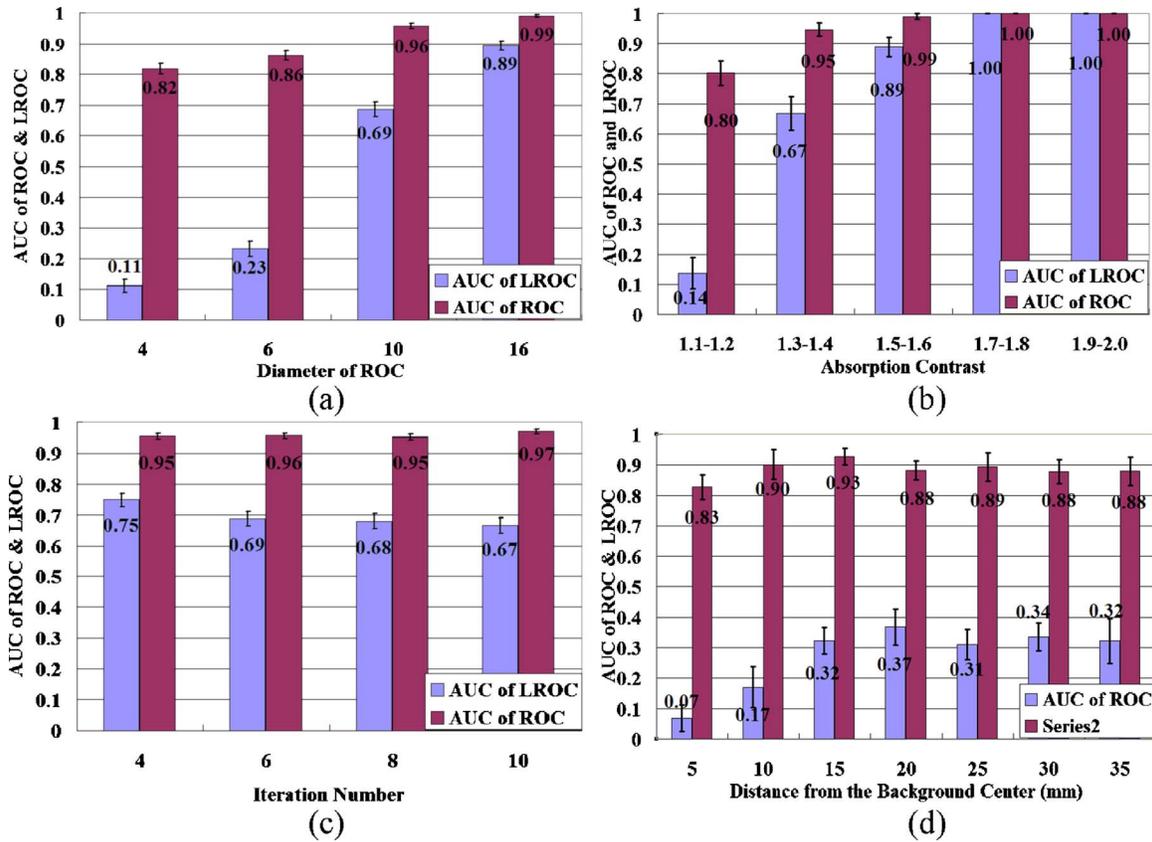


Fig. 11 ROC analysis of diffuse optical tomography images with area under the curve (AUC) values of ROC and LROC and their standard errors. In (a), the heterogeneity size study is shown. The objects all have the same contrast $C=2.0$, and six iterations were used. In (b), the heterogeneity contrast study is shown. The diameter of ROI was equal to 10 mm for all, and six iterations were used. In (c), the reconstruction process iteration number was studied. In (d), the results of a heterogeneity location study are shown, where the object location is varied from the center to the edge of the field. In (c) and (d), the object was a fixed diameter of 10 mm and the absorption contrast was varied randomly from 1.1 to 2.0.

al.¹⁴⁴ In this study, a six wavelength, multispectral imaging system was used to recover images of breast hemoglobin, oxygen saturation, water, and scattering properties in more than 100 patients. ROC analysis was not completed on this study; however, quantification of the contrast for each tumor region relative to the background breast tissue was completed. The tumor location information from mammographic images, such as size, distance from the chest wall, and depth within the breast, were provided prior to NIR imaging to allow quantification of the ROI values. Typical images of total hemoglobin (HbT), oxygen saturation (StO₂), water, scattering amplitude, and scattering power values are shown in Fig. 12(a).

Using the compiled values of contrast in the cancer and benign tumors, the curve of sensitivity and specificity was generated using a parametric analysis of the relative contrast data. The decision criterion was continuously varied to draw the complete ROC curve. Figure 12(b) shows the ROC curves based on the normalized hemoglobin values of both cancer tumors relative to benign tumors. The data in the graph were processed for two tumor sizes, those above 6 mm and those below 6 mm, to illustrate the fact that the size of the tumor plays a major factor in detectability of the cancer. If the tumor size was greater than 6 mm, the NIR tomography images have reasonably high diagnostic accuracy, with the AUC equal to 0.88. However, if the tumors were smaller than

6 mm, the system would fail to distinguish normal and abnormal cases.

The analysis presented here cannot strictly be termed ROC analysis, due to the lack of a human observer detecting the presence or estimating the likelihood of an object being present; however, the quantification of the relative contrast values allowed analysis of the sensitivity-specificity tradeoff, similar to what is determined in a formal ROC curve analysis. The inability to achieve high AUC performance for all tumor sizes indicates that the imaging modality may be inadequate as a screening tool to detect small cancers. Alternatively, relatively high AUC values for larger tumors indicate that the modality is efficient at differentiating malignant versus benign tumors above 6 mm in diameter. Further analysis of diffuse tomography used for an imaging medium to large sized cancers is likely to be a promising avenue for this modality. A similar analysis for fluorescence tomography would also be beneficial.

5 Summary

In summary, the tools for imaging system characterization, evaluation, and analysis of performance and use are well developed in the medical physics and radiology research communities. As new imaging modalities such as diffuse tomog-

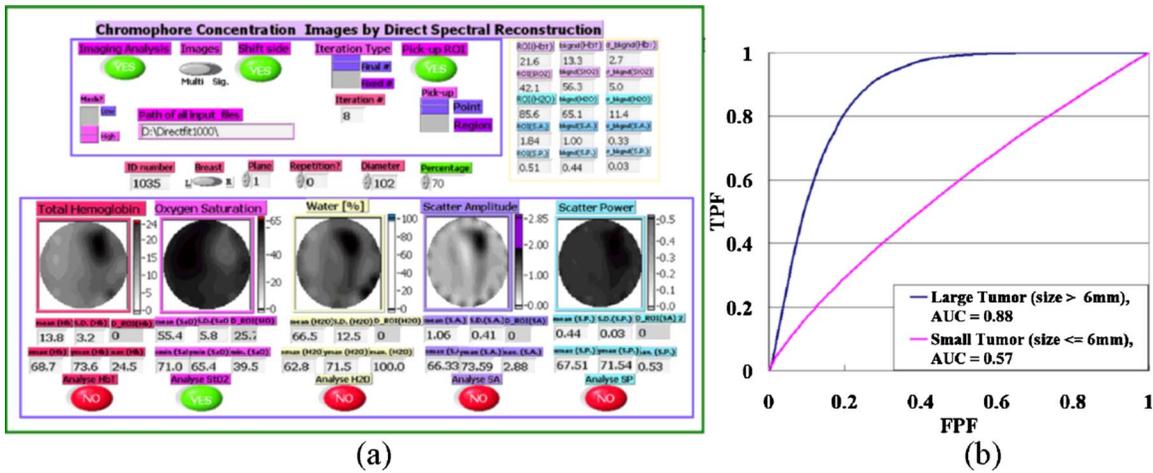


Fig. 12 (a) In this figure, a representative image is shown for a cancer case, with a focal increase in hemoglobin, decrease in oxygen saturation, increase in water, and increase in scattering at the site of the tumor. These values were quantified for a series of tumors used in the study, and an ROC curve was developed for classification of cancer tumors relative to benign tumors, based on the hemoglobin concentration value. The curve (b) is shown for both small (diameter ≤ 6 mm) and large (diameter > 6 mm) breast cancer tumors, to illustrate that detectability of smaller tumors is not feasible with this type of tomography, but that detection of larger tumors may be clinically beneficial.

raphy become tools that enter the clinical realm, it will be increasingly important to characterize the systems and images based on these well known and accepted standards. In this review, the initial stages are introduced and discussed in the context of diffuse tomography. The tools of image resolution are perhaps least well suited for analysis of the imaging system performance, as the systems generally have poor and spatially variant resolution, complicating the analysis considerably. If the value of diffuse tomography lies in the area of characterizing contrast due to hemoglobin, water, lipids, scattering, or luminescence, it will likely become increasingly important to use contrast-detail analysis to assess performance and compare individual systems and algorithms. Objective and automated use of contrast-detail analysis is possible using CNR thresholding; however, complete analysis will require the use of observers to analyze multiple images.

Implementation of ROC analysis will be useful when the imaging systems enter clinical trials, and are being evaluated in sufficient numbers of subjects to warrant use of this methodology. Current implementations of NIR tomography for breast cancer imaging are perhaps the most clinically advanced of any diffuse imaging application, yet insufficient numbers exist today to systematically evaluate these systems based on ROC analysis. The predominant role of the type of analysis is assessing an imaging modality used in a screening or detection mode. If diffuse imaging is not used in a screening mode, but rather a characterization or quantification mode, then ROC analysis has little role in assessing the system performance.

Further attention to these and other image analysis tools is imperative for diffuse imaging to grow and blend into current radiology practice, as well as provide the language for translating these systems into the clinical research world.

Acknowledgments

This study has been supported through National Cancer Institute grants PO1CA80139, RO1CA69544, and RO1CA-109558. The authors would like to acknowledge important

discussions and collaboration with the other people in our near-infrared tomography group, including Shudong Jiang, Subhadra Srinivasan, Xin Wang, and Phaneendra Yalavarthy.

References

1. W. M. Star, J. P. A. Marijnissen, and M. J. C. van Gemert, "Light dosimetry in optical phantoms and in tissues: I. Multiple flux and transport theory," *Phys. Med. Biol.* **33**(4), 437–454 (1988).
2. W. M. Star, "Diffusion theory of light transport," in *Optical-Thermal Response of Laser-Irradiated Tissue*, A. J. Welch and M. J. C. van Gemert, Eds., pp. 131–206, Plenum Press, New York (1995).
3. W. M. Star, "Light dosimetry in vivo," *Phys. Med. Biol.* **42**(5), 763–788 (1997).
4. S. R. Arridge, "Optical tomography in medical imaging," *Inverse Probl.* **15**(2), R41–R93 (1999).
5. M. J. C. van Gemert, S. L. Jacques, H. Sterenborg, and W. M. Star, "Skin optics," *IEEE Trans. Biomed. Eng.* **36**(12), 1146–1154 (1989).
6. S. T. Flock, M. S. Patterson, B. C. Wilson, and D. R. Wyman, "Monte Carlo modeling of light propagation in highly scattering tissues I: Model predictions and comparison with diffusion theory," *IEEE Trans. Biomed. Eng.* **36**(12), 1162–1168 (1989).
7. W. M. Star, *Comparing the P3-Approximation with Diffusion Theory and Monte Carlo Calculations of Light Propagation in a Slab Geometry*, pp. 146–154, SPIE Institute Series, IS 5, SPIE Press, Bellingham, WA (1989).
8. L. Wang, S. L. Jacques, and L. Zheng, "MCML—Monte Carlo modeling of light transport in multi-layered tissues," *Comput. Methods Biomed.* **47**(2), 131–146 (1995).
9. L. Wang and S. L. Jacques, "Optimized radial and angular positions in Monte Carlo modeling," *Med. Phys.* **21**(7), 1081–1083 (1994).
10. E. M. Sevick-Muraca, "Computations of time-dependent photon migration for biomedical optical imaging," *Methods Enzymol.* **240**, 748–781 (1994).
11. S. Behin-Ain, T. van Doorn, and J. R. Patterson, "Spatial resolution in fast timeresolved transillumination imaging: an indeterministic Monte Carlo approach," *Phys. Med. Biol.* **47**(16), 2935–2945 (2002).
12. S. A. Boppart, T. F. Deutsch, and D. W. Rattner, "Optical imaging technology in minimally invasive surgery. Current status and future directions," *Surg. Endosc.* **13**(7), 718–722 (1999).
13. E. Berber and A. E. Siperstein, "Understanding and optimizing laproscopic videosystems," *Suom Hammaslaak Toim* **15**(8), 781–787 (2001).
14. D. C. Sullivan, "Challenges and opportunities for in vivo imaging in oncology," *Technol. Cancer Res. Treat.* **1**(6), 419–422 (2002).
15. C. E. Hooper, R. E. Ansorge, H. M. Browne, and P. Tomkins, "CCD

- imaging of luciferase gene expression in single mammalian cells," *J. Biolumin. Chemilumin.* **5**(2), 123–130 (1990).
16. C. H. Contag and M. H. Bachmann, "Advances in *in vivo* bioluminescence imaging of gene expression," *Annu. Rev. Biomed. Eng.* **4**, 235–260 (2002).
 17. T. Troy, D. Jekic-McMullen, L. Sambucetti, and B. Rice, "Quantitative comparison of the sensitivity of detection of fluorescent and bioluminescent reporters in animal models," *Mol. Imaging* **3**(1), 9–23 (2004).
 18. S. Fantini, M. A. Franceschini, G. Gaida, E. Gratton, H. Jess, W. W. Mantulin, K. T. Moesta, P. M. Schlag, and M. Kaschke, "Frequency-domain optical mammography: edge effect corrections," *Med. Phys.* **23**, 149–157 (1996).
 19. V. E. Pera, E. L. Heffer, H. Siebold, O. Schutz, S. Heywang-Kobrunner, L. Gotz, A. Heinig, and S. Fantini, "Spatial second-derivative image processing: an application to optical mammography to enhance the detection of breast tumors," *J. Biomed. Opt.* **8**(3), 517–524 (2003).
 20. J. Ripoll and V. Ntziachristos, "Imaging scattering media from a distance: Theory and applications of noncontact optical tomography," *Mod. Phys. Lett. B* **18**(28–29), 1403–1431 (2004).
 21. V. Ntziachristos, A. G. Yodh, M. Schnall, and B. Chance, "Concurrent MRI and diffuse optical tomography of breast after indocyanine green enhancement," *Proc. Natl. Acad. Sci. U.S.A.* **97**(6), 2767–2772 (2000).
 22. B. W. Pogue, S. P. Poplack, T. O. McBride, W. A. Wells, K. S. Osterman, U. L. Osterberg, and K. D. Paulsen, "Quantitative hemoglobin tomography with diffuse near-infrared spectroscopy: Pilot results in the breast," *Radiology* **218**(1), 261–266 (2001).
 23. H. Jiang, Y. Xu, N. Iftimia, J. Eggert, K. Klove, L. Baron, and L. Fajardo, "Three-dimensional optical tomographic imaging of breast in a human subject," *IEEE Trans. Med. Imaging* **20**(12), 1334–1340 (2001).
 24. Y. Xu, N. Iftimia, H. Jiang, L. L. Key, and M. B. Bolster, "Three-dimensional diffuse optical tomography of bones and joints," *J. Biomed. Opt.* **7**(1), 88–92 (2002).
 25. X. Wang, B. W. Pogue, S. Jiang, X. Song, K. D. Paulsen, C. Kogel, S. P. Poplack, and W. A. Wells, "Approximation of Mie Scattering parameters from near-infrared tomography of health breast tissue *in vivo*," *J. Biomed. Opt.* **10**, 051704 (2004).
 26. V. Ntziachristos, C. H. Tung, C. Bremer, and R. Weissleder, "Fluorescence molecular tomography resolves protease activity *in vivo*," *Nat. Med.* **8**(7), 757–760 (2002).
 27. V. Ntziachristos, C. Bremer, and R. Weissleder, "Fluorescence imaging with near-infrared light: new technological advances that enable *in vivo* molecular imaging," *Eur. Radiol.* **13**(1), 195–208 (2003).
 28. R. B. Schulz, J. Ripoll, and V. Ntziachristos, "Experimental fluorescence tomography of tissues with noncontact measurements," *IEEE Trans. Med. Imaging* **23**(4), 492–500 (2004).
 29. V. Ntziachristos, E. A. Schellenberger, J. Ripoll, D. Yessayan, E. Graves, A. Bogdanov, Jr., L. Josephson, and R. Weissleder, "Visualization of antitumor treatment by means of fluorescence molecular tomography with an annexin VCy5.5 conjugate," *Proc. Natl. Acad. Sci. U.S.A.* **101**(33), 12294–12299 (2004).
 30. S. R. Arridge and M. Schweiger, "Inverse methods for optical tomography," in *Information Processing in Medical Imaging*, H. H. Barrett, Ed., pp. 259–277, Springer-Verlag, Flagstaff, AZ (1993).
 31. K. D. Paulsen and H. Jiang, "Enhanced frequency-domain optical image reconstruction in tissues through total-variation minimization," *Appl. Opt.* **35**(19), 3447–3458 (1996).
 32. S. R. Arridge and J. C. Hebden, "Optical imaging in medicine: II. Modelling and reconstruction," *Phys. Med. Biol.* **42**(5), 841–853 (1997).
 33. S. R. Arridge and M. Schweiger, "Image reconstruction in optical tomography," *Philos. Trans. R. Soc. London, Ser. B* **352**, 717–726 (1997).
 34. H. Dehghani, B. W. Pogue, S. P. Poplack, and K. D. Paulsen, "Multiwavelength three-dimensional near-infrared tomography of the breast: initial simulation, phantom, and clinical results," *Appl. Opt.* **42**(1), 135–145 (2003).
 35. H. Dehghani, B. W. Pogue, S. Jiang, B. Brooksby, and K. D. Paulsen, "Three dimensional optical tomography: Resolution in small object imaging," *Appl. Opt.* **42**(1), 135–145 (2003).
 36. S. Srinivasan, B. W. Pogue, B. Brooksby, S. Jiang, H. Dehghani, C. Kogel, W. A. Wells, S. Poplack, and K. D. Paulsen, "Near-infrared characterization of breast tumors *in-vivo* using spectrally-constrained reconstruction," *Technol. Cancer Res. Treat.* **13**, 195–202 (2006).
 37. B. Brooksby, S. Jiang, H. Dehghani, B. W. Pogue, K. D. Paulsen, J. B. Weaver, C. Kogel, and S. P. Poplack, "Combining near infrared tomography and magnetic resonance imaging to study *in vivo* breast tissue: implementation of a Laplacian-type regularization to incorporate MR structure," *J. Biomed. Opt.* **10**(5), 050504-1-10 (2005).
 38. E. Hecht, *Optics*, 4th ed., Boston, Addison Wesley, (2002).
 39. J. T. Bushberg, J. A. Seibert, E. M. Leidholdt, and J. M. Boone, *The Essential Physics of Medical Imaging*, 2nd ed., Lippencott Williams and Wilkins, Philadelphia (2002).
 40. G. Vargas, E. K. Chan, J. K. Barton, H. G. Rylander, and A. J. Welch, "Use of an agent to reduce scattering in skin," *Lasers Surg. Med.* **24**(2), 133–141 (1999).
 41. K. Dowling, M. Dayel, S. Hyde, P. French, M. Lever, J. Hares, and A. Dymoke-Bradshaw, "High resolution time-domain fluorescence lifetime imaging for biomedical applications," *J. Mod. Opt.* **46**(2), 199–209 (1999).
 42. S. Hyde, R. Jones, N. Barry, J. Dainty, P. French, K. Kwolek, D. Nolte, and M. Melloch, "Depth-resolved holography through turbid media using photorefraction," *IEEE J. Sel. Top. Quantum Electron.* **2**(4), 965–975 (1996).
 43. N. J. Kline and P. J. Treado, Raman chemical imaging of breast tissue," *J. Raman Spectrosc.* **28**(2-3), 119–124 (1998).
 44. B. W. Pogue and T. Hasan, "Targeting in PDT and Photoimaging," *Opt. Photonics News* **August**, 36–43 (2003).
 45. S. L. Jacques, J. C. Ramella-Roman, and K. Lee, "Imaging skin pathology with polarized light," *J. Biomed. Opt.* **7**(3), 329–340 (2002).
 46. V. Chernomordik, A. Gandjbakhche, M. Lepore, R. Esposito, and I. Delfino, "Depth dependence of the analytical expression for the width of the point spread function (spatial resolution) in time-resolved transillumination," *J. Biomed. Opt.* **6**(4), 441–445 (2001).
 47. J. P. Culver, R. Choe, M. J. Holboke, L. Zubkov, T. Durduran, A. Slemple, V. Ntziachristos, B. Chance, and A. G. Yodh, "Three-dimensional diffuse optical tomography in the parallel plane transmission geometry: evaluation of a hybrid frequency domain/continuous wave clinical system for breast imaging," *Med. Phys.* **30**(2), 235–247 (2003).
 48. D. A. Boas, M. A. O'Leary, B. Chance, and A. G. Yodh, "Scattering of diffuse photon density waves by spherical inhomogeneities within turbid media: analytic solution and applications," *Proc. Natl. Acad. Sci. U.S.A.* **91**(11), 4887–4891 (1994).
 49. E. M. Sevick, C. L. Burch, and B. Chance, "Near-infrared optical imaging of tissue phantoms with measurement in the change of optical path lengths," *Adv. Exp. Med. Biol.* **345**, 815–823 (1994).
 50. S. R. Arridge, M. Cope, and D. T. Delpy, "The theoretical basis for the determination of optical pathlengths in tissue: temporal and frequency analysis," *Phys. Med. Biol.* **37**(7), 1531–1560 (1992).
 51. M. S. Patterson, B. W. Pogue, and B. C. Wilson, "Computer simulation and experimental studies of optical imaging with photon density waves," in *Medical Optical Tomography: Functional Imaging and Monitoring*, G. Muller, editor, pp. 87–120, SPIE Press, Bellingham, WA (1993).
 52. R. L. Barbour, H. L. Graber, Y. Wang, J. H. Chang, and R. Aronson, "A perturbation approach for optical diffusion tomography using continuous-wave and time resolved data," in *Medical Optical Tomography: Functional Imaging and Monitoring*, G. Muller, Ed., pp. 87–120, SPIE Press, Bellingham, WA (1993).
 53. B. J. Tromberg, L. O. Svaasand, T. T. Tsay, and R. C. Haskell, "Properties of photon density waves in multiple-scattering media," *Appl. Opt.* **32**, 607–616 (1993).
 54. M. S. Patterson, S. Andersson-Engels, B. C. Wilson, and E. K. Osei, "Absorption spectroscopy in tissue-simulating materials: a theoretical and experimental study of photonpaths," *Appl. Opt.* **34**, 22–30 (1995).
 55. E. M. Sevick-Muraca, G. Lopez, J. S. Reynolds, T. L. Troy, and C. L. Hutchinson, "Fluorescence and absorption contrast mechanisms for biomedical optical imaging using frequency-domain techniques," *Photochem. Photobiol.* **66**(1), 55–64 (1997).
 56. J. A. Moon, R. Mahon, M. D. Duncan, and J. Reintjes, "Resolution limits for imaging through turbid media with diffuse light," *Opt. Lett.* **18**(19), 1591–1593 (1993).
 57. J. M. Kaltenbach and M. Kaschke, "Frequency and time-domain modelling of light transport in random media," in *Medical Optical Tomography: Functional Imaging and Monitoring*, G. Muller, Ed.,

- SPIE Press, Bellingham, WA (1993).
58. J. B. Fishkin and E. Gratton, "Propagation of photon-density waves in strongly scattering media containing an absorbing semi-infinite plane bounded by a straight edge," *J. Opt. Soc. Am. A* **10**(1), 127–140 (1993).
 59. D. J. Hall, J. C. Hebden, and D. T. Delpy, "Evaluation of spatial resolution as a function of thickness for time-resolved optical imaging of highly scattering media," *Med. Phys.* **24**(3), 361–368 (1997).
 60. H. L. Liu, C. L. Matson, K. Lau, and R. R. Mapakshi, "Experimental validation of a backpropagation algorithm for three-dimensional breast tumor localization," *IEEE J. Sel. Top. Quantum Electron.* **5**(4), 1049–1057 (1999).
 61. B. W. Pogue, T. O. McBride, J. Prewitt, U. L. Osterberg, and K. D. Paulsen, "Spatially variant regularization improves diffuse optical tomography," *Appl. Opt.* **38**(13), 2950–2961 (1999).
 62. J. M. Boone and J. A. Seibert, "An analytical edge spread function model for computer fitting and subsequent calculation of the LSF and MTF," *Med. Phys.* **21**(10), 1541–1545 (1994).
 63. D. T. Delpy, M. Cope, P. Van der Zee, S. R. Arridge, S. Wray, and J. S. Wyatt, "Estimation of optical pathlength through tissue from direct time of flight measurement," *Phys. Med. Biol.* **33**, 1433–1442 (1988).
 64. A. H. Gandjbakhche, R. Nossal, and R. F. Bonner, "Resolution limits for optical transillumination of abnormalities deeply embedded in tissues," *Med. Phys.* **21**(2), 185–191 (1994).
 65. D. R. Kirkby and D. T. Delpy, "Measurement of tissue temporal point spread function (TPSF) by use of a gain-modulated avalanche photodiode detector," *Phys. Med. Biol.* **41**(5), 939–949 (1996).
 66. B. W. Pogue, T. O. McBride, J. Prewitt, U. L. Osterberg, and K. D. Paulsen, "Spatially variant regularization improves diffuse optical tomography," *Appl. Opt.* **38**(13), 2950–2961 (1999).
 67. D. A. Boas, M. A. O'Leary, B. Chance, and A. G. Yodh, "Detection and characterization of optical inhomogeneities with diffuse photon density waves: a signal-to-noise analysis," *Appl. Opt.* **36**, 75–92 (1997).
 68. J. Chang, H. L. Graber, and R. L. Barbour, "Luminescence optical tomography of dense scattering media," *J. Opt. Soc. Am. A* **14**(1), 288–299 (1997).
 69. S. R. Arridge and M. Schiewer, "Gradient-based optimisation scheme for optical tomography," *Opt. Express* **2**(6), 212–226 (1998).
 70. H. Dehghani, B. W. Pogue, J. Shudong, B. Brooksby, and K. D. Paulsen, "Three-dimensional optical tomography: resolution in small-object imaging," *Appl. Opt.* **42**(16), 3117–3128 (2003).
 71. M. J. Eppstein, D. E. Dougherty, T. L. Troy, and E. M. Sevick-Muraca, "Biomedical optical tomography using dynamic parameterization and Bayesian conditioning on photon migration measurements," *Appl. Opt.* **38**(10), 2138–2150 (1999).
 72. H. B. Jiang, K. D. Paulsen, U. L. Osterberg, B. W. Pogue, and M. S. Patterson, "Simultaneous reconstruction of optical-absorption and scattering maps in turbid media from near-infrared frequency-domain data," *Opt. Lett.* **20**(20), 2128–2130 (1995).
 73. H. Jiang, K. D. Paulsen, U. L. Osterberg, B. W. Pogue, and M. S. Patterson, "Optical image reconstruction using frequency-domain data: simulations and experiments," *J. Opt. Soc. Am. A* **13**(2), 253–266 (1996).
 74. T. F. Budinger, D. A. Benaron, and A. P. Koretsky, "Imaging transgenic animals," *Annu. Rev. Biomed. Eng.* **1**, 611–648 (1999).
 75. C. H. Contag and M. H. Bachmann, "Advances in *in vivo* bioluminescence imaging of gene expression," *Annu. Rev. Biomed. Eng.* **4**, 235–260 (2002).
 76. C. H. Contag and B. D. Ross, "It's not just about anatomy: *In vivo* bioluminescence imaging as an eyepiece into biology," *J. Magn. Reson. Imaging* **16**, 378–387 (2002).
 77. T. Troy, D. Jekic-McMullen, L. Sambucetti, and B. Rice, "Quantitative comparison of the sensitivity of detection of fluorescent and bioluminescent reporters in animal models," *Mol. Imaging* **3**(1), 9–23 (2004).
 78. G. Cohen, "Contrast-detail-dose analysis of six different computed tomographic scanners," *J. Comput. Assist. Tomogr.* **3**(2), 197–203 (1979).
 79. G. Cohen, L. K. Wagner, S. R. Amtey, and F. A. Di Bianca, "Contrast-detail-dose and dose efficiency analysis of a scanning digital and a screen-film-grid radiographic system," *Med. Phys.* **8**(3), 358–367 (1981).
 80. G. Cohen and F. A. DiBianca, "The use of contrast – detail – dose evaluation of image quality in a computed tomographic scanner," *J. Comput. Assist. Tomogr.* **3**(2), 189–195 (1979).
 81. K. Faulkner and B. M. Moores, "Contrast-detail assessment of computed tomography scanners," *Phys. Med. Biol.* **31**(9), 993–1003 (1986).
 82. R. T. Constable and R. M. Henkelman, "Contrast resolution and detectability in MR imaging," *J. Comput. Assist. Tomogr.* **15**(2), 297–303 (1991).
 83. S. W. Smith and H. Lopez, "A contrast-detail analysis of diagnostic ultrasound imaging," *Med. Phys.* **9**(1), 4–12 (1982).
 84. K. J. Robinson, C. J. Kotre, and K. Faulkner, "The use of contrast-detail test objects in the optimization of optical density in mammography," *Br. J. Radiol.* **68**, 277–282 (1995).
 85. H. Liu, L. L. Fajardo, J. R. Barrett, and R. A. Baxter, "Contrast-detail detectability analysis: comparison of a digital spot mammography system and an analog screen-film mammography system," *Acad. Radiol.* **4**(3), 197–203 (1997).
 86. M. J. Tapiovaara, "Efficiency of low-contrast detail detectability in fluoroscopic imaging," *Med. Phys.* **24**(5), 655–664 (1997).
 87. H. G. Chotas and C. E. Ravin, "Digital chest radiography with a solid-state flatpanel x-ray detector: contrast-detail evaluation with processed images printed on film hard copy," *Radiology* **218**(3), 679–682 (2001).
 88. S. Peer, U. Neitzel, S. M. Giacomuzzi, R. Peer, E. Gassner, I. Steingruber, and W. Jäschke, "Comparison of low-contrast detail perception on storage phosphor radiographs and digital flat panel detector images," *IEEE Trans. Med. Imaging* **20**(3), 239–242 (2001).
 89. S. Suryanarayanan, A. Karellas, S. Vedantham, H. Ved, S. P. Baker, and C. J. D'Orsi, "Flat-panel digital mammography system: contrast-detail comparison between screen-film radiographs and hard-copy images," *Radiology* **225**(3), 801–807 (2002).
 90. G. Cohen, D. L. McDaniel, and L. K. Wagner, "Analysis of variations in contrast detail experiments," *Med. Phys.* **11**(4), 469–473 (1984).
 91. C. A. Kelsey, R. D. Moseley, Jr., J. F. Garcia, F. A. Mettler, Jr., T. W. Parker, and J. H. Juhl, "ROC and contrast detail image evaluation tests compared," *Radiology* **154**(3), 629–631 (1985).
 92. H. H. Barrett, C. K. Abbey, and E. Clarkson, "Objective assessment of image quality. III. ROC metrics, ideal observers, and likelihood-generating functions," *J. Opt. Soc. Am. A* **15**(6), 1520–1535 (1998).
 93. H. H. Barrett, "Objective assessment of image quality: effects of quantum noise and object variability," *J. Opt. Soc. Am. A* **7**(7), 1266–1278 (1990).
 94. B. W. Pogue, C. Willscher, T. O. McBride, U. L. Osterberg, and K. D. Paulsen, "Contrast-detail analysis for detection and characterization with near-infrared diffuse tomography," *Med. Phys.* **27**(12), 2693–2700 (2000).
 95. X. Song, B. W. Pogue, S. Jiang, M. M. Doyley, H. Dehghani, T. D. Tosteson, and K. D. Paulsen, "Automated region detection based on the contrast-to-noise ratio in near-infrared tomography," *Appl. Opt.* **43**(5), 1053–1062 (2004).
 96. S. C. Davis, B. W. Pogue, H. Dehghani, and K. D. Paulsen, "Contrast-detail analysis characterizes diffuse optical fluorescence tomography image reconstruction," *J. Biomed. Opt.* **10**(5), 050501-1-3 (2005).
 97. W. A. Kalender, *Computed Tomography: Fundamentals, System Technology, Image Quality, Applications*, Publicis MCD Werbeagentur, Munich (2000).
 98. K. J. Robson, C. J. Kotre, and K. Faulkner, "The use of a contrast-detail test object in the optimization of optical density in mammography," *Br. J. Radiol.* **68**(807), 277–282 (1995).
 99. R. Hendrick and E. Berns, "Optimizing mammographic techniques," in *Radiological Society of North America*, Radiological Society of North America (1999).
 100. F. Fischbach, T. Freund, M. Pech, M. Werk, C. Bassir, B. Stoeber, R. Felix, and J. Ricke, "Comparison of indirect CsI/a:Si and direct a:Se digital radiography. An assessment of contrast and detail visualization," *Acta Radiol.* **44**(6), 616–621 (2003).
 101. O. W. Hamer, M. Volk, Z. Zorger, S. Feuerbach, and M. Strotzer, "Amorphous silicon, flat-panel, x-ray detector versus storage phosphor-based computed radiography: contrast-detail phantom study at different tube voltages and detector entrance doses," *Invest. Radiol.* **38**(4), 212–220 (2003).
 102. Z. F. Lu, E. L. Nickoloff, J. C. So, and A. K. Dutta, "Comparison of computed radiography and film/screen combination using a contrast-detail phantom," *J. Appl. Clin. Med. Phys.* **4**(1), 91–98 (2003).

103. X. J. Rong, C. C. Shaw, X. Liu, M. R. Lemacks, and S. K. Thompson, "Comparison of an amorphous silicon/cesium iodide flat-panel digital chest radiography system with screen/film and computed radiography systems—a contrast-detail phantom study," *Med. Phys.* **28**(11), 2328–2335 (2001).
104. W. J. Veldkamp, M. A. Thijssen, and N. Karssemeijer, "The value of scatter removal by a grid in full field digital mammography," *Med. Phys.* **30**(7), 1712–1718 (2003).
105. S. Peer, S. M. Giacomuzzi, R. Peer, E. Gassner, I. Steingruber, and W. Jaschke, "Resolution requirements for monitor viewing of digital flat-panel detector radiographs: a contrast detail analysis," *Eur. Radiol.* **13**(2), 413–417 (2003).
106. G. C. Weatherburn and J. G. Davies, "Comparison of film, hard copy computed radiography (CR) and soft copy picture archiving and communication (PACS) systems using a contrast detail test object," *Br. J. Radiol.* **72**(861), 856–863 (1999).
107. O. Kocsis, L. Costaridou, L. Varaki, E. Likaki, C. Kalogeropoulou, S. Skiadopoulos, and G. Panayiotakis, "Visually lossless threshold determination for microcalcification detection in wavelet compressed mammograms," *Eur. Radiol.* **13**(10), 2390–2396 (2003).
108. G. Cohen, J. O. Barnes, and P. M. Pena, "The effects of the film/screen combination on tomographic image quality," *Ultrasound Med. Biol.* **129**(2), 515–520 (1978).
109. S. W. Smith, H. Lopez, and W. J. Bodine, Jr., "Frequency independent ultrasound contrast-detail analysis," *Ultrasound Med. Biol.* **11**(3), 467–477 (1985).
110. S. W. Smith, M. F. Insana, and H. Lopez, "New contrast-detail phantoms for improved precision in lesion detection measurements," *Ultrasound Med. Biol.* **15**(4), 383–393 (1989).
111. T. J. Hall, M. F. Insana, L. A. Harrison, N. M. Soller, and K. J. Schlehr, "Ultrasound contrast-detail analysis: a comparison of low-contrast detectability among scanhead designs," *Med. Phys.* **22**(7), 1117–1125 (1995).
112. T. J. Hall, M. F. Insana, N. M. Soller, and L. A. Harrison, "Ultrasound contrast detail analysis: a preliminary study in human observer performance," *Med. Phys.* **20**(1), 117–127 (1993).
113. C. J. Kotre, N. W. Marshall, and K. Faulkner, "An alternative approach to contrast-detail testing of X-ray image intensifier systems," *Br. J. Radiol.* **65**(776), 686–690 (1992).
114. A. G. Davies, A. R. Cowen, S. M. Kengyelics, R. F. Bury, and T. J. Bruijns, "Threshold contrast detail detectability measurement of the fluoroscopic image quality of a dynamic solid-state digital x-ray image detector," *Med. Phys.* **28**(1), 11–15 (2001).
115. N. W. Marshall and C. J. Kotre, "Measurement and correction of the effects of lag on contrast-detail test results in fluoroscopy," *Phys. Med. Biol.* **47**(6), 947–960 (2002).
116. J. S. Tsuruda and W. G. Bradley, "MR detection of intracranial calcification: a phantom study," *AJNR Am. J. Neuroradiol.* **8**(6), 1049–1055 (1987).
117. B. W. Pogue, T. McBride, U. Osterberg, and K. Paulsen, "Comparison of imaging geometries for diffuse optical tomography of tissue," *Opt. Express* **4**(8), 270–286 (1999).
118. H. Lopez, M. H. Loew, P. F. Butler, M. C. Hill, and R. M. Allman, "A clinical evaluation of contrast-detail analysis for ultrasound images," *Med. Phys.* **17**(1), 48–57 (1990).
119. A. Godavarty, A. B. Thompson, R. Roy, M. Gurfinkel, M. J. Eppstein, C. Zhang, and E. M. Sevick-Muraca, "Diagnostic imaging of breast cancer using fluorescence-enhanced optical tomography: phantom studies," *J. Biomed. Opt.* **9**(3), 488–496 (2004).
120. A. Godavarty, M. J. Eppstein, C. Zhang, and E. M. Sevick-Muraca, "Detection of single and multiple targets in tissue phantoms with fluorescence-enhanced optical imaging: feasibility study," *Radiology* **235**(1), 148–154 (2005).
121. C. H. Contag and B. D. Ross, "It's not just about anatomy: in vivo bioluminescence imaging as an eyepiece into biology," *J. Magn. Reson. Imaging* **16**(4), 378–387 (2002).
122. G. Wang, Y. Li, and M. Jiang, "Uniqueness theorems in bioluminescence tomography," *Med. Phys.* **31**(8), 2289–2299 (2004).
123. K. Vishwanath, B. Pogue, and M. A. Mycek, "Quantitative fluorescence lifetime spectroscopy in turbid media: comparison of theoretical, experimental and computational methods," *Phys. Med. Biol.* **47**(18), 3387–3405 (2002).
124. J. Ripoll, D. Yessayan, G. Zacharakis, and V. Ntziachristos, "Experimental determination of photon propagation in highly absorbing and scattering media," *J. Opt. Soc. Am. A* **22**(3), 546–551 (2005).
125. B. W. Pogue and K. D. Paulsen, "High resolution near infrared tomographic imaging simulations of rat cranium using a priori MRI structural information," *Opt. Lett.* **23**(21), 1716–1718 (1998).
126. J. Chang, H. L. Graber, P. C. Koo, R. Aronson, S. S. Barbour, and R. L. Barbour, "Optical imaging of anatomical maps derived from magnetic resonance images using time independent optical sources," *IEEE Trans. Med. Imaging* **16**, 68–77 (1997).
127. B. Brooksby, H. Dehghani, B. W. Pogue, and K. D. Paulsen, "Near infrared (NIR) tomography breast image reconstruction with a priori structural information from MRI: algorithm development for reconstructing heterogeneities," *IEEE J. Sel. Top. Quantum Electron.* **9**(2), 199–209 (2003).
128. B. Brooksby, B. W. Pogue, S. Jiang, C. Kogel, H. Dehghani, J. B. Weaver, S. P. Poplack, and K. D. Paulsen, "Analysis of adipose and fibroglandular tissue composition with MRI-coupled NIR tomography," *Proc. Natl. Acad. Sci. U.S.A.* **103**(23), 8820–8833 (2006).
129. Q. Zhu, S. H. Kurtzma, P. Hegde, S. Tannenbaum, M. Kane, M. Huang, N. G. Chen, B. Jagjivan, and K. Zarfos, "Utilizing optical tomography with ultrasound localization to image heterogeneous hemoglobin distribution in large breast cancers," *Neoplasia* **7**(3), 263–270 (2005).
130. N. G. Chen, M. Huang, H. Xia, D. Piao, E. Cronin, and Q. Zhu, "Portable near-infrared diffusive light imager for breast cancer detection," *J. Biomed. Opt.* **9**(3), 504–510 (2004).
131. M. Huang, T. Xie, N. G. Chen, and Q. Zhu, "Simultaneous reconstruction of absorption and scattering maps with ultrasound localization: feasibility study using transmission geometry," *Appl. Opt.* **42**(19), 4102–4114 (2003).
132. C. E. Metz, "ROC methodology in radiologic imaging," *Invest. Radiol.* **21**(9), 720–733 (1986).
133. H. H. Barrett, C. K. Abbey, and E. Clarkson, "Objective assessment of image quality: III. ROC metrics, ideal observers and likelihood-generating functions," *J. Opt. Soc. Am. A* **15**(6), 1520–1535 (1998).
134. X. H. Zhou and J. Harezlak, "Comparison of bandwidth selection methods for kernel smoothing of ROC curves," *Stat. Med.* **21**(14), 2045–2055 (2002).
135. D. P. Chakraborty, "Statistical power in observer-performance studies: comparison of the receiver operating characteristic and free-response methods in tasks involving localization," *Acad. Radiol.* **9**(2), 147–156 (2002).
136. K. H. Zou, W. J. Hall, and D. E. Shapiro, "Smooth non-parametric receiver operating characteristic (ROC) curves for continuous diagnostic tests," *Stat. Med.* **16**(19), 2143–2156 (1997).
137. J. A. Swets, "Sensitivities and specificities of diagnostic tests," *J. Am. Med. Assoc.* **248**(5), 548–550 (1982).
138. J. A. Swets, R. M. Pickett, S. F. Whitehead, D. J. Getty, J. A. Schnur, J. B. Swets, and B. A. Freeman, "Assessment of diagnostic technologies," *Science* **205**(4408), 753–759 (1979).
139. R. G. Swensson, "Unified measurement of observer performance in detecting and localizing target objects on images," *Med. Phys.* **23**(10), 1709–1725 (1996).
140. D. Chakraborty, "Statistical power in observer-performance studies: comparison of the receiver operating characteristic and free-response methods in tasks involving localization," *Acad. Radiol.* **9**(2), 147–156 (2002).
141. D. J. Goodenough and K. E. Weaver, "Factors related to low contrast resolution in CT scanners," *Comput. Radiol.* **8**(5), 297–308 (1984).
142. A. K. Scheel, M. Backhaus, A. D. Klose, B. Moa-Anderson, U. J. Netz, K. G. A. Hermann, J. Beuthan, G. A. Muller, G. R. Burmester, and A. H. Hielscher, "First clinical evaluation of sagittal laser optical tomography for detection of synovitis in arthritic finger joints," *Ann. Rheum. Dis.* **64**(2), 239–245 (2005).
143. X. Song, B. W. Pogue, H. Dehghani, S. Jiang, K. D. Paulsen, and T. D. Tosteson, "Location receiver operating characteristic analysis of near-infrared tomography images" (submitted for publication).
144. S. P. Poplack, K. D. Paulsen, A. Hartov, P. M. Meaney, B. W. Pogue, T. D. Tosteson, S. K. Soho, and W. A. Wells, "Electromagnetic breast imaging—pilot results in women with abnormal mammography," *Radiology* (to be published).