Published in final edited form as:

Proc IEEE Inst Electr Electron Eng. 2008 March ; 96(3): 500–511. doi:10.1109/JPROC.2007.913553.

A Task-Based Approach to Adaptive and Multimodality Imaging:

Computation techniques are proposed for figures-of-merit to establish feasibility and optimize use of multiple imaging systems for disease diagnosis and treatment-monitoring.

Eric Clarkson, Matthew A. Kupinski, Harrison H. Barrett [Fellow IEEE], and Lars Furenlid [Fellow IEEE]

The authors are with the College of Optical Sciences and the Department of Radiology, The University of Arizona, Tucson, AZ 85721 USA

Abstract

Multimodality imaging is becoming increasingly important in medical imaging. Since the motivation for combining multiple imaging modalities is generally to improve diagnostic or prognostic accuracy, the benefits of multimodality imaging cannot be assessed through the display of example images. Instead, we must use objective, task-based measures of image quality to draw valid conclusions about system performance. In this paper, we will present a general framework for utilizing objective, task-based measures of image quality in assessing multimodality and adaptive imaging systems. We introduce a classification scheme for multimodality and adaptive imaging systems and provide a mathematical description of the imaging chain along with block diagrams to provide a visual illustration. We show that the task-based methodology developed for evaluating single-modality imaging can be applied, with minor modifications, to multimodality and adaptive imaging. We discuss strategies for practical implementing of task-based methods to assess and optimize multimodality imaging systems.

Keywords

Adaptive imaging; image quality; multimodality

I. INTRODUCTION

Multimodality imaging is becoming increasingly common in both the clinic and the laboratory [1]–[3]. By combining both anatomical and functional images, one can, for example, localize uptake of a tracer or contrast agent to a particular tissue or organ within a subject. In the clinic, multimodality imaging can be used for the diagnosis of disease, staging to help select between therapies, or to assess the efficacy of a course of treatment. In the laboratory, multimodality imaging is used to correlate genomic and phenomic types, make quantitative data more reliable, quantify the damage due to induced disease states or injuries, and assess the usefulness of treatment options.

The motivation for combining imaging modalities is to improve diagnostic or prognostic accuracy, or to provide more accurate answers to scientific questions about particular diseases or treatments. In other words, there is usually a specific task to be performed when using multimodality imaging. Thus, the benefits of multimodality imaging cannot be assessed through the display of example images. Instead, the benefits of combining imaging methods must be assessed using objective, task-based measures of image quality [4]–[6]. This approach is especially important when comparing multimodality imaging to a single modality to justify the added expense of using more than one imaging system. The task-based approach can also

be used to objectively compare the various approaches for combining the data from complementary imaging modalities.

As we will discuss below, there are fundamental similarities between multimodality imaging and adaptive imaging. In adaptive imaging [7], [8], multiple images are taken with a single imaging system that can change its configuration during acquisition in response to acquired data. We show below that the mathematical descriptions of multimodality and adaptive imaging systems are similar. For this reason, we discuss adaptive as well as multimodality imaging.

In the next section, we introduce a classification scheme for multimodality and adaptive imaging systems. For each class of systems, we provide a general mathematical description of the imaging chain and a block diagram to provide a visual illustration. In Section III, we briefly review the fundamentals of task-based assessment of image quality, including the basic tasks and observers that have been used for single-modality imaging. In Section IV, we extend these task-based ideas to multimodality and adaptive imaging systems. In this case, the tasks remain more or less the same but the observers become more complicated. In Section V, we discuss strategies for practical implementation of these task-based methods to assess and optimize multimodality imaging systems. In Section VI, we summarize our results.

II. CLASSIFICATION OF SYSTEMS

There are two fundamental distinctions we use to classify imaging systems where more than one modality or configuration are used. The first distinction, to be elucidated below, is between *isoscopic* and *polyscopic* imaging systems. The second distinction is between adaptive and multimodality imaging systems. In this section, we discuss these fundamental distinctions and provide block diagrams and imaging-chain equations for various classes of systems.

A. Isoscopic and Polyscopic Systems

The essence of multimodality imaging is the use of two or more imaging methods to acquire data on the same patient to perform some specific task of clinical interest. Often one system will be used to acquire anatomical information [e.g., computed tomography (CT), ultrasound, magnetic resonance imaging (MRI)] and the other to acquire functional information [e.g., positron emission tomography (PET), single photon emission computed tomography (SPECT), functional MRI]; in these cases, the modalities measure fundamentally different physical properties of the object. We can, however, interpret the term modality more broadly; for example, two MRI acquisitions with different pulse sequences or two SPECT images with different isotopes could be considered as multimodality imaging.

If the same object property is imaged in both modalities, as when images are taken of the same radiotracer distribution with different imaging systems or acquisition protocols, the imaging is *isoscopic*. This word is derived from the root word “scope,” which means “to look;” and “iso,” which is a prefix indicating equality. Thus, for an isoscopic multimodality system, the two imaging systems are “looking” at the same property. When two distinctly different physical quantities are imaged in the same patient, as in CT/SPECT, then the imaging is *polyscopic*. The prefix “poly” indicates a multiplicity; thus, for a polyscopic multimodality imaging system, multiple properties are being imaged. One interesting example of isoscopic imaging is dual-detector imaging of ^{123}I , which emits 159 keV gamma rays as well as iodine X-rays at around 30 keV. In contrast, dual-isotope SPECT imaging is polyscopic. Different tracers define different objects, even if the same exact imaging system is used.

MR imaging with two different pulse sequences is isoscopic if we think of the object as a three-dimensional vector field, with components consisting of the spin density $\rho(\mathbf{r})$ and the relaxation

times $T_1(\mathbf{r})$ and $T_2(\mathbf{r})$. This formulation, however, requires a nonlinear system description. We can treat MR as a linear imaging systems if we regard the object as the time-dependent magnetization $M(\mathbf{r}, t)$, but in this view, the object depends on the pulse sequence; hence, the imaging is polyscopic.

In the imaging-system equations below, the object of study for an isoscopic study will be a function, possibly vector valued, $f(\mathbf{r})$ or $f(\mathbf{r}, t)$. For polyscopic systems, the object of study for the j th imaging modality will be $f_j(\mathbf{r})$ or $f_j(\mathbf{r}, t)$. To simplify notation, these functions will be indicated by the vectors f and f_j , respectively, in some appropriate function space. This notation is general in that f or f_j may represent scalar or vector-valued functions of $\mathbf{r}$ and/or t .

B. Adaptive and Multimodality Imaging

The distinction between adaptive and multimodality imaging relates to how the image data are used. A traditional multimodality system is polyscopic and generates data vectors

$$g_j = H_j f_j + n_j \quad (1)$$

where $j = 1, 2, \dots, J$ indicates one of the J modalities. The operator H_j is the imaging system operator, g_j is the image data vector, and n_j is the measurement noise, all for the j th modality. On the other hand, an isoscopic multimodality system is described by the equation

$$g_j = H_j f + n_j \quad (2)$$

where the object f remains the same for all modalities.

In contrast, an adaptive system uses the data from one or more acquisitions to modify the system operator for subsequent acquisitions. The simplest form of isoscopic adaptive imaging is described by

$$g_1 = H_1 f + n_1 \quad (3)$$

$$g_2 = H_2(g_1) f + n_2. \quad (4)$$

The image data vector g_1 is generated first using a predetermined system configuration represented by H_1 . This data vector is then used to determine the configuration of the imaging-system operator for the second acquisition. Thus, the H_2 operator is a function of the first data vector g_1 . A polyscopic adaptive system would be described in a similar fashion by the equations

$$g_1 = H_1 f + n_1 \quad (5)$$

$$g_2 = H_2(g_1) f_2 + n_2. \quad (6)$$

Here, the property being imaged varied from the first to the second acquisition.

C. Examples and Block Diagrams

To help illustrate the generality of this mathematical formalism for multimodality and adaptive imaging, we provide example imaging situations with corresponding block diagrams. In all of these diagrams, the symbol $\hat{f}_j$ represents a reconstructed image from the data vector g_j . In the normal use of the system, the observer produces a decision or an estimate, the end result of the process. However, to assess the quality of any of these systems, one must image many subjects and obtain a figure of merit that uses all of the decision outcomes.

The first example is traditional, polyscopic multimodality imaging (Fig. 1). Here, the same subject is imaged using both a functional system such as SPECT or PET along with an anatomical system such as CT. The data are acquired and processed independently to produce two separate reconstructions. These two reconstructions are either combined into a single display or shown separately to the observer who uses this information to make a decision. An isoscopic version of this scheme is also possible with, for example, dual-energy SPECT imaging with a single tracer. In this case, f_1 and f_2 are replaced with f , and all other steps remain the same.

The second example is that of isoscopic, adaptive imaging (Fig. 2). Here, the same object f is imaged in both Modality 1 and 2. The resulting data or reconstruction from Modality 1 is often called a “scout” image. The configuration of the second modality is a function of this scout image. Thus, the system adapts based on initial measurements from Modality 1. Modality 1 and 2 may, in fact, correspond to different configurations of a single imaging system. An example of this type of system is an adaptive SPECT system developed by Freed *et al.* [7], [8]. As another example, many modern clinical X-ray systems adaptively adjust tube voltage and current based on an initial low-dose scan. Adaptive imaging systems may also be polyscopic. For example, a CT image could be used to guide a SPECT or PET acquisition. In this case, f in Fig. 2 would be replaced with f_1 for Modality 1 and f_2 for Modality 2.

A third example, shown in Fig. 3, is that of polyscopic, assisted data analysis. The task of interest in this example is that of parameter estimation. The data from both modalities are used to estimate the parameters. An example of this type of analysis is to use a CT reconstruction $\hat{f}_2$ to define a region of interest from which the tracer uptake can be estimated using the first modality. Typically, the estimation step is performed with an algorithm that may use either the raw data g_1 or a reconstruction $\hat{f}_1$. An example of an estimator that uses raw data is the Wiener estimator discussed below. The distinction between this example and the first example is that the reconstruction from Modality 2 is used to modify the algorithm that operates on the data from Modality 1.

A final example, shown in Fig. 4, is that of polyscopic, assisted reconstruction. Here, the reconstruction for Modality 2 is used to assist the reconstruction step for Modality 1. In modern SPECT imaging, a CT reconstruction is often used to create an attenuation map that will aid in the SPECT reconstruction algorithm. The observer may use both reconstructions or just the reconstruction from Modality 1 to make the final decision.

III. EXAMPLE USES OF MULTIMODALITY AND ADAPTIVE IMAGING

In this section, we present some uses of multimodality and adaptive imaging to illustrate the importance of the task-based assessment approach discussed in the following section.

Example 1

A CT system can be used to define the boundary of a lesion. This boundary map can then be used with a registered PET or SPECT reconstruction to estimate the specific activity within the lesion. This type of analysis can be used to determine the degree of up-regulated glucose metabolism and the likelihood of malignancy. This example fits into the polyscopic, assisted-analysis paradigm of Fig. 3.

Example 2

A CT scan can be used to provide an attenuation map that improves a SPECT reconstruction of the heart, yielding a more reliable estimate of the cardiac ejection fraction, the fraction of blood pumped out of the left ventricle at the end of each cycle. This example is a case of polyscopic, assisted reconstruction as shown in Fig. 4.

Example 3

A CT or MRI scan could provide a prior probability distribution on tumor parameters in a parametric tumor model. This prior can then be used for a PET or SPECT scan to perform Bayesian estimation of these tumor parameters, such as tumor volume. This is an example of polyscopic assisted analysis.

Example 4

For functional SPECT or PET studies, we are often interested in pharmacokinetic parameters such as uptake and washout rates. This task is made easier if the compartments in a compartmental model are spatially localized. For example, a tumor compartment would be localized in the tumor. An MRI or CT scan could be used to provide the anatomical data to localize the compartments. This is an example of polyscopic assisted analysis.

Example 5

For the detection of suspected metastases, a CT scan could show which organ a lesion is in. Then, a functional imaging modality such as PET or SPECT could be used to determine whether this lesion is a malignancy or a benign growth. Prior knowledge of the organ allows the classification task to be optimized for the statistics of that tissue type. This optimization could be at the system level by, for example, choosing an optimal tracer for the specific organ—an example of polyscopic, adaptive imaging. The optimization could also be at the analysis stage by adapting the observer to the statistics of the data from that tissue type—an example of polyscopic assisted analysis.

Example 6

A similar strategy may be used where both modalities are MRI but the pulse sequences differ. A standard pulse sequence could be used to provide anatomical data to locate suspected lesions. Then, a pulse sequence specialized for T2 estimation could be used to classify the lesion as benign or malignant. The estimation of T2 from this type of pulse sequence is improved when the location of the lesion is known. This is an example of isoscopic assisted analysis, or isoscopic assisted reconstruction when the object is composed of the triple $(\rho(\mathbf{r}), T_1(\mathbf{r}), T_2(\mathbf{r}))$.

Example 7

To detect and localize a tumor, it may be advantageous to use a functional modality such as PET or SPECT for the detection and the anatomical modality for the localization if a tumor is detected. This is an example that fits best within the traditional polyscopic paradigm.

Example 8

Another example of traditional polyscopic imaging is MRI diffusion tensor imaging. In this case, the low-resolution diffusion tensor map can be superimposed on a higher resolution anatomical MRI. This type of imaging can be used to estimate the extent of the neural damage due to a stroke.

IV. FUNDAMENTALS OF TASK-BASED ASSESSMENT

In this section, we will briefly review the fundamental aspects of task-based assessment methodology. For this review, the vector $\mathbf{g}$ will represent the data and/or reconstructed images used to make the clinical decision. The application of task-based assessment of image quality to multimodality and adaptive imaging will be detailed in the next section. The four basic components needed to implement task-based assessment are the task, object ensemble, observer, and figure of merit.

A. Tasks

In medical imaging, images are acquired so that specific tasks can be performed. The tasks generally fall into one of three categories: classification, estimation, or combination. A classification task is one in which the image data are used to classify the patient into one of L classes. For example, in tumor detection, there are two classes: tumor present and tumor absent. Staging a tumor is a classification task with more than two classes. For an estimation task, the image data are used to estimate one or more parameters of clinical interest. For example, SPECT images of the heart are often taken to estimate the left ventricle cardiac ejection fraction. In tumor imaging, position estimation is an example of a multiparameter estimation task. Combined tasks involve both classification and estimation. An example of a combined task is the detection and localization of a small tumor.

B. Object Ensembles

Observer performance is limited by the sources of variability in the imaging chain. These sources of variability include detector noise and object variability. Detector noise is usually included as part of the imaging-system model. Object variability is the variation in the patients being imaged and is often the factor that most limits observer performance. Thus, to properly measure task-based image quality, one must accurately account for object variability by using the data from a large number of objects drawn from a well-specified ensemble. One method for generating a large dataset is through simulation. In this case, the simulated objects are drawn from a stochastic ensemble of mathematical objects that models the patient ensemble. For example, the nonuniform rational B-spline-based cardiac-torso (NCAT) phantom has been used extensively to model realistic and variable patient anatomy [9].

Another approach to generating an object ensemble is to use real data from a large number of patients. This has the advantage of directly sampling the ensemble of interest. The principal disadvantage is the time and expense needed to collect these data. One may also combine real and simulated data. For example, images collected from an ensemble of patients may have simulated tumors added to the data.

C. Figures of Merit

A figure of merit is a number that quantifies the average performance of an observer on the given task. Task-based assessment of image quality depends on our ability to compute a meaningful figure of merit for a given object ensemble, imaging system, task, and observer.

Any observer for a signal-detection task can be thought of as computing a test statistic $t(\mathbf{g})$ and comparing this number to a threshold. If the test statistic is greater than the threshold, then the signal is declared to be present. Otherwise, the signal is said to be absent. For a given threshold, the true-positive fraction (TPF) is the probability of deciding that the signal is present when it is actually present. The false-positive fraction (FPF) is the probability of deciding that a signal is present when it is actually absent. These two numbers also determine the true-negative fraction (TNF) and the false-negative fraction (FNF). If the threshold is varied and the TPF is plotted as a function of FPF, then the result is a receiver operating characteristic (ROC) curve. One useful measure of observer performance is the area under the ROC curve (AUC). A typical ROC curve is shown in Fig. 5. An observer with an AUC of one is always correct, while an observer with an AUC of 0.5 is guessing. One reason the AUC is a useful performance measure is that it can be measured with a two-alternative forced-choice (2AFC) test. Here, the observer is shown many pairs of signal-present and signal-absent images. He or she must determine which image in each pair contains the signal. The fraction of correctly identified signal-present images is an unbiased estimate of the AUC.

An alternative to the AUC is the signal-to-noise ratio (SNR) of the test statistic given by

$$\text{SNR}^2 = \frac{(\bar{t}_1 - \bar{t}_0)^2}{\frac{1}{2}[\sigma_1^2 + \sigma_0^2]}. \quad (7)$$

Here, $\bar{t}_1$ is the mean of the test statistic over the signal-present ensemble and $\bar{t}_0$ is the mean of the test statistic over the signal-absent ensemble. Similarly, σ_1^2 and σ_0^2 are the variances of the test statistic over the signal-present and signal-absent ensembles, respectively. If the distributions of the signal-present and signal-absent test statistics are normal, then SNR^2 is monotonically related to AUC.

If a meaningful cost can be assigned to the four possible outcomes (i.e., true positive, false positive, true negative, false negative), then the Bayes risk may be used as a figure of merit [10]. The Bayes risk is the cost of each possible decision averaged over system noise and object variability. Unlike AUC and SNR, the Bayes risk depends on the decision threshold. For any given observer, there is usually an optimum threshold that minimizes the Bayes risk. In medical imaging, it is considered difficult to assign costs to decision outcomes that directly affect patient health.

For estimation tasks, the observer produces an estimate $\hat{\theta}(\mathbf{g})$ of an unknown set of parameters θ . As with signal-detection tasks, one approach is to define a cost function $C(\hat{\theta}(\mathbf{g}), \theta)$ that measures the “cost” of using the estimate $\hat{\theta}(\mathbf{g})$ when the true value is θ . When the cost function is averaged over both image-data $\mathbf{g}$ and parameters θ , we arrive at the corresponding Bayes risk. If θ is a vector parameter and

$$C(\hat{\theta}(\mathbf{g}), \theta) = |\hat{\theta}(\mathbf{g}) - \theta|^2 \quad (8)$$

then the Bayes risk is called the ensemble mean squared error (EMSE).

For combined tasks, ROC analysis and Bayes risk can be combined. First, we define a utility function that is essentially the negative of a cost function. Then, instead of plotting TPF versus FPF as with an ROC curve, we plot the average utility of our estimate for the true-positive cases versus FPF. The result is the estimation ROC curve (EROC) [11]. If the parameters $\hat{\theta}(\mathbf{g})$ to be estimated are signal location and the utility function is one within a specified neighborhood of the true location and zero outside that neighborhood, then this analysis reduces to the standard localization ROC curve (LROC) [12]. The area under the EROC curve (AEROC) can be used as a figure of merit for combined detection and estimation tasks. As with ROC analysis, the AEROC can be estimated from the 2AFC test where the observer must decide which image has the signal and then estimate the parameters for that signal. The average utility for the correctly detected signals is an estimate of the observer’s AEROC.

D. Observers

For reconstructed images, the observers performing the given task are typically human. Human observer performance can be measured directly through psychophysical studies where the observers are shown the images under controlled conditions and asked to record their decisions. These types of studies are often impractical due to time and financial constraints. An alternative observer to use for reconstructed images is a mathematical model of a human observer [13]. These observers have been developed for relatively simple tasks and can accurately predict average performance of humans on these tasks.

For imaging system comparison and optimization, it may be better to use an observer that is mathematically ideal to quantify the information content produced by the system. We have previously demonstrated the use of these observers in system design and optimization for a

single-modality nonadaptive imaging system [14], [15]. Ideal observers and the application of these observers to multimodality and adaptive imaging will be the main topic in the following sections. For this reason, we will present the general mathematical formalism here.

For detection tasks, the data vector $\mathbf{g}$ is drawn from one of two ensembles described by probability density functions $pr(\mathbf{g}|1)$ (signal present) and $pr(\mathbf{g}|0)$ (signal absent). If these are known, they can be used to make a test statistic that maximizes the AUC and minimizes Bayes risk. This test statistic is called the likelihood ratio and is given by

$$\Lambda(\mathbf{g}) = \frac{pr(\mathbf{g}|1)}{pr(\mathbf{g}|0)}. \quad (9)$$

The observer that uses this (or an equivalent) test statistic is known as the Bayesian ideal observer. Unfortunately, the likelihood ratio is usually difficult to compute except for simplified tasks and object ensembles.

An alternative to the Bayesian ideal observer is the ideal linear (or Hotelling) observer. The Hotelling observer computes a test statistic using only linear manipulations of the data $\mathbf{g}$ and is the observer that maximizes the SNR of the test statistic among all linear observers. If we define the signal vector $\mathbf{s}$ as the difference in the mean data vector under the two hypotheses and the covariance $\mathbf{K}$ as the average covariance matrix of the data under the two hypotheses, the Hotelling test statistic is given by

$$t_H(\mathbf{g}) = \mathbf{s}^\dagger \mathbf{K}^{-1} \mathbf{g}. \quad (10)$$

Thus, the Hotelling observer only requires knowledge of the first- and second-order statistics of the data as opposed to the complete probability density functions.

For large-dimensional data vectors, even the Hotelling observer test statistic may be difficult to compute. In these cases, an alternative linear observer is the channelized Hotelling observer. The channelized Hotelling observer reduces the dimension of the data vector $\mathbf{g}$ using a channel matrix $\mathbf{T}$. The means and covariances of the channelized data vector are used to compute the test statistic using an equation similar to (10). If the channels $\mathbf{T}$ are chosen so that the SNR of the channelized Hotelling observer is close to the SNR of the Hotelling observer, then we call these channels efficient [16]. For certain tasks, it has been shown that Gabor-like channels can be used to predict average human performance. These types of channels are called anthropomorphic channels.

For estimation tasks, the data vector $\mathbf{g}$ is drawn from an ensemble described by the conditional probability density function $pr(\mathbf{g}|\boldsymbol{\theta})$ for some value of $\boldsymbol{\theta}$. The parameters to be estimated are also considered random with probability density function $pr(\boldsymbol{\theta})$ —often called the prior distribution. For certain convex cost functions [10], there is an estimator that minimizes the Bayes risk, the posterior-mean estimator, given by

$$\widehat{\boldsymbol{\theta}}_{PM}(\mathbf{g}) = \int \boldsymbol{\theta} pr(\boldsymbol{\theta}|\mathbf{g}) d\boldsymbol{\theta}. \quad (11)$$

The posterior distribution $pr(\boldsymbol{\theta}|\mathbf{g})$ is given by

$$pr(\boldsymbol{\theta}|\mathbf{g}) = \frac{pr(\mathbf{g}|\boldsymbol{\theta})pr(\boldsymbol{\theta})}{pr(\mathbf{g})}. \quad (12)$$

As with the Bayesian ideal observer, the posterior mean is difficult to compute since it requires full knowledge of the probability density functions $pr(\mathbf{g}|\boldsymbol{\theta})$ and $pr(\boldsymbol{\theta})$.

If the cost function is zero in a small neighborhood around the true value of the parameter and has a large constant value outside of this neighborhood, then the Bayes risk is minimized by the maximum *a posteriori* (MAP) estimator given by

$$\hat{\theta}_{\text{MAP}}(g) = \arg \max_{\theta} \{pr(g|\theta)pr(\theta)\}. \quad (13)$$

If the parameter has a finite range and the prior is assumed to be flat, then we arrive at the maximum-likelihood (ML) estimator, given by

$$\hat{\theta}_{\text{ML}}(g) = \arg \max_{\theta} \{pr(g|\theta)\}. \quad (14)$$

One reason for using the ML estimator is that it is known to be asymptotically efficient [10].

All of the above estimators assume knowledge of the conditional probability function $pr(g|\theta)$. If this density is difficult to compute or not known, an alternative is the optimal linear (or Wiener) estimator. The Wiener estimator is the linear estimator that minimizes the EMSE and is given by

$$\hat{\theta}_w(g) = \bar{\theta} + K_{\theta,g} K^{-1} (g - \bar{g}). \quad (15)$$

In this equation, $\bar{\theta}$ is the mean value of the prior distribution and $\bar{g}$ is the mean value of $pr(g)$. The matrix $K_{\theta,g}$ is the cross-covariance between θ and g and the matrix K is the covariance matrix for the distribution $pr(g)$. As with Hotelling observers, a channelized Wiener estimator can be used for certain types of parameters [17].

For combined detection and estimation tasks, there is an ideal observer that maximizes the AEROC [11]. For the detection part of the task, this observer uses a test statistic that is a weighted average of likelihood ratios. For the estimation part of the task, the observer uses an estimator similar to the posterior mean estimator. If Gaussian assumptions about the image-data statistics are valid, then this ideal AEROC observer is equivalent to a scanning Hotelling observer [11].

V. OBSERVERS FOR ADAPTIVE AND MULTIMODALITY IMAGING

To describe the ideal observer for multimodality and adaptive imaging, we will first introduce a simplified notation. For estimation tasks, $pr(g|\theta)$ and $pr(\theta)$ will be conditional and prior distributions as described above. We will also use $pr(g|\theta)$ for detection tasks, in which case θ only takes the values zero and one and the prior distribution $pr(\theta)$ consists of two numbers P_0 and P_1 , the prior probabilities for the two classes. With this in mind, we may write the conditional probability for a traditional polyscopic, dual-modality imaging system (see Fig. 1) in the following form:

$$pr(g_1, g_2|\theta) = \iint pr(g_1|f_1)pr(g_2|f_2)pr(f_1, f_2|\theta)df_1df_2. \quad (16)$$

The double integral in this equation, and similar integrals in subsequent equations, represents an expectation over the object ensemble (for a given value of θ) of the conditional probabilities. This equation assumes that the system noise in Modality 1 described by $pr(g_1|f_1)$ is independent of the system noise in Modality 2 described by $pr(g_2|f_2)$. The objects f_1 and f_2 being imaged by the two modalities are not independent since they both arise from the same patient. For a traditional isoscopic imaging system (e.g., dual-energy SPECT), the conditional probability density function is given by

$$pr(g_1, g_2 | \theta) = \int pr(g_1 | f) pr(g_2 | f) pr(f | \theta) df. \quad (17)$$

Here again, we are assuming that the detector noise in g_1 is independent from that of g_2 .

For adaptive imaging, the acquisition of data vector g_1 is affected by the initial data g_2 . This implies that (16) for the polyscopic case becomes

$$pr(g_1, g_2 | \theta) = \iint pr(g_1 | g_2, f_1) pr(g_2 | f_2) pr(f_1, f_2 | \theta) df_1 df_2. \quad (18)$$

Similarly, for the isoscopic adaptive case, (17) becomes

$$pr(g_1, g_2 | \theta) = \int pr(g_1 | g_2, f) pr(g_2 | f) pr(f | \theta) df. \quad (19)$$

Again, we assume that the system noise in g_1 is independent from the system noise in g_2 . The conditional probability density functions for assisted reconstruction and assisted-analysis systems (Figs. 3 and 4) are the same as those for traditional systems.

By setting

$$g = \begin{bmatrix} g_1 \\ g_2 \end{bmatrix} \quad (20)$$

the discussion in the previous section may be applied to multimodality and adaptive imaging systems. Methods for computing likelihood ratios and posterior-mean estimators using Markov chain Monte Carlo techniques have been developed for single-modality imaging systems [18]. Using (16)–(19), these methods can be extended for use with multimodality or adaptive systems.

For linear observers, the covariance matrix has the form

$$K = \begin{bmatrix} K_{11} & K_{12} \\ K_{12}^\dagger & K_{22} \end{bmatrix} \quad (21)$$

where

$$K_{ij} = \langle (g_i - \bar{g}_i)(g_j - \bar{g}_j)^\dagger \rangle \quad (22)$$

and

$$\bar{g}_i = \langle g_i \rangle \quad (23)$$

All expectations are performed using the appropriate conditional probability $pr(g_1, g_2 | \theta)$ [see (16)–(19)] and the prior density $pr(\theta)$. Similarly, the cross-covariance used in the Wiener estimator has the form

$$K_{\theta, g} = [K_{\theta, g_1} \quad K_{\theta, g_2}]. \quad (24)$$

With these expressions in mind, the formulas for the Hotelling test statistic and Wiener estimator (10) and (15) are unchanged for multimodality systems.

Using the appropriate equation above for the conditional probability density, each of the covariance and cross-covariance matrices can be decomposed into a sum of matrices. Each

matrix in the sum corresponds to the average contribution to the covariance from a particular noise source. These noise sources are system noise, object variability, and parameter variability. This type of analysis is discussed extensively in a previous publication [7].

There is another observer option that presents itself when we have two modalities. In terms of our block diagrams, this option fits into the adaptive-imaging paradigm (Fig. 2). The idea is to use a linear observer with the data from Modality 1 where the vectors and matrices that define the linear observer are allowed to depend on the data from Modality 2. These observers are called adaptive Hotelling observers for detection tasks and adaptive Wiener observers for estimation tasks.

The adaptive Hotelling observer test statistic is given by

$$t_{AH}(g_1, g_2) = w(g_2)^{\dagger} g_1 + c(g_2). \quad (25)$$

The adaptive template $w(g_2)$ and the shift $c(g_2)$ are determined using the posterior distribution $pr(f|g_2)$ and the conditional distribution $pr(g_1|f)$ [7]. The adaptive Wiener estimator has a similar form

$$\hat{\theta}(g_1, g_2) = W(g_2)g_1 + c(g_2). \quad (26)$$

Again, the posterior distribution, the conditional distribution, and the prior are used to determine the matrix $W(g_2)$ and the vector constant $c(g_2)$ [7]. These adaptive observers may also be used for assisted analysis systems (Fig. 3) and assisted-reconstruction systems (Fig. 4). One way to view these adaptive observers is that the “scout data” g_2 are being used to reduce the object ensemble from the full ensemble to a subset consistent with g_2 . This reduced ensemble is then used when computing means, covariances, and cross-covariances for the linear observers [7].

VI. STRATEGIES FOR ASSESSMENT AND OPTIMIZATION

In this section, we will outline the general procedure for implementing task-based assessment of image quality on multimodality or adaptive imaging systems. The first stage is implementing a protocol for performance assessment based on tasks, observers, and object ensembles. The second stage is using the resulting figure of merit for system optimization. Lastly, we will discuss some computational issues with carrying out this program.

A. Performance Assessment

We will discuss performance assessment as it relates to both existing systems and simulated imaging systems. For existing systems, real data are available and should be used to assess image quality. For detection tasks, computing a figure of merit requires the collection of a large dataset containing both signal-absent and signal-present cases. It is often difficult to acquire enough signal-present cases. In this case, the signal-present dataset may be synthesized by incorporating simulated lesions (or other abnormalities) into normal images [19]. For estimation tasks, a large dataset of abnormal images is required.

It is unlikely that any Bayesian observer could be used to assess image quality for existing systems since the conditional probability distribution is unknown. There are methods to estimate probability density functions from sample image data vectors, but these methods would require a prohibitively large number of samples since the dimension of the data vectors is very large. For this reason, a more practical alternative is a linear observer, i.e., the Hotelling observer for detection tasks or the Wiener estimator for estimation tasks. Again, however, the large dimensionality of the data vector may also make these observers impractical since a large

covariance matrix must be estimated and inverted. If this is the case, then channelized linear observers become the option of choice. Fortunately, efficient channels [16] have been developed so that figures of merit from channelized observers are meaningful measures of image quality.

Given a large number of sample images and a linear observer, we may estimate a figure of merit. For detection tasks, this figure of merit may be SNR or AUC. These two measures can be expected to be monotonically related for linear observers due to the central limit theorem. The Bayes risk is also an option for performance assessment if the cost matrix is known for the particular detection task. For estimation tasks, the EMSE can be estimated from samples as can the Bayes risk if a relevant and meaningful cost function is available.

Simulation is an important tool used in imaging-system design. It is necessary to be able to simulate the imaging system so that the design of the system can be optimized. It is easier to assess observer performance on simulated imaging systems and adjust imaging-system parameters accordingly than to wait until the system is built, in which case it may be too late to make design adjustments. To perform task-based assessment for simulated systems, a simulated object ensemble is needed. These ensembles vary from very simple object models such as Gaussian lesions on stationary backgrounds [20] to more realistic object models based on four-dimensional anatomical phantoms that can be randomly perturbed and statistical characterizations of tissue texture [21].

For the simpler object models, it is feasible to compute Bayesian ideal observers such as the likelihood ratio (9), the posterior-mean estimator (11), or the ideal AEROC observer [11]. As the object model becomes more complex, the time needed to compute the figure of merit increases and may become prohibitive. Thus, linear observers may be necessary for more complex object models. Since the objects are being simulated, we may be able to generate enough samples to accurately estimate the full covariance matrix needed for linear observers. This reduces the computational difficulty to inverting a large matrix. If this inversion is not feasible, then we may, again, fall back on channels to ease this burden. The figures of merit and their computation in this case are the same as for real systems.

B. System Optimization

It should be clear to the reader that computing observer performance may be computationally intensive. Thus, optimizing the design of a medical imaging system for task-based figures of merit is possible only if relatively few parameters are being adjusted. Usually, derivatives of the figure of merit are not available or are even more difficult to compute than the figure of merit itself. For this reason, derivative-free search algorithms are probably the only feasible optimization methods. Many deterministic search algorithms exist, including the Nelder–Mead algorithm [22] and other such pattern-search algorithms. Stochastic search algorithms, such as simulated annealing [23] or genetic algorithms [24], can also be employed. To the authors' knowledge, task-based optimization has not been carried out for any adaptive or multimodality imaging system. The first advances in this area will likely employ adaptive Hotelling observers and pattern-search optimization algorithms to optimize SPECT/CT or PET/CT systems.

Instead of optimizing a given system, we are often more interested in comparing a multimodality imaging system to a single-modality imaging system for a given task. This is easier than optimization, as it requires only two estimates of the figure of merit. Task-based assessment can be useful here, not only to determine if the multimodality imaging system is better but also to quantify the increase in task performance when modalities are combined. Finally, adaptation rules for adaptive imaging systems may also be evaluated using task-based measures of image quality [7]. This is also more easily accomplished than system optimization since the adaptation rules are often heuristic and we only need to measure their effectiveness

in terms of task performance. We are currently using the adaptive Hotelling observer to compare an adaptive SPECT system, which uses simple adaptation rules to a fixed-geometry SPECT system.

C. Computational Issues

Computing figures of merit for Bayesian observers is a two-stage process. The first stage is to compute the test statistic or estimate for a given dataset. The second stage is to repeat this computation many times to estimate the final figure of merit. Computing the test statistic or the estimate requires averaging over object ensembles generated from the posterior distribution of the objects given the image data. Markov chain Monte Carlo techniques have been developed to compute the necessary expectations [18]. Markov chain Monte Carlo techniques do, however, require many iterations to converge to a solution. For the relatively simple case of stationary backgrounds, the number of iterations is on the order of 10^5 to 10^6 . To compute a figure of merit, the Markov chain must be run for each sample image. There are two sources of error in the estimate of the figure of merit. The first source is due to the random nature of the end result of the Markov chain. The second source is due to the finite number of images used to estimate the figure of merit. These two sources of error are often called reader variability and case variability. The variance in the estimate of the figure of merit can be analyzed as a function of the number of readers and number of cases using multiple-reader multiple-case analysis [25]–[28].

For linear observers, the principal computational issues are computing covariance matrices and inverting them. One method for computing the covariance matrix involves estimating directly from the samples. To obtain an invertible matrix from this procedure, the number of samples must be at least the dimension of the data vector. Since data vectors for imaging systems tend to have dimensions on the order of 10^5 to 10^6 , this is a very large number of samples indeed. When we take into account the fact that, to obtain an accurate estimate of the covariance matrix, we probably need ten times that amount, computing a covariance matrix directly from samples may be infeasible. There is a way around this problem that takes advantage of an algebraic relation known as the matrix inversion lemma [10]. This lemma takes advantage of the fact that the covariance matrix of an imaging system may be decomposed as

$$K = \bar{K}_n + K_b \quad (27)$$

where $\bar{K}_n$ relates to the detector noise and is easy to estimate and is invertible. The term K_b relates to the object variability and can be estimated (in simulation) from noise-free samples. Because $\bar{K}_n$ is invertible, the sum of $\bar{K}_n$ with K_b is also invertible even if K_b is not full rank because too few samples are used. If the number of samples used to estimate K_b is L , then the matrix inversion lemma can be used to reduce the matrix inversion of K to the computation of an inverse of an $L \times L$ matrix [10]. Care must be used with this technique if the background variability is the dominant source of noise.

Finally, using channels poses no computational difficulties at all since the dimension of the vectors and matrices is generally small. Sample channel vectors can readily be used to estimate the channel covariance matrix used in the computation of the figure of merit. The main difficulty with channelized observers is determining whether the channels are efficient, e.g., whether the channelized SNR is close to the Hotelling SNR.

VII. SUMMARY AND CONCLUSIONS

Multimodality and adaptive imaging systems fall into general classes and may be distinguished by the object functions being imaged and the interaction between the two modalities. For each class, there is a general statistical model that can be used to generate Bayesian ideal observers,

ideal linear observers, or ideal adaptive linear observers for use in task-based assessment. These observers may be used to compute figures of merit to optimize multimodality systems, compare competing multimodality systems, or quantify the advantage of multimodality systems over single-modality systems. Computing these figures of merit may be computationally intensive, but there are strategies for addressing these concerns. In conclusion, the task-based assessment of multimodality and adaptive imaging systems is necessary and feasible to objectively determine their performance for clinically relevant tasks.

References

1. Park JM, Gambhir SS. Multimodality radionuclide, fluorescence, and bioluminescence small-animal imaging. *Proc IEEE* 2005;93(4):771–783.
2. Pietrzyk U, Herholz K, Schuster A, von Stockhausen HM, Lucht H, Heiss WD. Clinical applications of registration and fusion of multimodality brain images from PET, SPECT, CT, and MRI. *Eur J Radiol* 1996;21(3):174–182. [PubMed: 8777907]
3. Saoudi A, Lecomte R. A novel APD-based detector module for multi-modality PET/SPECT/CT scanners. *IEEE Trans Nucl Sci* 1999;46(3):479–484.
4. Barrett HH. Objective assessment of image quality: Effects of quantum noise and object variability. *J Opt Soc Amer A* 1990;7(7):1266–1278. [PubMed: 2370589]
5. Barrett HH, Denny JL, Wagner RF, Myers KJ. Objective assessment of image quality. II: Fisher information, Fourier crosstalk, and figures of merit for task performance. *J Opt Soc Amer A* 1995;12:834–852.
6. Barrett HH, Abbey CK, Clarkson E. Objective assessment of image quality—III: ROC metrics, ideal observers, and likelihood-generating functions. *J Opt Soc Amer A* 1998;15:1520–1535.
7. Barrett HH, Furenlid LR, Freed M, Hesterman JY, Kupinski MA, Clarkson E, Whitaker MK. Adaptive SPECT. *IEEE Trans Med Imag.* submitted for publication
8. Freed M, Kupinski MA, Furenlid LR, Barrett HH. A prototype instrument for adaptive SPECT imaging. *Med Phys.* submitted for publication
9. Segars WP, Tsui BMW. Study of the efficacy of respiratory gating in myocardial SPECT using the new 4D NCAT phantom. *IEEE Trans Nucl Sci* 2002;49(3):675–679.
10. Barrett, HH.; Myers, KJ. *Foundations of Image Science*. New York: Wiley; 2004.
11. Clarkson E. The estimation receiver operating characteristic curve and ideal observers for combined detection/estimation tasks. *J Opt Soc Amer A*.
12. Khurd, P.; Gindi, G. *Proc Med Imag 2005: Image Perception, Observer Perform, Technol Assess.* 5749. SPIE; 2005. Decision strategies maximizing the area under the LROC curve; p. 150-161.
13. Abbey CK, Barrett HH. Human- and model-observer performance in ramp-spectrum noise: Effects of regularization and object variability. *J Opt Soc Amer A* 2001;18:473–488.
14. Hesterman JY, Kupinski MA, Clarkson E, Barrett HH. Task-based hardware evaluation of small-animal SPECT systems. *Med Phys.*
15. Chen, L.; Barrett, HH. Optimizing lens-coupled digital radiographic imaging systems based on model observers performance. In: Chakraborty, DP.; Krupinski, EA., editors. *Proc Med Imag 2003: Image Perception, Observer Perform, Technol Assess.* SPIE; 2003. p. 63-70.
16. Barrett, HH.; Abbey, C.; Gallas, B.; Eckstein, M. Stabilized estimates of Hotelling-observer detection performance in patient-structured noise. In: Kundel, HL., editor. *Proc Med Imag 1998: Image Perception.* 3340. SPIE; 1998. p. 27-43.
17. Hesterman JY, Kupinski MA, Clarkson E, Barrett HH. Assessment of M3R apertures in tasks of signal detection and activity estimation using reconstructed images. *IEEE Trans Med Imag.* submitted for publication
18. Kupinski MA, Hoppin JW, Clarkson E, Barrett HH. Ideal-observer computation in medical imaging with use of Markov-chain Monte Carlo techniques. *J Opt Soc Amer A* 2003;20(3):430–438.
19. Burgess AE, Jacobson FL, Judy P. Mass discrimination in mammography: Experiments using hybrid images. *Acad Radiol* 2005;10:1247–1256. [PubMed: 14626299]

20. Rolland JP, Barrett HH. Effect of random background inhomogeneity on observer detection performance. *J Opt Soc Amer A* 1992;9:649–658. [PubMed: 1588452]
21. Segars WP, Tsui BMW, Frey EC, Johnson GA, Berr SS. Development of a 4D digital mouse phantom for molecular imaging research. *Mol Imag Biol* 2004;6(3):149–159.
22. Avriel, M. *Nonlinear Programming: Analysis and Methods*. New York: Dover; 2003.
23. Kirkpatrick S, Gelatt CD, Vecchi MP. Optimizing by simulated annealing. *Science* 1983;220:671–680. [PubMed: 17813860]
24. Goldberg, DE. *Genetic Algorithms in Search, Optimization and Machine Learning*. Boston, MA: Kluwer Academic; 1989.
25. Barrett, HH.; Kupinski, MA.; Clarkson, E. *Proc Med Imag 2005: Image Perception, Observer Perform, Technol Assess*. 2005. Probabilistic foundations of the MRMC method; p. 21-31.
26. Clarkson E, Kupinski MA, Barrett HH. A probabilistic model for the MRMC method, Part 1: Theoretical development. *Acad Radiol* 2006;13(11):1410–1421. [PubMed: 17070460]
27. Kupinski MA, Clarkson E, Barrett HH. A probabilistic model for the MRMC method, Part 2: Validation and applications. *Acad Radiol* 2006;13(11):1422–1430. [PubMed: 17070461]
28. Gallas BD. One-shot estimate of MRMC variance: AUC. *Acad Radiol* 2006;13:353–362. [PubMed: 16488848]

Biographies

Eric Clarkson received the bachelor's degree in mathematics, physics, and philosophy from Rice University, Houston, TX, in 1976, the master's degree in physics and the Ph.D. degree in mathematics from Arizona State University, Tempe, in 1980 and 1985, respectively, and the master's degree in optical sciences from the University of Arizona, Tucson, in 1996.

After teaching at Murray State University, Murray, KY, for six years, he joined the Faculty of the University of Arizona in 1996, where he is a Research Professor of optical sciences and radiology. His research interests include theoretical image science, image-quality assessment, and image reconstruction techniques.

Matthew A. Kupinski received the Ph.D. degree from the University of Chicago, Chicago, IL, in 2000.

He joined the Faculty of the University of Arizona, Tucson, in 2002, where he is an Assistant Professor of optical sciences and radiology. He has published numerous papers and book chapters on image quality and image science. His research interests include observer models, ideal-observer computations, and imaging hardware optimization.

Harrison H. Barrett (Fellow, IEEE) received the B.S. degree from Virginia Polytechnic Institute, Blacksburg, the M.S. degree from the Massachusetts Institute of Technology, Cambridge, and the Ph.D. degree from Harvard University, Cambridge, MA.

He is currently a Regents Professor at the University of Arizona, Tucson, with appointments in the College of Optical Sciences, the Department of Radiology, and the programs in Applied Mathematics and Biomedical Engineering. He is Director of the Center for Gamma-Ray Imaging.

Prof. Barrett is coauthor (with K. J. Myers) of *Foundations of Image Science*, which in 2006 received the First Biennial J. W. Goodman Book Writing Award from OSA and SPIE.

Lars Furenlid (Member, IEEE) received the B.S. degree from the University of Arizona, Tucson, and the Ph.D. degree from the Georgia Institute of Technology, Atlanta.

He is currently a Professor at the University of Arizona and Associate Director of the Center for Gamma-Ray Imaging, with appointments in the Department of Radiology and the College of Optical Sciences. He was a Staff Scientist with the National Synchrotron Light Source, Brookhaven National Laboratory. His major research area is the development and application of detectors, electronics, and systems for biomedical imaging.

Fig. 1.

An illustration of traditional polyscopic, multimodality imaging. Here, f_1 might represent the distribution of a tracer in SPECT, while f_2 might represent the X-ray attenuation coefficient as measured by CT. To compute a figure of merit, many subjects must be imaged.

Fig. 2.

An illustration of isoscopic, adaptive imaging. Here, the object f being imaged in both Modality 1 and 2 is the same. The data from the first modality (either g_1 or $\hat{f}_1$) are used to determine the configuration (e.g., $\mathcal{H}_2(g_1)$) of the second modality. The observer may use data from both modalities or just the second modality to make a decision.

Fig. 3.

An illustration of polyscopic, assisted data analysis. Here, the purpose for imaging the subject is to estimate parameters of interest. The reconstruction in Modality 1 is optional since raw data can often be used to estimate the parameters. The reconstruction from Modality 2 is used to assist in the data analysis.

Fig. 4.

An illustration of polyscopic, assisted reconstruction. Here, the reconstruction from Modality 2 is used to assist the reconstruction step for Modality 1. The observer may use both reconstructions or simply the reconstruction from Modality 1.

Fig. 5.
An example ROC curve used to measure observer performance.