

Analysis and Classification of Optical Tomographic Images of Rheumatoid Fingers with ANOVA and Discriminate Analysis

Ludguier D. Montejo^{*a}, Hyun K. Kim^a, Yrjö Häme^a, Jingfei Jia^a, Julio D. Montejo^b,
Uwe J. Netz^c, Sabine Blaschke^d, Paul Zwaka^d, Gerhard A. Müller^d, Jürgen Beuthan^e,
Andreas H. Hielscher^{*a,f,g}

^aDepartment of Biomedical Engineering, Columbia University, New York, NY 10027

^bDepartment of Mathematics, Harvard University, Cambridge, MA 02138

^cLaser-und Medizin-Technologie GmbH, Berlin-Dahlem, Germany

^dNephrology and Rheumatology, University Medicine of Göttingen, Göttingen, Germany

^eDept. of Medical Physics and Laser-Medicine, Charité – University Medicine Berlin, Germany

^fDepartment of Radiology, Columbia University Medical Center, New York, NY 10032

^gDepartment of Electrical Engineering, Columbia University, New York, NY 10027

ABSTRACT

We present a study on the effectiveness of computer-aided diagnosis (CAD) of rheumatoid arthritis (RA) from frequency-domain diffuse optical tomographic (FDOT) images. FDOT is used to obtain the distribution of tissue optical properties. Subsequently, the non-parametric Kruskal-Wallis ANOVA test is employed to verify statistically significant differences between the optical parameters of patients affected by RA and healthy volunteers. Furthermore, quadratic discriminate analysis (QDA) of the absorption (μ_a) and scattering (μ'_s) distributions is used to classify subjects as affected or not affected by RA.

We evaluate the classification efficiency by determining the *sensitivity* (Se), *specificity* (Sp), and the *Youden* index (Y). We find that combining features extracted from μ_a and μ'_s images allows for more accurate classification than when μ_a or μ'_s features are considered individually on their own. Combining μ_a and μ'_s features yields values of up to $Y = 0.75$ ($Se = 0.84$ and $Sp = 0.91$). The best results when μ_a or μ'_s features are considered individually are $Y = 0.65$ ($Se = 0.85$ and $Sp = 0.80$) and $Y = 0.70$ ($Se = 0.80$ and $Sp = 0.90$), respectively.

Keywords: Diffuse optical tomography, computer aided diagnosis, rheumatoid arthritis, discriminate analysis, ANOVA

1. INTRODUCTION

Over the last two decades, work in the field of diffuse optical tomography (DOT) has progressed from purely theoretical studies and bench-top experiments to first clinical trials that explore the utility in breast cancer diagnosis,¹ brain imaging,² and arthritis detection.³ While substantial advances have been made in building clinically useful instruments, and developing image reconstruction algorithms, much less effort has been dedicated to developing image analysis tools. Other medical imaging fields such as magnet resonance imaging (MRI), computer tomographic imaging (CT), and ultrasound imaging (US) frequently make use of advanced image analysis methods that enhance sensitivity and specificity in many cases.⁴

Our group has recently introduced application of CAD to DOT images from studies involving images of arthritic finger joints.⁵⁻⁹ However, these initial studies all focused on the analysis of absorption (μ_a) images only. In this work we expand on previous research by exploring the ability to diagnose RA by adding the analysis of scattering coefficient (μ'_s) distributions. We extract image features such as maximum and minimum μ_a and μ'_s as well as the variance across each volumetric parameter distribution.

Corresponding authors: ldm2106@columbia.edu (L.D. Montejo) and ahh2004@columbia.edu (A.H. Hielscher)

Advanced Biomedical and Clinical Diagnostic Systems IX, edited by Anita Mahadevan-Jansen, Tuan Vo-Dinh,
Warren S. Grundfest, Proc. of SPIE Vol. 7890, 78900L · © 2011 SPIE · CCC code: 1605-7422/11/\$18
doi: 10.1117/12.875977

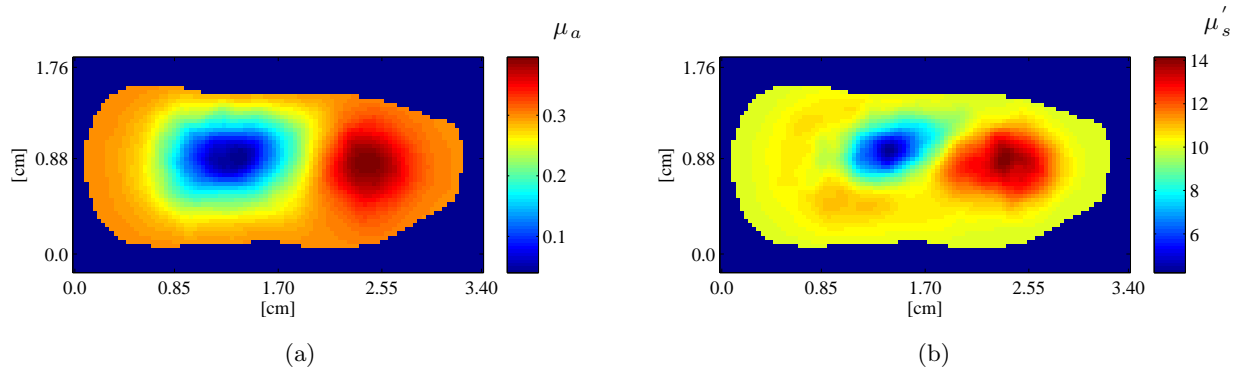


Figure 1. (a-b) Sample transverse cross-sections (perpendicular to plane of transillumination) from volumetric distributions of μ_a and μ'_s , respectively.

In this study, the non-parametric (or distribution-free) Kruskal-Wallis one-way analysis of variance (ANOVA) is used to test for statistically significant differences in extracted parameters between the two groups. The p-value is reported as a measure of confidence that observed differences in optical parameters between FDOT images of affected and healthy patients are statistically significant.¹³

These image features are subsequently used to classify the images as affected or healthy by rheumatoid arthritis (RA) with quadratic discriminate analysis (QDA). In QDA, absorption and scattering features are combinatorially combined, classified, and evaluated so as to determine what combinations yield the highest sensitivities and specificities.

2. METHODS

2.1 Clinical data and feature extraction

Data used in this work was obtained from a clinical trial in which 20 healthy volunteers and 33 patients with RA participated. Imaging of proximal interphalangeal (PIP) joints II-IV were performed on the right and left hand of each of the 20 healthy volunteers, yielding measurements of 120 healthy fingers. Imaging of PIP joints II-IV was performed on the dominant hand of patients with RA, yielding a total of 99 images of fingers with RA.

The instrument used for imaging is a frequency domain DOT sagittal laser scanner.⁵ The data used in this analysis was acquired with a source modulation frequency of 600 MHz. Optical parameters were reconstructed using a PDE-constrained optimization algorithm, where the radiative transfer equation (RTE) is used to model propagation of near infra-red light in tissue.¹⁰

Each FDOT reconstruction yields volumetric distributions of the μ_a and μ'_s coefficients within a given finger. A region of interest (ROI) is defined within the volume that includes all data points except points less than 2mm from the boundary. Then, three parameters are extracted from the ROI: $max(\mu)$, $min(\mu)$, and $var(\mu)$. An additional parameter, $ratio(\mu)$, is defined as $max(\mu)/min(\mu)$. The variable μ can be either μ_a or μ'_s . Sample cross sections from volumetric distributions of μ_a and μ'_s reconstructed from measurements on a healthy volunteer are presented in Fig. 1.

2.2 Kruskal-Wallis ANOVA

The (*non-parametric*) Kruskal-Wallis ANOVA is used to determine if observed differences are statistically significant. Unlike classical ANOVA, the Kruskal-Wallis test is used to determine confidence levels for observed differences between multiple groups through analysis of data rank instead of the actual observations.¹³

The ranking system is as follows: each observation (or feature) is ranked from smallest to largest, without regard to group, with the smallest observation having rank 1. Then, the general Kruskal-Wallis statistic, H , is defined as

$$H = \frac{12}{N(N+1)} \sum_{t=1}^T n_t (\bar{R}_t - \bar{R})^2, \quad (1)$$

where \bar{R} is the average rank of all observation, \bar{R}_t is the mean rank observed in group t , and n_t is the size of group t , N is the total number of observations, and T is the total number of groups. The critical value of the H -statistic to establish the presence of statistically significant differences is determined from a χ^2 distribution table with degrees of freedom $\nu = T - 1$.

2.3 Classification with discriminate analysis

Subsequent to feature extraction, we performed a retrospective classification of the images into two classes: joints affected by RA and healthy joints. The classification was performed with quadratic discriminant analysis (QDA). The QDA classifier was tested using the *leave-n-out* method, where 90% of the data is used to train the classifier while the remaining 10% of data is used for testing.

The Se and Sp of the method are computed by comparing QDA classification decisions on the testing data to the true clinical diagnosis of these patients.¹¹

$$Se = \frac{TP}{TP + FN} \quad (2)$$

$$Sp = \frac{TN}{TN + FP} \quad (3)$$

Where TP , TN , FP , and FN are true positives, true negatives, false positives, and false negative, respectively. Patients with RA are considered TP , while healthy patients are considered TN . The *Youden* index ($Y = Se + Sp - 1$) is also reported as a measure of classification efficiency. The training data is randomly selected from the entire data set; thus, to remove any possible bias each classification experiment is performed 100 times. The mean value over all iterations is reported.

In general, for classification with discriminant analysis, the posterior probability $p(\omega_i|\mathbf{x})$ of feature vector \mathbf{x} originating from class ω_i is defined by *Bayes theorem*,

$$p(\omega_i|\mathbf{x}) = \frac{p(\mathbf{x}|\omega_i)P(\omega_i)}{P(\mathbf{x})}, \quad (4)$$

where $P(\omega_i)$ and $P(\mathbf{x})$ represent the prior probabilities for class ω_i and feature vector \mathbf{x} , respectively.¹² The classification is done using the maximum a posteriori estimate,

$$\max_{\omega_i} \hat{p}(\omega_i|\mathbf{x}) \quad (5)$$

and setting the prior probability for each feature vector equal,

$$\hat{p}(\omega_i|\mathbf{x}) \propto \hat{p}(\mathbf{x}|\omega_i)P(\omega_i). \quad (6)$$

In this study, the prior probabilities for each class were defined to be equal, $P(\omega_i) = P(\omega_j) \forall i, j$, so that the classification depended only on the likelihood function estimate, $\hat{p}(\mathbf{x}|\omega_i)$. Furthermore, the likelihood functions are assumed to follow the general multivariate normal distribution,

$$p(\mathbf{x}|\omega_i) = \frac{1}{(2\pi)^{l/2}|\Sigma_i|^{1/2}} \exp \left[-\frac{1}{2}(\mathbf{x} - \vec{\mu}_i)^T \Sigma_i^{-1}(\mathbf{x} - \vec{\mu}_i) \right], \quad i = 1, \dots, M, \quad (7)$$

where l is the dimensionality of \mathbf{x} , M is the number of possible classes (here $M = 2$), $\vec{\mu}_i$ represents the mean value, and Σ_i represents the covariance matrix of class ω_i . Estimates for the mean and the covariance matrix are computed from the training data, using the maximum likelihood estimation.

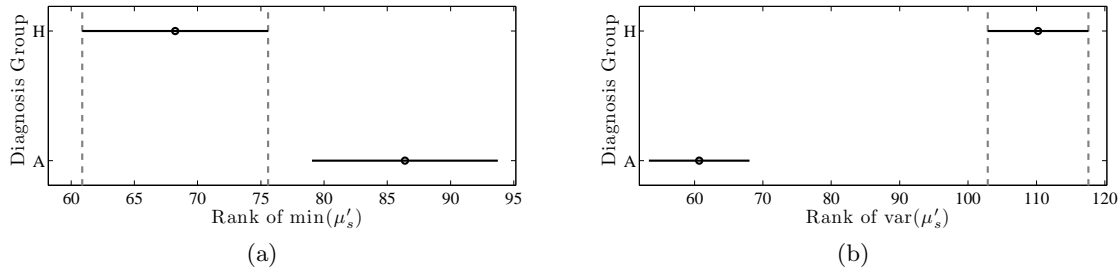


Figure 2. (a-b) Mean rank with standard error bars. (a-b) In these examples, the observed differences between the healthy (H) and affected (A) groups are statistically significant in parameters $\min(\mu'_s)$ and $\text{var}(\mu'_s)$ at the 95% confidence level. (b) The observed differences between groups H and A in $\text{var}(\mu'_s)$ are significant at the 99% confidence level.

3. RESULTS

3.1 Kruskal-Wallis ANOVA

Results from the Kruskal-Wallis test are summarized in Table 1. The number of distinct groups (k) is 2, thus H -statistic values greater than 3.841 and 6.635 are necessary to establish that observed differences between groups are statistically significant ($P < 0.05$ and $P < 0.01$, respectively). The H -statistic is greater than 3.841 for all extracted parameters, and therefore, the observed differences between the affected and healthy groups are statistically significant ($P < 0.05$). Furthermore, with the exception of $\min(\mu'_s)$, the observed differences between diagnosis groups are significant at the 99% confidence level (i.e. $H > 6.635$). Representative examples, Fig. 2, show that observed differences between affected and healthy for parameters $\min(\mu'_s)$ and $\text{var}(\mu'_s)$.

Table 1. Summary of H-statistic values.

Parameter Name	H -statistic
$\max(\mu_a)$	51.34
$\min(\mu_a)$	20.97
$\text{var}(\mu_a)$	51.96
$\text{ratio}(\mu_a)$	28.23
$\max(\mu'_s)$	39.67
$\min(\mu'_s)$	5.87
$\text{var}(\mu'_s)$	43.64
$\text{ratio}(\mu'_s)$	8.71

3.2 Diagnosis with either absorption or scattering coefficients

Results from classifying fingers as affected or not affected with RA with QDA using only either μ_a or μ'_s features are presented in this section. A typical QDA decision boundary that separates a 2-dimensional parameter space into 2 disjoint subsets is shown in Fig. 3. Each subset defines a region in parameter space as “healthy” or “affected.” Thus, images of finger joints whose extracted parameters are in the domain of the healthy or affected region are classified as “healthy” or “affected,” respectively.

In each case, an individual finger is represented by the four parameters of interest discussed in Section 2.1. The mean classification results are listed in Table 2. For the absorption images, the highest mean sensitivity values were obtained using $\min(\mu_a)$ and $\text{ratio}(\mu_a)$, resulting in $Se = 0.88$. The best mean specificity value was $Sp = 0.82$, with $\max(\mu_a)$, $\text{var}(\mu_a)$, and $\text{ratio}(\mu_a)$. The mean Youden index ($Y = Se + Sp - 1$) was maximized with $\min(\mu_a)$, $\text{var}(\mu_a)$, and $\text{ratio}(\mu_a)$, with the corresponding values being $Se = 0.85$, $Sp = 0.80$, and $Y = 0.65$.

Using the scattering coefficient images, the mean results were lower for sensitivity and notably higher for specificity. The

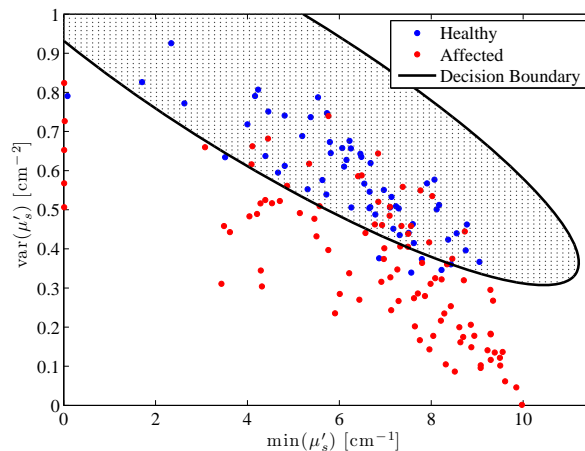


Figure 3. Distribution of and extracted from images of healthy (blue) and affected (red) patients. Points in the “shaded” region are classified healthy, while points outside the shaded region are classified affected.

Table 2. Classification using μ_a or μ'_s features independently.

Parameter Name	μ_a			μ'_s		
	<i>Se</i>	<i>Sp</i>	<i>Y</i>	<i>Se</i>	<i>Sp</i>	<i>Y</i>
$max(\mu), min(\mu)$	0.77	0.74	0.51	0.57	0.97	0.54
$max(\mu), var(\mu)$	0.80	0.68	0.48	0.63	0.85	0.49
$max(\mu), ratio(\mu)$	0.84	0.75	0.58	0.62	0.99	0.61
$min(\mu), var(\mu)$	0.83	0.74	0.57	0.80	0.90	0.70
$min(\mu), ratio(\mu)$	0.88	0.62	0.50	0.63	0.62	0.24
$var(\mu), ratio(\mu)$	0.80	0.80	0.60	0.75	0.76	0.52
$max(\mu), min(\mu), var(\mu)$	0.84	0.74	0.58	0.79	0.89	0.68
$max(\mu), min(\mu), ratio(\mu)$	0.80	0.76	0.56	0.61	0.98	0.60
$max(\mu), var(\mu), ratio(\mu)$	0.82	0.82	0.64	0.67	0.88	0.55
$min(\mu), var(\mu), ratio(\mu)$	0.85	0.80	0.65	0.79	0.91	0.70
$max(\mu), min(\mu), var(\mu), ratio(\mu)$	0.87	0.75	0.62	0.82	0.88	0.70

highest mean sensitivity value was $Se = 0.82$ using all four features, and highest mean specificity was $Sp = 0.99$ with $max(\mu'_s)$ and $ratio(\mu'_s)$. Three different feature combinations gave the highest mean *Youden* index of $Y = 0.70$. This result was reached using $min(\mu'_s)$ and $var(\mu'_s)$, with $Sp = 0.80$ and $Sp = 0.90$. The mean *Youden* index remained at the same value when $ratio(\mu'_s)$ was added to the feature combination, or when using all four features.

3.3 Diagnosis with scattering and absorption coefficients combined

Optical parameters from absorption and scattering coefficient distributions were combined so that an individual finger is represented by 8 parameters: $max(\mu_a)$, $min(\mu_a)$, $var(\mu_a)$, $ratio(\mu_a)$, $max(\mu'_s)$, $min(\mu'_s)$, $var(\mu'_s)$, and $ratio(\mu'_s)$. Multi-parameter classification can be performed using 2, 3, 4, 5, 6, 7, and 8 dimensional combinations of the extracted parameters. The number of different possible combinations is 28, 56, 70, 56, 28, 8, and 1, respectively. Quadratic discriminant analysis was performed on the total number of possible combinations (247). The highest *Youden* indices (including the corresponding *Se* and *Sp* values) obtained within each dimension are presented in Table 3.

The largest individual values for *Se* and *Sp* are 0.86 and 0.91, respectively. The largest *Youden* index, $Y = 0.75$ (with $Se = 0.84$, $Sp = 0.91$) was obtained in 4D, using $ratio(\mu_a)$, $min(\mu'_s)$, $var(\mu'_s)$, and $ratio(\mu'_s)$. The second largest *Youden* index is $Y = 0.73$ (with $Se = 0.81$ and $Sp = 0.91$), obtained in 3D from a combination of $max(\mu_a)$, $min(\mu'_s)$, and $var(\mu'_s)$. Interestingly, the smallest *Youden* index in Table 3 is $Y = 0.66$, obtained from classification in 8D.

Table 3. Summary of 2-8D quadratic discriminant analysis classification.

Dimension	Parameter Name	<i>Se</i>	<i>Sp</i>	<i>Y</i>
2	$min(\mu'_s), var(\mu'_s)$	0.83	0.90	0.72
3	$max(\mu_a), min(\mu'_s), var(\mu'_s)$	0.81	0.91	0.73
4	$ratio(\mu_a), min(\mu'_s), var(\mu'_s), ratio(\mu'_s)$	0.84	0.91	0.75
5	$min(\mu_a), ratio(\mu_a), min(\mu'_s), var(\mu'_s), ratio(\mu'_s)$	0.80	0.91	0.72
6	$max(\mu_a), min(\mu_a), var(\mu_a), max(\mu'_s), var(\mu'_s), ratio(\mu'_s)$	0.86	0.83	0.70
7	$max(\mu_a), min(\mu_a), var(\mu_a), ratio(\mu_a), min(\mu'_s), var(\mu'_s), ratio(\mu'_s)$	0.82	0.87	0.69
8	$max(\mu_a), min(\mu_a), var(\mu_a), ratio(\mu_a), max(\mu'_s), min(\mu'_s), var(\mu'_s), ratio(\mu'_s)$	0.80	0.85	0.66

4. CONCLUSION

Volumetric distributions of absorption (μ_a) and scattering (μ'_s) coefficients from FDOT imaging were used to diagnose RA in PIP joints II-IV. Four features of interest were extracted from each volumetric image $\{max(\mu), min(\mu), var(\mu), ratio(\mu)\}$. Thus, each imaged finger joint was represented by a total of 8 extracted features.

The non-parametric Kruskal-Wallis ANOVA test was employed to establish that observed differences between healthy and affected joints for all features considered were statistically significant at the 95% confidence level ($P < 0.05$). Computer aided diagnostics of finger joints as affected or not affected with RA was performed with QDA classification. First, classification was performed using only either μ_a or μ'_s features. Subsequently, classification was performed using all possible combinations of μ_a and μ'_s features.

Compared to absorption coefficient images, the scattering coefficient images provided better classification through very high specificity. All four features, $\{max(\mu), min(\mu), var(\mu), ratio(\mu)\}$, provided good results when included in a suitable combination. However, classification was significantly improved when μ_a and μ'_s features were combined. The best *Youden* index from combining features from μ_a and μ'_s images was $Y = 0.75$ ($Se = 0.84$, $Sp = 0.91$), while the best *Youden* indices when μ_a or μ'_s were used separately for classification were $Y = 0.65$ and $Y = 0.70$, respectively.

ACKNOWLEDGMENTS

This work was supported grants from the National Cancer Institute (5U54CA126513-03: Tumor Microenvironment Network - The role of inflammation and stroma in digestive cancer, Imaging Core) and the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS - 2R01 AR46255), which are part of the National Institutes of Health. Yrjö Häme is funded by the International Fulbright Science and Technology Award at the U.S. Department of State.

REFERENCES

1. Tromberg, B.J., Pogue, B.W., Paulsen, K.D., Yodh, A.G., Boas, D.A. and Cerussi, A.E., Assessing the future of diffuse optical imaging technologies for breast cancer management, *J. Med. Phys.* 35(6), 2443-2451, (2008).
2. Huppert, T.J., Diamond, S.G. and Boas, D.A., Direct estimation of evoked hemoglobin changes by multimodality fusion imaging, *J. Bio. Optics* 13(5), 195-201 (2008).
3. Lasker, J.A., Fong, C.J., Ginat, D.T., Dwyer, E. and Hielscher, A.H., Dynamic optical imaging of vascular and metabolic reactivity in rheumatoid joints, *J. Bio. Opt.* 12(5), 052001 (2007).
4. Meinel, L.A., Stolpen, A.H., Berbaum, K.S., Fajardo, L.L. and Reinhardt, J.M., Breast MRI lesion classification: Improved performance of human readers with a back propagation neural network computer aided diagnosis (CAD) system, *J. Magnetic Resonance Imag.*, 25(1), 89-95 (2007).
5. Hieslcher, A.H., Kim, H.K., Montejo, L.D., Blaschke, S., Netz, U.J., Zwaka, P.A., Illing, G., Mller, G.A. and Beuthan, J., Frequency-Domain Optical Tomographic Imaging of Arthritic Finger Joints, *IEEE Transactions in Medical Imaging*, (2011).
6. Klose, C.D., Klose, A.D., Netz, U., Beuthan, J. and Hielscher, A.H., Multi-parameter classifications of optical tomographic images," *J. Bio. Optics* 13(5), 050503 (2008).
7. Montejo, L.D., Montejo, J.D., Kim, H.K., Netz, U.J., Klose, C.D., Blaschke, S., Zwaka, P.A., Mller, G.A., Beuthan, J. and Hielscher, A.H., "Comparison of Classification Methods for Detection of Rheumatoid Arthritis with Optical Tomography," in *Biomedical Optics*, OSA Technical Digest (CD) (Optical Society of America, 2010), paper BWF2 (2010).
8. Klose, C.D., Klose, A.D., Netz, U.J., Beuthan, J. and Hielscher, A.H., Optical tomographic detection of rheumatoid arthritis with computer-aided classification schemes, *Proc. SPIE* 7171, 71710C (2009).
9. Klose, C.D., Klose, A.D., Netz, U.J., Scheel, A., Beuthan, J. and Hielscher, A.H., Computer-aided interpretation approach for opticaltomographic images, *J. Bio. Opt.* (2011).
10. Kim, H.K. and Hielscher, A.H., A PDE-constrained SQP algorithm for optical tomography based on the frequency-domain equation of radiative transfer, *Inverse Problems* 25, 015010 (2009).
11. Altman, D.G. and Bland, J.M., *Statistics Notes: Diagnostic tests 1: sensitivity and specificity*. *British Med. J.*, 308(6943), 1552-1552 (1994).
12. Theodoridis, S. and Koutroumbad, K., [Pattern Recognition], Elsevier, USA, (2006).
13. Glantz, S.A., [Primer of Biostatistics, Sixth Edition], McGraw-Hill, New York, (2005).