Numerical observer for cardiac motion assessment using machine learning

Thibault Marin[a], Mahdi M.Kalayeh[a], P. Hendrik Pretorius[b], Miles N. Wernick[a], Yongyi Yang[a], Jovan G. Brankov[a]
[a]Medical Imaging Research, Illinois Institute of Technology,
3440 S. Dearborn St., Chicago, IL, 60616, USA;
[b]Deptartment of Radiology, Division of Nuclear Medicine, University of Massachusetts Medical School, 55 Lake Ave. N., Worcester, MA, 01655, USA

ABSTRACT

In medical imaging, image quality is commonly assessed by measuring the performance of a human observer performing a specific diagnostic task. However, in practice studies involving human observers are time consuming and difficult to implement. Therefore, numerical observers have been developed, aiming to predict human diagnostic performance to facilitate image quality assessment. In this paper, we present a numerical observer for assessment of cardiac motion in cardiac-gated SPECT images. Cardiac-gated SPECT is a nuclear medicine modality used routinely in the evaluation of coronary artery disease. Numerical observers have been developed for image quality assessment via analysis of detectability of myocardial perfusion defects (e.g., the channelized Hotelling observer), but no numerical observer for cardiac motion assessment has been reported. In this work, we present a method to design a numerical observer aiming to predict human performance in detection of cardiac motion defects. Cardiac motion is estimated from reconstructed gated images using a deformable mesh model. Motion features are then extracted from the estimated motion field and used to train a support vector machine regression model predicting human scores (human observers' confidence in the presence of the defect). Results show that the proposed method could accurately predict human detection performance and achieve good generalization properties when tested on data with different levels of post-reconstruction filtering.

Keywords: Numerical observer; cardiac motion assessment; motion estimation; deformable mesh model; support vector machine.

1. INTRODUCTION

Image quality assessment is an important step in evaluation of new image processing techniques. In medical imaging, it has now become widely accepted that image quality assessment must adopt a task-based approach and measure the diagnosis performance of a human observer (physician) in a specific diagnostic task. Due to time and cost constraints, however, human observers are often replaced by numerical observers aiming at predicting the human observers' performance in a given diagnostic task. In cardiac imaging, single photon emission computed tomography (SPECT) is used for cardiac perfusion imaging to detect and evaluate coronary artery disease. In addition, cardiac-gated SPECT provides an image sequence which allows for cardiac motion assessment and assisting myocardium viability evaluation[1]. Diagnosis relies on detection of myocardial perfusion defects as well as assessment of cardiac motion. Therefore, image quality assessment for cardiac-gated SPECT images must include evaluation of perfusion defects detectability and visibility of cardiac motion.

For detection of perfusion defects, the channelized Hotelling observer (CHO)[2,3] is a widespread numerical observer used as a surrogate for human observers.[4–6] It relies on a set of non-overlapping bandpass filters to extract features, which are then classified using a linear discriminant (Hotelling observer or prewhitening matching filter). Since it can overestimate human performance[7], the CHO is often coupled with an internal noise model[7–9] to improve its agreement with human observers. The selection and tuning of the internal noise model resembles a machine learning procedure which led our group to develop a learning numerical observer using a support vector machine (SVM).[10]

CHO and SVM based numerical observer have been developed for image quality assessment via analysis of detectability of myocardial perfusion defects, but no numerical observer for cardiac wall motion assessment has been reported. In this

Medical Imaging 2011: Image Perception, Observer Performance, and Technology Assessment,
edited by David J. Manning, Craig K. Abbey, Proc. of SPIE Vol. 7966, 79660G · © 2011 SPIE
CCC code: 1605-7422/11/$18 · doi: 10.1117/12.878186

work, we present a method to design a numerical observer aiming to predict human performance in detection of cardiac motion defects.

In a clinical study of cardiac-gated SPECT images, a human observer (HO) scores the heart wall motion on the following semi-quantitative scoring system, as recommended in Ref. 1: 0 for normal, 1 for mild hypokinesis, 2 for moderate hypokinesis, 3 for severe hypokinesis, 4 for akinesis, and 5 for dyskinesis. In this paper we propose a numerical observer designed to predict HO confidence in the presence of the motion defect. During a HO study, reconstructed images were scored according to the level of confidence in presence of a motion defect. The location of the defect was known to both the human and numerical observers.

The proposed numerical observer relies on features extracted using our previously described cardiac motion estimation technique.[11,12] This method uses a deformable mesh model fitted to the left ventricle,[13] and deformed to track cardiac motion while accounting for myocardial brightening.[14] The estimated motion field is then used for feature extraction. Displacements along the axial, radial and tangential direction (with respect to the heart's long axis) are used to extract features in different regions of the myocardium. Features and human scores are used to train a support vector machine[15] designed to predict human scores. We have previously reported a similar approach using displacement norm for feature extraction[16] and a Hotelling observer for classification. The novelty of the proposed work mainly lies in the use of non-linear regression using support vector machines instead of a simple linear discriminant as in Refs. 16, 17.

Our results show that the proposed numerical observer can accurately predict human scores and that it can also offer good generalization performance. In practice, numerical observers are used to validate new image reconstruction or filtering methods, where they are used on images which differ greatly from the ones used for training. This results in a crucial need for numerical observers achieving high prediction performance when used on unseen images, i.e. images produced by a different method than the one used to train the numerical observers. To test this ability, in this work we study the numerical observer's performance obtained when testing the trained numerical observer on images filtered with different levels of temporal smoothing.

2. METHODS

In this section, we describe the design of the proposed numerical observer. The proposed method extracts motion features from reconstructed images and defines a nonlinear regression function to predict human scores from the set of motion features. First, we present the motion estimation procedure followed by the feature extraction step. Finally, we introduce the regression function used to predict human scores given the set of extracted features using support vector machines (SVM).

2.1 Cardiac motion estimation using a volumetric deformable mesh model

Here we briefly describe the motion estimation procedure. The reader is referred to our earlier work on motion estimation for a complete description.[11,12] Dense field cardiac motion is estimated using a volumetric left ventricular mesh model obtained from pre-reconstructed summed image.[13] This initial mesh structure is then deformed to track myocardial displacement during the image sequence. Motion estimation deforms the initial mesh structure by displacing mesh nodes in order to minimize an intensity miss-matching term within each mesh tetrahedron that composes the mesh structure. The intensity matching criterion accounts for myocardial brightening[14], an imaging artifact manifested by an increase in image intensity as the heart wall contracts.[6] The resulting nodal displacement is used for (linear) interpolation of the dense motion field to the regular Cartesian pixel grid (with dimension $64 \times 64 \times 64$). An example of a mesh structure and resulting dense motion field is shown on Fig. 1.

2.2 Feature extraction

The estimated dense motion field between end diastole and end systole is used to extract features in each region of the 17 segments map described in Ref. 1. While in our previous work[11] features were extracted from the displacement magnitude, in this work we use axial, radial and tangential components of the estimated motion field. For each component, four features are considered: mean, maximum, median and standard deviation. This configuration yields twelve features per segment. In addition, another feature was added based on myocardial brightening, which measures the change in intensity when the heart wall contracts. Note that the brightening does not utilize the estimated motion field. The final set of features is denoted by $\mathbf{x}$.

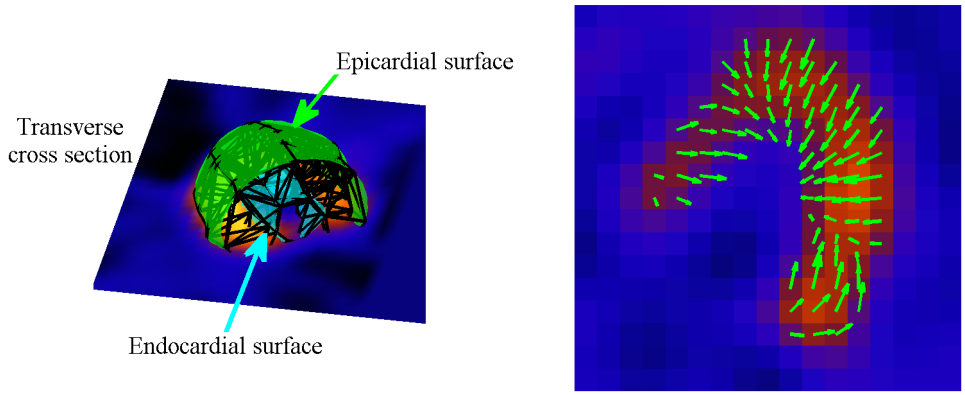


Figure 1. (left) Mesh structure fitted to the left ventricle (endocardial and surfaces are displayed), (right) Transverse view of the estimated dense motion field between end diastole and end systole.

2.3 Linear Hotelling observer for prediction of human scores

In our previous work[16,17], a linear Hotelling observer was used to perform regression from the set of extracted features **x**. In this work, we compare our new proposed SVM numerical observer to the linear observer described. The Hotelling discriminant between the two hypotheses (H_0 defect absent and H_1 defect present) defines the regression function, $f_{Hot}(.)$, using the Hotelling discriminant, $\mathbf{w}_{Hot}$, as:

$$f_{Hot}(\mathbf{x}) = \mathbf{w}_{Hot}^T \mathbf{x} = \Delta\boldsymbol{\mu}^T \mathbf{K}^{-1} \mathbf{x}\,, \tag{1}$$

where $\Delta\boldsymbol{\mu} = \boldsymbol{\mu}_1 - \boldsymbol{\mu}_0$ is the difference between the mean feature vector for each class and **K** is the covariance matrix obtained by averaging the covariance matrices under hypotheses H_0 and H_1. Comparing $f_{Hot}(\mathbf{x})$ to a variable threshold results in the receiver operating curve (ROC) which can be used to summarize the performance of a given observer.[18] In this work, we characterize an observer's performance by measuring the area under the ROC curve (AUC). Note that the Hotelling observer does not depend on the set of training human scores and therefore cannot adapt to different human observers.

2.4 Support vector machine for prediction of human scores

Here we describe the proposed model for prediction of human performance using SVM. Support vector machines are used to predict the human score Y from the set of extracted features **x**. To this end, the regression function $f_{SVM}(.)$ is expressed as:

$$f_{SVM}(\mathbf{x}) = \mathbf{w}_{SVM}{}^T \Phi(\mathbf{x}) + b\,, \tag{2}$$

where $\Phi(\mathbf{x})$ represents the nonlinear mapping of the features to a high-dimensional space where a linear regression is performed. $\mathbf{w}_{SVM}$ and b are parameters obtained from the training dataset $\{(\mathbf{x}_i, Y_i), i = 1,\ldots,N_s\}$, where $\mathbf{x}_i$ is the feature vector for training sample i, Y_i is the human score for training sample i, and N_s is the total number of training samples. The optimal values $\mathbf{w}_{SVM}^*$ and b^* correspond respectively to the weights and bias of the regression model and are given by:

$$\mathbf{w}_{SVM}^*, b^* = \arg\max_{\mathbf{w},b}\left(\frac{1}{2}\mathbf{w}^T\mathbf{w} + C\sum_{i=1}^{N_s} L_\varepsilon(\mathbf{x}_i)\right), \tag{3}$$

$$L_\varepsilon(\mathbf{x}_i) = \begin{cases} \left|Y_i - f_{SVM}(\mathbf{x}_i)\right| - \varepsilon, \left|Y_i - f_{SVM}(\mathbf{x}_i)\right| \geq \varepsilon \\ 0, otherwise \end{cases}, \tag{4}$$

where ε delimits the ε-*insensitive* band of the loss function L_ε. The optimal weights $\mathbf{w}^*_{SVM}$ are expressed in terms of support vectors, $\mathbf{s}_p$, $p = 1,\ldots,N_{SV}$ where N_{SV} is the total number of support vectors: $\mathbf{w}^*_{SVM} = \sum_{p=1}^{N_{SV}} \gamma_p \Phi(\mathbf{s}_p)$. The final regression is then expressed as:

$$f_{SVM}(\mathbf{x}) = \sum_{p=1}^{N_{SV}} \gamma_p \mathrm{K}(\mathbf{s}_p, \mathbf{x}) + b^* , \tag{5}$$

$$\mathrm{K}(\mathbf{s}_p, \mathbf{x}) = \Phi(\mathbf{s}_p)^T \Phi(\mathbf{x}) = \exp\left(-\frac{\|\mathbf{s}_p - \mathbf{x}\|^2}{2\sigma_K^2} \right), \tag{6}$$

where $\mathrm{K}(.,.)$ is the kernel function. In this work, we used Gaussian radial basis function kernels (RBF), parameterized by their width, σ_K. The optimal values of the parameters C, ε and σ_K were found by a three-dimensional grid search and selecting the parameter set yielding the lowest error in predicted AUC (compared to the human AUC) using a 5-fold cross-validation.[10,19] In our experiments SVM computations were performed using The Spider MATLAB toolbox.[20]

3. RESULTS

In this section we describe the validation strategy adopted to evaluate the proposed numerical observer. We first present the dataset used and the human observers studies performed to obtain a set of reference human scores. Then we present and discuss the results obtained using the linear Hotelling observer and the proposed SVM-based numerical observer.

3.1 Dataset

To generate imaging data we utilized the SIMIND software[21] which is a realistic Monte Carlo simulator of the SPECT imaging system. We used the MCAT phantom[22] (with size 64 × 64 × 64 in 16 time frames) where a cardiac motion defect was introduced in the anterior region as described in Ref. 23. Projections were 64 × 64 in 64 camera angles with voxel size 0.634 cm and myocardial activity corresponding to 500,000 detected counts originating from the heart region. Two datasets (one with no motion defect, one with an anterior motion defect, Fig. 2) were obtained, each containing 100 images corresponding to different noise realizations. Images were reconstructed using filtered back-projection followed by a low-pass Butterworth filter of order 5 and cutoff frequency 0.22 cycles per pixel and three different levels of temporal filtering (without motion compensation). Temporal filtering consists in applying a window on the temporal dimension of the image intensity, denoted by $\rho(\mathbf{r})$ using:

$$\tilde{\rho}^{(k)}(\mathbf{r}) = \sum_{l=1}^{K} h_{(l-k)} \rho^{(l)}(\mathbf{x}), \tag{7}$$

$$h_{(l-k)} = \frac{1}{A}\left(1 - \frac{2}{K} \min\left(|k-l|, K - |k-l| \right) \right)^{\gamma}, \tag{8}$$

where A is a normalization term ensuring unit DC gain, K is the total number of time frames and γ controls the strength of the filter. In this work, we considered three temporal filter strengths: no filter, light temporal filter and strong temporal filter, corresponding to $\gamma = \{\infty, 5, 1\}$. The three levels of temporal filtering will affect myocardium motion defect visibility.

3.2 Human reader study

Two hundred short-axis reconstructed image sequences (100 with no motion defect, 100 with anterior motion defect) were shown to a human observer for scoring the presence of the anterior motion defect (scores ranging from 0 to 5). The total study contained 600 images (200 per temporal filter level) for the human observer to score. Sample images at the end diastole and end systole for each dataset are shown on Fig. 2.

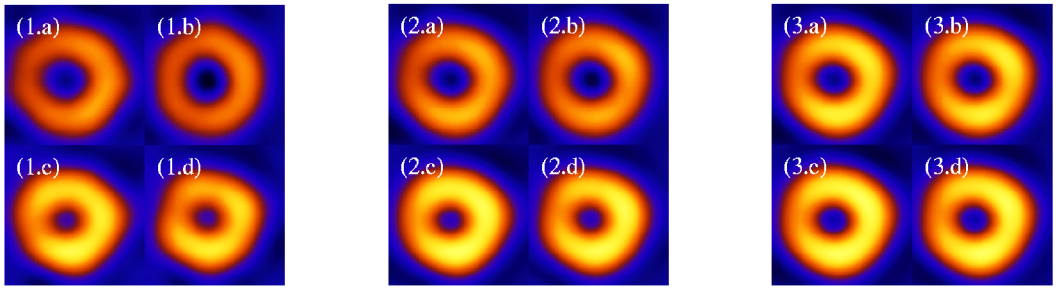


Figure 2. Images (MCAT phantom) for different level of temporal filtering (1-no temporal filtering, 2-light temporal filtering and 3-strong temporal filtering). Images with no cardiac motion defect (a,c) and with an anterior cardiac motion defect (b,d) are shown at the end diastole (a,b) and the end systole (c,d).

3.3 Prediction accuracy

The human scores obtained from data with no temporal filtering were used for training of a SVM. They correspond to the most simple reconstruction method, i.e. FBP followed by frame-by-frame spatial filtering. The resulting trained machine was used to predict the human scores of the three datasets (with different levels of temporal filtering). Results are analyzed through the receiver operator characteristic (ROC) curve for the human scores and the predicted scores. Note that the ROC curves were obtained using the online JROCFIT tool.[24] For comparison we also added the linear Hotelling observer (NO-HO) from Refs. 16, 17 using the same set of features. The ROC curves for different levels of temporal filtering are shown on Figs. 3, 4. The curves in Fig. 3 show the ROC curves obtained when applying the numerical observers (Hotelling and SVM) on the FBP data with no temporal filtering. This dataset is the one used for training of the SVM numerical observer (NO-SVM). Therefore, Fig. 3 shows the ability of the SVM observer to fit a given training dataset. As explained earlier, this ability is necessary but not sufficient to evaluate a numerical observer. Fig. 3 shows that the numerical observer (NO-SVM) closely matches the human observer (HO) when trained and tested on the same dataset (no temporal filtering applied). The proposed numerical observer achieves better prediction accuracy than the linear Hotelling observer (NO-HO).

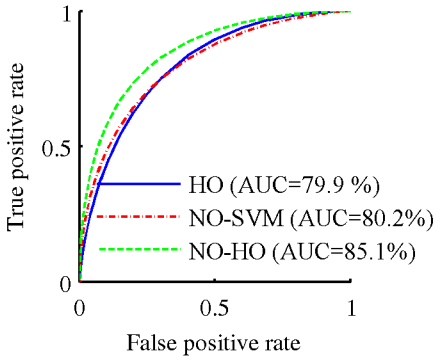
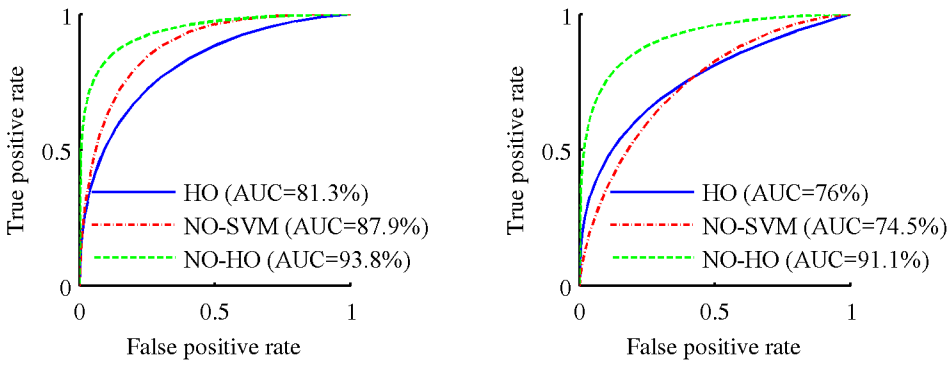


Figure 3. ROC curves for predicted scores from data with no temporal filtering (training data for the SVM numerical observer).

However, the ability to predict human performance in the training data does not guarantee that a numerical observer can accurately generalize to unseen data. Evaluation on unseen data is reported on Fig. 4, where the proposed numerical observer is tested on dataset with two types of temporal filtering. One can see that the proposed numerical observer generalizes well to unseen data as shown by the ROC curves obtained when tested with dataset generated using temporal filter (light and strong filter in Figs. 4(a) and 4(b), respectively). The proposed NO-SVM outperforms the Hotelling observer (NO-HO) from Refs. 17 in predicting the human observer's detection performance. In addition, one should note that the proposed SVM observer can adapt to different human observers while the Hotelling observer depends only on the images class and therefore cannot be adapted to model individual human observers.



(a) (b)

Figure 4. ROC curves for predicted scores from data with (a) light temporal filtering and (b) strong temporal filtering when the SVM numerical observer is trained on data with no temporal filtering. Curves and AUC values show that the proposed SVM numerical observer (NO-SVM) predicts better human detection performance.

4. CONCLUSION

We have proposed a new numerical observer for cardiac motion assessment. In cardiac SPECT, image quality assessment should include performance in detection of perfusion defect and assessment of cardiac motion. While numerical observers have been proposed for prediction of human scores in a perfusion defect detection task, automatic assessment of cardiac motion has not been thoroughly investigated. In this work, we propose a numerical observer for image quality via cardiac motion assessment. The proposed method relies on our 3D motion estimation technique and on regional features extracted from the estimated motion. We have shown that our proposed numerical observer can accurately predict human performance in a cardiac motion assessment task.

ACKNOWLEDGEMENTS

This work was supported by the National Institutes of Health under grants HL065425 and HL091017.

REFERENCES

[1] Hansen, C. L., Goldstein, R. A., Berman, D. S., Churchwell, K. B., Cooke, C. D., Corbett, J. R., Cullom, S. J., Dahlberg, S. T., Galt, J. R., Garg, R. K., Heller, G. V., Hyun, M. C., Johnson, L. L., Mann, A., McCallister, B. D., Taillefer, R., Ward, R. P., Mahmarian, J. J., and of the American Society of Nuclear Cardiology, Q. A. C., "Myocardial perfusion and function single photon emission computed tomography," Journal of Nuclear Cardiology 13, e97–e120 (2006).

[2] Myers, K. J. and Barrett, H. H., "Addition of a channel mechanism to the ideal-observer model," Journal of the Optical Society of America A 4, 2447–2457 (1987).

[3] Yao, J. and Barrett, H. H., "Predicting human performance by a channelized Hotelling observer model," in [Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series], Wilson, D. C. and Wilson, J. N., eds., Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference 1768, 161–168 (1992).

[4] Gifford, H. C., King, M. A., de Vries, D. J., and Soares, E. J., "Channelized Hotelling and human observer correlation for lesion detection in hepatic SPECT imaging," Journal of Nuclear Medicine 41, 514–521 (2000).

[5] Narayanan, M. V., Gifford, H. C., King, M. A., Pretorius, P. H., Farncombe, T. H., Bruyant, P. P. and Wernick, M. N., "Optimization of iterative reconstructions of 99mTc cardiac SPECT studies using numerical observers," IEEE Transactions on Nuclear Science 49, 2355–2360 (2002).

[6] Tang, J., Rahmim, A., Lautamäki, R., Lodge, M. A., Bengel, F. M. and Tsui, B. M. W., "Optimization of Rb-82 PET acquisition and reconstruction protocols for myocardial perfusion defect detection," Physics in Medicine and Biology 54, 3161–3171 (2009).

[7] Abbey, C. K. and Barrett, H. H., "Human- and model-observer performance in ramp-spectrum noise: effects of regularization and object variability," Journal of the Optical Society of America A 18, 473–488 (2001).
[8] Oldan, J., Kulkarni, S., Xing, Y., Khurd, P. and Gindi, G., "Channelized Hotelling and human observer study of optimal smoothing in SPECT MAP reconstruction," IEEE Transactions on Nuclear Science 51, 733–741 (2004).
[9] Zhang, Y., Pham, B. T. and Eckstein, M. P., "Evaluation of internal noise methods for Hotelling observer models," Medical Physics 34, 3312–3322 (2007).
[10] Brankov, J. G., Yang, Y., Wei, L., Naqa, I. E., and Wernick, M. N., "Learning a channelized observer for image quality assessment," IEEE Transactions on Medical Imaging 28, 991–999 (2009).
[11] Marin, T. and Brankov, J. G., "Deformable left-ventricle mesh model for motion-compensated filtering in cardiac gated SPECT," Medical Physics 37, 5471–5481 (2010).
[12] Brankov, J. G., Yang, Y., and Wernick, M. N., "Spatiotemporal processing of gated cardiac SPECT images using deformable mesh modeling," Medical Physics 32, 2839–2849 (2005).
[13] Feng, B., Modeling of the left ventricle (LV) by using mechanical models and image data, PhD thesis, University of Utah (2002).
[14] Galt, J. R., Garcia, E. V., and Robbins, W. L., "Effects of myocardial wall thickness on SPECT quantification," IEEE Transactions on Medical Imaging 9, 144–150 (1990).
[15] Vapnik, V. N., [Statistical learning theory], Wiley, New-York, NY (1998).
[16] Brankov, J. G., Marin, T., Pretorius, P. H., Yang, Y., and Wernick, M. N., "Numerical observer for cardiac motion assessment," in [Computational Imaging VIII], Bouman, C. A., Pollak, I., and Wolfe, P. J., eds., 7533, 75330G, SPIE, San Jose, CA, (2010).
[17] Marin, T., Pretorious, P. H., Yang, Y., Wernick, M. N., and Brankov, J. G., "Numerical observer for cardiac-motion," in [Proc. IEEE Nuclear Science Symposium], (2010 [Accepted for publication]).
[18] Barrett, H. H. and Myers, K. J., [Foundations of image science], Wiley-Interscience, Hoboken, NJ (2004).
[19] Duda, R. O., Hart, P. E. and Stork, D. G., [Pattern classification], Wiley, New-York, NY (2001).
[20] Weston, J., Elisseeff, A., Bakir, G., and Sinz, F., "The spider machine learning toolbox," [Available: http://www.kyb.tuebingen.mpg.de/bs/people/spider] (2005 [Accessed May 2010]).
[21] Ljungberg, M. and Strand, S.-E., "A Monte Carlo program for the simulation of scintillation camera characteristics," Computer Methods and Programs in Biomedicine 29, 257–272 (1989).
[22] Pretorius, P. H., King, M. A., Tsui, B. M. W., LaCroix, K. J., and Xia, W., "A mathematical model of motion of the heart for use in generating source and attenuation maps for simulating emission imaging," Medical Physics 26, 2323–2332 (1999).
[23] Pretorius, P. H., King, M. A., and Gifford, H. C., "A five-dimensional mathematical model for regional and global changes in cardiac uptake and motion," IEEE Transactions on Nuclear Science 51, 2634–2640 (2004).
[24] Eng, J., "ROC analysis: web-based calculator for roc curves," [Available: http://www.jrocfit.org] (2006 [Accessed June 2010]).