Automatic Noise Removal in MR Images Using Bilateral Filtering Associated with Artificial Neural Networks

Yu-Ju Lin and Herng-Hua Chang
Computational Biomedical Engineering Laboratory (CBEL)/Department of Engineering Science and Ocean Engineering/National Taiwan University, Daan 10617 Taipei, Taiwan
Email: r9887799@hotmail.com, herbertchang@ntu.edu.tw

***Abstract*—Noise removal in Magnetic Resonance (MR) images is important and essential for a wide variety of subsequent processing applications. Among the abundant denoising algorithms, the bilateral filter has been widely used in many image preprocessing procedures. However, it requires laborious tuning of parameters to obtain optimal filtering results, which is tedious and time-consuming. To address this problem, this paper is in an attempt to automate the bilateral filter based on an artificial neural network. Seven most significant attributes among 60 image attributes are used as the network input arguments. The BrainWeb image data with various scenarios of noise level, intensity non-uniformity, and slice thickness were adopted to evaluate this new system. Experimental results indicated that our automatic bilateral filter accurately predicted the denoising parameters and effectively removed the noise in MR images.**

***Index Terms*—bilateral filter, MRI, neural networks, denoise, Automation**

I. Introduction

Magnetic resonance imaging (MRI) has been one of the most frequently used medical imaging modalities due to its high contrast among different soft tissues, high spatial resolution across the entire field of view, and multi-spectral characteristics. In MR image analysis, noise is one of the main sources of quality deterioration not only for visual inspection but also in computerized processing such as tissue classification, segmentation and registration. Consequently, noise removal in MR images is important and essential for a wide variety of subsequent processing applications.

Over the decades, Gaussian filters have been widely used in many MR image processing applications for its simplicity [1]. Although the Gaussian filter smoothes noise quite efficiently edges are blurred significantly. To preserve the sharpness, the bilateral filter [2] has been proposed that performed effectively in MR image noise suppression and it has been the object of further studies [3]. However, the bilateral filter requires laborious tuning of parameters to obtain optimal filtering results, which is tedious and time-consuming. Automation of these parameters through artificial intelligence techniques will be highly beneficial. To address this problem, this paper proposes to automate the bilateral filter based on an artificial neural network.

Manuscript received January 10, 2015; revised March 10, 2015.

II. Related Works

A. Bilateral Filter

As shown in Fig. 1, the idea of the bilateral filter [2] is to combine gray levels based on both the geometric closeness and photometric similarity that is in favor of near values to distant values in both domain and range. More specifically, let (θ_x,θ_y) be the location of the pixel centered at a $(2N+1)\times(2N+1)$ neighborhood and

$$\Psi_{\theta_x,\theta_y}=\left\{(\mu_x,\mu_y):(\mu_x,\mu_y)\in[\theta_x-N,\theta_x+N]\times[\theta_y-N,\theta y+N]\right\} \quad (1)$$

be the pixels in the neighborhood of (θ_x,θ_y). The weighting functions for the spatial and radiometric components are defined respectively as

$$W^S_{\theta_x,\theta_y}(\mu_x,\mu_y)=\exp[-\frac{\left|(\mu_x,\mu_y)-(\theta_x,\theta_y)\right|^2}{2\sigma_s^2}] \quad (2a)$$

$$W^R_{\theta_x,\theta_y}(\mu_x,\mu_y)=\exp[-\frac{\left|I(\mu_x,\mu_y)-I(\theta_x,\theta_y)\right|^2}{2\sigma_R^2}] \quad (2b)$$

where $I(\cdot,\cdot)$ is the intensity value at the given position. The ensemble weight in the bilateral filter is the product of W^S and W^R:

$$W_{\theta_x,\theta_y}(\mu_x,\mu_y)=W^S_{\theta_x,\theta_y}(\mu_x,\mu_y)W^R_{\theta_x,\theta_y}(\mu_x,\mu_y) \quad (3)$$

In practice, each pixel is filtered using normalized weights to obtain the filtered image using

$$\tilde{I}(\theta_x,\theta_y)=\frac{\sum_{(\mu_x,\mu_y)\in\Psi}W_{\theta_x,\theta_y}(\mu_x,\mu_y)I(\mu_x,\mu_y)}{\sum_{(\mu_x,\mu_y)\in\Psi}W_{\theta_x,\theta_y}(\mu_x,\mu_y)} \quad (4)$$

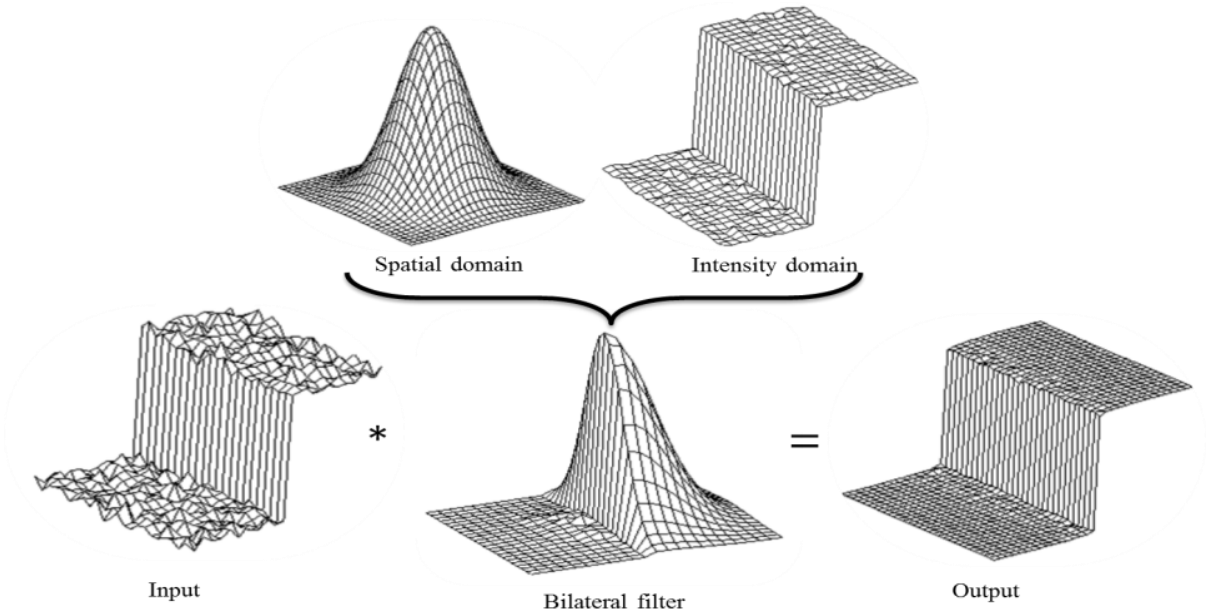

Figure 1. Illustration of the bilateral filter [2].

B. Gray Level Co-Occurrence Matrix

The gray level co-occurrence matrix (GLCM) [4] describes some easily computable textural features based on gray tone spatial dependencies using

$$M(i,j)=\sum_{x=1}^{Wx-dx}\sum_{y=1}^{Wy-dy}\begin{cases}1, \text{ if } W(x,y)=i \text{ and } W(x+dx,y+dy)=j\\ 0, \text{ otherwise}\end{cases} \quad (5)$$

where $M(i, j)$ is the quantized gray tone at position (i, j), Wx and Wy are the dimension of the resolution cells of the image ordered by their row-column designations, $W(x, y)$ is the gray level value in the cell, dx and dy are the spatial relation between two adjacent pixels defined by the angle θ and distance d from the cell origin. This texture-content information is then normalized to obtain the matrix of relative frequencies $P(i, j)$ as

$$P(i,j)=\frac{M(i,j)}{\sum_{i=0}^{W_x-1}\sum_{j=0}^{W_y-1}M(i,j)} \quad (6)$$

The derived textural features based on (6) include angular 2nd moment (ASM), contrast (CON), entropy (ENT), homogeneity (HOM), dissimilarity (DIS), mean, standard deviation (SD), and correlation (COR).

C. 2-D Discrete Wavelet Transform

By wavelet transform, we mean the decomposition of an image with a family of real orthonormal bases obtained through translation and dilation of a kernel function [5], [6]. Four subbands, namely LL1, LH1, HL1, and HH1, are obtained by the 1st order horizontal and vertical transformations sequentially. To obtain more detail information, the LL1 subband is further decomposed into four 2nd order subbands, LL2, LH2, HL2, and HH2 as shown in Fig. 2. After decomposing the images, the local wavelet coefficients in each subband can be computed based on the following energy equations [7]:

Norm-1 energy: $$e_1=\frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\left|x(m,n)\right| \quad (7)$$

Norm-2 energy: $$e_2=\frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\left|x(m,n)\right|^2 \quad (8)$$

Standard deviation: $$e_3=\frac{1}{MN}\sqrt{\sum_{m=1}^{M}\sum_{n=1}^{N}\left|x(m,n)-\bar{x}\right|^2} \quad (9)$$

where x represents the subband under consideration, M and N represent the dimension of the subband with $1 \le m \le M$ and $1 \le n \le N$, and $\bar{x}$ is the arithmetic mean of $x(m,n)$.

III. Methods

The proposed automatic bilateral filtering associated with the neural network framework consists of two major phases: training and testing, as shown in Fig. 3.

A. Feature Extraction

Three different categories are used to extract image features as shown in Fig. 4:

- Image statistics: compute the mean intensity (Mean), standard deviation (SD), variance (VAR), and entropy (ENT) of the input gray-level image.
- GLCM: first compute the difference image $I_n = I - I_D$, which is the difference between the input image I and

its Gaussian filtered image I_D. Then compute the textural features of GLCM using I_n with $d = 1$.

- 2-D DWT: first compute the normalized image I' based on Eq. (10). Then take the one and two stages of I' for the wavelet features using the Haar wavelet transform. Finally, compute the wavelet energy coefficients using (7)-(9)

$$I'(i,j) = \frac{I(i,j)}{\left(\frac{1}{MN}\sum_{k=1}^{M}\sum_{l=1}^{N} I(k,l)^2\right)^{1/2}} \quad (10)$$

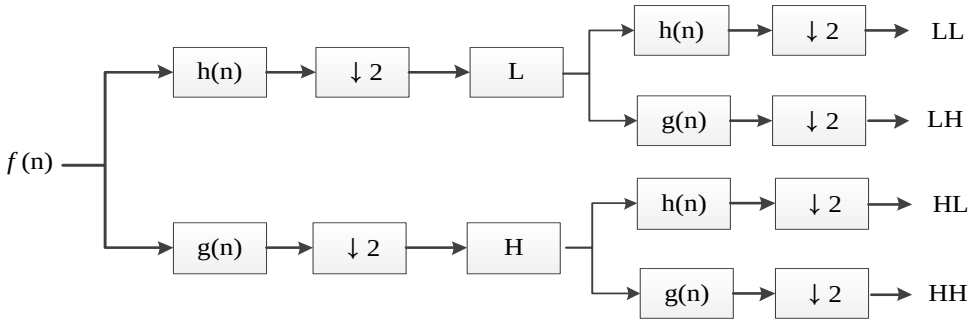

Figure 2. Illustration of 2-D DWT.

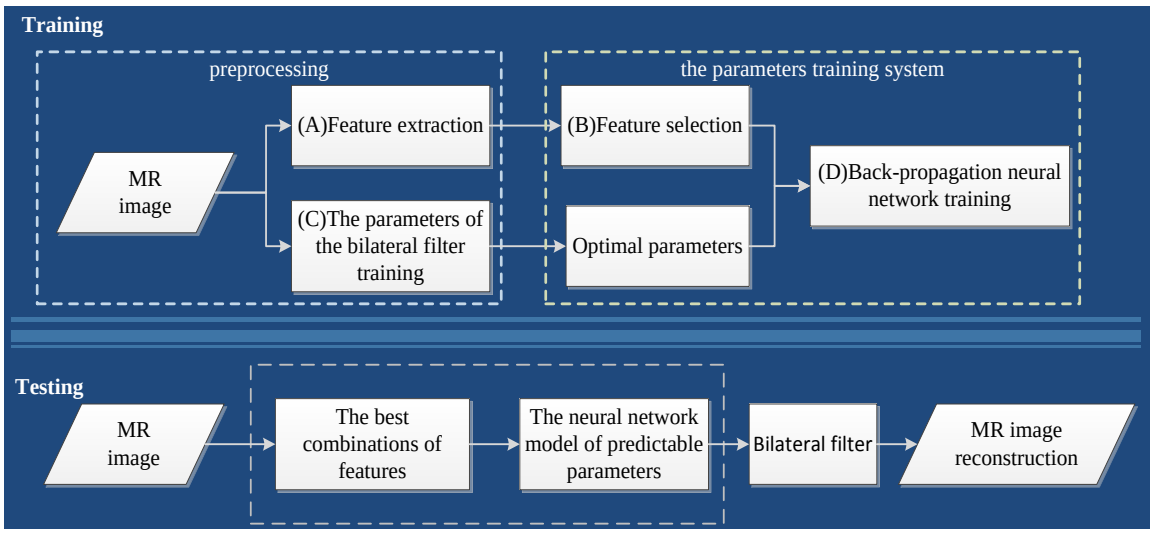

Figure 3. Flowchart of the proposed scheme.

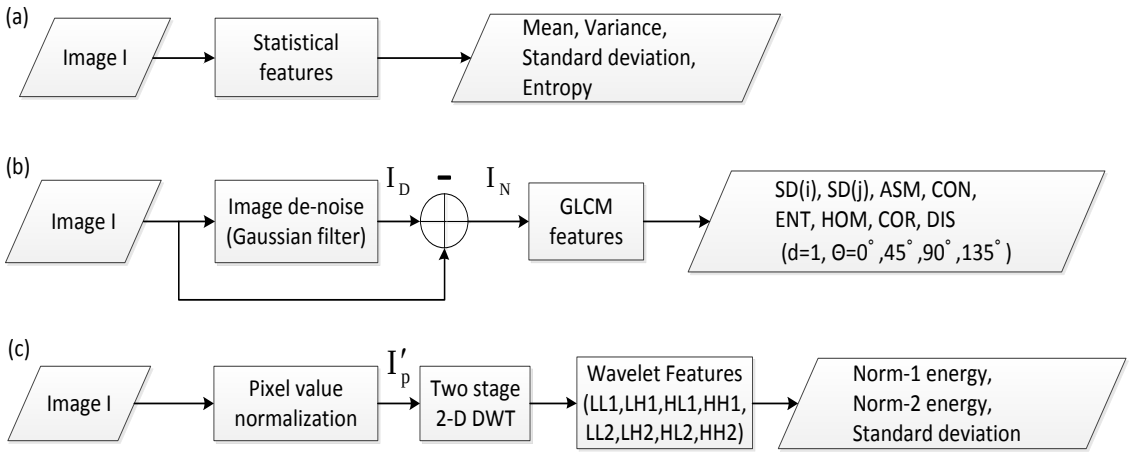

Figure 4. Image feature extraction: (a) image statistics, (b) GLCM, and (c) 2-D DWT.

B. Feature Selection

There are 60 different image features that are obtained based on the three feature extraction methods in every image. To obtain the most significant attributes, a paired-samples T-test [8], [9] is then applied to each individual image features to evaluate the ability of discrimination in two categories: Noise level and slice position. The evaluation is based on the distinguishing ability between noise levels, intensity distributions, and anatomical geometries according to the average *p*-value.

C. Optimal Parameter Selection

As described previously, there are three parameters in the bilateral filter: N, σ_S and σ_R. A brute-force method is conducted to find the optimal parameter settings on each individual image based on the peak signal-to-noise ratio (PSNR). The higher the PSNR values the better the restoration results. These optimal parameters are then used in the learning stage to train the neural network system as well as for the evaluation of the proposed automatic bilateral filtering system.

D. Back Propagation Network

The back propagation network (BPN) [10] with multilayer feedforward and error back propagation is adopted for the training of the automatic bilateral filtering system and for the testing stage also.

TABLE I. T-TEST RESULTS BASED ON THE *P*-VALUE: NOISE LEVEL

Classification: noise level	
p-value < 0.05	
p-value	Feature
[0.01, 0.02)	CON(90°)
[0.02, 0.03)	DIS(90°)
[0.03, 0.035)	e3(HL1), SD(x, 0°), SD(y, 0°), SD(x, 90°), SD(y, 90°), SD(x, 135°), SD(y, 135°), SD(x, 45°), SD(y, 45°), ASM(90°), CON(135°)
[0.035, 0.04)	ASM(45°), CON(0°), ASM(135°, 0°), HOM(45°)
[0.04, 0.05)	CON(45°), DIS(135°), ENT(0°), HOM(135°)
p-value ≥ 0.05	
p-value	Feature
[0.05, 0.07)	ENT(135°, 45°), DIS(0°), HOM(0°), DIS(45°), ENT(90°), COR(0°, 45°, 90°, 135°)
[0.07, 0.09)	e3(HH1, HH2), e1(HL1), HOM(90°), e1(HH1), e2(HH1), SD, VAR
[0.09, 0.2)	e1(HH2), ENT, e3(LL2, LL1), e1(LH1), e3(LH1)
[0.2, 0.4)	e2(HL2, LH2), e1(HL2), e2(HL1), e3(HL2), e2(HH2, LL2, LL1)
≥ 0.4	e1(LL1, LL2), e3(LH2), Mean, e1(LH2), e2(LH1)

IV. EXPERIMENTAL RESULTS

We have adopted the famous BrainWeb [11] image data of T1-weighted 1mm and 5mm MR image volumes with various levels of noise and intensity non-uniformity to evaluate the proposed system. Table I and Table II present the order of significance based on the average p-value of each individual feature using the T-test in noise level and slice position, respectively.

First, the image features were inserted into the classification tree of the classification and regression tree (CART) algorithm [12] for evaluation. Seven most significant image features from the 60 candidate features were obtained based on the average p-value using the T-test in noise level and slice position: CON(90°), DIS(90°), e3(LH1), e1(LL2), e3(HL2), e1(LL1), and e2(HL1). Subsequently, these features were used as the network input arguments in both training and testing phases.

Fifteen combinations of five noise levels and three intensity non-uniformities of normal scans with the same 1 mm slice thickness were used as the training dataset. Other different datasets with 1 mm / 5mm normal and multiple sclerosis scans were used as the testing evaluation. Table III presents the representative performance of our algorithm on the 1 mm multiple sclerosis dataset in comparison with the optimal results based on the PSNR values.

TABLE II. T-TEST RESULTS BASED ON THE *P*-VALUE: SLICE POSITION

Classification: slice position	
p-value < 0.05	
p-value	Feature
[0.01, 0.02)	e1(LL2)
[0.02, 0.03)	e3(LH2), e1(LL1), e2(LH1, LH2)
[0.03, 0.035)	e3(HL2), Mean, e1(LH2)
[0.035, 0.04)	e2(HH2, LL1)
[0.04, 0.05)	e2(LL2), e1(HL2)
p-value ≥ 0.05	
p-value	Feature
[0.05, 0.07)	e2(HL2)
[0.07, 0.09)	e2(HL1), e3(LL1)
[0.09, 0.2)	e3(LL2), e1(LH1), VAR, SD, e1(HH2)
[0.2, 0.4)	e3(HH2, LH1), e1(HL1), ENT, e3(HL1)
≥ 0.4	e2(HH1), SD(x, 135°), SD(y, 135°), SD(x, 90°), SD(y, 90°), COR(45°, 135°), SD(x, 0°), SD(y, 0°), SD(x, 90°), SD(y, 90°), COR(0, 90°), e1(HH1), e3(HH1), CON(45°, 135°), DIS(45°, 135°), ENT(45°, 135°, 90°, 0°), HOM(135°, 45°), CON(0°, 90°), ASM(135°), DIS(0°), ASM(45°), DIS(90°), ASM(0°, 90°), HOM(90°, 0°)

V. CONCLUSIONS

We have introduced a fully automatic system for bilateral filtering MR images based on a back-propagation neural network. We have systematically investigated significant attributes from various image features and textures as the network input arguments to facilitate subsequent

automation process. A wide variety of simulated T1-weighted MR images from the BrainWeb dataset were used to train and evaluate the proposed automatic filtering system. Experimental results indicated that our automatic bilateral filter accurately predicted the denoising parameters and effectively removed the noise in MR images. We believe that this new automatic filtering system is promising in a wide variety of MR image preprocessing applications that require automation.

TABLE III. COMPARISONS BETWEEN THE AUTOMATIC FILTERING RESULTS (PSNR_A) AND THE OPTIMAL FILTERING RESULTS (PSNR_O) FOR THE DATASET OF 1MM MULTIPLE SCLEROSIS SCANS

Noise	3%		5%	
Slice	PSNR_A	PSNR_O	PSNR_A	PSNR_O
1	34.2511	34.2468	29.3060	29.3092
21	35.2634	35.2804	30.8922	30.8894
32	32.3336	32.3351	28.0481	28.0490
55	35.0416	35.0471	31.2533	31.2657
66	34.4042	34.4034	31.1722	31.1740
96	33.3780	33.3779	30.5580	30.5585
105	35.4922	35.4881	30.8620	30.8629
118	33.8897	33.8964	30.1769	30.1767
126	34.8893	34.9016	30.6338	30.6358
140	33.9056	33.9137	29.9117	29.9199
Error	$\overline{\varepsilon_r}$ =0.0174±0.0143 (%)		$\overline{\varepsilon_r}$ =0.0108±0.0121 (%)	

ACKNOWLEDGMENT

This work was supported by the National Science Council under Research Grant No. NSC100-2320-B-002-073-MY3.

REFERENCES

[1] L. He and I. R. Greenshields, "A nonlocal maximum likelihood estimation method for Rician noise reduction in MR images," *IEEE Trans. Med. Imag.*, vol. 28, no. 2, pp. 165-172, 2009.

[2] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," *IEEE Proc. Int. Conf. Comput. Vis.*, Bombay, pp. 839-846, 1998.

[3] S. A. Walker, D. Miller, and J. Tanabe, "Bilateral spatial filtering: Refining methods for localizing brain activation in the presence of parenchymal abnormalities," *NeuroImage*, vol. 33, no. 2, pp. 564-569, 2006.

[4] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-3, pp. 610-621, 1973.

[5] J. Morlet, G. Arens, E. Fourgeau, and D. Glard, "Wave propagation and sampling theory-Part I: Complex signal and scattering in multilayered media," *Geophysics*, vol. 47, pp. 203-221, 1982.

[6] A. Grossmann and J. Morlet, "Decomposition of Hardy functions into square integrable wavelets of constant shape," *SIAM journal on mathematical analysis*, vol. 15, pp. 723-736, 1984.

[7] T. Chang and C. C. J. Kuo, "Texture analysis and classification with tree-structured wavelet transform," IEEE *Trans. Image Processing*, vol. 2, no. 4, pp. 429-441, Oct. 1993.

[8] R. A. Fisher, S. Genetiker, S. Genetician, G. Britain, and S. Généticien, *Statistical Methods for Research Workers*, 14, London: Oliver and Boyd Edinburgh, 1970.

[9] Student, "The probable error of a mean," *Biometrika*, pp. 1-25, 1908.

[10] R. Hecht-Nielsen, "Theory of the backpropagation neural network," *IJCNN*, vol. 1, pp. 593-605, 1989.

[11] BrainWeb: Simulated Brain Database, http://brainweb.bic.mni.mcgill.ca/brainweb.

[12] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, "Classification and regression trees, Wadsworth," *Journal of the American Statistical Association*, vol. 81, 1984.

Yu-Ju Lin received her M.S. in 2013 from the Department of Engineering Science and Ocean Engineering at National Taiwan University, Taipei, Taiwan. She was formerly a graduate student in the Computational Biomedical Laboratory (CBEL) directed by Professor Herng-Hua Chang in the Department of Engineering Science and Ocean Engineering at National Taiwan University. Her research interests include artificial intelligence, image denoising, and statistics analysis.

Herng-Hua Chang received his Ph.D. in biomedical engineering in 2006 from the University of California at Los Angeles (UCLA). He was formerly a postdoctoral scholar with the Laboratory of Neuro Imaging (LONI) and a member of the Center for Computational Biology (CCB) at UCLA. Currently, he is an assistant professor of the Department of Engineering Science and Ocean Engineering at National Taiwan University, Taipei, Taiwan. Dr. Chang founded the Computational Biomedical Laboratory (CBEL), whose goal is to build computational bridges between engineering and medicine. His research interests include variational methods for image restoration, registration and segmentation, computational biology and radiology, biomechanics modeling and biosystem simulation, pattern analysis and computer graphics, as well as medical informatics for healthcare applications.