Four-Neighborhood Clique Kernel: A General Framework for Bayesian and Variational Techniques of Noise Reduction in Magnetic Resonance Images of the Brain

Michael Osadebey,[1] Nizar Bouguila,[2] Douglas Arnold,[3] for the Alzheimer's Disease Neuroimaging Initiative

[1] Department of Electrical and Computer Engineering, Concordia University, Montreal, Quebec, Canada H3G 2W1

[2] Concordia Institute for Information Systems Engineering, Concordia University, Montreal, Quebec, Canada H3G 2W1

[3] NeuroRx Research Inc., Montreal, Quebec, Canada H2X 4B3

Received 5 December 2013; accepted 8 May 2014

ABSTRACT: Several algorithms have been proposed in the literature for image denoising but none exhibit optimal performance for all range and types of noise and for all image acquisition modes. We describe a new general framework, built from four-neighborhood clique system, for denoising medical images. The kernel quantifies smoothness energy of spatially continuous anatomical structures. Scalar and vector valued quantification of smoothness energy configures images for Bayesian and variational denoising modes, respectively. Within variational mode, the choice of norm adapts images for either total variation or Tikhonov technique. Our proposal has three significant contributions. First, it demonstrates that the four-neighborhood clique kernel is a basic filter, in same class as Gaussian and wavelet filters, from which state-of-the-art denoising algorithms are derived. Second, we formulate theoretical analysis, which connects and integrates Bayesian and variational techniques into a two-layer structured denoising system. Third, our proposal reveals that the first layer of the new denoising system is a hitherto unknown form of Markov random field model referred to as single-layer Markov random field (SLMRF). The new model denoises a specific type of medical image by minimizing energy subject to knowledge of mathematical model that describes relationship between the image smoothness energy and noise level but without reference to a classical prior model. SLMRF was applied to and evaluated on two real brain magnetic resonance imaging datasets acquired with different protocols. Comparative performance evaluation shows that our proposal is comparable to state-of-the-art algorithms. SLMRF is simple and computationally efficient because it does not incorporate a regularization parameter. Furthermore, it preserves edges and its output is devoid of blurring and ringing artifacts associated with Gaussian-based and wavelet-based algorithms. The denoising system is potentially applicable to speckle reduction in ultrasound images and extendable to three-layer structure that account for texture features in medical images. Int J Imaging Syst Technol, 24, 224–238, 2014; Published online in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/ima.22098

Key words: magnetic resonance imaging; Markov random field; single-layer Markov random field; total variation regularization; Tikhonov regularization; Euler-lagrange equation

Correspondence to: N. Bouguila; e-mail: nizar.bouguila@concordia.ca

Data used in preparation of part B of the experimental section of this report were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

Grant sponsor: Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI ; National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: Alzheimer's Association; Alzheimer's Drug Discovery Foundation; BioClinica, Inc.; Biogen Idec Inc.; Bristol-Myers Squibb Company; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; GE Healthcare; N. V. Innogenetics; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Medpace, Inc.; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Synarc Inc.; and Takeda Pharmaceutical Company. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego

I. INTRODUCTION

Medical images are the most important patient data for diagnostic, therapeutic, surgical, and prognostic procedures in clinical settings. Images acquired from magnetic resonance imaging (MRI) systems are important components of clinical trials of drugs for the treatment

of multiple sclerosis, Alzheimer, and other neurodegenerative diseases (Thulborn and Uttecht, 2000; Gaspari et al., 2009; Gold et al., 2012). The efficacy of image analysis performed either by visual inspections or computer aided systems is strongly dependent on level of noise in the image. Noise is unavoidably present in images because of imperfections of device components and the trade-offs in the operating parameters of imaging systems; signal-to-noise ratio (SNR), resolution, and length of scan time (Pizurica et al., 2003). Noise reduces diagnostic utility of images because it changes pixel configurations, weakens and can even obliterate what is supposed to be sharp edges, causes low spatial resolution as well as low contrast within and between anatomical structures. In the field of medical imaging and computer vision, noise reduction is at the crossroad between image processing and image analysis. Its invasive nature makes it the most delicate fundamental problem because there is high risk of introducing extraneous features which can further degrade image quality and reverse the gains derived from previous image processing tasks.

Many image denoising approaches have been proposed in the past. However, no particular algorithm can perform optimally for all range and types of noise and for all the different image acquisition modes. Careful review of their design reveals that sophisticated and state-of-the-art techniques can be classified into two major classes, global and specific schemes. The global schemes are general frameworks for specific ones, and their formulation generally results in a new specific scheme. They include geometrical framework proposed by Sochen et al. (1998) that combines linear heat flow, anisotropic diffusion, and mean curvature flow techniques into a single scheme, deformation by curvature technique proposed by Kimla and Siddiqi (1996), which unifies geometric heat equation and anisotropic diffusion, unification of probabilistic and variational techniques proposed by Hamza et al. (2002), and the global approach for solving the heat equation proposed by Auclair-Fortier and Ziou (2006). The specific schemes adopt one of three basic functions as core component. Gaussian and wavelets are two of the three basic functions. In the following two subsections, we describe how some of the state-of-the-art techniques are derived from the Gaussian and wavelets functions.

A. Gaussian-Based Algorithms. Gaussian filtering is the earliest classical technique for reducing noise in images (Deng and Cahill, 1993; Russo, 2005). A single two-dimensional Gaussian kernel is cast on a window moving in two orthogonal directions over the noisy image in a convolution process. At low noise levels when image structures dominate over noise signals, Gaussian filters in a single operation are excellent noise filters on images with pixels that are Gaussian distributed. At significant level of noise when image structures compete for visibility alongside noise signals and multiple operations of the denoising process are required, its rotationally invariant property results in smoothing both edges and homogeneous regions and this reduces the utility of the image for diagnosis. It is from this characteristic that the filter earns the name of smoothing kernel in literature (Monir and Siyal, 2011).

The convolution process in Gaussian filtering is the solution of two-dimensional diffusion equation of a physical system (Weickert, 1998; Wei, 2005), where the physical system under consideration is an image. This is the basis of anisotropic diffusion method pioneered by Perona and Malik (1990). In this method, a parameterized family of images satisfying partial differential equation describing diffusion process is generated from the noisy image. To reduce blurring, the diffusion tensor is configured as a symmetric positive-definite diffusion tensor chosen in such a way to encourage smoothing in homogeneous regions and discourage smoothing in the direction of edges. Variants of anisotropic diffusion can be found in Ma et al. (2012), Ford and El-Fallah (1997), and Gerig et al. (1992).

The bilateral filter introduced in Tomasi and Manduchi (1998) is a two-in-one Gaussian kernel filter. The first is the domain filter. It filters a pixel located in a square neighborhood within an image by assigning a closeness coefficient based on spatial proximity of the pixel to its neighbors. The other is the range filter which assigns a corresponding range coefficient based on similarity of its intensity values with its neighbors. The combination of both filters determines the denoised pixel by weighting the filter coefficients in the spatial domain with their corresponding pixel intensities. Detailed theoretical formulation behind the bilateral filter can be found in Hu et al. (2004) and Elad (2002).

Nonlocal means algorithm introduced in Buades et al. (2005) determines a denoised pixel at position in a neighborhood of an image from Gaussian weighted linear combination of all the pixels outside the neighborhood of the pixels under consideration. The degree of filtering and similarity features are measured using similarity weights determined by the parameters of the Gaussian kernel. Nonlocal means denoising was implemented for brain MRI images by Manjon et al. (2008) and Aksam Iftikhar et al. (2013).

In the kernel regression framework for image denoising introduced by Takeda et al. (2007), the authors formulate relationship between the number of noisy samples, estimator of the denoised pixels, and noise in a 2D image. Taylor series expansion around each data sample for estimation of the denoised pixels is implemented using spatially adaptive Gaussian kernel having steerable feature.

B. Wavelet-Based Algorithms. Wavelet function is family of wavelets generated by scalings and translations of a mother wavelet (Chui, 1992). In image processing task such as denoising, a family of wavelets is convolved with a degraded image to generate wavelet coefficients. In the transformed domain, the image signals are sparse, the coefficients are assumed Gaussian distributed, and the noise signal is spread out equally along all the coefficients. These properties are exploited to remove noise by reconstructing the image from the coefficients by retaining and eliminating coefficients of the image which are stronger and insignificant, respectively, relatively to a defined threshold energy of the Gaussian noise. This technique known as wavelet soft and hard thresholding method was pioneered by Donoho and Johnstone (1994, 1995a, 1995b). The threshold energy is a function of the noise level and the number of pixels in the image. Other thresholding methods have been formulated by manipulating the parameters of the threshold energy. They include VISU shrink (Donoho and Johnstone, 1995b), SURE shrink (Zhang and Desai, 1998; Luisier et al., 2007), and Bayes shrink (Chang et al., 2000). The authors in Nowak (1999), Pizurica et al. (2003), Wen et al. (2013), and Elad and Aharon (2006) have developed robust denoising methods by combining sparsity property with prior information of the image such as spatial smoothness for effective description and classification of the coefficients when the image is corrupted by various types of noise.

C. Limitations of Current Algorithms. Three major drawbacks common to current algorithms are high computational cost, introduction of artifacts into the denoised images, and the fact that their operations are restricted to nontexture images degraded by Gaussian-related noise. For some algorithms, such as nonlocal means, optimal performance is dependent on accurate estimate of

noise level. These setbacks are consequences of the strong parametric assumption inherent in their design. Current algorithms tend to attain optimal performance by adaptive estimate of Gaussian parameters that is robust to noise at each local region or wavelet scale. This makes the algorithms computationally intensive particularly when deployed to 3D MRI data of a single patient that may contain as much as 50 images (Hu et al., 2004; Coupe et al., 2008). The computed Gaussian parameters are not the ideal expectation because local structures of the image and wavelet coefficients at each scale cannot be well described by Gaussian distribution (Rajwade et al., 2013) resulting in blurring and ringing artifacts for Gaussian-based and wavelet-based algorithms, respectively.

D. Our Contribution. In this article, we propose a new two-layers structured general framework for image denoising which overcomes some of the problems of current algorithms. It is based on four-neighborhood clique system, an idea borrowed from graphical models (Koller and Friedman, 2009). We show that the four-neighborhood clique is the third basic filter which is a core component in the formulation of Markov random field model (Geman and Geman, 1984), Rudin, Osher, and Fatemi (ROF) total variation (Rudin et al., 1992), and Tikhonov regularization (Tikhonov and Arsenin, 1978) techniques for reducing noise in images. Our proposal results in a new algorithm which we refer to as the single-layer Markov random field (SLMRF) model. In the model, we detach the classical Markov random field model from Bayesian inference beliefs so that the observed image such as brain MRI image is the only physical system under consideration. SLMRF preserves edges in images and is computationally efficient as it denoises an image without the need to compute regularization parameter or reference to a prior model. The general framework can also be extended to three-layers structure to preserve texture features and reduce multiplicative noise in ultrasound images.

E. Outline. In section II, we describe the sources of data used for our experiments and explain the theoretical background behind our proposal. Section III displays the experimental results obtained by comparing our proposed SLMRF model with three popular denoising techniques. The results of the experiments are discussed in section IV. Section V concludes this article.

II. MATERIALS AND METHOD

A. Materials. The materials used for this article are 300 real MRI data provided by two international clinical trial study management centers: NeuroRx Research Inc. and the Alzheimer's Disease Neuroimaging Initiative (ADNI). Data from both sources were processed T2, T1, FLAIR, and PD weighted images formatted as 16 bit unsigned integer DICOM file of dimension 256 $\times$ 256. The two organizations that provided the data are described below.

A.1. NeuroRx Research. NeuroRx Research (http://www.neurorx.com/en/home.htm) is dedicated to working with the pharmaceutical industry to facilitate clinical trials of new drugs for multiple sclerosis (MS) and other neurological diseases. NeuroRx provides professional management of all MRI-related study activities and promptly delivers precise MRI outcome measurements that are performed in a regulatory compliant environment. The organization specializes in logistics of scan handling and tracking and can provide this service exclusively, if needed. NeuroRx uses advanced image analysis techniques to provide precise outcome data that maximize study power. Images are corrected for inhomogeneity and coregistered for perfect realignment and increased precision. Analysis is conducted in 3D, rather than on slices, so that information can be properly related to structures that span multiple slices. Customized automatic segmentation techniques are combined with expert supervision to maximize the precision of outcome measures related to both lesional and nonlesional pathology, as well as brain volume changes.

A.2. ADNI. The ADNI (adni.loni.usc.edu) was launched in 2003 by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies, and nonprofit organizations, as a $60 million, 5-year public-private partnership. The primary goal of ADNI has been to test whether serial MRI, positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD). Determination of sensitive and specific markers of very early AD progression is intended to aid researchers and clinicians to develop new treatments and monitor their effectiveness, as well as lessen the time and cost of clinical trials.

The Principal Investigator of this initiative is Michael W. Weiner, MD, VA Medical Center and University of California, San Francisco. ADNI is the result of efforts of many coinvestigators from a broad range of academic institutions and private corporations, and subjects have been recruited from over 50 sites across the U.S. and Canada. The initial goal of ADNI was to recruit 800 subjects but ADNI has been followed by ADNI-GO and ADNI-2. To date, these three protocols have recruited over 1500 adults, ages 55–90, to participate in the research, consisting of cognitively normal older individuals, people with early or late MCI, and people with early AD. The follow up duration of each group is specified in the protocols for ADNI-1, ADNI-2, and ADNI-GO. Subjects originally recruited for ADNI-1 and ADNI-GO had the option to be followed in ADNI-2. For up-to-date information, see www.adni-info.org.

B. Methods. The specific algorithms within our proposed generalized framework are presented as layered-structured schemes. When only the test image is the physical system under consideration the scheme is regarded as single layer, and as double layer when both the test image and a classical prior are considered. We begin this section with description of the four-neighborhood kernel, the building block of our proposed global scheme, followed by explanation of how the total clique potential of an MRI slice image varies with noise level. Next, we describe a new specific scheme which we refer to as SLMRF and how assignment of different smoothness costs to the kernel configures a test image for Bayesian and variational techniques. The four-neighborhood kernel was used as basic function to formulate our proposed SLMRF model, ROF total variation, Tikhonov, and classical Markov random field model techniques. We also provide theoretical formulation on how to "*navigate*" from ROF total variation to the classical Markov random field model, and vice versa. The section concludes by integrating the aforementioned specific schemes into a global scheme.

B.1. The 4-Neighborhood Kernel. Given a 2D image I of size $m \times m$, the four-neighborhood clique system $\mathbf{N}$ having clique size of two is the set consisting of four pixels that are closest in distance to any pixel located at (x, y):

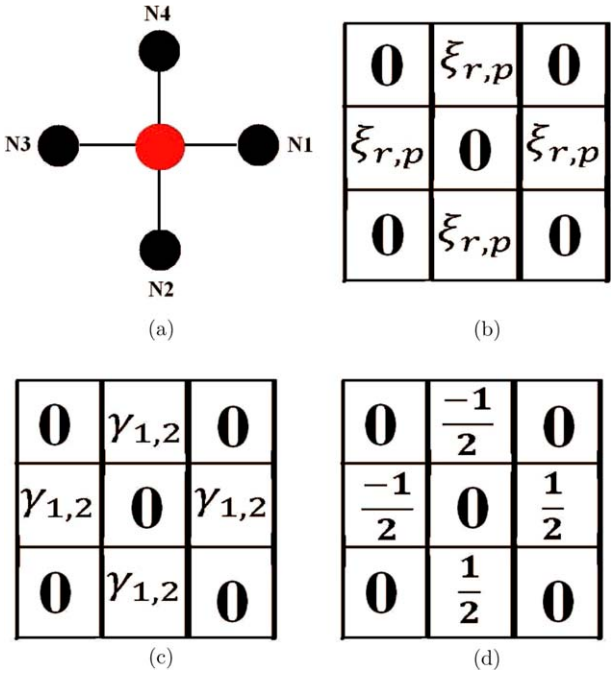

Figure 1. (a) Structural description of the four-neighborhood clique system. The four nearest neighbor of a pixel (in red color) located at (*x*, *y*) are the black colored pixels labeled N1, N2, N3, and N4. (b) The general form of the kernel for the four-neighborhood clique system. (c) Kernel for configuring an image for Bayesian mode of operation. (d) Kernel for configuring an image for variational mode of operation. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

$$\mathbf{N}=\{\mathbf{N_1}=(x,y+h),\mathbf{N_2}=(x+h,y),\mathbf{N_3}=(x,y-h),\mathbf{N_4}=(x-h,y)\}, \tag{1}$$

where x and y are indices of the rows and columns within the image, **N1, N2, N3**, and **N4** are labels identifying each neighbor, and h is the interval between each grid. The structure of the system is displayed in Figure 1a. The red colored pixel is the reference pixel and the four neighbors are the black colored pixels. The intensity levels of pixels that constitute the image belongs to a family of random variables $F=\{F_1,\ldots,F_M\}$ defined on the set of pixel locations called sites:

$$S=\{(x,y)|1\le x,y\le m\}, \tag{2}$$

where $|S|=M=m\times m$. Assignment of intensity levels, for example, for a grey level image, from the set $L=\{0,1,\ldots,255\}$ to each site S is an event called configuration $f=\{f_1,\ldots,f_M\}$. In the clique system, the reference pixel located at (x,y) is denoted i and any of its neighbors N1, N2, N3, and N4 located as expressed in Eq. (1) is denoted i'. They are assigned configurations f_i and $f_{i'}$, respectively. These configurations are adopted as indices to compute the Gibbs energy. The Gibbs energy $U(f)$ of an image with pixel configuration $f\in F$ is the sum of single site V_1 and pair-site V_2 cliques (Li, 2009):

$$U(f)=\sum_{i\in S}\alpha_1 V_1(f_i)+\sum_{i\in S}\sum_{i'\in N}\alpha_2 V_2(f_i,f_{i'})\equiv E_d+E_s, \tag{3}$$

where the first and second terms are the data E_d and spatially varying smoothness E_s terms, respectively, and α_1 and α_2 are the interaction coefficients. At each local clique N, where the reference pixel is $i\in S$, the contribution of each neighboring pixel $i'\in N$ to the local smoothness energy is determined according to the expression

$$V_2(f_i,f_{i'})=\begin{cases}\xi_r & \textbf{if } f_i=f_{i'}\\ \xi_p & \textbf{otherwise}\end{cases}, \tag{4}$$

where ξ_r is the reward and ξ_p the penalty for conformity $f_i=f_{i'}$ and violations $f_i\neq f_{i'}$ of the smoothness constraints, respectively. Henceforth, the two terms, Markov random field energy and smoothness energy, are considered equivalent and have same meaning for a given image. Matrix formulation of the kernel for computing the local clique potential energy is

$$\mathbf{H}=\begin{bmatrix}0 & \xi_{r,p} & 0\\ \xi_{r,p} & 0 & \xi_{r,p}\\ 0 & \xi_{r,p} & 0\end{bmatrix}, \tag{5}$$

where $\xi_{r,p}$ denotes assignment of either reward ξ_r if $f_i=f_{i'}$ or penalty ξ_p if $f_i\neq f_{i'}$. The physical structure of the kernel corresponding to this matrix is displayed in Figure 1b.

Assignment of scalar values γ_1 and γ_2 to ξ_r and ξ_p, respectively, in Eq. (4) results in a scalar valued smoothness cost at each local clique:

$$V_2(f_i,f_{i'})=\begin{cases}\xi_r=\gamma_1 & \textbf{if } f_i=f_{i'}\\ \xi_p=\gamma_2 & \textbf{otherwise}\end{cases}, \tag{6}$$

and **H** in Eq. (5) becomes

$$\mathbf{H_s}=\begin{bmatrix}0 & \gamma_{1,2} & 0\\ \gamma_{1,2} & 0 & \gamma_{1,2}\\ 0 & \gamma_{1,2} & 0\end{bmatrix}, \tag{7}$$

where $\gamma_{1,2}$ denotes assignment of either reward γ_1 if $f_i=f_{i'}$ or penalty γ_2 if $f_i\neq f_{i'}$. Assignment of vector values of ξ_{r_j} and ξ_{rk} to ξ_r and ξ_{p_j} and ξ_{p_k} to ξ_p in two orthogonal directions **j** and **k**, respectively in Eq. (4) results in a vector valued smoothness cost expressed as

$$V_2(f_i,f_{i'})=\begin{cases}\xi_{r_j}=\xi_{r_k}=0 & \textbf{if } f_i=f_{i'}\\ \xi_{p_j} & \textbf{otherwise, j direction}\\ \xi_{p_k} & \textbf{otherwise, k direction}\end{cases}. \tag{8}$$

The penalties ξ_{p_j} and ξ_{p_k} can be expressed using standard image location notation:

$$\begin{aligned}\xi_{p_j}&=\frac{I(j+h,k)-I(j-h,k)}{2h}\\ \xi_{p_k}&=\frac{I(j,k+h)-I(j,k-h)}{2h}\end{aligned}. \tag{9}$$

In this case **H** in Eq. (5) becomes

$$\mathbf{H_v}=\begin{bmatrix}0 & -\frac{1}{2} & 0\\ -\frac{1}{2} & 0 & \frac{1}{2}\\ 0 & \frac{1}{2} & 0\end{bmatrix}. \tag{10}$$

The physical structure of the kernel corresponding to $\mathbf{H_s}$ and $\mathbf{H_v}$ are displayed in Figures 1c and 1d, respectively.

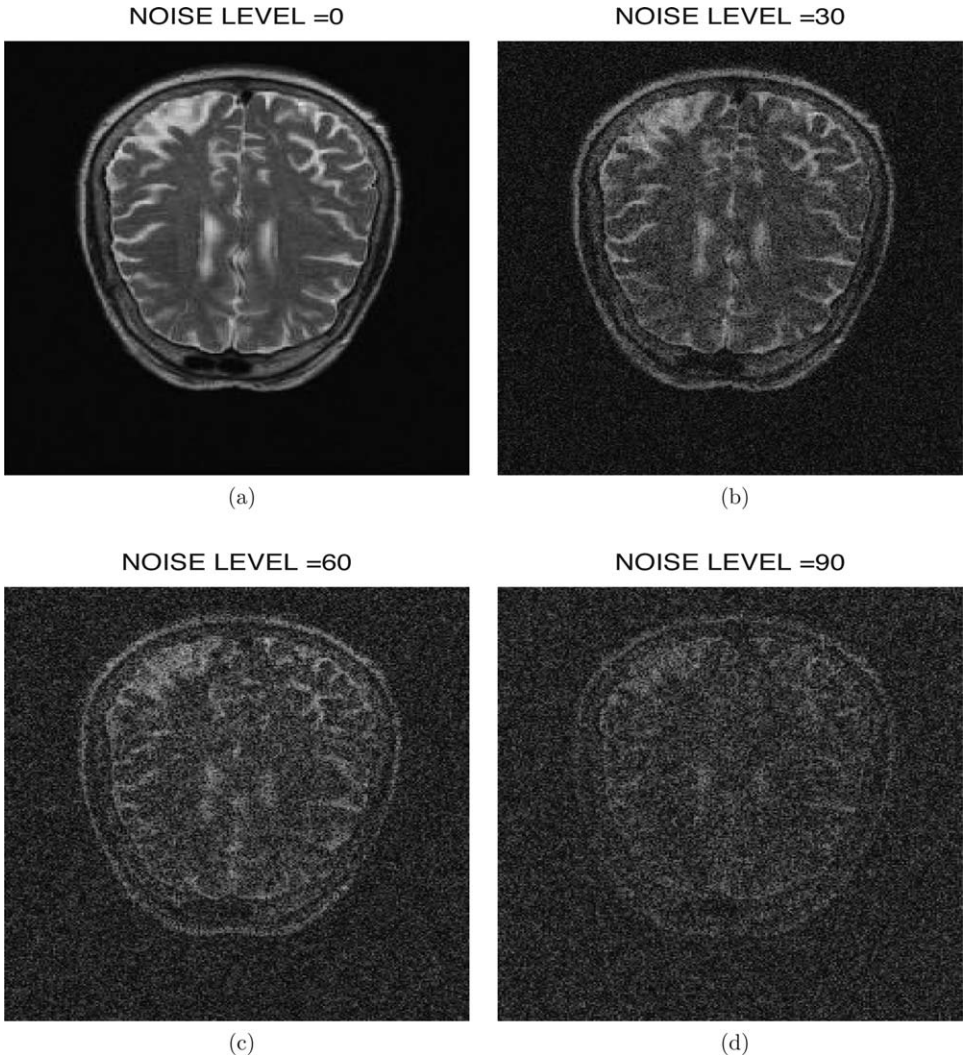

Figure 2. An MRI slice image at various levels of degradation by Rician noise. Observe the different image pixel configurations associated with noise levels of (a) $\sigma = 0$, (b) $\sigma = 30$, (c) $\sigma = 60$, and (d) $\sigma = 90$.

In the degradation process, different image pixel configurations $f \in F$ can be realized from an observed image $I_{\sigma_n}(f)$ degraded by noise σ_n of level n. The convolution of $I_{\sigma_n}(f)$ with $\mathbf{H_s}$ is scalar valued total smoothness energy $U(f)$ computed from sum of local clique potentials. We denote the scalar value as γ and express it as a function of image pixel configuration $\gamma(f)$:

$$U(f)=\sum_x \sum_y I_{\sigma_n}(x-j,y-k)\mathbf{H_s}(j,k)=\gamma(f). \tag{11}$$

Conversely, convolution with $\mathbf{H_v}$ results in total smoothness energy $U(f)$ that is vector valued and is expressed as function of image gradient $\nabla I(f)$:

$$U(f)=\sum_x \sum_y I_{\sigma_n}(x-j,y-k)\mathbf{H_v}(j,k)=\nabla I(f), \tag{12}$$

where j and k are the two orthogonal directions of the convolution process. The magnitude of the vector valued total energy can be measured either in L_1 or L_2 norm:

$$L_q=\begin{cases} (||\nabla I||) & \text{if } q=1 \\ (||\nabla I||^2) & \text{if } q=2 \end{cases}, \tag{13}$$

where q is the choice of norm.

B.2. SLMRF. In the SLMRF, the observed image is a specific image data such as a brain MRI slice image $I_{\sigma_n}(f)$ having pixel configuration f degraded by noise σ_n of level n. The observed image is the only physical system under consideration, and there is no reference to a classical prior model. The image is assumed to be initially clean and without noise $\sigma_0 = 0$ as shown in Figure 2a. The energy at the noise-free state is regarded as the equilibrium energy state. The observed noisy or higher energy state σ_n is attained by stepwise $\Delta\sigma$ incremental degradation by noise. Different levels of noise $\sigma_{30}=30$, $\sigma_{60}=60$ and $\sigma_{90}=90$ results in different pattern of arrangement of pixels, different spatial coherence and hence different image energy as shown in Figures 2b–2d. The mathematical model describing relationship between normalized smoothness energy $\{E : -\infty < E \leq 1\}$ and noise level σ for T2 weighted MRI image with background (bg) and foreground (fg) was reported in our previous work (Osadebey et al., 2013):

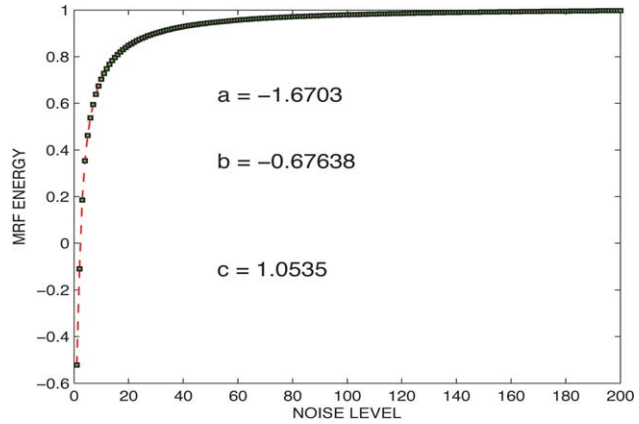

Figure 3. The plots of the proposed generalized mathematical models for describing the relationship between total clique potential energy and noise level of brain MRI slice images in the (a) background and (b) foreground modes. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

$$\begin{gathered} E=a\hat{\sigma}^{b}+c \\ a_{\mathrm{bg}}=-1.67, a_{\mathrm{fg}}=-0.6863 \\ b_{\mathrm{bg}}=-0.6764, b_{\mathrm{fg}}=-0.3663 \\ c_{\mathrm{bg}}=1.053, c_{\mathrm{fg}}=1.105 \end{gathered} . \tag{14}$$

From the plots of the model displayed in Figure 3, the minimum energies for background and foreground are $U_{\mathrm{bg}}=-0.6$ and $U_{\mathrm{fg}}=0.5$, respectively. The energies associated with the different levels of noise are computed by convolving the image with $\mathbf{H_v}$ according to Eq. (12). Different levels of noise σ_n result in different pixels configurations $I_{\sigma_n}(f)$, different image gradients $\nabla I(\sigma(f))$, and hence different levels of energy $U_n(\sigma(f))$. Thus, the smoothness energy U of the image I is a functional having three variables namely the image variance σ_n, the image pixel configurations $I_{\sigma_n}(f)$, and the image gradient $\nabla I(\sigma(f))$ measured in the L_1 norm:

$$U(\sigma_n, I_{\sigma_n}(f), \nabla I(\sigma(f)))=\sum_x \sum_y (||(\nabla I_{\sigma_n}(f))_{x,y}||). \tag{15}$$

The task of noise reduction is to return the image from the higher energy state σ_n to the equilibrium energy state σ_0. The Euler-Lagrange partial differential equation corresponding to Eq. (15) is (Chambolle, 2004; Lanczos, 1986):

$$\frac{\partial U}{\partial I}-\frac{\partial U}{\partial \sigma}\frac{\partial U}{\partial(||\nabla I||)}=0. \tag{16}$$

The classical method to derive the solution of Euler-Lagrange equation is by adopting finite difference method of approximating partial derivatives (Thomas, 1995). To make the expression in Eq. (16) well posed, the variable $||\nabla I||$ is perturbed by an amount $\{\beta : 0<\beta\leq 1\}$ and expressed as:

$$||\nabla I||=\sqrt{||\nabla I||^2+\beta}, \tag{17}$$

so that the solution of the perturbed problem will converge to a solution of the original expression as $\beta \rightarrow 0$ (Acar and Vogel, 1994). After perturbation, the solution $I_{\sigma\sigma}$ becomes:

$$I_{\sigma\sigma}=\frac{(I_x(x,y)^2+\beta)I_{yy}(x,y)-2I_x(x,y)I_y(x,y)I_{xy}(x,y)}{(I_x(x,y)^2+I_y(x,y)^2+\beta)^{3/2}}+\frac{I_{xx}(x,y)(I_y(x,y)+\beta)^2}{(I_x(x,y)^2+I_y(x,y)^2+\beta)^{3/2}} . \tag{18}$$

Based on the theoretical formulation of our algorithm, we adopt variance marching scheme with variance step of 0.1 in the gradient descent minimization scheme:

$$I_{\sigma+1}=I_\sigma-\eta_\sigma\nabla(I_\sigma). \tag{19}$$

Where the notations $I_{\sigma+1}, I_\sigma, \eta_\sigma$ and $\nabla(I_\sigma)=I_{\sigma\sigma}$ are the currently observed image, the initial image, the variance step and the gradient of the initial image. The minimization of the image gradient is subject to the mathematical model describing relationship between Markov random field energy and noise level in Eq. (14). At each iteration, the Markov random field energy expressed by the total clique potential is computed according to the formulation in Osadebey et al. (2013). Using the mathematical model described by the plots in Figure 3, the stopping criterion E_s is fixed at $E_s=K_{\mathrm{bg}}U_{\mathrm{bg}}$ and $E_s=K_{\mathrm{fg}}U_{\mathrm{fg}}$ for background and foreground modes, respectively, where K_{bg} and K_{fg} are arbitrary constants. These constants are multiplication factors, derived from experimental observations, which determines the number of iterations before the energy E_s is attained in the minimization process. The images we utilize for our experiments were with background and we choose $K_{bg}=2.5$.

B.3. Double-Layer Markov Randon Field: ROF Total Variation and Tikhonov Techniques. ROF total variation technique follows the same theoretical formulation as the SLMRF explained in the previous section but there is another physical system under consideration. It is an underlying clean image modeled as piecewise smooth and referred to as prior model I_c so that the minimization of Eq. (15) is subject to the observed image being closely matched to the prior. The estimate of the true image $\hat{I}_c$ is:

$$\hat{I}_c=\min_{I_c} \sum_x \sum_y ||(\nabla I_{\sigma_n}(f))_{x,y}||+\lambda||I_{\sigma_n}(f)-I_c||^2. \tag{20}$$

If the L_1 norm of the vector valued smoothness constraint in Eq. (20) is changed to L_2 norm the ROF total variation technique changes to Tikhonov or Sobolev regularization technique (Karl, 2005):

$$\hat{I}_c = \min_{I_c} \quad \sum_x \sum_y ||(\nabla I_{\sigma_n}(f))_{x,y}||^2 + \lambda ||I_{\sigma_n}(f) - I_c||^2. \tag{21}$$

The parameter $[\lambda : 0 \le \lambda < \infty\}$ is a positive constant referred to as the regularization parameter. This parameter is a weight which balances the level of the piecewise smooth property of the prior image that is impacted to the denoised image and the level of noise removal which is a measure of fidelity $(I_{\sigma_n}(f) - I_c)$ to the prior image I_c. The selection of λ in three different range of values results in denoised images with three different properties:

$$\lambda \quad \Rightarrow \begin{cases} \textbf{TooHigh} & \hat{\mathbf{I}}_\mathbf{c} \Rightarrow \textbf{denoised \& oversmooth} \\ \textbf{TooLow} & \hat{\mathbf{I}}_\mathbf{c} \Rightarrow \textbf{noisy \& piecewise smooth} \\ =0 & \hat{\mathbf{I}}_\mathbf{c} = \mathbf{I_d} \quad \textbf{denoised image same asobservation} \end{cases} . \tag{22}$$

If the value of λ tends to be too high, the algorithm is constrained to place more emphasize on noise removal but the denoised image is oversmooth because there is less emphasis on impacting piecewise smooth property of the prior image on the denoised image. Conversely, value of λ that tends to be too low results in denoised image that possesses the piecewise smooth property of the prior image but retains noise. A special case is when $\lambda = 0$ in which the denoised image is same as the noisy image. Thus, optimal performance of both ROF and Tikhonov techniques is strongly dependent on optimal selection of λ (Galatsanos and Katsaggelos, 1992).

B.4. Classical Double-Layer Markov Random Field Model. For convenience, we restate the minimization formulation in ROF total variation technique expressed as a convex functional in Eq. (20). The subtraction operator is inserted on both sides of the equation:

$$-\hat{I}_c = \min_{I_c} \quad -\left(\sum_x \sum_y ||(\nabla I_{\sigma_n}(f))_{x,y}|| + \lambda ||I_c - I_{\sigma_n}(f)||^2 \right). \tag{23}$$

Application of natural logarithm transformation maintains the convexity of the functional and equality of right-hand side (RHS) and left-hand side terms of Eq. (23):

$$\ln\left[\exp -(\hat{I}_c)\right] = \min_{I_c} \ln\left[\exp -\left(\sum_x \sum_y ||(\nabla I_{\sigma_n}(f))_{x,y}|| + \lambda ||I_c - I_{\sigma_n}(f)||^2 \right)\right]. \tag{24}$$

Now, we separate the exponential function on the RHS of Eq. (24) into product of two exponential functions:

$$\ln\left[\exp -(\hat{I}_c)\right] = \min_{I_c} \ln\left[\exp -\left(\sum_x \sum_y ||(\nabla I_{\sigma_n}(f))_{x,y}|| \right) \exp -\left(\lambda (I_c - I_{\sigma_n}(f))^2 \right)\right]. \tag{25}$$

With reference to Eq. (15), the first term on the RHS of Eq. (25) is the SLMRF energy $U(f)$ expressed as a function of the image pixel configuration f. We insert this term in Eq. (25):

$$\ln\left[\exp -(\hat{I}_c)\right] = \min_{I_c} \quad \ln\left[\exp -(U(f)) \exp -(\lambda (I_c - I_{\sigma_n}(f))^2)\right]. \tag{26}$$

Reformulation of Eq. (26) as strictly exponential function maintains equality of both sides of the equation but the functional changes from a convex functional to a concave functional. The optimization criteria also changes from minimization to maximization;

$$\left[\exp -(\hat{I}_c)\right] = \max_{I_c} \quad \left[\exp -(U(f)) \exp -(\lambda (I_c - I_{\sigma_n}(f))^2)\right]. \tag{27}$$

We know that the regularization parameter is a function of σ. Let us set $(\lambda = 1/2\sigma^2)$:

$$\left[\exp -(\hat{I}_c)\right] = \max_{I_c} \quad \left[(\exp -(U(f))) \left(\exp -\left(\frac{(I_c - I_{\sigma_n}(f))^2}{2\sigma^2} \right) \right) \right]. \tag{28}$$

The Gibbs distribution $P(f)$, the probability distribution of each possible pixel configuration of the prior model is proportional to the first term on the RHS of Eq. (28) (Li, 2009):

$$P(f) = P(I_c(f)) \propto \exp -(U(f)). \tag{29}$$

The Gaussian distribution P_N, the probability distribution of each realization of the observed image I_d from the prior I_c is proportional to the second term on the RHS of Eq. (28):

$$P_N = P(I_d(f)|I_c) \propto \exp -\left(\frac{(I_{\sigma_n}(f) - I_c)^2}{2\sigma^2} \right). \tag{30}$$

Each possible image pixel configuration of the prior and the likelihood are assumed to be independent and identically distributed. Thus, the probabilities $P(I_c)$ and $P(I_d|I_c)$ of the prior and likelihood are expressed as:

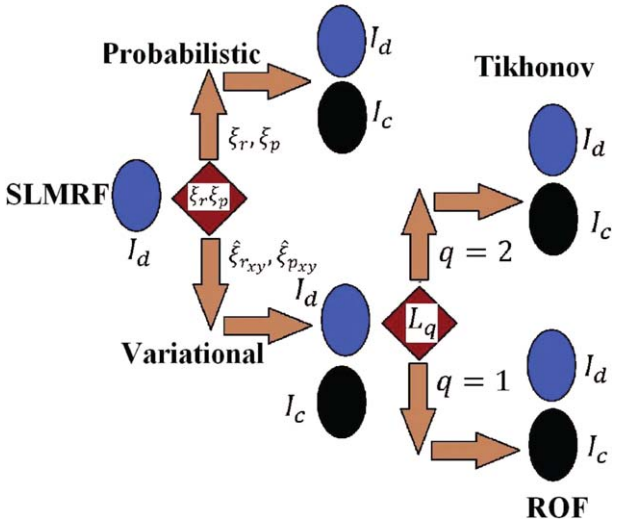

Figure 4. Two paths from our proposed SLMRF model having only the observed image I_d to double-layer Markov random field model having both observed and prior model images I_c. Scalar ξ_r, ξ_p and vector $\hat{\xi}_{r_{xy}}, \hat{\xi}_{p_{xy}}$ values assigned for reward and penalty for conformity and violations of the smoothness constraints configures the image for double-layer probabilistic and variational approaches, respectively. Choice of L_1 and L_2 norm for the image gradient leads to Tikhonov regularization and ROF total variation techniques, respectively. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

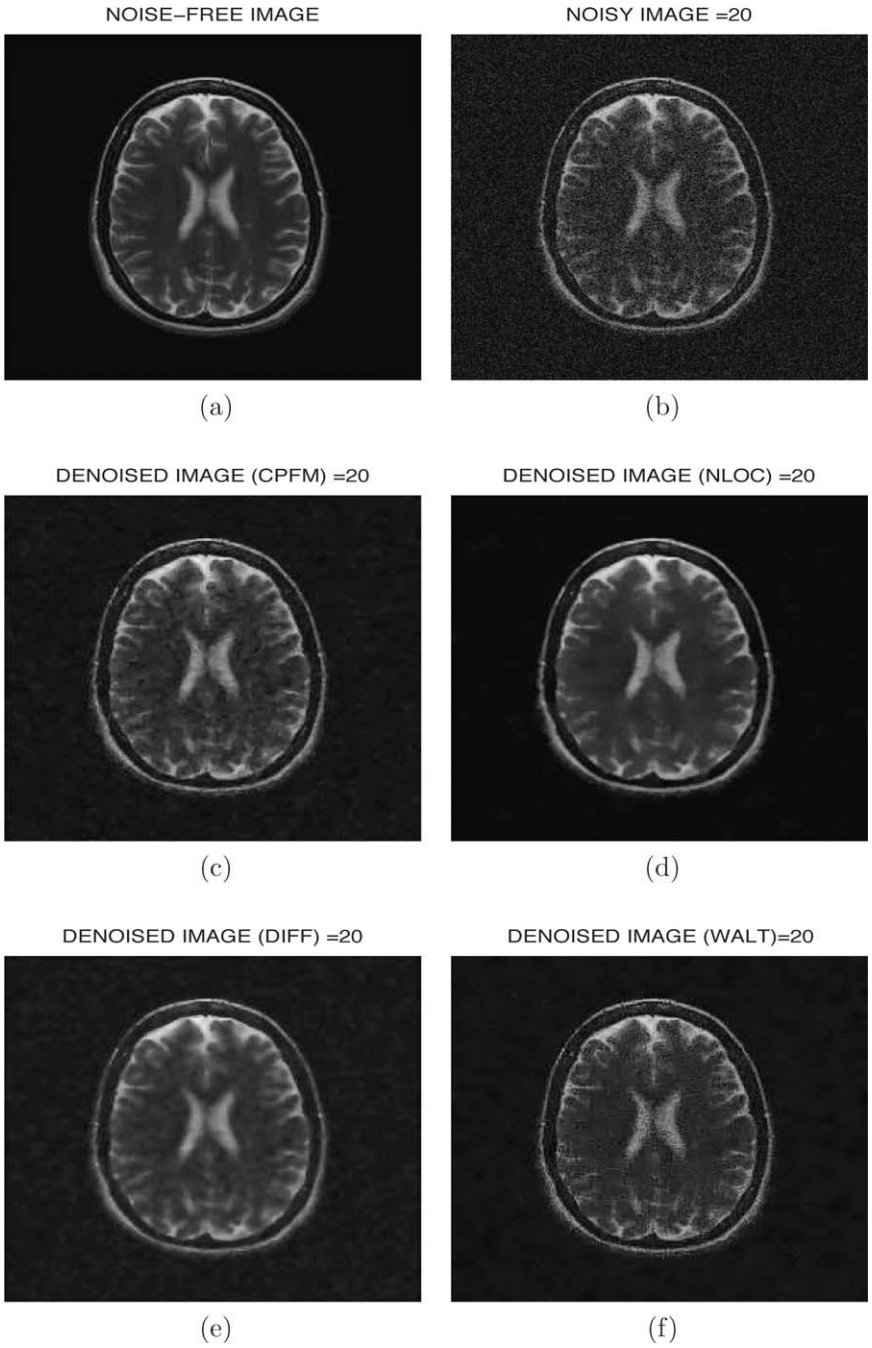

Figure 5. Comparative performance evaluation using MRI slice image from NeuroRx Research Inc. (a) The original and clean MRI. (b) The clean image degraded by noise level of $\sigma = 20$. (c) Denoised version of (b) by **CPFM.** (d) Denoised version of (b) by **NLOC.** (e) Denoised version of (b) by **DIFF.** (f) Denoised version of (b) by **WALT.**

$$P(I_c) \propto \prod_{f \in F} \exp -(U(f)), \tag{31}$$

$$P(I_d|I_c) \propto \prod_{f \in F} \exp -\left(\frac{(I_{\sigma_n}(f)-I_c)^2}{2\sigma^2}\right). \tag{32}$$

Inserting the probabilities in Eqs. (31) and (32) into the RHS of Eq. (28) gives Bayes posterior probability $P(I_c|I_d)$ formula (Mackay, 2003; Li, 2009):

$$P(I_c|I_d) \propto P(I_d|I_c)P(I_c), \tag{33}$$

and we arrive at the Bayesian formulation of the classical Markov random field model for image restoration introduced by Geman and Geman (1984):

$$\begin{aligned}\hat{I_c} &= \arg\max_{I_c} \quad P(I_c|I_d) = \arg\max_{I_c} \quad [P(I_d|I_c)P(I_c)] \\ &= \arg\max_{I_c} \quad \left\{\left[\prod_{f \in F} \exp -\left(\frac{(I_{\sigma_n}(f)-I_c)^2}{2\sigma^2}\right)\right]\left[\prod_{f \in F} \exp -(U(f))\right]\right\}.\end{aligned} \tag{34}$$

This formulation estimates the denoised image $\hat{I}_c$ from the image configuration which maximizes $P(I_c|I_d)$ referred to as maximum a

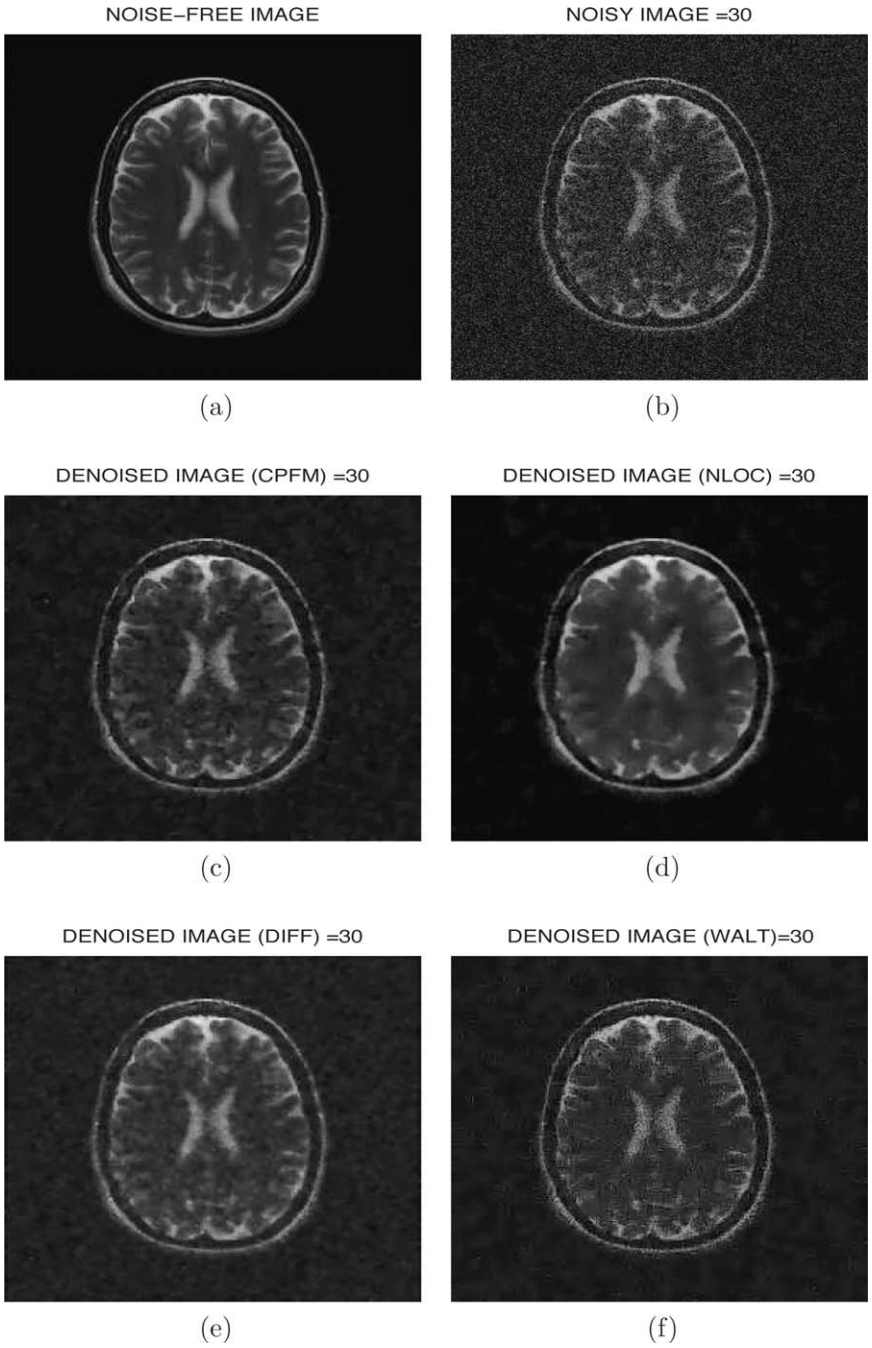

Figure 6. Comparative performance evaluation using MRI slice image from NeuroRx Research Inc. (a) The original and clean MRI. (b) The clean image degraded by noise level of $\sigma = 30$. (c) Denoised version of (b) by **CPFM.** (d) Denoised version of (b) by **NLOC.** (e) Denoised version of (b) by **DIFF.** (f) Denoised version of (b) by **WALT.**

posteriori probability. The probability distributions are functions of the image pixel configurations. The configurations are determined by the level of noise which is a function of the Markov random field energy expressed as the total clique potential computed using scalar valued kernel according to Eq. (11).

B.5. Algorithm Design. The design of the proposed general framework is illustrated by the flow chart in Figure 4. The system has three input parameters. The first is the number of layers ($\mathbf{SL}, \mathbf{DL}$) representing the number of images under consideration. The second is the type of value ($\mathbf{V}, \mathbf{S}$) assigned to ξ_r and ξ_p for conformity and violations of the smoothness constraints. The third parameter is the choice of norm ($\mathbf{q}=\mathbf{1}, \mathbf{2}$) for computing the magnitude of the image gradient. For a user's selection of single-layer **SL**, the denoising system does not make reference to a prior model I_c but a vector valued kernel of Eq. (12) is used for computing the gradient of the observed image I_d. For specific data such as brain MRI having pixels in the background, the energy minimization is controlled by the mathematical model stated in Eq. (14). At each iteration, the Markov random field energy is computed and the stopping criterion is fixed at ($E_s = K_b U_b = -1.5$) based on knowledge of mathematical function

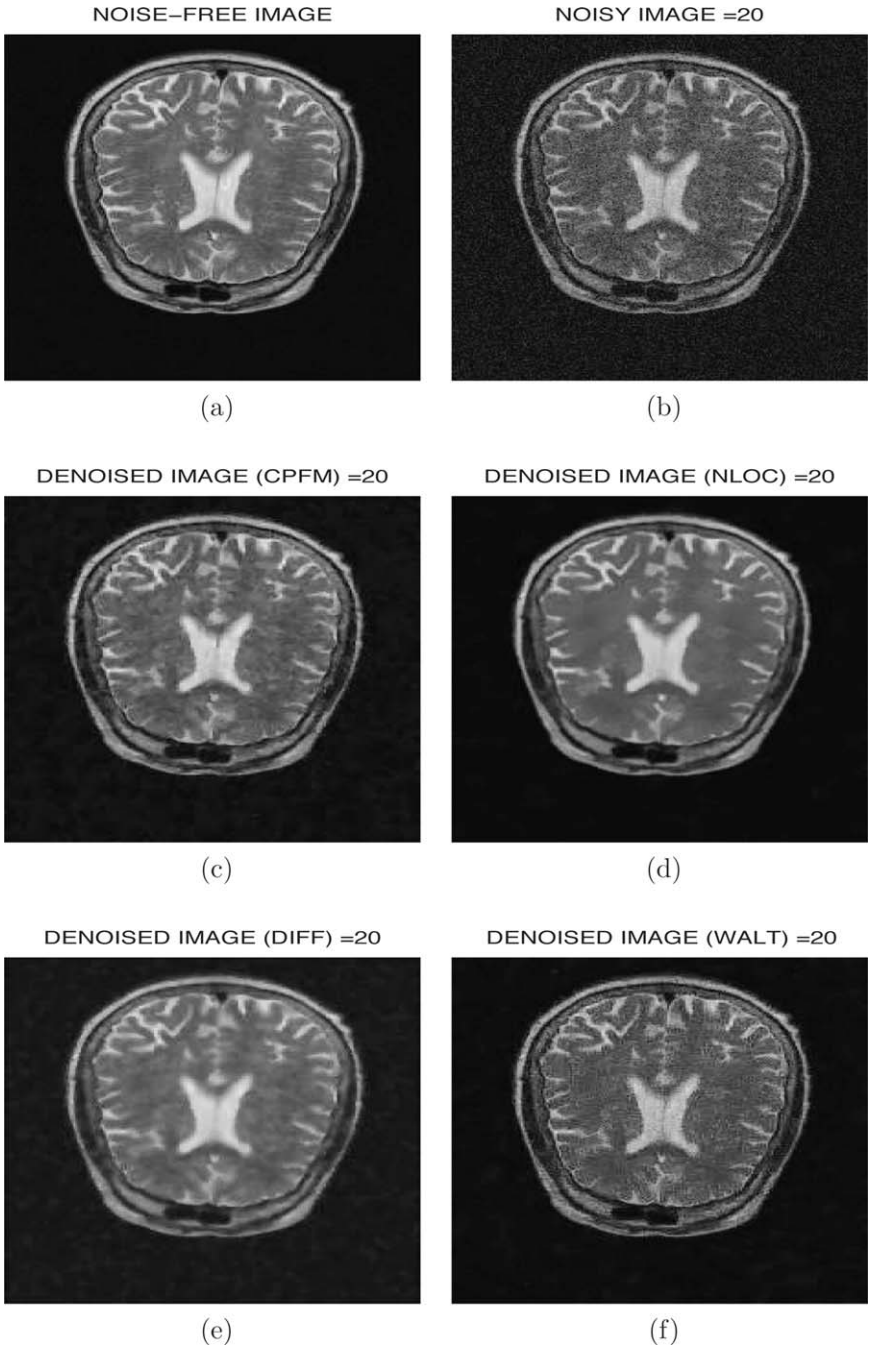

Figure 7. Comparative performance evaluation using MRI slice image from ADNI. (a) The original and clean MRI. (b) The clean image degraded by noise level of $\sigma = 20$. (c) Denoised version of (b) by **CPFM.** (d) Denoised version of (b) by **NLOC.** (e) Denoised version of (b) by **DIFF.** (f) Denoised version of (b) by **WALT.**

describing relationship between MRF energy and noise level (Osadebey et al., 2013).

The choice of double layer (**DL**) introduces an image prior as reference and requires the user to make additional choice of either scalar or vector valued smoothness constraint. With the choice of scalar value the classical Markov random field algorithm is deployed by the system to operate on the image and the energy of the image is computed according to Eq. (11). For the choice of vector valued smoothness, the system prompts the user to a choice of either L_1 or L_2 according to the selection of $\mathbf{q} = 1$ or $\mathbf{q} = 2$, respectively. Either choice deploys either total variation or Tikhonov technique to operate on the image which energy is computed according to Eq. (12).

III. RESULTS

We assume that each MRI slice is noise free, hence for evaluation, we induce artificial noise. For a m percent Rician noise level, the standard deviation of the equivalent normal distribution is given by

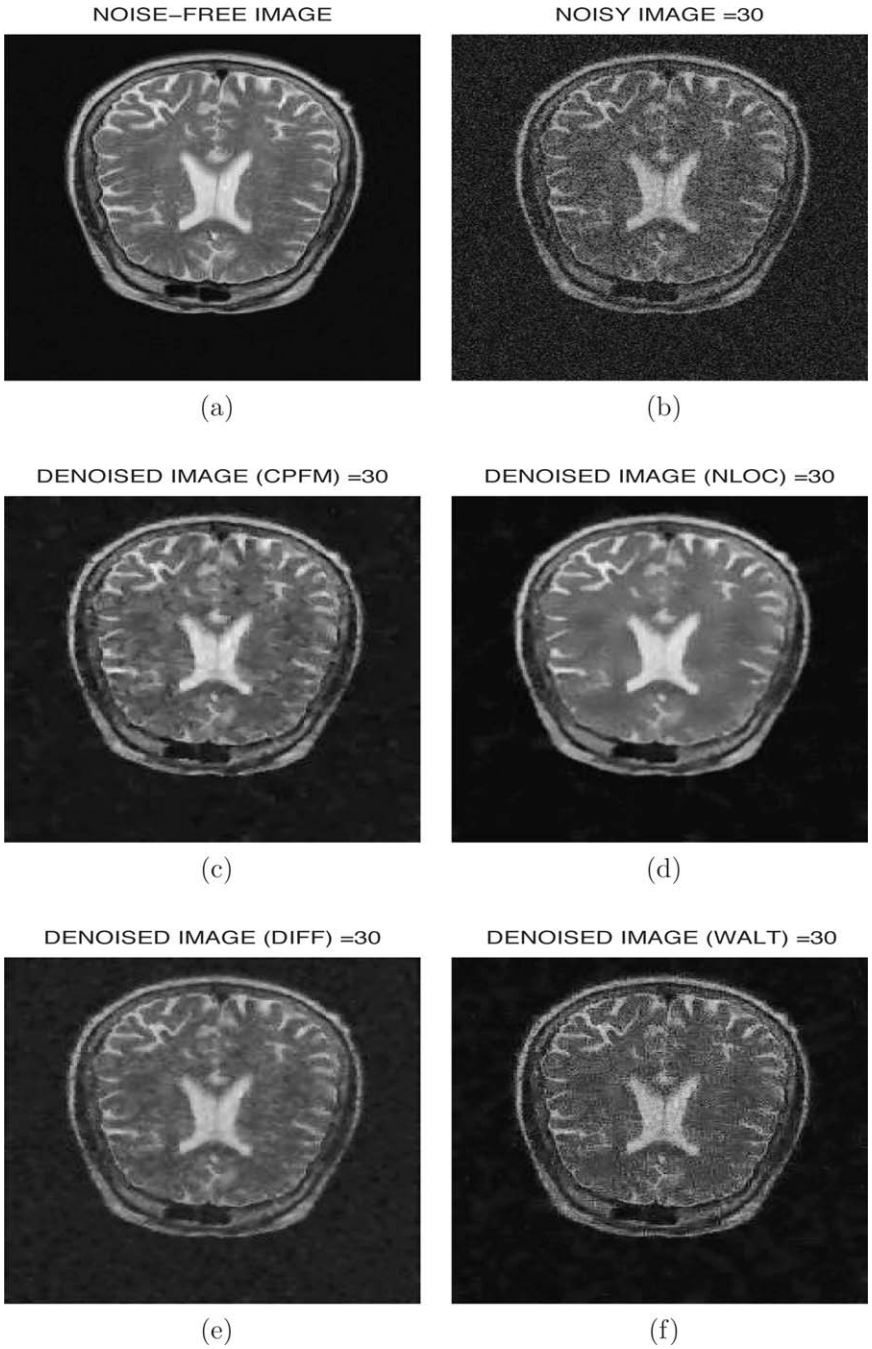

Figure 8. Comparative performance evaluation using MRI slice image from ADNI. (b) The clean image degraded by noise level of $\sigma = 30$. (c) Denoised version of (b) by **CPFM.** (d) Denoised version of (b) by **NLOC.** (e) Denoised version of (b) by **DIFF.** (f) Denoised version of (b) by **WALT.**

$$\sigma \approx N\left(0, \frac{\tau m}{100}\right), \tag{35}$$

where τ is the maximum pixel intensity (Coupe et al., 2010). The algorithms were evaluated using MRI slices from all the four types of MRI acquisitions and Rician noise levels in the range 0% ($\approx \sigma = 0$) to 30% ($\approx \sigma = 75$). However, to satisfy space constraints, we display only T2 weighted images for noise levels of $\sigma = 20$ and $\sigma = 30$.

On the evaluation figure plots that will be discussed in the next section our proposed approach is denoted **CPFM**. The state-of-the-art algorithms are nonlocal means algorithm proposed in Buades et al. (2005), denoted **NLOC**, anisotropic diffusion approach developed in Perona and Malik (1990), denoted **DIFF**, and implemented by D. S Lopes in http://web.ist.utl.pt/daniel.s.lopes/software and wavelet technique as detailed in Pizurica et al. (2003), denoted **WALT**.

For each evaluation parameter, the experimental results were categorized according to the source of MRI data. The images displayed in Figures 5 and 6 are comparative evaluations of the algorithms using NeuroRx data for noise levels of $\sigma = 20$ and $\sigma = 30$, respectively. Corresponding data from ADNI are shown in

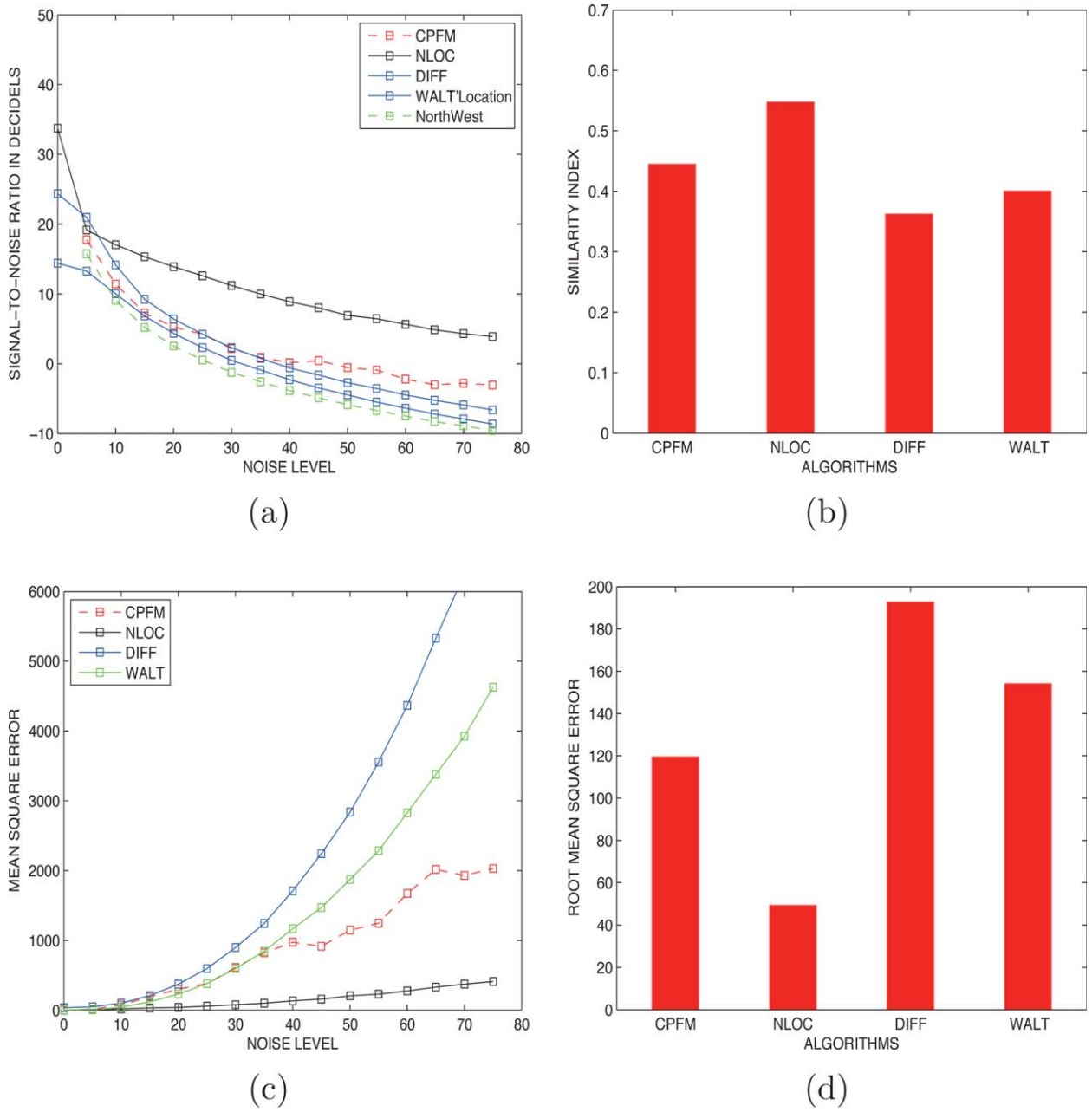

Figure 9. Comparative performance evaluation over the range of noise levels $0 \leq \sigma \leq 75$ using MRI slice image from NeuroRx Research Inc. (a) Signal-to-noise-ratio, (b) Structural similarity index, (c) Mean square error, and (d) Root mean square error. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Figures 7 and 8. Evaluation based on signal-to-noise-ratio, structural similarity index, mean square error, and root mean square error for NeuroRx and ADNI data are shown in Figures 9 and 10, respectively.

IV. DISCUSSION

This discussion section begins with comparative performance evaluation of our proposal with three popular state-of-the-art algorithms. Thereafter we explain the operating characteristics of our proposed algorithm, its limitations and the future work to extend its performance.

A. Comparative Performance Evaluation. The comparative performance evaluation experiment was biased in favor of **NLOC** because noise level, an input parameter for its optimal performance was assumed to be accurately estimated. Evaluation was based on the following five criteria: visual quality, signal-to-noise-ratio, structural similarity index, mean square error, and root mean square error. These criteria are discussed below.

A.1. Visual Quality. The images in Figures 5a and 6a are clean MRI slice images from NeuroRx. Clean MRI slice images from ADNI are displayed in Figures 7a and 8a. The degraded version of these images at noise levels of $\sigma = 20$ and $\sigma = 30$ are displayed in Figures 5b and 6b for NeuroRx data and Figures 7b and 8b for ADNI data. In both cases, the visual quality of the denoised version from our algorithm and the other algorithms can be said to be comparable to each other. Based on visual cues, it can be said that

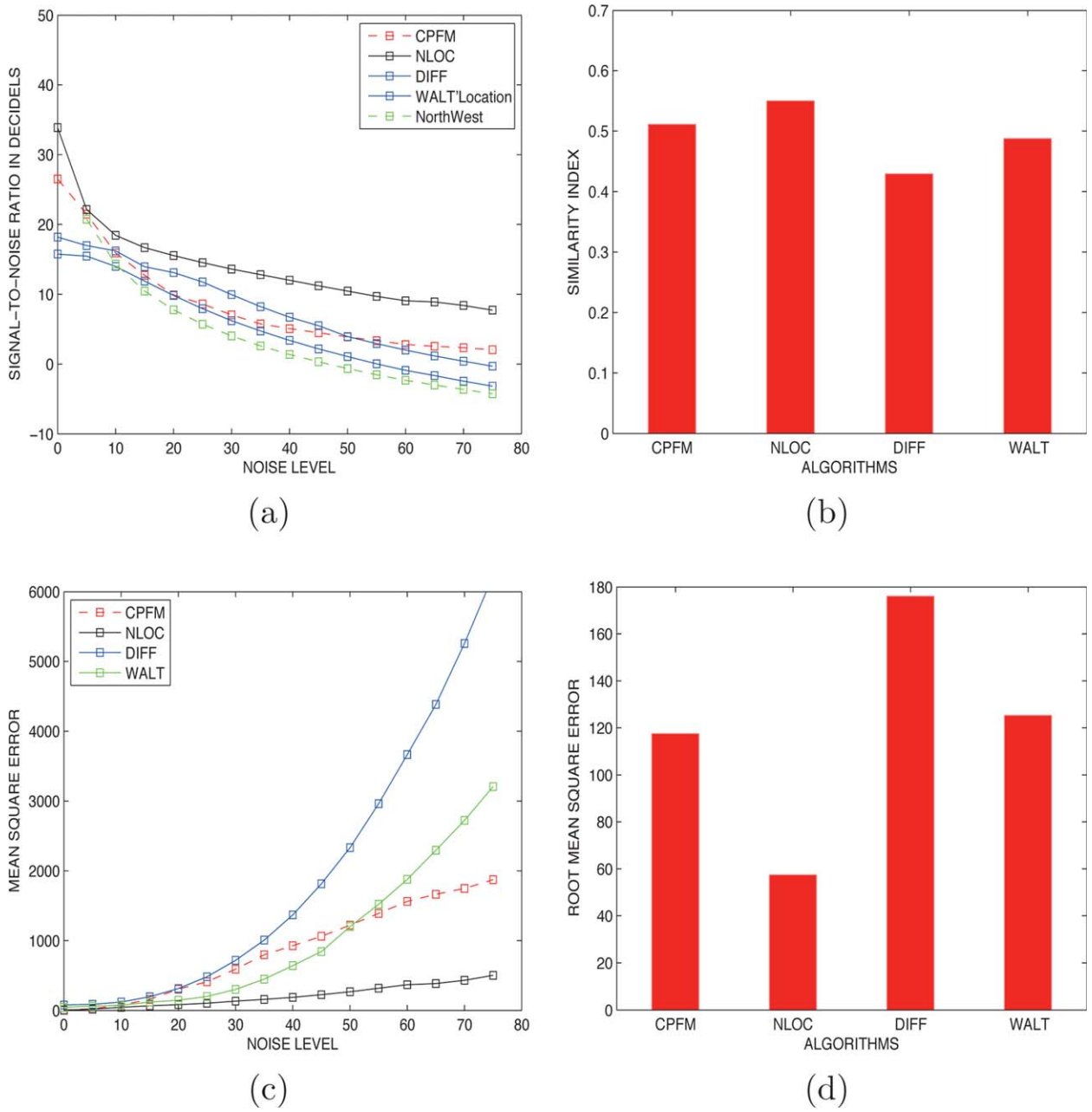

Figure 10. Comparative performance evaluation over the range of noise levels $(0 \leq \sigma \leq 75)$ using MRI slice image from ADNI. (a) Signal-to-noise-ratio, (b) structural similarity index, (c) mean square error, and (d) root mean square error. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

our algorithm and nonlocal means **NLOC** are competing for superiority. At first glance, **NLOC** can be said to be the best based on global view of the images. There is no doubt that **NLOC** performed excellently in removing noise from the background pixels, but the background region of the images are useless in medical diagnostics. A cursory examination of the foreground regions in all the images shows that our algorithm **CPFM** was the best in preserving the original structures of the images. The white and grey matter structures in the MRI images denoised by our algorithm are superior in visual quality because they are much more similar to the original images than the images denoised by **NLOC** which exhibit some degree of blurring.

A.2. SNR. Our proposed algorithm was comparable to the other algorithms in terms of SNR as shown in Figures 9a and 10a. The plot identifying our algorithm is raised above other algorithms but trailing behind **NLOC**. As shown in these figures, for noise level in the range $(0 \leq \sigma \leq 75)$, nonlocal means technique can be distinguished as a superior algorithm.

A.3. Structural Similarity Index. For evaluation based on structural similarity index (see plots in Figures 9b and 10b), all the algorithms are comparable. However, the best performance indicators of 0.57 and 0.55 were recorded by **NLOC**. Our proposal was trailing behind **NLOC** with 0.47 and 0.51 and ahead of the other state-of-the-art algorithms **DIFF** and **WALT**.

A.4. Mean Square Error. The plots in Figures 9c and 10c on mean square error show that the plot identifying our algorithm is lower than other algorithms but higher than **NLOC**, an indication that our proposal has the second best performance indicator after **NLOC** for noise level in the range $0 \leq \sigma \leq 75$.

A.5. Root Mean Square Error. In Figures 9d and 10d, our proposal recorded root mean square error of 120 in both cases, and was trailing behind nonlocal means, which recorded the best performance indicator of 50 and 60, respectively.

A.6. Operating Characteristics of Proposed Algorithm. Our proposed general framework is designed only for 2D medical images. Its operation can be extended to 3D if the individual 2D images that constitute the 3D data are fed separately to the system. The first layer, SLMRF operates on a specific medical image data because no prior model is required but the operation requires knowledge of the relationship between the image Markov random field energy and noise level. For a specific image data, such as brain MRI, the relationship between its Markov random field energy and noise level is described by the mathematical model in Eq. (14). Using a computer having 3 GB RAM and 1.66 GHz processor, it takes, on average, about 10 s to denoise a single MRI slice image. This translates to about 7 min for denoising a patient data consisting of 40 MRI slices. The second layer can operate on any type of medical image data degraded by Gaussian-related noise but its performance is dependent on optimal selection of regularization or model parameter because reference is made to a prior model.

A.7. Limitations of Proposed Algorithm. Operation of the proposed general framework is restricted to Gaussian and Gaussian-related noise such as Rician noise. Performance evaluation reveals that our proposed general framework exhibit acceptable level of performance at different noise levels for all the different modes of MRI acquisitions except T1-weighted MRI. Beyond noise level of $\sigma = 30$ there is significant decline in the quality of denoised T1-weighted MRI images. Operation of SLMRF on brain MRI images is restricted to MRI slice images acquired with or resized to dimension 256×256. This requirement must be satisfied because the mathematical model in Eq. (13) was derived using MRI images of dimension 256×256.

A.8. Future Work. We intend to extend the number of layers in the MRF model from two to three. This technique is an adaptation from the field of texture image segmentation where hierarchical MRF model is widely applied (Li, 2009; Cao et al., 2011). In the three-layer hierarchical MRF model, the image is assumed to be composed of disjoint texture regions besides smooth regions. The texture regions, the underlying piecewise smooth image regions, and the observed image are the first, second, and third layers, respectively. As in the double layer, the texture regions and the underlying image are computed using Bayesian framework. We also intend to formulate mathematical model that describes the relationship between MRF energy and different levels of speckle noise in ultrasound images. This will increase the types of specific medical images and noise in the operation of SLMRF.

V. CONCLUSIONS

In the literature (Li, 2009), different forms of variational techniques are identified as types of Markov random field model for image denoising but without details. In this article, we propose a new general framework that reveals the interconnection between Bayesian and variational techniques for image denoising. The proposed framework also integrates Bayesian and variational techniques into a layer-structured system. Its building block is the four-neighborhood system having clique size of two. Our proposed general framework reveals that the classical Markov random field model can be detached from Bayesian inference belief to create a new simple and computationally efficient noise reduction algorithm which we refer to as SLMRF. In the SLMRF, there is no reference to a prior model image, so the observed image is the only physical system under consideration. The proposed algorithm was applied to reducing noise in magnetic resonance images of the brain and can also be deployed to reduce noise in other medical images where there is knowledge of relationship between the image MRF energy and noise level. Comparative performance evaluation indicates that our proposed algorithm is comparable to popular and state-of-the-art algorithms. It is robust as it demonstrates steady denoising performance for very wide range of noise level $(0 \leq \sigma \leq 75)$. Visual quality assessment shows that it preserves the relevant image structures that are crucial in medical diagnostics and its output is devoid of blurring and ringing artifacts associated with some current algorithms.

ACKNOWLEDGMENT

More experimental results using computer simulated phantoms and real MRI images can be viewed at http://users.encs.concordia.ca/~m_osadeb/MnoiseE_DreductionO/results.htm. ADNI data are disseminated by the Laboratory for NeuroImaging at the University of California, Los Angeles.

REFERENCES

R. Acar and C.R. Vogel, Analysis of bounded variation penalty methods for ill-posed problems, Inverse Problems 10 (1994), 1217–1229.

M. Aksam Iftikhar, A. Jalil, S. Rathore, A. Ali, and M. Hussain, Brain MRI denoizing and segmentation based on improved adaptive nonlocal means, Int J Imaging Syst Technol 23 (2013), 235–248.

M. Auclair-Fortier and D. Ziou, A global approach for solving evolutive heat transfer for image denoising and impainting, IEEE Trans Image Process 15 (2006), 2558–2574.

A. Buades, B. Coll, and J. Morel, A non-local algorithm for image denoising, IEEE Conf on Comput Vision and Pattern Recogn (CVPR05), 2005, pp. 60–65.

Y. Cao, Y. Luo, and S. Yang, Image denoising based on hierarchical markov random field, Pattern Recognit Lett 32 (2011), 368–374.

A. Chambolle, An algorithm for total variation minimization and applications, J Math Imaging Vision 20 (2004), 89–97.

S. Chang, B. Yu, and M. Vetterli, Adaptive wavelet thresholding for image denoising and compression, IEEE Trans Image Process 9 (2000), 1532–1546.

C. Chui, An introduction to wavelets (wavelet analysis and its applications), Academic Press, San Diego, 1992.

P. Coupe, J. Manjon, E. Gedamu, D. Arnold, M. Robles, and D. Collins, Robust Rician noise estimation for MR images, Med Image Anal 14 (2010), 483–493.

P. Coupe, P. Yger, S. Prima, P. Hellier, C. Kervrann, and C. Barillot, An optimized blockwise nonlocal means denoising filter for 3-D magnetic resonance images, IEEE Trans Med Imaging 27 (2008), 425–441.

G. Deng and L. Cahill, An adaptive gaussian filter for noise reduction and edge detection, IEEE Nucl Sci Symp Med Imaging Conf, 1993, pp. 1615–1619.

D. Donoho and I. Johnstone, Ideal spatial adaptation by wavelet shrinkage, Biometrika 81 (1994), 425–455.

D. Donoho and I. Johnstone, Adapting to unknown smoothness via wavelet shrinkage, J Am Stat Assoc 90 (1995a), 1200–1224.

D. Donoho and I. Johnstone, Denoising by soft thresholding, IEEE Trans Inf Theory 41 (1995b), 613–627.

M. Elad, On the origin of the bilateral filter and ways to improve it, IEEE Trans Image Process 11 (2002), 1141–1150.

M. Elad and M. Aharon, Image denoising via sparse and redundant representations over learned dictionaries, IEEE Trans Image Process 15 (2006), 3736–3745.

G. E. Ford and A. I. El-Fallah, Inhomogeneous surface diffusion for image filtering, Int J Imaging Syst Technol 8 (1997), 450–456.

N. Galatsanos and A. Katsaggelos, Methods for choosing the regularization parameter and estimating the noise variance in image restoration and their relation, IEEE Trans Image Process 1 (1992), 322–336.

M. Gaspari, D. Saletti, C. Scandellari, and S. Stecchi, Refining an automatic EDSS scoring expert system for routine clinical use in multiple sclerosis, IEEE J Inf Technol Biomed 13 (2009), 501–511.

S. Geman and D. Geman, Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images, IEEE Trans Pattern Anal Machine Intelligence 6 (1984), 721–741.

G. Gerig, O. Kubler, R. Kikinis, and F. Jolesz, Nonlinear anisotropic filtering of MRI data, IEEE Trans Med Imaging 11 (1992), 221–232.

R. Gold, L. Kappos, D.L. Arnold, A. Bar-Or, G. Giovannoni, K. Selmaj, C. Tornatore, M. Sweetser, M. Yang, S.I. Sheikh, and K.T. Dawson, Placebo-controlled phase 3 study of oral bg-12 for relapsing multiple sclerosis, N Engl J Med 367 (2012), 1098–1107.

A. Hamza, H. Krim, and G. Unal, Unifying probabilistic and variational estimation, IEEE Signal Process Mag 19 (2002), 37–47.

Q. Hu, X. He, and J. Zhou, "Multi-scale edge detection with bilateral filtering in spiral architecture," In Pan-Sydney area workshop on visual information processing, Australian Computer Society Inc., Sydney, 2004, pp. 29–32.

W. Karl, "Regularization in image restoration and reconstruction," In Handbook of image and video processing, A. Bovik (Editor), 2nd ed., Elsevier Academic Press, Burlinton, 2005, pp. 183–202.

B. Kimla and K. Siddiqi, Geometric heat equation and nonlinear diffusion of shapes and images, Comput Vision Image Understanding 64 (1996), 305–322.

D. Koller and N. Friedman, Probabilistic graphical models: Principles and techniques, The MIT Press, Cambridge, Massachusetts, 2009.

C. Lanczos, The variational principles of mechanics, Dover Publications, New York, 1986.

S. Li, "Advances in pattern recognition," In Markov random field modeling in image analysis, S. Singh (Editor), Springer Publishing Company, London, 2009, pp. 13, 26–27, 146–151.

F. Luisier, T. Blu, and M. Unser, A new sure approach to image denoising: Inter-scale orthornomal wavelet thresholding, IEEE Trans Image Process 16 (2007), 593–6067.

W. Ma, Y. You, and M. Kaveh, Monotonically edge-sharpening anisotropic diffusion, J Electron Imaging 21 (2012), 013008-1–013008-9.

D. Mackay, Information theory, inference and learning algorithms, Cambridge University Press, Cambridge, 2003.

J. Manjon, J. Carbonell-Caballero, J. Lull, G. Garcia-Marti, L. Marti-Bonmati, and M. Robles, Mri denoising using non-local means, Med Image Anal 12 (2008), 514–523.

S.M. Monir and M.Y. Siyal, Iterative adaptive spatial filtering for noise-suppression in functional magnetic resonance imaging time-series, Int J Imaging Syst Technol 21 (2011), 260–270.

R. Nowak, Wavelet based Rician noise removal for magnetic resonance images, IEEE Trans Image Process 8 (1999), 1408–1419.

M. Osadebey, N. Bouguila, and D. Arnold (The Alzheimer's Disease Neuroimaging Initiative), The clique potential of markov random field in a random experiment for estimation of noise levels in 2D brain MRI, Int J Imaging Syst Technol 23 (2013), 304–313.

P. Perona and J. Malik, Scale-space and edge detection using anisotropic diffusion, IEEE Trans Pattern Anal Machine Vision 12 (1990), 629–639.

A. Pizurica, W. Philips, I. Lemahieu, and M. Acheroy, A versatile wavelet domain noise filteration technique for medical imaging, IEEE Trans Med Imaging 22 (2003), 323–331.

A. Rajwade, A. Rangarajan, and A. Banerjee, Image denoising using the higher order singular value decomposition, IEEE Trans Pattern Anal Machine Intelligence 35 (2013), 849–862.

L. Rudin, S. Osher, and E. Fatemi, Nonlinear total variation based noise removal algorithms, Physica D 60 (1992), 259–268.

F. Russo, A technique for image denoising based on adaptive piecewise linear filters and automatic parameter tuning, IEEE Conf Instrum Meas Technol, 2005, pp. 1119–1123.

N. Sochen, R. Kimmel, and R. Malladi, A general framework for low level vision, IEEE Trans Image Process 7 (1998), 310–318.

H. Takeda, S. Farsiu, and P. Milafar, Kernel regression for image processing and reconstruction, IEEE Trans Image Process 16 (2007), 349–366.

J. Thomas, Numerical partial difference equations: Partial difference methods, Springer, New York, 1995.

K.R. Thulborn and S.D. Uttecht, Volumetry and topography of the human brain by magnetic resonance, Int J Imaging Syst Technol 11 (2000), 198–208.

A.N. Tikhonov and V.Y. Arsenin, Solutions of ill-posed problems, J Am Math Soc 32 (1978), 1320–1322.

C. Tomasi and R. Manduchi, Bilateral filtering for gray and color images, IEEE Intl Conf Comput Vision, 1998, pp. 839–846.

J. Wei, Lebesgue anisotropic image denoising, Int J Imaging Syst Technol 15 (2005), 64–73.

J. Weickert, Anisotropic diffusion in image processing, B.G Teubner, Stuttgart, 1998.

J. Wen, Y. Li, and W. Wang, Wavelet-based denoising and its impact on analytical spect reconstruction with nonuniform attenuation compensation, Int J Imaging Syst Technol 23 (2013), 36–43.

X. Zhang and M. Desai, Adaptive denoising based on sure risk, IEEE Signal Process Lett 5 (1998), 265–267.