Method to Correct Intensity Inhomogeneity in MR Images for Atherosclerosis Characterization

Olivier Salvado, Claudia Hillenbrand, Shaoxiang Zhang, and David L. Wilson*, *Member, IEEE*

***Abstract*—We are developing methods to characterize atherosclerotic disease in human carotid arteries using multiple MR images having different contrast mechanisms (T1W, T2W, PDW). To enable the use of voxel gray values for interpretation of disease, we created a new method, local entropy minimization with a bicubic spline model (LEMS), to correct the severe (≈80%) intensity inhomogeneity that arises from the surface coil array. This entropy-based method does not require classification and robustly addresses some problems that are more severe than those found in brain imaging, including noise, steep bias field, sensitivity of artery wall voxels to edge artifacts, and signal voids near the artery wall. Validation studies were performed on a synthetic digital phantom with realistic intensity inhomogeneity, a physical phantom roughly mimicking the neck, and patient carotid artery images. We compared LEMS to a modified fuzzy c-means segmentation based method (mAFCM), and a linear filtering method (LINF). Following LEMS correction, skeletal muscles in patient images were relatively isointense across the field of view. In the physical phantom, LEMS reduced the variation in the image to 1.9% and across the vessel wall region to 2.5%, a value which should be sufficient to distinguish plaque tissue types, based on literature measurements. In conclusion, we believe that the correction method shows promise for aiding human and computerized tissue classification from MR signal intensities.**

***Index Terms*—Atherosclerosis, blood vessels, entropy, magnetic resonance imaging, splines.**

I. Introduction

MAGNETIC resonance imaging (MRI) can be used to diagnose and stage blood vessel disease by identifying lumen narrowing in a stenosis, by measuring blood flow, and by imaging plaque in the arterial wall [1]. At our institution, we are focusing on the latter in a comprehensive program that includes new imaging technologies such as intravascular coils, high field MRI, imaging agents, and advanced computer image analysis. Several studies have shown the pertinence of using MR imaging to characterize atherosclerosis lesions *in vivo* [2], [3]. There is good evidence that experts can identify and quantify plaque components such as lipid and fibrous tissue using multiple MR images having different contrast mechanisms [1], [4]. Our eventual goal is to create robust, accurate computer algorithms to perform tissue typing. Our initial application is the analysis of human carotid artery images.

A critical step for analysis methods relying on voxel gray values is the correction of the signal intensities across the MR image. The principal source of the degradation is the spatial inhomogeneity of coil sensitivity of the specially designed surface coils. A correction algorithm for carotid artery imaging faces many challenges. First, the receive coils suffer from a very steep sensitivity fall-off in the direction of increasing tissue depth that is much more significant than the variation across the brain when imaging with a head coil. If not well corrected, this can confound the examination of the vessel wall by experts and defeat automatic tissue classification algorithms. Second, the noise present in the MRI carotid images can disrupt algorithms. Third, there are many voxels close to the artery walls that are void of signal, either from fat suppression or from blood flow compensation. Such voxels do not provide information about the bias field. Fourth, there are relatively large skeletal muscle areas in the neck, near the carotid arteries, that do not include sufficient high spatial frequency content to separate the variations of the sensitivity inhomogeneity from the anatomical structures.

Although the predominant cause of the "shading" artifact in MR images is the sensitivity inhomogeneity of the RF receiver coils, other sources have been described such as inappropriate coil tuning, gradient eddy currents, RF standing wave effects, and RF penetration effects [5]. Simmons *et al.* [6] measured also artifacts induced by different repetition times and echo times.

A widely used model is to lump all the sources in one multiplicative factor: the bias field. The observed MRI signal Y is the product of the true signal X generated by the underlying anatomy and spatially varying field factor B and an additive noise N. At the pixel i, we get

$$y_i = x_i b_i + n_i. \tag{1}$$

Given the observed image Y, the problem is to estimate the true image X. The solution is not trivial since the bias field B is also unknown.

Manuscript received July 18, 2005; revised January 1, 2006. This work was supported in part by the Ohio Wright Center of Innovation under the Biomedical Research and Technology Transfer award "The Biomedical Structure, Functional and Molecular Imaging Enterprise." *Asterisk indicates corresponding author.*

O. Salvado is with the Department of Biomedical Engineering, Case western Reserve University, 10900 Euclid Ave., Cleveland, OH 44122 USA (e-mail: olivier.salvado@case.edu).

C. Hillenbrand was with the Radiology Department, University Hospitals of Cleveland, Cleveland, OH 44106 USA. She is now with Department of Radiological Sciences, St. Jude Children's Research Hospital, Memphis, TN 38105 USA (e-mail: claudia.hillenbrand@stjude.org).

S. Zhang was with the Radiology Department, University Hospitals of Cleveland, Cleveland, OH 44106 USA. He is now with Department of Radiology, Thomas Jefferson University Hospital, Philadelphia, PA 19107 USA (e-mail: Shaoxiong.Zhang@jefferson.edu).

*D. L. Wilson is with the Department of Biomedical Engineering, Case Western Reserve University, Cleveland, OH 44122 USA (e-mail: david.wilson @case.edu).

Digital Object Identifier 10.1109/TMI.2006.871418

When phased array coils are used, the shape of the bias field can be complex and the fall off very steep, as compared to a head cage coil. A low-order polynomial model for B does not describe accurately the intensity spatial inhomogeneity, and higher order polynomial functions can exhibit undesirable behaviors. A mechanical thin plate model can be used, but the optimization becomes computationally heavy. More flexible models for B increase the number of parameters to estimate, necessitating assumptions on X. The most common assumption on X is that there are piece-wise homogeneous regions in the image, and a classification scheme is commonly used. When a steep spatial dependence of the signal intensity and noise is present, classification methods can fail because the optimization is performed across the whole image including areas with low sinal-to-noise ratio (SNR). Most methods reported in the literature have been designed for brain imaging where the intensity inhomogeneity is relatively mild (typically around 20%, almost always $<$40%). There are relatively few validation studies of methods to accurately correct significant intensity inhomogeneity. Below, we review existing methods, especially with regard to carotid artery imaging and the special problems it presents.

Some correction techniques rely on measuring the coil sensitivity function using a physical phantom [7], [8] or model it using Biot and Savart's law [9]. These sensitivity functions can then be used to correct subsequent *in vivo* images. These approaches can correct for coil sensitivity but they do not correct for other sources of intensity inhomogeneity, as identified above. Another approach is to obtain an additional MR image from a body coil having a uniform sensitivity response. Dividing the image from the phased array coils by this image obtained with uniform sensitivity, one can obtain the bias field [10], [11]. Similar approximation is often used to estimate the sensitivity matrix in parallel imaging [12]. In the context of vessel wall imaging, the extra body coil images add significant imaging time to acquire and/or are too noisy to be used without additional processing.

Many image processing methods have been proposed to estimate the bias field directly from the image Y. To separate the underlying "true" image, X, from the bias field B, assumptions can be made on X and/or on B. We shall now focus our discussion on those assumptions and seek the most relevant ones for vessel wall imaging.

A key observation is that the bias field is smooth compared to a typical MR image and most, if not all, methods rely on this fact. The smoothness of B can be constrained in the frequency domain assuming that its spectrum contains low frequencies that do not significantly overlap with the relatively "high frequency" image spectrum. This method is very attractive since no other assumption on X or B is necessary and well-known low-pass filtering schemes can be implemented. One drawback is the edge artifacts that can result, but algorithms have been proposed to mitigate these effects [13]–[16]. For example, Vorkura *et al.* proposed a scheme based on integrating well-chosen derivatives of the image [16]. When phased array coils are used, the sensitivity inhomogeneity gradient can be very steep, and the requirement for nonoverlapping spectra may not be met. In addition, signal voids near the vessel wall challenge any method to reduce edge artifacts. Nevertheless, we considered low pass filtering techniques in our research because of their simplicity and robustness.

Attempts have been made to separate B and X in the image domain by constraining B to be smooth using a model. Proposed models, in order of increasing expressiveness, are: polynomial functions [17]–[20], discrete cosine transform [21], splines [22]–[25], thin plates like constrained membrane [26], [27], and smoothed residual [28]–[32]. As more degrees of freedom are added, the methods become more computationally demanding and local minima in the optimization can become problematic. Because the bias field obtained with phased array coils can vary as much as 80% across the image (as measured by us and others [33]) and are summed of multiple coils, its shape cannot be accurately modeled with a low-order polynomial. Higher order polynomial functions when fitted to strongly varying bias field can lead to spurious artifacts. When a high-order polynomial function is fitted to data that are not evenly spread across space, the polynomial can fit well in regions of closely packed data points but vary wildly in other regions where data are scarce. Better models are a bicubic spline or thin plate membrane. We used the latter in recent publications [27], [34], but to reduce the computational burden while maintaining expressiveness, we use bicubic spline in this paper.

Assumptions on X have also been used. Often one makes the reasonable assumption that the image consists of a finite number of tissue types and that piece-wise homogeneous regions are present. Classifications schemes can then be used that embed a model for B. Dawant *et al.* [20] manually segmented recognizable tissues; Wells [28] extended the framework of maximum likelihood classification to include a bias field, and this innovation led to many refinements [21], [23] [29]–[32]. Fuzzy c-means has also been modified to include a bias field [26], [35]; other tissue classification techniques have also been applied [18]–[20], [22]. All of these techniques determine clusters in feature space. Due to the very high intensity inhomogeneity present with phased array coils, clusters are spread over large regions in feature space, thus increasing the sensitivity to initial conditions and local minima. We have recently published a modification of the fuzzy c-means algorithm proposed by Pham et Prince [26], that circumvent these problems [27], [34]. We achieved good results, but some difficult cases lead us to develop the new method that will be quantitatively compared to fuzzy c-means.

Some researchers do not segment the image into classes; instead, they maximize the information content in X. Sled *et al.* [36] describe a deconvolution algorithm whereas Likar *et al.* [37] and Mangin *et al.* [25] propose an entropy minimization technique. The first method has become popular since minimal assumptions are required, and it achieves good results. However, the bias field is assumed to be normally distributed which is not the case when phased array coils are used. The two latter methods are also very attractive since no classification of the image is necessary. Likar *et al.* model the bias field with a polynomial function, which is insufficiently expressive for phased array coil imaging. Mangin *et al.* use a spline model for the bias field and the optimization is done with a stochastic annealing algorithm. In all three methods, optimization is performed over the entire dataset, regardless of SNR variation.

It is our experience that carotid artery images obtained with surface array coils challenge existing methods. Linear filtering methods gave significant edge artifacts in the vessel wall because of the step edge created by a lumen void of signal. With methods based on histogram modeling, such as Fuzzy c-means, we experienced convergence problems, especially in noisy images, and very long computation times. In the case of the method by Likar *et al.*, with our implementation we experienced convergence issues, and the corrected images often had visible inhomogeneity, probably because the optimization was trapped in a local minimum.

We propose a method that builds upon work in the literature and addresses issues with the carotid artery images. To avoid the drawbacks of classification schemes, we minimize entropy. To avoid the corruption from low SNR regions, we optimize entropy locally, starting with high SNR areas and merging areas with lower SNR in a sequential fashion. So as to be able to control expressiveness and account for the steep signal inhomogeneity, we use a bicubic spline model for B, which enables one to control smoothness with proper knot spacing. We carefully evaluate algorithms using synthetic, phantom, and patient images, and experimental methods and results are given in Sections III and IV. We next describe the new algorithm.

II. Algorithms

Below, we first give an overview of the algorithm. This is followed by an illustrative, one-dimensional (1-D) example in Section II-B. This is followed by many details of the implementation (Section II-C). We then describe how we choose algorithm parameters in Section II-D. In the last subsection, we briefly describe some other algorithms to which we will compare results.

A. Overview

We first identify all tissue voxels and filter the image to reduce noise. We fit a fourth-order polynomial function to the tissue voxels so as to provide a rough initial estimate of the bias field, $B0$. Air voxels in the background are excluded because they are void of signal. For the refined description, we model the bias field, B, as a bicubic spline with a rectangular grid of knots evenly spaced across the image. The spacing of knots is important: knots should be sufficiently close to ensure that the bias field can be adequately expressed and far enough apart that the estimated surface will not contain anatomical structures in the images. Spacing is related to the receiver coil geometry and is optimized in experiments described later. We use the values of the initial polynomial estimate of the bias field at the knot locations to initialize the bicubic spline bias field.

We now describe the piecewise optimization process which makes the use of a bicubic spline model tractable. We identify the knot $k1$ having the highest corresponding $B0$ value and begin optimization there. The signal from the coil at this location is high and the high local SNR ensures that we will get a good local estimate of B. For simplicity, we describe this knot as having the highest SNR, although this is not strictly true from (1), which gives $\mathrm{SNR} = BX/N$. The neighborhood of this knot, Ω_{k1}, is bordered by its eight neighboring knots in such a way that neighborhoods overlap. We optimize a single parameter, the amplitude at the knot location so as to minimize the entropy of the corrected image, $X = Y/B$, within Ω_{k1}. The knot $k2$ having the next highest SNR in its neighborhood, Ω_{k2}, is now identified. The entropy of X in the region $\Omega_{k1} \cup \Omega_{k2}$ is now optimized by adjusting the amplitude at $k2$. Next, the region with the third highest SNR region is determined, and again the knot amplitude is optimized so as to minimize the entropy within the three merged regions. The process continues until all knots have been optimized. Additional passes are performed, starting again with the region having the highest SNR and proceeding in the same piecewise manner. The process stops when either the knots or the image entropy do not change significantly, or the maximum number of passes is reached. Note that since the bias field is a relatively smooth function, there are typically no "holes" in the merged regions. We call this method LEMS for Local Entropy Minimization with bicubic Spline model.

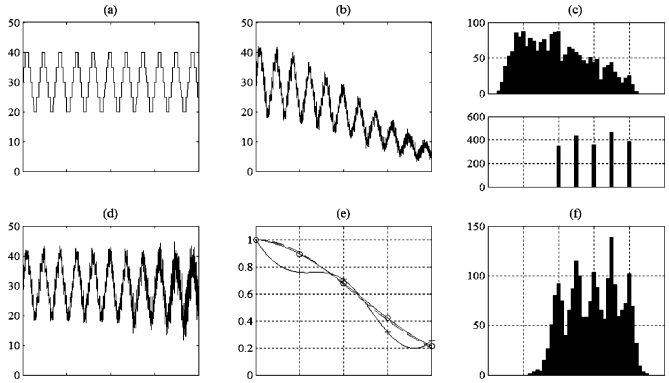


Fig. 1. One-dimensional example. The original signal X with five classes is shown in (a) and its histogram in the bottom of (c). A Gaussian bias field multiplies X to simulate the effect of the intensity inhomogeneity B of a RF coil located on the left of the data, and normal centered Gaussian noise with standard deviation 1 (SNR ranges from 40/1 on the left to 5/1 on the right) is added. The resulting signal $Y = BX + N$ is shown in (b) with its histogram in (c) top. (d) Corrected data, (e) initial bias field $B0$ with its knots marked as "+," the estimated bias field after two passes with its knots marked with "o", and the true bias in dashed line. (f) Represents the histogram of the corrected data showing the strong overlapping of the classes under noisy condition.

B. One-Dimensional Example

A 1-D illustrative example is shown in Fig. 1. The data X have been generated by digitizing a sine wave such that five classes are evenly spread across the y-axis. A Gaussian bias multiplies X to simulate the effect of the sensitivity inhomogeneity B of a RF coil located on the left of the data. Gaussian noise is added. Fig. 1 shows the result after two passes of the bias field correction algorithm. The corrected data [Fig. 1(d)] are very close to the true data [Fig. 1(a)] while the bias field estimated matches almost perfectly the true bias [Fig. 1(e)]. The five original classes are clearly seen as five peaks in the final histogram even when the classes overlap in the presence of noise. In this simple example the knots are optimized from left to right, since assuming an additive noise to the data (1), the SNR is higher where the RF coil sensitivity is higher, i.e., on the left of the graphs. This simulation exemplified the good behavior of the algorithm under noisy conditions. In Fig. 1(e), the estimated bias field is closer (as measured by percent error) to the true one at the left of the graph where the signal amplitude, and SNR, is highest. This property is exploited in our algorithm, as described previously.

C. Implementation Details

1) Tissue Segmentation: Image processing is used to automatically identify all the tissue voxels in the neck. First, we identify all the air voxels outside the neck. We use region growing with "regional" seeds at the top corners of the image, where air is always present. From the seed regions, we estimate the mean, μ_b, and the standard deviation, σ_b, of the air background. Connected pixels are included in the background if their value is less than $\mu_{\rm b} + 3\sigma_{\rm b}$. The algorithm stops when no more pixels fulfill this criterion.

Second, a fuzzy membership mask image, M, is created with a label of one for tissue voxels, zero for signal voids (e.g., air or flow suppressed voxels), and a value between zero and one for partial volume voxels. Criteria for these selections are listed below, where σ_b is an updated estimate following region growing

$$\begin{cases} y_i < \mu_b + a\sigma_b, & m_i = 0, \text{ background voxels} \\ \mu_b + a\sigma_b < y_i < \mu_b + b\sigma_b, & m_i = \frac{y_i - \mu_b + a\sigma_b}{(b-a)\sigma_b} \\ \mu_b + b\sigma_b < y_i, & m_i = 1, \text{ tissue voxels} \end{cases}. \tag{2}$$

The parameters a and b, are manually adjusted depending on the level of noise in the image. For most of the images that we tested, we used $a = 2$ and $b = 4$. For very noisy images we used $a = 1$ and $b = 2$.

Third, the fuzzy membership mask M is cleaned so as to remove incorrectly identified background voxels. M is thresholded to yield a binary mask for $M > 0$. The resulting binary image is processed using opening to remove small islands of pixels and closing to remove small holes. The largest connected component (tissue voxels in the neck) forms a binary mask. By multiplying M with this binary mask, the background is cleaned.

We use M in subsequent processing. We use only tissue voxels $(M = 1)$ in the calculation of the entropy for estimating the bias field. Signal void voxels are not used because there is no information about the bias field. When correcting the image with the estimated bias field, the signal at each voxel x_i will be corrected using

$$x_i = \frac{m_i y_i}{b_i} + (1 - m_i) y_i \tag{3}$$

which accounts for the partial volume effect

The noise standard deviation in the image is estimated from the noise in the external air background as obtained with region growing. When no signal is present, the noise has a Rayleigh probability distribution function [38], and the standard deviation of the background voxels σ_b is related to the standard deviation of the noise $\sigma_{\rm n}$, as $\sigma_b = \sigma_n \times 0.655$. Hereafter, we use this estimate for σ_n.

2) Noise Reduction Filtering: Noise reduction filtering is done using anisotropic diffusion [39], as modified by Black *et al.* [40]. The equations we use for filtering are given below where $B0$ is the initial bias field estimate

$$\frac{\partial Y}{\partial t} = -\nabla(g(\|\nabla Y - \nabla B0\|)(\nabla Y - \nabla B0)) \tag{4}$$

$$\text{where } g(x) = \begin{cases} \frac{\left[1 - \left(\frac{x}{\sigma_{AD}}\right)^2\right]^2}{2}, & |\mathrm{x}| \le \sigma_{AD} \\ 0, & \text{otherwise} \end{cases}. \tag{5}$$

We use $\sigma_{AD} = 2\sigma_n\sqrt{2}$ and nine iterations; other details are described in previous publications [27], [34]. We used the method by Black *et al.* which includes the bias field gradient because it reduces artifacts in the filtered data. And, as compared to the original method, it had no measurable effect on the bias field estimation.

3) Bias Field Model: A two-dimensional (2-D) polynomial function of order Np is used to provide an initial estimate of the bias field, $B0$

$$B0 = \sum_{n=0}^{Np} \sum_{m=0}^{Np-n} \hat{\theta}_{m,n} x^m y^n. \tag{6}$$

The polynomial function is fitted in a least square sense to the tissue pixels using a classic regression technique.

To improve upon the description of the bias field, we use a bicubic spline model [41] as implemented in the Matlab spline toolbox (The MathWorks, Natick, MA). To initialize this estimate, we take the value of $B0$ at the knot locations. After each optimization step, we divide the bias field B by a constant so as to ensure that its mean over the tissue voxels is 0.5. This ensures that the brightness and histogram of the corrected image is stable.

4) Entropy Optimization: The bicubic spline estimate of the bias field is optimized so as to minimize the entropy of the image using the piecewise, optimize and merge algorithm described previously. Entropy is given below where $PDF_X(l)$ is the probability density function of X, which is approximated by the histogram of the voxels in the area being optimized, divided by the number of voxels

$$H = -\sum_{l \in \{\text{gray levels}\}} \mathrm{PDF}_X(l) \log\left[\mathrm{PDF}_X(l)\right]. \tag{7}$$

Optimization is done using a golden section search and parabolic interpolation [42], as implemented in the Matlab optimization toolbox (The MathWorks Inc, Natick, MA). The probability density function in (7) is approximated at each step in the process from the gray value histogram from the corrected image, X, using a binning resolution of half a gray level value. The optimization of some knots requires special attention. If the knot region does not include at least 300 tissue voxels, the knot is left unchanged. This can occur at the border of an image as well as in large regions of air such as in the trachea. The value of 300 was experimentally determined to give a reasonable number of knots that can be modified while providing a statistically significant entropy measure. Very little difference was obtained with a value of 200, indicating that the algorithm is not overly sensitive to this parameter.

5) Algorithm Parameters: There are relatively few free parameters in the algorithm. A principal one is the knot spacing. On the one hand, knot spacing should be small so as to describe

the shape of the bias field. On the other hand, the distance between the knots should be greater than the size of anatomical structures. Else, when entropy is minimized, the estimated bias field will inappropriately contain anatomical structure. In addition, as one decreases knot spacing, the computational load increases. The relative position and geometry of each receiver coil affects the shape of the bias field; hence, it is more appropriate to measure the knot spacing in millimeters than voxels. We experimentally determined that bias field estimation was not overly sensitive to knot spacing; a value of 21 mm worked well in all cases for the receiver coils that we use. For tissue segmentation, we set $(a, b) = (2, 4)$ when the SNR was high (typically PDW and T1W data) and $(a, b) = (1, 2)$ for low SNR datasets (typically T2W). For filtering, we used $\sigma_{AD} = 2\sigma_n\sqrt{2}$ and nine iterations. For the initial estimate of the bias field, we used a 4th order, 2-D polynomial function. We used a stopping criterion of $<0.1\%$ entropy changes for optimization, in most of the cases the optimization stops upon reaching a fixed maximum number of iteration, typically four. For actual patient data, images were cropped to remove aliasing artifact and air region. For a typical patient dataset with a knot spacing of 21 mm and an in-plane resolution of 0.51 mm $\times$ 0.51 mm, the final number of knots was 4 rows by 8 columns, or 32 knots. To compute the bias field with the cubic spline, we used the variational or natural bicubic spline interpolant where the second derivatives at border knots are set to zero [43].

6) Other Algorithms: We compared the LEMS method to other algorithms in the literature. One was a modified version of the adaptive fuzzy c-means originally proposed by Pham and Prince [26] that we recently described [27], [34]. Briefly, the method models the bias field as an elastic thin plate membrane, which is closely related to our bicubic spline model. The method seeks to identify tissue types (typically four or five in our data set), defined by their mean intensities. It is an iterative algorithm that minimizes the functional

$$E_{\mathrm{mAFCM}} = \sum_{j\in\Omega}\left(u_{jb}^{p}\|y_j-\mu_n\|^2+\sum_{k=1}^{N_c}u_{jk}^{p}\|y_j-b_jv_k\|^2+uo_j\delta^2\right) +\alpha\sum_{j\in\Omega}\sum_{r=1}^{M}(D_r\otimes b)_j^2+\beta\sum_{j\in\Omega}\sum_{r=1}^{M}\sum_{s=1}^{M}(D_r\otimes D_s\otimes b)_j^2 \quad (8)$$

where y is the signal measured, u_{jk} is the membership function for the voxel j to the class k, N_c is the number of classes, b is the bias field, and v_k is the gray-scale center of cluster k. The first term of the LHS is the background class identified by thresholding. To make the original algorithm more robust to noise and voxels subject to partial volume effect, we have added an outlier class in the form of the third term of the LHS. The second line of the LHS describes an elastic membrane and a thin plate membrane terms that constrain the bias to be smooth. D_r and D_s are convolution kernels to compute the first and the second derivative. Unknowns u, v and b are optimized sequentially using zero-gradient conditions until class centers converge to stable values [34]. Hereafter, we call this method the modified fuzzy c-means segmentation based method (mAFCM). Parameters were chosen so as to give the best performances across the testing sets. When the maximum number of iterations (300) was reached, most often a bad solution was obtained with obvious misclassifications. Typically, the algorithm converged within 100 iterations to a good solution. The relatively stringent criterion for convergence was that the Euclidian distance between consecutive class centers did not change more than 0.01%. Parameters α and β were set to 5×10^5 and 5×10^4 respectively; p was set to 2; and initial class centers were spread linearly between the background and the maximal intensity.

A second method is a widely used linear filtering method described by Murakami *et al.* [11]. This method separates the signal X from the bias field on the frequency domain using linear filtering, and it will be called LINF. The method divides an image without bias field I_{BC} acquired with a body coil, by the same image acquired by the phased array coil I_{PA}. Prior to division only the low frequencies are kept: $B = \mathrm{LPF}(I_{\mathrm{PA}})/\mathrm{LPF}(I_{\mathrm{BC}})$, where $\mathrm{LPF}(\cdot)$ is a low-pass filter implemented by convolving with a Gaussian kernel with a standard deviation of 13.5 mm, which was manually determined as giving the best performance. I_{BC} is approximated by a binary image from the tissue segmentation step, and I_{PA} is the observed data Y.

III. Experimental Methods

A. Two-Dimensional Synthetic Phantom

A 256 $\times$ 256 image was generated with three tissue classes having gray values of 50, 100 and 150. Features in the image were created to mimic tissues distributed throughout a human neck. This image was the ground truth, X_{true}. MR scans were obtained of a homogeneous, saline-filled, cylindrical phantom about the size of a human neck. These images were acquired using the phased array surface coils and the resulting image, after normalizing the maximum value to one, gave the true bias field (B_{true}). The final synthetic image was obtained by multiplying X_{true} by B_{true}. Voxels were set to a low value of 20 to simulate air at four disks inside and in the region outside the pseudo neck. To add noise, the synthetic image was Fourier transformed, summed with a normally distributed noise on both real and imaginary parts, and inverse Fourier transformed. SNR was measured as the ratio of the image noise standard deviation to the highest class intensity (150) in the image domain. For every test case, ten noise realizations were generated, and results were displayed with means and standard deviation error bars. The corrected image and the estimated bias field were compared to X_{true} and B_{true}, respectively. Since we were most interested in the correction of the vessel wall, we quantified the correction by measuring the maximal signal variation in the corrected image, X, within a five pixel annulus surrounding each of the four lumens. This vessel wall error was expressed as a percentage of X_{true} : $100[\max((X - X_{\mathrm{true}})/X_{\mathrm{true}}) - \min((X - X_{\mathrm{true}})/X_{\mathrm{true}})]$. We also computed the global error across the entire image using the formula $100 \times \mathrm{STD}(X - X_{\mathrm{true}})/\mathrm{MEAN}(X_{\mathrm{true}})$, where $\mathrm{STD}(\cdot)$ and $\mathrm{MEAN}(\cdot)$ are the standard deviation and mean, respectively. The vessel wall error measures the maximum signal variation in a small area, whereas the global error measures the average variation across the corrected image.

B. Physical Phantom

We created a physical phantom so as to test the entire process of imaging and bias field correction. Four aqueous solutions with different concentrations of agar (0.5%, 1.0% 2.0% and 4.0%) and 10 mM of $CuSO_4$ were used to simulate physiological T1 and T2 values. Next, 15-mm plastic tubes were filled with the gel solution and placed into a cylindrical plastic container about the size of the human neck, which was then filled with a 10 mM $CuSO_4$ aqueous solution. Small empty plastic tubes (8 mm diameter) were added to simulate artery lumens. A tube filled with sesame oil was used to evaluate the fat suppression technique. The phantom was imaged with the same T2W sequence used for human imaging, as described in the next subsection.

The physical phantom lends itself to quantitative evaluation. After bias field correction, one expects to find the same MR signal value for each tube containing the same solution. To evaluate this, we manually segmented the tubes and computed for each tube the mean and standard deviation of its gray values. Statistical t-tests were subsequently performed to test the null hypothesis that the signal intensity of each tube was equal to the mean intensity of all the tubes of the same type, at the confidence level of 0.01. Since the ground truth image is not known, a *global region of interest (ROI) error* was computed by measuring the signal variations of each solution's ROIs across the image: $100 \times [\max(\text{ROIs}) - \min(\text{ROIs})]/\text{mean}(\text{ROIs})$. We excluded some outlier ROIs, which had very low signal intensity before correction and which are identified in the legend of Fig. 6.

C. Actual Patient Data Images

Sixteen patients with carotid artery stenosis, as documented by duplex ultrasound, were recruited for the study. Informed consent was obtained from all subjects under a protocol approved by the institutional review board for human investigation. All MR scans were conducted on a 1.5 T system (Magnetom Sonata; Siemens, Erlangen, Germany) with a custom-built phased array coil. Dark blood images were obtained using ECG-triggered double inversion recovery (DIR) turbo spin echo sequences. Imaging parameters (TR/TE/TI/NSA/thickness/FOV) were as follows: T1W: 1R-R/7.1 ms/500 ms/2/3 mm/13 cm; PDW: 2R-R/7.1 ms/600 ms/2/ 3 mm/13 cm; T2W: 2R-R/68 ms/600 ms/2/3 mm/13 cm. Fat saturation was applied. The in plane resolution was $0.51 \times 0.51\ \text{mm}^2$. Two phase array receiver coils were used. Each one is made of two overlapping rectangular loops (53 mm $\times$ 60 mm). The overall dimensions of each array are 60 mm $\times$ 103 mm. The two coil arrays are positioned on each side of the patient's neck and held in place with Velcro straps. Images were the magnitude of the sum of the squares of the four receiver coils. Images were cropped by removing the bottom 25% of the image corresponding to the posterior of the neck where low SNR and aliasing artifacts from the anterior portion of the image are found. Image quality was rated as poor, average, or good by an experienced reader. Poor quality image volumes with motion artifacts or very low SNR were removed from the study. This left 11 patient studies with a total of 288 images. Grouping images by SNR, there were 32%, 36%, 21%, and 11%, of images with SNRs near 10, 20, 30, and 40, respectively. SNR was computed as the ratio of the average sternocleidomastoid muscles signal to the noise standard deviation as estimated from air.

To quantify our ability to correct patient images, we assumed that skeletal muscle should have constant signal values across the neck. We manually created ROIs around the left and right sternocleidomastoid muscles as well as the left and right deep neck muscles. For each correction algorithm, mean and standard deviations were determined and used to compute a *global ROI error* in a manner similar to that for the physical phantom. We also measured the corrected signal percent difference between the left and right sternocleidomastoid muscles by computing the absolute difference between ROI means and dividing by the average. Box and whisker plot are used, where the box has lines at the lower quartile, median, and upper quartile values. Whiskers are lines extending from each end of the box to show the extent of the data $+/-150\%$ from the box. Outliers are data with values beyond the ends of the whiskers.

To further demonstrate the quality of the method, multiple corrected patient images (T1W, T2W, and PDW) were segmented using a maximum-likelihood classification technique. Both PDW and T1W histograms were modeled as a sum of three Gaussian distributions to represent the main tissues found in the neck (skeletal muscles, fat suppressed tissues, and connective tissue and bones); a Rayleigh distribution to model the background containing noise; and an uniformly distributed outliers class to encompass pixels subject to partial voluming, sparse high intensity normal tissues, and pathological tissues. This combination was found to fit the histograms. For T2W images, only two Gaussian tissue classes were used. Parameters for the distributions were estimated using the expectation-maximization algorithm [44]. Crisp segmentation was performed by labeling each pixel with its most likely tissue class. Resulting segmented image were visually checked for isointense skeletal muscles.

IV. Results

A. Results From the Synthetic Phantom Experiments

Fig. 2 shows the LEMS method applied to the synthetic phantom. To illustrate the algorithm, Fig. 2(a)–(f) are displayed following the optimization of the second knot in the second pass of optimization. The histogram clearly shows three tissue classes. The profile has been flattened, and there is no visible shading artifact on the corrected image in Fig. 2(b). Typically, we would run two more optimization passes before getting the final result.

Fig. 3 shows how algorithm parameters were optimized for LEMS. Fig. 3(a) shows vessel wall error as a function of the order of the 2-D polynomial function when it is fit to the bias field measured on a homogeneous physical phantom and not patient data. Error decreases with increasing order. A fourth order, 2-D polynomial function with 15 parameters gave an error of 3.2%, showing that a relatively high-order is needed to well describe the bias surface. In Fig. 3(b), vessel wall error following LEMS correction of the synthetic image data is plotted as a function of knot spacing. Results begin to deteriorate with a knot

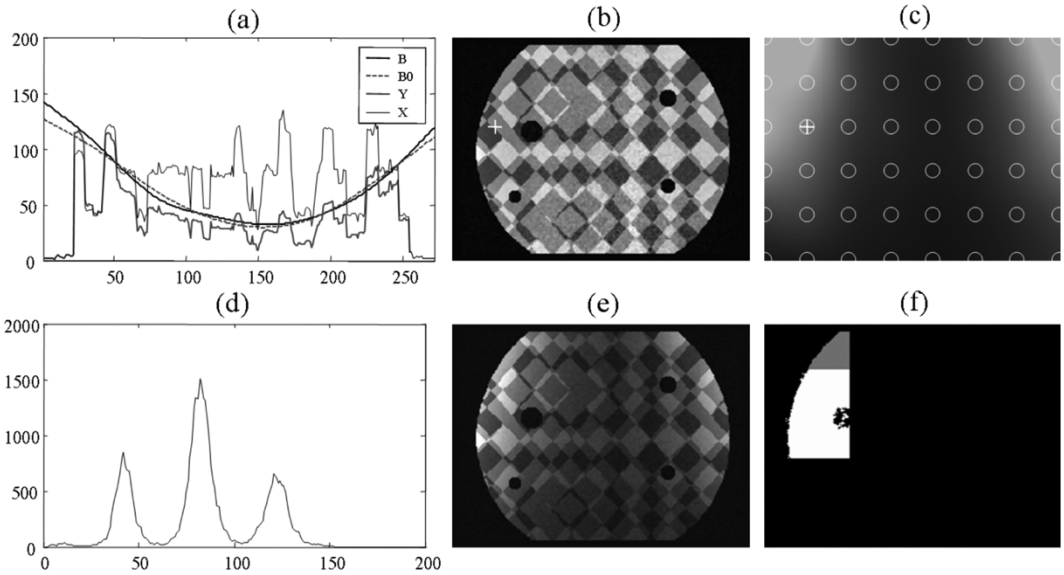


Fig. 2. Application of the LEMS correction method to the synthetic phantom. All results are shown at the beginning of the second pass so as to better illustrate the algorithm. The original image Y, in (e), is corrected by the current bias estimate in (c), which also includes a graphic overlay showing knot locations marked with an "∘" and the current knot being optimized marked with a "+". The current corrected image X is shown in (b). (f) Shows the areas where the entropy is computed at this time step; the new area corresponds to the current knot being optimized in white while the gray area shows the optimized area from the previous knot optimization. Both areas are merged to compute the entropy of the corrected image X in (b). The histogram in (d) shows the distinct peaks for the three classes, and these will continue to "sharpen" with subsequent iterations giving improved entropy, whereas the profiles (a) show the refinement of the original bias estimates ($B0$). The SNR for this simulation was 15 for the class having the highest gray value.

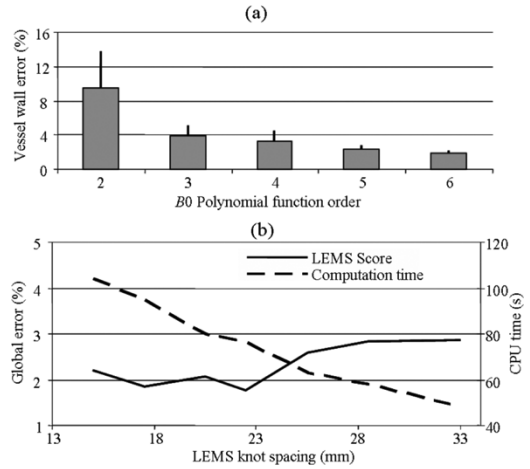
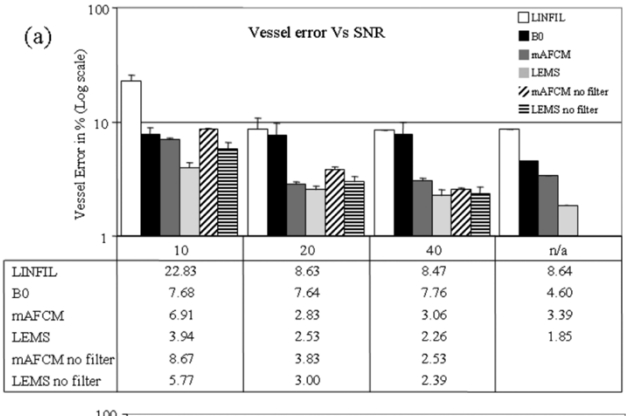


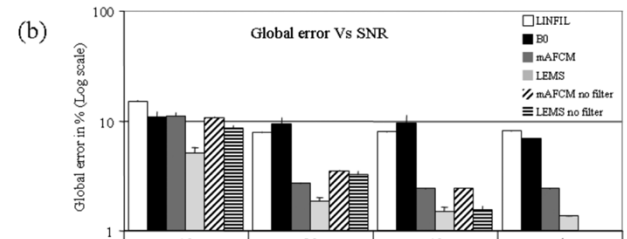
Fig. 3. Optimization of parameters in the LEMS algorithm. In (a), vessel wall error (in %), as defined in Section III, is plotted as a function of the order of the polynomial for the initial estimation of the bias field when fitted to the true bias field. This represents the best achievable bias modeling by a polynomial function. Error computed over the entire image reduces in a similar fashion with increasing order of the polynomial. In (b), lumen error (left axis in %) and computation time (right axis in seconds) are plotted for the LEMS method as a function of knot spacing. The image had a size of 192 × 256 voxels; the left most case of 15-mm knot spacing represents 63 knots, whereas the right most case with a knot spacing of 33 mm had 12 knots. Knot spacing has relatively little effect on error over a range from 13 mm to 24 mm. Computation time is for four passes of the optimization algorithm. Code is written in nonoptimized Matlab code and run on a 2.4-GHz Pentium IV; hence, absolute times are less important than the trends observed as one increases the number of knots.

spacing >25 mm. There is little difference over the range of 15 to 25 mm, indicating the relative insensitivity of the algorithm. Other tests on the physical phantom (not shown) lead us to consider a knot spacing of about 21 mm for the phased array coils used in the neck images. Computation time decreases as the number of knots decrease, as shown in Fig. 3. Code is written in nonoptimized Matlab code and run on a 2.4–GHz Pentium IV; hence, absolute times are less important than the trends observed as one increases the number of knots. Experiments in Fig. 3(a) and (b) were done on the same data, allowing us to compare a polynomial function to that of a bicubic spline. The bicubic spline optimized with LEMS achieves an error comparable to a fifth-order 2-D polynomial function having 21 parameters. In this experiment, the 2-D polynomial is fitted to a homogeneous phantom. When it is used on an anatomical MR image, its behavior is significantly degraded.

In Fig. 4(a), vessel wall errors are presented for corrections of the synthetic phantom with LEMS, LINF, and mAFCM, as well as correction with $B0$. Image SNR values vary across the horizontal axis of the graph. LEMS outperforms the other methods in every case, giving a vessel wall error of 4.13% or less for SNRs exceeding 10. This value after correction compares very favorably with the 40% vessel wall error *before* correction. For reference, the worst case SNR for the T2W patient images is about 10, whereas a worst case of 20 is more typical for T1W and PDW datasets. Across the entire field of view, LEMS outperformed the other methods and gave 2% global error, or better, for SNRs exceeding 10 [Fig. 4(b)]. Again, this is a remarkable improvement to the 44% before correction. Note that the global and vessel wall errors are not comparable because they are computed in different ways. Fig. 4 also shows the influence of the filtering step: errors are significantly decreased when image data are prefiltered, especially at low SNR. This improvement is true for LEMS but also for mAFCM. Typical processing times for



	10	20	40	n/a
LINFIL	22.83	8.63	8.47	8.64
B0	7.68	7.64	7.76	4.60
mAFCM	6.91	2.83	3.06	3.39
LEMS	3.94	2.53	2.26	1.85
mAFCM no filter	8.67	3.83	2.53	
LEMS no filter	5.77	3.00	2.39	



	10	20	40	n/a
LINFIL	15.03	7.78	7.88	8.06
B0	10.79	9.44	9.61	6.81
mAFCM	11.09	2.71	2.40	2.39
LEMS	5.15	1.88	1.50	1.36
mAFCM no filter	10.74	3.49	2.39	
LEMS no filter	8.58	3.26	1.54	

Fig. 4. Comparison of bias correction methods applied to the synthetic phantom. Vessel wall error (a) and Global error (b) are plotted on a logarithm scale for the three different bias correction methods, as well as for the initial bias $B0$. Errors are also reported for mAFCM and LEMS with no prefiltering. Values are obtained as a function of decreasing SNR for the class 150. The right most value, labeled N/A is the noise free case.

LINFIL, mAFCM and LEMS were 0.2 s, 130 s, and 70 s respectively for the 256×256 synthetic phantom. Our implementation of LINFIL uses optimized executable filtering functions from Matlab, whereas the two others are written with much slower Matlab scripts.

B. Validation Results With the Physical Phantom

We evaluated the imaging and correction algorithms using the physical phantom consisting of vials containing four different solutions providing four different "tissue" classes spread throughout the phantom. Fig. 5 shows the result of the LEMS correction. All the tubes filled with the same solution are isointense across the image as seen in the images as well as the horizontal profiles [Fig. 5(h)]. Distinct peaks arising from the different solutions are clearly seen on the histogram of the corrected image [Fig. 5(e)]. The peak with the highest intensity on the histogram corresponds to the doped water filling between the tubes. Two peaks overlap significantly around intensity level 600/650. The LEMS method achieved close to optimal performance in this experiment. The bottom of the corrected image shows a residual intensity shading artifact, where the measured signal intensity was the weakest. Thanks to the ordered, piecewise, merged region optimization method, these problematic areas do not significantly impact other high SNR regions. In our experience, other global optimization methods, such as mAFCM, can suffer from the problem areas.

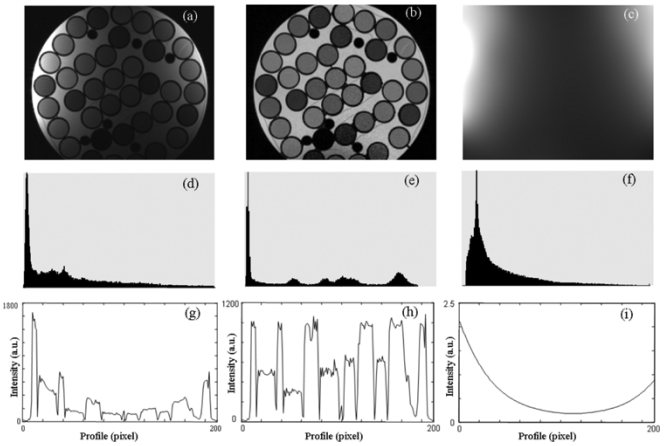


Fig. 5. Correction of physical phantom using the LEMS algorithm. Shown are (a), (b), and (c) images, (d), (e), and (f) histograms, and (g), (h), and (i) horizontal profiles from the middle of the images, for the original phantom [(a), (d), and (g)], the corrected image [(b), (e), and (h)], and the estimated bias field [(c), (f), and (i)]. After correction, (h) the profile is clearly leveled, and the four class values for the tubes as well as the class corresponding to the doped water filling are clearly seen on, despite the fact that two solutions overlap significantly in (e) the histogram. The faint curves surrounding the vials in the phantom are from pieces of tape used to construct the phantom. Algorithm parameters are given in Section II-D.

The performance of the three bias correction methods on the physical phantom images are compared in Fig. 6. LEMS is clearly superior to LINF and mAFCM, even in this case where the phantom consists of piecewise homogeneous areas from a known number of classes, a situation which should favor mAFCM. Global ROI errors for LEMS were 6.7%, 2.4%, 1.8%, and 3.5% for the 0.5%, 1.0%, 2.0%, and 4.0% solutions, respectively. Comparing the three method LEMS, mAFCM, and LINF, the global ROI error averaged over the four solutions were 3.6%, 18%, and 12.5%, respectively.

C. Patient Carotid Artery Images

PDW MR images before and after LEMS correction are shown for a healthy volunteer in Fig. 7. The corrected image is visually free of intensity inhomogeneity artifacts. Skeletal muscles are isointense across the neck and the carotid arteries are much more visible than in the original scan. Arteries are also isointense with respect to the skeletal muscles as expected from previous study [45].

The method was applied to PDW, T1W, and T2W MR patient images (Figs. 8–11). The LEMS method works well on the T1W and PDW images, but it works less well on very noisy data typical of T2W image. In Fig. 8, we determined that the bias field estimated from the PDW image can be used to correct the T2W image. Results are visually and quantitatively improved using this correction scheme. The average on the 10 slices of the signal variation between the two muscles reduces to 3% using the PDW estimate, from 5.71% using the T2W data.

We quantitatively evaluated correction methods by measuring their ability to make the skeletal muscles isointense across the patient images (Fig. 9). For all 288 patient images, the signal intensity variations between left and right sternocleidomastoid muscles are plotted as a function of the image SNR. The LEMS method performs best, and it is the only method improving upon

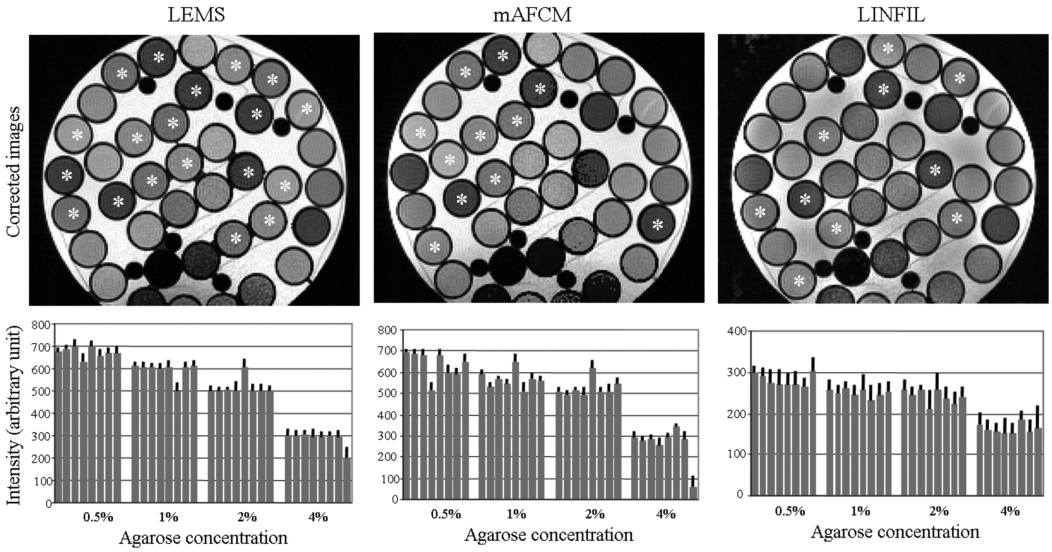


Fig. 6. Quantitative comparison of the three bias correction methods (LEMS, mAFCM, and LINF) applied to images of the physical phantom. The three methods are shown in the left, middle, and right columns, respectively. Bar graphs along the bottom row are plots of the means and standard deviations of gray values for ROIs within the vials. LEMS clearly creates the flattest image and bar graph values having the least variation within each of the four classes. (The class corresponding to water and the background are not shown.) To test the ability to discriminate classes, we applied a t-test to the data values in the bar graphs. Stars indicate ROIs for which the null hypothesis (ROI equals to mean of ROIs) cannot be rejected at the 0.01 significance level. LEMS has the most stars indicating good identification of the classes. Even in the case of LEMS, the correction at the bottom of the image is degraded because the original signal level is low here. This sub-optimal correction gives rise to outlier measurements in the graph which are omitted in the calculation of the global ROI error.

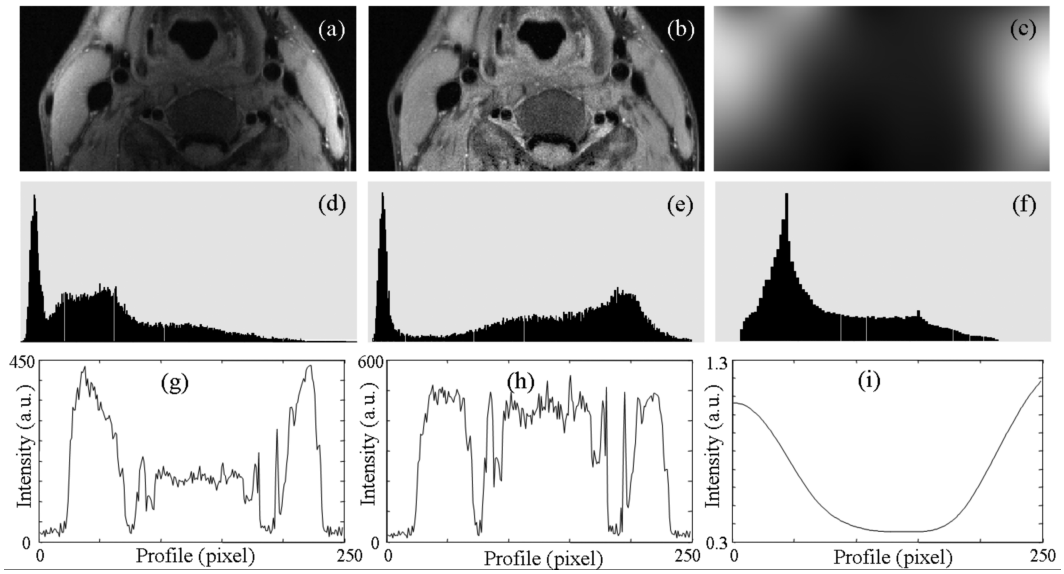


Fig. 7. Correction of a PDW MR scan of patient $B003$ with the LEMS method. Shown are images [(a), (b), and (c)], histograms [(d), (e), (f)], and horizontal profiles [(g), (h), and (i)] from the middle of the image, for the original image [(a), (d), and (g)], the corrected image [(b), (e), and (h)], and the estimated bias field [(c), (f), and (i)] after four passes of the algorithm. After correction, skeletal muscles are isointense across the image. The profile has been chosen to cross the carotid arteries on both sides. On this healthy volunteer the signal of the artery walls is isointense with the skeletal muscles as described in the literature.

the $B0$ estimate at the lowest SNR of 10. The LEMS method also gives fewer outliers at SNR values of 30 and 40.

We segmented tissues in the neck using a standard maximum-likelihood tissue classification algorithm, as described in Methods. Since we were interested in studying the influence of SNR and sequence type we did not use multi-spectral classification but classified images independently, moreover registration problems would have added another source of variability. Classification was applied to all images (288) for each of the correction methods ($B0$, LINFIL, mAFCM, and LINFIL). We visually checked crisp segmentation results as well as the quality of histogram fitting. Best results were found following LEMS preprocessing. A typical example is shown in Fig. 10, where all skeletal muscles belong to the same class (the highest intensity class of the three tissue classes). In almost all images, pathological tissues, probably corresponding to atheromateous lesions, were classified as outliers since they are hyperintense compared to skeletal muscles in PDW images

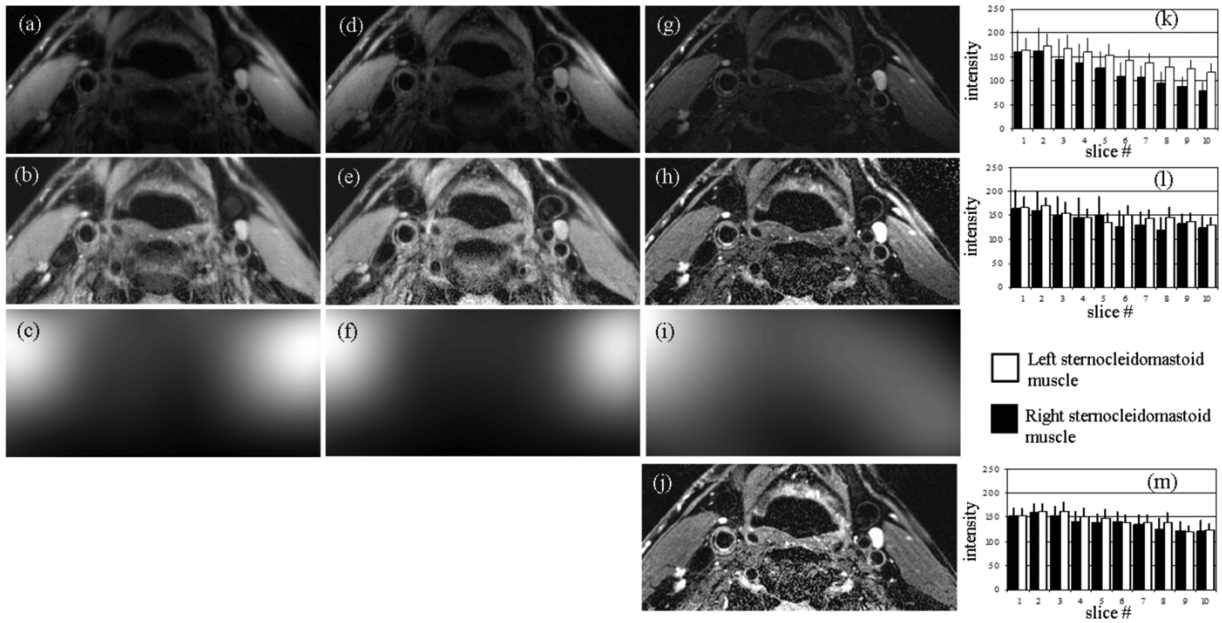


Fig. 8. Correction of patient A016 scans using the LEMS algorithm. The left, middle, and right columns of images correspond to PDW [(a), (b), (c)], T1W [(d), (e), (f)], and T2W [(g), (h), (i)]. Original images are shown on the top [(a), (d), (g)], corrected images with LEMS are shown in the middle row [(b), (e), (h)], whereas the bottom rows show the estimated bias field [(c), (f), (i)]. Bias fields estimated for PDW and T1W are very close [(c) and (f)], but the bias field for T2W (i) is different, probably because of the low SNR in this T2W data. The corrected image for T2W (h) shows intensity artifact in the top right corner of the image. (j) Original T2W image (g) corrected with the bias field estimated from the PDW data (c). Intensity inhomogeneity is visually reduced and quantitative evaluation shows improved correction: means and standard deviations of ROI on the sternocleidomastoid muscles on each side of the patient neck are plotted for the ten slices of the original data (k), the corrected T2W (l), and the T2W corrected with the bias field estimated from PDW (m).

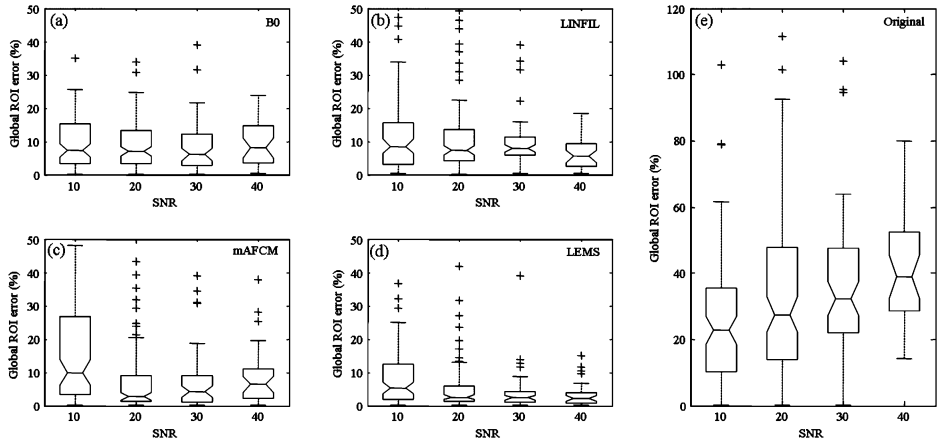


Fig. 9. Performance comparison on patient images. Shown are global errors after correction with $B0$ (a), LINFIL (b), mAFCM (c), and LEMS (d) as well as the global errors from the original data (e). Each boxplot represents the global errors measured by computing the signal variation between the right and left sternocleidomastoid muscles versus SNR. The total number of images is 288 from 10 patients and includes PDW, T1W and PDW in about equal proportion.

[46]. When intensity inhomogeneity was not corrected, classification yielded irregular results as shown in Fig. 10(c). Finally, Fig. 11 shows a typical example of patient images presenting atherosclerotic lesions.

V. Discussion

The new LEMS method outperformed the other methods that we tested for human carotid artery studies. For SNR of 20 global errors of 44% before correction were reduced to 1.9%, 2.7%, and 7.8%, for LEMS, mAFCM, and LINF, respectively, for the synthetic phantom. LEMS was relatively insensitive to noise, giving very good results (global error of 5.1%) with an SNR of 10, a comparable value than the noisy patient T2W images. Global ROI errors in the patient images, as measured from skeletal muscle ROIs, were also vastly improved following LEMS correction, having a value of less than 10% for most of the images when the error before correction could be as high as 95%. Of course our primary concern is that the vessel wall is well corrected, and the synthetic phantom provides a means for measuring this. Vessel wall errors for an SNR of 10 were improved from the 40% before correction to 3.9%, 6.9%, and 22.8%, for LEMS, mAFCM, and LINF, respectively.

There is evidence that the LEMS algorithm performs sufficiently well to enable the discrimination of atherosclerotic tissue types. For example, Clarke *et al.* [47] showed that about 25% signal intensity variation can be expected between fibrous tissue

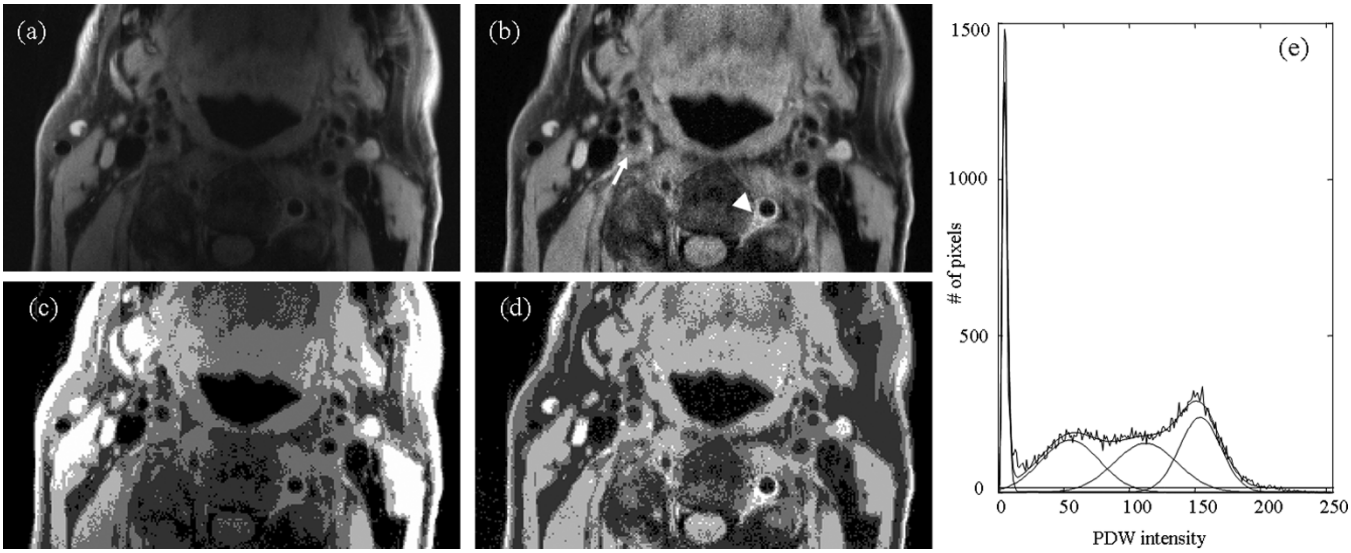


Fig. 10. Example of tissue classification. The original PDW image (a) was corrected with LEMS (b) and segmented by fitting the histogram (e) with three Gaussian distributions to model three tissues, a Rayleigh distribution to model the noise in the background, and a uniform distribution to identify the outliers (see text for details). A crisp classification is shown in (d) were the three tissue classes are displayed with three shades of gray, the background in black, and the outliers in white. Note that all the skeleton muscles belong to the same class across the image. The arrow shows a lesion in the right internal carotid artery of this patient that is much more visible in the corrected image. Note how outliers identify hyperintense pixels relative to muscular tissue in the lesion as well as in the left cervical artery that could have been missed in the original data (arrowhead). (c) Shows the classification result from the original image without correction: the segmentation is meaningless because of the strong intensity inhomogeneity.

and necrotic core in PDW images. Others studies measured similar variations in PDW images and also T1W and T2W images [4], [48]. It is encouraging that these values exceed the residual "vessel wall" error of 2.6% following correction of the synthetic phantom images. Moreover, our ultimate goal is to perform vector-based classification for tissue typing. Since we will use prior probabilities and multiple measurements simultaneously, discrimination of tissue types, in the presence of noise and residual inhomogeneity errors, should be improved. We have already shown that simple tissue classification algorithms can segment corrected images into meaningful tissue classes. In summary, we believe that corrections with the LEMS method are sufficient to proceed with evaluation of vector based classification studies.

The LEMS method was inspired by two algorithms in the literature. First, Likar *et al.* [37] used second- and fourth-degree polynomial functions having 18 and 34 free parameters, respectively, in three-dimension, and minimized the entropy over the entire image at once. Because of the severity of the signal inhomogeneity obtained with our phased array coils, a high-order ($\geq$4) polynomial function is required to describe the bias field. In our experience, high-order polynomial functions can give spurious artifacts when fitted to a strongly varying bias field. So instead of a polynomial function, we used a bicubic spline that can control such artifacts. Another great advantage of the spline is that expressiveness of the surface can be easily adjusted by varying the knot spacing. Second, Shattuck *et al.* [22] proposed a method to *locally optimize* a bias field described by a tricubic B-spline model that smoothes the residual of the difference between the classified and the original image. They used a parametric model of the histogram, with the limitations reviewed in Section I, rather than the minimum entropy used in LEMS. Since optimizing all knot values at once is computationally intensive, prone to local minima, and sensitive to regions of low signal intensity, we used a piecewise optimization approach similar to that of Shattuck *et al.* However, our method starts with regions having high SNR and proceeds by merging lower SNR areas.

We found that an edge-preserving prefiltering step increases the accuracy of the bias field estimation for LEMS and mAFCM. Since edge sharpening filtering like anisotropic diffusion acts to sharpen the histogram by reducing the number of outliers due to noise and partial volume effect, it is reasonable to expect that any method relying on histogram modeling (such as mAFCM) or histogram-based entropy measurement (such as LEMS) would benefit.

LEMS has noteworthy features making it suitable for carotid artery imaging. Because we use entropy rather than a classification technique, we avoid the need to specify the number of classes, and circumvent the problem of classification errors. The bicubic spline enables the modeling of the severe inhomogeneity from the neck surface coils. To keep optimization tractable, we locally optimize, starting with the area having the highest SNR. By progressively merging regions having lower SNR values, the bias field estimate in the high SNR areas is minimally affected by the noise in low signal areas. In preliminary experiments where this merging strategy was not used and entropy was independently minimized for each region surrounding a knot, local minima were a problem and poor results were achieved. This problem was acute when a strong intensity fall-off was present. For brain data with less than a 10% inhomogeneity, the merging process was unnecessary.

Significant differences in performance were obtained between the methods that we compared. We found that our method gave better results and was more robust than Likar's on our images that suffer from very steep intensity inhomogeneity

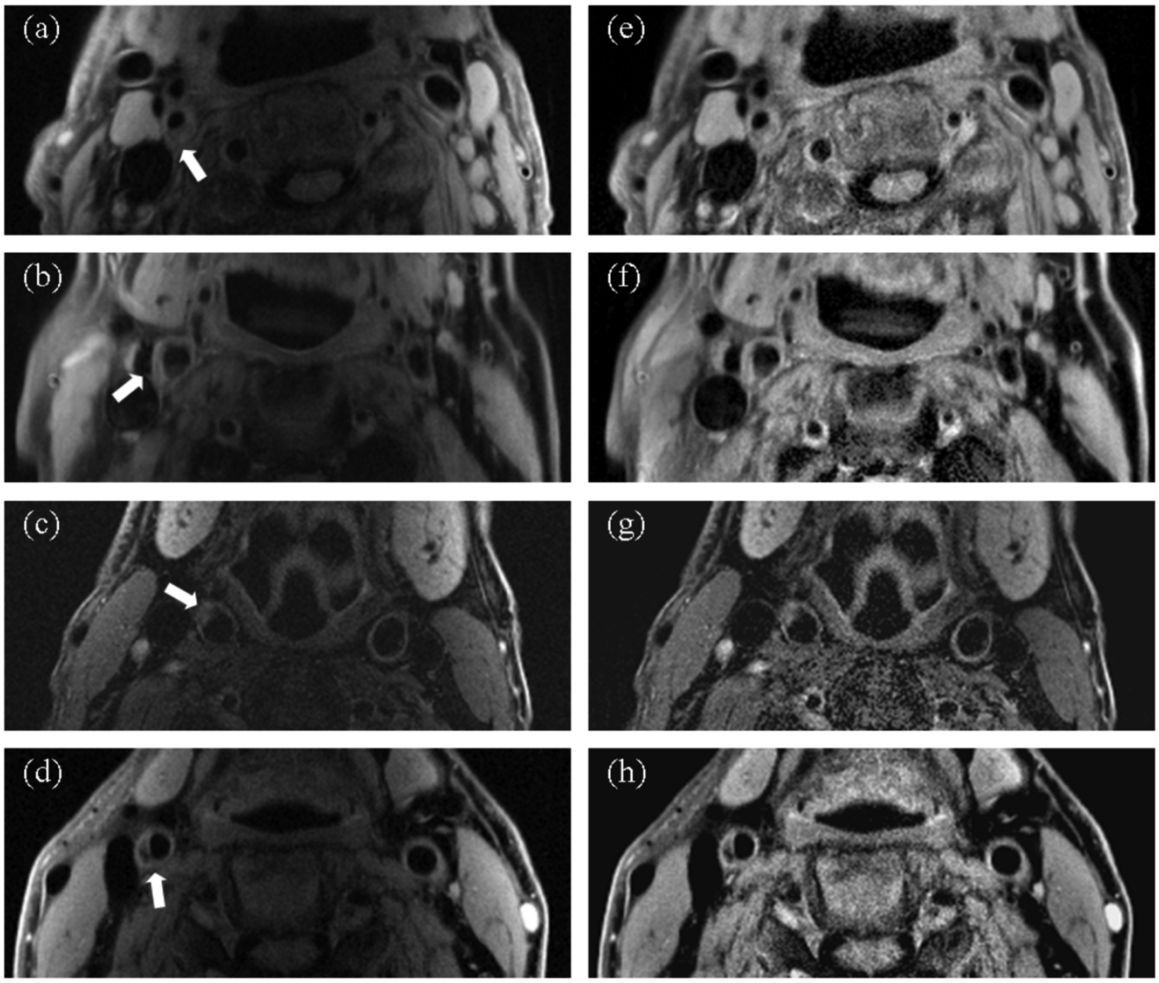


Fig. 11. Example patient images from four different patients. Original images on the left [(a)–(d])] are compared with respective images corrected with LEMS on the right [(e)–(h)]. These cases are illustrative of atheroma (arrows) as seen with PDW [(a), (e), (d), (h)], T1 [(b) and (f)], and T2W [(c) and (g)]. The corrected images contain much reduced intensity inhomogeneity, which should facilitate both visual diagnosis and automatic intensity-based segmentation.

and severe noise. From our experiments we think that local minima are exacerbated in the presence of noise and that low SNR areas corrupt the optimization: the performances degraded with increasing noise level. However, it should be noted that our data content (neck) differ markedly from the data corrected in Likar *et al.* (brain). Another difference is the optimization method: we used Matlab optimization algorithms that might not be optimally tuned to the problem at hand. These issues probably explain the difficulties that we experienced using Likar's method; convergence was often an issue, and inhomogeneity was still noticeable after correction.

Noise sensitivity was also a problem with the AFCM method, leading us to modify it to include a background and outlier classes to make it more robust, and address the problems at interface between tissue and pixels void of signal (the background class was not subject to bias field). Nonetheless, we experienced convergence issues on many images (more outliers in our results on patient images), and the method requires long computation time, although Pham and Prince use a multi-resolution approach that speeds up greatly the convergence [49]. More importantly, we obtained better quantitative performance with the LEMS method than the other methods including LINFIL, which is widely used because of its simple and fast implementation. A major drawback of the methods based on linear filtering is the presence of edge artifacts that hinders their ability to correct data at interfaces between tissue and air or, as in our case, voxels void of signal from fat or flow suppression. Another problem with this last class of methods is the assumption that data spectrum and the bias field spectrum do not overlap. This assumption does not hold when surface array coils are used because the signal drop-off can be very steep, generating relatively high frequencies comparable to the low frequencies generated by large homogeneous tissue areas such as the sternocleidomastoid muscles in the neck.

Extension to three dimensions could be implemented. However, we are interested in vascular imaging where data are acquired slice by slice. Some datasets have slices with varying quality; images can have ghosting artifacts inherent to sequences where k-space is filled in an interleaved way and motion artifacts from swallowing and head motion. Acquisitions are ECG gated, inducing different repetition times between slices, ensuing that the same tissue can have different signal levels at different slice. As a result of these issues, image slices have different histograms and the optimization of the histogram

for an entire volume might be problematic. Further, the phased array coils can be positioned at an oblique angle increasing the magnitude of bias field variations in the longitudinal direction, a potential problem with thick image slices. A three-dimensional (3-D) tricubic spline model would be more appropriate for correcting brain datasets where a head cage coil and 3-D acquisition MR sequences are used.

In the case of noisy images, such as those found in some T2W images, the LEMS algorithm, like other correction algorithms, sometimes fails (Fig. 8). One potential solution is to use the bias field estimated in another scan, such as a PDW scan, to correct the T2W image (Fig. 8). We did not observe the artifacts reported by others when this was done [6]. We are examining some potential algorithms that would automatically test for the suitability of a correction by comparing bias fields estimated from different image types. Very possibly, algorithms can also be created that would substitute a bias field from another scan type, use another bias field as an initialization, or compute a bias field from multiple MR scans.

LEMS is computationally demanding. On a Pentium 4 class desktop computer, correction of a 2-D MR image requires about 60 s, assuming four passes of the algorithm. With ten slices and three image types per patient, this gives about 30 minutes of computation time. However, this should be considered a worst-case scenario for multiple reasons. First, the bias estimated from one dataset could be used to correct another (Fig. 10) or at least provide an initial field for subsequent optimization. Second, the method has been implemented with Matlab's interpreted code and significant speed up should be expected upon compilation and code optimization. Third, a 3-D implementation for some applications where the intensity inhomogeneity changes smoothly in all three spatial directions would almost certainly reduce the number of knots, and the computation time. Simpler algorithms such as LINFIL require much less computation and can correct images in less than a second. However, this comes at the detriment of accuracy.

VI. Conclusion

We have presented a new method to correct intensity inhomogeneity in MR images subject to the strong shading artifact present when phased array reception coils are used to perform imaging of carotid arteries. Many methods exist to correct intensity inhomogeneity, each one providing trade off between computation burden and correction performance for a given application. Most of them have been designed for brain imaging. The method that we propose gives excellent results for surface coils imaging of atherosclerosis. After the intensity inhomogeneity is removed, it is easier to visually evaluate lesions by comparing gray scale values to those of skeletal muscle, as is often done by researchers in the field. Despite its relatively long processing time that we hope to reduce in the future, we believe that it will facilitate visual and/or computer diagnosis. Because entropy and not classification is used, this technique should be easily applied to a variety of medical imaging applications. The only parameter that might have to be adjusted for various applications is the knot spacing, and that should be related to the coil geometry. Future work includes the possible extension to 3-D, the normalization of intra- and inter-patient images, the optimization of code to reduce the computation time, and post-processing tissue classification of the arterial walls for atherosclerosis lesion assessment.

References

[1] H. H. Quick, J. F. Debatin, and M. E. Ladd, "MR imaging of the vessel wall," *Eur. Radiol.*, vol. 12, no. 4, pp. 889–900, 2002.

[2] J. M. Cai, T. S. Hatsukami, M. S. Ferguson, R. Small, N. L. Polissar, and C. Yuan, "Classification of human carotid atherosclerotic lesions with *in vivo* multicontrast magnetic resonance imaging," *Circulation*, vol. 106, no. 11, pp. 1368–1373, Sep. 2002.

[3] J. F. Toussaint, G. M. LaMuraglia, J. F. Southern, V. Fuster, and H. L. Kantor, "Magnetic resonance images lipid, fibrous, calcified, hemorrhagic, and thrombotic components of human atherosclerosis *in vivo*," *Circulation*, vol. 94, no. 5, pp. 932–938, Sep. 1996.

[4] J. Morrisett, W. Ick, R. Harma, G. Awrie, M. Eardon, E. Zell, J. Chwartz, G. Unter, and D. Orenstein, "Discrimination of components in atherosclerotic plaques from human carotid endarterectomy specimens by magnetic resonance imaging *ex vivo*," *Magn. Reson. Imag.*, vol. 21, no. 5, pp. 465–474, 2003.

[5] J. G. Sled and G. B. Pike, "Standing-wave and RF penetration artifacts caused by elliptic geometry: an electrodynamic analysis of MRI," *IEEE Trans. Med. Imag.*, vol. 17, no. 4, pp. 653–662, Aug. 1998.

[6] A. Simmons, P. S. Tofts, G. J. Barker, and S. R. Arridge, "Sources of intensity nonuniformity in spin-echo images at 1.5T," *Magn. Reson. Med.*, vol. 32, no. 1, pp. 121–128, 1994.

[7] G. Collewet, A. Davenel, C. Toussaint, and S. Akoka, "Correction of intensity nonuniformity in spin-echo T1-weighted images," *Magn. Reson. Imag.*, vol. 20, no. 4, pp. 365–373, 2002.

[8] M. Tincher, C. R. Meyer, R. Gupta, and D. M. Williams, "Polynomial modeling and reduction of RF body coil spatial inhomogeneity in MRI," *IEEE Trans. Med. Imag.*, vol. 12, no. 2, pp. 361–365, Jun. 1993.

[9] S. E. Moyher, D. B. Vigneron, and S. J. Nelson , "Surface coil MR-imaging of the human brain with an analytic reception profile correction," *J. Magn. Reson. Imag.*, vol. 5, no. 2, p. 139, Mar. 1995.

[10] W. W. Brey and P. A. Narayana, "Correction for intensity falloff in surface coil magnetic resonance imaging," *Med. Phys.*, vol. 15, no. 2, pp. 241–245, Jan. 1988.

[11] J. W. Murakami, C. E. Hayes, and E. Weinberger, "Intensity correction of phased-array surface coil images," *Magn. Reson. Med.*, vol. 35, no. 4, p. 585, 1996.

[12] P. K. Prussmann, M. Weiger, M. B. Scheidegger, and P. Boesiger, "SENSE: Sensitivity Encoding for Fast MRI," *Magn. Reson. Med.*, vol. 42, no. 5, pp. 952–962, 1999.

[13] B. H. Brinkmann, A. Manduca, and R. A. Robb, "Optimized homomorphic unsharp masking for MR grayscale inhomogeneity correction," *IEEE Trans. Med. Imag.*, vol. 17, no. 2, pp. 161–171, Apr. 1998.

[14] S. M. Cohen, R. M. DuBois, and M. M. Zeineh, "Rapid and effective correction of RF inhomogeneity for high field magnetic resonance imaging," *Hum. Brain Mapp.*, vol. 10, no. 4, pp. 204–211, 2000.

[15] C. Han, T. S. Hatsukami, and C. Yuan, "A multi-scale method for automatic correction of intensity nonuniformity in MR images," *J. Magn. Reson. Imag.*, vol. 13, no. 3, p. 428, 2001.

[16] E. A. Vokurka, N. A. Thacker, and A. Jackson, "A fast model independent method for automatic correction of intensity nonuniformity in MRI data," *J. Magn. Reson. Imag.*, vol. 10, no. 4, p. 550, 1999.

[17] K. Van Leemput, F. Maes, D. Vandermeulen, and P. Suetens, "Automated model-based bias field correction of MR images of the brain," *IEEE Trans. Med. Imag.*, vol. 18, no. 10, pp. 885–896, Oct. 1999.

[18] M. Styner, C. Brechbuhler, G. Szckely, and G. Gerig, "Parametric estimate of intensity inhomogeneities applied to MRI," *IEEE Trans. Med. Imag.*, vol. 19, no. 3, pp. 153–165, Mar. 2000.

[19] C. R. Meyer, P. H. Bland, and J. Pipe, "Retrospective correction of intensity inhomogeneities in MRI," *IEEE Trans. Med. Imag.*, vol. 14, no. 1, pp. 36–41, Mar. 1995.

[20] B. M. Dawant, A. P. Zijdenbos, and R. A. Margolin, "Correction of intensity variations in MR images for computer-aided tissue classification," *IEEE Trans. Med. Imag.*, vol. 12, no. 4, pp. 770–781, Dec. 1993.

[21] J. Ashburner and K. J. Friston, "Voxel-based morphometry—the methods," *Neuroimage*, vol. 11, no. 6, p. 805, 2000.

[22] D. W. Shattuck, S. R. Sandor-Leahy, K. A. Schaper, D. A. Rottenberg, and R. M. Leahy, "Magnetic resonance image tissue classification using a partial volume model," *Neuroimage*, vol. 13, no. 5, p. 856, 2001.

[23] J. L. Marroquin, B. C. Vemuri, S. Botello, E. Calderon, and A. Fernandez-Bouzas, "An accurate and efficient Bayesian method for automatic segmentation of brain MRI," *IEEE Trans. Med. Imag.*, vol. 21, no. 8, pp. 934–945, Aug. 2002.

[24] A. W. C. Liew and H. Yan, "An adaptive spatial fuzzy clustering algorithm for 3-D MR image segmentation," *IEEE Trans. Med. Imag.*, vol. 22, no. 9, pp. 1063–1075, Sep. 2003.

[25] J. F. Mangin, "Entropy minimization for automatic correction of intensity nonuniformity," in *Proc. IEEE Workshop Mathematical Methods in Biomedical Image Analysis*, Hilton Head Island, SC, 2000, pp. 162–169.

[26] D. L. Pham and J. L. Prince, "Adaptive fuzzy segmentation of magnetic resonance images," *IEEE Trans. Med. Imag.*, vol. 18, no. 9, pp. 737–752, Sep. 1999.

[27] O. Salvado, C. Hillenbrand, S. Zhang, J. S. Suri, and D. L. Wilson, "MR signal inhomogeneity correction for visual and computerized atherosclerosis lesion assessment," Proceeding of 2004 IEEE International Symposium on Biomedical Imaging pp. 1143–1146, 2004.

[28] W. M. I. Wells, W. E. L. Grimson, R. Kikinis, and F. A. Jolesz, "Adaptive segmentation of MRI data," *IEEE Trans. Med. Imag.*, vol. 15, no. 4, pp. 429–442, Aug. 1996.

[29] R. Guillemaud and M. Brady, "Estimating the bias field of MR images," *IEEE Trans. Med. Imag.*, vol. 16, no. 3, p. 238, Jun. 1997.

[30] J. C. Rajapakse and F. Kruggel, "Segmentation of MR images with intensity inhomogeneities," *Image Vis. Computing*, vol. 16, no. 3, pp. 165–180, 1998.

[31] Y. Zhang, M. Brady, and S. Smith, "Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm," *IEEE Trans. Magn.*, vol. 20, no. 1, pp. 45–57, Jan. 2001.

[32] K. Held, E. R. Kops, B. J. Krause, W. M. Wells, III, R. Kikinis, and H.-W. Muller-Gartner, "Markov random field segmentation of brain MR images," *IEEE Trans. Med. Imag.*, vol. 16, no. 6, pp. 878–886, Dec. 1997.

[33] C. E. Hayes, C. M. Mathis, and C. Yuan, "Surface coil phased arrays for high-resolution imaging of the carotid arteries," *J. Magn. Reson. Imag.*, vol. 6, no. 1, p. 109, 1996.

[34] O. Salvado, C. Hillenbrand, J. S. Suri, and D. L. Wilson, "MR coil sensitivity inhomogeneity correction for plaque characterization in carotid arteries," in *Medical Imaging 2004: Visualization, Image Guided Procedures, and Display*. San Diego, CA: SPIE, 2004.

[35] J. Suckling, T. Sigmundsson, K. Greenwood, and E. T. Bullmore, "A modified fuzzy clustering algorithm for operator independent brain tissue classification of dual echo MR images," *Magn. Reson. Imag.*, vol. 17, no. 7, p. 1065, 1999.

[36] J. G. Sled, A. P. Zijdenbos, and A. C. Evans, "A nonparametric method for automatic correction of intensity nonuniformity in MRI data," *IEEE Trans. Med. Imag.*, vol. 17, no. 1, pp. 87–97, Feb. 1998.

[37] B. Likar, M. A. Viergever, and F. Pernus, "Retrospective correction of MR intensity inhomogeneity by information minimization," *IEEE Trans. Med. Imag.*, vol. 20, no. 12, pp. 1398–1410, Dec. 2001.

[38] E. M. Haacke, R. W. Brown, M. R. Thompson, and R. Venkatesan, *Magnetic Resonance Imaging; Physical Principles and Sequence Design*. New York: Wiley-Liss, 1999.

[39] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 7, pp. 629–639, Jul. 1990.

[40] M. J. Black, G. Sapiro, D. H. Marimont, and D. Heeger, "Robust anisotropic diffusion," *IEEE Trans. Image Process.*, vol. 7, no. 3, pp. 421–432, Mar. 1998.

[41] E. W. Weisstein, Dec. 2005, Cubic Spline [Online]. Available: http://mathworld.wolfram.com/CubicSpline.html

[42] G. E. Forsythe, M. A. Malcolm, and C. B. Moler, *Computer Methods for Mathematical Computations*. Upper Saddle River, NJ: Prentice-Hall, 1976.

[43] MathWorks Inc., Dec. 2005, Matlab Bicubic Spline [Online]. Available: http://www.mathworks.com/products/splines/

[44] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed. New York: Wiley-Interscience, 2001, vol. 654.

[45] C. Yuan and W. S. Kerwin, "MRI of atherosclerosis," *J. Magn. Reson. Imag.*, vol. 19, no. 6, pp. 710–719, 2004.

[46] C. Yuan, L. M. Mitsumori, M. S. Ferguson, N. L. Polissar, D. Echelard, G. Ortiz, R. Small, J. W. Davies, W. S. Kerwin, and T. S. Hatsukami, "*In vivo* accuracy of multispectral magnetic resonance imaging for identifying lipid-rich necrotic cores and intraplaque hemorrhage in advanced human carotid plaques," *Circulation*, vol. 104, no. 17, pp. 2051–2056, Oct. 2001.

[47] E. S. Clarke, R. R. Hammond, J. R. Mitchell, and B. K. Rutt, "Quantitative assessment of carotid plaque composition using multicontrast MRI and registered histology," *Magn. Reson. Med.*, vol. 50, no. 6, pp. 1199–1208, 2003.

[48] X. Q. Zhao, C. Yuan, T. S. Hatsukami, E. H. Frechette, X. J. Kang, K. R. Maravilla, and B. G. Brown, "Effects of prolonged intensive lipid-lowering therapy on the characteristics of carotid atherosclerotic plaques *in vivo* by MRI: a case-control study," *Arterioscler. Thromb. Vasc. Biol.*, vol. 21, no. 10, pp. 1623–1629, Oct. 2001.

[49] D. L. Pham and J. L. Prince, "An adaptive fuzzy C-means algorithm for image segmentation in the presence of intensity inhomogeneities," *Pattern Recognit. Lett.*, vol. 20, no. 1, pp. 57–68, 1999.