A Variational Level Set Approach to Segmentation and Bias Correction of Images with Intensity Inhomogeneity

Chunming Li[1,*], Rui Huang[2], Zhaohua Ding[1], Chris Gatenby[1], Dimitris Metaxas[2], and John Gore[1]

[1] Vanderbilt University Institute of Imaging Science, USA
chunming.li@vanderbilt.edu
[2] Department of Computer Science, Rutgers University, USA

Abstract. This paper presents a variational level set approach to joint segmentation and bias correction of images with intensity inhomogeneity. Our method is based on an observation that intensities in a relatively small local region are separable, despite of the inseparability of the intensities in the whole image caused by the intensity inhomogeneity. We first define a weighted K-means clustering objective function for image intensities in a neighborhood around each point, with the cluster centers having a multiplicative factor that estimates the bias within the neighborhood. The objective function is then integrated over the entire domain and incorporated into a variational level set formulation. The energy minimization is performed via a level set evolution process. Our method is able to estimate bias of quite general profiles. Moreover, it is robust to initialization, and therefore allows automatic applications. The proposed method has been used for images of various modalities with promising results.

1 Introduction

A major problem for automatic segmentation of magnetic resonance (MR) images is the intensity inhomogeneity due to the bias field, which is caused by limitations in imaging devices and subject-induced susceptibility effect. The bias can cause serious misclassifications when intensity-based segmentation algorithms are used. Essentially, the misclassification is due to an overlap of the intensity range of different tissues introduced by the bias field, so that the voxels in different tissues are not separable based on their intensities. Intensity inhomogeneities also often occur in images of other modalities, such as X-ray and computed tomography images.

Bias correction has been extensively studied in the past two decades [1,2]. Methods of bias correction can be categorized into two classes: prospective methods and retrospective methods. Prospective methods aim to avoid intensity inhomogeneities in the image acquisition process. These methods, while capable of correcting intensity inhomogeneity induced by the imaging device, are not able to remove subject-induced intensity inhomogeneity. In contrast, retrospective methods only rely on the information in the acquired images. Therefore, they can also remove intensity inhomogeneities regardless of their sources. Early retrospective methods include those based on filtering [3],

---

surface fitting [4,5], and histogram [6]. Segmentation based methods [1,7,8,9] are more attractive, as they unify segmentation and bias correction within a single framework. These methods consist of the interaction between segmentation and bias correction, which benefit each other to yield final segmentation and bias correction.

The generally accepted assumption on the bias field is that it is slowly varying. It is necessary and beneficial to preserve the slowly varying property of the computed bias field in segmentation based methods. In Wells *et al.*'s method [1], the directly computed bias field is not smooth, which would lead to poor bias correction and segmentation results. An moving-average low pass filter is empirically used to force the bias field to be smooth. In [10], Pham and Prince proposed an energy minimization method for adaptive segmentation and estimation of the bias field. In their method, the smoothness of the bias field is ensured by adding a smoothing constraint term in their objective function, which leads to a highly expensive procedure to solve a space-varying difference equation. Such an expensive smoothing procedure is avoided in some well-known parametric methods (e.g. [8,7]) by modeling the bias field as a polynomial, which is smooth by nature. However, due to limited approximation capability of polynomials, these methods are not able to approximate bias fields of general profiles, such as those in 7T MR images (see Fig. 5 for example).

In this paper, we propose a variational level set approach to bias correction and segmentation for images corrupted with intensity inhomogeneities. A unique feature of our method is that the computed bias field is intrinsically ensured to be smooth by the data term in our variational formulation, without any additional effort to maintain the smoothness of the bias field, and it can approximate bias fields of more general profiles, such as those in 7T MR images. Moreover, our method is not sensitive to initialization, thereby allowing automatic applications.

2   Method

2.1   Model of Images with Intensity Inhomogeneity

Our method is based on the model commonly used to describe images with intensity inhomogeneity:

$$I = \tilde{b}\tilde{J} + n \tag{1}$$

where I is the measured image intensity, $\tilde{J}$ is the true signal to be restored, $\tilde{b}$ is the bias field, and n is noise. The superscript tilde in $\tilde{J}$ and $\tilde{b}$ is used to distinguish the unknown true signal $\tilde{J}$ and the bias field $\tilde{b}$ from their estimates, which will be denoted by J and b, respectively. The generally accepted assumption on the bias field is that it is smooth (or slowly varying). Ideally, the intensity $\tilde{J}$ in each tissue should take a specific value $\tilde{c}_i$ of the physical property being measured (e.g. the proton density for MR images).

In general, we assume that the true image $\tilde{J}$ and the bias field $\tilde{b}$ have the following properties:

(P1) The bias field $\tilde{b}$ is slowly varying in the entire image domain.
(P2) The true image intensities $\tilde{J}$ are approximately a constant within each class of tissue, i.e. $\tilde{J}(\mathbf{x}) \approx \tilde{c}_i$ for $\mathbf{x} \in \tilde{\Omega}_i$, with $\{\tilde{\Omega}_i\}_{i=1}^{N}$ being a partition of Ω.

2.2 Energy Formulation

Our ultimate goal is to separate the image domain Ω into N disjoint regions $\tilde{\Omega}_i$, $i = 1, \cdots, N$, based on the measured image I. However, due to the intensity inhomogeneity caused by the bias field $\tilde{b}$, the measured intensities are not separable by using traditional intensity based classification methods. In this section, we will propose a new method for joint segmentation and bias correction. Our method is based on an observation that the intensities in a relatively small region are separable, which can be verified by the above assumptions (P1) and (P2) as explained in the following.

We consider a circular neighborhood with a relatively small radius ρ centered at each point $\mathbf{x}$ in the image domain Ω, defined by $\mathcal{O}_{\mathbf{x}} \triangleq \{\mathbf{y} : |\mathbf{y} - \mathbf{x}| \le \rho\}$. The partition $\{\tilde{\Omega}_i\}_{i=1}^{N}$ induces a partition of the neighborhood $\mathcal{O}_{\mathbf{x}}$, i.e., $\{\mathcal{O}_{\mathbf{x}} \cap \tilde{\Omega}_i\}_{i=1}^{N}$. For a smooth function $\tilde{b}$, the values $\tilde{b}(\mathbf{y})$ for all $\mathbf{y}$ in the circular neighborhood $\mathcal{O}_{\mathbf{x}}$ can be well approximated by $\tilde{b}(\mathbf{x})$, which is at the center of $\mathcal{O}_{\mathbf{x}}$. Therefore, the intensities $\tilde{b}(\mathbf{y})\tilde{J}(\mathbf{y})$ in each subregion $\mathcal{O}_{\mathbf{x}} \cap \tilde{\Omega}_i$ are approximately the constant $b(\mathbf{x})\tilde{c}_i$. Thus, we have the following approximation

$$\tilde{b}(\mathbf{y})\tilde{J}(\mathbf{y}) \approx \tilde{b}(\mathbf{x})\tilde{c}_i \quad \text{for } \mathbf{y} \in \mathcal{O}_{\mathbf{x}} \cap \tilde{\Omega}_i \tag{2}$$

The constants $\tilde{b}(\mathbf{x})\tilde{c}_i$ can be considered as the approximations of the *cluster centers (or means)* of the clusters $\{I(\mathbf{y}) : \mathbf{y} \in \mathcal{O}_{\mathbf{x}} \cap \tilde{\Omega}_i\}$ within the neighborhood $\mathcal{O}_{\mathbf{x}}$. Therefore, the intensities in the neighborhood $\mathcal{O}_{\mathbf{x}}$ are around N distinct cluster centers $\tilde{m}_i \approx \tilde{b}(\mathbf{x})\tilde{c}_i$. The multiplicative components $\tilde{b}(\mathbf{x})$ and $\tilde{c}_i$ of the cluster centers $\tilde{m}_i \approx \tilde{b}(\mathbf{x})\tilde{c}_i$ can be estimated as the following.

Consider the task of classifying the intensities $I(\mathbf{y})$ in the neighborhood $\mathcal{O}_{\mathbf{x}}$ into N classes. In view of the separability of the intensities within the neighborhood $\mathcal{O}_{\mathbf{x}}$, this task can be performed by using the standard K-means clustering method. In this paper, we introduce a K-means clustering method based on the minimization of the following weighted objective function

$$\mathcal{E}_{\mathbf{x}} = \sum_{i=1}^{N} \int_{\mathcal{O}_{\mathbf{x}} \cap \Omega_i} K(\mathbf{x} - \mathbf{y})|I(\mathbf{y}) - b(\mathbf{x})c_i|^2 d\mathbf{y} \tag{3}$$

where $b(\mathbf{x})c_i$ are the cluster centers to be optimized, and $K(\mathbf{x} - \mathbf{y})$ is a non-negative weighting function such that $K(\mathbf{x} - \mathbf{y}) = 0$ for $|\mathbf{x} - \mathbf{y}| > \rho$ and $\int_{\mathcal{O}_{\mathbf{x}}} K(\mathbf{x} - \mathbf{y})d\mathbf{y} = 1$.

Although the choice of the weighting function is flexible, it is preferable to use a weighting function $K(\mathbf{x} - \mathbf{y})$ such that larger weights are assigned to the data $I(\mathbf{y})$ for $\mathbf{y}$ closer to the center $\mathbf{x}$ of the neighborhood $\mathcal{O}_{\mathbf{x}}$. In this paper, the weighting function K is chosen as a truncated Gaussian kernel

$$K(\mathbf{u}) = \begin{cases} \frac{1}{a}e^{-|\mathbf{u}|^2/2\sigma^2} & \text{for } |\mathbf{u}| \le \rho \\ 0 & \text{else} \end{cases}$$

where a is a constant such that $\int K(\mathbf{u}) = 1$. The above objective function $\mathcal{E}_{\mathbf{x}}$ can be rewritten as

$$\mathcal{E}_{\mathbf{x}} = \sum_{i=1}^{N} \int_{\Omega_i} K(\mathbf{x} - \mathbf{y})|I(\mathbf{y}) - b(\mathbf{x})c_i|^2 d\mathbf{x} \tag{4}$$

due to the fact that $K(\mathbf{x} - \mathbf{y}) = 0$ for $\mathbf{y} \notin \mathcal{O}_{\mathbf{x}}$.

As mentioned above, the measured intensities $I(\mathbf{y})$ within the neighborhood $\mathcal{O}_\mathbf{x}$ are separable, and therefore could be classified into N clusters by minimizing the objective function $\mathcal{E}_\mathbf{x}$, which results in the optimal cluster centers m_i and an optimal partition of $\mathcal{O}_\mathbf{x}$. However, we still cannot determine the components $b(\mathbf{x})$ and c_i of the computed cluster centers m_i. Moreover, our ultimate goal is to find an optimal set of a partition of the entire image domain Ω, the bias field b, and the constants c_i. The minimization of a single objective function $\mathcal{E}_\mathbf{x}$, which is defined for a point $\mathbf{x}$, does not achieve this goal. We need to minimize $\mathcal{E}_\mathbf{x}$ for all the points $\mathbf{x}$. This can be achieved by minimizing the integral of $\mathcal{E}_\mathbf{x}$ over Ω. Therefore, we define an energy $\mathcal{E} \triangleq \int \mathcal{E}_\mathbf{x} d\mathbf{x}$, i.e.

$$\mathcal{E} \triangleq \int \left(\sum_{i=1}^{N} \int_{\Omega_i} K(\mathbf{x}-\mathbf{y})|I(\mathbf{y})-b(\mathbf{x})c_i|^2 d\mathbf{y} \right) d\mathbf{x} \tag{5}$$

Directly minimizing the energy with the partition $\{\Omega_i\}_{i=1}^{N}$ as a variable is not convenient. We will use one or multiple level set functions to represent a partition $\{\Omega_i\}_{i=1}^{N}$. The energy minimization can thus be performed by solving a level set evolution equation.

3   Level Set Formulation

We first consider the case of $N = 2$. In this case, the image domain is partitioned into two regions $\{\Omega_i\}_{i=1}^{2}$. These two regions can be represented by the regions separated by the zero level contour of a function ϕ, i.e., $\Omega_1 \triangleq \{\phi > 0\}$ and $\Omega_2 \triangleq \{\phi < 0\}$. Using Heaviside function H, the energy $\mathcal{E}$ in Eq. (5) can be expressed as an energy in terms of ϕ, b, and $\mathbf{c}$ as below

$$\mathcal{E}(\phi, b, \mathbf{c}) = \int \left(\sum_{i=1}^{2} \int K(\mathbf{x}-\mathbf{y})|I(\mathbf{y})-b(\mathbf{x})c_i|^2 M_i(\phi(\mathbf{y}))d\mathbf{y} \right) d\mathbf{x} \tag{6}$$

where $M_1(\phi(\mathbf{x})) = H(\phi(\mathbf{x}))$ and $M_2(\phi(\mathbf{x})) = 1 - H(\phi(\mathbf{x}))$. In practice, we use a smoothed Heaviside function $H_\epsilon(x) = \frac{1}{2}[1 + \frac{2}{\pi}\arctan(\frac{x}{\epsilon})]$ to approximate the original Heaviside function H, with $\epsilon = 1$ as used in [11].

It is necessary to add a regularization term $\mathcal{R}(\phi)$ to the above energy in the following energy functional:

$$\mathcal{F}(\phi, b, c_1, c_2) \triangleq \mathcal{E}(\phi, b, \mathbf{c}) + \mathcal{R}(\phi) \tag{7}$$

where $\mathcal{R}(\phi) \triangleq \nu \int |\nabla H(\phi)|d\mathbf{x} + \mu \int (|\nabla \phi| - 1)^2 d\mathbf{x}$. The first term in $\mathcal{R}$ serves to regularize the zero level contour of ϕ as in typical level set methods [11], while the second term regularizes the entire level set function ϕ by penalizing its deviation from signed distance, as in the level set methods proposed by Li *et al.*[12,13]. The energy $\mathcal{E}(\phi, b, \mathbf{c})$ is the data term in our variational framework.

Similarly, we can use multiple level set functions $\phi_1, \cdots, \phi_n$ to represent regions $\{\Omega_i\}_{i=1}^{N}$ with $N = 2^n$ as in [14]. For convenience, we use a vector valued function $\Phi = (\phi_1, \cdots, \phi_n)$ to represent the functions $\phi_1, \cdots, \phi_n$. The energy for general multi-phase formulation of our method can be defined as

$$\mathcal{F}(\Phi, b, \mathbf{c}) \triangleq \int (\sum_{i=1}^{N} \int K(\mathbf{x}-\mathbf{y})|I(\mathbf{y})-b(\mathbf{x})c_i|^2 M_i(\Phi(\mathbf{y}))d\mathbf{y})d\mathbf{x} + \sum_{i=1}^{n} \mathcal{R}(\phi_i) \tag{8}$$

where $M_i(\Phi)$ are functions of Φ which are designed such that $\sum_{i=1}^{N} M_i(\Phi) = 1$. The definition of M_i in the four-phase case are given in [14]. For $N = 3$ and two level set functions ϕ_1 and ϕ_2, we can define $M_1(\phi_1, \phi_2) = H(\phi_1)H(\phi_2)$, $M_2(\phi_1, \phi_2) = H(\phi_1)(1 - H(\phi_2))$, and $M_3(\phi_1, \phi_2) = 1 - H(\phi_1)$ to obtain a three-phase formulation.

We only describe the energy minimization for the two-phase case in this paper (the multi-phase case can be solved with the similar procedure). For fixed $\mathbf{c}$ and b, the minimization of $\mathcal{F}(\phi, \mathbf{c}, b)$ consists in solving the level set evolution equation as the gradient descent equation

$$\frac{\partial \phi}{\partial t} = -\frac{\partial \mathcal{F}}{\partial \phi} \tag{9}$$

where $\frac{\partial \mathcal{F}}{\partial \phi}$ is the Gâteaux derivative (the first order functional derivative) of the energy $\mathcal{F}$. In numerical implementation, at each iteration according to Eq. (9), the variables $\mathbf{c}$ and b are updated according to the following procedure. For fixed ϕ and $\mathbf{c}$, we find an optimal bias field $\hat{b}$ that minimizes $\mathcal{F}(\phi, \mathbf{c}, b)$. It can be shown that the minimizer $\hat{b}$ is

$$\hat{b} = \frac{(IJ^{(1)}) * K}{J^{(2)} * K} \tag{10}$$

where $*$ is the convolution operation, and $J^{(1)} = \sum_{i=1}^{N} c_i M_i(\phi)$ and $J^{(2)} = \sum_{i=1}^{N} c_i^2 M_i(\phi)$. For fixed ϕ and b, we find an optimal $\hat{\mathbf{c}}$ that minimizes $\mathcal{F}(\phi, \mathbf{c}, b)$. By some calculus manipulations, it can be shown that the minimizer $\hat{\mathbf{c}} = (\hat{c}_i, \cdots, \hat{c}_N)$ is

$$\hat{c}_i = \frac{\int (b * K) I M_i(\phi) d\mathbf{x}}{\int (b^2 * K) M_i(\phi) d\mathbf{x}}, \quad i = 1, \cdots, N. \tag{11}$$

It is worth noting that the expression of $\hat{b}$ in Eq. (10) with convolutions shows that $\hat{b}$ is smooth. The smoothness of $\hat{b}$ is intrinsically ensured by the data term $\mathcal{E}(\phi, \mathbf{c}, b)$ in our variational framework. This is a desirable advantage of our method: there is no need for imposing a smoothing term to ensure the smoothness of the bias field.

4   Experimental Results and Validation

We use the parameters $\sigma = 4$, $\mu = 1$, and $\nu = 0.001 \times 255^2$ for all the images in this paper. Our method is robust to the initialization of the constants $\mathbf{c} = (c_1, \cdots, c_N)$, the bias field b, and the level set functions. For automatic applications, the constants $c_1, \cdots, c_N$ can be initialized as N equally spaced numbers between the minimum and maximum intensities of the original image, and the bias field b is initialized as $b = 1$. The level set functions can be automatically generated or manually initialized by the users. The number of phases N depends on the number of tissue types in the images, which is usually known in practice.

We first demonstrate our method in the two-phase case (i.e. $N = 2$). For example, Fig. 1 shows the result of our method for an X-ray image. Intensity inhomogeneity is obvious in this image. We use this example to show the desirable capability of our method in joint segmentation and bias correction. The bias corrected image is given by the quotient $I/\hat{b}$. It is worth noting that our method allows for flexible initialization
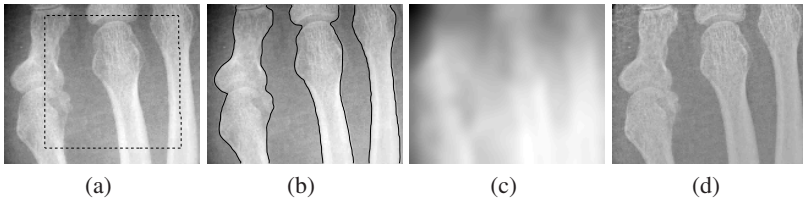
Fig. 1. Applications of our method to an X-ray image. (a) Original image and initial contour (dashed black line); (b) Segmentation result (black lines); (c) Computed bias field; (d) Bias corrected image.
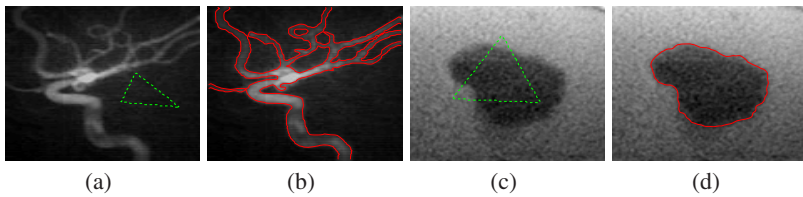

Fig. 2. Applications of our method to a CTA image and a CT image. (a) Original CTA image and initial contour (green dashed line); (b) Segmentation result; (c) Original CT image and initial contour; (d) Segmentation result.
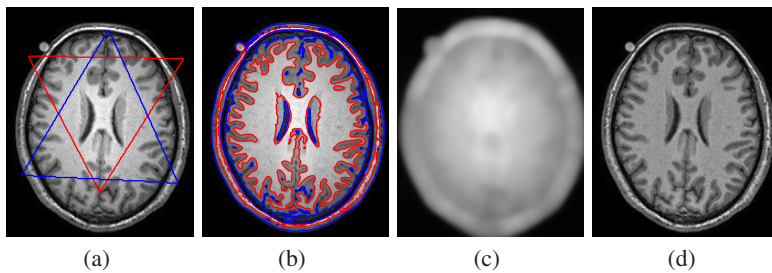
Fig. 3. Applications of our method to a 3T MR image. (a) Original image and initial contours: zero level contours of initial ϕ_1 (red) and ϕ_2 (blue). (b) Final zero level contours of ϕ_1 (red) and ϕ_2 (blue); (c) Computed bias field; (d) Bias corrected image.

of the level set function. The initial contour can be inside, outside, or cross the object boundaries. This can be seen from the results in Fig. 1 and those for a computed tomography angiography (CTA) image of vessel and a computed tomography (CT) image of a tumor in a liver shown in Fig. 2. The initial contours used to generate the initial level set functions are shown in Fig. 2(a) and 2(c), and the corresponding segmentation results are shown in Fig. 2(b) and 2(d).

Fig. 3 shows the result for a 3T MR brain image, which has obvious intensity inhomogeneity. The computed bias field and the segmentation result are simultaneously

(a)   (b)   (c)   (d)
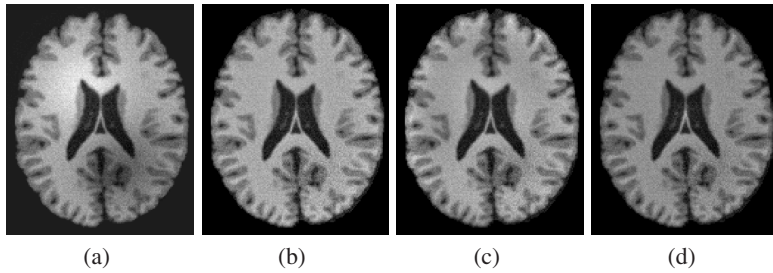
Fig. 4. Comparison with the methods of Wells *et al.*and Leemput *et al.*using simulated data obtained from BrainWeb. (a) Original image; (b) Our method; (c) Leemput *et al.*; (d) Wells *et al.*

Table 1. Coefficient of Variation

	Tissue	original	Bias corrected		
			Wells	Leemput	**Our method**
Image with bias	WM	20.16%	6.33%	6.81%	5.48%
	GM	23.65%	16.11%	16.97%	14.38%
Image with bias and noise	WM	18.93%	6.97%	7.92%	6.77%
	GM	21.20%	13.65%	14.68%	12.81%

(a)   (b)   (c)
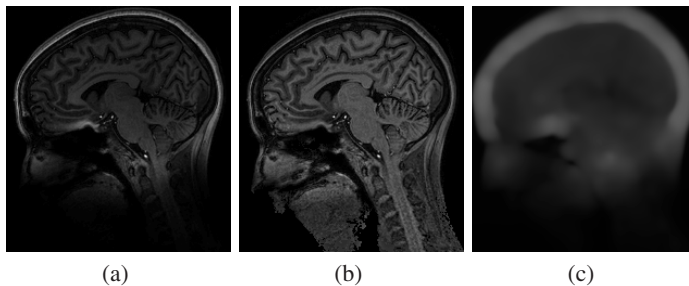
Fig. 5. Application of our method to a 7T MR image. (a) Original image; (b) Bias corrected image; (c) Computed bias field.

obtained, shown in Fig. 3(c) and 3(b), respectively. The bias corrected image is shown in Fig. 3(d). For comparison with other methods, we use coefficient of variance (CV) as a metric to evaluate the performance of the algorithms for bias correction and segmentation (c.f. [9,7]). Coefficient of variance is defined as a quotient between standard deviation and mean value of selected tissue class. A good algorithm for bias correction and segmentation should give low CV values for the bias corrected intensities within each segmented region. We tested our method and the methods of Leemput *et al.*and Wells *et al.*with two simulated images obtained from BrainWeb in the link `http://www.bic.mni.mcgill.ca/brainweb/`, one corrupted with bias field without noise and the other corrupted with both bias and noise. The CV values for the

two images are listed in Table 1. It can be seen that the CV values of our method are lower than those of Leemput's and Wells's methods, which indicates that the bias corrected images obtained in our method are more homogeneous than those of the other two methods. In Fig. 4, we show the bias corrected images for the noisy image as an example.

Our method has been tested on 7T MR images. At 7T, significant gains in image resolution can be obtained due to the increase in signal-to-noise ratio. However, susceptibility-induced gradients scale with main field, while the imaging gradients are currently limited to essentially the same strengths as used at lower field strengths (i.e., 3T). Such effects are most pronounced at air/tissue interfaces, as can be seen in Fig. 5(a) at the base of the frontal lobe. This appears as a localized and stronger bias, which is challenging to traditional methods for bias correction. This result shows the ability of our method to capture and correct such bias, as shown in Fig. 5(b) and 5(c).

5   Conclusion

We have presented a unified framework of bias correction and segmentation. A unique advantage of our method is that the smoothness of the computed bias field is intrinsically ensured by the data term in our variational formulation. Our method is able to capture bias of quite general profiles, and can be used for images of various modalities. Moreover, it is robust to initialization, thereby allowing automatic applications. Comparisons with two well-known bias correction methods demonstrate the advantages of the proposed method.

References

1. Wells, W., Grimson, E., Kikinis, R., Jolesz, F.: Adaptive segmentation of MRI data. IEEE Trans. Med. Imag. 15(4), 429–442 (1996)
2. Vovk, U., Pernus, F., Likar, B.: A review of methods for correction of intensity inhomogeneity in mri. IEEE Trans. Med. Imag. 26(3), 405–421 (2007)
3. Lewis, E., Fox, N.: Correction of differential intensity inhomogeneity in longitudinal MR images. Neuro. image 23(3), 75–83 (2004)
4. Dawant, B., Zijdenbos, A., Margolin, R.: Correction of intensity variations in MR images for computer-aided tissues classification. IEEE Trans. Med. Imag. 12(4), 770–781 (1993)
5. Meyer, C., Bland, P., Pipe, J.: Retrospective correction of intensity inhomogeneities in MRI. IEEE Trans. Med. Imag. 14(1), 36 (1995)
6. Sled, J., Zijdenbos, A., Evans, A.: A nonparametric method for automatic correction of intensity nonuniformity in MRI data. IEEE Trans. Med. Imaging 17(1), 87–97 (1998)
7. Leemput, K., Maes, F., Vandermeulen, D., Suetens, P.: Automated model-based bias field correction of MR images of the brain. IEEE Trans. Med. Imag. 18(10), 885–896 (1999)
8. Styner, M., Brechbuhler, C., Szekely, G., Gerig, G.: Parametric estimate of intensity inhomogeneities applied to MRI. IEEE Trans. Med. Imag. 19(3), 153–165 (2000)
9. Guillemaud, R., Brady, M.: Estimating the bias field of MR images. IEEE Trans. Med. Imaging 16, 238–251 (1997)
10. Pham, D., Prince, J.: Adaptive fuzzy segmentation of magnetic resonance images. IEEE Trans. Med. Imag. 18(9), 737–752 (1999)

11. Chan, T., Vese, L.: Active contours without edges. IEEE Trans. Imag. Proc. 10, 266–277 (2001)
12. C.,, Xu, C., Li, C.G., Fox, M.D.: Level set evolution without re-initialization: A new variational formulation. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), vol. 1, pp. 430–436 (2005)
13. Li, C., Kao, C., Gore, J., Ding, Z.: Implicit active contours driven by local binary fitting energy. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2007)
14. Vese, L., Chan, T.: A multiphase level set framework for image segmentation using the mumford and shah model. Int'l. J. Comp. Vis. 50, 271–293 (2002)