Automatic Correction of Intensity Nonuniformity from Sparseness of Gradient Distribution in Medical Images

Yuanjie Zheng[1], Murray Grossman[2], Suyash P. Awate[1], and James C. Gee[1]

[1] Penn Image Computing and Science Laboratory (PICSL),
Department of Radiology, University of Pennsylvania School of Medicine,
Philadelphia, PA, USA
[2] Department of Neurology, University of Pennsylvania School of Medicine,
Philadelphia, PA, USA

Abstract. We propose to use the sparseness property of the gradient probability distribution to estimate the intensity nonuniformity in medical images, resulting in two novel automatic methods: a non-parametric method and a parametric method. Our methods are easy to implement because they both solve an iteratively re-weighted least squares problem. They are remarkably accurate as shown by our experiments on images of different imaged objects and from different imaging modalities.

1 Introduction

Intensity nonuniformity is an artifact in medical images, perceived as a smooth variation of intensities across the image. It is also referred to as intensity inhomogeneity, shading or bias field. This artifact can be produced by different imaging modalities, such as magnetic resonance (MR) imaging, computer tomography (CT), X-ray, ultrasound, and transmission electron microscopy (TEM), etc. Although intensity inhomogeneity may not be noticeable to human observer, it can degrade many medical image analysis methods like segmentation, registration, and feature extraction, etc.

There exist different kinds of methods to correct the intensity nonuniformity in medical images. In the recent detailed review [1], they are classified into filtering based methods to remove image contents unrelated to the nonuniformity, surface fitting based methods based on the intensities of major tissues or the image gradients [2], segmentation based methods performed by iterating on the two processes of image segmentation and bias field's fitting [3,4], and image intensity histogram based methods through maximizing the high frequency of the image intensity distribution or minimizing the image entropy [5,6], etc.

In this paper, we propose to apply the sparseness property of the gradient probability distribution in medical images to the automatic estimation of the bias field. The sparse distribution is typically characterized by a high kurtosis and two heavy tails. From this prior knowledge, we obtain a non-parametric method and a parametric method. The two methods are simple to implement. They both

work by solving an iteratively re-weighted least squares (IRLS) problem. In each iteration, only a linear equations system needs to be solved. The two novel methods are also remarkably accurate, as shown by our results on simulated and real MR brain images, real CT lung images, and real TEM rabbit retina images.

2   Sparseness of Image Gradient Distribution

Recent research has shown that images of real-world scenes obey a sparse probability distribution in their gradients [7,8]. This sparseness property is extremely robust and characterized with a high kurtosis and two heavy tails in the gradient distribution. It stems from the assumption on the image that adjacent pixels have similar intensities unless separated by edges, i.e. the widely known piecewise constancy property of image.

This sparse distribution can be modeled with different ways, e.g. the exponential function [8], as expressed below

$$p(x) = e^{-|x|^{\alpha}} \tag{1}$$

where the parameter $\alpha < 1$ and can be fit from the gradient histogram.

We compute the image gradient histogram with the optimal bin size computed by the method in [9], and then use it to fit α in Eq. (1) with the maximum likelihood [10].

3   Methods

We use this sparseness property of gradient distribution in medical images to estimate the bias field in order to correct the intensity nonuniformity. To be concise, we provide the explanations of the 2D image case. However, the modifications to the 3D volume are straightforward.

3.1   Problem Definition

Considering the noise free case, a given 2D medical image Z is the product of the intensity nonuniformity free image I and a bias field B, as expressed below

$$Z(i,j) = I(i,j)B(i,j), \tag{2}$$

where (i,j) index pixel in the image, and if M and N represent the total numbers of rows and columns, respectively, we have $1 \leq i \leq M$ and $1 \leq j \leq N$.

In the process of correcting the intensity nonuniformity, our goal is to estimate B in Eq. (2) for each pixel. This is a classic ill-posed problem because the number of unknowns (I and B) is twice the number of equations. To make it solvable, we need to add constraints on both I and B.

Instead of computing B directly, we solve for its logarithm as in [8]. Let $\mathcal{Z} = \ln Z$, $\mathcal{I} = \ln I$, and $\mathcal{B} = \ln B$. Then, we have

$$\mathcal{Z}(i,j) = \mathcal{I}(i,j) + \mathcal{B}(i,j). \tag{3}$$

We denote the gradients of $\mathcal{Z}$, $\mathcal{I}$, and $\mathcal{B}$ for each pixel (i, j) by $\psi^{\mathcal{Z}}(i, j)$, $\psi^{\mathcal{I}}(i, j)$, and $\psi^{\mathcal{B}}(i, j)$, respectively. Then,

$$\psi^{\mathcal{Z}}(i, j) = \psi^{\mathcal{I}}(i, j) + \psi^{\mathcal{B}}(i, j). \tag{4}$$

Given an image $\mathcal{Z}$ with the intensity nonuniformity, we find a maximum a posteriori (MAP) solution to $\mathcal{B}$. Using Bayes' rule, this amounts to solving the optimization problem

$$\mathcal{B} = \arg\max_{\mathcal{B}} P(\mathcal{B}|\mathcal{Z}) \propto \arg\max_{\mathcal{B}} P(\mathcal{Z}|\mathcal{B}) P(\mathcal{B}). \tag{5}$$

Different specifications of $P(\mathcal{Z}|\mathcal{B})$ and $P(\mathcal{B})$ in Eq. (5) may lead to different algorithms to solve the bias field $\mathcal{B}$. The conditional probability $P(\mathcal{Z}|\mathcal{B})$ can be determined by some prior knowledge on the nonuniformity free image $\mathcal{I} = \mathcal{Z} - \mathcal{B}$, while $P(\mathcal{B})$ can be set from some prior information on the field $\mathcal{B}$.

To compute $P(\mathcal{Z}|\mathcal{B})$, we impose the sparseness prior of the image gradient distribution explained in Eq. (1) on $\mathcal{I}$ as below

$$P(\mathcal{Z}|\mathcal{B}) = P\left(\psi^{\mathcal{I}}\right) = e^{-\left|\psi^{\mathcal{I}}\right|^{\alpha}}, \quad \alpha < 1. \tag{6}$$

From Eq. (4), we have

$$\psi^{\mathcal{I}}(i, j) = \psi^{\mathcal{Z}}(i, j) - \psi^{\mathcal{B}}(i, j). \tag{7}$$

Substituting Eq. (7) into Eq. (6), yields

$$P(\mathcal{Z}|\mathcal{B}) = e^{-\sum_{(i,j)} \left|\psi^{\mathcal{Z}}(i,j) - \psi^{\mathcal{B}}(i,j)\right|^{\alpha}}. \tag{8}$$

To determine $P(\mathcal{B})$, there are basically two ways: non-parametric method and parametric method. Non-parametric method does not assume any model on $\mathcal{B}$ and estimate for each pixel while enforcing spatially local smoothness. $P(\mathcal{B})$ can be expressed explicitly by $\mathcal{B}$ values based on the smoothness constraints. Parametric method represents $\mathcal{B}$ with some model functions and estimate the parameters of the model. $P(\mathcal{B})$ is represented by the parameters of the model through enforcing some prior constraints on the model, or is eliminated when no prior constraint is necessary in the model.

3.2   Non-parametric Method

We impose a smoothness prior over $\mathcal{B}$:

$$P(\mathcal{B}) = e^{-\lambda_s \sum_{(i,j)} \left(\mathcal{B}_{xx}(i,j)^2 + \mathcal{B}_{yy}(i,j)^2\right)}, \tag{9}$$

where λ_s is a parameter (e.g. 0.8) determining how smooth the resulting $\mathcal{B}$ will be, and $\mathcal{B}_{xx}(i, j)$ and $\mathcal{B}_{yy}(i, j)$ are the second order derivatives at pixel (i, j) along the horizontal direction and vertical direction, respectively.
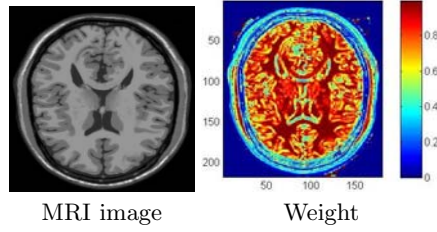
MRI image   Weight

Fig. 1. Computed weights (Eq. (13)) in the non-parametric method after the 3rd iteration of the IRLS algorithm, shown by color coding with the spectrum in the color bar (right).

The maximization of Eq. (5) becomes the the minimization of the below object function

$$\mathcal{O} = \sum_{(i,j)} \left| \psi^{\mathcal{Z}}(i,j) - \psi^{\mathcal{B}}(i,j) \right|^{\alpha} + \lambda_s \left(\sum_{(i,j)} \mathcal{B}_{xx}(i,j)^2 + \sum_{(i,j)} \mathcal{B}_{yy}(i,j)^2 \right). \quad (10)$$

The minimization of Eq. (10) does not have a closed-form solution because the exponent α in the first term is less than one as mentioned in Eq. (1). To solve it, most nonlinear optimization techniques [11] can be directly applied, however, they would be very time-consuming and easy to be trapped into a local minimum considering the large number of unknowns in Eq. (10) (it equals to the number of pixels). In order to get a good solution in less time to this minimization problem, we employ the iteratively re-weighted least squares (IRLS) technique [12,8]. IRLS poses the optimization as a sequence of standard least squares problems, each using a weight factor based on the solution of the previous iteration. Specifically, at the kth iteration, the energy function using the new weight can be written as

$$\mathcal{O} = \sum_{(i,j)} w_k(i,j) \left(\psi^{\mathcal{Z}}(i,j) - \psi^{\mathcal{B}}(i,j) \right)^2 + \lambda_s \left(\sum_{(i,j)} \mathcal{B}_{xx}(i,j)^2 + \sum_{(i,j)} \mathcal{B}_{yy}(i,j)^2 \right),$$
$$(11)$$

where weight $w_k(i,j)$ is computed in terms of the optimal $\mathcal{B}_{k-1}$ from the last iteration as

$$w_k(i,j) = e^{-S_1}(1 - e^{-S_2}),$$
$$S_1 = \left| \psi^{\mathcal{Z}}(i,j) - \psi^{\mathcal{B}_{k-1}}(i,j) \right|, \quad S_2 = \alpha (S_1)^{\alpha-1} \quad (12)$$

It is easy to see that minimizing $\mathcal{O}$ in Eq. (11) is to make the gradient of $\mathcal{B}$ equal to the gradient of $\mathcal{Z}$ at each pixel while enforcing spatially local smoothness on $\mathcal{B}$. However, it is well known that this kind of problem has no unique solution because adding a constant value on $\mathcal{B}$ will not influence $\mathcal{O}$ value. To tackle it, we add another constraint as shown in the new object function below

$$\mathcal{O} = \sum_{(i,j)} w_k(i,j) \left(\psi^{\mathcal{Z}}(i,j) - \psi^{\mathcal{B}}(i,j) \right)^2$$
$$+ \lambda_s \left(\sum_{(i,j)} \mathcal{B}_{xx}(i,j)^2 + \sum_{(i,j)} \mathcal{B}_{yy}(i,j)^2 \right) + \epsilon \sum_{(i,j)} \mathcal{B}(i,j)^2, \quad (13)$$

where ϵ is a very small value, e.g. 0.00001. This constraint makes $\mathcal{B}$ as small as possible (i.e. B as close to 1 as possible), leading $\mathcal{B}$ to a unique solution.

The object function in Eq. (13) is a standard least-squares problem. We can always obtain a closed-form optimal solution by solving a linear equations system. Details of the solution can be found in the supplementary file.

In our experiments, we initialize $\mathcal{B}(i,j) = 0$ (i.e. $B(i,j) = 1$) for all pixels (i,j), and find that it suffices to iterate three or four times to obtain satisfactory results. We also observed that the re-computed weights at each iteration k are higher at pixels whose gradients in $\mathcal{Z}$ are more similar to the ones in the estimated $\mathcal{B}_{k-1}$. Thus, the solution is biased towards smoother regions whose gradients are relatively smaller. Fig. 1 shows the weights recovered at the final iteration for an MR image.

3.3 Parametric Method

Different models are suitable to represent the nonuniformity field $\mathcal{B}$, considering its smoothly changing property, like the cubic B-splines [5], thin-plate splines, and the the bivariate polynomial, etc. We choose the bivariate polynomial for its efficiency shown by [2], for which the model in degree D (e.g. 4) is

$$\mathcal{B}(i,j) = \sum_{t=0}^{D}\sum_{l=0}^{t} a_{t-l,l} x_{(i,j)}^{t-l} y_{(i,j)}^{l} \tag{14}$$

where $\{a_{t-l,l}\}$ are parameters determining the polynomial, and $x_{(i,j)}$ and $y_{(i,j)}$ are the values of pixel (i,j) on the x-axis and y-axis, respectively. Note that the number of elements in $\{a_{t-l,l}\}$ is $(D+1)(D+2)/2$.

We apply the IRLS scheme used in the non-parametric method to estimating model parameters in the parametric method. Considering the fact that the model in Eq. (14) already incorporates the spatial smoothness on $\mathcal{B}$, the smoothness constraints on $\mathcal{B}$ values can be eliminated. The object function in each iteration of the IRLS is then written as

$$\mathcal{O} = \sum_{(i,j)} w_k(i,j) \left(\psi^{\mathcal{Z}}(i,j) - \psi^{\mathcal{B}}(i,j) \right)^2 + \epsilon \sum_{t=0}^{D}\sum_{l=0}^{t} a_{t-l,l}^2 \tag{15}$$

where we add the regularization on $\{a_{t-l,l}\}$ in order to get a unique solution as explained in Eq. (13), and the weights $w_k(i,j)$ are computed with Eq. (12).

The object function in Eq. (15) is also a standard least-squares problem relative to $\{a_{t-l,l}\}$, and its minimization also has a closed-form solution by solving a linear equations system. It can be seen from the fact that gradient on $\mathcal{B}$ is linear to $\{a_{t-l,l}\}$. More details of the solution can be found in the supplementary file. In our experiments, we initialize $\{a_{t-l,l}\}$ to zero, and it suffices for three or four times to obtain satisfactory results.

The proposed non-parametric and parametric methods are both easy to implement and run fast. The two methods iterate only three or four times on solving a weighted least square problem with the IRLS technique. The weighted least square problem is solved by resolving a linear equations system for which the solution can be represented by some operations of matrices. We note that the operations of sparse matrices are needed for solving Eq. (13).
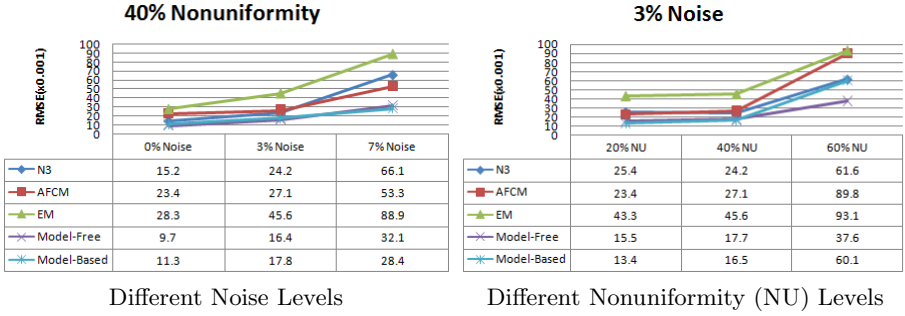
40% Nonuniformity

	0% Noise	3% Noise	7% Noise
N3	15.2	24.2	66.1
AFCM	23.4	27.1	53.3
EM	28.3	45.6	88.9
Model-Free	9.7	16.4	32.1
Model-Based	11.3	17.8	28.4

3% Noise

	20% NU	40% NU	60% NU
N3	25.4	24.2	61.6
AFCM	23.4	27.1	89.8
EM	43.3	45.6	93.1
Model-Free	15.5	17.7	37.6
Model-Based	13.4	16.5	60.1

Different Noise Levels   Different Nonuniformity (NU) Levels

Fig. 2. The statistics of Root Mean Squared Error (RMSE) in estimation of the bias field in the simulated MR brain data sets at different levels of noise and nonuniformity

4   Results

To implement our algorithms, the matrix operations in Eq. (13) were performed with the TAUCS [1], a library of sparse linear solvers.

We provide both quantitative evaluations with simulated data sets and visual evaluations with real data sets on our algorithms. For all experiments, we use $\alpha = 0.71$ in the sparse distribution model (Eq. (1)), that was estimated from 84 MR human brain images and 45 CT human lung images. These images were chosen to be free of the intensity nonuniformity guaranteed by visual inspections. In addition, we run our algorithms on a lower resolution image down-sampled from the given image and then reconstruct the resulting bias field to the original size by interpolation, as in [5]. In the experiments, the gradients along the two axis for 2D image and the three axes for 3D volume are all used.

4.1   Quantitative Evaluation

We test our algorithms on the 3D MR volumes obtained from the BrainWeb Simulated Brain Database[2], for which the ground truth of the bias field is known. The Root Mean Squared Error (RMSE) between the estimation and the ground truth is computed. The results of our algorithms are compared with the widely used N3 [5], AFCM [3], and EM based method [4].

The simulated data sets are obtained with the following settings: T1, T2 and PD modalities, slice thickness of 1 mm, 0%, 3% and 7% noise levels, and 20% and 40% intensity nonuniformity levels. As a preprocessing, the extra-cranial tissues were removed from all 3D volumes according to the ground-truth memberships of the tissues provided on the BrainWeb website. In order to get data sets with more severer intensity nonuniformity effects, we constructed the 60% intensity nonuniformity level by linearly scaling the range values of the ground truth to 0.70 ... 1.30 as explained on the website, and then enforcing it on the volumes download with the different noise levels and with 0% intensity nonuniformity.

---

[1] http://www.tau.ac.il/ stoledo/taucs/
[2] http://www.bic.mni.mcgill.ca/brainweb/

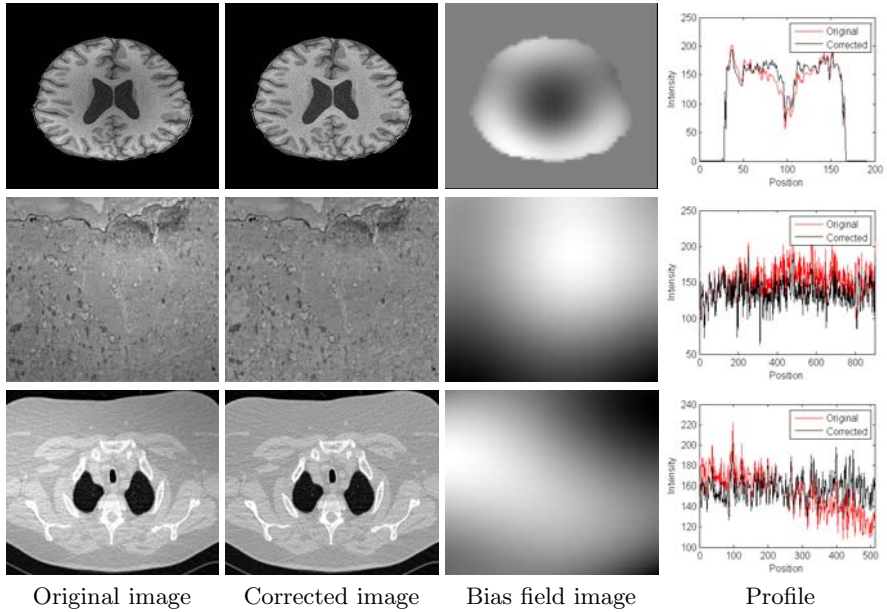| Original image | Corrected image | Bias field image | Profile |

Fig. 3. Corrections of the bias field by our non-parametric method on one MR brain image (up), one TEM image (middle) from rabbit retina, and one CT lung image (down). The profiles are drawn on a horizontal line of the image.

Before computing the RMSE statistics, the multiplicative factor [5] is removed from the resulting bias field by minimizing the mean square distance between the result and the ground truth. Therefore, only the shape differences account for the errors.

We found that the chosen methods all perform very similarly on different imaging modalities. Therefore, we averaged the RMSE statistics over the three modalities. From the results shown in Fig. 2, we can see that our methods can improve the estimation accuracies relative to the standard methods. Moreover, our methods seem more robust to noise. Compared between the two new methods, the parametric method resists noise better but may produce larger errors when the nonuniformity is very severer. It is because a severer bias field may go beyond the representation ability of the model in Eq. (14).

4.2   Visual Evaluation

We also run our algorithms on real data: 9 MR brain volumes, 2 TEM images from the rabbit retina, and 5 CT lung volumes, for which the intensity nonuniformity artifacts can be visually perceived. Due to the lacking of the ground truth of the bias fields, we evaluate the results by visual inspections through observing the intensity profile on selected lines.

We found that both of our two new methods can efficiently correct the bias field in images from different modalities and of different imaged objects, resulting in flatter profiles. We put some results by the non-parametric method in Fig. 3.

5   Conclusion and Future Work

Based on the sparseness of the gradient distribution in medical images, we proposed a non-parametric approach and a parametric approach for the automatic correction of the intensity nonuniformity. They are easy to implement and remarkably accurate. Similar strategy has already been used successfully vignetting correction in [8].

Considering the fact that the sparseness property can be treated as a robust prior knowledge of an ideal image or even an ideal deformation field, our paper may also inspire several more works following the line of using this property in medical image inpainting, image segmentation, and image registration.

References

1. Vovk, U., Pernus, F., Lika, B.: A review of methods for correction of intensity inhomogeneity in MRI. IEEE Transactions on Medical Imaging 26(3), 405–421 (2007)
2. Tasdizen, T., Jurrus, E., Whitaker, R.T.: Non-uniform illumination correction in transmission electron microscopy. In: MICCAI Workshop on Microscopic Image Analysis with Applications in Biology (2008)
3. Pham, D.L., Prince, J.L.: Adaptive fuzzy segmentation of magnetic resonance images. IEEE Transactions on Medical Imaging 18(9), 737–752 (1999)
4. Zhang, Y., Smith, S., Brady, M.: Hidden markov random field model and segmentation of brain MR images. IEEE Transactions on Medical Imaging 20, 45–57 (2001)
5. Sled, J.G., Zijdenbos, A.P., Evans, A.C.: A nonparametric method for automatic correction of intensity nonuniformity in MRI data. IEEE Transactions on Medical Imaging 17(1), 87–97 (1998)
6. Styner, M., Brechbuhler, C., Szekely, G., Gerig, G.: Parametric estimate of intensity inhomogeneities applied to MRI. IEEE Transactions on Medical Imaging 19, 153–165 (2000)
7. Olshausen, B.A., Field, D.J.: Emergence of simple-cell receptive field properties by learning a sparse code for natural images. Nature 381, 607–609 (1996)
8. Zheng, Y., Yu, J., Kang, S.B., Lin, S., Kambhamettu, C.: Single-image vignetting correction using radial gradient symmetry. In: CVPR (2008)
9. Shimazaki, H., Shinomoto, S.: A method for selecting the bin size of a time histogram. Neural Computation 19(6), 1503–1527 (2007)
10. Aldrich, J.: R.a. fisher and the making of maximum likelihood 1912-1922. Statistical Science 12(3), 162–176 (1997)
11. Nocedal, J., Wright, S.J.: Numerical Optimization. Springer, Heidelberg (2006)
12. Meer, P.: Robust techniques for computer vision, pp. 107–190. Prentice-Hall, Englewood Cliffs (2005)