

# A Fast and Automatic Method to Correct Intensity Inhomogeneity in MR Brain Images

Zujun Hou<sup>1,\*</sup>, Su Huang<sup>2</sup>, Qingmao Hu<sup>2</sup>, and Wieslaw L. Nowinski<sup>2</sup>

<sup>1</sup> Dept. of Interactive Media, Institute for Infocomm Research, Singapore 119613

<sup>2</sup> Biomedical Imaging Lab, Singapore Bioimaging Consortium, Singapore 138671

**Abstract.** This paper presents a method to improve the semi-automatic method for intensity inhomogeneity correction by Dawant et al. through introducing a fully automatic approach to reference points generation, which is based on order statistics and integrates information from the fine to coarse scale representations of the input image. The method has been validated and compared with two popular methods, N3 and BFC. Advantages of the proposed method are demonstrated.

## 1 Introduction

Intensity inhomogeneity (IIH) inherent in Magnetic Resonance (MR) images can substantially reduce the accuracy of segmentation and registration. With the advent of multichannel phased array coils and 3T units, the importance and prevalence of this problem is greatly increased, and a near-real time IIH correction is of primary importance.

A wealth of techniques has been devised to correct this artifact. The correction has been dealt prospectively by phantom scanning [1] or theoretical modeling [2,3]. More often, the issue is addressed retrospectively by deriving the solution from image data alone. Most of the mathematical modeling methods for IIH correction fall into the following three categories: low-pass filtering [4,5], surface fitting [6,7] and statistical modeling [8,9]. For a latest review, the interested readers are referred to [10].

In [11], a comprehensive comparison among six algorithms has been made and concludes that none of the six methods can perform ideally under all experiments that were carried out. In this paper, we present an improvement of the semi-automatic technique by [6], which fits a surface model among a few reference points selected from the white matter (WM). Sled et al. [12] has compared this simple method with the N3 [9] and the expectation maximization (EM) [8] method using simulated T1, T2 and PD weighted images. It is found that this method is comparable with the N3 method and superior to the EM method. In particular, its performance is encouraging in T1 weighted images. Nevertheless, the requirement of expert supervision on reference points generation significantly reduces the efficiency of this method. As an attempt to enhance this simple

---

\* The work is mainly done during his stay with the Biomedical Imaging Lab of Singapore Bioimaging Consortium.

method, an approach is proposed in this paper to automatically generate the reference points. The proposed method is demonstrated to be as fast as low-pass filtering methods, whereas as accurate as two popular methods, the N3 and the BFC method.

The rest of the paper is organized as follows. The method is described in Section 2, where an approach based on order statistics to automatically generating the reference points is detailed. Validation of the method as well as comparison with state-of-art methods are presented in Section 3. Finally, the paper concludes in Section 4.

## 2 An Approach to Automatic Reference Points Generation

In this section, we shall present an approach to automatic generation of the reference points. The method is based on order statistics and integrate information from fine to coarse scale.

Similar to many other IHH correction methods, the input of the system is a brain MR image with the skull/scalp removed. Several public domain software packages are available, such as the BrainSuite package (<http://neuroimage.usc.edu/>). Fig. 1 Row 1 shows an example, where on the left is an original scan and in the middle is the tissue after skull/scalp removal. The subsequent operations will be limited to the region with the tissue present.

To offset the impact of noise and partial volume effect, the reference points will be selected from a low-resolution image and the intensities will be determined from a high-resolution image. It is assumed that a point in the low-resolution image corresponds to a block of points in the high-resolution image. Thus, the identification of reference points in the low-resolution image is equivalent to identifying reference blocks in the high-resolution image.

### 2.1 Within-Block Homogeneity and the Trimmed Range Statistic

The first problem is to test the following hypothesis:

$$H_0: \text{For a given block, there is only one type of tissue.}$$

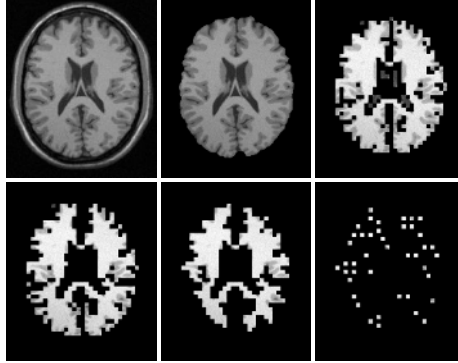
The test is based on order statistics as described below.

The image is raster-scanned and the *trimmed range* statistic is calculated block by block. Within a block, suppose that there are  $L$  tissue pixel/voxels. Firstly, let rank the pixel/voxels in a non-decreasing order and denote by  $x_{\{l\}}, l = 1, \dots, L$ . Next, the samples are normalized as follows:

$$\tilde{x}_{\{i\}} = x_{\{i\}} / (2x_{\{L/2\}}). \quad (1)$$

Let  $1 < j < k < L$ , then the trimmed range statistic  $r_{j,k}$  is computed by

$$r_{j,k} = \tilde{x}_{\{k\}} - \tilde{x}_{\{j\}}. \quad (2)$$



**Fig. 1.** An illustration of basic steps of the proposed method. Row 1: original image (left), result after skull/scalp removal (middle) and result after within block homogeneity testing (right); Row 2: result after inter-block homogeneity testing (left), result after between block homogeneity testing (middle) and result after downsampling (right).

If  $r_{j,k} < T_{range}$  for a given threshold  $T_{range}$ , then  $H_0$  will be accepted, i.e., the block is dominated by one tissue and its center is considered as a reference point candidate.

To make the statistic tolerant with possible outliers due to noise, tissue abnormality, limited tissue voxels of other types, etc., the values at two ends are excluded in computation, which is similar to the trimmed mean in deriving a robust mode estimate from samples. Readers interested in robust statistics as well as its application to image processing can refer to [13] for more detail. For the image in Row 1, Fig. 1, the result after this step is shown in Row 1 (right).

## 2.2 Inter-blocks' Median Homogeneity

Since the IIH field is slowly varying, it is reasonable to assume continuity between the *median* of a block and those of its neighboring blocks. The constraint on median's continuity analyzes the image in the neighborhood of block, which in scale space is coarser than the scale characterized by block size. In this coarser scale, it is assumed that the variation of IIH is also small, thus, oscillation between the the median of a block with respect to those of its neighbors would be attributed to the intensity variation of tissues: small variation means a possibly common tissue type and large oscillation indicates different tissue types, the former of which will be preserved and the latter will be suppressed by a thresholding.

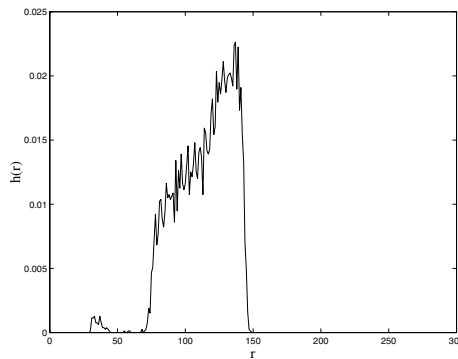
For the sample image in Fig. 1, the result after inter-block homogeneity testing is displayed in Row 2 (left) of the same figure. Compared with the image in Row 1 (right), some dark blocks adjacent to bright blocks are removed. It is noted that there exist some blocks which look “grey” and adjoin some bright blocks. Examination of the neighboring slices reveals that most voxels in these blocks look pretty “white”, thus, the actual median oscillation within these blocks is not as significant as what appears in this slice.

### 2.3 Between-Block Homogeneity

As aforementioned, a single candidate block can be dominated by a single type of tissue, but different blocks may contain tissue of different types. For the block candidates after last step, large oscillation in neighboring blocks could have been suppressed, and separated blocks with distinct tissue types could have been kept. To identify these tissue types and extract a sample of candidate blocks with the same type of tissue, one can resort to examining the distribution of these blocks' medians.

To extract the primary component, one can use a trimming procedure, which has been used in the previous section to test the tissue homogeneity within a block. Given a median histogram, let  $h(i)$ ,  $i = 0, \dots, 255$  denote the frequency of gray level. The  $Pr_L$  percentile to the left is defined by  $Pr_L = \sum_0^{T_{modeL}} h[i]$  and similarly the  $Pr_R$  percentile to the right is  $Pr_R = \sum_{T_{modeR}}^{255} h[i]$ . Then the mode analysis is to determine the range  $(T_{modeL}, T_{modeR})$  for given parameters  $Pr_L$  and  $Pr_R$ . In this investigation, the parameters are fixed:  $Pr_R = 0, Pr_L = 0.5$  for T1 weighted images, and  $Pr_R = 0.1, Pr_L = 0$  for T2 or PD weighted images.

As an instance, the median's histogram of the example in Fig. 1 is shown in Fig. 2. After thresholding, the blocks with median falling outside the range are removed. The remained blocks are displayed in Fig. 1 Row 2 (middle). Clearly, most blocks belong to WM and the IIH effect is also visible in the extracted blocks.



**Fig. 2.** The histogram of blocks' median for the example in Fig. 1 after inter-block homogeneity testing

### 2.4 Downsampling

After the above processing, the identified blocks usually constitutes a large portion of a contiguous tissue. However, it would suffice to approximate the smooth IIH function with a few sampled points. Then the number of blocks can be further reduced by downsampling the blocks, which can also speed up the computation for data regression in the following section. This process starts with a raster-scan, and if a candidate block is encountered, then its immediate neighboring (26-neighborhood in 3D) candidate blocks will be deleted. On the bottom right

of Fig. 1 is the result after this step. It should be pointed out that the image is 3D and only one slice is shown in this picture. Many blocks in the lower left are deleted because there are neighboring blocks that are preserved in neighboring slices.

### 3 Experiments and Results

To evaluate the performance of the proposed method, a number of experiments have been carried out using simulated and real scans. The simulated data are from the normal brain database, BrainWeb [14], The simulated sequences used in this study are T1, T2 and PD-weighted, with noise varying from 0 to 9% and IIH in 0, 20% and 40% level. The images are of size  $181 \times 217 \times 181$ , 8-bit per voxel and 1 mm slice thickness.

The proposed method has been compared with respect to N3 [9] and BFC [15], The parameters for both methods are tuned according to those suggested in the software, so that the results are consistent with those published by other groups. To quantify the performance of an IIH correction algorithm, the commonly used criteria, the coefficient of variation ( $cv$ ) and the coefficient of joint variation ( $cjv$ ) [7], are adopted in this study. As aforementioned, it is assumed that the presence of IIH increases the tissue variance, thus it is expected that an algorithm attaining smaller value of  $cv$  or  $cjv$  would be better.

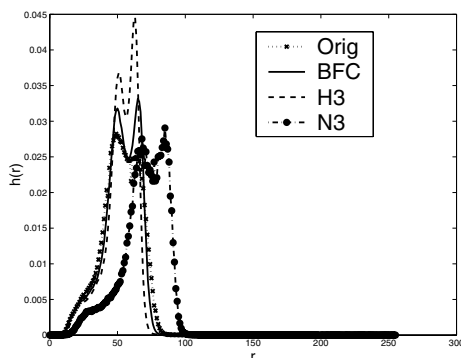
Table 1 shows the  $cv$  of GM, the  $cv$  of WM and the  $cjv$  between WM and GM by the three methods testing on BrainWeb data sets, where some outliers are highlighted. For T2 and PD weighted images, only results with 3% noise are displayed. One can observe that the performance of H3 is generally comparable

**Table 1.** Summary on correction of BrainWeb data sets, where some outliers are highlighted and the noise level for T2 or PD weighted images is 3%

		$cv(\text{GM})(\%)$				$cv(\text{WM})(\%)$				$cjv(\text{GM,WM})(\%)$			
IIH		Orig.	N3	BFC	H3	Orig.	N3	BFC	H3	Orig.	N3	BFC	H3
0	0	9.05	9.03	9.09	9.02	4.28	4.29	4.20	4.20	44.85	44.95	44.90	44.98
	20	10.25	9.07	9.14	9.04	6.05	4.30	4.31	4.32	52.20	45.14	44.85	45.26
	40	12.89	9.11	<b>10.23</b>	9.11	8.78	4.43	<b>5.38</b>	4.42	66.02	45.56	<b>50.32</b>	46.75
3	0	9.91	9.91	9.96	9.9	5.22	5.18	5.16	5.13	51.56	51.81	51.84	52.30
	20	11.01	9.91	10.01	9.91	6.74	5.20	5.25	5.29	57.05	51.67	51.29	52.03
	40	13.48	9.93	<b>11.34</b>	10.07	9.24	5.24	<b>6.4</b>	5.57	69.33	51.67	<b>57.76</b>	53.29
9	0	15.06	15.28	15.15	15.05	10.12	10.23	10.14	10.08	88.04	<b>96.35</b>	89.13	89.71
	20	15.70	15.45	15.16	15.03	10.76	10.09	10.00	9.99	86.64	<b>98.00</b>	86.30	87.66
	40	17.40	15.45	15.56	15.05	12.29	9.86	10.18	9.83	91.76	<b>96.47</b>	84.67	88.79
T2	0	17.71	17.72	18.14	17.90	11.95	11.93	12.01	11.87	83.19	83.58	87.79	85.85
	20	18.44	18.23	18.80	18.45	12.03	11.89	12.08	11.75	90.88	89.28	94.71	90.30
	40	20.30	19.81	20.63	19.47	13.32	12.93	13.75	12.13	106.39	102.81	111.83	99.03
PD	0	4.49	4.77	5.52	5.32	4.88	4.88	5.89	5.17	64.89	71.29	76.38	82.68
	20	6.27	4.81	5.15	5.58	5.49	4.78	4.97	5.11	94.97	71.96	74.24	<b>86.07</b>
	40	9.72	4.95	6.24	5.69	7.50	4.76	5.05	5.13	163.00	74.16	91.38	87.42

with that of N3 or BFC. It is also noted that the findings here are consistent with those in [12], where the Dawant et al. method [6] is compared with N3. Thus, it can be concluded that at least for BrainWeb data set the automatically generated reference points approximate the IIH map very closely with that by the manual selection.

Since N3 and H3 are run in the same platform, the system CPU time by the two methods has been compared. For the images tested above, the average CPU time by H3 is 0.9 second with deviation 0.03 second, which indicates that H3 can attain nearly real time IIH correction in spite of data quality. Correspondingly, N3 runs on average in 17 seconds with deviation about 33 seconds. Including time for file loading and output, H3 executes generally in less than 2 seconds.

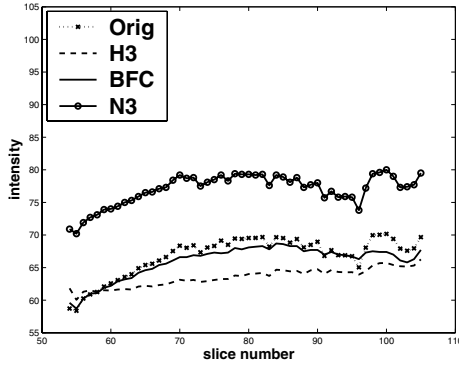


**Fig. 3.** Comparison of histograms on a real scan. The flat histogram due to IIH before correction and the effect of histogram sharpening by the three methods are clearly visible.

Besides the simulated images, the proposed method has been applied to about 21 real scans affected by IIH artifact. The data are acquired from Singapore and Japan using T1-weighted spin echo (SE) and spoiled gradient recalled echo (SPGR) sequences. The voxel size ranges from  $0.898 \times 0.898 \times 1.238 \text{ mm}^3$  to  $1.0 \times 1.0 \times 3.5 \text{ mm}^3$ . In general, image contrast can substantially be improved by the three IIH correction methods, and their effects are very similar.

Due to space limit, only one example is shown here. Fig. 3 plots the histogram of a volume (star dotted line), where the overlapping between the GM and the WM components is obvious, as results from the increase of tissue variances (the variance of GM and the variance of WM) due to IIH. In the histograms after IIH correction by three methods, BFC (solid line), H3 (dashed line) and N3 (circle dash-dotted line), the reduction on tissue variances is visible and the multi-modal shape for the tissue distribution is recovered.

As an attempt to quantify the correction for real data, a number of WM regions in the corona radiata and corpus callosum have been segmented manually on 52 slices for this volume. Fig. 4 plots the variation of segmented WM mean value over slices (from superior to inferior) before and after IIH correction by



**Fig. 4.** Variation of segmented white matter mean value vs slices (from superior to inferior) before (star dotted line) and after IIIH correction by H3 (dashed line), BFC (solid line) and N3 (circle solid line). H3 attains minimal reduction on the mean value variation across slices due to intensity inhomogeneity.

three methods. After IIIH correction by the three methods, one can see that there are little improvement on the variation pattern by N3 (circle solid line) or BFC (solid line), whereas the variation becomes much smoother by H3 correction (dash-dotted line).

## 4 Conclusion

We have presented a method to correct intensity inhomogeneity arising in brain MR images. The method automatically identifies the reference points in a coarse scale based on order statistics information, followed by surface fitting and inhomogeneity correction. It has been validated using simulated T1, T2 and PD weighted images as well as real scans. Experiments turn out that the presented method is capable of achieving intensity inhomogeneity correction in nearly real time as realized by the simple low-pass filtering methods, but with accuracy comparable with more complex methods such as N3 and BFC.

## References

1. Wicks, D. A. G., Barker, G. J. and Tofts, P. S.: Correction of intensity nonuniformity in MR images of any orientation. *Magnetic Resonance Imaging* 11 (1993) 183-196.
2. Alecci, M., Collins, C. M., Smith, M. B. and Jezzard, P.: Radio frequency magnetic field mapping of a 3 Tesla birdcage coil: experimental and theoretical dependence on sample properties. *Magnetic Resonance in Medicine* 46 (2001) 379-385.
3. Sutton, B. P., Noll, D. C. and Fessler, J. A.: Fast, iterative image reconstruction for MRI in the presence of field inhomogeneities. *IEEE Trans. Medical Imaging* 22 (2003) 178-188.

4. Brinkmann, B. H., Manduca, A. and Robb, R. A.: Optimized homomorphic unsharp masking for MR grayscale inhomogeneity correction. *IEEE Trans. Medical Imaging* 17 (1998) 161-171.
5. Gispert, J. D., Reig, S., Pascau, J., Vaquero, J. J., Garcia-Barreno, P. and Desco, M.: Method for bias field correction of brain T1-weighted magnetic resonance images minimizing segmentation error. *Human Brain Mapping* 22 (2004) 133-144.
6. Dawant, B. M., Zijdenbos, A. P. and Margolin, R. A.: Correction of intensity variations in MR images for computer-aided tissue classification. *IEEE Trans. Medical Imaging* 12 (1993) 770-781.
7. Likar, B., Viergever, M. A. and Pernus, F.: Retrospective correction of MR intensity inhomogeneity by information minimization. *IEEE Trans. Medical Imaging* 20 (2001) 1398-1410.
8. Wells, III, W. M., Grimson, W. E. L., Kikinis, R. and Jolesz, F. A.: Adaptive segmentation of MRI data. *IEEE Trans. Medical Imaging* 15 (1996) 429-442.
9. Sled, J. G., Zijdenbos, A. P. and Evans, A. C.: A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans. Medical Imaging* 17 (1998) 87-97.
10. Hou, Z.: A review on MR image intensity inhomogeneity correction, *International Journal of Biomedical Imaging*, special issue on mathematical modeling (in press).
11. Arnold, J. B., Liow, J., Schaper, K. A., Stern, J. J., Sled, J. G., Shattuck, D. W., Worth, A. J., Cohen, M. S., Leahy, R. M., Mazziotta, J. C. and Rottenberg, D. A.: Qualitative and quantitative evaluation of six algorithms for correcting intensity nonuniformity effects. *NeuroImage* 13 (2001) 931-943.
12. Sled, J. G., Zijdenbos, A. P. and Evans, A. C.: A comparison of retrospective intensity non-uniformity correction methods for MRI. *Lecture Notes in Computer Science* 1230 (1997) 459-464.
13. Hou, Z. and Koh, T. S.: Robust edge detection. *Pattern Recognition* 36 (2003) 2083-2091.
14. Collins, D. L., Zijdenbos, A. P., Sled, J. G., Kabani, N. J., Holmes, C. J. and Evans, A. C.: Design and construction of a realistic digital brain phantom. *IEEE Trans. Medical Imaging* 17 (1998) 463-468.
15. Shattuck, D. W., Sandor-Leahy, S. R., Schaper, K. A., Rottenberg, D. A. and Leahy, R. M.: Magnetic resonance image tissue classification using a partial volume model. *NeuroImage* 13 (2001) 856-876.