Published in final edited form as:

Magn Reson Imaging. 2007 February ; 25(2): 259–264.

A novel, fast entropy-minimization algorithm for bias field correction in MR images

Qing Ji^a, John O. Glass^a, and Wilburn E. Reddick^{a,b,*}

^aDivision of Translational Imaging Research, Department of Radiological Sciences (MS 210), St. Jude Children's Research Hospital, Memphis, TN 38105-2794, USA

^bDepartments of Electrical and Computer Engineering and Biomedical Engineering, University of Memphis, Memphis, TN 38152, USA

Abstract

A novel, fast entropy-minimization algorithm for bias field correction in magnetic resonance (MR) images is suggested to correct the intensity inhomogeneity degradation of MR images that has become an increasing problem with the use of phased-array coils. Four important modifications were made to the conventional algorithm: (a) implementation of a modified two-step sampling strategy for stacked 2D image data sets, which included reducing the size of the measured image on each slice with a simple averaging method without changing the number of slices and then using a binary mask generated by a histogram threshold method to define the sampled voxels in the reduced image; (b) improvement of the efficiency of the correction function by using a Legendre polynomial as an orthogonal base function polynomial; (c) use of a nonparametric Parzen window estimator with a Gaussian kernel to calculate the probability density function and Shannon entropy directly from the image data; and (d) performing entropy minimization with a conjugate gradient method. Results showed that this algorithm could correct different types of MR images from different types of coils acquired at different field strengths very efficiently and with decreased computational load.

Keywords

Magnetic resonance image; Inhomogeneity correction; Entropy minimization; Parzen window; Conjugate gradient method

1. Introduction

The measured magnetic resonance (MR) image is usually degraded by a bias field or intensity inhomogeneity (nonuniformity), which is induced primarily by the sensitivity profile of the radio frequency coil. The bias field is characterized by multiplicative smooth spatial variations that modulate the intensity of the true image data. Degradation resulting from these modulations in the intensity bias field presents a common problem for intensity-based imaging analysis techniques such as image segmentation and registration. Because independent measurement of the bias field is very difficult and time-consuming, most of the reported correction techniques are postprocessing or retrospective methods, in which the bias field is estimated from the image itself after acquisition. The separation of the bias field from the measured image data is generally a blind separation problem since the “true” image and the bias field are both unknowns. Previously reported methods have modeled the bias field, the true bias-free image

*Corresponding author. Division of Translational Imaging Research, Department of Radiological Sciences (MS 210), St. Jude Children's Research Hospital, Memphis, TN 38105-2794, USA. Tel.: +1 901 495 3270; fax: +1 901 495 5706. E-mail address: gene.reddick@stjude.org (W.E. Reddick).

or both by using some a priori knowledge of the bias field and bias-free image [1-7]. In entropy-minimization methods, the bias field is modeled as a polynomial with the coefficients estimated by minimizing the Shannon entropy without any assumption about the measured image and with little alteration of the “true” image data [8,9].

In this article, we describe another approach to minimize the Shannon entropy. We use the framework of Likar et al. [9] as a basis to further describe the Parzen window estimation of the probability distribution function (PDF) [10], the calculation of the derivatives of the entropy and the gradient-based entropy-minimization strategy [11]. Finally, we present the correction results for the simulated images (via the Brain Web software) and clinical images from pediatric oncology patients.

2. Methods

2.1. Correction strategy

Likar's entropy-minimization correction strategy is summarized here. The correction model can be expressed as:

$$u(x, y, z) = B(x, y, z) v(x, y, z) + n(x, y, z). \quad (1)$$

Here, $u(x, y, z)$ and $v(x, y, z)$ are the corrected and measured images, respectively, $B(x, y, z)$ is the correction function and $n(x, y, z)$ is the noise. In Eq. (1), the correction function compensates for the bias field. In the noise-free case, the last term in Eq. (1) can be ignored. Because the bias field is slowly varying in the image space, the correction function can also be modeled as an n th-order polynomial:

$$B(x, y, z) = \sum_{i,j,k} a_{i,j,k} x^i y^j z^k, \quad (2)$$

where $a_{i,j,k}$ ($i, j, k=0, 1, \dots, n$) are the coefficients of the polynomial and x , y and z are the image spatial coordinates. The coefficients $a_{i,j,k}$ are optimized to minimize the Shannon entropy under the constraints of a meanspreserving condition. The Shannon entropy of the corrected image can be expressed as

$$H = - \sum_u p(u) \log p(u), \quad (3)$$

where $p(u)$ is the PDF of the corrected image. The constraint applied on the $B(x, y, z)$ is the means-preserving condition, which can be written as

$$\frac{1}{N} \sum_{x,y,z} v(x, y, z) = \frac{1}{N} \sum_{x,y,z} B(x, y, z) v(x, y, z), \quad (4)$$

where N is the total number of voxels. Using the condition in Eq. (4), the correction function can be simplified as

$$B(x, y, z) = 1 + \sum_{i,j,k} a_{i,j,k} \left(f_{i,j,k}(x, y, z) - \bar{f}_{i,j,k} \right), \quad (5)$$

where $i, j, k=1, \dots, n$ and

$$f_{i,j,k}(x, y, z) = x^i y^j z^k. \quad (6)$$

The methods that involve tissue classification take advantage of a priori knowledge of tissue information in the image.

$$\bar{f}_{i,j,k} = \frac{1}{N} \sum_{x,y,z} f_{i,j,k}(x,y,z) v(x,y,z). \quad (7)$$

The PDF $p(u)$ in Eq. (3) is calculated by evaluating the corrected image's histogram $h(u)$:

$$p(u) = \frac{h(u)}{N}. \quad (8)$$

The histogram is smoothed by the interpolation

$$h(u) \leftarrow \sum_{i=-t}^{i=t} h(u+i) (t+1-|i|), \quad (9)$$

where the parameter t defines the size of a triangular window. Because the entropy defined in Eq. (3) is not an explicit function of the parameters $a_{i,j,k}$, entropy minimization is performed via Powell's multidimensional direction set method [11], which is a minimization algorithm based on a function value.

2.2. Gradient-based entropy-minimization algorithm

The algorithm outlined above has some disadvantages. Due to lack of entropy gradient information, the minimization is performed via Powell's multidimensional direction set method, which is not very efficient for this purpose, especially when the number of parameters is large. Powell's method requires successive line minimizations along all the directions of a multidimensional space to determine the next direction. If the entropy gradient can be calculated, this step is not necessary; the next minimization direction can be determined by the gradient calculation.

To calculate the gradient of the Shannon entropy, we used a Parzen window method to approximate the corrected image's PDF. We originally used this method to estimate the PDF from the limited sampling of a random variable. If we consider the image data to be "random," the PDF of the image can be approximated as

$$\begin{aligned} p(u) &= \frac{1}{N} \sum_{x,y,z} g_{\sigma}(u - u(x,y,z)) \\ &= \frac{1}{N} \sum_{x,y,z} g_{\sigma}(u - B(x,y,z) v(x,y,z)), \end{aligned} \quad (10)$$

where σ is the width parameter of g_{σ} , a kernel function. A Gaussian kernel function is used to take the derivative of the PDF and entropy. The Gaussian kernel can be expressed as

$$g_{\sigma}(x) = \frac{1}{\sqrt{\pi}\sigma} e^{-\frac{x^2}{\sigma^2}}, \quad (11)$$

where the partial derivatives of the PDF to the correction coefficients $a_{i,j,k}$ can be calculated:

$$\begin{aligned} \frac{\partial p(u)}{\partial a_{i,j,k}} &= \frac{1}{N} \sum_{x,y,z} g'_{\sigma}(u - B(x,y,z) v(x,y,z)) \\ &\quad \times \left(f_{i,j,k} - \bar{f}_{i,j,k} \right) v(x,y,z). \end{aligned} \quad (12)$$

The partial derivatives of the entropy H to $a_{i,j,k}$ can then be calculated as

$$\frac{\partial H}{\partial a_{i,j,k}} = - \sum_u (\log -x p(u) + 1) \frac{\partial p(u)}{\partial a_{i,j,k}}. \quad (13)$$

Eqs. (10), (12) and (13) can be computed directly from the image data; therefore, the entropy minimization can be performed using the gradient-based minimization algorithm. We implemented a multidimensional conjugate gradient method, an iterative algorithm that uses

the gradient information to determine the direction of line minimization. For each line, the minimization can be accelerated with the use of an extra first-derivative calculation.

2.3. Improved correction function

Using the base function of the polynomials in Eq. (6) is not very efficient. The polynomials are nonorthogonal and not fully independent; they are also limited in their ability to model complicated bias fields. We can improve its efficiency by using an orthogonal base function polynomial such as the one-dimensional Legendre polynomial of degree l , which has the base function in the form of

$$P_l(x) = \frac{1}{2^l l!} \frac{d^l}{dx^l} \left\{ (x^l - 1)^l \right\}, -1 \leq x \leq 1. \quad (14)$$

The 3D form base function can be written as

$$f_{i,j,k}(x, y, z) = P_i(x) P_j(y) P_k(z). \quad (15)$$

Each dimension of the image has to be scaled to $(-1, 1)$ to fully use the function range in Eq. (14). The use of the Legendre polynomial does not necessarily increase the computation time because the value of Eq. (14) for each image grid can be calculated and saved in a lookup table.

2.4. Sampling strategy

Optimized sampling of the image voxel data can greatly reduce the computation load. Because the number of parameters used to model the bias field is much less than the number of voxels in the measured image, it is not necessary to use all the measured voxels to estimate the bias field. However, the voxels used have to be uniformly distributed in the image space because the bias field is not localized. The methods that involve tissue classification take advantage of a priori knowledge of tissue information in the image. In reality, the number of voxels used depends on the complexity of the bias field and the measured image. We proposed a two-step sampling strategy for a 3D image data set:

1. Reduce the size of the measured image on each slice with a simple averaging method without changing the number of slices.
2. Use a binary mask generated by a histogram threshold method to define the sampled voxels in the reduced image.

The reason for keeping the same number of slices is that the patient's clinical image data are usually acquired as a stack of 2D images rather than as a 3D isotropic image. These 2D images have a limited number of slices with relatively large slice thickness and, usually, with a gap between slices. This strategy samples much more than 5000 voxels; however, the improved minimization technique can compensate for the extra computation time.

2.5. Algorithm implementation

Our purpose was not to fully validate the entropy-minimization algorithm, which has been previously done. Instead, we focused on improving its efficiency. We implemented the algorithm described above on an SGI Fuel Workstation (Silicon Graphics Incorporated, Mountain View, CA) by using the C programming language. A fifth-order Legendre polynomial with 55 coefficients was used to represent the correction function. We calculated the image's PDF by using a Parzen window approximation method with a Gaussian kernel that has a width of 8.0. The entropy-minimization algorithm was a multidimensional conjugate gradient method in which the line minimization procedure uses the first derivative of the entropy. We did not scale down the images' gray levels to 256. As a basis of comparison, we also implemented Powell's direction set entropy-minimization algorithm.

2.6. Validation strategy — image data sets

Four types of imaging data sets were analyzed to assess the validity of the current correction algorithm: simulated whole head, clinical 2D single section and clinical whole-head T_1 -, T_2 - and PD-weighted images from both a 1.5-T magnet using a CP head coil and a 3.0-T magnet using an 8-channel phased-array head coil.

We generated and downloaded simulated images from the Brain Web software [12]. They included brain T_1 -, T_2 - and PD-weighted images with 40%, 20% and 0% bias fields. We also downloaded the segmented brain white matter and gray matter maps. Each downloaded image has a resolution of $181 \times 217 \times 181$ and a depth of 12 bits. For each image, we chose 100 consecutive slices with substantial white matter and gray matter probabilities from the downloaded images for the test.

The clinical 2D image set, generated from a pediatric patient receiving treatment for a posterior fossa brain tumor, included a single section of T_1 -, T_2 - and PD-weighted images. We acquired the images as transverse 5-mm-thick slices with a 1-mm gap interleaved among them to avoid cross talk between slice excitations. T_1 images were acquired by using a gradient-echo 2D FLASH imaging sequence [repetition time (TR) between spin excitations=266 ms, echo time (TE)=6 ms, flip angle=90°; three acquisitions]. T_2 and PD images were acquired simultaneously by using a dual spin-echo sequence (TR/TE1/TE2=3500/19/93 ms; one acquisition). Those images were measured by a phased-array coil with a resolution of 256×256 and a pixel size of 0.8×0.8 mm.

A clinical 3D imaging set of stacked 2D images was acquired from a pediatric patient receiving treatment for leukemia. Images were acquired as 3-mm-thick contiguous transverse-imaging sets on a 1.5-T magnet using a CP head coil. T_1 -weighted images were acquired with a multiecho inversion-recovery imaging sequence (TR=8000 ms; TE=20 ms; time between inversion and the excitation pulse=300 ms; seven echoes; one acquisition). PD- and T_2 -weighted images were acquired simultaneously with a dual spin-echo sequence (TR/TE1/TE2=3500/17/102 ms; one acquisition). All images had voxel sizes of $0.8 \times 0.8 \times 3$ mm.

Another 3D imaging set of stacked 2D images was acquired from a volunteer on a 3.0-T magnet using an 8-channel phased-array head coil. Images were acquired as 3-mm-thick contiguous transverse-imaging sets. T_1 -weighted images were acquired with a 2D FLASH imaging sequence (TR=200 ms; TE=3 ms; four acquisitions). PD- and T_2 -weighted images were acquired simultaneously with a dual spin-echo sequence (TR/TE1/TE2=5110/15/118 ms; one acquisition). All images had voxel sizes of $0.9 \times 0.9 \times 3$ mm.

2.7. Validation strategy — statistical analyses

The correction results can be evaluated in many ways. If the bias-free image is available, the evaluation is fairly straightforward: we can directly compare the corrected image with the bias-free image. However, bias-free images are not usually available in clinical cases. We calculated three covariance measures: the covariance of white matter, the covariance of gray matter and the joint covariance of white matter and gray matter [9]. They are expressed as

$$CV(\text{white}) = \frac{S.D.(\text{white})}{\text{mean}(\text{white})} \quad (16)$$

$$CV(\text{gray}) = \frac{S.D.(\text{gray})}{\text{mean}(\text{gray})} \quad (17)$$

$$JCV(\text{gray/white}) = \frac{S.D.(\text{white}) + S.D.(\text{gray})}{|\text{mean}(\text{white}) - \text{mean}(\text{gray})|} \quad (18)$$

The calculation of these covariance measures requires independent segmentation of the gray matter and white matter. For simulated images, we obtained the segmentation by white matter and gray matter probability distribution maps; for the clinical images, we obtained the segmented white matter and gray matter maps by an automatic algorithm, which was verified by experienced observers [13,14].

3. Results

3.1. Simulated images

The results of the three covariance analyses for the bias-free, corrupted and corrected simulated images are shown in Table 1. The three covariances for the corrected T_1 -, T_2 - and PD-weighted images with 40% and 20% bias fields are close to those of the corresponding bias-free images (0% bias). The correction applied to the bias-free images modifies the original images only slightly. We conducted another test on the simulated images to correct them using both Powell's method and a conjugate gradient method. The number of times each algorithm updated the entropy was recorded during the correction process. The results listed in Table 2 show that the frequency of function evaluation was significantly reduced for the gradient-based minimization algorithm.

3.2. Clinical images

The results of correction on clinical whole-head images are listed in Table 3 for the 1.5-T images from the patient treated for leukemia and in Table 4 for the 3.0-T images from the volunteer. We observed a general reduction in the covariance and the joint covariance in both imaging sets. The quality of the corrected images is much better than that of the uncorrected images. A single representative section that displays the results of the correction for within-plane variation is shown in Fig. 1. We can visually observe that the bias field was removed from the image. The cross-slice bias field correction was evaluated by plotting intensity means of white matter and gray matter on each section for the 1.5-T images from the patient treated for leukemia (Fig. 2). The cross-slice means of both images were much more homogenous than those of the uncorrected images.

4. Discussion

We suggest a new implementation of the entropy-minimization algorithm to correct the bias field in MR images. This implementation improves the nonlinear minimization process by using the entropy gradient information. Although this implementation requires extra gradient calculations, the function evaluation time and line minimization time have been reduced significantly for each calculation. The reduction of the function evaluation time does not necessarily imply a corresponding reduction of the computation time. The computation time depends on the number of parameters used by the correction function, the image data size and the method used to sample the image data. Bias fields in MR images are very different. For the complicated bias field such as the one generated by the phased-array coil, high-degree polynomials are needed to model the underlying bias field, and more image data have to be sampled to fit the bias field. Insufficient modeling and sampling may result in an incorrect or suboptimal solution. The new implementation described here enabled us to correct complicated bias fields for large image data size in an acceptable time.

Although the entropy-minimization algorithm for bias field correction does not require tissue segmentation and is simple to implement, its accuracy and efficiency have not been established. The algorithm works well for images with initial multimodal distributions in which the tissue classes are fairly separated on the histogram, even with the bias field corruption. For the images in which the initial histogram or PDF does not show a multimodal distribution (which could

happen on an abnormal brain or the brain of young children), entropy minimization might result in the increase of the joint covariance between different tissue classes. This would result in the intensity means of the two different tissue classes being shifted toward each other. In these cases, a bias field correction algorithm based on tissue segmentation will also fail because a single spectrum-based segmentation algorithm strongly relies on how well the different tissue classes are separated in the intensity histogram. To improve the performance of the entropy-minimization method in correcting the image with the monomodal distribution, it is ideal to incorporate a constraint to prevent a shift in class means.

In summary, we have presented a new implementation of the entropy-minimization algorithm to correct complex bias fields in MR images. Four important modifications were made to the conventional algorithm: (a) implementation of a modified two-step sampling strategy for 3D image data sets, which included reducing the size of the measured image on each slice with a simple averaging method without changing the number of slices and then using a binary mask generated by a histogram threshold method to define the sampled voxels in the reduced image; (b) improvement of the efficiency of the correction function by using a Legendre polynomial as an orthogonal base function polynomial; (c) use of a nonparametric Parzen window estimator with a Gaussian kernel to calculate the probability density function and Shannon entropy directly from the image data; and (d) performing entropy minimization with a conjugate gradient method. The new implementation described here enabled us to correct complicated bias fields for large image data size in an acceptable time.

Acknowledgment

This work was supported in part by R01-CA90246, by Cancer Center Support (CORE) grant P30-CA21765 from the National Cancer Institute and by the American Lebanese Syrian Associated Charities (ALSAC).

References

- [1]. Brinkermann BH, Manduca A, Robb RA. Optimized homomorphic unsharp masking for MR grayscale inhomogeneity correction. *IEEE Trans Med Imaging* 1998;17:161–71. [PubMed: 9688149]
- [2]. Pham DL, Prince JL. Adaptive fuzzy segmentation of magnetic resonance images. *IEEE Trans Med Imaging* 1999;18:737–52. [PubMed: 10571379]
- [3]. Pham DL, Prince JL. An adaptive fuzzy C-means algorithm for image segmentation in the presence of intensity inhomogeneities. *Pattern Recogn Lett* 1999;20:57–68.
- [4]. Phillips WEII, Velthuisen RP, Phuphanich S, Hall LO, Clarke LP, Silbiger ML. Application of fuzzy C-means segmentation technique for tissue differentiation in MR images of a hemorrhagic glioblastoma multiforme. *Magn Reson Imaging* 1995;13:277–90. [PubMed: 7739370]
- [5]. Sled JG, Zijdenbos A, Evans AC. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans Med Imaging* 1998;17:87–97. [PubMed: 9617910]
- [6]. Van Leemput K, Maes F, Vandermeulen D, Suetens P. Automated model-based bias field correction of MR images of the brain. *IEEE Trans Med Imaging* 1999;18:885–96. [PubMed: 10628948]
- [7]. Wells WM III, Grimson WEL, Kikinis R, Jolesz FA. Adaptive segmentation of MRI data. *IEEE Trans Med Imaging* 1996;15:429–42. [PubMed: 18215925]
- [8]. Mangin, J-F. *IEEE workshop in mathematical methods in bioimage analysis*. IEEE Press; Hilton Head Island, S.C.: 2000. Entropy minimization for automatic correction of intensity nonuniformity.
- [9]. Likar B, Viergever MA, Pernus F. Retrospective correction of MR intensity inhomogeneity by information minimization. *IEEE Trans Med Imaging* 2001;20:1398–410. [PubMed: 11811839]
- [10]. Parzen E. On estimation of a probability density function and mode. *Ann Math Stat* 1962;33:1065.
- [11]. Press, WH.; Teukosky, SA.; Vetterling, WT.; Flannery, BP. *Numerical recipes in C*. Cambridge University Press; Cambridge, U.K: 1992.

- [12]. Collins DL, Zijdenbos AP, Kollokian V, Sled JG, Kabani NJ, Holmes CJ, et al. Design and construction of a realistic digital brain phantom. *IEEE Trans Med Imaging* 1998;17:463–8. [PubMed: 9735909]
- [13]. Reddick WE, Glass JO, Cook EN, Elkin TD, Deaton R. Automated segmentation and classification of multispectral magnetic resonance images of brain using artificial neural networks. *IEEE Trans Med Imaging* 1997;16:911–8. [PubMed: 9533591]
- [14]. Glass JO, Reddick WE, Reeves C, Pui C-H. Improving the segmentation of therapy-induced leukoencephalopathy in children with acute lymphoblastic leukemia using a priori information and a gradient magnitude threshold. *Magn Reson Med* 2004;52:1336–41. [PubMed: 15562471]

Fig. 1.

A single representative section that shows correction results for clinical images from a pediatric patient with a posterior fossa brain tumor. Left: an uncorrected proton density-weighted image. Right: corrected image.

Fig. 2.

White matter and gray matter intensity means across 22 sections in uncorrected (dashed line) and corrected (solid line) proton density-weighted images from a pediatric patient with leukemia imaged at 1.5 T with a CP head coil. The first two sections from the apex of the head are not shown since they contained very little tissue and substantial partial voluming artifacts.

Table 1
Coefficient of variation (CV%) and joint coefficient of variation (JCV%) for correction of simulated imaging data

Image set (bias)	CV% (white)		CV% (gray)		JCV% (white/gray)	
	Original	Corrected	Original	Corrected	Original	Corrected
T_1	None					
	20%	7.5	19.1	19.0	100.2	100.8
	40%	8.4	19.5	19.1	104.1	100.2
T_2	None					
	20%	10.0	20.3	19.1	111.9	101.8
	40%					
PD	None					
	20%	11.0	31.2	31.2	124.1	121.6
	40%	11.3	31.6	31.2	128.6	124.1
PD	None					
	20%	12.5	32.4	31.3	137.3	124.2
	40%					
PD	None					
	20%	4.8	6.2	6.2	93.9	93.9
	40%	5.4	7.0	6.2	114.8	95.3
PD	None					
	20%	7.0	8.8	6.2	155.6	94.2
	40%					

Table 2

Comparison of Powell's method and conjugate gradient method using the number of times (in thousands) that entropy was calculated

Image set (bias)	Powell's method	Conjugate gradient method
T_1		
None	6.0	0.5
20%	29.2	2.1
40%	36.6	2.3
T_2		
None	10.0	0.7
20%	23.0	1.2
40%	28.2	1.8
PD		
None	25.1	1.9
20%	43.6	2.8
40%	64.6	4.0

Table 3

Coefficient of variation (CV%) and joint coefficient of variation (JCV%) for correction of clinical imaging data from 1.5 T with a CP head coil

	T_2		PD	
	Original	Corrected	Original	Corrected
CV% (white)	7.4	7.0	3.9	3.6
CV% (gray)	16.4	15.7	5.2	4.5
JCV% (white/gray)	89.9	83.2	78.9	72.6

Table 4

Coefficient of variation (CV%) and joint coefficient of variation (JCV%) for correction of clinical imaging data from 3 T with an 8-channel phased-array coil

	T_2		PD	
	Original	Corrected	Original	Corrected
CV% (white)	25.8	16.0	12.1	5.0
CV% (gray)	35.5	27.9	12.9	9.8
JCV% (white/gray)	185.5	137.5	108.5	61.9