ELSEVIER

NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 23 (2004) 75–83

Correction of differential intensity inhomogeneity in longitudinal MR images

Emma B. Lewis* and Nicholas C. Fox

Dementia Research Group, Institute of Neurology, University College London, London, WC1N 3BG, UK

Received 4 December 2003; revised 10 March 2004; accepted 28 April 2004
Available online 15 July 2004

Longitudinal MR imaging is increasingly being used to measure cerebral atrophy progression in dementia and other neurological disorders. Differences in intensity inhomogeneity between serial scans can confound these measurements. This differential bias also distorts nonlinear registration and makes both manual and automated segmentation of tissue type less reliable. A technique is described for the correction of this differential bias that makes no assumptions about signal distribution, bias field or signal homogeneity. Instead, the bias field calculation is performed on the basis that the remaining structure in the difference image of registered serial scans has small-scale structure. The differential bias field is of much larger scale and can thus be obtained by applying an appropriate filter to the difference image. The serial scan pair is then corrected for the differential bias field and atrophy measurement can be performed on the corrected scan pair. Application of a known, simulated bias field to real serial MR images was shown to alter atrophy measurements significantly. The differential correction method recovered the applied differential bias field and thereby improved atrophy measurements. This method was then applied to serial imaging in patients with dementia using a set of serial scan pairs with visually identified, significant differential bias and a set of scan pairs with negligible differential bias. Differential bias correction specifically reduced the variance of the atrophy measure significantly for the scans with significant differential bias.

Keywords: Cerebral atrophy; Dementia; Imaging

Introduction

The increasing number of individuals with dementia represents one of the most pressing social and public health problems. Therapies to prevent or slow the progression of these disorders are urgently needed. There is great interest in using brain imaging to help meet this research challenge. The most important cause of dementia, Alzheimer's disease (AD), is characterised by an inexorable progression of cerebral atrophy. Magnetic resonance imaging (MRI) may be used to visualise and quantify this atrophy. In addition to its use in diagnosis (Fox et al., 1999), measurement of atrophy progression from serial MRI has been suggested as a surrogate marker of disease progression in disease modification trials in AD (Ashburner et al., 2003).

To detect changes in rates of atrophy, it will be crucial to have measurements of atrophy progression that are as precise as possible. Even small differences in MR image acquisition either within image or between images may confound these attempts. Any postprocessing methods that could remove or reduce these differences would therefore be very valuable.

One of the within-image artifacts caused by MR image acquisition is intensity inhomogeneity. Intensity inhomogeneity, or bias, is the slowly changing and smooth spatial variation in signal intensity that can occur within the scan. For example, often the grey and white matter in one part of the brain may have systematically lower signal intensity than similar tissue at the opposite extent of the brain. This effect can be in the order of 20% (Sled et al., 1998). This causes significant problems for the image analysis techniques subsequently applied to the MR images. Fig. 1 shows a pair of longitudinal scans, one of which has negligible intensity inhomogeneity and the other has significant intensity inhomogeneity. The relative difference in the signal intensity in the cerebellum is particularly noticeable.

This artifact is caused by several factors: inhomogeneity of the magnetic field, B_0, of the MR system; inhomogeneity of the radiofrequency (RF) pulse generated by the oscillating secondary magnetic field, B_1 (caused by either distortion of the RF field by the object being scanned or nonuniformity in the transmitter coil generating the RF pulse); or nonuniform sensitivity of the receiver coils used to detect the MR signal.

This bias has several implications for downstream processing. First, segmentation of an image into its constituent tissue types becomes very difficult. This applies to both manual segmentation, where it may become more difficult for the segmentor to identify tissue boundaries, and also automated segmentation techniques, which generally require a model for signal intensity for different tissue types; such models become invalid for images with large bias. Second, registrations performed on serial scans with significant differential bias, in particular nonlinear registrations, may be affected. The nonlinear registration could attempt to explain the

* Corresponding author. Dementia Research Group, Institute of Neurology, University College London, 8-11 Queen Square, London, WC1N 3BG, UK. Fax: +44-870-132-0447.
E-mail address: elewis@dementia.ion.ucl.ac.uk (E.B. Lewis).
Available online on ScienceDirect (www.sciencedirect.com.)

1053-8119/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2004.04.030

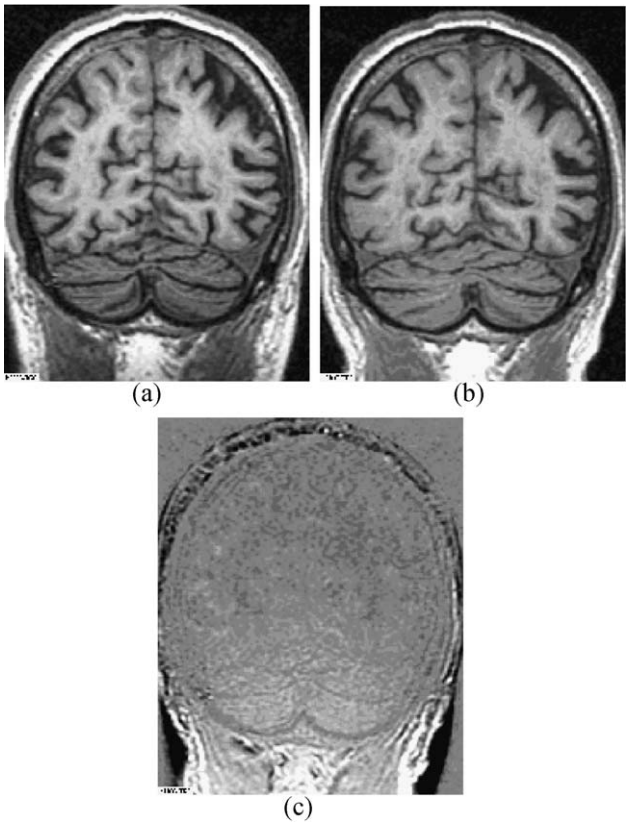

Fig. 1. Longitudinal MR scan pair with pronounced differential bias field: coronal slices through (a) the baseline image and (b) the registered repeat image. Significant bias can be seen from crown to neck with the cerebellum and neck appearing much brighter in (b) the repeat image; (c) coronal slice through difference image between baseline and registered repeat, revealing significant differential bias displayed as bright signal in inferior structures such as cerebellum and neck.

intensity inhomogeneity in the derived deformation field and in doing so produce unrealistic warps. Last, intensity-based quantification techniques will give inaccurate results for scans with significant intensity inhomogeneity. For example, the Brain Boundary Shift Integral (BBSI) (Freeborough and Fox, 1997), a technique used to quantify atrophy within longitudinal scan pairs, will be degraded by the bias field and more particularly by the differential bias field between the two scans.

Current techniques for intensity inhomogeneity correction fall into two categories. The first class of techniques use modified image acquisition protocols to handle bias correction at source. The second class are post processing methods. Both methods are designed to operate on a single scan.

Considering the modified acquisition protocols, correction of inhomogeneity due to nonuniformity of the transmitter or receiver coils can be performed by scanning a homogeneous test object to obtain the inhomogeneity field due to the coil (Wicks et al., 1993); this inhomogeneity field has also been obtained by scanning the object twice using two different RF coils (Narayana et al., 1998). However, neither of these techniques eliminate distortions caused by the object to be scanned. Alternatively, the RF field can be measured during the scan (Stollberger and Wach, 1996) to correct for effects due to the object being imaged; however, this does not remove effects due to nonuniform reception coil sensitivity.

Such modified acquisition protocols are not ideal, however, since they generally require increased scan acquisition time. Also, these techniques cannot be used to correct retrospective data.

Post-processing techniques include those based on the assumption of spatial homogeneity. By assuming tissue homogeneity within a region, the bias field can be calculated over that area and extrapolated to the rest of the image. For example, methods have been proposed to take an initial segmentation of the image into its component tissue types and model the intensity variation within regions of homogeneous tissue-type using polynomials (Tincher et al., 1993). Reference points of a typical tissue class can be selected by an expert and a bias field fitted to those values (Dawant et al., 1993).

Some techniques have assumed that the only low frequency component of an image is the bias field and hence attempted to obtain the bias field by applying a homomorphic filter to the image (a low-pass filter) and assumed what remains is the bias field (Axel et al., 1987; Lim and Pfefferbaum, 1989). Other more sophisticated techniques combine segmentation and intensity correction in a single algorithm, deriving the bias field by modeling tissue type to calculate the expected intensities and then calculating the bias field by comparing the expected and actual intensities. This has been done by maximising the expectations of the intensities (Guillemaud and Brady, 1997; Wells et al., 1996) and by formulating the problem as an energy minimisation one (Styner et al., 2000).

The most commonly used bias correction technique, N3 (Sled et al., 1998), is an automated technique which iteratively sharpens the intensity histogram by deconvolving Gaussian fields from subsequent estimates of the true signal and spline-smoothing the derived bias field. This technique assumes that the distribution of the bias fields is Gaussian, which may not be true for all coil geometries and anatomies (Vokurka et al., 1999). Also, the algorithm is necessarily iterative and is therefore computationally more expensive and may become trapped in local minima (Vokurka et al., 1999).

These post-processing techniques all require assumptions to be made: either a certain model for tissue type, the form of the bias field, regional homogeneity or the relative spatial frequency distribution of signal due to the structure of the brain and the effect of the bias field. The technique described here, however, has been developed specifically to correct differential intensity inhomogeneity to facilitate measurement of pathological changes such as atrophy. This technique avoids assumptions about tissue type, the form of the bias field or of regional homogeneity. Instead, the technique uses the difference image between longitudinal scans of a patient, which can be considered to consist only of noise and atrophy (both of which are small-scale features) and the differential bias field (relatively large scale compared to the noise and atrophy) to identify the differential bias field by the application of simple filters. It should be noted that this technique corrects differential bias, and therefore any bias common to the longitudinal images will persist after application of the technique. Such bias, if significant, will still affect automated segmentation, as described earlier.

The Materials and methods section details the implications of differential bias on atrophy quantification and describes the theory of the correction technique and the algorithm. The Results section presents the results obtained using this technique. The last section gives the conclusions and discussion.

Materials and methods

Theory

It is assumed that image formation is given by:

$$v(x) = u(x)b(x) + n(x) \quad (1)$$

where $v(x)$ is the measured intensity, $u(x)$ is the true signal intensity, $b(x)$ is the bias field and $n(x)$ is the noise, which is independent of signal intensity. This assumes that the bias field is multiplicative, which is generally accepted to be the case for MR images (Wells et al., 1996).

Baseline and repeat longitudinal images of the same individual are first registered using rigid registration with rescaling. The registered images with signal intensity $v_1(x)$ and $v_2(x)$ are log transformed because the bias field is multiplicative. The difference image between the two scans is then calculated to give:

$$\log v_1(x) - \log v_2(x) = \log[u_1(x)b_1(x) + n_1(x)] - \log[u_2(x)b_2(x) + n_2(x)] \quad (2)$$

which becomes:

$$\log v_1(x) - \log v_2(x) = \log u_1(x) - \log u_2(x) + \log b_1(x) - \log b_2(x) + \log[1 + n_1(x)/u_1(x)b_1(x)] + \log[1 + n_2(x)/u_2(x)b_2(x)] \quad (3)$$

In Eq. (3), the terms $\log[1 + n_i(x) / u_i(x)b_i(x)]$, where $i = 1$ and 2, represent the original additive noise of the system. The term $\log u_1(x) - \log u_2(x)$ represents the difference between the true signals in the images. Given that the images are longitudinal, and in register, this term contains no residual brain structure and comprises only registration error and genuine anatomical change between the two scans, namely, any atrophy that may have occurred. Assuming registration error is small and the disease process being studied has not caused any regional changes to the intensity of grey or white matter, then this term will have small-scale structure in the scale-space sense (Perona and Malik, 1990).

The term $\log b_1(x) - \log b_2(x)$ represents the differential bias field and is known to be of large-scale relative to the other factors contributing to signal (Axel et al., 1987; Lim and Pfefferbaum, 1989).

Eq. (3) therefore illustrates that the difference image consists of the differential bias field, noise, atrophy and registration error.

To obtain the differential bias field from the difference image, a median filter is applied to Eq. (3). The median filter removes the Gaussian noise (Petrou and Bosdogianni, 1999). It erases structure of size less than half that of the kernel (Paranjape, 2000), which largely prevents atrophy and small registration error from being mistaken for differential bias field. A kernel size of 5 was used ($11 \times 11 \times 11$ box). This choice was made since it appeared to erase any anatomical structure remaining in the difference image while still being sufficiently sensitive to obtain the differential bias field accurately.

Defining median (v) to be the median transform of image v and applying this median filter to Eq. (3) gives:

$$\log b_1(x) - \log b_2(x) \approx \text{median}\ [\log v_1(x) - \log v_2(x)] \quad (4)$$

Eq. (4) can be rewritten to express the ratio of the bias fields, that is, the differential bias field in terms of the actual signal:

$$b_1(x)/b_2(x) \approx \exp(\text{median}(\log v_1(x) - \log v_2(x))). \quad (5)$$

Implications for longitudinal measurements

Automated image analysis techniques to quantify change between longitudinal scans and/or to identify the location of significant change can be divided into two categories. Both types usually require initial rigid body registration and intensity normalisation of the scan pairs. The first set of techniques either derive a measurement of global atrophy by the comparison of image intensity profiles, for example, BBSI (Freeborough and Fox, 1997), which directly compares intensity values, and SIENA (Smith et al., 2002), which compares intensity gradient profiles, or identify areas of significant difference between scans by performing t tests on a voxel-by-voxel basis having previously removed gross anatomical differences using nonlinear registration (Ashburner and Friston, 2000).

The second set of techniques nonlinearly warp repeat image to baseline image and hold the information describing the atrophy in the deformation field itself, for example, voxel compression mapping (Fox et al., 2001) or deformation-based morphometry (Ashburner et al., 1998).

To analyse the implications of differential intensity inhomogeneity on such techniques, it is necessary to examine the effect of differential intensity inhomogeneity upon the image. Fig. 2 illustrates an idealised intensity profile, demonstrated on a one-dimen-

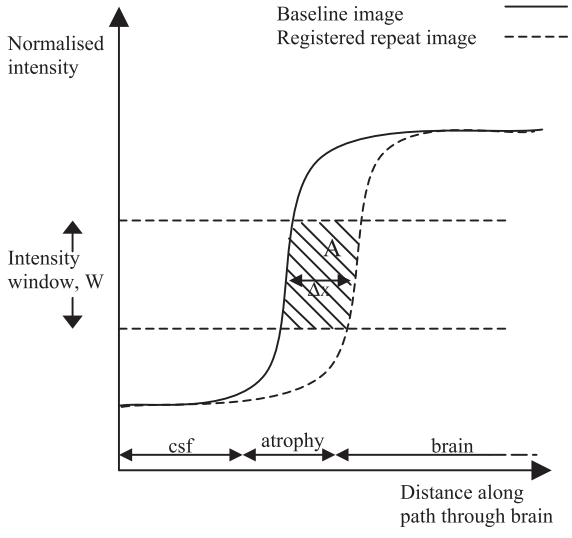

Fig. 2. Diagram illustrating the idealised intensity profiles in one dimension of a longitudinal scan pair across a cerebrospinal fluid (CSF) brain boundary for the case where there is negligible differential intensity inhomogeneity; W is the intensity window across which the boundary transition occurs, and Δx is the distance moved by an anatomical point between the two scans.

sional walk through the brain boundary of two registered and intensity normalised T1-weighted scans, with no differential intensity inhomogeneity. Fig. 3 illustrates the same intensity profiles, but for a longitudinal scan pair with significant differential intensity inhomogeneity.

It can be seen from Figs. 2 and 3 that the accuracy of measures of atrophy calculated using intensity differences will be severely decreased by the bias because of the artificial change in intensity value of the same anatomical point between the two scans. This will be the case for intensity-based techniques, which generally rely on subtraction images or gradients that will be altered by such an artificial change. It will also be the case for deformation-based atrophy quantification techniques where nonlinear registration of the biased images may be driven by the artificial intensity change into producing unrealistic warps.

Implications for Brain Boundary Shift Integral

In this paper, we will consider in particular the impact of differential bias on one of the intensity-based quantification techniques, the Brain Boundary Shift Integral (BBSI) (Freeborough and Fox, 1997). This technique gives a measure of the atrophy that has occurred between longitudinal scans.

Considering Fig. 2, the shift in the boundary of the brain can be approximated thus:

$$\Delta x \approx A/W \tag{6}$$

which is the average distance moved by voxels on the boundary of interest (i.e., partly inside the chosen intensity window), given that an anatomical point takes the same value in both the images.

This extends to three dimensions to give the BBSI, the volume moved through by the boundary, which should approximate the difference in brain volumes between the two scans.

In the case where there is significant differential intensity inhomogeneity between the longitudinal scans, the BBSI is affected as can be seen in Fig. 3. In this case, whichever intensity window is selected, the BBSI defined earlier does not give an indication of how far the boundary has moved. In addition, the intensity window is likely to be near the intensity level on one side of the boundary for one of the scans; thus, extra tissue that is not on the boundary will contribute to the BBSI and add extra noise. Hence, the BBSI does not give a valid measure of atrophy where there is significant differential intensity inhomogeneity between the scans.

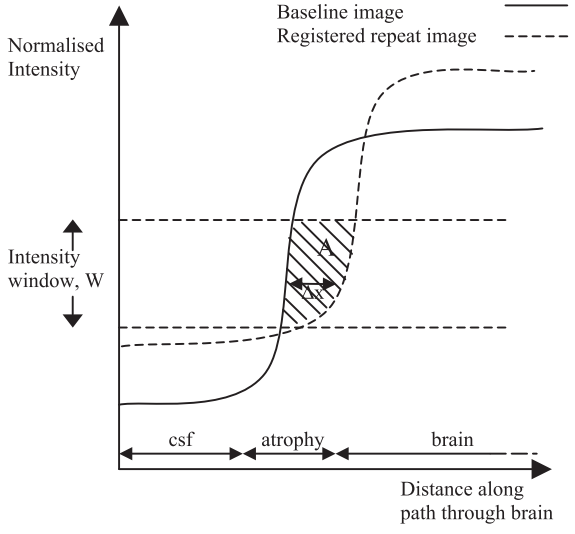

Fig. 3. Diagram illustrating the idealised intensity profiles in one dimension of a longitudinal scan pair across a CSF brain boundary for the case where there is significant differential intensity inhomogeneity; W is the intensity window across which the boundary transition occurs, and Δx is the distance moved by an anatomical point between the two scans.

Considering then, the BBSI calculated using the true signal, BSI_T, rather than the actual signal:

$$\begin{aligned} BSI_T = \sum & [K/(I_2 - I_1)][\mathrm{clip}(u_1(x), I_2, I_1) \\ & - \mathrm{clip}(u_2(x), I_2, I_1)] \end{aligned} \tag{7}$$

where: $\mathrm{clip}(a(x)) = I_2$ if $a(x) > I_2$; $\mathrm{clip}(a(x)) = a(x)$ if $I_1 < a(x) < I_2$; $\mathrm{clip}(a(x)) = I_1$ if $a(x) < I_1$

Substitution from Eq. (1), assuming a noise-free situation, gives:

$$\begin{aligned} BSI_T = & \sum_{I_1}^{I_2} K/(I_2 - I_1)\mathrm{clip}[v_1(x)/b_1(x), I_2, I_1] \\ & - \sum_{I_1}^{I_2} K/(I_2 - I_1)\mathrm{clip}[v_2(x)/b_2(x), I_2, I_1] \end{aligned} \tag{8}$$

which becomes:

$$\begin{aligned} BSI_T = & \sum_{I_1 b_1}^{I_2 b_1} K/[b_1(x)(I_2 - I_1)]\mathrm{clip}[v_1(x), b_1(x)I_2, b_1(x)I_1] \\ & - \sum_{I_1 b_1}^{I_2 b_1} K/[b_1(x)(I_2 - I_1)]\mathrm{clip}[v_2(x)b_1(x)/ \\ & b_2(x), b_1(x)I_2, b_1(x)I_1] \end{aligned} \tag{9}$$

Selecting $J_1 = I_1 / b_1(x)$ and $J_2 = I_2 / b_1(x)$ gives:

$$\begin{aligned} BSI_T = & \sum_{J_1}^{J_2} K/(J_2 - J_1)\mathrm{clip}[v_1(x), J_2, J_1] \\ & - \sum_{J_1}^{J_2} K/(J_2 - J_1)\mathrm{clip}[v_2(x)b_1(x)/b_2(x), J_2, J_1] \end{aligned} \tag{10}$$

This corresponds to calculating the BBSI over a "moving" window, given by limits $I_1 / b_1(x)$ and $I_2 / b_1(x)$. This moving window must meet the criteria for window selection, that is, for all areas of the scan the intensity window $I_1 / b_1(x)$ and $I_2 / b_1(x)$ must be contained within the relevant boundary intensity transition. Given that the maximum bias is of the order of 20%, this should be feasible.

It is thus clear that to correct the BBSI for intensity inhomogeneity problems, it is sufficient to correct for differential intensity homogeneity as obtained from Eq. (5).

The algorithm

Images are initially segmented into brain and nonbrain using MIDAS software (Freeborough et al., 1996). The repeat image is registered into the space of the baseline image using nine degrees of freedom rigid body registration incorporating scalings (Woods et al., 1992, 1993). The repeat brain mask is then

resliced into the same space using the transformation obtained from the registrations. Baseline and registered repeat images are normalised to their mean intensity within the interior brain region, where the interior brain region is the intersection of the baseline and registered repeat brain regions with one erosion applied (Freeborough and Fox, 1997). This normalisation ensures that, to a first approximation, mean intensity over tissue that has not changed between scans will have the same intensity in baseline and repeat images.

Baseline and registered repeat images are log transformed, and the difference image between the log-transformed baseline and registered repeat images is obtained. A median filter of kernel size 5 (i.e., over an 11 × 11 × 11 box) is applied to the difference image, over the union region of the baseline and registered repeat brains with two dilations applied, to remove noise, atrophy and registration errors, which should be of small magnitude because of the similarity of the serial images. The remaining field represents the log transform of the differential bias field of Eq. (4). The log-transform field is inverse log transformed to give the differential bias field. Since this field describes the relative, not absolute, bias, the bias is removed equally from the baseline and registered repeat images. This is achieved by calculating the square root of the differential bias field and correcting both images with this, resulting in both images having the "midway" bias field.

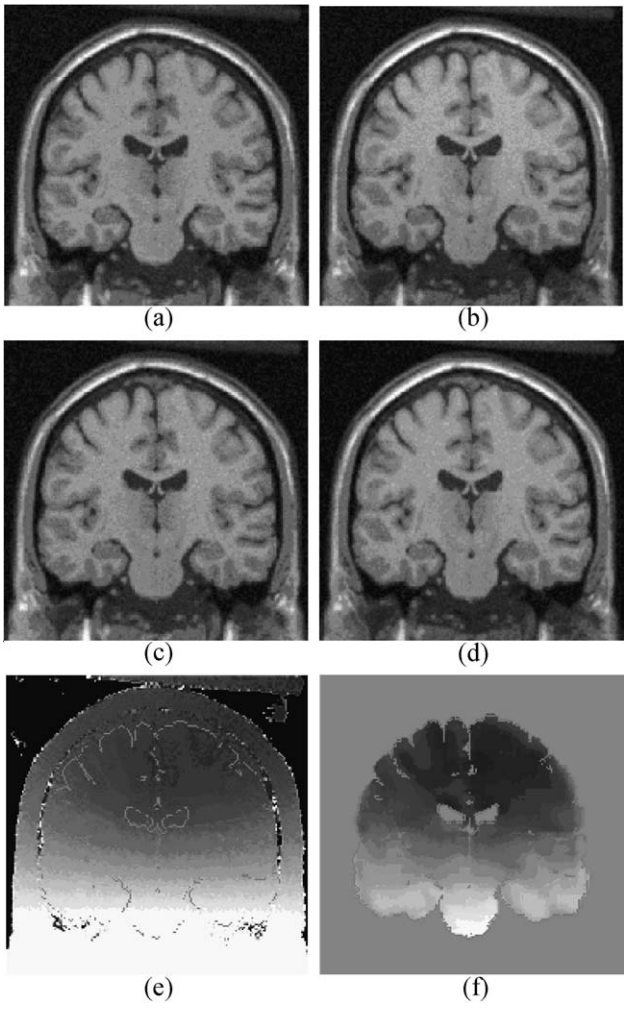

Fig. 4. (a and b) Coronal slices through two simulated images with differential bias fields; (c and d) coronal slices through the same two images after differential bias correction; (e and f) applied and calculated differential bias field.

Table 1
Mean squared difference, as a percentage of grey–white contrast, between applied differential bias field and the bias field obtained using the differential bias correction method for varying differential bias fields for simulated scan pairs, simulated scan pairs with simulated misregistration and simulated scan pairs with simulated atrophy

Scan pair	Mean squared difference
0–20%	0.0053
20–40%	0.0053
0–40%	0.0056
0–40% (atrophy)	0.040
0–20% (misregistered)	0.0099
20%–40% (misregistered)	0.0095
0%–40% (misregistered)	0.010
0%–20% (atrophy)	0.035
20%–40% (atrophy)	0.035

Experiments

The differential bias correction technique (dbc) was tested on both simulated data and real data from patients with Alzheimer's disease. The former enables genuine testing of the technique because the derived differential bias field obtained using the technique developed can be compared against the known, applied differential bias field. The latter is important because it tests the technique's ability to detect real bias fields and to assess how the technique may affect measurements where typical atrophy is occurring.

Simulated bias field applied to simulated images

Images were downloaded from the MNI BrainWeb database (http://www.bic.mni.mcgill.ca/brainweb/; Collins et al., 1998). These images were T1-weighted volumetric MR scans, voxel dimensions 1 × 1 × 1 mm, with 9% noise applied. Artificial bias fields, obtained from MNI BrainWeb, were applied to these images. The form of the simulated bias fields have been obtained from many real MR images. Several different magnitudes of bias fields were applied, namely, 0%, 20%, 40%, where for example, for a bias field of magnitude 20%, the true signal is multiplied by a field with values ranging from 0.9 to 1.1 (i.e., with values between 10% below true intensity and 10% above true intensity).

The technique was run on the scan pairs: 0% bias image to 20% bias image, 20% bias image to 40% bias image and 0% bias image to 40% bias image. The calculated differential bias field was compared with the known applied bias field. This was done both visually, and also by examining the mean squared difference between the known and calculated bias fields.

To analyse the effect of registration error, misregistration was simulated by applying a 0.2 voxel shift in each direction and reslicing the simulated MNI images: it has been shown that high quality registration achieves registration accuracy to within 0.01 voxels (±0.0004) (Hajnal et al., 1995). This technique was then run on scan pairs: 0% bias image to 20% bias misregistered image, 0%

bias image to 40% bias misregistered image and 20% bias image to 40% bias misregistered image.

To analyse the affect of atrophy, rescaling of 0.992 in each direction was applied to the simulated MNI images to simulate atrophy of 2.4% of brain volume (Fox et al., 2000). This technique was then run on scan pairs: 0% bias image to 20% bias rescaled image, 0% bias image to 40% bias rescaled image and 20% bias image to 40% bias rescaled image.

Simulated bias applied to real, same day scan pairs

T1-weighted volumetric MR scans, voxel dimensions 1 × 1 × 1.5 mm, were acquired on a 1.5-T Signa Unit (GE Medical Systems, Milwaukee), inversion recovery (IR)-prepared spoiled GRASS sequence; imaging parameters: TE, 6.4 ms; TI, 650 ms; TR 3000 ms; bandwidth 16 kHz. For each subject, two images were acquired on the same day. The only differences between the two scans should result from changes in head position, which may cause some small amount of differential bias, and noise, since no anatomical change would be expected between same day scans. The volume difference in each scan pair was calculated using the BBSI. This would be expected to be very small and distributed about zero. A simulated bias field in the order of 40% bias, as described in the last section, was then applied to the scan pairs and the BBSI calculated. Differential bias correction was run on each scan pair and the BBSI calculated for the corrected image pair. The BBSI for the image pair with differential bias and the BBSI for the corrected image pair were then compared with the original ground truth BBSI.

Real bias fields on real images

Longitudinal images were acquired of patients with Alzheimer's disease, all of scan interval of approximately 1 year. Thirty-eight pairs were selected, 19 with severe differential bias and 19 with relatively little change in bias, the degree of bias being assessed by an expert. BBSIs were calculated for all the pairs. The technique was then applied to the scan pairs and the BBSIs recalculated. The BBSIs before and after correction were then compared.

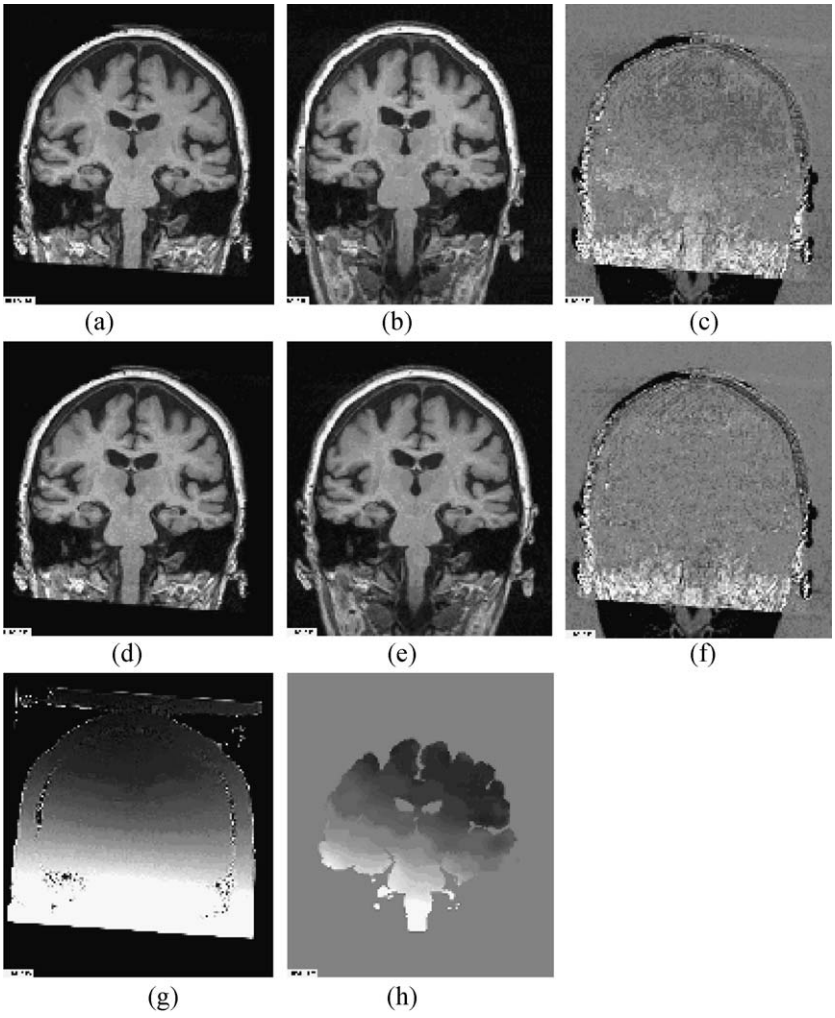

Fig. 5. (a and b) Coronal slices through artificially biased baseline image and registered repeat image, (c) their difference image showing differential bias; (d and e) coronal slices through differential bias corrected baseline and registered repeat image pair; (f) their difference image showing that the differential bias field has been removed; (g and h) applied and calculated differential bias field.

Results

Simulated bias field applied to simulated images

Fig. 4 illustrates the effect of correcting simulated differential bias on a simulated scan pair. The differential bias (of the order of 40%) can be seen between the scan pair (Figs. 4a and b). This differential bias is corrected in scans (Figs. 4c and d) where the differential bias no longer can be seen.

The applied differential bias field and differential bias field obtained from the correction are shown in Figs. 4e and f, respectively. It can be seen that dbc has clearly identified the differential bias field. To assess how accurately the correction technique has calculated the differential bias field, the mean squared difference between actual and obtained bias fields has been calculated for 0–20% bias, 20–40% bias and 0–40% bias for standard simulated image pairs, for image pairs with simulated misregistration (artificial shift) and for image pairs with simulated atrophy (artificial scaling) as a percentage of grey–white contrast, respectively. These results are shown in Table 1. It can be seen that dbc has accurately calculated the applied bias field. Its performance is slightly poorer for the cases of simulated misregistration, by a factor of approximately 1.8, and also poorer for the cases of simulated atrophy, by a factor of approximately 6.5. However, the mean squared difference values, which are a percentage of grey white contrast, show that even in cases of misregistration and atrophy, dbc can accurately obtain the differential bias field.

The algorithm took approximately 300 s to run on a SUNBlade 2000, 950 MHz processor, for an image of 181 × 181 × 217 voxels.

Simulated bias applied to real, same-day scan pairs

Figs. 5a and b illustrate one of the real T1-weighted, baseline and repeat image pairs used with simulated bias of magnitude 40% applied to the baseline image. The resulting differential bias field can be seen in Fig. 5c, which shows the difference image between baseline and registered repeat images. The corrected image pair are shown in Figs. 5d and e, and the corresponding difference image in Fig. 5f shows that the differential bias has been removed. The applied bias field is shown in Fig. 5g, and Fig. 5h shows the obtained differential bias field. It can be seen that the applied and obtained bias fields are similar within the baseline and registered repeat union brain mask, with only a slight difference in the gradient direction, which results from the initial relatively small differential bias field that can be seen between Figs. 5a and b, which will be in addition to the applied differential bias field.

The impact of the differential bias field upon atrophy measurement is shown in Fig. 6. The original calculated volume differences (atrophy) are negligible (zero mean with random noise) since the images were obtained on the same day, and therefore little anatomical change should be seen. The mean absolute volume difference was 0.93 cc (less than 0.1% of whole brain volume) and the variance was 0.82. The volume differences obtained after the application of the simulated bias field to the baseline image, shown in Fig. 5d, are less accurate, with a mean absolute volume difference of 2.51 cc and a greatly increased variance of 2.05, implying that differential bias adds noise to the atrophy quantification measure.

The differential bias correction technique made a significant improvement to the differential bias, as can be seen from the corrected pair shown in Figs. 5d and e. The difference image between the corrected pair, given in Fig. 5f, shows the bias is no longer apparent from visual inspection. The volume difference measurements are improved, with a mean absolute value of 1.09 cc, which is reduced in magnitude from the value for the biased scan pairs, and the variance was also reduced, with a value of 0.41. Note that the volume differences in this experiment result from noise rather than atrophy. Applying dbc therefore does not give obtained absolute volume differences close to the values for the corresponding unbiased scans, but rather reduces the mean and variance of absolute volume change.

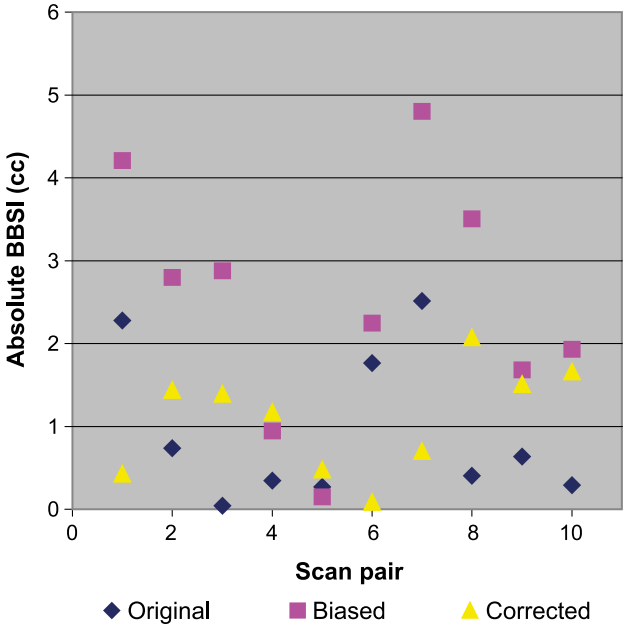

Fig. 6. Absolute amounts of volume difference between two scans (BBSI in cc) for same-day longitudinal image pairs with no differential bias, with applied differential bias and with corrected differential bias.

The algorithm took approximately 180 s to run on a SUNBlade 2000, 950 MHz processor, for an image of 256 × 256 × 124 voxels.

Real bias fields on real images

Figs. 7a and b illustrate one of the original, real T1-weighted, baseline and repeat image pairs used, showing significant real differential bias. Fig. 7c shows a slice from the bias field obtained using the differential bias correction technique. The corrected image pair are shown in Figs. 7d and e. It can be seen that the differential bias can no longer be identified. Figs. 7f and g show a colour representation of the BBSI with areas in red showing voxels where the repeat image has loss of signal (tissue) and green representing gain in signal; Fig. 7f shows the BBSI for the uncorrected pair with red highlighting the ventricular expansion but also with signal gain in the cerebellum shown in green; Fig. 7g shows the BBSI for the corrected pair, note how the artifactual change in the cerebellum for the uncorrected pair is removed by dbc, while the genuine expansion of the ventricles is unaffected.

Table 2 shows the effect of the differential bias correction upon the means and variances of the groups of scans, where group A is the set of scan pairs with significant differential bias and group B is the set of scans with negligible differential bias. It can be seen that the standard deviation is significantly reduced for group A scans

and is much less affected for group B scans. The means for both groups A and B are slightly reduced. This suggests that dbc might be removing a small amount of atrophy. However, the reduction in atrophy is very small relative to the change in standard deviation. In particular, atrophy measurements that were negative, indicating growth of the brain, which are more likely to be artifactual and due to the differential bias, are significantly improved after application of the differential bias correction. The atrophy measurements for scan pairs with little differential bias are very similar with and without correction.

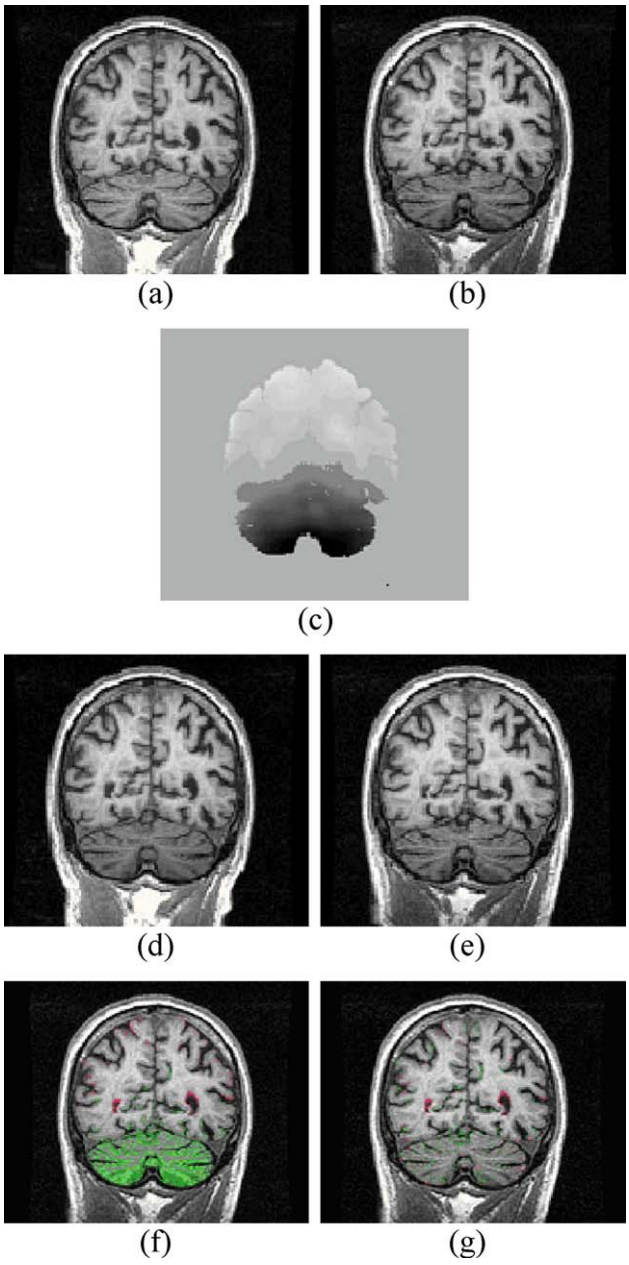

Fig. 7. Coronal slices through (a) baseline and (b) registered repeat of a longitudinal MR scan pair, with pronounced differential bias field; (c) obtained differential bias field; (d and e) same coronal slices in the differential bias-corrected images, note how much more similar to each other d and e are when compared with a and b; (f and g) colour representation of the BBSI: areas highlighted in red show voxels where the repeat image has loss of signal (tissue) and green represents gain in signal; (f) BBSI for the uncorrected pair with red highlighting the ventricular expansion but also with signal gain in the cerebellum shown in green; (g) BBSI for the corrected pair, note how the artifactual change in the cerebellum for the uncorrected pair is removed by the differential bias correction while the genuine expansion of the ventricles is unaffected.

Table 2
Mean and standard deviation of pre-dbc and post-dbc atrophy values (BBSI) for a group of scans with significant differential bias (group A), a group of scans with negligible differential bias (group B) and the two groups combined

	Pre-dbc	Post-dbc
Group A mean (cc)	13.7	13.4
Group A standard deviation	31.1	20.5
Group B mean (cc)	14.6	13.5
Group B standard deviation	7.5	6.7
Mean of combined group (cc)	14.1	13.4
Standard deviation of combined group	23.3	15.0

The algorithm took approximately 300 s to run on a SUNBlade 2000, 950 MHz processor, for an image of $256 \times 256 \times 124$ voxels.

Conclusions and discussion

Image analysis techniques developed for longitudinal images are particularly adversely affected by the difference in inhomogeneity fields, as demonstrated here with an atrophy quantification technique.

In this paper, a nonparametric, noniterative method is described for the calculation of the differential inhomogeneity field between longitudinal MR images. This technique makes none of the assumptions about tissue classes, intensity distributions, bias field form or spatial homogeneity of previous techniques and assumes only that the inhomogeneity field is of large scale (in scale-space sense) relative to the image noise, atrophy that has occurred and registration errors; large-scale components of registration error, atrophy and noise and any other large-scale changes from scan one to scan two may be removed by this technique.

This technique has been tested on simulated T1-weighted images with applied simulated inhomogeneity, on real longitudinal scan pairs onto which simulated inhomogeneity fields have been applied and on real longitudinal scan pairs with significant atrophy and significant differential inhomogeneity. The results of the experiments with simulated inhomogeneity show that the technique has correctly calculated the applied differential inhomogeneity field. Visual comparison of the corrected and uncorrected simulated and real image pairs shows that this technique has successfully removed much of the differential inhomogeneity field. In particular, running dbc on image pairs with simulated misregistration has shown that the technique handles registration error of up to 0.2 voxels; use of a high quality subvoxel registration algorithm, such as one based on mutual information or ratio of intensity uniformity, will register the scans well within this tolerance. However, if the initial registration was poor, the large registration errors would be picked up by the median filter and confused with differential bias, resulting in the technique failing. Running dbc on image pairs with simulated atrophy has shown that it successfully obtains the differential bias field in the presence of atrophy in the order of that seen in typical AD over an interval of 1 year; however, in the presence of larger amounts of large-scale atrophy (e.g., in the case of long intervals between scans), dbc may remove some of the atrophy.

By considering real longitudinal scans with significant atrophy, it has been shown that this technique only marginally removes atrophy, and atrophy values corrected with application of this technique show less variability. This technique can therefore be applied before atrophy quantification to improve the accuracy of atrophy quantification.

Further work will include more detailed analysis of the spectral form of bias, atrophy, registration error and noise with the aim of refining the method by distinguishing components of intensity resulting from noise, atrophy and misregistration from the differential inhomogeneity effects. The choice of filtering technique is expected to be critical in this. For example, one attractive approach could be to apply frequency domain filters to separate more explicitly the frequency components of the image. It would also be of interest to investigate the effect of application of the technique on the joint intensity histogram, which would be sharpened by removal of differential bias, with the possibility of incorporating joint intensity histogram sharpening directly into the algorithm. Future work also includes applying dbc to series of three or more longitudinal scans to analyse the effect upon series of atrophy rates.

This technique could perhaps be extended from correction of differential inhomogeneity to correction of the actual inhomogeneity field within a single scan by the creation of a difference image between that scan and an unbiased scan (or template) warped to the target scan. Nonlinear warping is nontrivial in the presence of bias. However, using an iterative approach with patient topology specific templates might enable this to be successful. This would then remove bias common to the longitudinal scans and hence improve automated segmentation.

Increasingly, longitudinal MRI studies are being analysed to identify disease-related structural changes and to detect treatment effects in therapeutic trials. The technique described here, by reducing spurious intensity differences due to changes in the inhomogeneity between scans, should improve the precision of these important analyses.

Acknowledgments

The authors would like to thank the imaging team of the Dementia Research Group for the manual brain segmentations they have performed on the data used in this paper and the MNI BrainWeb for simulated scans and bias fields. Research support from the MRC (UK) is greatly acknowledged. NCF holds an MRC Senior Clinical Fellowship. The authors also thank the reviewers for their useful comments and discussion.

References

Ashburner, J., Friston, K.J., 2000. Voxel-based morphometry—The methods. NeuroImage 11, 805–821.

Ashburner, J., Hutton, C., Frackowiak, R.S.J., Johnsrude, I.S., Price, C., Friston, K.J., 1998. Identifying global anatomical differences: deformation based morphometry. Hum. Brain Mapp. 6, 348–357.

Ashburner, J., Csernansky, J.G., Davatzikos, C., Fox, N.C., Frisoni, G.B., Thompson, P.M., 2003. Computer-assisted imaging to assess brain structure in healthy and diseased brains. Lancet Neurol. 2, 79–88.

Axel, L., Constantini, J., Listerud, J., 1987. Intensity correction in surface coil MR imaging. Am. J. Radiol. 148, 418–420.

Collins, D.L., Zijdenbos, A.P., Kollokian, V., Sled, J.G., Kabani, N.J., Holmes, C.J., Evans, A.C., 1998. Design and construction of a realistic digital brain phantom. IEEE Trans. Med. Imag. 17, 463–468.

Dawant, B.M., Zjidenbos, A.P., Margolin, R.A., 1993. Correction of intensity variations in MR images for computer-aided tissue classification. IEEE Trans. Med. Imag. 12, 770–781.

Fox, N.C., Warrington, E.K., Rossor, M.N., 1999. Serial magnetic resonance imaging of cerebral atrophy in preclinical Alzheimer's disease. Lancet 353, 2125.

Fox, N.C., Cousens, S., Scahill, R., Harvey, R.J., Rossor, M.N., 2000. Using serial registered brain magnetic resonance imaging to measure disease progression in Alzheimer disease. Arch. Neurol. 57, 339–344.

Fox, N.C., Crum, W.R., Scahill, R.I., Stevens, J.M., Janssen, J.C., Rossor, M.N., 2001. Imaging of onset and progression of Alzheimer's disease with voxel-compression mapping of serial magnetic resonance images. Lancet 358, 201–205.

Freeborough, P.A., Fox, N.C., 1997. The Boundary Shift Integral: an accurate and robust measure of cerebral volume changes from registered repeat MRI. IEEE Trans. Med. Imag. 16, 623–629.

Freeborough, P.A., Woods, R.P., Fox, N.C., 1996. Accurate registration of serial 3D MR brain images and its application to visualizing change in neurodegenerative disorders. J. Comput. Assist. Tomogr. 20, 1012–1022.

Guillemaud, R., Brady, M., 1997. Estimating the bias field of MR images. IEEE Trans. Med. Imag. 16, 238–251.

Hajnal, J.V., Saeed, N., Soar, E.J., Oatridge, A., Young, I.R., Bydder, G.M., 1995. A registration and interpolation procedure for subvoxel matching of serially acquired MR images. J. Comput. Assist. Tomogr. 19, 289–296.

Lim, K.O., Pfefferbaum, A., 1989. Segmentation of MR brain images into cerebrospinal fluid spaces, white and gray matter. J. Comput. Assist. Tomogr. 13, 588–593.

Narayana, P.A., Brey, W.W., Kulkarni, V., Sievenpiper, C.L., 1998. Compensation for surface coil sensitivity variation in magnetic resonance imaging. Magn. Reson. Imaging 6, 271–274.

Paranjape, R.B., 2000. Fundamental enhancement techniques. In: Bankman, I.N. (Ed.), Handbook of Medical Imaging: Processing and Analysis, vol. 7. Academic Press, San Diego, p. 7.

Perona, P., Malik, J., 1990. Scale-space and edge detection using anisotropic diffusion. IEEE Trans. Pattern Anal. Mach. Intell. 12, 629–639.

Petrou, M., Bosdogianni, P., 1999. Image Processing: The fundamentals. Wiley, New York, pp. 145–149.

Sled, J.G., Zijdenbos, A.P., Evans, A.C., 1998. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. IEEE Trans. Med. Imag. 17, 87–97.

Smith, S., Zhang, Y., Jenkinson, M., Chen, J., Matthews, P.M., Federico, N., De Stefano, N., 2002. Accurate, robust and automated longitudinal and cross-sectional brain change analysis. NeuroImage 17, 479–489.

Stollberger, R., Wach, P., 1996. Imaging of the active B1 field in vivo. Magn. Reson. Med. 35, 246–251.

Styner, M., Brechbuhler, C., Szekely, G., Gerig, G., 2000. Parametric estimate of intensity inhomogeneities. IEEE Trans. Med. Imag. 19, 153–165.

Tincher, M., Meyer, C.R., Gupta, R., Williams, D.M., 1993. Polynomial modeling and reduction of RF body coil spatial inhomogeneity in MRI. IEEE Trans. Med. Imag. 12, 361–365.

Vokurka, E.A., Thacker, N.A., Jackson, A., 1999. A fast model independent method for automatic correction of intensity nonuniformity in MRI data. J. Magn. Reson. Imaging 10, 550–562.

Wells III, W.M., Grimson, W.E.L., Kikinis, R., Jolesz, F.A. 1996. Adaptive segmentation of MRI data. IEEE Trans. Med. Imag. 15, 429–442.

Wicks, D.G., Barker, G.J., Tofts, P.S., 1993. Correction of intensity nonuniformity in MR images of any orientation. Magn. Reson. Imaging 11, 183–196.

Woods, R.P., Cherry, S.R., Mazziotta, J.C., 1992. Rapid automated algorithm for aligning and reslicing PET images. J. Comput. Assist. Tomogr. 16, 620–633.

Woods, R.P., Mazziotta, J.C., Cherry, S.R., 1993. MRI-PET registration with automated algorithm. J. Comput. Assist. Tomogr. 17, 536–546.