A Review of Methods for Correction of Intensity Inhomogeneity in MRI

Uroš Vovk, Franjo Pernuš, and Boštjan Likar*

***Abstract*—Medical image acquisition devices provide a vast amount of anatomical and functional information, which facilitate and improve diagnosis and patient treatment, especially when supported by modern quantitative image analysis methods. However, modality specific image artifacts, such as the phenomena of intensity inhomogeneity in magnetic resonance images (MRI), are still prominent and can adversely affect quantitative image analysis. In this paper, numerous methods that have been developed to reduce or eliminate intensity inhomogeneities in MRI are reviewed. First, the methods are classified according to the inhomogeneity correction strategy. Next, different qualitative and quantitative evaluation approaches are reviewed. Third, 60 relevant publications are categorized according to several features and analyzed so as to reveal major trends, popularity, evaluation strategies and applications. Finally, key evaluation issues and future development of the inhomogeneity correction field, supported by the results of the analysis, are discussed.**

***Index Terms*—Bias field, intensity inhomogeneity, intensity nonuniformity, magnetic resonance images (MRI), segmentation, shading.**

I. Introduction

MEDICAL image acquisition devices and protocols that have tremendously evolved over the last decades provide a vast amount of data out of which the information essential for diagnosis, therapy planning and execution, and monitoring the progress of disease or results of treatment has to be extracted. Automated extraction of clinically useful information usually requires a preprocessing step by which various image artifacts, which may degrade the results of subsequent image analysis algorithms, are removed. This paper addresses a class of preprocessing methods that deal with spurious smoothly varying image intensities, i.e., with the phenomenon that is usually referred to as intensity inhomogeneity, intensity nonuniformity, shading or bias field. This adverse phenomenon is apparent in images obtained by different imaging modalities, such as microscopy, computer tomography, ultrasound, and above all by magnetic resonance imaging (MRI). Intensity inhomogeneity in MRI, which arises from the imperfections of the image acquisition process, manifests itself as a smooth intensity variation across the image (Fig. 1). Because of this phenomenon, the intensity of the same tissue varies with the location of the tissue within the image. Although intensity inhomogeneity is usually hardly noticeable to a human observer, many medical image analysis methods, such as segmentation and registration, are highly sensitive to the spurious variations of image intensities. This is why a number of methods for intensity inhomogeneity correction of magnetic resonance (MR) images have been proposed in the past.

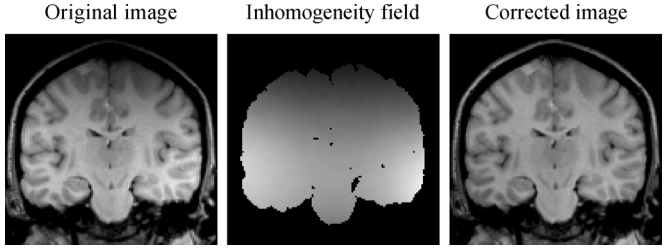

Fig. 1. Intensity inhomogeneity in MR brain image.

Early publications on MRI intensity inhomogeneity correction date back to 1986 [1], [2]. Since then, sources of intensity inhomogeneity in MRI have been studied extensively [3]–[6] and can be generally divided into two groups. Sources in the first group are related to the properties of the MRI device and include static field inhomogeneity, bandwidth filtering of the data, eddy currents driven by field gradients, and especially radio frequency (RF) transmission and reception inhomogeneity. These intensity inhomogeneities can be corrected for by shimming techniques [4], [7], special imaging sequences and different sets of coils, or by calibrating the MRI device by a phantom or a mathematical model [2], [8]–[10]. Sources in the second group are related to the imaged object itself, i.e., to the shape, position and orientation of the object inside the magnet, and to the specific magnetic permeability and dielectric properties of the imaged object. These intensity inhomogeneities are far more difficult to deal with. The impact of the imaged object is rather small in low magnetic field and more prominent in high magnetic field MR scanners. This is due to the linear increase of frequency needed to stimulate the nuclei under higher magnetic fields, which enhances the effects of RF standing waves and penetration [4], [11].

Very few reviews of methods for intensity inhomogeneity correction methods have been published in the past. In [12] and [13], the performances of four and six methods, respectively, have been compared, while a recent review of a large number of methods published until 2002 can be found in [14]. The present review is primarily focused on intensity inhomogeneity correction methods that are based on retrospective image analysis and less on methods dealing with specific hardware/acquisition aspects. Our goal was to provide a complete overview of the existing correction methods and evaluation strategies and to cate-

Manuscript received August 8, 2006; revised December 5, 2006. This work was supported by the Ministry of Higher Education, Science and Technology, Republic of Slovenia under under Grant P2-0232, Grant L2-7381, and L2-9758. *Asterisk indicates corresponding author.*

U. Vovk and F. Pernuš are with the University of Ljubljana, Faculty of Electrical Engineering, Tržaška 25, 1000 Ljubljana, Slovenia.

*B. Likar is with the University of Ljubljana, Faculty of Electrical Engineering, Tržaška 25, 1000 Ljubljana, Slovenia (e-mail: bostjan.likar@fe.uni-lj.si).

Digital Object Identifier 10.1109/TMI.2006.891486

gorize and analyze numerous publications so as to assess major trends, popularity, and applications of intensity inhomogeneity correction methods.

The rest of the paper is organized as follows. Section II presents the most common models of intensity inhomogeneity and noise. In Sections III–V different correction strategies are presented and classified. Section VI deals with the evaluation strategies. Results of analysis of 60 papers dealing with inhomogeneity correction methods are provided in Section VII, followed by discussion in Section VIII.

II. Models of Intensity Inhomogeneity

The generally accepted assumption on intensity inhomogeneity is that it manifests itself as a smooth spatially varying function that alters image intensities that otherwise would be constant for the same tissue type regardless of its position in an image. In its most simple form, the model assumes that intensity inhomogeneity is multiplicative or additive, i.e., the intensity inhomogeneity field multiplies or adds to the image intensities. Most frequently, the multiplicative model has been used as it is consistent with the inhomogeneous sensitivity of the reception coil. For modeling inhomogeneities that are due to induced currents and nonuniform excitation, the multiplicative model is less appropriate [15]. In addition to intensity inhomogeneity, the MR image formation model should incorporate high-frequency noise. This noise is known to have a Rician distribution [16]. However, as long as the signal-to-noise ratio (SNR) is not too low, noise can be approximated by a quasi-Gaussian distribution [17]. This approximation is appropriate for image areas corresponding to tissues but not for no-signal areas, such as air.

Different models of MR image formation have been proposed in the literature, depending on how the inhomogeneity-free image $u(\mathbf{x})$, intensity inhomogeneity field $b(\mathbf{x})$, and noise $n(\mathbf{x})$ interact. Two sources of noise were described in [16], [18], and [19], namely the biological noise, which corresponds to the within tissue inhomogeneity, and a scanner noise that arises from MR device imperfections. However, usually only one of these sources is modeled. The most common model of MR image formation assumes that the noise, approximated by Gaussian probability distribution, arises from the scanner and is therefore independent of the intensity inhomogeneity field $b(\mathbf{x})$ [20], [21]. According to this model, the acquired image $v(\mathbf{x})$ is obtained as

$$v(\mathbf{x}) = u(\mathbf{x})b(\mathbf{x}) + n(\mathbf{x}). \tag{1}$$

In another model of MR image formation only biological noise is considered, which is scaled by the intensity inhomogeneity field $b(\mathbf{x})$ so that the SNR is preserved [16], [22]

$$v(\mathbf{x}) = (u(\mathbf{x}) + n(\mathbf{x}))\, b(\mathbf{x}). \tag{2}$$

The third MR image formation model is based on log-transformed intensities, by which the multiplicative inhomogeneity field becomes additive [23]–[25]

$$\log v(\mathbf{x}) = \log u(\mathbf{x}) + \log b(\mathbf{x}) + n(\mathbf{x}). \tag{3}$$

In this model, the log-transformed noise $n(\mathbf{x})$ is still assumed to be Gaussian, which is methodologically convenient but inconsistent with the first model (1) that assumes the noise is Gaussian in the original nonlogarithmic domain. This inconsistency was apparently not considered relevant enough to be commented by the authors [23]–[25]. However, in most of inhomogeneity correction methods, the noise is handled by simple filtering, smooth model fitting, or some form of regularization and is therefore considered rather irrelevant.

The authors in [26] proposed a method that separated the intensity inhomogeneity correction field into multiplicative and additive components. The model used in [26] was similar to the first model (1) but with an extra term modeling the additive component of the correction field. The additional additive component was proved useful for microscopic images [27] but much less for MR images [26].

In [12] and [15], it was speculated whether a smooth multiplicative inhomogeneity correction field and additive noise are sufficient to accurately model the actual MR acquisition process and associated artifacts. Due to the susceptibility effects, intensity inhomogeneity might be tissue dependant and specially accentuated at tissue boundaries. In [28] it was suggested that for a certain MR sequences, such as inversion-recovery, the inhomogeneity is tissue dependant. The claim was supported by the appropriate simplified equation of the measured MR signal. Besides this theoretical claim, the authors in [29] built a special physical model and claimed that the multiplicative inhomogeneity field is not smooth in T1 MR images because the effects of RF transmission inhomogeneity depend not only on the location but also on T1 relaxation times of different tissues. The method proposed in [29] relied on a preacquired phantom image, flip angle measurement and special references. The experimental results showed that the proposed MR physics based mathematical model improved the results of a standard phantom calibration but, on the other hand, was too sensitive to input parameters and references. The results in [29] revealed that physics based modeling of higher magnetic field artifacts is still rather unreliable. This explains why the majority of the correction methods are based on simple and more reliable correction models (1)–(3). However, it should be stressed out that the most common smooth inhomogeneity correction models may not always yield satisfying corrections, especially when the properties of the tissues also vary smoothly across the anatomy, such as between the genu and splenium of the corpus callosum, and thereby mimic intensity inhomogeneity. Efficient separation of smooth inhomogeneity fields and smooth anatomical variations is far from trivial, requiring additional *a priori* knowledge on at least either of them, which makes the retrospective inhomogeneity correction problem less tractable than is commonly assumed.

III. Classification of Correction Methods

Numerous intensity inhomogeneity correction methods have been proposed in the last two decades. In the following, we propose a classification scheme, accordingly classify correction methods, and discuss the advantages and disadvantages of different correction strategies. From the very beginning, two main

approaches have been applied to minimize the intensity inhomogeneity in MR images, namely the prospective and retrospective approach. The first aims at calibration and improvement of the image acquisition process, while the latter relies exclusively on the information of the acquired image and sometimes also on some *a priori* knowledge. According to the classification proposed below, we have further classified the prospective methods into those that are based on phantoms, multicoils, and special sequences. The retrospective methods are further classified into filtering, surface fitting, segmentation, and histogram based.

Prospective

- Phantom
- Multicoil
- Special sequences

Retrospective

- Filtering
  - *Homomorphic*
  - *Homomorphic unsharp masking*
- Surface fitting
  - *Intensity*
  - *Gradient*
- Segmentation
  - *ML, MAP*
  - *Fuzzy c-means*
  - *Nonparametric*
- Histogram
  - *High frequency maximization*
  - *Information minimization*
  - *Histogram matching*
- Other

IV. Prospective Methods

Prospective methods treat intensity corruption as a systematic error of the MRI acquisition process that can either be minimized by acquiring additional images of a uniform phantom, by acquiring additional images with different coils, or by devising special imaging sequences.

A. Phantom Based

An estimate of the intensity inhomogeneity field can be obtained by acquiring an image of a uniform phantom with known physical properties and by scaling and smoothing of the acquired phantom image [3], [8]. Oil or water is usually used for phantoms and median filtering is applied for image smoothing. The phantom based approach cannot correct for patient-induced inhomogeneity, which is a major drawback of this approach. The remaining intensity inhomogeneity can be as high as 30% [30]. Another limitation of this approach is the temporal and spatial variation of the coil profile that calls for frequent acquisitions of the phantom image. To reduce the number of phantom acquisitions, the authors in [10] proposed to carefully record the orientation and position of the phantom and the coils so as to be able to geometrically transform the estimated inhomogeneity field to any acquired image. Attempts have been made to first mathematically model the inhomogeneity field by polynomials [31], curves [9] or by integration based on the Biot-Savart law [2], and then fit the obtained model to the phantom image. A method that minimizes the dependency of the inhomogeneity field on the scanner and object has been proposed in [29], deriving a mathematical model from the equation for T1 signal generation and using a phantom image, flip angle mapping, and reference objects. However, the usefulness of this method is limited by the specific imaging conditions and sensitivity to input parameters.

B. Multicoil

Two types of coils, surface and volume, are most frequently used in MRI. Surface coils usually have good SNR but induce severe intensity inhomogeneity, while volume coils exhibit less inhomogeneity but have poor SNR. Methods have been proposed that require acquisition and combination of surface and body coil images to obtain an intensity inhomogeneity-free image with good SNR [32], [33]. The intensity inhomogeneity field is obtained by dividing the filtered surface coil image with the body coil image and smoothing the resulting image. The main disadvantage of these methods is the prolonged acquisition time. Besides, the inhomogeneity of the body coil image remains in the corrected surface coil image. To shorten acquisition times, the authors in [34] used a low resolution body coil image, which was registered to the full resolution surface coil image. The final inhomogeneity field was modeled by a spline surface that had been fitted to a set of inhomogeneity field estimates. An energy minimization approach was proposed in [35], integrating one body coil and multiple surface coil images into the final image. This method could also handle multispectral images (T1, T2) if acquired by the same coils. In general, the usability of the multicoil methods is limited by prolonged acquisition times, special coil settings, and incomplete inhomogeneity correction.

C. Special Sequences

This group of techniques is predominantly related to specific acquisition (hardware) designs and is thus only briefly described. For certain pulse sequences, the spatial distribution of the flip angle can be estimated and used to calculate the intensity inhomogeneity. The mathematical model behind this approach requires the acquisition of two images, the second one with doubled nominal flip angle of the first [36]. Other techniques, such as sensitivity encoding by multiple receiver coils, also reduce the inhomogeneity artifact but they were mainly developed to speed up the scanning process [37]. In [38], echo planar imaging (EPI) phased modulation maps were estimated to remove the inhomogeneity field in the original image. Another method, proposed for the inhomogeneity field minimization of 4.7T images, modified certain pulses in the modified

driven equilibrium Fourier transform (MDEFT) imaging [39]. Besides, for MRI contrasts, such as diffusion weighting, magnetization transfer ratios, two point quantitative T1, and two point quantitative T2, in which the final image is a ratio, most of the multiplicative intensity inhomogeneity is cancelled and the images are relatively homogeneous. This is because the sensitivity variation of the receive coil remains unchanged even if the sequence parameters are varied.

V. Retrospective Methods

Retrospective methods are relatively general as only a few assumptions about the acquisition process are usually made. These methods mainly rely on the information of the acquired images in which useful anatomical information and information on the intensity inhomogeneity are integrated. *A priori* knowledge on spatial and/or intensity probability distribution of the imaged anatomy is used by some methods to facilitate extraction of information on intensity inhomogeneity. In contrast to the prospective methods, which can correct only the intensity inhomogeneity induced by an MR scanner, retrospective methods can also remove patient dependant inhomogeneity.

A. Filtering Methods

Filtering methods assume that intensity inhomogeneity is a low-frequency artifact that can be separated from the high-frequency signal of the imaged anatomical structures by low-pass filtering. However, this assumption is valid only if the imaged anatomical structures are relatively small and thus contain no low frequencies that might be mistakenly removed by low-pass filtering. For most of the anatomical structures imaged by MR this assumption does not hold, which results in overlap of anatomy and inhomogeneity frequency spectra. This limits the feasibility of filtering methods. Besides, high image contrasts generate filtering artifacts known as edge effects, manifesting themselves as a distortion of homogeneous tissues near the edges. The strongest edges that are usually at the object/background transitions can be removed by either masking out the background [40], replacing background pixels by average intensity values [41], [42], or by extrapolating tissue intensities over the background [1], [3], [43]. Nevertheless, substantial intensity inhomogeneity usually remains in an image after correction by these methods [13]. Two main filtering approaches, homomorphic filtering and homomorphic unsharp masking (HUM), have been proposed. Morphological filtering [44] and simple high-pass filtering also belong to this group of methods but, in contrast to some successful applications to microscopic images [45], they have not been found useful for MRI.

Homomorphic Filtering: Homomorphic filtering is conducted on log-transformed image intensities. The image background is usually altered prior to log-transformation by one of the approaches mentioned above [41]. The input image $\log v(\mathbf{x})$ is subtracted by its low-pass filtered (LPF) version, which estimates the inhomogeneity field

$$\log u(\mathbf{x}) = \log v(\mathbf{x}) - \mathrm{LPF}\left(\log v(\mathbf{x})\right) + C_N. \qquad (4)$$

The normalization constant C_N is added to preserve the mean or maximum intensity of the corrected image, while the corrected image $u(\mathbf{x})$ is finally obtained by exponentiation. The authors in [46] successfully modified this method to remove intensity inhomogeneity differences between two longitudinal images, thus enabling the measurement of cerebral atrophy progression.

Homomorphic Unsharp Masking: HUM was first proposed by Axel [8] and is probably the simplest and one of the most commonly used methods for intensity inhomogeneity correction in MRI. HUM, which is an approximation of classical homomorphic filtering, is very fast and easy to implement. The inhomogeneity correction field $b(\mathbf{x})$ is obtained by low-pass filtering of the input image $v(\mathbf{x})$, divided by the constant $\mathrm{C_N}$ to preserve mean or median intensity. The corrected image $u(\mathbf{x})$ is expressed as a division, which is equivalent to subtraction in log-domain, of the input image $v(\mathbf{x})$ by the inhomogeneity correction field

$$u(\mathbf{x}) = v(\mathbf{x})/b(\mathbf{x}) = v(\mathbf{x})C_N/\mathrm{LPF}\left(v(\mathbf{x})\right). \qquad (5)$$

Low-pass filtering can be mean [3], [43], or median based [47], or implemented by multiplication in Fourier domain [1], [42]. The authors in [40] showed that for the human brain mean filtering produced better results than median filtering.

B. Surface Fitting Methods

These methods fit a parametric surface to a set of image features that contain information on intensity inhomogeneity. The resulting surface, which is usually polynomial or spline based, represents the multiplicative inhomogeneity field that is used to correct the input image. According to the image features used for surface fitting, surface fitting methods are further sorted into intensity and gradient based methods.

Intensity Based: A parametric surface in the form of thin plate splines was least squares fitted to intensities of a set of pixels, which were assumed to belong to the same tissue and were distributed over the entire image [48]. Manual selection of pixels corresponding to a dominant tissue and automatic selection, which was based on neural network classification, were investigated. Although subjective and time-consuming, manual selection was shown to give better results. In [49], an automated iterative method was proposed, incorporating segmentation of homogeneous areas of the major tissue, followed by fitting a second order polynomial to intensities of the segmented tissue. A similar approach, using a Gaussian main tissue model, was applied to estimate the inhomogeneity fields and merge several surface coil (phased array) images [50]. The major drawback of these methods is that the inhomogeneity field is estimated only from intensities of one major tissue and then blindly extrapolated over the whole image. A polynomial surface fitting method [51] was also proposed in combination with multipass Gaussian filtering. A combination of *B*-spline surface fitting and a histogram based method was proposed in [52].

Gradient Based: The main assumption behind these methods is that sufficiently large homogeneous areas are evenly distributed over the entire image so that local gradients of intensity inhomogeneity can be estimated by local averaging of image intensity gradients. In [53], a polynomial surface was least squares fitted to underlying normalized intensities of homo-

geneous areas. Tissue independent segmentation was used to obtain the homogeneous areas. Because these were rather sparsely distributed over the entire volume, not all image information was used to estimate the inhomogeneity correction field. By a similar approach, a finite element surface model was fitted by minimizing the difference between derivatives of homogeneous areas and the corresponding surface model [54]. Instead of fitting a surface, some methods obtain the inhomogeneity field by integrating derivatives estimated inside homogeneous areas [55]–[57]. Simple polynomial models were used to extrapolate the gradients outside the homogeneous areas. The major drawback of these methods is that some adverse image information may be integrated. These methods are successful only if homogeneous image areas are large and distinctive, such as the white matter in brain images.

C. Segmentation Based Methods

Intensity inhomogeneity correction is often a necessary preprocessing step enabling better image segmentation. On the other hand, correct segmentation makes intensity inhomogeneity correction rather trivial. Intensity inhomogeneity correction and segmentation can thus be viewed upon as two intertwined procedures. In segmentation based intensity inhomogeneity correction methods these two procedures are merged so that they benefit from each other, simultaneously yielding better segmentation and inhomogeneity correction. These intensity inhomogeneity correction methods are further classified according to the image segmentation method utilized.

ML, MAP Based: The maximum-likelihood (ML) or the maximum *a posteriori* probability (MAP) criterion may be used to estimate the image intensity probability distribution by parametric models. Finite mixture and more frequently finite Gaussian mixture models are used and modified to incorporate intensity inhomogeneity. The models' parameters can be estimated by the expectation-maximization (EM) algorithm [25], [58], iterating between classification and intensity inhomogeneity correction. An additional class, named "other," which had a uniform intensity probability distribution was introduced in [23] to model the intensities not belonging to any of the main tissues. Another similar approach uses additional mixed tissue classes to model the partial volume effect [30], [59] and background by Rayleigh distribution [60]. Because the Gaussian model is only an approximation of a single tissue probability density, several Gaussians can be used per one main tissue, for example, 3 for white matter and 2 for grey matter in the brain, and much more for minor, less significant tissues [22]. A unique approach, first published in [61] and refined later in [30], upgraded the EM iterative scheme by adding a special step in which the resulting inhomogeneity field was scaled to minimize a new criterion, namely, the classification error rate (CER). In this way, the whole algorithm was primarily guided by the results that would be produced by a coarse segmentation. The criterion of minimum image entropy was used in [62] to estimate the inhomogeneity field, while the classical EM procedure was implemented to optimize ML in search for model parameters. To deal with the high number of searched parameters, e.g. when model parameter estimation or spatial constraints were added to the classification and inhomogeneity correction steps, or to speed up the algorithm, optimization schemes such as generalized EM (GEM) [24], iterative conditional modes (ICM) [59], [63] or expectation conditional minimization (ECM) [16], [64] have been proposed.

One of the main requirements of the EM based approaches is the initialization of explicitly modeled classes and spatial distribution of tissues. Initialization can be obtained by manual selection of representative points for each class [25], which is subjective and irreproducible. An automatic approach, initially classifying the tissues by intensity thresholding, was proposed in [58] where a registered set of images was needed to train the classifier. A self-adaptive vector quantization [59] and tree-structure k-means classification [63] were implemented to estimate the initial class means, standard deviations or brain mask. Other automatic approaches [22], [24], [60], [65] used a statistical probability atlas, initially registered to the processed image, to estimate the required parameters.

Because intensity probability models do not exploit the information about spatial connectedness of neighboring pixels belonging to the same class, (hidden) Markov random fields [(H)MRF] were frequently incorporated [59], [63], [64], [66], [67]. This resulted in improved segmentation, which was less sensitive to noise and had smoother tissue borders. However, to avoid over smoothing of tissue borders, spatial connectedness should not be too strong. The authors in [68] suggested to limit HMRF smoothing by penalizing only the neighboring intensity combinations that were implausible due to the specific topology of the imaged object.

Instead of using image registration only to estimate the initial model parameters, the authors in [22] proposed an iterative framework, interleaving segmentation, registration and intensity inhomogeneity correction to improve tissue segmentation. FGM (finite Gaussian mixture) was used as a probability model with intensity inhomogeneity dependant class means. The registration step in [22] was performed by deforming the tissue probability maps. The ICM scheme estimated different groups of parameters, using EM to find the mixture-classification parameters and Levenberg–Marquardt optimization for inhomogeneity field and registration steps. This algorithm, although rather complex and time-consuming, seems promising due to the integration of segmentation, registration and intensity inhomogeneity correction, which had been treated separately in the past.

Fuzzy C-Means: These methods use the standard fuzzy c-means (FCM) segmentation [69] and modify the objective function to adapt to intensity inhomogeneity. The main property of FCM methods is the soft classification model, which assumes that image voxels belong to more than one class. This is consistent with the partial volume effect observed in MR images and thus eliminates the explicit modeling of mixed classes, which is required by the abovementioned FGM models. The authors in [70] proposed an adaptive fuzzy c-means method, which multiplies the class centroid values by a function of location, estimating the intensity inhomogeneity. Spatial regularization terms, penalizing first and second derivatives of the inhomogeneity field, were added to the objective function to preserve its smoothness. Because the weights of these regularization terms are difficult to set, they have to be tuned manually. An automatic procedure based on histogram mode searching was

used to set the initial values of class centroids, while the objective function was minimized by the Jacobi iterative scheme applied on a multigrid algorithm to speed up the process [70]. The method was later generalized to 3-D multispectral images and accelerated by the same authors [20]. In [71] and refined later in [72], spatial information was incorporated by adding a spatial regularization term that enabled the class membership of a voxel to be influenced by its neighbors. This approach proved tolerant to salt and pepper noise, resulting in smoother segmentation. Again, a regularization parameter had to be tuned, determining the smoothness of segmentation and, implicitly, also the inhomogeneity field. Another approach was proposed in [73] where the inhomogeneity field was modeled by a *B*-spline surface, while the spatial voxel connectivity was implemented by a dissimilarity index, which enforced the connectivity constraint only in the homogeneous areas. In this way the tissue boundaries were better preserved. Yet another method that performed local fuzzy *c*-means classification and thereby completely avoided the need for modeling the intensity inhomogeneity function was presented in [74].

Nonparametric: Nonparametric segmentation based inhomogeneity correction methods were proposed in [75] and [76], using nonparametric maxshift or meanshift clustering. The methods did not require any *a priori* knowledge on the intensity probability distribution, e.g., tissue class means and variances, but blindly classified the voxels according to the main modes of the feature space that combined voxel intensities and corresponding second derivatives. The latter were incorporated to exploit the spatial voxel connectivity, i.e., to incorporate spatial information into classification. Inhomogeneity correction by a parametric polynomial model is based on iterative minimization of class square error, i.e., within class scatter, of the intensity distribution that is due to intensity inhomogeneity.

D. Histogram Based Methods

Histogram based methods operate directly on image intensity histograms and need little or no initialization and/or *a priori* knowledge on the intensity probability distribution of the imaged structures. This makes these methods fully automatic and highly general so that they can usually be applied to various images with or without pathology. Although a number of segmentation based methods also operate on image intensity histograms, the distinction between the segmentation based and histogram based methods is that the latter provide no segmentation results.

High-Frequency Maximization: A well-known intensity inhomogeneity correction method, known as the N3 (nonparametric nonuniformity normalization), was proposed in [15]. The method is iterative and seeks the smooth multiplicative field that maximizes the high frequency content of the distribution of tissue intensity. The method is fully automatic, requires no *a priori* knowledge and can be applied to almost any MR image. Interestingly, no improvements have been suggested for this highly popular and successful method.

Information Minimization: These methods are based on the assumption that intensity inhomogeneity corruption introduces additional information to the inhomogeneity-free image. The removal of intensity inhomogeneity is, therefore, based on constrained minimization of image information, which is estimated by image entropy. Image entropy can be computed from the original intensity distributions or from the log-transformed distributions. In the first case, multiplicative correction model has to be constrained so as to avoid uniform scaling of image intensities (contrast changing), which would otherwise result in completely uniform image with no anatomical information. On the other hand, in log-transformed intensity domain, the multiplicative correction model becomes additive and thereby requires no scaling constraints. However, numerical computation of entropy becomes far more difficult due to the nonlinear log-transformation of image intensities and associated problems with histogram binning. Nevertheless, the information minimization methods can be widely applied to different types of MR images because they use solely the information that is present in an image, without making assumptions on spatial and intensity distributions.

An information minimization based intensity inhomogeneity correction method was first considered in 1995 [77]. Refined applications on microscopic [27] and MR images [26], [28], [62], [78] followed in year 2000 and latter. The method in [28] utilized fast annealing to minimize a three part energy function, consisting of image entropy, field smoothness constraint and a mean preserving regularization term. In [26] and [78], image correction was performed by a linear model consisting of multiplicative and additive components, which were modeled by a combination of smoothly varying basis functions that were constrained to preserve the global intensity statistics. An interesting attempt to extent the information minimization method [26] was proposed in [79], optimizing the first-order conditional entropy. Inter-volume inhomogeneity correction, removing only the differences between two inhomogeneity fields to allow intervolume comparisons, was solved by minimization of joint volume entropy [80]. In [81], the authors proposed an iterative correction strategy in which intensity inhomogeneity correction forces, reducing the global entropy of the feature space, were estimated for each voxel. Besides intensities, spatial information in the form of second order derivatives was incorporated in the feature space. This method was further extended in [82] with the aim to integrate spatial and intensity information from multispectral MR images, i.e., from T1, T2, and proton density (PD) weighted images. Intensity inhomogeneities of such multispectral images were removed simultaneously in a four-step iterative procedure. The advantage of this method was its ability to incorporate complementary information of the multispectral images and their derived features for better inhomogeneity correction.

Integration of complimentary information from multispectral images requires that images are well registered, which is usually the case for multispectral MR images. In general, however, integration of multispectral images, e.g., computed tomography (CT) and MR images, requires multimodal image registration, which is also often solved efficiently in information theoretic framework, e.g., by maximizing mutual information. Therefore, the problem of registration and inhomogeneity correction, one concerned with transforming an image in spatial

domain to achieve spatial correspondence and the other concerned with transforming an image in intensity domain to restore intensity homogeneity, may both be solved simultaneously in an information theoretic framework. For example, by transforming a reference image, both in the spatial domain and in the intensity domain, such that mutual information of the reference and target images is maximized will bring the two images, being monomodal or multimodal, in spatial and in intensity correspondence. This suggests that simultaneous information theoretic registration and inhomogeneity correction are well worth further exploration.

Histogram Matching: A histogram matching method proposed in [21] divided the image into small subvolumes in which intensity inhomogeneity was supposed to be relatively constant. Local intensity inhomogeneity was estimated by least square fitting of the intensity histogram model to the actual histogram of a subvolume. The applied histogram model was a finite Gaussian mixture with seven parameters, initialized from the global histogram of the image. No segmentation was required as only intensity inhomogeneity estimation was needed for each subvolume. These local estimates were then tested for outliers and interpolated by a *B*-spline surface to produce the final inhomogeneity field. Another method [19] found the Legendre polynomial inhomogeneity model by nonlinear optimization based on a special valley function, which was shaped by the *a priori* given mean intensities and standard deviations of the main tissue classes. As histogram matching methods require several input parameters and a tissue model, they are far less general as, for example, the information minimization methods.

E. Other Methods

Three methods that cannot be easily classified into any of the above categories are briefly described. By the registration based method proposed in [83], an image undergoing intensity inhomogeneity correction was registered to a reference image. Normalized mutual information and a *B*-spline deformation model were used to perform multiscale rigid and nonrigid registrations. The inhomogeneity field was extracted by smoothing and dividing the two registered images. The major drawback of this approach is the requirement for an application-specific reference image. The second method relies on singularity function analysis [84]. The main idea behind this method was to correct an image in such a way that its high frequency spectrum remained unchanged while the intensity inhomogeneity corrupted low frequency part was removed and later reconstructed by a model that enforced piecewise intensity constancy in the image domain. The algorithm works on one dimensional image profiles, alternating between columns and rows. The final inhomogeneity field is obtained by smoothing the quotient of the original and the reconstructed image. The method requires no *a priori* knowledge or background removal but may be sensitive to a number of input parameters. The third method [85] combines a set of techniques for embedding the physics of the imaging process that generates a class of MR images into segmentation or registration algorithm, resulting in substantial invariance to acquisition parameters, as the effect of these parameters on the contrast properties of various brain structures is explicitly modeled in the segmentation. Besides, the integration of image acquisition with tissue classification allows the derivation of sequences that are optimal for segmentation purposes.

VI. Evaluation Strategies

The aim of evaluation is to experimentally show that a novel method is feasible for certain applications and that it performs favorably to other methods under various conditions. To further substantiate the importance of evaluation let us cite Jain and Binford [86]: "*The importance of theory cannot be overemphasized. But at the same time, a discipline without experiment is not scientific. Without adequate experimental methods, there is no way to rigorously substantiate new ideas and to evaluate different approaches*."

Many different evaluation approaches have been followed in the field of intensity inhomogeneity correction. The evaluation approaches can be classified into two major categories, namely into qualitative and quantitative evaluations, each of which can be further subdivided into two and three distinctive categories, respectively.

Qualitative:

Spatial domain;

Statistical.

Quantitative:

Inhomogeneity field;

Intensity variation;

Segmentation.

Qualitative evaluation is based on subjective visual inspection of the correction results, while quantitative evaluation relies on certain objective measures that are considered relevant for a given application. Although quantitative evaluation is considered more scientific, both qualitative and quantitative evaluations are almost equally important as they provide complimentary information on a method's performance.

A. Qualitative Evaluation

Qualitative evaluations provide useful information on the target application, the type and quality of images, the weaknesses of the correction method, and the results of individual steps of the method. For example, it is relatively easy to observe the severity of the inhomogeneity field in volume slices or to find any irregularities in the images of the estimated inhomogeneity field, such as the presence of anatomical structures. Qualitative evaluation can be carried out either in spatial domain or by using statistical tools. The former include 1-D or 2-D (mesh plot) intensity and inhomogeneity field profiles [54], original and corrected images, their differences, intensity coded inhomogeneity fields, and intensity coded segmentation results, which are usually used by segmentation based methods [63]. Other, less common qualitative evaluations use phase diagrams [2], manually selected reference points [48], homogeneous regions of support [53], volume surface rendering [25],

probability maps [24], isointensity lines [19] and segmentation results by tissue delineation [83]. The second class of qualitative evaluation methods relies on the statistics of the original and corrected images like inhomogeneity field histograms [15], scatter plots [13], profile of mean slice intensity [84], etc. Among these, intensity histograms are the most common and informative.

B. Quantitative Evaluation

Quantitative evaluation is essential for objective comparison of the results of different correction methods. Three approaches to quantitative evaluation are commonly used: direct comparison of inhomogeneity fields, estimation of intensity variations within tissues of interest, and indirect measurement of the effects of inhomogeneity correction via segmentation.

Inhomogeneity Field Based: By this approach, the extracted intensity inhomogeneity field is compared to a reference field, which is either a true inhomogeneity field or its close approximation. A measure of the quality of inhomogeneity correction can either be computed by comparing the extracted and reference inhomogeneity fields or by comparing intensities of the corrected and reference image. The image background is often masked out because it contains no information on intensity inhomogeneity.

The root mean square error (rms) was used in [16] and [40] to measure the difference between two inhomogeneity fields or image intensities, denoted by $b_1(\mathbf{x})$ and $b_2(\mathbf{x})$

$$\mathrm{rms}\left(b_1(x), b_2(x)\right) = \sqrt{\frac{1}{N}\sum_{x\in\Omega}\left(b_1(x) - b_2(x)\right)^2} \qquad (6)$$

where Ω is a region of interest of size N. Two similar measures, the mean square error (MSE) and the mean square distance (MSD) were used in [19], [35], and [46]. Because scaling of the b terms affects these measures, the b terms have to be normalized before these measures are calculated [21] and [49]. In this way, only the difference in shape of the inhomogeneity fields is measured and not their absolute values that contain no useful information on inhomogeneity correction. The measure that is inherently insensitive to the absolute values of inhomogeneity fields is the correlation coefficient (CC), which has been used in [13].

The drawback of direct comparison of inhomogeneity fields is that ideal inhomogeneity-free images or true inhomogeneity fields have to be known. This limits such evaluation to simulated images unless some approximation of the ground truth inhomogeneity-free real images is available. The authors in [83], for example, compared their corrected images to manually corrected images. In all these cases, however, the evaluations are influenced by the quality of the ground truth.

Intensity Variation Based: The results of inhomogeneity correction can be quantitatively assessed by estimating the mean and/or variance of intensities of a tissue of interest before and after correction. Assuming that the spatial intensity distribution of a tissue of interest is piecewise constant, its variance or standard deviation should be reduced if intensity inhomogeneity is removed. The mean intensity of a tissue, on the other hand, can be used to estimate global intensity scaling due to inhomogeneity correction or to measure the change in contrast between two tissues of interest. Means and/or variances can be used separately [41], [57] or combined into more informative measures, like coefficient of variation (CV) and coefficient of joint variation (CJV).

Coefficient of variation is defined as a quotient between standard deviation σ and mean value μ of selected tissue class (C)

$$\mathrm{CV}(C) = \frac{\sigma(C)}{\mu(C)}. \qquad (7)$$

CV is invariant to multiplicative uniform intensity transformation (intensity scaling or contrast), which an intensity inhomogeneity correction method could introduce. The quality of inhomogeneity correction has been quantitatively assessed by the change/reduction of the CV of a tissue of interest in [10], [15], [23] and [24]. A drawback of CV is its sensitivity to uniform additive intensity transformation (brightness), which changes the mean but not the variance of a tissue. Another drawback of CV is that it is computed for a single tissue class, which makes the assessment of the overall inhomogeneity correction difficult when CV improves for one tissue class but not for others. Yet another drawback of CV is that it does not provide any information on the overlap between intensity distributions of distinct tissue classes. For example, a correction method may transform a given image in such a way that the CV of two tissues is reduced, while the overlap between the intensity distributions of the two tissues is increased.

Because of the abovementioned drawbacks of the CV, the coefficient of joint variation, which estimates the overlap between two tissue classes (C_1 and C_2), was introduced in [26]. CJV, which is defined as

$$\mathrm{CJV}(C_1, C_2) = \frac{\sigma(C_1) + \sigma(C_2)}{|\mu(C_1) - \mu(C_2)|} \qquad (8)$$

is invariant to the uniform linear, i.e., multiplicative and additive, intensity transformation. The CJV, which approximates the classification error, was applied in [26], [30], [79], [81], [82], and [84]. Because CJV and CV can be used to compare different methods even if full gold standard segmentations are not available, they have been applied to a large number of real images in which only small subsets of tissue regions had been identified [15], [26]. However, CV and CJV do not provide information on the amount of inhomogeneity remaining after correction unless ideal inhomogeneity-free images are available.

Segmentation Based: By this approach, inhomogeneity correction is evaluated via segmentation results. Several measures derived from the true positive/negative and false positive/negative segmentations have been proposed. The misclassification ratio (MCR), a ratio between the number of misclassified versus all image voxels, has been used in [20], [25], [62], [66], and [73]. The MCR is calculated for all image voxels of interest, while the Jaccard similarity and the Dice coefficient estimate the segmentation of voxels of one segmented tissue. The Jaccard similarity, used in [21] and [68], is defined as the ratio between intersection and union of two sets S_1 and S_2, representing the obtained

and gold standard segmentations, respectively

$$J(S_1, S_2) = \frac{|S_1 \cap S_2|}{|S_1 \cup S_2|}. \quad (9)$$

The authors in [22], [41], and [73] used the Dice coefficient, which is a special case of the κ index, defined as

$$\kappa(S_1, S_2) = \frac{2|S_1 \cap S_2|}{|S_1| + |S_2|}. \quad (10)$$

In comparison to the Dice coefficient, the Jaccard similarity is more sensitive when sets are more similar. Other similar measures that have been used in the past include the accurate segmentation ratio (ASR) [4], segmentation accuracy (SA) [72], false positive (FPF) and false negative fraction (FNF) [59], false positives and true negatives [21], and measurement of the change of tumor volume [12]. All these segmentation based measures provide quantitative information on segmentation accuracy, assuming that gold standard segmentations are available, which is a highly demanding requirement, especially for real images. Besides, the quality of inhomogeneity correction itself is relatively obscured and, therefore, difficult to assess, both within and especially between different studies.

VII. Analysis of Publications

In the previous sections, we have addressed the classification of inhomogeneity correction methods and evaluation strategies. In this section, the methods are analyzed according to several features, like the methodological aspect, validation approach and application. The analysis provides an additional and a more general insight into the field of intensity inhomogeneity correction, yet, it is important to note that the most popular approaches may not necessary be the best ones.

A. Selected Publications

The following analyses was performed on a database of 60 selected publications, 48 journal, and 12 conference papers published between 1986 and 2005, dealing fully of partially with intensity inhomogeneity correction. The criterion for the inclusion of a publication in the database was that at least some results on inhomogeneity correction were provided. Conference publications were only included when no corresponding journal publication was available.

Fig. 2 shows the number of publications per year from 1986 to 2005. The field of intensity inhomogeneity correction became more active after 1996 and is active ever since. Altogether, the publications appeared in 23 different journals and conferences. Fig. 3 illustrates the journals and conferences where the publications most frequently appeared.

B. Methodology

Some interesting methodological aspects of the published inhomogeneity correction methods, namely correction strategy, optimization, modeling, additional processing and *a priori* knowledge, are analyzed in this subsection.

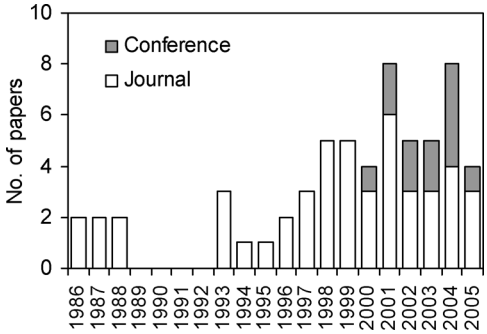

Fig. 2. Publications per year.

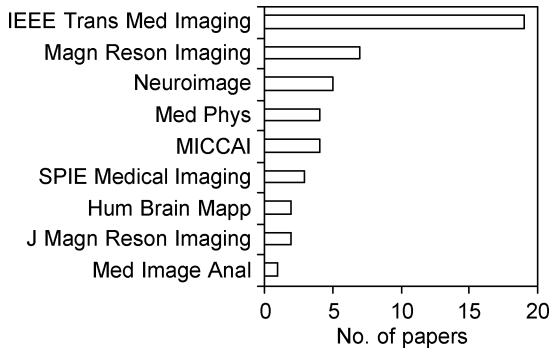

Fig. 3. Journals and conferences most frequently publishing intensity inhomogeneity correction methods.

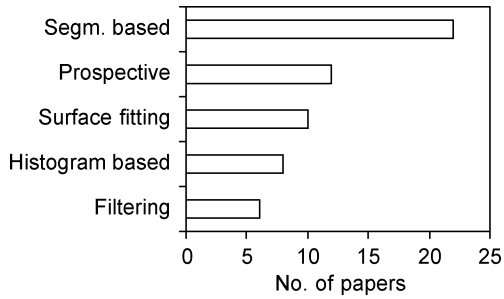

Fig. 4. The most popular correction strategies.

Correction Strategies: As described in Section III, the methods can be classified into several categories according to the basic correction strategy. The strategies that were most frequently followed over the past 20 years are illustrated in Fig. 4. Much more papers reporting on retrospective than prospective inhomogeneity correction methods have been published. The retrospective methods are further classified into segmentation based methods, surface fitting, histogram based methods and filtering methods. The segmentation based methods prevailed, which is probably due to the fact that segmentation is one of the most important procedures that heavily depends on the quality of images.

Fig. 5 gives the number of publications of the five most popular correction strategies per five-year periods. The prospective methods dominated in years 1986–1995. The more computationally demanding segmentation and histogram based methods dominate in the last decade, which is probably due to sufficient

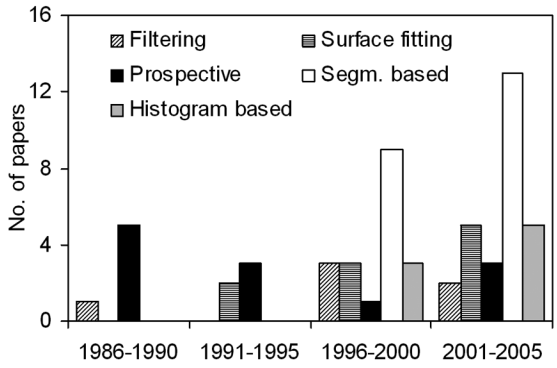

Fig. 5. The most popular correction strategies per three-year periods.

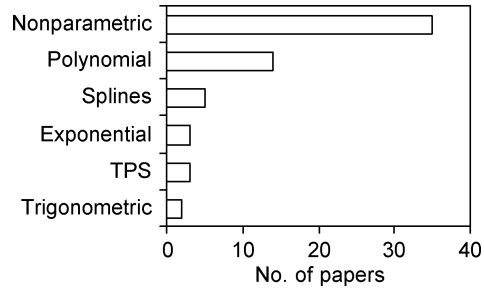

Fig. 6. The most popular inhomogeneity field models.

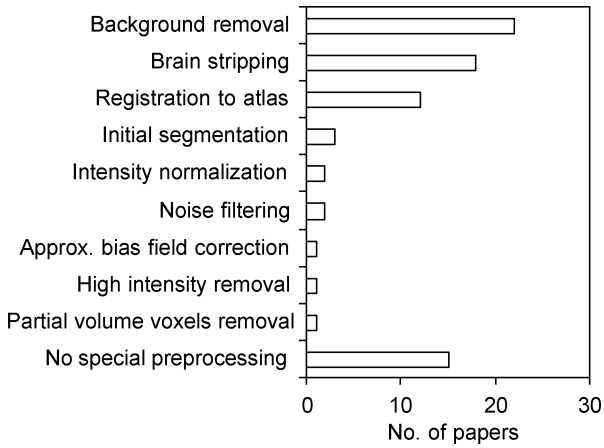

Fig. 7. Most frequent preprocessing and postprocessing steps.

computer hardware capabilities and the increased importance of quantitative MR image analysis.

Estimation and Modeling: In intensity inhomogeneity correction, the inhomogeneity model is defined either by direct estimation or by optimizing the model's parameters. The optimization based estimation appeared to be slightly more popular, as it was used in 60% of the published methods, typically incorporating the EM estimation scheme, coordinate descent, Powel's optimization, simulated annealing, or some other generally applicable optimization scheme. A more simple direct estimation of intensity inhomogeneity by which inhomogeneity is extracted in a single step by filtering, fitting, or by interpolation, was slightly less popular and usually incorporated into phantom based, filtering and surface fitting methods.

The basic assumption about the inhomogeneity field that all methods share is smoothness. Different parametric or nonparametric models were used to assure smoothness. Nonparametric models smooth the inhomogeneity field estimates either by extensive filtering or by using spatial regularization terms in the optimized objective function. The regularization terms usually penalize first-order or second-order derivatives of the inhomogeneity field. The parametric models, on the other hand, explicitly model the intensity inhomogeneity and thereby its smoothness, for example by a selection of basis functions, which were frequently polynomial terms. Some parametric bias models also use spatial regularization. The statistics of the most commonly used models is given in Fig. 6. Nonparametric models dominated in the past. The most common parametric models were polynomials of order 2 to 5 [24], [31], including the Lagrange polynomials [19]. Spline based models [28] and more specifically *B*-splines [15], [21] were also relatively popular, while exponential [9], thin-plate splines (TPS) [48], and trigonometric basis functions [22], [80] were seldom used.

Additional Processing and A Priori Knowledge: Many inhomogeneity correction methods require some additional preprocessing and/or postprocessing steps to either initialize the correction or to subsequently improve the results. The methods that usually do not include these additional steps are phantom based prospective methods [2], [29], surface fitting methods [48], [54] and to a lesser extent segmentation based methods [62]. Fig. 7 illustrates the most frequently used pre- and postprocessing steps. Background removal, usually executed by simple intensity thresholding, was the most common preprocessing step, closely followed by brain stripping, needed to remove all voxels not predicted by the FGM model. Registration to an atlas, required to initialize the probability density or spatial tissue distributions, was also a common step. Initial segmentation was sometimes used for the same purpose in [19], [59]. Postprocessing steps like intensity normalization [58], noise filtering [25], inhomogeneity field correction [35], high-intensity removal [43], nonmodeled tissue removal, and partial volume [19] removal were applied to enhance the final results. The authors in [87] studied the interaction between noise reduction and intensity inhomogeneity correction. They concluded that noise should be suppressed after intensity inhomogeneity correction and that noise filtering could benefit from the information on the inhomogeneity field.

Fig. 8 shows, that in many publications, reporting mainly on histogram based, surface fitting and filtering methods, no *a priori* knowledge was used to facilitate or improve the results. *A priori* knowledge that was most frequently included was the number of classes and corresponding means and standard deviations, usually required by segmentation based methods. A rich source of information is the atlas in the form of spatial probability maps or image intensities, to which the image that is to be corrected is registered. However, registration, which is a problem by itself, may introduce errors. Other *a priori* knowledge, like manual selection of reference points needed for surface fitting or preregistration, was seldom used [48], [58].

C. Validation

In this section, we analyze the publications with respect to the validation approaches presented in Section VI. Qualitative eval-

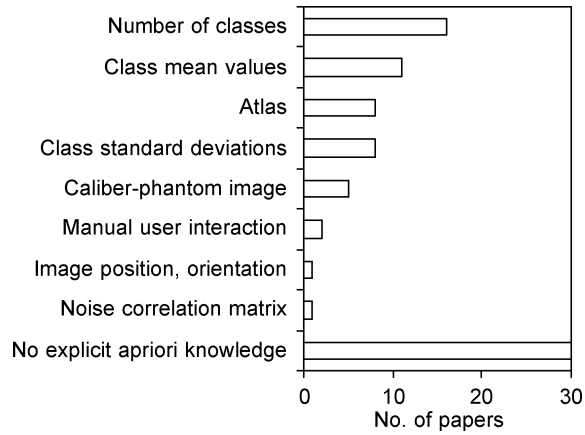

Fig. 8. Commonly used *a priori* knowledge.

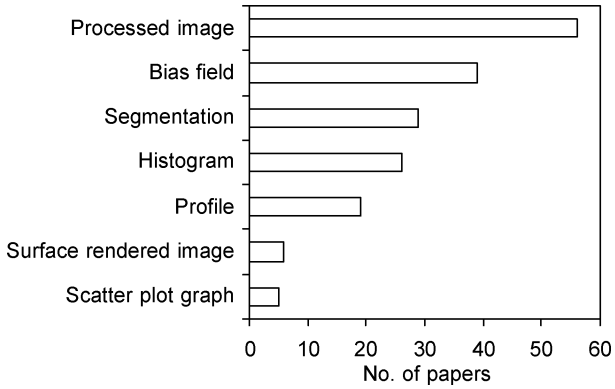

Fig. 9. The most common qualitative evaluation approaches.

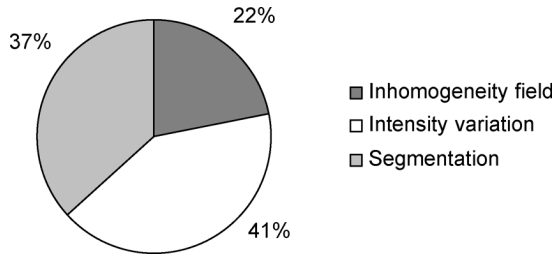

Fig. 10. The proportions of quantitative evaluation approaches.

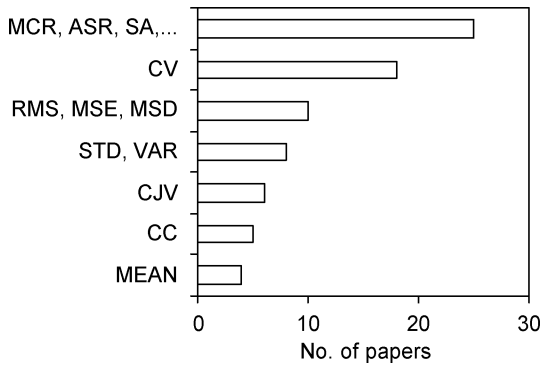

Fig. 11. The most common quantitative metrics.

uations are studied first, followed by the most common quantitative metrics, comparative evaluation, and finally the evaluation datasets.

Qualitative Evaluation: Fig. 9 gives the most frequent applied qualitative evaluation methodologies. In the majority of papers, correction results were illustrated by at least one figure, showing the original and corrected image side by side. Extracted inhomogeneity field, intensity coded segmentation results, and histograms of intensities or inhomogeneity fields demonstrated the correction results in about 50% of the papers. Although intensity or inhomogeneity field profiles are simple and illustrative, they have been used in only one third of publications, while scatter plots and volume rendering appeared even more rarely.

Quantitative Evaluation: The pie chart in Fig. 10 illustrates the proportions of the three quantitative evaluation approaches, i.e., inhomogeneity field, intensity variation and indirect segmentation based. Only in 22% of publications quantitative evaluations was performed on the inhomogeneity fields, most likely because the ground true field was not available. The segmentation based approach, evaluating the segmentation results produced by intensity inhomogeneity correction and segmentation, was more frequently (37%) used. The most frequent approach was the intensity variation based evaluation (41%). This was most likely due to its simplicity, requiring neither ground truth inhomogeneity fields nor full gold standard segmentations.

In Fig. 11, the quantitative metrics are sorted according to the number of publications they appeared in. The most frequent segmentation based measures, such as MCR, ASR, SA, and FPF, are summed up in the top column. The CV, which is an intensity variation based measure, was the second most frequently used, followed by the inhomogeneity field based measures (rms, MSE, and MSD). The latter measures require ground truth inhomogeneity data that limited evaluation to simulated images. The same holds for the correlation coefficient. The CJV is ranked quite high, considering that it was proposed lately in 2001.

Comparative Evaluation: Calculating a well defined quantitative evaluation metrics on a publicly available ground truth image dataset enables objective comparative evaluation of various methods. However, a correct and optimal implementation of others methods is not straightforward and may require interaction with different authors, even if the code is publicly available. Moreover, since the results of comparative evaluations may be highly data dependent, it is important that papers about new methods all use the same standard datasets and evaluation metrics.

Fig. 12 shows the inhomogeneity correction methods that were most frequently used for comparative evaluations. The nonparametric nonuniformity normalization (N3) method [15] has obviously become the standard method against which other methods are compared. The fuzzy segmentation method (FCM) and its adaptive version (AFCM) [20] were also often used for comparative evaluations. HUM [40], [47] was usually compared to variations of this method. Adaptive segmentation (AFGM-EM) [25], Adaptive Markov random field segmentation (AMRF) [88] and statistical parametric mapping based methods (SPM99) [89] have each been comparatively evaluated only twice. Most methods, 17 in total, have been comparatively evaluated only once and are thus not shown in Fig. 12.

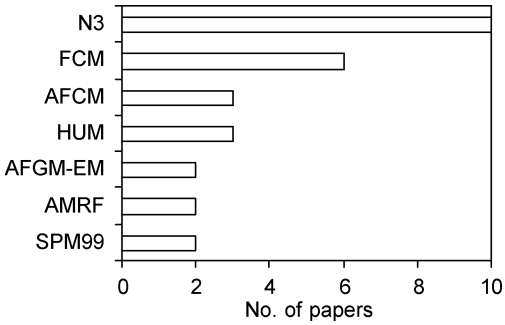

Fig. 12. Most frequently used methods for comparative evaluations.

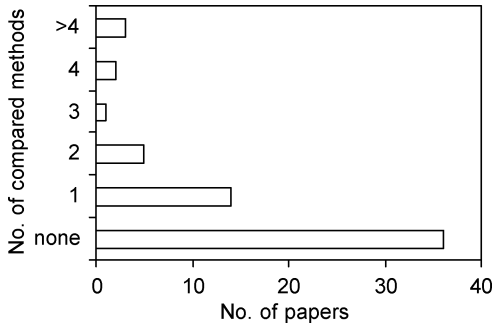

Fig. 13. The number of methods used for comparative evaluations.

Fig. 13 illustrates the number of methods that were compared in the analyzed publications. Surprisingly, in 35 out of 60 publications no comparative evaluations were reported. Studies including comparisons of three or more methods were very rare. The papers before 1994 contained no comparisons, while nowadays in more than two-thirds of publications the novel method is compared to at least one other method.

Evaluation Image Datasets: The reliability of quantitative evaluation depends on the quality of the gold standard. Simulated images seem attractive in this respect but usually real images, better representing anatomical and image quality variations, are needed to objectively estimate the quality and feasibility of the inhomogeneity correction methods. The form of the gold standard depends on the evaluation metrics. In inhomogeneity field based evaluations, gold standard inhomogeneity fields are usually obtained by fitting inhomogeneity fields to intensities at reference points manually selected within the same tissue. Intensity variation based evaluations require either partial image segmentation, i.e., extractions of small subregions of pure tissue regions, or complete segmentation. The latter is also required for segmentation based evaluations. Because complete segmentation is far more demanding then partial, it is obtained either manually or semiautomatically. The analysis of publications revealed that the number of manually labeled real images was small from the beginning and is lately even decreasing, indicating that quantitative evaluations are nowadays mostly conducted on simulated and much less on manually labeled real images.

Other important aspects of evaluation, directly related to the evaluation reliability, are the size of image datasets and type of images. Three types of images have been typically used: synthetic, containing abstract objects such as checkerboards, simulated (e.g., the BrainWeb simulator [90]), and real (e.g., internet brain segmentation repository[1]). Fig. 14 shows that in general the median number of images used for evaluation increases. However, the minimal number of images in some publications still remains surprisingly low.

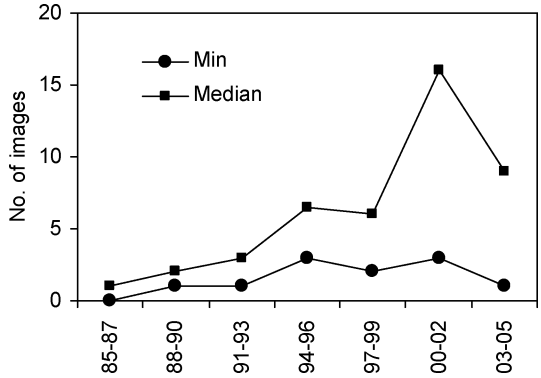

Fig. 14. Median and minimal number of all images used for evaluations.

The steep rise of the median number of images used for evaluations in the last decade is related to the development of the BrainWeb simulator [90]. Considering all the publications after 1997, 41% of the simulated images came from this simulator. In the last decade, the median number of real images remains constantly around four. Surprisingly, there are still publications not demonstrating the performance of novel inhomogeneity correction methods on real images. Although the number of papers with five or more real images used for evaluation is constantly increasing, the proportion of such papers does still not exceed 10% of the published papers.

D. Applications

All the 60 analyzed publications are summarized in Table I, listed in the order of publication year and categorized according to the inhomogeneity correction method, magnetic field strength, type and number of images used for evaluation, and imaged anatomical structures. Methods that are, to the best of our knowledge, available on the internet are marked by $(*)$ and corresponding internet addresses are given at the bottom of Table I.

Table I reveals that most of the published inhomogeneity correction methods were applied on T1 brain images acquired by 1.5T scanners. Applications on images of other anatomical structures, sequences and scanners were very seldom, which indicates that there is still much room for further development of inhomogeneity correction methods.

Numerous applications and evaluations of methods on MR images of human brain reflect the fact that the majority of quantitative image analysis research in the last two decades has been concentrated on the brain. As a result, many published methods are highly brain specific and cannot be applied to other anatomical structures without major modifications. Table I shows that 21 out of 22 segmentation based methods were applied exclusively to brain images. Therefore, the application of the most advanced segmentation based methods, especially those utilizing special *a priori* knowledge, to images of other

[1]http://www.cma.mgh.harvard.edu/ibsr/

TABLE I
SUMMARY OF THE ANALYZED PUBLICATIONS

Year	Ref	1st author	Method	Fields	Sequences	Images	Anatomical structures
1986	[1]	Haselgrove	filtering		PD	1	brain
	[2]	McVeigh	prospective	0.15T	T1	2	brain, lungs
1987	[8]	Axel	prospective	1.4T	T1	1	wrist
	[9]	Condon	prospective	0.15T	T1	3	brain, head
1988	[32]	Brey	prospective	2T		3	rabbit, tomato
	[33]	Narayana	prospective	1.5T	T2	1	spine
1993	[48]	Dawant	surface fitting	1.5T	T1,T2,PD	10	brain
	[31]	Tincher	prospective	0.5T	T1	2	body, liver
	[10]	Wicks	prospective	0.5T	PD	4	brain
1994	[3]	Simmons	prospective	1.5T	PD	3	brain
1995	[53]	Meyer	surface fitting	0.35T, 1.5T	T1	3	brain, abdominal, breast
1996	[41]	Johnston	filtering	1.5T	T2,PD	10	brain
	[25]	Wells	segmentation	1.5T	T2,PD	1105	brain
1997	[23]	Guillemaud	segmentation	1.5T	PD	5	brain, breast
	[55]	Koivula	surface fitting	1T	T1,T2	4	brain, spine, liver
	[66]	Rajapakse	segmentation	1.5T	T1	11	brain
1998	[36]	Mihara	prospective	7T		2	phantom
	* [15]	Sled	histogram	1.5T	T1,T2,PD	25	brain
	[70]	Pham	segmentation		T1	4	brain
	[63]	Rajapakse	segmentation	1.5T	T1	67	brain
	[40]	Brinkmann	filtering		T1,T2,PD	3	brain
1999	* [24]	van Leemput	segmentation	1.5T	T1,T2,PD	19	brain
	[54]	Lai	surface fitting		T1	6	brain, body, spine, breast
	[56]	Vokurka	surface fitting	1.5T	T1	6	brain, spine, shoulder
	[20]	Pham	segmentation		T1,T2,PD	3	brain
	[68]	van Leemput	segmentation	1T, 1.5T	T1,T2,PD	40	brain
2000	* [42]	Cohen	filtering	3T	T1	3	brain
	* [28]	Mangin	histogram	1.5T, 3T	T1	52	brain, head
	* [19]	Styner	histogram		T1,T2,PD	50	brain, breast
	* [89]	Ashburner	segmentation		T1,T2,PD	9	brain
2001	[79]	Solanas	histogram		T1,T2,PD	6	brain
	* [26]	Likar	histogram	1T	T1,T2,PD	71	brain, breast, prostate
	[16]	Prima	segmentation	1.5T	T1	16	brain
	[43]	Zhou	filtering	0.5T	T2	16	brain
	[57]	Vokurka	surface fitting	1.5T	T1	20	brain
	[60]	van Leemput	segmentation	1.5T	T1,T2,PD	600	brain
	* [67]	Zhang	segmentation		T1	6	brain
	* [21]	Shattuck	histogram		T1	35	brain
2002	[75]	Likar	segmentation		T1,T2,PD	42	brain
	[29]	Collewet	prospective	0.2T	T1	5	phantom, salmon
	[49]	Zhuge	surface fitting	1.5T	PD	1000	head, breast, abdominal,…
	[72]	Ahmed	segmentation	1.5T	T1,T2	4	brain
	[58]	Andersen	segmentation	1.5T	T1,T2,PD	3	brain
2003	[34]	Lai	prospective			4	heart
	[35]	Fan	prospective	1.5T	T1,T2	8	brain, cardiac, prostate
	[65]	Gerig	segmentation	3T	T1	10	brain
	[64]	Kim	segmentation		T1	4	brain
	[73]	Liew	segmentation		T1	5	brain
2004	[62]	Bansal	segmentation		T1	2	brain
	[83]	Studholme	other	1.5T	T1	9	brain
	[30]	Gispert	segmentation	0.5T, 1.5T	T1	29	brain
	[51]	Hernandez	surface fitting		T1	1	brain
	[80]	Learned	histogram		T1	26	brain
	[46]	Lewis	filtering	1.5T	T1	47	brain
	[81]	Vovk	histogram		T1,T2,PD	31	brain
	[52]	Milles	surface fitting		T1,T2	8	brain
2005	[84]	Luo	other	1.5T	T1,T2,PD	50	brain
	[59]	Li	segmentation	1.5T	T1,T2	4	brain
	* [22]	Ashburner	segmentation		T1,T2,PD	9	brain
	[50]	Vemuri	surface fitting	1.5T, 3T	T1,T2,PD	9	brain, neck
www	* [15]	Sled	*http://www.bic.mni.mcgill.ca/software/N3/*				
	* [24]	van Leemput	*http://www.medicalimagecomputing.com/*				
	* [42]	Cohen	*http://airto.bmap.ucla.edu/BMCweb/SharedCode/*				
	* [28]	Mangin	*http://brainvisa.info/*				
	* [19]	Styner	*http://www.itk.org/HTML/MRIBiasCorrection.htm*				
	* [26]	Likar	*http://lit.fe.uni-lj.si/shading*				
	* [67]	Zhang	*http://www.fmrib.ox.ac.uk/analysis/research/fast/*				
	* [21]	Shattuck	*http://brainsuite.usc.edu/*				
	* [22]	Ashburner	*http://www.fil.ion.ucl.ac.uk/spm/*				

anatomical structures and also to brain images with pathology may not be straightforward.

The applicability of a certain inhomogeneity correction method depends also on the number of different MR image

sequences that can be corrected without special modifications. Table I illustrates that most of the methods were tested on T1 and much less on T2 and PD weighted images. Only 17 out of 60 papers reported the results on T1, T2, and PD weighted images. Another important application issue is the strength of the magnetic field, which determines the magnitude and shape of intensity inhomogeneity. Table I shows that the great majority of images were acquired by 1.5T scanners. Applications of correction methods to images acquired by modern high magnetic field scanners of 3T and more, typically exhibiting more dynamic inhomogeneities, are still very few.

VIII. Discussion

An overview of methods for intensity inhomogeneity correction in MRI has been presented. Intense research in the field, reflected in the number of publications, started after 1996, when inhomogeneity correction methods were first combined with segmentation [25]. Some promising and fundamentally new approaches, such as high-frequency maximization [15], and information minimization [26], [77], appeared a few years latter. In the last two years, two promising approaches have been proposed. The first is based on merging inhomogeneity correction, segmentation and nonrigid registration [22], while the second is based on reconstructing the image from high-frequency spectra by singularity function analysis [84]. Recently, considerable effort has been put into automating the methods and minimizing the number of parameters that have to be tuned manually, thus making the methods less user-dependant, more practical, objective, and potentially applicable to various images. Because certain parameters will always have to be set empirically, parameter perturbations have been studied in the past [15], [20], [53], [59], [60], [83], however, many more systematical studies on this subject will have to be conducted in the future.

The analysis of numerous publications revealed that the majority of evaluations was conducted on human brain images. The choice of MRI brain images was most probably driven by the clinical interest, excellent image quality, lack of motion, image and ground truth availability, and by challenging image analysis problems and applications related to the brain. When intensity inhomogeneity correction methods had been applied to images of other structures than the brain they were usually only qualitatively evaluated. The main reason for less intensively using images of other anatomical structures is the lack of gold standard in the form of manually labeled tissues, selected regions of interest or inhomogeneity fields fitted to manually selected reference points [83]. The accuracy of a manually defined gold standard is generally problematic because a large number of voxels has to be manually labeled in 3-D. One practical solution to this problem is to first perform some supervised segmentation by a simple segmentation method [28], followed by manual correction of segmentation errors [12], [49]. Another solution is to manually segment or label only small regions or volumes of pure tissue [15], [26], [57]. That manual labeling had been problematic was acknowledged in [25], where the authors asked five human experts to segment brain tissues. However, interexpert variability was usually smaller than the accuracy of the automatic method [60], [63].

While the quality of real image gold standards remains questionable, image simulators, such as the BrainWeb [91], offer highly accurate reference data but inherently lack natural anatomical variability and image acquisition artifacts that are usually encountered in real images. The results of the methods published in the last few years indicate that simulated images can be almost perfectly corrected so that improvements in the performance of a novel method can only be revealed when tested on real images. The authors in [83] declined to use simulated data because these did not simulate tissue substructures and phenomena that were typical to the aging human brain. Therefore, better simulators that would model additional anatomical phenomena and image artifacts, especially the ones that are encountered in high magnetic field devices of 3T and more, are needed. There were already some efforts in this direction [6], [92].

The main goal of comparative quantitative evaluation of different methods is to objectively demonstrate the performances of existing and novel methods. However, the essential prerequisites for comparative evaluations are publicly available ground truth data, objective evaluation metrics, credible implementations of the methods and corresponding parameter selection, each of which is a challenging problem per se. A lack of publicly available ground truth data can be attributed to three reasons, namely to legal issues, to the high amount of required resources to obtain such data and to the lack of will to share the data, which is too often (un)justified by the first two reasons. The availability of evaluation metrics seems to be a rather well solved problem, although there is still much room for improvements, especially for highly task-specific purposes. Finally, obtaining credible implementations of the methods and corresponding parameter selection is highly problematic, requiring a lot of interaction with the authors, even when the code is publicly available. A list of some publicly available inhomogeneity correction methods can be found at the bottom of Table I. Some problems associated with code implementation can be avoided by assessing the original implementation via internet. An example of this approach is the implementation of the information minimization method[2] [26]. However, the most feasible solutions to the problems associated with comparative evaluations is either to conduct a joint evaluation by several authors of inhomogeneity correction methods as in [13], or to construct and maintain a blind evaluation site on the internet, which strongly calls for implementation.

Besides evaluation issues, possible future challenges of intensity inhomogeneity correction lie in the correction of high magnetic field MR images that exhibit severe intensity inhomogeneity [93]–[95], in multispectral intensity inhomogeneity correction, and in applications to MR images of other anatomical structures than the human brain.

IX. Conclusion

Methods for intensity inhomogeneity correction in MR images have been reviewed. The methods and validation approaches were classified according to various features. Additional insight into the field of intensity inhomogeneity

[2]http://lit.fe.uni-lj.si/shading

correction was provided by a detailed analysis of numerous publications that appeared in the last two decades. A number of important issues have been emphasized, indicating that intensity inhomogeneity correction is still not a completely solved problem. Because of this and also because of the evolving MRI technology and associated applications, the problem of intensity inhomogeneity correction will certainly continue to receive a lot of scientific attention in the future. Besides, validation issues should receive much more attention than in the past.

REFERENCES

[1] J. Haselgrove and M. Prammer, "An algorithm for compensating of surface-coil images for sensitivity of the surface coil," *Magn. Reson. Imag.*, vol. 4, pp. 469–472, 1986.

[2] E. R. McVeigh, M. J. Bronskill, and R. M. Henkelman, "Phase and sensitivity of receiver coils in magnetic resonance imaging," *Med. Phys.*, vol. 13, pp. 806–814, 1986.

[3] A. Simmons, P. S. Tofts, G. J. Barker, and S. R. Arridge, "Sources of intensity nonuniformity in spin echo images at 1.5 T," *Magn. Reson. Med.*, vol. 32, pp. 121–128, 1994.

[4] Z.-P. Liang and P. C. Lauterbur, *Principles of Magnetic Resonance Imaging: A Signal Processing Perspective*. Piscataway, NJ: IEEE Press, 1999.

[5] M. D. Keiper, R. I. Grossman, J. A. Hirsch, L. Bolinger, I. L. Ott, L. J. Mannon, C. P. Langlotz, and D. L. Kolson, "MR identification of white matter abnormalities in multiple sclerosis: A comparison between 1.5 T and 4 T," *AJNR Amer. J. Neuroradiol.*, vol. 19, pp. 1489–1493, 1998.

[6] M. Alecci, C. M. Collins, M. B. Smith, and P. Jezzard, "Radio frequency magnetic field mapping of a 3 Tesla birdcage coil: Experimental and theoretical dependence on sample properties," *Magn. Reson. Med.*, vol. 46, pp. 379–385, 2001.

[7] Z. Chen, S. S. Li, J. Yang, D. Letizia, and J. Shen, "Measurement and automatic correction of high-order B0 inhomogeneity in the rat brain at 11.7 Tesla," *Magn. Reson. Imag.*, vol. 22, pp. 835–842, 2004.

[8] L. Axel, J. Costantini, and J. Listerud, "Intensity correction in surface-coil MR imaging," *AJR Amer. J. Roentgenol.*, vol. 148, pp. 418–420, 1987.

[9] B. R. Condon, J. Patterson, D. Wyper, A. Jenkins, and D. M. Hadley, "Image non-uniformity in magnetic resonance imaging: Its magnitude and methods for its correction," *Br. J. Radiol.*, vol. 60, pp. 83–87, 1987.

[10] D. A. Wicks, G. J. Barker, and P. S. Tofts, "Correction of intensity nonuniformity in MR images of any orientation," *Magn. Reson. Imag.*, vol. 11, pp. 183–196, 1993.

[11] J. G. Sled and G. B. Pike, "Standing-wave and RF penetration artifacts caused by elliptic geometry: An electrodynamic analysis of MRI," *IEEE Trans. Med Imag.*, vol. 17, no. 4, pp. 653–662, Aug. 1998.

[12] R. P. Velthuizen, J. J. Heine, A. B. Cantor, H. Lin, L. M. Fletcher, and L. P. Clarke, "Review and evaluation of MRI nonuniformity corrections for brain tumor response measurements," *Med. Phys.*, vol. 25, pp. 1655–1666, 1998.

[13] J. B. Arnold, J. S. Liow, K. A. Schaper, J. J. Stern, J. G. Sled, D. W. Shattuck, A. J. Worth, M. S. Cohen, R. M. Leahy, J. C. Mazziotta, and D. A. Rottenberg, "Qualitative and quantitative evaluation of six algorithms for correcting intensity nonuniformity effects," *Neuroimage*, vol. 13, pp. 931–943, 2001.

[14] B. Belaroussi, J. Milles, S. Carme, Y. M. Zhu, and H. Benoit-Cattin, "Intensity non-uniformity correction in MRI: Existing methods and their validation," *Med. Image Anal.*, vol. 10, pp. 234–246, 2006.

[15] J. G. Sled, A. P. Zijdenbos, and A. C. Evans, "A nonparametric method for automatic correction of intensity nonuniformity in MRI data," *IEEE Trans. Med. Imag.*, vol. 17, no. 1, pp. 87–97, Feb. 1998.

[16] S. Prima, N. Ayache, T. Barrick, and N. Roberts, "Maximum likelihood estimation of the bias field in MR brain images: Investigating different modelings of the imaging process," in *Proc. Med. Image Comput. Computer-Assist. Intervention—MICCAI 2001: 4th Int. Conf.*, Utrecht, The Netherlands, 2001, vol. 2208, pp. 811–819.

[17] J. Sijbers, A. J. den Dekker, P. Scheunders, and D. Van Dyck, "Maximum-likelihood estimation of Rician distribution parameters," *IEEE Trans. Med Imag.*, vol. 17, no. 3, pp. 357–361, Jun. 1998.

[18] C. Brechbühler, G. Gerig, and G. Székely, "Compensation of spatial inhomogeneity in MRI based on a multi-valued image model and a parametric bias estimate," in *Proc. VBC '96 Visualization Biomed. Comput.*, New York, 1996, pp. 141–146.

[19] M. Styner, C. Brechbuhler, G. Szekely, and G. Gerig, "Parametric estimate of intensity inhomogeneities applied to MRI," *IEEE Trans. Med. Imag.*, vol. 19, no. 3, pp. 153–165, Mar. 2000.

[20] D. L. Pham and J. L. Prince, "Adaptive fuzzy segmentation of magnetic resonance images," *IEEE Trans. Med. Imag.*, vol. 18, no. 9, pp. 737–752, Sep. 1999.

[21] D. W. Shattuck, S. R. Sandor-Leahy, K. A. Schaper, D. A. Rottenberg, and R. M. Leahy, "Magnetic resonance image tissue classification using a partial volume model," *Neuroimage*, vol. 13, pp. 856–876, 2001.

[22] J. Ashburner and K. J. Friston, "Unified segmentation," *Neuroimage*, vol. 26, pp. 839–851, 2005.

[23] R. Guillemaud and M. Brady, "Estimating the bias field of MR images," *IEEE Trans. Med Imag.*, vol. 16, no. 3, pp. 238–251, Jun. 1997.

[24] K. Van Leemput, F. Maes, D. Vandermeulen, and P. Suetens, "Automated model-based bias field correction of MR images of the brain," *IEEE Trans. Med. Imag.*, vol. 18, no. 10, pp. 885–896, Oct. 1999.

[25] W. M. Wells, III, W. E. L. Grimson, R. Kikins, and F. A. Jolezs, "Adaptive segmentation of MRI data," *IEEE Trans. Med. Imag.*, vol. 15, no. 8, pp. 429–442, Aug. 1996.

[26] B. Likar, M. A. Viergever, and F. Pernuš, "Retrospective correction of MR intensity inhomogeneity by information minimization," *IEEE Trans. Med. Imag.*, vol. 20, no. 12, pp. 1398–1410, Dec. 2001.

[27] B. Likar, J. B. Maintz, M. A. Viergever, and F. Pernuš, "Retrospective shading correction based on entropy minimization," *J. Microsc.*, vol. 197, pt. 3, pp. 285–295, 2000.

[28] J.-F. Mangin, "Entropy minimization for automatic correction of intensity nonuniformity," presented at the IEEE Workshop Math. Methods Biomed. Image Anal. , Hilton Head Island, SC, 2000.

[29] G. Collewet, A. Davenel, C. Toussaint, and S. Akoka, "Correction of intensity nonuniformity in spin-echo T(1)-weighted images," *Magn. Reson. Imag.*, vol. 20, pp. 365–373, 2002.

[30] J. D. Gispert, S. Reig, J. Pascau, J. J. Vaquero, P. Garcia-Barreno, and M. Desco, "Method for bias field correction of brain T1-weighted magnetic resonance images minimizing segmentation error," *Hum. Brain Mapp.*, vol. 22, pp. 133–144, 2004.

[31] M. Tincher, C. R. Meyer, R. Gupta, and D. M. Williams, "Polynomial modeling and reduction of RF body coil spatial inhomogeneity in MRI," *IEEE Trans. Med. Imag.*, vol. 12, no. 2, pp. 361–365, Jun. 1993.

[32] W. W. Brey and P. A. Narayana, "Correction for intensity falloff in surface coil magnetic resonance imaging," *Med. Phys.*, vol. 15, pp. 241–245, 1988.

[33] P. A. Narayana, W. W. Brey, M. V. Kulkarni, and C. L. Sievenpiper, "Compensation for surface coil sensitivity variation in magnetic resonance imaging," *Magn Reson Imag.*, vol. 6, pp. 271–274, 1988.

[34] S. H. Lai and M. Fang, "A dual image approach for bias field correction in magnetic resonance imaging," *Magn. Reson. Imag.*, vol. 21, pp. 121–125, 2003.

[35] A. Fan, W. M. Wells, J. W. Fisher, M. Cetin, S. Haker, R. Mulkern, C. Tempany, and A. S. Willsky, "A unified variational approach to denoising and bias correction in MR," *Inf. Process Med Imag.*, vol. 18, pp. 148–159, 2003.

[36] H. Mihara, N. Iriguchi, and S. Ueno, "A method of RF inhomogeneity correction in MR imaging," *Magma*, vol. 7, pp. 115–120, 1998.

[37] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger, "SENSE: Sensitivity encoding for fast MRI," *Magn. Reson. Med.*, vol. 42, pp. 952–962, 1999.

[38] J. Y. Chiou, C. B. Ahn, L. T. Muftuler, and O. Nalcioglu, "A simple simultaneous geometric and intensity correction method for echo-planar imaging by EPI-based phase modulation," *IEEE Trans. Med Imag.*, vol. 22, no. 2, pp. 200–205, Feb. 2003.

[39] D. L. Thomas, E. De Vita, R. Deichmann, R. Turner, and R. J. Ordidge, "3-D MDEFT imaging of the human brain at 4.7 T with reduced sensitivity to radiofrequency inhomogeneity," *Magn. Reson. Med.*, vol. 53, pp. 1452–1458, 2005.

[40] B. H. Brinkmann, A. Manduca, and R. A. Robb, "Optimized homomorphic unsharp masking for MR grayscale inhomogeneity correction," *IEEE Trans. Med. Imag.*, vol. 17, no. 2, pp. 161–171, Apr. 1998.

[41] B. Johnston, M. S. Atkins, B. Mackiewich, and M. Anderson, "Segmentation of multiple sclerosis lesions in intensity corrected multispectral MRI," *IEEE Trans. Med. Imag.*, vol. 15, no. 2, pp. 154–169, Apr. 1996.

[42] M. S. Cohen, R. M. DuBois, and M. M. Zeineh, "Rapid and effective correction of RF inhomogeneity for high field magnetic resonance imaging," *Hum. Brain Mapp.*, vol. 10, pp. 204–211, 2000.

[43] L. Q. Zhou, Y. M. Zhu, C. Bergot, A. M. Laval-Jeantet, V. Bousson, J. D. Laredo, and M. Laval-Jeantet, "A method of radio-frequency inhomogeneity correction for brain tissue segmentation in MRI," *Comput. Med. Imag. Graph.*, vol. 25, pp. 379–389, 2001.

[44] J. Russ, *The Image Processing Handbook*, 2nd ed. Piscataway, NJ: IEEE Press, 1995.

[45] D. Tomazevic, B. Likar, and F. Pernus, "Comparative evaluation of retrospective shading correction methods," *J. Microsc.*, vol. 208, pp. 212–223, 2002.

[46] E. B. Lewis and N. C. Fox, "Correction of differential intensity inhomogeneity in longitudinal MR images," *Neuroimage*, vol. 23, pp. 75–83, 2004.

[47] P. A. Narayana and A. Borthakur, "Effect of radio frequency inhomogeneity correction on the reproducibility of intra-cranial volumes using MR image data," *Magn. Reson. Med.*, vol. 33, pp. 396–400, 1994.

[48] B. M. Dawant, A. P. Zijdenbos, and R. A. Margolin, "Correction of intensity variations in MR images for computer-aided tissues classification," *IEEE Trans. Med. Imag.*, vol. 12, no. 4, pp. 770–781, Dec. 1993.

[49] Y. Zhuge, J. K. Udupa, J. Liu, P. K. Saha, and T. Iwanage, "Scale-based method for correcting background intensity variation in acquired images," in *Proc. SPIE Med. Imag. 2002: Image Process.*, 2002, vol. 4686, pp. 1103–1111.

[50] P. Vemuri, E. G. Kholmovski, D. L. Parker, and B. E. Chapman, "Coil sensitivity estimation for optimal SNR reconstruction and intensity inhomogeneity correction in phased array MR imaging," presented at the Inf. Process. Med. Imag.: 19th Int. Conf. (IPMI 2005), Glenwood Springs, CO, 2005.

[51] J. A. Hernandez, M. L. Mora, E. Schiavi, and P. Toharia, "RF inhomogeneity correction algorithm in magnetic resonance imaging," presented at the 5th Int. Symp. Biol. Medical Data Anal. (ISBMDA 2004), Barcelona, Spain, 2004.

[52] J. Milles, Y. M. Zhu, N. Chen, L. P. Panych, G. Gimenez, and C. R. Guttmann, "MRI intensity nonuniformity correction using simultaneously spatial and gray-level histogram information," in *Proc. SPIE Med. Imag. 2004: Physiol., Function, Structure Med. Images*, 2004, vol. 5370, pp. 734–742.

[53] C. R. Meyer, P. H. Bland, and J. Pipe, "Retrospective correction of intensity inhomogeneities in MRI," *IEEE Trans. Med. Imag.*, vol. 14, no. 1, pp. 36–41, Mar. 1995.

[54] S. H. Lai and M. Fang, "A new variational shape-from-orientation approach to correcting intensity inhomogeneities in magnetic resonance images," *Med. Image Anal.*, vol. 3, pp. 409–424, 1999.

[55] A. Koivula, J. Alakuijala, and O. Tervonen, "Image feature based automatic correction of low-frequency spatial intensity variations in MR images," *Magn. Reson. Imag.*, vol. 15, pp. 1167–1175, 1997.

[56] E. A. Vokurka, N. A. Thacker, and A. Jackson, "A fast model independent method for automatic correction of intensity nonuniformity in MRI data," *J. Magn. Reson. Imag.*, vol. 10, pp. 550–562, 1999.

[57] E. A. Vokurka, N. A. Watson, Y. Watson, N. A. Thacker, and A. Jackson, "Improved high resolution MR imaging for surface coils using automated intensity non-uniformity correction: Feasibility study in the orbit," *J. Magn. Reson. Imag.*, vol. 14, pp. 540–546, 2001.

[58] A. H. Andersen, Z. Zhang, M. J. Avison, and D. M. Gash, "Automated segmentation of multispectral brain MR images," *J. Neurosci. Meth.*, vol. 122, pp. 13–23, 2002.

[59] X. Li, L. Li, H. Lu, and Z. Liang, "Partial volume segmentation of brain magnetic resonance images based on maximum a posteriori probability," *Med. Phys.*, vol. 32, pp. 2337–2345, 2005.

[60] K. Van Leemput, F. Maes, D. Vandermeulen, A. Colchester, and P. Suetens, "Automated segmentation of multiple sclerosis lesions by model outlier detection," *IEEE Trans. Med Imag.*, vol. 20, no. 8, pp. 677–688, Aug. 2001.

[61] J. D. Gispert, S. Reig, J. Pascau, R. M. Lazaro, J. J. Vaquero, and M. Desco, "Inhomogeneity correction of magnetic resonance images by minimization of intensity overlapping," presented at the 2003 Int. Conf. Image Process. (ICIP 2003), Barcelona, Catalonia, Spain, 2003.

[62] R. Bansal, L. H. Staib, and B. S. Peterson, "Correcting nonuniformities in MRI intensities using entropy minimization based on an elastic model," presented at the Medical Image Computing and Computer-Assisted Intervention (MICCAI 2004), Saint-Malo, France, 2004.

[63] J. C. Rajapakse and F. Kruggel, "Segmentation of MR images with intensity inhomogeneities," *Image And Vision Comput.*, vol. 16, pp. 165–180, 1998.

[64] S.-G. Kim, S.-K. Ng, G. J. McLachlan, and W. Deming, "Segmentation of brain MR images with bias field correction," presented at the APRS Workshop Digital Image Computing: WDIC 2003, St. Lucia, Brisbane, Australia, 2003.

[65] G. Gerig, M. Prastawa, W. Lin, and J. Gilmore, "Assessing early brain development in neonates by segmentation of high-resolution 3T MRI," presented at the 6th Int. Conf. Medical Image Comput. Computer-Assisted Intervention (MICCAI 2003), , Montreal, ON, Canada, 2003.

[66] J. C. Rajapakse, J. N. Giedd, and J. L. Rapoport, "Statistical approach to segmentation of single-channel cerebral MR images," *IEEE Trans. Med. Imag.*, vol. 16, no. 2, pp. 176–186, Apr. 1997.

[67] Y. Zhang, M. Brady, and S. Smith, "Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm," *IEEE Trans. Med Imaging*, vol. 20, no. 1, pp. 45–57, Jan. 2001.

[68] K. Van Leemput, F. Maes, D. Vandermeulen, and P. Suetens, "Automated model-based tissue classification of MR images of the brain," *IEEE Trans. Med. Imag.*, vol. 18, no. 10, pp. 897–908, Oct. 1999.

[69] J. C. Bezdek, L. O. Hall, and L. P. Clarke, "Review of MR image segmentation techniques using pattern recognition," *Med. Phys.*, vol. 20, pp. 1033–1048, 1993.

[70] D. L. Pham and J. L. Prince, "An adaptive fuzzy C-means algorithm for the image segmentation in the presence of intensity inhomogeneities," *Pattern Recognit. Lett.*, vol. 20, pp. 57–68, 1998.

[71] M. N. Ahmed, S. M. Yamany, A. A. Farag, and T. Moriarty, "Bias field estimation and adaptive segmentation of MRI data using amodified fuzzy C-means algorithm," presented at the IEEE Int. Conf. Comput. Vision Pattern Recognit. (CVPR'99), Fort Collins, CO, 1999.

[72] M. N. Ahmed, S. M. Yamany, N. Mohamed, A. A. Farag, and T. Moriarty, "A modified fuzzy C-means algorithm for bias field estimation and segmentation of MRI data," *IEEE Trans. Med. Imag.*, vol. 21, no. 3, pp. 193–199, Mar. 2002.

[73] A. W. Liew and H. Yan, "An adaptive spatial fuzzy clustering algorithm for 3-D MR image segmentation," *IEEE Trans. Med. Imag.*, vol. 22, no. 9, pp. 1063–1075, Sep. 2003.

[74] C. Zhu and T. Jiang, "Multicontext fuzzy clustering for separation of brain tissues in magnetic resonance images," *Neuroimage*, vol. 18, pp. 685–696, 2003.

[75] B. Likar, J. Derganc, and F. Pernuš, "Segmentation-based retrospective correction of intensity non-uniformity in multi-spectral MR images," in *Proc. SPIE Med. Imag. 2002: Image Process.*, San Diego, 2002, vol. 4684, pp. 1531–1540.

[76] J. Derganc, B. Likar, and F. Pernuš, "Nonparametric segmentation of multispectral MR images incorporating spatial and intensity information," in *Proc. SPIE Med. Imag. 2002: Image Process.*, San Diego, CA, 2002, vol. 4684, pp. 391–400.

[77] P. A. Viola, "Alignment by maximization of mutual information," Ph.D. dissertation, Mass. Inst. Technol., Cambridge, 1995.

[78] B. Likar, M. A. Viergever, and F. Pernuš, "Retrospective correction of mr intensity inhomogeneity by information minimization," presented at the 3rd Int. Conf. Med. Image Comput. Computer-Assisted Intervention (MICCAI 2000), Pittsburgh, PA, 2000.

[79] E. Solanas and J.-P. Thiran, "Exploiting voxel correlation for automated MRI bias field correction by conditional entropy minimization," presented at the 4th Int. Medical Image Comput. Computer-Assist. Intervent. (MICCAI 2001), Utrecht, The Netherlands, 2001.

[80] E. G. Learned-Miller and P. Ahammad, "Joint MRI bias removal using entropy minimization across images," presented at the Neural Inf. Proces. Syst. Conf. (NIPS), Vancouver, Canada, 2004.

[81] U. Vovk, F. Pernuš, and B. Likar, "MRI intensity inhomogeneity correction by combining intensity and spatial information," *Phys. Med. Biol.*, vol. 49, pp. 4119–4133, 2004.

[82] ——, "Intensity inhomogeneity correction of multispectral MR images," *Neuroimage*, vol. 32, pp. 54–61, 2006.

[83] C. Studholme, V. Cardenas, E. Song, F. Ezekiel, A. Maudsley, and M. Weiner, "Accurate template-based correction of brain MRI intensity distortion with application to dementia and aging," *IEEE Trans. Med Imag.*, vol. 23, no. 1, pp. 99–110, Jan. 2004.

[84] J. Luo, Y. Zhu, P. Clarysse, and I. Magnin, "Correction of bias field in MR images using singularity function analysis," *IEEE Trans. Med. Imag.*, vol. 24, no. 8, pp. 1067–1085, Aug. 2005.

[85] B. Fischl, D. H. Salat, A. J. van der Kouwe, N. Makris, F. Segonne, B. T. Quinn, and A. M. Dale, "Sequence-independent segmentation of magnetic resonance images," *Neuroimage*, vol. 23, no. Suppl 1, pp. S69–S84, 2004.

[86] R. C. Jain, T. O. Binford, M. A. Snyder, Y. Aloimonos, A. Rosenfeld, T. S. Huang, K. W. Bowyer, and J. P. Jones, "Ignorance, myopia, and naivete in computer vision systems," *Comput. Vision, Graphics, Image Process. : Image Understanding*, vol. 53, pp. 112–117, 1991.

[87] A. Montillo, J. K. Udupa, L. Axel, and D. N. Metaxas, "Interaction between noise suppression and inhomogeneity correction in MRI," in *Proc. SPIE Med. Imag. 2003: Image Process.*, 2003, vol. 5032, pp. 1025–1036.

[88] M. X. H. Yan and J. S. Karp, "An adaptive Bayesian approach to three-dimensional MR brain segmentation," in *Proc. Inf. Process. Med. Imag. (IPMI'95)*, 1995, vol. 20, pp. 12–13.

[89] J. Ashburner and K. J. Friston, "Voxel-based morphometry—the methods," *Neuroimage*, vol. 11, pp. 805–821, 2000.

[90] C. A. Cocosco, V. Kollokian, K. R. K. -S. Kwan, and A. C. Evans, "BrainWeb: Online interface to a 3-D MRI simulated brain database," *NeuroImage*, vol. 5, pp. 425–425, 1997.

[91] D. L. Collins, A. P. Zijdenbos, V. Kollokian, J. G. Sled, N. J. Kabani, C. J. Holmes, and A. C. Evans, "Design and construction of a realistic digital brain phantom," *IEEE Trans. Med Imag.*, vol. 17, no. 3, pp. 463–468, Jun. 1998.

[92] H. Benoit-Cattin, G. Collewet, B. Belaroussi, H. Saint-Jalmes, and C. Odet, "The SIMRI project: A versatile and interactive MRI simulator," *J. Magn. Reson.*, vol. 173, pp. 97–115, 2005.

[93] T. S. Ibrahim, C. Mitchell, P. Schmalbrock, R. Lee, and D. W. Chakeres, "Electromagnetic perspective on the operation of RF coils at 1.5–11.7 Tesla," *Magn. Reson. Med.*, vol. 54, pp. 683–690, 2005.

[94] T. S. Ibrahim, A. Kangarlu, and D. W. Chakeress, "Design and performance issues of RF coils utilized in ultra high field MRI: Experimental and numerical evaluations," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 7, pp. 1278–1284, Jul. 2005.

[95] H. Uematsu, L. Dougherty, M. Takahashi, Y. Ohno, M. Nakatsu, M. D. Schnall, and H. Hatabu, "Abdominal imaging at 4 T MR system: A preliminary result," *Eur. J. Radiol.*, vol. 47, pp. 161–163, 2003.