

# Nonrigid Registration of Joint Histograms for Intensity Standardization in Magnetic Resonance Imaging

Florian Jäger\* and Joachim Hornegger, *Member, IEEE*

**Abstract**—A major disadvantage of magnetic resonance imaging (MRI) compared to other imaging modalities like computed tomography is the fact that its intensities are not standardized. Our contribution is a novel method for MRI signal intensity standardization of arbitrary MRI scans, so as to create a pulse sequence dependent standard intensity scale. The proposed method is the first approach that uses the properties of all acquired images jointly (e.g., T1- and T2-weighted images). The image properties are stored in multidimensional joint histograms. In order to normalize the probability density function (pdf) of a newly acquired data set, a nonrigid image registration is performed between a reference and the joint histogram of the acquired images. From this matching a nonparametric transformation is obtained, which describes a mapping between the corresponding intensity spaces and subsequently adapts the image properties of the newly acquired series to a given standard. As the proposed intensity standardization is based on the probability density functions of the data sets only, it is independent of spatial coherence or prior segmentations of the reference and current images. Furthermore, it is not designed for a particular application, body region or acquisition protocol. The evaluation was done using two different settings. First, MRI head images were used, hence the approach can be compared to state-of-the-art methods. Second, whole body MRI scans were used. For this modality no other normalization algorithm is known in literature. The Jeffrey divergence of the pdfs of the whole body scans was reduced by 45%. All used data sets were acquired during clinical routine and thus included pathologies.

**Index Terms**—General intensity scale, intensity normalization, magnetic resonance imaging (MRI), nonrigid registration, signal intensity standardization, whole body MRI.

## I. MOTIVATION

**M**AGNETIC resonance imaging (MRI) is the preferred imaging modality of the brain and many other body regions due to its excellent soft tissue contrast. Susceptibility effects and local inhomogeneities of the coil system on the other

hand can influence signal intensity values. These intensity variations can be separated into two different classes. The first type of variation (*class I*) consists of intensities of the same tissue class which differ throughout a single volume. In literature, this is called intensity inhomogeneity and is generally caused by a gain/bias field. In order to deal with this problem, a variety of algorithms have been developed in the last decade. All of these approaches are based on the assumption that the gain field is very smooth over the whole image domain, and hence, it does not include high-frequency components. A detailed evaluation and summary of many types of algorithms for *inhomogeneity correction* is given in [1]–[3]. The most frequently used methods are variations of homomorphic unsharp masking [4], [5], which directly utilize the smoothness assumption. More sophisticated algorithms either use a segmentation step [6], [7], or make further statistical assumptions about the shape of the intensity probability distributions of the observed images, or the entropy of the images [8]–[11].

Furthermore, most of these methods do not explicitly solve the interscan signal intensity variation problem (*class II*): intensities vary between different scans with the same acquisition protocol even for inpatient studies; thus, a certain measured intensity cannot be associated with a specific tissue class. Solving *class II* problems is called *intensity standardization*. After a successful application of an inhomogeneity correction algorithm all tissue classes with theoretically identical signal intensities have identical gray values in the images. Unfortunately, it is not possible to assign an anatomical meaning to the observed signal intensities, as these depend on the acquisition and the applied bias field correction method. The distinction between intensity inhomogeneities and a missing standard signal intensity scale is illustrated in Fig. 1. It can be seen that inhomogeneity correction algorithms influence the signal intensities locally, whereas intensity standardization methods vary intensities globally.

In this paper, we will focus on the *class II* problem. In general, a standard intensity scale has no direct impact on medical diagnostics by experts; however, volume renderers cannot use standard presets (transfer functions) to visualize certain organs or tissue classes. The physician has to adjust the settings for every single scan. Furthermore, more sophisticated automatic segmentation and quantification methods are needed, as they have to adapt their parameters to the observed image intensities. Additionally, currently a new class of hybrid imaging systems combining MR and positron emission tomography (PET) is being developed. In order to increase the PET image quality,

Manuscript received April 21, 2008; revised July 24, 2008. First published August 15, 2008; current version published December 24, 2008. This work was supported by the Erlangen Graduate School in Advanced Optical Technologies (SAOT) by the German National Science Foundation (DFG) in the framework of the excellence initiative. *Asterisk indicates corresponding author.*

\*F. Jäger is with the Chair of Pattern Recognition, Friedrich-Alexander-University Erlangen-Nuremberg, 91058 Erlangen, Germany (e-mail: jaeger@informatik.uni-erlangen.de).

J. Hornegger is with the Chair of Pattern Recognition, Friedrich-Alexander-University Erlangen-Nuremberg, 91058 Erlangen, Germany (e-mail: hornegger@informatik.uni-erlangen.de).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TMI.2008.2004429

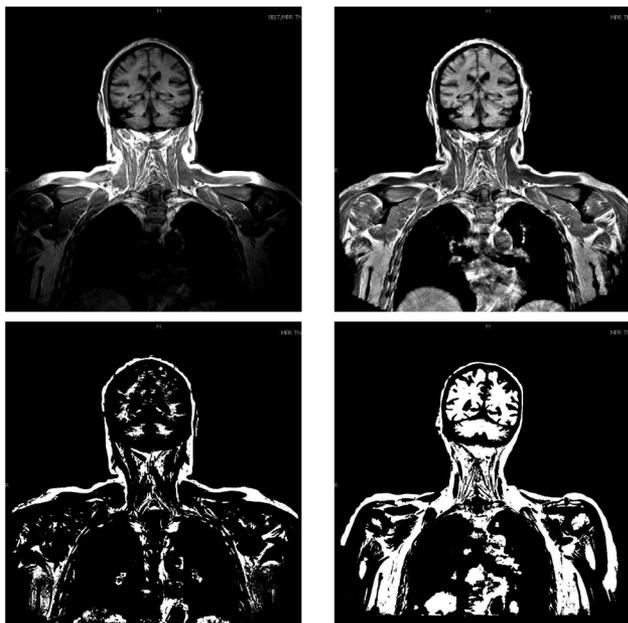


Fig. 1. The distinction of both types of variations (interscan and intrascan inhomogeneities). Upper Row: The left image shows the original FL2D scan of a patient. The right image shows the same slice after gain field correction. Lower Row: In the left image a threshold of 580 is applied to the gain field corrected slice of the upper row. The right image shows a FL2D scan of another patient after gain field correction and thresholding with the same threshold level (580).

an standardized attenuation correction utilizing the MR data has to be performed [12]. For this purpose the MR intensities have to be mapped to attenuation coefficients which correlate to tissue classes. Hence, a second class of approaches dealing with interscan intensity standardization has been developed by several authors. Note that although inhomogeneity correction has been researched considerably [2], intensity standardization has not received the same attention. How intensity standardization and bias correction influence each other is evaluated in [13]. The authors conclude that both steps are necessary but the correction of inhomogeneities has to be done beforehand.

Some intensity standardization methods, like the ones described in Nyúl *et al.* [14] and Ge *et al.* [15], use a 1-D histogram matching approach. First, they detect some landmarks (e.g., percentiles, modes) on a training set of histograms. Then all detected landmarks are averaged, in order to generate a standard landmark set. When a new image is acquired, the detected landmarks of its histogram are matched to the previously computed standard positions. Finally, in order to create a continuous intensity mapping, the histogram positions between the landmarks are linearly interpolated. This standardization method has been evaluated on brain scans but can be potentially applied to other regions of the body, too. Nevertheless, for every new body region and protocol, the corresponding histograms have to be analyzed and appropriate landmarks have to be chosen. Additionally, the intensities between the landmarks are linearly interpolated even though these intensity deformations are nonlinear in reality.

Another histogram based method is proposed by Pierre Hellier [16] who, given an image, estimates a mixture of Gaussians that approximates its histogram. The fitting is done utilizing the

approach presented in [17]. He then computes a polynomial correction function that aligns the mean intensities of the different tissue classes. Unfortunately, this approach is restricted to the head region, as this is the only region where the tissue classes can be estimated by a few distinct categories. Furthermore, it is highly dependent on the quality of the Gaussian fit.

Weisenfeld and Warfield propose a combined intensity standardization and inhomogeneity correction method in [18]. They estimate a multiplicative correction field that adapts the intensity statistics of an acquired MR volume to a previously created model. As model they employ the histogram of a previously acquired image with appropriate image properties. The correction is achieved by minimizing the Kullback–Leibler divergence between the model and the template intensity distribution. The standardization method has been evaluated on brain scans but can be potentially applied to other regions of the body, too. However, this method can only be applied to single band data sets.

A signal intensity standardization method using spatial tissue correlations between a reference and a template image is presented in [19]. In order to match the images, a nonlinear registration algorithm is used. Once the images are aligned, a scalar multiplicative correction weight is computed. Though potentially applicable to all body regions, the standardization is highly dependent on the result of the registration step. Thus, the results will be worse in regions of the body with significant anatomical differences and deformations, like the thorax or abdominal area.

Our contribution in this paper is an algorithm which utilizes all acquired images for intensity standardization by jointly using the pdfs of the acquired series. A nonparametric normalization function is computed between the corresponding intensity spaces. Although we assume that the intervolum signal intensity variations are independent for all acquired data sets, it is possible to separately correct tissue classes which have the same intensity in one image but distinct intensities in the other pulse sequences. Hence small structures (e.g., lesions) can be intensity standardized more reliably. Furthermore, the introduced approach does not rely on any assumptions about the shape of the joint histograms used. Thus, our method is completely independent of the application, region of interest and scanning protocol, as long as there are reference histograms available for the given task. These can be easily computed by choosing representative images for the new body regions and/or acquisition protocols. Because only pdfs are employed, no nonrigid spatial alignment of the data sets has to be done. For this reason, deformations or pathologies hardly influence the normalization results. Moreover, the proposed method is the only approach that can deal with whole body MRI data sets, which have much more complex statistical properties than spatially constrained images, like the head.

In the following section we present our approach for MRI signal intensity standardization. We first describe the theoretical principles underlying our method. We conclude the section by presenting an improvement of our approach regarding complex data sets like whole body MRI images. The next section is on experimental results. First, we present results of the standardization of pathological MRI head images. Then, we evaluate our method using artificially perturbed images. Finally, we show the applicability of the proposed approach to whole body

MRI scans. We have also included a short overview of nonrigid registration in the Appendix.

## II. INTENSITY STANDARDIZATION

The goal of the proposed intensity standardization approach is to find a mapping between the intensities of a  $n$ -tuple of images  $C = (C_1, C_2, \dots, C_n)$ , where  $n$  is the number of images acquired with different modalities and a reference  $n$ -tuple of images  $R = (R_1, R_2, \dots, R_n)$  so that an arbitrary intensity vector  $\mathbf{i} \in \mathcal{I}^n$  describes the same tissue class in both sets with  $\mathcal{I}^n \subseteq \mathbb{R}^n$  being the intensity space of the image sets. The main idea of our method is that this mapping can be approximated by the minimization of the distance between the joint histograms of the two  $n$ -tuples of images. The required joint histograms are of dimensionality  $n$ , corresponding to the number of images. The domain of the histograms is  $\mathcal{I}^n$ . In general, however, it can be scaled to  $[0; 1]^n$  due to the limited number of gray values observed. Because of signal intensity outliers, we do not use the full intensity range but the range up to the 99.8% intensity percentile. If the difference of the number of gray values in the images is big, then either the scaling of the intensities or the registration parameters have to be adapted for each dimension. The reason for this is that otherwise the smoothing of the mapping is stronger in the histogram direction with the larger number of gray values. Note that, at least for real data sets, no plausible transformation of the relative joint histograms can be found, such that, the difference is zero, because the volume of tissue classes in the image tuples  $C$  and  $R$  differ for interpatient as well as inpatient measurements (e.g., anatomical differences, partial volume averaging effects, positioning of the patient). Thus, the search for a mapping between the intensity spaces is equivalent to finding the transformation  $T$  between the corresponding joint histograms which minimizes a given distance measure  $\mathcal{D}$

$$\hat{T} = \arg \min_T \mathcal{D}(\mathcal{H}(R), \mathcal{H}(C); T) \quad (1)$$

with  $\mathcal{H}(R)$  and  $\mathcal{H}(C)$  being the joint histograms of the image tuples  $R$  and  $C$ . If the joint histograms are treated as images and if there are no constraints on  $T$ , this task can be viewed as a nonrigid image registration problem. Note that, although in theory histograms of arbitrary dimensionality can be used, in practice  $n$  should be smaller than five. Otherwise, due to the curse of dimension, the registration results may be no longer satisfactory. Techniques like Parzen estimation can solve the problem of insufficient samples; however, this leads to high computational costs.

For image registration a variety of algorithms are available. A survey about image registration is given in Maintz *et al.* [20] and Hill *et al.* [21]. We employed the variational nonrigid registration approach which was introduced by Modersitzki *et al.* [22]; however, other deformable registrations schemes are applicable, too. The result of this method's optimization is the transformation  $T: \mathbb{R}^n \mapsto \mathbb{R}^n$ . In the context of the registration of multi-dimensional joint histograms, it describes how to transform the gray values of one  $n$ -tuple of images  $C$  such that its intensity distribution best matches the reference distribution, with respect

to the used distance measure and smoother. The objective functional  $\mathcal{J}$  of the nonrigid registration can be written as

$$\mathcal{J}[\mathcal{H}(R), \mathcal{H}(C); \mathbf{u}] = \mathcal{D}[\mathcal{H}(R), \mathcal{H}(C); \mathbf{u}] + \alpha \mathcal{S}[\mathbf{u}] \quad (2)$$

with  $\mathcal{D}$  being the distance measure,  $\mathcal{S}$  being a smoother,  $\alpha$  defining the influence of the smoother on the optimization, and  $\mathbf{u}$  representing the deformation between the joint histograms. We used the sum of squared differences as distance measure and a curvature based smoother. A more detailed description of nonrigid image registration using a variational framework is given in the Appendix. The intensity standardization can be done by

$$\mathbf{i}_{\text{corr}} = \mathbf{i}_{\text{orig}} + \mathbf{u}(\mathbf{i}_{\text{orig}}) = \hat{T}(\mathbf{i}_{\text{orig}}) \quad (3)$$

where  $\mathbf{i}_{\text{orig}} \in \mathcal{I}^n$  describes the intensity vector in the original current image tuple  $C$  and  $\mathbf{i}_{\text{corr}} \in \mathcal{I}^n$  is the intensity vector in the corrected images, respectively. A schematic overview of the standardization process is given in Fig. 2. Here the relationship between the spatial and the intensity domain is illustrated.

If the computed transformation is applied to the joint pdf  $p$  of the current image tuple, it is not guaranteed that the resulting function is still a pdf, as the constraint  $\sum_{\mathbf{x}} (p \circ \hat{T})(\mathbf{x}) = 1$  might be invalid. However, as the derived mapping is applied to images, the resulting pdfs will fulfill the constraint again. Nevertheless, volume preserving nonrigid registration approaches can be used as well [23].

In a preprocessing step, the joint histograms were equalized [24]. As the histogram values in areas with small tissue support are very low, the equalization increases the performance of the registration of these regions by raising the values there and suppressing values of areas with high tissue support. This is very important for data sets acquired with protocols that highlight small structures (e.g., blood vessels or kidneys in TIRM images). Without the equalization step, areas in the joint histograms representing such structures are not treated satisfactorily in the registration process, as small histogram values hardly influence the distance measure. Thus, the registration concentrates on structures in the histograms with high tissue support.

For data sets being "statistically simple," like the head region, the proposed method returns satisfactory results (see Section III). However, the following problems may arise in more complex data sets: 1) tissue classes with a small number of voxels do not have enough support to be transformed in a reliable manner; 2) if a previous bias field correction step has failed, the histograms are blurred and the statistical information of a tissue class is spread to a broad range of gray values. Consequently, it is no longer possible to find a plausible global transformation of the intensity vectors. One straightforward solution to this is to split the data sets into smaller subvolumes. These subvolumes can then be intensity-standardized separately. However, this can still lead to problems if the statistical content of a subvolume is not sufficient for a reliable registration. In order to have sufficient statistical content, a partition should have the same dominating tissue classes as the corresponding partition in the reference image. Furthermore, the histogram has to have a similar morphology as the histograms of the neighboring partitions.

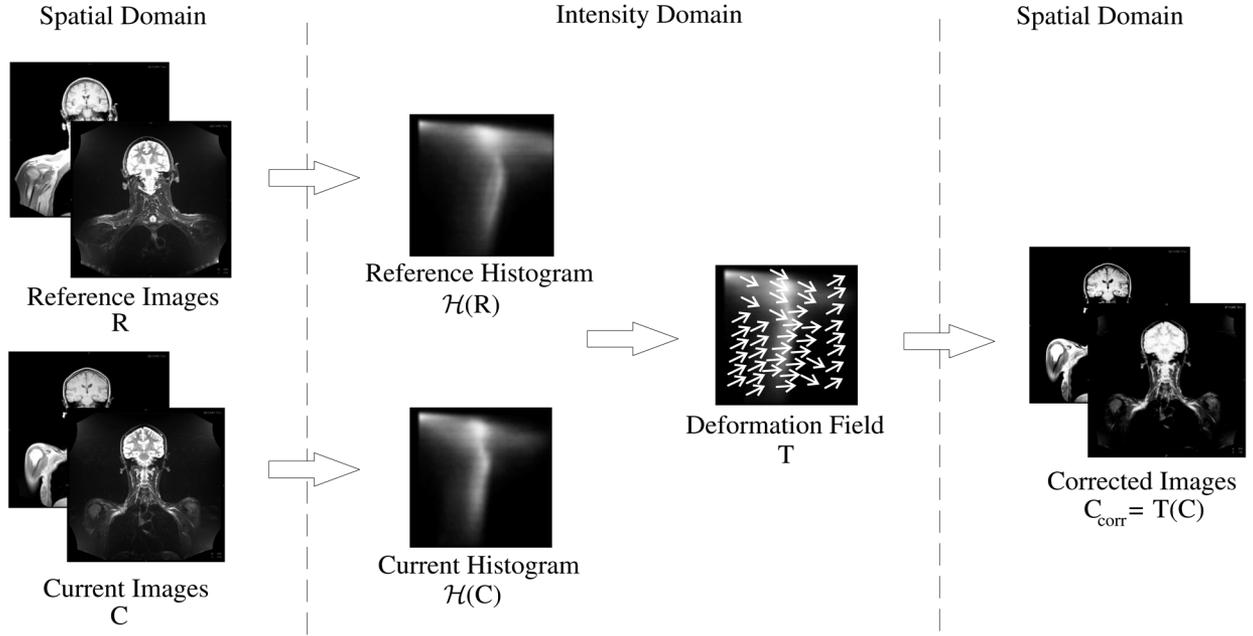


Fig. 2. Schematic illustration of the intensity standardization. First, from the reference images a reference joint histogram is created. This is the training component of the approach. Then from the current MRI images a joint histogram is generated. In the next step these histograms are nonrigidly registered. Using the gained transformation function, the current images are standardized. For visualization purposes, the proposed method is shown for  $n = 2$ , but it is applicable to any  $n$ .

If the sufficient content assumption does not hold, some tissue classes might be transformed in a false way. Furthermore, due to the independent standardization of the subvolumes, intensity discontinuities can occur at the junctions between the partitions [25].

In order to overcome these drawbacks we propose the following subvolume based technique. The core of our method is a new distance measure for the registration of the joint histograms which utilizes the statistical information of neighboring partitions in the joint histograms in regularizing the computed transformation of the intensities. The modified method can be summarized as follows. The data is split into  $K$  partitions. For each partition  $k \in \{1, \dots, K\}$  a joint histogram for the current and one for the reference volume is created. The partitioning of the input data sets is arbitrary. We use an equidistant partitioning along the  $z$ -axis; however, the subvolumes can also be identified using other strategies (e.g., segmentation). Each partition is independently registered, but all the remaining  $K - 1$  partitions are utilized for regularization of the registration. Consequently, the new distance measure can be described as a weighted linear combination of the simple distance measures of all  $K$  histograms. Thus, the deformation computed for partition  $j$  depends on all histogram partitions, as it utilizes their properties as well. This leads to the following distance measure:

$$D^j[\mathcal{H}(R), \mathcal{H}(C); \mathbf{u}_j] = \sum_{k=1}^K a_{j,k} \mathcal{D}[\mathcal{H}(R^k), \mathcal{H}(C^k); \mathbf{u}_j] \quad (4)$$

with

$$\sum_{k=1}^K a_{j,k} = 1 \quad (5)$$

where  $\mathcal{H}(R^k)$  and  $\mathcal{H}(C^k)$  are the reference histogram of subvolume  $k$  and the current histogram of subvolume  $k$ , respectively,  $\mathbf{u}_j$  is the deformation field of partition  $j$  corresponding to the transformation  $T_j$  and  $\mathcal{D}$  is a simple distance measure as mentioned above. Furthermore,  $j$  identifies the current subvolume and  $a_{j,k}$  is the influence of the force of partition  $k$  in the context of the standardization of subvolume  $j$ .

If all  $a_{j,k}$  are set to  $a_{j,k} = (1/K)$  the resulting deformation field of all partitions is the same. Setting  $a_{j,j} = 1$  for all  $j$  and  $a_{j,k} = 0, \forall j \neq k$  results in an unconstrained nonrigid registration of the joint histograms of all  $K$  blocks. The proposed approach can still lead to discontinuities at the junctions of the partitions. However, these can be significantly reduced, depending on the chosen weighting factors  $a_{j,k}$ . Furthermore, these discontinuities can be further reduced by interpolating the computed  $\mathbf{u}_j$  along the spatial positions of the corresponding intensities. Depending on the partitioning of the volume into blocks, different interpolation schemes have to be applied. As we decompose our volume along the  $z$ -axis only, we apply a cubic  $B$ -spline interpolation as follows. The number of bins of the single histograms of the  $n$  acquired images is denoted by  $l_1, \dots, l_n$ ; thus, the joint histograms have  $L = l_1 \times \dots \times l_n$  bins. Consequently, the transformation  $T_j$  corresponding to the deformation field  $\mathbf{u}_j$  is sampled at  $L$  positions. Thus,  $n \cdot L$  splines have to be computed using standard numerical methods [26]; one for each bin and for each dimension of the deformation field. Correcting an intensity  $i_m$  with  $z$ -coordinate  $z$  in an image  $C_m$  given an intensity vector  $\mathbf{i} = (i_1, \dots, i_m, \dots, i_n)^T$  results in

$$i_{m,\text{corr}} = i_m + S_{l,m}(z) \quad (6)$$

where  $m \in \{1, \dots, n\}$  and  $l \in \{1, \dots, L\}$  corresponds to the intensity vector  $\mathbf{i}$ . The relationship between the control points of the splines  $S_{l,m}$ , the  $z$ -coordinate of the volumes and the signal

intensity transformation is illustrated in Fig. 3. The approach is summarized in the following algorithm.

---

**Require:**  $\mathcal{H}(R^k), a_{j,k} \quad \forall k, j$

**Ensure:** Standard intensity scale

**for**  $k = 1$  to  $K$  **do**

compute current histogram  $\mathcal{H}(C^k)$

equalize  $\mathcal{H}(C^k)$

**end for**

**for**  $j = 1$  to  $K$  **do**

compute  $\mathbf{u}_j$ ; the previously computed histograms  $\mathcal{H}(R^k), \mathcal{H}(C^k)$  and  $a_{j,k}$  are required

**end for**

**for**  $l = 1$  to  $L$  **do**

**for**  $m = 1$  to  $n$  **do**

compute B-spline  $S_{l,m}$  using  $\mathbf{u}_j$

**end for**

**end for**

**for** all voxels **do**

transform the intensities using the interpolated transformations;  $S_{l,m}$  and the  $z$ -coordinate are required.

**end for**

---

A further advantage of the proposed intensity standardization method arises from the use of joint histograms. As only pdfs are utilized, the standardization results do not become worse if the region of interest in the two volumes differs slightly. For this reason just a coarse rigid registration of the reference with the current image data sets is required in the interpatient case to generate the  $K$  subvolumes for the proposed approach. However, if this is not done or cannot be guaranteed the joint histograms will not fit and hence the quality of the standardization will be worse. If the whole statistical information is used at once ( $K = 1$ ) then no alignment of the volume sets  $R$  and  $C$  has to be done. As a result of the flexibility concerning the region of interest, the method can deal with anatomical differences in interpatient scans or anatomical changes due to evolving lesions, for instance. However, if the source and target images are significantly different anatomically, or if there is a big variance in the localization of the regions of interest, the registration may fail. However, a possible solution for increasing the quality of the standardization results is to introduce additional reference histograms that cover the possible anatomical differences. Hence,

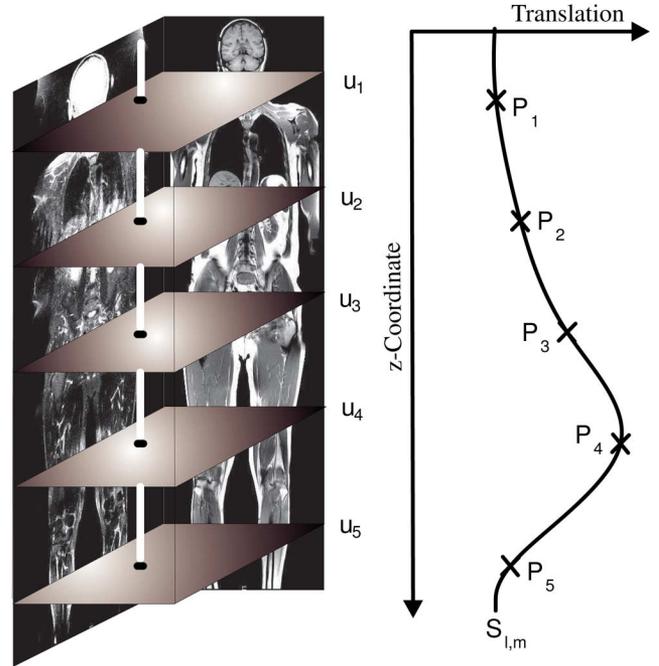


Fig. 3. Schematic illustration of the  $B$ -spline interpolation of the transformation vectors. The example uses  $K = 5$  partitions. The planes  $u_1$ - $u_5$  represent the magnitude of the deformation in a specific subvolume. The plot shows the relationship between the control points  $P_1$ - $P_5$  of the spline  $S_{l,m}$ , the  $z$ -coordinate within the volumes and the magnitude of the translation.

before the signal intensity standardization, the proper set of reference histograms has to be chosen. This can be achieved, for instance, by selecting the histograms with a minimal distance to the current histograms.

### III. EXPERIMENTS AND RESULTS

For evaluation of the proposed method various body regions, pulse sequences and different machines of the same brand were used. When possible the approach was evaluated using time-to-time studies (follow-up studies). However, in the case of whole body MRI, no voxel-wise evaluation was possible, as the deformation of the anatomical structures was too large even for follow-up studies, due to different positioning of the patients, nonrigid transformations of body regions, partial volume averaging effects, and pathologies.

#### A. Standardization of MRI Head Images

1) *Data Sets:* For the evaluation of the proposed intensity standardization method on MRI head data sets T1- and T2/FLAIR images were used. All data sets were acquired on a Siemens Symphony 1.5-T scanner. The T2-weighted FLAIR data sets had a resolution of  $408 \times 512 \times 19$ , an isotropic in-plane resolution of  $0.43 \text{ mm}^2$ ,  $7.2 \text{ mm}$  slice thickness, and  $\text{TE} = 143$  and  $\text{TR} = 9000 \text{ ms}$ . The T1-weighted images had a resolution of  $208 \times 256 \times 19$  with an isotropic in-plane resolution of  $0.86 \text{ mm}^2$ ,  $7.2 \text{ mm}$  slice thickness and  $\text{TE} = 14$  and  $\text{TR} = 510 \text{ ms}$ . In total 25 volumes from 11 different

TABLE I  
STANDARDIZATION RESULTS OF THE MRI HEAD IMAGES. ABSOLUTE MEAN AND STANDARD DEVIATION  
OF THE DIFFERENCE IMAGES WITH AND WITHOUT STANDARDIZATION

Volume	FLAIR				T1			
	original		standardized		original		standardized	
	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$
1a	<b>19.94</b>	62.41	27.51	62.27	90.98	57.82	<b>4.65</b>	55.01
1b	42.25	66.70	<b>6.30</b>	70.75	75.19	62.82	<b>3.17</b>	61.23
2	89.76	100.88	<b>3.64</b>	72.99	45.37	61.92	<b>9.36</b>	67.46
3	7.09	58.17	<b>1.02</b>	59.00	19.24	69.83	<b>1.64</b>	65.01
4a	80.11	79.58	<b>8.67</b>	58.95	53.39	74.52	<b>0.39</b>	77.93
4b	79.36	50.44	<b>2.14</b>	57.64	17.79	81.75	<b>10.18</b>	86.32
5	22.82	46.84	<b>0.32</b>	49.50	25.49	53.70	<b>9.74</b>	56.45
6	22.82	34.79	<b>13.39</b>	36.37	7.51	68.70	<b>1.81</b>	61.00
7	106.70	73.07	<b>7.70</b>	77.73	28.09	61.78	<b>4.50</b>	58.29
8	38.24	87.34	<b>30.33</b>	88.17	87.53	75.91	<b>12.27</b>	63.13
9	107.80	110.41	<b>12.48</b>	110.74	2.23	96.49	<b>0.90</b>	84.50
10	25.05	39.58	<b>1.84</b>	41.44	19.49	57.27	<b>5.56</b>	60.58
11a	<b>2.48</b>	56.81	3.02	55.87	<b>15.49</b>	52.84	16.53	51.74
11b	11.28	47.09	<b>2.72</b>	46.54	<b>8.47</b>	47.57	15.03	47.39
Means	46.84	36.41	8.65	9.17	35.45	29.02	6.84	5.12

patients were used where eight subjects had two scans and three subjects had three scans. All the images were of real patient data including evolving lesions.

2) *Evaluation Method:* For the experiments, a single T1-weighted volume and its corresponding FLAIR data set was chosen as reference for each patient. Then the follow ups were standardized to the intensities of the reference volumes. As the treated body region is small and the pdfs of brain images are not very complex, the volumes were not partitioned ( $K = 1$ ). By subtracting the two image tuples  $R$  and  $C$  intuitive quality measures can be computed.

In order to evaluate the intensity standardization the mean distance and the standard deviation between each patient's reference and current volumes were chosen as quality measures. The systematic error between the reference and the standardized images is expected to be smaller than that between the reference and the unprocessed images due to the signal intensity standardization. However, patient data with evolving structures from clinical routine was used and thus the anatomy of the brain changed. Consequently the evaluation method has the drawback that the difference between the volumes will never vanish. The lesions were removed by a segmentation step beforehand (just for evaluation, not for the signal intensity standardization step). Only those voxels that were classified as healthy brain tissue in both volumes were considered in computing the quality of the standardization. In order to be able to compute the difference between the volumes, a rigid registration using normalized mutual information [27] as distance measure had to be utilized, so that the reference and template volumes match each other.

For the signal intensity standardization, we used the following parametrization for the registration:  $\alpha = 0.001$  and 128 bins. As interpolation method, a bi-linear interpolation was utilized. The histograms were created by a partial volume technique. This means that the bins were filled relative to the distance to the intensity vector. This configuration was suitable for most of the experiments. In some cases the parameter  $\alpha$  had to be slightly adapted.

3) *Results:* The detailed results of the standardization are shown in Table I. For the FLAIR images the absolute mean

$\mu = (1/N_e) \sum_{j=1}^{N_e} |\mu_j|$  of all  $N_e$  experiments using the images without standardization was  $\mu_{\text{FLAIR}} = 46.84$  and the corresponding standard deviation was  $\sigma_{\text{FLAIR}} = 36.41$ . Using the proposed intensity standardization approach the mean was  $\mu'_{\text{FLAIR}} = 8.65$  with a standard deviation of  $\sigma'_{\text{FLAIR}} = 9.17$ . For the original T1-weighted images the absolute mean was  $\mu_{\text{T1}} = 35.45$  with a standard deviation of  $\sigma_{\text{T1}} = 29.02$ . Using the intensity standardization the mean was  $\mu'_{\text{T1}} = 6.84$  with a standard deviation of  $\sigma'_{\text{T1}} = 5.12$ . Although, in general, the method significantly decreases the difference, there are a few cases that this difference may increase. There are two reasons for this effect. First, the optimization of the registration might be stuck in a local minimum of the objective function. Second, if the structure of the histograms does not fit precisely, the resulting intensity mapping might be incorrect. However, this just happens if the distance between the histograms was small from the beginning; thus, the resulting differences are still very small. If the method proposed by Nyúl *et al.* [14] is employed the results are:  $\mu_{\text{FLAIR}}^{\text{Nyul}} = 10.73$ ,  $\sigma_{\text{FLAIR}}^{\text{Nyul}} = 8.30$ ,  $\mu_{\text{T1}}^{\text{Nyul}} = 11.27$ , and  $\sigma_{\text{T1}}^{\text{Nyul}} = 7.57$ . As the used images were bimodal, we utilized the second mode as histogram landmark. As mentioned before, a slight difference between the volume is expected as the anatomy of the brain changed [28]. Furthermore, we applied no bias correction method to the head images, because the intensity inhomogeneities of the acquired images were rather small. Fig. 4 shows the marginals of the joint histograms of the T1-weighted head images before and after standardization. The marginals of the joint histograms correspond to the histograms of the single volumes. Here six volumes were randomly selected and standardized to a single reference. The thick line represents the reference histogram in both plots.

In Fig. 5 the effect of the signal standardization on FLAIR images is shown. The first row shows the slices from the reference volume tuple. On the left side the original image slices are shown. On the right side the processed images are shown. All images are displayed with the same transfer function; thus, comparable tissue classes should have similar gray values in the corrected images. Furthermore, all images are from different acquisitions, whereas the images in the second and third row show

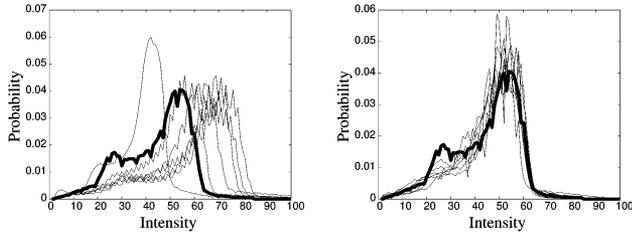


Fig. 4. Left: the marginals of the T1-weighted head images before standardization. Right: the marginals of the same images after signal intensity standardization. The thick line in both plots represents the reference histogram.

a follow-up study of the same patient. It can be clearly seen that the intensities are adapted to the intensities of the reference image.

### B. Parametric Perturbation of Intensities

1) *Data Sets*: In this experiment parametric perturbations were applied on a pair of T1- and T2/FLAIR weighted images. They were acquired on a Siemens Symphony 1.5-T scanner. The T2-weighted FLAIR data set had a resolution of  $408 \times 512 \times 19$ , an isotropic in-plane resolution of  $0.43 \text{ mm}^2$ , 7.2 mm slice thickness, and  $TE = 143$  and  $TR = 9000$  ms. The T1-weighted volume had a resolution of  $208 \times 256 \times 19$  with an isotropic in-plane resolution of  $0.86 \text{ mm}^2$ , 7.2 mm slice thickness, and  $TE = 14$  and  $TR = 510$  ms.

2) *Evaluation Method*: In all experiments in this section the used image pair was utilized as reference. Then the T1- and the T2/FLAIR- weighted images were perturbed by a parametric intensity mapping; thus, both data sets are transformed independently. With that setting, the joint standardization of all volumes has no advantage over the separate standardization. In this case the difference of the proposed method to Nyúl's approach [14] is reduced to the nonlinearity of our approach and that it is not necessary for us to locate any landmarks in the histograms. In order to evaluate the standardization methods, the perturbed images were mapped back to the previously chosen reference data sets. As a measure for the quality of the standardization, the mean distance between the reference and the intensity standardized images was used.

The first intensity mapping we evaluated was

$$i_{\text{new}} = i_{\text{ref}} \cdot \left( \frac{(\kappa - 1)}{i_{99.8}} \cdot i_{\text{ref}} + 1 \right) \quad (7)$$

with  $i_{\text{new}}$  being the perturbed intensity,  $i_{\text{ref}}$  the intensity in the reference image,  $i_{99.8}$  the 99.8% percentile of all intensities in the reference image and  $\kappa$  a factor to control the strength of the perturbation. This means that if  $\kappa = 1$ , the intensities do not change at all. If  $\kappa$  is smaller than one then  $i_{\text{new}} < i_{\text{ref}}$ ; if  $\kappa > 1$  then  $i_{\text{new}} > i_{\text{ref}}$ . Furthermore, the higher the value of the initial intensity the larger is the perturbation. We evaluated the signal intensity standardization for  $\kappa \in [0.3; 2.0]$  for Nyúl's method [14] and the proposed approach. As the used images were bimodal, in Nyúl's method the second mode was utilized as landmark.

Second, we evaluated

$$i_{\text{new}} = i_{\text{ref}} \cdot \left( 1 + c \cdot \sin \left( f \cdot \frac{i_{\text{ref}}}{i_{99.8}} \right) \right) \quad (8)$$

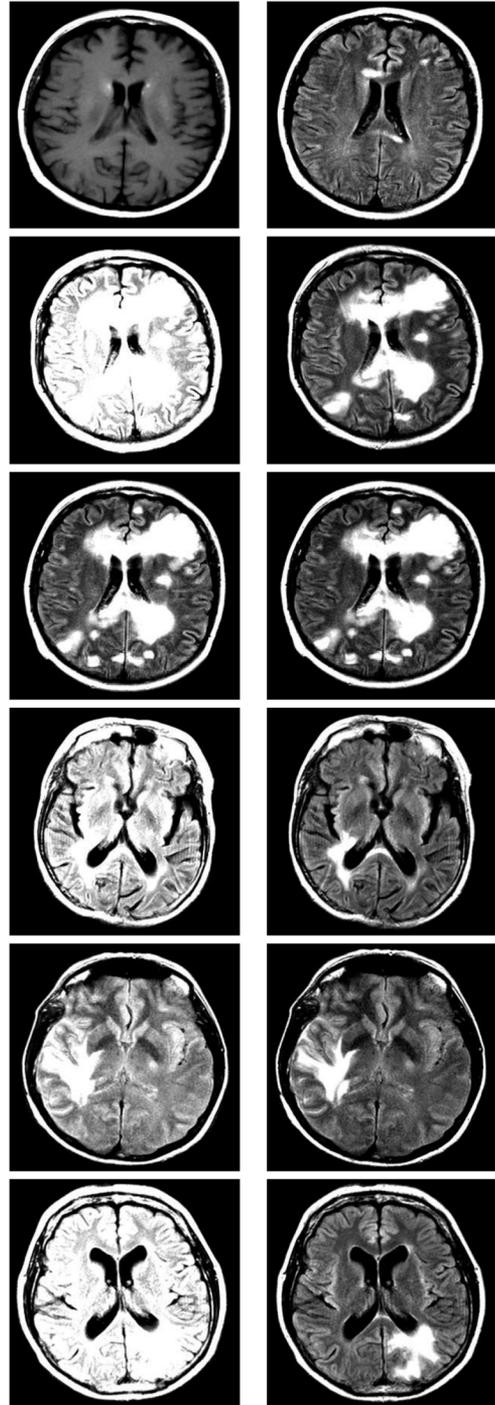


Fig. 5. First row: On the left side a slice of the reference T1-weighted volume is shown, on the right the FLAIR slice, respectively. Left column: the original FLAIR slices. Right column: the corrected FLAIR slices. All images (except the T1w slice) are displayed with the same transfer function: center 200 and width 200.

as an intensity mapping. Here the parameter  $c$  describes the amplitude of the distortion and  $f$  its frequency. We evaluated both algorithms with a frequency of  $f = 1$  and  $f = 4$ . For  $f = 1$  we evaluated the amplitude  $c \in [0; 0.5]$ ; for  $f = 4$  we chose  $c \in [0; 0.35]$ . An amplitude of  $c = 0$  yields no intensity distortion.

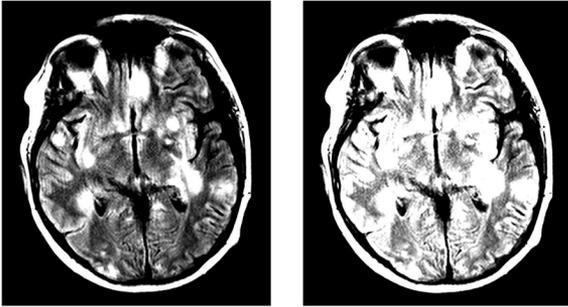


Fig. 6. The figure shows the effect of the artificial quadratic perturbation on a T2/FLAIR weighted image. The left image shows a slice of the original T2/FLAIR weighted volume. The right image shows the result of the quadratic perturbation. Both images are displayed with the same transfer function: center 300 and width 200.

3) *Results:* The effect of the quadratic perturbation on the T2/FLAIR weighted image is shown in Fig. 6. The results are illustrated in the three plots in Fig. 7. It can be seen that both approaches decrease the mean distance to the reference images. However, in general, the proposed algorithm yields better results than Nyúl's method [14]. Due to the piecewise linear intensity mapping between the chosen landmarks, the results of Nyúl's method get worse if the perturbations become more nonlinear. As the proposed signal intensity standardization method does not rely on the position of any landmarks, it is much more flexible and can easily adapt the correction to the nonlinearity. If there are small intensity distortions only, Nyúl's method slightly outperforms the proposed method. The reason for this is that if there are hardly any intensity distortions it is easy to exactly determine the same landmark positions in the histograms. As the proposed approach does not use any assumptions about the shape of the histograms and we are using a global regularization in the nonrigid registration, small changes in the perturbed histograms yield global changes in the intensity mapping.

### C. Artificial Perturbation of the Intensities of a Certain Tissue Class in Synthetic Images

1) *Data Sets:* For this experiment we used synthetic data sets from the McConnell Brain Imaging Centre (BIC) of the Montreal Neurological Institute, McGill University [29]. The simulated data sets had a resolution of  $181 \times 217 \times 181$  and a slice thickness of 1 mm. We chose the T1w and T2w-images with a noise level of 3% and no signal intensity inhomogeneities. In order to evaluate the proposed algorithm, the intensities were artificially perturbed by a warping of the joint histograms using thin plate splines. Slices from the synthetic images are illustrated in Fig. 8. Fig. 9(a) shows the joint histogram of the original T2w and T1w image. In Fig. 9(b) the joint histogram with the maximal intensity deformation is shown. For the experiments the intensity deformation is varied between no and maximal distortion. The major part of the perturbed intensities belongs to muscle tissue.

2) *Evaluation Method:* For the evaluation the mean distance of the reference image to the standardized intensity transformed

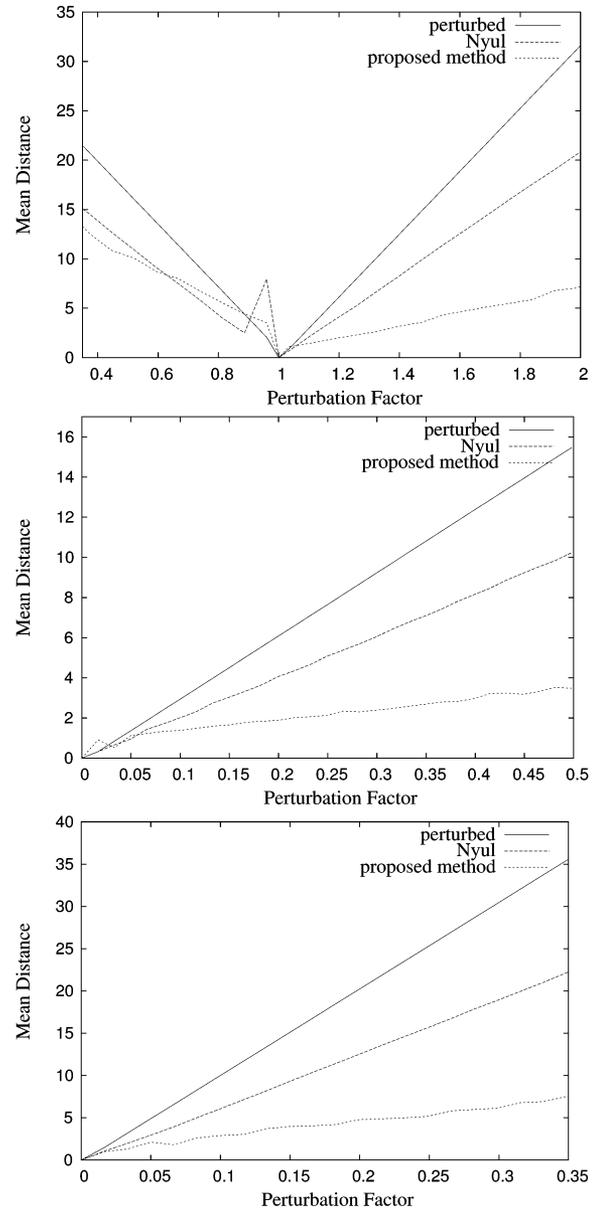


Fig. 7. The first plot shows the results of the standardization for the quadratic intensity perturbation of the images. In the second and third block the results for the sinusoidal perturbation are shown ( $f = 1$  and  $f = 4$ ). All mean distances are with respect to the FLAIR images.

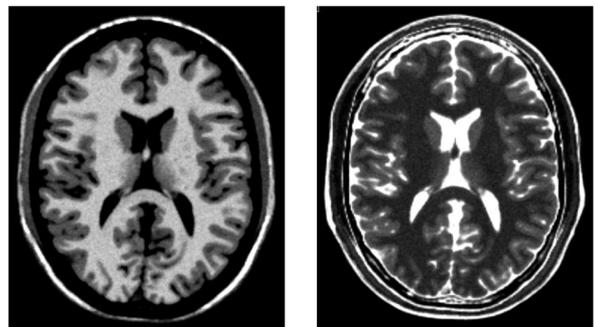


Fig. 8. The left image shows the used synthetic T1w image. On the right side a corresponding T2w slice is shown. The images are taken from the BrainWeb database [29].

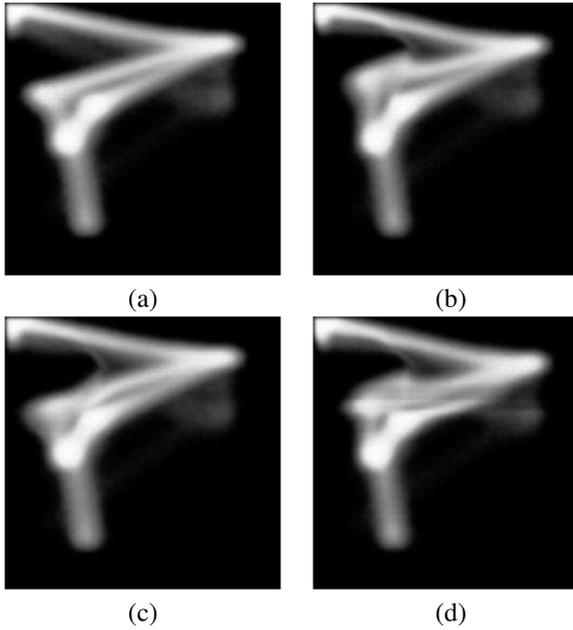


Fig. 9. Image (a) shows the original joint histogram of the synthetic T2w and T1w image. The histogram is equalized for presentation. The  $x$ -direction represents the T2w image and the  $y$ -direction the T1w image. (b) The joint histogram of the intensity transformed images is shown. (c) The correction result using the proposed method. (d) The correction result of Nyúl's method [14].

images was chosen. In order to create a set of transformed images the intensity deformation was scaled between no deformation and maximal deformation. The measures of the proposed method were compared to Nyúl's method [14]. Here, we tried different configurations and found, that using the percentile set  $P \in \{10, 20, \dots, 90\}$  where each element  $p_i \in P$  represents a  $p_i$ th percentile in the histogram yields the best results.

3) *Results:* This experiment demonstrates the main advantages of the signal intensity standardization using all acquired images jointly compared to the normalization of the images separately. It can be seen that if just the intensities of a certain tissue class are deformed, these changes cannot be corrected by aligning the 1-D histograms to a given reference. The reason for this is, that the translation of an intensity in the projection results in the translation of a complete row in the joint histogram. Thus, not only the disturbed tissue class is modified. The result of the experiment regarding the T1w image is shown in the plot of Fig. 10. As illustrated in Fig. 9, the plot shows that the proposed method outperforms Nyúl's method [14] given the artificial intensity distortion. Fig. 9(d) shows the standardization result of the images with maximal intensity deformation. The intensities of the transformed intensity class were not corrected properly. Consequently, the mean distances after standardization are worse than before. Furthermore, streak artifacts can be seen at the borders of the percentile positions. Using a continuous transformation instead of the piece-wise linear transformations could reduce these histogram artifacts. In Fig. 9(c) the normalization result of the proposed method is shown. The intensity deformation was reduced significantly. However, due to

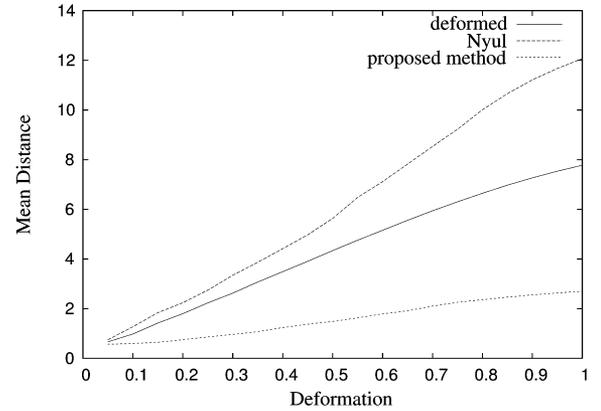


Fig. 10. The plot shows the mean distance of the artificial intensity perturbed T1w image, the corrected version using Nyúl's method [14] and the standardized images using the proposed method with respect to the degree of the intensity distortion.

the regularization and the loss of some intensities during the distortion step, no perfect standardization was reached.

#### D. Standardization of Head Data Sets Including Artificially Evolving Pathologies

1) *Data Sets:* For the evaluation of the influence of pathologies on the proposed signal intensity standardization T2/FLAIR weighted images and T1 weighted images with contrast bolus were used. All data sets were acquired on a Siemens Symphony 1.5 T scanner. The T2-weighted FLAIR data sets had a resolution of  $408 \times 512 \times 19$ , an isotropic in-plane resolution of  $0.43 \text{ mm}^2$ ,  $7.2 \text{ mm}$  slice thickness, and  $\text{TE} = 143$  and  $\text{TR} = 9000 \text{ ms}$ . The T1-weighted images had a resolution of  $208 \times 256 \times 19$  with an isotropic in-plane resolution of  $0.86 \text{ mm}^2$ ,  $7.2 \text{ mm}$  slice thickness,  $\text{TE} = 14$  and  $\text{TR} = 510 \text{ ms}$ , and  $14 \text{ ml}$  of *Magnevist* contrast bolus.

2) *Evaluation Method:* Two evolving pathologies were simulated in this experiment. First, the ventricles were increased and second, a lesion within the data set was artificially enlarged. In order to simulate the growth of the pathologies, a landmark based nonrigid registration approach was employed. For this purpose a set of landmarks  $L = \{l_1, \dots, l_n\}$ , where  $l_i \in \mathbb{R}^3$ ,  $i \in \{1, \dots, n\}$ , on the border of the structure (ventricles/lesion) were selected in the treated volume pair  $V$ . These  $n$  landmarks were manually assigned to new positions  $L' = \{l'_1, \dots, l'_n\}$ . Additionally, for each landmark  $l_i$  a fixed pair of landmarks at a distance  $\delta$  outside the structure was introduced to keep the transformation as local as possible. Finally, the eight corners of the volume were used as fixed landmarks. Thus, in total  $2n + 8$  landmarks were utilized. The registration approach makes use of thin plate splines [30]. After the alignment the landmarks are matched exactly. The values between the landmarks are interpolated smoothly. The result of the registration is a mapping  $\Gamma : \mathbb{R}^3 \mapsto \mathbb{R}^3$  that describes the transformation of a voxel. By multiplying  $\Gamma$  by a constant factor  $g_a \in [0, 1]$ , the resulting mapping  $\Gamma_a$  generates a new volume pair  $V_a$ , where the size of the treated structure lies between the size in the original volumes and the size in the volumes mapped by  $\Gamma$ . This is illustrated in Fig. 11.

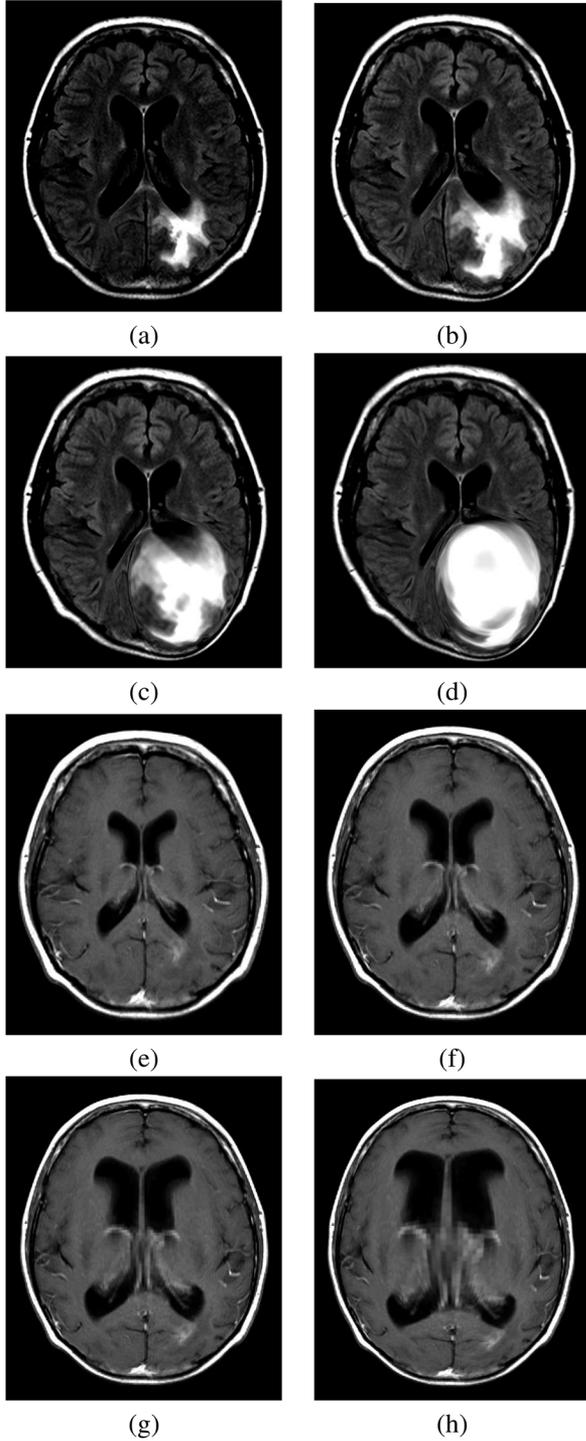


Fig. 11. In the images (a)–(d) the artificial growth of an lesion is shown. (a) Shows the original FLAIR slice, (d) the maximal deformation, (b) 33% of the maximal deformation, and (c) 66%, respectively. (e)–(h) Show the artificial growth of the first two ventricles: (e) original slice, (f) 33%, (g) 66%, and (h) maximal deformation.

Let  $\mathbf{V}$  be the original image tuple,  $\mathbf{V}_a$  the volumes with the artificially enlarged pathologies and  $\mathbf{R}$  a reference volume tuple defining the standard intensities. The computed mapping of the intensity space of a current volume tuple  $\mathbf{V}_a$  to the reference tuple  $\mathbf{R}$  is denoted by  $\gamma_a : \mathbf{V}_a \mapsto \mathbf{R}$ . In order to be able to

compare the result of the signal intensity standardization, the resulting intensity transformation  $\gamma_a$  is applied to the original volume pair  $\mathbf{V}$  yielding  $\mathbf{V}_{\gamma_a} = \gamma_a(\mathbf{V})$ . If pathologies have no influence on the signal standardization, all corresponding volume pairs  $\mathbf{V}_{\gamma_a}$  should have the same intensities for all voxels. Let  $\bar{\mathbf{V}}$  be the mean of all signal standardized volumes  $\mathbf{V}_{\gamma_a}$ . Thus, the pathology dependency can be measured by the mean distance  $\mu_p$  and its standard deviation of an observed volume pair  $\mathbf{V}_{\gamma_a}$  to the mean volume pair  $\bar{\mathbf{V}}$ .

3) *Results:* In the experiment using the artificially enlarged lesion, the volume of the lesion varied between approximately  $v_0 = 38 \text{ cm}^3$  for  $g_0 = 0$  to about  $v_1 = 136 \text{ cm}^3$  for  $g_1 = 1$ . This corresponded to a lesion to head ratio of about  $r_0 = 0.02$  to  $r_1 = 0.08$ . The results are illustrated in the upper plots of Fig. 12. The mean absolute difference of all transformed FLAIR images yielded  $\mu_l^{\text{FLAIR}} = 0.26$  with a standard deviation of  $\sigma_l^{\text{FLAIR}} = 0.13$ ; the mean absolute difference of the T1w images was  $\mu_l^{\text{T1}} = 0.54$  with a standard deviation of  $\sigma_l^{\text{T1}} = 0.37$ .

In the experiment using the artificially increased ventricles, the volume of the ventricles varied between approximately  $v_0 = 85 \text{ cm}^3$  for  $g_0 = 0$  to about  $v_1 = 316 \text{ cm}^3$  for  $g_1 = 1$ . This corresponded to a ventricle to head ratio of about  $r_0 = 0.04$  to  $r_1 = 0.14$ . The results are illustrated in the lower plots of Fig. 12. The mean absolute difference of all transformed FLAIR images yielded  $\mu_v^{\text{FLAIR}} = 0.40$  with a standard deviation of  $\sigma_v^{\text{FLAIR}} = 0.23$ ; the mean absolute difference of the T1w images was  $\mu_v^{\text{T1}} = 0.64$  with a standard deviation of  $\sigma_v^{\text{T1}} = 0.49$ .

The input images had gray values in the range between zero and approximately 1200. Thus, the maximal measured intensity deviations show that the proposed algorithm is robust against severe pathological changes within the volumes.

### E. Standardization of Whole Body Data Sets

1) *Data Sets:* All data sets were acquired on a Siemens Avanto 1.5 T whole body MRI scanner. The TIRM images had a resolution of  $512 \times 512 \times 30$  (each block) with an isotropic in-plane resolution of  $0.98 \text{ mm}^2$  and 5.5 mm slice thickness and  $\text{TE} = 83$ ,  $\text{TR} = 1660$  ms and the FL2D images had a resolution of  $512 \times 410 \times 30$  (each block) with an isotropic in-plane resolution of  $0.98 \text{ mm}^2$  and 5.5 mm slice thickness and  $\text{TE} = 4.7$  and  $\text{TR} = 291$  ms. The size of the composed whole body images was  $542 \times 1746 \times 20$  for both protocols. Only the composed volumes were used for the experiments. All images were acquired in clinical routine, thus, including pathologies. In total nine whole body MRI data sets were used for evaluation. Two of these were from the same patient, acquired with a time delay of six months. In order to reduce the effects of intensity inhomogeneities homomorphic unsharp masking was employed.

2) *Evaluation Method:* Due to the anatomical differences and the large deformations within follow-up studies, no voxel-wise evaluation method was used. As quality measure the relative distance between the reference and current joint histogram before and after the standardization was chosen

$$r = \frac{d_{\text{After}}}{d_{\text{Before}}}. \quad (9)$$

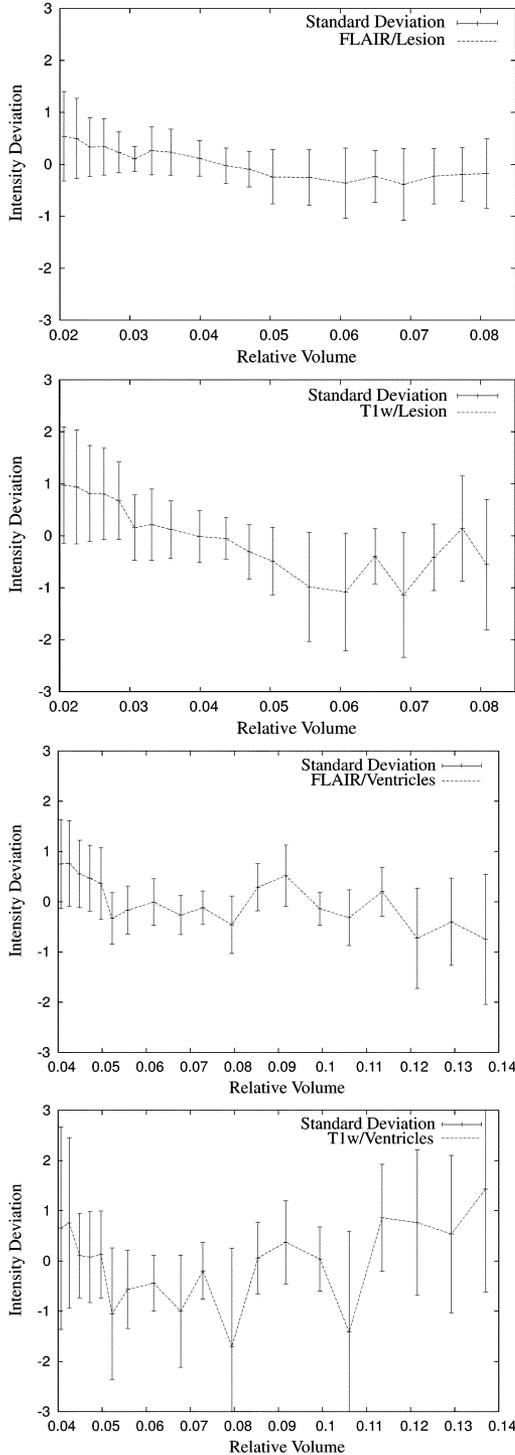


Fig. 12. The plots show the difference of the transformed volumes with different sizes of the structures to the “mean” transformed volume. The first two ones show the results of the artificially enlarged lesion. The latter two illustrate the results of an increased ventricle size. The intensity range of the images was  $[0, 1200]$ .

Consequently, if  $r < 1$  the distance between the images was reduced and if  $r \geq 1$  the standardization failed. As distance measure between the joint histograms the Jeffrey divergence

$$d(\mathcal{H}(R), \mathcal{H}(C)) = \sum_{i=1}^L \left( h_i^R \log \frac{h_i^R}{m_i} + h_i^C \log \frac{h_i^C}{m_i} \right) \quad (10)$$

where

$$m_i = \frac{h_i^R + h_i^C}{2} \quad (11)$$

was used, with  $h_i^R \in \mathcal{H}(R)$  and  $h_i^C \in \mathcal{H}(C)$ . Similar to the Kullback–Leibler–Divergence, the Jeffrey Divergence measures how efficient, on average, it would be to code one histogram using the other as code book [31]. Our assumption is that if the histograms of two images are as close as possible, a tissue class in both images covers the same intensity range in both volumes. Hence the distance between the joint histograms is a good measure for the quality of the results of the signal intensity standardization.

In the first experiment we evaluated the standardization quality related to the influence of neighboring histograms on the regularization. We used a uniform partitioning along the  $z$ -axis. Only direct neighbors to the current histogram are allowed to affect the registration. For this reason we introduced an influence parameter  $\beta \in [0; 1]$  with

$$a_{j,k} = \begin{cases} \beta, & \text{if } j = k \\ (1 - \beta)/2, & \text{if } k = j - 1 \vee k = j + 1 \\ 0, & \text{else.} \end{cases} \quad (12)$$

where  $j$  is the current sub-volume to be registered. Consequently,  $\beta = 1.0$  does not use statistical information of neighboring histograms and  $\beta = 0.0$  uses neighboring pdfs only for the standardization. The results showed that the parameter  $\beta$  has to be chosen between  $\beta \in [0.5; 0.7]$ , depending on the number of partitions. If fewer partitions are used, then the influence has to be reduced. Note that, because the first and the last sub-volume are mirrored, their influence on the registration result is not just  $\beta$  but  $\beta + (1 - \beta)/2$ . The standardization results using different numbers of partitions and a varying parameter  $\beta$  are illustrated in Fig. 13 (upper plot).

The second experiment deals with the number of partitions to be used. In order to evaluate this parameter, the best signal intensity standardization result for each number of partitions (the parameter  $\beta$  may vary) is selected. This shows that approximately ten subvolumes have to be chosen. If more partitions are used the pdfs cannot be estimated in a reliable manner anymore. If fewer partitions are used, the standardization is affected by intensity inhomogeneities and small structures are neglected in the registration. The standardization results are illustrated in Fig. 13 (lower plot).

3) *Results:* The results of both experiments show that it is best to use approximately  $K = 10$  partitions. They also show, that the statistical information of neighboring partitions has a significant impact on the quality of the standardization of a sub volume. For the given number of partitions a parameter  $\beta \approx 0.6$  provided the best results. Using the Jeffrey divergence it was possible to achieve a relative distance of  $r = 0.567$ . This means that the distance between the joint histograms has been reduced by about 45%. However, it is not possible to achieve significantly better results as the anatomical differences between the data sets are too large. Consequently, even if the signal intensities of all tissue classes are transformed in an ideal manner, there still have to be differences between the histograms. The

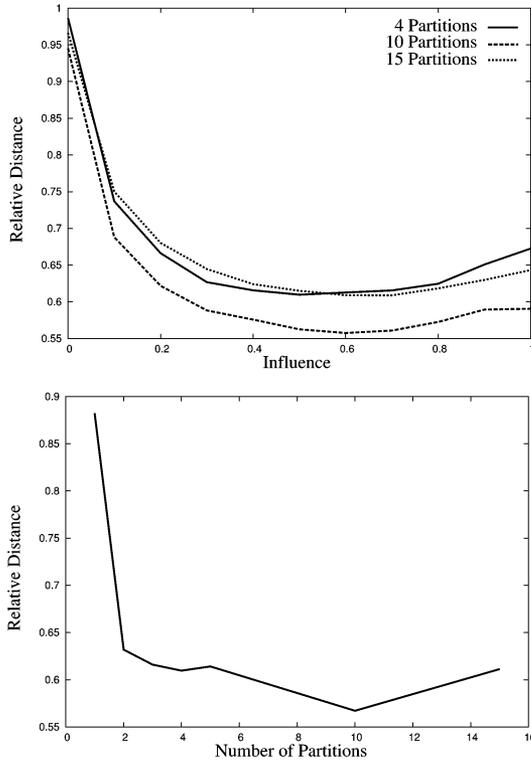


Fig. 13. The top plot shows the results of the intensity standardization with respect to the influence parameter  $\beta$ . The bottom plot illustrates the effect of different numbers of partitions on the standardization.

results of the whole body signal intensity standardization are illustrated in Fig. 14.

#### IV. SUMMARY AND CONCLUSION

In this paper, we present a new method for MRI signal intensity standardization of MRI scans. In contrast to most of the previously published methods, the proposed approach is independent from any prior knowledge about the structure of the data sets and it relies only on the image properties. Hence, no prior registration or segmentation of the data sets are necessary. Furthermore, the method is independent of the application, body region and pulse sequence, if reference joint histograms are available. The most important improvement is that the proposed method utilizes all acquired images jointly and does not do the standardization of individual images separately. Nevertheless, the proposed signal intensity standardization approach can also be used for single-feature normalization ( $n = 1$ ).

The presented standardization method is a reliable way to adjust pdfs of multiple series of MRI scans to a pulse sequence dependent, standard signal intensity scale. However, if the image properties are too distinct, the obtained results may not be satisfying. This is the case, for instance, for volumes disturbed by a strong bias field. This yields blurred histograms and thus no reliable registration is possible anymore. In general, however, the method produces satisfactory results. Thus, postprocessing methods, like automatic segmentation, the presentation of the data-sets, or the MR based attenuation correction for PET images can be done more easily.

Lastly, the proposed method is the only approach for whole body intensity standardization known in literature. Whole body MRI will gain more importance in the future, as it allows fast imaging of the whole body without the need of repositioning the patient or the coils. Moreover, it allows screening of patients with generalizing diseases, like malign lymphoma. However, due to the large amount of acquired information, (semi) automatic preprocessing and filtering of the data sets is necessary. A reliable processing of the acquired images needs a general signal intensity scale, like the Hounsfield units in CT imaging. For this reason signal intensity standardization of whole body MRI will gain more importance in the future.

#### APPENDIX NONRIGID REGISTRATION

Image registration can be summarized as the problem of finding a deformation between a reference image  $A$  and a current image  $B$  so that the deformed template image  $B_\varphi$  is similar with respect to a certain distance measure  $\mathcal{D}$ . In the approach presented in this article the reference image  $A$  corresponds to the joint histogram  $\mathcal{H}(R)$  representing an approximation of the multidimensional density function of  $R$  and, respectively,  $B = \mathcal{H}(C)$  represents the approximation of the density of  $C$ .

The used distance measure depends on the application. The most common ones are the sum of squared differences (SSD) for mono-modal applications, the normalized cross correlation (NCC), and mutual information (MI) for multimodal problems [32]. Because the function values of the joint histograms have identical meaning, the usage of SSD is sufficient in this paper. However, the minimization of the proposed distance measures yields an ill-posed optimization problem. Further regularization terms have to be added to smooth the objective function. These so-called smoothers restrict the deformation of the current image, in general. Mostly either elastic, fluid, curvature, or diffusion approaches are utilized. Here, the deformation is not expected to change very fast, thus we chose a curvature based regularizer.

The minimization problem to be solved can be formulated as in [22]

$$\mathcal{J}[A, B; \mathbf{u}] = \mathcal{D}[A, B; \mathbf{u}] + \alpha \mathcal{S}[\mathbf{u}] \quad (13)$$

where the function  $\mathbf{u} : \mathbb{R}^n \mapsto \mathbb{R}^n$  corresponds to the deformation field and  $n$  is the dimensionality of the images (in this article joint histograms). Furthermore,  $\mathcal{D}$  is the distance measure and  $\mathcal{S}$  represents the smoother. The factor  $\alpha$  defines the influence of the regularizer on the objective function. The deformed density function  $B_\varphi$  can be computed as  $B_\varphi(\mathbf{x}) = B \circ \varphi(\mathbf{x})$  with

$$\varphi(\mathbf{x}) = \mathbf{x} - \mathbf{u}(\mathbf{x}) \quad (14)$$

where  $\mathbf{x}$  defines a position within the images. As mentioned before, we use a similarity measure based on SSD to compute the distance between the joint histograms. In our context this can be formulated as

$$\mathcal{D}^{\text{SSD}}[A, B; \mathbf{u}] = \frac{1}{2} \int_{\Omega} (B_\varphi(\mathbf{x}) - A(\mathbf{x}))^2 d\mathbf{x}. \quad (15)$$

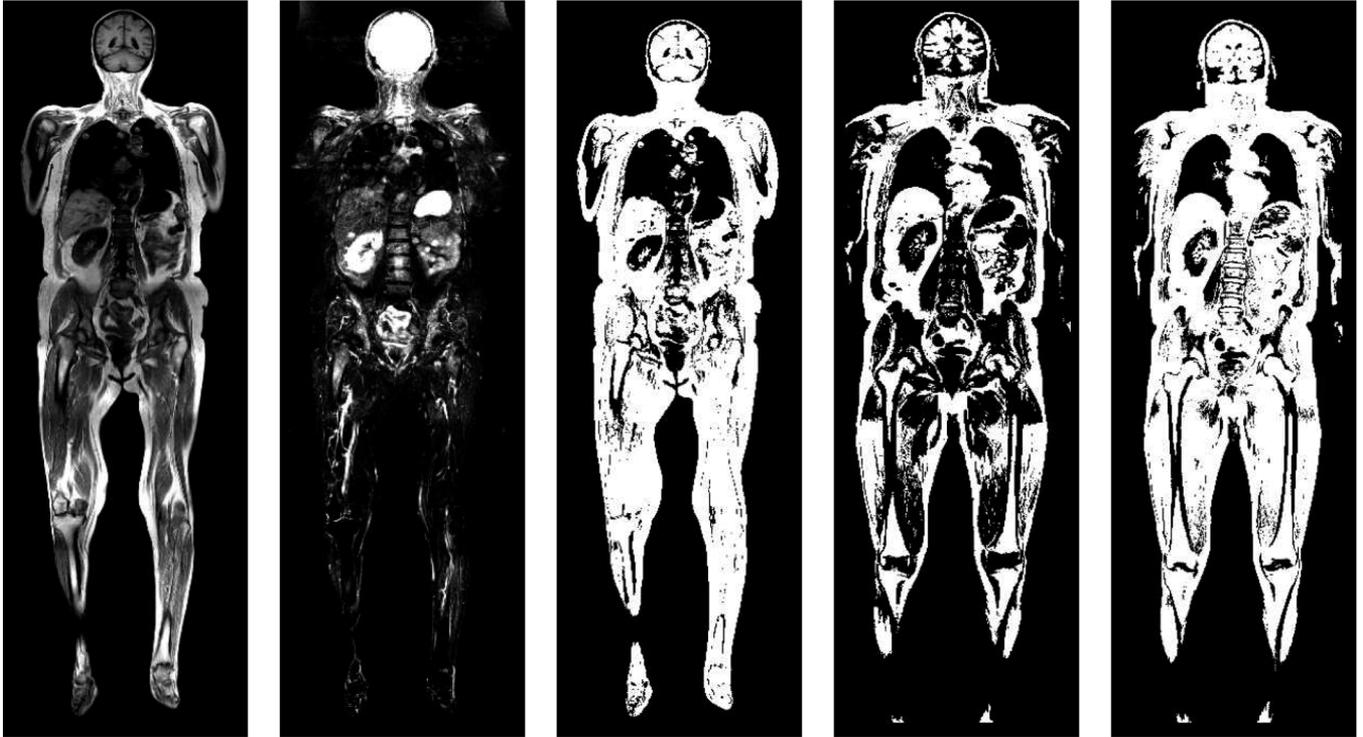


Fig. 14. From left to right: slice from the reference T1w volume; slice from the reference TIRM weighted data set. A threshold of 400 was applied to the third, fourth, and fifth image to better visualize the differences. The third image shows a reference T1w slice, the fourth the original T1w slice of a different patient, and the fifth image the standardized T1w slice.

It calculates the distance between the functions related to their function values at a position  $\mathbf{x}$ , with  $\Omega = [0, 1]^n$  representing the image domain.

The curvature based regularizer used in this paper can be formulated as

$$\mathcal{S}^{\text{curv}}[\mathbf{u}] := \frac{1}{2} \sum_{l=1}^n \int_{\Omega} (\Delta u_l)^2 d\mathbf{x} \quad (16)$$

with  $\Delta$  being the Laplacian operator. In order to find the minimum of the objective function  $\mathcal{J}$ , a variational problem of first order has to be solved. Therefore, the Gâteaux derivative has to be applied to  $\mathcal{J}$ . Thus, the variational gradient for the proposed distance measure yields

$$\begin{aligned} f^{\text{SSD}}(\mathbf{x}, \mathbf{u}(\mathbf{x})) &= d\mathcal{D}^{\text{SSD}}[A, B; \mathbf{u}] \\ &= (B_{\varphi}(\mathbf{x}) - A(\mathbf{x})) \nabla B_{\varphi}(\mathbf{x}) \end{aligned} \quad (17)$$

where the operator  $d\mathcal{D}$  is the Gâteaux derivative.

The derivative  $d\mathcal{S}^{\text{curv}}$  with respect to Neumann boundary conditions results in

$$\mathcal{A}^{\text{curv}}[\mathbf{u}] = d\mathcal{S}^{\text{curv}}[\mathbf{u}] = \Delta^2 \mathbf{u}. \quad (18)$$

Using the introduced similarity measures and the proposed smoother the solution to the variational optimization problem can be found by solving the Euler Lagrange equation

$$\mathcal{A}^{\text{curv}}[\mathbf{u}] - f^{\text{SSD}}(\mathbf{x}, \mathbf{u}(\mathbf{x})) = 0, \quad \text{for all } \mathbf{x} \in \Omega. \quad (19)$$

#### ACKNOWLEDGMENT

The authors would like to thank the following people for their support: F. Wacker and B. Frericks, Department of Radiology and Nuclear Medicine, Charité Berlin, CBF, for the whole body MRI data sets, R. Janka, Department of Diagnostic Radiology, University of Erlangen, for the MRI head images, L. Nyúl, University of Szeged, for his assistance, E. Angelopoulou, Chair of Pattern Recognition, University of Erlangen, for proofreading, and the company HipGraphics for providing the volume renderer InSpace3D.

#### REFERENCES

- [1] Z. Hou, "A review on MR image intensity inhomogeneity correction," *Int. J. Biomed. Imag.*, vol. 2006, p. ID 49515, 2006.
- [2] U. Vovk, F. Pernus, and B. Likar, "A review of methods for correction of intensity inhomogeneity in MRI," *IEEE Trans. Med. Imag.*, vol. 26, no. 3, pp. 405–421, Mar. 2007.
- [3] J. B. Arnold, J.-S. Liow, K. A. Shaper, J. J. Stern, J. G. Sled, D. W. Shattuck, A. J. Worth, M. S. Cohen, R. M. Leahy, J. C. Mazziotta, and D. D. Rottenberg, "Qualitative and quantitative evaluation of six algorithms for correcting intensity nonuniformity effects," *NeuroImage*, vol. 13, no. 5, pp. 931–943, May 2001.
- [4] L. Axel, J. Constantini, and J. Listerud, "Intensity correction in surface coil MR imaging," *Amer. J. Roentgenol.*, vol. 148, no. 2, pp. 418–420, Feb. 1987.
- [5] 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.
- [6] 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.
- [7] 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.

- [8] W. M. Wells III, W. E. L. Grimson, R. Kikinis, and F. A. Jolesz, "Adaptive segmentation of MRI data," *IEEE Trans. Med. Imag.*, vol. 15, no. 4, pp. 429–442, Aug. 1996.
- [9] B. Likar, M. A. Viergever, and F. Petruš, "Retrospective correction of MR intensity inhomogeneity by information minimization," *IEEE Trans. Med. Imag.*, vol. 20, no. 12, pp. 1398–1410, Dec. 2001.
- [10] O. Salvado, C. Hillenbrand, S. Zhang, and D. L. Wilson, "Method to correct intensity inhomogeneity in MR images for atherosclerosis characterization," *IEEE Trans. Med. Imag.*, vol. 25, no. 5, pp. 539–552, May 2006.
- [11] 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.
- [12] H. Zaidi, "Is MR-guided attenuation correction a viable option for dual-modality PET/MR imaging?," *Radiology*, vol. 244, no. 3, pp. 639–642, Sep. 2007.
- [13] A. Madabhushi and J. K. Udupa, "Interplay between intensity standardization and inhomogeneity correction in MR image processing," *IEEE Trans. Med. Imag.*, vol. 24, no. 5, pp. 561–576, May 2005.
- [14] L. G. Nyúl, J. K. Udupa, and X. Zhang, "New variants of a method of MRI scale standardization," *IEEE Trans. Med. Imag.*, vol. 19, no. 2, pp. 143–150, Feb. 2000.
- [15] Y. Ge, J. K. Udupa, L. G. Nyúl, L. Wei, and R. I. Grossman, "Numerical tissue characterization in MS via standardization of the MR image intensity scale," *J. Magn. Reson. Imag.*, vol. 12, no. 5, pp. 715–721, Oct. 2000.
- [16] P. Hellier, "Consistent intensity correction of MR images," in *Int. Conf. Image Process. (ICIP 2003)*, Sept. 2003, vol. 1, pp. 1109–12.
- [17] 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.
- [18] N. Weisenfeld and S. Warfield, "Normalization of joint image-intensity statistics in MRI using the Kullback-Leibler divergence," in *IEEE Int. Symp. Biomed. Imag.*, Arlington, VA, Apr. 2004, pp. 101–104.
- [19] M. Schmidt, A method for standardizing MR intensities between slices and volumes Univ. Alberta, Edmonton, AB, Tech. Rep. TR05-14, 2005.
- [20] J. B. A. Maintz and M. A. Viergever, "A survey of medical image registration," *Med. Image Anal.*, vol. 2, no. 1, pp. 1–36, Mar. 1998.
- [21] D. L. G. Hill, P. G. Batchelor, M. Holden, and D. J. Hawkes, "Medical image registration," *Phys. Med. Biol.*, vol. 46, no. 3, pp. R1–R45, Mar. 2001.
- [22] J. Modersitzki, *Numerical Methods for Image Registration*. New York: Oxford Univ. Press, 2004.
- [23] E. Haber and J. Modersitzki, "Numerical methods for volume preserving image registration," *Inverse Problems*, vol. 20, no. 5, pp. 1621–1638, Oct. 2004.
- [24] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 2nd ed. Upper Saddle Hill, NJ: Prentice-Hall, 2002.
- [25] F. Jäger, L. Nyúl, B. Frericks, F. Wacker, and J. Hornegger, "Whole body MRI intensity standardization," in *Bildverarbeitung für die Medizin 2007*, A. Horsch, T. Deserno, H. Handels, H. Meinzer, and T. Tolxdorff, Eds. Berlin, Germany: Springer, Mar. 2007, pp. 459–463.
- [26] A. Davies and P. Samuels, *An introduction to computational geometry for curves and surfaces*. New York: Oxford Univ. Press, 1996.
- [27] D. A. Hahn, J. Hornegger, W. Bautz, T. Kuwert, and W. Römer, "Unbiased rigid registration using transfer functions," *Proc. SPIE Med. Imag.*, vol. 5747, pp. 151–162, 2005.
- [28] F. Jäger, Y. Deuerling-Zheng, B. Frericks, F. Wacker, and J. Hornegger, "A new method for MRI intensity standardization with application to lesion detection in the brain," in *Vision Model. Visualization*, L. Kobbelt, T. Kuhlen, T. Aach, and R. Westermann, Eds. Köln, Germany: AKA GmbH, Nov. 2006, pp. 269–276.
- [29] D. Collins, A. Zijdenbos, V. Kollokian, J. Sled, N. Kabani, C. Holmes, and A. Evans, "Design and construction of a realistic digital brain phantom," *IEEE Trans. Med. Imag.*, vol. 17, no. 3, pp. 463–468, Jun. 1998.
- [30] K. Rohr, *Landmark-Based Image Analysis Using Geometric and Intensity Models*, ser. Computational Imaging and Vision, M. A. Viergever, Ed. Dordrecht, The Netherlands: Kluwer, 2001.
- [31] Y. Rubner, C. Tomasi, and L. J. Guibas, "The earth mover's distance as a metric for image retrieval," *Int. J. Comput. Vis.*, vol. 40, no. 2, pp. 99–121, Nov. 2000.
- [32] G. Hermosillo, C. Chef'd'hotel, and O. Faugeras, "Variational methods for multimodal image matching," *Int. J. Comput. Vis.*, vol. 50, no. 3, pp. 329–343, Dec. 2002.