

Effectiveness of Intensity Normalization on Human Brain MRIs with Multiple Sclerosis

Mohak Shah^{1,3}, Yiming Xiao¹, Simon Francis^{2,3}, Douglas Arnold^{2,3}, D. Louis Collins², and Tal Arbel¹

¹ Centre for Intelligent Machines, McGill University, Canada
{mohak,arbel}@cim.mcgill.ca, yiming.xiao@mail.mcgill.ca,

² Montreal Neurological Institute, McGill University, Canada
louis.collins@mcgill.ca, {simon,doug}@mrs.mni.mcgill.ca,

³ NeuroRx Research, Montreal, Canada.

Abstract. Tissue intensity standardization is an important preprocessing step in the study and analysis of Magnetic Resonance Images (MRI) of human brain. Sources of variations in the intensity ranges across different MRI volumes, even after intensity inhomogeneity correction, can result from heterogeneity of data due to difference in scanners, presence of Multiple Sclerosis (MS) lesions and the stage of disease progression in the brain. As a result, intensity normalization plays a significant role in standardizing the tissue intensity ranges across MRI volumes on which most automatic image analysis methods base their distributional assumptions. The method of Nyul *et. al.* [3] has become a widely used standard for intensity normalization. However, an extensive validation of this approach on multi-site multi-scanner data in the presence of MS has yet not been performed. We aim to undertake this validation in this work and show the effectiveness of this procedure on standardizing tissue intensities in the MRI volumes.

1 Introduction

Multiple Sclerosis (MS) is the most common neurological disease affecting young adults in North America. MS is a demyelinating disease of CNS where myelin, the insulation around nerve fibre (axon) is attacked resulting in focal MS lesions. Magnetic Resonance Imaging (MRI), because of its ability to visualize lesions with very high contrast sensitivity, has proven to be critical in identifying the MS lesions serving as disease markers. Most of the automatic and semi-automatic lesion identification and tissue type segmentation approaches applied to brain MRI volumes rely, explicitly or implicitly, on strong assumptions regarding the shape of the underlying distribution of various tissue type intensities. However, large variations in intensity ranges can violate these assumptions significantly and adversely affect their success. These variations stem not only from the differences in the protocols of various MRI scan acquisitions, the different manufacturers and scanner models, but also due the fact that scans from various sites can correspond to patients at varying stages of the disease, and the presence of pathology can impact the tissue intensity behaviors significantly. As a result, a lack of

standard intensity scale makes it difficult to generalize the relative behavior of various tissue types across different volumes. Hence, in addition to having a standardized protocol of data acquisition, it is further necessary to preprocess the data before any segmentation or lesion identification technique can be applied. There are two main steps in such pre-processing: (1) the intensity inhomogeneity correction [5] and (2) the intensity normalization. The former is generally aimed at addressing scanner specific variations in specific MRI images. In this paper, we address the latter and examine a technique to correct the non-uniformity in intensity scales across various images. We aim at assessing the effectiveness of this normalization over the variations mentioned above.

Various intensity normalization methods have been proposed to deal with intensity scale inhomogeneity of various tissue types and have met with varying degrees of success. Among the approaches aimed at MRI intensity normalization, some of the main techniques include those based on utilizing even-order histogram derivatives [7], template histogram matching using multiplicative correction field [6], region-specific standardization [21], and mapping the normalization problem to non-linear registration problem [20]. Although, some methods such as the one proposed in [20] show good results, they suffer from high computational complexity requirements, partially due to the non-linear registration requirements. Some of the issues with other faster methods include lack of adaptability to brain images (e.g., in the case of [21]), dependence on specific brain segmentation algorithm [7] and difficulty in obtaining standard templates (e.g., in the case of [6]).

One of the most widely accepted and utilized intensity normalization methods has been that of Nyul *et. al.* [1] which seems to give the best of both worlds in the sense that the method is easy to implement, adaptable to various body parts and gives results comparable to the more accurate, but slower, methods described above[19].⁴ The two stage method consist of a training stage to find the parameters, which includes the range and mode, of the standard scale and a transformation stage that maps the histograms of candidate volumes to the standard histogram scale. Nyul *et. al.* [1] also present a quantitative and qualitative evaluation of this intensity normalization method with regard to intra- and inter-patient variations. However, there is no evaluation available, to the best of our knowledge, of the effectiveness of this normalization method across scanners from different brands or manufacturers as well as variations occurring due to different machines of the same brand. Furthermore, it is not conclusively confirmed if the intensity normalization is equally effective in the presence of pathology such as MS in the brain and reflected in the resulting MRIs.

In this work we study the effect of these variations on the efficacy of the intensity normalization procedure of [1] both qualitatively and quantitatively. In addition to investigating the effect of the intensity normalization and its dependence on data from heterogenous sources. We also examine its behavior on the different MRI acquisition modalities. Finally, and probably more importantly, we assess the effect of the technique in the presence of MS lesions.

⁴ The earlier version of the Nyul approach appears in [3].

The rest of the paper is organized as follows: Section 2 briefly describes Nyul *et. al.*'s decile formulation of the intensity normalization of [1] utilized in this work. Section 3 gives details of the data acquisition and the image non-uniformity and inhomogeneity correction preprocessing. Section 4 presents the evaluation framework followed by qualitative and quantitative evaluation results in Sections 4.1 and 4.2 respectively. Finally, we conclude with a discussion of future work in Section 5.

2 Decile Formulation of Normalization method of Nyul *et. al.* [2]

The two-stage method of Nyul and Udupa [3, 1] has some significant advantages over other methods. It is both easier to customize to various body regions and is fast in practice. Moreover, the method does not sacrifice accuracy for gains in speed. Some of these methods were compared by [19]. Their results further confirms these conclusions. Another major advantage of this approach is that it does not rely heavily on specific statistical properties of tissue classes and has a lower computational complexity. As discussed by [2], the original formulation of the normalization procedure of [3] suffers from some limitations in the case of application to the MRIs of brains with MS. One of the main issues arose from the fact that the method relied on a single landmark over the intensity histogram (the mode) to identify the standard scale resulting in a switching behavior whereby incorrect tissue types can be mapped against each other.

The authors of [2] suggest variants to the proposed intensity normalization methods. However, rigorous validation of these variants has not been done so far. We now present a variant that is an adaptation of the method of [3, 1] with deciles piecewise linear mapping as described below.

2.1 Notations

We consider a volume image $\mathcal{I}$ consisting of voxel elements $e \in E$ and an intensity mapping function $f_{ei} : E \rightarrow \mathbb{N}_0$ that maps each voxel element e to an intensity value. Hence, $f_{ei}(e) \geq 0, \forall e \in E$ with the equality satisfied iff there is no measured data for element e , and $\mathbb{N}_0$ is the set $\{0, 1, 2, \dots\}$ of natural numbers. Hence, each volume image can be described jointly by the set of voxel elements E and the associated intensity mapping function f_{ei} $\mathcal{I} = (E, f_{ei})$. That is, $\mathcal{I}$ represents, in our case, three dimensional brain volume image acquired through MR Imaging. f_{ei} is element to intensity transformation function that assigns an integer intensity value $i \in \mathbb{N}_0$ to each element e . Note, however, that we have access only to the intensity values for each voxel element e . The function f_{ei} is considered to be the underlying intensity mapping function but this function is generally not known.

We denote by $\mathcal{A} = \{A_1, A_2, \dots, A_k\}$ the set of acquisition modalities of MRIs such that $|\mathcal{A}| = k$. In our case, we focus on mainly three image modalities: the PD, T1 and T2 weighted images. That is, we have $k = 3$.

Let $\mathcal{I}_{\mathcal{A}}$ denote the superset of all images $\mathcal{I}$ that can be acquired according to the protocols in $\mathcal{A}$. The histogram of an image $\mathcal{I}$ is denoted by $h = (I, f_{ic})$

such that $I \subset \mathbb{N}_0$ is the set of all possible intensity values in $\mathcal{I}$ and $f_{ic} : I \rightarrow \mathbb{N}_0$ is intensity to count mapping function such that $f_{ic}(i) = |\{f_{ei}(e) = i\}|$. That is, the function $f_{ic}(i)$ outputs the number of voxel elements that have an associated intensity value of i . Finally, let i_{min} and i_{max} denote the minimum and maximum intensity values respectively in an image $\mathcal{I}$. That is, $i_{min} = \min \{f_{ei}(e) | e \in E, f_{ei}(e) \neq 0\}$ and $i_{max} = \max \{f_{ei}(e) | e \in E, f_{ei}(e) \neq 0\}$.

2.2 Choosing Intensity range

The tails of the image histograms are generally pruned so as to make the algorithm robust against artifacts and outliers resulting in inter-patient and scanner variations. This pruning results in an intensity range called intensity of interest (IOI). Let p_{low} and p_{high} denote the minimum and maximum percentile values of the overall intensity range that bounds this IOI. Now, let i_{min}^h and i_{max}^h be the minimum and maximum intensity values in the histogram h corresponding to p_{low} and p_{high} respectively. We consider the more common bimodal histograms for the MR images as suggested by [3] and hence use the second mode, denoted as m corresponding to the main foreground object in the image as a histogram landmark.

We consider the following landmark configuration C_L :

$$C_L = [p_{low}, p_{high}, m_{10}, m_{20}, m_{30}, m_{40}, m_{50}, m_{60}, m_{70}, m_{80}, m_{90}]$$

where each $m_i, i \in \{10, 20, \dots, 90\}$ denotes the mode at each decile. Finally, let i_{min}^s and i_{max}^s be the minimum and maximum intensity values in the standard scale respectively (the superscript denoting the standard scale). Given, this setting, our intensity normalization algorithm is implemented in two stages as described below.

2.3 The Algorithm

The two-stage intensity standardization algorithm can be summarized as follows. For each image in the training set of images, we determine the image histogram. From this image intensity histogram, we determine the intensity value i_{min}^h corresponding to p_{low} and the intensity value i_{max}^h corresponding to p_{high} the upper percentile. Now, we identify the intensity values $m_{10}^i, \dots, m_{90}^i$ that correspond to each decile of this intensity range of interest. The minimum and maximum intensity values of the image i_{min}^h and i_{max}^h are then mapped to the corresponding minimum and maximum intensity values i_{min}^s and i_{max}^s of the standard intensity scale respectively. Note that the values forming the range of the standard intensity scale i_{min}^s and i_{max}^s are parameters of the learning algorithms and are chosen by the user. The choice generally depend on the trade-off between the desired resolution of the resulting intensity standardization scale and the desired efficiency in the intensity mapping in each segment of this intensity range under the decile formulation. Once the mapping from the actual image intensity range to the standard intensity scale is done, new landmarks corresponding to each decile of the mapped image $m_{10}^i, \dots, m_{90}^i$ are calculated. The landmarks for

the standard scale are then calculated from the rounded means of each of the landmarks from the mapped image from the training set of images.

The landmarks of the standard scale obtained from the training phase are then used, in the second stage of the algorithm, to perform the transformation of a new MRI image to the standard scale as follows. Given a new MRI image $\mathcal{I}_i$, we compute the histogram h_i of the new image and determine the p_{low} and P_{high} percentiles. We also determine the intensity values corresponding to each decile (landmark locations) in this image. These intensity values corresponding to the decile landmarks enables segmenting of the image in ten segments. Each of the segment is then linearly mapped to the corresponding segment of the standard image intensity scale determined in the training stage of the algorithm. This output the standardized image $\mathcal{I}_i^s$ of the input image $\mathcal{I}_i$.

3 Data Acquisition and pre-processing

To test the performance of MRI intensity normalization techniques by [1] across MRI scanners made by different manufactures, and different brands of MRI scanners from the same manufacture, a group of subjects' MRI scans (T1w, T2w, and PDw modalities), were obtained courtesy of NeuroRx Research (Montreal, Canada).

The MR images for study were selected based on a combination of diverse criterions. The selected group includes in total 21 MRI scans, with seven subjects acquired from MRI machines made by each of the three major MRI machine manufactures (GE, Phillips, and Siemens).⁵ Within each sub-group of 7 cases, the sampled patients also possess varying sizes of ventricles, and varying loads of MS lesions. An effort was made to standardize the protocols in a scanner-specific manner that resulted in comparable contrasts across images from different scanners. Each MRI volume has a resolution of $1 \times 1 \times 3$ mm for voxels in each slice with each volume containing 50 slices.

Different MRI modalities acquired may not correlate spatially upon acquisition due to various reasons such as patient movement, scanner behavior over time and so on. An alignment is required as a result so as to obtain spatial correlation between brain volumes. Therefore, image modalities are aligned to a common stereotaxic space so that the voxel locations between modalities have a deterministic mapping. Following the alignment, an intensity inhomogeneity correction is performed followed by the intensity standardization (normalization).⁶

Factors such as the Partial Volume effect resulting from the limited scanner resolution can affect the tissue intensity behavior significantly. However, the nature of these effects is not very well understood. Hence, we limit ourselves to studying the intensity behavior of various tissue types from voxels that are free

⁵ The scanner models included Signa, Genesis Signa and Signa Excite from GE, Intera and Infinion from Philips, and Symphony, Symphony Vision, Allegra and Avanto from Siemens.

⁶ The order of the intensity inhomogeneity correction and the intensity normalization can have significant impact as noted by [8]. In accordance with their observations, we follow the intensity inhomogeneity correction by intensity normalization step.

from such variations. We refer to such voxels as “pure” tissue voxels. Thus, we obtain “pure” tissue samples from each tissue type category. Subsequently, the effects of normalization are studied and evaluation performed on these “pure” tissue samples for each tissue category. Voxels corresponding to the tissues belonging to four tissue classes: white matter (WM), cortical gray matter (CGM), deep gray matter (DGM) and cerebrospinal fluid (CSF), were manually sampled uniformly from various anatomical locations. Further sampling is performed in alternate slices so as to obtain representative samples from the whole volume averaging about 1000 samples per tissue type from approximately 20 slices per volume. These samples are obtained from MRI images both before and after the intensity normalization is performed. Due to the fact that MS lesions are more difficult to be identified correctly, lesions were first identified automatically using heuristic-based classification technique [4] and are then further validated by five expert radiologists. The samples of the lesions accepted are the ones that are agreed upon by a consensus of experts.

4 Evaluation

We perform the training of the intensity normalization algorithm on a set of 100 separate MRI volumes with varying brain volume, scans at different stages of MS and varying lesion loads. The standard scale histogram parameters were obtained from this training set of 100 Multiple Sclerosis brain volumes and were the same for all image modalities. These parameters then defined the standard scale used for normalization of subsequent images.

The evaluation of the normalization algorithm was performed keeping in perspective the potential advantages and uses of the resulting normalized images. The main factors hence include testing whether the tissue intensities behave in accordance with the assumptions made by various automatic lesion segmentation and analysis techniques, whether the change in tissue intensity behavior helps in increasing the efficacy of such methods, the effect of MRIs coming from heterogeneous sources including scanners from different manufacturers as well as different scanner (models) from the same manufacturer. The dataset chosen for testing accordingly takes these factors into account. Further we also performed inter-patient analysis of intensity behaviors across different tissue types. Finally, unlike [2], we do not work with lesion removed images since we are interested in understanding the behavior of the tissue intensity normalization method in the presence of this MS pathology.

4.1 Qualitative Evaluation

We first present qualitative effect of the intensity normalization on the multi-spectral MRI volumes. Figure 1 shows the effect of normalization on axial image slices of the T1, T2 and PD modality MRI volume from an MS patient from among the ones used to test the procedure. We study both the overall effect and the effect of normalization on scanners from various manufacturers. Figure 2 presents the scatter plots of the multispectral volumes grouped by tissue type for all MRI volumes as well as for the MRI volumes grouped by scanner

manufacturer. Each point in the scatterplots in the first row represent the three dimensional intensity vector for the corresponding voxel with each value in the vector denoting the intensity value in one of the acquired image modalities (T1w, T2w, PD). The same data after intensity normalization is shown in the second row. From the first row of Figure 2 it can be seen how each set of tissue classes

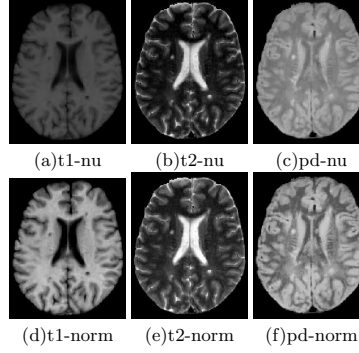


Fig. 1. Qualitative images of MRI slices before and after normalization. The three images in the first row give T1, T2 and Pd slices from an un-normalized MRI volume; the three images in the second row give the corresponding normalized slices.

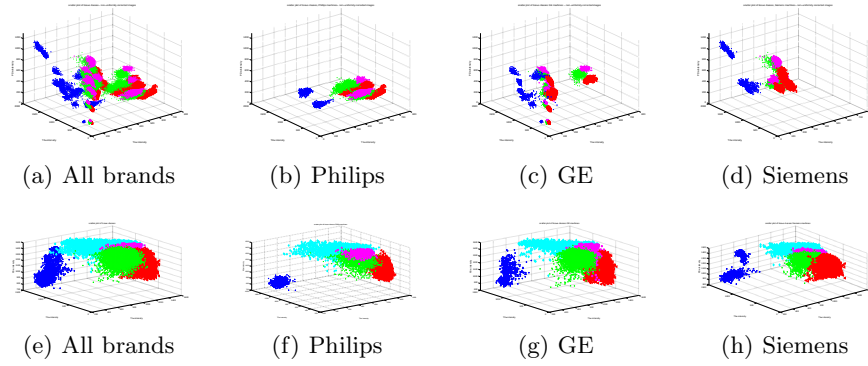


Fig. 2. Intensity distribution scatterplots of various brand machines. The first row gives intensities before normalization and the second row after normalization. The intensities in Red denote White matter, in Green denote Cortical Gray Matter, in Magenta denote Deep Gray Matter, and in Blue denote CSF. Lesions are not included in the unnormalized image intensities since they were labelled after the image normalization process and hence expert labeling was not available for the lesions.

(WM, Deep GM, Cortical GM, and CSF) belonging to the same subject has a characteristic spatial distribution. When all subjects are plotted together, the variations in imaging acquisition from different scanner brands and their sub-series are visible. As shown in the second row, these variations in tissue intensity

distributions' localization relative to each other are largely reduced as a result of the intensity normalization. The same effect can also be seen in the case of scanners from different manufacturers. A standardized characterization of intensity distributions for various tissue types on a standard scale has significant implication on the performance of both the automatic lesion identification as well as tissue segmentation algorithms.

4.2 Quantitative Evaluation

Various studies have suggested that the underlying distribution of various MRI tissue intensities from human brain can be approximated as Gaussian [10, 11]. Moreover, many parametric algorithms that perform tissue segmentation as well as automatic or semi-automatic lesion identification support and rely on the Gaussian assumption (see, for instance, [17]). We wish to investigate how representative are the voxel intensities from various tissue types of a Gaussian model built around the necessary statistics computed from the data. We will use the Jeffrey divergence measure to assess the impact of the intensity normalization on these assumptions as discussed below.

Since the intensity normalization is applied on a single modality basis, we restrict all our analysis for distributions of tissue intensities to single modalities as well. However, some studies do claim that the tissues maintain a Gaussian distribution in a multi-spectral space (see, for instance, [17] for T1w, T2w and PD modalities).

For the un-normalized MRI volumes, we first compute the necessary statistic of a Gaussian distribution from the data which would be the mean vector μ_t^a and the variance $(\sigma_t^a)^2$ over the MRI intensity values for tissue type t where $t \in \mathcal{T}$ such that $\mathcal{T}$ is the set of all tissue types and $a \in \mathcal{A}$ such that $\mathcal{A}$ denotes the set of acquisition modalities. We consider $\mathcal{T}$ containing CorticalGM, WM, DeepGM, CSF and Lesion tissue types. We can then generate samples of a Gaussian distribution centered at μ_t^a and with a variance $(\sigma_t^a)^2$, i.e., $\mathcal{N}(\mu_t^a, (\sigma_t^a)^2), \forall t \in \mathcal{T}, \forall a \in \mathcal{A}$. Next, for each tissue type $t \in \mathcal{T}$, we generate 100-bin sized histograms from the data over three modalities taking into account 98 data percentile (upper and lower 1 percentile data left out as noise and outliers). This gives an account of the actual model of the data for the given tissue type.

In order to verify whether the sample distributions differ significantly from Gaussians, we employ the Jeffrey divergence measure to calculate the distance between these two distributions. Jeffrey divergence is a symmetric measure of similarity between two distributions. Low values of divergence represent less difference between the two distributions. For any two continuous probability distributions $p(\cdot)$ and $q(\cdot)$, the Jeffrey Divergence [18] is defined as:

$$D(p, q) = \int_{-\infty}^{\infty} p(x) \log \frac{p(x)}{q(x)} dx + \int_{-\infty}^{\infty} q(x) \log \frac{q(x)}{p(x)} dx$$

In our case, under gaussian assumptions for $p(x)$ and $q(x)$ we can use a discretized approximation of the above metric.⁷ In order to assess the effect of the intensity normalization, we employ the same distribution generation method as discussed above but this time on intensity normalized MR images. We again calculate the Jeffrey divergences. The data statistics and the Jeffrey divergence values before and after intensity normalization are shown in Figure 3. Hence, We effectively calculate the divergences between a Gaussian distribution centered at sample mean with the sample variance, and the histogram model built from observed intensities. If the sample indeed comes from a Gaussian distribution then this difference would be small. Also, if intensity normalization does not play any role then the divergence value after normalization should not differ too much from the one before normalization.

It can be seen that clearly intensity normalization has a significant role to play in tissue modeling as a result of the reduced Jeffrey divergence measure over the distribution differences after normalization. We have also performed the same analysis to check for scanner-specific effects. To do so, we perform a similar analysis but this time only on MRI images from scanners from a specific manufacturer. As can be seen in Figure 3, the same observations and conclusions as above hold in these cases too. Intensity normalization, hence, results in data that is more homogenous for a given tissue type and that is more representative of the underlying distribution. This has significant potential impact in further automatic processing and analysis of the data including automatic image segmentation with regard to various tissue types as well as identifying pathological tissues.

4.3 Tissue Divergences

In order to investigate the effect of intensity normalization on the ability to discriminate between various tissue types, we again employ the Jeffrey divergence measure between the distributions of pairs of tissue types before and after normalization. In order to do this, we obtain the histogram models over the voxel intensities for the pairs of tissue to be compared. We calculate the distance between these two distributions using the Jeffrey divergence measure as described above. Next, we do the same for normalized images. If the intensity normalization improves the tissue contrast then we should see an increase in these divergence measures between the two distributions on the normalized images. That is, if we assume that the normalization has no effect in making intensities from the same tissue type more homogenous (and different from those of other tissue types) then we would not observe any change in the divergence values between the distributions of two different tissue types as a result of normalization.

For this analysis, we concern ourselves with the characteristic of primary tissue types, that is, the GM (combined Deep GM and Cortical GM), WM

⁷ In addition to considering various studies that claim that the tissue distribution can be approximated as Gaussian, we also performed goodness of fit tests with different known distributions. The results from these tests too suggested Gaussian distribution to be able to approximate the underlying data generating model most closely, even for un-normalized data.

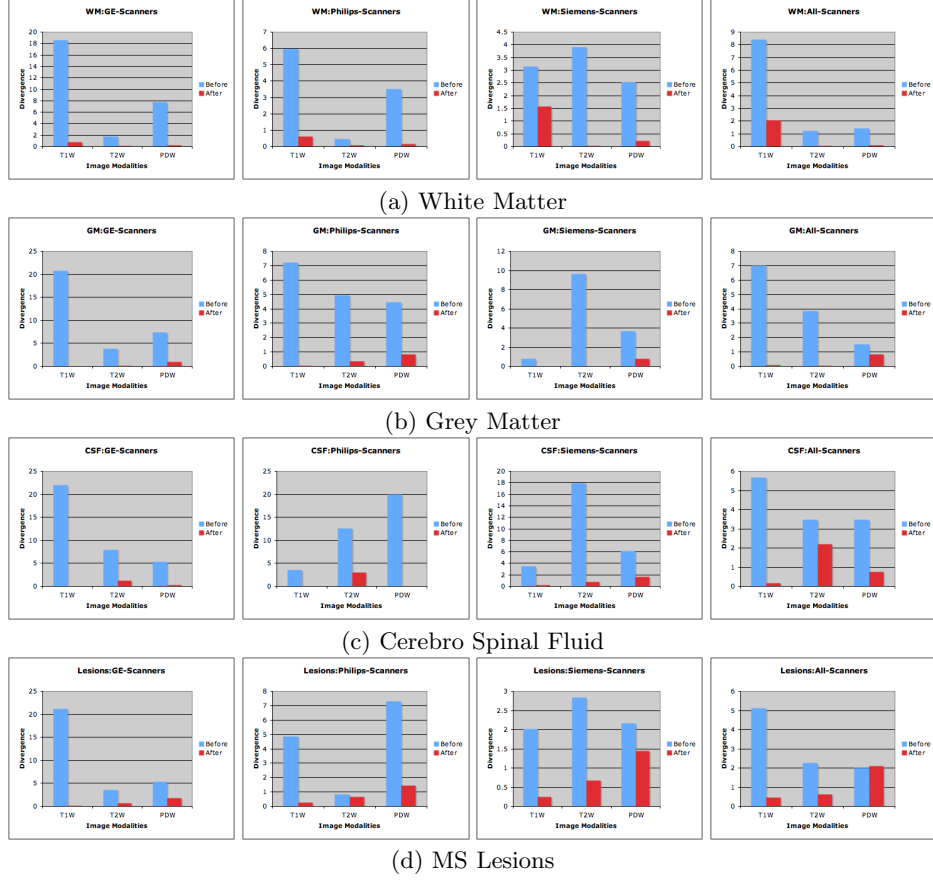


Fig. 3. Divergence measures between distribution fits before and after normalization. The blue bars in each graph denote, for a given modality and tissue type, the Jeffrey divergence measure between the Gaussian distribution centered at the sample mean with sample variance, and the histogram model generated from tissue intensities on the un-normalized volumes. The red bars correspond to the same measure over volumes after normalization. An absence of bar indicates that the divergence was too small to plot. Each row corresponds to measures for a given tissue type. Each column refers to divergence measures for MRI volumes from different machines: GE, Philips, and Siemens (for the first three). The last column gives the measure over all MRI volumes taken together.

and CSF. Since, we deal with pure samples only, we do not include the lesions in this analysis since these are affected by many factors such as the Partial volume effects that we do not take into account for this study. The results are presented in Table 4. As can be seen, the intensity normalization contributes to a better relative scale mapping of various tissue types contributing to tissue type separation in intensity space. The only discrepancy comes for the contrast between WM:GM contrast in T2w scans. One possible reason can be the effect of the global 98 percentile dynamic range mapping resulting in a range compression for WM and GM.

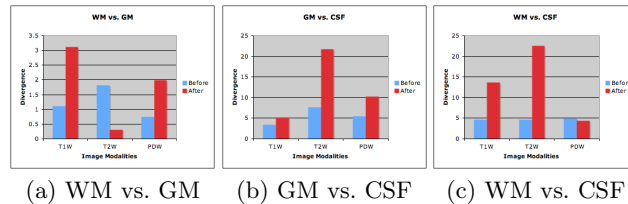


Fig. 4. Jeffrey Divergences between tissue types before (blue bar) and after (red bar) Normalization. The other notations are the same as previous figures.

5 Conclusion and Future Work

In this work we investigated the efficacy of the intensity standardization procedure of Nyul *et. al.* [3,2] on multi-site multi-scanner data from multi-spectral (T1, T2 and PD weighted) MRI volumes of patients with varying degrees of MS lesion load. Utilizing the Jeffrey divergence criteria, we also verified whether the assumptions of an underlying Gaussian distribution on different tissue types made by various automatic tissue segmentation and visualization approaches hold. Finally, we also studied the effect of the procedure on tissue separation among different tissue types.

Our results demonstrate that intensity normalization has a significant role to play in both enabling efficient tissue modeling and relative scale mapping. Rendering the tissue intensity values in a standardized range enables verification of the modeling and distributional assumptions made by many automatic image analysis techniques. Further, the intensity normalization also results in more homogenous intensity values for voxels of the same tissue type. Along with an effective relative scale mapping, this also contributes to a better tissue type separation in intensity space. Further, these conclusions seem to hold across a range of data from heterogenous sources with varying scanner types and, importantly, with varying degrees of MS pathology in the brain.

Although we examined the overall behavior of the intensity normalization procedure on MRIs of patients with varying degrees of MS lesion load, the natural extension of this analysis would be to explore its correlation with the amount of lesion load. We plan to investigate this in future. Since most of the automatic segmentation and visualization methods operate in multispectral tissue intensity space, it would also be interesting to study quantitatively the effect of intensity normalization on multispectral tissue intensity behavior and on tissue segmen-

tation. We also plan to extend our analysis to take this factor into account.

References

1. L. G. Nyul and J. K. Udupa, On standardizing the MR Image Intensity Scale, *Mag. Res. in Med.* 42, 1072-1081, 1999.
2. Laszlo G. Nyul *et. al.*, New Variants of a Method of MRI Scale Standardization, *IEEE TMI* 19(2), 143-150, 2000.
3. L. G. Nyul and J. K. Udupa, An Approach to Standardizing the MR Intensity Scale, *Med. Imag.* 1999(3658), 595-603, 1999.
4. S. Francis, Automatic Lesion Identification in MRI of Multiple Sclerosis Patients, MS thesis, McGill University, Montréal, 2004.
5. B. Belaroussia *et. al.*, Intensity non-uniformity correction in MRI: Existing methods and their validation. *MIA* 10(2), 234-246, 2006.
6. N.I. Weisenfeld and S.K. Warfield., Normalization of joint image-intensity statistics in mri using the kullback-leibler divergence. In *IEEE ISBI*, 2004.
7. J.D. Christensen, Normalization of brain magnetic resonance images using histogram even-order derivative analysis, *Mag Res Imag* 21, 819-442, 1996.
8. A. Madabhushi and J. K. Udupa, Interplay between intensity standardization and inhomogeneity correction inMR image processing, *IEEE TMI*, 24(5), 561-576, 2005.
9. J. Puzicha *et. al.*, Empirical evaluation of dissimilarity measures for color and texture. In *ICCV-1999*, 2, 1165-1173, 1999.
10. P. Santago and H. D. Gage, Quantification of MR brain images by mixture density and partial volume modeling. *IEEE TMI* 12(3), 566-574, 1993.
11. M. B. Cuadra *et. al.*, Comparison and validation of tissue modelization and statistical classification methods in T1-weighted MR brain images. *IEEE TMI* 24, 1548-1565, 2005.
12. Y. Zhang *et. al.*, Segmentation of brain MR images through a hidden Markov random field model and the expectation maximization algorithm. *IEEE TMI* 20, 45-57, 2005.
13. C. R. Wirtz *et. al.*, Intraoperative magnetic resonance imaging to update interactive navigation in neurosurgery. Method and preliminary experience. *Comput Aided Surg*, 2, 172-179, 1997.
14. L. O'Donnell and C. F. Westin, High-dimensional white matter atlas generation and group analysis. *MICCAI* 2006, 243-251, 2006.
15. R. W. Cox and A. Jesmanowicz., Real-time 3D image registration for functional MRI, *Magn. Reson. Med.*, 42, 1014-1018, 1999.
16. W.M. Wells III *et. al.*, Adaptive segmentation of MRI data, *IEEE TMI*, 15, 429442, 1996.
17. R. Harmouche, Bayesian MS lesion classification modelling using Regional and Local Spatial Information, Master's Thesis, McGill University, 2006.
18. Jan Puzicha *et. al.*, Non-parametric similarity measures for unsupervised texture segmentation and image retrieval. *CVPR '97*, 267, 1997.
19. J. Bergeest and F. Jäger, A Comparison of Five Methods for Signal Intensity Standardization in MRI, In *Bildverarbeitung fr die Medizin*, Berlin : Springer 2008, pp. 36-40, 2008.
20. F. Jäger *et. al.*, A new Method for MRI Intensity Standardization with Application to Lesion Detection in the Brain, *Vision Modeling and Visualization* 2006, pp. 296-276, 2006.
21. P. Hellier, Consistent intensity correction of MR images. *ICIP* 2003;1, pp. 1109-1112, 2003.
22. I. J. Cox *et. al.*, Dynamic histogram warping of image pairs for constant image brightness, *ICIP* 1995, 2366-2369, 1995.