A Hybrid Approach to Tissue-based Intensity Standardization of Brain MRI Images

Raghav Mehta *Jayanthi Sivaswamy*

CVIT, IIIT-Hyderabad, India

ABSTRACT

The variations in the intensity scale in Magnetic Resonance Images pose a problem for many tasks and Intensity Standardization (IS) aims to solve this problem. Existing methods generally use landmark values of the image histogram and match it to a standard scale. The landmarks are often chosen to be percentiles from different segmented tissues. We propose a method for IS in which tissue information (via segmentation) is needed during training but not during testing by using landmark propagation. A KL divergence-based technique is employed for identifying volumes from the training set, which are similar to a given non-standardized testing volume. The landmarks from the similar volumes are then propagated to the given test volume. Evaluation of the proposed method on 24 MRI volumes from 3 different scanners shows that the IS results are better than L4 and at par with a method which uses prior segmentation, to get percentile-based landmarks. The proposed method aids speeding up and expanding the scope of IS to volumes with no tissue information.

Index Terms— Magnetic Resonance Images, Intensity Standardization, KL Divergence

1. INTRODUCTION

Magnetic Resonance Images (MRI) of the brain is widely used for the diagnosis of many neuro disorder starting from brain aneurysms and stroke, to Dementia, MS lesions and cancer. Atlas construction is another application widely used to get population specific normative parameters of the brain. This requires Registration of MRI of different patients from different scanner manufacturers. Disease diagnosis is based on quantitative information derived from segmentation and registration (to atlas) of MR volumes.

A major problem in MRI is that it lacks standardization of voxel intensity values. There is inter-scanner, intra-scanner, inter-patient, intra-patient variations of intensity profile for same tissue type. It is well known that this kind of variation affects processing steps like segmentation, registration [1] etc., which in turn affects above mentioned applications.

Various solutions have been proposed for the Intensity Standardization (IS) task. These approaches can be categorised as A: independent of or B: dependent on prior knowledge of tissue labels. Earlier approaches to IS belong to category A. Here, IS is performed in the histogram space using mappings. These include one-to-one, one-to-many and many-to-one mapping between non-standard to a standard-volume [2]. The importance of preserving tissue distributions were recognised and a Gaussian Mixture Modelling was used to estimate different tissue types and their variation across volumes was minimised [3]. One of the most popular method for IS uses landmark-based matching and mapping [4]. Landmarks are based on modes of histogram [5] or percentiles of the foreground [4][6]. IS has also been cast as a non-rigid registration problem [7].

More recent methods are in category B. For example, [8] uses joint image histograms to determine intensity correspondence between the input and standard volumes in each tissue class, and uses it as landmark points. In contrast, [9] identifies the landmark points for each tissue class individually and matches them to standard scale. This method is reported to outperform [4].

Tissue distribution based IS (category B) is said to improve the performance in terms of tissue separability and tissue Gaussianity [9] [8]. However, the preprocessing steps of brain extraction and tissue segmentation are time consuming and hence these methods are slower than the methods in category A.

We propose a hybrid approach to IS which exploits the strengths of both categories. This is motivated by the fact that i) In a practical scenario, it is not possible to have tissue segmentation for every volume, ii) IS is a preprocessing step for further image analysis and hence needs to be computationally light. We propose to restrict the use of tissue labelled volumes to the training phase of IS. Given a new volume with unlabelled voxels, we propose a novel way to identify the landmarks, by propagating the landmark points from similar volumes, identified using KL divergence, in the training dataset.

We use the method proposed in [9] to illustrate our idea. Nevertheless, our approach of finding similar image from training dataset and use its information for testing data is equally applicable to [8] and other approaches based on prior knowledge of tissue distributions.

This work was partly funded by the Department of Science and Technology, Govt. of India, under Grant SR/CSRI/194/2013(G).

978-1-4799-2349-6/16/$31.00 ©2016 IEEE

2. PROPOSED METHOD

The proposed method consists of two stages: offline training and online testing. During training we have a set of non-standardized MRI volumes I_n; and corresponding Tissue Masks M_n^j where $n = 1, ..N$ and $j =$ Grey Matter (GM), White Matter (WM) and Cerebrospinal Fluid (CSF). During testing, we have a given MRI volume I_g with no tissue mask. All volumes (training and testing) are preprocessed to identify the foreground by thresholding the voxel intensity at a threshold value identified using Otsu's method [10].

2.1. Training Stage

Given I_n and M_n^j the main steps in the training phase are:

1) Calculation of grey level histograms and corresponding percentiles for each tissue type P_n^j.

2) Determination of transformations $T_n^j, \forall j$ by matching the percentile landmarks P_n^j of I_n with the corresponding standard-scale percentile landmarks P_s^j.

3) Derivation of the continuous mapping T_n for the whole volume via spline-fitting through T_n^j as in [9].

2.2. Testing Stage

Our main contribution is in this stage. Given a test volume I_g our aim is to derive a transformation T. The key information needed is the landmarks on the tissue histograms. Since, no tissue-level information is available for I_g, we proposed to derive it by finding the nearest neighbours of I_g in the training set I_n.

2.2.1. *Finding the Nearest Neighbours*

First, the intensity range of the training volume is remapped to that of testing volume intensity range by a linear transformation. We choose the KL divergence (KLD) [11] metric to identify volumes from the training set I_n which are similar to I_g. Given two Probability Distribution Functions(PDF) P and Q the KLD is given as

$$KLD(P,Q) = \sum_i P(i) \log \frac{P(i)}{Q(i)}$$

In our case, KLD is computed between the histogram of the given volume and all the volumes in the training set to obtain $KLD(n) = KLD(h_n, h_g)$, where h_x denotes the histogram of an volume I_x.

Low KLD values indicate a high degree of similarity between volumes. Hence, $KLD(n)$ is thresholded to extract the required set of L training volumes, which are the highly similar to I_g. The threshold value controls the size of the set L and the diversity within the set.

2.2.2. *Projecting the landmarks from Training volume*

The L nearest neighbours of I_g, are used to determine the histogram landmarks for I_g as follows. The landmarks P_i^j, $i \in [1, L]$ from the L volumes are projected onto the intensity range of I_g and then averaged to get the required landmarks for the testing volume. The final landmarks are found as

$$P_g^j = \sum_{i=1}^{L} W_i P_i^j, \forall j$$

$$where, W_i = \frac{w_i}{\sum_{i=1}^{L} w_i}, w_i = \frac{1}{KLD(i)}, \forall i$$

A *weighted* averaging, with weights being inversely proportional to KLD values, is chosen to accommodate the possibility that the training and testing volumes may be quite diverse. In this case, the threshold may have to be lowered to permit L of sufficient size to be identified for accurate landmark identification.

2.2.3. *Matching to standard scale percentiles*

The percentile landmarks P_g^j are matched to the corresponding standard-scale landmarks P_s^j for each tissue type individually to get corresponding transformations T_g^j and a spline fitting is performed to obtain a continuous transformation T_g.

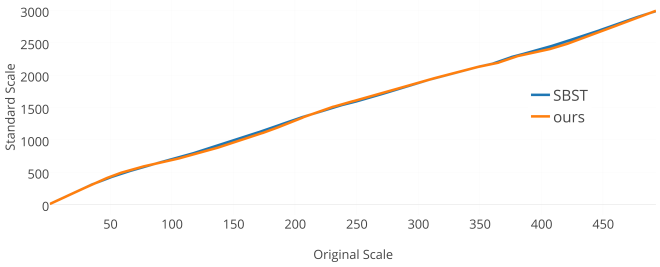

Fig. 1. Mapping function comparison

The effectiveness of this approach is illustrated in Fig.1 with a sample volume for which tissue masks were available. The transformation curves for mapping a given volume to the standard scale were derived with the proposed method by ignoring the tissue information and with tissue information using [9]. These are shown in red and blue, respectively, in Fig.1. The two maps are highly overlapping emphasising the effectiveness of our method.

3. EXPERIMENTS AND RESULTS

Dataset: T1 weighted MRI volumes from 1.5 T (GE, Siemens and Phillips) scanners were used to construct a set of $8 \times 3 = 24$ volumes. The volumes from GE and Siemens scanners were sourced from local hospitals while the ones from the

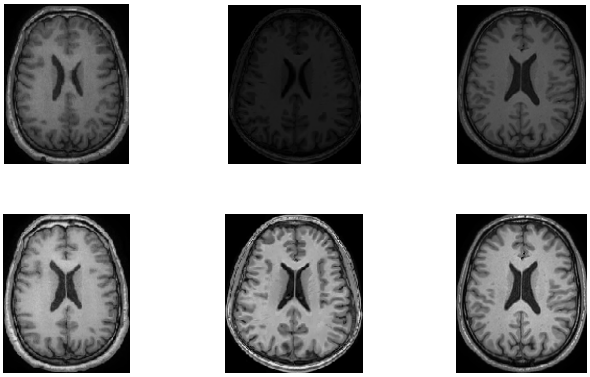

Fig. 2. Sample images from GE, Siemens and Phillips scanners (left to right) before (top row) and after (bottom row) IS.

	TE(ms)	**TR**(ms)	**TI**(ms)	**FA**(°)
GE	4.2	10.2	450	15
Siemens	2.9	2370	1000	7
Phillips	4.6	9.83	NA	8

Table 1. Scanning parameters.

Phillips scanner were sourced from [12]. Details of the acquisition parameters for all scanners is given in Table1.

Validation: Preprocessing steps like denoising and intensity inhomogeneity correction were done using [13] and [14] respectively, on all the MRI volumes. Tissue classification was done with the FAST tool of FSL Library [15]. Prior to tissue classification, skull stripping was done using BET[16].

Validation of the proposed method was done using a leave one out approach. All the computations were done on a PC with Intel $i5$ processor, with 4GB of RAM.

Landmarks for the standard scale were learnt from 30 MRI volumes drawn from different scanners, by projecting their landmarks to the voxel value range of 0:4096. The KLD values were thresholded to obtain (on average) L=3 similar volumes for a given test volume.

Sample images from 3 different scanners and results of IS with our method are shown in Fig.2. All images were shown at a fixed intensity window range of 0:4096. After IS, the images appear more uniform in terms of intensity.

A major issue in MRI is the intensity variations at the inter-scanner and intra-scanner levels. Our IS method reduces these variations as shown in Fig 4 and Fig 3 respectively.

Quantitative analysis of IS was done using the Jeffery Divergence (JD), which is defined for two PDFs P and Q as follow: $JD(P,Q) = 0.5[KLD(P,Q) + KLD(Q,P)]$

The JD was computed for each tissue class, both across volumes within a scanner and across scanners.

Let $I_n^j(s)$ denote the n^{th} volume from scanner s with tissue mask j; $n = 1,..,8$, both j and $s = 1,2,3$. The intra-scanner ($JD_{intra}(s,j)$) and inter-scanner ($JD_{inter}(s,j)$) divergence were computed as follow:

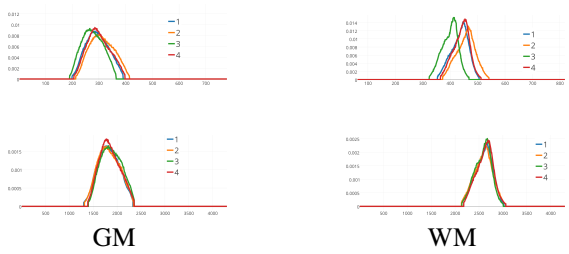

Fig. 3. Intra-Scanner variation in tissue PDF before (top row) and after (bottom row) IS for the Siemens scanner.

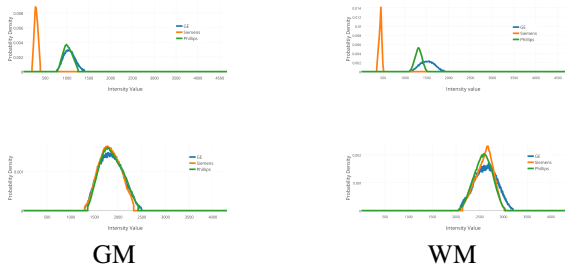

Fig. 4. Inter-Scanner variation in tissue PDF before (top row) and after (bottom row) IS.

$$JD_{intra}(s,j) = \frac{1}{mn}\sum_n\sum_m JD(I_n^j(s), I_m^j(s)),$$

where, $m \neq n, \forall n, m \in s$

$$JD_{inter}(s_a, s_b, j) = \frac{1}{mn}\sum_n\sum_m JD(I_n^j(s_a), I_m^j(s_b)),$$

where,$\forall n \in s_a, \forall m \in s_b$

These are listed in Table 2. Low values signify similarity among volumes which is desirable especially after IS. The effect of IS with 3 different methods are presented to aid comparison. L4 [4], provides a baseline method as it does not use tissue information at any stage. IS results are comparable for all 3 methods for CSF whereas for GM/WM classes, SBST [9] and our method are (comparable and) better than L4. Similar trend is also seen in the inter-scanner divergence.

We also computed the Standard Deviation (σ_{NMI}),% Coefficient of Variance (CV) and mean (μ_{NMI}) for Normalized Mean Intensity (NMI) before and after IS, for all tissue types for the entire dataset. The NMI of a tissue class is with respect to the maximum intensity in a volume. These are tabulated in Table 2. A similar trend as in the case of JD, is observed for σ_{NMI} and CV. However, the μ_{NMI} is comparable for all 3 IS methods across tissues. This is to be expected as the effect of IS is to minise variation rather than shift the mean intensity of a volume.

		Intra-scanner JD (in x 10^{-2})			Inter-scanner JD			NMI statistics (across all volumes)		
		G	**S**	**P**	**G vs S**	**G vs P**	**S vs P**	σ_{NMI}	μ_{NMI}	% **CV**
CSF	Before	7.99	6.06	2.04	1.25	0.28	1.10	0.0240	0.1444	16.621
	L4	3.87	2.74	0.97	0.08	0.03	0.05	0.0127	0.2228	5.7001
	ours	3.53	2.40	0.87	0.04	0.03	0.04	0.0055	0.2506	2.1942
	SBST	3.50	2.35	0.84	0.03	0.02	0.03	0.0049	0.2412	2.0315
GM	Before	16.32	9.28	3.34	1.38	0.48	1.23	0.0305	0.2676	11.3999
	L4	7.70	3.54	2.28	0.15	0.08	0.05	0.0128	0.4129	3.1001
	ours	5.78	2.32	1.35	0.04	0.05	0.03	0.0094	0.4444	2.1152
	SBST	5.67	2.28	1.32	0.03	0.03	0.02	0.0084	0.4427	1.8974
WM	Before	19.53	8.71	3.46	1.38	0.88	1.25	0.0285	0.4049	7.0391
	L4	7.56	5.25	2.95	0.44	0.25	0.14	0.0196	0.5792	3.3836
	ours	5.32	2.47	2.27	0.09	0.07	0.05	0.0119	0.6200	1.9193
	SBST	5.19	2.40	2.19	0.07	0.05	0.04	0.0106	0.6205	1.7082

Table 2. Quantitative Analysis for all tissue types, across 3 scanners [GE(G), Siemens(S), Phillips(P)]

4. CONCLUSIONS

We proposed a hybrid method for IS of MRI volumes which relies on tissue information only during training. A technique for handling the lack of this information in test data was proposed using a KL divergence. Our method outperforms [4] which does not use any tissue information but is at par with [9], which uses prior tissue segmentation for test volume; it is robust to change of scanner. IS of a new volume with our method is significantly fast (as skull stripping and tissue segmentation are not required) and it can be easily adapted to any existing method which requires tissue information.

5. REFERENCES

[1] Ulaş Bağcı et al., "The role of intensity standardization in medical image registration," *Pattern Recognition Letters*, vol. 31, no. 4, pp. 315–323, 2010.

[2] Ingemur J Cox et al., "Dynamic histogram warping of image pairs for constant image brightness," in *ICIP*. IEEE, 1995, vol. 2, pp. 366–369.

[3] Pierre Hellier, "Consistent intensity correction of mr images," in *ICIP*. IEEE, 2003, vol. 1, pp. I–1109.

[4] László G Nyúl et al., "New variants of a method of mri scale standardization," *IEEE TMI*, vol. 19, no. 2, pp. 143–150, 2000.

[5] Lászlo G Nyu et al., "On standardizing the mr image intensity scale," *image*, vol. 1081, 1999.

[6] Mohak Shah et al., "Evaluating intensity normalization on mris of human brain with multiple sclerosis," *Medical image analysis*, vol. 15, no. 2, pp. 267–282, 2011.

[7] Florian Jäger et al., "A new method for mri intensity standardization with application to lesion detection in the brain," .

[8] Nicolas Robitaille et al., "Tissue-based mri intensity standardization: application to multicentric datasets," *Journal of Biomedical Imaging*, vol. 2012, pp. 4, 2012.

[9] Giorgio De Nunzio et al., "Robust intensity standardization in brain magnetic resonance images," *Journal of digital imaging*, pp. 1–11, 2015.

[10] Nobuyuki Otsu, "A threshold selection method from gray-level histograms," *Automatica*, vol. 11, no. 285-296, pp. 23–27, 1975.

[11] Solomon Kullback et al., "On information and sufficiency," *The annals of mathematical statistics*, 1951.

[12] London Imperial College, "Brain development dataset," `http://brain-development.org/`.

[13] Antoni Buades et al., "A review of image denoising algorithms, with a new one," *Multiscale Modeling & Simulation*, vol. 4, no. 2, pp. 490–530, 2005.

[14] John G Sled et al., "A nonparametric method for automatic correction of intensity nonuniformity in mri data," *IEEE TMI*, vol. 17, no. 1, pp. 87–97, 1998.

[15] Yongyue Zhang et al., "Segmentation of brain mr images through a hidden markov random field model and the expectation-maximization algorithm," *IEEE TMI*, vol. 20, no. 1, pp. 45–57, 2001.

[16] Stephen M Smith, "Fast robust automated brain extraction," *Human brain mapping*, vol. 17, no. 3, pp. 143–155, 2002.