Longitudinal Intensity Normalization in Multiple Sclerosis Patients

Yogesh Karpate, Olivier Commowick, Christian Barillot, Gilles Edan

Longitudinal Intensity Normalization in Multiple Sclerosis Patients

Yogesh Karpate, Olivier Commowick, Christian Barillot and Gilles Edan

VISAGES: INSERM U746 - CNRS UMR6074 - INRIA - Univ. of Rennes I, France

Abstract. In recent years, there have been many Multiple Sclerosis (MS) studies using longitudinal MR images to study and characterize the MS lesion patterns. The intensity of similar anatomical tissues in MR images is often different because of the variability of the acquisition process and different scanners. This paper proposes a novel methodology for a longitudinal lesion analysis based on intensity standardization to minimize the inter-scan intensity difference. The intensity normalization maps parameters obtained using a robust Gaussian Mixture Model (GMM) estimation not affected by the presence of MS lesions. Experimental results demonstrate that our technique accurately performs the task of intensity standardization. We show consequently how the same technique can improve the results of longitudinal MS lesion detection.

1 Introduction

Multiple Sclerosis (MS) is an acquired inflammatory, demyelinating disease which causes disabilities in young adults and it is very common in the northern hemisphere. Quantitative analysis of longitudinal Magnetic Resonance Images (MRI) of subject taken at different time points provides a time varying analysis of the brain tissues which may lead to the discovery of new biomarkers of disease evolution. In MS, White Matter (WM) lesions are also present in addition to healthy brain tissues. Lesions can remain stationary, change volume, or disappear in later time points depending upon the state of MS. Due to protocol variations in the scanners, following the evolution of tissue intensities in a patient, e.g. changing appearance of lesions, makes quantitative evaluation of lesions difficult. In order to alleviate this problem, intensity normalization is necessary.

Histogram matching is a widely used technique in intensity standardization. In their seminal work, Nyul et al. [7] proposed landmark based methods. It consists of matching the input image histogram landmarks onto standard histogram landmarks, obtained in a training phase, performing a linear interpolation of intensities between the positions. The technique in [7] uses percentile landmarks, which is simple yet powerful. Jager et al. [5] extended this principle to two or more jointly used MRI sequences (e.g., T1-w and T2-w), matching multidimensional joint histograms with nonlinear registration. With this method, no prior registration of the reference and normalized MR images is required. An algorithm proposed by Wang et al. [12] expands or shrinks a windowed part of the

input image histogram with a multiplicative factor, found by minimizing the bin-count difference between the source and moving images histograms. The window is used to include only voxels of interest and exclude the background. This makes the technique linear in the intensity range of interest. Other techniques use parametric models, such as the technique proposed by Hellier [4]. It models histogram of a reference image and of the standardized image with two GMMs and aligns their means through a polynomial correction function. Weisenfeld et.al. [13] have proposed to estimate a multiplicative correction field that alters the intensity statistics of an image or set of images to best match those of a model. In that paper, the Kullback-Leibler divergence between the source and moving images is minimized iteratively to estimate the parameters of a model, thus histograms are equalized. All these methods may be affected by the presence of white matter lesions.

We propose a longitudinal intensity normalization algorithm for multichannel MRI in the presence of MS lesions, which provides consistent and reliable longitudinal detections. The tissue intensities from multichannel MRI are modeled with parametric transform using a robust GMM estimation based on γ divergence, thereby keeping the lesions unaffected. The proposed technique is built on ideas similar to Hellier [4] but taking into account the presence of pathological tissues in the intensity transformation function. It provides a technique that (1) uses tissue-specific intensity information by modeling them using a robust GMM; (2) provides a consistent intensity normalization between longitudinal images. Subsequently, we demonstrate its crucial role for further lesion analysis.

This paper is organized as follows. The modeling and parameter estimation of multi-sequence MRI with γ divergence followed by intensity normalization are reviewed in Section 2. The details of experiments and their results on longitudinal MS patients are discussed in section 3.

2 Methodology

Given two MR images of a single MS patient at time instant t_1 and t_2, we seek to estimate a correction factor such that corresponding anatomical tissues adopt the same intensity profile. We model the image intensities of a healthy brain with a 3-class GMM, where each Gaussian represents one of the brain tissues White Matter (WM), Gray Matter(GM) and Cerebrospinal fluid (CSF). We consider the m MR sequences as a multidimensional image with n voxels. Each voxel i is represented as $\mathbf{x}_i = [x_{i1}...x_{im}]$. The probability of intensity $\mathbf{x}_i$ is calculated as follows:

$$f(\mathbf{x}_i|\theta) = \sum_{k=1}^{3} \pi_k \mathcal{N}(\mu_k, \Sigma_k) \tag{1}$$

where the mean μ_k and covariance Σ_k define the parameters $\mathcal{N}(\mu_k, \Sigma_k)$ of each Gaussian of the model along with their mixing proportions π_k merged into parameter θ. If the proportions were known, θ could be estimated through the Maximum Likelihood Estimator (MLE):

$$\widehat{\theta} = \underset{\theta}{\operatorname{argmax}}\ L(\theta) = \underset{\theta}{\operatorname{argmax}}\ \prod_{i=1}^{n} f(\mathbf{x}_i|\theta) \tag{2}$$

Where $\mathbf{x}_i$ are considered as i.i.d. samples. However, as π_k are unknown, an Expectation Maximization (EM) algorithm [3] is used to estimate the parameters.

2.1 γ-loss Function for the Normal Distribution

The parameter estimation with classic MLE for GMM can deviate from its true estimation in presence of outliers. In MS patients, such outliers may be of crucial importance as they may denote appearing or disappearing lesions. Notsu et al. [6] proposed a modification of the MLE in order to make it more robust to outliers. The basic idea is to maximize (2) in the form of γ divergence. We consider the γ-loss function for the Normal distribution with mean vector μ and covariance matrix Σ.

$$L_\gamma(\mu, \Sigma) = \left|\Sigma^{-\frac{\gamma}{2(1+\gamma)}}\right| \sum_{i=1}^{n} \exp\big(-\frac{\gamma}{2}(\mathbf{x}_i - \mu)^T \Sigma^{-1}(\mathbf{x}_i - \mu)\big) \tag{3}$$

Where $|.|$ indicates the determinant. The bounded influence function of an estimator is an indicator of robustness to outliers. The influence function for GMM with γ loss function is bounded whereas the one for regular GMM is unbounded. As γ grows larger, bounds become tighter. For a sufficiently large γ, ($\gamma \geq 0.1$), the estimating equation has little impact from outliers contaminated in the data set. Equation (3) can be casted to yield an EM style algorithm as follows.

Expectation Step. In the case of a GMM, the latent variables are the point-to-cluster assignments $k_i, i = 1, ..., n$, one for each of n data points. The auxiliary distribution $q(k_i|\mathbf{x}_i) = q_{ik}$ is a matrix with $n \times K$ entries. Each row of q_i can be thought of as a vector of soft assignments of the data points $\mathbf{x}_i$ to each of the Gaussian modes.

$$q_{ik} = \frac{\pi_k \exp\big(-\frac{\gamma}{2}(\mathbf{x}_i - \mu_k)^T \Sigma_k^{-1}(\mathbf{x}_i - \mu_k)\big)}{\sum_{l=1}^{K} \pi_l \exp\big(-\frac{\gamma}{2}(\mathbf{x}_i - \mu_l)^T \Sigma_l^{-1}(\mathbf{x}_i - \mu_l)\big)} \tag{4}$$

Maximization Step. The maximization step estimates the parameters of the Gaussian mixture components and the mixing proportions π_k, given the auxiliary distribution on the point-to-cluster assignments computed in the expectation step. The mean μ_k of a Gaussian mode is obtained as the mean of the data points assigned to it (accounting for the strength of the soft assignments). The

other quantities are obtained in a similar manner, yielding to:

$$\mu_k = \frac{\sum_{i=1}^{n} q_{ik}\mathbf{x}_i}{\sum_{i=1}^{n} q_{ik}} \tag{5}$$

$$\Sigma_k = (1+\gamma)\frac{\sum_{i=1}^{n} q_{ik}(\mathbf{x}_i - \mu_k)(\mathbf{x}_i - \mu_k)^T}{\sum_{i=1}^{n} q_{ik}} \tag{6}$$

$$\pi_k = \frac{\sum_{i=1}^{n} q_{ik}}{\sum_{i=1}^{n}\sum_{l=1}^{K} q_{il}} \tag{7}$$

2.2 Selection of Parameter γ

The estimation of power index γ plays a critical role in our approach, since γ affects the estimated parameters in presence of outliers. Notsu et al. [6] suggested the selection of γ as a model selection problem based on Akaike information criterion (AIC). Let K be the number of clusters, p be the total numbers of parameters of a model and (μ_k, Σ_k), $k = 1, .., K$ be the means and the covariance matrices of the clusters respectively. From (1), the AIC is defined as follows:

$$\mathbf{AIC}_\gamma = -2\sum_{i=1}^{n} \log f_\gamma(\mathbf{x}_i|\theta) + 2\left\{K\frac{p(p+3)}{2} + K - 1\right\} \tag{8}$$

The value of γ which minimizes AIC is used as the optimal γ. For various values of γ, equation (8) is evaluated in cross validation manner and the γ which results in minimum value is chosen for the experiment.

2.3 Intensity Correction

We obtain the means and covariances of tissues for the source and target images using the procedure mentioned above. We chose a linear correction function such that $g(\mathbf{x}) = \Sigma_i \beta_i \mathbf{x}_i$. The coefficients β_i are estimated to minimize the following cost function: $\Sigma_{l=1}^{l=n}(g(\mu_{source,k}) - \mu_{target,k})^2$. This function can be solved by linear regression. Using the results of the linear regression, the intensity profiles of the two images are normalized by mapping the intensity of the source image to the target image. The resulting correction function is smooth and interpolates the intensity correction.

3 Experiments and Results

3.1 Dataset and Preprocessing

Whole-brain MR images were acquired on 18 MS patients. T1-w MPRAGE, T2-w and FLAIR modalities were chosen for the experiment. Expert annotations of lesions were carried out by an expert radiologist on all MS patients. The volume size for T1-w MPRAGE and FLAIR is 256×256×160 and voxel size is 1×1×1 mm^3. For T2-w, the volume size is 256×256×44 and voxel size is 1×1×3

mm^3. All imaging experiments for this study were performed on a 3T Siemens Verio (VB17) scanner with a 32-channel head coil. MR images from each patient are de-noised [2], bias field corrected [11] and registered with respect to T1-MPRAGE volume [9, 1]. All the images are processed to extract intra-cranial region using **BET** (Brain Extraction Tool) [10].

We show the effect of longitudinal intensity normalization followed by detection on both normal tissues and lesions for 18 MS subjects, having 4 time-points each, approximately separated by a period of three to six months. The first time point is considered as the reference point to which the subsequent time points (moving ones) are aligned using intensity normalization. First, the parameters of reference and moving images are estimated using γ likelihood estimator as described in section 2.1. Secondly, voxels of moving image are aligned with respect to reference image using the procedure in section 2.3. Each patient and each time point $t = 2, ..., t_n$, are rigidly registered to the T1-w MPRAGE of first time instance. The obtained difference image is processed further to obtain a soft detection by using heuristic thresholding iteratively (1) by Otsu's threshold [8]; (2) erosion of image by one voxel. The detections from this difference image are compared with difference image of ground truth at corresponding time points.

3.2 Intensity Correction Evaluation

To evaluate the quality of intensity normalization, we compare the histograms of reference, moving and intensity normalized moving image using chi-squared distance given by $\chi^2_{x,y} = \frac{1}{2}\sum\frac{(x_i-y_i)^2}{x_i+y_i}$. Lower values of this distance indicate better alignment of intensities. Table 1 reports the chi-squared distance for various imaging sequences. Different methods are compared against the proposed one. We report the mean χ^2 distance for our method as 0.18($\pm$0.045), 0.28($\pm$0.037) 0.32($\pm$0.038) for T1-w MPRAGE, T2-w and FLAIR respectively, outperforming other state of the art methods.

Table 1. Chi-squared distance analysis for histogram matching

	Before Normalization	After Normalization		
Modality		Proposed	Hellier	Nyul
T1-w	0.56 ($\pm$0.03)	**0.18 ($\pm$0.045)**	0.35 ($\pm$0.029)	0.3 ($\pm$0.019)
T2-w	0.62 ($\pm$0.029)	**0.28 ($\pm$0.037)**	0.414 ($\pm$0.03)	0.315 ($\pm$0.042)
FLAIR	0.56 ($\pm$0.027)	**0.32 ($\pm$0.038)**	0.45 ($\pm$0.051)	0.39 ($\pm$0.045)

Figure 1 shows the intensity correction results for T1-w MPRAGE, T2-w and FLAIR images. Three time points and their corresponding MR modalities of a subject are shown before and after normalization. Each row represents the imaging modality and each column depicts the first time point, second time point, the absolute difference image without and with intensity normalization respectively. This figure demonstrates visually the ability of our approach to

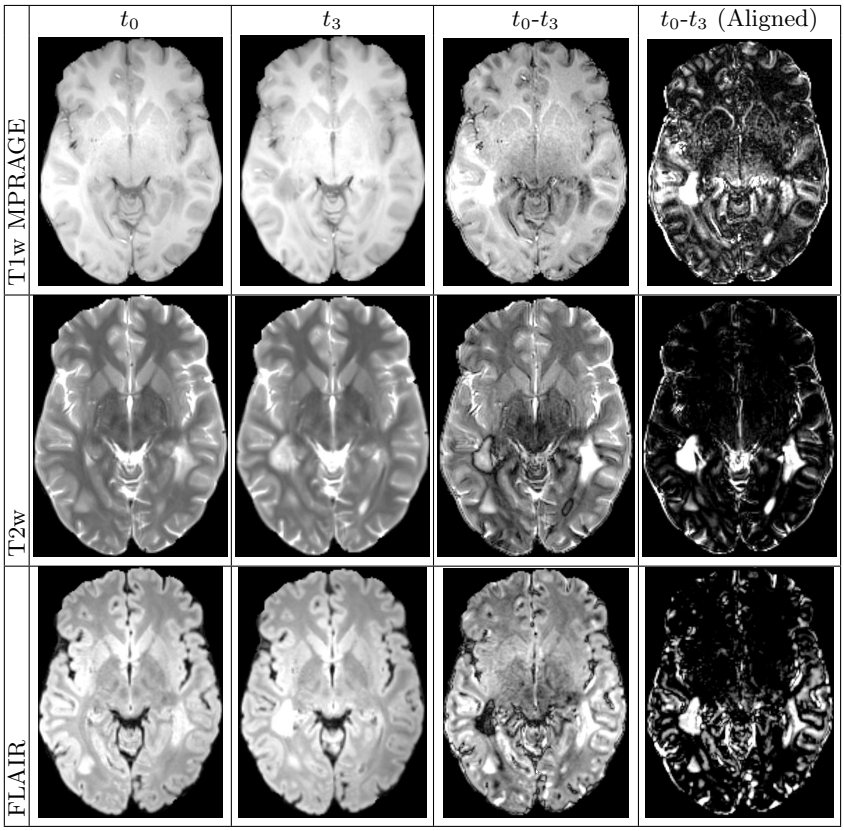

Fig. 1. Intensity correction

normalize intensities. As seen from the difference image of the first and second time points, intensity alignment reduces significantly the difference in intensities without affecting the lesion appearance. It will be easier to automatically detect evolving lesions on the images in the last column.

3.3 Longitudinal Lesion Detection

To show the quantitative improvement for identification of lesions, we report in Table 2 the precision (Positive Predicted Value) and recall (Sensitivity) of lesion detection averaged across the 18 patients for various overlap thresholds. The lesion is said to be detected if $\frac{R_c \cap R_{GT}}{R_{GT}} \geq \varphi$ where R_c, R_{GT} and φ are respectively the candidate region in the image, the ground truth and a threshold. Table 2 reports values of precision and recall for various thresholds. As from the figures, our approach outperforms other methods. We have a very high recall of 0.90 at $\varphi = 0.2$ and 0.82 even at $\varphi = 0.4$.

Figure 2 depicts the detected lesions for a representative image. The green label shows new lesions at t_3, orange shows stationary lesions which are also a

Table 2. Performance analysis for lesion detection.

	$\varphi = 0.2$		$\varphi = 0.3$		$\varphi = 0.4$	
Method	Precision	Recall	Precision	Recall	Precision	Recall
Nyul	0.63±0.01	0.60±0.02	0.61±0.04	0.67±0.02	0.58±0.03	0.64±0.03
Proposed	**0.73±0.04**	**0.90±0.05**	**0.68±0.03**	**0.85±0.04**	**0.63±0.03**	**0.82±0.01**
Hellier	0.65±0.02	0.74±0.03	0.64±0.06	0.68±0.04	0.62±0.03	0.59±0.05

part of t_1; blue shows false positive detections. We are able to accurately detect appearing and disappearing lesions thanks to the proposed method.

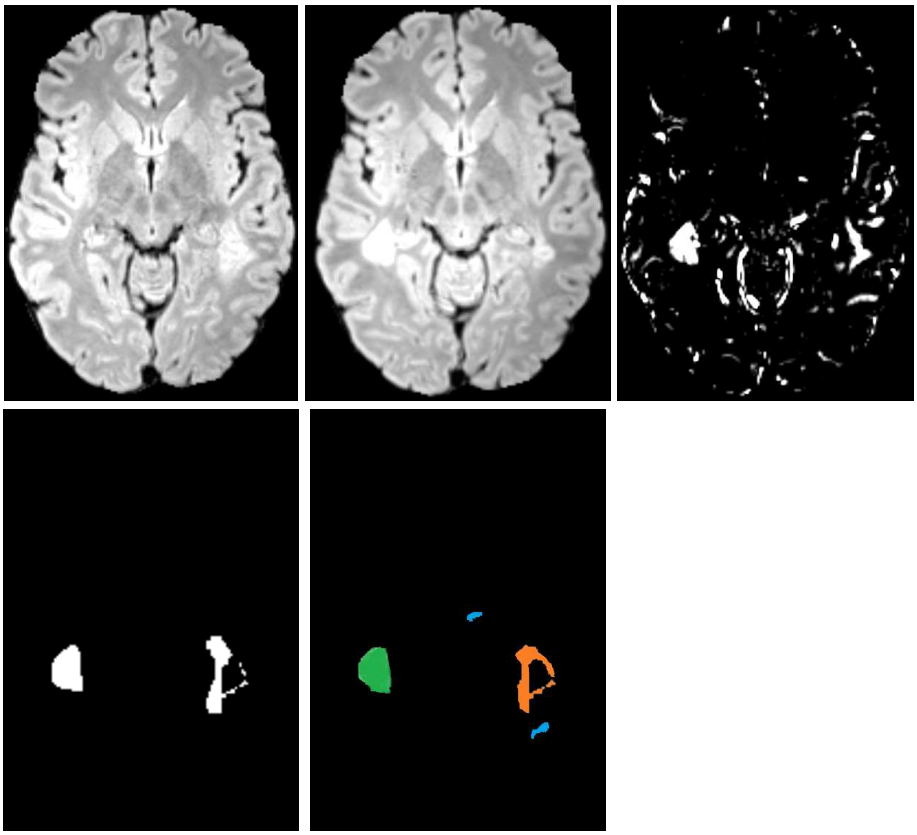

Fig. 2. Lesion detection examples. For top and bottom, from left to right: Slice of FLAIR for t_0, t_3, $|t_0 - t_{3(Normalized)}|$, ground truth and lesions detected by our algorithm.

4 Discussion and Conclusion

We proposed a new intensity normalization technique based on a robust GMM estimation with γ divergence. The efficacy of our method was evaluated through histogram matching distance method and longitudinal lesion detection. Compared to Nyul and Hellier method, our methodology is more suitable for longitudinal MS lesion analysis because of its ability to preserve the intensity variations caused by pathological changes. Our system relies heavily on a robust parametric modeling of tissue intensities based on γ divergence. The resulting system is both efficient and accurate, outperforming the state of the art methods. This performance suggests that it can provide valuable assistance in detecting the longitudinal MS lesions in clinical routine with high reliability. Our models are already capable of detecting highly variable lesion patterns, but we would like to move towards richer models. The framework described here allows for exploration of additional MR sequences with or without contrast agents. For example, one can consider infusing T1-w Gadolinium and DTI.

References

1. Commowick, O., Wiest-Daesslé, N., Prima, S.: Block-matching strategies for rigid registration of multimodal medical images. In: ISBI. pp. 700–703 (2012)
2. Coupé, P., Yger, P., Prima, S., Hellier, P., Kervrann, C., Barillot, C.: An optimized blockwise nonlocal means denoising filter for 3-D magnetic resonance images. IEEE TMI 27(4), 425–441 (2008)
3. Dempster, A., Laird, N., Rubin, D.: Maximum likelihood from incomplete data via the EM algorithm. J. of the Royal Statistical Society. Series B 39(1), 1–38 (1977)
4. Hellier, P.: Consistent intensity correction of MR images. In: ICIP. pp. 1109–1112 (2003)
5. Jäger, F., Nyul, L., Frericks, B., Wacker, F., Hornegger, J.: Whole body MRI intensity standardization. In: Bildverarbeitung für die Medizin. Springer (2007)
6. Notsu, A., Komori, O., Eguchi, S.: Spontaneous clustering via minimum gamma-divergence. Neural Computation 26(2), 421–448 (2014)
7. Nyul, L., Udupa, J., Zhang, X.: New variants of a method of MRI scale standardization. IEEE TMI 19(2), 143–150 (2000)
8. Otsu, N.: A threshold selection method from gray-level histograms. IEEE Trans. on Systems, Man and Cybernetics SMC-9(1), 62–66 (1979)
9. Ourselin, S., Roche, A., Prima, S., Ayache, N.: Block matching: A general framework to improve robustness of rigid registration of medical images. In: MICCAI. LNCS, vol. 1935, pp. 557–566 (2000)
10. Smith, S.: Fast robust automated brain extraction. HBM 17(3), 143–155 (2002)
11. Tustison, N., Avants, B., Cook, P., Zheng, Y., Egan, A., Yushkevich, P., Gee, J.: N4ITK: improved N3 bias correction. IEEE TMI 29(6), 1310–1320 (2010)
12. Wang, L., Lai, H., Barker, G., Miller, D., Tofts, P.: Correction for variations in MRI scanner sensitivity in brain studies with histogram matching. MRM 39(2), 322–327 (1998)
13. Weisenfeld, N., Warfield, S.: Normalization of joint image-intensity statistics in MRI using the Kullback-Leibler divergence. In: ISBI. pp. 101–104 (2004)