A Method for Standardizing MR Intensities between Slices and Volumes

Mark Schmidt
University of Alberta
Department of Computing Science
Edmonton, Alberta, T6G 2E8, CANADA
schmidt@cs.ualberta.ca

Abstract

The first section of this paper describes a method for correcting inter-slice intensity variations in MR images. The method does not rely on a tissue model or segmentation, and is not affected by the presence of abnormalities. The second section of this work extends this technique to the correction of inter-volume intensity variations in MR images of the brain, using a characterization of bi-lateral symmetry to confer robustness to the presence of large abnormalities such as tumors or edema.

1 Inter-Slice Intensity Variation Reduction

Due to gradient eddy currents and 'crosstalk' between slices in 'multislice' acquisition sequences, the two-dimensional slices acquired under some MRI acquisition protocols may have a constant slice-by-slice intensity offset [1]. It is noteworthy that these variations have different properties than the intensity inhomogeneity observed within slices, or typically observed across slices. As opposed to being slowly varying, these variations are characterized by sudden intensity changes in adjacent slices. A common result of inter-slice intensity variations is an interleaving between 'bright' slices and 'dark' slices [2], (the 'even-odd' effect). While most intensity inhomogeneity correction methods can correct for slowly varying intensity variations, most methods for intensity inhomogeneity reduction do not consider these sudden changes. This work, therefore, presents a simple method to reduce sudden intensity variations between adjacent slices.

In comparison to the estimation of slowly varying intensity inhomogeneities, correcting inter-slice intensity variations has received little attention in the medical imaging literature. One early attempt to correct this problem in order to improve segmentation was presented in [3]. This work presented a system for the segmentation of normal brains using Markov Random Fields, and presented two simple methods to re-estimate tissue parameters between slices (after patient-specific training on a single slice). One method thresholded pixels with high probabilities of containing a single tissue type, while the other used a least squares estimate of the change in tissue parameters. A similar approach was used in one of the only systems thus far to incorporate this step for tumor segmentation [4]. This system first used patient-specific training of a neural network classifier on a single slice. When segmenting an adjacent slice, this neural network was first used to classify all pixels in the adjacent slice. The locations of pixels that received the same label in both slices were then determined, and these pixels in the adjacent slice were used as a new training set for the neural network classifier used to classify the adjacent slice. Each of these approaches require not only a tissue model, but patient-specific training, making them unsuitable for use in automatic systems for detecting and segmenting large abnormalities.

One of the most impressive inter-slice intensity correction methods to date was presented in [1]. This work presented two methods to incorporate inter-slice variation correction within an EM segmentation framework. The first simply incorporated slice-by-slice constant intensity offsets into the inhomogeneity estimation, while the second method computed a two-dimensional inhomogeneity field in each slice and used these to produce a three-dimensional inhomogeneity field that allowed inter-slice intensity variations. The method used by the INSECT system for this step was presented in [5] to improve the segmentation of Multiple Sclerosis lesions. This method estimated a linear intensity mapping based on pixels at the same location in adjacent slices that were of the same tissue type. Unfortunately, despite the lack of patient-specific training, these methods each still require a tissue model (in each slice) that may be violated in data containing significant pathology.

A method free of a tissue model was presented in [6]. This method used a median filter to reduce noise, and pruned pixels from the intensity estimation by band thresholding of histogram, and removing pixels represent-

ing edges. The histogram was divided into bins and a parabola was fit to the heights of the 3 central bins, used to determine the intensity mapping. Although model-free, this method makes major assumptions about the distribution of the histogram, that may not be true in all modalities or in images with pathological data. In addition, this method ignores spatial information.

Inter-slice intensity variation correction can be addressed using the same techniques employed in Intensity Standardization, which will be discussed in the next section. However, most methods for Intensity Standardization employ a tissue model or a histogram matching method that will be sensitive to outliers. It was ultimately chosen not to use one of the existing histogram matching methods, since real data may have anisotropic pixels, where the tissue distributions can change significantly between slices. The methods in [5, 4] are more appealing since these methods use spatial information to determine appropriate pixels for use in estimation. However, these methods rely on a tissue model that could be inappropriate for data with significant pathology. Although the method of [6] is a histogram matching method, removing points from the estimation in a model-free way is appealing. We present in this section a simple method to identify good candidates for estimating the intensity between slices as in [5, 4], but without an explicit tissue model.

We will assume that the intensity mapping between slices can be described by a multiplicative scalar value w, a model commonly used [5, 1]. If we assume that the slices are exactly aligned such that each pixel in slice X corresponds to a pixel in slice Y of the same tissue type, then the scalar w could be estimated by solving the equation below (where X and Y are vectors of intensities and $X(i)$ has the same spatial location as $Y(i)$ within the image):

$$Xw = Y \tag{1}$$

However, since there will not be an exact mapping between tissue types at locations in adjacent slices, an exact value for w that solves this equation will not exist, and therefore the task becomes to estimate an appropriate value for w. One computationally efficient way to estimate a good value of w would be to calculate the value for w that minimizes the sum of squared errors between the elements of Xw and Y:

$$\min_w \sum_i (X(i)w - Y(i))^2 \tag{2}$$

The optimal value for w in this case can be determined by solving for w in the 'normal equations' [7] (we employ the matrix pseudoinverse):

$$w = (X'X)^{-1}X'Y \tag{3}$$

Unfortunately, this computation is sensitive to areas where different tissue types are not aligned, since these regions are given weight equal to that of pixels where tissue types are appropriately aligned in the adjacent slices. The value w thus simply minimizes the error between the intensities at corresponding locations in adjacent slices, irrespective of whether the intensities should be the same (possibly introducing additional inter-slice intensity variations). The objective must thus be modified to restrict the estimation of w to locations that actually *should* have the same intensity after the intensity transformation w is applied. This is difficult without the use of a tissue model or a segmentation of the image. However, an alternate approach to identifying tissues or performing a segmentation is to weight the errors based on the importance of having a small error between each corresponding location $(X(i), Y(i))$. Given a weighting of the importance for each pixel to have the same intensity between adjacent slices $R(i)$, the calculation of w would focus on computing a value that minimizes the squared error for areas that are likely to be aligned, while reducing the effect of areas where tissues are likely misaligned. Given $R(i)$ for each i, the least squares solution can be modified to use this weight by performing element-wise multiplication of both the vectors X and Y with R [8]. This scaling of both vectors modifies the error function to be proportional to the values in R (using $.*$ to denote element-wise multiplication):

$$\min_w \sum_i ((X(i).*R(i))w - Y(i).*R(i))^2 \tag{4}$$

The value w that minimizes the above relevance-weighted loss function can be computed as before:

$$w = ((X.*R)'(X.*R))^{-1}(X.*R)'(Y.*R) \tag{5}$$

If the image was segmented into anatomically meaningful regions, computing $R(i)$ would be trivial, it would be 1 at locations where the same tissue type is present in both slices and 0 when the tissue types differ. Without a segmentation, this can be approximated. An intuitive approximation would be to weight pixels based on a measure of similarity between their regional intensity distributions. A method robust to intensity-scaling to perform this approximation is to compute the (regional) joint entropy of the intensity values. The (Shannon) joint entropy is defined as follows [9]:

$$H(A_1, A_2) = -\sum_{i \epsilon A_1, j \epsilon A_2} p(i,j) \log p(i,j) \tag{6}$$

The value $p(i,j)$ represents the likelihood that intensity i in one slice will be at the same location as intensity j in

Figure 1. Toy example of weighting for inter-slice intensity variation reduction. The goal is to transform slice 1 so that its intensities match those of slice 2. Top row, left to right: Original slice 1, original slice 2, slice 1 scaled from unweighted linear regression. Since the objects in the two slices differ, the unweighted linear regression did not estimate a good transformation. Middle row: elements of the weighting. Left to right: Regional joint entropy, absolute difference, joint foreground pixels. Bottom left: Combined weighting (darker regions receive less weight). Note that the weighting focuses the estimation on regions that should have the same intensity. Bottom right: Linear regression using the combined weighting estimates a near-optimal linear scaling.

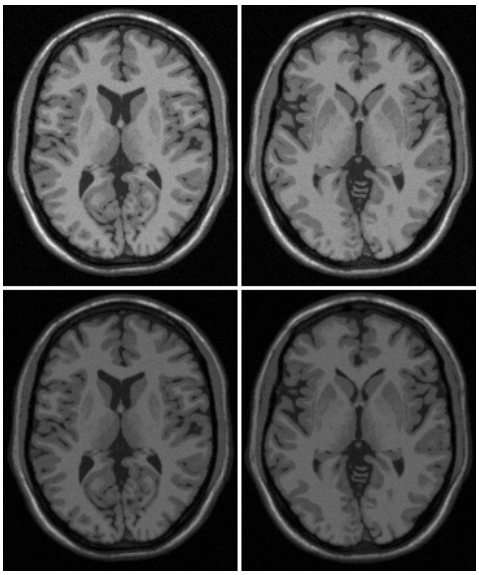

Figure 2. Inter-slice intensity variation reduction with simulated MR images [10, 11, 12, 13, 14] and an applied linear intensity offset. Top left: Simulated slice 1. Top right: Simulated slice 2 ($10mm$ away from slice 1). Bottom left: Slice 1 with an applied linear offset. Bottom right: Slice 2 transformed using the weighted linear regression between slice 2 and the darkened slice 1. The method successfully recovered a scale factor very close to the one applied.

the adjacent slice, based on an image region. We use a 25 pixel square window to compute the values $p(i, j)$ for a region, and divide the intensities into 10 equally spaced bins to make this computation. The frequencies of the 25 intensity combinations in the resulting 100 bins are used for the $p(i, j)$ values (smoothing these estimates could give a less biased estimate). The joint entropy computed over these values of $p(i, j)$ has several appealing properties. In addition to being insensitive to scaling of the intensities, it is lowest when the pixel region is homogeneous in both slices, will be higher if the intensities are not homogeneous in both slices but are spatially correlated, and will be highest when the intensities are not spatially correlated. After a sign reversal and normalization to the range $[0, 1]$, the regional joint entropy of the image regions could be used as values for $R(i)$, that would encourage regions that are more homogeneous and correlated between the slices to receive more weight in the estimation of w than heterogeneous and uncorrelated regions.

Joint entropy provides a convenient measure for the degree of spatial correlation of intensities, which is not dependent on the values of the intensities as in many correlation measures. However, the values of the intensities in the same regions in adjacent slices should also be considered, since pixels of very different intensity values should receive decreased weight in the estimation, even if they are both located in relatively homogeneous regions. Thus, in addition to assessing the spatial correlation of regional intensities, higher weight should be assigned to areas that have similar intensity values before transformation, and the weight should be dampened in areas where intensity values are different. The most obvious measure of the intensity similarity between two pixels is the absolute value of their intensity difference. This measure is computed for each set of corresponding pixels between the slices, and normalized to be

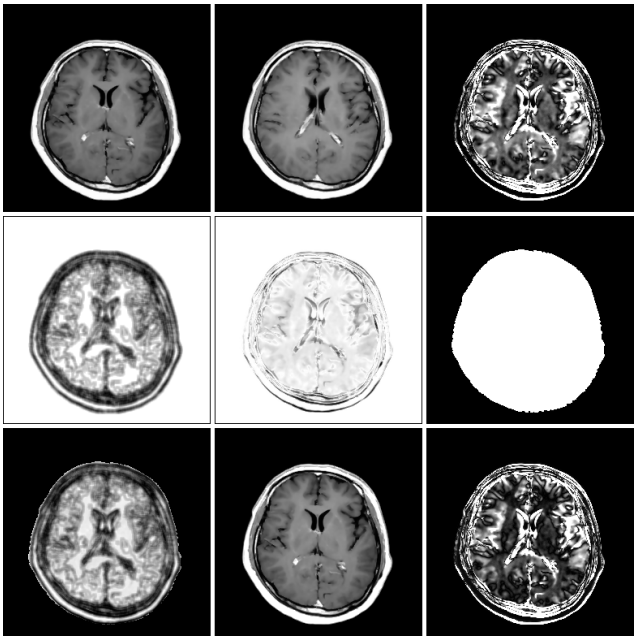

Figure 3. Inter-slice intensity variation reduction for real data. Top, left to right: Slice with an unknown intensity offset, adjacent slice, difference between the adjacent slices (multiplied by 10). Middle row, left to right: Entropy weighting, difference weighting, joint foreground pixels. Bottom, left to right: Combined weighting, slice 1 after transformation, difference between slice 1 after transformation and slice 2. The effect has not been completely removed, but has been noticeably reduced.

in the range $[0, 1]$ (after a sign reversal). Values for $R(i)$ that reflect both spatial correlation and intensity difference can be computed by multiplying these two measures. As a further refinement to this measure, the threshold selection algorithm from [15] (and morphological filling of holes) is used to distinguish foreground (air) pixels from background (head) pixels, and $R(i)$ is set to zero for pixels representing background areas in either slice (since they are not relevant to this calculation). Thus, each value in $R(i)$ is computed as follows (where N_1 and N_2 are normalizing constants, $H(X(i), Y(i))$ is the regional joint entropy centered at i, and P_{fore} is an indicator function that returns 1 if the pixel is part of the foreground and 0 otherwise):

$$r(i) = \frac{(N_1 - H(X(i), Y(i)))}{N_1} \frac{(N_2 - |X(i) - Y(i)|)}{N_2} P_{fore}(X(i)) P_{fore}(Y(i)) \quad (7)$$

Figure 1 demonstrates the advantage of weighting the estimation on a toy example. Figure 2 shows an example

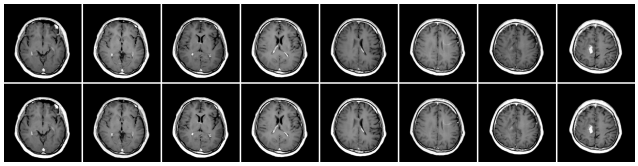

Figure 4. Inter-slice intensity variation reduction for an image series. Top, original image series (the even slices are noticeably brighter than the odd slices). Bottom, the same series after reduction of inter-slice intensity variations. The variations have not been completely corrected, but their effects have been reduced.

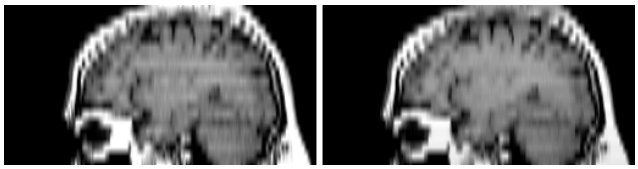

Figure 5. Inter-slice intensity variation reduction from a different angle. Left: Sagittal view of the original slices from the series in the figure above. Right: The same view after inter-slice intensity variation reduction. The brain area in the input image clearly shows the 'even-odd' offset effect, which has been noticeably reduced in the output image.

of the estimation being applied to simulated MR images to correct for an applied linear offset, while Figure 3 presents results on real data.

In our implementation, the weighted least squares estimation computes the linear mapping to the median slice in the sequence from each of the two adjacent slices. The implemented algorithm then proceeds to transform these slices, and then estimates the intensity mappings of their adjacent slices, continuing until all slices have been transformed. Figure 4 shows the results of this process, while Figure 5 shows the same results viewed from an orthogonal angle.

Future implementations could expand on this method by computing a non-linear intensity mapping between the slices. Our experiments with non-linear mappings showed that they were difficult to work with, since non-linear transformations tended to reduce image contrast. This process would thus need to be subject to more advanced regularization measures.

2 Intensity Standardization

Intensity standardization is a vital step that allows the intensity values between volumes of the same or different individuals to approximate an anatomical meaning. This subject has not received as significant of a focus in the literature as intensity inhomogeneity correction, but research effort in this direction has grown in the past several years. This is primarily due to the fact that it can remove the need for patient specific training or the reliance on tissue models in segmentation, which may not be available for some tasks or for some areas of the body. This section will survey the literature relating to intensity standardization, before presenting our approach. Although EM-based methods that use spatial priors are an effective method of intensity standardization, they will not be discussed here, since these methods can be sensitive to areas of abnormality [16], a case that we are interested in addressing.

The intensity standardization method used by the INSECT system [17] was (briefly) outlined in [5], in the context of improving Multiple Sclerosis lesions segmentation, and was discussed earlier in this document in the context of inter-slice intensity variation reduction. This method estimates a linear coefficient between the image and template based on the distribution of 'local correction' factors. Another study focusing on intensity standardization for Multiple Sclerosis lesion segmentation was presented in [18], that compared four methods of intensity standardization. The first method simply normalized based on the ratio of the mean intensities between images. The second method scaled intensities linearly based on the average white matter intensity (with patient-specific training). The third method computed a global scale factor using a "machine parameter describing coil loading according to reciprocity theorem", computing a transformation based on the voltage needed to produce a particular 'nutation angle' (that was calibrated for the particular scanner that was used). The final method examined was a simple histogram matching technique based on a non-linear minimization of squared error applied to 'binned' histogram data, after the removal of air pixels outside the head (this outperformed the other three). In [19], another histogram matching method was presented (later made more robust in [20]), that computed a piecewise intensity scaling based on 'landmarks' in the histogram. Similar to previous works on intensity standardization, this study also demonstrated that intensity standardization could aid in the segmentation of Multiple Sclerosis lesions. This method was later used in a study that evaluated the effects of inhomogeneity correction and intensity standardization [21], finding that these steps complemented each other, but that inhomogeneity correction should be done prior to intensity standardization. Another method of intensity standardization was presented in [22], that normalized white matter intensities using histogram derivatives. An intensity standardization method that was used as a preprocessing step in a tumor segmentation system was presented in [23]. This method thresholded background pixels, and used the mean and variance of foreground pixels to standardize intensities. A similar method was used in [24], comparing it to no standardization, scaling based on the intensity maximum, and scaling based on the intensity mean.

The methods discussed above are relatively simple and straightforward. Each method (with the exception of [5]) uses a histogram matching method that assumes either a simple distribution or at least a close correspondence between histograms. These assumptions can be valid for controlled situations, where the protocols and equipment used are relatively similar, and only minor differences exist between the image to be standardized and the template histogram. However, in practice this may not be the case, as histograms can take forms that are not well characterized by simple distributions, in addition to potential differences in the shapes of the input and template image histograms. This relates to the idea that a term like 'T1-weighted' does not have a correspondence with absolute intensity values, since there are a multitude of different ways of generating a T1-weighted image, and the resulting images can have different types of histograms. Furthermore, one 'T1-weighted' imaging method may be measuring a slightly different signal than another, meaning that tissues could appear with different intensity properties on the image, altering the histogram.

A more sophisticated recent method was presented in [25]. This method used the Kullback-Leibler (KL) divergence as a measure of relative entropy between an image intensity distribution and the template intensity distribution. This method computed an inhomogeneity field that minimized this entropy measure, and thus simultaneously corrected for intensity inhomogeneity and performed intensity standardization. Relative entropy confers a degree of robustness to the histogram matching, but even this powerful method fundamentally relies on a histogram matching scheme and ignores potentially relevant spatial information. Without the use of spatial information to 'ground' the matching by using the image-specific characteristics of tissues, standardizing the histograms does not necessarily guarantee a standardization of the intensities of the different tissue types. The EM-based approaches (that use spatial priors) can perform a much more sophisticated intensity standardization, since the added spatial information in the form of priors allows individual tissue types to be well characterized. By using spatial information to locate and characterize the different tissue types, the standardization method is inherently standardizing the intensities based on actual tissue characteristics in the image modalities, rather than simply aligning elements of the histograms. Further-

more, we are not aware of any existing methods that incorporate a means to reduce the effects of tumors and edema pixels (that are not present in the template image) on the estimation of the standardization parameters without the use of a tissue model. Thus, for this implementation, a simple method of intensity standardization was developed that is related to the proposed approach for inter-slice intensity variation reduction discussed earlier. The advantages of using this method are that it uses spatial information to ensure that similar tissue types are being matched, and that it uses symmetry to reduce the effects of tumors and edema on the estimation.

Our method for inter-slice intensity variation reduction uses spatial information between adjacent slices to estimate a linear mapping between the intensities of adjacent slices, but used simple measures to weight the contribution of each corresponding pixel location to this estimation. For intensity standardization, the problems that complicate the direct application of this approach are determining the corresponding locations between the input image and the template image, and accounting for outliers (tumor, edema, and areas of mis-registration) that will interfere in the estimation. Determining the corresponding locations between the input image and the template was trivial for inter-slice correction, since we assumed that adjacent slices would in general have similar tissues at identical image locations. This is not trivial for intensity standardization. For this stage, we used the non-linear regularized registration algorithm implemented in [26], and described in [27, 28, 29]. After this alignment, we assume that locations in the input image and the template will have approximately similar tissues.

In inter-slice intensity correction, the contribution of each pixel pair was weighted in the parameter estimation based on a measure of regional spatial correlation and the absolute intensity difference, which made the technique robust to areas where the same tissue type was not aligned. Since the input image will not be exactly aligned with the template image in the case of intensity standardization, these weights can also be used to make the intensity standardization more robust. However, intensity standardization is complicated by the presence of tumors and edema, areas that may be homogeneous and similar in intensity to the corresponding region in the template, but that should not significantly influence the estimation. To account for this, we use a measure of regional symmetry as an additional factor in computing the weights used in the regression. The motivation behind this is that regions containing tumor and edema will typically be asymmetric [30, 31]. Thus, giving less weight to asymmetric regions reduces the influence that abnormalities will have on the estimation.

A simple measure of symmetry is used, since the images have been non-linearly warped to the template where the line of symmetry is known. The first step in computing symmetry is computing the absolute intensity difference between each pixel and the corresponding pixel on the opposite side of the known line of symmetry. Since this estimation is noisy and only reflects pixel-level symmetry, the second step is to smooth this difference image with a 5 by 5 Gaussian kernel filter (the standard deviation is set to 1.25), resulting in a smoothly varying regional characterization of symmetry. Although symmetry is clearly insufficient to distinguish normal from abnormal tissues since normal areas may also be asymmetric, this weighting is included to decrease the weight of potentially bad areas from which to estimate the mapping, and thus it is not important if a small number of tumor pixels are symmetric or if a normal area is asymmetric.

The final factor that is considered in the weighting of pixels for the intensity standardization parameter estimation is the spatial prior 'brain mask' probability in the template's coordinate system (provided by the SPM2 software [26]). This additional weight allows pixels that have a high probability of being part of the brain area to receive more weight than those that are unlikely to be part of the brain area. This additional weight ensures that the estimation focuses on areas within the brain, rather than standardizing the intensities of structures outside the brain area, that are not as relevant to the eventual segmentation task.

The weighted linear regression is performed between the image and the template in each modality. The different weights used are the regional joint entropy, the absolute difference in pixel intensities, the regional symmetry measured in each modality, and the brain mask prior probability. These are each normalized to be in the range [0,1], and the final weight is computed by multiplying each of the weights together (assuming independence). The values for $R(i)$ are thus computed as follows (with $S(X(i))$ denoting the measure of symmetry at pixel i in image X, $P_{brain}(i)$ being the spatial prior probability that the pixel is part of the brain, and $H(X(i), Y(i))$ defined as before):

$$r(i) = \frac{(N_1 - H(X(i), Y(i)))}{N_1} \frac{(N_2 - |X(i) - Y(i)|)}{N_2} \frac{(N_3 - S(X(i)))}{N_3} P_{bra} \tag{8}$$

This method was implemented in Matlab [32], and is applied to each slice rather than computing a global factor to ease computational costs. The results of applying this technique to toy data and data with a synthetic tumor to recover a known intensity offset are shown in Figures 6 and 7, respectively. The application of this technique to real data (from different sites) to standardize the intensities between images is demonstrated in Figure 8.

There are several methods that could be explored to improve this step in future implementations. Different loss functions could be examined, since loss functions such as the absolute error and the Huber loss are more robust to out-

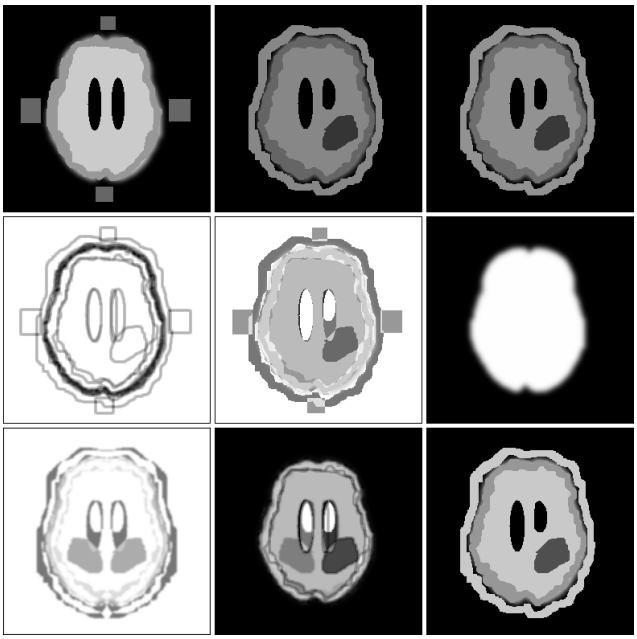

Figure 6. Intensity Standardization of a toy volume. Top left: Toy template image (consisting of gray matter, white matter, CSF, and fiducial markers. Top middle: Toy image to be standardized, that is slightly different anatomically, has fat visible outside of the skull, a large tumor, and no fiducial markers. Top right: The (poor) results obtained by unweighted linear regression. Middle row: Different elements of the pixel weighting. Left to right: Regional joint entropy, absolute difference, and brain area prior probability. The entropy and absolute difference have the same effect as before, but the brain probability allows restriction of the estimation to the brain area, rather than all foreground pixels. Bottom left: Symmetry weighting (note the low weight assigned to the tumor). Bottom middle: Combined weighting, indicating the estimation will place the largest weight on common gray matter, white matter, and csf regions. Bottom right: The results of weighted linear regression with the combined weighting for intensity standardization.

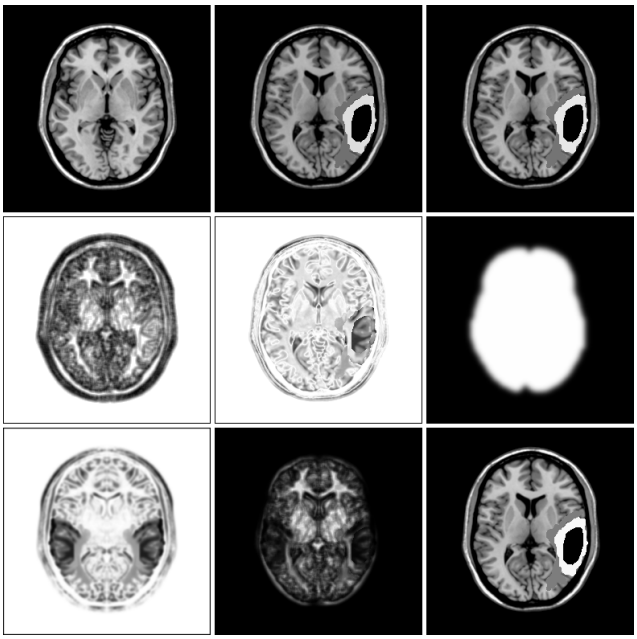

Figure 7. Intensity Standardization with a synthetic tumor to recover a known intensity offset. Top, left to right: Template image, image to be standardized with a synthetic tumor and an applied intensity offset, results of unweighted linear regression. Middle, left to right: Regional joint entropy weighting, absolute difference weighting, spatial brain prior weighting. Bottom, left to right: Symmetry weighting, combined weighting, result of weighted linear regression for intensity standardization. Note that the weighting makes the estimation primarily based on shared white matter regions, and reduces the tumor area's effect on the estimation.

liers than the squared error measure used here [33], though at a higher computational expense. In general, we found that non-linear transformations could further reduce the average error between the images, but this came at the cost of reduced contrast in the images. This occurred even when using a simple additive factor in addition to the linear scale factor. Future work could further explore non-linear methods that incorporate regularization to allow non-linear intensity standardization that is constrained to preserve image contrast. Although we have purposely avoided a tissue model in our inter-slice correction method, this may be a step that could benefit from a tissue model, especially if the technique will be applied for large data sets where intensity standardization will be a larger problem. One direction to explore with respect to this idea could be to use a method similar to the tissue estimation performed in [34], that used spatial prior probabilities for gray matter, white matter, and CSF to build a tissue model, but used an outlier detection scheme to make the estimation more robust. The weighting methods discussed in this section, and symmetry in particular, could be incorporated into an approach similar to this strategy to potentially achieve more effective intensity standardization.

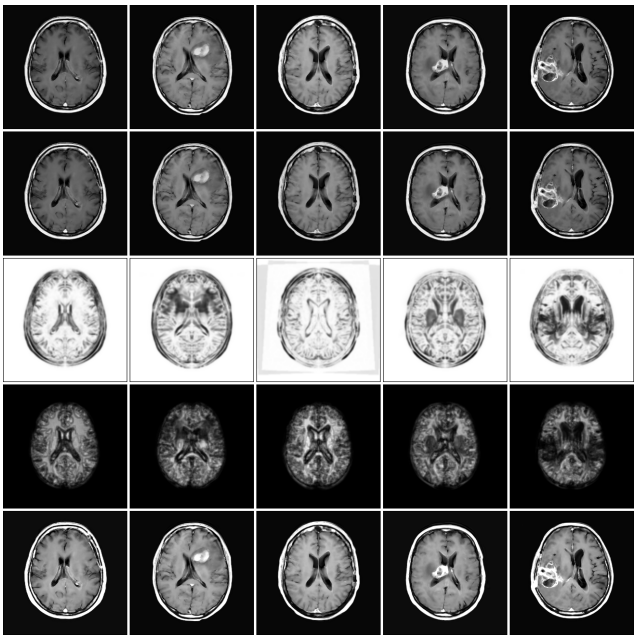

Figure 8. Intensity Standardization of real data. Top row: T1-weighted images from 5 patients. Second row: Intensity Standardization based on unweighted linear regression. Third row: Symmetry weighting based on T1-weighted and (coregistered) T2-weighted image. The three abnormal regions have clearly had their weight reduced. Fourth row: Combined weighting. The estimation for most of the images is primarily based on white matter regions, although some images also have high weights assigned to csf and gray matter regions. Bottom row: The results of Intensity Standardization. It is obvious that the differences in intensity between images have been significantly reduced.

References

[1] K.V. Leemput, F. Mase, D. Vandermeulen, and P. Suentens. Automated model-based bias field correction of mr images of the brain. *IEEE Transactions on Medical Imaging*, 18(10):885–896, October 1999.

[2] A. Simmons, P.S. Tofts, G.J. Barker, and S.R. Arridge. Sources of intensity nonuniformity in spin echo images at 1.5 t. *Magn Reson Med*, 32(1):121–128, Jul 1994.

[3] H.S. Choi, D.R. Haynor, and Y. Kim. Partial volume tissue classification of multichannel magnetic resonance images-a mixel model. *IEEE Transactions on Medical Imaging*, 10(3):395–407, Sep 1991.

[4] R.J. Maciunas M. Ozkan, B.M. Dawant. Neural-network-based segmentation of multi-modal medical images: a comparative and prospective study. *IEEE Transactions on Medical Imaging*, 12(3):534–544, Sep 1993.

[5] A.P. Zijdenbos, B.M. Dawant, and R.A. Margolin. Intensity correction and its effect on measurement variability in mri. In *Computer Assisted Radiology*, Jun 1995.

[6] E. Vokurka, N. Thacker, and A. Jackson. A fast model independent method for automatic correction of intensity nonuniformity in mri data. *JMRI*, 10(4):550–562, 1999.

[7] J. Shawe-Taylor and N. Cristianini. *Kernel Methods for Pattern Analysis*. Cambridge University Press, 2004.

[8] C. Moler. Numerical computing with matlab. http://www.mathworks.com/moler/, 2002.

[9] J.P. Pluim, J.B.A. Maintz, and M.A. Viergever. Mutual-information-based registration of medical images: a survey. *IEEE Transactions on Medical Imaging*, 22(8):986–1004, Aug 2003.

[10] Brainweb: a www interface to a simulated brain database (sbd) and custom mri simulations, http://www.bic.mni.mcgill.ca/brainweb/, Online.

[11] C.A. Cocosco, V. Kollokian, R.K.-S. Kwan, and A.C. Evans. Brainweb: Online interface to a 3d mri simulated brain database. *NeuroImage*, 5(S425), 1997.

[12] R.K.-S. Kwan, A.C. Evans, and G.B. Pike. Mri simulation-based evaluation of image-processing and classification methods. *IEEE Transactions on Medical Imaging*, 18(11):1085–1097, Nov 1999.

[13] R.K.-S. Kwan, A.C. Evans, and G.B. Pike. An extensible mri simulator for post-processing evaluation. *Lecture Notes in Computer Science*, 1131(11):135–140, 1996.

[14] D.L. Collins, A.P. Zijdenbos, V. Kollokian, J.G. Sled, N.J. Kabani, C.J. Holmes, and A.C. Evans. Design and construction of a realistic digital brain phantom. *IEEE Transactions on Medical Imaging*, 17(3):463–468, Jun 1998.

[15] N. Otsu. A threshold selection method from gray-level historgrams. *IEEE Trans. Systems, Man adn Cybernetics*, 9(1):62–66, 1979.

[16] J.G. Sled. *A nonparametric method for automatic correction of intensity nonuniformity in MRI data*. PhD thesis, McGill University, 1997.

[17] A.P. Zijdenbos, R. Forghani, and A.C. Evans. Automatic quantification of MS lesions in 3D MRI brain data sets: Validation of INSECT. *Medical Image Computing and Computer-Assisted Intervention*, pages 439–448, 1998.

[18] L. Wang, H.M. Lai, G.J. Barker, D.H. Miller, and P.S. Tofts. Correction for variations in mri scanner sensitivity in brain studies with histogram matching. *Magn Reson Med*, 39(2):322–327, Feb 1998.

[19] L.G. Nyul and J.K. Udupa. On standardizing the mr image intensity scale. *Magn Reson Med*, 42(6):1072–1081, Dec 1999.

[20] L.G. Nyul, J.K. Udupa, and Xuan Zhang. New variants of a method of mri scale standardization. *IEEE Transactions on Medical Imaging*, 19(2):143–150, Feb 2000.

[21] A. Madabhushi and J.K. Udupa. Evaluating intensity standardization and inhomogeneity correction in magnetic resonance images. In *IEEE 28th Annual Northeast Bioengineering Conference*, pages 137–138, April 2002.

[22] J.D. Christenson. Normalization of brain magnetic resonance images using histogram even-order derivative analysis. *Magn Reson Imaging*, 21(7):817–820, Sep 2003.

[23] S. Shen, W.A. Sandham, and M.H. Granat. Preprocessing and segmentation of brain magnetic resonance images. In *4th International IEEE EMBS Specific Topic Conference on Information Technology Applications in Biomedicine*, pages 149–152, Apr 2003.

[24] G. Collowet, M. Strzelecki, and F. Mariette. Influence of mri acquisition protocols and image intensity normalization methods on texture classification. *Magn Reson Imaging*, 22(1):81–91, Jan 2004.

[25] N.I. Weisenfeld and S.K. Warfield. Normalization of joint image-intensity statistics in mri using the kullback-leibler divergence. In *IEEE Interational Symposium on Biomedical Imaging*, 2004.

[26] Statistical parametric mapping, http://www.fil.ion.bpmf.ac.uk/spm/, Online.

[27] K.J. Friston, J. Ashburner, C.D. Fristh, J.B. Poline, J.D. Heather, and R.S.J. Frackowiak. Spatial registration and normalization of images. *Human Brain Mapping*, 2:165–189, 1995.

[28] J. Ashburner, P. Neelin, D.L. Collins, A.C. Evans, and K.J. Friston. Incorporating prior knowledge into image registration. *NeuroImage*, 6(4):344–352, November 1997.

[29] J. Ashburner and K.J. Friston. Nonlinear spatial normalization using basis functions. *Human Brain Mapping*, 7(4):254–266, 1999.

[30] D.T. Gering. Diagonalized nearest neighbor pattern matching for brain tumor segmentation. *R.E. Ellis, T.M. Peters (eds), Medical Image Computing and Computer-Assisted Intervention*, 2003.

[31] S. Joshi, P. Lorenzen, G. Gerig, and E. Bullitt. Structural and radiometric asymmetry in brain images. *Med Image Anal.*, 7(2):155–70, Jun 2003.

[32] Matlab - the language of technical computing, http://www.mathworks.com/products/matlab/, Online.

[33] T.J. Hastie, R.J. Tibshirani, and J. Friedman. *Elements of Statistical Learning: data mining, inference and prediction.* Springer-Verlag, 2001.

[34] M. Prastawa, E. Bullitt, S. Ho, and G. Gerig. A brain tumor segmentation framework based on outlier detection. *Medical Image Analysis*, 8(3):275–283, September 2004.