Interactive Non-Uniformity Correction and Intensity Standardization of MR Images

Yubing Tong, Jayaram K. Udupa, Dewey Odhner, Shobhit Sharma, Drew A. Torigian

Medical Image Processing Group, 423 Guardian Drive, 4th Floor Blockley Hall, Philadelphia PA 19104 [2]Department of Radiology, University of Pennsylvania, Philadelphia, PA 19104.

ABSTRACT

Image *non-uniformity* and *intensity non-standardness* are two major hurdles encountered in human and computer interpretation and analysis of magnetic resonance (MR) images. Automated methods for image non-uniformity correction (NC) and intensity standardization (IS) may fail because solutions for them require identifying regions representing the same tissue type for several different tissues, and the automatic strategies, irrespective of the approach, may fail in this task. This paper presents interactive strategies to overcome this problem: interactive NC and interactive IS. The methods require sample tissue regions to be specified for several different types of tissues. Interactive NC estimates the degree of non-uniformity at each voxel in a given image, builds a global function for non-uniformity correction, and then corrects the image to improve quality. Interactive IS includes two steps: a calibration step and a transformation step. In the first step, tissue intensity signatures of each tissue from a few subjects are utilized to set up key landmarks in a standardized intensity space. In the second step, a piecewise linear intensity mapping function is built between the same tissue signatures derived from the given image and those in the standardized intensity space to transform the intensity of the given image into standardized intensity. Preliminary results on abdominal T1-weighted and T2-weighted MR images of 20 subjects show that interactive NC and IS are feasible and can significantly improve image quality over automatic methods. Interactive IS for MR images combined with interactive NC can substantially improve numeric characterization of tissues.

Keywords: Non-uniformity correction, intensity standardization, magnetic resonance (MR) image analysis

1. INTRODUCTION

Image *non-uniformity* and *intensity non-standardness* are two major hurdles encountered in human and computer interpretation and analysis of magnetic resonance (MR) images. The former phenomenon is caused by imperfections in the imaging device, specifically the magnetic field, and has been studied extensively, where many solutions have been proposed in the literature [1-3]. Generally, the intensity non-uniformity component is assumed to be multiplicative or additive, i.e., the component is multiplicative or additive to an ideal image [3]. Most frequently, the multiplicative model has been used as it is consistent with the inhomogeneous sensitivity of the receiver coil of the magnetic resonance imaging (MRI) scanner. The N3 and N4 algorithms [1, 2] are two popular correction approaches which have adopted the multiplicative model.

Although to a far lesser extent, solutions have also been presented for the second problem of intensity non-standardness [8, 9]. This phenomenon causes the lack of a tissue specific numeric intensity meaning, even within the same MRI protocol, for the same body region, for images obtained on the same scanner, and for the same patient. It is also known that all non-uniformity correction methods introduce their own non-standardness as part of the correction process itself [12]. The interplay between the two processes, proper ways of handling the phenomena in combination, and the manner in which image segmentation performance improves substantially when these problems are addressed properly have also been demonstrated [12, 10, 4-7].

However, in many practical situations, automated methods fail to yield acceptable solutions for both problems. This is mainly because solutions for both problems require identifying regions representing the same tissue type for several different tissues, and the automatic strategies, irrespective of the approach, may fail in this task. The goal of this paper is to present interactive strategies to overcome this problem in both phenomena wherein the required high level knowledge is provided by an operator to improve image quality substantially in such situations. We describe novel interactive non-uniformity correction and intensity standardization approaches, and present early results on their performance and

improvements in the image based on MR images of the abdomen. For further reference, we will name these approaches *i*NC and *i*IS, respectively (for *interactive non-uniformity correction* and *interactive intensity standardization*).

2. MATERIALS & METHODS

Image Data: This retrospective study was conducted following approval from the Institutional Review Board at the Hospital of the University of Pennsylvania along with a Health Insurance Portability and Accountability Act waiver. Abdominal MR images of acceptable quality from 20 subjects (between 26 - 85 years) with normal appearing livers as per a board-certified radiologist's (DAT) review were utilized for illustrating the new *i*NC and *i*IS approaches. All MR images had been previously acquired on 1.5T MRI scanners. In-phase T1-weighted gradient recalled echo (TR = 150 msec, TE = 4 msec) and non-fat suppressed heavily T2-weighted fast spin echo (TR ~ 490 msec, TE = 90 msec) images were utilized for this study. Scene size is 512 x 432 x 20-30 and voxel size is 0.78 x 0.78 x 8.00 mm^3.

Notations: NC – non-uniformity correction. IS – intensity standardization. *a*NC – automated NC by the method described in [10]. N3 – NC method of [1]. *a*IS - automated IS by the method described in [9]. *i*NC, *i*IS – interactive NC and interactive IS, respectively, by the new approach presented in this paper. D – a given set of images. D_c – the set of images resulting from applying a NC method c to the images in D. $D_{c,s}$ – the set of images resulting from applying an IS method s to the images in D_c. If an IS method s is applied directly on the original images in D, the resulting set will be denoted by D_s. In our study, $c \in \{aNC, N3, iNC\}$ and $s \in \{aIS, iIS\}$. $t1, \ldots, tn$ – n different types of tissues. $D^{t1}, \ldots, D^{tn}$ – the sets of binary images corresponding to the images in D where the binary images represent masks indicating sample tissue regions for $t1, \ldots, tn$, respectively.

The basic idea behind both *i*NC and *i*IS approaches is to first decide on the different tissue types that can be used as reference and to subsequently delineate sample regions for these tissues in the images in D. The selection of tissue type should be guided by three criteria: (i) the tissue is normal or near normal, (ii) the tissue region is spatially well-distributed throughout the image volume, and (iii) the tissue region can be reliably sampled by an efficient interactive operation such as painting with broad brush strokes. The sampled tissue regions are subsequently utilized for performing both *i*NC and *i*IS operations.

For each image in the given set D of images, a binary mask file for each of n tissue types $t1,\ldots, tn$ is created by drawing tissue regions guided by the above criteria. The mask does not need to be the exact delineation of the whole boundary of the tissue region, but the marked regions should be pure in the sense of containing only one tissue type. Any combination of quick interactive segmentation tools can be used for this purpose such as painting by broad brush strokes. We perform this and all other operations in this paper by using the open source software system CAVASS [11]. We employ the same set of mask data sets $D^{t1}, \ldots, D^{tn}$ for both *i*NC and *i*IS processes as described below. In other words, user interaction is needed only once and not separately for each process.

*Interactive non-uniformity correction - i*NC

For *i*NC, we are given the original image set D and the associated binary image sets $D^{t1}, \ldots, D^{tn}$. For each image I in D and the tissue region R_i corresponding to ti in I, we estimate the degree of non-uniformity at each voxel in R_i as traditionally done by comparing the voxel value to the mean intensity in R_i for estimating either an additive or a multiplicative factor. We fit a second degree polynomial to the estimated non-uniformity factors in R_i. In this process, we treat each tissue region as an independent entity. Subsequently, we combine the fitted functions from all tissue regions into a global function. The image I is then corrected using the combined global function. For the input set D, the output image set produced by the *i*NC operation will be D_{iNC}.

*Interactive intensity standardization - i*IS

The *i*IS operation consists of two steps – a one-time *calibration step* and a *transformation step* that is performed for each image to be standardized. To reap the full benefits of standardization, as recommended in [12], we perform *i*IS after *i*NC. Otherwise, if *i*IS is performed first followed by *i*NC, as demonstrated in [12], residual non-standardness introduced by the (any) NC method will be left over. That is, we perform *i*IS on the set D_{iNC} resulting from applying *i*NC on D. Note that *i*IS can also be applied directly on D, or D_c for any $c \in \{aNC, N3, iNC\}$, or even D_{aIS}. Below we will use D_c to denote any input image set for the *i*IS operation.

The purpose of the *calibration step* is to estimate parameters of the *standard intensity scale*, see Figure 1. Some sample images from the input set D_c are used for this purpose. The parameters of the standard scale are the "signatures" μ_{t1}, ..., μ_{tn} of the different reference tissues. In this paper, we take μ_{t1}, ..., μ_{tn} to be the mean of the intensities in the n tissue regions R_i within the sample images of D_c. Instead, other tissue-specific signatures may also be employed. Additionally two other parameters are used – p_1 which denotes the lowest meaningful intensity, and p_2 that indicates the largest meaningful intensity. As in [8], we take p_1 to be the smallest intensity over all sample images and p_2 to be the intensity at the 99.8 percentile within the sample images. By observing 1000s of MR images under different protocols and body regions, it has been found in [8] that there is usually a distinct cut off at 99.8 percentile beyond which the intensities represent mostly outlier values.

In the *transformation step*, given any image $I \in D_c$, it is standardized by first finding its tissue signatures, in our case the mean intensities m_{t1}, ..., m_{tn} corresponding to its n tissue regions, and then performing a piecewise linear transformation, as illustrated in Figure 1. Instead of the piecewise linear mapping, any curve fitting method may also be used. The set of images that result will be, as per our notation, $D_{c,iIS}$. Note that the transformation extends beyond (p_2, μ_2) along the
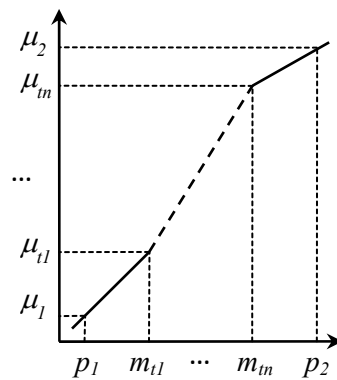

Figure 1.   Illustration of the *i*IS operation. The abscissa represents the mean intensities of the reference tissues in a test image and the ordinate denotes the standardized reference intensities.

positive axes directions and beyond (p_1, μ_1) along the negative axes directions following the corresponding linear segments. In this manner, no input intensities are lost and a 1:1 mapping can be guaranteed.

To sum up, user interaction is needed by the proposed approach only once for each image for performing both *i*NC and *i*IS operations. As pointed out earlier, many combinations of the NC and IS operations and their iterative executions are possible, although we will not explore them in this paper.

3.   RESULTS

We use the data set D described in the previous section for testing our *i*NC and *i*IS approaches. For the calibration step of *i*IS, we used 5 images from the set D_{iNC}. For comparison, we also applied the automatic methods *a*NC and *a*IS in that order to D to yield the non-uniformity corrected set D_{aNC} and corrected and standardized image set $D_{aNC,aIS}$. We used two tissue types ($n = 2$), fat and muscle, as reference for the correction and standardization operations.

Note that many different combinations of operations can be carried out and compared for the quality of the final images. Following our notation, the combinations which we will compare in this initial study are as follows. For NC: D, D_{N3}, D_{aNC}, and D_{iNC}. For IS: D, $D_{aNC,aIS}$, and $D_{iNC,iIS}$.

To assess the performance of an NC operation $c \in \{aNC, N3, iNC\}$, we estimate the percent coefficient of variation (%cv) of intensity within each specific tissue region in each image I in the resulting image set D_c and then describe the mean and standard deviation of this %cv value over all images in this set. Percent cv values within a tissue region in an image I in D_c give an indication of the amount of residual non-uniformity left over in I as attested by the tissue region.

For this evaluation, in addition to the fat and muscle tissue regions used as reference, we use the liver object region as an independent additional test tissue region. To facilitate this analysis, we segmented the liver 3D region in all images by using the user-driven live wire method. The image sets have been selected carefully to make sure that the liver region is normal in every image.

For assessing the performance of an IS operation s in conjunction with an NC operation c in the resulting image set $D_{c,s}$, the mean intensity is estimated in each tissue region in each image I in $D_{c,s}$, and the mean and standard deviation of this mean intensity are described over all images in the set. This is a sensible approach since non-standardness is a trait of not a given image but of how the same tissue property changes over a subject population [6, 8, 9]. We use the same three tissue regions for IS evaluation as for NC.

Tables 1 lists the mean and standard deviation of the %cv values for NC operations on T1-weighted in-phase images for the sets D, D_{N3}, D_{aNC}, and D_{iNC}. The p-values resulting from separate t-tests comparing D_{N3}, D_{aNC}, and D_{iNC} with D are also listed in the table. Similarly Table 2 lists %cv and p-values for the T2-weighted images. Tables 3 and 4 analogously display the mean and standard deviation of mean tissue intensities for T1- and T2-weighted images illustrating the performance of the IS operations.

Table 1. Results for non-uniformity correction. Mean and standard deviation of %*cv* values for T1-weighted in-phase MR images.						
	Liver	p-value	Fat	p-value	Muscle	p-value
D	17.66 (3.05)	-	16.82 (4.48)	-	17.27 (15.66)	-
D_{N3}	18.45 (2.86)	0.15	11.53 (5.59)	0.0	12.64 (4.89)	0.08
D_{aNC}	17.50 (7.31)	0.46	19.85 (9.70)	0.06	15.16 (6.26)	0.25
D_{iNC}	17.35 (5.26)	0.36	14.01 (5.87)	0.01	11.76 (3.40)	0.04

Table 2. Results for non-uniformity correction. Mean and standard deviation of %*cv* values for T2-weighted MR images.						
	Liver	p-value	Fat	p-value	Muscle	p-value
D	33.05 (9.47)	-	18.40 (4.97)	-	36.25 (7.58)	-
D_{N3}	33.44 (8.22)	0.35	9.65 (3.32)	0.0	36.79 (7.33)	0.2
D_{aNC}	33.64 (9.42)	0.46	17.17 (6.99)	0.06	45.48 (18.94)	0.25
D_{iNC}	30.83 (9.04)	0.36	13.38 (3.24)	0.01	34.89 (6.95)	0.04

Table 3. Intensity standardization. Mean and standard deviation of tissue mean intensities for T1-weighted in-phase MR images.									
	D			$D_{aNC, aIS}$			$D_{iNC, iIS}$		
	Liver	Fat	Muscle	Liver	Fat	Muscle	Liver	Fat	Muscle
Mean intensity	325.21	575.42	273.55	1460.8	2534.64	1248.32	1713.33	2567.20	1278.11
SD	191.24	282.72	133.92	161.47	264.65	294.91	183.11	87.02	24.38
cv%	58.81	49.13	48.96	11.06	10.44	23.62	10.69	3.39	1.91

Figure 2 displays one slice of the T1-weighted images from each of three subjects from the sets D, $D_{aNC, aIS}$, and $D_{iNC, iIS}$. The images from D shown in the first row are difficult to display at a fixed window gray level setting because of the intensity non-standardness present among subjects. The second row displays images from the set $D_{aNC, aIS}$ where automatic NC and IS have been applied. The images clearly demonstrate an improvement when displayed at a fixed gray level window setting. The last row shows corresponding images for the same three subjects drawn from $D_{iNC, iIS}$ resulting from interactive NC and IS. The fixed gray level window display seems to show the best portrayal of each type of tissue in the three subjects at about the same brightness level, clearly indicating the improved numeric characterization achieved for the tissues across subjects.

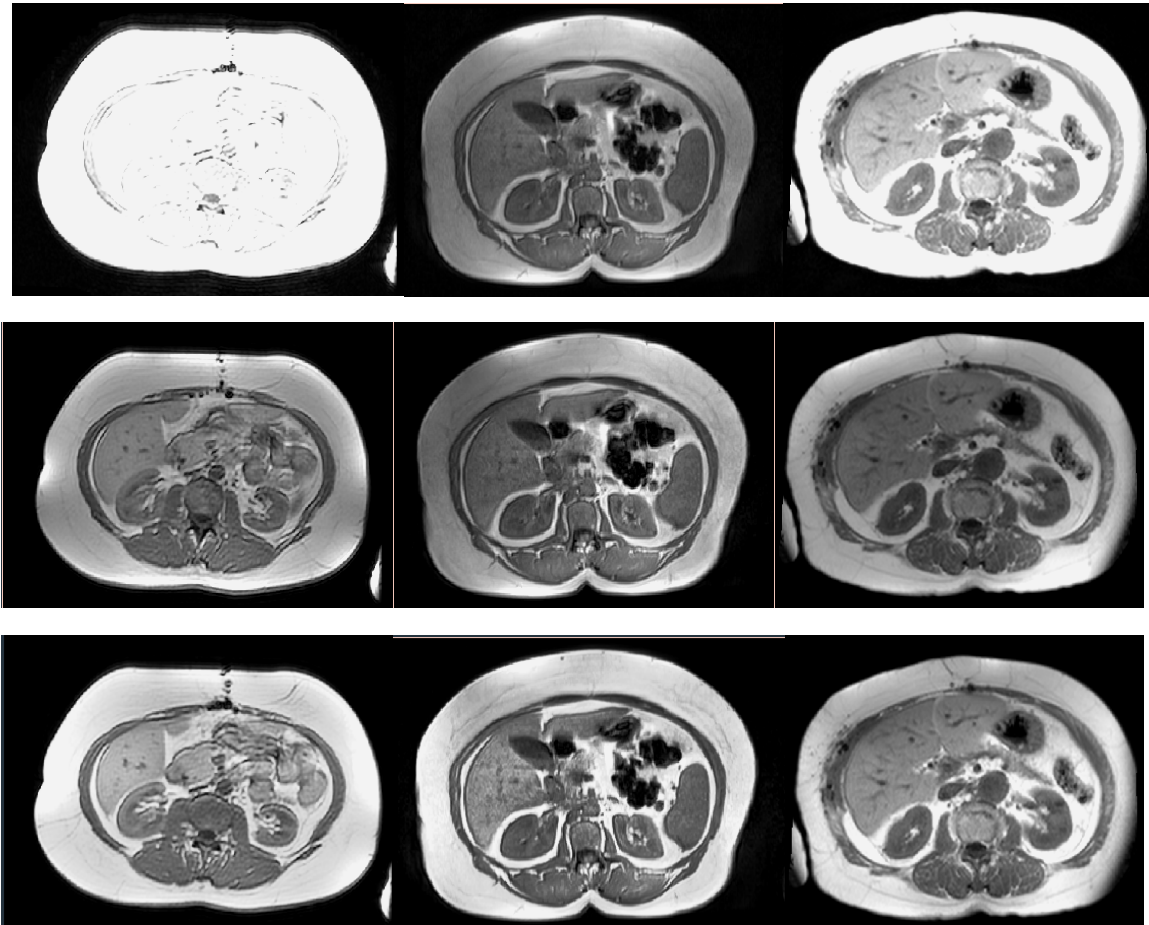| Table 4. Intensity standardization. Mean and standard deviation of tissue mean intensities for T2-weighted MR images.								
	D			$D_{aNC, aIS}$			$D_{iNC, iIS}$		
	Liver	Fat	Muscle	Liver	Fat	Muscle	Liver	Fat	Muscle
Mean intensity	128.40	678.08	117.51	531.33	2475.21	373.10	647.86	2483.35	497.19
SD	56.32	290.89	69.79	161.50	193.04	169.78	166.61	38.77	20.96
$cv\%$	43.87	42.90	59.39	30.40	7.80	45.51	25.72	1.56	4.22

Figure 2. A display of a sample slice from the images of three subjects. Top row: D. Middle row: $D_{aNC, aIS}$. Bottom row: $D_{iNC, iIS}$.

From Tables 1 and 2 we observe that iNC generally performs better than the other two automatic NC methods. More importantly, the quality of the images obtained after iNC and iIS operations is substantially better than the quality of the images after the automated operations aNC and aIS for all tissue regions.

4. CONCLUSIONS

In this paper, we presented an interactive strategy for correcting image intensity non-uniformities and non-standardness. The interaction required for both operations is only once in the form of regions of two or more tissue regions painted in each given image. Non-standardness correction has an initial one-time calibration step. We presented an evaluation of the two methods comparing them to some automated methods from the literature on T1- and T2-weighted abdominal

MR images. The overall quality improvement achieved by the two operations applied in tandem seems to be substantial in terms of the tissue specificity achieved by the numeric pixel values.

More extensive evaluations of other combinations of the NC and IS operations on other types of images from different body regions are underway. One disadvantage of the methods is the human interaction required on a per image basis. We are investigating ways of reducing this interaction to a minimum where we can still achieve substantial quality advantage for the resulting images. We are also studying severe cases of non-uniformity that can be handled satisfactorily by the proposed methods where other automatic methods may fail.

Acknowledgements

References

[1] Sled JG, et al. "A nonparametric method for automatic correction of intensity non-uniformity in MRI data," IEEE Trans. Medical Imaging, vol. 17, 87-97 (1998).
[2] Tustison, N.J., Avants, B.B., Cook, P.A., Zheng, Y., Egan, A., Yushkevich, P.A., Gee, J.C., "N4ITK: improved N3 bias correction," IEEE Transaction on Medical Imaging, 29 (6), 1310-1320 (2010).
[3] Vovk U, Pernus F, Likar B., "A review of methods for correction of intensity inhomogeneity in MRI," IEEE Trans Med Imaging, 26(3), 405-421 (2007).
[4] Bai W., Shi W., Ledig C., Rueckert D., "multi-atlas segmentation with augmented features for cardiac MR," Medical Image Analysis, 19(1), 98-109 (2015).
[5] Wu. S., Udupa J.K., Marinaki A., Weinstein S.P., Kontos D., "Intensity standardization in breast MR images improves tissue quantification," Medical Imaging: Proc. SPIE, vol. 8668, 866822-1-866822-6 (2013).
[6] Ge Y, Udupa J.K., et al. "Numerical tissue characterization in MS via standardization of the MR image intensity scale," Journal of Magn.Reson. Imag, vol. 12, 711-721 (2000).
7] Zhuge Y, Udupa J.K., "Membership-based multiprotocol MR brain image segmentation," Medical Imaging: Proc. SPIE, vol. 5032, 1572-1579 (2003).
[8] Nyúl LG, Udupa J.K., "On standardizing the MR image intensity scale," Magn. Reson. Med., 42 (6), 1072-1081 (1999).
[9] Nyúl LG, Udupa J.K., "New variants of a method of MRI scale standardization," IEEE, Trans. Medical Imaging, vol. 19, 143-150 (2000).
[10] Zhuge, Y., Udupa J.K., et al., "Image background inhomogeneity correction in MRI via intensity standardization," Comp. Med. Imag. and Graph., 33(1),7-16 (2009).
[11] Grevera G., Udupa J.K., et al., "CAVASS: A computer-assisted visualization and analysis software system," Journal of Digital Imaging, vol. 20, 101–118 (2007).
[12] Madabhushi A, Udupa J.K., "Interplay between intensity standardization and inhomogeneity correction in MR image processing," IEEE Trans. Medical Imaging, vol. 24, 561-576 (2005).