NeuroImage 49 (2010) 345–354

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Construction and assessment of a 3-T MRI brain template

Florent Lalys [a,b,c,*], Claire Haegelen [a,b,c,d], Jean-Christophe Ferre [a,b,c,e], Omar El-Ganaoui [a,b,c], Pierre Jannin [a,b,c]

[a] INSERM, U746, Faculté de Médecine CS 34317, F-35043 Rennes Cedex, France
[b] INRIA, VisAGeS Unité/Projet, F-35042 Rennes, France
[c] University of Rennes I, CNRS, UMR 6074, IRISA, F-35042 Rennes, France
[d] Department of Neurosurgery, Pontchaillou University Hospital, F-35043 Rennes, France
[e] Department of Radiology, Pontchaillou University Hospital, F-35043 Rennes, France

ARTICLE INFO

Article history:
Received 27 March 2009
Revised 15 July 2009
Accepted 5 August 2009
Available online 12 August 2009

Keywords:
Brain template
Basal ganglia
3-T MRI
Deep brain stimulation

ABSTRACT

New MR imaging protocols enable visualization of brain structures. However, for dedicated clinical applications such as targeting deep brain stimulation (DBS), a more accurate localization requires the use of atlases. We developed a three-dimensional digitized mono-subject anatomical template of the human brain based on 3-T magnetic resonance images (MRI). By averaging 15 registered T1 image acquisitions, we have shown that the final image corresponds to an optimal image, limited by the performance of the 3-T MR machine. We compared different preprocessing workflows for template construction. With the optimal strategy, along with validated existing processing methods, one T1 template and one T1–T2 mixing template were created in order to improve visualization of spatially complex deep structures. Reduction of voxel size to 0.25 mm^3 was also advantageous to observe fine structures and white matter/gray matter intensity crossings. Results demonstrated that such a template also improved inter-patient registration for population comparison in DBS. These MR templates are made freely available to our community (http://www.vmip.org/mritemplate) to serve as a reference for neuroimage processing methods.

Introduction

The development of medical imaging equipment is driving increased demand for reference data sets. Anatomical reference images are becoming of vital importance for comparison of results, to achieve optimal spatial and intensity resolution, and to allow better identification of structures. As reference images, templates are defined as anatomical models built from multiple volume averaging. They can be mono- or multi-subject ones and are generally based on a single modality. Atlases are derived from templates but can be built from different modalities and are most often characterized by specific structure labeling. Paper-based atlases were originally obtained from experts in anatomy who manually drew and labeled reference images. For instance, atlases of the human brain turn out to be very helpful for various procedures. Printed atlases by Schaltenbrand and Wahren (1977), Talairach and Tourmoux (1988), or Ono et al (1990) have been used with success in various computer-aided decision systems. Today, digital atlases are directly built from digital images and offer new capabilities and applications. Nowinski et al (1997) used a combination of these three digitized print atlases to develop a digital atlas, which has been used for surgical planning and intraoperative support. Such atlases are also used as anatomy teaching tools, by providing interactive labeling of structures and high image resolution and contrast. In image analysis, deformable atlases provide a powerful tool for image segmentation, by exploiting constraints derived from the image data together with a priori knowledge of structure parameters (Zhou and Rajapakse, 2005). Elastic registrations to a template of the organ of interest are also very useful, enhancing the objectivity of image interpretation (Friston et al., 1995). Finally, templates may serve as a common space for population comparisons studies.

Digital templates can be classified according to the number of subjects used for the computation. Multi-subject templates are mainly built from a single acquisition of different control subjects reflecting the population targeted by the clinical study (Seghers et al., 2004; Lee et al., 2005). Such multi-subject templates are primarily intended to serve as anatomical references for spatial normalization usually required before studying human anatomical or functional variability. A prominent example is the multimodal template used in SPM (Statistical Parametric Mapping, Institute of Neurology, University College of London, UK; Evans et al., 1993). The construction of smoothed multiple subject templates captures intersubject variability, but their use for alignment could hinder the representation of targets in a common space.

Single-subject templates strive to attain optimal spatial resolution. Different strategies have been followed for defining such templates. The first strategy was the use of coregistered monomodal or multimodal acquisitions of a single particular subject mainly based

* Corresponding author. Unite/Equipe U746 VISAGES, IRISA, Campus de Beaulieu, 35042 Rennes CEDEX, France. Fax: +33 2 99 84 71 71.
E-mail address: florent.lalys@irisa.fr (F. Lalys).

1053-8119/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2009.08.007

on MR acquisitions. The Colin27 MRI brain template (Holmes et al., 1998) has been used in various neurosurgical applications. For instance, St-Jean et al. (1998) created a deformable volumetric template of the basal ganglia and thalamus in combination with the Schaltenbrand and Wahren atlas in order to estimate template-to-patient transformations. It has also been used to create probabilistic functional templates, like the one by Finnis et al (2003) for the combination of intraoperative data with MRI data or to validate template warping techniques (Chakravarty et al., 2008). The second strategy was the combination of images and histological templates (Yelnik et al., 2007; Chakravarty et al., 2006). The interest of the first strategy is that the whole clinical acquisition setup is applied to healthy, living control subjects, and they prove to be quite feasible in practice. Histological templates have advantages like higher structural and spatial resolutions. However, they turn out to be very complicated and lengthy to produce. Finally, templates following the first strategy are based on the average of all registered volumes, which enhances the quality of the final template by increasing signal-to-noise ratio (SNR) and contrast, essentially between gray matter (GM) and white matter (WM). These templates are constrained by both the resolution and the quality of available imaging technologies.

Digital templates need to be evaluated and extensively validated as they are used for anatomical reference images in various applications. It is crucial to assess the quality of the images by quantifying their parameters, from the contrast resolution to the signal-to-noise ratio (SNR). Many parameters come into the creation of a template, including the MR machine's intrinsic settings, the number of scans to be averaged, the preprocessing methods, the order of preprocessing, and the processing parameters. All of these parameters have to be optimized in order to create the template with the optimal construction strategy. The Colin27 MRI brain template has been lightly validated in the original work of Holmes et al. (1998) by computing intensity profiles to demonstrate the improvement in image quality, but no such studies have been performed in depth. Moreover, templates have to be validated in a clinical context to assess its actual added value. In neurosurgery, they are widely used for surgical planning and targeting, especially in DBS (deep brain stimulation). DBS is a procedure for patients with movement disorders (e.g., Parkinson's disease) for which medical therapy is not effective. It uses electrical impulses to stimulate targets (often the subthalamic nucleus, STN) in the brain. For such neurosurgical procedures, identification of basal ganglia on patient specific images is not always possible because of the lack of contrast between structures. The use of digital atlases has helped in addressing this problem as deep brain structures are more visible and allow more accurate targeting.

In this article, we present how we built and assessed MRI templates using a one-subject average of volumes acquired with off-the-shelf medical image protocols, and processed with up-to-date image processing methods. The objective is to create one T1 template with an optimal construction strategy that we will validate in this paper. This framework relies on the fact that the quality of the template increases with the number of volumes averaged, as demonstrated by Holmes et al. (1998), on condition that every volume is perfectly defined in the same common space. The quality reaches a maximum that would theoretically correspond to an MRI without noise and intensity inhomogeneities. We chose to use image processing methods that have already been validated in the medical imaging context along with fixed MR machine parameters. However, we evaluated the impact of the order of image preprocessing, the number of scans required to reach an optimal image, and the overall evolution of image quality. We showed quantitative results using different complementary criteria demonstrating the effectiveness of the best strategy. Then we studied the impact of template choice in a nonlinear registration task in a deep brain stimulation (DBS) context. Finally, using the optimal strategy, we constructed a T2 template in order to create a multimodal brain template by mixing T1- and T2-weighted data for better visualization of deep brain structures. The basic principle was to retain helpful information from T2, i.e., deep brain structures (such as basal ganglia), and to merge this region of interest (ROI) with the remaining T1 MRI area. The T1 template aims at helping patient-to-template image registration applied in neurosurgical procedures, whereas the T1–T2 mixed template aims at identifying basal ganglia.

Material and methods

Image acquisition

The subject was a 45-year-old man without any clinical pathology. Absence of brain pathology was checked by a neuroradiologist on the two first sets of MR images. Fifteen T1-weighted sequences and 7 T2-weighted sequences were performed on a Philips Achieva 3T system (Philips Medical Systems, Best, The Netherlands) using an 8-channel head coil, on the dates in Table 1. For the selected T1-weighted sequence, the parameters were 3D Fast Field Echo Sequence, sagittal acquisition, 160 continuous slices, section thickness = 1 mm, field of view = 256 mm, TR/TE/TI = 9.8/4.6/915 ms, flip angle = 8°, SENSE factor = 1.5, matrix = 256 × 256 (acquisition), 512 × 512 (after zero-fill interpolation), voxel size = 0.5 × 0.5 × 1 mm, and acquisition time = 5 min 05 s. For the selected 2D T2-Weighted sequence, the parameters were 2D Turbo Spin Echo Sequence, coronal acquisition, 64 continuous slices, section thickness = 1 mm, field of view = 256 mm, TR/TE = 3035/80 ms, flip angle = 90°, echo train length = 15, matrix = 256 × 256, voxel size = 1 × 1 × 1 mm, and acquisition time = 7 min 17 s.

There was minimal patient motion between acquisitions within the same session, except for the eighth where the subject got out of the machine between the two scans in order to simulate two different sessions. In each session, a first Talairach repositioning was performed before launching scans to fit the *x*-axis with AC–PC. Each session took place between 1PM and 2PM, just after lunch, endeavouring to keep the same physiological patient conditions. DICOM images were stored on dedicated CDs.

Preprocessing

For the study of the preprocessing impact, MR images were denoised with the nonlocal means algorithm (Coupé et al., 2008), which has been successfully validated on 3-T MR images (Coupé et al., 2006). The bias correction algorithm used was based on intensity values (Mangin, 2000) and applied using BrainVisa software (CEA, Orsay, France, http://brainvisa.info), which was well adapted for 3-T MR images (Vovk et al., 2007). An intensity normalization step was

Table 1
Dates of scan sessions.

No. of session	1	2	3	4	5	6	7	8
Date of session	08/07/04	08/07/11	08/07/30	08/07/31	08/08/01	08/09/19	08/09/29	08/11/27
Number of T1 acquisitions	5	5	1	1	1	1	1	0
Number of T2 acquisitions	2	1	0	0	0	1	1	2

not necessary since all images were acquired in accordance with the same protocol, on the same subject.

Construction of the templates

Registration

One T1 native image was randomly chosen to be the target for registration. We realigned the target volume to have the AC–PC line on *y*-axis and the mid-sagittal plane on the *z*-axis. All volumes (T1 and T2) were then linearly registered to the target (rigid registration, cost function: Mutual Information, NEWUOA optimization (Wiest-Daesslé et al., 2007)). Every native image was defined in the same common space, which allowed us to perform the average.

Average

We used a method called kappa-sigma clipping average (Jorsater, 2006). This method, largely used in astronomy, finds and deletes pixels that are significantly different from the median. Median and σ were first computed for each voxel of the image. A threshold was then applied to reject all the voxels with an intensity superior to median$\pm$kapp*σ. This method helped reduce the influence of outliers, which are caused by the patient or external factors, e.g., scanner table vibrations (Bernstein et al., 2004). Values of kappa of two or three are usually used in the literature (Newberg et al., 1999). Moreover, internal studies have shown that the quality of the template was unchanged from a value of kappa of 1.8. We chose this value and rejected approximately 2% of all voxels and possible outliers.

Subsampling

Resulting averaged volumes were subsampled, using a cubic B-Spline algorithm. For the T1 template, a twofold reduction was computed for the *x*- and *y*-axes, and a threefold reduction for the *z*-axis. For the T2 template, a fourfold reduction was computed for all axes, so that we obtained a 0.25 mm^3 voxel volume size for both templates.

T1 + T2 fusion

T2-weighted images provide information on deep brain structures that are not visible in T1-weighted images. In order to take advantage of both MR protocols, we computed a mixed T1 + T2 template as follows. We first inversed the intensity of the T2 images in order to fit the WM/GM crossings and the global contrast of the T1 images. An expert defined a bounding box containing basal ganglia, similar to the one described in Chakravarty et al. (2008). Image intensities from T2 images were kept inside the bounding box; the remaining signal was derived from the T1 images. The final step included boundary smoothing. We deliberately applied a low level of smoothing to reveal the boundaries between both modalities.

Effectiveness of the strategy

In order to perform a first visual assessment on the contrast quality, we used the intensity profile. This nonquantitative criterion was extracted from the native nonpreprocessed image and from the 15-volume average. The section was taken in the coronal planes through WM and GM.

The image quality evaluation studies for the T1 template were based on five quality criteria that allowed comparing the final 15-volume average (used as a reference for the evaluation studies) with the intermediate *N*-volumes average in order to evaluate and quantify the evolution of improvement. We initially used image intensity correlation (QC1) and mutual information (QC2), defined in the whole brain. These criteria are both based on intensity values and are widely used in medical image processing to evaluate differences between two images. We then used the signal-to-noise ratio (SNR) (QC3), defined by:

$$\mathrm{SNR} = 10.\log\left(\frac{255}{\mathrm{MSE}}\right)$$

where MSE is the mean square error between the reference volume (the 15-volume average image) and the intermediate volume. Another possibility for estimating the SNR was to compute the signal on a region of interest (ROI) divided by the standard deviation of the background noise, which is usually defined outside the anatomy on a MR image. This one, after preprocessing and few volume averages, was null, which made impossible to perform a global study until the 15-volume average image. For this reason, we preferred using the first formula for the estimation of the SNR.

Another vital quality parameter is the contrast, essentially between WM and GM. Similarly to SNR, the usual contrast equation (Fushimi et al., 2007) supposes to compute the standard deviation of the background noise, so it is not applicable to our studies. To quantify it, we computed the acutance (QC4) (Choong et al., 2003), which is the edge contrast of an image:

$$\mathrm{Ac} = \frac{\overline{G}}{x_{\mathrm{GM}} - x_{\mathrm{WM}}}$$

where $\overline{G}$ is the mean density gradient between GM and WM, and $x_{GM}-x_{WM}$ is the distance between GM and WM regions. It represents the amplitude of the derivative of brightness with respect to space, and approaches the definition of the sharpness of the image. In addition, the gradient term adds a notion of distance that is essential in the concept of contrast. A simple $Cd = \frac{S_{GM} - S_{WM}}{D}$ contrast difference (QC5) was also computed, where D is the dynamic range of the image, S_{GM} is the signal in the GM region and S_{WM} the signal in the WM region. For the computation of the acutance and the contrast, GM and WM regions of frontal lobes and basal ganglia (e.g., putamen, thalamus) were selected as ROIs by a neuroradiologist. A total of 10 WM/GM intensity crossings were chosen in order to obtain an overall representative contrast. In each average step, the same ROIs were applied.

To study the impact of preprocessing, we performed the same studies for (1) the native images without preprocessing, (2) images with denoising only, (3) images with bias correction only, (4) images with bias correction followed by denoising, and (5) images with denoising followed by bias correction. The average was computed with the volumes in the same order as their acquisitions. All volumes were warped on the reference target with the transformations computed in the registration step used for the construction of the templates.

We performed two final studies using correlation criterion to assess the impact of the order of images in template construction. The first study compared the average of five images from the same session with five from different sessions. For this study, the evaluation reference was the final 5-volume average. Similarly, the second study compared a randomized selection of acquisitions with the temporal ordered selection.

Intrasubject registration validation

As the quality of the final template was dependent on the quality of the intrasubject image registration, we studied all native T1 registrations with the T1 target. Nine anatomical landmarks were defined by a neurosurgeon (C.H.) and identified on each floating and reference image. Five of them (Fig. 1) were defined within the bounding box describe in the construction of the templates: the anterior and posterior commissures (AC and PC, points 1 and 3, respectively), the interthalamic adhesion on the middle of the axial slice (point 2), the infundibular recesse (point 4), and the middle between the mamillary

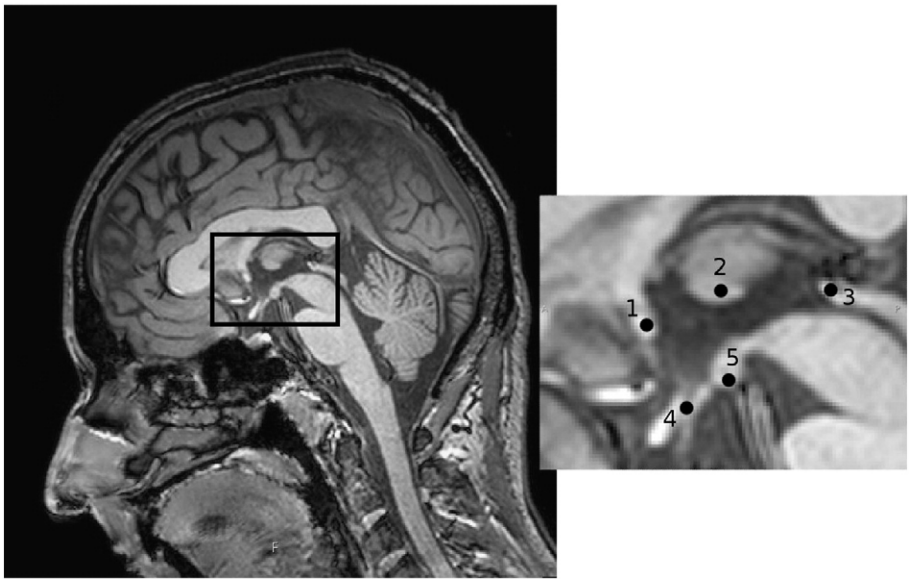

Fig. 1. One native image with the five anatomical landmarks of the bounding box. Point 1: the anterior commissure (AC). Point 2: the interthalamic adhesion on the middle of the axial slice. Point 3: the posterior commissure (PC). Point 4: the infundibular recesse. Point 5: the middle between the mamillary bodies on the axial slice just above the last visualization of the optic chiasma.

bodies on the axial slice just above the last visualization of the optic chiasma (point 5). The four others are the left and right carotid division into anterior and middle cerebral arteries, and the middle of the origin of the trigeminal nerve (also left and right). The Euclidean distance for each landmark defined in both images was computed, in order to obtain the global misplacement error. A similar study was performed to assess the T2 to T1 registration.

Validation for clinical studies

We studied the impact on patient-to-template image registration of our final template compared with the improved version of the Colin27 template (Aubert-Broche et al., 2006) and with a native 3-T scan. We applied the same registration workflow to a set of 15 patients with idiopathic Parkinson's diseases (Lalys et al., 2009). The preoperative T1 MR images of each patient were used. All subjects had STN DBS according to selected inclusion criteria (Lang and Lozano, 1998; The DBS for Parkinson's Disease Study Group, 2001). Examinations were performed on a 3-T whole-body imager (Achieva, Philips Medical Systems, Best, The Netherlands) by using a transmit–receive head coil and were acquired with a Fast Field Echo sequence after injection of Gadolinium. The acquisition parameters were as follows: TE/TR/Flip angle = 4.6 ms/9.9 ms/8°, acquired matrix size = 256 × 256 mm, field of view (FOV) = 256 mm, voxel size = 1 × 1 × 1 mm, volume = 182 axial 1-mm thickness slices, no SENSE factor, and acquisition time = 6 min 59 s. The T1 template was resampled (cubic B-Spline algorithm) to fit the Colin27 template (0.5 mm^3 voxel volume size). We used a strategy close to the one described by Sanchez Castro et al. (2006), which states that the best registration of patient images to the atlas in the DBS context was a global affine image to atlas registration, followed by a nonlinear registration using a Demons algorithm along with semiautomated segmentations of deep structures. The registration procedure proposed here has been adapted to be fully automated and not only to be close to the one proposed by Sanchez but also to be very similar to the work of Chakravarty et al. (2008). The following registration workflow was applied for each patient: affine T1-MR-to-atlas registration was first computed. Then a local affine registration was computed on a region of interest including the deep brain structures. The final step included a nonlinear local registration step using the Demons approach [www.itk.org] without presegmentation step, which estimated a 3D deformation field between a source volume and a target volume. A landmark-based validation study was applied, similar to the intrasubject registration validation

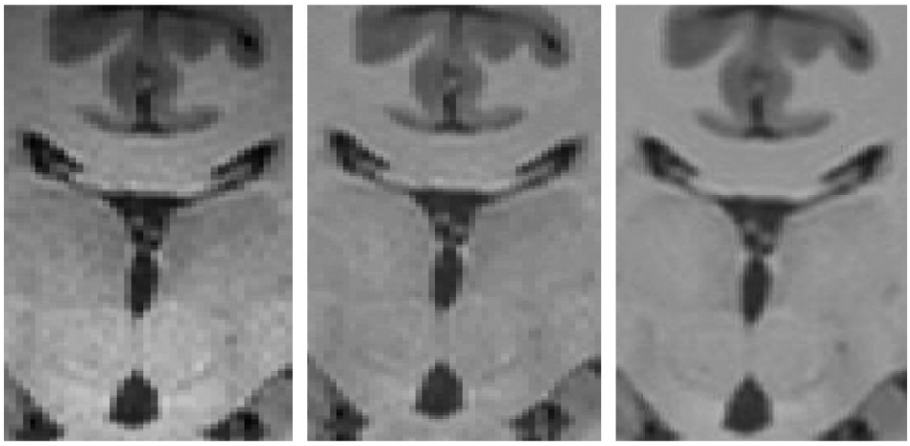

Fig. 2. Coronal slices of a native image (left), a native image after denoising and inhomogeneities correction (middle), and the T1 template (right).

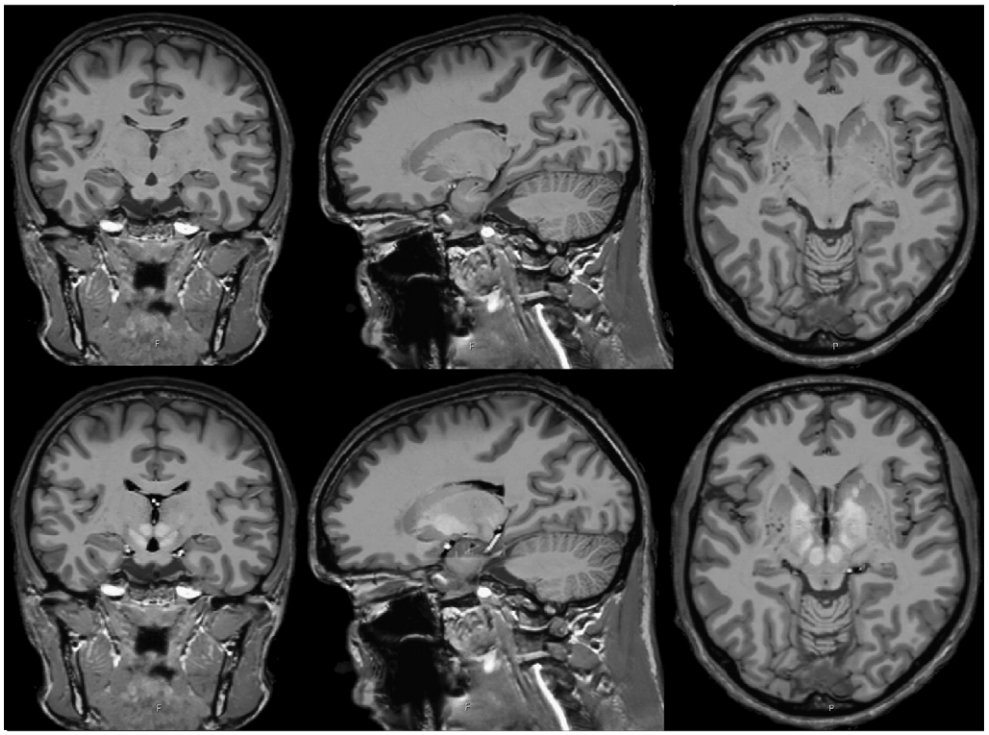

Fig. 3. Coronal, sagittal, and transversal slices of the T1 (above) and T1 + T2 (below) templates.

study. As we decided to limit this study to a DBS context, the registration validation was based only on the five anatomical landmarks in the bounding box (Fig. 1), which were manually defined by a neurosurgeon on the 15 volumes and on the two templates. The intrarater variability of landmark placement was also computed. The 5 landmarks used in this study were placed three times on five normal subjects. The intrarater variability was measured from the distances between the multiple placements; (dist (L_1, L_2) + dist (L_1, L_3) + dist (L_2, L_3)/3), in which L_{1-3} are landmark location from the three trials.

Results

Construction of the templates

Volume averaging visually enhances the global quality of the template (Fig. 2). Fig. 3 shows the homogenization of the signal and an accurate visualization of GM and WM regions on the T1 template. On the mixed T1–T2 template, the bounding box is visible but intensities of the inverse-T2 globally match the T1 boundaries. This allows visualization of deep brain structures with high resolution as well as a better demarcation of GM/WM intensity crossings.

Fig. 4 shows the 7–T2 volumes average before contrast inversion and the subsampling step. Deep brain structures are clearly visible in the mixed T1–T2 template, e.g., the transversal slices on Fig. 5.

Effectiveness of the strategy

Intensity profile

Fig. 6 shows a visual improvement from the native image to the template, as GM/WM intensity crossings are more significant in the averaged image. Each GM/WM intensity region is defined in lower/higher intensities, thereby increasing the differentiation between them. We observed a reduction in noise, as WM areas seem to be more homogeneous.

Figs. 7a–e shows image quality criteria for each *N*-intermediate average volume. Computations for Figs. 7a–c were based on a ground truth: the 15-volume image.

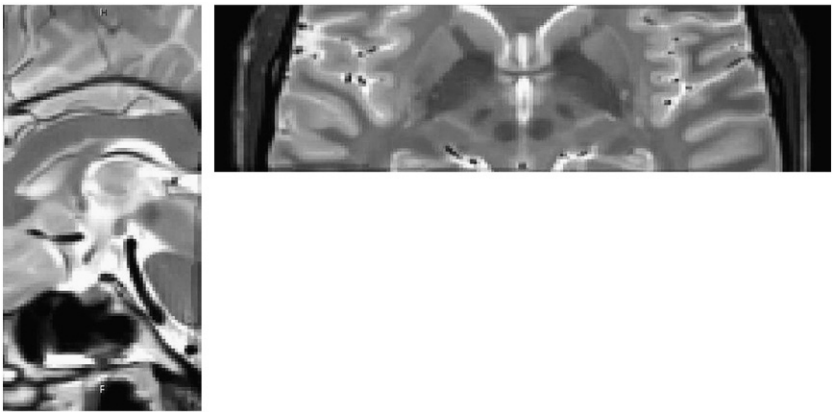

Fig. 4. Details of the T2 template.

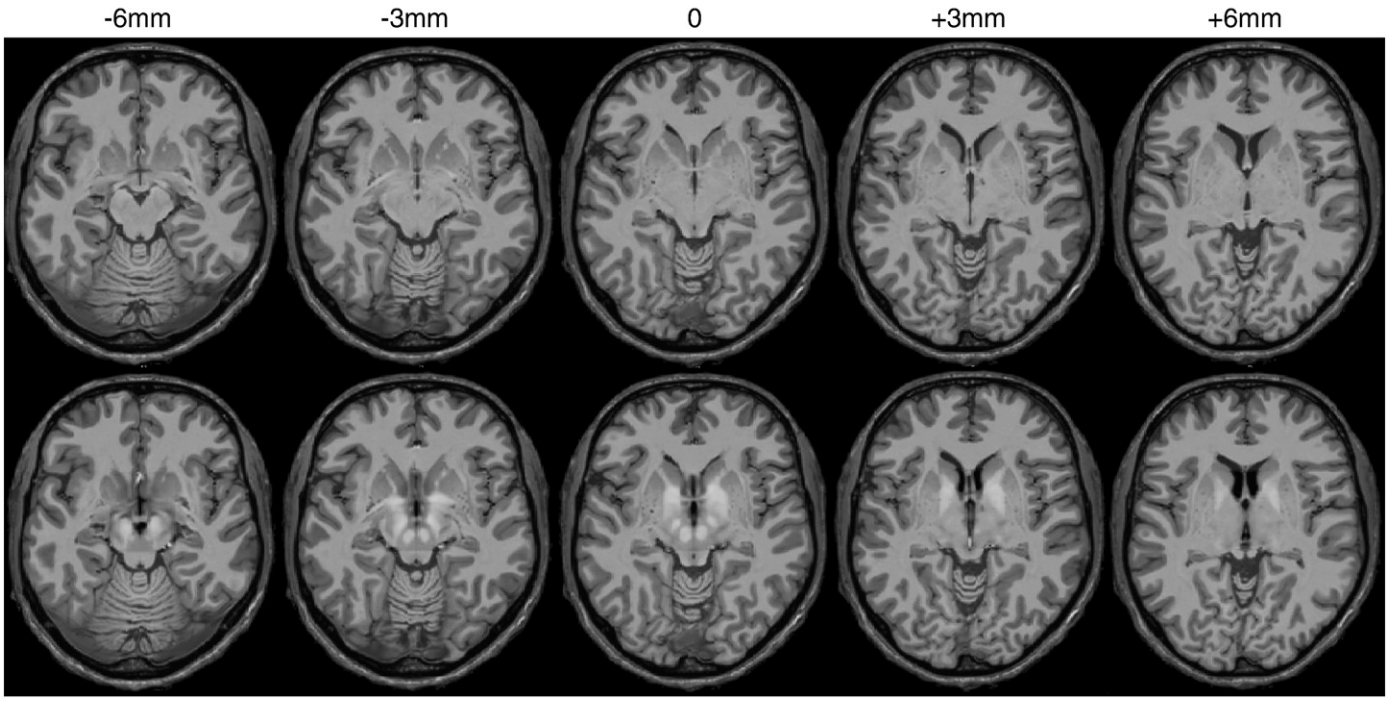

Fig. 5. Five transversal slices of the T1 (above) and T1 + T2 template (below).

For all criteria, the five first images (corresponding to acquisitions of the first session) did not have a significant impact on the final result and on the improvement of quality criteria. The profiles increased, but very slowly. When the acquisition session changed, at the fifth average, all criteria were suddenly raised, until the sixth or seventh average. The comments from the first session could also be applied, since from the fifth to tenth averages, it corresponded to the second session. The images subjected to a bias correction were superior to those without this preprocessing. In terms of acutance, this result was valuable starting from the sixth average. For correlation, acutance, SNR and contrast, the profiles seemed to reach a maximum, whereas it was not the case for mutual information. We found that an average with images first denoised, and then corrected, had almost the same profile as the average with images first corrected, and then denoised. Moreover, in all criteria, profiles from nonpreprocessed images were almost identical to those only denoised.

In Fig. 8, the averages based on acquisitions from different sessions evolved more than the ones based on acquisitions from the same session, which did not vary. In Fig. 9, the profile of randomized image averages had a more linear increase than the ordered image averages, without the sudden change at the fifth and tenth average.

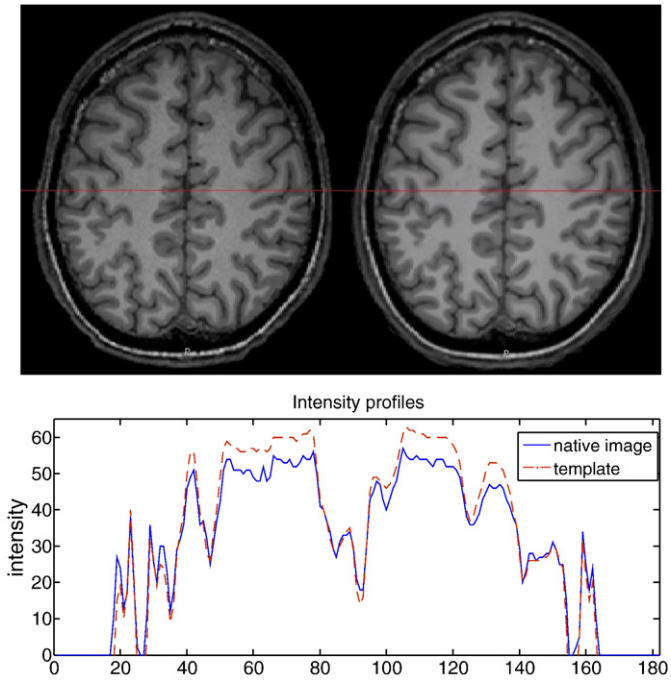

Fig. 6. Intensity profiles of a coronal slice of the native image (left) and the T1 template (right).

Intrasubject registration

For all 15 T1 volumes, the global placement error was null; all five anatomical landmarks perfectly matched their corresponding points on the native target. The same results were found with the T2-registered volumes.

Validation for clinical studies

We found a placement error of 1.58 ± 0.33 mm for the registration with the Colin27 template, 1.07 ± 0.36 mm with a native 3-T image and 0.78 ± 0.19 mm with our template. ANOVA and a multiple comparison test were computed at a 5% significance level for the 15-patient study on Fig. 10. Every group was found significantly different from each other. The intrarater variability (Table 2) was 0.66 ± 0.32 mm.

Discussion

In the current study, we constructed and validated MR mono-subject brain templates. This kind of template could serve as a teaching tool to visualize complex deep structures and as a spatial reference for neurological studies. Different criteria were studied, first in order to find the optimal strategy for building this template and then to demonstrate the impact of using such high-quality image templates in patient-to-template registration. Results demonstrated the quality of the resulting template and its impact on accurate registration.

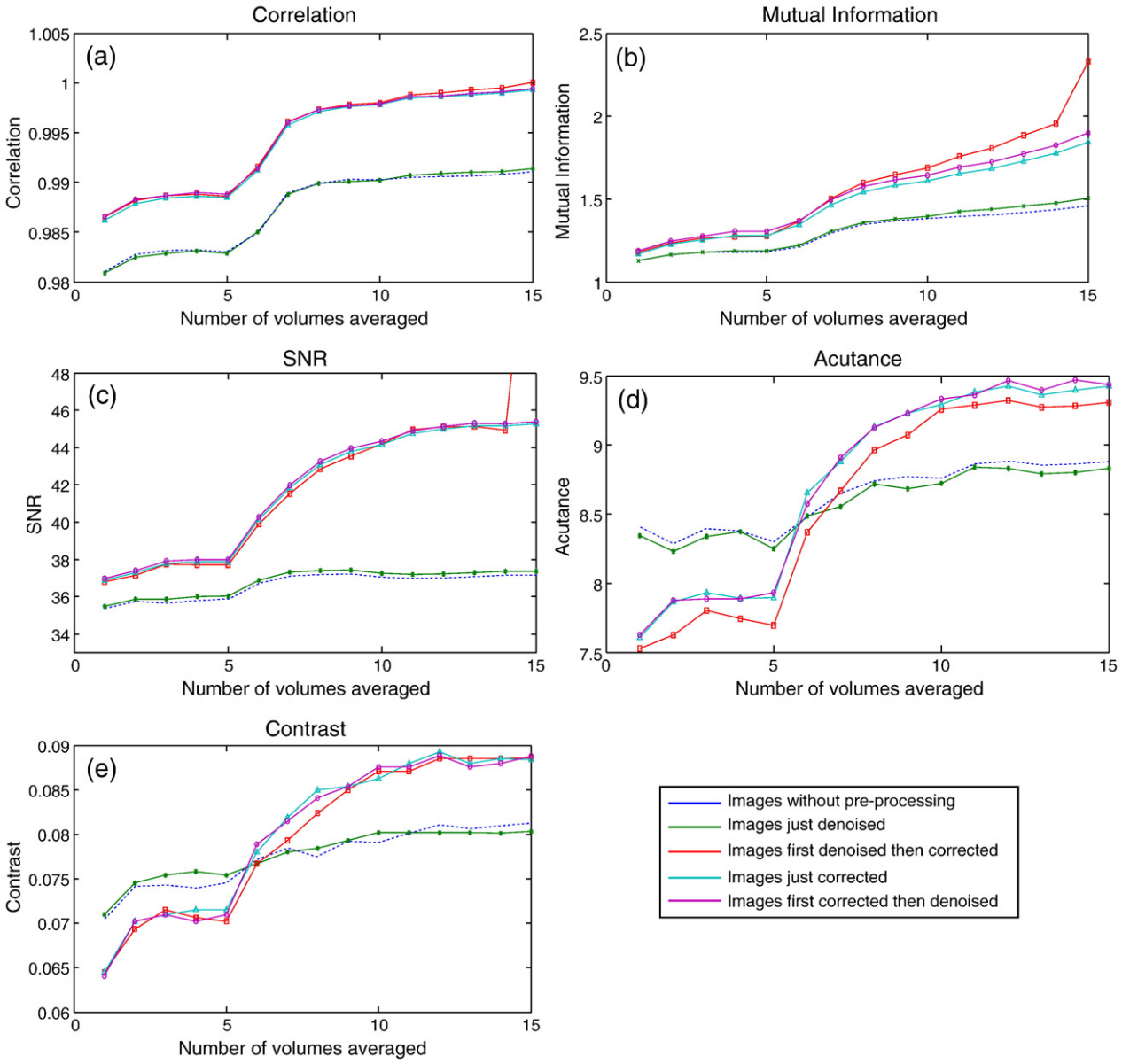

Fig. 7. Quality criteria for all N-intermediate average volumes. (a) Correlation, (b) mutual information, (c) SNR, (d) acutance, (e) contrast.

Influence of 3-T machine

It has recently been proven that signal-to-noise ratio is significantly better at 3 T than at 1.5-T MR imaging (Hoenig et al., 2005; Manka et al., 2005; Frayne et al., 2003; Schick, 2005; Yongbi et al., 2002). Other authors (Nobauer-Huhmann et al., 2002, Scarabino et al., 2003; Sasaki et al., 2003; Ross, 2004) have reported lower contrast between GM and WM at 3 T, but it was subjective, as it was based on visual assessment only. In Fushimi et al. (2007), differences of contrast

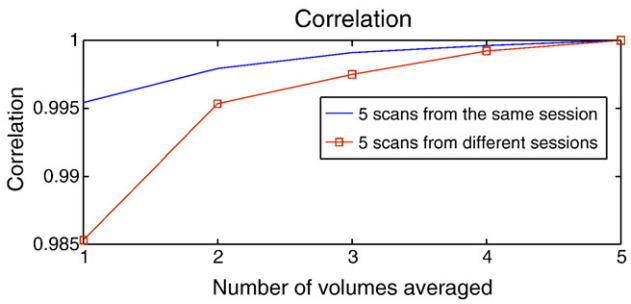

Fig. 8. Correlation between 5 volumes templates and the N-intermediate volumes for two cases: with acquisitions from the same session and with acquisitions from different sessions.

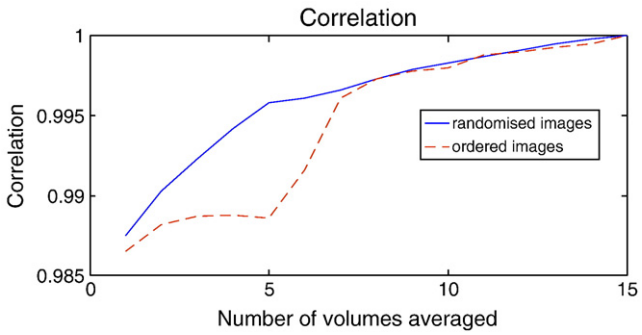

Fig. 9. Correlation between the final T1 template and the intermediate average of randomized images and ordered images (taken from the date of acquisition).

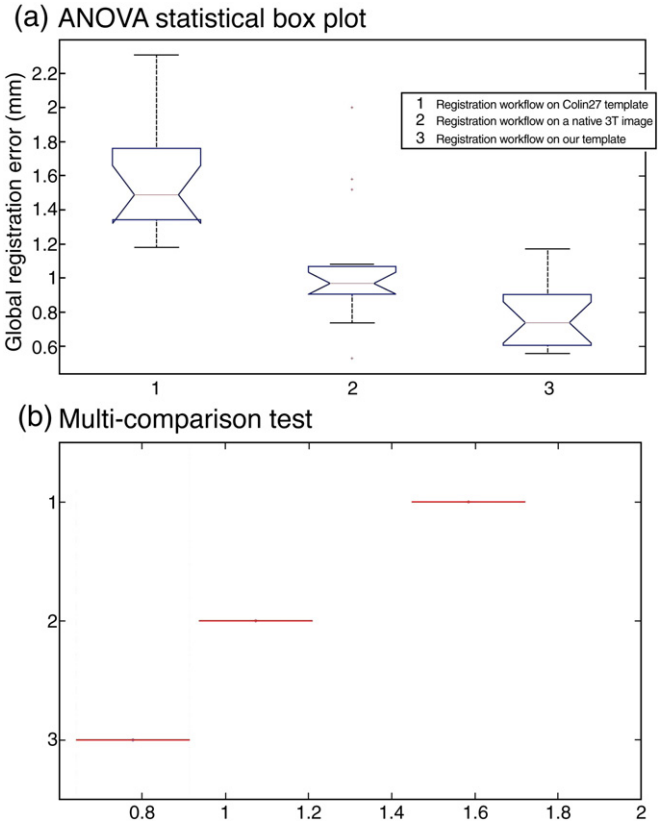

Fig. 10. Global registration error on a 15-patient study, with a workflow using the Colin27 template, a native 3-T image and our template. (a) ANOVA statistical box plot. (b) Multicomparison test.

between 1.5 and 3 T were quantitatively computed with the CNR and best results were shown for the 3-T images.

Patient motion

As mentioned earlier, the ideal result would be an image without noise, inhomogeneities, and partial volume effects, keeping in mind that the major source of improvement is due to the SNR increase. As already mentioned (Holmes et al., 1998), a small patient motion between each acquisition seems to be essential to increase the quality of the final template. It can be explained by the modification of the magnetic fields and the noise inside the machine. Intrinsic machine parameters do not significantly change between consecutive acquisitions. That means that averaging these images is comparable to averaging virtually identical images. Figs. 8 and 9 show that it is essential to acquire scans from different sessions in order to decrease the noise. To rapidly raise the quality of our template, starting from the tenth acquisitions, we chose to wait a few days between each scan. With this approach, we were assured that the subject did not have the same physical location in the machine and that intrinsic machine parameters had fluctuated. Moreover, it allows us to suggest that the intense signal regions of the magnetic field could have moved between acquisitions and that the global spatial energy could be reformed after the average. As we can see in Fig. 7, for almost all criteria, the global quality of the volumes does not significantly increase from the 12th to 13th average. This allows us to suggest that the template had reached a limit corresponding to an optimal image without noise and with a global contrast corresponding to the threshold of 3-T machine performances.

Table 2
Intrarater variability.

	Max	SD	Mean
Landmark placement variability (mm)	1.88	0.32	0.66

Registration validation

With current imaging technologies, intrasubject rigid registration appears to be very effective, especially when native acquisitions are subject to a reliable preprocessing step. A good spatial definition of the volumes in a common space is required to precisely compute the average. The main drawback of this method is a deterioration of the contrast with higher GM/WM intensity crossings. As we found no registration error, the quality of the final template is not distorted by inaccurate intrasubject registration.

Preprocessing influence

As already mentioned, the preprocessing workflow was divided into two principal components: inhomogeneity correction (IC) and noise correction using a nonlocal means (NLM) method. Without loss of generality, we have chosen the reference image as the image preprocessed by (NLM+IC).

In Figs. 7a–c, profiles that did not include a preprocessing step of IC were always higher than those including this step. This can be explained by the fact that the image histogram was significantly changed by the IC and was therefore more influential on intensity values than a simple NLM. This argument also explains why the two profiles with NLM preprocessing alone and without preprocessing are relatively similar in all figures. In Figs. 7d and e, profiles that did not include a preprocessing step of IC were above those including this step until the sixth acquisition. This is probably due to the dynamic spread that decreases the contrast and the acutance, which is closely related to the contrast definition.

The order of the two preprocessing steps is often related to the application (Montillo et al., 2003) and can modify the results of the image processing applied after these treatments. For Figs. 7a–e, the profiles NLM+IC and IC+NLM are nearly identical. Only Fig. 7d shows a significant difference between the two profiles, with IC+NLM higher than NLM+IC. These observations fit the previous assumption that the preprocessing has to comply with the image processing. When noise removal is applied first, the weak edges can be smoothed and will not be recovered by inhomogeneity correction. For image processing dedicated to segmentation, the IC is usually used before denoising, which is often unnecessary to obtain good segmentation (Pham and Prince, 1999). In our case, the order of the two operations has no consequences. Indeed, acutance is the edge contrast of an image and is therefore different from simple contrast and from the three other criteria. IC may remove some anatomical information, but we were not interested in recovering the exact bias field, just a delineation of the tissue's gray level distribution (Mangin, 2000).

It is important to point out that all results found here are unique to the set of preprocessing algorithms that we used, and that results would obviously vary depending on the use of other commonly used methods.

Usefulness of T1–T2 mixed template

Multi-subject MR templates have been shown to allow better visual inspection of deep brain structures due to better contrast, since averaging improved the signal-to-noise ratio (D'Haese et al., 2005; Bardinet et al., 2008). Mono-subject MR templates further increase contrast since there is no anatomical variability between scans. Our T1–T2 mixed template further improves contrast of basal ganglia, such as the subthalamic nucleus and the globus pallidus, which are targets of DBS in patients with movement disorders. Our templates would be useful in the preoperative workflow for improved DBS

targeting (Stancanello et al., 2008; Dawant et al., 2007; Guo et al., 2007), as well for postoperative assessment (Lalys et al., 2009). Such high-resolution low-noise templates could also be used to evaluate and improve atlas-based image processing methods, such as segmentation (Scherrer et al., 2008; Pohl et al., 2006).

Results of the comparative patient-to-template image registration study have also shown that the T1 template was a good reference for registration methods. Finally, these templates could serve as a teaching tool to visualize complex structures, which are barely visible in current MR images.

Validation for clinical studies

Results of the patient-to-template registration comparison have shown that our T1 template increases the accuracy of a patient-to-template basal ganglia registration. A recent study of patients with Parkinson's diseases (Rocha Vasconcellos et al., 2009) has proved that no significant anatomical differences exist for these patients when compared to a control group. However, limitations of this validation study can be found in the usual limitations of such landmark-based approaches. The main limitation is that such validations are very susceptible to human factor. Landmark-based validation is also difficult for neurosurgeons because of the homogeneous regions within the brain. Reproducible landmarks are very hard to define, which limits its applicability. One other uncertainty concerns the spatial localization of anatomical landmarks, which is limited by the image quality and resolution. In order not to bias the landmark-based validation, we performed a resampling step on our T1 template to achieve the same image resolutions for both templates. One additional limitation is the number of experts and subjects used. Even if significant differences between both templates were found in our study, a larger one is required to completely validate the contribution of our new template. Nevertheless, the improved accuracy found when using our 3-T template compared to the 1.5T Colin27 one is in large part explained by the better signal available with 3-T MR images (see Fig. 10). In addition, we expect that registration is more accurate when using same field MR images. Results of this patient-to-template registration validation were unique to a specific registration workflow, and it would be interesting to try other warping techniques on our new template. Original DICOM images are made freely available to experiment other approaches.

Conclusion

In this article, we have reported on the construction of in vivo healthy human neuroanatomy MR templates. Our objective was to optimize spatial and intensity resolutions. The reduction of noise enhanced the visibility of fine structures, as both contrast and SNR increased with the number of volumes averaged. Such image quality is not available with current imaging protocols. Small movements between scans turned out to be a vital condition to enhance the quality of the final template. We demonstrated that the two computed templates allow visualization of spatially complex structures as well as increased contrast between GM and WM. We also showed that they greatly improved the accuracy of template-based registration. Both resulting templates are freely available online (http://www.vmip.org/mritemplate).

References

Aubert-Broche, B., Evans, A.C., Collins, L., 2006. A new improved version of the realistic digital brain phantom. NeuroImage 32, 138–145.

Bardinet, E., Bhattacharjee, M., Dormont, D., Pidoux, B., Malandain, G., Schüpbach, M., Ayache, N., Cornu, P., Agid, Y., Yelnik, J., 2008. A three-dimensional histological atlas of the human basal ganglia. II. Atlas deformation strategy and evaluation in deep brain stimulation for Parkinson disease. J. Neurosurg. 31, 1–12.

Bernstein,, M.A., King,, K.F., Zhou,, X.J., 2004. Handbook of MRI pulse sequence. Elsevier, Amsterdam.

Chakravarty, M., Bertrand, G., Hodge, C.P., Sadikot, A.F., Collins, D.L., 2006. The creation of a brain atlas for image guided neurosurgery using serial histological data. NeuroImage 30 (2), 359–376.

Chakravarty, M., Sadikot, A.F., Germann, J., Bertrand, G., Collins, D.L., 2008. Towards a validation of atlas warping techniques. Med. Image Anal. 12, 713–726.

Choong, Y.F., Rakebrfsandt, F., Morgan, J.E., 2003. Acutance, an objective measure of retinal nerve fibre image clarity. Br. J. Ophthalmol. 87, 322–326.

Coupé, P., Yger, P., Barillot, C., 2006. Fast non local means denoising for 3D MR images. MICCAI'06, International Conference on Medical Image Computing and Computer Assisted Intervention, in Lecture Notes in Computer Science, Springer, Copenhaguen, Denmark, 4191, 33-40.

Coupé, P., Yger, P., Prima, S., Hellier, P., Kervram, C., Barillot, C., 2008. An optimized blockwise non local means denoising filter for 3D magnetic resonance images. IEEE Trans. Med. Imag. 24 (4), 425–441.

D'Haese, P.F., Cetinkaya, E., Konrad, P.E., Kao, C., Dawant, M.B., 2005. Computer-aided placement of deep brain stimulators: from planning to intraoperative guidance. IEEE Trans. Med. Imag. 24 (11), 1469–1478.

Dawant, B., D'Haese, P.F., Pallavaram, S., Li, R., Yu, H., Spooner, J., Davis, T., Kao, C., Konrad, P., 2007. The VU- DBS project: integrated and computer-assisted planning, intra-operative placement, and post-operative programming of deep-brain stimulators. Proceedings of SPIE: Medical Imaging, San Diego, CA, 6509, 2007. 701–710.

Evans, A.C., Collins, D.L., Mills, S.R., Brown, E.D., Kelly, R.L., Peters, T.M., 1993. 3D statistical neuroanatomical models from 305 MRI volumes, 1993. Proc IEEE Nucl Science Symp Medl Imaging Conf, 1813–1820.

Finnis, K.W., Starreveld, Y.P., Parrent, A.G., Sadikot, A.F., Peters, T.M., 2003. Three-dimensional database of subcortical electrophysiology for image-guided stereotactic functional neurosurgery. IEEE Trans. Med. Imag. 22 (1), 93–104.

Frayne, R., Goodyear, B.G., Dickhoff, P., Lauzon, M.L., Sevick, R.J., 2003. Magnetic resonance imaging at 3.0 Tesla: challenges and advantages in clinical neurological imaging. Invest. Radiol. 38, 385–402.

Friston, K.J., Holmes, A.P., Worsley, K.J., Poline, J.P., Frith, C.D., Frackowiak, R.S.J., 1995. Statistical parametric maps in functional imaging. Hum. Brain Mapp. 2, 189–210.

Fushimi, Y., Miki, Y., Urayama, S., Okada, T., Mori, N., Hanakawa, T., Fukuyama, H., Togashi, K., 2007. Gray matter-white matter contrast on spin-echo T1-weighted images at 3 T and 1.5 T: a quantitative comparison study. Eur. Radiol. 17 (11), 2921–2925.

Guo, T., Parrent, A.G., Peters, T.M., 2007. Surgical targeting accuracy analysis of six methods for subthalamic nucleus deep brain stimulation. Comp. Aid Surg. 12 (6), 325–334.

Hoenig, K., Kuhl, C.K., Scheef, L., 2005. Functional 3.0-T MR assessment of higher cognitive function: are there advantages over 1.5-T imaging? Radiology 234, 860–868.

Holmes, C.J., Hoge, R., Colin, L., Woods, R., Toga, A.W., Evans, A.C., 1998. Enhancement of MR images using registration for signal averaging. J. Comput. Assist. Tomogr. 22 (2), 324–333.

Jorsater, S., 2006. Methods in astronomical image processing with special applications to the reduction of CCD data. Lect. Notes Phys. 413, 193–207.

Lalys, F., Haegelen, C., Abadie, A., Jannin, P., 2009. Post-operative assessment in deep brain stimulation based on multimodal images: registration workflow and validation. Proceedings of SPIE: Medical Imaging, Orlando, FL, 7261.

Lang, A.E., Lozano, A.M., 1998. Parkinson's disease. First of two parts. N. Engl. J. Med. 339 (15), 1044–1053.

Lee, J.S., Lee, D.S., Kim, J., Kim, Y.K., Kang, E., Kang, H., Kang, K.W., Lee, J.M., Kim, J.J., Park, H.J., Kwon, J.S., Kim, S.I., Yoo, T.W., Chang, K.H., Lee, M.C., 2005. Development of Korean Standard Brain templates. J. Korean Med. Sci. 20, 461–483.

Mangin, J.-F., 2000. Entropy minimization for automatic correction of intensity non uniformity. IEEE Press, Hilton Head Island, SC, pp. 162–169.

Manka, C., Traber, F., Gieseke, J., Schild, H.H., Kuhl, C.K., 2005. Three-dimensional dynamic susceptibility-weighted perfusion MR imaging at 3.0 T: feasibility and contrast agent dose. Radiology 234, 869–877.

Montillo, A., Udupa, J.K., Axel, L., Metaxas, D.N., 2003. Interaction between noise suppression and inhomogeneity correction in MRI. Proceedings of SPIE: Medical Imaging, San Diego, CA, 5032, 1025–1036.

Newberg, H.J., Richards, G.T., Richmond, M., Fan, X., 1999. Catalog of four-color photometry of stars, galaxies, and QSOS using SDSS filters. Astrophys. J., Supl. Ser. 123 (2), 377–435.

Nobauer-Huhmann, I.M., Ba-Ssalamah, A., Mlynarik, V., Barth, M., Schoggl, A., Heimberger, K., Matula, C., Kaider, A., Trattnig, S., 2002. Magnetic resonance imaging contrast enhancement of brain tumors at 3 Tesla versus 1.5 Tesla. Invest. Radiol. 37, 114–119.

Nowinski, W.L., Fang, A., Nguyen, B.T., Raphel, J.K., Jagannathan, L., Raghavan, R., Bryan, N.R., Miller, G.A., 1997. Multiple brain atlas database and atlas based neuroimaging system. Comput. Aided Surg. 2 (1), 42–66.

Ono, M., Kubik, S., Abernathy, C.D., 1990. Atlas of the Cerebral Sulci. Georg Thieme Verlag/Thieme medical publishers, Stuttgart, Germany.

Pham, D.L., Prince, J.L., 1999. An adaptive fuzzy C-means algorithm for image segmentation in the presence of intensity inhomogeneities. Pattern Recogn. Lett. 20 (1), 57–68.

Pohl, K.M., Fisher, J., Grimson, W.E., Kikinis, R., Wells, W.M., 2006. A Bayesian model for joint segmentation and registration. NeuroImage 31, 228–239.

Rocha Vasconcellos, L.F., Pereira Novis, S.A., Moreira, D.M., Rosso, A.L.Z., Leite, A.C.C.B., 2009. Neuroimaging in Parkinsonism: a study with magnetic resonance and spectroscopy as tools in the differential diagnosis. Arq. Neuro-Psiquiatr. 67 (1).

Ross, J.S., 2004. The high-field-strength curmudgeon. AMJN Am. J. Neuroradiol. 25, 168–169.

Sanchez Castro, F., Pollo, C., Meuli, R., Maeder, P., Culsenaire, O., Bach Cuadra, M., Villemure, J., Thiran, J., 2006. A cross validation study of deep brain stimulation targeting: from experts to atlas-based, segmentation-based and automatic registration. IEEE Trans. Med. Imag. 25 (11), 1440–1450.

Sasaki, M., Inoue, T., Tohyama, K., Oikawa, H., Ehara, S., Ogawa, A., 2003. High-field MRI of the central nervous system: current approaches to clinical and microscopic imaging. Magn. Reson. Med. Sci. 2, 133–139.

Scarabino, T., Nemore, F., Giannatempo, G.M., Bertolino, A., Di Salle, F., Salvolini, U., 2003. 3.0 T magnetic resonance in neuroradiology. Eur. J. Radiol. 48, 154–164.

Schaltenbrand, G., Wahren, W., 1977. Atlas for stereotaxy of the human brain. Thieme, Stuttgart, Germany.

Scherrer, B., Forbes, F., Garbay, C., Dojat, M., 2008. Fully Bayesian joint model for MR brain tissue and structure segmentation. MICCAI'08, International Conference on Medical Image Computing and Computer Assisted Intervention, in Lecture Notes in Computer Science, Springer, New York, USA, 5242, 1066–1074.

Schick, F., 2005. Whole-body MRI at high field: technical limits and clinical potential. Eur. Radiol. 15, 946–959.

Seghers, D., D'Agostino, E., Maes, F., Vandermeulen, D., Suetens, P., 2004. Construction of a brain template from MR images using state-of-the-art registration and segmentation techniques. Lect. Notes Comput. Sci. 3216, 696–703.

St-Jean, P., Sadikot, A.F., Collins, L., Clonda, D., Kasrai, R., Evans, A.C., Peters, T.M., 1998. Automated atlas integration and interactive three-dimensional visualization tools for planning and guidance in funcionnal neurosurgery. IEEE Trans. Med. Imag. 17 (5), 672–680.

Stancanello, J., Muacevic, A., Sebastiano, F., Modugno, N., Cerveri, P., Ferrigno, G., Uggeri, F., Romanelli, P., 2008. TTT3T MRI evaluation of the accuracy of atlas-based subthalamic nucleus identification. Med. Phys. 35 (7), 3069–3077.

Talairach, J., Tournoux, P., 1988. Co-Planar Stereotaxic Atlas of the Human Brain. Thieme, Stuttgart, Germany.

The Deep-Brain Stimulation for Parkinson's Disease Study Group, 2001. Deep-brain stimulation of the subthalamic nucleus or the pars interna of the globus pallidus in Parkinson's disease. N. Engl. J. Med. 345 (13), 956–963.

Vovk, U., Pernus, F., Likar, B., 2007. A review of methods for correction of intensity inhomogneity in MRI. IEEE Trans. Med. Imag. 26 (3), 405–421.

Wiest-Daesslé, N., Yger, P., Prima, S., Barillot, C., 2007. Evaluation of a new optimization algorithm for rigid registration of MRI data. Proc. SPIE: Med. Imag. 6512, 601–610.

Yelnik, J., Bardinet, E., Dormont, D., Malandain, G., Ourselin, S., Tandé, D., Karachi, C., Ayache, N., Cornu, P., Agid, Y., 2007. A three-dimensional, histological and deformable atlas of the human basal ganglia. I. Atlas construction based on immunohistochemical and MRI data. NeuroImage 34 (2), 618–638.

Yongbi, M.N., Fera, F., Yang, Y., Frank, J.A., Duyn, J.H., 2002. Pulsed arterial spin labelling: comparison of multisection baseline and functional MR imaging perfusion signal at 1.5 and 3.0 T: initial results in six subjects. Radiology 222, 569–575.

Zhou, J., Rajapakse, J.C., 2005. Segmentation of subcortical brain structures using fuzzy templates. NeuroImage 28 (4), 915–924.