ELSEVIER

NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 32 (2006) 138 – 145

A new improved version of the realistic digital brain phantom

Berengere Aubert-Broche,* Alan C. Evans, and Louis Collins

McConnell Brain Imaging Center, Montreal Neurological Institute, McGill University, 3801 University Street, Montreal, Quebec, Canada H3A 2B4

Received 12 November 2005; revised 23 February 2006; accepted 7 March 2006

Image analysis methods must be tested and evaluated within a controlled environment. Simulations can be an extremely helpful tool for validation because ground truth is known. We created the digital brain phantom that is at the heart of our publicly available database of realistic simulated magnetic resonance image (MRI) volumes known as BrainWeb. Even though the digital phantom had l mm^3 isotropic voxel size and a small number of tissue classes, the BrainWeb database has been used in more than one hundred peer-reviewed publications validating different image processing methods.

In this paper, we describe the next step in the natural evolution of BrainWeb: the creation of digital brain phantom II that includes three major improvements over the original phantom. First, the realism of the phantom, and the resulting simulations, was improved by modeling more tissue classes to include blood vessels, bone marrow and dura mater classes. In addition. a more realistic skull class was created. The latter is particularly useful for SPECT, PET and CT simulations for which bone attenuation has an important effect. Second, the phantom was improved by an eight-fold reduction in voxel volume to 0.125 mm^3. Third, the method used to create the new phantom was modified not only to take into account the segmentation of these new structures, but also to take advantage of many more automated procedures now available. The overall process has reduced subjectivity and manual intervention when compared to the original phantom, and the process may be easily applied to create phantoms from other subjects.

MRI simulations are shown to illustrate the difference between the previous and the new improved digital brain phantom II. Example PET and SPECT simulations are also presented.

Introduction

Image processing methods need evaluation data sets to characterize, evaluate and optimize their performance. Three main types of evaluation data sets can be distinguished: real acquisitions of subjects, real acquisitions of physical phantoms and simulations from numerical phantoms. By using real acquisitions, the whole acquisition set up is taken into consideration. Although physical phantoms can provide a gold standard because the ground truth of the object is perfectly known, the data obtained have unrealistic complexity due to the limited number of compartments that they include. Unlike physical phantom acquisitions, subject acquisitions have a realistic complexity, but the ground truth is not available. Simulations provide a way of generating data where ground truth is known and where realistic complexity may be taken into account when a realistic numerical phantom used as input to the simulator and the simulator reproduces the physics of data generation. Contrary to the real acquisitions, the contributions of errors from different sources in acquisition can be separated and evaluated independently. Simulations allow the control of various acquisition parameters, whereas real acquisitions are limited to the types of scan acquired (parameters, slice thickness…). Since some subtle real effects may not be included in simulations, these simulations are not sufficient to evaluate image processing methods. However, the known ground truth and the realistic complexity of the simulated data allow the simulations to be an extremely helpful tool during the process of method evaluation.

A set of realistic simulated anatomical brain magnetic resonance imaging (MRI) volumes, known as BrainWeb, is available to the neuroimaging community[1]. It is also possible for external groups to customize the simulator parameters, run the simulator on our computer system and download the resulting 3D MR simulations along with the digital phantom. These simulations have been used by more than one hundred external groups, e.g. Grabowskia et al. (2000), Arnold et al. (2001), Schnack et al. (2001), Cardenasa et al. (2001) and Tzourio-Mazoyer et al. (2002). These groups have incorporated these simulations to study the performance of techniques such as non-linear co-registration, cortical surface extraction, correction of MRI intensity non-uniformity and tissue classification. For example, Arnold et al. (2001) used the database to study six different algorithms for the correction for MRI intensity non-uniformity.

These MR simulations, and simulations of different modalities as positron emission tomography (PET) and single photon emission computed tomography (SPECT), were generated by varying specific imaging parameters for each tissue type in the simulator. The spatial distribution of these different tissues (gray matter, white matter, cerebrospinal fluid, muscles, skull, skin and fat) was defined on volumetric fuzzy volumes. These fuzzy volumes, where voxel

---

* Corresponding author. Fax: +1 514 3398 2975.
E-mail address: broche@bic.mni.mcgill.ca (B. Aubert-Broche).

[1] http://www.bic.mni.mcgill.ca/brainweb/.

1053-8119/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2006.03.052

intensity is proportional to the fraction of tissue within the voxel, define the digital brain phantom (Collins et al., 1998).

This digital brain phantom was created by registering and averaging 27 T1-weighted, 12 PD-weighted and 12 T2-weighted MRI scans from a single subject. As a direct result of the high signal-to-noise ratio, these single subject average volumes exhibit fine anatomical details and enable a faithful representation of the brain's complex anatomical structures.

In this paper, we describe the important modifications made to this phantom and its construction. The goal of the first modification is to improve the realism of the phantom and thus of the simulated data. The voxel size of the phantom was decreased to allow the discrimination of finer detail and enable a better classification. To improve the realism of MRI or fMRI simulations, more structures were segmented such as marrow, dura mater and blood vessels. To improve SPECT, PET or CT simulations, the phantom required a better skull class, with cranium and facial bones. These improvements were made possible with computed tomography (CT) and MR angiography (MRA) of this subject, acquired after the creation of the initial phantom in 1998.

The second modification concerns the method used to construct the phantom. This method has been modified not only to take into account the creation of new segmented structures, but also to be mostly automated in order to be easily applied to create phantoms from other subjects.

The following sections detail phantom construction and these improvements.

Methods

Data acquisition

The anatomical phantom is derived from high-quality T1-, T2- and PD-weighted images, formed from averages of 27, 12 and 12 scans respectively, of the same normal subject (Holmes et al., 1998) using a 1.5 T Phillips clinical scanner. The acquisition protocol consisted of a series of 20 1 mm^3 T1-weighted MR scans acquired using a T1-weighted spoiled GRASS [TR/TE = 18 ms/10 ms, FA = 30°, 256 * 256 matrix, 256 mm field of view, NSA 1], a series of 7 0.78 mm^3 T1-weighted MR scans acquired using a T1-weighted spoiled GRASS [TR/TE = 20 ms/12 ms, FA = 40°, 256 * 256 matrix, 200 mm field of view, NSA l] and a series of 12 2-mm thick T2/PD-weighted MR scans acquired with dual echo fast SE [2D multiple slice, TR/TE = 3300 ms/35,120 ms, 256 * 256 matrix, 256 mm * 204 mm rectangular field of view, NSA 1, 12 echoes].

In addition, a computed tomography (CT) scan and a MR angiography of this subject were performed to assist the segmentation of bone and vascular structures. The CT scan [0.98 * 0.47 * 0.47 mm^3 voxel size] was obtained on a Picker spiral CT scanner. The MRA scan was acquired using a 0.45 mm thick T1-weighted FFE sequence [TR/TE = 48.2 ms/6.7 ms, FA = 15°, 512 * 512 matrix, NSA 1]. The CT and MRA data were not available when the original phantom was created.

Construction of the T1, T2 and PD average volumes

Preprocessing

Intensity non-uniformity in MR images is reduced by applying a 3D non-uniformity correction algorithm based on the deconvolution of the non-uniformity blurring kernel from the intensity histogram of the image (Sled et al., 1998). This process, independent of pulse sequence, is applied to each of the native T1-, T2- and PD-weighted MR images independently before the registration and averaging described below.

Stereotaxic registration

One T1 native image was chosen at random to serve as a target for registration. This image was linearly registered (9 parameters using cross correlation) to the International Consortium for Brain Mapping (ICBM) average brain space (Evans et al., 1993) using the mritotal automatic stereotaxic registration procedure (Collins et al., 1994). This registration algorithm proceeds with a coarse-to-fine approach by registering subsampled and blurred MRI volumes with the stereotaxic target.

The remaining images were linearly registered (rigid body, 6 parameters) to this first T1 native image. T1 volumes were registered using cross correlation, and the T2/PD pair, with mutual information. The two previous transformations were concatenated so that only one resampling operation was performed to map each native data volume into stereotaxic space.

These resulting transformations were used to resample all native images into the ICBM space. The resampling was done using a trilinear interpolation at a sampling grid of 0.5 mm.

Averaging

All 27 T1-, 12 T2- and 12 PD-weighted resampled images were intensity normalized (using the volume mean value) and averaged to create T1, T2 and PD average volumes respectively.

Construction of the fuzzy tissue volumes

Classification

A fuzzy minimum distance classification was used to classify gray matter (GMc, the "c" means that the volume comes from the classifier), white matter (WMc), cerebrospinal fluid (CSFc), fat (FATc) and background (BKGc). In this process, voxel intensities from T1, T2 and PD average volumes were considered. Prior to classification, approximately four thousand manually selected training points were used to estimate the means (m_{t_1}, m_{t_2} and m_{pd_i}) for each tissue class i and to drive this automatic classification procedure.

In the resulting classified volumes, each voxel intensity v_i of a fuzzy volume represents the fraction of the tissue i (i from 0 to 4, for BKGc, GWc, WMc, FATc or CSFc) within the voxel, between zero and one. If g_{t_1}, g_{t_2} and g_{pd} are the intensities of the voxel considered in T1, T2 and PD average volumes respectively:

$$v_i = \frac{f_i}{\sum_{j=0}^{4} f_j} \text{ with } f_i = \frac{1}{\sqrt{\left(g_{t_1} - m_{t_1 i}\right)^2 + \left(g_{t_2} - m_{t_2 i}\right)^2 + \left(g_{pd} - m_{pd_i}\right)^2}}$$

GMc, WMc or CSFc fuzzy volumes generated by the classifier contained extra-cerebral structures such as muscle or scalp. To separate brain from non-brain structures, automatically generated mask volumes were used, as described in the next section.

Brain and skull extraction

Different masks are required to separate tissues within the intracranial cavity from tissues outside the skull. To begin, we need a mask to identify the brain's surface. This surface is defined by a fully automated surface-model-based method (BET[2] created by Smith (2002)). A deformable model evolves to fit the brain's surface using a set of locally adaptative model forces. The method was applied to the T2 average volume. As cerebrospinal fluid is bright on this imaging sequence, it is included in the brain segmentation. The brain mask obtained (B) contains gray matter, white matter and cerebrospinal fluid. A small bright region of the dura matter was included in this initial brain mask and was removed by manual editing. The complement of this brain mask (named B') was also created to identify non-brain and CSF structures.

BET2 (Brain Extraction Tool version2) is a new fully automated tool (Jenkinson et al., 2005) based on BET (Smith, 2002) for extracting inner and outer skull and scalp surfaces. From the T1 and T2 average volumes, inner and outer skull surfaces are defined. From these surfaces, inner skull volume (IS) and outer skull volume (OS) were computed, as well as skull volume (S) defined as the space between IS and OS. The difference between inner skull volume (IS) and brain mask volume (B) is that inner skull volume (IS) includes the dura matter, whereas the brain mask volume (B) does not include the dura when it is thick (e.g. near the sagittal sinus). When the dura is thin, it is unfortunately included in the CSF due to partial volume effects.

Two skull volumes were defined. The previous skull volume (S), after a single erosion, was used as seen later to identify bone marrow. A second volume (Sct) was used to define the bones of the skull by applying a threshold to the registered CT scan. This second volume (Sct) allows a better segmentation of the facial bones and features, such as the zygomatic, maxilliary, nasal bones and the foramen magnum. These masks will be used in the next section to define different tissues.

Tissue volumes definition

The brain masks were used to separate each of the five volumes output from the classifier into 10 tissue volumes: gray matter (GM), white matter (WM), cerebrospinal fluid (CSF), skull (SKULL), marrow within the bone (MARROW), dura (DURA), fat (FAT), tissue around the fat (FAT2), muscles (MUSCLES) and muscles/skin (SKIN/MUSCLES).

The WMc voxels within the brain mask (B) became the final WM volume, those outside this brain mask formed the muscles (MUSCLES) volumes. The GMc voxels within the brain mask (B) became the final GM class. The remaining GMc voxels outside the brain mask became the class muscle and skin (MUSCLES/SKIN).

Voxels from the CSFc within the brain mask (B) formed the final CSF volume containing both ventricular and subarachnoid CSF. Voxels from the CSFc outside the brain mask (B) but inside the inner skull mask (IS) formed the dura volume (DURA). A small manual intervention was needed in this step to remove some incorrectly classified voxels from the DURA and to add them to CSF. Voxels from the CSFc outside the inner skull mask (IS) but inside the outer skull volume (OS), i.e. inside the skull volume (S), formed the bone marrow volume (MARROW).

Voxels from the background (BCKc) within the skull mask (Set) defined from the CT scan form the skull (SKULL).

The relation between these volumes is summarized in Table 1. After the creation of these ten classes, the sum of all tissue fractions for a voxel is not always equal to 1.0. These voxels were simply normalized by dividing their respective tissue components by their sum.

Table 1
Relationship between the original classified, binary masks and final tissue volumes that define the phantom

Class name	Creation
GM	$\text{GMc} \cap B$
WM	$\text{WMc} \cap B$
CSF	$\text{CSFc} \cap B$
SKULL	$\text{BCKc} \cap \text{Sct}$
MARROW	$\text{CSFc} \cap S$
DURA	$\text{CSFc} \cap (B' \cap \text{IS})$
FAT	FATc
FAT2	$\text{WMc} \cap B'$
MUSCLES	$\text{GMc} \cap B'$
SKIN/MUSCLE	$\text{CSFc} \cap \text{IS}'$

The $\cap$ symbol indicates intersection. B is the brain mask, B' is its complement, IS is the inner skull mask, IS′ is its complement, S is the skull mask defined by the surface-model-based method, Sct is the skull mask defined from CT scan.

Vessel segmentation

Two different methods were used, depending on the location of the vessels. For vessels within the brain, we used a new multi-scale geometric-flow-based method (Descoteaux et al., 2004) applied on the PD-weighted average volume. A multi-scale 'vesselness' measure based on the eigenvalues of the Hessian matrix was used to detect centerlines of tubular structures. This multi-scale measure is distributed to create a vector field orthogonal to vessels boundaries so that the flux maximizing flow algorithm (Vasilevskiy and Siddiqi, 2002) can be applied to recover them. This procedure yields a binary label volume for the internal vessels.

Superficial vessels do not have a tubular appearance in PD-weighted scans since they are not completely surrounded by a brighter signal and appear as indentations or troughs on the surface of the brain. These vessels were segmented by hand using the registered MRA data for guidance and merged with the vessel label volume described above. In order to have a continuous representation of the vessels (higher values in vessel center, lower values at vessel edge), the combined binary volume was blurred with a 1 mm FWHM Gaussian kernel. The resulting volume made the VESSEL class.

Addition of the vessel class into the phantom

As vessels may be segmented inside other classes, the following condition was applied for all previous classes: if $x_{i,j}$ is the fraction of ith tissue within the voxel j, and $x_{v,j}$ the fraction of vessel within this voxel, $x_{i,j} = x_{i,j} * (1 - x_{v,j})$. For example, if the vessels' fraction is 90%, the fraction of the ith tissue becomes 10% of its initial value.

The eleven fuzzy volumes (the vessel class plus the 10 classes in Table 1) define the digital brain phantom II.

Simulations

MR simulations

The MR simulator (Kwan et al., 1999) uses first principles modeling based on the Bloch equations to implement a discrete event simulation of NMR signal production. Each tissue class was described by its nuclear magnetic resonance (NMR) relaxation properties (T1, T2, T2*) and proton density (PD). The appropriate

---

[2] Brain Extraction Tool, http://www.fmrib.ox.ac.uk/analysis/research/bet.

Table 2
Each tissue class was described by its nuclear magnetic resonance (NMR) relaxation properties (T1, T2, T2*) and proton density (PD) in the MRI simulations and by attenuation (in cm^{-1}) and activity (in Bq/cm^3) for PET and SPECT

Class name	MRI simulation				PET simulation		SPECT simulation	
	T1 (ms)	T2 (ms)	T2* (ms)	PD	Attenuation (cm^{-1})	Activity (Bq/cm^3)	Attenuation (cm^{-1})	Activity (Bq/cm^3)
GM	833	83	69	0.86	0.09853	22,990	0.1551	11,764
WM	500	70	61	0.77	0.09853	8450	0.1551	8194
CSF	2569	329	58	1	0.0956	0	0.1508	0
SKULL	0	0	0	0	0.151108	0	0.3222	0
MARROW	500	70	61	0.77	0.09853	0	0.1551	0
DURA	2569	329	58	1	0.09853	0	0.1551	0
FAT	350	70	58	1	0.087718	8450	0.1394	0
FAT2	500	70	61	0.77	0.087718	8450	0.1394	0
MUSCLES	00	47	30	1	0.098731	8450	0.1553	0
SKIN/MUSLE	569	329	58	1	0.098731	8450	0.1553	0
VESSELS	0	0	0	0	0.0956	0	0.1508	0

values came from the literature and were optimized by minimizing the difference between the real and simulated images (cf. Table 2). These NMR relaxation parameters are uniform for all regions of a tissue class.

Scan parameters such as slice thickness, field of view, receiver bandwidth, number of signal averages, scan matrix size and pulse sequences parameters can all be controlled.

Coil parameters, related to the scanner's system, can also be controlled. They include noise models and RF field inhomogeneity. The noise models could generate percentage noise with a standard deviation given as a percentage of the signal for a reference tissue or intrinsic noise based on the intrinsic SNR model of the scanner and image acquisition parameters. The noise models could also generate images with a required image SNR. Receive and transmit RF inhomogeneity maps related to different MR scanners could also be used.

To compare T1-, T2- and PD-weighted simulated brain images to real scans, scan parameters used were identical to those detailed in the data acquisition paragraph. A percentage noise level of 3% is applied, but no receive and transmit inhomogeneity is used.

PET simulations

The Monte-Carlo-based SORTEO-PET[3] simulator software (Reilhac et al., 2004) was used to simulated PET images with the new digital phantom. It was configured to generate emission and transmission projections for the [^{18}F]FDG radiotracer and for the Ecat Exact HR+ scanner (Siemens). The 3D emission protocol consisted in the collection of data onto a single time frame over a 10-min period starting 45 min post-injection, during which an average of 16 billion disintegration histories were simulated.

A functional model describes the activity within each tissue class of the phantom. For the [^{18}F]FDG functional model, the different regions of the brain phantom were merged onto 2 compartments: gray matter and white matter/skin/fat/muscle. Constant activity levels (in Bq/cm^3) assigned to the two compartments are those published by Reilhac et al. (2005). These values were derived from actual human PET data which were fully corrected (scatter, random, attenuation, system dead-time and radioisotope decay) (Reilhac et al., 2005) and are uniform for all regions of these compartments.

A map of photoelectric linear attenuation coefficients was obtained by assigning a tissue type together with the associated attenuation coefficient v at the 511 keV energy emission to each class of the phantom (cf. Table 2).

Emission data obtained from the simulator were normalized and corrected as detailed in Reilhac et al. (2005). Data were reconstructed using the 3D filtered back projection (Hanning filter, cutoff frequency of 0.3 mm^{-1}).

SPECT simulations

The Monte-Carlo-based SimSET[4] simulator software (Harrison et al., 1993) was used to simulated SPECT images with the new digital phantom. The SimSET software was configured to generate emission projections for the ^{99m}Tc-HMPAO radiotracer and for the three-head gamma camera equipped with ultra high-resolution parallel collimators (IRIX, Philips Medical Systems).

Monte Carlo simulations of SPECT data require the definition of an activity map representing the 3D spatial distribution of the radiotracer and of the associated attenuation map describing the attenuation properties of the body.

An uniform activity value (in Bq/cm^3) is assigned to the white matter and gray matter classes of the phantom (cf. Table 2). These activity values come from the study of Grova et al. (2003) in which a theoretical model of brain perfusion was established from real SPECT data.

A theoretical map of photon attenuation coefficients was obtained by assigning a tissue type together with the associated attenuation coefficient v at the 140 keV energy emission of ^{99m}Tc to each class of the phantom for the radiotracer ^{99m}Tc-HMPAO (cf. Table 2).

Emission data obtained from the simulator were normalized and corrected as detailed in Grova et al. (2003). Reconstruction was performed by filtered backprojection using ramp filter (Nyquist frequency cutoff). The reconstructed data were post-filtered with an 8 mm FWHM Gaussian filter.

Results

Digital brain phantom II

The average of many single image volumes enhances the quality of MR images. While the process enhances the resolution of the phantom, the resulting volume does not have 0.5 mm^3

---

[3] http://sorteo.cermep.fr/.

[4] http://depts.washington.edu/simset/html/simset main.html.

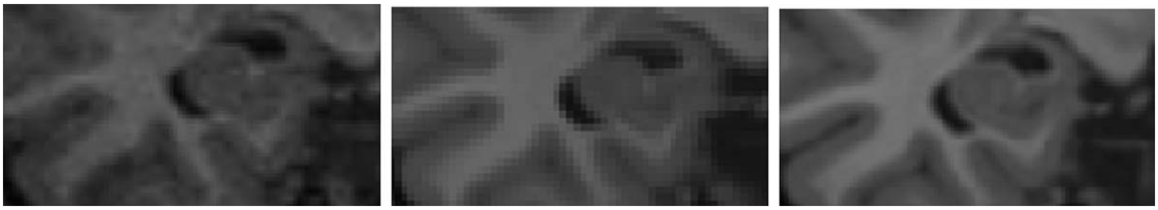

Fig. 1. Coronal slice of a single native T1 MR (left) through the left temporal lobe, the T1 average volume with a voxel size of 1 mm (center) and 0.5 mm (right) showing improved signal-to-noise and neuroanatomical detail, as in the hippocampus or claustrum.

resolution, but the quality is better than the 1 mm^3 data used to create the phantom. The decrease of the voxel size of the phantom could modify partial volume effects at structure borders but allows a better delineation of fine structures (see Fig. 1).

All the fuzzy volumes that define the brain phantom are shown in Fig. 2. A discrete version of the phantom was also created by storing the label of the most important fraction class at each voxel location. This discrete phantom was not used for simulation but rather to aid the visualization of the spatial relationships between the different tissue volumes.

The skull was well detected on the previous and on the new phantom, but the bones and cartilage of the face such as the zygomatic, nasal, maxilla or sphenoid bone were better detected on the new phantom (see Fig. 3A), thanks to the CT scan.

The multi-scale geometric flow method allows a good segmentation of internal vasculature like the anterior, middle and posterior arteries and segments (Fig. 3B). The registered angiography was helpful for the segmentation of superficial vasculature and principally the superior sagittal sinus, the straight sinus, the transverse sinus and some superior cerebral veins. The superficial segmented vessels are also shown in Fig. 3B.

Simulations

To show the impact of the vessels class in the simulation, T2- and PD-weighted images were simulated from the previous and the new phantom. In Fig. 4, sagittal slices are chosen to show the middle cerebral artery. The vessels of the middle cerebral arteries

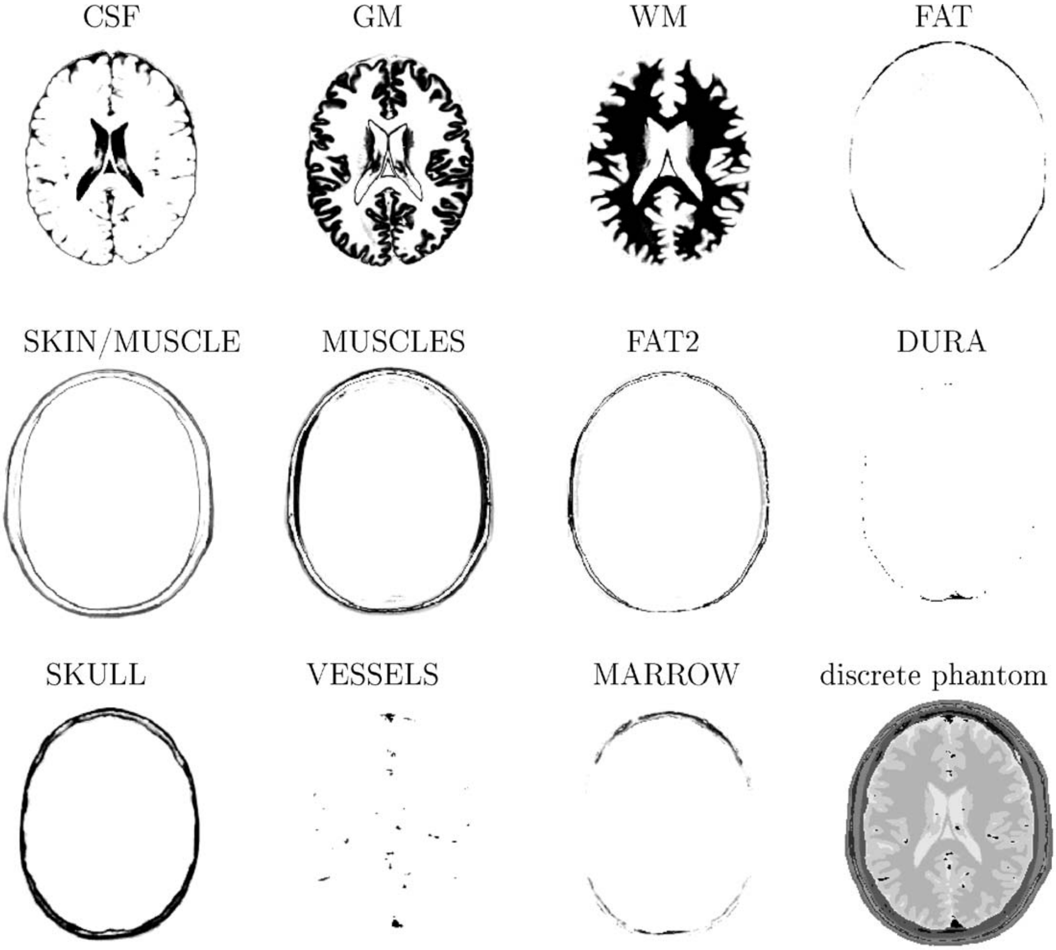

Fig. 2. Transverse slices through the CSF, GM, WM, FAT, SKIN/MUSCLES, MUSCLES, FAT2, DURA, SKULL, VESSELS, MARROW classes and the discrete phantom.

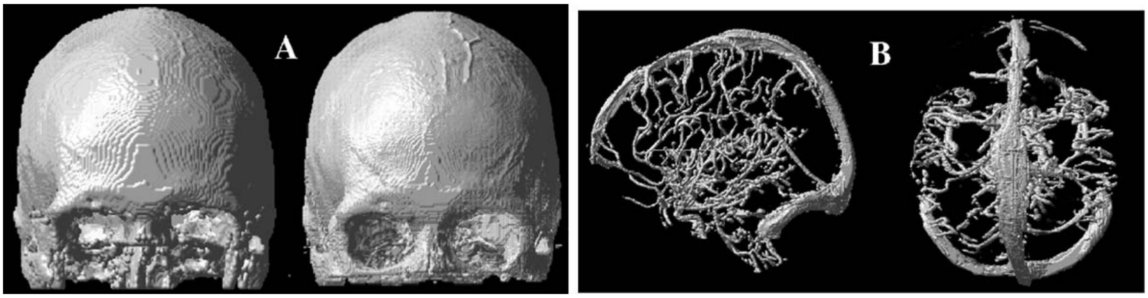

Fig. 3. (A) Skull classes from the previous phantom (on the left) and from the new phantom (on the right) (B) 3D left view (on the left) and top view (on the right) of the vessels class including superficial and internal vessels.

were included in the CSF class in the simulations from the previous phantom, while they were well defined in the simulations from the new phantom. Some superficial vessels could also be seen in the simulations from the new phantom. Furthermore, the dura matter appears darker and more realistic in the new simulations compared to previous simulations.

This phantom has been used to generate PET and SPECT images with the simulators of Reilhac et al. (2004) and Harrison et al. (1993) respectively. The activity and attenuation maps obtained from the anatomical classes of the phantom and filled with values of the Table 2 are shown Fig. 5 for PET data and Fig. 6 for SPECT data.

Discussion

The major improvements made to the phantom are decreased voxel size and the addition of four tissue classes: vessels, dura, bone marrow and bone. The final resolution of the phantom is due to the resolution and sampling rate of the original data, the quality of the inter-volume registrations, the interpolation kernel used in the resampling steps and the sampling rate of the final phantom. Interpolation with a kernel that does not have infinite support will decrease the resolution of the interpolated data. By choosing a small kernel (0.5 mm), the degradation to the interpolation step is minimized. The original data is band-limited, and the registration and averaging process does not increase the resolution of the phantom. However, the signal to noise (SNR) and contrast to noise (CNR) ratios are improved in the average volume, thus enabling visualization and separation of smaller and finer structures when compared to any of the individual MRI data sets. The digital brain phantom II is more anatomically realistic. Defined by “fuzzy” volumes, it models partial volume effects between tissues. Note that, while there may remain errors in classification when compared to the true anatomy used to define the data, we define the digital phantom as truth for simulations. The goal is to generate anatomically realistic images and not to exactly classify the original data. This phantom is a good gold standard to measure the performance of image processing algorithms since it can be used to generate realistic simulations for validation studies.

We work on the assumption that the phantom is the “ground truth” and defines truth in the simulations. Even though we attempted to build the most realistic anatomical phantom possible, a few unrealistic partionings are present in the phantom. We list below some discrepancies between the real brain of the phantom and the possible impact on eventual validations. First, the VESSEL class could be different from the real vessel structure for two reasons: the diameter of the vessels in the phantom may be

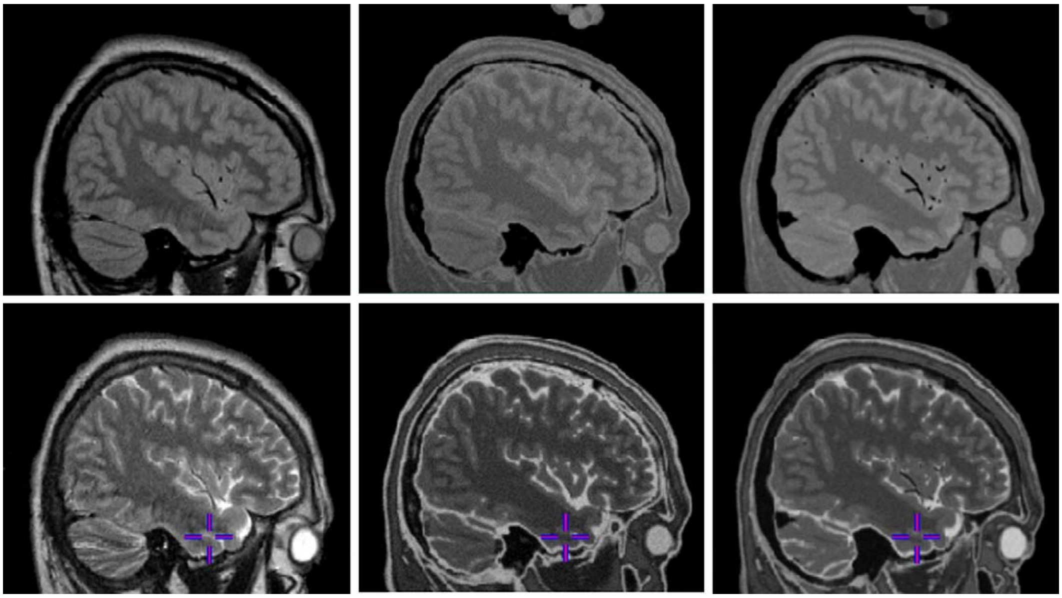

Fig. 4. Comparison between real and simulated images from the previous and the new phantoms. Top row: PD images, bottom row: T2 images. Left: real MRI, center: simulation from previous phantom, right: simulation from new phantom. Note improvements due to vessel class in the region of the insular cortex.

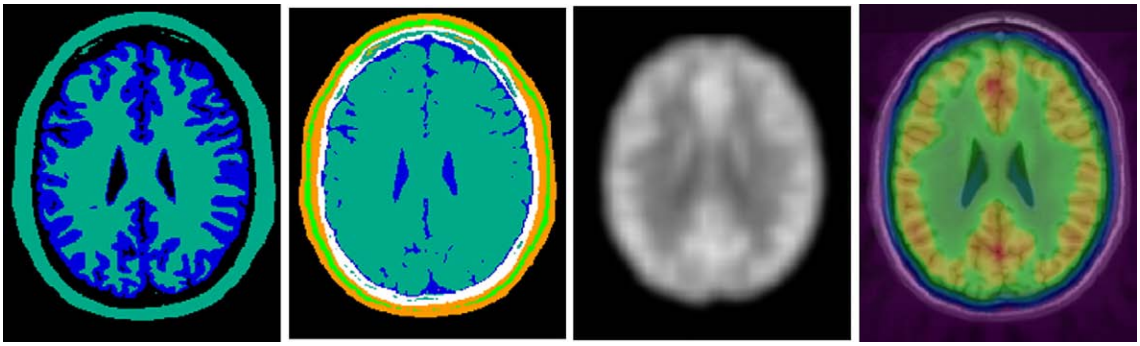

Fig. 5. Transverse slice through an activity image of the simulated FDG PET (A), an attenuation image of the simulated FDG PET (B) and a simulated FDG PET image (C) superimposed on the T1 averaged image (D).

different because of the 1 mm blurring and some small vessels were not detected in the vessels class. This should be taken into account if the MR simulations are used to validate a vessel segmentation method. Second, the CSF class could include non-CSF tissues such as small vessels not detected in the VESSEL class, tiny amounts of dura matter not classified in the DURA class and internal structures such as choroid plexus included in the temporal horns of the lateral ventricles. We should consider these inclusions for classifier performance evaluation on MR simulations when the phantom is used to define truth. Third, partial volume effects could be modified due to the 0.5 mm resampling. It could affect classification algorithms at structure borders or cortical thickness measures. Last, we cannot assert the exact classification of all tissues outside the skull. Some layers are very thin (such as muscles, skin or fat) and could be mixed together. Therefore, the MR simulations are not adapted for the validation of the segmentation of such structures. However, they could be used for the validation of bone segmentation that was improved due to the CT scan.

Another assumption is that each tissue class is homogeneous, with uniform NMR tissue parameters, activity and attenuation values. This assumption could be realistic for fat, muscles or skull classes but could be a limitation with respect to white or gray matter complexity. For example, it is well-known that the cortex is layered, and this layering is not uniform throughout the brain. Cortical structures could be classified in five fundamental types based primarily on the relative development of granular and pyramidal cells (Economo and Koskinas, 1929). A better model could be obtained by taking voxels currently labeled as ‘gray matter’ and reclassifying them into these five gray matter subclasses. Another possible sub-classification could be a lobe-based classification, as some studies have reported significant differences in T1 (Steen et al., 2000), T2 (Wansapura et al., 1999; Zhou et al., 2001) or perfusion (Grova et al., 2005) from one location in the cortex or in the white matter to another. The anatomical variation of the NMR parameters within the brain is the subject of a future research project within our group and will lead to an improved phantom for MRI-specific simulations.

The digital brain phantom II has been used to generate realistic MR, SPECT and PET images. Only a few simulations are shown in this paper to illustrate the possible uses of the phantom. Many different simulations could be computed (PET and SPECT with different tracers, MRI with different sequences…). Some modifications such as adding pathologies or highlighting activations could also be easily performed.

The parameters assigned to each class of the phantom to compute MRI, SPECT and PET simulations were previously published (Reilhac et al., 2005; Grova et al., 2003). For the vessel class, the NMR relaxation properties of the blood (T1, T2, T2* and PD) have to be optimized to take into account flow (for the moment, these parameters are considered as being null in the vessel class). The visual comparison of both simulated and real data provides qualitative evaluation of the simulation realism. A quantitative evaluation and an optimization of these parameters are beyond the scope of this paper which focused on the phantom improvements.

The current anatomical phantom is based on a single brain. While there are many aspects of an image processing method’s performance which can be evaluated within this restriction, it would be desirable to have equivalent simulated data from more

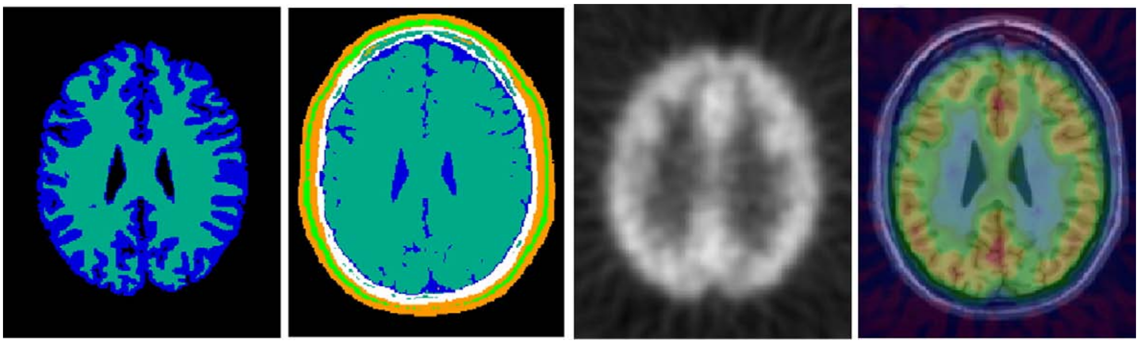

Fig. 6. Transverse slice through an activity image of the simulated HMPAO SPECT (A), an attenuation image of the simulated HMPAO SPECT (B) and a simulated HMPAO SPECT image (C) superimposed on the T1 averaged image (D).

than one subject to take into account inter-subject anatomical variabilities. However, if simulations are to produce scans typical of a population of brains, anatomical inter-subject variabilities must be considered. This can only be done by testing the method to validate on a database of simulations from different anatomical models. In this way, it also possible to validate the robustness of a method in the presence of interindividual anatomical variability. Furthermore, a set of different brains enables not only the study of anatomical difference but also differences in tissue composition that give rise to variability in NMR tissue parameters. It should be useful to model tissue parameter variations across a population of subjects.

The creation of the previous phantom (Collins et al., 1998) required significant manual intervention which made fastidious its application to the creation of a family of phantoms from different subjects. The method presented here requires little intervention and could therefore be applied to the creation of a group of digital phantoms using data from different subjects. This will be the subject of further work in our laboratory.

Conclusion

The digital brain phantom was improved using mostly automated techniques. The new tissue classes added (vessels, dura matter and marrow) improve the realism of the resulting simulations. These data are invaluable as it can be used to drive simulators for different modalities including MRI, fMRI, PET, SPECT and CT. As the method proposed to build the phantom is mostly unsupervised, it will be easily used to create phantoms from other subjects.

References

Arnold, J.B., Liow, J.-S., Sehaper, K.A., Stern, J.J., Sled, J.G., Shattuck, D.W., Worth, A.J., Cohen, M.S., Leahy, R.M., Mazziotta, J.C., Rottenberg, D.A., 2001. Qualitative and quantitative evaluation of six algorithms for correcting intensity nonuniformity effects, NeuroImage, 13 (5), 931–943.

Cardenasa, V.A., Ezekielb, F., Selafanid, V.D., Gomberge, B., Feind, G., 2001. Reliability of tissue volumes and their spatial distribution for segmented magnetic resonance images, Psychiatry Res.: NeuroImaging, 106 (3), 193–205.

Collins, D., P, N., Peters, T., Evans, A., 1994. Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space, J. Comput. Assist. Tomogr., 18 (2), 192–205.

Collins, D., Zijdenbos, A., Kollokian, V., Sled, J., Kabani, N., Holmes, C., Evans, A., 1998. Design and construction of a realistic digital brain phantom, IEEE Trans. Med. Imag., 17 (3), 463–468.

Descoteaux, M., Collins, D.L., Siddiqi, K., 2004. Geometric flows for segmenting vasculature in MRI: theory and validation, Medical Image Computing and Computer-Assisted Intervention Conference, Springer Verlog, St-Malo, Fance, pp. 500–507.

Economo, C., Koskinas, G., 1929. The Cytoarchitectonics of the Human Cerebral Cortex, Oxford Univ. Press, Oxford.

Evans, A., Collins, D., Mills, S., Brown, E., Kelly, R., Peters, T., 1993. 3D statistical neuroanatomieal models from 305 MRI volumes, IEEE Nuclear Science Symposium and Medical Imaging Conference, pp. 1813–1817 (San Francisco, USA).

Grabowskia, T.J., Frankb, R.J., Szumskia, N.R., Browne, C.K., Damasio, H., 2000. Validation of partial tissue segmentation of single-channel magnetic resonance images of the brain, NeuroImage, 12 (6), 640–656.

Grova, C., Jannin, P., Biraben, A., Buvat, I., Benali, H., Bernard, A., Scarabin, J., Gibaud, B., 2003. A methodology for generating normal and pathological brain perfusion spect images for evaluation of mri/spect fusion methods: application in epilepsy, Phys. Med. Biol., 48 (24), 4023–4043.

Grova, C., Jannin, P., Buvat, I., Benali, H., Bansard, J., Biraben, A., Gibaud, B., 2005. From anatomic standardization analysis of perfusion SPECT data to perfusion pattern modeling: evidence of functional networks in healthy subjects and temporal lobe epilepsy patients, Acad. Radiol., 12 (5), 554–565.

Harrison, R., Vannoy, S., Haynor, D., Gillispie, S., Kaplan, M., Lewellen, T., 1993. Preliminary experience with the photon history generator module of a public-domain simulation system for emission tomography, IEEE Nucl. Sci. Symp., 2, 1154–1158.

Holmes, C., Hoge, R., Collins, D., Woods, R., Toga, A., Evans, A., 1998. Enhancement of mr images using registration for signal averaging, J. Comput. Assist. Tomogr., 22, 324–333.

Jenkinson, M., Peehaud, M., Smith, S., 2005. BET2-MR-based estimation of brain, skull and scalp surfaces, International conference on Human Brain Mapping, Toronto, Canada.

Kwan, R.-S., Evans, A., Pike, G., 1999. MRI simulation-based evaluation of image-processing and classification methods, IEEE Trans. Med. Imag., 18 (11), 1085–1097.

Reilhac, A., Lartizien, C., Costes, N., Sans, S., Comtat, C., Gunn, R., Evans, A., 2004. PET-SORTEO: a Monte Carlo-based simulator with high count rate capabilities, IEEE Trans. Nucl. Sci., 51 (11), 46–52.

Reilhac, A., Batan, G., Miehel, C., Grova, C., Tohka, J., Collins, D., Costes, N., Evans, A., 2005. PET-SORTEO: validation and development of database of simulated PET volumes, IEEE Trans. Nucl. Sci., 52 (5), 1321–1328.

Schnack, H.G., Pol, H.E.H., Baare, W.F.C., Staal, W.G., Viergever, M.A., Kahn, R.S., 2001. Automated separation of gray and white matter from mr images of the human brain, NeuroImage, 13 (1), 230–237.

Sled, J., Zijdenbos, A., Evans, A., 1998. A nonparametric method for automatic correction of intensity nonuniformity in MRI data, IEEE Trans. Med. Imag., 17 (l), 87–97.

Smith, S., 2002. Fast robust automated brain extraction, Hum. Brain Mapp., 17, 143–155.

Steen, R., Reddick, W., Ogg, R., 2000. More than meets the eye: significant regional heterogeneity in human cortical t1, Magn. Reson. Imaging, 18 (4), 361–368.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the MNI MRI single-subject brain, NeuroImage, 15 (l), 273–289.

Vasilevskiy, A., Siddiqi, K., 2002. Flux maximizing geometric flows, IEEE Trans. Pattern Anal. Mach. Intell., 24 (12), 1565–1578.

Wansapura, J., Holland, S., Dunn, R., Ball, W., 1999. Nmr relaxation times in the human brain at 3.0 tesla, J. Magn. Reson. Imaging, 9 (4), 531–538.

Zhou, J., Golay, X., Zijl, P.V., Silvennoinen, M., Kauppinen, R., Pekar, J., Kraut, M., 2001. Inverse t2 contrast at 1.5 t between gray matter and white matter in the occipital lobe of normal adult human brain, Magn. Reson. Med., 46 (2), 401–406.