Chapter 7

Automatic 3D regional MRI segmentation and statistical probability anatomy maps

AC Evans, DL Collins, CJ Holmes
Montreal Neurological Institute, McGill University
McConnell Brain Imaging Centre
3801, University St., Montreal, Canada H3A 2B4

Abstract

In the rapidly-growing field of human brain mapping, the use of a standardized, or stereotaxic, coordinate space is a central factor in the interpretation and comparison of results from different centers. The use of some variant of the Talairach space is almost universal but this procedure has a number of weaknesses arising from the facts that it (i) is derived from an unrepresentative single 60-yr old female brain, (ii) ignores left-right hemispheric differences, (iii) has variable slice separation, up to 4mm. We have previously reported the creation of a 305-member, finely-sampled averaged MRI volume which can be used alongside the Talairach atlas. Although visually indicative of the anatomical variability across the young normal population usually scanned in brain mapping experiments, this dataset did not express that variability quantitatively. Such information requires the accurate and reproducible labelling of all voxels in any given structure for each brain within the MRI ensemble. Manual labelling is prohibitively time-consuming and error-prone. We therefore describe an automated procedure for 3D labelling of individual brain structures based on non-linear deformation of individual MRI volumes to match a previously-labelled target volume, followed by numerical inverse transformation of the labels to the native MRI space. Subsequent processing allows these labels to be mapped into stereotaxic space, now using a simple linear transformation, and combined to generate probability maps for the location of each pre-defined structure. The procedure is described along with validation experiments and selected examples of such maps. The general principle described is applicable to any scale and any object which can be defined within the target volume. Importantly, since the labels themselves take no part in the image-matching process, any number of labelling-schemes, representing different brain atlas formulations, can co-exist and their corresponding probability maps generated from a single determination of the mapping transformation.

7.1 Introduction

The quantitative comparison of functional and structural brain image data across subjects requires a common, so-called stereotaxic, coordinate frame with respect to which the spatial variability of corresponding features from different individuals can be expressed. Since that positional variability will depend on the degrees of freedom (DOF) allowed for mapping the image data from its original, native space into stereotaxic space, it is essential that the stereotaxic transformation be well-defined and reproducible. For simplicity and ease of use of interpretation, it is also desirable that the transformation have as few DOF as possible while accounting for most of the positional variability. The problem of **quantifying** anatomical variability is thus considered to be distinct from the practical issue of **removing** anatomical variability. In the latter task, the goal is to render all brains identical after transformation, a goal which, in principle, could require as many DOF as there are independent samples in the data volume. Numerous non-linear methods have been employed to increase the stereotaxic correspondence among different brains .e.g.[Evans et al., 1991, Roland et al., 1994, Minoshima et al., 1994] Unfortunately, since they vary in the number of effective DOF, any direct comparison of residual anatomical variability is difficult, if not impossible, and quotations for anatomical variability have little absolute meaning.

A related but separate problem is the accurate identification, or labelling, of 'correspondent' points which are to be optimally aligned within the constraints of the the stereotaxic transformation's allowed DOF. There is never an exact one-to-one mapping between two brains even at the gross morphological level owing to, for instance, multiple sulcal patterns while at finer spatial scale the notion of correspondence becomes meaningless. The quantification of neuroanatomical variability is then an empirical process which is useful provided the positional uncertainty in labelling any point is small compared with the true positional variability among individuals. This uncertainty/variability trade-off varies widely, with deep grey structures and ventricular features being much easier to define than cortical features. Typically, structural labelling has been a manual process which, while likely to be more accurate than automated procedures for any one version of labelling within a single brain, will less precise (reproducible) within that brain, less systematic across brains and more time-consuming. Also, the labelling scheme will differ among investigators, introducing further methodological inconsistencies which will mask the true neuroanatomical variability being sought.

We have considered the above issues in defining a strategy for the creation of a probabilistic atlas of human neuroanatomy. The strategy uses automated techniques for 3-D regional image segmentation and non-linear volume deformation and comprises two conceptual stages **LABELLING** and **STEREOTAXY**, as schematically illustrated in Fig. 7.1 . For the **LABELLING** stage, the native data is matched to a previously-labelled target volume already resident in stereotaxic space using a 3-D non-linear deformation described below. The inverse transformation is used to map the voxel labels, defined on the target volume, back on to the native volume. Each newly-labelled native volume can than be transformed into some simply defined coordinate frame such that the spatial variability in size, shape, or centroid location of any individual structure with respect to that coordinate frame/transformation pair can be quantified in the **STEREOTAXY** stage. The first stage allows unlimited DOF in matching a labelling template, or atlas, with the native image volume. The second allows only linear stereotaxic transformation of the labelled volume when quantifying anatomical variability. Note that it is not essential for the labelling process that the target volume and its associated labels be resident in stereotaxic space, merely convenient for purposes of implementation and later image comparison. The two perspectives, linear **STEREOTAXY** and non-linear **LABELLING** are combined within the concept of probabilistic neuroanatomical atlases. Here, large ensembles of brain image data are first labelled in their native space and then mapped into some stereotaxic space using a limited number of degrees of freedom, e.g, 9- or 12-parameter affine transform. The likelihood of any structure label occurring at any stereotaxic voxel gives rise to a statistical probability anatomy map (SPAM) for each structure which, when expressed for all structures, gives rise to a probabilistic anatomical atlas. [1]

The perspective of segmentation as registration followed by delineation yields one powerful advantage, that of atlas-independent segmentation. The geometric contours defined on the target image volume are not used to determine the match between source and target volumes. The non-linear transformation is established by matching the two volumes directly. Therefore, any atlas defined in the target space can be used for segmentation. this allows for the co-existence of multiple atlases, each of which is simultaneously mappable to the native MR image volume without the need to recalculate the non-linear matching transformation. The underlying structure of the MRI is used as a frame, onto which the atlas contours are mapped. Thus, any structure can be defined in the atlas and need not follow edges present in the tomographic image volume. For example, functional neuroanatomy areas, whose limits can be inferred from the underlying anatomical substrate, can be incorporated into the atlas definitions with affecting the spatial transformation defined by the algorithm.

7.2 Methods

7.2.1 Algorithm

The procedure for automatic segmentation, termed ANIMAL (Automated Nonlinear Image Matching and Anatomical Labelling) [Collins et al., 1995a, Collins et al., 1995b], builds up a 3D non-linear deformation field by sequential linear matching of many small data cubes in one volume to the other volume. These cubes are arranged in a 3-D grid to fill the volume and each cube can only move within a range defined by the grid spacing. The algorithm is applied iteratively in a multi-scale hierarchy, so that image blurring and cube size are reduced after each grid pass. The initial fit is obtained rapidly since only lower spatial frequencies and gross distortions are considered but later iterations at finer scales become increasingly computationally intensive as finer details are matched. ANIMAL can be terminated at any intermediate spatial scale if faster, approximate results are satisfactory. The algorithm is an extension of a previously reported linear method for stereotaxic normalization [Collins et al., 1994]. Here we concentrate on the use of this non-linear extension for automatic labelling of brain structures. Both linear and non-linear registration procedures rely on the extraction of geometrically invariant features from the original volumetric images. We have chosen to use blurred versions of the image intensity and image gradient magnitude as features that are independent of the original position, scale and orientation of one data set with respect to the other. These features are calculated by convolution of the original data with zeroth and first order 3D isotropic Gaussian derivatives. Convolution with such an operator maintains linearity, shift-invariance and rotational-invariance in the detection of features [tar Haar Romeny et al., 1991, Koenderink and van Doorn, 1987]. We use the full-width-half-max FWHM $= 2.35\sigma$) of this Gaussian kernel as the parameter to measure the spatial scale.

The optimization procedure used to recover the best spatial transformation is accomplished in a hierarchical multi-scale fashion. Registration is performed at different spatial scales, starting with very blurred data and increasing detail at each step by using less blurred images, refining the registration at each stage. As well as increasing the speed of calculation, compared with performing all operations at the finest scale, this strategy avoids local minima in the objective function hyper-surface. The linear component of the

[1] The above rationale does not conflict with the notion of non-linear stereotaxic mapping. The latter has the pragmatic goal of improved alignment of individual morphological features in the pursuit of increased signal from multi-subject functional activation experiments rather than the quantitative characterization of individual anatomical differences being discussed here

required spatial mapping is found by estimating the best affine transformation matrix that maximizes the similarity of invariant features extracted from both target and source volumes and is described in detail in [Collins et al., 1994]. The feature similarity criterion used can be selected from a set including a) cross-correlation, b) minimum variance of feature ratio [Woods et al., 1993], c) z-score. All results reported here used cross-correlation as the similarity criterion. The remaining non-linear morphological differences are again estimated with a hierarchical multi-scale strategy. The approach is an extension of the linear registration method, except that now the target volume is a small cubic neighbourhood of the whole brain. Also, only 3 translational parameters are varied to obtain optimal neighbourhood similarity, using a Simplex optimization procedure. The process is applied to all parts of the brain volume, with successive cubes being recursively selected by stepping through every node of the entire volume in a 3-D grid pattern [Collins et al., 1995a, Collins et al., 1995b]. The size of the cube at each scale step is set at 1.5 FWHM with a node-spacing set at 0.5 FWHM in each direction. After each scale step, the 3 translational parameters at each node define a 3D non-linear deformation field representing the non-linear component of inter-subject neuroanatomical variability.

7.2.2 Validation

The performance of the ANIMAL algorithm has previously been validated against experimental or digital brain phantom data and manual labelling of MRI data [Collins et al., 1995a, Sorlie et al., 1995]. Some representative work, included for illustration, used a digital brain phantom created by identifying different structures and tissue types from an MR data set of a single volunteer. The volume is sampled on a $1 \times 1 \times 1\text{mm}^3$ matrix, and covers the entire head. The resulting labeled volume could then be endowed with the appropriate contrast for each tissue type, and Rayleigh noise added, to resemble a true MRI acquisition. This 3D data set provides a gold-standard with which to evaluate the linear registration algorithm. It should be noted that it does not matter that some voxels from the original MRI data may have been mis-labeled. The labelled volume is now a phantom defined as truth which, while no different in principle than a simple geometric phantom, closely mimics the convoluted nature of a real brain and whose intensities and noise characteristics match those dealt with in MRI. An arbitary 3D distortion was applied to the T_1-weighted data using a thin-plate spline (TPS) transformation to produce a topologically equivalent, but spatially warped target data set. Ten percent noise, corresponding to the typical signal-to-noise ratio of our MRI scanner, was added to the source volume. In addition to the voxels classified as grey or white matter, the voxels contained in five anatomical structures (head of left and right caudate, lateral and third ventricles) were manually identified on the original unwarped phantom data. Application of the known TPS transformation to the labelled voxels identified corresponding voxels in the TPS-warped volume. ANIMAL was then applied to map the original volume to the TPS-warped volume with the goal of identifying the same voxels in the TPS-warped volume. The results were evaluated by calculating the percent difference, δ, and the percent overlap, Δ, between the known voxels within each TPS-warped structure and those identified by ANIMAL.

7.2.3 Statistical Probability Anatomy Maps (SPAMs)

Comparison with Manual Labelling: As described above, a primary motivation for the development of the ANIMAL algorithm is the need for automatic labelling of individual brain structures in each of a set of 3D MRI volumes, and ultimately the automatic generation of probability maps for each structure. The utility of the algorithm for this was evaluated by statistical comparison of SPAMs for the frontal cortex labelled in 10 individual MRI volumes, generated by manual labelling or by ANIMAL. Individual volumes were collected as 160 1-mm thick contiguous slices with a 3D gradient-echo T1-weighted sequence (T_R/T_E=18/10 msec). Manual labelling was performed using a triplane display with real-time update of transverse, sagittal and coronal planes through a manually-driven 3D cursor. At each voxel, the manual and automated SPAMs were compared by calculation of a z-statistic of the form:-

$$z = \left| \frac{P_a - P_b}{\sqrt{P(1-P)(1/N_a + 1/N_b)}} \right| \tag{7.1}$$

where P_a, P_b and N_a, N_b are probabilities and subject numbers at each voxel and,

$$P = \frac{N_a P_a + N_b P_b}{N_a + N_b} \tag{7.2}$$

Generation of SPAMs for multiple regions: ANIMAL was applied as illustrated in Fig. 7.1 to a total of 45 normal MRI volumes, obtained from young normal volunteers using the same pulse sequence. The 3D spatial transformation between each volume and the target volume was first obtained. A set of 3D neuroanatomical regions, both deep grey structures and lobar cortex zones, were labelled within the target volume by a trained neuroanatomist (CJH). The inverse spatial transformations were applied to label each volume in its native space. Finally, a 9-parameter linear stereotaxic transformation was applied to the individually-labelled MRI volumes, now using a 305-MRI average MRI volume as the target for ANIMAL, running in linear mode [Evans et al., 1992, Evans et al., 1993, Evans et al., 1994, Collins et al., 1994].

7.3 Results

As Fig. 7.2 shows, the structures in the TPS-warped data are outlined well by the ANIMAL-warped contours, when compared to a simple 9-parameter linear transformation. Since a preclassified data set was used to create the brain phantom, the set of grey-matter and white-matter voxels were known and the overlap indices could be evaluated in the same way for the tissue classes as for the brain structures. Table 7.1 shows that on average, δ was reduced by more than half, from 4.2% to 1.7%, and Δ was improved 38%, from 70.4% to 97.0% when comparing non-linear to linear matching.

Fig. 7.3 compares the the manual and automated SPAMs for the frontal cortex from 10 subjects. Fig. 7.4 shows the statistical z-field expressing significant differences between the two, thresholded at z=2.0 and z=3.0. The z-field is shown after convolution of the original SPAMs with a 2mm FWHM 3D Gaussian kernel, since the z-field should carry no signal spatial frequencies higher than the cut-off resolution used for the ANIMAL algorithm. The deformation field derived for each source volume applies equivalently to all internal structures. Hence, and any new label can be defined in the target space and used to label the corresponding structure in each source MRI volume without a time-consuming re-calculation of the deformation field. Figs 7.5, 7.6 and 7.7 show SPAMs for frontal cortex, temporal cortex and caudate nucleus, each derived from application of the procedure outlined in Fig. 7.1 to forty-five MRI volumes.

7.4 Discussion

The automatic segmentation method presented here reverses the standard segmentation strategy, from the matching of geometric contours directly to image data to regional labelling through non-linear image deformation. The structures identified by this method are comparable to those segmented manually and the experimental results on both simulated and real MRI data have shown that the method is accurate and robust (see Figs 7.2, 7.3 and 7.4). The ANIMAL algorithm allows a mapping of any previously-defined segmentation scheme on to individual MRI volumes so that this information can be obtained in a reproducible fashion. Once the deformation field has been defined, it is possible to use the strategy outlined in Fig. 7.1 to generate any number of SPAMs for structures pre-defined in the target volume as illustrated by the selected examples in Figs 7.5, 7.6 and 7.7. It must be acknowledged that the automated approach may never be able to match human capabilities for outlining the finest details of individual brain structures, particularly in areas of high variance for location, sulcal pattern or surface detail. It remains to be seen whether the objectivity and precision of an automated approach outweighs this decreased ability to delineate fine structure when investigating possible group differences between normal subjects or patients.

The goal of a probabilistic 3-D description of normal neuroanatomy will take many years to achieve. It is most probable that the major components, i.e. the target volumes/features, the matching/segmentation procedure and the database being segmented will all be repeatedly updated during that time. However, technical improvements in computers and brain imaging equipment combined with the knowledge gained at each iteration should produce an accelerating cycle. The work described here is part of a broader program to create a probabilistic atlas of human neuroanatomy by the International Consortium for Brain Mapping (ICBM)[Mazziotta et al., 1995]. This effort will combine large numbers of normal MRI volumes with smaller numbers of high-resolution (100 micron) macrocryotome volumes in a stereotaxic space. These data will be labelled anatomically in different ways and statistical maps of structural variance will be generated in ways similar to that described here. Functional information will be added as voxel attributes as well as pointers to non-voxel databases. The ICBM perspective is that any single segmentation scheme, or atlas, will always be unacceptable to some sectors of the scientific community or become obsolete with time. Therefore it is vital that an automated regional labelling is independent of the specific form of parcellation employed so that many atlases can be applied simultaneously to any given brain. The segmentation strategy employed with the ANIMAL algorithm offers that independence.

Acknowledgments

We wish to thank Dr. Keith Worsley for advice on statistical matters as well as Peter Neelin, David MacDonald and Greg Ward for help in data processing. The authors would like to express their appreciation for support from the Canadian Medical Research Council (SP-30), the McDonnell-Pew Cognitive Neuroscience Center Program, the U.S. Human Brain Project (HBP), NIMH and NIDA. This work forms part of a continuing project of the HBP-funded International Consortium for Brain Mapping (ICBM) to develop a probabilistic atlas of human neuroanatomy.

Bibliography

[Collins et al., 1995a] Collins, D., Holmes, C., Peters, T., and Evans, A. (1995a). Automatic 3d model-based neuroanatomical segmentation. *Human Brain Mapping (in press)*.

[Collins et al., 1995b] Collins, D. L., Evans, A. C., Holmes, C., and Peters, T. M. (1995b). Automatic 3D segmentation of neuro-anatomical structures from MRI. In Bizais, Y., Barillot, C., and Paola, R. D., editors, *Information Processing in Medical Imaging (IPMI)*, pages 139–152. Kluwer.

[Collins et al., 1994] Collins, D. L., Neelin, P., Peters, T. M., and Evans, A. C. (1994). Automatic 3D inter-subject registration of MR volumetric data in standardized talairach space. *Journal of Computer Assisted Tomography*, 18(2):192–205.

[Evans et al., 1993] Evans, A. C., Collins, D. L., Mills, S. R., Brown, E. D., Kelly, R. L., and Peters, T. M. (1993). 3d statistical neuroanatomical models from 305 mri volumes. In *Proc. IEEE-Nuclear Science Symposium and Medical Imaging Conference*, pages 1813–1817.

[Evans et al., 1992] Evans, A. C., Collins, D. L., and Milner, B. (1992). An MRI-based stereotactic atlas from 250 young normal subjects. *Soc.Neurosci.Abstr.*, 18:408.

[Evans et al., 1991] Evans, A. C., Dai, W., Collins, D. L., Neelin, P., and Marrett, T. (1991). Warping of a computerized 3-D atlas to match brain image volumes for quantitative neuroanatomical and functional analysis. In *Proceedings of the International Society of Optical Engineering: Medical Imaging V*, volume 1445, San Jose, California. SPIE.

[Evans et al., 1994] Evans, A. C., Kamber, M., Collins, D. L., and MacDonald, D. (1994). An mri-based probabilistic atlas of neuroanatomy. In Shorvon, S., editor, *Magnetic Resonance Scanning and Epilepsy*, pages 263–274. Plenum Press, New York.

[Koenderink and van Doorn, 1987] Koenderink, J. and van Doorn, A. (1987). Representation of local geometry in the visual system. *Biol. Cybern.*, 55:367–375.

[Mazziotta et al., 1995] Mazziotta, J. C., Toga, A. W., Evans, A. C., Fox, P. T., and Lancaster, J. (1995). A probabilistic atlas of the human brain: Theory and rationale for its development. *Neuroimage*, 2:89–101.

[Minoshima et al., 1994] Minoshima, S., Koeppe, R., Frey, K., and Kuhl, D. (1994). Anatomic standardization: linear scaling and nonlinear warping of functional brain images. *Journal of Nuclear Medicine*, 35(9):1528–37.

[Roland et al., 1994] Roland, P., Graufelds, C., Wahlin, J., Ingelman, L., Andersson, M., Ledberg, A., Pedersen, J., Ackerman, S., Dabringhaus, A., and Zilles, K. (1994). Human brain atlas: For high-resolution functional and anatomical mapping. *Human Brain Mapping*, pages 173 – 184.

structure	*linear* δ	*linear* Δ	non-linear δ	non-linear Δ
caudate:	1.3%	65.7%	4.9%	96.4%
ventricle:	7.0%	77.6%	-0.0%	99.1%
grey:	3.4%	72.0%	1.0%	97.0%
white:	5.0%	66.1%	-0.7%	95.6%
average:	**4.2%**	**70.4%**	**1.7%**	**97.0%**

Table 7.1: Segmentation results for brain phantom. This table shows the percent volume difference, δ, and the percent volume overlap, Δ, for the individual structures defined on the digital brain phantom and segmented from a manually warped data set. These values are normalized to the true structure volume.

[Sorlie et al., 1995] Sorlie, C., Collins, D., Worsley, K., and Evans, A. (1995). Neuroanatomical variability: A quantitative study of young normal brains in stereotaxic space. *Human Brain Mapping (in press).*

[tar Haar Romeny et al., 1991] tar Haar Romeny, B., Florack, L. M., Koenderink, J. J., and Viergever, M. A. (1991). Scale space: its natural operators and differential invariants. In Colchester, A. C. F. and Hawkes, D. J., editors, *Information Processing in Medical Imaging*, page 239, Wye, UK. IPMI.

[Woods et al., 1993] Woods, R., Mazziotta, J., and Cherry, S. (1993). MRI-PET registration with automated algorithm. *Journal of Computer Assisted Tomography*, 17:536–546.

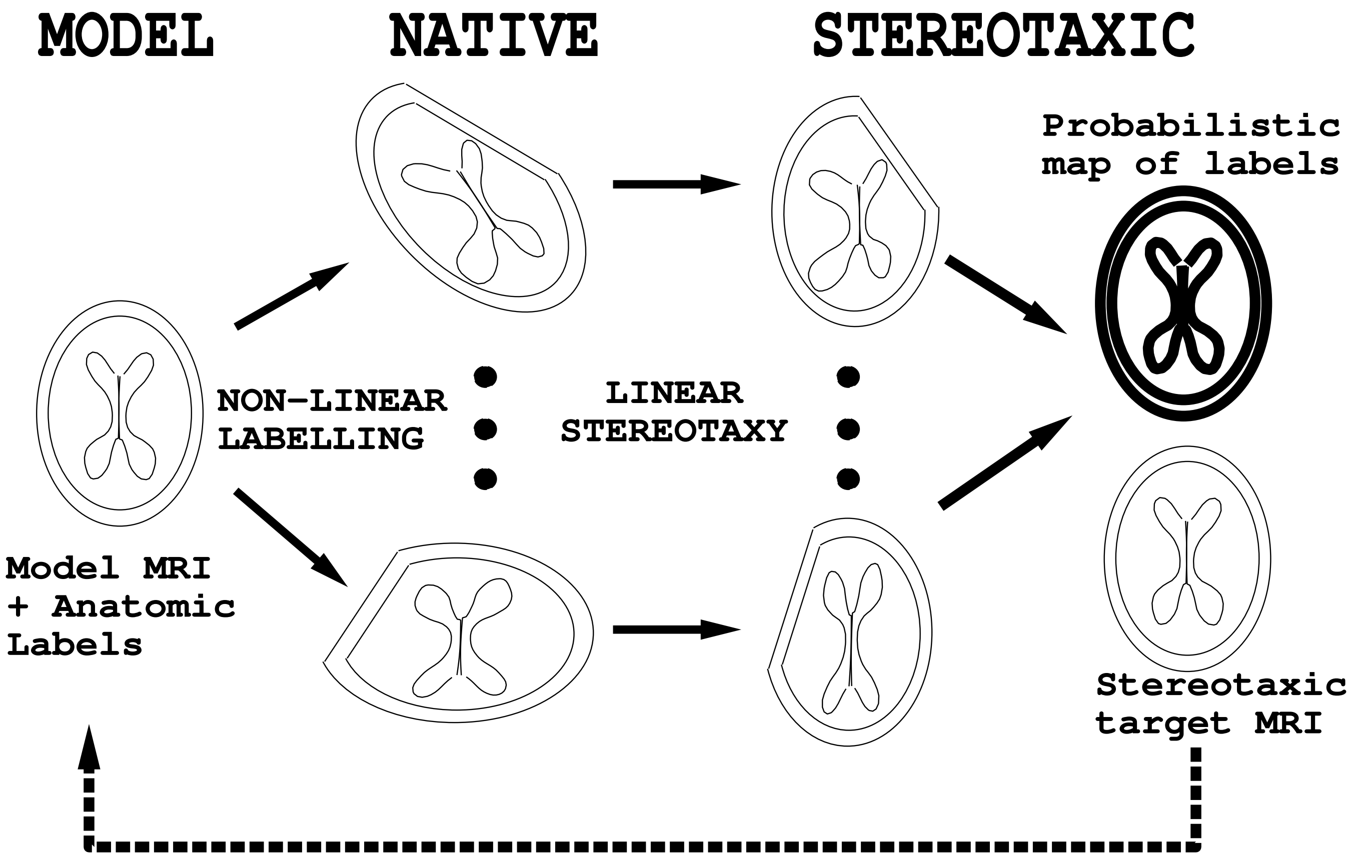

Figure 7.1: Schematic of principles underlying creation of probabilistic neuroanatomy maps

The LABELLING and STEREOTAXIC steps can be considered to be independent operations. The result of the LABELLING step is the transfer of anatomic labels from a target space to the native space of the source dataset. A subsequent stereotaxic mapping can be applied to investigate the spatial variability of those labels with respect to a well-defined coordinate frame. In practice, as indicated by the dotted line, the target volume and the stereotaxic target volume are in the same space although this is not strictly necessary. Also in practice, a 9-parameter transformation is applied for stereotaxic mapping and probabilistic map creation.

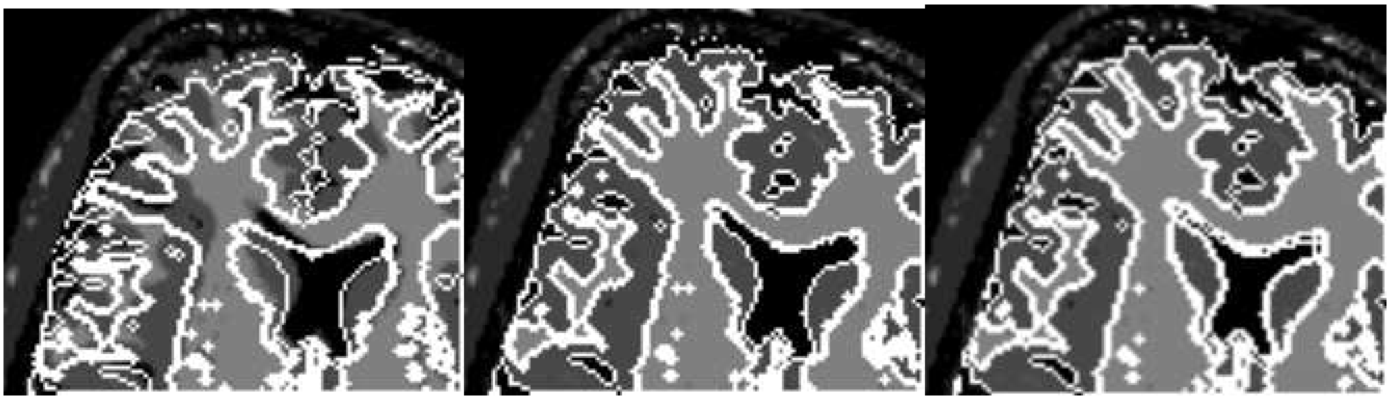

Figure 7.2: Segmentation of digital brain phantom

Three transverse images of the 3D digital brain phantom through the ventricles at the level of the basal ganglia are shown here with the contours of the resulting segmentation superposed. The underlying image, the same in each case, was derived by application of a 3D thin-plate spline (TPS) deformation to the initial brain phantom volume. The contours, defined initially on the unwarped volume, were then mapped to the warped volume in three ways. They correspond to the head of caudate, the lateral ventricles and the grey and white matter boundaries. On the left, the contours are mapped through the linear transformation only. In the center are the correctly-warped contours, generated by mapping the original atlas contours through the known TPS transformation. On the right, are the contours resulting from the ANIMAL segmentation. Almost no difference is visible between the two rightmost sets of contours.

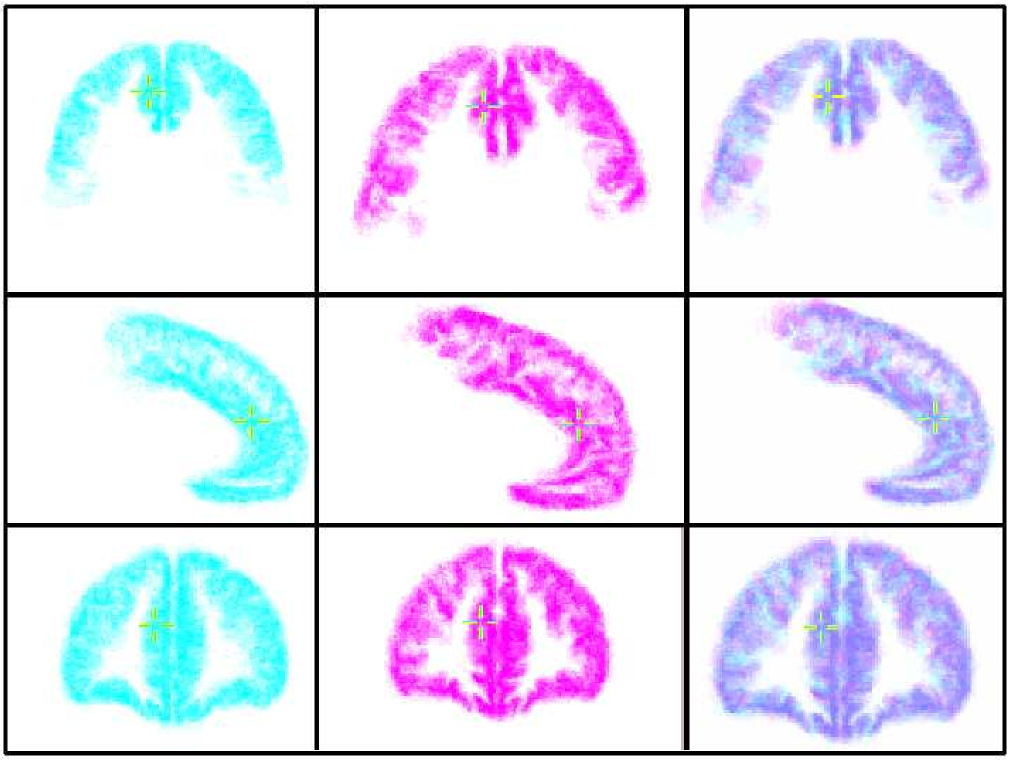

Figure 7.3: Visual comparison between manual and automic generation of probabilistic frontal cortex from 10 brains

The left column shows three orthogonal cuts through the probabilistic map for frontal cortex derived from manual labelling of frontal cortex in 10 MRI volumes from young normal volunteers. The middle column shows corresponding cuts through the ANIMAL-generated equivalent. The right column shows the colour overlay of the two (Note: that each panel has independent pan/zoom capability and so images in the left and middle columns may look incompatible. When images are combined, as in the rightmost column, pan/zoom parameters are identical).

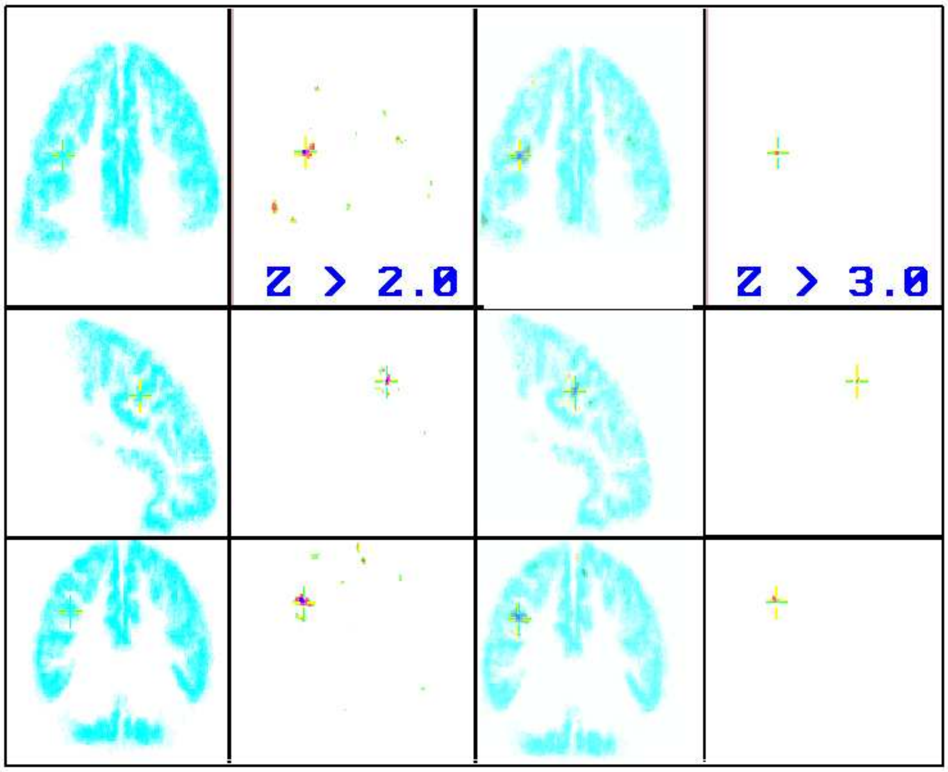

Figure 7.4: z-statistic map for differences between manual and automatic generation of probabilistic frontal cortex from 10 brains

Statistical comparison of manual- and ANIMAL-generated probability maps for frontal cortex where the z-statistic used is defined in equation 1. The z-map, smoothed with a 2mm-FWHM 3D-Gaussian kernel, is thresholded at z=2.0 and z=3.0. Interleaved columns show the ANIMAL-generated probability map for comparison. The 3D cursor, defining the intersection of the three orthogonal planes displayed, is located on the single region exceeding the z=2.0 threshold. No regions in the z-field exceeded z=3.0.

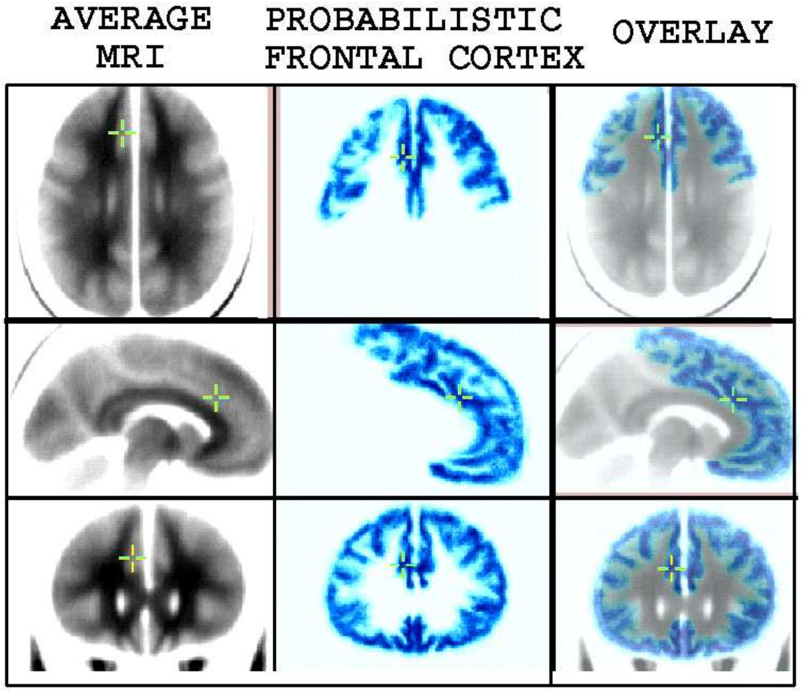

Figure 7.5: Probabilistic frontal cortex from 45 brains

Triplane display through the ANIMAL-generated probability map, derived from 45 individual MRI volumes, for frontal cortex (middle column), the 305-member average MRI (left column) and the overlay of the two (righ column). Note pan/zoom comments for Fig. 7.3

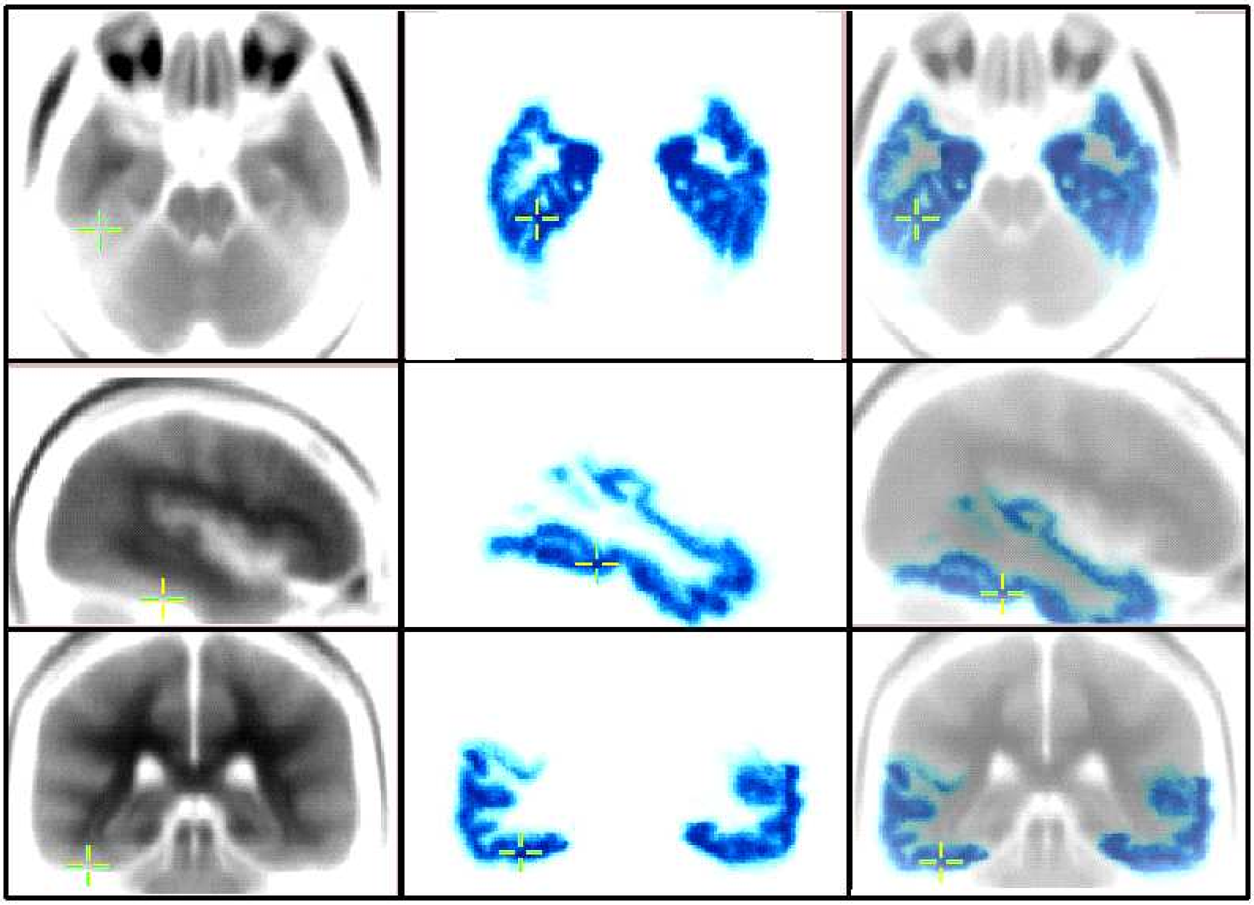

Figure 7.6: Probabilistic temporal cortex from 45 brains

Triplane display through the ANIMAL-generated probability map, derived from 45 individual MRI volumes, for temporal cortex (middle column), the 305-member average MRI (left column) and the overlay of the two (righ column). Note pan/zoom comments for Fig. 7.3

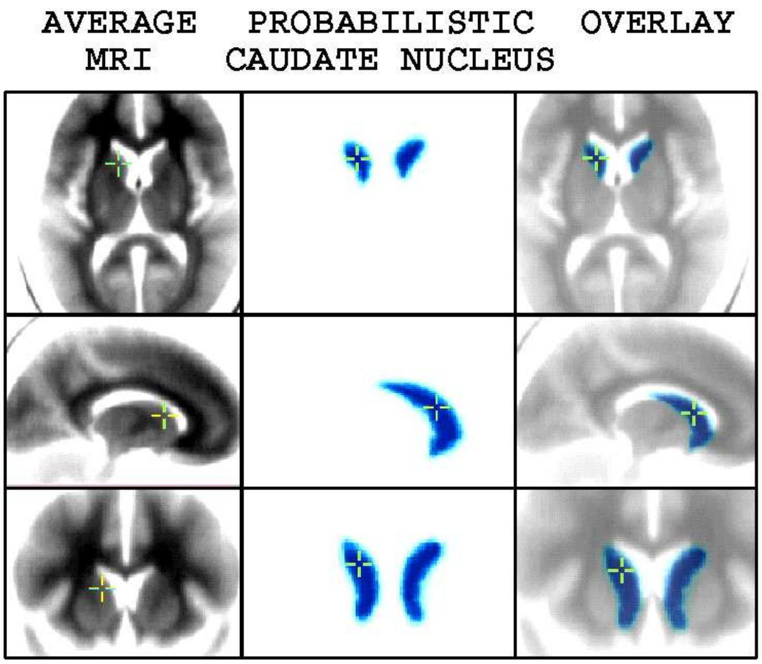

Figure 7.7: Probabilistic caudate nucleus from 45 brains

Triplane display through the ANIMAL-generated probability map, derived from 45 individual MRI volumes, for the caudate nucleus (middle column), the 305-member average MRI (left column) and the overlay of the two (right column). Note pan/zoom comments for Fig. 7.3