Stratified Voxel-Based Morphometry (sVBM)

M. Jorge Cardoso[1], Ivor Simpson[2], Marc Modat[1], and Sebastien Ourselin[1,2]

[1] Centre for Medical Image Computing (CMIC), University College London, UK,
[2] Dementia Research Centre (DRC), University College London, UK

Abstract. In neuroimaging, voxel based morphometry (VBM) has been a valuable tool of identifying brain-wide differences between populations. One of the key elements of VBM is to define a space for voxelwise comparisons. However, errors in this mapping to common space and variations of brain morphology, both natural and pathologic can result in false positives. In this work we explore a new framework, where a spatially varying morphological similarity graph is created between pairs of images. This graph is then used to stratify natural and pathological variability in a VBM-like setting. In contrast to VBM, which describes the group differences on an average brain morphology, sVBM describes how different brain morphologies are independently affected by pathology. Due to its pairwise nature, this technique provides smoother and better localised differences between populations, possibly providing novel insights into the homogeneity of pathological effects for different brain morphologies.

1 Introduction

The characterisation of morphometric differences between healthy and pathological populations is one of the cornerstones of medical imaging. Brain morphometry has been extensively studied in diverse populations, ranging from dementia (e.g. Alzheimer's and Huntington's disease) and schizophrenia, to autism and even normal ageing.

The morphometric characterisation of populations explores differences in volume, mass, shape, cortical thickness and tissue density. Due to the advent of high resolution anatomical imaging and increased computational power, voxel-based techniques have become the *de facto* tool for morphometric analysis. These techniques range from the classic voxel-based morphometry [1], characterising differences in tissue density, to tensor- [2] and deformation-based [3] morphometry, characterising differences in the mapping tensor and deformation parameters respectively. Voxel based morphometry (VBM) involves a voxel-wise comparison of the local concentration of grey matter between two groups of subjects when mapped to a common space. In order to reduce the bias towards the choice of template, a mean shape/appearance space, known as the groupwise space, is often used for the comparison[4]. Despite their seeming optimality for morphometric comparison, groupwise spaces suffer from three main problems: (1) their construction is highly dependent on the choice of image similarity metric and regularisation [5]; (2) the mapping errors to the groupwise space can result

Fig. 1. The "Apples and Oranges" painting by Paul Cézanne, illustrating the discrimination complexity between different but similar objects.

in morphometric mismatch, a problem which has generated wide criticism [6, 7]; (3) and groupwise spaces mix naturally occourring morphologies (e.g. sulcal patterns, brain shape) with pathological (e.g. atrophy) effects. This work focuses on the latter effect.

Some of the above effects have been studied to some degree. Lepore *et al.* [8] and Koikkalainen *et al.* [9] have looked at multi-template TBM but have not explored the stratified space statistics. Other techniques based on pattern-based, feature-based morphometry, patch-based morphometry have also been used to overcome some of the VBM/TBM limitations by attempting to improve the power to detect pathological differences at the population level. However, these techniques do not explore how pathology affects each different brain morphology independently.

By mixing pathological effects with different morphologies, the VBM analysis is to some degree "comparing apples to oranges", i.e. when two items or groups of items cannot or should not be meaningfully compared (see Fig. 1). The use of a groupwise space for morphometric comparison has two underlying assumptions: first, that pathology affects all morphologies in a similar manner, i.e. it assumes that the pathological and natural variability in the high dimensional space of all brain morphologies are linearly separable, even though it is well known that different morphologies have different connectivity and functional activation patterns [10]; second, that the point estimate of the mean and variance of the population tissue densities assumes a monomodal distribution and accurately summarises the full population (see Fig. 2-right), i.e. VBM combines all modulated tissue densities in a groupwise space and then obtains a point estimate of the difference in tissue density between populations. However, certain phenotypes might have protective or deleterious effects on pathology, indicating that the tissue density differences might not be monomodal.

Some groups have explored the idea of stratifying populations in different subgroups for the purpose of segmentation [11, 12]. However, this stratification

normally attempts to separate the groups according to their pathological status, not taking into account the fact that there are different non-pathology-related morphological subgroups. Subsequently, further stratification into different local morphological subgroups is necessary. In the limit, this population stratification process considers each subject's morphological subgroup as a stratum. Interestingly, this limit situation can be interpreted as the same problem solved by the multi-atlas propagation and label fusion community [13, 14], where the morphological similarity between subjects is used as a hub for label propagation. This stratification strategy was used in [15] for the purpose of classification, producing one of the highest reported accuracy results for AD prediction.

This work explores a novel approach for stratified voxel-based morphometry (sVBM). We re-interpret the question behind VBM as a morphology specific process. This approach is built on the Geodesic Information Flows (GIF) implicit-graph construct proposed by Cardoso *et al.* [16], which stratifies the subjects in a population with the use of an implicit spatially variant graph (see Fig. 2-right). This stratified space is then used for morphometric comparison. Overall, this work attempts to answer the question: how is each brain morphology affected by pathology?

2 Methods

This section will first introduce the mathematical framework and the undirected graph for geodesic information flow (GIF), followed by the introduction of the morphological distance metric between images. The GIF framework will then be used for voxel-based tissue density comparison in order to separate naturally-occurring from pathological effects. These stratified effects are then combined in a groupwise space for population analysis.

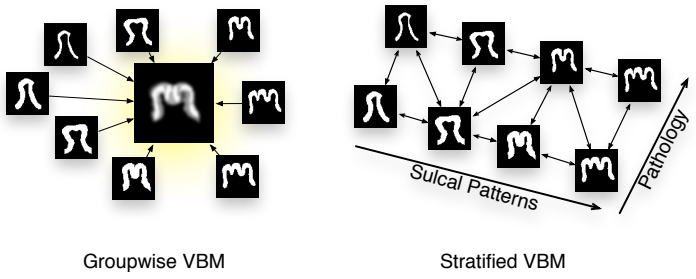

Fig. 2. Left) Representation of how the tissue density flows between subjects in a groupwise approach, representing the average shape and appearance of the population. Right) In a pairwise approach, a spatially varying implicit graph is constructed. At a certain voxel location, the neighbourhood of each subject is dependent on the morphological similarity.

2.1 The geodesic information flows framework

Let a set Y of N images be the full set of observed T1-weighted MRI data with the i-th image of this set denoted by Y_i. Each image Y_i is a vector of size L_i, with the voxel $\boldsymbol{v}$ denoted by $Y_i(\boldsymbol{v})$. In this framework, all images in the database need to be independently non-rigidly mapped to every other image (pairwise registration). The mapping between image Y_i and image Y_j at voxel $\boldsymbol{v}$ is denoted by $T_{i\rightarrow j}(\boldsymbol{v})$. As the framework requires a symmetric, inverse consistent and diffeomorphic coordinate transformation, we use a symmetric variant of a non-rigid free-form registration algorithm as described in [17].

As in [16], let D be a set of distance matrices characterising the morphological similarity between images, with the i-th matrix of this set denoted by D_i. Here, D_i will be an $L_i \times (N-1)$ matrix describing the distance between the image Y_i and each one of the remaining $(N-1)$ images at every voxel position $\boldsymbol{v}$. More specifically, $D_{ij}(\boldsymbol{v})$ will contain the distance between the $\boldsymbol{v}$-th voxel of image Y_i and its corresponding location in image Y_j. Under this assumption, a heat kernel decay function $W_{ij}(\boldsymbol{v})$ is then used as a weight to diffuse information between images. This kernel is defined as $W_{ij}(\boldsymbol{v}) = \exp(-\frac{D_{ij}(\boldsymbol{v})}{t})$, with t being the temperature of the heat kernel that will determine the distance and the speed of information diffusion and $D_{ij}(\boldsymbol{v})$ is a morphological distance defined as in Cardoso *et al.* [16]. Here, t is set to 1 and its optimisation is out of the scope of this work. Note that the graph is undirected only if the distances are a semi-metric (subadditivity is not required). In short, the morphological neighbourhood of a voxel in a certain subject will be restricted to the most similar subjects from each group. In this work, α (as defined in [16]) will be set to 0.9, reducing the influence of the co-ordinate mapping in the morphological similarity metric. This information will be later introduce through Jacobian modulation. Also, no truncation is applied to $W_{ij}(\boldsymbol{v})$, making it spann the full space of morphologies.

2.2 Tissue Segmentation

All N_i images are segmented into their 3 constituent tissue classes: white-matter (WM), grey-matter (GM) and cerebrospinal-fluid (CSF). The tissue segmentation is obtained using AdaPT [18], an EM-based probabilistic segmentation algorithm. This algorithm reduces the bias introduced in the tissue segmentation by the choice of tissue priors through a patient-specific prior sampling procedure. Here, $S_i^c(\boldsymbol{v})$ represents the probability for voxel $\boldsymbol{v}$ in image i to belong to tissue class c. All images are skull-stripped using STEPS [19].

2.3 Stratified VBM

In VBM framework, all tissue segmentations are mapped to a groupwise space and commonly modulated by the Jacobian determinants of the transformation. The modulated segmentation of image i at voxel $\boldsymbol{v}$ in the groupwise space for tissue X is denoted by $M^c_{i\leftarrow \mathrm{GW}}(\boldsymbol{v}) = S_i^c(T_{i\leftarrow \mathrm{GW}}(\boldsymbol{v})) \times |\mathrm{Jac}(T_{i\leftarrow \mathrm{GW}}(\boldsymbol{v}))|$. The

global affine transformation between the images was removed from the Jacobian estimation in order to reduce the influence of head size.

Unlike VBM, which compares all the segmentations in a groupwise space, sVBM compares the segmentations in the space of every image in the database. For a subject i, the segmentation of tissue type C of all other subjects $j \in N \setminus \{i\}$ is mapped to the space of i and modulated by the Jacobian determinants. The modulated segmentations are denoted by $M^c_{i \leftarrow j}$. The same process is repeated for all $i \in N$. Thus, $N \times (N-1)$ modulated probability images are generated per tissue type. Each modulated probability is smoothed by a $2mm$ isotropic full width at half maximum (FWHM) Gaussian filter. This filtering is intrinsically smaller than in VBM as the pairwise registration and the weighting scheme reduces the registration errors.

As previously mentioned, $W_{ij}(\boldsymbol{v})$ characterises the morphological similarity between image i and j at voxel $\boldsymbol{v}$. Thus, $W_{ij}(\boldsymbol{v})$ can be used to weight the modulated tissue densities by the morphological similarity between subjects.

Let the N subjects be divided into two sub-populations of interest: e.g. healthy-controls (N^{HC}), and diseased subjects (N^P). Within the space of a specific subject i, the difference in modulated tissue densities between the two groups at voxel $\boldsymbol{v}$ and for tissue X, here denoted by $D^c_i(\boldsymbol{v})$, will be given by:

$$D^c_i(\boldsymbol{v}) = \frac{\sum_{\forall j \in N^P \setminus \{i\}} W_{ij}(\boldsymbol{v}) \times M^c_{i \leftarrow j}(\boldsymbol{v})}{\sum_{\forall j \in N^P \setminus \{i\}} W_{ij}(\boldsymbol{v})} - \frac{\sum_{\forall j \in N^{HC} \setminus \{i\}} W_{ij}(\boldsymbol{v}) \times M^c_{i \leftarrow j}(\boldsymbol{v})}{\sum_{\forall j \in N^{HC} \setminus \{i\}} W_{ij}(\boldsymbol{v})}$$

These density difference maps are obtained for every subject i and tissue c, resulting in $N \times C$ modulated tissue difference maps between the two populations. Each one of these maps, D^c_i, provides voxel-wise information about the difference in the density of tissue type c between the two populations, weighted by the local similarity between each subject j and the subject i. Note that the current implementation does not take into account the presence of covariates like age and gender, as the testing data is balanced. However, this can be easily added through a weighted least-squares GLM, using W_{ij} as weights.

2.4 Groupwise projection

The values of D^c_i already characterise the differences between the populations for each of the subjects' morphological cluster. Thus, one can now map them to the groupwise space without entangling the pathological with the naturally-occurring variability. For all subjects and tissue types, the tissue density differences are then mapped to the groupwise space, in a similar fashion to the VBM approach. However, in contrast to VBM, where only a point estimate of the populations tissue density difference is obtained at each voxel, in sVBM, a non-parametric distribution of tissue density differences for the different morphologies and between the diseased and heathy populations is obtained at each voxel $\boldsymbol{v}$ of the groupwise space. The vector of tissue density differences at voxel $\boldsymbol{v}$ in the groupwise space is denoted by $D^c_{GW}(\boldsymbol{v}) = \{D^c_1(T_{1 \leftarrow \mathrm{GW}}(\boldsymbol{v})), \ldots, D^c_N(T_{N \leftarrow \mathrm{GW}}(\boldsymbol{v}))\}$.

2.5 Sampling the null distribution

In order to test the significance of the findings, a non-parametric strategy similar to permutation testing is employed. Permutation methods (also known as randomisation methods) can be used for inference and thresholding of statistic maps when the null distribution is not known due to noise, population sample and the use of non-standard statistics. Ideally for sVBM, the null distribution should be constructed in the space of each patient's morphology in order to provide information about the null distribution of density differences for each subject's morphological cluster. For a relatively large number N, this process would be computationally very time consuming. Furthermore, the multiple null distributions on the space of each subject's morphology would have to be propagated to the groupwise space using $T_{i \leftarrow \mathrm{GW}}(\boldsymbol{v})$, breaking the null distribution's independence criterion for each voxel. Instead, an approximation of the true null distribution of D^c_{GW} is obtained by permuting the labels of $M^c_{i \leftarrow \mathrm{GW}}$ [20]. A $6mm$ isotropic FWHM Gaussian filter smoothing is necessary here to correct for the problem of low variance [21] and also in order to account for minor mapping errors to the groupwise space. In this work, 20,000 random label samples were used per voxel to build the null distribution $\mathscr{H}_0$. An example of the null distribution for a voxel in the parahippocampal gyrus can be seen in purple in Fig. 3. The probability that a sample of D^c_{GW} comes from $\mathscr{H}_0$ will thus be given by $\int f(D^c_{GW}) \times f(\mathscr{H}_0)$, with $f(D^c_{GW})$ and $f(\mathscr{H}_0)$ being the probability density function of D^c_{GW} and $\mathscr{H}_0$ respectively.

3 Data

Due to the computational complexity of the pairwise registration strategy, only 100 ADNI2 (`http://adni.loni.ucla.edu`) data sets (50 healthy-controls and

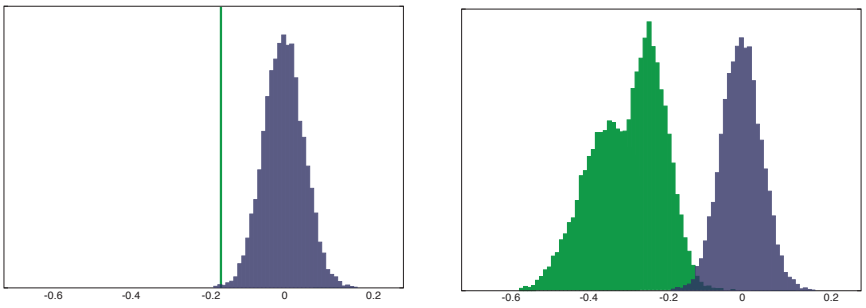

Fig. 3. The null distribution $\mathscr{H}_0$ (in purple) for a voxel in the parahippocampal gyrus. Left) The green line represents the voxel's GM tissue density difference; Right) A smoothed probability density function of the samples in D^c_{GW} (in green) suggesting a non-gaussian distribution of pathological effects for the different brain morphologies.

50 AD subjects) were used for validation. The 50 subjects of each group were chosen as the 50 first subjects ranked by scan number. The list of subjects used in this work will be made available on the author's website at the time of publication. No age and gender statistical differences were found between the two populations.

4 Validation

All subjects were affinely mapped to a template's physical space without resampling by updating the header information. All these affinely-aligned images were then non-rigidly mapped to each other in order to obtain $T_{i\to j}(\boldsymbol{v})$, $W_{i\to j}(\boldsymbol{v})$ and $M_{i\to j}(\boldsymbol{v})$, requiring 9900 registrations (100×99). All density difference maps D^c_{GW} and segmentations were also mapped to a mean deformation space, here called the groupwise space. The segmentations were modulated by the Jacobian determinant and smoothed with a 6mm FWHM Gaussian filter ($M^c_{i\leftarrow \mathrm{GW}}$), similar to VBM, accounting for possible mapping errors. These modulated segmentations were then used to sample $\mathscr{H}_0$ and the p-values (P-map) of the sVBM's population differences map. For comparison purposes, the same smoothed modulated segmentations were also used to estimate a regular VBM-based difference map between the AD and HC populations and its associated p-values. All p-maps are presented uncorrected for multiple comparisons and masked to regions with density above 0.1. Results for the GM are shown in Fig. 4.

5 Discussion

The proposed method shows a well localised and anatomically plausible pattern of atrophy for AD when compared to HC. By stratifying the naturally-occurring image morphologies from the estimation of pathological effects, the mean density differences between the populations become shaper and more pronounced. More interesting than the accuracy of density difference localisation are the characteristics of the non-parametric distribution D_{GW} (see Fig. 4-bottom). The mean, std, 10th and 90th percentiles of the density difference distribution D_{GW} at each voxel are presented in Fig. 4. Note that the least pathologically-affected brain morphologies (90th percentile) still suffer from atrophy in the temporal lobe and thalami area. However, the most pathologically-affected morphologies (10th percentile) suffer from atrophy in nearly all brain regions. The skewness and standard deviation (std) of D_{GW} suggest that the disease process does not affect all morphologies in a similar manner, with some specific local brain morphologies showing almost no pathological effects (see Fig. 4-bottom right).

6 Conclusion

The proposed work presents a novel technique for morphometric characterisation of population differences. In contrast to VBM, which provides an answer to

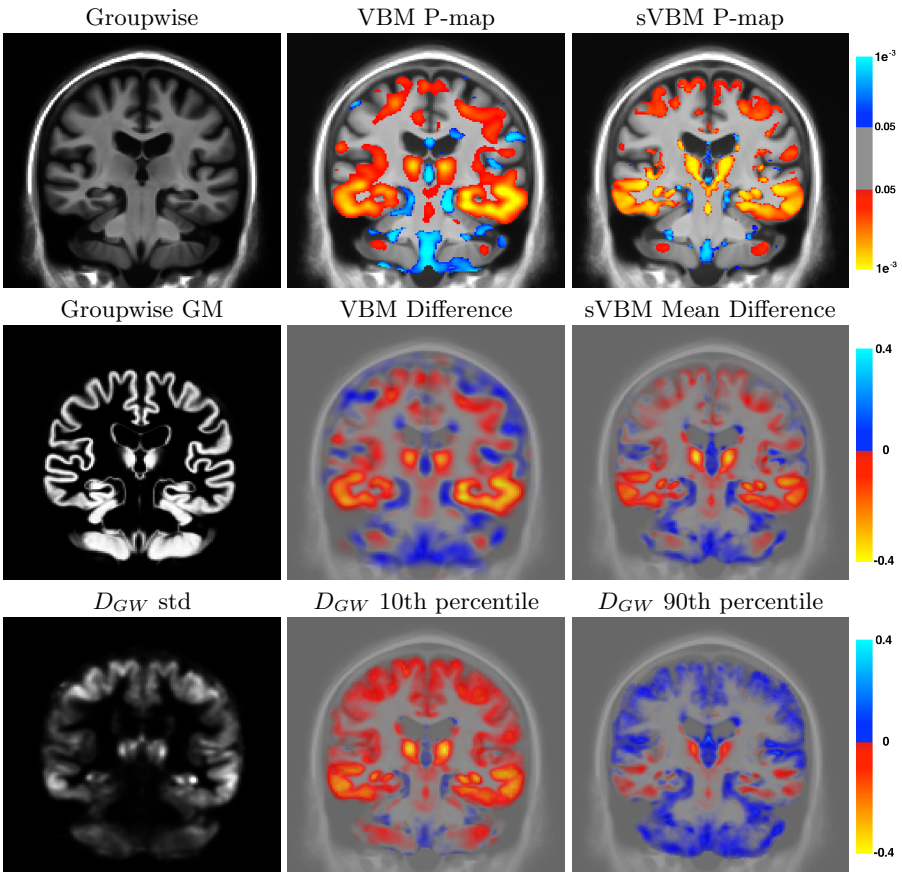

Fig. 4. The first row shows the groupwise mean image, and the positive and negative VBM and sVBM p-maps thresholded at 0.05, showing sharper atrophy localisation. The red and blue colours represent atrophy and expansion respectively. Note the improved localisation accuracy of the pathological differences in sVBM. The second row shows the mean GM segmentation, the VBM GM density difference and the sVBM GM density difference mean, both between the AD and HC populations. The last row shows the std and the 10th and 90th percentile of the density difference distribution D_{GW}.

the question "How does pathology affect an average brain morphology?", the proposed work attempts to answer the question "How does pathology affect each individual brain morphology?". By answering a different question, sVBM could provide a complementary and possibly richer source of information about neurodegeneration.

Acknowledgements

The Dementia Research Centre is an Alzheimer's Research Trust Co-ordinating centre and has also received equipment founded by the Alzheimer's Research Trust. MJC, MM and SO are funded by both the EPSRC (EP/H046410/1) and the CBRC Strategic Investment Award (Ref. 168).

References

1. Ashburner, J., Friston, K.J.: Voxel-based morphometry - The methods. NeuroImage **11**(6) (2000) 805–821
2. Woods, R.P.: Characterizing volume and surface deformations in an atlas framework: theory, applications, and implementation. NeuroImage **18**(3) (March 2003) 769–788
3. Ashburner, J., Hutton, C., Frackowiak, R., Johnsrude, I., Price, C., Friston, K.: Identifying global anatomical differences: deformation-based morphometry. Human Brain Mapping **6**(5-6) (1998) 348–357
4. Ashburner, J., Friston, K.J.: Unified segmentation. NeuroImage **26**(3) (2005)
5. Yeo, B.T.T., Sabuncu, M.R., Desikan, R.S.R., Fischl, B., Golland, P.: Effects of registration regularization and atlas sharpness on segmentation accuracy. MedIA **12**(5) (October 2008) 603–615
6. Bookstein, F.L.: "Voxel-Based Morphometry" Should Not Be Used with Imperfectly Registered Images. NeuroImage **14**(6) (December 2001) 1454–1462
7. Davatzikos, C.: Why voxel-based morphometric analysis should be used with great caution when characterizing group differences. NeuroImage **23**(1) (September 2004) 17–20
8. Leporé, N., Brun, C., Chou, Y.Y., Lee, A., Barysheva, M., De Zubicaray, G.I., Meredith, M., Macmahon, K., Wright, M., Toga, A., Thompson, P.M.: Multi-Atlas Tensor-Based Morphometry and its Application to a Genetic Study of 92 Twins. 2nd MICCAI Workshop ... (October 2008) 48–55
9. Koikkalainen, J., Lötjönen, J., Thurfjell, L., Rueckert, D., Waldemar, G., Soininen, H., Alzheimer's Disease Neuroimaging Initiative: Multi-template tensor-based morphometry: application to analysis of Alzheimer's disease. NeuroImage **56**(3) (June 2011) 1134–1144
10. Amiez, C., Neveu, R., Warrot, D., Petrides, M., Knoblauch, K., Procyk, E.: The Location of Feedback-Related Activity in the Midcingulate Cortex Is Predicted by Local Morphology. Journal of Neuroscience **33**(5) (January 2013) 2217–2228
11. Sabuncu, M.R., Balci, S.K., Shenton, M.E., Golland, P.: Image-driven population analysis through mixture modeling. IEEE TMI **28**(9) (September 2009) 1473–1487
12. Ribbens, A., Hermans, J., Maes, F., Vandermeulen, D., Suetens, P.: SPARC: unified framework for automatic segmentation, probabilistic atlas construction, registration and clustering of brain MR images. In: IEEE International Symposium on Biomedical Imaging, IEEE (2010) 856–859
13. Rohlfing, T., Russakoff, D.B., Maurer Jr, C.R.: Performance-based classifier combination in atlas-based image segmentation using expectation-maximization parameter estimation. IEEE TMI **23**(8) (August 2004) 983–994
14. Warfield, S.K., Zou, K.H., Wells III, W.M.: Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation. IEEE TMI **23**(7) (July 2004) 903–921

15. Coupé, P., Eskildsen, S.F., Manjón, J.V., Fonov, V.S.: Scoring by nonlocal image patch estimator for early detection of Alzheimer's disease. NeuroImage: Clinical (2012)
16. Cardoso, M.J., Wolz, R., Modat, M., Fox, N., Rueckert, D., Ourselin, S.: Geodesic Information Flows. In Ayache, N., Delingette, H., Golland, P., Mori, K., eds.: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2012. Springer Berlin / Heidelberg, Berlin, Heidelberg (2012) 262–270
17. Modat, M., Cardoso, M.J., Daga, P., Cash, D., Fox, N.C., Ourselin, S.: Inverse-Consistent Symmetric Free Form Deformation. In: WBIR. (July 2012)
18. Cardoso, M.J., Melbourne, A., Kendall, G.S., Modat, M., Robertson, N.J., Marlow, N., Ourselin, S.: AdaPT: An adaptive preterm segmentation algorithm for neonatal brain MRI. NeuroImage (August 2012)
19. Jorge Cardoso, M., Leung, K.K., Modat, M., Keihaninejad, S., Cash, D., Barnes, J., Fox, N.C., Ourselin, S., for the Alzheimer's Disease Neuroimaging Initiative: STEPS: Similarity and Truth Estimation for Propagated Segmentations and its application to hippocampal segmentation and brain parcelation. MedIA (March 2013)
20. Nichols, T.E., Holmes, A.P.: Nonparametric permutation tests for functional neuroimaging: a primer with examples. Human Brain Mapping **15**(1) (January 2002) 1–25
21. Ridgway, G.R., Omar, R., Ourselin, S., Hill, D.L.G., Warren, J.D., Fox, N.C.: Issues with threshold masking in voxel-based morphometry of atrophied brains. NeuroImage **44**(1) (2009) 99–111