Label propagation using group agreement – DISPATCH

Rolf A. Heckemann[1,2], Christian Ledig[3], Paul Aljabar[3,4], Katherine R. Gray[3], Daniel Rueckert[3], Joseph V. Hajnal[4], and Alexander Hammers[1,2]

[1] The Neurodis Foundation, Lyon, France, soundray@fondation-neurodis.org,
[2] Centre for Neuroscience (Hammersmith), Imperial College London, UK
[3] Department of Computing, Imperial College London, UK,
[4] Division of Imaging Sciences and Biomedical Engineering, Kings College London, UK

Abstract. To address the challenge of applying expert anatomical knowledge captured in brain atlases to unseen brain images, we propose "DISPATCH". **D**ISPATCH **i**s **s**egmentation by **p**ropagating **at**lases with **c**oerced **h**armony; a multi-level, multi-resolution label propagation approach that exploits per-level groupwise agreement.
The method relies on pairwise label propagation at a given resolution level. The resulting segmentations are combined using selective vote-rule decision fusion. The consolidated label set serves as a common target for a label-based registration using label consistency as the similarity metric. The resulting transformations are used as the starting point for iterating the registration-fusion-registration sequence at the next resolution level. We participate in the "MICCAI 2012 Grand Challenge and Workshop on Multi-Atlas Labeling" with this novel approach.

1 Introduction

Magnetic resonance (MR) scanning of the human brain with state-of-the-art equipment generates large quantities of data. These are typically represented as 3D grey scale images that map the spatial signal distribution. For many applications, visual section-by-section review of these images is still the preferred method of processing such data. It does not, however, scale to large numbers of images. To extract information from large sets of images or multi-centre image repositories, such as ADNI, AIBL, IXI, OASIS, etc., efficient automatic methods are required.

Automatic anatomical segmentation provides an important avenue towards dimensionality reduction, feature extraction, and identification of imaging biomarkers. To segment a brain image of a given study subject or patient anatomically, most approaches rely on atlases generated by experts through manual segmentation of equivalent images. The optimal strategy for transferring this expert knowledge from the atlases to the new image is a matter of scientific debate. However, label propagation with decision fusion has consistently been among the best performing methods.

The approach presented here, "DISPATCH" (**D**ISPATCH **i**s **s**egmentation by **p**ropagating **at**lases with **c**oerced **h**armony) is a multilevel labelling procedure relying on forcing agreement of all atlases at each refinement step. It was developed from a brain extraction method called "pincram" (**p**yramidal **in**tra**cra**nial **m**asking). On the Segmentation Validation Engine (http://sve.loni.ucla.edu/), pincram is currently ranked second after the gold-standard labels for the site's test set (accessed 5 July 2012).

2 Method

2.1 Material

Data provided in the course of the "MICCAI 2012 Grand Challenge and Workshop on Multi-Atlas Labeling" were used. A total of 35 images was supplied, originating from the OASIS project (http://www.oasis-brains.org/). Training data consisted of 15 T1-weighted images, with spatially corresponding, expertly generated maps identifying 138 labels, 113 of which had been declared relevant for the challenge[1]. Testing data consisted of 20 T1-weighted images, with labels that remained hidden from the participants.

2.2 Iterative labelling procedure

For a given target, labels were generated at progressive levels of refinement, termed affine, coarse, medium and fine, according to the detail level of the image registration step. At each level, the following calculations were carried out:

1. Label sets were generated from all 15 atlases using a standard label propagation approach based on image registration (cf. Section 2.3).
2. The 15 label sets were consolidated in target space using vote-rule decision fusion, generating fused set F1.
3. The 15 individual label sets were ranked in descending order of label agreement (Jaccard index) with F1.
4. These ranked values were used to determine an acceptability threshold $\theta = L_1 - 2(L_1 - L_7)$. L_7 was a coarse approximation of the median.
5. Individual segmentations passing the threshold were consolidated, generating fused set F2 and a mask M indicating non-unanimously labelled target voxels. Both F2 and M are in target space.
6. M was applied to the target image, creating an image with original grey scale values in voxels that were non-unanimously labelled and zero beyond. This ensured that only regions where labels had not been unanimously assigned were considered for the similarity calculation during the subsequent iteration.
7. The 15 atlas label sets (in their original space) were registered to F2, maximizing label consistency (cf. Section 2.3). The resulting 15 transformations from each individual atlas space to the target space were retained.

[1] Label sets were provided by Neuromorphometrics, Inc. (http://neuromorphometrics.com/) under academic subscription.

The masked target image and the retained transformations were used as starting points to iterate the procedure at the next level of refinement.

2.3 Image registration

Each pair of images (T1 or label sets) was affine-registered (first iteration) or nonrigidly registered (subsequent iterations). Nonrigid registration consisted in applying displacements to the atlas image via a lattice of control points, blended using B-spline basis functions, maximizing a similarity metric (see below) [1]. The stopping condition for the optimization was either no further improvement in similarity or the reaching of a maximum number of iterations.

The similarity metric for registering pairs of T1-weighted images was normalized mutual information [2]. To ensure the effectiveness of the data reduction step (6.), only values greater than zero were considered in the calculation.

The similarity metric for registering pairs of label sets was label consistency. It is defined as the fraction of voxels with agreed classifications, i.e. if n_{ij} represents the number of voxels given label i by one image and label j by the other, then label consistency is measured as

$$\frac{\sum_i n_{ii}}{\sum_{i,j} n_{ij}}. \tag{1}$$

All registration steps were carried out using the Image Registration Toolkit (IRTK, www.doc.ic.ac.uk/~dr/software/).

3 Discussion

The DISPATCH procedure combines proven techniques in a novel fashion. Previous multi-atlas label propagation methods produced multiple segmentations from individual atlases independently and only combined them in the final step of the procedure (e.g. MAPER, [3]). DISPATCH generates instead a consensus labeling at each of several levels of refinement, using information from the T1-weighted images. Starting estimates for the subsequent level are then created by registering all atlas label sets to the consensus set. These starting estimates constrain and inform the subsequent level of T1-pair registrations by driving them towards the consensus. In addition, efficiency is achieved through data reduction: at subsequent refinement levels, only those parts of the target image are considered which are likely to contain label boundaries. Such a data reduction step has previously been described for BEaST, an accurate brain extraction method that similarly constrains the algorithm's boundary search to regions which, based on information taken from preceding iterations, are likely to contain that boundary [4].

DISPATCH performs best on full head images. Brain extraction, a prerequisite step for many other approaches, is unnecessary.

Another related approach that has previously been described is "SIMPLE" [5]. It also relies on estimating the performance of an individual atlas by estimating agreement of its propagated label set with a fused label set. SIMPLE discards atlases that fail to pass an agreement threshold, whereas DISPATCH rescues such atlases by registering their labels to the consensus set. The performance estimation in SIMPLE is carried out after a detailed registration has been performed, whereas in DISPATCH, a performance estimation step is inserted at each detail level, integrating the fusion step into each successive refinement procedure.

The evaluation results indicate that the DISPATCH approach works in principle. While the accuracy is somewhat inferior to, for example, MAPER, we are nevertheless encouraged by the findings. Since the development of DISPATCH is still in its infancy, no sophisticated refinement has been attempted yet. We expect that incorporating tissue class information, employing more elaborate label fusion techniques, and choosing more appropriate atlas selection strategies will yield substantial improvements. Also, due to the large number of registrations involved in DISPATCH, the parameter space is large, and further improvements are likely to emerge from parameter optimizations.

References

1. Rueckert, D., Sonoda, L.I., Hayes, C., Hill, D.L., Leach, M.O., Hawkes, D.J.: Nonrigid registration using free-form deformations: application to breast MR images. IEEE Transactions on Medical Imaging **18**(8) (August 1999) 712–721
2. Studholme, C., Hill, D.L.G., Hawkes, D.J.: An overlap invariant entropy measure of 3D medical image alignment. Pattern Recognition **32**(1) (January 1999) 71–86
3. Heckemann, R.A., Keihaninejad, S., Aljabar, P., Rueckert, D., Hajnal, J.V., Hammers, A.: Improving intersubject image registration using tissue-class information benefits robustness and accuracy of multi-atlas based anatomical segmentation. Neuroimage **51**(1) (May 2010) 221–227
4. Eskildsen, S.F., Coupé, P., Fonov, V., Manjón, J.V., Leung, K.K., Guizard, N., Wassef, S.N., Østergaard, L.R.R., Collins, D.L., Alzheimer's Disease Neuroimaging Initiative: BEaST: brain extraction based on nonlocal segmentation technique. Neuroimage **59**(3) (February 2012) 2362–2373
5. Langerak, T.R., van der Heide, U.A., Kotte, A.N., Viergever, M.A., van Vulpen, M., Pluim, J.P.: Label fusion in atlas-based segmentation using a selective and iterative method for performance level estimation (SIMPLE). IEEE Transactions on Medical Imaging **29**(12) (December 2010) 2000–2008