OPTIMAL ATLAS SELECTION USING IMAGE SIMILARITIES IN A TRAINED REGRESSION MODEL TO PREDICT PERFORMANCE.

Akin Akinyemi, Costas Plakas, Jim Piper, Colin Roberts, Ian Poole.

Toshiba Medical Visualization Systems, Europe

ABSTRACT

An atlas in the context of atlas-based segmentation refers to a pre-selected image with labelled anatomical regions of interest. Atlas-based segmentation is the propagation of these labels to a novel image after both images have been registered. The goal of an atlas is to be representative of an anatomical category, but in practice there exists variability in human anatomy. One solution to maintain consistent segmentation accuracies is to use multiple atlases, with a system for selecting the most appropriate atlas at the time of segmentation. This paper describes a method for selecting an atlas using a linear regression model to predict the segmentation accuracy based on image similarity measures. It goes further to present an offline method for automatically selecting a set of atlases, representative of the training set to be used during segmentation; all of this illustrated by segmentation of the heart and kidneys in 3D CT images.

Index Terms— Atlas-based segmentation, multi-atlas, optimal atlas selection, registration

1. INTRODUCTION

Atlas based segmentation is the process of mapping a set of labelled anatomical regions from a representative image, termed the atlas, onto a novel image following an alignment process. Multi-atlas based segmentation extends this by using a set of atlases, selecting the most suitable atlas based on the quality of alignment to the novel image (see figure 1).

Multiple atlases are needed because of the variability that exists within human anatomy. This variation in appearance of an organ across different patients may be due to pathology. It may also be due to demographics, size and subject's lifestyle.

The evaluation of selection strategies in [1] shows that selecting the atlas which produces the highest normalised mutual information after non-rigid registration to the novel dataset, results in higher average segmentation accuracy on the experiment set compared with the highest average accuracy obtained by using any atlas in the set on its own. Subsequent approaches in [2] extend this method by computing the similarity measure in a region of interest; this is particularly useful in situations where certain sub-anatomies are likely to vary across subjects.

A common approach to multi-atlas based segmentation is decision-fusion from multiple classifiers as shown in [3], [1] and [4], where each atlas is treated as a separate classifier and each pixel in the novel image assigned a class label following registration. The task is to find a combination of the individual segmentations that produces the most accurate overall segmentation.

A problem with vote counting strategies arises when the anatomical category represented by the novel image has minority representation amongst the group of atlases used. In [3] and [5], this issue is addressed by limiting the set of atlases to those that are most similar to the novel dataset using image similarity and meta-information, such as patient's age, gender and clinical status.

The method proposed in this paper draws upon the findings reported in [3], [6] and [1], stating that an atlas selected on the basis of image similarity produces better results than is possible with a single atlas used across all novel datasets. It extends this idea, showing how calibration of similarity measures such as cross-correlation can be used to predict the segmentation accuracy.

2. METHOD

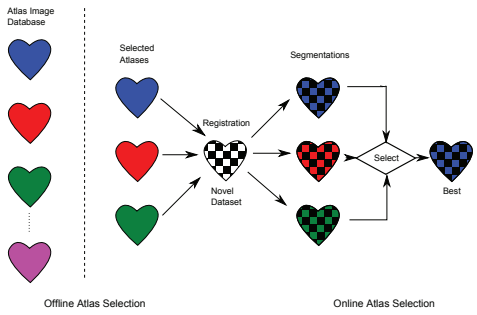


Fig. 1. Multi-Atlas Based Segmentation
The atlases in use in the application are selected offline from a database of image atlases. Each of the chosen atlases produces a segmentation when aligned with the novel image, therefore a mechanism capable of predicting the best segmentation is required.

2.1. Multi-Atlas Based Segmentation

Assuming a set $A = \{a_0, ..., a_{M-1}\}$ of M atlases has been provided, segmentation of the required anatomical region in a novel dataset t begins by registering each $a_m (m = 0, ..., M-1)$ with t in order to produce a set of warp-fields $W = \{W_0, ..., W_{M-1}\}$, where W_m maps a_m onto t.

Using the warp-field W_m, a number of similarity measures between $W_m(a_m)$ and t can be evaluated. As mentioned in [3], there is no strict requirement for the selection measure to be the same as that used to drive registration; we believe that for fairness, they should be different. The rest of this paper focuses on cross-correlation as the selection criteria.

The cross-correlation between any $W_m(a_m)$ and t is not used directly to determine the segmentation accuracy, because this measure by itself is not directly comparable across all possible pairs of $W_m(a_m)$ and t. In order to get a more standard measure for selecting the best out of M possible segmentations, a linear regression model specific to each atlas is used. The function is of the form $J = \hat{\beta}_1 x + \hat{\beta}_0$, where x is the value of the cross-correlation and J is the predicted segmentation accuracy, using the Jaccard overlap (1); $\hat{\beta}_1$ and $\hat{\beta}_0$ are stored as part of the atlas.

$$Jaccard(A, B) = \frac{Vol.(A \cap B)}{Vol.(A \cup B)}, \quad (1)$$

where A and B are the generated and ground-truth segmentations, respectively.

The labelled anatomical region from the a_m that predicts the highest segmentation accuracy, is propagated onto t using W_m, resulting in a segmentation. The overall offline and online atlas selection is depicted in figure 1.

$$a^* = \underset{a_m}{\operatorname{argmax}}\{\hat{\beta}_{1_{a_m}} corr(W_m(a_m), t) + \hat{\beta}_{0_{a_m}}\}$$

Section 2.2 describes how these regression functions are obtained, while section 2.3 describes a method for selecting an appropriate atlas set A based on available training data.

2.1.1. Registration

The aim of registration is to find a mapping W, between a reference R and a target image T, that maximises some similarity measure. The registration process used here is a multi-scale approach, performing 2 passes of rigid registration, followed by 2 passes of non-rigid registration, with all stages run until convergence of the cost function.

The rigid registration finds the best 3-D translation, anisotropic scaling and roll, pitch and yaw components to transform the atlas image R onto the novel image T. Mutual information (MI) is used as the similarity measure to drive the registration. Powell's optimisation method ([7]) is used to find the rigid transformation matrix $\mathbf{A}$ that maximises MI between the transformed R and T.

$$\mathbf{A}^* = \underset{\mathbf{A}}{\operatorname{argmax}}\{\mathbf{MI}(\mathbf{A}.R, T)\}$$

The non-rigid phase aims to find the warp-field W that maximises MI between T and the transformed reference R'. Non-rigid registration allows non-linear deformation of the image, making it an ill-posed problem, therefore some form of regularisation is required. Our method finds the W that maximises MI between R' and T within fluid and elastic constraints.

The warp-field is initialised from the rigid phase and evolved to convergence, with the warp-field at iteration i given by:

$$W_i = E \otimes (W_{i-1} + kV \otimes F_{i-1}),$$

where F_i is the current force-field, i.e. the derivative of MI at each point in R computed from a joint intensity histogram using the method described in [8]. V and E are viscous fluid ad elastic constraints, respectively; both modelled by Gaussian filters.

2.2. Training

Each dataset in turn is used as an atlas to segment an organ in all the other datasets using the registration method outlined in section 2.1. The cross-correlation between the two aligned volumes measured on a region of interest (ROI), and the Jaccard overlap (1) are collected from each segmentation.

These measurements make up the training parameters and they are presented in seperate cross-segmentation matrices, **J** and **CC**(figure 2), where $\mathbf{J}(i, j)$ is the Jaccard overlap produced when atlas a_j is used to segment the organ in image t_i, and $\mathbf{CC}(i, j)$ is the cross-correlation measured at the optimal alignment between atlas image a_j and novel image t_i.

	1798_006	1803_025	1805_003	1831_004	1686_001	1688_001
01798_006	0.928	0.855	0.911	0.794	0.878	0.911
01803_025	0.822	0.92	0.901	0.85	0.846	0.874
01805_003	0.838	0.795	0.927	0.804	0.842	0.854
01831_004	0.882	0.829	0.907	0.934	0.879	0.869
01686_001	0.816	0.754	0.855	0.828	0.918	0.817
01688_001	0.863	0.786	0.854	0.8	0.821	0.917

Cross-Correlation Matrix

Fig. 2. Training Parameter Matrices for Heart Segmentation *The image identifiers are displayed as row and column headers in the matrix. The columns represent the atlas images, while the rows represent the novel images; therefore, the diagonal elements are self-segmentation parameters.*

The parameter model used in this system is a univariate least-squares regression function measured on each similarity measure for each atlas in the form $\hat{y} = \hat{\beta}_1 x + \hat{\beta}_0$, correlating cross-correlation with segmentation accuracy (figure 3). The

predicted segmentation accuracy when using atlas a_j to segment the organ in image t_i based on the measured correlation coefficient between the two at the ROI is obtained using 2.

$$J^*_{a_j}(t) = \hat{\beta}_{1a_j}\mathbf{CC}(i,j) + \hat{\beta}_{0a_j} \quad (2)$$

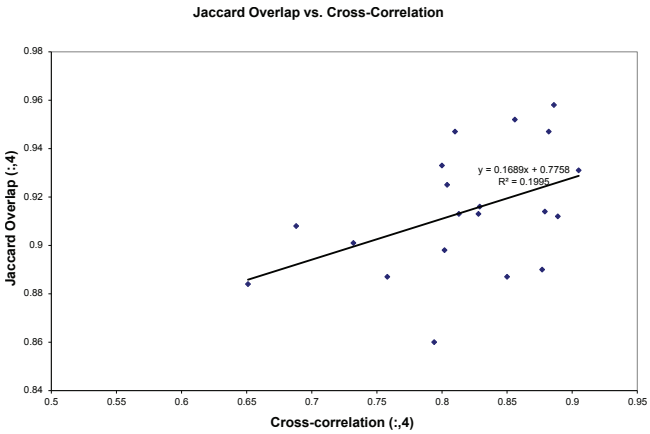


Fig. 3. Segmentation Overlap vs. Similarity Measure
This regression model is built up for each atlas based on values collected from the relevant column in the cross-segmentation matrix.

2.3. Offline Multi-Atlas Selection

The aim of this step is to select an optimal set A_{opt} of M atlases, $A = \{a_0, ..., a_{M-1}\}$. The optimal set is that which maximises the segmentation accuracy as measured by the average Jaccard overlap, using the multi-atlas based segmentation method outlined in section 2.1.
Formally:

$$A_{opt} = \underset{A \in D^M}{\operatorname{argmax}} \frac{1}{N-M} \sum_{i=1, d_i \notin A}^{n} J'_A(d_i), \quad (3)$$

where $N = |D|$ is the number of training images available,

$$J'_A(d_i) = \mathbf{J}(d_i, \underset{a \in A}{\operatorname{argmax}}\, J^*_a(d_i)), \text{and} \quad (4)$$

$$J^*_a(d_i) = \hat{\beta}_{1a}.\mathrm{CC}(d_i, a) + \hat{\beta}_{0a} \quad (5)$$

3. EXPERIMENTS

3.1. Heart Segmentation

Cardiac CT angiography (CCTA) datasets obtained from hospitals in Japan, U.S.A, and the U.K. are used in this experiment. The scanners used varied according to the data sources.

The images were 3-D volumes containing at least 300 slices with 512x512 pixels per slice, and a slice thickness of 0.5mm with typical intra-slice pixel spacing of 0.4mm. Each image therefore contained the entire heart and thoracic cavity.

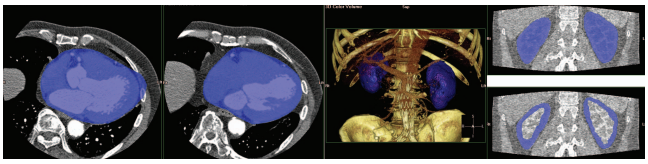


Fig. 4. Heart and Kidney Segmentation Results
Left to Right: Heart segmentation using best single atlas. Heart segmentation using optimal multi-atlas with $M = 2$. VR showing kidney segmentation in low-dose CT. MPR showing kidney segmentation.

Whole heart segmentations, represented as masks containing the entire heart as one fully-connected domain, were collected by a clinically qualified individual on 21 of these images. Atlases were created from each image and mask combination by constraining the image to an ROI containing the heart mask morphologically dilated by 6mm to contain a degree of context.

3.2. Kidney Segmentation

Abdominal contrast-enhanced 4-D CT studies from 24 subjects are used in this experiment, each study containing 20 phases. The scans were taken using Toshiba's Aquilion ONE 320-slice scanner, which covers a vertical range of 16cm, hence both kidneys in one gantry rotation. Manual segmentations of each kidney and its renal cortex were collected on the early arterial phase for all 24 studies, and atlases were created from each image by constraining them to a region containing a single kidney mask morphologically dilated by 2mm.

4. RESULTS AND DISCUSSION

Example segmentations are shown in figure 4. Figure 5 shows an increase in the average heart segmentation accuracy for M ranging from 1 to 5. The top line shows the best possible results that can be generated if the most suitable atlas was selected from the set of M in every case, although in actual use this will be unknown. This chart suggests that the ROI based measures, particularly cross-correlation are better predictors than their global equivalents. For heart segmentation, the region containing the heart apex and the rib-wall was used. In the case of kidney segmentation, the region bounded by the whole kidney was the ROI used.

Figure 6 highlights the benefit of using the offline atlas selection (equation 3), as indicated by the upper line. The lower line refers to the expectation E of the overlaps averaged over all training images for a multi-atlas set selected at random:

$$E = \frac{1}{\binom{N}{M}} \sum_{A \in D^M} [\frac{1}{N-M} \sum_{i=1, d_i \notin A}^{N} J'_A(d_i)] \quad (6)$$

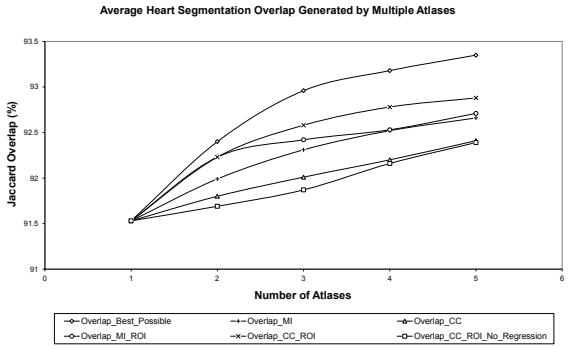


Fig. 5. Average Heart Segmentation Overlap Generated by Multiple Atlases

The segmentation accuracy increases as the size of the multi-atlas set increases. The topmost line shows the average heart segmentation accuracy generated if the atlases were selected based on the actual overlaps, rather than predicted overlaps.

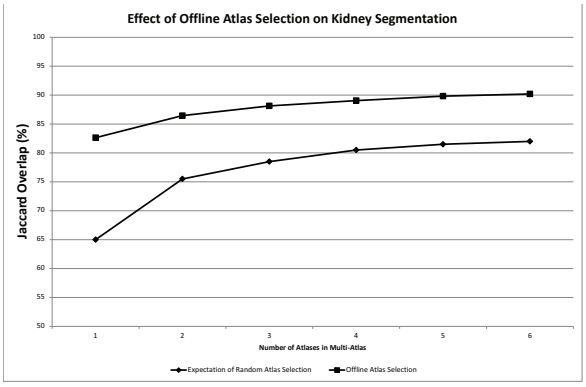


Fig. 6. Effect of Offline Atlas Selection on Kidney Segmentation

The upper line shows the Jaccard overlap averaged over all validation data when selecting the optimal atlases for each set size (3), while the lower line shows the mean of the Jaccard overlap averaged over all training data for all possible random selections of atlases (6).

5. CONCLUSION

This paper presents a method for optimal atlas selection from a set of candidate atlas images. The underlying method is applied at two stages in multi-atlas based segmentation. In the first stage, it is used offline to select an optimal multi-atlas set from a database of potential atlases, while in the second stage it is used during application to select an optimal atlas from within the set.

Each atlas contains an image, a mask containing the labelled anatomy of interest and a pair of linear regression parameters. The parameters are obtained from a training set of images by co-registering each pair of images and recording the resulting similarity measures and Jaccard overlaps of the generated segmentation. The optimal atlas from a set is selected following registration by measuring the cross-correlation between each atlas image and the novel image and applying the respective linear regression slopes and intercepts to predict the Jaccard overlap. The atlas image that predicts the highest segmentation accuracy is selected. The method was applied to heart and kidney segmentation from 3D CTA data.

The results demonstrate that a multi-atlas approach performs better than a single atlas method, and that calibration of the atlas selection criteria improves segmentation accuracy further.

6. REFERENCES

[1] Torsten Rohlfing, Robert Brandt, Randolf Menzel, Daniel B. Russakoff, and Calvin R. Maurer Jr., "Quo vadis, atlas-based segmentation?," in *The Handbook of Medical Image Analysis – Volume III: Registration Models*, Jasjit Suri, David L. Wilson, and Swamy Laxminarayan, Eds., chapter 11, pp. 435–486. Kluwer Academic / Plenum Publishers, New York, NY, Aug 2005.

[2] Minjie Wu, Caterina Rosano, Pilar Lopez-Garcia, Cameron S. Carter, and Howard J. Aizenstein, "Optimum template selection for atlas-based segmentation," *Neuroimage*, vol. 34, pp. 1612–1618, 2006.

[3] P. Aljabar, R.A. Heckemann, A. Hammers, J.V. Hajnal, and D. Rueckert, "Multi-atlas based segmentation of brain images: Atlas selection and its effect on accuracy," *NeuroImage*, vol. 46, no. 3, pp. 726–738, 2009.

[4] S. K. Warfield, K. H. Zou, and W. M. Wells, "Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation," *Medical Imaging, IEEE Transactions on*, vol. 23, pp. 903–921, 2004.

[5] D. Blezek and J. Miller, "Atlas stratification," *Med Image Anal*, 10 2007.

[6] J.M.P. Lötjönen, Robin Wolz, J.R. Koikkalainen, Lennart Thurfjell, Gunhild Waldemar, Hilkka Soininen, and Daniel Rueckert, "Fast and robust multi-atlas segmentation of brain magnetic resonance images," *NeuroImage*, 2009.

[7] Jorge Nocedal and Stephen J. Wright, *Numerical Optimization*, Springer, August 2000.

[8] William R. Crum, Derek L. G. Hill, and David J. Hawkes, "Information theoretic similarity measures in non-rigid registration," in *IPMI*, 2003, pp. 378–387.