Available online at www.sciencedirect.com
SCIENCE DIRECT®
Pattern Recognition Letters 26 (2005) 2070–2079
Pattern Recognition Letters
www.elsevier.com/locate/patrec

Multi-classifier framework for atlas-based image segmentation

Torsten Rohlfing *, Calvin R. Maurer Jr.

Department of Neurosurgery, Stanford University, 300 Pasteur Drive, MC 5327, Stanford, CA 94305-5327, USA

Received 20 July 2004; received in revised form 2 December 2004
Available online 29 April 2005
Communicated by Y.J. Zhang

Abstract

Three different systematic approaches to generate multiple classifiers in atlas-based biomedical image segmentation are compared. Different atlases, as well as different parametrization of the registration algorithm, lead to different atlas-based classifiers. The classifier outputs are combined and compared to a manual ground truth segmentation. Classifier combination consistently improved classification accuracy with the biggest improvement from multiple atlases. We conclude that multi-classifier techniques have a natural application to atlas-based segmentation and increase classification accuracy in real-world segmentation problems.

Keywords: Atlas-based segmentation; Multiple classifier system; Bagging; Non-rigid registration

1. Introduction

Combinations of multiple independent classifiers can be substantially more accurate than the individual classifications. Numerous applications of this principle have been reported in the pattern recognition literature over the past dozen years. Xu et al. (1992) evaluated different combination schemes of classifiers for the recognition of unconstrained handwritten numerals. They observed an increase of the recognition rates from 93% (best out of four individual classifiers) to 98.6% for several of their combination methods. Similarly, Kittler et al. (1998) combined four classifiers for optical character recognition of uppercase letters and digits, and found the classification rate to improve from 95% (best individual classifier) to over 98% for a combined classifier.

A common problem with the application of this principle is the construction of independent classifiers. A generic solution for this problem has been

* Corresponding author. Address: SRI International, Neuroscience Program, 333 Ravenswood Ave., Menlo Park, California 94025-3493, USA. Tel.: +1 650 859 3379; fax: +1 650 859 2743.
E-mail addresses: torsten@synapse.sri.com (T. Rohlfing), crmaurer@stanford.edu (C.R. Maurer Jr.).

0167-8655/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.patrec.2005.03.017

proposed by Breiman (1996) under the name "bootstrap aggregation", or simply "bagging". If the training of a classifier is unstable, that is, sensitive to the training set, multiple classifiers can be generated by using different training sets. In this paper, we propose and evaluate a classifier approach to atlas-based image segmentation. This approach allows for the systematic generation of a virtually unlimited number of classifiers in a natural way, similar to the bagging method.

The particular classification method we look at in this paper, atlas-based segmentation, has proven especially successful in current applications to segment three-dimensional (3D) biomedical images (Collins et al., 1995). The segmentation of an image is performed by computing an anatomically correct coordinate transformation between the image and an already segmented image, the so-called atlas. The process of computing the transformation between the two images is called registration, and is itself the subject of active research. A good overview of the field is given, for example, by Hajnal et al. (2001).

When comparing atlas-based segmentation to classifier techniques, the atlas assumes the role of the training set, and training of the classifier is achieved by computing the registration between image and atlas. We will take a more detailed look at the parallels between atlas-based segmentations and classifiers in Section 2.

Analogously to the bagging method, multiple independent classifiers arise in a natural way by using multiple training sets, that is multiple different atlases. Most published work on atlas-based segmentation uses atlases that were generated by manually segmenting the image of one individual. If two atlases are generated from different individuals, then the classifiers are inherently independent. Likewise, a different training method, e.g., a different registration algorithm, can be used to generate multiple different classifiers using the same atlas. In the absence of multiple working registration algorithms, multiple different parametrizations of the same algorithm may be used.

We evaluate three ways to generate multiple atlas-based classifiers. In Section 3, confocal microscopy images from 17 bee brains are segmented using three individual atlases, or using one atlas and three different values for each of two parametrizations of the registration algorithm. In each case, the resulting three individual segmentations are combined into a final segmentation. For each individual and for each combined segmentation, its recognition rate (relative number of correctly labeled image voxels) is computed using a manual ground truth segmentation.

In Section 4 we repeat the same experimental protocol with publicly available human brain magnetic resonance (MR) images from 10 subjects. Again, three subjects are used as atlases, and the images of the remaining 7 are segmented and the results compared to a manual segmentation. We conclude this paper in Section 5 with a discussion of our results, their relevance, and their practical implications.

2. Atlas-based classifiers

This section introduces a classifier view of atlas-based segmentation. Based on this foundation we then demonstrate the application of multi-classifier principles to atlas-based segmentation.

2.1. Analogies

Let us first take a look at the analogies between a generic classifier, say a neural network, and an atlas-based classifier. Table 1 gives an overview of the related concepts. From an operational perspective, an atlas-based classifier takes as its input a coordinate x inside the domain of the unsegmented target image, in the 3D case $\mathbb{R}^3$, and produces as its output the label assigned to that coordinate. It therefore computes a mapping from the image domain to a label set $\Lambda = \{1, \ldots, L\}$ with L labels (i.e., classes). Internally, this label is determined by lookup from a discrete label field, or atlas, $\mathscr{A}$, which is a mapping from $\mathbb{R}^3$ to Λ. The target image and the atlas are related to each other by a coordinate transformation $\boldsymbol{T}$. A simple atlas-based classifier thus implements the mapping

$$\boldsymbol{x} \mapsto \mathscr{A}(\boldsymbol{T}(x)). \tag{1}$$

The parameters of the transformation are the internal parameters of the classifier, so that training the

Table 1
Analogies between neural networks and atlas-based classifiers

	Neural network	Atlas-based classifier
Input	Pattern	Coordinate in image domain
Output	Class	Segmentation label
Internal structure	Network topology	Transformation model
Parameters	Connection weights	Transformation parameters
Parameter adjustment	Training	Registration
Learning input	Training set	Atlas
Typical learning algorithm	Backpropagation	Intensity-based non-rigid registration

classifier is equivalent to performing a registration between the image and the atlas.[1]

The coordinate transformation $\boldsymbol{T}$ of an atlas-based classifier is continuous, typically even smooth, and does not in general map grid points of the image to grid points in the atlas. Label look-up therefore requires an interpolation of labels of some sort.

The simplest interpolation that can be applied to non-numerical data such as labels is nearest neighbor (NN) interpolation, which returns the unique label of the nearest atlas grid point as the classifier output. A more sophisticated technique that is applicable to label data is partial volume (PV) interpolation, originally introduced by Maes et al. (1997) for histogram generation in entropy-based image registration. Using PV interpolation, a vector of weights is returned as the classifier output, where each weight represents the relative share of one label. The classifier therefore becomes a function that maps to $[0, 1]^{\Lambda}$, i.e.,

$$\boldsymbol{x} \mapsto (w_1, \ldots, w_L), \tag{2}$$

where w_l for $l = 1, \ldots, L$ is the relative weight (confidence) of x belonging to class l. In a multi-classifier system, each classifier returns a weight vector of dimension L. In this paper, we use PV interpolation, and we compute the joint classifier output for any given input by sum fusion, i.e., by adding all weight vectors and selecting the class with the highest weight in the sum.

From an image processing perspective, PV interpolation of the label map avoids jagged edges when generating oblique slices. From the classifier perspective, the fractional label weights resulting from PV interpolation in combination with sum fusion virtually eliminate the possibility of a tie in the classifier voting, thus greatly reducing the number of rejected inputs.

2.2. Classifier training: image-to-atlas registration

We use an intensity-based registration method based on the image similarity measure normalized mutual information (NMI), introduced by Studholme et al. (1999). The NMI measure quantifies the similarity of an image $\mathscr{I}$ with the gray-value component of an atlas $\mathscr{A}$ under a transformation $\boldsymbol{T}$.

In order to prevent unrealistic transformations due to incomplete or noisy data, the optimization function of the non-rigid registration algorithm is regularized with a constraint term $E_{\text{constraint}}$ that enforces smoothness of the deformation field. The overall optimization function E_{total} is a weighted combination of the NMI image similarity measure and the smoothness constraint:

$$E_{\text{total}}[\mathscr{I}, \mathscr{A}, \boldsymbol{T}] = (1 - w)E_{\text{NMI}}[\mathscr{I}, \mathscr{A} \circ \boldsymbol{T}] - wE_{\text{constraint}}[\boldsymbol{T}], \tag{3}$$

where $w \in \mathbb{R}$ with $0 \leqslant w < 1$. The negative sign of the constraint term reflects the need to minimize the constraint while maximizing the similarity term.

The constraint weight can be thought of as controlling the "stiffness" of the non-rigid transformation, i.e., the resistance of the coordinate space to deformation. Larger weights increase this

[1] Note that in fact an atlas consists of a gray-value image and a pixel-wise corresponding label field. We use the gray-value component for registration, but the label field for labeling the image that is being segmented.

resistance and lead to smoother deformations, but make it more difficult to model actual shape differences. Smaller weights decrease the resistance and make it easier to model the actual differences between both images, but increase the danger of incorrect transformations, or even folding of the coordinate space.

The coordinate mapping $\boldsymbol{T}$ between image and atlas is computed by a non-rigid registration algorithm introduced by Rueckert et al. (1999).[2] The transformation model is a free-form deformation (FFD) defined on a uniformly spaced grid of discrete control points (Sederberg and Parry, 1986). The degrees of freedom of this transformation model are the 3D positions of the control points, which move independently of each other. Control points in the image background are identified using a simple entropy-based criterion (Rohlfing and Maurer, 2001) and are kept fixed during registration to reduce the number of degrees of freedom and reduce computation time. A smooth deformation field is interpolated between the control points using approximating third-order B-splines. As detailed by Rohlfing et al. (2003), the smoothness constraint $E_{\text{constraint}}$ can be efficiently computed from the derivatives of the B-spline polynomials.

In the classifier training step, we compute the transformation $\hat{\boldsymbol{T}}$ between image $\mathscr{I}$ and the gray-value component of atlas $\mathscr{A}$ as

$$\hat{\boldsymbol{T}} = \arg\max_{\boldsymbol{T}} E_{\text{total}}[\mathscr{I}, \mathscr{A}, \boldsymbol{T}] \tag{4}$$

by optimizing over the space of all FFD transformations with a given control point lattice. The search for the optimal parameters in our implementation is a simple but efficient steepest-ascent optimizer.

To improve segmentation accuracy and robustness, and to be able to model both small and large deformations, we employ a multi-resolution deformation strategy. A coarse initial control point grid is repeatedly refined until a final resolution is reached. This final resolution, i.e., the final spacing between the FFD control points, is one important parameter of the algorithm.

It is important to note that other non-rigid registration algorithms can also be used for atlas-based segmentation. The method used for the present paper is merely one example of a technique that we have empirically found to be accurate and computationally efficient for a wide variety of biomedical applications.

2.3. Bagging of atlas-based classifiers

The principle idea of bagging as originally described by Breiman (1996) is to generate multiple independent classifiers by exploiting instability of classifier learning under changes to the learning set. In addition, one can exploit instability under different internal parametrization of the classifier.

In the context of atlas-based segmentation, using a different learning set means using a different atlas to register the unsegmented image to. The non-rigid registration algorithm used in this paper has operating parameters that influence its behavior. Two of these parameters are the spacing of the control point grid (or rather, the final spacing in a multi-resolution strategy), and the relative weight of the regularization term in the overall cost function (Eq. (3)). Changing either of these parameters changes the registration outcome, and therefore results in a different classifier, even when applied to the same atlas.

In particular, we evaluate the following strategies for multiple classifier generation:

(1) Multiple training sets (atlases): closely following the original idea of bagging, we generate multiple classifiers by using different independent training sets. In our context, this means using multiple atlases derived from different reference subjects. Each of the images to be segmented is registered independently to each of three atlases.
(2) Multiple regularization constraint weights: this results in variation of the "stiffness" of the coordinate transformation.

[2] Non-rigid registration is an active field of research and many algorithms have been reported (Bro-Nielsen and Gramkov, 1996; Thirion, 1998; Christensen and Johnson, 2001; Shen and Davatzikos, 2002). We chose the Rueckert algorithm because it is the one that we have the most experience with, and we have found it to be accurate, robust, and efficient for a wide range of applications. Any other effective non-rigid registration algorithm could be applied instead of the one we are using.

(3) Multiple FFD control point grid spacings: in addition to a variation of the spatial resolution of the deformation, changing the spacing also leads to a shift of control points.

As mentioned above, the classifiers internally use PV interpolation from their respective label maps to obtain for each input a weight vector that assigns a confidence value to each of the possible output classes. All classifier combinations are computed by sum rule fusion of these weights.

3. Evaluation study: confocal microscopy images

3.1. Methods

We first evaluate bagging of multiple atlas-based classifiers by segmenting 3D confocal microscopy images from 20 bee brains. The image acquisition details were previously described in (Rohlfing et al., 2004). In summary, each image contained 84–114 slices with thickness 8 μm and each slice had 610–749 pixels in x direction and 379–496 pixels in y direction with pixel size 3.8 μm. The image dimensions and pixel sizes are relevant insofar as the image resolution determines the ranges of meaningful values for various parameters of the non-rigid registration, most notably the FFD control point spacing. For each image in the population, a trained expert performed a complete manual segmentation into 22 anatomical structures (see Fig. 1 for an example microscopy image with corresponding ground truth segmentation). This segmentation can either serve as an atlas to segment other images, or as a ground truth to quantify the accuracy of an automatic segmentation.

From all individuals, a random subset of three individuals was designated for use as atlases. The images from the remaining subjects serve as the images to be segmented. The manual segmentations for these subjects serve as the ground truth used to evaluate the atlas-based segmentations. The registration parameters were 60 μm final control point spacing, and 0.1 as the smoothness constraint weight. We previously found these values to produce satisfactory segmentation accuracies (Rohlfing et al., 2004).

Out of the three selected atlases, one was randomly selected for the single-atlas experiments. In the first single-atlas experiment, three registrations were run for each image with smoothness constraint weight 0.05, 0.1, and 0.3, respectively. The final control point spacing was 60 μm.

For the second single-atlas experiment, three different control point spacings of the non-rigid registration algorithm were used, 50, 60, and 70 μm (which translated to control point numbers between 16,000 and 51,000). The smoothness constraint weight was 0.1 for all registrations in this experiment.

3.2. Results

The results of the three bagging methods applied to confocal microscopy segmentation are

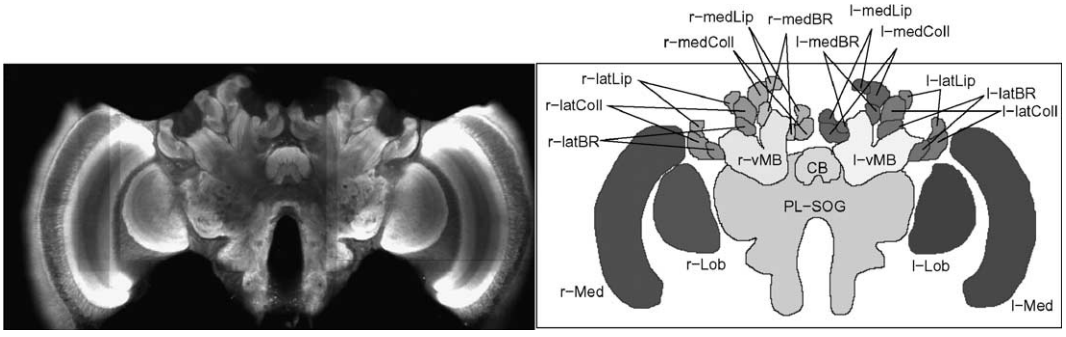

Fig. 1. Example of bee brain confocal microscopy image (left) and corresponding label image generated by manual segmentation (right). Each gray level in the label image represents one of 22 different anatomical structures. Due to limitations of reproduction, different gray levels may look alike. From Rohlfing et al. (2004).

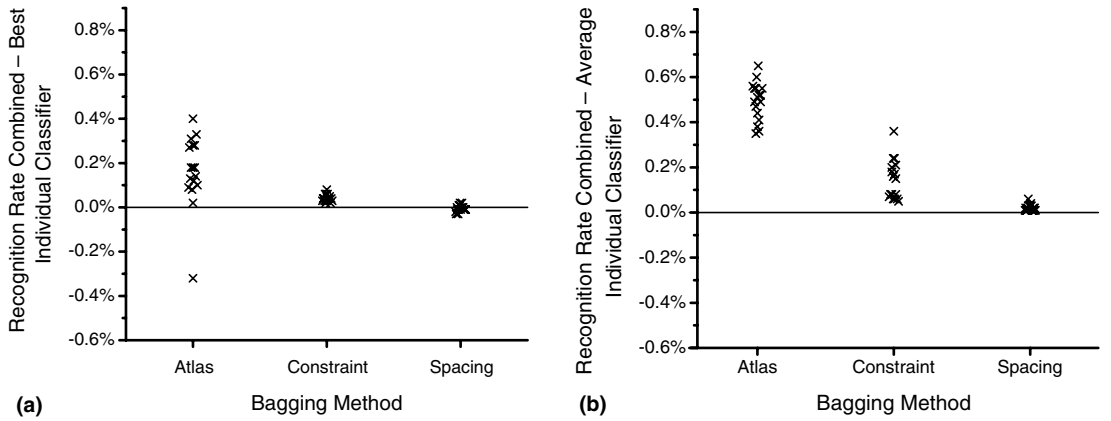

Fig. 2. Relative improvements of recognition rate by classifier combination for confocal microscopy image segmentation: (a) compared to the best of three individual classifiers; (b) compared to the average of three individual classifiers.

compared in Fig. 2. The actual recognition rates of mean and best individual classifiers as well as the combined classifications are given in Table 2. Combining the results from three different atlases, the combined classifier had a higher recognition rate than the *best* individual classifier for all but one segmented image. For all subjects, the combined classifier had a better recognition rate than the average of the three individual classifiers. On average, the recognition rate improved by 0.16%.

For different smoothness constraint weights, the combined classifier had a better or equal recognition rate compared to the best individual classifier for all 17 segmented images. However, the recognition rate on average improved by only 0.04%.

Using different control point spacings, the combined classifier had a better or equal recognition rate compared the best individual classifier for 8 out of 17 segmented images. On average, the combined classifier performance was virtually identical to the best individual classifier performance.

Table 2
Total recognition rates for confocal microscopy image segmentation

Bagging method	Individual		Combined
	Mean	Best	
Atlas	0.9724	0.9757	0.9773
Constraint	0.9723	0.9733	0.9738
Spacing	0.9725	0.9728	0.9727

Each value is the average over segmented images from 17 different subjects.

4. Evaluation study: human brain MR images

4.1. Methods

As a second evaluation study, we apply bagging of atlas-based classifiers to publicly available human brain MR images from the Internet Brain Segmentation Repository (IBSR). For this purpose, we obtained T_1-weighted anatomical images from ten subjects. We also obtained the corresponding manual segmentations with 43 anatomical structures. All images had the same size, $256 \times 256 \times 128$ voxels, with coronal slice orientation. The in-plane pixel size was either 0.9 mm or 1.0 mm. The slice thickness of all images was 1.5 mm. For an example MR image with corresponding segmentation see Fig. 3.

As before, three subjects were randomly selected as atlases. The images from the remaining seven subjects served as test cases for the segmentation. The manual segmentations of the test cases provided the ground truth for performance evaluation. Since the image dimensions and properties the human brain MRI are fundamentally different from confocal microscopy images of bee brains, the parameters of the non-rigid registration algorithm were adapted accordingly. The registration

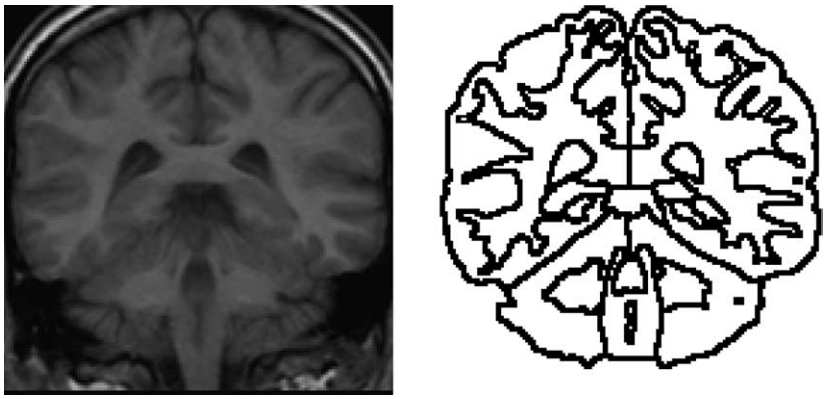

Fig. 3. Example of human brain MR image (left) and outlines of the corresponding ground truth segmentation (right). The MR image and the ground truth segmentation were obtained from the Internet Brain Segmentation Repository (IBSR).

parameters for multi-atlas segmentation were 30 mm final control point spacing and smoothness constraint weight 0.1.

In the first single-atlas experiment, three registrations were run for each image with smoothness constraint weight 0.03, 0.05, and 0.1, respectively. The initial control point spacing was 30 mm, which was refined three times to a final spacing of 3.75 mm.

For the second single-atlas experiment, three different control point spacings of the non-rigid registration algorithm were used, 3.125, 3.75, and 4.375 mm (which translated to control point numbers between 178,000 and 555,000). The smoothness constraint weight was 0.1 for all registrations in this experiment.

4.2. Results

The results of the three bagging methods applied to human brain MR images are compared in Fig. 4. The actual recognition rates of mean and best individual classifiers as well as the combined classifications are given in Table 3. Combining the results from three different atlases, the combined classifier had a higher recognition rate

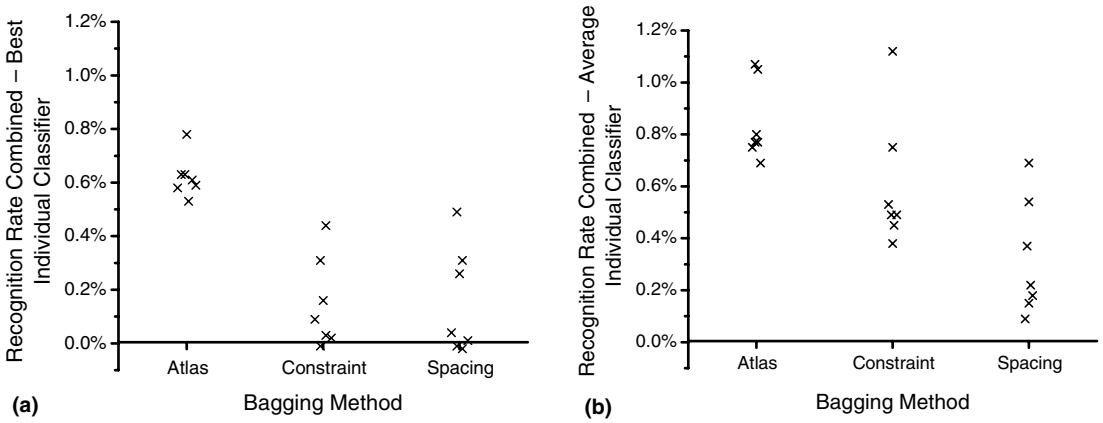

Fig. 4. Relative improvements of recognition rate by classifier combination for human brain MR image segmentation: (a) compared to the best of three individual classifiers; (b) compared to the average of three individual classifiers.

Table 3
Total recognition rates for human brain MR image segmentation

Bagging method	Individual		Combined
	Mean	Best	
Atlas	0.9651	0.9673	0.9735
Constraint	0.9622	0.9676	0.9682
Spacing	0.9646	0.9663	0.9678

Each value is the average over segmented images from 7 different subjects.

than the best individual classifier for all 7 segmented images. On average, the recognition rate improved by 0.6% over the best classifier.

For different smoothness constraint weights, the combined classifier had a better or equal recognition rate compared the best individual classifier for 6 out of 7 segmented images. The recognition rate on average on average improved by 0.1%.

Using different control point spacings, the combined classifier had a better recognition than the best individual classifier for 5 out of 7 segmented images. On average, the combined classifier performance improved by 0.2%.

5. Discussion

The main advantage of a classifier view of atlas-based segmentation is that it opens the field to the application of multi-classifier decision fusion techniques. We have shown in previous work (Rohlfing et al., 2004) that segmentation accuracy can be significantly improved when more than a single atlas is used. However, multiple atlases are not always available, since their generation is time consuming and tedious. This makes the message of the present paper ever more important—it has demonstrated that improvements of segmentation accuracy are possible with only a single atlas, but different parametrization of the non-rigid registration technique.

This paper has evaluated three different ways of generating multiple atlas-based classifiers: use of multiple atlases, multiple deformation control point spacings, and multiple regularization constraint weights. Our results suggest that there is a quantitatively bigger accuracy gain from using multiple atlases than there is from using different regularization weights, which itself produced bigger gains than multiple control point spacings. We believe that the reason for this is that atlases obtained from multiple reference individuals are truly independent and also reflect, to some extent, the variability of a population. Whenever possible, multiple atlases should therefore be used. Of course, in addition to using multiple atlases, further improvements may be possible by registering to each atlas several times with different registration parameters.

Between the two registration parameters we considered, control point spacing and smoothness constraint weight, the latter appeared to have a bigger impact on the segmentation result. The fact that decision fusion effectively benefits from multiple classifiers resulting from variations of this parameter is particularly important as there is no a priori correct value for it. Since the constraint weight relates fundamentally unrelated quantities, image similarity and deformation energy, classifier fusion allows us to cover a range of possible values without having to pick one. The effectiveness of variations of the constraint weight is also relevant, because, unlike the control point spacing, this type of parameter is also present in any other registration algorithms based on computing an equilibrium between external forces (i.e., image similarity) and internal forces (i.e., a physics-based constraint such as elasticity or volume preservation).

For control point spacing, on the other hand, we can basically say that ''more is better'', that is, the finer the control point grid, the more successfully the transformation can model inter-individual shape differences, and the more accurate the registration and thus the segmentation will be (unpublished results). At least this is true for segmentation of the confocal microscopy images of bee brains. For human brain MRI, different control point spacings seem to be an effective bagging method. We suspect that this is a result of the more complex shapes of structures segmented in this data (compare Figs. 1 and 3), which presumably makes the segmentation more sensitive to the control point spacing.

Overall, the relative improvements achieved through classifier combination reported in this paper may appear marginal, but there are several factors that illustrate the effort may still be worthwhile. First and most important, we considered combinations of only three classifiers, which is the minimal number that can be combined in a meaningful way. In previous work, the use of a larger number of classifiers produced substantially better combined classification results (Rohlfing et al., 2004). Second, the improvements reported here are relative to the best individual classifier. However, the best individual classifier can hardly ever by identified without access to the ground truth classification. Therefore, from a practical point of view, the typical application performance improves more than our numbers may suggest.

Finally, atlas-based segmentation, due to the continuous and smooth transformation model, preserves topology. Errors therefore tend to occur on and around the surface of segmented structures. If the fraction of surface voxels among the total image voxels is small, segmentation improvements only lead to small improvements of the total recognition rate. This is supported by an additional analysis, where we computed the recognition rate improvement for foreground voxels only. For the confocal microscopy data, the improvement of combining segmentations using three different atlases over the best individual classifier was 0.9% (2.3% compared to average of individual classifiers). For the human brain MR data, this value was 2.4% (4.5%).

The obvious disadvantage of the suggested procedure is the repeated application of a computationally expensive non-rigid registration step. We have therefore limited our consideration in this paper to three segmentations per subject. Even so, we observed noticeable and consistent accuracy improvements. In the future, with ever increasing computer speeds, we consider registration times to be not too serious of an issue. Furthermore, non-rigid registration can be parallelized with very low overhead (Rohlfing and Maurer, 2003), and multiple independent registrations can easily be performed on a cluster of inexpensive computation nodes.

Moving past the combination of multiple atlas-based classifiers, our framework provides a theoretical foundation for combining atlas-based segmentations with segmentations obtained from fundamentally different methods. They could be combined, for example, with results generated by level set techniques, and with such originating from active contour methods. The more dissimilar the methods combined, the more likely their errors are independent.

Acknowledgments

The authors thank Robert Brandt (Mercury Computer Systems GmbH, Berlin, Germany) and Randolf Menzel (Freie Universität Berlin, Berlin, Germany) for providing segmented bee brain microscopy images. The normal MR brain data sets and their manual segmentations were provided by the Center for Morphometric Analysis at Massachusetts General Hospital and are available at http://www.cma.mgh.harvard.edu/ibsr/. All non-rigid registrations were performed on an SGI Origin 3800 supercomputer in the Stanford University Bio-X core facility for Biomedical Computation.

This paper is an extended version of work presented at the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), Washington, DC, June 26–July 2, 2004 (Rohlfing and Maurer, 2004).

References

Breiman, L., 1996. Bagging predictors. Mach. Learn. 24 (2), 123–140.

Bro-Nielsen, M., Gramkov, C., 1996. Fast fluid registration of medical images. In: Proc. SPIE Medical Imaging. SPIE, pp. 267–276.

Christensen, G.E., Johnson, H.J., 2001. Consistent image registration. IEEE Trans. Med. Imag. 20 (7), 568–582.

Collins, D.L., Holmes, C.J., Peters, T.M., Evans, A.C., 1995. Automatic 3-D model-based neuroanatomical segmentation. Human Brain Mapp. 3 (3), 190–208.

Hajnal, J.V., Hill, D.L.G., Hawkes, D.J. (Eds.), 2001. Medical Image Registration. CRC Press, Boca Raton, FL.

Kittler, J., Hatef, M., Duin, R.P.W., Matas, J., 1998. On combining classifiers. IEEE Trans. Pattern Anal. Mach. Intell. 20 (3), 226–239.

Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., Suetens, P., 1997. Multimodality image registration by maximisation of mutual information. IEEE Trans. Med. Imag. 16 (2), 187–198.

Rohlfing, T., Brandt, R., Menzel, R., Maurer Jr., C.R., 2004. Evaluation of atlas selection strategies for atlas-based image segmentation with application to confocal microscopy images of bee brains. NeuroImage 21 (4), 1428–1442.

Rohlfing, T., Maurer Jr., C.R., 2001. Intensity-based non-rigid registration using adaptive multilevel free-form deformation with an incompressibility constraint. In: Niessen, W., Viergever, M.A. (Eds.), Proc. Fourth International Conf. on Medical Image Computing and Computer Assisted Intervention (MICCAI 2001), Lecture Notes in Computer Science, vol. 2208. Springer-Verlag, Berlin, pp. 111–119.

Rohlfing, T., Maurer Jr., C.R., 2003. Nonrigid image registration in shared-memory multiprocessor environments with application to brains, breasts, and bees. IEEE Trans. Inform. Technol. Biomed. 7 (1), 16–25.

Rohlfing, T., Maurer Jr., C.R., 2004. Multi-classifier framework for atlas-based image segmentation. In: Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, Washington, DC, USA, June 27–July 2, 2004, Part I. IEEE Press, Los Alamitos, CA, pp. 255–260.

Rohlfing, T., Maurer Jr., C.R., Bluemke, D.A., Jacobs, M.A., 2003. Volume-preserving nonrigid registration of MR breast images using free-form deformation with an incompressibility constraint. IEEE Trans. Med. Imag. 22 (6), 730–741.

Rueckert, D., Sonoda, L.I., Hayes, C., Hill, D.L.G., Leach, M.O., Hawkes, D.J., 1999. Nonrigid registration using free-form deformations: Application to breast MR images. IEEE Trans. Med. Imag. 18 (8), 712–721.

Sederberg, T.W., Parry, S.R., 1986. Free-form deformation and solid geometric models. Comput. Graphics 20 (4), 151–160.

Shen, D., Davatzikos, C., 2002. HAMMER: Hierarchical attribute matching mechanism for elastic registration. IEEE Trans. Med. Imag. 21 (11), 1421–1439.

Studholme, C., Hill, D.L.G., Hawkes, D.J., 1999. An overlap invariant entropy measure of 3D medical image alignment. Pattern Recognition 32 (1), 71–86.

Thirion, J.P., 1998. Image matching as a diffusion process: An analogy with Maxwell's demons. Med. Image Anal. 2 (3), 243–260.

Xu, L., Krzyzak, A., Suen, C.Y., 1992. Methods of combining multiple classifiers and their applications to handwriting recognition. IEEE Trans. Syst. Man Cybernet. 22 (3), 418–435.