Ensemble registration: Combining groupwise registration and segmentation

Ensemble registration: Combining groupwise registration and segmentation

Sri Purwani[1,2]
sri.purwani@postgrad.manchester.ac.uk
Tim Cootes[1]
t.cootes@manchester.ac.uk
Carole Twining[1]
carole.twining@manchester.ac.uk

[1] Imaging Science, The University of Manchester, UK
[2] Mathematics Department, Faculty of Mathematics and Natural Sciences, Padjadjaran University, Indonesia

Abstract

We investigate incorporating structural information from segmentation into a groupwise registration framework. Previous work by Petrovic et al., using MR brain images, showed that using tissue fractions to help construct an intensity reference image gives better results than just using intensity images alone. In their work, a Gaussian Mixture Model (GMM) was fitted to the 1D intensity histogram, then used to construct tissue fraction images for each example. The mean fraction images were then used to create an artificial intensity reference for the registration.

By using only the mean, this discarded much of the structural information. We retain all this information, and augment each intensity image with its set of tissue fraction images (and also intensity gradient images) to form an image ensemble for each example. We then perform groupwise registration using these ensembles of images.

This groupwise ensemble registration is applied to the same real-world dataset as used by Petrovic et al. Ground-truth labels enable quantitative evaluation to be performed. It is shown that ensemble registration gives quantitatively better results than the algorithm of Petrovic et al., and that the best results are achieved when more than one of the three types of images (intensity, tissue fraction and gradient) are included as an ensemble.

1 Introduction

Previous work [1,6,7] has shown that integrating segmentation and registration in a unified framework gives better results. Asburner et al. [1] showed that repeated iteration among segmentation, bias correction, and registration of brain images gives better results than serial applications of each component. On the other hand, Konukoglu et al. [6] integrated segmentation into registration to improve robustness of the registration. The segmentation automatically labels structures of human anatomies in its found regions. This integration shows improvement on CT image alignments.

In Petrovic et al. [7], a Gaussian Mixture Model (GMM) is fitted to the intensity distribution of MR brain images. Each image is taken to consist of three main tissues (CSF, grey matter (GM) and white matter (WM)) plus partial-volume voxels. This model then enables the construction of a set of tissue fraction images for each image in the set. The mean fraction images across the group are then used to create an artificial intensity

reference to be registered with its original intensity image during registration. Their results out-performed those which just used intensity images alone.

However, the use of a 1D intensity image histogram in their method sometimes suffers from the problem of failing to find the correct parameters, such as the tissue mean intensities. Hence segmentation fails for some images, as is shown in Figure 1(a). However, since just the ***mean*** fraction images were used to create the reference, their algorithm was not overly sensitive to these failures. However, this also means that only a limited amount of the structural information is actually used.

Our current method is designed to use all of the structural information available. This means including not just the tissue fraction images, but also intensity and gradient images to form an ensemble of images for registration. This integration hence makes better use of all available information for alignment of structures.

However, this requires a more robust GMM fitting algorithm than that used by Petrovic et al. [7]. We achieve this by using the 2D intensity value-and-gradient histogram [2], which gives a more robust estimate of the GMM intensity parameters. We show that our method improves on that of Petrovic et al. [7]. Tanimoto overlap values [3] for the ground truth labels are used to evaluate the registration, and show that combining image types into an ensemble gives the best results.

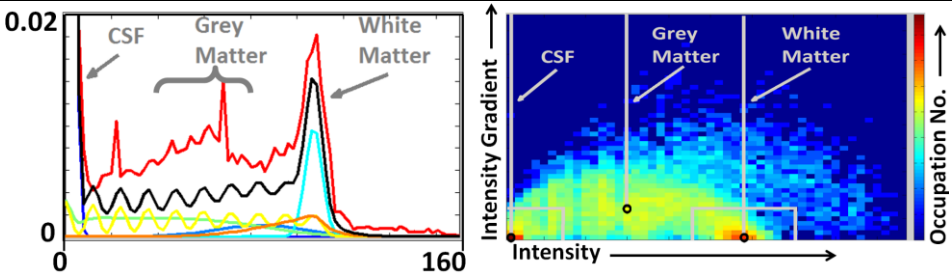
2 Method

Given a set of images, we are going to perform groupwise ensemble registration from each example image in the set. The way to do this is by segmenting each image into tissue fraction images and then constructing an ensemble of images consisting of those fraction images and the original intensity image (plus intensity gradient image derived from the original image). This can be viewed in two ways: either as each example consisting of a stack (or ensemble) of images, or as each example image now being a vector-valued feature image, with intensity, fraction, gradient etc. being components of the vector. This will be further explained in the following sections.

2.1 Tissue Segmentation

We consider the case where different tissues in an image have different intensity distributions hence tissue segmentation can be performed by analysis of the intensity histogram. In Figure 1(a), we show a typical 1D intensity histogram (red). The GMM fitting algorithm used by Petrovic et al. [7] gives the model shown in black. In this case the 1D algorithm ***fails*** to correctly locate the pure-grey-matter Gaussian, and the segmentation fails. We hence use the 2D intensity value-and-gradient histogram instead [2]. As can be seen in Figure 1(b), the pure tissue peaks are located at the feet of the arcs that correspond to partial-volume mixtures of tissues. The gradient, which describes intensity changes within images, is highest at the boundary between tissues, hence can separate partial-volume boundary pixels from the pure-tissue pixels (see [2] for further details). So it could be said that as well as gradient information being useful for registration [5], this shows that it is useful for segmentation as well.

A greedy algorithm is used to locate the peaks/means for model fitting (Figure 1(b)). After finding the first peak, this algorithm defines an excluded zone (grey lines) before finding the next peak. This is to avoid having multiple local peaks, and is repeated until all



(a) 1D histogram (b) 2D histogram

Figure 1: Intensity 1D histogram (Left) and the new intensity-gradient 2D histogram (Right), to the same horizontal scale. (a) The red line is the image histogram, and the black line is the ***failed*** GMM fit, which consists of a sum over 3 pure tissue Gaussians, plus 3 mixtures (coloured lines); (b) The 2D Intensity-gradient vs Intensity histogram. The yellow/red regions at the feet of the yellow arcs give the new estimates of the pure-tissue means (shown by the vertical grey lines and the black circles).

three peaks are found. These peaks give the means for the three Gaussians, which are kept fixed during the optimisation of the fit, and we optimise all the other parameters such as variance and weights by using Sum of Squared Differences.

We have found that using this procedure, it was possible to correctly fit all the image histograms in the set, including those where the 1D histogram fitting procedure is failed.

2.2 Ensemble Registration

Now we have robust segmentation of the histogram, this can be used to relate image intensities to tissue content [2], hence segment the image to obtain 3 tissue-fraction images. These can then be combined with the original intensity image. As noted above, we also include the intensity gradient image [5]. This then gives a 5 component image ensemble for each example in the set. Registration can then be performed to the reference in the usual way, but with a separate reference image for each component. The objective function then must be summed across components as well as across the set. Using Sum of Absolute Differences, we then have an objective function of the form:

$$\mathcal{L} = \sum_{\alpha=1}^{n} \sum_{j=1}^{m} \sum_{A} |I_{\alpha}^{j}(x_A) - \tilde{I}_{ref}^{j}(x_A)| \qquad \textbf{(1)}$$

where $\alpha = 1, 2, \ldots, n$, $j = 1, 2, \ldots, m$ are the indices for example number from the set, and image ensemble component respectively, whereas A runs over all pixels in the image. For each component, $\tilde{I}_{ref}^{j}$ is the reference image for that component, and is taken to be the mean over the set. Each image in the ensemble is scaled to lie in the range 0 to 1, with the full range of values retained, so that the different components are commensurate. Each component hence contributes equally to the objective function. It would obviously also be possible to weight the individual components, but here we are just interested in using a component compared to not using it, hence we give all components equal weight.

The same multi-resolution image warping framework (piece-wise affine deformation) as was used in [7] is applied, to enable a true comparison of the registration objective functions. The algorithm is applied to the same dataset as in [7] (37 MR brain image slices, from 37 different individuals, of size 190 × 190 pixels).

2.3 Evaluation: Ground Truth Labels

On this set of MR brain images there are 64 ground truth tissue labels. However, some labels are of very small volume. Hence to get a more informative comparison, Petrovic et al. [7] consider just the main labels having the most contribution to the image set, which are the 8 labels consisting of white matter, grey matter, caudate, and lateral ventricle, left and right in each case. For analysis and comparison purposes with other datasets, we arrange our labels into 4 tissue sets, CSF (not including background), GM, WM, and other tissues. Some samples of ground truth labels from this image set are given in Figure 2.

We perform quantitative evaluation of the registration using Tanimoto overlaps on these different tissue sets, which we found to be more informative than the single variously-weighted combination used in [7]. These evaluation results are presented in the experiment section.

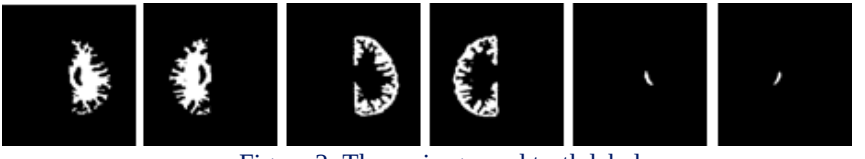


Figure 2: The main ground truth labels.

3 Experiments

For our initial comparisons, we consider two modifications to the original method of Petrovic et al. [7] (results in Table 1). It was noted that they had three pairwise mixtures in their method, including the mixture WM-CSF. Given the problems with segmentation (see Section 2.1), we first restricted mixtures to those sorted by intensity[1], so we include CSF-GM, and GM-WM, but ***not*** CSF-WM [4]. Together with pure tissues, this then gives 5 possible classes. We also modify slightly the mappings from intensity values to fraction values. In [7] (see Eqn. (4)), a hard assignment of pixels to the most probable class was made, whereas we use a softer assignment [4], which gives a smoother, continuous mapping. In the second modification, we additionally replace the 1D histogram method with the 2D histogram method (see Section 2.1). As well as these slight modifications of the original algorithm, we also include the Ensemble registration (1). For completeness, we consider all seven possible combinations of components in the ensemble, where F denotes the inclusion of the 3 fraction images, I denotes intensity images, and G intensity gradient.

It is slightly surprising that the first two modifications, which could be viewed as having a better theoretical justification, either make things worse, or have a negligible effect. More importantly, it can be seen that the Ensemble registration can produce slightly better results than the algorithm of Petrovic et al. [7]. In terms of single-component-type registration, in general gradient alone is the worst, followed by intensity, but with fraction images being the best. This is perhaps explained by the fact that the single-tissue mean intensities vary for this dataset (which is why the example-specific reconstructed intensity reference images were created by Petrovic et al. originally). This might mean that gradient (which removes this factor) should perform better than intensity is. Except the largest gradients are in the region of the skull (see Figure 3), but there are no labels (see Figure 2) associated with the skull, so that this region does not even enter into the evaluation!

[1] See Figure 1b), where the two prominent yellow arcs correspond to the two included mixtures, with little evidence of a CSF-WM arc (the other arc corresponds to skull mixtures, which are not included in the GMM).

The pairs-of-component-types results respect this ordering, so that fraction-plus-gradient outperforms intensity-plus-gradient, intensity-plus-fraction outperforms intensity plus-gradient, and fraction-with-intensity is better than fraction-with-gradient. Finally, the combination of all three component types is always ranked first or second.

These results clearly show that the explicit structural information in the tissue fraction images is needed, and that it seems to be of better value than simpler quantities such as intensity gradient (since we take the absolute magnitude of the gradient, rather than a full vector-valued gradient image, some information is lost when we take the gradient).

Figure 3 shows initial and final component reference images respectively for the full ensemble registration (FIG). This shows that as registration progresses the alignment of structures improve, and the reference images consequently become sharper for all 5 components of the ensemble.

Registration	Tanimoto Values for Tissue Set			
	CSF	grey	white	others
Petrovic et al. [7]	0.6854(9)	0.5804(6)	0.700(1)	0.458(3)
Modification 1	0.665(4)	0.576(1)	0.6884(7)	0.457(7)
Modification 2	0.675(2)	0.571(2)	0.698(1)	0.459(4)
F	0.678(4)	0.575(1)	0.7021(8)	0.507(4)
I	0.646(3)	0.5393(7)	0.652(3)	0.32(1)
G	0.524(4)	0.468(1)	0.6211(9)	0.499(9)
I F	**0.690(5)**	**0.582(1)**	**0.7069(1)**	0.508(5)
G F	0.675(5)	0.577(1)	0.704(1)	**0.515(4)**
G I	0.666(2)	0.553(1)	0.6792(6)	0.462(4)
F I G	**0.686(3)**	**0.584(1)**	**0.709(1)**	**0.509(2)**

Table 1: Tanimoto values by tissue set, for various registration methods. The best two values in each column are highlighted. Bracketed numerals show the estimated errors.

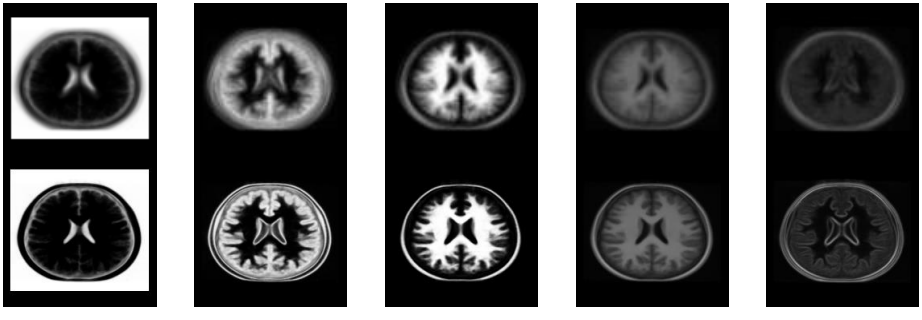


Figure 3: The top and bottom rows show initial and final component reference images respectively for the full ensemble registration (FIG). From Left to Right, the components are CSF & background, GM, and WM, intensity, and intensity gradient magnitude.

4 Discussion and Conclusions

We have demonstrated that integrating segmentation and registration can improve the registration results over using intensity images alone. As known correspondences achieved by registration can help segmentation to find the structures, on the other hand the found structures from segmentation can speed up the registration. Hence their integration can benefit each other.

In order to achieve our integration, we had to improve the performance of the segmentation stage. This was done by using a 2D intensity value-and-gradient histogram rather than a simple 1D intensity histogram. This gave a more robust segmentation of the images into tissue-types as shown in Figure 3.

We have also shown that adding further structural information, in the form of tissue fraction images, further improved registration results. We have shown that Ensemble registration using combinations of more than one type of images (intensity, fraction, and intensity gradient) can also produce slightly better results than previous algorithms [7]. This previous work also used segmentation but in a rather different way, in that it was used to reconstruct an artificial intensity reference image from reference fraction images.

Finally, it should be noted that whilst our results clearly show that combinations of the various types of images are needed to get the best Ensemble result, these results only used a naïve equal-weighting of components (1), and it is possible that tuning the weighting of the components could give a further improvement.

5 Future Work

In order to enable a direct comparison with previous work, we have reported extensive experiments on the set of 37 MR brain image slices as used previously. However, we also have available 5 other similar (but larger) datasets, with ground-truth annotation, and we are in the process of performing a similar series of experiments on these other datasets. In particular, there are no theoretical reasons why the algorithm cannot be applied to 3D data.

Our current experiments just use a repeated pairwise matching to the mean method of groupwise registration. Our next extension of our method is to consider all the image set during registration, by using a fully groupwise algorithm.

Furthermore, we also intend to consider extending our method to other types of images, such as radiographs of hands. This will be more challenging, since the tissue segmentation needed will be different. Such radiographs may introduce other problems, since they are created from projections of 3D objects, unlike MR images which are truly volumetric.

Acknowledgment: Ms Purwani would like to thank the Directorate General of Higher Education of Indonesia (DIKTI) for providing funding.

References

[1] J. Ashburner and K. J. Friston. Unified segmentation. *NeuroImage*, 26: 839-851, 2005.

[2] P. A. Bromiley and N. A. Thacker, Multi-dimensional medical image segmentation with partial volume and gradient modelling, Annals of the BMVA 2008, vol. 2, 2008, pp. 1–22.

[3] W. Crum, O. Camara, and D. Hill. Generalised overlap measures for evaluation and validation in medical image analysis. *IEEE Trans. Medical Imaging*, 25:1451-1461, 2006.

[4] H. D. Gage, W. E. Snyder, and P. Santago. Quantification of brain tissue through incorporation of partial volume effects. *Medical Imaging*, 1652:8496, 1992.

[5] E. Haber and J. Modersitzki, Intensity gradient based registration and fusion of multi-modal images, MICCAI 2006. Springer Berlin Heidelberg, 2006. 726-733.

[6] E. Konukoglu, A. Criminisi, S. Pathak, D. Robertson, S. White, and K. Siddiqui, Robust linear registration of CT images using random regression forests, *SPIE Medical Imaging*, vol. 7962, 2011, p. 79621X.

[7] V. Petrovic, T. F. Cootes, C. J. Twining, and C. J. Taylor, Simultaneous registration, segmentation, and modelling of structure in groups of medical images. In *4th IEEE International Symposium on Biomedical Imaging (ISBI), From Nano to Macro*, pages 1–4, 2007.