

Geodesic Shape-Based Averaging

M. Jorge Cardoso¹, Gavin Winston², Marc Modat¹, Shiva Keihaninejad³,
John Duncan², and Sebastien Ourselin^{1,3}

¹ Centre for Medical Image Computing (CMIC), UCL, UK

² Epilepsy Society MRI Unit, Dep. of Clinical and Experimental Epilepsy, UCL, UK

³ Dementia Research Centre (DRC), UCL, UK

Abstract. A new method for the geometrical averaging of labels or landmarks is presented. This method expands the shape-based averaging [1] framework from an Euclidean to a geodesic based distance, incorporating a spatially varying similarity term as time cost. This framework has unique geometrical properties, making it ideal for propagating very small structures following rigorous labelling protocols. The method is used to automate the seeding and way-pointing of optic radiation tractography in DTI imaging. The propagated seeds and waypoints follow a strict clinical protocol by being geometrically constrained to one single slice and by guaranteeing spatial contiguity. The proposed method not only reduces the fragmentation of the propagated areas but also significantly increases the seed positioning accuracy and subsequent tractography results when compared to state-of-the-art label fusion techniques.

1 Introduction

Diffusion magnetic resonance imaging (diffusion MRI) is an imaging technique that provides insights about the pattern of diffusion of water molecules in the brain, frequently represented by a tensor. Tractography algorithms have been proposed to characterise and delimit white matter fibre bundles [2]. However, diffusion imaging and subsequently tractography techniques are prone to imaging artefacts and algorithmic limitations [3]. Due to these limitations, the most commonly used technique for tractography initialisation is still manual localisation of tractography seeds, waypoints and exclusion zones. Nevertheless, diffusion based image analysis techniques are increasingly used for their ability to characterise white matter connectivity and microstructure [4], possibly leading to the development of biomarkers for neurodegenerative disease progression[5].

Ideally, one would like to reduce both human interaction time and inter-rater variability by standardising the tractography procedure using automated methods. A basic step which is common to many recently proposed automated seeding and way-pointing procedures is multi-atlas based information propagation through inter-subject alignment of multiple image volumes. This alignment involves either a single [6] or multiple reference anatomies [7] and fusion of candidate segmentations using a multitude of methods normally adapted from computer vision. For example, Suarez *et al.* [7] proposes to use non-rigid registration

based multi-atlas propagation combined with the STAPLE (Simultaneous Truth and Performance Level Estimation) fusion framework for automated seed propagation. Note that, contrarily to T1 anatomical label fusion frameworks, none of these automated methods have explored the use of image similarity metrics to improve the fusion accuracy. Furthermore, the label fusion methods used in previous works do not permit geometrically restrictive protocols where only a very limited number of voxels and/or slices are allowed.

In this work we present an extension to Shape-based Averaging (SBA). The Euclidean distance used in SBA is replaced by a geodesic distance in order to integrate a image similarity term in the label fusion algorithm. The geodesic SBA enables the robust propagation of sparse and non-overlapping labels and allows the inclusion of geometrical and volumetric constrains. To the best of our knowledge, this paper presents the first framework that incorporates both an image similarity metric and shape-based label fusion.

2 Methods

This section first introduces the classical SBA framework and its unique geometrical properties, followed by a reformulation of the framework using a geodesic distance metric. Finally, the similarity based geodesic time cost function is described within the geodesic SBA framework.

2.1 SBA and Its Geometrical Properties

The intuitive concept behind the original Shape-based Averaging method [1] is that if one wants to fuse a set of candidate labels from a number of classifiers, the fused label would be the one which has the smallest Euclidean distance to the boundaries of all the candidate labels, i.e. the one with the mean shape.

For K input images over R^m and $k = 1, \dots, K$, let $s_k(i)$ be the label value at location i . Each label is a number in the set $\Lambda = \{0, 1, \dots, L - 1\}$, where zero without loss of generality represents the image background. Now let a distance maps $d_{k,l}(i)$ be the signed Euclidean distance transform (EDT) from the pixel i in image k to label l . The signed EDT for each label is computed as the distance between pixel i and the edge of label l , with its value being positive if pixel i is on the inside of the edge, and negative if i is on the outside. Formally, $d_{k,l}(i)$ is negative if $s_k(i) = l$ and $d_{k,l}(i)$ is positive if $s_k(i) \neq l$. It is important to note that $d_{k,l}$ is a convex function.

Then, for each label one can then calculate the mean Euclidean distance from pixel i to the edge of all candidate labels in all K images, as

$$D_l(i) = \frac{1}{K} \sum_{k=1}^K d_{k,l}(i) \quad (1)$$

These Euclidean distance transforms from different labels are then fused by minimizing the mean distance over all labels, $S(i) = \operatorname{argmin}_{l \in \Lambda} D_l(i)$. A key feature

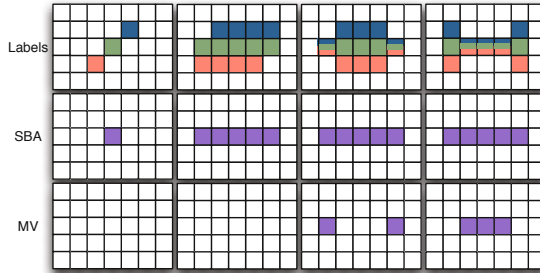


Fig. 1. Top) Four examples with three labels to fuse (in blue, red and green). Overlapping labels are represented by a partial pixels. Middle) Results from the SBA algorithm in purple. Note that the fused label is continuous even if the labels do not overlap. Bottom) Results from the Majority Voting algorithm in purple. Majority voting only places labels in overlapping areas, leading to empty or non-connected results.

to note is that D_l is convex as it is the sum of a series of convex functions $d_{k,l}$. Due to this particularity, SBA is known to produce smooth fusions with reduced structural fragmentations [1]. Also, it is possible to obtain a value of distance from the shape-based average boundary at every pixel and for every label. The combination of the convexity of the function and the fact that one can sample this distance at any point provides unique geometrical properties to the SBA framework. It allows the fusion of labels that do not overlap and labels with completely different shapes. A graphical depiction is show in Fig. 1. These geometrical properties are not present in majority voting, weighted voting and STAPLE label fusion techniques.

2.2 Geodesic Distance Transform

A natural extension of the above described algorithm is the introduction of classifier performance weights. If the proposed algorithm is seen in the perspective of multi-label propagation and fusion, where a series of labels from a template database are propagated to a new unseen image using image registration, then each propagated label should have a different fusion weight depending on how similar the unseen image is to the registered image after transformation. The introduction of classifier performance can help the overall label propagation performance by giving insights about registration accuracy and morphological similarity between images.

In this perspective, the Eq. 1 can then be extended to

$$D_l(i) = \frac{\sum_{k=1}^K W_k(i) d_{k,l}(i)}{\sum_{k=1}^K W_k(i)} \quad (2)$$

where $W_k(i)$ represent the similarity between the propagated template images k and the current image at position i . If this similarity $W_k(i)$ is spatially varying, one cannot guarantee that $W_k(i) d_{k,l}(i)$ will be monotonically increasing. This

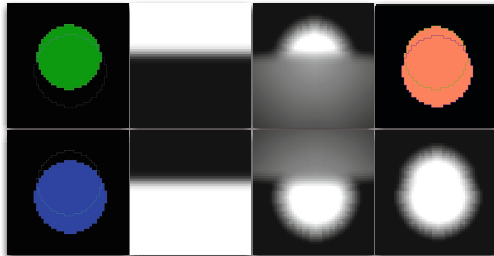


Fig. 2. a) Two labels in blue and green; b) A similarity function $W_k(i)$ for each label with white being similar and black being dissimilar; c) The resultant geodesic distance transforms $d_{k,l}(i)$; d) The final set $S(i)$ on the top and the distance $D_l(i)$ at the bottom. All distances are thresholded between -4 and 4 for the purpose of visualisation.

would only occur if $W_k(i)$ is constant for every i , where $D_l(i)$ is reduced to a weighted sum of Euclidean distance transforms.

In order to maintain the monotonicity of $d_{k,l}(i)$ while introducing a spatially variant similarity term, one can replace the Euclidean distance transform by a Eikonal equation based geodesic distance transform. The Eikonal equation is of the form $|\nabla d_{k,l}(i)| = W_k(i)$, with $d_{k,l}(i)|_l = 0$ and $W_k(i) > 0$ and with ∇ representing the gradient and $|\cdot|$ representing the Euclidean norm. Physically, $d_{k,l}(i)$ is the shortest time needed to travel from the boundary of label l to i with $W_k(i)$ being the time cost at i . This distance can be solved using the fast marching method [8]. As $d_{k,l}(i)$ is the result of a first order PDE with $W_k(i) > 0$, $d_{k,l}(i)$ is guaranteed to be monotonically increasing. An example of this distance can be seen in Fig. 2.

2.3 Label Propagation and Similarity Metric

Without loss of generality, the proposed method will be used to automate the seeding and way-pointing of optic radiation tractography. Starting from a database of 40 datasets containing both T1 and DTI MRI imaging modalities, each T1 image was rigidly aligned to the fractional anisotropy (FA) of the DTI image using a block matching approach [9]. An expert human rater manually placed the seeds and waypoints on the DTI image, following the protocol defined in [10].

Assuming a new unseen image, all the multimodal datasets in the database can be propagated to this new dataset by performing an affine registration, using the same block matching approach, followed by a non-rigid alignment step. To make use of the multimodal nature of the available data for the non-rigid registration step, a multi-modal fast free-form registration algorithm [11] was used. As in Daga *et al.* [12], the multimodal data used for the registration step was comprised of the T1 image and the FA image. The resulting transformations were then used to propagate the manually placed seeds and waypoints from each

atlas to the new image under study. These manual labels were resampled using a nearest-neighbour interpolation in order to maintain their binary nature.

The similarity metric used in this work is based on the local sum of squared differences (LSSD) within a gaussian kernel with standard deviation of 5mm. This similarity term is necessary to characterise both the local differences in tissue appearance due to pathology, registration errors and morphological variability between subjects that could not be captured by the registration. The similarity is calculated independently between the propagated T1 and FA images from the atlas and the unseen T1 and FA images, respectively. This results in a FA derived metric ($LSSD_{FA}$) and an T1 derived metric ($LSSD_{T1}$). The LSSD can then be combined into a common similarity metric by using an inverse exponential function. Thus, $W_k(i)$ will be defined as $W_k(i) = e^{(-LSSD_{FA} - LSSD_{T1})}$, where higher values of $W_k(i)$ means higher similarity.

2.4 The Clinical Protocol

The seeding and waypoint placement protocol is described in [10]. In short, the seed mask consists of a series of voxels antero-lateral to the lateral geniculate nucleus at the base of Meyers loop, positioned in the white matter in one single coronal slice. These voxels should be contiguous in order to ensure that the entire coronal cross-section of Meyers loop was encompassed. The volumes of seed masks were standardised to 15 voxels for all subjects in order to reduce inter-subject tractography variability. In addition to the seed, a waypoint was placed in the lateral wall of the occipital horn of the lateral ventricle. This waypoint was also restricted to one single coronal slice.

From this clinical protocol, three main constrains are necessary: both the seeds and waypoints should be continuous and restricted to one single coronal slice and the number of voxels in the seed is standardised to $N_l = 15$. Due to the geometrical nature of the proposed method, these constrains are easily integrated into the current framework by constraining the space of solutions of $S(i)$. Because we are not interested in the background label, instead of finding what is the label that minimises the sum of the geodesic distances, one can find what is the coronal slice with a set $S(i)$ of N_l voxels that minimises the sum of the geodesic distances. As the geodesic $d_{k,l}(i)$ is not necessarily convex, contiguity of $S(i)$ is not mathematically guaranteed. However, the monotonically increasing nature of $d_{k,l}(i)$ will result in a smooth solution.

3 Validation

The validation of the proposed methodology will first quantify the error in automated placement of seeds and waypoints when compared to the manually placed ones using a leave-one-out cross validation approach. The propagated seeds and waypoints will also be characterised in terms of volume and contiguity. Secondly, because small errors in seed positioning can result in drastically different tractography results, the probabilistic overlap between manual and automatically

Table 1. Automated seed and waypoint accuracy results using MV, STAPLE, SBA and the proposed method. Column 'Manual' contains the gold standard statistics. Distance is not available for both MV and SBA because the fused output was empty.

Metric		Seed Fusion					Waypoint Fusion				
		MV	STAPLE	SBA	GSBA	Manual	MV	STAPLE	SBA	GSBA	Manual
Distance	Mean	-	1.70	-	1.60	-	-	1.86	-	1.86	-
	Std	-	1.81	-	1.41	-	-	1.12	-	0.97	-
Voxels #	Mean	0	5.8	0	15	15	0	31.89	0	113.15	136.17
	Std	0	2.89	0	0	0	0	18.85	0	17.62	85.6
Mean #	Conn.Comp.	0	1.2	0	1	1	0	1.9	0	1	1
Mean Euler		0	3.2	0	2	2	0	7.2	0	2	2

generated tracts is assessed. For all experiments, the geodesic SBA fusion performance is also compared to majority voting (MV), STAPLE [7], Euclidean SBA [1] with the original cost function for $S(i)$.

3.1 Seed and Way-Point Placement Accuracy

This section aims at assessing both accuracy of seed and waypoint placement and their geometrical properties. In order to do so, both seeds and waypoints were propagated as described in Section 2.3 and then fused using the above described methods. Due to the very small size of the seeds and waypoints and because they are limited in thickness to one single coronal slice, overlap measures like the Dice coefficient would not provide insightful information about the placement accuracy. Instead, we use the mean Euclidean distance from all the points in the automated seed and waypoints to the manually placed ones as a measure of accuracy. We also calculate the number of voxels, the number of connected components and the Euler characteristic of each propagated seed. The later describes topology of the propagated seed, where an Euler characteristic of 2 means that the seed is homotopic to a closed disk.

Results are presented in Table 1. Due to the label positioning variability, both MV and SBA methods resulted in an empty set. Note that SBA also fails because the original metric for $S(i)$ was used. Both GSBA and STAPLE obtain similar results regarding positioning accuracy, but the geometrical characteristics of the STAPLE's seeds and waypoints are discrepant when compared to the gold standard. The mean number of voxels, number of connected components and the Euler characteristic for the seeds and waypoints from STAPLE was significantly different ($p < 10^{-4}$) from the gold standard, while no significant differences were found between the GSBA and the gold standard.

3.2 Automated Tractography Validation

As a geometrically accurate positioning of the seeds and waypoints is meaningless if they lead to different tractography results, a further assessment of the similarity between manually generated and automatically generated tracks is also preformed. In order to do so, the Probabilistic Index of Connectivity (PICo)

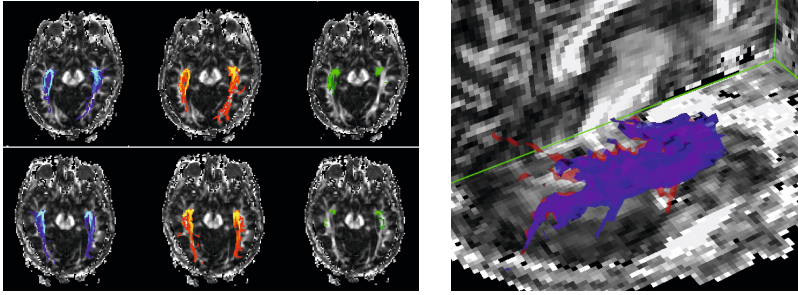


Fig. 3. Left) Optic radiation tractography for two patients. Tractography from the manual seeds, GSBA and STAPLE are in blue, red and green respectively. Right) A 3D reconstruction of the optic radiation for the manual and geodesic SBA methods.

algorithm [2] was used for tractography. Voxels in which a single tensor fitted the data poorly were identified using a spherical-harmonic voxel-classification algorithm [13]. In these voxels, a two tensor model was fitted, with the principal diffusion directions of the two diffusion tensors providing estimates of the orientations of the crossing fibres. In all other voxels, a single tensor model was fitted. Tracking from the seed was performed using 50000 Monte Carlo iterations, an angular threshold of 180 (sufficient angular flexibility to allow tracking of Meyer’s loop) and a fractional anisotropy threshold of 0.1. Because both MV and SBA method fail to generate usable seeds and waypoints, only STAPLE and the geodesic SBA were used for the comparison.

As the output from PICO is a probabilistic segmentation and because thresholding a probabilistic segmentation introduces too much boundary variability, the probabilistic Dice coefficient [14] was used as a measure of similarity between tracts. Results show that the mean(std) Dice coefficient between the manually seeded tracts and the automated tracts using the geodesic SBA is 0.547(0.138), while with STAPLE method is 0.226(0.114). Using a two-tailed paired t-test for statistical comparison, the geodesic SBA method provides highly statistically significant improvements ($p < 10^{-4}$) in tractography accuracy. Note that this is the overlap of the full tract from the Meyer’s loop to the visual cortex.

4 Conclusion

This work presents an geodesic extension the Euclidean Shape-based Averaging (SBA) framework, integrating an image similarity term in the label fusion algorithm. The geodesic SBA enables the robust propagation of sparse and non-overlapping labels and naturally permits the inclusion of geometrical and volumetric constrains into the propagated label sets. Application to DTI tractography shows statistically significant improvements in both seed and waypoint placement accuracy, geometric characteristics and topology when compared to state-of-the-art methodology, leading to improvements in tractography accuracy.

Acknowledgements. MJC, MM and SO were supported by the EPSRC (EP/H046410/1) and the CBRC (Ref. 168). GW was supported by the MRC - CRT Fellowship (G0802012). Scan acquisition was funded by a Wellcome Trust Programme Grant (083148). We are grateful to the Big Lottery Fund, Wolfson Trust and the Epilepsy Society for supporting the Epilepsy Society MRI scanner. This work was undertaken at UCLH/UCL who received a proportion of funding from the Department of Health's NIHR BRC funding scheme.

References

1. Rohlfing, T., Maurer, C.R.: Shape-based averaging. *IEEE TIP* 16(1), 153–161 (2007)
2. Parker, G.J.M., Haroon, H.A., Wheeler-Kingshott, C.A.M.: A framework for a streamline-based probabilistic index of connectivity (PICo) using a structural interpretation of MRI diffusion measurements. *J. Magn. Reson. Imaging* 18(2) (2003)
3. Bammer, R., Acar, B., Moseley, M.E.: In vivo MR tractography using diffusion imaging. *European Journal of Radiology* 45(3), 223–234 (2003)
4. Parker, G.J.M., Alexander, D.C.: Probabilistic Monte Carlo Based Mapping of Cerebral Connections Utilising Whole-Brain Crossing Fibre Information. In: Taylor, C.J., Noble, J.A. (eds.) *IPMI 2003*. LNCS, vol. 2732, pp. 684–695. Springer, Heidelberg (2003)
5. Goldberg-Zimring, D., Mewes, A.U.J., Warfield, S.K.: Diffusion Tensor Magnetic Resonance Imaging in Multiple Sclerosis. *Neuroimaging* (2005)
6. Zhang, W., Olivi, A., Hertig, S.J., van Zijl, P., Mori, S.: Automated fiber tracking of human brain white matter using diffusion imaging. *NeuroImage* 42(2) (2008)
7. Suarez, R.O., Commowick, O., Prabhu, S.P., Warfield, S.K.: Automated delineation of white matter fiber tracts with a multiple region-of-interest approach. *NeuroImage* 59(4), 3690–3700 (2012)
8. Sethian, J.A.: A fast marching level set method for monotonically advancing fronts. *PNAS* 93(4), 1591–1595 (1996)
9. Ourselin, S., Roche, A., Prima, S., Ayache, N.: Block Matching: A General Framework to Improve Robustness of Rigid Registration of Medical Images (2000)
10. Yogarajah, M., Focke, N.K., Bonelli, S., Cercignani, M., Acheson, J., Parker, G.J.M., Alexander, D.C., McEvoy, A.W., Symms, M.R., Koepp, M.J., Duncan, J.S.: Defining Meyer's loop-temporal lobe resections, visual field deficits and diffusion tensor tractography. *Brain* 132(6), 1656–1668 (2009)
11. Modat, M., Ridgway, G.R., Taylor, Z.A., Lehmann, M., Barnes, J., Hawkes, D.J., Fox, N.C., Ourselin, S.: Fast free-form deformation using graphics processing units. *CMPB* 98(3), 278–284 (2010)
12. Daga, P., Winston, G., Modat, M., White, M., Mancini, L., Cardoso, M., Symms, M., Hawkes, D., Duncan, J., Ourselin, S.: Accurate Localisation of Optic Radiation during Neurosurgery in an Interventional MRI Suite. *IEEE TMI* (December 2011)
13. Alexander, D.C., Barker, G.J., Arridge, S.R.: Detection and modeling of non-Gaussian apparent diffusion coefficient profiles in human brain data. *Magn. Reson. Med.* 48(2), 331–340 (2002)
14. Crum, W.R., Camara, O., Hill, D.L.G.: Generalized Overlap Measures for Evaluation and Validation in Medical Image Analysis. *IEEE TMI* 25(11) (2006)