
Multimodal Registration of White Matter Brain Data via Optimal Mass Transport

Tauseef ur Rehman¹, Eldad Haber², Kilian M. Pohl³, Steven Haker³, Mike Halle³, Florin Talos³, Lawrence L. Wald⁴, Ron Kikinis³ and Allen Tannenbaum³

May 31, 2008

¹Schools of Electrical & Computer and Biomedical Engg., Georgia Institute of Technology, Atlanta, GA.

²Department of Mathematics and Computer Science, Emory University, Atlanta, GA.

³Surgical Planning Laboratory, Department of Radiology, Brigham & Women's Hospital, Harvard Medical School, Boston, MA.

⁴Martinos Center, Department of Radiology, Massachusetts General Hospital, Harvard Medical School, Charlestown, MA.

tauseef, tannenba@ece.gatech.edu, haber@mathcs.emory.edu, pohl@csail.mit.edu, kikinis@bwh.harvard.edu

Abstract

The elastic registration of medical scans from different acquisition sequences is becoming an important topic for many research labs that would like to continue the post-processing of medical scans acquired via the new generation of high-field-strength scanners. In this note, we present a parameter-free registration algorithm that is well suited for this scenario as it requires no tuning to specific acquisition sequences. The algorithm encompasses a new numerical scheme for computing elastic registration maps based on the minimizing flow approach to optimal mass transport. The approach utilizes all of the gray-scale data in both images, and the optimal mapping from image A to image B is the inverse of the optimal mapping from B to A . Further, no landmarks need to be specified, and the minimizer of the distance functional involved is unique. We apply the algorithm to register the white matter folds of two different scans and use the results to parcellate the cortex of the target image. To the best of our knowledge, this is the first time that the optimal mass transport function has been applied to register large 3D multimodal data sets.

Contents

1	Introduction	2
2	Obtaining the descent direction	3
3	Deriving an efficient numerical method	4
3.1	Conservative discretization	4
3.2	Computation of a step	5
3.3	Orthogonal projection into the mass preserving constraint	5
4	Registration of Brain Data	6
5	Conclusions	8

1 Introduction

Registration is an important pre-processing step for many automatic approaches that extract cortical structures from Magnetic Resonance Images (MRI) [9, 22, 11]. Common approaches for aligning the atlas of the segmenter to the patient MRI are based on the B-spline representation [11, 19] and continuum and fluid mechanics, [7, 15, 6, 21]. The accuracy of these approaches generally depends on how well they are tuned to the sequence of patient scan. Tuning these algorithms often requires expertise about the underlying algorithm. Clinicians scanning with new acquisition sequences are therefore often concerned on how to post-process these scans. In this paper, we propose a parameter-free algorithm for the registration of MRIs.

We model the registration of images as an optimal mass transport problem. Introduced by Monge and Kantorovich [13], the solution to the problem is an optimal mapping \hat{u} (in some sense) between two densities $\mu_0 > 0$ and $\mu_1 > 0$. If we now define d as the dimension of the image domain, $\det(\cdot)$ as the determinant, u as a mapping from $\Omega \rightarrow \Omega$ with Ω a subdomain of \mathbb{R}^d , and represent by $\rho(\cdot, \cdot) : \Omega \times \Omega \rightarrow \mathbb{R}^+$ a distance function between two points in Ω , then the problem can be formalized as

$$\hat{u} \triangleq \min_{u \in \mathcal{U}} \frac{1}{2} \int_{\Omega^d} \mu_0(x) \rho(u(x), x) dx, \quad (1.1)$$

where $\mathcal{U} = \{u : \Omega \rightarrow \Omega \mid c(u) = \det(\nabla u) \mu_1(u) - \mu_0 = 0\}$.

We refer to the constraint $c(u) = 0$ as the *mass preserving* (MP) property.

For the remainder of this note, we take $\rho(\cdot, \cdot)$ to be the squared distance function $\rho(u(x), x) \triangleq \|u(x) - x\|^2$. Even for the simple L^2 -norm, (1.1) defines a highly non-linear optimization problem. While there exists a large body of literature which deals with the analysis of the problem, such as [1, 8], only a smaller number of papers discuss efficient *numerical* solutions for the problem. Benamou and Brenier [5] estimate \hat{u} by relating Equation (1.1) to the minimization of a certain kinetic energy functional with a space-time transport partial differential equation (PDE) constraint. Their approach not only estimates the optimal mapping but also provides the transportation path between the densities. A computationally faster solution to (1.1) was proposed by Angenent et al. [3]. Their algorithm directly estimates \hat{u} by first computing a transformation u_0 that fulfills the MP property. Afterwards, the algorithm improves u_0 by concatenating the mapping with the transformation,

$$\hat{s} \triangleq \min_{s \in \mathcal{S}} \frac{1}{2} \int_{\Omega^d} \mu_0(x) (u_0(s^{-1}(x)) - x)^2 dx, \quad (1.2)$$

where $\mathcal{S} = \{s : \Omega \rightarrow \Omega \mid \tilde{c}(s) = \det(\nabla s) \mu_0(s) - \mu_0 = 0\}$.

We refer to the second equation in (1.2) as the \tilde{c} constraint. This means that $s \in \mathcal{S}$ is an MP mapping from μ_0 to itself. The authors in [3] show that \hat{s} can be estimated via a steepest descent flow. To register 2D MRIs, they implement the method using forward Euler equation scheme for time stepping and a simple finite difference discretization of the spatial derivatives. The approach, however, does not enforce the MP constraint at each step of the numerical algorithm, so that the final solution generally does not fulfill the MP property. In addition, steepest descent is very slow in estimating the solution to Equation (1.2). For these reasons it would be very challenging to efficiently register 3D medical images with this approach. To overcome this hurdle, this paper describes a faster numerical solution to Equation (1.2) that enforces the MP constraint.

Unlike [3], we solve the optimization problem via an approach where we choose a direction other than steepest descent and show that it converges faster (see Section 2). Furthermore, we derive a numerical approach that uses a consistent conservative discretization method and enforces the MP constraint at each update of the solution (Section 3). In Section 4, we test the robustness of our approach by registering the white matter folds of two MRIs. The first MRI scan is part of a publicly available atlas [14] with detailed anatomical information about the scan. The second scan was acquired using a very different scanning protocol. Our approach accurately aligns the two scans. We then use the aligned atlas to outline the cortical folds in the new scan.

We end this section with the comment that our approach most closely relates to those registration approaches based on fluid mechanics. The optimal warping map of the L^2 Monge-Kantorovich equation may be regarded as the velocity vector field which minimizes a standard energy integral subject an Euler continuity equation constraint [5]. In particular, in the fluid mechanics framework, this means that the optimal Monge-Kantorovich solution is given as a *potential flow*.

2 Obtaining the descent direction

We now quickly review the derivation presented in [12] but within a variational framework. Assuming that the MP constraint condition is valid, we take a perturbation in s which stays on the MP constrained manifold. This leads to

$$\begin{aligned} 0 &= c(s + \delta s) - c(s) = \det(\nabla(s + \delta s))\mu_0(s + \delta s) - \det(\nabla s)\mu_0(s) \\ &= \det(\nabla s)(\nabla \cdot (\delta s(s^{-1}))(s))\mu_0(s) + \det(\nabla s)\nabla\mu_0(s) \cdot \delta s. \end{aligned}$$

This expression can be simplified as long as the constraint is valid. Since $\det(\nabla u) > 0$ we can divide, and rearranging we have

$$\begin{aligned} 0 &= (\mu_0 \nabla \cdot (\delta s(s^{-1}))(s) + \nabla\mu_0(s) \cdot \delta s \\ &= \mu_0 \nabla \cdot (\delta s(s^{-1})) + \nabla\mu_0 \cdot \delta s(s^{-1}) = \nabla \cdot (\mu_0 \delta s(s^{-1})). \end{aligned}$$

Defining $\delta\zeta = \mu_0 \delta s(s^{-1})$, we see that $\nabla \cdot \delta\zeta = 0$. Next, looking at $u = u_0(s^{-1})$, we can write $u(s) = u_0$ which implies that,

$$(\nabla u(s))\delta s + \delta u(s) = 0 \Rightarrow \delta u = -(\nabla u)\delta s(s^{-1}).$$

Using the definition of $\delta\zeta$ we obtain that as long as the constraint is valid and that $u(s) = u_0$ we have

$$\delta u = -\mu_0^{-1}(\nabla u)\delta\zeta, \quad (2.3a)$$

$$0 = \nabla \cdot \delta\zeta. \quad (2.3b)$$

Let M be defined as the objective function in (1.2) then it can be shown that

$$\delta M = \int_{\Omega} u \cdot \delta\zeta \, dx. \quad (2.4)$$

In the the original paper [12], it is suggested to use the Helmholtz decomposition in order to obtain a descent direction. Here we employ a different approach. First, we note that the divergence constraint can be eliminated by selecting $\delta\zeta = \nabla \times \delta\eta$, and thus to reduce M we need to obtain a direction that yields a negative δM , that is we seek a direction, $\delta\eta$ such that $\delta M = \int_{\Omega} u \cdot \nabla \times \delta\eta \, dx < 0$. Using the Gauss theorem, we obtain that

$$\int_{\Omega} u \cdot \nabla \times \delta\eta \, dx = \int_{\Omega} \nabla \times u \cdot \delta\eta \, dx + \int_{\partial\Omega} (u \cdot (\delta\eta \times \vec{n})) \, d\vec{n}$$

and therefore the steepest descent direction is given by

$$\delta\eta = \nabla \times u, \quad \delta\eta \in \Omega; \quad \delta\eta \times \vec{n} = 0, \quad \delta\eta \in \partial\Omega$$

which leads to the update $\delta\zeta = \nabla \times \nabla \times u$, and finally to the steepest descent direction in u $\delta u = -\frac{1}{\mu_0}(\nabla u)\nabla \times \nabla \times u$ or, in symmetric form

$$\mu_0(\nabla u)^{-1}\delta u = -\nabla \times \nabla \times u. \quad (2.5)$$

The operator $-\nabla \times \nabla \times$ is negative and elliptic thus, the equation can be thought of as a parabolic PDE as long as real part of the eigenvalues of ∇u are positive. Using the above decomposition a family of different directions may be obtained. Note that in order to reduce the objective $\int_{\Omega} \nabla \times u \cdot \eta dx$ any vector field of the form $\delta\eta = A \nabla \times u$ can be used where A is a symmetric positive definite matrix. For example, a choice that leads to a similar method to the one derived in the original paper [12] in 2D is $A = -\Delta^{-1}$, which leads to the update

$$\mu_0(\nabla u)^{-1}\delta u = \nabla \times \Delta^{-1}\nabla \times u. \quad (2.6)$$

Using the above calculation it is easy to see that the flow (2.6) is valid also in 3D. Moreover, it is easy to verify that given a smooth u the second formulation (2.6) leads to a more stable method that should converge faster compared with the first formulation (2.5), because the operator $\nabla \times \Delta^{-1}\nabla \times$ is compact while the $\nabla \times \nabla \times$ operator is unbounded. In this work, we therefore derive a numerical method for (2.6) rather than for (2.5).

3 Deriving an efficient numerical method

In this section, we derive an efficient numerical method for the solution of the flow. The proposed method has three main components: a conservative discretization of differential operators, a criterion to choose step size, and a method to correct steps that deviate from the mass preservation constraint.

3.1 Conservative discretization

The applications we have in mind derive from medical imaging where images are discretized on a regular grid. We therefore construct our discretization based on a finite volume/difference approach. To derive and analyze our discretization we introduce a new variable $\delta p = \Delta^{-1}\nabla \times u$ and rewrite (2.6) as

$$\begin{pmatrix} \mu_0(\nabla u)^{-1} & \nabla \times \\ 0 & \Delta \end{pmatrix} \begin{pmatrix} \delta u \\ \delta p \end{pmatrix} = \begin{pmatrix} 0 \\ \nabla \times u \end{pmatrix}. \quad (3.7)$$

In order for the discrete system to be well posed we need consistent discretizations for Δ , ∇u and $\nabla \times u$. There are a number of possible discretizations that lead to a well-posed system.

We divide Ω into $n_1 \times \dots \times n_d$ cells, each of size $h_1 \times \dots \times h_d$ where d is the dimension of the problem. We discretize all the components of u at the nodes of each cell to obtain d grid functions $\hat{u}^1, \dots, \hat{u}^d$. Since δp is connected to u by the curl operator, we employ a staggered grid and place δp at cell centers. To approximate ∇u at each node, we use long differences. Thus, in 3D, the discretized (1,1) block in (3.7) is a matrix of the form

$$(\nabla_h \hat{u}) = \frac{1}{h} \begin{pmatrix} \text{diag}(D_1 \hat{u}^1) & \text{diag}(D_2 \hat{u}^1) & \text{diag}(D_3 \hat{u}^1) \\ \text{diag}(D_1 \hat{u}^2) & \text{diag}(D_2 \hat{u}^2) & \text{diag}(D_3 \hat{u}^2) \\ \text{diag}(D_1 \hat{u}^3) & \text{diag}(D_2 \hat{u}^3) & \text{diag}(D_3 \hat{u}^3) \end{pmatrix}, \quad (3.8)$$

where D_j is a matrix of long differences in the j^{th} direction. To obtain a consistent discretization of the Laplacian we use a standard discretization (5 point stencil in 2D and 7 point stencil in 3D) with Dirichlet boundary conditions. Finally, we employ short differences in one direction averaged in the other direction to obtain a cell centered approximation of $\nabla \times u$.

3.2 Computation of a step

The computation of each step requires two parts. Firstly, the solution of (3.7) and secondly, a way to determine if it is an acceptable step. The solution of the system (3.7) is straightforward. Any fast Poisson solver can be used for the task. Here we have used a standard multigrid method with weighted Jacobi smoothing, bilinear prolongation and its adjoint as a restriction.

The validity of the update is determined using the following procedure. Assume that at iteration n we have \hat{u}_n as an approximation to u and that we computed $\delta\hat{u}$. The update is then performed using,

$$\hat{u}_{n+1} = \mathcal{P}(\hat{u}_n + \alpha\delta\hat{u}), \quad (3.9)$$

where \mathcal{P} is an orthogonal projection discussed in Section 3.3 below, that projects $\hat{u}_n + \alpha\delta\hat{u}$ into the mass preserving manifold. The step size α is then chosen such that the objective function is decreased and that the real part of the eigenvalues of $(\nabla_h \hat{u})$ is positive. The whole algorithm is outlined Algorithm 1.

Algorithm 1 Solution of OMT:

$\hat{u} \leftarrow \text{OMTsol}(\mu_0, \mu_1);$

Use μ_0 and μ_1 to compute a mass preserving u_0
while true **do**
 Solve (3.7) for $\delta\hat{u}$
 line search: set $\alpha = 1$
 while true **do**
 $\hat{u}_{n+1} = \mathcal{P}(\hat{u}_n + \alpha\delta\hat{u})$
 if $\|\hat{u}_{n+1} - x\|_{\mu_0} < \|\hat{u}_n - x\|_{\mu_0}$ and $\text{Re}(\lambda(\nabla_h \hat{u}_{n+1})) > 0$ **then**
 Break
 end if
 $\alpha \leftarrow \alpha/2$
 end while
end while

3.3 Orthogonal projection into the mass preserving constraint

Assume that we have computed a mass preserving mapping \hat{u}_n , and that we have updated it to obtain $v_n = \hat{u}_n + \alpha\delta\hat{u}$. It should be noted that an infinitesimal $\delta\hat{u}$ does not guarantee mass preservation. Furthermore, we aim to take large steps in $\delta\hat{u}$, and therefore the MP constraint is likely to be invalid. To correct for this we use orthogonal projections. The goal is to compute a vector field δv such that $c(v + \delta v) = 0$. Obviously, δv is non-unique and therefore we seek a minimum norm solution that is we seek δv such that

$$\min_v \frac{1}{2} \|\delta v\|_{\mu_0}^2 \quad \text{subject to} \quad c(\delta v) = \mu_0(v + \delta v) \det(\nabla(v + \delta v)) - \mu_1 = 0.$$

It is easy to verify that a correction for δv can be obtained by solving the system $\delta v \approx c_v^\top (c_v c_v^\top)^{-1} c(v)$. The system $c_v c_v^\top$ can be thought as an elliptic system of equations. The system is solved using preconditioned conjugate gradient with an incomplete Cholesky preconditioner.

4 Registration of Brain Data

Our goal is the identification of cortical structures by mapping a publicly available atlas[14] to the scan of a patient. In our scenario, the scanning sequence of the atlas is very different from the one of the patient. The MRI of the atlas is a spoiled gradient recalled image acquired on a 1.5-Tesla General Electric Signa System (GE Medical Systems, Milwaukee) with $256 \times 256 \times 124$ voxels and voxel dimension of $0.92 \times 0.92 \times 1.5$ mm. The patient scan is a MPRAGE acquired on a Siemens 3T long bore machine using a 8 channel head coil. The resolution of the scan is $256 \times 256 \times 144$ with voxel dimension $0.54 \times 0.54 \times 1.0$ mm (See Figure 1(b)).

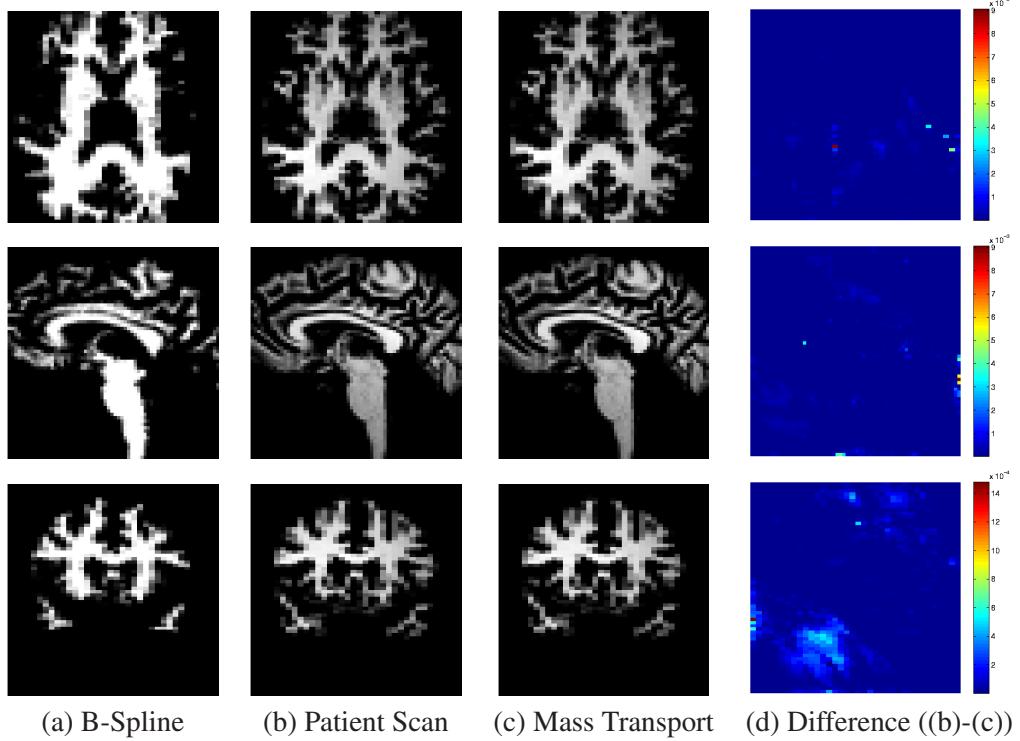


Figure 1: Registration results.

The parcellation of the cortex can be encoded by partitioning the boundary between cortex and white matter into anatomical regions [9]. The label map of cortical structures can then be inferred from this partition by propagating the labeling along the boundary to the entire cortex. The pipeline described below will apply this concept for the parcellation of the cortex to the high resolution scan.

The input of the pipeline consist of the atlas, the high resolution scan as well as a segmentation of the scan into the major tissue classes. In the first step, we coarsely align the atlas to the image data using the B-spline implementation by Rohlfinger[19] with a final spacing of the grid nodes of 2.5 mm. This results in a coarse alignment of the scans. The algorithm has difficulties in mapping the folds of the white matter due to the inherent constraints of the B-spline representation. We then reduce the atlas to the white matter including the parcellation of the cortex along the boundary between gray and white matter (see Figure 1(a)). Afterwards, we refine the alignment of this new atlas to the white matter of the high resolution scan using our Optimal Mass Transport registration approach. Registration using Optimal Mass Transport is a highly flexible approach that is, unlike B-Splines, not constrained to a set of control points. The intensities in the two input datasets are first normalized and rescaled to make sure that both have the same total mass. The

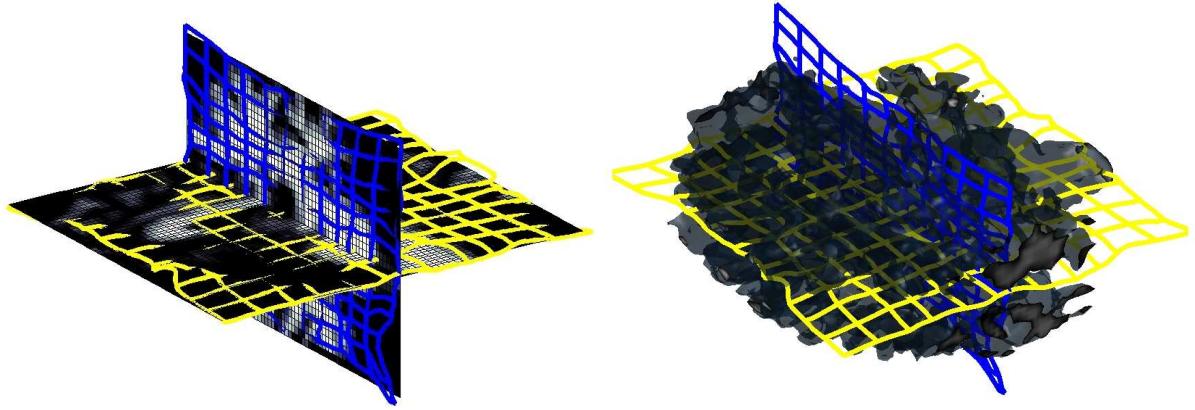


Figure 2: Deformed Grid on white matter Slices (left) and 3D volume (right).

white matter registration with the proposed algorithm took just 12 iterations to converge with 2 iterations of the projection to constraint per iteration. This is a huge improvement over algorithm proposed in [12] where thousands of iterations were required for convergence with roughly the same computational complexity per iteration. The $\nabla \times u$ (convergence metric) was reduced to an order of 10^{-3} indicating an optimal map. Figure 1(c) shows the resampled images with 3D views of the corresponding deformation grid in Figure 2. The difference (Figure 1(d)) between target (Figure 1(b)) and resampled image indicates that our approach accurately aligned the folds. After this local alignment, the folds of the atlas should perfectly align with

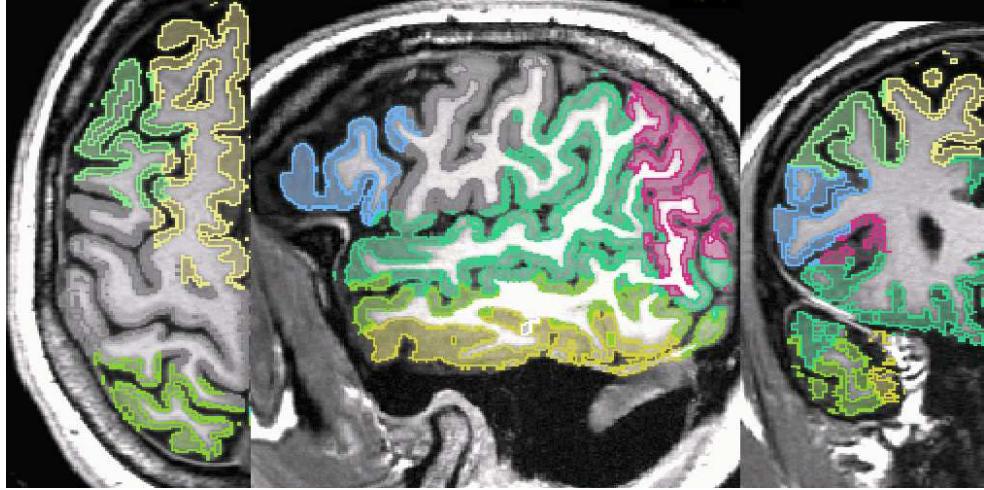


Figure 3: Parcellation results

the ones of the high resolution scan. The parcellation of the folds of the atlas, therefore, also encodes the parcellation of the same region in the high resolution scan. We then complete the cortex parcellation of the high resolution scan by confining the Voronoi diagram of the aligned atlas to the gray matter mask of the high resolution scan. The results in Figure 3 show the corresponding segmentation when applying the deformation map of the B-Spline registration and our approach to the the label map of [14], and propagating the labels to the cortex via the Voronoi diagram.

We are aware of the variety of other methods for registering and segmenting cortical structures. We also note that our segmentation results are by no means perfect. However, to the best of our knowledge, this

is the first time in medical imaging that a parameter-free registration tool has been used for registering the cortical folds of 3D MRIs.

5 Conclusions

The difficulties of aligning cortical folds is reflected by the large body of literature discussing this topic. Registration approaches based on continuum and fluid mechanics are often applied to this problem. However, the accuracy of these approaches generally depends on how well they are tuned to the sequence of patient scan. We view Optimal Mass Transport (OMT) as part of these types of registration approaches. Unlike the current state of the art, OMT is parameter free. It is, therefore, especially suited to align new acquisition sequences, which the other methods have not yet been tuned to.

In this paper we presented an efficient variational methodology for the computation of the optimal L^2 mass transport mapping based on the formulation of [12]. Although, the theory was rigorous in [12], the proposed numerics were problematic. All of these problems have been addressed in our approach. This has lead to an efficient robust elastic deformation algorithm which is guaranteed to converge to the optimal solution of the Monge-Kantorovich problem. We applied the approach to register the white matter between two MRI datasets. We then use the results to resample the label map of the source providing us with a parcellation of the cortex of the target image. We note that the approach is applicable to a whole range of registration and image morphing problems where the mass preservation constraint makes sense. Based on deriving this numerical framework, we are quite sure that in the near future we will be able to provide cases in which we show superior performance to other well established tools in the community. Finally, the set-up can be extended directly to optimal transport on a manifold as in [4].

Acknowledgements

This work was supported in part by grants from NSF, AFOSR, ARO, MURI, as well as by a grant from NIH (NAC P41 RR-13218) through Brigham and Women’s Hospital. This work is part of the National Alliance for Medical Image Computing (NAMIC), funded by the National Institutes of Health through the NIH Roadmap for Medical Research, Grant U54 EB005149. Information on the National Centers for Biomedical Computing can be obtained from <http://nihroadmap.nih.gov/bioinformatics>. Tannenbaum was also supported by a Marie Curie Grant through the Technion, Israel Institute of Technology.

References

- [1] L. Ambrosio, “Lecture Notes on Optimal Transport Problems,” Lectures given at Euro Summer School, July 2000. Available on <http://cvgmt.sns.it/papers/amb00a/>. [1](#)
- [2] S. Angenent, S. Haker, R. Kikinis, A. Tannenbaum, “Laplace-Beltrami operator and brain surface flattening,” *IEEE Trans. on Medical Imaging*, 18:700-711, 1999.
- [3] S. Angenent, S. Haker, and A. Tannenbaum, “Minimizing flows for the Monge-Kantorovich problem,” *SIAM J. Math. Anal.*, 35:61–97, 2003. [1](#), [1](#)
- [4] A. Dominitz, S. Angenent, and A. Tannenbaum, “On the computation of optimal transport maps using gradient flows and multiresolution analysis,” in *Recent Advances in Learning and Control* edited by V. Blondel, S. Boyd, and H. Kimura, Springer-Verlag, New York, 2008. [5](#)

[5] J. D. Benamou and Y. Brenier, “A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem,” *SIAM J. Math. Analy.*, 35:61–97, 2003. [1](#), [1](#)

[6] M. Bro-Nielsen and C. Gramkow, “Fast fluid registration of medical images,” in *Visualization in Biomedical Imaging* edited by K. Höhne and R. Kikinis, Lecture Notes in Computer Science **1131**, pp. 267–276, Springer-Verlag, New York, 1996. [1](#)

[7] G. E. Christensen, R. D. Rabbit, M. Miller, “Deformable templates using large deformation kinematics,” *IEEE Trans. on Image Processing* **5** (1996), pp. 1435–1447. [1](#)

[8] L. C. Evans, *Partial Differential Equations and Monge-Kantorovich Transfer*. Lecture Notes, 1989. [1](#)

[9] B. Fischl *et al.*, “Automatically Parcellating the Human Cerebral Cortex,” *Cerebral Cortex*, 14(11):11–22, 2004. [1](#), [4](#)

[10] A. Guimond, A. Roche, N. Ayache, and J. Meunier, “Three-dimensional multimodal brain warping using the demons algorithm and adaptive intensity corrections,” *IEEE TMI*, 20(1):58–69, 2001.

[11] R. A. Heckemann, J. Hajnal, P. Aljabar, D. Rueckert, and A. Hammers, “Automatic anatomical brain MRI segmentation combining label propagation and decision fusion,” *NeuroImage*, 33(1):115–126, 2006. [1](#)

[12] S. Haker, L. Zhu, A. Tannenbaum, and S. Angenent, “Optimal mass transport for registration and warping,” *Int. Jour. Computer Vision*, 60(3):225–240, 2004. [2](#), [2](#), [2](#), [4](#), [5](#)

[13] L. Kantorovich, “On a problem of Monge,” *Uspekhi. Mat. Nauk.*, 3:225–226, 1948. [1](#)

[14] R. Kikinis *et al.*, “A digital brain atlas for surgical planning, model-drivensegmentation, and teaching,” *IEEE Trans. on Vis. and Comp. Graph.*, 2(3):232–241, 1996. [1](#), [4](#), [4](#)

[15] M. Miller, G. Christensen, Y. Amit, U. Grenander, “Mathematical textbook of deformable neuroanatomies,” *Proc. Nat. Acad. of Science* **90**, pp. 11944–11948. [1](#)

[16] K. Pohl, S. Bouix, R. Kikinis, and W. Grimson, “Anatomical guided segmentation with non-stationary tissue class distributions in an expectation-maximization framework,” *IEEE ISBI*, 81–84, 2004.

[17] J. Nocedal and S. Wright, *Numerical Optimization*, Springer, New York, 1999.

[18] S. Rachev and L. Rüschendorf, *Mass Transportation Problems*, Volumes I and II, Probability and Its Applications, Springer, New York, 1998.

[19] T. Rohlfing and C. R. Maurer, Jr., “Nonrigid image registration in shared-memory multiprocessor environments with application to brains, breasts, and bees,” *IEEE Trans. on Inf. Techn. in Biomed.*, vol. 7, no. 1, pp. 16–25, 2003. [1](#), [4](#)

[20] Y. Rubner, C. Tomasi, and J. Guibas, “The earth mover’s distance as a metric for image retrieval,” Technical Report STAN-CS-TN-98-86, Department of Computer Science, Stanford University, September 1998.

[21] J.-P. Thirion, “Fast non-rigid matching of non-rigid images,” *INRIA Technical Report 2547*, Project Epidaure, INRIA, France, 1995. [1](#)

[22] A. Toga, *Brain Warping*, Academic Press, San Diego, 1999. [1](#)