

# Mass-Preserving Maps for Registration and Visual Tracking

Steven Haker

Department of Radiology, Surgical Planning Laboratory  
Brigham and Women's Hospital  
Boston, MA 02115; haker@bwh.harvard.edu

Allen Tannenbaum

Department of Electrical and Computer Engineering  
Georgia Tech  
Atlanta, GA 30332-0250; tannenba@ece.gatech.edu

## Abstract

In this note, we consider a new method for an important aspect of the visual tracking problem. Tracking in the presence of a disturbance is a classical control issue, but because of the highly uncertain nature of the disturbance, this type of problem is very difficult. A key issue in many visual tracking tasks is that of *registration*. Image registration is the process of establishing a common geometric reference frame among several data sets taken at different times. In this note, we propose a method of registration based on the Monge-Kantorovich problem of optimal mass transport. We argue that such an approach can also be very useful for several problems in controlled active vision.

**Key words:** Visual tracking, image registration, mass-preserving maps, optimal transport, gradient descent.

## 1 Introduction

Image tracking is one of the key tasks in any visual control system. The method of registration, i.e., establishing a common geometric reference frame between two or more data sets from the same or different times, is essential in such a task. In this paper, we propose a new method based on the theory of mass-preserving maps and the optimal mass transport problem. This problem was first formulated by Monge in 1781, and concerned finding the optimal way, in the sense of minimal transportation cost, of moving a pile of soil from one site to another. The problem was given a modern formulation in the work of Kantorovich [12], and so is now known as the *Monge-Kantorovich problem*. We should note that very recently, optimal transport has been used by Anders Rantzer [18] to study nonlinear system stability.

The registration problem is one of the great challenges that must be addressed in order to make image-guided tracking a practical reality. Registration has a huge literature devoted to it with numerous approaches ranging from statistical to computational fluid dynamics to various types of warping methodologies. See [20] for a number of recent papers on the subject as well as an extensive set of references.

We should note that multimodal registration methods play a central role in image-guided systems. First, they allow for the fusing of information from each imaging modality, providing better and more accurate information than can be obtained from each image viewed separately. Second, they allow quantitative comparison of images taken at different times, from which information about evolution over time can be inferred. Finally, they allow for the updating of a precomputed image or model using real-time tracking data.

Multimodal registration proceeds in several steps. First, each image or data set to be matched should be individually calibrated, corrected for imaging distortions and artifacts, and cleared of noise. Next, a measure of similarity between the data sets must be established, so that one can quantify how close an image is from another after transformations are applied. Next, the transformation that maximizes the similarity between the transformed images is found. Often this transformation is given as the solution of an optimization problem where the transformations to be considered are constrained to be of a predetermined class. Finally, once an optimal transformation is obtained, it is used to fuse the image data sets.

The technique we propose in this paper is designed for elastic registration, and is based on an optimization problem built around the  $L^2$  Kantorovich-Wasserstein distance taken as the similarity measure. The constraint that we will put on the transformations con-

sidered is that they obey a mass preservation property. Thus, we will be matching *mass densities* in this method, which may be thought of as weighted areas in 2D or weighted volumes in 3D. We will assume that a rigid (non-elastic) registration process has already been applied before applying our scheme.

## 2 Optimal Transport Problem

We now give a modern formulation of the Monge-Kantorovich problem. Let  $\Omega_0$  and  $\Omega_1$  be two subdomains of  $\mathbf{R}^d$ , with smooth boundaries, each with a positive density function,  $\mu_0$  and  $\mu_1$ , respectively. We assume

$$\int_{\Omega_0} \mu_0 = \int_{\Omega_1} \mu_1$$

so that the same total mass is associated with  $\Omega_0$  and  $\Omega_1$ . We consider diffeomorphisms  $\tilde{u}$  from  $(\Omega_0, \mu_0)$  to  $(\Omega_1, \mu_1)$  which map one density to the other in the sense that

$$\mu_0 = |D\tilde{u}| \mu_1 \circ \tilde{u}, \quad (1)$$

which we will call the *mass preservation* (MP) property, and write  $\tilde{u} \in MP$ . Equation (1) is called the *Jacobian equation*. Here  $|D\tilde{u}|$  denotes the determinant of the Jacobian map  $D\tilde{u}$ . In particular, Equation (1) implies, for example, that if a small region in  $\Omega_0$  is mapped to a larger region in  $\Omega_1$ , then there must be a corresponding decrease in density in order for the mass to be preserved. A mapping  $\tilde{u}$  that satisfies this property may thus be thought of as defining a redistribution of a mass of material from one distribution  $\mu_0$  to another distribution  $\mu_1$ .

There may be many such mappings, and we want to pick out an optimal one in some sense. Accordingly, we define the  $L^p$  Kantorovich-Wasserstein metric as follows:

$$d_p(\mu_0, \mu_1)^p := \inf_{\tilde{u} \in MP} \int \|\tilde{u}(x) - x\|^p \mu_0(x) dx. \quad (2)$$

An *optimal MP map*, when it exists, is one which minimizes this integral. This functional is seen to place a penalty on the distance the map  $\tilde{u}$  moves each bit of material, weighted by the material's mass.

The case  $p = 2$  has been extensively studied and will be the one proposed in this paper for registration. The  $L^2$  Monge-Kantorovich problem has been studied in statistics, functional analysis, and the atmospheric sciences; see [5, 3] and the references therein. A fundamental theoretical result [13, 4, 9], is that there is a unique optimal  $\tilde{u} \in MP$  transporting  $\mu_0$  to  $\mu_1$ , and that this  $\tilde{u}$  is characterized as the gradient of a convex function  $w$ , i.e.,  $\tilde{u} = \nabla w$ . Note that from Equation (1), we have that  $w$  satisfies the *Monge-Ampère* equation

$$|Hw| \mu_1 \circ (\nabla w) = \mu_0,$$

where  $|Hw|$  denotes the determinant of the Hessian  $Hw$  of  $w$ .

Hence, the Kantorovich-Wasserstein metric defines the distance between two mass densities, by computing the cheapest way to transport the mass from one domain to the other with respect to the functional given in (2), the optimal transport map in the  $p = 2$  case being the gradient of a certain function. The novelty of this result is that like the Riemann mapping theorem in the plane, the procedure singles out a particular map with preferred geometry.

## 3 Algorithms for Computing The Transport Map

There have been a number of algorithms considered for computing an optimal transport map. For example, methods have been proposed based on linear programming [17], and on Lagrangian mechanics closely related to ideas from the study of fluid dynamics [3]. An interesting geometric method has been formulated by Cullen and Purser [5].

In this section, we will employ a natural solution based on the equivalent problem of *polar factorization*; see [4, 8, 14] and the references therein. We will work with the general case of subdomains in  $\mathbf{R}^d$ , and point out some simplifications that are possible for the  $\mathbf{R}^2$  case.

As above, let  $\Omega_0, \Omega_1 \subset \mathbf{R}^d$  be subdomains with smooth boundaries, with corresponding positive density functions  $\mu_0$  and  $\mu_1$  satisfying  $\int_{\Omega_0} \mu_0 = \int_{\Omega_1} \mu_1$ . Let  $u : (\Omega_0, \mu_0) \rightarrow (\Omega_1, \mu_1)$  be an initial mapping with the mass preserving (MP) property. Then according to the generalized results of [4, 8], one can write

$$u = (\nabla w) \circ s, \quad (3)$$

where  $w$  is a convex function and  $s$  is an MP mapping  $s : (\Omega_0, \mu_0) \rightarrow (\Omega_0, \mu_0)$ . This is the *polar factorization* of  $u$  with respect to  $\mu_0$ . In [8], just the case of area preservation is considered, i.e.,  $\mu_0$  is assumed constant, but the general case goes through as well.

Our goal is to find the polar factorization of the MP mapping  $u$ , according to the following strategy. We consider the family of MP mappings of the form  $\tilde{u} = u \circ s^{-1}$  as  $s$  varies over MP mappings from  $(\Omega_0, \mu_0)$  to itself. If we consider  $\tilde{u}$  as a vector field, we can always find a function  $w$  and another vector field  $\chi$ , with  $\text{div}(\chi) = 0$ , such that

$$\tilde{u} = \nabla w + \chi,$$

i.e., we can decompose  $\tilde{u}$  into the sum of a curl-free and divergence-free vector field [19]. Thus, what we try to do is find a mapping  $s$  which will yield a  $\tilde{u}$  without any

curl, that is, such that  $\tilde{u} = \nabla w$ . Once such an  $s$  is found, we will have  $u = \tilde{u} \circ s = (\nabla w) \circ s$  and so we will have found the polar factorization (3) of our given function  $u$ .

Now, here is the key point. As we discussed above, the unique optimal solution of the  $L^2$  Monge–Kantorovich problem has the form  $\tilde{u} = \nabla w$ , and so the problem of finding the polar factorization of  $u$  and finding the optimal Monge–Kantorovich mapping  $\tilde{u}$  are equivalent. In essence, to solve the Monge–Kantorovich problem we create a “rearrangement” of an initial vector field  $u$  using a map  $s$ , so that the resulting vector field  $\tilde{u} = u \circ s^{-1}$  has no curl. We can now give the technical details.

### 3.1 Initial Mapping

We will now propose an explicit algorithm to solve the Monge–Kantorovich problem. So we want to minimize the  $L^2$  Kantorovich–Wasserstein distance functional over MP functions from  $(\Omega_0, \mu_0)$  to  $(\Omega_1, \mu_1)$ . We will try to do this by finding an initial MP mapping  $u$  and then minimizing over  $\tilde{u} = u \circ s^{-1}$  by varying  $s$  over MP mappings from  $\Omega_0$  to  $\Omega_0$ , starting with  $s$  equal to the identity map. Our first task is to find an initial MP mapping  $u$ . This can be done for general domains using a method of Moser [15, 6], or for simpler domains using the following algorithm. For simplicity, we work in  $\mathbf{R}^2$  and assume  $\Omega_0 = \Omega_1 = [0, 1]^2$ , the generalization to higher dimensions being straightforward. We define a function  $a = a(x)$  by the equation

$$\int_0^{a(x)} \int_0^1 \mu_1(\eta, y) dy d\eta = \int_0^x \int_0^1 \mu_0(\eta, y) dy d\eta \quad (4)$$

which gives by differentiation with respect to  $x$

$$a'(x) \int_0^1 \mu_1(a(x), y) dy = \int_0^1 \mu_0(x, y) dy. \quad (5)$$

We may now define a function  $b = b(x, y)$  by the equation

$$a'(x) \int_0^{b(x, y)} \mu_1(a(x), \rho) d\rho = \int_0^y \mu_0(x, \rho) d\rho, \quad (6)$$

and set  $u(x, y) = (a(x), b(x, y))$ . Since  $a_y = 0$ ,  $|Du| = a_x b_y$ , and differentiating (6) with respect to  $y$  we find

$$\begin{aligned} a'(x) b_y(x, y) \mu_1(a(x), b(x, y)) &= \mu_0(x, y) \\ |Du| \mu_1 \circ u &= \mu_0, \end{aligned}$$

which is the MP property we need. This process can be interpreted as the solution of a one-dimensional Monge–Kantorovich problem in the  $x$  direction followed by the solution of a family of one-dimensional Monge–Kantorovich problems in the  $y$  direction.

### 3.2 Gradient Descent

Once an initial MP  $u$  is found, we need to apply the process which will remove its curl. We begin with the following elementary and intuitive property of MP mappings, the proof of which is a simple calculation.

*Lemma:* The composition of two mass preserving (MP) mappings is an MP mapping. The inverse of an MP mapping is an MP mapping.

Thus, since  $u$  is an MP mapping, we have that  $\tilde{u} = u \circ s^{-1}$  is an MP mapping if and only if  $s$  is, that is, if and only if

$$\mu_0 = |Ds| \mu_0 \circ s.$$

In particular, when  $\mu_0$  is constant, this equation requires that  $s$  be area or volume preserving.

Next, rather than working with  $s$  directly, we solve the polar factorization problem via gradient descent. Accordingly, we will assume that  $s$  is a function of time, and then determine what  $s_t$  should be to decrease the  $L^2$  Monge–Kantorovich functional. This will give us an evolution equation for  $s$  and in turn an equation for  $\tilde{u}_t$  as well, the latter being the most important for implementation. By differentiating  $\tilde{u} \circ s = u$  with respect to time, we get

$$\tilde{u}_t = -D\tilde{u} \tilde{s}_t$$

where we’ve abused notation to define  $\tilde{s}_t := s_t \circ s^{-1}$ . We need to make sure that  $s$  maintains its MP property. Differentiating  $\mu_0 = |Ds| \mu_0 \circ s$  with respect to time, we derive

$$\operatorname{div}(\mu_0 \tilde{s}_t) = 0,$$

from which we see that  $\tilde{s}_t$ ,  $s_t$  and  $\tilde{u}_t$  should have the following forms:

$$\tilde{s}_t = \frac{1}{\mu_0} \zeta, \quad (7)$$

$$s_t = \left( \frac{1}{\mu_0} \zeta \right) \circ s, \quad (8)$$

$$\tilde{u}_t = -\frac{1}{\mu_0} D\tilde{u} \zeta, \quad (9)$$

for some vector field  $\zeta$  on  $\Omega_0$ , with  $\operatorname{div}(\zeta) = 0$  and  $\langle \zeta, \vec{n} \rangle = 0$  on  $\partial\Omega_0$ ,  $\vec{n}$  being the normal to the boundary of  $\Omega_0$ . This last condition ensures that  $s$  remains a mapping from  $\Omega_0$  to itself, by preventing the flow of  $s$ , given by  $s_t = \left( \frac{1}{\mu_0} \zeta \right) \circ s$ , from crossing the boundary of  $\Omega_0$ . This also means that the range of  $\tilde{u} = u \circ s^{-1}$  is always  $u(\Omega_0) = \Omega_1$ .

Consider now the problem of minimizing the Monge–Kantorovich functional:

$$M = \int \|\tilde{u} - x\|^2 \mu_0 \quad (10)$$

$$= \int \|\tilde{u}\|^2 \mu_0 - 2 \int \langle \tilde{u}, x \rangle \mu_0 + \int \|x\|^2 \mu_0 \quad (11)$$

The last term is obviously independent of time. Interestingly, so is the first,

$$\begin{aligned} \int \|\tilde{u}\|^2 \mu_0 &= \int \|u \circ s^{-1}\|^2 \mu_0 \\ &= \int \|u \circ s^{-1}\|^2 |Ds^{-1}| \mu_0 \circ s^{-1} \\ &= \int \|u\|^2 \mu_0 \end{aligned}$$

where  $\mu_0 = |Ds^{-1}| \mu_0 \circ s^{-1}$  since  $s^{-1}$  is an MP map.

Turning now to the middle term, we do a similar trick,

$$\begin{aligned} \int \langle \tilde{u}, x \rangle \mu_0 &= \int \langle u \circ s^{-1}, s \circ s^{-1} \rangle \mu_0 \\ &= \int \langle u \circ s^{-1}, s \circ s^{-1} \rangle |Ds^{-1}| \mu_0 \circ s^{-1} \\ &= \int \langle u, s \rangle \mu_0, \end{aligned}$$

and taking  $s_t = \left(\frac{1}{\mu_0} \zeta\right) \circ s$ , we compute

$$-\frac{1}{2} M_t = \int \langle u, s_t \rangle \mu_0 = \int \langle \tilde{u}, \zeta \rangle.$$

Now decomposing  $\tilde{u}$  as  $\tilde{u} = \nabla w + \chi$ , we have

$$-\frac{1}{2} M_t = \int \langle \nabla w + \chi, \zeta \rangle \quad (12)$$

$$= \int \langle \nabla w, \zeta \rangle + \int \langle \chi, \zeta \rangle \quad (13)$$

$$= \int (\operatorname{div}(w\zeta) - w \operatorname{div}(\zeta)) + \int \langle \chi, \zeta \rangle \quad (14)$$

$$= \int_{\partial\Omega_0} w \langle \zeta, n \rangle + \int \langle \chi, \zeta \rangle \quad (15)$$

$$= \int \langle \chi, \zeta \rangle. \quad (16)$$

Thus, in order to decrease  $M$ , we can take  $\zeta = \chi$  with corresponding formulas (7)-(9) for  $s_t$ ,  $\tilde{s}_t$ , and  $\tilde{u}_t$ , provided that we have  $\operatorname{div}(\chi) = 0$  and  $\langle \chi, \vec{n} \rangle = 0$  on  $\partial\Omega_0$ . Thus it remains to show that we can decompose  $\tilde{u}$  as  $\tilde{u} = \nabla w + \chi$  for such a  $\chi$ .

**Gradient Descent:  $\mathbf{R}^d$ :**

We let  $w$  be a solution of the Neumann-type boundary problem

$$\operatorname{div}(\tilde{u}) = \Delta w \quad (17)$$

$$\langle \nabla w, \vec{n} \rangle = \langle \tilde{u}, \vec{n} \rangle \text{ on } \partial\Omega_0, \quad (18)$$

and set  $\chi = \tilde{u} - \nabla w$ . It is then easily seen that  $\chi$  satisfies the necessary requirements.

Thus, by (9), we have the following evolution equation for  $\tilde{u}$ :

$$\tilde{u}_t = -\frac{1}{\mu_0} D\tilde{u} (\tilde{u} - \nabla \Delta^{-1} \operatorname{div}(\tilde{u})). \quad (19)$$

This is a first order non-local scheme for  $\tilde{u}_t$  if we count  $\Delta^{-1}$  as minus 2 derivatives. Note that this flow is consistent with respect to the Monge-Kantorovich theory in the following sense. If  $\tilde{u}$  is optimal, then it is given as  $\tilde{u} = \nabla w$ , in which case  $\tilde{u} - \nabla \Delta^{-1} \operatorname{div}(\tilde{u}) = \nabla w - \nabla \Delta^{-1} \operatorname{div}(\nabla w) = 0$  so that by (19),  $\tilde{u}_t = 0$ .

**Gradient Descent:  $\mathbf{R}^2$ :**

The situation is somewhat simpler in the  $\mathbf{R}^2$  case, due to the fact that a divergence free vector field  $\chi$  can in general be written as  $\chi = \nabla^\perp h$  for some scalar function  $h$ , where  $\perp$  represents rotation by 90 deg, so that  $\nabla^\perp h = (-h_y, h_x)$ . In this case, (16) becomes

$$-\frac{1}{2} M_t = \int \langle \nabla^\perp f, \nabla^\perp h \rangle = \int \langle \nabla f, \nabla h \rangle \quad (20)$$

where the decomposition of  $\tilde{u}$  is  $\tilde{u} = \nabla w + \nabla^\perp f$ , and we can take  $h = f$ . The function  $f$  can be found by solving the Dirichlet-type boundary problem

$$-\operatorname{div}(\tilde{u}^\perp) = \Delta f, \quad (21)$$

$$f = 0 \text{ on } \partial\Omega_0, \quad (22)$$

which gives us the evolution equation

$$\tilde{u}_t = \frac{1}{\mu_0} D\tilde{u} \nabla^\perp \Delta^{-1} \operatorname{div}(\tilde{u}^\perp). \quad (23)$$

We may also derive a second order *local* evolution equation for  $\tilde{u}$  by using the divergence theorem with (20) to get

$$\tilde{u}_t = -\frac{1}{\mu_0} D\tilde{u} \nabla^\perp \operatorname{div}(\tilde{u}^\perp). \quad (24)$$

### 3.3 Computing the Optimal Warping Map

Typically in elastic registration and optimal flow, one wants to see an explicit warping which smoothly deforms one image into the other. This can easily be done using the solution of the Monge-Kantorovich problem. Thus, we assume now that we have applied our gradient descent process as described above and that it has converged to the Monge-Kantorovich mapping  $\tilde{u}_{MK}$ .

Following the work of Benamou and Brenier, [3], (see also [9]), we consider the following related problem:

$$\inf \int \int_0^1 \mu(t, x) \|v(t, x)\|^2 dt dx$$

over all time varying densities  $\mu$  and velocity fields  $v$  satisfying

$$\frac{\partial \mu}{\partial t} + \operatorname{div}(\mu v) = 0, \quad (25)$$

$$\mu(0, \cdot) = \mu_0, \quad \mu(1, \cdot) = \mu_1. \quad (26)$$

It is shown in [3] that this infimum is attained for some  $\mu_{min}$  and  $v_{min}$ , and that it is equal to the  $L^2$  Kantorovich–Wasserstein distance between  $\mu_0$  and  $\mu_1$ . Further, the flow  $X = X(x, t)$  corresponding to the minimizing velocity field  $v_{min}$  via

$$X(x, 0) = x, \quad X_t = v_{min} \circ X$$

is given simply as

$$X(x, t) = x + t(\tilde{u}_{MK}(x) - x). \quad (27)$$

Note that when  $t = 0$ ,  $X$  is the identity map and when  $t = 1$ , it is the solution  $\tilde{u}_{MK}$  to the Monge–Kantorovich problem. This analysis provides appropriate justification for using (27) to define our continuous warping map  $X$  between the densities  $\mu_0$  and  $\mu_1$ .

We tested the method on a sequence of brain data used in medical imaging. In Figures 1 through 4 we show a brain deformation sequence. Here, the first and last images were given, and the intermediate two were found using our process. This type of elastic brain deformation occurs during surgery, after the skull is opened.

#### 4 Conclusions

In this paper, we presented a natural method for image registration based on the classical problem of optimal mass transportation. We showed that for an  $L^2$  version of the problem, one could derive easily-implementable gradient descent equations to carry out the method. Our approach allows us to have non-uniform densities on both the domain and range.

Although applied here to the Monge–Kantorovich problem, the method used to enforce the mass preservation constraint is general and has other applications. In particular, the concept a harmonic mapping, defined as a minimizer of the Dirichlet integral, can be combined with a mass preservation constraint to obtain a new approach to mass-preserving diffeomorphisms [1]. We state the results for Euclidean space even though they apply more generally to Riemannian surfaces. As above, let  $\Omega_0, \Omega_1 \subset \mathbf{R}^2$  be subdomains equipped with positive densities  $\mu_0$  and  $\mu_1$ , respectively, and consider the minimization of the Dirichlet integral over all MP maps:

$$\min_{\tilde{u} \in MP} \int_{\Omega_0} \|D\tilde{u}\|^2. \quad (28)$$

A minimizer, when it exists, is called an MP map of *minimal distortion*.

Non-local and local gradient descent methods for computing such a map of minimal distortion can be derived in a manner very similar to that described above for the Monge–Kantorovich functional [1]. These methods

have applications to optical flow for tracking, and can be naturally combined with the Monge–Kantorovich methodology described above.

#### References

- [1] S. Angenent, S. Haker, A. Tannenbaum, and R. Kikinis, "On area preserving maps of minimal distortion," in *System Theory: Modeling, Analysis, and Control*, edited by T. Djaferis and I. Schick, Kluwer, Holland, 1999, pages 275-287.
- [2] S. Angenent, S. Haker, A. Tannenbaum, and R. Kikinis, "Laplace-Beltrami operator and brain surface flattening," *IEEE Trans. on Medical Imaging* **18** (1999), pp. 700-711.
- [3] J.-D. Benamou and Y. Brenier, "A computational fluid mechanics solution to the Monge–Kantorovich mass transfer problem," *Numerische Mathematik* **84** (2000), pp. 375-393.
- [4] Y. Brenier, "Polar factorization and monotone rearrangement of vector-valued functions," *Com. Pure Appl. Math.* **64** (1991), pp. 375-417.
- [5] M. Cullen and R. Purser, "An extended Lagrangian theory of semigeostrophic frontogenesis," *J. Atmos. Sci.* **41** (1984), pp. 1477-1497.
- [6] B. Dacorogna and J. Moser, "On a partial differential equation involving the Jacobian determinant," *Ann. Inst. H. Poincaré Anal. Non Linéaire*, **7** (1990), pp. 1-26.
- [7] D. Fry, *Shape Recognition Using Metrics on the Space of Shapes*, Ph.D. Thesis, Harvard University, 1993.
- [8] W. Gangbo, "An elementary proof of the polar factorization of vector-valued functions," *Arch. Rational Mechanics Anal.* **128** (1994), pp. 381-399.
- [9] W. Gangbo and R. McCann, "The geometry of optimal transportation," *Acta Math.* **177** (1996), pp. 113-161.
- [10] W. Gangbo and R. McCann, "Shape recognition via Wasserstein distance," Technical Report, School of Mathematics, Georgia Institute of Technology, 1999.
- [11] S. Haker, S. Angenent, A. Tannenbaum, and R. Kikinis, "Nondistorting flattening maps and the 3D visualization of colon CT images," *IEEE Trans. of Medical Imaging*, July 2000.
- [12] L. V. Kantorovich, "On a problem of Monge," *Uspekhi Mat. Nauk.* **3** (1948), pp. 225-226.
- [13] M. Knott and C. Smith, "On the optimal mapping of distributions," *J. Optim. Theory* **43** (1984), pp. 39-49.
- [14] R. McCann, "Polar factorization of maps on Riemannian manifolds," preprint 2000. Available at <http://www.math.toronto.edu/mccann>.

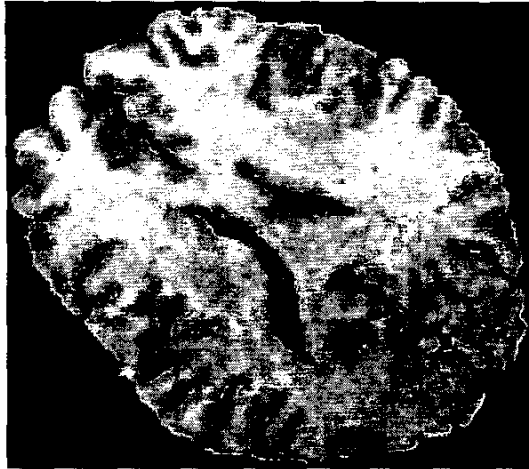


Figure 1: Brain Warping:  $t = 0.00$



Figure 2: Brain Warping:  $t = 0.33$

[15] J. Moser, "On the volume elements on a manifold," *Trans. Amer. Math. Soc.* **120** (1965), pp. 286-294.

[16] W. Press, S. Teukolsky, W. Vetterling and B. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd Edition, Cambridge University Press, Cambridge U.K., 1992.

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

[18] A. Rantzer, "A dual to Lyapunov's stability theorem," *Systems and Control Letters* **42** (2001), pp. 161-168

[19] G. Strang, *Introduction to Applied Mathematics*, Wellesley-Cambridge Press, Wellesley, Mass., 1986.

[20] A. Toga, *Brain Warping*, Academic Press, San Diego, 1999.

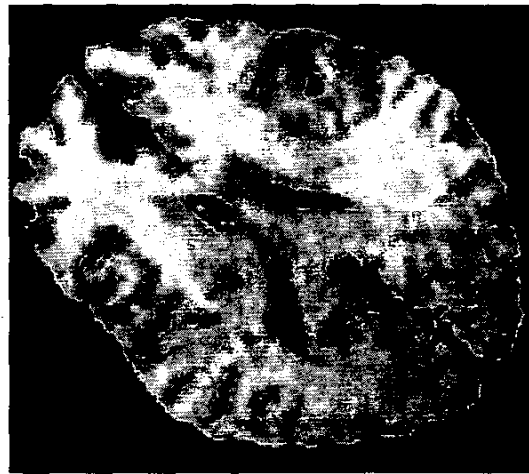


Figure 3: Brain Warping:  $t = 0.66$

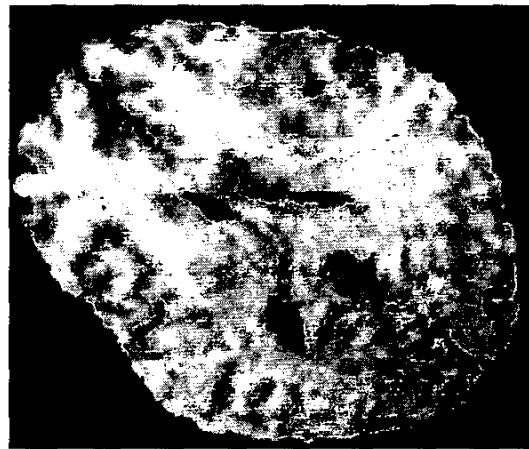


Figure 4: Brain Warping:  $t = 1.00$