Int J Comput Vis (2010) 90: 255–266
DOI 10.1007/s11263-010-0367-1

Population Shape Regression from Random Design Data

Brad C. Davis · P. Thomas Fletcher · Elizabeth Bullitt · Sarang Joshi

Received: 22 January 2008 / Accepted: 16 June 2010 / Published online: 13 July 2010

Abstract Regression analysis is a powerful tool for the study of changes in a dependent variable as a function of an independent regressor variable, and in particular it is applicable to the study of anatomical growth and shape change. When the underlying process can be modeled by parameters in a Euclidean space, classical regression techniques (Hardle, Applied Nonparametric Regression, 1990; Wand and Jones, Kernel Smoothing, 1995) are applicable and have been studied extensively. However, recent work suggests that attempts to describe anatomical shapes using *flat Euclidean spaces* undermines our ability to represent natural biological variability (Fletcher et al., IEEE Trans. Med. Imaging 23(8), 995–1005, 2004; Grenander and Miller, Q. Appl. Math. 56(4), 617–694, 1998).

In this paper we develop a method for regression analysis of general, manifold-valued data. Specifically, we extend Nadaraya-Watson kernel regression by recasting the regression problem in terms of Fréchet expectation. Although this method is quite general, our driving problem is the study anatomical shape change as a function of age from random design image data.

We demonstrate our method by analyzing shape change in the brain from a random design dataset of MR images of 97 healthy adults ranging in age from 20 to 79 years. To study the small scale changes in anatomy, we use the infinite dimensional manifold of diffeomorphic transformations, with an associated metric. We regress a representative anatomical shape, as a function of age, from this population.

B.C. Davis (✉) · E. Bullitt
University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
e-mail: brad.davis@unc.edu

P.T. Fletcher · S. Joshi
University of Utah, Salt Lake City, UT, USA

Keywords Spatio-temporal shape analysis · Kernel regression · Deformable atlas building

1 Introduction

An important area of medical image analysis is the development of methods for automated and computer-assisted assessment of anatomical change over time. For example, the analysis of structural brain change over time is important for understanding healthy aging. These methods also provide markers for understanding disease progression.

A number of longitudinal growth models have been developed to provide this type of analysis to time-series imagery of a single subject (e.g., Beg 2003; Clatz et al. 2005; Miller 2004; Thompson et al. 2000). While these methods provide important results, their use is limited by their reliance on longitudinal data, which can be impractical to obtain for many medical studies. Also, while these methods allow for the study of an *individual's* anatomy over time, they do not apply when the *average* growth for a population is of interest.

Random design data sets, which contain anatomical data from many different individuals, provide a rich environment for addressing these problems. However, in order to detect time-related trends in such data, two distinct aspects of anatomical variation must be separated: individual variation and time effect. For measurements that naturally form Euclidean vector spaces, this separation can be achieved by *regressing* a representative value over time from the data.

For example, in Fig. 1 we apply kernel regression to measurements reported in a study by Mortamet et al. (2005) on the effect of aging on gray matter and ventricle volume in the brain. The regression curves demonstrate the average volume, as a function of patient age, of these structures. These trends—on average there is a loss of gray matter

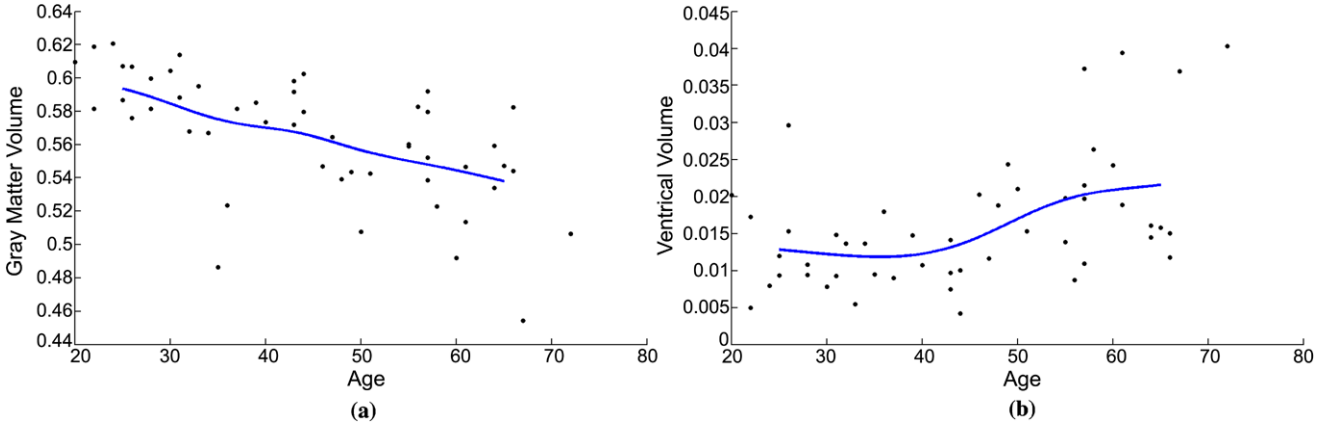


Fig. 1 (Color online) Illustration of univariate kernel regression: the effect of aging on gray matter (**a**) and ventricle volume (**b**) in the brain. *Circles* represent volume measurements relative to total brain volume. Kernel regression is used to estimate the relationship between patient age and structure volume (*filled lines*)

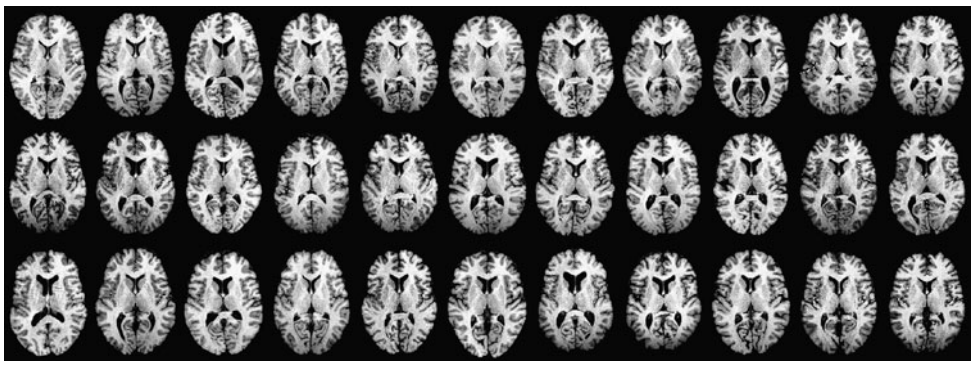


Fig. 2 Brain image database. To demonstrate the extent of natural brain shape variability within a population of healthy subjects, a mid-axial slice is presented for a sample of images used in this study. The images are arranged in order of increasing patient age from 30 (*top left*) to 60 (*bottom right*). Because of the complexity of the shapes and the high level of natural shape variability, it is extremely difficult to visually discern any patterns within these data

and expansion of the ventricles—have been widely reported in the medical literature on aging (Guttmann et al. 1998; Matsumae et al. 1996; Mortamet et al. 2005). While volume-based regression analysis is important, it does not provide any information about the detailed *shape* changes that occur in the brain, on average, as a function of age. This has motivated us to study regression of shapes.

Recent work has suggested that representing the geometry of shapes in *flat Euclidean vector spaces* limits our ability to represent natural variability in populations (Fletcher et al. 2004; Grenander and Miller 1998; Miller 2004). For example, Fig. 2 demonstrates the amazing *non-linear* variability in brain shape among a population of healthy adults. The analysis of transformation groups that describe shape change are essential to understanding this shape variability. These groups vary in dimensionality from simple rigid rotations to the infinite-dimensional group of diffeomorphisms (Miller and Younes 2001). These groups are not generally vector spaces and are instead naturally represented as manifolds.

A number of authors have contributed to the field of statistical analysis on manifolds (see Pennec 2006 for a more detailed history). Early work on manifold statistics includes directional statistics (Bingham 1974; Jupp and Mardia 1989) and statistics of point set shape spaces (Kendall 1984; Le and Kendall 1993). The large sample properties of sample means on manifolds are developed in Bhattacharya and Patrangenaru (2002, 2003). Jupp and Kent (1987) describe a method of regression of spherical data that 'unwraps' the data onto a tangent plane, where standard curve fitting methods can be applied. In Fletcher et al. (2004), Joshi et al. (2004), Pennec (2006), statistical concepts such as averaging and principal components analysis were extended to manifolds representing anatomical shape variability. Many of the ideas are based on the method of averaging in metric spaces proposed by Fréchet (1948).

In this paper we use the notion of Fréchet expectation to generalize regression to manifold-valued data. We use this method to study spatio-temporal anatomical shape change in a random design database consisting of three-dimensional MR images of healthy adults. Our method generalizes Nadaraya-Watson kernel regression in order to compute representative images of this population over time. To determine the shape change in the population over time, we apply a diffeomorphic growth model (Miller 2004) to this time-indexed population representative image.

2 Methods

2.1 Review of Univariate Kernel Regression

Univariate kernel regression (Hardle 1990; Wand and Jones 1995) is a non-parametric method used to estimate the relationship, on average, between an independent random variable T and a dependent random variable Y. The estimation is based on a set of observations $\{t_i, y_i\}_{i=1}^N$ drawn from the joint distribution of T and Y. This relationship between T and Y can be modeled as $y_i = m(t_i) + \epsilon_i$, where ϵ_i describes the random error of the model for the ith observation and m is the unknown function that is to be estimated.

In this setting, $m(t)$ is defined by the conditional expectation

$$m(t) \equiv E(Y|T=t) = \int y \frac{f(t,y)}{f_T(t)}\, dy \tag{1}$$

where $f_T(t)$ is the marginal density of T and $f(t,y)$ is the joint density function of T and Y. For random design data, both $f(t,y)$ and $f_T(t)$ are unknown and so m has no closed-form solution. A number of estimators for m have been proposed in the kernel regression literature.

One such estimator—the Nadaraya-Watson kernel regression estimator (Nadaraya 1964; Watson 1964)—can be derived from (1) by replacing the unknown densities with their kernel density estimates

$$\hat{f}_T^h(t) \equiv \frac{1}{N}\sum_{i=1}^N K_h(t-t_i) \quad \text{and}$$

$$\hat{f}^{h,g}(t,y) \equiv \frac{1}{N}\sum_{i=1}^N K_h(t-t_i)K_g(y-y_i). \tag{2}$$

In these equations, K is a function that satisfies $\int_{\mathbb{R}} K(t)\,dt = 1$. $K_h(t) \equiv \frac{1}{h}K(\frac{t}{h})$ and $K_g(t) \equiv \frac{1}{g}K(\frac{t}{g})$ are kernel functions with bandwidths h and g respectively.

Plugging these density estimates into (1) gives

$$\hat{m}_{h,g}(t) = \int y \frac{\frac{1}{N}\sum_{i=1}^N K_h(t-t_i)K_g(y-y_i)}{\frac{1}{N}\sum_{i=1}^N K_h(t-t_i)}\, dy. \tag{3}$$

Finally, assuming that K is symmetric about the origin, integration of the numerator leads to

$$\hat{m}_h(t) = \frac{\sum_{i=1}^N K_h(t-t_i)y_i}{\sum_{i=1}^N K_h(t-t_i)}. \tag{4}$$

Intuitively, the Nadaraya-Watson estimator returns the weighted average of the observations y_i, with the weighting determined by the kernel. Note that $\hat{f}^{h,g}(t,y)$ is factored out of the estimator—the weights only depend on the values t_i.

In Fig. 1 we illustrate univariate kernel regression by applying it to demonstrate the effect of aging on ventricle volume and gray matter volume in the brain. This illustration is based on data collected by Mortamet et al. (2005). Each point represents a volume measurement, relative to total brain volume, for a particular patient. These measurements were derived from 3D MR images of 50 healthy adults ranging in age from 20 to 72 using an expectation-maximization based automatic segmentation method (Leemput et al. 1999). We used kernel regression to estimate the relationship, on average, between volume and patient age (filled lines). A Nadaraya-Watson kernel estimator with a Gaussian kernel of width $\sigma = 6$ years was used.

2.2 Kernel Regression on Riemannian Manifolds

In this section we consider the regression problem in the more general setting of manifold-valued observations. Let $\{t_i, p_i\}_{i=1}^N$ be a collection of observations where the t_i are drawn from a univariate random variable T, but where p_i are points on a Riemannian manifold $\mathcal{M}$. The classical kernel regression methods presented in Sect. 2.1 are not applicable in this setting because they rely on the vector space structure of the observations. In particular, the addition operator in (4) is not well defined.

The goal is to determine the relationship, on average, between the independent variable T and the distribution of the points $\{p_i\}$ on the manifold. This relationship can be modeled by

$$p_i = \mathrm{Exp}_{m(t_i)}(\epsilon_i) \tag{5}$$

where $m : \mathbb{R} \to \mathcal{M}$ defines a curve on $\mathcal{M}$. The error term $\epsilon_i \in T_{m(t_i)}\mathcal{M}$ is a tangent vector that is interpreted as the displacement along the manifold of each observation p_i from the curve $m(t)$. The exponential mapping, Exp, returns the point on $\mathcal{M}$ at time one along the geodesic flow beginning at $m(t_i)$ with initial velocity ϵ_i.

Following the univariate case, we define the regression function $m(t)$ in terms of expectation. However, in this case we generalize the idea of expectation of real random variables to manifold-valued random variables via Fréchet expectation (Fréchet 1948; Karcher 1977). Let $f(p)$, $p \in \mathcal{M}$

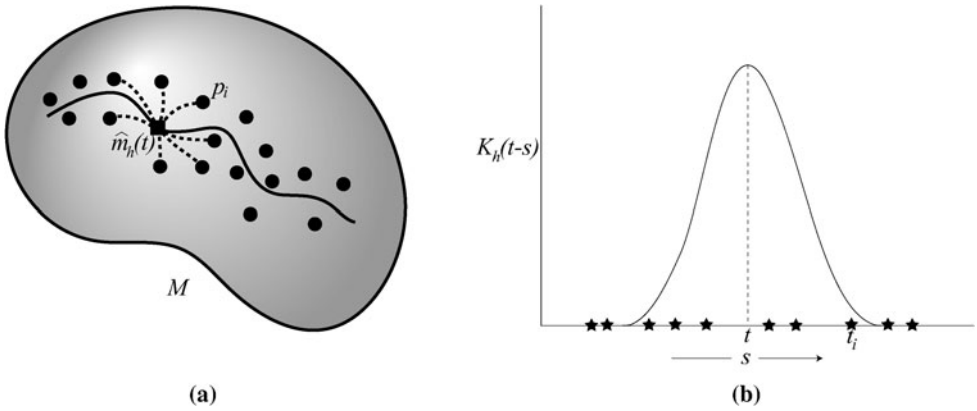


Fig. 3 Manifold kernel regression schematic. (**a**) For any value of the predictor variable t, the manifold-valued observations p_i are summarized by the weighted Fréchet mean point $\hat{m}_h(t)$. (**b**) As in the univariate case, the weights are determined by the predictor values t_i and the kernel K_h

be a probability density on the manifold. The Fréchet expectation is defined as

$$\mathbb{E}_f[p] \equiv \operatorname*{argmin}_{q\in\mathcal{M}} \int_{\mathcal{M}} d(q,p)^2 f(p)dp \tag{6}$$

where $d(q,m)$ is the metric on the manifold $\mathcal{M}$. This definition is motivated by a minimum variance characterization of the mean, where variance is defined in terms of the metric. Note that Fréchet expectation might not be unique (Karcher 1977). Using the above definition, an empirical estimate of the Fréchet mean, given a collection of observations $\{p_i, i = 1\ldots N\}$ on a manifold $\mathcal{M}$, is defined by

$$\mu = \operatorname*{argmin}_{q\in\mathcal{M}} \frac{1}{N}\sum_{i}^{N} d(q,p_i)^2. \tag{7}$$

Motivated by the definition of the Nadaraya-Watson estimator as a weighted averaging, we define a *manifold kernel regression estimator* using the weighted Fréchet empirical mean estimator as

$$\hat{m}_h(t) = \operatorname*{argmin}_{q\in\mathcal{M}} \left(\frac{\sum_{i=1}^{N} K_h(t-t_i)d(q,p_i)^2}{\sum_{i=1}^{N} K_h(t-t_i)} \right). \tag{8}$$

This estimator is illustrated in Fig. 3. Notice that when the manifold under study is a Euclidean vector space, equipped with the standard Euclidean norm, the above minimization results in the Nadaraya-Watson estimator.

2.3 Bandwidth Selection

It is well known within the kernel regression literature that kernel width plays a crucial role in determining regression results (Wand and Jones 1995). In particular, it is important to select a bandwidth that captures relevant population-wide changes without either oversmoothing and missing relevant changes or undersmoothing and biasing the results based on individual noisy data points. The 'Goldie Locks' method of tuning the bandwidth until the results are most pleasing is a common subjective method for bandwidth selection. However, non-subjective methods may be required, for example, when kernel regression is part of a larger statistical study. A number of automatic kernel bandwidth selection techniques have been proposed for this purpose (Wand and Jones 1995; Jones et al. 1996; Loader 1999).

One classic method for automatic bandwidth selection is based on least squares cross-validation. This method is easily extended to the manifold regression setting in the following way. For observations $\{t_i, p_i\}_{i=1}^{N}$, with $t_i \in \mathbb{R}$ and $p_i \in \mathcal{M}$, the least squares cross-validation estimate for the optimal bandwidth h is defined as

$$\hat{h}_{\text{LSCV}} \equiv \operatorname*{argmin}_{h\in\mathbb{R}^+} \frac{1}{N}\sum_{i=1}^{N} d(\hat{m}_h^{i-}(t_i), p_i)^2 \tag{9}$$

where

$$\hat{m}_h^{i-}(t) \equiv \operatorname*{argmin}_{q\in\mathcal{M}} \left(\frac{\sum_{j=1,j\neq i}^{N} K_h(t-t_j)d(q,p_j)^2}{\sum_{j=1,j\neq i}^{N} K_h(t-t_j)} \right) \tag{10}$$

is the manifold kernel regression estimator with the i-th observation left out.

It is important to note that (9) may achieve multiple local minima; this is true even in Euclidean space (Hall and Marron 1991).

2.4 Regression of Rotational Pose ($SO(3)$)

Before we present results of the study of brain growth, we exemplify the methodology in detail on the finite-dimensional Lie group of 3D rotations, $SO(3)$.

Following the approach in Buss and Fillmore (2001), we solve the weighted averaging problem in (8) by a gradient descent algorithm. The tangent space of $SO(3)$ at the identity is the Lie algebra of 3×3 skew-symmetric matrices, denoted $\mathfrak{so}(3)$. We equip $SO(3)$ with the standard bi-invariant metric, given by the Frobenius inner product on $\mathfrak{so}(3)$. The tangent space at an arbitrary rotation $R \in SO(3)$ is given by either left or right multiplication of $\mathfrak{so}(3)$ by R.

The Lie group exponential map and its inverse, the log map, are used to compute geodesics and distances. The exponential map for a tangent vector $X \in \mathfrak{so}(3)$ is given by

$$\exp(X) = \begin{cases} I, & \theta = 0, \\ I + \frac{\sin\theta}{\theta} X + \frac{1-\cos\theta}{\theta^2} X^2, & \theta \in (0, \pi), \end{cases} \tag{11}$$

where $\theta = \sqrt{\frac{1}{2}\operatorname{tr}(X^T X)}$. A geodesic $\gamma(t)$ starting at a point $R \in SO(3)$ with initial velocity RX is given by $\gamma(t) = R\exp(tX)$. The Lie group log map for a rotation matrix $R \in SO(3)$ is given by

$$\log(R) = \begin{cases} I, & \theta = 0, \\ \frac{\theta}{2\sin\theta}(R - R^T), & |\theta| \in (0, \pi), \end{cases} \tag{12}$$

where $\operatorname{tr}(R) = 2\cos\theta + 1$. The distance between two rotations $R_1, R_2 \in SO(3)$ is given by $d(R_1, R_2) = \|\log(R_1^{-1} R_2)\|$.

Now consider the weighted averaging problem with rotation data $R_i \in SO(3)$ and corresponding weights $w_i = K_h(t - t_i)/\sum_{j=1}^N K_h(t - t_j)$. The regression problem in (8) minimizes the weighted sum-of-squared distance function of the form $f(R, \{R_i, w_i\}) = (1/2)\sum_i w_i d(R, R_i)^2$. The gradient for this function at a point $R \in SO(3)$ is given by $\nabla_R f = -\sum_i w_i R \log(R^{-1} R_i)$. Therefore, given the estimate $\hat{R}_k$ for the weighted average, the gradient descent update to solve (8) is given by $\hat{R}_{k+1} = \hat{R}\exp(-R^{-1}\nabla_{\hat{R}_k} f)$.

2.5 Kernel Regression for Populations of Brain Images

In this section we apply our shape regression methodology to study the effect of aging on brain *shape* from random design image data. We have observations of the form $\{t_i, I_i\}_{i=1}^N$ where t_i is the age of patient i and I_i is a three-dimensional image that we identify with the anatomical configuration of patient i. We seek the unknown function m that associates a representative anatomical configuration, and its associated image $\hat{I}$, with each age.

Let $\Omega \subset \mathbb{R}^3$ be the underlying coordinate system of the observed images I_i. Each image $I \in \mathcal{I}$ can be formally defined as an L^2 function from Ω to the reals. However, it is important to point out that we cannot rely on the natural L^2 structure of the images themselves for our analysis. While images can be added voxel-wise, the result is a loss of any identification with the anatomical configuration.

Instead, we represent anatomical differences in terms of transformations of the underlying image coordinates. This approach is common within the shape analysis literature (Grenander and Miller 1998; Miller et al. 1997). Because we are interested in capturing the large, natural geometric variability evident in the brain (cf. Fig. 2), we represent shape change as the action of the group of diffeomorphisms, denoted by $\mathcal{H}$. In the rest of this section, we formalize this notion and define a distance between shapes that is valid in this setting and will allow us to apply our regression methodology.

Let $\mathcal{H}$ be the group of diffeomorphisms that are isotopic to the identity. Each element $\phi : \Omega \to \Omega$ in $\mathcal{H}$ deforms an image according to the following rule

$$I_\phi(x) = I(\phi^{-1}(x)). \tag{13}$$

We apply the theory of large deformation diffeomorphisms (Beg et al. 2005; Dupuis and Grenander 1998; Joshi and Miller 2000; Miller and Younes 2001) to generate deformations ϕ that are solutions to the Lagrangian ODEs $\frac{d}{ds}\phi_s(x) = v_s(\phi_s(x))$ for a simulated time parameter $s \in [0, 1]$. The transformations are generated by integrating the time-varying velocity fields v_s forward in time.

We introduce a metric on $\mathcal{H}$ using a Sobelev norm via a partial differential operator A applied to v where $\|v_s\|_V^2 \equiv \int_\Omega \langle A v_s, v_s\rangle\, dx$. Let $e \in \mathcal{H}$ be the identity transformation. We define the squared metric $d_{\mathcal{H}}(e, \phi)^2$ as

$$d_{\mathcal{H}}(e, \phi)^2 = \min_{v:\dot{\phi}_s = v_s(\phi_s)} \int_0^1 \|v_s\|_V^2\, ds \tag{14}$$

subject to

$$\phi(x) = x + \int_0^1 v_s(\phi_s(x))\, ds \quad \text{for all } x \in \Omega. \tag{15}$$

The distance between any two diffeomorphisms is defined by

$$d_{\mathcal{H}}(\phi_1, \phi_2)^2 = d_{\mathcal{H}}(e, \phi_1^{-1} \circ \phi_2)^2. \tag{16}$$

This distance satisfies all of the properties of a metric: it is non-negative, symmetric, and satisfies the triangle inequality (Miller et al. 2002).

Using this metric on $\mathcal{H}$, we can define the distance between two images as

$$d_{\mathcal{I}}(I_1, I_2)^2 \equiv \min_{v:\dot{\phi}_s = v_s(\phi_s)} \left[\int_0^1 \|v_s\|_V^2\, ds + \frac{1}{\sigma^2}\|I_1(\phi^{-1}) - I_2\|_{L_2}^2 \right] \tag{17}$$

where the second term accounts for the noise model of the image (Joshi et al. 2004). While this construction is motivated by the metric on $\mathcal{H}$, it does not strictly define a Riemannian metric on the space of anatomical images (because of the second term). In the future we plan to define distance in terms of the elegant construction described in Trouvé and Younes (2005).

Having defined a metric on the space of images that accommodates anatomical variability, we can apply that metric to regress a representative anatomical configuration, with

associated image, from our observations $\{t_i, I_i\}$

$$\hat{I}_h(t) = \operatorname*{argmin}_{I\in\mathcal{I}} \left(\frac{\sum_{i=1}^N K_h(t-t_i) d_{\mathcal{I}}(I, I_i)^2}{\sum_{i=1}^N K_h(t-t_i)} \right). \tag{18}$$

Equation (18) expresses the following intuitive idea: For any age t, the population can be represented by the anatomical configuration that is centrally located, according to $d_{\mathcal{I}}$, among the observations that occur near in time to t. As in the univariate case, the weights are determined by the kernel K.

2.6 Diffeomorphic Growth Model

Having regressed a population representative anatomical image, as a function of age, we can now study the local shape changes evident—for the population—as a function of age. We use the manifold kernel regression estimator to extend a single-subject longitudinal growth model in order to study *population-average* geometric change. In particular, we estimate the age-indexed diffeomorphism that quantifies the fine scale anatomical shape change of the population representative $\hat{I}$.

2.6.1 Single-Subject Growth Model

The dynamic growth model described in Miller (2004) associates a single subject with a collection of image observations $J_t \in \mathcal{I}$, which are acquired over a period of time $t \in [0, 1]$. The goal is to determine the diffeomorphic flow g_t that deforms an exemplar image J_α through time in such a way that it matches these image observations. In practice J_0 is used as the exemplar image. This methodology has been applied, for example, to measure growth or atrophy of structures within the brain.

The formalization of the growth problem is similar to the definition of the image metric $d_{\mathcal{I}}$ (cf. (17)) in that it is defined as a minimization problem that seeks to find a solution g_t that requires the least amount of deformation according to the metric $d_{\mathcal{H}}$ on the space of diffeomorphisms:

$$\operatorname*{argmin}_{v:\dot{g}_t = v_t(g_t)} \left[\int_0^1 \| v_t \|_V^2 \, dt + \frac{1}{\sigma^2} \int_0^1 \| J_\alpha(g_t^{-1}) - J_t \|_{L_2}^2 \, dt \right]. \tag{19}$$

A primary difference between this equation and (17) is that in the case of the growth model the second term is integrated over time. This enforces the requirement that the deforming exemplar image $J_\alpha(g_t^{-1})$ match the observed imagery J_t throughout the growth period.

It has been shown using the calculus of variations (Miller et al. 2002) that the solution to (19) satisfies

$$Av_t = -\frac{1}{2\sigma^2} \nabla(J_\alpha \circ g_t^{-1}) \int_t^1 (J_u(g_u \circ g_t^{-1}) - J_\alpha(g_t^{-1})) |D(g_u \circ g_t^{-1})| \, du \tag{20}$$

where $\nabla(J_\alpha \circ g_t^{-1})$ is the gradient of the deformed exemplar image and $D(g_u \circ g_t^{-1})$ is the Jacobian of the diffeomorphic transformation that maps the anatomical configuration at time t to the configuration at time u. The discrete version of this equation is used to construct an iterative solution for v_t. g_t is initially set to the identity map for all t. At each iteration $v_t, t \in [0, 1]$ is updated according to the observed images J_t and the current estimate of $g_t, t \in [0, 1]$.

2.6.2 Population Growth Model

In order to extend this growth model to apply to a population of subjects, we replace the subject-specific collection of observed imagery J_t with the *expected* observed imagery, as a function of time, for the population (cf. Fig. 4). This is achieved by combining the manifold kernel regression estimator (see (18)) with the growth model (see (19)):

$$\operatorname*{argmin}_{v:\dot{g}_t = v_t(g_t)} \int_0^1 \| v_t \|_V^2 \, dt + \frac{1}{\sigma^2} \int_0^1 \left\| I_\alpha(g_t^{-1}) - \operatorname*{argmin}_{I\in\mathcal{I}} \left(\frac{\sum_{i=1}^N K_h(t-t_i) d_{\mathcal{I}}(I, I_i)^2}{\sum_{i=1}^N K_h(t-t_i)} \right) \right\|_{L_2}^2 dt. \tag{21}$$

In this way the population representative images serve as a collection of *population average* time-sequence imagery.

In order to solve (21) we first solve the interior minimization problem for a discrete collection of time points. This is legitimate since this problem does not depend on the growth deformation g_t. Once these population representative images are computed, the time-indexed deformation g_t is computed using the iterative method based on (20). We use $I_\alpha \equiv \hat{I}(0)$ as our population exemplar image. In order to speed convergence, we apply the growth model within a three-level multi-resolution framework where initial solutions at coarser scale levels are used to initialize the optimization procedure at finer scale levels.

Once g_t is computed, it can be analyzed to determine local, age-indexed geometric change for the population. For example, instantaneous local growth and atrophy can be measured via the log-determinant of the Jacobian of the velocity field defined by

$$\log \begin{vmatrix} \frac{\partial \dot{g}_t^1}{\partial x^1}(x) & \frac{\partial \dot{g}_t^1}{\partial x^2}(x) & \frac{\partial \dot{g}_t^1}{\partial x^3}(x) \\ \frac{\partial \dot{g}_t^2}{\partial x^1}(x) & \frac{\partial \dot{g}_t^2}{\partial x^2}(x) & \frac{\partial \dot{g}_t^2}{\partial x^3}(x) \\ \frac{\partial \dot{g}_t^3}{\partial x^1}(x) & \frac{\partial \dot{g}_t^3}{\partial x^2}(x) & \frac{\partial \dot{g}_t^3}{\partial x^3}(x) \end{vmatrix}. \tag{22}$$

Values of the log-Jacobian greater than zero indicate local expansion; values less than zero indicate local contraction.

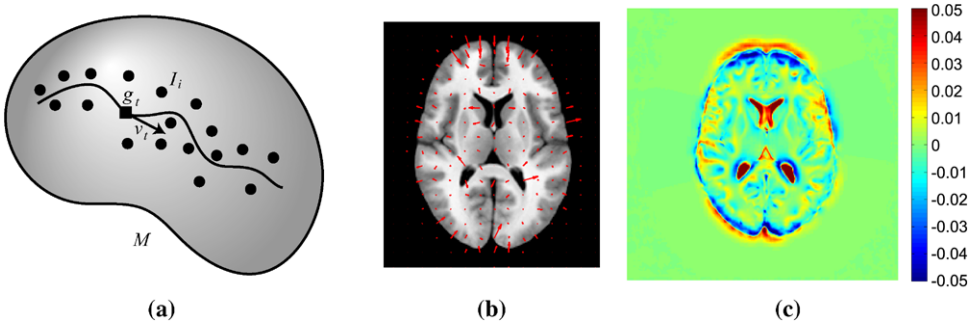


Fig. 4 (Color online) Population growth model schematic. (**a**) The diffeomorphism g_t quantifies the geometric change of $\hat{I}$ throughout the growth period. (**b**) The velocity field that is identified with the tangent vector $v_t = \dot{g}_t$ is overlaid on the underlying anatomical image $\hat{I}(t)$. The *arrows* indicate instantaneous shape change at age t. (**c**) This colormap identifies regions of local expansion and contraction of the underlying anatomy at time t. *Red* indicates expansion; *blue* indicates contraction. See text for details

2.6.3 Algorithm

In summary, the algorithm for computing a population growth model for a collection of age/image pairs, $\{t_i, I_i\}_{i=1}^N$, consists of two steps.

1. *Compute the kernel regression*. Choose a bandwidth parameter, h, and a discrete sampling of the age range, $s_1, \ldots, s_T$, and compute the corresponding regressed images, $\hat{I}_h(s_k)$, as the weighted average given in (18). For the results in this paper we approximate solutions to (18) using an iterative greedy algorithm that is similar to the method described in Joshi et al. (2004).
2. *Compute the population growth model*. Solve the growth model (21), plugging in the regressed images, $I_h(s_k)$, as the data. This is done using the gradient descent given by (20). The result is a time-varying deformation represented by a set of age-indexed vector fields v_{s_k}. Changes in shape over time can be quantified using these vector fields, for instance, by analyzing the log of the deformation Jacobian.

3 Results

3.1 Synthetic Data Experiment

Before describing the anatomical study, we present a proof of concept experiment based on synthetic data. In this experiment, we apply our manifold regression method to a database of synthetic 2D images that were generated from a known, underlying geometric process. Or goal is to recover, from the imagery alone, the underlying geometric change.

The database consists of two cohorts that each contain 100 256×256 2D bulls-eye images. The cohorts, B_1 and B_2, differ by the amount of *random* geometric variation present. Each image is associated with a particular value of the synthetic predictor variable $t \in [0, 1]$; the values of t for the database were drawn from a uniform random distribution on $[0, 1]$. For the i-th image there are three disks which independently change in radii according to

$$\begin{aligned} r_1(t_i) &= f_1(t_i) + \epsilon_i + \epsilon_{i,1} \\ r_2(t_i) &= f_2(t_i) + \epsilon_i + \epsilon_{i,2} \\ r_3(t_i) &= f_3(t_i) + \epsilon_i + \epsilon_{i,3} \end{aligned} \tag{23}$$

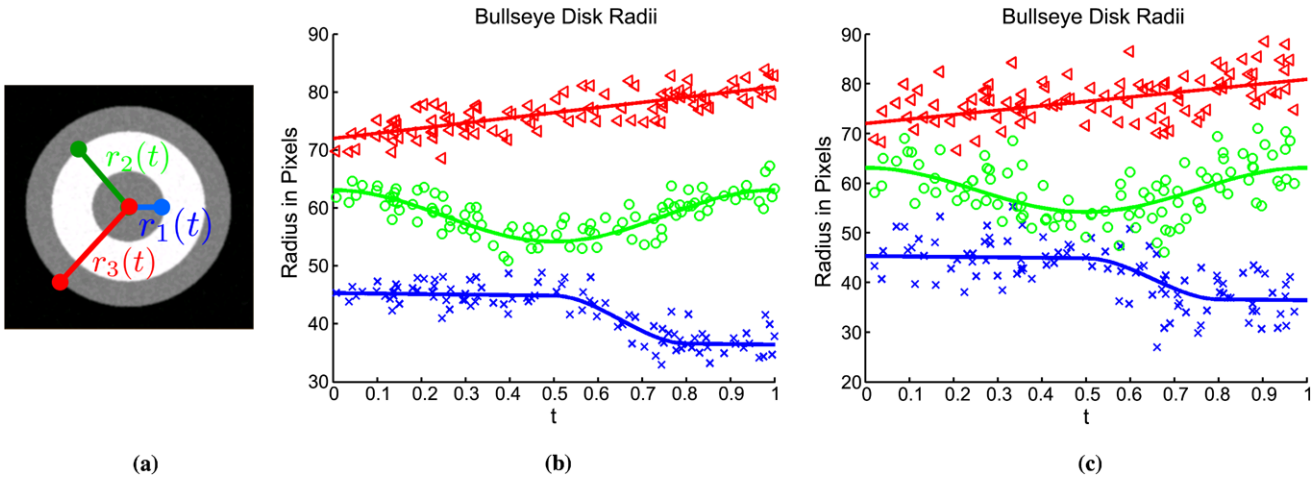
subject to

$$r_1(t) < r_2(t) < r_3(t) \quad \text{for all } t \in [0, 1]. \tag{24}$$

The functions f_1, f_2, and f_3 are known; they *define* the noise-free, ground-truth geometric change as a function of t. Noise is added to these radius functions via the zero mean Gaussian random variables $\epsilon_i, \epsilon_{i,1}, \epsilon_{i,2}$ and $\epsilon_{i,3}$. For cohort B_1, $\epsilon_i \sim N(\mu = 0, \sigma^2 = 4$ pixels) and $\epsilon_{i,1}, \epsilon_{i,2}, \epsilon_{i,3} \sim N(\mu = 0, \sigma^2 = 1$ pixels). For cohort B_2, $\epsilon_i \sim N(\mu = 0, \sigma^2 = 16$ pixels) and $\epsilon_{i,1}, \epsilon_{i,2}, \epsilon_{i,3} \sim N(\mu = 0, \sigma^2 = 4$ pixels). Once the image geometries are fixed i.i.d Gaussian noise is added to the image intensities. Figure 5 contains a schematic of the image generation process. Figure 6 displays a sample of the images from this database.

For each cohort, we applied our algorithm in order to regress a population representative bulls-eye image for 8 equally spaced values of t. A kernel bandwidth of $\sigma = 0.045$ was used. For this experiment, the solutions to (17) were computed using MATLAB codes based on the LDDMM algorithm described in Beg (2003), Beg et al. (2005).

Figure 7 contains the results of this experiment. The regressed images are shown in the background. The ground truth radii values, f_1, f_2, and f_3, are depicted as colored overlays. The close agreement with the regressed images and the overlays indicates that the underlying geometric process was recovered from the image database—that is, the underlying time effect was separated from the random geometric variation. Comparing the results for the two cohorts, the regression of the geometries is rather robust to level of the geometrical noise. Only a slight degradation in accuracy



(a) (b) (c)

Fig. 5 (Color online) Synthetic bulls-eye data set construction. The bulls-eye database contains 200 total 2D images; each images is associated with a value of t drawn from a uniform distribution on $[0, 1]$. (**a**) Images are generated from three independent, noisy radius values: r_1, r_2, and r_3. (**b**) Cohort B_1: each observed radius value (markers) is a function of t and is determined by adding random noise to the ground truth functions f_1, f_2, and f_3, which are depicted by the *solid curves*. (**c**) The second cohort, B_2, was generated using a higher level of random geometric variation

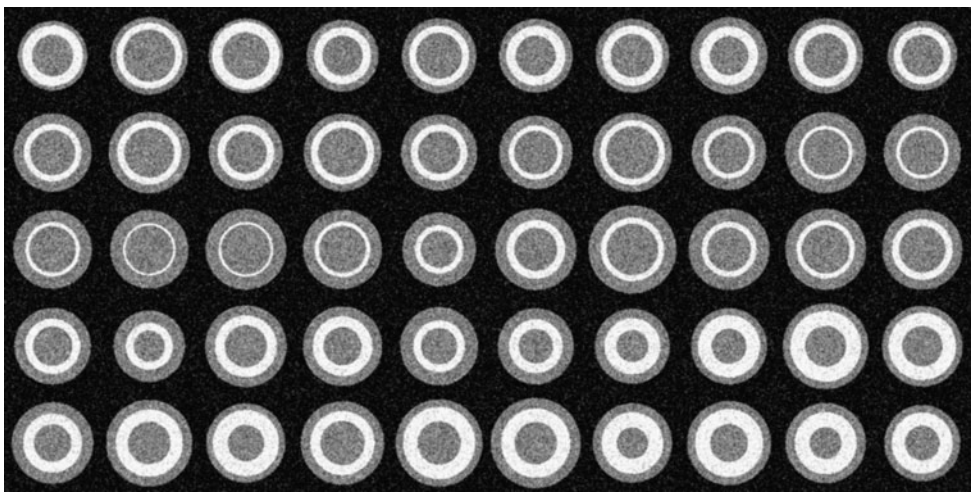


Fig. 6 Random design database of 2D bulls-eye images. These images are taken from cohort B_2. Associated time measurements increase from left to right and from top to bottom. Inner, middle, and outer disk radii are generated by adding noise to the underlying curves depicted in Fig. 5(b) and (c)

of the estimate is seen with a four-fold increase in the radii noise variance.

3.2 Regressing Average Change of the Healthy Brain from 3D MR Images

To demonstrate our method for estimating cross-sectional growth, we applied the algorithm to a database of 3D MR images. The database contains MRA, T1-FLASH, T1-MPRAGE, and T2-weighted images from 97 healthy adults ranging in age from 20 to 79 (Lorenzen et al. 2006). For this study we only utilized the T1-FLASH images; these images were acquired at a spatial resolution of 1 mm × 1 mm × 1 mm using a 3 Tesla head-only scanner. The tissue exterior to the brain was removed using a mask generated by a brain segmentation tool described in Prastawa et al. (2004). This tool was also used for bias correction. In the final preprocessing step, all of the images were spatially aligned to an atlas using affine registration.

We applied our algorithm separately for males and females. We selected only patients for which T1-Flash data was available. The final size of the male cohort is 38 subjects ranging in age from 22 to 72; the final size of the female cohort is 46 subjects ranging in age from 20 to 66. Midaxial slices for a sample of these subjects are shown in Fig. 2.

We applied the manifold kernel regression estimator (18) to compute representative anatomical images for each cohort. Images were computed for ages 30 to 60 at increments of 1 year using a Gaussian kernel with $\sigma = 6$ years. This bandwidth was subjectively determined. Figures 8 and 9 contain slices from these representative images.

We applied the diffeomorphic growth estimation algorithm described in Sect. 2.6 to determine the anatomical

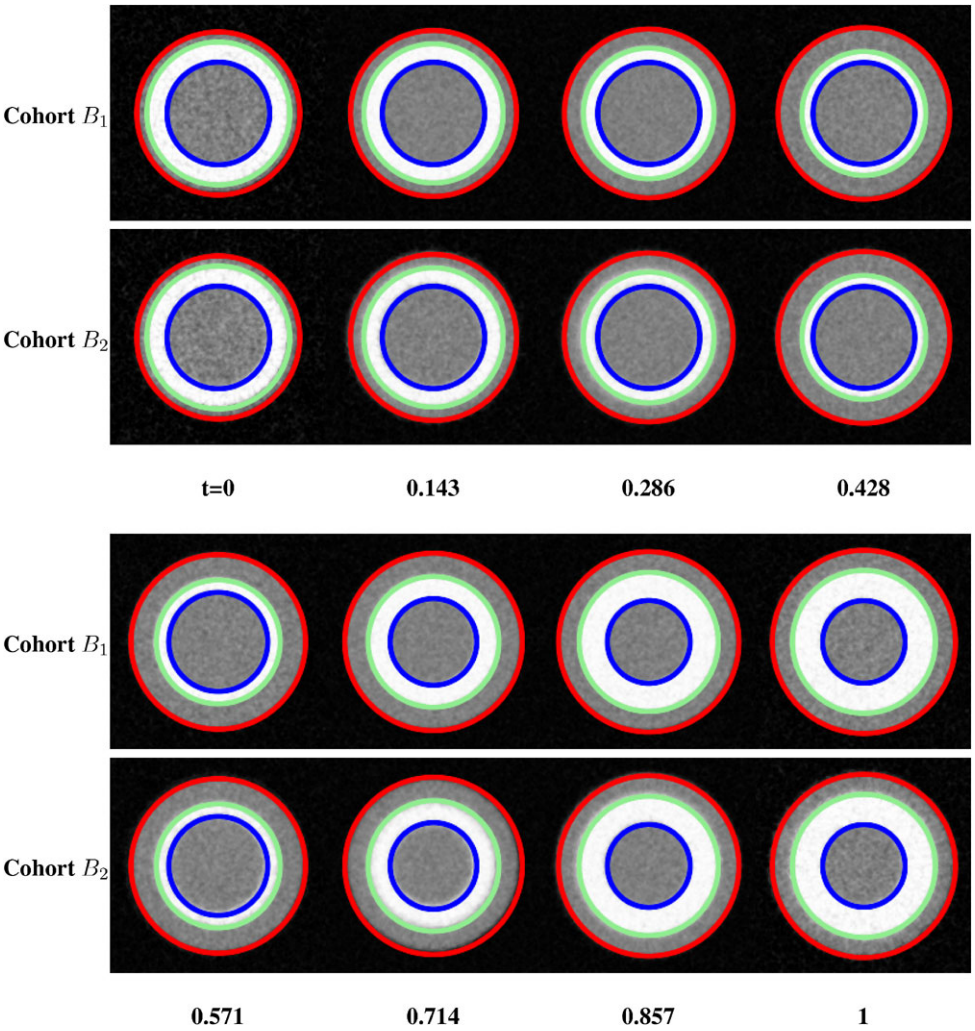


Fig. 7 (Color online) Regression results for synthetic bulls-eye database. These images show the regression results for the bulls-eye database at 8 equally-spaced time points for cohorts B_1 and B_2. Colored overlays denote the ground truth radii as determined by the underlying curves in Fig. 6(b) and (c)

shape change over time for each cohort. Figure 10 illustrates the instantaneous change in the deformation at 8 different ages. More precisely, the figure shows the log-determinant of the Jacobian of the time-derivative of the deformation. In these images, red pixels indicate *expansion* of the underlying tissue, at the given age, while blue pixels indicate *contraction*. According to these determinant maps, expansion of the ventricles is evident for each age group. However, the expansion is accelerated for ages 50 to 60. Note that this finding agrees well with volume-based regression analysis from Fig. 1.

3.2.1 Computational Strategy

For this study we approximate solutions to (18) using an iterative greedy algorithm that is similar to the method described in Joshi et al. (2004). Results were computed using a multithreaded C++ implementation on an 8 processor (16 core) 3 GHz system with approximately 64 gigabytes of RAM. Processing time averaged 116 minutes per $256 \times 256 \times 256$ regressed image volume.

When computing each representative image $\hat{I}(x)$, we use a multi-resolution approach that generates images at progressively higher resolutions, where each level is initialized by the results at the next coarsest scale. This strategy has the dual benefits of (a) addressing the large scale shape changes first and (b) speeding algorithm convergence.

The dominating computation at each iteration is a Fast Fourier Transform. The order of the algorithm is $MNn \log n$ where M is the number of iterations, N is the number of images, and n is the number of voxels along the largest dimension of the images. Therefore, the complexity grows linearly with the number of observations, making this algorithm suitable for application to large data sets.

4 Conclusion and Future Work

We have proposed a method for *population shape regression* that enables novel analysis of population shape and growth from random design data when the underlying shape model is non-Euclidean. While the method is quite general, in this

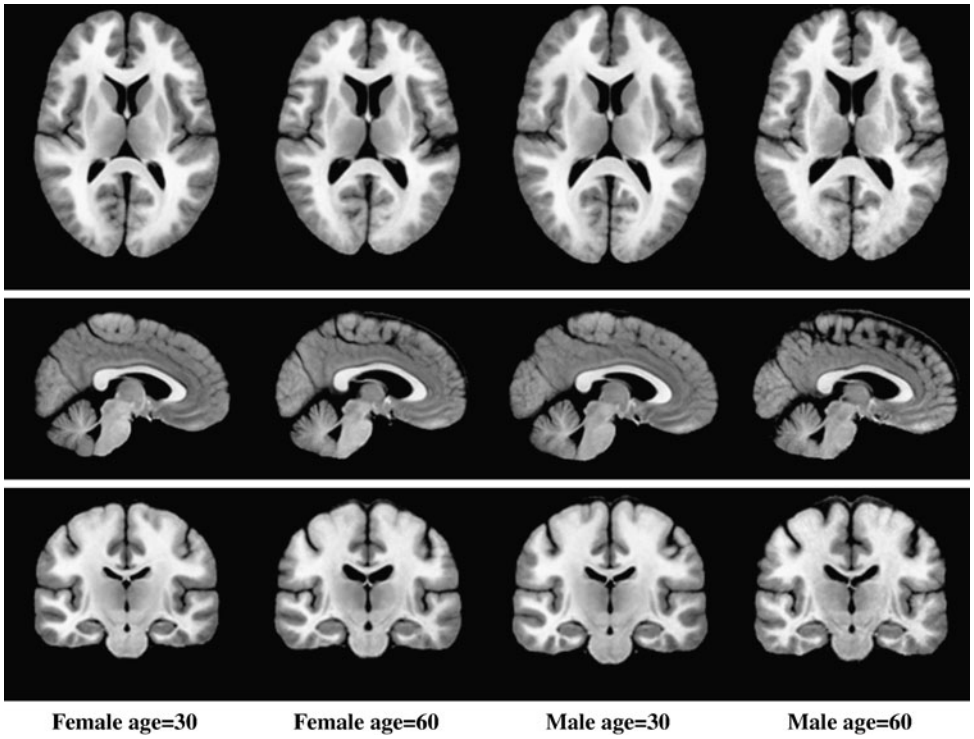


Fig. 8 Regressed brain images. Representative anatomical images for each cohort at ages 30 (*left*) and 60 (*right*). These images were generated from the random design 3D MR database using the shape regression method described in Sect. 2

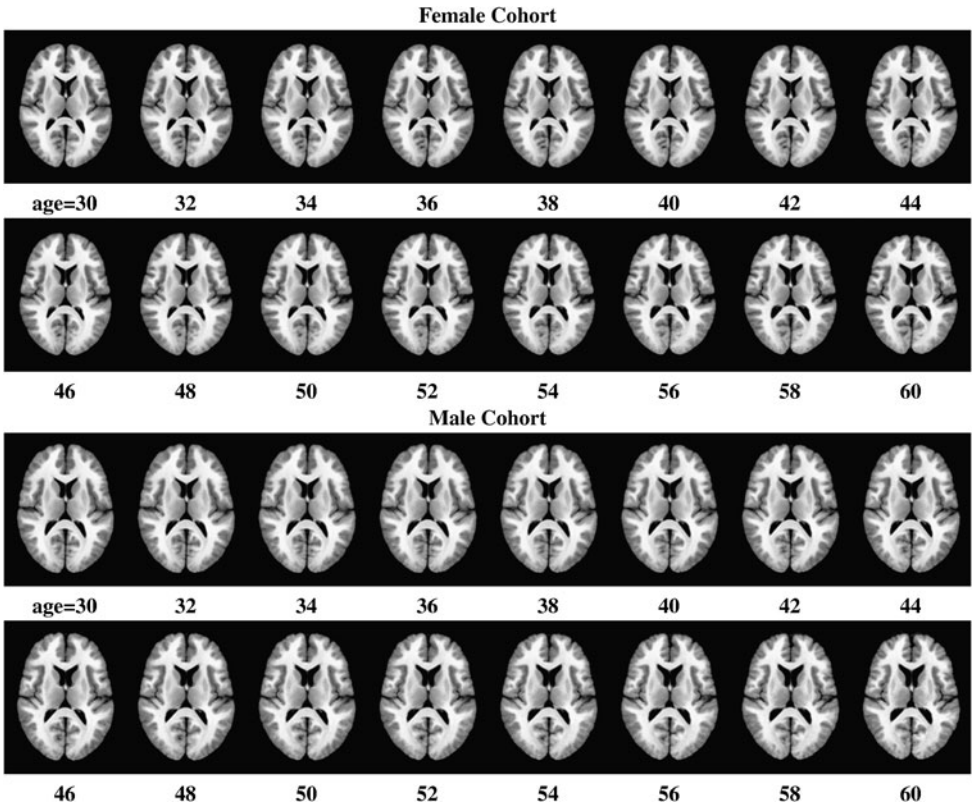


Fig. 9 The average, aging brain. These images show the average brain shape as a function of age for the female (*top*) and male (*bottom*) cohorts. These are not images from any particular patient—they are computed using the regression method proposed in this paper (18). Noticeable expansion of the lateral ventricles is clearly captured in both the image data and the determinant maps (Fig. 10). All 2D slices are extracted from the 3D volumes that were used for computation

paper we apply this method to study the effect of aging on the brain. We regress a population representative shape, indexed by age, from a database of MR brain images. Finally, we apply a longitudinal growth model to these representative images to study the detailed local shape change that occurs, on average, as a function of age.

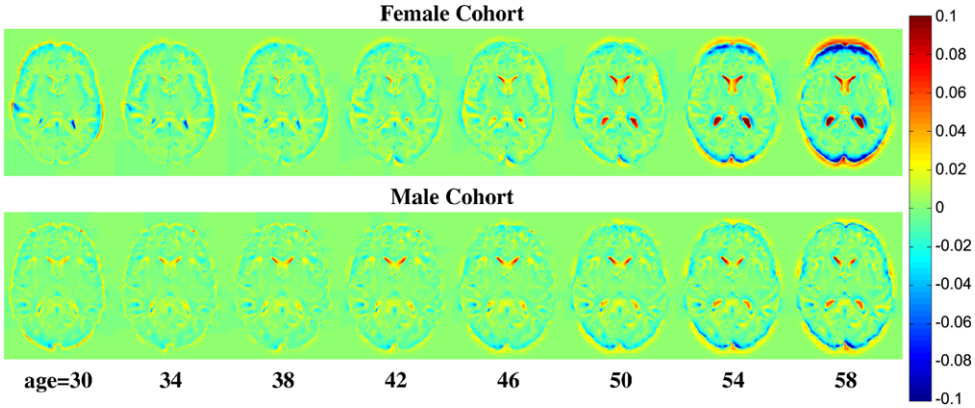


Fig. 10 (Color online) Age-indexed local expansion and contraction. Illustration of the local brain shape change as a function of age for the female (*top*) and male (*bottom*) cohorts. These data were generated by applying a diffeomorphic growth 21 model to the representative images that were computed using manifold regression. *Red voxels* indicate local expansion; *blue voxels* indicate local contraction

The regression approach presented in this paper produces purely descriptive trends of brain shape trajectories. Future work will investigate methods for quantifying these resulting trends. One question that one would like to ask is if the shape changes in a population are statistically significant. This can be answered using permutation tests, where the ages of the subjects are reassigned randomly. This tests the null hypothesis that there is no relationship between the independent variable (age) and the dependent variable (brain shape). The difficulties with this approach are designing an appropriate test statistic. Since this is a nonparametric regression method, there is no single "slope" to test. Therefore, we have to either design a statistic that tests the overall trend or use statistics that test the local slopes within particular age ranges. Another question that one would like to investigate is whether two populations have significantly different brain shape trajectories. Again, permutation tests can be used in this case, and test statistics must be developed that test the differences between the population trends.

Acknowledgements We thank Bénédicte Mortamet for providing tissue volume data and Peter Lorenzen for image preprocessing. We gratefully acknowledge our funding sources including NIH grants R01 EB000219-NIH-NIBIB, and R01 CA124608-NIH-NCI.

References

Beg, M., Miller, M., Trouvé, A., & Younes, L. (2005). Computing large deformation metric mappings via geodesic flows of diffeomorphisms. *International Journal of Computer Vision*, *61*(2).

Beg, M. F. (2003). *Variational and computational methods for flows of diffeomorphisms in image matching and growth in computational anatomy*. PhD thesis, The Johns Hopkins University.

Beg, M. F., Miller, M. I., Trouvé, A., & Younes, L. (2005). Computing large deformation metric mappings via geodesic flows of diffeomorphisms. *International Journal of Computer Vision*, *61*(2), 139–157.

Bhattacharya, R., & Patrangenaru, V. (2002). Nonparametric estimation of location and dispersion on Riemannian manifolds. *Journal of Statistical Planning and Inference*, *108*, 23–36.

Bhattacharya, R., & Patrangenaru, V. (2003). Large sample theory of intrinsic and extrinsic sample means on manifolds I. *Annals of Statistics*, *31*(1), 1–29.

Bingham, C. (1974). An antipodally symmetric distribution on the sphere. *The Annals of Statistics*, *2*(6), 1201–1225.

Buss, S. R., & Fillmore, J. P. (2001). Spherical averages and applications to spherical splines and interpolation. *ACM Transactions on Graphics*, *20*(2), 95–126.

Clatz, O., Sermesant, M., Bondiau, P. Y., Delingette, H., Warfield, S. K., Malandain, G., & Ayache, N. (2005). Realistic simulation of the 3D growth of brain tumors in MR images coupling diffusion with mass effect. *IEEE Transactions on Medical Imaging*, *24*(10), 1334–1346.

Dupuis, P., & Grenander, U. (1998). Variational problems on flows of diffeomorphisms for image matching. *Quarterly of Applied Mathematics*, *LVI*(3), 587–600.

Fletcher, P. T., Joshi, S., Ju, C., & Pizer, S. M. (2004). Principal geodesic analysis for the study of nonliner statistics of shape. *IEEE Transactions on Medical Imaging*, *23*(8), 995–1005.

Fréchet, M. (1948). Les elements aleatoires de nature quelconque dans un espace distancie. *Annales de L'Institut Henri Poincare*, *10*, 215–310.

Grenander, U., & Miller, M. I. (1998). Computational anatomy: An emerging discipline. *Quarterly of Applied Mathematics*, *56*(4), 617–694.

Guttmann, C., Jolesz, F., Kikinis, R., Killiany, R., Moss, M., Sandor, T., & Albert, M. (1998). White matter changes with normal aging. *Neurology*, *50*(4), 972–978.

Hall, P., & Marron, J. S. (1991). Local minima in cross-validation functions. *Journal of the Royal Statistical Society, Series B*, *53*(1), 245–252.

Hardle, W. (1990). *Applied nonparametric regression*. Cambridge: Cambridge University Press.

Jones, M. C., Marron, J. S., & Sheather, S. J. (1996). A brief survey of bandwidth selection for density estimation. *Journal of the American Statistical Association*, *91*(433), 401–407.

Joshi, S., Davis, B., Jomier, M., & Gerig, G. (2004). Unbiased diffeomorphic atlas construction for computational anatomy. *NeuroImage*, *23*, S151–S160. (Supplemental issue on Mathematics in Brain Imaging).

Joshi, S. C., & Miller, M. I. (2000). Landmark matching via large deformation diffeomorphisms. *IEEE Transactions on Image Processing*, *9*(8), 1357–1370.

Jupp, P., & Mardia, K. (1989). A unified view of the theory of directional statistics, 1975–1988. *International Statistical Review*, *57*(3), 261–294.

Jupp, P. E., & Kent, J. T. (1987). Fitting smooth paths to spherical data. *Applied Statistics*, *36*(1), 34–46.

Karcher, H. (1977). Riemannian center of mass and mollifier smoothing. *Communications on Pure and Applied Mathematics*, *30*, 509–541.

Kendall, D. G. (1984). Shape manifolds, Procrustean metrics, and complex projective spaces. *Bulletin of the London Mathematical Society*, *16*, 18–121.

Le, H., & Kendall, D. (1993). The Riemannian structure of Euclidean shape spaces: A novel environment for statistics. *The Annals of Statistics*, *21*(3), 1225–1271.

Leemput, K. V., Maes, F., Vandermeulen, D., & Suetens, P. (1999). Automated model-based tissue classification of mr images of the brain. *IEEE Transactions on Medical Imaging*, *18*(10), 897–908.

Loader, C. R. (1999). Bandwidth selection: Classical or plug-in? *The Annals of Statistics*, *27*(2), 415–438.

Lorenzen, P., Prastawa, M., Davis, B., Gerig, G., Bullitt, E., & Joshi, S. (2006). Multi-modal image set registration and atlas formation. *Medical Image Analysis*, *10*(3), 440–451.

Matsumae, M., Kikinis, R., Mórocz, I., Lorenzo, A., Sándor, T., Sándor, T., Albert, M., Black, P., & Jolesz, F. (1996). Age-related changes in intracranial compartment volumes in normal adults assessed by magnetic resonance imaging. *Journal of Neurosurgery*, *84*, 982–991.

Miller, M. (2004). Computational anatomy: shape, growth, and atrophy comparison via diffeomorphisms. *NeuroImage*, *23*, S19–S33.

Miller, M., & Younes, L. (2001). Group actions, homeomorphisms, and matching: A general framework. *International Journal of Computer Vision*, *41*, 61–84.

Miller, M., Banerjee, A., Christensen, G., Joshi, S., Khaneja, N., Grenander, U., & Matejic, L. (1997). Statistical methods in computational anatomy. *Statistical Methods in Medical Research*, *6*, 267–299.

Miller, M. I., Trouve, A., & Younes, L. (2002). On the metrics and Euler-Lagrange equations of computational anatomy. *Annual Review of Biomedical Engineering*, *4*, 375–405.

Mortamet, B., Zeng, D., Gerig, G., Prastawa, M., & Bullitt, E. (2005). Effects of healthy aging measured by intracranial compartment volumes using a designed mr brain database. In *Lecture notes in computer science (LNCS): Vol. 3749. Medical image computing and computer assisted intervention (MICCAI)* (pp. 383–391).

Nadaraya, E. A. (1964). On estimating regression. *Theory of Probability and its Applications*, *10*, 186–190.

Pennec, X. (2006). Intrinsic statistics on Riemannian manifolds: Basic tools for geometric measurements. *Journal of Mathematical Imaging and Vision*, *25*, 127–154.

Prastawa, M., Bullitt, E., Ho, S., & Gerig, G. (2004). A brain tumor segmentation framework based on outlier detection. *Medical Image Analysis*, *8*(3), 275–283.

Thompson, P. M., Giedd, J. N., Woods, R. P., MacDonald, D., Evans, A. C., & Toga, A. W. (2000). Growth patterns in the developing brain detected by using continuum mechanical tensor maps. *Nature*, *404*(6774), 190–193. doi:10.1038/35004593.

Trouvé, A., & Younes, L. (2005). Metamorphoses through lie group action. *Foundations of Computational Mathematics*, *5*(2), 173–198.

Wand, M. P., & Jones, M. C. (1995). *Kernel smoothing: Vol. 60. Monographs on statistics and applied probability*. London: Chapman & Hall/CRC.

Watson, G. S. (1964). Smooth regression analysis. *Sankhya*, *26*, 101–116.