Peer reviewed conference paper, final version, in print Proc. SPIE 2012, to appear Feb. 4, 2012

A framework for longitudinal data analysis via shape regression

James Fishbaugh[a], Stanley Durrleman[a], Joseph Piven[b] for the IBIS Network*, Guido Gerig[a]

[a]Scientific Computing and Imaging Institute
University of Utah, Salt Lake City, Utah

[b]Carolina Institute for Developmental Disabilities
University of North Carolina, Chapel Hill, North Carolina

ABSTRACT

Traditional longitudinal analysis begins by extracting desired clinical measurements, such as volume or head circumference, from discrete imaging data. Typically, the continuous evolution of a scalar measurement is estimated by choosing a 1D regression model, such as kernel regression or fitting a polynomial of fixed degree. This type of analysis not only leads to separate models for each measurement, but there is no clear anatomical or biological interpretation to aid in the selection of the appropriate paradigm. In this paper, we propose a consistent framework for the analysis of longitudinal data by estimating the continuous evolution of shape over time as twice differentiable flows of deformations. In contrast to 1D regression models, one model is chosen to realistically capture the growth of anatomical structures. From the continuous evolution of shape, we can simply extract any clinical measurements of interest. We demonstrate on real anatomical surfaces that volume extracted from a continuous shape evolution is consistent with a 1D regression performed on the discrete measurements. We further show how the visualization of shape progression can aid in the search for significant measurements. Finally, we present an example on a shape complex of the brain (left hemisphere, right hemisphere, cerebellum) that demonstrates a potential clinical application for our framework.

1. INTRODUCTION

During the last several years, there has been an increased emphasis on longitudinal analysis in clinical studies. Specifically, longitudinal analysis has lead to advances in our understanding of developmental disabilities such as autism,[1] neurodegenerative diseases such as Huntington's disease,[2] and neonatal-pediatric brain tissue development.[3] The framework for most longitudinal studies is as follows. Clinically relevant measurements are extracted from imaging data and a continuous evolution is estimated by fitting a regression model to the discrete measures. Typical choices for regression include kernel regression, polynomials of fixed degree, or other parametric functions such as the logistic[4] or Gompertz function.[5] Any further statistical analysis is conducted using the trajectories (or parameters) estimated during regression.

In this work, we present a framework for longitudinal analysis centered around the estimation of *shape* evolution. Rather than extracting measurements from discrete data, we propose modeling the continuous evolution of one or more anatomical shapes. From the resulting growth scenario, we can simply extract any measurements of interest. We model the evolution of biological tissue over time as the twice differentiable flow of deformations, guaranteeing the estimated growth is smooth in both space and time. This growth model is chosen to realistically capture the growth of anatomical surfaces, whereas there is no clear anatomical or biological interpretation of

* The IBIS Network. Clinical Sites: University of North Carolina: J. Piven (IBIS Network PI), H.C. Hazlett, C. Chappell; University of Washington: S. Dager, A. Estes, D. Shaw; Washington University: K. Botteron, R. McKinstry, J. Constantino, J. Pruett; Childrens Hospital of Philadelphia: R. Schultz, S. Paterson; University of Alberta: L. Zwaigenbaum; Data Coordinating Center: Montreal Neurological Institute: A.C. Evans, D.L. Collins, G.B. Pike, P. Kostopoulos; Samir Das; Image Processing Core: University of Utah: G. Gerig; University of North Carolina: M. Styner; Statistical Analysis Core: University of North Carolina: H. Gu; Genetics Analysis Core: University of North Carolina: P. Sullivan, F. Wright.

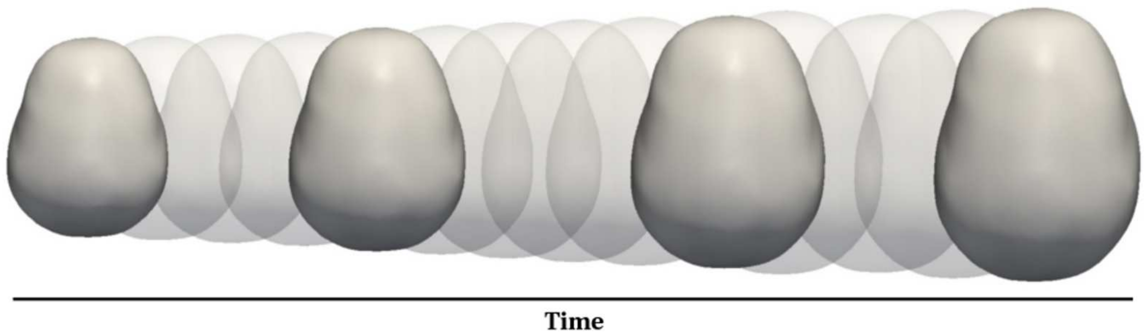

Figure 1. An illustration of shape regression. Here we have four observations of the intracranial surface over time, shown as solid surfaces. The objective of shape regression is to infer the continuous evolution of shape (transparent surfaces) which best explains the observed data.

common 1D regression models. Furthermore, our framework involves the selection of only one regression model, in contrast to traditional longitudinal studies that may involve separate models for each measurement.

We demonstrate on real anatomical surfaces extracted from a longitudinal database the power and flexibility of our proposed framework. Two case studies are presented as a proof of concept. First, a subject specific analysis is explored by estimating continuous shape evolution for each subject independently. We show how viewing the evolution as a movie is a valuable data exploration tool. Finally, a group analysis is conducted by comparing average growth scenarios for each population using a bootstrap procedure.

2. SHAPE REGRESSION

Shape regression involves inferring the continuous evolution of shape to closely match a set of target shapes over time, illustrated in Figure 1. Here we consider shape in a general sense, represented as point sets, curves, or surfaces. The problem is often posed as the trade off between fidelity to data and regularity, with the most likely shape evolution estimated based on a regularized least-square criterion.

Several shape regression algorithms have been proposed, such as the extension of kernel regression to general manifolds.[6] Also, large deformation diffeomorphic metric mapping (LDDMM) registration has been extended to time-series data.[7,8] The method can be considered piecewise-linear regression in the space of diffeomorphisms, and is commonly referred to as piecewise-geodesic regression. Similarly, linear regression has been extended to geodesic regression for image time-series[9] and general manifold spaces.[10] A stochastic growth model based on perturbations from geodesic paths has been proposed, demonstrating better interpolation on several synthetic experiments on 2D landmarks, as compared to piecewise-geodesic regression.[11]

Recently, an acceleration controlled growth model based on the twice differentiable flow of deformations was proposed under the assumption that anatomical growth is temporally smooth, but doesn't necessarily follow a geodesic path, or close to a geodesic path.[12] Experiments on real anatomical surfaces demonstrated improved interpolation properties and robustness to noise, as compared to piecewise-geodesic regression. Due to these considerations, we have selected acceleration controlled regression for our framework. However, any of the regression methods discussed would be viable options.

3. METHOD

Here we briefly summarize acceleration controlled regression. Consider a set of shapes S_{t_i}, observed at times t_i. Shape evolution is modeled by continuously deforming S_0 in order to match the remaining target shapes. The trajectories of the shape points over time, written as $x_i(t)$, are determined by the integration of the 2nd-order ODE

$$\ddot{\phi}_t(x_i(t)) = a(x_i(t), t) \tag{1}$$

where ϕ_t is the time varying deformation and a is a smooth, time varying acceleration field defined at the location of the shape points, which can be considered an indication of the underlying forces acting on the anatomy.

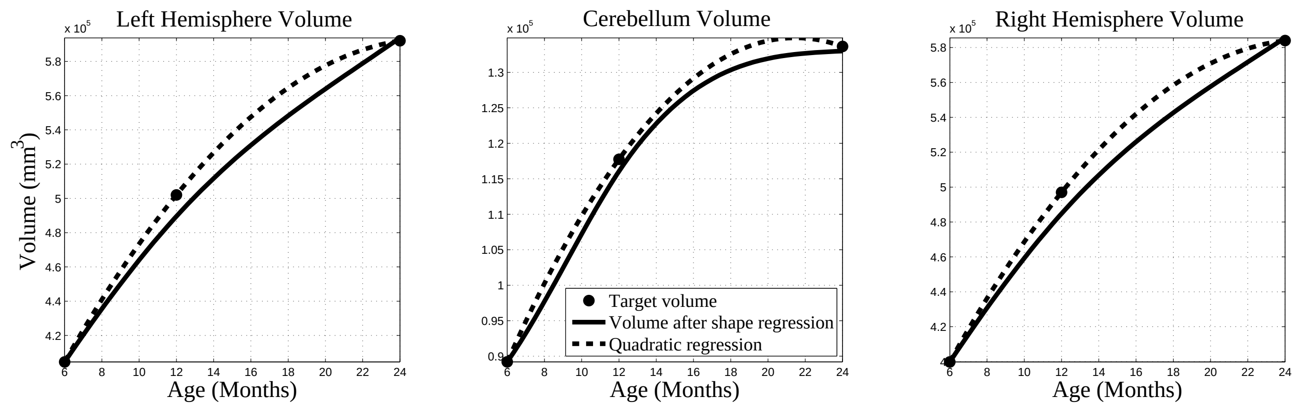

Figure 2. Volume measurements extracted after shape regression compared with quadratic regression on the discrete volume measurements. The evolution of shape was estimated jointly on all shapes whereas three independent curves were estimated in the 1D regression case.

The method seeks to find optimal point forces that drive particles on the anatomical surface through space to match target data, with the resulting trajectories being twice differentiable.

The continuous evolution of shape is then the starting point for any further analysis. From here, the shape evolution can be viewed as an animation, allowing for the investigation of potentially significant measures. Any desired measurement can simply be extracted from the collection of shapes. As we will show in the next section, this framework allows for either individual or group based analysis.

4. EXPERIMENTS

4.1 Dataset and Preprocessing

To demonstrate the application of our framework, we use a longitudinal database of 14 control children and 12 children with positive ADOS score (likely to be diagnosed with autism spectrum disorder or autism), from here on referred to as the autism group. Each child has been scanned three times, at approximately 6, 12, and 24 months old. The 6 and 12 month images are first rigidly co-registered to the 24 month image using IRTK.[13] Next, the left hemisphere, right hemisphere, and cerebellum are segmented via deformable registration with a template. Triangular meshes are extracted via marching cubes and are simplified and smoothed, resulting in a quality mesh for each subject. Finally, the 6 and 12 month shape complexes are rigidly co-registered to the 24 month shape complex using `gmmreg`[14] to further reduce the amount of translation and rotation between time points.

4.2 Subject Specific Shape Regression

A growth scenario is jointly estimated for each subject's left hemisphere, right hemisphere, and cerebellum. We set the standard deviation of the Gaussian kernel λ_V controlling the deformation to 10 mm in order to capture very detailed shape changes. For λ_W, the scale at which shape differences are considered noise, we choose 4 mm for the hemispheres and 2 mm for the cerebellum. We weight regularity by 0.01 and discretize time into increments of 0.37 months.

First, we investigate the application of extracting scalar measurements from shape regression, as the framework relies on the trajectories of these measurements being realistic. Figure 2 shows volume from one particular control subject extracted after shape regression as well as quadratic regression estimated using the sparse volume measurements. For the sparse measurements, separate 1D regressions were computed for each shape. In contrast, our shape regression paradigm provides a consistent framework for dealing with multiple shapes as well as multiple measurements. The regression only needs to be estimated *once*. Additionally, because the estimation is done jointly on several shapes, it incorporates potentially important spatial relationships between shapes. The decrease in cerebellum volume at 22 months estimated by quadratic regression highlights the downside of using such models, as it is unlikely the true anatomy decreased in volume.

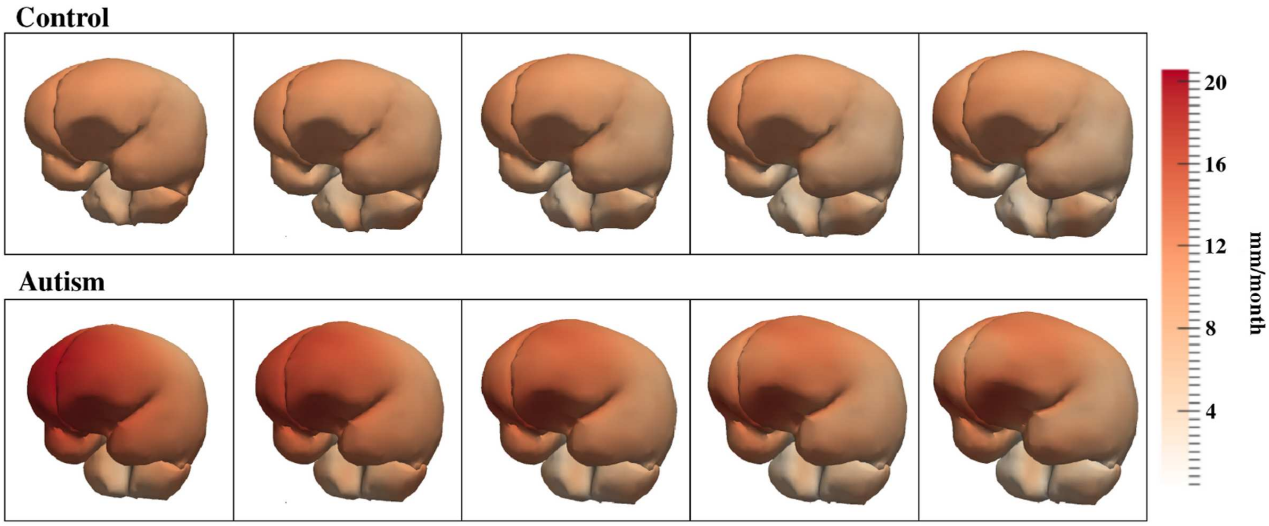

Figure 3. Snapshots of shape evolution from 6 to 24 months (left to right) for one control (**top**) and one subject from the autism group (**bottom**) with color denoting magnitude of velocity. The evolution of the control shows a more isotropic change of scale with respect to the subject from the autism group, who is growing much faster along the anterior/posterior axis.

In addition to volume, circumference, and other common clinical measurements, other less obvious features may be of significance. From the multitude of available measurements, its is difficult or impossible to determine which landmarks are most salient from the image data alone. In contrast, the visualization of shape evolution is a powerful exploratory tool which allows a researcher to quickly and intuitively explore potentially significant measurements. Several snapshots of shape evolution for one control and one child from the autism group is shown in Figure 3. This visualization makes it clear there is a difference in the way the brains of these two children develops. The child from the autism group experiences considerably faster growth as the forebrain expands outward while the control undergoes a mostly isotropic change in scale over time. Please note that this serves only as an illustration of our shape regression framework and should not be interpreted as a significant finding.

From this observation, we investigate the evolution in the direction of the three major axes of the brain. These measurements are easily extracted by computing the length, width, and height of the bounding box for each shape complex over time. A 'brain-box' analysis of one control and one child from the autism group (the same two children as the previous example) is shown in Figure 4, which shows the normalized distance over time in the three major directions. First, we consider growth at the earliest age, from 6 to 10 months old.

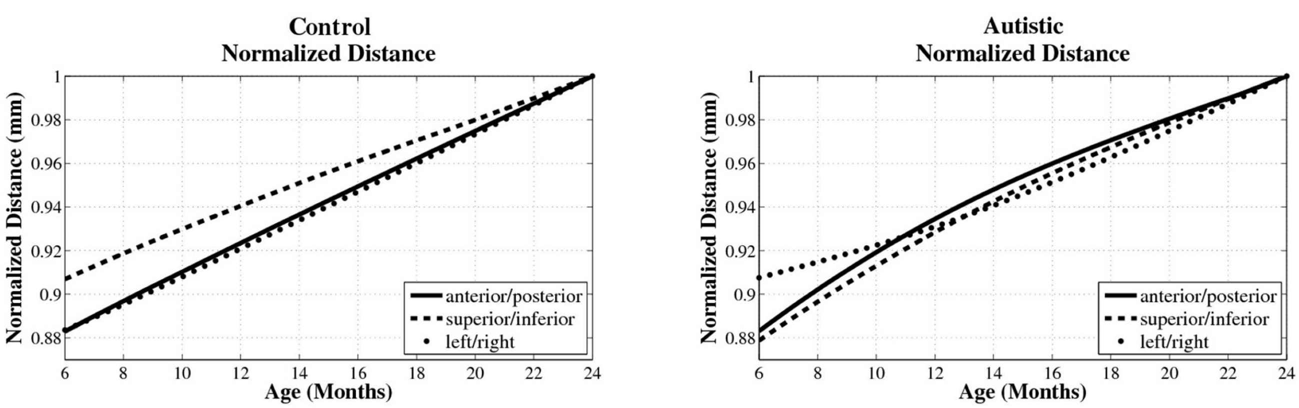

Figure 4. Evolution of the distance along the three major axes of the brain for a control (**left**) and a child from the autism group (**right**).

The slopes of the curves for the control over this time period are all approximately 0.012, confirming that the evolution is mostly isotropic. In contrast, the child from the autism group has an initial slope of 0.018 for the anterior/posterior and superior/inferior axes, but a slope of 0.007 for the left/right axis. This corresponds to accelerated initial growth as the forebrain expands outward and upward, growing 2.5 times faster than along the left/right axis.

Though the child from the autism group undergoes anisotropic growth initially, it appears to stabilize after 12 months old. From 12 to 24 months, the speed of growth is roughly the same in all directions. Furthermore, the control subject shows nearly linear growth in all directions over the entire time interval, with no major change after 12 months old. In contrast, the child from the autism group shows non-linear growth, as the anterior/posterior speed decreases over time while the left/right speed increases. This is an interesting finding that provides an avenue for further research in a full scale clinical study.

4.3 Population Based Shape Regression

In addition to estimating shape evolution for each subject independently, we can also consider one regression for each population or group. In this case, the regression takes into account all data from all subjects simultaneously. The resulting evolution is a mean growth scenario which best represents the population, which is also commonly referred to as an average or an atlas.

4.3.1 Bootstrapping

We consider two populations as before, 14 controls and 12 children from the autism group. The bootstrapping procedure involves sampling each group, with replacement, creating a new data set with the same number of samples as the original. A mean scenario is then estimated for each group based on the new samples, and the process is repeated 100 times to simulate the variability within each group. In the end, we have 100 average growth scenarios for each population.

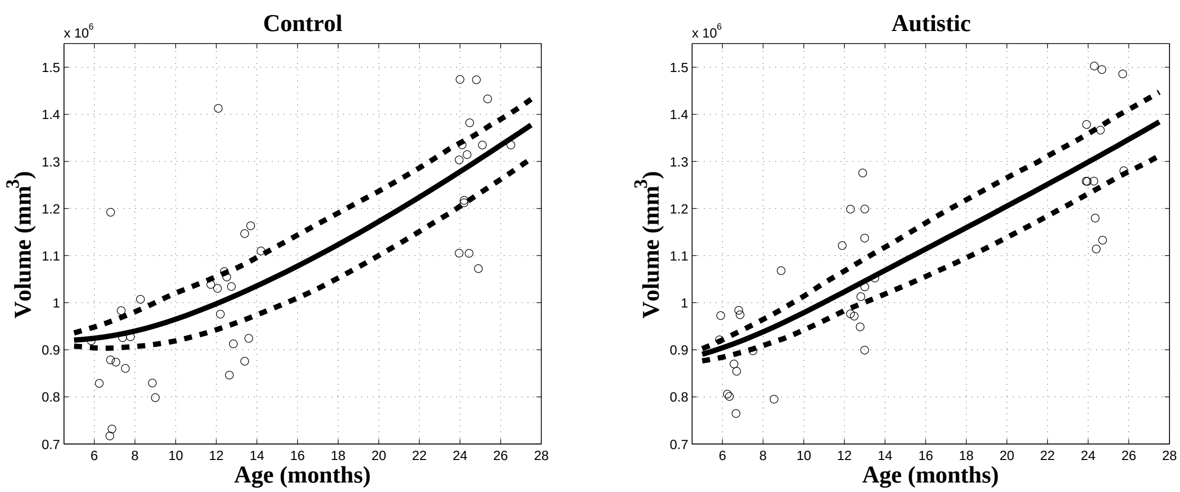

Figure 5. Bootstrap 90% confidence intervals of volume for control and autism group, measured after shape regression. Small circles represent the volume of the target shapes.

From the mean growth scenarios, we extract 90% confidence intervals of volume, by discarding the largest and smallest 5%, shown in Figure 5. We observe that there is a large amount of variability in volume present in both groups. The intervals overlap considerably, showing there is no significant difference in volume between the two populations. However, there is an interesting difference in the initial rate of volume growth between the two groups, as the volume in the autism group increases faster than the controls. As with the individual regression analysis, this implies that the period of time around 6 months old is an important direction for future study. Additionally, this experiment provides further evidence that rate of growth may be a more relevant measure than size.

5. CONCLUSION

In this paper, we have introduced a framework for the analysis of longitudinal data based on shape regression. Rather than fitting curves to scalar measurements extracted from imaging data, we advocate modeling the evolution of the shapes of interest. After shape regression, any number of measurements can simply be extracted from the shapes continuously. In contrast to traditional longitudinal studies of scalar measures, shape regression provides a generic and flexible framework that allows for consistent treatment of multiple measurements and multiple shapes simultaneously. The growth model in our framework was chosen specifically to realistically capture the evolution of anatomical shapes. Furthermore, shape regression takes into account the spatial relationships between shapes that are essentially ignored in traditional regression. As demonstrated in the experiments, the extracted measurements may be either linear or non-linear, with no prior constraint on linearity.

We have demonstrated on real anatomical surfaces that volume measurements derived from shape evolution are compatible with a 1D regression of volume. The visualization of shape evolution was shown to improve data exploration, highlighting significant measurements that might have otherwise been overlooked. Two case studies were explored, showing the application of shape regression to both individual and population based analysis. Future work will focus on applying this framework to a statistically rigorous clinical study.

Acknowledgments

This work was supported by NIH grant RO1 HD055741(ACE, project IBIS) and by NIH grant U54 EB005149 (NA-MIC).

REFERENCES

[1] H. Hazlett, M. Poe, G. Gerig, M. Styner, C. Chappell, R. Smith, C. Vachet, and J. Piven, "Early brain overgrowth in autism associated with an increase in cortical surface area before age 2 years," *Arch Gen Psychiatry* **68**, pp. 467–76, May 2011.

[2] S. Tabrizi, D. Langbehn, B. Leavitt, R. Roos, A. Durr, D. Craufurd, C. Kennard, S. Hicks, N. Fox, R. Scahill, B. Borowsky, A. Tobin, H. Rosas, H. Johnson, R. Reilmann, B. Landwehrmeyer, and J. Stout, "Biological and clinical manifestations of huntington's disease in the longitudinal track-hd study: cross-sectional analysis of baseline data.," *Lancet Neurol* , 2009.

[3] S. Xu, M. Styner, J. Gilmore, and G. Gerig, "Multivariate longitudinal statistics for neonatal-pediatric brain tissue development," in *Proc. of SPIE Medical Imaging 2008*, **6914**, 2008.

[4] S. Xu, M. Styner, J. Gilmore, J. Piven, and G. Gerig, "Multivariate nonlinear mixed model to analyze longitudinal image data: Mri study of early brain development," in *Computer Vision and Pattern Recognition Workshops*, June 2008.

[5] N. Sadeghi, M. Prastawa, J. Gilmore, W. Lin, and G. Gerig, "Spatio-temporal analysis of early brain development," in *Proc. IEEE Asilomar Conference on Signals, Systems and Computers*, 2010.

[6] B. Davis, P. Fletcher, E. Bullitt, and S. Joshi, "Population shape regression from random design data," in *Proc. of ICCV*, pp. 1–7, Oct. 2007.

[7] S. Durrleman, X. Pennec, A. Trouvé, G. Gerig, and N. Ayache, "Spatiotemporal atlas estimation for developmental delay detection in longitudinal datasets," in *Proc. of the 14th international conference on Medical image computing and computer-assisted intervention*, G.-Z. Yang, D. Hawkes, D. Rueckert, A. Noble, and C. Taylor, eds., *LNCS* **5761**, pp. 297–304, Springer, 2009.

[8] A. Khan and M. Beg, "Representation of time-varying shapes in the large deformation diffeomorphic framework," in *Proc. of ISBI*, pp. 1521–1524, 2008.

[9] M. Niethammer, Y. Huang, and F.-X. Vialard, "Geodesic regression for image time-series," in *Proc. of the 14th international conference on Medical image computing and computer-assisted intervention*, G. Fichtinger and T. Peters, eds., *LNCS* **6982**, pp. 655–662, Springer-Verlag, (Berlin, Heidelberg), 2011.

[10] T. Fletcher, "Geodesic Regression on Riemannian Manifolds," in *Proc. of the Third International Workshop on Mathematical Foundations of Computational Anatomy - Geometrical and Statistical Methods for Modelling Biological Shape Variability*, X. Pennec, S. Joshi, and M. Nielsen, eds., pp. 75–86, (Toronto, Canada), Sept. 2011.

[11] F. Vialard and A. Trouvé, "Shape splines and stochastic shape evolutions: A second-order point of view," *Quarterly of Applied Mathematics* , 2010.

[12] J. Fishbaugh, S. Durrleman, and G. Gerig, "Estimation of smooth growth trajectories with controlled acceleration from time series shape data," in *Proc. of the 14th international conference on Medical image computing and computer-assisted intervention*, G. Fichtinger and T. Peters, eds., *LNCS* **6982**, pp. 401–408, Springer, 2011.

[13] J. A. Schnabel, D. Rueckert, M. Quist, J. M. Blackall, A. D. Castellano-Smith, T. Hartkens, G. P. Penney, W. A. Hall, H. Liu, C. L. Truwit, F. A. Gerritsen, D. L. G. Hill, and D. J. Hawkes, "A generic framework for non-rigid registration based on non-uniform multi-level free-form deformations," in *Proceedings of the 4th International Conference on Medical Image Computing and Computer-Assisted Intervention*, *MICCAI '01*, pp. 573–581, Springer-Verlag, 2001.

[14] B. Jian and B. C. Vemuri, "A robust algorithm for point set registration using mixture of gaussians," in *Proc. of the Tenth IEEE International Conference on Computer Vision*, *ICCV '05*, pp. 1246–1251, IEEE Computer Society, 2005.