

Hierarchical Pointset-Based Statistical Shape Modeling and Applications

Ph.D. Thesis
Submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy
by

Saurabh Shigwan
(Roll No. 144050001)



9 August 2020

Acceptance Certificate

**Department of Computer Science and Engineering
Indian Institute of Technology, Bombay**

The PhD Thesis entitled “Hierarchical Pointset-Based Statistical Shape Modeling and Applications” submitted by Saurabh Shigwan (Roll No. 144050001) may be accepted for being evaluated.

Date: 9 August 2020

Declaration

I declare that this written submission represents my ideas in my own words and where others' ideas or words have been included, I have adequately cited and referenced the original sources. I declare that I have properly and accurately acknowledged all sources used in the production of this report. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be a cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Saurabh Shigwan

(Roll No. 144050001)

Date:

Acknowledgements

Foremost, I would like to express the sincere gratitude to my advisor Prof. Suyash P. Awate (CSE, IITB) for the continuous support in the Ph.D study and related research, for his patience, motivation, and immense knowledge. Besides my advisor, I would like to thank the RPC committee members Prof. Ganesh Ramakrishnan (CSE, IIT Bombay), Prof. Shabbir Merchant(EE, IIT Bombay) and Prof. Siddhartha Chaudhuri(CSE,IIT Bombay) for their insightful comments and encouragement, and also for the valuable questions and suggestions during APS which helped to widen my research from various perspectives. I am honored to be part of Vision Graphics and Imaging Lab(ViGIL). I give special thanks to my coauthor and labmate Akshay Gaikwad for sharing his work in shape model and segmentation for my research work. I also thank my fellow labmates: Nitin Kumar, Preeti Gopal, Kratika Gupta, Rahul Mitra, Meghshyam Prasad, Ravi Chalasani, Pratik Kalshetti, Rohit Saluja, PS Viswanath and Jerin Jecob, for the stimulating discussions and help. I would also like to thank office staff from the CSE department for their support. I am also grateful for funding from MHRD and IIT Bombay Seed Grant 14IR-CCSG010. Finally, I would like to thank my parents, for their patience with me that almost never wore out.

Saurabh Shigwan

IIT Bombay

Abstract

Statistical shape models play a key role in medical image analysis, e.g., for hypothesis testing, object segmentation and shape clustering. Several approaches to shape modeling, e.g., diffeomorphic deformation models, medial models, point distribution models, etc., are active research areas in their own right, each with its benefits and limitations. Challenges to current model learning schemes stem from many factors including (i) inevitable *human errors in segmentation*, where the individual shape models may overfit to the errors and lead to inflated group variability, and (ii) *low sample sizes for groups*, where the group distribution model may overfit and incorrectly indicate reduced group variability. We propose to address these issues, in the context of learning point distribution models, through (i) a novel *generative model* of the data that *accounts for human errors in segmentations*, (ii) a novel *hierarchical model that regularizes group-distribution fits* using population-level shape variables, and (iii) Riemannian analysis in Kendall shape space. We propose a novel method for *sampling* shapes from distributions in Kendall shape space, and leverage that for inference using Monte-Carlo expectation maximization. We propose a deep neural net based object segmentation incorporating a novel *Riemannian shape prior* with a novel inference method. We also propose to address these issues, in context of learning a multi-modal point distribution model for shape clustering, through (i) a novel *generative mixture model* of the data that *accounts for human errors in segmentations*, (ii) a novel *hierarchical mixture model that regularizes group-distribution fits* using population-level shape variables, and (iii) Riemannian analysis in Kendall shape space. The results show some benefits of our framework in hypothesis testing, object segmentation and shape clustering on simulated and clinical datasets.

Table of Contents

Acknowledgements	iii
Abstract	iv
List of Figures	vii
1 Introduction	1
1.1 Shape Modeling Approaches and Applications	2
1.2 Contributions	4
1.3 Organization of Thesis	5
2 Background and Related Work	6
2.1 Shape Modeling with Diffeomorphism	6
2.2 Shape Modeling with Medial Models	7
2.3 Shape Modeling with Mixture Models	7
2.4 Application: Hypothesis Testing using Shape distribution	8
2.5 Application: Object Segmentation	8
2.5.1 Non-Parametric Shape Models: Deep Boltzmann Machine, Deep Neural Network	9
2.6 Application: Shape Clustering	10
3 A Hierarchical Generative Shape Model for Multigroup Shape Data	11
3.1 Mathematical Notation	11
3.2 Modeling Smooth Shapes	12
3.3 Modeling Label-Map Likelihoods given Shapes	13
3.4 Our Hierarchical Statistical Shape Model	13
3.5 Hierarchical Model Inference using Monte-Carlo Expectation Maximization	14
3.6 Robust Efficient MCMC Sampling on Riemannian Manifolds	16

4 Hierarchical Riemannian Shape Modeling for Hypothesis Testing	18
4.1 Validation: Benefits of Hierarchical Modeling	19
4.2 Results: Hypothesis Testing	20
4.2.1 Results: Hypothesis Testing on Simulated Ellipsoidal Data	21
4.2.2 Results: Hypothesis Testing on Carpal Bones Data	24
5 Riemannian Shape Prior for Bayesian Image Segmentation	30
5.1 Object-Shape Representation and Segmentation Prior	30
5.2 Generative Model for Observed Image Data	31
5.3 Bayesian Formulation for Object Segmentation	32
5.4 Optimization Algorithm	32
5.5 Results: Segmentation	32
5.5.1 Results: Segmentation on Simulated MRI data	32
5.5.2 Results: Segmentation in Clinical Brain Subcortical MRI	35
6 Hierarchical Shape Clustering in Shape Space	43
6.1 Mathematical Notation	43
6.2 Likelihood model	44
6.3 Model fitting	45
6.4 Results: Riemannian Mixture Shape Clustering	47
6.4.1 Results: Shape Clustering on Simulated Ellipsoidal Data	48
7 Conclusion	50
A Appendices	52
A.1 Gradients of the Log-Likelihood With Respect to Similarity Transform Parameters	52
A.2 Gradients of Mahalanobis distance in Riemannian Space	52
A.3 Approximating the Normalizing Factor in Riemannian Space	54
References	55
List of Publications	62

List of Figures

3.1	Hierarchical Model	14
4.1	Benefits of Hierarchical Modeling. (a) <i>Regularizing Effect of Population-Level Variables:</i> Distribution of geodesic distances between estimated and true group means, with varying strengths α (see text) of the population-level prior. (b) <i>Reducing False Positives in Hypothesis Testing (Shrinkage Effect):</i> Distribution of the test statistic between two group samples drawn from the same shape distribution, with varying strengths α (see text) of the population-level prior.	20
4.2	Results: Hypothesis Testing on Simulated Ellipsoidal Data. (a) Ground truth segmentation. (b) Corrupted segmentation (data), mimicking human errors. (c)-(d) Our method's MAP estimates for group means z_1, z_2 . (e)-(f) ShapeWorks's estimates for group means z_1, z_2	21
4.3	Results: Hypothesis Testing on Simulated Ellipsoidal Data. Estimated principal mode of variation of the first group with mean z_1 for our method and ShapeWorks, with λ_{11} as first principal eigenvalue and v_{11} as first principal eigenvector of group covariance C_1	21
4.4	Results: Hypothesis Testing on Simulated Ellipsoidal Data. Estimated principal mode of variation of the second group with mean z_2 for our method and ShapeWorks, with λ_{21} as first principal eigenvalue and v_{21} as first principal eigenvector of group covariance C_2	22

4.5 Results: Hypothesis Testing on Simulated Ellipsoidal Data. Compactness of point correspondences across fitted shapes in one group, indicated pointwise by the variability in the locations of each set of corresponding points (i.e., $\{y_{mij}\}_{i=1}^{N_m}$) across the group, where the variability measure is the sum of the smallest two eigenvalues of the 3D covariance of each point on the mesh. Each mesh-point's variability in location (y_{mij}) across the group (i.e., over all shapes $i = 1, \dots, N_m$) for (a) our model and (b) ShapeWorks, visualized on the estimated mean shape (i.e., z_m) for each group. (c)-(d) Eigenspectra for group covariances C_1 and C_2 ; error bars show variation across multiple experiments with different corruption instances.	22
4.6 Results: Hypothesis Testing on Simulated Ellipsoidal Data. (a)-(b) Sampled individual shapes y_{mi}^s from our method. (c)-(d) MAP estimates of individual shapes y_{mi} from our method. (e)-(f) Estimated shapes of individuals from ShapeWorks.	22
4.7 Results: Hypothesis Testing on Simulated Ellipsoidal Data. Cohen's d effect size between the two group distributions based on estimates of group means z_1 and z_2 and covariances C_1 and C_2 , computed per mesh point j based on the j -th components of the means and the j -th diagonal elements of the covariance matrices, and visualized on the population mean μ , for: (a) ground truth data (large sample size; 2000 per group) and (b)-(c) our method and ShapeWorks, respectively (small sample size, 36 per group). Permutation tests showing the histogram of the test statistic under the null hypothesis and the unpermuted test-statistic values (error bars represent results from repeated experiments) for: (d) ground truth data (large sample size; 2000 per group) and (e)-(f) our method and ShapeWorks, respectively (small sample size, 36 per group).	23
4.8 Results: Hypothesis Testing on Carpal Bones - Hamate. (a) Ground truth segmentation. (b) Corrupted segmentation (data), mimicking human errors. (c)-(d) Our method's MAP estimates for group means z_1, z_2 . (e)-(f) ShapeWorks's estimates for group means z_1, z_2	24
4.9 Results: Hypothesis Testing on Carpal Bones - Hamate. Estimated principal mode of variation of the first group with mean z_1 for our method and ShapeWorks, with λ_{11} as first principal eigenvalue and v_{11} as first principal eigenvector of group covariance C_1	25

4.10 Results: Hypothesis Testing on Carpal Bones - Hamate. Estimated principal mode of variation of the second group with mean z_2 for our method and ShapeWorks, with λ_{21} as first principal eigenvalue and v_{21} as first principal eigenvector of group covariance C_2	25
4.11 Results: Hypothesis Testing on Carpal Bones - Hamate. Compactness of point correspondences across fitted shapes in one group, indicated pointwise by the variability in the locations of each set of corresponding points (i.e., $\{y_{mij}\}_{i=1}^{N_m}$) across the group, where the variability measure is the sum of the smallest two eigenvalues of the 3D covariance of each point on the mesh. Each mesh-point's variability in location (y_{mij}) across the group (i.e., over all shapes $i = 1, \dots, N_m$) for (a) our model and (b) ShapeWorks, visualized on the estimated mean shape (i.e., z_m) for each group. (c)-(d) Eigenspectra for group covariances C_1 and C_2 ; error bars show variation across multiple experiments with different corruption instances.	26
4.12 Results: Hypothesis Testing on Carpal Bones - Hamate. (a)-(b) Sampled individual shapes y_{mi}^s from our method. (c)-(d) MAP estimates of individual shapes y_{mi} from our method. (e)-(f) Estimated shapes of individuals from ShapeWorks.	26
4.13 Results: Hypothesis Testing on Carpal Bones - Hamate. Cohen's d effect size between the two group distributions based on estimates of group means z_1 and z_2 and covariances C_1 and C_2 , computed per mesh point i based on the i -th components of the means and the i -th diagonal elements of the covariance matrices, and visualized on the population mean μ , for: (a) our method and (b) ShapeWorks. Permutation tests showing the histogram of the test statistic under the null hypothesis and the unpermuted test-statistic values (error bars represent results from repeated experiments) for: (c) our method and (d) ShapeWorks.	26
4.14 Results: Hypothesis Testing on Carpal Bones - Hamate. (a)-(b) Ground truth object segmentations of an example male and female bones, respectively. (c)-(d) Our method's MAP estimates for the bone shapes in (a) and (b), respectively. (e)-(f) ShapeWorks's estimates for the bone shapes in (a) and (b), respectively, <i>when ShapeWorks additionally smooths the corrupted data to get rid of the added bumps.</i>	27
4.15 Results: Hypothesis Testing on Carpal Bones - Trapezoid. Captions as in Figure 4.13.	27

4.16	Results: Hypothesis Testing on Carpal Bones - Capitate. Captions as in Figure 4.13.	27
4.17	Results: Hypothesis Testing on Carpal Bones - Lunate. Captions as in Figure 4.13.	27
4.18	Results: Hypothesis Testing on Carpal Bones - Pisiform. Captions as in Figure 4.13.	28
4.19	Results: Hypothesis Testing on Carpal Bones - Scaphoid. Captions as in Figure 4.13.	28
4.20	Results: Hypothesis Testing on Carpal Bones - Trapezium. Captions as in Figure 4.13.	28
4.21	Results: Hypothesis Testing on Carpal Bones - Triquetrum. Captions as in Figure 4.13.	28
5.1	Results: Segmentation on Simulated Data. (a) A 2D slice of a simulated 3D image. (b) Expert segmentation. Box plots showing methods' performances across dataset, measured by (c) Dice similarity coefficient and (d) average of the shortest inter-surface distances between each point on estimated object surface and ground-truth. Estimated 3D segmentation's surface rendering, with each surface location colored by the shortest signed distance to ground-truth surface (in terms of percentage of ground-truth diameter), for: (e) our method, (f) ShapeWorks (SW) shape model, (g) multitalas (MA) segmentation with nonlinear nonparametric registration, and (h) ground-truth segmentation.	34
5.2	Results: Segmentation on Carpel Bone Shapes: Trapezoid. Captions as in Figure 5.1.	35
5.3	Results: Segmentation on Carpel Bone Shapes: Hamate. Captions as in Figure 5.1.	35
5.4	Results: Segmentation on Carpel Bone Shapes: Pisiform. Captions as in Figure 5.1.	36

5.5 Results on Clinical Brain MRI: Subcortical Structure Thalamus.	
(a) A 2D slice of the 3D test image. Box plots showing methods' performances across dataset, measured by (b) Dice similarity coefficient, (c) mean of the shortest inter-surface distances between each point on estimated object surface and ground-truth. (d) High-quality ground-truth segmentation. Estimated segmentation's surface rendering, with each surface location colored by the shortest signed distance to ground-truth surface (in terms of percentage of ground-truth diameter), for (e) Our method; (f) Unet (Ronneberger et al. (2015)); (g) Shape-regularized Unet (Ravishankar et al. (2017)), extended to 3D; (h) Unet + ShapeWorks prior model; (i) Multiatlas segmentation.	37
5.6 Results on Clinical Brain MRI: Subcortical Structure Caudate.	
Captions (a)–(i) same as in Figure 5.5. (j) Segmentation with Unet+SWs, where ShapeWorks's prior learning extra smooths segmentations in low-quality training set.	38
5.7 Results on Clinical Brain MRI: Subcortical Structure Globus.	
Captions (a)–(i) same as in Figure 5.5.	39
5.8 Results on Clinical Brain MRI: Subcortical Structure Hippocampus.	
Captions (a)–(i) same as in Figure 5.5.	40
5.9 Results on Clinical Brain MRI: Subcortical Structure Putamen.	
Captions (a)–(i) same as in Figure 5.5.	41
6.1 Hierarchical Model	45
6.2 Results: Shape Clustering on Simulated Ellipsoidal Data. (a) Ground truth segmentation. (b) Corrupted segmentation (data), mimicking human errors. (c) Accuracy boxplot over 5 noisy instances of ellipsoid data for <i>Ours</i> and VBMixPCA. (d) Estimated population cluster mean from our method.	48
6.3 Results: Shape Clustering on Simulated Ellipsoidal Data. (a)-(c) Our method's MAP estimates for group means z_1 , z_2 and z_3 . (d)-(f) VBMixPCA's estimates for group means z_1 , z_2 and z_3	48
6.4 Results: Shape Clustering on Simulated Ellipsoidal Data. Estimated principal mode of variation of group-1 for our method and VBMixPCA. .	49
6.5 Results: Shape Clustering on Simulated Ellipsoidal Data. Estimated principal mode of variation of group-2 for our method and VBMixPCA. .	49
6.6 Results: Shape Clustering on Simulated Ellipsoidal Data. Estimated principal mode of variation of group-3 for our method and VBMixPCA. .	49

A.1 Difference between the Riemannian distance and the Euclidean distance as a function of the arc length in radians.	53
--	----

Chapter 1

Introduction

Statistical shape models play a key role in a variety of applications in medical image analysis. *Generative* statistical shape models are important because they rely on shape representations that can lead to visualizations of individual shapes, group average shapes, or the modes of shape variation within groups. Such visualizations typically enhance the interpretability of the results underlying the study, which can provide insights to the medical experts. Applications where generative shape models play an important part include testing for differences in the distributions of shapes in two groups, segmenting objects in images, clustering a population of shapes into groups, etc. *Hypothesis testing* between groups of object shapes becomes challenging when the data from multiple groups have subtle differences that can get camouflaged when the learned statistical shape models are non-compact (depicting inflated variability) or when the model learning overfits to the (inevitable) corruptions in the data in the form of image noise or human errors in object segmentations. In such cases, (i) *corruption-aware learning* strategies, which can help infer the true shapes underlying the corruptions, (ii) *Riemannian* analysis, which can adapt to the nonlinear manifold structure underlying the true shapes, and (iii) *hierarchical modeling*, which can help regularize model fits, together have the potential to reduce overfitting and provide reliable compact per-group models that inform the inter-group differences. *Object segmentation* in medical images becomes challenging when the data differentiates poorly between the object and its surroundings, e.g., because of low contrast, blur, imaging noise, motion artifacts, etc. In such cases, information from a *learned statistical (prior) model of object shapes* can be crucial for accurate segmentation. *Shape Clustering* of population of shapes of particular medical organ is challenging when shapes are not well registered or landmarks are not provided. In such case, fitting a *hierarchical mixture model of shapes* directly to the population helps us getting accurate clusters without much of prior knowledge.

1.1 Shape Modeling Approaches and Applications

Statistical shape analysis can rely on several different modeling strategies, such as (i) boundary point distribution models (Kendall (1989); Cootes et al. (1995); Freifeld & Black (2012)), (ii) nonlinear dense diffeomorphic warps (Leventon et al. (2000); Glasbey & Mardia (2001); Vaillant & Glaunes (2005); Durrleman et al. (2009); Allassonniere et al. (2010); Durrleman et al. (2014); Zhang et al. (2017); Bone et al. (2018)), (iii) medial models (Fletcher et al. (2004); Siddiqi & Pizer (2008); Pouch et al. (2015)), (iv) spherical harmonics (SPHARM) based parameterization (Gerig et al. (2001)), or (v) implicit models (Tsai et al. (2003); Dambreville et al. (2008); Chen & Radke (2009)). All these modeling approaches continue to be active areas of research in their own right. Models based on distance transforms efficiently handle topology changes, but the statistical analysis in the space of distance transforms presents several challenges regarding defining, computing, and visualizing the means and modes of variation. Medial models and warp-based models represent shape as an equivalence class of object boundaries and lead to statistical analyses in the associated Riemannian shape spaces. While medial representations are typically limited to non-branching genus-0 objects, methods based on diffeomorphisms involve very large dimensional Riemannian spaces where the analysis can be expensive and challenged by noise and limited sample sizes of training data (Pizer et al. (2012)). On the other hand, *boundary point distribution models* can succinctly represent shapes, and we extend these models for hierarchical modeling in a Riemannian shape space, namely Kendall shape space, and we formulate the model learning to adapt to the corruption in the data. This work shows the benefits of our proposed modeling and inference scheme for hypothesis testing, Bayesian object segmentation and shape clustering.

Many methods for learning multigroup statistical shape models, e.g., (Terriberry et al. (2005); Cates et al. (2007); Schulz et al. (2016)), first pool the data from multiple groups to fit a single model that describes the variation across all groups, and then use the correspondences learned in the first step to estimate parameters of the distribution for each group. However, when data is naturally organized in groups, we employ *partial pooling* through *hierarchical modeling* (Gelman (2006)) that leads to a unified model whose inference leads to the *joint* estimation of (i) the optimal point correspondences and (ii) the distribution parameters for each group. Hierarchical models offer benefits including regularized model fits (e.g., for individual shapes or group-mean shapes) and reduced risk of overfitting per-group distributions (e.g., when the sample size is small or outliers are present) through *shrinkage*. In this work, we have quantitatively analyzed the benefits of hierarchical modeling for point distribution modeling schemes.

Object segmentation of any modality image in which object under consideration has a well defined shape, then using this information becomes very important under noisy acquisition, sparse acquisition, low resolution acquisition or combination of all three, which are very common scenario in normal clinical imaging systems. Various object segmentation method can be categorized into two major areas, as (i) curve or surface evolution methods, which involve fitting curve or surface to the boundaries, defined by energy functions (Cootes et al. (1995); Kass et al. (1988); Leventon et al. (2000)), (ii) labeling individual pixel or patch in the image as part of object or not (Aljabar et al. (2009); Wang et al. (2013a); Bai et al. (2013); Ronneberger et al. (2015)), which includes multi-atlas and deep neural nets. It is well understood that while performing image segmentation for an object with well defined shape, it is useful to incorporate shape information in the method itself. Methods like (Leventon et al. (2000)) tries to embed shape information in curve evolution itself to get improved performance. But from last few years, majority of object segmentation literature focused on multi-atlas, deep neural-nets and random forest (Wang et al. (2013a); Bai et al. (2013); Ronneberger et al. (2015); Badrinarayanan et al. (2017); Cuingnet et al. (2012)). All three method mentioned previously do not model shape explicitly. Recently many segmentation method based on deep neural-net incorporate data-driven shape information inside the model to improve performance (Ravishankar et al. (2017); Oktay et al. (2018)). Unlike previously mentioned deep neural-nets, we used formal definition of shape space (i.e. Kendall shape space) to define our distribution of shapes. On the similar line with this work (Li et al. (2018)), which combines PCA-based shape model with random forest segmentation, we have formulated a novel deep neural-net segmentation using our shape prior in Riemannian shape space. we have also quantitatively analyzed benefits of our shape prior for segmentation framework.

Along with object segmentation, clustering segmented object into set of categories without or very less human intervention is a challenging task, especially because shape of any medical entity under observation, always has complex and subtle information, which need humongous human effort and training forest for categorizing them into desired groups. Clustering set of data points into certain number of categories without supervision, is an old problem which has been addressed by many useful algorithms upto certain extent. Clustering algorithms can be categorized into two major areas, as (i) hard boundary clustering, in which each data-point will be labeled exactly single category (e.g. K-means, Graph-cut), (ii) soft boundary clustering, in which each data-point will be labeled a set of probabilities or memberships related to respective categories in consideration (e.g. Fuzzy C-means, Gaussian Mixture Model). Both areas has it's own significance in many applications. In our case, each data point is a shape representation from dataset

which does not have so well separated categories. To address such a shape clustering problem, we have formulated a novel Gaussian mixture model based hierarchical Riemannian shape clustering model. We have also quantitatively analyzed benefits of our model over state of the art shape clustering algorithm VBMixPCA (Gooya et al. (2018)).

1.2 Contributions

This work presents a novel framework for statistical shape analysis using point distribution models. In the context of *pointset* shape representations, we propose perhaps the first *hierarchical generative model* for multigroup data. We design our model to handle *human errors in segmentations* that are inevitable in practice. Together with *Riemannian* analysis in Kendall shape space, our model aims to learn reliable and compact shape distributions. We propose a novel method for *sampling* shapes from distributions in Kendall shape space, and use that for inference using Monte-Carlo expectation maximization (EM). We propose a novel Bayesian method for *object segmentation using shape priors*. we have formulated a shape prior using our Riemannian shape model, which can be incorporated with state of the art methods like deep neural net based image segmentation and multi-atlas based segmentation using our novel inference method. We propose a novel hierarchical Riemannian mixture clustering method for shape clustering. The results demonstrate the benefits of our framework in hypothesis testing, object segmentation, and shape clustering on simulated and clinical data. We have designed a simulated ellipsoids experiments to show two benefits of hierarchical model, (i) regularizing effect of population level variable, (ii) reducing false positives in hypothesis testing (shrinkage effect). For applications of hypothesis testing, we have fit our hierarchical Riemannian shape model to two-group simulated ellipsoid's noisy segmentation and clinical two-group (male and female) carpal bone segmentations. Then we have performed hypothesis testing using our novel test statistic on final estimated model. We have compared our hypothesis testing results with state of the art shape model, ShapeWorks (SCI Institute (2013)). For applications of object segmentation, we have performed two types of experiments, (i) *Simulating MRI* in which, we have simulated MRI like images from segmentations of ellipsoids and carpal bones. Then we performed our shape prior based segmentation method extending multi-atlas based segmentation method (Bai et al. (2013)). We have compared our results with state of the art shape model priors and pure multi-atlas based segmentation. (ii) *Clinical brain MRI* in which, we have actual clinical brain MRI scans and respective human-expert segmentation of sub-cortical structure is available. Then we performed our shape prior based segmentation method extending Unet (a deep neural nets) (Ronneberger

et al. (2015)) based segmentation method. We have compared our results with state of art shape priors, pure multi-atlas and deep neural networks. For applications of shape clustering, we have simulated three-group ellipsoid's noisy segmentation with overlapping group distributions. Then we estimated clusters using our hierarchical Riemannian mixture clustering method. We have quantitatively compared our results with state of the art shape clustering method, VBMixPCA (Gooya et al. (2018)).

1.3 Organization of Thesis

The rest of the report is organized as follows. Chapter 2 describes the related literature on shape modeling for hypothesis testing, object segmentation and shape clustering. Chapter 3 describes our methods for corruption-aware hierarchical Riemannian statistical shape modeling, its inference algorithms, and its applications to hypothesis testing. Chapter 4 describes our novel test statistic and presents results of hypothesis testing for simulated and clinical image data. Chapter 5 describes object segmentation using shape prior based on our Riemannian statistical shape modeling and presents results of segmentation using shape prior for simulated and clinical image data. Chapter 6 describes our newly designed shape clustering model using Riemannian statistical mixture shape model, its inference algorithms and presents results of shape clustering for simulated data. Chapter 7 concludes the thesis.

Chapter 2

Background and Related Work

This chapter describes the literature on shape modeling related to single group and multi-group dataset and discusses it's application to image segmentation, hypothesis testing and shape clustering. It also places the contributions in this report within that context.

2.1 Shape Modeling with Diffeomorphism

Some pioneering works on shape analysis through diffeomorphisms on images model nonlinear warps as latent variables (Allassonniere et al. (2007, 2010)). They treat the segmentation as error free. On the other hand, we model individual shapes as latent variables (treating data-to-shape similarity transforms as parameters) and allow for errors in segmentations. Nevertheless, (Allassonniere et al. (2007, 2010)) do *not* use a hierarchical model for multigroup data. Some recent approaches use hierarchical models in the context of nonlinear mixed-effect models to model longitudinal shape data in the space of diffeomorphisms on intensity images (Bone et al. (2018)). Another recent work (Zhang et al. (2017)) proposes an efficient probabilistic model of anatomical variability in a linear space of initial velocities of diffeomorphic transformations. It models principal modes as a multivariate complex Gaussian distribution on the initial velocities in a bandlimited space. In the context of hypothesis testing, methods based on diffeomorphisms involve very large dimensional Riemannian spaces, where the analysis can be expensive and challenged by noise and limited sample sizes of training data (Pizer et al. (2012)). On the other hand, pointset based shape representations typically model shapes in spaces that are much smaller dimensional.

2.2 Shape Modeling with Medial Models

Recent advances in medial representations of shape, e.g., the s-rep (Pizer et al. (2013); Hong et al. (2016); Tu et al. (2018)), improve over traditional medial models (Siddiqi & Pizer (2008)). While medial models have been popular for segmentation, they continue to be best suited to non-branching objects; a recent work (Pouch et al. (2015)) applies continuous medial models to segment the branching structure of the aortic valve, in a first such application. In contrast, point distribution models are applicable to various topologies, including multi-object complexes where the *joint* model can represent modes of joint variation across the objects in the complex. While some skeletal models infer modes of variation through composite principal nested spheres, we represent the modes of variation as eigenvectors of a covariance matrix in the tangent space of Kendall shape space at the estimated Frechet mean, by leveraging an extension of the Normal law on Riemannian manifolds (Pennec (2006)) to shape space. The focus in this work is on *hierarchical* (pointset-based) shape models that are aware of the corruptions in the data, for use in hypothesis testing, object segmentation and shape clustering.

2.3 Shape Modeling with Mixture Models

Very few works including (Gooya et al. (2018)) actually models shape space using mixture models. Mixture models has special importance when it comes to learning categories in unsupervised way. Unsupervised learning deals with learning patterns in unlabeled data through various methods, including clustering, density estimation, and latent variable models involving principal component analysis (PCA) and expectation maximization (EM). In case of clustering unlabeled data, Gaussian mixture models (GMMs) (Reynolds (2015); Rasmussen (2000)) have shown state-of-the-art performance in various domains, e.g., speech and text (Reynolds & Rose (1995); Reynolds et al. (2000)), image analysis and vision (Permuter et al. (2006); Greenspan et al. (2006); Zivkovic (2004)), etc. GMMs belong to a class of nonparametric probability density functions that can estimate complex densities as (weighted) superposition of multivariate Gaussian distributions. By using a sufficient number of Gaussians, and by proper adjustment of their respective parameters, almost any continuous density can be approximated to arbitrary accuracy (Nguyen & McLachlan (2018)).

2.4 Application: Hypothesis Testing using Shape distribution

Hypothesis testing using shape distributions has a long history in the contexts of medial models (Terriberry et al. (2005); Schulz et al. (2016)), point distribution models (Cates et al. (2008)), diffeomorphism based models (Durrleman et al. (2014); Joshi et al. (2016)), or contour representations (Hagwood et al. (2013)). Approaches based on constrained nonlinear deformations, e.g., constrained to be sufficiently smooth, invertible, inverse-consistent, etc., capture shape changes between objects in the combination of (i) the deformation and (ii) the residual inter-object mismatch modulo the random noise. On the other hand, our method, like many point distribution models, allows only similarity transforms on objects and thereby capture shape changes between objects using only the residual inter-object mismatch. In this way, our approach entails hypothesis testing in a much lower-dimensional space of pointsets, compared to the space of dense diffeomorphisms. We show that our pointset based hierarchical model, which models the corruptions in the data, performs better than the popular pointset based modeling scheme in (Cates et al. (2007); SCI Institute (2013)).

2.5 Application: Object Segmentation

There are many applications of active shape models (ASMs) for image segmentation (Heimann & Meinzer (2009)), including a recent one for 2D segmentation in cardiac ultrasound (Li et al. (2018)). A major limitation of the ASM scheme is the need for manual placement of landmarks, which is laborious and often infeasible for 3D objects having non-trivial shapes. Later methods for shape-model based segmentation (Tsai et al. (2003); Dambreville et al. (2008); Chen & Radke (2009)) rely on the level-set framework that represents object boundaries without the need for landmarks and has the ability to handle topological changes with ease. (Tsai et al. (2003)) relies on principal component analysis (PCA) on signed distance transforms of level sets, which represent object boundaries. Recently, (Veni et al. (2017)) uses graph cuts to impose local and global shape priors, where the latter uses distance transforms. However, the space of signed distance transforms is a nonlinear manifold and the linear combinations underlying PCA can fail to produce valid or useful shapes. To improve over linear PCA, (Dambreville et al. (2008)) proposes kernel PCA to model the distribution on the nonlinear manifold and (Chen & Radke (2009)) proposes a joint prior on shape and intensity for object segmentation using kernel density estimation that theoretically resembles kPCA. A recent method (Zhang

et al. (2012)) represents object shape as a sparse linear combination of image templates in a training set, but it assumes knowledge of a one-to-one correspondence between templates before learning. Another method models shape as a union of convex polytopes, where the polytopes are represented by intersections of half spaces (Mesadi et al. (2018)), but it requires the images to be segmented to be aligned to a fixed coordinate system as a pre-processing step. All aforementioned segmentation approaches rely on learning shape models from *highly-curated* error-free expert segmentations that are infeasible to obtain in practical deployment scenarios because of constraints on expert time and budgets.

Medial models have been used for object segmentation (Siddiqi & Pizer (2008); Pouch et al. (2014)), where they typically impose geometrical constraints on the estimated segmentation of non-branching sheet-like objects. Implicitly, the single learned medial-model template coupled with validity constraints and regularization terms acts as a topology-preserving shape “prior”, even though it is learned from only a single example and does *not* explicitly model a distribution of shapes. In contrast, our method builds a shape prior based on a learned distribution of object shapes.

Multiatlas segmentation (Aljabar et al. (2009); Wang et al. (2013a); Bai et al. (2013); Awate & Whitaker (2014)) typically relies on a large database of high-quality template images and segmentations. When the atlas database has limited size (Asman & Landman (2013)) or is imperfectly curated, it may reduce the reliability of the underlying image registration or label fusion, resulting in segmentation errors, e.g., altering its topology.

2.5.1 Non-Parametric Shape Models: Deep Boltzmann Machine, Deep Neural Network

Deep learning with Boltzmann machines (BMs) hierarchically models a distribution on shapes represented as discrete label maps (Chen et al. (2013)). However, the discretization makes it difficult for deep BMs to model fine-scale structural features; if they increase spatial resolution of the image, then they need larger training sets for reliable learning. Deep neural networks (DNNs) (Ronneberger et al. (2015); Oktay et al. (2018); Shin et al. (2016)) can perform complex segmentation tasks, but they often require either a *large* training set (e.g., with a few hundred expert segmentations; before data augmentation) or a *highly-curated* training set (e.g., with several tens of expert segmentations; before data augmentation) for effective learning to prevent overfitting. The need for large training sets with high-quality expert segmentations limits applicability of deep models in many clinical applications. Some methods (Ravishankar et al. (2017); Arif et al. (2017)) propose to modify the 2D Unet (Ronneberger et al. (2015)) to explicitly model global-level information in the form of 2D shapes for objects. In the context of medical image

segmentation, very recent methods published this year (Ataloglou et al. (2019); Zhu et al. (2019); Nogovitsyn & et al. (2019)) propose some schemes to handle the dearth of expert segmentations. For instance, (Zhu et al. (2019)) trains the DNN on 3D patches, at the risk of losing information about global shape regularity during learning. While (Nogovitsyn & et al. (2019)) train a DNN for hippocampus segmentation using 2500 samples, (Ataloglou et al. (2019)) use transfer learning to segment the hippocampus in one dataset by building upon a DNN trained on a large set of highly-curated segmentations in the ADNI dataset. However, such segmented datasets would be unavailable for many anatomical structures of interest (this work shows results on the thalamus and the caudate).

Our Bayesian model learning coupled with hierarchical Riemannian analysis produces compact shape (distribution) models even using imperfect expert segmentations. Consequently, ours results show that our method segments more accurately than other shape models. Our method extends and enhances multiatlas segmentation by combining a multiatlas-based likelihood model with our pointset-based Riemannian prior model for object shape.

2.6 Application: Shape Clustering

There are recent works on Shape clustering using pointset based models, which are mainly based on gaussian mixture models over shapes (Gooya et al. (2018)). But they are doing it in Euclidean manifold, which has drawbacks indicated by our results. Our method of clustering combines our Riemannian shape model with Gaussian mixture model formulation to get benefits of both worlds.

Chapter 3

A Hierarchical Generative Shape Model for Multigroup Shape Data

We model shape as an equivalence class of pointsets, where the equivalence is defined using similarity transforms comprising translation, rotation, and isotropic scaling. We describe a hierarchical statistical model that includes: (i) population-level variables modeling a distribution over group distributions of object shapes, (ii) group-level variables modeling a distribution over individual object shapes, (iii) a likelihood model on an individual’s segmentation (label map) conditioned on the individual’s object shape, (iv) the observed data that comprises expert generated segmentations that exhibit manual errors in the 3D labeling.

Our hierarchical framework is elaborated in figure 3.1.

3.1 Mathematical Notation

Consider a population comprising M groups, where group m has N_m individuals. For individual i in group m , let data x_{mi} represent the binary image (mask) corresponding to the segmented anatomical object. For individual i in group m , let y_{mi} be the (unknown) pointset representing object shape. For group m , let z_m be the (unknown) pointset representing the mean object shape and let C_m model the (unknown) covariance of shapes. We model the variability of the group-mean shapes z_m using population-level variables (i) μ that models the (unknown) pointset representing the population-level mean object shape and (ii) C that models the associated (unknown) covariance.

Each shape-representing pointset, or *shape pointset*, has J points in \mathbb{R}^3 . Let $y_{mij} \in \mathbb{R}^3$, $z_{mj} \in \mathbb{R}^3$, and $\mu_j \in \mathbb{R}^3$ denote the j -th points in the shape pointsets y_{mi} , z_m , and μ , respectively. We rely on the *Procrustes* framework that assumes a correspondence

between the J points across all shape-pointset variables, allowing us to represent each pointset by a vector in \mathbb{R}^{3J} . We model the population mean μ and covariance C and the group covariances $\{C_m\}_{m=1}^M$ as *parameters*. We model the shape pointsets Z_m and Y_{mi} as *latent random variables*. Each shape pointset also lies in preshape space (Kendall (1989); Goodall (1991)), with centroid at the origin and unit L^2 norm. For shape pointsets a and b , the Procrustes distance is $d_{\text{Pro}}(a, b) := \min_{\mathcal{R}} d_g(a, \mathcal{R}b)$, where operator \mathcal{R} applies a rotation to each point in the pointset b and $d_g(\cdot, \cdot)$ is the geodesic distance on the unit hypersphere in preshape space.

3.2 Modeling Smooth Shapes

We model a probability density function (PDF) to capture the variability across (i) group means Z_m and (ii) individual shapes Y_{mi} within each group, by extending the approximate Normal law on Riemannian manifolds (Pennec (2006)) to shape space, as motivated in (Gaikwad et al. (2015)). For a and b on the unit hypersphere, let $\text{Log}_a(b)$ be the logarithmic map of b with respect to a . Considering the tangent space of Kendall shape space at μ to relate to preshapes that are rotationally aligned to μ , we model the logarithmic map of shape a to the tangent space of the shape space at μ as $\text{Log}_\mu^\S(a) := \text{Log}_\mu(\mathcal{R}^*a)$, where $\mathcal{R}^* := \arg \min_{\mathcal{R}} d_g(\mathcal{R}a, \mu)$ optimally aligns the preshape-space pointset a to μ . Extending the Procrustes distance, we define the squared Mahalanobis shape-distance of shape a with respect to μ and C as $d_{\text{Mah}}^2(a; \mu, C) := \text{Log}_\mu^\S(a)^\top C^{-1} \text{Log}_\mu^\S(a)$. To model a PDF that gives larger probabilities to smoother shapes a , we use a prior that penalizes distances between each point a_j and its neighbors within the pointset a . Let the neighborhood system $\mathcal{N} := \{\mathcal{N}_j\}_{j=1}^J$, where set \mathcal{N}_j has the neighbor indices of j -th point. In practice, we get \mathcal{N} by fitting a triangular mesh to the segmented object boundary. Thus, the probability for shape a is

$$P(a|\mu, C, \beta) := \frac{1}{\eta(C, \beta)} \exp \left(-\frac{d_{\text{Mah}}^2(a; \mu, C)}{2} - \frac{\beta}{2} \sum_{j=1}^J \sum_{k \in \mathcal{N}_j} \|a_j - a_k\|_2^2 \right), \quad (3.1)$$

where $\beta \geq 0$ controls the strength of the smoothness prior and $\eta(C, \beta)$ is the normalizing constant. The second term in the exponent equals $0.5a^\top \Omega a$, where $\Omega \in \mathbb{R}^{3J \times 3J}$ is a sparse precision matrix with diagonal elements 2β and the only non-zero off-diagonal elements equaling $(-\beta)$ when the corresponding points are neighbors. We use this model for (i) the conditional PDF $P(z_m|\mu, C, \beta)$ of group mean shapes z_m and (ii) the conditional PDF $P(y_{mi}|z_m, C_m, \beta_m)$ of individual shapes y_{mi} .

3.3 Modeling Label-Map Likelihoods given Shapes

We design a novel measure of dissimilarity between the binary image x_{mi} and individual shape pointset y_{mi} . First, for each point y_{mij} in the shape pointset y_{mi} , we measure the Euclidean distance to the closest point on the object boundary given by the segmentation x_{mi} . We can compute such distances efficiently using the signed distance transform of the object boundary indicated by the segmentation x_{mi} . Let $\mathcal{D}_{x_{mi}}(v)$ be the value of this signed distance transform, at voxel v . Second, for each voxel on the zero level set of the distance transform $\mathcal{D}_{x_{mi}}$, we measure the Euclidean distance to the nearest point in the pointset y_{mi} . Let $\mathcal{Z}_{x_{mi}}^l$ be the coordinate of l -th voxel on the zero-level set of $\mathcal{D}_{x_{mi}}$. Third, we compute the aforementioned distances after aligning the pointset and the segmentation through a similarity transform. Let the operator \mathcal{T}_{mi} apply a similarity transform to the binary image x_{mi} . Then, our dissimilarity measure is:

$$\Delta(x_{mi}, y_{mi}) := \min_{\mathcal{T}_{mi}} \left(\sum_{j=1}^J (\mathcal{D}_{\mathcal{T}_{mi}x_{mi}}(y_{mij}))^2 + \sum_{l=1}^L \min_j \|\mathcal{Z}_{\mathcal{T}_{mi}x_{mi}}^l - y_{mij}\|_2^2 \right). \quad (3.2)$$

In practice, it is less efficient to apply similarity transforms to images x_{mi} instead of shape pointsets y_{mi} . Let the operator \mathcal{S}_{mi} represent a similarity transform applied to each point in the pointset y_{mi} . Then, our dissimilarity measure becomes

$$\Delta(x_{mi}, y_{mi}) := \min_{\mathcal{S}_{mi}} \left(\sum_{j=1}^J (\mathcal{D}_{x_{mi}}(\mathcal{S}_{mi}y_{mij}))^2 + \sum_{l=1}^L \min_j \|\mathcal{Z}_{x_{mi}}^l - \mathcal{S}_{mi}y_{mij}\|_2^2 \right). \quad (3.3)$$

Methods (Durrleman et al. (2009); Yu et al. (2014)) using current distance (Vaillant & Glaunes (2005)) have quadratic complexity in either pointset's cardinality. $\Delta(x, z)$ can be efficiently approximated using algorithms $O((J + L) \log(J))$, described later. We model $P(x_{mi}|y_{mi}) \propto \exp(-\gamma_m \Delta(x_{mi}, y_{mi}))$.

3.4 Our Hierarchical Statistical Shape Model

For multigroup data of object segmentations, PDF is $P(x, Y, Z|\mu, C, \{C_m\}_{m=1}^M) :=$

$$\prod_{m=1}^M P(Z_m|\mu, C) \prod_{i=1}^{N_m} P(Y_{mi}|Z_m, C_m; \beta_m) P(x_{mi}|Y_{mi}; \gamma_m), \quad (3.4)$$

where all $Z := \{Z_m\}_{m=1}^M$ and $Y := \{\{Y_{mi}\}_{i=1}^{N_m}\}_{m=1}^M$ are latent random variables, $x := \{\{x_{mi}\}_{i=1}^{N_m}\}_{m=1}^M$ are the observed data, the parameter set to be optimized is $\theta := \{\mu, C, \{C_m\}_{m=1}^M\}$. In this work, parameters $\beta_m := \beta, \gamma_m := \gamma, \forall m$, and β and γ are user defined.

Our optimization function becomes

$$\begin{aligned}\theta^* &:= \arg \max_{\theta} P(x|\theta) \\ &= \arg \max_{\theta} \int P(x, Y, Z|\theta) dY dZ\end{aligned}$$

which can be solved using Monte-Carlo Expectation Maximization(MCEM) optimization. Since PDFs involved in this optimization are not analytical, we choose MCEM over variational-EM.

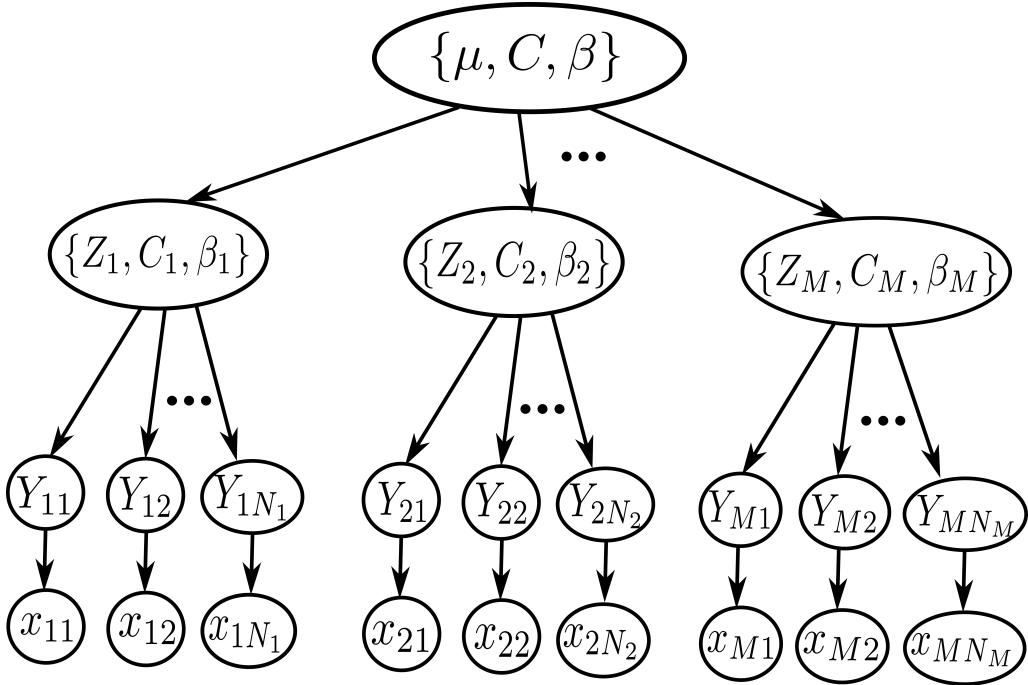


Figure 3.1: Hierarchical Model

3.5 Hierarchical Model Inference using Monte-Carlo Expectation Maximization

We use EM to fit the hierarchical model to the data $x := \{\{x_{mi}\}_{i=1}^{N_m}\}_{m=1}^M$. At iteration t within EM, let the parameter estimates be $\theta^t := \{\mu^t, C^t, \{C_m^t\}_{m=1}^M\}$. At iteration t , the *E step* defines the optimal lower bound to the observed-data likelihood function as $Q(\theta; \theta^t) := E_{P(Y, Z|x, \theta^t)}[\log P(x, Y, Z|\theta)]$, where $P(x, y, z|\theta)$ is the complete-data likelihood. We propose a Monte-Carlo approximation for the (intractable) expectation as

$$\widehat{Q}(\theta; \theta^t) := \frac{1}{S} \sum_{s=1}^S \log P(x, y^s, z^s|\theta), \text{ where } (y^s, z^s) \sim P(Y, Z|x, \theta^t). \quad (3.5)$$

Section 3.6 describes our algorithm for sampling shapes (y, z) from $P(Y, Z|x, \theta^t)$.

The *M step* updates the parameters to $\theta^{t+1} := \arg \max_{\theta} \widehat{Q}(\theta; \theta^t)$. Given data x and the sampled shape pairs $\{(y^s, z^s)\}_{s=1}^S$, we alternately optimize the shape-distribution parameters θ and the internal parameters \mathcal{S}_{mi} until convergence.

Updating Similarity Transforms \mathcal{S}_{mi} . Typically, the number of data points L on the zero level-set are at least an order of magnitude fewer than the cardinality J of the shape pointsets. While aligning the binary images to the sampled shape pointsets $\{y_{mi}^s\}_{s=1}^S$, we can ensure that the parameter updates are sufficiently small such that, for most data points $\mathcal{Z}_{x_{mi}}^l$, the nearest shape point $y_{mi,j'}^s$ (where the shape point indexed by j' is the one closest to the data point indexed by l) is the same before and after the update. We use (projected) gradient descent for optimization. We approximate the gradients of the desired objective function by first finding the pairs of nearest points (l, j') and then taking the gradients of the resulting function

$$\mathcal{J}(\mathcal{S}_{mi}) := \sum_{s=1}^S \left(\sum_{j=1}^J (\mathcal{D}_{x_{mi}}(\mathcal{S}_{mi} y_{mi,j}^s))^2 + \sum_{l=1}^L \|\mathcal{Z}_{x_{mi}}^l - \mathcal{S}_{mi} y_{mi,j'}^s\|_2^2 \right), \quad (3.6)$$

with respect to the parameters underlying the similarity transform \mathcal{S}_{mi} , which are (i) translation $\mathbf{t} \in \mathbb{R}^3$ represented as a 3×1 column vector, (ii) scaling parameter $\mathbf{s} \in \mathbb{R}_{>0}$, and (iii) rotation matrix $\mathbf{R} \in \text{SO}(3)$, such that, for a point $a \in \mathbb{R}^3$ represented as a 3×1 column vector, the similarity-transformed point is $\mathcal{S}_{mi}(a) := \mathbf{t} + \mathbf{s}\mathbf{R}a$. The rotation matrix update relies on a reparameterization of the rotation matrix as the matrix exponential of a skew symmetric matrix. To update \mathbf{R} , we first compute the gradient of the objective function with respect to \mathbf{R} , then project the gradient onto the tangent space of $\text{SO}(3)$ at the identity matrix (i.e., the space of skew symmetric matrices), scale it by the step-size underlying gradient descent, take the matrix exponential (which is also the exponential map) of the scaled projection, and right-compose the resulting rotation matrix update with the original rotation matrix \mathbf{R} . A.1 gives the gradients of $\mathcal{J}(\mathcal{S}_{mi})$ with respect to parameters \mathbf{t} , \mathbf{s} , and \mathbf{R} .

The objective function value depends on the nearest point in a shape pointset $\mathcal{S}_{mi} y_{mi}$ (of cardinality J) to each point in the set $\mathcal{Z}_{x_{mi}}$ (of cardinality L). We can find the required pairs of nearest neighbors in $O((J+L) \log J)$ time by building the k-d tree for each pointset (taking $O(J \log J)$ computation) followed by L nearest-neighbor searches (taking $O(L \log J)$ computation). Computing $\mathcal{D}_{x_{mi}}$ is a one-time pre-processing task. To evaluate the objective function, we need to access the values within $\mathcal{D}_{x_{mi}}$ at J locations given by $\mathcal{S}_{mi}(y_{mi})$, taking $O(J)$ computation. So, for each individual, computing the likelihood function has $O(J + (J+L) \log J)$ complexity.

Updating Population-Level Mean μ . The update for μ is given by

$$\min_{\mu} \sum_{s=1}^S \sum_{m=1}^M \text{Log}_{\mu}(z_m^s)^\top C^{-1} \text{Log}_{\mu}(z_m^s) \text{ such that } \|\mu\|_2 = 1. \quad (3.7)$$

We use projected gradient descent (A.2 derives the gradient), where the projection rescales the updated mean to unit norm. Here all gradients should lie in tangent hyper-plane at μ in Riemannian space, but calculated gradients can have components normal to tangent hyper-plane, so we project gradient on tangent hyper-plane at μ . By doing this, next update, by this gradient, always lies in tangent hyper-plane at μ and has assured unit norm. This update can be get by exponential map of projected gradient at μ , which is nothing but Riemannian version of gradient descent update.

Updating Group-Covariance Parameters C_m . We update C_m as

$$\arg \min_{C_m} \sum_{i=1}^{N_m} \sum_{s=1}^S \text{Log}_{z_m^s}^{\mathbb{S}}(y_{mi}^s)^\top C_m^{-1} \text{Log}_{z_m^s}^{\mathbb{S}}(y_{mi}^s) + (y_{mi}^s)^\top \Omega y_{mi}^s + 2 \log \eta(C_m, \beta_m). \quad (3.8)$$

Although the normalization term $\eta(C_m, \beta_m)$ is difficult to evaluate analytically, we find that it can often be well approximated in practice. Assuming that the shape distribution $P(y_{mi}^s | z_m^s, C_m)$ has sufficiently low variance, the tangent vector $\text{Log}_{z_m^s}^{\mathbb{S}}(y_{mi}^s)$ is close to the difference vector $y_{mi}^s - z_m^s$ (details in A.3), in which case $P(y_{mi}^s | z_m^s, C_m)$ appears as a product of a multivariate Gaussian $G(y_{mi}^s; z_m^s, C_m)$ with another multivariate Gaussian $G(y_{mi}^s; \mathbf{0}, \Omega)$. The product distribution equals $G(y_{mi}^s; z_m^s, C_m^{\text{reg}})$ where the regularized covariance $C_m^{\text{reg}} := (C_m^{-1} + \Omega)^{-1}$ restricts all variability to the tangent space at the mean z_m^s and the normalization term $\eta(C_m, \beta_m) \approx (2\pi)^{D/2} |C^{\text{reg}}|^{0.5}$. Then, the optimal covariance $\widehat{C}_m^{\text{reg}}$ is the sample covariance of tangent vectors $\text{Log}_{z_m^s}^{\mathbb{S}}(y_{mi}^s)$ in the tangent spaces at z_m^s . So, the group covariance \widehat{C}_m update takes a closed form.

Updating Population-Level Covariance Parameter C . The strategy for updating C is analogous to the one just described for estimating C^m . We first compute \widehat{C}^{reg} as the sample covariance of tangent vectors $\text{Log}_{\mu}^{\mathbb{S}}(z_m^s)$ in the tangent space at μ . Then, the optimal $\widehat{C} = ((\widehat{C}^{\text{reg}})^{-1} - \Omega)^{-1}$.

3.6 Robust Efficient MCMC Sampling on Riemannian Manifolds

Monte-Carlo EM entails sampling shape-pointset pairs (y^s, z^s) from their posterior PDF $P(Y, Z|x, \theta^t)$ in shape space. We propose a generic scheme for efficient sampling in high-dimensional spaces on a Riemannian manifold and adapt it for sampling in shape

space. Standard Metropolis-Hastings or Gibbs MCMC samplers are inefficient in high-dimensional spaces (MacKay (2012)) where the data typically shows strong correlations across dimensions. We propose to adapt Skilling’s multistate leapfrog method (MacKay (2012)), an efficient MCMC sampler, to Riemannian spaces. We find that alternate efficient MCMC methods, e.g., Hamiltonian Monte Carlo used in (Yu et al. (2014)), can be more sensitive to the tuning of the underlying free parameters (Wang et al. (2013b)) that involve parameters underlying solving the coupled partial differential equations related to the Hamiltonian. The choice of the free parameters underlying the HMC can be critical in controlling the numerical stability of the differential equation solver. On the other hand, we propose a sampler that we found to be more robust in practice, requiring lesser parameter tuning.

We adapt Skilling’s multistate leapfrog method to a Riemannian manifold as follows. Consider a multivariate random variable F taking values f on a Riemannian manifold \mathbb{F} with associated PDF $P(F)$. We initialize the MCMC sampler with a set of states $\{f^q \in \mathbb{F}\}_{q=1}^Q$. We propose to leapfrog a randomly-chosen current state f^{q_1} over another randomly-chosen state f^{q_2} to give a proposal state $f^{q_3} := \text{Exp}_{f^{q_2}}^{\mathbb{F}}(-\text{Log}_{f^{q_2}}^{\mathbb{F}}(f^{q_1}))$, where the logarithmic and exponential maps are defined with respect to the manifold \mathbb{F} . The proposal state f^{q_3} is accepted (replacing the original state f^{q_1}), according to the Metropolis-Hastings method, with probability equal to the ratio $P(f^{q_3})/P(f^{q_1})$. The sampler only needs to evaluate probability density ratios, without needing the gradients of the PDF $P(F)$. Such leapfrog jumps are repeated and after sufficient burn in, the set of Q states are considered to be a sample from $P(F)$.

We adapt the proposed leapfrog sampling scheme to shape space for sampling from the Normal law $P(z|\mu, C, \beta)$ that leads to a Gaussian PDF in the tangent space of shape space at μ , where the tangent space comprises all shapes aligned to μ . We initialize the set of states to the pointsets $\{z^q\}_{q=1}^Q$ in preshape space and rotationally aligned to the mean μ . In shape space, we propose the leapfrog step

$$z^{q_3} := \arg \min_{c:=\mathcal{R}b} d_g(c, \mu), \text{ where } b := \text{Exp}_{z^{q_2}}(-\text{Log}_{z^{q_2}}(z^{q_1})), \quad (3.9)$$

which approximates the geodesic from z^{q_1} to z^{q_2} and uses that to propose the “leap” to z^{q_3} in shape space. The Metropolis-Hastings rule then accepts or rejects z^{q_3} . To increase more independent samples, as suggested by (MacKay (2012)), we reject 4 consecutive set of states before accepting new set of state. Since we have to sample shape-pointset pairs (y^s, z^s) from their posterior PDF $P(Y, Z|x, \theta')$, we perform iteration between y^s and z^s according to Gibbs sampling mechanism, and inside each iteration we sample y^s and z^s individually using our Riemannian leapfrog sampling.

Chapter 4

Hierarchical Riemannian Shape Modeling for Hypothesis Testing

Given multigroup shape data, in the form of manual segmentations of objects across individuals and groups, we can use our MCEM hierarchical model fitting to test the hypothesis that the distributions of object shapes in two groups is identical. We use permutation testing for this task. Unlike parametric hypothesis tests, permutation tests are nonparametric, rely on the generic assumption of exchangeability, lead to stronger control over Type-1 error, and are more robust to random errors in the measurements / post-processing of image data. After estimating the group means and covariances $\{z^m, C^m\}_{m=1}^M$, we propose a test statistic to measure the differences between the shape distributions arising from two cohorts, say, A and B , by summing up the squared Mahalanobis shape-distances from each individual shape $y_{\bullet i}$ in both groups to the mean z^\bullet in the other group (using the other group's covariance C^\bullet), i.e.,

$$T := \frac{1}{N_A} \sum_{i=1}^{N_A} d_{\text{Mah}}^2(y_{Ai}; z^B, C^B) + \frac{1}{N_B} \sum_{i=1}^{N_B} d_{\text{Mah}}^2(y_{Bi}; z^A, C^A). \quad (4.1)$$

When the two groups have similar means and covariances, the test statistic T takes relatively smaller values, compared to the case when the group means and covariances differ. The test-statistic distribution is unknown analytically and we infer it using bootstrap sampling, using 300 repeats.

We first present results to show the benefits of hierarchical modeling, where we use simulated data that gives us a ground truth for comparison. Subsequently, we show results comparing our method with other methods, on simulated and clinical Carpel bone image data, for the tasks of hypothesis testing.

4.1 Validation: Benefits of Hierarchical Modeling

Hierarchical modeling, which learns multigroup distributions jointly with population-level variables, offers some benefits in (i) regularizing model fits and (ii) reducing the occurrences of false positives during hypothesis testing.

Regularizing Effect of the Population-Level Variables. When we have data from multiple groups of shapes such that one group has a much lower sample size than the others, the population-level variables can help regularize the parameter estimation for that group by reducing the sensitivity of the parameter estimation to the particular random sample observed for that group. To evaluate this regularizing effect, we simulated data from three groups of ellipsoidal shapes, where all groups keep two of their ellipsoidal axes lengths fixed to 1 and vary the third axis from (i) 0.35 to 0.55 for group 1, (ii) 0.4 to 0.6 for group 2, and (iii) 0.45 to 0.65 for group 3. For the first and third group, we create data for 1000 ellipsoids randomly drawn from the distribution. For the second group, we create data for 10 ellipsoids randomly drawn from the distribution. Then, we estimated all model parameters by introducing and varying the weight, say, α , of the hierarchical prior $\alpha \log P(z_m; \mu, C, \beta)$, with $\alpha \in [0, 1]$. We evaluated the model fits by computing the geodesic distance between estimated mean and true mean for the low-sample group. We repeated this experiment 500 times under varying random draws of the data for all groups. The cumulative distribution function (CDF) of the distances (Figure 4.1(a)), across experiments, shows that the estimates have decreasing variation (CDF rises faster) as α increases to 1. This is the regularizing effect we expect to see. In this set of experiments, we also find that the estimates gain accuracy with increasing α , because the population mean and the mean of the low-sample group are close to each other by design. In general, a hierarchical model can reduce the variation / sensitivity of the estimates at the cost of introducing a bias.

Reducing False Positives in Hypothesis Testing (Shrinkage Effect). During hypothesis testing, when the group sample sizes are small, there is a tendency of models to overfit to the data, which can cause the fitted distributions to differ severely from the actual group distributions. However, the presence of the population-level variables in the hierarchical model have a regularizing effect on the group-model parameters, by pulling group mean estimates towards the population mean and, thereby, reducing the distance between them. This “shrinkage” effect prevents overfitting and reduces the occurrence of false positives. Thus, a hypothesis test indicating statistically significant group differences using hierarchical modeling can be more reliable than the one without hierarchical modeling. To evaluate the shrinkage effect, we simulate ellipsoidal data from two groups, with identical shape distributions, by keeping two ellipsoidal axes fixed to 1 and varying

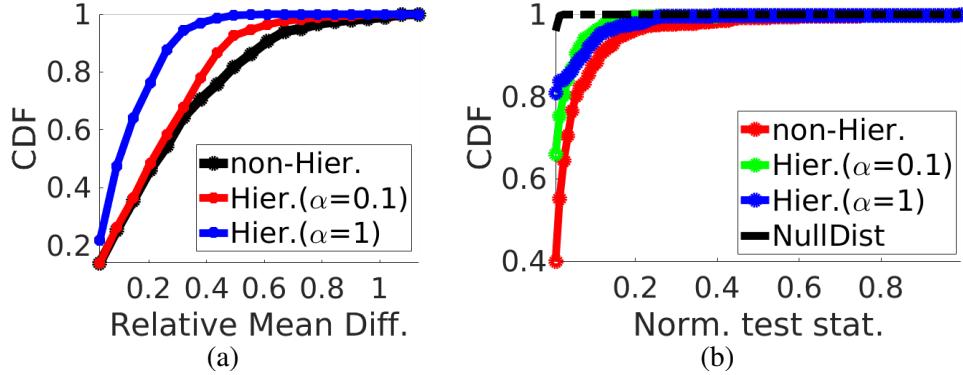


Figure 4.1: **Benefits of Hierarchical Modeling.** (a) *Regularizing Effect of Population-Level Variables*: Distribution of geodesic distances between estimated and true group means, with varying strengths α (see text) of the population-level prior. (b) *Reducing False Positives in Hypothesis Testing (Shrinkage Effect)*: Distribution of the test statistic between two group samples drawn from the same shape distribution, with varying strengths α (see text) of the population-level prior.

the third axis in the range 0.4 to 0.8. We randomly generate two groups of sample size 10 (small), fit the hierarchical model with varying α , and compute the test statistic. We repeat this experiment 500 times to get the CDF of the test statistic with the small sample sizes of 10. As α increases, we find the test statistic PDF mass shifting left towards zero, which is desirable because the two groups of data are actually drawn from the same shape distribution. We also conduct another 500 experiments with a large sample size of 1000, where the CDF (Figure 4.1(b)) clearly indicates all test statistic values to be close to zero, as expected.

4.2 Results: Hypothesis Testing

We present the results of hypothesis testing using our method for hierarchical shape modeling and inference. We have implemented our model and hypothesis testing framework in MATLAB programming environment without any external library support. We have compared our method with a state-of-the-art pointset-based shape analysis approach named ShapeWorks (SCI Institute (2013)) that differs from our approach in many ways as follows. First, ShapeWorks constraints the points to be placed (at real-valued sub-voxel locations) on the zero level set associated with the distance transform of the object boundary represented by the segmentation; in this way, it becomes sensitive to the (inevitable) errors in manual segmentation. Second, it does *not* rely on a hierarchical modeling approach. Third, it assumes an initial alignment of the segmentations and does *not* align them during inference. Fourth, ShapeWorks treats the group means as parameters in-

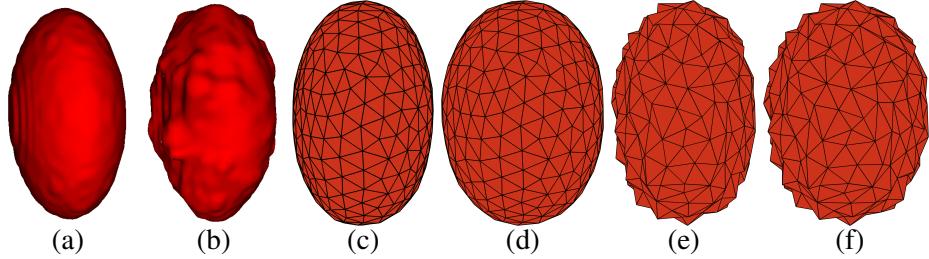


Figure 4.2: **Results: Hypothesis Testing on Simulated Ellipsoidal Data.** (a) Ground truth segmentation. (b) Corrupted segmentation (data), mimicking human errors. (c)-(d) Our method’s MAP estimates for group means z_1, z_2 . (e)-(f) ShapeWorks’s estimates for group means z_1, z_2 .

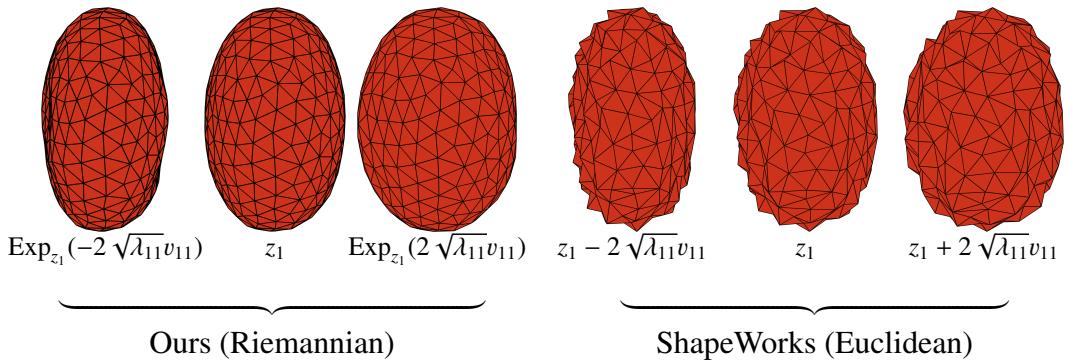


Figure 4.3: **Results: Hypothesis Testing on Simulated Ellipsoidal Data.** Estimated principal mode of variation of the first group with mean z_1 for our method and ShapeWorks, with λ_{11} as first principal eigenvalue and v_{11} as first principal eigenvector of group covariance C_1 .

stead of random variables. To evaluate ShapeWorks, we first align all the images using similarity-transform based registration, and then give the aligned segmentations, of all groups, to ShapeWorks to compute the optimal point placement. Then, given the aligned pointsets of all groups, we compute the shape mean for each group. Now, we make the scale of the mean-shape pointsets and individual-shape pointsets obtained by ShapeWorks commensurate to that of our approach as follows. For each group, we center the mean shape so that its centroid is at the origin, rescale the mean shape to unit norm, and then apply the same group-specific centering and rescaling to each member of the group. We evaluate the methods on simulated data as well as real-world clinical data including bone shapes obtained from computed tomography (CT) imaging and subcortical brain shapes obtained from magnetic resonance imaging (MRI).

4.2.1 Results: Hypothesis Testing on Simulated Ellipsoidal Data

We simulate two groups of segmentations of ellipsoidal shapes, each having 36 samples with variation along single axis, keeping the other two axes lengths the same over the

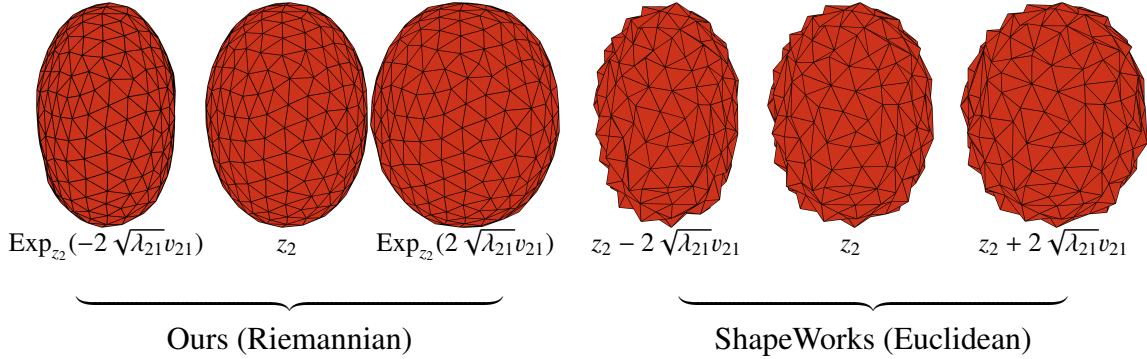


Figure 4.4: **Results: Hypothesis Testing on Simulated Ellipsoidal Data.** Estimated principal mode of variation of the second group with mean z_2 for our method and ShapeWorks, with λ_{21} as first principal eigenvalue and v_{21} as first principal eigenvector of group covariance C_2 .

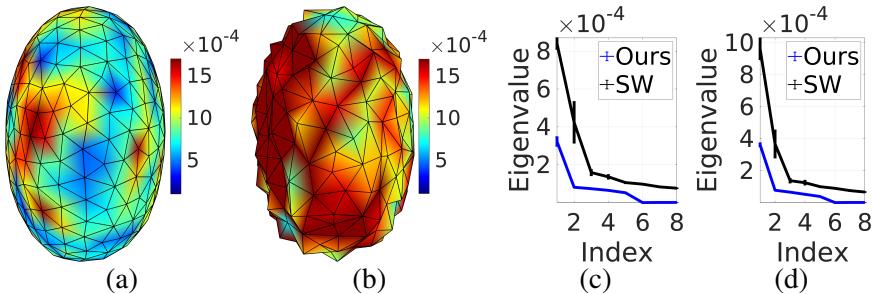


Figure 4.5: **Results: Hypothesis Testing on Simulated Ellipsoidal Data.** Compactness of point correspondences across fitted shapes in one group, indicated pointwise by the variability in the locations of each set of corresponding points (i.e., $\{y_{mij}\}_{i=1}^{N_m}$) across the group, where the variability measure is the sum of the smallest two eigenvalues of the 3D covariance of each point on the mesh. Each mesh-point’s variability in location (y_{mij}) across the group (i.e., over all shapes $i = 1, \dots, N_m$) for (a) our model and (b) ShapeWorks, visualized on the estimated mean shape (i.e., z_m) for each group. (c)-(d) Eigen-spectra for group covariances C_1 and C_2 ; error bars show variation across multiple experiments with different corruption instances.

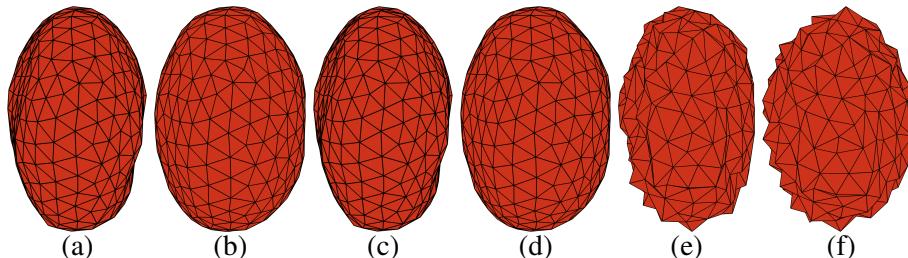


Figure 4.6: **Results: Hypothesis Testing on Simulated Ellipsoidal Data.** (a)-(b) Sampled individual shapes y_{mi}^s from our method. (c)-(d) MAP estimates of individual shapes y_{mi} from our method. (e)-(f) Estimated shapes of individuals from ShapeWorks.

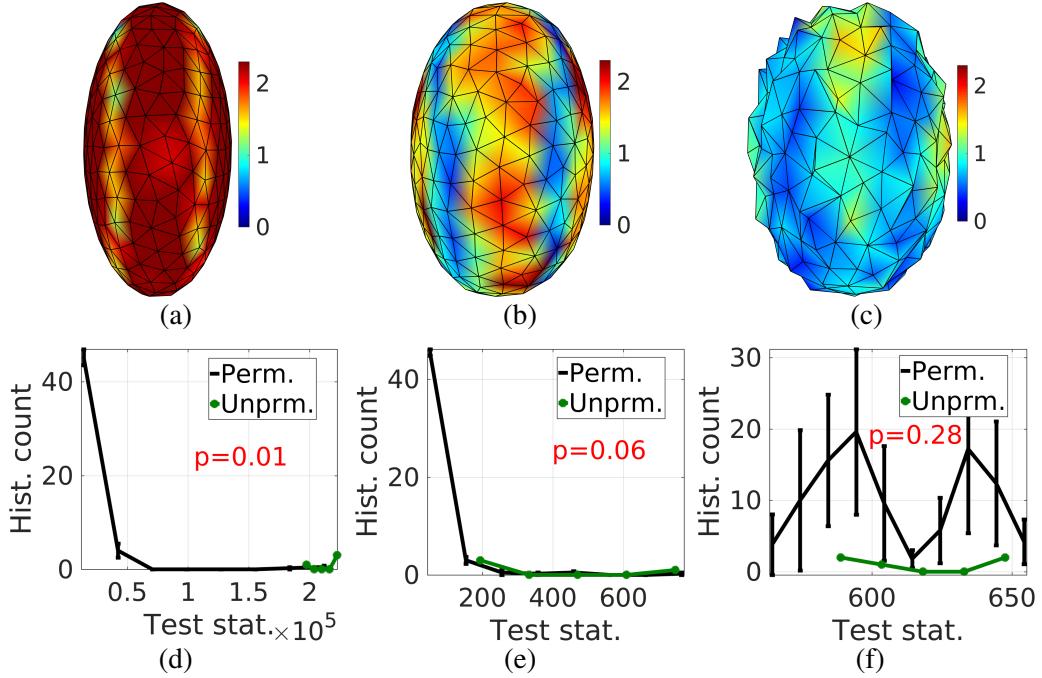


Figure 4.7: Results: Hypothesis Testing on Simulated Ellipsoidal Data. Cohen’s d effect size between the two group distributions based on estimates of group means z_1 and z_2 and covariances C_1 and C_2 , computed per mesh point j based on the j -th components of the means and the j -th diagonal elements of the covariance matrices, and visualized on the population mean μ , for: (a) ground truth data (large sample size; 2000 per group) and (b)-(c) our method and ShapeWorks, respectively (small sample size, 36 per group). *Permutation tests* showing the histogram of the test statistic under the null hypothesis and the unpermuted test-statistic values (error bars represent results from repeated experiments) for: (d) ground truth data (large sample size; 2000 per group) and (e)-(f) our method and ShapeWorks, respectively (small sample size, 36 per group).

entire population. The range of the variations across both groups overlap. We also introduce surface perturbations in the form of bumps and pits to these binary segmentations to mimic human errors in labeling (Figure 4.2(a)-(b)). To generate the ground truth values for the group-mean and group-covariance parameters, we simulate two groups of ellipsoidal segmentations as before, but with 2000 shapes in each group without any errors in the segmentations. We then use the large-sample uncorrupted data to fit the model to each group to get its “true” means and covariances. The estimates of group means (Figure 4.2) and modes of variation (Figures 4.3, 4.4) from our method indicate (i) smoother and more accurate group means and (ii) modes of variation that indicate smoother and more accurate shape variations, compared to ShapeWorks. The variability in the locations of each set of corresponding mesh points, within a group, is much smaller for our model, indicating that this distribution was more compact (mainly concentrated in the first eigenvalue) com-

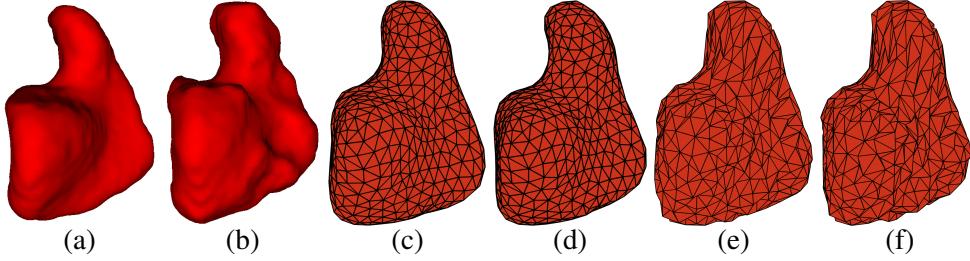


Figure 4.8: **Results: Hypothesis Testing on Carpal Bones - Hamate.** (a) Ground truth segmentation. (b) Corrupted segmentation (data), mimicking human errors. (c)-(d) Our method’s MAP estimates for group means z_1, z_2 . (e)-(f) ShapeWorks’s estimates for group means z_1, z_2 .

pared to ShapeWorks (Figure 4.5(a)-(b)). The group covariances obtained from our model (Figure 4.5(c)-(d)) are more compact than those from ShapeWorks. The compactness of our model fit stems from (i) points *not* being forced to lie on the (corrupted) segmentation boundary, unlike ShapeWorks, and (ii) the use of hierarchical modeling and Riemannian geometrical modeling. Consequently, the sampled shapes (Figure 4.6(a)-(b)) and MAP estimates of individual shapes from our method (Figure 4.6(c)-(d)) are smooth and realistic, unlike the individual shape estimates from ShapeWorks (Figure 4.6(e)-(f)). Cohen’s d effect sizes (Figure 4.7(a)–(c)) show that the surface pattern of inter-group differences inferred using method is closer to the ground truth (learned using a very large sample size of 2000 per group), as compared to ShapeWorks. Consequently, the permutation-test p values from our method (Figure 4.7(e)) are significantly smaller, as desired, than those from ShapeWorks (Figure 4.7(f)), thereby correctly indicating of a significant different between the distributions of the two groups as indicated by the ground truth (Figure 4.7(d)).

4.2.2 Results: Hypothesis Testing on Carpal Bones Data

This dataset (Moore et al. (2007)) includes segmentations of eight different carpal bones (namely, hamate, trapezoid, pisiform, capitate, trapezium, lunate, scaphoid, and triquetrum) for 15 male subjects and 15 female subjects. Unlike the simulated ellipsoidal dataset, the bone shapes in this dataset are non-trivial. These segmentations inherently have a small amount of bumpy surface perturbations. To create a more challenging task, we introduced additional surface perturbations in the form of bumps and pits to these segmentations to mimic human errors in labeling, e.g., in Figure 4.8(a)-(b).

Analogous to the ellipsoidal dataset, for the hamate bone, we find that the means (Figure 4.8(c)–(f)) and the modes of variation (Figures 4.9,4.10) are more realistic for our method, compared to ShapeWorks. Compared to ShapeWorks, the variability in the locations of each set of corresponding mesh points, within a group, is more com-

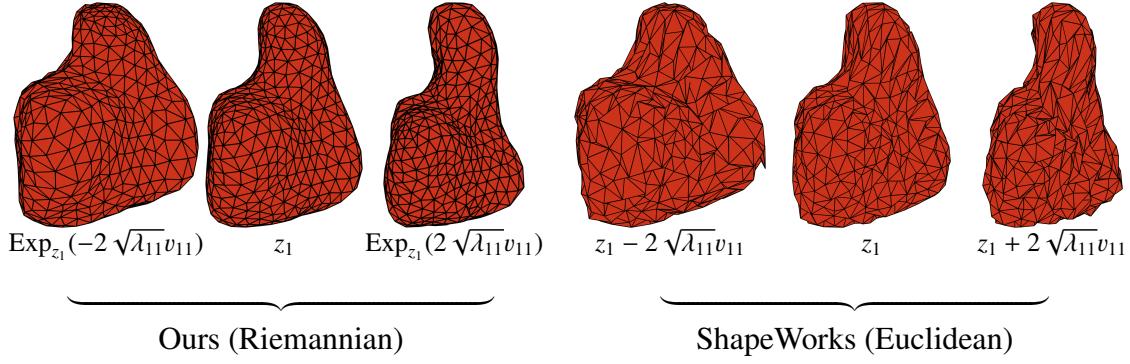


Figure 4.9: **Results: Hypothesis Testing on Carpal Bones - Hamate.** Estimated principal mode of variation of the first group with mean z_1 for our method and ShapeWorks, with λ_{11} as first principal eigenvalue and v_{11} as first principal eigenvector of group covariance C_1 .

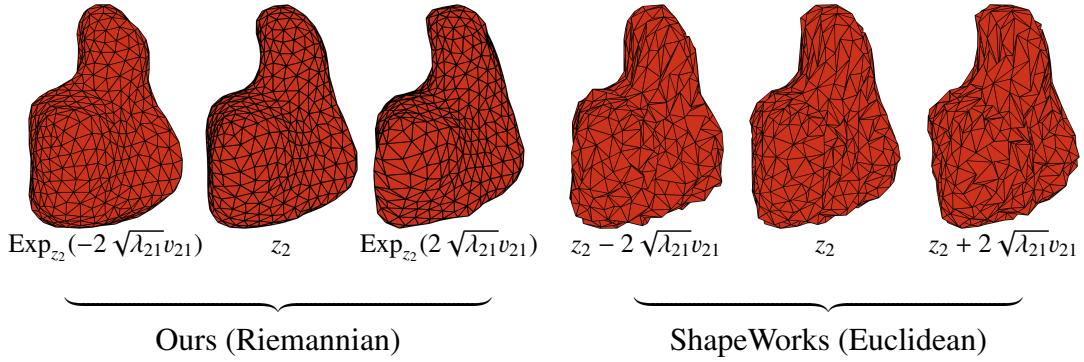


Figure 4.10: **Results: Hypothesis Testing on Carpal Bones - Hamate.** Estimated principal mode of variation of the second group with mean z_2 for our method and ShapeWorks, with λ_{21} as first principal eigenvalue and v_{21} as first principal eigenvector of group covariance C_2 .

pact for our model (Figure 4.11(a)-(b)) and, consequently, so are the group covariances (Figure 4.11(c)-(d)). The sampled shapes (Figure 4.12(a)-(b)) and MAP estimates (Figure 4.12(c)-(d)) of individual shapes from our method are more smooth and realistic, compared to ShapeWorks (Figure 4.12(e)-(f)). Some literature points to shape differences between the carpal bone shapes in male and female hands. Our method shows the differences, in terms of Cohen's d effect sizes (Figure 4.13(a)-(b)) in a more spatially focused manner. The permutation test using our statistic tends a little more towards the indication of differences between the groups (Figure 4.13(c)-(d)). We also explored utility of our model that allows the optimal point placements off the segmented object boundary. For ShapeWorks, if we additionally smoothed the segmentation data in order to eliminate the perturbations in the segmentation, then there is a risk of oversmoothing the features themselves, leading to parameter estimates, e.g., group mean shapes (Figure 4.14(e)-(f)), and individual shapes that lose the characteristic biological features, which is undesirable. In

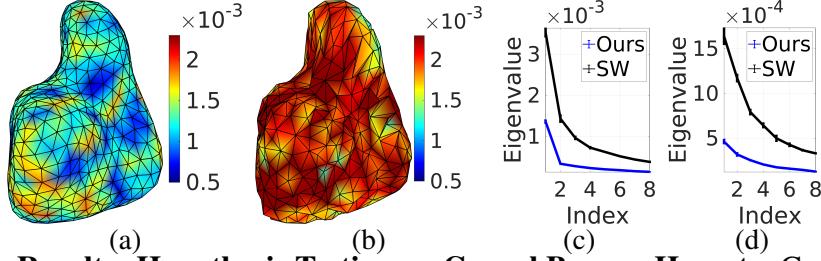


Figure 4.11: **Results: Hypothesis Testing on Carpal Bones - Hamate.** Compactness of point correspondences across fitted shapes in one group, indicated pointwise by the variability in the locations of each set of corresponding points (i.e., $\{y_{mij}\}_{i=1}^{N_m}$) across the group, where the variability measure is the sum of the smallest two eigenvalues of the 3D covariance of each point on the mesh. Each mesh-point’s variability in location (y_{mij}) across the group (i.e., over all shapes $i = 1, \dots, N_m$) for (a) our model and (b) ShapeWorks, visualized on the estimated mean shape (i.e., z_m) for each group. (c)-(d) Eigenspectra for group covariances C_1 and C_2 ; error bars show variation across multiple experiments with different corruption instances.

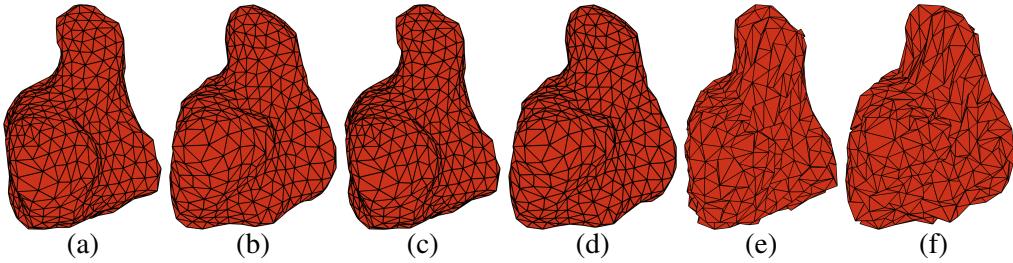


Figure 4.12: **Results: Hypothesis Testing on Carpal Bones - Hamate.** (a)-(b) Sampled individual shapes y_{mi}^s from our method. (c)-(d) MAP estimates of individual shapes y_{mi} from our method. (e)-(f) Estimated shapes of individuals from ShapeWorks.

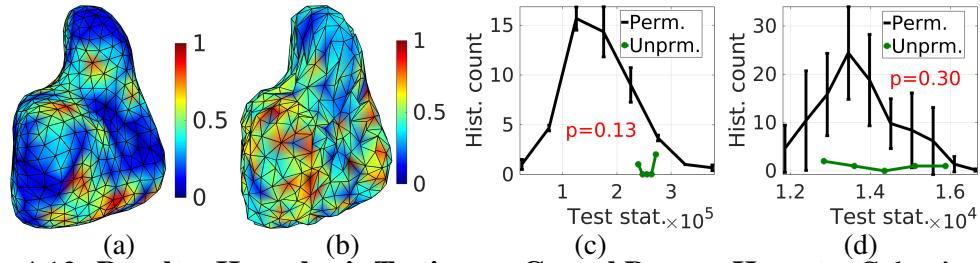


Figure 4.13: **Results: Hypothesis Testing on Carpal Bones - Hamate.** Cohen’s d effect size between the two group distributions based on estimates of group means z_1 and z_2 and covariances C_1 and C_2 , computed per mesh point i based on the i -th components of the means and the i -th diagonal elements of the covariance matrices, and visualized on the population mean μ , for: (a) our method and (b) ShapeWorks. Permutation tests showing the histogram of the test statistic under the null hypothesis and the unpermuted test-statistic values (error bars represent results from repeated experiments) for: (c) our method and (d) ShapeWorks.

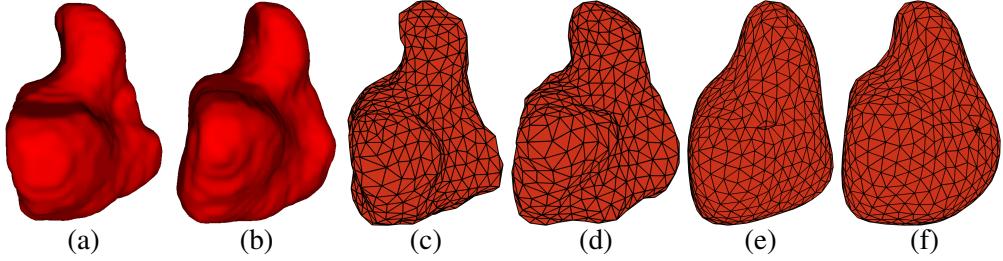


Figure 4.14: **Results: Hypothesis Testing on Carpal Bones - Hamate.** (a)-(b) Ground truth object segmentations of an example male and female bones, respectively. (c)-(d) Our method’s MAP estimates for the bone shapes in (a) and (b), respectively. (e)-(f) ShapeWorks’s estimates for the bone shapes in (a) and (b), respectively, *when ShapeWorks additionally smooths the corrupted data to get rid of the added bumps*.

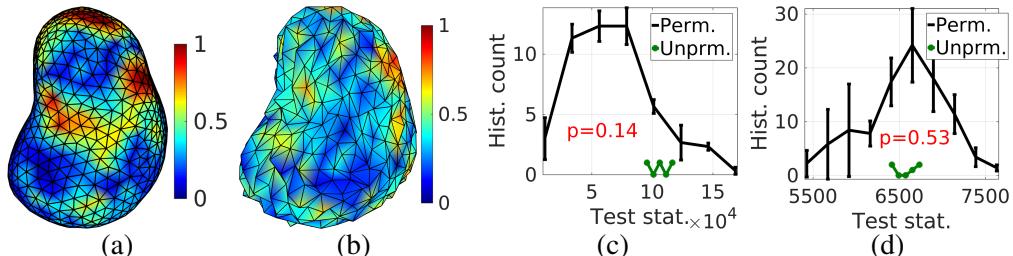


Figure 4.15: **Results: Hypothesis Testing on Carpal Bones - Trapezoid.** Captions as in Figure 4.13.

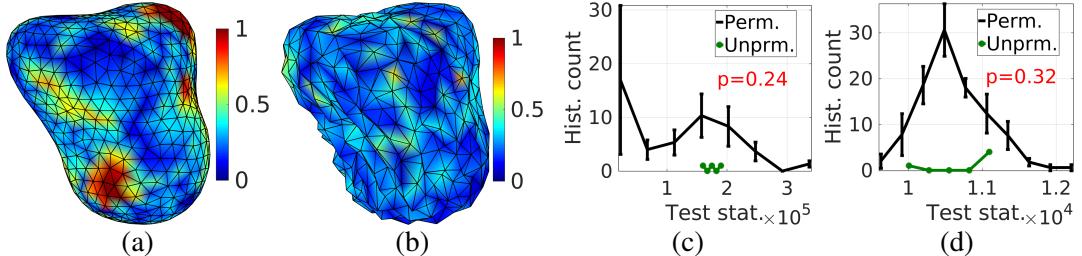


Figure 4.16: **Results: Hypothesis Testing on Carpal Bones - Capitate.** Captions as in Figure 4.13.

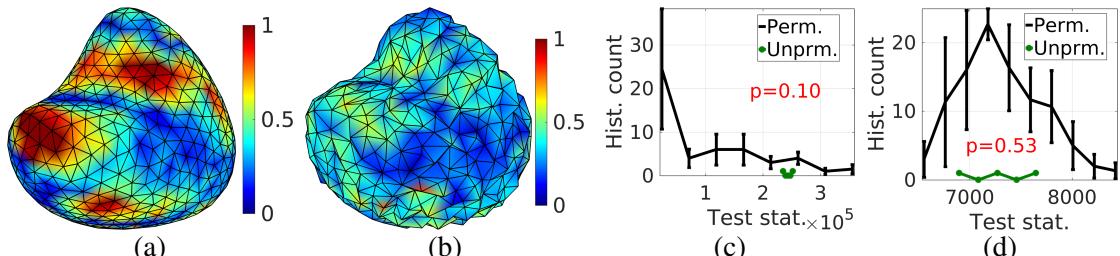


Figure 4.17: **Results: Hypothesis Testing on Carpal Bones - Lunate.** Captions as in Figure 4.13.

contrast, our method (Figure 4.14(c)-(d)) avoids adhoc smoothing using pre-processing, but rather accounts for the corruptions in the segmentations in the statistical model, and retains they key biological bone characteristics.

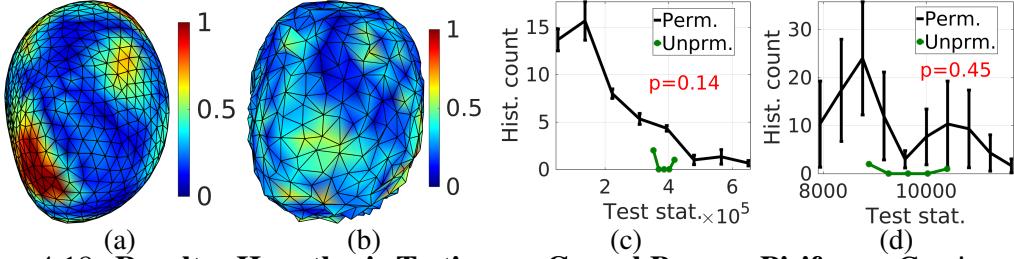


Figure 4.18: **Results: Hypothesis Testing on Carpal Bones - Pisiform.** Captions as in Figure 4.13.

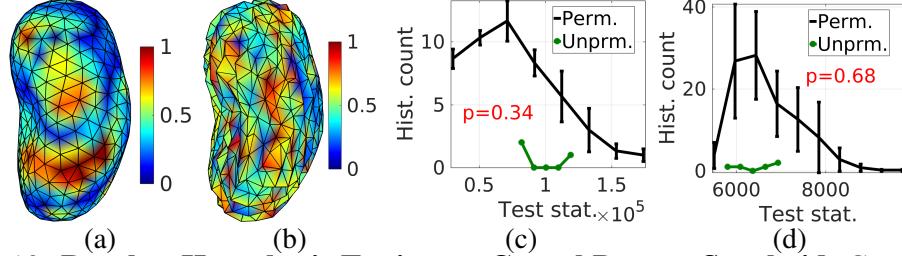


Figure 4.19: **Results: Hypothesis Testing on Carpal Bones - Scaphoid.** Captions as in Figure 4.13.

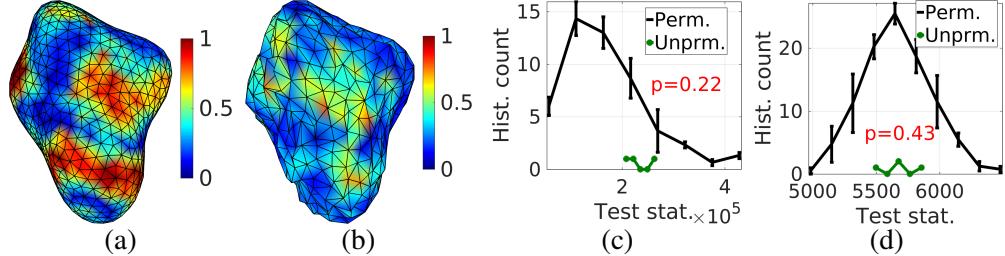


Figure 4.20: **Results: Hypothesis Testing on Carpal Bones - Trapezium.** Captions as in Figure 4.13.

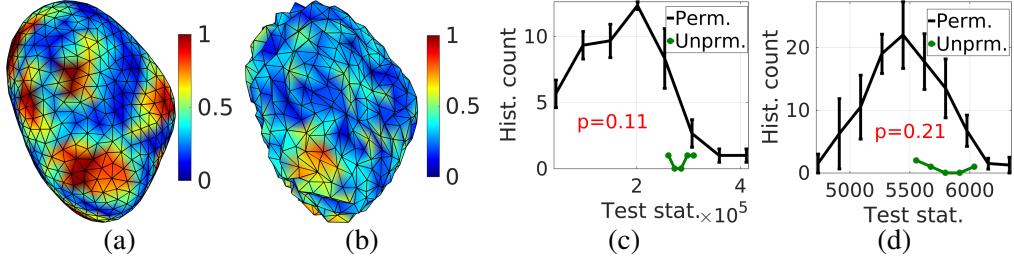


Figure 4.21: **Results: Hypothesis Testing on Carpal Bones - Triquetrum.** Captions as in Figure 4.13.

We perform the hierarchical shape model fitting and hypothesis testing on the other bone groups, i.e., trapezoid (Figure 4.15), capitate (Figure 4.16), lunate (Figure 4.17), pisiform (Figure 4.18), scaphoid (Figure 4.19), trapezium (Figure 4.20), and triquetrum (Figure 4.21). Overall, as seen in the figures, we find that our model leads to smoother means and smoother modes of variation, which leads to greater regularity in the Cohen's d effect size maps compared to ShapeWorks. Our compact models also lead to smaller p values than ShapeWorks. As discussed in Section 4.1, hierarchical modeling inherently

incorporates a safeguard against false positives in detecting inter-group differences. Thus, the lower p values from our approach can be considered to be reliable.

Chapter 5

Riemannian Shape Prior for Bayesian Image Segmentation

For a group of shapes, the learned shape model, say, $P(\cdot|z^m, C^m)$, gives (i) the mean shape z^m in Kendall shape space and (ii) the principal modes of variation of shape through the principal eigenvectors of C^m defined in the tangent space at z^m . We propose a novel formulation for object segmentation that constrains the estimated object shape to be represented within a subspace of the tangent space at the mean spanned by the principal eigenvectors of the covariance.

5.1 Object-Shape Representation and Segmentation Prior

In this section, for simplicity, we use the notation of μ being the *group* mean and C being the *group* covariance, unlike the previous sections where we used these symbols for the population level parameters. Let the top- K eigenvalues and eigenvectors of C be $\{\lambda_k\}_{k=1}^K$ and $\{v_k\}_{k=1}^K$. We propose to model true object shapes y using (i) a linear combination of the top eigenvectors $\{v_k\}_{k=1}^K$ with coefficients $\{w_k \in \mathbb{R}\}_{k=1}^K$, i.e., $\sum_{k=1}^K w_k v_k$, followed by the (ii) the exponential map $\text{Exp}_\mu(\cdot)$ with respect to the mean μ (Pennec (2006)). This work uses $K \leq 45$. Shapes y model an equivalence class of pointsets through similarity transforms \mathcal{S} , i.e., $\mathcal{S}\text{Exp}_\mu(\sum_{k=1}^K w_k v_k)$, which model object boundaries. We propose to model a PDF on true object shapes that promotes shapes that are representable using the top few eigenvectors of C , by penalizing coefficients w_k proportional to their magnitude and inversely proportional to the standard deviation $\sqrt{\lambda_k}$. Our novel shape prior model is $P(y = \text{Exp}_\mu(\sum_{k=1}^K w_k v_k)) := \zeta \exp(-\tau \sum_{k=1}^K |w_k| / \sqrt{\lambda_k})$, with normalizing constant ζ and free parameter $\tau \in \mathbb{R}_{>0}$.

5.2 Generative Model for Observed Image Data

We need to fit our model $P(y)$ to the observed data that is in the form of an image I having intensity-based feature F_p at a voxel p . Let the set of voxel locations be \mathcal{P} . Let functional entity F which take image I as input and produces Because we represent the object boundary by a pointset $\mathcal{S} \circ y$, we propose to first convert this representation into an image representation by assigning binary values to all voxels \mathcal{P} in the image, by assigning a value of 1 for p inside the object and a value of 0 for p outside the object. Let \mathcal{B} be this binarization operator, and let $\mathcal{B}(\mathcal{S} \circ y)_p$ denote the binary value at voxel p . Subsequently, we propose to measure the quality of fit of the object segmentation $\mathcal{B}(\mathcal{S} \circ y)$ to the observed feature image I as follows. Let image L model the segmentation labels, where $L_p \in \{0, 1\}$ is the label at voxel p , label 1 denotes the object's interior, and label 0 denotes the object's exterior. Then we can express likelihood of label L as $P_F(L|I, \theta)$, where F represents underlying model, which take image I as input and produces $P_F(L|I, \theta)$ as output with θ as parameters. Here θ is learned through training of model F with the training data of image I and corresponding label map L .

For segmentation purpose, we have two kinds of dataset, as (i) *simulated MRI data* containing MRI image generated from actual binary segmentations of ellipsoids and carpel bone, (ii) *clinical brain MRI data* containing actual brain MRI scans with respective expert segmentations of subcortical structures. For simulated MRI data we propose to model F as trained multi-atlas segmentation method with parameters θ_{MA} , and for clinical brain MRI data we model F as trained DNN with weight parameters θ_{DNN} . We propose to model the posterior PDF of object shape y , given the (i) learned Riemannian shape prior model parameterized by the mean μ , eigenvalues Λ , and eigenvectors V , (ii) learned F model on labels parameterized by θ , and (iii) observed image I , as

$$\log P(y = \text{Exp}_\mu(\sum_{k=1}^K w_k v_k) | \mathcal{S}, w, \mu, \Lambda, V, \theta, I) := \sum_{p \in \mathcal{P}} \left(\begin{array}{l} \mathcal{B}(\mathcal{S} \circ y)_p \log P_F(l_p = 1 | \theta, I) \\ + (1 - \mathcal{B}(\mathcal{S} \circ y)_p) \log P_F(l_p = 0 | \theta, I) \\ + (-\tau \sum_{k=1}^K |w_k| / \sqrt{\lambda_k}), \end{array} \right)$$

upto an additive term, i.e., the log normalizing constant. In this PDF term $\mathcal{B}(\mathcal{S} \circ y)_p \log P_F(l_p = 1 | \theta, I)$ will be zero if voxel p is in highest probability region of $P_F(L|I, \theta)$ otherwise a large negative number if voxel p is inside $\mathcal{S} \circ y$ and in lower probability region $P_F(L|I, \theta)$. Similarly, term $(1 - \mathcal{B}(\mathcal{S} \circ y)_p) \log P_F(l_p = 0 | \theta, I)$ will be zero if voxel p is in lowest probability region of $P_F(L|I, \theta)$ otherwise a large negative number if voxel p is outside $\mathcal{S} \circ y$ and in higher probability region $P_F(L|I, \theta)$.

5.3 Bayesian Formulation for Object Segmentation

We formulate object segmentation as maximum-a-posteriori (MAP) estimation on object shapes y , as

$$\arg \max_{\mathcal{S}, \{w_k\}_{k=1}^K} \log P(y = \text{Exp}_\mu(\sum_{k=1}^K w_k v_k) | \mathcal{S}, w, \mu, \Lambda, V, \theta, I)$$

The solution gives the object boundary as $\mathcal{S}^* \circ \text{Exp}_\mu(\sum_{k=1}^K w_k^* v_k)$, where \mathcal{S}^* is the optimal similarity transform and $\{w_k^*\}_{k=1}^K$ are the optimal coefficients.

5.4 Optimization Algorithm

We propose to optimize the posterior by alternately optimizing the object-boundary representation parameters \mathcal{S} and $\{w_k\}_{k=1}^K$. We propose a novel brute-force search for the parameters underlying object-boundary representation. We have (i) K parameters for the object-shape representation and (ii) the parameters underlying the similarity transform \mathcal{S} that we model using 3 parameters for translation, 1 parameter for isotropic scaling, and 3 parameters for rotation. We model rotation in 3D by an orthogonal matrix of determinant +1, which we model as the matrix exponential of a 3×3 skew symmetric matrix (Taylor & Kriegman (1994)) parameterized by its 3 off-diagonal terms. On these $(7 + K)$ scalar parameters representing the object boundary, we propose alternating coordinate descent using brute-force optimization over a chosen discretized set of the scalar parameter values within typical ranges.

5.5 Results: Segmentation

We present the results of object segmentation using our method for our shape prior. We have implemented our method and shape prior model in MATLAB programming environment without external library support. DNN models are implemented in python programming environment using Tensorflow-Keras. Multiatlas method is implemented in MATLAB programming environment.

5.5.1 Results: Segmentation on Simulated MRI data

We have simulated MRI image from binary segmentation by adding combinations of gaussian noise and salt-and-pepper noise. As explained in previous sections, we have model F as trained multiatlas segmentation method (Bai et al. (2013); Awate & Whitaker (2014)). We trained multiatlas segmentation using noisy version of binary segmentation to test robustness of our method.

We compare against two other methods: (i) Segmentation using a prior learned using ShapeWorks (SCI Institute (2013)): During learning, unlike our method, ShapeWorks assumes manual segmentations to be perfect, it constraints the point placement to lie only on the object boundary indicated by the manual segmentation, and it ignores the manifold structure of the shape space. During learning and segmentation, and it heuristically pre-aligns intensity images to a common image coordinate space, instead of (iteratively) optimally aligning the shape-model reference frame to the object in the image. (ii) Multi-atlas segmentation (Bai et al. (2013); Awate & Whitaker (2014)): that relies on nonlinear nonparametric diffeomorphic registration of a database of atlases to the test image and, then, performing patch-based label fusion to segment the test image. Multiatlas segmentation, however, fails to account for the human errors in image segmentation in the atlas database, which can lead to losses in performance, especially for small database sizes. We measure segmentation quality using (i) the Dice similarity coefficient (DSC) and (ii) the average of the shortest inter-surface distances between each point on the estimated object surface and the ground truth (or the expert segmentation, when “truth” unavailable for clinical data).

Results: Segmentation on Simulated Ellipsoids Data

Our hierarchical shape models learned on the ellipsoidal dataset in Section 4.2.1 provides a model for each group of shapes. We use one of the group models to segment corrupted images comprising objects with ellipsoidal shapes drawn from the same group. As described in Section 4.2.1, we simulate a group of 36 ellipsoids represented as binary segmentation images, where the true single mode of shape variation changes the length of the major axis of the ellipsoid over a specified range. Given the true binary segmentations, we simulate intensity images by blurring the binary (0–1) image and adding zero-mean, independent and identically distributed Gaussian random noise of standard deviation 0.3. To mimic errors in expert segmentations in practice, we introduce randomly generated coarse-scale and fine-scale perturbations (bumps and pits at multiple scales) on the true ellipsoidal surface. This group of pairs of (noisy) intensity and (perturbed) label images forms our *training set* that multiatlas segmentation uses as the atlas database, while our method and ShapeWorks learn the shape prior from the perturbed segmentations. Our *test set* is another group of 30 ellipsoid intensity images (Figure 5.1(a)) created using the same procedure as our training set. We treat output of multiatlas segmentation on test set, using training set atlas database, as $P_F(L|I, \theta)$ as mentioned in 5.2. Compared to other methods (Figure 5.1(c)-(d)), our framework produces the highest DSC values and the lowest inter-surface distances, evaluated against the ground truth (Figure 5.1(h)), signifi-

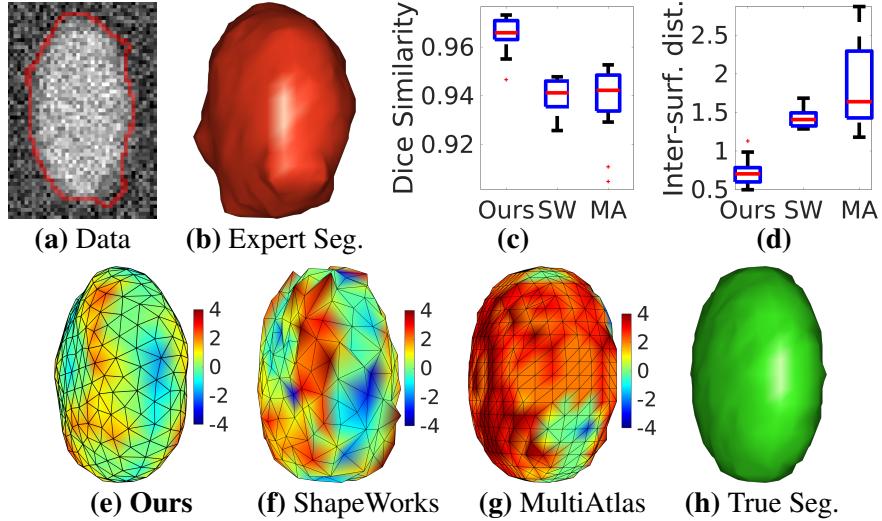


Figure 5.1: Results: Segmentation on Simulated Data. (a) A 2D slice of a simulated 3D image. (b) Expert segmentation. Box plots showing methods’ performances across dataset, measured by (c) Dice similarity coefficient and (d) average of the shortest inter-surface distances between each point on estimated object surface and ground-truth. Estimated 3D segmentation’s surface rendering, with each surface location colored by the shortest signed distance to ground-truth surface (in terms of percentage of ground-truth diameter), for: (e) our method, (f) ShapeWorks (SW) shape model, (g) multiatlas (MA) segmentation with nonlinear nonparametric registration, and (h) ground-truth segmentation.

cantly improving over ShapeWorks and multiatlas segmentation. Qualitatively, the object boundaries from our segmentations (Figure 5.1(e)) are closest to the ground truth (Figure 5.1(h)), compared to ShapeWorks (Figure 5.1(f)) and multiatlas segmentation (Figure 5.1(g)). ShapeWorks’s estimate of the object boundary lacks smoothness because of the unreliable learning stemming from erroneous expert segmentations, which leads to errors in its prior model’s mean and modes of variation. Multiatlas segmentation also suffers because of (i) errors in measuring patch similarity, stemming from image noise and artifacts, and (ii) errors in label fusion stemming from the errors in expert segmentations.

Results: Segmentation on Carpal Bone Data

Analogous to the experiments with ellipsoidal shapes in Section 5.5.1, we use the carpal bone segmentations to generate corrupted intensity images as data (note: this dataset does *not* provide the original medical images, but only their high-quality segmentations). We then use the shape models learned in Section 4.2.2 to segment the bone shapes from the generated images. We perform these experiments on three examples of carpal bones, i.e., trapezoid, hamate, and pisiform. For trapezoid shapes (Figure 5.2), hamate shapes

(Figure 5.3), and pisiform shapes (Figure 5.4), we find that our method produces segmentations that have higher DSC values and lower inter-surface distances. Qualitatively, our method also produces smoother and more realistic shapes.

5.5.2 Results: Segmentation in Clinical Brain Subcortical MRI

We evaluate using 2 databases with clinical brain MRI (IRB approved studies) and radiologist-provided segmentations of the thalamus, the caudate, the globus, the hippocampus, and the putamen. One dataset, acquired at Tata Memorial Hospital in Mumbai, has 30 images with low-quality segmentations because of a constrained time budget—this is the *training* database for all models. Another dataset, from the National

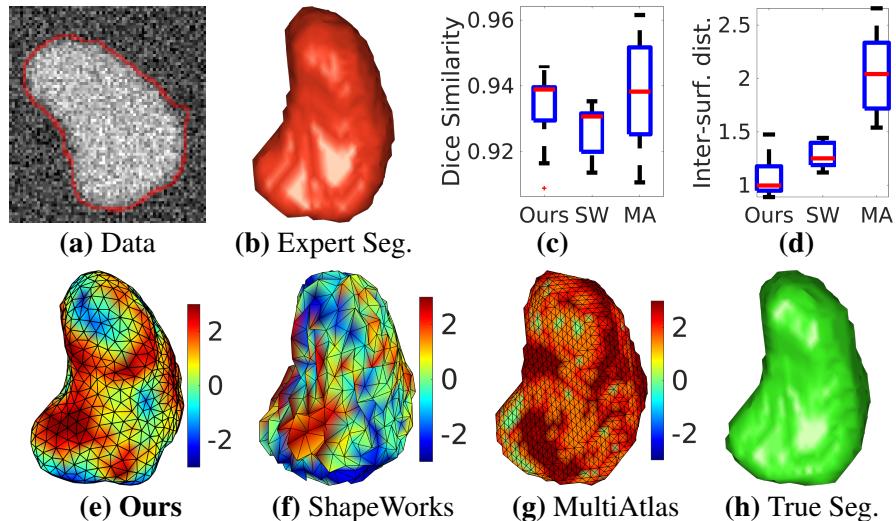


Figure 5.2: **Results: Segmentation on Carpal Bone Shapes: Trapezoid.** Captions as in Figure 5.1.

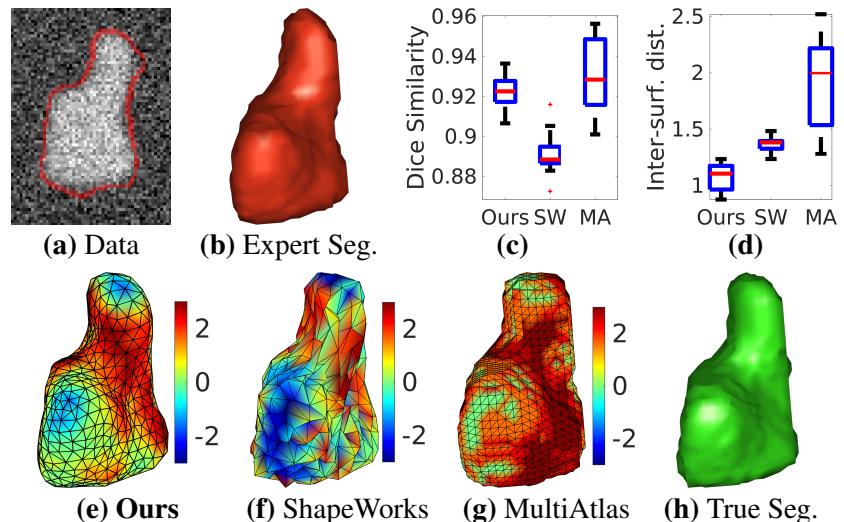


Figure 5.3: **Results: Segmentation on Carpal Bone Shapes: Hamate.** Captions as in Figure 5.1.

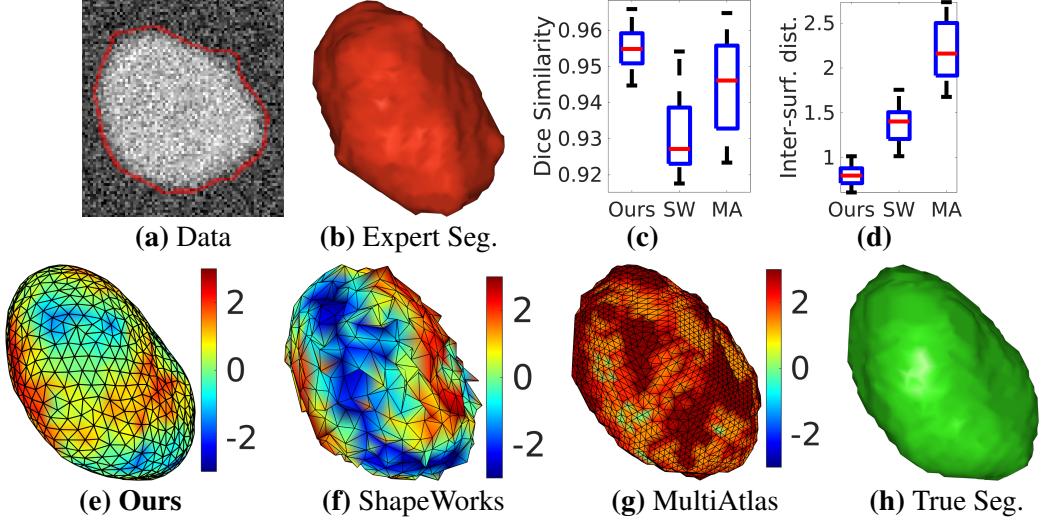


Figure 5.4: **Results: Segmentation on Carpel Bone Shapes: Pisiform.** Captions as in Figure 5.1.

Alliance on Medical Image Computing (www.na-mic.org), has 70 images with high-quality segmentations—we use 10 images as *validation set* (to tune hyperparameters for each method) and 60 images for *evaluation*. We compare our framework that couples a 3D Unet with the sparse Riemannian shape prior, as in chapter 5, with 4 other frameworks: (i) *Unet*: 3D Unet (Ronneberger et al. (2015)) with typical data augmentation; (ii) *SR-Unet*: Shape-regularized Unet (Ravishankar et al. (2017)), which we extend to 3D; (iii) *Unet+SW*: Unet coupled with a shape prior learned from ShapeWorks (SCI Institute (2013)); and (iv) *MA*: Multiatlas segmentation that relies on nonlinear nonparametric diffeomorphic registration of the (training) database of atlases to the test image and, then, performing patch-based label fusion to segment the test image. We evaluate performance using: (i) Dice similarity coefficient (DSC) and (ii) mean of the shortest inter-surface distances between each point on the estimated object surface and the surface indicated by the high-quality segmentation (ground truth).

As expected, the Unet faces challenges in learning from a training set that is small sized and has low-quality segmentations. Thus, we see reduced regularity (Figures 5.5(f), 5.6(f), 5.7(f), 5.8(f), 5.9(f)) and quantitative performance (Figures 5.5(b)-(c), 5.6(b)-(c), 5.7(b)-(c), 5.8(b)-(c), 5.9(b)-(c)), compared to the typical cases in the literature where the training set is higher quality and larger.

The SR-Unet (Ravishankar et al. (2017)) infuses additional information in the learning, which aims to implicitly capture the statistics of shape variability and regularity, similar to other such approaches (Arif et al. (2017)). While for the thalamus, SR-Unet’s results improve over those of the Unet, SR-Unet’s performance for the caudate is statistically similar to that of Unet (Figures 5.5(b)-(c), 5.6(b)-(c)). For globus, hippocampus, and

putamen, SR-Unet's quantitative performance is similar to that of Unet (Figures 5.7(b)-(c), 5.8(b)-(c), 5.9(b)-(c)), but in case of globus, it produces over-smoothed shape (Fig-

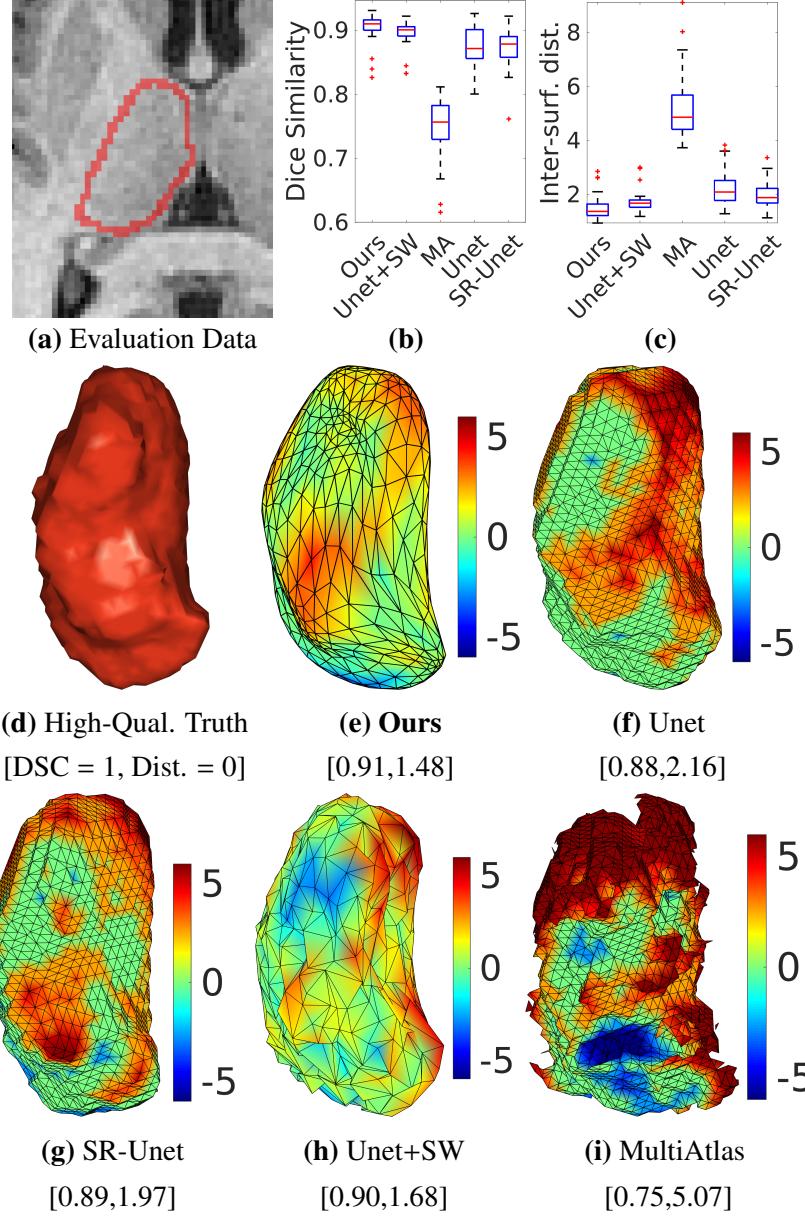


Figure 5.5: Results on Clinical Brain MRI: Subcortical Structure Thalamus. (a) A 2D slice of the 3D test image. Box plots showing methods' performances across dataset, measured by (b) Dice similarity coefficient, (c) mean of the shortest inter-surface distances between each point on estimated object surface and ground-truth. (d) High-quality ground-truth segmentation. Estimated segmentation's surface rendering, with each surface location colored by the shortest signed distance to ground-truth surface (in terms of percentage of ground-truth diameter), for (e) **Our method**; (f) Unet (Ronneberger et al. (2015)); (g) Shape-regularized Unet (Ravishankar et al. (2017)), extended to 3D; (h) Unet + ShapeWorks prior model; (i) Multiatlas segmentation.

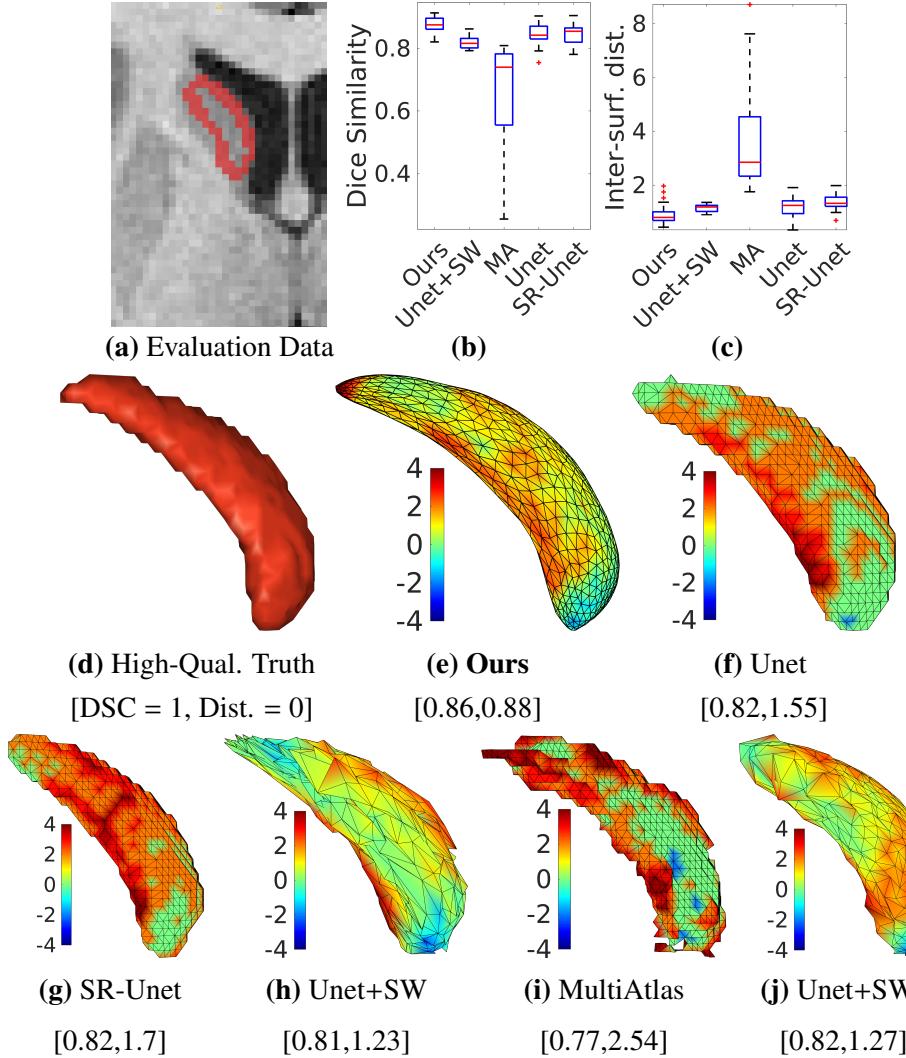


Figure 5.6: **Results on Clinical Brain MRI: Subcortical Structure Caudate.** Captions (a)–(i) same as in Figure 5.5. (j) Segmentation with Unet+SWs, where ShapeWorks’s prior learning extra smooths segmentations in low-quality training set.

ure 5.7(f)-(g)) compared to that of UNet. In case of hippocampus and putamen, SR-UNet produces segmentations with less regularity and smoothness compared to that of UNet (Figures 5.8(f)-(g), 5.9(f)-(g)). In this way, DNN approaches can face difficulty in optimizing the large number of underlying parameters, especially from imperfectly-curated small-sized training sets. While transfer-learning schemes can, in principle, alleviate this problem to some extent, typical such schemes (Ataloglu et al. (2019); Nogovitsyn & et al. (2019)) mainly focus on hippocampus segmentation for which large high-quality training sets are available—thus, in practice it is unlikely that DNN features and mappings learned for segmenting one anatomical structure will generalize well to other structures associated with different contrasts, shapes, and sizes.

We find that Unet+SW significantly improves quantitative performance over the Unet for the Thalamus (Figure 5.5(b)–(c)), but the quantitative improvement is unclear in

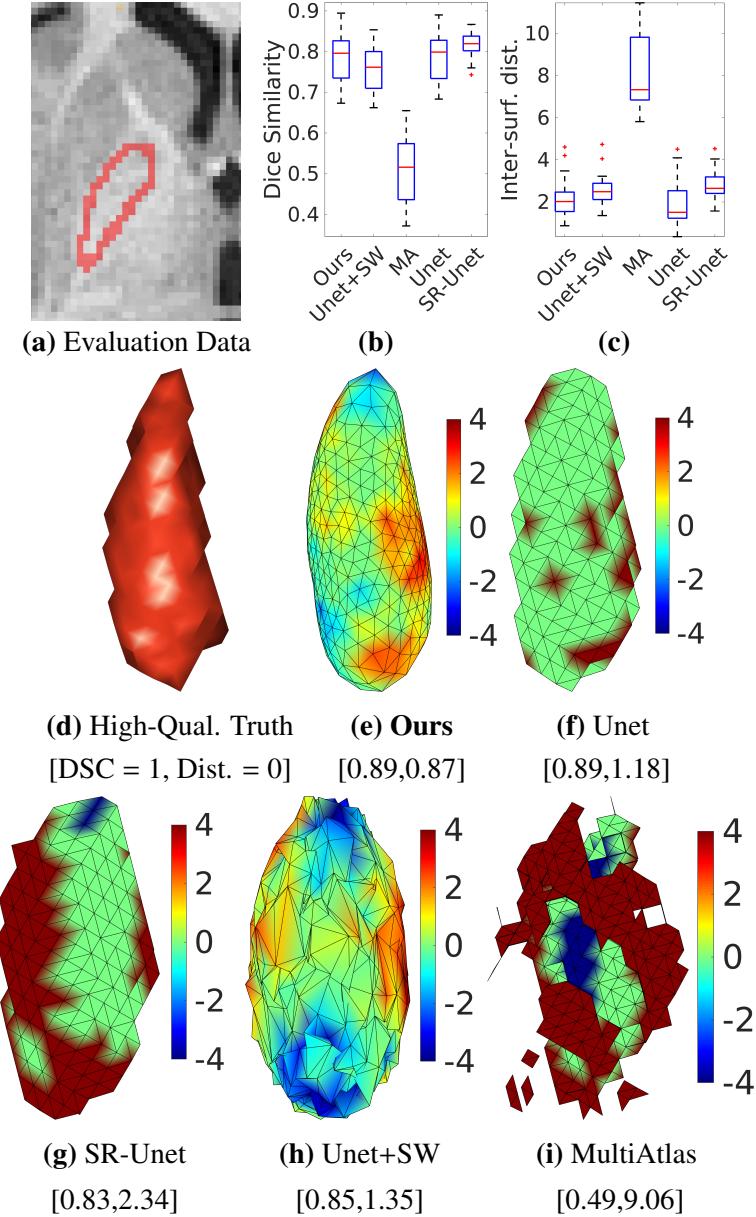


Figure 5.7: **Results on Clinical Brain MRI: Subcortical Structure Globus.** Captions (a)–(i) same as in Figure 5.5.

case of caudate, globus, hippocampus, and putamen (Figures 5.6(b)-(c), 5.7(b)-(c), 5.8(b)-(c), 5.9(b)-(c)). Qualitatively, Unet+SW produces segmentations with significantly less regularity and smoothness for all five brain subcortical structures in study (Figure 5.5(h), 5.6(h), 5.7(h), 5.8(h), 5.9(h)). The learning scheme for ShapeWorks, unlike our method, (i) assumes segmentations to be high-quality, (ii) forces points to lie on the object boundary indicated by the (low-quality, slightly erroneous) segmentation, and (iii) models shape space as a linear manifold. During learning and inference, ShapeWorks pre-aligns intensity images to a common image coordinate space, unlike our approach of (iteratively) aligning the shape-model reference frame to the object during MAP optimization. This,

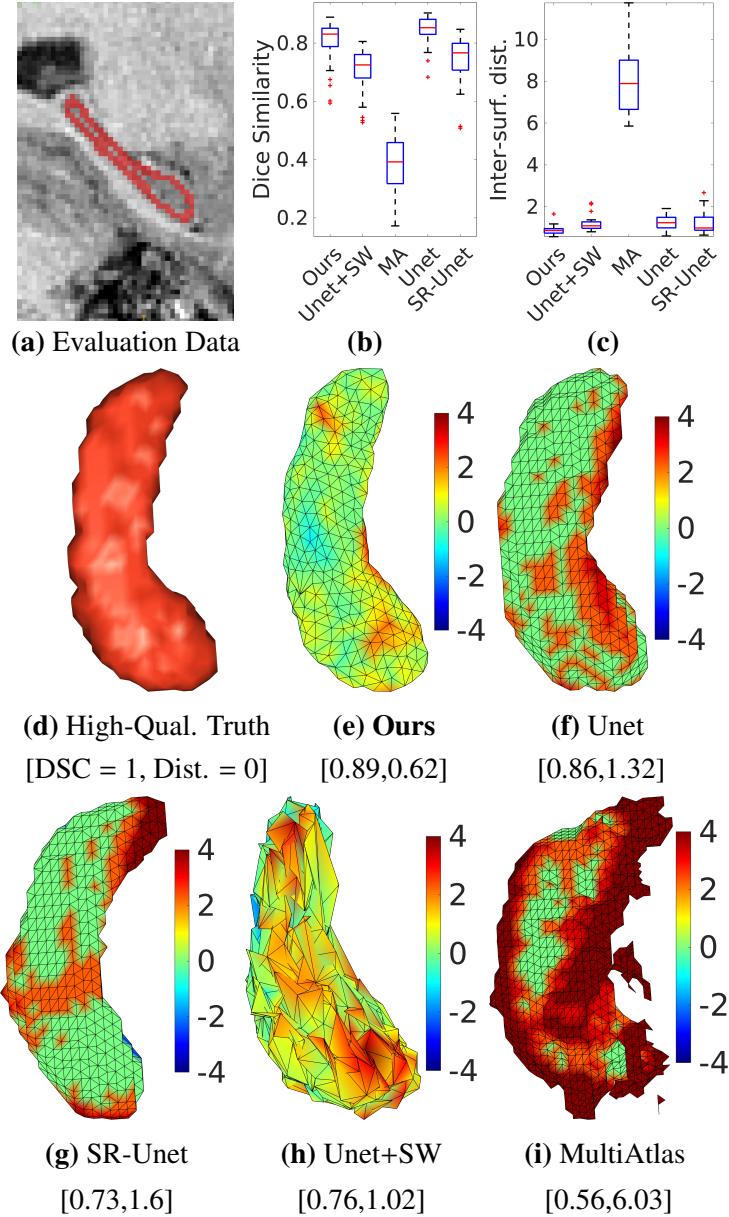


Figure 5.8: Results on Clinical Brain MRI: Subcortical Structure Hippocampus.

Captions (a)–(i) same as in Figure 5.5.

we find that the inherent (albeit small) errors in expert segmentations cause problems for ShapeWorks by leading to artificially inflated estimates of shape variability (Shigwan & Awate (2016)), unlike our approach that uses a Bayesian approach during prior learning and allows points placements slightly off the expert-segmented object boundary. We find that when ShapeWorks learns a prior from extra-smoothed low-quality segmentations, using surface smoothing to compensate for curation errors, it can lead to the loss of the ability of the prior model (mean shape and covariance) to model subtle features and variations in shapes, e.g., the high curvature region and the bending in the caudate (Figure 5.6(j)). Multiatlas segmentation (Figures 5.5(i), 5.6(i), 5.7(i), 5.8(i), 5.9(i)) introduces

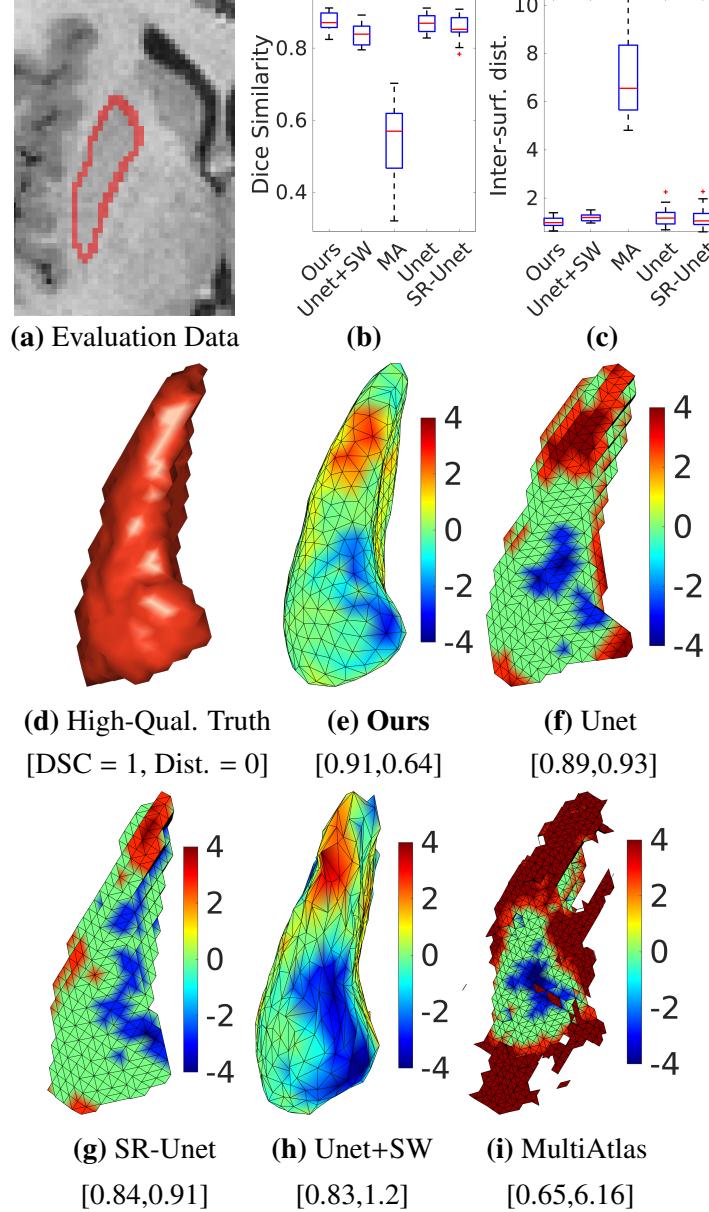


Figure 5.9: **Results on Clinical Brain MRI: Subcortical Structure Putamen.** Captions (a)–(i) same as in Figure 5.5.

some errors from imperfect label-fusion because of the low-quality expert segmentations in the atlas database; consistent with the analysis using a small-sized (even though good-quality) atlas database in (Asman & Landman (2013)).

Unlike the pure-DNN methods (Unet, SR-Unet) and the DNN coupled with a shape prior from ShapeWorks (Unet+SW), our method couples the Unet with a sparse Riemannian prior that explicitly handles errors in low-quality segmentations in the training set. DSCs and inter-surface distances, evaluated against high-quality expert segmentations (Figures 5.5(d), 5.6(d), 5.7(d), 5.8(d), 5.9(d)), for our method are better than those from other methods (Figures 5.5(b)–(c), 5.6(b)–(c), 5.7(b)–(c), 5.8(b)–(c), 5.9(b)–(c)). Qual-

itatively, the object boundaries estimated by our method (Figures 5.5(e), 5.6(e), 5.7(e), 5.8(e), 5.9(e)) show improved regularity, smoothness, and accuracy than those estimated by other methods (Figures 5.5(f)–(i), 5.6(f)–(i), 5.7(f)–(i), 5.8(f)–(i), 5.9(f)–(i)). Some DNN-based results show a bias in the estimated boundaries, distorting the object’s true shape and pose.

Chapter 6

Hierarchical Shape Clustering in Shape Space

In this chapter, we will see a modeling method for clustering given set of ordered pointsets representing shape. After modeling the distribution over a group of shapes, we will model a distribution over group mean, which effectively acts as a prior over group means. This is a hierarchical model over shapes. Finally, we will introduce EM algorithm for hierarchical clustering model in shape space.

Until now we have seen single probabilistic Gaussian shape models over set of shapes, now we will build a hierarchical shape model of Gaussian mixture distribution. Figure 6.1 shows the model.

6.1 Mathematical Notation

We consider N individuals in population. Let this population have M (unknown) groups/clusters. Let data $\mathbf{x} := \{x_n\}_{n=1}^N$ represent the binary image of the segmented anatomical object in individual n . Let $\mathbf{y} := \{y_n\}_{n=1}^N$ be the (unknown) pointset representing object shape for individual n . Let $\mathbf{Y} := \{Y_n\}_{n=1}^N$ be corresponding hidden object shape variable for individual n . Let $\mathbf{z} := \{z_m\}_{m=1}^M$ be the (unknown) pointset representing the mean object shape for cluster m . Let $\bar{C} := \{\bar{C}_m\}_{m=1}^M$ (unknown) model the covariance of shapes in cluster m . Let μ be the (unknown) pointset representing the population-level mean object shape and let (unknown) C model the associated covariance; μ and C capture the variability of the group-mean shapes z_m . Let $\bar{v} := \{v_n \in \{1, 2, \dots, M\}\}_{n=1}^N$ be the (hidden-unknown) cluster-label associated data x_n . Let $\bar{w} := \{\bar{w}_m\}_{m=1}^M$ (unknown) be mixture weight for m -th cluster, with constraints $\sum_{m=1}^M w_m = 1$ and $w_m \geq 0$. Let β (user-defined) repre-

sent smoothness prior of distribution on mean shapes \mathbf{z} . Let β_m (user-defined) represent smoothness prior of mixture component for cluster m .

Each shape-representing pointset (or *shape pointset*) has J points in 3D; so $y_n \in \mathbb{R}^{3J}$, $z_m \in \mathbb{R}^{3J}$.

Let $\theta := \{\mu, C, \bar{C}, \bar{w}\}$. Since we have special hierarchical structure, as illustrated in figure 6.1, we optimize probability of data given parameters as $\max_{\mathbf{z}, \theta} P(\mathbf{x}|\mathbf{z}, \theta)P(\mathbf{z}|\theta)$ instead of $\max_{\mathbf{z}, \theta} P(\mathbf{x}|\mathbf{z}, \theta)$, so that we can take advantages of hierarchical structure, as explained in section 4.1. This additional term $P(\mathbf{z}|\theta)$ provides benefits of hierarchical model along with common coordinate system of shape registration, as we can refer alignment of μ universal. This optimization can be further elaborated as follow,

$$\max_{\mathbf{z}, \theta} P(\mathbf{x}|\mathbf{z}, \theta)P(\mathbf{z}|\theta) = \max_{\mathbf{z}, \theta} \int P(\mathbf{x}, \mathbf{Y}, \bar{v}|\mathbf{z}, \theta)P(\mathbf{z}|\theta)d\mathbf{Y}d\bar{v} \quad (6.1)$$

Now by Bayes rule and given each data individual is independent and identically distributed, we get following split.

$$P(\mathbf{x}, \mathbf{Y}, \bar{v}|\mathbf{z}, \theta)P(\mathbf{z}|\theta) = \prod_{n=1}^N P(x_n|y_n)P(y_n|v_n, \mathbf{z}, \bar{C})P(v_n|\theta)P(\mathbf{z}|\mu, C, \beta) \quad (6.2)$$

Now, by Markov independent assumption, objective function in eq(6.1) can be written as

$$\prod_{n=1}^N \sum_{m=1}^M \int P(x_n, Y_n, v_n = m|\mathbf{z}, \theta)P(\mathbf{z}|\theta)dY_n = \prod_{n=1}^N \sum_{m=1}^M \int P(x_n, Y_n, v_n = m|z_m, \theta)P(\mathbf{z}|\theta)dY_n \quad (6.3)$$

$$= \prod_{n=1}^N \sum_{m=1}^M \int P(x_n|Y_n)P(Y_n|v_n = m, z_m, C_m)P(v_n = m|\theta)P(\mathbf{z}|\mu, C, \beta)dY_n \quad (6.4)$$

We use equation (3.1) to model (i) conditional PDF $P(z_m|\mu, C, \beta)$ of group mean shapes z_m (ii) the conditional PDF $P(y_n|z_m, C_m, \beta_m)$ of individual shapes y_n (Shigwan & Awate (2016)). We model $w_m := P(v_n = m|\theta)$ as a parameter of model, which is same for each individual data x_n .

6.2 Likelihood model

The proposed model relies on a dissimilarity measure $\Delta(x_n, y_n)$ between binary segmentation x_n and shape pointset y_n , as defined in equation (3.3).

We model $P(x_n|y_n) := \exp(-\Delta(x_n, y_n))/\kappa$, with normalization constant κ .

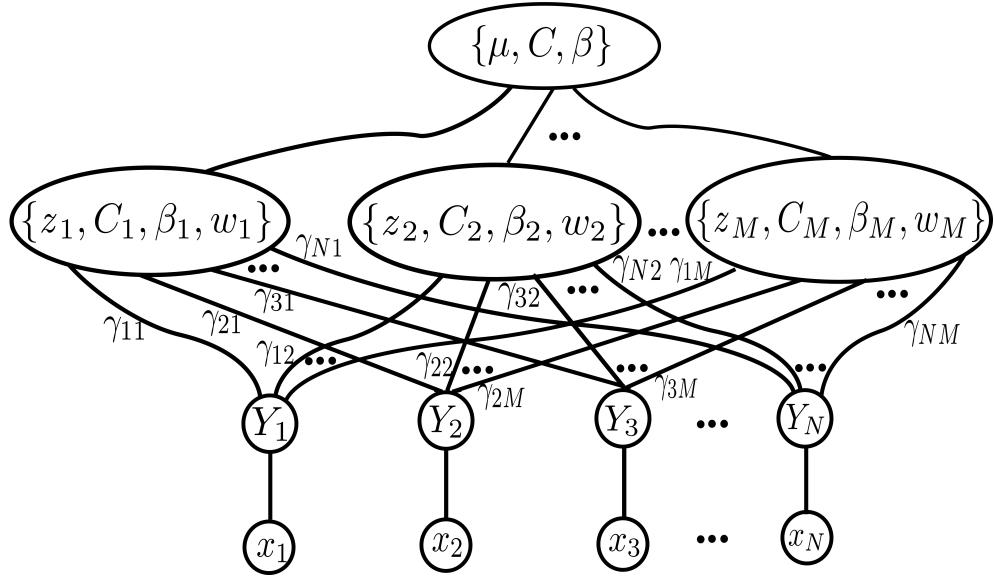


Figure 6.1: Hierarchical Model

6.3 Model fitting

We use Monte-Carlo Expectation-Maximization (MC-EM) to find optimal parameters of log-likelihood of objective function in eq.(6.4).

E-step.

$$\begin{aligned}
 Q(\mathbf{z}, \theta; \mathbf{z}^i, \theta^i) &= \sum_{n=1}^N \mathbb{E}_{P(y_n, v_n | x_n, \mathbf{z}^i, \theta^i)} [\log(P(x_n, y_n, v_n | \mathbf{z}, \theta) P(\mathbf{z} | \theta))] \\
 &= \sum_{n=1}^N \sum_{m=1}^M \int_{y_n} P(y_n, v_n = m | x_n, z_m^i, \theta^i) [\log(P(x_n | y_n) P(y_n | v_n = m, z_m, C_m) P(v_n = m | \theta) P(\mathbf{z} | \mu, C))] dy_n \\
 &= \sum_{n=1}^N \sum_{m=1}^M \int_{y_n} P(y_n, v_n = m | x_n, z_m^i, \theta^i) [\log(P(x_n | y_n) P(y_n | v_n = m, z_m, C_m) P(v_n = m | \theta) P(z_m | \mu, C))] dy_n + \text{constant}.
 \end{aligned}$$

By Bayes rule, we get

$$\begin{aligned}
 P(y_n, v_n = m | x_n, z_m^i, \theta^i) &= \left[\frac{P(x_n, y_n | v_n = m, z_m^i, \theta^i)}{P(x_n | z_m^i, \theta^i)} \right] P(v_n = m | \theta^i) \\
 &= \left[\frac{P(x_n, y_n | z_m^i, v_n = m, \theta^i)}{P(x_n | \theta^i)} \right] w_m^i
 \end{aligned}$$

Hence by neglecting constant term,

$$\begin{aligned} Q(\mathbf{z}, \theta; \mathbf{z}^i, \theta^i) &= \sum_{n=1}^N \sum_{m=1}^M \int_{y_n} \frac{P(x_n, y_n | z_m^i, \nu_n = m, \theta^i)}{P(x_n | z_m^i, \theta^i)} w_m^i [\log(\\ &\quad P(x_n | y_n) P(y_n | \nu_n = m, z_m, C_m, \beta_m) w_m P(z_m | \mu, C, \beta))] dy_n \end{aligned} \quad (6.5)$$

Let γ_{nm} represent membership of data x_n to cluster m .

$$\begin{aligned} \gamma_{nm} &= P(\nu_n = m | x_n, z_m^i, \theta^i) \\ &= \int_{y_n} P(y_n, \nu_n = m | x_n, z_m, \theta^i) dy_n \\ &= \int_{y_n} \frac{P(x_n, y_n | \nu_n = m, z_m^i, \theta^i)}{P(x_n | z_m^i, \theta^i)} P(\nu_n = m | \theta^i) dy_n \\ &= \int_{y_n} \frac{P(x_n, y_n | \nu_n = m, z_m^i, \theta^i)}{P(x_n | z_m^i, \theta^i)} w_m^i dy_n \\ &= \int_{y_n} \frac{P(x_n | y_n) P(y_n | \nu_n = m, z_m^i, \theta^i)}{P(x_n | z_m^i, \theta^i)} w_m^i dy_n \end{aligned} \quad (6.6)$$

Now membership γ_{nm} in eq(6.6) is intractable to calculate exactly, so we approximate integration by Monte-Carlo Sampling. Let S -samples drawn from joint distribution $\{y_n^s\}_{s=1}^S \sim P(Y_n | \nu_n = m, z_m^i, \theta^i)$, then we define γ_{nm}^s as follow.

$$\gamma_{nm}^s := \frac{P(x_n | y_n^s) w_m^i}{\sum_{m'=1}^M \sum_{s'=1}^S P(x_n | y_n^{s'}) w_{m'}^i} \quad (6.7)$$

Finally $Q(\mathbf{z}, \theta; \mathbf{z}^i, \theta^i)$ becomes,

$$\begin{aligned} Q(\mathbf{z}, \theta; \mathbf{z}^i, \theta^i) &= \sum_{n=1}^N \sum_{m=1}^M \sum_{s=1}^S \gamma_{nm}^s [\log(\\ &\quad P(x_n | y_n^s) P(y_n^s | \nu_n = m, z_m, C_m, \beta_m) w_m P(z_m | \mu, C, \beta))] \end{aligned} \quad (6.8)$$

For sampling from $P(Y_n | \nu_n = m, z_m^i, \theta^i)$, we used our novel Riemannian leapfrog sampling method, introduced in section 3.6.

M-step.

We will find optimal parameters $\{\mathbf{z}, \theta\}$ by maximizing $Q(\mathbf{z}, \theta; \mathbf{z}^i, \theta^i)$ in eq(6.8). Update for $\{\mu, C, \{C_m\}_{m=1}^M\}$ are similar to updates in (Shigwan & Awate (2016)).

Update μ is given by

$$\min_{\mu} \frac{1}{\sum_{n'=1}^N \sum_{m'=1}^M \sum_{s'=1}^S \gamma_{n'm'}^{s'}} \sum_{s=1}^S \sum_{m=1}^M \left(\sum_{n=1}^N \gamma_{nm}^s \right) \text{Log}_{\mu}(z_m)^{\top} C^{-1} \text{Log}_{\mu}(z_m) \text{ such that } \|\mu\|_2 = 1. \quad (6.9)$$

Update z_m is given by

$$\min_{z_m} \frac{1}{\sum_{n'=1}^N \sum_{s'=1}^S \gamma_{n'm}^{s'}} \sum_{s=1}^S \left(\left(\sum_{n=1}^N \gamma_{nm}^s \text{Log}_{z_m}(y_n^s)^\top C_m^{-1} \text{Log}_{z_m}(y_n^s) \right) + \text{Log}_\mu(z_m)^\top C^{-1} \text{Log}_\mu(z_m) \right)$$

such that $\|z_m\|_2 = 1$.

(6.10)

We use projected gradient descent (A.2 derives the gradient), where the projection rescales the updated means μ and $\{z_m\}_{m=1}^M$ to unit norm, as explained in section 3.5.

Update C is given by,

$$C'^{new} = \frac{1}{\sum_{n'=1}^N \sum_{m'=1}^M \sum_{s'=1}^S \gamma_{n'm'}^{s'}} \sum_{m=1}^M \sum_{s=1}^S \left(\sum_{n=1}^N \gamma_{nm}^s \right) \text{Log}_\mu(z_m) \text{Log}_\mu(z_m)^\top$$

$$C^{new} = [(C'^{new})^{-1} - \beta \Omega]^{-1}$$

(6.11)

Update C_m is given by,

$$C_m'^{new} = \frac{1}{\sum_{n'=1}^N \sum_{s'=1}^S \gamma_{n'm}^{s'}} \sum_{s=1}^S \sum_{n=1}^N \gamma_{nm}^s \text{Log}_{z_m}(y_n^s) \text{Log}_{z_m}(y_n^s)^\top$$

$$C_m^{new} = [(C_m'^{new})^{-1} - \beta_m \Omega]^{-1}$$

(6.12)

Update w_m is given by,

w_m has additional constraints that $\sum_{m=1}^M w_m = 1$ and $w_m \geq 0$. If we maximize $Q(\mathbf{z}, \theta; \mathbf{z}^i, \theta^i)$ in eq(6.8) with these constraints, then we get following update.

$$w_m = \frac{\sum_{n=1}^N \gamma_{nm}}{\sum_{m'=1}^M \sum_{n'=1}^N \gamma_{n'm'}} = \frac{\sum_{n=1}^N \gamma_{nm}}{N}$$

(6.13)

6.4 Results: Riemannian Mixture Shape Clustering

We present the results of shape clustering using our method for hierarchical shape modeling and inference. We have implemented our model in MATLAB programming environment without any external library support. We have compared our method with a state-of-the-art pointset-based shape analysis approach named VBMixPCA (Gooya et al. (2018)) that differs from our approach in many ways as follows. First, VBMixPCA constraints the points to be placed (at real-valued sub-voxel locations) on the zero level set associated with the distance transform of the object boundary represented by the segmentation; in this way, it becomes sensitive to the (inevitable) errors in manual segmentation. Second, it does *not* rely on a hierarchical modeling approach. Third, it assumes an initial alignment of the segmentations and does *not* align them during inference. Fourth, VBMixPCA treats the group means as parameters instead of random variables. To evaluate

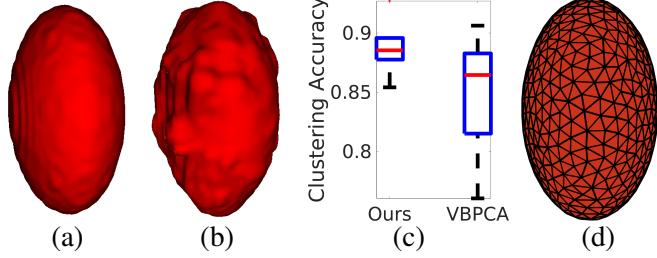


Figure 6.2: **Results: Shape Clustering on Simulated Ellipsoidal Data.** (a) Ground truth segmentation. (b) Corrupted segmentation (data), mimicking human errors. (c) Accuracy boxplot over 5 noisy instances of ellipsoid data for *Ours* and VBMixPCA. (d) Estimated population cluster mean from our method.

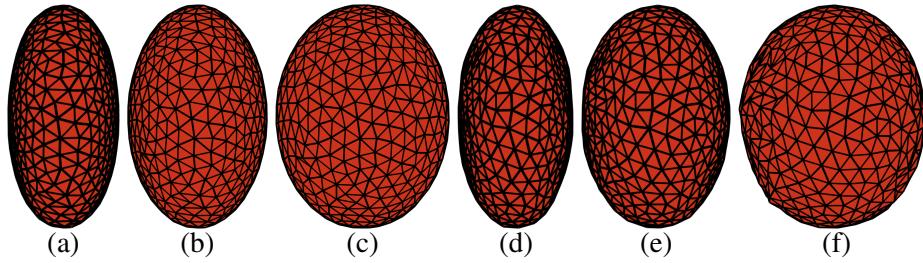


Figure 6.3: **Results: Shape Clustering on Simulated Ellipsoidal Data.** (a)-(c) Our method’s MAP estimates for group means z_1 , z_2 and z_3 . (d)-(f) VBMixPCA’s estimates for group means z_1 , z_2 and z_3 .

VBMixPCA, we first align all the images using similarity-transform based registration, and then give the aligned segmentations, of all groups, to VBMixPCA to compute the optimal point placement and shape mean, modes for each group. Now, we make the scale of the mean-shape pointsets and individual-shape pointsets obtained by VBMixPCA commensurate to that of our approach as follows. For each group, we center the mean shape so that its centroid is at the origin, rescale the mean shape to unit norm, and then apply the same group-specific centering and rescaling to each member of the group. We evaluate the methods on simulated ellipsoid data.

6.4.1 Results: Shape Clustering on Simulated Ellipsoidal Data

As described in Section 4.2.1, we simulate three groups of 32 ellipsoids represented as binary segmentation images, where the true single mode of shape variation changes the length of the major axis of the ellipsoid over a specified range.

The range of the variations across consecutive groups slightly overlap. We also introduce surface perturbations in the form of bumps and pits to these binary segmentations to mimic human errors in labeling (Figure 6.2(a)-(b)). The estimates of group means (Figure 6.3) and modes of variation (Figure 6.4, 6.5, 6.6) from our method indicate (i) smoother and more accurate group means and (ii) modes of variation that indicate

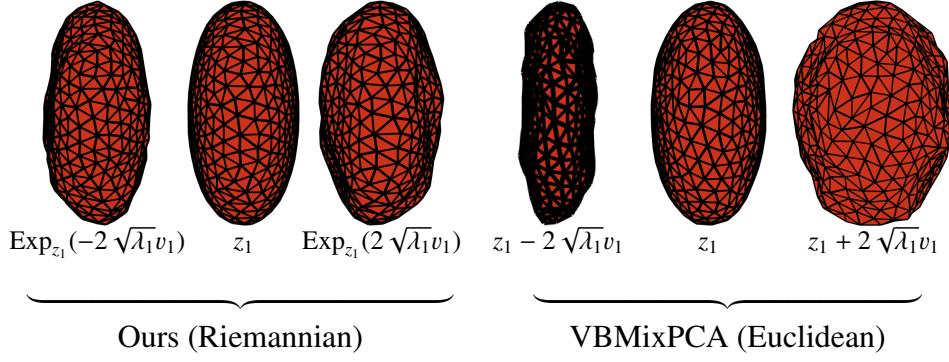


Figure 6.4: Results: Shape Clustering on Simulated Ellipsoidal Data. Estimated principal mode of variation of group-1 for our method and VBMixPCA.

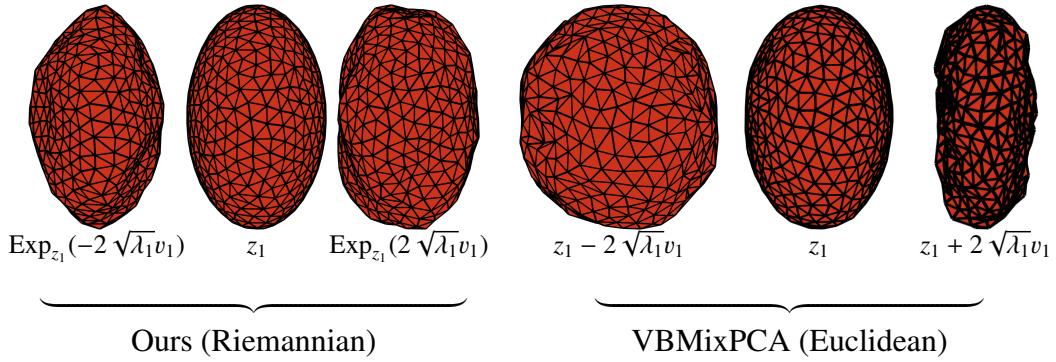


Figure 6.5: Results: Shape Clustering on Simulated Ellipsoidal Data. Estimated principal mode of variation of group-2 for our method and VBMixPCA.

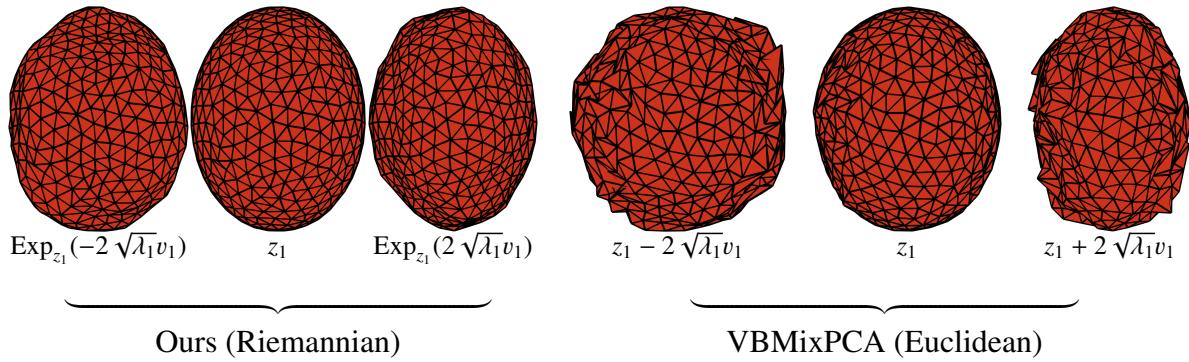


Figure 6.6: Results: Shape Clustering on Simulated Ellipsoidal Data. Estimated principal mode of variation of group-3 for our method and VBMixPCA.

smoother and more accurate shape variations, compared to VBMixPCA, which has captured more noisy due to absence of internal regularization. Clustering accuracy are shown in Figure 6.2(c). Clustering accuracy Acc is defined as follows

$$Acc = \frac{\text{count of correctly predicted label}}{\text{count of total individuals}}, \quad (6.14)$$

and predicted label of individual is nothing but a label, which has maximum membership for that individual.

Chapter 7

Conclusion

Several approaches to shape modeling, e.g., diffeomorphic deformation models, medial models, point distribution models, etc., are active research areas in their own right, each with its benefits and limitations. This work proposes a novel *hierarchical generative model for statistical shape analysis using point distribution models*, which leverages the characteristics of the pointsets lying in Kendall shape space. Our model is aware that, during learning, the manual segmentations will naturally exhibit human errors in practical deployment scenarios. The hierarchical modeling can help regularize the estimates of group distributions, through the effect of population-level variables. We propose a novel Monte-Carlo EM algorithm for model fitting to the multigroup data, which leverages our MCMC algorithm for sampling shapes from distributions in Kendall shape space.

We evaluate the proposed *hierarchical modeling framework for hypothesis testing*. Results on simulated data show the benefits of hierarchical modeling in (i) reducing the risk of false positives in rejecting the null hypothesis and (ii) reducing the sensitivity of the estimated per-group shape distributions to low sample sizes. For clinical datasets involving CT imaging of carpal bones, our approach leads to lower p-values even when we introduce errors in the manual segmentations to mimic human errors, compared to a popular framework for point-distribution based shape models. For *Bayesian object segmentation using statistical shape priors*, learned from a training set of manual segmentations using our hierarchical modeling and inference scheme, we propose a novel formulation that extends popular multiatlas segmentation methods and deep neural-net methods by introducing shape-distribution based regularization in Riemannian shape space. For *hierarchical shape clustering framework based on Riemannian mixture component*, we have performed our analysis on simulated ellipsoidal data. Results on simulated data show the benefits of hierarchical modeling in (i) getting high accuracy in very noisy conditions (ii) reducing sensitivity of estimated clusters and mean shape of clusters to noisy perturba-

tions. Our ongoing and future work involves exploring the applications of the proposed framework for other applications in shape analysis, e.g., shape completion and classification. One could also explore a “fully” Bayesian scheme that treats the group covariances C_m also as latent variables, introducing a hyper-prior on them at the population level.

Appendix A

Appendices

A.1 Gradients of the Log-Likelihood With Respect to Similarity Transform Parameters

The gradients of $\mathcal{J}(\mathcal{S}_{mi})$ with respect to parameters \mathbf{t} , \mathbf{s} , and \mathbf{R} are as follows.

$$\nabla_{\mathbf{s}} \mathcal{J}(\mathcal{S}_{mi}) = \sum_{s=1}^S \left(\sum_{j=1}^J 2\mathcal{D}_{x_{mi}}(\mathcal{S}_{mi}y_{mij}^s) \nabla \mathcal{D}_{x_{mi}}(\mathcal{S}_{mi}y_{mij}^s)^\top (\mathbf{R}y_{mij}^s) \right. \\ \left. + \sum_{l=1}^L 2(\mathcal{S}_{mi}y_{mij'}^s - \mathcal{Z}_{x_{mi}}^l)^\top (\mathbf{R}y_{mij}^s) \right). \quad (\text{A.1})$$

$$\nabla_{\mathbf{t}} \mathcal{J}(\mathcal{S}_{mi}) = \sum_{s=1}^S \left(\sum_{j=1}^J 2\mathcal{D}_{x_{mi}}(\mathcal{S}_{mi}y_{mij}^s) \nabla \mathcal{D}_{x_{mi}}(\mathcal{S}_{mi}y_{mij}^s)^\top \right. \\ \left. + \sum_{l=1}^L 2(\mathcal{S}_{mi}y_{mij'}^s - \mathcal{Z}_{x_{mi}}^l) \right). \quad (\text{A.2})$$

$$\nabla_{\mathbf{R}} \mathcal{J}(\mathcal{S}_{mi}) = \sum_{s=1}^S \left(\sum_{j=1}^J 2\mathcal{D}_{x_{mi}}(\mathcal{S}_{mi}y_{mij}^s) \nabla \mathcal{D}_{x_{mi}}(\mathcal{S}_{mi}y_{mij}^s) (\mathbf{s}y_{mij}^s)^\top \right. \\ \left. + \sum_{l=1}^L 2(\mathcal{S}_{mi}y_{mij'}^s - \mathcal{Z}_{x_{mi}}^l) (\mathbf{s}y_{mij}^s)^\top \right). \quad (\text{A.3})$$

A.2 Gradients of Mahalanobis distance in Riemannian Space

According to (Pennec (2006)), given μ and z are on unit hypersphere,

$$\text{Log}_\mu(z) := \cos^{-1}(\mu^\top z) \frac{(z - (\mu^\top z)\mu)}{\|z - (\mu^\top z)\mu\|_2}. \quad (\text{A.4})$$

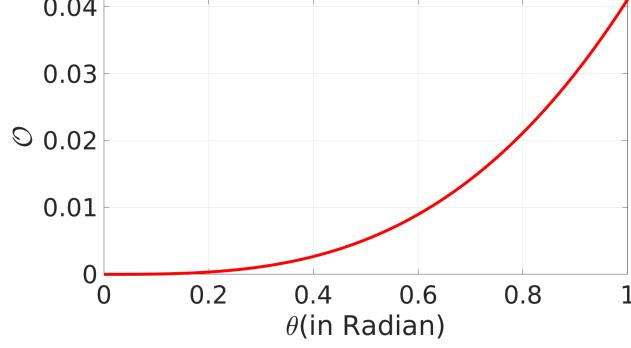


Figure A.1: Difference between the Riemannian distance and the Euclidean distance as a function of the arc length in radians.

For given covariance C squared Mahalanobis distance is

$$\begin{aligned}
 & \text{Log}_{\mu}(z)^{\top} C^{-1} \text{Log}_{\mu}(z) \\
 &:= (\cos^{-1}(\mu^{\top} z))^2 \frac{(z - (\mu^{\top} z)\mu)^{\top} C^{-1} (z - (\mu^{\top} z)\mu)}{\|(z - (\mu^{\top} z)\mu)\|_2^2} \\
 &:= (\cos^{-1}(\mu^{\top} z))^2 \frac{(z^{\top} C^{-1} z - 2(\mu^{\top} z)(\mu^{\top} C^{-1} z) + (\mu^{\top} z)^2(\mu^{\top} C^{-1} \mu))}{1 - (\mu^{\top} z)^2}
 \end{aligned} \quad . \quad (\text{A.5})$$

$$\begin{aligned}
 & \frac{\partial \text{Log}_{\mu}(z)^{\top} C^{-1} \text{Log}_{\mu}(z)}{\partial \mu} \\
 &:= \frac{-2z \cos^{-1}(\mu^{\top} z)}{\sqrt{1 - (\mu^{\top} z)^2}} \frac{(z^{\top} C^{-1} z - 2(\mu^{\top} z)(\mu^{\top} C^{-1} z) + (\mu^{\top} z)^2(\mu^{\top} C^{-1} \mu))}{1 - (\mu^{\top} z)^2} \\
 &+ (\cos^{-1}(\mu^{\top} z))^2 \frac{(-2(1 - \mu^{\top} z)(z(\mu^{\top} C^{-1}(z - \mu)) + (\mu^{\top} z)C^{-1}(z - \mu)))}{1 - (\mu^{\top} z)^2} \\
 &+ \frac{2z(\mu^{\top} z)(z^{\top} C^{-1} z - 2(\mu^{\top} z)(\mu^{\top} C^{-1} z) + (\mu^{\top} z)^2(\mu^{\top} C^{-1} \mu))}{(1 - (\mu^{\top} z)^2)^2}
 \end{aligned} \quad . \quad (\text{A.6})$$

$$\begin{aligned}
 & \frac{\partial \text{Log}_{\mu}(z)^{\top} C^{-1} \text{Log}_{\mu}(z)}{\partial z} \\
 &:= \frac{-2\mu \cos^{-1}(\mu^{\top} z)}{\sqrt{1 - (\mu^{\top} z)^2}} \frac{(z^{\top} C^{-1} z - 2(\mu^{\top} z)(\mu^{\top} C^{-1} z) + (\mu^{\top} z)^2(\mu^{\top} C^{-1} \mu))}{1 - (\mu^{\top} z)^2} \\
 &+ (\cos^{-1}(\mu^{\top} z))^2 \frac{2C^{-1} z - 2\mu(\mu^{\top} C^{-1} z) - 2(\mu^{\top} z)(C^{-1} \mu) + 2(\mu^{\top} z)\mu(\mu^{\top} C^{-1} \mu)}{1 - (\mu^{\top} z)^2} \\
 &+ \frac{2\mu(\mu^{\top} z)(z^{\top} C^{-1} z - 2(\mu^{\top} z)(\mu^{\top} C^{-1} z) + (\mu^{\top} z)^2(\mu^{\top} C^{-1} \mu))}{(1 - (\mu^{\top} z)^2)^2}
 \end{aligned} \quad . \quad (\text{A.7})$$

A.3 Approximating the Normalizing Factor in Riemannian Space

According to (Pennec (2006)), given μ and z are on unit hypersphere,

$$\text{Log}_\mu(z) := \cos^{-1}(\mu^\top z) \frac{(z - (\mu^\top z)\mu)}{\|z - (\mu^\top z)\mu\|_2}. \quad (\text{A.8})$$

The distance between μ and z on the unit hypersphere is $\|\text{Log}_\mu(z)\|_2 = \cos^{-1}(\mu^\top z)$. If μ and z are sufficiently close, then we can approximate $\text{Log}_\mu(z)$ by $z - \mu$. We now analyze the quality of this approximation. Let the residual be $O := \|\text{Log}_\mu(z)\|_2 - \|z - \mu\|_2 = \cos^{-1}(\mu^\top z) - \sqrt{2 - 2\mu^\top z}$. Let $\theta = \cos^{-1}(\mu^\top z)$ be the angle between the two unit vector μ and z . Then $O = \theta - \sqrt{2 - 2\cos(\theta)}$. Figure A.1 shows the graph of O versus θ , for θ between 0 to 1 radians. Thus, for shape pointsets that are within ± 0.6 radians from the mean shape, the residual between the two lengths is less than 1% the radius of the hypersphere.

References

- Aljabar, P., Heckemann, R., Hammers, A., Hajnal, J., & Rueckert, D. (2009). Multi-atlas based segmentation of brain images: atlas selection and its effect on accuracy. *NeuroImage*, 46, 726–38.
- Alloussonniere, S., Amit, Y., & Trouve, A. (2007). Toward a coherent statistical framework for dense deformable template estimation. *J. Royal Statistical Society, Series B*, 69, 3–29.
- Alloussonniere, S., Kuhn, E., & Trouve, A. (2010). Construction of Bayesian deformable models via a stochastic approximation algorithm: A convergence study. *Bernoulli*, 16, 641–78.
- Arif, M., Knapp, K., & Slabaugh, G. (2017). SPNet: Shape prediction using a fully convolutional neural network. In *Med. Imag. Comput. Comp.-Assist. Interv.* (pp. 203–11).
- Asman, A., & Landman, B. (2013). Non-local statistical label fusion for multi-atlas segmentation. *Med. Imag. Anal.*, 17, 194–208.
- Ataloglou, D., Dimou, A., Zarpalas, D., & Daras, P. (2019). Fast and precise hippocampus segmentation through deep convolutional neural network ensembles and transfer learning. *Neuroinformatics*, .
- Awate, S. P., & Whitaker, R. T. (2014). Multitlas segmentation as nonparametric regression. *IEEE Transactions on Medical Imaging*, 33, 1803–17.
- Badrinarayanan, V., Kendall, A., & Cipolla, R. (2017). Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39, 2481–2495.
- Bai, W., Shi, W., O'Regan, D. P., Tong, T., Wang, H., Jamil-Copley, S., Peters, N. S., & Rueckert, D. (2013). A probabilistic patch-based label fusion model for multi-atlas

- segmentation with registration refinement: Application to cardiac MR images. *IEEE Transactions on Medical Imaging*, 32, 1302–1315.
- Bone, A., Colliot, O., & Durrleman, S. (2018). Learning distributions of shape trajectories from longitudinal datasets: a hierarchical model on a manifold of diffeomorphisms. In *Computer Vision and Pattern Recognition*.
- Cates, J., Fletcher, P. T., & Whitaker, R. (2008). A hypothesis testing framework for high-dimensional shape model. In *Mathematical Foundations of Computational Anatomy Workshop* (pp. 170–81).
- Cates, J., Fletcher, T., Styner, M., Shenton, M., & Whitaker, R. (2007). Shape modeling and analysis with entropy-based particle systems. In *Proc. Info. Process. Med. Imag.* (pp. 333–45).
- Chen, F., Yu, H., Hu, R., & Zeng, X. (2013). Deep learning shape priors for object segmentation. In *IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1870–1877).
- Chen, S., & Radke, R. J. (2009). Level set segmentation with both shape and intensity priors. In *International Conference on Computer Vision* (pp. 763–770).
- Cootes, T., Taylor, C., Cooper, D., & Graham, J. (1995). Active shape models - their training and application. *Computer Vision and Image Understanding*, 61, 38–59.
- Cuingnet, R., Prevost, R., Lesage, D., Cohen, L. D., Mory, B., & Ardon, R. (2012). Automatic detection and segmentation of kidneys in 3d ct images using random forests. In *Med. Imag. Comput. Comp.-Assist. Interv.* (pp. 66–74).
- Dambreville, S., Rathi, Y., & Tannenbaum, A. (2008). A framework for image segmentation using shape models kernel space shape priors. *IEEE Trans. Pattern Anal. Mach. Intell.*, 30.
- Durrleman, S., Pennec, X., Trouve, A., & Ayache, N. (2009). Statistical models of sets of curves and surfaces based on currents. *Med. Imag. Anal.*, 13, 793–808.
- Durrleman, S., Prastawa, M., Charon, N., Korenberg, J. R., Joshi, S., Gerig, G., & Trouv , A. (2014). Morphometry of anatomical shape complexes with dense deformations and sparse parameters. *NeuroImage*, 101, 35–49.
- Fletcher, T., Lu, C., Pizer, S., & Joshi, S. (2004). Principal geodesic analysis for the study of nonlinear statistics of shape. *IEEE Trans. Med. Imag.*, 23, 995–1005.

- Freifeld, O., & Black, M. J. (2012). Lie bodies: A manifold representation of 3D human shape. In *European Conf. on Computer Vision* (pp. 1–14). Springer-Verlag.
- Gaikwad, A., Shigwan, S., & Awate, S. P. (2015). A statistical model for smooth shapes in Kendall shape space. In *Proc. Med. Image Comput. Comp. Assist. Interv.* (pp. 628–35). volume 3.
- Gelman, A. (2006). Multilevel (hierarchical) modeling: What it can and cannot do. *Am. Stat. Assoc.*, 48, 432–5.
- Gerig, G., Styner, M., Shenton, M., & Lieberman, J. (2001). Shape versus size: Improved understanding of the morphology of brain structures. In *Medical Image Computing and Computer-Assisted Intervention* (pp. 24–32). Springer Berlin Heidelberg.
- Glasbey, C. A., & Mardia, K. V. (2001). A penalized likelihood approach to image warping. *J Royal Statistical Society, Series B*, 63, 465–514.
- Goodall, C. (1991). Procrustes methods in statistical analysis of shape. *Journal of the Royal Statistical Society*, 53, 285–339.
- Gooya, A., Lekadir, K., Castro-Mateos, I., Pozo, J. M., & Frangi, A. F. (2018). Mixture of probabilistic principal component analyzers for shapes from point sets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40, 891–904.
- Greenspan, H., Ruf, A., & Goldberger, J. (2006). Constrained gaussian mixture model framework for automatic segmentation of mr brain images. *IEEE Transactions on Medical Imaging*, 25, 1233–45.
- Hagwood, C., Bernal, J., Halter, M., Elliott, J., & Brennan, T. (2013). Testing equality of cell populations based on shape and geodesic distance. *IEEE Transactions on Medical Imaging*, 32, 2230–2237.
- Heimann, T., & Meinzer, H.-P. (2009). Statistical shape models for 3d medical image segmentation: A review. *Medical Image Analysis*, 13, 543–563.
- Hong, J., Vicory, J., Schulz, J., Styner, M., Marron, J., & Pizer, S. M. (2016). Non-euclidean classification of medically imaged objects via s-reps. *Medical Image Analysis*, 31, 37–45.
- Joshi, A. A., Leahy, R. M., Badawi, R. D., & Chaudhari, A. J. (2016). Registration-based morphometry for shape analysis of the bones of the human wrist. *IEEE Transactions on Medical Imaging*, 35, 416–426.

- Kass, M., Witkin, A., & Terzopoulos, D. (1988). Snakes: Active contour models. *International Journal of Computer Vision*, 1, 321–331.
- Kendall, D. (1989). A survey of the statistical theory of shape. *Statist. Sci.*, 4, 87–99.
- Leventon, M. E., Grimson, W. E. L., & Faugeras, O. (2000). Statistical shape influence in geodesic active contours. In *IEEE Conf. on Com. Vis. Pat. Rec.* (pp. 316–323). volume 1.
- Li, Y., Ho, C. P., Toulemonde, M., Chahal, N., Senior, R., & Tang, M. X. (2018). Fully automatic myocardial segmentation of contrast echocardiography sequence using random forests guided by shape model. *IEEE Transactions on Medical Imaging*, 37, 1081–1091.
- MacKay, D. (2012). *Information theory, inference, learning algorithms*. Cambridge Univ. Pr.
- Mesadi, F., Erdil, E., Cetin, M., & Tasdizen, T. (2018). Image segmentation using disjunctive normal bayesian shape and appearance models. *IEEE Transactions on Medical Imaging*, 37, 293–305.
- Moore, D., Crisco, J., Trafton, T., & Leventhal, E. (2007). A digital database of wrist bone anatomy and carpal kinematics. *J. Biomech.*, 40, 2537–42.
- Nguyen, H., & McLachlan, G. (2018). On approximations via convolution-defined mixture models. *Communications in Statistics-Theory and Methods*, (pp. 1–14).
- Nogovitsyn, N., & et al., R. S. (2019). Testing a deep convolutional neural network for automated hippocampus segmentation in a longitudinal sample of healthy participants. *NeuroImage*, 197, 589–97.
- Oktay, O., Ferrante, E., Kamnitsas, K., Heinrich, M., Bai, W., Caballero, J., Cook, S. A., de Marvao, A., Dawes, T., OâĂŹRegan, D. P., Kainz, B., Glocker, B., & Rueckert, D. (2018). Anatomically constrained neural networks (acnns): Application to cardiac image enhancement and segmentation. *IEEE Transactions on Medical Imaging*, 37, 384–395.
- Pennec, X. (2006). Intrinsic statistics on Riemannian manifolds: Basic tools for geometric measurements. *J. Mathematical Imaging and Vision*, 25, 127–154.

- Permuter, H., Francos, J., & Jermyn, I. (2006). A study of gaussian mixture models of color and texture features for image classification and segmentation. *Pattern Recognition*, 39, 695–706.
- Pizer, S., Jung, S., Goswami, D., Vicory, J., Zhao, X., Chaudhuri, R., Damon, J., Huckemann, S., & Marron, J. (2012). Nested sphere statistics of skeletal models. *Innov. Shape Anal.*, (pp. 93–115).
- Pizer, S. M., Jung, S., Goswami, D., Vicory, J., Zhao, X., Chaudhuri, R., Damon, J. N., Huckemann, S., & Marron, J. S. (2013). Nested sphere statistics of skeletal models. In *Innovations for Shape Analysis: Models and Algorithms* (pp. 93–115). Springer.
- Pouch, A., Tian, S., Takebe, M., Yuan, J., Gorman, R., Cheung, A., Wang, H., Jackson, B., Gorman, J., Gorman, R., & Yushkevich, P. (2015). Medially constrained deformable modeling for segmentation of branching medial structures: Application to aortic valve segmentation and morphometry. *Med. Imag. Anal.*, 26, 217–31.
- Pouch, A., Wang, H., Takabe, M., Jackson, B., Gorman, J., Gorman, R., Yushkevich, P., & Sehgal, C. (2014). Fully automatic segmentation of the mitral leaflets in 3D transesophageal echocardiographic images using multi-atlas joint label fusion and deformable medial modeling. *Med. Imag. Anal.*, 18, 118–29.
- Rasmussen, C. (2000). The infinite gaussian mixture model. In *Advances in Neural Information Processing Systems* (pp. 554–60).
- Ravishankar, H., Venkataramani, R., Thiruvenkadam, S., & Vaidya, S. (2017). Learning and incorporating shape models for semantic segmentation. In *Med. Imag. Comput. Comp.-Assist. Interv.* (pp. 203–11).
- Reynolds, D. (2015). Gaussian mixture models. *Encyclopedia of Biometrics*, (pp. 827–832).
- Reynolds, D., Quatieri, T., & Dunn, R. (2000). Speaker verification using adapted gaussian mixture models. *Digital Signal Processing*, 10, 19–41.
- Reynolds, D., & Rose, R. (1995). Robust text-independent speaker identification using gaussian mixture speaker models. *IEEE Transactions on Speech and Audio Processing*, 3, 72–83.
- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *Med. Imag. Comput. Comp.-Assist. Interv.*.

- Schulz, J., Pizer, S. M., Marron, J. S., & Godtliebsen, F. (2016). Non-linear hypothesis testing of geometric object properties of shapes applied to hippocampi. *J. Math. Imaging Vis.*, 54, 15–34.
- SCI Institute (2013). ShapeWorks: An open-source tool for constructing compact statistical point-based models of ensembles of similar shapes that does not rely on specific surface parameterization. <http://www.sci.utah.edu/software/shapeworks.html>.
- Shigwan, S. J., & Awate, S. P. (2016). Hierarchical generative modeling and monte-carlo em in riemannian shape space for hypothesis testing. *Med. Image Comput. Computer Assist. Interv. (MICCAI)*, .
- Shin, H., Roth, H., Gao, M., Lu, L., Xu, Z., Nogues, I., Yao, J., Mollura, D., & Summers, R. (2016). Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE Trans. Med. Imag.*, 35.
- Siddiqi, K., & Pizer, S. (2008). *Medial Representations: Mathematics, Algorithms and Applications*. Springer.
- Taylor, C. J., & Kriegman, D. J. (1994). Minimization on the lie group $\text{so}(3)$ and related manifolds. *Technical Report Yale University*, 9405.
- Terriberry, T., Joshi, S., & Gerig, G. (2005). Hypothesis testing with nonlinear shape models. In *Proc. Info. Process. Med. Imag.* (pp. 15–26).
- Tsai, A., Yezzi, A., Wells, W., Tempany, C., Tucker, D., Fan, A., Grimson, W. E., & Willsky, A. (2003). A shape-based approach to the segmentation of medical imagery using level sets. *IEEE Transactions on Medical Imaging*, 22, 137–154.
- Tu, L., Styner, M., Vicory, J., Elhabian, S., Wang, R., Hong, J., Paniagua, B., Prieto, J. C., Yang, D., Whitaker, R., & Pizer, S. M. (2018). Skeletal shape correspondence through entropy. *IEEE Transactions on Medical Imaging*, 37, 1–11.
- Vaillant, M., & Glaunes, J. (2005). Surface matching via currents. In *Proc. IPMI* (pp. 381–92).
- Veni, G., Elhabian, S. Y., & Whitaker, R. T. (2017). Shapecut: Bayesian surface estimation using shape-driven graph. *Medical Image Analysis*, 40, 11–29.

- Wang, H., Suh, J. W., Das, S. R., Pluta, J. B., Craige, C., & Yushkevich, P. A. (2013a). Multi-atlas segmentation with joint label fusion. *IEEE Trans. Pattern Anal. Mach. Intell.*, 35, 611–623.
- Wang, Z., Mohamed, S., & Freitas, N. (2013b). Adaptive Hamiltonian and Riemann manifold Monte Carlo samplers. In *Proc. Int. Conf. Mach. Learn.* (pp. 1462–70). volume 3.
- Yu, Y., Fletcher, P., & Awate, S. (2014). Hierarchical Bayesian modeling, estimation, and sampling for multigroup shape analysis. In *Med. Image Comput. Comp. Assist. Interv.* (pp. 9–16).
- Zhang, M., Wells, W. M., & Golland, P. (2017). Probabilistic modeling of anatomical variability using a low dimensional parameterization of diffeomorphisms. *Medical Image Analysis*, 41, 55–62.
- Zhang, S., Zhan, Y., Dewan, M., Huang, J., Metaxas, D. N., & Zhou, X. S. (2012). Towards robust and effective shape modeling: Sparse shape composition. *Medical Image Analysis*, 16, 265 – 277.
- Zhu, H., Shi, F., Wang, L., Hung, S., Chen, M., Wang, S., Lin, W., & Shen, D. (2019). Dilated dense U-net for infant hippocampus subfield segmentation. *Frontiers in Neuroinformatics*, 13, 30.
- Zivkovic, Z. (2004). Improved adaptive gaussian mixture model for background subtraction. In *Proceedings of the 17th International Conference on Pattern Recognition* (pp. 28–31). volume 2.

List of Publications

- Shigwan, S.J.,and Awate, S.P., 2016, October. Hierarchical Generative Modeling and Monte-Carlo EM in Riemannian Shape Space for Hypothesis Testing. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 191–200). Springer, Cham.
- Gaikwad, A.V. and Shigwan, S.J. and Awate, S.P., 2015, October. A Statistical Model for Smooth Shapes in Kendall Shape Space. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 628–635). Springer, Cham.
- Shigwan, S.J.,and Awate, S.P., Object Segmentation With Deep Neural Nets Coupled with a Shape Prior, When Learning from a Training Set of Limited Quality and Small Size. In IEEE 17th International Symposium on Biomedical Imaging (ISBI) 2020 (pp. 1149-1153)
- Shigwan, S.J.,and Awate, S.P., A Hierarchical Multigroup Riemannian Model on Smooth Shapes using Pointset-based Representations: Applications in Hypothesis Testing and Object Segmentation. (journal to-be-submitted)