SCIENTIFIC COMPUTING AND IMAGING INSTITUTE

THE UNIVERSITY OF UTAH

ShapeWorks Documentation

ShapeWorks : Particle-based Shape Correspondence Software

Shireen Elhabian, Praful Agrawal, Riddhish Bhalodia, Joshua Cates, Manasi Datar, Ross Whitaker

June 28, 2018

Abstract

This document outlines the basic theory behind shapeworks alogorithm.

Contents

1 License

2 Acknowledgement

If you use ShapeWorks and/or ShapeWorksStudio in work that leads to published research, we humbly ask that you add the following to the 'Acknowledgments' section of your paper: *"This project was supported by the National Institute of General Medical Sciences of the National Institutes of Health under grant number P41 GM103545-18."*

3 Introduction

This software is an open source distribution of the algorithm for constructing correspondence-based statistical models of sets of similar shapes, originally created and maintained by the research group of Dr. Whitaker at the University of Utah [1, 2, 3, 4, 5]. The scientific and clinical effectiveness of ShapeWorks and its underlying scientific premise of particle-based shape modeling have been demonstrated in a range of applications, including neuroscience [3, 6, 7, 8, 9, 10], biological phenotyping [11, 12], orthopedics

[12, 13, 14], and cardiology [15, 16]. This document is intended to provide necessary background information for understanding and using the ShapeWorksStudio software, including a basic tutorial and reference. This section gives a technical overview of correspondence shape models and the correspondence optimization algorithm that is implemented in the ShapeWorks software. For a more detailed explanation of the ShapeWorks optimization algorithm, please refer to [2, 5].

3.1 Particle-based correspondence models

Manually defined landmarks have been the most popular choice for a lightweight, general shape representation that is suitable for statistical analysis and visual communication of the results [17, 18]. ShapeWorks provides an important paradigm shift from manually defined anatomical homologies to computationally derived *correspondence shape models* that are constructed automatically from 3D images of anatomy (i.e., computed tomography (CT) and magnetic resonance imaging (MRI)). Without relying on any surface parameterization, ShapeWorks enables a generalized and flexible approach to handling general anatomy with arbitrary topology compared to existing tools for surface-based analysis that are either tailored to specific anatomical structures or restrict modeling to spherical topologies. In particular, ShapeWorks computes a statistically optimal representation of the population variability by choosing landmark positions that minimize the overall information content of the model while maintaining a good sampling of surface geometry.

Correspondence point models represent shape by sampling each shape surface in a consistently ordered fashion so as to define homologous surface points across the population of shapes. These homologous surface points are called *correspondences*. Once chosen, the 3D positions of all m correspondences on a shape sample can be encoded as a $3m$ shape vector, which is the point-based representation of that shape. The ordering of the particles on each shape implies a correspondence among shapes, and the distribution of all of the shape samples in this $3m$-dimensional vector space (*shape space*) gives rise to the statistical analysis. The problem of how to choose the correspondence positions, and thus the shape-space distribution of the sample population, is a model selection problem. Model selection problems in statistics are typically resolved by choosing the simplest model that explains the observed data (see, e.g., [19]). Consistent with this idea, the ShapeWorks correspondence optimization algorithm seeks a compact distribution for its correspondences in the shape space (for model simplicity), while simultaneously seeking good surface samplings in order to accurately represent the data. Fig 1 illustrates these basic concepts for a population of hand shapes. The figure shows the mapping of all the surface point samples from a single shape to a single point in the higher- dimensional shape space, and the relative compactness of the distribution of all samples in that shape space after optimization.

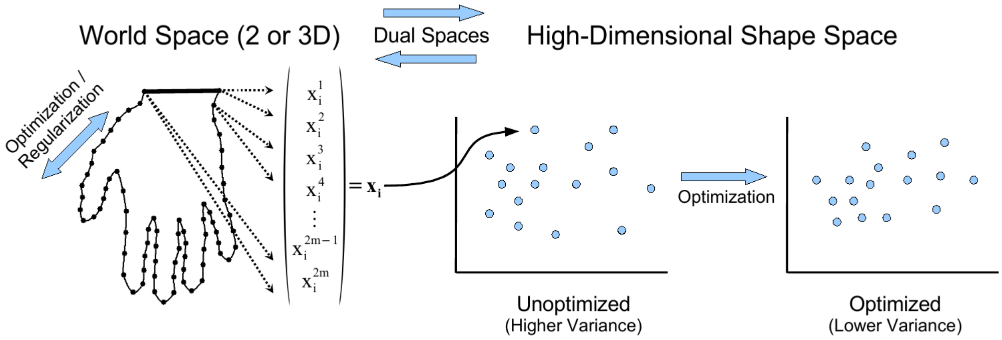

Figure 1: An illustration of the basic concepts the ShapeWorks point-based correspondence optimization.

3.2 Correspondence optimization

The correspondence optimization currently implemented in ShapeWorks works by modeling the correspondence positions as sets of dynamic particles that are constrained to lie on the set of sample shape surfaces. The positions of the correspondences are optimized by gradient descent on an energy function that balances the negative entropy of the distribution of particles on each shape surface with the positive entropy of the distribution of the shape samples shape space. More specifically, and with reference to Fig 1, we consider $\mathbf{x}_k \in \mathbb{R}^{3m}$ as an instance of a random variable $\mathbf{Z}$ and minimize the energy function

$$Q = \alpha H(\mathbf{Z}) - \sum_k H(\mathbf{x}_k) \tag{1}$$

where H is an estimation of differential entropy. Minimizing the first term in Q produces a compact distribution of samples in shape space, while the second term seeks uniform surface samplings for accurate shape representation. This latter term can also be modified to adaptively over-sample in response to local surface features such as curvature. The free parameter α balances the tradeoff between model compactness and accurate shape representation. Since correspondence points in this formulation are not tied to a specific parameterization, the method operates directly on volumetric data and extends easily to arbitrary shapes, even non-manifold surfaces.

3.3 Shape modeling pipeline

This software distribution consists of several applications that together make up a full processing pipeline for correspondence computation and visualization. The stages of the pipeline are as follows, (1) preprocessing, (2) alignment, (3) initialization, (4) optimization, and (5) visualization & analysis. This section gives an overview of the steps involved in the pipeline.

Preprocessing: Any set of implicitly defined surfaces, such as a set of binary segmentations, is appropriate as input to this pipeline. By default, the software will assume a closed surface, but can also handle open boundaries (specified by cutting planes). A binary mask, such as the output of an image segmentation process, for example, contains an implicit shape surface at the interface of the labeled pixels and the background. Binary masks contain aliasing artifacts, however, that should first be removed. Typically we follow an antialiasing step with a very slight Gaussian blurring to remove the high-frequency artifacts that can occur as a result of numerical approximations.

Alignment: A collection of shape segmentations must often be aligned in a common coordinate frame for modeling and analysis. The ShapeWorks software does not provide full support for this stage of the pipeline, since every new dataset may require a different process of alignment. Some basic alignment tools that are included with ShapeWorks are the ability to align the segmentations with respect to their centers of mass and the orientation of their first principal eigenvectors. For some classes of data, this method is effective as a rough alignment and is followed, during the optimization process, by iterative Procrustes alignments based on the correspondence point positions [20]. ShapeWorks distinguishes between two separate coordinate frames: a local coordinate frame for each shape sample, and a common coordinate frame in which all shape samples are co-registered. For applications in which the input data is completely registered, these two coordinate frames are the same. Where there is a distinction, ShapeWorks will maintain and output appropriate transformations to the common coordinate frame for each shape sample, as well as separate correspondence files for each coordinate frame.

Correspondence initialization: Typically, we initialize the correspondence optimization with a single particle that finds the nearest zero of the implicit surface, and then splits under optimization (producing a new, nearby particle) at regular intervals until a specific number of particles are produced and reach a steady state. This initialization procedure is supported within ShapeWorks and its Studio software. We have found this procedure to be generally applicable to classes of closed shapes with reasonably smooth surfaces.

Correspondence optimization: The optimization stage proceeds as described earlier starting with an initial set of particle positions and a set of implicit surfaces (e.g., signed distance transforms). Several types of optimization procedures are supported: adaptive gradient descent, Jacobi updates, and Gauss-Seidel updates. By default, gradient descent with Gauss-Seidel updates is used. Important parameters to consider for the optimization are the regularization on the covariance matrix for the shape-space entropy estimation, and the parameter α from Equation 1 which controls the trade-off between the uniformity of the surface sampling and the compactness of the optimized model. Also consider whether iterative Procrustes alignment is necessary and how often to it can be performed.

Visualization and analysis: Statistical analysis of point-based shape models is difficult because point-wise statistical tests require multiple-comparisons corrections that significantly reduce statistical power [21]. Analysis in the full shape space, however, is problematic due to the high number of dimensions and the difficulty of obtaining sufficient samples. A common solution is to employ dimensionality reduction reduction by choosing a subspace in which to project the data for traditional multivariate analysis (see, e.g., [3, 21]). This software does not include specific code for statistical analysis of the correspondences. It does, however, provide principal component analysis (PCA), which can be used for dimensionality reduction prior to subsequent statistical analysis, and visualization of orthogonal modes of variation in the model (see, e.g,. [3, 21]). ShapeWorks also includes methods for computing and visualizing the Euclidean mean shapes and the median shapes of populations. Typically, we use a separate statistical package, such as R (the R Foundation, www.r-project.org), for analysis of the PCA loadings of the correspondence positions. The user-friendly Studio package provides tools to export different optimization outputs for subsequent statistical analysis. For more ideas regarding visualization and statistical analysis of correspondence models, see the various publications listed in the bibliography section of this document.

References

[1] J. Cates, M. Meyer, T. Fletcher, and R. Whitaker, "Entropy-based particle systems for shape correspondence," in "1st MICCAI Workshop on Mathematical Foundations of Computational Anatomy: Geometrical, Statistical and Registration Methods for Modeling Biological Shape Variability," (2006), pp. 90–99.

[2] J. Cates, P. T. Fletcher, M. Styner, M. Shenton, and R. Whitaker, "Shape modeling and analysis with entropy-based particle systems,"

in “Information Processing in Medical Imaging,” (Springer, 2007), pp. 333–345.

[3] J. Cates, P. T. Fletcher, M. Styner, H. C. Hazlett, and R. Whitaker, “Particle-based shape analysis of multi-object complexes,” in “Medical Image Computing and Computer-Assisted Intervention–MICCAI 2008,” (Springer, 2008), pp. 477–485.

[4] I. Oguz, J. Cates, M. Datar, B. Paniagua, T. Fletcher, C. Vachet, M. Styner, and R. Whitaker, “Entropy-based particle correspondence for shape populations,” International Journal of Computer Assisted Radiology and Surgery pp. 1–12 (2015).

[5] J. Cates, S. Elhabian, and R. Whitaker, “Shapeworks: Particle-based shape correspondence and visualization software,” in “Statistical Shape and Deformation Analysis: Methods, Implementation and Applications, 1st Edition,” , G. Zheng, S. Li, and G. Szekely, eds. (Academic Press, 2017), chap. 10.

[6] I. Oguz, J. Cates, T. Fletcher, R. Whitaker, D. Cool, S. Aylward, and M. Styner, “Cortical correspondence using entropy-based particle systems and local features,” in “Biomedical Imaging: From Nano to Macro, 2008. ISBI 2008. 5th IEEE International Symposium on,” (IEEE, 2008), pp. 1637–1640.

[7] I. Oguz, M. Niethammer, J. Cates, R. Whitaker, T. Fletcher, C. Vachet, and M. Styner, “Cortical correspondence with probabilistic fiber connectivity,” in “Information Processing in Medical Imaging,” (Springer, 2009), pp. 651–663.

[8] M. Datar, J. Cates, P. T. Fletcher, S. Gouttard, G. Gerig, and R. Whitaker, “Particle-based shape regression of open surfaces with applications to developmental neuroimaging,” in “Medical Image Computing and Computer-Assisted Intervention–MICCAI 2009,” (Springer, 2009), pp. 167–174.

[9] M. Datar, I. Lyu, S. Kim, J. Cates, M. A. Styner, and R. Whitaker, “Geodesic distances to landmarks for dense correspondence on ensembles of complex shapes,” in “Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013,” (Springer, 2013), pp. 19–26.

[10] S. Sultana, P. Agrawal, S. Elhabian, R. Whitaker, T. Rashid, J. Blatt, J. Cetas, and M. Audette, “Towards a statistical shape-aware deformable contour model for cranial nerve identification,” in “Workshop on Clinical Image-Based Procedures,” (Springer, 2016), pp. 68–76.

[11] J. Cates, P. T. Fletcher, Z. Warnock, and R. Whitaker, “A shape analysis framework for small animal phenotyping with application to mice with a targeted disruption of hoxd11,” in “Biomedical Imaging: From

Nano to Macro, 2008. ISBI 2008. 5th IEEE International Symposium on," (IEEE, 2008), pp. 512–515.

[12] K. B. Jones, M. Datar, S. Ravichandran, H. Jin, E. Jurrus, R. Whitaker, and M. R. Capecchi, "Toward an understanding of the short bone phenotype associated with multiple osteochondromas," Journal of Orthopaedic Research **31**, 651–657 (2013).

[13] M. D. Harris, M. Datar, R. T. Whitaker, E. R. Jurrus, C. L. Peters, and A. E. Anderson, "Statistical shape modeling of cam femoroacetabular impingement," Journal of Orthopaedic Research **31**, 1620–1626 (2013).

[14] P. R. Atkins, S. Elhabian, P. Agrawal, M. D. Harris, R. T. Whitaker, J. A. Weiss, C. L. Peters, and A. E. Anderson, "Quantitative comparison of cortical bone thickness using correspondence-based shape modeling in patients with cam femoroacetabular impingement," Journal of Orthopaedic Research (2016).

[15] J. Cates, E. Bieging, A. Morris, G. Gardner, N. Akoum, E. Kholmovski, N. Marrouche, C. McGann, and R. S. MacLeod, "Computational shape models characterize shape change of the left atrium in atrial fibrillation," Clinical Medicine Insights. Cardiology **8**, 99 (2015).

[16] G. Gardner, A. Morris, K. Higuchi, R. MacLeod, and J. Cates, "A point-correspondence approach to describing the distribution of image features on anatomical surfaces, with application to atrial fibrillation," in "Biomedical Imaging (ISBI), 2013 IEEE 10th International Symposium on," (IEEE, 2013), pp. 226–229.

[17] N. Sarkalkan, H. Weinans, and A. A. Zadpoor, "Statistical shape and appearance models of bones," Bone **60**, 129–140 (2014).

[18] S. Zachow, "Computational planning in facial surgery." Facial Plastic Surgery: FPS **31**, 446 (2015).

[19] M. H. Hansen and B. Yu, "Model selection and the principle of minimum description length," Journal of the American Statistical Association **96**, 746–774 (2001).

[20] C. Goodall, "Procrustes methods in the statistical analysis of shape," Journal of the Royal Statistical Society. Series B (Methodological) pp. 285–339 (1991).

[21] J. Cates, T. Fletcher, and R. Whitaker, "A hypothesis testing framework for high-dimensional shape models," in "2nd MICCAI Workshop on Mathematical Foundations of Computational Anatomy," (2008), pp. 170–181.