Population-specific human biomechanical models for medical training

Nathan Drenkow[1], Nathanael Kuo[1], Jason Harper[1], Manuel Uy[1], Dean Kleissas[1], Duane Cornish[1], Gaurav Thawait[2], Jan Fritz[2], Catherine Carneal[1], Brian Corner[3], Michael Maffeo[3]
[1]The Johns Hopkins University Applied Physics Laboratory
[2]The Johns Hopkins University School of Medicine
[3]U.S. Army Natick Soldier Research, Development and Engineering Center (Natick, MA)

1. Introduction

The inherent variability of the human anatomy precludes the use of average geometries when developing population-representative human models. For the purpose of human modeling for injury prediction or medical training, models must ideally account for internal differences that exist due to patient demographics (e.g., race, age, gender) and/or external physical anthropometry (e.g., height and weight). Along this vein, allometry is a method for identifying how an organism's internal anatomical structures scale with its overall size. Most prior work has focused on studying allometric laws for single organs in the human body. However, given the complexity of the human body, a multi-organ approach is likely to produce a better understanding of how the internal structures vary collectively.

Conventional approaches to allometry focus on producing statistical shape atlases for anatomies of interest. Such atlases are designed to capture how a particular structure's size, shape, and location information varies within or across populations. Expanding on previous work, this effort developed a computational pipeline to process a set of thoracic patient Computed Tomography (CT) images representative of a population of interest (e.g., military), and produce a multi-organ statistical shape atlas to better quantify shape differences. The atlas provides a means for identifying relative differences in organ shape, size, and location between patient demographic and anthropometric data, as well as the ability to make predictions about the internal geometries using external information. The current method and pipeline analyzes the lung, liver, kidneys, and spleen, but is extensible to other anatomies.

Several statistical shape atlas techniques and systems have been developed and tend to fall into three categories. First, single-organ atlases highlight the variation of an individual organ's shape[1] (others also combine multiple single-organ atlases[2 3]). Next, aggregate atlases generate a single model that represents an "average" organ shape[4]. Finally, segmentation-driven atlases discover inter-organ relationships[5 6 7]. However, these atlases are primarily used to inform automated segmentation processes, but unlike our approach, do not enable investigations of allometric scaling laws or prediction of geometries from anthropometric measurements. Prior human modeling work has provided limited insight into the allometric laws for multi-organ systems. This paper intends to address some of these limitations.

This paper provides an overview of the computational pipeline including a discussion of the overall architecture and key components. Additionally, a dataset of subjects representative of the United States Army population was collected and processed through the pipeline to produce multi- and single-organ statistical shape atlases. A number of statistical analyses were conducted using these atlases to investigate allometric trends within this military-representative population. This report provides a discussion of the results and recommendations for future work and utilization of the pipeline.

2. Methods

2.1. Overview

The following sections outline the key components of the statistical shape atlas pipeline and the associated tools for atlas evaluation and study. At a high-level, the approach consists of four main

components: data collection/pre-processing, the computational pipeline, data analysis/post-processing, and visualization.

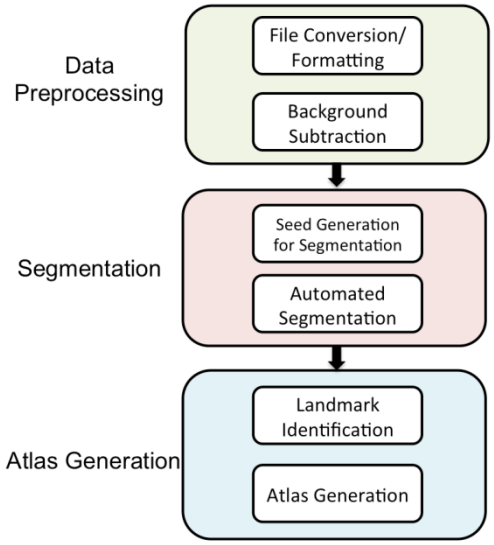

Figure 1. Computational pipeline design
Raw CT images are converted to a common data format, processed to remove artifacts, segmented automatically, and segmentations are used to generate statistical shape atlases.

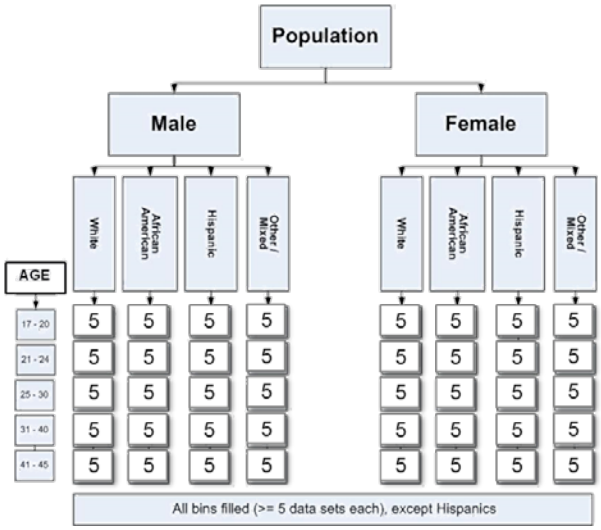

Figure 2 Target dataset demographic distributions selected to span a military-relevant population.

2.2. Data Collection

One objective for developing a pipeline for computing multi-organ statistical shape atlases is to enable population-specific analyses. Therefore, in order to study shape and size properties of different organs, data for a representative sample of the population of interest should be collected using a 3-dimensional medical imaging technique. Our process currently uses X-ray CT images, but could be extended to other modalities (e.g., MRI). We also capture external demographic and anthropometric data to enable future analysis and prediction based on external covariates instead of atlas mode weights directly.

An application was submitted to the institutional review board (IRB) and accepted to collect and analyze an existing CT image dataset at the Johns Hopkins Medical Institute. This project focused on identifying allometric trends related to organs of the thorax (i.e., lung, liver, kidneys, spleen, and bladder) and as such, a dataset consisting of thoracic CT scans of subjects representative of the military population was required. Only strictly normal scans of the thorax were collected and included in this study (normal by report and inspection). Any scan with obvious or minimal pathology was excluded. Scans that showed organs without disease but with findings different from normal were also excluded. In order to maintain consistency with the military population, a set of demographic bins (see Figure 2) were identified based on the Army Anthropometric Survey (ANSUR II)[8] and archived images were selected in an attempt to fill those bins.

For all CT images collected, height and weight were also extracted from the medical record archives. Six other measurements, more specifically – waist front length, chest span cranial-caudal (CC), chest depth, chest breadth, waist breadth, and bispinous breadth – were manually approximated from landmarks on the CT images. Visualizations of all eight of these external measurements are shown in Figure 3.

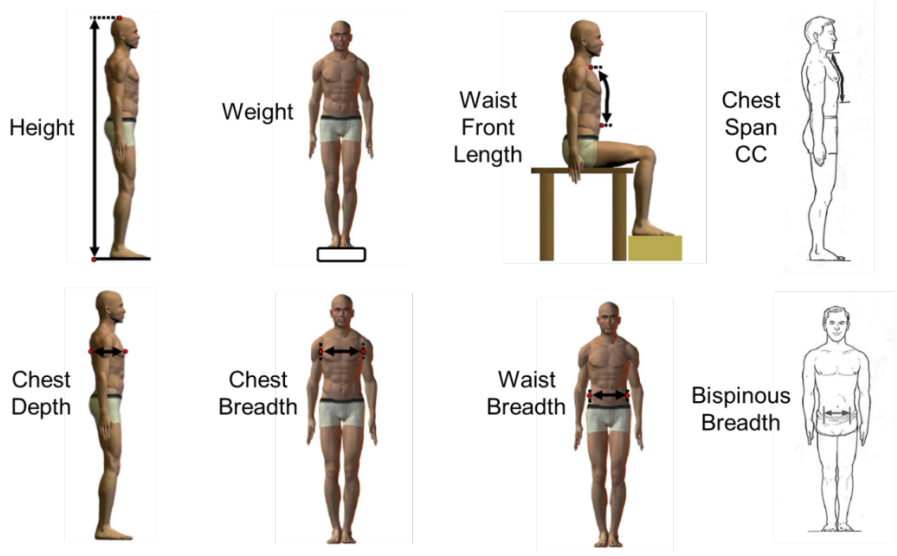

Figure 3. Visualization of anthropometric measurements. Height and weight were extracted from medical record archives while the remaining measurements were measured from the CT scans.

2.3. Computational Pipeline

2.3.1. Data preprocessing

In order to facilitate downstream computation and ensure consistency across the dataset, all CT images were exposed to several pre-processing steps. In particular, the images were re-sliced to have uniform voxel dimensions (i.e., 1mm on all sides) and padded so that volume dimensions were fixed across the dataset. Background subtraction was also performed on each image to remove all external image artifacts (e.g., scanning table, jewelry, etc.) introduced during the scan.

2.3.2. Isomap

Our method also requires a set of manually segmented scans in order to generate biological priors required for automated segmentation. Unlike previous methods that select a reference CT randomly, we take a more principled approach by selecting a set of scans that span the population's variability.

We use the Isomap method[9] in a novel iterative approach to embed the high-dimensional CT images in a lower dimensional space where each dimension captures major modes of variation. This iterative method is geared toward finding an embedding of the CT images in a lower dimensional space whereby the CT image closest to the population mean is the image whose lower-dimensional representation is closest to the origin (see Figure 4). A set of additional subjects are selected as local references by performing K-means clustering in Isomap-space and selecting the subjects closest to the K cluster centers. All references, global and local, are then manually segmented by an expert radiologist.

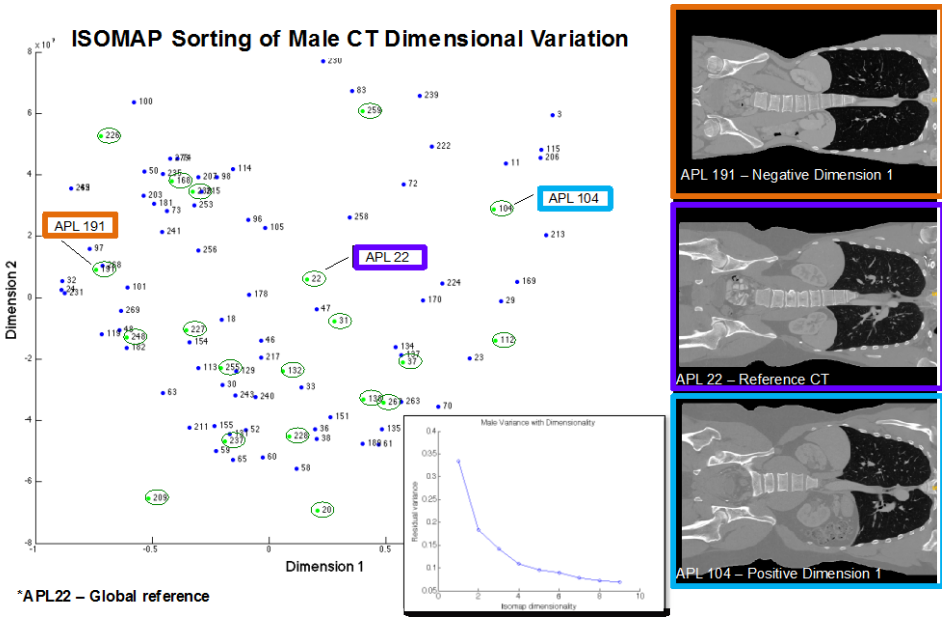

Figure 4. Isomap embedding of male CT images
Note: APL22 lies closest to the origin and was selected as the male reference. The first Isomap dimension appears to approximately capture overall subject size (small to large).

2.3.3. Automated Segmentation

In order to identify shape differences across the population, 3D shape information from each organ must be extracted from the CT images from a large number of individuals representative of that population. The cost of using a radiologist to perform the organ segmentation is prohibitively high as the number of CT images grows beyond just a few. As such, automated approaches using computer vision must be adopted to provide a scalable solution.

We have developed a novel segmentation method that combines machine learning and traditional segmentation to extract organ shapes automatically from CT volumes. We apply a two-stage approach for each organ of interest. In the first stage, we train a set of Random Forest[10] classifiers to label volumetric pixels (i.e., voxels) as organ or non-organ using the closest manually-segmented local reference in Isomap-space as ground truth for our classifier. The classifiers are then applied pixel-wise to the un-segmented test subject. In the second stage, the segmentation produced using the Random Forest classification is used as a seed for traditional Random Walker[11] segmentation. The result is the extraction of volumetric information for each organ of interest.

2.3.4. Mesh Transfer

A statistical shape atlas is generated once a set of manual or automated segmentations is available. The computational pipeline must first identify corresponding landmark points on all of the segmentation surfaces and then align those points in the same reference space. The atlas measures how these points vary across the population dataset.

On the global reference (i.e., the subject closest to the population mean), a mesh is generated for each organ where the nodes of the mesh become the set of surface points that we'll identify on all other subjects in the dataset. For each subject in the dataset, we generate a higher resolution mesh and use a series of point-cloud registration techniques (i.e., Procrustes and ICP) to align the subject mesh with the global reference mesh. When alignment is complete, for each surface point in the global reference mesh, we select the nearest surface point in the high-resolution subject mesh.

Once point correspondence is achieved, we rigidly align each subject into a common global reference

frame using Procrustes Analysis. This is akin to putting all of the subjects into a common coordinate system where they are all aligned via translation and rotation.

2.3.5. Principal Component Analysis

With each subject defined by a list of corresponding points defined in a common coordinate system, we analyze the set of subjects using Principal Component Analysis (PCA)[12].

The input data matrix to the PCA algorithm is formed by first unraveling a subject's landmark point locations into a single row vector followed by stacking row vectors for multiple subjects. The final statistical atlas is comprised of the PCA components. By modifying the weights for individual components, new geometries can be synthesized or existing structures can be reconstructed. When PCA is conducted using N subjects, the atlas will consist of a mean vector (representing the average geometry) and N-1 principal components.

2.4. Regression Analysis

We sought to explore the potential of predicting internal organ geometry from a combination of demographic and anthropometric information using our multi-organ shape atlas. Previous work[13] presented an initial feasibility study of predicting single organ geometry from nine demographic and anthropometric variables using a statistical shape atlas. Here we extend that work to predicting multiple organ geometry from eleven demographic and anthropometric variables.

As shown in Figure 5, our proposed method has three main steps. The first step is to perform PCA to more easily characterize our atlas of mesh points. Thus, each subject mesh, $\boldsymbol{x_i}$ $(i = 1, \dots, N)$, may be represented by the equation

$$\boldsymbol{x_i} = \overline{\boldsymbol{x}} + \sum_{j=1}^{M} a_{ij}\boldsymbol{e_j} \qquad (1)$$

where $\overline{\boldsymbol{x}}$ represents the mean mesh shape, M is the number of principal components, a_{ij} represents the weights, and $\boldsymbol{e_j}$ represents the principal component vectors.

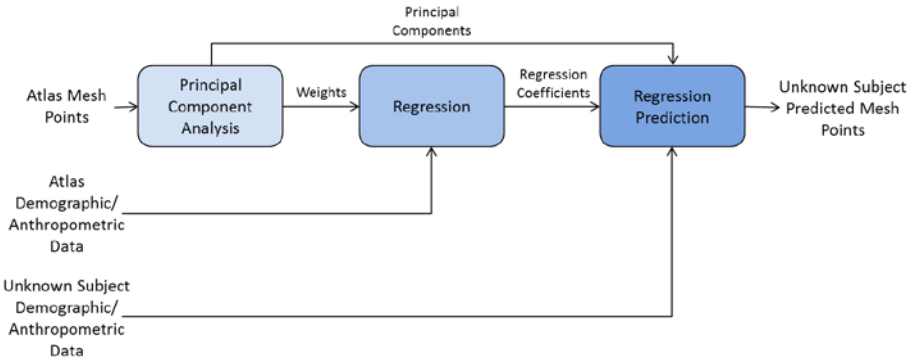

Figure 5. Overview of internal geometry prediction.

The next step in the prediction pipeline is to train a regression between the principal component weights computed from the previous step and the demographic/anthropometric data of our atlas subjects. Generally speaking, regression applies an equation model to compute a best fit between the independent variables and the dependent variable. In our case, the independent variables come from the demographic/anthropometric data and the dependent variable is a single principle component weight. This is repeated for all principle component weights, amounting to M regression equations. Using the common linear regression technique, the principal component weights may be represented as

$$\begin{cases} a_1 = b_{10} + \sum_{k=1}^{K} b_{1k} x_k \\ \vdots \\ a_M = b_{M0} + \sum_{k=1}^{K} b_{Mk} x_k \end{cases} \quad (2)$$

where a_j $(j = 1, \dots, M)$ are the dependent variables (i.e., principal component weights), x_k $(k = 1, \dots, K)$ are the independent variables (i.e., measurement data), and b_{jk} (including $k = 0$) are the coefficients for linear regression. In addition to linear regression, stepwise regressions using both constant and quadratic initial models were also explored.

Once PCA has been applied and the coefficients from training our regression model have been obtained, the final step is prediction. Given demographic/anthropometric data from an unknown subject, we may now predict the mesh points of the unknown subject by first computing the principal component weights using the regression coefficients (as in (2)). We then reconstruct the mesh points using the principal components and mean mesh shape as done in (1). Reconstruction using stepwise regression is done similarly, except regression coefficients are computed using the optimized equation model.

2.5. Atlas Viewer

Since the atlas data is very complex with hundreds of thousands of vertices per organ and many-dimensional component vectors, it is impossible to fully comprehend the data directly without some visual aid. Therefore, an atlas viewer was built to display the atlas data as well as allow interactive exploration in order to better grasp the meaning of the data. Key features of the atlas include selection of atlas data based on demographics (i.e., age, sex, ethnicity), visualization of internal organ geometry predicted in real-time from arbitrary external anthropometric measurements, computation of various internal metrics, and additional visualization tools to compare various geometries.

3. Results

3.1. Data Collection

In total, 180 CT images (96 males, 84 females) were included in the final dataset. These patient CTs met the established demographic requirements, and were confirmed to have absence of disease or abnormality by an expert radiologist. The demographic distribution is described in Figure 6**.** An even gender balance was successfully achieved, but Hispanic subjects, and subjects in the 17-20 age range were slightly under-represented due to difficulty finding appropriate medical scans.

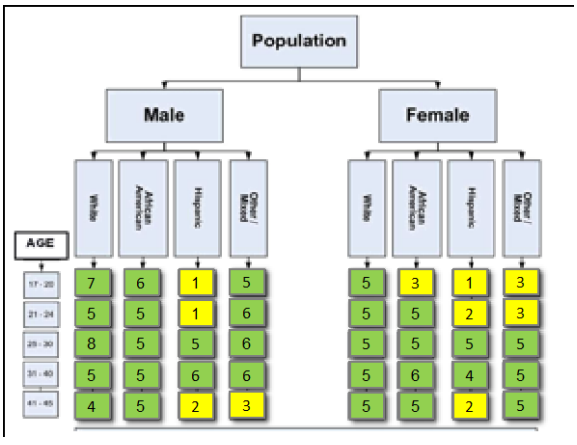

Figure 6. Subject demographic distribution included in shape atlas. Green indicates greater than or equal to five images obtained. Yellow indicates at least one image but less than five

obtained.

3.2. Computational Pipeline

The dataset was split into male and female subsets and each subset of images was passed through the computational pipeline. By leveraging the pipeline framework, individual organs were segmented in parallel and merged (per patient) at the last step. As such, the pipeline takes a set of CT images and produces a corresponding set of multi-organ segmentations.

Since the objective of the pipeline is to produce a multi-organ statistical shape atlas, only segmentations whose qualitative accuracy was sufficiently high for all organs were selected for input into the atlas generation pipeline (see Section 2.4 for more details). The results of the qualitative assessment are described in Table 1. Figure 7 provides examples of automated segmentation outputs. Overall, the computational pipeline was able to process the male and female scans within 24 hours on our cluster of 10 High Performance Computing nodes. In contrast, expert radiologists required, on average, 6 hours per *individual* subject.

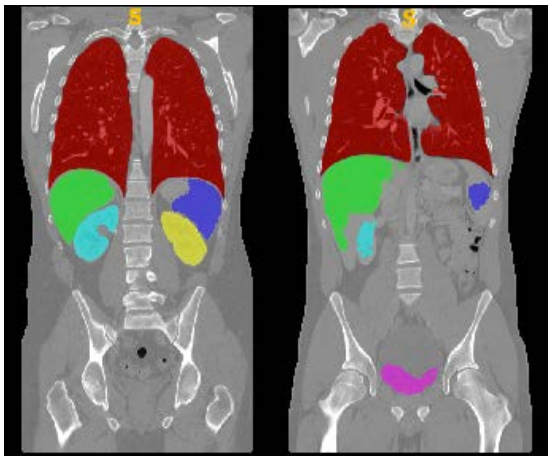

Figure 7. Automated segmentation outputs: (left) Acceptably accurate segmentation; (right) Segmentation with noticeable inaccuracies (e.g., spleen, bladder)

Table 1
Qualitative assessment of automated segmentations
Each cell represents the number of segmentations that were deemed to have few or no segmentation errors. Total number of male, female segmentations was 96, 84 respectively.

	Lungs	Liver	Spleen	Left Kidney	Right Kidney
Female	72	66	53	60	54
Male	94	86	73	85	79
Total	**166**	**152**	**126**	**145**	**133**

3.3. Regression Analysis

Results were obtained by taking the mean errors produced by leave-one-out validation. This form of validation removes a subject from the atlas when applying the training steps of PCA and regression, but uses that subject as the unknown subject for the testing step of prediction. Errors are computed on a point-by-point basis when the predicted mesh is compared to the actual mesh, and an overall mean error is computed when all subjects have been left out and tested. Our multi-organ statistical shape atlas was constructed from the 50 automated segmentations that contained accurate segmentations for all organs (lung, liver, spleen, and kidneys). The regression results presented in this section are based

on the atlas built using the Procrustes analysis for mesh alignment, and are summarized in Table 2.

Stepwise regression using a quadratic initial model also resulted in relatively large mean errors. We discovered that this is due to overfitting, i.e., too many terms were used to train the regression model, thus producing low residual errors on the training data, but high errors on new testing data. We proposed various techniques to predict internal organ geometry based on demographic/anthropometric information, and conclude that stepwise regression using a constant initial model produces the least overall mean error for our six atlases (see Table 2).

Table 2
Mean errors (mm) for leave-one-out validation

	Linear	Stepwise constant	Stepwise quadratic
Lungs	14.43	13.41	257.73
Liver	13.87	13.61	181.58
Spleen	14.11	12.34	215.69
Left kidney	11.04	11.31	235.63
Right kidney	11.37	10.45	224.02
All	20.35	19.45	314.41
Mean error	**14.20**	**13.43**	**238.18**

In order to get a better visual sense of how well stepwise constant regression performs in predicting meshes, the best and worst subjects corresponding to the minimum and maximum mean errors, respectively, are shown for each atlas in Figure 8. Note that in the best-case scenarios, errors rarely reach above 25 mm and shapes are rather smooth. In the worst-case scenarios, however, the errors may reach up to 100 mm, and shapes tend to be rougher. While the current effort focused on developing a single regression model to simultaneously predict the anatomies of the organs, it is hypothesized that developing separate independent models for all organs may yield greater accuracy in anatomical results, although the project completed before this approach could be investigated.

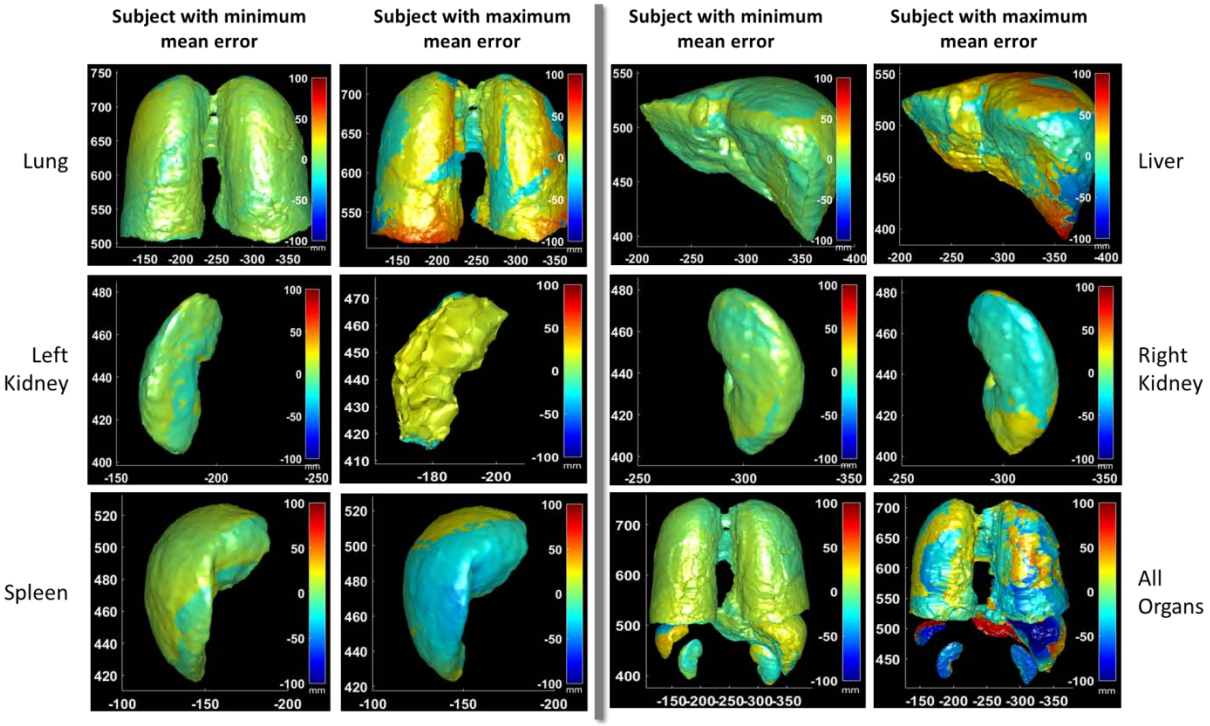

Figure 8. Visualizations of best-case (minimum) and worst-case (maximum) predicted geometries colored by error (mm). Positive error reflects a difference from ground truth in the surface outward-facing normal direction, while negative error reflects a difference from ground truth in the surface internal-facing normal direction.

3.4. Atlas Viewer

A screenshot of the developed organ atlas visualization tool is shown below in Figure 9. The Euclidean distance is computed between points on the surface of the comparison mesh (white, transparent) to the equivalent point on the surface of the baseline mesh (colored). The distance is measured in millimeters and a color is applied so that green indicates 0mm distance and red indicates >20mm. The viewer can toggle the user interface to display a predictive geometry based on specific subject demographic and anthropomorphic metrics. The baseline mesh can be set to any other subject demographic or the overall average of a population to easily comprehend the differences in internal anatomy between the two conditions.

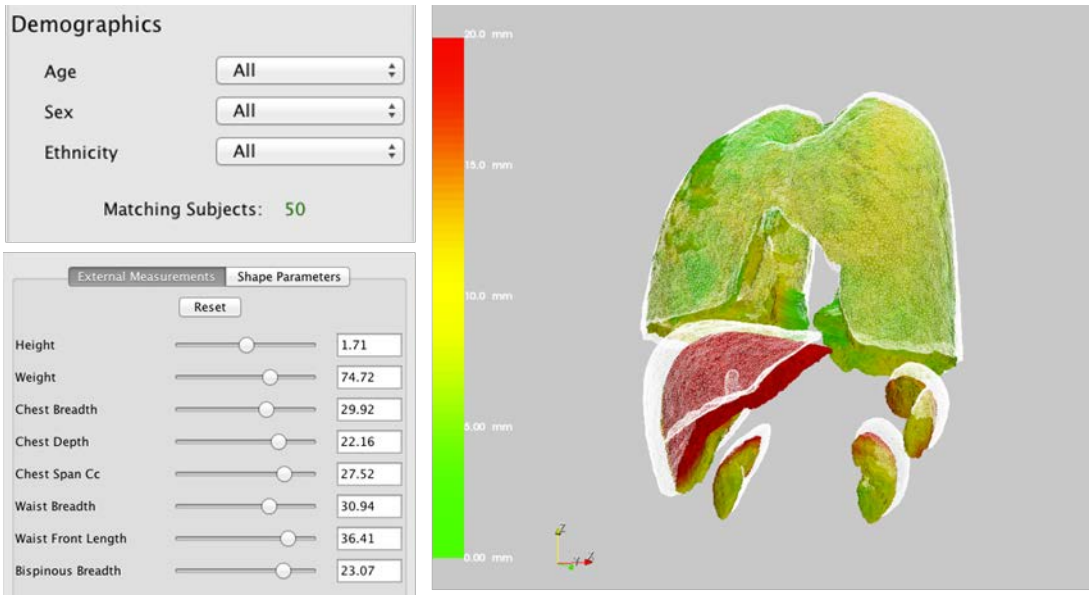

Figure 9. Comparison to baseline geometry based on distance

4. Discussion and Conclusion

This project has successfully designed and implemented a computational pipeline for generating multi-organ statistical shape atlases. The utility of the pipeline was demonstrated on a military-representative subject population which was collected and processed through the computational framework. Furthermore, a regression approach was identified to be able to predict internal structure from only external anatomical information. Finally, a visualization tool was developed to provide a way to qualitatively examine and analyze geometrical shape and statistical differences between demographic bins.

This work demonstrates the ability to generate a population-representative multi-organ statistical shape atlas, and initial feasibility using this atlas to predict the internal anatomy of specific population or individual subjects. The current statistical analysis is based on quantitative properties extracted from the data, such as gender information, demographic information, and discrete anatomical measurements calculated from boney landmarks. Ideally, an approach would be implementable where internal organ prediction models could be automatically generated from subject surface scans.

Future work should focus on improving key elements of the pipeline (e.g., segmentation), utilizing other imaging modalities (e.g., MRI), and expanding the training CT dataset to build more robust atlases. The goal is to utilize the statistical atlases generated by the pipeline as well as the statistical analysis tools to improve capabilities for developing highly-accurate human anatomical models employable for a variety of applications ranging from injury prediction models to improved methods for medical simulation, training, and treatment.

5. Acknowledgements

The authors would like to express their appreciation to the U.S. Army Natick Soldier Research, Development and Engineering Center (NSRDEC) for funding this work under Naval Sea Systems Command Contract No. N00024-13-D-6400, Task Order #VKW02. This paper is released under Distribution Statement A – Approved for Public Release; Distribution is unlimited.

---

[1] Otake, Y., Carneal, C.M., Lucas, B.C., Thawait, G., Carrino, J.A., Corner, B.D., Carboni, M.G., DeCristofano, B.S., Maffeo, M.A., Merkle, A.C., et al.: Supervised learning of anatomical structures using demographic and anthropometric informa- tion. In: Pattern Recognition Applications and Methods, pp. 225–240. Springer (2015).

[2] Saito, A., Nakada, M., Oost, E., Shimizu, A., Watanabe, H., Nawano, S.: A statis- tical shape model for multiple organs based on synthesized-based learning. In: Ab- dominal Imaging. Computation and Clinical Applications, pp. 280–289. Springer (2013).

[3] Oshima, R., Saito, A., Watanabe, H., Shimizu, A., Nawano, S.: Relaxed condi- tional hierarchical statistical shape model of multiple organs. In: Computing and Networking (CANDAR), 2013 First International Symposium on. pp. 288–293. IEEE (2013).

[4] Guimond, A., Meunier, J., Thirion, J.P.: Average brain models: A convergence study. Computer vision and image understanding 77(2), 192–210 (2000).

[5] Okada, T., Linguraru, M.G., Hori, M., Suzuki, Y., Summers, R.M., Tomiyama, N., Sato, Y.: Multi-organ segmentation in abdominal ct images. In: Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE. pp. 3986–3989. IEEE (2012).

[6] Linguraru, M.G., Li, Z., Shah, F., Chin, S., Summers, R.M.: Auto- mated liver segmentation using a normalized probabilistic atlas (2009), http://dx.doi.org/10.1117/12.810938.

[7] Linguraru, M.G., Sandberg, J.K., Li, Z., Shah, F., Summers,
R.M.: Automated segmentation and quantification of liver and spleen from ct images using normalized probabilistic atlases and enhancement estimation. Medical Physics 37(2), 771–783 (2010), http://scitation.aip.org/content/aapm/journal/medphys/37/2/10.1118/1.3284530.

[8] Gordon, Claire C., et al. 2012 Anthropometric Survey of US Army Personnel: Methods and Summary Statistics. No. NATICK/TR-15/007. ARMY NATICK SOLDIER RESEARCH DEVELOPMENT AND ENGINEERING CENTER MA, 2014.

[9] J. B. Tenenbaum, V. de Silva, J. C. Langford, A Global Geometric Framework for Nonlinear Dimensionality Reduction, Science 290, (2000), 2319–2323.

[10] Breiman, Leo. "Random forests." Machine learning 45.1 (2001): 5-32.

[11] Grady, Leo. "Random walks for image segmentation." IEEE transactions on pattern analysis and machine intelligence 28.11 (2006): 1768-1783.

[12] Pearson, K. (1901). "On Lines and Planes of Closest Fit to Systems of Points in Space". Philosophical Magazine. 2 (11): 559–572.

[13] Otake, Yoshito, et al. "Prediction of Organ Geometry from Demographic and Anthropometric Data based on Supervised Learning Approach using Statistical Shape Atlas." ICPRAM. 2013.