MODELING VARIABILITY IN TORSO SHAPE FOR CHAIR AND SEAT DESIGN

Matthew P. Reed
University of Michigan Transportation Research Institute
Ann Arbor, MI 48109
mreed@umich.edu

Matthew B. Parkinson
Engineering Design Program
Department of Mechanical Engineering
The Pennsylvania State University
University Park, PA 16802
parkinson@psu.edu

ABSTRACT
Anthropometric data are widely used in the design of chairs, seats, and other furniture intended for seated use. These data are valuable for determining the overall height, width, and depth of a chair, but contain little information about body shape that can be used to choose appropriate contours for backrests. A new method is presented for statistical modeling of three-dimensional torso shape for use in designing chairs and seats. Laser-scan data from a large-scale civilian anthropometric survey were extracted and analyzed using principal component analysis. Multivariate regression was applied to predict the average body shape as a function of overall anthropometric variables. For optimization applications, the statistical model can be exercised to randomly sample the space of torso shapes for automated virtual fitting trials. This approach also facilitates trade-off analyses and other the application of other design decision-making methods. Although seating is the specific example here, the method is generally applicable to other designing for human variability situations in which applicable body contour data are available.

INTRODUCTION
General Problem
Products intended to “fit” their users must be designed with careful consideration of the size and shape of the user population. Clothing is the most apparent example, but many other products, including bicycles, automobiles, and mobile phones, are developed with reference to anthropometric data that describe the distributions of human size and shape. For most applications, designers draw on reference texts that tabulate percentiles of various dimensions of interest [1]. More generally, limiting cases are identified that associate particular body dimensions with design variables [2]. For example, the width between the arms of an office chair should be larger than most users’ seated hip breadth.

However, most anthropometric data are only approximately related to the design variables of interest. Anthropometric data have customarily been gathered using tape measures, calipers, and other simple apparatus to obtain repeatable point-to-point and circumferential measures [3]. These dimensions are usually gathered in only one or two standardized postures, and consequently do not represent well the body dimensions that are important for many applications, such as automobile interior design [4, 5, 6]. Standard anthropometric data also do not represent body shape well and are of minimal value for designing the contours of products that are to fit snugly over part of the body. These interfaces between user and designed artifact can affect fit and other measures of performance such as safety [7].

In recent years, three-dimensional scanning technology has made feasible rapid recording of whole body shape. Most applications of body scan data are focused on the shapes of individuals. Body scans have been used to fit personalized garments [8, 9] and as part of mass customization of footwear [10]. Large-scale surveys of populations have provided the opportunity to use shape information directly in design. However, methods for using whole-body scan data for design purposes are not as well developed as those applied with standard anthropometric measures. This paper presents a systematic approach to analyzing and modeling body shape for design purposes, with a focus on
representing torso shape, which is not well described by conventional anthropometric measures.

The current analysis uses data from the Civilian American and European Anthropometric Resource (CAESAR), which contains information on 2400 U.S. men and women measured between 1999 and 2002 [11]. The participants were scanned in one standing and two seated postures while wearing minimal, tight-fitting clothing. The models presented here are based on data from a relaxed, unsupported seated posture.

Applications of the CAESAR body scan data have primarily relied on the selection of a small number of individuals from the dataset to use as virtual test subjects [12]. But these individuals have idiosyncrasies that may make them unsuitable to evaluate particular features of a design, and the appropriate way to weight evaluations performed with a small number of manikins is unclear. Boundary manikin approaches have been proposed to facilitate the evaluation process (e.g., [13, 14]) but these methods do not produce the desired quantification of accommodation if the selected anthropometric dimensions are not uniformly related to accommodation [15, 16].

For the design of products to be used by large populations, virtual evaluations using the body shapes of a large number of people would be preferred. One approach would use a random sample of body shapes from the target population, but this is impractical and inefficient. Body scan data are expensive to gather and the existing datasets (e.g., CAESAR) are generally not representative of any particular target population. Product evaluations or design criteria could be based on a weighted analysis, in which the results from each individual tested are weighted according to the likelihood of such an individual being sampled from the target population. However, some individuals of considerable anthropometric interest, such as those with high body weight for their stature, are not well-represented in the CAESAR data, and hence would have high sampling weights in such a scheme.

An alternative approach presented in this paper is to develop a statistical model of body shape that can be used to generate a wide range of synthetic torso shapes representing populations with varying characteristics. The resulting shapes can be selected to avoid the idiosyncrasies of individuals while accurately preserving the relationships between overall body dimensions, such as stature (standing height) and weight, and body shape. This approach parameterizes the data in a way that readily accommodates the application of rigorous design tools such as optimization.

Specific Problem: Backrest Design

The backrests of office chairs and vehicle seats are critical determinants of comfort in seating. A poorly designed backrest can lead to discomfort and poor product acceptance. One challenge in backrest design is that few of the standard anthropometric variables are relevant. For example, the anthropometric database used by the Business and Institutional Furniture Manufacturers Association to develop dimensional recommendations for office chairs [17] contains only a few linear dimensions, such as chest breadth and erect seated shoulder height, that are relevant to backrest design.

Backrest design for both office and vehicle seating is currently guided primarily by the subjective judgements of designers, by copying successful designs, and through repeated subjective trials of physical prototypes. This approach is time-consuming, expensive, and often yields uncomfortable seats. This paper describes some of the foundational activity needed to move to a system in which most of the evaluations for backrest fit are performed using virtual sitting trials, using computer manikins to represent a wide spectrum of potential users. Common problems, such as bolsters that are too narrow or upper-back contours that encroach on the shoulders, can be readily identified and eliminated prior to building expensive physical prototypes.

Virtual fitting has been conducted for more than two decades using scalable digital human figure models to represent occupants [18, 12]. A central problem with this approach has been that the manikin body shapes have been generated by coarse scaling of a standard, midsize body form, and the resulting shapes have been distinctly unrealistic. Moreover, a relatively small number of manikins are typically used (fewer than 20), leading to inaccurate and imprecise accommodation estimates.

In the current paper, we describe an approach to representing human size and shape for design applications that provides considerably more useful detail than previous methods. A model of human torso shape is developed that can provide a large number of statistically representative torso shapes to represent a target population characterized by gender mix and the distributions of a few overall body dimensions. Populations of virtual sitters developed using this model can be used to perform automated fitting trials of seats or chairs, providing an opportunity to conduct trade-off studies and optimize three-dimensional fit.

METHODS

The current methods were developed from those presented by [19]. A uniform mesh is fit to body scan data and a principal component analysis is performed on the mesh. Regression analysis is used to predict principal component scores from population descriptors, such as stature and body weight. Finally, virtual individuals are sampled from a statistical representation of the desired population and their body shape is generated from the principal-component model.

Generating a Standard Mesh

Prior to conducting a statistical analysis of body shape, the scan data must be mapped to a uniform representation. The raw CAESAR data are a high-resolution (mm-scale) irregular poly-
gon mesh with a large number of holes and other irregularities. Due to the design of the laser scanning system, the vertices of the mesh are organized in horizontal slices. The data reduction process used here, which differs substantially from that used by [19], exploits the slice-oriented structure of the data, as follows:

1. Extract slices at 10-mm vertical intervals between the seat surface and the top of the head.
2. Using a semi-automated procedure, strip off the data for the arms below the shoulders and the legs below the knees.
3. “Shrinkwrap” each data slice. This procedure uses an automated algorithm to create a continuous planar contour that interpolates across holes while preserving the real concavities in the surface contour.
4. Resample each slice, taking 120 points on each side of the body.
5. Align the slices laterally by shifting to compensate for body lean in the raw data.
6. Create a cylindrical mesh and resample to 60 slices between the seat and the top of the head.
7. Downsample each slice to 60 vertices using spline interpolation.

The result is that each subject’s torso data is represented by a $60 \times 60$ vertex mesh. Figure shows examples of torso data from several subjects, including a set of body landmarks that were recorded along with the body scans. These 92 landmarks are useful for estimating the locations of skeletal features and, in the current work, for tracking the statistical performance of the analysis and reconstruction algorithms. The results in this paper are based on analysis of 315 male and 449 female torsos from the CAESAR dataset. The sample includes all of the obese (BMI $\geq 30$ kg/m$^2$) participants, since their body dimensions are often limiting for backrest design. Note that this sample, and indeed the overall CAESAR dataset, was not designed to be representative of any design population of interest.

**Principal Component Analysis**

When confronting data with a high number of dimensions, principal component analysis (PCA) is a widely used tool to (1) express the data on an orthogonal basis that can be more readily analyzed, and (2) achieve data compression [20]. The latter results when most of the variance in the data is contained in the first few principal components, which occurs when many data dimensions are highly correlated. Geometrically, the first PC is the direction in the space of the data with the highest variance, the second is the direction orthogonal to the first PC with the highest variance, and so on. PCA is sometimes carried out using correlations, particularly when the data to be analyzed have different units or widely different scales. However, for the current application, analysis using the covariance matrix is more appropriate because it preserves scale [20].

The current PCA approach follows the methods presented by [19]. Separate analyses are conducted for men and women, since we expect gender specific patterns of body shape variation. The $60 \times 60$ vertex mesh ($60 \times 60 \times 3$ floating point numbers) is concatenated with the coordinates of 92 landmarks to obtain a geometry vector $g$ of length $l = (60 \times 60 + 92) \times 3 = 11076$ for each individual in the database. The geometry vectors for each individual are append to form the geometry matrix $G$ with dimensions for $N = 315$ men of $315 \times 11076$. Without loss of generality, the $G$ matrix is assumed to be centered by subtracting from each subject’s $g$ vector the overall mean $\overline{g}$.

The PCA can be computed as the eigendecomposition of the covariance matrix, but this can be computationally expensive (the covariance matrix is $11076 \times 11076$, in this case). Instead, we use an PCA algorithm by Turk and Pentland [21] that computes only the first $k$ principal components in descending order of variance. Although $k$ could take a value as large as $l$, the algorithm calls for setting $k = N$, significantly reducing the size of the matrix on which operations are performed. This choice is justified by the results below, which show that 99 percent of the variance in the data is accounted for by the first 60 principal components. For $k = l$,

$$G = SP,$$

where $S$ is an $N \times l$ matrix of principal component (PC) scores and $P$ is an $l \times N$ matrix, each column of which is a normalized principal component. In effect, the PCA projects the data into a new basis given by the principal components.
The PCA for the first $k$ principal components (i.e., the $k$ with the highest associated variance) yields

$$G^* = S_k P_k,$$  \hspace{1cm} (2)

where $G^*$ is an approximation of $G$, $S_k$ is the first $k$ columns of $S$, and $P_k$ is the first $k$ columns of $P$. With the Turk and Pentland algorithm, we compute $k = N$ PCs, so $S_k$ is square.

An individual’s geometry can be approximated by

$$g^* = \bar{g} + P_k^T S_k^T \varepsilon$$  \hspace{1cm} (3)

where $S_k$ is the row of the $S_k$ matrix corresponding to the $i$th individual’s principal component scores. Considerable compression is achieved by using only the first, say, 60 PC scores, rather than the individual’s full 11076 element geometry vector.

Regression Analysis

The PCA expresses the data in a relatively compact, orthogonal basis, which is amenable to statistical analysis. To predict the PC scores associated with a particular set of characteristics (such as body mass or sitting height), we create a linear model of the form

$$S_k = CF + \varepsilon^T$$  \hspace{1cm} (4)

where $F$ is a feature matrix with rows of vectors

$$f_i = [v_{1i}, v_{2i}, ..., 1]^T$$  \hspace{1cm} (5)

such that $v_{1i}$ is the value of the first subject descriptor for the $i$th subject and $\varepsilon^T$ is a column of residuals. The coefficient matrix $C$ is estimated using standard least-squares techniques by taking the Moore-Penrose pseudoinverse $F^+$ of $F$ such that

$$C = S_k F^+.$$  \hspace{1cm} (6)

For the current application, the design population can usually be described by gender mix (fraction of males and females in the population) and by the distributions within gender of stature (erect standing height) and body mass index (BMI). BMI is computed as body weight in kg divided by stature in meters squared and is widely used for characterizing weight-for-stature. Unlike body weight, BMI is only weakly correlated with stature, and hence is a better predictor to include in regression analyses along with stature. Although it is sometimes of interest to examine the statistical significance of these models using conventional procedures, those analyses do not ultimately affect these results and hence are not presented.

The outcome of the regression analysis is a model in the form

$$\hat{s} = Cf^T$$  \hspace{1cm} (7)

where $f$ is $[v_1, v_2, ..., 1]$ that gives the predicted PC scores conditioned on values of the predictors. The predicted torso shape is then constructed from Eq. 3. For the current work, one model is generated for each gender.

The residual variance not accounted for by the predictors ($\varepsilon$) can be considerable and is important to retain. We represent the residual variance for the current models by computing the standard deviation across subjects of the residuals on each PC. The residuals are independent and, for the current work, assumed to be normally distributed with zero mean.

Model Application

The male and female models are exercised to generate torso shapes by (1) inputting a stature and BMI to compute a PC-score vector, (2) optionally adding a random component obtained by sampling $k$ zero-mean normal distributions with standard deviations given by the $\varepsilon$ vector, (3) multiplying the PC-score vector by the PC matrix and adding $\bar{g}$ to obtain $g$, and partitioning $g$ to obtain coordinates of the torso mesh.

RESULTS

Number of Principal Components

As expected, an individual’s body shape could be represented to acceptable accuracy by a relatively small number of PCs. Figure 2 shows the cumulative fraction of variance accounted for by number of PCs for men and women. Using the current parameterization of body shape, 99% of variance is captured by the first 60 PCs for both men and women. In the results that follow, 60 PCs were included. Using more PCs preserves more individual variability, but for the current design application no more than 60 PCs are needed.

More important than the cumulative fraction of variance is how well a selected number of PCs preserves key body dimensions. A good error metric for this application is the torso width at key landmark locations. Table 1 lists the median error in the torso width at the hips, top of the pelvis, and shoulders for reconstructions using 60, 120, and 180 PCs. The median width error is approximately halved when going from 60 to 180 PCs. For men, the median error is below 5 mm for all three dimensions with 120 PCs. The median errors for women are larger, reflecting greater shape variability. Figure 3 shows a qualitative
Figure 2. Cumulative fraction of variance accounted for by \( k \) principal components for men (solid line) and women (dashed line).

Comparison of the reconstruction and original data for six women using 60 principal components. Idiosyncratic aspects of individuals’ body shapes are smoothed while the overall size and shape is retained.

**Generating Virtual Torsos**

The PCA/regression model can be exercised in a variety of ways to create torsos representative of people with desired body characteristics. Figure 4 shows a series of torso models created by holding stature constant and varying body mass index. Figure 5 shows a similar series for men. The male and female models readily capture the different effects of increasing body weight for men and women.

The residual variance can also be used to create a set of virtual people who share selected overall body characteristics. Figure 6 shows a set of torsos for women who have median stature and body weight. Note that the figures vary in body shape, yet are all approximately the same size.

**Virtual Fitting Trials for Target Populations**

As noted above, the CAESAR dataset and the subsample of 764 individuals used for the current modeling do not represent any design population of interest. To evaluate a product, a set of torso models that have a known relationship to the target population must be created, which usually is accomplished by creating a set of models that represent a random sample. First, the population is characterized by the gender mix and the within-gender distributions of stature and BMI. Second, these distributions are randomly sampled to obtain the desired number of individuals. Third, torsos are generated for each individual, using the target gender, stature, and BMI, and including the residual variance represented by random from a set of zero-mean normal distributions with variance given by the \( \varepsilon \) vector.

The process of using the torsos to evaluate products depends

Table 1. Median Reconstruction Error Between Pairs of Landmarks (mm)

<table>
<thead>
<tr>
<th>Landmark</th>
<th>Number of PCs</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
<td>60</td>
<td>120</td>
<td>180</td>
</tr>
<tr>
<td>Iliocristale (top of pelvis)</td>
<td>4.6</td>
<td>3.4</td>
<td>2.5</td>
<td>12.5</td>
</tr>
<tr>
<td>Trochanter (hips)</td>
<td>6.0</td>
<td>3.6</td>
<td>2.8</td>
<td>11.9</td>
</tr>
<tr>
<td>Acromion (shoulders)</td>
<td>6.6</td>
<td>4.9</td>
<td>3.1</td>
<td>8.9</td>
</tr>
</tbody>
</table>

Figure 3. Comparison of original (thin, green lines) and reconstructed (thick, blue lines) torsos. Reconstruction was performed using 60 PCs.

Figure 4. Synthesized female torsos generated by holding stature constant at 1621 mm and setting body mass index from 18 (upper left) to 45 (lower right) kg/m\(^2\) using 60 PCs.
DISCUSSION

This work presents the first application of human body scan data to seat and chair design that incorporates the variance in a large dataset, rather than extracting a few individuals. The relatively large number of people in the dataset, and their anthropometric diversity, means that the resulting models capture a large percentage of the variance in body shape that is important for chair and seat design. The approach is broadly applicable and can be readily extended to other problems for which body-scan data are available. It also facilitates trade-off analysis and has been developed to be readily implemented in a workflow incorporating optimization.

The analysis methods presented here are derived principally from [19], but the model construction, analysis, and application methods differ in several important ways. We use a cylindrical mesh model that simplifies the parameterization of the scan data. In principle, any consistent parameterization can be used. For example, we have conducted similar modeling using the control points of parametric surfaces (e.g., Non-Uniform Rational B-Splines). The cylindrical parameterization can also be generalized to represent the whole body [22], although the more complex surface mesh model used by Allen et al. (2003) has advantages for visualization. Our emphasis is to create models that can be represented simply in any CAD system and can also be subjected to additional statistical analysis. The simplicity of the parameterization used here has allowed with modest effort the extraction of data from 764 scans, rather than the 250 used by [19]. While we cannot easily define how many scans is “enough” outside of the context of a particular design analysis, adding more individuals provides higher confidence that the population variability in body size and shape is adequately represented. For example, the current dataset includes all of the obese individuals in the U.S. CAESAR dataset, who were included to improve the model performance for simulating the larger people who are often the limiting cases for chair and seat design. Our PCA methods follow [19] closely, including the selection of the Turk and Pentland algorithm. We differ in applying the results of the regression analysis to the modeling of target populations using the residual variance from the regressions, a natural but important extension of the earlier work. This methodology is motivated by functional anthropometric models developed for posture prediction [4]. Finally, the application of this analysis and modeling approach to ergonomics analysis and virtual fitting trials is novel, and represents an important advancement in the application of anthropometric data to seat and chair design.

The use of principal component analysis to reduce large datasets is a widely used technique that has been applied previously to modeling anatomical shapes (e.g., [23]), but the dependence of the resulting models on the selected parameterization of the geometry has not been widely discussed. In particular, the relationship between the high-variance PCs and the shape are strongly dependent on the distribution of the landmarks, ver-
Figure 7. Synthesized torsos used for virtual fitting trials in an office chair.
tices, or other features represented by the parameterization. The current work used a simple parameterization that distributed vertices evenly on each slice on each side of the body. Body landmark locations are parameterized separately from the mesh, so that the location of a particular landmark relative to the mesh can vary across individuals. This approach simplifies the fitting process, but means that detailed features are not well preserved. This tradeoff is acceptable for the current work, because the body surfaces of interest have relatively smooth contours, but a parameterization with more homologous points would be appropriate for modeling face contours, for example [24].

The modeling methods presented here differ from the standard approaches to shape analysis, which remove size (scale) prior to analyzing shape [25]. A typical shape analysis is intended to test hypotheses regarding group membership based on the relative positions of landmarks [26]. Removing scale allows shape to be differentiated among individuals who differ widely in size. Generalized Procrustes Analysis (GPA) is often a first step in such shape analyses. GPA standardizes scale and computes an optimal (in a least-squares sense) alignment or registration between the homologous landmarks of each individual in the sample [25]. PCA is often used to analyze the remaining deviations from the average landmark locations. The current method bypasses the GPA and instead uses an anatomical coordinate system to align the data from individuals prior to subtracting off the mean. Our approach is dramatically faster and is justified in place of Procrustes analysis when (1) a reasonable coordinate system can readily be defined, in this case exploiting bilateral symmetry, and (2) when the errors in the locations of the landmarks used to establish the coordinate system are small relative to the effects to be analyzed. But the most important difference between the current work and typical shape analysis is that our objective is not testing hypotheses regarding shape differences, but rather simultaneously modeling both size and shape, and hence we do not normalize scale prior to analysis.

The results presented here are based on analysis of the torso shape as scanned, but sitters’ postures when using chairs and seats can be somewhat different. Methods for morphing the torso to represent a range of seated postures have been developed [22] and statistical models to predict seated posture are available for some environments [27]. For virtual fitting trials, representing variance in posture can be as important as modeling anthropometric variability [15, 28, 16].

The development of statistical models of body contours provides the opportunity for more robust and automated design tools for creating products that fit their users well. The technology of automated fit assessments is still evolving and most applications have been in the apparel industry [29, 30]. As these methods are refined, we anticipate that virtual fitting trials will be used to improve product design and to reduce development time and cost.

ACKNOWLEDGMENT

This work was supported in part by Herman Miller, Incorporated.

REFERENCES


