Statistical Learning of Human Body through Feature Wireframe

Jida HUANG¹, Tsz-Ho KWOK²*, Chi ZHOU¹

¹ Industrial and Systems Engineering, University at Buffalo, SUNY, Buffalo NY, USA;
² Mechanical, Industrial and Aerospace Engineering, Concordia University, Montreal QC, Canada

DOI: 10.15221/17.337   http://dx.doi.org/10.15221/17.337

Abstract

Statistical learning of human body shape can be used for reconstructing/estimating body shapes from incomplete data, semantic parametric design, modifying images or videos, or simulation. It is applicable in many areas including computer vision & graphics, ergonomic design, personalized design, or virtual try-on. A digital human body is normally represented in a high-dimensional space, and the number of vertices in a mesh is far larger than the number of human bodies in publicly available databases, which results in a model learnt by Principle Component Analysis (PCA) can hardly reflect the true variety in human body shapes. Furthermore, if the number of vertices and size of database are large, it will be very challenging to perform PCA on such a huge problem. This paper presents a hierarchical method for statistical learning of human body by using feature wireframe as one of the layers to separate the whole problem into smaller and more solvable sub-problems. The feature wireframe is a collection of feature curves which are semantically defined on the mesh of human body, and it is consistent to all human bodies. A set of patches can then be generated by clustering the whole mesh surface to separated ones that interpolate the feature wireframe. Since the surface is separated into patches, PCA only needs to be conducted on each patch but not on the whole surface. The spatial relationship between the patches and the wireframe are learnt by linear regression. An application of semantic parametric design is used to demonstrate the capability of the method, where the semantic parameters are linked to the feature wireframe instead of the mesh directly. Under this hierarchy, the feature wireframe acts like an agent between semantic parameters and the mesh, and also contains semantic meaning by itself. The proposed method of learning human body statistically with the help of feature wireframe is scalable and has a better quality.

Keywords: Statistical learning, parametric design, 3D human body, PCA, feature

1. Introduction

Statistical learning of human body is a fundamental problem in many engineering applications such as biometric analysis, gaming animation and customized apparel design. Traditional way of acquisition of human body model is using 3D scanner to attain scans of realistic human model and then register the scans together to generate a consistent model. However, this process is expensive and time consuming, and would be rather difficult to obtain realistic model when there is lack of scanning device and reconstruction platform. Therefore, it would be convenient and applicable with a parametric design tool for human bodies so that the shape of a 3D human body can be generated from a set of given semantic parameters such as body height, chest-girth and waist-girth.

A lot of studies have been investigated on body shape and their poses. For example, Allen et al. [1] proposed a parametric freeform mesh design to reconstruct human model from range scans. Seo et al. [2] proposed an example-based approach to human body manipulation, which utilizes a modeling synthesizer that learns from preprocessed examples to interpolate new body geometry. In order to generate and animate realistic humans, Hasler et al. [3] proposed a learning based approach which could accurately model muscle deformations not only as a function of pose but also dependent on the physique of the subject. While Hasler et al. [4] presented a learning method for estimating a rigid skeleton for shape and pose, and Anguelov et al. [5] proposed a data-driven method – Shape Completion and Animation for People (SCAPE) – for building a human shape model that spans variation in both subject shape and pose. Another aspect of study focuses on the parametric design of human related product. Wang et al. [6] investigated a feature-based human model for digital apparel design. Kwok et al. [7] proposed an optimization algorithm for the complexes and the shape of common base domains in cross parameterization for reducing the distortion of the bijective mapping. Hasler et al. [8] studied a method to estimate the detailed 3D body shape of dressed humans. The approach is based on a space of human shapes learnt from a large database of registered body scans. Similarly, Li et al. [9] proposed a method for fitting a given 3D garment model onto human models of various body shapes and poses. Baek et al. [10] developed parametric human body shape modeling framework for human-centered product design.

* tszhokwok@concordia.ca ;   +1 (514) 848-2424 x3807
Our research is motivated by the work of Chu et al. [11], which is a linear system correlating the semantic parameters and human body. In their work, the PCA is conducted to reduce computational complexity. However, it may lose important feature information by conducting PCA on all the vertices of human body directly. In order to address this problem, this paper presents a hierarchical method for statistical learning of human body. Instead of correlating semantic parameters and human model directly, we use feature wireframe as an intermediate layer between the semantic parameters and human body model. By introducing the feature wireframe, a set of feature patches can be generated by clustering the whole mesh surface to separated ones that interpolate the feature wireframe. Since the surface is separated into patches, PCA only needs to be conducted on each patch but not on the whole surface. Under this hierarchy, we can construct a new model by fitting new wireframe to generate new patches of human model from a given set of semantic parameters. The proposed method can preserve feature information of database and integrate the wireframe and patch information into the generation of new models. Experimental results show that our method is less sensitive to system setting, more robust to extreme cases, and higher accuracy in model reconstruction compared to the previous work [11]. Our study is using the same human body database (78 female subjects) as used by Chu et al. [11], so that the conclusions are correct and accurate.

2. Parametric design model

This paper aims to develop a semantic parametric design method which constructs a human model conformed to a set of semantic parameters by learning through feature wireframe. Figure 1 sketches the idea of proposed hierarchical learning method for parametric design of human model.

![Flowchart of the proposed parametric design method]

It can be seen that there are three layers in the learning system. Instead of a complete human model, the third layer is a set of mesh patches, which is separated by the wireframe. Therefore, a set of semantic parameters is correlated with the wireframe of human models, and the boundary of each patch as a part of wireframe determines the relationship between the wireframe and all points in that patch. Hence, we construct a new statistical model based on a set of semantic parameters by integrating the feature wireframe. The proposed method not only can capture the characters of wireframe but also can reflect the feature of each patch. Therefore, it can avoid losing feature information of human model while solving smaller sub-problems.

Figure 2 presents the details of proposed algorithm in this paper. The proposed method includes two phases: statistical learning phase and parametric design phase. In the statistical learning phase, a set of human models in a database and the corresponding semantic parameters are given. Firstly, a set of wireframes is defined on the human model, which is detailed in Section 2.1. Secondly, with the wireframes, the human model is clustered and separated into a number of patches. Thirdly, we perform PCA on three sets of data: semantic parameters, wireframes and mesh patches, and obtain three batches of principle components: $SP$, $WF$ and $MP$, respectively. Finally, the correlation between each layer, i.e., the relation between $SP$ and $WF$, as well as $WF$ and $MP$, are learnt by linear regression, which will be discussed in Section 2.2 and 2.3.
In the parametric design phase, a set of user-designed new parameters are given as input to conduct the parametric design. Firstly, the input semantic parameters are projected onto the same eigenspace (E1) as that of the semantic parameters in the learning phase. Secondly, we use the relationship between $SP$ and $WF$ to obtain the $NewWF$ and subsequently get the $NewMP$ by using the correlation between $WF$ and $MP$. Thirdly, new patches of the predicted model can be reconstructed by an inverse projection on the eigenspace (E3) mapping mesh patches to $MP$. After we determined each patch of the new model, a smoothing step is conducted on the boundary of the patches to ensure continuity between patches. As a result, a new design model would be generated by a set of given semantic parameters. The technical details will be presented in Section 2.4.

2.1. Feature wireframe and patches extraction

The parametric design method relies on training models to capture 3D body shape with semantic parameters through wireframes. The training data of this paper are the full body 3D scans obtained from [11]. The feature-based human body parameterization technique [12] is applied to extract the semantic parameters for all the 78 subjects, and there are 48 parameters for each subject. Since all the human models have the consistent mesh connectivity, the bijective mapping or correspondence of points among human bodies is given. The synthesis of a new human body can be completed by generating a new mesh with the same connectivity but different vertex positions.

The wireframe is defined by a set of feature curves based on the anthropometrical rules. Specifically, the selected feature curves selected are the girths of head, neck, neck-base, elbow, wrist, bust, under-bust, thigh, knee, and ankle, as well as armhole circumstance. The human models are then separated into feature patches according to the defined wireframe. Figure 3 depicts the process of wireframe definition and patch generation.
2.2. Correlation between wireframe and semantic parameters

The goal in this section is to learn the relationship between the semantic parameters and the wireframe of human models as shown in Fig. 4. Demonstrated by Fig. 2, the correlation between any two data are done on their principal components, so the details of Principal Component Analysis (PCA) will be briefly discussed with wireframe first, and it is followed by the correlation. One of the reasons to apply PCA on the data is for dimensionality reduction, which is a very useful step for processing high-dimensional datasets, while still retaining as much of the variance in the dataset as possible.

Assume there are \( m \) scanned models in the database, the wireframe on human model defined by a set of vertices can be collected as 

\[
\mathbf{h} = \begin{bmatrix} h_1 \ h_2 \ldots \ h_m \end{bmatrix}_{3n_1 \times 1},
\]

where \( h_i \) is a vector with \( n_1 \) vertices from the wireframe of the \( i \)th human model in the database. Letting

\[
\mathbf{\bar{h}} = \frac{1}{m} \sum_{i=1}^{m} h_i,
\]

we have \( \bar{H} = \begin{bmatrix} (h_1 - \bar{h}) & (h_2 - \bar{h}) & \ldots & (h_m - \bar{h}) \end{bmatrix} \), and its covariance is \( V = \bar{H}(\bar{H})^T \), which dimension is \( 3n_1 \times 3n_1 \). Since \( 3n_1 \gg m \), we instead compute the transpose of its covariance \( W = (\bar{H})^T\bar{H} \). Then by eigenvalue decomposition, we could have \( Wx = \lambda x \). Then \( m \) eigenvectors \( \{x_j\}_{m \times 1} \) could be obtained, with \( x_j \), the \( j \)th eigenvector of \( C \) can be determined: \( y_j = \bar{H}x_j \), here \( y_j \) is a \( 3n_1 \times 1 \) vector.

The normalized eigenvectors \( \tilde{y}_j = y_j/\|y_j\| \) (\( j = 1, ..., m \)) are the principal vectors of \( \bar{H} \), where each is associated with a variance \( \sigma_j \). The vectors are sorted according to \( \sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_m \). The largest variance means the corresponding vector \( y_j \) has the most dominant effect in the model space. We keep the first \( k_1 \) principal components according to the percentage of the total variance explained \( r \) by each principal component.

\[
r = \frac{\lambda_1 + \lambda_2 + \ldots + \lambda_{k_1}}{\lambda_1 + \lambda_2 + \ldots + \lambda_m} \geq 0.95
\]

Wireframe of each human model of all scanned models serving as training data set can be projected onto \( k_1 \)-dimensional points by:

\[
g_j = \begin{bmatrix} \tilde{y}_j^T_1 \\ \tilde{y}_j^T_2 \\ \vdots \\ \tilde{y}_j^T_{k_1} \end{bmatrix} (h_j - \bar{h})
\]

Thus, we map \( H_{3n_1 \times m} \) into a reduced matrix \( WF_{k_1 \times m} = [g_j] (k_1 \ll 3n_1) \) spanning the linear space of exemplar patches of human bodies, named as the reduced exemplar matrix. After PCA, a wireframe \( h_i \) is projected to a \( k_1 \)-dimensional point \( g_i \) and its collection is the principal components \( WF \).

Similarly, suppose we have \( Z \) semantic parameters, PCA can be applied to the semantic parameters to obtain \( z \) principle components, i.e., \( SP = [s_1 \ s_2 \ldots \ s_m]_{z \times m} \) where \( s_i \) is a \( z \times 1 \) vector. The correlation between \( SP \) and \( WF \) is represented by a system of linear equations:

\[
Rg_j + \mu = s_j \quad (j = 1, ..., m)
\]

where \( R \) is the relation matrix (\( z \times k_1 \)) and the vector \( \mu \) is a corresponding residual (\( z \times 1 \)). By picking \( k_1 \) such that \( m > k_1 + 1 \), the value of \( R \) and \( \mu \) can be determined by least-square method. In this step, the relation between \( SP \) and \( WF \) is constructed.
2.3. Correlation between patches and wireframe

As aforementioned, the wireframe is used to cluster and separate the human body into mesh patches, so that the whole problem can be divided and solved by a number of sub-problems. Together with the separated patches \( \{mp_i\} \in MP \), the wireframe is also separated into different parts \( \{wf_i\} \in WF \) correspondingly as the boundary of each patch, and one example is shown in Fig. 5.

\[
\begin{align*}
\text{Figure 5 Relation between patch and boundary wireframes}

\text{The correlation is then built separately on each of the patches to its own boundary. Noted that the semantic parameters are linearly linked with the whole wireframe in the previous section, but a mesh patch is linked with only part of the wireframe here.}

\text{Similar to the process for wireframe discussed above, we could map each patch } p_j \text{ with } n_2 \text{ vertices into a } k_2\text{-dimensional point } q_j. \text{ Accordingly, we map each patch of human models } P_{3n_2 \times m} \text{ with } \bar{p} = \frac{1}{m}\sum_{j=1}^{m} p_j \text{ into a reduced matrix } MP_{k_2 \times m} = [q_j] (k_2 \ll 3n_2) \text{ spanning the linear space of exemplar patches of human bodies by}
\end{align*}
\]

\[
q_j = \begin{bmatrix}
\hat{X}_j^T \\
\hat{Y}_j^T \\
\hat{Z}_j^T
\end{bmatrix} (p_j - \bar{p})
\]

Suppose we have \( t \) points on the boundaries of a patch, listing the parameters for all \( m \) examples forms a boundary matrix: \( B = [b_1, b_2, \ldots, b_m]_{|\times m} \) with \( b_i \) a \( 1 \times 1 \) vector. The correlation between \( MP \) and \( B \) can be represented by a system of linear equations:

\[
Kd_j + v = b_j \quad (j = 1, \ldots, m),
\]

where \( K \) is the relation matrix \( (t \times k_2) \) and the vector \( v \) is a corresponding residual \( (t \times 1) \). For \( m > k_2 + 1 \), the value of \( K \) and \( v \) can be determined by the least-square method. This is repeated for each of the patches, and thus each patch will give one linear equation system, i.e.,

\[
K_i \cdot wf_i + u_i = mp_i \quad (i = 1, \ldots, t),
\]

where the number of patches is \( t \), and \( \{wf_i \in WF\} \) and \( \{mp_i \in MP\} \).

2.4. Parametric design model

The previous sections have discussed how to build the correlations between semantic parameters and wireframes, as well as between patches and wireframe, in the learning phase. Therefore, a hierarchical learning model for human body is constructed, and the database is used to train the model and record the related parameters. In this section, we will discuss the details in the design phase making use of the trained parameters to synthesize new models.

Given a new set of new semantic parameters projecting to the same eigenspace in the learning process to get the new principal components, \( s_{\text{new}} \), the corresponding coefficient in the reduced linear space \( WF \) can be computed by the inverse of Eq. (4), i.e.:

\[
g_{\text{new}} = (s_{\text{new}} - \mu)R^T
\]

Then, the new wireframe model could be predicted using Eq. (3):

\[
h_{\text{new}} = [\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_{k_2}]g_{\text{new}} + \bar{h}
\]
After the new wireframe has been computed, all of the boundaries could be extracted. Given this new set of boundary points, \( b_{\text{new}} \), the corresponding coefficient in the reduced linear space \( MP \) can be computed by Eq. (6):

\[
q_{\text{new}} = (b_{\text{new}} - v)K^T \quad (i = 1, 2, ..., T)
\]  

(10)

Then, the new patch model can be reconstructed using Eq. (5):

\[
\hat{p}_{\text{new}} = [\hat{x}_1 \quad \hat{x}_2 \quad ... \quad \hat{x}_{k2} | q_{\text{new}} + \tilde{p}] \quad (i = 1, 2, ..., T)
\]  

(11)

Hence, given a new set of semantic parameters, we can generate the new wireframes. Once the new wireframes are determined, the boundary wireframes can also be extracted, with the regression method of boundaries and patch. The new vertices in each patches can then be computed, and they are assembled to an intact human model. However, since we generate the new human model by synthesizing each vertex on different patches, this would lead to slight discontinuity between patches. As a post-processing step, we apply Laplacian smoothing operator on the boundary of the patches in reconstructed model to smooth out the artifacts caused by separated reconstruction of patches.

3. Experimental results

In this paper, we have proposed a new hierarchy method for statistical learning of human models. In order to verify the validness and effectiveness of the method, several experiments are conducted to investigate the performance of the proposed methodology compared with the previous method [11].

3.1. Sensitivity to System Setting

In the dataset, there are 47 semantic parameters like body height, neck girth, bust, waist and etc. have been collected to represent the human bodies. Intuitively, more input semantic parameters would characterize the human model more precisely. However, on the other hand, since the correlations are built by linear systems, too many parameters may over affect the generation of wireframe and impact the synthesis of human model. Therefore, as reported by Chu et al. [11], the learning system is very sensitivity to the number of parameters used.

To study the effect of number of semantic parameters and PCA on semantic parameters, we randomly select a model from the database, and then use 4, 8 and 12 semantic parameters to reconstruct the model respectively. The results are compared with the previous method [11] in Fig. 6.

![Figure 6: Reconstructing human model with different number of semantic parameters.](image)

It can be seen that the previous method [11] is indeed very sensitive to the number of parameters used to train the system. When the number of parameters is 4 and 12, their method cannot synthesize a reasonable result. In contrast, the results generated by our method are better than theirs with different number of parameters, so our method is less sensitive to the setting of the system, and thus has a better repeatability and is more practical. When the number of parameters is 8, both methods can obtain a reasonable model, but our method is closer to the original model. It can be also seen that too many semantic parameters would affect the generation of human model, this is mainly because the correlations between each layers are linear, too much parameters could over affect the generation of next layer. In the rest of experiments, the number of parameters used is set as 8.
3.2 Robustness to Extreme Case

In this experiment, we test the robustness of methods under some extreme cases in terms of design parameters. The range of each design parameters in the dataset are listed in Table 2, and we conduct two cases. The first one is the inseam design: from the dataset the range of inseam parameter is [63, 84], and we test whether the methods can still work when the input semantic parameters are not in the range of original dataset. In the second test, we use the largest body height and body width, and the smallest neck girth in the design. With the parametric design values in two cases as shown in Table 2, we conduct the extreme case study and compared the robustness of our method with the previous method [11].

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Body height</th>
<th>Neck girth</th>
<th>Width</th>
<th>Bust</th>
<th>Under bust</th>
<th>Waist</th>
<th>Hip</th>
<th>Inseam</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Range (cm)</td>
<td>[145, 181]</td>
<td>[29, 44]</td>
<td>[29, 63]</td>
<td>[79, 117]</td>
<td>[70, 107]</td>
<td>[60, 115]</td>
<td>[85, 126]</td>
<td>[63, 84]</td>
</tr>
<tr>
<td>Case 1</td>
<td>175</td>
<td>32</td>
<td>32</td>
<td>85</td>
<td>76</td>
<td>67</td>
<td>91</td>
<td>62</td>
</tr>
<tr>
<td>Case 2</td>
<td>181</td>
<td>29</td>
<td>63</td>
<td>90</td>
<td>86</td>
<td>88</td>
<td>85</td>
<td>70</td>
</tr>
</tbody>
</table>

Figure 7 shows the comparison results of two methods. In case 1, the previous method [11] generates a model with two legs overlapped because of the small inseam value. In case 2, their method gives a result with the arms stretched because of large width value. In both cases, our method is able to obtain reasonable models. From this experiment we can see that the proposed method is more robust than the previous method [11] and can work better with extreme cases such as parameters do not lie in the range of known dataset.

3.3. Accuracy of Synthesis

To validate our trained system can capture more feature information, we compare the accuracy of synthesis results with the previous method [11]. To compare the methods on the dataset comprehensively, we select 20 models from the scanned human dataset and then use the same parameters of original model to synthesize new model. In order to depict the synthesis accuracy of two methods, the $L^2$ norm is used:

$$L^2(g) = \| \mathbf{h} - h \|^2$$

where $\mathbf{h}$ is the synthesized human model by the same parameters of model $h$ in the dataset. Table 2 shows the comparison of $L^2$ norm for both methods on 20 models.
Table 2 Comparison of $L^2$ error

<table>
<thead>
<tr>
<th>Model</th>
<th>Chu et al. [11]</th>
<th>Our</th>
<th>Improve</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>578.1772</td>
<td>224.6850</td>
<td>61.14%</td>
</tr>
<tr>
<td>2</td>
<td>435.5650</td>
<td>313.7772</td>
<td>27.96%</td>
</tr>
<tr>
<td>3</td>
<td>764.6166</td>
<td>502.7911</td>
<td>34.24%</td>
</tr>
<tr>
<td>4</td>
<td>688.7536</td>
<td>287.7786</td>
<td>58.22%</td>
</tr>
<tr>
<td>5</td>
<td>607.6612</td>
<td>301.7146</td>
<td>50.35%</td>
</tr>
<tr>
<td>6</td>
<td>443.0983</td>
<td>287.1589</td>
<td>35.19%</td>
</tr>
<tr>
<td>7</td>
<td>581.0557</td>
<td>330.4439</td>
<td>43.13%</td>
</tr>
<tr>
<td>8</td>
<td>467.8834</td>
<td>204.9625</td>
<td>56.19%</td>
</tr>
<tr>
<td>9</td>
<td>582.6971</td>
<td>416.9059</td>
<td>28.45%</td>
</tr>
<tr>
<td>10</td>
<td>613.53</td>
<td>182.1251</td>
<td>70.32%</td>
</tr>
<tr>
<td>11</td>
<td>703.9862</td>
<td>158.6499</td>
<td>77.46%</td>
</tr>
<tr>
<td>12</td>
<td>537.6945</td>
<td>276.5605</td>
<td>48.57%</td>
</tr>
<tr>
<td>13</td>
<td>626.0728</td>
<td>245.2357</td>
<td>60.83%</td>
</tr>
<tr>
<td>14</td>
<td>680.7657</td>
<td>207.6034</td>
<td>69.50%</td>
</tr>
<tr>
<td>15</td>
<td>585.2951</td>
<td>365.9956</td>
<td>37.47%</td>
</tr>
<tr>
<td>16</td>
<td>621.4558</td>
<td>272.3597</td>
<td>56.17%</td>
</tr>
<tr>
<td>17</td>
<td>587.4541</td>
<td>296.7489</td>
<td>49.49%</td>
</tr>
<tr>
<td>18</td>
<td>596.1429</td>
<td>319.2569</td>
<td>46.45%</td>
</tr>
<tr>
<td>19</td>
<td>577.7599</td>
<td>224.2557</td>
<td>61.19%</td>
</tr>
<tr>
<td>20</td>
<td>449.3290</td>
<td>365.0120</td>
<td>18.77%</td>
</tr>
</tbody>
</table>

It can be seen from Table 2 that our proposed method performs better than the previous method [11] on all the selected models, which decrease the synthesis error from 18.77~77.46%. Figure 8 shows the synthesis results of model 4 by the two methods, from which we can visually see that the result by our method is closer to the original model.

![Figure 8 Comparison of parametric design on Model 4.](image)

Figure 9 shows the Hausdorff distance using color map on the model 7 synthesized by two methods, from which we can see that the model by the previous method [11] have deformations especially on the head and arms, while our method generates a very close result.
The results show that the proposed method outperforms the previous method, and this is because our new hierarchy method uses the wireframe as an interlayer and conducts PCA on patches instead of on whole surface of human body directly, hence it can avoid losing feature information of original model and better represent the surface of human model. As a result, our method has a higher accuracy.

4. Conclusion

In this paper, we proposed a hierarchy method for learning the correlation of semantic parameters and the human bodies, instead of building relationship between semantic parameters and vertices of human body, we added a wireframe of human body as intermedia layer. The correlation between the semantic parameters, wireframes as well as patches are studied and a hierarchy learning model is investigated from the database. The proposed framework enables the new human model design for a given set of new parameters. The experimental results show the method outperforms the previous method and could capture more feature information of human bodies. Our method is less sensitive to system settings, more robust to extreme cases, and higher accuracy in model reconstruction.

Future work could be studied in several aspects: more data can be integrated in the method which enables a generalized parametric design of human model. In current method, we only considered linear relationship between each layer, a comprehensive deep learning method could be studied for the relationship between each layers. In addition, only mesh is studied in this work, if pose information and shape deformation are considered simultaneously, a more detailed framework could be built and investigated.

References


