Part-Based 3D Face Morphable Model with Anthropometric Local Control

Abstract
We propose an approach to construct realistic 3D facial morphable models (3DMM) that allows an intuitive facial attribute editing workflow. Current face modeling methods using 3DMM suffer from the lack of local control. We thus create a 3DMM by combining local part-based 3DMM for the eyes, nose, mouth, ears, and facial mask regions. Our local PCA-based approach uses a novel method to select the best eigenvectors from the local 3DMM to ensure that the combined 3DMM is expressive while allowing accurate reconstruction. The editing controls we provide to the user are intuitive as they are extracted from anthropometric measurements found in the literature. Out of a large set of possible anthropometric measurements, we filter the ones that have meaningful generative power given the face data set. We bind the measurements to the part-based 3DMM through mapping matrices derived from our data set of facial scans. Our part-based 3DMM is compact yet accurate, and compared to other 3DMM methods, it provides a new trade-off between local and global control. We tested our approach on a data set of 135 scans used to derive the 3DMM, plus 19 scans that served for validation. The results show that our part-based 3DMM approach has excellent generative properties and allows intuitive local control to the user.

CCS Concepts
• Computing methodologies → Shape modeling;

1. Introduction
Authoring realistic 3D faces with intuitive controls is used in a broad range of computer graphics applications such as video games, person identification, facial plastic surgery, and virtual reality. This process is particularly time-consuming given the intricate details found in the eyes, nose, mouth, and ears. Consequently, it would be convenient to use high-level controls, such as anthropometric measurements, to edit human-like character heads.

Many methods use 3D morphable face models (3DMM) for animation (blend shapes), face capture, and face editing. Even though face animation concerns are important, our work focuses on the editing of facial meshes. 3DMMs are typically constructed by computing a Principal Component Analysis (PCA) on a data set of scans sharing the same mesh topology. New 3D faces are generated by changing the relative weights of the individual eigenvectors. These methods are popular due to the simplicity and efficiency of the ap-
We propose a measurement selection technique to bind the essential worst than our approach for modeling purposes (considering a data set composed of animations for a single person), but perform worst than our approach for modeling purposes (considering a data set made of neutral faces from different persons).

We propose an approach to constructs realistic 3DMMs. We increase the controllability of our faces by segmenting them into independent subregions and selecting the most dominant eigenvectors per part. Furthermore, we rely on facial anthropometric measurements to derive useful controls to use our 3DMM for editing faces. We propose a measurement selection technique to bind the essential measurements to the 3DMM eigenvectors. Our method allows the user to edit faces by adjusting the facial parts using sliders controlling the values of anthropometric measurements. The measurements are mapped to eigenvector weights, allowing us to compute the individual parts matching the values selected by the user. Finally, the reconstructed parts are seamlessly blended together to generate the desired 3D face. We present experimental evidence to demonstrate how these tailored 3DMMs are preferred over the global PCA models.

2. Related Work

3D morphable models are powerful statistical models widely used in many applications in Computer Vision and Computer Graphics. One of the most well-known previous work is proposed by Blanz and Vetter [BV99]. Their pioneer work proposes a model using PCA from face scans. Although they propose a multi-segment model and decompose a face into four parts to augment expressiveness, the PCA decomposition is computed globally on the whole face. Other global PCA methods have been proposed [ACP03; CWZ*14; BRZ*16; LBB*17; BRP*18; LGJV18]. A downside of global PCA-based methods is that they exhibit global support: adjusting parameters to change one part has undesirable changes on other regions such as the torso and arms. In contrast, another downside is a lack of intuitive user control for face editing. While the eigenvectors are good at extracting the dominant modes of variation of the data, they provide weak intuitive interpretation.

To address the former problem, several local methods have been proposed that segment the face into independent sub-regions and select the most dominant eigenvectors per part. Tena et al. [TDM11] propose a method to create localized clustered PCA models for animation. They select the location of the basis using a spectral clustering on the geodesic distance and correlation of vertex displacement considering variation in the expressions. Their method requires a manual step to adjust the boundaries of the segments, making it not so different than ours where the parts are user specified. Chi et al. [CGZ17] adaptively segments the face model into soft regions based on user-interaction and coherency coefficients. Afterwards, they estimate the blending weights which satisfy the user constraints as well as the spatio-temporal properties of the face set.

Here too, the required user intervention renders the segmentation not so different than our user provided segments. SPLOCS [NVW*13] propose the theory of sparse matrix decompositions to produce localized deformation from an animated mesh sequence. They use vertex displacement in the Euclidean coordinates to select the basis in a greedy fashion. We noticed that when considering variation in identity instead of variation in expression, the greedy selection leads to basis which are far less local than those from the method of Tena et al. [TDM11] and ours. These papers address facial animation instead of face modeling and therefore assume large yet localized deformations caused by facial expressions which are different from our context where each face is globally significantly different from the others.

Cao et al. [CCWL18] segment the face with the same spectral clustering followed by manual adjustment as Tena et al. [TDM11]. While their method focuses mostly on expression, they also provide some identity modeling, as they rely on the FaceWarehouse [CWZ*14] global model, which they decompose using the segments defined by the spectral clustering. In their case, the goal is to do the real-time adaptation of a 3DMM to a face from a video feed. While their method works remarkably well for the real-time application of “virtual makeup”, it does not provide a very detailed facial model compared to ours, and it does not support a face editing workflow.

Other papers supplement decomposition approaches with the extraction of fine details, allowing to reconstruct a faithful facial model [CBZB15; GZC*16; SWTCC14]. The major problem with these approaches is that they work for a specific person and do not provide editing capabilities. The Phace [IKKP17] method allows the user to edit fat or muscle maps in texture space on the face. While this provides a physically-based adjustment, the control is implicit. The user modifies the texture and then the system simulates muscles and fat to get the result.

Wu et al. [WBGB16] propose an anatomically-constrained local deformation model to improve the fidelity of monocular facial animation. Their model uses 1000 overlapping parts and then decouple the rigid pose of the part from its non-rigid deformation. While this approach works particularly well for reconstruction, the parts are too small for editing semantic face parts such as the nose or the eyes.

As opposed to the methods described so far, the methods of Allen et al. [ACP03] and BodyTalk [SQH*16] greatly ease the editing by mapping intuitive features to modifications of global 3DMM eigenvector weights. In particular, BodyTalk [SQH*16] relates transformations of the meshes to keywords such as “fit” and “sturdy”. While the mapping between the words and the deformations is not perfect, it still makes it reasonably intuitive to edit the mesh of the body. One problem with this method, is that it provides words for bodies, not faces. A second major problem is the inability to do local adjustments; adjustments that increase the length of the legs will result in changes to other regions such as the torso and arms. In contrast, for our approach, we aim at providing local control in the editing.

A downside of global PCA based methods is that they exhibit global support: adjusting parameters to change one part has unwanted effects on other unrelated parts. To address this problem, we segment the face into independent sub-regions and provide a...
We separate the face into regions and compute independent PCA where eigenvectors, $P$ and $\mathbf{v}$, are the number of vertices of the face mesh, $n_d$ is the number of vertices for part $d$, $V_j$ is on the ground truth and $V_j'$ is the corresponding point on the blended face. We compute averages over all vertices and per part to ensure that parts with more vertices do not end up using most of the eigenvector budget at the expense of parts with fewer vertices. We do so for the entire data set and for a set of 19 validation faces that were not part of the training data set. We compute the median data set error as well as the median per part reconstruction and the whole face reconstruction. To evaluate the accuracy of our selection, we construct the facial parts (Eq. 1) and blend them together (Sec. 5) to generate the whole face. Afterwards, we assess the accuracy of reconstruction by calculating the average of the geometric error $D_{GE}$ between the ground truth and the blended face. We first do a rigid alignment step (rotation and translation) between the facial parts of the ground truth and the blended result. We then record the average per-vertex Euclidean distance over all vertices and per part:

$$D_{GE_{\text{part}}} (S') = \frac{1}{n_{\text{det}}^{\text{part}}} \sum_{d=1}^{5} \sum_{V_j \in S_d} \left| V_j - \left( R_d V_j' + T_d \right) \right|^2$$

$$D_{GE_{\text{whole}}} (S') = \frac{1}{n_{\text{det}}} \sum_{d=1}^{5} \sum_{V_j \in S_d} \left| V_j - \left( R_d V_j' + T_d \right) \right|^2$$

$$D_{GE} (S) = \frac{D_{GE_{\text{part}}} (S') + D_{GE_{\text{whole}}} (S')}{2}$$

where $n_{\text{det}}$ is the number of vertices of the face mesh, $n_{\text{det}}^{\text{part}}$ is the number of vertices for part $d$, $V_j$ is on the ground truth and $V_j'$ is the corresponding point on the blended face. We compute averages over all vertices and per part to ensure that parts with more vertices do not end up using most of the eigenvector budget at the expense of parts with fewer vertices. We do so for the entire data set and for a set of 19 validation faces that were not part of the training data set. We compute the median data set error as well as the median per part error, and we average the two in a global error.

At each step of our incremental eigenvector selection, we decide which of the five parts will get a new eigenvector added to its set. We compare the geometric errors resulting from each of the five parts with the corresponding point on the blended face. We compute averages over all vertices and per part to ensure that parts with more vertices do not end up using most of the eigenvector budget at the expense of parts with fewer vertices. We do so for the entire data set and for a set of 19 validation faces that were not part of the training data set. We compute the median data set error as well as the median per part error, and we average the two in a global error.

At each step of our incremental eigenvector selection, we decide which of the five parts with the greatest impact on decreasing the error. When eigenvectors from multiple parts result in similar decreases in error, instead of systematically picking the eigenvector based on lowest error, we select by sampling from a discrete probability density function (PDF) created from the respective decreases in error of the five candidate eigenvectors. This PDF selection process creates a more even distribution of eigenvectors across the parts and it maintains a low error. As we iterate, the reconstruction error decreases. For the female and male data set faces, the average reconstruction errors are 2.00 and 2.13 mm when considering zero eigenvector. The errors decrease to 0.75 and 0.74 mm after 80 iterations, and when considering all eigenvectors the errors are 0 mm. We chose an error threshold of 0.75 and 0.74 mm after 80 iterations, and when considering all eigenvectors the errors are 0 mm. We chose an error threshold of

process to select the best set of eigenvectors given a target number of eigenvectors. Methods that segment the face in subregions target facial animation instead of modeling. We will demonstrate that our approach is better suited to the task of face editing than these methods. Another problem of most of the previous work is that it does not allow facial model editing through the adjustment of objective measurements. In contrast, our method relies on anthropometric measurements used as controls for the editing. Furthermore, we propose a process to select the right set of anthropometric measurements for each facial part.

3. Overview

In this paper, we introduce a pipeline for constructing a 3DMM. We separate the face into regions and compute independent PCA decomposition on each region. We then combine the per region 3DMMs, paying particular attention to select the most dominant eigenvectors across the eigenvectors of the different regions. While the eigenvectors are good at extracting the dominant modes of variation of the data, they provide weak intuitive interpretation. We thus use anthropometric measurements to provide human understandable adjustments of the face. The reconstruction from the measurements is done through a mapping from the measurements to the weights that need to be applied to each eigenvector. Out of the set of measurements we extracted from our survey of the literature, we selected a subset which proved to provide the least reconstruction error. The overview of our approach can be found in Fig. 1.

The remainder of this paper is organized as follows: Sec. 4 describes how 3DMMs are constructed, including the face decomposition and selecting the most dominant eigenvectors. Afterwards, we discuss how to reconstruct a face by smooth blending of different facial parts (Sec. 5). In Sec. 6, the selection of the anthropometric measurements, and the mapping between these measurements and the PCA eigenvectors are discussed. We demonstrate the results in Sec. 7 and discuss them in Sec. 8.

4. 3D Morphable Face Model

We employ PCA on a data set of scans (Sec. 8.3) to construct our 3DMMs. We propose to segment the face into different parts in order to focus the decomposition on a part-by-part basis instead of computing the PCA decomposition on the whole face. We compute the decomposition separately for the male and female subsets. As is shown in Fig. 1, we decompose the face into five parts: eyes, nose, mouth, ear, and what we refer to as the facial mesh (grouping the remaining areas such as cheeks, jaws, forehead, and chin). We further discuss this design choice in Sec. 8.2. This face decomposition allows us to have eigenvectors for each part. The geometry of the facial parts is presented with a shape-vector $S_d = [V_1 \ldots V_n] \in \mathbb{R}^{n_d}$, where $n_d$ is the number of vertices of $d$th facial part, $d \in \{1, \ldots, 5\}$, and $V_i = [x_i, y_i] \in \mathbb{R}^2$ defines the $x$, $y$, and $z$ coordinates of the $i$th vertex. After applying PCA, each facial part $d$ is reconstructed as:

$$S'_d = S_d + \sum_{j=1}^{n_e} P_j b_j$$

where $S_d$ is the mean shape of $d$th facial part, $n_e$ is its number of eigenvectors, $P_j$ is an eigenvector of size $3n_e$, $b$ is a $n_e \times 1$ vector containing the weights of the corresponding eigenvectors, and $S_d'$ is the reconstruction, which will be an approximation when not using all eigenvectors.

Our approach selects the smallest set of eigenvectors that still reconstructs accurately the shape. We accomplish this by incrementally adding the eigenvectors in the order of their significance to the reconstruction until a certain accuracy is met. Even though we rely on the eigenvalues to sort the eigenvectors for each part, we provide the user with a measurable error (in mm) which is more precise than relying solely on eigenvalues across different parts. We determine the best set of eigenvector to achieve a balance between the quality of the per part reconstruction and the whole face reconstruction. To evaluate the accuracy of our selection, we construct the facial parts (Eq. 1) and blend them together (Sec. 5) to generate the whole face. Afterwards, we assess the accuracy of reconstruction by calculating the average of the geometric error $D_{GE}$ between the ground truth and the blended face. We first do a rigid alignment step (rotation and translation) between the facial parts of the ground truth and the blended result. We then record the average per-vertex Euclidean distance over all vertices and per part:

$$D_{GE_{\text{whole}}} (S') = \frac{1}{n_{\text{det}}} \sum_{d=1}^{5} \frac{1}{n_d} \sum_{V_j \in S_d} \left| V_j - \left( R_d V_j' + T_d \right) \right|^2$$

$$D_{GE_{\text{part}}} (S') = \frac{1}{n_{\text{det}}} \sum_{d=1}^{5} \frac{1}{n_d} \sum_{V_j \in S_d} \left| V_j - \left( R_d V_j' + T_d \right) \right|^2$$

$$D_{GE} (S) = \frac{D_{GE_{\text{part}}} (S') + D_{GE_{\text{whole}}} (S')}{2}$$

where $n_{\text{det}}$ is the number of vertices of the face mesh, $n_{\text{det}}^{\text{part}}$ is the number of vertices for part $d$, $V_j$ is on the ground truth and $V_j'$ is the corresponding point on the blended face. We compute averages over all vertices and per part to ensure that parts with more vertices do not end up using most of the eigenvector budget at the expense of parts with fewer vertices. We do so for the entire data set and for a set of 19 validation faces that were not part of the training data set. We compute the median data set error as well as the median per part error, and we average the two in a global error.

At each step of our incremental eigenvector selection, we decide which of the five parts will get a new eigenvector added to its set. We compare the geometric errors resulting from each of the five candidate eigenvectors, and we select a candidate eigenvector which has a great impact on decreasing the error. When eigenvectors from multiple parts result in similar decreases in error, instead of systematically picking the eigenvector based on lowest error, we select by sampling from a discrete probability density function (PDF) created from the respective decreases in error of the five candidate eigenvectors. This PDF selection process creates a more even distribution of eigenvectors across the parts and it maintains a low error. As we iterate, the reconstruction error decreases. For the female and male data set faces, the average reconstruction errors are 2.00 and 2.13 mm when considering zero eigenvector. The errors decrease to 0.75 and 0.74 mm after 80 iterations, and when considering all eigenvectors the errors are 0 mm. We chose an error threshold of
1 mm which balances out the cost associated with considering too many eigenvectors and the accuracy of the reconstruction. Table 1 shows the resulting eigenvector distribution after achieving our 1 mm reconstruction accuracy.

Table 1: Number of selected eigenvectors for each part.

<table>
<thead>
<tr>
<th>Nb. of eigenvectors</th>
<th>46 steps for female</th>
<th>46 steps for male</th>
</tr>
</thead>
<tbody>
<tr>
<td>Facial mask</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>Eye</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>Nose</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Mouth</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>Ear</td>
<td>14</td>
<td>16</td>
</tr>
</tbody>
</table>

5. Face Generation through Parts Blending

This step focuses on the problem of constructing a realistic new face through blending the five segmented parts together. As opposed to methods such as those of Tena et al. [TDM11] and Cao et al. [CCWL18] which handle the transition between the parts by relying on the adjustment of a single strip of vertices, we diffuse the transition across three strips of vertices. In contrast to other methods that adjust the transition by vertex averaging [CCWL18] or least-squares fitting [TDM11], we use Laplacian blending [SCL*04] of the parts and the transition, resulting in a smooth yet faithful global surface. The vertex positions are solved by an energy minimization which reduces the surface curvature discontinuities at the junction between the parts while maintaining the desired surface curvature. To this end, we define a transition zone made of quadrilateral strips around the parts. In our experiments, we interpolated a band of two quadrilaterals (three rings of green vertices in Fig. 2) to provide good results. We interpolate the Laplacian \( \mathcal{L} \) (the cotangent weights) of the five facial parts weighted by \( \beta_d \), which has values of \( \beta_d = 1 \) inside the part, \( \beta_d \in \{0.75, 0.5, 0.25\} \) going outward of the part in the transition zone, \( \beta_d = 0 \) elsewhere. We normalize these weights so that they sum to one for each vertex. These soft constraints allow for some leeway in the transition zone. The boundary conditions of our system are set to the ring of blue vertices in Fig. 2, and we solve for the remaining vertices. To this end, we minimize the following energy function:

\[
E(V') = \sum_{i \in \text{inner}} \left\| T_{i} \mathcal{L}(V'_i) - \frac{\sum_{d=1}^{3} \beta_{i,d} R_d \mathcal{L}(V_i,d)}{\sum_{d=1}^{3} \beta_{i,d}} \right\|^2,
\]

where “inner” is the set of vertices of the five parts excluding the vertices of the boundary conditions. \( T_i \) is an appropriate transformation for vertex \( V'_i \) based on the eventual new configuration of vertices \( V'_i \) and \( R_d \) is the rotation of part \( d \).

We solve Eq. 5 in a similar fashion to ARAP [SA07] by alternating solving for the vertex position and rotation matrices until the change is small. Fig. 3 shows that the rotation quickly converges as the Frobenius norm of consecutive iterations is large only for the first few iterations. Given our experiments, we decided to stop iterating when the Frobenius norm falls below 0.01 or after 6 iterations. Fig. 4 shows an example of blended face. In this case, the Frobenius norm was below 0.01 after five iterations. Fig. 5 shows the evolution of the geometric error \( D_{GE} \) between the ground truth parts and their blended counterparts for the example of Fig. 4. As can be seen, the error quickly reaches a plateau as the rotation stabilizes.

6. Significance of Anthropometric Measurements

PCA eigenvectors characterize the space of the data variation but they do not provide a clear intuitive interpretation. In this paper, we mainly focus on constructing linear regression models from data using a set of intuitive facial anthropometric measurements. Facial anthropometric measurements provide a quantitative description by...
means of measurements taken between specific surface landmarks defined with respect to anatomical features. We use 33 parameters listed in Table 2. Each measurement corresponds to either an Euclidean distance or ratios of Euclidean distances between surface landmarks of measurements taken between specific surface landmarks.

### 6.1. Mapping Method

We evaluate the measurements on the facial parts of the data set, where a row of ones is appended to the measurement matrix \( F \). To construct a new facial part based on specific measurements, we use \( \Delta f = \Delta f \) in Eq. 1, as follows:

\[ M_d = B_d F_d^T, \]

where a row of ones is appended to the measurement matrix \( F_d \) for \( y \)-intercepts in the regression and \( F_d \) is the pseudoinverse of \( F_d B_d \) a \( (n_x) \times (n_m + 1) \) matrix containing the corresponding eigenvector of the related facial part. Afterwards, we can solve for a weight vector \( b \) for a set of measurement values:

\[ b = M_d [f_{11} \ldots f_{m_{11}}]^T. \]

Table 2: Anatomical terms and corresponding abbreviations of our selected and discarded measurements.

<table>
<thead>
<tr>
<th>Selected measurements</th>
<th>Abbrev</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nasal Width Root Width</td>
<td>NWRW</td>
<td>[EE12]</td>
</tr>
<tr>
<td>Nasal Width Length of Bridge</td>
<td>NWLB</td>
<td>[EE12]</td>
</tr>
<tr>
<td>Nasal Width Width of Nostril</td>
<td>NWWN</td>
<td>[EE12]</td>
</tr>
<tr>
<td>Nasal Root Width Tip Protrusion</td>
<td>NRTP</td>
<td>[EE12]</td>
</tr>
<tr>
<td>Length of Nasal Bridge Tip Protrusion</td>
<td>NBTP</td>
<td>[EE12]</td>
</tr>
<tr>
<td>Nasal width Tip protrusion</td>
<td>NWTP</td>
<td>[EE12]</td>
</tr>
<tr>
<td>Nasal Root width Length of Bridge</td>
<td>NRLB</td>
<td>[EE12]</td>
</tr>
<tr>
<td>Nasal Root width Width of Nostril</td>
<td>NRRW</td>
<td>[EE12]</td>
</tr>
<tr>
<td>Length of Nasal Bridge Width of Nostril</td>
<td>NBWN</td>
<td>[EE12]</td>
</tr>
<tr>
<td>Width of Nose Tip protrusion</td>
<td>WTNP</td>
<td>[EE12]</td>
</tr>
<tr>
<td>Philtrum Width</td>
<td>PW</td>
<td>[Far94]</td>
</tr>
<tr>
<td>Face Height</td>
<td>FH</td>
<td>[FKF05]</td>
</tr>
<tr>
<td>Orbits Intercanthal Width</td>
<td>OIW</td>
<td>[FKF05]</td>
</tr>
<tr>
<td>Orbits Fissure Length</td>
<td>OFL</td>
<td>[FKF05]</td>
</tr>
<tr>
<td>Orbits Biocural Width</td>
<td>OBW</td>
<td>[FKF05]</td>
</tr>
<tr>
<td>Nose Height</td>
<td>NH</td>
<td>[FKF05]</td>
</tr>
<tr>
<td>Face Width</td>
<td>FW</td>
<td>[LHP*15]</td>
</tr>
<tr>
<td>Bitragion Width</td>
<td>BW</td>
<td>[LHP*15]</td>
</tr>
<tr>
<td>Ear Height</td>
<td>EH</td>
<td>[LHP*15]</td>
</tr>
<tr>
<td>Bigonial Breadth</td>
<td>B</td>
<td>[LHI*06]</td>
</tr>
<tr>
<td>Bigygomic Breadth</td>
<td>BB</td>
<td>[LHI*06]</td>
</tr>
<tr>
<td>Facial Index</td>
<td>F</td>
<td>[RC06]</td>
</tr>
<tr>
<td>Nasal Index</td>
<td>N</td>
<td>[RC06]</td>
</tr>
<tr>
<td>Mouth-Face Width Index</td>
<td>MFW</td>
<td>[RC06]</td>
</tr>
<tr>
<td>Biocular Width-total Face Height Index</td>
<td>BWFH</td>
<td>[RC06]</td>
</tr>
<tr>
<td>Lip Length</td>
<td>LL</td>
<td>[ZLB*10]</td>
</tr>
<tr>
<td>Maximum Frontal Breadth</td>
<td>Max FB</td>
<td>[ZLB*10]</td>
</tr>
<tr>
<td>Interpupillary Distance</td>
<td>ID</td>
<td>[ZLB*10]</td>
</tr>
<tr>
<td>Nose Protrusion</td>
<td>NP</td>
<td>[ZLB*10]</td>
</tr>
<tr>
<td>Nose Length</td>
<td>NL</td>
<td>[ZLB*10]</td>
</tr>
<tr>
<td>Nose Breadth</td>
<td>NB</td>
<td>[ZLB*10]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Discarded measurements</th>
<th>Abbrev</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anatomical term</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eye Fissure Index</td>
<td>EF</td>
<td>[RC06]</td>
</tr>
<tr>
<td>Minimum Frontal Breadth</td>
<td>Min FB</td>
<td>[ZLB*10]</td>
</tr>
</tbody>
</table>

### 6.2. Measurement Selection

We propose a novel technique for automatically detecting the most effective and relevant anthropometric measurements. Some might be redundant with respect to others, some might not make sense for a specific part (e.g., the “Ear Height” might not be relevant for the mouth), and some might even lead to mapping matrices that generate worst results. During our investigations, we discovered that considering more anthropometric measurements does not necessarily lead to a lower reconstruction error. Fig. 6 illustrates that a higher error occurs considering all the measurements in comparison with our selected combination. Fig. 7 shows two odd-looking examples from using all the measurements of nose (a) and facial mask (b). We thus
We then evaluate a geometric error, $D_{GE}$, with the per vertex distance per face. Considering a facial part $F$, on Body Talk [SQH*16], we use Pearson’s correlation coefficient on $F$ to evaluate the relationship between the anthropometric measurements. Considering a facial part $d$, the Pearson’s correlation coefficient $\text{Cor}_{jk}$ for measurements $j$ and $k$ is expressed as:

$$\text{Cor}_{jk} = \frac{\sum_{i=1}^{n} (f_{ij} - \bar{f}_j)(f_{ik} - \bar{f}_k)}{\sqrt{\sum_{i=1}^{n} (f_{ij} - \bar{f}_j)^2} \sqrt{\sum_{i=1}^{n} (f_{ik} - \bar{f}_k)^2}}$$

(11)

where $f_{ij}, f_{ik} \in f_{d,i}$ are measurements of scan $S_i$, $\bar{f}_j$ and $\bar{f}_k$ are the mean values of measurements $j$ and $k$ respectively, and $n$ is the number of scans. The coefficient is a value between $-1$ and $1$ that represents the correlation. When adjusting measurement $k$
Accordingly, we can evaluate the influence of one measurement on the others as well as conditioning on one or more measurements and creating the most likely ratings of the other measurements.

7. Results

Compared to global 3DMM methods that compute one set of eigenvectors for the whole face, our 3DMM computes a set of eigenvectors for each part. This brings one of the advantages of our approach: its ability to locally adjust the faces. We compare our approach to other methods that rely on local 3DMM. We created mapping matrices (Eq. 6) for global 3DMMs, SPLOCS [NVW*13], clustered PCA [TDM11], as well as our part-based 3DMMs, and tested the adjustment of measurements with these models. We used 46 eigenvectors for global 3DMM, SPLOCS, and our part-based 3DMM. For clustered PCA, we first tested using 13 clusters as is reported in the paper, but found that it lead to a non-symmetrical result (Fig. 8b). By checking other clusterings, we selected 12 clusters (Fig. 8a). As clustered PCA does not allow for different number of eigenvectors for each cluster, and to avoid having too few eigenvectors per part, we used 46 eigenvectors for each cluster (selecting the 46 with largest eigenvalues).

To compare our approach and the use of measurements with other methods, we derived a way to use our measurements with SPLOCS and clustered PCA. We furthermore demonstrate that with SPLOCS and our approach, we can have more local or global control. For our approach, Table 3 shows that some measures influence more than one part. For example, the “Lip Length” is found in the lists for both mouth and nose. When a measurement is shared between different facial parts, our method allows to decide to have more localized changes by adjusting the measure for only one part, or have more coherence across the parts by adjusting all of the parts that involve the measurement. When comparing with SPLOCS, we can also balance between locality and globality of the influence of the measurement. Each measurement is based on computations involving specific measurement vertices (such as the corner of the mouth and the tip of the nose). To enforce locality, when considering a measurement, we check which SPLOCS “eigenvectors” infer significant movement at the related measurement vertices. We compute this by checking if the eigenvector displacement vector at a measurement vertex is large enough compared to the maximum displacement vector of the eigenvector (we check if it is larger than 1% of the maximum displacement of all vertices of the eigenvector). A SPLOCS eigenvector is considered for a measurement only if it meets the criterion for one of the measurement vertices of a specific measurement. To enforce more globality with SPLOCS, we use the mapping matrices for all of the eigenvectors. Fig. 9 compares global PCA eigenvectors, local and global SPLOCS, clustered PCA, and our local and global approaches. Note how our strategy for global SPLOCS and our global approach create a useful continuum from the global PCA to our local approach.

We will now focus on local editing. Fig. 10-12 show the adjustment of the same anthropometric measurement using global 3DMM, local SPLOCS, clustered PCA, and our local approach. The color coding shows the per vertex Euclidean distance. Note that the colors do not represent errors, but vertex movement. Thus, the goal is to have warmer colors around the location where the editing is intended, and colder colors in unrelated regions. In Fig. 10f and 10g, we can see that even though we wanted to adjust the “Nose Breadth”, the adjustment using the global eigenvectors and local SPLOCS resulted in significant deformation all over the face while clustered PCA and our approach can focus the deformation around the nose as expected (Figs. 10h and 10i). We can observe similar unwanted global deformation of the face in Fig. 11f and 11g. Also note the automatic segmentation of clustered PCA does not provide desired deformation for some cases such as Figs. 11b and 12b. We consistently outperform clustered PCA in terms of local deformation where expected. The results shown in Fig. 10-12 highlight the difficulty of locally controlling the face deformation, and the power of our approach in locally adjusting the face with respect to the anthropometric measurements.

In the accompanying video, we show multiple edits on multiple parts starting from the average face, while Fig. 13 shows edits starting from four real faces. We can see that our approach allows to capture the essence of the anthropometric measurements, providing an easy to use workflow.

8. Discussion

In this section we discuss different aspects of our approach. We present different comparisons highlighting the impact of the eigenvector and measurement selection. We then discuss the choice of
face segmentation, and we end by describing the procedure used to bring all of our scans to a common face mesh.

8.1. Measurements Error

To verify the robustness of our 3DMMs and of our set of selected measurements, we reconstruct real faces relying on their anthropometric measurements to compute their eigenvector weights (Eq. 7).

Figure 10: Results adjusting “Nose Breadth”: (a) nose of a female from validation faces adjusted using (b) global PCA eigenvectors, (c) SPLOCS, (d) clustered PCA, and (e) our approach. The color mapped renderings (f)-(i) indicate respective per vertex Euclidean distance (blue = 0 mm, red = 5 mm).

Figure 11: Results increasing “Lip Length”: (a) mouth of a male from validation faces edited using (b) global PCA eigenvectors, (c) SPLOCS, (d) clustered PCA, and (e) our approach. The color mapped renderings (f)-(i) indicate respective per vertex Euclidean distance (blue = 0 mm, red = 8 mm).

Figure 12: Results increasing “Bizygomatic Breadth” (the bizygomatic width of the face): (a) bizygomatic of a male from validation faces edited using (b) global PCA eigenvectors, (c) SPLOCS, (d) clustered PCA, and (e) our approach. The color mapped renderings (f)-(i) indicate respective per vertex Euclidean distance (blue = 0 mm, red = 14 mm).

Figure 13: We generated random faces (left faces (a)-(d)) and edited them by increasing (“+”) or decreasing (“-”) the value of some of the indicated anthropometric measurement.

We then get the face with our approach, including the blending procedure (Sec. 5) and compute its resulting anthropometric measurements. We compute the quality of the reconstruction through the absolute value of the difference between the ground truth measurement and the measurement from the reconstructed face. Since measurements correspond to either an Euclidean distance or ratios of Euclidean distances, we normalized all the measurements to the range [0%, 100%]. Fig. 14 shows that the average percentage of error is low when using “our measurements”. This means that both the selection of eigenvectors and the mapping matrix work

...
8.2. Face Decomposition

Our face segmentation was motivated by several facial animation artists with whom we worked that strongly prefer that they have control over the face patches in order to make sure they match the morphology of the face and muscle locations. This type of control is impossible to achieve with an automatic method that is typically agnostic to the underlying anatomical structure. It is important to note that this manual way of selecting the regions is not more cumbersome than the current state of the art methods. According to the authors and confirmed by our own experiments, the state of the art method of Tena et al. [TDM11] requires a post-processing step to fix occasional artifacts in their segmentation method. Furthermore, as illustrated in Fig. 8b, the segmentation boundaries can occasionally occur across important semantic regions such as the eyes, leading to complications further down the pipeline.

8.3. Data Set

The quality of the input mesh data set is important to the reconstruction of good 3D face models. As many existing methods, we assume that the meshes share a common mesh topology. Mapping the raw 3D scans to a common base mesh is typically done by a surface mapping method [RGdL*18; ZB13; ARV07]. We established this correspondence with the commercial solution R3DS WRAP.

9. Conclusion

In this paper, we designed a new local 3DMM used for face editing. We demonstrated the difficulty to locally edit the faces with global 3DMMs: we thus segmented the face in five parts and combined the 3DMMs for each part into a single 3DMM by selecting the best eigenvectors through prediction error measurements. We then posed the use of established anthropometric measurements as a basis for the face editing. We mapped the anthropometric measurements to the 3DMM through a mapping matrix. We proposed a process to select the best set of anthropometric measurements, leading to an improve reconstruction accuracy and the removal of conflicting measurements. From the list of 33 anthropometric measurements that we surveyed from the literature, we identified 31 which lead to an improvement of the reconstruction and we rejected 2 as they decreased the quality of the reconstruction. Note that the anthropometric measurement selection process would apply as well even if using a different 3DMM than the one proposed in this paper, as well as if considering a different set of anthropometric measurements. We demonstrated this by applying our set of measurement to both SPLOCS [NVW*13] and clustered PCA [TDM11]. This also demonstrated that our approach produces results superior to those of established methods proposing automatic segmentation and different ways to construct the eigenvector basis. We also presented

---

Figure 14: Using our subset of measurements on the data set and validation faces leads to lower errors (percentage) compared to using “all measurements”.

Figure 15: Starting from one face, we adjust one of its measurements to match the value of that measurement for another face. We then compute the difference between the prescribed measurement value and the measurement value calculated from the mesh. We do so for 1,000 such edits. Our approach leads to smaller error (percentage) compared to cluster PCA and slightly better results when compared to local SPLOCS.

Figure 16: (a) and (b) show that we sometimes get odd looking results when editing just the region of the cheeks or just region of the chin.
different experimental evidence to show the superiority of our approach, especially in terms of local control, compared to the typical global 3DMM.

A limitation of our approach is the mapping matrices assuming a linear relationship between anthropometric measurements and the eigenvector weights. An interesting avenue for future work would be to apply machine learning to identify non-linear mappings. Our set of anthropometric measurements contains too few measures for the ears due to the scarcity of measurements within the ear compared to the nose. It would be interesting to identify more anthropometric measurements for ears, as well as, considering the measurements that specify the distribution of curvature over the face, such as the measurement specifying the angle formed at the tip of the chin. Another limitation comes from the blending of the different parts of the skull and the measurements. Another avenue for future research based on an energy minimization of the skin thickness considering allow as much deformation for the shape of the head. An additional measurement specifying the angle formed at the tip of the chin.

to apply machine learning to identify non-linear mappings. Our set approach, especially in terms of local control, compared to the typical

References


[CGZ17] CHI, JING, GAO, SHANSHAN, and ZHANG, CAIMING. “Interactive facial expression editing based on spatio-temporal coherency”. The Visual Computer 33.6 (June 2017), 981–991 2.


© 2020 The Author(s)

Computer Graphics Forum © 2020 The Eurographics Association and John Wiley & Sons Ltd.