3D Face Synthesis Based on the Information of Words Expressing Facial Features

Futoshi Sugimoto
Department of Information and Computer Sciences
Toyo University
Feelings and sensitivity are indispensable elements in establishing a human-centered system

Information expressed by the human face plays a pivotal role in human communication

Various kinds of research have recently been conducted that deal with various types of sensitivity information about face

We focused on the synthesis of individual faces based on verbal expressions with a high degree of freedom
Goal

- In the past studies on the synthesis of a face based on word information, the words are very restrictive
  - words expressing physical quantities, such as the height of nose or the size of eye
  - others expressing their degree, such as slight, a little, and very

- Our research attempts to synthesize a face based on abstract, sentimental and metaphoric expressions and descriptions
  - gimlet eye, piggy nose, and haggard face
Outline of the facial synthesis system

- Text expressing facial features
- Collection of feature words
  - Word space
    - Nose, Mouth, Cheeks
  - Mapping
  - Physical model space
    - Nose, Mouth, Cheeks
- Combining
## Facial elements and Feature words

<table>
<thead>
<tr>
<th>Facial element</th>
<th>The number of words</th>
<th>Examples of feature words</th>
</tr>
</thead>
<tbody>
<tr>
<td>mouth</td>
<td>27</td>
<td>big mouth, thin mouth</td>
</tr>
<tr>
<td>nose</td>
<td>52</td>
<td>long nose, plump nose, snub nose</td>
</tr>
<tr>
<td>eyes</td>
<td>23</td>
<td>eyes slanting upward, puffy eyes</td>
</tr>
<tr>
<td>eye-brows</td>
<td>26</td>
<td>thin eye-brows, thick eye-brows</td>
</tr>
<tr>
<td>cheeks</td>
<td>16</td>
<td>plump cheeks, haggard cheeks</td>
</tr>
<tr>
<td>jaw</td>
<td>25</td>
<td>sharp jaw, protruding jaw</td>
</tr>
<tr>
<td>profile</td>
<td>23</td>
<td>egg-shaped face, firm face</td>
</tr>
</tbody>
</table>
A similarity matrix among the feature words is obtained based on the similarity of the shape recalled by the words.

A spatial layout of the feature words is obtained by inputting the similarity matrix into Multi-Dimensional Scaling method (MDS).

This spatial layout of the feature words is the word space.

The word space is made for each facial element.

Word space
Characteristics of word space

- The origin is neutral
- The farther away a feature word is from the origin, the greater the characteristic of the feature word
- Similar words are arranged close together
- Dissimilar words are arranged further away from each other
- Every word space of the feature words has six dimensions
Word space of nose

- plump nose
- fat nose
- big nose
- crooked nose
- high nose
- downward nose
- long nose
- upward nose
- thin nose
- small nose
- short nose
- low nose
Physical model space

- Concrete shapes of the facial elements are determined by wire frame model on 3-dimensional computer graphics.
- A set of xyz coordinates of all the apexes in the wire frame model becomes the parameters of the physical model space.
- The number of parameters is different from each facial element.
- The physical model space is made for each facial element.
A feature word $W_i$ of a facial element is described as follows:

$$W_i = (w_1, w_2, ..., w_6) \quad (1)$$

$i = 1, ..., \text{the number of feature words}$

A physical model $M_i$ of the same facial element is described as follows:

$$M_i = (P_{i1}, P_{i2}, ..., P_{in}) \quad (2)$$

$$P_{ij} = (x_{ij}, y_{ij}, z_{ij}), j = 1, 2, ..., n \quad (3)$$

The mapping function is described as follows:

$$M_i = f(W_i) \quad (4)$$

$$P_{ij} = f_j(W_i) \quad (5)$$

$$x_{ij} = f_{xj}(W_i), y_{ij} = f_{yj}(W_i), z_{ij} = f_{zj}(W_i) \quad (6)$$
## Training data of nose

<table>
<thead>
<tr>
<th>feature word</th>
<th>physical model</th>
<th>feature word</th>
<th>physical model</th>
</tr>
</thead>
<tbody>
<tr>
<td>thick nose</td>
<td><img src="image" alt="thick nose" /></td>
<td>thin nose</td>
<td><img src="image" alt="thin nose" /></td>
</tr>
<tr>
<td>big nose</td>
<td><img src="image" alt="big nose" /></td>
<td>upward nose</td>
<td><img src="image" alt="upward nose" /></td>
</tr>
<tr>
<td>crooked nose</td>
<td><img src="image" alt="crooked nose" /></td>
<td>short nose</td>
<td><img src="image" alt="short nose" /></td>
</tr>
<tr>
<td>downward nose</td>
<td><img src="image" alt="downward nose" /></td>
<td>small nose</td>
<td><img src="image" alt="small nose" /></td>
</tr>
<tr>
<td>high nose</td>
<td><img src="image" alt="high nose" /></td>
<td>low nose</td>
<td><img src="image" alt="low nose" /></td>
</tr>
<tr>
<td>long nose</td>
<td><img src="image" alt="long nose" /></td>
<td>plump nose</td>
<td><img src="image" alt="plump nose" /></td>
</tr>
</tbody>
</table>
Calculation to compound the feature words
Individual model 1

- upward-slanting eyes
- a thin nose
- an oblong mouth
- sink cheeks
- a sharp jaw
Individual model 2

very thin eyes
a straight lined nose
a little edge upped mouth
a little chubby cheeks
a sturdy jaw
He has very downward-slanting eyes and eyebrows, a little big nose, and chubby cheeks. His jaw is round.

He has very thin eyes, a straight lined nose, and a little edge upped mouth. His jaw is sturdy and cheeks a little chubby.

He has upward-slanting eyes, a thin nose, and an oblong mouth. His jaw is sharp and cheeks sink.
Experiment for evaluation (2)
Experimental results

Table 1

<table>
<thead>
<tr>
<th>Synthesized face ID</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>Rate of Correct Answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a)</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>17</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0.773</td>
</tr>
<tr>
<td>(b)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>21</td>
<td>0.955</td>
</tr>
<tr>
<td>(c)</td>
<td>4</td>
<td>0</td>
<td>17</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0.773</td>
</tr>
</tbody>
</table>

Table 2

<table>
<thead>
<tr>
<th>Synthesized face ID</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>Rate of Correct Answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>15</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>4</td>
<td>0.682</td>
</tr>
<tr>
<td>(b)</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>17</td>
<td>0</td>
<td>0.773</td>
</tr>
<tr>
<td>(c)</td>
<td>1</td>
<td>2</td>
<td>18</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.818</td>
</tr>
</tbody>
</table>
Future works

- Collect feature words about hair style and texture of skin, and make their physical models
- Make more elaborate physical models using the face database of 300 people
- Cope with differences in gender and age by changing the standard facial model