Expressions are Expressive

- Facial expression is one of primary means by which we communicate with each other.
- Nowadays, more and more people are “talking” without seeing each other, so…
  
  :-) , :-( , :-8, ^_^ , * ^ ,
  😁😊😃😄😆❤

- The popularity of these representations has proven the great needs for facial expressions in numerous applications.
- No representations are more nature and friendly than the facial images of the real human.
Main Challenges

- Acquisition: It costs user too much effort to take photos for every expression.
- Transference: Images are too big to transfer.
- Display: Screen is too small, and Images are too large.

Our Goal

- A solution to synthesize realistic facial expression images from photographs for the portable devices such as cell phones and PDAs, which have
  - Limited processing power
  - Limited network bandwidth
  - Limited display area

Referred as----

the “LLL” environment
Overview of our approach

- Face Alive Icons (FAI) are small size realistic facial images
- Facial Icon Profile (FIP) is compact in size;
- Expression code (ECode) is just one byte.
- Expression synthesis is a light load operation.

<table>
<thead>
<tr>
<th>Step</th>
<th>Host</th>
<th>Input</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1: Expression Decomposition</td>
<td>Server</td>
<td>Photograph</td>
<td>Facial Icon Profile</td>
</tr>
<tr>
<td>Step 2: Icon Synthesis</td>
<td>Terminal Device</td>
<td>Facial Icon Profile</td>
<td>Face Alive Icons</td>
</tr>
</tbody>
</table>

Expression Decomposition

- Observations from Japanese Female Facial Expression (JAFFE) database (216 images, 10 persons):
  - Inactive and active parts
  - Parts share the same appearances across expressions
- Decomposition rule FAI:=FF+EF
  - FF: Facial Features, the inactive parts of face in expressions
  - EF: Expressional Features, the active parts of face in expressions
- FAI:=FF+<left eye>+<right eye>+<mouth>
Discrete Model of EF

Built through Principle Component Analysis (PCA) – Distribution is investigated and **standard states** are defined. Assuming \( n \) sample data items in training set and \( m \) variables for landmark points of each item. The \( i \)th data item \( X_i \)

\[
X_i = (x_{i,1}, x_{i,2}, \ldots, x_{i,m})
\]

Where \( x_{i,k} \) could be either the coordinates or grayscales of the landmark points

\[
X_i = \bar{X} + P \cdot b
\]

In which,
- \( \bar{X} \) is the average of the training samples
- \( P \) is the matrix of unit eigenvectors of the covariance of deviation.
- \( b \) is a vector of eigenvector weights referred as **Model Parameters**.

Discrete Model of EF

- Categorized by expressions, the vector \( b \) of features for \( e_i \) can be represented by the averages

\[
S = \{ \bar{b}_{e_1}, \bar{b}_{e_2}, \ldots, \bar{b}_{e_p} \}
\]

- Assuming uniform distribution,

\[
delta = \max \frac{|b_i - b_j|}{(n-1)}
\]

- Keep merging the closest items \( b_i \) and \( b_j \) if \( |b_i - b_j| < \delta \)
- A unique semantic name is given for each of the standard states according their appearances.

Figure 1: The Eye Area with 18 Landmark Points

Figure 2: States of the right eye: (a) \( b_{e_1} \) normal (b) \( b_{e_2} \) up (c) \( b_{e_3} \) wide-open (d) \( b_{e_4} \) down

Figure 3: The States of the mouth: (a) \( b_{e_5} \) normal (b) \( b_{e_6} \) down-close (c) \( b_{e_7} \) up-open (d) \( b_{e_8} \) down-open
FAI Synthesis

- Synthesis rules are determined by statistical analysis on the distance from the standard states.
- Face Icon Profile (FIP) includes only one facial image and several standard states of expressional features.

<table>
<thead>
<tr>
<th></th>
<th>left eye</th>
<th>right eye</th>
<th>mouth</th>
</tr>
</thead>
<tbody>
<tr>
<td>normal</td>
<td>2.37</td>
<td>2.34</td>
<td>4.37</td>
</tr>
<tr>
<td>up</td>
<td>3.06</td>
<td>2.28</td>
<td>3.56</td>
</tr>
<tr>
<td>down</td>
<td>2.29</td>
<td>2.27</td>
<td>2.98</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>left eye</th>
<th>right eye</th>
<th>mouth</th>
</tr>
</thead>
<tbody>
<tr>
<td>normal</td>
<td>0.22</td>
<td>0.24</td>
<td>0.24</td>
</tr>
<tr>
<td>up</td>
<td>0.33</td>
<td>0.35</td>
<td>0.35</td>
</tr>
<tr>
<td>down</td>
<td>2.06</td>
<td>2.06</td>
<td>2.06</td>
</tr>
</tbody>
</table>

Face Icon Profile (FIP) includes only one facial image and several standard states of expressional features.

Prototype System

- An experimental prototype has been implemented.
  - decomposition process: simulated on PC (Window XP)
  - synthesis process: simulated on Palm Zire (Palm OS)
- More than 30 meaningful expressions can be generated.
- Suppose
  - 20 buddies on your list
  - 30 expressions

Our approach
- FIP size = 20*128K = 2.56 Mbytes
- Approaches which transfer Images in JPEG files:
  - Total size = 20 * 30 * 30K = 18 Mbytes

The average distance / distance between training data and the standard states.

Approaches which transfer Images in JPEG files:
- Total size = 20 * 30 * 30K = 18 Mbytes
Contribution

- Propose a system to synthesize realistic facial expressions from photographs for portable devices in the “LLL” environment.

- Reveal the inner relations among expressions by the semantic synthesis rules assessed through the statistical analysis on training data.

Acknowledgement

- This work is advised by Prof. Shi-Kuo Chang.

- We would like to thank Dr. Lyons who kindly provides the JAFFE database, ChiehChih Chang and Jui-Hsin Huang for his work on implementation of the experimental prototype system, and Dr. Carl Kuo for the valuable comments.

- The work is pending to the patents in Taiwan and US.