Ubiquitous and Mobile Computing
CS 403x: Visage

Ross Foley, Chris Hanna, Dan True

Computer Science Dept.
Worcester Polytechnic Institute (WPI)
What is Visage?

- Real-time Face Interpretation Engine
- Input to Apps
  - 3D Head Poses
  - Mood Interpretation
Motivation

- Front-Facing Camera Enables New Possibilities
- Search for New Solution for Onboard Facial Recognition
  - Resource Aware
  - Mobile Camera
Vision

● Onboard System
  - Without need for backend server

● Resource-limited Mobile Devices
Street View +

- Makes use of head pose inference
- Provide user with navigation on-the-go
Mood Profiler

- Senses users’ expressions
- Visualized summaries while users are interacting with specific applications

(a) YouTube

(b) Email
Related Work

- Sense Cam
- MoVi
- Recognizer
- PEYE
Methodology

- Preprocessing Stage
- Tracking Stage
- Inference Stage
Methodology

- Preprocessing Stage
- Tracking Stage
- Inference Stage
Preprocessing Stage

Phone Posture Component

- Raw data from sensors
- Estimates direction of gravity
Preprocessing Stage

*Face Detection with Tilt Compensation*

- Locates face
- Normalizes the face angle

http://eclecti.cc/files/2008/03/face.jpg
Preprocessing Stage

Adaptive Exposure Component
Methodology

- Preprocessing Stage
- Tracking Stage
- Inference Stage
Tracking Stage

*Feature Point Tracking*

- Searches for Facial Feature Points
  - Looks for lips and eyes
  - Multiple frames of video required
  - Uses spatial relationships for further calculation

Tracking Stage

Pose Estimation

- Estimates 3D pose of head
- Models head as cylinder
- Compensates for error
Tracking Stage

Pose Estimation
Methodology

- Preprocessing Stage
- Tracking Stage
- Inference Stage
Inference Stage

Active Appearance Model

- Machine Learning Algorithm
- Generates Triangular Mesh over Image
Inference Stage

Expression Classification

- Support Vector Machine Classifier
  - Angry, Disgust, Fear
  - Happy, Neutral, Sad
  - Surprise
Results

- Pose Estimation
- Expression Estimation
- Computation Expenses
Results

Pose Estimation

- Head in constant position
- Phone rotated
Results

Pose Estimation

- Head rotated
- Phone constant
Results

*Expression Estimation*

- 5 Volunteers make facial expressions
- Visage categorized resulting expressions

<table>
<thead>
<tr>
<th>Expressions</th>
<th>Anger</th>
<th>Disgust</th>
<th>Fear</th>
<th>Happy</th>
<th>Neutral</th>
<th>Sadness</th>
<th>Surprise</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy(%)</td>
<td>82.16</td>
<td>79.68</td>
<td>83.57</td>
<td>90.30</td>
<td>89.93</td>
<td>73.24</td>
<td>87.52</td>
</tr>
</tbody>
</table>

- Overall accuracy = 83.78%
## Results

### Expression Estimation

<table>
<thead>
<tr>
<th>Expressions</th>
<th>Anger</th>
<th>Disgust</th>
<th>Fear</th>
<th>Happy</th>
<th>Neutral</th>
<th>Sadness</th>
<th>Surprise</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anger</td>
<td>93.33</td>
<td>6.67</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Disgust</td>
<td>6.90</td>
<td>75.86</td>
<td>17.24</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Fear</td>
<td>0</td>
<td>7.41</td>
<td>92.54</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3.23</td>
</tr>
<tr>
<td>Happy</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>87.10</td>
<td>6.45</td>
<td>3.23</td>
<td>0</td>
</tr>
<tr>
<td>Neutral</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>90.00</td>
<td>10.00</td>
<td>0</td>
</tr>
<tr>
<td>Sadness</td>
<td>0</td>
<td>6.45</td>
<td>9.68</td>
<td>3.23</td>
<td>9.68</td>
<td>70.97</td>
<td>0</td>
</tr>
<tr>
<td>Surprise</td>
<td>0</td>
<td>0</td>
<td>3.33</td>
<td>3.33</td>
<td>0</td>
<td>0</td>
<td>93.33</td>
</tr>
</tbody>
</table>
## Results

*Computation Estimation*

<table>
<thead>
<tr>
<th>Tasks</th>
<th>Avg. CPU usage</th>
<th>Avg. memory usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>GUI only</td>
<td>&lt; 1%</td>
<td>3.18MB</td>
</tr>
<tr>
<td>Pose estimation</td>
<td>58%</td>
<td>6.07MB</td>
</tr>
<tr>
<td>Expression inference</td>
<td>29%</td>
<td>4.57MB</td>
</tr>
<tr>
<td>Pose estimation &amp; expression inference</td>
<td>68%</td>
<td>6.28MB</td>
</tr>
</tbody>
</table>
Conclusions

- Succeeded in creating an onboard facial recognition platform
  - Comparable to traditional cloud-based image analysis systems

- Could be Useful?
Questions ?
References

- The Visage Face Interpretation Engine for Mobile Phone Applications Xiaochao Yang, Chuang-Wen You, Andrew Campbell, in Proc MobiCase 2012
Insert your Title here

- Insert your stuff here