Multimodal affective human-machine interaction
Towards the design and implementation of an affect-sensitive empathetic agent

Tessa Verhoef

January 15, 2009
Master Thesis Colloquium
Overview

• General information

• Scope of the project

• Background/Design considerations

• Affect recognition

• Affect generation

• Discussion/Conclusion
General information

• Dr. Fokie Cnossen & Drs. Tijn van der Zant

• Dr. Christine Lisetti

• Florida International University, Miami

• Affective Social Computing Group
Scope

• Humans and machines living together
  • E-learning
  • Therapy
  • Elderly care
  • Social Companion
  • Service / Household

• Natural and effective human-machine interaction necessary
Emotions in Artificial Intelligence

- Affective Computing *(Picard, 1997)*

- Emotions versus rational thinking
  - *(Bechara et al., 1997)*

- Human-human communication
Human-machine interaction

“People naturally express emotion to machines, but machines do not naturally recognize it” (Picard, 1997)

Office Space (1999)
Background/Design considerations

- Important questions:
  - What is an affective state?
  - What kinds of evidence warrant conclusions about affective states?
  - How can emotion be elicited?
What is an affective state?

- Goal: finding a structured representation
- No consensus among emotion theorists
- Different issues on which they disagree
  - Most important: Discrete versus continuous
    - Discrete: set of basic emotion categories
    - Continuous: dimensional space
Discrete emotions

• Paul Ekman (1992)

• Finding universal facial expression

• Six basic emotions: Angry, Happy, Sad, Surprise, Disgust, Fear

• Creating a taxonomy of facial expressions: Action Units (Ekman and Friesen, 1978)

<table>
<thead>
<tr>
<th>AU 1</th>
<th>AU 2</th>
<th>AU 4</th>
<th>AU 5</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Image of AU 1" /></td>
<td><img src="image2.png" alt="Image of AU 2" /></td>
<td><img src="image3.png" alt="Image of AU 4" /></td>
<td><img src="image4.png" alt="Image of AU 5" /></td>
</tr>
</tbody>
</table>
Continuous emotion space

- Two-dimensional Valence-Arousal space

- Appraisal theory *(Scherer, 2001)*

- Sequential Evaluation Checks

- Suitable for computational purposes

---

**Scope**

**Background**

**Affect recognition**

**Affect generation**

**Conclusion**
What kinds of evidence warrant conclusions about affective states?

- Facial expressions
- Vocal expressions
- Physiological signals
- Pupil diameter (Barreto et al., 2007)
- Thermal imaging (Puri et al., 2005)
- Body posture (Kapoor & Picard, 2005)

- Limitations to all > multimodal fusion
How can emotion be elicited?

- Posed versus spontaneous

- Actors
  - easy labelling, but exaggerated (not real)

- Expose to stimuli
  - more difficult to label, but genuine

- Pictures (IAPS), Movies, Music, Games/puzzles
Present research

- Two building blocks
  - Affect recognition
  - Affect generation
Affect recognition

- Data Collection
  - Emotion elicitation
  - Signal measuring from multiple modalities
  - Feature extraction
- Training and testing classifiers
Data collection

- Elicitation with movies
- Movie clips from data base
- Small user study

- Six clips:

<table>
<thead>
<tr>
<th>Movie</th>
<th>Affect</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Champ</td>
<td>Sadness</td>
</tr>
<tr>
<td>Schindler’s List</td>
<td>Anger</td>
</tr>
<tr>
<td>Fear Factor</td>
<td>Disgust</td>
</tr>
<tr>
<td>Pixar - Boundin’</td>
<td>Happiness</td>
</tr>
<tr>
<td>The Ring</td>
<td>Fear</td>
</tr>
<tr>
<td>Capricorn one</td>
<td>Surprise</td>
</tr>
</tbody>
</table>
Signal measuring

• Modalities:
  • Facial expressions
  • Galvanic Skin Response
  • Blood Volume Pressure

• Features

Scope | Background | Affect recognition | Affect generation | Conclusion
Procedure

- 25 test subjects, age 21 to 41
- 16 male, 9 female
- Questionnaire: self-report in multiple choice and emotion wheel
- Excluding data

Adapted from Scherer (2005)
Classifier training and testing

• Classifiers for discrete categorization of 5 emotions
  • Static Bayesian Network
  • Dynamic Bayesian Network
  • k-Nearest Neighbours
• Alternative method proposed by Peter & Herbon (2006)
  • Abandons the use of discrete categorization
• Decision level fusion with FaceReader decision
Alternative method (Peter & Herbon, 2006)

• Four steps:
  • Using a self-report method with dimensional structure
  • Assigning physiological measurements to ratings
  • Cluster emotions based on physiology and ratings.
  • Identify characteristic patterns in physiology for each cluster.
Results

• FaceReader: 39.6 % no face found, 71.7 % recognized as Anger

• Results without Facial Data:

<table>
<thead>
<tr>
<th>Performance</th>
<th>SBN</th>
<th>DBN</th>
<th>kNN</th>
<th>PH</th>
</tr>
</thead>
<tbody>
<tr>
<td>43.5 %</td>
<td>31.9 %</td>
<td>44.4 %</td>
<td>60.0 %</td>
<td></td>
</tr>
</tbody>
</table>

SBN = Static Bayes Network, DBN = Dynamic Bayesian Network, kNN = k-Nearest Neighbours, PH = Alternative Peter & Herbon method
Discussion

• No facial data, losing one important modality
  • Enough discriminative ability in other modalities?

• Humans are no perfect affect recognizers either

• Small data set

• Other reasons

• A lot can be improved
Affect generation

- Providing feedback to the users
- Anthropomorphic avatar (Haptek Inc)
- Psychologically grounded facial expressions
- Small user study to test recognizability and believability
Psychologically grounded facial expressions

- Scherer’s Sequential Evaluation Checks

- Translation into Action Units

- 12 emotions

**Appraisal sequence**

<table>
<thead>
<tr>
<th>Event</th>
<th>Relevance</th>
<th>Implication</th>
<th>Coping</th>
<th>Normative Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effect of each check</td>
<td>1+2</td>
<td>4</td>
<td>17+23</td>
<td>10+14</td>
</tr>
<tr>
<td>Cumulative effect</td>
<td>1+2</td>
<td>1+2+4</td>
<td>1+2+4+17+23</td>
<td>1+2+4+17+23+10+14</td>
</tr>
</tbody>
</table>
Avatar expressions

Happiness  Disgust  Pride  Despair

Fear  Interest  Sadness  Contempt

Surprise  Boredom  Anger  Shame
Examples
Examples
Examples
Examples
Examples
User Study

- 22 test subjects, ages 21 to 62
- 11 male, 11 female
- **Multiple choice:** Anger, Happiness, Surprise, Pride, Fear, Disgust, Interest, Contempt, Shame, Boredom, Sadness, Despair, Neutral, No idea
- Believability scale 1 to 5
## Results

<table>
<thead>
<tr>
<th>Emotion</th>
<th>Recognition score</th>
<th>Believability</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Happiness</td>
<td>50.0 %</td>
<td>3.20</td>
<td>1.04</td>
</tr>
<tr>
<td>Disgust</td>
<td>36.4 %</td>
<td>3.25</td>
<td>1.04</td>
</tr>
<tr>
<td>Pride</td>
<td>36.4 %</td>
<td>3.38</td>
<td>1.19</td>
</tr>
<tr>
<td>Fear</td>
<td>40.9 %</td>
<td>2.89</td>
<td>0.78</td>
</tr>
<tr>
<td>Interest</td>
<td>54.5 %</td>
<td>2.92</td>
<td>1.08</td>
</tr>
<tr>
<td>Sadness</td>
<td>54.5 %</td>
<td>3.75</td>
<td>0.97</td>
</tr>
<tr>
<td>Surprise</td>
<td>81.8 %</td>
<td>3.44</td>
<td>0.92</td>
</tr>
<tr>
<td>Boredom</td>
<td>50.0 %</td>
<td>3.00</td>
<td>1.26</td>
</tr>
<tr>
<td>Anger</td>
<td>63.6 %</td>
<td>3.23</td>
<td>1.17</td>
</tr>
<tr>
<td>Shame</td>
<td>13.6 %</td>
<td>3.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Contempt</td>
<td>18.2 %</td>
<td>2.50</td>
<td>0.58</td>
</tr>
<tr>
<td>Despair</td>
<td>22.7 %</td>
<td>2.00</td>
<td>0.71</td>
</tr>
</tbody>
</table>
### Results

<table>
<thead>
<tr>
<th></th>
<th>Ha</th>
<th>Di</th>
<th>Pr</th>
<th>Fe</th>
<th>In</th>
<th>Sa</th>
<th>Su</th>
<th>Bo</th>
<th>An</th>
<th>Sh</th>
<th>Co</th>
<th>De</th>
<th>Ne</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ha</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>18</td>
<td>0</td>
<td>32</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Di</td>
<td>0</td>
<td>36</td>
<td>0</td>
<td>5</td>
<td>14</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>23</td>
<td>0</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>14</td>
</tr>
<tr>
<td>Pr</td>
<td>5</td>
<td>0</td>
<td>36</td>
<td>0</td>
<td>23</td>
<td>0</td>
<td>5</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>0</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>Fe</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>41</td>
<td>9</td>
<td>0</td>
<td>14</td>
<td>4</td>
<td>0</td>
<td>5</td>
<td>14</td>
<td>4</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>In</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>55</td>
<td>0</td>
<td>18</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>18</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sa</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>55</td>
<td>0</td>
<td>18</td>
<td>5</td>
<td>14</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Su</td>
<td>5</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>82</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bo</td>
<td>0</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>5</td>
<td>5</td>
<td>50</td>
<td>0</td>
<td>9</td>
<td>14</td>
<td>0</td>
<td>5</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>An</td>
<td>0</td>
<td>18</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>64</td>
<td>0</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sh</td>
<td>4</td>
<td>0</td>
<td>36</td>
<td>5</td>
<td>23</td>
<td>0</td>
<td>0</td>
<td>14</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Co</td>
<td>0</td>
<td>18</td>
<td>0</td>
<td>9</td>
<td>14</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>18</td>
<td>0</td>
<td>14</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td>De</td>
<td>0</td>
<td>9</td>
<td>0</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>23</td>
<td>4</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>

---

**Scope**

**Background**

**Affect recognition**

**Affect generation**

**Conclusion**
Discussion

- One out of 5 > 94% correct, one out of 12 > 44% correct
- More choices
- More similar emotions
- Missing context, missing gestures
- Timing information missing in theory
Overall conclusion

- State of the Art
- A lot to be improved
- Affective Computing is a young field
- But growing...
Questions?


References (II/II)


