Affective Computing
Overview of Theory, Techniques and Applications

http://www.cs.unibo.it/difelice/

Context Aware Systems

Prof. Marco Di Felice
Department of Computer Science and Engineering
University of Bologna
Affective Computing

Affective computing is an emerging field of research that aims to enable intelligent systems to recognize, feel, infer and interpret human emotions (emotion-aware computing).

- Interdisciplinary research area: computer science (computer vision, IA, NLP, ..), psychology, cognitive science, social science.
- Include sentiment analysis (not covered in this presentation) and emotion recognition (focus on mobile applications here).
Driver state monitoring
- Monitor levels of driver fatigue and distraction to enable appropriate alerts and interventions that correct dangerous driving.
- Monitor driver anger to enable interventions or route alternatives that avoid road rage.
- Address handoff challenge between driver and car in semi-autonomous vehicles.

Occupant Experience Monitoring
- Personalize content recommendations
- Adapt environmental conditions
Affective Computing

USE CASES: RECOMMENDER SYSTEMS (II)

- Emotion-aware recommender systems
- Architecture proposed in [1]
- Architecture proposed in [2]


[2] Yong Feng, Qiana Yin, Zhang Xiao, Han, Yu, Limei Peng, "EARS: Emotion-aware recommender system based on hybrid information fusion", Information Fusion, 46, March 2019, 141-146
Affective Computing

- **USE CASES: SMART EDUCATION (III)**

INSTRUMENTS USED in AFFECTIVE COMPUTING STUDIES [1]

- Questionnaires 26%
- Heart Beat 9%
- Blood oxygen 1%
- Blood pressure 3%
- Skin conductance response 16%
- EEG 6%
- EMG 4%
- Respiration 1%
- Facial 11%
- Others 9%
- Pressure mouse 1%
- fMRI 1%
- Blink 1%
- Audio-visual 2%
- Speech 3%
- Verbalization 6%

Figure 3: Proportion of physiological signals and instruments used in affective computing (AC) studies. EEG, electroencephalography; EMG, electromyography; fMRI, functional magnetic resonance imaging.

**Facial Analysis Software Spots Struggling Students**

A computer can learn to recognize, and respond intelligently to, users’ emotional state.

by Will Knight
Jul 1, 2013

https://www.technologyreview.com/s/516606/facial-analysis-software-spots-struggling-students/

Affective Computing

USE CASES: SOCIAL ROBOTS (IV)

“Octavia is a social humanoid robot with an expressive face and dexterous hands. It's designed to understand how humans perceive, think, and act, using her knowledge to interact naturally with people.”[1]

CREATOR Naval Research Laboratory & Xitome Design
Affective Computing

- How to describe the emotions
- How to model affective data
- Which emotions can be recognized
- How to perform emotion detection
- Research issues
- Demo: Affective Tool

Soujanya Poria, Erik Cambria, Rajiv Bajpai, Amir Hussain, "A review of affective computing: From unimodal analysis to multimodal fusion", Information Fusion (37), 2017
Affective Computing

Terminology

- **Core affect**: “Most elementary accessible affective feelings that need not to be directed at anything” (Russell and Burrett).

- **Emotions**: elicited by something, reactions to something.

- **Moods**: last longer than emotions, are less specific, less intensive, less likely to be triggered by a particular event.

- **Sentiments**: thoughts, opinions, idea based on the feeling about a situation, or a way of thinking about something,
Affective Computing

PROBLEM 1: How to describe the emotions?

- **Categorical** model

  **Six** basic emotions (culturally universal)

  - Facial Action Coding System (FACS) → anatomically based system for describing all observable facial movement for every emotion.

  ![Emotional Expressions](https://www.paulekman.com)
Affective Computing

 alunos

PROBLEM 1: How to describe the emotions?

.Dimension Model

Circumplex Model of Affect (proposed by Russell)

All affective states arise from two fundamental neurophysiological systems:

- **Valence** (pleasure-displeasure), x axis
- **Arousal** (alertness), y axis

Each state is a linear combination of the two basic dimensions with different intensities (degrees).
Affective Computing

- PROBLEM 1: How to describe the emotions?
  - Personality model
    - Big Five model (proposed by Goldberg)
      - Hierarchical model of personality traits.
      - Five broad categories represent personality at the highest level of abstraction.
      - Each dimension summarises a large number of specific personality characteristics.
      - Rating tools available (e.g. IPIP, NEO-FFI)

http://emotiondevelopmentlab.weebly.com/circumplex-model-of-affect.html
Affective Computing

- How to describe the emotions
- How to model affective data
- Which emotions can be recognized
- How to perform emotion detection
- Research issues
- Demo: Affective Tool
Affective Computing

PROBLEM 2: How to model the emotions?

EmotionML (https://www.w3.org/TR/emotionml/#s1.1) Emotion Markup Language, proposed by the W3C working group

- Manual annotation of material involving emotionality, such as annotation of videos, of speech recordings, of faces, of texts, etc
- Automatic recognition of emotions from sensors, including physiological sensors, speech recordings, facial expressions.
- Generation of emotion-related system responses, which may involve reasoning about the emotional implications of events.
Affective Computing

PROBLEM 2: How to model the emotions?

EmotionML (https://www.w3.org/TR/emotionml/#s1.1)

Main elements:
- <emotionml> → root of the XML document
- <emotion> → single emotion annotation
- <category> → Description of an emotion or a related state according to a vocabulary set.
- <dimension> → Description of an emotion or a related state according to an emotion dimension vocabulary.
- <action-tendency> → Description of an emotion or a related state according to an emotion action tendency vocabulary.
**PROBLEM 2: How to model the emotions?**

**EmotionML** ([https://www.w3.org/TR/emotionml/#s1.1](https://www.w3.org/TR/emotionml/#s1.1))

```xml
<emotion dimension-set="http://www.w3.org/TR/emotion-voc/xml#pad-dimensions">
  <dimension name="arousal" value="0.8" confidence="0.9/>
  <dimension name="pleasure" value="0.6" confidence="0.3"/>
</emotion>

<emotion category-set="http://www.w3.org/TR/emotion-voc/xml#big6">
  <category name="surprise" confidence="0.95"/>
</emotion>

<emotion action-tendency-set="http://www.example.com/custom/action/robot.xml">
  <action-tendency name="charge-battery" value="0.9"/>
</emotion>
```
Affective Computing

- How to describe the emotions
- How to model affective data
- Which emotions can be recognized
- How to perform emotion detection
- Research issues
- Demo: Affective Tool
PROBLEM 3: Which emotions can be recognized?

DISTINCT STATES MODEL RECOGNITION

EmotionSense → a mobile sensing framework for recognising user’s emotions based on built-in smartphone sensors [1].

- Infer participants’ emotional states in the Ekman’s basic set
- Emotion recognition based on speech techniques
- Neighbour awareness via Bluetooth scanning
- Position awareness via GPS sampling
- Activity awareness via accelerometer sampling
Affective Computing

PROBLEM 3: Which emotions can be recognized?

DISTINCT STATES MODEL RECOGNITION

Table 1. Emotion clustering

<table>
<thead>
<tr>
<th>Broad emotion</th>
<th>Narrow emotions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Happy</td>
<td>Elation, Interest, Happy</td>
</tr>
<tr>
<td>Sad</td>
<td>Sadness</td>
</tr>
<tr>
<td>Fear</td>
<td>Panic</td>
</tr>
<tr>
<td>Anger</td>
<td>Disgust, Dominant, Hot anger</td>
</tr>
<tr>
<td>Neutral</td>
<td>Neutral normal, Neutral conversation, Neutral distant, Neutral tete, Boredom, Passive</td>
</tr>
</tbody>
</table>

ACCURACY

LOCAL vs REMOTE COMPUTATION
Affective Computing

PROBLEM 3: Which emotions can be recognized?

STRESS CONDITION RECOGNITION

- Automatic recognition of daily stress as a 2-class classification problem (non-stressed vs stressed) [1]
- Inputs: (i) people activities, as detected through their smartphones; (ii) weather conditions; (iii) personality traits (Big Five Model is used).

- Features extracted from call and sms logs and from Bluetooth hits, e.g.: (i) the amount of calls, of sms and of proximity interactions; (ii) the diversity of calls, of sms, and of proximity interactions; and (iii) regularity in user behaviors.
Affective Computing

PROBLEM 3: Which emotions can be recognized?

Table 6: Model Metrics Comparison for Feature Subsets

<table>
<thead>
<tr>
<th>Model</th>
<th>Accuracy</th>
<th>Kappa</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Our Multifactorial Model</td>
<td>72.28</td>
<td>37.52</td>
<td>52.72</td>
<td>83.35</td>
<td>57.89</td>
</tr>
<tr>
<td>Baseline Majority Classifier</td>
<td>63.84</td>
<td>0.00</td>
<td>100.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Weather Only</td>
<td>36.16</td>
<td>0.00</td>
<td>100.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Personality Only</td>
<td>36.16</td>
<td>0.00</td>
<td>100.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Bluetooth + Call + Sms</td>
<td>48.59</td>
<td>6.80</td>
<td>73.80</td>
<td>34.32</td>
<td>50.94</td>
</tr>
<tr>
<td>Personality + Weather</td>
<td>43.55</td>
<td>2.96</td>
<td>81.90</td>
<td>21.83</td>
<td>51.20</td>
</tr>
<tr>
<td>Personality + Bluetooth + Call + Sms</td>
<td>46.40</td>
<td>7.01</td>
<td>83.17</td>
<td>25.57</td>
<td>52.88</td>
</tr>
<tr>
<td>Weather + Bluetooth + Call + Sms</td>
<td>49.60</td>
<td>-5.45</td>
<td>38.45</td>
<td>55.91</td>
<td>35.55</td>
</tr>
</tbody>
</table>

Table 2: List of Basic Features

<table>
<thead>
<tr>
<th>General Phone Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Total Number of Calls (Outgoing+Incoming)</td>
</tr>
<tr>
<td>2. Total Number of Incoming Calls</td>
</tr>
<tr>
<td>3. Total Number of Outgoing Calls</td>
</tr>
<tr>
<td>4. Total Number of Missed Calls</td>
</tr>
<tr>
<td>5. Number of SMS received</td>
</tr>
<tr>
<td>6. Number of SMS sent</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Diversity</th>
</tr>
</thead>
<tbody>
<tr>
<td>7. Number of Unique Contacts Called</td>
</tr>
<tr>
<td>8. Number of Unique Contacts who Called</td>
</tr>
<tr>
<td>9. Number of Unique Contacts Communicated (Incoming+Outgoing)</td>
</tr>
<tr>
<td>10. Number of Unique Contacts Associated with Missed Calls</td>
</tr>
<tr>
<td>11. Entropy of Call Contacts</td>
</tr>
<tr>
<td>12. Call Contacts to Interactions Ratio</td>
</tr>
<tr>
<td>13. Number of Unique Contacts SMS received from</td>
</tr>
<tr>
<td>14. Number of Unique Contacts SMS sent to</td>
</tr>
<tr>
<td>15. Entropy of SMS Contacts</td>
</tr>
<tr>
<td>16. Sms Contacts to Interactions Ratio</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Active Behaviors</th>
</tr>
</thead>
<tbody>
<tr>
<td>17. Percent Call During the Night</td>
</tr>
<tr>
<td>18. Percent Call Initiated</td>
</tr>
<tr>
<td>19. Sms response rate</td>
</tr>
<tr>
<td>20. Sms response latency</td>
</tr>
<tr>
<td>21. Percent SMS Initiated</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Regularity</th>
</tr>
</thead>
<tbody>
<tr>
<td>22. Average Inter-event Time for Calls (time elapsed between two events)</td>
</tr>
<tr>
<td>23. Average Inter-event Time for SMS (time elapsed between two events)</td>
</tr>
<tr>
<td>24. Variance Inter-event Time for Calls (time elapsed between two events)</td>
</tr>
<tr>
<td>25. Variance Inter-event Time for SMS (time elapsed between two events)</td>
</tr>
</tbody>
</table>

SOURCE COMPARISON

Table 1: List of Basic Weather Features

<table>
<thead>
<tr>
<th>Weather Feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Daily Mean Temperature</td>
</tr>
<tr>
<td>2. Daily Mean Humidity</td>
</tr>
<tr>
<td>3. Daily Mean Wind Speed</td>
</tr>
<tr>
<td>4. Daily Mean Visibility</td>
</tr>
<tr>
<td>5. Daily Mean Visibility</td>
</tr>
</tbody>
</table>

Table 3: List of Basic Bluetooth Proximity Features

<table>
<thead>
<tr>
<th>Bluetooth Feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Bluetooth IDs seen for more than k time slots</td>
</tr>
<tr>
<td>2. Times most common Bluetooth ID is seen</td>
</tr>
<tr>
<td>3. Bluetooth ID diversity</td>
</tr>
<tr>
<td>4. Bluetooth ID entropy</td>
</tr>
<tr>
<td>5. Time interval for which a Bluetooth ID is seen</td>
</tr>
</tbody>
</table>

Table 4: List of Basic Personality Features

<table>
<thead>
<tr>
<th>Personality Feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Openness</td>
</tr>
<tr>
<td>2. Conscientiousness</td>
</tr>
<tr>
<td>3. Extraversion</td>
</tr>
<tr>
<td>4. Agreeableness</td>
</tr>
<tr>
<td>5. Neuroticism</td>
</tr>
</tbody>
</table>

Table 5: List of Basic SMS Features

<table>
<thead>
<tr>
<th>SMS Feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Sms response rate</td>
</tr>
<tr>
<td>2. Sms response latency</td>
</tr>
<tr>
<td>3. Sms response latency</td>
</tr>
<tr>
<td>4. Sms initiation</td>
</tr>
<tr>
<td>5. Sms initiation</td>
</tr>
</tbody>
</table>

Table 7: List of Basic Bluetooth Call Features

<table>
<thead>
<tr>
<th>Call Feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Call Contacts</td>
</tr>
<tr>
<td>2. Call Contacts to Interactions Ratio</td>
</tr>
<tr>
<td>3. Call Contacts to Interactions Ratio</td>
</tr>
<tr>
<td>4. Call Contacts to Interactions Ratio</td>
</tr>
<tr>
<td>5. Call Contacts to Interactions Ratio</td>
</tr>
</tbody>
</table>

Harvard University, Center for the Study of Learning and Development, 2014
Affective Computing

PROBLEM 3: Which emotions can be recognized?

BIG FIVE PERSONALITY TRAITS RECOGNITION

Personality traits can explain patterns of mobile phone usage (e.g. extraverted, neurotic, disagreeable, and unconscientious individuals spent more time writing and receiving SMS) [1].

<table>
<thead>
<tr>
<th>Predictors</th>
<th>n</th>
<th>M</th>
<th>SD</th>
<th>Min.</th>
<th>Max.</th>
<th>Skew</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neuroticism</td>
<td>112</td>
<td>20.73</td>
<td>8.04</td>
<td>2</td>
<td>38</td>
<td>.12</td>
</tr>
<tr>
<td>Extraversion</td>
<td>112</td>
<td>30.26</td>
<td>6.57</td>
<td>14</td>
<td>43</td>
<td>−.14</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>112</td>
<td>31.82</td>
<td>6.14</td>
<td>8</td>
<td>46</td>
<td>−1.0</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>112</td>
<td>34.5</td>
<td>6.54</td>
<td>15</td>
<td>48</td>
<td>−.45</td>
</tr>
<tr>
<td>SEI</td>
<td>112</td>
<td>75.32</td>
<td>16.5</td>
<td>32</td>
<td>100</td>
<td>−.58</td>
</tr>
</tbody>
</table>

(* N = 112).
Affective Computing

PROBLEM 3: Which emotions can be recognized?

HAPPINESS OR INDIVIDUAL STATE RECOGNITION

✧ Build an AI system that can automatically detect when a college student is becoming vulnerable to depression [1].

✧ Input sources

- Physiological data: electrodermal activity (EDA) (a measure of physiological stress), and 3-axis accelerometer (a measure of steps and physical activity)
- Survey data: questions related to academic activity, sleep, drug and alcohol use...
- Phone data: phone call, SMS, and usage patterns
- Location data: coordinates logged throughout the day

Affective Computing

PROBLEM 3: Which emotions can be recognized?

HAPPINESS OR INDIVIDUAL STATE RECOGNITION

✧ Classify each student as: happy or sad.
✧ Test different sources/ML classification algorithms[1].

<table>
<thead>
<tr>
<th>Modality</th>
<th>Dataset Size</th>
<th># Features</th>
<th>Classifier</th>
<th>Parameter Settings</th>
<th>C=100.0, RBF kernel, $\beta = 0.0001$</th>
<th>C=100.0, RBF kernel, $\beta = 0.01$</th>
<th>Num trees = 40, Max depth = infinite</th>
<th>C=100.0, RBF kernel, $\beta = 1$</th>
<th>C=0.1, Linear kernel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physiology</td>
<td>933</td>
<td>426</td>
<td>SVM</td>
<td></td>
<td>68.37%</td>
<td>71.26%</td>
<td>66.67%</td>
<td>69.95%</td>
<td>72.84%</td>
</tr>
<tr>
<td>Survey</td>
<td>1110</td>
<td>32</td>
<td>SVM</td>
<td></td>
<td>53.65%</td>
<td>50.86%</td>
<td>51.98%</td>
<td>53.65%</td>
<td>53.94%</td>
</tr>
<tr>
<td>Phone</td>
<td>1072</td>
<td>289</td>
<td>RF</td>
<td></td>
<td>71.26%</td>
<td>50.86%</td>
<td>55.95%</td>
<td>65.10%</td>
<td>68.48%</td>
</tr>
<tr>
<td>Mobility</td>
<td>905</td>
<td>15</td>
<td>SVM</td>
<td></td>
<td>51.79%</td>
<td>62.50%</td>
<td>55.95%</td>
<td>65.10%</td>
<td>68.48%</td>
</tr>
<tr>
<td>All</td>
<td>768</td>
<td>200</td>
<td>SVM</td>
<td></td>
<td>64.62%</td>
<td>62.50%</td>
<td>55.95%</td>
<td>65.10%</td>
<td>68.48%</td>
</tr>
</tbody>
</table>

Classification Accuracy

Table III

Table IV

Table VIII
Affective Computing

PROBLEM 3: Which emotions can be recognized?

Table 3
Affect recognition objective.

<table>
<thead>
<tr>
<th>Objective</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stress</td>
<td>35.7%</td>
</tr>
<tr>
<td>Emotion</td>
<td>33.3%</td>
</tr>
<tr>
<td>Wellbeing and user behaviour</td>
<td>14.3%</td>
</tr>
<tr>
<td>Personality</td>
<td>11.9%</td>
</tr>
<tr>
<td>Mood</td>
<td>4.8%</td>
</tr>
</tbody>
</table>

Table 2
Used representation models.

<table>
<thead>
<tr>
<th>Big-Five model</th>
<th>11.9%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distinct states model 61.9%</td>
<td></td>
</tr>
<tr>
<td>Ekman’s distinct states model</td>
<td>7.1%</td>
</tr>
<tr>
<td>Stress related model</td>
<td>35.7%</td>
</tr>
<tr>
<td>Happiness, boredom and other distinct states models</td>
<td>19.0%</td>
</tr>
<tr>
<td>Wellbeing and human behaviour models</td>
<td>14.3%</td>
</tr>
<tr>
<td>Dimensional models</td>
<td>11.9%</td>
</tr>
</tbody>
</table>
Affective Computing

- How to describe the emotions
- How to model affective data
- Which emotions can be recognized
- How to perform emotion detection
- Research issues
- Demo: Affective Tool
PROBLEM 4: How does emotion recognition work?

- **Unimodal** features (single input source)
  - Visual modality
  - Audio modality
  - Textual modality

- **Multi-modal** features (multiple input sources)
PROBLEM 4: How does emotion recognition work?

- **Facial expressions** are the primary cues for understanding emotions and sentiments.
- **FACS** (Facial Action Coding System) → database of facial expressions in terms of Actions Units (AU), i.e. contraction or relaxation of one or more muscles.

FACS example:

- E.g., Action code: 1, 2, 4, 5, 7, 20, 1C Inner brow raise
- 2C Outer brow raise
- 4B Brow lower
- 5D Upper lid raise
- 7B Lower lid tighten
- 20B Lip stretch
- 26B Jaw drop

http://mplab.ucsd.edu/grants/project1/research/face-detection.html
HAR systems involve two steps: training and classification.
PROBLEM 4: How does emotion recognition work?

Class of **deep neural networks**

Connectivity pattern between neurons resembles the organization of the **animal visual cortex**.

A ConvNet is able to successfully capture the Spatial and Temporal dependencies in an image through the application of relevant filters.

**Convolutional Neural Network (CNN)**

[Diagram of ConvNet]

Affective Computing

PROBLEM 4: How does emotion recognition work?

- **Unimodal** features (single input source)
  - Visual modality
  - Audio modality
  - Textual modality

- **Multi-modal** features (multiple input source)
PROBLEM 4: How does emotion recognition work?

Acoustic parameters are dependent on personality traits.
- Mel Frequency Cepstral Coefficients (MFCC)
  Short-term power spectrum of a voice or of a sound
- Spectral Flux
  Measure of how quick the power spectrum is changing
- Spectral Centroid
  Center of mass of the power spectrum, brightness of the sound
- Pause duration and Pitch
PROBLEM 4: How does emotion recognition work?

✧ OpenSMILE (https://www.audeering.com/opensmile/)

✧ Modular and flexible feature extractor for signal processing and machine learning applications

✧ Emotion recognition from audio parameters.
Affective Computing

PROBLEM 4: How does emotion recognition work?

✧ Unimodal features (single input source)
  ❏ Visual modality
  ❏ Audio modality
  ❏ Textual modality

✧ Multi-modal features (multiple input source)
PROBLEM 4: How does emotion recognition work?

Sentiment analysis systems can be categorized into two main classes:

Knowledge-based systems
✧ Use of linguistic patterns to understand the sentence structure based on its lexical dependency tree.
✧ Employ semantic analysis of text which allows the aggregation of conceptual and affective information (bag of concepts).

Statistics-based systems
✧ Assume the availability of a large dataset annotated with polarity or emotion labels; employ statistical and/or machine-learning algorithms.
Affective Computing

PROBLEM 4: How does affection recognition work?

(b) The old way: averaging over a bag of sentiment words. The overall polarity of a sentence is given by the algebraic sum of the polarity values associated with each affect word, divided by the total number of words.

(a) Dependency tree of a sentence.
Affective Computing

PROBLEM 4: How does emotion recognition work?

(d) The electronic circuit metaphor: sentiment words are "sources" while other words are "elements", e.g., very is an amplifier, not is a logical complement, rather is a resistor, but is an OR-like element that gives preference to one of its inputs.

(e) The final sentiment data flow of the "signal" in the "circuit".
Affective Computing

PROBLEM 4: How does affection recognition work?

DATA SOURCES IN AFFECTIVE COMPUTING STUDIES [1]

SENSOR DATA SOURCES IN AFFECTIVE COMPUTING STUDIES [1]
PROBLEM 4: How does emotion recognition work?

- **Unimodal** features (single input source)
  - Visual modality
  - Audio modality
  - Textual modality

- **Multi-modal** features (multiple input source)
Affective Computing

PROBLEM 4: How does emotion recognition work?

Early fusion

- It fuses the features extracted from various modalities, such as visual features, text, audio as a single vector which is then processed.

Decision-level fusion

- The features of each modality are examined and classified independently and results are fused as a decision vector.

Rule-based fusion

- Multimodal information is fused by statistical rule based methods such as linear weighted fusion, majority voting or custom-defined rules.
Multimodal sentiment analysis, consisting in harvesting sentiments from Web videos via a model that uses audio, visual and textual modalities [1].

These rules dealt with verbs which had as complements, either an adjective or a closed clause (i.e., a clause, usually finite, with its own subject).

**Trigger:** when the active token was head verb of one of the complement relations.

**Behavior:** if a word \( h \) was in a direct nominal object relationship with a word \( t \), then the concept \( h \rightarrow t \) was extracted.

**Example:** in (4), \( \text{smells} \) was the head of a clausal complement dependency relation with \( \text{bad} \) as the dependent.

(4) This meal smells bad. In this example, the concept \((\text{smell, bad})\) was extracted.

**Trigger:** when the active token was found to be the syntactic subject of a verb.

**Behavior:** if a word \( h \) was in a subject noun relationship with a word \( t \) then the concept \( t \rightarrow h \) was extracted.

**Example:** in (1), \( \text{movie} \) was in a subject relation with \( \text{boring} \).

(1) The movie is boring.

---

**Features**

- Distance between right eye and left eye
- Distance between the inner and the outer corner of the left eye
- Distance between the upper and the lower line of the left eye
- Distance between the left iris corner and the right iris corner of the left eye
- Distance between the inner and the outer corner of the right eye
- Distance between the upper and the lower line of the right eye
- Distance between the left iris corner and the right iris corner of the right eye
- Distance between the left eyebrow inner and the outer corner
- Distance between the right eyebrow inner and the outer corner
- Distance between top of the mouth and bottom of the mouth

**Visual Features**

**Audio Features**

**Text Features** (linguistic patterns)

- MFCC
- Spectral Centroid
- Spectral Flux
- Beat Histogram
- ....
Affective Computing

PROBLEM 4: How does emotion recognition work?

Two fusion schemes are investigated in [1]:

- **Early** fusion (i.e. combine all features within a single vector)
- **Linear weighted** fusion, with equal weights for each source

<table>
<thead>
<tr>
<th>Table 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Results of feature-level fusion.</td>
</tr>
<tr>
<td>Combination of modalities</td>
</tr>
<tr>
<td>---------------------------</td>
</tr>
<tr>
<td>Accuracy of the experiment carried out on Textual Modality</td>
</tr>
<tr>
<td>Accuracy of the experiment carried out on Audio Modality</td>
</tr>
<tr>
<td>Accuracy of the experiment carried out on Video Modality</td>
</tr>
<tr>
<td>Experiment using only visual and text-based features</td>
</tr>
<tr>
<td>Result obtained using visual and audio-based features</td>
</tr>
<tr>
<td>Result obtained using audio and text-based features</td>
</tr>
<tr>
<td>Accuracy of feature-level fusion of three modalities</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Results of decision-level fusion.</td>
</tr>
<tr>
<td>Combination of modalities</td>
</tr>
<tr>
<td>---------------------------</td>
</tr>
<tr>
<td>Experiment using only visual and text-based features</td>
</tr>
<tr>
<td>Result obtained using visual and audio-based features</td>
</tr>
<tr>
<td>Result obtained using audio and text-based features</td>
</tr>
<tr>
<td>Accuracy of decision-level fusion of three modalities</td>
</tr>
</tbody>
</table>
Affective Computing

- How to describe the emotions
- How to model affective data
- Which emotions can be recognized
- How to perform emotion detection
- Research issues
- Demo: Affective Tool
Affective Computing

Open Research Issues

- Privacy
  - Issues might arise not only regarding users of the emotion-aware systems but also others in its vicinity.
  - Data might not only describe the current affective state of the individuals but also his/her predicted future state.

- Implementation on resource-constrained devices
- Theoretical foundations of affective computing
- Build cultural transparent systems
Affective Computing

- How to describe the emotions
- How to model affective data
- Which emotions can be recognized
- How to perform emotion detection
- Research issues
- Demo: Affective Tool
Affective Computing

Automated facial coding has been applied to each facial bounding box and 34 landmarks to detect facial landmark detection. Using this data, we collected videos of hundreds of thousands of individuals. These videos were coded by expert FACS coders to provide a rich dataset of facial expression examples. In addition to facial action and emotion expression coding of facial expressions have been based framework using HOG. 3) Classification of the key facial measures. The classifiers have two operating modes: static and causal. In the bounding box is ignored. For details of the training and testing, we trained the SDK has classifiers for determining gender and whether the person is wearing glasses. The static classifiers allow classification of any facial action and emotion expression. The emotion expressions are given a similar score from 0 to 100. The EMFACS is an emotional facial action coding system.

Face and landmark detection

Oriented Gradient (HOG) features extraction

Classification via RNNs and CNNs


Figure 1 shows an overview.

Face and facial landmark detection.

Face texture feature extraction.

Emotion state classification.

Surprise, Anger …, Fear

Facial action classification.

Support Vector Machine (SVM), Random Forest (RF), and Artificial Neural Networks (ANNs) are used. The classifiers are based on multiple faces within a video. The causal classifiers leverage temporal information available in video sequences. The static classifiers allow classification of any facial action and emotion expression. The emotion expressions are given a similar score from 0 to 100. The classifiers have two operating modes: static and causal. In the bounding box is ignored. For details of the training and testing, we trained the SDK has classifiers for determining gender and whether the person is wearing glasses. The static classifiers allow classification of any facial action and emotion expression. The emotion expressions are given a similar score from 0 to 100. The EMFACS is an emotional facial action coding system.

https://www.affectiva.com
Measures 7 emotions and 20 facial expressions.

In addition to facial action and emotion expression classifiers, the SDK provides classifiers for determining gender and whether the person is wearing glasses.

Two operating modes: static and causal.
World’s largest emotion data repository
87 countries, 6.5M faces analyzed, 3.8B facial frames
Includes people emoting on device, and while driving

Top Countries for Emotion Data

USA 1,166K
MEXICO 150K
UNITED KINGDOM 265K
GERMANY 148K
BRAZIL 194K
THAILAND 184K
INDIA 1,363K
CHINA 562K
JAPAN 61K
VIETNAM 148K
PHILIPPINES 159K
INDONESIA 325K

😊 Affectiva
public class MyActivity implements Detector.FaceListener {
    @Override
    protected void onCreate(Bundle savedInstanceState) {
        detector.setFaceListener(this);
    }
}

detector.setDetectAllExpressions(true);
detector.setDetectAllEmotions(true);
detector.setDetectAllEmojis(true);
detector.setDetectAllAppearances(true);
Affective Computing

**INITIALIZE THE DETECTOR**

detector.start();

**PROCESS A PHOTO**

```java
byte[] frameData;
int width = 640;
int height = 480;
Frame.ByteArrayFrame frame = new Frame.ByteArrayFrame(frameData, width, height,
Frame.COLOR_FORMAT.YUV_NV21);
detector.process(frame);
```
public void onImageResults(List<Face> faces, Frame image, float timestamp) {
    // For each face found
    for (int i = 0; i < faces.size(); i++) {
        Face face = faces.get(i);
        int faceId = face.getId();
        // Appearance
        Face.GENDER genderValue = face.appearance.getGender();
        Face.GLASSES glassesValue = face.appearance.getGlasses();
        Face.AGE ageValue = face.appearance.getAge();
        Face.ETHNICITY ethnicityValue = face.appearance.getEthnicity();
        // Some Emotions
        float joy = face.emotions.getJoy();
        float anger = face.emotions.getAnger();
        float disgust = face.emotions.getDisgust();
        ....
    }
}
In questo capitolo viene trattato il funzionamento del sistema e la sua progettazione specificando la sua architettura e le componenti utilizzate.

2.1 Architettura e scelte progettuali

In questa sezione vengono presentati l'architettura dell'applicazione e scelte metodologiche che sono state adottate. L'applicazione segue lo schema presentato nella Fig. 2.1.

Il modello di classificazione è stato creato sulla base di un dataset di dati scaricato da internet [35], utilizzato in [36]. Si è preferito adottare questa soluzione in quanto non si aveva a disposizione un'automobile e perché il rilevamento di dati di guida presupone la prova di manovre pericolose e azzardate, e quindi più adatte a piloti esperti. Anche nei documenti consultati [18], [16], [8] si è adottata questa soluzione per garantire l'incolumità del guidatore.

Guida dell'app

1. La pressione del bottone "Start Trip" indica al sistema di cominciare a tracciare i sensori. L'utente può cliccando il bottone impostazioni o quello di spiegazione dell'app.

2. L'utente premendo il bottone impostazioni può decidere di iniziare a tracciare il proprio viaggio con le soglie di default oppure cambiarle secondo le sue esigenze.

Le soglie sono state impostate con tali valori dopo aver svolto varie prove ed aver analizzato i diversi risultati riportati dal sistema.

2.2 Funzionalità dell'applicazione

3. L'utente può leggere una guida dell'applicazione cliccando il bottone "?".

4. Durante la fase di rilevazione del comportamento dell'utente, egli/ella visualizza sulla carta anteriore la classe predetta dal sistema.

5. Rilevata la situazione pericolosa si attiva la fotocamera che traccia il volto dell'utente e lo avverte nel caso i dati rilevati superino le soglie impostate nel sistema.