Emotion Recognition and Personality Perception Using Acoustic/Prosodic Information and Semantic Labels

Speaker: Chia-Jui Liu
Advisor: Chung-Hsien Wu

Department of Computer Science and Information Engineering, National Cheng Kung University, Tainan, Taiwan
Outline

• Introduction
• Emotion Recognition
  • Acoustic/Prosodic Information
  • Semantic Labels
• Automatic Personality Perception
  • Background
  • ANN-based BFI detector
• Experimental Results
• Conclusion
Emotion Recognition

1. Acoustic/Prosodic Information
2. Semantic Labels

Emotion Recognition of Affective Speech Based on Multiple Classifiers Using Acoustic-Prosodic Information and Semantic Label, IEEE Trans. on Affective Computing, Jan-Mar, 2011
Introduction to Emotion Recognition

- The application of Spoken Dialog Systems are still limited to simple informational dialogue systems.
- To enable more complex systems, new capabilities such as affective interaction are needed.
- Emotion recognition is one approach to achieve the goal of affective interaction.
- Different kinds of affective information have been widely investigated for emotion recognition.
Related Work

- **Speech-based emotion recognition**
  - Acoustic features
  - Prosodic features
  - Spectrum
  - Cepstrum

- **Emotion recognition from Text** focused on discovery and utilization of emotional keywords (EK)
  - Keyword-based systems have following problems
    - Ambiguity in EK definition
    - Difficult to recognize the emotion of the sentence with no EK
    - Lack of affect-related semantic and syntactic knowledge base
Related Work

- For emotional state modeling
  - Support vector machine (SVM)
  - Gaussian mixture model (GMM)
  - Multilayer perceptron (MLP)
  - Decision tree

- Mixture of experts
Framework of Emotion Recognition
Emotional Salient Segment (ESS)

- Generally, an entire utterance is comprised of pause/breath segments and salient segments

- Steps of ESS extraction
  - Using Praat software to extract the pitch values of input speech
  - Legendre polynomial-based curve fitting for contour smoothing
  - Three types of ESS according to accent tones

- Only the ESS with largest duration used to feature extraction
Acoustic/Prosodic Feature Extraction

- Related work showed that acoustic and prosodic features reveal different performance due to different data and different task design

- In our system, a broad variety features from speech signal
  - Mean, Median, Max, Min, and Std. Dev. of Pitch (5)
  - Mean, Median, Max, Min, and Std. Dev. of Intensity (5)
  - Mean, Median, Max, Min, and Std. Dev. of formant 1-4 (20)
  - Mean, Median, Max, Min, and Std. Dev. of formant bandwidth 1-4 (20)
  - four types of jitter-related features (4)
  - six-types of shimmer-related features (6)
  - three-types of harmonicity-related features (3)
  - Mean, Median, Max, Min, and Std. Dev. of 12 dim MFCC + Energy, delta and acceleration coefficients (5x39)
Base-level Classifiers

- For $j$-th emotional state, given feature vector $\mathbf{f}$
  - GMM-based classifier
    \[
    P_{GMM}^j(\mathbf{f}) = \sum_k \omega_{j,k} P(\mathbf{f} \mid \mathbf{u}_{j,k}, \Sigma_{j,k})
    \]
  - SVM-based classifier
  - MLP-based classifier
Attributes in MDTs

- Attributes are derived from the emotional state probability distributions $P_L(f)$ predicted by the base-level classifier $L$ given the feature vector $f$

$$P_L(f) = \{P_L^1(f), P_L^2(f), ..., P_L^J(f)\}$$

- $J$ is the number of emotional state
- Attr. 1: Maximum probability over all emotional states
  $$L_{\text{Max Prob}} = \max_{j=1}^{J} P_L^j(f)$$
- Attr. 2: Entropy of emotional state probability distribution
  $$L_{\text{Entropy}} = -\sum_j P_L^j(f) \log_2 P_L^j(f)$$
- Attr. 3: Weight, $L_{\text{Weight}}$ is the fraction of the training data used to estimate the class probability of test data
Physical meaning of three attributes in MDTs

- MaxProb and Entropy can be interpreted as the estimates of the recognition confidence of the model
- Weight quantifies how reliable the model’s estimate of its own confidence

To induce MDT, the measure used in MLC4.5 is defined as

\[
info(S) = 1 - \max_{L \in L^\#} \text{accuracy}(L, S)
\]
Automatic Personality Perception

1. Introduction to Personality Perception
2. Background
3. Related Work
4. Problems
5. System Framework
6. Experiments
Personality Trait Chart

- Two dimensions

![Personality Trait Chart Diagram]

- 悠鬱的
- 難定的
- 易怒的
- 樂觀的

- Introversion
- Extraversion
- choleric
- sanguine
- melancholic
- phlegmatic

- moody
- anxious
- rigid
- sober
- pessimistic
- reserved
- quiet
- restless
- excitable
- changeable
- impulsive
- irresponsible
- outgoing
- talkative
- responsive
- easygoing
- lively
- carefree
- leadership
- calm
Introduction to Personality Perception

- Spontaneous and unaware processes influence our behavior to a large extent in social interactions
  - People make social inferences without intentions, awareness, or effort, i.e., spontaneously
- Whenever we listen to a voice for the first time, we attribute personality traits to the speaker
  - It significantly influences our behavior toward others
- Human sciences showing that nonverbal vocal behavior significantly influences personality perception
Introduction to Personality Perception

- Interactions with unknown individuals are frequent in our everyday life and include

- We attribute traits not only to people, but also to machines
  - Exhibiting human-like features and behaviors, including robots, embodied conversational agents, animated characters, etc.

- Perceived traits correlate with a wide spectrum of personal characteristics better than self-assessed traits as the actual personality of an individual
Introduction to Personality Perception

- Prosodic-based Automatic Personality Perception (APP) is automatically mapping acoustic/prosodic aspects of speech into personality traits attributed by human listeners.
- Unlike Automatic Personality Recognition (APR), which is expected to predict the real personality of an individual.
  - Perception indicate how you feel when listening to a speech clip.
- The goal of APP and APR is different.
  - APP: predict the personality as perceived by observers.
  - APR: predict the real personality of an given person.
- In other words, APP is not expected to predict the real personality of a given person, but the personality that others attribute to her in a given situation.
The most common personality representation relies on the Big Five (BF), five broad dimensions:

- Openness: Artistic, Curious, Imaginative, etc.
- Conscientiousness: Efficient, Organized, Planful, etc.
- Extroversion: Active, Assertive, Energetic, etc.
- Agreeableness: Appreciative, Kind, Generous, etc.
- Neuroticism: Anxious, Self-pitying, Tense, etc.
Background

- **Big Five Inventory** [John et al. 1999]
- 44-item inventory that measures an individual on the Big Five Factors of personality

<table>
<thead>
<tr>
<th>Disagree strongly</th>
<th>Disagree a little</th>
<th>Neither agree nor disagree</th>
<th>Agree a little</th>
<th>Agree Strongly</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

**Extraversion**: 1, 6R, 11, 16, 21R, 26, 31R, 36

**Agreeableness**: 2R, 7, 12R, 17, 22, 27R, 32, 37R, 42

**Conscientiousness**: 3, 8R, 13, 18R, 23R, 28, 33, 38, 43R

**Neuroticism**: 4, 9R, 14, 19, 24R, 29, 34R, 39

**Openness**: 5, 10, 15, 20, 25, 30, 35R, 40, 41R, 44

R: reverse-score (1→5, 2→4, 3, 4→2, 5→1)

1. Is talkative
2. Tends to find fault with others
3. Does a thorough job
4. Is depressed, blue
5. Is original, comes up with new ideas
6. Is reserved
7. Is helpful and unselfish with others
8. Can be somewhat careless
9. Is relaxed, handles stress well
10. Is curious about many different things
11. Is full of energy
12. Starts quarrels with others
13. Is a reliable worker
14. Can be tense
15. Is ingenious, a deep thinker
16. Generates a lot of enthusiasm
17. Has a forgiving nature
18. Tends to be disorganized
19. Worries a lot
20. Has an active imagination
21. Tends to be quiet
22. Is generally trusting
23. Tends to be lazy
24. Is emotionally stable, not easily upset
25. Is inventive
26. Has an assertive personality
27. Can be cold and aloof
28. Perseveres until the task is finished
29. Can be moody
30. Values artistic, aesthetic experiences
31. Is sometimes shy, inhibited
32. Is considerate and kind to almost everyone
33. Does things efficiently
34. Remains calm in tense situations
35. Prefers work that is routine
36. Is outgoing, sociable
37. Is sometimes rude to others
38. Makes plans and follows through with them
39. Gets nervous easily
40. Likes to reflect, play with ideas
41. Has few artistic interests
42. Likes to cooperate with others
43. Is easily distracted
44. Is sophisticated in art, music, or literature
Background

- The main instruments for score assignment are questionnaires
  - Big Five Inventory (BFI) is the most commonly questionnaires for personality assessment
    - Complete BFI (44 items)
    - BFI-10 (10 items)

The BFI-10 Questionnaire
Used in the Experiments of This Work

<table>
<thead>
<tr>
<th>ID</th>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>This person is reserved</td>
</tr>
<tr>
<td>2</td>
<td>This person is generally trusting</td>
</tr>
<tr>
<td>3</td>
<td>This person tends to be lazy</td>
</tr>
<tr>
<td>4</td>
<td>This person is relaxed, handles stress well</td>
</tr>
<tr>
<td>5</td>
<td>This person has few artistic interests</td>
</tr>
<tr>
<td>6</td>
<td>This person is outgoing, sociable</td>
</tr>
<tr>
<td>7</td>
<td>This person tends to find fault with others</td>
</tr>
<tr>
<td>8</td>
<td>This person does a thorough job</td>
</tr>
<tr>
<td>9</td>
<td>This person gets nervous easily</td>
</tr>
<tr>
<td>10</td>
<td>This person has an active imagination</td>
</tr>
</tbody>
</table>

- **Extroversion:** $Q_6 - Q_1$.
- **Agreeableness:** $Q_2 - Q_7$.
- **Conscientiousness:** $Q_8 - Q_3$.
- **Neuroticism:** $Q_9 - Q_4$.
- **Openness:** $Q_{10} - Q_5$. 
Related Work

- Both personality perception and personality recognition [Mairesse et al. 2006, 2007]
  - Feature: LIWC, MRC and prosodic features (separately and in combination)
  - Approaches: C4.5 decision tree learning, nearest neighbor, Naïve Bayes, Ripper, Adaboost, and SVM with linear kernels

- Use the statistical functions of the main prosodic features
  - Some traits can be classified satisfactory from the non-acted, speaker-independent data [Mohammadi et al. 2010]

- Feature selection approach [Polzehl et al. 2010]

- Addressing time-dependency and trait interplay [Polzehl et al. 2010]

- Perceived traits correlate with a wide spectrum of personal characteristics [Mohammadi et al. 2012]
Related Work

- The INTERSPEECH 2012 Speaker Trait Challenge [Schuller et al. 2012]
  - Speaker Personality Corpus (SPC)
  - Personality baseline results (by linear SVM and random forests)

### Partitioning of Speaker Personality Corpus

<table>
<thead>
<tr>
<th>SPC Sub-Task</th>
<th>#</th>
<th>Train</th>
<th>Devel</th>
<th>Test</th>
<th>Σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Openness</td>
<td>O</td>
<td>97</td>
<td>70</td>
<td>80</td>
<td>247</td>
</tr>
<tr>
<td></td>
<td>NO</td>
<td>159</td>
<td>113</td>
<td>121</td>
<td>393</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>C</td>
<td>110</td>
<td>81</td>
<td>99</td>
<td>290</td>
</tr>
<tr>
<td></td>
<td>NC</td>
<td>146</td>
<td>102</td>
<td>102</td>
<td>350</td>
</tr>
<tr>
<td>Extraversion</td>
<td>E</td>
<td>121</td>
<td>92</td>
<td>107</td>
<td>320</td>
</tr>
<tr>
<td></td>
<td>NE</td>
<td>135</td>
<td>91</td>
<td>94</td>
<td>320</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>A</td>
<td>139</td>
<td>79</td>
<td>105</td>
<td>323</td>
</tr>
<tr>
<td></td>
<td>NA</td>
<td>117</td>
<td>104</td>
<td>96</td>
<td>317</td>
</tr>
<tr>
<td>Neuroticism</td>
<td>N</td>
<td>140</td>
<td>88</td>
<td>90</td>
<td>318</td>
</tr>
<tr>
<td></td>
<td>NN</td>
<td>116</td>
<td>95</td>
<td>91</td>
<td>322</td>
</tr>
<tr>
<td>Σ</td>
<td></td>
<td>256</td>
<td>183</td>
<td>201</td>
<td>640</td>
</tr>
</tbody>
</table>

### Personality baseline results by linear SVM and random forests

<table>
<thead>
<tr>
<th>Task</th>
<th>C</th>
<th>SVM Devel UA (WA)</th>
<th>SVM AUC</th>
<th>Test UA (WA)</th>
<th>Test AUC</th>
<th>Random Forests</th>
<th>Devel UA (WA)</th>
<th>AUC</th>
<th>S_{opt}</th>
<th>Test UA (WA)</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(N)O</td>
<td>10^{-3}</td>
<td>60.4 (62.8)</td>
<td>67.6</td>
<td>57.8 (59.7)</td>
<td>62.9</td>
<td></td>
<td>57.7 ± 2.3 (64.4)</td>
<td>67.0</td>
<td>15</td>
<td>59.0 (63.7)</td>
<td>67.4</td>
</tr>
<tr>
<td>(N)C</td>
<td>10^{-2}</td>
<td>74.5 (74.9)</td>
<td>80.0</td>
<td>80.1 (80.1)</td>
<td>84.5</td>
<td></td>
<td>74.9 ± 0.9 (74.8)</td>
<td>81.2</td>
<td>25</td>
<td>79.1 (79.1)</td>
<td>83.7</td>
</tr>
<tr>
<td>(N)E</td>
<td>10^{-2}</td>
<td>80.9 (80.9)</td>
<td>90.5</td>
<td>76.2 (76.6)</td>
<td>84.1</td>
<td></td>
<td>82.8 ± 0.9 (82.8)</td>
<td>92.0</td>
<td>28</td>
<td>75.3 (75.6)</td>
<td>85.2</td>
</tr>
<tr>
<td>(N)A</td>
<td>10^{-3}</td>
<td>67.6 (65.6)</td>
<td>71.1</td>
<td>60.2 (60.2)</td>
<td>62.1</td>
<td></td>
<td>67.2 ± 1.4 (64.6)</td>
<td>71.6</td>
<td>5</td>
<td>64.2 (64.2)</td>
<td>66.7</td>
</tr>
<tr>
<td>(N)N</td>
<td>10^{-2}</td>
<td>68.0 (68.3)</td>
<td>71.9</td>
<td>65.9 (65.7)</td>
<td>71.8</td>
<td></td>
<td>68.9 ± 0.6 (68.9)</td>
<td>73.5</td>
<td>10</td>
<td>64.0 (63.7)</td>
<td>71.6</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>70.3 (70.5)</td>
<td>76.2</td>
<td>68.0 (68.5)</td>
<td>73.1</td>
<td></td>
<td>70.3 (71.1)</td>
<td>77.1</td>
<td></td>
<td>68.3 (69.3)</td>
<td>74.9</td>
</tr>
</tbody>
</table>
Related Work

- From Speech to Personality: Mapping Voice Quality and Intonation into Personality Differences [Mohammadi et al. 2012]
  - Jitter, shimmer, glissando likelihood
  - Predict the mutual position of two speakers in the personality space

- Speaker Personality Classification Using Systems Based on Acoustic-Lexical Cues and an Optimal Tree-Structured Bayesian Network [Audhkhasi et al. 2012]
  - Many pairs of OCEAN dimensions have appreciable correlation
  - Finding the tree with maximum sum of pairwise mutual information

\[
(o,c,e,a,n)_{MAP} = \arg \max_{(o,c,e,a,n) \in \{0,1\}^5} \left[ P(N = n|x) P(A = a|N = n, x) P(E = e|A = a, x) P(C = c|E = e, x) P(O = o|E = e, x) \right]
\]
Problems

- Passed research about APP or APR is predicting personality from speech features to results directly, but our personality assessment is done by BFI scoring.
- It would be a gap between BFI score and prediction result:
  - BFI score come from the speech
  - Prediction result come from speech features
- BFI is not only for scoring the personality but also should be a cue to predict personality perception.
Framework of Personality Perception

Training
Speech Personality Corpus → Feature Extraction → Acoustic-Prosodic Features → ANN-based BFI Detector → BFI Score computing

Testing
Speech Clip → Feature Extraction → Personality Prediction → BFI Detection Model → Prediction Result
ANN-based BFI Detector

Speech Clip

Feature Vectors

ANN-based BFI Detector

\[
\begin{bmatrix}
P(B_1 | o_1) \\
P(B_2 | o_1) \\
\vdots \\
P(B_M | o_1) \\
P(B_1 | o_2) \\
P(B_2 | o_2) \\
\vdots \\
P(B_M | o_2) \\
\vdots \\
P(B_1 | o_N) \\
P(B_2 | o_N) \\
\vdots \\
P(B_M | o_N)
\end{bmatrix}
\]

\(M = 44\) or \(10\), \(N = \) frame number

Result
ANN-based BFI Detector

- Feature Extraction
  - Low-Level Feature Descriptor (frame)
    - Zero-crossing rate
    - Root means square (RMS) frame energy
    - Pitch
    - Harmonics-to-noise (HNR) by autocorrelation function
    - MFCC 1-12
    - Delta coefficient of above features
  - Frame length is 25ms with 10ms overlapped
  - OpenSMILE toolkit [Eyben et al. 2009] is used for feature extraction
Experiments

- **Databases**
  - Berlin Emotional Speech Database (EMO-DB) [Burkhardt et al. 2005]
    - 535 sentences, about 120000 frames
    - 10 speakers (5 male, 5 female)
    - German

- **Assessment**
  - Listen to a speech clips, then filling BFI for observes
  - Score turn into below (1, 2, 3) and above (4, 5)

- **Speech Example**

  ![Audio Controls](sound_icon)
Experiments

- **ANN Setup statistics**
  - Hidden nodes: 300
  - 100 iterations
  - Features --> Hidden Layer, 串 -4~+4 frames
  - Hidden Layer --> Output, 串 0~3 frames
  - Feedback 連結,串 1~9 frames

- **K-fold cross-validation (K=5)**
Experiments

- Number of instances per BFI in our corpus

<table>
<thead>
<tr>
<th>BFI item</th>
<th>Numbers of High/Low class</th>
<th>BFI item</th>
<th>Numbers of High/Low class</th>
</tr>
</thead>
<tbody>
<tr>
<td>BFI_1</td>
<td>351/184</td>
<td>BFI_23</td>
<td>282/253</td>
</tr>
<tr>
<td>BFI_2</td>
<td>242/293</td>
<td>BFI_24</td>
<td>370/165</td>
</tr>
<tr>
<td>BFI_3</td>
<td>391/144</td>
<td>BFI_25</td>
<td>415/120</td>
</tr>
<tr>
<td>BFI_4</td>
<td>380/155</td>
<td>BFI_26</td>
<td>176/359</td>
</tr>
<tr>
<td>BFI_5</td>
<td>368/167</td>
<td>BFI_27</td>
<td>140/395</td>
</tr>
<tr>
<td>BFI_6</td>
<td>217/318</td>
<td>BFI_28</td>
<td>412/123</td>
</tr>
<tr>
<td>BFI_7</td>
<td>383/152</td>
<td>BFI_29</td>
<td>306/229</td>
</tr>
<tr>
<td>BFI_8</td>
<td>221/314</td>
<td>BFI_30</td>
<td>351/184</td>
</tr>
<tr>
<td>BFI_9</td>
<td>287/248</td>
<td>BFI_31</td>
<td>429/106</td>
</tr>
<tr>
<td>BFI_10</td>
<td>480/55</td>
<td>BFI_32</td>
<td>364/171</td>
</tr>
<tr>
<td>BFI_11</td>
<td>435/100</td>
<td>BFI_33</td>
<td>298/237</td>
</tr>
<tr>
<td>BFI_12</td>
<td>173/362</td>
<td>BFI_34</td>
<td>344/191</td>
</tr>
<tr>
<td>BFI_13</td>
<td>259/276</td>
<td>BFI_35</td>
<td>380/155</td>
</tr>
<tr>
<td>BFI_14</td>
<td>374/161</td>
<td>BFI_36</td>
<td>303/232</td>
</tr>
<tr>
<td>BFI_15</td>
<td>434/101</td>
<td>BFI_37</td>
<td>146/389</td>
</tr>
<tr>
<td>BFI_16</td>
<td>419/116</td>
<td>BFI_38</td>
<td>419/116</td>
</tr>
<tr>
<td>BFI_17</td>
<td>383/152</td>
<td>BFI_39</td>
<td>362/173</td>
</tr>
<tr>
<td>BFI_18</td>
<td>235/300</td>
<td>BFI_40</td>
<td>374/161</td>
</tr>
<tr>
<td>BFI_19</td>
<td>364/171</td>
<td>BFI_41</td>
<td>217/318</td>
</tr>
<tr>
<td>BFI_20</td>
<td>410/125</td>
<td>BFI_42</td>
<td>362/173</td>
</tr>
<tr>
<td>BFI_21</td>
<td>195/340</td>
<td>BFI_43</td>
<td>267/268</td>
</tr>
<tr>
<td>BFI_22</td>
<td>193/342</td>
<td>BFI_44</td>
<td>182/353</td>
</tr>
</tbody>
</table>
Experiments

- Number of instances per trait in our corpus

<table>
<thead>
<tr>
<th>Personality Traits</th>
<th>Numbers of High/Low class</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Based on BFI</strong></td>
<td></td>
</tr>
<tr>
<td>Extraversion</td>
<td>347/188</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>262/273</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>377/158</td>
</tr>
<tr>
<td>Neuroticism</td>
<td>377/158</td>
</tr>
<tr>
<td>Openness</td>
<td>420/115</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>1783/892</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Personality Traits</th>
<th>Numbers of High/Low class</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Based on BFI-10</strong></td>
<td></td>
</tr>
<tr>
<td>Extraversion</td>
<td>338/197</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>421/114</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>346/189</td>
</tr>
<tr>
<td>Neuroticism</td>
<td>350/185</td>
</tr>
<tr>
<td>Openness</td>
<td>290/245</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>1745/930</td>
</tr>
</tbody>
</table>
Experiments

- Experimental Results

<table>
<thead>
<tr>
<th>Results (frame)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>BFI num.</td>
<td>44</td>
<td></td>
</tr>
<tr>
<td>Total Frame</td>
<td>129965</td>
<td></td>
</tr>
<tr>
<td>All Right Frame acc.</td>
<td>0.742 %</td>
<td></td>
</tr>
<tr>
<td>Average BFI Frame acc.</td>
<td>70.01 %</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Results (frame)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>BFI num.</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Total Frame</td>
<td>129965</td>
<td></td>
</tr>
<tr>
<td>All Right Frame acc.</td>
<td>10.848 %</td>
<td></td>
</tr>
<tr>
<td>Average BFI Frame acc.</td>
<td>71.563 %</td>
<td></td>
</tr>
</tbody>
</table>
Experiments

- Prediction Rate of Personality Perception

<table>
<thead>
<tr>
<th>BFI-based Results (Sentence)</th>
<th>BFI-10-based Results (Sentence)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Openness acc.</td>
<td>Openness acc.</td>
</tr>
<tr>
<td>67.85 %</td>
<td>70.27 %</td>
</tr>
<tr>
<td>Conscientiousness acc.</td>
<td>Conscientiousness acc.</td>
</tr>
<tr>
<td>66.49 %</td>
<td>68.68 %</td>
</tr>
<tr>
<td>Extroversion acc.</td>
<td>Extroversion acc.</td>
</tr>
<tr>
<td>71.94 %</td>
<td>73.86 %</td>
</tr>
<tr>
<td>Agreeableness acc.</td>
<td>Agreeableness acc.</td>
</tr>
<tr>
<td>72.41 %</td>
<td>81.63 %</td>
</tr>
<tr>
<td>Neuroticism acc.</td>
<td>Neuroticism acc.</td>
</tr>
<tr>
<td>70.56 %</td>
<td>72.46 %</td>
</tr>
<tr>
<td>Average acc.</td>
<td>Average acc.</td>
</tr>
<tr>
<td>69.85 %</td>
<td>73.38 %</td>
</tr>
</tbody>
</table>
Conclusions

• This work has presented experiments on acoustic-prosodic based Automatic Personality Perception using an ANN classifier
• Experimental results show an accuracy about 70% in predicting the perception of observers which outperform most baseline system
• The effectiveness of our ANN-based BFI detector for predicting personality perception is confirmed, but several future studies are needed
  • More speech features
  • The relationship between BFI items
Thanks for your attention