“Speech and Personality”

Florian Metze
National Science Foundation, May 7, 2015
Florian Metze

“What’s in Speech That’s Not in Text”

• Analysis of Personality in Speech
• New paradigms: articulatory features & deep learning
• Customized Speech Recognition
  • Low resource ASR, ASR for non-experts
• Speech in Healthcare
• Large-scale Multi-media Analysis
• Analysis of non-speech audio
Personality in Speech

- No role whatsoever in current speech technology (recognition or synthesis)
State of the Art
Personality in Speech

• No direct way of measuring personality, instead using questions like “I like to have people around me”

• Classification and Regression of manually labeled data
  • Self-assessment or external assessment?

• Very few databases available

• We can always engineer a specific, single “good enough” solution, but we’ll never solve the problem (even with more data & labels)
Automatic Prediction of Personality Ratings

![Graph showing SVM prediction correlation vs number of features for different personality traits: Extroversion, Neuroticism, Agreeableness, Conscientiousness, and Openness.](image)
Analysis of Personality Impressions in Speech

- Could help uncover the structure of personality traits
- Only N-/ O- assignments and absolute magnitude of distances differ
- 4 “natural” classes emerge:
  - All “decreased” (-)
  - C+ and A+
  - E+ and O+
  - N+
Challenges

• Building isolated “good enough” solutions won’t advance understanding and never result in scalable approaches

• We need more than just speech and audio information, we want
  • Text, Image, Video
  • Non-English data, cultural differences
  • Synthesis ↔ recognition

• For “personality” to generalize, we need richer data, and scale opportunity to work for and with non-speech experts
  • HCI folks, roboticists, doctors/ therapists, ethnologists, …
Solution 1: “Robustness to User”

- Speech processing must become “universal cultural skill”
  - Same as text processing
- Not every college graduate will have the opportunity to attend classes at CMU/ LTI
  - Step up speech processing (self-)education
  - Stop thinking about tools, data, labels, etc. in isolation
  - Think about experiments & solutions, “work-benches”
Solution 2: “MM Twitter”

Can you fly this thing?
Thank You!

Questions?