449.004 - Augmenting Natural Communication in Nonverbal Individuals with Autism

Background:

Despite technological and usability advances, some individuals with minimally verbal autism (mvASD) still struggle to convey affect and intent using current augmentative communication systems. In contrast, their non-speech vocalizations are often affect and context rich and accessible in almost any environment. Our system uses primary caregivers' unique knowledge of an individual's vocal sounds to label and train machine learning models in order to build holistic communication technology (see Figure 1).

Objectives:

This work involves the development of 4 major research outputs:

- Scalable data collection methods to capture naturalistic nonverbal communication, including intuitive in-the-moment labeling by primary caregivers
- Signal processing techniques to accurately characterize the real-world audio, including signal-label alignment and spectral classification
- Personalized machine learning methods to elucidate an individual's unique utterances
- Interactive augmentative communication interfaces to increase agency and improve understanding

Methods:

We first identified the needs of the community through interviews (n=5) and surveys (n=18) with mvASD individuals and their families. We then conducted an eight-month case study with an elementary-aged nonverbal child and his family. Through a highly participatory design process, we refined the data collection process, created an inexpensive wearable audio recording system, and developed and deployed an open-source Android app for primary caregivers to label communicative and affective exchanges in real-time with minimal burden or interference. We collected over 13 hours of unprompted vocalizations from the child in his everyday environments with more than 300 labeled instances. The labeled signals were then used to classify states of affect, interaction, or communicative intent using multiple machine learning methods.

Results:

Initial machine learning results using a multi-class support vector machine (SVM) with 6 researcher-labeled states produced a weighted f1-score of 0.67, suggesting these states could be differentiated with audio only. Visual inspection of the audio waveforms highlighted the diversity and distinct characteristics of the individual's vocalizations, which varied in tone, pitch, and duration depending on the individual's emotional or physical state and intended communication. (Canonical examples from this individual are shown in Figure 2.) We then trained a deep learning model using a long short-term memory (LSTM) recurrent neural network (RNN) and a zero-shot transfer learning approach. This method employed a generic audio database to classify three categories of caregiver-labeled...
vocalizations: laughter, negative affect, and self-soothing sounds. We identified laughter and negative affect with 70% and 69% accuracy, respectively, but classification of the self-soothing sounds produced accuracies around chance.

Conclusions:

These results highlight both the need for specialized, naturalistic databases and novel computational methods and their potential to enhance communicative and affective exchanges between mvASD individuals and the broader community. Future work includes 1) the development of a database of naturalistic mvASD vocalizations to engage the machine learning community, 2) data processing pipelines for signal accuracy and alignment, and 3) personalized, semi-supervised machine learning methods that leverage the sparsely labeled structure of the real-world data. Each of these steps is built on an iterative co-designing process with autistic stakeholders with the goal of increasing agency and
enhancing dialogue between individuals with mvASD and the world.

Figure 1: An augmentative, total communication system for nonverbal individuals using naturalistic vocalizations

Figure 2: Example high-quality waveforms from an mvASD individual’s vocalizations for six caregiver-labeled classes

Authors

K T. Johnson
MIT Media Lab
Detection of Autism Related Behaviors from Video of Infants Using Machine Learning

S. Liaqat¹, C. Wu², C. N. Chuah³, S. C. S. Cheung¹ and S. Ozonoff⁴, (1)University of Kentucky, Lexington, KY, (2)Computer Science, UC Davis, Davis, CA, (3)University of California, Davis, Davis, CA, (4)Psychiatry and Behavioral Sciences, University of California at Davis, MIND Institute, Sacramento, CA

Digital Assessment of Social Communication Skill, Arousal and Gross Motor Functioning in Infants, Toddlers, Children and Adults

R. T. Schultz¹, K. Bartley², C. J. Zampella¹, E. Sariyanidi², W. Guthrie¹, J. Pandey¹, J. D. Herrington³, B. Tunc¹ and J. Parish-Morris¹, (1)Center for Autism Research, Children's Hospital of Philadelphia, Philadelphia, PA, (2)Center for Autism Research, The Children's Hospital of Philadelphia, Philadelphia, PA, (3)Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA

Innervoice: Emotional Communication Combined with Artificial Intelligence

L. J. Brady, iTherapy, LLC, Martinez, CA

Toward a Digital Phenotype of Autism: An Exploratory Study Using Twitter Data

A. Jutla¹, M. R. Donohue², H. E. Reed³ and J. Veenstra-Vander Weele⁴, (1)1051 Riverside Drive, New York State Psychiatric Institute / Columbia University, New York, NY, (2)Washington University School of Medicine, St. Louis, MO, (3)Columbia University Medical Center, New York, NY, (4)Psychiatry, New York State Psychiatric Institute / Columbia University, New York, NY

Magical Musical Mat: Augmenting Communication with Touch and Music

R. Chen¹, A. Ninh² and B. Yu³, (1)Graduate School of Education, UC Berkeley, Berkeley, CA, (2)University of California, Berkeley, Berkeley, CA, (3)San Francisco State University, Berkeley, CA