

Call for Papers

INTERSPEECH 2009 Emotion Challenge

humaine
emotion-research.net

Feature, Classifier, and Open Performance Comparison for Non-Prototypical Spontaneous Emotion Recognition

Deutsche Telekom Laboratories
An-Institut der Technischen Universität Berlin



Organisers:

Bjoern Schuller (Technische Universität München, Germany)

Stefan Steidl (FAU Erlangen-Nuremberg, Germany)

Anton Batliner (FAU Erlangen-Nuremberg, Germany)

Sponsored by:

HUMAINE Association

Deutsche Telekom Laboratories

The Challenge

The young field of emotion recognition from voice has recently gained considerable interest in Human-Machine Communication, Human-Robot Communication, and Multimedia Retrieval. Numerous studies have been seen in the last decade trying to improve on features and classifiers. However, in comparison to related speech processing tasks such as Automatic Speech and Speaker Recognition, practically no standardised corpora and test-conditions exist to compare performances under exactly the same conditions. Instead, a multiplicity of evaluation strategies employed such as cross-validation or percentage splits without proper instance definition, prevents exact reproducibility. Further, to face more realistic use-cases, the community is in desperate need of more spontaneous and less prototypical data.

In these respects, the **INTERSPEECH 2009 Emotion Challenge** shall help bridging the gap between excellent research on human emotion recognition from speech and low compatibility of results: the FAU Aibo Emotion Corpus of spontaneous, emotionally coloured speech, and benchmark results of the two most popular approaches will be provided by the organisers. Nine hours of speech (51 children) were recorded at two different schools. This allows for distinct definition of test and training partitions incorporating speaker independence as needed in most real-life settings. The corpus further provides a uniquely detailed transcription of spoken content with word boundaries, non-linguistic vocalisations, emotion labels, units of analysis, etc.

Three sub-challenges are addressed in two different degrees of difficulty by using non-prototypical five or two emotion classes (including a garbage model):

- The **Open Performance Sub-Challenge** allows contributors to find their own features with their own classification algorithm. However, they will have to stick to the definition of test and training sets.
- In the **Feature Sub-Challenge**, participants are encouraged to upload their individual best features per unit of analysis with a maximum of 100 per contribution. These features will then be tested by the organisers with equivalent settings in one classification task, and pooled together in a feature selection process.
- In the **Classifier Sub-Challenge**, participants may use a large set of standard acoustic features provided by the organisers for classifier tuning.

The labels of the test set will be unknown, but each participant can upload instance predictions to receive the confusion matrix and results up to 25 times. As classes are un-balanced, the measure to optimise will be mean recall. The organisers will not take part in the sub-challenges but provide baselines.

Overall, contributions using the provided or an equivalent database are sought in (but not limited to) the areas:

- Participation in any of the sub-challenges
- Speaker adaptation for emotion recognition
- Noise/coding/transmission robust emotion recognition
- Effects of prototyping on performance
- Confidences in emotion recognition
- Contextual knowledge exploitation

The results of the Challenge will be presented at a Special Session of [Interspeech 2009](#) in Brighton, UK.

Prizes will be awarded to the sub-challenge winners and a best paper.

If you are interested and planning to participate in the Emotion Challenge, or if you want to be kept informed about the Challenge, please send the organisers an [e-mail](#) to indicate your interest and visit the [homepage](#):

<http://emotion-research.net/sigs/speech-sig/emotion-challenge>