

# SMAD: A tool for automatically annotating the smile intensity along a video record

Stéphane Rauzy and Mary Amoyal

Laboratoire Parole et Langage  
CNRS & Aix Marseille Université, France

`stephane.rauzy@univ-amu.fr`, `mary.amoyal@univ-amu.fr`

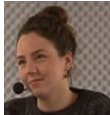
# Analysis of smiles in conversational corpus

Smile considered as an “interactive facial gesture” (Bavelas and Gerwing, 2007)

- Convey some information during the collaborative process (e.g. the co-construction of Common Ground during the interaction)
- Require a smile representation more subtle than presence versus absence of smile, e.g. the 5 levels Smiling Intensity Scale of Gironzetti et al. (2016)

Manual annotation of smile is a time consuming task : 1 hour of manual annotation for 1 minute of video

Any automatic tool reducing the annotation time is welcome!



## **Neutral facial expression (S0):**

No smile, no flexing of the zygomaticus (no AU12)

## **Closed mouth smile (S1):**

Flexing of the zygomaticus (AU12), may show flexing of the orbicularis oculi (AU6 or AU7).

## **Open mouth smile (S2):**

Showing upper teeth (AU25), plus AU12, AU6 or AU7.

**Wide open mouth smile (S3):** Showing lower and upper teeth (AU25 and AU26, plus AU12, AU6 or AU7.

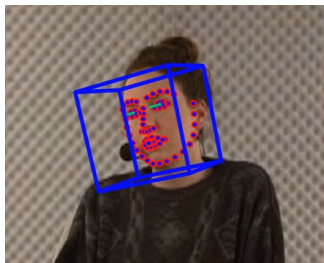
## **Laughing smile (S4):**

Jaw dropped (AU25 and AU26 or AU27), plus AU12, AU6 or AU7.

# The OpenFace toolkit

OpenFace (Baltrušaitis et al., 2018) is a state-of-the-art open source software coming from the field of Computer Vision and Machine Learning.

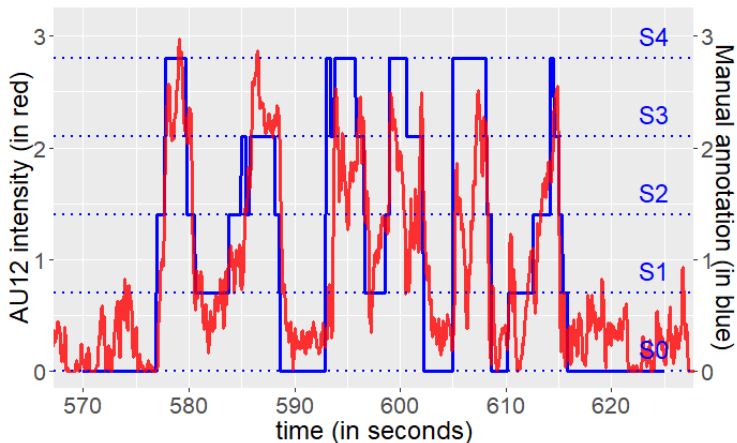
From a video record, OpenFace performs head tracking, facial landmark detection, head pose estimation, facial action unit recognition and eye-gaze estimation.



- Input: A video record
- Output for each video frame:
  - Facial landmark positions
  - Head pose angles
  - Intensities of the facial Action Units (AUs)

Figure: Example of a frame capture of the OpenFace processed video.

# AU OpenFace measurements vs manual annotations



**Figure:** An example of the temporal correlation of the zygomaticus AU12 intensity measured by OpenFace (the red curve) with the manually annotated smiles on the SIS system (blue curve).

# Building up the automatic annotation engine

A “small” training dataset with SIS manual annotations and AU measurements (1300 labeled intervals from  $S_0$  to  $S_4$ ).

- A 4 steps HMM-like process:
  - Find time intervals with  $S_0$  label vs group  $S_1S_2S_3S_4$  label
  - Among the remaining group, find intervals  $S_1$  vs  $S_2S_3S_4$
  - Iterate until step 4 which split the intervals in  $S_3$  or  $S_4$
- The label prediction at time  $t$  depends on the AU measurements and of the previous label at time  $t - 1$
- Each AU intensity contributes differently to each step (see table of coefficients below)

Name	Step 1	Step 2	Step 3	Step 4
AU06	-	0.45	0.67	0.62
AU07	0.26	-	0.15	0.21
AU12	0.64	0.22	-	0.17
AU25	0.10	0.33	0.18	-

# The smiles automatic annotation engine SMAD

Output of SMAD: A sequence of adjusted time intervals labeled from  $S_0$  to  $S_4$  following the SIS system.

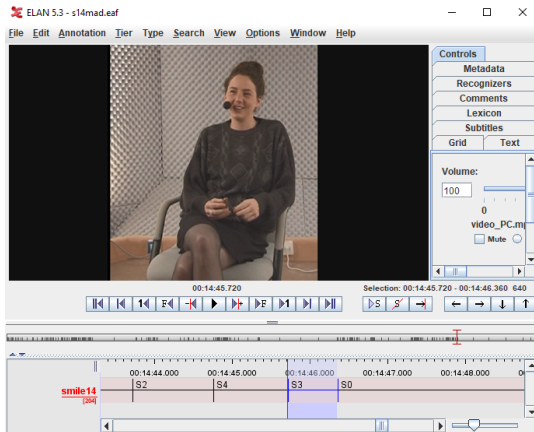


Figure: Example of SMAD output formatted for the ELAN software.

# Evaluation of the tool

We performed two kinds of evaluation:

- Using the 1 hour manually annotated corpus and the standard metrics (confusion matrix, precision, recall, Cohen kappa, ...).  
The overall accuracy is of 68%. Classes  $S_0$  and  $S_4$  obtain very good scores, the distinction between classes  $S_1$ ,  $S_2$  and  $S_3$  is more fuzzy.
- Evaluation of the time saving by comparing:
  - The time spent for manually correcting the labels and interval time boundaries of the automatic SMAD output
  - The time for manually annotate the data without pretreatment

A gain of a factor 10 in annotation time (6 hours to correct 1 hour of video as compared to 60 hours of annotation without pretreatment)



# Conclusions and perspectives

- The methodology (manual correction of the SMAD automatic smile output) has been applied successfully to PACO (Amoyal et al., 2020), a 5 hours corpus of conversational data built up for analyzing the impact of common ground in spontaneous face-to-face interaction
- SMAD is an open source and collaborative project, so feel free to contribute!  
The SMAD scripts and documentation are available at the github url:

`https://github.com/srauzy/HMAD`