AMHUSE
A MULTIMODAL DATASET FOR HUMOR SENSING

Giuseppe Boccignone[1], Donatello Conte[2], Vittorio Cuculo[1], Raffaella Lanzarotti[1]
[1]Università degli Studi di Milano, Italy - [2]Université de Tours, France

UNIVERSITÀ DEGLI STUDI DI MILANO

UNIVERSITÉ FRANÇOIS-RABELAIS TOURS

PHUSE Lab — Perceptual computing and HUman SEnsing

AMHUSE Dataset

A **M**ultimodal dataset for **HU**mour **SE**nsing is the result of an experiment concerning **amusement** elicitation.

Involved **36 subjects**: 9 females and 27 males, with an age varying from 18 to 54 years old ($\mu = 26.7$ and $\sigma = 8.8$).
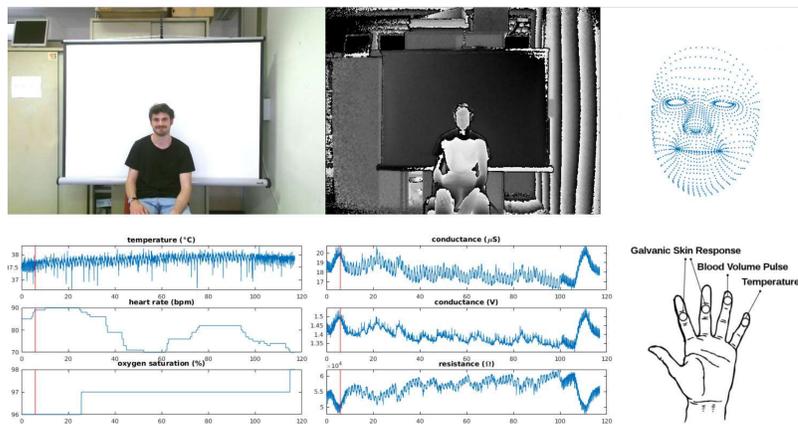

Fig.1: *Samples of RGB recorded videos.*

Fig.2: *Samples of all recorded signals, and detail of the physiological sensors placement (bottom-right).*

The **stimuli** (video clips) have been chosen being aware that people have different sense of humour. Thus each clip features a different kind of comicality:

- **parody**
- **political satire**
- **sketch comedy**

Real-time **recorded signals**, relevant for emotion detection, include:

- **RGB video** (+ face landmarks)
- **depth video** (+ face landmarks)
- **electrodermal activity (EDA)**
- **blood volume pulse (BVP)**
- **skin temperature (SKT)**

adopting a Kinect camera system and low-cost wearable sensors.

Self reports were collected in terms of **discrete PAD** annotations from all the 36 subjects after each stimulus. The agreement between subjects leads to ICC(3,k) $= 0.844$, obtained as the mean of the ICC for:

- pleasure ($= 0.664$),
- arousal ($= 0.921$),
- dominance ($= 0.948$).

On average, the subjects had **remarkable agreement** on their ratings, validating stimulus' effectiveness.

To request AMHUSE dataset
http://amhuse.phuselab.di.unimi.it

Dataset	Subjects	Audio	RGB	Depth	EEG	ECG	BVP	EDA	EMG	R	T	G	FP	Pose	AU
MAHNOB-HCI [4]	30	✓	✓	-	✓	✓	-	✓	-	✓	✓	✓	-	-	-
RECOLA [2]	46(18)	✓	✓	-	-	✓	-	✓	-	-	-	-	-	✓	✓
OPEN_EmoRec_II [3]	30	✓	✓	-	-	-	✓	✓	✓	✓	-	-	-	-	-
AMHUSE	36	-	✓	✓	-	-	✓	✓	-	-	✓	-	✓	-	✓

Tab. 1: *Comparison with existing multimodal datasets. In brackets the number of complete data. EEG = electroencephalography, ECG = electrocardiogram, BVP = blood volume pulse, EDA = electrodermal activity, R = respiration, T = temperature, G = eye gaze tracking, FP = Facial Points, AU = Action Units.*

DANTE Annotation Tool

Dimensional **AN**notation **T**ool for **E**motions, is a novel **web-based** annotation tool for **valence and arousal** continuous value in presence of recorded **audio/videos**.
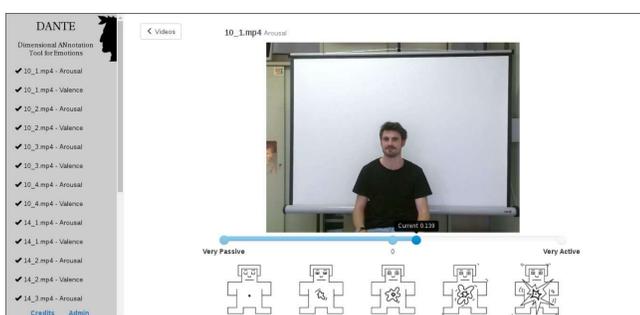

Fig.4: *An annotation session using DANTE.*

Problems with existing annotation tools include:

- **recording rate** not customisable;
- **annotations saved** only on text files;
- missing **administration** interface to manage annotators and videos;
- unable to create **groups** of annotators;
- absence of **visual indication** for the videos already annotated.

External annotations of AMHUSE dataset were made by 4 annotators (3 males and 1 female), and aggregated adopting the **Evaluator Weighted Estimator (EWE)** [1]. It weights the annotations of each rater by her respective agreement with the others, defining an individual evaluator confidence score.
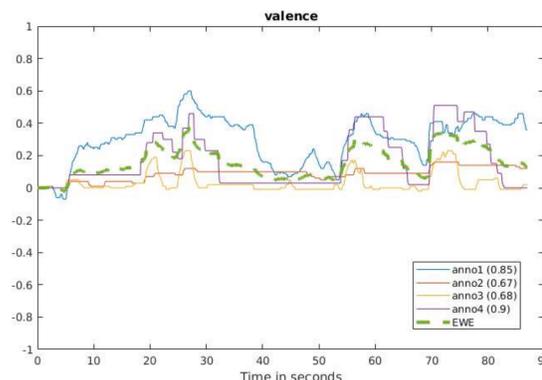

Fig.3: *Annotations of valence from 4 annotators and the corresponding gold standard in dashed green line. In brackets it is shown the evaluator score, namely the inter-rater agreement.*

Results indicate good **inter-rater reliability** for the **valence**, whilst, as expected, a poorer value for **arousal**. Indeed, it is well known that the level of arousal is more difficult to distinguish than valence, resulting in a lower agreement between the annotators. The Cronbach's α of valence annotations falls in the range $0.8 \leq \alpha < 0.9$, which is to be considered a **good internal consistency**.

	α	CC	CCC
Arousal	0.325	0.310	0.093
Valence	0.842	0.742	0.509

To download DANTE annotation tool
https://github.com/phuselab/DANTE

References

[1] M. Grimm and K. Kroschel. Evaluation of natural emotions using self assessment manikins. In *Automatic Speech Recognition and Understanding, 2005 IEEE Workshop on*, pages 381–385. IEEE, 2005.

[2] F. Ringeval, A. Sonderegger, J. Sauer, and D. Lalanne. Introducing the recola multimodal corpus of remote collaborative and affective interactions. In *Automatic Face and Gesture Recognition (FG), 2013 10th IEEE International Conference and Workshops on*, pages 1–8. IEEE, 2013.

[3] S. Rukavina, S. Gruss, S. Walter, H. Hoffmann, and H. C. Traue. OPEN EmoRec II – a multimodal corpus of human-computer interaction. *International Journal of Computer, Electrical, Automation, Control and Information Engineering*, 9(5):1181–1187, 2015.

[4] M. Soleymani, J. Lichtenauer, T. Pun, and M. Pantic. A multimodal database for affect recognition and implicit tagging. *IEEE Transactions on Affective Computing*, 3(1):42–55, 2012.

Acknowledgements