

# DANTE

## Dimensional ANnotation Tool for Emotions

Progetto di Interazione Naturale e Computazione Affettiva  
Università degli studi di Milano  
Milan, Italy

Lorenzo Franceschi 886473  
Filippo Lovat 895085

April 7, 2018

**D**ANTE (Dimensional ANnotation Tool for Emotions) is an emotional annotation tool to annotate any kind of video in terms of valence and arousal continuous dimensions.

DANTE  
Dimensional ANnotation  
Tool for Emotions



In other terms, it requires: Apache HTTP Server, MySQL and PHP.

To install, configure and test Apache 2.4 and PHP 7 for development on Windows follow the instructions in this link <sup>2</sup>.



### 1.2 Installation

To install DANTE follow these steps:

- Clone the repository using `git clone https://github.com/PHuSeLab/DANTE.git`.
- Import in a MySQL database the file `annotationdb.sql`, which is responsible of the creation of all the needed tables.
- Edit the file `config.php` according to the environment. (For more details, see section configuration.)
- Open the browser to the page `http://localhost/login.php` or wherever you installed DANTE.
- Login with default credentials `admin:admin`.
- Follow the configuration instructions and change the default password!

### 1.3 Configuration

The configuration file contained in `config.php` is very simple and permits to configure some basics parameters, such as:

- `db_host` correspond to the hostname of the database (default: `localhost`).

<sup>2</sup><https://danielarancibia.wordpress.com/2015/09/27/installing-apache-2-4-and-php-7-for-development-on-windows/>

## 1 DANTE

DANTE is a project developed by the PHuSe Lab<sup>1</sup> from Università degli Studi di Milano, Italy, in collaboration with Département Informatique de Ecole Polytechnique de l'Université François Rabelais de Tours, France. The tool is part of a multimodal dataset acquired with the aim to study emotional response in presence of amusement stimulus.



UNIVERSITÀ  
DEGLI STUDI  
DI MILANO

### 1.1 Requirements

The annotation tool is a web-based software, and for this reason it simply requires a LAMP/MAMP/WAMP software bundle, depending on the operating system.

<sup>1</sup><http://phuselab.unimi.it/>

- db\_name correspond the database name to use and is the one where was imported the annotationdb.sql file (default: annotationdb).
- db\_user database username.
- db\_pass database password.
- anno\_rate is the recording frequency of the annotation. DANTE will record a value every 1/anno\_rate seconds (default: '25').
- save\_mode DANTE is able to save the annotations to the database 'db' or to a text file 'file' (default: 'db').

If you have some problems with *msqli*'s function see section 4.2.

## 2 The theory behind DANTE

With DANTE project, PHuseLab wants to understand how people feel watching videos with different characters and emotions. The idea is to find a easy and user-friendly tool to annotate precisely the emotions of watchers. First of all, we have to understand what is an emotion and how we can measure it.

## 3 Define and measure emotions

Our emotions play an important role throughout the span of our lives but define "emotion" is a notorious problem.



Aristotle defined emotions as *things on account of which the ones altered differ with respect to their judgments, and are accompanied by pleasure and pain: such are anger, pity, fear, and all similar emotions and their contraries* (ret. 1378a).

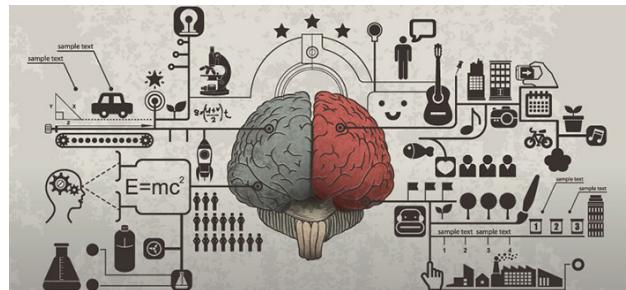
Descartes, in his last published work (1649) defined emotion as "*things which inform us about important changes related to our soul*". Darwin think emotions are a complex expression, that only human being have. So emotions are a result of evolution.

A particularly unfortunate example is William James's asking the question "What is an emotion?" when he really meant "feeling", a misnomer that started a debate which is still ongoing, more than a century later. James think emotion are born in ANS (Automatic Nervous System). James' theory is showed with a

simple example: "you don't run from a bear because you're afraid; you're afraid because you run from a bear".

In recent years, neuroscience has introduced a new way of thinking about our emotions. The scientists behind the latest brain-imaging studies say they can now pinpoint with precision where these feelings are located within our heads. In 2013, for instance, a team of psychologists published a study in which they claimed that they had found neural correlates for nine very distinct human emotions: anger, disgust, envy, fear, happiness, lust, pride, sadness, and shame.

The conclusion is that emotions are difficult to explain but people can try to measure them.



Now let's move on the difference between measuring emotion of ourselves and emotion of somebody else.

To measure emotions of somebody else we can use the Facial Action Coding System (FACS)<sup>3</sup> whom is a system to taxonomize human facial movements by their appearance on the face. It's a common standard to systematically categorize the physical expression of emotions. The Action Units (AUs) are the fundamental actions of individual muscles or groups of muscles. As AUs are independent of any interpretation, they can be used for any higher order decision making process including recognition of basic emotions, or pre-programmed commands for an ambient intelligent environment. The FACS Manual is over 500 pages in length and provides the AUs interpretation of their meaning.

Let's see which are action units:

AU1	AU2	AU4	AU5	AU6
Inner brow raiser	Outer brow raiser	Brow Lowerer	Upper lid raiser	Cheek raiser
AU7	AU9	AU12	AU15	AU17
Lid tighten	Nose wrinkle	Lip corner puller	Lip corner depressor	Chin raiser
AU23	AU24	AU25	AU26	AU27
Lip tighten	Lip presser	Lips part	Jaw drop	Mouth stretch

<sup>3</sup>[https://en.wikipedia.org/wiki/Facial\\_Action\\_Coding\\_System](https://en.wikipedia.org/wiki/Facial_Action_Coding_System)

and how we can describe the most common emotions using them:

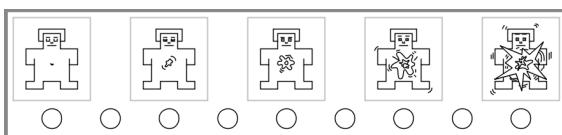
Emotion ↴	Action units ↴
Happiness	6+12
Sadness	1+4+15
Surprise	1+2+5B+26
Fear	1+2+4+5+7+20+26
Anger	4+5+7+23
Disgust	9+15+16

To measure emotions of ourselves we can use different ways. The most common used are:

- Likert scale: it is the scale commonly involved in research that employs questionnaires. It offers a range of answer option from one extreme attitude to another. Typically, it includes a moderate or neutral midpoint. Let's see an example:



- SAM: the main problem of using Likert scale is that "non-English" people don't understand what are values in the scale. So another common scale used is the SAM (the Self-Assessment Makin), in which the values are represented by images. Usually the SAM have 5 values with a neutral midpoint. In DANTE we use the SAM annotation.

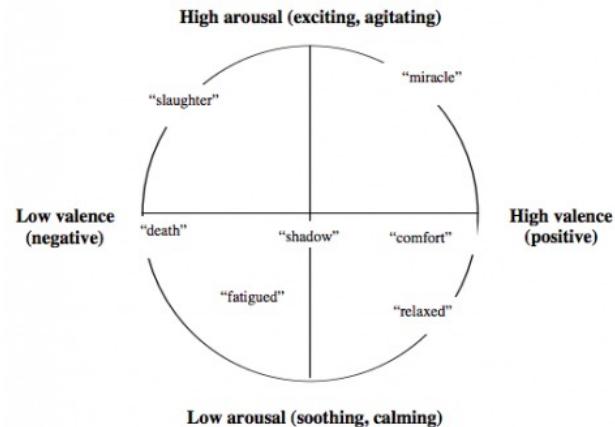


### 3.1 Valence and Arousal

In general, emotional experiences can be described by 2 terms. Though the two terms are often confused, the difference is pretty simple:

- Valence is positive or negative affectivity. Emotion valence could be described by bipolar scale that, in aggregate, defines a continuum from pleasantness (happy, pleased, hopeful, etc.) to unpleasantness (unhappy, annoyed, desiring, etc.)
- Arousal measures how calming or exciting the information is. Arousal comes from our reptilian brain. It inspires a fight-or-flight response, that, evolutionary, aided our survival

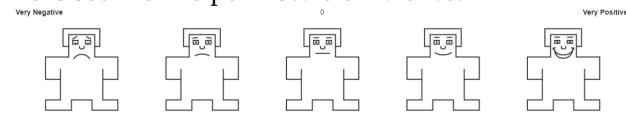
The common framework for dealing with emotional experience is characterized in a two-dimensional space. Valence ranges from highly negative to highly positive, and arousal ranges from calming/soothing to exciting/agitating. Remember this visual:



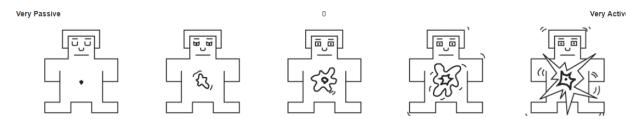
### 3.2 SAM (the Self-Assessment Makin)

The Self-Assessment Makin is a non-verbal pictorial assessment technique that directly measures the pleasure, arousal and dominance with a person's affective reaction to a wide variety of stimuli. SAM is an emotion assessment tool that uses graphic scales, depicting cartoon characters expressing three emotion elements: pleasure, arousal and dominance. SAM has been used often in evaluations of advertisements, and increasingly also in evaluations of products. SAM is based on the PAD emotion model of Mehrabian<sup>4</sup> to describe and measure emotional states using numerical dimension (Pleasure, Arousal and Dominance).

Let's see the five point scale of valence:



and of arousal:



## 4 How DANTE works

DANTE's goal is to annotate how people felt themselves watching other people with different emotions. So it is

<sup>4</sup>[https://en.wikipedia.org/wiki/PAD\\_emotional\\_state\\_model](https://en.wikipedia.org/wiki/PAD_emotional_state_model)

a way to annotate and save the reactions of different people to study the common behaviours.



DANTE's database contains certain videos representing few people (usually only one) in foreground which talk and move. For each video the watcher can annotate valence and arousal values, one per time. On the left side there is a list of video the watcher can choose. Once selected, the video will open and when the watcher plays the video and starts to annotate (by clicking on the toolbar under the videoplayer), the annotator starts to save values until the video is over. If you are annotating, in the top-right side of the videoplayer there must be the "REC" green light switch on, otherwise, if the REC light is red you are only watching video without annotating. To understand if you are annotating you can also control if there is the emoticon in the screen. To not bore watchers, the video are quite short but once started, you cannot stop it until the end.

When the video ends, annotations saved correctly. You can re-annotate the same video re-play it. The old annotated values are deleted and the new one are saved.

In DANTE's website users can choose between two option:

- YOU: where you can annotate your experience using the toolbar and the Emoji or SAM scale. It's what we have discuss about before.
- EXPERT: users can see the "expert" experience, if any. We have used two different python scripts to calculate expert values , the first one calculates the arithmetic average of all annotators. In the other hand, we use the *evaluator weight* (EWE) discussed in Grim and Kroschel<sup>5</sup>:

$$x_n^{\text{EWE},(i)} = \frac{1}{\sum_{k=1}^K r_k^{(i)}} \sum_{k=1}^K r_k^{(i)} \hat{x}_{n,k}^{(i)}$$

We calculate the individual amount of disturbance during the annotations compared with the arithmetic average calculate before. This value is the *evaluator-dependent weight*:

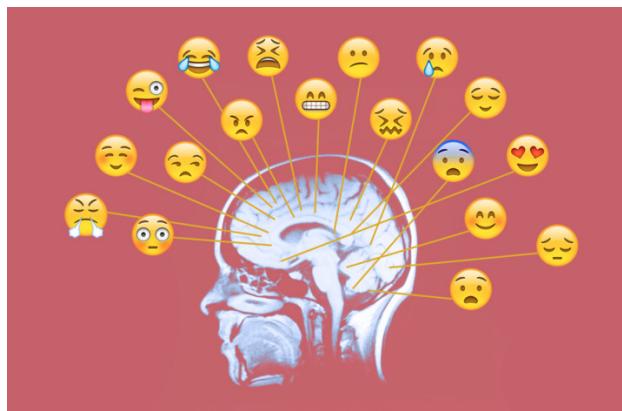
$$r_k^{(i)} = \frac{\sum_{n=1}^N \left( \hat{x}_{n,k}^{(i)} - \frac{1}{N} \sum_{n'=1}^N \hat{x}_{n',k}^{(i)} \right) \left( \bar{x}_n^{(i)} - \frac{1}{N} \sum_{n'=1}^N \bar{x}_{n'}^{(i)} \right)}{\sqrt{\sum_{n=1}^N \left( \hat{x}_{n,k}^{(i)} - \frac{1}{N} \sum_{n'=1}^N \hat{x}_{n',k}^{(i)} \right)^2} \sqrt{\sum_{n=1}^N \left( \bar{x}_n^{(i)} - \frac{1}{N} \sum_{n'=1}^N \bar{x}_{n'}^{(i)} \right)^2}}$$

Thus, it's possible have a more accurate expert than arithmetic average's expert because it's important how the annotators do them job. The scripts assign an "influence" value calculating the expert value, normalize all of them and write a csv file with the EWE's values. This will be the file to get valence or arousal values to represent the right emoji on the screen. So, you can watch the expert annotation without annotate anything and start to have confidence with DANTE.

To create the expert component, we have created a Python script whos for all videos, check if there are Valence and / or Arousal values and use them to evaluate and create the expert in a csv file. Once created the csv expert files, the php class "csv reader" reads server values and passes them to the client using the javascript class "encoding json" whos refresh the emoji values and position. Of course, you cannot change expert and you experiece once the video has started.

#### 4.1 Innovations introduced

Our work is about helping the annotator to understand if he is annotating in a good way. In other words we want to lead the annotator through an immediate feedback on the screen. While the annotator is moving the toolbar to explain his emotions during the annotation phase, he can see a little and transparent emoji near the face of the character on the screen which represents the emotion he is annotating. We have tried to modernize the SAM introducing a common feature of recent years: the emoji (or emoticon). People when text a message try to explain his emotion through a mobile phone or a personal computer. The emoticons are a good way to try to do it. We do the same, instead of the classical SAM example we have used a sequence of emoticons.



<sup>5</sup><http://robotics.usc.edu/~mower/Papers/GrimSpeechComm2007.pdf>

# DANTE

## Dimensional ANnotation Tool for Emotions

So in the new DANTE model, the watcher can choose between the classic SAM annotator and the new one with emoticons.

Let's see how are valence emoticons:

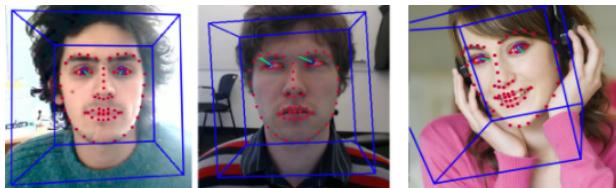


and arousal emoticons:



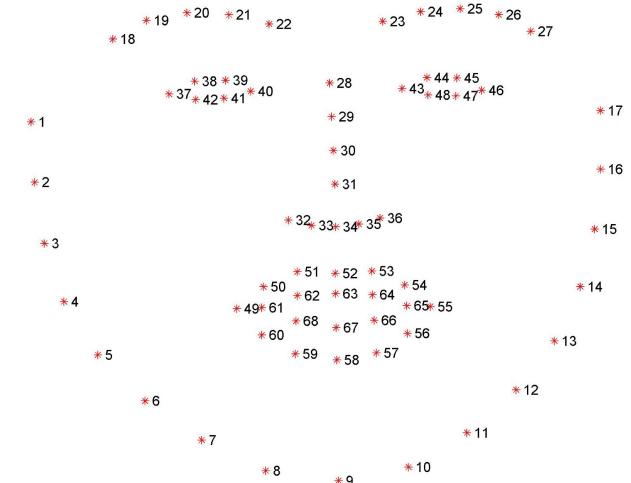
To create the emoji which follow the head of the video's character, we used the open-source software OpenFace<sup>6</sup> to create a huge csv file with landmarks and other features of character. To understand this step, let's move on what are landmarks and other features. OpenFace software, create a csv file with:

- frames number
- timestamp
- confidence
- gaze direction and angle
- 56 eyes landmarks in 2D (pixels) and 3D (millimeters) coordinates
- 68 face landmarks in 2D and 3D coordinates
- some Action Units



To make the emoji follow the character face we have written a python script which extract from the huge csv file:

- frames number
- timestamp
- x\_1, x\_17, y\_9 and y\_22 2D face landmarks



The python script control which OpenFace files have not already been reduced and reduce them. We extract only these 4 points because we want to understand where the face is, so we create something like a bounding box where the face is included. We extract the frame and timestamp from the videoplayer using a github repo<sup>7</sup> and comparing these data with the OpenFace return values, we put the emoji always closed to character face. The emoji is transparent to not disturb the watcher but is useful to give an immediate feedback about what he is annotating.

## 4.2 Problems occurred

A problem we have occurred is when we decrease the browser window's size. We cannot know the screen resolution or browser's window dimension to resize the emoji.

If you have downloaded the github repo and during the configuration have a fatal error with an undefined mysql's function you have to rename the `php.develop` in `php.ini` and modify it adding the extension `extension=mysqli.dll`.

## 5 Possible future developments

We have used Big Brother's videos because we thought that the big brother's competitors inside the confessional can show a wide variety of behaviours and emotions in front of nonexistent problems. So a future possible development will be extend the variety of database's videos adding some more cultural content like film clips or something like that.

Another future development can be increase the variety of psychometric scale like Likert's to have more people annotating as possible.

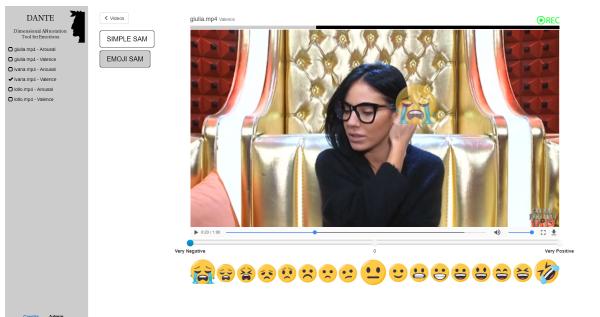
<sup>6</sup><https://github.com/TadasBaltrusaitis/OpenFace>

<sup>7</sup><https://github.com/allensarkisyan/VideoFrame>

## 6 Conclusion

The repo of DANTE is uploaded on github at link  
<https://github.com/phuselab/DANTE>.

This is the new version of DANTE with the emoticons' new features:



This has been a useful and funny social experiment to understand, with people's help, how they react watching emotions of others to try to define what emotions are and continue the evolution of the definition of emotions starts with Aristotele until the recent psychologists' definition. So we would like to define emotions starts with the common features and Action Units. DANTE is the training set to arrive to the goal.