p-ISSN: 2184-2205 e-ISSN: 2184-3414 ISBN: 978-989-35106-0-5 © 2023
https://doi.org/10.36315/2023inpact146

NON-BIASED CFACS MEASUREMENT TOOL: FROM IDEA TO SOFTWARE APPLICATION

Mikhail Baev[1], Alexey Gusev[2], & Alexander Kremlev[2]
[1]LLC Uchet-N, St. Petersburg (Russia)
[2]Faculty of Psychology of Lomonosov Moscow State University (Russia)

Abstract

We would like to draw the readers' attention to our method of precise facial expressions analysis. Our research pursued two main objectives: (1) developing the methods and algorithms for the direct evaluation of facial expressions using computer vision (CV) technology; (2) creating software for conducting empirical research in psychology and related subjects. We have developed an authentic implementation of cFACS (Rosenberg, Ekman, 2020) as an approach to computer facial expressions analysis. We would like to stress that at present, the neural network approach is de facto dominant. The approach is based on the basic emotions concept, and on the assumption that a neural network can be trained to recognize emotions using a respective set of pictures, or, more rarely, videos. In our opinion, implementing this approach is limited (Baev et al., 2021). We have applied a comprehensive approach to the analysis of facial activity, based on detecting and analyzing time-defined FACS action units' (AUs) combinations. The chosen approach has obvious advantages: 1) employing AUs as facial activity analysis units; 2) the feasibility of assessment of individual features of facial activity; 3) racial and age biases becoming irrelevant. While developing the software, two major methodological principles have been used: 1) the direct assessment of facial surface movements aiming to detect AUs, coupled with the deliberate rejection of employing neural networks for facial events classification; 2) modelling an expert's perception of facial surface movements' peculiarities while detecting certain AUs, taking into account the general topography of facial surface movements. We have developed the EmoRadar software, using our proprietary CV procedures that allow us to evaluate facial surface changes in the areas corresponding to different AUs. In the process, there was created a system of rules, based on which the "raw" data on lighting changes transform into surface movements, and those, in turn, convert into AUs, and, further on, into basic emotions and patters of facial activity. In the course of our software's empirical verification, the method of automated analysis of job interviews was created. We have also performed automatic evaluation of video recordings for forensic psychological examinations and analyzed participants` emotional dynamics, participating in the "SIRIUS-21" project of the Institute of Biomedical Problems of the Russian Academy of Sciences. We believe that the implementation of the above-mentioned principles of the direct approach to facial activity registering and the comprehensive approach to its analysis provides tangible advantages:
• High accuracy of facial surface movements assessment.
• Feasibility of conducting all kinds of facial analysis (basic emotions, basic emotions emblems, facial activity patterns) based on reliable allocation of 22 main AUs.
• Complete absence of racial bias.
• Feasibility of assessing children's facial activity (up to 5-7 years old).
• Independence from theoretical approaches to the facial expressions analysis (P. Ekman, K. Scherer, L. Feldman Barrett, H. Oster).

Keywords: *FACS, emotions, affective computing, facial expressions, computer vision*.

1. Introduction

In the psychological studies of today, a toolkit for automatic analysis of a person's emotional state by his/her facial expressions is becoming more and more acclaimed. In our opinion, somewhat of a paradox has emerged: each year the volume of videos increases, while no software, allowing to reliably assess facial expressions on an expert level, has been developed. Specialists able to professionally code mimic activity and categorize various facial expressions as emotional states' manifestations, are critically few; given this deficit, we observe an obvious lack of software intended for research and empirical work that require exact and objective assessment of an emotional state of the person whose face appears in the video.

We believe that the quality of the existing software is, for now, rather low, and the hope that neural networks will soon learn to do that, is yet to be fulfilled. Our own research, and our fellow psychologists' experience have proved that the work by the leading world companies in the sphere of affective computing, such as Microsoft Azure, Amazon Rekognition, and Noldus FaceReader, does not provide the required precision and reliability of the automatic assessment of either the basic emotions, or separate Facial Action Coding System (FACS) Action Units (AUs).

Thus, we present the results of our own software development for precise and non-biased facial expressions analysis, and the preliminary outcomes of the empirical research carried out with the help of our software. We also propose a substantiation of the approach used in the process of our software development.

We have applied cFACS (computer FACS) (Ekman, 2020) in the course of facial expressions analysis based on our CV procedures. We'd like to emphasize that the neural network approach is currently de facto dominant. It is based on the discrete emotions concept and an assumption that consistently applying a training set of pictures it is possible to train a neural network to recognize emotions. In our opinion, this approach has limitations (Baev et al., 2021).

2. Substantiation of our approach

Our approach to mimic activity assessment is based on the FACS, created by P. Ekman and his colleagues (Ekman, Friesen, Hager, 2002), which was widely recognized as a method of universal description of facial expressions. The base element of the analysis by FACS is an Action Unit (AU), understood as the characteristic change of the face surface in the form of a single movement of, e.g., lips, brows, eyelids, nasolabial folds. A certain combination of AUs allows to single out an expression of seven basic emotions (Ibid.). It should be underlined that FACS does not define the theoretical base for singling out and assessment of the basic emotions.

Now in the sources there is a discussion on what emotions are reflected in the facial expressions, and how they are expressed (Rosenberg, Ekman, 2020, Davidson, Scherer, Goldsmith, 2003, Feldman Barrett et al., 2019); thus, we focus on the fact that FACS and its computer adaptation, cFACS is the system that can be used to detect expressions of emotions and other facial activity patterns on the face, regardless of the way different authors understand the task of detecting emotions on a person's face. The latter is especially important, because even professional psychologists quite often blend these two problems, i.e., *describing* facial expressions, and *categorizing* them as emotional expressions.

In the sources, there are two main approaches to the solution of the task of automated analysis of facial expressions presented on a video recording (Rosenberg, 2020). The *selective* approach is based on establishing correspondence between the facial expression seen in the video, and a sample of basic emotions and/or AUs, beforehand annotated by experts.

Its peculiarities and flaws are as follows:

1. It applies a neural network trained by a limited, and, as a rule, low-variant selection of the images of people's faces;

2. Because the training datasets are limited, fast facial expressions of the face (microexpressions) are either detected badly, or not at all;

3. To train a neural network, one needs datasets of images of the faces belonging to people of various sex, age and ethnicity;

4. A manual tuning of the sensitivity of triggering of the used neural network algorithms is required.

The *comprehensive* approach is based on detecting and analyzing time-defined FACS AUs combinations. Its peculiarities are as follows:

1. Using FACS allows to describe all the possible facial expressions;

2. Feasibility of assessing asymmetrical facial movements and microexpressions of emotions;

3. Feasibility of assessing the individual variability of facial expressions, i.e., instances of hyperkinesis, mannerisms, mimical dialects;

4. It is free of racial bias;

5. While analyzing a video recording, one assesses the individual dynamics of facial expressions of a certain given person, which is why there is no need to tune the computer system in accordance with the specifics of the task in question;

6. There are no age limitations, because it is AUs emergence that is described, but not emotions.

The algorithms of the facial expressions computer analysis that we have developed, are based on the comprehensive approach and the following principles:

1. The direct assessment of the facial surface aimed at detecting AUs and deliberate rejection of employing neural networks for facial events classification;

2. Modelling an expert's perception of facial surface movements' peculiarities while detecting certain AUs.

To implement the comprehensive approach, we have developed authentic CV procedures, that were specifically focused on the movement of certain parts of the facial surface and characteristic peculiarities of various AUs' emergence.

To detect various AUs, we have applied the principle of multi-layer analysis of mimic activity:

• Within the frames of the video, analyzing the change of light distribution over the various segments of the face, singling out the meaningful changes and rendering them on the facial surface as bulges, pouches, furrows and wrinkles appearing;

• Matching the singled-out changes with each other in accordance with the anatomical structure of the face and the FACS scheme;

• Defining the begin and end of certain AUs;

• Singling out the emotions and other complex mimic patterns;

• Calculating the integral indexes of facial activity, e.g., frequency of AUs per time unit, valency etc.;

• Complex categorization of facial activity patterns, i.e., true emotions, emblems of emotions, mannerisms, singling out the significant speech fragments, etc.

3. The software empirical testing

The EmoRadar LS software has been empirically tested in a number of applied studies.

3.1. Personnel assessment during a video job interview

Employees of large industrial companies (413 men and 242 women, aged 42.6 in average) remotely participated in a structured 6-16 questions video interview. 4038 videos were selected for the analysis. Individual AUs, basic emotions and facial expressions patterns had been detected automatically.

Based on the data from the sources and our own experience, we have constructed the following facial activity patterns: Fear/Stress, Tension, Displeasure, Negative Attitude in communication.

Following the criteria proposed by HR specialists, six rules for automatic analysis were developed. The rules were based on the combinations of different AUs patterns and characterized the respondent behavior while answering an interview question as corresponding/inconsistent with the expert's expectation. Sixty videos were selected for comparative analysis. These videos were evaluated by six experts on the following scales: inclusiveness, stress, confidence, lightness, activity, strength. The obtained results show good agreement between the results of the computer facial expressions analysis and expert assessments. Thus, the possibility of practical use of the original technology of facial expressions analysis for evaluating of a video job interview has been confirmed.

3.2. Forensic psychology

Forty-three office employees of a large Russian company were being tested for involvement in fraudulent activity. The security service asked them to answer the 15 questions of a standardized video interview. The respondents were told that the aim of the interview was clearing the circumstances of committing offenses, and their working relations with the employees who had been officially accused.

Based on the facial activity patterns, characterizing outward expression of negative emotions, all the videos were assessed according to the level of the risk of the possible involvement in the crime:

1. Risk level 1 (low possibility of the involvement) marked the videos with the minimal indexes of coping with stress, absence of the negative attitude in communication, minimal signs of experiencing danger and frustration. In this group, there were 62.8% of the videos.

2. Risk level 2 (there is a possibility of involvement, additional research is recommended) marked the videos with medium to low indexes of coping with stress, presence of the negative attitude in communication, experiencing danger and frustration. In this group, there were 27.9% of the videos.

3. Risk level 3 (the possibility of involvement is high) marked the videos with extremely high, or high signs of stress, with the negative attitude in communication simultaneously discovered, as well as frustration and discontent. In this group, there were 9.3% of the videos.

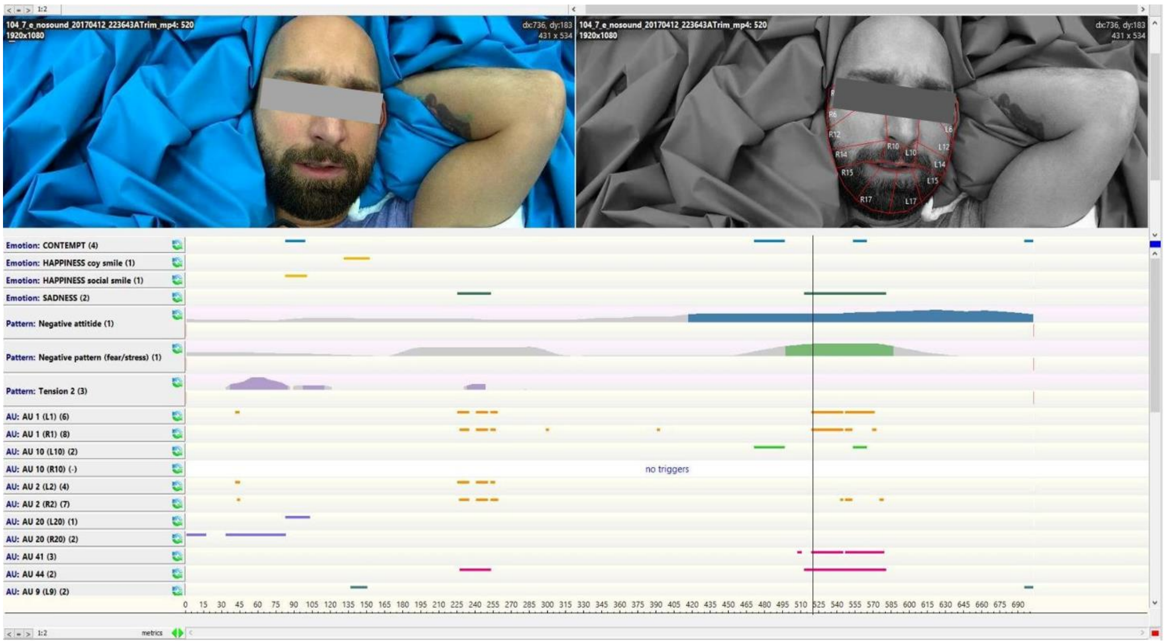
Here is a sample fragment of an expert's assessment concerning the respondent A.K., whose video interview was marked as Risk Level 3: "Obvious negative emotional reaction to the introductory part of the interview [is demonstrated]. In the course of the interview, high-level emotional tension is seen.

[The respondent shows] clear difficulties in coping with stress and controlling his own behavior. The facial activity is intensive and non-coordinated. An emotion of fear was found at the time of the question on his attitude to the fact of his colleagues' detention. There are stereotypical movements of the tic type. An additional polygraph examination is recommended."

3.3. Aerospace psychology

Within the SIRIUS-21 experiment, we studied the change of the participants' emotional state influenced by lower gravity. The paradigm of the so-called "dry immersion" imitating the state of relative weightlessness, where a person is immersed in water without immediate contact with the liquid, was applied. For 7.5 days, in the morning and in the evening, 10 participants recorded a brief video interview about their state of health and emotion. With the help of EmoRadar LS, the basic emotions and above-mentioned facial expression patterns were assessed. *Figure 1* demonstrates a screenshot of EmoRadar LS interface presenting the results of facial expressions' analysis of Participant 104 in the evening of the seventh day of the experiment. The vertical line (*Figure 1*) is used to mark a complex time-defined combination of various facial events, i.e., Sadness and negative AUs patterns of the Negative Attitude, Fear/Stress. We emphasize the emergence of AUs 1, 10, 14, 41, 44, corresponding to these patterns, within the area of the marked frame.

Figure 1. EmoRadar LS software screenshot. There are different indexes of facial expression patterns, emotion expressions and AUs on the timelines under the participant's photos. The lower line shows the frame number in the video.



4. Conclusion

Our technology of facial expression analysis allows for high-precision assessment of a person's face videos, detecting within them the outer expressions of emotional state change. We see the perspectives of our research in the integration of facial activity with the results of voice analysis, psychosemantic analysis of speech and the psychophysiological studies data.

References

Baev, Mikhail, et al. (2021). Unbiased Mimic Activity Evaluation: F2F Emotion Studio Software. IEEE, 2021 9th International Conference on Affective Computing and Intelligent Interaction (ACII), https://doi.org/10.1109/ACIIW52867.2021.9666319

Barrett, Lisa Feldman, et al. "Emotional Expressions Reconsidered: Challenges to Inferring Emotion from Human Facial Movements." Psychological Science in the Public Interest, vol. 20, no. 1, 2019, pp. 1–68., https://doi.org/10.1177/1529100619832930.

Davidson, R. J, Scherer K. R., Goldsmith H. H. (Eds.). (2003). Handbook of Affective Sciences. Oxford: Oxford University Press. (Journal Articles with Continuous Pagination format:) Last Name, Initials of Author, & followed by last names and initials of other authors if any. (Year of Publication). Title of article. *Title of Journal*, *Volume*, Page Numbers.

Ekman, P. (2020). FACS: Yesterday and Today. In Rosenberg, E. L., Ekman, P. (Eds.), What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS) (pp. 614-618). New York: Oxford University Press.

Rosenberg, E. L., Ekman, P. (Eds.). (2020). What the Face Reveals Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS). New York: Oxford University Press.

Rosenberg, E. L. (2020). FACS in the 21st Century. In Rosenberg, E. L., Ekman, P. (Eds.), What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS) (pp. 1-22). New York: Oxford University Press.