Detecting True and Declarative Facial Emotions by Changes in Nonlinear Dynamics of Eye Movements

Albert Śledzianowski[1(✉)], Jerzy P. Nowacki[1], Andrzej W. Przybyszewski[1], and Krzysztof Urbanowicz[2]

[1] The Faculty of Information Technology, Polish-Japanese Academy of Information Technology, Warsaw 02-008, Poland
albert.sledzianowski@gmail.com, przy@pjwstk.edu.pl
[2] Move2Mind, Boston, MA, USA

Abstract. In our previous work we have showed that we can improve classifications of facial emotions (FE) by extending a dataset with chaotic dynamics parameters of eye movements (EM). This time we wanted to confirm our results using public and independently created video sources and for this purpose we chose the Affectiva-MIT Facial Expression Dataset (AM-FED). Our purpose was to find out whether we can estimate Happiness through non-linear dynamics of EM also in independent video data. We have calculated EM chaotic dynamics in video recordings of the AM-FED database and performed estimations of Happiness calculated with parameters provided by the Open Face library (OF). We also calculated correlation between these parameters and parameters attached to the AM-FED database using our own method based on sliding windows and proposed a method of using its output parameter with a short algorithm. We have observed that true Happiness was connected to a moderate value of negative correlation with EM chaotic dynamics in the case when smile was not present, while for declarative "Smile" parameters we observed a moderate positive value. By using EM chaotic dynamics correlation we have estimated the difference between posed smiles and true Happiness with the XGBoost classifier, with accuracy results of 0.75 (ROC-AUC 0.9) and precision of 0.8 (tested with dataset of 0.33). We are proposing EM chaotic dynamics parameters as an extension for estimations of Happiness based only on facial muscles activity. We think that this approach can confirm the authenticity of Happiness in various cases and also introduce the distinction between real and declarative FE into Computer Vision. It also can bring solution in cases when lower part of the face is hidden, i.e. when is covered by a protective mask.

Keywords: Machine learning · Eye moves · Smile · Happiness · Facial emotions · Mimicry · Emotion authenticity · Noise · Chaos · Nonlinear dynamical systems · Eye move chaosity · Eye movement chaosity

N. T. Nguyen et al. (Eds.): ACIIDS 2022, LNAI 13757, pp. 106–116, 2022.
https://doi.org/10.1007/978-3-031-21743-2_9

1 Introduction

There is lack of publications related to the EM chaotic dynamics during different emotions, especially in the context of their use in estimation of the facial emotions. This is the continuation of our research on the complex dynamical system describing EM presented in our previous article [1], where we have proved a positive, statistically significant correlation between the value of EM chaotic dynamics and the intensity of the Happiness (as FE).

In our previous experiment we have simultaneously recorded FE and EM in 49 subjects and analyzed responses to the video stimuli with Happiness and Contempt [1]. We have described EM by nonlinear dynamical system and observed that Happiness was positively correlated with EM chaotic behavior, while both the linear and the noise components were mostly negatively correlated with this FE [1]. In the case of our data, we have observed that while Happiness intensity increases, the EM becomes more chaotic and behaves less noisy [1].

As a continuation of our research project, we wanted to test our findings on public domain videos with FE classified by Facial Action Coding System (FACS) [2]. For this purpose we have chosen the "AM-FED" database created by researchers of the Affectiva Inc, the MIT Media Lab and the Robotics Institute of Carnegie Mellon University, providing both human ("Smile") and algorithmic ("Smile V2") types of classification, showing the faces of people watching amusing video clips [3]. We wanted to compare the differences in EM chaotic dynamical properties between people who show the true (Happiness) or only the declarative FE (Smile). We have used the OpenFace (OF) library for Action Units (AU) estimations and for estimations of gaze vectors used to calculate the chaotic dynamics of the EM [4]. We used the FACS algorithm for Happiness estimations in the AM-FED videos. For Smile parameters, estimations have been provided with this database [3].

Seeing is an active process. Our eyes are continuously moving and searching the environment in order to capture interesting objects within the fovea which is the high resolution region of the retina. As these interactions are complex, EM must response with fast complex movements. Therefore, describing the EM behavior as a nonlinear dynamic system seems natural. The EM changes may have different attractors that determine their behavior in time. As we have previously demonstrated, in natural conditions EM behaviors are dominating by chaos or by noise. Chaotic behaviors were described in different parts of the human body: in the heart rate [5], in the brain activity recorded by EEG and described by fractal dimension or in the gait kinematics of healthy subjects that have chaotic properties [6] which are lost in Parkinson's diseases [7]. However, still not many researchers describe EM by nonlinear dynamics, even if processes in the retina show a complex, multi-attractor behaviors [8].

E. Paulson discussed the reading processes and concluded that it can be described as a self-similar, nonlinear dynamical system dependent on the reader and text characteristics that can be explained by the chaos theory principle. Chaotic dynamics can explain difficulty in predicting the nature of a reader's eye movement regressions [9]. K.M.Hampson and E.A.H. Mallen observed that the eye aberration dynamics are chaotic and they characterized it by using techniques from the chaos theory by measuring the monochromatic aberration [10]. W. Richards et al. have used the EM data collected from perceptual tasks: binocular rivalry, fixation patterns during search, simple multi-stable percepts, and perceptual segments in several movies suggested that the mechanisms underlying our percepts might be modeled as nonlinear, deterministic systems that exhibit chaotic behavior. The eye scanning strategy appears to be controlled by nonrandom, dynamical spatial representation, but not a temporal one [11]. Harezlak et al. classifies EM as a signal exposing features characteristic of the chaotic natures, simultaneously emphasized that obtaining confidence in differentiating chaotic and noise behaviors requires the application of various approaches [12]. S. S. M. Chanijani analyzed 3 types of students and concluded that information entropy of EM is higher for novice people. C. Astefanoaei et al. found properties of chaos in the saccadic EM temporal series collected from a healthy subjects by estimating the correlation dimensions [13,14]

In our work, we have compared results of Smile and Happiness classification in context of the EM chaotic dynamics. We propose to use the EM chaotic parameters as an additional data in the method of FE classification. In particular, we were interested if EM chaotic dynamics can improve estimations of authenticity of expressed Happiness and if it can help to solve the common mistakes in the automated FE classifications, like mistaking a worry face of pain with a smile.

2 Methods

The AM-FED database consists of 242 facial videos (168,359 frames) recorded by web-application at the frequency of 14 FPS and with the resolution of 320×240 in a real user's conditions. The viewers watched one of three intentionally amusing videoclips [3]. All AM-FED's videos were estimated by certified FACS coders and by automated facial expression analysis [3]. Each video was labeled frame-by-frame for the presence of AUs, by at least 3 FACS trained specialists (chosen from 16) [3]. The AM-FED coders labeled activity of AU2, AU4, AU5, AU9, AU12, AU14, AU15, AU17, AU18 and AU26. This database also provides its own index for the "Smile" and according to the authors, smiles were labeled and in this dataset are distinct from the labels for AU12. The AU12 may occur also in an expression that would not necessarily be given the label of smile (e.g. a grimace) and this is why they labeled for presence of the "smile" rather than AU12 [3]. Researchers agreed that AU activity is present in case of consent of over 50% of the coders and assumed that a label is not present if 100% agreed (the mean percentage agreement across classifications was 0.98) [3]. This database also includes the results of the baseline estimation for smile ("smile V2") using an algorithm based on landmarks detection and tracking, Histogram of Orientated Gradients computed for the face regions and classification based on the Support Vector Machine with the RBF kernel [3].

As previously mentioned, we used the OF library to extract from AM-FED videos the information about eye directions expressed as the gaze vectors (x,y,z) and also to detect, set and track landmarks describing the shape of face elements in the video and evaluate the presence and the intensity of specific AU [1,4].

We used this data to estimate Happiness, first by calculating its presence (HP) as:

$$HP = P(AU06) \cdot P(AU12) \tag{1}$$

and next, for non-zero presence outputs calculating its intensity (HI) as:

$$HI = \frac{I(AU06) + I(AU12)}{2} \cdot HP \tag{2}$$

where P stands for a function which expresses the presence of emotion in AU, I - function which expresses intensity of given AU, n is the number of AU's defined for happiness.

Our methods of the EM chaotic dynamical analysis described in our previous article calculates 3 parameters: chaos, noise and linear [1]. The EM chaotic dynamics parameters were calculated in the window which size was determined experimentally and then we calculated Pearson's correlation coefficient for the AM-FED's "Smile", "smile V2" and for "Happiness. We used a simplified model to calculate the noise level, which does not require correlation entropy [15]. We basically counted the average line in the recurrence diagram, which must be greater than 1. In our case, the noise level was in the range of NTS = up to 50% i.e. the standard noise deviation is 50% of the standard deviation of the data. From the equation in previous noise estimation article [15], to estimate 50% of the NTS, we need an average line of length <n> and from this we calculate minimal number of data Min(no-data) as stated in Eq. 3.

$$\begin{aligned} \langle n\rangle(NTS = 0.5) &= \frac{2 - 0.5^p}{1 - 0.5^p} = \frac{2 - 0.5^{0.3441717}}{1 - 0.5^{0.3441717}} = 5.71 \\ Min(no_{data}) &= e^{5.71} = 301 \end{aligned} \tag{3}$$

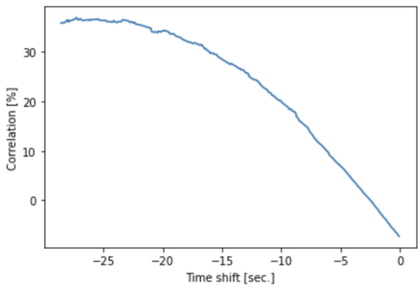
From the Eq. 22 presented in the article on Noise-level estimation we can calculate the minimum value of parameter p = 0.3441717, so the minimum amount of data to estimate 50% of the noise level is 301 items [15]. The standard deviation and the mean change over time in our non-stationary data, thus we needed the window as small as possible but larger than 301, so that the stationarity within the window will be maintained to minimize the error.

We also wanted to keep the window data as stationary as possible, because the noise estimation error is also due to the non-stationarity. In order to reveal the optimal window size we calculated Noise, Linear and Chaos separately for different widths and the optimal window size of the stationary data was around 400 items. In our previous article we performed analysis to make sure that the number of points is not too small, thus we calculated cross-correlations to different reference window widths between 300 and 1000 items. After this analysis we concluded that the window of 300 points was too small, but width of 400 brought statistically significant results and showed the optimal reference point for the entire estimation in the center of the window [1].

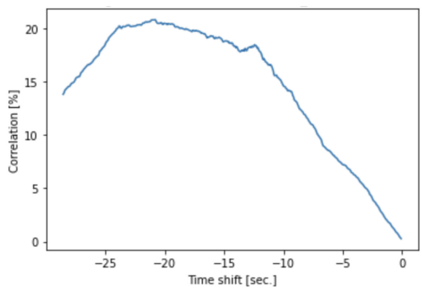
In our dataset we also wanted to include different correlations and provided it by recreating correlation data through the dependency between EM chaotic dynamics and chaos means shifted to different positions in the window.

3 Results

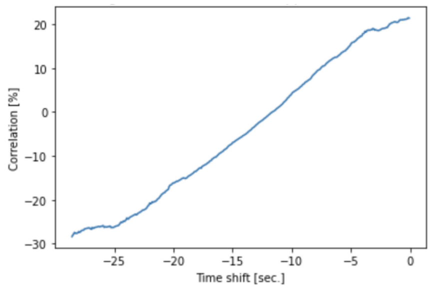
As previously mentioned, we calculated the EM chaotic dynamics parameters basing on gaze vectors time series obtained from the AM-FED videos. Then we calculated the Pearson's correlation coefficient between EM chaotic dynamics parameters shifted to the center of the window and AM-FED's "Smile", "Smile V2" and "Happiness" calculated on the basis of the OF's AU classifications. Figure 1 presents results of this correlation calculations performed on averaged data from AM-FED videos, where the p-value of correlation calculated for a particular data was <0.05. For Smile it was 161 AM-FED videos, for "Smile V2" 240 and for Happiness 190. The total number of all video frames with calculated p-value <0.05 was 317783. As we can see on both Figs. 1a and 1b a smile is accompanied by a positive correlation with the chaotic EM, while in the case of Happiness we can see negative correlation (while "Smile" and "Smile



(a) "Smile (AM-FED)"



(b) "Smile V2 (AM-FED)"



(c) "Happiness (OpenFace)"

Fig. 1. Average correlations of EM chaotic dynamics

V2" are simultaneously not present). Because we assumed that both "Smile" parameters could simultaneously occur with Happiness, we calculated this factor and speculated that Happiness could occur in 65% of smile frames (73008 of AU06's non-zero intensity for 111708 non-zero intensity data frames of AU12).

Just to remind, according to the AMFED database authors, smiles are distinct from the labels of AU12, as AU12 may occur also in other expressions, like a grimace. For both Smile and Happiness the opposite direction is visible when analyzing the correlation results with the chaotic EM (while maintaining similar low-average levels). In our opinion, this parameter distinguishes between posed smile and true FE of Happiness. The chaotic EM is visible for all data positively correlated with Happiness, negative correlation is only visible for Happiness not confirmed by results of "Smile" estimations.

This observation was earlier confirmed with very similar positive statistically significant correlations levels [1]. Therefore, we propose to introduce the chaotic EM parameters into the FACS-based automated methods of Happiness estimation along with a decision mechanism that would determine whether the intensity of AU06 and AU12 can be classified as true or not. The resolving mechanism could be based on short algorithm realizing the following formula:

$$\begin{gathered}[AU12]Activity = TrueSmile = \\ [AU06]Intensity \cdot [Correlation(EyeChaos, [AU06]Intensity) > 0]\end{gathered} \tag{4}$$

Figure 2 presents proposed model modifying to the standard FACS approach by adding parameters of chaotic EM to the estimation of the emotion. The final decision should be based on previously described window size estimation, method of windows centring and calculation of correlation coefficient between EM chaotic dynamics and intensity of Happiness.

We tested this new approach on a simplified dataset and for this purpose we used both Smile parameters and additionally: happiness intensity, centered EM chaos mean and EM chaos mean shifted by different thresholds (including different multiplications as described in the "Methods"). We normalized dataset with the min-max approach and as classification target we chose "Smile" or "Smile V2" binarized for zero and non-zero intesities. We tested estimations on our data with the XGBoost classifier, which might be described as a decision tree with gradient boosting optimization [16].

When we were building the dataset for estimations, we first checked if classification of smile parameters by Happiness data was random and after this confirmation, we started to add additional EM chaos parameters. Interestingly, when we started to do it, the accuracy started then to rise to a level of 0.78. Fig3 presents confusion matrixes for both types of smile estimation and for both types of dataset (with and without the chaotic dynamics data).

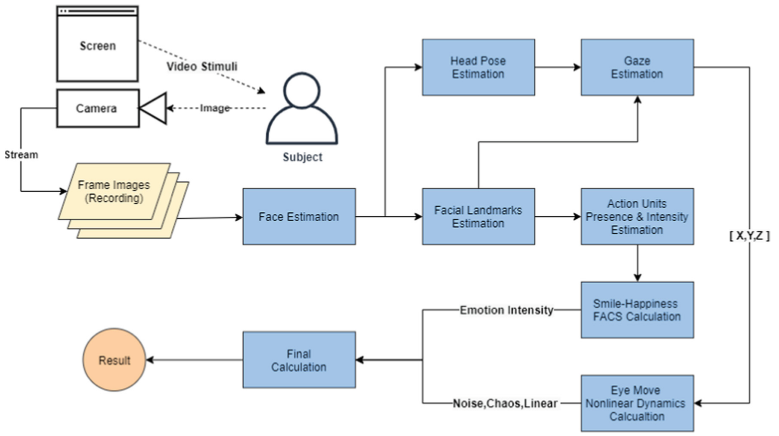
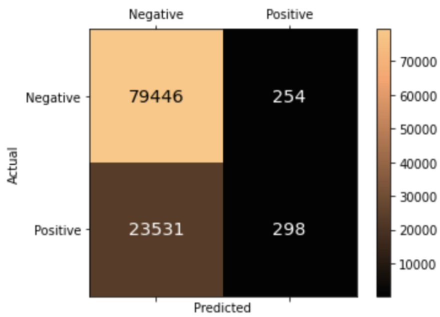
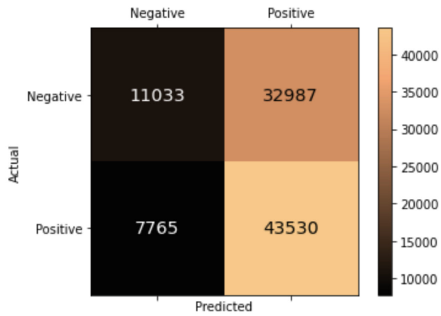


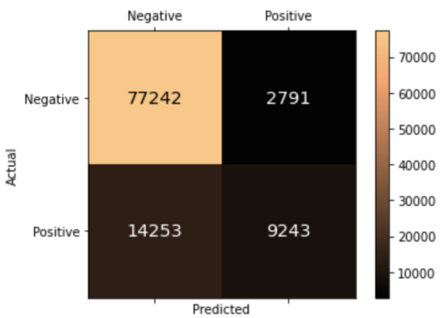
Fig. 2. Proposed improvement to Happiness estimation model basing on FACS and EM chaos calculations



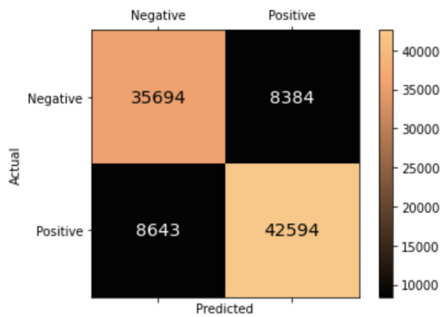
(a) "Smile"



(b) "Smile V2"



(c) "Smile and parameters of EM chaos"



(d) "Smile V2 and parameters of EM chaos"

Fig. 3. Confusion matrixes.

For a test dataset of 0.33 of the data and the XGBoost's classification we achieved "Smile" accuracy of 0.76 (ROC-AUC = 0.86) with precision of 0.44 and "Smile V2" accuracy of 0.74 (ROC-AUC = 0.93) with precision of 0.80.

The confusion matrixes show a significant difference between the results of classification of "Smile" and "Smile V2". We think that it is easier for an algorithm to estimate the results of another algorithm and this could be a reason for this differences. The results of human analysis could be more subjective and based on observations that are intuitive for humans, including complexity of negotiations carried out by human coders. This aspects could make classification more difficult for a machine learning methods, using this set of data.

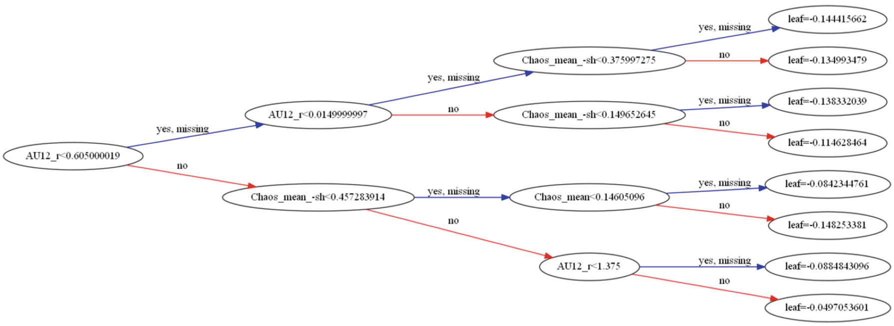


Fig. 4. The plot of decision tree for "Smile" attribute with Xgboost's regression and 2 levels of depth.

Figure 4 presents the decision tree generated for the "Smile" attribute with the Xgboost regression (XGBRegressor) with 2 levels of depth and the dataset used for classifications. We used different methods for visualisation, but we think that regression have a greater degree of freedom and thus are more reliable and conveys the best tree information. As we can see, two parameters are taking major part in the estimation process: the mean of EM chaos shifted to the window center ("Chaos_meansh") and the intensity of AU12 ("AU12_r"). However, at the first tree node the AU12 is taken into account only at very low intensity (<0.01) and later correlation between 'Chaos_meansh' and "Smile" is well visible. If 'Chaos_meansh' is small, then "Smile" is also low and if is greater than the threshold, the "Smile" value is greater too. From this decision decomposition we can see that EM chaos plays a more important role than the intensity of AU12.

4 Discussions

It should be noted, that both Smile and Happiness were estimated from the same video recordings, but partly from different time series when only AU12 (Smile) or both AU06 and AU12 are active (Happiness). Stewart, Bucy, and Mehu by using FACS Action Units definitions differentiate between Posed Smile (also called “non-felt” or false when only AU12 is active - lip corners pulled up and at an angle only) and true Happiness or Enjoyment (also called “felt” or “Duchenne”) when both AU12 and AU06 are active (lip corners pulled up and at an angle, muscles surrounding the eyes contracted) [17]. We think that this approach has a reference to the differences between AM-FED smile parameters and Happiness calculated over the OF’s estimations.

Happiness, besides the AU12 activity, is additionally accompanied by activity of the AU06. At the same time, this activity (AU06) seems to be connected to the change in correlation (from negative to positive) between the EM chaotic dynamics and the intensity of this emotion.

Researches so far differentiated between “Smile” and “Happiness” through the AU activity and postulated various emotional bases and social functions for each type of smile. As smiles are varying in relation to the social context, they have different consequences for observers and different diversity of expression [17].

Stewart et al. listed 5 different kinds of expressions: posed, enjoyment, amusement, controlled and contempt smiles. They might be identified by morphological characteristics of the face, like the direction of lip corner pulling, different muscular “controls” in the lower face that influence the shape of the mouth and by co-activation of muscles surrounding the eyes [17]. By distinguishing in the AU06 activity, authors found differences between posed smile and face expression related to the Happiness. According to authors, the most prominent signal of smiling is the pulling of lip corners, prototypically up and at an angle by the zygomatic muscle (AU12), but the activation of the orbicularis oculi muscle (AU6) is the most common facial component that could reflect the pleasantness of an emotional experience as well as its intensity and authenticity [17].

Posed smiles occur when an individual attempts to, either just signal positive emotion or to mask other emotions. In both cases the AU06 is generally not contracted [17]. In contrast, enjoyment smiles in addition to activation of AU12, activates also the upper face AUs and this way regulates the eye aperture and reinforces the impression in the recipient, that the smile was “felt”. This is why enjoyment smiles have been strongly associated with feelings of amusement and happiness, as well as behaviors like cooperation. It might lead to “facial feedback” as the attribution of emotion by a viewers will be affected by specific facial display morphology [17].

We might say, that posed smiles might not be as strongly associated with the felt emotion of happiness as enjoyment smiles. In our study, we found that they are additionally accompanied by a change in the correlation between the chaotic EM and estimated emotion intensity. Perhaps this reaction has a communication dimension of social signaling. It might be related to the expectation of feedback on the observer’s face, which is occurring by searching the view field and thus, in a change in the dynamics of EM. Here, facial expressions connected to Happiness should be interpreted as displaying different levels of emotional information, on the other hand, posed smiles seem to be a lighter signals, not triggering a feedback reaction in our social and emotional system.

We think that such observations could be also used in medical analysis. One example could be for Bradykinesia with Parkinson's disease. The Bradykinesia can be defined as slow motions and is often characterized by difficulties in initiating, maintaining or voluntarily synchronizing a motions [18]. One of the most difficult aspects of the Bradykinesia, is the fact that it ultimately affects all striated muscles [18]. Therefore, as the disease progresses, people with PD experience chewing difficulties, dysphagia, and difficulty speaking and expressing their faces [18]. Such a person can give the impression of being stiff and wooden-faced. In this regard, we think that the analysis of the EM chaotic dynamics could provide be a support in the interpretation of facial emotions.

5 Conclusions

Happiness allows individuals to build and strengthen social connections and facial expressions are very important features of communication for most of the primates [19,20]. The humans through their social and emotional abilities are excellent in distinguishing between posed and true Happiness, but existing methods in computer science often are wrong and rather unable to define the real FE. The method presented in this article uses EM chaotic dynamics correlation together with the FACS estimations, which can bring significant support and improvement in these problematic processes. We see the sense of further research in this area, as it can help in FE classification of people with mimicry problems, like Parkinson's disease patients with facial effects of the Bradykinesia ("poker face") or just people having problems with proper Happiness mimicry. This method can also bring solution in Happiness detection for partially covered face i.e. covered by mask worn during the epidemic).

Declaration of Competing Interest
OpenFace 2.0 Facial Behavior Analysis Toolkit contributed to the development of the software methods used in the analysis presented in this article. The authors declare no conflict of interests.

References

1. Sledzianowski, A., et al.: Face emotional responses correlate with chaotic dynamics of eye movements. In: Procedia Computer Science, vol. 192, pp. 2881–2892 (2021). Knowledge-Based and Intelligent Information And Engineering Systems: Proceedings of the 25th International Conference KES (2021)
2. P. Ekman and W.V. Friesen. Facial Action Coding System. Consulting Psychologists Press, Sunnyvale (1978)
3. McDuff, D., el Kaliouby, R., Senechal, T., Amr, M., Cohn, J.F., Picard, R.: Affectiva-MIT facial expression dataset (AM-FED): naturalistic and spontaneous facial expressions collected "in-the-wild". In 2013 IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 881–888 (2013)
4. Baltrusaitis, T., Zadeh, A., Lim, Y.C., Morency, L.-P.: OpenFace 2.0: facial behavior analysis toolkit. In: 2018 13th IEEE International Conference on Automatic Face Gesture Recognition (FG 2018), pp. 59–66 (2018)

5. Glass, L.: Synchronization and rhythmic processes in physiology. Nature **410**, 277–284 (2001)
6. Piórek, M., Josiński, H., Michalczuk, A., Świtoński, A., Szczesna, A.: Quaternions and joint angles in an analysis of local stability of gait for different variants of walking speed and treadmill slope. Inf. Sci. **384**, 263–280 (2017)
7. Afsar, O., Tirnakli, U., Marwan, N.: Recurrence quantification analysis at work: Quasi-periodicity based interpretation of gait force profiles for patients with parkinson disease. Sci. Rep. **8**, 1–12 (2018)
8. Przybyszewski, A.W., Linsay, P.S., Gaudiano, P., Wilson, C.M.: Basic difference between brain and computer: integration of asynchronous processes implemented as hardware model of the retina. IEEE Trans. Neural Netw. **18**, 70–85 (2007)
9. Paulson, E.J.: Viewing eye movements during reading through the lens of chaos theory: how reading is like the weather. Read. Res. Q. **40**(3), 338–358 (2005)
10. Hampson, K.M., Mallen, E.A.H.: Chaos in ocular aberration dynamics of the human eye. Biomed. Opt. Express. **3**(5), 863–877 (2012)
11. Richards, W., Wilson, H.R., Sommer, M.A.: Chaos in percepts? Biol. Cybern. **70**, 345–349 (2004)
12. Harezlak, K., Kasprowski, P.: Searching for chaos evidence in eye movement signals. Entropy **20**(1), 1–11 (2018)
13. Chanijani, S.S.M., Klein, P., Bukhari, S.S., Kuhn, J., Dengel, A.: Entropy based transition analysis of eye movement on physics representational competence. In: UbiComp Adjunct, pp. 1027–1034 (2016)
14. Astefanoaei, C., Pretegiani, E., Optican, L.M., Creanga, D., Rufa, A.: Eye movement recording and nonlinear dynamics analysis - the case of saccades. Rom J. Biophys. **23**, 81–92 (2013)
15. Urbanowicz, Krzysztof, Hołyst, Janusz A.: Noise-level estimation of time series using coarse-grained entropy. Phys. Rev. E **67**, 1–14 (2003)
16. Chen, T., Guestrin, C.: XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD 2016, pp. 785–794. ACM, New York, (2016)
17. Stewart, P.A., Bucy, E.P., Mehu, M.: Strengthening bonds and connecting with followers: a biobehavioral inventory of political smiles. Polit. Life Sci. **34**(1), 73–92 (2015)
18. Galvan, A., Wichmann, T.: Pathophysiology of parkinsonism. Clin. Neurophysiol. Off. J. Int. Fed. Clin. Neurophysiol. **119**(7), 1459–1474 (2008)
19. Taylor, S.: Tend and befriend biobehavioral bases of affiliation under stress. Current Direct. Psychol. Sci. **15**, 273–277 (2006)
20. Taylor, S., Klein, L., Lewis, B., Gruenewald, R.G., Updegraff, J.: Biobehavioral responses to stress in females: tend-and-befriend, not fight-or-flight. Psychol. Rev. **107**, 411–429 (2000)