Predicting Brainwaves from Face Videos

Christian S. Pilz
CanControls, Aachen, Germany

Ibtissem Ben Makhlouf
CanControls, Aachen, Germany

Ute Habel
Uniklinik RWTH Aachen, Germany

Steffen Leonhardt
RWTH Aachen University, Germany

Abstract

We investigate the regulation of human brain arousal in the central nervous system and its synchronization with the autonomic nervous system affecting the facial dynamics and its behavioral gestalt. A major focus is made on the sensing observable during natural human eye to eye communication. Although the inner state of the autopoietic system is deterministic, its outer facial behavioral component non-deterministic. Beside the introduction of general validity of the classical empirical interpretation of the vigilance continuum during open eyes, we show that the facial behavior can be used as suitable surrogate measurement for specific states of mind. As a consequence we predict brainwaves from face videos formulated as inverse problem of the underlying stochastic process. Finally, we discuss the impact and range of application field.

1. Introduction

Based on the concept of association of entropy production in living systems, it is generally accepted that an increase in order within an organism is compensated for by an increase in disorder outside this organism [31]. Although it seems that the dynamics of life are at odds with the second law of thermodynamics, which states that the entropy of an isolated system can only increase, this paradox is apparently resolved by the open nature of living systems. This basically states that those kind of systems can exchange either heat or matter or both with their environment. In order to sustain vital processes, living organisms continuously consume energy, produces heat and entropy via metabolic pathways and transfers entropy to the environment through various waste channels to maintain a permanent thermal biological state. As a result, the human body can be described as an open thermodynamic system. Since entropy and its related variables vary with time, the dynamic of entropy typically describes the human physiological and developmental processes in this context. This allows to measure the various states of a living system, for example differentiating between states of health and disease.

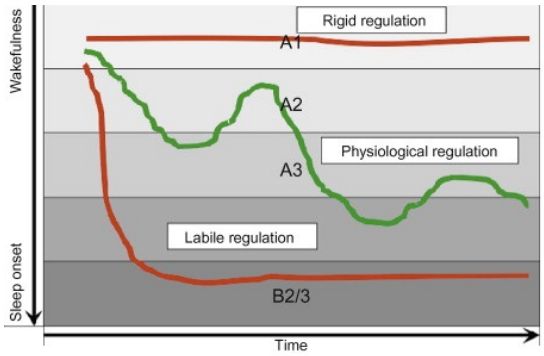

Figure 1. Brain arousal regulation of the vigilance continuum as synopsis of commonly used viso-morphologically defined classifications of the EEG-stages occurring between relaxed wakefulness during closed eyes and deep sleep. The vigilance is measured as the connectivity in terms of EEG alpha waves between the visual cortex and the frontal lobe. The stages A1 to B2 are defined by the seminal works of Bente and Roth and are further investigated by Ulrich, Olbrich and Hegerl [5, 15, 20, 25, 33, 34]. A decrease in vigilance can lead to two different behaviors, 1) the organism decides to go to sleep and the vigilance reverts to sleep stages or, 2) the organism exhibits auto-stabilization behavior to counter regulate their vigilance level.

The challenge, however, is to understand how these human individual processes are involved into this physical energy regulation. In human machine interaction the role of the central and vegetative nervous system is usually not observable. However, these inner states are important mandatory factors building sophisticated advanced machines able to reproduce anthropomorphic associative abilities. Here, mind and emotions are properties which machines are not able to understand and to reproduce reasonable well yet. In this study, we attempt to investigate the basic principle of the thermodynamic regulation during sensing observable during natural human eye to eye communication. From the theoretical background, we form the mathematical model and

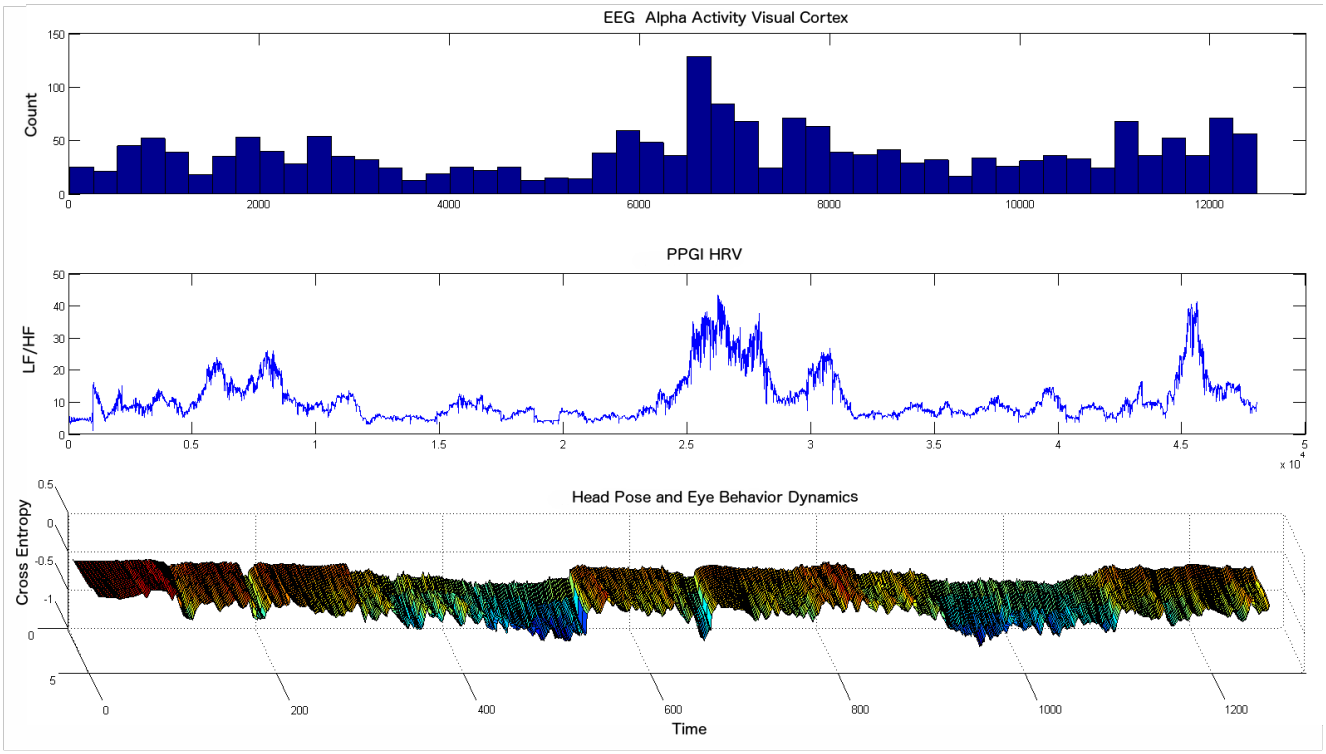

Figure 2. The physiological and behavioral feature time series of an user over the entire recording duration of approximately three hours. The first row represents the dominant occipital alpha power as histogram count. At the beginning, in the middle and at the end we see periods of higher alertness. These periods exactly correspond to the exhausting reaction time tests performed for measuring alertness and sustained attention of each individual user. The second row represents the Photoplethysmography Imaging [21, 22, 23] based heart rate variability as low frequency/ high frequency ratio. Again, during the periods corresponding to the reaction time tests the activity of the autonomic nervous system increases significantly like the EEG vigilance. By comparing the dynamics of the facial features in the last row with the physiological brain arousal and the autonomic nervous system there's again some similarity. In the multi-scale entropy [1, 9] graph the facial features produce plateaus during the reaction time tests. Particular they seem to be correlated.

we show how learning can be interpreted with linear complexity. In substantial experiments we verify the hypothesis and finally predict brainwaves from face videos.

2. The Vigilance

If we summarize under vigilance such different terms, like alertness, sustained attention and selective attention, we are in fundamental danger of a categorical post-hoc fallacy. Sustained and selective attention are certain measurable performances, resulting from the interaction between a living organism and its environment. Therefore these terms refer to the psychology and include all kinds of behavior, outward as well as inward directed. Alertness with its tonic or phasic part does not refer to some concrete performance. It refers to the mechanism enabling a certain performance. This is part of the physiology domain of organismic functions which has to be strictly distinguished from performances. We differentiate between the physiological and psychological phenomena because both phenomena are neither caused by physiological phenomena nor reducible to them. The term vigilance dates back to the seminal work of Head [14]. In its origin Head's vigilance is neither a function nor a performance. It has more a semantic meaning like Bateson's idea of *explanatory principles* [4] and these principles are synonymous to Carnap's ideas of a *theoretical construct* [7], finally Kant's metaphysical *Regulative Ideen* [17] as foundation of the transcendental dialectic [33]. We particular emphasize that Head's vigilance is primarily useful only for scientists, who are interested in the reality behind the empirical phenomena. If someone assumes the strictly mainstream position, accepting objectively measurable facts by senses and introspection, Head's vigilance appears nonsensical [32]. Electroencephalography (EEG) changes of vigilance by a complex spatio-temporal process were first pointed out by Loomis and Davis[10, 18]. The authors distinguished two EEG stages between resting wakefulness and sleep. Half a century ago, the practical usefulness was further investigated by Bente and Roth [5, 25]. They defined the stages into substages (see Figure 1). The first machine learning approach for the classification of these stages is given by the work of Ulrich, Olbrich

and Hegerl [15, 20, 33, 34]. The vigilance is measured as the connectivity in terms of EEG alpha waves between the visual cortex and the frontal lobe. In this research the term vigilance is used as a multivariate construct which is investigated in the domain of observable behavior.

3. The Model Space

The primary goal is to infer from a given initial facial dynamics to the target brain arousal evolution describing the state of mind. This can be expressed as an inverse problem

$$O = G(M) \tag{1}$$

where the operator G describes the relation between the observation O and the model M. The problem reduced to estimating from an observed random process x another related process y. The two processes are related by a probabilistic model $p(x, y)$ which is unknown. The solution is given by the unconstrained optimization problem

$$\underset{f(x)}{\operatorname{argmin}} E\big[\|y - f(x)\|^2\big]. \tag{2}$$

If $p(y)$ and $p(x|y)$ are Gaussian and $E\big[x|y\big]$ is linear in y, the solution is linear. For zero mean y and x the solution yields to the conditional mean estimate

$$E\big[y|x\big] = w^T x \tag{3}$$

where

$$w = \underset{w}{\operatorname{argmin}} E\big[(y - w^T x)^2\big] \tag{4}$$

$$w = (E\big[xx^T\big])^{-1} E\big[xy\big] \tag{5}$$

In case x is not linearly related to y or y is not Gaussian distributed, the conditional estimate of y given x is nonlinear. The computation of the nonlinear conditional mean estimate leads to over-fitted or sub-optimal solutions, since there are no convergence guarantees or the regularized estimation yields to restricted distributions. According to the Hardamand criterions the problem is ill-posed [12].

If we assume the model function $f(x_*)$ with a known input $x_* \in \mathbb{R}^d$ to be a realization of a prior represented by a Gaussian random process

$$f \sim \mathcal{GP}(0, k(x, x')) \tag{6}$$

and the observations corrupted by Gaussian noise

$$y_k = f(x_k) + \varepsilon_k \tag{7}$$

with $\varepsilon_k \sim \mathcal{N}(0, \sigma_n^2)$. For a given training data set $D = \{(x_k, y_k)|k = 1, 2, ..., n\}$, we are able to find the closed form solution for the prediction [24]

$$p(f(x_*)|D) = \mathcal{N}(\mathbb{E}[f(x_*)], \mathbb{V}[f(x_*)]) \tag{8}$$

with

$$\mathbb{E}[f(x_*)] = k_*^T (K + \sigma_n^2 I)^{-1} y \tag{9}$$

and

$$\mathbb{V}[f(x_*)] = k(x_*, x_*) - k_*^T (K + \sigma_n^2) I)^{-1} k_*. \tag{10}$$

Usually the hyperparameters of the covariance function $k(x, x')$ and the noise variance σ_n^2 have to be learnt by maximizing the marginal likelihood function using gradient descent methods. However, due to the recent progress in the interpretation of stochastic processes, a more convenient way of constructing the Gaussian process is given as solution to the following nth order linear stochastic differential equation [16, 28]

$$a_n \frac{\mathrm{d}^n f(t)}{\mathrm{d}t^n} + ... + a_1 \frac{\mathrm{d}f(t)}{\mathrm{d}t} + a_0 f(t) = w(t), \tag{11}$$

with $w(t)$ a zero-mean Gaussian white noise process. According to Wiener's spectral factorization [36] the Fourier transform of the differential equation yields to

$$F(i\omega) = \underbrace{(\frac{1}{a_n(i\omega)^n + ... + a_1(i\omega) + a_0})}_{G(i\omega)} W(i\omega). \tag{12}$$

Essentially this is a rational function and the causal transfer function. The form of an all-pole filter. From the Wiener-Khinchin theorem we know that a stationary covariance function is given by the inverse Fourier transform of the corresponding spectral density [8, 35]

$$C(t) = \mathcal{F}^{-1}[S(\omega)] = \frac{1}{2\pi} \int S(\omega) \exp(i\omega t) \mathrm{d}\omega, \tag{13}$$

where the spectral density of the process is given by

$$S(\omega) = q_c |G(i\omega)|^2 \tag{14}$$

with $q_c = |W(i\omega)|^2$ the spectral density of the white noise process. Converting the transfer function to the corresponding state space formulation results in

$$\frac{\mathrm{d}f(t)}{\mathrm{d}t} = \underbrace{\begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ -a_0 & -a_1 & \dots & -a_{n-1} \end{pmatrix}}_{\mathrm{F}} f(t) + \underbrace{\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}}_{\mathrm{L}} w(t) \tag{15}$$

and

$$y_k = \underbrace{\begin{pmatrix} 1 & 0 & \dots & 0 \end{pmatrix}}_{\mathrm{H}} f(t) + \varepsilon_k, \tag{16}$$

with the feedback matrix F, the noise effect matrix L and the observation matrix H. This kind of inference problem can be

solved with linear complexity using the traditional Kalman filtering. Here, the optimization of the hyperparameters is solved implicitly. For some standard covariance functions the approximation of the transfer function is given by its Taylor series expansion [16].

4. Experiments

We performed substantial experiments under laboratory conditions in order to determine the statistical relationship between the brain arousal, the cardiovascular activity and the facial behavioral dynamics. Here, we focused on the physiological continuous measurement of vigilance by EEG sensors positioned on the frontal lobe and the visual cortex. The facial behavioral dynamics were monitored by a common RGB camera based facial alignment [26, 37]. The cardiovascular activity by the camera based Photoplethysmography Imaging (PPGI) [21, 22, 23]. As standard psychological performance-based surrogate measurement for the brain arousal dynamics we instructed all user during the experiments to participate in sustained attention and alertness measurements to enable their maintenance of response persistence and continuous effort over extended periods of time [11]. We completed the sensor acquisition by measure the situational subjective level of sleepiness using the Karolinska sleepiness scale (KSS) [2]. The completely synchronized recordings were then automatically analyzed to extract the features for performing the final statistical regression analysis.

Data Collection

Uncompressed video, EEG and reaction-time of 100 users were recorded during the experimental laboratory sessions. Each single sessions consists of a time length of approximately three hours. The gender balanced population was selected to fit a normal based age-range between 18 and 65 years old persons. In Figure 3 three younger persons of the database are shown wearing the EEG sensors and a headset. All video data was recorded with a 25 frames per second using an IDS RGB camera. The brain signals were captured with EEG wet electrodes using a VARIOPORT analog to digital converter with a sample rate of 128 Hz. We divided the three hours recording time into different parts where each user performed different tasks. The two main parts were selected with a duration about one hour. In the first part all users played an action sport computer game and in the second part a train simulator game. Every five minutes we asked the users via the headset to state their subjective level of sleepiness on the KSS. Before and after each game session we instructed the users to perform psychomotor vigilance tasks (PVT) [11] over a time period of approximately 15 minutes. At the start and end of each PVT the users additionally stated their KSS score. To guarantee data consistency, a manual pre-preprocessing check was carried out after each recording. Utilizing recorded timestamp markers we were able to segment the different tasks automatically.

Feature Space

All EEG signals were analysed by standard Fourier based spectral method. To compensate eye blink artefacts in the EEG signals we applied a common unsupervised ICA approach. We selected a sliding time window of 60 seconds with one seconds overlap for the short time Fourier transform. In order to extract the alpha activity in the corresponding brain regions, we band-pass filtered the frames of spectral power in the range between 8 and 12 Hz. The EEG vigilance level is determined based upon the ratio of high frontal Alpha power and high occipital alpha power. High occipital alpha power is assumed to reflect vigilance stage A1 alertness states and high frontal alpha power the stage of transition to more sleepy states during closed eyes (see Figure 1). In Figure 2 the first row represents the dominant occipital alpha power as histogram count over the time period of the entire recording of an user. At the beginning, in the middle and at the end we see periods of higher alertness. These periods exactly correspond to the exhausting reaction time tests performed for measuring alertness and sustained attention of each individual user.

For measuring the activity of the autonomic nervous system we concentrated on the relationship between the sympathetic and parasympathetic activity. Here, we used the camera based PPGI method [21, 22, 23] to extract the RR-interval of the quasi periodic blood volume changes out of registered facial skin regions [26, 37]. Each obtained pulse waveform is re-sampled to 256 Hz first. In the next step a frequency domain based method [3, 29] using the Lomb–Scargle periodogram [19, 30] is applied to extract high and low frequency bands out of the RR-traces in order to determine the heart rate variability. We used the bands from 0.15 to 0.4 Hz as high frequency (HF) and from 0.04 to 0.15 Hz as low frequency (LF) regions. As activity level the LF/HF ratio is used. High LF/HF values indicates a stronger sympathetic activity analog to the fight or flight response [6]. In Figure 2 the second row represents the PPGI LF/HF ratio over the time period of the entire recording of an user. During the periods corresponding to the reaction time tests, the activity of the autonomic nervous system increases significantly like the EEG vigilance. This seems to be correlated.

The facial behavioral dynamics can be represented by several different features. We focused on the ocular movements in terms of eye gaze and eye blink behavior. We dropped all head pose features since during the experiments all users are instructed to follow the course of action on the nearby computer screen. Thus, head pose dynamics will not occur in a sufficient large manner. Initially the face was captured

Figure 3. Three images from the data collection. These younger persons are wearing the EEG sensor cap and a headset for experiment instructions. The camera is positioned below the computer monitor. During the experiments all participants performed several different tasks like playing computer games and performing reaction time tests. After specific time periods they were asked to provide their situational subjective sleepiness rating.

by a common face finder. The rough face region is used to initialize the facial alignment using a cascaded regression approach [26, 37]. From the set of registered landmarks the head pose can be determined by either predicting it using a separately trained pose regressor [37] or fitting a statistical 3D point distribution model to the target landmarks in the image plane [27]. Over the set of eye landmarks we computed the eye lid opening and fitted a 3D geometric eye model to determine the eye gaze [13]. In Figure 2 the last row represents the facial behavioral dynamics over the time period of the entire recording of an user. We employed a cross entropy analysis to provide robust measures of the deterministic or stochastic content of the time series (regularity), as well as the degree of structural richness (complexity), through operations at multiple data scales [1, 9]. By comparing the dynamics of the facial features with the physiological brain arousal and the autonomic nervous system activity we find again some similarity. In the multiscale graph the facial features produce plateaus during the reaction time tests. This gave us a reason to conduct a more detailed statistical analysis.

Statistical Analysis

To act on the assumption of some correlation between the activity of the central and autonomic nervous system and its impact on facial behavioral dynamics, we performed a detailed statistical analysis of the pre-computed features described in the previous section. Our primary concern was to establish whether we are able to reconstruct brain arousal states using video based facial features and what feature behavior is forming the basis of the hypothesized underlying predictive power.

In the first step we computed the correlation between the dominant EEG occipital alpha power and the PPGI LF/HF ratio over the entire database. To analyse we calculated the cumulative distribution function (CDF) of these correlation values. The corresponding CDF ist visualized in the following Figure 4. The correlation values are varying from -0.8 to 0.8. with nearly exactly 50 percent coverage for all negative correlations and 50 percent coverage for all positive correlations. The observable anti-correlation behavior appears to describe the loss of energy transport via oxygen through the blood flow to the brain in order to maintain states of hypervigilance. This can be interpreted as transition to hypovigilance, e.q states of higher sleep onset or drowsiness. The human body is starting to deactivate the visual cortex and can't pay attention anymore. To verify the established brain-heart synchronization, we

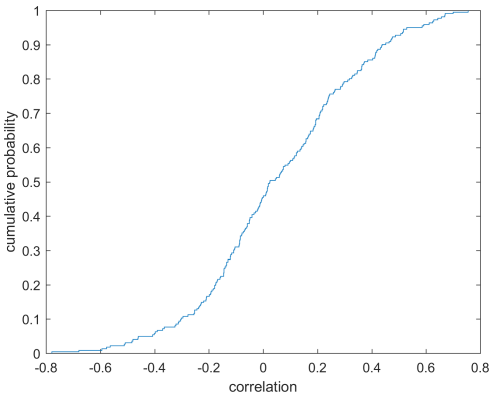

Figure 4. The cumulative distribution function of correlation values between the dominant EEG occipital alpha power and the PPGI heart rate variability in terms of LF/HF ratio. A strong correlation versus anti-correlation behavior.

computed the feature behavior of both instances over the entire range of correlations. We decided to visualize each feature using a box-plot diagram for each correlation bin. Since the amount of samples in the absolute correlation range between 0.5 and 0.8 is rather small, we decided to assign these feature values to the 0.5 correlation bin. In Figure 5 the box plot of the dominant EEG occipital alpha power over the correlation between the PPGI heart rate variability in terms of LF/HF ratio is visualized. A slightly trend from positive to negative correlations. The occipital alpha power increases for higher negative correlations in

contrast to the positive ones. For the range of positive correlations exists a higher amount of uncertainty. In

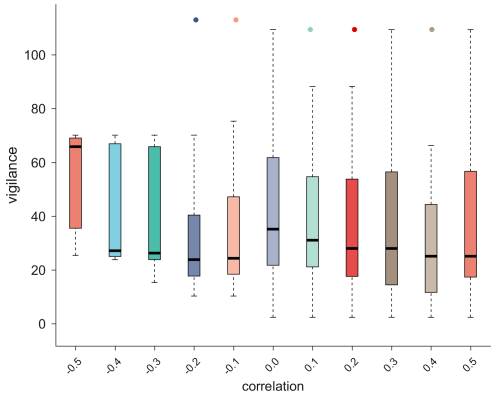

Figure 5. The box plot visualization of the dominant EEG occipital alpha power over the correlation range between the PPGI heart rate variability in terms of LF/HF ratio.

Figure 6 the box plot of the PPGI heart rate variability in terms of LF/HF ratio is visualized. A strong decreasing trend from positive to negative correlations. The PPGI LF/HF ratio is higher for positive correlations in contrast to the negative ones. Again, for the range of positive correlations there exists a higher amount of uncertainty. Essentially, for high correlation values the EEG occipital

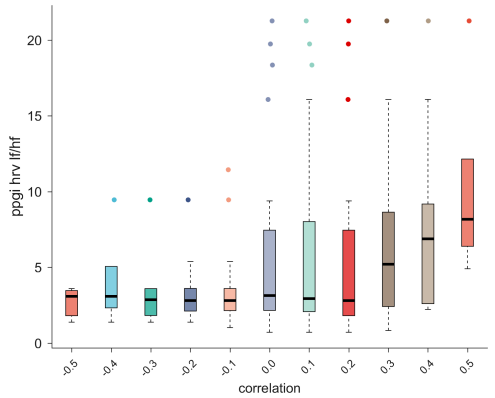

Figure 6. The box plot visualization of the PPGI heart rate variability in terms of LF/HF ratio over the correlation range between the dominant EEG occipital alpha power.

alpha power is low and the sympathetic activity of the autonomic nervous system is high. Here, the human body attempts to preserve the energy supply to the brain in order to maintain states of higher wakefulness. Once this energy supply gets regulated by the parasympathetic activity, the PPGI LF/HF ratio changes to a smaller values. The EEG occipital alpha power increases shutting down the brain region of the visual cortex. As a result the human should feel sleepy and should not be able to pay attention anymore. This is the state of hypovigilance. Since we monitored the users subjective sleepiness rating, we decided to verify this against the KSS scores. Analog to the previous processing we computed the KSS score behavior over the entire range of correlations. In Figure 7 the box plot of the KSS scores as metric for the subjective sleepiness is visualized. We are able to observe an increasing sleepiness trend from positive to negative correlations. This acts as strong indicator for the validity of the previous interpretation of the vigilance and heart rate variability connection. Indeed, the human body seems to regulate the vigilance level by energy support activated through the autonomic nervous system. In contrast to the classical vigilance interpretation

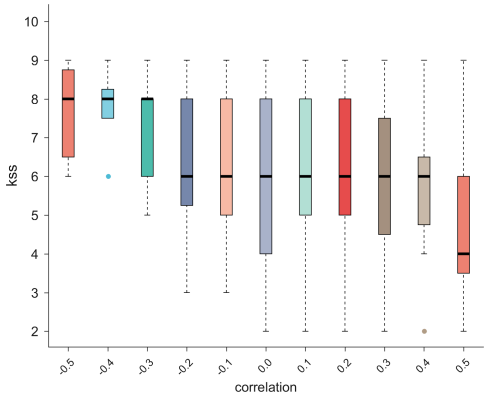

Figure 7. The box plot visualization of the subjective level of sleepiness expressed as KSS score over the correlation range between the dominant EEG occipital alpha power and the PPGI heart rate variability in terms of LF/HF ratio.

during closed eyes a high level of EEG alpha power doesn't correspond to wakefulness states during open eyes. It does more reflect the deactivation of the specific brain region towards the state of sleep onset. In this case the visual cortex followed by the frontal lobe. Thus, we arrive at the conclusion that these states noticeable impact the human performance carried out during the reaction time tests in our experiments. We computed the reaction time for the sustained attention and corresponding tap error behavior for both features over the entire range of correlations again. In Figure 8 the box plot of the sustained attention reaction time in seconds is visualized. For higher negative correlation values the reaction time is larger. In Figure 9 the box plot of the sustained attention tap errors as count of wrong decision is visualized. There's an increasing amount of errors from positive to negative correlation values observable. As expected, during states of higher hypovigilance the users aren't able to maintain attention anymore as good as during fully wakefulness. In the final step of the conducted feature analysis we performed the same procedure for the facial ocular movements in terms of eye blink and eye gaze dynamics. Since eye behavior can

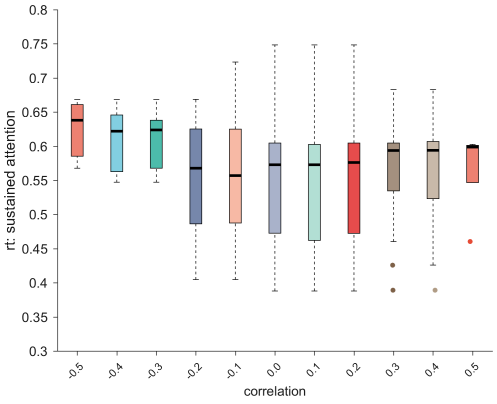

Figure 8. The box plot visualization of the sustained attention reaction time in seconds over the correlation range between the dominant EEG occipital alpha power and the PPGI heart rate variability in terms of LF/HF ratio.

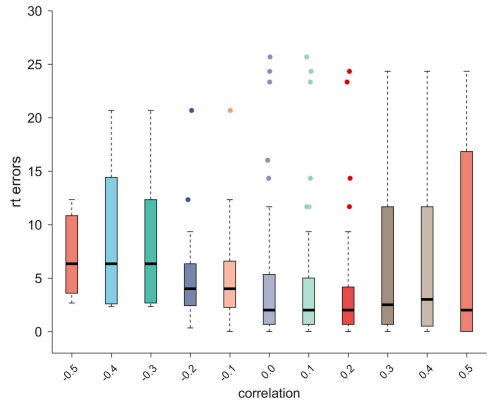

Figure 9. The box plot visualization of the sustained attention reaction time errors as wrong decision over the correlation range between the dominant EEG occipital alpha power and the PPGI heart rate variability in terms of LF/HF ratio.

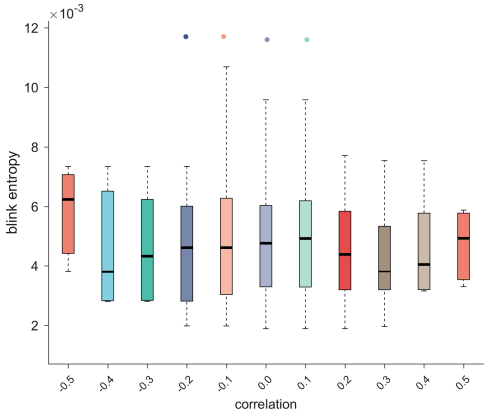

Figure 10. The box plot visualization of the eye blink dynamics as a single scale of the multivariate sample entropy over the correlation range between the dominant EEG occipital alpha power and the PPGI heart rate variability in terms of LF/HF ratio.

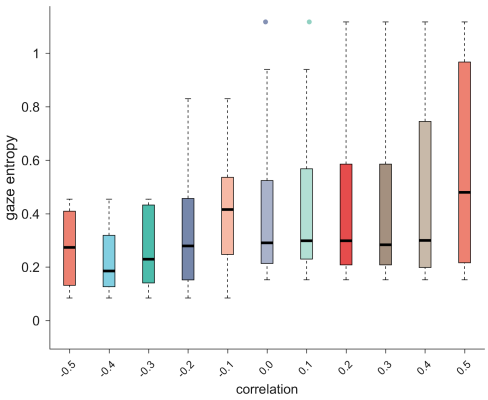

Figure 11. The box plot visualization of the eye gaze entropy over the correlation range between the dominant EEG occipital alpha power and the PPGI heart rate variability in terms of LF/HF ratio.

be parameterized differently, we decided to compute the eye blink frequency and the individual duration. The eye gaze is parameterized by the eye gaze angular displacement and velocity. We aggregated both instances of the eyes and their features using a single scale of the multivariate sample entropy [1, 9]. In Figure 10 the box plot of the eye blink entropy is visualized. An increased entropy for high negative correlation values. This reflects an increase of complexity which can be interpreted as a loss of deterministic behavior during states of hypovigilance. We tried to inspect what exactly happens with the individual eye blink parameters. We often found an irregular increase of blink frequency and longer blink duration. In Figure 11 the box plot of the eye gaze entropy is visualized. We're able to observe a stronger change of the entropy values from positive to negative correlations. This time the entropy decreases towards high negative correlations with a loss of uncertainty. Unlike the blink dynamics, this means an increase of deterministic behavior like the eyes are staring into space.

In the final experiment we attempted to predict states of vigilance as well as the common surrogate measurements for the performance based indicators by means of machine learning. We defined the dominant EEG occipital alpha power, the KSS score, the PVT reaction time and tap error as individual target variables. The predictor variables are selected as the camera based features computed out of facial regions solely. This pool of predictor variables includes the PPGI based heart rate variability and the ocular movements in terms of blink frequency, blink duration, the eye gaze angular displacement and the eye gaze velocity. By applying the coarse grained sample entropy approach [1, 9] over the set of features we obtained the fused complexity dynamics of the facial skin blood perfusion and

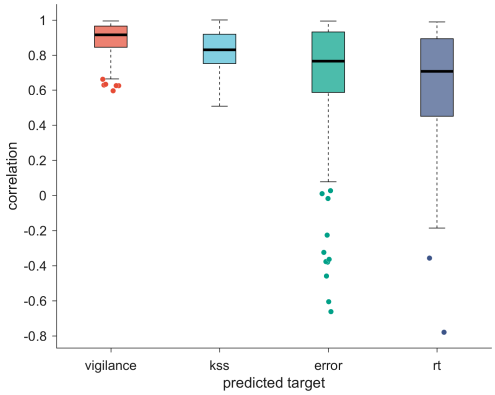

Figure 12. The box plot visualization of correlation values for the regression cross-validation against different targets. The predictor variables are based on the camera based features computed out of facial regions. From left to right: Vigilance corresponds to the dominant EEG occipital alpha power. Kss corresponds to the situational subjective level of sleepiness, e.q. KSS score. Error corresponds to the false tap answers during the reaction time tests and rt corresponds to the measured PVT reaction time.

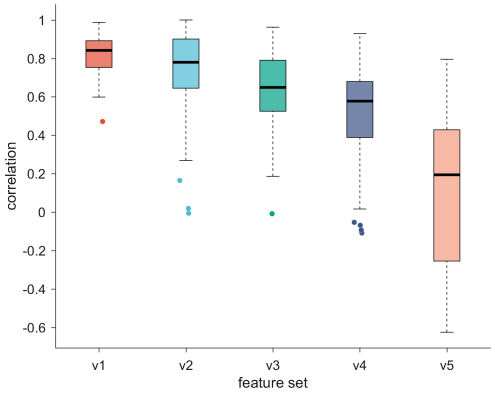

Figure 13. The box plot visualization of correlation values for the regression cross-validation using different kinds of features against the KSS target. From left to right: V1 corresponds to the PPGI based heart rate variability, the ocular movements in terms of blink frequency, blink duration, the eye gaze angular displacement and the eye gaze velocity. V2 corresponds to the ocular movements in terms of blink frequency, blink duration, the eye gaze angular displacement and the eye gaze velocity. V3 corresponds to the eye gaze angular displacement and the eye gaze velocity. V4 corresponds to the ocular movements in terms of blink frequency, blink duration. V5 corresponds to the PPGI based heart rate variability.

facial behavior. We solved the inverse problem by applying the described Kalman filtering approach on state space representation of the Taylor series approximated squared exponential covariance function [16]. We performed a leave-one-out cross validation on the entire dataset of users. In Figure 12 the box plot of the correlation values for the regression cross validation is visualized. A decreasing performance from left to right. In the box plots vigilance corresponds to the dominant EEG occipital alpha power, KSS corresponds to the situational subjective level of sleepiness. Error corresponds to the false tap answers during the reaction time tests and RT corresponds to the measured PVT reaction time. To determine the predictive power of the individual features, we repeated the cross-validation for different kinds of feature sets. To simulate a potential application scenario, we selected the KSS score to predict the user's sleepiness as target. In Figure 13 the box plot of the correlation values for the regression cross validation using the different features against the KSS score is visualized. Interestingly but not surprised regarding the anti-correlation behavior presented in Figure 4, the heart rate variability alone is not able to predict KSS states reasonable well. However, the ocular dynamics perform well and benefit from the heart rate variability.

5. Discussion

We studied the regulation of human brain arousal in the central nervous system and its synchronization with the autonomic nervous system affecting the facial dynamics. We performed a detailed statistical analysis of the individual signal behavior and predicted states of vigilance using a video based surrogate measurement out of facial regions solely. We demonstrated the possibility to reconstruct brainwaves directly from face videos. Furthermore, it was shown the obtained results can also be used to derive other kind of markers. Sustained attention and sleepiness are only one of these direct derivatives which have a broad application field in the transport, control room and marketing sector. In the medical sector the rigid and labile regulation of the vigilance reflects a suitable biomarker for the diagnosis of affective disorders. Particular in this sector, such kind of markers are the essential requirements for the development of new therapeutic methods and investigation of new pharmaceuticals. The proper combination with the emotional brain activity would pave the way for an entire new interpretation of the dimensional valence-arousal model in affective computing.

References

[1] M. U. Ahmed and D. P. Mandic. Multivariate multiscale entropy analysis. *IEEE Signal Processing Letters*, 19(2):91–94, 2012. 2, 5, 7

[2] T. Akerstedt and M. Gillberg. Subjective and objective sleepiness in the active individual. *International Journal of Neuroscience*, 52:29–37, 1990. 4

[3] S. Akselrod, D. Gordon, J.B. Madwed, N.C. Snidman, D.C. Shannon, and R.J. Cohen. Hemodynamic regulation: investigation by spectral analysis. *Am J Physiol*, 249:867–875, 1985. 4

[4] G. Bateson. Mind and nature: A necessary unity. *New York: E. P. Dutton*, 1979. 2

[5] D. Bente. Die Insuffizienz des Vigilitästonus. *Habilitationsschrift*, 1964. 1, 2

[6] S. Bracha. Freeze, flight, fight, fright, faint: Adaptationist perspectives on the acute stress response spectrum. *CNS Spectrums*, pages 679–685, 2004. 4

[7] R. Carnap. The methodological character of theoretical concepts. *The Foundations of Science and the Concepts of Psychology and Psychoanalysis, University of Minnesota Press*, pages 38–76, 1956. 2

[8] A. Chintchin. Korrelationstheorie der stationären stochastischen Prozesse. *Mathematische Annalen*, 109(1), 1934. 3

[9] M. Costa and A.L. Goldberger C.-K. Peng. Multiscale entropy analysis of complex physiologic time series. *PPhysical Review Letters*, 89(2), 2002. 2, 5, 7

[10] H. Davis, P. A. Davis, A. L. Loomis, E. N. Harvey, and G. Hobart. Changes in human brain potentials during the onset of sleep. *Journal of Experimental Psychology*, 1:24–38, 1938. 2

[11] I. D. Dinges and J. W. Powell. Microcomputer analysis of performance on a portable, simple visual reaction time task sustained operations. *Behavior Research Methods, Instrumentation, and Computers*, 17:652–655, 1985. 4

[12] J. Hadamard. Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton University Bulletin*, pages 49–52, 1902. 3

[13] D. W. Hansen and Q. Ji. In the eye of the beholder: A survey of models for eyes and gaze. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(3):478–500, 2010. 5

[14] H. Head. The conception of nervous and mental energy - vigilance: a physiological state of the nervous system. *British Journal of Psychology*, 14:126–147, 1923. 2

[15] U. Hegerl and S. Olbrich. The vigilance regulation model of affective disorders and adhd. *Neuroscience and Biobehavioral Reviews*, 44:45–57, 2014. 1, 3

[16] J.Hartikainen and S. Särkkä. Kalman filtering and smoothing solutions to temporal gaussian process regression models. *IEEE International Workshop on Machine Learning for Signal Processing*, pages 379–384, 2010. 3, 4, 8

[17] I. Kant. Kritik der reinen Vernunft. *Verlag von L. Heimann*, 1870. 2

[18] A. l. Loomis, E. N. Harvey, and G. A. Hobart. Celebral states during sleep as studied by human brain potentials. *Journal of Experimental Psychology*, 21:127–144, 1937. 2

[19] N.R. Lomb. Least-squares frequency analysis of unequally spaced data. *Astrophysics and Space Science*, 39(2):447—-462, 1976. 4

[20] S. Olbrich, C. Sander, H. Matschinger, R. Mergl, M. Trenner, P. Schoenknecht, and U. Hegerl. Brain and body: Associations between eeg-vigilance and the autonomous nervous system activity during rest. *Journal of Psychophysiology*, 25(4):190–204, 2011. 1, 3

[21] C. S. Pilz, J. Krajewski, and V. Blazek. On the diffusion process for heart rate estimation from face videos under realistic conditions. *Pattern Recognition. GCPR 2017. Lecture Notes in Computer Science, vol 10496. Springer*, 10496:361–373, 2017. 2, 4

[22] C. S. Pilz, I. B. Makhlouf, V. Blazek, and S. Leonhardt. On the vector space in photoplethysmography imaging. *The IEEE International Conference on Computer Vision (ICCV) Workshops*, pages 1254–1262, 2018. 2, 4

[23] C. S. Pilz, S Zaunseder, J. Krajewski, and V. Blazek. Local group invariance for heart rate estimation from face videos in the wild. *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, pages 1254–1262, 2018. 2, 4

[24] C. E. Rasmussen and C. K. I. Williams. Gaussian processes for machine learning. *The MIT Press*, 2006. 3

[25] B. Roth. The clinical and theoretical importance of eeg rhythms corresponding to states of lowered vigilance. *Clinical Neurophysiology*, 13:395–399, 1961. 1, 2

[26] E. Sanchez-Lozano, B. Martinez, G. Tzimiropoulos, and M. Valstar. Cascaded continuous regression for real-time incremental face tracking. *The 14th European Conference on Computer Vision (EECV)*, pages 532–539, 2016. 4, 5

[27] J. M. Saragih, S. Lucey, and J. F. Cohn. Deformable model fitting by regularized landmark mean-shift. *International Journal of Computer Vision*, 91(2):200–2015, 2011. 5

[28] S. Särkkä, A. Solin, and J.Hartikainen. Spatio-temporal learning via infinite-dimensional bayesian filtering and smoothing. *IEEE Signal Processing Magazine*, 30(4), 2013. 3

[29] B. M. Sayers. Analysis of heart rate variability. *Ergonomics*, 16:17–32, 1973. 4

[30] J.D. Scargle. Studies in astronomical time series analysis. ii - Statistical aspects of spectral analysis of unevenly spaced data. *Astrophysical Journal*, 263(1):835–853, 1982. 4

[31] E. Schrödiner. What is Life? The Physical Aspect of the Living Cell. *Cambridge University Press*, 1944. 1

[32] G. Ulrich. Addendum to: The importance of the concept of vigilance for psychophysiological research. Held at University of Leipzig 2008. *Medical Hypotheses*, 27:227–229, 1988. 2

[33] G. Ulrich. The theoretical interpretation of electroencephalography (eeg). *BMED Press LLC*, 2013. 1, 2, 3

[34] G. Ulrich and K. Frick. A new quantitative approach to the assessment of stages of vigilance as defined by spatiotemporal eeg patterning. *Perceptual and Motor Skills*, 62(2):567–576, 1986. 1, 3

[35] N. Wiener. Generalized harmonic analysis. *Acta Mathematica*, 55(1), 1930. 3

[36] N. Wiener and P. Masani. The prediction theory of multivariate stochastic processes. *Acta Math*, pages 111–150, 1957. 3

[37] X. Xiong and F. De la Torre. Supervised descent method and its applications to face alignment. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 532–539, 2013. 4, 5