

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/51530664>

EEG potentials predict upcoming emergency braking during simulated driving

Article in *Journal of Neural Engineering* · July 2011

DOI: 10.1088/1741-2560/8/5/056001 · Source: PubMed

CITATIONS

114

READS

849

6 authors, including:



Stefan Haufe

Charité Universitätsmedizin Berlin

71 PUBLICATIONS 3,278 CITATIONS

[SEE PROFILE](#)



Matthias Treder

Cardiff University

70 PUBLICATIONS 2,592 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Cam-CAN project: uncovering brain-behaviour relationships across the lifespan [View project](#)



Age-Related Delay in Visual and Auditory Evoked Responses is Mediated by White- and Gray-matter Differences [View project](#)

EEG potentials predict upcoming emergency brakings during simulated driving

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

2011 J. Neural Eng. 8 056001

(<http://iopscience.iop.org/1741-2552/8/5/056001>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 134.226.252.160

The article was downloaded on 04/10/2011 at 15:21

Please note that [terms and conditions apply](#).

EEG potentials predict upcoming emergency brakings during simulated driving

Stefan Haufe^{1,2}, Matthias S Treder¹, Manfred F Gugler³,
Max Sagebaum⁴, Gabriel Curio³ and Benjamin Blankertz^{1,2,4}

¹ Machine Learning Group, Department of Computer Science, Berlin Institute of Technology, Franklinstraße 28/29, D-10587 Berlin, Germany

² Bernstein Focus Neurotechnology, Berlin, Germany

³ Neurophysics Group, Department of Neurology, Campus Benjamin Franklin, Charité University Medicine Berlin, D-12203 Berlin, Germany

⁴ Intelligent Data Analysis Group, Fraunhofer Institute FIRST, Kekuléstraße 7, D-12489 Berlin, Germany

E-mail: stefan.haufe@tu-berlin.de

Received 8 March 2011


Accepted for publication 1 July 2011

Published 28 July 2011

Online at stacks.iop.org/JNE/8/056001

Abstract

Emergency braking assistance has the potential to prevent a large number of car crashes. State-of-the-art systems operate in two stages. Basic safety measures are adopted once external sensors indicate a potential upcoming crash. If further activity at the brake pedal is detected, the system automatically performs emergency braking. Here, we present the results of a driving simulator study indicating that the driver's intention to perform emergency braking can be detected based on muscle activation and cerebral activity prior to the behavioural response. Identical levels of predictive accuracy were attained using electroencephalography (EEG), which worked more quickly than electromyography (EMG), and using EMG, which worked more quickly than pedal dynamics. **A simulated assistance system using EEG and EMG was found to detect emergency brakings 130 ms earlier than a system relying only on pedal responses.** At 100 km h⁻¹ driving speed, this amounts to reducing the braking distance by 3.66 m. This result motivates a neuroergonomic approach to driving assistance. Our EEG analysis yielded a characteristic event-related potential signature that comprised components related to the sensory registration of a critical traffic situation, mental evaluation of the sensory percept and motor preparation. While all these components should occur often during normal driving, we conjecture that it is their characteristic spatio-temporal superposition in emergency braking situations that leads to the considerable prediction performance we observed.

 Online supplementary data available from stacks.iop.org/JNE/8/056001/mmedia

(Some figures in this article are in colour only in the electronic version)

In our increasingly complex world, many safety-critical systems are operated by computers. Wherever human control is required, it may constitute a safety risk as human failures can have dramatic consequences. Traffic accidents, for instance, rank third among the causes of death in the USA and are largely caused by human errors [1].

Driving assistance systems aim to complement human control and can prevent failures in several ways. A recent development is hybrid approaches combining information from vehicle/surround sensors and human behaviour. Basic safety measures are adopted once external (radar or laser) sensors indicate a potential upcoming crash. If further 'panic'

activity at the brake pedal is detected, it is interpreted as the driver's confirmation of the criticality of the situation. This allows the system to go into an emergency braking procedure as soon as the brake pedal is touched by the driver, which saves time. However, the brake pedal response is only the very last event in the cascade of behavioural responses triggered during an emergency braking situation. Consequently, there have been attempts to retrieve the driver's braking intent earlier by considering additional behavioural inputs, such as gas pedal release, steering angle, foot position and head movements [2, 3].

This study takes the next two steps in monitoring human behaviour during driving. First, by measuring myoelectric (EMG) activity caused by muscle tension in the lower leg, we detect leg motion before it can be registered by pedal sensors. The existence of a characteristic EMG pattern preceding crashes has recently been noted [4], but no evaluation of its predictiveness regarding emergency braking detection has yet taken place. Second, more interestingly, for the first time we tap the driver's intention at the time and locus of its very generation—the human brain. Observable behaviour in an emergency braking situation is preceded by a cascade of neural events, including sensory registration, mental evaluation and motor preparation. As witnessed by the recent expansion of the field of neuroergonomics (for an overview, see [5]), monitoring cerebral activity during realistic tasks affords unprecedented insights into human functioning. The advent of powerful machine learning and signal processing techniques, largely developed in the context of brain–computer interfacing, makes it possible to reliably decode and track mental states non-invasively in real time using electroencephalography (EEG) [6–9].

To date, neurophysiological research for safety-critical applications has mainly focused on the detection of fatigue [10–17] and high mental workload [18]. The present study is the first one to investigate EEG correlates of intended emergency braking. We conducted a driving simulator study ($N = 18$) in which participants had to follow a computer-controlled lead vehicle within 20 m distance at a speed of 100 km h⁻¹. The simulation included sharp curves and dense oncoming traffic to make the driving task both realistic and demanding. Occasionally, emergency braking situations were induced by means of a rapid braking of the lead vehicle.

Under a unified analysis framework, we assessed the extent to which features derived from different input channels characterize emergency braking situations. Besides neurophysiological recordings, these channels also comprised the technical parameters used in commercial assistive systems^{5,6}. The results reported here were obtained in four consecutive steps. First, univariate features were investigated, leading to the discovery of a characteristic event-related potential (ERP) signature preceding executed emergency brakings. Using a classification approach, we then assessed how much better critical traffic situations can be distinguished from normal driving when the prediction is based on multivariate spatio-temporal instead of univariate

features. Both analyses were carried out separately for each channel and each stage of emergency braking situation. Thus, it is possible to compare channels in terms of their predictiveness at each stage. In the third step, we investigated potential redundancy among input channels by comparing the performance of a model that integrates all channels to that of models that exclude EEG and/or EMG. In the last step, we applied prototypical emergency braking intention detectors, either integrating neurophysiology or not, to the data in a pseudo-online fashion. This gives us the opportunity to estimate and compare the utility of both approaches in practice.

1. Material and methods

1.1. Experimental setup

Eighteen healthy right-handed participants (age 30.6 ± 5.4 years, four females) participated in the study. All possessed a driver's licence and had normal or corrected-to-normal vision. Subjects were paid for their participation. The experiment was conducted in accordance with the Declaration of Helsinki and written informed consent was given by all participants. The participants were seated in a driving simulator rack (Virtual Performance Parts, Germany) in front of a wide screen composed of three 19" monitors (the experimental apparatus can be seen in figure 1). On the screen, a customized version of the open-source racing software TORCS (The Open Racing Car Simulator⁷) was shown.

The participants' task was to drive a virtual racing car using the steering wheel and gas/brake pedals (automatic clutch), and to tightly follow a computer-controlled lead vehicle at a driving speed of 100 km h⁻¹. While the participants were within the desired maximal distance of 20 m, the lead vehicle occasionally (20–40 s inter-stimulus-interval, randomized) decelerated abruptly to between 60 and 80 km h⁻¹ (randomized). The manoeuvre was accompanied by a flashing of the brakelights. The driver was instructed to perform immediate emergency braking in these situations in order to avoid a crash. In the following, we refer to the onset of the lead vehicle's braking (and brakelight flashing) as the *stimulus*, while the first notable post-stimulus brake pedal deflection is called the *response*. Three seconds after a deceleration, the lead vehicle accelerated again to 100 km h⁻¹. When the distance between both vehicles exceeded 100 m (e.g. after crashes), the lead vehicle would wait for the driver's own vehicle to catch up. The lead vehicle would only ever brake to induce an emergency situation. The distance between vehicles was indicated by a coloured circle next to the speedometer, which was green for distances under 20 m and otherwise yellow.

The setting included oncoming traffic, such that the participants did not have the opportunity to avoid a potential accident by switching to the neighbouring lane. Three blocks (45 min each) of driving were conducted with rest periods of 10–15 min in between.

⁵ http://en.wikipedia.org/wiki/PreCrash_system.

⁶ http://en.wikipedia.org/wiki/Brake_Assist.

⁷ <http://torcs.sourceforge.net/>.



Figure 1. Snapshot of the experimental setup.

1.2. Data acquisition and preprocessing

EEG was acquired at 1000 Hz sampling frequency (low cutoff 0.1 Hz, high cutoff 250 Hz) from 64 scalp sites (extended international 10–20 system [19], reference at nose) using Ag/AgCl electrodes mounted on a cap (Easycap, Germany). Furthermore, EMG was recorded using a bipolar montage at the tibialis anterior muscle and the knee of the right leg. The impedances were below 20 k Ω for EEG electrodes and below 50 k Ω for the EMG.

The EEG and EMG signals were amplified and digitized using BrainAmp hardware (BrainProducts, Germany). Technical and behavioural markers such as stimulus onset times, brake and gas pedal deflection, acceleration of the lead vehicle and the driver's own vehicle, as well as the distance between vehicles, were provided by the TORCS software at 67 Hz sampling rate. Braking response times were defined based on the first noticeable (above noise-level) braking pedal deflection after an induced braking manoeuvre.

A schematic illustration of the data analysis process is provided in the supplementary material available from stacks.iop.org/JNE/8/056001/mmedia. All data analysis was conducted using Matlab (The Mathworks, USA). The EEG data were lowpass-filtered (tenth-order causal Chebychev type II filter) at 45 Hz. The EMG data were bandpass-filtered between 15 and 90 Hz (sixth-order causal Elliptic filter) with an additional notch filter (second-order digital) at 50 Hz for removing line noise, and rectified. Physiological and technical channels were synchronized and (causally) down-/upsampled to a common sampling rate of 200 Hz. This yielded one multivariate multi-modal time series per subject, from which we extracted the parts that were recorded during induced emergency situations, as well as parts reflecting normal driving.

Target situations were defined as those in which the braking response was given no earlier than 300 ms and no later than 1200 ms after the lead vehicle's braking onset.

Target segments were obtained in a stimulus-aligned version ([−300 ms, 1200 ms] intervals around the lead vehicle's brakelight flashing), as well as a response-aligned one ([−1300 ms, 200 ms] around the braking response). Note that most of the subsequently described analyses were carried out both for stimulus- and response-aligned data. However, here we only report results obtained from stimulus-aligned data, while the corresponding results obtained from response-aligned data are provided as supplementary material available from stacks.iop.org/JNE/8/056001/mmedia.

Non-targets were obtained by collecting all data blocks (1500 ms duration, 500 ms equidistant offset) that were at least 3000 ms apart from any stimulus. Baseline correction of EEG data was performed segment-wise by subtracting the average EEG amplitude in the first 100 ms.

The two target datasets (stimulus- and response-aligned segments), as well as the set of non-targets, were split into training and test parts, such that the training sets contain only data from the first half of driving and the test sets only data from the second half. Training data were only used to tune certain parameters, while the reported results were exclusively obtained on test data. Compared to an interleaved splitting scheme, our sequential validation approach ensures that the presented features are robust against slowly fluctuating processes such as fatigue, which is important regarding the applicability of our system in practice. The average number of targets and non-targets contained in the test sets were $N_t = 99 \pm 20$ and $N_{nt} = 6742 \pm 213$, respectively.

1.3. Event-related potentials and area under the curve analysis

We computed grand-average emergency-related signals by taking the arithmetic mean of EEG and auxiliary (technical and peripher-physiological) signals across the extracted target segments of all 18 subjects. Moreover, we investigated

how well univariate features—that is, single-time single-sensor readings—classify segments as targets or non-targets. This was done by evaluating the area under the specificity–sensitivity (i.e. receiver-operating characteristics) curve (AUC), obtained by varying the detection threshold [20]. The AUC is a non-parametric statistic, which is equivalent to Mann–Whitney U and Wilcoxon’s rank-sum tests for assessing differences in two populations, and Kendall’s τ measuring correlation between feature values and binary class labels [21]. Besides, it is an estimate of the probability that a randomly chosen target has a higher feature value than a randomly chosen non-target [20, 21].

To compute AUC scores, non-targets was sampled from the normal driving periods and compared to the targets. The AUC is symmetric around 0.5. That is, scores above (below) 0.5 indicate that targets have higher (lower) feature values than non-targets, while 0.5 is attained for completely uninformative features. Grand-average AUC scores were calculated as the arithmetic mean across subjects.

Based on the AUC, we defined a class separation (CS) score $CS = 2|AUC - 0.5|$, which varies between 0 for no CS and 1 for perfect CS.

1.4. Classification

We were also interested in the area under curve attainable at each stage of emergency braking when the (spatio-) temporal dynamics in the preceding interval could be used for prediction. To this end, we constructed additional sets of target segments with a constant length of 1500 ms. The endpoints of these segments were varied in steps of 20 ms from –200 to 1180 ms relative to the stimulus and from –1200 to 180 ms relative to the response. Furthermore, separate datasets were created for the seven channels EEG, EMG, Gas, Brake, deceleration of lead vehicle and own vehicle (abbreviated DecL and DecO) and distance (Dist), the combination of the two traffic-related channels DecL and Dist, as well as for four combinations of channels reflecting the driver’s emergency braking intent. These combinations were EEG+EMG+Gas+Brake+DecO, EMG+Gas+Brake+DecO (no EEG), EEG+Gas+Brake+DecO (no EMG) and Gas+Brake+DecO (no physiology at all).

For each channel and target/non-target combination, ten discriminating time intervals (not necessarily of equal length) were determined using a heuristic based on biserial correlation scores [22]. EEG electrodes showing extraordinarily high spectral power (in a broad 5–40 Hz frequency band) or abnormally high variability in power on training data were discarded. The criterion used to identify abnormal observations was based on the interquartile range. The average signals of the remaining electrodes in the selected time intervals were stacked into feature vectors and fed into regularized linear discriminant analysis (RLDA) [23, 24] classifiers, along with the corresponding class label.

The discriminant function of RLDA has the form $y(\mathbf{x}) = \mathbf{w}^\top \mathbf{x} + b$, where \mathbf{x} is a feature vector and $\mathbf{w} = \hat{\Sigma}^{-1}(\hat{\mu}_2 - \hat{\mu}_1)$ is a vector that is orthogonal to the class-separating hyperplane. The vectors $\hat{\mu}_{1/2}$ and the matrix $\hat{\Sigma}$ are empirical estimators of the two feature class means and the common feature covariance

matrix, while the scalar b is the decision threshold. Without loss of generality, we define $\hat{\mu}_1$ to be the mean of the non-target samples and $\hat{\mu}_2$ to be the mean of the targets. We applied the transformation $\mathbf{w} \leftarrow 2\mathbf{w}/[\mathbf{w}^\top(\hat{\mu}_2 - \hat{\mu}_1)]$ and set $b = 1 - \mathbf{w}^\top \hat{\mu}_2$ in order to achieve that $y(\hat{\mu}_1) = -1$ and $y(\hat{\mu}_2) = 1$.

To mitigate numerical instability when inverting singular or poorly conditioned covariance matrices, the shrinkage technique [25, 26] was applied. That is, the true covariance matrix was approximated by a convex combination of the empirical covariance and the identity matrix. The shrinkage coefficient was chosen such that the mean-squared-error between $\hat{\Sigma}$ and the true unknown covariance matrix was minimized [27, 28].

Data from the first half of driving were used for selecting time intervals and for estimating class means and covariance matrices. The trained discriminant functions were applied to the corresponding test data and the resulting outputs were used to calculate unbiased estimates of the AUC.

1.5. Statistical testing

For sufficiently large numbers of targets and non-targets ($N_t + N_{nt} > 10$), the AUC is approximately normal distributed with variance $\text{Var}(AUC, N_t, N_{nt}) = [AUC(1 - AUC) + (N_t - 1)(Q_1 - AUC^2) + (N_{nt} - 1)(Q_2 - AUC^2)]/[N_t N_{nt}]$, where $Q_1 = AUC/(2 - AUC)$ and $Q_2 = 2AUC^2/(1 + AUC)$ [21]. We used this to test whether AUC scores are significantly different from 0.5 in the grand average. This was done by calculating subject-wise z -scores $z_k = (AUC_k - 0.5)/\sqrt{[\text{Var}(0.5, N_t, N_{nt})]}$, summing them up to a grand-average z -score $\bar{z} = (z_1 + z_2 + \dots + z_{18})/\sqrt{18}$ and performing a two-sided z -test to obtain grand-average p -values. Differences between two AUC scores, $AUC_{k,j}$ and $AUC_{k,i}$, were tested by treating $z_{k,ji} = (AUC_{k,j} - AUC_{k,i})/\sqrt{[\text{Var}(AUC_{k,j}, N_t, N_{nt}) + \text{Var}(AUC_{k,i}, N_t, N_{nt})]}$ as standard normal distributed quantities and proceeding as above to obtain grand-average z -scores and p -values. In order to compare one AUC score to multiple others, the minimal absolute value of the z -scores obtained in the binary comparisons was used to calculate the p -value.

The p -values obtained in exploratory analyses were Bonferroni-corrected to account for multiple hypothesis testing [29]. The correction factor used in the ERP analysis was 280 (time instants) \times 57 (electrodes) $= 15\,960$. In the analysis of auxiliary (non-neural) parameters it was 280 (time instants). All p -values smaller than 10^{-10} are reported as $p \approx 0$. Statistical significance is assumed if (corrected) p -values are smaller than 0.05.

1.6. Pseudo-online emergency braking detection

We constructed driver-specific emergency braking detectors that were designed to work on continuous data without knowledge about the timing of stimulus and response. Since stimulus-related signal components (such as the VEP induced by the brake-light flashing) are best recognized by classifiers trained on stimulus-aligned target intervals, and response-related signal components (such as motor readiness potential) are best recognized by classifiers trained on response-aligned target intervals, both types of classifier were used for the

detector. We selected those RLDA classifiers trained on response-aligned target intervals ending between 260 ms pre-response and the response onset. These are called *response-classifiers*. Analogously, we designated a set of *stimulus-classifiers* that had been trained to detect the stimulus-aligned targets. To account for the variable response time, we used classifiers trained on segments, the endpoints of which ended between $P_{25} - 260$ ms and P_{75} (P_{25} and P_{75} denoting the 25th and 75th percentile of the pooled response times of all subjects, respectively).

We calculated continuous classifier outputs on test data by extracting the predetermined features and applying the pre-selected RLDA discriminant functions every 20 ms. The classifier outputs were divided by their standard deviation, as estimated from training data. For any point in time, the maximum output of the stimulus-classifiers was added to the maximum output of the response-classifiers to yield a continuous detector time series. Note that the calculation of this time series does not require any knowledge about the timing of stimulus and/or response, which is not available in practice. Rather, each time the detector is evaluated, the ‘most likely’ timings of stimulus and response are estimated by means of the maximum operation.

We considered four different rules for generating ‘alarms’ from detector times series. In *driver-based detection*, an alarm was issued whenever the driver-intent-related detector time series exceeded a critical threshold. In *driver-confirmed detection*, alarms were raised only when there were two consecutive (at most 1000 ms after another) alarms derived from traffic- and driver-intent-related channels. Braking intent was measured either based on behavioural channels (Gas, DecO and Brake, denoted BHV) alone, or on combined physiological (EEG and EMG, denoted PHY) and behavioural channels. Traffic-related channels included DecL and Dist.

The percentage of detected critical situations, the average alarm time relative to the response onset and the number of false alarms was calculated for every possible detection threshold. A correct detection was counted if an alarm was raised within 1500 ms post-stimulus of a target event, otherwise the trial was classified as a miss. Alarm times were assessed for correct detections. Alarms occurring outside any 5000 ms post-stimulus interval were counted as false alarms.

Variation of the threshold affected all three performance measures, not necessarily in a monotonic way. Nevertheless, to be able to compute a grand average, false alarm rates and the response times of each participant were treated as functions of the detection rate and evaluated at 100 equidistant points using linear interpolation.

2. Results

2.1. Driving performance

The participants stayed $93 \pm 7\%$ (mean \pm standard deviation) of the time within the required maximal distance of 20 m behind the lead vehicle. During these periods, 225 ± 17 critical (emergency braking) situations were artificially induced. The distribution of pooled response times in target situations

was skewed with percentiles $P_5 = 505$ ms, $P_{25} = 595$ ms, $P_{50} = 665$ ms (median), $P_{75} = 750$ ms and $P_{95} = 910$ ms. Collisions with the lead vehicle occurred in $17 \pm 10\%$ of the critical situations.

2.2. Behavioural signature of emergency braking

In general, the emergency situations induced by our paradigm were not resolvable without immediate emergency braking. In most cases the participants had to release the gas pedal and relocate the right foot in order to press the brake pedal. This behavioural pattern was observed indirectly through gas and brake pedal deflections (designated as Gas and Brake channels, respectively). Moreover, **the foot movement was accompanied by a burst of EMG activity at the tibialis anterior muscle, while the pedal dynamics naturally lead to measurable deceleration of the driver’s own vehicle** (DecO channel).

We report the relative post-stimulus times at which moderate class separation ($CS > 0.1$) was attained. For the EMG channel this was the case 335 ms post-stimulus. The EMG was followed by the Gas (430 ms), DecO (470 ms) and Brake (595 ms) channels. **The maximal CS scores were 0.992 for Brake (at 1100 ms), 0.940 for EMG (at 645 ms), 0.834 for DecO (at 840 ms) and 0.726 for Gas (at 630 ms).** All reported effects are significant (AUC scores significantly different from 0.5 with $|z| > 14.9$, $p_{\text{corr.}} \approx 0$).

An illustration of the temporal development of the stimulus-aligned grand-average feature and AUC scores is shown in figure 2(A). The figure also shows the progression of the deceleration of the lead vehicle (DecL channel) and the distance between vehicles (Dist channel). Note that these two variables are either entirely (DecL) or partly (Dist) experimentally controlled, which causes them to be (artificially) correlated with emergency braking events (DecL is almost always zero during inter-stimulus-intervals, while Dist is always smaller than 20 m during emergency situations).

The CS scores of the two indexes exceeded 0.2 at 10 ms (DecL) and 360 ms (Dist) post-stimulus, respectively. The maximal CS observed for DecL was 0.790 (at 70 ms), while it was 0.784 for Dist (at 1200 ms). These effects are also significant ($|z| > 13.8$, $p_{\text{corr.}} \approx 0$). However, the estimated predictive power of both channels is partly attributable to our experimental design and may not be achievable in practice.

2.3. ERP signature of emergency braking

The grand-average stimulus-aligned analysis of the EEG revealed a number of distinct spatio-temporal ERP complexes. Among the three most prominent was **a symmetric negative deflection in occipito-temporal areas**. This component peaked 345 ms post-stimulus at electrode P9 and 350 ms post-stimulus at electrode P10. The peak amplitudes were $-7.4 \mu\text{V}$ (P9) and $-9.4 \mu\text{V}$ (P10). The CS score exceeded 0.2 at 265 ms for P9 and 255 ms for P10. The maximal CS observed for P9 was 0.290 (at 345 ms), while for P10 it was 0.308 (at 350 ms).

The second ERP component was a negativity at central scalp sites, which was most pronounced around electrode FCz. The EEG peak time at that electrode was 1120 ms

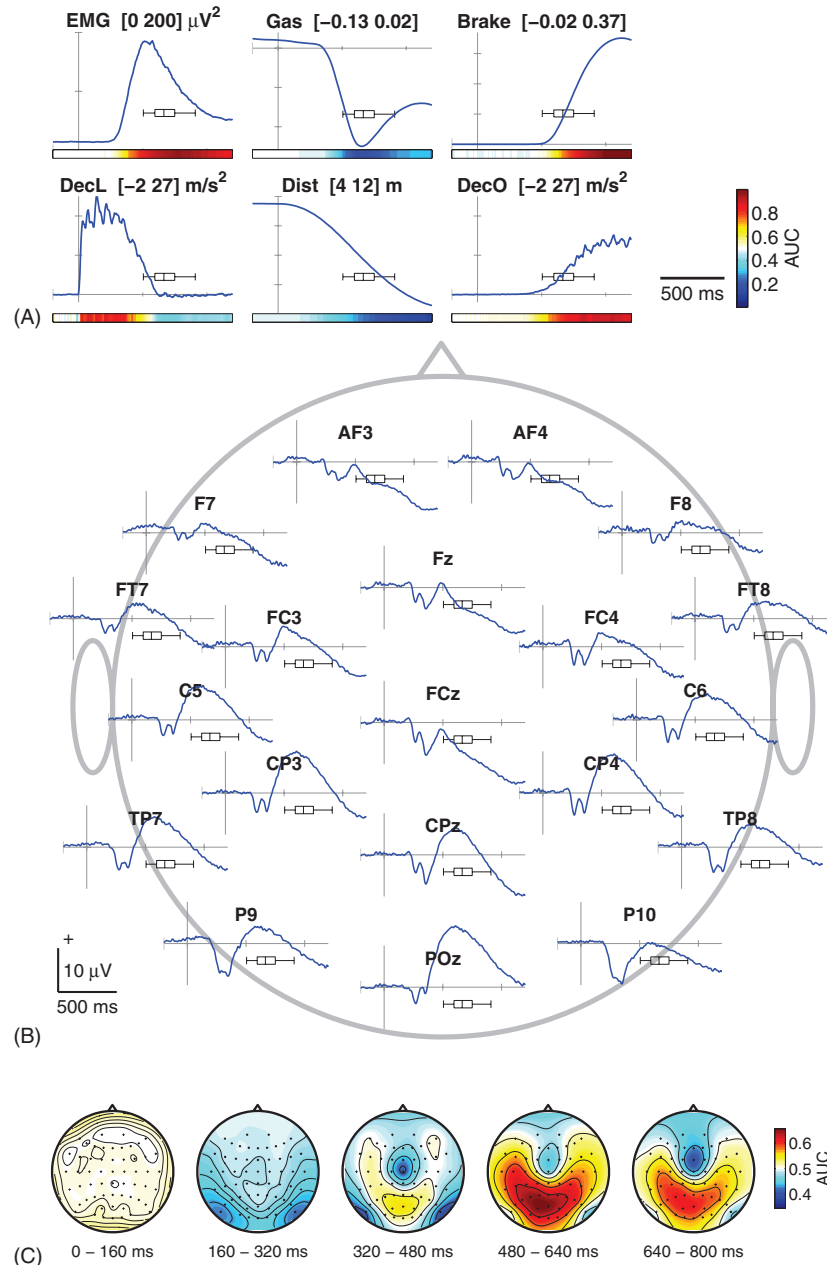


Figure 2. Grand-average stimulus-aligned signals before and during emergency braking. STIM denotes the onset of braking (brakelight flashing) of the lead vehicle. (A) Technical, behavioural and EMG channels. Upper part: EMG at the tibialis anterior muscle, gas and brake pedal deflections. Lower part: deceleration of the lead vehicle (DecL) and the driver's own vehicle (DecO), and distance between vehicles. Colour-coded bars depict grand-average area under the curve (AUC) scores measuring differences in feature values between target (critical) and non-target (normal driving) situations. Yellow and red colour (AUC > 0.5) indicates that a feature attains higher values in targets than in non-targets, while cyan and blue colour (AUC < 0.5) indicates the opposite case. The distribution of pooled braking response times is indicated by box plots showing 5th, 25th, 50th (median), 75th and 95th percentile. (B) Grand-average ERP curves. (C) Topographical maps of grand-average AUC scores calculated from mean EEG activity (ERP) in five temporal intervals.

post-stimulus, while the peak amplitude was $-14.7 \mu V$. The CS score exceeded 0.2 at 725 ms and attained a maximum of 0.298 at 1125 ms.

Finally, a positive deflection could be seen in spatially extended centro-parietal areas centred around electrode CPz. The peak amplitude of this component was $14.1 \mu V$, and was reached 610 ms post-stimulus. The CS score exceeded 0.2 at 450 ms post-stimulus, and the CS was maximal at

580 ms with CS = 0.3. Again, all these effects are significant ($|z| > 14.6$, $p_{\text{corr.}} \approx 0$).

The superposition of the three components gave rise to a characteristic spatio-temporal ERP sequence, the temporal and spatial aspects of which are depicted in figure 2 (panels B and C). Panel B shows time-resolved ERP curves, while panel C depicts topographical (scalp) maps of average AUC scores in five subsequent time intervals of 160 ms length.

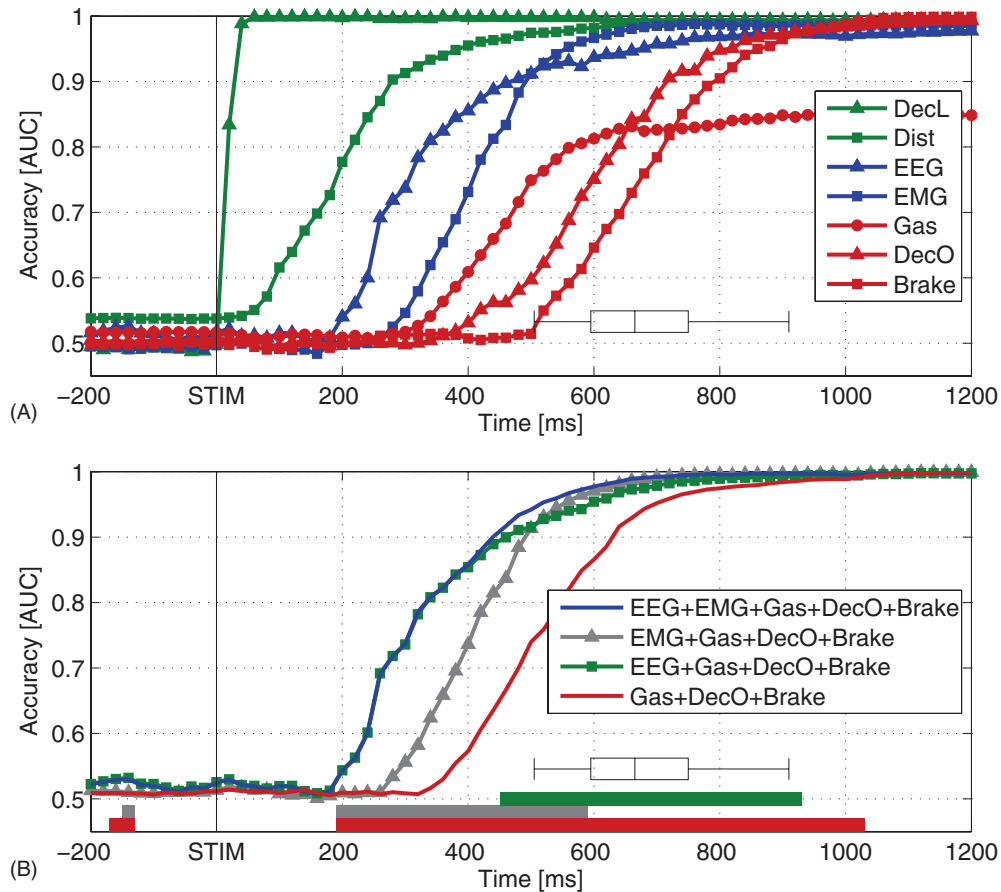


Figure 3. Grand-average area under the curve (AUC) scores calculated from the outputs of linear classifiers that were optimized to distinguish normal driving intervals from stimulus-aligned target intervals representing different stages of emergency braking situations. STIM denotes the onset of braking (brakelight flashing) of the lead vehicle. The distribution of pooled braking response times is indicated by box plots showing 5th, 25th, 50th (median), 75th and 95th percentile. Classification was based on (spatio-) temporal features observed prior to the respective decision points. (A) Performance of single input channels. Green curves (traffic-related channels): deceleration of the lead vehicle (triangle markers), distance between vehicles (square markers). Blue curves (physiological channels): EEG (triangle markers), EMG (square markers). Red curves (behaviour-related channels): gas pedal deflection (circle markers), deceleration of the driver's own vehicle (triangle markers), brake pedal deflection (square markers). (B) Performance of different combinations of input channels. Blue: EEG+EMG+Gas+DecO+Brake (all driver-intent-related features). Grey, triangle markers: EMG+Gas+DecO+Brake (no EEG). Green, square markers: EEG+Gas+DecO+Brake (no EMG). Red: Gas+DecO+Brake (no physiology at all).

2.4. Multivariate decoding

We analysed the extent to which the CS could be improved using the observed spatio-temporal structure in the data. To this end, we evaluated the AUC of outputs of linear classifiers trained to distinguish targets from non-targets using multivariate (spatio-) temporal features. In the evaluation, classifiers could only use information that was available at the time of prediction. Note that AUC scores obtained from classifier outputs are always greater than or not significantly different from 0.5. Furthermore, the AUC scores increase monotonically over time, as more data become available.

The results for the seven previously considered channels are presented in figure 3(A). The CS score exceeded 0.2 after 20 ms (DecL), 100 ms (Dist), 240 ms (EEG), 340 ms (EMG), 400 ms (Gas), 520 ms (DecO) and 560 ms (Brake) post-stimulus, respectively. That is, compared to the aforementioned results on univariate features, the same classification accuracy was obtained at an earlier stage for the Dist, EEG and Gas channels. Moreover, the maximal CS was

increased for all channels. The channels DecL and Brake attained maximal scores of 0.999, while the maximum was 0.988 for the DecO, 0.978 for the Dist, 0.976 for the EMG, 0.954 for the EEG and 0.7 for Gas channel. The benefit of the multivariate approach was most pronounced for the EEG channel, where the maximal score attained for a univariate feature (0.3 at the electrode POz) had already been exceeded 320 ms earlier (at 260 ms post-stimulus).

Among the five channels that reflect the driver's intention to go into emergency braking (EEG, EMG, Gas, DecO and Brake), the EEG channel performed best from 200 to 480 ms post-stimulus (significant with $p < 0.05$; largest difference at 320 ms, $z_{\min} = 23.4$, $p \approx 0$). The EMG channel performed significantly better than all other channels from 560 to 940 ms (largest difference at 660 ms, $z_{\min} = 5.3$, $p < 10^{-6}$), while Brake performed best from 1100 to 1200 ms (largest difference at 1100 ms, $z_{\min} = 3.4$, $p < 0.005$).

In a complementary analysis, classifiers trained on combined driver-intent-related channels (EEG, EMG, Gas,

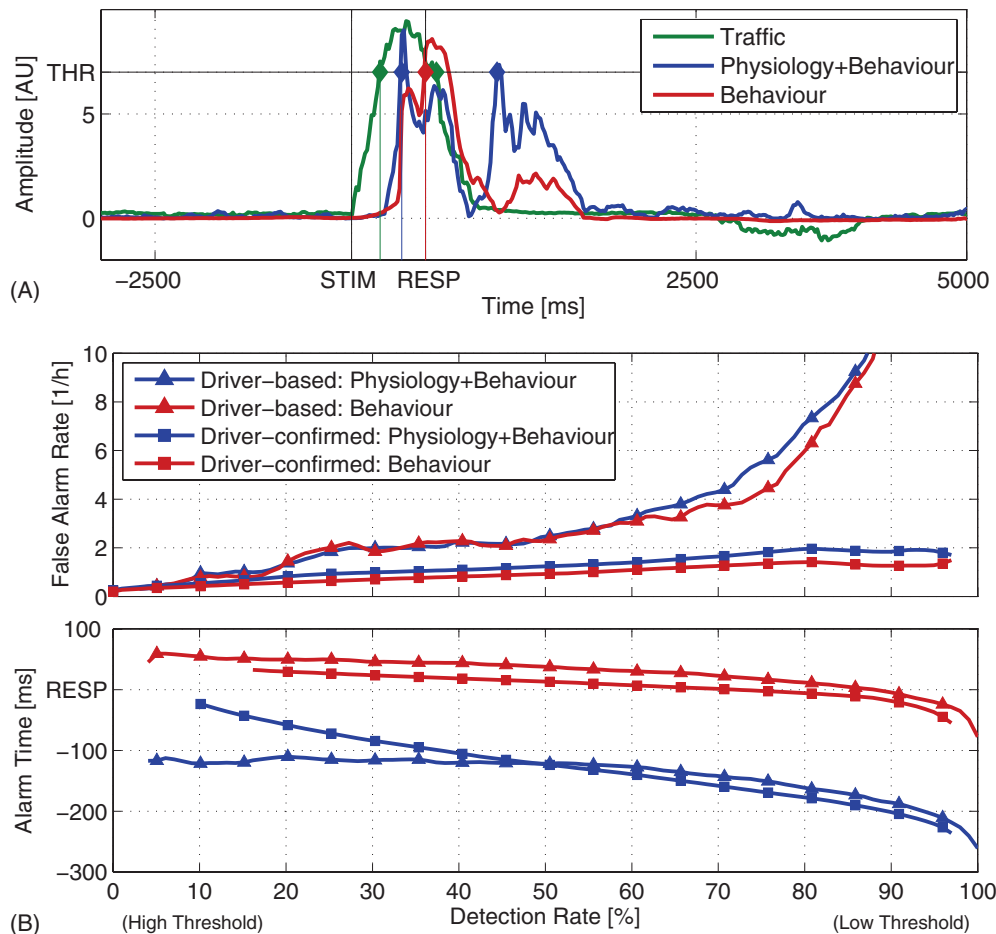


Figure 4. (A) Emergency braking detector time series derived from traffic- (green), behaviour- (red) and combined physiology- and behaviour-related channels (blue). STIM denotes the onset of braking (brakelight flashing) of the lead vehicle. A hypothetical detection threshold (THR) is represented by a thin black line, and its exceedance by the detector time series is marked by coloured dots and thin vertical lines. (B) Grand-average performance of emergency braking detectors that were continuously evaluated on test data. The plot shows alarm time (relative to the braking response onset) and false alarm rate as functions of the detection rate. Driver-based approach (triangle markers): detection based solely on behaviour- (red) or combined physiology- and behaviour-related channels (blue). Driver-confirmed approach (square markers): detection using additional evidence from traffic-related channels.

DecO and Brake) were compared to classifiers omitting either the EMG or EEG channel, or both. The results are shown in figure 3(B). The performance without EEG dropped significantly between 200 and 560 ms post-stimulus (largest difference at 320 ms, $z = 23.3$, $p \approx 0$). Without the EMG channel, there was a significant performance drop between 460 and 900 ms post-stimulus (largest difference at 580 ms, $z = 5.3$, $p < 10^{-6}$). When both physiological channels were left out, the performance was significantly reduced between 200 and 1000 ms post-stimulus (largest difference at 360 ms, $z = 35.1$, $p \approx 0$). Note that, although in principle possible, the combination of features did not raise the classification performance above that of the best performing included feature at any time. For example, the AUC related to the combined EEG+EMG+Gas+Brake+DecO channels resembles that of the EEG channel in the interval from 200 to 400 ms post-stimulus. These results confirm that EEG and EMG contain information that is independent of the entirety of the remaining channels in the sense that it can neither be reproduced nor enhanced by combinations of these channels.

2.5. Pseudo-online emergency braking detection

The performance of stimulus- and response-uninformed emergency-braking detectors was assessed in terms of the percentage of correct alarms, the alarm time relative to the braking response and the frequency of false alarms for each possible value of the detection threshold. The grand-average results are shown in figure 4(B) as a function of the detection rate. Figure 4(A) shows the temporal development of the traffic-, behaviour- and combined behaviour- and physiology-derived detector time series during a single emergency braking event. We also provide a video of a 10 min driving scene and synchronous pseudo-online detector time series, a sample of which is provided as supplemental material available from stacks.iop.org/JNE/8/056001/mmedia.

In the driver-based approach, the alarm time relative to the response when using behaviour-related channels was 29 ms at a detection rate of 20%. At a detection rate of 80% the relative alarm time dropped to -6 ms. The inclusion of physiological channels (EEG and EMG) led to earlier alarms (-58 ms at 20% detection rate and -177 ms at 80%). In the

driver-confirmed approach, alarm times were determined by the later (driver-related) detector time series and hence similar to those of the solely driver-based approach. The observed alarm times were 50 ms (BHV) and -110 ms (BHV+PHY) at 20% detection rate. At 80% detection rate, we measured 12 ms (BHV) and -160 ms (BHV+PHY). The inclusion of EEG and EMG channels yielded an improvement of relative alarm time (measured as the mean value across all detection rates occurring) of 167 ± 85 ms for the driver-based approach and 132 ± 55 ms for the driver-confirmed approach.

Omitting EEG/EMG channels did not lead to significantly different false alarm rates. In the driver-based approach, it increased from 1.37 h^{-1} (at 20% detection rate) to 7.05 h^{-1} (at 80% detection rate) for combined physiological and behaviour-related channels and from 1.45 to 5.91 h^{-1} for solely behaviour-related channels. The false alarm rate was lower in the driver-confirmed approach: it did not exceed 1.96 h^{-1} for combined physiological and behaviour-related channels and 1.49 h^{-1} for solely behaviour-related channels, even in regimes with high detection rates.

3. Discussion

3.1. Neurophysiological relevance

The present investigation revealed a characteristic spatio-temporal ERP pattern observable before and during emergency brakings. This signature was found to be a superposition of at least three ERP components reflecting the cognitive processes between stimulus perception and actual braking. The neurophysiological relevance of the individual components can be deduced from their spatio-temporal properties as follows.

The lead vehicle's braking manoeuvre involves flashing brakelights, which is a strong visual stimulus. Moreover, due to the rapidly decreasing distance to the braking vehicle in front, the visual input is magnified continuously over time. The initial negative deflection observed in occipito-temporal EEG electrodes may therefore be attributed to visual stimulus perception.

A broad parieto-occipital positivity is typical for a P300 component. Classically related to higher-order processing of an 'oddball' stimulus [30, 31], the occurrence of a P300 complex is explained here by the relative rareness and importance of the critical traffic situation.

Finally, the central negativity is attributed to planning processes in the motor system before and during the act of switching the right foot from gas to braking pedal. The preparatory (pre-movement) component of this negativity was first described in [32], where it was called 'Bereitschaftspotential' (readiness potential). The later part of the signal (starting from EMG onset) is better described as a 'movement-related potential'. In our case the Bereitschafts- and movement-related potentials are observed at central electrodes. This is consistent with the assumed generators in the foot-area of the motor cortex [33].

Normal driving situations frequently require sensory registration of stimuli, processing of unexpected situations or

motor preparation/execution. However, in our experiment the excellent predictive power of EEG-derived features indicated that it is mainly the sequence of these overlapping cerebral processes that robustly characterizes critical traffic situations.

3.2. Relevance for driving assistance systems

Assistance systems that require the driver to (re)act before initiating a safety program currently rely on behavioural markers such as brake pedal deflection and gas pedal release. While this approach ensures that safety measures are not taken against the driver's will, waiting for the driver's response can lead to a slow response in emergency situations. Therefore, in order to obtain a faster confirmation, our study suggests that it is feasible to detect a driver's *intention* to brake, which naturally precedes any observable actions. In fact, we found that EEG and EMG recordings reach the predictive accuracy of behavioural channels at earlier stages of the emergency situation. In principle, this paves the way for a neuroergonomic approach to driving assistance, including peripher-physiological sensors. We evaluated a simplistic implementation of such a system in our simulation environment and estimated that the time saved is around 130 ms. While this seems to be a tiny improvement, it may actually have a significant impact on accident prevention. Consider, for example, driving a car in dense traffic at 100 km h^{-1} . If an emergency braking manoeuvre is detected 130 ms earlier, the braking distance will be reduced by as much as 3.66 m, or the length of a compact car.

This study was conducted in an idealized simulation environment. Real-world driving, of course, yields a much more dynamic scenario comprising more complex and more diversified traffic situations. These may include, for example, switching traffic lights or changing speed limits. The distinction between normal and exceptional traffic events are more intricate under these circumstances. We are currently planning a real-world driving study using a similar experimental setup to test the transferability of our result to this setting.

A crucial issue to be investigated is the feasibility of conducting EEG and EMG measurements under realistic conditions. While myoelectricity is a very strong signal that is readily acquired using only two electrodes, the situation is more difficult for brain wave recordings. The EEG system has to cope with a multitude of artefacts that are stronger than the neural signals. For example, there are mechanical artefacts induced by electrode motion, which can result from unsteady direction of motion of the vehicle (e.g. when entering a sharp curve) or by head and body movements of the driver. Moreover, electrical activity of head muscles, e.g. caused by chewing, raising eyebrows or eye blinking might be more prevalent during real driving than under our controlled conditions. Presently, it is unclear whether (a) the signal quality in real-world recordings will be significantly degraded compared to the laboratory setting, (b) this would also lead to deteriorating detection performance and (c) in this case, advanced data analysis methods such as independent component analysis artefact reduction [34–36] may help to

overcome such problems. Our current system does not include any artefact-removal strategy apart from the omission of EEG sensors with apparent bad signal quality.

Another issue regarding EEG technology concerns its practical applicability. Current systems are based on electrode caps that are uncomfortable to wear, unattractive and involve the application of abrasive gel to the skin. Furthermore, a 64-electrode system as used here requires time-consuming preparation. However, significant advances have been made recently, e.g. in the development of dry electrodes [37]. With respect to wearing comfort, new miniature electrodes that can be mounted capless (using only one droplet of gel per electrode) promise to be virtually unnoticeable in practice [38]. Moreover, commercial wireless EEG systems are already available.

4. Conclusion

In this study, we have demonstrated the viability of a neuroergonomic approach to driving assistance which allows us to reduce response times by detecting a driver's intention to brake before any actions become observable. The implementation of this system in practice depends on advances in engineering and data analysis that are the focus of ongoing research.

Acknowledgments

We acknowledge financial support by the BMBF grant nos 01GQ0850, 01GQ0851, 01IB001A, DFG grant no MU 987/3-2 and the FP7-ICT Programme of the European Community, under the PASCAL2 Network of Excellence, ICT-216886. We thank Klaus-Robert Müller for valuable discussions, Arne Ewald for help in conducting experiments and Paul von Büna for proofreading the manuscript.

References

- [1] Subramanian R 2007 Motor vehicle traffic crashes as a leading cause of death in the USA, 2004 *NHTSA-Report* DOT HS 810 742
- [2] Trivedi M M and Cheng S Y 2007 Holistic sensing and active displays for intelligent driver support systems *Computer* **40** 60–8
- [3] McCall J C and Trivedi M M 2007 Driver behavior and situation aware brake assistance for intelligent vehicles *Proc. IEEE* **95** 374–87
- [4] Behr M, Poumarat G, Serre T, Arnoux P J, Thollon L and Brunet C 2010 Posture and muscular behaviour in emergency braking: an experimental approach *Accid. Anal. Prev.* **42** 797–801
- [5] Parasuraman R and Rizzo M 2008 *Neuroergonomics: The Brain at Work* 1st edn (New York: Oxford University Press)
- [6] Blankertz B, Dornhege G, Krauledat M, Müller K-R and Curio G 2007 The non-invasive Berlin brain–computer interface: fast acquisition of effective performance in untrained subjects *NeuroImage* **37** 539–50
- [7] Müller K-R, Tangermann M, Dornhege G, Krauledat M, Curio G and Blankertz B 2008 Machine learning for real-time single-trial EEG-analysis: from brain–computer interfacing to mental state monitoring *J. Neurosci. Methods* **167** 82–90
- [8] Blankertz B et al 2010 The Berlin brain–computer interface: non-medical uses of BCI technology *Front. Neurosci.* **4** 198
- [9] Haufe S, Tomioka R, Dickhaus T, Sannelli C, Blankertz B, Nolte G and Müller K-R 2011 Large-scale EEG/MEG source localization with spatial flexibility *NeuroImage* **54** 851–859
- [10] Beatty J, Greenberg A, Deibler W P and O'Hanlon J F 1974 Operant control of occipital theta rhythm affects performance in a radar monitoring task *Science* **183** 871–3
- [11] Kecklund G and Åkerstedt T 1993 Sleepiness in long distance truck driving: an ambulatory EEG study of night driving *Ergonomics* **36** 1007–17
- [12] Makeig S and Inlow M 1993 Lapses in alertness: coherence of fluctuations in performance and EEG spectrum *Electroencephalogr. Clin. Neurophysiol.* **86** 23–35
- [13] Makeig S and Jung T-P 1996 Tonic, phasic, and transient EEG correlates of auditory awareness in drowsiness *Cogn. Brain Res.* **4** 15–25
- [14] Jung T P, Makeig S, Stensmo M and Sejnowski T J 1997 Estimating alertness from the EEG power spectrum *IEEE Trans. Biomed. Eng.* **44** 60–9
- [15] Lal Saroj K L and Craig A 2001 A critical review of the psychophysiology of driver fatigue *Biol. Psychol.* **55** 173–94
- [16] Papadelis C, Chen Z, Kourtidou-Papadeli C, Bamidis P D, Chouvarda I, Bekiaris E and Maglaveras N 2007 Monitoring sleepiness with on-board electrophysiological recordings for preventing sleep-deprived traffic accidents *Clin. Neurophysiol.* **118** 1906–22
- [17] Schmidt E A, Schrauf M, Simon M, Fritzsche M, Buchner A and Kincses W E 2009 Drivers' misjudgement of vigilance state during prolonged monotonous daytime driving *Accid. Anal. Prev.* **41** 1087–93
- [18] Kohlmorgen J, Dornhege G, Braun M, Blankertz B, Müller K-R, Curio G, Hagemann K, Bruns A, Schrauf M and Kincses W 2007 Improving human performance in a real operating environment through real-time mental workload detection *Toward Brain–Computer Interfacing* ed G Dornhege, J del R Millán, T Hinterberger, D McFarland and K-R Müller (Cambridge, MA: MIT Press) pp 409–422
- [19] Sharbrough F, Chatrian G E, Lesser R P, Lüders H, Nuwer M and Picton T W 1991 American Electroencephalographic Society guidelines for standard electrode position nomenclature *J. Clin. Neurophysiol.* **8** 200–2
- [20] Fawcett T 2006 An introduction to ROC analysis *Pattern Recognit. Lett.* **27** 861–74
- [21] Hanley J A and McNeil B J 1982 The meaning and use of the area under a receiver operating characteristic (ROC) curve *Radiology* **143** 29–36
- [22] Blankertz B, Tomioka R, Lemm S, Kawanabe M and Müller K-R 2008 Optimizing spatial filters for robust EEG single-trial analysis *IEEE Signal Proc. Mag.* **25** 41–56
- [23] Fisher R A 1936 The use of multiple measurements in taxonomic problems *Ann. Eugen.* **7** 179–88
- [24] Duda R O, Hart P E and Stork D G 2000 *Pattern Classification* (New York: Wiley)
- [25] Friedman J H 1989 Regularized discriminant analysis *J. Am. Stat. Assoc.* **84** 165–75
- [26] Blankertz B, Lemm S, Treder M S, Haufe S and Müller K-R 2010 Single-trial analysis and classification of ERP components—a tutorial *NeuroImage*
- [27] Ledoit O and Wolf M 2004 A well-conditioned estimator for large-dimensional covariance matrices *J. Multivariate Anal.* **88** 365–411
- [28] Schäfer J and Strimmer K 2005 A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics *Stat. Appl. Genet. Mol. Biol.* **4** 32
- [29] Bonferroni C E 1936 Teoria statistica delle classi e calcolo delle probabilità *Pubblicazioni del R Istituto Superiore di*

- Scienze Economiche e Commerciali di Firenze* **8** 3–62
- [30] Chapman R M and Bragdon H R 1964 Evoked responses to numerical and non-numerical visual stimuli while problem solving *Nature* **203** 1155–7
- [31] Sutton S, Braren M, Zubin J and John E R 1965 Evoked-potential correlates of stimulus uncertainty *Science* **150** 1187–8
- [32] Kornhuber H H and Deecke L 1965 Hirnpotentialänderungen bei Willkürbewegungen und passiven Bewegungen des Menschen: Bereitschaftspotential und reafferente Potentiale *Pflügers Arch.* **284** 1–17
- [33] Penfield W and Boldrey E 1937 Somatic motor and sensory representation in the cerebral cortex of man as studied by electrical stimulation *Brain* **60** 389–443
- [34] Vigário R, Jousmäki V, Hämäläinen M, Hari R and Oja E 1998 Independent component analysis for identification of artifacts in magnetoencephalographic recordings (*Proc. NIPS vol 10*) (Cambridge, MA: MIT Press) pp 229–35
- [35] Jung T P, Makeig S, Humphries C, Lee T W, McKeown M J, Iragui V and Sejnowski T J 2000 Removing electroencephalographic artifacts by blind source separation *Psychophysiology* **37** 163–78
- [36] Ziehe A, Müller K-R, Nolte G, Mackert B-M and Curio G 2000 Artifact reduction in magnetoneurography based on time-delayed second-order correlations *IEEE Trans. Biomed. Eng.* **47** 75–87
- [37] Popescu F, Fazli S, Badower Y, Blankertz B and Müller K R 2007 Single trial classification of motor imagination using 6 dry EEG electrodes *PLoS ONE* **2** e637
- [38] Nikulin V V, Kegeles J and Curio G 2010 Miniaturized electroencephalographic scalp electrode for optimal wearing comfort *Clin. Neurophysiol.* **121** 1007–14