This is the author's final draft. The final publication is in Zander, T. O. et al. (2017), *Journal of Medical Robotics Research, 2*(1), 1750003. DOI: 10.1142/S2424905X17500039

Automated Task Load Detection with Electroencephalography: Towards Passive Brain-Computer Interfacing in Robotic Surgery

Abstract

Automatic detection of the current task load of a surgeon in the theatre in real time could provide helpful information, to be used in supportive systems. For example, such information may enable the system to automatically support the surgeon when critical or stressful periods are detected, or to communicate to others when a surgeon is engaged in a complex manoeuvre and should not be disturbed. Passive brain-computer interfaces infer changes in cognitive and affective state by monitoring and interpreting ongoing brain activity recorded via an electroencephalogram. The resulting information can then be used to automatically adapt a technological system to the human user. So far, passive brain-computer interfaces have mostly been investigated in laboratory settings, even though they are intended to be applied in real world settings. In this study, a passive brain-computer interface was used to assess changes in task load of skilled surgeons performing both simple and complex surgical training tasks. Results indicate that the introduced methodology can reliably and continuously detect changes in task load in this realistic environment.

1 Introduction

The last thirty years have witnessed a radical transformation in the operative environment with the introduction of Minimally Invasive Surgery (MIS), propelled by patient demand, smaller operative incisions and faster recovery. However, surgeons take significantly longer to reach proficiency in MIS[1] and find it cognitively more burdensome compared to traditional open operative procedures.[2] For example, open hand knot tying (OHKT), a routinely performed surgical manoeuvre, is far less challenging than knot tying performed in minimally invasive surgery (MISKT), and an inability to perform MISKT efficiently limits the surgeon from performing advanced MIS procedures.[3] This is attributed to poorer instrument ergonomics such as the loss of depth perception, reduced degrees of freedom of movement, amplification of tremors from using long instruments, a lack of tactile feedback, and paradoxical movements as a result of the fulcrum effect.[2] Additionally, the influx of new technology that supports MIS necessitates the operators' vigilance to attend to auditory alarms that alerts the surgeon to a faulty technical device or declining status of the monitored patient. If an alarm suggesting failure of a device is undetected, it may cause potential harm to the patient.

Surgeons are unique amongst doctors because they are often required to make decisions based on the presentation of the real-time problem whilst operating, and therefore should not just possess the technical capability but also the cognitive resources to deal with unanticipated scenarios.[4] Moreover whilst operating, surgeons on average are interrupted 13.5 times[5] and at times these interruptions warrant immediate decisions on management of a critically unwell patient outside the theatre. The impact of such secondary tasks (decision-making, detection of sensory stimuli) has been observed to degrade performance of the primary task (technical performance) albeit to a lesser extent in experts who have presumably achieved automaticity.[6–9]

According to resource theory, humans have a finite pool of resources that can be allocated or shared across tasks[10] with the assumption that the more challenging a task, the greater the resources it consumes. Therefore, a cognitively chal-

This is the author's final draft. The final publication is in Zander, T. O. et al. (2017), *Journal of Medical Robotics Research, 2*(1), 1750003. DOI: 10.1142/S2424905X17500039

lenging surgical task can not only impair situational awareness but also degrade optimal performance, which ultimately could jeopardize patient safety. To enhance surgical ergonomics with the view of improving patient safety it is imperative that we be able to characterize these variations in cognitive demand, brought about by different tasks or different contingencies during the same task.

Compared to traditional cognitive state measures such as behavioural correlates[11] or subjective questionnaires such as the NASA Task Load Index (NASA-TLX),[12] the analysis of neural mechanisms[13] has a number of advantages. The measurement can be done continuously and in real time, i.e. in the very moment the relevant events take place, and the measurement does not interfere with the actual task. Also influences of memory, e.g. primacy and recency effects, do not play a role in such measurements. Most importantly however, measures of brain activity may provide more detailed and fine-grained insights into the state of the surgeon than traditional measures. Neural correlates may be used to accurately identify current tasks, but potentially also carry information on e.g. task load, attention, and error detection.

Studying brain behaviour in surgeons enables the impact of novel technologies on operators' cognitions to be assessed. Non-invasive neuroimaging technologies such as electroencephalography (EEG)[14] and functional optical brain imaging have previously been applied to assess technical expertise,[15] skills acquisition,[15,16] cognitive burden[17] and fatigue.[18] These technologies can be applied to assess brain dynamics underlying cognitive processes even in actively moving participants.[19,20]

Robotic MIS platforms can conceivably acquire and learn, in situ, operator-specific motor and cognitive behaviour through human-robot interaction. This novel concept termed 'perceptual docking'[21] can be realised from emerging multimodal sensing and feedback rather than one aspect of surgeon-robot interaction. Surgical robotic platforms with tremor filtration capabilities, image magnification and improved actuation offer a high degree of surgical precision and may offload a degree of cognitive burden placed on the operator during MIS. Critically, analyses of operator brain function when synergized with online learning algorithms may enable the robot to benefit from human surgical intelligence and learn to better assist the surgeon, preventing errors and enhancing patient safety. Brain activity may provide an additional seamless communication channel between the surgeon and robot, and thus improve the robot's understanding of the cognitive challenges faced by the operator. This secondary, implicit interaction loop would provide valuable information to the robot, without demanding extra cognitive load from the operator.[22] Technology that adapts to the users in this way based on their brain activity can be termed *neuroadaptive technology*, and the required form of communication between the brain and the machine can be implemented using *passive* brain-computer interface (BCI).[23]

In order to evaluate the potential value of passive BCI to robotic surgery we employed an EEG-based BCI system to detect tasks and assess task load over surrogate measures that have been employed in other studies.

Traditionally, BCI is defined as a non-muscular communication and control channel for people suffering from diseases that disrupt the neural pathways through which the brain communicates with and controls its external environment.[24] Brain activity associated with the user's intent is measured and translated in real time into control signals for communication systems or other external devices. A recent development within the field of BCI broadens this general BCI approach by substituting the user's volitionally generated command (e.g. intentionally imagined hand movements) with passively conveyed implicit information.[23,25] Such a passive BCI derives its input from brain activity arising without the purpose of voluntary control, e.g. spontaneous activity indicative of task-induced cognitive or affective states.[23,26] This activity reflects covert aspects of the user's state. Therefore, it carries task-relevant information, and can thus be used to goal-directedly enrich the human-machine interaction without requiring any overt communication from the user.[23] This extends the potential field of application of BCI technology to include users without disabilities[25]—users, for example, in critical working environments where mental state measurements can provide tangible benefits. Indeed, passive BCIs have proven to be a valuable tool for detecting cognitive load[27] and working memory load.[28] These effects have been demonstrated under laboratory conditions (e.g. Gevins et al., 1998) as well as for more complex and natural tasks such as simulated driving (Brookhuis and de Waard, 1993) or flight (Sterman and Mann, 1995; Sterman et al., 1994) (as cited elsewhere[29]).

The vast majority of these studies evaluate spectral power differences in certain frequency bands that have shown to be sensitive to different cognitive load conditions. In particular, an increase of frontal theta activity is observed while parietal alpha power decreases as more cognitive resources are allocated to the task.[30–33] Hence, we hypothesize that passive BCI can be used to reliably differentiate between OHKT and MISKT, and that during performance of technically challenging tasks (MISKT, dual-tasking) an increased frontal theta activity with concurrent decreased parietal alpha activity will be observed in the EEG.

Figure 1: A participant wearing the mobile EEG cap. Data recorded at each electrode is transferred wirelessly to the amplifier connected to a standard PC. The participant is performing the MISKT task in a laparoscopic box trainer.

2 Methods

2.1 Participants

Nine participants were recruited from Imperial College London after seeking consent and screening for neuropsychiatric illnesses. Participants comprised of eight residents (PGY3-8) and one attendee aged 32.2 ± 3.3 years (mean ± S.D.). All participants underwent 15 minutes of mandatory practice for warm up of the relevant surgical tasks.

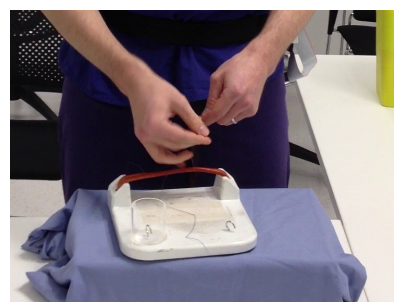

Figure 2: An example of the OHKT task on the bench knot-tying trainer.

2.2 Experimental Tasks and Conditions

The primary task was to perform one set of OHKT and MISKT alone for a fixed period of 200 seconds. The OHKT task involved the formation of as many reef knots as possible within 200 seconds using suture material (2/0 Polysorb) on a bench knot-tying trainer (Ethicon Ltd, Sommerville, New Jersey, USA) as illustrated in Figure 2. MISKT was performed in a laparoscopic box trainer (iSurgicals, UK) using 2/0 Polysorb suture material with laparoscopic needle holders (model 26173KC; Karl Storz GmbH and Co, Tuttlingen, Germany) as illustrated in Figure 1.

In order to simulate a more realistic setting, a secondary auditory task was added to the experiment. In addition to performing OHKT/MISKT, participants were required to simultaneously count the number of high-toned beeps randomly introduced amongst a series of low tone beeps (secondary task). The auditory stimuli were generated at a standardized decibel level from a speaker placed at a set distance away, on the right side of the participant (simulating a stationary auxiliary device). A tone was presented every second. 20% of tones were high-toned target stimuli.

There were five experimental conditions: Secondary (auditory) task only, primary (OHKT, MISKT) task only, and dual-task OHKT and MISKT plus auditory.

Following the warm-up phase, the secondary task was performed once for a baseline measurement. The other four conditions were performed in three sets (i.e. 12 sets in total), the order of which was randomized to minimize learning and other temporal effects such as fatigue.

2.3 Subjective Measure: NAXA-TLX

Subjective cognitive load of the baseline, mono, and dual tasks were all assessed by the NASA-TLX questionnaire.[12] After each set, participants provided ratings on six subscales (mental demands, physical demands, temporal demands, own performance, effort, and frustration), which were then combined into a single task load index, ranging from 0 to 100. Higher NASA-TLX scores indicate higher mental workload experienced by the participant.

2.4 Objective Measure: EEG Set-Up and Processing

32-channel EEG was recorded from all participants with a BrainAmp system (Brain Products GmbH, Gilching, Germany). Data was transferred wirelessly between cap and amplifier by a BrainAmp MOVE system (Brain Products GmbH, Gilching, Germany), in order not to restrict the free movement of the participants (Figure 1). Due to technical failure, EEG data from the first three participants was discarded, leaving six participants. This is the minimum required number for statistical significance in our tests

A passive BCI was set up using the BCILAB toolbox[34] with the intention to discriminate continuously, based on the EEG, whether a participant is working in OHKT or MISKT mode. In brief, each second of the continuous EEG data was transformed into a set of features of lower dimensionality. The transformation was optimised to discriminate between the two primary conditions (OHKT/MISKT), allowing continuous task detection. We used data from the dual-task sessions, as these represent the most realistic condition. Specifically, features were extracted by the spectrally-weighted Common Spatial Patterns (SpecCSP) method[35] and classification was done with a Linear Discriminant Analysis (LDA)[36] regularized by shrinkage.[37] SpecCSP generated 16 pairs of spatio-temporal filters for each participant based on all data recorded for this participant in the dual-task session. The spatial and temporal parameters of these filters were optimized iteratively to optimally discriminate between the two classes (OHKT/MISKT) based on brain activity including the theta and alpha bands (5–18 Hz). 32-dimensional features were generated by projecting one second of data recorded after the onset of each tone with each of the filters generated by SpecCSP. This resulted in 512 normally distributed features per class. LDA is the optimal classifier for this decision problem, as it provides an optimized decision plane and suffices a very low Vapnik-Chervonenkis (VC) dimension.[38]

An estimate of the online (real-time) accuracy of the resulting classifier was derived by calibrating the classifier on one part of the data, and applying it to the remaining part. Thus, for each second of the data that was not used for calibration, the classifier indicated whether or not the surgeon was at that time performing OHKT or MISKT. This was done a number of times, calibrating on and applying to different parts of the available data. The resulting accuracy indicates the percentage of correct indications. Specifically, we used a (5,5)-times nested cross-validation[36] with margins of 5. These margins were selected to guarantee the IIDness of the features. The outcomes of the 5-fold outer runs, regularized by the one shrinkage parameter derived in the appropriate inner runs, gave the estimates for the reliability of each model. The overall reliability (Estimated Classification Accuracy, ECA) was then given by the mean of these single runs' reliability. The validity of this estimate is supported by the low probability of overfitting of classifiers with low VC dimension, by the fact that the suboptimal ratio between feature dimensionality and number of trials can be counterbalanced by a well-chosen shrinkage regularization, and lastly by the fact that a nested cross-validation was applied properly.

For an additional accuracy measure, testing how well the classifier performed on data recorded in a different context, each participant's classifier, trained on dual-task EEG data, was then tested on the EEG data recorded for this respective participant during the mono-task session.

Features were generated as above from each (non-overlapping) second of this session, and classified as being either OHKT or MISKT task periods. This was then compared to the real task labels, resulting in a pseudo-online classification accuracy (POCA), simulating an online application of the classifier on a separate data set.

In essence, the SpecCSP patterns generated by the classifier can be interpreted as filters, isolating specific, maximally discriminative processes.[35] They thus identify different cognitive processes, which can be investigated further. To investigate these electrocortical processes underlying classification, EEGLAB[39] was used to cluster patterns resulting from each SpecCSP filter by multiplying with the inverse covariance matrix of the underlying data. These patterns represent the scalp projections of the underlying generator sources. Clustering is a method to identify sources that consistently aided classification across participants. To this end, the patterns' spatial distributions were first reduced to 25 dimensions by means of PCA and subsequently clustered using k-means specifying 28 clusters.[36] Further analysis focused on both the clusters and the individual SpecCSP patterns. The time course of the event-related spectral perturbation (ERSP,[40]) and the frequency power spectrum were calculated for 0.5 seconds before the onset of the auditory stimulus until 1.5 seconds after. The timing was chosen to demonstrate repeated stimulus related effects. Permutation tests, corrected for false discovery rate (FDR), were used to test for significant differences between the ERSPs, as implemented in EEGLAB.[39]

Task	NASA-TLX score
Mono OHKT	16.99 ± 2.85
Mono MISKT	48.79 ± 6.17
Auditory	24.33 ± 6.07
Dual OHKT	51.04 ± 7.43
Dual MISKT	67.55 ± 6.92

Table 1: Group averaged NASA-TLX scores (mean ± S.E.) for each task.

Test	P-value
Dual OHKT vs Mono OHKT	.018
Dual MISKT vs Mono MISKT	.012
Dual OHKT vs Auditory	.017
Dual MISKT vs Auditory	.011
Mono OHKT vs Mono MISKT	.012
Mono MISKT vs Auditory	.025
Mono OHKT vs Auditory	.553

Table 2: Significance values for across-subjects comparisons of NASA-TLX scores using Wilcoxon signed-rank tests.

Participant ID	ECA	POCA
1	82.4% (6.2%)	72,5%
2	77.8% (11.4%)	77.2%
3	95.5% (4.1%)	91.7%
4	90.7% (5.9%)	95.8%
6	99.7% (0.4%)	97.8%

Table 3: Estimated classification accuracy (ECA) and pseudo-online classification accuracy (POCA) for each participant.

3 Results

3.1 Questionnaire Results

The raw questionnaire results are listed in Table 1. The scores were compared using Wilcoxon signed-rank tests. The alpha level was set at 0.05. The results of these comparisons are listed in Table 2. It is clear that dual-task MISKT was found to be the most challenging task, followed by dual-task OHKT and mono-task MISKT. The two least challenging tasks were the secondary task in itself, and mono-task OHKT. These two did not differ significantly from each other.

3.2 BCI Results

3.2.1 Classification Accuracy

Table 3 provides the estimated classification accuracies from the calibration session and the pseudo-online classification accuracies from the test session for each participant. High accuracies in both measures attest to a high reliability of the passive

This is the author's final draft. The final publication is in Zander, T. O. et al. (2017), *Journal of Medical Robotics Research, 2*(1), 1750003. DOI: 10.1142/S2424905X17500039

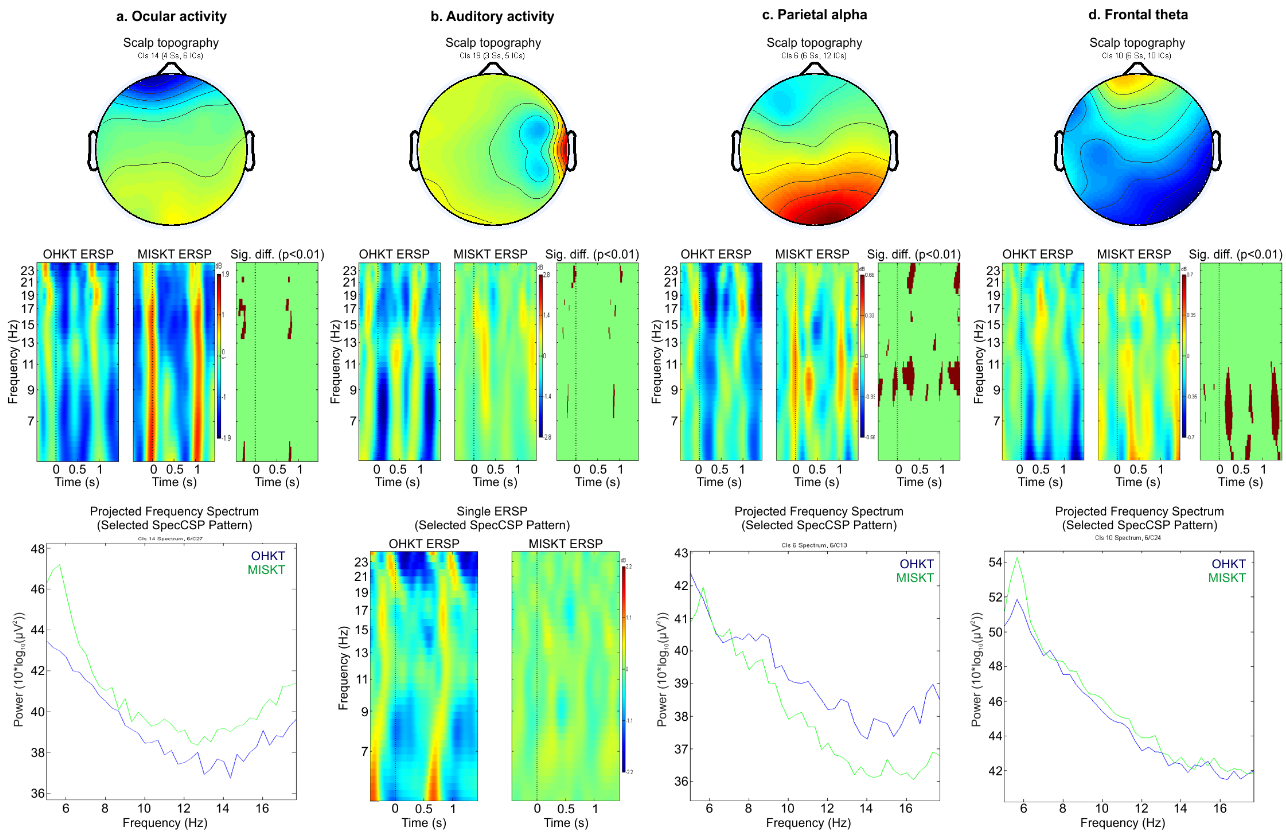

Figure 3: Analysis of a number of patterns consistently identified by the BCI across participants show different activities that contributed to classification. *Top row:* Representation of clustered SpecCSP pattern means showing the topography of the identified influences. *Middle row:* Cluster mean ERSP time-locked to tone onset in both conditions, illustrating the spectral activity. *Bottom row (a,c,d):* Frequency spectrum projected from a single pattern illustrating non-time-locked differences between conditions. *Bottom row (b):* Selected component ERSP, illustrating individual differences within this cluster.

BCI system applied here. As there is no significant difference (tested with a Student's t-test) between the ECA and the POCA measures, it is unlikely that the BCI definition overfitted on unknown factors during calibration. Together, these results provide a promising step towards robust, real-time task detection in the theatre.

3.2.2 Contributing EEG Activity

Figure 3 highlights a number of findings from the neurophysiological analysis of the EEG activity that contributed to classification. The SpecCSP patterns that were generated and used by the classifier, can be interpreted as isolating different discriminative processes.[35] Across participants, the classifier identified a number of consistent patterns reflecting task-relevant processes. Based on the topography of their scalp projections, these processes can be localised within the three-dimensional space covered by the electrodes. We highlight four clusters that illustrate the variety of cortical and non-cortical information that passive BCI systems can potentially make use of to enable real-time user state measurement.

The top row of each panel shows the mean spatial distribution of the cluster patterns. Below that, the cluster's mean 2-second ERSP up to 25 Hz, time-locked to the auditory stimuli, is shown for both the OHKT condition (blue) and the MISKT, higher task load condition (green), as well as a plot highlighting significant differences between the two computed by permutation statistics. Exemplary single-participant data (frequency spectra or ERSP) is illustrated on the bottom row.

Panel *a* of Figure 3 shows a cluster associated with ocular activity, as identifiable by the spatial distribution of the weights. The stimulus-locked increase in broadband spectral power indicates that this cluster entails slow eye movements (horizontal movements associated with the task) as well as high frequency muscle activity related to muscular components of the eye movement in the MISKT condition. The increase in the low frequency range of the individual spectral plot clearly shows an increase of blink activity during higher task load.

A cluster most likely representing the processing of auditory information, with high weights near the right auditory cortex, is presented in panel *b* of Figure 3. This cluster's mean ERSP reveals stronger desynchronization in a wide frequency range up to 25 Hz for the lower task load condition (OHKT) time locked to the auditory stimulus onset. This is reflective of increased processing of auditory information (i.e. more resources available for auditory processing) during lower task load levels. However, these differences are only marginally significant and are present only for very restricted periods of time. The bottom part of this panel shows the ERSP projected by a single pattern for one participant, revealing strong differences for this individual.

Increased blink activity and modulated auditory processing can be well explained given the current conditions, but may not be specifically related to task load in general. For more generic, context-independent effects, we had hypothesised to find changes in parietal alpha and frontal theta activity. These effects of task load are further investigated in panels *c* and *d* of Figure 3.

Panel *c* represents a cluster of SpecCSP topographies with high positive weights over occipito-parietal areas. The mean ERSPs demonstrate, contrary to the hypothesis, increased alpha desynchronization in the low as compared to the high load condition. However, individual spectra of certain participants, as e.g. illustrated in the bottom row, do demonstrate the expected, opposite effect in the alpha band. Here one must note that the topographic plot shows two separate weight centres over frontal and posterior areas, indicating that this component may reflect two, probably related processes. The observed effects are time-locked to the onset of the auditory stimulus, and may thus partially represent a cognitive process related to the secondary task, rather than the expected more generic load effect.

Panel *d*, finally, shows three plots generated from a cluster representing theta activity, most likely the frontal theta component described in.[31] The mean topographic distribution of the patterns' SpecCSP weights shows a fronto-posterior polarity inversion with positive fronto-central weight and negative weights over the posterior cortex, with a slight right-lateralization. In line with our hypothesis, the ERSPs show a clear increase in the theta band during the MISKT condition. The frequency spectrum projected from a single SpecCSP pattern demonstrates how the energy in the theta band increases with higher task load. There also are supportive effects in the alpha band, but these fail to reach significance here.

4 Discussion and Conclusion

This study simulated a realistic operative environment using both a traditional and an MIS version of an established surgical task. These tasks were combined with an additional auditory task to increase load. A noninvasive EEG system that allowed free movement was used to record the participants' brain activity throughout the experiment. Based on this data, a passive BCI was calibrated in order to automatically discriminate between periods of OHKT (lower load) and MISKT (higher load). We had hypothesised that a) this discrimination was reliably possible, and b) that it would take parietal alpha and frontal theta effects into account.

In similar simulations of high-risk scenarios, such as military operations or air traffic control, using fNIRS optical brain imaging, an increase in task complexity was linked to an increase in prefrontal cortex activity.[41,42] The prefrontal cortex is a region within the frontal lobe associated with e.g. executive control and error detection. When an auditory secondary task is introduced, cognitive load increases, leading to a reduction in alertness, as e.g. observed by prolonged brake reaction times in a driving task.[43] Similarly, the impact of auditory distractors on the surgeon whilst operating, as used here, may increase load especially during critical, technically challenging moments or near the end of the procedure as fatigue sets in.

The passive BCI used here was calibrated on data from these dual-task scenarios. It then was capable of discriminating between EEG data recorded during mono-task OHKT and MISKT sessions with a high reliability. The cross-validated accuracy estimations for each participant are not significantly different than the accuracies achieved on the test set, strongly supporting the BCI's ability to automatically discriminate between conditions even in real time. In practice, EEG does not only contain brain activity, but also first order artefacts related to bodily activity (e.g. muscle tension, eye movements) and second order artefacts (e.g. external electromagnetic fields, unstable electrodes, mechanical forces working on the cables or electrodes[44, 45]). Because of this, a range of different types of activity is available for the BCI to aid the discrimination between conditions, both cortical and non-cortical. Further analysis of the features underlying the BCI's operation provides information with respect to the selected dynamics that contributed to classification. Indeed, a number of different activity sources are reflected in the features.

The processes underlying the demonstrated successful discrimination include cortical activity in line with previous findings related to alpha and theta activity, as hypothesised, although some individual differences remain. We also found other processes that may not be uniquely related to generic task load, but may still be relevant to track in real time, as BCI technology allows.The analysis presented here does not definitively exclude the possibility that other processes played a role in classification, but the BCI's level of performance across training and testing conditions attests that the classifier was unlikely to be overfitted on circumstantial activity. It is clear that BCI enables a number of different task-relevant cortical and non-cortical measurements to take place simultaneously.

A focus on cortical as opposed to artefactual activity is advantageous as it is likely to be more robust across different situations and context. Figure 3 shows that relevant cortical information was indeed available, and taken into account for classification. It is thus possible to monitor and detect task-relevant cortical activity from the EEG of a surgeon actively involved in a surgical procedure.

An ability to detect task load in real time enables a number of potentially important improvements to be implemented to aid the surgeon and the operative team. Regarding robotic surgery, this pBCI framework has implications that may help maximize patient safety. For example, at times of escalating load burden, the system may initiate active dampening which may change the ratio of instrument motion scaling, or engage dynamic active constraints to prevent instruments from entering critical 'no go' anatomical zones to restrict the chance of injury to vital organs. Mundane tasks may be relegated to the robot as a means of adaptive automation to enable the surgeon to focus solely on critical tasks requiring higher level decision-making. Such an intelligent system may also be used for surgical training and assessment of technical and cognitive skills, and opens avenues for neural feedback training to improve performance.

5 Outlook

This study represents a first proof of principle that passive BCIs can identify task-relevant cognitive processes in established surgery scenarios, and differentiate between different tasks based on continuous EEG data. The estimated and pseudo-online classification accuracies support the idea that passive BCIs may be used for real-time task load detection in the theatre during robotic surgery scenarios. This opens up a variety of applications for BCI technology, and gives a perspective on new types of assistive technology in these and other environments. Aside from task load, other cognitive and affective processes can potentially be assessed using the same BCI methodology, providing richer, more detailed information about the current state of the operator. This information can then be used to automatically improve the interaction between the operator, machine, and operative assistants. Transferring knowledge from neuroscientific research into the field of passive BCI should enable many different cognitive and affective aspects to be detected automatically. Robotic surgery in particular could benefit from information about the current level of attention of a surgeon, or by detecting specific intentions (e.g. the need for a specific tool, which can then be prepared).

Future steps continuing the work presented here could include an investigation of the task specificity of the presented approach. Can it be transferred easily to tasks that are significantly different from OHKT and MISKT? Practical aspects of the set-up may also be investigated: Can the amount of training data be reduced without losing accuracy, and what is the minimum amount of electrodes needed for reliable results?

A final, highly interesting step would be to investigate the application of other developments from the field of human-computer interaction and passive BCI to robotic surgery, in particular, passive BCI-based implicit control.[46] Implicit control could open up an additional, goal-oriented communication channel for the surgeon without placing any additional load on them. The fusion of two worlds, that of passive BCI research and robotic surgery, shows the potential of leading us into a fascinating new world of brain-based interaction.

Acknowledgments

The authors would like to thank Brain Products GmbH for supporting us with a BrainAmp DC and a BrainAmp Move system, such that we could record the EEG for this study.

References

[1] P. J. Guillou, P. Quirke, H. Thorpe, J. Walker, D. G. Jayne, A. M. H. Smith, R. M. Heath and J. M. Brown, Short-term endpoints of conventional versus laparoscopic-assisted surgery in patients with colorectal cancer (mrc 5clasicc6 trial): multicentre, randomised controlled trial, *The Lancet* **365**(9472) (2005) 1718–1726.

[2] R. Berguer, W. D. Smith and Y. H. Chung, Performing laparoscopic surgery is significantly more stressful for the surgeon than open surgery, *Surgical Endoscopy* **15**(10) (2001) 1204–1207.

[3] J. R. Korndorffer, J. B. Dunne, R. Sierra, D. Stefanidis, C. L. Touchard and D. J. Scott, Simulator training for laparoscopic suturing using performance goals translates to the operating room, *Journal of the American College of Surgeons* **201**(1) (2005) 23–29.

[4] A. G. Gallagher, R. M. Satava and G. C. Osullivan, Attentional capacity: An essential aspect of surgeon performance, *Annals of Surgery* (2013).

[5] A. N. Healey, N. Sevdalis and C. A. Vincent, Measuring intra-operative interference from distraction and interruption observed in the operating theatre, *Ergonomics* **49**(5-6) (2006) 589–604.

[6] D. Stefanidis, M. W. Scerbo, J. R. Korndorffer and D. J. Scott, Redefining simulator proficiency using automaticity theory, *American Journal of Surgery* **193**(4) (2007) 502–506.

[7] K. E. Hsu, F.-Y. Man, R. A. Gizicki, L. S. Feldman and G. M. Fried, Experienced surgeons can do more than one thing at a time: effect of distraction on performance of a simple laparoscopic and cognitive task by experienced and novice surgeons, *Surgical Endoscopy* **22**(1) (2008) 196–201.

[8] B. Zheng, M. A. Cassera, D. V. Martinec, G. O. Spaun and L. L. Swanström, Measuring mental workload during the performance of advanced laparoscopic tasks, *Surgical Endoscopy* **24**(1) (2010) 45–50.

[9] B. Zheng, G. Tien, S. M. Atkins, C. Swindells, H. Tanin, A. Meneghetti, K. A. Qayumi, O. Neely and M. Panton, Surgeon's vigilance in the operating room, *American Journal of Surgery* **201**(5) (2011) 673–677.

[10] M. A. Staal, Stress, cognition, and human performance: A literature review and conceptual framework, *NASA Technical Memorandum* **212824** (2004).

[11] M. Fafrowicz and T. Marek, Quo vadis, neuroergonomics?, *Ergonomics* **50**(11) (2007) 1941–1949.

[12] S. G. Hart and L. E. Staveland, Development of nasa-tlx (task load index): Results of empirical and theoretical research, *Human Mental Workload*, eds. P. A. Hancock and N. Meshkati (North-Holland, Amsterdam, 1988), pp. 139–183.

[13] M. A. Just, P. A. Carpenter and A. Miyake, Neuroindices of cognitive workload: Neuroimaging, pupillometric and event-related potential studies of brain work, *Theoretical Issues in Ergonomics Science* **4**(1-2) (2003) 56–88.

[14] F. F. Zhu, J. M. Poolton, M. R. Wilson, Y. Hu, J. P. Maxwell and R. S. W. Masters, Implicit motor learning promotes neural efficiency during laparoscopy, *Surgical Endoscopy* **25**(9) (2011) 2950–2955.

[15] D. R. Leff, C. E. Elwell, F. Orihuela-Espina, L. Atallah, D. T. Delpy, A. W. Darzi and G.-Z. Yang, Changes in prefrontal cortical behaviour depend upon familiarity on a bimanual co-ordination task: An fnirs study, *NeuroImage* **39**(2) (2008) 805–813.

[16] K. Ohuchida, H. Kenmotsu, A. Yamamoto, K. Sawada, T. Hayami, K. Morooka, S. Takasugi, K. Konishi, S. Ieiri, K. Tanoue, Y. Iwamoto, M. Tanaka and M. Hashizume, The frontal cortex is activated during learning of endoscopic procedures, *Surgical Endoscopy* **23**(10) (2009) 2296–2301.

[17] D. R. C. James, F. Orihuela-Espina, D. R. Leff, G. P. Mylonas, K.-W. Kwok, A. W. Darzi and G.-Z. Yang, Cognitive burden estimation for visuomotor learning with fnirs, *MICCAI 2010*, *Lecture Notes in Computer Science* **6363** (Springer Berlin Heidelberg, Berlin and Heidelberg, 2010), pp. 319–326.

[18] D. R. Leff, F. Orihuela-Espina, T. Athanasiou, V. Karimyan, C. Elwell, J. Wong, G.-Z. Yang and A. W. Darzi, Circadian cortical compensation: a longitudinal study of brain function during technical and cognitive skills in acutely sleep-deprived surgical residents, *Annals of Surgery* **252**(6) (2010) 1082–1090.

[19] S. Makeig, K. Gramann, T.-P. Jung, T. J. Sejnowski and H. Poizner, Linking brain, mind and behavior, *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology* **73**(2) (2009) 95–100.

[20] K. Gramann, D. P. Ferris, J. Gwin and S. Makeig, Imaging natural cognition in action, *International Journal of Psychophysiology* (2013).

[21] G.-Z. Yang, G. P. Mylonas, K.-W. Kwok and A. Chung, Perceptual docking for robotic control, *Medical Imaging and Augmented Reality*, eds. T. Dohi, I. Sakuma and H. Liao, **5128** (Springer, Berlin, 2008), pp. 21–30.

[22] M. Rötting, T. Zander, S. Trösterer and J. Dzaack, Implicit interaction in multimodal human-machine systems, *Industrial Engineering and Ergonomics*, ed. C. M. Schlick (Springer, Berlin, 2009), pp. 523–536.

[23] T. O. Zander and C. Kothe, Towards passive brain–computer interfaces: applying brain–computer interface technology to human–machine systems in general, *Journal of Neural Engineering* **8**(2) (2011) p. 025005.

[24] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller and T. M. Vaughan, Brain–computer interfaces for communication and control, *Clinical Neurophysiology* **113**(6) (2002) 767–791.

[25] T. O. Zander, C. Kothe, S. Jatzev and M. Gaertner, Enhancing human-computer interaction with input from active and passive brain-computer interfaces, *Brain-Computer Interfaces: Applying our Minds to Human-Computer Interaction*, eds. D. S. Tan and A. Nijholt, *Human-Computer Interaction Series* (Springer, London, 2010), pp. 181–199.

[26] T. O. Zander and S. Jatzev, Context-aware brain–computer interfaces: exploring the information space of user, technical system and environment, *Journal of neural engineering* **9**(1) (2012) p. 016003.

[27] J. Kohlmorgen, G. Dornhege, M. Braun, B. Blankertz, K. R. Muller, G. Curio, K. Hagemann, A. Bruns, M. Schrauf and W. Kincses, Improving human performance in a real operating environment through real-time mental workload detection, *Toward brain-computer interfacing*, eds. G. Dornhege, J. R. Millán, T. Hinterberger, D. J. McFarland and K.-R. Müller (MIT Press, Cambridge, 2007), pp. 409–422.

[28] D. Grimes, D. S. Tan, S. E. Hudson, S. Pradeep and R. P. N. Rao, Feasibility and pragmatics of classifying working memory load with an electroencephalograph, *Pro*

SIGCHI conf on Human factors in computing systems, ed. M. Czerwinski (ACM, New York and NY, 2008), pp. 835–844.

[29] C. D. Wickens and J. S. McCarley, *Applied attention theory* (CRC Press, Boca Raton, 2008).

[30] A. Gevins, Neurophysiological measures of working memory and individual differences in cognitive ability and cognitive style, *Cerebral Cortex* **10**(9) (2000) 829–839.

[31] A. Gevins and M. E. Smith, Neurophysiological measures of cognitive workload during human-computer interaction, *Theoretical Issues in Ergonomics Science* **4**(1-2) (2003) 113–131.

[32] A. Holm, K. Lukander, J. Korpela, M. Sallinen and K. M. I. Müller, Estimating brain load from the eeg, *TheScientificWorldJOURNAL* **9** (2009) 639–651.

[33] M. E. Smith, A. Gevins, H. Brown, A. Karnik and R. Du, Monitoring task loading with multivariate eeg measures during complex forms of human-computer interaction, *Human Factors: The Journal of the Human Factors and Ergonomics Society* **43**(3) (2001) 366–380.

[34] C. A. Kothe and S. Makeig, Bcilab: a platform for brain-computer interface development, *Journal of neural engineering* **10**(5) (2013) p. 056014.

[35] R. Tomioka, G. Dornhege, G. Nolte, B. Blankertz, K. Aihara and K.-R. Müller, Spectrally weighted common spatial pattern algorithm for single trial eeg classification, *Mathematical Engineering Technical Reports, University of Tokyo, Tokyo, Japan* **40** (2006).

[36] R. O. Duda, P. E. Hart and D. G. Stork, *Pattern classification*, 2nd ed edn. (Wiley, New York, 2001).

[37] B. Blankertz, S. Lemm, M. Treder, S. Haufe and K.-R. Müller, Single-trial analysis and classification of erp components — a tutorial: Multivariate decoding and brain reading, *NeuroImage* **56**(2) (2011) 814–825.

[38] V. N. Vapnik and A. Y. Chervonenkis, On the uniform convergence of relative frequencies of events to their probabilities, *Theory of Probability & Its Applications* **16**(2) (1971) 264–280.

[39] A. Delorme and S. Makeig, Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis, *Journal of ods* **134**(1) (2004) 9–21.

[40] S. Makeig, Auditory event-related dynamics of the eeg spectrum and effects of exposure to tones, *Electroencephalography and Clinical Neurophysiology* **86**(4) (1993) 283–293.

[41] H. Ayaz, P. A. Shewokis, S. Bunce, K. Izzetoglu, B. Willems and B. Onaral, Optical brain monitoring for operator training and mental workload assessment, *NeuroImage* **59**(1) (2012) 36–47.

[42] K. Izzetoglu, S. Bunce, B. Onaral, K. Pourrezaei and B. Chance, Functional optical brain imaging using near-infrared during cognitive tasks, *International Journal of Human-Computer Interaction* **17**(2) (2004) 211–227.

[43] A. Sonnleitner, M. S. Treder, M. Simon, S. Willmann, A. Ewald, A. Buchner and M. Schrauf, {EEG} alpha spindles and prolonged brake reaction times during auditory distraction in an on-road driving study, *Accident Analysis and Prevention* **62**(0) (2014) 110 – 118.

[44] J. T. Gwin, K. Gramann, S. Makeig and D. P. Ferris, Electrocortical activity is coupled to gait cycle phase during treadmill walking, *NeuroImage* **54**(2) (2011) 1289 – 1296.

[45] K. Gramann, J. T. Gwin, N. Bigdely-Shamlo, D. P. Ferris and S. Makeig, Visual evoked responses during standing and walking, *Frontiers in Human Neuroscience* **4**(202) (2010).

[46] T. O. Zander, J. Broenstrup, R. Lorenz and L. R. Krol, Towards bci-based implicit control in human-computer interaction, *Advances in Physiological Computing*, eds. S. Fairclough and K. Gilleade, *Human-Computer Interaction Series* (Springer, London, in press).

This is the author's final draft. The final publication is in Zander, T. O. et al. (2017), *Journal of Medical Robotics Research, 2*(1), 1750003. DOI: 10.1142/S2424905X17500039

Dr Thorsten O. Zander leads the research group "Team PhyPA" at the chair for Biological Psychology and Neuroergonomics at the Technische Universität Berlin, Germany. His research focuses on the development of Neuradaptive Technology by integrating passive Brain-Computer Interfaces (BCI) into Human-Machine Systems. He defined the field of passive BCI in 2008 and founded the "Community for Passive BCI research" in 2013. In 2010, he co-defined the notion of Hybrid Brain-Computer Interfaces, integrating BCI technology into multimodal Human-Machine Systems. He is credited for investigating the combination of eye tracking and active/passive BCI in Human-Computer Interaction, the use of dry sensors for electroencephalography in real-world settings as well as for investigating the use of human error responses for adaptive automation. Next to application oriented research, Dr Zander is interested in investigating brain activity in real-world environments utilizing the mobile brain/body imaging approach, for example in flight simulators or during driving. Next to Technische Universität Berlin he is affiliated to the Leibniz Institute Knowledge Research Center in Tuebingen, to the Excellence Cluster Graduate School for Learning, Educational Achievement, and Life Course Development, the Swartz Center for Computational Neuroscience and the Cognitive Science Department at the University of California San Diego, USA.

Professor Guang-Zhong Yang (FREng, FIEEE, FIET, FAIMBE, FIAMBE, FMICCAI, FCGI) is director and co-founder of the Hamlyn Centre for Robotic Surgery, Deputy Chairman of the Institute of Global Health Innovation, Imperial College London, UK. Professor Yangs main research interests are in medical imaging, sensing and robotics. In imaging, he is credited for a number of novel MR phase contrast velocity imaging and computational modelling techniques that have transformed in vivo blood flow quantification and visualization. These include the development of locally focused imaging combined with real-time navigator echoes for resolving respiratory motion for high-resolution coronary-angiography, as well as MR dynamic flow pressure mapping for which he received the ISMRM I. I Rabi Award. He pioneered the concept of perceptual docking for robotic control, which represents a paradigm shift of learning and knowledge acquisition of motor and perceptual/cognitive behaviour for robotics, as well as the field of Body Sensor Network (BSN) for providing personalized wireless monitoring platforms that are pervasive, intelligent, and context-aware. He is a Fellow of the Royal Academy of Engineering, fellow of IEEE, IET, AIMBE, IAMBE, MICCAI, City & Guilds and a recipient of the Royal Society Research Merit Award and listed in The Times Eureka Top 100 in British Science.