Parsimonious Identification of Physiological Indices for Monitoring Cognitive Fatigue

Lance J. Myers and J. Hunter Downs

Archinoetics, LLC. 700 Bishop St, Suite E, Honolulu, HI, 96813

Abstract. The objective of this study was to identify a parsimonious set of physiological measures that could be used to best predict cognitive fatigue levels. A 37 hour sleep deprivation study was conducted to induce reduced levels of alertness and cognitive impairment as measured by a psychomotor vigilance test. Non-invasive, wearable and ambulatory sensors were used to acquire cardio-respiratory and motion data during the sleep deprivation. Subsequently 23 potential predictors were derived from the raw sensor data. The least absolute shrinkage and selection operator, along with a cross validation strategy was used to create a sparse model and identify a minimum predictor subset that provided the best prediction accuracy. Final predictor selection was found to vary with task and context. Depending on context selected predictors indicated elevated levels of sympathetic nervous system activity, increased restlessness during engaging tasks and increased cardio-respiratory synchronization with increasing cognitive fatigue.

Keywords: cognitive fatigue, heart rate variability, feature selection, wearable sensors.

1 Introduction

Fatigue is a growing problem in modern society. Although sleep experts have found that most adults need 8 hours of sleep per night [1], the average American adult is sleeping only 6.8 hours per night [2], and as much as 20% of the population appears to be acquiring only 6.5 hours of sleep per night [3].

In general terms, losing even small amounts of sleep each night will exert cumulative adverse effects on waking performance which include vigilance decrements, increased lapses of attention, cognitive slowing, short-term memory failures, deficits in frontal lobe functions, and rapid and involuntary sleep onsets [4]. Studies comparing the effects of increased blood alcohol concentrations (BAC) to the effects of sleep loss illustrate the seriousness of insufficient sleep on alertness and performance. Investigations have shown that sustained wakefulness of 20-24 hours produces decrements equal to those observed with BAC levels of between 0.08%-0.10% on tests of psychomotor performance, grammatical reasoning, vigilance, and simulated driving performance [5]. Operator fatigue is frequently responsible for costly accidents and mishaps in driving, aviation, shift health care workers and other similar industries.

D.D. Schmorrow et al. (Eds.): Augmented Cognition, HCII 2009, LNAI 5638, pp. 495–503, 2009.

If it were possible to accurately predict when an individual was becoming overly fatigued, timely mitigation strategies could be employed to prevent accidents and other costly fatigue related problems. Several computer models currently calculate performance-readiness predictions that are generally accurate, but none are capable of accounting for individual differences in fatigue vulnerability [6].

Clearly, sleep-wake times are insufficient to model the impact of individual differences in fatigue vulnerability. Therefore we hypothesized that a more accurate assessment of cognitive fatigue could be made by using measured physiology of an individual. To address this, we first ask the question of which physiological variables would be the best predictors of cognitive fatigue. This paper focuses on this question and seeks to identify a parsimonious subset of physiological variables that best track changes in cognitive fatigue and vigilance due to chronic sleep restriction.

2 Methods

2.1 Physiological Measurements

2.1.1 Sensor Technology

There are many candidate physiological measurements that may have strong predictive power of cognitive fatigue. Given that real-world monitoring applications require a product that may be used in non-laboratory, ambulatory contexts, we focused only on those sensor technologies that met criteria of non-invasive, ambulatory, wearable, unobtrusive and artifact resistant. We thus selected a representative commercially available ambulatory, wearable monitoring system called the BioHarness (ZephyrTech, NZ). This system is a chest strap that is capable of non-invasively, continuously and simultaneously monitoring electrocardiography (ECG), respiration, motion, posture and skin temperature. The strap is light-weight, comfortable and uses proprietary smart-fabric sensing technology. Data is continuously logged using onboard flash memory and subsequently downloaded to a PC for further offline processing. Data may also be wirelessly acquired over a BlueTooth connection for further real-time development.

2.1.2 Candidate Feature Derivation

There is a strong link between cognitive fatigue and cortical arousal as measured with electroencephalographic (EEG) activity [7]. However, due to the fact that EEG measurements are better suited for laboratory conditions, we sought to investigate whether systemic autonomic arousal demonstrated similar progression to cortical arousal. Studies have demonstrated that the fatigue state is associated with a shift of sympathovagal balance toward sympathetic predominance and reduced vagal tone [7]. Thus measurements indicative of autonomic balance could have sufficient predictive power to discriminate between differing levels of cognitive fatigue.

Heart rate variability

The analysis of heart rate variability provides a way to non-invasively study the autonomic nervous system (ANS) by acting as a dynamic window into autonomic function and balance. Over the years, a variety of metrics have been proposed to succinctly quantify HRV and the associated respiratory sinus arrhythmia (RSA). Due to the variety of metrics, in 1996 the Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology suggested standard mathematical procedures for short-term HRV evaluation [8]. However the task force report only considers HRV from the ECG and a number of laboratory studies have documented that changes in respiratory parameters can seriously confound the association of RSA and cardiac vagal tone [9]. Thus when respiratory measures are also available, several adaptations to these methods, as well as new methods may be included. In addition further indices derived from non-linear dynamical analysis are also included. Table 1 lists all the candidate predictor features derived from HRV and RSA measures.

Table 1. Candidate HRV features

Feature name	***Feature description***
Time domain features	
ANN	Average heart rate
RMSSD	Root mean square of successive differences
SDNN	Standard deviation
Frequency domain features	These were derived using the Welch method.
HF	Spectral power in high frequency range[a]
LF	Spectral power in the low frequency range[b]
VLF	Spectral power in the very low frequency range[c]
LF/HF	Ratio of LF to HF
HFnorm	HF power normalized to total power
LFnorm	LF power normalized to total power
Non-linear features	
SampEn	Sample Entropy [10]
CD	Correlation Dimension [10]
MLE	Maximum Luypanov Exponent [10]
Cardio-resp features	
RSA	Correlation coefficient[d]
PSI	Phase Synchrony Index [11][e]

[a]Default range is 0.15-0.4Hz. However, the respiratory rate was calculated from the respiration band and the HF range was centered at the respiration frequency, with a width of 0.15Hz on either side of this.

[b]Default range is 0.04-0.15Hz

[c]Default range is 0.01-0.04Hz

[d]This was the normalized correlation coefficient between the respiratory signal and the heart rate signal over the specified time epoch. The HR signal was resampled using a linear interpolation to an even rate of 2Hz. The respiration signal was downsampled to 2Hz.

[e]This is a newly developed data analysis technique based on the mathematics of nonlinear dynamics and allows any interaction that does occur in even weakly coupled complex systems to be observed [11]. It is especially well suited to probe the weak interactions between irregular and non-stationary oscillators such as the human heart and respiratory system. Phase locking of respiratory and the cardiac rhythms,

Table 2. Candidate respiratory features

Feature name	*Feature description*
General respiration	
RR	Respiration rate
Vol	Tidal volume (uncalibrated)
Mvol	Minute Ventilation
DC	Duty Cycle
Respiratory irregularity	
TVI	Tidal Volume Instability [12]
DCvar	Duty Cycle Variability[12]
BRV	RMS of sucessive differences of breath period [12]

and respiratory modulation of heart rate (RSA), are two competing aspects of cardio-respiratory interaction.

Respiratory features
Everyday observation suggests that psycho-physiologic state is a determinant of respiration and this relationship has been actively investigated dating as early as the beginning of this century. However, contemporary research showing how respiration may be used as a surrogate of autonomic balance is relatively sparse. Despite this, there is evidence that changes in autonomic balance in general do influence respiration, and that different states give rise to different breathing patterns. There are certainly obvious markers such as respiration rate and volume that are typically used to index autonomic changes. However in addition to this several studies point to dysregulated breathing as having the most potential to index arousal. For example, studies examining the relationship between respiration and state in the clinical context of anxiety, panic disorder and chronic pain all point to irregularity in breathing as a key marker of anxiety [12]. Furthermore, it has been demonstrated that irregularity in breathing appears to increase under conditions of sympathetic arousal such as emotional upset and excitement. Table 2 lists all the candidate predictor features derived from respiratory measures.

Motion features
In addition to the cardio-respiratory features, two features derived from the trunk accelerometer were also used. These features were fairly simple features and were the overall level of motion and the variability in motion. These are listed in table 3.

Table 3. Candidate motion features

Feature name	*Feature description*
motion	Mean acceleration
motionVar	Standard deviation of acceleration

2.2 Vigilance Assessment

Objective cognitive performance evaluations were accomplished with the Psychomotor Vigilance Task (PVT). This is a portable, low-voltage, battery-powered reaction-time test known to be sensitive to sleep loss [13]. The PVT is commonly utilized to track changes in vigilance or sustained attention; attributes that underlie the successful accomplishment of many types of more complex cognitive tasks. A variety of data is generated from the PVT, but of primary interest were reaction-time measures, accuracy measures, and attention-lapse indications since these reflect cognitive slowing and response failures.

2.3 Study Design

Six subjects were recruited for a continuous 37 hour sleep deprivation study. Participants were instructed to obtain a minimum of 8 hours of sleep prior to reporting the research facility. They were instructed to awaken at 0700 on day 1 and they subsequently had to remain awake until at least 1900 hours on the second day of the study. Participants were outfitted with the BioHarness and data quality was verified to ensure correct fit and function.

Starting at 1000 on Day 1, each subject had to complete the first 10-minute test session on the PVT. Subsequent PVT sessions occurred every hour until 1800 on Day 2. Before and after each PVT, participants reported to the testing room where they were seated in a comfortable chair and asked to remain relaxed, still, and quiet with eyes open for 3 full minutes to stabilize autonomic activity. In between PVT sessions, participants will be free to play computer games, watch TV, or engage in any other type of sedentary activity. Following completion of the study, data was downloaded off the on-board flash memory storage to a PC for subsequent analysis.

2.4 Parsimonious Feature Selection

The objective of this study was to identify a parsimonious set of physiological measures that could be used to best predict cognitive fatigue levels. The importance of parsimony in feature selection is emphasized as it tends to improve prediction performance and simpler models are preferred for the sake of scientific insight and interpretation of the chosen features. In general, statistical learning theory poses a structural risk minimization criterion that balances the trade-off between good empirical performance (i.e., classification accuracy on training data) and good generalization ability (i.e., classification accuracy on unseen data). Most classifiers will generalize badly in the situation of many irrelevant features. Unfortunately, a frequently encountered constraint when working with physiological data is that the numbers of potential or candidate predictors tend to be of the same order of magnitude or larger than the number of available observations for training. We therefore sought to perform effective feature subset identification given this constraint.

A complete set of candidate features that compactly represent the original physiological data set were identified. A total of $M=23$ features were selected for the candidate feature set. All features were extracted over a 3 minute quiet period preceding and following the PVT test and also the 10 minute period during the PVT test. Each feature was subsequently normalized to the initial baseline period.

For each subject, and for each of time period, a vector of features plus a constant term was formed:

$$\boldsymbol{x}_i = [1, x_{i1}, x_{i2}, \dots, x_{iM}]^T \tag{1}$$

The mean reaction times from the PVT tests were pooled together for each subject and each test to form a response vector:

$$\boldsymbol{y} = [y_1, y_2, \dots, y_N]^T \tag{2}$$

where N is the number of time periods multiplied by the number of subjects. All features were pooled together into a single N by M+1 matrix, $\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}, \dots, \boldsymbol{x}_N]^T$.

When the number of predictors M exceeds the number of training samples N, the modeling problem is underdetermined, or ill-posed in the Hadamard sense. In this instance, it is desirable to find a model with significantly fewer predictors and in fact, the more sparse the model, the more likely that the predictors are causally related to the dependent variable. In order to achieve this for ill-posed data, the values of the regression coefficients $\boldsymbol{b}$, can be constrained via a shrinkage function.

$$\boldsymbol{b} = \arg\min_b \|\boldsymbol{y} - \boldsymbol{X}b\|^2 + \lambda \sum_{i=1\dots M} |b_i| \tag{3}$$

where λ is a tuning parameter.

This method is called the Least Absolute Shrinkage and Selection Operator (LASSO), and is a regularization and selection method [14]. For each value of the tuning parameter, the LASSO will generate a sparse solution by setting many of the parameters to 0. The LASSO is a particularly attractive algorithm as it uses L^1 norm which can be viewed as the most selective shrinkage function that remains convex. Since a convex function has a global minimum and no local minima, convexity guarantees that we can find the one global solution for a given dataset. We used a highly efficient algorithm for solving the LASSO, termed Least Angle Regression (LARS) [15]. This algorithm converges to the final solution in M steps.

For each solution returned by the LARS algorithm, we evaluated the accuracy of the result using a leave-one-out cross validation procedure. We iterated through each subject, forming a test vector of PVT scores for that subject and a training matrix with the remaining subject. At each iteration we calculated a first order correlation coefficient between the predicted PVT scores and the test PVT scores. These were then averaged to provide a single statistic. The set of features that provided the largest statistic was subsequently selected as the final reduced feature set.

Both predictor and response data were centered and normalized to unit deviation prior to running the feature selection algorithms.

3 Results

Table 4 shows the LASSO selected predictors for each studied epoch. Also displayed are the regression accuracies using the leave-one-out cross validation procedure. No more than two features were selected for each task and this may be due to the fact that many features are well correlated amongst themselves.

As an illustration of prediction performance, figure 1 shows the prediction results for the 3 minute post PVT test period. Note that the data in this figure is normalized and centered to unit deviation.

Table 4. Selected features and prediction accuracy

Task	Selected features	Accuracy
3 minutes prior to test	BRV; LFnorm	0.52
PVT test	Motion	0.74
3 minutes post test	BRV; PSI	0.72

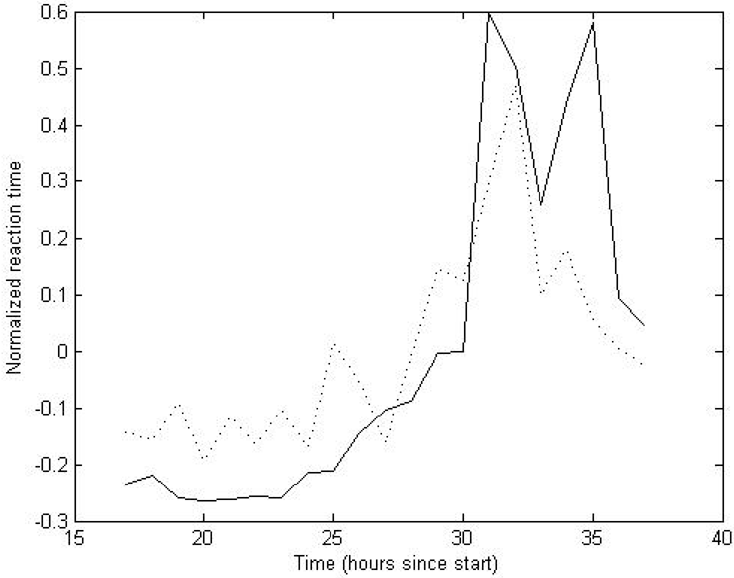

Fig. 1. Prediction results for the 3 minute period post PVT test. The solid line shows the normalized and centered mean PVT reaction time for each continuous hour of wake across subjects. The dotted line shows the regression prediction using the LASSO selected features (BRV; PSI).

4 Discussion

In-between PVT sessions, the participants were free to engage in any type of activity. As a result of this, the levels of autonomic activity during the 3 minute quiet period prior to the PVT test may have been dominated by the activities during the free periods. This is potentially why this period showed the poorest correlations. However, it remains interesting to note that the features selected during this period both strongly represented sympathetic arousal (the breath rate variability and the normalized low frequency power of the HRV). Although it is expected that sleep restriction increases sympathetic system outflow, the weak prediction accuracy may have been due to additional changes in this system caused by uncontrolled activity during the non-test periods.

The feature selection results during the actual PVT test initially appear to be surprising. Here the mean level of body motion was selected as the strongest predictor of cognitive fatigue. Taking a PVT test demands a degree of concentration and during alert periods subjects were more focused and able to concentrate on the task with minimal distraction. However, as the subjects became more fatigued, they tended to become increasingly restless during administration of the PVT test. This restlessness was captured by background body movements as recorded by the accelerometer and was directly correlated with PVT performance.

In the 3 minute stationary rest period following the test, the selected features were the breath rate variability (BRV) and the phase synchrony index (PSI). The BRV represents increases in sympathetic outflow and it has previously been noted that this increases during sustained attention tasks with increasing sleep restriction. The PSI represents the synchronization between the cardiac and respiratory systems and this increase in synchronization only manifested during the rest period following a demanding attention task. It is has been shown that synchronization and modulation (RSA) are two different competing aspects of cardio-respiratory interaction. Often when synchronization goes up, RSA goes down [11] and thus it may indicate changes in sympathetic activity. It is further possible that synchronization plays a homeostatic role in returning the system to baseline levels following increases in sympathetic system outflow. This would explain why there is an increased synchronization drive following high attention tasks when sleep restricted. However further research is required to better understand this feature.

This study demonstrates the power of the LASSO based feature selection paradigm to select a parsimonious physiological feature set. With appropriate individual sleep restriction data, this method could be used to perform individualized feature selection accounting for individual differences in fatigue vulnerability. It is also important to note that accurate context identification is of fundamental importance for any automated fatigue prediction system.

5 Conclusion

The LASSO feature selection technique allows one to select en-masse, via a continuous subset optimization, the set of variables that together are effective predictors of operator alertness status. This technique combined with commercially-available, wearable physiologic monitoring systems is a further step toward a system that can improve operational safety and effectiveness by accurately assessing cognitive fatigue levels during stressful day to day conditions.

References

1. Kleitman, N.: Duration of sleep. In: Sleep and Wakefulness, pp. 114–121. University of Chicago Press, Chicago (1963)
2. National Sleep Foundation. In: 2005 Sleep in America poll. Washington, National Sleep Foundation (2005)
3. Kripke, D.F., Garfinkel, L., Wingard, D.L., Klauber, M.R., Marler, M.R.: Mortality associated with sleep duration and insomnia. Arch Gen Psychiat 59, 131–136 (2002)
4. Carskadon, M.A., Roth, T.: Sleep restriction. In: Monk, T.H. (ed.) Sleep, sleepiness and performance. John Wiley & Sons, New York (1991)

5. Arnedt, J.T., Wilde, J.S., Munt, P.W., Maclean, A.W.: Simulated driving performance following prolonged wakefulness and alcohol consumption: Separate and combined contributions to impairment. J. Sleep Res. 9, 233–241 (2000)
6. Gundel, A., Marsalek, K., ten Thoren, C.A.: Critical review of existing mathematical models for alertness. Somnologie - Schlafforschung und Schlafmedizin 11, 148–156 (2007)
7. Zhang, C., Zheng, C., Yu, X.: Automatic recognition of cognitive fatigue from physiological indices by using wavelet packet transform and kernel learning algorithms. Expert Systems with Applications 36, 4664–4671 (2009)
8. Task Force of the European Society of Cardiology the North American Society of Pacing Electrophysiology: Heart Rate Variability. Standards of Measurement, Physiological Interpretation, and Clinical Use.: Circulation 93, 1043–1065 (1996)
9. Grossman, P., Kollai, M.: Respiratory sinus arrhythmia, cardiac vagal tone, and respiration: Within- and between-individual relations. Psychophysiology 30, 486–495 (1993)
10. Rajendra Acharya, U., et al.: Heart rate variability: a review. Medical and Biological Engineering and Computing 44, 1031–1051 (2006)
11. Lotric, M.B., Stefanovska, A.: Synchronization and modulation in the human cardiorespiratory system. Physica A: Statistical Mechanics and its Applications 283, 451–461 (2000)
12. Wilhelm, F.H., Trabert, W., Roth, W.T.: Physiologic instability in panic disorder and generalized anxiety disorder. Biological Psychiatry 49, 596–605 (2001)
13. Dinges, D.F., et al.: Cumulative sleepiness, mood disturbance, and psychomotor vigilance performance decrements during a week of sleep restricted to 4-5 hours per night. Sleep 20, 267–277 (1997)
14. Tibshirani, R.: Regression Shrinkage and Selection via the Lasso. Journal of the Royal Statistical Society. Series B (Methodological) 58, 267–288 (1996)
15. Efron, B., et al.: Least Angle Regression. The Annals of Statistics 32, 407–451 (2004)