Development of operator functional state evaluation methods in the maritime domain

Wu, Yanbin

(Degree)
博士（工学）
(Date of Degree)
2017-09-25
(Date of Publication)
2018-09-01
(Resource Type)
doctoral thesis
(Report Number)
甲第7015号
(URL)
https://hdl.handle.net/20.500.14094/D1007015

（別紙様式３）

論文内容の要旨

氏　名　　毋 岩斌 (WU Yanbin)

専　攻　　海事科学専攻

論文題目（外国語の場合は，その和訳を併記すること。）

Development of operator functional state evaluation methods in the maritime domain

(船舶運航管理実務者の機能評価手法の開発)

指導教員　　内田　誠

（注）2,000字～4,000字でまとめること。

Chapter 1 Research background

In human machine systems where the risk of accident is closely related to human component such as in maritime operation, it is crucial to maintain an optimal Operator Functional State (OFS). The complexity of these modern systems, the rapid update of contemporary technologies and reduced manning level all contribute to the high cognitive demands experienced by on-board crewmembers. The core element of ERM and BRM is to appropriately allocate resources, among which personnel resource is the most important one, considering the priority of tasks. While the condition of machinery and information resources are relatively straightforward, the condition of every crewmember is implicitly difficult to quantify. With knowing the functional state of operators, the allocation of personnel resources can be conducted more effectively. The purpose of this research is to categorize the human factors that lead to suboptimal operator performance and to develop different methods to recognize those suboptimal factors. The evaluation methods of OFS fall into three general categories: subjective rating scales, physiological signal based objective measures, and performance based measures.

Chapter 2 Preprocessing of physiological signals

In the work environment of maritime domain, operators have to complete required tasks in an ambulatory situation. The necessity of pre-processing physiological signals is emphasized by the fact that many artifacts would decrease the signal quality, especially when the subject has many body movements. In chapter 2, a combination of recursive percentage filter and median filter is used to detect and replace outliers of RR interval series. In an attempt to eliminate artifact of single channel EEG, which is always contaminated across all relevant power bands, an accelerometer was directly attached to EEG electrodes to measure electrode vibration. We used a portable EEG device (Digital electronic, Japan) with two channels. Channel 1 is for scalp voltage measurement (EEG electrodes), and channel 2 is an accelerometer attached directly to the electrodes to measure electrodes vibration. EEG epochs that are contaminated by movement artefact can be detected based on the power of channel 2. A linear model that based on calculating covariance and maximizing independence is proved effective in reducing artifact of small amplitude across wide range of power bands (1-40 Hz). If it is hypothesized that the signal measured by EEG electrode E (channel 1) is the linear sum of clean EEG signal S and artifact caused by electrode movement measured accelerometer V (channel 2), we can have the following linear model with two unknown constants k and b. Calculate the covariance between E and V,

$$cov(E,V) = cov(S,V) + k \cdot cov(V,V) + cov(b,V)$$

As S and V are from different resource and if we maximize the independence between S and V, we can have

$$k = cov(E,V) / D(V)$$

$$S + b = E - k \cdot V$$

In actual situation, both head movement and verbal communication happen occasionally, resulting intermittent contaminated signal. Therefore, the solution of k and b can be different when signals are contaminated by different source. Epochs are recognized as contaminated by body movement artifact when the power of channel 2 exceeds a pre-set threshold value. As shown in Figure 1, the subject shake his head for one time and the power of original EEG is abnormally high (up to 28 db) in 0-5Hz band. The large difference between green line and red line indicates that the regression effect is obvious. The processed signal (green line) is more close to the nature of clean EEG power spectrum.

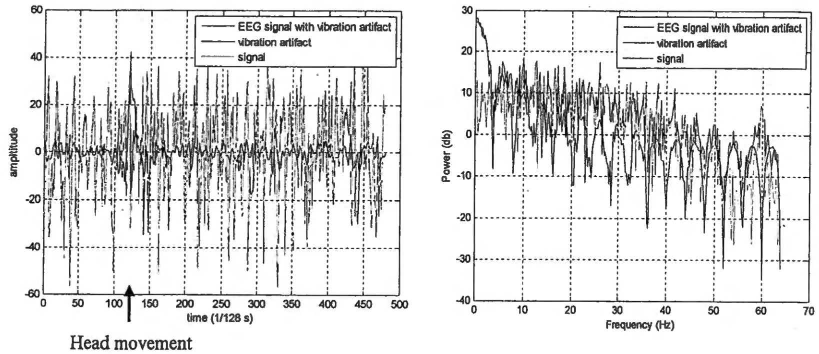

Figure 1 Processing of contaminated EEG signals: head movement happened once in reading a book loudly k=0.552, b=59.5. Left: time domain; Right: frequency domain

Chapter 3 Evaluation of mental workload of standard task and simulator task

Chapter 3 is an experiment study of using physiological features to evaluate operator' s mental workload in conducting two kinds of tasks: standard reference task and engine-room simulator task. Ten male and one female participants were recruited to carry out standard 4-level n-back tasks (E-prime) and simulated 4-level maritime operation tasks. The task demand of n-back task was manipulated by increasing the number of 'N' from zero to three, and the task demand of maritime operation was manipulated by involving a different number of operation procedures and types of pipes. EEG and HRV signal were measured by ambulatory devices. EEG measures: Alpha wave (8-13 Hz) rate; Beta wave (20-25 Hz) rate; Theta wave (4-7 Hz) rate; Time domain data were sampled in 128 Hz and discrete fast Fourier transformed (DFT) to frequency band. Signal epochs contaminated by body movement were detected by power of vibration signal and removed from analysis. Heart rate measures: Mean value of heart beat interval, Standard deviation of heart beat interval, and LF (0.04-0.15 Hz) to HF (0.15-0.4 Hz) ratio. The response accuracy and time consumed to fulfil operation target were used as performance measure, respectively. Six physiological features, a subjective rating scale (NASA-TLX), and performance measures are used to correlate with task demand.

Table 1. Pearson' s correlation coefficients between performance measures, NASA-TLX and Physiological indices

Measures	n-back task		MEPS task	
	Performance	NASA-TLX	Performance	NASA-TLX
NASA-TLX	0.634	1.000	0.717	1.000
Alpha wave rate	-0.393	-0.653	-0.254	-0.063
Beta wave rate	0.241	0.404	-0.525	-0.565
Theta wave rate	0.170	0.237	0.568	0.531
MHBI	0.062	-0.219	-0.263	-0.149
SDHBI	0.060	-0.081	-0.008	0.093
LF/HF	0.275	0.313	0.032	-0.016

The conclusions of Chapter 3 are:

(1) Alpha band wave suppression and subjective self-report MWL are sensitive to n-back task demand while heart ratio related measures are not. Alpha (8-13Hz) band wave suppression, beta wave band (20-25Hz), and LF/HF correlate with subjective MWL in n-back task

(2) Three EEG features are sensitive to MEPS task demand. Beta band (20-25Hz) and theta band (4-7Hz) correlate with subjective MWL in MEPS task.

(3) Ceiling effect of using physiological metrics to infer human MWL were found.

Chapter 4 Evaluation of operator fatigue and performance in pipeline work

Human fatigue caused by either physical exertion or mental strain is one of the most significant factors that constrain operator's functional capability to fulfil specific tasks. To ensure working performance and to improve occupational health, Chapter 4 aims to develop a quantitative method to evaluate operator fatigue during conducting pipeline works. A Japanese version of RPE scale and heart inter-beat intervals are measured in an experiment study. Hurst exponent (HE) is extracted from detrended fluctuation analysis to define the fractal structures of RR interval time series. DFA was widely used to study the fractal properties and long-term autocorrelations of nonstationary time series. Fractal geometry originally depicts the roughness of a surface and is applicable to time series data in the following essence: a process with stronger fractal characteristics does not adhere to equilibrium around any specific scale such as a constant heart rate. HE of biomedical signals generally ranges from 0.5 to 1.5. HE is 0.5 when the signal is white noise, while 1.0 indicates a pink noise and 1.5 indicates a brown noise. HR of rest, pipeline work, and recovery baseline were 72.5±8.0, 98.0±11, and 83.7±12, respectively. After cropping RR interval series into epochs of tightening every two flanges, Borg's RPE highly correlated with the decrease of RR interval. It corresponds with the former study that RPE scale is able to track the changes in HR. SD of tightened torque of each flange was considered as one performance measure.

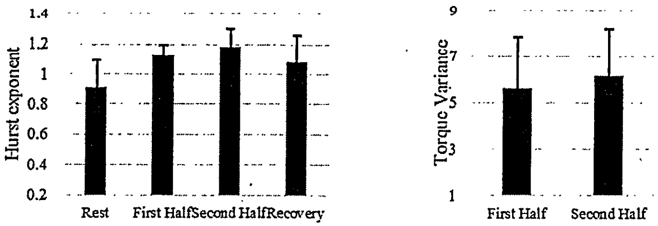

Figure 2 a. HE of rest, first half of pipeline work, second half of pipeline work, and recovery condition. b. Torque variance during first half and second half of pipeline work.

The first and latter half of torque variance were 5.6±2.1 Nm and 6.2±2.0 Nm (Figure 2 b), indicating that the bolts were more evenly tightened in the first half when participants' degree of fatigue was lower. HE of rest, first half of pipeline work, second half of pipeline work, and recovery baseline were 0.91±0.17, 1.12±0.07, 1.19±0.10, and 1.08±0.18, respectively (Figure 2 a). HE of working condition is significantly bigger than that of rest condition (p=0.003). HE during the second half of pipeline work was significantly higher than first half (p=0.04), which might be caused by fatigue.

RR interval series were cropped into segments that corresponded with the onset of tightening each flange and ended 40 seconds later after accomplishment of tightening each flange. A weak positive correlation (Figure 3) between HE and torque variance was found for five participants. The average correlation coefficient was 0.26±0.1. It indicates that the participant's working accuracy was lower when HE was higher, although this cause-and-effect relationship was unclear.

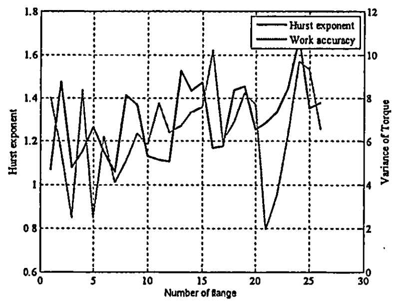

Figure 3 Correlation between Hurst exponent and work accuracy

Chapter 5 Estimation of mental workload in actual ship operation

In chapter 3, experiments were conducted in laboratory and simulator environment. Physiological signals were successfully measured under a relative controlled situation. However, real on-board environment is subject to many noise resources such as vibration and ship motions. For different individual subject, reference tasks whose difficulty is manipulated in a same range and pace, may elicit different arousal level. This is partially reflected by the results in chapter 3 that subjects show different tendency of ceiling effect. In addition, task of different difficulty level can elicit similar arousal level, represented by close values of physiological features. Therefore, in Chapter 5, a real

world experiment that studies the mental workload of the first engineer of training ship Fukaemaru is described. During the leaving port and entering port condition, the subject was in charge of operating and monitoring the engine system in the engine control-console area. The subject had to verbal communicate with other engineers inside the engine room, the on-duty navigation officers, and the chief engineer. He was free to walk around in the ship bridge, where the engine control console was set. One information flow model that divides operator's mental capacity into four channels: visual, auditory, cognitive, and psychomotor (VACP) is used to analyze subject's behavioral information. The weight of each channel is assigned with an orderly scale according to the mental workload exposed to the subject. The total VACP score is calculated by summing up the weight of four channels. Individualized combination of physiological features are decided based on the clustering quality in n-back task, which can be quantitatively evaluated from an I-index $I(\mathrm{K})=\left(\frac{1}{K}\times\frac{E_1}{E_K}\times D_K\right)^p$. Two methods are used to calculate the distance. First, Euclidean distance is generally defined as point-to-point distance. Euclidean distance is also known as geometrical distance. Before calculating Euclidean distance, the variables of input features must be normalized to eliminate the effect of different scales and units. Mahalanobis distance is a multi-dimensional generalization of the idea of measuring how many standard deviations away a point is from the mean of a sample. Through a selection process of calculating I-index based on Euclidean distance, feature vector of [r_α, r_θ, meanHBI, sdHBI, LFHF] under 0-back, 1-back, and 2-back task are considered as low, normal, and high MWL training data set respectively. The input features for Mahalanobis distance based classifier is [r_α, r_β, meanHBI, sdHBI, LFHF]. The k nearest neighbour (k-NN) classifier is one of the most basic pattern recognition methods. The basic principle of k-NN is the intuitive idea that data points belong to a same cluster should be close to each other in the feature space. To decide the choice of parameter k, ten-fold cross validation is used to check the within n-back task classification accuracy. Ten-fold cross validation is to randomly and averagely divide the sample into ten sets, nine sets of data as training data and the left one set is used as test data set. This validation process is repeated ten times to ensure all data has been used as test data. The parameter k varies from 1 to 19, and the result shows that the error rate is minimum when k=3.

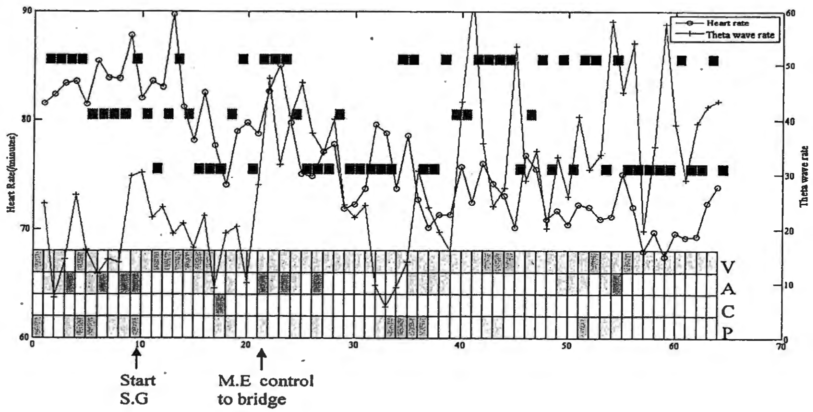

Figure 4 Estimated mental workload based on Euclidean distance: Leaving port

Mental workload is estimated by Euclidean distance based classification method and Mahalanobis distance, respectively. Figure 4 shows one analysis result of leaving port condition. The classification period is 23:00-27:16, lasted 256 seconds. By using the training data of n-back task, MWL in conducting maritime operation tasks is labeled based on k-NN. The red, green, and blue rectangle represent high, normal, and low MWL. The blue and red line show the instantaneous change of heart rate and theta wave rate. The rectangles with different gray scale represent different weights of VACP score. Higher gray scale represents higher weight of VACP. The linear correlation between estimated MWL and VACP score is checked by calculating Pearson' correlation coefficient r. A low positive correlation is found in leaving port condition, with r=0.16.

Chapter 6 Conclusions, summary, and future research

The effectiveness and safety of many real-world complex human-machine systems rely on the normal functional states of both human operators and the machinery systems, and their cooperative interactions. This paper focuses on the research of modelling and quantitatively evaluating the human operator functional state and its interaction with machinery systems. The findings of this paper provide a potential solution to reduce human error and to improve human operator's comfort in the maritime domain.

To manipulate the subjects' cognitive state, standard psychological experiment schemes (e.g. auditory/visual n-back) were developed by using E-prime Software. Simulator based operation scenario, on-board field study, and quasi real-world tasks were designed to ensure the successful manipulation of' physical and mental state in various environment. In Chapter 5, human functions were decomposed into four channels: Visual, Auditory, Cognitive, and Psychomotor (VACP), whose quantitative integration was used as one of the ground truth measure. Subjective questionnaires (NASA-TLX, Borg's RPE scale etc.) were also collected as reference information.

Body movement, vehicle vibration, loose electrodes contact, and verbal communication can easily

contaminate EEG and HRV signals. A combination of recursive percentage filter and median filter was used to detect and replace outliers of RR interval series. In my attempt to eliminate artifact of single-channel EEG, which is always contaminated across all relevant power bands, an accelerometer was directly attached to EEG electrodes to measure electrodes vibration. A linear model that based on calculating covariance and maximizing independence has been proved effective in reducing artifact of small amplitude across wide range of power bands (1-40 Hz). In an experimental study, the subject's EEG signal was measured in three situations, keep still, verbal speaking, and walking around. The linear model eliminates body movement artefact caused by verbal communication.

Both time-domain and frequency domain (FFT based) features are extracted and their effectiveness are firstly studied by using ANOVA, T-test, and Correlation analysis. In choosing proper combination of these features, clustering quality is determined by I-index, which is based on the ratio between cross clusters distance and within cluster distance. The distance between cluster centers are calculated by Euclidean distance and Mahalanobis distance. The best combination of features are chosen when I-index reaches its maximum value. In using RR interval data to evaluate operator fatigue, Hurst exponent is extracted by Detrended fluctuation analysis (DFA), it is found that RR interval shows more autocorrelation structures during work situation than in rest situation, representing a larger Hurst exponent in work situation. Furthermore, average RR interval highly correlates with the subjective fatigue scale, Borg's RPE scale.

The findings of this paper would become more applicable to real world scenarios if the following aspects can be improved. First, in EEG measurements, eliminating artifacts caused by body movement is still a troublesome issue. According to the finds in this paper, I plan to focus on developing decomposition methods that utilize information of electrode vibration measured by accelerometer. Second, wearable sensors are rapidly being improved; a possible future direction may be to integrate these sensors into objects of daily use, e.g. EEG measuring safety helmet, heart rate monitoring undergarments, and view tracking glasses. Third, gain high-quality training data by having subjects conduct standard psychological task schemes that are able to elicit physiological responses close to real-world tasks.

（別紙１）　　　　　　　　　　　　論文審査の結果の要旨

氏名	毋　岩斌（WU YANBIN）		
論文題目	Development of operator functional state evaluation methods in the maritime domain （船舶運航管理実務者の機能評価手法の開発）		
審査委員	区　分	職　名	氏　　　名
	主　査	教　授	内田　誠
	副　査	教　授	堀口　知也
	副　査	教　授	段　智久
	副　査		印
	副　査		印

要　　旨

IMO（国際海事機関）により採択され日本が批准している STCW（船員の訓練及び資格証明並びに当直の基準に関する国際条約）の 2010 年マニラ改正で、ERM（機関室資源管理）が強制要件化され、Non-technical Skills（非専門技術）に関する教育訓練ならびにその評価手法の重要性が高まった。本研究では、船舶機関管理を焦点に、船舶運航管理実務を担う海技者の機能について、生理学的信号、主観的負担、行動記録から普遍的な評価手法の開発を試みるものである。

本論文は 6 章で構成されており、第 1 章では研究背景と研究目的が述べられ、第 6 章では総括として本研究の成果と今後の展望がまとめられている。研究内容の骨格となる第 2 章から第 5 章の概要は以下の通りである。

第 2 章では、第 3 章から第 5 章で分析対象とする生理学的信号の 1 つである脳波に関し、分析前の信号処理について論述している。船舶運航管理実務者を対象とした本研究では脳波計測対象者の行動（特に、頭部の動き）などの影響が、脳活動による電位変化にノイズとして影響を及ぼす。頭部加速度信号を活用した共分散分析による計測信号の汚染除去手法が構築され、その効果が検証された。

第 3 章では、舶用機関プラントシミュレーター環境において、船舶機関管理実務を模擬したタスクを設定し、船舶機関士を想定した被験者の心的負担（MWL：Mental Workload）の種類及びレベルと、脳波、心拍の生理学的特長との相関を議論するための実験が行われた。負荷レベルを 4 段階に設定可能な標準タスクにおける計測結果を参照情報とした分散分析および複数指標によるマハラノビスの汎距離に基づくクラスター分析から、船舶機関管理実務者の機能適応状態を生理学的信号から評価する可能性が示された。

第 4 章では、船舶機関管理実務者の運動負担による疲労に注目し、配管作業における疲労の定量的評価の開発が試みられた。持続的作業中の被験者の心拍変動についてトレンド除去変動分析を行い、作業の進行に伴い心拍間隔が短くなる傾向と、被験者の RPE（知覚努力レベル）スケールが上昇する傾向に強い相関があることが示された。また、ハースト指数（トレンドを測る指標）と作業精度の相関が示され、疲労による作業への影響を定量的に評価する可能性が示された。

氏名	毋　岩斌（WU YANBIN）

第 5 章では、実船環境における船舶機関管理実務を焦点に、出入港時の機関操作および監視に携わる機関士の MWL を生理学的信号から把握することが試みられた。また、行動記録（ビデオ映像）から 4 つの指標（VACP：視覚、聴覚、認知、精神運動）による行動分析の時系列と生理学的信号の相関が議論された。

以上のとおり、複雑な人間ー機械システムの典型である船舶機関管理における実務者の Non-technical Skills に関する機能評価を客観的かつ普遍的に実施する手法を開発するために必要な課題について、シミュレータ環境および実船環境において実験研究が行われ、生理学的信号、主観的負担および行動の相関を明らかにするなど、研究内容の独自性及び明確な研究成果が認められる。本研究では、船舶機関管理を焦点に、船舶運航管理実務を担う海技者の機能評価手法の開発に取り組まれたが、生理学的信号処理、主観的負担の把握および行動記録から機能評価を試みる手法は、適用対象を船舶に留まることなく、自動車、鉄道、航空機などの交通輸送管理実務および陸上プラント管理実務など、広範な人間ー機械システムにも適用可能であり、人と機械が絡む複雑システムにおける安全向上の観点で重要な知見を得たものとして価値のある集積である。

提出された論文は海事科学研究科学位論文評価基準を満たしており、学位申請者の毋岩斌（WU YANBIN）は、博士（工学）の学位を得る資格があると認める。