A new methodology to evaluate Human Workload State Inference

A. Malagoli, M. Corradini, S. Fonda
Department of Life Sciences, University of Modena and Reggio Emilia

***Abstract*— Electronics and automation are increasingly part of our daily lives, and led to the introduction of systems and intelligent machines to which the human work is delegated and that collaborate and support the user in the conduct of man-critical operations.**

The University of Modena and Reggio Emilia is partner of the european project "Designing Dynamic Distributed Cooperative Human-Machine Systems" (D3CoS, www.d3cos.eu), that aim is the definition of affordable methods, techniques and tools which go beyond assistance systems and consequently address the specification, development and evaluation of cooperative systems from a multi-agent perspective where human and machine agents are in charge of common tasks, assigned to the system as a whole. The key on which to base the cooperation between the machine and the human is the amount of workload of the human operator.

So we were involved into investigate aspects of the functional state of human operators interacting with the demonstrator in the unmanned aerial vehicle (UAV) and maritime domains. We analyzed and correlated objective psycho-physiological measures: eye blinks, respiration rate and amplitude, electro dermal activity, heart rate variability, blood pressure; with subjective self-assessed measure evaluated with two questionnaires: NASA-TLX and Rating Scale Mental Effort (RSME); with the aim to realize a mathematical regression model for classifying the mental operators workload.

***Keywords*— Workload, psycho-physiological measures, statistical analysis, cooperative systems.**

I. Introduction

Mental workload (WL) is a multidimensional and complex construct for which there is no clearly defined and universally accepted definition in literature: there are inconsistencies relative to its sources, mechanisms, consequences that, in consequence, issue in a variety of assessment methods. In general it can be considered as a "quantitative function" of the relationship between the mental demands required by a task and the available resources of the human operator.

In order to gather a valid and reliable assessment of WL it is necessary to employ a set of measures according to: (1) psycho-physiological measures and (2) self-assessed measures (rating scales).

For (1) we decided to acquire information by:

- **Autonomic nervous system measures**: with EDA (Electro Dermal Activity) which is linearly correlated to arousal and measures stress and frustrations too.
- **Ocular activity**: eye activity reflects central nervous system activity and indicates also mental WL . In particular the eye blink rate yields information about task demands, level of fatigue, memory and response demands. [3]
- **Cardiovascular** measures: they have been reported to be sensitive to WL , and emotional activation which was evaluated by means of self-assessed measures. ECG signal, Heart Rate, HRV (Heart Rate Variability), LFHF ratio of HRV spectrum and Photoplethysmography (PPG) were measured directly. Blood Pressure (BP) was indirectly calculated from ECG R-peak and PPG peak delay with single calibration.
- **Respiration**: respiration rate, inspiration and expiration time, complete cycle time, volume and flow rate was measured.

For the (2) we used two different questionnaires: a NASA-TLX (National Aeronautics and Space Administration Task Load Index) and RSME (Rating Scale Mental Effort) . The NASA-TLX produces a numeric index as result of a sum of six values: Mental Demand, Physical demand, Temporal demand, Effort, Performance, Frustration that is indicated as Overall UnWeighted Score (OUWS); while the RSME is a single index checked by the operator.[1]

II. Methods

We performed two different experiments in which we tested our methodologies and tools in the UAV and Maritime scenario. In the first case at Selex Galileo (UAV Scenario), Ronchi dei Legionari - Italy, on four UAV pilots; and in the second one at Kongsberg (Maritime Scenario), Horten - Norway, also on five Vessel Traffic Service (VTS) operators.

In each experiments we individuate, according with our partners, four different workload scaled scenarios (underload, medium, high and overload task). At Selex Galileo, one trainer provided instructions to UAV pilot. The workload change in this case was due to the increasing number of of tasks. Instead, in the maritime domain, the workload modification was due to different factors: increasing number of vessels in the monitored area, increasing number of vessel reporting, increasing number of software alerts and increasing number of communication between VTS operator and vessels. In this case an external person simulated the vessels communication.

The duration for each scenario was about 10 minutes.

All the psycho-physiological measures was recorded by BioSemi Mod. Active Two and sampled at 2048Hz, expect ocular activity that was recorded by Jazz Novo – Ober Consulting and sampled at 1000Hz. At the end of each scenario a NASA-TLX and RSME questionnaires were compiled by the operator.

In addition to conventional digital filtering techniques, Empirical Mode Decomposition (EMD) [2] processing was applied to all the recorded signals, in order to identify and separate signal information from noise; we also used this technique to separate the tonic and phasic in EDA responses and to evaluate the respiration rate and amplitude; furthermore we introduced a linear model to get systolic blood pressure as a function of the delay between R peak of ECG and blood "pulse wave transit time" peak, detected with plethysmograph [4] . We also developed a particular algorithm to detect blinks based on xy eyes movement components. The heart rate variability was analyzed estimating its spectrum behaviour in low (0.04-0.15Hz) and high frequency (0.15-0.4Hz) band, focusing into the low frequency over high frequency (LF/HF) ratio. The LFHF was calculated by the RHRV package in R Studio [5] .

For every WL phase all the psycho-physiological recorded data was split into 30s intervals in order to get many sets of equivalent measures and associated with those obtained in NASA-TLX and RSME questionnaires. The duration of 30s has been chosen because it is the minimum time value to give a reliable evaluation of many parameters, e.g. HRV, LFHF, EDA. All the data was analyzed with MANOVA methodology, to check the significant “variability” of parameters subsequently a Linear Discriminant Analysis (LDA) and Canonical Correlation Analysis (CCA) was performed with R studio software.

TABLE I
PSYCHOPHYSIOLOGICAL MEASURES

Parameter	*Definition*
MeanHR	Mean value of heart rate [bpm]
LFHRratio	Low frequency to hight frequency ratio, from spectral analysis of heart rate variability (HRV) [number]
MeanSyst	Mean value of systolic blood pressure, base on pulse wave transit time R-PPG peak [mmHg]
MeanTonicEDA	Mean value of tonic component EDA [µSiemens]
MeanRatePhasicEDA	Mean rate value of phasic component of EDA [µSiemens]
MeanBreathRate	Mean respiration rate per interval [Bpm]
MeanBreathAmp	Mean respiration “peak-valley” amplitude [arbitrary units]
MeanBlinkRate	Mean value of the eye blinking per interval rate [EB]

§

III. Results

The results presented here are not in their final version and come from one subject; but they are significant to introduce the new approach on investigating human workload correlated to self-assessed and psycho-physiological measures.

Statistical analysis were performed on the collected measures to detect:

- Correlations (1) among single measures and the task demand, (2) among all the recorded measures;
- Psycho-physiological variables clusters related to specific task load/performance/subjectively experienced conditions;
- Psycho-physiological factors as main responsible for the performance outcome as revealed by previous clusters.

In first analysis both the p-values got from MANOVA test , for the couples variable-factor, than the CCA correlation between subjective and objective measures, show that practically every parameter contribute to the variance measured between the four workload phases .

Introducing the subjective variables provided by TLX and RSME questionnaires, the exploratory analysis shows interesting relationships between objective and subjective measurements Fig. 1. The Fig. 4, contain the parameters behavior vs questionnaire scale and represent a preparatory result to further accommodate a more finalized analysis.

The CCA coefficient [Fig. 3] also shows that NASA-TLX and RSME questionnaires are quite similar and the MeanHR and the Me are high related with subjective measurements. Furthermore, going from low to high workload, variability of the LF/HF ratio decreases along the phase duration. [Fig. 2]

IV. Conclusion

Actually, we are completing the algorithm for statistical analysis in order to create a model to classify the status of the human operator's workload on the basis of few psychological signals.

The threshold between acceptable/unacceptable subject states was defined on the basis of the results of the described analysis and the integration of the different methods results.

It is difficult to reach significant statistical conclusions with such a small number of subjects associated to so complex ambition as measuring human workload. On the other hand these results show the possibility to combine various methods (from signal processing by empirical mode decomposition, to advanced statistical analysis) in a single procedure and must be taken as a support towards the possibility of measuring an operator’s workload in an objective and quantitative scale.

Acknowledgement

We would like to thanks our D3CoS partners Selex Galileo and Kongsberg for their support.

V. References

[1] **Zheng, B.; Jiang, X.; Tien, G.; Meneghetti, A.; Panton, O. N. M. and Atkins, M. S.** (**2012**). *Workload assessment of surgeons: correlation between NASA TLX and blinks.*, Surg Endosc 26 : 2746-2750.

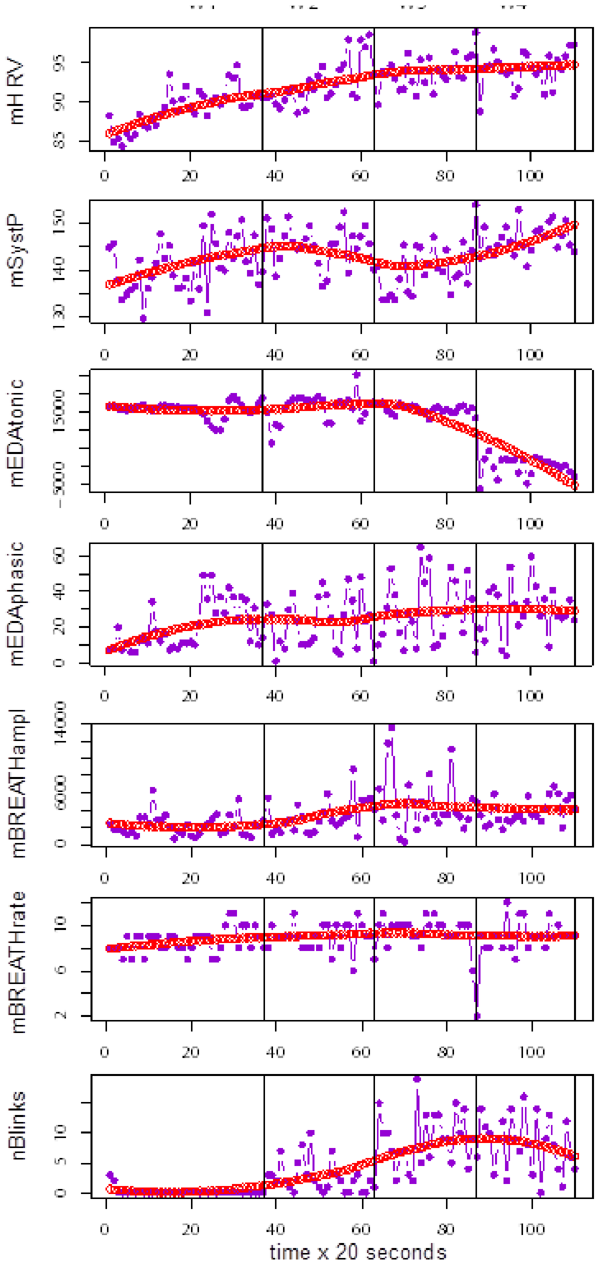
[2] **Li, Y.; Wang, Q. and Deng, Q.** (**2005**). *[Research on EMD and its application in biomedical signal processing].*, Sheng Wu Yi Xue Gong Cheng Xue Za Zhi 22 : 1058-1062.

[3] **Cardona, G. and Quevedo, N.** (**2014**). *Blinking and driving: the influence of saccades and cognitive workload.*, Curr Eye Res 39 : 239-244.

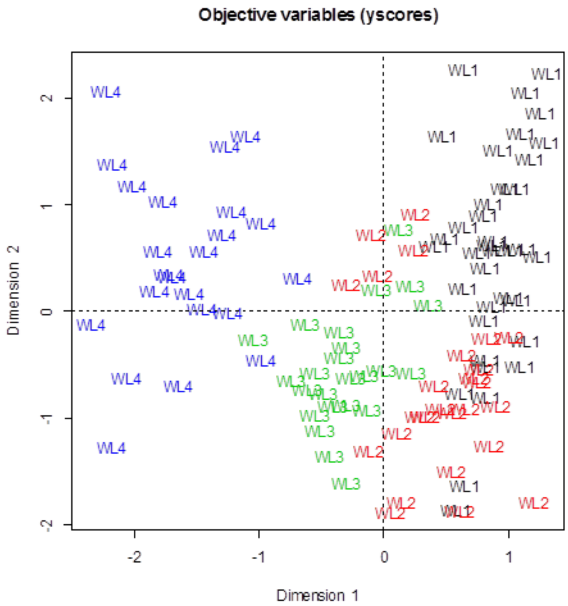
[4] **M. K. Ali Hassan Mashor M.Y., M. S. A. and S., M.** (**2011**). *Non-invasive Continuous Blood Pressure Monitoring Based on PWTT*, 1 : 63-73.

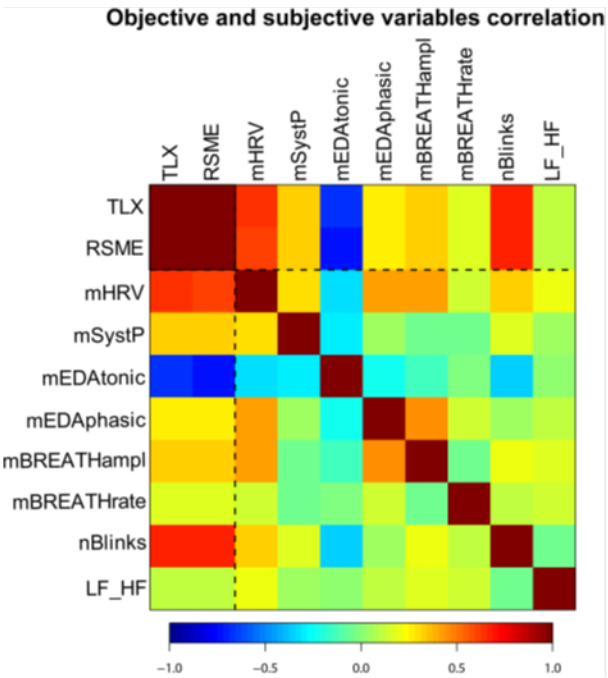
[5] **Rodríguez-Liñares, L.; Méndez, A. J.; Lado, M. J.; Olivieri, D. N.; Vila, X. A. and Gómez-Conde, I.** (**2011**). *An open source tool for heart rate variability spectral analysis.*, Comput Methods Programs Biomed 103 : 39-50.

Fig. 1 : *Comparison of LFHF ratio parameter of heart rate variability spectra, obtained in exercises W1(a), underload, and W4(b), overload.* **a:***The LFHF ratio in underload phase W1 is few minutes longer than the other phases.* **b:***The LFHF ratio in overload phase (W4) presents significant differences if compared to the underload phase.*



t in blue colour represents the mean value of the signal into the 30s interval. The red line is the result of LOESS interpolation algorithm





(CCA). No misclassified workload point

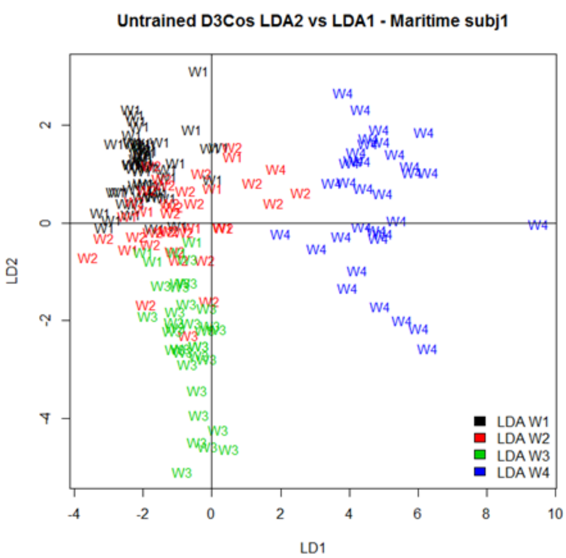


Fig. 2 *Linear Discriminant Analysis (LDA). Few misclassified workload points*