EEG-Based Human Factors Evaluation of Conflict Resolution Aid and Tactile User Interface in Future Air Traffic Control Systems

Xiyuan Hou, Fitri Trapsilawati, Yisi Liu, Olga Sourina, Chun-Hsien Chen, Wolfgang Mueller-Wittig and Wei Tech Ang

Abstract Currently, Air Traffic Control (ATC) systems are reliable with automation supports, however, the increased traffic density and complex air traffic situations bring new challenges to ATC systems and air-traffic controllers (ATCOs). We conduct an experiment to evaluate the current ATC system and test conflict resolution automation and tactile user interface to be the inputs of the future ATC system. We propose an Electroencephalogram (EEG)-based system to monitor and analyze human factors measurements of ATCOs in ATC systems to apply it in our experiment. The EEG-based tools are used to monitor and record the brain states of ATCOs during the experiment. Real-time EEG-based human factors evaluation of an ATC system allows researchers to analyze the changes of ATCOs' brain states during the performance of various ATC tasks. Based on the analyses of the objective real time data together with the subjective feedback from ATCOs, we are able to reliably evaluate current ATC systems and refine new concepts of future ATC system.

X. Hou (✉) · Y. Liu · O. Sourina · W. Mueller-Wittig
Fraunhofer IDM@NTU, Nanyang Technological University, Singapore, Singapore
e-mail: houxy@ntu.edu.sg

Y. Liu
e-mail: liuys@ntu.edu.sg

O. Sourina
e-mail: eosourina@ntu.edu.sg

W. Mueller-Wittig
e-mail: askwmwittig@ntu.edu.sg

F. Trapsilawati · C.-H. Chen · W.T. Ang
School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore, Singapore
e-mail: fitritra001@e.ntu.edu.sg

C.-H. Chen
e-mail: mchchen@ntu.edu.sg

W.T. Ang
e-mail: askwmwittig@ntu.edu.sg

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_73

Keywords EEG · Human factors · Brain states monitoring · Air traffic control

1 Introduction

Although performance and reliability of current Air Traffic Control (ATC) systems has been improved with automation supports, the increased traffic density and complex air traffic situations bring additional requirements and new challenges to ATC systems and air-traffic controllers (ATCOs). To design a future ATC platform which can provide more effective and robust handling of heavy traffic situations, states of art technologies like touch and tactile human computer interface, interactive 3D situation displays and advanced CRA software will be integrated to current ATC platform. However, the standard evaluation method that uses questionnaires after each assessment can only give an overall rating of the performed task. It cannot tell designers how the workload and emotions are changing during the task performance in a complex traffic situation. This information may be estimated from some other performance factors, but it cannot be done in high time resolution using traditional methods. So there is a need for the tools to objectively estimate how novel interfaces affect ATCOs during operations. To solve this problem, we propose to use reliable brain computer interface (BCI) to measure performance of ATCOs in different ATC experiments. By using such bio-signal technology with standard evaluation methods, we can enhance and refine design and development of a new ATC platform. To our best knowledge, we are the first to develop real-time workload, emotion and stress recognition algorithms that use fewer electrodes and have good accuracy that could be used for ATM system evaluation. A real-time brain states monitoring system in which emotion, attention, workload, and stress recognition algorithms can be recognized in real-time is applied for evaluation of the future workplace of ATCOs.

We conduct an experiment to evaluate the costs and benefits of conflict resolution automation and tactile user interface in future ATC systems. In the user study, we evaluate the current ATC system with conflict resolution automation and tactile user interface used as inputs. ATCOs and students with ATC knowledge are instructed to complete ATC tasks in three conflict resolution aid scenarios including reliable, unreliable, and manual conditions with or without tactile user interface. During this user study, objective human factors measurements including mental workload, stress, and emotion of ATCOs while performing ATC tasks are obtained real time using an Electroencephalogram (EEG) device.

In this paper, we propose an EEG-based system to monitor and analyze human factors measurements of ATCOs in ATC systems. The EEG-based tools are used to monitor and record the brain states of ATCOs during the experiment. In subjective human factors studies, the data of mental workload, stress, emotion et al. are obtained through questionnaires that are administered upon completion of each task or/and after an experiment. However, this method only offers the overall evaluation of ATCOs performance. Real-time EEG-based human factors evaluation of an ATC system allows researchers to analyze the changes of ATCOs' brain states during the performance of various ATC tasks. The data can be analyzed during or at any time interval starting from 1/32 s. Machine learning techniques are applied to the EEG data to recognize levels of mental workload, stress and emotion during each ATC task.

2 Related Work

Air traffic controllers (ATCOs) have to handle a significant amount of information that has to be interpreted and analyzed in a time critical manner. The increase in air traffic density is becoming a major issue in air traffic control. As reported by Sheridan [1] and International Civil Aviation Organization [2], worldwide traffic density will be up to double in 2025 compared to 2006. Given the limited airspace available, the possibility of having more air traffic conflict is unavoidable [3]. Under this circumstance, ATCOs will inevitably need a better support to overcome their cognitive limitations in handling more aircrafts in airspace. Providing support through automation is considered to be an effective solution to minimizing the workload imposed on ATCOs [4]. Endsley and Rodgers [5] discovered that when the air traffic increased, the controllers' awareness of each aircraft declined rapidly and when the workload was excessive, operational errors appeared. Recent research showed that conflict resolution aid (CRA) software has the potential to support ATCOs in resolving air traffic conflict in an effective and efficient manner [6] regardless of its imperfection [7] by advising ATCOs all the possible maneuvers. The future work place should therefore be designed to reduce mental workload and stress of ATCOs for optimal performance.

In research and development of human-machine interfaces, the evaluation of workload is a key point. Workload is described as a noticeable relationship between the human cognitive capacity and the effort required to process a particular task [8]. There are mainly three classifications for measurement of workload: subjective, physiological, and performance-based measures [9, 10]. Subjective measurement of levels of workload is based on the use of question-answer type response to measure the amount of workload a person feels during a task.

Currently, there are many subjective measure procedures designed to evaluate the mental workload as NASA Task Load Index (TLX) [11], Subjective Assessment Technique (SWAT) [12], and Cooper–Harper Scale [13]. NASA-TLX uses mental workload, physical demand, temporal demand, performance, effort, and frustration as six dimensions scales to evaluate mental workload. SWAT uses different three dimensional scales time load, mental workload, and psychological stress load as three discrete levels. But, Hill et al. [14] have proved that NASA-TLX is superior to SWAT in terms of measurement sensitivity especially for measurement of low workload.

Physical workload is the measurable portion of physical response of body when performing a given task and is affected by a range of factors. These physical responses include brain activity, cardiac activity, respiratory activity, and eye activity. Performance-based measurement of workload relies on examining some key parameters during a specific task which can reflect the capacity of a subject. In physiological measurement, electroencephalogram (EEG) interface is more suitable for monitoring people's mental workload because that EEG signals are directly captured from brain activity. EEG-based analyses have been widely used in clinical diagnosis of mental diseases and in bioengineering research. A number of EEG-based methods and corresponding applications are designed and implemented in order to recognize the user's workload levels [15, 16]. In [17], mental workload is evaluated in online EEG monitoring during the security surveillance task. Comparing the mental workload index with the error rate for the subjects, the correlation coefficient is approximately 0.7, which indicates that when the workload increases people have a tendency to make more errors. The correlation between workload and EEG signals has been proved in [18, 19]. In [18], the driver's mental workload is significantly correlated with theta band power and alpha band power. In different driving tasks, the frontal theta activity shows significant increases when working memory load increases. In another experiment studying the workload and fatigue in aircraft pilots [19], increased EEG theta band power and decreased alpha band power are observed in high mental workload comparing with the low mental workload. Additionally, in [19] it is shown that when the pilots have high mental workload and mental fatigue, their EEG theta band power as well as the delta and alpha bands power increases.

In this research, the EEG-based workload recognition, subjective user studies, and task performance are used together for evaluation of ATCOs' workload in different scenarios.

3 EEG-Based Workload Recognition

3.1 Feature Extraction

In our previous work [20], the real-time EEG-based brain states monitoring system CogniMeter is proposed to recognized emotion, workload, and stress. So, in this paper, we implemented the same algorithm for ATCOs' mental workload recognition based on FD and statistical features.

FD measures the complexity and irregularity of time series [21]. It can be used as an index for characterizing the complexities of EEG signals. For a regular signal, the fractal dimension value is low. If the signal becomes irregular, the fractal dimension value increases accordingly. Wang et al. [22] proposed to use Higuchi fractal dimension to recognize different arithmetic mental tasks from EEG. It is also used in EEG-based serious games to identify attention level. In this paper, the Higuchi algorithm is used to calculate FD feature for real-time workload recognition.

Statistical features are widely used in EEG based brain states recognition including emotion recognition algorithms [23]. Six statistical features such as mean, standard deviation, mean of absolute values of the first differences, mean of absolute values of the first differences of normalized signals, mean of absolute values of the second difference, and mean of the second differences of the normalized signals are extracted from EEG for emotion recognition.

3.2 Mental Workload Recognition

The mental workload recognition algorithm has been proposed in [24], in which the algorithm has been tested on the EEG database with different feature combinations and classifiers. For different feature combinations, the average accuracy of SVM classifier is 9.56 % higher than k-NN classifier based on mental workload EEG data. By combining statistical and FD features and using SVM classifier, the best accuracy is 90.39 % for 2 levels mental workload recognition and 80.09 % for 4 levels mental workload recognition. Therefore, in this experiment, we use FD and statistical features calculated from 14 channels and SVM classifier for mental workload recognition.

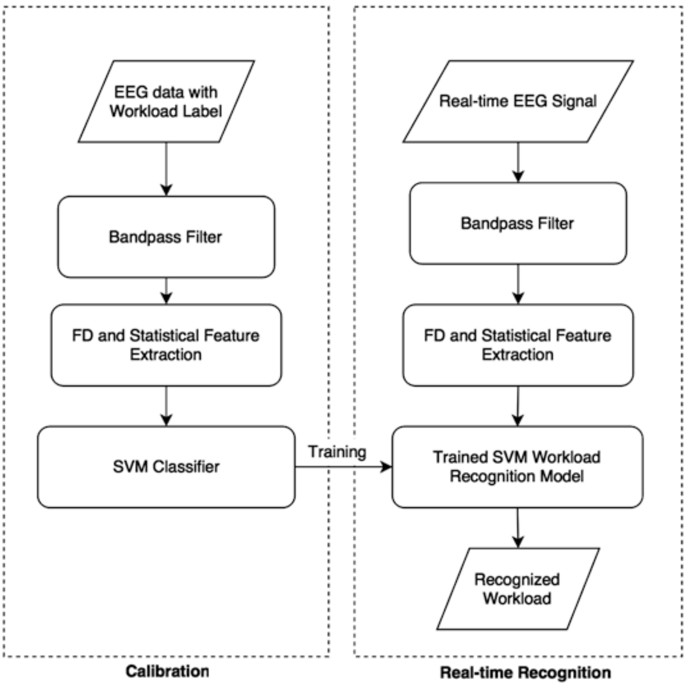

Fig. 1 The overall diagram of calibration and real-time brain states recognition algorithms for evaluation of mental workload [20]

The subject-dependent real-time workload recognition system consists of two parts: calibration and real-time mental workload recognition algorithm. The overall diagram of calibration and real-time workload recognition algorithms is shown in Fig. 2. In calibration, EEG data are labeled with different levels of mental workload for workload recognition correspondingly. Then, the EEG data are filtered, the corresponding features are extracted and the support vector machine (SVM) classifier is trained. After that, during real-time workload recognition, EEG signals are filtered and the FD and statistical features are extracted using a 4 s sliding window with 3 s overlapping. Next, new data features are input into the SVM classifier model trained in calibration. The classifier can recognize mental workload level based on each 4 s EEG signals input (Fig. 1).

4 Experiment

A preliminary experiment is designed and implemented to study the human factor in current ATC work place with some new features. 31 ATCOs and 5 students with ATC knowledge participated in the current user study and provided a signed-consent form that was approved by NTU IRB. All of them have received training of air traffic control and none of them has history of mental illness.

All participants were equally divided into three groups: Non-Display, Display, and Trajectory Prediction. Non-Display group was the baseline condition where participants were only equipped with the CRA. Participants in Display group were provided with the CRA and an additional display that depicted aircraft profile. In Trajectory Prediction group, participants were equipped with the CRA as well as an additional display that showed the prediction of aircraft trajectory including climb and descend rate information.

In every group, each participant performed ATC tasks in three CRA conditions: Manual, Reliable and Unreliable. In the reliable condition, the CRA was able to provide correct advisories to all the potential conflicts. In the unreliable condition, the maneuvering advisory provided an incorrect resolution advice that led to a conflict. In both reliable and unreliable CRA conditions, there was a conflict resolution advisory for each conflict and participants were free to either accept or reject the advisory by clicking a respective button. In the manual condition, participants were asked to resolve the potential conflicts by providing their own resolution maneuvering instructions.

In the experiment, there were three one-hour ATC scenarios corresponding to the three different CRA conditions. A balanced Latin square was adopted for the counterbalancing of CRA conditions to deal with any carry-over effects. In each scenario, participants were required to communicate with the pseudo-pilots to issue appropriate altitudes, to maintain separation between aircraft, to accept all aircraft that entered their sector, to hand-off aircraft that left their sector and to issue the correct radio frequency change.

4.1 Brain Computer Interface

In our experiment, Emotiv headset [25] is used to capture the users' EEG signals wirelessly with the USB receiver. It is a popular low-cost EEG device widely used for research including usability testing, neural marketing, serious games, etc. Emotiv EPOC has 14 channels located at AF3, F7, F3, FC5, T7, P7,O1, O2, P8, T8, FC6, F4, F8, and AF4 as shown in Fig. 2. During experiment, the EEG-based mental workload recognition system records EEG signals and recognizes ATCOs' workload in real time.

4.2 Workload Calibration and Recognition

As real-time EEG-based brain state recognition algorithms are subject dependent, calibration is required before real-time recognition. For calibration, four Stroop color-word test with different settings (congruent/incongruent ink colour or time limit) is used to induce different levels of workload. Each part of the test lasts for 1 min, and subject needs to fill a prompted questionnaire to evaluate his/her mental workload level on the scale from 1 to 9 and to describe his/her feelings in words as shown in Fig. 3. The calibration protocol using the Stroop color-word test is shown in Fig. 4. In the "Introduction" section, the subjects are briefly explained about the Stroop color-word test and get familiar with it; followed by the "Rest" section, which is used to record EEG data when the subjects are in the relaxed state. Then the subjects perform the Stroop test with three different levels and the self-assessment for each level is done at the end of each section as described above. To induce low workload, the word's meaning is the same with the word's font color (Congruent Section). To induce medium workload, the word's meaning is not the same with the word's font color (Incongruent Section 1). To increase workload to a higher level, the subject needs to react to the incongruent word within the limited time (Incongruent Section 2).

After calibration, the EEG data recorded during four tests were used for training of classifiers for each subject. Then, the EEG-based workload monitoring system

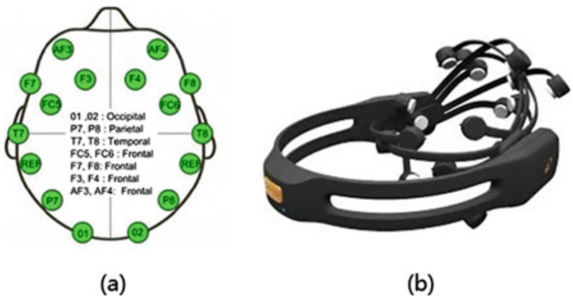

Fig. 2 The Emotiv brain computer interface. **a** The location map of 14 electrodes based on international 10–20 system. **b** Emotiv EPOC device records EEG signal at sampling rate 128 Hz with frequency response between 0.16 and 43 Hz

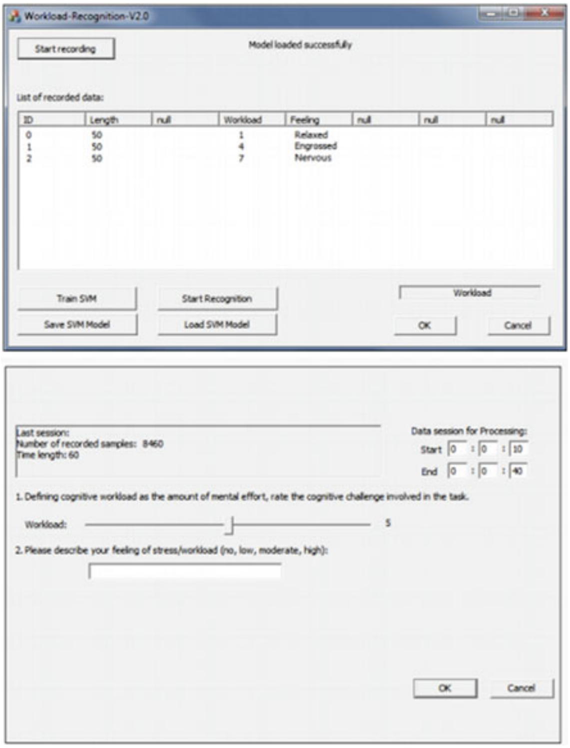

Fig. 3 Screenshots of the mental workload calibration interface and questionnaire interface

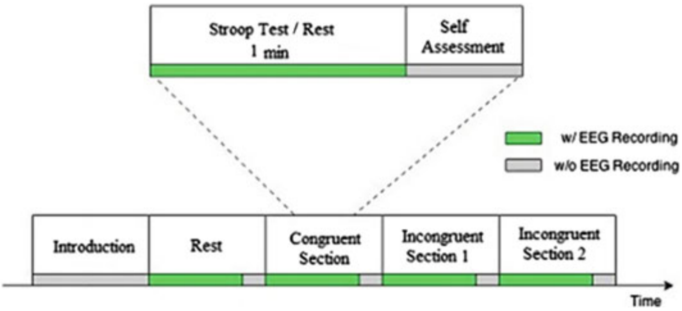

Fig. 4 The calibration protocol for EEG-based mental workload recognition

can recognize subject's mental workload each second. In Fig. 5, on the left, the subject's workload is visualized on the dynamic meter in real time. Besides color representation such as "red" color used for high workload and "green" color used for low workload, there is a word in the center of each meter to describe current workload level. After a completing of monitoring, a workload levels distribution

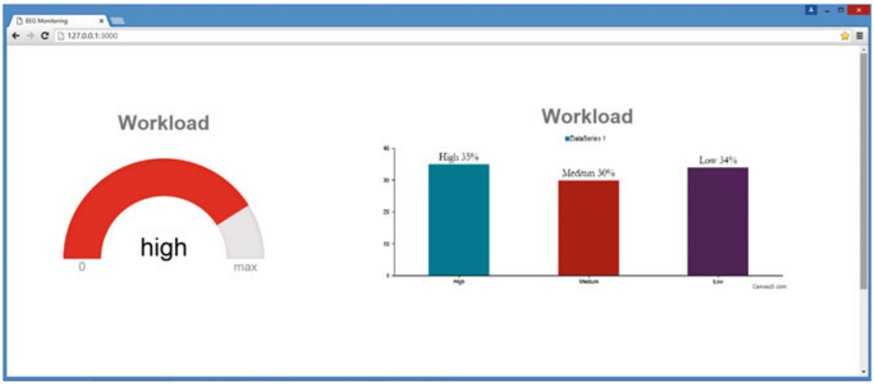

Fig. 5 Screenshot of EEG-based workload monitoring system. *Left meter* shows that current workload level is high. *Right diagram* shows overall distribution of different workload levels during the task performance

diagram can be generated to summarize the overall workload level during the task performance as it is shown in Fig. 5 on the right. This real-time workload monitoring system helps researcher to do more insightful analysis of subject's performance during ATC experiment.

4.3 ATC Simulation

In our current user study, we evaluate the current ATC work place and test interactive touch display that is the input of the development of the future ATC work place. The ATC simulator that used in our experiment is shown in Fig. 6. The middle monitor is used to display the primary radar informaiton. On the right side of the radar display, the monitor shows the Flight Progress Strips (FPS). FPS is an automation tool that provide aircraft updates including the latest altitude clearance, flight route as well as estimated outbound and inbound time for all departing and arriving aircraft, respectively. On the left side, a conflict resolution aid (CRA) display is intergrated to the current work place to support ATCOs in resolving conflict. The CRA is an automation aid that could advise ATCOs on the resolution of a potential conflict about 2 min in advance. In front of ATCOs, there is an interactive touch display to help ATCOs to understand the airspace situation. This display provides ATCOs with the information of aircraft speed profile, climb and descend rate along the time axes. During the experiment, the performance of percentage of resolved conflict and conflict resolution time are measured automatically through the data obtained from the simulator. Upon completion of each experiment scenario, the NASA-TLX questionnaire is used to measure mental workload of ATCOs. The EEG data are recorded throughout the experiment. The results of the

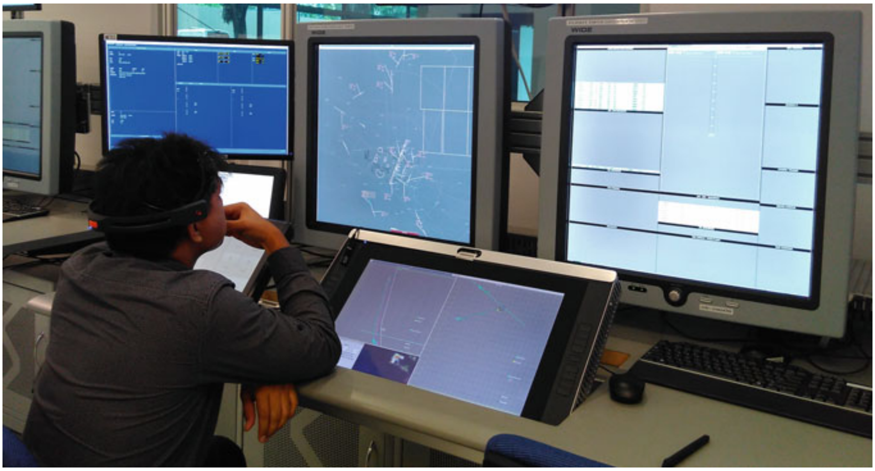

Fig. 6 The air traffice control work place integrated with interactive touch display and brain computer interface

study will drive the refinement and further development of both hardware configuration and software development.

5 Preliminary Results

Currently, we analyze the user study data of the ATC work place for three groups (Non-Display, Display, and Trajectory Prediction) in the three CRA conditions (Manual, Reliable and Unreliable). We studied the relation between the data received using traditional NASA-TLX method and the workload rating method used to label EEG data in the proposed EEG-based system for human factor study. Both methods were administered after each scenario. This analysis allowed for direct comparison between NASA-TLX and the proposed EEG-based evaluation system. Table 1 shows the correlation of workload rating received in 1–9 scale and NASA-TLX workload calculated after completion of each scenario of the experiment. Generally, the two evaluation methods were found to be highly correlated in most of the simulations. Only in unreliable CRA condition of non-display group and trajectory prediction group, the correlation between workload rating and NASA-TLX resulting data was not significant; however, the trend of positive correlation between the two methods' data in this condition could still be observed. These findings confirm that the method used for labeling of the EEG data with workload levels produces labels which are correlated with NASA-TLX workload evaluation data. Furthermore, with the reference to the labeling of the EEG data with workload levels, the EEG-based workload recognition algorithm can be used to calculate the workload levels in real time through all recorded EEG data with

Table 1 Correlation analysis between workload rating and NASA-TLX in the three groups (non-display, display, and trajectory prediction) and three CRA conditions (manual, reliable and unreliable)

	Manual	Reliable	Unreliable
Non-display	$r = 0.847$ $p = 0.001$*	$r = 0.748$ $p = 0.004$*	$r = 0.415$ $p = 0.180$
Display	$r = 0.661$ $p = 0.019$*	$r = 0.590$ $p = 0.043$*	$r = 0.716$ $p = 0.009$*
Trajectory Prediction	$r = 0.669$ $p = 0.017$*	$r = 0.529$ $p = 0.077$**	$r = 0.258$ $p = 0.419$

*Significant at $\alpha = 0.05$; **Significant at $\alpha = 0.1$

high time resolution. Thus, the EEG-based workload evaluation has been proven to validly assess workload and has a strong benefit since it could provide real-time workload data corresponding to the tasks performed throughout the experiment.

6 Conclusion

In this paper, we propose novel EEG-based tools for human factor study in Air Traffic Control (ATC) systems. The proposed system allows for recognition of mental workload, stress, emotions of subjects during the performance of ATC tasks to assess novel automation systems for future ATC. The EEG-based brain state recognition algorithms are implemented using machine learning techniques. We analyzed relation between mental workload calculated using traditional NASA-TLX method and the method used to label EEG data with different workload levels. It was found that the data are highly correlated in most of the simulations. Thus, the EEG-based system can be used to recognize workload during the task performance at any time. By utilizing the proposed EEG-based system, true understanding of ATCOs' working pattern can be obtained. Based on the analyses of the objective real time data together with the subjective feedback from ATCOs, we are able to reliably evaluate current ATC systems and refine new concepts of future ATC system.

Acknowledgments This research was supported by Civil Aviation Authority of Singapore (CAAS) and Air Traffic Management Research Institute (ATMRI) Project ATMRI: 2014-R5-CHEN.

References

1. Sheridan, T.B.: Next generation air transportation systems: human-automation interaction and organizational risks. In: Proceedings of the 2nd Symposium on Resilience Engineering, pp. 272–282 (2006)
2. Intensifying Asia-Pacific Collaboration to Address Efficiency and Safety.: International Civil Aviation Organization (2010)

3. Murphy, E.D., Albert, H.A., Chen, J.M., Anderson, G.G.: The role of mental computations in current and future en route air traffic control. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 56, pp. 110–114. Sage Publications, London (2012)
4. Endsley, M., Sollenberger, R., Stein, E.: The use of predictive displays for aiding controller situation awareness. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 43, pp. 51–55. SAGE Publications, London (1999)
5. Endsley, M.R., Rodgers, M.D.: Distribution of attention, situation awareness, and workload in a passive air traffic control task: implications for operational errors and automation. Federal Aviation Administration Office of Aerospace Medicine (1997)
6. Prevot, T., Homola, J.R., Martin, L.H., Mercer, J.S., Cabrall, C.D.: Toward automated air traffic control—investigating a fundamental paradigm shift in human/systems interaction. Int. J. Human-Comput. Inter. **28**, 77–98 (2012)
7. Trapsilawati, F., Qu, X., Wickens, C.D., Chen, C.-H.: Human factors assessment of conflict resolution aid reliability and time pressure in future air traffic control. Ergonomics 1–12 (2015)
8. Hart, S.G., Staveland, L.E.: Development of a multi-dimensional workload rating scale. Human Mental Workload 139–183 (1988)
9. De Waard, D., Studiecentrum, V.: The measurement of drivers' mental workload. Groningen University, Traffic Research Center Netherlands (1996)
10. Miller, S.: Workload measures. National Advanced Driving Simulator. Iowa City, United States (2001)
11. Hart, S.G., Staveland, L.E.: Development of NASA-TLX (task load index): results of empirical and theoretical research. Adv. Psychol. **52**, 139–183 (1988)
12. Reid, G.B., Nygren, T.E.: The subjective workload assessment technique: a scaling procedure for measuring mental workload. Adv. Psychol. **52**, 185–218 (1988)
13. Cooper, G.E., Harper Jr, R.P.: The use of pilot rating in the evaluation of aircraft handling qualities. DTIC Document (1969)
14. Hill, S.G., Iavecchia, H.P., Byers, J.C., Bittner, A.C., Zaklade, A.L., Christ, R.E.: Comparison of four subjective workload rating scales. Human Factors J Human Factors Ergon. Soc. **34**, 429–439 (1992)
15. Berka, C., Levendowski, D.J., Ramsey, C.K., Davis, G., Lumicao, M.N., Stanney, K., Reeves, L., Regli, S.H., Tremoulet, P.D., Stibler, K.: Evaluation of an EEG workload model in an Aegis simulation environment. Defense and security, pp. 90–99. International Society for Optics and Photonics (2005)
16. Kamzanova, A.T., Kustubayeva, A.M., Matthews, G.: Use of EEG workload indices for diagnostic monitoring of vigilance decrement. Human Factors J. Human Factors Ergon. Soc. **56**, 1136–1149 (2014)
17. Müller, K.-R., Tangermann, M., Dornhege, G., Krauledat, M., Curio, G., Blankertz, B.: Machine learning for real-time single-trial EEG-analysis: from brain–computer interfacing to mental state monitoring. J. Neurosci. Methods **167**, 82–90 (2008)
18. Lei, S., Roetting, M.: Influence of task combination on EEG spectrum modulation for driver workload estimation. Human Factors J. Human Factors Ergon. Soc. **53**, 168–179 (2011)
19. Borghini, G., Astolfi, L., Vecchiato, G., Mattia, D., Babiloni, F.: Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness. Neurosci. Biobehav. Rev. **44**, 58–75 (2014)
20. Hou, X., Liu, Y., Sourina, O., Mueller-Wittig, W., Lim, W.L., Lan, Z., Wang, L.: CogniMeter: EEG-based brain states monitoring. In: LNCS Transactions on Computational Science (2016)
21. Higuchi, T.: Approach to an irregular time series on the basis of the fractal theory. Physica D **31**, 277–283 (1988)
22. Qiang, W., Sourina, O.: Real-time mental arithmetic task recognition from EEG signals. IEEE Trans. Neural Syst. Rehabil. Eng. **21**, 225–232 (2013)
23. Lan, Z., Sourina, O., Wang, L., Liu, Y.: Real-time EEG-based emotion monitoring using stable features. The Vis. Comput. 1–12 (2015)

24. Lim, W.L., Sourina, O., Wang, L., Liu, Y.: EEG-based mental workload recognition related to multitasking. In: 10th International Conference on Information, Communications and Signal Processing (ICICS) (2015)
25. The Emotiv EEG Device.: http://www.emotiv.com/epoc.php