AN EVALUATION FRAMEWORK IN MARITIME EDUCATION AND TRAINING FOR E-LEARNING USER'S ENVIRONMENT SATISFACTION

[1]HU QUIYOU, [2]NIKITAKOS NIKITAS, [3]PAPACHRISTOS DIMITIROS

[1]Merchant Marine College, Shanghai Maritime University
[2,3]Shipping, Trade and Transport Department, University of Aegean

ABSTRACT

This paper presents the development of an evaluation research framework of user's satisfaction based on gaze tracking and voice recording for sentiment analysis and applies this to the evaluation of learning in maritime education and training. The research methodology that will be used has its roots in the Neuroscience field and connects the behavioral evaluation with the gaze track and head movements' data, sentiment and opinion analysis of speech and traditional methods (interview, questionnaire). The research purpose is defined in detecting, recognizing and interpreting the emotional information in conjunction with other information created during the execution of a scenario in a maritime e-learning system (simulators or training software). The proposed method and experiment (ECDIS course in SMU) contributes to the enhancement of evaluation method in adult maritime education and can be taken up by maritime learning systems developers to improve their learning content and process in educational project.

Keywords: *Evaluation, e-learning, gaze tracking, sentiment analysis, neuroscience*

1. INTRODUCTION

In the maritime education and training, the user's satisfaction based on objective criteria poses an important research subject because via this we can determine the background that explains the satisfaction phenomena, recommending at the same time new considerations that will expand the up-to-date educational conclusions on the adult education in educational programs and software development ([1],[2],[3]).

Eye observation on handiness tests is a rather but promising new field especially for the system designers, as it may offer information on what may attract the user's attention and which are the problematic areas during using the system. Also, another factor that can be investigated in relation to the emotional experience (satisfaction) is the sentiment analysis (language processing). The research area on use of the optical recording tools is the quest for an exact interpretation of the optical measurements and voice recording for sentiment/opinion analysis, which is their connection to the satisfaction and the learning effectiveness of the users. Suggested research aims at this direction with the use of neuroscience methods in combination with the use of qualitative-quantitative researches aiming at the extraction of useful conclusion that will help simulator system designers to develop the systems (especially the interface, delivering and organizing education material), the class designers to better organize of the material and modern tools use (better planed educational scenarios that thriftily develop the trainees abilities but also can offer a more objective evaluation of their abilities and their function as future captains or mechanics) and finally the expansion of the adult education field by offering new conclusions regarding the e-learning use (introduction modes, evaluation) and possible revision of the marine education models of the respective apposite organizations (IMO) ([2],[3],[4],[5]).

The phenomenon of the subjective satisfaction of the user is the ultimate goal (Fig.1). It is complicated in its nature as it is affected by many factors varying from situation to situation and from individual to individual. The ultimate goal is to find an average, concerning the maritime education and training and any special factors concerning: the education suitability (maritime educational models), the ability to simulate the functional abilities regarding the actual maritime environment, the realism of situations in correspondence with the real-time situations, the controlled functional environment, the satisfactory interface for the users and the ability to conduct full-time system control by the educator-trainer.

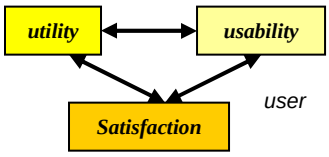

Figure 1 Structural vision of the users-students "subjective satisfaction" phenomenon

We propose a research evaluation framework for Satisfaction Evaluation of e-learning user's environment. The main elements of the proposed approach include ([5],[6]):

- Natural parameters' measurement (head movement, gaze tracking, sentiment/opinion analysis) and
- Registering user opinion/viewpoint/view (statistical analysis).

This procedure is a primary effort to research the educational and usability evaluation with emotion analysis (satisfaction phenomenon) of the users-students

in maritime e-learning environments.

2. LITERATURE REVIEW AND SCOPE

Usability testing procedures used in user-centered interaction design to evaluate a product by testing it on users. This can be seen as an irreplaceable usability practice, since it gives direct input on how real users use the system. Usability testing focuses on measuring a human-made product's capacity to meet its intended purpose ([4],[7]).

Usability has been defined by ISO 9241 as "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use". Effectiveness means accuracy and completeness with which users achieve specified goals. Efficiency means resources expended in relation to the accuracy and completeness with which users achieve goals. Satisfaction means freedom from discomfort and positive attitudes towards the use on the product. It is widely acknowledged that the first two goals of usability (efficiency and effectiveness) can be measured in an objective manner (usability tests), but he third goal (personal satisfaction) is subjective in nature and depends on the characteristics of the user groups addressed [5].

International bibliography provides many sources on the Eye-tracking research in education. For instance, Schiessl et al. (2003) used an eye-tracker to investigate gender interfaces in attention behavior for textual vs. pretorial stimuli on websites. An investigate outcome was that, when the participants were asked where in the interface they thought they looked their perceptions often differed from reality, showing that accurate attention patterns could only be found with an eye-tracker. In Jakob (1998) explores issues surrounding the real-time processing of eye data such as efficient noise reduction and the organization of gaze information into tokens from which relevant data may be extracted. He then discusses the potential of eye-tracking as a tool in several forms of interface manipulation, including object selection/movement, scrolling text and navigating menus. Salvucci and Anderson (2000) applied these ideas to design IGO (Intelligent Gaze-added Operating system), a system that allows users to use their eyes to perform interface operations such as opening, closing and dragging windows. Sibert et al., (2000) describe the use of gaze trucking to assess reading performance in the Reading Assistant, a system for automated reading remediation that provides visual and auditory cues if user gaze patterns indicate difficulties in reading a word. Qu and Johnson (2005) use eye-tracking for interaction adaptation within the Virtual Factory teaching systems (VFTS), an computer tutor for teaching engineering skills. Eye-tracking is used to discern the time the user spends reading something from the time the user spends thinking before taking action, with the goal of assessing and adapting to the motivational states of student effort and confusion. Gluck and Anderson (2001) studied the use of eye-tracking to assess student problem-solving behaviors within the PAT Algebra I tutor, including attention shifts, disambiguation of problem statements and errors, processing of error messages and other information critical to problem solving [8].

Investigating the emotional gravity of words spoken by a speaker and defined its emotional state (current or past) constitutes a state of the art issue. Most of the emotional state categorization suggested concern the English language. To overcome this problem, studies have been conducted that approach the matter cross-culturally and study the assignment of the categories to various languages. This assignment has conceptual traps since the manner in which an emotional state is apprehensible; an emotional state is influenced by cultural factors as well.

International bibliography contains various approaches – techniques (sorting algorithms) concerning linguistic emotional analyses, which are followed and are based mainly in the existence of word lists or dictionaries with labels of emotional gravity along with applications in marketing, cinema, internet, political discourse etc. There are studies also concerning sorting English verbs and French verbs that state emotions based on conceptual and structural-syntactical characteristics ([9],[10],[11]).

The major idea of this paper is propose a general approach for satisfaction evaluation of the students-users' (subjective) satisfaction of the marine education via user interface evaluation of several types of educational software (i.e. simulator). The proposed research is a combination of qualitative – quantitative methodology, on one hand, and a use of neuroscience tools (gaze tracker, sentiment analysis), on the other hand.

3. THE EVALUATION RESEARCH FRAMEWORK

The suggested *E*valuation *REsearch FRamework (EREFR*) aims at interpreting, determining and evaluating the figures of the gaze tracker and voice recording in combination with the conventional methods (qualitative, quantitative) results based on the factors that is possible to influence the user's satisfaction (Fig.2). Therefore it is essential to continuously and thoroughly analyze factors and parameters that contribute in the determination of the user's satisfaction level aiming at the evolution of the interpretation framework into a complete interpretation and evaluation model of the students-users' satisfaction in maritime education and Training.

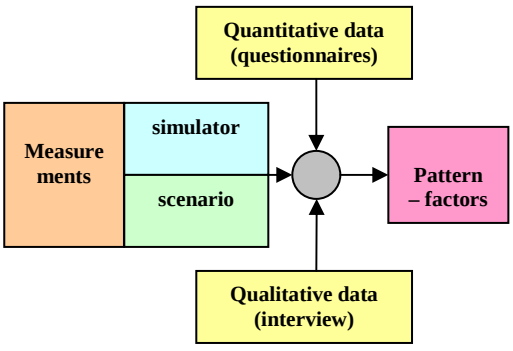

Figure 2 Interpretation procedure of EREFR

The structure of proposed EREFR concerns the following sections ([4],[5]):

(a) Mood/Emotion before the scenario/exercise (oral text)

(b) Behavioral action (head movement, gaze) during the scenario and

(c) Emotional post-experience – satisfaction (oral text).

Measuring the emotional information will be realized using the following processes:

(a) Natural parameters' measurement: Movement parameters (head movement, gaze movement) and oral text as text and

(b) Registering user opinion/viewpoint/view.

The EREFR is comprised by the following parts (Fig. 3) ([4],[5]):

- Influence Part (IP)*:* We suppose based on Action Tendency Theory (concern view) and on Practical Reasoning Theory by M. Bratman that there are possible interactions (influences) in the user's emotional state / satisfaction through intentions-desires-predisposition.
- Emotion Measurement Part (EMP): The emotional analysis theories based on measurement process concern the meta-experience of the emotional state, either during measurement before the experiment (mood before the educational act) or after the experiment (sensation/view/opinion/stance).
- Appraisal Part (AP): In this section, the Satisfaction recording takes place but also commenting related to the day when measurement is taking place and in total up to that moment, as far as the software tool is concerned, after the experimental conduct of the scenario/exercise (usability), personal self-evaluation, scenario evaluation (benefits) in combination with the weighed usability measurement tool (DEC SUS Tool).

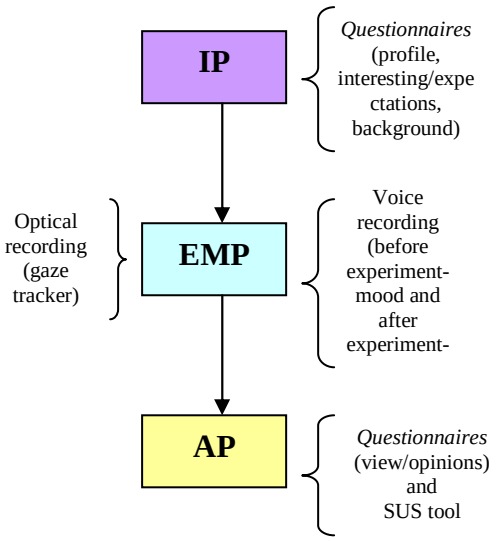

Figure 3 Internal structure of EREFR

Defining satisfaction concerns the following issues, which are being investigated ([2],[3],[4],[5]):

- concerning software and educational scenarios,
- system usability as far as the system per se is concerned (total functionality),

as well as the individual training and technical characteristics that complete the teaching act.

The detection of emotional information (satisfaction) will be realized using the Technical and Theoretical Tools ([4],[5],[12],[13]):

- Tool-I (T-I*):* In the protocol the optical data registration will be conducted by the 'Face Analysis' software that was developed by the IVML Lab of the National Technical University of Athens, in connection with a Web camera set on the computer in which there is the subject of the research (educational software i.e. ECDIS, Engine Simulator). That particular software records a large number of variables but we focus on the following parameters that refer to the user's eyes and head movement: (i) eyes movement: vertical and horizontal eye movements (Eye gaze vector), (ii) user's head position in regard to the eyes up/down – right/left movement (Head Pose Vector: pitch, yaw), (iii) eye distance from the computer screen (Dist_monitor) and (iv) rolling of the head (eye angle from a horizontal level) (Fig.4).

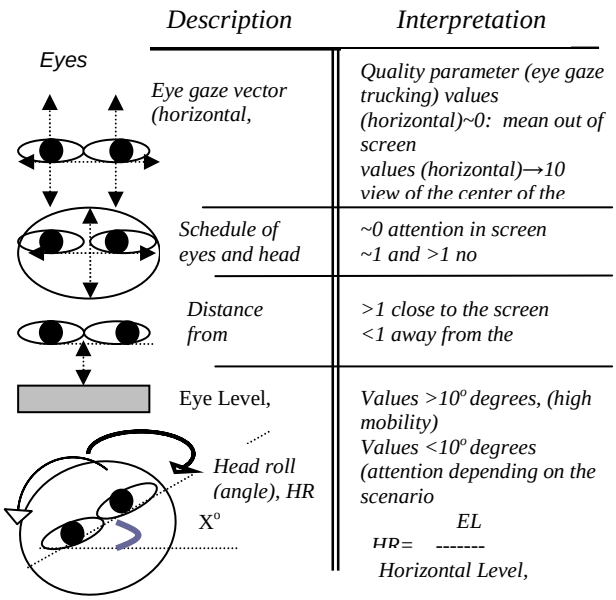

Figure 4 Parameters interpretation of 'Face Analysis' Tool-I

- Tool-II (T-II)*:* Use of a microphone for voice recording of spoken words (speech-text). This will be used for the registration of 3 temporal marks: (i) First Point (T_1) - the temporal mark before recording for measuring mood. This executes the voice recording (1 file) of the user where the user explains how he/she feels and why, (ii) Second Point (T_2) - a temporal mark after the recording for the measurement of mood-emotion after the recording. This executes the voice recording (1 file) of the user where the user explains how he/she feels and why and (iii) Third Point (T_3) - a temporal mark after the recording where the satisfaction choices are justified (software, scenario). The voice recordings

consisting of 3 .WAV type files will be analyzed further during the processing section in three dimensions: Lexicological (emotional analysis), Style analysis of linguistic characteristics, and Qualitative analysis of the spoken word so that the user emotional state/satisfaction can be justified (Fig.5).

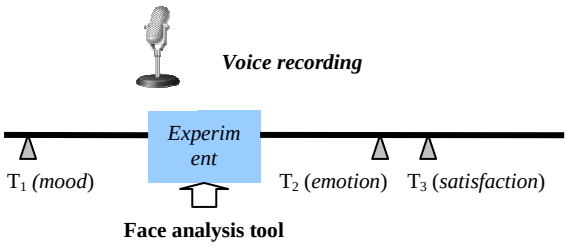

Figure 5 Voice recording

- Tool-III (T-III*):* Questionnaires using for opinion/attitudes/expectation/self-evaluation. It concerns 3 questionnaires: (a) T-III-1, influence data, (b) T-III-2, mood, scenario label, recording data and (c) T-III-3, User appraisal concerning software, scenario, educational environment, errors, total assessment and suggestions (Fig.6).

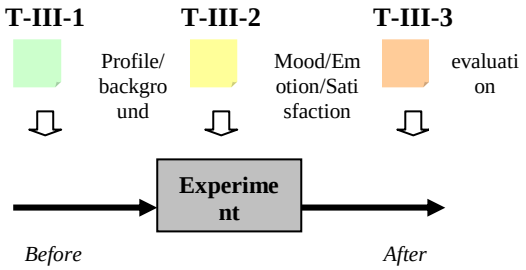

Figure 6 The structure of T-III

- SUS Tool-IV (T-IV): The System Usability Scale (SUS) is a simple, ten-item scale giving a global view of subjective assessments of usability. The SU scale is generally used after the respondent has had an opportunity to use the system being evaluated, but before any debriefing or discussion takes place. Respondents should be asked to record their immediate response to each item, rather than thinking about items for a long time. The SUS scores have a range of 0 to 100 (Table 1).

Table 1. System Usability Scale (DEC-1986)

No	Question	Coef
1	I think that I would like to use this system frequently	4
2	I found the system unnecessarily complex	1
3	I thought the system was easy to use	1
4	I think that I would need the support of a technical person to be able to use this system	4
5	I found the various functions in this system were well integrated	1
6	I thought there was too much inconsistency in this system	2
7	I would imagine that most people would learn to use this system very quickly	1
8	I found the system very cumbersome to use	1
9	I felt very confident using the system	4
10	I needed to learn a lot of things before I could get going with this system	3

The data processing concerns the composition of all the above mentioned relationships so that patterns of the natural parameters in relation to emotional states (Satisfaction) can be found. The composition of elements from questionnaires, SUS, optical and voice recording data had as final goal the support or rejection of hypothesis Research (Fig.7):

Hypothesis-1: There is an assigned relationship between the emotional state to the eye and head movement or meta-emotional experience via spoken words?

Hypothesis-2: There is a relation between the personal data of user (gender, place of work, experience, education, and age) and the user satisfaction by the equipment and the work operation?

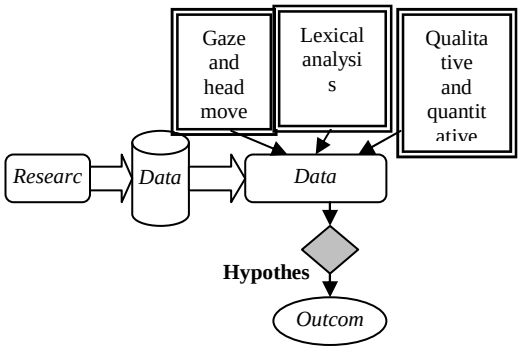

Figure 7 the data processing flow

Finally, the main purpose of the research via EREFR is the investigation of usability-satisfaction of a user of maritime e-learning systems via the assistance of biometric tools, language tools but also other traditional methods (questionnaires, SUS tool).

4. OBJECT

The experimental procedure presented here is a primary effort to research the satisfaction phenomenon of the users-students in e-navigation environments (Electronic Chart Display and Information Systems-ECDIS) by using a combination of qualitative – quantitative techniques with a biometric tool (gaze tracking tool) and sentiment analysis (voice recording)(Fig.8). The 'Goals' ECDIS are set as 'to plan and display the ship's route for the intended voyage and to plot and monitor positions throughout voyage' based on SOLAS V/19.2.1.4 [14].

Especially the case study aims the following:

- the evaluation of the user satisfaction from using the *ECDIS* software and scenario and

- Educational evaluation of *ECDIS* from the user's point of view (opinions).

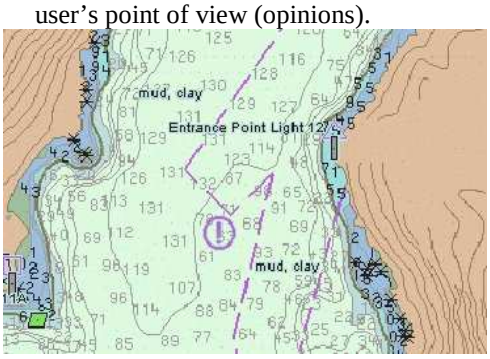

Figure 8 ECDIS image (Electronic navigational chart – NOAA, source: Wikipedia)

5. PARTICIPANTS

The sampling will be done in the course ECDIS in Shanghai Maritime University (SMU) (Fig.9) [15]. This course is provided in the third year. The ECDIS course contains two parts: theory and simulator lab practice. The next table shows the students profile for Age. 29 maritime Students of SMU (ECDIS course) will take part in the experiment (Fig.10).

Figure 9 The Merchant Marine College of SMU (web site)

Table 1. Student profile of age

Sex	Age	
	Mean	*STD*
Male	22.6	21.5
Female	1.4	0.7
Total	22.5	1.38

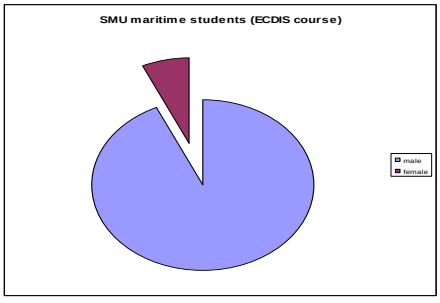

Figure 10 SMU maritime students (ECDIS course) profile of sex (male, female)

6. PROCEDURE

The experiment research process will include the following levels (Fig.11):

- *First Level 1st*: Information about the experiment, Presentation of the acceptance document by the user-student (estimated time duration 5 - 10 minutes).
- *Second Level 2nd*: Completion of a user's profile and of the assessment survey concerning educational and technical characteristics (questionnaire, T-III-1) and voice recording T-II and T-III-2 for mood recording by the student (estimated time period 10 - 15 minutes).
- *Third Level 3rd*: Equipment installation ('Face Analysis' T-I) and configuring the parameters.
- *Four Level 4th*: Optical recording (T-I) (estimated time duration 20-25 minutes).
- *Five Level 5th*: Completion of the process (T-I disconnect) through a structured interview (T-II and T-III-2) and questionnaire (T-III-3 and T-IV) with the user (estimated at 5 – 10 minutes).

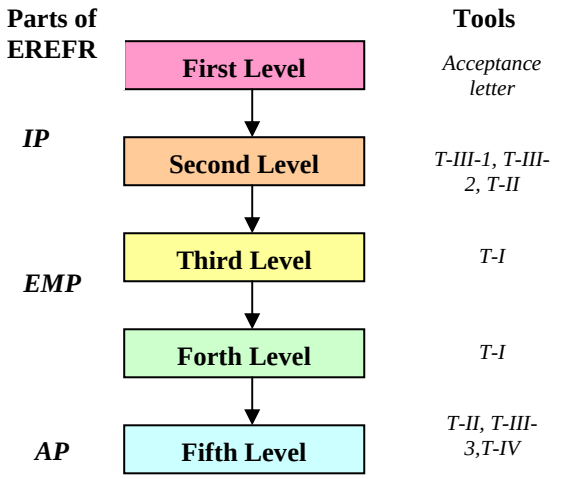

Figure 11 The levels of proposed research procedure

7. DISCUSSION

In Marine Education (ME), the use of neuroscience methods and tools (biometric tools) is a useful contribution in its amelioration. ME follows certain education standards (STCW'95) for each specialty (Captain, Engineer) and for each level (A', B', C'). Its scope is the acquisition of basic scientific knowledge, dexterities on execution (navigation, route plotting, administering the engine etc.) as well as protecting the ship and crew (safety issues and environment protection issues) [16].

The main purpose of the research via EREFR is the investigation of satisfaction of a user of maritime education equipment (engine-navigation simulators, ECDIS) via the assistance of biometric tools, language tools for sentiment/opinion analysis but also other traditional methods (questionnaires-interviews).

The suggested research framework (EREFR) focuses in the following dimensions of research work: (i)

investigation of basic emotional state (happiness-sad) of a user of a marine system of electronic learning as far as the dipole happiness-sad through interaction, (ii) investigation of the emotional state connection to *satisfaction* as far as the educational use of such systems is concerned, of training programs of scenario-exercises, (iii) investigation of possible detection means for emotional states – physiological parameters in electronic learning environments (visual and language recording), and (iv) interactions of all of the above with relation to Maritime education and training (standards, etc.) (16],[17]).

In particular, the biometric tools and the neural sciences methods will help expand the adult education field and especially marine education opening new broad ways to the learning and tutoring sectors within the new educational framework established by the use on new technologies.

8. REFERENCES

[1] IMO, *Issues for training seafarers resulting from the implementation on board technology*, STW 34/INF, 2003.

[2] PAPACHRISTOS, D., NIKITAKOS, N., *Application Methods and Tools of Neuroscience, in Marine Education,* Conference Proceedings "Marine Education and Marine Technology" (ELINT), Athens, Greece, 2010.

[3] PAPACHRISTOS, D., NIKITAKOS, N., *Evaluation of Educational Software for Marine Training with the Aid of Neuroscience Methods and Tools*, Symposium Proceedings, TransNav2011 Symposium, Gdynia Maritime University, Poland, 15-17 June, 2011.

[4] DIX, A., FINLAY, J., ABOWD, G., BEALE, R., *Human-Computer Interaction*, UK, Pearson Education Limited, 2004.

[5] PAPACHRISTOS, D., KOUTSABASIS, P., NIKITAKOS, N., *Usability Evaluation at the Ship's Bridge: A Multi-Method Approach*, in Proceedings of The 4th International Symposium on "Ship Operation, Management and Economics - SOME12" SNAME, Greek Section, Athens, Greece, Dec 8-9, 2012.

[6] PAPACHRISTOS, D., NIKITAKOS, N., ALAFODIMOS, C., *Emotion Evaluation of Simulation System in Educational Practice*, In Proceedings of The International Conference on E-Learning in the Workplace 2012 (ICELW2012), www.icelw.org, New York, NY, USA, June 13th-15th, 2012.

[7] NIELSEN, J., *Usability Engineering*, Academic Press Inc, 1994.

[8] CONATI, C., MERTEN, C., *Eye-tracking for user modeling in exploratory learning environments: An empirical evaluation*, Knowledge-Based Systems 2007, 20:557-74.

[9] LAMBOV, D., PAIS, S., DIAS, G., *Merged Agreement Algorithms for Domain Independent Sentiment Analysis, Pacific Association*, For Computational Linguistics (PACLING 2011), Procedia - Socila and Behavioural Sciences, 27, pp 248-257, 2011.

[10] LI, N., WU, D.D., *Using text mining and sentiment analysis for online forums hotspot detection and forecast*, Decision Support Systems, 48, pp 354-368, 2010.

[11] PANG, B., LEE, L., *Opinion Mining and Sentiment Analysis*, Foundations and Trends in Information Retrieval, Vol. 2, pp 1-135, DOI:10.1561/1500000001, 2008.

[12] ASTERIADIS, S., TZOUVELI, P., KARPOUZIS, P., KOLLIAS, S., *Estimation of behavioral user state based on eye gaze and head pose-application in a e-learning environment*, Multimedia Tools and Applications, Springer, Vol. 41, No3, pp 469-493, 2009.

[13] BROOK, J., *SUS, - A quick and dirty usability scale", Usability Net*, EU project, 2003, Available online:http://www.usabilitynet.org/tools/r_questionnaire.htm, (last access 5 April 2013).

[14] IMO, *Development of an e-Navigation strategy implementation plan*, NAV 58/INF.13, 27 April 2012.

[15] SMU, Merchant Marine College, official web site http://mmc.shmtu.edu.cn/index.php?lang=en, (last access 14 April 2013).

[16] IMO, *Role of the Human Element Report on investigation into near misses*, MSC 74/15/1, 2001.

[17] TSOUKALAS, V., PAPACHRISTOS, D., MATTHEU, E., TSOUMAS, N., *Marine Engineers' Training: Educational Assessment of Engine Room Simulators*, WMU Journal of Maritime Affairs, Vol.7, No2, 2008, pp. 429-448.