Scientific Journals
of the Maritime University of Szczecin

Zeszyty Naukowe
Akademii Morskiej w Szczecinie

2016, 46 (118), 95–100
ISSN 1733-8670 (Printed)
ISSN 2392-0378 (Online)
DOI: 10.17402/124

Received: 31.08.2015
Accepted: 18.02.2016
Published: 27.06.2016

Application of Markov chains to a navigator visual attention model

Bartosz Muczyński, **Maciej Gucma**, **Mateusz Bilewski**

Maritime University of Szczecin, Marine Traffic Engineering Centre
1–2 Wały Chrobrego St., 70-500 Szczecin, Poland
e-mails: {b.muczynski; m.gucma; m.bilewski}@am.szczecin.pl
corresponding author

Key words: eye tracking, usability studies, stochastic model, marine simulator, Markov chains, navigation

Abstract

This paper presents initial results from a series of eye-tracking experiments on a Full Mission Bridge simulator. The aim of this research was to develop a stochastic model of a navigator's attention distribution during their navigational watch. Such model could be used as a tool for workload and usability studies for navigators and navigational equipment interfaces. A structure of the model is discussed together with the evaluation of Markov chains as a main modelling tool. Initial results are presented and discussed. It is suggested that 1st order Markov chains are not fully applicable for this problem. A combination of the 1st and higher-order Markov chains will be applied in the next stage of research.

Introduction

Human factors are said to be the main reason for over 90% of all ship collisions (Zhengjiang & Zhaolin, 2003). Although the frequently-occurring unsafe actions in collisions have been roughly identified, very little is known about the human element and the reasons why those mistakes occur repeatedly. Many studies emphasize the critical role of fatigue and lack of sleep as well as an inadequate lookout and a lack of experience and knowledge (Hetherington, Flin & Mearns, 2006). At the same time, very few studies take into consideration human cognitive processing and marine interfaces design.

A study that analysed 177 maritime accidents reports from accidents occurring from 1987–2000 reported that 71% of all human errors on ships are situation-awareness related problems (Hetherington, Flin & Mearns, 2006). Some researchers emphasize the role that cognitive load plays in situations of both monitoring and collision avoidance (Glandrup, 2013).

On the 14th of July, 2015, the International Maritime Organization (IMO) adopted the "Guideline on Software Quality Assurance and Human-Centred Design for e-navigation" that was developed by the Sub-Committee on Navigation, Communications, and Search and Rescue (NCSR) and accepted at its second session in March 2015. With this guideline, the IMO recognized the importance of workload on the navigator and on the design of navigational interfaces, which very often serve as the information source. The guideline states that the *"Usability Testing (UT) is a key component of Human Centred Design (HCD) and uses methods that rely on including users to test the ability of systems to support user needs. UT helps to identify potential problems and solutions during design and development stages by using an iterative approach to testing where the design evolves through rounds of prototyping, testing, analysing, refining and testing again."* (IMO MSC, 2015).

Although it recognizes the need for further studies in the area of usability and human-centred design, it does not provide a framework for such testing and it recognized neither the bridge environment as a unique set of interfaces and information sources nor the watch keeping as a well-defined process based in the bridge environment and conducted by the navigator. Authors hold that it is crucial to describe all the components of the watch-keeping process properly and in detail, using advanced techniques developed and used in the fields of interface design and cognitive psychology. The most basic activity for each navigator is the process of data acquisition, and this paper presents the initial model of a navigator's visual attention based on the eye-tracking data and Markov-Chain models.

Eye tracking data

Eye-tracking, as the name suggests, focuses on tracking the position and movement of an eye. In general, we can distinguish two types of eye movement monitoring techniques: those that measure the position of the eye relative to the head; and those that measure the orientation of the eye in space (or *Point of Regard*, POR) (Duchowski, 2007). For human factor studies, it is very important to understand the connection between eye movement and the visual scene, which is necessary to measure where, and for how long, the subject was focusing his/her attention, and how often certain areas in visual field were revisited, etc. Such an approach is widely used in usability studies (Jacob & Karn, 2003), interface design (Goldberg & Kotval, 1999), ergonomic evaluation of a workspace, and in many other fields where subjects need to acquire information from specific areas in the visual field. Two main measures that are used in this field are called fixations and saccades.

A fixation is one of the most basic events related to movement of the eye and it occurs when the eye remains still over a period of time (i.e., it is fixating on a specific point in the visual field). During a fixation, three distinct types of eye movements occur: tremor, microsaccades, and drifts, but those are mainly used in studies of human neurology and have not yet found any application in human factor research. Fixation itself, as an event during which visual information is acquired, is strongly connected to cognitive processing (Holmqvist et al., 2011). Thus, the distribution of fixations in space: shows the main sources of navigation information for an officer; allows for identification of the main distractors, both on the bridge and in the manoeuvring area; helps to understand how the navigational and the hydro-meteorological situations influence the behaviour of an officer; and shows differences in the decision-making process between experienced and inexperienced crew (Muczyński, Gucma & Gucma, 2013).

The duration of fixations is directly related to mental workload. Subjects tend to fixate longer on the areas that are critical for a given task but also when the visual information is more complex or requires additional mental tasks (e.g. calculations). Also, experienced subjects show shorter fixations in the same task, compared to novices. Some researchers point out that shorter fixations can also indicate high mental workload, due to the stress level and the complexity of the task (Holmqvist et al., 2011).

A saccade is a rapid motion of an eye between one fixation and another one. It is the fastest movement that the body can produce and it is assumed that visual information is not acquired during this movement (Holmqvist et al., 2011). Since a saccade takes place between two fixations, the number and proportions of both events are strictly connected. Saccadic measures are widely used, mainly in studies with a static stimulus.

A visualization of saccades and fixations on a stimuli picture creates the so called scanpath, which helps to identify information-seeking patterns and is very useful for the initial inspection of data. Observing a dynamic scanpath from a recording with a mobile eye-tracker allows for quick evaluation of an officer and his performance by, for example, showing when exactly and based on which information a risk of collision situation has been identified properly.

Markov chains

Markov chains have already been used in the field of eye tracking. So far, this method has been mainly used for fixation clustering and for modelling human behaviour through identification of visual scene properties like contrast, shape, colour, etc. (Bagci et al., 2004; Kimura et al., 2008). A prime example of such research is a work by Kimura et al. (2008). They proposed a stochastic model of human visual attention based on a Bayesian network with four layers:

1. A saliency map that shows the average saliency response at each position of a video frame;
2. A stochastic saliency map that converts the saliency map into a natural human response through a state-space model;
3. An eye movement pattern that predicts the human viewing pattern using a hidden Markov model (HMM); and

4. An eye position density map that estimates the probable human attended region.

Such models have already been proved useful in many general visual recognition tasks yet what authors suggest is a need for a model that is dedicated to a very limited set of tasks. It is related to a common and natural process of specialization that occurs when a subject is in a set working environment. When that happens, visual categories related to the graphical features of the scene become less significant. A subject is trained to pay attention to features and object that are relevant for the given task. This way, semantic or cognitive categories become increasingly significant for the visual attention distribution.

When considering a navigator, two separate areas can be distinguished:

1. An outside area, where the navigator performs classic search tasks, identifying objects by their visual features, which is mostly shape, colour (e.g. lights at night), and movement;
2. A bridge area, where the navigator has access to all navigational data available in multiple different forms: graphical, vector, gauges, text, voice, and numbers.

Information gathered from the second area has the highest impact on the decision-making process when considering safety of navigation and safety manoeuvres. At the same time, it is not possible to consider those two areas independently. When, for any reason, the navigator focuses attention on a given ship, care is taken to identify this ship on the radar and the ECDIS/AIS. It is required to establish all relevant parameters that are necessary for the decision process. There is also an opposite relation that when a dangerous target is identified on the radar or on the ECDIS, it is important to locate this target in the outside area. This process itself disrupts the visual attention model that is based purely on the image characteristics.

Authors propose a model that will include two levels describing distribution attention. The first level is related to available sources of information and the second on the fixation distribution in the area of a given source.

For the first level, it is required to define so-called areas of interest (AOI). Each area has to be specified by its border and its relevance for the specific navigation task (e.g. voyage monitoring or collision avoidance). The second level is concerned with a given AOI and its cognitive category; i.e., what type and what form of information is available in this area. It is also relevant to consider the description of the complexity of a given area. This will have an application to software interfaces where specific functions and information require direct interaction by the navigator.

Analysing the data from the previous research study (Muczyński, Gucma & Gucma, 2013), authors made the assumption that the visual attention process, or more specifically, fixation distribution in the visual scene, during the navigation task can be described as a stochastic process. First-order Markov Chains were chosen as the most appropriate description of this process. A state of the process is given by the location of i-th fixation, as indicated by the AIO in which the fixation appeared. According to the Markov process definition, the probability that a navigator's attention will be focussed in a given AOI in the i-th step, depends only on the location of the fixation in step i–1. That means that during the decision making process, a selection of the next source of information is dependent only on the present one.

This assumption is a simplified one and stands in opposition with working memory theory. Working memory itself is responsible for the transient holding and processing of new and already stored information and is connected with reasoning, comprehension, learning, and memory updating. In practice, it is to be expected that the navigator holds a certain amount of information in the working memory area. Thus, certain information sources do not need to be revisited for a time period that depends on the capacity of the working memory and a complexity of a given task (which directly influence the cognitive workload). At the moment, no research has been done to establish the capacity of a navigator's working memory and hence it is not possible to include this variable in the presented model. This assumption considers low cognitive processing capabilities, related directly to low capacity of a working memory.

To construct a complete Markov chain, an initial state vector and a transition matrix is required. The transition matrix defines the probabilities of the system changing state from one to another (1).

In this case, the transition matrix describes the probability of choosing the next information source; i.e., the probability of changing from one AOI to another. Such a matrix can be directly calculated for each subject by taking the number of two-consecutive fixations between each pair of AOIs and dividing it by a total number of unique AOIs in a given row or column.

$$\mathbf{T} = \begin{bmatrix} P(\mathrm{AOI}_{11}) & P(\mathrm{AOI}_{12}) & \cdots & P(\mathrm{AOI}_{1(n-1)}) & P(\mathrm{AOI}_{1n}) \\ P(\mathrm{AOI}_{21}) & P(\mathrm{AOI}_{22}) & \cdots & P(\mathrm{AOI}_{2(n-1)}) & P(\mathrm{AOI}_{2n}) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ P(\mathrm{AOI}_{(m-1)1}) & P(\mathrm{AOI}_{(m-1)2}) & \cdots & P(\mathrm{AOI}_{(m-1)(n-1)}) & P(\mathrm{AOI}_{(m-1)n}) \\ P(\mathrm{AOI}_{m1}) & P(\mathrm{AOI}_{m2}) & \cdots & P(\mathrm{AOI}_{m(n-1)}) & P(\mathrm{AOI}_{mn}) \end{bmatrix} \tag{1}$$

In a standard experiment procedure, an initial state vector is based on locations of the first fixations taken from all participants. In the simulator, setting such approaches is not practicable – each participant has to be briefed about all available equipment and navigational situations. After familiarization, a short amount of time is required to switch the eye tracker on and to start the simulation. This completely blurs the meaning of a first fixation. For this reason, the initial state vector is calculated by averaging the first 10 fixations from all participants.

By multiplying the initial state vector $\mathbf{S_0}$ and the transition matrix **P**, the next state of the process can be calculated (S_1 S_2 S_3 …). For example, having the transition matrix **P** and the initial state vector $\mathbf{S_0}$, future states of the process can be calculated using vector-matrix multiplication.

Analysing data from the first experiment showed that the resulting Markov chain is an example of a stationary process or so called time-homogeneous Markov chain; i.e., after a finite number of iterations, the state of the process does not change. In a stationary Markov chain, an initial state is irrelevant since a stationary state vector depends only on the values of the transition matrix. This is given as:

$$P\left(X_{n+1} \le y \middle| X_0, X_1, X_2, ..., X_n\right) = P\left(X_{n+1} \le y \middle| X_n\right) \tag{2}$$

At the second level, the model is concerned with the navigator's attention on a single interface and mainly the fixation characteristics of duration and quantity. This will allow modelling of not only the sources of information but also the complexity of a given interface. Such a model could also be used as a baseline for measuring cognitive workload during different navigational scenarios.

Results

Initial calculations of the transition matrices showed a very strong effect of fixations repeated in a given AOI (Figure 1), which is considered natural since a single glance with a 300 ms duration would not be sufficient to acquire significant data from any navigational equipment. This effect hampered the initial model and lead to a number of repeated fixations that did not correspond with the observed data. Two solutions were considered. The first assumed clustering of all fixations repeated in the given AOI (Figure 2).

	AOI	weight
1	Ship_B	34
2	Controls	3
3	Conning	3
4	Controls	19
5	Conning	50
6	Ship_B	12
7	Conning	6
8	Ship_B	33
9	Radar	47
10	Ship_B	9
11	Controls	86
12	Conning	25
13	Ship_B	32

Figure 2. Grouping of repeated fixations in a given AOI

$$\begin{array}{r c} & \begin{array}{cccccc} \text{Birds\_eye} & \text{Conning} & \text{Controls} & \text{Radar} & \text{Ship\_B} & \text{Ship\_C} \end{array} \\ \begin{array}{r} \text{Birds\_eye} \\ \text{Conning} \\ \text{Controls} \\ \text{Radar} \\ \text{Ship\_B} \\ \text{Ship\_C} \end{array} & \begin{bmatrix} \mathbf{0.9806949} & 0.0015444 & 0.0015444 & 0.00154440 & 0.012355 & 0.00231660 \\ 0.0016741 & \mathbf{0.9709821} & 0.0089285 & 0.00000000 & 0.017857 & 0.00055803 \\ 0.0061162 & 0.0366972 & \mathbf{0.9311926} & 0.00000000 & 0.025993 & 0.00000000 \\ 0.0044052 & 0.0044052 & 0.0000000 & \mathbf{0.97797356} & 0.013215 & 0.00000000 \\ 0.0041691 & 0.0068493 & 0.0071471 & 0.00059559 & \mathbf{0.973198} & 0.00804050 \\ 0.0107913 & 0.0071942 & 0.0143884 & 0.00359712 & 0.075539 & \mathbf{0.8884892} \end{bmatrix} \end{array}$$

Figure 1. Transition matrix

$$
\begin{array}{r c}
 & \begin{array}{cccccc} \text{Birds\_eye} & \text{Conning} & \text{Controls} & \text{Radar} & \text{Ship\_B} & \text{Ship\_C} \end{array} \\
\begin{array}{r} \text{Birds\_eye} \\ \text{Conning} \\ \text{Controls} \\ \text{Radar} \\ \text{Ship\_B} \\ \text{Ship\_C} \end{array} &
\begin{bmatrix}
0.00000000 & 0.08000000 & 0.0800000 & 0.08000000 & 0.6400000 & 0.12000000 \\
0.05769231 & 0.00000000 & 0.3076923 & 0.00000000 & 0.6153846 & 0.01923077 \\
0.08888888 & 0.53333333 & 0.0000000 & 0.00000000 & 0.3777778 & 0.00000000 \\
0.20000000 & 0.20000000 & 0.0000000 & 0.00000000 & 0.6000000 & 0.00000000 \\
0.15555556 & 0.25555556 & 0.2666667 & 0.22222222 & 0.0000000 & 0.30000000 \\
0.09677419 & 0.06451613 & 0.1290323 & 0.03225806 & 0.6774194 & 0.00000000
\end{bmatrix}
\end{array}
$$

Figure 3. Transition matrix calculated from grouped fixation data

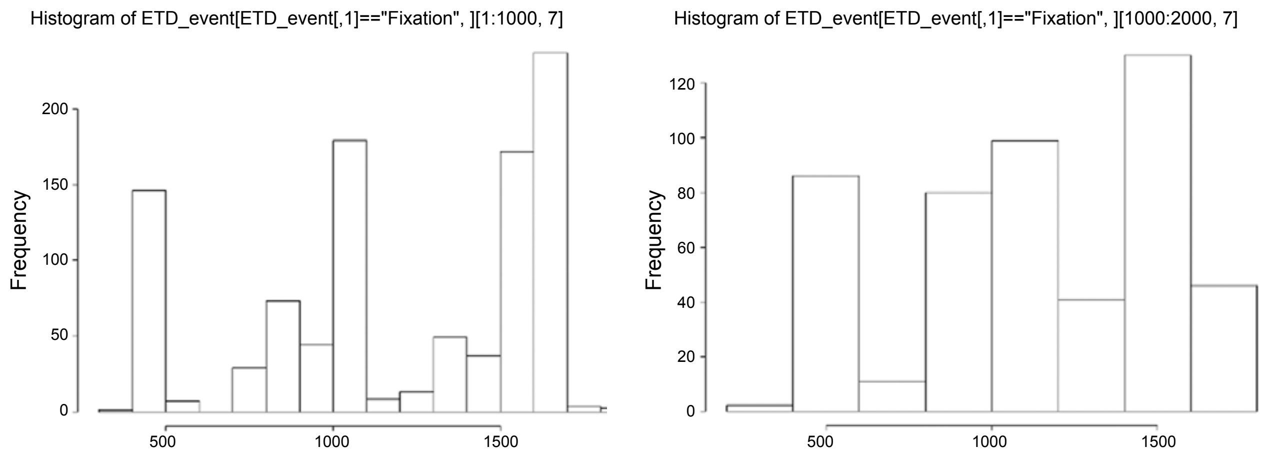

Figure 4. Examples of the fixation distributions on two different interfaces: ECDIS (on the left) and Radar (on the right). The vertical axis shows the number of fixations and the horizontal – fixations duration in microseconds

The transition matrix calculated from such data (Figure 3) shows a higher dispersion of transition probabilities but at the same time the average number of repeated fixations per AOI was not modelled properly.

The second solution assumes dynamically-calculated probabilities of transition. The probabilities depend on the number of consecutive fixations in the given AOI. Such a model includes Markov chains of higher order and is currently being developed.

To model the fixation characteristics, it is important to identify the type of distribution that is related to a given AOI/Interface. Initial results show a large variability (Figure 4) and the distribution has not been yet specified to a degree where it could be implemented in the model.

Conclusions

Markov chain models show promising results and proved useful in modelling the most basic aspects of the visual attention distribution. Such an approach can be used to develop a model that is based not on the visual characteristics of the perceived scene but on the cognitive categories. This opens up the possibility for a model that will be dedicated for a particular setting and thus could be used to provide a baseline for both workload measurements and usability testing.

Acknowledgments

This article is part of a study in the RepSail project and has been co founded by EraNET in decision: ENT III/Future_Traveling/2/2014.

References

1. Bagci, A.M., Ansari, R., Khokhar, A. & Cetin, E. (2004) *Eye tracking using Markov models*. in Pattern Recognition. ICPR 2004. Proceedings of the 17th International Conference, vol. 3, pp. 818–821, August 23–26, 2004.
2. Duchowski, A. (2007) *Eye Tracking Methodology*. London: Springer.
3. Glandrup, M. (2013) *Improving Situation Awareness in the Maritime Domain, in Situation Awareness with Systems of Systems*. New York: Springer Science+Business Media.
4. Goldberg, J.H. & Kotval, X.P. (1999) Computer interface evaluation using eye movements: methods and constructs. *International Journal of Industrial Ergonomics* 24, 6. pp. 631–645.
5. Hetherington, C., Flin, R. & Mearns, K. (2006) Safety in shipping: The human element. *Journal of Safety Research* 37, 4. pp. 401–411.

6. Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H. & van de Weijer, J. (2011) *Eye tracking: A comprehensive guide to methods and measures*. Oxford: Oxford University Press.
7. IMO MSC (2015) *Guideline on Software Quality Assurance and Human-Centred Design for e-navigation*. London.
8. Jacob, R.J.K. & Karn, K.S. (2003) Eye tracking in human-computer interaction and usability research: Ready to deliver the promises (Section Commentary). In J. Hyona, R. Radach & H. Deubel (Eds.), *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research.* pp. 573–605. Amsterdam: Elsevier Science.
9. Kimura, A., Pang, D., Takeuchi, T., Yamato, J. & Kashino, K. (2008) *Dynamic Markov random fields for stochastic modelling of visual attention.* Pattern Recognition. ICPR 2008. 19th International Conference, pp. 1–5, 8–11 Dec. 2008, Tampa, FL.
10. Muczyński, B., Gucma, M. & Gucma, L. (2013) Method of officer of the watch performance analysis by gaze data and eye metrics with use of marine ship's simulator researches. *European Journal of Navigation* 11, 3. pp. 11–17.
11. Pang, D., Kimura, A., Takeuchi, T., Yamato, J. & Kashino, K. (2008) *A stochastic model of selective visual attention with a dynamic Bayesian network.* Multimedia and Expo, IEEE International Conference, pp. 1073–1076, June 23–26, 2008, Hannover.
12. Zhengjiang, L. & Zhaolin, W. (2003) *The Human Elements in Ship Collisions at Sea*. Asia Navigation Conference, 4th September, Kobe, Japan.