Observing Human Behaviour to Identify Risk in Task Performance

Ronald Poppe and Mark ter Maat

Information and Computing Sciences, University of Utrecht / Human Media Interaction, University of Twente
Princetonplein 5, 3584 CC Utrecht, r.w.poppe@uu.nl / Drienerlolaan 5, 7500 AE Enschede

Introduction

Small mistakes in task performance can have large consequences. Administering a wrong dosage of a medicine, ignoring a system warning or failing to check whether certain values are set can have severe negative effects for patient and process safety. The repetitive nature of such tasks and the often high cognitive load of the worker can easily lead to decreased attention, with inaccurate or incorrect task performance as a result. These effects are even stronger when the normal workflow is interrupted, for example when the worker receives a phone call or when other unexpected events occur. These deviations from a typical task performance can be indicative of risky task performance. By observing the behaviour of a worker, we aim at identifying these moments in time and to alert the worker. We specifically focus on training scenarios in which we can provide understandable feedback to the worker, so the risky behaviour can be avoided in the future. In this paper, we present the Patria system, which observes gestures, head movement, facial expressions, eye gaze and key strokes from subjects operating a simple medical device.

User Observation

The Patria system's recording setup consists of several visible-light and infra-red cameras, infra-red lights, a 19" desktop monitor and computers to process and log all input and to run the simulation software (or any other software that the worker needs to work with). Figure 1 shows an overview of the setup.

Two visible-light cameras below the screen are used to analyse facial expressions and head orientation, whereas the two infra-red cameras are employed for gaze estimation. The two pairs of cameras are calibrated and synchronized to allow for stereo-images to be recorded. Especially for the estimation of gaze, the distance from the camera to the eye is of great importance for robust analysis. The two cameras above the screen (see Figure 1) are used to analyse the hand gestures of the worker. The placement of the cameras is such that the worker can freely move around in a relatively large space.

Figure 1. Overview of the Patria system.

For the processing of these image streams, several software programs are employed. For the gaze estimation, we rely on the ITU gaze tracker (San Agustin et al., 2010), which proved to be reasonably robust in the detection of whether a person looks at the screen or not. We have not attempted more accurate detection and consider this a viable next step in the further development of the setup.

The analysis of facial expressions and 6 degrees of freedom head pose (location and rotation) is performed with VicarVision FaceReader 5.0 (Den Uyl & Van Kuilenburg, 2005). Gestures are analysed by fitting a parametric body model to the stereo-images obtained from the top cameras. Finally, all inputs, including videos and estimated behaviour parameters, are time-stamped and logged in Noldus Observer XT 11.0 (Zimmerman et al., 2009), and can be replayed and further analysed when desired.

Proceedings of Measuring Behavior 2014, (Wageningen, The Netherlands, August 27-29, 2014).
Editors: A.J. Spink, L.W.S. Loijens, M. Woloszynowska-Fraser & L.P.J.J. Noldus. www.measuringbehavior.org

Workflow Monitoring and Risk Assessment

One important aspect of the Patria system is its ability to estimate how risky the observed behaviour of the worker is. For this, we employ two different mechanisms:

Work-flow monitoring: for the analysis of risk, we take into account the notion of a correctly performed task. Many workflows, notably those that can have severe consequences for safety, have strict regulations to minimise these risks. Deviating from this normal task performance is considered risky behaviour. To this end, we model a task flow as a state diagram. Each state in the diagram corresponds to a state in the workflow. Transitions between states are triggered by detected behaviour events such as looking at the screen, pressing "enter" or picking up a form with the left hand. Transitions can also occur based on a timer. For example, if a certain action does not occur within 5 seconds after arriving in a certain state. Each transition has an associated risk value assigned to it, between 0 and 100%. As both the states and the actions triggering the transitions between them are interpretable, we can present specific information to the worker about why the risk value was increased. This can help in understanding and improving the task performance, both in training situations as well as in online use.

User state monitoring: based on the observed facial expressions, described as action unit activation scores, we try to find out whether or not they correlate with these moments in time where risky behaviour occurs. Intuitively, one might expect that a stressed, bored or frustrated person is more likely to make mistakes. Also, we expected that people would show observable facial reactions to unexpected events, which are also more likely to result in abnormal task performance. To estimate the level of risk from a worker's facial expressions, we have trained a support vector machine (SVM) classifier based on pre-recorded data in which we manually identified risky behaviour. The SVM classifier is then evaluated on a frame-by-frame basis to arrive at a user state estimate over time.

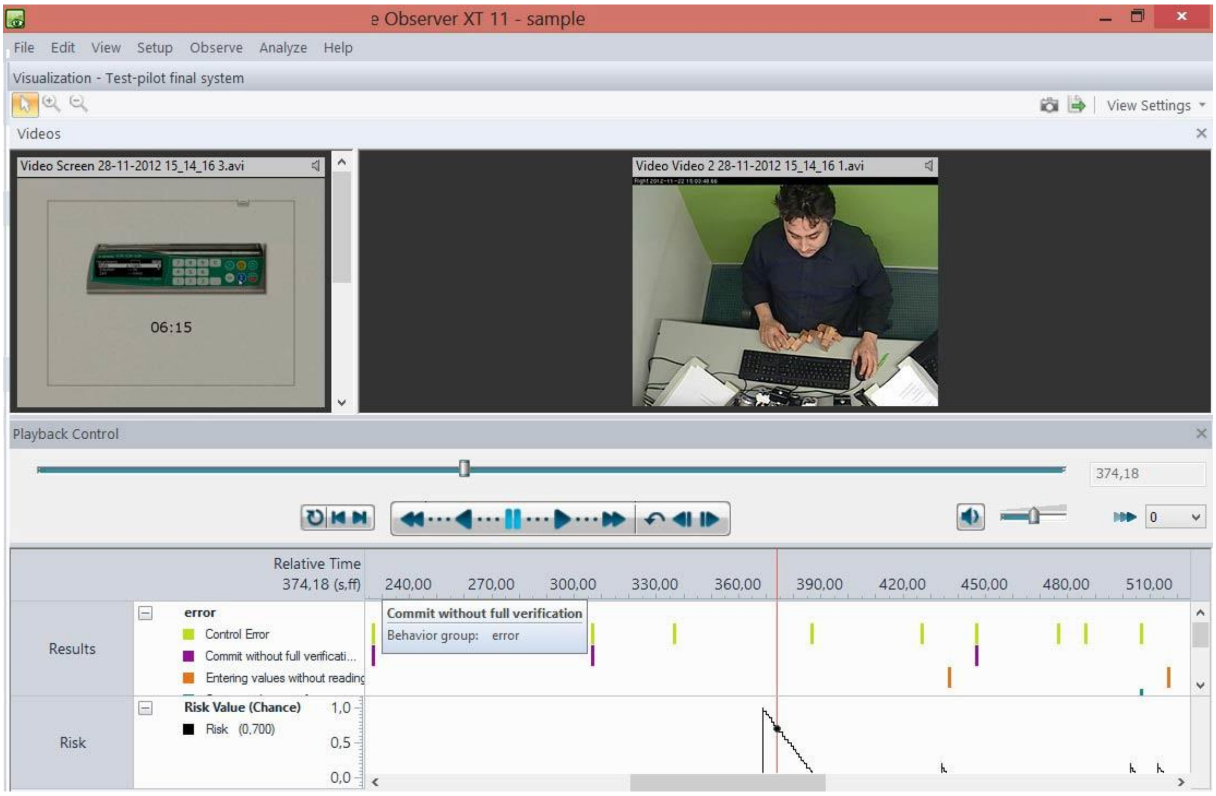

Figure 2. Screenshot of the risk assessment module in Noldus Observer XT 11.0.

Risk values are summed over time, and small risk values added in rapid succession lead to an increasing risk value over time. A decay function is used to lower the risk after each frame. Apart from the analysis of these values over a task performance, a threshold is set on this risk value to allow for the trigger of alerts to the worker or a supervisor. See Figure 2 for a screenshot of the risk estimation module in Noldus Observer XT 11.0.

Discussion and Future Work

We have applied the Patria system in a training scenario where subjects had to set the values of an infusion pump (see Figure 2, left-upper corner) based on patient information. We considered three different executions of the task: (1) read the correct value from a screen and enter it, (2) read the correct value from a form and enter it, (3) read the correct value from the screen, fill it in on a form and enter it in the simulator. This workflow was modelled in a state diagram. Risky actions were assigned when the subject did not check the number that was filled in, when the subject failed to finish the task in time, when the form was not taken or stored at the right location, or when several of the actions were repeated within one session. At semi-random moments, the subject was prompted with a beep to perform one of the three tasks. In the meantime, to increase the cognitive load of the participants and to ensure that enough non-task-related behaviour was displayed, the subjects were to solve a wooden snake puzzle. Currently, we make risk decisions mainly at the end of a “turn”. However, our rules could also have a temporal component, such as action B should be performed within 10 seconds after action A.

We tested the system using 20 participants. Overall, risk values corresponded with episodes in which the subjects made mistakes, usually by not checking the entered values, or by not finishing before the next beep. In some cases, a missing detection of gaze towards the screen resulted in an estimated risk of not having verified what was entered on screen. More robust eye gaze estimation could solve this issue. We found that the use of facial expressions did not add to the performance of the system, typically because subjects were often not showing much variation in their facial behaviour. We expect that the monitoring of user state from the face becomes more important and informative when the task performance is measured over extended periods of time. Boredom and frustration then might be more salient.

In the future, we plan to reduce the number of components in the system, notably the number of processing units. Also, we will combine most of the equipment in a single device. Eventually, we aim at designing a robust low-cost worker observation system that is portable and can be adjusted to different workplaces. From a processing point of view, we will address automatically learning the state diagrams for normal and abnormal task performance, and facilitate the process of assigning risk values to state transitions. Especially when several tasks are performed in parallel, this is a challenging process.

Acknowledgements

This work is supported by the INTERREG IV A project Patria2 (http://patria2.eu/), and partly financed with the European Fund for Regional Development (EFRO), coordinated by Euregio. The authors like to thank the partners for their collaboration in the project: Noldus (Tobias Heffelaar), VicarVision (Marten den Uyl, Jordi Bieger, Mark de Greef), Fachhochschule Münster (Uvo Hölscher, Michael Lindenthal, Heiko Verwold).

References

1. Den Uyl, M.J., Van Kuilenburg, H. 2005. The FaceReader: Online facial expression recognition. Proceedings of the International Conference on Methods and Techniques in Behavioural Research (Measuring Behaviour), Wageningen, The Netherlands, pp. 589-590.

2. San Agustin, J., Skovsgaard, H., Mollenbach, E., Barret, M., Tall, M., Hansen, D. W., and Hansen, J. P. 2010. Evaluation of a low-cost open-source gaze tracker. In Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications (ETRA). Austin, TX, pp. 77-80.

3. Zimmerman, P. H., Bolhuis, J. E., Willemsen, A., Meyer, E. S., & Noldus, L. P. 2009. The Observer XT: A tool for the integration and synchronization of multimodal signals. Behavior Research Methods, 41(3), pp. 731-735.