Towards measuring mental workload from facial expressions

Andre Potenza
Center for Applied Autonomous Sensor Systems (AASS), Örebro University
Email: andre.potenza@oru.se

***Abstract*— Multitasking is a common issue negatively impacting performance in robotic teleoperation and in particular, as we argue, in telepresence. Operating a telepresence robot typically involves engaging in a social interaction with other people who are collocated with the robot, while simultaneously having to control the robot, possibly resulting in an elevated mental workload. One way to mitigate this adverse effect is to have the telepresence robot execute certain tasks autonomously - when necessary. In this extended abstract, we discuss how mental workload measurements can contribute towards dynamically allocating tasks between user and robot, so that a high performance can be achieved, ideally throughout all ongoing tasks. To this end, we are proposing a method for estimating users' mental workload from facial expressions via learned models.**

INTRODUCTION

Remotely controlling a robot in a distant environment requires training with the control interface, and insufficient sensory input causes users to operate with limited information about the robot's surroundings. Moreover, this control task is often accompanied by additional activities, such as social interaction in the case of robotic telepresence [1]. The result is an increased mental workload and possibly diminished situation awareness, as users may have difficulties taking in and processing all relevant information that is available to them. If the operator's workload capacity is exceeded, this, in turn, can lead to reduced performance in one or all of the tasks being performed [2].

Indeed, some of the above-mentioned challenges can be mitigated by upgrading the teleoperated robot with capable sensors and efficient user interface design [3], by way of which a high level of situation awareness (SA) can be attained and retained with less effort [4]. However, the problem of multitasking persists regardless, and with the expectation of future telepresence robots providing enhanced actuation capabilities beyond navigation, users' workload is projected to increase even further. We argue that one potential solution to this issue could be found in mixed-initiative adjustable autonomy [5], in which the robot can decide to reallocate a subset of its functions, if deemed necessary.

A mixed-initiative adjustable autonomy system allows both human and robot to initiate a handover or takeover of functions or entire tasks. While the human user may trigger such a shift for any reason and at any point, the robot is required to have clear, predefined and measurable criteria to decide when and which task should be reallocated.

If a given task can be performed reasonably well by both agents, i.e., human operator and robot, in at least a subset of all possible situations (e.g., navigation), it is eligible for dynamic assignment between them. In an adjustable autonomy system, we identify two primary sets of criteria for determining how the combined total of the system's functionalities should be distributed:

1) Task-specific: If a task is eligible for automation, the robot needs to monitor its execution perpetually, regardless of which agent is currently in control of it. If the task performance falls below a preset threshold (for longer than a preset duration), its control authority may be shifted.
2) User-specific: Humans possess an intrinsic, yet variable capacity for processing information that is available to them. Human factors research involves a set of cognitive constructs that describe users' capacity to understand and process information available to them based on a multitude of cognitive constructs. Those constructs include, among others, situation awareness (SA) [6], [7], mental workload [8], stress and fatigue.

Obviously, these two classes are not entirely exhaustive, as the difficulty and criticality of tasks may vary in dynamic environments, with implications for the preference of the respective agent being in control. That notwithstanding, for the above-mentioned use case of teleoperated robots, a combination of these two classes is required. For instance, it would not make sense to assign all tasks to the human operator just because they are better at all of them, as it could cause mental overload and consequentially result in a loss of overall system performance. Hence, a key requirement of an effective mixed-initiative adjustable autonomy system is a reliable evaluation of users' mental states.

Here, we discuss two constructs from the human factors and ergonomics research and their suitability as metrics for monitoring of the operator state.

HUMAN FACTORS MEASUREMENT

As a cognitive construct, SA plays a vital role in automation - particularly when it comes to deciding on the appropriate level of automation (LOA) of a task or functionality. In fact, SA, together with mental workload, are typically in a complex interplay with the LOA and the impact on all three of these factors needs to be carefully considered when designing a system [9]. As a general rule of thumb, both workload and SA can be expected to decrease as the LOA is increased. While a low level of workload is in most cases desired, a low level of SA can be detrimental to system performance, as automation is often imperfect and expected to fail in some situations. When this happens, a high degree of maintained SA allows users to assess the situation quickly and take appropriate measures to guide the system back to a nominal state.

While its relevance in system design is well established, its essence and the ways in which it is commonly measured [10] are very closely related to task performance rather than the operator's general mental state. As such, it does not add much information to our user-specific criteria class. Since measurement of the operator state should be unintrusive, implicit and objective, for this purpose it appears worthwhile to examine mental workload more closely.

Mental workload is a well-studied cognitive concept, researched in a variety of areas ranging from cognitive psychology to applied sciences such as user design. Yet, and even though almost everybody has an intuitive idea of what it denotes, there exists no single universally accepted definition of it in the literature [8], [2]. For most purposes, it could be described as the relative degree to which an individual's personal mental processing capacity is exhausted by the entirety of the mental processing that they are performing at a given time. Thus, if their capacity is exceeded and another task occupying the same mental resources is added, the performance in at least one of the currently performed tasks is expected to decline. In fact, it has been shown and is worth noting that tasks of disparate nature (e.g., spatial vs. verbal, visual vs. auditory) do not necessarily occupy the same attentional resources and may be performed simultaneously without interfering with one another [11].

In experimentation, a common way of recording subjects' workload is subjective self-reports at several points throughout an experiment. Arguably the most widely used tool for such reports is the NASA-TLX (Task Load Index) questionnaire [12], which allows subjects to rate perceived task difficulty and workload across multiple dimensions.

On the other hand, various physiological measures have been used to estimate workload objectively [13], ranging from slightly invasive (e.g., heart rate variability [14], skin conductance etc.) to more invasive (EEG [15]). While some of these methods have shown success, they lack practicability for casual users of telepresence robots. Since in any telepresence robot a camera is recording the operator's face by design, we intend to investigate the possibility of estimating users' workloads from facial expressions.

Proposed Method and Experiment

In recent years, deep neural networks have been applied to detect a variety of features from facial expressions, such as emotions [16], gender and age [17], arousal, etc.

In the proposed experiment we plan to have participants perform mental tasks of varying difficulty levels and intermittently report their experienced workload levels via the NASA-TLX [12]. Throughout the experiment, their faces are recorded with a regular RGB camera. From the collected time series data (video footage) and the reports serving as ground truth, we will train a recurrent convolutional neural network [18] whose purpose will be to classify the workload of individuals over time series segments.

Conclusion and Future Outlook

In teleoperation, automation should be partial and selective, dynamically adapting to the current user's condition, skills and needs. We have discussed the two broad means based on which performance can be evaluated - task-specific or user-specific. The latter type, if measured accurately enough, can function as a support for autonomous decision making in overall task allocation between human and robot. For mental workload, several different approaches exist as they can be either subjective or objective, as well as more or less invasive. In this paper, we argued for an approach that better suits the typical requirements found in telepresence robotics. This is what we aim to investigate in an upcoming user study.

Acknowledgement

This work has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 721619 for the SOCRATES project.

References

[1] M. Desai, K. M. Tsui, H. A. Yanco, and C. Uhlik, "Essential features of telepresence robots," in *Technologies for Practical Robot Applications (TePRA), 2011 IEEE Conference on*. IEEE, 2011, pp. 15–20.

[2] M. S. Young, K. A. Brookhuis, C. D. Wickens, and P. A. Hancock, "State of science: mental workload in ergonomics," *Ergonomics*, vol. 58, no. 1, pp. 1–17, 2015.

[3] A. Kristoffersson, S. Coradeschi, and A. Loutfi, "A review of mobile robotic telepresence," *Advances in Human-Computer Interaction*, vol. 2013, p. 3, 2013.

[4] M. R. Endsley, "Designing for situation awareness in complex systems," in *Proceedings of the Second International Workshop on symbiosis of humans, artifacts and environment*, 2001, pp. 1–14.

[5] B. Sellner, F. W. Heger, L. M. Hiatt, R. Simmons, and S. Singh, "Coordinated multiagent teams and sliding autonomy for large-scale assembly," *Proceedings of the IEEE*, vol. 94, no. 7, pp. 1425–1444, 2006.

[6] M. R. Endsley, "Measurement of situation awareness in dynamic systems," *Human factors*, vol. 37, no. 1, pp. 65–84, 1995.

[7] ——, "Toward a theory of situation awareness in dynamic systems: Situation awareness," *Human factors*, vol. 37, no. 1, pp. 32–64, 1995.

[8] B. Cain, "A review of the mental workload literature," Defence Research And Development Toronto (Canada), Tech. Rep., 2007.

[9] R. Parasuraman, T. B. Sheridan, and C. D. Wickens, "Situation awareness, mental workload, and trust in automation: Viable, empirically supported cognitive engineering constructs," *Journal of Cognitive Engineering and Decision Making*, vol. 2, no. 2, pp. 140–160, 2008.

[10] P. Salmon, N. Stanton, G. Walker, and D. Green, "Situation awareness measurement: A review of applicability for c4i environments," *Applied ergonomics*, vol. 37, no. 2, pp. 225–238, 2006.

[11] C. D. Wickens, "Multiple resources and mental workload," *Human factors*, vol. 50, no. 3, pp. 449–455, 2008.

[12] S. G. Hart and L. E. Staveland, "Development of nasa-tlx (task load index): Results of empirical and theoretical research," in *Advances in psychology*. Elsevier, 1988, vol. 52, pp. 139–183.

[13] S. H. Fairclough, "Fundamentals of physiological computing," *Interacting with computers*, vol. 21, no. 1-2, pp. 133–145, 2009.

[14] A. Hoover, A. Singh, S. Fishel-Brown, and E. Muth, "Real-time detection of workload changes using heart rate variability," *Biomedical Signal Processing and Control*, vol. 7, no. 4, pp. 333–341, 2012.

[15] A.-M. Brouwer, M. A. Hogervorst, J. B. Van Erp, T. Heffelaar, P. H. Zimmerman, and R. Oostenveld, "Estimating workload using eeg spectral power and erps in the n-back task," *Journal of neural engineering*, vol. 9, no. 4, p. 045008, 2012.

[16] A. Mollahosseini, D. Chan, and M. H. Mahoor, "Going deeper in facial expression recognition using deep neural networks," in *Applications of Computer Vision (WACV), 2016 IEEE Winter Conference on*. IEEE, 2016, pp. 1–10.

[17] G. Levi and T. Hassner, "Age and gender classification using convolutional neural networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2015, pp. 34–42.

[18] O. Vinyals, A. Toshev, S. Bengio, and D. Erhan, "Show and tell: A neural image caption generator," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3156–3164.