Stop Helping Me - I'm Bored! Why Assembly Assistance needs to be Adaptive.

Markus Funk
University of Stuttgart
Stuttgart, Germany
markus.funk@vis.uni-stuttgart.de

Tilman Dingler
University of Stuttgart
Stuttgart, Germany
tilman.dingler@vis.uni-stuttgart.de

Jennifer Cooper
University of Minnesota
Minneapolis, MN, USA
cooperj@umn.edu

Albrecht Schmidt
University of Stuttgart
Stuttgart, Germany
albrecht.schmidt@vis.uni-stuttgart.de

Ubicomp/ISWC'15 Adjunct, September7–11, 2015, Osaka, Japan.
 ISBN 978-1-4503-3575-1/15/09 .. $15.00
DOI: http://dx.doi.org/10.1145/2800835.2807942

Abstract
With the demographic change and a generally increasing product complexity, there is a growing demand for assistance technology to cognitively support workers during industrial production processes. Many approaches including head-mounted displays, smart gloves, or in-situ projections have been suggested to provide cognitive support for the workers. Recently, research focused on improving the cognitive feedback by using activity recognition to make it context-aware. Thereby an assistance technology is able to detect work steps and provide additional feedback in case the worker makes mistakes. However, when designing feedback for a rather monotonous task, such as product assembly, it should be designed in a way it does neither over-challenge nor under-challenge the worker. In this paper, we sketch out requirements for providing cognitive assistance at the workplace that can adapt to the worker's needs in real-time. Further, we discuss challenges and provide design suggestions.

Author Keywords
context-aware instructions; cognitive assistance; adaptive user interfaces;

ACM Classification Keywords
H.5.m [Information interfaces and presentation (e.g., HCI)]: Miscellaneous

Introduction

Assistive technology is influencing the evolution of industrial workplaces due to the outcomes of reducing errors and improving workflow. With the average age of laborers on the rise, it is important to be proactive in meeting the needs of production lines. By integrating intelligent machinery into factories, commands can be delivered to workers on a step-by-step basis. Frequent commands can benefit workers in training as well as workers with disabilities; however, there are times when the necessity for direction falls, as is the case with expert workers. When the workers' need for assistive technologies changes, the feedback from the machine should also reduce, while continuously checking for errors.

The benefit of creating adaptive assistive systems is that they conform to the cognitive needs of the worker in real-time. When a worker is over-taxed, they exhibit symptoms of stress and anxiety. These symptoms can be monitored through bio-signals, such as heart pulse rate, galvanic skin response (GSR), electroencephalography (EEG), electromyography (EMG). Furthermore, eye movements can be used as indicators for cognitive workload. Just and Carpenter [6] investigated the relation between eye fixations and cognitive processes. This data can be used to deduct workers' current workload in real-time. When combined with system analytics data, such as task completion time and error rates in task assembly, we can deduct a holistic picture of the worker's current state.

The vision for this project is to provide an adaptive technology to assist workers based on level of experience (e.g., beginner, advanced, expert). The ultimate goal is to provide a tool that can be used in different scenarios and different user groups.

Related Work

Assisting workers during complex assembly tasks has been a topic of various research projects. Bannat *et al.* [1] use a projector that is mounted on top of a workplace to highlight the bin that the worker needs to pick the next part from. They use a camera to track if the worker's hand is in the correct bin. Further, the worker is equipped with a grasping sensor that is mounted on the workers hand. The system only triggers a correct pick if the hand is in the correct bin and the user does a grasping gesture. To give assembly hints, they are using pictorial instructions that show both start and end of an assembly step. On the other hand, Funk *et al.* [4] evaluated which projected in-situ feedback is perceived best by cognitively impaired workers. Their results suggest that using a contour-based visualization that highlights the position and orientation of work-pieces to assemble, increases the task completion time and reduces the number of errors that are made. Further, a contour-based-visualization leads to less perceived cognitive effort (measured using the NASA-TLX [5]). Such a contour-based visualization is also used by Zhou *et al.* [11], who apply projectors to highlighting welding spots. In their approach, they highlight the spot with a green color. Additionally, to easily notice the location of the welding spot they propose the use of different visualizations, e.g. an arrow pointing to the next spot, shrinking circles, and moving arrows around the target position.

However, all these different visualizations compete for the attention of the user. But cognitive resources are limited. Especially when assembly tasks become more complex due to multiple working steps. Here, an adaptive system could break down instructions for those steps into digestible chunks: Miller [7] found the feasibility of about 7 chunks (plus or minus 2) that can be held in working memory, which points to the capacity limitations when processing informa-

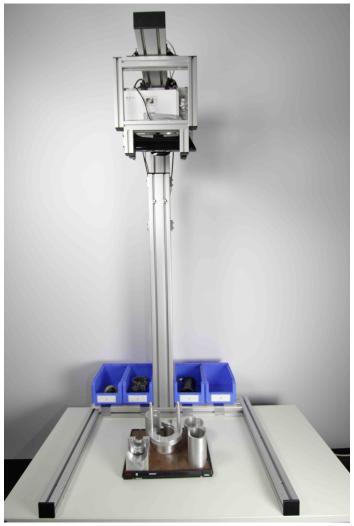

Figure 1: We equipped a workplace with a camera-projector pair. It is able to project assembly instructions and react upon worker's actions.

Figure 2: A worker is placing a workpiece at the working area according to a highlighted position.

tion. To decrease complexity and facilitate learnability of assembly tasks, additional interaction dimensions could be added to a workspace: Dingler *et al.* [3] used different interaction zones to map functionality across space thereby reducing potentially ambiguous interactions. Using such interaction spaces can further be used to provide an adaptive work space where monotonous body posture is prevented and workers are encouraged to physically move in between working blocks.

Depending on their nature, multiple tasks can be performed concurrently or not. Wickens [10] describes this *multiple resource theory* in detail. So while workers are focused on their task at hand, they use their auditory channel to implicitly monitor their work environment while receiving instructions through the visual channel. This sensory multi-tasking affects cognitive workload, which again has an effect on workers' performance. Although it is not always the case that increased workload necessarily decreases performance, but when workloads become too high or too low performance is affected negatively [8]. While low workload (underload) can lead to boredom, loss of situation awareness and reduced alertness, too high workloads can cause stress, anxiety, and therefore negatively impact the worker's health. Pielot *et al.* [9] investigated how boredom can be detected and taken advantage of in a mobile context. A similar approach might be possible for assembly task lines where data collected by observing workers can be used to build predictive models to infer the worker's state of attention and intervene accordingly.

An Adaptive Assistive System

We equipped a regular assembly workplace with a top-mounted projector to provide in-situ feedback during manufacturing (see Figure 1). Further, we added a *Microsoft Kinect* depth camera that is able to sense activities that are performed at the workplace. By sensing the worker's activities, the system is aware of the current state of the assembly and can give feedback according to the performed assembly step. Our system can distinguish between three activities that are performed at the workplace: first, it detects whether the worker picks an assembly part from the correct box. Second, it detects if the part was correctly assembled or if it is still in an intermediate state. And third, it detects whether a tool was taken from its defined spot using the SURF [2] object recognition algorithm.

Our prototype implementation can adapt the feedback to the user's performance as it provides three levels of feedback: beginner mode which provides video instructions for every step, advanced mode which just highlights the contour of the pick and assembly locations, and expert mode which does not show any feedback. While assembling, the system counts the amount of correctly performed steps and the number of errors, which consist of assembly errors and picking errors. Every time the user performs a step correctly or makes an error the system calculates a $correct\ steps/errors$ ratio. When the ratio is below a threshold of 0.1 the feedback level is switched to advanced mode. Accordingly, if the ratio drops below a threshold of 0.01, the system displays the expert mode feedback. Thereby the system assumes, if the worker does not make any errors, the worker has learned the workflow and does not need feedback anymore. The values for the threshold were determined empirically to fit our example product.

Discussion

While this approach is an easy way to provide visual feedback that is adapted to the user's performance as it is only using data from one sensor i.e. the Kinect depth camera. However, in real production environments errors can be triggered by the worker as quick as hovering a hand over

Figure 3: The position of the part that needs to be picked next is highlighted by the system.

the wrong box, which might cause the system to switch to a lower feedback level. This level might sustain for several working steps until a lower ratio is reached again. Therefore we argue to take more variables into account.

The system could not only take the worker's actions for calculating the feedback level, but additionally use vital parameters: **galvanic skin response** can be used to measure how much a worker is sweating. **Heart rate** and **breathe frequency** can be measured to see how physically exhausting a task is for the worker. Additionally, the system could collect the workers **EEG** data to measure an attention and meditation values. Finally, the worker's **blinking rate** can be measured to react upon fatigue effects.

We believe that collecting a worker's vital parameters to adjust the provided feedback brings us closer to real-time adaptive assembly workplaces. Using the real-time sensing of worker's performance, vital parameters, and cognitive load, the workplace could use different modalities such as visual, audio or haptic feedback according to the worker's stress level. Additionally, even within the modalities, the stress level can be taken into account by adjusting the frequency and intensity of the provided feedback. We believe that this approach has some benefits for both workers and companies. First, work tasks could be dynamically queued according to a worker's stress level. This would ensure product quality, increase the production efficiency and increase worker happiness. Second, our approach lowers workers' stress levels by adapting the task to their current capacity. This could have positive implications on workers' health by avoiding stress-related symptoms, such as burnout syndromes. Third, by retaining a constant concentration level, the risk for accidents at the workplace occurring due to a lack of focus, might be lower.

However, there are a few challenges for our approach. First, the worker needs to wear additional sensors to measure the worker's vital signs. This would interfere with the work task and might be an additional load for the workers. Therefore, the measuring of the vital signs should be designed as unintrusive as possible. Second, it should be mentioned that our approach comes with privacy issues. An employer could get access to the vital parameters of the workers and infer information about their health. Therefore, when designing such a system, we need to ensure that all parameters are only used to calculate an overall feedback level and that none of the parameters can be accessed from outside.

Conclusion & Future Work

In this position paper we described our prototype for providing context-aware in-situ assembly instructions at the workplace using a camera-projector pair. We outline benefits and challenges for making the projected instructions adaptive by considering worker's stress levels and vital parameters. Additionally we suggest different strategies to counteract high stress levels and cognitive over-load and under-load by changing the projected feedback or dynamically changing the currently produced product.

In future work, we want to conduct experiments to evaluate the effects of the proposed strategies and changes in the feedback levels on the vital parameters of the worker. We believe that adaptive feedback and adaptive queuing strategies can help to maintain a solid cognitive load and can keep work challenging even when performing monotonous tasks.

Acknowledgments

This work is funded by the German Federal Ministry for Economic Affairs and Energy in the project motionEAP, grant no. 01MT12021E. The authors further acknowledge the financial support of the Future and Emerging Technologies (FET) programme within the 7th Framework Programme for Research of the European Commission, under FET grant number: 612933 (RECALL).

REFERENCES

1. A Bannat, F Wallhoff, G Rigoll, F Friesdorf, H Bubb, S Stork, HJ Müller, A Schubö, M Wiesbeck, and MF Zäh. 2008. Towards optimal worker assistance: a framework for adaptive selection and presentation of assembly instructions. In *Proceedings of the 1st international workshop on cognition for technical systems, Cotesys*.
2. Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. 2006. Surf: Speeded up robust features. In *Computer vision–ECCV 2006*. Springer, 404–417.
3. Tilman Dingler, Markus Funk, and Florian Alt. 2015. Interaction Proxemics: Combining Physical Spaces for Seamless Gesture Interaction. In *Proceedings of the 4th International Symposium on Pervasive Displays (PerDis '15)*. ACM, New York, NY, USA, 107–114.
4. Markus Funk, Andreas Bächler, Liane Bächler, Oliver Korn, Christoph Krieger, Thomas Heidenreich, and Albrecht Schmidt. 2015. Comparing Projected In-Situ Feedback at the Manual Assembly Workplace with Impaired Workers. *Proc. PETRA'15* (2015).
5. Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in psychology* 52 (1988), 139–183.
6. Marcel Adam Just and Patricia A Carpenter. 1976. Eye fixations and cognitive processes. *Cognitive psychology* 8, 4 (1976), 441–480.
7. George A Miller. 1956. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological review* 63, 2 (1956), 81.
8. Friedhelm Nachreiner. 1995. Standards for ergonomics principles relating to the design of work systems and to mental workload. *Applied Ergonomics* 26, 4 (1995), 259–263.
9. Martin Pielot, Tilman Dingler, Jose San Pedro, and Nuria Oliver. 2015. When Attention is not Scarce - Detecting Boredom from Mobile Phone Usage. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM.
10. Christopher D Wickens. 1991. Processing resources and attention. *Multiple-task performance* (1991), 3–34.
11. Jianlong Zhou, Ivan Lee, Bruce Thomas, Roland Menassa, Anthony Farrant, and Andrew Sansome. 2011. Applying spatial augmented reality to facilitate in-situ support for automotive spot welding inspection. In *Proceedings of the 10th International Conference on Virtual Reality Continuum and Its Applications in Industry*. ACM, 195–200.