Neuroergonomic Assessment of Human Robot Interaction

Master Thesis

Systems Neuroscience & Neurotechnology Unit
Saarland University of Applied Sciences
Faculty of Engineering

Submitted by	:	Dominik Limbach, B.Sc.
Matriculation Number	:	3662306
Course of Study	:	Biomedical Engineering (Master)
Specialisation	:	Neural Engineering
First Supervisor	:	Prof. Dr. Dr. Daniel J. Strauss
Second Supervisor	:	Dr. Lars Haab

Saarbrücken, December 9, 2019

Abstract

In recent years wearables have become a staple in our society. In the form of Smartwatches and fitness-tracking devices wearables have made their way into nearly a third of german households. Aside from the basic functions of a phone or a watch these devices offer a wide array of functionalities, many of which revolve around monitoring the user's vital parameters for health reasons as well as improving their physical performance in recreational activities.
Research suggests that there is great potential in the application of psycho-physiologic monitoring in highly demanding workplaces of today's industry, an area in which wearables are still largely underrepresented.
Therefore, we built a new system based on a wrist worn device capable of monitoring workers in stressful workplaces, specifically collaborative workplaces involving Human-Robot-Interaction (HRI), and predicting their current mental state on a real-time basis. With the deployment of such a system we are able to create neuroergonomic workplaces that are sensitive to a persons mental capability and capable of eliminating stress as one of the leading causes of injury and disease in the working population.

Zusammenfassung

In den letzten Jahren sind Wearables zu einer festen Größe in unserer Gesellschaft geworden. In der Form von Smartwatches und Fitness-Armbändern haben Wearables Einzug in annähernd dreißig Prozent aller deutschen Haushalte gehalten. Neben den Grundfunktionen eines Telefons oder einer Uhr bieten diese praktischen Geräte eine Bandbreite an Funktionen. Viele dieser Zusatzfunktionen befassen sich mit dem Messen der Vitalparameter des Trägers zur Früherkennung von Krankheiten oder zur Optimierung der sportlichen Leistungsfähigkeit. Forschungsarbeiten der letzten Jahre lassen das Potential der Anwendung von psychophysiologischem Monitoring in besonders stressvollen Arbeitsplätzen erahnen, ein Gebiet in dem Wearables nur selten anzutreffen sind. Aus diesem Grund haben wir ein System für Arbeiter an besonders stressvollen Arbeitsplätzen, insbesondere collaborative Arbeitsplätze mit Mensch-Roboter-Interaktion (HRI). Das System basiert auf einem Wearable, welches am Handgelenk getragen wird, und ist in der Lage den mentalen Zustand des Trägers zu beurteilen. Durch den Einsatz eines solchen Systems sind wir in der Lage neuroergonomische Atbeitsplätze zu schaffen die sich an die cognitive Leistungsfähigkeit eines Arbeiters anpassen und somit den Einfluss von Stress, als einen der führenden Gründe für Krankheiten und Verletzungen unter der arbeitenden Bevölkerung, drastisch zu senken.

Declaration

I hereby declare that I have authored this work independently, that I have not used other than the declared sources and resources, and that I have explicitly marked all material which has been quoted either literally or by content from the used sources. This work has neither been submitted to any audit institution nor been published in its current form.

Saarbrücken, December 9, 2019

Dominik Limbach, B.Sc.

Contents

1 Introduction

Neuroergonomics is often described as the study of brain and behavior at work. As the name suggests, neuroergonomics is comprised of two disciplines, neuroscience and ergonomics (also known as human factors). Neuroscience is concerned with the structure and the function of the brain. It is a highly interdisciplinary body of research spanning disciplines such as physiology, psychology, medicine, computer science, or mathematics. Human factors on the other hand is focused on examining the human use technology at work or other real-world settings. As the intersection of these two fields, neuroergonomics addresses both the brain and humans at work, but even more so their dynamic interaction [1] . By understanding the neural bases of perceptual and cognitive functions, such as seeing, hearing, planing and decision making in relation to technology and settings in the real world, neuroergonomics strives to develop new optimization methods for various areas of applications. The additional value neuroergonomics can provide, compared to 'traditional' neuroscience and 'conventional' ergonomics, promises substantial economical benefits as well as significant improvements to health care and therefore society at large. In the scope of this thesis, we will focus primarily on applications in work settings such as modern automated systems. An area in which the effects of neuroergonomics are expected to be even greater, considering the difficulty of obtaining measures of overt behavior [1].

Automated systems, in one form or another have been present in almost every branch of industry for the better part of what has been two centuries. Automation itself can be defined as the creation and application of technology by which the production and delivery of various goods and services is controlled and monitored with minimal human assistance. Automation can be encountered in many different places, from a simple control loop in a hydraulic system up to an artificial intelligence handling emergency breaking in cars, based on scans of the car's environment. Static automation is the original form of automation. In this form, automation is an all-or-none technology either performing a task for us or not [2]. Although, there are numerous benefits (i.e. relieving workers from the strain of performing repetitive tasks) to it, there also is increasing evidence that static automation comes at the price of impaired decision making, manual skill degradation, loss of situational awareness, and monitoring inefficiency [2]. These problems stem from the significant change of roles that automation causes in a work environment. What this means is that workers, once active operators of machines and technology, are now passive monitors who often face monitoring workloads that are inherently different and often significantly higher when compared to the manual control conditions of a non automated work environment. The continuous exposure to workloads that are either too high or too low can have dramatic effects on human-system performance, thereby potentially compromising safety [3]. This has already been indicated by the work of Parasuraman et al. (1993, 1994) who tested subjects in a multi-task flight simulator with optionally automatable components and reported a substantial decrease of human operator detection of automation failures after short periods of time in a static automation scenario with constant task assignment of both operator and the automated system [2]. However, we must not forget the consequences these highly automated and ever so stressful workplaces have on human operators specifically. Studies by Cooper et al. showed that working in a stressful environment increases the risk of suffering physical illness or symptoms of psychological distress, as well as work related accidents and injuries [4].

We will now take a look at what is called adaptive automation. Adaptive automation has been introduced to resolve the abovementioned issues of statically automated systems. Whereas static automation is considered to be an agent working for the operator, adaptive automation is viewed as an interactive aid working with the operator [2]. It attempts to optimize system performance by adjusting the task assignment between the human operator and automation dynamically. This task reallocation is based on task demands, user capabilities, and system requirements. Meaning that during high task load conditions or emergencies the use of automation is increased and decreased during normal operations [3]. Another advantage of adaptive over static automation is the ability to reconstruct the task environment in terms of what is automated, how it is automated, what tasks may be shared, and when changes occur [2]. However, for an adaptive system to be efficient we require both the operator and the automated system to have sufficient knowledge of each other's current capabilities, performance and state [2]. As discussed by Byrne et al. (2006), there are three main approaches to address this issue, which either use model-based prediction or continuous measurements to determine the operators current state. In the scope of this thesis we will focus on the latter of the two types. Usually, the way to measure human performance at work is to use physiological measures that reflect, more or less directly, aspects of brain functions. There is however a group of measures that is particularly favored among neuroergonomic researches. Those are the ones that are derived from the brain itself such as electroencephalography (EEG), magnetencephalography (MEG), and event-related potentials (ERPs), as well as measures related to the brain's metabolic and vascular responses such as positron emission tomography (PET), and functional magnetic resonance imaging (fMRI) [1]. Although, there are many advantages to this group of measures, such as the temporal resolution of ERPs and the spatial resolution of fMRI, there is still one major disadvantage. In fact, the majority of the abovementioned measures are either too expensive, impose too much restrictions on the movement of the subject, or are simply unfit to be used in a portable system, which ultimately prevents their application in real-world settings. Also, the lack of comfort and therefore low operator acceptance of a certain measure could have detrimental effects on the success of an adaptive automated system.

An alternative could by provided by systems that use psychophysiological indices to trigger changes in automation. In general, pychophysiology is focused on physiological measures and their psychological correlates [1], but there are many psychophysiological indices that reflect underlying cognitive activity, arousal levels, and external task demand. Some of these include cardiovascular measures (e.g. heart rate, heart rate variability), respiration, galvanic response, ocular motor activity, and speech [5]. Even though, research examining the utility of psychophysiology in adaptive automation has been rare, physiological measures are likely to be considered in the design of adaptive systems, either in isolation or in combination with other measures [2]. In addition psychophysiological measures are usually well accepted, due to their non-invasive nature and easy of application.

In conclusion, human-machine interaction in highly automated workplaces can be optimized by creating a work environment that is sensitive to the mental state of human operators. Using psychophysiological measures, a constant feedback in the form of a neuroergonomic assessment of the operator condition is provided to the machine agent. Consequently, adaptation mechanisms can be deployed to alter the quantity, or quality of the workload according to operator capability.

Within the scope of this thesis we designed a wearable system that facilitates this process. We built our system upon the Empatica E4 wristband, a medical-grade wearable device capable of real-time physiological data acquisition. We then conducted a pilot experiment, deploying our system in near real-world conditions. We monitored 14 subjects performing two cognitive tasks, simulating high and medium workload conditions, and two sets of visual stimulation, designed to elicit specific emotional states. Finally, evaluated different machine learning algorithms using the acquired dataset. With our system we address a distinct need for a reliable, and unobstrusive method of handling neuroergonomic assessment of human-machine interaction in collaborative work environments. Therefore, we are able to create neuroergonomic workplaces that are sensitive to a persons mental capability and capable of eliminating stress as one of the leading causes of injury and disease in the working population.

1.1 Acknowledgments

2 Problem Analysis and Goals

2.1 Theoretical Background

2.2 Related Work

3 Materials and Methods

3.1 Materials

All experiments and measurements were conducted within the facilities of Systems Neuroscience and Neurotechnology Unit, particularly the Green Lab, located at the University Hospital Saarland, and the Mindscan Lab, located at the HTW Saar (Technikum).

3.1.1 System Components

3.1.2 Hardware

Empatica E4 Wristband The Empatica E4 wristband is a wearable wireless device designed for comfortable, continuous, real-time data acquisition. It is a class IIa medical device in the EU, according to CE Cert. No. 1876/MDD (93/42/EEC Directive) and was designed for daily life usage [6].

Figure **??** shows an overview of the entire E4 wristband from either side indicating key attributes as wells as a total of four different sensors that will be discussed briefly in the following:

- **Photoplethysmography (PPB)** to provide blood volume pulse (BVP), from which heart rate, heart rate variability and other cardiovascular features may be derived
- **Electrodermal activity (EDA)** is used to measure sympathetic nervous system arousal and to derive features related to stress, engagement and excitement
- **3-Axis accelerometer** to capture motion-based activity
- **Infrared thermopile** for reading skin temperature

As the E4 is intended to be worn on the wrist these sensors are set up in a specific way to provide for optimal use. As can be seen on **??** the majority of the sensors are located on the backside of the main unit not including the EDA-sensor, which is located on the wristband itself.

Wearing the E4 wristband is equally intrusive to wearing a watch and therefore providing a high level of convenience compared to other physiologic measures such as electrocardiogram ECG or electroencephalogram EEG.

Sampling Specifications All recordings were performed using only software licensed by Empatica. Using the approved streaming application and the compatible Bluetooth receiver, the recorded data was streamed directly to an operator's personal computer via a Bluetooth connection.

EDA sensor

- Sampling frequency: 4 Hz (Non customizable).
- Resolution: 1 digit 900 pSiemens.
- Range: 0.01 μSiemens – 100 μSiemens.
- Alternating current (8Hz frequency) with a max peak to peak value of 100 μAmps (at 100 μSiemens).
- Electrode(Placement): on the ventral (inner) wrist.
- Electrode(Build): Snap-on, silver (Ag) plated with metallic core.
- Electrode(Longevity): 4–6 months

PPG sensor

- Sampling frequency 64 Hz (Non customizable).
- LEDs: Green (2 LEDs), Red (2 LEDs) Photodiodes: 2 units, total 15.5 mm2 sensitive area.
- Sensor output: Blood Volume Pulse (BVP) (variation of volume of arterial blood under the skin resulting from the heart cycle).
- Sensor output resolution 0.9 nW / Digit.
- Motion artifact removal algorithm: Combines different light wavelengths. Tolerates external lighting conditions.

Infrared Thermopile

- Sampling frequency: 4 Hz (Non customizable).
- Range(Ambient temperature): -40...85degC (if available).
- Range(Skin temperature): -40...115degC.
- Resolution: 0.02degC.
- Accuracy ±0.2degC within 36-39degC.

Real-time clock

- Resolution(Recording mode): 5s synchronization resolution. Average of 6 seconds in 6 million seconds drift.
- Resolution(Streaming mode): Temporal resolution up to 0.2 seconds with connected device.

3.2 Methods

3.2.1 Participants

3.2.2 Stroop Test

3.2.3 Visual Stimulation

3.2.4 Procedure

One of the most important parts to this project was the collection of authentic data that could be used later on to develop a reliable classifier for our system. For that reason a experiment, specifically designed to elicit certain emotional and cognitive states in a subject, was conducted. The following section is focused on the procedure applied in this experiment.

The procedure was comprised of a total of five sessions. Every experiment was initiated with a short briefing session. Containing a short questionnaire, covering personal information of the participant as well as habits that may have a influence on the measurement. Further a series of questions, regarding their handedness, use and frequency of use of watches or other wearables was posed, to estimate the additional influence that may be caused by wearing the Empatica E4 wristband. Concluding the first session, the participants were given a coarse outline of the experiment covering the structure and a basic description of their responsibilities.

The second session consisted of a baseline measurement used to log the participants form of the day and also to be able to account for environmental influences in the following processing steps. Before the start of the measurement the subject was placed on a chair in front of a monitor (24 inches, Resolution: 1080p) with a approximated distance of 1m. The Empatica E4 was then put on the wrist of the non-dominant hand and secured in a position that caused minimal light leakage to the PPG-sensor and provided optimal contact for the EDA electrodes. After the participants were comfortable with the device a one minute test sequence was measured to verify the functionality of the system. Consequently the paradigm was displayed on the monitor and the session was started. After reading the instructions, in which the subjects were asked to relax and remain still, and confirmation with the participant the measurement was initiated with a ten second countdown to give some additional time for preparation. During the measurement the EDA, BVP, and temperature of the subject were measured for a duration of five minutes. Afterwards, to conclude the second session, the participants had to give a subjective rating of their current mental state, regarding their stress level, ranging from 1 (completely relaxed) to 10 (stressed out).

The third session was comprised of three separate measurements, two cognitive tasks and a relaxation segment. As before a rating followed the recording. The ratings consisted of a subjective assessment by the subjects regarding their stress level. Additionally subjects had to rate the test difficulty on a scale from 1 (very easy) to 10 (very difficult) for both tasks. For the first measurement the subjects were instructed to count down aloud from 700 in steps of 7 while maintaining a certain pace. The counting rhythm was indicated by a flashing dot on the instruction screen, for a duration of five minutes. The dot's color and flashing frequency were altered during the experiment to further increase difficulty at the three and four minute mark. If the participants were to slow down or loose track an instructor would intervene to help. The second task consisted of a Stroop-Word-Color test. The test featured 11 different colors, resulting in a total of 220 trials the subjects had to work through. Although the color palette seems rather extensive when compared to the standard 3 color variation of the test, this was a conscious decision to guarantee a test time of at least 5 minutes to mitigate monotony. Each trial presented the subject with a colored word in the center of the screen and one possible answer to either side. The participants then had to choose the right answer based on the color of the word.Each decision was recorded via a key press on the keyboard.

3.2.5 Signal Analysis

Heart Rate Variability

GSR

Temperature

4 Results

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

5 Discussion

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

6 Conclusions and Future Work

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

List of Figures

List of Tables

Bibliography

[1] R. Parasuraman. Neuroergonomics: Research and practice. *Theoretical Issues in Ergonomics Science*, 2003.

[2] E. A. Byrne and R. Parasuraman. Psychophyiology and adaptive automation. *Biological Psychology*, 2006.

[3] R. K. Mehta and R. Parasuraman. Neuroergonomics: a review of applications to physical and cognitive work. *frontiers in Human Neuroscience*, 2013.

[4] S. Clarke and C. L. Cooper. *Managing the Risk of Workplace Stress. Health and safety hazards.* Routledge. London and New York, 2004.

[5] R. Parasuraman and Rizzo M. *Neuroergonomics*, volume 3 of *Human-Technology Interaction Series*. Oxford University Press, 2008. ISBN 0195368657, 9780195368659.

[6] *E4 wristband from empatica. User's manual.* Empatica, Via Stendhal 36, 20144 Milano (MI). URL www.empatica.com.