CareCam: Towards user-tailored Interventions at the Workplace using a Webcam

Dimitri Kraft
University of Rostock
Rostock, Germany
dimitri.kraft@vcric.igd-r.fraunhofer.de

Angelina Schmidt
University of Rostock
Rostock, Germany
angelina.schmidt@uni-rostock.de

Lea Büttner
Fraunhofer IGD
Rostock, Germany
lea.buettner@igd-r.fraunhofer.de

Frederike Marie Oschinsky
Fraunhofer IGD
Rostock, Germany
frederike.oschinsky@igd-r.fraunhofer.de

Fabienne Lambusch
University of Rostock
Rostock, Germany
fabienne.lambusch@uni-rostock.de

Kristof van Laerhoven
University of Siegen
Siegen, Germany
kvl@eti.uni-siegen.de

Gerald Bieber
Fraunhofer IGD
Rostock, Germany
gerald.bieber@igd-r.fraunhofer.de

Michael Fellmann
University of Rostock
Rostock, Germany
michael.fellmann@uni-rostock.de

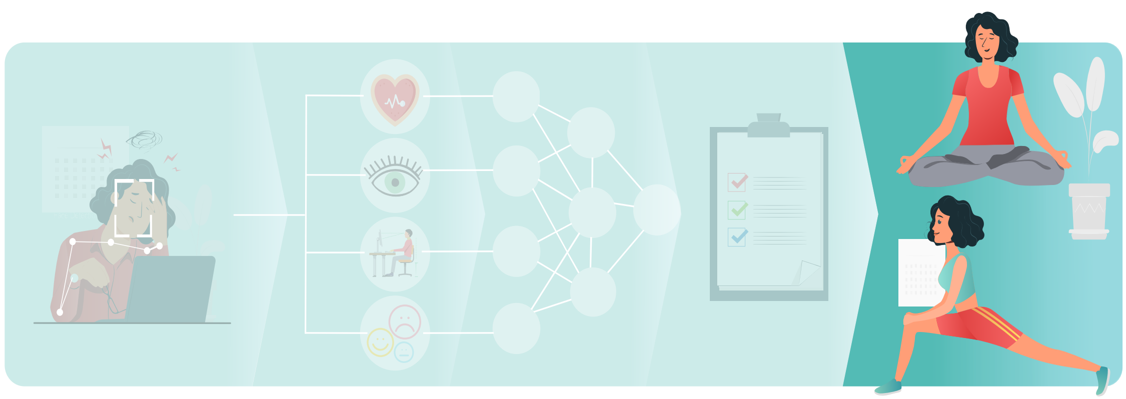


Figure 1: CareCam, a concept for measuring mental and physical health in the workplace using a simple webcam. In this paper we focus on the last part of the processing pipeline, the interventions.

ABSTRACT

High visual and cognitive demands characterize computer work. Therefore, preventive, health-promoting support at the workplace plays a central role in the competitiveness of companies. The evaluation of camera images, e.g. with a simple webcam, offers the possibility of gaining health-relevant data and using them for corporate health management. We, therefore, present the CareCam, a concept that can be used to record various data such as blink rate, pulse rate and upper body posture at the workplace. These data allow conclusions to be drawn about the health burden at the workplace and can be used for personalized interventions or health-promoting recommendations. However, how the collected data can best be used for interventions to minimize health risks at the VDU (Visual Display Unit) workplace in the long term has not yet been clarified. Therefore, this work outlines a basic concept for health interventions at the VDU workplace based on camera data. We utilize the facial expression, blinking rate, eye distance to screen and upper body posture to provide a targeted reminder and

PETRA '22, June 29-July 1, 2022, Corfu, Greece

ACM ISBN 978-1-4503-9631-8/22/06.
https://doi.org/10.1145/3529190.3534778

use their frequency as a marker to provide targeted user-tailored interventions at the workplace using a simple webcam.

CCS CONCEPTS

• **Human-centered computing** → *Ubiquitous and mobile computing theory, concepts and paradigms*; Empirical studies in HCI;

KEYWORDS

Remote health monitoring, Digital corporate health management, Stress detection, Imaging photoplethysmography, Pervasive computing

ACM Reference Format:

Dimitri Kraft, Angelina Schmidt, Lea Büttner, Frederike Marie Oschinsky, Fabienne Lambusch, Kristof van Laerhoven, Gerald Bieber, and Michael Fellmann. 2022. CareCam: Towards user-tailored Interventions at the Workplace using a Webcam. In *The 15th International Conference on PErvasive Technologies Related to Assistive Environments (PETRA '22), June 29-July 1, 2022, Corfu, Greece.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3529190.3534778

1 INTRODUCTION

Workplace interventions recently emerged as an occupational health strategy to improve work-related outcomes through structural, organizational changes, education-based interventions, physical activity programs, or other multi-component interventions. Especially sedentary employees (e.g., office workers) are prone to musculoskeletal disorders since most spend 3/4 of the day sitting. The lack of physical activity that occurs with excessive sitting has been shown to lead to various chronic diseases, as well as cardiovascular and metabolic diseases [3].

Office workers tend to interact with the computer about 7 hours a day, whether at home or work [10]. Working on the computer is characterized by high visual and cognitive demands. Health problems of employees due to excessive sitting, mental overload or bad posture during computer work result in reduced productivity, more extended absences and even early retirement. For the competitiveness of companies, preventive, health-promoting support at the workplace, therefore, plays a central role. Our approach to supporting office workers consists of contactless technology based solely on optical sensors (e.g., camera). The evaluation of camera images, e.g., using a simple webcam, either built-in or external, offers the possibility of health-relevant (e.g., heart rate, posture, eye blink frequency) data and use them for corporate health management and user-centered interventions to reduce mental and physical stress at the workplace. These crucial data allow conclusions to be drawn about the strain at the workplace and can be used for personalized interventions or health-promoting recommendations to increase general acceptance of interventions. Workplace related health programs are now widely accepted. They range from free fitness memberships to technical solutions, e.g., smartwatches, specialized depth cameras, inertial measurement units attached to the neck or other wearable solutions. Companies may also provide height-adjustable desks or ergonomic chairs to promote dynamic sitting during the day. However, these technical interventions are not necessarily solving the problem of excessive and static sitting since only 20% of the users use the function of adjusting the height of the table [13]. Further, it has not yet been determined how the data collected can best be used for interventions to minimize health risks at the workplace in the long term. Therefore, this work aims to outline a basic concept for individualized, user-tailored health interventions at the workplace based on this camera data. First, a field study will be conducted at the computer workstation, where test persons receive and evaluate these interventions to test the developed interventions. Subsequently, the test subjects will be interviewed about the experiment to make possible change requests or other comments and to evaluate the proposed interventions accordingly.

2 LITERATURE REVIEW

2.1 Camera-based monitoring systems for VDU workplaces

Camera-based health monitoring systems in VDU workplaces are still the subject of research activities today. 2011 Mary et al. [6] already recognized the need for monitoring and based assessment of health problems among IT professionals as they achieve exceptionally high screen time. Common computer-related health problems include visual problems from eyestrain and musculoskeletal problems such as back pain, wrist pain, and muscle fatigue. Stress and headaches can also be triggered. Monitoring with health risk detections offers the potential to take appropriate action at the early stages. The authors, therefore, propose an automatic system that focuses on avoiding stressful postures. For this purpose, gestures are recognized from body movements or states originating from the eyes, the neck or the hand. To enable continuous monitoring, webcam recordings will be used to capture these gestures from employees and then processed using several different techniques: The video signals are converted into single frames and processed by foreground segmentation. As a result, the features relevant to the analysis (e.g., speed, position and orientation) can be extracted from the image. After that, gestures, such as wrist position, are estimated. The recognized gestures are then compared with an ideal: possible ideal postures are stored in a system table. If there is no match in the template matching, a warning is attracted in the computer system. If there is a discrepancy, a warning is sent to the user. Other systems have been developed that use a webcam, external cameras and additional sensors to record data to monitor stress or other health issues for office workers.

In [5], the concept of *COSMS* is outlined. *COSMOS* captures physiological information about the standard equipment of a VDU workstation: the camera measures vital signs such as heart rate, facial expression and eye blinking. In addition, a mouse with embedded sensors is used to measure temperature and humidity as potential environmental stresses. Mouse movements and keystroke volume are taken into account in a complementary manner. The collected data is stored in the InfluxDB time-series database and displayed using a visualization application. The system collects this information continuously without disturbing the user while working and without attaching sensors to the user. In the future, the collected data could be analyzed to find correlations between different work activities and mental health and set appropriate triggers in the visualization application. No health interventions have been implemented yet.

Chen et al. [1] use the webcam and additional cameras at the workplace to provide feedback on the ergonomic state of the worker. For this purpose, parameters such as the average working and break times, the distance to the screen, head movements, gaze directions, and blinking frequencies are recorded. Furthermore, the additional cameras installed at the workplace provide information on general posture (standing vs sitting) and social interactions with others. The user can thus be provided with statistics for self-reflection, which an ergonomics expert can also evaluate. The data includes information on activity levels, posture, and eye fatigue. By manually entering the user's schedule, contextual information can also be included. This allows the user to determine which activities are associated with unhealthy habits. After analyzing the data, personalized recommendations can then be made. The system learns the user's behaviour pattern and provides ergonomic reminders. This probabilistic model provides semantic interpretations such as *close to the screen, low head mobility* or *absent*. Thus, personalized reminders can be sent to the user. The system is compared to rule-based reminders, as planned for this work. For this purpose, users classify reminders into *useful* and *disruptive*. The proposed method outperforms rule-based reminders by sending reminders tailored to the user's behaviour pattern and personal schedule. As a result, more reminders are classified as useful.

Vildjiounaite et al. [12] propose in their work a person-specific stress monitoring system that captures users' motion trajectories using depth cameras in the office. A method based on discrete hidden Markov models is used for the analysis. The system correctly identified users' most stressful work periods in the study.

Paliyawan et al. [9] focus their system on detecting prolonged sitting of office workers. A Kinect camera is set up to provide integrated skeletal tracking. A systematization in the form of a point system with three danger levels was introduced to classify the health risk. This classification is based on a traffic light: green symbolizes a healthy state, yellow a warning state and red an unhealthy state. Points are added or subtracted depending on time spent sitting and moving. The number of points is then used to determine the level of danger, and appropriate warnings are sent to the user as real-time feedback. A visualization to create a daily summary for reflection is also available [9]. The system also includes features to detect unhealthy sitting postures. Mathematics is applied to convert postures to body angles, and a threshold model detects posture changes. This allows the detection of head tilts and contorted body positions. The thresholds for these angles come from the *RULA* and *REBA* procedures. To begin, the user must define a baseline body posture: a healthy posture from which the relative changes in body postures can be calculated [8]. To make ergonomic interventions effective over the long term, a continuous process that includes frequent feedback is necessary. In [11] the authors have confirmed in a study the effectiveness of continuous feedback using webcam photos of posture at work to reduce musculoskeletal risk among workers at VDT workstations. A photo-training group received traditional office training plus an automated feedback system that displayed a photo of current sitting posture and a photo of correct posture. The risk for musculoskeletal disorders was assessed using the *RULA* method. The training method achieved short-term improvements in posture and sustained improvements, as the application of this method can be used over a more extended period, unlike office training alone. Further, the method achieved better results in women, the elderly, and those with more pain. A combination of complementary feedback for different target groups should therefore be considered.

3 CONCEPTION AND METHOD CARECAM

The related work shows that objective data, long-term data over weeks or months, are necessary for the user's self-reflection. A non-contact monitoring tool is essential to ensure that a measuring procedure does not influence the measurement itself. For a comprehensive analysis of the user's condition during this daily work, we need the evaluation of the following parameter:

- Heart rate and heart rate variability
- Blinking frequency and blinking pattern
- Posture
- Time spend without movement
- Distance to camera
- Facial expression
- Respiration

The CareCam is a tool that measures these parameters using only a standard webcam and aggregates the parameters into a stress level by applying artificial intelligence. The aggregated parameters are control variables for a control circuit and the use of interventions. To give an example, as illustrated in Figure 2: If the heart rate is higher than desired, the intervention (e.g., deep breathing) will affect the vital data and lower the heart rate. The CareCam measures the new heart rate and compares it to the control variable again. This mechanism ensures that we track the success of an intervention and may also apply for eye-blinking rate and breathing rate. The same approach may enable the integration of biofeedback techniques as well. In addition, short questionnaires at the end of an intervention enable additional subjective feedback, further tailoring the interventions to the user.

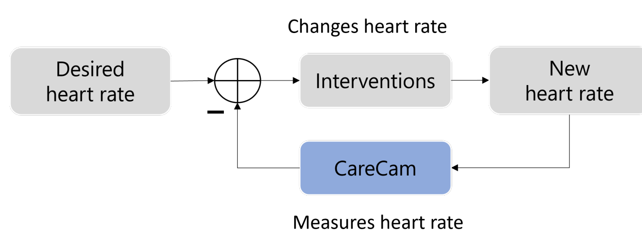


Figure 2: Control Circuit of CareCam, example of heat rate influencing by interventions.

The overall concept of the CareCam has already been outlined in [4]. We use various computer vision concepts to extract vital signs from simple RGB images (Figure 3), which are used to provide helpful reminders and interventions to the user during the workday (Figure 4). The interventions and reminders are tailored to the user as we measure the vital signs in real-time and adjust them based on the current situation at work.

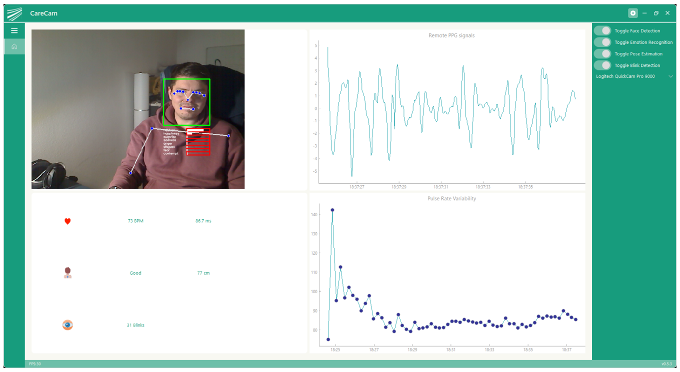


Figure 3: Snapshot of the CareCam user interface. The upper left panel shows the current camera image with the detected face, pose markers and facial expression detection. The upper right panel shows the pulse, which contains some artefacts due to poor lighting. The lower left panel shows aggregated information (e.g., pulse rate, pose classification, and distance to the camera). The lower right panel currently shows the variability of the pulse rate.

4 INTERVENTIONS

4.1 Reminder

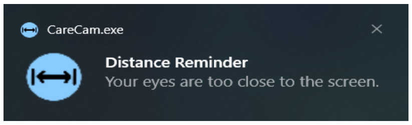
Reminders are short messages that should not interrupt the workflow. They, therefore, contain only a short headline and a short descriptive sentence and appear as a pop-up notification via the operating system. In addition, the reminders contain various icons so that the messages have a recognition value and can thus be grasped more quickly (Figure 4). Table 1 shows all implemented reminders. When the condition for a reminder is met, the triggers are used to decide whether to display a reminder. This notification is sent directly to the user via the operating system. When a reminder is triggered, it is stored in a buffer: only after five minutes have elapsed can this reminder lead to a notification again by the trigger. This prevents a reminder from appearing too frequently, thus interrupting and disturbing the user in his work. The decisions for the reminders and their triggers are based on the literature listed earlier. The blink reminder is intended to help the user maintain a blink frequency high enough to prevent tearing the tear film even during periods of concentration on the screen. Since the blink frequency is subject to many fluctuations and influencing factors, the limit value of 12 blinks per minute was selected and according to [2] the normal spontaneous blink rate is between 12 and 15 blinks per minute. For a healthy posture, reducing the duration of sitting by incorporating dynamic sitting techniques is essential. Since the webcam only records the upper body posture, it is not so easy to distinguish between sitting and standing activities. For this reason, a reminder for dynamic sitting was chosen: The sitting position should be changed frequently. For this purpose, the distance of the posture points recorded by the CareCam is measured. If this distance changes by 12 pixels, this is classified as a movement. If there is no movement after 300 captured images, the reminder is triggered. This corresponds to approximately two minutes (pose estimation is not triggered in each frame). The distance of the eyes to the screen depends on the size of the screen or font size, depending on the activity. Since 50 cm emerged as the smallest value in the recommended distances of [7], a minimum distance of 50 cm was simplified for this reminder. It was decided not to query the screen size for the time being since the viewing distance may be smaller for reading tasks. Determining the currently performed activity to find out whether the user is trying to capture the entire screen or is performing a reading activity was considered too complex within this work. Maintaining the distance to the screen is important, as this has an influence on the posture as well as the eyes. An attempt to counteract negative emotions is implemented through motivational messages. These messages are also sent as short reminders in the operating system's notifications: Once per second, the emotional state is saved. After one minute, when 60 emotion states have been stored, the predominant emotion is determined. This is the emotional state that has appeared most frequently during this time. The reminder is triggered if the facial expression is predominantly characterized by a negative emotion (sadness or fear). This randomly outputs different motivational messages. No Reminder depends on pulse rate or pulse rate variability. Although these metrics offer the potential for health interventions, particularly for detecting stress, it was decided not to use them within this work. Pulse and pulse rate variability are highly individual and complex, as they depend on many factors. Pulse rate variability is also subject to significant inter-day variations. In addition, the system is tested in a natural working environment: The qualitative measurement of the pulse values depends heavily on the right lighting conditions, a good webcam and the movements of the test subjects. Accordingly, meaningful use of these measured values requires a more complex system, including a calibration phase and outlier management.

Table 1: Overview of reminder with corresponding trigger description and required function.

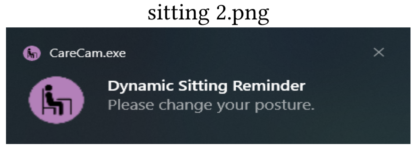
Reminder	Trigger	Condition
Blinking	Blinking rate <12 bpm	Blink detection
Dynamic sitting	No movement within 2 minutes	Pose estimation
Distance to camera	Distance to camera less than 50 cm for 30 seconds	Blink detection
Motivational message	Predominantly negative facial expression within one minute	Facial expression recognition

4.2 Breaks

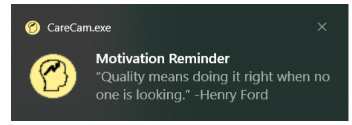
Break interventions can be divided into three groups: *eye exercises, physical breaks* and *mindfulness*. To decide which intervention group has priority during the break, the reminders are counted during the 55-minute work period. The assignment can be seen in Table 2 2: since blink frequency and maintaining distance from the screen have an impact on eye health, the counter for this intervention group is increased by one each time these reminders are sitting 2.png



(a) Reminder: Distance to camera



(b) Reminder: Dynamic sitting



(c) Reminder: Motivating message

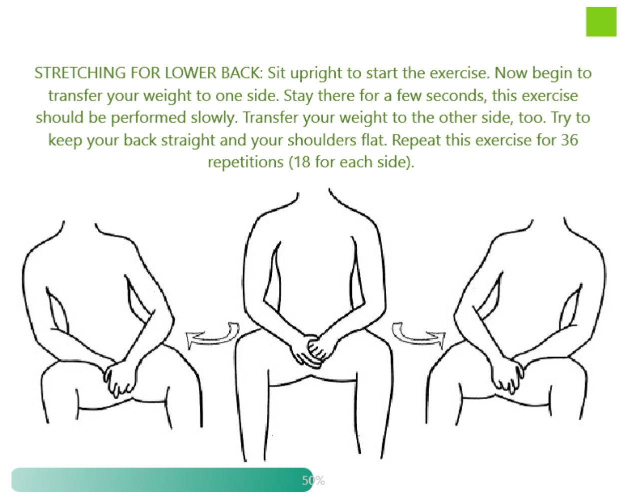
Figure 4: Example for reminders (Windows 10). The reminder are implemented as operating system notifications.

triggered. Furthermore, since a short distance to the screen may involve forward bending of the head and thus favour the development of a text Neck, the counter for the intervention group of physical breaks (i.e. posture-related interventions) is also increased by one. Additionally, reminders to sit dynamically will increase the numerator for this intervention group by one, as static muscle activity can lead to tension. Finally, the motivational message reminder increases the mindfulness intervention group. Furthermore, the counter increases by one when the user has completed two work phases. This was chosen to include high work periods. The exercises are to be performed within the scheduled five-minute break. The performance of the exercises does not fill the five minutes: This allows that the break time does not have to be exceeded, and the user, on the one hand, has enough time to read through the intervention accordingly and, on the other hand, can use the rest of the break individually, for example, to get a glass of water, exchange with colleagues or have an additional mental break without a task. After an intervention group is triggered, the counter for that group is reset. The counter of the other groups remains the same. Thus, compensation occurs since the intervention supported the group and the associated health aspect. An intervention buffer manages the interventions' frequency and occurrence, preventing repeating interventions and thus reducing monotony. When the software is restarted - for example, on the next working day - the counters for all groups are back to zero.

Table 2: Overview of active breaks with corresponding trigger description and required function.

Intervention	Trigger	Condition
Relax eyes, Change viewpoint	Blink reminder, distance from screen	Blink detection
Standing pause, stretching and relaxation	Sitting position, distance from screen	Pose estimation
Deep breathing, meditation, mental break	Motivational messages every two work periods	Face expression recognition

Examples for active breaks are depicted in Figure 5. These interventions are designed to take around two minutes and may be followed within a five minute break.



(a) Active break: Lower back stretching

Figure 5: The instructions for active pauses are displayed as an additional widget consisting of a progress bar and a small icon for closing the intervention.

5 LIMITATIONS

Our multimodal approach consists of several components (e.g., pulse rate estimation, facial expression recognition), each of which works independently and supports during the workday. However, these components are themselves subject to several limitations (e.g., lighting, motion, camera noise, eyeglasses) that reduce their accuracy and thus the credibility of reminders and interventions. Therefore, an intelligent mechanism must be created that assigns a credibility value to each measurement and only measures and intervenes when the circumstances are right (ideal lighting conditions, user movement) to achieve sufficient accuracy. A balance between ideal conditions, measurement time and accuracy, should be sought. In addition, the user must be informed why these interventions and reminders are occurring and why they are essential to enhance adherence to the software.

6 CONCLUSION

In this paper, we have shown that a simple sensor, a webcam, is capable of detecting almost any internal state of the user in the work environment and providing interventions tailored to the user and based on the context. We hypothesize that long-term monitoring and analysis of the user can lead to a reduced physiological f̈ootprintöf high work demands leading to healthier work. Also, the system could be used to establish healthier working habits both in terms of physiological and mental habits and thus help to increase the resilience and overall long-term health of employees. This should increase happiness and keep us all healthier. We also provided an overview of camera-based monitoring systems in the workplace and adapted our approach to interventions based on various workplace safety guidelines. Finally, we outlined how and what interventions (reminders and active breaks) can be integrated during work and under what conditions these interventions should occur. A qualitative study highlighting the benefits of these interventions is still in progress. A field study will be conducted at the VDU workplace, where subjects will receive and evaluate these interventions to test the intervention and reminder concept. Subsequently, the subjects will be interviewed about the experiment to obtain any requests for changes or other comments and evaluate the proposed interventions accordingly.

REFERENCES

[1] C. Chen, T. Määttä, Kevin Bing-Yung Wong, and H. Aghajan. 2012. A collaborative framework for ergonomic feedback using smart cameras. In *2012 Sixth International Conference on Distributed Smart Cameras (ICDSC)*. 1–6.

[2] Michael J Doughty. 2001. Consideration of three types of spontaneous eyeblink activity in normal humans: during reading and video display terminal use, in primary gaze, and while in conversation. *Optometry and Vision Science* 78, 10 (2001), 712–725.

[3] Genevieve N. Healy, Katrien Wijndaele, David W. Dunstan, Jonathan E. Shaw, Jo Salmon, Paul Z. Zimmet, and Neville Owen. 2008. Objectively measured sedentary time physical activity and metabolic risk: the Australian Diabetes Obesity and Lifestyle Study (AusDiab). *Diabetes care* 31, 2 (2008), 369–371.

[4] Dimitri Kraft, Kristof van Laerhofen, and Gerald Bieber. 2021. CareCam: Concept of a new tool for Corporate Health Management. In *The 14th PErvasive Technologies Related to Assistive Environments Conference*. 585–593.

[5] Naoki Maeda, Yuko Hirabe, Yutaka Arakawa, and Keiichi Yasumoto. 2016. COSMS: Unconscious stress monitoring system for office worker. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*, Paul Lukowicz, Antonio Krüger, Andreas Bulling, Youn-Kyung Lim, and Shwetak N. Patel (Eds.). ACM, New York, NY, USA, 329–332.

[6] G. Maria Justin Mary and P. Munipriya. 2011. Health monitoring of IT industry people. In *2011 3rd International Conference on Electronics Computer Technology*. IEEE, 144–148. https://doi.org/10.1109/ICECTECH.2011.5941875

[7] Melvin Mohokum and Julia Dördelmann. 2018. Ergonomie am Arbeitsplatz mit Praxisbeispielen. In *Betriebliche Gesundheitsförderung*, Melvin Mohokum and Julia Dördelmann (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 175–213.

[8] Pujana Paliyawan, Chakarida Nukoolkit, and Pornchai Mongkolnam. 10/1/2014 - 10/3/2014. Office workers syndrome monitoring using kinect. In *The 20th Asia-Pacific Conference on Communication (APCC2014)*. IEEE, 58–63. https://doi.org/10.1109/APCC.2014.7091605

[9] Pujana Paliyawan, Chakarida Nukoolkit, and Pornchai Mongkolnam. 5/14/2014 - 5/17/2014. Prolonged sitting detection for office workers syndrome prevention using kinect. In *2014 11th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*. IEEE, 1–6.

[10] Susan A. Randolph. 2017. Computer Vision Syndrome. *Workplace Health & Safety* 65, 7 (2017), 328–328. https://doi.org/10.1177/2165079917712727 arXiv:https://doi.org/10.1177/2165079917712727 PMID: 28628753.

[11] Meirav Taieb-Maimon, Julie Cwikel, Bracha Shapira, and Ido Orenstein. 2012. The effectiveness of a training method using self-modeling webcam photos for reducing musculoskeletal risk among office workers using computers. *Applied Ergonomics* 43, 2 (2012), 376–385.

[12] Elena Vildjiounaite, Ville Huotari, Johanna Kallio, Vesa Kyllönen, Satu-Marja Mäkelä, and Georgy Gimel'farb. 2018. Detection of Prolonged Stress in Smart Office. In *Intelligent Computing*, Kohei Arai (Ed.). Advances in Intelligent Systems and Computing Ser, Vol. 857. Springer, Cham, 1253–1261. https://doi.org/10.1007/978-3-030-01177-2{_}90

[13] Stephen Wilks, Monica Mortimer, and Per Nylén. 2006. The introduction of sit–stand worktables; aspects of attitudes compliance and satisfaction. *Applied Ergonomics* 37, 3 (2006), 359–365.

ACKNOWLEDGMENTS

We thank our students and study participants for their support and patience.