Guest Editorial Camera-Based Monitoring for Pervasive Healthcare Informatics

MEASURING physiological signals from the human face and body using video cameras is an emerging research topic that has grown rapidly in the last decade. Remote cameras (in both visible and infrared wavelengths) can be used to measure vital signs from a human body based on skin optics or body movements thereby avoiding mechanical contact with the skin. Recent advances in machine vision, as well as sensor technology and AI, have acted as catalysts for camera-based physiological measurement, significantly improving its robustness in real applications with the practical challenges of patient monitoring. The current measurements include heart rate (including variability), respiration rate, blood oxygenation saturation, blood pressure, and skin temperature, but can be extended to other physiological markers/indicators that have diagnostic capabilities, such as the detection of arrhythmia, atrial fibrillation, apnea, breathing cessation, hypertension, fever, etc.

Imaging methods for recovering vital signs also present new opportunities for machine vision applications that require better understanding of human physiology (e.g., affective computing). AI-enabled machine vision technologies, in turn, provide the analysis of high-level image/video semantics, facilitating the understanding of human behaviors and scene context for health monitoring, such as facial expression analysis for detecting pain in the newborn or delirium for a patient in the ICU, emotion recognition for depression analysis, sleep posture detection for sleep quality assessment, body motion tracking for bed exit/fall detection and patient actigraphy, activity recognition for clinical workflow monitoring and optimization (e.g., hygiene compliance), etc. Context monitoring and understanding are unique advantages delivered by camera sensors as compared to contact-based biomedical sensors.

Camera-based health monitoring will bring a rich set of compelling healthcare applications that directly improve upon contact-based monitoring solutions and impact people's care experience and quality of life in various scenarios, such as in hospital care units, sleep/senior centers, assisted-living homes, telemedicine and e-health, baby/elderly care at home, fitness and sports, driver monitoring in automotive applications, cardiac/respiratory gating for MRI/CT, AR/VR based therapy and clinical training, etc. It is unquestionable that camera-based health monitoring will evolve into a key technology in healthcare, advancing medical diagnosis and prognosis, patient care and treatment, and management of well-being and chronic diseases.

This special issue of the IEEE Journal of Biomedical and Health Informatics journal is meant to present and highlight some of the latest developments on applying camera-based techniques for pervasive healthcare informatics. It has attracted a great number of submissions from researchers both from academia and from industry. After critical peer-review and selection, 13 manuscripts were accepted for publication in this special issue, covering a variety of topics in the category of physiological measurement and contextual measurement with clear application goals in healthcare.

I. Camera Based Physiological Measurement for Healthcare

Five articles presented in this special issue focus on camera-based physiological measurement, applying novel signal- and image-processing techniques to improve the robustness of vital-sign measurement and having the systems validated in real-world use cases such as automotive and clinical practice.

A general video-processing framework for camera-based pulse-rate measurement includes two common parts: front-ends that pre-process images or post-process measured signals, and the core algorithm that extracts the photoplethysmography (PPG) signal from the optical signals. *Woyczyk et al.* [item 1) in the Appendix] aimed at improving the pre-processing step in a pulse-rate monitoring framework and presented a method to segment and track a suitable skin "Region of Interest" (RoI) to optimize the pulse measurement. The proposed method yields a time-varying RoI constructed from multiple homogeneous skin areas under morphological constraints of the RoI. It can be integrated into any pulse-extraction framework as an image pre-processing step, and its effectiveness was demonstrated by the achievement of 2nd place in an open-benchmark competition. *Song et al.* [item 2) in the Appendix] focused on the post-processing step and introduced a "Generative Adversarial Network"-based method to refine the pulse signal measured from a video, such that the refined pulse signal resembles the characteristics of cardiac waveforms of contact-based PPG. The improved pulse waveform is useful for calculating more accurate inter-beat interval features, which also enables additional measurements beyond pulse rate.

In addition to the measurement of pulse rate, *Aziz et al.* [item 3) in the Appendix] measured blood hemoglobin levels based on pulse signals generated from fingertip videos captured by a

Digital Object Identifier 10.1109/JBHI.2021.3072439

smartphone camera. Such a ubiquitous and portable setup will make the hemoglobin measurement widely accessible and affordable to a large population. The proposed solution is intended for spot-check of hemoglobin levels in a short interval (e.g., few seconds) instead of long-term continuous monitoring. The large-scale experiments on 167 patients show a strong linear correlation between the proposal and reference, with a mean absolute percentage error of 5%.

Considering the applications of vital-sign monitoring, *Huang et al.* [item 4) in the Appendix] took camera-based pulse-rate monitoring into the automotive context for driver monitoring. To address the challenges that are specific to automotive such as driver/vehicle motions and significant lighting changes, the authors proposed a novel algorithm called "adaptive spectral filter banks" to strike the balance between robustness and sensitivity of pulse-rate tracking, and had it rigorously tested in various automotive scenarios. The results show that the measured pulse-rates have improved robustness to disturbances without sacrificing responsiveness to spontaneous changes in pulse rate, which is crucial for driver monitoring. *van Gastel et al.* [item 5) in the Appendix] built a near-infrared camera setup for sleep monitoring and investigated its potential of replacing polysomnography for the diagnosis of sleep disorders (e.g., apnea). Three types of vital signs (pulse rate, respiration rate and blood oxygen saturation) were measured in a clinical trial involving patients with sleep apnea. Though the study did not establish a direct relationship between the camera-measured vital signs and sleep disorders, it reported that the camera can detect pulse rate with a Mean Absolute Error (MAE) of 2 beats/min for 92% of the time and respiratory rate with a MAE of 2 breaths/minute for 91% of the time. The blood oxygen saturation estimates (SpO_2) are within 4% of the values from the finger-oximeter for 89% of the time.

II. Camera Based Context Measurement for Healthcare

Eight articles in the special issue are related to camera-based context measurement. *Weber et al.* [item 6) in the Appendix] used an infrared camera to detect the presence of a preterm newborn in the incubator and open-bed scenarios. It leverages the deep transfer learning technique to detect the presence of infants, which essentially exploits the function of context monitoring of a surveillance camera. Following presence detection, *Cheng et al.* [item 7) in the Appendix] proposed a video-based infant monitoring system that analyzes infant facial expressions (discomfort, unhappiness, joy and neutral state) and status (sleep, pacifier and open mouth), and further correlated the discomfort event with a symptom-related disease such as gastroesophageal reflux. The system utilized a fast R-CNN and Hidden Markov Model to analyze the infant's facial expressions. Average precision for discomfort detection reaches 90%. Based on both RGB and thermal video signals, *Awis et al.* [item 8) in the Appendix] employed a combination of a deep convolutional neural network with a SVM classifier for neonatal sleep-wake classification. In a study on 19 Chinese newborns, the authors demonstrated high classification accuracy (93.8%), a performance superior to that in 9 other cited studies.

In addition to the monitoring of infants, *Rezaei et al.* [item 9) in the Appendix] applied video-based pain monitoring to elderly individuals with dementia, as this population is often undertreated for pain due to their inability to inform carers of their pain levels. The authors presented a deep learning based solution that detects facial expressions associated with pain, and validated their solution on a dataset acquired from elderly individuals with and without dementia (including human annotated ground-truth). The introduced pairwise comparative inference method and fast contrastive training method improved the sensitivity of pain measurement for individuals, which shows clear improvements over the cited baseline methods, typically used with elderly individuals with dementia.

Opportunities to use wearable cameras embedded in glasses for healthcare were also explored. *Dousty et al.* [item 10) in the Appendix] proposed a novel method to detect tenodesis grasp based on egocentric videos of individuals, which allows clinicians to monitor the rehabilitation outcomes of a patient suffering from cervical spinal cord injury and give remote therapeutic guidance. The detection of tenodesis grasp is via the analysis of wrist angle data in egocentric videos by hand detection, pose and arm orientation estimation. *Qiu et al.* [item 11) in the Appendix] also used egocentric videos to assess the dietary intake of an individual, called "passive dietary monitoring", which is considered to be more objective than self-reports in epidemiological studies. The authors developed a deep neural network based method that estimates the individual dietary intake from egocentric videos by recognizing the consumed food items and counting the number of bites taken by the person.

Beyond continuous monitoring, cameras have also been used as optical sensors to screen skin patterns for diagnosing malignant skin tumors. *Jiang et al.* [item 12) in the Appendix] presented a light-weight framework based on the attention mechanism of deep learning to differentiate between 11 types of skin diseases in histopathological images. The method was validated on a real histopathological image dataset collected by the authors in the last 10 years, and shows significant improvements over baseline methods that use conventional deep learning models. The proposed system generates both the classified disease label and a diagnostic report that visualizes possible skin areas affected by the disease, helpful for medical diagnosis.

The pivotal role of cameras in healthcare is not limited to that of a diagnostic tool. It can further be used for clinical training and education. *Guo et al.* [item 13) in the Appendix] prototyped an ultrasound-guided surgery training system focused on renal biopsy. The system was built on AR technology, whereby a camera is used to create an immersive environment including virtual annotations and real surgical scenes for optimizing the operation (e.g., puncture path planning and puncture training assistance). The AR-based system was benchmarked against a standard VR-based system used in renal puncture surgery training, and the AR-based option was shown to be more suitable for the task, with improved accuracy and reliability of surgical training.

All 13 papers tackle different but relevant domain vectors of Camera-based Monitoring for Pervasive Healthcare Informatics. We believe this special issue will raise awareness in both scientific and industrial communities that video cameras are ubiquitous and valuable sensors for pervasive healthcare, with a multidisciplinary research path with joint endeavours from different fields (sensors, machine vision, biomedical informatics, healthcare, and AI) to create the momentum necessary to build on the promise of this field.

Acknowledgment

The guest editors have many people to thank for their contributions to this special issue. We would like to thank all the authors for their high-quality contributions, and all the reviewers for their professional service. Also, we would like to thank the editorial team of IEEE J-BHI. We hope you will enjoy this collection of papers and that the special issue can spark and stimulate further research and development in this area.

Wenjin Wang
Eindhoven University of Technology
5600 MB, The Netherlands
wenjin.wang@philips.com

Steffen Leonhardt
RWTH Aachen University
D-52074, Germany
leonhardt@hia.rwth-aachen.de

Lionel Tarassenko
University of Oxford
OX1 3PJ, U.K.
lionel.tarassenko@eng.ox.ac.uk

Caifeng Shan
Shandong University of Science
and Technology
266590, China
caifeng.shan@gmail.com

Daniel McDuff
Microsoft Research
WA 98052, USA
damcduff@microsoft.com

Appendix
Related Work

1) A. Woyczyk, V. Fleischhauer, and S. Zaunseder, "Adaptive gaussian mixture models driven level set segmentation for remote pulse rate detection," *IEEE J. Biomed. Health Inform.*, to be published, doi: 10.1109/JBHI.2021.3054779.
2) R. Song, H. Chen, J. Cheng, C. Li, Y. Liu, and X. Chen, "PulseGAN: Learning to generate realistic pulse waveforms in remote photoplethysmography," *IEEE J. Biomed. Health Inform.*, 2021. doi: 10.1109/JBHI.2021.3051176.
3) M. H. Aziz, M. K. Hasan, A. Mahmood, R. R. Love, and S. I. Ahamed, "Automated cardiac pulse cycle analysis from photoplethysmogram (PPG) signals generated from fingertip videos captured using a smartphone to measure blood hemoglobin levels," *IEEE J. Biomed. Health Inform.*, to be published, doi: 10.1109/JBHI.2021.3068658.
4) P. Huang, B. Wu, and B. Wu, "A heart rate monitoring framework for real-world drivers using remote photoplethysmography," *IEEE J. Biomed. Health Inform.*, to be published, doi: 10.1109/JBHI.2020.3026481.
5) M. van Gastel, S. Stuijk, S. Overeem, J. P. van Dijk, M. M. van Gilst, and G. de Haan, "Camera-based vital signs monitoring during sleep – A proof of concept study," *IEEE J. Biomed. Health Inform.*, to be published, doi: 10.1109/JBHI.2020.3045859.
6) R. Weber, S. Cabon, A. Simon, F. Poree, and G. Carrault, "Preterm newborn presence detection in incubator and open bed using deep transfer learning," *IEEE J. Biomed. Health Inform.*, to be published, doi: 10.1109/JBHI.2021.3062617.
7) C. Li, A. Pourtaherian, L. van Onzenoort, W. E. Tjon a Ten, and P. H. N. de With, "Infant facial expression analysis: Towards a Real-time video monitoring system using R-CNN and HMM," *IEEE J. Biomed. Health Inform.*, to be published, doi: 10.1109/JBHI.2020.3037031.
8) M. Awais et al., "A hybrid DCNN-SVM model for classifying neonatal sleep and wake states based on facial expressions in video," *IEEE J. Biomed. Health Inform.*, 2021. doi: 10.1109/JBHI.2021.3073632.
9) S. Rezaei1, A. Moturu, S. Zhao, K. M. Prkachin, T. Hadjistavropoulos, and B. Taati, "Unobtrusive pain monitoring in older adults with dementia using pairwise and contrastive training," *IEEE J. Biomed. Health Inform.*, to be published, doi: 10.1109/JBHI.2020.3045743.
10) M. Dousty and J. Zariffa, "Tenodesis grasp detection in egocentric video," *IEEE J. Biomed. Health Inform.*, to be published, doi: 10.1109/JBHI.2020.3003643.
11) J. Qiu, F. P.-W. Lo, S. Jiang, Y. Tsai, Y. Sun, and B. Lo, "Counting bites and recognizing consumed food from videos for passive dietary monitoring," *IEEE J. Biomed. Health Inform.*, to be published, doi: 10.1109/JBHI.2020.3022815.
12) S. Jiang, H. Li, and Z. Jin, "A visually interpretable deep learning framework for histopathological Image-based skin cancer diagnosis," *IEEE J. Biomed. Health Inform.*, 2021. doi: 10.1109/JBHI.2021.3052044.
13) Z. Guo, Y. Tai, J. Du, Z. Chen, Q. Li, and J. Shi, "Automatically addressing system for ultrasound-guided renal biopsy training based on augmented reality," *IEEE J. Biomed. Health Inform.*, doi: 10.1109/JBHI.2021.3064308.