

# A Broader Look: Camera-Based Vital Sign Estimation across the Spectrum

Christoph Hoog Antink\*, Simon Lyra\*, Michael Paul, Xinchu Yu, Steffen Leonhardt  
 Medical Information Technology (MedIT), Helmholtz-Institute for Biomedical Engineering,  
 RWTH Aachen University, Aachen, Germany

## Summary

**Objectives:** Camera-based vital sign estimation allows the contactless assessment of important physiological parameters. Seminal contributions were made in the 1930s, 1980s, and 2000s, and the speed of development seems ever increasing. In this survey, we aim to overview the most recent works in this area, describe their common features as well as shortcomings, and highlight interesting “outliers”.

**Methods:** We performed a comprehensive literature research and quantitative analysis of papers published between 2016 and 2018. Quantitative information about the number of subjects, studies with healthy volunteers vs. pathological conditions, public datasets, laboratory vs. real-world works, types of camera, usage of machine learning, and spectral properties of data was extracted. Moreover, a qualitative analysis of illumination used and recent advantages in terms of algorithmic developments was also performed.

**Results:** Since 2016, 116 papers were published on camera-based vital sign estimation and 59% of papers presented

results on 20 or fewer subjects. While the average number of participants increased from 15.7 in 2016 to 22.9 in 2018, the vast majority of papers ( $n=100$ ) were on healthy subjects. Four public datasets were used in 10 publications. We found 27 papers whose application scenario could be considered a real-world use case, such as monitoring during exercise or driving. These include 16 papers that dealt with non-healthy subjects. The majority of papers ( $n=61$ ) presented results based on visual, red-green-blue (RGB) information, followed by RGB combined with other parts of the electromagnetic spectrum ( $n=18$ ), and thermography only ( $n=12$ ), while other works ( $n=25$ ) used other mono- or polychromatic non-RGB data. Surprisingly, a minority of publications ( $n=39$ ) made use of consumer-grade equipment. Lighting conditions were primarily uncontrolled or ambient. While some works focused on specialized aspects such as the removal of vital sign information from video streams to protect privacy or the influence of video compression, most algorithmic developments were related to three areas: region of interest selection, tracking, or extraction of a one-dimensional signal. Seven papers used

deep learning techniques, 17 papers used other machine learning approaches, and 92 made no explicit use of machine learning.

**Conclusion:** Although some general trends and frequent shortcomings are obvious, the spectrum of publications related to camera-based vital sign estimation is broad. While many creative solutions and unique approaches exist, the lack of standardization hinders comparability of these techniques and of their performance. We believe that sharing algorithms and/or datasets will alleviate this and would allow the application of newer techniques such as deep learning.

## Keywords

Photoplethysmography imaging; photoplethysmography; thermography; camera-based sensing; vital signs

Yearb Med Inform 2019:102-14

<http://dx.doi.org/10.1055/s-0039-1677914>

## 1 Introduction and History

Vision is the most developed sense in humans, and the majority of interactions with our surroundings is based on visual information. Thus, it comes to no surprise that extensive engineering efforts were invested into the development of technologies related to “(tele)vision”. In the beginning of the 20th century, video camera tubes emerged that allowed the acquisition and, in combination with appropriate means of transmission and reproduction, the real-time display of visual information at separate locations. By

the 1980s, solid-state technologies began to replace video camera tubes, first in the form of so-called charge coupled devices (CCDs), later with complementary metal-oxide-semiconductor (CMOS) sensors. The advent of these technologies significantly helped to transform “(tele)vision” (in the most general sense) from a mainly entertainment-oriented technology to a widely available, quantitative tool for science and medicine [1, 2].

The first scientific field to make major use of digital images acquired with solid-state sensors was astronomy, where CCDs replaced photographic plates in telescopes [2]. In the medical world, digital images and their processing began to play an increasingly important role. This is most obvious in non-optical

imaging modalities such as computed tomography, electrical impedance tomography, or electron microscopy, which are virtually and literally unimaginable without digital image processing. However, even classical optical modalities such as light microscopy were significantly influenced by the advent of digital image acquisition and processing, as well as the increased availability of computers and appropriate software [1].

As computing power advanced further, it became feasible to transit from the static to the dynamic, i.e. from digital image to digital video processing. While this transition has had many consequences in several areas such as surveillance, industrial processes, or navigation, it opened up a completely new path,

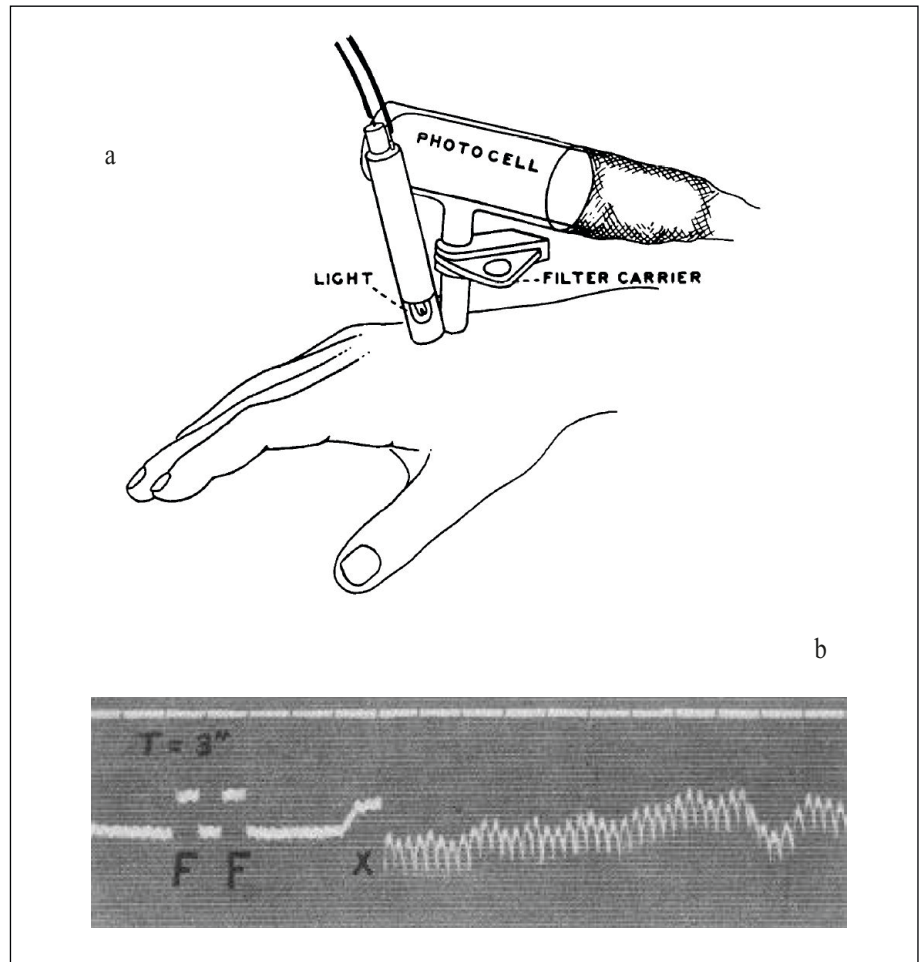
\* Both authors contributed equally

the unobtrusive monitoring of vital signs, usually defined as blood pressure (BP), heart rate (HR), breathing rate (BR), and body temperature [3]. Their estimation dates back to 1625, when Santorio published methods to time the pulse with a pendulum and measure body temperature with a spirit thermometer. However, only after rediscovery in the 18<sup>th</sup> and 19<sup>th</sup> centuries and subsequent improvements of watch and thermometer, HR, BR, and body temperature became standard vital signs. With the introduction of the sphygmomanometer by Riva-Rocci in 1896, BP joined as the fourth vital sign [4].

With the exception of the slow-varying body (core) temperature, vital signs are dynamic quantities. In particular heart rate and breathing rate are, by definition, frequencies, which lie in the range of about 6 to 240 events per minute, i.e. 0.1 to 4 Hz. According to the Nyquist-Shannon sampling theorem, this frequency content can be represented in discrete time at 8 Hz if (perfect) sampling is performed. Fortunately, most commercial video equipment is designed for flicker-free playback, which is achieved at frame rates of at least 20 to 30 Hz, thus covering the frequency content of vital signs.

At the time the first functioning video camera tubes were presented in the 1930s, Hertzman published a paper entitled “The blood supply of various skin areas as estimated by the photoelectric plethysmograph” [5]. Fig. 1a describes his setup, in which light is emitted from a miniature lightbulb into the skin of a subject. At an adjacent area, a photocell is used to detect “only light which has passed through the skin”.

Using this setup, Hertzman recorded “photoelectric plethysmograms” from various subjects and areas of the skin, see for example Fig. 1b recorded from the forehead. Due to the pulsatile nature of blood flow and pressure, the microvasculature expands and contracts with every heartbeat, which leads to a change in relative blood volume. Since blood has different optical properties than the surrounding tissue, this ultimately causes a change in the intensities of transmitted and reflected light. While the word “electric” was later dropped, the general principle of photoplethysmography (PPG) was demonstrated by Hertzman and, over the years, became an invaluable clinical tool. It allowed the non-invasive (yet



**Fig. 1** PPG apparatus as described by Hertzman (a), “photoelectric plethysmogram” (b), from [5].

contact-based) monitoring of the heart rate, and later even oxygen saturation. This method termed pulse oximetry was developed in the 1970s by Aoyagi [6] and has its roots in discoveries made by Matthes in the 1930s [7]. It derives the oxygen saturation of the arterial blood (SpO<sub>2</sub>) from the different spectral properties of oxygenated and deoxygenated hemoglobin, typically at two wavelengths of 660 and 940 nm.

More than 60 years after Hertzman, the authors Wu, Blažek, and Schmitt published a “well-ignored” paper “Photoplethysmography Imaging: A New Noninvasive and Non-contact Method for Mapping of the Dermal Perfusion Changes” [8]. In it, they presented a system for Photoplethysmography Imaging (PPGI), which consisted of:

- a cooled CCD-camera;
- a light source made up of near-infrared (NIR) light emitting diodes (LEDs);
- a “high performance PC (Pentium II 350 MHz, 256 MB RAM and 9 GB hard disk)”.

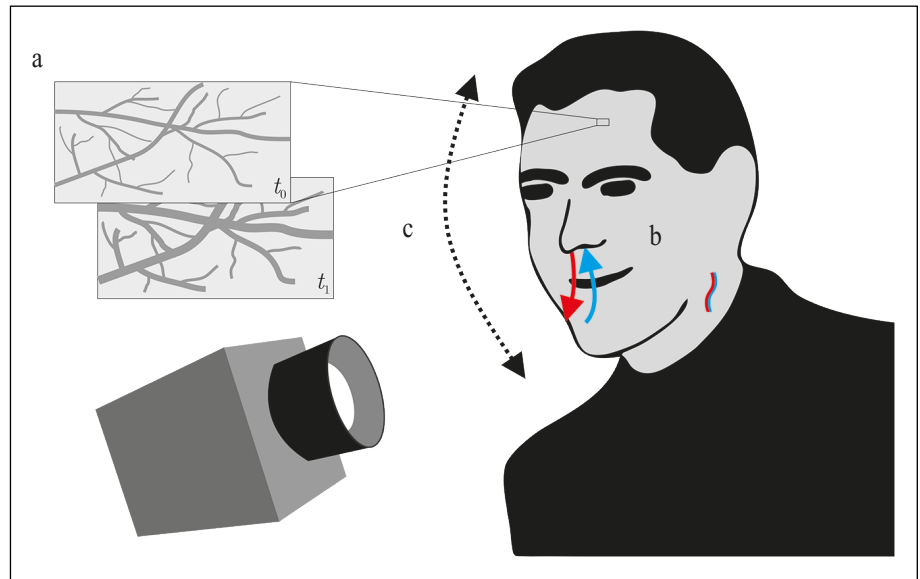
With this pioneering system, Wu, Blažek, and Schmitt demonstrated that camera-based technology could indeed be used for non-contact measurement of heart rate and breathing rate from dermal perfusion, fusing Hertzman’s work with the technological advancements in video processing of the new millennia. Fig. 2a shows the concept, demonstrated for a patch on the forehead. As with contact-based PPG, the pulsatile changes in blood pressure lead to rhythmic changes in the microvascu-

lature and thus to changes in the optical properties of the skin. Instead of using a single photo detector in contact with the skin, a camera records these changes from a distance.

In 2004, Murthy, Pavlidis, and Tsi-amyrztis first reported the use of a cooled mid-wave infrared (MWIR) thermography camera for the non-contact measurement of breathing rate [9]. This work is fundamentally different from previous approaches as it describes a completely passive technology, i.e. one neither depending on additional illumination nor ambient light, but on thermal radiation. Moreover, as the system exploits respiration-induced temperature changes, it relies not only on different parts of the electromagnetic spectrum (usually with wavelengths above 3,000 nm) but also on a different physiological effect. Three years later, Garbey et al. showed that minute changes in skin temperature atop major superficial vessels can be monitored with an MWIR camera and used for heart rate estimation [10]. The concept is shown in Fig. 2b for both the rhythmic cooling/heating in the nose and mouth region as well as the pulsatile changes in temperature that can be recorded from the neck.

Although SpO<sub>2</sub> is not one of the four basic vital signs according to the common definition, it is often considered the “fifth vital sign” [11]. In addition, its derivation is directly linked to the pulsatile PPG information [12]. Starting in 2005 with a pioneering work by Wieringa, Mastik, and van der Steen [13], researchers have investigated whether SpO<sub>2</sub> can be derived from PPGI information.

The following statement is found in Herzman’s 1937 publication: “In applying the plethysmograph to the skin of the face and forehead, it is necessary to mount it on a counter-weighted head strap due to movements of the head with the heart beat and with respiration.” Although Wu et al. also mentioned the need for movement compensation in their work, in 2013 Balakrishnan, Durand, and Gutttag were the first to explicitly show that this “artifact” can be a useful signal, often termed video ballistocardiography [14]. While PPGI relies on superficial perfusion, motion-



**Fig. 2** The three principles of camera-based vital sign monitoring: a) PPGI detects slight variations of the optical properties of the skin, b) systems using MWIR / LWIR exploit temperature changes due to respiration and blood flow, and c) motion-based methods detect slight motion of the head (or other parts of the body) related to cardiac or respiratory activity.

based methods do not require exposed skin in the camera’s field of view for heart rate or respiratory rate estimation (see Fig. 2c).

## 2 Today’s Situation and Scope of the Review

Almost 20 years after the work of Wu, Blažek, and Schmitt, PPGI and other camera-based vital sign estimation methods have seen tremendous, accelerating growth, and several groups around the globe are working on various aspects of the technology, often using different names for the very same method, see for example Table 4 in [15]. Together with the historic background provided above, the reader is cautioned that anyone pursuing work in the area is in some danger of reinventing the wheel. The problem was somewhat alleviated by several excellent review papers on the topic that have been published in recent years. Thus, instead of sketching the complete history, the next part of the paper will start with a “review of reviews”. After that, and in accordance with the scope of this journal, only the most recent works are presented and

analyzed, i.e. publications from 2016 and following. In particular, this survey intends to highlight the broad spectrum of research currently undertaken, its commonalities, and interesting outliers.

For our analysis, an extensive search of the current literature was performed using several approaches. First, existing reviews were consulted. Next, four major online databases (IEEE Xplore, Google Scholar, PubMed, and ResearchGate) were queried with the following search terms: PPGI, Photoplethysmography Imaging, Remote PPG, IRT, Infrared Thermography, Video Monitoring, Non-Contact Pulse Measurement, Heart Rate Video Extraction, Remote Vital Sign Estimation, Real-Time PPG, Vital Sign Monitoring, Thermography Vital Sign, Heart Rate Camera, Vital Sign Camera, Contactless Heart Rate, and Non-Contact Blood Perfusion. In addition, works citing the fundamental works on camera-based vital sign estimation were identified using Google Scholar. Finally, only papers presented in established journals and conference proceedings were accepted for further analysis. For this, each citation was reviewed individually by the authors and a subjective decision was made.

### 3 A Review of Reviews

The first major review on PPGI is presented in 2015 by McDuff et al., in which motion-based methods are explicitly excluded [16]. They report that from over 60 papers they found, more than a third are published within the previous 12 months, indicating the fast growth of the field. As one major technical challenge, they identify motion tolerance and suggested improvements in the areas of (i) algorithmic development, (ii) spatial redundancy, and (iii) multi-wavelength approaches. As another, tolerance to changes in ambient light is identified. In addition to those technical problems, McDuff et al. identify the following problems: lack of common datasets, small population sizes, and no forum for publishing “null results”. Moreover, it is stated that quantities other than heart rate (they mentioned pulse rate variability, pulse transit time, respiration, and SpO2) are seldom evaluated in terms of performance. Laboratory works are most commonly cited. With respect to applications outside of the laboratory, the authors mention applications in neonatal intensive care units and during dialysis.

After McDuff et al., Sun, and Thakor published an extensive review, focusing on PPGI but also covering non-contact (but non-camera) or wearable PPG [17]. Sikdar et al. also made an interesting contribution by reviewing and implementing various published works [18]. Based on a dataset recorded from 14 volunteers under varying conditions, they evaluated the different algorithms’ capabilities for heart rate estimation. At the end of their systematic analysis, they conclude that “none of the existing techniques can predict accurate PR [pulse rate] at all types of conditions and, thus, may not be suitable for clinical purposes unless improved further.” [18]. A recent review by Rouast et al. also focused on algorithmic aspects [19]. It stands out as the authors analyzed 35 papers in terms of “signal extraction” (six subcategories), “signal estimation” (“filtering” and “dimensionality reduction”), “heart rate estimation”, and “contribution / deficiencies”, and presented their results in a large table [19].

The most recent review paper by Zauneder et al. is the most comprehensive one so far with more than 220 references, giving an excellent

overview with background and current realizations of PPGI [20]. In particular, they analyzed the state of the art of unobtrusive heart rate (variability) estimation in terms of:

- populations and experimental protocols;
- hardware;
- image processing;
- color channels;
- signal processing;
- heart rate extraction.

In addition, a separate section is dedicated towards “physiological measures beyond HR”. In the discussion section, the authors conclude that major improvements in PPGI were made in “region of interest (ROI) definitions and combination of color channels”. However, the authors also identify two major shortcomings. First, the applicability of most work outside the laboratory is still questionable, in particular since large public datasets recorded under real-world conditions including pathological conditions are missing. Second, they claim that “available knowledge in many cases is not considered properly”, in particular when it comes to color transformations and blind source separation.

### 4 Survey Taxonomy

In our search, we found 116 papers that were published within the last three years, and that (in our opinion) contributed to the field of camera-based vital sign estimation. As previous review papers have identified, research papers currently exhibit many similarities, which one might in some cases consider shortcomings of the state of the art. In Table 2 of the Zauneder review [20], papers are broadly grouped and counted by their main focus. Expanding on this concept, the first goal of this survey is to conduct a semi-quantitative analysis of the state of the art by asking the following, specific questions:

- How many subjects participated?
- Did the subjects consist of healthy volunteers or patients?
- Were recordings made in the laboratory or under real-world conditions?
- What type of camera was used?
- What type of illumination was used?

- What part of the spectrum was used?
- How many cameras were used?
- What algorithms were used?

As a result, quantitative information about the most recent publications was obtained that may be used to better understand the current state of the art in the field of camera-based monitoring of vital-signs without diving too deep into the technical details of specific implementations. This information is condensed into straightforward graphical representations that provide an intuitive understanding of the state of the art. Thus, this survey is intended to be an addition to comprehensive reviews that were recently published [20] instead of a newer version of the same concept.

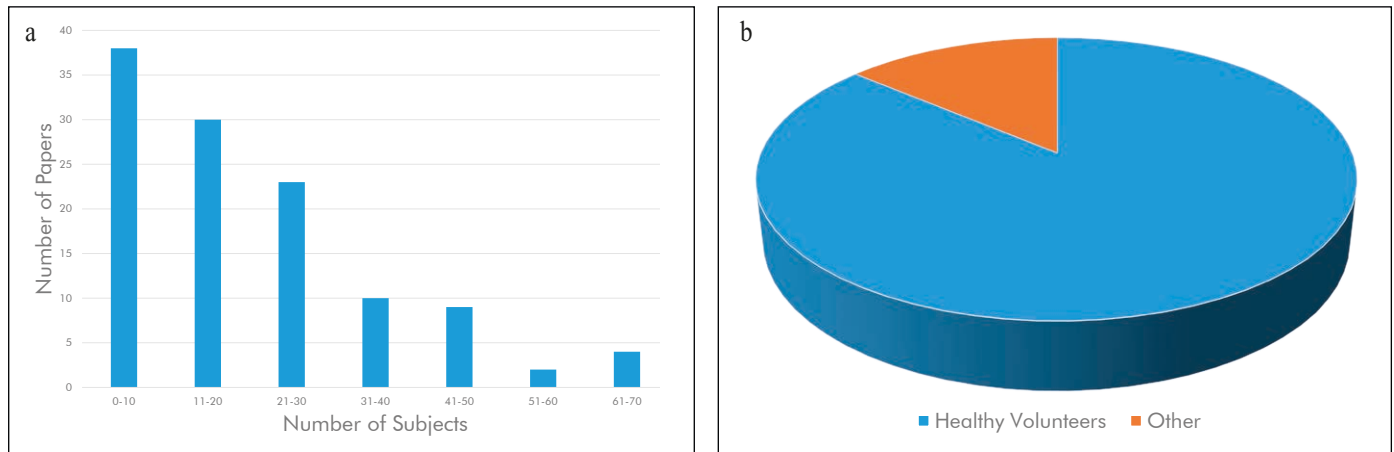
The second goal of the survey is to highlight works that stand out from the majority of approaches and might indicate interesting future perspectives.

#### 4.1 From Healthy to Sick, from Infants to Elders, from Lab to Real World

Fig. 3a shows the distribution of studied subjects in the reviewed literature. Some of the reviewed papers used multiple datasets for evaluation. In those cases, we considered the dataset with highest number of subjects. Fig. 3b visualizes the fraction of papers that used healthy volunteers only.

As observed in the previous literature, the majority of papers present results on a few (20 or fewer) healthy subjects. However, notable exceptions do exist. For example, [21] and [22] present results from a video dataset of 70 patients “during the immediate recovery after cardiac surgery” [21] with an average duration of 28.6 minutes. Elphick et al. report results on 61 sleeping subjects (41 adults, 20 children) [23]. Although using healthy subjects under laboratory conditions, three published papers present results on more than 50 subjects, namely Fan and Li (65 subjects, one minute each) [24], Cerina et al. (60 subjects, five minutes each) [25] and Vieira Moco et al. (54 subjects for model calibration, duration not reported) [26]. It





**Fig. 3** (a) Histogram of number of subjects per paper, (b) distribution of studies with healthy volunteers compared to studies including pathological conditions (b). The majority of current studies includes fewer than 20 subjects and deal with healthy volunteers.

is interesting to note that the average number of participating subjects has increased every year: from 15.7 in 2016, over 21.7 in 2017, to an average of 22.9 subjects for papers published in 2018.

Another encouraging observation we made is that data sharing seems to be increasing. Camera-based technologies for vital sign estimation suffer from the obvious drawback that proper anonymization of raw data is, at least in the visual spectrum, virtually impossible. In addition, the amount of data associated with raw video streams makes sharing cumbersome. Nevertheless, we found that four datasets with video and reference data were publicly available and used in 10 research papers:

1. UBFC-RPPG (42 subjects, [27]), used in [28-30];
2. MMSE-HR (40 subjects, [31]), used in [30, 32];
3. MAHNOB-HCI (30 subjects, [33]), used in [32, 34, 35];
4. DEAP (22 subjects, [36]), used in [37, 38].

While studies with healthy subjects certainly have their place in this evolving field, some researchers are already working with subjects diagnosed with different conditions. For example, PPGI was applied in an interoperative [39] as well as a postoperative setting [40]. In [41], the authors reported results on 40 patients

undergoing hemodialysis treatment, and in [42], respiratory rate has been estimated in 28 patients in the post-anesthesia care unit with infrared thermography. Kamshilin et al. analyzed microcirculation in migraine patients and healthy controls [43], while Rubīns, Spīgulis, and Miščuks used PPGI for continuous monitoring of regional anesthesia [44]. Another growing field of application for PPGI is non-contact monitoring of neonates, particularly in the neonatal intensive care unit (NICU). To this end, several groups have presented work to extract respiratory [45] and cardiac information [46-48]. Data from the NICU is also presented in [49] (HR estimation, one subject) as well as in [50] (automated skin segmentation, 15 subjects).

In the following, the published papers are categorized in terms of studies that present laboratory results only and studies that include scenarios that can be described as other, “real world” applications (see Fig. 4). According to our definition, all studies presented above that include subjects other than healthy individuals belong to this category. Nevertheless, there are also works that approach real world measurement scenarios with healthy subjects only. For example, the area of monitoring during sportive exercise is challenging due to the presence of motion artifacts and was addressed by Wang et al. with PPGI for HR extraction [51-54].

Similarly, with subjects on a stationary exercise bike, Chauvin et al. used thermography for respiratory rate estimation [55], while Capraro et al. estimated vital signs with thermography and an RGB camera [56]. Another promising yet challenging environment for camera-based monitoring is the car. For this scenario, two groups have presented results on HR estimation for one subject each [57, 58]. Focusing on a motion resistant spectral peak-tracking framework, Wu et al. have evaluated their approach during fitness as well as driving scenarios [59]. An NIR setup for driver monitoring was also used by Nowara et al. [60]. An initial evaluation of using PPGI for triggering magnetic resonance imaging was presented by Spicher et al. [61].

The previous focus on non-healthy subjects, large cohorts, and/or real-world applications does not imply that other works contribute less to the advancement of the field of camera-based vital sign estimation. On the contrary, important fundamental insights might be gained from laboratory, analytical, or simulative studies. For example, several works have contributed to the understanding of the true origin of the PPGI signal, its disturbances and modulations [62-67], effects of vasomotor activity [68, 69], assessment of arterial stiffness [70], or the fusion of information from a camera array [71].

## 4.2 From Smartphones to High-End Cameras, from Ambient Light to Kilowatts of Illumination

To analyze current research with respect to the equipment used, we chose two categories. In the category “consumer grade”, we subsume works that make use of of-the-shelf camera equipment, such as webcams, Internet Protocol (IP)-based cameras, camcorders, and digital photo cameras. Moreover, approaches that used smartphone-integrated cameras are also part of this category. Finally, we understand “specialized cameras” as laboratory-grade or industrial-grade equipment, i.e. cameras that are generally not used for entertainment purposes. Often, these cameras provide features such as interchangeable lenses, a high signal-to-noise ratio (SNR), bit depth, quantum efficiency, and/or frame rate. While we acknowledge that these categories are neither sharp nor comprehensive (a high-end, of-the-shelf digital single lens reflex (DSLR) camera or camcorder might be both more expensive and more powerful than a specialized camera), we believe that this is the most useful categorization. The distribution is shown in Fig. 5.

Out of the 116 analyzed papers, we found to our surprise that only 39 works used consumer grade equipment. This was unexpected as the low-cost aspect of camera-based vital sign estimation was often highlighted in previous reviews [16-18, 20]. Half of those publications made use of dedicated webcams [72-75], other types of cameras are regular or “action” camcorders [76-78], DSLRs [79], and the Microsoft Kinect [80-82]. Notably, the public databases DEAP [36] and UBFC-RPPG [27] were also recorded using an inexpensive camcorder/webcam respectively. Using an action camcorder, Al-Naji, Perera, and Chahl demonstrated the estimation of vital signs from a hovering unmanned aerial vehicle [83]. While Szankin et al. used a webcam to estimate vital signs from a distance of six meters [84], Blackford et al. evaluated the possibility to extract HR with a mirrorless camera (Panasonic GH4 camera + Rokinon 650Z-B 650-1300 mm Super Telephoto Zoom Lens) from up to 100 m [85, 86]. Jeong and Finkelstein presented a work on non-contact blood pressure estimation, in which a consumer-grade high-speed camera was used (Casio EX-FH20, 224×168 pixels

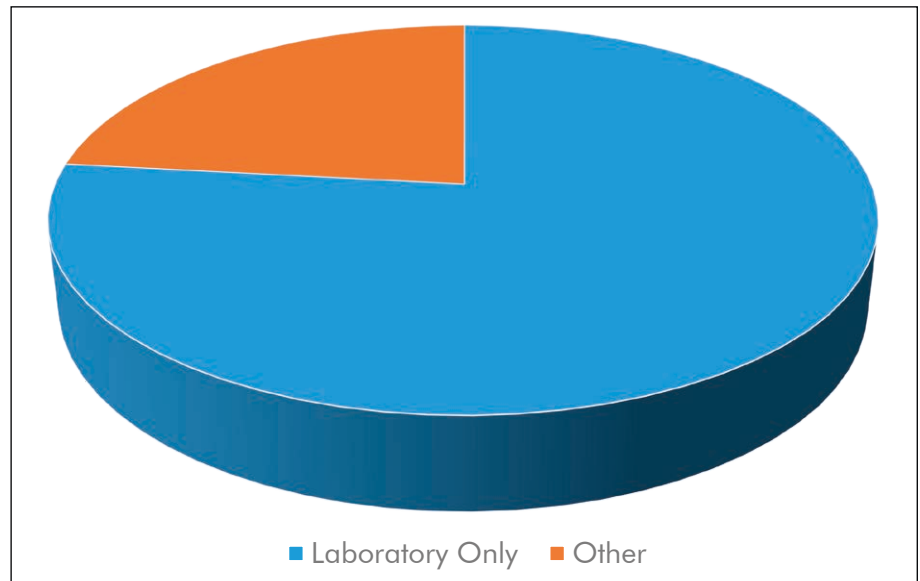


Fig. 4 Distribution of laboratory studies and studies that include real-world applications such as hospital environments, exercising, or driving subjects, etc.

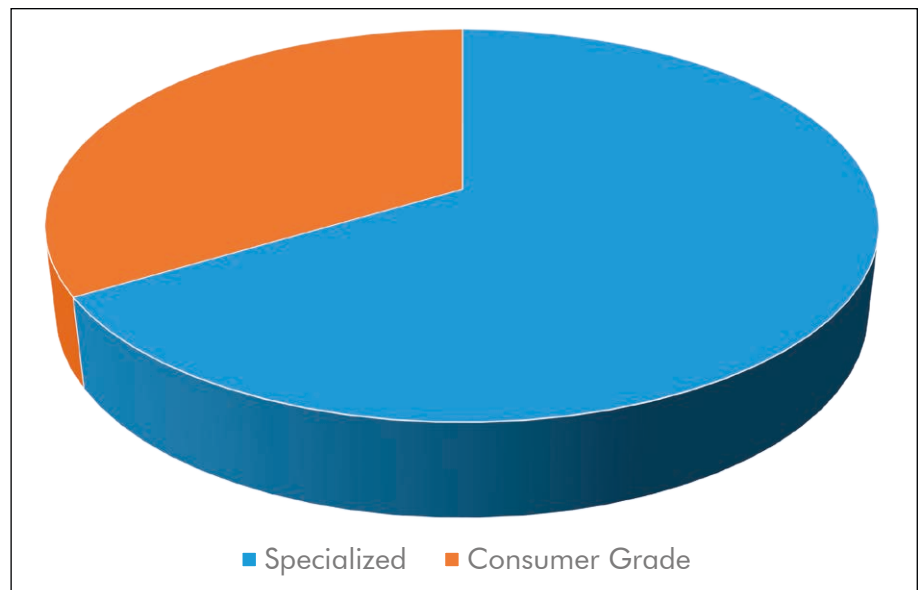


Fig. 5 Distribution of publications using consumer grade equipment vs. studies using specialized camera systems.

at 420 frames per second) [87]. An approach was suggested by Zhang et al., where contact PPG would in the future be fused with PPGI from a smartwatch for blood pressure estimation via pulse transit time [73], although the presented measurements were performed with a stationary camera. Other groups have

evaluated the usage of smartphone cameras for the estimation of HR and RR [88] as well as heart rate variability [89]. The approach presented by Nam et al. used both cameras of the smartphone, one as a contact bases PPG sensor and one to capture chest motion [90]. There are several other research projects and

even commercial software that extract a PPG signal from the user's finger when put on the smartphone's camera [91], which is out of the scope of this review. The vast majority of works on thermography use specialized and expensive equipment such as [55, 92] and are discussed in the following section. However, exceptions do exist that make use of relatively inexpensive smartphone-attached thermal cameras [93, 94].

In our opinion, it is difficult to further subdivide cameras technologies into meaningful subcategories due to the heterogeneity in terms of resolution, frame rate, focal length, quantum efficiency, etc. For an overview, we suggest Table 1 in [86] and Table 2 in [17] to interested readers.

With the exception of methods based on thermography, suitable illumination is necessary for PPGI. In our analysis, we found that the situation in more than 80% of publications can broadly be described as ambient light, which may consist of sunlight, light from fluorescent tubes, incandescent bulbs or a mixture thereof. Some authors state that challenging illumination conditions are included in their dataset [95] or have explicitly analyzed the effect of fluctuating lighting conditions [72]. Nevertheless, setups may use very specialized illumination schemes. For example, high-power quartz-halogen illumination is used for the multispectral [96] and for the long-range [74, 75] setups, 3kW in the former and 2kW in the latter case. Iakovlev et al. introduced a custom light source based on narrow-bandwidth high power LEDs for PPGI applications [97]. Moreover, Moco et al. demonstrated that homogenous and orthogonal illumination can reduce ballistocardiographic artifacts in PPGI measurements [26], while Trumpp et al. found a significant improvement when in addition perpendicular polarization is used [98].

### 4.3 From Thermography to (Ultra)Violet, from a Single Camera to Eighteen

The majority of recent works on camera-based estimation of vital signs uses a single RGB camera (see Fig. 6). In addition to the high availability of devices, important algorithmic advantages have been made that

rely on RGB signals, such as blind source separation and skin segmentation, which we will discuss later on. One disadvantage of RGB imaging is that some sort of illumination in the visual domain is necessary. Depending on the application, this may be ambient light (and thus no additional light source is needed) but still prohibits applications that require dark environments such as sleep monitoring. However, works have shown that the usage of NIR illumination [99] or imagers with NIR filters [100-103] can be used to estimate HR and BR. There are also works that have analyzed the performance of RGB and NIR systems when recorded in parallel [39, 104].

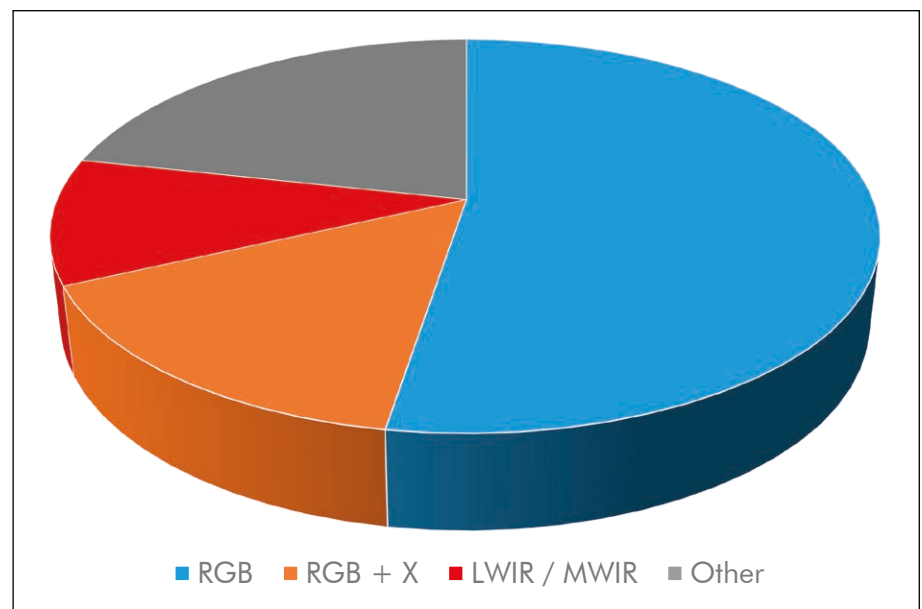
If complete passiveness is necessary or desired, thermography, i.e. imaging in the LWIR or MWIR domains, offers a promising alternative [105]. The pioneering works by Garbey et al. had shown that the extraction of cardiac information is possible under laboratory conditions, and some works are continuing in this direction [106]. Still, the majority of recent works use thermography for extraction of respiratory information [107-109].

A different motivation for the usage of non-RGB systems lies in the estimation of oxygen saturation. As mentioned above, derivation of SpO<sub>2</sub> requires PPG(I) information

from different wavelengths. Thus, works in this domain use, for example, individual cameras with filters and center frequencies of 675 and 842 nm [110], 760, 800, and 840 nm [111], or 760, 800, and 890 nm [112]. Using a different approach, Shao et al. used a monochromatic camera and illumination at 611 and 880 nm [113]. Both Würtenberger et al. [114] and Rapczynski et al. [115] used a hyperspectral video system in the NIR range (600 – 975 nm) to assess HR estimation performance in different NIR bands. These works and those using monochromatic imaging are subsumed in the "Other" category in Fig. 6.

Another promising trend is the combination of RGB information with thermography to increase robustness of ROI selection [116, 117], the estimation of HR [118] and BR [119], or to extract additional information such as increased perfusion [120].

Finally, the most complex multi-camera systems so far was presented by Blackford and Estep [96]. It consists of 12 individual monochromatic cameras equipped with bandpass filters of 50 nm width distributed from 400-750 nm, a dual-CCD RGB/near-infrared imager as well as a spatial array of five RGB imagers placed around the subject.



**Fig. 6** The majority of studies (53%) rely on RGB information. 16% of studies combine RGB with additional spectral information such as NIR / LWIR / MWIR, while about 10% rely on LWIR / MWIR (thermal imaging) only. The remaining 22% of studies use other mono- or poly-chromatic information.

Together with spectral and physiological reference measurement systems, the authors state that the system will be used for a “full-scale evaluation of the spectral components” of PPGI as well as its SNR, spatial perfusion, and blood flow dynamics.

#### 4.4 From Bandpass-Filtering to Machine and Deep Learning

When analyzing the recent literature on camera-based vital sign assessment, almost all publications claim to make some sort of algorithmic contribution to the field. On the one hand, this can be attributed to the nature of the problem which suggests that no optimal solution has been found yet. It may also suggest an active interest of the image processing community to this medical domain. A closer look reveals that datasets or algorithms are not usually shared. As a consequence, complex algorithms need to be re-implemented from one research team to the other and numeric results are seldom comparable. Nonetheless, several important contributions have been made in recent years, and some trends can be observed as well. The most generic and simplistic description of most algorithms for video-based vital sign extraction can be described by the following steps:

1. Select pixels in space and time that contain useful information;
2. Extract a one or multi-dimensional signal from the selected pixels;
3. Calculate the desired quantity from the extracted signal(s).

The first step usually implies the selection of a region of interest (ROI) and its tracking over time. While the most basic approach would consist of a manually selected, time-invariant ROI, several advanced concepts have been developed in recent years. For example, Tulyakov et al. proposed an approach termed “Self-Adaptive Matrix Completion”, which dynamically selects individual face regions and “outputs the HR measurement while simultaneously selecting the most reliable face regions for robust HR estimation” [32]. Similarly, approaches presented by Bobbia et al. have used the concept of superpixels for dynamic

ROI selection and tracking [27-29]. Using convolutional neural networks, Chaichulee et al. were able to detect patients and select skin regions from NICU recordings [50]. In [121], a set of stochastically sampled points from the cheek region was used to estimate the PPG waveform via a Bayesian approach. To estimate BR, Lin et al. automatically selected a salient region from the torso [122].

ROI selection is both crucial and difficult in applications using thermography [123]. For one, it is often impossible to use existing face detectors without intense tuning due to significant differences to RGB images (which are generally used for the development and test of these algorithms). For another, a common source of the signal is the nostril region and is thus much smaller compared to, for example, the forehead. In consequence, specific detection and tracking algorithms have to be used [55, 92].

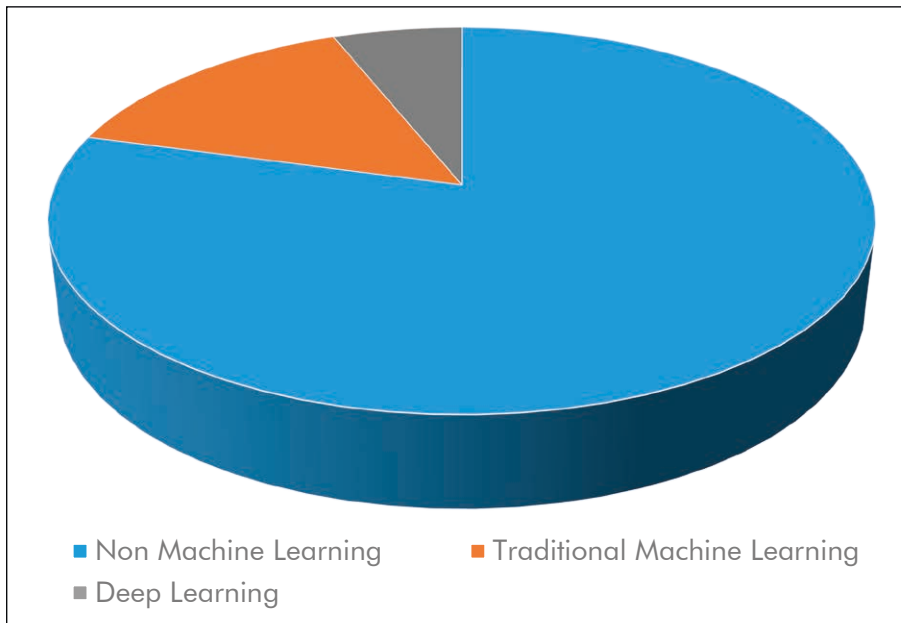
Once the ROI is selected, the extraction of one or more signals is performed. In “Algorithmic Principles of Remote PPG”, Wang et al. categorized this step into blind source separation (BSS), model-based, and data-driven methods [124]. Traditionally, BSS methods using principal or independent component analysis (PCA and ICA respectively) have been widely used in spatial as well as spectral dimensions, see for example reviews presented in [19, 20]. Wedekind et al. found that BSS methods may actually improve in monochromatic conditions, specifically if they are used only on pixels of the green channel [40]. To improve the performance of ICA, constraints based on skin tone and periodicity were introduced in [30]. Non-linear mode decomposition was used by Demirezen and Erdem for HR estimation [125]. In a different approach, Qi et al. used joint blind source separation (JBSS) to extract a set of sources from multiple RGB regions [38]. In [72], JBSS was combined with ensemble empirical mode decomposition to increase robustness towards illumination variation. Unlike BSS, model-based methods such as BPV and CHROM use the knowledge of the color vectors of the different components to control the demixing and were found to be superior to PCA and ICA methods [124]. However, results were on average only marginally better than the

basic method of subtracting the red from the green channel [124]. Further improvements were achieved by the data-driven “Spatial Subspace Rotation” [95] and by its combination with the CHROM approach [124]. To fuse several chrominance-features into a single PPG curve, Liu et al. introduced self-adaptive signal separation combined with a weight-based scheme to select “interesting sub-regions” [126].

In addition to these works that can, in the broadest sense, be summarized as preprocessing, other algorithmic works have focused on distinct aspects of camera-based vital sign estimation. For example, some recent works have addressed the important topic of the effect of video compression on PPGI [25, 127-129]. Others have focused on specific concepts such as the Eulerian video magnification (EMV) [84, 130], the exploitation of the ballistocardiographic effect [34], the derivation of signal quality indices [131], or a performance evaluation of several existing methods on a publicly available dataset [37]. Yet other works deal with specific aspects that reach beyond the basic vital signs such as the estimation of blood pressure variability [132], pulse wave delay [133], the jugular venous pulse waveform [134], and venous oxygen saturation [135].

Fig. 7 shows the distribution of the usage of machine learning in camera-based vital sign estimation. While the majority of papers ( $n=92$ ) do not explicitly use machine learning techniques, works that do are increasing (24 in total, 11 since 2018). As of now, traditional approaches [136] are the most common ( $n=17$ ). Given the general trend towards deep learning in computer vision, it comes to no surprise that this tool has found use in camera-based vital sign estimation ( $n=7$ ) [94, 137-139]. Deep learning has proven useful for image classification and segmentation in several fields, including medical image analysis [140]. First applications of deep learning in camera-based vital sign estimation applied the concept for detection of neonates and skin regions in an incubator by using a multi-task convolutional neural network (CNN) [50]. Similarly, Pursche et al. used CNNs to optimize ROIs [137]. Qiu et al. combined EMV and CNN approaches to extract HR from facial video data [138]. After pre-processing, a regression CNN is applied





**Fig. 7** The majority of studies (79%) make no explicit use of machine learning techniques. However, 15% of studies employ some traditional machine learning techniques, while 6% of papers use recent approaches from the realm of deep learning such as convolutional neural networks.

to so-called “feature-image” to extract HR. Cho et al. also used a CNN-based approach for the analysis of breathing patterns acquired with thermography [94]. However, their CNN architecture was applied to extracted spectrogram data and not the raw thermal images.

The fact that vital sign estimation is possible with ubiquitous, inexpensive, consumer-grade equipment even from great distances raises privacy concerns that have been addressed by Chen and Picard with an approach able to eliminate physiological information from facial videos [141]. Focusing on privacy issues in camera-based sensing, Wang, den Brinker, and de Haan introduced a single element approach termed “SoftSig” [142].

## 5 Lessons Learned, Limitations, and Future Work

Since 2016, the “average publication” on camera-based vital sign estimation is made of a laboratory study including about 20 subjects, filmed with an industrial-grade or laboratory RGB camera under ambient light conditions

and presents some algorithmic advancement in terms of ROI selection, tracking, or 1D signal extraction. At the same time, the variety of recent works is broader. It ranges from LWIR to violet in terms of electromagnetic spectrum usage, from low-cost webcams to cooled thermal cameras, from one camera to 18, from one healthy subject to 70 patients, from monitoring the palm to the whole body, and from sitting in front of the webcam over exercising subjects to recordings made from a distance of 100 meters.

In our opinion, some limitations of the state-of-the-art need to be addressed. First, the focus on young, healthy, stationary individuals may offer insights into the fundamental properties of camera-based vital sign estimation. Nevertheless, our own experiences suggest that signals from other types of subjects such as infants or elders may be quite different and offer a wide range of challenges. Second, algorithms developed for RGB images may not be applicable to data with other spectral properties such as thermography, which can be acquired without illumination in the visible domain and thus expands the range of application. Third, sharing of algorithms and data is relatively uncommon. This is

partly understandable for data due to privacy or bandwidth issues. Nevertheless, it would make results much more comparable and could thus severely help the development of the field. Unfortunately, our findings with respect to standardization of setups and data sharing hardly differ from the ones reported in previous review papers, even those made by McDuff et al. in 2015 [16].

It was surprising to find that most publications use laboratory or industrial-grade equipment, since camera-based vital sign estimation is often marketed as a low-cost technology. Nevertheless, several works have shown the usability of consumer-grade equipment and we suspect that the reasons for using specialized cameras are more of pragmatic nature. Having the control over internal camera parameters and frame-by-frame synchronization with other devices is a major advantage compared to consumer grade equipment. As it is usually the recording that is the most cumbersome, it is understandable to use the best equipment possible for data acquisition. Additional future work should focus on establishing practical boundaries by, starting from high-quality data, systematically decreasing spatial resolution, bit depth, frame rate, SNR, compression quality, etc. Several works, including, but not limited to, those of our own group, have shown the usefulness of infrared thermography as a completely passive camera-based method. Thus, it is encouraging to see that even these specialized cameras have reached the consumer market with prices lower than 250 € and that their quality is sufficient for vital sign estimation. The same is obviously true in terms of processing power, considering that the average smartphone today has more processing power than the “high performance PC” used by Wu, Blažek, and Schmitt in 2000 [8].

In the following, we highlight three publications that, in our view, had a major impact on the field since 2016. First, the paper “Camera-based photoplethysmography in critical care patients” by Rasche et al. presents an analysis of the performance of camera-based HR estimation on 70 critical care patients [21]. It provides results on a large cohort of non-healthy subjects for whom vital sign estimation is crucial. HR could be detected with an accuracy of  $\pm 5$  BPM for

83% of the measurement time. Moreover, low arterial blood pressure was found to have a negative impact on the SNR. Second, the paper “Algorithmic Principles of Remote PPG” by Wang et al. provides a systematic analysis of different methods to extract HR from RGB video data [124]. Among other findings and insights into the PPGI signal, the authors showed that neither PCA nor ICA could, on average, provide better results than those obtained with the straightforward method of subtracting green and red channels to extract the pulsatile signal. Finally, the paper “Unsupervised skin tissue segmentation for remote photoplethysmography” by Bobbia et al. presents an approach to implicitly select skin tissues based on their distinct pulsatility feature [27]. In addition to providing the algorithm, the authors make a major contribution to the field by providing free access to the largest PPGI database in terms of subjects so far.

“Systematic reviews and meta-analyses are essential to summarize evidence relating to efficacy and safety of health care interventions accurately and reliably” [143]. Thus, it would be interesting to see a statistical meta-analysis for camera-based vital sign estimation that would correlate study attributes and results. Unfortunately, we believe that this is impossible at the current point in time. For one, very few publications use the same datasets. For another, study cohorts are usually way too small to allow for meaningful comparison. Finally, the measurement setups and algorithms usually differ from one group to the next, sometimes even from publication to publication. Thus, a straightforward quantification as performed in this paper, which, as of now, was not performed in any other review, is the only form of analysis we consider feasible at this point in time.

Four public databases are currently available, UBFC-RPPG [27], MMSE-HR [31], MAHNOB-HCI [33], and DEAP [36]. Unfortunately, their usage is relatively limited and as of now, no publication has made use of more than two databases simultaneously. In general, results are usually presented on either a private or a public dataset, an exception being [35]. We believe it would be good practice for future algorithm-oriented publications to present

results on multiple, independent datasets, and sharing algorithms would be beneficial as well. As data sharing may be problematic, it would be helpful to provide a common interface so that interested researchers could test novel algorithms on other datasets by exchanging code without needing to access to potentially sensitive data. A consequent next step would be the establishment of a large benchmark dataset made of a public and a hidden fraction. To access the latter, code would be submitted and executed on a server, which would minimize overfitting and help make results comparable. Finally, sharing and standardizing datasets would allow current and future methods from the realm of deep learning to be applied to camera-based vital sign monitoring.

As stated above, two papers specifically address privacy issues arising from camera-based vital sign estimation. Challenges become even more severe with the increased popularity of deep learning. On the one hand, these methods have shown superior performances in several domains. On the other hand, they require large datasets for training while “[...] privacy issues make it more difficult to share medical data than natural images” [144]. New legislations such as the General Data Protection Regulation (GDPR) of the European Union are likely to exacerbate these challenges and call for approaches such as “Privacy-Preserving Deep Learning” that would allow “[...] multiple parties to jointly learn an accurate neural-network model for a given objective without sharing their input datasets” [145].

To conclude, camera-based monitoring of vital signs is an extremely broad field in terms of hardware setups, applications, and algorithms. In our opinion, this shows both the great versatility of the concept as well as the lack of standardization and comparability. To overcome this, standardized reporting of results, and sharing of algorithms and data (or at least access to data) are necessary. In addition, we believe that the processing of non-visual and multi-spectral video data constitutes a particularly interesting area of research, as it could lead to less obtrusive, more robust monitoring and the estimation of the “fifth vital sign” oxygen saturation. Finally, future work should focus on motion tolerance, non-

healthy subjects, and most important, the combination of the two. Only if pathological episodes can be detected robustly and distinguished from motion artifacts, clinically relevant contributions can be expected.

## References

1. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* 2012;9(7):671-5.
2. Burke BE, Gregory JA, Cooper M, Loomis AH, Young DJ, Lind TA, et al. CCD Imager Development for Astronomy. *Lincoln Lab J* 2007;16(2):393-412.
3. Goldfain A, Smith B, Arabandi S, Brochhausen M, Hogan WR. Vital Sign Ontology. In: *Proceedings of the Workshop on Bio-Ontologies, ISMB, Vienna, June 2011*; 2011. p. 71-4.
4. Glaeser DH, Thomas LJ. Computer Monitoring in Patient Care. *Annu Rev Biophys Bioeng* 1975;4(1):449-76.
5. Hertzman AB. The blood supply of various skin areas as estimated by the photoelectric plethysmograph. *Am J Physiol* 1938;124:328-40.
6. Severinghaus JW. Takuo Aoyagi: Discovery of Pulse Oximetry. *Anesth Analg* 2007 Dec;105(6 Suppl):S1-4.
7. Matthes K. Untersuchungen über die Sauerstoffsättigung des menschlichen Arterienblutes [Studies on the oxygen saturation of human arterial blood]. *Naunyn. Schmiedebergs. Arch Exp Pathol Pharmacol* 1935;179(6):698-711.
8. Wu T, Blažek V, Schmitt HJ. Photoplethysmography Imaging: A New Noninvasive and Non-contact Method for Mapping of the Dermal Perfusion Changes. In: *Optical Techniques and Instrumentation for the Measurement of Blood Composition, Structure, and Dynamics 2000*;4163:62-71.
9. Murthy R, Pavlidis I, Tsiamyrtzis P. Touchless monitoring of breathing function. In: *The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* 2004;3:1196-9.
10. Garbey M, Sun N, Merla A, Pavlidis I. Contact-Free Measurement of Cardiac Pulse Based on the Analysis of Thermal Imagery. *IEEE Trans Biomed Eng* 2007;54(8):1418-26.
11. Neff TA. Routine Oximetry. *Chest* 1988;94(2):227.
12. Allen J. Photoplethysmography and its application in clinical physiological measurement. *Physiol Meas* 2007;28(3):R1-R39.
13. Wieringa FP, Mastik F, van der Steen AFW. Contactless Multiple Wavelength Photoplethysmographic Imaging: A First Step Toward ‘SpO2 Camera’ Technology. *Ann Biomed Eng* 2005;33(8):1034-41.
14. Balakrishnan G, Durand F, Guttag J. Detecting Pulse from Head Motions in Video. In: *2013 IEEE Conference on Computer Vision and Pattern Recognition*; 2013. p. 3430-7.
15. Leonhardt S, Leicht L, Teichmann D. Unobtrusive Vital Sign Monitoring in Automotive Environ-

- ments—A Review. *Sensors* 2018;18(9):3080.
16. McDuff DJ, Estep JR, Piasecki AM, Blackford EB. A survey of remote optical photoplethysmographic imaging methods. In: Proc of the 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); 2015. p. 6398–404.
  17. Sun Y, Thakor N. Photoplethysmography Revisited: From Contact to Noncontact, From Point to Imaging. *IEEE Trans Biomed Eng* 2016;63(3):463–77.
  18. Sikdar A, Behera SK, Dogra DP. Computer-Vision-Guided Human Pulse Rate Estimation: A Review. *IEEE Rev Biomed Eng* 2016;9:91–105.
  19. Rouast PV, Adam MTP, Chiong R, Cornforth D, Lux E. Remote heart rate measurement using low-cost RGB face video: a technical literature review. *Front Comput Sci* 2018;12(5):858–72.
  20. Zauneder S, Trumpp A, Wedekind D, Malberg H. Cardiovascular assessment by imaging photoplethysmography – a review. *Biomed Tech (Berl)* 2018;63(5):617–34.
  21. Rasche S, Trumpp A, Waldow T, Gaetjen F, Plötze K, Wedekind D, et al. Camera-based photoplethysmography in critical care patients. *Clin Hemorheol Microcirc* 2016;64(1):77–90.
  22. Trumpp A, Rasche S, Wedekind D, Rudolf M, Malberg H, Matschke K, et al. Relation between pulse pressure and the pulsation strength in camera-based photoplethysmograms. *Curr Dir Biomed Eng* 2017;3(2):489–92.
  23. Elphick HE, Alkali AH, Kingshott RK, Burke D, Saatchi R. Exploratory Study to Evaluate Respiratory Rate Using a Thermal Imaging Camera. *Respiration* 2019;97(3):205–12.
  24. Fan Q, Li K. Non-contact remote estimation of cardiovascular parameters. *Biomed Signal Process Control* 2018;40:192–203.
  25. Cerina L, Iozzia L, Mainardi L. Influence of acquisition frame-rate and video compression techniques on pulse-rate variability estimation from vPPG signal. *Biomed Tech (Berl)* 2017;64(1):53–65.
  26. Moco AV, Stuijk S, de Haan G. Ballistocardiographic Artifacts in PPG Imaging. *IEEE Trans Biomed Eng* 2016;63(9):1804–11.
  27. Bobbia S, Macwan R, Benezeth Y, Mansouri A, Dubois J. Unsupervised skin tissue segmentation for remote photoplethysmography. *Pattern Recognit Lett* 2017;:1–9.
  28. Bobbia S, Luguern D, Benezeth Y, Nakamura K, Gomez R, Dubois J. Real-Time Temporal Superpixels for Unsupervised Remote Photoplethysmography. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW); 2018. p. 1341–8.
  29. Bobbia S, Benezeth Y, Dubois J. Remote photoplethysmography based on implicit living skin tissue segmentation. In: Proc of the 23rd International Conference on Pattern Recognition (ICPR); 2016. p. 361–5.
  30. Macwan R, Benezeth Y, Mansouri A. Remote photoplethysmography with constrained ICA using periodicity and chrominance constraints. *Biomed Eng Online* 2018;17(1):22.
  31. Zhang Z, Girard JM, Wu Y, Zhang X, Liu P, Ciftci U, et al. Multimodal spontaneous emotion corpus for human behavior analysis. In: Proc of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2016. p. 3438–46.
  32. Tulyakov S, Alameda-Pineda X, Ricci E, Yin L, Cohn JF, Sebe N. Self-Adaptive Matrix Completion for Heart Rate Estimation from Face Videos under Realistic Conditions. In: Proc of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2016. p. 2396–404.
  33. Soleymani M, Lichtenauer J, Pun T, Pantic M. A Multimodal Database for Affect Recognition and Implicit Tagging. *IEEE Trans Affect Comput* 2012;3(1):42–55.
  34. Hassan MA, Malik AS, Fofi D, Saad NM, Ali YS, Meriaudeau F. Video-Based Heartbeat Rate Measuring Method Using Ballistocardiography. *IEEE Sens J* 2017;17(14):4544–57.
  35. Hassan MA, Malik AS, Fofi D, Saad N, Meriaudeau F. Novel health monitoring method using an RGB camera. *Biomed Opt Express* 2017;8(11):4838.
  36. Koelstra S, Mühl C, Soleymani M, Lee J-S, Yazdani A, Ebrahimi T, Paet al. DEAP: A Database for Emotion Analysis Using Physiological Signals. *IEEE Trans Affect Comput* 2012;3(1):18–31.
  37. Unakafov AM. Pulse rate estimation using imaging photoplethysmography: generic framework and comparison of methods on a publicly available dataset. *Biomed Phys Eng Express* 2018;4(4):045001.
  38. Qi H, Guo Z, Chen X, Shen Z, Wang JZ. Video-based human heart rate measurement using joint blind source separation. *Biomed Signal Process Control* 2017;31:309–20.
  39. Trumpp A, Lohr J, Wedekind D, Schmidt M, Burghardt M, Heller AR, et al. Camera-based photoplethysmography in an intraoperative setting. *Biomed Eng Online* 2018;17(1):33.
  40. Wedekind D, Trumpp A, Gaetjen F, Rasche S, Matschke K, Malberg H, et al. Assessment of blind source separation techniques for video-based cardiac pulse extraction. *J Biomed Opt* 2017;22(3):035002.
  41. Villarroya M, Jorge J, Pugh C, Tarassenko L. Non-Contact Vital Sign Monitoring in the Clinic. In: Proc of the 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017); 2017. p. 278–85.
  42. Hochhausen N, Barbosa Pereira C, Leonhardt S, Rossaint R, Czaplík M. Estimating Respiratory Rate in Post-Anesthesia Care Unit Patients Using Infrared Thermography: An Observational Study. *Sensors* 2018;18(5):1618.
  43. Kamshilin AA, Volynsky MA, Khayrutdinova O, Nurkhametova D, Babayan L, Amelin AV, et al. Novel capsaicin-induced parameters of microcirculation in migraine patients revealed by imaging photoplethysmography. *J Headache Pain* 2018;19(1):43.
  44. Rubins U, Spigulis J, Miščuks A. Photoplethysmography imaging algorithm for continuous monitoring of regional anesthesia. In: Proceedings of the 14th ACM/IEEE Symposium on Embedded Systems for Real-Time Multimedia - ESTIMedia'16; 2016. p. 67–71.
  45. Jorge J, Villarroya M, Chaichulee S, Guazzi A, Davis S, Green G, et al. Non-Contact Monitoring of Respiration in the Neonatal Intensive Care Unit. In: Proc 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017); 2017. p. 286–93.
  46. Blanik N, Heimann K, Pereira C, Paul M, Blazek V, Venema B, et al. Remote vital parameter monitoring in neonatology – robust, unobtrusive heart rate detection in a realistic clinical scenario. *Biomed Tech (Berl)* 2016;61(6): 631–43.
  47. Chaichulee S, Villarroya M, Jorge J, Arteta C, Green G, McCormick K, et al. Localised photoplethysmography imaging for heart rate estimation of pre-term infants in the clinic. In: Optical Diagnostics and Sensing XVIII: Toward Point-of-Care Diagnostics. International Society for Optics and Photonics; 2018. p. 105010R.
  48. van Gestel M, Balmaekers B, Verkruysse W, Bambang Oetomo S. Near-continuous non-contact cardiac pulse monitoring in a neonatal intensive care unit in near darkness. In: Optical Diagnostics and Sensing XVIII: Toward Point-of-Care Diagnostics. International Society for Optics and Photonics; 2018. p. 1050114.
  49. Wang W, den Brinker AC, de Haan G. Full video pulse extraction. *Biomed Opt Express* 2018;9(8):3898.
  50. Chaichulee S, Villarroya M, Jorge J, Arteta C, Green G, McCormick K, et al. Multi-Task Convolutional Neural Network for Patient Detection and Skin Segmentation in Continuous Non-Contact Vital Sign Monitoring. In: Proc of the 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017); 2017. p. 266–72.
  51. Wang W, Balmaekers B, de Haan G. Quality metric for camera-based pulse rate monitoring in fitness exercise. In: Proc of IEEE International Conference on Image Processing (ICIP); 2016. p. 2430–4.
  52. Wang W, den Brinker AC, Stuijk S, de Haan G. Color-Distortion Filtering for Remote Photoplethysmography. In: Proc of the 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017); 2017. p. 71–8.
  53. Wang W, den Brinker AC, Stuijk S, de Haan G. Robust heart rate from fitness videos. *Physiol Meas* 2017;38(6):1023–44.
  54. Wang W, den Brinker AC, Stuijk S, de Haan G. Amplitude-selective filtering for remote-PPG. *Biomed Opt Express* 2017;8(3):1965.
  55. Chauvin R, Hamel M, Brière S, Ferland F, Grondin F, Létourneau D, et al. Contact-Free Respiration Rate Monitoring Using a Pan-Tilt Thermal Camera for Stationary Bike Telerehabilitation Sessions. *IEEE Syst J* 2016;10(3):1046–55.
  56. Capraro G, Etebari C, Luchette K, Mercurio L, Merck D, Kirenko I, et al. No Touch' Vitals: A Pilot Study of Non-contact Vital Signs Acquisition in Exercising Volunteers. In: Proc of the IEEE Biomedical Circuits and Systems Conference (BioCAS); 2018. p. 1–4.
  57. Blocher T, Schneider J, Schinle M, Stork W. An online PPGI approach for camera based heart rate monitoring using beat-to-beat detection. In: Proc of the IEEE Sensors Applications Symposium (SAS); 2017. p. 1–6.
  58. Zhang Q, Wu Q, Zhou Y, Wu X, Ou Y, Zhou H. Webcam-based, non-contact, real-time measurement for the physiological parameters of drivers. *Measurement* 2017;100:311–21.
  59. Wu B-F, Huang P-W, Lin C-H, Chung M-L, Tsou T-Y, Wu Y-L. Motion Resistant Image-Photoplethysmography Based on Spectral Peak Tracking Algorithm. *IEEE Access* 2018;6:21621–34.
  60. Nowara EM, Marks TK, Mansour H, Veeraraghava



- van A. SparsePPG: Towards Driver Monitoring Using Camera-Based Vital Signs Estimation in Near-Infrared. In: Proc of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW); 2018. p. 1385–94.
61. Spicher N, Kukuk M, Maderwald S, Ladd ME. Initial evaluation of prospective cardiac triggering using photoplethysmography signals recorded with a video camera compared to pulse oximetry and electrocardiography at 7T MRI. *Biomed Eng Online* 2016;15(1):126.
  62. Moço AV, Stuijk S, de Haan G. Skin inhomogeneity as a source of error in remote PPG-imaging. *Biomed Opt Express* 2016;7(11):4718.
  63. Moço AV, Stuijk S, de Haan G. New insights into the origin of remote PPG signals in visible light and infrared. *Sci Rep* 2018;8(1):8501.
  64. Addison PS, Jacquel D, Foo DMH, Borg UR. Video-based heart rate monitoring across a range of skin pigmentations during an acute hypoxic challenge. *J Clin Monit Comput* 2018;32(5):871–80.
  65. Moço AV, Stuijk S, de Haan G. Motion robust PPG-imaging through color channel mapping. *Biomed Opt Express* 2016;7(5):1737.
  66. Butle, M J, Crowe JA, Hayes-Gill BR, Rodmell PI. Motion limitations of non-contact photoplethysmography due to the optical and topological properties of skin. *Physiol Meas* 2016;37(5):N27–N37.
  67. Sidorov IS, Romashko RV, Koval VT, Giniatullin R, Kamshilin AA. Origin of Infrared Light Modulation in Reflectance-Mode Photoplethysmography. *PLoS One* 2016;11(10): e0165413.
  68. Trumpp A, Schell J, Malberg H, Zaunseder S. Vasomotor assessment by camera-based photoplethysmography. *Curr Dir Biomed Eng* 2016;2(1):199–202.
  69. Zaytsev VV, Miridonov SV, Mamontov OV, Kamshilin AA. Contactless monitoring of the blood-flow changes in upper limbs. *Biomed Opt Express* 2018;9(11):5387.
  70. Moço AV, Mondragon LZ, Wang W, Stuijk S, de Haan G. Camera-based assessment of arterial stiffness and wave reflection parameters from neck micro-motion. *Physiol Meas* 2017;38(8):1576–98.
  71. McDuff DJ, Blackford EB, Estep JR. Fusing Partial Camera Signals for Noncontact Pulse Rate Variability Measurement. *IEEE Trans Biomed Eng* 2018;65(8):1725–39.
  72. Cheng J, Chen X, Xu L, Wang ZJ. Illumination Variation-Resistant Video-Based Heart Rate Measurement Using Joint Blind Source Separation and Ensemble Empirical Mode Decomposition. *IEEE J Biomed Health Inform* 2017;21(5):1422–33.
  73. Zhang G, Shan C, Kirenko I, Long X, Aarts R. Hybrid Optical Unobtrusive Blood Pressure Measurements. *Sensors* 2017;17(7):1541.
  74. Kwasniewska A, Ruminski J, Szankin M, Czuszyński K. Remote Estimation of Video-Based Vital Signs in Emotion Invocation Studies. *Conf Proc IEEE Eng Med Biol Soc* 2018; 2018:4872–6.
  75. Massaroni C, Schena E, Silvestri S, Taffoni F, Merone M. Measurement system based on RGB camera signal for contactless breathing pattern and respiratory rate monitoring. In: Proc of the IEEE International Symposium on Medical Measurements and Applications (MeMeA); 2018. p. 1–6.
  76. Alghoul K, Alharthi S, Al Osman H, El Saddik A. Heart Rate Variability Extraction From Videos Signals: ICA vs. EVM Comparison. *IEEE Access* 2017;5:4711–9.
  77. He X, Goubran R, Knoefel F. IR night vision video-based estimation of heart and respiration rates. In: Proc of the IEEE Sensors Applications Symposium (SAS); 2017. p. 1–5.
  78. Blackford EB, Estep JR. Using consumer-grade devices for multi-imager non-contact imaging photoplethysmography. In: *Optical Diagnostics and Sensing XVII: Toward Point-of-Care Diagnostics* 2017;10072:100720P.
  79. Al-Naji A, Chahl J. Simultaneous Tracking of Cardiorespiratory Signals for Multiple Persons Using a Machine Vision System With Noise Artifact Removal. *IEEE J Transl Eng Health Med* 2017;5:1–10.
  80. Procházka A, Charvátová H, Vyšata O, Kopal J, Chambers J. Breathing Analysis Using Thermal and Depth Imaging Camera Video Records. *Sensors* 2017;17(6):1408.
  81. Lewis GF, Davila MI, Porges SW. Novel Algorithms to Monitor Continuous Cardiac Activity with a Video Camera. In: Proc of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW); 2018. p. 1395–403.
  82. Gambi E, Ricciuti M, Spinsante S. Sensitivity of the Contactless Videoplethysmography-Based Heart Rate Detection to Different Measurement Conditions. In: Proc of the 26th European Signal Processing Conference (EUSIPCO) 2018; 2018. p. 767–71.
  83. Al-Naji A, Perera AG, Chahl J. Remote monitoring of cardiorespiratory signals from a hovering unmanned aerial vehicle. *Biomed Eng Online* 2017;16(1):101.
  84. Szankin M, Kwasniewska A, Sirlapu T, Wang M, Ruminski J, Nicolas R, et al. Long Distance Vital Signs Monitoring with Person Identification for Smart Home Solutions. *Conf Proc IEEE Eng Med Biol Soc* 2018 Jul;2018:1558–61.
  85. Blackford EB, Piasecki AM, ad Estep JR. Measuring pulse rate variability using long-range, non-contact imaging photoplethysmography. *Conf Proc IEEE Eng Med Biol Soc* 2016 Aug;2016:3930–6.
  86. Blackford EB, Estep JR, Piasecki AM, Bowers MA, Klosterman SL. Long-range non-contact imaging photoplethysmography: cardiac pulse wave sensing at a distance. In: *Optical Diagnostics and Sensing XVII: Toward Point-of-Care Diagnostics* 2016;9715: 971512.
  87. Jeong I, Finkelstein J. Introducing Contactless Blood Pressure Assessment Using a High Speed Video Camera. *J Med Syst* 2016;40(4):77.
  88. Wei B, He X, Zhang C, Wu X. Non-contact, synchronous dynamic measurement of respiratory rate and heart rate based on dual sensitive regions. *Biomed Eng Online* 2017;16(1):17.
  89. Huang R-Y, Dung L-R. Measurement of heart rate variability using off-the-shelf smart phones. *Biomed Eng Online* 2016;15(1):11.
  90. Nam Y, Kong Y, Reyes B, Reljin N, Chon KH. Monitoring of Heart and Breathing Rates Using Dual Cameras on a Smartphone. *PLoS One* 2016;11(3):e0151013.
  91. Poh M-Z, Poh YC. Validation of a Standalone Smartphone Application for Measuring Heart Rate Using Imaging Photoplethysmography. *Telemed J E Health* 2017;23(8):678–83.
  92. Barbosa Pereira C, Yu X, Czaplik M, Blazek V, Venema B, Leonhardt S. Estimation of breathing rate in thermal imaging videos: a pilot study on healthy human subjects. *J Clin Monit Comput* 2017 Dec;31(6):1241–54.
  93. Cho Y, Julier SJ, Marquardt N, Bianchi-Berthouze N. Robust tracking of respiratory rate in high-dynamic range scenes using mobile thermal imaging. *Biomed Opt Express* 2017 Sep 13;8(10):4480–503.
  94. Cho Y, Bianchi-Berthouze N, Julier SJ. Deep-Breath: Deep learning of breathing patterns for automatic stress recognition using low-cost thermal imaging in unconstrained settings. In: Proc of the Seventh International Conference on Affective Computing and Intelligent Interaction (ACII) 2017; 2018:456–63.
  95. Wang W, Stuijk S, de Haan G. A Novel Algorithm for Remote Photoplethysmography: Spatial Subspace Rotation. *IEEE Trans Biomed Eng* 2016;63(9):1974–84.
  96. Blackford EB, Estep JR. A multispectral testbed for cardiovascular sensing using imaging photoplethysmography. In: *Optical Diagnostics and Sensing XVII: Toward Point-of-Care Diagnostics* 2017;10072:100720R.
  97. Iakovlev D, Dwyer V, Hu S, Silberschmidt V. Noncontact blood perfusion mapping in clinical applications. In: *Biophotonics: Photonic Solutions for Better Health Care V* 2016:988: 988712.
  98. Trumpp A, Bauer PL, Rasche S, Malberg H, Zaunseder S. The value of polarization in camera-based photoplethysmography. *Biomed. Opt. Express*, vol. 8, no. 6, p. 2822, Jun. 2017.
  99. Li, M.H., Yadollahi, A., and Taati, B. Noncontact Vision-Based Cardiopulmonary Monitoring in Different Sleeping Positions. *IEEE J Biomed Health Inform* 2017;21(5):1367–75.
  100. Amelard R, Clausi DA, Wong A. Spectral-spatial fusion model for robust blood pulse waveform extraction in photoplethysmographic imaging. *Biomed Opt Express* 2016;7(12):4874.
  101. Amelard R, Clausi DA, Wong A. Spatial probabilistic pulsatility model for enhancing photoplethysmographic imaging systems. *J Biomed Opt* 2016;21(11):116010.
  102. Wang W, Stuijk S, de Haan G. Living-Skin Classification via Remote-PPG. *IEEE Trans Biomed Eng* 2017;64(12):2781–92.
  103. Amelard R, Hughson RL, Greaves DK, Clausi DA, Wong A. Assessing photoplethysmographic imaging performance beyond facial perfusion analysis. In: *Optical Diagnostics and Sensing XVII: Toward Point-of-Care Diagnostics* 2017;10072:100720Q.
  104. van Gestel M, Stuijk S de Haan G. Robust respiration detection from remote photoplethysmography. *Biomed Opt Express* 2016;7(12):4941.
  105. Lahiri BB, Bagavathiappan S, Jayakumar T, Philip J. Medical applications of infrared thermography: A review. *Infrared Phys Technol* 2012;55(4):221–35.
  106. Hamedani K, Bahmani Z, Mohammadian A. Spatio-temporal filtering of thermal video sequences for heart rate estimation. *Expert Syst Appl* 2016;54:88–94.
  107. Basu A, Routray A, Mukherjee R, Shit S. Infra-



- red imaging based hyperventilation monitoring through respiration rate estimation. *Infrared Phys Technol* 2016;77:382–90.
108. Mutlu K, Rabell JE, del Olmo PM, Haesler S. IR thermography-based monitoring of respiration phase without image segmentation. *J Neurosci Methods* 2018;301:1–8.
  109. Barbosa Pereira C, Czaplik M, Blazek V, Leonhardt S, Teichmann D. Monitoring of Cardio-respiratory Signals Using Thermal Imaging: A Pilot Study on Healthy Human Subjects. *Sensors* 2018;18(5):1541.
  110. Verkruyse W, Bartula M, Bresch E, Rocque M, Meftah M, Kirenko I. Calibration of Contactless Pulse Oximetry. *Anesth Analg* 2017 Jan;124(1):136–45.
  111. van Gastel M, Stuijk S, de Haan G. New principle for measuring arterial blood oxygenation, enabling motion-robust remote monitoring. *Sci Rep* 2016 Dec 7;6:38609.
  112. Vogels T, van Gastel M, Wang W, de Haan G. Fully-Automatic Camera-Based Pulse-Oximetry During Sleep. In: Proc of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW); 2018. p. 1462–70.
  113. Shao D, Liu C, Tsow F, Yang Y, Du Z, Iriya R, et al. Noncontact Monitoring of Blood Oxygen Saturation Using Camera and Dual-Wavelength Imaging System. *IEEE Trans Biomed Eng* 2016 Jun;63(6):1091–8.
  114. Wurtenberger F, Haist T, Reichert C, Faulhaber A, Bottcher T, Herkommer A. Optimum wavelengths in the near infrared for imaging photoplethysmography. *IEEE Trans Biomed Eng* 2019;9294.
  115. Rapczynski M, Zhang C, Al-Hamadi A, Notni G. A Multi-Spectral Database for NIR Heart Rate Estimation. In: Proc of the 25<sup>th</sup> IEEE International Conference on Image Processing (ICIP); 2018. p. 2022–6.
  116. Scelba G, Dragas J, Hu S, Karlen W. Improving ROI detection in photoplethysmographic imaging with thermal cameras. In: Proc of the 39<sup>th</sup> Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); 2017. p. 4285–8.
  117. Negishi T, Sun G, Liu H, Sato S, Matsui T, Kirimoto T. Stable Contactless Sensing of Vital Signs Using RGB-Thermal Image Fusion System with Facial Tracking for Infection Screening. In: Proc of the 40<sup>th</sup> Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) 2018;2018:4371–4.
  118. Gupta O, McDuff D, Raskar R. Real-Time Physiological Measurement and Visualization Using a Synchronized Multi-camera System. In: Proc of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW); 2016. p. 312–9.
  119. Hu M-H, Zhai G-T, Li D, Fan Y-Z, Chen X-H, Yang X-K. Synergetic use of thermal and visible imaging techniques for contactless and unobtrusive breathing measurement. *J Biomed Opt* 2017;22(3):036006.
  120. Blazek V, Blanik N, Blazek CR, Paul M, Pereira C, Koeny M, Venema B, Leonhardt S. Active and Passive Optical Imaging Modality for Unobtrusive Cardiorespiratory Monitoring and Facial Expression Assessment. *Anesth Analg* 2017;124(1):104–19.
  121. Chwyl B, Chung AG, Amelard R, Deglint J, Clausi DA, Wong A. SAPPHERE: Stochastically acquired photoplethysmogram for heart rate inference in realistic environments. In: Proc of the IEEE International Conference on Image Processing (ICIP) 2016;2016:1230–4.
  122. Lin K-Y, Chen D-Y, Tsai W-J. Image-Based Motion-Tolerant Remote Respiratory Rate Evaluation. *IEEE Sens J* 2016;16(9):3263–71.
  123. Fei J, Pavlidis I. Virtual Thermistor. *Conf Proc IEEE Eng Med Biol Soc* 2007;2007:250–3.
  124. Wang W, den Brinker AC, Stuijk S, and de Haan G. Algorithmic Principles of Remote PPG. *IEEE Trans Biomed Eng* 2017;64(7):1479–91.
  125. Demirezen H, Erdem CE. Remote Photoplethysmography Using Nonlinear Mode Decomposition. In: Proc of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2018; 2018: p. 1060–4.
  126. Liu X, Yang X, Jin J, Li J. Self-adaptive signal separation for non-contact heart rate estimation from facial video in realistic environments. *Physiol Meas* 2018 Jun 28;39(6):06NT01.
  127. Hanfland S, Paul M. Video Format Dependency of PPGI Signals. In: Proc of the International Conference on Electrical Engineering; 2016. p. 1–6.
  128. Zhao C, Lin C, Chen W, Li Z. A Novel Framework for Remote Photoplethysmography Pulse Extraction on Compressed Videos. In: Proc of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW); 2018. p. 1412–21.
  129. McDuff DJ, Blackford EB, Estep JR. The Impact of Video Compression on Remote Cardiac Pulse Measurement Using Imaging Photoplethysmography. In: Proc of the 12<sup>th</sup> IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017); 2017. p. 63–70.
  130. He X, Goubran RA, Liu XP. Wrist pulse measurement and analysis using Eulerian video magnification. In: 2016 Proc of the IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI); 2016. p. 41–4.
  131. Fallet S, Schoenberger Y, Martin L, Braun F, Moser V, Vesin J-M. Imaging Photoplethysmography: a Real-time Signal Quality Index. *Comput Cardiol* 2017;44:1–4.
  132. Sugita N, Yoshizawa M, Abe M, Tanaka A, Homma N, Yambe T. Contactless Technique for Measuring Blood-Pressure Variability from One Region in Video Plethysmography. *J Med Biol Eng* 2019;39(1):76–85.
  133. Kamshilin AA, Sidorov IS, Babayan L, Volynsky MA, Giniatullin R, Mamontov OV. Accurate measurement of the pulse wave delay with imaging photoplethysmography. *Biomed Opt Express* 2016;7(12):5138.
  134. Amelard R, Hughson RL, Greaves DK, Pfisterer KJ, Leung J, Clausi DA, et al. Non-contact hemodynamic imaging reveals the jugular venous pulse waveform. *Sci Rep* 2017 Jan 9;7:40150.
  135. van Gastel M, Liang H, Stuijk S, de Haan G. Simultaneous estimation of arterial and venous oxygen saturation using a camera. In: *Optical Diagnostics and Sensing XVIII: Toward Point-of-Care Diagnostics*, 2018, Vol. 10501. p. 105010W.
  136. Ghanadian H, Ghodratioghar M, Al Osman H. A Machine Learning Method to Improve Non-Contact Heart Rate Monitoring Using an RGB Camera. *IEEE Access* 2018;6:57085–94.
  137. Pursche T, Claus R, Tibken B, Moller R. Using neural networks to enhance the quality of ROIs for video based remote heart rate measurement from human faces. In: Proc of the IEEE International Conference on Consumer Electronics (ICCE); 2019. p. 1–5.
  138. Qiu Y, Liu Y, Arteaga-Falconi J, Dong H, El Saddik A. EVM-CNN: Real-Time Contactless Heart Rate Estimation from Facial Video. *IEEE Trans Multimed* 2019;9210:1–1.
  139. Dagdanpurev S, Sun G, Choimaa L, Abe S, Matsui T. Clinical Application of Multiple Vital Signs-Based Infection Screening System in a Mongolian Hospital: Optimization of Facial Temperature Measurement by Thermography at Various Ambient Temperature Conditions Using Linear Regression Analysis. In: 2018 Proc of the 40<sup>th</sup> Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) 2018;2018:5313–6.
  140. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. *Med Image Anal* 2017;42:60–88.
  141. Chen W, Picard RW. Eliminating Physiological Information from Facial Videos. In: Proc of the 12<sup>th</sup> IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017) 2017;21(3):48–55.
  142. Wang W, Den Brinker AC, De Haan G. Single Element Remote-PPG. *IEEE Trans Biomed Eng* 2018. (Early Access).
  143. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol* 2009;62(10):e1–e34.
  144. Greenspan H, van Ginneken B, Summers RM. Guest Editorial Deep Learning in Medical Imaging: Overview and Future Promise of an Exciting New Technique. *IEEE Trans Med Imaging* 2016;35(5):1153–9.
  145. Shokri R, Shmatikov V. Privacy-preserving deep learning. In: 2015 Proc of the 53rd Annual Allerton Conference on Communication, Control, and Computing (Allerton); 2015:909–10.

Correspondence to:  
Christoph Hoog Antink  
E-mail: hoog.antink@hia.rwth-aachen.de