A Survey of Remote Optical Photoplethysmographic Imaging Methods

Daniel J. McDuff 1, Member, IEEE, Justin R. Estepp 2, Member, IEEE, Alyssa M. Piasecki 3 and Ethan B. Blackford 4, Member, IEEE

Abstract—In recent years researchers have presented a number of new methods for recovering physiological parameters using just low-cost digital cameras and image processing. The ubiquity of digital cameras presents the possibility for many new, low-cost applications of vital sign monitoring. In this paper we present a review of the work on remote photoplethysmographic (PPG) imaging using digital cameras. This review specifically focuses on the state-of-the-art in PPG imaging where: 1) measures beyond pulse rate are evaluated, 2) non-ideal conditions (e.g., the presence of motion artifacts) are explored, and 3) use cases in relevant environments are demonstrated. We discuss gaps within the literature and future challenges for the research community. To aid in the continuing advancement of PPG imaging research, we are making available a website with the references collected for this review as well as information on available code and datasets of interest. It is our hope that this website will become a valuable resource for the PPG imaging community. The site can be found at: http://web.mit.edu/~djmcduff/www/remote-physiology.html

I. INTRODUCTION

In recent years researchers have presented a number of new methods for non-contact physiological measurement using digital cameras and image processing. These approaches have several strengths; they avoid the need for a device to be in contact with the skin, cameras (particularly webcams) are ubiquitous and often inexpensive, and measurements from multiple sites or multiple people can be performed concomitantly. These methods present the possibility for many novel applications of vital sign monitoring. The greatest amount of research on non-contact measurement of physiology using digital cameras has focused on measurement via photoplethysmographic (PPG) imaging. Photoplethysmography is the measurement of light reflected from, or transmitted through, the body. This signal captures the volumetric flow of blood within the vessels [1]. In most examples of remote PPG imaging, the signal has been measured from images of the face or hand, at least in part because the skin on these regions of the body is often naturally exposed.

In this paper we present a review of the work on remote PPG imaging using digital cameras. We discuss gaps within the literature and future challenges for the research community. This review specifically focuses on the state-of-the-art in PPG imaging where: 1) measures beyond pulse rate are evaluated, 2) non-ideal conditions (e.g., the presence of motion artifacts) are explored, and 3) use cases in relevant environments are demonstrated. In our search of the literature, we used the following key words: ‘remote’, ‘imaging’, ‘non-contact’, ‘physiology’, ‘digital camera’, and ‘photoplethysmography’. We found over 60 papers that have been published over the past five years related to remote PPG imaging using digital cameras. Almost a third of these were published in the latest 12 months. To aid in the continuing advancement of PPG imaging research we are making available a website with the references collected for this review as well as information on available code and datasets of interest. It is our hope that this website will become a valuable resource for the PPG imaging community. The site can be found at: http://web.mit.edu/~djmcduff/www/remote-physiology.html.

Figure 1 shows a schematic of the components within a generic PPG imaging system. A camera is used to collect images of a person, or part of a person; image processing is then performed on a computer to recover the PPG signal. In most cases illumination is from ambient light; however,
a dedicated light source may be used. Multiple imagers or light sources may be used in some cases, as well. The imaged region of interest can be segmented manually or automatically (e.g., using a face detector). Following this, the camera color channel signals are processed, typically to recover the blood volume pulse (BVP) signal. Physiological parameters can then be extracted from the recovered BVP signal (e.g., pulse rate, inter-beat intervals, pulse rate variability measures, respiration rate).

There are a number of terms used in the literature to describe this class of approach. The most common are remote PPG (rPPG), non-contact PPG (ncPPG), imaging PPG (iPPG) and PPG imaging (PPGi/PPPG). For consistency, we will use PPGi to distinguish between remote imaging methods and traditional contact PPG measurement methods.

Previous review articles [2], [3], [4], [5] have covered different aspects of the non-contact monitoring of vital signs. Hu et al. [2] surveyed developments in the measurement of PPG and contrasted traditional contact methods with burgeoning methods for non-contact and camera-based measurement. Two recent papers [3], [5] presented broad reviews of non-contact monitoring techniques which included some examples of the use of digital cameras; however, these did not consider details of all the relevant work. Liu et al. [4] presented a review of low-cost, non-contact PPG measurement techniques. There have been many new works published since 2012; we found over 30 highly relevant papers in our literature review. Measurement of ballistocardiographic (BCG) signals using digital cameras has also been proposed as another method for capturing vital signs using digital cameras [6]. We will not cover these methods in this review.

In the remainder of the paper we cover: 1) background on physiological measurement using PPGi, 2) advancements in motion and ambient lighting tolerance, 3) image optimization (including multi-spectral imaging) and region of interest optimization, 4) applications and future challenges for PPGi.

II. MEASURING PHYSIOLOGICAL SIGNALS

Estimation of physiological parameters via remote PPGi can be performed using different methods. Figure 2 shows a block diagram representing methods for calculating parameters. The BVP can be recovered from the camera signals (a) and pulse rate estimated as the average frequency or via the Inter-Beat Intervals (IBI) (b + d). Pulse rate variability (PRV) can be measured from the dynamic IBI information (e). Respiration rate can be estimated from the recovered BVP wave or via the frequency of the highest PRV spectrum peak between 0.15-0.4Hz (c + f). Blood oxygenation is typically measured directly from ratios of the raw color channel signals without estimating the BVP.

Fig. 2. Estimation of physiological parameters via remote PPGi can be performed using different methods. The BVP signal is recovered from the camera channel(s) (a) and pulse rate estimated as the average frequency or by calculation of the Inter-Beat Intervals (IBI) (b + d). Pulse rate variability (PRV) can be measured from the dynamic IBI information (e). Respiration rate can be estimated from the recovered BVP wave or via the frequency of the highest PRV spectrum peak between 0.15-0.4Hz (c + f). Blood oxygenation is typically measured directly from ratios of the raw color channel signals without estimating the BVP.
III. ADVANCEMENTS IN REMOTE PPG IMAGING

Contemporary studies in PPGi often focus on improvement over prior art given one or more imaging acquisition factors such as subject motion, ambient illumination, and image sensor spectrum sensitivity. Each of these factors, and many others, presents a particular challenge to recovering accurate physiological data given any demonstrated methodology. While the “best” methods for improving PPGi data quality may not be common for all circumstances (i.e., imaging in infrared may not be strictly comparable to visible-light imaging), it is worthwhile to consider the range of elements that could affect PPGi. Study of, and advancements made within, the following areas are necessary to understand how PPGi may be applied in natural environments.

A. Motion Tolerance

The effects of subject motion have been among the most studied dynamics in PPGi measurement. Early investigations into PPGi focused on rigid, stationary regions of interest [7], [22], [27], [28], [8]. Some later studies focusing on imaging of the subject’s face allowed for limited naturalistic motion [9], but many implemented protocols to strictly limit motion during imaging [19], [11]. Other studies have investigated algorithm performance under limited translational motion using techniques such as simple region of interest focused object tracking [13] and color difference, or chrominance-based, signals from RGB color space [29]. Approaches for estimating the contribution of motion artifacts and correcting the PPGi signal using adaptive filtering have also been explored [10].

Although the impact of subject motion on PPGi signal quality has been observed and characterized [9], studies implementing a systematic investigation of subject motion and its impact on PPGi signal quality have been lacking in relevant literature until recently. De Haan and van Leest [14] examined pulse rate measurement during exercise on five different fitness devices and presented a new approach for PPGi based on linear decomposition of color channel information constrained by optical properties of the imager setup and illumination source spectrum. This new measure, termed the blood volume pulse vector ($\vec{P}_{bh}$), was also tested on a large corpus (N=117) of stationary subjects and shown to improve pulse signal quality over the previously presented chrominance method [29]. Estepp, Blackford, and Meier [12] presented one of the first carefully controlled subject motion studies focused on rigid motion rotation, as opposed to translation, of the subject’s head. In this work, a novel, multi-imager array was used in conjunction with blind source separation to reduce the effects of rigid head motion artifact in the measurement of pulse rate using PPGi. Alternative approaches taking advantage of spatial redundancy within a single imager, as opposed to multiple imagers, have also been proposed for mitigating a variety of motion-induced artifacts including translation, scaling, rotation, and talking [30]. A comparable framework has been extended to include multi-imagers in the infrared spectrum, as well [31]. In other work, Chung et al. [17] presented a novel, motion-compensated PPGi method based on motion information from location tracking, showing that accounting for motion information, as opposed to not, improved results.

Motion artifact should continue to be a focus of exploration for PPGi, with specific improvements concerned with three major areas: 1) image processing algorithm development, 2) spatial redundancy (in both single- and multi-imager sensor systems), and 3) the use of integrated, multi-band approaches in visible and infrared wavelengths. While systematically varying motion artifacts is of value to explore, use cases in applied environments including but not limited to exercise [14], cognitive/affective states [32], and clinical care [33], [34] should also be considered with a mind toward transition out of the laboratory environment. Lastly, as expressed elsewhere in this review, the effects of motion artifact should be extended to cardio-pulmonary measures beyond pulse rate alone.

B. Ambient Lighting Tolerance

While ambient lighting conditions may be considered relatively constant in most applications, there are some potential use cases (e.g., driving, computer simulation and virtual reality environments, etc.) where environmental lighting can be highly variant. While changes in illumination intensity have been shown to affect absolute, but not relative, magnitude of the PPGi waveform [11], its effects on any currently available PPGi method are largely unknown. In a limited, uncontrolled study, Li et al. [35] used an adaptive filtering approach, with an isolated background region of interest serving as the input noise reference signal, to compensate for background illumination. Their correction for varying illumination using a publicly available video database [36] resulted in a modest overall reduction in heart rate error (when compared to a contact electrocardiogram (ECG)); however, pairwise comparisons of their multi-stage processing approach (region of interest detection, illumination variation correction, data
pruning, and temporal filtering) were not made available in order to sufficiently determine the effectiveness of any single stage. Amelard et al. [18] recently presented results of PPGi measurement and pulse rate estimation in dynamic ambient lighting conditions. Their system, however, required a temporally-coded light source and synchronized camera, which is novel in comparison to most imaging testbeds developed for PPGi measurement. Aside from uncertainty resulting from dynamic illumination conditions, the qualities and properties of constant illumination used for PPGi measurement that are necessary to produce physiological measurements of adequate quality are not well known. From a practical perspective, these properties may not be different than those that would ensure a reasonable image quality given the ability to adjust lens aperture and focus, sensor sensitivity (ISO), and individual frame exposure time (integration, shutter speed). To our knowledge, PPGi in imagery that would be considered of “poor quality” by any one of the aforementioned factors has not yet been explored but could be of interest in some application areas.

C. Image Optimization

The qualities of an imager, or image sensor, that allow for robust PPGi are not easily understood given the variety of image sensor properties that could be considered. These properties include, but are not limited to, basic sensor type (e.g. CCD, CMOS, and other custom designs), color filter array (e.g. Bayer, Fovenon X3, and RGBE), bit depth, imager size, and number of pixels. Qualities of the overall image not necessarily attributable to the image sensor itself, such as lens type and quality, spectral properties of the illumination source, and image aperture/shutter speed/ISO, additionally affect the overall content of any acquired image. Although not directly related to image sensor properties, the frame rate at which the video is captured should also be considered. How the imager setup is used in data acquisition is also relevant to PPGi quality; pixel density of the imaged tissue can vary dramatically with the distance from imager to subject. Evidence toward the detrimental effects of image/video compression algorithms on PPGi signal recovery have likely been seen in unpublished data, but published data in this topic area has yet to make it to the seminal literature. Some studies exist on the comparison of multiple imagers running in parallel during data collection (e.g. [11]) and offer some confidence that PPGi signal recovery can be robust over widely varying imager properties. Similarly, some promising work also exists on the study of image size (pixel density) and frame rate in single-imager [20] and multi-imager [37] sensor designs. Needless to say, a systematic variation of the fully-crossed parameters discussed here, as well as potentially numerous others, would likely be too cumbersome to be practical (and, ultimately, may provide very little utility).

A great contribution would be the creation and standardization of an explicit benchmark test, and related metrics, that could be performed with a variety of imagers to better understand and compare results across studies and methods. Benchmark test results could be provided in text, or as supplementary material, and potentially explored in aggregate as a future meta-analysis. Instead of taking the approach of describing detailed results for imaging methods that produce successful results, a forum for publishing “null results” (those particular imagers, techniques, and methodologies that do not successfully demonstrate PPGi measurement capability) could be of value to the community, as well. With a sizeable null result database, possibly paired with the same benchmark test for positive results, additional future meta-analyses could prove to be as informative for determining worst case imaging scenarios as with those considered to be the best case. Specifically a benchmark dataset should include detailed imaging device, lighting and participant demographic information. In addition to videos and gold-standard contact measurements (ideally both BVP and ECG).

D. Multi-Spectral Imaging

Many contemporary PPGi studies focus on RGB imaging in the visible light spectrum. While green/orange visible bands are the most absorptive for common derivatives of oxy- and deoxy-hemoglobin [38], multi-spectral imaging from a single image sensor has often been used for PPGi methods involving linear decomposition or blind source separation using multiple data channels. As demonstrated by Martinez et al. [39] using spectrophotometry of the forehead, certain wavebands are better for PPGi-derived pulse rate and respiration rate measurement than others when considering signal-to-noise ratio of the desired physiological phenomenon. Given these results, one can reason that spectral bands common in most digital imagers may be somewhat limited in resolution, range, and sensitivity for PPGi.

While hyperspectral imaging, or imaging spectroscopy, would be most ideal for determining the most suitable multi-spectral montages for PPGi, imagers with this capability are limited in practice due to their cost, mismatched temporal resolution (push-broom vs. snapshot, or global shutter, image acquisition) and lack of ubiquity (in comparison to standard digital cameras in the visible and near-infrared ranges). Despite these limitations, investigations into multi-band imaging, as well as imaging outside of the visible and near-infrared ranges, have proved fruitful. As an example, a five-band visible band camera, with cyan and orange pixels present (in addition to standard RGB) produced results that were optimal for waveband combinations including orange, green, and cyan [40]. Mid-wave infrared imagers used in imaging thermography applications have also been shown as capable for primarily carotid-focused pulse rate measurement [41]. Future work may seek to explore multi-band, and potentially multi-imager, applications that leverage spectrally-tuned approaches to integrating multiple wavebands. For example, this could be achieved through the use of specifically tuned optical filters across a small, spatially-redundant array of both visible and infrared imagers.

E. Region of Interest Optimization

Identifying regions of interest for PPGi measurement has been used both for isolating relevant patient tissue (e.g. [27],
With reasonably static, stationary subjects, pulse waveform amplitude (indexed by pulse rate frequency band power in the frequency-domain) can be mapped to small regions of the imaged tissue. Those spatial locations with the largest pulse waveform amplitude are those that have the largest overall contribution to the PPGi waveform. Lempe et al. [42] demonstrated this technique and revealed the most salient facial locations for PPGi measurement to be the carotids (on the neck), suborbital regions, and the central forehead superior to the nasion. Although individual subjects exhibited some variability in the exhibition and relative strength of the PPGi waveform, isolation of these regions improved the signal-to-noise ratio of the pulse rate peak observed in the frequency domain. It reasons to believe that other methodologies for obtaining the PPGi waveform would also benefit from an improvement in region of interest selection in facial imagery. Lempe proposed using adaptive facial regions for this purpose.

IV. IMAGE ENHANCEMENT & VISUALIZATION

The cyclic changes in light absorption and reflection resulting from the recurrent delivery of the BVP to the superficial microvasculature are very small and often imperceptible to the unaided eye. Image processing techniques have been proposed to support visualization of hemodynamics. The techniques rely on analyzing the spatial or temporal dynamics of the PPG signal and amplifying those dynamics. One of the most notable visualization techniques is Eulerian Video Magnification (EVM) [26]. This method applies a spatial decomposition, followed by temporal filtering and selective frequency amplification for successive video frames and can be used to highlight subtle color changes and motions. The spatial decomposition pools pixel regions and thus improves the temporal signal-to-noise ratio to overcome image sensor and quantization noise. The selective amplification of temporal frequency bands magnifies these dynamics and allows them to be easily perceived in the video sequence. For example, amplying frequencies in the range of 0.4-4Hz (corresponding to 24-240 beats per minute) has been effective at highlighting the color changes associated with the BVP and produces a pulsatile, red flushing of the face associated with the cardiac cycle. The resulting video can be visually interpreted for spatial and temporal dynamics of the BVP, such as symptoms of underlying circulatory issues such as asymmetry of pulse wave propagation [26].

Another method for visualizing the hemodynamics of the BVP uses Fourier analysis of successive pixels or pixel regions over time. The resulting Fourier coefficients, power spectrum, and phase for each pixel or region are then mapped to a false color image, highlighting the intensity differences within the image. Similar to EVM, implementations may incorporate spatial filtering, averaging, pooling or region tracking and alignment to increase signal-to-noise ratio and minimize artifacts. The resulting power spectrum and phase maps highlight the perfusion and hemodynamics of the skin and can be used to provide clinically relevant information. This approach was shown by Wu, Blazek, and Schmitt [43] during vein occlusion and muscle pump tests. Wieringa, Mastik, and Van der Steen [7] combined this technique with multiple wavelength imaging to allow the visualization of spatial power and phase dynamics of blood oxygen saturation [7]. Other applications have used this method to analyze the spatial hemodynamics of: vascular skin lesions (e.g., port wine stains) [8], skin regions under local anesthetic [44], dynamics during individual cardiac cycles and local irritations (e.g., small scratches) [45], for identifying regions of interest on the face with greatest SNR for the BVP signal [42], and during allergic response testing [46].

V. APPLICATIONS

There are many potential real-world applications of remote PPG imaging. Klaessens et al. [34] presented the possibility of a baby-friendly non-contact measurement of various vital signs (including PR, breathing, skin temperature, and SpO2) in infants and neonates located in an Intensive Care Unit (NICU). The study employed the use of RGB color magnification to measure PR and IR-thermography to measure RR of a total of seven infants with gestational ages of 24 to 39 weeks. Their results provided preliminary evidence of the feasibility of remote monitoring of vital signs in infants. Aarts et al. [33] also addressed applications of remote PPG measurement in NICU environments. The study involved vital sign recordings of 19 infants with gestational ages of 25 to 42 weeks. The experimental setup involved data collection with a standard color digital camera. A region of interest (ROI) was manually selected in an initial video frame which was then used as a global tracker for the remaining frames in order to track the movement of the infant. Results of the study indicated successful monitoring of PR; however, noise caused by motion and/or poor illumination conditions interfered with continuous measurement.

Tarassenko et al. [23] presented the application of PPGi for vital sign monitoring of PR, RR, and SpO2 of hemodialysis patients in a Kidney Unit. A total of 46 patients were monitored using RGB color imaging during 133 dialysis sessions in order to attain vital signs, with the goal of accurate vital sign measurement over a four hour dialysis session. A non-parametric Bayesian image segmentation algorithm was used to localize the ROIs, and novel methods of autoregressive (AR) modeling and pole cancellation were used to remove aliased frequency components associated with artificial light flicker. Results showed the efficacy of AR modeling in order to eliminate unwanted frequency components of strong fluorescent lighting often found in clinical environments. Additionally, accurate vital signs were attained from patients in the clinic who may have had extraneous factors affecting vital signs, such as diabetes or obstructive sleep apnea. RR was estimated via breathing-synchronous changes in PPG amplitude, as opposed to PR which may have fluctuated in response to the patients’ additional comorbidities.

Kwon et al. [47] presented a smartphone application for recording vital signs, providing a low-cost, easily accessible
TABLE II
SUMMARY OF APPLICATIONS OF REMOTE PPG IMAGING. THE TABLE DESCRIBES THE APPLICATION, SUBJECTS TESTED, PHYSIOLOGICAL PARAMETERS MEASURED AND THE DEVICE USED FOR MEASUREMENT.

<table>
<thead>
<tr>
<th>Paper</th>
<th>Application</th>
<th>Subjects</th>
<th>Phys. Params</th>
<th>Device</th>
</tr>
</thead>
<tbody>
<tr>
<td>Klaessens et al. [34]</td>
<td>Monitoring in NICU</td>
<td>7 infants (ages from 24 to 39 weeks)</td>
<td>HR, RR</td>
<td>RGB: DSC-H9 Sony</td>
</tr>
<tr>
<td>Aarts et al. [33]</td>
<td>Monitoring in ICU/NICU</td>
<td>19 infants (ages from 25 to 42 weeks)</td>
<td>HR</td>
<td>uEye, IDS imaging</td>
</tr>
<tr>
<td>Tarassenko et al. [23]</td>
<td>Kidney Unit</td>
<td>46 patients during 133 4-hour dialysis sessions</td>
<td>HR, RR, SpO2</td>
<td>Grasshopper2 GigE</td>
</tr>
<tr>
<td>Kwon et al. [47]</td>
<td>iPhone app</td>
<td>10 healthy subjects</td>
<td>HR</td>
<td>iPhone camera</td>
</tr>
<tr>
<td>McDuff et al. [32]</td>
<td>Cognitive stress measurement</td>
<td>10 healthy subjects</td>
<td>PR, RR, PRV</td>
<td>Five-band camera</td>
</tr>
<tr>
<td>Burzo et al. [48]</td>
<td>Emotion valence recognition</td>
<td>14 healthy subjects</td>
<td>HR</td>
<td>Webcam</td>
</tr>
</tbody>
</table>

VI. FUTURE CHALLENGES

Remote measurement of vital signs using digital cameras and ambient light presents significant promise. As highlighted by this survey a number of applications have been proposed that use the approaches that have been developed. However, there still remain a number of key challenges for the community. Whilst many have presented results for PR measurement, only a few studies have evaluated performance for PRV, PTT, respiration and SpO2. Accurate measurement of average PR is not likely to be enough for many of the compelling applications that remote PPGi presents. One of the main hinderances is data. As mentioned above, benchmark datasets are essential for comparative analysis of different algorithmic approaches. Soleymani et al. [36] have released the one of the few video datasets with accompanying physiological "gold standard" measurements. However, this dataset was not explicitly released for the purposes of PPGi comparison. A dataset with varying amounts of subject and camera motion, examples of static and dynamic illumination and a diverse subject pool would be extremely beneficial to the community. Several datasets may be necessary as a range of imagers and set-ups should also be tested.

VII. CONCLUSIONS

In recent years researchers have presented a number of new methods for recovering physiological parameters using just low-cost digital cameras and image processing. In this paper, we have presented a survey of the work on remote PPG imaging using digital cameras. Motion tolerance, illumination tolerance and imager optimization have all received attention and considerable advancements in performance have been made in only a few years.

There remain a number of future challenges for the field. More work should address physiological parameters beyond average PR, as this will greatly increase the utility of PPGi methods. A dataset with varying amounts of subject and camera motion, examples of static and dynamic illumination, and a diverse subject pool would be extremely beneficial to the community. Such a resource would allow for comparative analysis of different algorithmic approaches. The ubiquity of digital cameras presents the possibility for many new low-cost applications of vital sign monitoring. Several studies have validated the efficacy of this technology in real-life settings; however, these are still limited to proof-of-concept type experiments and thorough clinical trials are necessary.

REFERENCES
