ARL-TN-0962 ● SEP 2019

DEVCOM ARMY RESEARCH LABORATORY

Remote Photoplethysmography for Cardiovascular Monitoring

by Steven Gutstein

NOTICES

Disclaimers

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Citation of manufacturer's or trade names does not constitute an official endorsement or approval of the use thereof.

Destroy this report when it is no longer needed. Do not return it to the originator.

ARL-TN-0962 ● SEP 2019

Remote Photoplethysmography for Cardiovascular Monitoring

by Steven Gutstein
Human Research and Engineering Directorate, CCDC Army Research Laboratory

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED (From - To)
September 2019	Technical Note	November 2018–July 2019

4. TITLE AND SUBTITLE
Remote Photoplethysmography for Cardiovascular Monitoring

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Steven Gutstein

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
CCDC Army Research Laboratory
ATTN: FCDD-RLH-FC
Aberdeen Proving Ground, MD 21005-5425

8. PERFORMING ORGANIZATION REPORT NUMBER
ARL-TN-0962

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
ORCID ID: Steven Gutstein, 000-0001-7537-4523

14. ABSTRACT
Half of the deaths during Operations Iraqi Freedom and Enduring Freedom due to potentially survivable injuries were caused by noncompressible hemorrhaging. Photoplethysmography can diagnose this condition well before it becomes critical, without the use of magnetic resonance imaging, computed tomography scans, or ultrasound imaging. Remote photoplethysmography (rPPG) offers the possibility of doing this with a single webcam rather than a specialized piece of equipment that must be worn by the Soldier being monitored. rPPG may also be used to monitor other individual vital signs. This technical note examines state of the art for rPPG along with some actual and potential applications.

15. SUBJECT TERMS
photoplethysmography, remote photoplethysmography, PPG, rPPG, noncompressible hemorrhaging

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Steven Gutstein
Unclassified	Unclassified	Unclassified	UU	37	19b. TELEPHONE NUMBER (Include area code) 410-278-5856

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

Contents

List of Figures

List of Tables

1. Introduction

1.1 Army Applicability

Half of the deaths during Operations Iraqi Freedom and Enduring Freedom due to potentially survivable injuries were caused by noncompressible hemorraghing[1] (i.e., approximately 60 deaths per year on average). This bleeding tends to occur within the torso, where it can neither be seen nor readily stopped. Furthermore, because the human body is very resilient to blood loss, this life-threatening condition is not reliably detectable by most readily obtainable vital signs (e.g., blood pressure, pulse rate, blood oxygenation levels, and mental state). Although blood loss can be detected via magnetic resonance imaging (MRI), computed tomography (CT), or ultrasound imaging, the necessary equipment is generally limited to a hospital setting, and, in the case of MRI and CT scanning, is not an option for a patient with embedded shrapnel. As a result, one must wait until a person loses about 40%–50% of their blood before standard vital signs give a reliable indication of this condition.[2,3] These indicators tend to be quickly followed by the onset of shock and then death. Early detection would be the key to reduced mortality.

Photoplethysmography (PPG) is an optical technique used to monitor blood flow. Because human skin is slightly translucent, incident light interacts with both the skin's surface and subcutaneous tissue. As a result, the skin's hue oscillates with each heartbeat. These oscillations create the PPG waveform and may be measured with a photosensor placed directly on the skin, or remotely with a camera. The PPG waveform reflects the arterial pressure waveform, which indicates the state of a person's cardiovascular system and can give early warning of dangerous medical conditions.

It has been observed that machine-learning techniques can be applied to both the PPG waveform and the arterial pressure waveform to determine when an individual has lost a significant fraction of their blood (i.e., hypovolemia) or is in the process of doing so.[2,4,5] In order to take advantage of this, the Air Force has been developing field-wearable PPG sensors.[3] A remote PPG (rPPG) approach would not only relieve individual Soldiers of the need to wear their own monitoring device but also enable medics to remotely diagnose noncompressible hemorrhaging for several patients with only a single, lightweight piece of equipment. Furthermore, recent work has indicated that these measurements could be obtained via a camera mounted on an unmanned aerial vehicle (UAV).[6] This would increase the ability to begin quickly monitoring and diagnosing injured Soldiers.

PPG waveforms can also be used to measure several standard vital signs:

- Pulse rate[7,8]
- Pulse variability/interbeat interval (IBI)[9–11]
- Respiratory rate[12,13]
- Glucose levels[14]
- SpO_2–peripheral capillary oxygenation saturation (percentage of oxygenated hemoglobin)[15]
- Blood pressure[16–18]

The accuracy of PPG measurements tends to be as good as, or comparable to standard electrocardiogram (ECG or EKG) measurements for vital signs such as pulse rate, variability, and blood pressure.[10,18,7,19]

PPG may also be used to infer these two characteristics:

- Biometric ID[20,21]
- Changes in stress, both physiologic and psychologic[22,23]

This information can be used not only to identify people but also to determine how well they are reacting to their immediate situation and how comfortable and capable they are of handling it. This would help to optimally match individuals to tasks.

PPG measurements are most commonly made with a pulse oximeter, which is usually placed on a fingertip or ear lobe (Fig. 1). This entails shining light through the skin and monitoring the transmitted light's spectrum.

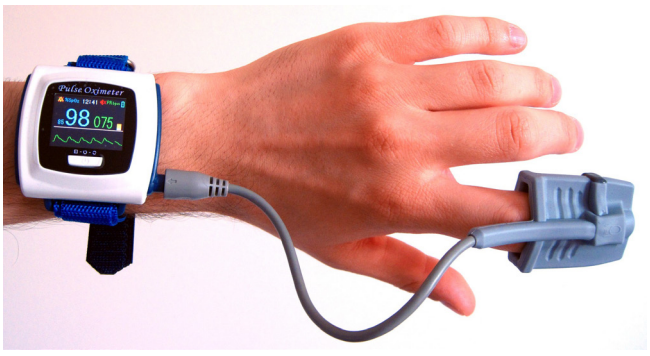

Fig. 1 Finger cuff pulse oximeter. Wrist monitor shows blood oxygenation, pulse rate, and PPG signal. (Wikimedia Commons contributors. File:Wrist-oximeter.jpg. Wikimedia Commons, the free media repository; 2018 Nov 27, 15:59 UTC [accessed 2019 July 29]. https://commons.wikimedia.org/w/index.php?title=File:Wrist-oximeter.jpg&oldid=329545296.)

Although pulse oximeters are relatively noninvasive, they require direct contact with the subject. rPPG sensors do not. They encompass sensors such as standard webcams, thermal imaging cameras (near and far infrared), cameras with nonstandard color responses (e.g., RGBCO, where R = red, G = green, B = blue, C = cyan, O = orange), and radar.

When cameras are used, subjects are no longer encumbered by any sensor(s). Furthermore, one camera can easily do the work of several sensors, meaning less equipment needs to be provided, maintained, and transported for a group of individuals. Finally, the use of cameras creates the possibility of surreptitiously making these measurements. Because standard webcams are easily available, they have the greatest potential for general utility. For this reason, the focus of this technical note is on the methods that use them and the results obtained.

1.2 Overview

The explanation and descriptions of rPPG techniques in this technical note are at a relatively high level. Mathematical detail is eschewed in favor of providing heuristic explanations of current techniques. References are provided so that the interested reader may acquire the detail needed to implement any method desired. Furthermore, because of the challenges inherent in comparing results among different experimenters developing different methods under various conditions, results and comparisons are included in the final section of this report.

rPPG may also be referred to as “Imaging PPG” (iPPG), “Video PPG” (vPPG), or other equivalently descriptive terms. Herein the term rPPG will be exclusively used. Some rPPG techniques rely on motion detection, either observing the pulsation of veins such as the carotid or gross motion of the head caused by the force of blood being pumped into it.[24] These approaches will also be ignored in favor of concentrating on more robust methods that rely solely upon observing changes in the hue of a subject’s skin.

In order to obtain a signal from a video, the rPPG method has three main steps:

1) Detect and track a region of interest (ROI).
2) Extract the PPG signal.
3) Make the desired measurement.

Because the research portion of rPPG analysis focuses on the first two steps, this technical note focuses on them more thoroughly in the Sections 2 and 3.

2. ROI Detection

Generally, a meaningful signal is most easily extracted from a person's face. So, the first task of an rPPG method is usually face detection. Although this can be done with a combination of manual labeling and tracking methods, significant progress has been made in face detection (and recognition) over the past two decades. One of the more common methods is the Viola–Jones[25] object detection approach, which was used by Verkruysse et al.[13] This technique relies on splitting an image into many subregions and using several rectangular filters as feature detectors. The filters are ranked in terms of their ability to eliminate subregions that do not contain a face without mistakenly eliminating those that do. As each filter is applied to the various subregions, more and more get rejected as not containing a face. Eventually, the surviving subregions all have high probabilities of containing a face.

Additionally, recent advances in deep learning allow for several new approaches to face detection. These generally involve training deep convolutional neural nets, which do not need handcrafted filters (e.g., Liu et al.[26] and Zhang et al.[27]) Deep-learning-based approaches are rapidly becoming standard approaches for image-based tasks like face detection.

Yet, the real issue is not solely face detection. The key goal is to identify a large region of skin pixels possessing strong pulse signals. The larger this region, the more camera quantization noise will be mitigated.[8] Certain regions of the face give stronger, clearer signals than others.[22] By concentrating only on the regions with the strongest signal, one can improve the signal-to-noise ratio (SNR). In addition to detecting and tracking faces, one can track facial landmarks to better divide a face into regions with strong signal.[28,29]

Rather than directly looking for a face, which may or may not have its subregions with stronger signal extracted, one can also directly look for the pixels exhibiting intensity variations at near pulsatile temporal frequencies in the expected spectral range. Schmitz's[30] initial work along these lines was cited by van Luijtelaar et al.[29] in their paper extending Schmitz' work. van Luijtelaar et al. used feature corners (as determined with a Harris Corner detector)[31] to divide an image into triangles, which they could then track.

In a similar vein, Bobbia et al.[32] divided an image into temporal superpixels (i.e., superpixels that had temporal as well as spatial similarities) and then selected the superpixels displaying the greatest signal. An advantage of these approaches is that they make it easier to get signals without a frontal face view.

3. Extracting the PPG Signal

Once the ROI has been located, the signal must be extracted. The signal of interest is fairly faint, as is evidenced by the fact that it is not visible to the naked eye. SNR has a narrow peak of about 1:24 (i.e., –2.76 dB)[33,34] near the yellowish end of the green spectral region, which can be seen in Fig. 2. There are several factors that contribute to the total signal[8]:

1) Intensity of light source and the spatial relations among it, the target skin and detector
2) Specular reflection–reflection from the skin's surface
3) Diffuse reflection–reflections from beneath the skin's surface
    a) Pulsate reflection–reflection due to absorption and scattering from blood (i.e., rPPG signal)
    b) Stationary skin-tissue reflection–reflection due to absorption and scattering from other subcutaneous tissue
4) Noise–variations in environmental light and in the sensor's response to the light

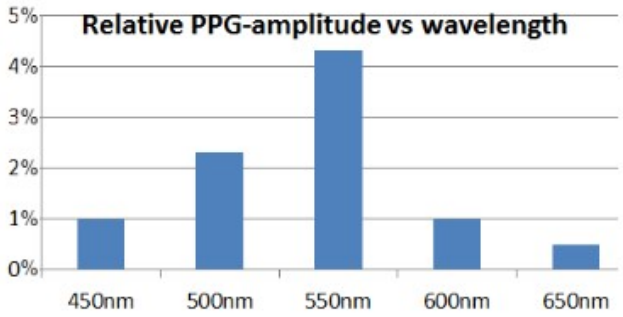

Fig. 2 Graph of light reflected by skin as a function of wavelength as reported by Crowe and Damianou[35] and reproduced in de Haan and Jeanne.[33] Note the strong peak at green-yellow color and decrease at red wavelengths. Since at peak, the rPPG signal is one twenty-fifth of the full signal, the SNR is about 1:24 (i.e., –2.76 dB).

The main methods for extracting the pulsate signal from the reflected light are as follows:

1) Blind source separation techniques: Express the total signal as a linear combination of separate components, then select the component most nearly matching expected pulse temporal frequencies.

2) Color filtering: Use known absorption properties of hemoglobin to extract the color frequencies of the total signal that primarily carry the pulsate information.

3) Signal amplification: Magnify temporal changes in the signal so as to increase the pulsate signal.

4) Machine-learning techniques: Using a large number of known total signals, train an algorithm to extract the known pulsate signal.

3.1 Blind Source Separation

If we assume a relatively steady incident light intensity, then the problem can be viewed as one of blind source separation (BSS)—identifying one component of a linear combination of sources. There are at least four BSS approaches that have been applied to the problem of extracting rPPG signal extraction:

1) Independent component analysis (ICA)

2) Principal component analysis (PCA)

3) Zero-phase component analysis (ZCA)

4) Joint blind source separation (JBSS)

3.1.1 Independent Component Analysis

ICA assumes that non-Gaussian, statistically independent components create the received signal. In this case, it is a fairly reasonable assumption, and ICA has been used by several groups to remotely extract a PPG signal.[22,36–39]

There are, however, difficulties inherent with ICA. First, it is a computationally expensive technique; second, it requires one to both guess the correct number of sources comprising the signal and have at least the same number of separate sampling sources as of independent components; and finally, there is the ”sorting” problem. ICA will break a signal up into a number of independent components, but it will not determine which component is the one of interest. This invariably leads to searching for a component exhibiting a strong frequency response in the expected range of heart rates. However, if the subject is displaying some periodic motion with the same frequency, or there is some other environmental noise with the expected frequency, then the correct signal is hard to disentangle.

Poh et al.[37] used the RGB channels of a single webcam to provide three source components. They then assumed that the non-DC component whose power spectrum had the highest peak was the desired signal. A pulse’s strong periodicity

makes this a reasonable assumption. Zhao et al.[39] use embedded time delay vectors to provide an arbitrary number of mixed source vectors, which allowed them to find an arbitrary number of source components. Also, by moving away from using the RGB color planes, they were able to work both in darkened conditions and with gray scale images. Liu et al.[22] took a different approach and divided the face into 10 regions, which provided 10 sources for their ICA, but only found 5 significant individual components. The main contributions to the independent components came from the forehead, left cheek, and left chin. This result gives a good indication of which ROIs should be used if one does not average over the entire face.

The method of constrained ICA (cICA)[40] can be used both to place additional restrictions on the derived independent components and to impose an ordering, thus avoiding the sorting problem. These constraints can involve the a priori knowledge of the pulse's temporal frequency. This approach was used by Tsouri et al.,[38] who used cICA to find a component with a primary frequency as close as possible to a predefined reference frequency. By sweeping through a range of plausible reference frequencies, they found the independent component with the best fit to its chosen reference. Although greater accuracy was achieved, there was also an increase in computational effort required. Macwan et al.[36] avoided this extra cost, by instead requiring ICA to extract the most auto-correlated independent component possible. Again, this uses the a priori knowledge that pulses are highly periodic. Additionally, Macwan et al.[36] used information about the color bands most affected by the subcutaneous blood flow.

3.1.2 Principal Component Analysis

PCA takes a signal, in an n-dimensional space, and rotates it to maximize the variation along each principal axis. The signal will have its greatest variation along the first principal axis, its next greatest along the second principal axis, and so on, until n principal axes have been determined. Again, since the subject's pulse is expected to be the major cause of variation in the received signal, the first principal component should be the desired PPG signal.

This approach was used successfully by Lewandowska et al.[41] However, although ICA gives better results than PCA, it requires longer signals in order to determine the statistically independent components. This makes PCA more attractive for real-time applications.[8,42]

3.1.3 Zero-Phase Component Analysis

ZCA is used to "whiten" data, so that the variance along each principal axis is normalized, before returning to the original coordinate axis. There are several machine-learning techniques for which data whitening improves learning. Iozzia et

al.[43] used ZCA to extract pulse information, and they obtained slightly better accuracy than with the CHROM[33] method, which is popular color-filtering method to be described later herein.

3.1.4 Joint Blind Source Separation

In general, most BSS techniques can only accommodate one data set. However, some of the nonpulsatile lighting variations that appear in a given ROI also appear in the background. The presence of these variations in the background is, in fairly obvious sense, an indicator that they are nonpulsatile. JBSS is notable in that it is capable of meaningfully integrating several data sets as it seeks to isolate the various sources for a given signal. Cheng et al.[28] used this technique to identify dependent components between the facial ROI and nonfacial regions of the image. This makes it possible to remove variations caused by variations in the lighting intensity.

3.2 Color Filtering

BSS techniques use a priori knowledge of human pulse frequencies to extract the desired rPPG signal. This makes them particularly susceptible to motion-based noise and other periodic noise sources. However, in addition to the expected pulse frequency, another piece of a priori information can be used——the absorption properties of hemoglobin. Hemoglobin has its peak absorption in the green-to-yellow range of the spectrum. There are several approaches that use this prior knowledge to more effectively extract the pulse.

At first glance, relying purely on chromatic information seems inherently risky. As seen in Fig. 2, the rPPG signal is only a fairly small percentage of the total signal (i.e., light reflected from the skin). Most of the skin's reflected light is "noise" with respect to the information we are trying to obtain. The steady-state components of this noise come from the specular reflection, the DC component of the diffuse reflection, and the spectra of the incident light. The time-varying components come from relative motion of the light source, subject, and camera.

An additional source of complexity is that people can have widely varying skin tones. The Fitzpatrick scale divides human skin tone into six classes ranging from Class I (the lightest) to Class VI (the darkest). Fortunately, there are two reasons these variations are not significant obstacles to obtaining the rPPG signal with spectral filtering. First, the rPPG signal is determined by the absorption properties of hemoglobin, which are independent of skin tone. Second, the major factor creating skin tone is the amount of melanin in the skin. Although different amounts will cause skin to be lighter or darker, the absorbed spectra are still the same.[44]

However, darker skin tones tend to lead to lower SNR since extra melanin will reduce the diffuse reflection, which contains the rPPG signal, but will not affect the specular reflection, which is for us the noise.[33]

There are three main approaches to color filtering examined in this technical note:

1) RGB color filtering: Finding the RGB color space vector most nearly aligned to the rPPG signal.
2) Non-RGB filtering: Using a non-RGB color space to find a vector most nearly parallel to the rPPG signal.
3) Chromatic-temporal filtering: Combining chromatic and temporal a priori knowledge to isolate the rPPG signal.

3.2.1 RGB Color Filtering

Green light is both absorbed strongly by hemoglobin (unlike red light) and can penetrate deeply through skin (unlike blue light).[13] So, the simplest RGB color-filtering technique is to use only the G color plane to extract the PPG signal.

Yet, this ignores all the known sources of noise. The DC component can be removed by dividing each color channel by its temporal average and then subtracting 1, to center each color channel's signal at 0.[8,33]

Motion-induced noise is independent of color channel (i.e., if the subject moves, then the multiplicative factor by which the diffusely reflected light changes is independent of color). So, movement-induced noise can be mitigated by measuring the ratio between the reflected G and R signals, which de Haan et al.[33] say is related to a method suggested by Huelsbusch.[45] They refer to this method as RoverG. (Note: This is somewhat counterintuitive nomenclature, since the published ratio has G in the numerator and R in the denominator.)

However, specularly reflected light contributes an additive component to noise. If one assumes the illumination source is white light, then specular reflection affects each color channel by adding an identical increment to each channel. de Haan et al.[33] eliminated this by using color differences (i.e., chrominance). Instead of taking the ratio of the G and R signals, they constructed two orthogonal chrominance signals:

1) $X = R - G$
2) $Y = 0.5R + 0.5G - B$

These were not randomly chosen. X was designed to be parallel to the pulsatile signal, whereas Y is perpendicular to it. This gives rise to the XoverY method, which, as stated, assumes white light illumination.

It is possible to mitigate the white light assumption, by scaling the R, G, and B data to account for the skin's melanin, prior to creating the X and Y vectors. Furthermore, these ratio techniques can be closely approximated by a linear combination of signals, which is more similar to the BSS methods. This is the basis of the three following techniques for which a geometric interpretation in color space gives a strong heuristic understanding of their efficacy:

1) CHROM: Chrominance[33]
2) POS: Plane orthogonal to skin[8]
3) PBV: Pulse blood volume[43]

A unifying view of these approaches was given by Wang et al.[8] In this work, the RGB color space is explicitly thought of as a geometric space in which the signal moves. All three approaches involve first normalizing out the temporal mean of the signal in order to remove its DC component, thereby centering the signal mean at 0. They differ in the way they proceed from there.

CHROM, which takes its name from its use of chrominance signals, next normalizes the signal with respect to an average skin tone. Then, the signal is projected onto a plane perpendicular to the specular reflection, thus eliminating that noise source. As principal axes of this plane, an axis approximately parallel to the pulsate variations (i.e., S1) is chosen and one that is approximately perpendicular to them (i.e., S2). Through a process referred to as "alpha-tuning", one weights S2 by the ratio $\alpha = \frac{\sigma(S_1)}{\sigma S_2}$ and subtracts it from S1 to get an estimate of the PPG signal.[8]

POS is a variation of CHROM, which first projects the signal onto a plane perpendicular to the temporally normalized skin tone, instead of the specular reflection. This not only removes motion-induced noise, which is a major noise source, but also affects the RGB channels equally. Then, as before, two principal axes are defined in such a way that one expects the PPG signal to be concentrated along one of the axes. Then, through alpha-tuning, the PPG signal is obtained.[8]

PBV is conceptually the simplest of the three techniques. It projects the signal component along the pulsatile vector and ignores any variations not in the pulsatile direction. In essence, PBV is a "greedy" method, which uses knowledge of the chromatic direction along which hemoglobin exhibits the most absorption. In contrast, CHROM and POS both attempt to first remove sources of noise, before attempting to capture the desired signal. CHROM first removes the specular noise

by projecting the signal onto a chromatic plane perpendicular to it. Then, using "alpha-tuning" removes noise due to the illuminant source. POS does the same things, but in the reverse order.

Both CHROM and PBV suffer still from a sensitivity to the illumination specifics. Illumination can affect both the vector in RGB color space along which the pulsatile variations are most pronounced and average skin tone used to help normalize a signal. To counteract this, Wang et al.[46] introduced the spatial subspace rotation (2SR) technique. The core of this approach is to consider the color distribution of a subject's skin pixels as defining a subspace of the RGB color space. By observing how the eigenvectors of a covariance-type matrix change in time, one can extract a pulse signal with better precision than either ICA, CHROM, or PBV approaches.[46] This approach is conceptually similar to observing the changes in the average hue (from hue, saturation, value [HSV] space) of the skin pixels.

A more detailed approach to color filtering can be found in Feng et al.[34] Using a detailed model of light's interaction with the skin, they calculate the difference between the intensity of the G and R signals, after normalizing to account for the product of the illumination and reflected spectra. This has the effect of eliminating two noise sources, thus making the resultant rPPG measurement more robust. They refer to their method as "Adaptive Green Red Difference". It would be interesting to know how similar their approach is to the GR method that Wang et al.[8] attribute to Huelsbusch,[45] which also takes a weighted difference of the green and red channels. But, since Huelsbusch's work only seems to appear in a difficult-to-obtain, untranslated German thesis, for this work, it will remain a point of curiosity.

3.2.2 Non-RGB Color Filtering

One facet that all of these color-filtering techniques share is use of the RGB color space. Most of the work involves removing intensity variations caused by signals with different chromatic signatures than the pulsatile signal. By working in the CIELab color space, Yang et al.[47] were able to show that motion induced artifacts appeared mainly in the L∗ (Luminance) channel and not in the a∗ or b∗ channels, which contained the chrominance information. They were able to demonstrate markedly better SNR in the a∗ channel than in RGB's G channel. It seems possible that had they tried to find a linear combination of the a∗ and b∗ channels using an approach similar to alpha tuning, they could have achieved even better results.

3.2.3 Chromatic Frequency Filtering

Researchers have also combined color filtering with temporal frequency filtering. A fairly straightforward approach was employed by Bal,[48] who used the CHROM

technique to map to chromatic coordinates, then applied Dual-Tree Complex Wavelet Transforms to filter out temporal frequencies inconsistent with heart rates.

A more imaginative use of frequency filtering was employed by Wang et al.,[49] who observed that the relative magnitude of the pulsatile signal is known for each RGB color channel. This led them to develop the preprocessing technique of "amplitude-selective filtering", which takes a Fourier transform of the temporally normalized signal (i.e., divide by mean value and subtract 1 to remove DC component), and imposes an amplitude based bandpass. Frequencies with amplitudes larger than the expected pulsatile amplitudes are dampened since they are most likely displaying motion-induced noise. This signal then undergoes denormalization and inverse Fourier transforming to complete the preprocessing.

3.3 Signal Amplification

All of the filtering techniques described previously are needed because the pulsatile signal is a subtle temporal feature of living skin. Commonly, we magnify subtle spatial features using physical techniques (i.e., lenses). In the course of several papers from the Massachusetts Institute of Technology's Computer Science and Artificial Intelligence Laboratory (CSAIL), techniques to magnify temporal features using computational techniques were developed.[50–55] Computational approaches for temporal magnification can be divided into two main categories based upon analogies with fluid mechanics, Lagrangian and Eulerian. Both approaches view an image as a fluid of pixels. The Lagrangian approach focuses on tracking the motion of each pixel, whereas the Eulerian approach divides the image into a grid and focuses on the movement of pixels in and out of each grid element. The CSAIL techniques belong to the Eulerian family and each involves three distinct phases[50]:

1) Spatial decomposition: Each cell has a temporal function fit to its variations.
2) Motion isolation: Undesired temporal frequencies and DC components are filtered out. Linear terms of the temporal function are then multiplied by a magnification factor.
3) Representation denoising: Smoothing of output.

The earlier approaches[51,55] would use local Taylor expansions as their temporal functions. Since only the first term of the expansion was kept, this is referred to as a linear magnification. The phase-based magnifications[53,54] are more robust. They involve a transformation to frequency space either using steerable pyramids[56] or Riesz transforms,[54] with the temporally linear portion of the phase term being multiplied by a magnification term. The efficacy of these methods can be seen in a

video posted by Rubenstein on YouTube at https://www.youtube.com/watch?v=EDt7ylmR7FQ. The clip demonstrates this family of algorithms at the stage where they won an Honorable Mention for the 2012 NSF International Science and Engineering Visualization Challenge.

Finally, to avoid the hand-tuning of features involved with both the linear and phase-based methods, Oh et al.[50] introduce an end-to-end deep learning approach to learn how to magnify small temporal variations in images.

3.4 Machine-Learning Approaches

Over about the past two decades there has been an explosive increase in the use and efficacy of machine-learning techniques for a wide range of problems. Machine-learning methods use iterative optimizations to learn how to manipulate input data to accomplish a desired task. This input data may consist of either hand-selected features or literal raw data streams.

For example, to measure heart rate, Hsu et al.[57] used combinations of the Fourier transform of the raw signal, the Fourier transform of the ICA-processed feature used by Poh et al.[37] (i.e., presumably, the ICA component with the highest power spectrum peak), and a CHROM-style feature used by de Haan and Jeanne[33] as raw input features for support vector regression (SVR). The choice of features was driven by human knowledge and intuition. The manner in which they were combined was determined by an algorithmic learning method (i.e., SVR). Unsurprisingly, the best single feature by far was the CHROM-based feature. Somewhat surprisingly, although adding the raw Fourier transform gave a slight improvement, adding the ICA component resulted in slightly decreased performance.

Another application is Jindal et al.'s[20] use of the PPG signal as a two-stage biometric identifier. Again, after much preprocessing, 11 features were extracted from the PPG signal. These features are first grouped into three clusters, then, for each cluster, a deep belief net trained to identify individuals within it.

Each of these machine-learning techniques required hand selection of features, reflecting the human designer's prior knowledge and intuition. When taken to their extreme, machine-learning techniques employ end-to-end learning. This entails working with just a raw signal and depending upon the learning algorithm to extract and use relevant features. Chen and McDuff[58] rely on end-to-end learning for their DeepPhys—a deep convolutional neural net designed to measure heart and respiration rate using rPPG signals. This approach employed rigorous end-to-end learning, even requiring the net to learn the ROI. Their "trick" was to use two Visual

Geometry Group (VGG)-style nets. The first would take two frames, separated by a single time step as input, and output the difference in signal attributable to the pulse at those two time steps. The second net had the same architecture as the first, only without the fully connected output layers and the final layer of feature maps. The two nets were joined at two different layers, and feature maps of the second net were used as masks for the first net. In this way, the second net was trained to detect the ROI, and the first net was trained to extract the changes in intensity due to the pulse.

The only somewhat debatable aspect of their work is how labels for the training data were obtained. The authors state that they used "gold standard" physiological readings to label their data sets. Specifically, this appears to have been the PPG signals recorded by a finger pulse PPG probe. It is not immediately clear (to this author) how the readings from a finger probe correspond to the desired readings from a remote image of the face. There are two issues here:

1) *Phase mismatch.* There is a time lag between pulse arrival at a fingertip and at a face. The phase of the "gold standard" signal and the face's rPPG signal will not be the same.

2) *Amplitude mismatch.* The finger probe has its own light source and sensor, which are each in contact with the fingertip. The rPPG measurement comes from a camera, not in contact with the face, and a different illuminating source. The amplitude of the "gold standard" signal and the rPPG signal will not be the same.

Nevertheless, the authors do report fairly good results, especially in comparison to other published work. So, either the above issues were handled implicitly or this is another example of the "unreasonable effectiveness" of neural nets, where the net somehow learned to estimate the PPG signal at a fingertip by using the rPPG signal from a subject's face.

4. Standard Data Sets

The lack of standard, publicly available data sets hampers the ability of researchers to independently verify and compare each other's work, or even to enter this area. Some publicly available data sets are not suitable because the compression algorithms used to record them corrupts the rPPG data. In a footnote, Wang et al.[8] state that they use an uncompressed bitmap format for storing their data to avoid such degradation. They specifically claim that compression makes the MAHNOB-HCI[59] unsuitable for rPPG research. This may be why the MAHNOB-HCI[59] and

COHFACE[60] had poor results for various rPPG algorithms when tested by Heusch et al.[60]

Yet, the volume of data that can easily be generated with videos creates significant incentive to use compression. The main focus of most compression algorithms is preservation of human-noticeable features. Since the rPPG signal is very clearly not something people notice, there is generally no attempt to preserve it, thus rendering it extremely susceptible to degradation during compression and decompression. McDuff et al.[61] performed experiments designed to quantify the effects of compressive loss on rPPG signals for the popular x264 and x265 codecs. They found an essentially linear degradation as the constant rate factor (CRF) increased from 0 to about 30, after which the degradation seemed to stabilize, perhaps indicating that virtually all of the signal had been lost. Yet, one can also compress the data by either downsampling or taking fewer frames per second, which results in less loss of signal.[61] Additionally, spatially or temporally downsampled data may be still have an rPPG signal extracted from it with greater accuracy using deep-learning-based super-resolution techniques.[62]

The Dataset for Emotion Analysis using Physiologic Signals (DEAP)[63] was used by Unakov[7] to verify rPPG methods for pulse estimation. The best results obtained were using the POS method to extract the signal and continuous wavelet transform (CWT) to measure the pulse rate. The mean average error in beats per minute (BPM) was 1.99, the root-mean-squared error (RMSE) was 3.25, and the measurement was within 3.5 BPM 87% of the time.

The MIMIC-III data set[64] was used by Liang et al.[17] to assess whether a patient's blood pressure was indicative of hypertension or prehypertension. However, one of the features they used, in addition to rPPG signal was pulse arrival time (PAT) (i.e., the time between the beat of the heart to arrival of pulse at the measurement point). The PAT was measured with an ECG, so their experiment was not a true instance of a remote measurement. It would be interesting to see what results DeepPhys[58] would obtain with that data set.

The ability of researchers to obtain reasonable results with these data sets suggests that they at least contain usable rPPG signals.

As a complement to standard data sets, one group of researchers has created an open source toolbox, "iPhys", of rPPG approaches, which can be downloaded from https://github.com/danmcduff/iphys-toolbox.[65] However, since the code is in MATLAB, it is only usable by possessors of the appropriate MATLAB licenses.

5. Applications and Accuracies

5.1 Detection of Noncompressible Hemorrhaging

Detecting noncompressible hemorrhaging, low blood volume, and blood loss is likely the application of most significance for the Army due to its potential of directly saving lives on the battlefield. To the best of my knowledge, all currently performed investigations for these detections have involved contact PPG measurements, not rPPG. Nevertheless, once it is established that PPG measurements can help to quickly identify Soldiers who are losing significant amounts of blood internally, then the advantages of rPPG are clear: the ability to perform remote diagnoses, not needing to attach a PPG device to each Soldier to get readings, and not requiring each Soldier to carry their own PPG reader.

The main hurdle in developing methods to detect blood loss is obtaining data. For obvious reasons, it is difficult to enroll subjects undergoing sudden, severe blood loss for a study. As a result, one must either rely upon some standard data sets that have been gathered for hospital patients or upon healthy volunteers. Fortunately, these healthy volunteers are not required to subject themselves to massive blood loss; instead, they undergo lower body negative pressure (LBNP). This involves subjecting their lower body (i.e., legs) to negative pressure, which draws more blood away from the upper body but which may also be quickly reversed without (I assume) injury or significant risk of injury to the participants. Some of the earliest work on this was performed by Convertino et al.,[2] where they established that PPG readings could be used to measure LBNP with a correlation coefficient of 0.95. More significantly, Hatib et al.[4] used existing databases (Edwards Lifesciences, MIMIC II, and UCI) to get arterial pressure waveform data on patients undergoing surgery. Using hand-selected features and logistic regression, they were able to predict hypotension (mean arterial pressure: 65 mmHg) 15 min in advance with an area under curve (AUC) of 0.95 and 5 min in advance with an AUC of 0.97). Using these curves they chose operating points with sensitivity/specificity of 88%/87% and 92%/92%.

Additionally, Reljin et al.[66] used support vector machines to analyze features obtained from PPG recordings both of trauma patients and healthy volunteers subjected to LBNP. They were able to identify those who had lost 0.9 L of blood versus no loss with 88% accuracy and 0.45 L versus no loss with 71% accuracy.

Although PPG can be used for early identification of progressing blood loss, there does not seem to have been much research on using rPPG techniques for this application. The rest of this note focuses on rPPG applications.

5.2 Pulse Rate

Comparison of various approaches across different publications is difficult. In addition to differing experimental conditions, extraction of physiologic information from a PPG signal involves several decisions:

1) ROI selection, detection, and tracking
2) Signal preprocessing
3) rPPG signal extraction
4) Signal postprocessing
5) Extraction of physiologic information

Most rPPG papers present a unique contribution for only one of these five facets. However, they make nontrivial individual decisions about the other four. This makes comparisons of different published results something of an apples-to-oranges affair. Unakafov[7] performed a comparison of various rPPG signal extraction and pulse rate extraction approaches using the same ROI and signal pre- and postprocessing methods with data from DEAP[63]. Unakafov's results are shown in Table 1.[7] He examined several PPG methods:

1) G: Green channel only
2) GRD: Difference between green and red channels
3) GRD: Adaptive difference between green and red channels[34]
4) CHROM: Chrominance[33]
5) POS: Plane orthogonal to skin[8]

Table 1 Different accuracy metrics using BPM (mean average error [MAE], RMSE and probability of error ≤ 3.5) for various combinations of PPG extraction methods and of extracting pulse rate from a PPG signal.[50]

iPPG extraction	MAE, BPM				RMSE, BPM				$PE_{3.5}$, %			
	IBI	DFT	AR	CWT	IBI	DFT	AR	CWT	IBI	DFT	AR	CWT
G	5.91	6.54	5.62	**5.35**	7.42	9.62	**6.78**	7.62	44	54	44	**58**
GRD	4.05	3.82	3.99	**3.07**	5.30	6.36	5.11	**4.96**	60	73	60	**78**
aGRD	4.41	4.41	4.28	**3.59**	5.70	7.07	**5.45**	5.55	57	70	58	**74**
ICA	3.91	3.61	3.71	**2.94**	5.41	6.03	4.87	**4.77**	64	75	63	**79**
CHROM	3.46	2.70	3.05	**2.08**	4.64	4.70	4.04	**3.46**	65	81	69	**86**
POS	3.13	2.61	2.91	**1.99**	4.30	4.50	3.80	**3.25**	70	81	71	**87**

The techniques used to then infer a pulse rate from the PPG signal were as follows:

1) IBI: interbeat interval
2) DFT: discrete Fourier transform
3) AR: autoregressive modeling
4) CWT: continuous wavelet transform

The overall best combination was POS and CWT, which achieved pulse rate measurements within 3.5 BPM 87% of the time and had an average error of either 1.99 or 3.25, depending upon whether one used a mean average error metric or RMSE metric. The error was determined by comparing the rPPG measurement to the measurement obtained from a contact (i.e., standard) PPG signal. Although Unakafov would have preferred to use an ECG measurement, the DEAP data set only provided the contact PPG measurement as a ground truth reference.

5.3 Blood Pressure

Recent work published by Mousavi et al.[18] uses some traditional machine-learning techniques (i.e., not deep learners) to measure blood pressure using rPPG. Their measurements were of diastolic blood pressure (DBP), systolic blood pressure (SBP), and mean arterial pressure (MAP). Ground-truth blood pressures were recorded from the MIMIC II database, which used ECG measurements to establish the "true" pressures. Mousavi et al. achieved accuracies well within those established by the Association for the Advancement of Medical Instrumentation (AAMI) for DBP and MAP, and just on the border of the standards for SBP, though with an average error of close to zero. The data they used came from the MIMIC II[67] database (version 3, 2015), which gave then access to 1323 samples from 441 subjects (Table 2).

Table 2 Measuring blood pressure. Table shows mean error, standard deviation, and number of subjects for DBP, MAP, and SBP, along with the accuracy guidelines promulgated by the AAMI. (Table adapted from Mousavi et al.[18])

		Mean (mmHg)	SD (mmHg)	Subject
Our results	DBP	0.187	4.173	441
	MAP	0.067	4.911	441
	SBP	–0.050	8.901	441
AAMI	BP	≤5	≤8	≥85

A more detailed breakdown of Mousavi et al.'s results can be seen in their table of cumulative errors in Table 3.

Table 3 Cumulative errors achieved by Mousavi et al.[18] in measuring blood pressure. Table shows percentages of measurements that fall at or below specific error limits.

	Cumulative Error Percentage			
	≤5 mmHg	≤10 mmHg	≤15 mmHg	Total
DBP	1118 (84%)	1227 (92%)	1295 (97%)	1323
MAP	1055 (79%)	1166 (83%)	1236 (93%)	1323
SBP	953 (71%)	1027 (77%)	1113 (84%)	1323

5.4 Respiration Rates

Although respiration rates are more difficult to accurately measure, van Gastel et al.[12] performed two independent series of experiments—one using guided breathing, wherein the subjects attempted to breathe at a predefined rate, and the other with neonates, who breathed at whatever rate they desired. Comparisons were made under differing conditions (i.e., different breath rates for guided breathers and different camera angles for neonates, using various rPPG signal extraction techniques). Both the CHROM and POS methods were improvements over the current state-of-the-art algorithms, indicated by BM_F and BM_I, where BM stands for "benchmark". Ground truth for the guided breathing experiments was the targeted breath rates; for the neonates, ground truth was determined with ECG-determined respiratory rates. The ECG reference was not used for the targeted breathing groups due to its sensitivity to motion-induced noise.

However, since the average adult has a respiratory rate of 12–20 breaths per minute and newborns tend to breathe 30–60 times per minute, the given RMSEs, as seen in Tables 4 and 5, correspond to error rates of, at best, approximately 10%–20%. It appears that rPPG techniques are not yet capable of truly accurate respiration rate measurements.

Table 4 Pearson correlation coefficient, Bland–Altman plot slopes, MAE, RMSE, and standard deviation of error for respiration rate measurements obtained for guided breathing subjects by van Gastel et al.[12]

Breathing scenario	Method	r	B	MAE	RMSE	σ
Overall	CHROM	**0.962**	0.971	**1.74**	**2.67**	**2.61**
	PBV	0.957	**0.973**	1.99	2.92	2.85
	BM_I	0.875	0.904	3.65	5.21	4.85
	BM_F	0.862	0.897	4.47	5.90	5.59

Table 5 Pearson correlation coefficient, Bland–Altman plot slopes, MAE, RMSE, and standard deviation of error for respiration rate measurements obtained for neonates by van Gastel et al.[12]

Camera distance	Method	r	B	MAE	RMSE	σ
Close-view	CHROM	**0.886**	**0.898**	**4.67**	**9.24**	**9.13**
	PBV	0.871	0.868	5.31	9.53	9.49
	BM_I	0.819	0.831	7.36	11.5	11.2
	BM_F	0.783	0.808	9.09	14.7	14.3
Wide-view	CHROM	**0.869**	**0.890**	**4.76**	**9.49**	**9.29**
	PBV	0.855	0.861	5.46	9.72	9.77
	BM_I	0.814	0.827	7.43	11.7	11.3
	BM_F	0.776	0.813	9.17	14.8	14.4
Overall	CHROM	**0.872**	**0.892**	**4.72**	**9.35**	**9.22**
	PBV	0.862	0.864	5.39	9.66	9.67
	BM_I	0.817	0.829	7.40	11.6	11.3
	BM_F	0.780	0.811	9.13	14.8	14.4

6. Conclusion

This technical note presents an overview of state-of-the-art techniques, results, uses, and problems for a very specific approach to rPPG, which uses ordinary RGB cameras to monitor spectral changes from human skin. The selection of rPPG was based on these three areas:

1) Easily available hardware
2) Ability to use video that had not been originally intended as a source of rPPG data
3) Robustness to noise

Evidence of this can be seen in the recent commercialization of this technology, though primarily in nonmedical applications. Examples of such commercialization include the Philips VitalSigns Camera and Microsoft's CardioLens,[68] which is a wearable headset that displays rPPG data on people interacting with the user. This commercialization is occurring in conjunction with active research, as can be seen in the distribution of dates for the references used herein.

The potential military usefulness of this technology ranges from diagnosing noncompressible hemorrhaging prior to the onset of shock, to easily and noninvasively monitoring peoples' physical condition, verifying their identity, or even gaining insight into their cognitive and emotional state.

7. References

1. Kelly JF, Ritenour AE, McLaughlin DF, Bagg KA, Apodaca AN, Mallak CT, Pearse L, Lawnick MM, Champion HR, Wade CE, et al. Injury severity and causes of death from operation Iraqi Freedom and Operation Enduring Freedom: 2003–2004 versus 2006. Journal of Trauma and Acute Care Surgery. 2008;64(2):S21–S27.

2. Convertino VA, Moulton SL, Grudic GZ, Rickards CA, Hinojosa-Laborde C, Gerhardt RT, Blackbourne LH, Ryan KL. Use of advanced machine-learning techniques for noninvasive monitoring of hemorrhage. Journal of Trauma and Acute Care Surgery. 2011;71(1):S25–S32.

3. Swiston AJ, Raj M. FY12 line-supported bio-medical initiative program: advanced photoplethysmography (PPG) sensors for operational and casualty care medicine. Lexington (MA): Lincoln Labs, Massachusetts Institute of Technology; 2013 Feb 11. Report No.: LSP-40; ESC-EN-HA-TR-2012-116.

4. Hatib F, Jian Z, Buddi S, Lee C, Settels J, Sibert K, Rinehart J, Cannesson M. Machine learning algorithm to predict hypotension based on high-fidelity arterial pressure waveform analysis. Anesthesiology: the Journal of the American Society of Anesthesiologists. 2018;129(4):663–674.

5. van der Ster BJP, Westerhof BE, Stok WJ, van Lieshout JJ. Detecting central hypovolemia in simulated hypovolemic shock by automated feature extraction with principal component analysis. Physiological Reports. 2018;6(22):e13895.

6. Al-Naji A, Perera AG, Chahl J. Remote monitoring of cardiorespiratory signals from a hovering unmanned aerial vehicle. Biomedical Engineering Online. 2017 Dec;16(1):101.

7. Unakafov AM. Pulse rate estimation using imaging photoplethysmography: generic framework and comparison of methods on a publicly available dataset. Biomedical Physics and Engineering Express. 2018;4(4):045001.

8. Wang W, den Brinker AC, Stuijk S, de Haan G. Algorithmic principles of remote PPG. IEEE Transactions on Biomedical Engineering. 2017;64(7):1479–1491.

9. Gil E, Orini M, Bailon R, Vergara JM, Mainardi L, Laguna P. Photoplethysmography pulse rate variability as a surrogate measurement of heart rate variability during non-stationary conditions. Physiological Measurement. 2010;31(9):1271.

10. Jeyhani V, Mahdiani S, Peltokangas M, Vehkaoja A. Comparison of HRV parameters derived from photoplethysmography and electrocardiography signals. In: Proceedings of the 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); 2015 Aug 25–29; Milan, Italy. p. 5952–5955.

11. Wei J, Luo H, Wu SJ, Zheng PP, Fu G, Lee K. Transdermal optical imaging reveal basal stress via heart rate variability analysis: a novel methodology comparable to electrocardiography. London (England): Digital Science and Research Solutions, Inc.; c2019 [accessed 2018 Nov 29]. https://app.dimensions.ai/details/publication/pub.1100894218.

12. van Gastel M, Stuijk S, de Haan G. Robust respiration detection from remote photoplethysmography. Biomedical Optics Express. 2016;7(12):4941–4957.

13. Verkruysse W, Svaasand LO, Stuart J. Nelson. Remote plethysmographic imaging using ambient light. Optics Express. 2008;16(26):21434–21445.

14. Lincy AP, Rajasekaran K, Jothi ESJ. Continuous monitoring of blood glucose using photophlythesmograph signal. In: Proceedings of the 2017 International Conference on Innovations in Electrical, Electronics, Instrumentation and Media Technology (ICEEIMT); 2017 Feb 3–4; Coimbatore, India. p. 187–191.

15. Guazzi AR, Villarroel M, Jorge J, Daly J, Frise MC, Robbins PA, Tarassenko L. Non-contact measurement of oxygen saturation with an RGB camera. Biomedical Optics Express. 2015;6(9):3320–3338.

16. Jeong IC, Finkelstein J. Introducing contactless blood pressure assessment using a high speed video camera. Journal of Medical Systems. 2016;40(4):77.

17. Liang Y, Chen Z, Ward R, Elgendi M. Hypertension assessment via ECG and PPG signals: an evaluation using mimic database. Diagnostics. 2018;8(3):65.

18. Mousavi SS, Firouzmand M, Charmi M, Hemmati M, Moghadam M, Ghorbani Y. Blood pressure estimation from appropriate and inappropriate PPG signals using a whole-based method. Biomedical Signal Processing and Control. 2019;47:196–206.

19. Weiler DT, Villajuan SO, Edkins L, Cleary S, Saleem JJ. Wearable heart rate monitor technology accuracy in research: a comparative study between PPG and ECG technology. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting; 2017 Oct 9–13; Austin, TX. Thousand Oaks (CA): SAGE Publications Sage. Vol. 61, p. 1292–1296.

20. Jindal V, Birjandtalab J, Pouyan MB, Nourani M. An adaptive deep learning approach for PPG-based identification. Proceedings of the 2016 IEEE 38th Annual International Conference of the Engineering in Medicine and Biology Society (EMBC); 2016 Aug 16–20; Orlando, FL. p. 6401–6404.

21. Seepers RM, Wang Wenjin, de Haan G, Sourdis I, Strydis C. Attacks on heartbeat-based security using remote photoplethysmography. IEEE Journal of Biomedical and Health Informatics. 2017;22(3):714–721.

22. Liu J, Luo H, Zheng PP, Wu SJ, Lee K. Transdermal optical imaging revealed different spatiotemporal patterns of facial cardiovascular activities. Scientific Reports. 2018;8(1):10588.

23. McDuff D, Gontarek S, Picard R. Remote measurement of cognitive stress via heart rate variability. In: Proceedings of the 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); 2014 Aug 26–30; Chicago, IL. p. 2957–2960.

24. Balakrishnan G, Durand F, Guttag J. Detecting pulse from head motions in video. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2013 June 23–28; Portland, OR. p. 3430–3437.

25. Viola P, Jones M. Rapid object detection using a boosted cascade of simple features. In: Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2001); 2001 Dec 8–14; Kauai, Hawaii. vol. 1, p. I–I.

26. Liu Z, Luo P, Wang X, Tang X. Deep learning face attributes in the wild. In: Proceedings of the IEEE international conference on computer vision; 2015 Dec 7–13; Santiago, Chile. p. 3730–3738.

27. Zhang Z, Luo P, Loy CC, Tang X. Facial landmark detection by deep multi-task learning. In: Proceedings of the 13th European Conference on Computer Vision (ECCV 2014); 2014 Sep 6–12; Zurich, Switzerland. Basel (Switzerland): Springer; c2014. p. 94–108.

28. 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 Biomedical and Health Informatics. 2017;21(5):1422–1433.

29. van Luijtelaar R, Wang W, Stuijk S, de Haan G. Automatic ROI detection for camera-based pulse-rate measurement. In: Proceedings of the Asian Conference on Computer Vision (ACCV 2014); 2014 Nov 1–2; Singapore, Singapore. Basel (Switzerland): Springer; 2014. vol 9009. p. 360–374.

30. Schmitz GLLH. Video camera based photoplethysmography using ambient light. [MS thesis dissertation]. [Eindhoven (The Netherlands)]: Technische Universiteit; 2011.

31. Harris C, Stephens M. A combined corner and edge detector. Alvey Vision Conference. 1988:15(50).

32. Bobbia S, Macwan R, Benezeth Y, Mansouri A, Dubois J. Unsupervised skin tissue segmentation for remote photoplethysmography. Pattern Recognition Letters. 2017 Oct 18.

33. de Haan G, Jeanne V. Robust pulse rate from chrominance-based rPPG. IEEE Transactions on Biomedical Engineering. 2013;60(10):2878–2886.

34. Feng L, Po L-M, Xu X, Li Y, Ma R. Motion-resistant remote imaging photoplethysmography based on the optical properties of skin. IEEE Transactions on Circuits and Systems for Video Technology. 2015;25(5):879–891.

35. Crowe JA, Damianou D. The wavelength dependence of the photoplethysmogram and its implication to pulse oximetry. In: Proceedings of the 1992 14th Annual International Conference of the IEEE Engineering in Medicine and Biology Society; 1992 Oct 29–Nov 1; Paris, France. Vol. 6, p. 2423–2424.

36. Macwan R, Benezeth Y, Mansouri A. Remote photoplethysmography with constrained ICA using periodicity and chrominance constraints. Biomedical Engineering Online. 2018;17(1):22.

37. Poh M-Z, McDuff DJ, Picard RW. Advancements in noncontact, multiparameter physiological measurements using a webcam. IEEE Transactions on Biomedical Engineering. 2011;58(1):7–11.

38. Tsouri GR, Kyal S, Dianat SA, Mestha K. Constrained independent component analysis approach to nonobtrusive pulse rate measurements. Journal of Biomedical Optics. 2012;17(7):077011.

39. Zhao F, Li M, Qian Yi, Tsien JZ. Remote measurements of heart and respiration rates for telemedicine. PloS One. 2013;8(10):e71384.

40. Lu W, Rajapakse JC. Constrained independent component analysis. In: Advances in Neural Information Processing Systems; 2000; Denver, CO. p. 570–576.

41. Lewandowska M, Ruminski J, Kocejko T, Nowak J. Measuring pulse rate with a webcam—a non-contact method for evaluating cardiac activity. In: Proceedings of the 2011 Federated Conference on Computer Science and Information Systems (FedCSIS); 2011 Sep 18–21; Szczecin, Poland. p. 405–410.

42. Holton BD, Mannapperuma K, Lesniewski PJ, Thomas JC. Signal recovery in imaging photoplethysmography. Physiological Measurement. 2013;34(11):1499.

43. Iozzia L, Cerina L, Mainardi L. Relationships between heart-rate variability and pulse-rate variability obtained from video-PPG signal using ZCA. Physiological Measurement. 2016;37(11):1934.

44. de Haan G, van Leest A. Improved motion robustness of remote-PPG by using the blood volume pulse signature. Physiological Measurement. 2014;35(9):1913.

45. Huelsbusch M. Ein bildgestuetstes, funktionelles verfahren zur optoelektronischer erfassung der hautperfusion [PhD thesis]. [München (Germany)]: Fakultaet fuer Elektrotechnik un Informationstechnik, Technische Universität München; 2008.

46. Wang W, Stuijk S, de Haan G. A novel algorithm for remote photoplethysmography: spatial subspace rotation. IEEE Transactions on Biomedical Engineering. 2016;63(9):1974.

47. Yang Y, Liu C, Yu H, Shao D, Tsow F, Tao N. Motion robust remote photoplethysmography in CIELab color space. Journal of Biomedical Optics. 2016;21(11):117001.

48. Bal U. Non-contact estimation of heart rate and oxygen saturation using ambient light. Biomedical Optics Express. 2015;6(1):86–97.

49. Wang W, den Brinker AC, Stuijk S, de Haan G. Amplitude-selective filtering for remote-PPG. Biomedical Optics Express. 2017;8(3):1965–1980.

50. Oh T-H, Jaroensri R, Kim C, Elgharib M, Durand F, Freeman WT, Matusik W. Learning-based video motion magnification. Ithaca (NY): Cornell University; 2018. http://arXiv preprint arXiv:1804.02684.

51. Rubinstein M et al. Analysis and visualization of temporal variations in video [PhD thesis]. [Cambridge (MA)]: Massachusetts Institute of Technology; 2014.

52. Wadhwa N Revealing and analyzing imperceptible deviations in images and videos. [PhD thesis]. [Cambridge (MA)]: Massachusetts Institute of Technology; 2016.

53. Wadhwa N, Rubinstein M, Durand F, Freeman WT. Phase-based video motion processing. ACM Transactions on Graphics (TOG). 2013;32(4):80.

54. Wadhwa N, Rubinstein M, Durand F, Freeman WT. Riesz pyramids for fast phase-based video magnification. In: Proceedings of the 2014 IEEE International Conference on Computational Photography (ICCP 2014); 2014 May 2–4; Santa Clara, CA. p. 1–10.

55. Wu HY, Rubinstein M, Shih E, Guttag J, Durand F, Freeman W. Eulerian video magnification for revealing subtle changes in the world. ACM Transactions in Graphics (TOG). 2012;31(4). Article 65.

56. Simoncelli EP, Freeman WT. The steerable pyramid: a flexible architecture for multi-scale derivative computation. In: Proceedings of the International Conference on Image Processing; 1995 Oct 23–26; Washington, DC. Vol. 3, p. 444–447.

57. Hsu Y, Lin Y-L, Hsu W. Learning-based heart rate detection from remote photoplethysmography features. In: Proceedings of the 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP); 2014 May 4–9; Florence, Italy. p. 4433–4437.

58. Chen W, McDuff D. Deepphys: video-based physiological measurement using convolutional attention networks. Ithaca (NY): Cornell University; 2018. http://arXiv preprint arXiv:1805.07888.

59. Soleymani M, Lichtenauer J, Pun T, Pantic M. A multimodal database for affect recognition and implicit tagging. IEEE Transactions on Affective Computing. 2012;3(1):42–55.

60. Heusch G, Anjos A, Marcel S. A reproducible study on remote heart rate measurement. Ithaca (NY): Cornell University; 2017. http://arXiv preprint arXiv:1709.00962.

61. McDuff DJ, Blackford EB, Estepp JR. The impact of video compression on remote cardiac pulse measurement using imaging photoplethysmography. In: Proceedings of the 2017 12th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2017); 2017 May 30–June 3; Washington, DC. p. 63–70.

62. McDuff D. Deep super resolution for recovering physiological information from videos. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW); 2018 June 18–22; Salt Lake City, UT. p. 1367–1374.

63. Koelstra S, Muhl C, Soleymani M, Lee J-S, Yazdani A, Ebrahimi T, Pun T, Nijholt A, Patras I. DEAP: a database for emotion analysis; using physiological signals. IEEE Transactions on Affective Computing, 2012;3(1):18–31.

64. Johnson AEW, Pollard TJ, Shen L, Liwei HL, Feng M, Ghassemi M, Moody B, Szolovits P, Celi LA, Mark RG. MIMIC-III, a freely accessible critical care database. Scientific Data. 3:160035, 2016.

65. McDuff D, Blackford E. Iphys: an open non-contact imaging-based physiological measurement toolbox. Ithaca (NY): Cornell University; 2019. http://arXiv preprint arXiv:1901.04366.

66. Reljin N, Zimmer G, Malyuta Y, Shelley K, Mendelson Y, Blehar DJ, Darling CE, Chon KH. Using support vector machines on photoplethysmographic signals to discriminate between hypovolemia and euvolemia. PloS One. 2018;13(3):e0195087.

67. Saeed M, Villarroel M, Reisner AT, Clifford G, Lehman L-W, Moody G, Heldt T, Kyaw TH, Moody B, Mark RG. Multiparameter intelligent monitoring in intensive care II (MIMIC-II): a public-access intensive care unit database. Critical Care Medicine; 2011;39(5):952.

68. Hurter C, McDuff D. Cardiolens: remote physiological monitoring in a mixed reality environment. In: Proceedings of the ACM SIGGRAPH 2017 Emerging Technologies; 2017 July 30–Aug 3; Los Angeles, CA. p. 6.

List of Symbols, Abbreviations, and Acronyms

2SR	spatial subspace rotation
AAMI	Association for the Advancement of Medical Instrumentation
AR	autoregressive modeling
BPM	beats per minute
BSS	blind source separation
CICA	constrained ICA
CRF	constant rate factor
CSAIL	Computer Science and Artificial Intelligence Laboratory
CT	computed tomography
CWT	continuous wavelet transform
DBP	diastolic blood pressure
DC	direct current
DEAP	Dataset for Emotion Analysis using Physiologic Signals
DFT	discrete Fourier transform
ECG	electrocardiogram
IBI	interbeat interval
ICA	independent component analysis
ID	identification
iPPG	imaging PPG
JBSS	joint blind source separation
LBNP	lower body negative pressure
MAP	mean arterial pressure
MAE	mean average error
MRI	magnetic resonance imaging
NSF	National Science Foundation

PAT	pulse arrival time
PBV	pulse blood volume
PCA	principal component analysis
POS	plane orthogonal to skin
PPG	photoplethysmography
RMSE	root-mean-squared error
ROI	region of interest
rPPG	remote photoplethysmography
SBP	systolic blood pressure
SNR	signal-to-noise ratios
SVR	support vector regression
VGG	Visual Geometry Group
vPPG	video PPG
ZCA	zero-phase component analysis

1 (PDF)	DEFENSE TECHNICAL INFORMATION CTR DTIC OCA
1 (PDF)	CCDC ARL FCDD RLD CL TECH LIB
1 (PDF)	GOVT PRINTG OFC A MALHOTRA
1 (PDF)	CCDC ARL FCDD RLH FC S GUTSTEIN