PAPER

Enhanced vascular perfusion mapping and heart rate estimation via spatio-temporal rPPG with optical and motion compensation techniques

To cite this article: Faisal Farhan and Yannick Benezeth 2026 *Biomed. Phys. Eng. Express* **12** 025013

Biomedical Physics & Engineering Express

PAPER

RECEIVED
16 July 2025

REVISED
30 October 2025

ACCEPTED FOR PUBLICATION
12 November 2025

PUBLISHED
18 February 2026

Enhanced vascular perfusion mapping and heart rate estimation via spatio-temporal rPPG with optical and motion compensation techniques

Faisal Farhan* **and Yannick Benezeth**

Université Bourgogne Europe, Laboratory ImViA, 9 Avenue Alain Savary, 2100 Dijon, France
* Author to whom any correspondence should be addressed.

E-mail: faisal_farhan@etu.u-bourgogne.fr

Keywords: heart rate estimation, remote photoplethysmography, perfusion mapping

Abstract

Remote photoplethysmography (rPPG) offers a non-contact method for monitoring physiological signals using camera-based systems. The goal of this research is to estimate heart rate and spatial distributions of vascular perfusion using spatio-temporal rPPG (ST-rPPG) and to evaluate the impact of polarization, spectral filtering, and motion compensation on perfusion map quality and heart rate estimation. Two acquisition setups were used: an RGB camera with and without cross-polarization, and a monochrome camera combined with spectral filters. A motion compensation strategy was implemented that combined optical flow-based stable segment selection and temporal video stabilization to reduce motion artifacts. Four rPPG algorithms (GREEN, CHROM, POS, and G-R) were evaluated using three performance metrics: Absolute Error (AE), Signal Quality Index (SQI), and Signal-to-Noise Ratio (SNR) under cross polarized and non polarized lighting in 20 subjects to assess their suitability for perfusion mapping. GREEN and G-R method stood out giving the best results. In the second setup, nine spectral filters were tested across three anatomical regions using the GREEN method, to investigate the influence of wavelength selection on spatial perfusion signal quality. Green, orange, and blue wavelengths produced the best results in terms of AE, SQI and SNR, particularly in the palm region. Visualizations like the spatial perfusion maps, confirmed the superiority of motion-compensated, polarized, and spectrally optimized conditions for enhancing non-contact vascular perfusion assessment. Prior rPPG studies focused primarily on facial datasets or single optical factors, while this work provides the systematic evaluation of polarization, spectral filtering, and motion compensation in a unified hand-based framework, extending established rPPG methods toward high-resolution perfusion mapping.

1. Introduction

Contact photoplethysmography (cPPG), a technique established in the 1930s, involves the use of light to detect variations in blood volume caused by rhythmic pulsations of the circulatory system. Since 2008, studies have shown that photoplethysmography (PPG) can be conducted remotely (referred to as rPPG) by using ambient light as the optical source (Verkruysse *et al* 2008a). Following this discovery, numerous research efforts have been dedicated to extracting heart rate information using cameras (Hassan *et al* 2017, Rouast *et al* 2018). PPG and rPPG techniques have been increasingly investigated as non-invasive tools for monitoring cardiovascular health and assessing risk factors associated with cardiovascular disease (Charlton *et al* 2022). The cardiovascular system, composed of the heart and blood vessels, is responsible for circulating blood throughout the body, supplying oxygen and nutrients to tissues, and eliminating waste products. In a healthy adult, the pulse rate typically falls between 60 and 100 beats per minute, while the respiratory rate (RR) ranges from 12 to 20 breath cycles per minute, each cycle including an inhale and exhale. The oxygen saturation level (SpO2), which represents the ratio of

oxygenated to deoxygenated hemoglobin, is generally 95% or higher. Furthermore, normal blood pressure during heartbeats is considered to be below 120/80 mmHg (B. *et al* 2018). Abnormalities in the cardiovascular system can lead to a variety of cardiovascular diseases, including cardiomyopathy, hypertensive heart disease, and arrhythmia, all of which can have a negative impact on overall health (Virani *et al* 2021).Heart rate and blood pressure are key physiological indicators that are widely used to monitor cardiovascular health and detect potential cardiovascular diseases. These parameters can be measured through a variety of physiological sensors and medical devices.

In general, methods for monitoring heart rate can be categorized into contact-based and noncontact-based measurements. Contact-based methods require sensors to be placed directly on the skin or other body contact surfaces such as electrocardiography (ECG), photoplethysmography (PPG), ballistocardiography (BCG), and impedance cardiography (ICG) (Chan *et al* 2019). ECG is widely regarded as the gold-standard contact technique for recording cardiac signals (Tran *et al* 2024). It records the heart's electrical activity through chest electrodes, capturing the sequential depolarization and repolarization of the cardiac chambers, which appear as characteristic waveform components (P wave, QRS complex, and T wave). Heart rate (HR) is derived from the temporal interval between consecutive R-peaks (the RR interval). However, ECG-based HR measurement requires the placement of multiple adhesive electrodes at specific anatomical locations, which can restrict patient mobility and may cause discomfort or skin irritation during long-term or ambulatory monitoring (Alizadeh-Meghrazi *et al* 2021).

Photoplethysmography (PPG) (Kim and Baek 2023) is a popular alternative that overcomes limitations and provides accurate measurements. A photoplethysmography (PPG) device utilizes a combination of a light source and a photodetector to measure blood volume changes in the microvascular bed of tissue. By detecting variations in blood volume pulse (BVP) within vessels beneath the skin, PPG accurately captures vital information about blood circulation. Based on Beer–Lambert's law (Swinehart 1962), when light at a certain wavelength penetrates the skin, it undergoes diffusion and scattering, resulting in light attenuation. The amount of light absorbed by the skin depends on the penetration depth and the hemoglobin concentration within the blood vessels. Throughout each cardiac cycle, pulsatile changes in blood volume cause subtle variations in the intensity of reflected and transmitted light. A photodetector captures these changes as a photoplethysmography (PPG) signal, which is subsequently used for heart rate (HR) measurement. This principle is widely employed in wearable technologies like pulse oximeters, smartwatches, and smartphone sensors. Despite its advantages of comfort, low cost, and ease of integration into compact devices, PPG cannot replace ECG in clinical diagnostics, as its applications remain limited primarily to vital sign monitoring. Traditional PPG sensors are limited by spot measurements, lack of continuous perfusion data, and the need for direct skin contact, making them unsuitable for sensitive or injured skin.

Non-contact techniques for heart rate measurement, including radar, thermal imaging, and remote photoplethysmography (rPPG), have gained considerable interest in recent years due to their ability to function without direct physical contact with the body (Debnath and Kim 2025). Remote photoplethysmography (rPPG) is the most convenient and widely accessible method, as it simply requires a camera, along with a light source like an LED to capture skin videos. Compared to radar and thermal imaging, which rely on specialized equipment, rPPG offers a more affordable and less equipment-dependent alternative. Verkruysse *et al* (Verkruysse *et al* 2008b) initially introduced this technique in 2008, utilizing a consumer-grade camera and ambient light to successfully estimate heart rate. Following this, various traditional methods were developed to mitigate the effects of motion artifacts and changes in illumination, aiming to enhance the accuracy of HR estimation. This method involves performing PPG without contacting the subject, increasing the distance between the sensor and the light source. Extensive research in both contact-based and contactless PPG has focused on classifying different methodologies and evaluating their suitability under various conditions. The standard remote PPG process starts with capturing a video of a person's skin region using an imaging sensor, such as a camera, smartphone, or webcam. The lighting can be provided by either a dedicated source or ambient light. The recorded video maintains a frame rate suitable for capturing subtle variations in light reflection from the skin. The process involves selecting regions of interest (ROIs) within the video frames, such as the face, forehead, chest, or palm, either manually or by automated methods. The objective is to identify an area that provides a strong and reliable signal for effective extraction. The raw rPPG signal is derived from the pixels within the selected ROI using different techniques Lee *et al* (2023). Following this, signal optimization processes are applied, such as filtering out unwanted frequencies, to enhance quality. Finally, heart rate estimation is performed through mathematical analysis. Most existing studies in the field have concentrated primarily on developing and refining algorithms for rPPG signal extraction and analysis. These algorithms are often tested using recordings obtained in controlled laboratory environments, where variables such as lighting, background, and subject positioning are carefully managed. However, this controlled setup may not fully represent the challenges encountered in real-world scenarios,

where unpredictable movements, varying lighting conditions, and background clutter can significantly affect signal quality. As a result, there is a growing need for research that addresses rPPG performance in more dynamic, naturalistic environments. Most studies in rPPG have focused on healthy individuals, but interest in applying the technique to patients with cardiovascular conditions is growing. Research has demonstrated its potential in clinical settings, including monitoring after heart surgery, detection of atrial fibrillation, and identification of arrhythmias. Remote photoplethysmography offers a versatile solution for heart rate monitoring across various environments, including home settings and remote locations. Additionally, it can be integrated with other physiological data, such as video, audio, and temperature, to create a more holistic assessment of an individual's health. The development of precise, reliable, and practical rPPG techniques has been a major area of research, leading to the proposal and evaluation of various methods. While most rPPG research focuses on global metrics like heart rate, its potential for assessing spatially-resolved microvascular perfusion is still underexplored. Mapping perfusion blood flow in tiny vessels is crucial for tissue health, wound healing, and personalized care. Recent advances in spatiotemporal rPPG (ST-rPPG) now enable affordable, non-contact estimation of high-resolution perfusion maps, broadening the clinical applications of rPPG in areas such as dermatology, vascular assessment, and wound care. Although these issues are important, research in this area is still limited, which makes it difficult to develop effective solutions. Traditionally, both contact Photoplethysmography (cPPG) and remote Photoplethysmography (rPPG) have been widely used for non-invasive heart rate monitoring. More recently, however, advances in signal processing and artificial intelligence (AI) have expanded their applications to include the estimation of other clinically valuable vital signs such as blood pressure and heart rate variability (HRV) (Slapničar *et al* 2019).

This study focuses on the estimation of spatially-resolved vascular perfusion using ST-rPPG. Unlike prior studies that primarily evaluated existing rPPG algorithms on face datasets or under fixed illumination, this work introduces a novel acquisition pipeline and dataset focusing on hand-based recordings. The dataset was recorded using both RGB and monochrome cameras, with controlled polarization and spectral filtering conditions. Furthermore, a motion compensation framework was implemented prior to signal extraction, combining automated stable-segment selection with temporal stabilization using the VidStab library in python. This preprocessing ensured that the subsequent application of four established rPPG methods (GREEN, CHROM, POS, and G-R) was conducted on motion-optimized video segments. In addition to heart rate estimation, the study extends these methods toward spatial microvascular perfusion mapping, providing new insights into blood flow distribution under different optical conditions. It further investigates the influence of polarization, spectral filtering, and motion compensation on the quality and reliability of the generated perfusion maps, with the aim of advancing non-contact, affordable techniques for microvascular assessment.

2. State of the art

This section reviews prior research in rPPG signal extraction, optimization, and applications. Initially, for remote photoplethysmography researchers first used industrial cameras, but now they are often relying on low-cost, widely available devices such as webcams and phone cameras. Most studies using camera data for rPPG rely on red, green, and blue (RGB) color information. Different techniques have been developed to combine these color channels. Near-infrared (NIR) cameras have also been utilized in some studies (van Gastel *et al* 2015). However, a key limitation of NIR is its lower absorption by hemoglobin, which leads to a reduced signal-to-noise ratio (SNR). Polarization techniques, such as using polarized illumination and filters, have been investigated to reduce artifacts, particularly surface reflections. While earlier studies showed mixed results—Hülsbusch (Hülsbusch 2008) reported no significant improvement, whereas Sidorov *et al* (Sidorov *et al* 2016) observed increased reliability with orthogonal polarization. More has demonstrated that orthogonal polarization filtration can enhance signal quality, help distinguish between blood volume and ballistocardiographic effects, and improve the understanding of iPPG signal origins. These findings support the potential value of using polarization filters in rPPG applications. Despite these advancements, most studies still rely on standard ambient lighting—whether artificial, natural, or a mix of both. In rPPG, both single-channel and multi-channel approaches are commonly used, with the green channel often preferred due to its typically higher signal quality (Verkruysse *et al* 2008a). When combining color channels, two main strategies exist. One method is knowledge-based, which uses known differences in how pulse signals and noise appear in different color channels or spaces. The other is data-driven, using blind source separation (BSS) to automatically combine color channels based on the signal itself, without needing prior knowledge. De Haan and Jeanne (de Haan and Jeanne 2013) proposed empirically derived weighted combinations of the red, green, and blue (RGB) channels to generate a chrominance signal, a method that has since been widely adopted in the field (Huang and Dung 2016). Blind source separation (BSS) has played a central role in iPPG signal processing ever since its initial application by Poh *et al* (Poh *et al* 2010). This technique is designed to extract

the meaningful physiological signal, such as the heartbeat, while filtering out unwanted noise and artifacts. Common implementations include methods such as Principal Component Analysis (PCA) and Independent Component Analysis (ICA), JADE and FastICA, along with more advanced variations like joint BSS, spatio-temporal ICA, constrained ICA, radical ICA , robust ICA, and Zero Phase Component Analysis (ZCA) have been employed in iPPG (Zaunseder *et al* 2018).

Signal processing techniques are often used either prior to, following, or on both sides of the color channel combination step. Common procedures include detrending and band-pass filtering (Botina-Monsalve *et al* 2022). Band-pass filters are typically designed to include a specific frequency range that corresponds to expected heart rates, such as 0.7 to 3.0 Hz, 0.75 to 4.5 Hz or 0.7 to 4 Hz. Several studies have introduced specialized signal processing methods to enhance iPPG signals. Wu *et al* (Wu *et al* 2017) applied the continuous wavelet transform (CWT), enhancing the signal by filtering and weighting wavelet coefficients before reconstructing it. Huang *et al* (Huang and Dung 2016) also used CWT but took a different approach—they identified the best scale for inverse transformation, effectively creating an adaptive band-pass filter. Early research primarily concentrated on accurate heart rate (HR) detection. However, since HR alone offers limited diagnostic insight, camera-based heart rate variability (HRV) analysis is becoming increasingly relevant and is attracting interest. Heart rate (HR) is commonly estimated using fixed time windows, with 10 seconds being a standard choice. However, studies have also explored longer intervals, such as 18 seconds, 20 seconds, and even up to 60 seconds (Zaunseder *et al* 2018). These windows allow for frequency-domain and time-frequency analysis, often using transforms such as FFT or its variants, fractional Fourier transform (FrFT), or spectral methods, and sometimes autoregressive models for estimating HR or HRV. Alternatively, time-domain approaches like autocorrelation can also be used to compute HR. In addition to conventional methods, some studies have explored the use of machine learning techniques to estimate heart rate directly from signal segments or their spectral representations (Hsu *et al* 2014). Recent advancements in remote photoplethysmography (rPPG) have extended its applications beyond heart rate and HRV estimation. Studies have demonstrated the feasibility of SpO_2 estimation using RGB cameras and adaptive processing techniques (Tarassenko *et al* 2014). Research has also explored vascular assessment using wavelength-dependent responses to monitor vasomotor activity (Trumpp *et al* 2016). In addition, pulse transit time (PTT) linked to blood pressure has been estimated using camera setups (Jeong and Finkelstein 2016). These developments highlight rPPG's potential for broader applications in non-contact physiological monitoring.

3. Material and methods

3.1. Experimental Setup

The experiments were conducted using a standard remote photoplethysmography (rPPG) configuration at the Imaging and Artificial Vision Laboratory (ImVIA), Université Bourgogne Europe (uB), France. A total of 24 volunteers, aged between 24 and 40 years, participated in the study. Data was collected using two distinct acquisition setups equipped with different camera systems. Prior to the recordings, informed consent was obtained from each participant. In both setups, the cameras were positioned approximately 65 cm from the participant's hand, which was placed on a table. The camera angle was adjusted to directly face the hand, ensuring a consistent viewpoint. To provide uniform illumination, a DÖRR DLP-600 LED lighting kit was used, featuring a panel measuring 31.7 × 27 × 4.5 cm and comprising 588 LEDs. The recordings were performed in a dark room without any external lighting sources to minimize interference with ambient light. Each camera was mounted on an adjustable tripod to maintain a fixed position throughout the acquisition. Participants were instructed to minimize hand movements during recording sessions. Both the camera and lighting panel were aligned at a 90-degree and 45-degree angles to the hand, respectively, with careful positioning to prevent shadows from affecting the region of interest.

To investigate the influence of specular reflections in the rPPG signals, a cross-polarization setup was also implemented. This involved attaching a linear polarizer (300 × 200 × 0.2 mm) to the LED panel, while a polarization filter was placed in front of the camera lens. These components helped reduce reflected glare and enhance the signal quality from the skin surface. Camera parameters for both setups were systematically optimized to achieve the best signal acquisition quality. All video recordings were saved in Audio Video Interleave (`.avi`) format and stored on a connected computer for subsequent processing and analysis. The `.avi` format was chosen as it ensures lossless compression support and compatibility with frame-accurate physiological signal extraction.

3.2. Experimental procedure

For the first acquisition system, 20 volunteers participated in the study. All participants reported no history of cardiovascular, neurological, or dermatological conditions. The imaging device used was a USB 3.0 RGB industrial camera (UI-3060CP), equipped with a Sony 2.3 MP sensor, selected for its high image quality and robust performance in low-light environments. A 25 mm C-mount lens with a VIS-NIR spectral range

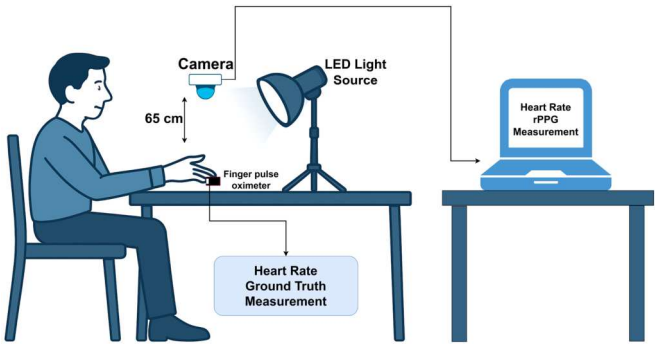

Figure 1. Schematic representation of the experimental setup.

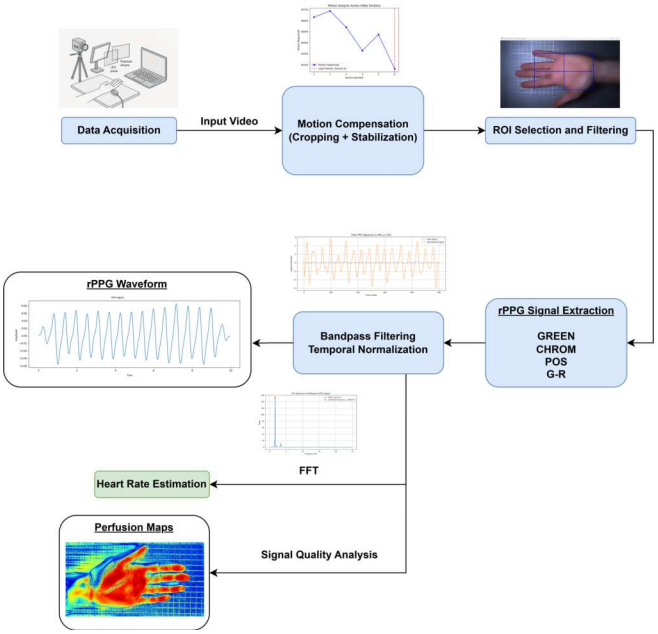

Figure 2. Workflow diagram of the experiment.

was mounted on the camera. During the experiment, each participant was comfortably seated on a chair and asked to place one hand on the table with the palm facing the camera. To obtain the ground truth heart rate, a pulse oximeter was attached to the index finger of the participant's opposite hand. The experiment was conducted in two stages. In the first stage, no polarization film was applied to the LED lighting, and no polarizing filter was used on the camera. A 60-second video of the participant's palm was recorded under these non-polarized lighting conditions. The camera recorded at a frame rate of 50 frames per second (fps), and the video was saved directly to the computer's hard drive in an uncompressed format. In the second stage, a linear polarization film was placed on the LED lighting kit, and a polarizing filter was attached to the camera lens. The recording procedure was repeated under these polarized lighting conditions, using the same setup and duration as before. This two-stage recording protocol was followed consistently for all 20 participants. The experimental setup and workflow diagram of the experiment are illustrated in figure 1 and 2 respectively. Sample raw images from two adjacent frames—captured under both polarized and non-polarized lighting conditions —are presented in figure 3 to illustrate the visual differences in illumination.

For the second acquisition system, a monochrome CMOS camera (UI-3240ML-NIR-GL) was employed. All recordings in this setup were conducted under cross-polarized illumination to reduce specular reflections and enhance signal quality. Three healthy volunteers with no known medical conditions participated in this study. As in the previous setup, the participants were seated in a chair and asked to place one hand on a table. For each participant, recordings were taken from three distinct anatomical regions: the palm, the forearm, and the dorsal side (back) of the hand. To explore the influence of different wavelengths on the acquisition of rPPG signal, nine optical

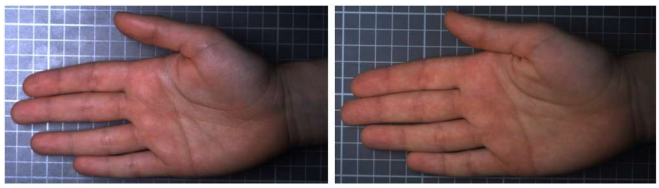

Figure 3. Sample raw images of the same palm recorded (a) without polarization filtration and (b) with cross polarization filtration.

Table 1. Technical specifications of optical bandpass filters used in the second acquisition system.

Filter name	Central wavelength (nm)	Spectral range (nm)
Light Red	650	620–680
Orange	600	580–620
NIR-UV	850	800–900
Cyan	500	480–520
Blue	470	450–490
NUV	380	360–400
Dark Red	700	680–720
UV-VIS	400	380–420
Green	540	520–560

bandpass filters were used. These filters -light red, orange, near-infrared-ultraviolet (NIR-UV), cyan, blue, near-ultraviolet (NUV), dark red, ultra-violet-visible (UV-VIS) and green were sequentially mounted on the camera lens during acquisition. The detailed technical specifications of each filter are presented in table 1. Description of data is displayed in table 2. For each participant, recordings were taken from all three anatomical positions using each of the nine filters, resulting in a total of 27 recordings per subject. Each video was recorded for 60 seconds under polarized lighting conditions only. The camera operated at a frame rate of 50 frames per second (fps), and the recorded data was stored on the computer hard drive in uncompressed format for further analysis.

In addition to the primary datasets, videos of two participants (with and without polarization filtration) were recorded to evaluate the robustness of the system against motion and to assess the effectiveness of motion compensation techniques. Each subject was instructed to perform both scaling and translational hand movements at varying speeds. Each recording lasted 60 seconds. The data acquisition procedure followed the same protocol as described in the first acquisition system, with the key distinction being the intentional introduction of motion. This motion robustness data set was specifically designed to test and validate the motion compensation algorithm used during the preprocessing phase of the experiment. Based on the evaluation, the motion compensation algorithm was subsequently applied to all datasets in the final experimental pipeline to ensure consistency and reduce motion artifacts.

Table 2. Dataset description.

Dataset name	Participants	Videos	Camera
Motion Compensation Dataset	2	4	RGB
RGB Dataset	20	40	RGB
Monochrome Dataset	2	54	Monochrome

3.3. Hardware setup

The implementation and tests of this research were performed on a PC with an Intel^{TM} i7-10700K processor and 32 GB RAM, running the Microsoft Windows 11 operating system and Python 3.10 environment. The cameras and lenses used for video recording and data collection were provided by IDS Imaging Development Systems. The ground truth heart rate was measured using a Contec^{TM} CMS 50E pulse oximeter.

3.4. Data preprocessing

To enhance the quality and reliability of remote photoplethysmography (rPPG) signals, it is essential to minimize motion-induced artifacts. The motion-induced artifacts occur due to both translational displacement of the hand and frame-to-frame jitter that can degrade temporal signal consistency. During the pre-processing phase, a motion-based video segmentation technique was implemented to automatically identify and extract the most stable 10 second segment from each recorded video. The approach involved analyzing the optical flow across consecutive video frames to estimate the magnitude of motion throughout the first 60 seconds of each video. To quantitatively define motion, the optical flow field between consecutive frames was computed using the Farnebäck algorithm (Farnebäck 2003). For each pixel, the motion vector was defined as:

$$m(x, y, t) = \sqrt{u(x, y, t)^2 + v(x, y, t)^2} \quad (1)$$

where $u(x, y, t)$ and $v(x, y, t)$ are the horizontal and vertical flow components, respectively. The motion magnitude for a frame was then calculated as the spatial average of $m(x, y, t)$ across the ROI. The full 60-second video was divided into six non-overlapping 10-second sections based on the frame rate. For each section, the accumulated motion magnitude was computed and the section with the lowest total motion was assumed to have the least movement, and

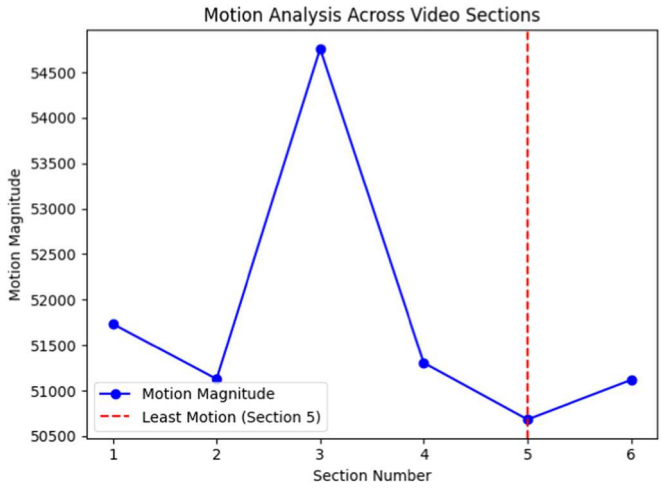

Figure 4. Motion magnitude analysis across six 10-second segments of the video for Subject 7. Section 5 exhibits the lowest overall motion, identified by the red dashed line, and was selected for further rPPG signal extraction due to its minimal motion artifacts.

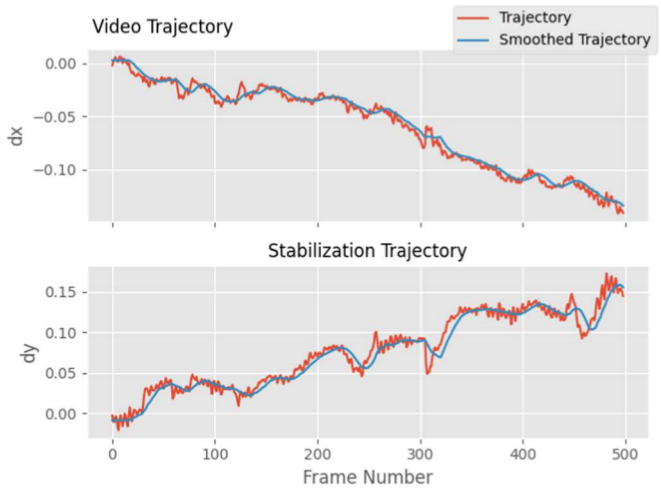

Figure 5. Trajectory analysis before and after video stabilization. The top plot shows the original horizontal displacement (dx) trajectory and its smoothed version, while the bottom plot illustrates the corresponding vertical displacement (dy) trajectory. The smoothed trajectories demonstrate reduced jitter, indicating effective stabilization.

consequently the least noise introduced by motion artifacts, making it ideal for rPPG signal extraction (see figure 4).

For further motion compensation, the best 10-second segment was stabilized using a video stabilization library called vidstab (Boer 2018). Although the lowest-motion segment was selected by optical flow analysis, it is still possible for small jitter or subtle drift to remain. To address this, a temporal video stabilization algorithm was implemented to enhance spatial consistency between frames and improve the quality of the input used for signal extraction. VidStab is a Python library that stabilizes videos by estimating and smoothing camera movement using feature points tracked between frames with the Kanade-Lucas-Tomasi (KLT) tracker (Yongyong *et al* 2020). It applies a moving average filter to reduce jitter and computes corrective transforms to align frames with a stable path. Warping is then used to adjust the frames, with borders added to fill any gaps. An example of the smoothed motion trajectory before and after stabilization is illustrated in figure 5. The final stabilized video frames are then fed into the rPPG extraction pipeline for further analysis.

This allowed us to capture both large-scale displacements (e.g., finger movements, hand shifts) and subtle micro-motions (e.g., tremors or drift). The subsequent stabilization step (VidStab) addressed residual jitter by aligning trajectories of tracked feature points.This motion compensation technique was applied to the motion robustness data set to evaluate its effectiveness by comparing the motion magnitude before and after stabilization. The results showed that the method not only automated the selection of stable video segments and reduced the magnitude of the overall motion, but also improved the quality of the

physiological signals extracted. Significant improvements in SNR and SQI were observed after applying the proposed motion compensation technique, which is further discussed in the results section.

3.5. ROI selection and filtering

In the first step to process the stabilized recorded frames, a Region of Interest (ROI) is manually selected, covering most parts of the skin region such as the palm, forearm, or dorsal side of the hand, from which the remote photoplethysmography (rPPG) signal is extracted. ROI selection is crucial for isolating areas that are rich in physiological signals and free from motion artifacts or background noise. Once the ROI is selected, a bilateral filter is applied to each ROI frame to reduce spatial noise while preserving important edge information (Chen *et al* 2017). The bilateral filter maintains the integrity of local spatial structures, which is critical in rPPG as it helps preserve subtle color variations caused by changes in blood volume in the skin. The bilateral filter operates based on both spatial distance and pixel intensity difference. For any given pixel, the filter considers not just the spatial proximity of the surrounding pixels (controlled by the parameter `sigmaSpace`) but also their photometric similarity (controlled by `sigmaColor`). The `d` parameter defines the diameter of the neighborhood of pixels used during filtering. The bilateral filter used in this study can be defined by:

$$I_{\text{filtered}}(x) = \frac{1}{W_p}\sum_{x_i\in\Omega} I(x_i) \cdot f_s(\|x_i - x\|) \cdot f_r(|I(x_i) - I(x)|) \quad (2)$$

Here, f_s is the spatial Gaussian function controlled by `sigmaSpace`, and f_r is the range Gaussian function controlled by `sigmaColor`. The term W_p is the normalization factor that ensures that the weights are equal to one. By tuning these three parameters, the filter smooths out random noise in homogeneous areas of the skin while retaining sharp boundaries between anatomical structures such as veins or creases. In this experiment, a `sigmaSpace` value of 75, and a pixel neighborhood diameter (`d`) of 55 were used. This enhances the quality of temporal pixel signals extracted from the video frames, allowing for more accurate frequency domain analysis in subsequent rPPG signal processing stages.

3.6. Rppg signal extraction

Remote photoplethysmography (rPPG) signal extraction is highly influenced by the way light interacts with human skin. According to the dichromatic reflection model (DRM) (Weijer and Beigpour 2011), the light captured comprises both specular and diffuse reflection components. Specular reflection, which occurs on the surface of the skin, does not carry physiological information, whereas diffuse reflection penetrates the skin, interacts with blood and tissue, and carries valuable signals related to changes in blood volume.

The observed skin color at a given pixel k and time t is modeled as:

$$C_k(t) = I(t) \cdot (\mathbf{v}_s(t) + \mathbf{v}_d(t)) + \mathbf{v}_n(t) \quad (3)$$

where $I(t)$ is the time-varying illumination, $\mathbf{v}_s(t)$ is the specular reflection, $\mathbf{v}_d(t)$ is the diffuse reflection, and $\mathbf{v}_n(t)$ is the camera noise.

The specular reflection $\mathbf{v}_s(t)$ and the diffuse reflection $\mathbf{v}_d(t)$ can be further modeled as:

$$\mathbf{v}_s(t) = \mathbf{u}_s \cdot (s_0 + s(t)) \quad (4)$$

$$\mathbf{v}_d(t) = \mathbf{u}_d \cdot d_0 + \mathbf{u}_p \cdot p(t) \quad (5)$$

where $\mathbf{u}_s$ and $\mathbf{u}_d$ are the unit color vectors for the light source and skin tissue, respectively; s_0 and d_0 are the steady state reflections; $s(t)$ accounts for motion-induced changes; $\mathbf{u}_p$ denotes the direction of pulse-related color variation, and $p(t)$ is the desired photoplethysmography signal.

Combining all terms, the full DRM expression becomes:

$$C_k(t) = I_0 \cdot (1 + i(t)) \cdot (\mathbf{u}_c \cdot c_0 + \mathbf{u}_s \cdot s(t) + \mathbf{u}_p \cdot p(t)) + \mathbf{v}_n(t) \quad (6)$$

The goal of rPPG is to isolate the physiological signal $p(t)$ from this mixture. The techniques used for this purpose fall into two broad categories: model-based methods, which leverage prior knowledge such as skin color or light spectra, and blind source separation (BSS)-based methods, which extract $p(t)$ statistically without needing prior assumptions.

In this study, four state-of-the-art model-based rPPG algorithms—GREEN (Verkruysse *et al* 2008b), CHROM (de Haan and Jeanne 2013), POS (Trumpp *et al* 2017), and G-R (Haugg *et al* 2023) were implemented to extract pulse waveform signals from video data. For the first acquisition system, which utilized an RGB camera, all four methods were applied to evaluate their performance under standard color imaging conditions. For the second acquisition system, which employed a monochrome camera, only the GREEN method was used. This limitation arises because CHROM, POS, and G-R methods rely on multichannel color information to perform signal extraction. Since monochrome cameras capture only luminance data without separating color channels, these methods are incompatible with such input and were therefore excluded.

3.6.1. GREEN method

The GREEN rPPG method is recognized as one of the most straightforward and widely adopted approaches for estimating heart rate using remote photoplethysmography. This method is based on the observation that the green channel of RGB video data contains the most prominent photoplethysmography signal because of the higher absorption of green light by

hemoglobin. In this approach, the pulse signal $S(t)$ is directly derived from the temporal variation of the green channel $x_g(t)$, as expressed by the equation:

$$S(t) = x_g(t) \tag{7}$$

Hemoglobin absorbs green light more effectively than red or blue light. As blood volume in the capillaries changes with each heartbeat, the reflectance of green light on the skin also varies. These small variations are captured by the green channel in a video sequence and can be processed to extract the pulse signal. The GREEN method assumes a relatively motionless subject and stable lighting conditions, as it lacks mechanisms for compensating for motion or color distortions.

3.6.2. CHROM method

The CHROM method, introduced by de Haan and Jeanne (de Haan and Jeanne 2013), was developed to address limitations in traditional rPPG techniques that are prone to specular reflections and color normalization problems. These issues arise due to unpredictable surface reflections that introduce noise into the signal. CHROM mitigates this by leveraging chromatic differences between color channels under the assumption that pulse-induced variations differ chromatically from illumination-induced artifacts.

First, the mean color signals from the Region of Interest (ROI) are computed for each frame over time and normalized to remove global intensity variations. Normalized RGB signals, denoted as $x_r(t)$, $x_g(t)$, and $x_b(t)$ for the red, green, and blue channels, respectively, are projected into two orthogonal chrominance signals:

$$X_C(t) = 3x_r(t) - 2x_g(t) \tag{8}$$

$$Y_C(t) = 1.5x_r(t) + x_g(t) - 1.5x_b(t) \tag{9}$$

Then a scaling factor α is calculated as the ratio of the standard deviations of the two components:

$$\alpha = \frac{\sigma(X_C(t))}{\sigma(Y_C(t))} \tag{10}$$

The final rPPG signal $S(t)$ is obtained by linearly combining the chrominance signals using α:

$$S(t) = X_C(t) - \alpha Y_C(t) \tag{11}$$

This method effectively suppresses lighting-related artifacts while improving the physiological signal, making it more robust to changes in motion and illumination.

3.6.3. POS method

The Plane-Orthogonal-to-Skin (POS) method, introduced by Wang *et al* (Wang *et al* 2017), was developed to improve the quality of the remote photoplethysmography (rPPG) signal by minimizing the effects of specular reflections that often contaminate signals from the skin surface. Like the CHROM method, POS also operates in a temporally normalized RGB space. However, instead of using chrominance projections, it defines a plane orthogonal to the skin tone direction to suppress static and motion-induced specular components while retaining the pulsatile signal.

The POS method follows a three-stage process. First, temporal normalization is applied to the RGB signals to remove illumination-related intensity variations. Then, two axes orthogonal to the skin-tone direction are computed from the normalized signals:

$$X_{\mathrm{POS}}(t) = x_g(t) - x_b(t) \tag{12}$$

$$Y_{\mathrm{POS}}(t) = x_g(t) + x_b(t) - 2x_r(t) \tag{13}$$

where $x_r(t)$, $x_g(t)$, and $x_b(t)$ denote the red, green, and blue channel signals over time.

Finally, the two components are combined using a weighting factor α, calculated similarly to the CHROM method, to produce the final rPPG signal:

$$S(t) = X_{\mathrm{POS}}(t) + \alpha Y_{\mathrm{POS}}(t) \tag{14}$$

Unlike CHROM, which often produces anti-phase chrominance components, the POS method ensures that the projected signals are in phase, thus improving robustness and interpretability.

3.6.4. G-R method

The G-R method (Haugg *et al* 2023), short for Green minus Red, is a simple yet effective approach for extracting rPPG signals from RGB video data. This method is based on the empirical observation that the green channel typically contains the strongest photoplethysmography signal because of hemoglobin absorption characteristics, while the red channel is more susceptible to noise and motion-related artifacts.

By subtracting the red channel from the green, the G-R method emphasizes the pulsatile component while reducing common-mode noise such as illumination variations and minor motion artifacts that affect both channels similarly.

The rPPG signal $S(t)$ is computed as:

$$S(t) = x_g(t) - x_r(t) \tag{15}$$

where $x_g(t)$ is the average intensity of the green channel in the selected region of interest (ROI) at time t, $x_r(t)$ is the average intensity of the red channel in the same ROI at time t, and $S(t)$ is the resulting photoplethysmography signal over time.

The extracted signal $S(t)$ is typically passed through a temporal band pass filter to retain frequency components within the physiological heart rate range. Compared to the basic GREEN method, the G-R method offers improved robustness under minor lighting fluctuations and motion disturbances by subtracting the red channel, which helps reduce non-pulsatile variations shared by both channels.

3.7. Temporal bandpass filtering

After extracting the raw rPPG signals from the video data, a temporal band-pass filter (BPF) (Blocher *et al* 2018) is applied to isolate the frequency components

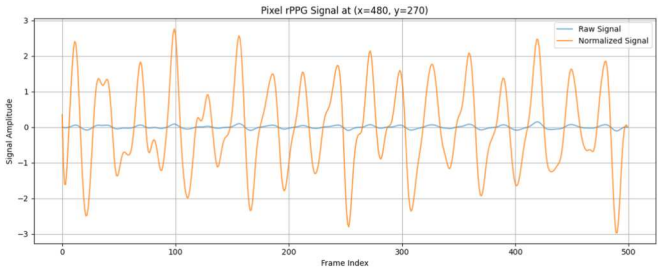

Figure 6. Comparison of the raw and normalized rPPG signals extracted from a single pixel.

corresponding to the heart rate. The main goal of this filtering step is to suppress unrelated noise and motion artifacts while retaining only physiological oscillations within the typical heart rate range. In this implementation, a second-order Butterworth band-pass filter is used, characterized by a low cut-off frequency of 0.7 Hz and a high cut-off frequency of 4 Hz. These frequency bounds were chosen to cover the standard range of human heart rates, which typically fall between 42 bpm and 240 bpm. The Butterworth filter is preferred due to its flat frequency response in the pass-band and smooth roll-off, making it well-suited for physiological signal processing. The filter is applied to the temporal traces obtained from the manually selected region of interest.

3.8. Window-based temporal normalization

To further improve the robustness of the extraction of rPPG signals against illumination variations and camera exposure fluctuations, a window-based temporal normalization (de Haan and Jeanne 2013) strategy was implemented. This approach is especially critical in model-based methods like CHROM and POS, where accurate estimation of chrominance or projection vectors relies on consistent color channel dynamics over time.

In this method, the input signals are segmented into sliding temporal windows of length L, forming a matrix $C \in \mathbb{R}^{3\times L}$, where each row corresponds to one of the RGB color channels and each column represents a time step within the window. The RGB signals within each window are normalized by their temporal mean to remove global intensity variations, ensuring only relative chromatic fluctuations are retained. This is computed using the transformation:

$$C_n(t) = \text{inv}(\text{diag}(\boldsymbol{\mu}_C)) \cdot C(t) - \mathbf{1} \tag{16}$$

where $\boldsymbol{\mu}_C$ is the mean RGB vector of the reference ROI over the window, $\text{diag}(\boldsymbol{\mu}_C)$ is a diagonal matrix constructed from it, and $\mathbf{1}$ is a vector of ones used for mean centering. This normalization ensures that the pulse-induced variations in color are preserved while removing the influence of ambient lighting changes.

3.9. Heart rate estimation

In this study, heart rate (HR) was estimated from the extracted rPPG signals using a frequency domain approach based on the Fast Fourier Transform (FFT). This method computes the average heart rate throughout the duration of the signal by identifying the dominant frequency component within the physiological heart rate range, typically between 0.7 Hz and 4 Hz.

First, the one-dimensional rPPG signal $s(t)$ is transformed into the frequency domain using FFT to obtain its power spectrum $P(f)$. This spectrum reflects the energy distribution of signal frequencies over the duration of the measurement. The signal is typically zero-padded to three times its original size to improve frequency resolution. The dominant frequency f_{peak} is identified as the frequency with the highest power in the spectrum:

$$f_{\text{peak}} = \arg\max_f(|\,\text{FFT}(s(t))|^2) \tag{17}$$

The corresponding heart rate in beats per minute (bpm) is then computed as:

$$\text{HR} = 60 \times f_{\text{peak}} \tag{18}$$

This FFT-based heart rate computation method has been widely adopted in rPPG research due to its computational efficiency and effectiveness under constrained conditions (Sun and Thakor 2016). Figures 6, 7 and 8 illustrate the Pixel rPPG sinal, FFT spectrum and the final rPPG signal, respectively, for one of the subjects in the collected data set.

3.10. Performance evaluation metrics

To evaluate the quality and reliability of the extracted rPPG signals and the derived heart rate (HR) estimates, three key performance metrics Absolute Error (AE) in beats per minute (bpm), Template Correlation-Based Signal Quality Index (SQI) and Signal-to-Noise Ratio (SNR) are used. These metrics assess both the accuracy of HR estimation and the signal quality of the extracted waveforms.

3.10.1. Absolute error

Absolute Error (AE) is used in this study to evaluate the accuracy of heart rate (HR) estimation by comparing the value derived from the rPPG signal with the

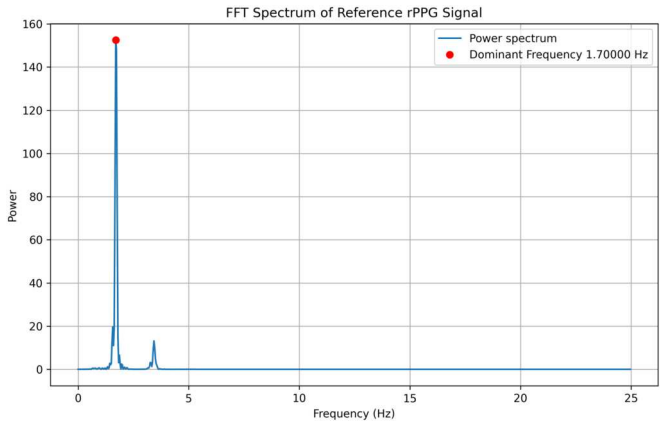

Figure 7. FFT of the reference signal (Subject 5).

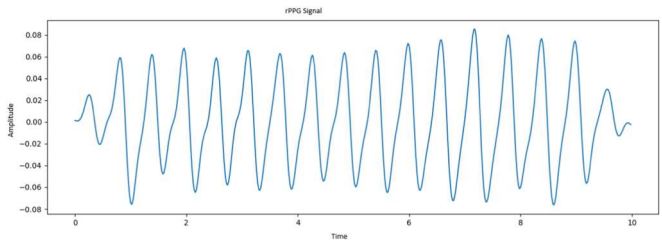

Figure 8. Final rPPG extracted signal (Subject 5).

ground truth measurement obtained via a pulse oximeter during data acquisition. It is calculated as the absolute difference between the estimated HR and the reference HR for each measurement window.

$$\mathrm{AE} = |HR_{\mathrm{rPPG}} - HR_{\mathrm{GT}}| \tag{19}$$

where:

- HR_{rPPG} denotes the heart rate estimated from the rPPG signal, and
- HR_{GT} represents the ground truth heart rate obtained from the reference device.

Lower AE values indicate a smaller discrepancy between the estimated and reference HR, reflecting improved accuracy in rPPG-based heart rate estimation.

3.10.2. Signal-to-Noise Ratio (SNR)

The Signal-to-Noise Ratio (SNR) is a key performance metric used to evaluate the quality of rPPG signals, particularly for estimating the spatial distribution of vascular perfusion. In this work, the SNR is calculated in the frequency domain by analyzing the power of the main pulsatile component relative to the background noise.

The rPPG signal of each pixel or region, $S_i(t)$, is transformed into the frequency domain using the Fourier Transform:

$$\mathcal{F}\{S_i(t)\} = S_i(f) \tag{20}$$

The physiological heart rate range of interest is defined between f_1 and f_2, corresponding to 40–240 bpm (approximately 0.67–4 Hz in this study).

To estimate the SNR, a double-step frequency mask is constructed to separate the signal and noise components. The signal mask $h^i_{\mathrm{signal}}(f)$ is defined as:

$$h^i_{\mathrm{signal}}(f) = [\delta(f - f_0^i) + \delta(f - 2f_0^i)] * \Pi(\pm f_r) \tag{21}$$

where δ is the Dirac delta function, f_0^i is the fundamental frequency (peak of the periodogram), $2f_0^i$ is its second harmonic, and $\Pi(\pm f_r)$ is a rectangular window of half-width f_r, used to allow a finite bandwidth around each harmonic.

The corresponding noise mask is defined as the complement:

$$h^i_{\mathrm{noise}}(f) = 1 - h^i_{\mathrm{signal}}(f) \tag{22}$$

The final SNR is computed as the ratio of signal power to noise power, integrated over the physiological range:

$$SNR_i = 10 \cdot \log_{10}\left(\frac{\int_{f_1}^{f_2} h^i_{\mathrm{signal}}(f)|S_i(f)|^2\, df}{\int_{f_1}^{f_2} h^i_{\mathrm{noise}}(f)|S_i(f)|^2\, df}\right) \tag{23}$$

This formulation ensures that the SNR reflects the prominence of the pulse-related components (fundamental and harmonic) relative to the rest of the spectral content, providing a robust measure of signal

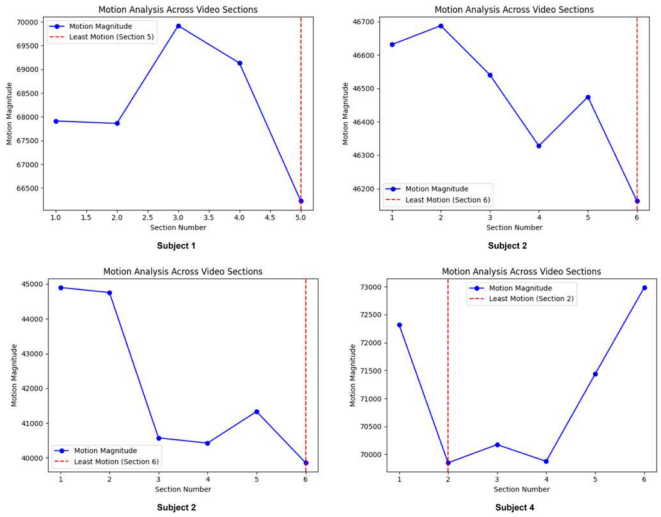

Figure 9. Average motion analysis for the motion compensation dataset.

quality. High SNR values indicate strong, clean pulsatile signals (typically over skin regions), while low SNR values correspond to background or noise-dominated regions.

Spatially resolved SNR heatmaps are generated by applying this computation pixel-wise across the Region of Interest (ROI), enabling the visualization of areas with high-quality perfusion signals.

3.10.3. Vascular perfusion mapping

To obtain vascular perfusion maps, the pixel-wise rPPG signals extracted from the Region of Interest (ROI) were analyzed in the frequency domain. For each pixel, the signal-to-noise ratio (SNR) and signal quality index (SQI) were computed using the procedures described in sections 3.10.2 and 3.10.3. These values were then spatially arranged to form *SNR maps* and *SQI (template-correlation) maps*, providing a two-dimensional visualization of perfusion quality across the skin surface. Regions with high SNR and SQI correspond to areas where blood-volume pulsations are more strongly represented, reflecting better vascular perfusion.

To enhance interpretability, the maps were normalized to a consistent intensity scale and displayed as heatmaps, with intense colors indicating higher perfusion quality. This allowed not only quantitative comparison of methods and acquisition conditions but also qualitative visualization of how perfusion varies across different anatomical regions (palm, forearm, dorsal hand).

4. Results

4.1. Motion compensation results

To assess the effectiveness of the proposed motion compensation strategy, the average motion magnitude was calculated for four video sequences (two with cross polarization filtration and two without polarization filtration) before and after the application of the stabilization pipeline. The results, summarized in table 3, indicate a consistent reduction in motion magnitude across all videos following the application of the segmentation and stabilization procedures. The motion-based segmentation technique successfully identified the most stable 10-second interval of each of video by analyzing the optical flow magnitudes throughout the video segment as illustrated in figure 9. This automated selection step alone contributed to a significant reduction in overall motion energy by focusing on periods with minimal subject movement. Subsequently, temporal video stabilization was applied to the selected segments using the VidStab library, which further minimized residual motion such as jitter or subtle drifts. By aligning frames more consistently, the stabilization step enhanced temporal coherence and reduced spatial displacement artifacts that can degrade signal quality in pixel-wise rPPG extraction.

Table 3. Average motion magnitude for the **motion compensation dataset** before and after motion compensation.

Subject	Before motion compensation	After motion compensation
Subject 1	2837.49	1816.95
Subject 2	2016.88	1915.04
Subject 3	6309.89	5995.27
Subject 4	1192.79	1090.70

The reduction in motion magnitude not only validates the effectiveness of the stabilization approach but also correlates with improved signal integrity in the rPPG processing pipeline. To assess the impact of motion compensation on signal quality, two key performance metrics were analyzed: Signal-to-Noise Ratio (SNR) and Signal Quality Index (SQI). These

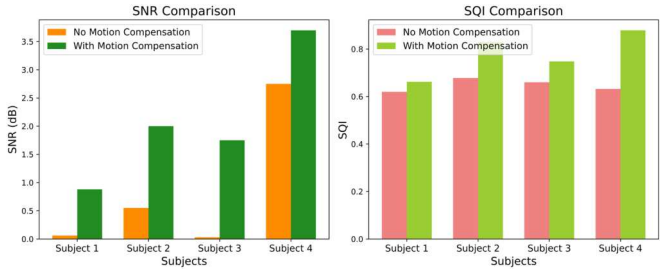

Figure 10. SNR and SQI Comparison for the motion compensation dataset.

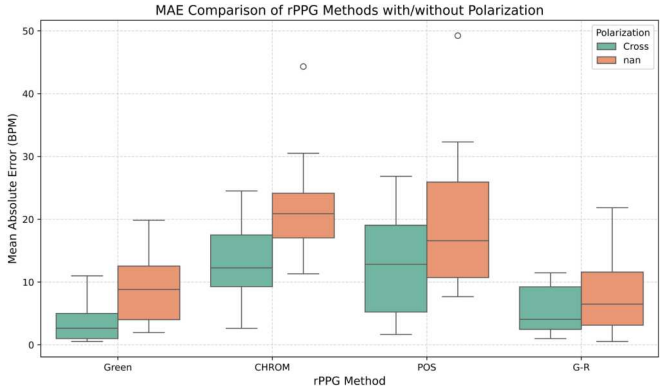

Figure 11. Comparison of Mean Absolute Error (MAE) across different rPPG signal extraction methods (GREEN, CHROM, POS, G-R) under cross-polarized and non-polarized conditions. Lower MAE values indicate higher accuracy in heart rate estimation.

metrics were computed from the spatially averaged reference rPPG signals, extracted using the GREEN method over the selected ROI. As shown in figure 10, both the SNR and the SQI values exhibited consistent improvements after motion compensation, confirming enhanced signal fidelity and greater morphological consistency of the extracted pulse waveforms.

4.2. rPPG signal quality under polarized and non-polarized lighting conditions

This section presents the performance evaluation of the extraction of rPPG signals under two different lighting conditions - nonpolarized and crosspolarized using video recordings from 20 subjects acquired in the first acquisition system. The recordings were processed using four rPPG algorithms: GREEN, CHROM, POS, and G-R. The quality and reliability of the extracted physiological signals were quantitatively evaluated using three performance metrics: Mean Absolute Error (MAE), Signal Quality Index (SQI), and Signal-to-Noise Ratio(SNR).

Figure 11 illustrates the comparison of Absolute Error (AE) across the four rPPG methods under cross-polarized and non-polarized lighting conditions. AE was computed as the absolute difference between the estimated rPPG-based heart rate and the ground truth heart rate obtained from the pulse oximeter. As shown in the figure, all methods demonstrate lower AE values when cross-polarized illumination is used. The reduction in estimation error is particularly significant for the GREEN and G-R algorithms, highlighting the role of cross-polarization in improving signal accuracy. This improvement is primarily attributed to the suppression of specular reflections and illumination-induced artifacts, which otherwise degrade the signal quality in non-contact heart rate monitoring.

Figure 12 presents the subject-wise comparison of mean SNR values for the four rPPG algorithms. SNR quantifies the ratio of pulsatile signal strength to noise and is an important indicator of signal fidelity. From the figure, it is evident that GREEN and G-R methods consistently produce higher SNR values across most subjects, demonstrating their ability to extract cleaner signals even under varying motion and illumination conditions. POS achieves moderate SNR performance, while CHROM exhibits relatively lower SNR, suggesting it is more sensitive to ambient disturbances.

To further evaluate the morphological consistency of the extracted rPPG signals relative to the reference signal derived from the spatially averaged ROI, the Signal Quality Index (SQI) was calculated using template correlation. Figure 13 compares the SQI values in 20 subjects under both lighting conditions. The results show that cross-polarization consistently yields higher SQI values, indicating better alignment and structural coherence of the pulse waveforms This improvement is attributed to the

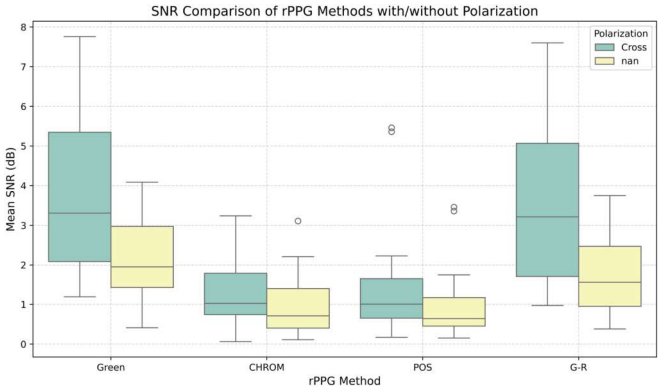

Figure 12. Mean Signal-to-Noise Ratio (SNR) across 20 subjects for four different rPPG algorithms: GREEN, POS, CHROM, and G-R under cross-polarized and non-polarized conditions. Higher SNR values indicate better signal fidelity and robustness against noise.

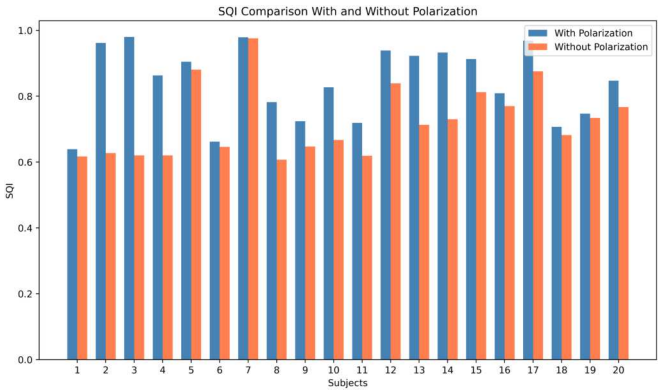

Figure 13. Comparison of Signal Quality Index (SQI) across 20 subjects with and without cross-polarization. Each subject's signal quality is evaluated under both illumination setups to assess the impact of polarization on morphological consistency of extracted rPPG signals.

Table 4. Average performance metrics for the 20 subjects under different polarization conditions.

Method	Average error	Average SNR	Average SQI	Polarization
GREEN	8.20	2.12	0.74	No Polarization
CHROM	21.86	0.97	0.69	
POS	18.70	0.99	0.63	
G-R	9.25	1.82	0.73	
GREEN	3.79	3.82	0.81	Cross Polarization
CHROM	13.79	1.34	0.70	
POS	12.52	1.45	0.71	
G-R	5.49	3.45	0.80	

reduction of glare and specular highlights, enabling clearer detection of color variations driven by changes in blood volume. These findings confirm the benefits of cross-polarization in improving the morphological integrity of rPPG signals.

The average values of AE, SNR, and SQI across all 20 subjects are summarized in table 4. Overall, the application of cross-polarized illumination results in marked improvements in all performance metrics. In particular, GREEN achieves the highest average SQI, while G-R demonstrates the best performance in terms of SNR and AE under polarized conditions. Although the GREEN method is straightforward and computationally efficient, its performance is comparatively more affected by specular distortions, particularly under non- polarized lighting.

4.2.1. Spatial signal quality analysis

The effect of using cross-polarized illumination on spatial signal quality in remote photoplethysmography (rPPG) is demonstrated in figures 14 and 15. These figures present comparative Signal-to-Noise

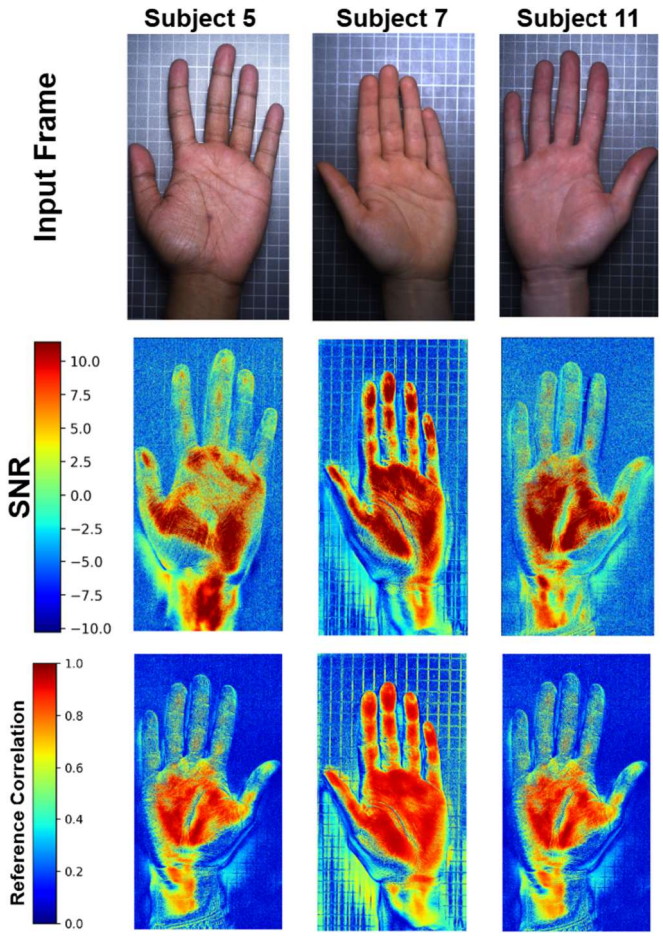

Figure 14. SNR Map and Reference Correlation map of three representative subjects under non-polarized conditions.

Ratio (SNR) maps and reference correlation maps for three subjects under scenarios involving cross-polarization and non-polarized conditions. Under cross-polarized illumination, the SNR maps exhibit higher and more spatially uniform SNR values across the palm region. This improvement is attributed to the suppression of specular reflections, which allows the camera to capture diffuse reflections that carry physiologically relevant pulse information of the blood volume. The corresponding reference correlation maps similarly reveal higher correlation values over more extensive regions of the palm, suggesting that the extracted pulse waveforms exhibit greater temporal consistency and morphological coherence relative to the reference pulse template. In contrast, under non-polarized conditions (figure 14), the SNR maps reveal more heterogeneous patterns, with regions of low or negative SNR, particularly in areas affected by glare (e.g., fingertips, central palm). Similarly, the reference correlation maps show noticeably lower correlation values, indicating greater variability and distortion in the pulse waveforms extracted from these regions.

The significance of achieving higher SNR and higher correlation in rPPG signal extraction lies in their direct impact on signal fidelity and analysis robustness. A higher SNR indicates a stronger ratio of the desired pulsatile signal to unwanted noise, thus enhancing the detectability of heart rate and its harmonics in the frequency domain, which in turn improves the accuracy and reliability of heart rate estimation. Similarly, a higher reference correlation means that the extracted pulse waveforms remain consistent across time and spatial regions, an essential characteristic for advanced applications such as heart rate variability (HRV) analysis, waveform morphology studies, and the detection of subtle physiological changes. Insufficient SNR can result in noise that can distort the extracted signal, leading to inaccurate or unstable heart rate estimates, while poor correlation can cause the waveforms to deviate from the true physiological oscillations. Overall, these results confirm that cross-polarized illumination substantially improves the robustness and reliability of rPPG signal extraction by mitigating specular reflections and enhancing the contrast of blood volume-induced color variations. This highlights the necessity of integrating optical filtering methods, like cross-polarization, into rPPG collection processes to ensure reliable and high-quality physiological monitoring.

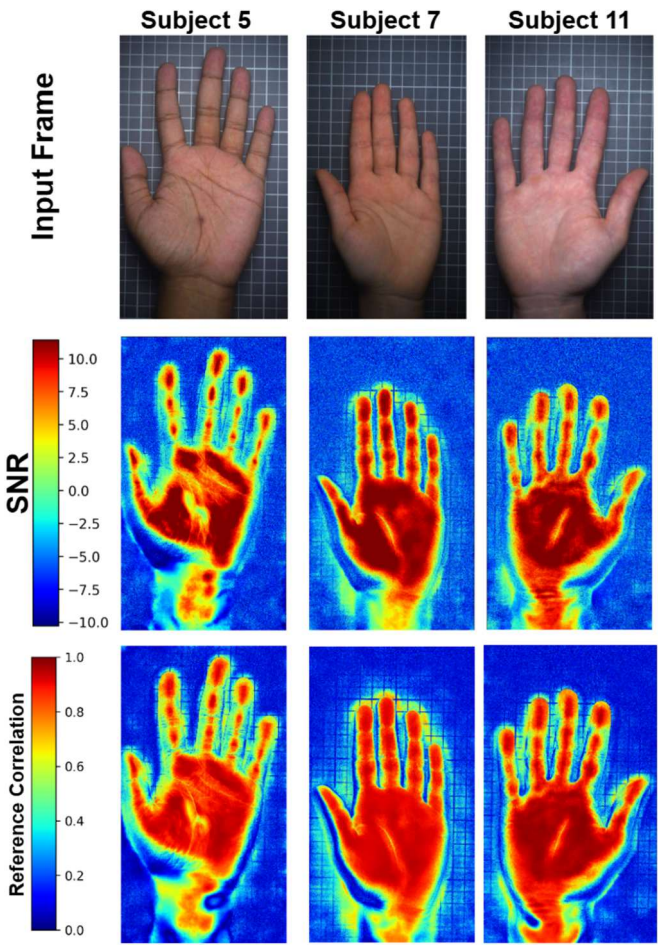

Figure 15. SNR Map and Reference Correlation map of three representative subjects under cross-polarized conditions.

4.3. Results of the monochrome acquisition system with spectral filtering

This section presents the experimental results obtained using the second acquisition system, which was designed to investigate the effects of wavelength-specific optical filtering on the quality of the rPPG signal. A monochrome CMOS camera (UI-3240ML-NIR-GL, IDS Imaging) paired with cross-polarized illumination was used to reduce specular reflections and enhance signal fidelity. The main goal of this experimental setup was to see how different spectral bands influence the accuracy of the quality of the extracted rPPG signal and perfusion maps, using the well-established GREEN method for signal extraction.

To quantitatively evaluate the performance of each spectral condition, the same three complementary metrics for the first acquisition system were computed: Absolute error (AE) between the estimated heart rate and the ground truth (GT) reference obtained from the pulse oximeter, the Signal-to-Noise Ratio (SNR) of the reference rPPG signal and the signal quality index (SQI), computed using template correlation (TC).

The results for the palm region (table 5) reveal that filters corresponding to green, cyan, and orange wavelengths consistently gave better performance, with lower AE values and higher SQI and SNR scores. In particular, the Green filter achieved the lowest AE (4.44 BPM) and the highest SNR (−0.74 dB), accompanied by a robust SQI of 0.781. In contrast, filters in the UV and NIR spectral ranges (NUV, UV-VIS, NIR-UV) gave higher AE values and degraded SNR, which indicates weaker or noisier rPPG signals. These findings suggest that visible wavelengths centered around green to orange bands optimally capture the pulsatile blood volume variations in the palm region.

Table 6 summarizes the results for the forearm region. The UV-VIS filter achieved the best performance, with the lowest Absolute Error and highest SQI (0.810). The dark red, green, and orange filters also showed competitive SQI values. Compared to the palm, longer wavelengths performed better in the forearm, likely due to deeper vascular structures and greater tissue thickness that favor signal penetration at these wavelengths.

The dorsal side of the hand exhibited the greatest variability in performance across filters as illustrated in (table 7). The Green filter achieved the lowest AE (13.66 BPM), although the Blue filter provided the highest SQI (0.823), indicating a strong signal quality. In contrast, filters in the NUV and Dark Red ranges

Table 5. Average performance metrics monochrome dataset for different optical bandpass filters on palm region.

Filters	Reference BPM	GT (BPM)	Absolute error	SNR reference (dB)	SQI (T.C)
Blue	68	73.21	5.21	−1.74	0.669
Cyan	68	74	6	−7.70	0.815
Dark Red	104	74	30	−5.96	0.688
Green	68	72.44	4.44	−0.74	0.781
Light Red	66	73	7	−3.96	0.673
NIR-UV	62	73	11	−4.53	0.829
NUV	124	77	47	−2.87	0.592
Orange	60	74.22	14.22	−5.30	0.598
UV-VIS	58	81	23	−5.79	0.781

Table 6. Average performance metrics on monochrome dataset for different optical bandpass filters on forearm region.

Filters	Reference BPM	GT (BPM)	Absolute error	SNR reference (dB)	SQI (T.C)
Blue	68	78.00	10.00	−3.13	0.611
Cyan	64	79.10	15.10	−3.83	0.707
Dark Red	77	80.00	3.00	−4.50	0.715
Green	74	81.55	7.55	−2.87	0.676
Light Red	72	81.36	9.36	−4.34	0.740
NIR-UV	64	82.00	18.00	−2.28	0.757
NUV	70	81.22	11.22	−7.39	0.673
Orange	66	81.11	15.11	−5.04	0.764
UV-VIS	82	81.00	1.00	−7.88	0.810

Table 7. Average performance metrics monochrome dataset for different optical bandpass filters on dorsal side of the hand.

Filters	Reference BPM	GT (BPM)	Absolute error	SNR reference (dB)	SQI (T.C)
Blue	58	81.55	23.55	−3.13	0.823
Cyan	60	80.00	20.00	−2.90	0.715
Dark Red	108	82.00	26.00	−10.71	0.699
Green	68	81.66	13.66	−3.21	0.581
Light Red	123	81.00	42.00	−7.28	0.751
NIR-UV	101	82.00	19.00	−6.41	0.707
NUV	109	81.33	27.67	−7.99	0.661
Orange	66	81.11	15.01	−4.84	0.754
UV-VIS	101	81.00	20.00	−6.64	0.736

resulted in higher AE and degraded SNR, confirming that shorter wavelengths and certain longer wavelengths are suboptimal for this anatomical site, likely due to thicker skin and reduced vascular visibility on the dorsal surface.

These results underscore the critical role of wavelength selection in optimizing rPPG signal acquisition. While green and orange bands performed robustly across all regions—particularly in the palm—longer wavelengths (e.g., dark red, NIR-UV) showed improved penetration and signal stability in the forearm and dorsal hand. Figures 16 and 17 shows the comparison metrics and figure 18 presents the spatial SNR maps for the palm region across the different optical bandpass filters evaluated in this study. These visualizations provide qualitative confirmation of the quantitative results reported earlier highlighting how the choice of wavelength impacts both the signal quality and the spatial distribution of usable rPPG signals.

5. Discussion

This study introduces a new approach that integrates motion-compensated preprocessing, dual-camera acquisitions (RGB and monochrome), and optical filtering (polarization and spectral bands) into a unified framework for heart rate estimation and spatial perfusion mapping. To the best of our knowledge, this is the first work to systematically evaluate these combined factors using a hand-based dataset. We explored how polarization, spectral filtering, and motion compensation affect the quality of rPPG signals and the accuracy of spatial perfusion mapping using spatio-temporal rPPG (ST-rPPG).

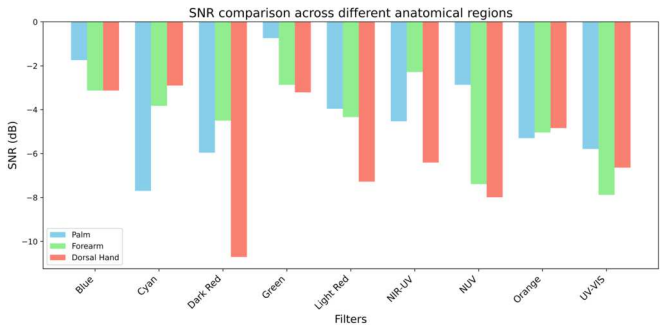

Figure 16. Comparison of Signal-to-Noise Ratio (SNR) across different anatomical regions (palm, forearm, dorsal hand) and optical bandpass filters.

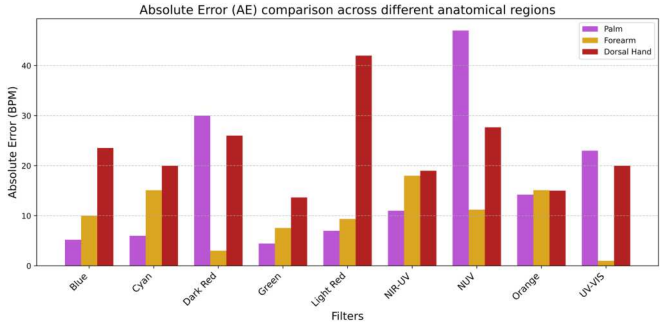

Figure 17. Comparison of Absolute Error (AE) in heart rate estimation across different anatomical regions (palm, forearm, dorsal hand) and optical bandpass filters.

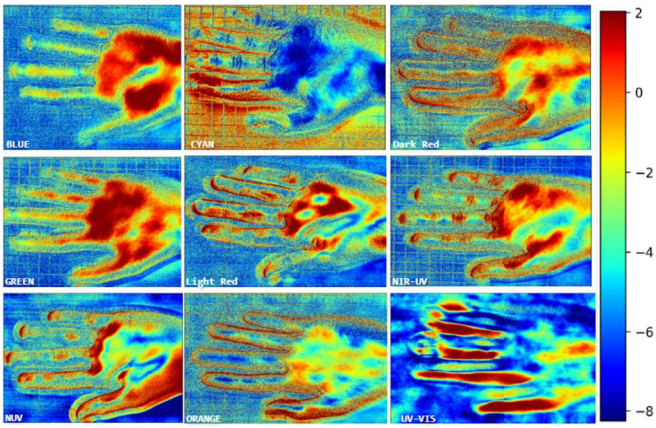

Figure 18. Spatial SNR maps of the palm region across different optical bandpass filters.

The combined approach of motion-based segmentation and temporal video stabilization significantly reduced average motion magnitude across all datasets. This reduction in motion was consistent and measurable, validating the effectiveness of the proposed preprocessing strategy. Importantly, this decrease in motion correlated with substantial improvements in both the Signal-to-Noise Ratio (SNR) and the Signal Quality Index (SQI). The improved SNR reflects a clearer physiological signal in the frequency domain, while the higher SQI values indicate better morphological consistency in the temporal waveform. The optical flow-based segment selection ensured that only the most stable portion of each video (typically 10 seconds) was used for rPPG extraction, minimizing gross motion distortions. The subsequent stabilization step (using VidStab) further reduced residual jitter and intra-frame displacement, which is especially important for pixel-wise processing where even minor frame-to-frame misalignment can degrade spatial signal integrity. These improvements translated to clearer pulsatile signals and more morphologically consistent pulse waveforms.

Under the first acquisition system, performance comparisons between polarized and non-polarized lighting conditions revealed that cross-polarization significantly enhances rPPG signal quality. Specifically, the Absolute Error (AE) decreased, and both SQI and SNR increased across all four tested rPPG algorithms (GREEN, CHROM, POS, and G-R) when polarization was applied. This outcome is consistent with the known benefits of polarization in suppressing specular reflections and emphasizing subsurface scattering, which is more directly linked to blood volume changes. Visual analysis of SNR maps and correlation maps further confirmed that cross-polarization reduces spatial heterogeneity and glare-induced artifacts, leading to more uniform and reliable perfusion maps. Among the four algorithms tested on RGB data, the G-R and GREEN methods consistently outperformed CHROM and POS in terms of SNR and MAE. This is likely due to their simplicity and stronger reliance on the green channel, which, as expected, contains the most prominent rPPG signal due to hemoglobin absorption characteristics. POS and CHROM, while more sophisticated and robust to some color distortions, showed lower SNR and higher sensitivity to motion and lighting changes, possibly due to their reliance on more complex projections of chrominance components that may be more vulnerable to residual noise in real-world recordings.

The results of the second acquisition system (monochrome camera + spectral filters) highlight the critical importance of wavelength selection to optimize the quality of the rPPG signal across different anatomical regions. These results are consistent with the optical properties of skin and blood. Green wavelengths are optimal for detecting superficial perfusion signals due to high absorption by hemoglobin and moderate scattering, making them ideal for thin skin regions like the palm. Longer wavelengths penetrate deeper, making them more suitable for thicker tissues or regions with deeper vasculature, such as the forearm. UV and NIR wavelengths generally performed poorly, confirming that these ranges either penetrate too shallowly (UV) or are not sufficiently absorbed by hemoglobin (NIR), leading to weaker pulsatile signals.Spatial SNR maps visually confirmed these trends, with green and blue filters producing broader and more uniform high-SNR regions on the palm. These vascular perfusion maps are not direct measurements of absolute blood flow but rather spatial quality maps of rPPG-derived pulsatile signals, which serve as a proxy for vascular perfusion. This representation provides a practical and non-invasive method to visualize microvascular distribution patterns.

This study demonstrates that high-resolution spatial perfusion mapping is achievable with relatively affordable hardware, including industrial cameras, LED panels, and spectral filters, which broadens the potential for clinical adoption of rPPG. Motion compensation emerges as an essential component for ensuring spatial signal fidelity and should be a standard part of any processing pipeline. Cross-polarization significantly enhances signal quality while adding minimal system complexity, making it highly beneficial. Future research can be focused on automating ROI selection using Artificial Intelligence to streamline workflows, exploring dynamic wavelength selection or multispectral fusion to improve signal extraction across diverse skin types and anatomical sites, validating system performance on patients with cardiovascular conditions, and developing real-time implementations to enable online perfusion monitoring.

6. Conclusions

This study investigated strategies to enhance the quality and spatial resolution of short-term remote photoplethysmography (ST-rPPG) for non-contact vascular perfusion assessment. The work addressed three key challenges—motion artifacts, specular reflections, and wavelength selection by integrating motion compensation, cross-polarization, and spectral filtering into the ST-rPPG acquisition and processing pipeline. By combining motion compensation through optical flow-based segmentation and video stabilization, the study demonstrated a considerable enhancement in signal fidelity, reflected by high SNR and SQI values across the data sets. Cross-polarization was shown to further suppress specular reflections, yielding higher spatial uniformity and accuracy in perfusion maps. The application of cross-polarized illumination resulted in precise and robust heart rate measurements by four robust rPPG algorithms. Among them, the GREEN and G-R method demonstrated excellent robustness under different lighting conditions. The spectral filter analysis utilizing a single-color camera indicated that visible green to orange wavelengths give bertter rPPG signals confirming the value of wavelength selection in system design. This study validated that high-resolution spatial perfusion mapping can be achieved with accessible hardware, supporting the feasibility of deploying rPPG systems in broader clinical and home health applications. The findings provide clear guidelines for system design and suggest that optimized ST-rPPG offers a promising tool for applications ranging from cardiovascular monitoring to dermatology and wound evaluation. Overall, these results show that careful optimization of acquisition parameters such as motion, illumination, and wavelength can significantly improve the reliability and usability of rPPG systems.

Acknowledgments

The research reported in this publication was supported by the Imaging and Artificial Vision Laboratory (ImVIA), Universite Bourgogne Europe (uB), France. The content was solely the author's responsibility and does not necessarily represent the official views of the sponsors. This work was supported by the French National Research Agency(ANR) under grant ANR-24-CE45-7356.

Data availability statement

The data cannot be made publicly available upon publication because no suitable repository exists for hosting data in this field of study. The data that support the findings of this study are available upon reasonable request from the authors.

Author contributions

Faisal Farhan 0000-0002-4741-4812
Conceptualization (lead), Data curation (lead), Formal analysis (lead), Methodology (lead), Validation (lead), Visualization (lead), Writing – original draft (lead), Writing – review & editing (lead)

References

Alizadeh-Meghrazi M *et al* 2021 Evaluation of dry textile electrodes for long-term electrocardiographic monitoring *BioMedical Engineering OnLine* **20** 82

B N, R K S, B L P, M R and N J 2018 *Pearson's Comprehensive Medical Assisting: Administrative and Clinical Competencies* 4 ed. (Pearson)

Blocher T, Zhou K, Krause S and Stork W 2018 An adaptive bandpass filter based on temporal spectrogram analysis for photoplethysmography imaging, in: 2018 *IEEE 20th International Workshop on Multimedia Signal Processing (MMSP), IEEE* 16

Boer C 2018 vidstab: Video stabilization using opencv https://github.com/adamrehn/vidstab Accessed: 2025-05-28

Botina-Monsalve D, Benezeth Y and Miteran J 2022 Performance analysis of remote photoplethysmography deep filtering using long short-term memory neural network *BioMedical Engineering OnLine* **21** 1–20

Chan G, Cooper R and HMWKKPZDAJADLNFREM 2019 Multi-site photoplethysmography technology for blood pressure assessment: challenges and recommendations *J Clin Med* (https://doi.org/10.3390/jcm8111827)

Charlton P H, Kyriaco P A and MJMVCPAJ 2022 Wearable photoplethysmography for cardiovascular monitoring *Proc IEEE Inst Electr Electron Eng* **110** 355–81

Chen Y, Xu Y and He J 2017 Bilateral filtering for biomedical image denoising: a review *Biomed. Signal Process. Control* **33** 69–79

Debnath U and Kim S 2025 A comprehensive review of heart rate measurement using remote photoplethysmography and deep learning *Biomed Eng Online* **24**

Farnebäck G 2003 *Two-Frame Motion Estimation Based on Polynomial Expansion* (Springer) 363–70

de Haan G and Jeanne V 2013 Robust pulse rate from chrominance-based rppg *IEEE Trans. Biomed. Eng.* **60** 2878–86

Hassan M, Malik A, Fofi D, Saad N, Karasfi B, Ali Y and Meriaudeau F 2017 Heart rate estimation using facial video: a review *Biomed. Signal Process. Control* **38** 346–60

Haugg F, Elgendi M and Menon C 2023 Grgb rppg: an efficient low-complexity remote photoplethysmography-based algorithm for heart rate estimation *Bioengineering* **10** 243

Hsu Y, Lin Y L and Hsu W 2014 Learning-based heart rate detection from remote photoplethysmography features2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) IEEE **15** 4433–37

Huang R Y and Dung L R 2016 Measurement of heart rate variability using off-the-shelf smart phones *BioMedical Engineering OnLine* **15**

Hülsbusch M 2008 Ein bildgestütztes, funktionelles Verfahren zur optoelektronischen Erfassung der Hautperfusion *Ph.D. thesis. RWTH Aachen University. Aachen.* https://publications.rwth-aachen.de/record/50278zsfassung in dt. und engl. Sprache; Aachen, Techn. Hochsch., Diss., 2008

Jeong I C and Finkelstein J 2016 Introducing contactless blood pressure assessment using a high speed video camera *Journal of Medical Systems* **40**

Kim K B and Baek H J 2023 Photoplethysmography in wearable devices: a comprehensive review of technological advances, current challenges, and future directions *Electronics* **12**

Lee R J, Sivakumar S and Lim K H 2023 Review on remote heart rate measurements using photoplethysmography *Multimedia Tools Appl.* **83** 44699–728

Poh M Z, McDuff D J and Picard R W 2010 Non-contact, automated cardiac pulse measurements using video imaging and blind source separation *Opt. Express* **18** 10762

Rouast P V, Adam M T P, Chiong R, Cornforth D and Lux E 2018 Remote heart rate measurement using low-cost rgb face video: a technical literature review *Frontiers of Computer Science* **12** 858–72

Sidorov I S, Volynsky M A and Kamshilin A A 2016 Influence of polarization filtration on the information readout from pulsating blood vessels *Biomedical Optics Express* **7** 2469

Slapničar G, Mlakar N and Luštrek M 2019 Blood pressure estimation from photoplethysmography using a spectro-temporal deep neural network *Sensors* **19** 3420

Sun Y and Thakor N 2016 Photoplethysmography revisited: From contact to noncontact, from point to imaging *IEEE Trans. Biomed. Eng.* **63** 463–77

Swinehart D F 1962 The beer-lambert law *J. Chem. Educ.* **39** 333

Tarassenko L, Villarroel M, Guazzi A, Jorge J, Clifton D A and Pugh C 2014 Non-contact video-based vital sign monitoring using ambient light and auto-regressive models *Physiol. Meas.* **35** 807–31

Tran T, Ma D and Balan R 2024 Remote multi-person heart rate monitoring with smart speakers: overcoming separation constraint *Sensors* **24**

Trumpp A, Bauer P L, Rasche S, Malberg H and Zaunseder S 2017 The value of polarization in camera-based photoplethysmography *Biomedical Optics Express* **8** 2822

Trumpp A, Schell J, Malberg H and Zaunseder S 2016 Vasomotor assessment by camera-based photoplethysmography *Current Directions in Biomedical Engineering* **2** 199–202

van Gastel M, Stuijk S and de Haan G 2015 Motion robust remote-ppg in infrared *IEEE Trans. Biomed. Eng.* **62** 1425–33

Verkruysse W, Svaasand L O and Nelson J S 2008a Remote plethysmographic imaging using ambient light *Opt. Express* **16** 21434–45

Verkruysse W, Svaasand L O and Nelson J S 2008b Remote plethysmographic imaging using ambient light *Opt. Express* **16** 21434

Virani S S *et al* 2021 Heart disease and stroke statistics—2021 update: A report from the american heart association *Circulation* **143** e254–743

Wang W, den Brinker A C, Stuijk S and de Haan G 2017 Algorithmic principles of remote ppg *IEEE Trans. Biomed. Eng.* **64** 1479–91
Weijer J and Beigpour S 2011 The dichromatic reflection model - future research directions and applications 011
Wu B F, Huang P W, Tsou T Y, Lin T M and Chung M L 2017 Camera-based heart rate measurement using continuous wavelet transform *2017 International Conference on System Science and Engineering (ICSSE)* (IEEE) 711
Yongyong D, Xinhua H, yujie Y and Zongling W 2020 Image stabilization algorithm based on klt motion tracking *2020 International Conference on Computer Vision, Image and Deep Learning (CVIDL)* (IEEE) 4447
Zaunseder S, Trumpp A, Wedekind D and Malberg H 2018 Cardiovascular assessment by imaging photoplethysmography - a review *Biomedical Engineering / Biomedizinische Technik* **63** 617–34