

# Novel Algorithms to Monitor Continuous Cardiac Activity with a Video Camera

Gregory F. Lewis  
Indiana University  
lewigr@iu.edu

Maria I. Davila  
University of North Carolina  
maria\_davila@med.unc

Stephen W. Porges  
Indiana University  
sporges@indiana.edu

## Abstract

*Recent advances in computer vision methods have made physiological signal extraction from imaging sensors feasible. There is a demand to translate current post-hoc methods into real-time physiological monitoring techniques. Algorithms that function on a single frame of data meet the requirements for continuous, real-time measurement. If these algorithms are computationally efficient they may serve as the basis for an embedded system design that can be integrated within the vision hardware, turning the camera into a physiological monitor. Compelling results are presented derived from an appropriate algorithm for extracting cardiac pulse from sequential, single frames of a color video camera. Results are discussed with respect to physiologically relevant features of variability in beat-to-beat heart rate.*

## 1. Introduction

Encoded in physiological signals are neural indicators of health status and regulation of visceral organs. Simple applications of cardiac monitoring, such as vital sign tracking, can be achieved with only an estimate of heart rate (HR). Neurophysiological systems that regulate the heart create defined patterns in beat-to-beat heart rate that reflect dynamic control in response to physical and psychological processes. These dynamic patterns of variations in the time between sequential heartbeats may be quantified with heart rate variability (HRV) metrics. Systems and tools for quantifying HRV during dynamic psychological and health-related processes, such as attention, fatigue, and deception are of use to clinicians, researchers, commercial products, and security operations.

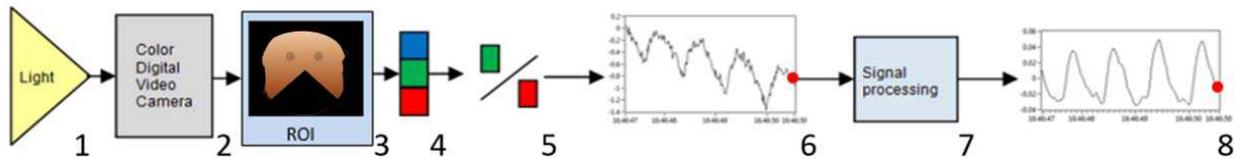
Due to a need for specialized bio-amplification equipment, assessment of neurophysiological regulation of the heart, via HRV metrics, has been limited generally to laboratory and clinical settings. To extract the precise timing of interbeat intervals (IBI), an accurate digital representation of the electrocardiogram (ECG) is considered best practice [2, 10]. Recent advances in Micro-Electro-Mechanical Systems (MEMS) have led to an

explosive growth in inexpensive, portable sensor technologies that have provided new opportunities to monitor neural regulation of the heart, via HRV parameters, outside the laboratory and clinic. However, these technologies are often constrained by an inability to dynamically monitor beat-to-beat heart rate with the requisite measurement accuracy to quantify HRV metrics.

Computer vision methods for extracting cardiovascular signals are driving the next wave in physiological assessment tools. Extraction of cardiac pulse from color video data has advanced from blind-source separation [11] of the Red-Green-Blue components of a webcam video to include Eulerian [18] magnification of rhythmic components of video that include pulse related color, motion tolerance estimation of published techniques [7, 8], and implementation of spatial redundancy to improve the motion robustness [17]. Despite the success of these research tools, the most promising applications of this technology will require embedded systems that can generate near real-time physiological parameter estimates. Such systems require computational algorithms that are sufficiently fast and efficient to process digital images at sampling frequencies capable of capturing transient physiological signals.

When HRV parameters are derived from the ECG, optimal output requires a robust cardiac signal and a sampling rate of at least 250 Hz [2]. The interaction between the rapid ventricular depolarization event and vascular transmission results in a smoothed pulse wave without high-frequency fiducial marks, yet precision in tracking beat to beat interval changes is still required.

In this manuscript we describe a method that builds on the system presented by Davila, et al. [5] to monitor beat-to-beat heart rate using a CCD color video camera as a photosensor, the PhysC. We present a new Physiological Camera for real-time (PhysCRT) applications with capabilities to continuously monitor the arterial pulse with sufficient precision to derive HRV metrics. Data are presented contrasting the analysis strategy described in [5] with a new analysis strategy that can be implemented in real-time. The PhysCRT embodies several enhancements of the PhysC that were empirically determined to improve motion tolerance, reduce artifacts due to speaking, and optimize estimation of HRV parameters.



**Figure 1:** Process to obtain arterial pulse from a video source, the PhysC [4]. Steps to transform one frame of live video into one sample of the cardiac pulse wave.

As shown in Figure 1, the PhysC methodology uses the ambient light (1) that is captured by the camera (2) in front of the participant's face to apply a combination of a geometric and a skin color mask ROI (3). The methodology then separates the image in the RGB color plane (4) to utilize the ratio of mean Green to Red level to obtain a value of arterial pulse per frame (6). Additional signal processing (7) filters the signal to produce a clear representation of the arterial pulse wave (8).

In addition, future innovations of this technology are outlined. Critically, the algorithms and analytic strategies described here are suitable for deployment to an embedded camera system that functions without delay or video archiving. In this study, no video was recorded, all processing of pixel level data occurred in real-time, and extracted parameters were analyzed to create these results.

The PhysC<sub>RT</sub> combines an external head tracker (Microsoft Kinect) and a CCD digital color video camera as shown in Figure 2. The participant's face tracking is done by the Kinect device and the capturing of the physiological information (arterial pulse) is done by the color camera. The use of two devices improves the face tracking and allows for faster sampling frequencies. The tracking feature is a stand-in for a secondary processing pipeline in an embedded system that would locate faces for processing.



**Figure 2:** The Physiological Camera in real-time system (PhysC<sub>RT</sub>): 1) Digital Color Video Camera, 2) Microsoft Kinect.

## 2. Material and Methods

When HRV is deconstructed through statistical procedures, it is possible to quantify rhythmic components that reflect specific pathways of neural regulation. The most salient components are: 1) a respiratory oscillation known as respiratory sinus arrhythmia (RSA) assumed to reflect

cardiac vagal tone via myelinated pathways originating in a brainstem area known as nucleus ambiguus [6] and 2) a slower frequency (LFHRV) assumed to be related to blood pressure regulation via the baroreceptors and peripheral vasomotor activity [13, 16]. Heart period (HP) the time between heart beats, was also monitored, since it reflects the sum of neural, neurochemical, and intrinsic influences on the heart. To validate the variables derived by the PhysC<sub>RT</sub>, an experiment was conducted to identify and describe the similarities and differences among the heart rate indices derived from signals monitored by the PhysC<sub>RT</sub>, an earlobe photoplethysmogram (ePPG), and an ECG. The camera data were reduced to a single dimension time-series in real-time, making direct comparison to *post hoc* video analysis methods impossible. Video storage would have reduced the achievable frame rate of the current system below tolerances necessary to achieve sufficient precision in the timing of heart rate in the real-time processing pipeline.

### 2.1. Experiment Design

Participants were seated in front of a computer, 3 m away from the camera lens. Criterion signals were collected with contact sensors, ECG and ePPG. The protocol consisted of synchronously collecting pulse wave activity and ECG data during a ten minute period including two minutes for task preparation, three minutes of baseline, two minutes of watching a video chosen to elicit focused attention, and three minutes of answering questions related to the video. Of interest for this manuscript are the three-minute baseline and the two-minute video.

### 2.2. Participants

The study was approved by the Institutional Review Board of the University of North Carolina at Chapel Hill (UNC) as protocol # 14-1560 with the title "PhysioCam: A noncontact system to monitor heart rate.". The IRB authorized the recruitment of 30 subjects from the student subject pool and volunteers. Twenty-seven participants between the ages of 18 and 33 ( $M = 20.11$ ,  $SD = 2.98$ ) were recruited. The gender mix was, 56% female and 44% male. Participants self-identified as White or Caucasian (62.96%), Asian (18.52%), Black or African American (3.70%), American Indian or Alaskan Native (0%),

Hispanic or Latino (3.70%), Asian and Hispanic or Latino (3.70%), Asian and White or Caucasian (3.70%), and White or Caucasian and Hispanic or Latino (3.70%).

### 2.3. Hardware and Software

The PhysC<sub>RT</sub> used a Pt.Grey Grasshopper video camera (model GRAS-03K2C-C) with Fujinon manual zoom optics and MS Kinect for Windows Academic. The face tracking was implemented with a modified version of the Kinect SDK v1.8 program “Face Tracking Basics-WPF”. The modification made the tracking results available to Labview via a UDP port connection. Algorithms to extract the arterial pulse signal and derive the HRV parameters were implemented on NI Labview 2014 (64-bit).

The criterion contact signals, ECG and ePPG, were collected using the BIOPAC MP150 and the AcqKnowledge 4.1 software.

### 3. Data Quantification

Since the HRV parameters of interest are derived from the IBI obtained from the ECG or the pulse waves, the data was quantified in four phases.

#### 3.1. Phase 1: Raw PhysC<sub>RT</sub> Signal Processing

The arterial pulse from the face was captured by the PhysC<sub>RT</sub> system [15]. The system was controlled in stages that included manual calibration, as a stand-in for future automated initialization processes, and a continuous collection stage that is our model for embedded device operation. The pipeline critically maintained the frame-level analysis procedures that were selected to optimize signal stability during natural head movement, including during talking.

**Camera-Kinect registration:** this stage is done when the system is initialized to collect data from individual participants. The Kinect for Windows Software Development Kit (SDK) offers the Microsoft Face Tracking Software Development Kit (Face Tracking SDK). The SDK and the Face Tracking SDK, in conjunction with the Kinect device, facilitate the tracking of the subject’s face in real-time. The Face Tracking SDK uses the CANDIDE-3 [1] model to parameterize the human face into output information data. X and Y coordinates of several vertices of the parameterized face are available in an array of 121 set of points; each point corresponds to an X and Y pixel position on the Kinect image output. For example: Point #6 will be the (X,Y) pair that relates to the pixel position of the bottom middle edge of nose; point #3 will be the (X,Y) pair that relates to the pixel position of the midpoint between eyebrows. Points from the CANDIDE-3 array are translated to the color camera frame dimensions

by a “camera registration” procedure; correcting for differences in the field of view (FOV) between the Kinect device and the color camera. These points were used to track the subject’s face in real-time in the color camera space. This spatial information feeds the coordinates to create the “Geometric Mask”, which is the region of interest (ROI). The ROI representing features of the participant’s face is formed by a semi-annulus centered on the participant’s nose covering most of the face and the subtraction of an oval shape where the mouth is. Using the Kinect sensor for motion tracking facilitates processes associated with: 1) identifying the subject’s face rapidly on initialization, 2) quickly and automatically recalibrating following movements that disrupt detection of the face, 3) adjusting to variations in the distance between the participant and the camera, and 4) improving the real-time PhysC<sub>RT</sub> frame rate. Overall incorporating the Kinect face tracking provided an efficient tracking solution that reduced computational load on the PhysC<sub>RT</sub> processor and allowed an operational frame rate of 50 Hz with only minimal restriction of subject movement across all tasks. Under these conditions, natural body sway and head nodding resulted in the Geometric Mask tracking the participant’s face as it dynamically moved across the field-of-view. The PhysC<sub>RT</sub> system reported here can be used in near field (2m or less) or far field (3-10m) settings.

**Customizing camera attributes:** before starting pulse acquisition the parameters of gain, gamma, saturation, white balance, and shutter were manually adjusted for each participant to optimize the distribution of light captured by the sensor over the selected integration time. Parameters were adjusted to maximize the distribution of Red and Green pixel levels of skin across the 8-bit range of the sensor. Empirical testing indicated that this increased the motion tolerance of the system. After setting the camera parameters, values for each frame were constrained in the Hue-Saturation-Luminance color plane to correspond to the participant’s particular skin color, creating a “Skin Mask” that assured that the pixels being analyzed were highly likely to only include skin regions. This step maximizes the information content of the subsequent images processed; minimizes the computational load, allowing faster frame rates; and assures prioritization of the physiological information of interest.

**PhysC<sub>RT</sub> functioning and pulse extraction:** the image of the participant is captured at a 50 Hz sampling frequency; each frame is processed independently to continuously generate samples to the 1D arterial pulse output signal. Figure 3 depicts the processes that are applied to each frame: Ambient light (1) is reflected by the participant’s face and captured by the camera (2), the ROI (3) is generated by combining a geometric mask obtained by the CANDIDE 3 points and the skin mask obtained by the skin

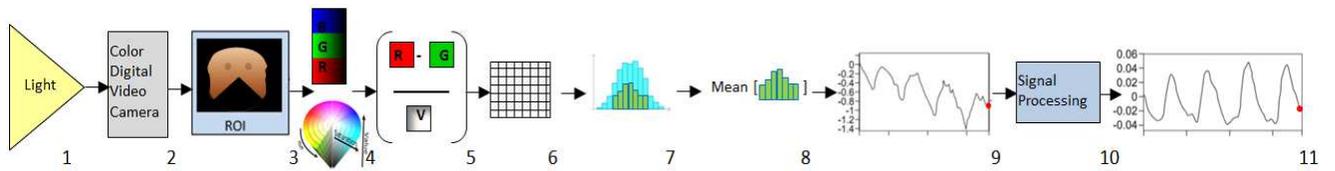


Figure 3: The Physiological Camera (PhysC<sub>RT</sub>) functioning. Steps to transform one frame of live video into one sample of the cardiac pulse wave time-series.

color optimization, the hardware encoded Bayer map decomposition is used to estimate RGB values per pixel from which the (4) HSV plane is calculated in software, pixel level values of R, G, and V are used to compute eq. (1) the Preprocessed Pixel Value (PPV) (5):

$$PPV = 5 + \log_{10} \left( 2 + \frac{|\text{Red}-\text{Green}|}{\text{Value}} \right) \quad (1)$$

the resulting matrix (6) is down-sampled to  $\frac{1}{4}$  resolution then flattened, sorted and binned with a histogram (7), the trimmed mean PPV value, user adjustable but default set to central 40% of initial frame PPV values, is calculated (8) to generate a single sample of the raw cardiac pulse waveform (9). The raw signal (10) is continuously filtered to eliminate very low frequencies not related to the pulse wave and transformed into the discreet difference between samples to stabilize the detection of pulse wave arrival times (11). This process yields a final arterial pulse wave at 50 Hz. Steps 4-6 are computationally inefficient due to the onboard Bayer map decomposition architecture, but could be eliminated in an optimized embedded system design. Down-sampling of the image is used to reduce calculation time in the per-pixel quantification of PPV, in order to approximate calculations on discreet Red, Green or Blue pixels on a planned sensor that does not include Bayer map decoding.

Compared to the PhysC, the PPV algorithm and subsequent exclusion of PPV outliers, were empirically determined to increase the motion tolerance due to natural body sway, head nodding, and talking. The arbitrary offsets of 5 and 2 were selected for convenience. Consistent with the PhysC, the pulse signal was derived from the local difference between red and green light absorption on the skin. Weighting these differences by the local illumination (Value) reduced the influence of outlier pixels near saturation level. A log transformation normalized the PPV distributions for most subjects in most illumination settings.

### 3.2. Phase 2: Raw Criterion Signal Processing

**Earlobe Photoplethysmogram (ePPG):** the earlobe pulse was monitored with ePPG. The pulse wave sampled at 1 kHz through a DC amplifier with no filter setting to preserve the trend of the vasomotor activity that reflects sympathetic influences.

**Electrocardiogram (ECG):** ECG was sampled at a 1 kHz using the preferred ECG BIOPAC MP-150 settings.

### 3.3. Phase 3: Dynamic Heart Rate Measures.

**Interbeat Interval (IBI):** IBI is the time between consecutive heart beats expressed in milliseconds. IBIs are calculated by the consecutive differences of the time component of the ECG R-wave peaks or of the arterial pulse peaks. The R-wave and pulse peak coordinates were extracted using the Cardio Peak-Valley Detector (CPVD) [5]. The result is the IBI event series.

**5 Hz instantaneous IBI (instIBI):** An innovative strategy was implemented to obtain heart rate (HR) measures from the pulse signals (PhysC<sub>RT</sub> and ePPG); this approach will facilitate the transition to a real-time system. A time-frequency method was applied to monitor the energy distribution of the pulse wave to determine the confidence that could be assigned to each instantaneous heart rate measure. This reduces the delay in detecting aberrant pulse wave intervals that could contaminate the HRV parameters in a real-time estimation procedure, while eliminating the time-sampling procedure required for a beat detection method. Figure 4 outlines the proposed algorithm: (a) 5-second portion of the pre-processed pulse signal obtained from the PhysC<sub>RT</sub> or the ePPG; (b) a Blackman-Harris scaled window is applied to the portion of data for the autoregressive (AR) spectral analysis model to facilitate the extraction of the predominant frequency; (c) the AR spectrum identifies the instantaneous HR, an estimated 'instIBI' (i.e., the inverse of the instantaneous HR). The 5-second window moves 200 ms to cover the following portion of 5-seconds data and obtain the next HR estimate, each portion of the data overlaps by 4,800 ms. The result is the 5 Hz instIBI signal.

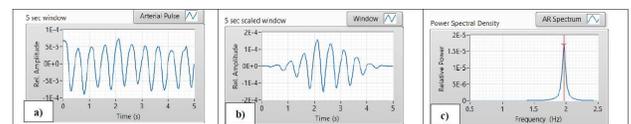


Figure 4: Steps to obtain the instantaneous IBI (instIBI): a) 5 second window of the arterial pulse; b) a window function; c) predominant frequency is obtained by an AR-Spectrum function, this frequency corresponds to heart rate.

An editing step was added to prevent spurious instIBI values, usually due to participant's movement or spontaneous arrhythmias. A Signal to Noise Ratio (SNR) is obtained from the AR-Spectrum. Applying the SNR threshold helps to discriminate between good (Pass) instIBI or bad (Fail) instIBI; bad instIBI are replaced by a linear spline, producing the final edited 5 Hz instIBI pulse signal. The SNR was set as 1 for instIBI values less or equal to 950 ms and 2 for instIBI greater than 950 ms. These values were the result of previous observations; future work will include an auto set SNR informed by the instIBI dynamics.

### 3.4. Phase 4: Quantification of HRV Parameters.

**Time Sampled Mean IBIs from 2 and 5 s Windows (IBI 2sW and IBI 5sW):** The IBI event series was resampled at 2 Hz to generate an equally spaced intervals time series. Sequential IBIs were averaged within 2- and 5-s windows to create smoothed IBI trends.

**Time Sampled Mean 5 Hz instantaneous IBI from 1, 2, and 5 s Windows (instIBI 1sW, instIBI 2sW, and instIBI 5sW):** Sequential instIBIs were group and average within 1-, 2- and 5-s windows to create smoothed instIBI trends.

**Heart Period (HP):** is the average value of the 2 Hz IBI or the 5 Hz instIBI time series within a specific task.

**Respiratory Sinus Arrhythmia (RSA):** Calculations were based on the Porges-Boher method [12, 14], applying a third-order moving polynomial filter (MPF) with a duration of approximately 10.5 seconds on the two HR time series (2 Hz IBI and 5 Hz instIBI) to remove low frequency oscillations and slow trend. The residual detrended output of the MPF is filtered with a Kaiser FIR windowed filter with cutoff frequencies that remove variance not related to spontaneous breathing in adults (0.12–0.40 Hz). The filtered detrended output is divided into sequential 30-s epochs and the variance within each epoch is transformed by a natural logarithm [ $\ln(\text{ms}^2)$ ], the mean of these epoch values is used as the estimate of RSA.

**Low Frequency Heart Rate Variability (LFHRV):** Calculations were made for both HR time series, based on the Porges-Boher method [12, 14] that uses a moving polynomial window of approximately 25 seconds to remove extremely low frequencies, and a Kaiser FIR windowed filter with cutoff frequencies between 0.04–0.10 Hz. The output was divided into sequential 30-s epochs and the variance within each epoch is transformed by a natural logarithm [ $\ln(\text{ms}^2)$ ], the mean of these epoch values is used as the estimate of LFHRV for each task.

### 3.5. Data Analysis

Statistical analyses were approached on three levels: 1) Replicate results previously reported [5] by comparing and contrasting traditional offline analyses based on IBI values derived from the PhysC<sub>RT</sub> and the ECG and ePPG criterion signals; 2) Compare and contrast the new instIBI analysis between the PhysC and the ePPG to evaluate feasibility of a real-time system; 3) Compare HRV parameters obtained by IBI and instIBI for the different sensor combinations. Statistical analyses were processed using IBI SPSS Statistics for Windows, Version 24.0. Armonk, NY, USA: IBI Corp.

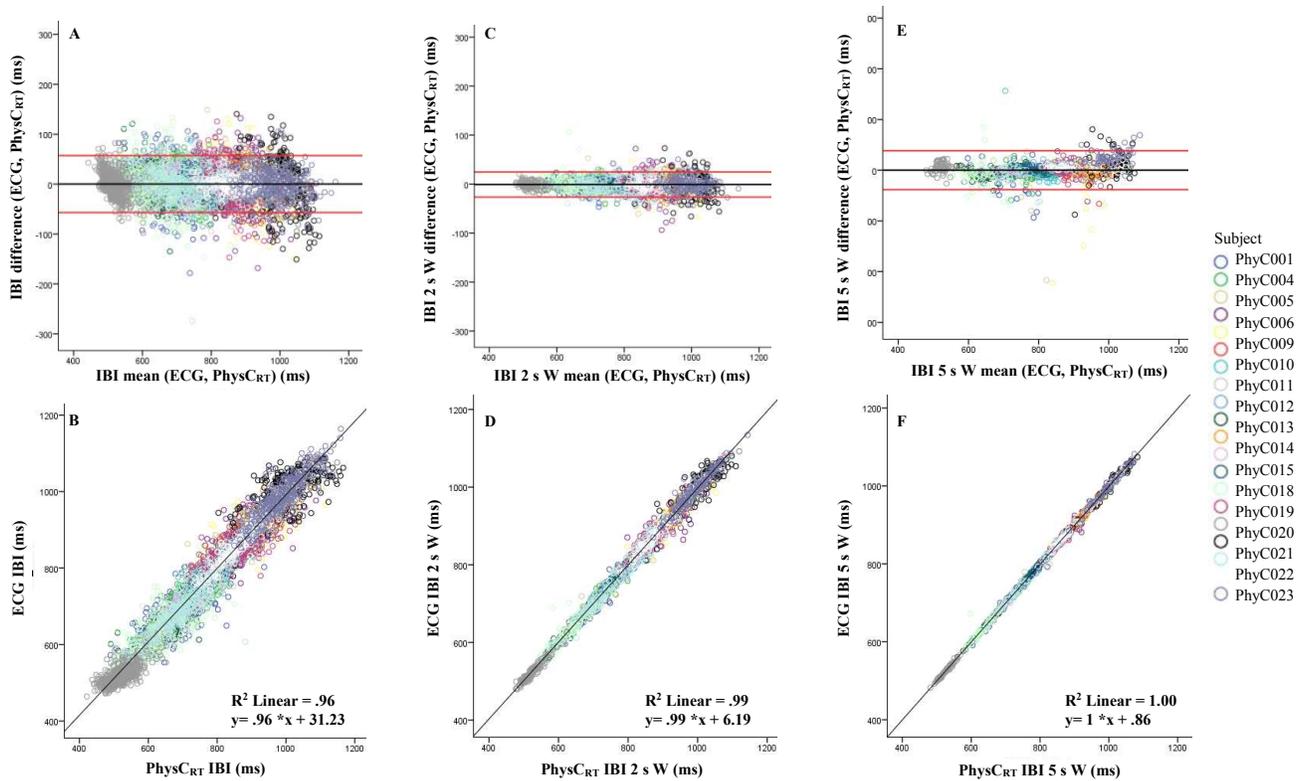
**Bland-Altman (B-A) Plots:** B-A plots [3] were used to compare IBI from the PhysC<sub>RT</sub> with IBI from the criterion signals (ECG and ePPG) and also to compare values generated by the ePPG and ECG signals. Heart rate parameters obtained with the new instIBI approach were evaluated using B-A plots between measures of the PhysC<sub>RT</sub> and the ePPG. B–A plots enable the determination of agreement between two sensors, by plotting the mean between pairs of measurements against its difference. Visual inspection of the B–A plots was used to identify systematic biases and possible outliers.

**Scatter Plots and Linear Regressions:** Scatter plots and linear regression analyses were used to visualize and calculate the level of convergence between the physiological measures derived from the different sensors and from the two approaches to process heart rate data, the traditional IBI and the proposed instIBI. For the IBI, PhysC<sub>RT</sub> was evaluated with each of the criterion signals (ECG and ePPG), and the two criterion signals between each other. For the instIBI, the PhysC<sub>RT</sub> was evaluated with the ePPG. Derived measures of HRV from the two approaches were evaluated between sensors combinations (ECG-PhysC<sub>RT</sub>, ePPG-PhysC<sub>RT</sub>, and ECG-ePPG).

## 4. Results

### 4.1 IBI

Visual inspection of the Bland-Altman (B-A) plot located in panel A of Figure 5 indicate excellent agreement and minimal bias between the sequential IBIs measured with ECG and PhysC<sub>RT</sub> (color coded by participant), panels C and E show B-A plots of the 2 s and 5 s IBI windowing, as the window widens the agreement between the two sensors increases, the differences between measures approaches zero and the 95 % confidence interval decreases. Panels B, D, and F show the linear regression for IBI, 2 s and 5 s IBI windowing,  $R^2$  increases to show complete agreement as the window widens.



**Figure 5:** B-A and scatter plots for IBI, IBI 2 s W, and IBI 5 s W from the ECG and PhysC<sub>RT</sub>, color coded by participant. (A and B) B-A and scatter plots for IBI, (C and D) B-A and scatter plots for IBI 2 s W, and (E and F) B-A and scatter plots for IBI 5 s W.

		Differences			LR	
		IBI	N	Mean	SD	R <sup>2</sup>
PhysC <sub>RT</sub> vs ECG	IBI	8088	0.09	29.13	.96	
	2 s W	3165	-0.90	13.07	.99	
	5 s W	1266	-0.92	5.90	1.00	
	HP	40	-0.88	0.80	1.00	
PhysC <sub>RT</sub> vs ePPG	IBI	8088	0.10	29.16	.96	
	2 s W	3165	-0.96	12.29	.99	
	5 s W	1266	-0.98	5.78	1.00	
	HP	40	-0.93	0.79	1.00	
ePPG vs ECG	IBI	8088	0.00	6.19	1.00	
	2 s W	3165	0.06	5.64	1.00	
	5 s W	1266	0.06	2.30	1.00	
	HP	40	0.05	0.23	1.00	

**Table 1:** Difference means and standard deviations for all subjects. All tasks between the ECG and PhysC<sub>RT</sub>, ePPG and PhysC<sub>RT</sub>, and ECG and ePPG for IBI, IBI 2 s windows (W), IBI 5 s windows (W), and HP 30 s epoch. R<sup>2</sup> from linear regressions between PhysC<sub>RT</sub> and ECG, PhysC<sub>RT</sub> and ePPG, and ePPG and ECG.

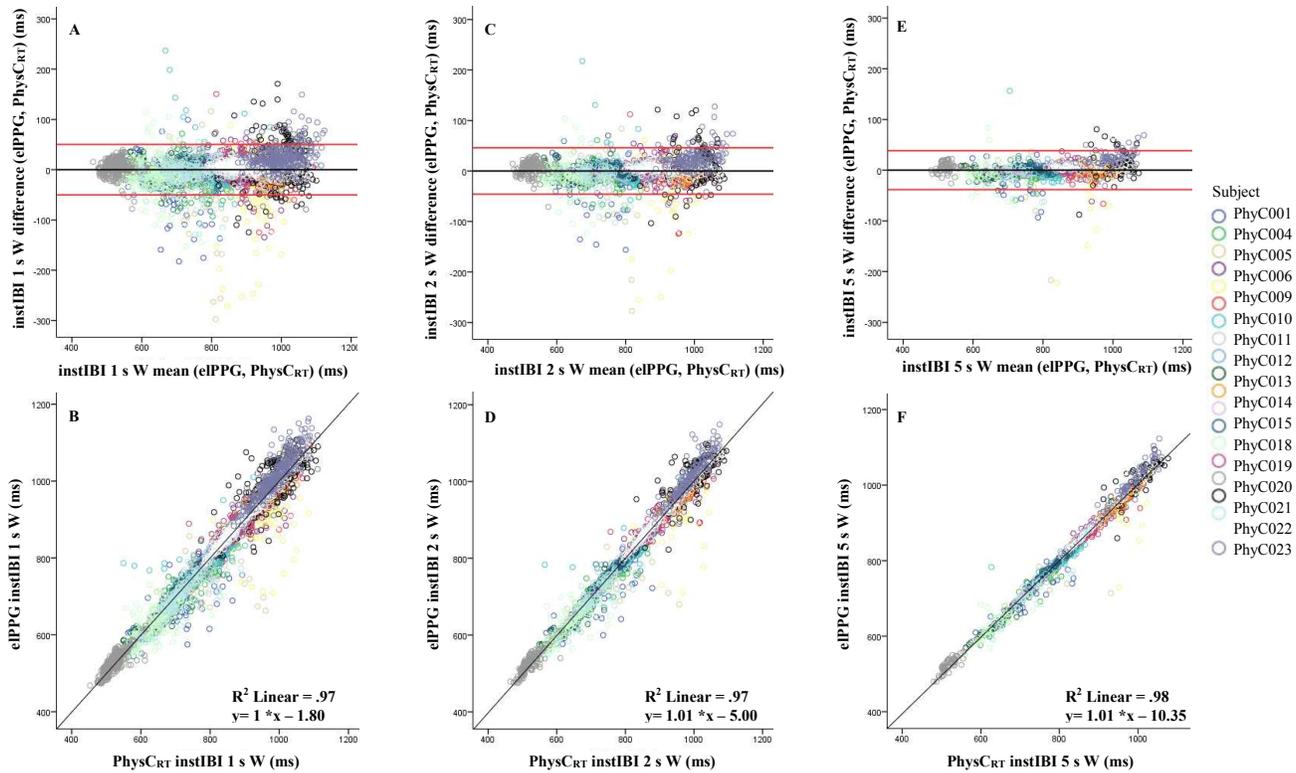
Table 1 reaffirms this observation, as the window to average IBI measures widens from individual IBIs to 2 s and 5 s windows, the standard deviation of the differences between the PhysC<sub>RT</sub> and the ECG decreases and the R<sup>2</sup> increases, eventually reaching unity.

Note that the means of the differences are less than 1 ms and that there are no significant differences between the metrics in central tendency.

The PhysC<sub>RT</sub> performs acceptably when compared with traditional contact sensors, ECG and ePPG, that collect the data at 1kHz under the same real-time conditions.

#### 4.2 5 Hz Instantaneous IBI

Visual inspection of the Bland-Altman (B-A) plots located in panel A of Figure 6 indicate excellent agreement and minimal bias between the sequential instIBIs from 1-second windows of the ePPG and PhysC<sub>RT</sub> (color coded by participant), panels C and E show B-A plots of the 2 s and 5s IBI windowing, as the window widens the agreement between the two sensors increases, the differences between measures and the 95% confidence interval decrease; panels B, D, and F show the linear regressions indicating agreement between the sensors which again approaches unity as the window widens from 1 s to 2 s and to 5 s. Table 2 reaffirms this observation, as the window to average instIBI measures widens from 1 s to 2 s and 5 s windows, the standard deviation of the differences between



**Figure 6:** B-A and scatter plots for instIBI from the ePPG and PhysCRT, color coded by participant. (A and B) B-A and scatter plots for instIBI 1 s W, (C and D) B-A and scatter plots for instIBI 2 s W, and (E and F) B-A and scatter plots for instIBI 5 s W.

the PhysCRT and the ePPG decreases and the R<sup>2</sup> increases reaching unity in the case of the HP average on 30 s epoch.

		Differences			LR
		N	Mean	SD	
PhysCRT vs ePPG	instIBI 1sW	6308	-0.05	25.62	.97
	instIBI 2sW	3153	-0.05	23.52	.97
	instIBI 5sW	1261	-0.02	19.59	.98
	HP	40	-0.22	10.24	1.00

**Table 2:** Difference means and standard deviations for all subjects. All tasks between the ePPG and PhysCRT for instIBI 1 s window (W), instIBI 2 s windows (W), instIBI 5 s windows (W), and HP 30 s epoch. R<sup>2</sup> from linear regression between PhysCRT and ePPG.

Note that the means of the differences are less than 0.5 ms and that there are no significant differences between the metrics in central tendency.

### 4.3 Quantification Methods Comparison: IBI versus 5 Hz instIBI

Table 3 shows R<sup>2</sup> values between the different sensors for the HRV parameters of HP, RSA, and LFHRV obtained by the traditional IBI method and the new instIBI method.

Sensors	HRV Parameter	R <sup>2</sup>	
		IBI	5 Hz instIBI
ECG- PhysCRT	HP	1.00	1.00
	RSA	.85	.74
	LFHRV	.97	.91
ePPG- PhysCRT	HP	1.00	1.00
	RSA	.86	.74
	LFHRV	.96	.84
ECG-ePPG	HP	1.00	1.00
	RSA	.99	.93
	LFHRV	.99	.90

**Table 3:** HRV parameters linear regression (R<sup>2</sup>) between sensors for measures derived from IBI and 5 Hz instIBI.

Measures obtained with ECG and PhysCRT of HP, RSA and LFHRV were highly correlated [9]. Similar strong correlations were observed when comparing the HP, RSA and LFHRV monitored with ePPG and the PhysCRT or

with ECG and ePPG. Considering that the PhysC<sub>RT</sub> is a non-contact sensor performing in this embodiment that is currently camera limited to 50 Hz, the results are compelling, since data from the contact sensors, ECG and ePPG, were sampled at 1kHz.

## 5. Discussion and Future Directions

This paper has described an innovative algorithm for extraction of HRV parameters from color video data that are computationally appropriate for applications in an embedded system design. The system, as illustrated in this paper, transforms a camera into a photosensor for physiological monitoring. Although video data are not stored, the PhysC<sub>RT</sub> has the capabilities to continuously monitor pulse wave activity and to extract three key metrics of dynamic heart rate patterns: HP, RSA, and LF. In addition, the system incorporates an artifact rejection feature that imposes minimal lag (5 seconds) on the output of HRV components.

The procedures embedded in this system provide a continuous estimation of interbeat interval, the instIBI, which is more appropriate for dynamic tracking of HRV in a noncontact photosensor system, due to its robustness against transient noise or signal loss. The dependence of traditional HRV analysis on peak detection and artifact correction imposes significant constraints on physiological monitoring. The comparable sensitivity of instIBI to IBI (post-hoc) based evaluations provides validation for the method. Furthermore, the results are similar with respect to contact HRV correlation as obtained by Davila, et al. [5] in *post hoc* analysis.

The results of this study validate the current approach to data reduction through simple arithmetic operations conducted on the pixel readout level of the sensor. The excellent convergence between contact and noncontact HRV parameters suggests that an acceptable noise level can be achieved with this approach. Unfortunately, the real-time extraction of the PPV values, combined with project constraints, precluded direct comparison of this method with established *post hoc* methods (e.g., of [11] or [17]).

Applications that demand real-time monitoring of physiological changes through noncontact sensing will be possible through integration of this algorithm into a camera system. Currently, several parameters of the color sensor and optical system are manually optimized prior to measurement. Motion tracking is also handled by an external system (Kinect). An optimal PhysC<sub>RT</sub> system would integrate these operations within a single processing structure. The power requirements and computational demand of these operations must be evaluated through further testing.

This study demonstrates the feasibility of our approach to a continuous, real-time physiological monitor based on analysis of sequential, single frames. We continue to refine

algorithms involved in instIBI estimation, integration of the instIBI parameter into a peak detection system, and extraction of noise features from the short-term cardiac pulse signal to inform editing decisions.

We continue to innovate to optimize function of the PhysC<sub>RT</sub> system. In order to enhance the performance of future versions of the PhysC<sub>RT</sub> platform we are currently conducting research to modify the wavelength segmentation of the color filter array to prioritize physiological signal extraction over visual fidelity, minimizing the impact of melanin on sensitivity of the sensor, and integrating multiple overlapping sensor FOVs to enhance subject identification and tracking.

Although future comparison studies may reveal precision advantages to previously disclosed methods, such as Blind Source Separation, we believe that the tradeoff for real-time estimation of HRV is worthwhile in many applications. Our research approach will deploy computer vision based physiological measurement methods in health screening, critical care, security, psychological assessment, basic and applied research, and training settings. Faster frame rates, improved artifact rejection, and robust signal processing algorithms that optimize physiological features of interest without compromising stability will be possible in an embedded system design that integrates increased processing power with intelligent design decisions.

## Acknowledgment:

This research was funded by the Defense Advanced Research Projects Agency (DARPA) grant W911NF-14-1-0158.

## References:

- [1] J. Ahlberg, "Candide-3-an updated parameterized face," Technical report, Sweden, 2001.
- [2] G. Berntson, J. Thomas Bigger, J. Eckberg, P. Grossman, P. Kaufmann, M. Malik, H. Nagaraja, S. Porges, J. Saul, P. Stone and M. Der Molen, "Heart rate variability: origins, methods, and interpretive caveats," *Psychophysiology*, vol. 34, no. 6, pp. 623-648, 1997.
- [3] J. Bland and D. Altman, "Statistical methods for assessing agreement between two methods of clinical measurement," *The Lancet*, vol. 372, no. 8476, pp. 307-310, 1986
- [4] M. Davila, G. Lewis and W. Porges, "The physiocam: cardiac pulse, continuously monitored by a color video camera," *Journal of Medical Devices*, vol. 10, no. 2, p. 020951, 2016.
- [5] M. Davila, G. Lewis and S. Porges, "The Physiocam: a novel non-contact sensor to Measure heart rate Variability in clinical and Field applications,"

- Frontiers in Public Health, vol. 5, no. 300, p. <http://doi.org/10.3389/fpubh.2017.00300>, 2017.
- [6] J. Denver, S. Reed and S. Porges, "Methodological issues in the quantification of respiratory sinus arrhythmia," *Biological Psychology*, vol. 74, no. 2, pp. 286-294, 2007.
- [7] J. Estepp, J. Blackford and C. Meier, "Recovering pulse rate during motion artifact with a multi-imager array for non-contact imaging photoplethysmography," in *In Systems, Man and Cybernetics (SMC), 2014 IEEE International Conference*, Sydney, Australia, 2014.
- [8] X. Li, J. Chen, G. Zhao and M. Pietikainen, "Remote heart rate measurement from face videos under realistic situations," in *In Proceedings of the IEEE conference on computer vision and pattern recognition*, Columbus, 2014.
- [9] Mukaka, "A guide to appropriate use of correlation coefficient in medical research," *Malawi Medical Journal*, vol. 24, no. 3, pp. 69-71, 2012.
- [10] M. Malik, J. Thomas Bigger, A. Camm, R. Kleiger, A. Malliani, A. Moss and J. Schwartz, "Heart rate variability, standards of measurement, physiological interpretation, and clinical use," *European Heart Journal*, vol. 14, no. 1, pp. 354-381, 1996.
- [11] M. Poh, D. McDuff and R. Picard, "Non-contact, automated cardiac pulse measurements using video imaging and blind source separation," *Optics express*, vol. 18, no. 10, pp. 10762-10774, 2010.
- [12] S. Porges, "Method and Apparatus for Evaluating Rhythmic Oscillations in Aperiodic Physiological Response Systems". United States of America Patent 4,510,944, 16 April 1985.
- [13] S. Porges, "The polyvagal perspective" *Biological Psychology*, vol. 74, no. 2, pp. 116-143, 2007.
- [14] S. Porges and R. Bohrer, "Analyses of periodic processes in psychophysiological research," in *Principles of Psychophysiology: Physical, Social, and Inferential Elements*, New York, Cambridge University Press, 1990, pp. 708-753.
- [15] S. Porges, M. Davila and G. Lewis, "System and methods for measuring physiological parameters". United States of America Patent 15/105,674, 3 November 2016.
- [16] G. Reyes del Paso, W. Langewitz, L. Mulder, A. Roon and S. Duschek, "The utility of low frequency heart rate variability as an index of sympathetic cardiac tone: a review with emphasis on a reanalysis of previous studies," *Psychophysiology*, vol. 50, no. 5, pp. 477-487, 2013.
- [17] W. Wang, S. Stuijk, and G. De Haan, "Exploiting spatial redundancy of image sensor for motion robust rPPG". *IEEE Transactions on Biomedical Engineering*, 62(2), pp.415-425, 2015.
- [18] H. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand and W. Freeman, "Eulerian video," *ACM Transactions on Graphics*, vol. 31, no. 4, pp. 1-8, 2012.