Cardiac Inter Beat Interval and Atrial Fibrillation Detection using Video Plethysmography

Jairo Hernandez Guzman
jh1206@rit.edu

Follow this and additional works at: http://scholarworks.rit.edu/theses

Recommended Citation

This Thesis is brought to you for free and open access by the Thesis/Dissertation Collections at RIT Scholar Works. It has been accepted for inclusion in Theses by an authorized administrator of RIT Scholar Works. For more information, please contact ritscholarworks@rit.edu.
Cardiac Inter Beat Interval and Atrial Fibrillation Detection

using Video Plethysmography

by

Jairo Hernandez Guzman

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of
Master of Science in Electrical Engineering

Supervised by

Dr. Gill R. Tsouri
Department of Electrical and Microelectronic Engineering
Kate Gleason College of Engineering
Rochester Institute of Technology
Rochester, NY
June 2016

Approved By:

Dr. Gill R. Tsouri
Primary Advisor – R.I.T. Dept. of Computer Engineering.

Dr. Panos P. Markopoulos
Committee Member – R.I.T. Dept. of Electrical and Microelectronic Engineering.

Dr. Jean-Philippe Couderc
Committee Member – U.R.M.C. Heart Research Follow-up Program, Dept. of Cardiology.

Dr. Sohail A. Dianat
Department Head – R.I.T. Dept. of Electrical and Microelectronic Engineering.
Abstract

Facial videoplethysmography provides non-contact measurement of heart activity based on blood volume pulsations detected in facial tissue. Typically, the signal is extracted using a simple webcam followed by elaborated signal processing methods, and provides limited accuracy of time-domain characteristics. In this study, we explore the possibility of providing accurate time-domain pulse and inter-beat interval measurements using a high-quality image sensor camera and various signal processing approaches, and use these measurements to diagnose atrial fibrillation. We capture synchronized signals using a high-quality camera, a simple webcam, an earlobe photoplethysmography sensor, and a body-surface electrocardiogram from a large group of subjects, including subjects diagnosed with cardiac arrhythmias. All signals are processed using both blind source separation and color conversion. We then assess accuracy of IBI detection, heart rate variability estimation, and atrial fibrillation diagnose by comparing to a body-surface electrocardiogram. We present a new heart variability indicator for blood volume pulsating signals. Our results demonstrate that the accuracy of a facial VPG system is greatly improved when using a high-quality camera. Coupling the high-quality camera with color conversion from RGB to Hue provides a level of accuracy equivalent to that of commercially available photoplethysmography sensors, and offers a non-contact alternative to current technology for heart rate variability assessment and atrial fibrillation screening.
# Table of Contents

ABSTRACT ..........................................................................................................................II  

LIST OF FIGURES .............................................................................................................. V  

LIST OF TABLES ................................................................................................................... VII  

SUMMARY OF CONTRIBUTIONS ....................................................................................... VIII  

INTRODUCTION ................................................................................................................... 1  

CHAPTER 1. LITERATURE REVIEW ..................................................................................... 4  

1.1 Photoplethysmography ............................................................................................... 4  

1.2 Videoplethysmography ............................................................................................. 5  

1.3 Atrial fibrillation diagnosis ....................................................................................... 8  

CHAPTER 2. PROPOSED APPROACH ................................................................................. 11  

2.1 The sensing environment .......................................................................................... 11  

2.2 Signal processing for VPG ....................................................................................... 14  

2.3 Materials and apparatus .......................................................................................... 18  

2.4 Detection of Inter Beat Intervals ............................................................................. 25  

2.5 Detection of Atrial Fibrillation .................................................................................. 27  

CHAPTER 3. CLINICAL STUDY .......................................................................................... 30  

3.1 Subjects ...................................................................................................................... 31  

3.2 Data Gathering .......................................................................................................... 33  

CHAPTER 4. RESULTS AND DISCUSSION ....................................................................... 35  

4.1 IBI detection .............................................................................................................. 35  

4.2 Atrial fibrillation detection using VPG ...................................................................... 47
4.3 Image sensor comparison ........................................................................................................ 52

CHAPTER 5. CONCLUSIONS ........................................................................................................ 53

REFERENCES .................................................................................................................................... 54

ACKNOWLEDGEMENTS ................................................................................................................. 63
List of Figures

Fig. 1 – Electrical activity of a normal heart (left) and a heart with atrial
fibrillation (right) ................................................................. 1

Fig. 2 – ECG signal of atrial fibrillation subject .................................. 2

Fig. 3 – Transmission (left) and Reflection (right) mode PPG configurations .... 4

Fig. 4 – Sensing platform .................................................................. 12

Fig. 5 – Frequency response of bandpass filter ....................................... 17

Fig. 6 – Logitech Quickcam Pro 9000 .................................................. 18

Fig. 7 – On Semiconductor MT9D131 (Webcam) image sensor Quantum 
Efficiency Vs Wavelength ................................................................. 19

Fig. 8 – Basler Ace . ........................................................................ 20

Fig. 9 – Sony IMX174 Image sensor Quantum Efficiency Vs Wavelength. .... 21

Fig. 10 – Mortara H12+ ECG Holter recorder ........................................ 21

Fig. 11 – Binar Heart Sensor HRS-07UE ............................................... 23

Fig. 12 – Dell Latitude E6440 ............................................................ 24

Fig. 13 – Schematic description of the experimental setup of the study .......... 31

Fig. 15 – Bland-Altman plot of the IBIs estimated from Gree-Basler compared 
to the IBIs estimated from the ECG Holter. .......................................... 36

Fig. 14 – Bland-Altman plot of the IBIs estimated from PPG Sensor compared to 
the IBIs estimated from the ECG Holter. ............................................. 36

Fig. 17 – Bland-Altman plot of the IBIs estimated from ICA-Basler compared to 
the IBIs estimated from the ECG Holter ............................................. 37
Fig. 16 – Bland-Altman plot of the IBIs estimated from Hue-Basler compared to the IBIs estimated from the ECG Holter. .......................................................... 37

Fig. 19 – Bland-Altman plot of the IBIs estimated from ICA-Webcam compared to the IBIs estimated from the ECG Holter. .......................................................... 38

Fig. 18 – Bland-Altman plot of the IBIs estimated from Hue-Webcam compared to the IBIs estimated from the ECG Holter. .......................................................... 38

Fig. 20 – 15.5 seconds excerpt of the signals in time. .......................................................... 40

Fig. 21 – Mean and variance of absolute error of estimated HRV parameters using video and PPG compared to ECG (Blue is mean; Orange is SD). .......... 42

Fig. 22 – Ectopic beat with visible hemodynamic response .............................................. 44

Fig. 23 – Ectopic beat with unnoticeable hemodynamic response .................................. 44

Fig. 24 – IBIs as a function of beat count (accurate detection). ...................................... 46

Fig. 25 – IBIs as a function of beat count (inaccurate detection) .................................... 46

Fig. 26 – ROC Logitech Webcam RMSSD ......................................................................... 50

Fig. 27 – ROC Basler ACE RMSSD .................................................................................. 50

Fig. 28 – ROC Basler ACE Occupied Bandwidth .............................................................. 51

Fig. 28 – ROC Logitech Webcam Occupied Bandwidth .................................................. 51
List of Tables

TABLE 1 - Webcam specifications ................................................................. 19
TABLE 2 - Basler Ace specifications ............................................................. 20
TABLE 3 - Mortara H12+ specifications ....................................................... 22
TABLE 4 - Binar Heart Sensor HRS-07UE specifications ............................. 23
TABLE 5 - Dell Latitude E6440 specifications ............................................ 24
TABLE I – Absolute error of estimated HRV parameters using VPG and PPG
compared to ECG ...................................................................................... 41
TABLE II - Critical image sensor parameter comparison ............................. 52
TABLE III – Classification results using normalized RMSSD method ......... 48
TABLE IV – Classification results using normalized RMSSD method .......... 48
TABLE V – Classification results using occupied bandwidth method ........... 48
TABLE VI – Classification results using occupied bandwidth method ........... 48
Summary of Contributions

- **Atrial fibrillation detection** – We propose and evaluate a system to detect atrial fibrillation using facial videos.

- **Comparison of two RGB cameras for facial videoplethysmography** – We assess the performance obtained by evaluating cardiac pulse measurements using two RGB cameras, and present the benefits obtained by using a higher end sensor.

- **Comparison of Hue and ICA approaches** – We present the comparison of Hue and ICA approaches for the extraction of cardiac pulse from facial RGB video data.

- **Heart variability indicator** – A new heart variability indicator is developed to identify episodes of atrial fibrillation.

- **Synchronized database** – A database is built, including synchronized data from two RGB cameras, an earlobe PPG sensor, and a body surface ECG recording.

- **Publications** –
Introduction

Atrial fibrillation is the most common clinically significant cardiac arrhythmia [1], affecting 33.5 million [2] patients worldwide. It is more common within the aging population, as men 75 to 79 years of age show two times more prevalence when compared to men 65 to 69 years of age, and five times more prevalence when compared to men 55 to 59 years of age [1]. Furthermore, the results presented in [1] demonstrate a progressive increase in the overall incidence, prevalence, and atrial fibrillation associated mortality between 1990 and 2010, which justifies the growing concern around atrial fibrillation.

Through atrial fibrillation, electrical impulses responsible for the contraction of the atria during a heart beat are triggered from many areas in and around the atria rather than just one area, which results in an irregular tremble of the atrial walls instead of an actual contraction (depicted in Fig. 1). This irregularity in the atria is associated with the loss of

Fig. 1– Electrical activity of a normal heart (left) and a heart with atrial fibrillation (right) [3]
atrial contribution to ventricular filling, thus decreasing the ventricular stroke volume up to 20 percent [4].

Atrial Fibrillation has been correlated with significant morbidity and mortality, increasing the risk of stroke up to 5 times, doubling the risk for dementia, and tripling the risk for heart failure [5]. Nevertheless, these risks can be reduced with treatment, lowering the mortality by 25% and the risk of ischaemic stroke by 68% with anticoagulant therapy [6].

![ECG signal of atrial fibrillation subject.](image)

Atrial Fibrillation is usually diagnosed upon examination of an electrocardiographic (ECG) signal. This signal is characterized by the replacement of consistent P waves by rapid oscillations or fibrillatory waves that vary in amplitude, shape, and timing, associated with an irregular, frequently rapid ventricular response [7].

The overall heart variability is usually quantized using the Heart Rate Variability (HRV) indicators standardized by the Task Force of the European Society of Cardiology and The North American Society of Pacing and Electrophysiology [8]. Such indicators provide a quantitative assessment of autonomic activity based on the variation of Inter Beat Intervals (IBI). HRV indicators such as the average IBI between normal heart beats (MeanRR), the standard deviation of the IBIs between normal heart beats (SDNN), and the
root mean square difference between consecutive normal IBIs (RMSSD), are often used as a statistical approach for the detection of lethal arrhythmias, like atrial fibrillation.

While ECG continues to be the gold standard for the diagnose of atrial fibrillation, acquiring such signal usually requires the subject to wear a cumbersome array of electrodes to capture the heart’s electrical activity. Furthermore, in a general sense, the apparatus required to record ECG signals is not widely accessible to the general population.

This work explores a contactless approach to provide accurate pulse measurements for the diagnose of atrial fibrillation using a facial video recorded with an RGB camera as the source for the signal. Our proposal aims to provide a more accessible means of diagnosing atrial fibrillation that does not require contact with the subject.
Chapter 1. Literature review

1.1 Photoplethysmography

Introduced in 1937 by A. B. Hertzman [9], *PhotoPlethysmoGraphy* (PPG) is a low cost non invasive technique used to detect blood volume changes in the tissue. Essentially, it comprises a light source which illuminates the tissue, and a photodetector which measures the subtle variations in light absorption caused by the changes in perfusion in the catchment volume [10]. PPG is widely used in clinical physiological monitoring, as it offers information regarding blood oxygen saturation, heart rate, blood pressure, and cardiac output.

PPG has two operational modes: transmission and reflection (presented in Fig. 3). The transmission mode requires the tissue to be placed between the light source and the photodetector to measure the changes in transmitted light. On the other hand, the reflection mode measures the changes in reflected light. The transmission mode limits the measurement site to areas such as the finger tip and the earlobes.

![Fig. 3 – Transmission (left) and Reflection (right) mode PPG configurations [11].](image-url)
1.2 Videoplethysmography

Following the same principle as reflection mode PPG, *VideoPlethysmoGraphy* (VPG) measures the small variations in light intensity as perceived by a camera, which acts as the photo detector, recording the tissue. The technique used to obtain the VPG signal measures the average brightness per frame in order to represent the changes in light intensity as a function of time. Just like a PPG signal, the resulting VPG signal would have a pulsatile component, with a fundamental frequency representing the heart rate, combined with a much slower varying signal that varies as a function of respiration rate, thermoregulation, and vasomotor activity.

Initial efforts exploring the possibility of acquiring a remote VPG signal often used a dedicated light source and a special camera sensor array. The work presented by T. Wu et al. [12] (2000) made use of a near infra-red light source coupled with an IR filtered monochrome *Charge-Coupled Device* (CCD) consumer handheld camera to evaluate changes in venous blood. Similarly, Wieringa et al. [13] (2005) used near infra red and infra red light sources coupled with an experimental, filtered, monochrome *Complementary Metal-Oxide Semiconductor* (CMOS) camera to estimate arterial oxygen saturation. Unlike the aforementioned studies, which used recordings obtained from limbs, Takano et al. [14] (2007) used ambient light paired with a commercially available monochrome CCD consumer handheld camera to estimate heart and respiration rates from facial recordings. Along the same line of thought, Verkruysse et al. [15] (2008) estimated heart and respiration rates from facial videos recorded using ambient light paired with a commercially available RGB CCD consumer handheld camera, and evaluated the
performance obtained by the RGB bands. In the discussion section of this latter study, the authors acknowledge that the G channel yields the strongest plethysmographic signal, and noted that the R and B channels may contain complementary information.

Following the remarks from Verkruysse et al. [15], the study presented by Poh et al. [16] (2010) introduced the analysis of the average RGB values recorded per frame as independent sources for the underlying plethysmographic signal. This latter study proposed a system to estimate average heart rate using ambient light paired with a standard RGB CMOS webcam, using a blind source separation technique named Independent Component Analysis, using the RGB averages as the independent sources to obtain the plethysmographic signal. Furthermore, this study utilized face tracking as a means to improve the estimation in the presence of motion. The results obtained using the technique described by Poh et al. [16] were very encouraging and could be replicated using widely available webcams.

As noted by Sun et al. [17] (2016), there has been a significant increase in literature related to VPG. These recent efforts have demonstrated the feasibility of the use of VPG signals extracted from facial videos as an alternative to conventional PPG sensors. The majority of these studies employs some form of blind source separation as part of the signal processing needed to extract the pulsatile plethysmographic signal from the RGB traces.

The work presented by Poh et al. [18] (2011), continues their previous effort and explores the estimation of Heart Rate Variability with their VPG system. Along the same line, the work presented by Shan et. al. [19] (2013), using a CMOS RGB smartphone camera and ambient light, explores additional face detection and tracking algorithms and
compares the performance obtained by using Independent Component Analysis (ICA) versus Principal Component Analysis (PCA). Similarly, Scalise et. al. [20] (2012) evaluated the feasibility of using a VPG system, based on a commercially available RGB CMOS webcam and the use of ICA, to monitor neonatal patients in an intensive care unit.

Despite the fact that these aforementioned studies obtained favorable results for average heart rate estimations, it is known that the output of the ICA method returns the resulting Independent Components in a random order, thus effectively limiting the accurate reconstruction of the plethysmographic signal. The work presented by Tsouri et. al. [21] (2012) expands on the performance limitations introduced by the sorting problem inherent to ICA methods, and proposes a Constrained ICA algorithm as the solution. However, the proposed algorithm requires prior knowledge of the parameter being extracted, thus limiting its application to assess signals with significant variability.

Some other recent efforts have concentrated on improving the overall performance of the VPG system by adding signal processing steps to compensate for movement. Such is the case of the study presented by Wang et. al. [22] (2015), in which they use multiple segments of the facial area, recorded with a commercially available RGB CCD machine vision camera using ambient light, containing skin (identified through a skin classificator) as separate sensors, thus enabling a spatial diversity analysis to improve the signal to noise ratio of the plethysmographic signal. Their approach includes the use of color space pruning, adaptive band-pass filtering, and Principal Component Analysis. Additionally, their approach uses a standardized skin tone assumption and the chrominance VPG model introduced by Haan et. al. [23] (2013).
In an effort to reduce the computational effort required to extract the VPG signal, Tsouri et al. [24] (2015) explored the use of alternative color spaces derived from RGB as a surrogate to ICA. After analyzing facial videos recorded with a commercially available RGB webcam, their results indicate that the Hue component from the HSV color space provides a suitable alternative to ICA from RGB.

1.3 Atrial fibrillation diagnosis

Atrial fibrillation is usually diagnosed upon examination of a body surface ECG signal in which the absence of P waves, and a rapid ventricular response, would mark the occurrence of an Atrial Fibrillation episode. However, the P wave can be easily obscured in the presence of noise, as it usually has a very low amplitude when compared to the R waves. Because of this reason, the automatic detection of atrial fibrillation is currently divided into two main categories: the analysis of P waves [25,26,27,28], and the analysis of RR intervals (interval between consecutive peaks of the QRS complex) [29,30,31,32,33].

Given that the P wave would only be visible in ECG signals, as it depicts the electrical impulse that triggers the atria, and that such wave is not present in the blood volume pulsations recorded via plethysmography (PPG, VPG), we will focus on the atrial fibrillation detection methods that rely on the analysis of RR intervals.

Tateno et al. [30] used the RR intervals, and the difference between consecutive RR (ΔRR) intervals to establish standard density histograms and compared the histograms using the Kolmogorov-Smirnov test to measure the difference between distributions. Additionally, they compared the classification obtained by using a Coefficient of Variation
test, being the coefficient equal to the standard deviation of RR (or ∆RR) intervals divided by the mean RR (or ∆RR) interval, to the Kolmogorov-Smirnov test. The best classification was with the Kolmogorov-Smirnov test on the ∆RR intervals, obtaining a sensitivity of 94.4%, for a specificity of 97.2%.

Duverney et al. [31] identified atrial fibrillation episodes using wavelet and fractal analysis of RR intervals. The first step of their analysis is to identify periods of high variability using a discrete wavelet transform and then further classify the high variability periods in sinus rhythm or atrial fibrillation using fractal analysis. Their method provided a sensitivity of 96.1%, for a specificity of 92.6%.

Dash et al. [32] used a combination of a turning point test, a normalized root mean square of successive RR intervals (RMSSD/MeanRR), and Shannon Entropy derived from RR intervals to evaluate the presence of atrial fibrillation episodes. Their method, applied to segments of 128 beats, yielded a sensitivity of 94.4%, for a specificity of 95.1%.

Zhou et al. [33] used the RR intervals to derive an instantaneous heart rate that was then translated into a 64-bit symbol (max heart rate = 315 bpm). These symbols were then used to form a word (with 3 consecutive symbols), for which a value would be assigned (0 to 262143). The Shannon entropy of 127 consecutive words was evaluated to identify atrial fibrillation episodes. Their method provided 96.14% sensitivity, for 95.73% specificity.

It is worth noting that the aforementioned studies used the MIT arrhythmia database to train and test their algorithms. This database contains a total of 48 half-hour excerpts of two-channel, 24-hour, ECG recordings obtained from 47 subjects [34]. The database is not
exclusively dedicated to atrial fibrillation, as it contains annotated beats for a variety of arrhythmias.

Nevertheless, the fact that R waves can be correlated to the blood volume pulsations present in plethysmographic signals enables the use of atrial fibrillation classification algorithms derived from RR intervals on plethysmographic signals. An example of such use can be found in the work presented by Lee et al. [35]. In addition to the MIT database, they used IBIs extracted via VPG data captured from the fingertip (contact) of 25 atrial fibrillation subjects, using the green channel from the camera in a smartphone, to identify atrial fibrillation episodes. They compared the classification results obtained by evaluating the normalized root mean square of successive IBIs (RMSSD/Manabí), the Shannon Entropy, and the Sample Entropy, obtaining a sensitivity of 97.63%, for a specificity of 99.61%, using the RMSSD method on 64 beats segments.

Another notable study exploring the possibility of detecting atrial fibrillation using videoplethysmography is the research presented by Couderc et al. [36]. They analyzed the data captured from the facial videos 11 atrial fibrillation subjects, recorded using a standard webcam and the constrained ICA technique introduced by Tsouri et al., to identify atrial fibrillation episodes. For that purpose, they introduced a quantifier of pulse variability named Pulse Harmonic Strength, obtained by measuring the energy of the fundamental frequency and its main harmonics within 0.05- and 3Hz bandwidth divided by the energy of the remaining spectral space. Their method reported a 20% classification error rate for 15 seconds segments.
Chapter 2.Proposed Approach

2.1 The sensing environment

As noted earlier, at its core, the VPG system comprises a light source to illuminate the tissue and a photodetector to measure the small variations in light intensity resulting from the blood volume changes in the tissue. The VPG system must take into consideration the characteristics of the environment in which the camera will acquire the signal and how the signal is perceived by the camera in that environment. The spectral distribution of the radiation from the light source, along with its overall consistency and uniformity, have a tremendous impact on the information captured by the camera. Frequently utilized light sources, such as fluorescent and LED lamps, aim to resemble the light projected by the sun with the use of an additive combination of RGB. Each of these colors have a different interaction with the skin based on their wavelengths, where the wavelengths ranging from 510 to 590 nm provide the maximum blood pulsation modulation for reflected light [37]. This interaction between the light source, medium, tissue, and photodetector is depicted in Fig. 4.

The VPG signal is present in the small variations around a mean value of color intensity. These variations usually present a fundamental frequency corresponding to the heart rate, and modulate a much slower varying signal that is influenced by variables such as: respiration, temperature, and vasomotor activity. The camera serving as the
photodetector for the VPG system must be sensitive enough to capture these small variations, even more so in the presence of an irregular rhythm with an inconsistent pulsation, such as atrial fibrillation.

The irregular cardiac output caused by atrial fibrillation is reflected as a variable amplitude of the pulsatile plethysmographic signal. Furthermore, the IBIs that characterize this irregularly irregular arrhythmia circumvent the periodicity of the cardiac pulse. These conditions introduced by atrial fibrillation create a pulsatile signal with random period and amplitude, thus posing an additional hindrance to identify the pulse, even in the absence of noise.

Fig. 4 – Sensing platform.
Considerations for an ideal VPG camera sensor

- **Large pixel size**: larger pixels are able to capture more light, thus providing a better signal to noise ratio when compared to smaller pixels. Additionally, a larger pixel can also provide a larger dynamic range.

- **Global shutter**: with a global shutter all pixels are exposed to the scene at the same time, thus avoiding an overlap between instances of time in the same image/frame.

- **Low temporal dark noise**: temporal dark noise is the noise inherent to the camera and it is caused by its electronic components and circuits.

- **High dynamic range**: a wide spectrum of values for the analog to digital converter to chose from aids an accurate representation of the subtle variations in the tissue.

- **High quantum efficiency**: quantum efficiency is the ability to convert photons to electrons. High quantum efficiency translates to higher sensitivity as the sensor requires less light to output a pixel value.
2.2 Signal processing for VPG

2.2.1 Raw signal extraction

The raw traces used in VPG systems are obtained by computing the average value of all pixels per color band, per frame, thus rendering a time vector of the average intensity changes as perceived by each band. In order to improve the signal to noise ratio of the raw traces, these averages are usually computed from the pixels contained within a region of interest. In the case of facial videoplethysmography, a face detection and tracking algorithm can be implemented for this purpose, as shown in the work presented in [16,18].

Once the raw R, G, and B traces have been computed, a raw Hue trace can be obtained through a nonlinear combination of these values, as noted in [24], using the following formula:

\[
H' = \begin{cases}
0, & \text{if } C = 0 \\
\frac{G-B}{C} \text{ mod 6}, & \text{if } M = R \\
\frac{B-R}{C} + 2, & \text{if } M = G \\
\frac{R-G}{C} + 4, & \text{if } M = B
\end{cases}
\]

(1)

\[
H = 60 \text{ deg } xH',
\]

(2)

where

\[
M = \max(R,G,B),
\]

\[
m = \min(R,G,B),
\]

\[
C = M - m.
\]
As mentioned earlier, each color band acquired by the camera captures the variation of light as a function of the blood volume pulsations as experienced in different layers of the skin. The shorter wavelength, blue, reaches the epidermis, while the dermis and hypodermis are reached by green and red, respectively. Since Hue is obtained through a nonlinear combination of RGB, it offers a comprehensive assessment of the overall blood volume pulsations in the skin, as perceived by the RGB bands. Furthermore, a change in overall light intensity, which would equally affect all color bands, would not have any effect on the Hue trace as the pure color value would be maintained. The idea of utilizing color conversion comes from the premise that the pulsating heart might present itself better in a color space where variation in color is emphasized rather than variation in specific color intensity.

The raw traces are normalized using the standard score method using the equation below:

\[ x'_i(t) = \frac{x_i(t) - \mu_i}{\sigma_i} \]

where \( i = R, G, B, H; \mu_i \) and \( \sigma_i \) are the mean and standard deviation of \( x'_i(t) \).
2.2.2 Detrending

The raw plethysmographic signals contain linear and nonlinear trends produced by changes in light intensity, thermoregulation, respiration, and other variables. Removing such baseline shift improves the signal to noise ratio of the pulsatile component corresponding to the heart rate. There are multiple detrending techniques, from which the following could be highlighted: 1) first differencing, 2) curve fitting, 3) digital filtering, 4) piecewise polynomials, and 5) smoothness priors approach.

Given the nature of the trends present in the VPG signal, the most frequently used methods include the curve fitting and smoothness priors approach. While the former identifies the trend as a polynomial fitted with a least square error approach, which is then subtracted from the VPG data, the latter uses a regularized square solution to identify the trend. The computational effort required for the smoothness priors method is considerably higher than the least square error curve fitting method.
2.2.3 Filtering

The VPG signal is filtered to remove frequencies residing outside the expected bandwidth of a PPG signal. A bandpass filter with cutoff frequencies set at 0.7 and 4 Hz, corresponding to 40 and 240 bpm, respectively, is applied to the VPG in order to focus the analysis on the pulsatile signals that fall within the range of possible human heart rates.

![Fig. 5 – Frequency response of bandpass filter.](image)

An additional, narrower, adaptive bandpass filter could be applied to the VPG signal in order to further isolate the pulsatile component from other signals. However, the use of such filter could hinder the ability of the VPG system to accurately represent signals with high variability, such as atrial fibrillation.
2.3 Materials and apparatus

As noted earlier, a VPG system is essentially comprised by a light source and a photodetector. In order to validate our proposal, we explore the use of different photodetectors contrasted with standard cardiac monitoring devices described below.

Logitech Quickcam Pro 9000 – Standard Webcam

Released in 2007, the Quickcam Pro 9000 (Fig. 6) is a standard USB 2.0, 2 Megapixel camera. The quality of the images provided by this camera is still acceptable under today’s standards. Since most recent cameras provide better quality images, the performance provided by this camera is expected to represent the lower bound of the performance obtained using our proposal.
Table 1 - Webcam specifications [39]

<table>
<thead>
<tr>
<th>Optical format</th>
<th>1/3.2-inch</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full resolution</td>
<td>1,600 x 1,200 pixels (UXGA)</td>
</tr>
<tr>
<td>Pixel size</td>
<td>2.8 µm x 2.8 µm</td>
</tr>
<tr>
<td>Active pixel array area</td>
<td>4.73 mm x 3.52 mm</td>
</tr>
<tr>
<td>Shutter type</td>
<td>Electronic Rolling Shutter (ERS) with global reset</td>
</tr>
<tr>
<td>Maximum frame rate</td>
<td>15 fps at full resolution, 30 fps in preview mode (800 x 600)</td>
</tr>
<tr>
<td>Maximum data rate</td>
<td>8m MB/s</td>
</tr>
<tr>
<td>ADC resolution</td>
<td>10-bit, on-die</td>
</tr>
<tr>
<td>Responsivity</td>
<td>1.0/lux-sec @ 550nm</td>
</tr>
<tr>
<td>Dynamic range</td>
<td>71 dB</td>
</tr>
<tr>
<td>SNRmax</td>
<td>42.3 dB</td>
</tr>
<tr>
<td>Power consumption</td>
<td>348 mW @ 15 fps, full resolution, 223 mW @ 30 fps, preview mode (800 x 600)</td>
</tr>
<tr>
<td>Sensor type</td>
<td>CMOS</td>
</tr>
<tr>
<td>Sensor</td>
<td>On Semiconductor MT9D131 (or equivalent)</td>
</tr>
</tbody>
</table>

Fig. 7 – On Semiconductor MT9D131 (Webcam) image sensor Quantum Efficiency Vs Wavelength [39].
Basler Ace acA1920-155 uc – High performance camera

Introduced in 2014, the Basler Ace (Fig. 8) is a USB 3.0 area scan camera often utilized for machine vision purposes in different industries. It features a high frame rate image sensor with additional features that enable accurate color reproduction and high efficiency.

Table 2 - Basler ace specifications [41]

<table>
<thead>
<tr>
<th>Specification</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Optical format</strong></td>
<td>1/1.2-inch</td>
</tr>
<tr>
<td><strong>Full resolution</strong></td>
<td>1,920 x 1,200 pixels</td>
</tr>
<tr>
<td><strong>Pixel size</strong></td>
<td>5.86 µm x 5.86 µm</td>
</tr>
<tr>
<td><strong>Sensor size</strong></td>
<td>11.33 mm x 7.10 mm</td>
</tr>
<tr>
<td><strong>Temporal dark noise (e-)</strong></td>
<td>6.78 e-</td>
</tr>
<tr>
<td><strong>Shutter</strong></td>
<td>Global shutter</td>
</tr>
<tr>
<td><strong>Absolute Sensitivity Threshold</strong></td>
<td>10.44 γ</td>
</tr>
<tr>
<td><strong>Maximum frame rate</strong></td>
<td>164 fps at full resolution</td>
</tr>
<tr>
<td><strong>ADC resolution (max)</strong></td>
<td>12 bits</td>
</tr>
<tr>
<td><strong>Dynamic range (max)</strong></td>
<td>73.07 dB</td>
</tr>
<tr>
<td><strong>SNRmax</strong></td>
<td>45.15 dB</td>
</tr>
<tr>
<td><strong>Power consumption (typical)</strong></td>
<td>~3.4 W</td>
</tr>
<tr>
<td><strong>Sensor type</strong></td>
<td>CMOS</td>
</tr>
<tr>
<td><strong>Sensor</strong></td>
<td>Sony IMX174</td>
</tr>
</tbody>
</table>
The Basler camera was paired with a 1” 16mm lens for 600mm wide image capture at an approximate distance of 85cm.

**Mortara H12+ - ECG Holter**

The Mortara H12+ (Fig. 10) is a compact ECG Holter recorder that weighs just 4 oz., it was designed to be worn for extended periods of time, and allows up to 48 hours of...
The Mortara H12+ was paired with the Mortara LeadForm standard patient cable to acquire a 12 lead resolution using the 10 electrode configuration.

Table 3 - Mortara H12+ specifications [44]

<table>
<thead>
<tr>
<th>Instrument type</th>
<th>12-lead ECG Holter recorder</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input channels</td>
<td>simultaneous acquisition of all leads</td>
</tr>
<tr>
<td>Standard leads acquired</td>
<td>I, II, III, aVR, aVL, aVF, V1, V2, V3, V4, V5, V6</td>
</tr>
<tr>
<td>Input impedance</td>
<td>&gt;10MΩ</td>
</tr>
<tr>
<td>Digital sampling rate</td>
<td>10,000 s/sec/channel used for pacemaker spike detection. 180 s/sec/channel used for recording and storage</td>
</tr>
<tr>
<td>Special functions</td>
<td>Pacemaker detection, ECG display, Lead quality check</td>
</tr>
<tr>
<td>ADC resolution</td>
<td>20 bit</td>
</tr>
<tr>
<td>Storage</td>
<td>Compact flash memory card</td>
</tr>
<tr>
<td>Battery</td>
<td>1 AA alkaline required</td>
</tr>
</tbody>
</table>

continuous recording. This device meets all the requirements of ANSI/AAMI EC38 for medical instrumentation for ECG measurements.
Binar Heart Sensor HRS-07UE – Earlobe PPG sensor

The Binar Heart Sensor HRS-07UE (Fig. 11) is a small USB ear-clip reflective PPG sensor designed for long term pulse measurements. It is a simple and low cost sensor that provides high quality measurements.

<table>
<thead>
<tr>
<th>Table 4 - Binar Heart Sensor HRS-07UE specifications [45]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Weight (with cable)</strong></td>
</tr>
<tr>
<td><strong>Dimensions</strong></td>
</tr>
<tr>
<td><strong>Signal Bandwidth</strong></td>
</tr>
<tr>
<td><strong>A/D conversion</strong></td>
</tr>
<tr>
<td><strong>Sampling rate</strong></td>
</tr>
<tr>
<td><strong>Housing</strong></td>
</tr>
<tr>
<td><strong>Operating temperature</strong></td>
</tr>
<tr>
<td><strong>Data interface</strong></td>
</tr>
</tbody>
</table>
Dell Latitude E6440 – Data gathering computer

Launched in 2014, the Dell Latitude E6440 (Fig. 12) serves as a robust data gathering computer. Its 4 USB 3.0 ports, upgraded hard drive, expanded RAM, and fast processor enable simultaneous acquisition of data from different sources.

Table 5 - Dell Latitude E6440 specifications.

<table>
<thead>
<tr>
<th>Specification</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Processor</td>
<td>Intel Core i7 4610M @ 3.0GHz, 4 MB SmartCache, 64-bit instruction set, 22 nm Lithography</td>
</tr>
<tr>
<td>Operating System</td>
<td>Microsoft Windows 7 Enterprise 64-bit</td>
</tr>
<tr>
<td>Memory</td>
<td>16GB DDR3L @ 1600MHz</td>
</tr>
<tr>
<td>Chipset</td>
<td>Mobile Intel QM87</td>
</tr>
<tr>
<td>Graphics Card</td>
<td>Intel integrated HD Graphics 4600</td>
</tr>
<tr>
<td>Display</td>
<td>14” (1366x768)</td>
</tr>
<tr>
<td>Hard Drive</td>
<td>Samsung SSD 850 EVO SATA 6GBs, Sequential Read Max. 540 MB/s, Sequential Write Max. 520MB/s</td>
</tr>
<tr>
<td>Optical Drive</td>
<td>DVD+/-RW</td>
</tr>
<tr>
<td>Connectivity</td>
<td>10/100/1000 Gigabit Ethernet, Dell Wireless 1506 (802.11g/n 1x1)</td>
</tr>
<tr>
<td>Battery</td>
<td>9-cell (97Wh) Lithium Ion battery with ExpressCharge</td>
</tr>
<tr>
<td>Ports</td>
<td>4 USB 3.0, HDMI, VGA, RJ-45</td>
</tr>
</tbody>
</table>
2.4 Detection of Inter Beat Intervals

IBIs are extracted from the VPG signal, after the heartbeats have been identified, by measuring the difference between contiguous beats. The nature of the VPG signal allows for its pulsatile component, which represents a heartbeat, to be automatically identified by using a peak detection algorithm. However, as noted in [38], this task is more challenging for plethysmographic signals than it is for ECG signals, as the morphology of the pulse in plethysmographic signals is less evident than the QRS complex present in ECG signals. Similarly, the varying hemodynamic response regulated by the autonomic nervous system makes plethysmographic signals much more fluctuating than their ECG counterparts.

The majority of automatic peak detection algorithms that have been developed for the analysis of biomedical signals revolve around the extraction of features from ECG signals [47, 48]. Nevertheless, there have been a few notable studies that explored the peak detection problem present in plethysmographic signals. An example of such case is the work presented by Aboy et al. [49], in which the authors propose an algorithm to automatically detect the systolic peak in SpO₂ signals, among other features from pressure signals. However, the authors indicate that the algorithm is designed for subjects without significant cardiac arrhythmias, meaning its performance would not be maintained in the presence of atrial fibrillation.

Another notable contribution was made by Liu et al. [50], who presented an automatic beat detection for blood pressure signals using a combination of a windowed and weighted slope sum function, the mean shift algorithm, and a peak calibration procedure.
Their results showed a significant level of agreement between automatic detection and expert annotation, nearing 99% on the majority of their evaluations.

The contactless approach of facial VPG poses an additional layer of complexity to the beat detection problem. The pulsatile signal measured on the facial tissue must traverse a medium to reach the image sensor on the camera, which makes the VPG signal more susceptible to noise.

The complexity of peak detection in VPG signals escalates further with the irregularity introduced by atrial fibrillation. For that reason, our approach steers away from further signal processing and performs peak detection based on thresholds. We used the first 10 seconds of the processed signals to obtain an RMS value for each signal/subject. The peaks are initially identified by locating the transitions from rising to dropping, noticeable in the derivative of the VPG signal. These peaks are then marked as heartbeats if their amplitude is greater than 0.61*RMS value of the signal, and they are located at least 0.33 seconds away from the preceding peak.
2.5 Detection of Atrial Fibrillation

2.5.1 Using IBI Detection

As noted in [32], a normalized RMSSD value can be utilized to determine the occurrence of an atrial fibrillation episode using IBI data. The same principle was used in the analysis presented by Lee et al. [35]. Despite the fact that the same parameter was used to identify atrial fibrillation episodes, these studies do not share the same value for the threshold that differentiates between normal sinus rhythm and atrial fibrillation. While Dash et al [32] found the optimal threshold value to be 0.1, Lee et al. [35] found the value to be 0.13 for the data from the MIT database, and 0.8494 for the data captured using the smartphone camera. Our approach will assess the occurrence of atrial fibrillation episodes setting the threshold for normalized RMSSD at 0.113 using the 5.5 minutes recording, regardless of the beat count. A subject is said to be in atrial fibrillation if the value of the normalized RMSSD is greater than the threshold.

2.5.2 Using Occupied Bandwidth

Atrial fibrillation classification algorithms based on IBIs are completely dependent on the accuracy of the peak detection algorithm utilized to estimate the IBIs. As mentioned earlier, the estimation of the IBIs on facial VPG signals is much more challenging in the presence of atrial fibrillation, and misdetection of heartbeats would considerably impact the performance of IBI based classification methods. These considerations lead us to explore the alternative of exploiting other characteristics of the VPG signal to identify atrial fibrillation episodes.
What remains after the VPG signal has been detrended, normalized, and filtered, is the pulsatile component, corresponding to the heart rate, with a morphology that resembles a sine wave with a mean value of zero. The information needed to identify atrial fibrillation episodes lays on the variability of the period of this sine wave.

The main frequency of the pulsatile component corresponds to the heart rate. The variability of the IBIs around this main heart rate resembles a frequency modulated signal. In this model, the IBI information we are interested in modulates the carrier, represented by the main heart rate. This analogy allows for us to hypothesize that the bandwidth occupied by the VPG signal widens as a function of the variability of the IBIs. Consequently, the bandwidth occupied by the signal, would provide a heart rate variability indicator that can be obtained without performing peak detection.

Let $x(t)$ and $P_{xx}(f)$ be the pulsatile plethysmographic signal representing the pulse and its power spectral density. The total power of the signal within the $[0.7, 4]$ Hz range would be given by:

$$\int_{0.7}^{4} P_{xx}(2\pi f)df = \text{Total Signal Power (TSP)}$$

It is important to note that the total power in this frequency range contains the power the harmonics of the main pulsatile signal. For this reason, we choose to evaluate the bandwidth occupied by a portion of the signal, which corresponds to the main pulsatile component but does not include other harmonics.
Symmetry in bandwidth distribution

\[
\begin{cases}
\int_{0.7}^{f_{\text{low}}} P_{xx}(2\pi f) df = P_{xx}(0.7) + \frac{1}{2} (1 - p) * TSP, \\
\int_{0.7}^{f_{\text{high}}} P_{xx}(2\pi f) df = P_{xx}(0.7) + \frac{1}{2} (1 + p) * TSP,
\end{cases}
\]

Where \( p \) represents the portion of the signal to be measured (i.e. \( \frac{1}{4} \)), and the occupied bandwidth is obtained by solving for \( f_{\text{low}} \) and \( f_{\text{high}} \), and subtracting \( f_{\text{low}} \) from \( f_{\text{high}} \).

Our approach will assess the presence of atrial fibrillation based on the bandwidth occupied by 30% of the total signal power within the [0.7, 4] Hz range. If the occupied bandwidth is above 100 mHz, the subject is said to experience AF.
Chapter 3. Clinical Study

We developed custom software using Matlab (Mathworks, Natick, MA) and C++ to capture and analyze data. The software simultaneously captures synchronized data from the following sensors:

- Logitech (Logitech International, Romanel-sur-Morges, Switzerland) Quickcam Pro 9000 (RGB webcam). 640x480 24-bit RGB frames sampled at 30Hz were stored in binary format with no compression. All automatic adjustments (gain, contrast, exposure, etc.) were disabled and manually configured for each capture. A manual region of interest containing the subject’s face was selected for analysis.

- Basler (Basler AG, Ahrensburg, Germany) ACE 1920-155uc (RGB camera). A manual region of interest (typical 600x500) containing the face was selected before each capture, enabling a sampling frequency of 180Hz while outputting 12 bits per channel (36 bits in total per frame, Bayer RG 12) to an uncompressed binary file. All automatic adjustments (gain, black level, exposure, etc.) were disabled and manually configured for each capture.

- Binar (Binar, Poulsbo, WA) HeartSensor HRS-07UE (PPG sensor). Reflective PPG sensor sampled at 300Hz with a 16-bit per sample output to an uncompressed binary file.

The H12+ (Mortara Instruments, Milwaukee, MN) Holter ECG recorder was used to record the body surface ECG signal, capturing 12 leads with a resolution of 20 bits per lead sampled at 180Hz. In order to synchronize the ECG data with the data captured
from the other sensors using our custom software, a synchronization event was generated by momentarily interrupting lead V5 to mark the start of the video acquisition.

3.1 Subjects

3.1.1 Healthy cohort

Fifty healthy subjects (22 males and 28 females) between the ages of 18-59 were enrolled to participate in a clinical study approved by the Internal Review Committees for Protecting Human Subjects at the University of Rochester Medical Center (URMC, Rochester, NY) and the Rochester Institute of Technology (RIT, Rochester, NY). All subjects were recorded for a period of 5 minutes and 30 seconds in an examination room in the clinic located at URMC. The room was illuminated by fluorescent light, standard in all patient rooms, and no additional light source was used. Subjects were asked to assume a supine position in an examination bed, as presented in Fig. 13, and minimize movement while being recorded.

![Schematic description of the experimental setup of the study.](image-url)
The 10 electrodes of the Mortara H12+ recorder were positioned on the subject’s torso using a Mason-Likar configuration to record the ECG signals, and ECG recording was started before the acquisition from all other sensors was initiated. Additionally, the PPG sensor was clipped to the left ear of the subject while the cables of the Holter were connected to the skin electrodes. The cameras were then placed above the subject’s head, approximately 85 cm away.

3.1.2 Cohort of patients with Atrial Fibrillation

Thirty-four patients (29 males and 5 females) between the ages of 38-82 years, who underwent direct current electrical cardioversion at Strong Memorial Hospital (Rochester, NY), were enrolled to participate in a clinical study approved by the Internal Review Committees for Protecting Human Subjects at the University of Rochester Medical Center (URMC, Rochester, NY) and the Rochester Institute of Technology (RIT, Rochester, NY). Subjects were excluded if an implanted cardio-defibrillator or cardiac resynchronization therapy device was present, if they refused to sign the consent form, or they were unable to cooperate with the protocol. All subjects were recorded for a period of 5 minutes and 30 seconds, in a room in the cardiac catheterization and electrophysiology lab located at Strong Memorial Hospital, before and after cardioversion. The room was illuminated by fluorescent light, standard in all patient rooms, and no additional light source was used. Subjects were asked to assume a supine position in an examination bed, as presented in Fig. 13, and minimize movement while being recorded. The acquisition of the signals followed the same procedure that was described for the healthy subject cohort, with the exception of having two captures: 1) pre cardio version, and 2) post cardioversion.
After the pre cardioversion recording was completed, the PPG sensor and the cameras were removed for the cardioversion to take place. Once the cardioversion was completed, the cameras were repositioned, and the PPG sensor was clipped to the subjects left earlobe to capture the post cardioversion recording.

### 3.2 Data Gathering

The raw color filter array frames captured by the Basler camera were “debayered” in Matlab to generate a truecolor RGB frame before processing. The average value of all the pixels, contained in the region of interest established at the time of the capture, per color channel was obtained per frame to generate a vector of R, G, and B per camera. These vectors were used to compute a raw Hue vector per camera using the equations (1) and (2) shown in Chapter 2.

The nonlinear trends present in the raw traces (R, G, B, and H) obtained from each camera and the raw trace recorded by the PPG sensor were eliminated by fitting a polynomial of order 6 to the raw signal and then subtracting it, as described in [51]. The detrended traces were normalized using the standard score method described in Chapter 2.

For the purpose of comparison, the JADE [52] implementation of the ICA algorithm was applied to the normalized detrended traces, selecting R, G, and B as three independent sources for the analysis. The output of the ICA and the normalized detrended traces were filtered using a 1024 points bandpass [0.75, 4] Hz finite impulse response filter. This output was then smoothened by passing it through a 5-point moving average. All processed signals were interpolated to 300Hz using the cubic spline method.
The IBIs were computed after the beats were identified based on overall height, and distance relative to neighboring peaks, as described in Chapter 2. For ICA, the IBIs extracted from the output with the highest spectral peak, within the [0.75, 4] Hz range, were used for the comparison as was done in [18]. Using this criteria, we were able to achieve 100% beat detection.

Given that all sensors experience a different delay, relative to the electric impulse recoded by the ECG Holter on a per subject basis, the correlation between the IBI data from each sensor and the ECG data was determined to ensure proper alignment of the signals for IBI comparison.
Chapter 4. Results and Discussion

4.1 IBI detection

4.1.1 Healthy subjects

We use the IBI data acquired from all subjects in the healthy cohort to illustrate the overall IBI detection performance of the different sensors/channels for normal sinus rhythm, represented with Bland-Altman plots [53] in which the data extracted from the ECG Holter monitor is the baseline. We omitted twelve subjects from the data used to populate the graph for which data was corrupted during recording: three subjects due to removal of the PPG sensor, two subjects due to corrupted ECG data, two subjects due to videos cut shorter than four minutes, and five subjects that clearly exhibited ectopic beats. The data captured from the Basler camera was downsampled to match the frame rate and bit depth of the webcam in order to focus the comparison on the actual efficiency of the image sensors.
Fig. 14 – Bland-Altman plot of the IBIs estimated from Green-Basler compared to the IBIs estimated from the ECG Holter.

Fig. 15 – Bland-Altman plot of the IBIs estimated from PPG Sensor compared to the IBIs estimated from the ECG Holter.
Fig. 16 – Bland-Altman plot of the IBIs estimated from Hue-Basler compared to the IBIs estimated from the ECG Holter.

Fig. 17 – Bland-Altman plot of the IBIs estimated from ICA-Basler compared to the IBIs estimated from the ECG Holter.
Fig. 19 – Bland-Altman plot of the IBIs estimated from ICA-Webcam compared to the IBIs estimated from the ECG Holter.

Fig. 18 – Bland-Altman plot of the IBIs estimated from Hue-Webcam compared to the IBIs estimated from the ECG Holter.
Figs. 14-19 depict the difference of IBI detection compared to ECG in milliseconds as a function of the average value of the IBIs in milliseconds. Fig. 16, which corresponds to Hue-Basler, provides the closest resemblance to the result presented in Fig. 14, which corresponds to the PPG sensor. Both, Fig. 14 and Fig. 16, exhibit a mean error that is close to zero. Examining Figs. 14-19, it is evident that Hue-Basler provides the most accurate estimation among all estimations obtained from both cameras. We observed that, in most cases, ICA would provide a level of accuracy comparable to Hue, yet its accuracy was limited by the sorting problem, as the component with the highest spectral peak did not always provide the best result.

While the average error is zero for all sensors, a linear trend can be observed for some IBIs estimated from the webcam and the Basler camera. The trend is a product of instantaneous decrease of the SNR in the signal, which produces an erroneous detection of noise peaks as heartbeats. With a regular heart rate, each individual error in peak detection results in two IBI estimation errors (before and after the additional noise induced peak), where the IBI to the left of the noise induced peak would be shorter than the reference IBI, and the subsequent IBI would be longer than the reference in a proportion that compensates for its shorter predecessor. Multiple instances of this event produce a visible linear trend in the Bland-Altman comparison.

When comparing the results obtained from the Basler camera (presented in Fig. 15, Fig. 16, and Fig. 17) to the results obtained from the Logitech webcam (presented in Fig. 18 and Fig. 19) we can clearly establish that the Basler camera provides a higher level of
accuracy. The error is closer to zero on all the estimations extracted from the Basler camera when compared to their respective webcam counterparts.

It is worth mentioning that the morphology of the time based signals extracted from the Basler camera is almost identical to that of the time signal extracted from the PPG sensor. Fig. 20 presents a 15.5 seconds interval of the time based signals. Out of all traces, Hue-Basler is the one that consistently exhibits the closest resemblance to the morphology of the signal from the PPG sensor. In most cases ICA would provide an output almost identical to Hue, as it is the case with the pair Hue-ICA obtained from the webcam for this subject. However, the ICA signal extracted from the Basler camera suffered from the sorting problem in this case, as the component with the highest spectral peak does not render the best representation of the signal.

Fig. 20 – 15.5 seconds excerpt of the signals in time.
<table>
<thead>
<tr>
<th>Subject ID</th>
<th>PPG</th>
<th>Hue-Basler</th>
<th>ICA-Basler</th>
<th>Green-Basler</th>
<th>Hue-Webcam</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MeanRR SDNN RMSSD</td>
<td>MeanRR SDNN RMSSD</td>
<td>MeanRR SDNN RMSSD</td>
<td>MeanRR SDNN RMSSD</td>
<td>MeanRR SDNN RMSSD</td>
</tr>
<tr>
<td>6</td>
<td>3.6 1.3 3.8</td>
<td>0.2 1.0 3.8</td>
<td>0.2 1.0 3.3</td>
<td>0.2 4.4 17.4</td>
<td>7.3 11.8 38.0</td>
</tr>
<tr>
<td>7</td>
<td>3.6 0.1 3.3</td>
<td>0.1 0.4 3.7</td>
<td>5.2 21.7 21.8</td>
<td>39.7 112.5 169.2</td>
<td>1.1 17.2 40.7</td>
</tr>
<tr>
<td>8</td>
<td>4.3 1.1 3.0</td>
<td>0.1 1.0 2.9</td>
<td>0.1 0.7 2.2</td>
<td>0.1 0.7 2.2</td>
<td>2.6 7.9 15.7</td>
</tr>
<tr>
<td>9</td>
<td>2.9 0.2 0.8</td>
<td>0.4 1.8 7.0</td>
<td>0.4 2.2 8.3</td>
<td>3.8 28.0 62.5</td>
<td>0.1 19.1 48.0</td>
</tr>
<tr>
<td>10</td>
<td>3.5 2.3 5.2</td>
<td>4.9 16.8 28.1</td>
<td>15.8 60.1 99.0</td>
<td>21.3 66.6 117.7</td>
<td>41.6 109.9 191.0</td>
</tr>
<tr>
<td>11</td>
<td>3.8 0.5 1.7</td>
<td>0.1 0.1 0.7</td>
<td>0.1 0.5 1.6</td>
<td>0.1 0.6 2.4</td>
<td>2.1 4.7 8.9</td>
</tr>
<tr>
<td>12</td>
<td>3.4 5.3 11.8</td>
<td>0.1 2.6 4.5</td>
<td>2.5 47.6 54.6</td>
<td>27.8 103.6 145.2</td>
<td>124.5 191.5 260.2</td>
</tr>
<tr>
<td>13</td>
<td>2.8 0.7 2.1</td>
<td>6.1 61.3 90.1</td>
<td>227.5 322.9 462.4</td>
<td>487.5 320.1 518.5</td>
<td>2.7 181.1 264.9</td>
</tr>
<tr>
<td>15</td>
<td>3.5 0.3 0.2</td>
<td>0.0 0.4 1.2</td>
<td>0.0 0.4 1.6</td>
<td>0.0 5.0 14.8</td>
<td>1.0 11.5 27.5</td>
</tr>
<tr>
<td>16</td>
<td>3.2 0.7 3.0</td>
<td>0.2 27.1 31.5</td>
<td>26.3 95.4 146.8</td>
<td>7.4 147.4 226.3</td>
<td>4.8 224.2 328.6</td>
</tr>
<tr>
<td>17</td>
<td>3.1 0.4 0.7</td>
<td>0.1 1.0 2.7</td>
<td>0.5 5.6 15.2</td>
<td>4.1 30.7 57.2</td>
<td>1.7 8.9 22.2</td>
</tr>
<tr>
<td>18</td>
<td>3.3 3.9 8.1</td>
<td>0.2 0.6 4.5</td>
<td>0.2 0.5 4.5</td>
<td>0.2 1.3 0.3</td>
<td>0.5 10.6 19.1</td>
</tr>
<tr>
<td>19</td>
<td>3.4 0.8 2.4</td>
<td>0.1 0.3 1.5</td>
<td>16.8 57.0 93.0</td>
<td>19.5 60.4 97.2</td>
<td>1.2 10.1 29.6</td>
</tr>
<tr>
<td>21</td>
<td>3.9 0.0 0.1</td>
<td>0.2 0.5 0.1</td>
<td>0.2 0.5 0.1</td>
<td>3.3 16.0 22.9</td>
<td>0.3 34.3 82.5</td>
</tr>
<tr>
<td>22</td>
<td>3.4 0.6 1.8</td>
<td>0.1 0.0 0.8</td>
<td>0.0 0.1 0.6</td>
<td>2.6 11.5 30.8</td>
<td>2.2 17.0 47.0</td>
</tr>
<tr>
<td>23</td>
<td>4.0 0.6 0.7</td>
<td>0.1 0.1 0.5</td>
<td>0.1 0.2 0.1</td>
<td>0.1 0.0 0.2</td>
<td>0.6 9.5 35.1</td>
</tr>
<tr>
<td>24</td>
<td>3.0 5.2 10.3</td>
<td>0.8 1.9 6.3</td>
<td>6.1 20.1 24.4</td>
<td>16.6 75.5 107.6</td>
<td>153.0 196.8 248.5</td>
</tr>
<tr>
<td>25</td>
<td>4.2 0.9 2.9</td>
<td>0.1 0.9 3.4</td>
<td>0.1 0.7 2.5</td>
<td>0.1 0.2 0.2</td>
<td>4.3 27.6 47.4</td>
</tr>
<tr>
<td>26</td>
<td>2.9 0.4 1.3</td>
<td>0.2 3.8 15.3</td>
<td>3.7 29.7 59.1</td>
<td>3.3 50.9 103.4</td>
<td>53.7 51.2 113.2</td>
</tr>
<tr>
<td>30</td>
<td>3.6 0.2 0.5</td>
<td>0.2 0.6 0.2</td>
<td>0.2 0.5 0.1</td>
<td>0.2 0.4 0.0</td>
<td>1.9 4.6 13.9</td>
</tr>
<tr>
<td>31</td>
<td>3.4 1.1 2.3</td>
<td>0.0 0.5 0.1</td>
<td>0.0 0.6 0.7</td>
<td>0.0 1.0 1.6</td>
<td>1.5 6.4 14.4</td>
</tr>
<tr>
<td>32</td>
<td>2.8 1.8 7.8</td>
<td>0.1 1.3 5.9</td>
<td>3.0 42.2 73.7</td>
<td>5.6 96.6 156.9</td>
<td>0.8 28.0 61.3</td>
</tr>
<tr>
<td>33</td>
<td>3.6 1.0 2.6</td>
<td>0.2 0.5 0.3</td>
<td>0.3 1.3 2.2</td>
<td>0.3 6.4 14.8</td>
<td>0.8 30.0 60.6</td>
</tr>
<tr>
<td>34</td>
<td>3.8 1.3 2.9</td>
<td>0.9 1.2 1.0</td>
<td>0.9 2.2 1.7</td>
<td>0.2 1.3 2.7</td>
<td>1.8 6.9 16.6</td>
</tr>
<tr>
<td>35</td>
<td>2.8 0.1 0.4</td>
<td>0.1 3.4 10.7</td>
<td>1.5 9.4 17.0</td>
<td>0.0 37.1 72.9</td>
<td>3.5 43.3 81.6</td>
</tr>
<tr>
<td>36</td>
<td>4.6 2.6 5.4</td>
<td>0.1 2.7 5.9</td>
<td>0.0 2.0 3.4</td>
<td>4.3 11.7 20.6</td>
<td>4.8 14.4 25.2</td>
</tr>
<tr>
<td>37</td>
<td>2.8 2.0 8.5</td>
<td>0.1 8.7 25.2</td>
<td>1.5 20.6 41.8</td>
<td>1.5 53.2 95.0</td>
<td>6.0 88.9 155.0</td>
</tr>
<tr>
<td>38</td>
<td>4.2 2.4 5.8</td>
<td>0.1 1.2 4.0</td>
<td>0.4 1.1 3.5</td>
<td>0.4 1.1 3.7</td>
<td>0.2 0.1 1.4</td>
</tr>
<tr>
<td>39</td>
<td>4.8 0.9 4.8</td>
<td>0.1 0.0 0.3</td>
<td>0.1 0.6 3.3</td>
<td>0.1 0.9 5.6</td>
<td>2.5 6.0 22.7</td>
</tr>
<tr>
<td>41</td>
<td>3.6 49.5 83.6</td>
<td>0.0 0.4 2.0</td>
<td>0.0 4.5 14.8</td>
<td>2.4 23.0 43.3</td>
<td>0.3 17.5 40.5</td>
</tr>
<tr>
<td>42</td>
<td>3.4 1.1 3.3</td>
<td>0.1 0.3 0.5</td>
<td>0.0 3.6 13.9</td>
<td>0.0 4.4 16.9</td>
<td>5.0 83.8 160.6</td>
</tr>
<tr>
<td>43</td>
<td>3.7 0.3 0.2</td>
<td>0.9 2.4 5.1</td>
<td>0.9 2.1 4.5</td>
<td>0.9 1.9 4.4</td>
<td>3.2 27.3 68.6</td>
</tr>
<tr>
<td>45</td>
<td>3.1 0.0 0.6</td>
<td>0.9 9.3 32.8</td>
<td>0.0 1.3 5.4</td>
<td>1.9 25.2 61.2</td>
<td>2.9 31.4 75.4</td>
</tr>
<tr>
<td>46</td>
<td>3.2 0.0 0.3</td>
<td>0.5 1.7 1.4</td>
<td>0.5 1.8 1.4</td>
<td>6.4 48.7 77.3</td>
<td>0.5 7.5 18.0</td>
</tr>
<tr>
<td>48</td>
<td>3.0 0.1 0.2</td>
<td>0.2 9.0 30.1</td>
<td>3.4 25.4 56.1</td>
<td>7.0 55.3 113.8</td>
<td>44.1 224.1 356.5</td>
</tr>
<tr>
<td>49</td>
<td>2.9 0.6 1.2</td>
<td>0.1 1.9 4.2</td>
<td>0.1 1.2 1.9</td>
<td>1.9 14.8 27.9</td>
<td>1.6 25.3 49.7</td>
</tr>
</tbody>
</table>

Mean: 3.5 2.5 5.4
Mean: 0.5 8.0 13.5

Std. Dev.: 0.5 4.6 9.4
Std. Dev.: 1.2 10.9 16.7

Table VI – Absolute error of estimated HRV parameters using VPG and PPG compared to ECG
Fig. 21 presents a summary of the data presented on Table I. The figure presents the Mean (blue) and SD (orange) from estimation error. Note that Hue of the Basler camera provides comparable performance to that of the PPG sensor while clearly outperforming Green. The webcam offers much poorer performance than the Basler camera for both Green and Hue.

When just considering Hue-Basler and PPG, it seems that both offer a reasonable alternative to ECG. While the PPG average absolute error for meanRR, SDNN and RMSSD is 3.5, 2.5 and 5.4 milliseconds respectively, the corresponding Hue-Basler errors are 0.5, 4.6 and 9.4 milliseconds respectively.
4.1.2 Ectopic beats in Healthy subjects

A total of 32 ectopic or irregular beats were identified upon examination of the ECG signals recorded from 5 different subjects in the healthy cohort. Such heartbeats produce an irregular hemodynamic response based on the nature of the contraction that triggered the beat.

Figs. 22-23 illustrates two scenarios in which an ectopic beat is present in the ECG signal. The ectopic beat presented in Fig. 22 is visible in the VPG signal in the same way is perceived by the contact PPG sensor. However, the ectopic beat presented in Fig. 23 did not produce a hemodynamic response that could be captured by the contact PPG sensor, nor the VPG system.

30 (93.75%) of the identified ectopic beats produced a hemodynamic response that was visible in the Hue signal obtained through the Basler camera and the PPG signal from the earlobe sensor.
While ectopic beats are mostly asymptomatic and usually have no clinical significance, there are some cases in which the occurrence of ectopic beats can indicate susceptibility towards abnormal heart rhythms, such as ventricular tachycardia, especially if the patient has a heart disease [54].

Fig. 22 – Ectopic beat with visible hemodynamic response

Fig. 23 – Ectopic beat with unnoticeable hemodynamic response
4.1.3 Patients with Atrial Fibrillation

Even though peak detection is more challenging in the presence of atrial fibrillation, we were able to obtain accurate IBI measurements from some subjects, thanks to the high signal to noise ratio obtained by using the Basler camera and the hue approach. We will use the data extracted from two atrial fibrillation subjects to illustrate the performance obtained by estimating IBIs in the presence of atrial fibrillation.

Fig. 24 shows an example of a subject in which the IBIs were accurately detected. In this figure, the IBIs measured from Hue-Basler and the IBIs measured from the ECG Holter are simultaneously plotted. Notice how the IBIs measured from Hue-Basler (plotted in blue) show from minimal to no deviation from the IBIs measured from the ECG Holter (plotted in red).

Unlike in Fig. 24, the data presented in Fig. 25 shows an example of a subject in which the IBIs were poorly detected. When compared to the IBIs from the ECG Holter (plotted in red), The IBIs obtained via Hue-Basler (plotted in blue) indicate that there were multiple instances in which a beat was not detected. These instances are marked with black arrows.
Fig. 24 – IBIs as a function of beat count (accurate detection).

Fig. 25 – IBIs as a function of beat count (inaccurate detection).
4.2 Atrial fibrillation detection using VPG

Using the methods described in Chapter 2, we computed the IBI’s and occupied bandwidth for all the captures (healthy cohort, and atrial fibrillation cohort). The atrial fibrillation diagnose was then obtained based on the comparison of the value to its corresponding threshold, where it would be said that a given trace was AF if the exceeded its corresponding threshold: \( \frac{\text{RMSSD}}{\text{meanRR}} \geq \text{th}\_\text{rmssd} \), Bandwidth occupied by 30% of the signal’s power \( \geq \text{th}\_\text{obw} \). Using this criteria, we found the number of true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN). This values were used to compute the specificity \( \frac{\text{TN}}{\text{TN}+\text{FP}} \), sensitivity \( \frac{\text{TP}}{\text{TP}+\text{FN}} \), and accuracy \( \frac{\text{TP}+\text{TN}}{\text{TP}+\text{TN}+\text{FP}+\text{FN}} \).

Tables VII-X below summarize AF detection results for the entire AF and healthy subject cohort. We excluded the data from a total of four subjects as the PPG data from the earlobe sensor was not available for these captures. The accuracy, sensitivity and specificity of AF detection based on the RMSSD method applied to the detected IBIs are presented in tables III and IV. We investigate RMSSD for the Red (R), Green (G), Blue (B), Hue (H) traces and ICA method all from both cameras. The corresponding results using the Occupied Bandwidth method follow in tables V and VI.
### Table VII – Classification results using normalized RMSSD method

<table>
<thead>
<tr>
<th></th>
<th>Logitech webcam</th>
<th></th>
<th></th>
<th></th>
<th>ICA</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Accuracy</strong></td>
<td>40.78%</td>
<td>65.05%</td>
<td>33.01%</td>
<td>67.96%</td>
<td>60.19%</td>
</tr>
<tr>
<td><strong>Sensitivity</strong></td>
<td>100.00%</td>
<td>100.00%</td>
<td>100.00%</td>
<td>100.00%</td>
<td>100.00%</td>
</tr>
<tr>
<td><strong>Specificity</strong></td>
<td>11.59%</td>
<td>47.83%</td>
<td>0.00%</td>
<td>52.17%</td>
<td>40.58%</td>
</tr>
</tbody>
</table>

### Table VIII – Classification results using normalized RMSSD method

<table>
<thead>
<tr>
<th></th>
<th>Basler ace</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Accuracy</strong></td>
<td>46.60%</td>
<td>66.99%</td>
<td>33.01%</td>
<td>81.55%</td>
<td>73.79%</td>
</tr>
<tr>
<td><strong>Sensitivity</strong></td>
<td>100.00%</td>
<td>100.00%</td>
<td>100.00%</td>
<td>100.00%</td>
<td>100.00%</td>
</tr>
<tr>
<td><strong>Specificity</strong></td>
<td>20.29%</td>
<td>50.72%</td>
<td>0.00%</td>
<td>72.46%</td>
<td>63.87%</td>
</tr>
</tbody>
</table>

### Table IX – Classification results using occupied bandwidth method

<table>
<thead>
<tr>
<th></th>
<th>Logitech webcam</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Accuracy</strong></td>
<td>63.11%</td>
<td>85.44%</td>
<td>41.75%</td>
<td>87.38%</td>
<td>83.50%</td>
</tr>
<tr>
<td><strong>Sensitivity</strong></td>
<td>100.00%</td>
<td>97.06%</td>
<td>100.00%</td>
<td>100.00%</td>
<td>100.00%</td>
</tr>
<tr>
<td><strong>Specificity</strong></td>
<td>44.93%</td>
<td>79.71%</td>
<td>13.04%</td>
<td>81.16%</td>
<td>75.36%</td>
</tr>
</tbody>
</table>

### Table X – Classification results using occupied bandwidth method

<table>
<thead>
<tr>
<th></th>
<th>Basler ace</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Accuracy</strong></td>
<td>62.14%</td>
<td>86.41%</td>
<td>44.66%</td>
<td>95.15%</td>
<td>92.23%</td>
</tr>
<tr>
<td><strong>Sensitivity</strong></td>
<td>100.00%</td>
<td>100.00%</td>
<td>100.00%</td>
<td>100.00%</td>
<td>100.00%</td>
</tr>
<tr>
<td><strong>Specificity</strong></td>
<td>43.48%</td>
<td>79.71%</td>
<td>17.39%</td>
<td>92.75%</td>
<td>88.41%</td>
</tr>
</tbody>
</table>
The RMSSD method, which is derived from IBIs, was applied to the IBI data extracted from the earlobe PPG sensor, and the ECG Holter recorder. The PPG sensor provided 85.44% accuracy, and 100% sensitivity, for 78.26% specificity. Similarly, the ECG Holter provided 86.41% accuracy, and 100% sensitivity, for 79.71% specificity.

Since the PPG and VPG waveforms share the same morphology, we applied our occupied bandwidth classification method to the PPG signal. Using this method, the PPG sensor provided 94.17% accuracy, and 97.06% sensitivity, for 92.75% specificity.

4.2.1 ROC curves

Receiver Operating Characteristic (ROC) curves for atrial fibrillation detection using the RMSSD method and the occupied bandwidth method were generated for Green, Hue and ICA using both Webcam and Basler. Each ROC curve was generated by performing 100 runs over the entire database using different values of the detection threshold. The range of values evaluated for RMSSD goes from 0.1 to 1, and the range of values for the occupied bandwidth method goes from 10 mHz to 1 Hz.

These graphs confirm the superiority of Hue, when compared to Green and ICA methods. Furthermore, the performance gap between the webcam and the Basler camera is further accentuated upon examination of these plots. Additionally, the graphs show that the occupied bandwidth approach outperformed its IBI based counterparts on both cameras.
Fig. 26 – ROC Logitech webcam RMSSD.

Fig. 27 – ROC Basler ace RMSSD.
Fig. 29 – ROC Logitech webcam occupied bandwidth.

Fig. 28 – ROC Basler ace occupied bandwidth.
4.3 Image sensor comparison

Table II summarizes the most relevant specifications from the image sensors in the cameras used for the study. The performance gap between the Webcam and the Basler Ace camera becomes evident upon examination of the results we have presented. The Basler Ace camera consistently provides a better result than the Logitech webcam. The Basler camera uses a Sony (Sony Corp., Tokyo, Japan) IMX174 sensor, while the Logitech webcam is presumed to use a sensor equivalent to the ON Semiconductor (ON Semiconductor, Phoenix, AZ) MT9D131.

Despite the fact that the Basler camera surpasses the Webcam on all the specifications presented on Table II, it could be inferred that the improved accuracy obtained by the Basler camera can be attributed to its larger pixel size and better quantum efficiency for the 525nm wavelength, which falls within the previously noted optimal range to capture blood volume pulsations.

Table XI - Critical image sensor parameter comparison

<table>
<thead>
<tr>
<th>Critical Image Sensor Parameter Comparison</th>
<th>Basler ACE</th>
<th>Webcam</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dynamic range [dB]</td>
<td>73.07</td>
<td>71</td>
</tr>
<tr>
<td>SNRmax [dB]</td>
<td>45.15</td>
<td>42.3</td>
</tr>
<tr>
<td>Pixel size [μm]</td>
<td>5.86x5.86</td>
<td>2.8x2.8</td>
</tr>
<tr>
<td>Full resolution [pixel]</td>
<td>1920x1200</td>
<td>1600x1200</td>
</tr>
<tr>
<td>Quantum Efficiency @525nm</td>
<td>74%</td>
<td>37.50%</td>
</tr>
</tbody>
</table>
Chapter 5. Conclusions

The use of VPG as a surrogate to ECG and PPG for cardiac monitoring should not compromise the accuracy of the measurement. The results presented in this study demonstrate that our VPG approach, which encompasses an efficient capture and the use of Hue, provides for an accurate depiction of cardiovascular activity, as we obtained an average error below 10 milliseconds on the estimation of meanRR, SDNN, and RMSSD using IBI detection on normal sinus rhythm. Additionally, we introduced a new heart variability indicator, designed for plethysmographic signals, that can be obtained without performing peak detection and serves as a classifier to identify atrial fibrillation episodes. We obtained above 90% accuracy, sensitivity, and specificity with a combination of the proposed VPG approach and occupied bandwidth classifier.
References


[40] [Basler ace stock photo]. Retrieved June 01, 2016, from: http://www.automacionenews.it/nuove-telecamere-basler-ace-con-interfaccia-usb-3-0/


Acknowledgements

I can’t stress enough how grateful I am for the guidance Dr. Gill Tsouri has given me. He believed even more than I believed in myself, and the completion of this work is something couldn’t have done without his help. He raised the bar of “teaching” as he did much more than that.

I am thankful to Dr. Jean Phillipe Couderc and the whole Heart Research Follow-up at the University of Rochester. Your accessibility and collaboration made me feel like part of the team every step of the way.

I would like to thank Diane Ellison for making it possible for me to concentrate on completing this work. Her encouragement and assistance was of big help.

I would also like to thank my family and friends for their support. For some reason they all have this superhuman impression of me, and think I can do all kinds of things.

Last but not least, I want to express my gratitude towards all the members of the Electrical Engineering department at RIT for pushing me to the next level.