12-2020

Camera-Based Remote Photoplethysmography for Estimation of Heart Rate using Single Board Computers

Benjamin Sweely
University of Tennessee, Knoxville, bsweely@vols.utk.edu

To the Graduate Council:

I am submitting herewith a thesis written by Benjamin Sweely entitled "Camera-Based Remote Photoplethysmography for Estimation of Heart Rate using Single Board Computers." I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Mechanical Engineering.

Xiaopeng Zhao, Major Professor

We have read this thesis and recommend its acceptance:

Xiaopeng Zhao, Tami H. Wyatt, Jeffrey A. Reinbolt

Accepted for the Council:
Dixie L. Thompson

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Camera-Based Remote Photoplethysmography for Estimation of Heart Rate Using Single Board Computers

A Thesis Presented for the
Master of Science
Degree
The University of Tennessee, Knoxville

Benjamin Sweely
December 2020

ACKNOWLEDGEMENTS

Thank you to Dr. Xiaopeng Zhao, Mr. Soheil Borhani, and my thesis committee members for your guidance and support. Additionally, thank you to Dr. Zhao for believing in me even when it was difficult. It was a great pleasure working with the students of the 2020 Summer Undergraduate Research Group Experience (SURGE) program. Thanks for your support and collaboration. Special thanks go to Austin Park for his technical support. And, of course a big thank you to my friends, family, and girlfriend for their help and support.

ABSTRACT

The objective of this project was to develop a wireless, noncontact monitoring system that measures multiple physiological parameters in human faces from a distance using a camera. Compared to traditional physiological sensors, this monitoring system does not use wires or adhesives, thus providing a safer, more user-friendly application. As an illustration of the monitoring systems, experiments were conducted to estimate heart rate (HR). The current practices of measuring HR involve collecting electrocardiogram (ECG) signals from adhesive electrodes placed on various parts of the body or using a pulse oximeter (PO) typically placed on the ear lobe or finger. We developed 2 monitoring systems and compared their results to that from a PO. Both monitoring systems are low-cost at less than $200. Neither system has been shown to exist in literature thus making them novel implementations. In conclusion, we were able to estimate HR from a distance using a camera-based system. The developed systems may have many useful applications, in both clinical and home health settings.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION

The ability to measure heart rate (HR) noninvasively is important in both a hospital and a home setting due to the role this vital sign plays in the diagnosis of health and wellbeing. There have been many studies (e.g., Wieringa et al. 2005; Humphreys et al. 2007; Verkruysse et al. 2008; Zheng et al. 2008; Poh et al. 2010; Poh et al. 2011; Kamshilin et al. 2011; Holton et al. 2013) showing the measurement of this parameter using noncontact methods. This type of technology has the potential to change the healthcare industry and bring more utility to the home setting.

Despite great advancements and improvements in recent years, safety remains a challenging issue in neonatal intensive care unit (NICU). Traditional sensors found in NICU incubators involve the use of adhesives and wires; see Figure 1. Wires can become tangled and cut off circulation, causing potential harm to the infant. Adhesives, widely used to secure electrocardiogram (ECG) leads, temperature probes and/or medical equipment, can damage the fragile skin of a newborn, often leaving scars (Johnson 2016). Approximately 10 to 15% of neonatal graduates leave with physical scars, which can contribute to later body image issues (Boyar 2019). These adhesives also have the ability to pull off the nipple of an infant, which will never grow back leaving the infant forever physically altered. The epidermis of the premature infant is 3 layers thick compared to the adult epidermis which is 20 layers. Additionally, the spaces between the layers of epidermis are fluid filled without tight connections, making tissue disruption common. Preterm infants' skin is so translucent that one can see through it and see all the veins. The skin will remain severely thin until about 2 weeks after birth (Oranges et al. 2015); therefore, infants are extremely susceptible to skin damage during this time. Tissue damage causes pain and disrupts the infant's barrier defense against infections. Wires and traditional sensors may limit the amount of physical contact between the infant and its parents, which can lead to issues with bonding (Chung et al. 2019). Studies have shown that neonates that receive parental touch have displayed stronger vital signs and lower mortality rates (Smith 2012). While infants are more likely to experience a rise in vital statistics when skin to skin parent contact is possible, current technologies often do not take this finding into consideration (Morris et al. 2019; Chung et al. 2019).

Figure 1.1 Infants in NICU often wear multiple sensors, which are attached to their bodies using adhesives; figure copied from (About CDH).

1.1 Prior Art

Over the last decade, considerable research effort has gone into development of non-contact camera-based photoplethysmography (PPG) measurement system (e.g., Wieringa et al. 2005; Humphreys et al. 2007; Verkruysse et al. 2008; Zheng et al. 2008; Poh et al. 2010; Poh et al. 2011; Kamshilin et al. 2011; Holton et al. 2013). Initially, external arrays of LEDs at different wavelength (red, infrared) were used to illuminate a region of tissue for measuring HR and peripheral oxygen saturation (SpO2) using a monochrome CMOS camera (Wieringa et al. 2005; Humphreys et al. 2007; Zheng et al. 2008). Good correspondence for PR estimation has been established quantitatively, but not much evidence was provided for reliable SpO2 measurements. Later, Verkruysse et al. (2008) showed that HR and respiratory rate (RR) can be determined using simply a color camera and ambient illumination. The authors reported that the green channel of the RGB camera gives the best results for detecting PR and BR. More recently, Poh et al. (2010; 2011) used a webcam under ambient illumination to detect simultaneous PR, heart rate variability (HRV) and RR of multiple people in a video by employing independent component analysis (ICA) on three color channels (R,G,B) to extract the PPG signal. This approach was further investigated in Lewandowska et al.

(2011), where they claimed that principal component analysis (PCA) of the three-color channel can be as effective as ICA.

Most of the past work, however, did not report how the system performs on individuals having different skin tones. It is well-known that the melanin present in darker skin tones absorbs a lot of light, and thus attenuates the SNR of the acquired PPG signal, making the system ineffective for extracting vital signs.

To counter the motion artifact challenge, Poh et al. (2010) used automatic face detection in consecutive frames to track the face. In Sun et al. (2011), the authors computed 2D shifts between consecutive frames using image correlation to model the motion. Both these approaches (Sun et al. 2011; Poh et al. 2010) could only captures translation motion of face, and do not cater to more natural motions like tilting of the face, smiling or talking generally found in in-situ scenarios.

The results obtained from ICA and PCA show that those two methods extract the "pulse" component with similar accuracy. The obtained results show that when only HR is considered there is no need to use the more computationally complex ICA method. The accuracy of extraction of cardiac pulse signal by PCA is comparable to that obtained by ICA and is sufficient for our purposes. Moreover, the time of calculation obtained for PCA suggests possibility of using this approach in real-time applications (Lewandowska et al. 2011). Sun et al. (2012) compared the results from a low-cost webcam and an expensive CMOS camera and showed the results were comparable.

1.2 Main Contributions

This thesis addresses 2 challenges in acquiring HR using a camera: 1) high cost of current remote monitoring systems, and 2) lack of flexibility to implement the camera system in different environments to meet specific needs. We developed 2 camera-based HR monitoring systems using Raspberry Pi (Pi) and Jetson Nano (Nano), respectively. Both systems can estimate the PPG signal in-situ using a camera and are under $200. In terms of usability, both systems can be setup in a variety of environments due to their small size. For both systems, all that is needed is the camera and the computer board.

The contribution of the work presented here is significant because it provides a user-friendly, low-cost approach to recording HR. Noncontact monitoring has the potential to increase safety, efficiency, and convenience in NICU. There have been many recent advances in noncontact monitoring. Most of these advancements have come in the form of better motion tracking and better artifact removal. This project has implemented these state-of-the-art techniques and algorithms to estimate HR. The focus of the system is to be more affordable and user friendly. Our system utilizes low-cost equipment and software to ensure the product is accessible to as many health practitioners and educators as possible. A Raspberry Pi module that costs $60 and a Jetson Nano that costs $100 were used in the

making of the 2 systems. Attached to both is a Raspberry Pi camera which comes in at $30. Therefore, both systems can be produced for less than $200. Due to its open source and low cost, our system is easily scalable and adaptable for further health innovation. The systems are suitable to be used in telemedicine applications and as an education tool (Villarroel et al. 2019).

The future goal of the monitoring systems in this project is to eliminate potential risks associated with wires and adhesives. It is hoped that these created systems can solve some of these issues by eliminating some of the dangerous wires and adhesives through non-contact monitoring techniques. The systems incorporate machine learning algorithms to detect faces quickly and accurately. This contribution is significant, because it provides an affordable and user-friendly approach to recording HR compared to its predecessors. These devices may be used as a framework in future research to aid in early detection of diseases such as sepsis in NICU infants. For example, the accurate and timely diagnosis of apnea and sepsis remains a major issue in neonates in the NICU. These issues exponentially increase medical costs in our country. The systems in this proposal can act as prototypes to be used for more advanced systems, providing quicker and more accurate diagnosis of sepsis in NICU infants. The systems can run automatically when turned on and output the estimated HRs to a connected computer via WiFi connection.

CHAPTER TWO

BACKGROUND

2.1 PPG Signal

Photoplethysmography (PPG) is a commonly used, simple, and low-cost optical technique that is applied at the skin surface to assess blood volume changes in the microvascular bed of tissue (Allen 2007). These blood volume changes are assessed through a photodetector that detects perfusion by means of small variations in light intensity found through using a light source, often in the red or infrared spectrum, to illuminate the tissue (Allen 2007; Wang et al. 2018). The volume of blood in the arteries and capillaries changes in concurrence with the cardiac cycle, which leads to small changes in skin color (Kumar et al. 2015). Similarly, respirations also cause skin color variations (van Gastel et al. 2016). Even the smallest blood vessels of the skin have blood volume variations that change the absorption of light, which causes reflected or transmitted light to be captured by the PPG, resulting in a PPG waveform (van Gastel et al. 2016). PPG waveforms, representing changes in skin color, can be utilized to detect vital signs such as RR, blood pressure, HR, and HRV (Kumar et al. 2015).

While these subtle color changes are invisible to human eyes, they can be seen and recorded through digital cameras (Wang et al. 2018). However, the interaction of light and the biological tissue is complex and includes the optical processes of fluorescence, transmission, absorption, reflection, and multiple scattering (Allen 2007). Although blood absorbs more light than the surrounding tissues (Wang et al. 2018), the amount of light measured by the photodetector can also be influenced by the placement of red blood cells, blood volume, and movement of the blood vessel wall (Allen 2007). While PPG has the benefit of measuring heart rate and respiratory rate concurrently, the derived PPG respiratory rate depends on a regular breathing pattern and normal circulation (Harford et al. 2019). Two major challenges in using PPG accurately include motion artifacts from an individual moving in front of the camera as well as low signal strength of the color change, especially in low light conditions and darker skin tones (Kumar et al. 2015).

2.2 History of PPG

The first use of PPG dates back to 1936, when two research groups used the technology in their study of rabbit ears. Molitor and Kniazuk (1936) and Hanzlik et al. (1936) “described similar instruments used to monitor the blood volume changes in the rabbit ear following venous occlusion and with administration of vasoactive drugs” (Allen 2007). Molitor and Kniazuk furthered the advancement of PPG by beginning research into human use of the technology with a reflection

mode PPG system that was tested on human fingers. The PPG technique was established by Hertzman and Dillon (1940a), who split the AC and DC components with separate electronic amplifiers and monitored vasomotor activity. Hertzman and colleagues were the first to use PPG technology to measure blood volume changes in the fingers induced by the Valsalva maneuver, exercise and with exposure to cold (Allen 2007). This paper was the first research that proved that the PPG technology has the ability to be used in the clinical setting.

Recent advancement in PPG technology has been motivated by the need for low cost, user friendly, and non-invasive methods of assessment. These needs have driven the re-establishment of PPG (Allen 2007). Another factor that has driven recent advancement in PPG is the development of new technology to supplement PPG. Advances in technology include development in opto-electronics and clinical instrumentation, as well as development in semiconductor technology. A major advance in the clinical use of a PPG-based technology came with the introduction of the pulse oximeter as a non-invasive method for monitoring patients' arterial oxygen saturation (Aoyagi et al 1974; Yoshiya et al 1980). There have also been considerable developments in computer-based digital signal processing and pulse wave analysis. "Semiconductor technology such as light emitting diodes, photodiodes and phototransistors have made improvements to the sensitivity and reliability of the PPG probe design" (Allen 2007).

2.3 PPG Waveform Features

Recent studies have emphasized the need for more thorough analysis of the PPG waveform to be made beyond its current applications of pulse oximetry and heart rate (Elgendi 2012). Furthermore, the pulsatile nature of the blood pulse waveform is not yet fully understood (Amelard et al. 2016), and errors in the PPG waveform due to insufficient light conditions and motion is present in a much greater magnitude in a camera-based system than a contact pulse oximeter (Kumar et al., 2015). Moreover, the PPG waveform is, in general, vulnerable to sudden amplitude changes by the automatic gain controller, which changes the gain of the amplifier based on the amplitude of the input signal (Elgendi 2012). Amelard et al. (2016) described a typical blood pulse waveform signal as having a quasi-periodic nature with dicrotic characteristics and periodicity resulting in harmonic frequencies in the power spectral density. The blood volume waveform generally comprises of a pulsatile ('AC') physiological waveform attributed to synchronous changes in the blood volume with each heartbeat, and is superimposed on a slow altering ('DC') baseline with differing lower frequency components attributed to respiration, sympathetic nervous system activity, and thermoregulation (Allen 2007; Burton 1939; Burton and Taylor 1940; Hertzman and Dillon 1940b; Hertzman and Roth 1942a; 1942b; 1942c; Hertzman and Flath 1963; Hyndman et al 1971; Pen˜az´ 1978; Ahmed et al 1982; Harness and Marjanovic 1989; Nitzan et al 1996b; 1996a;

Shusterman et al. 1997; Schultz-Ehrenburg and Blazek 2001; Nitzan et al. 2001). With correct filtering the AC component can be extracted for pulse wave analysis.

PPG waveform has broad clinical applications. Through the use of the PPG signals, it has been possible to notice several vascular stiffness and aging indices, which to be indicators of related cardiovascular variables (Elgendi 2012). Additionally, PPG has been used to create a blood pulse waveform signal from physiological information. This PPG waveform was able to detect a cardiac arrhythmia, using the visual waveform of PPG (Amelard et al. 2016). This remarkable discovery shows the ability of the PPG waveform, and the possible uses that it presents to the clinical field.

Lewandowska et al. (2011) found that the selected R and G channels contained most of the information about color changes corresponding to the blood volume pulse, which suggests the possibility to reduce the number of analyzed signals from three to two. Decrease in ROI's size or number of channels increases the level of noise. However, taking into account that these measurements are destined for daily monitoring of vital signs not for clinical purposes, the high accuracy is not the most important factor.

CHAPTER THREE

CAMERA BASED PPG

This chapter outlines how the monitoring systems were created and function. The 2 systems differ in Hardware and Software but have similar theory behind their approaches. The following section lays out the general pipeline used in both systems. The other sections lay out how each system differ in the details of their pipelines. There were two strategies employed to achieve the desired goal. The hardware and software differ greatly between the two, but the general methodology remains the same. One monitoring system was created using a Raspberry Pi programmed in Matlab. The other system used a Jetson Nano programmed in Python. A main contribution of the experiment was to test the functionality of both systems.

3.1 General System

In this section, a general pipeline is described for both monitoring systems. Specific details on the algorithms and implementations will be mentioned in each system's sections.

PPG Pipeline

Each system starts with face detection. After a face is detected, the system will lock onto a relevant region-of-interest (ROI). For every ROI, 3 channels are extracted (Red, Green, Blue) and the intensity of the pixel values were calculated for each frame. This is done by summing all the pixel values for each color channel in the frame. The 3 color channels are then detrended and normalized to remove any linear trend along the axis, overcome nonstationarity from the signal, and make them optimal for the feature selection algorithm.

Feature selection can be done in a few ways. Firstly, a color channel can be chosen to represent the PPG signal. This is typically the Green channel because the wavelength of green is highly absorbed in the hemoglobin of the blood. Other typical feature selection algorithms include PCA and ICA. They both create new components that are linear combinations of the 3 color channels. The source that has the highest power in the relevant frequency range is passed on. PCA is faster and ICA is said to be more accurate (Lewandowska et al. 2011). A study by Lewandowska et al. found that the results using the two approaches are similar, suggesting that choosing PCA is a better choice since it consumes less computational power and time (Lewandowska et al. 2011). These signals then passed to the HR estimation pipeline. This process is shown in Figure 3.1 below.

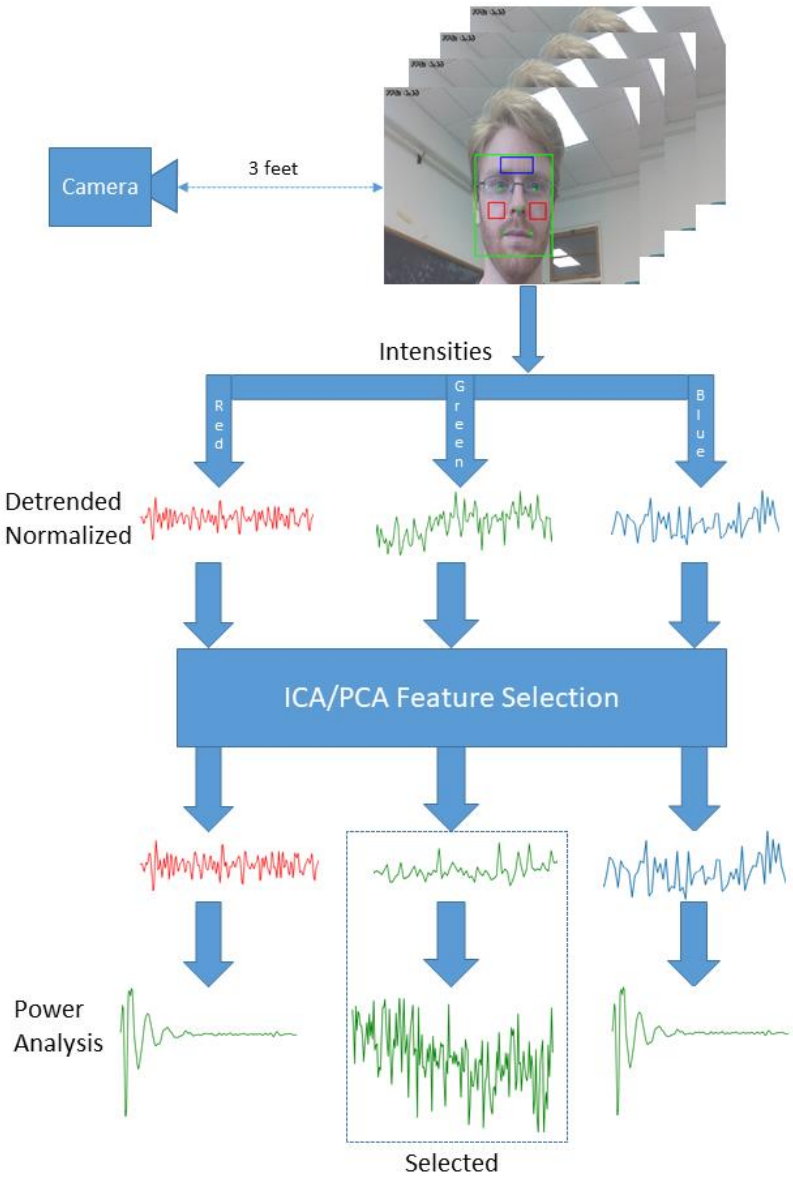


Figure 3.1 shows the general PPG estimation pipeline used in both systems.

HR Pipeline

The signal was then averaged on a sliding window of 5 frames. The signal was then filtered using a bandpass filter with the range 0.8 to 3 Hz. This range of frequency was selected based on the literature. This frequency range captures any signal that correlates to a HR from 48 to 180 beats per minute (BPM).

The average HR was calculated by finding the frequency at which the peak power occurs in the PSD and multiplying that frequency by 60 seconds to give BPM. The typical window size for the number of frames was varied but most success was found around 300 frames which at a frame rate of 15 frames per second (fps) equates to 20 seconds worth of data. This is enough data to accurately estimate the HR from what has been observed in previous studies and pulse oximeters. This estimated HR is the average HR during the time interval of 20 seconds.

3.2 Raspberry Pi with Matlab

3.2.1 PPG Estimation Pipeline

The face detection model uses the Viola-Jones algorithm (Viola and Jones 2001). This common object detection algorithm uses edge detection to find the contours of the objects. The new features this algorithm brought was quicker detection, the use of AdaBoost (Freund and Schapire 1997) as the classifier, and the use of a cascade to quickly filter out insignificant regions (Viola and Jones 2001). Custom ROI detection was developed to lock on to the forehead and cheeks. The code that implements these algorithms is listed in the Appendix.

Next, the feature selection algorithm ICA was performed on the 3 color channels (red, green, blue) to find the independent combinations. ICA was performed using the Matlab function called kurtosis ICA (kICA). During the kICA process, the data is whitened to make the unmixing matrix always orthogonal. Next the fourth order statistics are mapped down into a matrix such that the inner product of a unit vector with this matrix gives the kurtosis in that direction. The maximal eigenvectors of this new matrix are the directions of maximum kurtosis, which give a good estimate for the unmixing directions.

Power spectral density (PSD) was then performed on each signal by taking Fast Fourier Transform (FFT) of the signals and then calculating the power of the transformed signals. This power was plotted over the frequency range. The signal that had the highest peak in power over the 0.8 to 3 Hz frequency range was selected as the best feature.

3.2.2 HR Estimation Pipeline

Continuing with the PPG signal from the last pipeline, the signal was run through a bandpass filter (type = butterworth, order = 5) passing the signal in the frequency range of 0.8 to 3 Hz. PSD of the filtered signal was performed as described above.

The max peak of the PSD was detected which indicated the frequency at the highest power of the signal. This frequency was multiplied by 60 which gave the estimated average HR over the 300 frames passed into the algorithm.

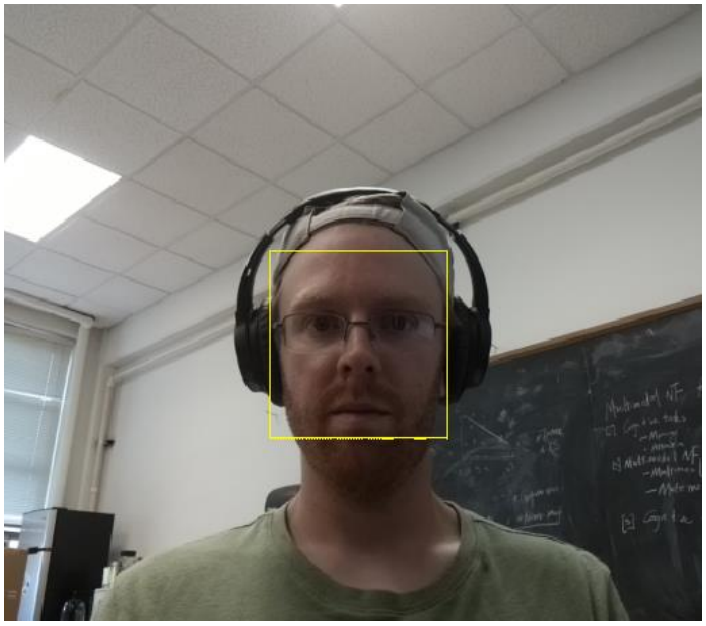
3.3 Jetson Nano with Python

3.3.1 PPG Estimation Pipeline

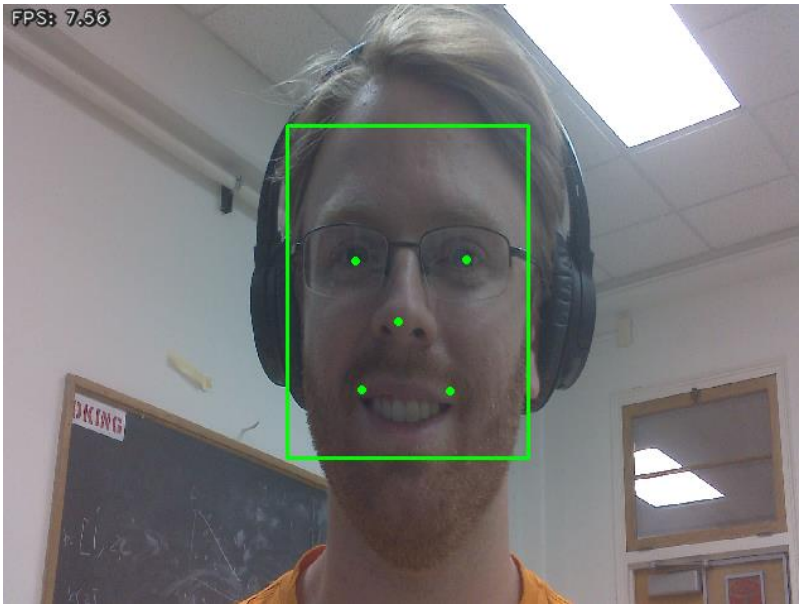
The PPG estimation pipeline begins with face detection. The Multi-task Cascaded Convolutional Networks (MTCNN) face detection model was used and contains 3 stages of neural networks that each improve the predicting power (Zhang et al. 2016). The image is initially resized to the correct format. Stage 1 is a Proposal Network (P-Net) that gathers the facial windows and the corresponding vectors for the bounding box. This is followed by employing a non-maximum suppression (NMS) to combine all candidates that overlap. In Stage 2, the candidates are inputted into a Refine Network (R-Net), which rejects false candidates, calibrates, and conducts NMS. Finally, Stage 3 aims to detect specific face regions and outputs landmarks of these regions. These include the eyes, nose and ends of the mouth as seen in Figure 3.2.

The created ROI algorithm uses the outputted landmarks of the MTCNN algorithm to find the coordinates of the cheeks and forehead. The ROI algorithm then detected when the cheeks and forehead were not visible. This was done to preserve the PPG signal even if the participant's face was turned. Albeit, our model still has issues with the cheek ROI's if the face is turned. This will be discussed more in the limitations section. The Python code that uses these algorithms is shown in the Appendix.

The 3 color channels were extracted from the ROI. Their intensity was calculated by summing up all the pixel values in each color channel. Therefore, each channel had one value per frame. These 3 channels were then detrended and normalized. Detrending takes away any slope in the data and normalizing makes the mean zero and each standard deviation away from the mean equal to a whole number (1, 2, etc.). ICA was run on the 3 signals to create 3 separate sources. We used an ICA algorithm known as FastICA, available in the scikit-learn 0.23.1 toolbox. To find the most relevant of the sources, FFT was performed on the signals and power was calculated. These 3 powers were then plotted over the frequency domain to produce PSD. We found the PSD plot with the highest peak in the frequency range of 1 to 3 Hz. This PSD indicated which of the ICA sources is most relevant. The selected source had a moving average window of size 5 performed on it. The outcome was the final PPG estimation.

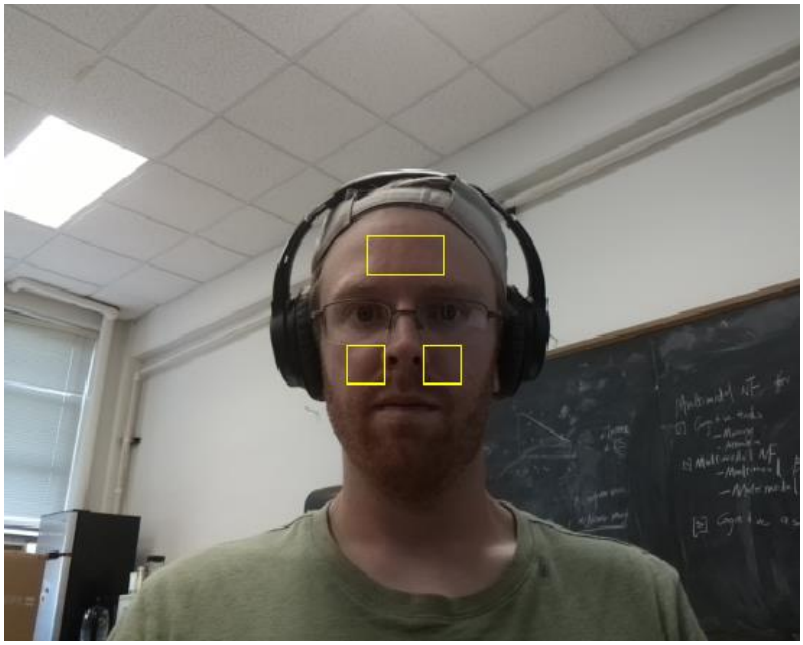


(a)

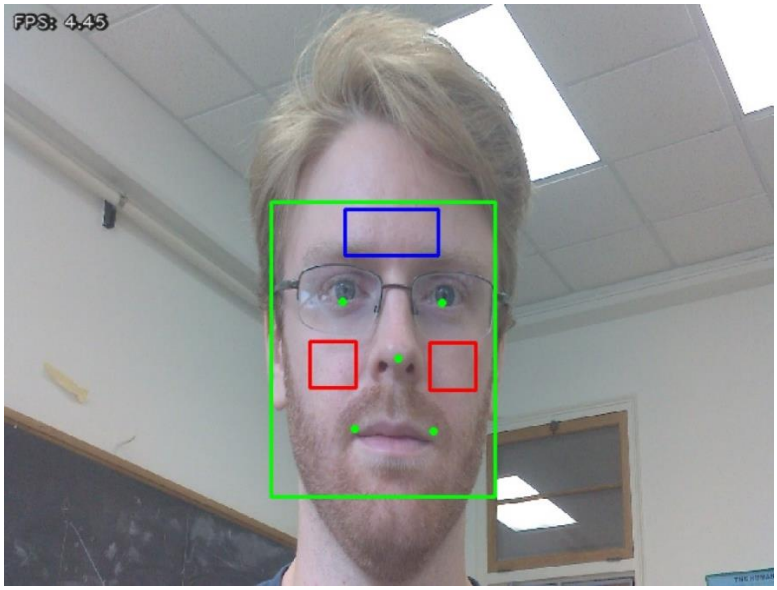


(b)

Figure 3.2 shows how the 2 face detection models, (a) Viola-Jones of the Pi and (b) MTCNN of the Nano, lock onto key features of the face that are useful when computing the ROI.



(a)



(b)

Figure 3.3 shows the results of the two face detection models used. (a) The Pi system using the Viola-Jones cascade model. (b) The Nano using MTCNN model.

3.3.2 HR Estimation Pipeline

Continuing with the PPG signal from the last pipeline, the signal was run through a bandpass filter (type = butterworth, order = 3) passing the signal that is between 0.8 and 3 Hz. The PSD of this signal was created just as before and the max peak of the PSD was detected indicating the frequency at the highest power of the signal. This frequency was multiplied by 60 which gave the estimated average HR over the number of frames passed into the algorithm which was 300.

CHAPTER FOUR

EXPERIMENT AND RESULTS

An experiment was carried out as a feasibility for future research in NICU. For this experiment, the 2 systems were compared to a HR monitor under similar conditions (i.e. lighting, tasks, etc.). The HR from the Nano system, Pi system and HR monitor were all compared.

4.1 Experiment

4.1.1 Setup

6 healthy adult volunteers (mean age = 21) participated in this study. In both systems the camera was placed 3 feet away positioned to look at the participant's face. A HR monitor (HRM) was placed on the wrist of the participants. Both systems recorded 200 frames of video data at a time. The resulting estimated HR of the 2 systems was compared to the HR of the HRM.

4.1.2 Procedure

The participants completed tasks in 3 stages, where their HR was estimated in between those tasks using 300 frames of data spanning about 20 seconds since the frame rate of the camera being set to 15 frames per second (fps). In the first stage, the participants sat in a chair for 1 minute. During that time, participants were invited to read, use their phone, write, or engage in any other activity normally done in a chair. After 1 minute, the participants looked straight ahead at the camera for data acquisition (about 20 seconds). For the second stage, the participants closed their eyes and sat quietly for 1 minute. They were instructed to focus on their breathing. Afterwards, the participants looked straight ahead for data acquisition. Finally, the third stage involved the participant standing up and performing a mild to moderate exercise for 1 minute—either walking or running in place. After 1 minute, the participant sat in the chair and looked forward for data acquisition. When all 3 stages were completed, this was considered 1 trial. Each participant went through 10 trials which took about 40 minutes. This concluded the experimental procedure. The HRM was then promptly removed and the study procedure involving the participant was complete. Figure 4.1 shows a schematic of the procedure.

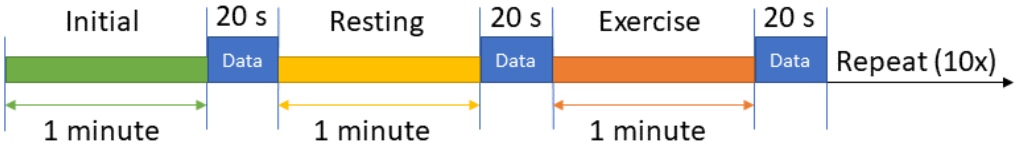


Figure 4.1 shows the experimental procedure performed. Each trial took about 4 minutes which meant the procedure took about 40 minutes to complete.

4.1.3 Analysis

Using the results from the models, we compared the HR estimated by our system to the HR of the PO for each trial during each task on every patient. We measured the difference of the measurements and found the mean error, absolute mean error, and the standard deviation of error or each trial.

4.1.4 Evaluation

RMSE

Root Mean Square Error (RMSE) is the standard deviation of the residuals. Residuals are a measure of how far from the regression line the data points are. RMSE is a measure of how spread out these residuals are and tells us how concentrated the data is around the line of best fit. This is calculated using the equation, Equation 1. Where n is the number of trials, y_i is the estimated HR, and $\hat{y}_i$ is the actual HR from the PO.

Equation 1

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(\hat{y}_i - y_i)^2}{n}}$$

Bland-Altman and Correlation

The Bland–Altman plot is a method of plotting data for analyzing the agreement between two different measurements of the same variable (Bland and Altman 1999). The y-axis consists of the difference between the variables, and the x-axis is the average of the variables. Limits of agreement were found by taking the mean

and adding and subtracting the standard deviation across all the infants multiplied by 1.96 as seen in (Bolkhovsky et al. 2012) for the upper and lower bound respectively. The Pearson's correlation plot shows the two variables plotted against each other in a scatter plot with the trendline drawn. The slope of the trendline is the correlation coefficient. A coefficient close to 1 means the two variables are highly correlated, and one close to zero shows very little correlation.

4.2 Results

The developed monitoring systems have been used and tested in a research project in the 2020 Summer Undergraduate Research Group Experience (SURGE) program at the University of Tennessee, Knoxville. Six undergraduate students with diverse background (mechanical engineering, biomedical engineering, computer science, nursing, and kinesiology) participated in the SURGE project. The undergraduates worked on the Pi system which provided a nice learning platform for them to acquire technical skills such as programming, computer vision, machine learning, and more. A picture of the Nano system is shown in Figure 1.2 below. Results of the SURGE 2020 project served as proof of success of the system being used as a research and education tool.

Figure 4.2 shows the Jetson Nano system. It includes a mini tripod and custom camera holder to keep the camera pointing straight.

4.2.1 Jetson Nano System

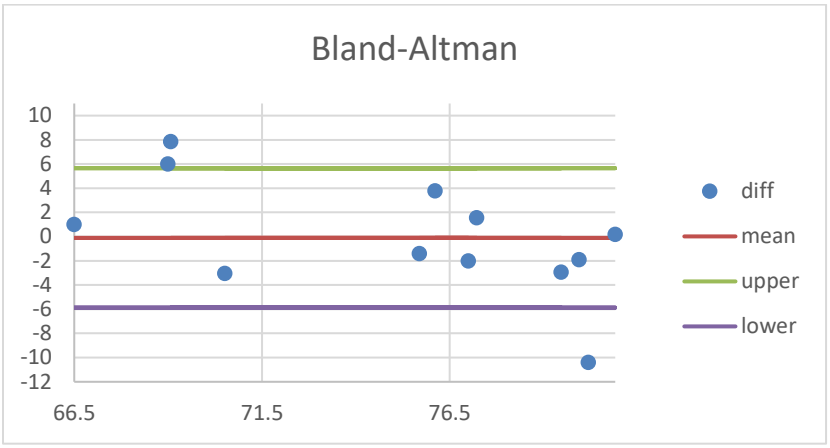
Table 4.1 shows the RMSE, absolute mean error, mean error, and standard deviation in error after each Stage in version 6 of the system. The results are fairly good for the Resting and Initial Stages but Exercise tended to highly underpredict the actual HR.

Stage	RMSE	Abs Mean Error	Mean Error	SD Error
Initial	3.32	2.58	0.362	3.14
Resting	2.91	2.43	0.434	2.88
Exercise	9.42	5.78	8.35	5.37

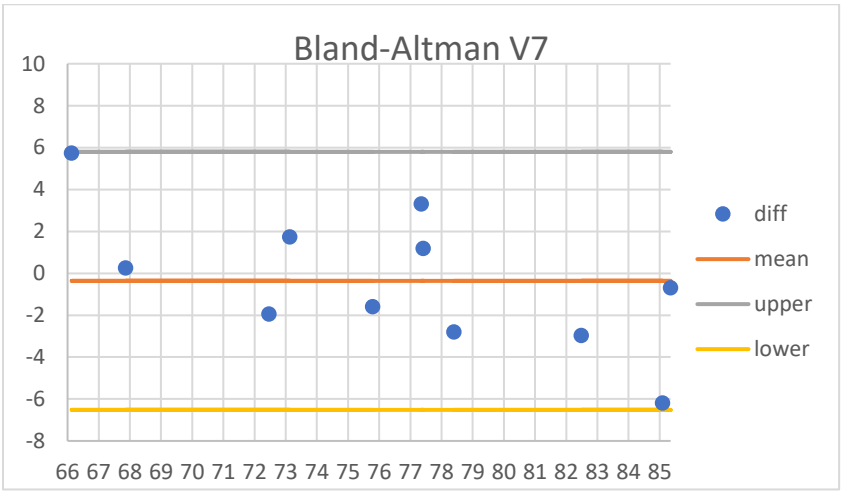
Table 4.1 shows that the results in the Exercise stage are worse, probably due to drastic changes of HR during the Exercise stage. Future research shall calculate real-time HR and compare the estimated PPG signal to the PPG from the PO, as seen in (Kumar et al. 2015) for better comparison with the gold standard obtained using PO. This way we could estimate the HR at the same rate as the PO and have a more direct comparison of the results. It would no longer be an average HR estimation of 20 seconds. Instead, it would be a continuously updating HR estimation (every second) that should show the same HR as the PO at every time point if it is accurate.

Figure 4.3 shows that most of the data points in (a) are within one SD of the mean. The Bland-Altman plot for V7 (b) show better result with all the data points falling within one SD of the mean.

Figure 4.4 shows that the r^2 value of 0.498 in V6 relates to a correlation of 0.706 which is an average correlation, but the value of 0.8501 in V7 indicates a good correlation of 0.922. In the literature, the correlations reported are usually between 0.8 and 0.9 which indicates that the improvement of the system has been significant, and the current results are relevant.

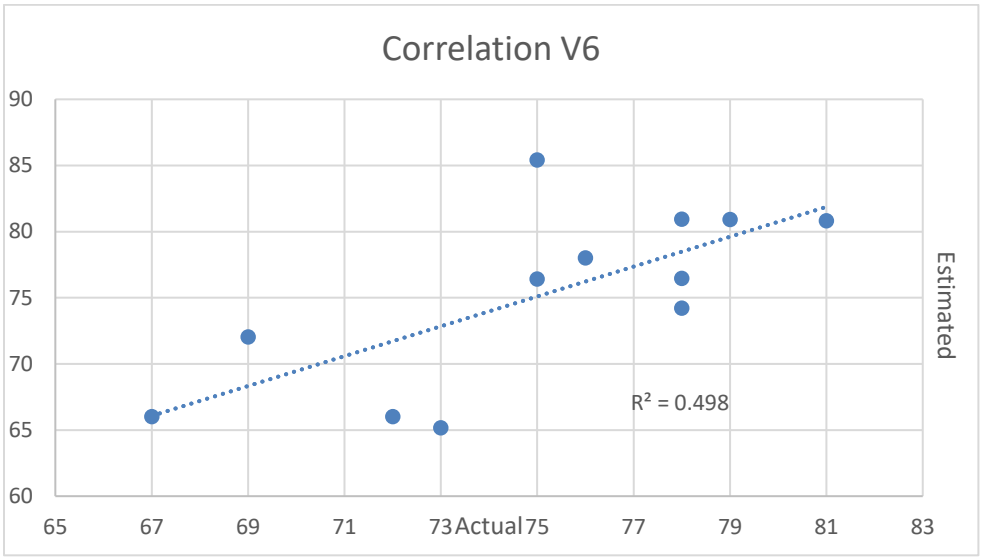


(a)

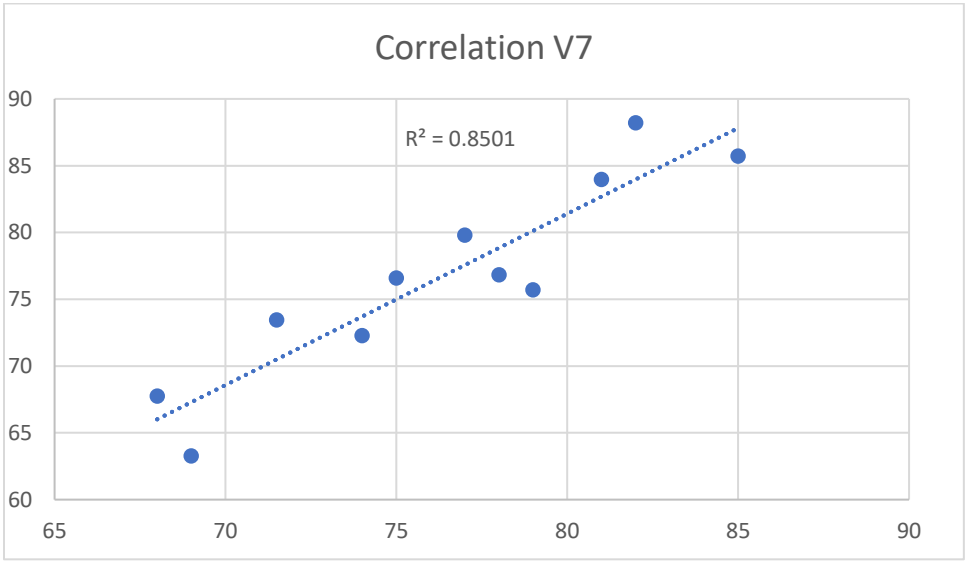


(b)

Figure 4.3 shows the Bland-Altman plot of the Initial Stage for one participant. (a) Bland-Altman for Nano V6 and (b) is Nano V7.



(a)



(b)

Figure 4.4 shows the Correlation plots of the data for one participant during the Initial Stage. (a) In version 6, the correlation coefficient is 0.498, and in (b) version 7, coefficient is 0.711.

4.2.2 Raspberry Pi System

The trend in results vs Stage shown in Table 4.2 is the same as what happened in the Nano system where the Resting Stage and Initial Stage had the best results and Exercise had the worst. We theorize this happened for the same reason as in the Nano system.

Figure 4.5 shows that all the data points lie within the upper and lower bounds which indicates good replicability. Since there are few data points, statistical significance cannot be inferred.

Figure 4.6 shows that R^2 is 0.6608, which yields a correlation coefficient of 0.813. This indicates a decent correlation between the 2 measurements. The first set of results using this system had a correlation of 0.15, but compared to the literature (around 0.9), the current correlation is still low and indicates more improvement needs to be made. Also, there are only 5 data points being used in Figure 4.5 which indicates much more data has to be gained in order to call these results significant.

Table 4.2 shows the results of the Pi System after each Stage.

Stage	RMSE	Abs Mean Error	Mean Error	SD Error
Initial	5.9	6.14	4.85	5.88
Resting	6.48	5.33	3.56	7.79
Exercise	11.4	14.11	11.95	14.66

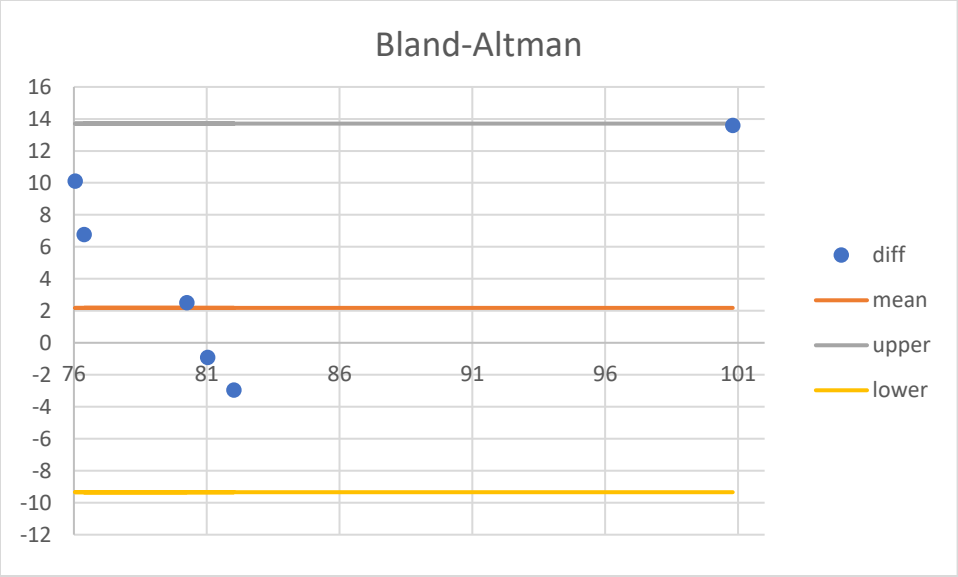


Figure 4.5 shows the Bland-Altman plot of the Initial Stage for one participant using the Pi V4.

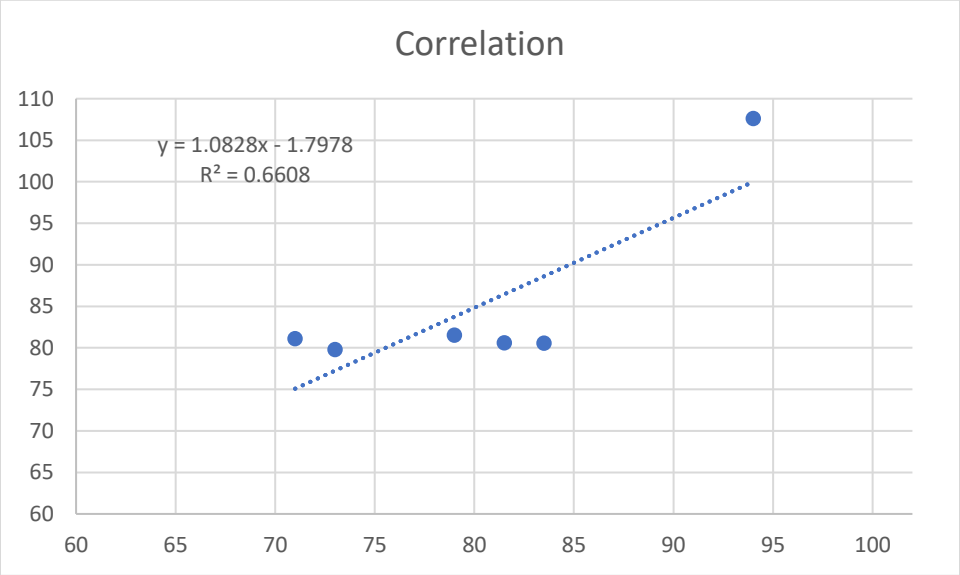


Figure 4.6 shows the Correlation plot of the data for one participant during the Initial Stage. The x-axis is the actual HR from the PO and the y-axis is the estimated HR from our system. R^2 is used to find the correlation coefficient by taking the square root of this number.

CHAPTER FIVE

DISCUSSION AND CONCLUSION

This chapter starts by showing the performance of oth systems over each iteration. Then it summarizes the problems that arose during this project. Some problems were solved, others have yet to be. This chapter will also give a conclusion on the results as well as discuss suggestions for future work. Both systems are discussed in greater detail.

5.1 Performance

In this section the performance of both systems is documented over each iteration. Performance is based on RMSE between the current standard for measuring HR and the HR values obtained in this study. Table 5.1 shows the RMSE, absolute mean error, mean error, and standard deviation in error for each system over all iterations. Each iteration had at least 10 trials recorded to ensure the results were relevant. A brief description of current issues and changes made between each iteration is included in the Notes column. After each iteration, the results almost always improved which is very promising. The latest iteration has results that are significant when compared to current literature.

Tables 5.1 and 5.2 show that there was significant improvement in both systems over the iterations. A major reason for improvement in the Nano system was solving the fps issue which allowed each frame to be captured at a consistent rate. This made analyzing the signals more accurate. An important improvement was the acquisition of a good consistent ROI over the cheeks and forehead.

In the Pi system, the biggest improvement came from switching the software from Matlab to Python. This is probably the case due to Python having better build in libraries for computer vision. Plus running the MTCNN face detection model in Python was now doable on the Pi, which allowed more reliable and accurate face detection.

Table 5.1 shows the documented results of the Nano over each iteration.

Iteration	RMSE	Abs Mean Error	Mean Error	SD of Error	Notes
V1	22.07	18.88	9.53	19.91	Doesn't remove missing frames; lighting is not perfect; camera being held by hand; low and inconsistent fps
V2	16.07	13.56	7.19	14.37	Updated feature selector to pick the green channel; camera mount added; missing frames are less common; better lighting (consistent); need better Roi placement; fps is still an issue
V3	27.63	19.68	17.64	21.26	Added moving average, using ICA source, separation of face recording & face detection, FPS
V4	8.95	6.9	4.08	7.97	Chose new FPS based on rate of image collection, there is an error where image collection lags at the end
V5	13.5	8.29	-2.93	10.65	Adjusted ROI of forehead (smaller); sped up image collection loop which made fps higher
V6	5.01	3.5	-0.109	2.94	adjusted bandpass to 0.8-3Hz; made forehead taller; improved image collection system; took out ICA source
V7	3.32	2.58	0.362	3.14	Made forehead wider and higher up on the face; made cheeks further from the nose

Table 5.2 shows the documented results of the Pi over each iteration.

Software	Iteration	RMSE	Abs Mean Error	Mean Error	SD of Error
Matlab	V1	10.49	13.25	12.95	10.36
Matlab	V2	11.46	11.6	-2.57	13.46
Python	V3	7.35	7.24	2.18	8.51
Python	V4	5.9	6.14	4.85	5.88

5.2 Limitations

One of the biggest challenges was an inconsistent FPS during data acquisition. This can be solved with better hardware and more precise programming. There is debate in the literature over whether to use the frame rate of the camera or the frequency at which the software can process the data. We tried doing both and had better results with the camera fps. We separated the face detection model from the acquisition pipeline to speed up image collection which helped control the fps, but there were still fluctuations and inconsistent face detections. The rest of the signal processing pipeline seemed to be more accurate when the fps was on the lower end of the spectrum, so we controlled the fps of the image collection to stay around 15 frames per second.

A limitation in the potential use of this system as a clinical tool is the practice of kangaroo care. Kangaroo care is defined as developmental care by holding an infant skin-to-skin against the parent's bare chest. While the infant is being held, our system cannot currently lock onto his/her face. Thus, our system would not provide any physiological monitoring while this type of care is being performed.

Another limitation was the PO used. For testing, we used a low-cost commercial grade PO that was highly rated on Amazon.com. There are many reviews stating they experienced inaccurate readings. The refresh rate of the readings was also slow which made for few points of reference when averaging the HR over the image collection period.

Another limitation is when the face is turned, the ROI for the cheeks sometimes collects pixels that are not over the face. To overcome this problem, participants in this study were requested to always look straight at the camera. Future work shall consider a solution that will make the ROI area smaller as the person turns their face and eventually the ROI will disappear if the cheek can no longer be seen by the camera.

5.3 Future Work

Much work can be done to improve the performance for both systems. One possible issue is lack of controlled, consistent lighting in the experiments. Controlling the number of lumens in the experiment could lead to more consistent results. Another improvement can be made in the quality of the camera used. We used a relatively cheap camera that was easy to implement; however, better results may require a camera that has a higher resolution, frame rate, and higher sensitivity should be used.

Currently, motion tracking is one of the new features found in the state-of-the-art literature. While it was never implemented in this project, it was researched. In the

future, we plan to implement such algorithms to account for motion. One such algorithm is the Kanade-Lucas-Tomasi (KLT) feature tracker (Tomasi and Kanade 1991).

Future work shall also adopt a higher quality pulse oximeter to improve the reliability of the reference results. Comparing the estimated PPG signal of our system to a PO would allow us to create a signal-to-noise ratio (SNR). This SNR would allow the us to better understand how much of the processed signal is correlated to the true underlying PPG signal as seen in Kumar et al. (2015).

In the future, we will make our systems give an estimated HR in real time. We will do so by using a moving window to update the HR continuously.

Our current implementations use frequency analysis to find the estimated HR. Another method that will be explored is peak-to-peak analysis. This method finds the peaks in the estimated PPG signal. The time between each peak is the inter-beat interval (IBI). The average IBI can be taken over a section of frames. 60 divided by the average IBI would give the estimated average HR over the section of frames.

5.4 Discussion

Both systems used the same camera. The Pi camera V2 is a solid camera, but it had some issues as well. The quality of the image was sometimes lacking. The exposure could be set, but it was not up to par as some more expensive counterparts. The camera was easy to use and user-friendly though. It was very compatible with both the Nano and Pi and there was plenty of documentation which made debugging easier.

5.4.1 Pi System

The Raspberry Pi was chosen due to its adaptability and low cost. In order to use Matlab, we had to use a laptop with the correct support packages installed. The laptop was connected to the Pi over WiFi and libraries were downloaded onto the Pi's operating system. A positive aspect of using Matlab was the quick trouble-shooting and easy-to-implement models.

The system is very lightweight and smaller than the Nano. Processing power was not an issue when using the Viola-Jones cascade model. The system was able to estimate HR. However, the system's face detection was an issue. It was slow and inconsistent. This led to missing frames which throw off the HR estimation. The Pi was also lacking some power at running the face detection algorithm at better camera settings, so a compromise had to be made in the image quality of the inputted data which led to a weaker PPG estimation.

5.4.2 Nano System

A Jetson Nano was first considered as a solution to some much-needed power to run a more advanced Face Detection model. The Jetson Nano has a great GPU that could run the MTCNN model at a quick frame rate (10 fps). There was less documentation on using the Nano vs the Pi, but it was surprisingly adaptable. It was easy to create many different versions of the system and each consecutive version made good progress.

5.5 Conclusion

In conclusion, both systems accomplished the goal they set to achieve. One lesson learned in this study is that different hardware may be compatible with different software. Also, estimating physiological parameters from video footage is a challenging task as there are many variables, which are hard to control. Moreover, insights are gained as to how the blood flow changes in the body and how skin and blood interacts with outside light.

LIST OF REFERENCES

Aarts L. A. M., V. Jeanne, J. P. Cleary, C. Lieber, J. S. Nelson, S. Bambang Oetomo, and W. Verkruysse, “Non-contact heart rate monitoring utilizing camera photoplethysmography in the neonatal intensive care unit - a pilot study," Early human development, vol. 89, no. 12, pp. 943-948, Dec. 2013, PMID: 24135159.

About CDH. (n.d.). Retrieved August 10, 2020, from http://www.cdhgenetics.com/congential-diaphragmatic-hernia.cfm

Ahmed A K, Harness J B and Mearns A J 1982 Respiratory control of heart rate Eur. *J. Appl. Physiol.* 50 95–104

Allen J., “Photoplethysmography and its application in clinical," Physiological Measurement, vol. 28, no. 3, p. R1, Mar. 2007. [Online]. Available: http://iopscience.iop.org/0967-3334/28/3/R01

Amelard, R., Clausi, D. A., & Wong, A. (2016). Spectral-spatial fusion model for robust blood pulse waveform extraction in photoplethysmographic imaging. Biomedical optics express, 7(12), 4874-4885.

Aoyagi, T. (1974). Improvement of the earpiece oximeter. Abstracts of the Japanese Society of Medical Electronics and Biological Engineering, 1974, 90-91.

Bland JM, Altman DG. Measuring agreement in method comparison studies. Stat Methods Med Res. 1999;8:135–60. 10.1191/096228099673819272

Bolkhovsky, J. B., Scully, C. G., & Chon, K. H. (2012). Statistical analysis of heart rate and heart rate variability monitoring through the use of smart phone cameras. *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 1610-1613). IEEE.

Boyar, V. 2019. 'Medical Adhesive-related Skin Injury and Chemical Injury in a Preterm Neonate', *Wound Management & Prevention*, 65: 8-12.

Burton A C 1939 The range and variability of the blood flow in the human fingers and the vasomotor regulation by body temperature *Am. J. Physiol.* 127 437–53

Burton A C and Taylor R A 1940 A study of the adjustment of peripheral vascular tone to the requirements of the regulation of body temperature *Am. J. Physiol.* 129 565–77

Chung, H. U., B. H. Kim, J. Y. Lee, J. Lee, Z. Q. Xie, E. M. Ibler, K. Lee, A. Banks, J. Y. Jeong, J. Kim, C. Ogle, D. Grande, Y. Yu, H. Jang, P. Assem, D. Ryu, J. W. Kwak, M. Namkoong, J. B. Park, Y. Lee, D. H. Kim, A. Ryu, J. Jeong, K. You, B. W. Ji, Z. J. Liu, Q. Z. Huo, X. Feng, Y. J. Deng, Y. S. Xu, K. I. Jang, J. Kim, Y. H. Zhang, R. Ghaffari, C. M. Rand, M. Schau, A. Hamvas, D. E. Weese-Mayer, Y. G. Huang, S. M. Lee, C. H. Lee, N. R. Shanbhag,

A. S. Paller, S. Xu, and J. A. Rogers. 2019. 'Binodal, wireless epidermal electronic systems with in-sensor analytics for neonatal intensive care', *Science*, 363: 947-+.

Elgendi M., “On the analysis of _ngertip photoplethysmogram signals," Current Cardiology Reviews, vol. 8, no. 1, pp. 14{25, Feb. 2012, PMID: 22845812 PMCID: PMC3394104. [Online]. Available: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3394104/

Freund, Yoav and Robert E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. Journal of Computer and System Sciences, 55(1):119–139, August 1997

Harford, M., Catherall, J., Gerry, S., Young, J. D., & Watkinson, P. (2019). Availability and performance of image-based, non-contact methods of monitoring heart rate, blood pressure, respiratory rate, and oxygen saturation: a systematic review. *Physiological measurement*, 40(6), 06TR01.

Harness J B and Marjanovic D Z 1989 Low-frequency photoplethysmograph signals *Clin. Phys. Phys. Meas.*10 365–8

Hertzman A B and Dillon J B 1940a Distinction between arterial, venous and flow components in photoelectric plethysmography in man *Am. J. Physiol.* 130 177–85

Hertzman A B and Dillon J B 1940b Applications of photoelectric plethysmography in peripheral vascular disease *Am. Heart J.* 20 750–61

Hertzman A B and Dillon J B 1940b Applications of photoelectric plethysmography in peripheral vascular disease. Am. Heart J. 20 750–61

Hertzman A B and Flath F 1963 The continuous simultaneous registration of sweating and blood flow in a small skin area *Aerospace Med.* 34 710–3

Hertzman A B and Roth L W 1942a The vasomotor components in the vascular reactions in the finger to cold *Am. J. Physiol.* 136 669–79

Hertzman A B and Roth L W 1942b The reactions of the digital artery and minute pad arteries to local cold *Am J. Physiol* 136 680–91

Hertzman A B and Roth L W 1942c The absence of vasoconstrictor reflexes in the forehead circulation. Effects of cold *Am. J. Physiol.* 136 692–7

Holton B. D., K. Mannapperuma, P. J. Lesniewski, and J. C. Thomas, Signal recovery in imaging photoplethysmography," Physiological Measurement, vol. 34, no. 11, pp. 1499{1511, Nov. 2013. [Online]. Available: http://m.iopscience.iop.org/0967-3334/34/11/1499

Humphreys K., T. Ward, and C. Markham, \Noncontact simultaneous dual wavelength photoplethysmography: a further step toward noncontact pulse

oximetry," The Review of scienti_c instruments, vol. 78, no. 4, p. 044304, Apr. 2007, PMID: 17477684.

Hyndman B W, Kitney R I and Sayers B M 1971 Spontaneous rhythms in physiological control systems *Nature* 233 339–41

Johnson, D. E. 2016. 'Extremely Preterm Infant Skin Care A Transformation of Practice Aimed to Prevent Harm', *Advances in Neonatal Care*, 16: S26-S32.

Kamshilin A. A., S. Miridonov, V. Teplov, R. Saarenheimo, and E. Nippolainen, 54 Photoplethysmographic imaging of high spatial resolution," Biomedical Optics Express, vol. 2, no. 4, pp. 996-1006, Mar. 2011, PMID: 21483621 PMCID: PMC3072138. [Online]. Available: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3072138/

Lewandowska M., J. Ruminski, T. Kocejko, and J. Nowak, “Measuring pulse rate with a webcam - a non-contact method for evaluating cardiac activity," in 2011 Federated Conference on Computer Science and Information Systems (FedCSIS), Sep. 2011, pp. 405-410.

Lucas B. D. and T. Kanade, “An iterative image registration technique with an application to stereo vision," 1981, p. 674679.

Molitor, H., & Kniazuk, M. (1936). A new bloodless method for continuous recording of peripheral circulatory changes. Journal of Pharmacology and Experimental Therapeutics, 57(1), 6-18.

Morris, A., and A. Paller. “New Technology in the NICU.” Northwestern Medicine Magazine, 22 May 2018, magazine.nm.org/2019/05/28/new-technology-in the-nicu/.

Nitzan M, Babchenko A, Milston A, Turivnenko S, Khanokh B and Mahler Y 1996a Measurement of the variability of the skin blood volume using dynamic spectroscopy *Appl. Surface Sci.* 106 478–82

Nitzan M, Turivnenko S, Milston A, Babchenko A and Mahler Y 1996b Low-frequency variability in the blood volume pulse measured by photoplethysmography *J. Biomed. Opt.* 1 223–9

Nitzan M., B. Khanokh, and Y. Slovik, “The difference in pulse transit time to the toe and _nger measured by photoplethysmography," Physiological measurement, vol. 23, no. 1, pp. 85{93, Feb. 2002, PMID: 11876244.

Oranges, T., V. Dini, and M. Romanelli. 2015. 'Skin Physiology of the Neonate and Infant: Clinical Implications', *Advances in Wound Care*, 4: 587-95.

Poh M. Z., D. J. McDu_, and R. W. Picard, “Non-contact, automated cardiac pulse measurements using video imaging and blind source separation," Optics Express, vol. 18, no. 10, pp. 10,762-10,774, May 2010. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-18-10-10762

Poh M. Z., D. McDuff, and R. Picard, “Advancements in noncontact, multiparameter physiological measurements using a webcam," IEEE Transactions on Biomedical Engineering, vol. 58, no. 1, pp. 7-11, Jan. 2011.

Shusterman V, Anderson K P and Barnea O 1997 Spontaneous skin temperature oscillations in normal human subjects *Am. J. Physiol.* 273 (3 Pt 2) R1173–81

Smith, J. R. 2012. 'Comforting Touch in the Very Preterm Hospitalized Infant An Integrative Review', *Advances in Neonatal Care*, 12: 349-65.

Sun Y., S. Hu, V. Azorin-Peris, S. Greenwald, J. Chambers, and Y. Zhu, “Motion-compensated noncontact imaging photoplethysmography to monitor cardiorespiratory status during exercise," Journal of biomedical optics, vol. 16, no. 7, p. 077010, Jul. 2011, PMID: 21806290.

Tomasi C. and T. Kanade, “Detection and tracking of point features," Carnegie Mellon University, Technical Report MU-CS-91-132, 1991.

Verkruysse W., L. O. Svaasand, and J. S. Nelson, “Remote plethysmographic imaging using ambient light," Optics express, vol. 16, no. 26, pp. 21,434-21,445, Dec. 2008, PMID: 19104573 PMCID: PMC2717852.

Villarroel, M., Chaichulee, S., Jorge, J., Davis, S., Green, G., Arteta, C., ... & Tarassenko, L. (2019). Non-contact physiological monitoring of preterm infants in the Neonatal Intensive Care Unit. NPJ digital medicine, 2(1), 1-18.

Viola P. and M. Jones, “Rapid object detection using a boosted cascade of simple features," in Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001. CVPR 2001, vol. 1, 2001, pp. I511-I518 vol.1.

Wang, G., Atef, M., & Lian, Y. (2018). Towards a continuous non-invasive cuffless blood pressure monitoring system using PPG: Systems and circuits review. IEEE Circuits and Systems Magazine, 18(3), 6-26.

Wieringa F. P., F. Mastik, and A. F. W. van der Steen, Contactless multiple wavelength photoplethysmographic imaging: a first step toward "SpO2 camera technology," Annals of biomedical engineering, vol. 33, no. 8, pp. 1034-1041, Aug. 2005, PMID: 16133912.

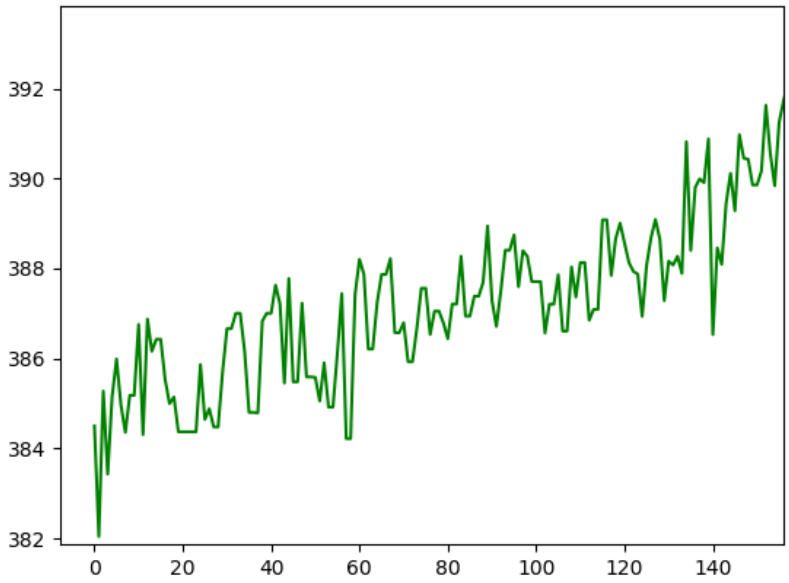
Yoshiya, I., Shimada, Y., & Tanaka, K. (1980). Spectrophotometric monitoring of arterial oxygen saturation in the fingertip. Medical and Biological Engineering and Computing, 18(1), 27-32.

Zhang, K., Zhang, Z., Li, Z., and Qiao, Y. (2016). Joint face detection and alignment using multitask cascaded convolutional networks. IEEE Signal Processing Letters, 23(10):1499–1503.

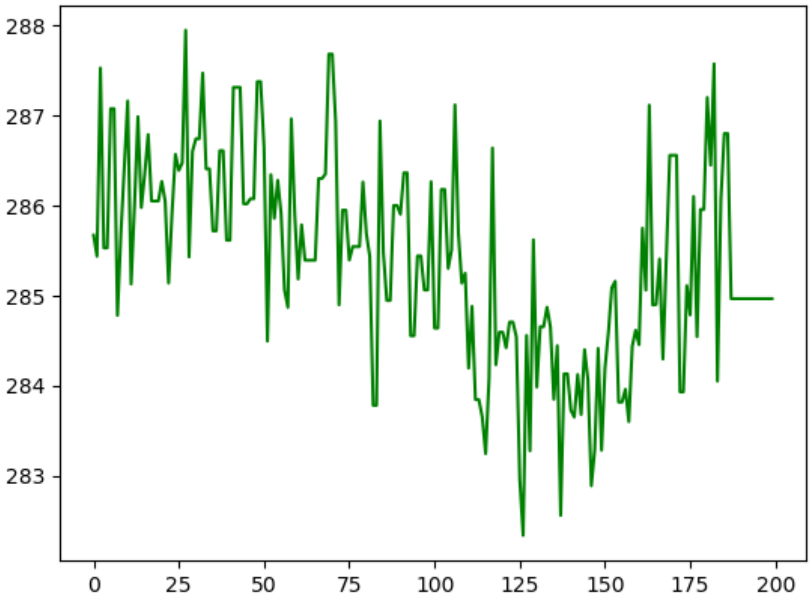
Zheng J., S. Hu, V. Azorin-Peris, A. Echiadis, V. Chouliaras, and R. Summers, “Remote simultaneous dual wavelength imaging photoplethysmography: a further step towards 3-d mapping of skin blood microcirculation</title>," F. S. Azar and X. Intes, Eds., Feb. 2008, pp. 68 500S-68 500S-8. [Online]. Available: http://spie.org/x648.html?product id=761705

APPENDIX

The following graphs are of the signals produced at each stage of the experiment by the Nano System. (a) Initial Stage and (b) Exercise Stage.

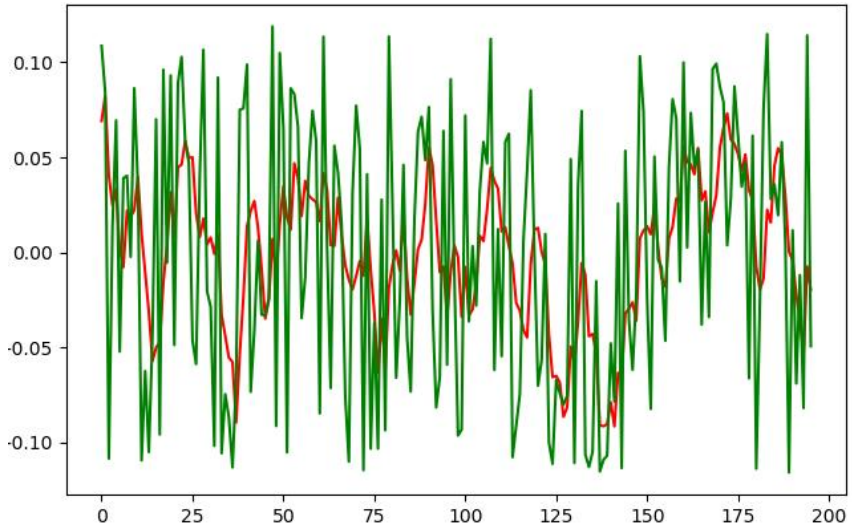


(a)

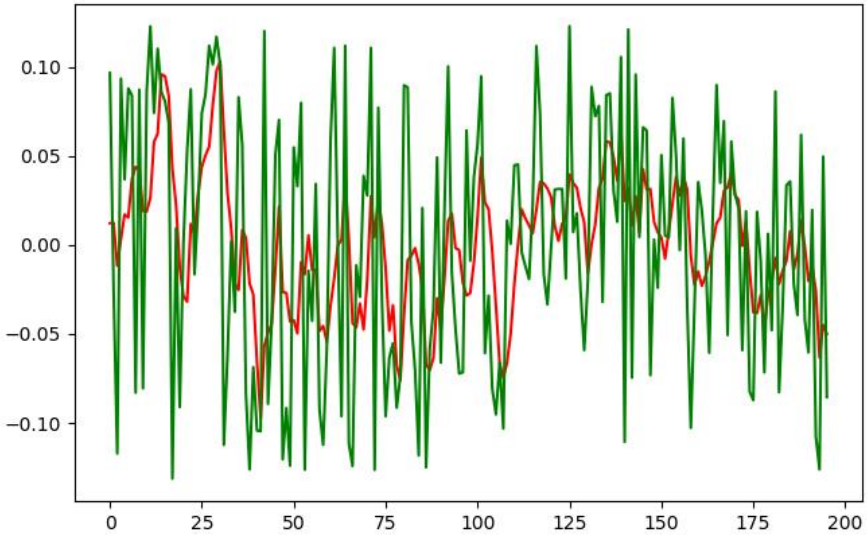


(b)

Figure A.1 shows the unprocessed PPG signals from the Nano following a trial from each Stage. These are plotted on the x-axis of frame number.

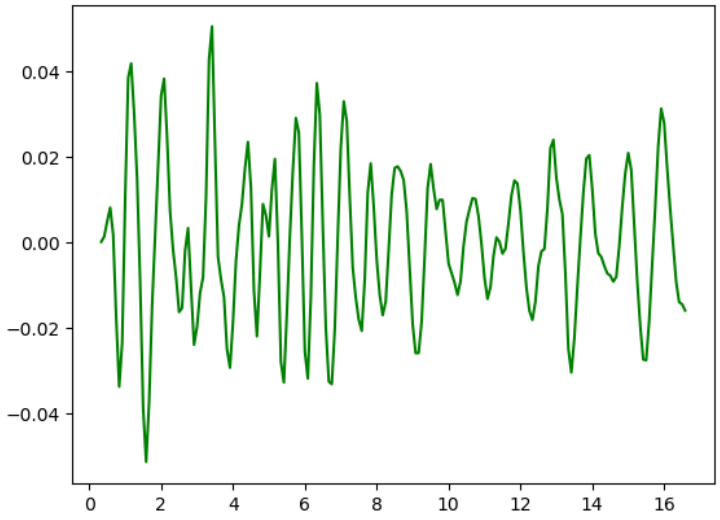


(a)

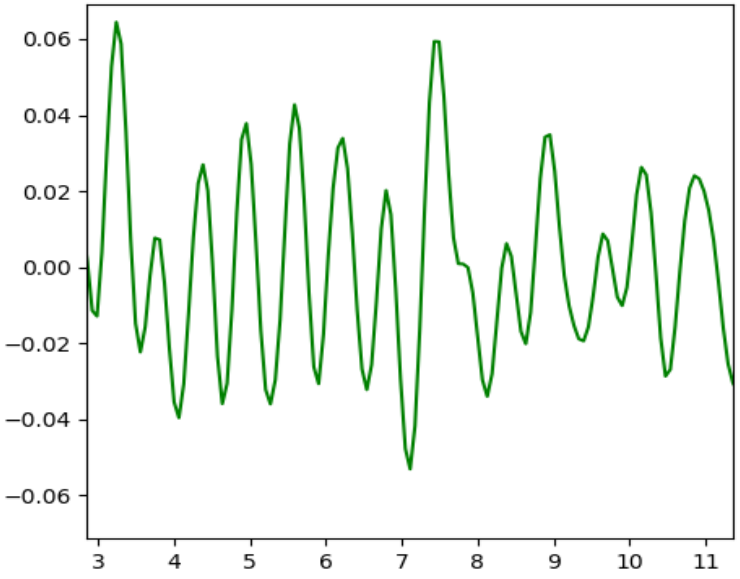


(b)

Figure A.2 shows the Running Mean after each Stage. These are plotted on the x-axis of frame number.

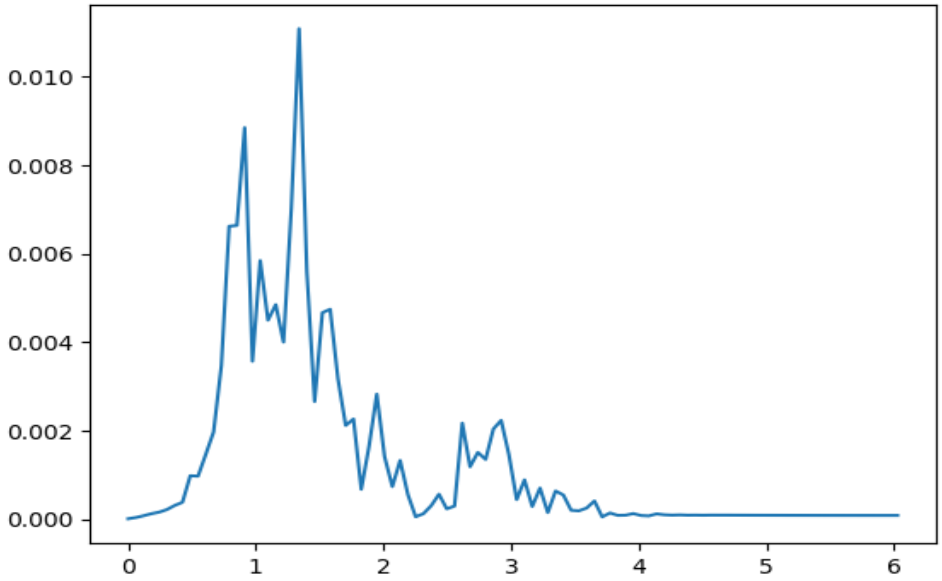


(a)

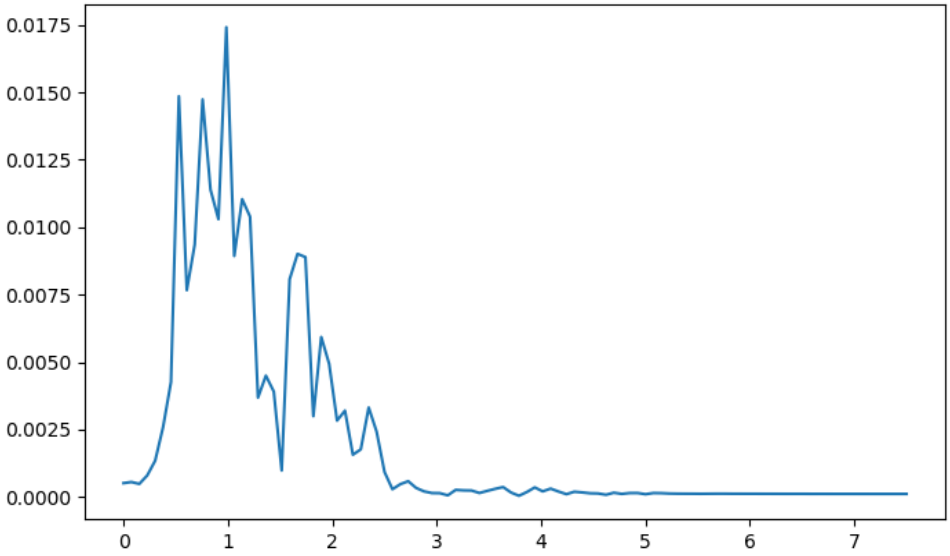


(b)

Figure A.3 shows the filtered signal after each Stage. These are plotted on the x-axis of time.



(a)



(b)

Figure A.4 shows the PSD (Power vs Frequency) plot after each Stage.

This Matlab code is for the ROI estimation for the Pi system

```
x = roiHead(:,1);
y = roiHead(:,2);
roii = cell(3,1);
faceWidth = max(x) - min(x);
faceHeight = max(y) - min(y);
A = sqrt(faceWidth*faceHeight*0.04); % area of cheeks and forehead

bottomOfFH = round(min(y)); % the bottom of forehead
topOfFH = round(min(y) + faceHeight*0.2); % the top of forehead

roiForeheadX = ((min(x) + max(x)) / 2);
roiForeheadY = ((bottomOfFH + topOfFH) / 2);
roiF = [roiForeheadX roiForeheadY]; % forehead center coordinates

roii{1} = [roiF(1)-A roiF(2)+A/2; roiF(1)-A roiF(2)-A/2; roiF(1)+A
roiF(2)+A/2; roiF(1)+A roiF(2)-A/2];

% cheek detection
leftCheekX = (min(x) + faceWidth*0.3);
leftCheekY = (min(y) + faceHeight*0.65);
roiL = [leftCheekX leftCheekY]; % left cheek center coordinates

roii{2} = [roiL(1)-A/2 roiL(2)+A/2; roiL(1)-A/2 roiL(2)-A/2;
roiL(1)+A/2 roiL(2)+A/2; roiL(1)+A/2 roiL(2)-A/2];

rightCheekX = (min(x) + faceWidth*0.7);
rightCheekY = (min(y) + faceHeight*0.65);
roiR = [rightCheekX rightCheekY]; % right cheek center coordinates

roii{3} = [roiR(1)-A/2 roiR(2)+A/2; roiR(1)-A/2 roiR(2)-A/2;
roiR(1)+A/2 roiR(2)+A/2; roiR(1)+A/2 roiR(2)-A/2];
```

This Python code is for the ROI estimation for the Nano system.

```
x1, y1, x2, y2 = int(bb[0]), int(bb[1]), int(bb[2]), int(bb[3])

#using face coordinates, detect cheek and forehead region
l_eye = (ll[0], ll[5]) # x and y coordinate of the left eye
r_eye = (ll[1], ll[6])
nose = (ll[2], ll[7])
l_lip = (ll[3], ll[8])
r_lip = (ll[4], ll[9])

roi = np.array([[0,0,0,0],[0,0,0,0],[0,0,0,0]])

A = np.sqrt((x2 - x1) * (y2 - y1)) * 0.2

l_cheek = (np.average([l_eye[0], l_lip[0]]), np.average([l_eye[1],
l_lip[1]])) #left cheek center coordinates

r_cheek = (np.average([r_eye[0], r_lip[0]]), np.average([r_eye[1],
r_lip[1]])) #right cheek center coordinates

forehead = (np.average([l_eye[0], r_eye[0]]), np.average([l_eye[1],
r_eye[1]]) - A) #forehead center coordinates

        #determine if side of face is showing
if nose[0] > np.average([l_eye[0], l_lip[0]]):
       roi[0] = np.array([int(l_cheek[0] - A/2), int(l_cheek[1] -
       A/2), int(l_cheek[0] + A/2), int(l_cheek[1] + A/2)])

if nose[0] < np.average([r_eye[0], r_lip[0]]):
       roi[1] = np.array([int(r_cheek[0] - A/2), int(r_cheek[1] -
       A/2), int(r_cheek[0] + A/2), int(r_cheek[1] + A/2)])

if nose[0] > np.average([l_eye[0], l_lip[0]]) and nose[0] <
np.average([r_eye[0], r_lip[0]]):
       roi[2] = np.array([int(forehead[0] - A/2), int(forehead[1] -
       A/2), int(forehead[0] + A/2), int(forehead[1] + A/2)])
```

VITA

Benjamin Sweely was born in Johnson City, TN on September 22, 1994. He finished a Bachelor's degree in Aerospace at the University of Tennessee. He has been pursuing a Master's degree in Mechanical Engineering since August 2018. He enjoys playing Badminton and reading books in his free time.