8-2009

Non-contact pulse measurement for use in smart room applications.

Nicholas R. Blumenthal
University of Louisville

NON-CONTACT PULSE MEASUREMENT FOR USE
IN SMART ROOM APPLICATIONS

By

Nicholas R. Blumenthal
B.S., University of Louisville, 2008

A Thesis
Submitted to the Faculty of the
University of Louisville
J. B. Speed School of Engineering
as Partial Fulfillment of the Requirements
for the Professional Degree

MASTER OF ENGINEERING

Department of Electrical and Computer Engineering

August 2009

NON CONTACT PULSE MEASUREMENT FOR USE
IN SMART ROOM APPLICATIONS

Submitted by:_____
                    Nicholas R. Blumenthal

A Thesis Approved on

_____
                    (Date)

by the Following Reading and Examination Committee:

_____
        Dr. Aly A. Farag, Thesis Director

_____
              Dr. Thomas L. Starr

_____
                Dr. John Naber

_____
            Dr. Michael Cunningham

ACKNOWLEDGEMENTS

ABSTRACT

Nicholas R. Blumenthal

August 2008

The past decade has brought about major advances in the realms of both healthcare and biometrics. New technologies and techniques in the field of biometrics has allowed for the quick and efficient identification of people, while new healthcare technologies are allowing for less invasive monitoring of a person's physiological state indicators. The combination of these two fields has allowed for a whole new frontier of science to be explored in the form of smart rooms.

Smart rooms use the fusion of biometrics and healthcare monitoring to provide a solution to the problem of how best to monitor a patient to ensure their best possible health, safety, and comfort. This work attempts to push the boundaries of the field by creating a smart room technology that can provide nonintrusive monitoring of patient heart rate. This type of room could be used to revolutionize patient monitoring in terms of both comfort and safety.

In this thesis, the culmination of several technological advances in the Computer Vision and Image Processing Lab were utilized to develop a methodology for the non-contact detection of a subject's pulse. The algorithms and methodology employed in this thesis resulted in a system that was able to identify a subject's pulse with 91.2% accuracy

in a partially automated system. This demonstrates the proof of concept and shows a

wide new range of possibilities for the world of medicine and patient monitoring.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I:

INTRODUCTION

Few things in the world are as important as one's physical well-being. To this end, trillions of dollars are spent on healthcare, fitness, and medicine every year for the sole purpose of ensuring that the physical body continues to operate at its peak efficiency [1]. In almost every medical situation there are several key indicators of person's physiological state that help the medical professional determine a person's level of health and fitness.

These indicators, commonly referred to as vital signs, convey different information about a person's physical and psychological state. The four main vital signs are respiration, temperature, blood pressure, and pulse. These indicators give medical professionals a starting point from which to base their diagnosis as well as important indicators with which to measure the continuing progress of patients after treatments or injuries [2].

As the name indicates, vital signs are integral to any diagnosis and there are a plethora of methods for measuring each vital sign with a range of accuracy and invasiveness for each measurement technique. Cuffs and thermometers can be used for blood pressure and temperature respectively, while all-purpose monitoring equipment such as electrocardiograms can be used to monitor respiration and pulse. Indeed, the

pulse is a very important human physiological measurement with a huge variety of techniques to detect it. These techniques include palpation, photoplethysmography, and even measurement of the electrical signals guiding the heart such as in electrocardiograms. Invariably however, all of these available techniques use some form of contact for measurement, which yields an integral problem that this thesis work attempts to solve.

Under certain circumstances, contact measurement of vital signs is not a large problem for either the patient of the medical professional. There are however, a large number of medical situations where the use of contact measurement devices could create discomfort for the patient or create a sustained burden on the medical professional in charge of patient care. These circumstances often arise in extended care situations, such as long hospital stays for surgeries or in elderly care scenarios, where the patient needs to have his or her vital signs monitored for an lengthy period of time. In these situations, it can often be uncomfortable for the patient to be connected to monitoring equipment such as an ECG device for extended periods of time. In addition, these devices can often be dislodged by normal patient activity such as movement while sleeping. This creates an extra burden on the medical staff as they are often called to reattach the monitoring devices. In extreme situations, such as severe burn victims, the attachment of measuring devices may be impossible due to the excessive pain they cause the patient. These types of injuries are not so rare as one might expect, with over 20,000 burn victims suffering severe burns over 25% of their bodies in the United States each year [3]. The solution to

this dilemma is the subject of this thesis, which is the creation of a non-contact form of vital sign measurement.

Due to the disparate nature of the vital signs, only the pulse was chosen as the target vital sign to achieve a form of non-contact measurement. Indeed, non-contact measurement of the pulse has been achieved through the use of active monitoring via a radar measurement device as explored in Geisheimer's 1998 article: "Radar Vital Signs Monitor." Unfortunately, this equipment was an active measurement technique and easily corrupted by subject movement. Conversely, the goal of this project is to create a passive measuring device with which the pulse can be ascertained from a distance from the subject, even in the presence of a certain threshold of movement.

To achieve this goal, this project utilizes thermal imaging techniques to peek underneath the human skin and view the heat patterns due to the perfusion of blood, which is related to the pulse rate. The methodology then employs novel mapping, segmenting, and signal processing techniques to extract a final heart rate from the thermal videos of the subjects. The end result is a methodology which can serve as a basis for non-contact pulse measurement and which can vastly improve patient comfort while reducing the burden placed on healthcare providers.

CHAPTER II:

LITERATURE REVIEW

2.1 PULSE OVERVIEW

An understanding of the basic mechanics of the pulse and the cardiovascular system is necessary to understand the various measurement techniques. The cardiovascular system is a collection of organs and tissues which are designed to transport nutrients and chemicals to the tissues of the body while collecting waste products that need to be expelled from the body [4]. This system can be visualized to be similar to a train system. The train is analogous to the blood and plasma on which the passengers of nutrients, gases, and metabolic wastes are transported to their assigned destinations. These destinations might be any of the various tissues and organs of the body, such as the lungs or brain. The tracks of this body-wide train network are the blood vessels which consist of arteries that carry the blood away from the heart, veins that carry blood to the heart, and capillaries which are the small vessels that are the location for the exchange of the nutrients and waste. The motor of this system is the heart, which serves to provide the force necessary to circulate the blood and nutrients throughout the entire system. Figure 1 shows a simplified version of the entire system as it permeates the entire breadth of the human body [4].
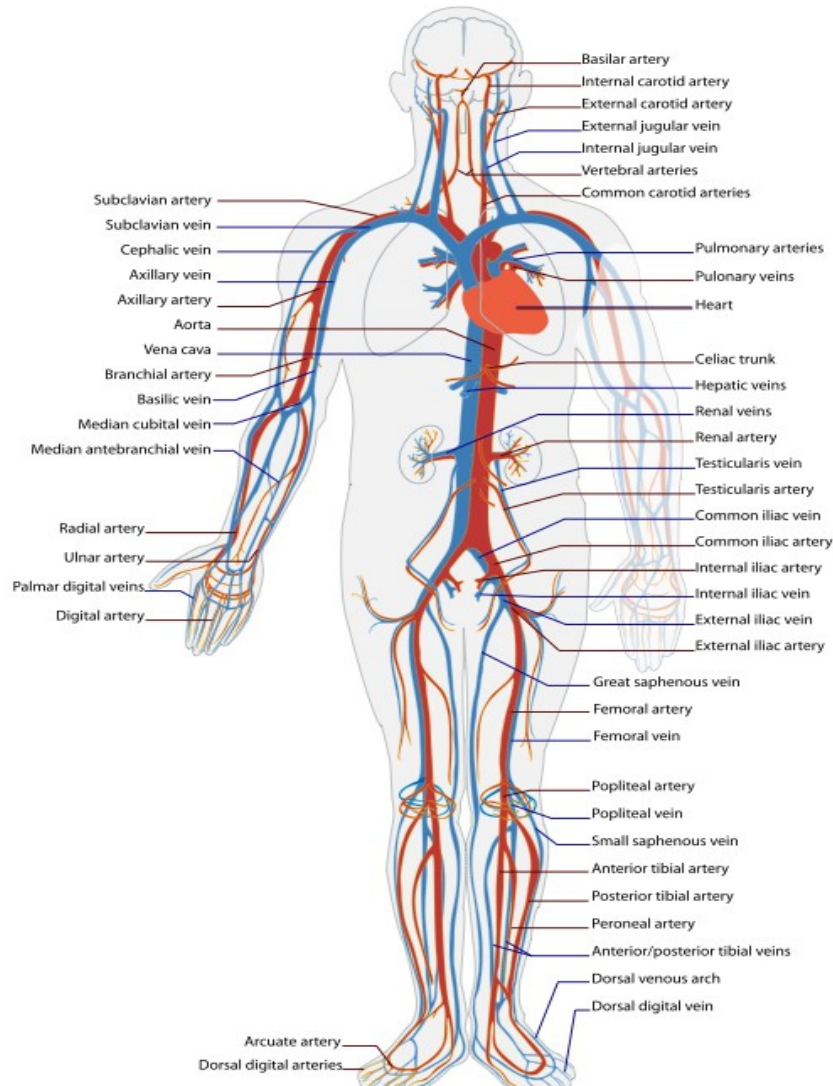
Figure 1: Overview of the circulatory system. Taken from [5].

The heart is a vital piece of the entire system and is an intricately designed pump which serves to circulate blood throughout the system for a human's entire life-span, equating to upwards of over 2.5 billion beats [6]. The heart accomplishes this amazing

feat through a designed honed by millennia of evolution. The heart consists of four

chambers as seen in Figure 2
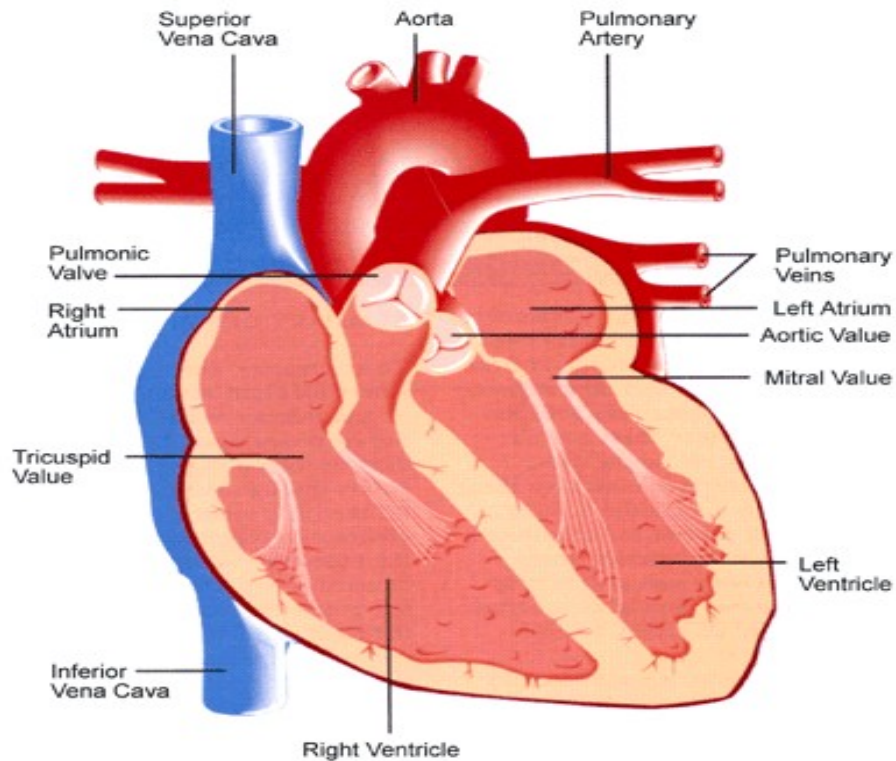
Figure 2: Diagram of the heart. Taken from [6].

These chambers are the left and right atria and ventricles. During a normal heartbeat,

these chambers work in unison to cycle the blood throughout the body. Each heartbeat

actually consists of two stages. The first stage is the systole stage. In this stage, the

ventricles contract and force blood out of the heart. The right ventricle forces oxygen

poor blood to the lungs where waste gases can be exchanged for oxygen. At the same

time, the left ventricle forces out oxygen rich blood fresh from the lungs into the

circulatory system. During the diastole phase, the ventricles relax and receive blood from

their respective atria. The entire process is controlled by the sinoatrial and

atrioventricular nerves which send electrical signals to the muscles of the heart in order to coordinate the contraction of the muscles in a smooth and continuous motion [4].

Each contraction of the heart creates a pressure wave as blood is expelled from the chambers and into the circulatory system. This is because the walls of the blood vessels are elastic and each contraction of the heart muscles causes a distention of these walls as the non-compressible liquid is expelled into the closed system. This pressure wave is transmitted throughout the entire circulatory system and can travel almost fifteen times faster than the actual speed of the fluid in the system. This pulse pressure wave creates the distention that can be manually palpated as one methodology for measuring the pulse. In healthy hearts, this rate of occurrence of pressure waves is a direct measure of the heart rate, although heart abnormalities, such as arrhythmia can create misleading readings because the heart may not create a noticeable pressure wave during some beats [7]. In addition to creating a pressure wave that causes a physical distention of arterial walls, the pulse also creates a heat signature. This is because the blood is pumped from the core of the body and is at a hotter temperature than the outer tissues which are experiencing heat loss due to exposure to the outside elements. Thus hot blood is transmitted to cooler external tissues. Additionally, the physical distention of the artery can create a small heat signature as the heat is transmitted to a wider surface area and then is transmitted to the surrounding tissues. This heat signature forms the basis of the non-contact approach to pulse measurement that is pursued in this thesis work [8].

This leads to the question of where best to try and detect these heat signatures for use in non-contact pulse measurement. Certain conditions would need to be met for a

16

vascular structure to qualify as a good measurement location. First, the structure needs to be sufficiently large to create a heat distribution which can be detected with the current available instrumentation. Second, the vascular structure needs to be sufficiently close to the surface of the skin to produce a distinct a heat signature. And finally, the structure would need to be consistently present in the vast majority of the human population in order to make the overall methodology applicable for use in general medical and industrial practices. A study of the human anatomy shows three distinct locations which would be ideal for viewing the heat signature. The first location would be the radial artery of the arm as this artery can often be found close to the surface of the skin. However, this artery can also be covered by larger amounts of tissues deposits depending on the subject's body type of the subject. Another possible measuring site is the carotid artery which runs up through the neck and into the head. As with the radial artery, this vessel can also be covered by varying amounts of tissue depending on the body composition of the subject. The final solution lies in the superficial temporal artery which lies in the forehead region. This artery is an eventual branch off the carotid artery. It benefits from being highly exposed in the majority of subjects and also it tends to be covered by very little body tissue. Anatomy studies of 27 subject show that all 27 of them contained the vascular structure [9] [10]. Thus, the superficial temporal artery meets the necessary requirements of creating a large heat signature in an exposed location which is present in the vast majority of the population. The next exploration lies in the area of acquiring the heat signature.
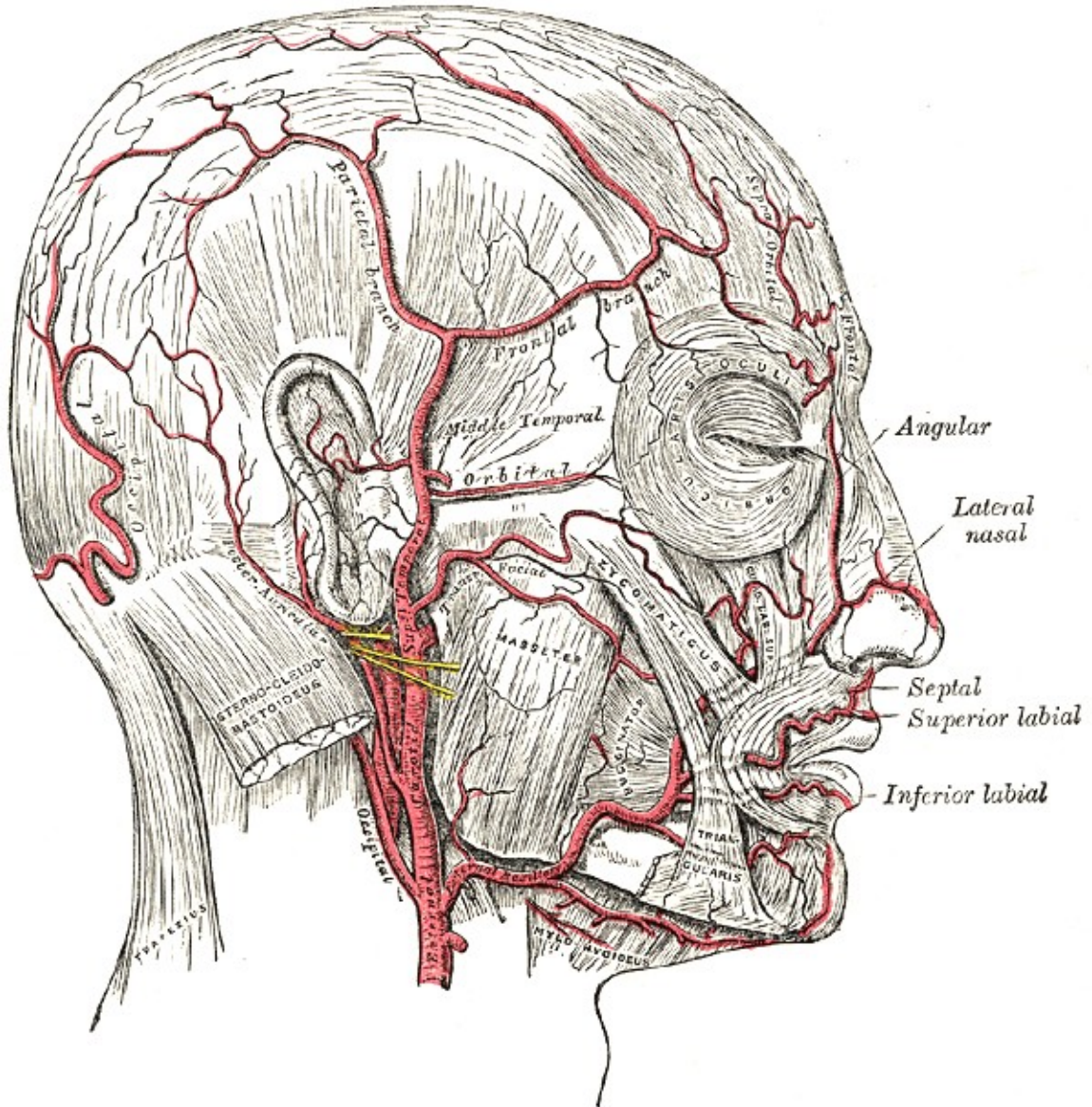
Figure 3: Anatomy of the vascular structure of the face. Note the superficial temporal artery runs next to the ear and branches into two segments in the forehead. The frontal branch is the targeted area for signal acquisition. Taken from [11].

18

2.2 THERMAL IMAGING

Thermal imaging is a novel technology that allows humans to expand their vision into the world of infrared energy. Nature has constrained humans to view their environment by perceiving the electromagnetic spectrum in a limited range commonly called the visible spectrum, however, this is a very small portion of the information that is truly conveyed. The infrared spectrum consists of electromagnetic energy just below the threshold of human vision, with wavelengths spanning from roughly 750 nanometers up to 1 millimeter.
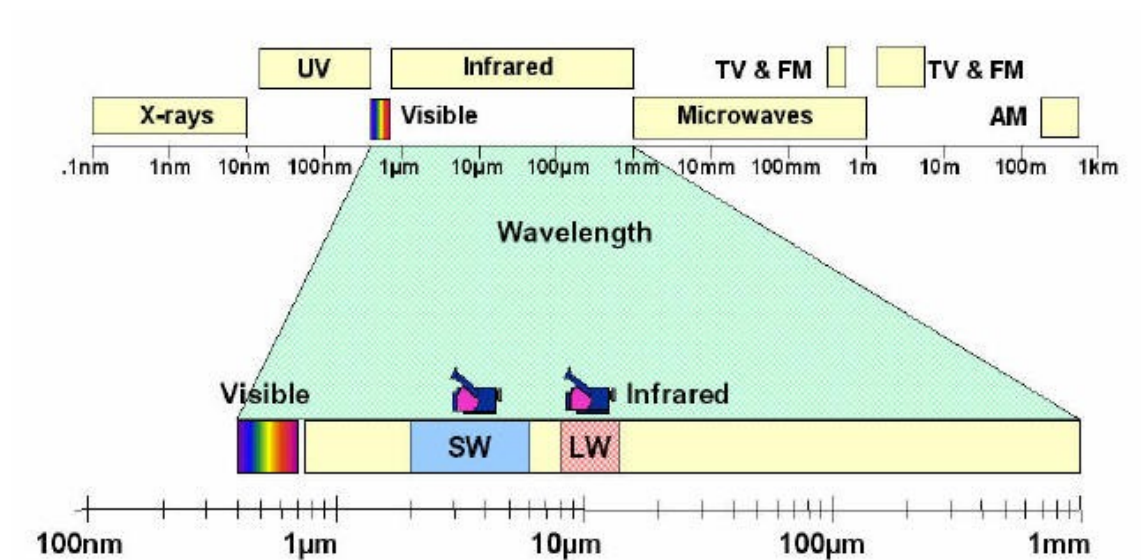
Figure 4: Electromagnetic spectrum showing infrared band . Taken from [12].

The spectrum can be further broken down into narrower bands as illustrated in Table 1:

Table 1: Infrared Spectrum Breakdown

Near Infrared: .75-1.4 um
Short Wave Infrared: 1.4-3 um
Mid Wave Infrared: 3-8 um
Long Wave Infrared: 8-15 um
Far Infrared: 15-1000 um

The full extent of these bands is not always available in certain atmospheres, as the presence of certain chemicals can absorb specific portions of the spectrum. The transmittance of the infrared spectrum can be seen in the figure below. The figure shows that there are several gaps in transmission, where the infrared energy is absorbed by gases in the atmosphere.
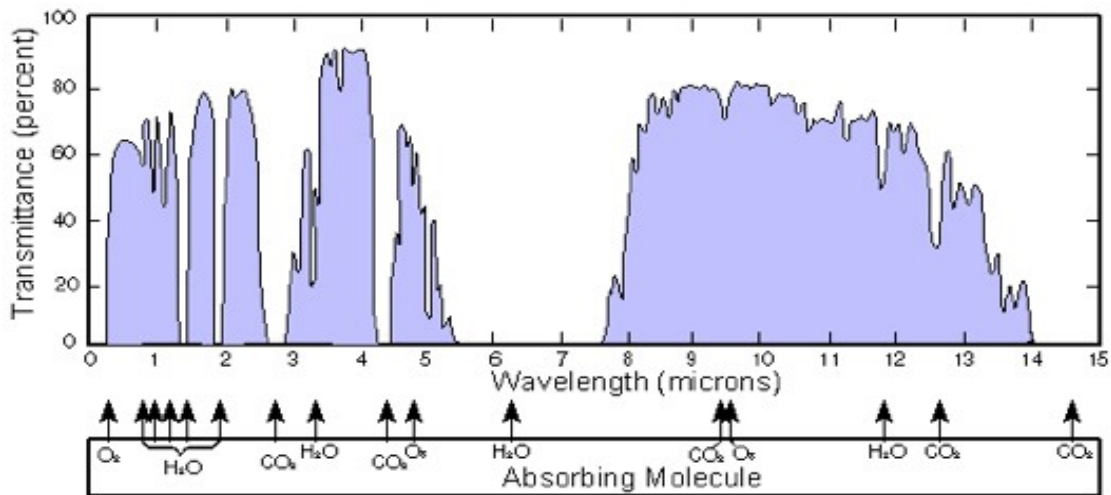
Figure 5: Transmittance of infrared energy through atmosphere. Absorption from certain elements reduces transmittance in specific bands. Taken from [12].

Thermal imaging is useful because all objects radiate infrared energy with a continuum of frequencies that are proportional to the temperature of the object. The spectrum of frequencies emitted is determined by Planck's Law of Radiation as shown below in equation 1:

$$I(\nu, T) = \frac{2h\nu^3}{c^2} \frac{1}{e^{\frac{h\nu}{kT}} - 1} \qquad (1)$$

where I is the spectral radiance of a body for a given frequency, ν, and a given temperature, T. Thus, as the temperature of a radiating body increases, the main frequency of its radiated energy increases as shown in the figure below [13].
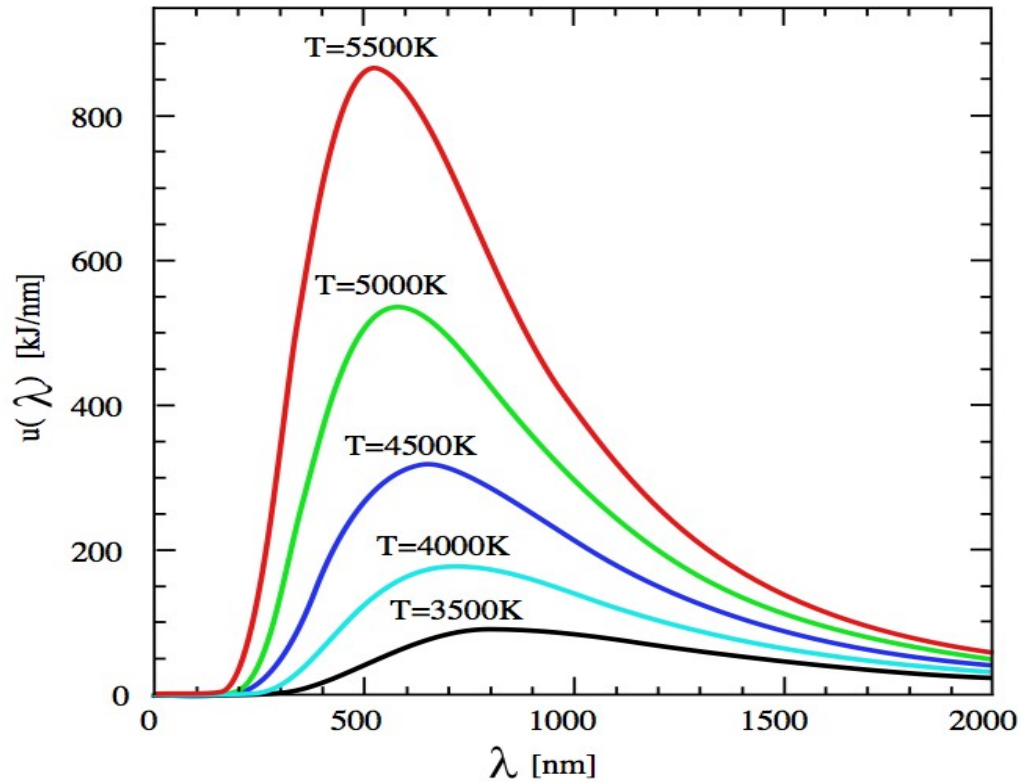
Figure 6: Radiation spectrum of a blackbody at varying temperatures. Taken from [13]

The figure shows that as the temperature of the blackbody increases, the average frequency of the emitted radiation increases as well. Thus, with the proper sensor and environmental conditions, it is possible to detect infrared radiation from any object.

The basic process of thermal imaging is much like digital imaging in the visible spectrum. The purpose of the imaging process is to create a representation of a scene based off of the radiation emitted, reflected, and transmitted by the objects within the scene. Thus, any thermal imaging system must be able to collect and focus the appropriate spectral energy on a sensor that can then analyze the incoming data and represent it in a format appropriate for the human visual system.

There are many varieties of sensors that can be used for the detection of infrared radiation. Commonly called focal plane arrays, or FPAs, these sensor consist of mixtures of materials that make them sensitive to radiation within a specific bandwidth. Some common FPAs are listed with their bandwidths below in Table 2:

Table 2 : Focal Plane Arrays

Indium Gallium Arsenide - Near IR
Indium Antimonide -Mid Wave IR
Quantum Well Infrared Photodetector - Long Wave IR

While the FPAs may be constructed with different materials, they all serve the purpose of receptor to infrared signals, which allows the quantization of incident energy to occur.

Focal plane arrays do not actually directly measure the temperature of an imaged object. Instead, they measure the radiosity of the target. Radiosity is "the infrared energy coming from a target modulated by the intervening atmosphere, and consists of emitted, reflected and sometimes transmitted IR energy" [14]. As indicated by the definition, several components come together to create the radiosity of an object. The first component is emissivity. Emissivity is a measure of how easily an object emits its infrared radiation. Highly emissive objects will readily emit their thermal energy, while objects with a low emissivity will not readily emit radiation. Thus for two objects at the same temperature, the more emissive one will appear 'brighter' to the focal plane array because it emits more radiation. The converse of emissivity is reflectance. This refers to the ability of objects to reflect infrared radiation emitted by other objects around them. Highly emissive objects will have low reflectance, while objects with low emissivity will

23

have a high reflectance. This can be seen in Figure 7, which shows metal cans at

different temperatures. The cans have highly emissive tape on them which is at the same

temperature as the cans. Since the tape is more emissive than the cans, it appears 'hotter'

in the thermograph, while in truth it is simply more emissive. (Orlove, 2003)
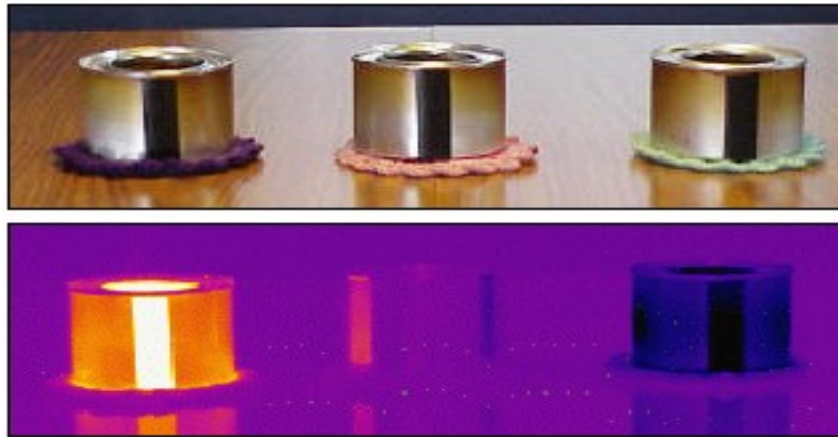
Figure 7: Demonstration of emissivity. The tape and cans are at the same temperature,

however the tape appears brighter or colder relatively, due to it's high level of emission.

Taken from [14].

Since different receptors are sensitive to different bandwidths of infrared radiation

it is important to select an imaging system which is tailored for the infrared emitted by

the subject matter during thermal imaging. Humans naturally emit the bulk of their

radiation in the range of 8-9um, which makes them ideal subjects for long wave IR

sensitive systems. While they can be imaged in other spectral bands, the greatest

sensitivity is available in the long wave IR band [10]. Due to the fact that the radiation is

from internal body heat, this means that human subjects can be imaged even without any

ambient lighting, which is useful for a variety of surveillance and identification purposes.

24

Additionally this property of humans makes it very useful for medical imaging, as the measure of a person's body temperatures can be useful in diagnosing conditions and vital signs.

In humans, thermal images are formed due to the convergence of several factors. First, humans naturally emit radiation in the 8-9 um range due to the fact that core temperatures in humans tend to be a constant 98.6 degrees Fahrenheit. This core temperature is then transferred to the skin through the tissues of the body. Muscles, bone, and fat tend to serve as insulators and keep hotter temperatures near the core, while the circulatory system tends to provide the main source to convect heat to the body's surface where it can be detected through thermal imaging systems [8]. The skin directly above the vessels of the circulatory system tends to be hotter than the surrounding skin because the blood is heated in the core and then carries that heat to the surface of the skin. The skin itself is subject to a variety of factors which can affect its appearance in thermal imaging systems. This is because the skin is susceptible to perspiration for heat regulation as well as thermal influences from the surrounding atmosphere, such as convection. However, in controlled environments noise due to external factors tends to be minimized. A lack of perspiration and convection heat losses will help ensure there is a maximal signal to noise ratio when taking thermal images of human subjects. Despite this though, techniques still need to be created to filter out noise which will invariably creep into any imaging system due to things such as atmospheric disturbances.

2.3 CONTINUOUS WAVELET TRANSFORM ANALYSIS

One of the main challenges in recovering the pulse signal from thermal imagery is the problem of noise. Noise comes from a variety of influences and can greatly corrupt the pulse signal's magnitude, which is often small in comparison to the magnitude of the noise. Factors that contribute to noise include motion, atmospheric interference, and extraneous heat signals due to physiological responses. In order to combat the effect of noise, continuous wavelet transform analysis is utilized in order to find the appropriate pulse signal.

Continuous wavelet transform analysis is a method for analyzing signals. In many ways it is similar to the Fourier transform, however it has some distinct advantages over the Fourier transform. The Fourier transform does an excellent job of indicating the frequencies present in a signal, however it does not indicate where in time those given frequencies occurred. Thus, for any time varying signal, the results of the Fourier transform can be somewhat ambiguous, as there is no indication of when the frequencies occurred. This creates problems when trying to transfer back to the time domain from a set of filtered frequencies in the frequency domain. Unlike the Fourier transform though, the continuous wavelet transform preserves the spatial or time location of the frequencies present in the signal. This makes it ideal for analyzing time variant signals such as pulse.

The methods will by which this is achieved will be examined in the next few pages (Polikar, 2009)[15].

To gather a better understanding of the Continuous Wavelet Transform, we will first look at its formulation and then break it down into its component parts to see how analysis through this method benefits the detection of a pulse signal. The formulation for the Continuous Wavelet Transform is shown in Equation 2:

$$X_w(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t)\psi^* \left(\frac{t-b}{a}\right) dt \qquad (2)$$

Where, $x(t)$ is the signal in time, ψ^* is the complex conjugate of the mother wavelet $\psi(t)$, and a and b are the scaling and translation of the mother wavelet, respectively. The mother wavelet refers to a compactly supported or finite duration oscillating function. (citation needed) There are a multitude of possible mother wavelets such as the ones shown in figures 8-10 below:
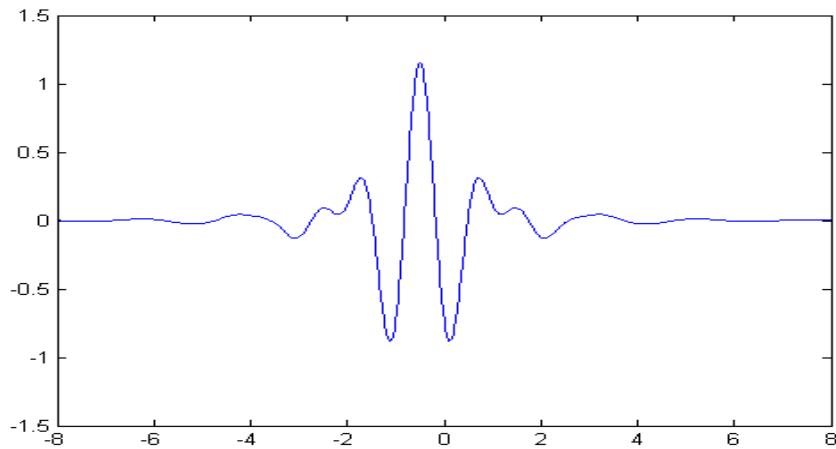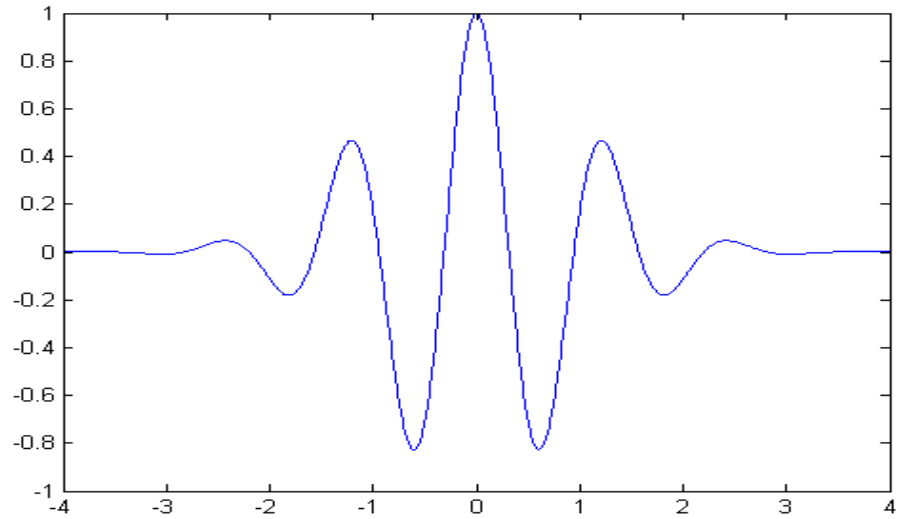

Figure 8: Meyer Wavelet.
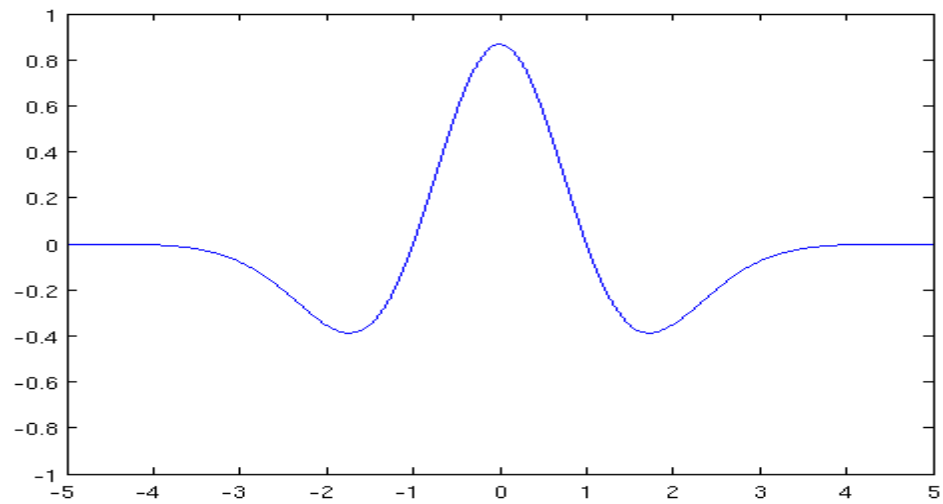
Figure 9: Morelet Wavelet.

Figure 10: Mexican Hat Wavelet.

A little exploration into the idea and mathematics of wavelets is necessary to gain complete understanding of the subject. The main idea of a wavelet is that it can be used to analyze a signal at different scales in order to detect differing features because the

wavelets are compactly supported in time; that is they do not extend to infinity so they are useful for analyzing non-periodic signals. In order to create a meaningful mathematical way of dealing with wavelets, the idea of mother wavelets is introduced. The mother wavelet is the template for the daughter wavelets, which are produced using equation 3:

$$\psi_{s,\tau}(t) = \frac{1}{\sqrt{s}}\psi\left(\frac{t-\tau}{s}\right) \qquad (3)$$

In this equation, the daughter wavelet $\psi_{s,\tau}(t)$ is made up of a version of the mother wavelet, ψ, which has been scaled by s and translated by τ. The scaling factor in front of the equation is for energy normalization [16] [17]. The wavelets themselves have to satisfy two conditions known as the "admissibility and regularity conditions." The admissibility condition is given in equation 4 and from this equation, one can determine that equation 5 is also necessary.

$$\int \frac{|\Psi(\omega)|^2}{|\omega|}d\omega < +\infty \qquad (4)$$

$$\int \psi(t)dt = 0 \qquad (5)$$

This condition is necessary because square integrable waves can be used to "analyze and reconstruct a signal without loss of information" [16]. Equation 5 indicates that the wavelet has an area of zero underneath the curve which indicates that the wave indeed does have a wavelike feature of oscillation. The regularity conditions are imposed to cause the signal to decay rapidly with scale, which is desirable for the purposes of analyzing signals with a compactly supported wavelet. [16]

For the purposes of the thesis, the mother wavelet that was used for analysis was the Mexican hat wavelet which is formulated using equation 6, which is simply the negative of the second derivative of the Gaussian equation:

$$\psi(t) = \frac{1}{\sqrt{2\pi\sigma^3}} \left(1 - \frac{t^2}{\sigma^2} \right) e^{\frac{-t^2}{2\sigma^2}} \qquad (6)$$

It is the mother wavelet that will be scaled and translated to produce the daughter which will be used to analyze the signal.

Each mother wavelet is controlled by the scaling and translation factors to produce daughter wavelets. Scaling serves to compress or dilate the mother wavelet, while translation moves the central location of the wavelet in time as seen in figures 11-12 below:
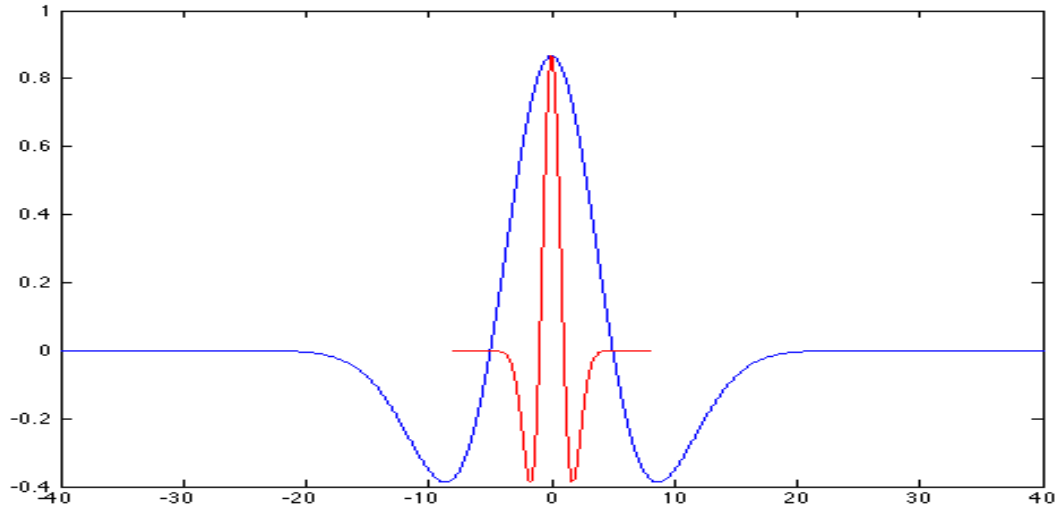
Figure 11: Dilation of the mexican hat wavelet, a = 5.
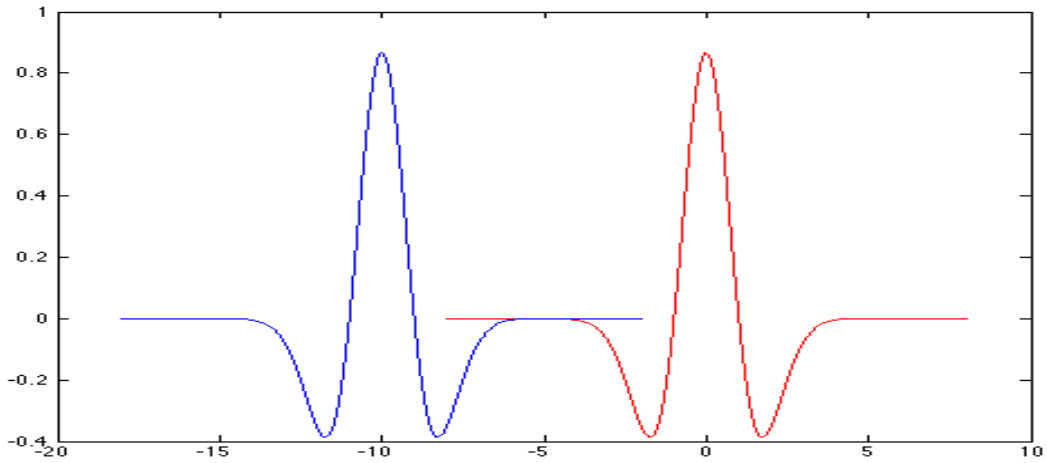
*Figure 12: Translation of the mexican Hat Wavelet, b = 10.*

Thus, it can be seen from the formulation that for every point in time, a scaled and translated version of the mother wavelet is convolved with the signal. If the signal is

31

similar to the wavelet with a given scaling and time translation, it will produce a strong

response which is captured as X_W. The location of X_W is given by *a* and *b* in 2-D space

and indicates what scaling and translation are used in order to generate the given

response. It can be seen from simple observation that scale correlates to the frequency of

the signal. The relationship between scale and frequency is an inverse one, with higher

scales correlating to lower frequencies and lower scales correlating to high frequencies.

In many ways, this formulation is much like the Fourier transform. However, one

of the major differences between the two formulations is that the continuous wavelet

transform utilizes a wavelet which has a frequency component within some discrete

portion of time. In contrast, the Fourier transform convolves the input signal with infinite

sinusoidal signals, whose frequency components are not localized to a specific time scale.

The difference between the two methodologies is what gives the continuous wavelet

transform its specific strengths. By utilizing a finite duration pulse, the location of

specific frequency components can be determined in time. This is in contrast to the

Fourier Transform, where the strengths of frequency components in a signal can be

measured, but their location in time cannot be given. Thus, a signal that has a time-

varying frequency component can not be accurately reconstructed using the Fourier

transform, while it can be simulated with the coefficients derived from the continuous

wavelet transform.

Consider the example of the signal plotted below in figure 13:
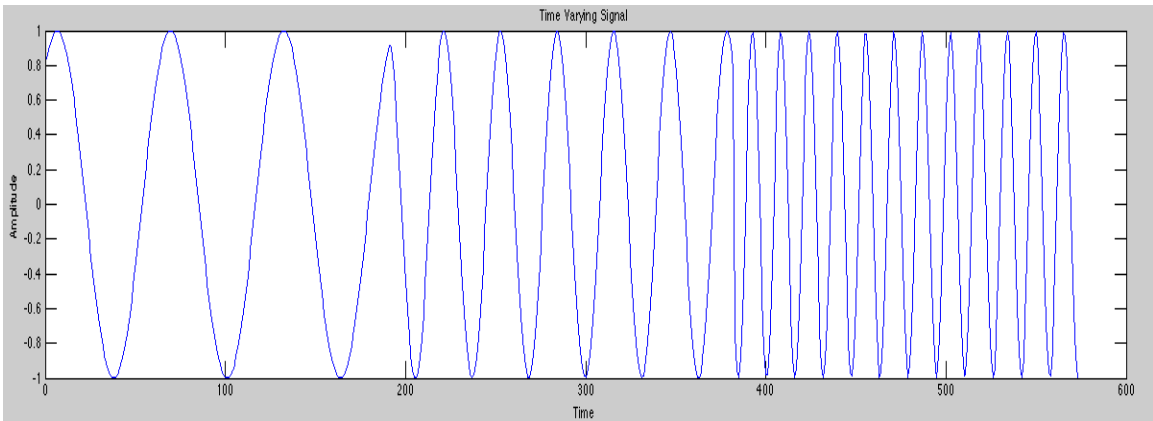
Figure 13: Signal with time varying frequency.

It can be seen from observation that the signal's frequency changes over time. An

examination of the Fourier transform reveals the following data presented in figure 14:
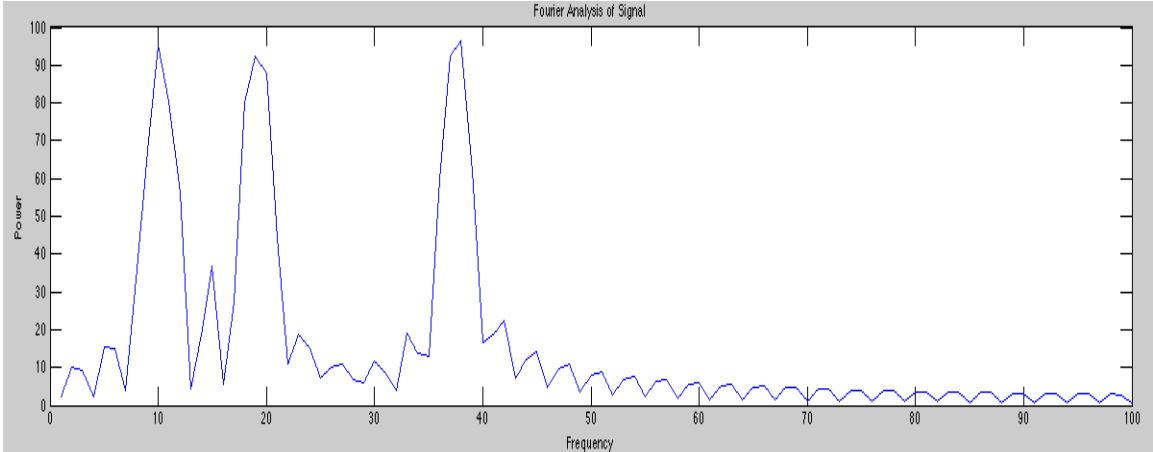
Figure 14: Fourier transform of signal in figure 13.

The data reveals the three distinct frequencies that occur in the signal, however it does

not reveal when those frequencies occurred. Thus, a reconstruction of the signal from the

absolute coefficients derived from the Fourier transform yields the following signal, shown in figure 15:
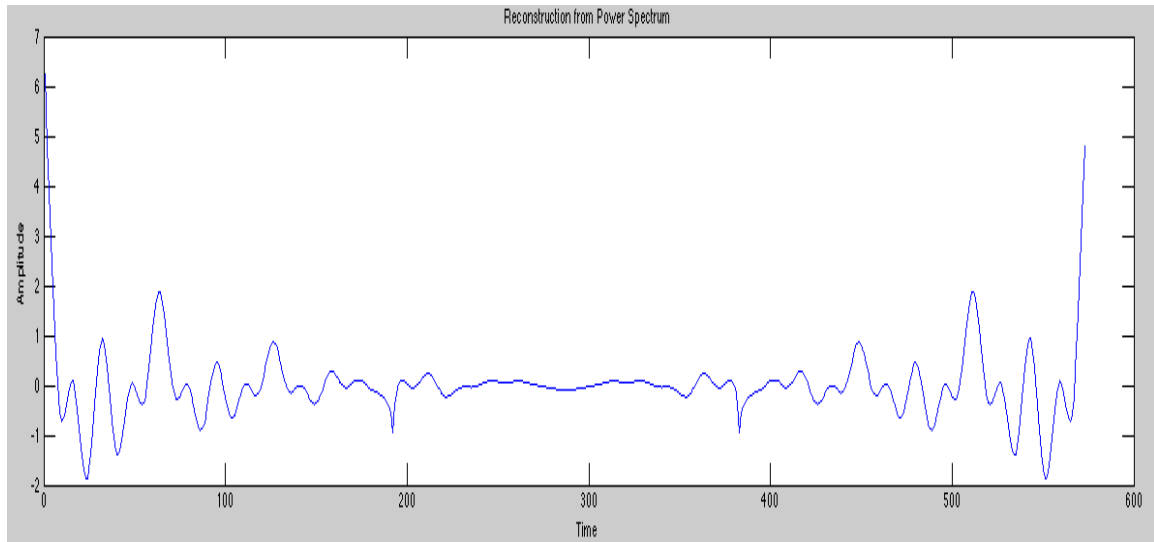
Figure 15: Reconstruction of signal from inverse Fourier transform and Fourier

Transform in figure 14.

These results are obviously different from the original signal and this is due to the inability of the Fourier transform to localize frequencies in time. However, a look at the results of the continuous wavelet transform in Figure 16 reveals far more information. The transform shows the three distinct frequencies as well as their location in time. This is seen as the three distinct peaked sections that shift over time. The left most section corresponds to the lower frequency (higher scales) portion of the signal and then transitions to higher frequency (lower scales) over time. This additional information allows for the accurate reconstruction of the signal as seen in figure 16 below.
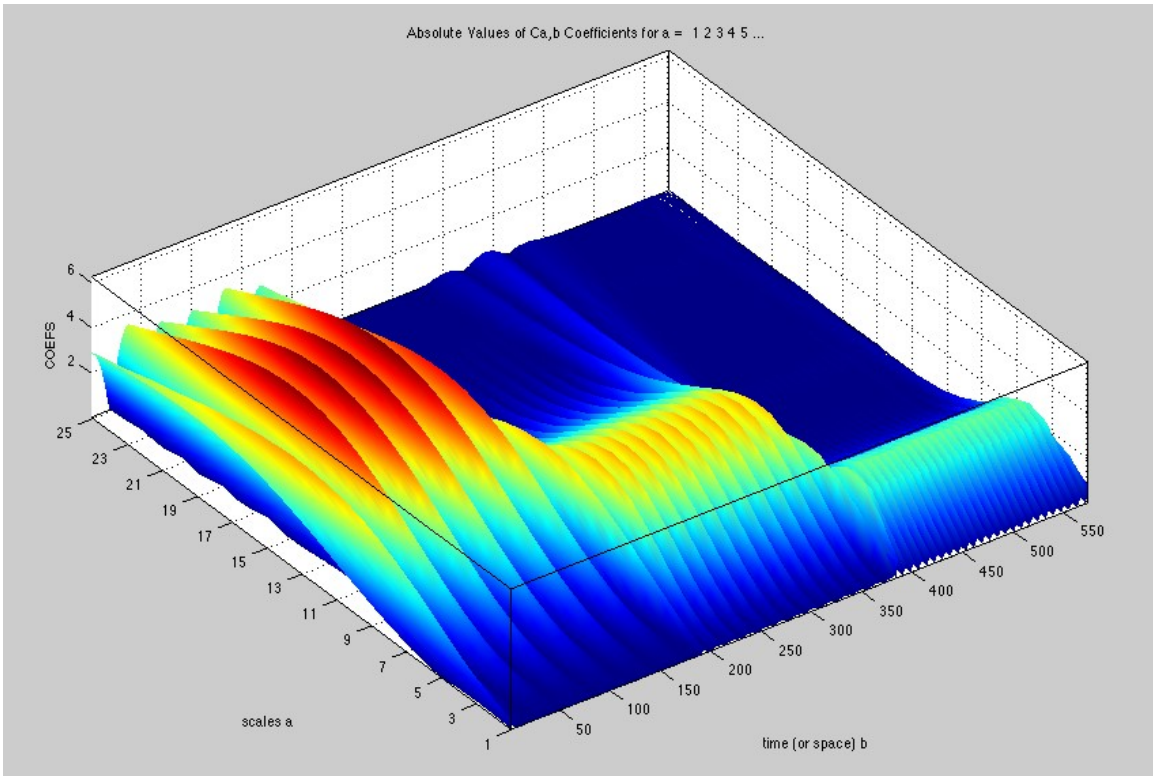
Figure 16: Continuous wavelet transform of signal showing changing frequencies over time.
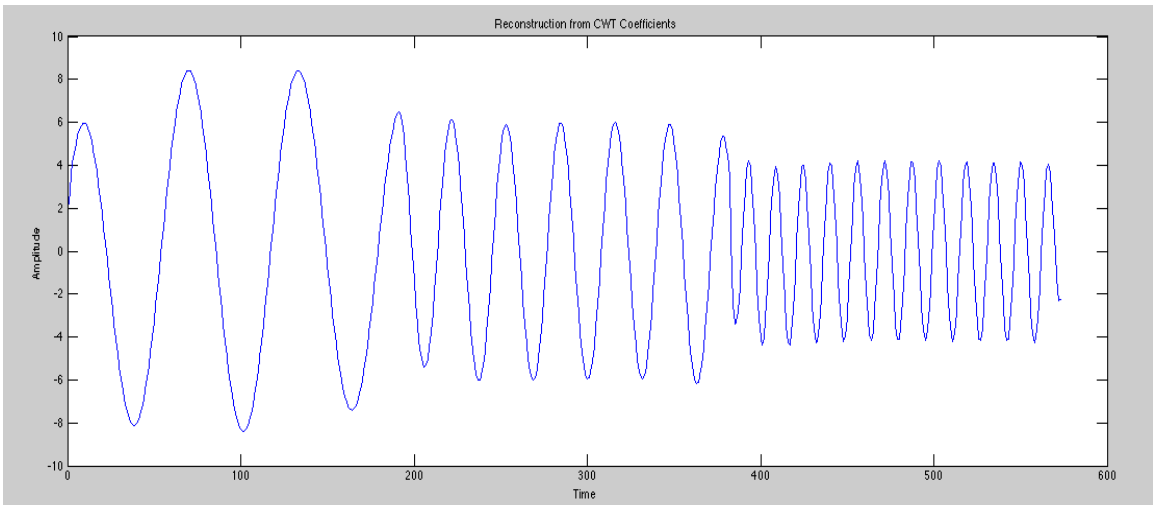

Figure 17: Reconstruction of signal from inverse continuous wavelet transform.

35

There is an intrinsic trade-off present in the formulation of the continuous wavelet transform in terms of the accuracy of the temporal location of the frequencies. Higher scales, which correspond to lower frequencies, lead to better resolution in the frequency domain, but poorer resolution in the time domain. This is because the lower frequency wavelets require a larger temporal basis and thus reduce the accuracy in pinpointing when in time a frequency component occurs. Higher frequency wavelets need much smaller temporal basis which allows for a more precise time location. However, there is more ambiguity as to what the exact frequency value is as the scales get closer and closer to zero; this is what causes the loss in frequency resolution. This trade-off is well-suited for the vast majority of signal analysis applications though, since signals of interest are usually composed of low frequency terms for the majority of the time, with higher frequency pulses coming at specific time segments [18].

2.4 SUMMARY

This chapter provided an introduction to a variety of topics that will be used in the solution of how to achieve non-contact pulse measurement. It first introduced the cardiovascular system and the elements that make up the pulse. Then, the concepts related to thermal imaging were introduced, including the spectral components of the infrared band as well as the concepts of emissivity and how thermal energy is generated and modulated in humans. The final concept introduced in this chapter was the idea of Continuous Wavelet Transformation. The chapter developed an understanding of the concept as well as it's strengths and weaknesses in terms of its capabilities in the signal processing field. Additionally, the chapter looked at the mathematical background behind the construction and use of wavelets.

CHAPTER III: SYSTEM DESIGN

3.1 NON-CONTACT PULSE MEASUREMENT PROCESS

The process of performing non-contact pulse measurement is a multistep process that covers a wide variety of mathematical and scientific topics. This section will give an overview of the challenges in pulse measurement and a broad overview of the techniques used to overcome them. Subsequent sections will delve deeper into the specifics of these techniques to provide a deep understanding of the entire process.

A general overview of the process is shown in figure 18 below:
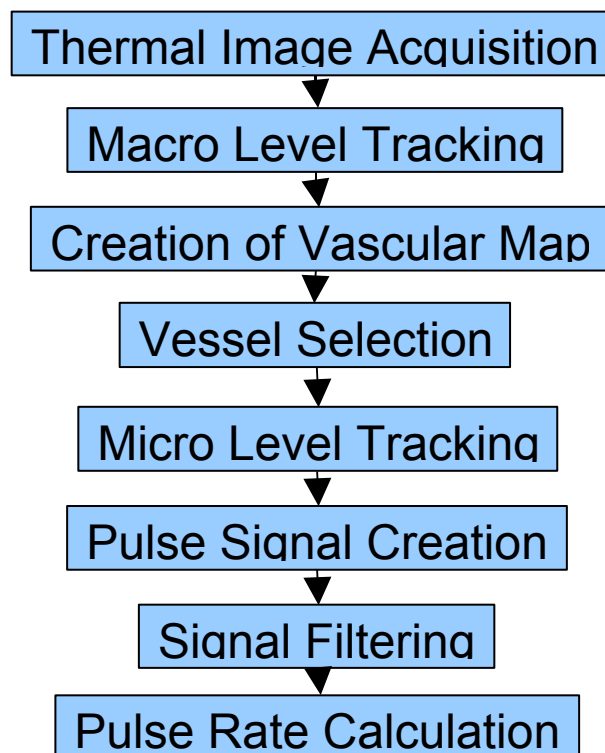
Figure 18: Non-Contact Pulse Measurement Process.

The first step in the process is image acquisition followed by generalized facial tracking, which is termed "macro-level tracking". The ultimate goal of this work is to develop a system that can improve the conditions of those in long term medical care. Therefore, the system must be able to track the area of measurement as the patient exercises a certain degree of mobility necessary for comfort. In order to accommodate this need, a system of particle filters using game theory is employed to determine the location of the region of interest. In the case of pulse detection, the general area of interest is the forehead, as the vessel of interest, the superficial temporal artery lies in this region. Therefore, the forehead tracker is used to locate this region in each frame and serves as an input into the next portion of the procedure.

The next step in the process is vessel tracking or "micro-level tracking". Within each frame, the specific vessel of interest needs to be located and tracked throughout time. This is of extreme importance because even slight movements due to natural muscle contractions and breathing can destroy the measured signal. Micro-level tracking ensures that each measurement in time is taken at the precise location inhabited by the vessel. The tracking is accomplished through a few steps. The first step is the manual selection of the vessel of interest by drawing a line over the general vicinity of the clearest branch of the superficial temporal artery. The program then employs a snakes algorithm to contour the measurement points to the shape of the vessel. This is necessary because the vessel usually follows a unique and tortuous route through the forehead region and therefore it is necessary to employ a more complex shape than just a simple line to map the path of the vessel. After the vessel has been segmented, the micro-level

tracking process employs a routine to determine the best scaling, translation, and rotation that will yield the nearest fit from the initial frame to all following frames.

The next step in the process is to extract the raw and unfiltered version of the arterial pulse signal from each measurement point along the vessel. This is done by using the tracking information to provide the location of points of interest then by taking the thermal measurements as the raw signal. This raw signal is subject to a variety of noise influences, however it holds the true data within it.

The final step is to filter the raw data to extract the true pulse rate. This can be accomplished using the techniques available with the continuous wavelet transform. Additionally, the combination of multiple measurement points helps yield a higher level of accuracy as the bulk of the measurement points should yield a consensus as to what the pulse truly is. A final Fourier analysis of the filtered signal is then used to give an average pulse rate over the time frame of image acquisition.

3.2 MACRO-LEVEL FACIAL TRACKING

The goal of macro-level facial tracking is to provide a general search area for the micro-level tracking routine. Therefore, it must provide an input region which includes the vessel of interest which is to be measured. This is accomplished by creating an algorithm which tracks the entire forehead region. The algorithm used is a version of the coalition tracking network developed and explained by Dowdall *et. al* in the 2007 paper, "Coalition Tracking" [19].

The coalition tracking method uses a network of simple particle filters to achieve high performance tracking without extreme complexity. Each particle filter consists of a simple sub-sampling block as seen in figure 19 below:
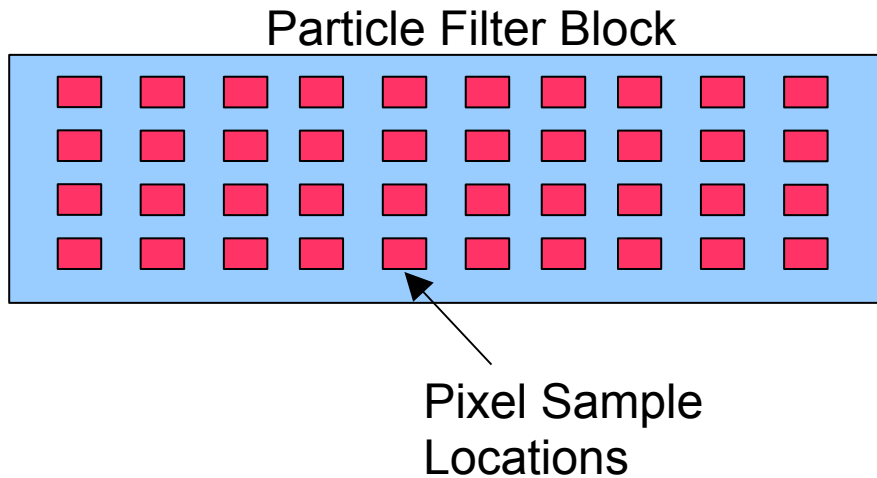

Figure 19: Individual particle filter used for forehead tracking.

A small exploration of the mathematics of particle filters will reveal the theory behind their physical workings. To begin, the general goal of this process is to estimate some process X with only access to a process Y. In other words, we are attempting to know the state of the entire image X with only the sampling of pixel values which can be considered to be the process Y. Mathematically, this can be thought of as attempting to determine the distribution given in equation 7:

$$\{p(x_n|y_{1:n})\}_{n>0} \qquad (7)$$

In essence we are attempting to determine the probability distribution of the state of process X at a specific time n given the observation of process Y from the beginning of the process until time n. As time progresses, we continue to use the newly acquired values from the process Y to update our understanding of process X. In this way, the process is an offshoot of typical Monte Carlo algorithms which use random samplings to guide the program to an approximation of a density function for a process at a given time using the previous samplings over time. In order to determine the distribution given in equation 8, we have to break down the steps mathematically. To determine tracking motion, we need to determine which set of motion parameters has the highest probability of occurring at a given time. We can represent this as seen in equation 8:

$$\pi_n\left(x_n \mid x_{1:n-1}\right) = p\left(x_n \mid y_n, x_{n-1}\right) \qquad (8)$$

Here, the probability of a given motion set occurring given that a previous history of motion sets have occurred is equal to the probability of that set occurring given the current sampling of pixel values and all past motion set histories. To approximate this probability, we use a set of weights, which are issued based on the closest match between pixel subsamplings from observation y_{n-1} to y_n given some set of possible motions X_n. The highest weighted motion set is then considered to have the maximum probability of occurring and becomes the new motion set, x_n [20].

Physically speaking, each individual particle filter serves to take a sub-sampling of the pixel values at it's location. In addition, each particle filter has a set of parameters which include position, scale, and orientation. Each particle filter tries to track the pixels within it by a simple error-minimization algorithm. The algorithm attempts to minimize the error between pixel values between each frame by finding the translation, rotation, and scaling which yields the minimum difference between sampled pixel values. This method works well for objects where lighting conditions are held constant, as the introduction of shadow or new light sources can cause error for this simple tracking routine. Since thermal video relies only on radiation emitted from the subject and not on external lighting, this makes it an ideal visual system on which to perform this style of tracking.

Taken individually, each particle filter is of only moderate accuracy due to its simplified tracking design. However, the use of a network of particle filters can serve to improve the overall accuracy of the tracking process. In this instance, the network of particle filters is controlled by factors determined via the use of game theory. Game

theory is an attempt to "mathematically capture behavior in strategic situations, in which an individuals success in making choices depends on the choices of others" (Game Theory, 2009). This type of mathematical modeling has been explored since the early part of the 20th century and has been very successful in describing a wide range of situations, ranging from economic conditions to political scenarios. In it's simplest form game theory explores the actions, known as strategies, that a set of players makes for a given set of conditions. Generally, each action has a result which can be thought of as payoff that will be dependent on the state of the other players. Each player seeks to find the action that will maximize his or her payoff [19].

This formulation is a perfect fit for the natural problem posed by having a network of particle filters. Each individual particle filter is considered to be a player in the game and seeks to find the set of translations, rotations, and scalings that will maximize its payoff in terms of overall tracking performance. In order to further improve performance, groups of high performing trackers will form coalitions and propagate their influence to other, lower performing trackers. This ensures that accurate trackers will keep lost trackers in a position such that they can regain tracking capabilities.

In order to determine the best coalition, a set of four criteria are established and used to quantify the tracking performance of all possible coalition combinations. The first criterion is known as the template match score. This score can be calculated using : equation 9.

$$\alpha_{C_k}^{(t)} = \frac{1}{k} \sum_{i=1}^{k} \alpha_{m_i}^{(t)} \qquad (9)$$

where, t is the time, C_k is the given coalition consisting of the set of k individual particle filters, with each individual particle filter being designated by m_i. The value α_{mi} refers to the template match score of an individual tracker and is given by equation 10.

$$\alpha_{\mathbf{p}}^t = \sum [\mathbf{T_p(c)} - \mathbf{T(c^0)}]^2 \qquad (10)$$

where $\mathbf{T(c)}$ are the values of the pixel samples at the transformation which minimizes the error between those samples and the values at the previous time step, given in $\mathbf{T(c^0)}$. Essentially, this is just a distance measure between the pixel samples at time t and time t-1.

The second criterion is the geometric alignment score, $\beta^{(t)}$. This score is used to ensure that coalitions have the best possible geometric alignment between trackers by ensuring that each tracker in a given coalition gives similar translation, scaling, and rotational corrections for the entire network.

The third criterion is the inter-frame projection agreement score, $\gamma^{(t)}$. This score is very similar to the geometric alignment score, except it rewards geometric alignment over time instead of between different particle trackers. The score calculates the alignment between the projections of a tracker at time t and time t-1. Again, good tracking performance is usually indicated by a high correlation between alignments and the score rewards those coalitions that have highly aligned trackers over the course of time.

The final criterion is the inter-frame membership retention score, $\delta^{(t)}$. In general, well performing coalitions tend to retain their members and this criterion is used to reward those coalitions that have constant membership over time. Therefore, coalitions that have the same members are awarded additional points, whereas coalitions that are swapping points do not receive the scoring benefit and are less likely to be viewed as winners in the 'game' scenario.

All possible combinations of coalitions are considered in terms of the previous scoring programs. The coalition that maximizes the payoff is considered the 'winner' and the translation, scaling, and rotational factors from that specific coalition are applied to all possible trackers in the network. The results of this process can be seen in the figure below, where a box tracks the forehead region as the subject moves across the field of view of the camera.
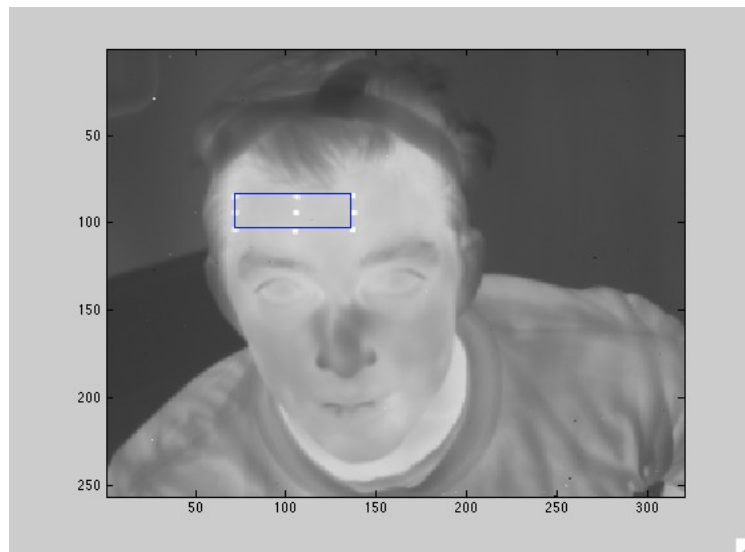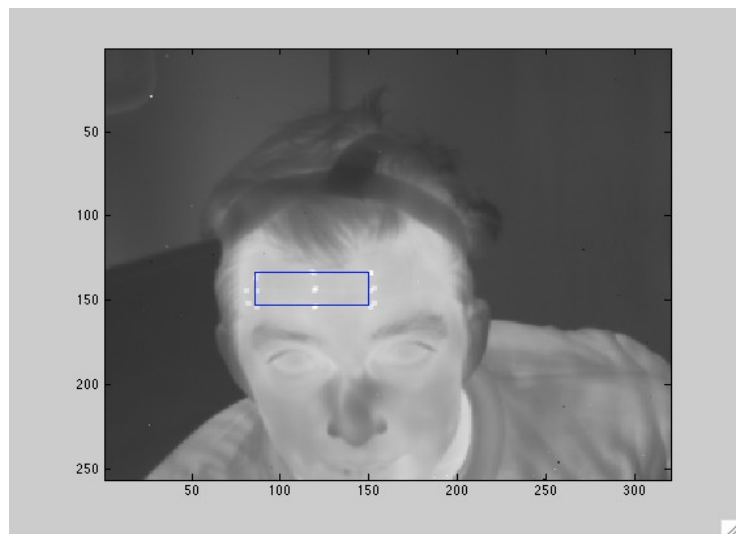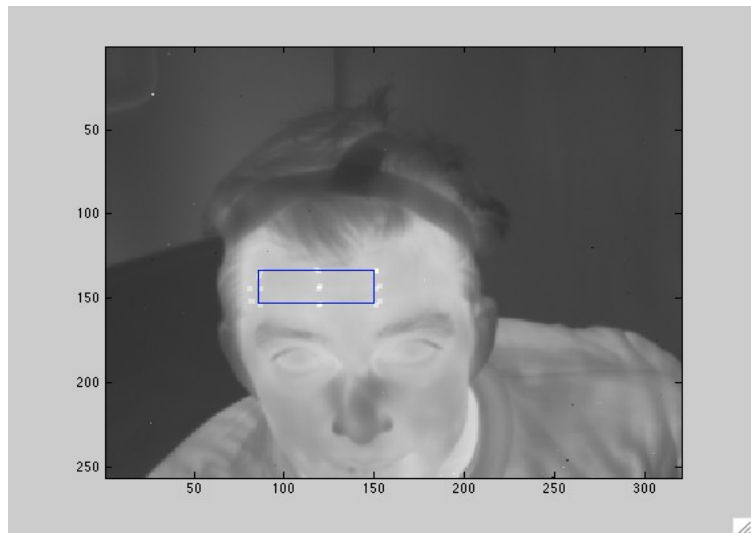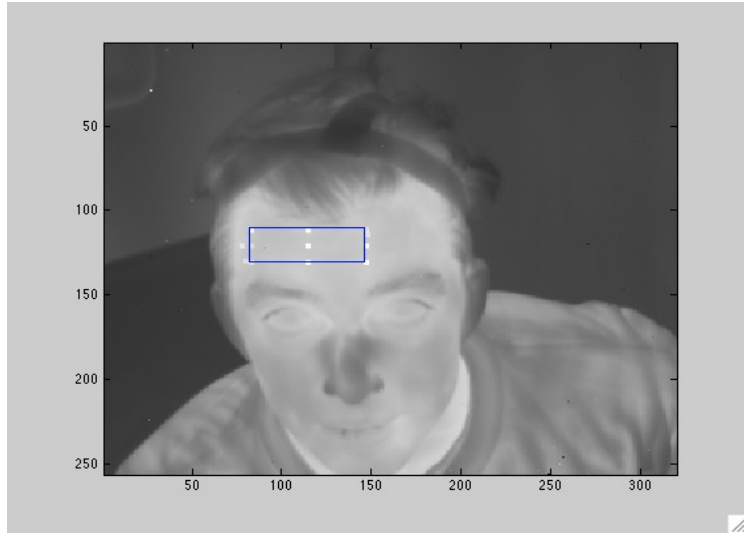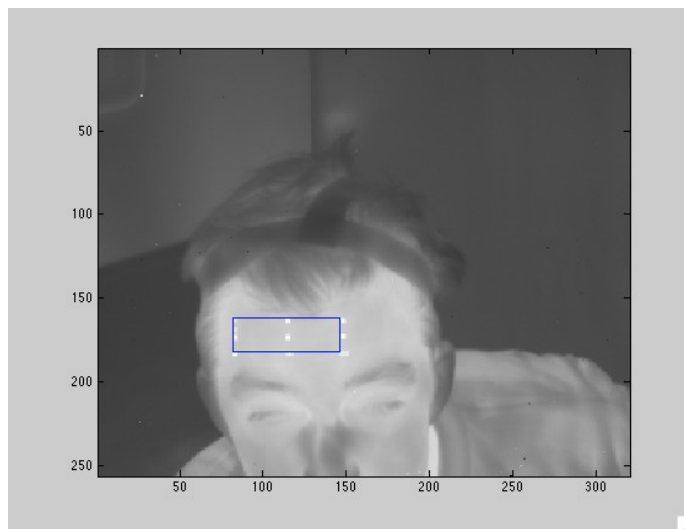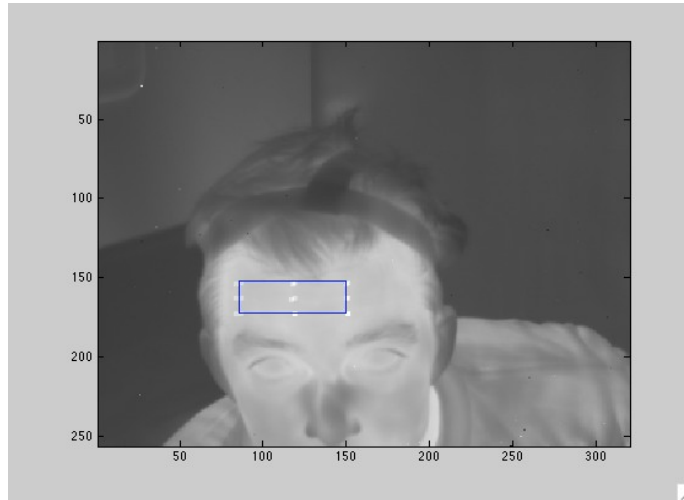
Figure 20: Macro tracking example. Shows the progression of the tracking block (blue) over the course of 30 seconds of thermal imaging as it tracks the forehead region.

3.3 MICRO-LEVEL VESSEL TRACKING

Macro-level tracking serves to give a large-scale search space for finding the required vessels of interests. Once these desired vessels have been selected through an initial manual segmentation though, a much finer tracking system is needed to ensure that all measurements are taken as accurately and precisely as possible. To meet this need, micro-level tracking must be implemented

There are three main steps to the micro-tracking process: vascular mapping, segmentation, and tracking. The starting step is the creation of a vascular map. This map shows the location of the vascular structures as derived from the thermal images. The creation of the vascular map is in itself a multi-step process. The first step is the segmentation of skin tissue from either background or non-skin tissue such as hair. Since the input into the micro-scale tracking step will consist of a frame filled entirely by the forehead, segmentation of the head from the background is unnecessary. Therefore, simple thresholding segmentation can be performed to remove the pixels that show hair, since hair will be cooler than skin tissue and can be easily and quickly segmented with this methodology. If this step was not taken, the hair would show up as vascular structures in the following steps.

49

The next step in the vascular mapping process is the transformation of the image into the blood perfusion domain. This transformation essentially enhances the visibility and contrast of the vessels and should cause areas that experience a large amount of blood flow to show up 'hotter' on the thermal image. This transformation takes the form of a simple gamma transformation which enhances contrast changes in the hotter portion of the image while suppressing change in the colder portions of the image. This helps reduce noise related to cooler thermal process that should not be related to the hot effects of the pulse [8].

After these preprocessing steps have been accomplished, the bulk of the vascular algorithm can be performed. First, the image must be smoothed. However, the image cannot be indiscriminately smoothed as this would destroy the edges that result from the vessels themselves. Thus, anisotropic filtering is employed to filter the image while maintaining edges. Anisotropic filtering is useful because it performs 'intra-region' smoothing, 'as opposed to inter-region smoothing' which is commonly found in filters such as Gaussian blurring [21]. This preserves the edges developed by the vessels while ridding the image of noise due to other processes. After blurring has been performed, white top-hat segmentation can be performed. White top-hat employs a combination of 'erosion and dilation operations' to segment out bright or 'hot' objects from an image. This final step produces an image that shows the center line of vessel passages that can serve as a map. The map can then be utilized in the segmentation and tracking portions of the process [21].

The second step is the segmentation portion, where an accurate estimate of the shape and position of the vessel of interest must be created in order to proceed onto the tracking portion of the process. The most basic solution to this problem is to manually pick points out along the pathway of the vessel of interest in order to provide a structural and positional model of the vessel. However, this process can be time consuming, so an easier solution was devised using a semi-automated method. This method requires the user to select the general vessel of interest by drawing a line along the length of the artery. The line is then conformed to the vessel's pathway via a deformable contours style algorithm.

Deformable contours are an important concept within image processing and are commonly referred to as snakes. They are especially important for image segmentation purposes, making the subject ideally suited for the problem of segmenting out a vessel of interest. In this simple version of the concept, the model to be deformed consists of a basic line. The line is acted upon by various forces which are allowed to deform the line into a curve. The forces can be grouped into two categories: external and internal forces. External forces consist of the forces generated by the image on the line itself and are easily modeled as the gradient of the image at the location of the line. Internal forces can be used to create certain characteristics within the line as it deforms, such as curvature and bending properties. In this case, the internal forces were used to try to keep the line at its original length as it deformed, thus preventing it from expanding to infinity or contracting into a single point.
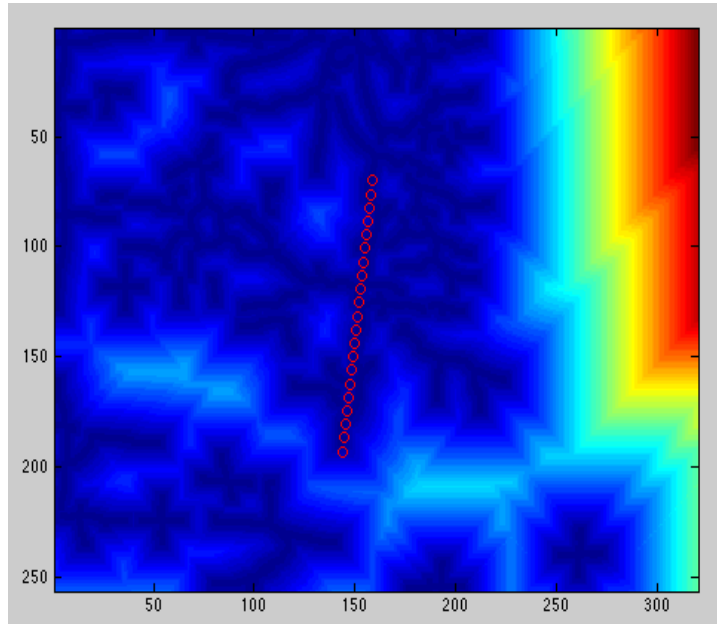
Figure 21: Manual selection of artery.
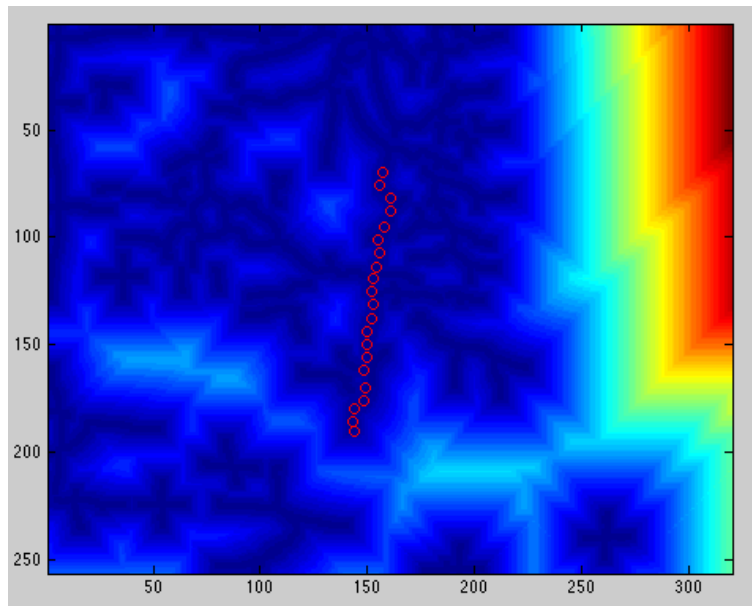
Figure 22: Artery segmentation after snakes algorithms.

Once the artery segmentation has been performed, the second part of the process is initialized. This is the actual tracking portion of the procedure and it relies on a few basic assumptions. The first assumption is that the tracking procedure must compensate for translational, scaling, and rotational effects. The second assumption is that despite the translation, scaling, and rotation of the imaged subject, the vessel of interest itself will still maintain a constant relative shape. Thus, the initial segmentation should be able to be matched to some similar vascular shape at a future time so long as the apparent translation, scaling, and rotation effects are accounted for.

In order to do the micro-level tracking, the initial segmentation is transformed by a set of possible translations, rotations, and scalings. For each transformation, the error between the initial segmentation and the vascular map at time t are compared. The transformation that gives the least amount of error then becomes the specified transformation for time t. Figure 23 shows an example of the measuring points tracking the vascular map at some random time, t.
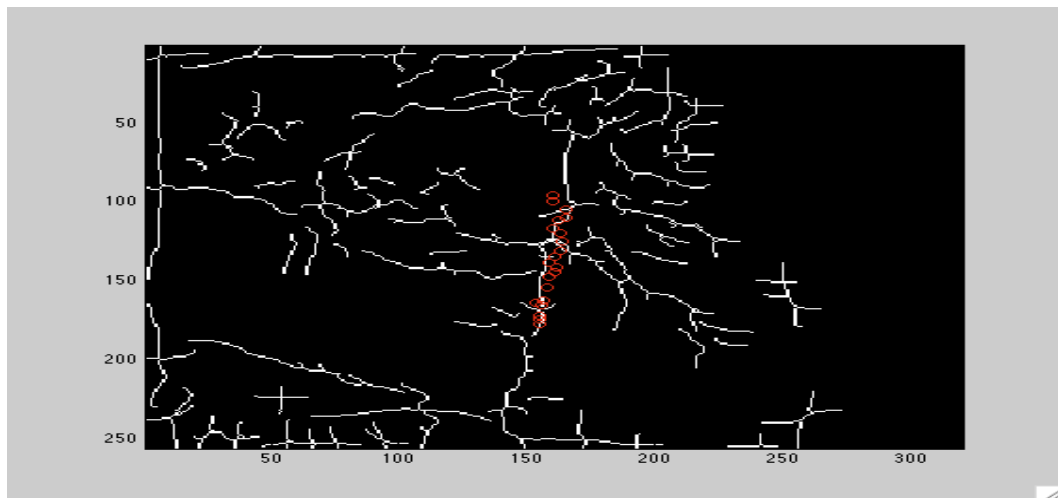
Figure 23: Vascular tracking points (shown in red) on vascular map.

53

3.4 ARTERIAL PULSE SIGNAL CONSTRUCTION AND FILTERING

Once the preprocessing segments have been completed, the true task of pulse detection can be approached. In this portion of the project, the goal is to extract the pulse signal and quantify the pulse rate. This portion has two main subject areas, the construction of the raw arterial pulse signal and the filtering of the detected signal to remove noisy components which are not related to the true pulse signal.

The first portion, consisting of the construction of the raw arterial pulse, is relatively straightforward. The previous steps constructed a set of initial measuring positions which have an x and y location in the thermal image. This set of points will be denoted as L_n with $n = \{1 : Number\ of\ measuring\ locations\}$. Additionally, each set of points has a corresponding transformation which consists of a scaling, rotation, and translations for each step in time. This will be denoted as $T_n(t)$. Finally, the images essentially form a three dimensional array with x and y dimensions and the third dimension consisting of time. Therefore, a specific pixel will be denoted as $I(x,y,t)$. Putting these components together, we can formulate the raw arterial pulse signal, $AP_n(t)$ as shown in equation 11:

$$AP_n(t) = I(T_n(t) * L_n, t) \qquad (11)$$

54

Essentially, this can be visualized as taking a single pixel location and gathering it's values over time while compensating for the motion artifacts naturally introduced during the image acquisition process.

The second part of this task is the far more challenging aspect of filtering the raw signal. Corruption of the arterial pulse signal can occur from a variety of sources. Motion artifacts can be a major source of noise, even with the compensatory tracking measures in place. If the measurement locations are not consistently in the same position relative to the vascular network, they may pick up on different components of the pulse wave and thus, add noise to the signal. In addition, the environment can be a source of noise to the signal as well. Air currents can cause convective cooling which can be another source of noise interference. Finally, other sources of infrared radiation in the imaging environment can be reflected from the skin. Fortunately, the emissivity of human skin is roughly .98 [22]. Since reflectance is the converse of emissivity, this means that the reflectance of thermal energy from other sources by the skin is relatively low and not likely to introduce considerable noise into the process.

Therefore, the challenge of this project is to separate the noisy non-periodic elements from the periodic pulses due to the redistribution of heat to the surface of the skin with every beat of the heart. One of the more robust approaches for filtering this type of signal is the application of the continuous wavelet transform. As described in the section continuous wavelet transform Analysis, the CWT method is very useful for the analysis of time-varying signals. The results of the transform yield a detailed view of the spectral components of the signal over time which can then be used for filtering.

Once the CWT of the signal has been acquired, the art of filtering can take place. An exploration of the transforms of various signal types revealed patterns that allow filtering to take place. Pulse type signals tend to have high energy components that run from high to low scales in the CWT with the high energy corresponding to the peaks of the signal in time domain [23] [24]. On the other hand, extraneous noise is typically higher frequency and manifests itself as very low scale signals which do not consistently run from high to low scales. Therefore, the first step of the filtering process is to remove the low scale components of the transform which serves to filter out the high frequency noise. The next step of the process is to remove components which do not traverse from high to low scales as does the typical pulse signal. This is accomplished by examining each column in the two dimensional CWT transform and finding the maximum. If the maximum is not above a certain threshold it is considered to be too weak to be a pulse signal and the entire column is set to zero. The minimum threshold is the average of the entire CWT data set. If the maximum is above this threshold, the algorithm then examines the coefficients from high to low scales until it finds a minimum below the threshold and zeros the column out after that minimum. In this way components with a strong tendency to contain energy from high to low scales for a given position in time are preserved, while transient signals and noise are filtered out [23] [24].

Once these filtering steps have been performed, the signal can be reconstructed from the filtered transform, yielding a signal whose frequencies should be more closely related to the heart rate than the original unfiltered signal. This signal is then analyzed with Fourier analysis to see what the dominant frequencies are over the time span of the

56

acquired signal. While these steps reduce a large bulk of the noise, the Fourier analysis can still be somewhat inconclusive for a single point of measurement. To solve this problem, the entire process is performed on multiple measuring points along the length of the vessel of interest. Due to the propagation of the pulse itself, the initial signals cannot be combined in the time domain because the pulse peaks could occur at different times as the pulse travels along the length of the vessel. However, a difference in the time domain becomes a simple phase shift in the frequency domain, which is removed when the absolute value of the frequency is determined. Thus, the frequency components of the multiple measuring positions can be combined without problem. By averaging these frequency readings together, the overall level of accuracy is improved as a smaller range of dominate frequencies is revealed.

To further improve the level of accuracy, the averaging of the frequencies is weighted by the apparent performance of each signal measurement point. A simple averaging without weighting could potentially corrupt the answer by providing equal weighting for measurement points that returned both clear and noisy signals. In order to quantify the performance of a signal, the concept of periodicity is used. Periodicity, as it's name indicates, is a measure of how close a signal is to being truly periodic. A true pulse signal will be highly periodic even if the pulse rate varies slightly over time. Conversely, a signal that is composed of mostly noise elements will have little evidence of strong periodicity. Therefore, signals that are highly periodic should contain more true signal than ones that are not highly periodic. To quantify what the periodicity measure is, the following algorithm as proposed by Chekmenev in his 2008 dissertation is used:

- Determine number of local maxima above a set threshold in the signal

- Determine the average distance between peaks for a perfectly periodic signal as: $D_{avg}=(t_1-t_n)/n$, where t_i is the time location of the first maximum and t_n is the location of the last maximum and there are n maxima

- Determine the difference between each consecutive maxima as $L_n = t_n-t_{n-1}$, for all n

- Set the periodicity score as $PM = \sum |L_n - D_{avg}|$, for all n.

Thus, it can be seen that highly periodic signals will have small differences between the consecutive differences between the maxima and the desired distance for a truly periodic signal. Conversely, signals which are not highly periodic will tend to acquire larger and larger values in the periodicity measurement score. Therefore, it is desirable to have a lower periodicity measurement score, with the best possible score being zero. Signals with low periodicity measurement scores are considered to be better indicators of the pulse signal and thus the signals' corresponding frequency components are given a higher weighting in the averaging process. This further increases the overall accuracy of the returned pulse measurement.

The combination of all of these techniques serves to provide a method for filtering the raw data that is input into the system and returns a value that indicates the average pulse rate over a predefined period of time. The next stage involves designing an experiment which can be used to test and validate the methodologies described above.

3.5 SUMMARY

This chapter gives an overview of the design of system with the end goal being non-contact pulse measurement. The chapter gives a high level overview of the entire process and then first delves into an introduction to large scale or macro-level tracking via the use of a network of particle filters. It next explores how to create a vascular map in order to select a region of interest for signal acquisition. Thirdly, the chapter explores the use of a micro-level tracking program to compensate for tiny, involuntary motions while acquiring thermal images of the subject. The last section of the chapter shows the process used for acquiring the raw signal and then processing it in order to extract the final estimate of the pulse rate.

CHAPTER 4:

EXPERIMENTAL DESIGN

In order to test the accuracy and capabilities of the pulse measurement process, an experimental setup was created with the goal of creating a low noise environment for acquiring controlled thermal images, as well as ground truth vital signs, for validating the experimental results. This section will describe the setup and the progression of experimental designs which were utilized to explore the concepts in this thesis.

The cornerstone to the entire project is naturally the thermal imaging system, since this piece of equipment is utilized to acquire the pulse signal itself. The thermal imaging system that was utilized was the Indigo Phoenix Camera System with Real Time Imaging Electronics as seen in figure 24 (Product #420-0011-007, REV 120, S/N 0013). In order to capture the thermal radiation emitted by the human body, the long wave IR camera was employed. This camera detects radiation in the range of $8.0 - 9.2$ μm. The camera uses a quantum well infrared photodetector (QWIP) as the focal plane array in order to detect incoming thermal radiation. It also uses a Stirling cycle cooling system to cool the camera equipment down to roughly 70 Kelvin. This helps reduce noise incident on the focal plane array from hot components in the camera itself which might normally emit radiation at room temperature. The readings from the focal plane array are digitized and sent to the computing system via a frame grabber. Talon Ultra 5.2 image acquisition

software, provided by Image-Pro Plus (Ver. 4.5.1.27) is used to grab the required number of frames, with the frame rate being 30 frames per second.

Figure 24: Phoenix Indigo LWIR camera.

The camera itself is mounted on a gantry which rotates around a central platform, on which the subject can be seated for image acquisition. The gantry has black drapes on the opposite side of the imaging system to prevent noise from background objects interfering with the signal. Likewise, foam padding was placed over the warmer electrical equipment in order to prevent thermal reflections from providing another source of noise.

The center platform of the gantry is equipped with a chair so the subject can be seated during the image acquisition process. In order to provide stability for initial

testing, a chin rest was also employed as shown in Figure 25; this helped in reducing

motion artifacts. However, later algorithmic improvements have allowed for the free

seated acquisition of thermal images without any additional stabilization devices.
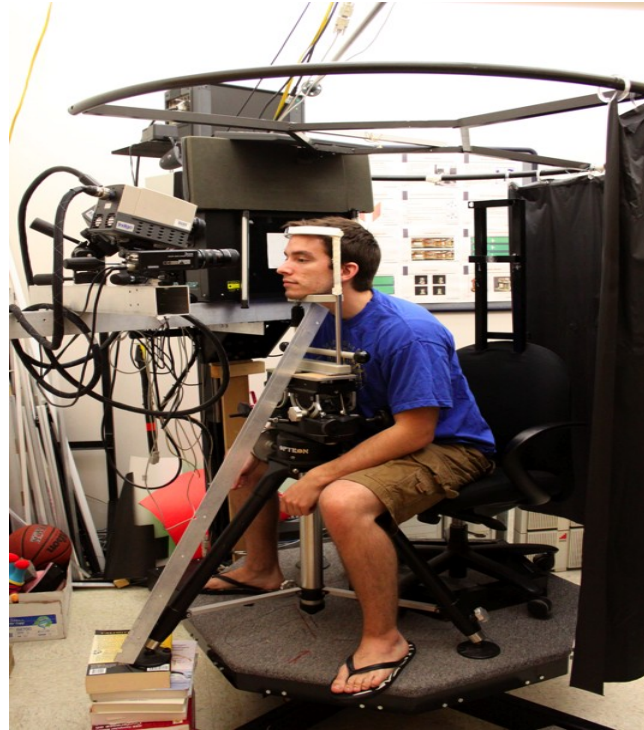
Figure 25: Image acquisition system with chin rest for stability.

The system utilizes the Advisor Vital Signs Monitor to determine the ground truth signal

with which the computed pulse rate is compared. The Advisor Vital Signs Monitor offers

the capability of measuring pulse rate through either electrocardiography (ECG) or

photoplethysmography (pulse oximetry). The photoplethysmography device employs a

finger clip which can easily be attached to the subject and then the heart rate can be

measured over the period of image acquisition. This device has a usable range of 30-250

beats per minute with an accuracy of the larger of either +- 2 beats per minutes or +-2%

of the heart rate. The ECG has a slightly extended range of 30-300 accuracy and an identical accuracy. Due to the somewhat invasive nature of attaching electrodes to the subject's chest and abdomen region with the ECG, the pulse oximeter (photoplethysmography) method was employed for ground truth observations.

The process for experimentation is as follows. The subject is seated in a comfortable position in the chair in the center of the gantry. The thermal camera is then focused on the upper left corner of subject's forehead and the pulse oximeter is attached to his or her finger. The subject is then asked to remain as still as possible, while 1024 frames of thermal data are collected at 30 frames per second. 1024 frames are collected to provide an easier analysis using the Fourier transform as described in the previous sections. This corresponds to 34.13 seconds of thermal imaging data. The pulse reading is noted at 0, 10, 20, and 30 seconds, and those readings are averaged and rounded to the nearest beat to determine the overall ground truth value for accuracy analysis. The only exception to this procedure was for the first five test trials where the subject used the chin rest without straps for stabilization. However even under this scenario the motion tracking software was used to alleviate the noise due to small motions that inevitably occur.

The collected data is then input into the processing algorithm. Once the thermal data set is input into the algorithm, a vascular map is automatically generated for each frame. From the initial frame, the user then selects a general vessel of interest by drawing a straight line over the vessel of interest in the graphical user interface. The snakes algorithm is utilized to precisely segment the vessel from the user indicated area

of interest by contouring the straight line segment to the tortuous route of the vessel. Once this is accomplished, the tracking routine is utilized to compensate for scaling, translation, and rotation effects through time. Then the pulse signal is extracted and filtered to provide an overall estimate of the pulse rate for the duration of the acquired signal. This is then compared to the ground truth measurement and an accuracy measurement in generated. The Complement of the Absolute Normalized Difference (CAND) measure, in equation 12, was used to determine accuracy:

$$CAND = 1 - |Ground\ Truth - Measured| / Ground\ Truth \quad (12)$$

The image acquisition procedure was performed on a variety of subjects in the Computer Vision and Image Processing lab over the course of several weeks. Subjects where imaged multiple times and at a variety of points throughout the day in order to acquire a broad base of possible normal heart rates. Traditionally, normal heart rates are characterized as being between 60-80 beats per minute and the heart rates of the test subjects ranged from 58 to 90 beats per minute according to the Advisor Vital Signs Monitor, which represents a good range of test data. There were 13 total subjects. 11 were male and 2 were female. In terms of ethnicities, 6 were Caucasian, 1 was African American, 1 was Asian, and 6 were Middle Eastern. Ages ranged from 18 to 38. This study is approved for human subject under IRB 457 study titled "Biometric Study of Heart Rate, Breathing Rate, and Gait" at the University of Louisville.

4.1 SUMMARY

Chapter 4 develops the methodology used for validating the system design presented in chapter 3. The acquisition equipment and setup are described, as is the process for the experimental setup. Finally, the chapter describes the method for calculating the accuracy of the results. In addition, this section describes the data protocol in terms of the IRB process.

CHAPTER 5:

5.1 RESULTS

The initial results of the experimentation show great promise for the procedure. An initial experiment was performed on nine different individuals with multiple data sets being taken for each individual. A follow up experiment was then conducted on a larger pool of 31 individuals with one data set for each person. As described in the section titled "Experimental Design," each person was imaged while ground truth data was gathered using a pulse oximeter. The results of the experimentation are shown below in table 3 and table 4 as well as the accuracy as calculated with the Complement of the Absolute Normalized Difference (CAND).

Table 3: Experimental Results for Group 1

Subject #	Trial #	Ground Truth (BPM)	Calculated (BPM)	CAND
1	1	70	61.8	88.3%
1	2	70	65.9	94.1%
1	3	80	70.0	87.5%
1	4	76	78.2	97.1%
1	5	79	67.1	84.9%
2	1	75	67.1	89.5%
3	1	85	86.4	98.4%
3	2	58	74.1	72.2%
4	1	81	74.1	91.5%
5	1	75	74.1	98.8%
6	1	81	78.2	96.5%
6	2	60	65.8	94.2%
7	1	60	61.8	97.2%
8	1	90	90.6	99.3%
8	2	85	61.8	72.7%
9	1	88	78.2	88.9%
9	2	86	82.4	95.8%
10	1	70	74.1	94.1%
11	1	64	74.1	84.2%
12	1	76	82.4	91.6%
13	1	74	74.1	99.9%

Table 4: Experimental Results for Group 2

Subject #	Trial #	Ground Truth (BPM)	Calculated (BPM)	CAND
1	1	79.2	70.0	88.4%
2	1	82.1	90.6	89.7%
3	1	67.5	70.0	96.2%
4	1	64.5	57.7	89.4%
5	1	64.5	74.1	85.1%
6	1	61.6	65.9	93.0%
7	1	79.2	74.1	93.6%
8	1	79.2	90.6	85.6%
9	1	105.6	107.1	98.6%
10	1	70.4	61.8	87.8%
11	1	67.5	65.9	97.7%
12	1	73.3	78.2	98.9%
15	1	82.1	86.5	94.7%
16	1	67.5	65.9	97.7%
17	1	82.1	78.2	95.3%
18	1	82.1	86.5	94.7%
19	1	58.7	57.7	98.3%
20	1	99.7	94.7	95.0%
21	1	64.5	74.0	85.3%
22	1	87.9	94.7	92.4%
23	1	70.4	78.2	88.8%
24	1	76.3	74.1	97.2%
25	1	73.3	74.1	98.9%
26	1	82.1	78.2	95.3%
27	1	96.8	94.7	97.9%
28	1	96.8	86.5	89.4%
30	1	64.5	78.2	78.7%
31	1	61.6	78.2	73.0%
32	1	70.4	78.2	88.8%

The experiment yielded an average CAND accuracy of 91.2% with a standard deviation

of 8.00% for group 1. This means that the average error from ground truth in beats per

minute was approximately 6.69 BPM. For group 2, the average CAND was 91.9% and

the variance was 6.30%. The similarity of results in both groups indicates the strength of

the results.

In order to judge the value of these results, it is beneficial to compare them to

similar works. Within the Computer Vision and Image Processing Lab, previous work on

the challenge of non-contact pulse measurement had yielded an average accuracies of

86.3%. Additionally, this methodology was further constrained by the fact that the

subject had to be stabilized with the use of a chin rest as demonstrated in the initial

experimental setup section. These results also compare well with the work of Garby, et.

al., published in 2007 which performed non-contact pulse measurement through infrared

imaging. That study utilized a Fourier-based signal filtering method with motion tracking

and achieved an average CAND accuracy of 88.5% across 34 trials. Thus, the results

indicate that the outcomes of this work stand well in light of the body of work available

for comparison.

Additionally, the results need to be judged in light of their statistical significance.

To determine this, the Pearson product-moment correlation coefficient is calculated via

equation 13:

$$r_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{(n-1)s_x s_y} \qquad (13)$$

Since certain subjects had differing numbers of trials, only the first trial for each subject was considered in the determination of statistical significance. This yields a correlation coefficient, r, equal to 0.76 indicating strong correlation between the ground truth and the calculated values. To test the significance of this, equation 13 is used to determine the t value. If this t value is larger than 3.055, the null hypothesis that r = 0.0 can be rejected and it can be claimed that this is statistically significant at p = 0.01.

$$t = r\sqrt{\frac{n-2}{1-r^2}}$$ (14)

The resultant t value from equation 8 is 3.878, indicating that the results from group 1 are indeed highly significant. Likewise, group 2 has an r value of .83 which yields a t value of 7.73 which indicates the results are highly significant.

The results showed that the filtering works to bring out the frequency content related to the pulse rate, however visual inspection of the signal before and after filtering does not show noticeable information. Figure 26 shows the raw signal on top and the filtered signal on bottom.
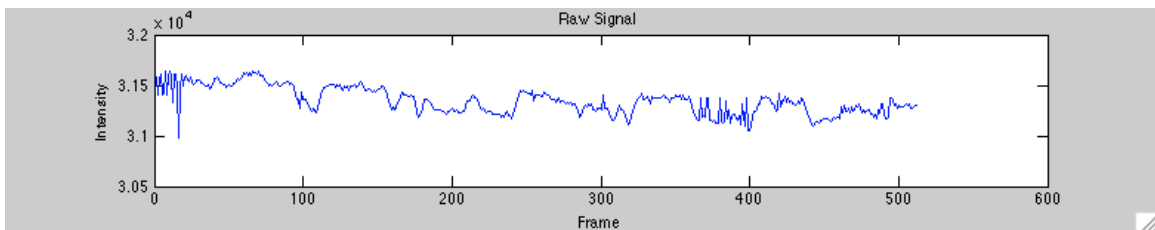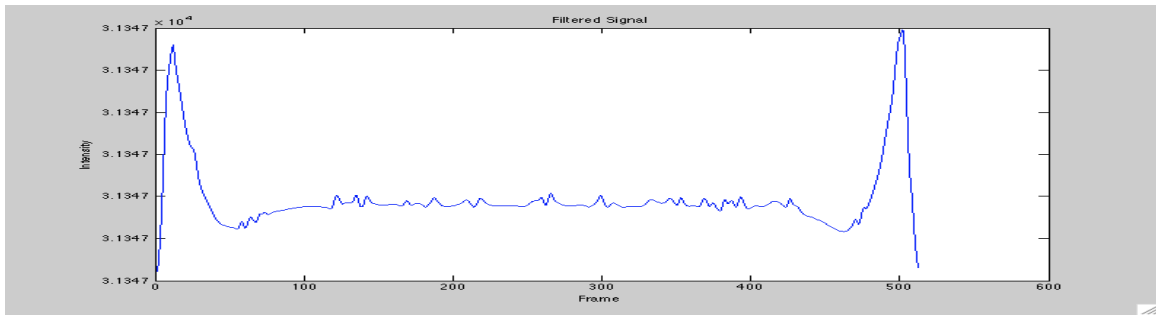
Figure 26: Top shows raw pulse signal and bottom shows filtered signal.

The visual appearance of the raw signal does indeed show it is noisy, while the filtered signal is misleading because the discontinuities at the beginning and ending of the signal produce large spikes. However, the frequency content is detectable and indeed in the filtered signal we can see the smaller spikes which are correlated to the actual pulse in terms of frequency.

The results also show that measured values tend to cluster around distinct measurement values as a result of the filtering method. Since the final analysis after filtering relies on Fourier analysis, the accuracy of the returned measurements is limited by the sampling frequency and number of samples. In the case of this thesis, the sample rate is 30 Hz and 512 samples are used for calculating the heart rate. This means the best resolution of the final Fourier analysis is the sample rate divided by the number of samples or 3.5 BPM. This effectively yields 'binning' of results, such that the returned values lie in bins spaced roughly three beats per minute apart from each other. This binning effect could be reduced by including more samples in the data for the given sample rate. However, this would extend the measurement time and increases heart rate variability, which is less useful for the application of this technology to medical

71

applications. Another solution would be the use of more advanced technology that had a faster sample rate. This would be the more advantageous solution, however, it would require the use of expensive technology not currently available for use. Thus, in the end it is concluded that this was an acceptable level of error in the trial measurements to ascertain the viability of the work.

Another aspect that was revealed by the results was the importance of proper vessel selection. Some vascular networks were the result of noise due to hair and other data that managed to remain unfiltered throughout the algorithm. Selection of those false vessels yielded vastly inaccurate results. However, selecting the proper vessel along the superficial temporal artery brought about a much closer match with the ground truth value. An examination of vessel selection choices in the Results section of the Appendix shows the high level of variability in vessel selection. Consistently, the choices appear to be along a branch of the Superficial Temporal Artery, however the exact location along the artery differs from trial to trial. Likewise, the location of the artery itself has a high level of variability within individuals and contributes to the diverse array of measurement locations.

Additionally, the results indicate that the methodology works even with the small, inevitable motions that could often cause noise during the signal acquisition process. This represents a definitive step forward from the previous work in this area and brings the technology one step closer to feasible application. This demonstrates that the tracking technology utilized in the project is useful in decreasing the amount of detected noise in the thermal imagery. Investigation revealed that, for people who were not constrained

but requested to remain stationary during the acquisition process, the macro-level facial tracking was not necessary for vascular tracking. This is because the experimental protocol required the person to sit passively and therefore no subjects experienced involuntary motions of such magnitude where macro-level tracking was required for compensation. Therefore, only micro-level tracking was utilized for vessel tracking. This helped lower computational overhead and decreased runtime.

Another observation during the course of acquiring the results was that little rotational and scaling motion was perceived during vascular tracking. Thus, it was possible to remove both the rotational and scaling tracking factors with minimal error to the tracking performance. Since the algorithm relies on a looping structure to check all combinations of translations, rotations, and scaling, this also allowed for a vast reduction in computational time.

The problem of runtime was revealed to be a major issue during the experimentation. Due to the programming of the algorithm in Matlab, the entire process , from tracking all the way through to final pulse rate calculation, could take up to 45 minutes to complete. Obviously, this result reveals that for medical applications, the computational process would need to be streamlined and implemented in a format which would allow for realtime calculation of pulse rate.

Another issue that was revealed in experimentation was the problem of thermal camera focus. In order to create a proper vascular map, the camera needs to have an excellent focus on the subject which will reveal the fine details that allow the construction of an accurate map. While the tracking software compensated for a variety

of motions there are very few ways to compensate for an out-of-focus acquisition scenario in terms of software. Thus, the subject was required to attempt to not drift closer or farther away from the camera, as this would greatly hamper the tracking process. Fortunately, very few subjects had enough motion to produce focus problems. One subject had to be imaged again because the focal problems became so great that the vascular map failed entirely and thus tracking was not possible and subjects 13, 14, and 29 were not used in the testing due to acquisition errors which made vascular mapping impossible.

5.1 SUMMARY

This chapter describes the results that came about due to the effort invested in this thesis. It details the results and their statistical significance as well as the many qualitative results that arose from this research. It explores possible reasonings for these results and lays the groundwork for the future work which is discussed in chapter 6.

CHAPTER 6:

CONCLUSIONS AND FUTURE WORK

In summary, the work of this thesis proved the concept of using thermal imaging to provide a new physiological indicator in the form of pulse rate for the use in smart room applications. An experimental system was designed which used thermal imaging capabilities combined with tracking and filtering software to produce an estimation of an average heart rate over a short time period. With an average accuracy of roughly 91%, the experiment shows that high accuracy measurements can be obtained even with the presence of small motion artifacts associated with uncontrollable muscle motion. This technology could be used to revolutionize the comfort and effectiveness of smart rooms worldwide.

Naturally, this work leads to several other areas of work that could be pursued for vast performance improvements. One major area would be the automatic selection of the vessel of interest. This work sacrifices true autonomy for decreased calculation time. With improved segmentation technology or a faster processing platform, automatic vessel selection is implementable. This would yield a truly automated process, since the current work requires the user to provide an initial manual selection.

Additionally, the motion tracking software was used to deal only with small-scale motion artifacts, not large scale movement. Future work should test the viability of this

technology under more challenging motion options and other visual scenarios such as partial occlusion. This might require the use of multiple imaging systems to compensate for various pose issues which were held constant under the current body of work.

Mechanical improvements could also be made to the system which would help with both the tracking and focal issues which were discussed in the Results section. To achieve consistent proper focus, an auto-focus system would need to be developed which could keep the thermal camera in constant proper focus. This could be achieved by using a range finder which could monitor subject distance and calibrate the camera appropriately. Additionally, the camera system could also employ a mechanical system to keep it directed on the patient which could allow for much broader tracking capabilities than even a software system which would require the subject to always be within the camera's field of view.

Naturally, for true future use in medical scenarios, the system will have to be optimized for speed. Under the current body of work, the Matlab programming language is used and while this is useful for ease of coding and development, it is not created for optimal speed. This is because Matlab is an interpreted programming language which needs to be reduced to machine code at run time. Future implementations should utilize compiled languages such as C#, which could dramatically decrease calculation times because the code is compiled before execution. Eventually, this process needs to be optimized to an extent that it could be run in real time, as the eventual application for it in medical scenarios would require near-instantaneous results. For instance, the Advisor Vital Signs Monitor utilized for ground truth measurement has a response time to heart

rate changes of six seconds. Therefore, any future application of this work would have a similar goal in terms of computation speed. This would allow the device to be implemented in smart room scenarios where it could be used to replace the current monitoring technology.

Finally, future work should also have the end goal of higher accuracy. An accuracy of 91% translates into an error of roughly 6 – 10 BPM for the range of heart rates tested under this work.   However, current ECG and pulse monitoring devices typically have an accuracy that allows them to calculate the heart rate within +/- 2 BPM. This yields an average accuracy of 97 – 99% for the heart beats in the range of 60 – 80 beats per minute. This should obviously be the target accuracy for future systems, although this would represent an enormous leap in technology as the process also has to compensate for a variety of factors not present in contact measurement techniques such as pulse oximetry. Still, this can be achieved in several forms. Naturally, longer acquisition periods can allow for less error during the Fourier analysis. However, this would not typically be feasible for medical work, where high accuracy measurements are required over short time durations. Another method for increasing the accuracy is to increase the sampling rate via better image acquisition systems with faster frame rates.

Considering the work done in this thesis and future improvements, this technology could be a great benefit to health systems around the world. The thesis served as a conclusive demonstration that non contact pulse measurement is indeed a feasible reality. It sets the basis for a new arena of research with the hopes of creating a fully

functional smart room that can utilize the benefits of technology to improve the health

and safety of its users.

REFERENCES

[1] "Facts About Healthcare." Health Care Insurance Costs. Web. 18 Aug. 2009.
        <http:/www.nchc.org/facts/cost.shtml>.

[2] "Vital Signs." University of Virginia Health System. Web. 20 Aug. 2009.   <http://
        www.healthsystem.virginia.edu/uvahealth/adult_cardiac/vital.cfm>.

[3] Ehline, Michael. "What Are the Burn Injury Statistics in the United States?"
        EzineArticles. Web. 22 May 2009. <http://ezinearticles.com/?What-Are-the-
        Burn-Injury-Statistics-in-the-United-States?&id=1266763>.

[4] Cotterill, Simon. "The Cardiovascular System." CancerIndex. University of
        Newcastle upon Tyne, 31 Aug. 2000. Web. 11 Aug. 2009.
        <http://www.cancerindex.org/medterm/medtm8.htm>.

[5] "Circulatory System." Wikipedia. Web. 11 Aug. 2009. <http://en.wikipedia.org
        /wiki/Circulatory_system>.

[6] "Amazing Heart Facts." Cut to the Heart. PBS, 1997. Web. 11 Aug. 2009.   <http://
        www.pbs.org/wgbh/nova/heart/heartfacts.html>.

[7] Milnor, William R. "Cardiovascular physiology". New York: Oxford UP, 1990.

[8] Wu, Shiqian, Weisi Lin, and Shoulie Xie. "Skin Heat Transfer Model of Facial
        Thermograms and its Application in Face Recognition." Pattern Recognition
        41 (2008): 2718-729.

[9] Atamaz Y. and Govsa F. Anatomy of the superficial temporal artery and its
        branches: its importance for surgery. Surgical and Radiologic Anatomy,
        28(3):248–253, 2006.

[10] Chekmenev, Sergey. "Non-Contact and Passive Computer Vision Methodology for
        Measurement of Arterial Pulse using Thermal IR Imaging." Diss. University of
        Louisville, 2008.

[11] "Superficial Temporal Artery." Wikipedia. Web. 11 Aug. 2009.
        <http://en.wikipedia.org/wiki/Superficial_temporal_artery>.

[12] "Infrared." Wikipedia. Web. 10 Jan. 2009. <http://en.wikipedia.org/wiki/Infrared>.

[13] "Planck Law." ScienceWorld. Wolfram Research. Web. 25 Aug. 2009. <http://scienceworld.wolfram.com/physics/PlanckLaw.html>.

[14] Orlove, Gary L., and Robert P. Madding. "Answers to the Common Questions People Ask about Infrared Thermography." IR Thermography Primer. 23 June 2003. Web. 21 Mar. 2009. <http://www.itcnewsletter.com/Newitc/ir_primer.htm>.

[15] Polikar, Robi. "Multiresolution Analysis and the Continuous Wavelet Transform." THE WAVELET TUTORIAL PART III. Rowan University. Web. 11 Feb. 2009. <http://users.rowan.edu/~polikar/WAVELETS/WTpart3.html>.

[16] Valens, C. "A Really Friendly Guide To Wavelets." Web. 23 Aug. 2009. <http://pagesperso-orange.fr/polyvalens/clemens/wavelets/wavelets.html>.

[17] Graps, Amara. "An Introduction to Wavelets." IEEE Computational Science and Engineering 2.2 (1995).

[18] Polikar, Robi. "Multiresolution Analysis and the Continuous Wavelet Transform." THE WAVELET TUTORIAL PART III. Rowan University. Web. 11 Feb. 2009. <http://users.rowan.edu/~polikar/WAVELETS/WTpart3.html>.

[19] Dowdall, Jonathan, Ioannis T. Pavlidis, and Pangiotis Tsiamyrtzis. "Coalition Tracking." Computer Vision and Image Understanding 106 (2007): 205-19.

[20] Doucet, Arnaud, and Adam M. Johansen. A Tutorial on Particle Filtering and Smoothing: Fifteen years later. Tech.

[21] Buddharaju, Pradeep, Ioannis Pavlidis, Panagoitis Tsiamyrtzis, and Mike Bazakos. "Physiology-Based Face Recognition in the Thermal Infrared Spectrum." IEEE Transactions on Pattern Recognition and Machine Intelligence 29.4 (2007): 613-26. IEEE Xplore. Web. 11 Aug. 2009. <http://ieeexplore.ieee.org>.

[22] Steketee, J. "Spectral Emissivity of Skin and Pericardium." Physical Medical Biology 18.5 (1973): 686-94.

[23] Chekmenev, Sergey, Aly Farag, and Edward Essock. "Multiresolution Approach for Non-Contact Measurements of Arterial Pulse using Thermal Imaging." Proceedings of the 2006 Conference on Computer Vision and Pattern Recognition Workshop (2006). IEEE Xplore. Web. 05 Feb. 2009. <http://ieeexplore.ieee.org>.

[24] Chekmenev, Sergey, Aly Farag, and Edward Essock. "Thermal Imaging of the Superficial Temporal Artery: An Arterial Pulse Recovery Model." IEEE Conference on Computer Vision and Patern Recognition (2007). IEEE Xplore. Web. 12 Jan. 2009. <http://ieeexplore.ieee.org>.
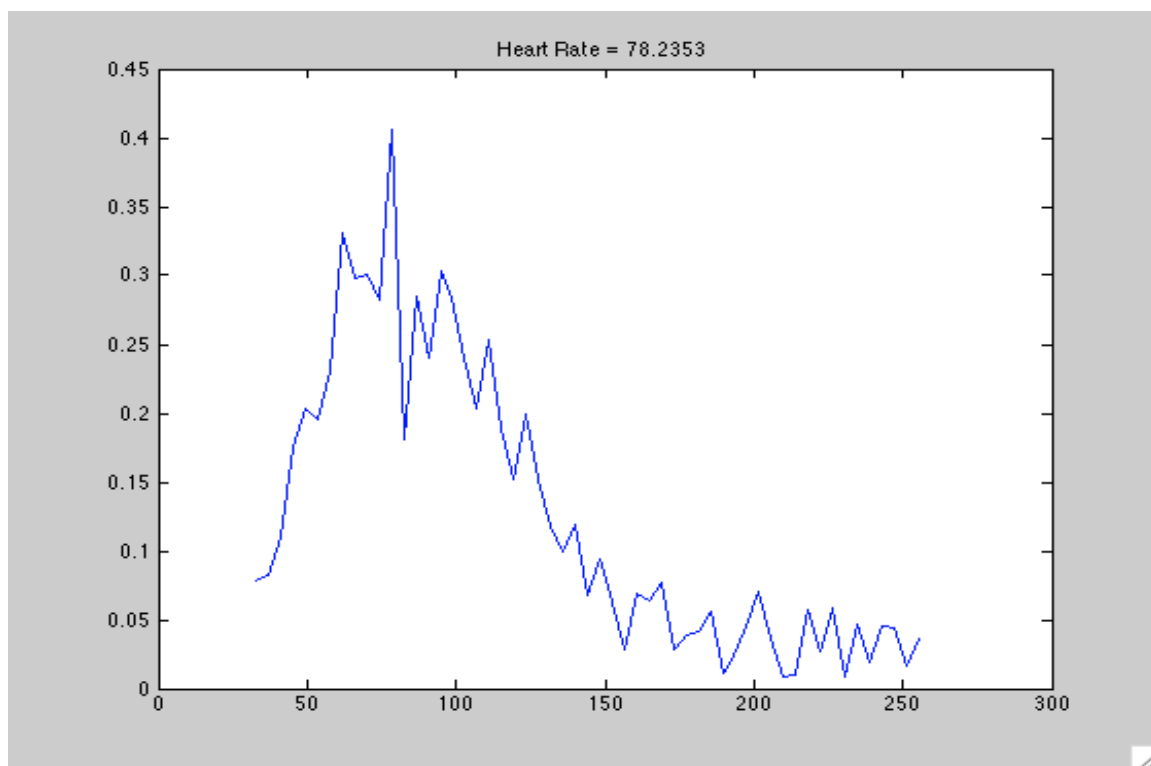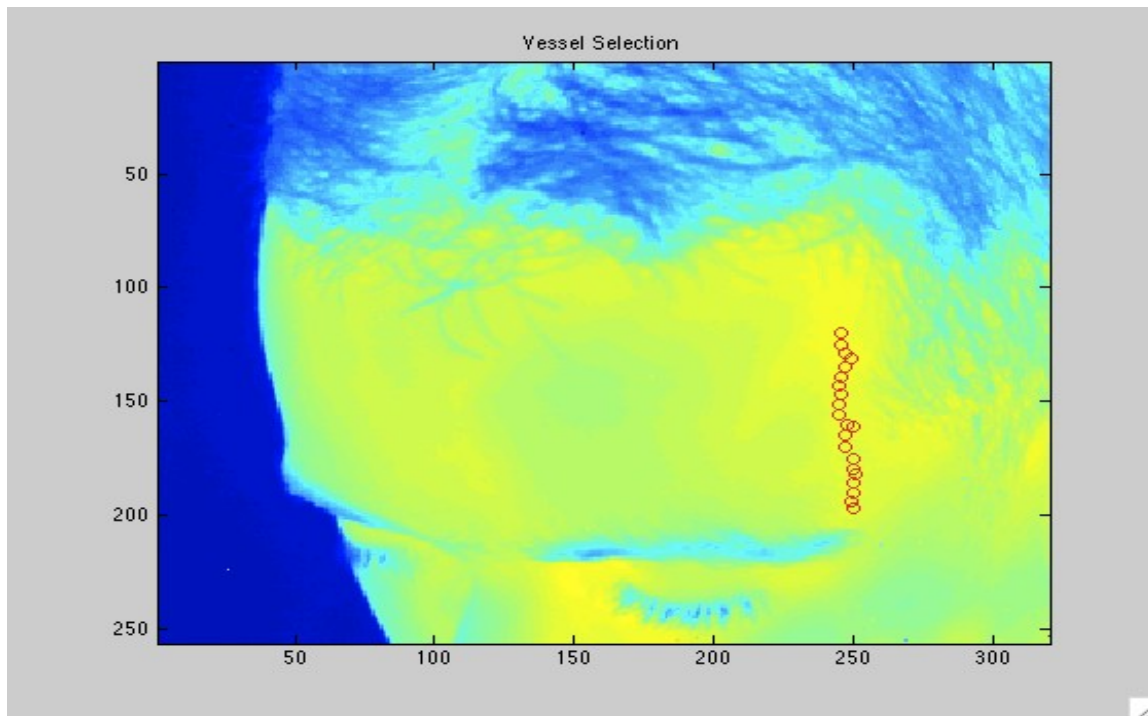
APPENDICES

Trial Results from Group 1

Subject 1, Trial 1, Ground Truth = 70 BPM, Calculated = 61.8 BPM
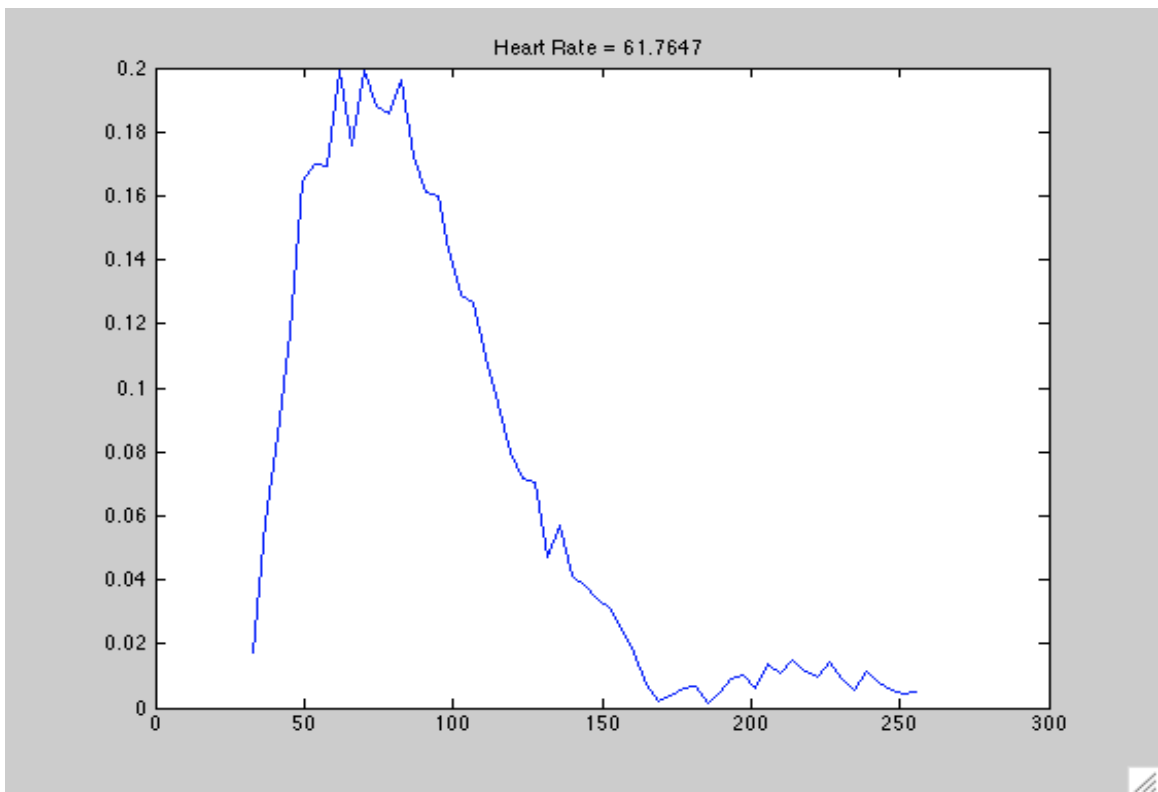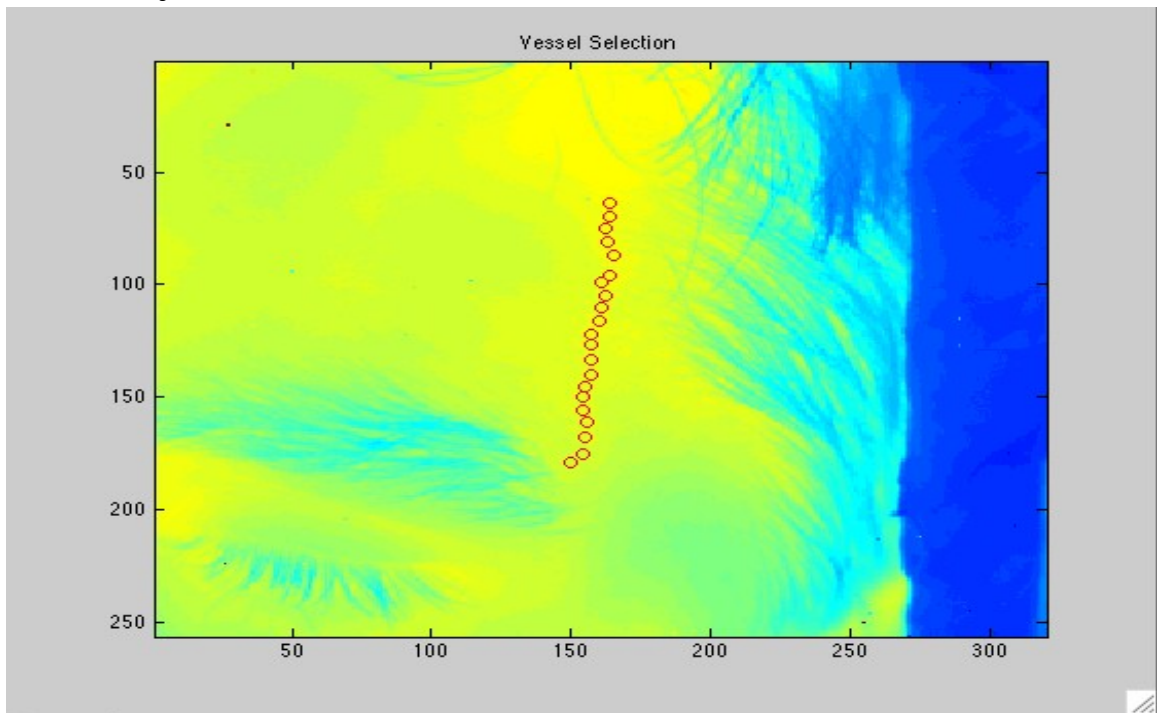
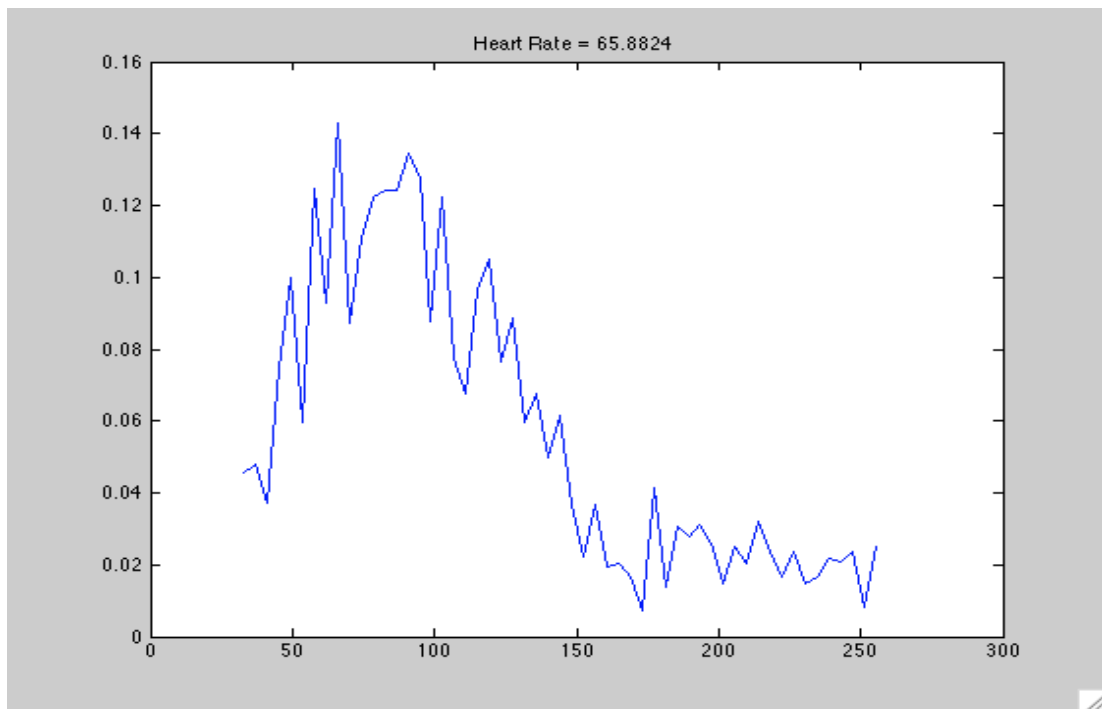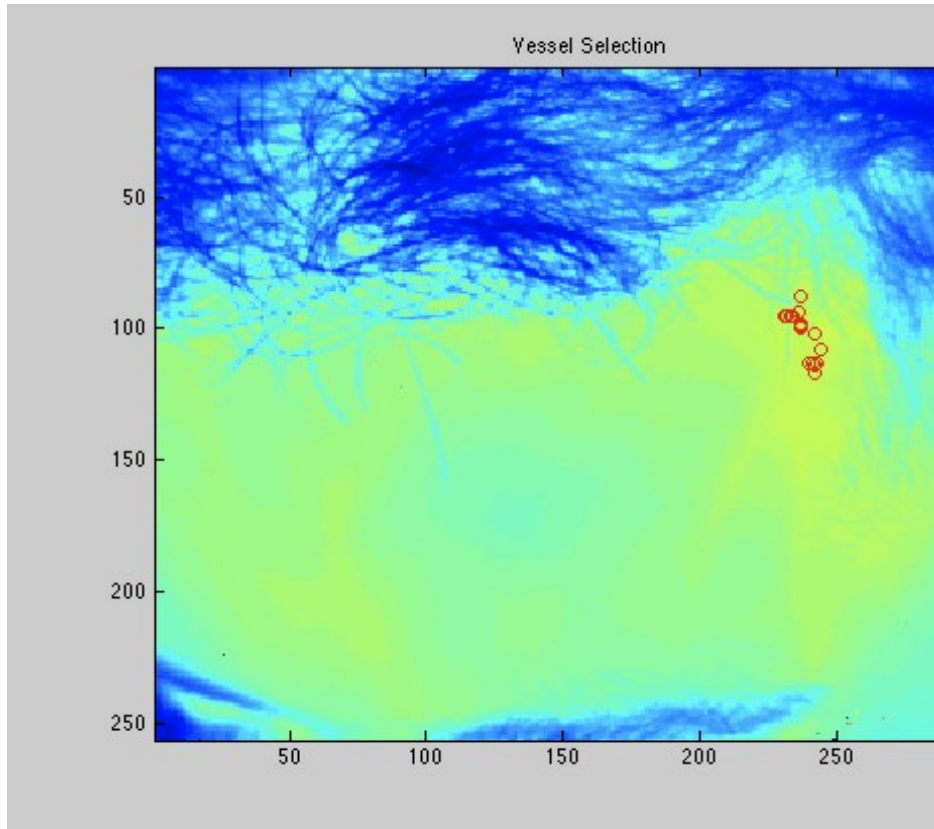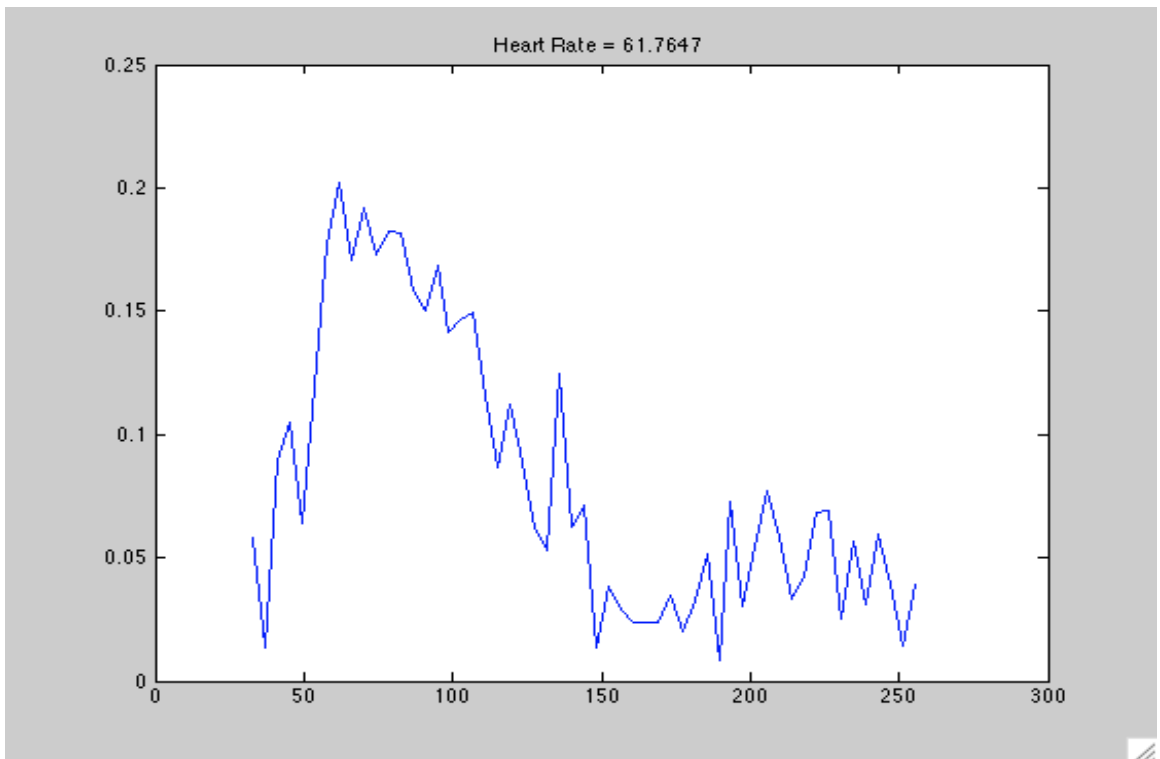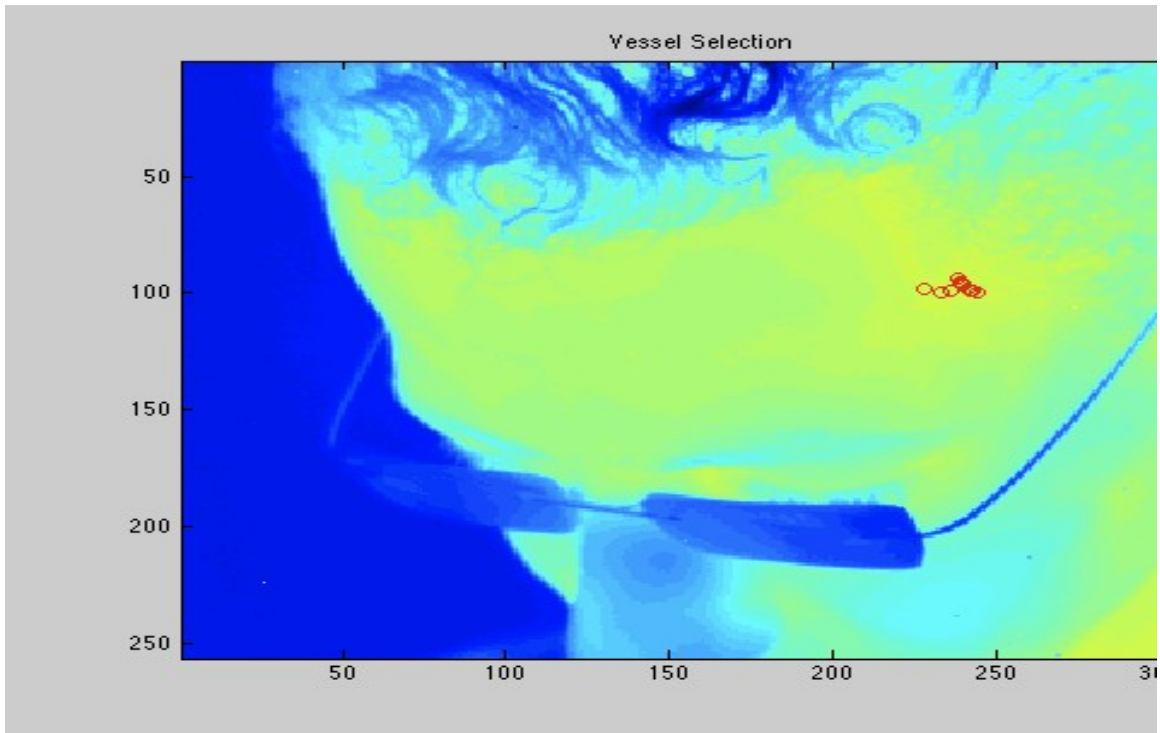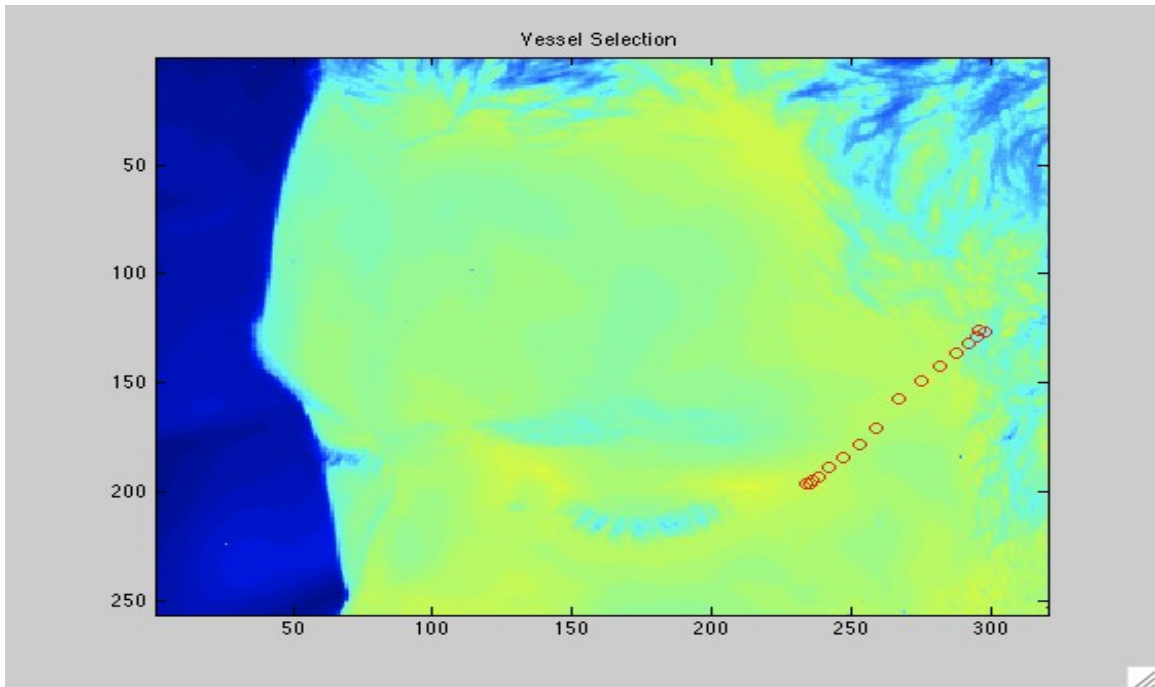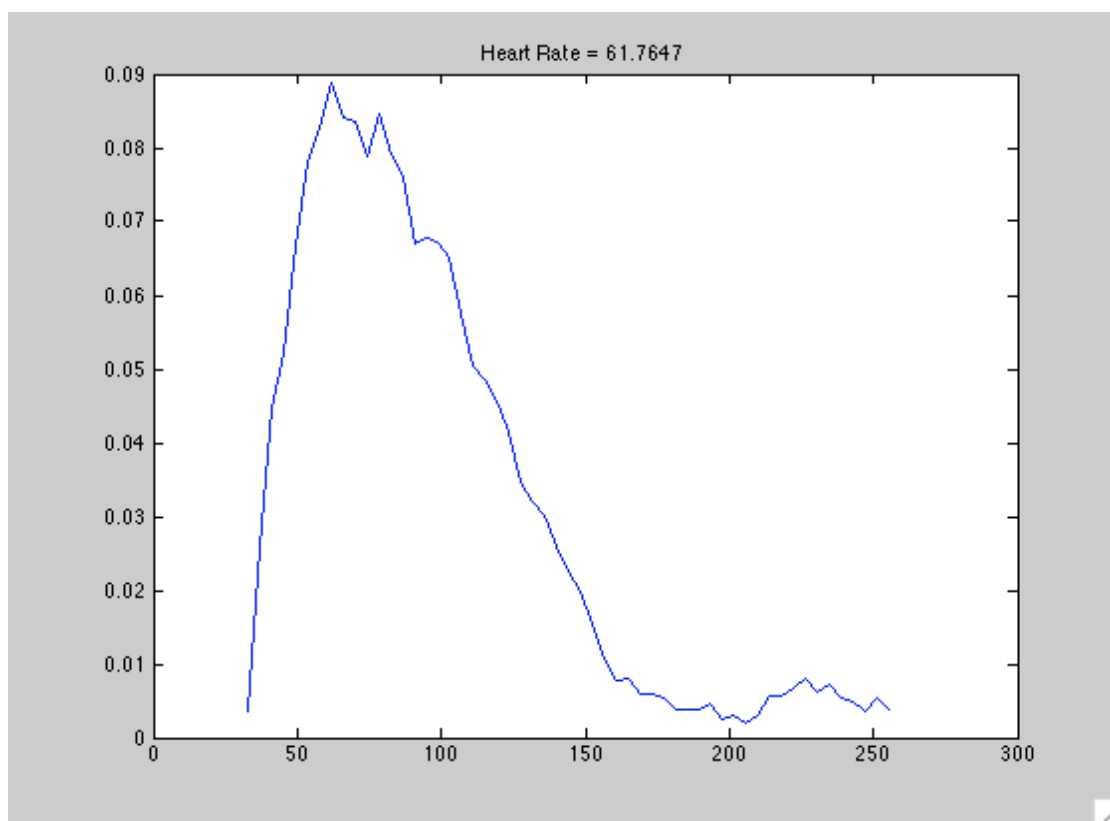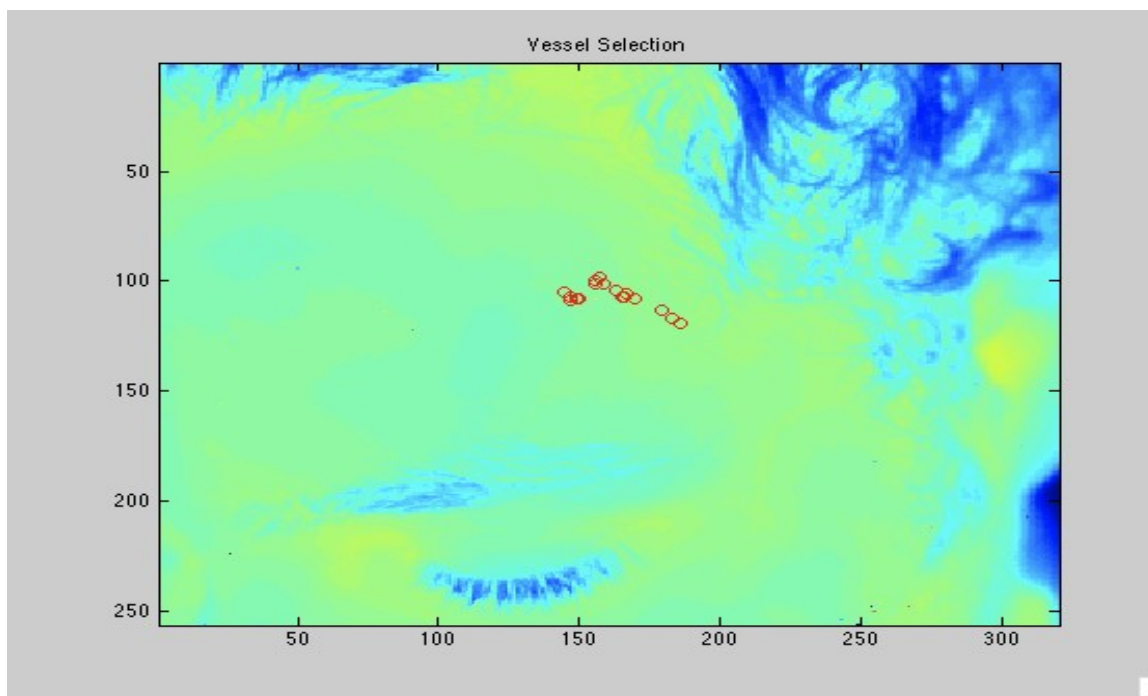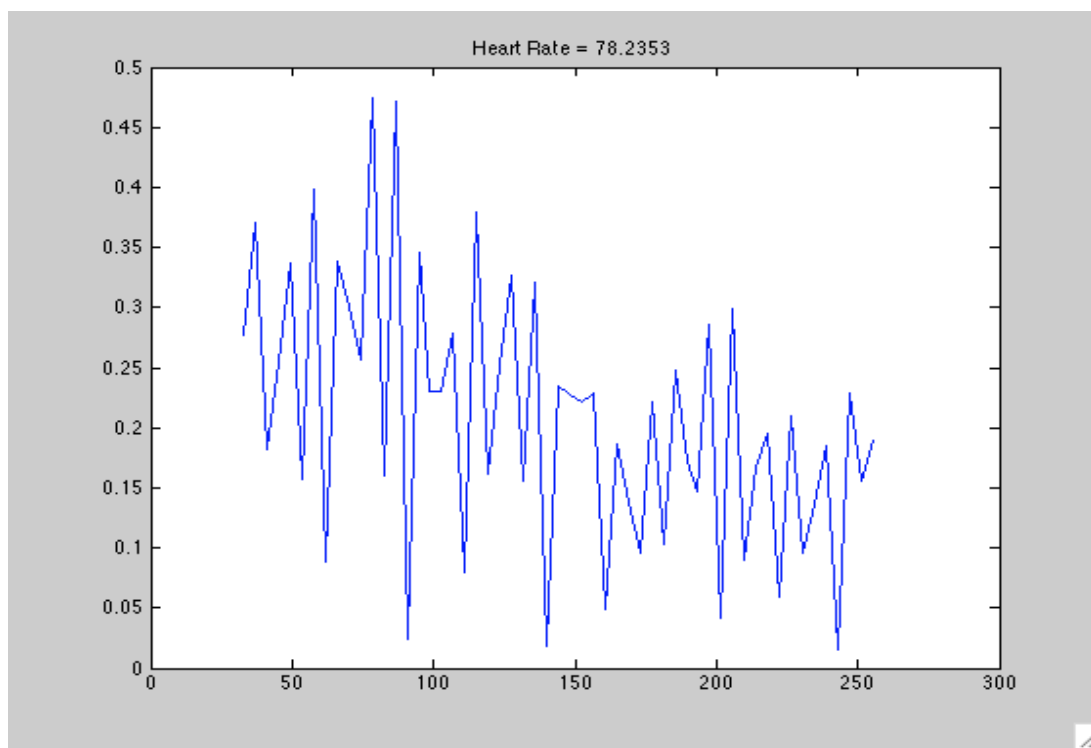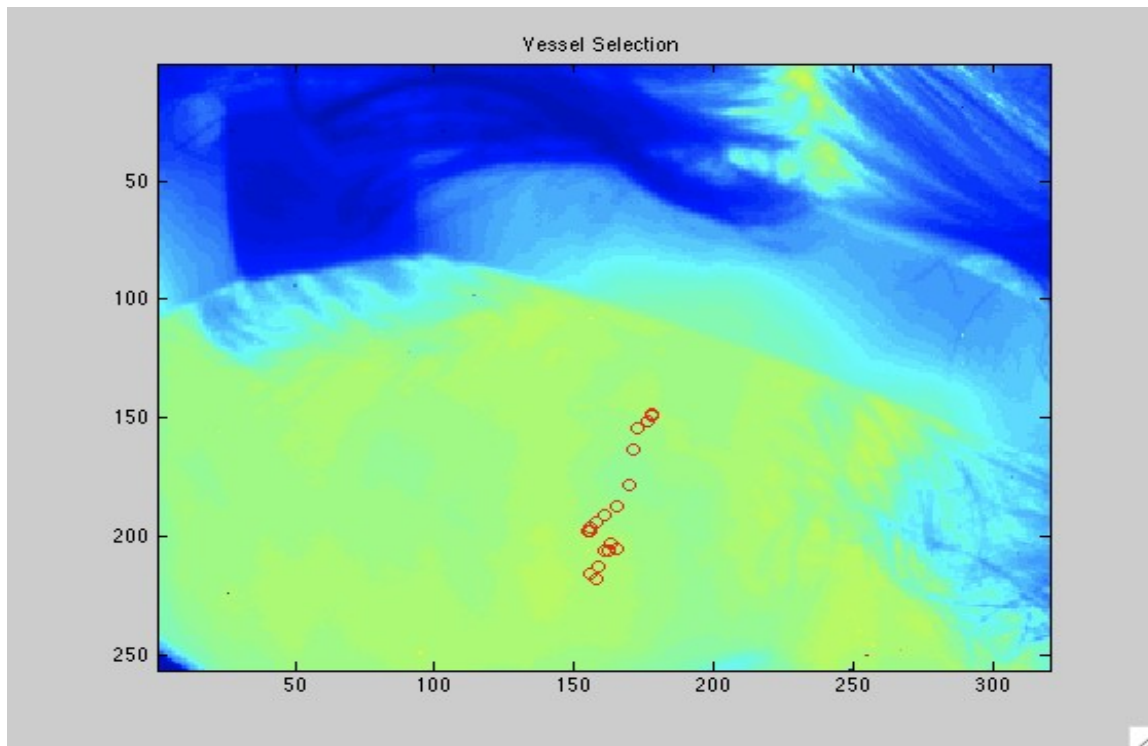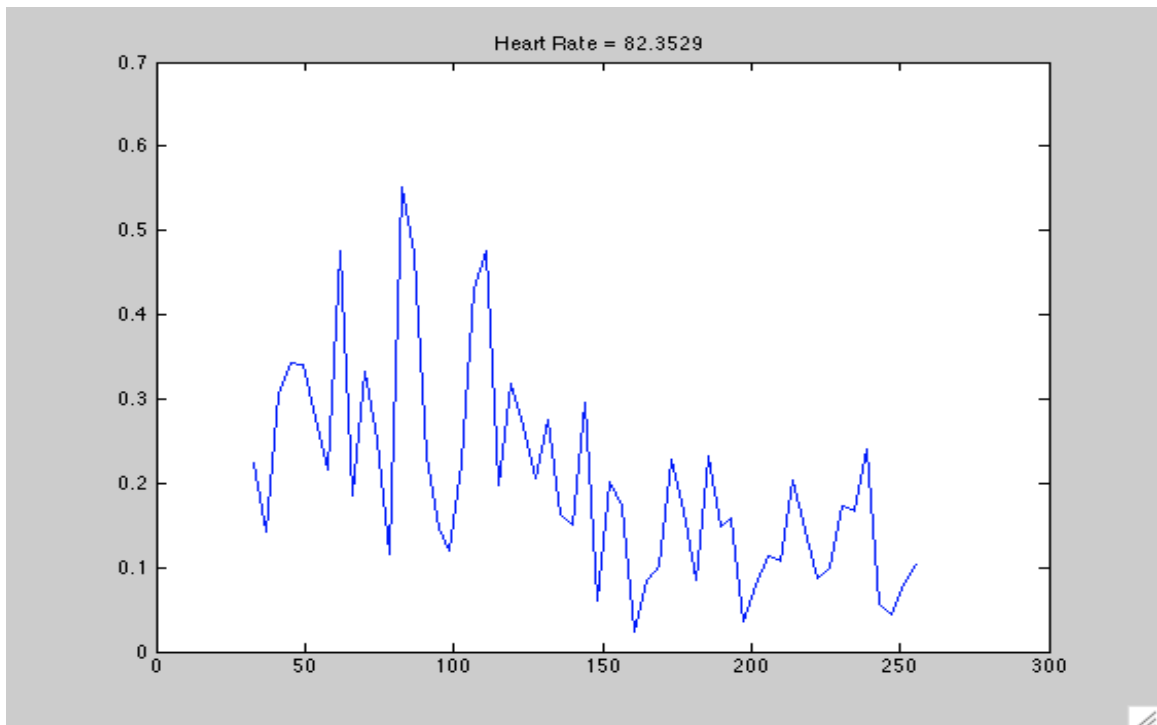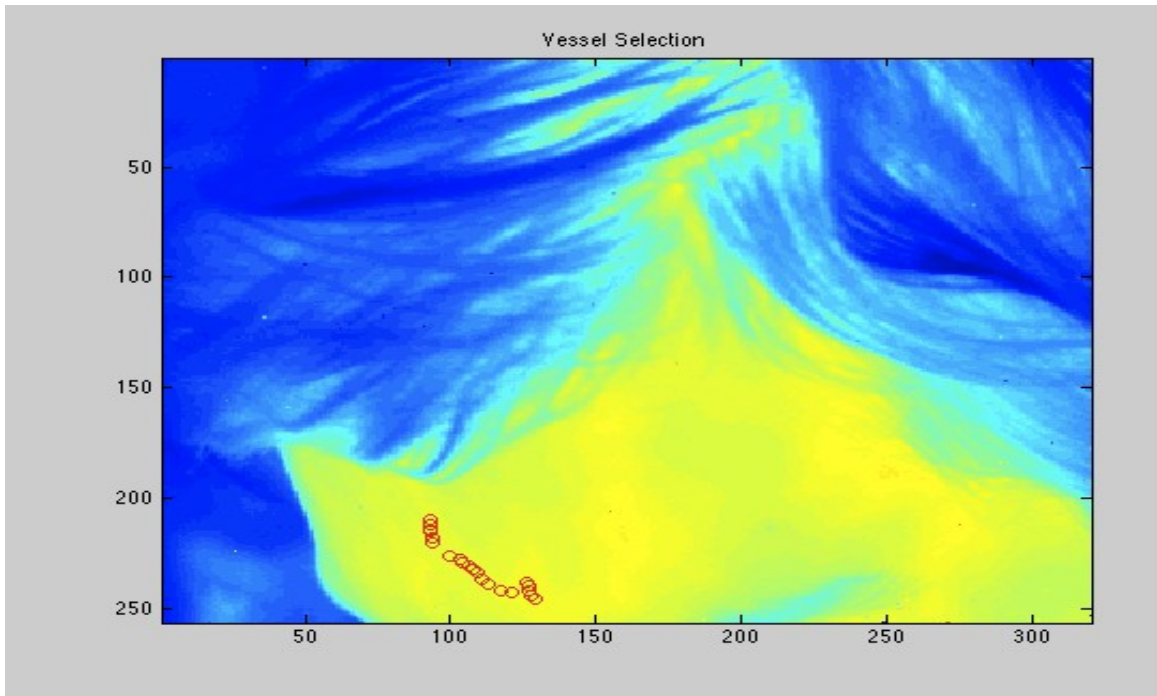**Subject 1, Trial 2, Ground Truth 70 BPM, Calculated 65.9 BPM**
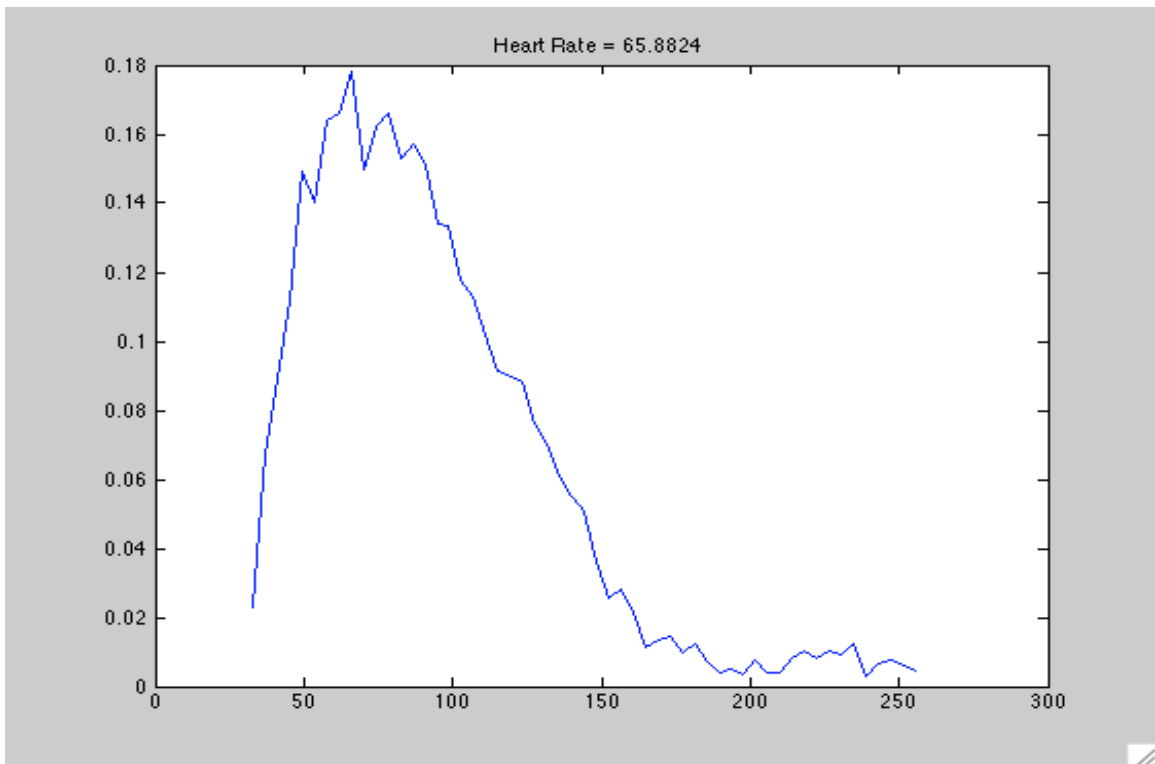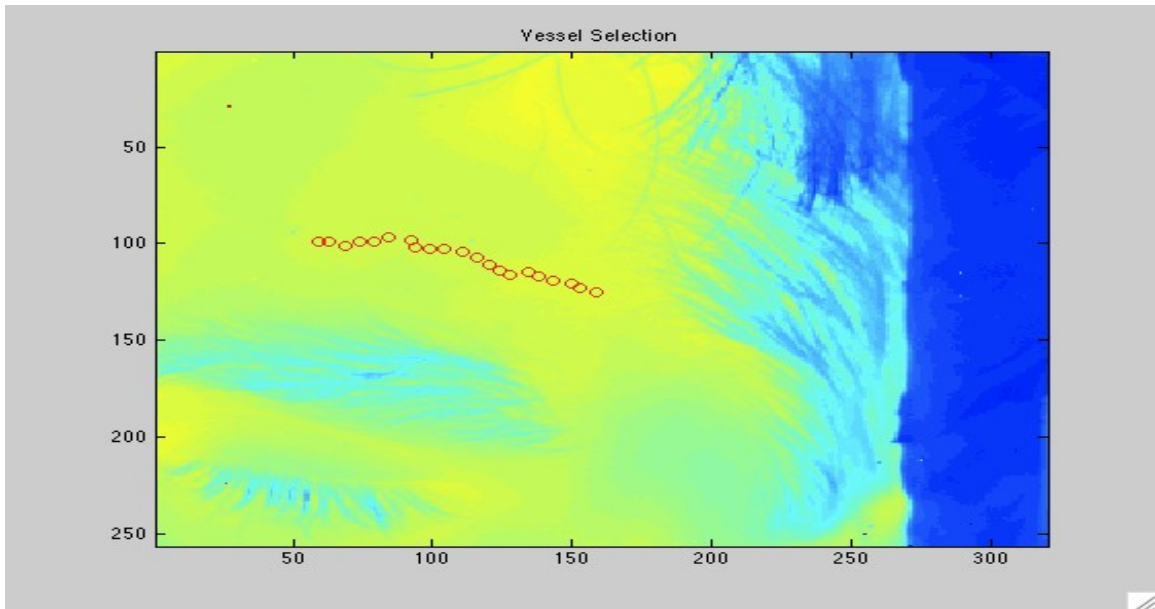
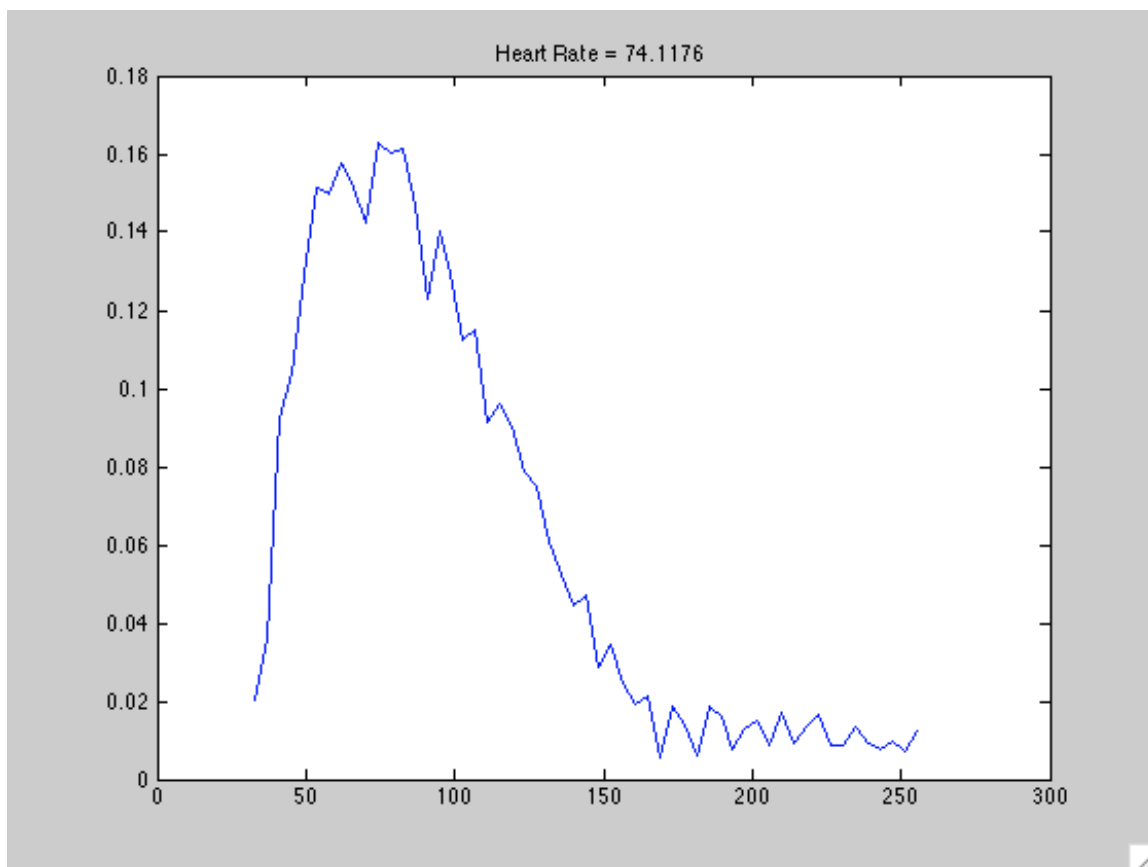**Subject 1, Trial 3, Ground Truth = 80 BPM, Calculated = 70.0 BPM**
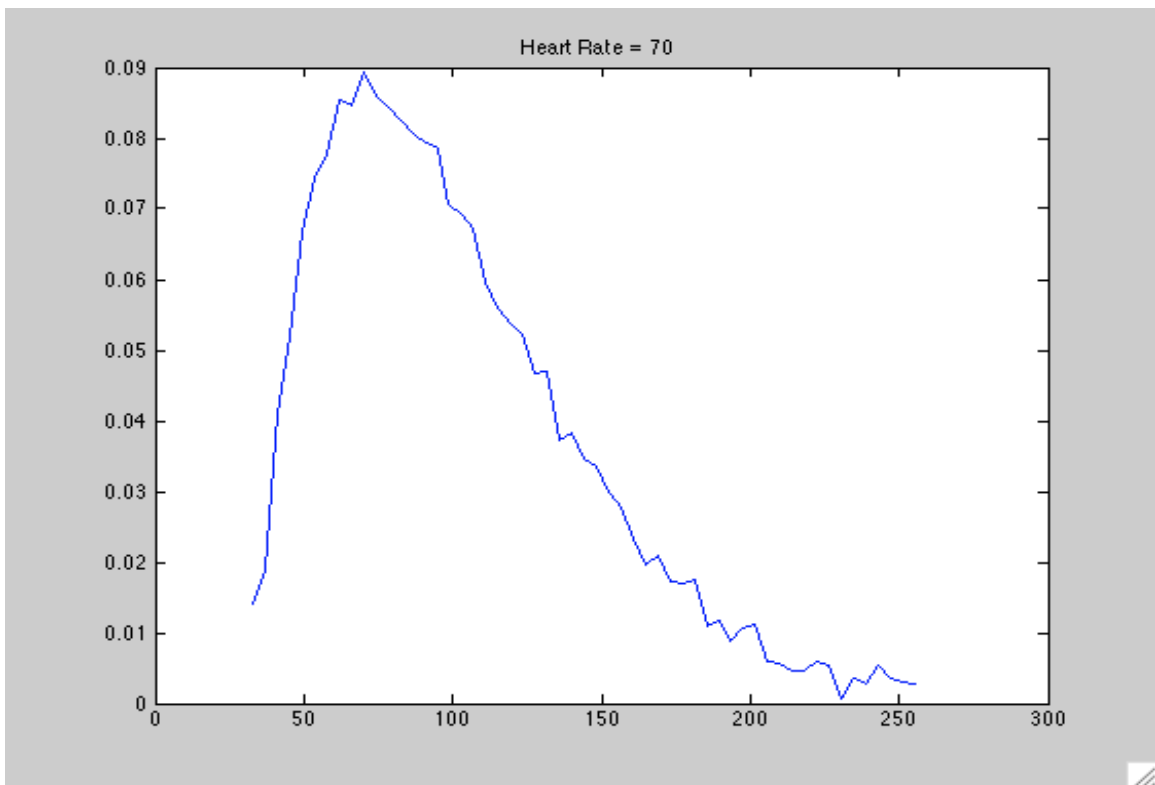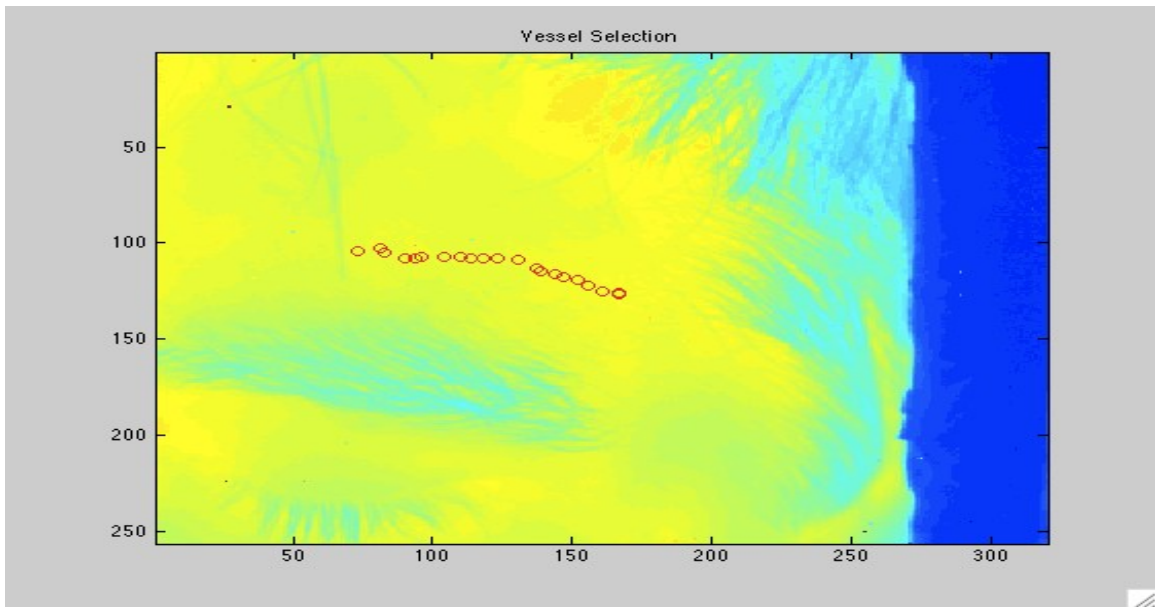
Subject 1, Trial 4, Ground Truth = 76 BPM, Calculated = 78.2 BPM
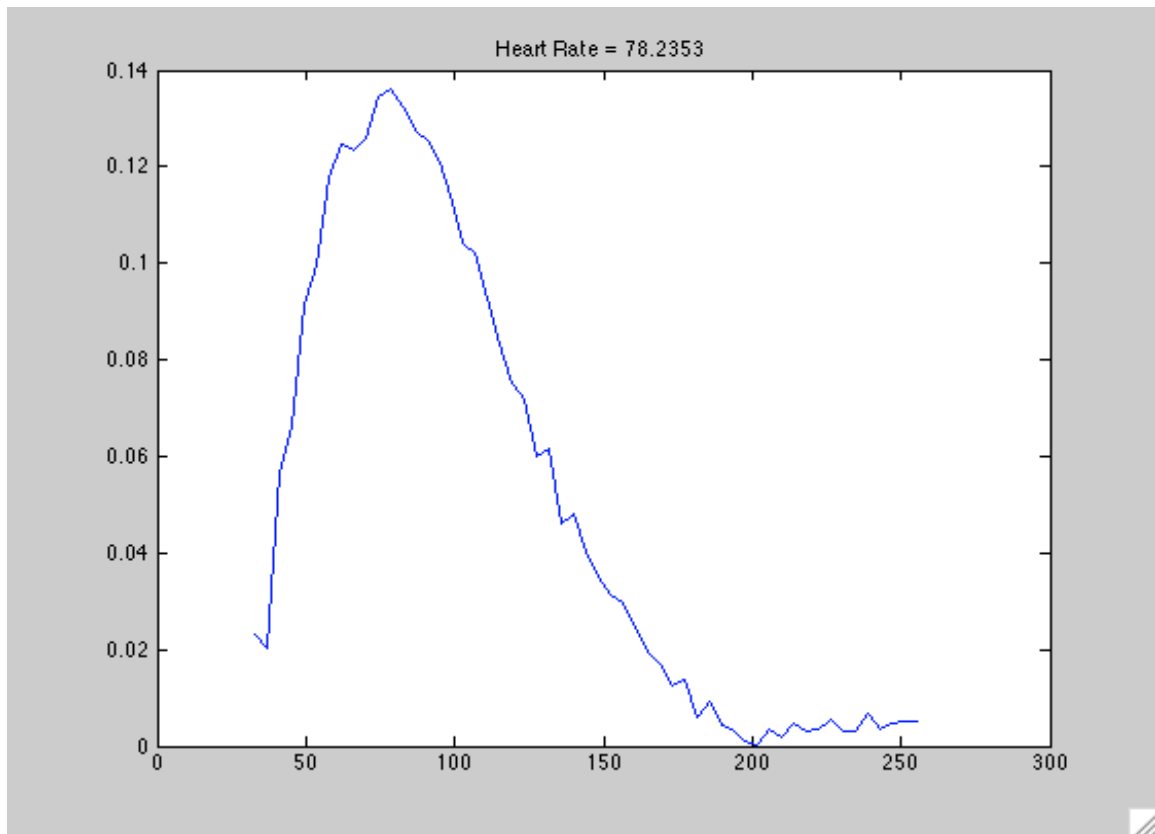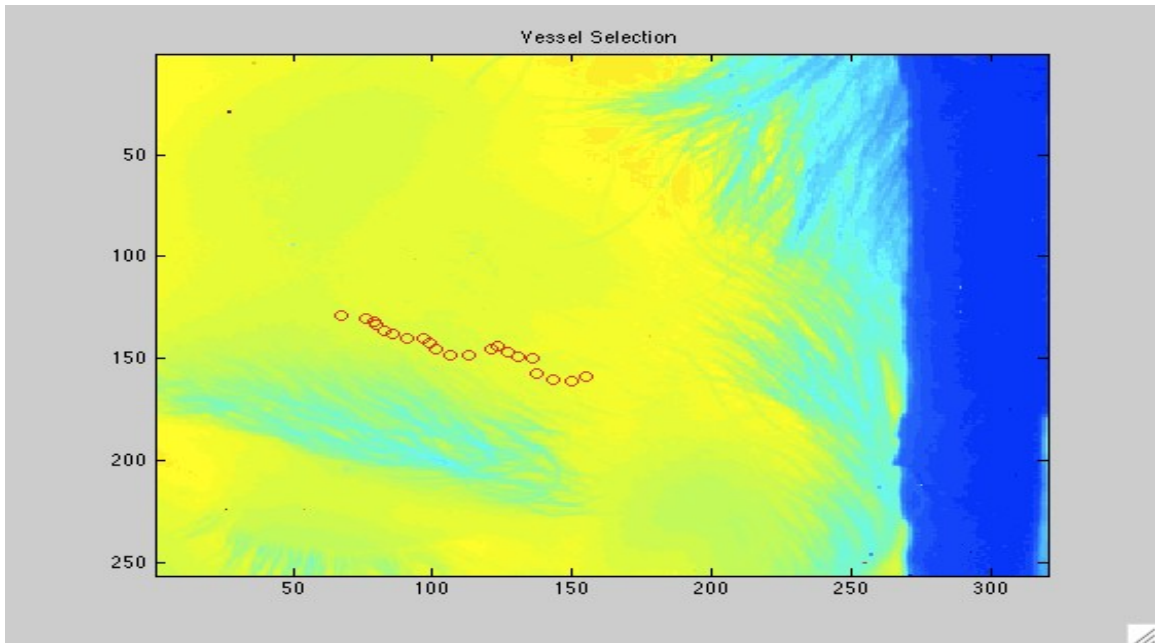
Subject 1, Trial 5, Ground Truth = 79, Calculated = 67.1
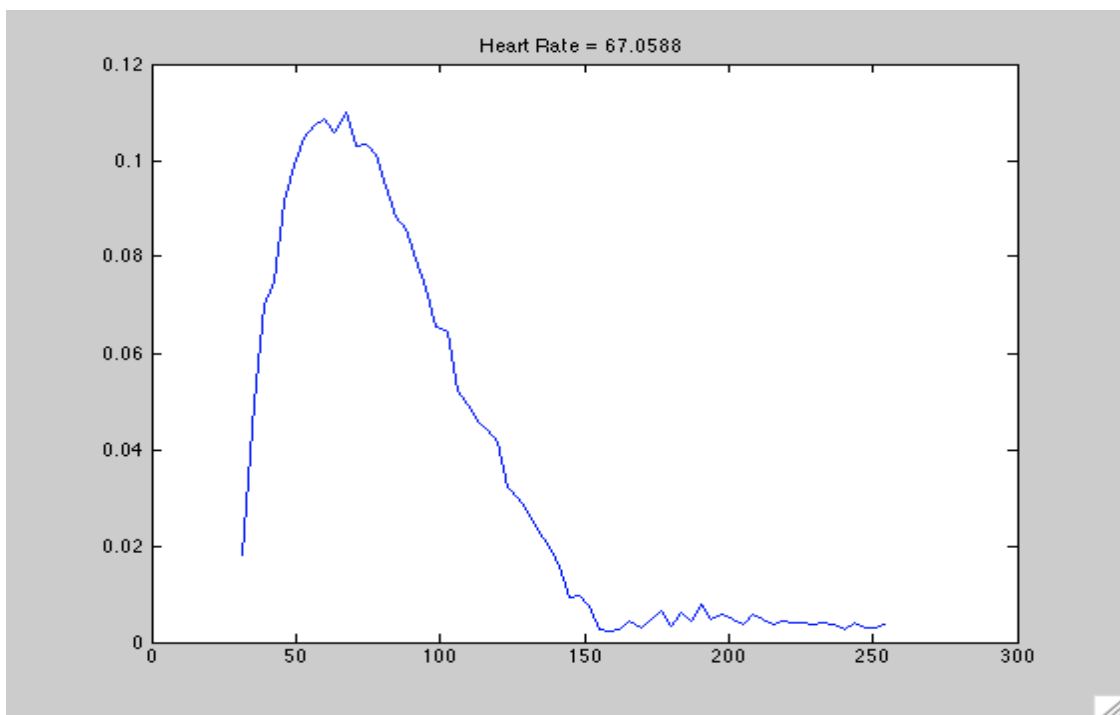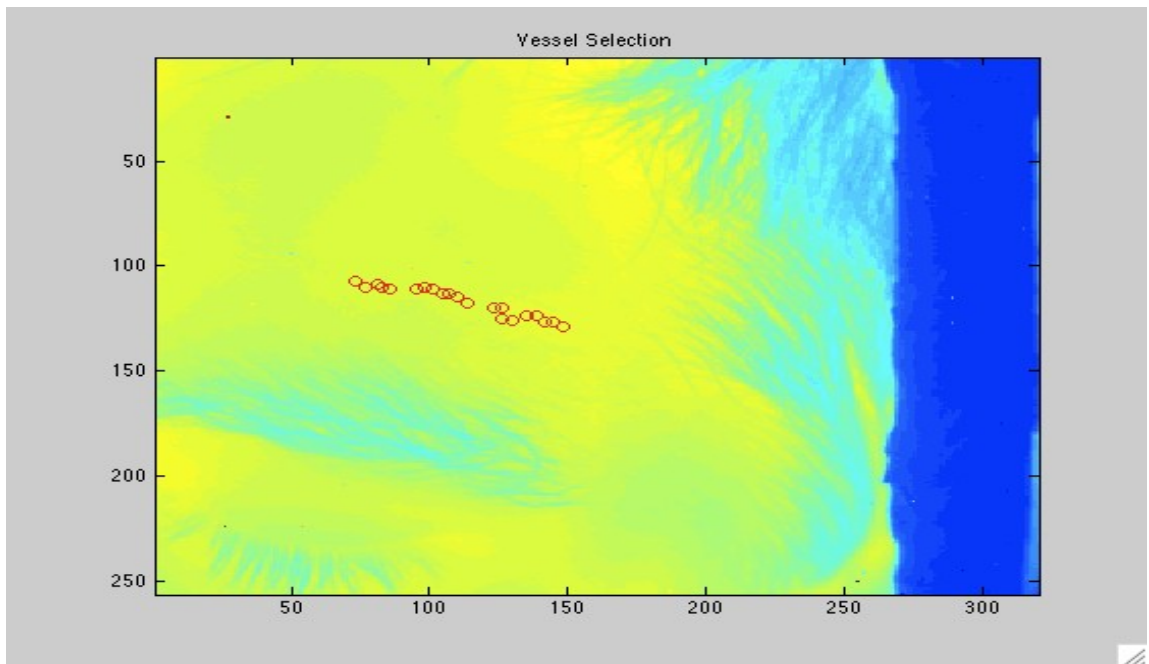
Subject 2, Trial 1, Ground Truth = 75 BPM, Calculated = 67.1 BPM
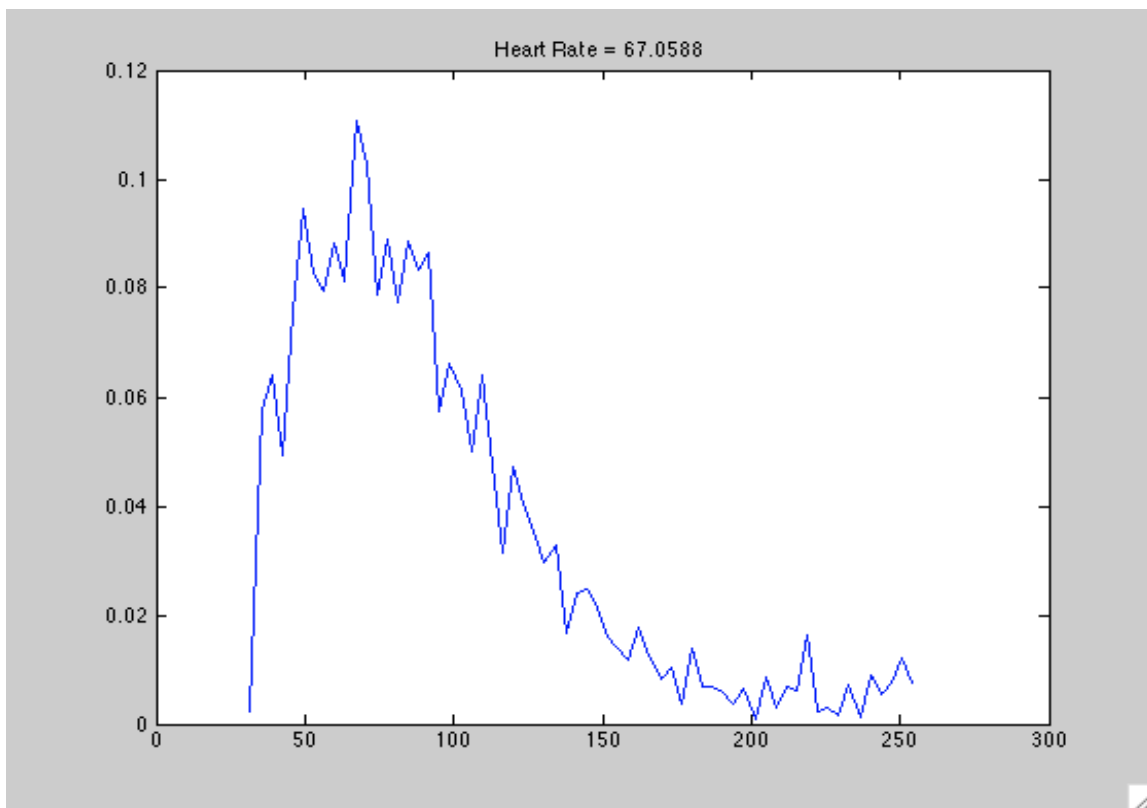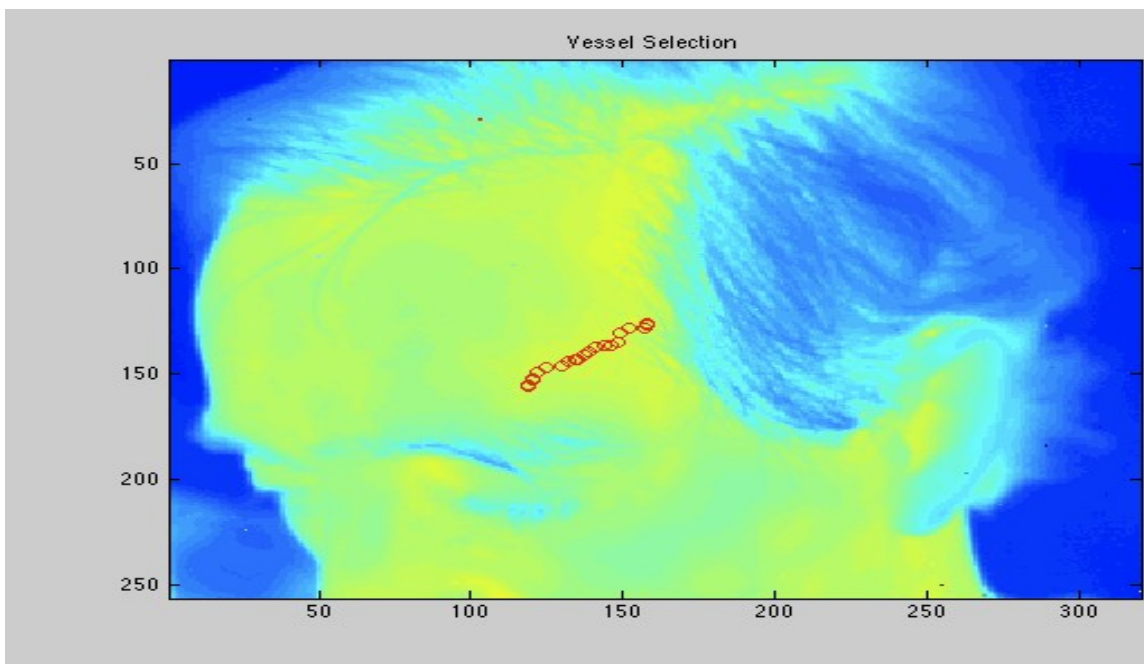
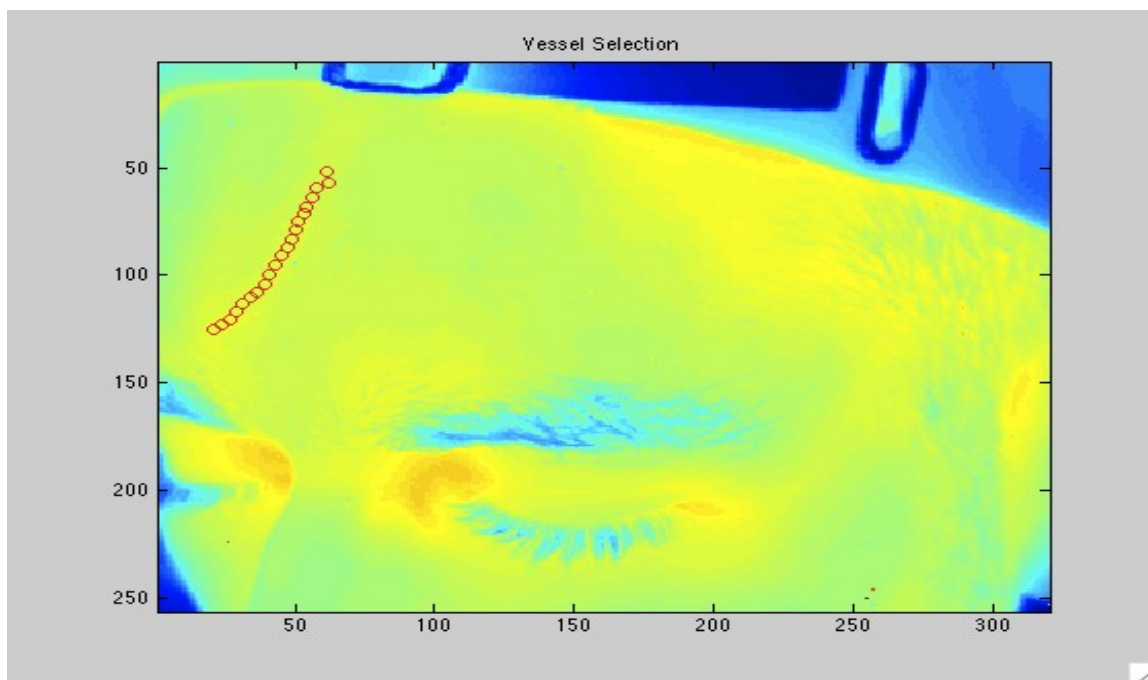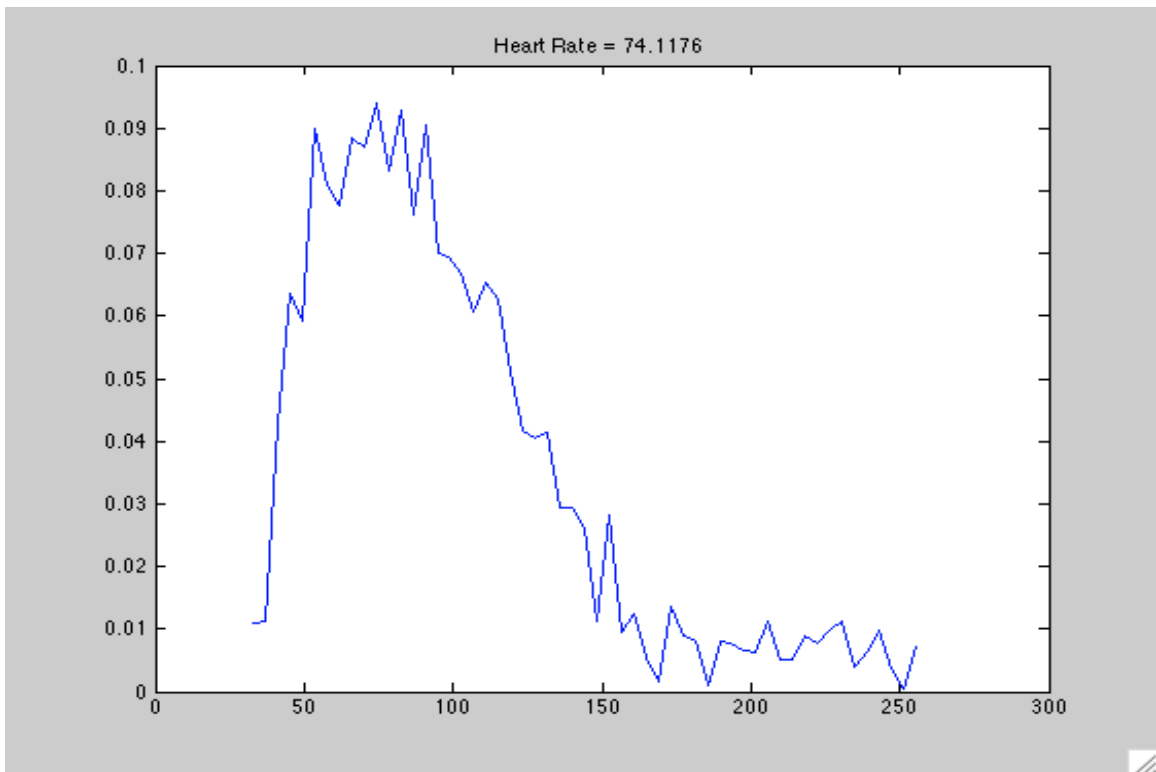**Subject 3, Trial 1, Ground Truth = 85 BPM, Calculated = 86.4 BPM**
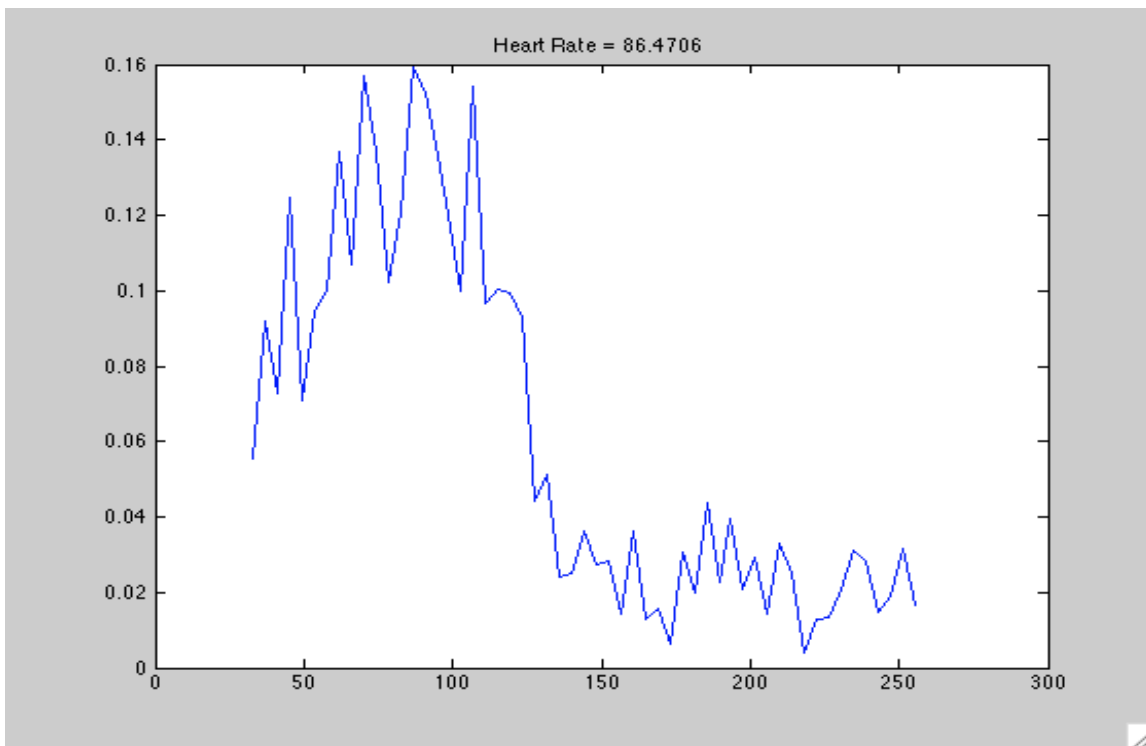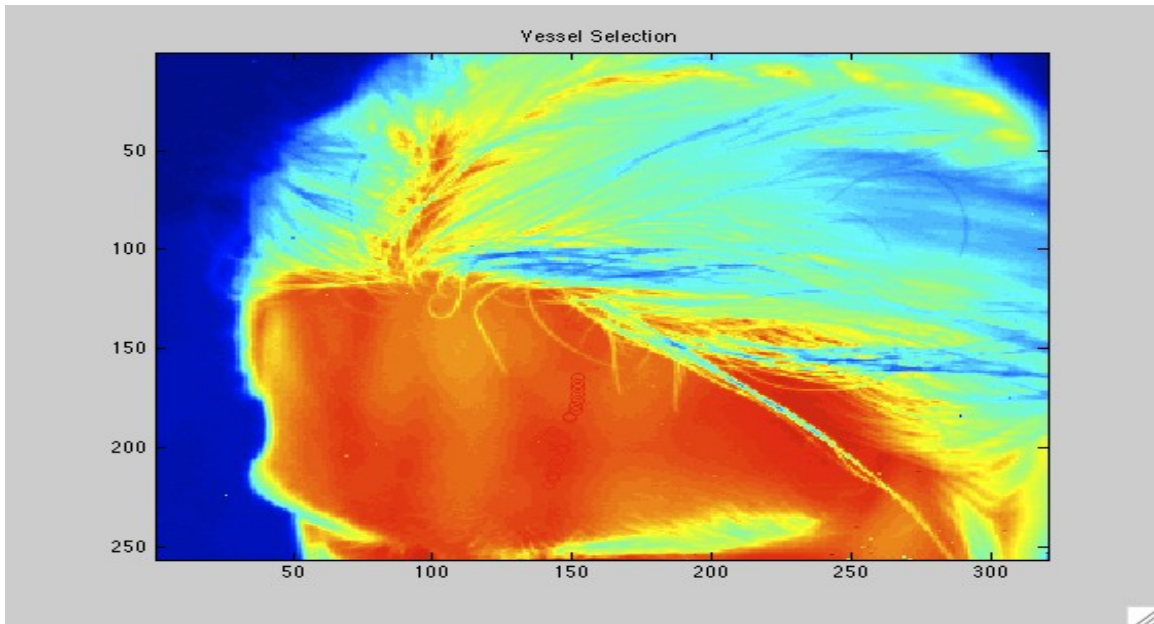
Subject 3, Trial 2, Ground Truth = 58 BPM, Calculated = 74.1 BPM
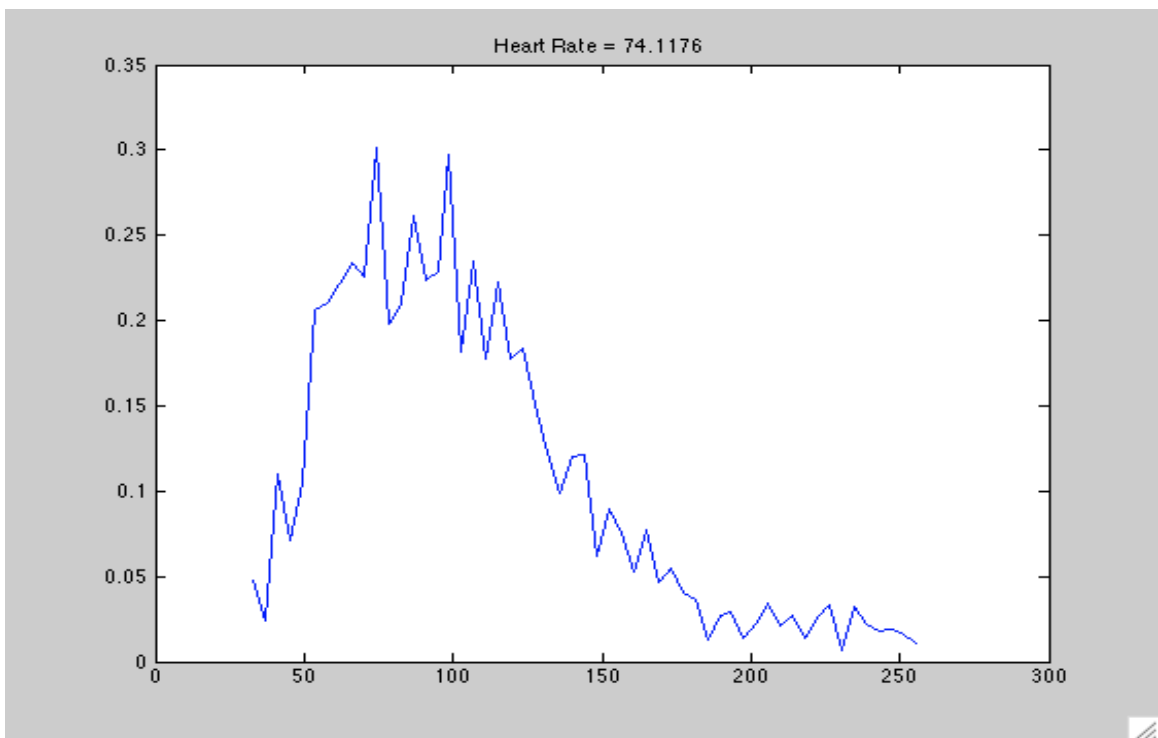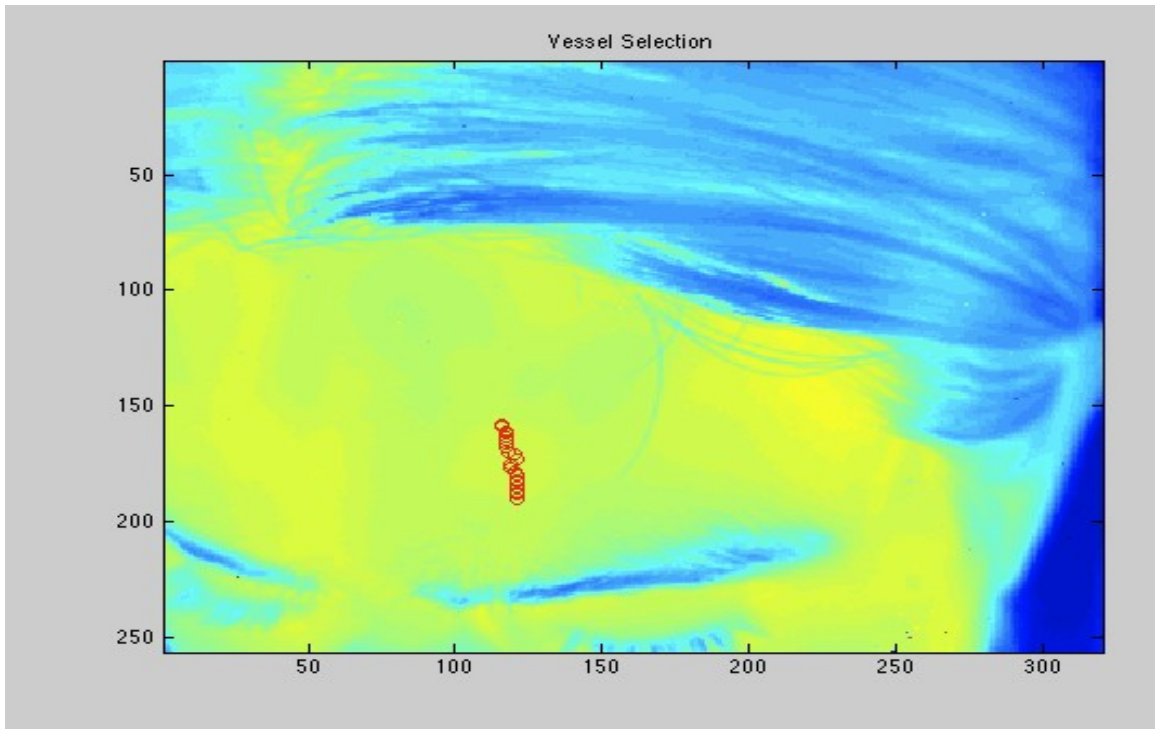
Subject 4, Trial 1, Ground Truth = 81 BPM, Calculated = 74.1 BPM
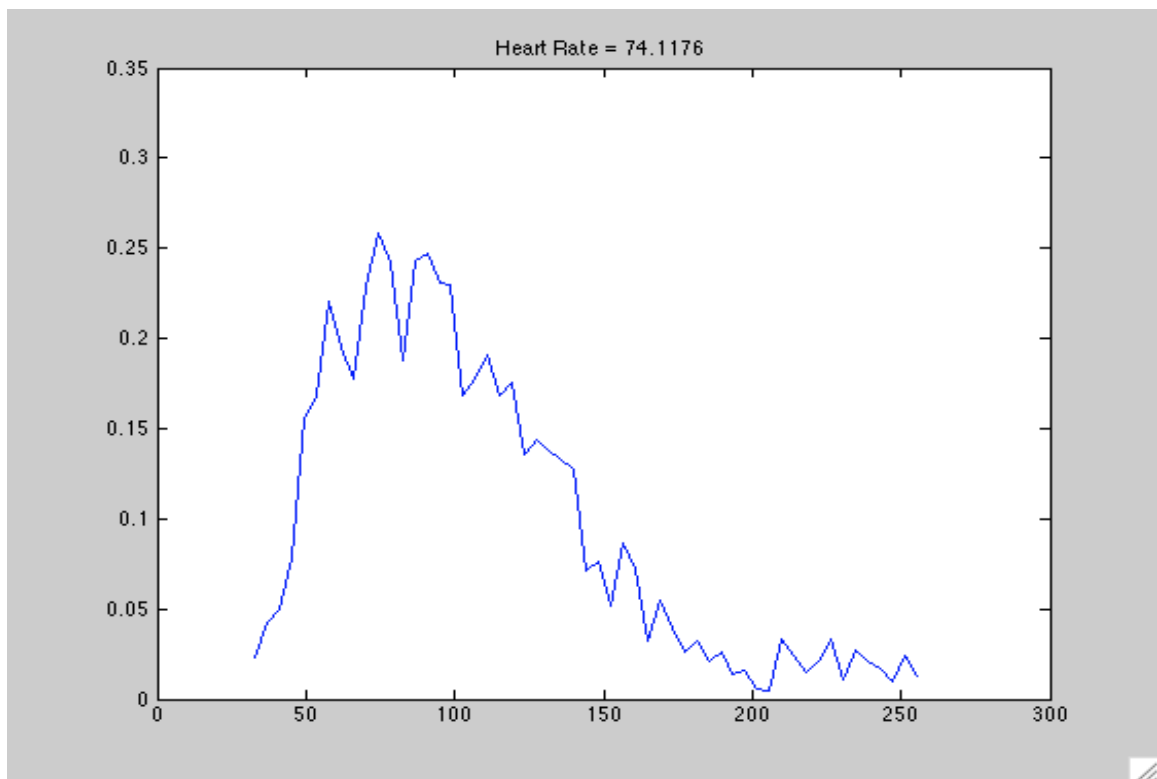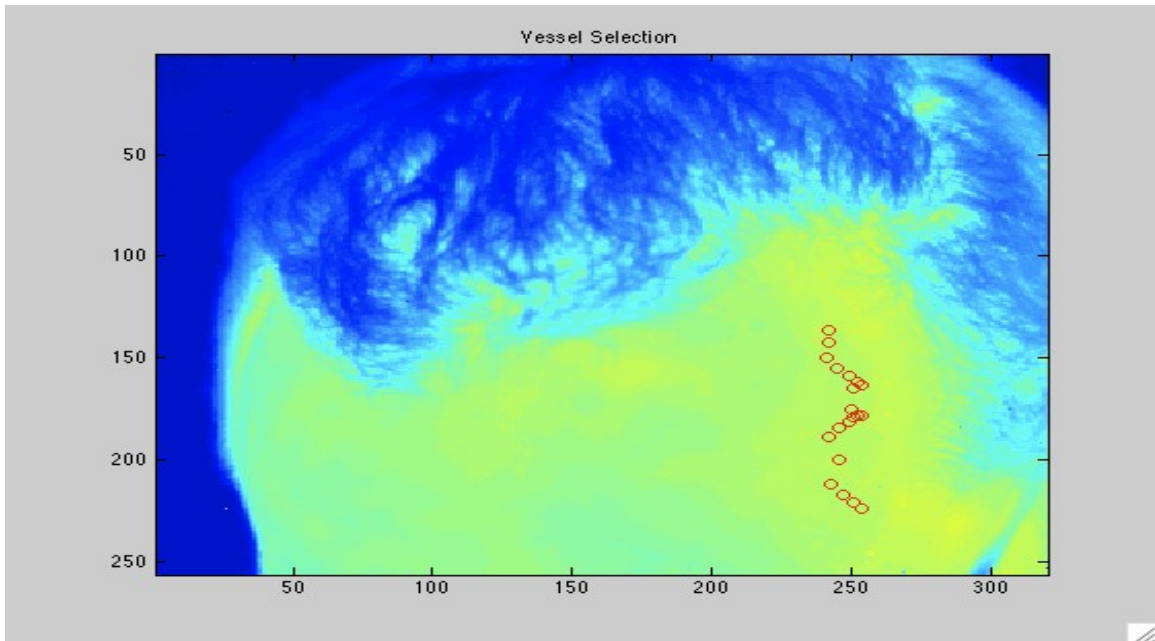
Subject 5, Trial 1, Ground Truth = 75 BPM, Calculated = 74.1 BPM
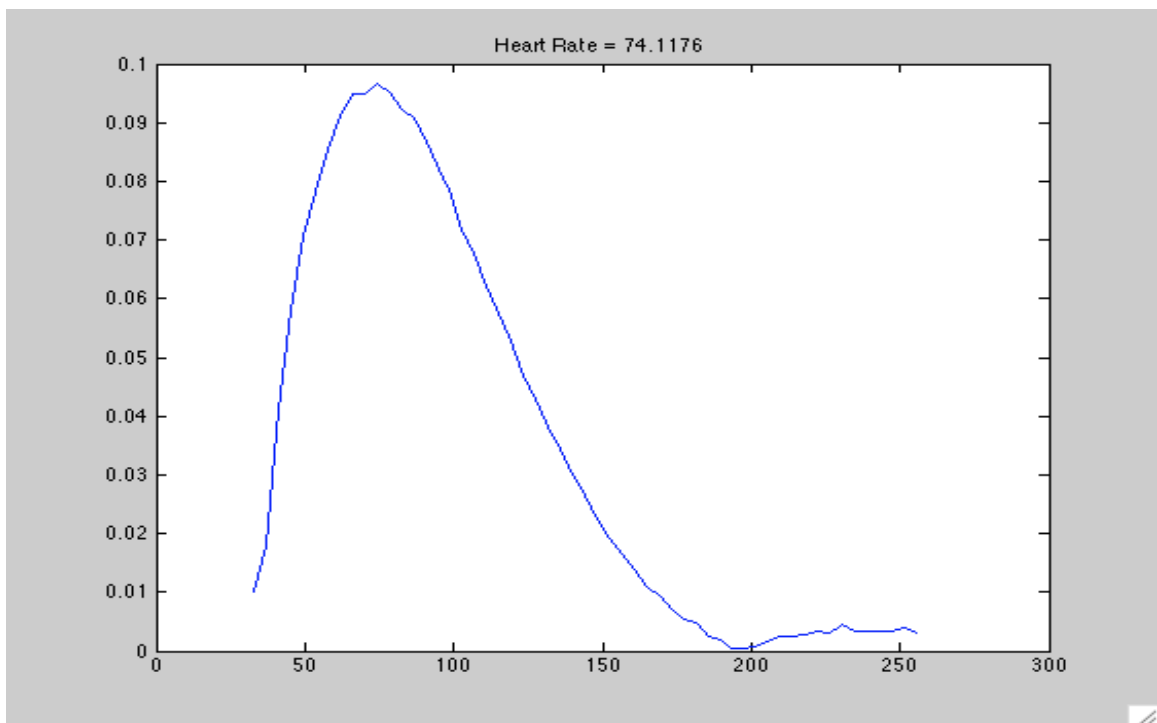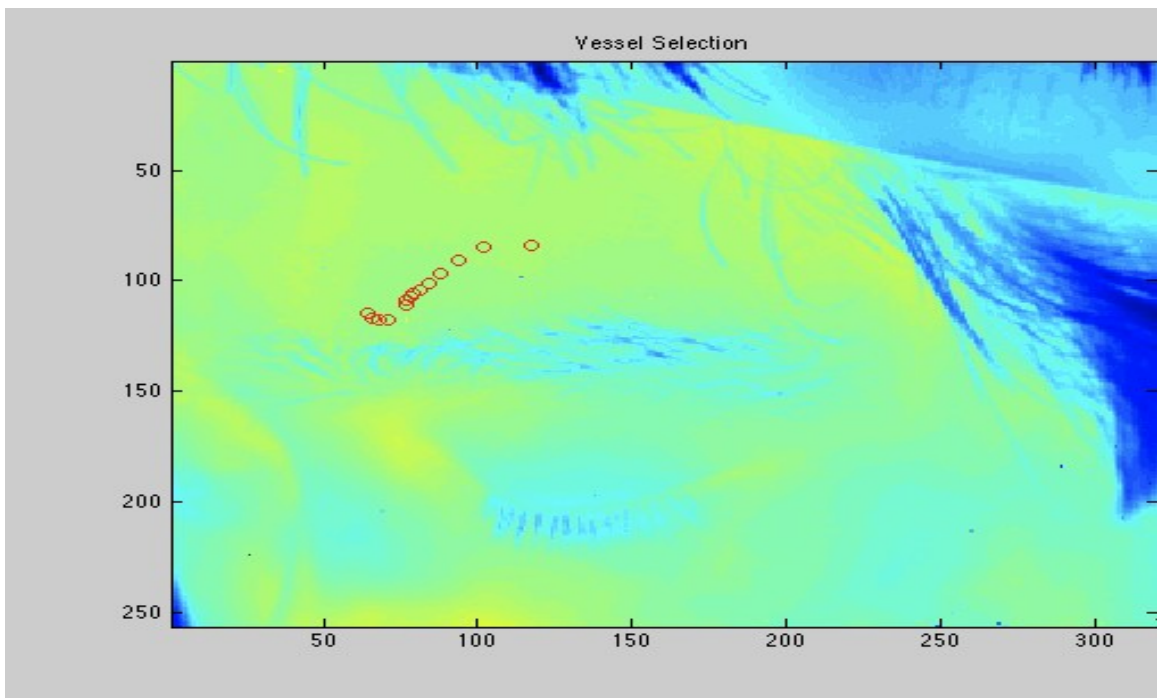
Subject 6, Trial 1, Ground Truth = 81, Calculated = 78.2

Subject 6, Trial 2, Ground Truth = 60.0, Calculated = 65.8

Subject 7, Trial 1, Ground Truth = 60 BPM, Calculated = 61.8 BPM

Subject 8, Trial 1, Ground Truth = 90 BPM, Calculated = 90.6 BPM
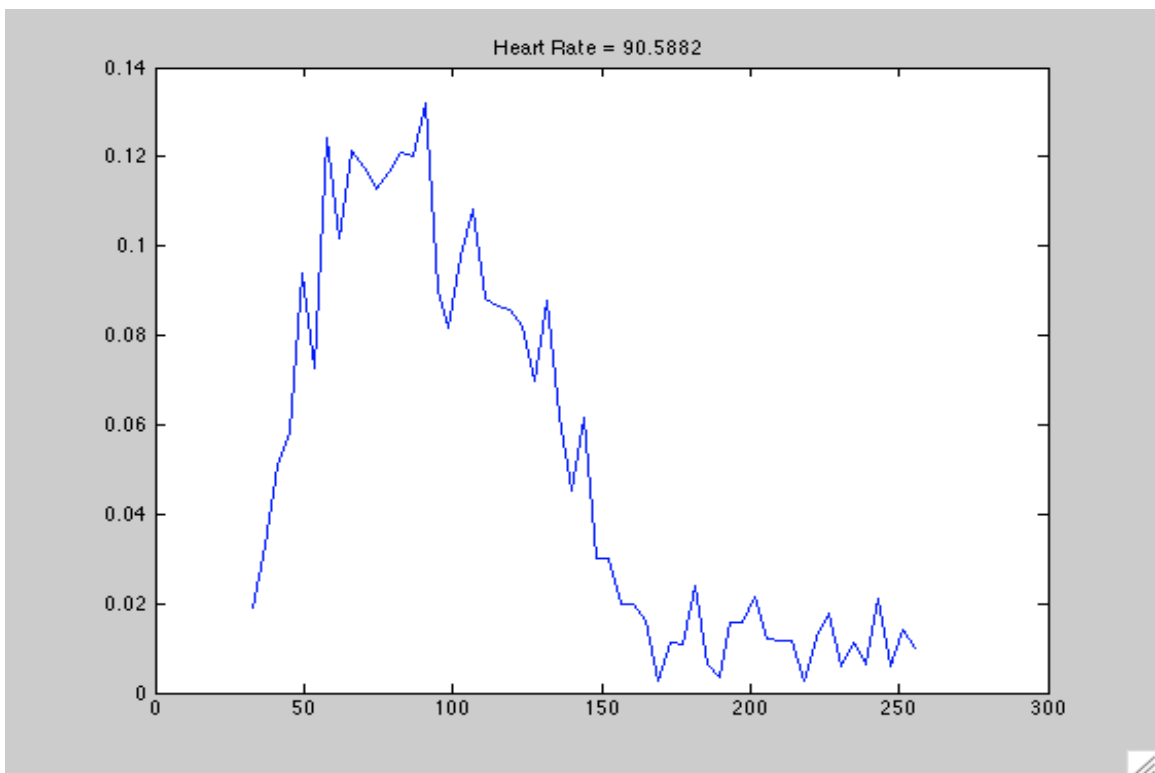
Subject 8, Trial 2, Ground Truth = 85 BPM, Calculated = 61.8 BPM

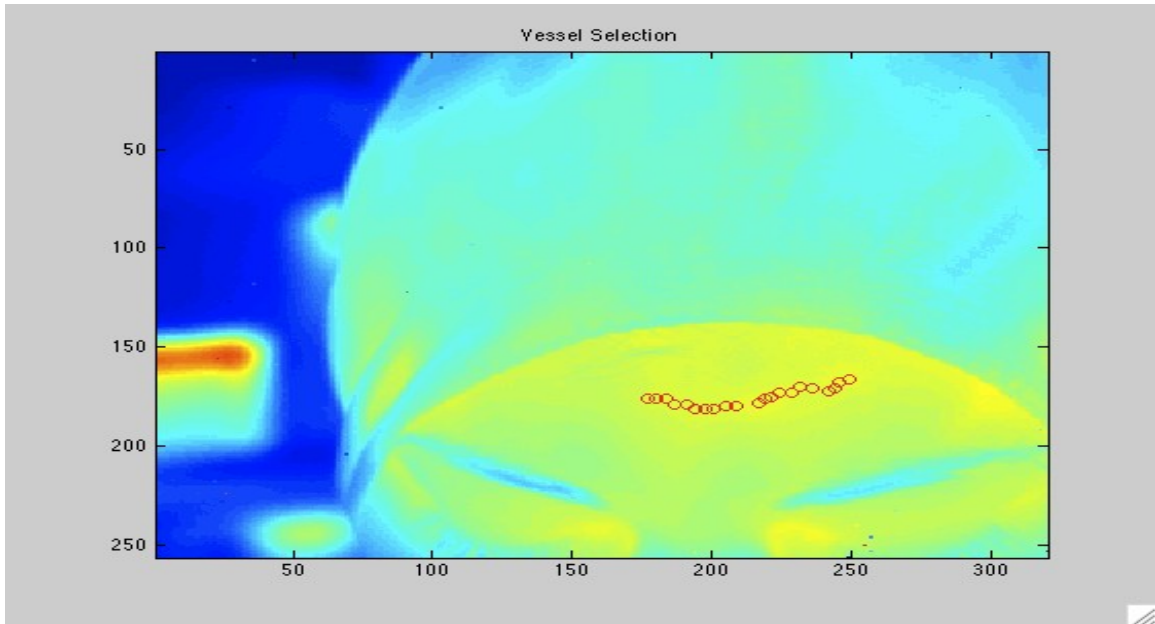**Subject 9, Trial 1, Ground Truth = 88 BPM, Calculated = 78.2 BPM**

100

Subject 9, Trial 2, Ground Truth = 86 BPM, Calculated = 82.4 BPM

Vessel Selection

Heart Rate = 82.3529

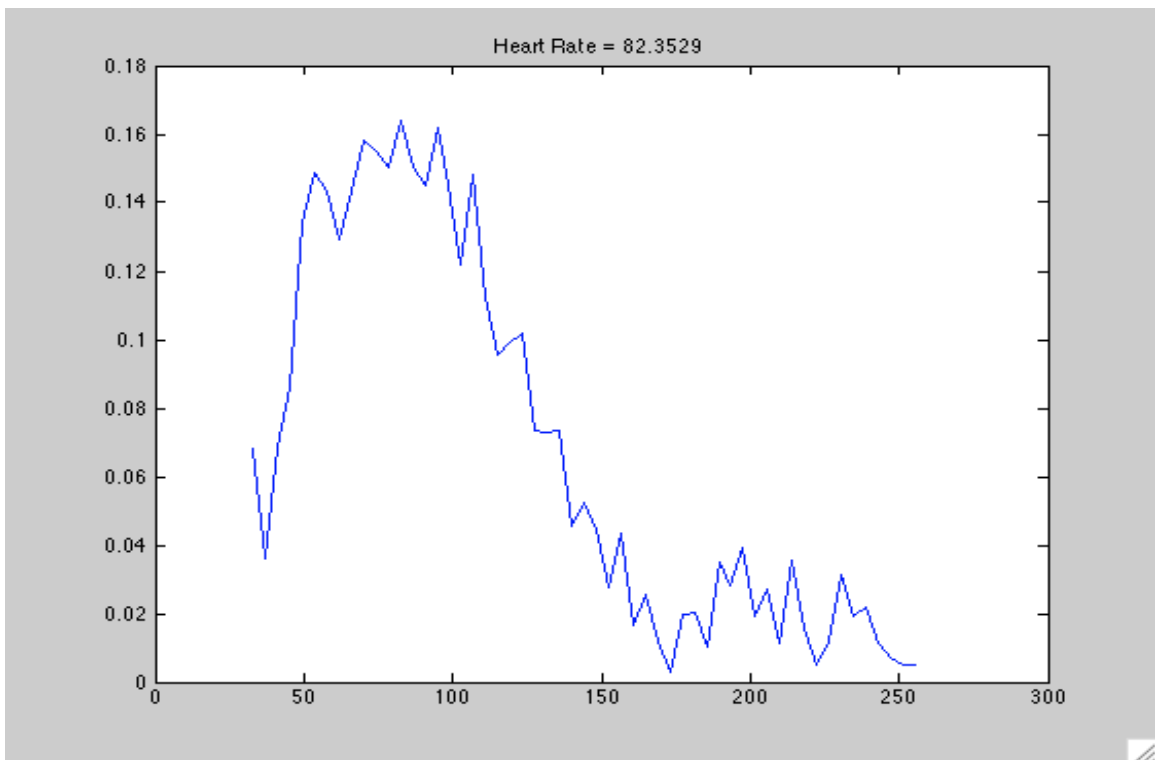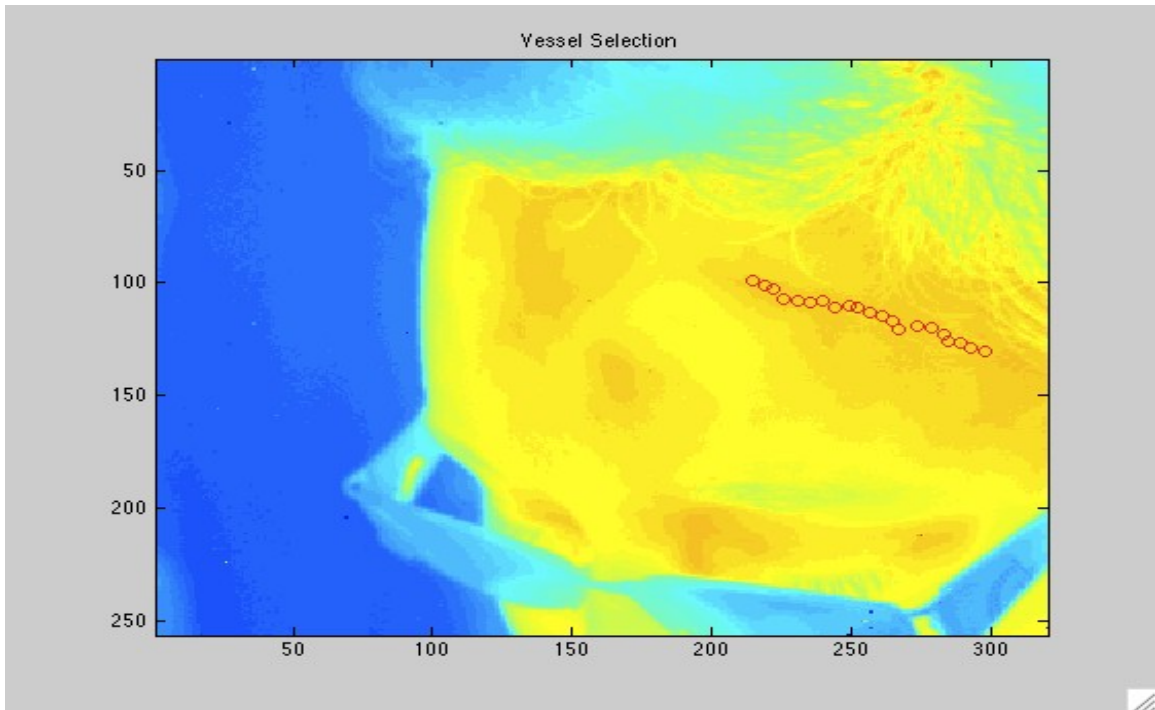**Subject 10, Trial 1, Ground Truth = 70 BPM, Calculated = 74.1 BPM**

Subject 11, Trial 1, Ground Truth = 64, Calculated = 74.1

Subject 12, Trial 1, Ground Truth = 76, Calculated = 82.4

Subject 13, Trial 1, Ground Truth = 74, Calculated = 74.1