Extracting Physiological Measurements from Thermal Images

by

Christian Hessler

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science
(Computer and Information Science)
in the University of Michigan-Dearborn
2018

Master's Thesis Committee:

Assistant Professor Mohamed Abouelenien, Chair
Professor William Grosky
Associate Professor Luis Ortiz

Table of Contents

List of Figures

List of Tables

Abstract

Multiple techniques are used to extract physiological signals from the human body. These signals provide a reliable method to identify the physical and mental state of a person at any given point in time. However, these techniques require contact and cooperation of the individual as well as human effort for connecting the devices and collecting the needed measurement. Moreover, these methods can be invasive, time-consuming, and infeasible in many cases. Recent efforts have been made in order to find alternatives to extract these measurements using non-contact and efficient techniques. One of these alternatives is the use of thermal cameras for health monitoring. Our work explores reliable methods for extracting respiration rate, skin temperature and heart rate from thermal video. These methods leverage a combination of image processing and signal processing techniques in order to extract and filter physiological signals from the thermal domain. Finally, we review the use of thermal imaging in several applications, such as deception detection, stress detection and emotion recognition.

Chapter 1

Introduction

In recent years there has been a growing interest in developing automated systems that are capable of monitoring human physiological responses in order to provide a real-time assessment of a person's general health and well-being. Such measurements include heart rate, temperature, respiration rate, among other skin responses. With the proper assessment, these physiological measurements can identify the physical and mental state of a person. In addition, the fact that the human body often exhibits unique physiological characteristics in response to external stimuli, made it possible to detect and predict a person's behavior or psychological state, such as emotions, mood, stress level, distraction, and deceit. Hence, different studies are exploring the feasibility of incorporating physiological monitoring into a wide array of different applications.

However, there are limitations to the traditional methods and devices used to collect physiological measurements, such as the requirement to connect the devices and sensors to the human body. Attaching these sensors can be time consuming, uncomfortable, and impractical for certain applications. Devices, such as ECG sensors require electrodes to be attached to specific areas of the body. These devices can cause discomfort and may require the presence of trained personnel to set up the device. Other sensors may introduce noise if the leads do not have solid contact with the skin. Even worse, some sensors may not provide reliable measurements outside of a controlled environment. [24] designed a network for monitoring patients' vital signs during health emergencies. The authors noted that exposure to cold temperatures restricts blood flow to the fingers which can disrupt pulse oximeter readings collected from a finger sensor. Therefore, new approaches are proposed to avoid the usage of wearable sensors to collect such data.

In particular, thermal image processing has been proposed as a potential method for acquiring physiological data.

Vital sign monitoring systems generally monitor blood glucose level, blood pressure, pulse rate, electrocardiograph patterns, respiration rate, and temperature [69]. Certain vital signs are thought to be better indicators of specific physiological abnormalities than others. Researchers explored ways to harness physiological data for applications in a number of areas, such as health care, sports, military, and surveillance. Moreover, physiological monitoring may be more effective at diagnosing certain disorders that are difficult to diagnose from external symptoms alone.

For instance, heart rate is useful in diagnosing cardiovascular disease (CVD), which is a leading cause of death worldwide. In particular, there is evidence linking resting heart rate to CVD risk factors, such as hypertension, obesity, family history and work stress [62]. Another example can be seen in studies suggesting that changes in the respiratory rate may be a more effective measure for discriminating between stable patients and those that are at risk. In fact, evidence suggests that an adult with a respiration rate of over 20 breaths per minute (bpm) is probably unhealthy, while an adult with a respiration rate of over 24 bpm is likely to be critically ill [18]. Irregular increases in respiration rates have been observed in patients suffering from panic attacks and sleep bruxism (teeth grinding) [46, 35]. Taking more than one vital sign into account has also proven to be beneficial in diagnosing certain ailments. There is evidence indicating that elevated heart and respiration rates observed immediately after trauma are acute predictors of delayed post traumatic stress disorder [12].

Skin is another vital organ that receives signals from control centers in the brain to maintain the body's core temperature through a process called thermoregulation [19]. Physiological thermoregulation in humans comprises changes in heat dissipation (sweating) and heat generation (shivering) in response to various internal and external thermal stimuli [16]. Thermal imaging utilizes this principle to detect natural thermal radiation emitted by the skin, which can be interpreted in terms of physiological changes [32]. Skin conductance is another physiological measurement that refers to the varying electrical properties of the skin in response to sweat secreted from eccrine sweat glands [63]. The skin becomes more conductive as sweat accu-

mulates. This process reflects the arousal of the sympathetic autonomic nervous system which accompanies various psychological processes [22]. While the usage of thermal images to detect peripheral skin temperature is apparent, recent research has shown interesting potential of using thermal images to extract multiple physiological signals from the human body.

Chapter 2

Extracting Physiological Features

This section presents different methods for extracting heart rate, respiratory rate, skin temperature, and skin conductance from thermal videos. Many of these techniques use a procedure called Eulerian Video Magnification (EVM), which can reveal hidden information by magnifying subtle color changes and imperceptible motions using spatio-temporal processing [67]. This process can indicate subtle variations in the blood flow through the face.

2.1 Heart Rate

Several studies have proposed methodologies to extract heart rate from thermal images by tracking superficial blood vessels on the face. Blood flow regulates skin temperature due to heat exchange between vessels and the surrounding tissue. These changes in skin temperature are most prominent along superficial blood vessels. Extracting the blood vessels from the face is often challenging due to the low contrast between the edges of the blood vessels and the surrounding facial tissue. This is a result of heat diffusion, which creates a smooth gradient temperature between hot and cold areas. Fortunately, there are several methods for segmenting blood vessels from the face to create what is known as a vascular map. One of these methods is called top hat segmentation. There are two forms of top hat segmentation: white top segmentation enhances bright objects and black top hat segmentation enhances dark objects. White top segmentation is effective for enhancing the ridge-like structures of the blood vessels, which are represented by hot or bright areas in the image [13].

Despite the effects of heat diffusion, edge detection methods can still prove to be effective. We applied several well known edge detection algorithms including Canny, Prewitt, Roberts

and Sobel. In our experiments Canny's method proved to be the most effective. Figure 2.1 demonstrates the result of applying the Canny edge detection algorithm to a gray-scale thermal image. The edges detected in the image clearly resemble a vascular structure in the forehead region of interest. However, edge detection alone, may fail to capture the center of the vein where the effect of heat transfer due to blood perfusion is most pronounced. For that reason, we expanded the edges by a factor of one pixel in every direction to ensure that we extract the heat radiating from the center of each vein near the surface of the skin. Figure 2.2 shows the result after applying this technique.

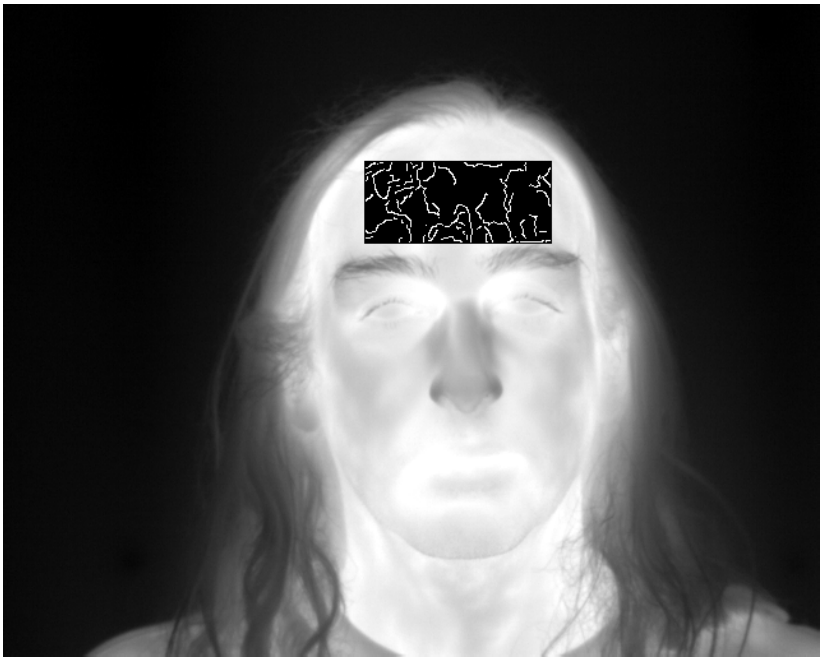

Figure 2.1: Vascular map extraction using Canny edge detection

The thermal signal detected along a blood vessel presents a composite signal that includes extraneous physiological and environmental signals in addition to the pulse [25]. [60] proposed a method to extract the pulse by applying a Fast Fourier Transform (FFT) to several points along the blood vessel in order to isolate the thermal propagation component. They followed this by using an adaptive estimation function to quantify the pulse based on current and past measurements. The authors were able to achieve an overall accuracy of 92.1% based on ground truth measurements collected from a piezoelectric pulse transducer.

In [26], the authors introduced several improvements based on previous work [13, 17,

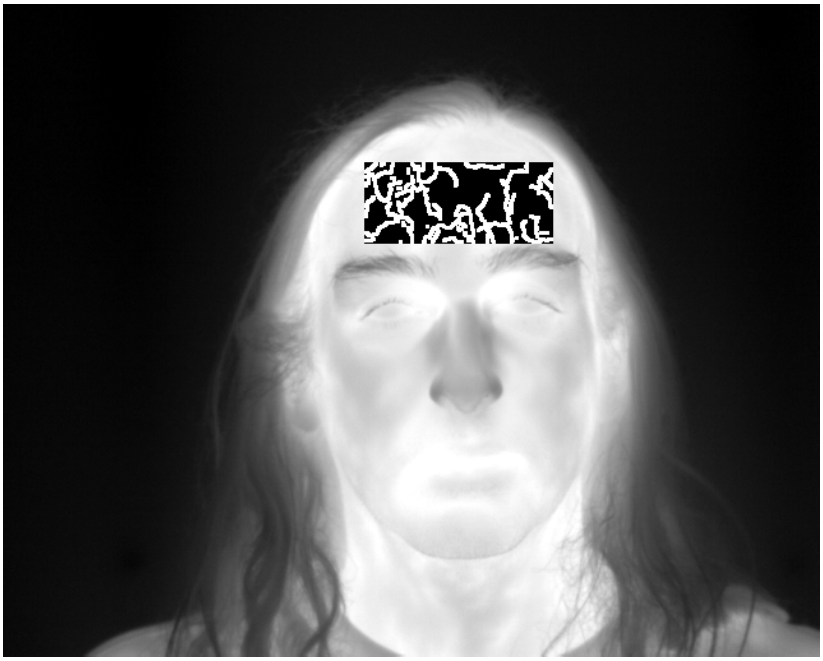

Figure 2.2: Canny edge detection after expansion

27, 28, 60, 59]. First, they incorporated a blood-perfusion model to more accurately create vascular maps, segment the forehead, and enhance the raw thermal data. Second, once they identified suitable blood vessels on the face, they applied wavelet based filtering in place of FFT analysis. In the final step they were able to automate the entire process by presenting a systematic approach to select appropriate vessel segments from the vascular map.

A slightly different approach was taken in [10] for extracting heart rate by applying the Eulerian Video Magnification method to thermal videos. The goal of their research was to remedy the fact that EVM may amplify indiscriminate noise in addition to the true heart rate signal. In their experiment, the subject wore a smart shirt (a shirt containing various textile sensors) to capture the ECG signal while a thermal camera recorded video of the subject. They applied two passes of EVM. The first pass applied a wide band pass filter with a low amplification factor to identify the region of interest (ROI) most likely to reveal the true heart rate. In this case, the subject's chest was defined as the region of interest. The second pass applied a narrow band pass filter with a high amplification factor to the signal acquired from the ROI.

2.2 Respiration Rate

Many methods have been proposed to extract respiration rate from thermal videos using different combinations of image processing and facial tracking techniques. Figure 2.3 depicts the procedure we followed for extracting physiological features from thermal videos. This begins with image correction and enhancement in order to make certain features more distinguishable. Examples of image enhancement techniques were briefly discussed in the previous section. A facial detection algorithm is often employed to segment the face from the background image. Once the face has been isolated, regions of interest (ROI) are defined in order to focus on particular areas of the face that are known to display the desired thermal characteristics. Finally, a variety of image processing techniques are applied to the ROI in an attempt to find a correlation between the temporal features within the thermal and physiological domains. [11] compared temperature-based methods to motion-based methods for extracting respiration rate from thermal videos. The temperature-based method employed segmentation-based image processing and image tracking algorithms to capture temperature variations over time. They presented a variety of pre-processing methods including image enhancement, noise removal, edge-detection, and facial recognition, all of which were used to identify the subject's nostrils as the ROI. The respiration signal was then calculated as the mean intensity within each ROI subjected to low pass filtering to remove noise.

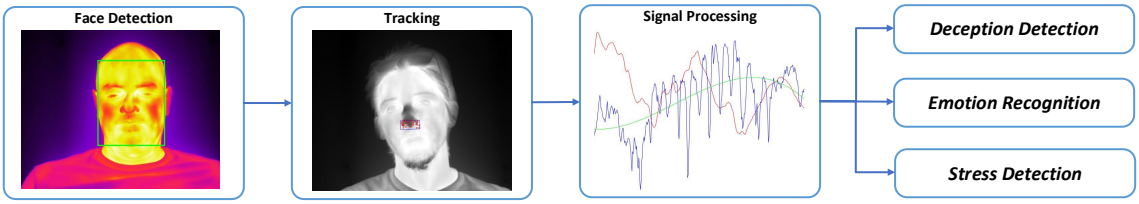

Figure 2.3: Image Processing Procedure

The motion-based analysis was carried out by simply calculating the absolute differences between the first frame and all succeeding frames, then once again applying a low pass filter to remove noise. The temperature-based analysis worked better for detecting the volume of airflow while the motion-based analysis provided better results for detecting irregular breathing, such as hyperventilation and the absence of breath. Neither method outperformed the other in detecting the respiration rate for all breathing patterns. Hence, the authors recommended the development

of fusion algorithms that could combine multiple methods for extracting respiration rates from thermal videos.

[9] developed a facial tracking method to monitor respiration rate in real time. One departure from previous work was the use of an Otsu-based thresholding algorithm to segment the face from the background image. Tsai's method under-segmented areas below the neck and even part of the face. Kapur's method performed slightly better but still under-segmented areas below the face. The Otsu method proved to be the most effective and was even efficient enough to allow for real time face detection and tracking. Lastly, they applied noise filtering techniques and FFT to extract the respiration rate from the ROI. This system was able to process each frame in 40ms, making the system feasible for deployment in real-time applications. Other studies followed a similar procedure to extract the respiration rate using different methods to perform noise removal and signal processing. Additional methodologies for extracting the respiration signal include clustering and harmonic analysis [68], wavelet analysis [21] and high pass filtering [36].

2.3 Skin Temperature

Body temperature depends on core temperature and skin temperature. Generally speaking, core temperature is the temperature of the blood in circulation, which is regulated by the brain, whereas skin temperature is primarily influenced by blood flow and environmental conditions [37]. Heat stress is a condition in which skin blood flow increases, followed by a rise in skin temperature, which releases heat from the body. Cold stress describes the opposite effect in which skin blood flow and temperature decrease, actively conserving heat in the body. Heat stress and cold stress are also respectively referred to as vasodilation and vasoconstriction. This is the process by which the human body is able to maintain a constant core temperature. Modern high resolution thermal cameras have given researchers the ability to observe physiological thermal regulatory response in real time.

A significant variation in body temperature is often an indication of illness such as fever or hypothermia. In the interest of preventing the spread of disease, several studies have explored the feasibility of designing fever-based detection systems for use in airports and mass transits

[58, 43]. The literature also discusses the many challenges involved in designing such a system. As of writing, the only reliable way to acquire an accurate core temperature reading is to measure temperature from the rectum or esophagus [40]. In spite of this, several studies have reported the inner corners of the eyes to be the most suitable area for fever detection [44].

There is also a desire to better understand the relationship between thermoregulation and athletic performance in sports medicine. [61] recorded thermal videos of athletes running on a treadmill as well as their resting states before and after the exercise. Surprisingly, the skin temperature of the athletes began to decline immediately upon starting to run even at low speeds. A continuous increase in exercise intensity caused the skin temperature to decrease even further. On the other hand, thermal images of the athletes during motionless recovery revealed a rapid increase in skin temperature as well as the appearance of hyper-thermal spots.

The hyper-thermal spots are most likely a sign of vasodilation caused by a reduction of warm blood flow to active muscles. These findings are supported by the results of previous work done by [41], in which the skin temperature of trained and untrained subjects was recorded during exercise. Their results revealed that the minimal skin temperature of trained subjects was significantly lower than those of untrained subjects when they stopped exercising. [14] studied the thermal comfort of subjects during and after a 30 minute exercise. Their findings also indicate that the skin temperature of most subjects continues to increase during motionless recovery due to the delay of the heat transfer process.

Thermal images clearly reveal variable amounts of heat radiating from different areas of the human body during periods of rest and physical labor. However, research shows little variation in facial skin temperature in response to cold stimuli and corresponding changes in core temperature. In [31], subject's skin and core body temperatures were measured using thermocouples and an ingestible thermometer pill. When the subjects were exposed to a cold environment, skin temperatures of the hands and feet decreased substantially while the forehead remained reasonably constant. This poses a difficulty for applications that hope to extract body temperature from thermal facial images alone.

Chapter 3

Applications

In this section, we review some of the applications that extracted multiple features from thermal images in order to achieve their goal. Most of these applications are related to modeling of human behavior. Some of them integrated these features with contact-based physiological measurements and, hence, they can potentially benefit from the aforementioned techniques.

3.1 Deception Detection

Polygraph testing remains the standard tool used by law enforcement in the U.S. to verify whether or not a subject is telling the truth during questioning. Polygraph tests monitor the subject's blood volume pulse, respiratory changes, and electrodermal activity. Employing polygraph tests was shown to be unreliable in many cases as it requires decisions from human experts, which is subject to bias and error [20, 23]. Reports dating back three decades indicated that polygraph results were false one third of the time [38].

Hence, research was conducted to find alternatives, including the usage of thermal imaging as a mean for deception detection. Most experiments in this field begin by establishing the baseline physiological characteristics of the subject prior to the interview. This generally involves asking the subject a series of control questions designed to elicit a particular physiological response.

However, research suggests that guilty subjects who are trained on using physical or mental countermeasures are able to defeat polygraph tests by corrupting the initial baseline measurements [30]. Hence, additional information collected from thermal images have the potential to improve the reliability of deception detection models.

In [47], a method was described for classifying a person's responses as deceitful or truthful based on changes in blood flow rate as observed from thermal images of the person's face. In this method, raw thermal data was transformed into blood flow rate data using a number of different processing techniques, such as segmentation algorithms and heat transfer modeling. Although different regions might be used, [48] found that the periorbital region (area around the eyes) carried the most significant discriminating power. They observed that the slope of the periorbital blood flow rate as a function of time grows steeper during a deceptive answer.

Accuracy of thermal imaging as a lie detection tool in airport screening was tested in [65]. Their results revealed that the skin temperature of liars rose significantly during the interview, whereas, the skin temperature of truth tellers remained constant.

Baseline measurements did not reveal any significant difference between passengers who were instructed to tell the truth and those who were instructed to lie. Therefore, the authors concluded that deception detection systems based on skin temperature alone would not be suitable for rapid screening of passengers at an airport.

[54] is another study that analyzed the periorbital region of the face to perform automated deception detection. They tracked the eye corner regions, concatenated the ROI data across all frames within the response time-line, and finally applied principal component analysis to obtain thermal features. One unique aspect of their research was the fact that they compared the predictive ability of a within-person classification to a between-person classification. A between-person approach was shown to have poor predictive performance. The authors explain that a leave-one-person-out cross validation method assumes that behavior and physiological responses are common traits among people of various ages, genders, culture, etc. On the other hand, the within-person approach trains a classifier specific to each subject using the aforementioned baseline measurement as training data. Their model was able to achieve an overall accuracy of 87% using a k-nearest neighbor classifier.

More recent studies propose the use of fusion models that incorporate features from more than one modality [4, 15, 1, 8, 5, 6, 51, 50, 3]. The authors analyzed thermal videos, facial expressions, and other visual features to identify areas of the face that are the most indicative of deceptive behavior. Their approach generated feature vectors by transforming each ROI into

a thermal map represented by the Hue Saturation Values pixel representation. In contrast with previous work by Pavlidis, they found that thermal features extracted from the forehead region were the most effective for discriminating between truth and deceit. This may be attributed to the different methods that were used to extract thermal features; heat transfer modeling versus thermal mapping.

3.2 Emotion Recognition

Many studies in the literature have explored the use of thermal imaging for classifying human emotions. The study of affect states and arousal levels is an emerging topic of interest in both neuroscience and affective computing. However, there are conflicting theories that attempt to explain how nuerophysiological systems activate different emotional states. Recent studies in affective computing have designed classification methods based on a relatively recent idea in nueroscience known as the circumplex model. "The circumplex model of affect proposes that all affective states arise from cognitive interpretations of core neural sensations that are the product of two independent neurophysiological systems" [52]. This model is based on the idea that emotional states are not discrete categories but rather a result of varying degrees of arousal and valence as shown in Figure 3.1.

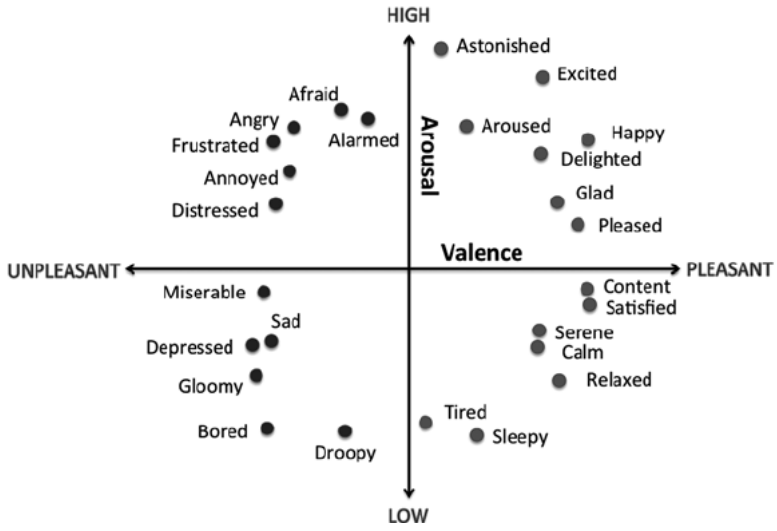

Figure 3.1: Two-dimensional model of valence and arousal. Figure from [64].

[45] designed a binary classifier to distinguish baseline thermal states from affective states. Facial thermal infrared data, blood volume pulse, and respiration rate were recorded while sub-

jects were shown visual stimuli designed to elicit different affective states. Arousal and valence levels during stimulus onset were measured using the International Affective Picture System. The Periorbital, supraorbital, and nasal regions of the face were selected and tracked as regions of interest. Wavelet analysis was used to extract features and remove noise from the thermal infrared data. Finally, a genetic algorithm was used to select optimal features to be used for training a linear discriminate analysis classifier. This classification procedure was able to achieve accuracies of 80% and 75% in classifying high and low levels of arousal and valence from the baseline, respectively.

Other studies took advantage of the fact that different facial expressions are generally associated with certain emotional states. Research has demonstrated that thermal cues may provide a more effective means for recognizing facial expressions compared to visual cues. In [66], a thermal-based facial expression classifier outperformed a visual based classifier due to the fact that thermal images are unaffected by variations in illumination and skin complexion.

[34], leveraged the findings of these two studies to develop a unique classification algorithm. Instead of using a binary classifier, they chose to use a clustering algorithm to model affective states as clusters in a multi-affect and multi-arousal discriminant space. The thermal images were analyzed using accompanying visual images to find points along major facial muscles that displayed the greatest thermal variation. Previous research explains this method for acquiring Facial Thermal Feature Points [33]. Principal Component Analysis and Linear Discriminate Analysis were used to perform dimensionality reduction and feature selection. The resulting facial thermal vectors are used to construct a smaller feature space. The authors used the distance between the optimal feature vectors and the centroids of arousal levels for affect assessment. Their thermal expression recognition system was able to correctly classify approximately 96% of happy and sad expressions.

3.3 Facial Recognition

Facial recognition technology has advanced rapidly within the last decade. In fact, many of today's smart phones offer features that rely on facial recognition and tracking software. However, most of the research in the field has focused on detecting and recognizing faces in the

visible spectrum. Visual facial recognition systems read visible light reflected off the surface of the skin to track facial features. As a result, these systems usually do not perform well in variable lighting conditions. Facial recognition in the thermal infrared spectrum has been proposed as an alternative to overcome these problems. Thermal imaging is insensitive to illumination changes and capable of detecting unique physiological characteristics beneath the skin.

The forehead region is useful for facial recognition since it is a uniform surface that overlays several superficial arterial branches [28]. There are a number of methods for segmenting the thermal imprint of these supraorbital vessels [71]. The vascular mapping can be used to classify the subject's face and to track the subject's movements [27]. The authors in [13] proposed a method for thermal facial recognition based on the fact that the contrast between the superficial vasculature and the surrounding tissue is a physiological characteristic that does not change over time.

3.4 Stress Detection

"Following the perception of an acute stressful event, there is a cascade of changes in the nervous, cardiovascular, endocrine, and immune systems" [55]. Furthermore, clinical studies have demonstrated relationships between psycho-social stressors and diseases, such as cardiovascular disease, upper respiratory diseases, immunodeficiency and depression [39, 55]. Physiological responses to stress may include an increase in blood pressure, redirected blood flow, and vasoconstriction as well as dilated pupils, accelerated heart rate, paling or flushing in the face, and an increase in perspiration [70].

Research has successfully demonstrated the use of thermal imaging to detect the onset of stress from observing physiological changes in subjects' faces. Different activities produce distinct facial thermal patterns. Thermal videos of anxious subjects who were exposed to stressful situations revealed an increase in temperature around the eyes and forehead as a result of heat dissipation caused by increased blood flow [49, 53].

Recent studies have classified these thermal physiological markers to develop automated stress detection algorithms. [57] analyzed spatio-temporal facial patterns in videos captured from both the thermal (TS) and visible (VS) spectrums. Subjects were recorded watching stress-

ful and calming video clips. The authors extracted features from the videos using a technique known as local binary patterns on three orthogonal planes (LBP-TOP). This method was specifically used to analyze the temporal dynamics of muscle movements by extracting features, which incorporated appearance and motion. In addition, they proposed a new feature set to model thermal images, which captured normalized dynamic thermal patterns in histograms (HDTP). The goal of this method is to enhance participant-independent recognition of symptoms for stress and to reduce individual bias. The HDTP features extracted from the thermal videos produced better stress recognition rates compared to the LBP-TOP features used for binary classification. A fusion of HDTP and LBP-TOP features extracted from TS and VS video, respectively, achieved the best results with a recognition rate of 72%.

In [2], contact-based physiological measures and facial thermal images were used to train a stress detection classifier. Ground truth measurements were based on the perceived stress of subjects in stressful situations. Thermal features were extracted using a variety of methods to perform face segmentation, tracking, and transformation. Facial bounding boxes were manually defined, the Shi-Tomasai corner detection algorithm was used to identify discriminating points within the face. A fast Kanade-Lucas-Tomasi (KLT) tracking algorithm was used to track the points throughout the entire response. The background in the image was discarded using image binarization and cropping. Lastly, features were extracted by creating a thermal map in which hue saturation value (HSV) colors represented temperature values. HSV values were organized in a histogram and normalized to form a probability distribution over all bins. Moreover, the thermal features were integrated with the contact-based physiological features including the heart rate, respiration rate, skin temperature, and skin conductance. They trained a decision tree classifier using their features, which was able to detect stress with an accuracy of 75%.

Chapter 4

Dataset Description

We used our own dataset consisting of multimodal recordings collected from 104 undergraduate and graduate students with a gender distribution of 53 females and 51 males. The subjects were between 20 to 45 years of age and had different ethnic backgrounds. Seven sessions were recorded per subject. Two of these sessions were used as baseline measurements, in which the subjects were instructed to silently sit still. The remaining five sessions feature the subjects engaging in conversation. Each session consists of a thermal video recording of the subject's face in addition to several contact-based sensor readings. Therefore, a total of 728 thermal videos with accompanying sensor readings were analyzed.

Chapter 5

Experimental Setup

The camera used to record the videos was a FLIR SC6700 thermal camera with a resolution of 640x512 and 7.2 M electrons capacity, reaching a frame rate of approximately 100 frames/second. Physiological data was collected using Thought Technology's FlexComp Infiniti sensors. These four bio-sensors were used as the ground truth; a blood volume pulse sensor, skin conductance sensor, skin temperature sensor and an abdominal respiration band. The first three sensors were attached to the fingers of subject's non-dominant hand. The abdominal respiration band was placed around the thoracic region.

Our experimental station consists of recording devices, the physiological sensors, two desktop computers, and a chair placed at a fixed distance from the cameras. The experimental setup and procedure were explained to the subjects and they were asked to avoid excessive movements to keep them in the field of view of the cameras. [7]

Chapter 6

Methodology

We perform a series of steps to extract physiological features from the thermal domain. First, we define and track multiple regions of interest throughout the video to compensate for any displacement caused by subject movement. Secondly, we construct a raw thermal signal by sampling the temperature within the region of interest defined for each frame. Finally, we filter the raw thermal signals in order to extract various physiological signals.

6.1 Region of Interest Identification & Tracking

We begin by defining regions on the face which are known to display significant temperature variations based on related work in the literature. These regions are generally the maxillary region (surrounding the nose), periorbital region (surrounding the eyes) and supraorbital region (forehead). We manually created bounding boxes to define each region within the first frame of every video. Tracking was performed using the point tracking feature included in the Matlab Computer Vision System Toolbox. Good Features to Track within the ROI were identified using the Shi-Tomasi corner detection algorithm. These points are located by calculating image derivatives based on pixel intensity values. If the change in intensity is greater than a certain threshold in both the x and y directions then the point is labeled as a corner. These features were then passed to the point tracker object, which uses the Kanade-Lucas-Tomasi (KLT) feature-tracking algorithm to stabilize the region for the duration of the video.

After detecting 'interesting points' to track we apply a geometric transformation, which estimates the location of these points from one frame to the next using a variant of the Random Sample Consensus (RANSAC) algorithm. Considering a frame-rate of 100fps, we do not expect

any rapid motion between successive frames. Therefore, the distance between a point and its projection in the next frame should be minimal. For that reason, we limit this distance to five pixels. Secondly, if the number of mapped points between two successive frames is less than 95% we skip the current frame and resume tracking in the next frame. This is a precaution to account for potential occlusion.

We found that the tracking algorithm was robust against most head movements. However, a few of the subjects either turned their head 90 degrees left or right as shown in figure 6.1 or looked down as shown in figure 6.2. These types of movements would cause the bounding box to go outside of the camera view. In many cases, the point tracker can still reliably track a region that is partially off-screen. Nonetheless, our intention is to capture the entire region of interest to extract the desired signal. This is especially true for the heart rate signal in which accurate vein segmentation within the ROI is a critical part of our proposed method. For the sake of avoiding corrupt data collection, we simply discarded videos in which tracking errors were detected. This was only necessary for extracting heart rate signals. Two videos were discarded when extracting heart rates from the forehead region. One video was discarded when extracting heart rates from the inner eye corner region.

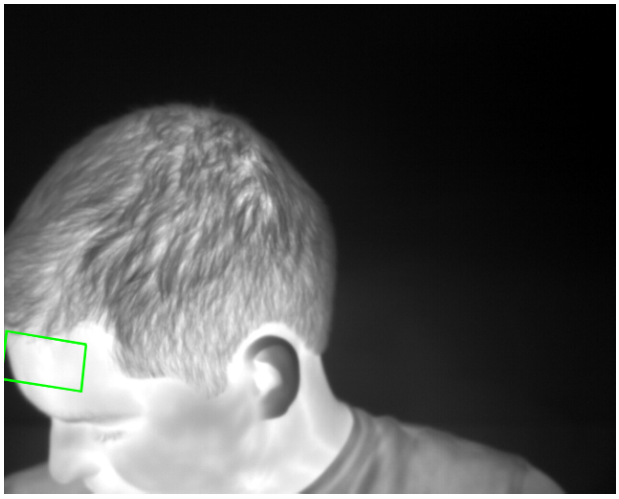

Figure 6.1: Tracking Error: Head Turn (ROI outlined in green)

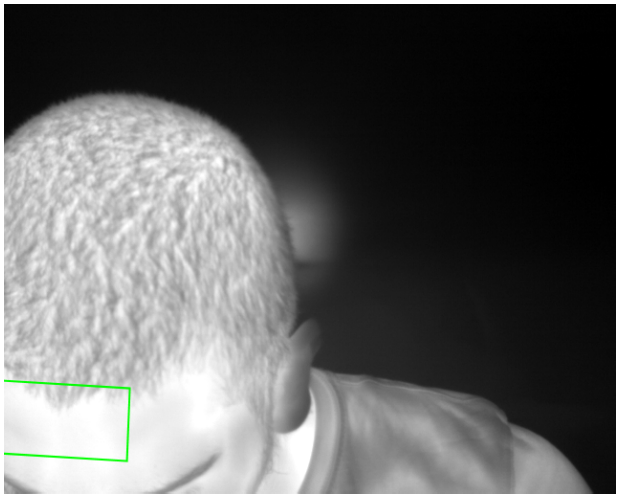

Figure 6.2: Tracking Error: Looking Down (ROI outlined in green)

6.2 Feature Extraction

The previous tracking output data defines the size and location of a bounding box within each frame for every video. Each bounding box was masked against the raw thermal data in the corresponding frame to extract temperature values within the region of interest. Different segmentation methods were performed to extract the desired temperature values. A thermal signal was then constructed by averaging these values for every frame in the video. The resulting signal has a sampling frequency of 100Hz. A series of signal processing methods were then used to filter the thermal signal and isolate the physiological signal of interest.

6.3 Respiration Rate

We developed a method for extracting respiration rates from the maxillary (nose) region of a thermal image. We tested our method using a novel dataset, which includes recordings of subjects while they are speaking and sitting silently. Thus, we were able to observe the influence that speaking has on nasal breathing and how it affects thermal image based respiration rate estimation. We also calculated the correlation between breathing signals extracted from the thermal images and those measured by the abdominal chest strap sensor. The raw thermal signal was constructed by calculating the average pixel value within the maxillary of interest for each frame of the video.

1) Differencing: We calculate the differences between adjacent elements of the signal $S(t)$ to produce the transformed signal $\hat{S(t)}$

$$\hat{S(t)} = S(t) - S(t-1) \tag{6.1}$$

2) Normalization: We normalized the signal amplitude as follows. μ and σ are the mean and standard deviation of S_t respectively. The transformed signal $\hat{S(t)}$ has mean $\mu = 0$ and standard deviation $\sigma = 1$.

$$\hat{S(t)} = \frac{S(t) - \mu}{\sigma} \tag{6.2}$$

3) Averaging: We down-sampled the signal to 25Hz by calculating the average across every $k = \nu/25$ samples.

$$\hat{S(t')} = \frac{\sum_{i=t}^{t+k} S(t)}{k} \tag{6.3}$$

4) Continuous Wavelet Transform: Mexican Hat a.k.a. Ricker Wavelet was used as the mother wavelet $\psi(t)$. Equation 6.4 describes the Ricker Wavelet in which, σ is the standard deviation and *t* represents time. Equation 6.5 describes the continuous wavelet transform in which *S* is the input signal function, *t* is time, *a* is the scale value and *b* is the translation value.

$$\psi(t) = \frac{2}{\sqrt{3\sigma}\pi^{1/4}} \left(1 - \left(\frac{t}{\sigma}\right)^2\right) e^{-\frac{t^2}{2\sigma^2}} \tag{6.4}$$

$$S_w(a,b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} S(t)\psi\left(\frac{t-a}{b}\right) dt \tag{6.5}$$

5) Breathing Waveform: Equation 6.6 was used to select the scale which best represents the breathing component in which, *WT* is the wavelet transform function, *i* is the scale index and *t* is time. This scale is defined as $s_{\max}$ which corresponds to a local maximum of the energy wavelet coefficients $WT_s(t)$ [21]. Lower scales are likely to contain noise, while higher scales contain metabolic contributions.

$$s_{\max} = \text{argmax}\left\{ \sum |WT_i(t)|^2 \right\} \tag{6.6}$$

Applying the s_{max} formula to the ground truth signal consistently indicated a scale corresponding to an appropriate waveform. However, applying the s_{max} formula to the thermal signal would sometimes indicate a scale corresponding to an under-filtered or over-filtered signal. We attribute this to the fact that the thermal signal contains noise and metabolic contributions in addition to breathing. In contrast, we consider the ground truth signal collected from the bio-sensors to be an accurate representation of the breathing waveform. In order to address this, we randomly sampled 25% of the thermal data, applied the continuous wavelet transform, performed the s_{max} calculation, and plotted the resulting scale values. Based on the box plot, we introduced a constraint which states that the scale value must lie within the interquartile range (index 7 - 11 or scale 0.6076 - 2.3475) shown in Figures 6.3 and 6.4. If the s_{max} formula returns a scale outside of this range we set the scale equal to the mean. The mean value was found to be the 8th value, which corresponds to a scale of 0.8518 as shown in table 6.1.

Index	Scale
1	0.0800
2	0.1122
3	0.1573
4	0.2205
5	0.3091
6	0.4334
7	0.6076
8	0.8518
9	1.1943
10	1.6744
11	2.3475
12	3.2913
13	4.6144
14	6.4695
15	9.0703
16	12.7166
17	17.8289
18	24.9963
19	35.0452
20	49.1338

Table 6.1: CWT Scale Parameter Values

6) Rate Calculation: Finally we calculate the respiration rate based on the resulting waveform. We begin by counting the number of peaks in the wave. However, some waveforms

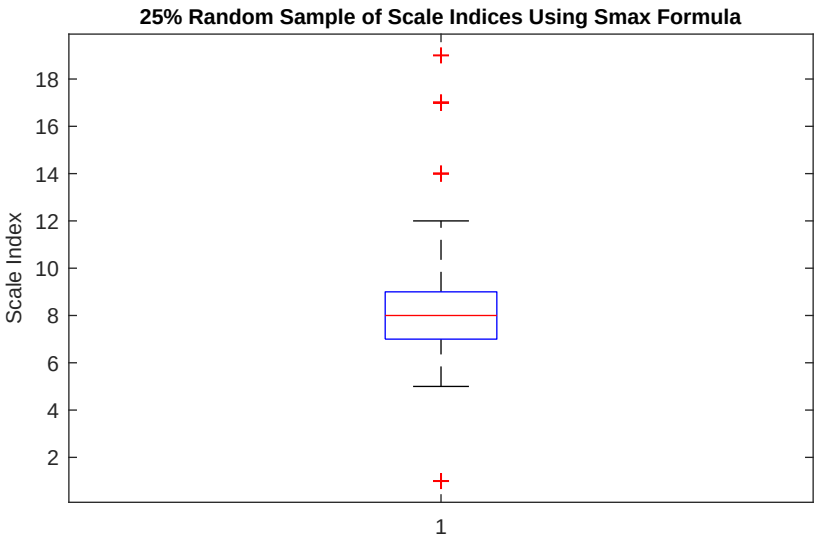

Figure 6.3: Scale Indices

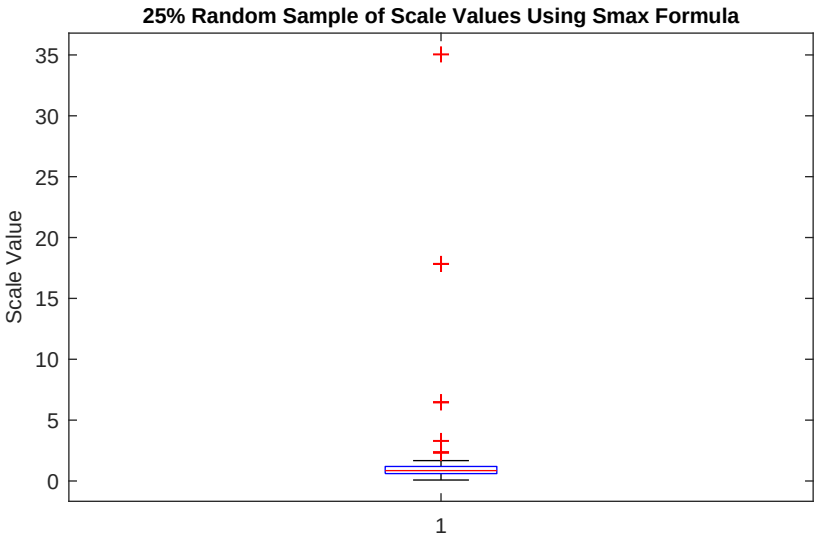

Figure 6.4: Scale Values

contain several smaller peaks, which are not consistent with the breathing function. To address this issue we define a constraint known as *MinPeakProminence* when selecting the peaks. *MinPeakProminence* measures the height of a peak relative to other nearby surrounding peaks. This option is particularly useful for our dataset since the signal level varies across subjects as well as between the sensor and thermal signals. We set this parameter to the standard deviation of the signal in order to filter out noise relative to the signal level. Figure 6.5 depicts an example of employing the approach on the filtered thermal respiration rate signal.

7) Cross Correlation: Cross correlation was used to find the maximum correlation between shifted copies of the sensor and thermal signals. We then calculate the Pearson correlation coef-

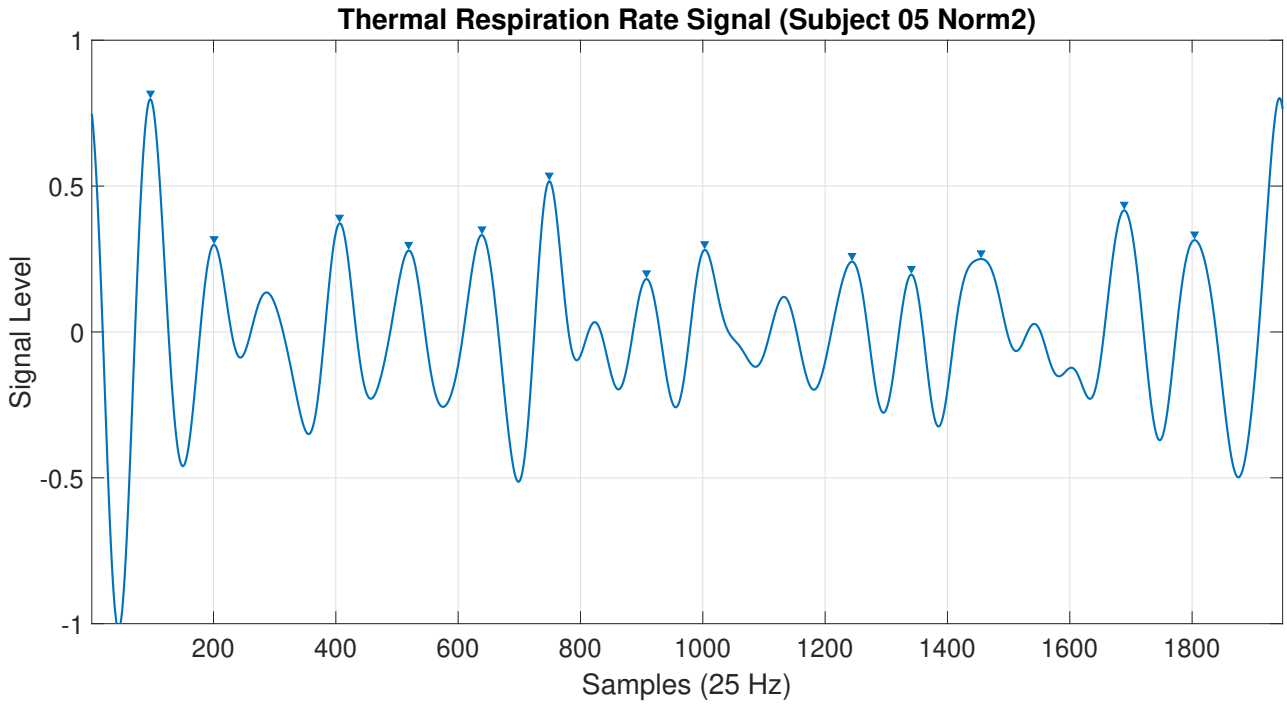

Figure 6.5: Example of Identifying Peaks in the Thermal Signal

ficient of the shifted signals. Table 6.2 provides a general guide for interpreting the correlation coefficient in the context of medical applications [42]. In this study we consider a correlation above 0.5 to be statistically significant.

Size of Correlation	**Interpretation**
0.9 to 1.0	Very high correlation
0.7 to 0.9	High correlation
0.5 to 0.7	Moderate correlation
0.3 to 0.5	Low correlation
0.0 to 0.3	Negligable correlation

Table 6.2: Rule of Thumb for Interpreting the strength of a Correlation Coefficient [29]

6.4 Experimental Trials

We tested several different methods and settings before deciding to apply the solution outlined in the previous section. Earlier in our studies we averaged the signals to reduce the sampling rate in order to remove noise and other extraneous frequency information. We analyzed signals with sampling rates of 1Hz, 5Hz, 25Hz, 50Hz and 100Hz. Furthermore, we performed several trials in which we reduced the sampling rate before and after applying the continuous

wavelet transform.

The non-stationary nature of the breathing signal introduced challenges in calculating an accurate respiration rate. We evaluated a number of methods to overcome this problem. The first of which was to divide the signal into 15-second windows, calculate the rate within each window, then calculate the average rate across all windows. Our second approach was to use regression analysis to fit a sine wave to the breathing signal to construct a consistent respiration signal from which we could extract a stable respiration rate. However, after conducting many trials utilizing different frequencies and order of operations, we found that reducing the frequency to 25Hz before applying the continuous wavelet transform produces better results.

We came to this conclusion based on the following findings. Reducing the sampling rate to 25Hz cleans the signals without risk of neglecting the breathing signal. The continuous wavelet transform analyzes the signal in both the time and frequency domains. Finally we count the number of peaks in the signal and scale that number by the duration of the signal to calculate the respiration rate. In doing so, we also experimentally determined that we only wish to include peaks with an amplitude greater than one standard deviation from the mean of the signal.

6.5 Skin Temperature

Our research explores the effects of thermoregulation across different parts of the body by investigating heat exchange rates in the sensor and thermal domains. We used the ProComp Infiniti Temp-Flex/Pro sensor to measure the skin temperature of the subjects as the ground truth. This device converts changes in temperature to changes in an electrical current. The temperature sensor was strapped to the little finger. Studies in the literature suggest that changes in the core and overall skin temperature occur very gradually over a period of minutes or even hours. For that reason, we employed multiple linear regression to capture the long term trend of the thermal signal extracted from the forehead region. Our regression routine performs several iterations to find the best fit. We begin with a linear model consisting of one intercept and one predictor variable. Upon each iteration we add an additional predictor variable which becomes the leading order term. Thus, the model goes from being linear to quadratic to cubic etc... We used the coefficient of determination (R-squared value) to determine the 'goodness of fit'.

Equation 6.7 is the proportion of the total sum of squares explained by the model, in which SSE is the sum of squared error and SST is the sum of squared total.

$$R^2 = 1 - \frac{SSE}{SST} \tag{6.7}$$

We evaluate the rate of change in the R-squared value as leading order terms are added to the model. The 'adjusted R-squared' metric is often used to determine the optimal number of predictor variables to include in a model. However, using the adjusted R-squared metric led to over-fitting for our purposes. To solve this problem, we experimentally determined a threshold of 25% to prevent over-fitting. Therefore, we stop adding additional terms when the R-squared value increases by less than 25% from the previous iteration.

6.6 Heart Rate

We developed our own method for estimating the resting heart rates of subjects using thermal data collected from the inner corner of the eyes. The blood volume pulse measured from the index finger was used as the ground truth. We used the Procomp Infiniti BVP-Flex/Pro to measure the blood volume pulse of the subjects. Blood volume pulse is also known as photoplethysmography in the literature. This type of sensor measures the amount of blood present in the skin by emitting an infra-red light against the surface of the skin. Blood reflects red light but absorbs other colors. It is therefore possible to acquire the blood volume pulse by observing the amount of light that is reflected over time. The Procomp Infiniti device provides the raw blood volume pulse signal, which we then use to calculate the heart rate. Our heart rate extraction method follows the same principles used by the respiration rate method outlined in section 6.3. We performed some additional steps and modified some of the signal transformations to account for attributes specific to the heart rate signal.

1) Segmentation: The Periorbital region was used as the region of interest. This is a departure from previous works, which often use other regions of either the face or neck to extract heart rate measurements. Image binarization was used to segment areas within the ROI, which exhibit thermal characteristics related to blood flow. For each video we define a threshold based

on the 25% hottest pixels in the image. This threshold is required to ignore temperature fluctuations which result from the effects of heat diffusion explained in section 2.1. Figure 6.6 shows an example in which white pixels within the bounding box are used to construct the thermal signal. Black pixels indicate temperature values below the threshold, which are ignored.

Figure 6.6: Image Binarization Applied to the Periorbital Region

2) Averaging: We down-sampled the signal to 100Hz by calculating the average across every $k = \nu/100$ samples.

$$\hat{S(t')} = \frac{\sum_{i=t}^{t+k} S(t)}{k} \tag{6.8}$$

3) Sensor Signal Maximal Overlap Discrete Wavelet Transform: We applied a discrete wavelet transform to the sensor signal as a precaution to remove noise due to sensor error. This step may not be required if the sensor is securely attached to the finger under normal conditions. Daubechies 10 (db10) wavelet was used as the mother wavelet. We selected the DB10 wavelet because in most cases it closely resembled the characteristics of the raw BVP signals. Figure 6.7 provides a visual representation of the DB10 wavelet with corresponding scaling function.

4) Thermal Signal Continuous Wavelet Transform: Morse Wavelet was used as the mother wavelet $\psi(t)$, where $U(\omega)$ is the unit step, $\alpha_{\beta,\gamma}$ is a normalizing constant, γ characterizes the

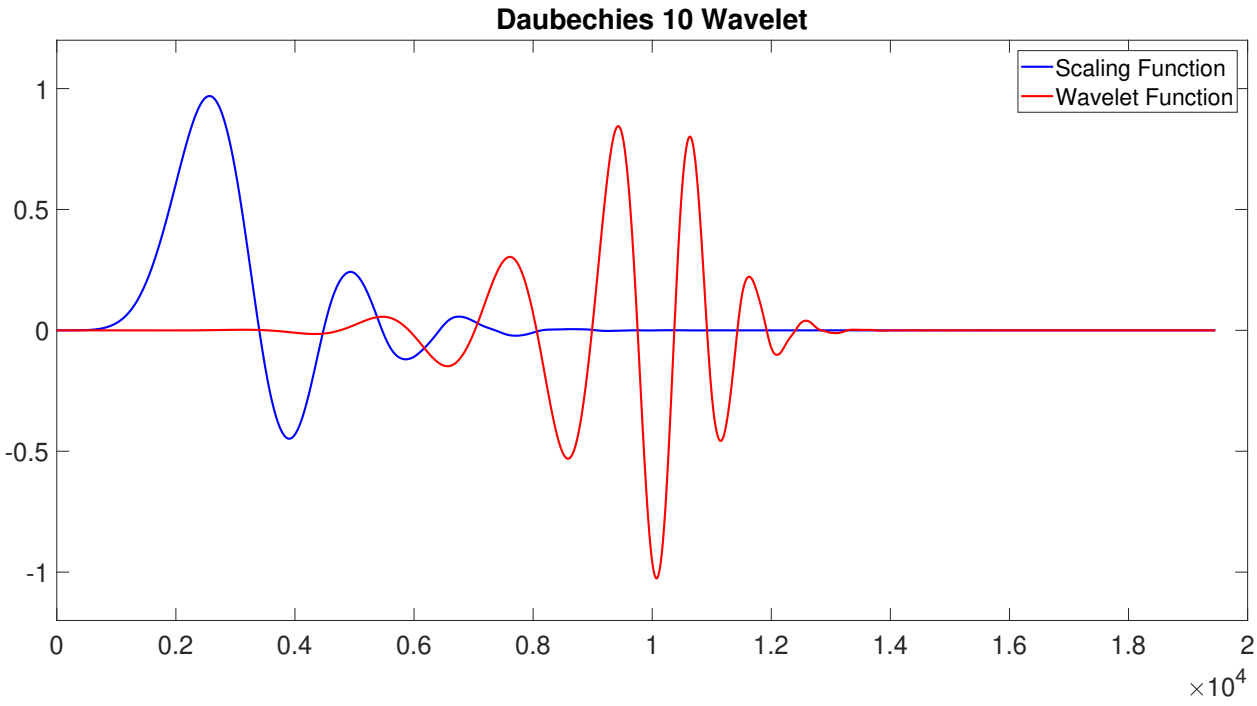

Figure 6.7: DB10 Wavelet & Scaling Functions

symmetry of the Morse wavelet and β is the decay parameter.

$$\psi_{\beta,\gamma}(\omega) = U(\omega)\alpha_{\beta,\gamma}\omega^{\beta}e^{\omega^{\gamma}} \tag{6.9}$$

5) Wavelet Scale Selection: We experimentally found that discrete wavelet scale 32 provided the best set of coefficients to use for calculating an accurate heart rate from the sensor signal. For the thermal signal we select scales which correspond to frequencies in the range of 1-1.67 Hz, which approximately corresponds to 60-100 beats per minute. We then reconstruct the signal from this frequency band using the inverse continuous wavelet transform function. In equation 6.10, $<>$ denotes the inner product and Re$\{\ \}$ denotes the real part of the function.

$$f(t) = 2Re\left\{\frac{1}{C_{\psi,\delta}}\int_0^{\infty} < f(t), \psi_{a,b}(t) > \frac{da}{a}\right\} \tag{6.10}$$

6) Rate Calculation: We count the number of peaks in the wave then scale that value by the length of the recording to calculate the heart rate in units of beats per minute. This calculation is identical to the rate calculation outlined in section 6.3.

Chapter 7

Results

7.1 Respiration Rate

We use the ground truth measurement as a baseline to evaluate the respiration rate which we extract from the thermal domain. Consequently, an accurate baseline measurement is critical for assessing the accuracy of our results. The data set was divided into two sets, which we will refer to as 'Norm' and 'Talking'. In the Norm videos, subjects are silently sitting still without being exposed to any stimuli. In this case, the subjects are likely breathing through their noses. However, in the Talking videos, the subjects are speaking most of the time. Speaking disrupts nasal breathing in order to blow air against the vocal chords. This will affect the thermal respiration rate we extract from the maxillary region of interest.

The literature reports a valid respiration rate to be between 12 and 24 breaths per minute (bpm). We assume the respiration rates of all subjects to fall within this range since they were all relatively young and healthy college students. We consider respiration rates outside of this range to be outliers. Samples containing invalid ground truth respiration rates were therefore eliminated. We analyzed the Norm set first since we expect subjects to exhibit natural breathing patterns while they are in a relaxed state. Eight videos produced respiration rates below 12bpm, while 43 videos produced respiration rates above 24bpm. A total of 51 videos were eliminated from a total of 209, which accounts for 24% of the total videos.

Comparison of the ground truth and thermal respiration rates extracted from the Talking set reveal a higher variation and a lower correlation between the two signals as seen in Table 7.2. 113 videos produced respiration rates less than 12bpm, while 34 produced respiration rates

above 24bpm. A total of 147 videos were eliminated from a total of 519, which accounts for 28% of the total videos.

Table 7.1 shows that the majority of the eliminated videos had respiration rates which are considered to be too low. This supports our assumption that invalid measurements are most likely the result of an ill fitted respiration sensor. Examples of this could include excess clothing worn underneath the sensor or the chest strap not fitting snug against the patient's chest. Either of which, could explain a lower reported respiration rate. Furthermore, the Institutional Review Board (IRB) for Protection of Human Subjects in Research defines guidelines which limit the constraints we can impose on the subjects. For that reason, subjects were left to their discretion in deciding what clothes they wore during the study or how tight they chose to adjust the chest strap.

	Category	
Respiration Rate (bpm)	Norm	Talking
< 12	8	113
> 24	43	34
[12 24]	158	372
Total	209	519

Table 7.1: Video Categorical Breakdown

Table 7.2 displays the average correlation rates before and after removing outliers. Removing outliers resulted in a three percent increase for both sets of videos.

Category	Correlation Before	Correlation After
Norm	0.58	0.61
Talking	0.40	0.43

Table 7.2: Average Correlation Rates Before & After Removing Outliers

Video	Ground Truth Respiration Rate	Thermal Respiration Rate	Cross Correlation
050 Norm2	15.70	16.39	0.98
025 Norm1	19.48	19.48	0.98
096 Norm1	18.39	18.39	0.97
048 Norm2	15.67	15.67	0.97
039 Norm1	22.70	22.03	0.97
040 Norm1	18.25	18.25	0.97
034 Norm2	19.36	18.70	0.97
031 Norm1	21.50	21.50	0.96
097 Norm1	15.85	15.18	0.96
047 Norm1	16.62	15.98	0.96

Table 7.3: Ten Highest Correlated Respiration Rates

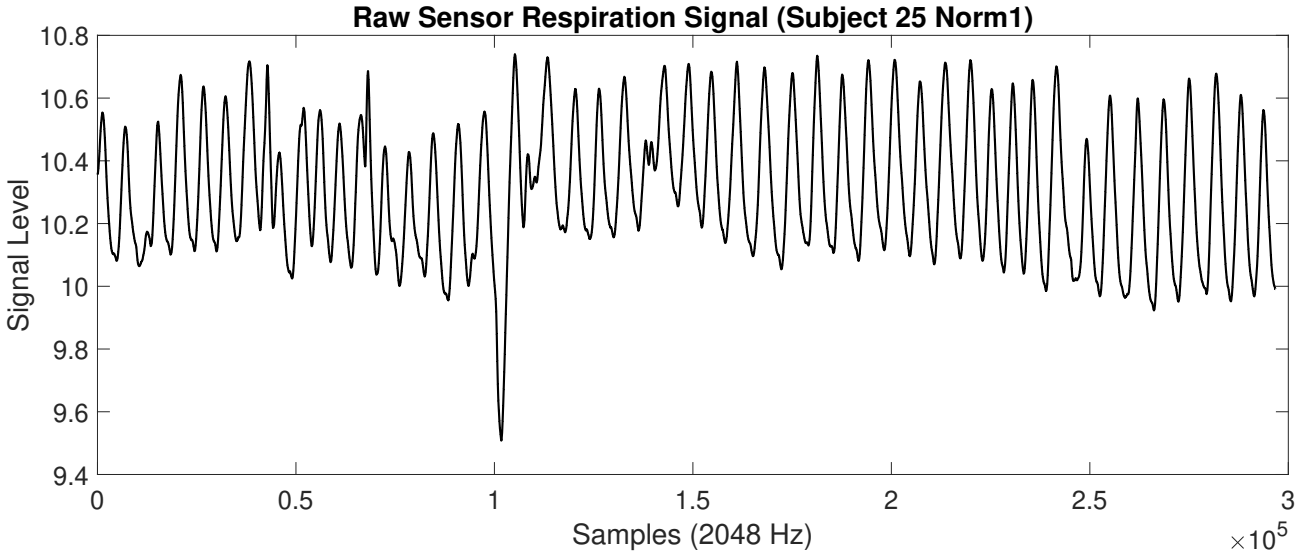

Figure 7.1: Subject 25 Raw Sensor Signal

Schuirmann's Two One Sided Tests (TOST) Procedure was used to assess equivalence between the ground truth and thermal respiration rates [56]. The TOST procedure is based on the following assumptions:

- The data is normally distributed
- The within-subject variances of the test and reference populations are the same (not a critical assumption)
- The study is a balanced crossover study; there is an equal number of subjects in each sample and there are no missing observations from any subject
- The equivalence interval is symmetric about zero (not a critical assumption)

Our data meets these assumptions since the respiration rates of humans are expected to be nor-

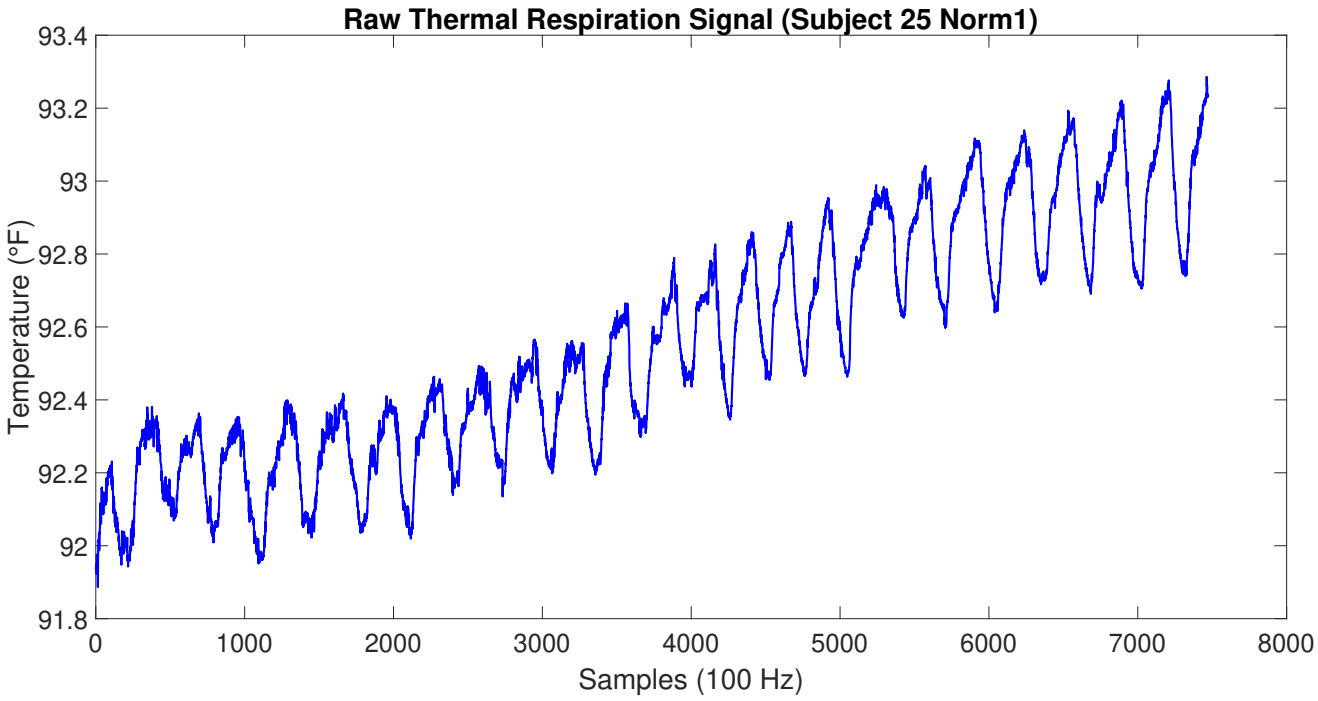

Figure 7.2: Subject 25 Raw Thermal Signal

mally distributed according to medical studies in the literature. Our results show that the within-subject variances for the sensor and thermal domains are the same within one standard deviation, though this is not a critical assumption. Our study is balanced since we extracted a ground truth signal and thermal signal from every subject in each video. For the sake of simplicity we chose a symmetric interval of ± 1 breath per minute.

The interval $[\theta_1, \theta_2]$ is the "interval hypothesis." Let μ_T be the mean of the test population and μ_R be the mean of the reference population. The null hypothesis, H_0 states that μ_T and μ_R are not equivalent. The alternative hypothesis, H_1 states that they are equivalent.

$$H_0: \quad \mu_T - \mu_R \leq \theta_1 \quad \text{or} \quad \mu_T - \mu_R \geq \theta_2$$

$$H_1: \quad \theta_1 < \mu_T - \mu_R < \theta_2$$

The two sets of one sided hypothesis can be tested using ordinary one-sided t-tests. Equivalency can be concluded if and only if both conditions of the null hypothesis are rejected at a chosen level of significance α. s is the estimated within-subject standard deviation defined as

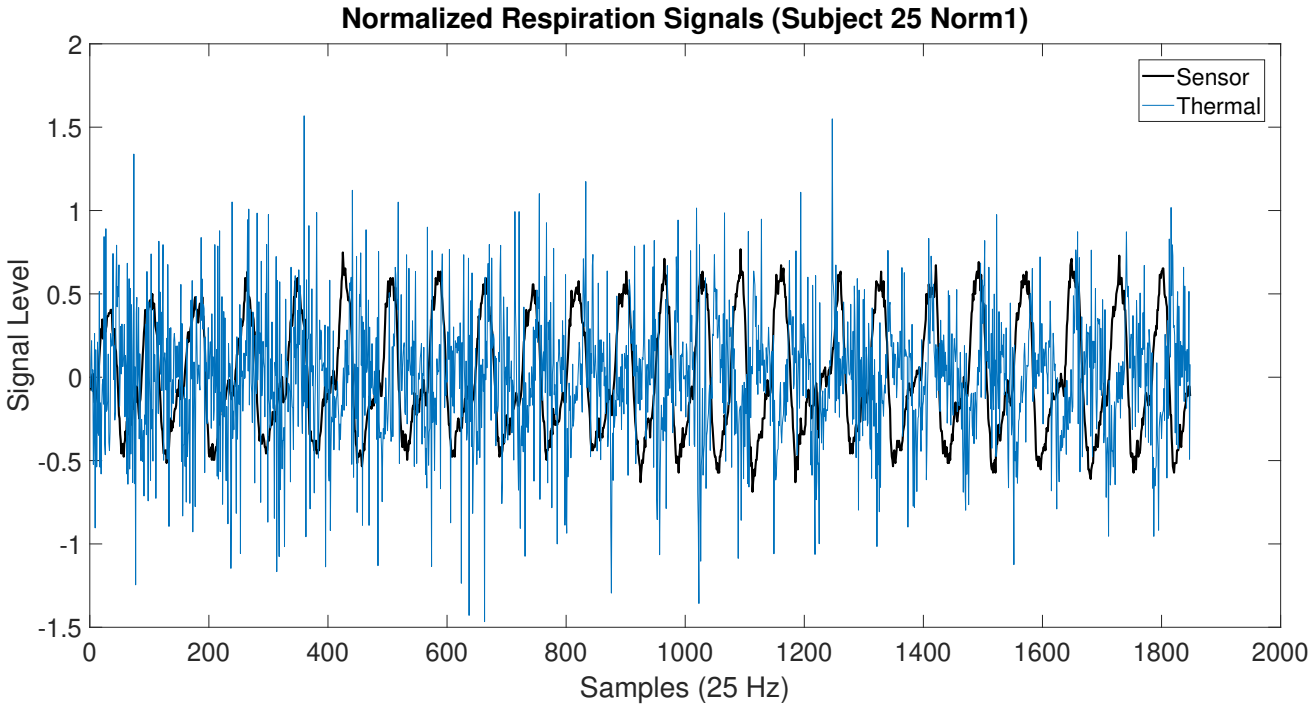

Figure 7.3: Subject 25 Normalized Signals

the square root of the mean squared error. The quantity $s\sqrt{2/n}$ is the standard error of $\bar{X_T} - \bar{X_R}$ based on *v* degrees of freedom.

$$t_1 = \frac{(\bar{X_T} - \bar{X_R}) - \theta_1}{s\sqrt{2/n}} \geq t_{1-\alpha(v)} \quad \text{and} \quad t_2 = \frac{\theta_2 - (\bar{X_T} - \bar{X_R})}{s\sqrt{2/n}} \geq t_{1-\alpha(v)} \tag{7.1}$$

We defined our hypothesis interval to be [-1, 1] bpm. Using this interval, we were able to conclude that the average Norm ground truth and thermal respiration rates are equal within ± 1 bpm and a 99% confidence interval. Furthermore, the Talking ground truth and thermal respiration rates are equal within a ± 1 bpm and an 80% confidence interval. A comparison of Figures 7.9 and 7.10 show that the confidence interval for the Talking videos is noticeably smaller than that of the Norm videos. This makes sense considering the higher variation in the thermal signal as well as the higher difference between the sample means. The higher number of talking videos also results in 371 degrees of freedom compared to only 157 for the Norm videos. This will effectively exaggerate the differences between the two populations thus reducing the confidence interval.

The results of the two one sided T-test are listed below. *t* is the point that isolates probability α in the upper tail of the Student's t distribution with *df* degrees of freedom and *p* is

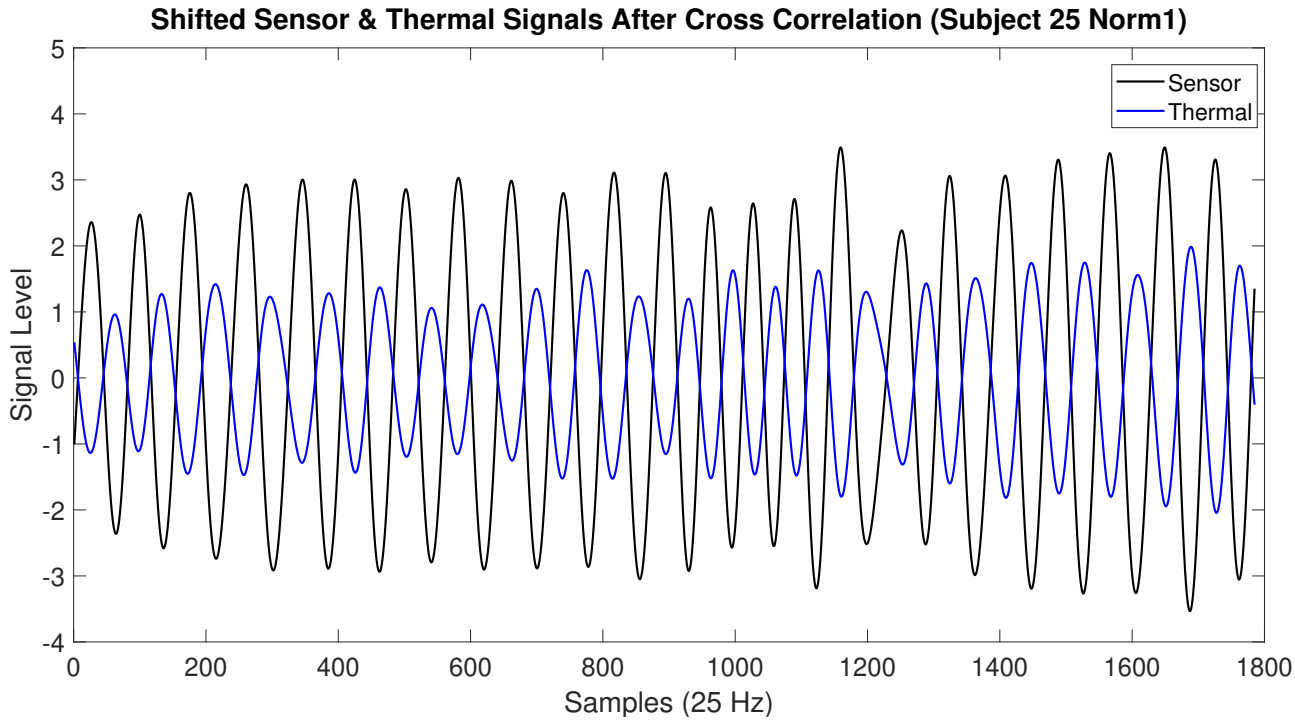

Figure 7.4: Subject 25 Transformed Signals (Correlation = 0.98)

the associated p-value. The confidence intervals are listed in the Equivalence Bounds table, in which Low/High show the user defined interval and Lower/Higher show the TOST confidence interval centered around the difference of the sample means. *Cohen's d* is the standardized difference between the means while Raw is simply the mean difference. The results strongly suggest that a relationship exists between the sensor and thermal domains and that this relation is not simply due to chance. The test also reveals that the rates extracted from the normal videos are more consistent than those extracted from the talking videos. This supports our assertion that our respiration rate estimation is more accurate when subjects are breathing through their noses without speaking.

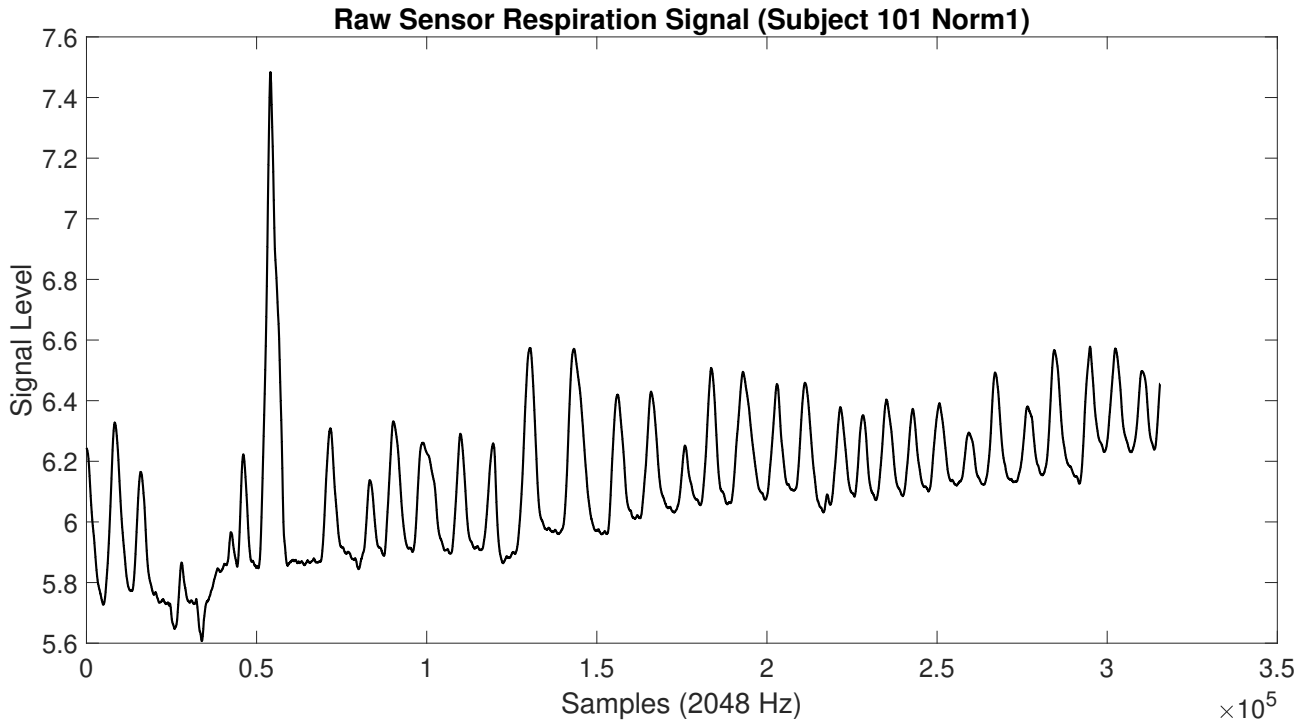

Figure 7.5: Subject 101 Raw Sensor Signal

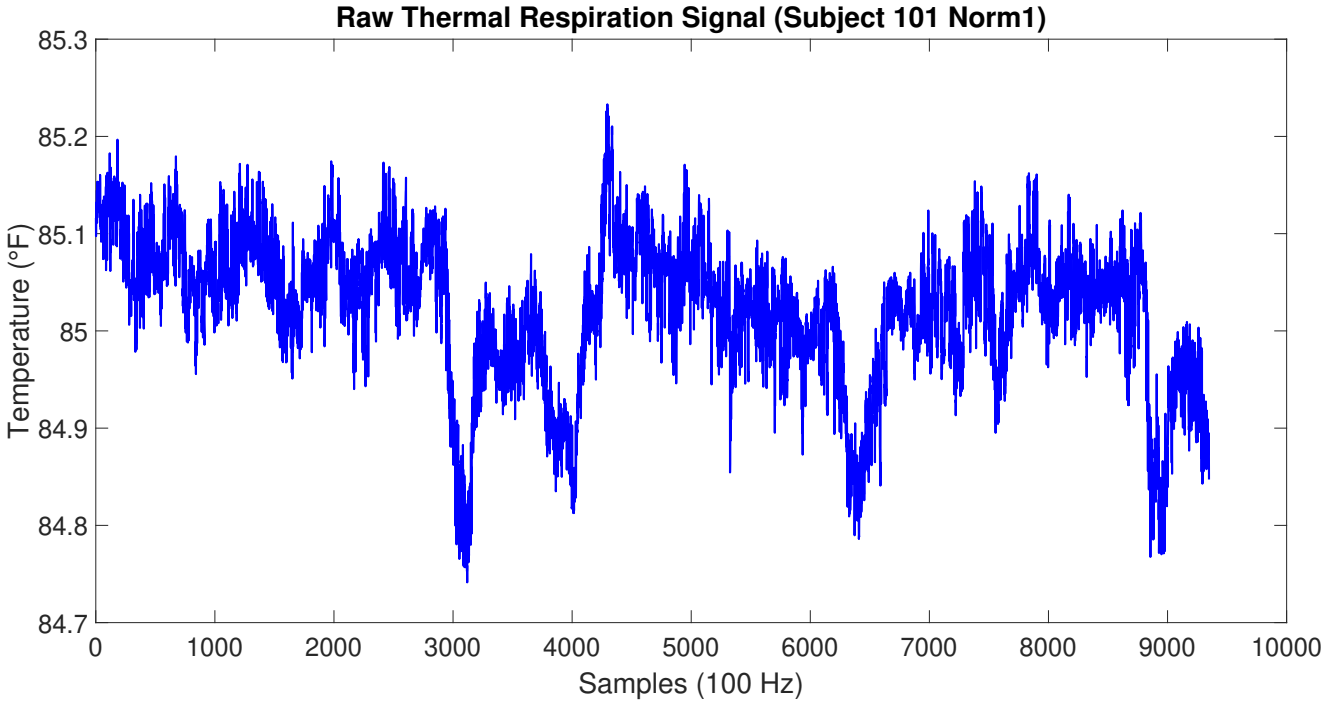

Figure 7.6: Subject 101 Raw Thermal Signal

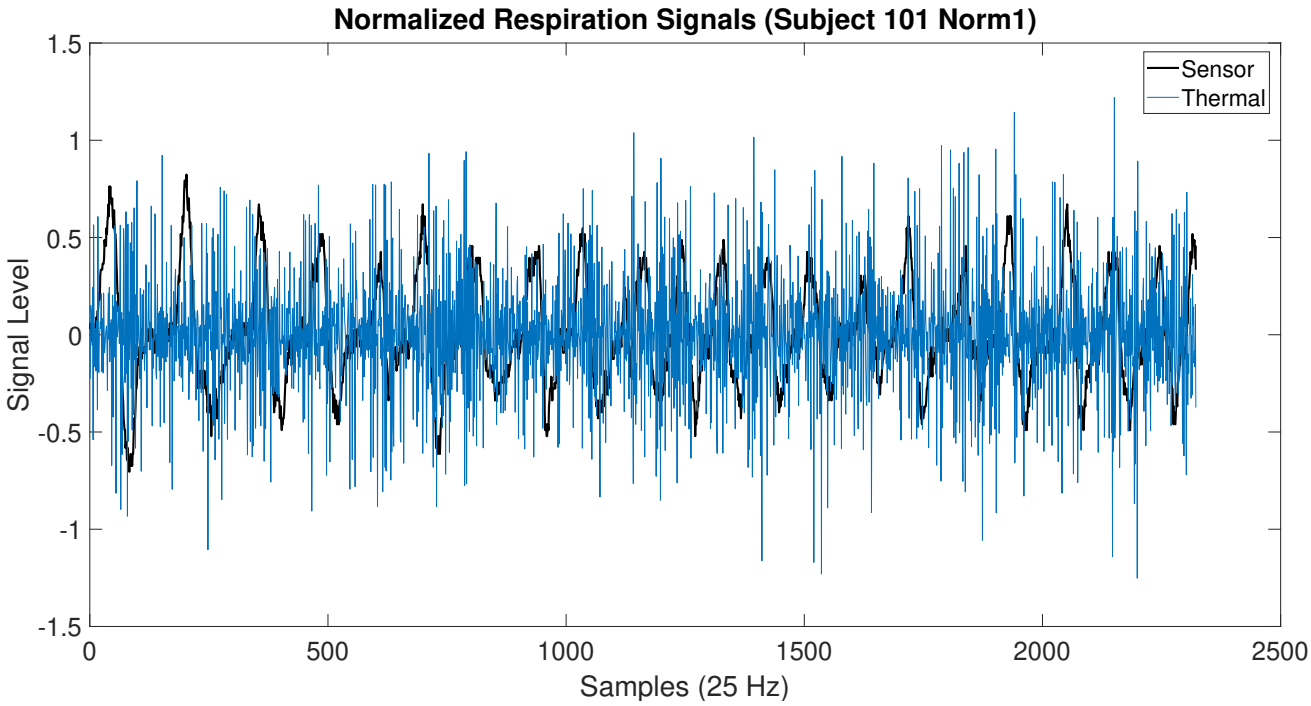

Figure 7.7: Subject 101 Normalized Signals

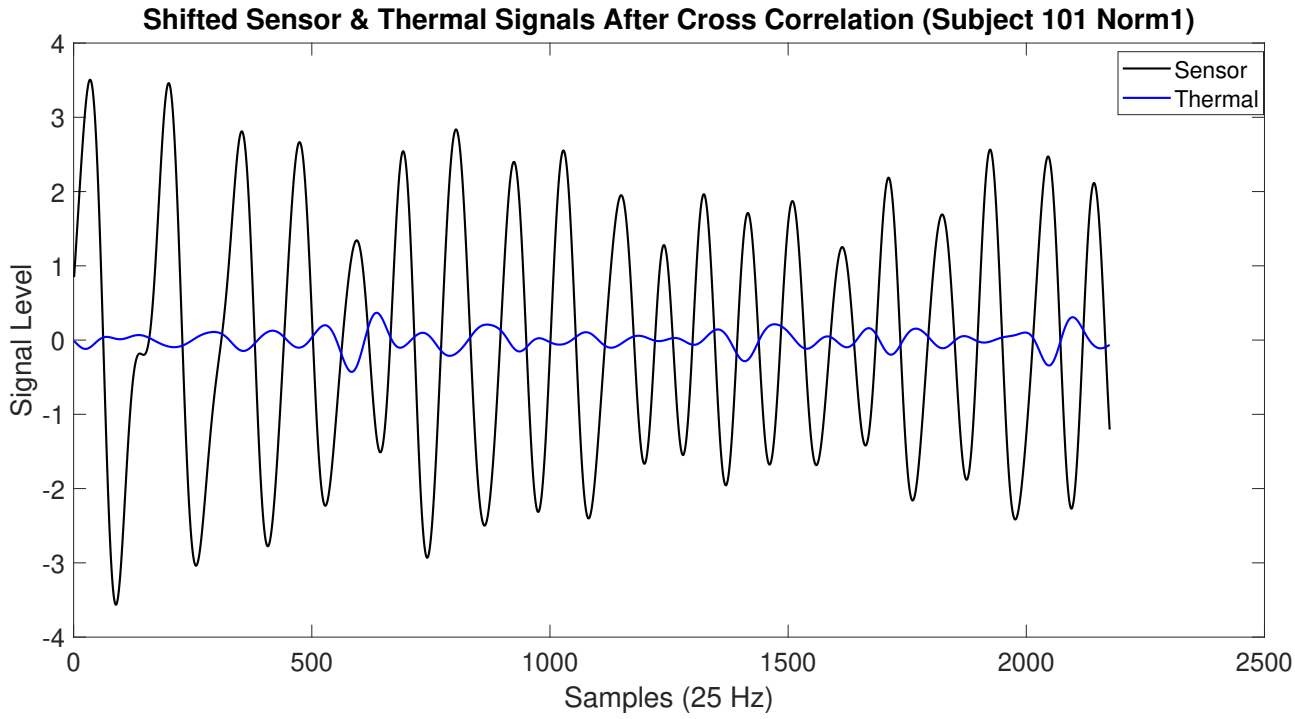

Figure 7.8: Subject 101 Transformed Signals (Correlation = 0.60)

NORM TOST PAIRED SAMPLES T-TEST Results

			t	df	p
sensor	thermal	t-test	0.665	157	0.507
		TOST Upper	-2.91	157	0.002
		TOST Lower	4.24	157	< .001

Equivalence Bounds

			Low	High	Lower	Upper
sensor	thermal	Cohen's d	-0.285	0.285		
		Raw	-1.00	1.00	-0.471	0.842

Descriptives

	N	Mean	Median	SD	SE
sensor	158	17.3	17.0	2.85	0.227
thermal	158	17.1	17.0	3.84	0.305

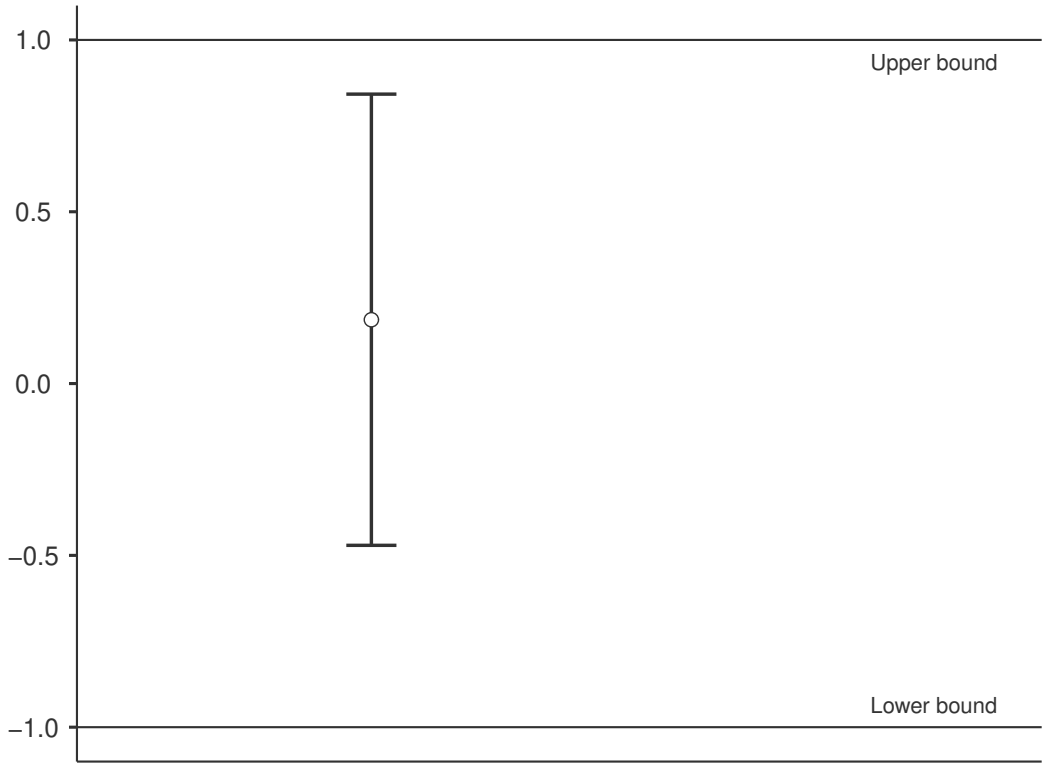

Figure 7.9: NORM Two One Sided T-Test: Confidence Interval

TALKING TOST PAIRED SAMPLES T-TEST Results

			t	df	p
sensor	thermal	t-test	3.35	371	< .001
		TOST Upper	-0.828	371	0.204
		TOST Lower	7.53	371	< .001

Equivalence Bounds

			Low	High	Lower	Upper
sensor	thermal	Cohen's d	-0.217	0.217		
		Raw	-1.00	1.00	0.600	1.00

Descriptives

	N	Mean	Median	SD	SE
sensor	372	16.9	16.3	3.29	0.171
thermal	372	16.1	16.2	3.44	0.178

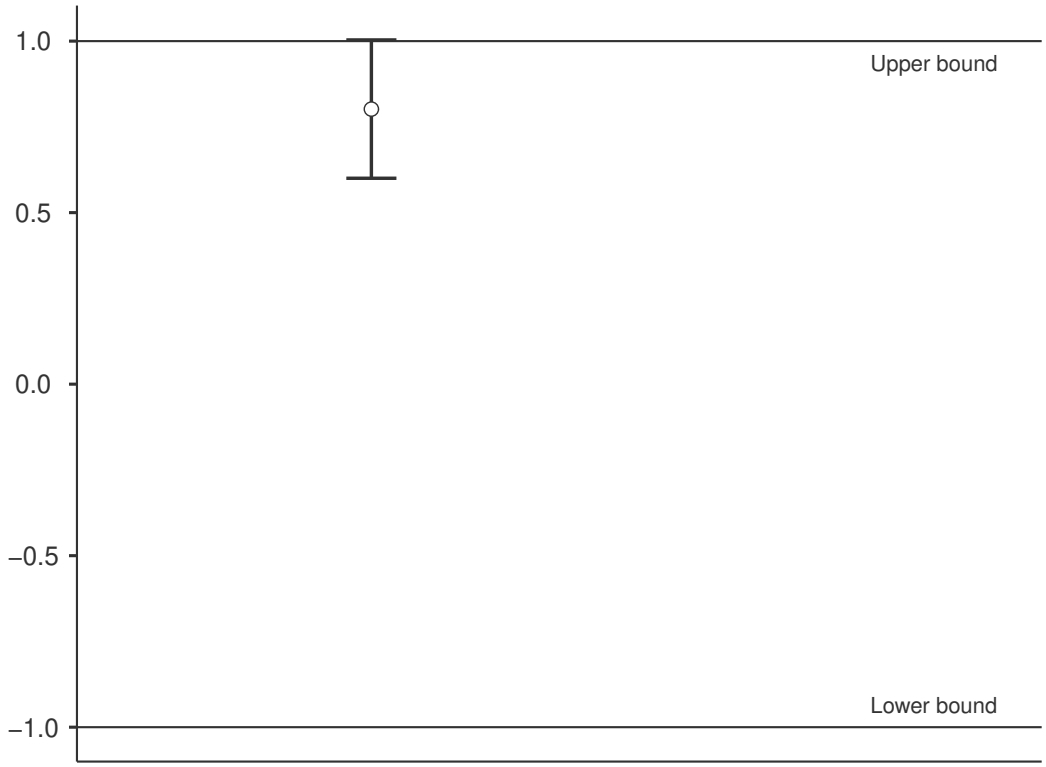

Figure 7.10: TALKING Two One Sided T-Test: Confidence Interval

7.2 Skin Temperature

We applied multiple linear regression to the entire dataset consisting of 104 subjects and a total of 728 videos. We calculated the average temperature within the forehead region of interest to measure the skin temperature. This measurement is less prone to tracking errors since we can still measure the average temperature of the ROI even if the bounding box is partially obscured. For that reason, we were able to extract the skin temperature from all videos without excluding any of them. The average cross correlation between the regression line and the sensor signal is -0.17, which is statistically insignificant. However, the average absolute cross correlation is +0.81. This implies that some relationship exists between the thermal signal extracted from the forehead and the sensor signal measured from the finger. Figures 7.11, 7.12, 7.13 and 7.14 depict examples of strongly correlated skin temperature signals. The fact that the correlation is strong in both positive and negative directions suggests that the rate of heat exchange varies between both subjects and body parts. Related Work previously mentioned in the skin temperature section support this finding. In particular, [41]'s results "show for trained

subjects a better capability of losing heat, thus of shifting through the blood the heat from the muscle to the skin."

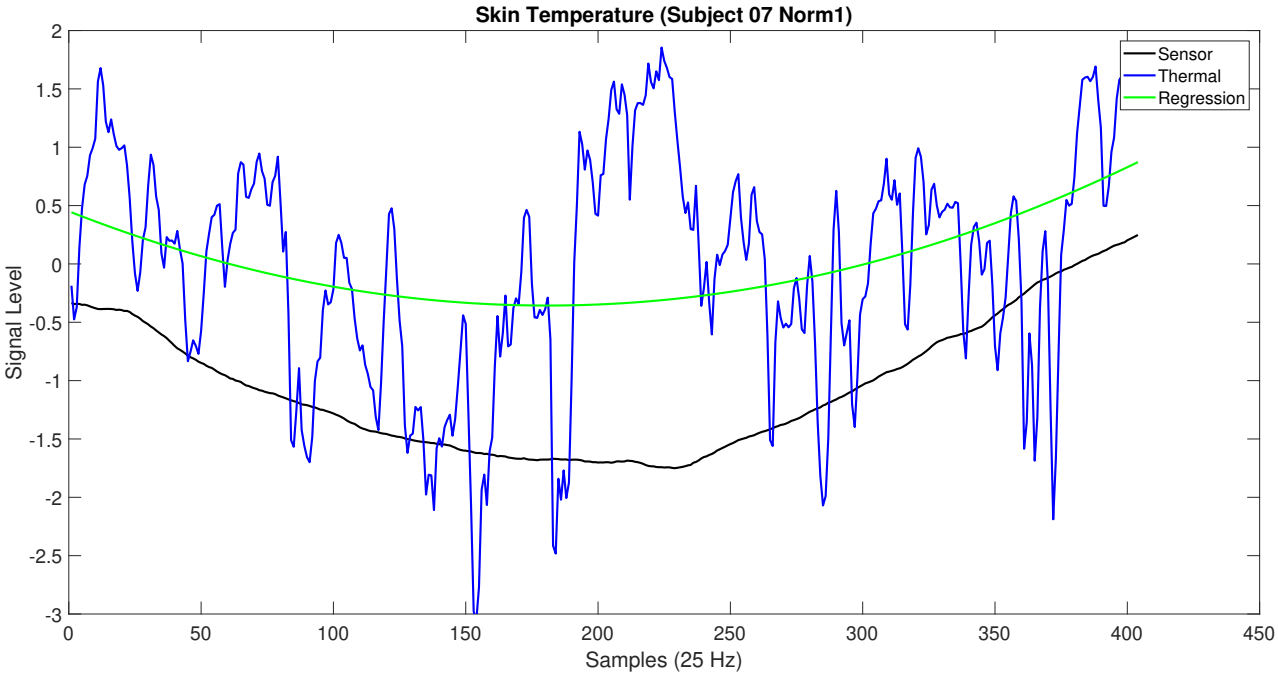

Figure 7.11: Example 1: Strong Positive Correlation

Secondly, we wish to quantify the relationship between the skin temperature of the forehead and the finger, which were collected from the thermal camera and contact based sensor respectively. We use equation 7.2 to derive a factor to translate between forehead and finger temperatures in degrees Fahrenheit. For each video *i*, μ_T is the mean of the thermal signal, μ_S is the mean of the sensor signal and n is the number of videos. Equation 7.2 is simply the average of the ratio of the signal means.

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\mu_{Si}}{\mu_{Ti}} \tag{7.2}$$

The result of equation 7.2 is 0.895. Therefore, equation 7.3 can be used to calculate the sensor based temperature from the thermal temperature. S_S and S_T are the sensor and thermal temperatures respectively.

$$S_S = 0.895 S_T \tag{7.3}$$

The margin of error was calculated using equation 7.4, in which t is the inverse t-distribution, α is the confidence interval, n is the sample size, and σ_{ratio} is the standard deviation of the ratio

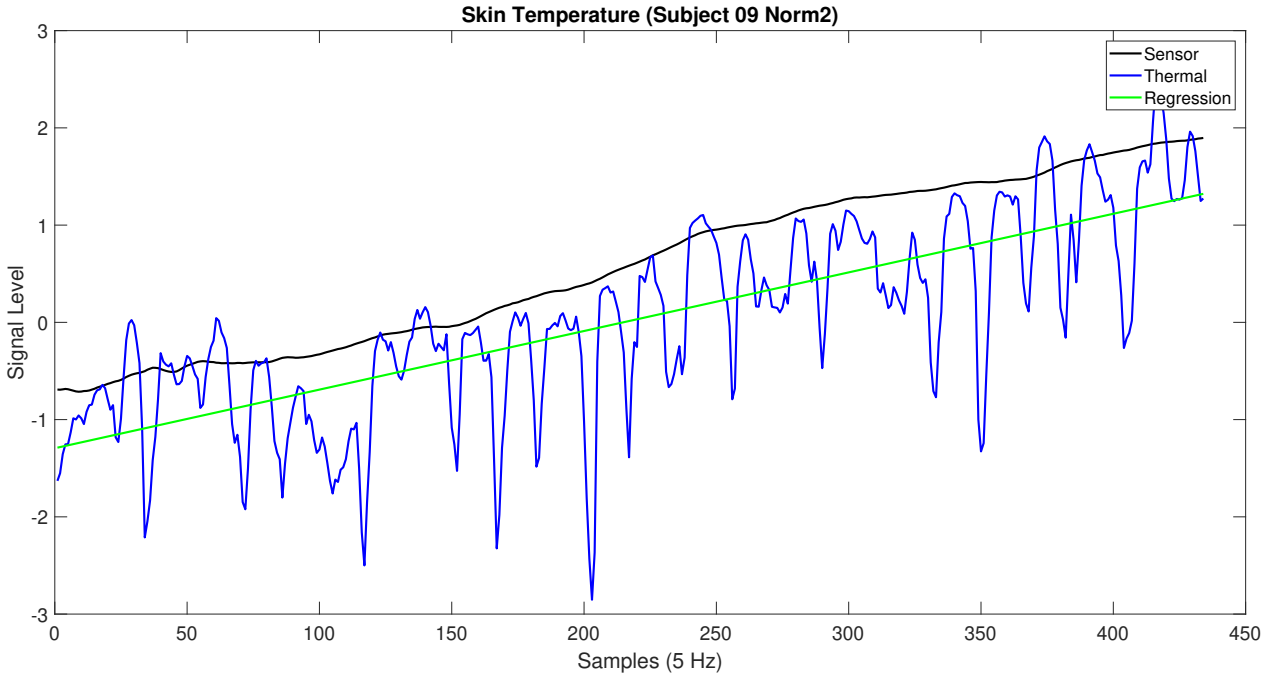

Figure 7.12: Example 2: Strong Positive Correlation

of the signal means. Given $\alpha = 95$, $\sigma_{\text{ratio}} = 0.07$ and $n = 728$, the margin of error for a 95% confidence interval is only 0.5%. This suggests that the amount of sampling error is negligible due to our rather large sample size. Therefore, the average ratio between forehead and finger skin temperatures is likely valid for a larger population and not merely a result of chance.

$$t(1 - \alpha, n - 1) \times \frac{\sigma_{\text{ratio}}}{\sqrt{n}} \tag{7.4}$$

For the sake of comparison we also applied a linear regression to evaluate equation 7.3 as well as to visualize the temperature samples collected from both the sensor and thermal domains. In this regression model, the predictor variable is defined by the mean thermal temperature values and the response variable is defined by the mean sensor values. The least squares fit is given by equation 7.5. S_S and S_T are the sensor and thermal temperatures respectively.

$$S_S = 0.007 \times S_T + 83.734 \tag{7.5}$$

Figure 7.15 reveals that the temperature samples do not appear to follow a linear trend. On the other hand, it shows that the samples are clustered around a point with sensor mean equal

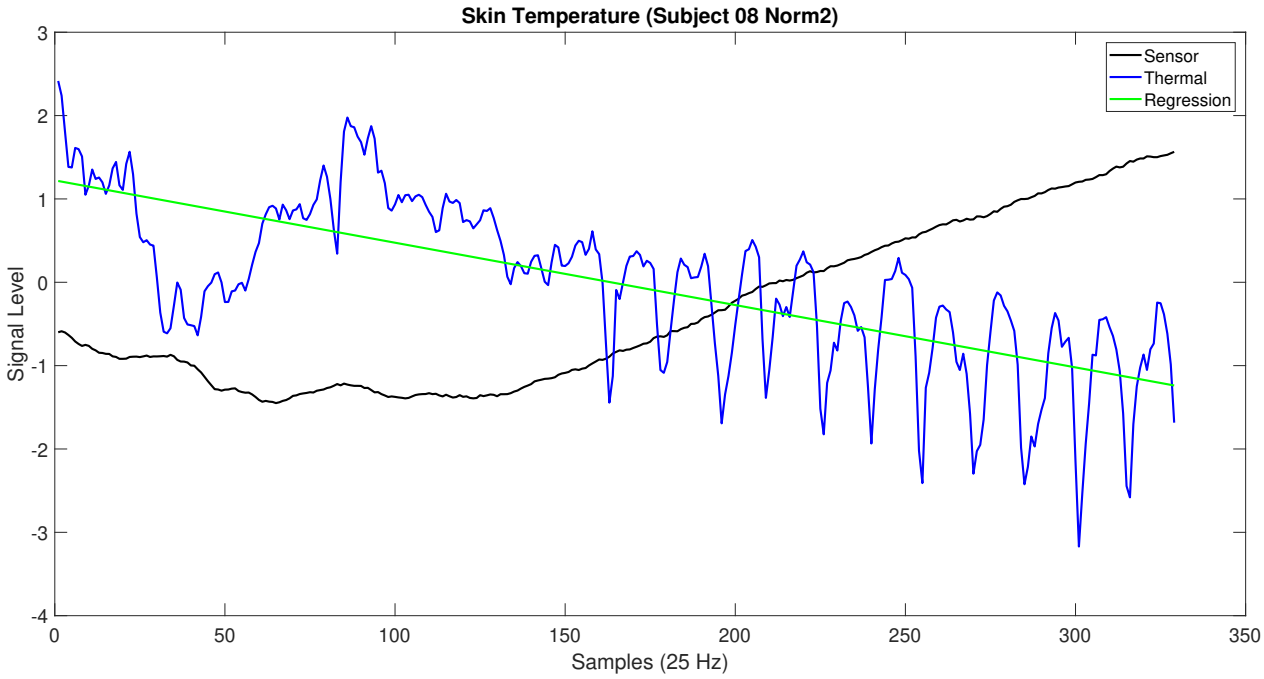

Figure 7.13: Example 3: Strong Negative Correlation

to 84.4°F and thermal mean equal to 94.3°F. This observation also supports the validity of equation 7.2. However, equation 7.2 is based on the skin temperature of subjects who were sitting comfortably at room temperature. An additional study would be required to determine if a similar relationship holds for subjects with skin temperatures outside of the ranges observed in our study.

7.3 Heart Rate

We analyzed a total of 727 videos of 104 subjects. The inner corners of the eyes were used as the region of interest for extracting the thermal heart rate signal. In this section we no longer make a distinction between the 'Normal' and 'Talking' videos as we did in section 7.1 for calculating respiration rates. The reason for this is that measurements extracted from the eye region should not be affected by talking in the same way the nose region was. Table 7.4 shows that the sensor rates are more spread out compared to the thermal heart rates. The small standard deviation of the thermal rates is likely a result of the narrow pass-band we selected when computing the inverse continuous wavelet transform. Widening the pass-band led to results which were less consistent with the ground truth sensor rates. Therefore, we experienced

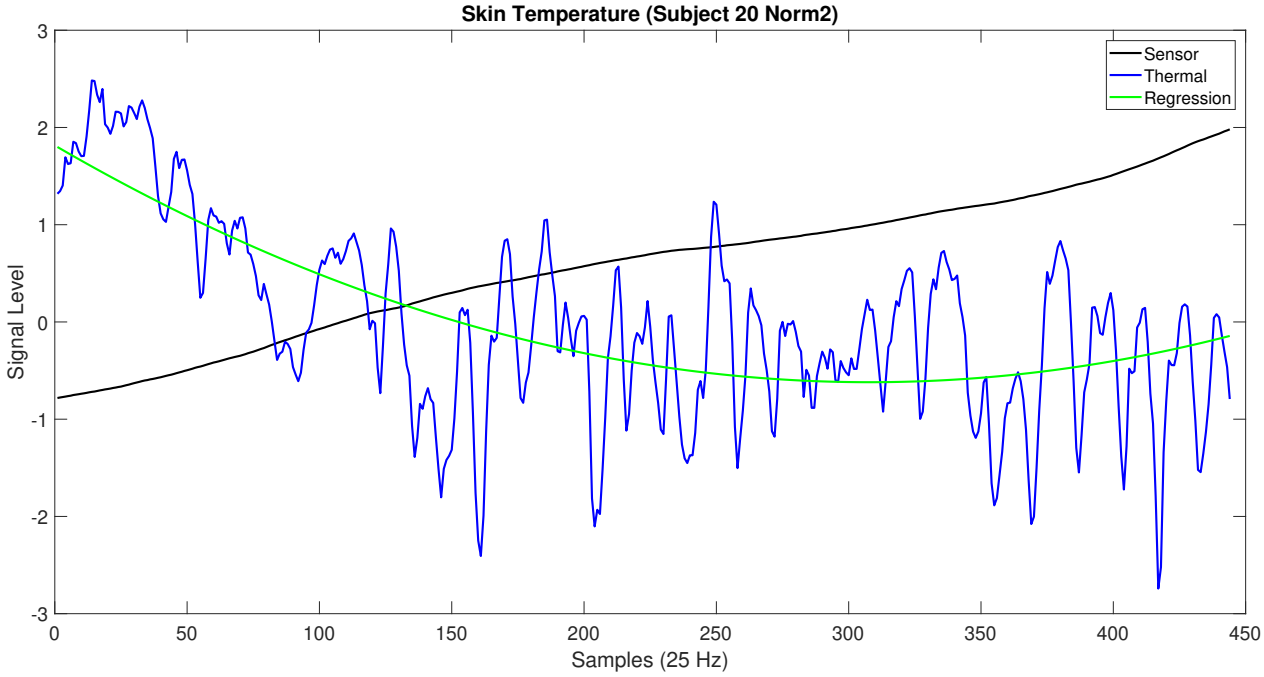

Figure 7.14: Example 4: Strong Negative Correlation

a trade-off between the scope of the frequency pass-band and heart rate accuracy. Normal resting heart rates are generally considered to be in the range of 60-100 beats per minute. For that reason, we consider heart rates above 100bpm to be outliers. Removing outliers reduced the standard deviation of the sensor rates from 12bpm to 10bpm.

Table 7.5 also shows a decrease in the average difference between the sensor rate and the thermal rate from 11bpm to 9bpm. A future study could explore methods to successfully detect heart rates outside of the normal range.

Domain	Min	Max	Average	St. Dev.
Sensor	54	128	84	12
Thermal	67	88	78	3

Table 7.4: Heart Rate Statistical Measures in beats per minute

Outlier Removal	Difference (bpm)
Before	11
After	9

Table 7.5: Average Heart Rate Difference Between Sensor & Thermal Rates Before & After Removing Outliers

Figures 7.16, 7.17, 7.18 and 7.19 display examples of the sensor and thermal heart rate

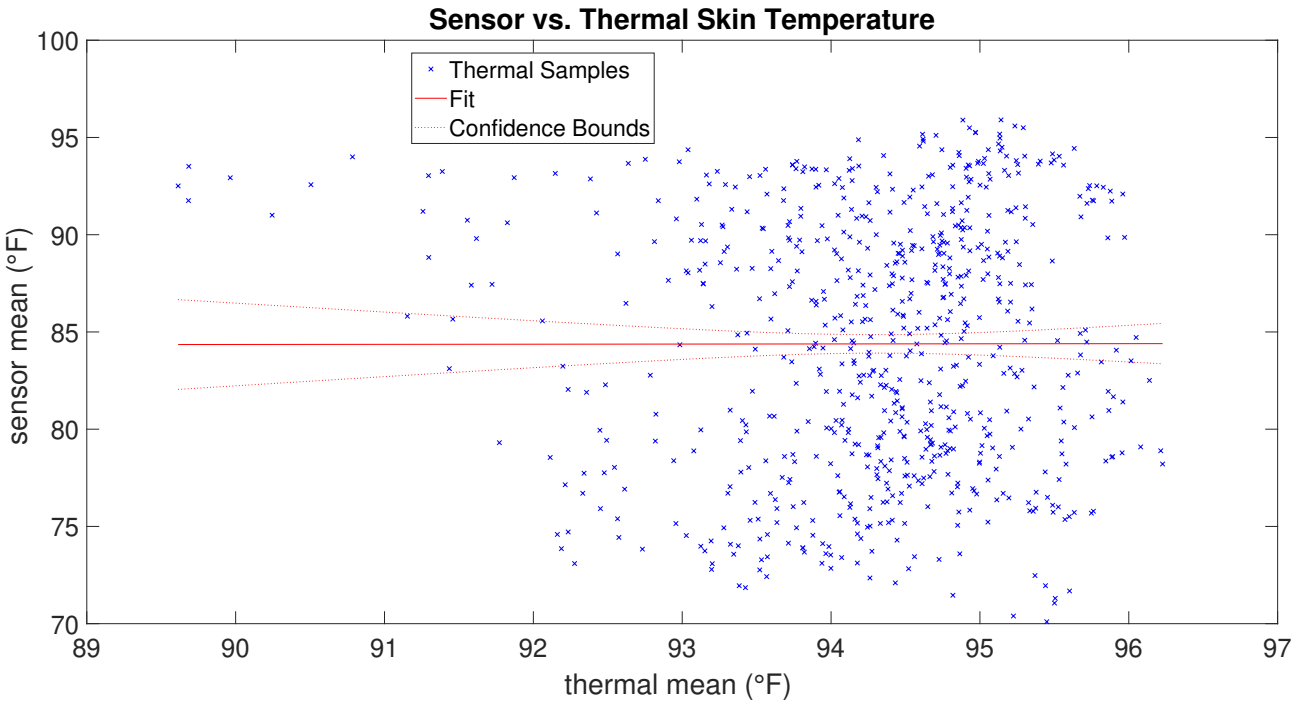

Figure 7.15: Skin Temperature Linear Regression Plot

signals during various stages of processing.

Finally, we perform a two one sided T-test on the heart rates to test the equivalency between the sensor and thermal rates. 657 samples out of the total 727 were used after eliminating outliers. The test reveals that the two sets are equivalent in the interval [-5,+5] for a 99% confidence interval. Figure 7.20 shows that the TOST interval is within the user defined interval [-5,5]. These results suggest that the true mean difference between the heart rates is less than 5 beats per minute. Assuming a normal distribution, this statistic should remain true for larger sample sizes and/or different samples from the same population, in this case university students.

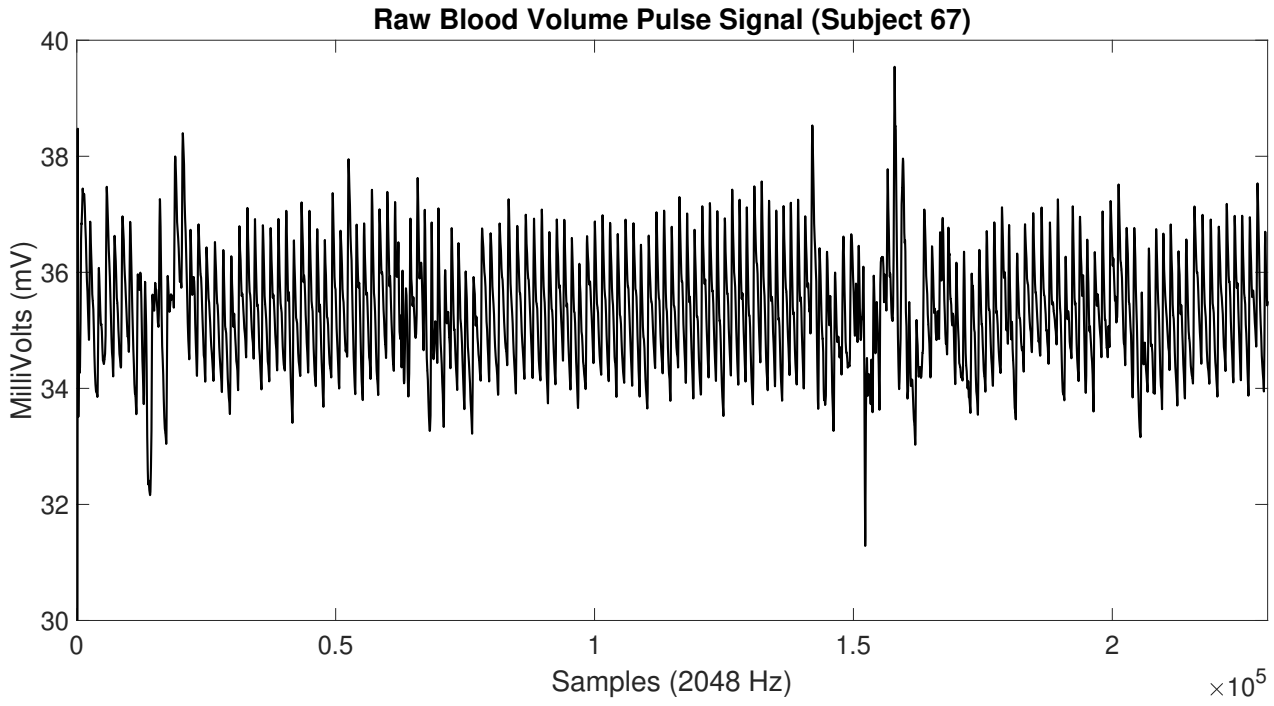

Figure 7.16: Subject 67 Raw Sensor Heart Rate Signal

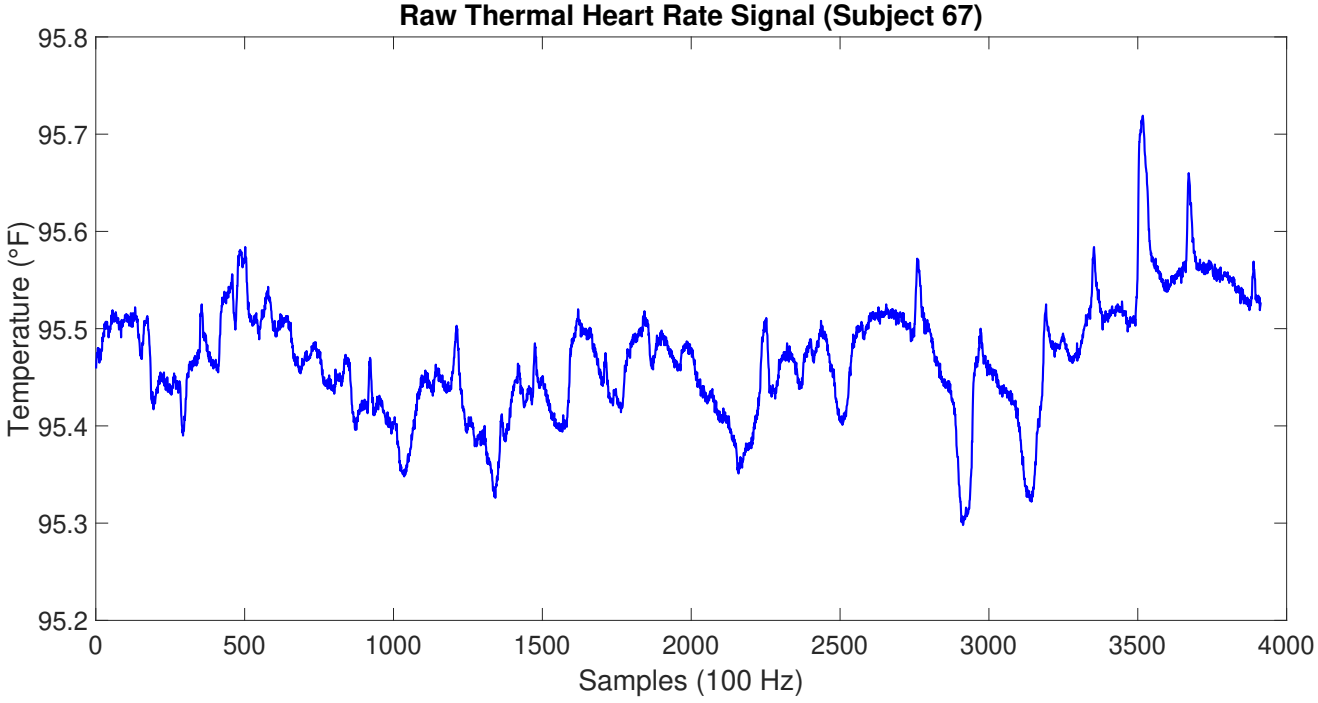

Figure 7.17: Subject 67 Raw Thermal Heart Rate Signal

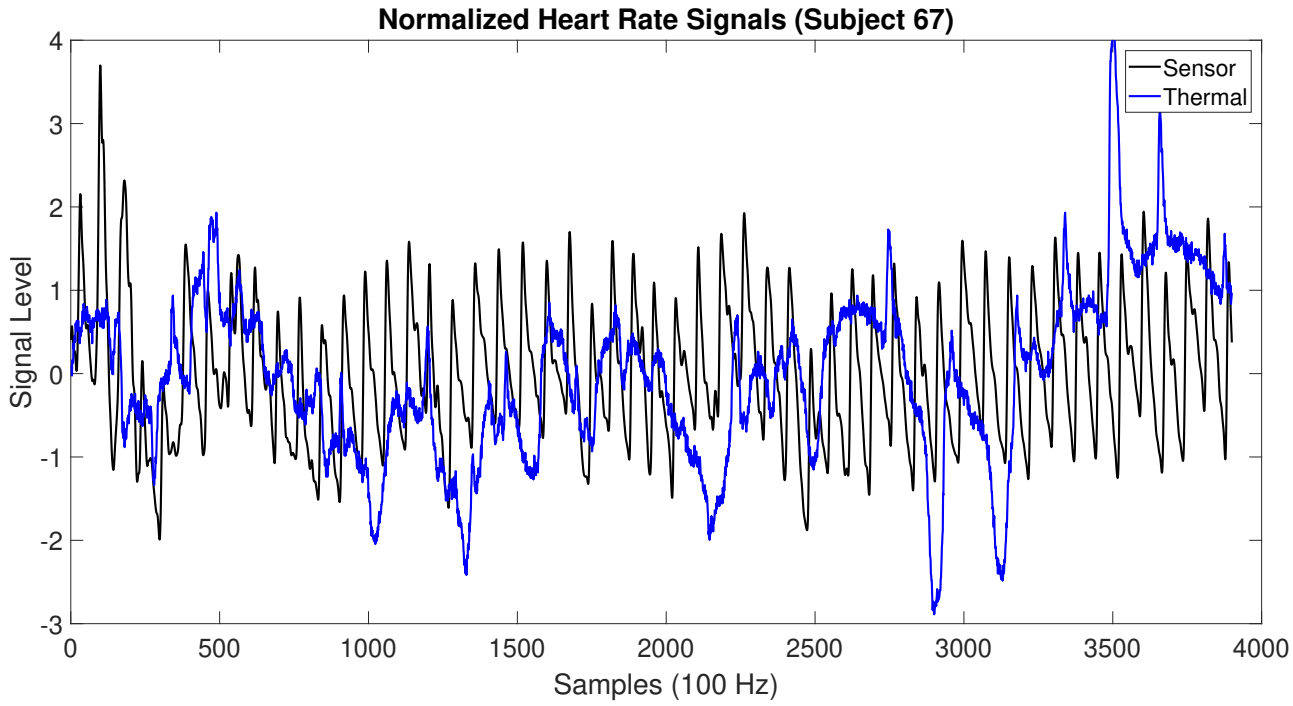

Figure 7.18: Subject 67 Normalized Heart Rate Signals

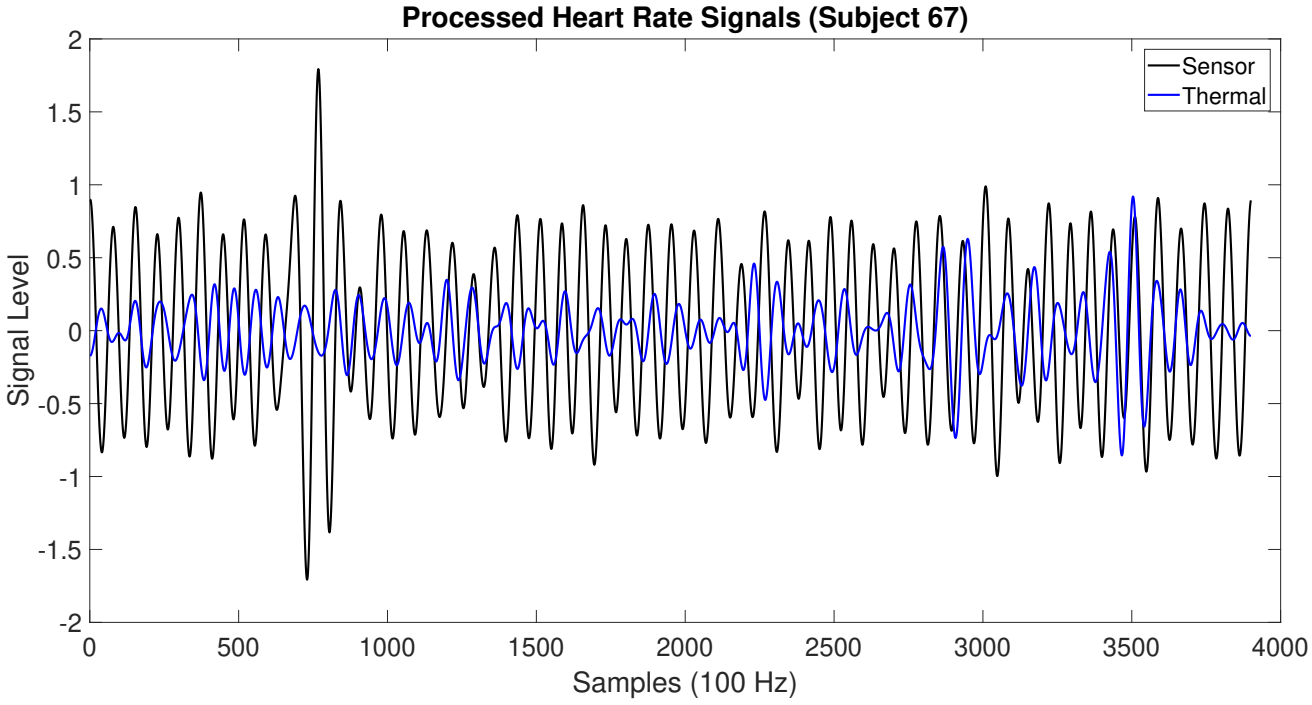

Figure 7.19: Subject 67 Processed Heart Rate Signals

HEART RATE TOST PAIRED SAMPLES T-TEST

TOST Results

			t	df	p
sensor	thermal	t-test	8.21	656	< .001
		TOST Upper	-3.91	656	< .001
		TOST Lower	20.3	656	< .001

Equivalence Bounds

			Low	High	Lower	Upper
sensor	thermal	Cohen's d	-0.473	0.473		
		Raw	-5.00	5.00	2.42	4.35

Descriptives

	N	Mean	Median	SD	SE
sensor	657	81.6	81.4	10.3	0.403
thermal	657	78.2	78.3	2.61	0.102

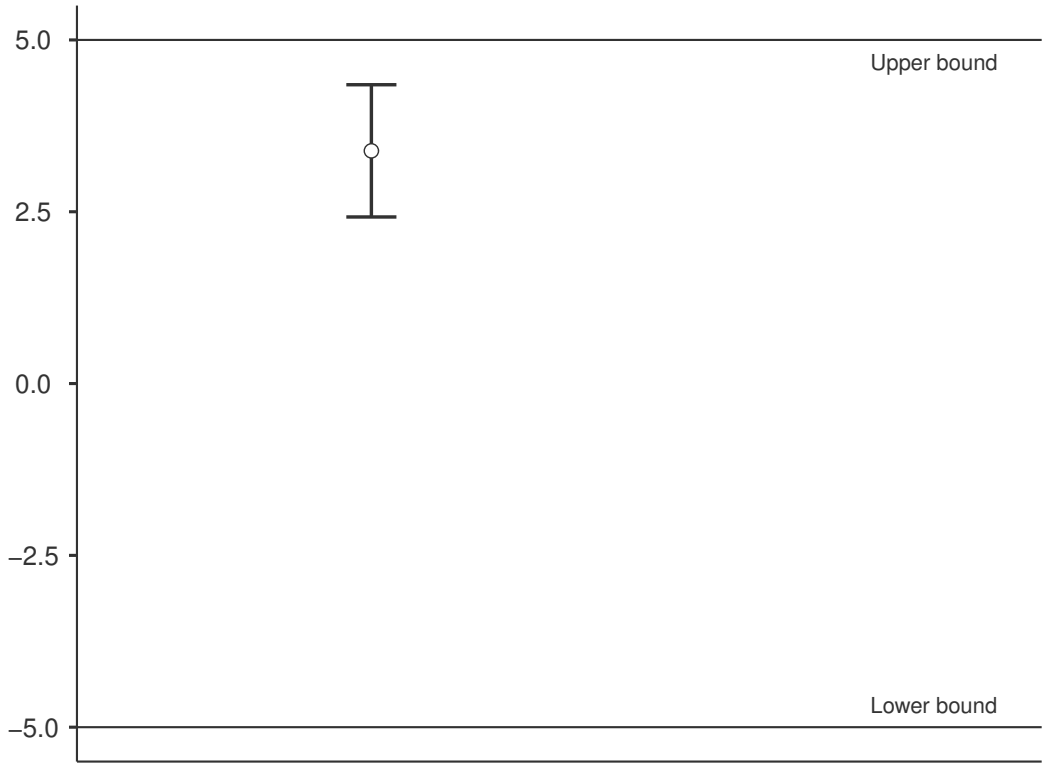

Figure 7.20: Heart Rate Two One Sided T-Test: Confidence Interval

Chapter 8

Conclusion

There are numerous benefits associated with the use of thermal imaging for health monitoring and modeling of human behavior. Recording thermal video is far more convenient compared to attaching multiple sensors to the body. Contact-based sensors and electrodes attached to the body may cause discomfort, which can alter the subject's physiological state. Thermal cameras, on the the other hand, are unobtrusive and therefore less likely to influence the subject or introduce bias. Moreover, thermal cameras can potentially screen people in seconds compared to the time-consuming task of outfitting someone with an array of sensors. Modern technology is facilitating the development of real-time thermal imaging systems that are sensitive to minute variations in skin temperature.

We have developed a solution which integrates image and signal processing techniques to extract various physiological signals from thermal images. Specifically, we proposed different methods for capturing temporal features in the face and for filtering thermal signals. Our results show a significant relation between the respiration rates extracted from thermal imaging and those recorded by contact-based sensors. Mean respiration rates extracted from the thermal domain were within a range of one breath per minute from the mean of the sensor-based measurements. We were also able to establish an average correlation of 0.6 between the thermal and sensor breathing signals. Secondly, we found that heat exchange rate varies between the finger and the forehead. We also derived an equation to estimate the skin temperature of the finger based on the thermal skin temperature signal extracted from the forehead. Finally, we extracted heart rates from the thermal domain. The mean difference between the sensor and thermal heart rates was five beats per minute. This research has shown that non-contact approaches

using thermal imaging represent a reliable alternative to using contact-based sensors.

Bibliography

[1] Abouelenien, M., Burzo, M., and Mihalcea, R. Cascaded multimodal analysis of alertness related features for drivers safety applications. In *Proceedings of the 8th ACM International Conference on PErvasive Technologies Related to Assistive Environments* (New York, NY, USA, 2015), PETRA '15, ACM, pp. 59:1–59:8.

[2] Abouelenien, M., Burzo, M., and Mihalcea, R. Human acute stress detection via integration of physiological signals and thermal imaging. In *Proceedings of the 9th ACM International Conference on PErvasive Technologies Related to Assistive Environments* (2016), ACM, p. 32.

[3] Abouelenien, M., Mihalcea, R., and Burzo, M. Analyzing thermal and visual clues of deception for a non-contact deception detection approach. In *Proceedings of the 9th ACM International Conference on PErvasive Technologies Related to Assistive Environments* (2016), ACM, p. 35.

[4] Abouelenien, M., Pérez-Rosas, V., Mihalcea, R., and Burzo, M. Deception detection using a multimodal approach. In *Proceedings of the 16th International Conference on Multimodal Interaction* (New York, NY, USA, 2014), ICMI '14, ACM, pp. 58–65.

[5] Abouelenien, M., Pérez-Rosas, V., Mihalcea, R., and Burzo, M. Multimodal gender detection. In *Proceedings of the 19th ACM International Conference on Multimodal Interaction* (New York, NY, USA, 2017), ICMI 2017, ACM, pp. 302–311.

[6] Abouelenien, M., Pérez-Rosas, V., Zhao, B., Mihalcea, R., and Burzo, M. Gender-based multimodal deception detection. In *Proceedings of the Symposium on Applied Computing* (New York, NY, USA, 2017), SAC '17, ACM, pp. 137–144.

[7] Abouelenien, M., Pérez-Rosas, V., Zhao, B., Mihalcea, R., and Burzo, M. Gender-based multimodal deception detection. In *Proceedings of the Symposium on Applied Computing* (2017), ACM, pp. 137–144.

[8] Abouelenien, M., Pérez-Rosas, V., Mihalcea, R., and Burzo, M. Detecting deceptive behavior via integration of discriminative features from multiple modalities. *IEEE Transactions on Information Forensics and Security 12*, 5 (May 2017), 1042–1055.

[9] Alkali, A. H., Saatchi, R., Elphick, H., and Burke, D. Facial tracking in thermal images for real-time noncontact respiration rate monitoring. In *Modelling Symposium (EMS), 2013 European* (2013), IEEE, pp. 265–270.

[10] Bennett, S. L., Goubran, R., and Knoefel, F. Adaptive eulerian video magnification methods to extract heart rate from thermal video. In *Medical Measurements and Applications (MeMeA), 2016 IEEE International Symposium on* (2016), IEEE, pp. 1–5.

[11] Bennett, S. L., Goubran, R., and Knoefel, F. Comparison of motion-based analysis to thermal-based analysis of thermal video in the extraction of respiration patterns. In *Engineering in Medicine and Biology Society (EMBC), 2017 39th Annual International Conference of the IEEE* (2017), IEEE, pp. 3835–3839.

[12] Bryant, R. A., Creamer, M., O'Donnell, M., Silove, D., and McFarlane, A. C. A multisite study of initial respiration rate and heart rate as predictors of posttraumatic stress disorder. *The Journal of Clinical Psychiatry* (2008).

[13] Buddharaju, P., Pavlidis, I. T., Tsiamyrtzis, P., and Bazakos, M. Physiology-based face recognition in the thermal infrared spectrum. *IEEE transactions on pattern analysis and machine intelligence 29*, 4 (2007), 613–626.

[14] Burzo, M., Abouelenien, M., Perez-Rosas, V., Wicaksono, C., Tao, Y., and Mihalcea, R. Using infrared thermography and biosensors to detect thermal discomfort in a building's inhabitants. In *ASME 2014 International Mechanical Engineering Congress and Exposition* (2014), American Society of Mechanical Engineers, pp. V06BT07A015–V06BT07A015.

[15] Burzo, M., Abouelenien, M., Perez-Rosas, V., Wicaksono, C., Tao, Y., and Mihalcea, R. Using infrared thermography and biosensors to detect thermal discomfort in a building's inhabitants. In *ASME 2014 International Mechanical Engineering Congress and Exposition* (2014), American Society of Mechanical Engineers, pp. V06BT07A015–V06BT07A015.

[16] Charkoudian, N. Skin blood flow in adult human thermoregulation: how it works, when it does not, and why. In *Mayo Clinic Proceedings* (2003), vol. 78, no. 5, Elsevier, pp. 603–612.

[17] Chekmenev, S. Y., Farag, A. A., Miller, W. M., Essock, E. A., and Bhatnagar, A. Multiresolution approach for noncontact measurements of arterial pulse using thermal imaging. In *Augmented vision perception in infrared*. Springer, 2009, pp. 87–112.

[18] Cretikos, M. A., Bellomo, R., Hillman, K., Chen, J., Finfer, S., and Flabouris, A. Respiratory rate: the neglected vital sign. *Medical Journal of Australia 188*, 11 (2008), 657.

[19] Dawson, M., Schell, A., and Filion, D. The electrodermal system. *Handbook of psychophysiology 2* (2007), 200–223.

[20] Derksen, M. Control and resistance in the psychology of lying. *Theory and Psychology 22*, 2 (2012), 196–212.

[21] Fei, J., and Pavlidis, I. Thermistor at a distance: unobtrusive measurement of breathing. *IEEE Transactions on Biomedical Engineering 57*, 4 (2010), 988–998.

[22] Figner, B., Murphy, R., et al. Using skin conductance in judgment and decision making research. *A handbook of process tracing methods for decision research* (2011), 163–184.

[23] Gannon, T., Beech, A., and Ward, T. *Risk Assessment and the Polygraph*. Nova Science Publishers, 2009, pp. 129–154.

[24] Gao, T., Greenspan, D., Welsh, M., Juang, R. R., and Alm, A. Vital signs monitoring and patient tracking over a wireless network. In *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the* (2006), IEEE, pp. 102–105.

[25] Garbey, M., Sun, N., Merla, A., and Pavlidis, I. Contact-free measurement of cardiac pulse based on the analysis of thermal imagery. *IEEE Transactions on Biomedical Engineering 54*, 8 (2007), 1418–1426.

[26] Gault, T., and Farag, A. A fully automatic method to extract the heart rate from thermal video. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* (2013), pp. 336–341.

[27] Gault, T. R., Blumenthal, N., Farag, A. A., and Starr, T. Extraction of the superficial facial vasculature, vital signs waveforms and rates using thermal imaging. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on* (2010), IEEE, pp. 1–8.

[28] Gault, T. R., and Farag, A. A. Computationally light forehead segmentation from thermal images. In *Image Processing (ICIP), 2012 19th IEEE International Conference on* (2012), IEEE, pp. 169–172.

[29] Hinkle, D. E., Wiersma, W., and Jurs, S. G. Applied statistics for the behaviors science (4* ed.), 1998.

[30] Honts, C. R., and Kircher, J. C. Mental and physical countermeasures reduce the accuracy of polygraph tests. *Journal of Applied Psychology 79*, 2 (1994), 252.

[31] Huizenga, C., Zhang, H., Arens, E., and Wang, D. Skin and core temperature response to partial-and whole-body heating and cooling. *Journal of Thermal Biology 29*, 7-8 (2004), 549–558.

[32] Jones, B. F., and Plassmann, P. Digital infrared thermal imaging of human skin. *IEEE Engineering in Medicine and Biology Magazine 21*, 6 (2002), 41–48.

[33] Khan, M., Ward, R., and Ingleby, M. Distinguishing facial expressions by thermal imaging using facial thermal feature points. In *Proceedings of HCI* (2005), pp. 5–9.

[34] Khan, M., Ward, R., and Ingleby, M. Toward use of facial thermal features in dynamic assessment of affect and arousal level. *IEEE Transactions on Affective Computing 8*, 3 (July 2017), 412–425.

[35] Khoury, S., Rouleau, G. A., Rompré, P. H., Mayer, P., Montplaisir, J. Y., and Lavigne, G. J. A significant increase in breathing amplitude precedes sleep bruxism. *CHEST Journal 134*, 2 (2008), 332–337.

[36] Lewis, G. F., Gatto, R. G., and Porges, S. W. A novel method for extracting respiration rate and relative tidal volume from infrared thermography. *Psychophysiology 48*, 7 (2011), 877–887.

[37] Lim, C. L., Byrne, C., and Lee, J. K. Human thermoregulation and measurement of body temperature in exercise and clinical settings. *Annals Academy of Medicine Singapore 37*, 4 (2008), 347.

[38] Lykken, D. T. Polygraphic interrogation. *Nature 307*, 5953 (1984), 681–684.

[39] McGonagle, K. A., and Kessler, R. C. Chronic stress, acute stress, and depressive symptoms. *American journal of community psychology 18*, 5 (1990), 681–706.

[40] Mercer, J. B., and Ring, E. F. J. Fever screening and infrared thermal imaging: concerns and guidelines. *Thermology International 19*, 3 (2009), 67–69.

[41] Merla, A., Iodice, P., Tangherlini, A., De Michele, G., Di Romualdo, S., Saggini, R., and Romani, G. Monitoring skin temperature in trained and untrained subjects throughout thermal video. In *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the* (2006), IEEE, pp. 1684–1686.

[42] Mukaka, M. M. A guide to appropriate use of correlation coefficient in medical research. *Malawi Medical Journal 24*, 3 (2012), 69–71.

[43] Nakayama, Y., Sun, G., Abe, S., and Matsui, T. Non-contact measurement of respiratory and heart rates using a cmos camera-equipped infrared camera for prompt infection screening at airport quarantine stations. In *Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA), 2015 IEEE International Conference on* (2015), IEEE, pp. 1–4.

[44] Ng, E. Y., Kawb, G., and Chang, W. Analysis of ir thermal imager for mass blind fever screening. *Microvascular research 68*, 2 (2004), 104–109.

[45] Nhan, B. R., and Chau, T. Classifying affective states using thermal infrared imaging of the human face. *IEEE Transactions on Biomedical Engineering 57*, 4 (2010), 979–987.

[46] Papp, L. A., Martinez, J. M., Klein, D. F., Coplan, J. D., Norman, R. G., Cole, R., de Jesus, M. J., Ross, D., Goetz, R., and Gorman, J. M. Respiratory psychophysiology of panic disorder: three respiratory challenges in 98 subjects. *American Journal of Psychiatry 154*, 11 (1997), 1557–1565.

[47] Pavlidis, I. System and method using thermal image analysis for polygraph testing, Feb. 15 2005. US Patent 6,854,879.

[48] Pavlidis, I., and Levine, J. Thermal image analysis for polygraph testing. *IEEE Engineering in Medicine and Biology Magazine 21*, 6 (2002), 56–64.

[49] Pavlidis, I., Levine, J., and Baukol, P. Thermal image analysis for anxiety detection. In *Image Processing, 2001. Proceedings. 2001 International Conference on* (2001), vol. 2, IEEE, pp. 315–318.

[50] Pérez-Rosas, V., Abouelenien, M., Mihalcea, R., and Burzo, M. Deception detection using real-life trial data. In *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction* (New York, NY, USA, 2015), ICMI '15, ACM, pp. 59–66.

[51] Pérez-Rosas, V., Abouelenien, M., Mihalcea, R., Xiao, Y., Linton, C., and Burzo, M. Verbal and nonverbal clues for real-life deception detection. In *2015 Conference on Empirical Methods in Natural Language Processing, EMNLP 2015* (2015), Association for Computational Linguistics, pp. 2336–2346.

[52] Posner, J., Russell, J. A., and Peterson, B. S. The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and psychopathology 17*, 3 (2005), 715–734.

[53] Puri, C., Olson, L., Pavlidis, I., Levine, J., and Starren, J. Stresscam: non-contact measurement of users' emotional states through thermal imaging. In *extended abstracts on Human factors in computing systems* (2005), pp. 1725–1728.

[54] Rajoub, B. A., and Zwiggelaar, R. Thermal facial analysis for deception detection. *IEEE transactions on information forensics and security 9*, 6 (2014), 1015–1023.

[55] Schneiderman, N., Ironson, G., and Siegel, S. D. Stress and health: psychological, behavioral, and biological determinants. *Annu. Rev. Clin. Psychol. 1* (2005), 607–628.

[56] Schuirmann, D. J. A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. *Journal of pharmacokinetics and biopharmaceutics 15*, 6 (1987), 657–680.

[57] Sharma, N., Dhall, A., Gedeon, T., and Goecke, R. Modeling stress using thermal facial patterns: A spatio-temporal approach. In *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on* (2013), IEEE, pp. 387–392.

[58] Sun, G., Saga, T., Shimizu, T., Hakozaki, Y., and Matsui, T. Fever screening of seasonal influenza patients using a cost-effective thermopile array with small pixels for close-range thermometry. *International Journal of Infectious Diseases 25* (2014), 56–58.

[59] Sun, N., Garbey, M., Merla, A., and Pavlidis, I. Imaging the cardiovascular pulse. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on* (2005), vol. 2, IEEE, pp. 416–421.

[60] Sun, N., and Pavlidis, I. Counting heartbeats at a distance. In *Engineering in Medicine and Biology Society, 2006. EMBS'06. 28th Annual International Conference of the IEEE* (2006), IEEE, pp. 228–231.

[61] Tanda, G. The use of infrared thermography to detect the skin temperature response to physical activity. In *Journal of Physics: Conference Series* (2015), vol. 655, no. 1, IOP Publishing, p. 012062.

[62] Thayer, J. F., Yamamoto, S. S., and Brosschot, J. F. The relationship of autonomic imbalance, heart rate variability and cardiovascular disease risk factors. *International journal of cardiology 141*, 2 (2010), 122–131.

[63] van Dooren, M., Janssen, J. H., et al. Emotional sweating across the body: Comparing 16 different skin conductance measurement locations. *Physiology & behavior 106*, 2 (2012), 298–304.

[64] Walla, P., Nesbitt, K., Blackmore, K., Hookham, G., and Kay-Lambkin, F. Using the startle eye-blink to measure affect in players. In *Serious Games Analytics* (06 2015), pp. 401–434.

[65] Warmelink, L., Vrij, A., Mann, S., Leal, S., Forrester, D., and Fisher, R. P. Thermal imaging as a lie detection tool at airports. *Law and human behavior 35*, 1 (2011), 40–48.

[66] Wesley, A., Buddharaju, P., Pienta, R., and Pavlidis, I. A comparative analysis of thermal and visual modalities for automated facial expression recognition. In *International Symposium on Visual Computing* (2012), Springer, pp. 51–60.

[67] Wu, H.-Y., Rubinstein, M., Shih, E., Guttag, J., Durand, F., and Freeman, W. Eulerian video magnification for revealing subtle changes in the world. *ACM Transactions on Graphics 31*, 4 (2012), 1–8.

[68] Yang, M., Liu, Q., Turner, T., and Wu, Y. Vital sign estimation from passive thermal video. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on* (2008), IEEE, pp. 1–8.

[69] Yilmaz, T., Foster, R., and Hao, Y. Detecting vital signs with wearable wireless sensors. *Sensors 10*, 12 (2010), 10837–10862.

[70] Yuen, P., Hong, K., Chen, T., Tsitiridis, A., Kam, F., Jackman, J., James, D., Richardson, M., Williams, L., Oxford, W., et al. Emotional & physical stress detection and classification using thermal imaging technique, 2009.

[71] Zhu, Z., Tsiamyrtzis, P., and Pavlidis, I. The segmentation of the supraorbital vessels in thermal imagery. In *Advanced Video and Signal Based Surveillance, 2008. IEEE Fifth International Conference on* (2008), IEEE, pp. 237–244.