A Non-contact Method for Extracting Heart and Respiration Rates

Christian Hessler
Computer and Information Science
University of Michigan, Dearborn
Dearborn, MI, USA
cahessle@umich.edu

Mohamed Abouelenien
Computer and Information Science
University of Michigan, Dearborn
Dearborn, MI, USA
zmohamed@umich.edu

Mihai Burzo
Mechanical Engineering
University of Michigan, Flint
Flint, MI, USA
mburzo@umich.edu

***Abstract*—Physiological signals provide a reliable method to identify the physical and mental state of a person at any given point in time. Accordingly, there is a myriad of techniques used to extract physiological signals from the human body. However, these techniques often require direct contact with the body. This demands the cooperation of the individual as well as the human effort required to connect devices and collect measurements.**

In this paper, we propose reliable, non-contact based methods for extracting respiration rate and heart rate from thermal images using a large dataset of human thermal recordings. These methods leverage a combination of image and signal processing techniques in order to extract and filter physiological signals from the thermal domain. Our results evidently show that features extracted from thermal images highly correlate with the ground truth measurements as well as indicate the feasibility of developing non-contact based methods to extract physiological signals.

***Keywords*-thermal; physiological; heart rate; respiration rate;**

I. Introduction

There is a burgeoning research focus in developing automated systems capable of monitoring human physiological responses to provide a real-time assessment of a person's general health and well-being. Such measurements include heart rate, temperature and respiration rate. In addition, the fact that the human body exhibits unique physiological characteristics in response to external stimuli, offers the potential to detect and predict a person's behavior or psychological state, such as emotion, mood, stress level, distraction, and deceit. Hence, different studies are exploring the feasibility of incorporating physiological monitoring into a wide array of applications.

However, there are limitations to the traditional methods and devices used to collect physiological measurements, such as the need to connect the devices and attach sensors to the human body. Physically 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 patient's skin. Even worse, some sensors may fail to provide reliable measurements outside of a controlled environment. [1] 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 sensor placed on the finger. For that reason, several newer approaches attempt to altogether avoid the use of wearable sensors for data collection. In particular, thermal image processing has been recently proposed as an alternative method for acquiring physiological data.

Works in the literature suggest that health monitoring may be more effective at diagnosing disorders which do not manifest external symptoms. Heart rate, for example, is a useful measure for diagnosing cardiovascular disease (CVD), which is a leading cause of death worldwide. Specifically, there is evidence linking resting heart rate to CVD risk factors, such as hypertension, obesity, family history and work stress [2]. Additional evidence suggests that elevated heart and respiration rates observed alongside each other immediately after trauma are acute predictors of delayed post traumatic stress disorder [3]. Therefore, examining multiple vital signs may be more effective for diagnosing specific ailments.

Thermal imaging utilizes a principle called thermoregulation to detect natural thermal radiation emitted by the skin, which can be interpreted in terms of physiological changes [4]. The skin receives signals from control centers in the brain to maintain the body's core temperature [5]. Therefore, physiological thermoregulation in humans comprises changes in heat dissipation (sweating) and heat generation (shivering) in response to various internal and external thermal stimuli [6]. Research exploring the potential of extracting physiological measurement from thermal images has been very limited due to the cost associated with thermal camera sensors. However, the cost has been decreasing significantly recently with improvements in the thermal sensors technology.

This paper explores various methods to extract multiple physiological measurements, which could prove useful for a variety of multimodal applications, using a dataset collected from more than 100 subjects using a state-of-the-art thermal

978-1-7281-9891-0/20/$31.00 ©2020 IEEE
DOI 10.1109/CRV50864.2020.00009

camera. Moreover, our methodology explores the benefits of using wavelet transforms to filter physiological signals. This departs from many previous works in the literature, which largely relied on traditional signal processing methods, such as Fourier transforms and/or band-pass filtering, which are less effective at localizing features of non-stationary signals. Finally, we calculated cross correlation to measure the similarity between the sensor and thermal signal waveforms. This provides a much higher level of granularity for the purpose of evaluation as opposed to simply comparing frequencies, which was a common practice used in previous work.

II. Respiration Rate

Methods have been proposed to extract respiration rate from thermal videos using different combinations of image processing and facial tracking techniques. Figure 1 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. Secondly, a facial detection algorithm is 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.

[7] 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. They extracted the respiration signal by first calculating the mean intensity within each ROI for every frame. This was followed by low pass filtering to remove noise from the signal.

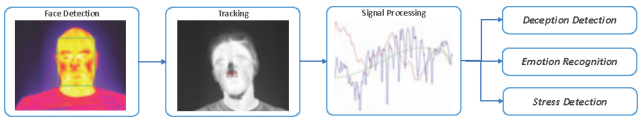

Figure 1. Image Processing Procedure

III. Heart Rate

Several studies have proposed methods to extract heart rate from thermal images by tracking superficial blood vessels within 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 blood vessels and surrounding tissue. This is a result of heat diffusion, which creates a smooth gradient temperature between hot and cold areas. Fortunately, there is a mean to overcome this challenge by 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 [8].

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 demonstrates the result of applying the Canny edge detection algorithm to a gray-scale thermal image from our dataset. 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 3 shows the result after applying this technique.

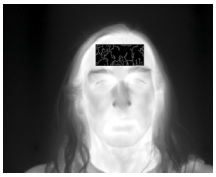

Figure 2. Vascular map extraction using Canny edge detection

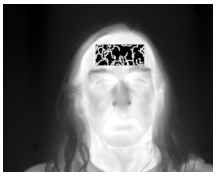

Figure 3. Canny edge detection after expansion

The thermal signal detected along a blood vessel presents a composite signal that includes extraneous physiological and environmental signals in addition to the pulse [9]. Therefore, the thermal signal must be filtered to extract the signal(s) of interest. [10] proposed 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.

In [11], the authors introduced several improvements based on previous work [8], [12], [13], [14], [10], [15]. 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 filtering. 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.

IV. Dataset Description

We used our own dataset consisting of thermal recordings collected from 104 participants with a gender distribution of 53 females and 51 males 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 collected.

V. Experimental Setup

The camera used to record the videos was a state-of-the-art 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. The two bio-sensors were used as to provide the ground truth; a blood volume pulse sensor and an abdominal respiration band. The first sensor was attached to the fingers of the subject's non-dominant hand. The abdominal respiration band was placed around the thoracic region.

Our experimental station consists of the recording device, the physiological sensors, two desktop computers, and a chair placed at a fixed distance from the camera. 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 camera.

VI. 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 the respiration rate and heart rate signals.

A. 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.

Features to Track within each 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 the tracked interesting points, we applied a geometric transformation method, 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 did not expect any rapid motion between successive frames. Therefore, we assumed the distance between a point and its projection in the next frame to be minimal. For that reason, we limited 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.

B. Thermal Signal Construction

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.

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.

C. Respiration Rate

We developed a method for extracting respiration rates from the maxillary (nose) region of a thermal image. We tested our method using our dataset, which includes recordings of subjects while they are speaking as well as 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, which is a direction that, to our knowledge, was not explored before. 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 $\hat{S(t)}$

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

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 $\sigma = 1$.

$$\hat{S(t)} = \frac{S(t) - \mu}{\sigma} \tag{2}$$

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

$$\hat{S(t')} = \frac{\sum_{i=t}^{t+k} S(t)}{k} \tag{3}$$

4) Continuous Wavelet Transform: Mexican Hat a.k.a. Ricker Wavelet was used as the mother wavelet $\psi(t)$. Equation 4 describes the Ricker Wavelet in which, σ is the standard deviation and t represents time. Equation 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.

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

$$S_w(a,b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} S(t)\psi\left(\frac{t-b}{a}\right) dt \tag{5}$$

5) Breathing Waveform: Equation 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 $WT_s(t)$ [16]. Lower scales are likely to contain noise, while higher scales contain metabolic contributions.

$$s\text{max} = \text{argmax}\left\{ \sum |WT_i(t)|^2 \right\} \tag{6}$$

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 biosensors 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.

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 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 4 depicts an example of employing the approach on the filtered thermal respiration rate signal.

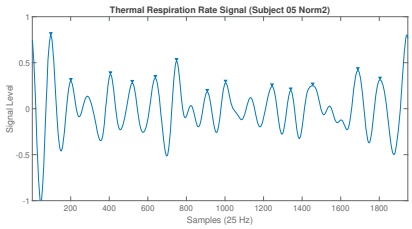

Figure 4. Example of Identifying Peaks in the Thermal Signal

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 coefficient of the shifted signals. Table I provides a general guide for interpreting the correlation coefficient in the context of medical applications [17]. 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 I
RULE OF THUMB FOR INTERPRETING THE STRENGTH OF A CORRELATION COEFFICIENT [18]

D. Experimental Trials

We tested several different methods and settings on a sample of our dataset 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. After conducting multiple trials utilizing different frequencies and order of operations, we found that reducing the frequency to 25Hz before applying the continuous wavelet transform produces the best results without the risk of neglecting the breathing signal.

Finally, in addition to measuring correlation, we counted the number of peaks in the signal and scaled 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.

E. 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. 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.

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 III. Figure 5 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 5. 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.

$$\hat{S(t')} = \frac{\sum_{i=t}^{t+k} S(t)}{k} \tag{7}$$

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 provides a visual representation of the DB10 wavelet with corresponding scaling function.

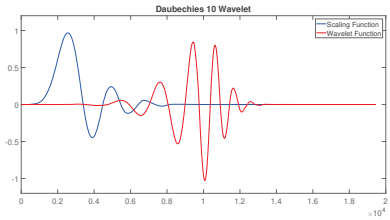

Figure 6. DB10 Wavelet & Scaling Functions

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

$$\psi_{\beta,\gamma}(\omega) = U(\omega)\alpha_{\beta,\gamma}\omega^{\beta}e^{\omega^{\gamma}} \tag{8}$$

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 9, $\langle\rangle$ denotes the inner product and Re$\{\ \}$ denotes the real part of the function.

$$f(t) = 2Re\left\{\frac{1}{C_{\psi,\delta}}\int_0^{\infty} < f(t), \psi_{a,b}(t) > \frac{da}{a}\right\} \tag{9}$$

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 VI-C.

VII. Results

A. 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

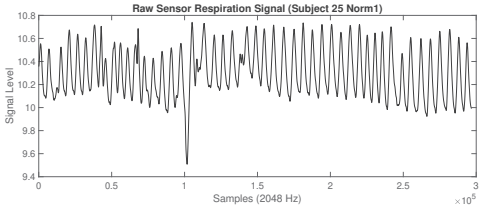

Figure 7. Subject 25 Raw Sensor Signal

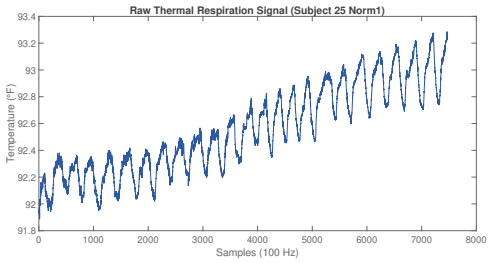

Figure 8. Subject 25 Raw Thermal Signal

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. Figures 7, 8, and 9 display examples of the sensor and thermal respiration rate signals during various stages of processing. These figures display the respiration signal extracted from a subject breathing normally without speaking.

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

Table II
VIDEO CATEGORICAL BREAKDOWN

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

Category	Before	After
Norm	0.58	0.61
Talking	0.40	0.43

Table III
AVERAGE CORRELATION RATES BEFORE & AFTER REMOVING OUTLIERS

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 most

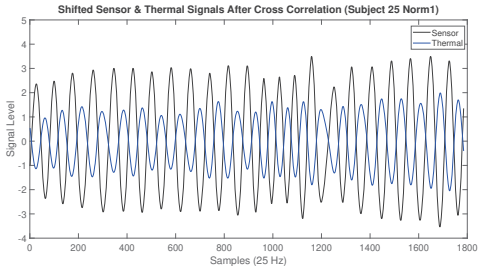

Figure 9. Subject 25 Transformed Signals (Correlation = 0.98)

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

Table IV
RESPIRATION RATE: TWO ONE-SIDED T-TEST RESULTS

of them were 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 III. 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 II 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.

Schuirmann's Two One Sided Tests (TOST) Procedure was used to assess equivalence between the ground truth and thermal respiration rates [19]. 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

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

Table V
RESPIRATION RATE: TWO ONE-SIDED T-TEST EQUIVALENCE BOUNDS

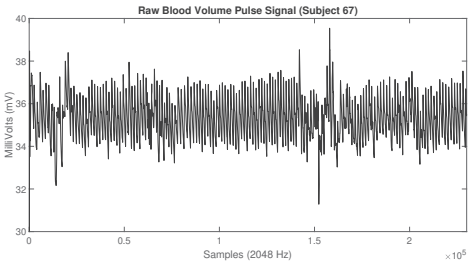

Figure 10. Subject 67 Raw Sensor Heart Rate Signal

truth and thermal respiration rates are equal within a ± 1 bpm and an 80% confidence interval.

The results of the two one sided T-test are listed in table IV, in which 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 the associated p-value. The confidence intervals are listed in table V, 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.

B. 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 'Norm' and 'Talking' videos as we did in section VII-A 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 VI 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.

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

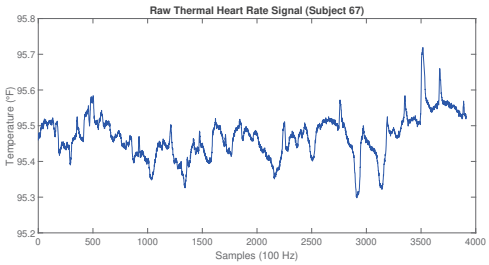

Figure 11. Subject 67 Raw Thermal Heart Rate Signal

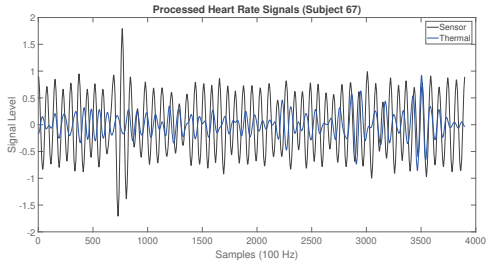

Figure 12. Subject 67 Processed Heart Rate Signals

from 12bpm to 10bpm as shown in table VI. This also proves our point that contact-based sensors do not necessarily provide accurate measurements and a non-contact approach can not only be more feasible, but more reliable as well. Figures 10, 11, and 12 display examples of the sensor and thermal heart rate 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 results listed in Table VII reveal that the two sets are equivalent in the interval [-5,+5] for a 99% confidence interval. Table VIII 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.

VIII. CONCLUSION

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 by first defining and tracking our thermal ROI, second constructing a thermal signal from the tracked ROI, and third filtering the thermal signal to extract physiological measurements.

Our results show a significant relation between the respiration rates extracted from thermal imaging and those recorded by contact-based sensors. Mean respiration rates

	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

Table VI
HEART RATE: DESCRIPTIVES

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

Table VII
HEART RATE: TWO ONE-SIDED T-TEST RESULTS

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

Table VIII
HEART RATE: TWO ONE-SIDED T-TEST EQUIVALENCE BOUNDS

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 0.61 between the thermal and sensor breathing signals for the Norm recordings. Finally, we extracted heart rates from the thermal domain. The mean difference between the sensor and thermal heart rates was five beats per minute. Hence, this research has shown that non-contact approaches using thermal imaging represent a reliable alternative to using contact-based sensors.

REFERENCES

[1] T. Gao, D. Greenspan, M. Welsh, R. R. Juang, and A. Alm, "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*. IEEE, 2006, pp. 102–105.

[2] J. F. Thayer, S. S. Yamamoto, and J. F. Brosschot, "The relationship of autonomic imbalance, heart rate variability and cardiovascular disease risk factors," *International journal of cardiology*, vol. 141, no. 2, pp. 122–131, 2010.

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

[4] B. F. Jones and P. Plassmann, "Digital infrared thermal imaging of human skin," *IEEE Engineering in Medicine and Biology Magazine*, vol. 21, no. 6, pp. 41–48, 2002.

[5] M. Dawson, A. Schell, and D. Filion, "The electrodermal system," *Handbook of psychophysiology*, vol. 2, pp. 200–223, 2007.

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

[7] S. L. Bennett, R. Goubran, and F. Knoefel, "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*. IEEE, 2017, pp. 3835–3839.

[8] P. Buddharaju, I. T. Pavlidis, P. Tsiamyrtzis, and M. Bazakos, "Physiology-based face recognition in the thermal infrared spectrum," *IEEE transactions on pattern analysis and machine intelligence*, vol. 29, no. 4, pp. 613–626, 2007.

[9] M. Garbey, N. Sun, A. Merla, and I. Pavlidis, "Contact-free measurement of cardiac pulse based on the analysis of thermal imagery," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 8, pp. 1418–1426, 2007.

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

[11] T. Gault and A. Farag, "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.

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

[13] T. R. Gault, N. Blumenthal, A. A. Farag, and T. Starr, "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*. IEEE, 2010, pp. 1–8.

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

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

[16] J. Fei and I. Pavlidis, "Thermistor at a distance: unobtrusive measurement of breathing," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 4, pp. 988–998, 2010.

[17] M. M. Mukaka, "A guide to appropriate use of correlation coefficient in medical research," *Malawi Medical Journal*, vol. 24, no. 3, pp. 69–71, 2012.

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

[19] D. J. Schuirmann, "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*, vol. 15, no. 6, pp. 657–680, 1987.