

Thermal Vision for Sleep Apnea Monitoring

Jin Fei¹, Ioannis Pavlidis¹, and Jayasimha Murthy²

¹ Computational Physiology Lab,
University of Houston, Houston, TX, USA
jin.fei@gmail.com, ipavlidis@uh.edu

² Division of Pulmonary, Critical Care and Sleep Medicine,
University of Texas Health Science Center, Houston, TX USA
Jayasimha.Murthy@uth.tmc.edu

Abstract. The present paper proposes a novel methodology to monitor sleep apnea through thermal imaging. First, the nostril region is segmented and it is tracked over time via a network of cooperating probabilistic trackers. Then, the mean thermal signal of the nostril region, carrying the breathing information, is analyzed through wavelet decomposition. The experimental set included 22 subjects (12 men and 10 women). The sleep-disordered incidents were detected by both thermal and standard polysomnographic methodologies. The high accuracy confirms the validity of the proposed approach, and brings non-obtrusive clinical monitoring of sleep disorders within reach.

1 Introduction

Sleep apnea is a respiratory disorder in which the breath stops repetitively during sleep. It can occur hundreds of times during a single night and each breath pause takes more than ten seconds. Sleep apnea is a common disorder and its prevalence can be as high as 30% among middle aged adults. It is associated with the development of high blood pressure and other cardiovascular diseases, and may lead to metabolic, organic, central nervous system, and endocrine ailments [1]. Therefore, there is a strong need for unobtrusive breathing measurement methods, where lengthy sleep studies with the minimum amount of discomfort are required.

Various contact modalities have been developed to assess the likelihood of apnea that capitalize on different aspects of the breathing phenomenon. Polysomnography (PSG) is the most reliable diagnostic method for the detection of sleep apnea syndrome. It is a multi-channel wired signal acquisition system which typically records ECG, nasal airflow, abdominal and thoracic movements, and blood oxygen saturation SpO_2 . The sensor probes and cables are uncomfortable for a patient under monitoring. They may interfere with the usual sleep pattern of the patient and influence the results of the test.

Human breathing consists of expiration and inspiration phases. The expired air has higher temperature than the inspired air due to heat exchange in the lungs and respiratory passageways [2]. This thermal nature of breath around

the nostril area creates an opportunity for a thermal measurement. Thermal infrared imaging is a passive contact-free modality. The sensing element itself can be viewed as a 2D array of contact-free thermistors, and therefore has a principle of operation similar to the nasal thermistor probe in PSG.

In this paper, we introduce a novel thermal imaging method to detect sleep apnea in the clinical sleep lab. It is based on automatic tracking/localization of the nasal region and wavelet analysis. The sleep apnea events were automatically detected from the onset of higher energy at lower frequency wavelets. In Section 2, we describe the tracking and localization algorithms as well as the wavelet-based detection method. We discuss the experimental setup in Section 3.1 and present the experimental results in Section 3.2. Section 4 concludes the paper.

2 Methodology

To measure the breathing function in thermal video we need to track the motion of subject, localize the measurement region, and analyze the extracted signal. We address each of these issues in detail in the following subsections.

2.1 Tracking

We chose the coalitional tracking algorithm [3] to track facial tissue during breath measurements. It optimizes multi-tracker interaction via game theory. The coalitional tracking method was developed to address the conflicting goals of generality and accuracy that arise in the context of thermo-physiological measurements on the face. Thermal imaging is functional imaging that depicts an evolving physiological process. The dynamic nature of thermal imaging poses a modeling challenge to tracking. Particle filter trackers [4] overcome this challenge because they are general and adapt well to changes. The accuracy of these trackers peaks when the real estate they cover is neither too large nor too small. By optimizing the behavior of a spatially distributed cluster of particle filter trackers (coalition), one gains in accuracy without sacrificing adaptability.

We use a coalition grid composed of four particle filter trackers. The grid outline is drawn interactively by a click and drag operation on the first frame.

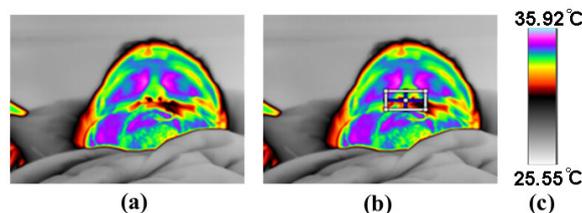


Fig. 1. (a) Thermal snapshot of a subject's face. (b) Initialization of coalitional tracker (TROI). (c) Thermal color map.

It encompasses comfortably the nostril region without any stricter specification. The grid outline constitutes the Tracking Region of Interest (TROI) - see white rectangle in Figure 1. The localization algorithm (see Section 2.2) operates within TROI and determines more rigorously the nostril region of measurement.

2.2 Localization

The source of thermal signal is the nostrils where we select the Measurement Region of Interest (MROI) for breathing. This region features both spatial and temporal variances. First, the shape of the nostril region is different for different individuals. Thermal imaging is a functional imaging modality that records the changing image physiology. In the case of breathing, thermal imagery registers the temperature fluctuation between the inspiration and expiration phases. This, however, increases the segmentation difficulty, as the shape of nostrils varies temporally due to the varying thermal signature of inspiration and expiration.

Figure 2 shows how within TROI the nostrils are separated from the rest of the facial tissue due to colder boundaries formed by cartilage. This feature can help to localize MROI. The contrast at the boundary of the nose is quite strong not only in thermal imagery but also in visual imagery, for different reasons. The nose is a distinct 3D feature in an otherwise 2D facial surface and forms strong edges at the seams. Due to similar nose boundary properties, we can leverage some of the work performed in the visual spectrum for thermal imaging purposes. Specifically, Brunelli and Poggio [5] showed that the horizontal gradients are useful in detecting the left and right visual boundaries of nose, whereas vertical gradients can detect the nose base. Kotropoulos and Pitas [6] demonstrated that the vertical and horizontal projection profiles of human nose are obtained by summing-up visual pixel intensities row-wise and column-wise, respectively. We have used elements of these approaches transplanted in the thermal infrared domain.

Let $I(x, y)$ be the original thermal image and $E_X(x, y)$ and $E_Y(x, y)$ be the edge images after applying the Sobel edge detectors. We perform integral

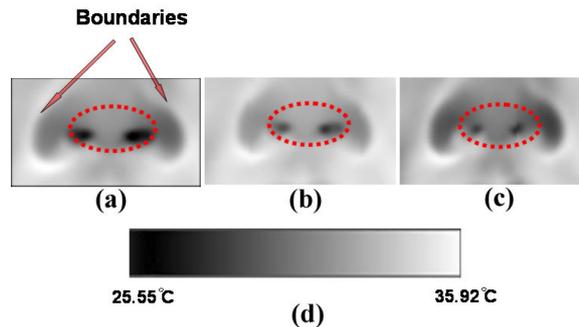


Fig. 2. Temporal variance of nostril region in thermal imagery during breathing. (a) Inspiration phase. (b) Transition phase. (c) Expiration phase. (d) Thermal color map.

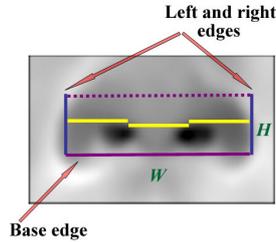


Fig. 3. MROI based on mean nose edges and anthropometric estimates

projections on the edge images to extract the outer edges of nostrils. The vertical and horizontal integral projections are:

$$P_v(x) = \sum_{y=y_1}^{y_m} E_X(x, y), \quad (1)$$

$$P_h(y) = \sum_{x=x_1}^{x_n} E_Y(x, y). \quad (2)$$

The left- and right-most peaks of $P_v(x)$ yield the left and right outer nostril edges correspondingly. The maximum of $P_h(y)$ yields the base edge.

The MROI selection varies from frame to frame. Some projections are weak and the locations vary as well. We use a time window of 4 – 5 *sec* to compute the mean vertical and horizontal projections. Such window is representative of the full spatio-temporal evolution, as it covers both expiration and inspiration phases.

Based on anthropometric knowledge [7], the distance between left and right alares is about 30 *mm* and the distance between subnasale and columella is about 10 *mm*. Therefore, we estimate the nostril's height H as one third of its base edge segment W , which is delineated by the left and right outer nostril edges. Thus, we construct the MROI $W \times H$ (see Figure 3). We compute the mean temperature within MROI in every frame. This produces a quasi-periodic thermal signal along the timeline, which is indicative of the breathing function.

2.3 Wavelet Analysis

We perform wavelet analysis on the imaged thermal signal to detect the incidents of sleep apnea. Wavelet is the appropriate analysis tool as breathing is a non-stationary process.

The thermal video sampling rate fluctuates. A constant sampling rate is necessary for optimal results in wavelet decomposition. We choose the re-sampling rate of the thermal signal as 10 *fps*.

We normalize and perform wavelet analysis on sliding segments (windows) of the re-sampled thermal signal. As the sliding window travels along the evolving timeline of the re-sampled and normalized signal, we compute a series of wavelet

energy coefficients. This renders apnea incident detection real time. The details of each algorithmic step are as follows:

Normalization. We define as $S(t)$, $t \in \{0, \dots, N\}$, the re-sampled breathing signal. We normalize the signal amplitude as follows:

$$S'(t) = \frac{S(t) - \mu}{\sigma}, \quad (3)$$

where μ and σ are the mean and standard deviation of $S(t)$ respectively. The normalization transforms signal $S(t)$ to $S'(t)$ with mean $\mu' = 0$ and standard deviation $\sigma' = 1$.

Wavelet Transform. We perform Continuous Wavelet Transformation (CWT) on the resampled and normalized thermal signal:

$$\Psi_{S'}^{\psi}(\tau, s) = \frac{1}{\sqrt{|s|}} \int S'(t) \psi\left(\frac{t - \tau}{s}\right) dt, \quad (4)$$

where ψ is the ‘mother wavelet’, τ represents the translation parameter, while s denotes the scale at which the signal is examined. We use the Mexican Hat (MH) as the mother wavelet.

Wavelet Energy. CWT allows analysis at all scales, hence, facilitating the extraction of the signal component of interest (i.e., breathing and apnea). We assume that the wavelet energy at scale s_i corresponding to the wavelet coefficients $WT_i(t)$ is:

$$P_i = \sum |WT_i(t)|^2. \quad (5)$$

We define P_n and P_o the wavelet energies of normal and obstructive breathing respectively. In normal breathing, P_n is larger than P_o . However, P_o increases at low frequency (higher scale) during sleep apnea incidents. Hence, we choose to monitor P_o for detection of sleep apnea events.

3 Experiments

3.1 Experimental Setup

The center-piece of the imaging system we used in our experiments is a FLIR SC6000 Mid-Wave Infra-Red (MWIR) camera with an Indium Antimonite (InSb) detector operating in the range $3 - 5 \mu m$ [8]. The camera has a focal plane array (FPA) with maximum resolution of 640×512 pixels. The sensitivity is $0.025^\circ C$. The camera is outfitted with a MWIR 100 mm lens $f/2.3$, $Si : Ge$, bayonet mount from FLIR Systems [8]. The MWIR camera was calibrated with a two-point calibration at $28^\circ C$ and $38^\circ C$, which are the end points of a typical temperature distribution on a human face.

The experiments took place in a climate controlled room according to an approved protocol by the Institutional Review Board of the University of Texas

Health Science Center. Subjects, lying in a comfortable bed, were positioned 10 *ft* away from the imaging system. The subjects were fitted with standard polysomnography (PSG) to ground-truth the imaging measurements. We recorded approximately one-hour long thermal clips (and the corresponding PSG signals) for each of the twenty-two subjects (twelve men, ten women). The age range was 24 to 66 years and the Body Mass Index (BMI) 19.71 to 45.57 kg/m^2 . The control group included twelve subjects who had no history of obstructive sleep apnea. The pathological group had ten subjects with clinical diagnosis of obstructive sleep apnea. The ground-truth apnea events were detected by clinical specialists and reviewed by clinical doctors.

3.2 Results

The thermal signal produced in the vicinity of the nasal area is detectable when subjects are at sleep and can be recovered in pristine form with the help of the tracker. The interweaved inspiration and expiration phases produce periodic breathing signals. Figure 4 depicts a breathing signal from a patient suffering from obstructive sleep apnea. We report the presence of apnea (events) in 30 sec epochs, following established standards for clinical diagnosis. The computerized method features redundant epochs that overlap consecutive epochs by a half epoch. Therefore, we can detect the apnea event in $Epoch_{i+3}$, in the example shown in Figure 4. This event may be missed if only the consecutive epochs $Epoch_{i+2}$ and $Epoch_{i+4}$ are considered.

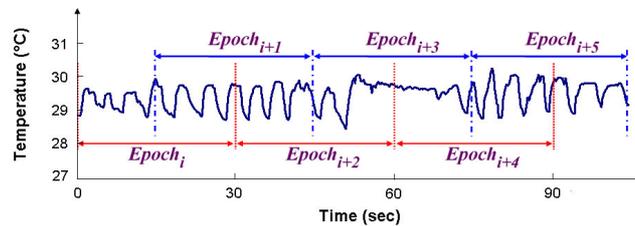


Fig. 4. Analysis of thermal breathing signals with overlapped epochs

We compute the wavelet energy from the thermal breathing curves. Apnea incidents are detected when P_o drops at the frequency of 6 *cpm*. In Figure 5, the spikes represent the epochs with apnea incidents. The wavelet energy threshold corresponding to the chosen P_o threshold is 2000 - it is computed from the simulated signal (similar to $Epoch_{i+3}$ in Figure 4).

Professional sleep medicine specialists manually scored the sleep events from PSG. Physicians certified in the specialty of Sleep Medicine reviewed scoring of apneas and finalized the interpretation of PSG. the apnea reports. The detected apnea events based on the computerized method are nearly the same as in the clinical report. The main difference is that the scoring process has been automated.

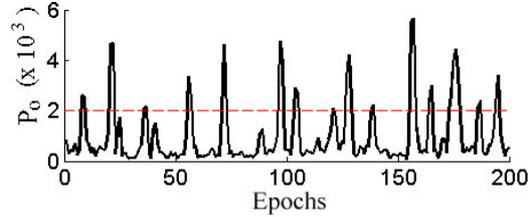


Fig. 5. Energy P_o for all the epochs

Table 1. Performance

Group	PSG Event	TP	TN	FP	FN	Accuracy	Precision	Recall
Control	22	19	2553	92	3	96.44	86.36	96.52
Pathological	145	132	1337	123	13	91.53	91.03	91.58
Overall	167	151	3890	215	16	94.59	90.42	94.76

We evaluated the measurement accuracy with three performance indicators: Accuracy, precision, and recall. Table 1 compares the results of true positive (TP), true negative (TN), false positive (FP), and false negative (FN) from both the control and pathological groups. Overall, the doctors reported 167 apnea events from PSG - 151 of these were detected by thermal imaging. There were 16 ‘false negative’ cases. Most inconsistencies arose from ‘false positive’. Such faults mainly resulted from tracking failures, when the tracker was misdirected to the background and away from the periodic effect of breathing. This resulted in a flat signal, which resembled an apnea event. The accuracy, precision, and recall were 94.59%, 90.42%, and 94.76%, respectively. The high performance proves that the thermal imaging method is promising in detecting incidents of sleep apnea. Besides the contact-free benefits, thermal imaging analysis could also assist in automating the diagnosis of sleep apnea.

4 Discussion and Conclusion

The current mean edge localization algorithm has been designed to deal with front views only. When the subject is on lateral posture (subject #3 and #21), we need to manually select the MROI. A pose estimation enhancement may fix this problem in the future. In case where both the nasal and oral areas are blocked by the blanket no measurement is feasible. Fortunately, such incidents happen rarely and do not last very long.

In general, the imaging method accurately recorded incidents of sleep apnea concomitantly with the standard detection instrument. Where the imaging method really shines is the highly automated and totally unobtrusive nature of its operation and the potential for improved post-processing, following the development of improved algorithms. By contrast, the subject needs to be outfitted

with obtrusive sensors and cables in traditional PSG. The discomfort caused by such a monitoring method may interfere with the sleep routine of the subject, thus, biasing the experiment. This imaging method may be especially beneficial in pediatric patients or in patients with facial trauma.

In this paper, for the first time we have described a new methodology based on passive imaging to detect sleep apnea in clinical studies. The present method can automatically localize the nasal region. After applying Wavelet Transform, the method detects sleep apnea events based on the onset of high energy at low frequency wavelets. The sensing system can operate as a computer peripheral, which opens the way for home-based sleep monitoring in the future.

Acknowledgments. This material is based upon work supported by the National Institutes of Health Clinical and Translational Sciences Award (UL1RR024148; PI: Frank C. Arnett, M.D.) at the University of Texas Health Science Center, a research contract by the Defense Academy for Credibility Assessment, and National Science Foundation Awards (IIS-0812526, CNS-0521527, and IIS-0414754) - PI: Ioannis Pavlidis, Ph.D. - at the University of Houston. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding agencies.

References

1. Pack, A.: Centennial review: advances in sleep-disordered breathing. *Am. J. Respir. Crit. Care Med.* 173, 7–15 (2006)
2. Zhang, Y.: Quantitative measurement of radiation properties for opaque and semi-transparent greybodies. *Infrared Physics* 30(2), 149–153 (1990)
3. Dowdall, J., Pavlidis, I., Tsiamyrtzis, P.: Coalitional tracking. *Computer Vision and Image Understanding* 106(2-3), 205–219 (2007)
4. Isard, M., Blake, A.: Condensation-conditional density propagation for visual tracking. *International Journal of Computer Vision* 29(1), 5–28 (1998)
5. Brunelli, R., Poggio, T.: Face recognition: Features versus templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 15(10), 1042–1052 (1993)
6. Kotropoulos, C., Pitas, I.: Rule-based face detection in frontal views. In: *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Munich, Germany, April 21-24, vol. 4, pp. 2537–2540 (1997)
7. Farkas, L.G.: *Anthropometry of the Head and Face in Medicine*. Elsevier, New York (1981)
8. FLIR Systems (2007), <http://www.flir.com>