

Face Tracking and Respiratory Signal Analysis for the Detection of Sleep Apnea in Thermal Infrared Videos with Head Movement

Marcin Kopaczka^(✉), Özcan Özkan, and Dorit Merhof

Institute of Imaging and Computer Vision, RWTH Aachen University,
Templergraben 55, 52074 Aachen, Germany
Marcin.kopaczka@lfb.rwth-aachen.de

Abstract. Infrared Thermography as imaging modality has gained increased attention over the last years. Its main advantages in human action monitoring are illumination invariance and its ability to monitor physiological parameters such as heart and respiratory rates. In our work, we present a novel approach for detecting respiratory-related data, in our case apnea events, from thermal infrared recordings. In contrast to already published methods where the subjects were required not to move, our approach uses state-of-the-art thermal face tracking technology to allow monitoring of subjects showing head movement, which is an important aspect for real-world applications. We implement different methods for apnea detection and face tracking and test them on videos of different subjects against a ground truth acquired with an established breathing rate monitoring system. Results show that our proposed approach allows robust apnea detection with moving subjects. Our methods allow using already presented or novel vital sign monitoring systems under conditions where the monitored persons are not required to keep their heads in a given position.

1 Introduction

Thermal Infrared or long-wave infrared (LWIR) thermography cameras detect electromagnetic waves with a wavelength between 7 and 14 μm . This is the energy window in which thermal radiation at room temperature is emitted, thereby allowing LWIR detectors to detect body heat without requiring an external light source, eliminating many problems connected to illumination variance. At the same time, thermal infrared is a relevant modality for human condition observation as many physiological parameters such as heart rate and breathing rate can be determined by analyzing the thermal infrared videos. However, while several authors have shown that vital signs can be derived from such video data, most of them have only presented results from lab studies where subject head movement was highly constrained, and pointed out that robust face tracking technology would be required in order to achieve applicability in unconstrained conditions. To this end, we present a comparison of different approaches for face tracking in thermal infrared images and show that they can be used to improve

robustness of algorithms for vital sign monitoring. Our chosen application is the detection of apnea events which has already been proven to work under constrained conditions.

2 Previous Work

There has been extensive previous work in the fields of sleep apnea detection, thermal infrared face tracking and thermal infrared vital sign extraction.

Obstructive sleep apnea is a common sleeping disorder that results in reduced blood oxygen levels and is mainly caused by obstruction of the upper airway [1]. Usually, sleep apnea is diagnosed using polysomnography, a method where different vital parameters such as EEG, EMG, EKG, air flow through mouth and nose and breathing movements are recorded during sleep and analyzed subsequently [4]. A commonly used, versatile and efficient method that can also be part of a polysomnographic recording is a thoracic-abdominal band that allows to measure upper body circumference changes and thereby allows extraction of the breathing movement [3].

Thermal imaging for medical purposes has been proven beneficial in different scenarios, for example for fever detection in airports [5], breast cancer detection [6] or inflammation [7]. A recent overview of different applications can be found in [2]. Contributions for the analysis and extraction of vital signs from facial images using thermal infrared recordings include methods for the monitoring of respiratory rate of newborns [9] and adults [8,10], heart rate [12,13] and more currently the thermal signatures of psychophysiological phenomena [14–16].

A common property of all literature listed above is that the presented approaches make only limited use of tracking technologies, in fact in most cases no tracking is applied at all. While this is sufficient for fundamental research or low-throughput measurements where the regions of interest (ROI) for thermal signature analysis can be updated manually on a frame-per-frame basis, any measurement that should allow head movement requires tracking. Only limited work has been published in the field of face and facial landmark tracking in thermal infrared images. Notably, [17] introduces a set of particle filters for tracking in thermal images, while the approach shown in [18] uses feature-based active appearance models for precise tracking of facial landmarks.

3 Materials and Methods

In this section we describe the tracking methods used to allow adapting ROI positions to head movements and the methods for apnea detection and respiratory rate measurement.

3.1 ROI Tracking

We implemented two state-of-the-art tracking mechanisms in order to allow tracking of facial regions:

- **TLD tracking** - TLD (Track, Learn, Detect) [19] is a general-purpose tracker making heavy use of online learning strategies. Constant updates of the target templates improve tracking accuracy over time, while a set of local and global correlation filters ensure robustness even for facial areas that strongly vary in appearance due to head movements. So far, TLD has not been applied to face tracking in thermal infrared images.
- **Feature-based active appearance models** - Feature-based active appearance models (AAMs) combine the well-established tracking approach of active appearance models with image feature descriptors for improved tracking robustness. They have been proven to show good performance in the tracking of faces in thermal infrared videos [18]. We used an AAM trained with a database of 2500 manually annotated thermal infrared images.

For TLD, the ROI for respiratory rate extraction was defined manually in the 1st frame of the video by drawing a box covering the nostrils. The tracker learned the ROI appearance and tracked it in the subsequent frames. For the feature-based AAM, the ROIs were defined automatically by using the two landmarks on the detected nostril positions and using them as centers of rectangular boxes with a width of 15 pixels. Figure 1 shows the results of both approaches on the same image.

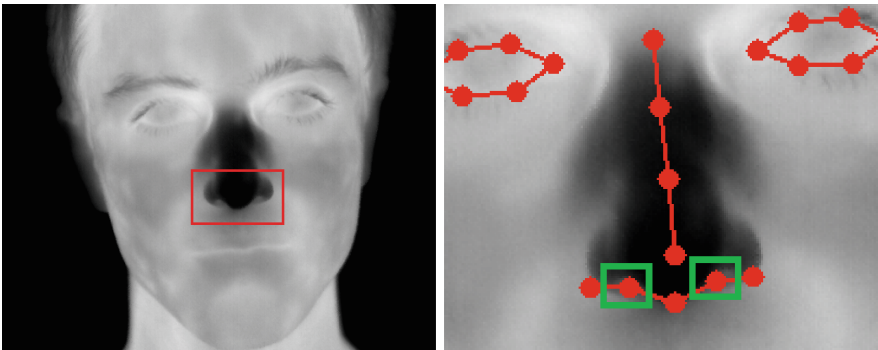


Fig. 1. ROI definition. Left: Manually defined bounding box for the static ROI and TLD tracking. Right: automatically defined ROIs computed from AAM tracking results.

3.2 Apnoe Detection

We developed and implemented different methods for apnea detection that all use the thermal signal extracted by computing the average or minimal temperature in the ROIs defined above. In a preprocessing step, all temporal temperature curves were filtered with a spectral lowpass defined using a Gaussian kernel with a width of 0.25s (7 frames at a frame rate of 30 fps) to reduce high-frequency noise. Subsequently, the following methods for apnea detection were applied:

- **Gradient Sum** - By assuming that regular breathing results in stronger signal change and thereby higher gradients, we computed the moving sum of the absolute temperature gradient curves over the past 4 s. Apnea events are considered as regions where the gradient value is below 0.6 times the average of the whole video sequence. An example of the output of the gradient analysis can be found in Fig. 2
- **Variance Analysis** - Similar to gradient analysis, the variance analysis method also relies on the fact that signal changes during apnea events occur less frequently than during regular breathing. For variance analysis, the temperature variance over the past 7 s is computed, subsequently all areas with a variance lower than 0.4 times the average variance of the analyzed video sequence are considered to be apnea events. Figure 3 shows the output of the variance analysis.
- **Spectral Analysis** - apnea events can be detected in the spectral domain as well. To this end, we analyze the temperature curve and subtract the average temperature of the past 1.3 s from each signal value to reduce low-frequency noise. Subsequently, Short-Time Fourier Transform with a window length of 10 s is applied to the filtered signal. We analyze the frequency window between 0.2 and 0.8 Hz over the last 5 s, as our preliminary studies have shown that the respiratory signal is dominant in this spectral range. When applying spectral analysis, the threshold for an apnea is set to 0.1 times the average signal energy of the sequence. An example result is shown in Fig. 4
- **Wavelet Transform** - The wavelet transform is similar to the Short-Time Fourier Transform since it also allows signal localization in both temporal and spectral space. The general applicability of the wavelet transform to apnea detection in thermal infrared images has already been shown in [11], in our work we introduce a slightly adapted and extended version that transforms the resulting wavelet into a set of one-dimensional values, thereby allowing the use of 1D signal processing methods as in the methods introduced above. In a first step, we use the method from [11], where we apply wavelet transform using the Mexican hat wavelet and compute 50 scales equidistantly between 0.21 and 0.75 Hz. Subsequently, we expand the original method by first applying thresholding to the result with a threshold value equal to the mean value of the wavelet coefficients. In order to extract a curve from the thresholded wavelet signal, we compute the sum of all coefficients for the past 5 s. The resulting curve has high similarity with the spectral curve shown in 4, and similar to the spectral analysis method introduced above we define an apnea event as areas where the extracted signal is lower than 0.1 times the average signal value.

4 Experiments and Results

To evaluate the implemented algorithms, we acquired thermal infrared recordings of 10 healthy subjects under laboratory conditions. The used camera provided a resolution of 1024×768 pixels with a thermal sensitivity of 0.03 K.

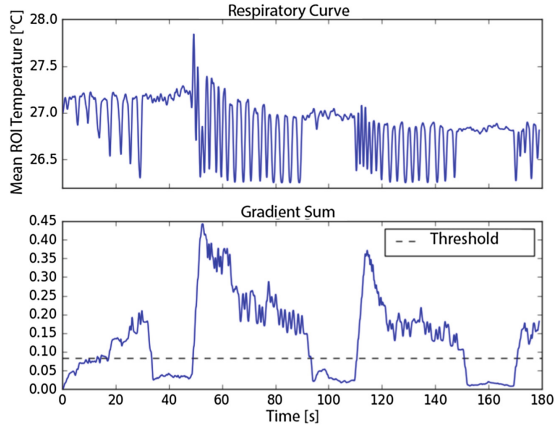


Fig. 2. Gradient analysis. Top: original temperature curve extracted from the ROI. Bottom: computed gradient sum.

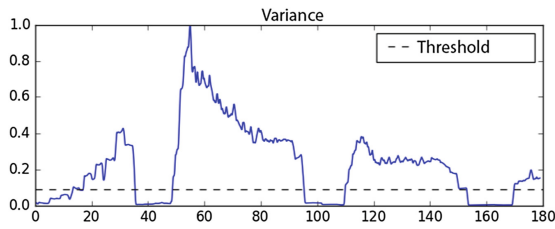


Fig. 3. Variance analysis results

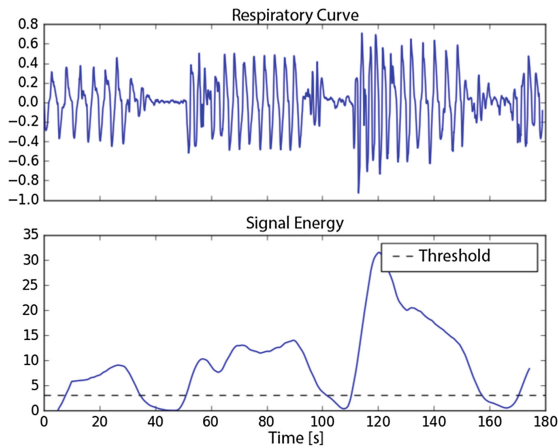


Fig. 4. Spectral analysis results for a window length of 10 s.

Each participant was filmed frontally for 5 min, see Fig. 5 for a sample frame. The persons were instructed to breath normally except for simulated apnea events that started after 60, 150 and 240 s of the recording. Since apnea usually occurs during rest, the head remained still during the apnea. Between the events, free head movement with increasing speed was allowed. The reference for apnea estimation was acquired by additionally utilizing a clinically approved thoracic-abdominal band as described in [3] that allowed measurement of thorax circumference and its changes. Apnea events were manually marked in the output signal of the belt.

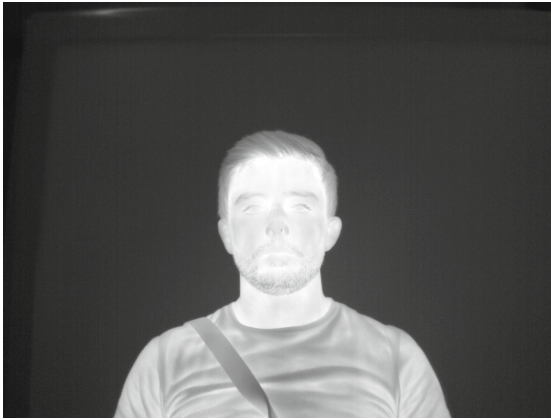


Fig. 5. Sample frame from an experimental infrared recording. Note that a holding strap of the chest belt is visible in the recording as well.

All acquired video sequences were subsequently tracked using the TLD and AAM method. For TLD, the initial ROI was drawn manually and tracked in subsequent frames. In the AAM, the head position was also defined manually in the 1st frame and tracked by the algorithm afterwards. For comparison with previously published work, we also analyzed a constant ROI (defined as the ROI used for initialization of the TLD tracker) as this is the method of choice in most available literature. The results are given in Table 1.

The results show that both tracking methods clearly outperform a non-tracked analysis. Of the two implemented trackers, the AAM constantly provides better results than the TLD method. All four apnea detection algorithms show similar performance, with the spectral methods being more robust towards misdetections than the two time-based approaches.

Table 1. Apnoe detection results using different trackers and detectors.

Detector	Tracker	True Positive	False Positive	False Negative
Gradient Sum	Static	16	27	9
Gradient Sum	TLD	26	4	3
Gradient Sum	AAM	30	3	0
Variance Analysis	Static	17	33	6
Variance Analysis	TLD	25	25	3
Variance Analysis	AAM	29	9	1
Spectral Analysis	Static	18	23	3
Spectral Analysis	TLD	25	2	3
Spectral Analysis	AAM	30	2	0
Wavelet Transform	Static	20	14	6
Wavelet Transform	TLD	25	2	5
Wavelet Transform	AAM	29	1	1

5 Conclusion

In this work, we introduced different algorithms for the detection of sleep apnea in thermal infrared video sequences. To improve robustness of the detection methods, we also implemented two algorithms for face region tracking in thermal infrared recordings. Results show that the presented methods allow reliable apnea detection in thermal infrared recordings and that modern face tracking algorithms clearly improve the robustness of the apnea detection.

Future work should include a real-time implementation of the described algorithms and a validation of the proposed method in a clinical scenario.

References

1. Somers, V.K., White, D.P., Amin, R., Abraham, W.T., Costa, F., Culebras, A., Daniels, S., Floras, J.S., Hunt, C.E., Olson, L.J., et al.: Sleep apnea and cardiovascular disease. *Circulation* **118**(10), 1080–1111 (2008)
2. Lahiri, B.B., Bagavathiappan, S., Jayakumar, T., Philip, J.: Medical applications of infrared thermography: a review. *Infrared Phys. Technol.* **55**(4), 221–235 (2012)
3. Denissova, S.I., Yewondwossen, M.H., Andrew, J.W., Hale, M.E., Murphy, C.H., Purcell, S.R.: A gated deep inspiration breath-hold radiation therapy technique using a linear position transducer. *J. Appl. Clin. Med. Phys.* **6**(1), 61–70 (2005)
4. Bloch, K.E.: Polysomnography: a systematic review. *Technol. Health Care* **5**(4), 285–305 (1997)
5. Nguyen, A.V., Cohen, N.J., Lipman, H., Brown, C.M., Molinari, N.-A., Jackson, W.L., Kirking, H., Szymanowski, P., Wilson, T.W., Salhi, B.A., et al.: Comparison of 3 infrared thermal detection systems and self-report for mass fever screening. *Emerg. Infect. Dis.* **16**(11), 1710 (2010)

6. Arora, N., Martins, D., Ruggerio, D., Tousimis, E., Swistel, A.J., Osborne, M.P., Simmons, R.M.: Effectiveness of a noninvasive digital infrared thermal imaging system in the detection of breast cancer. *Am. J. Surg.* **196**(4), 523–526 (2008)
7. Lasanen, R., Piippo-Savolainen, E., Remes-Pakarinen, T., Kröger, L., Heikkilä, A., Julkunen, P., Karhu, J., Töyräs, J.: Thermal imaging in screening of joint inflammation and rheumatoid arthritis in children. *Physiol. Meas.* **36**(2), 273 (2015)
8. Pereira, C.B., Yu, X., Czaplik, M., Blazek, V., Venema, B., Leonhardt, S.: Estimation of breathing rate in thermal imaging videos: a pilot study on healthy human subjects. *J. Clin. Monit. Comput.* October 2016
9. Abbas, A.K., Heimann, K., Jergus, K., Orlikowsky, T., Leonhardt, S.: Neonatal non-contact respiratory monitoring based on real-time infrared thermography. *Bio-med. Eng. Online* **10**(1), 93 (2011)
10. Fei, J., Pavlidis, I.: Thermistor at a distance: unobtrusive measurement of breathing. *IEEE Trans. Biomed. Eng.* **57**(4), 988–998 (2010)
11. Fei, J., Pavlidis, I., Murthy, J.: Thermal vision for sleep apnea monitoring. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) *MICCAI 2009*. LNCS, vol. 5762, pp. 1084–1091. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-04271-3_131
12. Gault, T.R., Farag, A.A.: A fully automatic method to extract the heart rate from thermal video. In: *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, June 2013
13. Jing, B., Li, H.: A novel thermal measurement for heart rate (2013)
14. Panasiti, M.S., Cardone, D., Pavone, E.F., Mancini, A., Merla, A., Aglioti, S.M.: Thermal signatures of voluntary deception in ecological conditions. *Sci. Rep.* **6** (2016)
15. Cardone, D., Merla, A.: New frontiers for applications of thermal infrared imaging devices: computational psychophysiology in the neurosciences. *Sensors* **17**(5), 1042 (2017)
16. Paolini, D., Alparone, F.R., Cardone, D., van Beest, I., Merla, A.: The face of ostracism: The impact of the social categorization on the thermal facial responses of the target and the observer. *Acta Psychol.* **163**, 65–73 (2016)
17. Dowdall, J., Pavlidis, I.T., Tsiamyrtzis, P.: Coalitional tracking. *Comput. Vis. Image Underst.* **106**(2–3), 205–219 (2007)
18. Kopaczka, M., Acar, K., Merhof, D.: Robust facial landmark detection and face tracking in thermal infrared images using active appearance models. In: *Proceedings of the 11th Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2016), VISAPP, Rome, Italy, 27–29 February 2016*, vol. 4, pp. 150–158 (2016)
19. Kalal, Z., Mikolajczyk, K., Matas, J.: Tracking-learning-detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **34**(7), 1409–1422 (2012)