

Spatiotemporal Reconstruction of the Breathing Function

D. Duong¹, D. Shastri², P. Tsiamyrtzis³, and I. Pavlidis¹

¹ Department of Computer Science, University of Houston, Houston, TX 77024, USA

² Department of Computer and Mathematical Sciences, University of
Houston-Downtown, Houston, TX 77002, USA

³ Department of Statistics, Athens University of Economics and Business,
Athens 10434, Greece

dcduong@cs.uh.edu, shastrid@uhd.edu, pt@aub.gr, ipavlidis@uh.edu

Abstract. Breathing waveform extracted via nasal thermistor is the most common method to study respiratory function in sleep studies. In essence, this is a temporal waveform of mean temperatures in the nostril region that at every time step collapses two-dimensional data into a single point. Hence, spatial heat distribution in the nostrils is lost along with valuable functional and anatomical cues. This article presents the construction and experimental validation of a spatiotemporal profile for the breathing function via thermal imaging of the nostrils. The method models nasal airflow advection by using a front-propagating level set algorithm with optimal parameter selection. It is the first time that the full two-dimensional advantage of thermal imaging is brought to the fore in breathing computation. This new multi-dimensional measure is likely to bring diagnostic value in sleep studies and beyond.

Keywords: Breathing, data visualization, sleep studies, thermal imaging.

1 Introduction

Sleep studies require overnight monitoring of the patient's breathing function which is typically accomplished via contact-sensors. A widely used sensor is the nasal thermistor which extracts the temporal breathing waveform by sensing the average temperatures in the nostril region at every point in time. The sensor is placed inside the nostril, a non-comfortable arrangement for patients who have problems with breathing and sleep in the first place. As an alternative to this clinical practice, a thermal imaging method has been proposed recently [1][2]. The method could be characterized as a 'virtual thermistor', because it produces a temporal breathing waveform by averaging emission values in the nostrils at every time step. The comparative advantage lies only in its non-contact nature. Although thermal imaging carries inherently spatial information, this is never recovered and used. Evolution of spatial heat distribution in the nostrils can reveal subtle breathing abnormalities that may hint at anatomical

and functional problems. These problems by and large go undetected due to the averaging nature of the existing measurement methods [3]. Relevant examples include small nasal polyps that locally affect airflow or hypopnea, where airflow is curtailed but not totally suppressed [4].

In this article, we describe a method for spatiotemporal reconstruction of the breathing function via thermal imaging.¹ A level set algorithm captures the spatial evolution of nostril emission, as affected by inspiratory and expiratory airflow. A registration algorithm that accounts for nostril motion ensures the meaningful application of the level set computation. The soundness of the method is verified experimentally. Interestingly, the method captures subtle pathophysiological incidents in the data set that escape detection by the 'virtual thermistor', thus, bringing to the fore its potential clinical value.

2 Methodology

2.1 Temporal Registration

Breathing is a physiological process that continuously modulates the spatial heat distribution inside the nostrils (see Fig. 1). We capture this spatial evolution by employing the Chan-Vese active contour without edge model [5] (see Section 2.2). Prior this step though, we nullify the subject's head motion that gradually translates and rotates the nostril region over time, introducing artifacts in the spatiotemporal visualization. We correct the motion error by registering the perinasal region in every frame to a global reference frame. We use the FFT-based phase correlation algorithm proposed by Reddy *et.al* [6] for that purpose. The main advantage of this algorithm is that it computes any amount of translation and rotation in fixed time for images of the same size. The algorithm has been designed for visual images where low-level features, such as object boundaries, are clearly distinguishable. Boundaries in thermal images, however, are fuzzy because of thermal diffusion. For this reason we support Reddy's algorithm with a Laplacian boundary enhancement function.

Translation. Let f represent the thermal image of the perinasal region (see the insets of Fig 1). Let also the perinasal image f_1 be translated by the vector (x_0, y_0) , producing the image f_2 (i.e., $f_2(x, y) = f_1(x - x_0, y - y_0)$). As per the Fourier shift theorem, their corresponding Fourier transforms F_1 and F_2 are related via the following equation:

$$F_2(\xi, \eta) = e^{-j2\pi(\xi x_0 + \eta y_0)} F_1(\xi, \eta). \quad (1)$$

The cross-power spectrum of the images is defined as:

$$\frac{F_1(\xi, \eta) F_2^*(\xi, \eta)}{|F_1(\xi, \eta) F_2^*(\xi, \eta)|} = e^{-j2\pi(\xi x_0 + \eta y_0)}, \quad (2)$$

¹ Spatiotemporal visualization clips of breathing function from the experimental set can be accessed at: <http://www.cpl.uh.edu/miccai-2012/>

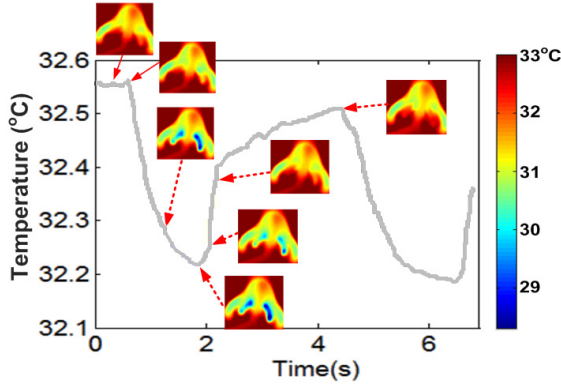


Fig. 1. Spatiotemporal heat distribution at the nostril region

where F^* is the complex conjugate of F . Taking the inverse Fourier transform of this representation will yield an impulse function at the displacement (x_0, y_0) and zero everywhere else.

Rotation. Let the image f_1 be translated by the vector (x_0, y_0) and rotated by angle θ_0 , producing the image f_2 (i.e., $f_2(x, y) = f_1(x \cos \theta_0 + y \sin \theta_0 - x_0, -x \sin \theta_0 + y \cos \theta_0 - y_0)$). According to the Fourier translation and rotation property, their Fourier transforms F_1 and F_2 are related by:

$$F_2(\xi, \eta) = e^{-j2\pi(\xi x_0 + \eta y_0)} F_1(\xi \cos \theta_0 + \eta \sin \theta_0, -\xi \sin \theta_0 + \eta \cos \theta_0). \quad (3)$$

The rotation without translation can be represented as a translation displacement in polar coordinates. Using phase correlation, one can find the angle θ_0 easily. Let us denote f'_2 to be the motion corrected image.

2.2 Nostril Segmentation

In this step we localize the left and right nostrils inside the perinasal region of interest. The nostril region features a non-uniform heat distribution of the breathing function which evolves over time. This dynamic nature of the spatiotemporal heat distribution poses a modeling challenge to the segmentation task. In particular, the nostrils' temperature elevates during the breathing expiration phase because the expired air absorbs heat in the lungs and respiratory passageways. In contrast, their temperature lowers during the breathing inspiration phase. We have adopted the active contour model framework because it is suitable for tracking this kind of spatiotemporal dynamic behavior [7]. In particular, we use the Chan-Vese active contour without edge modeling algorithm that is appropriate for fuzzy thermal boundaries [5].

Let f'_1, f'_2, \dots, f'_n be continuous thermal images of the perinasal region that are corrected for motion error as discussed in Section 2.1. Given a feature vector \mathbf{I} , the Chan-Vese active contour model is defined by:

$$\begin{aligned}
\frac{\partial \phi}{\partial t} = & \delta_\epsilon(\phi) \left[\mu \nabla \cdot \left(\frac{\nabla \phi}{\|\nabla \phi\|} \right) \right. \\
& - \frac{1}{N} \sum_{i=1}^N \lambda_i^+ (I_i(x, y) - c_i^+)^2 \\
& + \frac{1}{N} \sum_{i=1}^N \lambda_i^- (I_i(x, y) - c_i^-)^2 \\
& \left. - v \right],
\end{aligned} \tag{4}$$

where c^+ and c^- are mean values of regions inside and outside the evolving curve and N is the length of the feature vector. λ_i^+ and λ_i^- are the scaling parameters. μ controls the smoothness of the contour. v refines the level of the contour.

We construct a feature vector of size $N = 2$ that includes the pixel temperature in the current frame and temporal variation of pixel temperature in previous frames. Specifically, for a pixel at location (x, y) in frame n , the feature vector $\mathbf{I}_n(x, y) = (f'_n(x, y), \alpha \sigma_n^2(x, y))$, where α is the weight that we set dynamically using an exponential decaying function [8]. The decay constant of the function is set to 0.0743, which takes into account the temperature variation of half the normal breathing cycle (at the data acquisition rate of 25 frames per second). This arrangement guarantees that the contour evolution in the current frame takes into account the most recent temporal changes due to breathing.

2.3 Spatiotemporal Construction

Outcomes from the segmentation step are collected into a 3D array. Each X-Y plane of the array stores the segmented nostril region. The planes are stacked along the Z-axis of the array. To comply with sleep study metrics, every 30 seconds of data (i.e., 1 epoch), is stacked into one array. Finally, the 3D sets are supplied as a time-series model to the Avizo 6.2 software visualization tool.

3 Validation Analysis

3.1 Registration Validation

Thermal imaging records both physical motion (e.g., head motion) and physiological process (e.g., breathing). To find out how well our registration algorithm corrects the motion error in presence of breathing, we performed a simulation study. To build the simulation we used as a reference frame the thermal image of a subject's face. Within the reference frame we modulated the nostrils' temperature by applying a spatiotemporal evolution of a normal breathing cycle. We generated 20 simulated images equidistantly spaced across the breathing cycle. Each image was then translated and rotated by randomly generated values $((x_0, y_0), \theta_0)$. The values were treated as ground truth. The transformed images were supplied to our registration algorithm for motion correction. The resultant

images were compared with the corresponding reference frames for qualitative analysis (see Fig. 2). The translation and rotation parameters estimated through the algorithm were compared against their respective ground truth values. The ANOVA test performed on the data shows that there no statistically significant difference between the estimated translations and their ground truth values ($P > 0.01$) as well as the estimated rotations and their ground truth values ($P > 0.01$). This validates the registration algorithm.

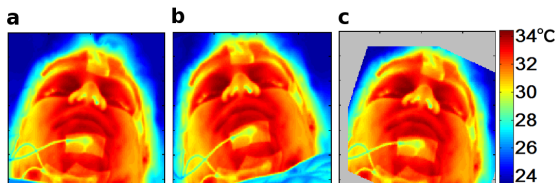


Fig. 2. (a) Reference image. (b) Simulated image. (c) Registered image. The orientation and position of the registered image is in agreement with that of the reference image.

3.2 Segmentation Validation

In a heterogeneous region, such as the perinasal region in the thermal imagery, a zero-level contour of convergence ϕ often leads to suboptimal delineation of the region of interest. Therefore, it is required to train the regulating parameter v to achieve optimal delineation.

Training the Parameter v : We used manual segmentations of the nostril region and a probabilistic scoring mechanism for training the parameter. In particular, three experts were asked to manually delineate in the thermal images the breathing evolution inside the left and right nostrils. A total of six sets of the thermal images were prepared from the six subjects. Each set comprised of 100 consecutive thermal images that represented 2-3 normal breathing cycles. Each expert repeated the delineation task twice per set. Thus, we acquired a total of six ground truth sets of manual segmentation per subject.

The performance of the segmentation algorithm was assessed against these ground truth sets by computing the Probabilistic Rand Index (PRI) [10]. PRI finds a common agreement between the multiple ground truth values and the segmentation output. Higher PRI indicates better performance of the segmentation algorithm. For every thermal image we varied v from 0 to 4 in steps of 0.05 and computed the PRI for each v value. The v value corresponding to the highest PRI was recorded as the optimal value for that particular image. The process was repeated for all 100 images in each set. Their optimal v values were averaged and designated as the tuned parameter v_t for that set. Corresponding PRI values were also averaged and recorded for comparison with the testing dataset. (see column-1 in Table 1).

Testing the Trained Parameter: The trained parameter was tested on different sets of manual segmentations generated by three experts different from

the one used in the parameter training. This time however, the six sets were prepared from randomly selected 100 images from the thermal videos. We supplied every set of images along with its tuned ν value to the segmentation algorithm and computed the PRI.

Table 1 summarizes the mean and standard deviation of the PRI values for every subject. The ANOVA test performed on the data concludes that there are no statistically significant difference between the PRIs of training and testing images ($P > 0.01$). This confirms that the regulating parameter ν was trained optimally for every subject.

Table 1. Mean and Std. of PRI for training and testing sets

	$\mu(PRI)$ (Training)	$\mu(PRI)$ (Testing)	$\sigma(PRI)$ (Training)	$\sigma(PRI)$ (Testing)
Subject 1	0.88	0.78	0.07	0.13
Subject 2	0.88	0.82	0.04	0.05
Subject 3	0.95	0.96	0.05	0.05
Subject 4	0.92	0.96	0.03	0.05
Subject 5	0.90	0.85	0.02	0.09
Subject 6	0.87	0.76	0.05	0.13

4 Experiments

4.1 Experimental Setup

The experiment was conducted in a controlled room environment at a Sleep Research Center. Six subjects (1 female and 5 males) participated in the experiment. The mean age of the subject pool was 25 ± 1.86 years. The experiment lasted 45 minutes. During the experiment the subjects were fitted with the standard polysomnography sensors to ground truth the imaging measurements. The subjects were asked to lay prostrate in a comfortable bed for the experiment period. They were positioned 2.5 m away from a thermal imaging system focused on their faces (see Fig. 2). The thermal imaging system consisted of a Thermo Vision SC6000 Mid-Wave Infrared (MWIR) camera from FLIR [9], a MWIR 100 mm lens, and a HP Pavilion m9040n desktop.

4.2 Experimental Results

Fig. 3 illustrates the spatiotemporal reconstruction of the breathing function. The 3D clouds in the figure represent the inhalation phases. The clouds' inhomogeneous colors depict the nonuniformity of the breathing function. In particular, the core of each cloud has marginally higher temperature than the peripheral region. The temperature gradient exists because the nasal cavities get narrower as they run from the mandibular to the periorbital region. This progressive stenosis amplifies heat convection which elevates the core's temperature. The gap in between two clouds represents the exhalation phase that our algorithm is unable

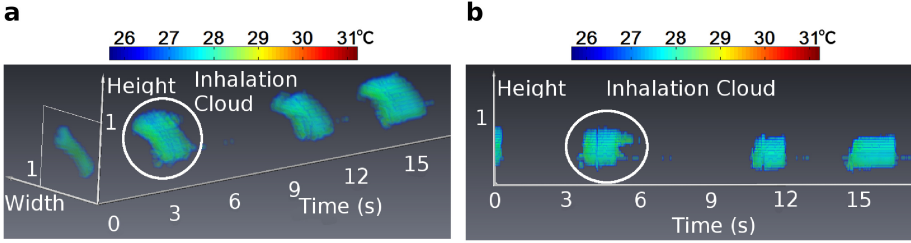


Fig. 3. The plot illustrates an angular view (a) and a projection view (b) of Subject 1's breathing function

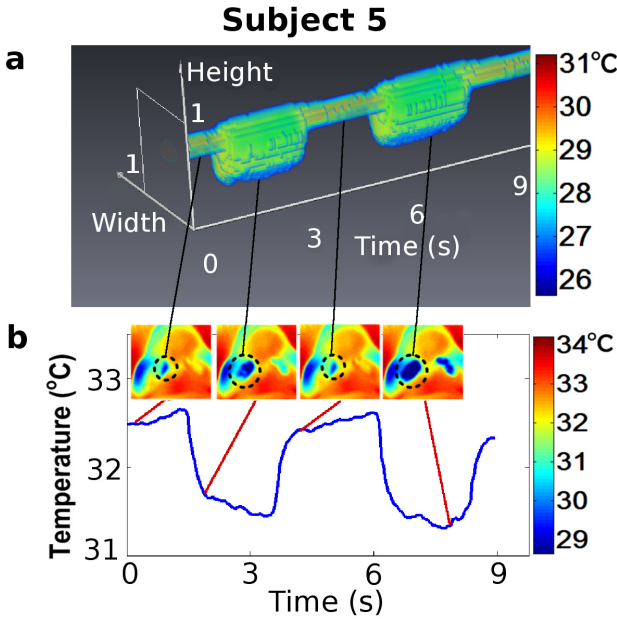


Fig. 4. (a) Angular view of the multi-dimensional visualization. (b) 1D breathing waveform. Conditions of reduced air-flow during exhalation on the upper part of the right nostril are evident by the bridging 'pipe' between the successive exhalation clouds. This spatiotemporal pattern is lost in the 1D breathing signal.

to capture at this point. This happens because the hot air from the exhalation phase has similar thermal profile as the nostril cartilage. In the future we plan on applying probabilistic methods, such as Bayesian classification, to separate the more dynamic breathing function from the relatively stationary thermal profile of the nostril cartilage.

An advantage of the spatiotemporal reconstruction of the breathing function over the mean temperature waveform [1][2] is the localization of subtle patho-

logical patterns. These patterns are usually obliterated by the averaging process at work in the 1D-breathing signal formation. Fig. 4(a) illustrates reduced flow during exhalation on the upper part of Subject 5's right nostril. This is evident by a bridging 'pipe' in the visualization that connects the two inhalation clouds and reveals that in the upper part of the nasal cavity, inhalation conditions persist even during exhalation. This finding is all but obscured if one considers Fig. 4(b) that shows the classic 1D breathing waveform, where the breathing waveform appears normal.

Acknowledgments. This material is based upon work supported by a National Science Foundation (NSF) grant (# IIS-1049004) titled 'EAGER: Improving Human Engagement and Enjoyment in Routine Activities.' We are grateful to Dr. Max Hirshkowitz, and Dr. Amir Sharafkhaneh from the Baylor Sleep Research Center of the Veterans Affairs Hospital for their help.

References

1. Murthy, J., Van Jaarsveld, J., Fei, J., Pavlidis, I., Harrykissoon, R., Lucke, J., Faiz, S., Castriotta, R.: Thermal Infrared Imaging: A Novel Method to Monitor Airflow During Polysomnography. *Sleep* 32(11), 1521–1527 (2009)
2. Fei, J., Pavlidis, I.: Thermistor at a Distance: Unobtrusive Measurement of Breathing. *IEEE Transactions on Biomedical Engineering* 57(4), 988–998 (2010)
3. Rappai, M., Collop, N., Kemp, S., de Shazo, R.: The Nose and Sleep-disordered Breathing: What We Know and What We Do Not Know. *Chest* 124(6), 2309–2323 (2003)
4. Farre, R., Montserrat, J.M., Rotger, M., Ballester, E., Navajas, D.: Accuracy of Thermistors and Thermocouples as Flow-measuring Devices for Detecting Hypopnoeas. *European Respiratory Journal* 11(1), 179–182 (1998)
5. Chan, T.F., Vese, L.A.: Active Contours Without Edges. *IEEE Transactions on Image Processing* 10(2), 266–277 (2001)
6. Reddy, B., Chatterji, B.: An FFT-based Technique for Translation, Rotation, and Scale-invariant Image Registration. *IEEE Transactions on Image Processing* 5(8), 1266–1271 (1996)
7. Kass, M., Witkin, W., Terzopoulos, D.: Active Contour Models. *International Journal of Computer Vision* 1(4), 321–331 (1988)
8. Nowlan, S.: Soft Competitive Adaptation: Neural Network Learning Algorithms Based on Fitting Statistical Mixtures. PhD thesis, School of Computer Science, Carnegie Mellon University, Pittsburgh (1991)
9. FLIR Systems, <http://www.flir.com>
10. Unnikrishnan, R., Pantofaru, C., Hebert, M.: Toward Objective Evaluation of Image Segmentation Algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29(1), 929–944 (2007)