

# Journal of Biomedical Optics

BiomedicalOptics.SPIEDigitalLibrary.org

## **Synergetic use of thermal and visible imaging techniques for contactless and unobtrusive breathing measurement**

Meng-Han Hu  
Guang-Tao Zhai  
Duo Li  
Ye-Zhao Fan  
Xiao-Hui Chen  
Xiao-Kang Yang

**SPIE.**

Meng-Han Hu, Guang-Tao Zhai, Duo Li, Ye-Zhao Fan, Xiao-Hui Chen, Xiao-Kang Yang, "Synergetic use of thermal and visible imaging techniques for contactless and unobtrusive breathing measurement," *J. Biomed. Opt.* **22**(3), 036006 (2017), doi: 10.1117/1.JBO.22.3.036006.

# Synergetic use of thermal and visible imaging techniques for contactless and unobtrusive breathing measurement

Meng-Han Hu, Guang-Tao Zhai,\* Duo Li, Ye-Zhao Fan, Xiao-Hui Chen, and Xiao-Kang Yang  
Shanghai Jiao Tong University, Institute of Image Communication and Information Processing, Shanghai, China

**Abstract.** We present a dual-mode imaging system operating on visible and long-wave infrared wavelengths for achieving the noncontact and nonobtrusive measurements of breathing rate and pattern, no matter whether the subjects use the nose and mouth simultaneously, alternately, or individually when they breathe. The improved classifiers in tandem with the biological characteristics outperformed the custom cascade classifiers using the Viola–Jones algorithm for the cross-spectrum detection of face and nose as well as mouth. In terms of breathing rate estimation, the results obtained by this system were verified to be consistent with those measured by reference method via the Bland–Altman plot with 95% limits of agreement from  $-2.998$  to  $2.391$  and linear correlation analysis with a correlation coefficient of  $0.971$ , indicating that this method was acceptable for the quantitative analysis of breathing. In addition, the breathing waveforms extracted by the dual-mode imaging system were basically the same as the corresponding standard breathing sequences. Since the validation experiments were conducted under challenging conditions, such as the significant positional and abrupt physiological variations, we stated that this dual-mode imaging system utilizing the respective advantages of RGB and thermal cameras was a promising breathing measurement tool for residential care and clinical applications. © The Authors. Published by SPIE under a Creative Commons Attribution 3.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/1.JBO.22.3.036006](https://doi.org/10.1117/1.JBO.22.3.036006)]

Keywords: thermal imaging; dual-mode imaging; respiration measurement; noncontact techniques.

Paper 160731R received Oct. 27, 2016; accepted for publication Feb. 21, 2017; published online Mar. 6, 2017.

## 1 Introduction

Breathing rate, along with blood oxygen saturation, heart rate, and blood pressure, is considered as one of four main vital physiological signs.<sup>1</sup> The breathing rate is 12 to 18 breaths per minute (bpm) for a healthy adult at rest,<sup>2</sup> whereas it will increase to the range of 35 to 40 bpm when this person is undergoing or has just done exercise.<sup>3</sup> Alterations in the breathing rate and pattern are known to occur with serious adverse events<sup>4</sup> or early clues of the pathology processes.<sup>5</sup> Some diseases such as sleep disorders cause abnormal breathing rhythms such as Kussmaul breathing.<sup>6</sup> Moreover, the observation of breathing plays a crucial role in many other applications and research, including sport studies,<sup>7</sup> quarantine, and security inspections.<sup>8,9</sup>

Current noninvasive breathing measurement approaches contain electrical impedance tomography, respiratory inductance plethysmography, capnography, tracheal sound measurement, spirometers, respiratory belt transducer, and electrocardiography-derived method.<sup>10–12</sup> Nonetheless, the above devices carry out breathing rate estimation in a contact way, which leads to discomfort, stress, and even to soreness of a subject.<sup>12</sup> Increasing daily and clinical demands for contactless and unobtrusive yet accurate breathing measurement alternatives in uncontrolled environments have spurred considerable interest among researchers on the application of innovative tools for breathing observation solutions. Doppler radar was used in the noncontact and through-clothing breathing evaluation via

the measurement of chest wall motion.<sup>13</sup> This method is yet limited by the potential radiation and high sensitivity to motion artifacts. A laser Doppler vibrometer determined the breathing rate by the assessment of the chest wall displacements;<sup>14</sup> however, its result will not be accurate when, for example, improper measurement points are selected on the thoracic surface. Min et al. developed an ultrasonic proximity sensing approach to measure breathing signatures by means of calculating time intervals between the transmitted and received sound waves during the abdominal wall fluctuation.<sup>15</sup> The subjects under this test are required to remain still and refrain any other movements. In addition, owing to the mature image processing techniques, visible imaging sensors have attracted much attention for breathing evaluations.<sup>16,17</sup> Shao et al. determined the breathing patterns using the cameras in the visible region to track the small shoulder movements associating with breathing.<sup>18</sup> Although the random body movements can be corrected by the motion-tracking algorithm, breathing rate estimation based on visible imaging is by nature sensitive to the slight movements, thus not being appropriate for the long-term monitoring.

Compared with the aforementioned active sensors, the passive thermal infrared imaging that records the emitted energy from the objects does not need any harmful radiation and light source.<sup>19</sup> The principle of thermal imaging for breathing estimation is based on the fact that the changes of temperature around the nostrils and mouth will accompany the inhalation and exhalation.<sup>20</sup> Temperature variation is, in contrast to the displacement change, more significant and thus more suitable for deriving breathing signature. Despite many face recognition algorithms in the visible band,<sup>21</sup> locating and tracking face and

\*Address all correspondence to: Guang-Tao Zhai, E-mail: [zhaiguangtao@sjtu.edu.cn](mailto:zhaiguangtao@sjtu.edu.cn); [zhaiguangtao@gmail.com](mailto:zhaiguangtao@gmail.com)

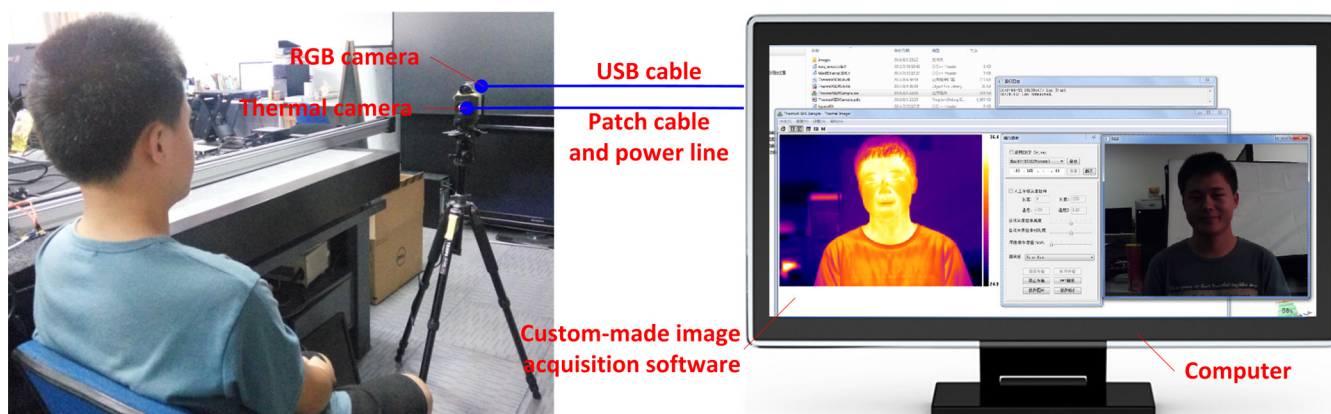


Fig. 1 Schematic illustration of experimental setup.

facial tissues in the thermal band are highly challenging due to the few geometric and textural details as well as the various physiological changes in a thermal image. Basu et al. manually selected the nasal area and it was afterward tracked by the corner detection in conjunction with the registration process.<sup>22</sup> The hyperventilation was therefore successfully monitored by the thresholding technique. Abbas et al. extracted the breathing signals from manually-selected region in thermal images, and the good performance of the proposed method had been shown on the breathing measurements of eight adults and one infant.<sup>23</sup> Several investigators applied the template matching method to track the predefined nasal tissues in thermal infrared images.<sup>24,25</sup> Some literature has attempted to automatically identify the nasal region. Fei and Pavlidis first determined the nasal contour by the use of horizontal and vertical projection profiles in spatial dimension, and subsequently, the nostril regions were found by taking the temporal variances into account.<sup>2</sup> A retained boosted cascade classifier based on the temperature feature was utilized to detect the nasal cavities,<sup>26</sup> while the classification accuracy seemed to be unacceptable for the purpose of breathing measurement. The salient physical features of the human face in a thermal image were used to segment the nasal region;<sup>27</sup> however, it will not succeed when wearing glasses or conducting open-mouth respiration. Moreover, the direct identifications of facial tissues in thermal images are camera-dependent and extremely interfered with by the abrupt physiological changes such as perspiration. To face these challenges, the cross-spectrum face and facial tissue recognitions may provide the possibility to locate and track regions of interest exactly in thermal images.

The objectives of the current study are to: (1) establish and register the thermal and visible dual-mode imaging system; (2) develop a cross-spectrum face and facial tissue recognition algorithm for long-wave infrared and visible bands and obtain the temperature variation signal; and (3) validate the dual-mode imaging system and the proposed algorithm for breathing rate and pattern measurements.

## 2 Methods

The thermal imager for breathing rate and pattern measurement is based on the fact that the temperature around the nose and mouth fluctuates throughout the inspiration and expiration cycles. The disadvantage compared to the RGB image is that, due to few geometric and textural facial details, the thermal image is at present inadequate to design fast and reliable face

detection algorithms.<sup>28</sup> Therefore, in this study, the visible imaging technique is adopted to aid in the automatic recognitions of face and facial tissue in thermal images.

The steps of the development of a dual-mode imaging system, image registration, detection of face, and its tissue in two spectral domains, region of interest (ROI) tracking, computation of breathing signal, and validation experiment are elaborated in the following sections.

### 2.1 Thermal and Visible Dual-Mode Imaging System

The experimental setup is shown in Fig. 1. A thermal imager (MAG62, Magnity Electronics Co. Ltd., Shanghai, China) with the resolution of  $640 \times 480$  and the pixel pitch of  $17 \mu\text{m}$  is stabilized on a tripod to prevent vibration during the experiments. The spectral range and thermal sensitivity of thermal imager are  $7.5$  to  $14 \mu\text{m}$  and  $0.5^\circ\text{C}$ , respectively. An RGB camera (USBFHD01M, RERVISION Co. Ltd., Shenzhen, China) with the resolution of  $640 \times 480$  is fixed on the top of the thermal imager. These two cameras are parallel to each other in such a way that the field of view is almost the same. To connect a computer, a USB and patch cables are used for RGB and thermal cameras, respectively. A custom-made image acquisition software is executed to generate two trigger signals, thus allowing the simultaneous acquisition of thermal and visible videos. The recorded videos are then imported into the MATLAB<sup>®</sup> R2014a (The Mathworks, Inc., Natick, Massachusetts) for further analysis.

### 2.2 Image Registration

After establishing the dual-mode imaging system, the affine transformation is required to register thermal and visible images.<sup>29</sup> The first step is to select the strongly correlated points in the first frame of bimodal videos so as to define the control point pairs, viz., the fixed points in the thermal image and the moving points in the RGB image. Subsequently, these points are adjusted by the cross-correlation to obtain the transformation matrix. We align each frame from RGB video according to

$$I_{\text{visr}} = I_{\text{vis}} \mathbf{T}, \quad \text{where } \mathbf{T} = \begin{bmatrix} s \cos \theta & s \sin \theta & 0 \\ -s \sin \theta & s \cos \theta & 0 \\ \mathbf{b}_x & \mathbf{b}_y & 1 \end{bmatrix}, \quad (1)$$

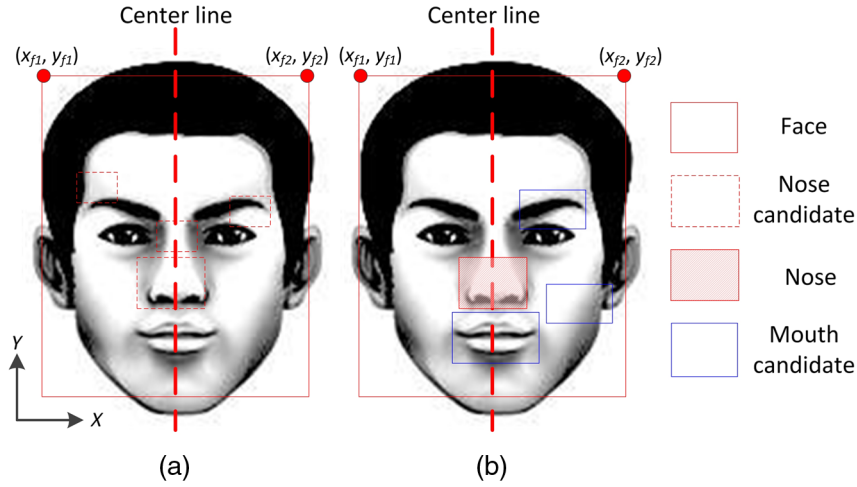


Fig. 2 Illustrations of the further screening of (a) nose and (b) mouth regions.

where  $I_{\text{vis}}$  is the original RGB image and its corresponding transformed RGB image is  $I_{\text{vis}r}$ ;  $\mathbf{T}$  represents the transformation matrix;  $\mathbf{s}$ ,  $\boldsymbol{\theta}$ , and  $\mathbf{b}$  denote the scaling, rotation, and translation vectors, respectively.

The row and column of the registered RGB images are resized to be equal to those of thermal images.

### 2.3 Cross-Spectrum Face and Facial Tissue Detection

The cascade object detector using the Viola–Jones algorithm<sup>30</sup> coupled with the screening technique based on biological characteristics was used to detect face and nose as well as mouth in the RGB image, and subsequently, the linear coordinate mapping was conducted to determine the corresponding regions in the thermal image.

The Haar-like features are extracted from the integral images and afterward served as the input of the custom cascade classifier. This algorithm can be summarized as follows.<sup>30</sup>

Let us assume that there is a dataset  $U = \{x_1, \dots, x_N\}$ , and each data sample  $x_i \in U$  carries a label variable  $y_i \in Y (Y = \{1, -1\})$ , where  $i = 1, \dots, N$ . Hence, the initial distribution for the samples in training set can be represented as  $D_1(i) = 1/N$ . For every weak classifier,  $h_t: U \rightarrow Y$ , the error of distribution  $D_t$  can be denoted as  $\epsilon_t = P_{D_t}[h_t(x_i)] \neq y_i$  and therefore the weight of the weak classifier as  $\alpha_t = 0.5 \ln[(1 - \epsilon_t)/\epsilon_t]$ , where  $t = 1, \dots, T$ .  $T$  is the number of weak classifiers. The final strong classifier is

$$H_{\text{final}}(x) = \text{sign} \left[ \sum_{t=1}^T \alpha_t h_t(x) \right], \quad (2)$$

where  $H_{\text{final}}$  represents the final strong classifier.

When more than one region is considered as a face using the custom Viola–Jones algorithm, we design the algorithm for searching the facial tissues such as nose in these regions and the region inclusive of facial contents is chosen as the real face. Once the face position has been confirmed, the above procedure will be repeated within the face region to detect the locations of nose and mouth. Nevertheless, several potential nose and mouth regions may be found by the custom cascade classifiers. To solve this problem, the biological characteristic of nose that locates on the center line of face is utilized [Fig. 2(a)].

The minimum distance between center lines of face and nose candidate regions is calculated by

$$n_{\text{final}} = \arg \min_{n, n'=1, \dots, k} \left( \left| \frac{x_{f1} + x_{f2} - x_n - x_{n'}}{2} \right| \right). \quad (3)$$

In the equation,  $n_{\text{final}}$  is the final nose region;  $x_{f1}$  and  $x_{f2}$  are the horizontal ordinates of two corners of the face region's top side;  $x_n$  and  $x_{n'}$  are the horizontal ordinates of two corners of the top side of the nose region;  $k$  is the number of nose candidate regions.

If there still exist several nose candidate regions, Eq. (4) is applied to find the largest nose candidate region as the real nose region

$$n_{\text{final}} = \arg \max_{n, n'=1, \dots, k} (|x_n - x_{n'}|). \quad (4)$$

In the case of further screening of the mouth region, due to the biological characteristics of facial tissues, the vertical ordinate of mouth should be smaller than that of nose. Simultaneously, the horizontal ordinate of mouth should be near to that of mouth [Fig. 2(b)]. This step can be expressed by

$$m_{\text{final}} = \arg \min_{m=1, \dots, i} \left[ \left\| \begin{pmatrix} x_n \\ y_n \end{pmatrix} - \begin{pmatrix} x_m \\ y_m + \alpha h \end{pmatrix} \right\|_2 \right] \quad (5)$$

subject to  $y_m < y_n$ ,

where  $m_{\text{final}}$  is the final mouth region;  $(x_n, y_n)$  and  $(x_m, y_m)$  are the coordinates of the confirmed nose region and mouth candidate regions, respectively;  $m$  denotes the number of mouth candidate regions;  $h$  is the distance between the mouth and nose;  $\alpha \in [0, 1]$  is the arbitrary value defined by the priori knowledge.

The algorithm of searching mouth is further refined by introduction of Eq. (4) to eliminate the small interfering blocks near to the center line.

Later, the corresponding positions of nose and mouth can be automatically found in the thermal images via the linear coordinate mapping.



## 2.4 ROI Tracking

The Shi–Tomasi corner detection algorithm<sup>31</sup> derived from the Harris–Stephens method<sup>32</sup> is applied to extract the interest points from nose and mouth ROIs in the visible gray images. For each pixel in input image, the covariance matrix  $\mathbf{M}$  corresponding to its neighborhood  $S(p)$  is

$$\mathbf{M} = \sum_{x,y \in S(p)} w(x,y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix}, \quad (6)$$

where  $w(x,y)$  represents the given feature window;  $I_x$  and  $I_y$  are the differences of  $x$  and  $y$  directions, respectively.

The strongest key features  $C_s$  is calculated by

$$C_s = ((i,j) | \min[\det(\mathbf{M})] \geq k \max_{i,j \in I} \{\min[\det(A)]\}), \quad (7)$$

where  $\mathbf{M}$  is the covariance matrix of the pixel to be detected;  $\mathbf{A}$  denotes the vector containing the covariance matrices in input image  $I(i,j)$  preprocessed by a Gaussian filter, and  $k$  is the empirical constant for tuning the threshold (here is 0.01).

Next, the ROI is tracked via the Kanade–Lucas–Tomasi algorithm<sup>33,34</sup>

$$\varepsilon(\mathbf{X}) = \iint_w [A(x - \Delta x, y - \Delta y, t) - B(x, y, t + \tau)] \omega(\mathbf{X}) d\mathbf{X}, \quad (8)$$

where  $\varepsilon$  is the sum of squared intensity difference between the local image model  $A$  at the current time  $t$  and local image model  $B$  at time  $t + \tau$ ;  $\Delta x$  and  $\Delta y$  are the displacements in the  $x$  and  $y$  directions, respectively;  $\mathbf{X}$  is the vector including the displacement and time variables;  $W$  is the given window and  $\omega$  is the weighting function (here is 1).

Based on the above equation, the tracking of ROI in video sequences can be realized by the use of the displacement  $(\Delta x, \Delta y)$ , which is determined by minimizing the  $\varepsilon$ . Furthermore, the tracking procedure is refined by the forward-backward error.<sup>35</sup> This method invalidates the tracked points if their errors exceed the setting value, thus enabling the selection of more reliable trajectories among the consecutive frames. In this study, the threshold is set as 2 pixels. The cross-spectrum ROI tracking is achieved by the linear coordinate mapping.

## 2.5 Extraction of Breathing Signature and Pattern

Because the shape of original ROI may change from the rectangle to the polygon during the tracking operation, the equation listed below is available to acquire the average pixel intensity  $\bar{s}(k)$  within the ROI of thermal image

$$\bar{s}(k) = \frac{1}{n} \sum_{i,j \in \mathbf{N}} s(i,j,k), \quad (9)$$

where  $s(i,j,k)$  is the pixel intensity of thermal image at pixel  $(i,j)$  and video frame  $k$ ;  $\mathbf{N}$  is the vector of pixel coordinates in ROIs and  $n$  is its number.

The raw pixel intensities of ROIs in all the frames are smoothed by the moving average filter with the data span of 5. The abrupt change in the breathing waveform can be eliminated by the above smoothing method. The breathing rate and

pattern as a function of video frame can be therefore simply estimated from the smoothed intensity data by computing the number of the obvious peaks. Considering both the speed and purposes of the analysis, it is not necessary to carry out the unit conversion to make the breathing signature a function of measurement time.

## 2.6 Validation of Proposed Method

To evaluate the performance of the cascade classifier using the Viola–Jones algorithm in tandem with the biological characteristic screening for the cross-spectrum face and facial tissue detection, a total of 66 image pairs cross the two domains, collected from 11 volunteers aged 23 to 28 under various breathing conditions, were used for the validation experiment through comparing the detection accuracy to that of the custom cascade classifier.

A database of thermal and visible dual-mode videos is constructed to quantitatively and qualitatively verify the proposed cross-modal breathing measurement method. The videos were captured with the frame rate of 30 frames per second under the uncontrolled illumination and room temperature. All volunteers involved in the experiments consented to be subjects, and were instructed to breathe using the nose and mouth simultaneously, alternately, or individually. Moreover, in the current work, the different breathing situations, such as the translations and rotations of body and the variations of facial expression when laughing, yawning, and speaking, were allowed (and even encouraged) during the experiments to guarantee that there are larger variations in the obtained videos. The distance between the volunteer and cameras is about 150 cm.

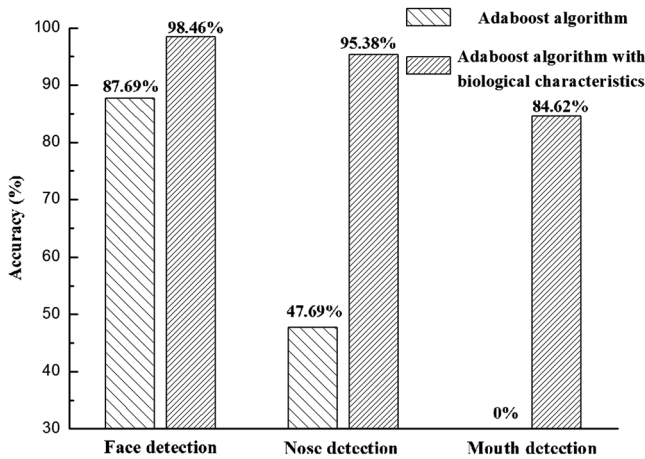
For the quantitative validation, the volunteers were asked to breathe at their own pace for 1 min, and this procedure was repeated six times for each volunteer. At the same time, the reference breathing rate was recorded by two dedicated and qualified human observers. The Bland–Altman plot<sup>36,37</sup> and linear correlation analysis were used to check the effectiveness of our approach.

In terms of qualitative testing, the volunteers were required to complete two intended breathing sequences: (I) eupnea (normal breathing) and tachypnea, followed by apnea and Kussmaul breathing (deep breathing); and (II) Cheyne–Stokes respiration,<sup>38</sup> which is an abnormal periodic breathing pattern containing the progressively deeper and sometimes faster breathing, followed by the gradual decrease and temporary apnea at the bottom of Fig. 8. The recorded breathing sequences extracted using a dual-mode imaging system were visually compared with the intended breathing patterns.

## 3 Results and Discussion

### 3.1 Detection Accuracy of Face and Facial Tissue

Prior to the breathing signature extraction, the locations of face and facial tissue should be determined in the RGB and thermal images. We introduced the biological characteristics into the classification framework, aiming to attain higher accuracy than the custom cascade classifier using the Viola–Jones algorithm. The detection accuracies of face and nose as well as mouth are shown in Fig. 3. Overall, compared to the custom cascade classifier, the modified classifier gave the relatively good performance for face detection (98.46% versus 87.69%). The accuracies had been remarkably improved for detection of



**Fig. 3** Histogram of the detection accuracies of face and nose as well as mouth using the Viola–Jones algorithm and the custom cascade classifier coupled with biological characteristics. (Here, if a detected object is not uniquely contained in a return result, it is considered to be misclassified.)

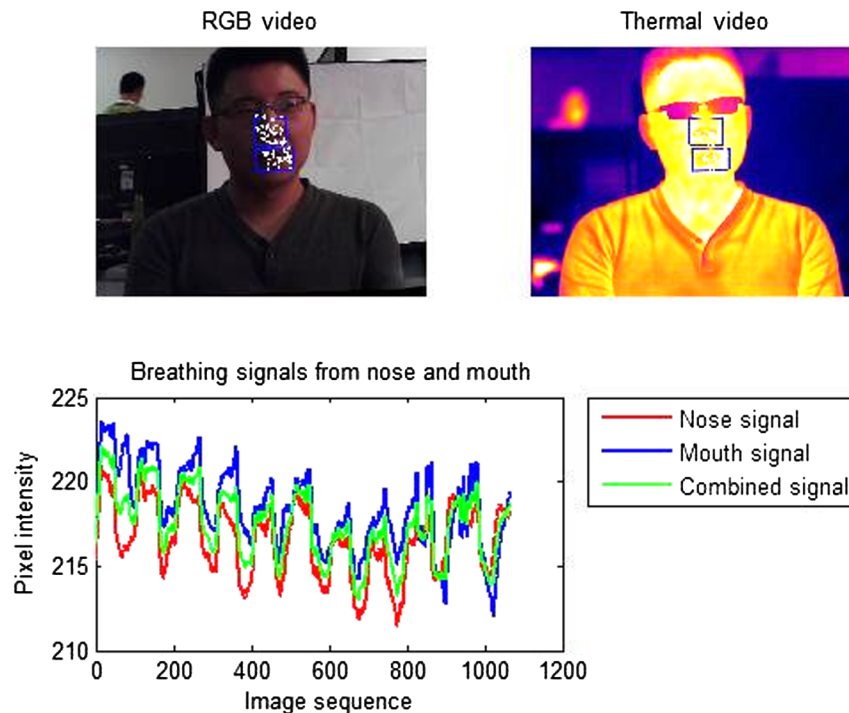
nose and mouth separately from 47.69% to 95.38% and 0% to 84.62%.

For breathing measurement, previous literature on detecting the face and its tissue usually applied image processing and analysis in thermal images directly. Pereira et al. segmented the thermal image by the multilevel Otsu’s method and identified the largest area of the remaining regions in the binary image as the real face region.<sup>12</sup> The human anatomy and physiology that limited the nose search window in the hottest region

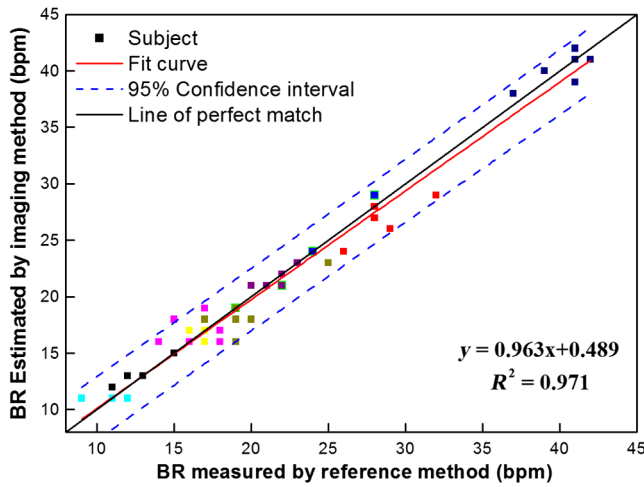
were also utilized by them to locate the nose. However, this algorithm will fail when the other large or hot objects appear in the thermal image. Deepika et al. implemented the thresholding operation in the green component of thermal images to extract the nose region.<sup>39</sup> This method cannot work if the subject breathes through the mouth. Despite being advantageous over the single mode imaging, the visible-thermal imaging system had been scarcely reported for recognition of face and facial tissue in breathing estimation applications,<sup>28,40</sup> due in part to the relatively complicated imaging architecture and data processing. Although there exist more effective and efficient approaches to find the face region in the face detection domain,<sup>41,42</sup> the proposed face and facial tissue detection method can achieve the acceptable accuracy for breathing measurement using triple coordinate calculation operations based on the traditional algorithm, thus having met the objectives of the current study. Consequently, considering the results and discussion mentioned above, we state that the visible and thermal dual-mode imaging framework and related algorithm in this study offer an alternative or complementary solution to face and nose as well as mouth detection in breathing research.

### 3.2 Validation of Breathing Rate Measurement

Figure 4 shows the breathing signature processing interface for a visible and thermal dual-mode imaging system. This screenshot was one frame extracted from the short video series in the video, which was an attempt to illustrate the robustness of our system and the corresponding algorithms. As shown in the video, the imaging system on visible and long-wave infrared wavelengths, associated with the proposed object detection and tracking



**Fig. 4** Screenshot of the breathing signature processing interface for visible and thermal dual-mode imaging system. A short video illustrating the effectiveness of the presented system and algorithm is demonstrated in Videos 1–3. (Video 1, 6 MB, MOV [URL: <http://dx.doi.org/10.1117/1.JBO.22.3.036006.1>], Video 2, 3 MB, MOV [URL: <http://dx.doi.org/10.1117/1.JBO.22.3.036006.2>], and Video 3, 6 MB, MOV [URL: <http://dx.doi.org/10.1117/1.JBO.22.3.036006.3>].)

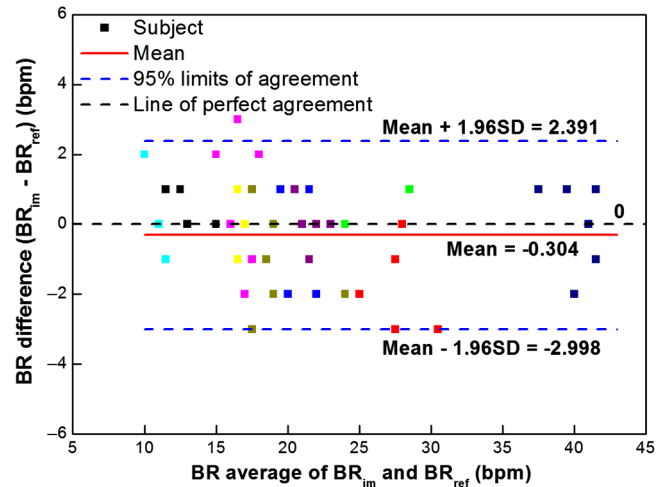


**Fig. 5** Correlation between the breathing rate measured from the reference method and thermal-visible dual-mode imaging system. (The different colors denote the different subjects.)

algorithms, was capable of following the ROIs regardless of the actions of the other persons such as walking into the field of view (Fig. 4). In addition, this system was robust against the translations and rotations of body (e.g., head) within the angle of 90 deg in the field of view, as well as the abrupt physiological variations (e.g., yawning, swallowing, and speaking). Hence, with the help of the RGB camera, the thermal imaging-based breathing measurement device can detect and track the nose and mouth accurately, thus being able to maximally avoid the erroneous measurement of breathing signature.

To test the performance of breathing measurement with the dual-mode imaging technique in a contactless and unobtrusive manner, the statistical analysis approaches viz., linear correlation analysis and Bland–Altman plot were used for the validation of data from the small-scale pilot experiment. The scatter plot and regression line of estimated and reference BR are shown in Fig. 5. As shown in Fig. 5, most of the scatter points were close to the line of perfect match (slope = 1) and within the 95% confidence intervals. By means of linear correlation analysis, the strong relationship ( $R^2 = 0.971$ ) between the simultaneously acquired measured and reference BR was found over the range from 9 to 42 bpm, indicating that the proposed method was acceptable for BR estimation.

The corresponding Bland–Altman plot in respect to two techniques is demonstrated in Fig. 6 with the mean of differences of  $-0.304$  bpm and limits of agreement of  $-2.998$  and  $2.391$  bpm. It could be observed that the majority of points were dispersed around the line of perfect agreement (BR differences = 0), and the 12 points approximately located on this line and were considered to be fully consistent. There was one point (magenta) out of the upper limits of agreement with the offset about 0.6 bpm. By checking the original video, we inferred that the reason for this might be that the testing subject conducted very significant and frequent as well as irregular body motion during the experiment. This is also illustrated in Video 3. Three points from two subjects (one in dark yellow and the others in red) approximately fell on the lower limits of agreement, perhaps because of the alternative use of mouth and nose and the changes of facial expression when breathing. The distribution of points in the Bland–Altman plot in Fig. 6 was to great extent similar to that of scatter points in Fig. 5. In general, the result of the



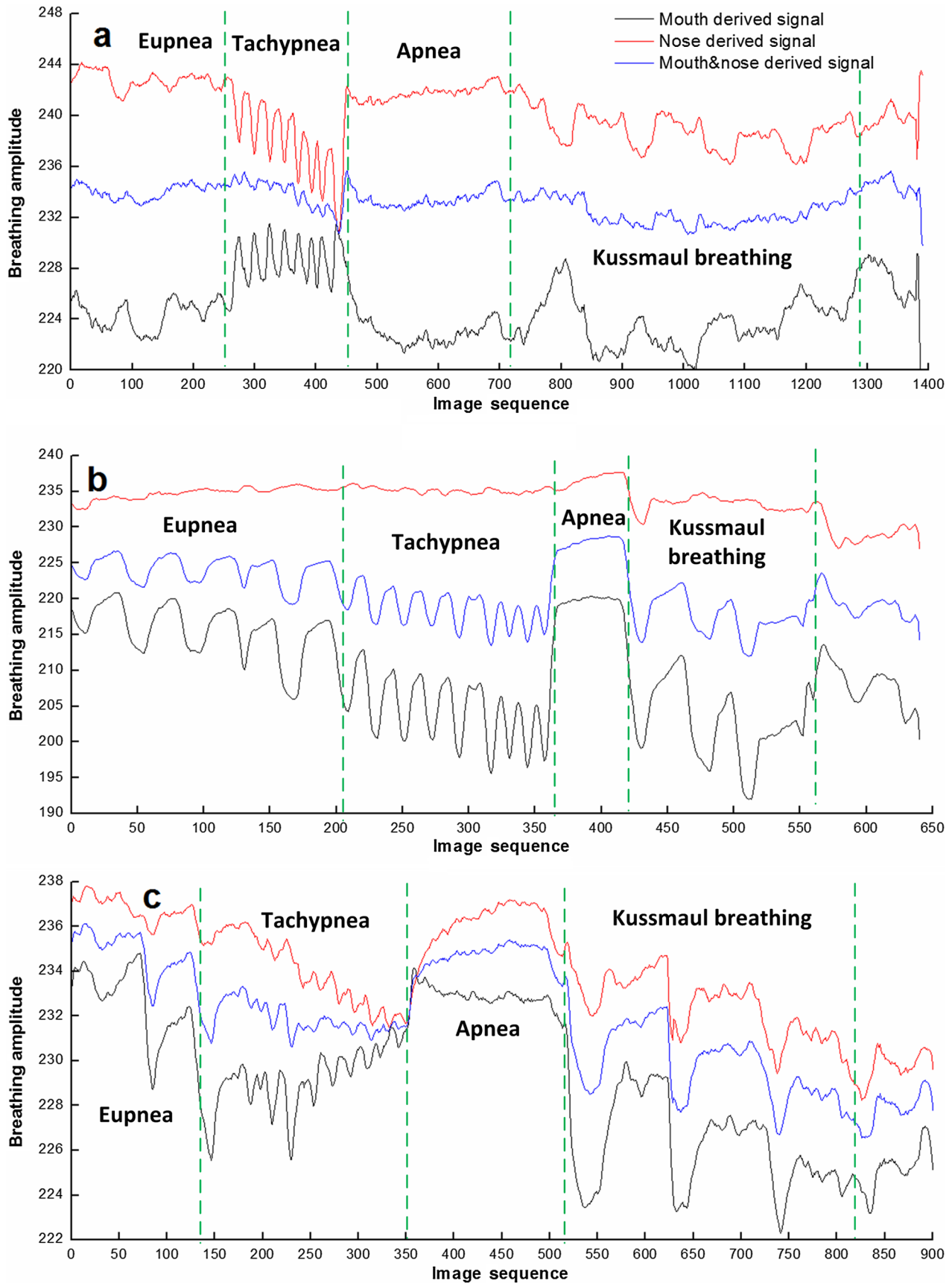
**Fig. 6** Bland–Altman plot ( $N = 56$  measurement times) of the difference against average for breathing rate using the reference method ( $BR_{ref}$ ) and thermal-visible dual-mode imaging system ( $BR_{im}$ ). (The different colors denote the different subjects and SD is the standard deviation.)

Bland–Altman plot demonstrated the feasibility of using the visible and thermal dual-mode imaging system in tandem with the proposed algorithm for the contactless and unobtrusive BR estimation.

A group of investigators measured the breathing rate by the application of dual RGB cameras installed in a smartphone.<sup>17,43</sup> They extracted the BR from the recorded chest movement signals, and the lower and upper limits of agreements were  $-0.850$  and  $0.802$  bpm, respectively. In the other research, the ranges of limits of agreement between  $-1.4$  and  $1.3$  bpm were obtained from the thermal image sequences, but the values had increased from  $-3.7$  to  $3.9$  bpm since the subjects followed the complex breathing profile.<sup>12</sup> The manually defined ROIs in the RGB and infrared images were selected for prompt infection screening at airports,<sup>8</sup> and the limits of agreements varied between  $-1.0$  and  $0.9$  bpm for the measurement of breathing rate. Compared to the published literature, though limits of agreement covered a relatively wider range, the dual-mode imaging system proved to be more immune to various variations for BR extraction via adding the camera operating at visible wavelengths. Notice that the imaging system coupled with the proposed algorithm that can minimize the BR measurement mistakes cannot eliminate the errors caused by a variety of uncontrolled variations.

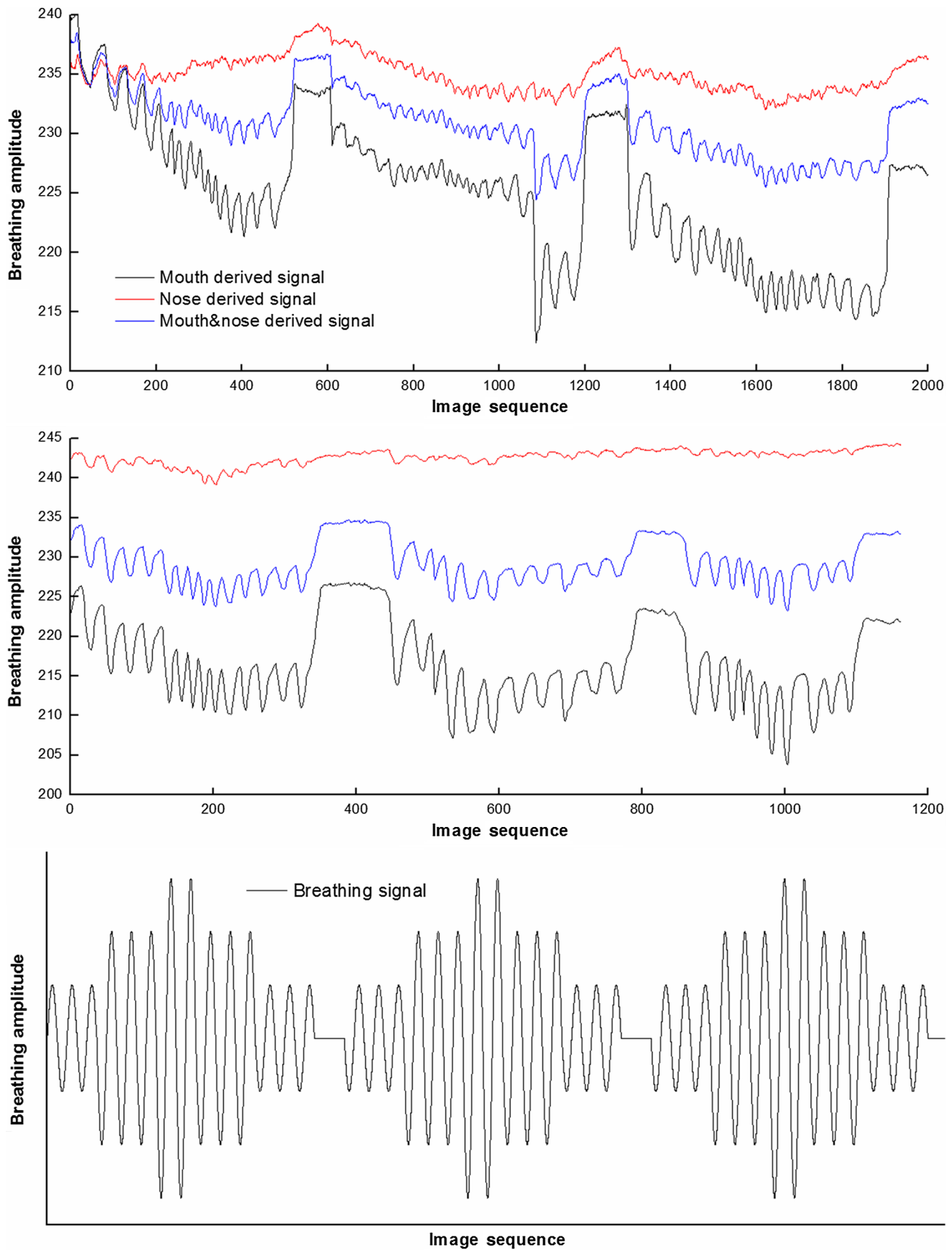
### 3.3 Validation of Breathing Pattern Measurement

The breathing pattern sequences (I) of three subjects, corresponding to the use of nose-dominated, mouth-dominated, and nose and mouth combined breathing manners, are shown in Fig. 7. According to the labels in Fig. 7, it could be intuitively observed that the waveforms from the dual-mode imaging system successfully reproduced the predefined breathing sequences, containing eupnea, tachypnea, apnea, and Kussmaul breathing patterns. In fact, some noise events existing in the extracted signature led to the distortion of waveforms, which might in turn cause the misclassifications of breathing patterns. These unwanted signatures mainly attributed to alternately breathing through mouth and nose, for example, in the case of nose-dominated breathing [Fig. 7(a)], the waveform in the eupnea phase was largely affected by occasional open-mouth

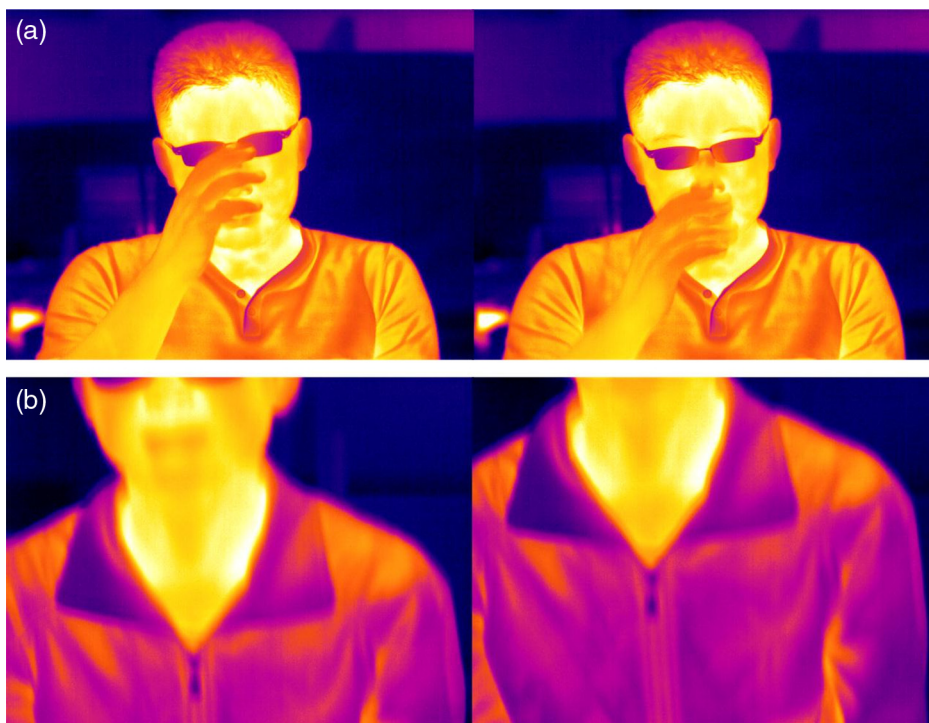


**Fig. 7** Breathing pattern sequences containing the eupnea, tachypnea, apnea, and Kussmaul breathings measured by the thermal-visible dual-mode imaging system: (a) nose-dominated breathing, (b) mouth-dominated breathing, and (c) nose and mouth breathing.





**Fig. 8** Two typical examples of the intended breathing pattern of Cheyne–Stokes respiration measured using the thermal-visible dual-mode imaging system. (The bottom is the standard Cheyne–Stokes breathing profile.)



**Fig. 9** Two failed breathing measurement cases due to the loss of tracked points of ROI caused by (a) adjusting the glasses and (b) moving out of plane.

breathing. Fortunately, we could still correctly classify the different breathing patterns in the obtained sequences.

For the sake of further validating the reliability of the proposed method to identify the breathing pattern, the more complex breathing sequence (II), called Cheyne–Stokes respiration, was applied in this study. Figure 8 exhibits the Cheyne–Stokes respiration sequences of two subjects and the corresponding standard profile. In Fig. 8, the top and middle subplots are the breathing sequences of two subjects breathing through nose and mouth simultaneously and mouth primarily, respectively. Overall, the Cheyne–Stokes breathing sequences obtained using the visible and thermal dual-mode imaging system were basically consistent with the standard profile.

### 3.4 Failure Measurement Case Analysis

The validation experiments had demonstrated the robust performance of our dual-mode imaging system for breathing rate and pattern measurements. Nonetheless, this would be insufficient when, for example, the tracked points of ROI were completely obscured. Figure 9 displays two failed breathing measurement cases resulting from losing the tracked points. In Fig. 9(a), the subject pushed his glasses during the experiment, thus leading to the failure of cross-spectrum ROI tracking. For the second case, the targeted points were missing because of the out-of-plane movement. The relevant improvement of the algorithm should in the future be made to let the measurement continue after losing the tracked points.

## 4 Conclusion

A dual-mode imaging system, on visible and long-wave infrared wavelengths, has a capability of being used as a noncontact and

nonobtrusive measurement tool to estimate breathing rate and pattern, instead of the conventional methods. The addition of RGB images allowed the more accurate and faster detection and tracking of face and facial tissue in thermal images. Moreover, integrating the biological characteristics into the custom cascade classifier using the Viola–Jones algorithm yielded the superior accuracies for detecting face, nose, and mouth with classification accuracies of 98.46%, 95.38%, and 84.62%, respectively. For breathing rate estimation, the dual-image derived results were in agreement with those measured by the reference method, regardless of whether the subjects used nose and mouth simultaneously, alternately, or individually when they breathed. Taking the open-mouth breathing into account made the system highly adaptable for home care and clinical applications. Through visual comparison, the different breathing patterns could be clearly revealed by the extracted pixel intensities of thermal images. Apart from the situations requiring recovering the ROIs, the proposed system proved to be robust against challenging conditions such as significant positioning and abrupt physiological variations.

### Disclosures

Dr. Hu, Dr. Zhai, Mr. Li, Mr. Fan, Mr. Chen, and Dr. Yang have nothing to disclose. The authors have no relevant financial interests in the manuscript and no other potential conflicts of interest to disclose.

### Acknowledgments

The authors would like to acknowledge the China Postdoctoral Science Foundation funded project (No. 2016M600315) and the financial support from the National Science Foundation of China under Grant Nos. 61422112, 61371146, and

61221001. The authors thank all the volunteers who agreed to participate in this study.

## References

- G. Zamzmi et al., "Machine-based multimodal pain assessment tool for infants: a review," arXiv preprint arXiv 1607.00331 (2016).
- J. Fei and I. Pavlidis, "Virtual thermistor," in *29th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society*, pp. 250–253, IEEE (2007).
- S. Telles et al., "Oxygen consumption and respiration following two yoga relaxation techniques," *Appl. Psychophysiol. Biofeedback* **25**(4), 221–227 (2000).
- F. Q. AL-Khalidi et al., "Respiration rate monitoring methods: a review," *Pediatr. Pulmonol.* **46**(6), 523–529 (2011).
- P. Young and M. Boentert, "Early recognition of Pompe disease by respiratory muscle signs and symptoms," *J. Neuromuscular Dis.* **2**(s1), S3 (2015).
- R. L. Riha, "Diagnostic approaches to respiratory sleep disorders," *J. Thorac. Dis.* **7**(8), 1373 (2015).
- J. Naranjo et al., "A nomogram for assessment of breathing patterns during treadmill exercise," *Br. J. Sports Med.* **39**(2), 80–83 (2005).
- Y. Nakayama et al., "Non-contact measurement of respiratory and heart rates using a CMOS camera-equipped infrared camera for prompt infection screening at airport quarantine stations," in *IEEE Int. Conf. on Computational Intelligence and Virtual Environments for Measurement Systems and Applications*, pp. 1–4, IEEE (2015).
- I. Pavlidis et al., "Human behaviour: seeing through the face of deception," *Nature* **415**(6867), 35–35 (2002).
- M. Folke et al., "Critical review of non-invasive respiratory monitoring in medical care," *Med. Biol. Eng. Comput.* **41**(4), 377–383 (2003).
- K. van Loon et al., "Non-invasive continuous respiratory monitoring on general hospital wards: a systematic review," *PLoS One* **10**(12), e0144626 (2015).
- C. B. Pereira et al., "Remote monitoring of breathing dynamics using infrared thermography," *Biomed. Opt. Express* **6**(11), 4378–4394 (2015).
- A. D. Droitcour et al., "Non-contact respiratory rate measurement validation for hospitalized patients," in *Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society*, pp. 4812–4815, IEEE (2009).
- L. Scalise et al., "Non-contact laser-based human respiration rate measurement," in *Advances in Laserology-Selected Papers of Laser Florence 2010: The 50th Birthday of Laser Medicine World*, pp. 149–155, AIP Publishing (2011).
- S. D. Min et al., "Noncontact respiration rate measurement system using an ultrasonic proximity sensor," *IEEE Sens. J.* **10**(11), 1732–1739 (2010).
- B. A. Reyes et al., "Towards the development of a mobile phonopneumogram: automatic breath-phase classification using smartphones," *Ann. Biomed. Eng.* **44**(9), 2746–2759 (2016).
- Y. Nam et al., "Monitoring of heart and breathing rates using dual cameras on a smartphone," *PLoS One* **11**(3), e0151013 (2016).
- D. Shao et al., "Noncontact monitoring breathing pattern, exhalation flow rate and pulse transit time," *IEEE Trans. Biomed. Eng.* **61**(11), 2760–2767 (2014).
- A. K. Abbas et al., "Neonatal non-contact respiratory monitoring based on real-time infrared thermography," *Biomed. Eng. Online* **10**(1), 93–17 (2011).
- C. B. Pereira et al., "Robust remote monitoring of breathing function by using infrared thermography," in *37th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society*, pp. 4250–4253, IEEE (2015).
- S. Hu et al., "A polarimetric thermal database for face recognition research," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition Workshops*, pp. 119–126, IEEE (2016).
- A. Basu et al., "Infrared imaging based hyperventilation monitoring through respiration rate estimation," *Infrared Phys. Technol.* **77**, 382–390 (2016).
- A. K. Abbas et al., "Non-contact respiratory monitoring based on real-time IR-thermography," in *World Congress on Medical Physics and Biomedical Engineering*, pp. 1306–1309, Springer (2009).
- A. H. Alkali et al., "Facial tracking in thermal images for real-time noncontact respiration rate monitoring," in *European Modelling Symp. (EMS '13)*, pp. 265–270, IEEE (2013).
- Y. Zhou et al., "Spatiotemporal smoothing as a basis for facial tissue tracking in thermal imaging," *IEEE Trans. Biomed. Eng.* **60**(5), 1280–1289 (2013).
- D. Hanawa et al., "Nose detection in far infrared image for non-contact measurement of breathing," in *Proc. of 2012 IEEE-EMBS Int. Conf. on Biomedical and Health Informatics*, pp. 878–881, IEEE (2012).
- F. Q. Al-Khalidi et al., "Tracking human face features in thermal images for respiration monitoring," in *ACS/IEEE Int. Conf. on Computer Systems and Applications-AICCSA*, pp. 1–6, IEEE (2010).
- M. S. Sarfraz and R. Stiefelhagen, "Deep perceptual mapping for thermal to visible face recognition," arXiv preprint arXiv:1507.02879 (2015).
- F. Lamare et al., "Respiratory motion correction for PET oncology applications using affine transformation of list mode data," *Phys. Med. Biol.* **52**(1), 121–140 (2007).
- P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proc. of the 2001 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR '01)*, pp. 511–518, IEEE (2001).
- J. Sh and C. Tomasi, "Good features to track," in *Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR '94)*, pp. 593–600, IEEE (1994).
- C. Harris and M. Stephens, "A combined corner and edge detector," in *Alvey Vision Conf.*, pp. 1306–1309, Citeseer (1988).
- B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *Proc. of the 7th Int. Joint Conf. on Artificial Intelligence (IJCAI '81)*, pp. 674–679 (1981).
- C. Tomasi and T. Kanade, *Detection and Tracking of Point Features*, School of Computer Science, Carnegie Mellon University, Pittsburgh (1991).
- Z. Kalal et al., "Forward-backward error: automatic detection of tracking failures," in *20th Int. Conf. on Pattern recognition (ICPR '10)*, pp. 2756–2759, IEEE (2010).
- J. M. Bland and D. G. Altman, "Measuring agreement in method comparison studies," *Stat. Methods Med. Res.* **8**(2), 135–160 (1999).
- A. Carkeet, "Exact parametric confidence intervals for Bland-Altman limits of agreement," *Optom. Vision Sci.* **92**(3), e71–e80 (2015).
- O. Oldenburg et al., "Cheyne-Stokes respiration in heart failure: friend or foe? Hemodynamic effects of hyperventilation in heart failure patients and healthy volunteers," *Clin. Res. Cardiol.* **104**(4), 328–333 (2015).
- C. Deepika et al., "An efficient method for detection of inspiration phase of respiration in thermal imaging," *J. Sci. Ind. Res.* **75**, 40–44 (2016).
- F. Nicolo and N. A. Schmid, "Long range cross-spectral face recognition: matching SWIR against visible light images," *IEEE Trans. Inf. Forensics Secur.* **7**(6), 1717–1726 (2012).
- A. Asthana et al., "Robust discriminative response map fitting with constrained local models," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 3444–3451, IEEE (2013).
- W. J. Baddar et al., "A deep facial landmarks detection with facial contour and facial components constraint," in *IEEE Int. Conf. on Image Processing*, pp. 3209–3213, IEEE (2016).
- B. Reyes et al., "Tidal volume and instantaneous respiration rate estimation using a smartphone camera," *IEEE J. Biomed. Health Inf.* PP(99), 2168–2194 (2016).

**Meng-Han Hu** received his PhD with honors in biomedical engineering at the University of Shanghai for Science and Technology in 2016. Currently, he is a postdoctoral researcher at Shanghai Jiao Tong University.

**Guang-Tao Zhai** received his BE and ME degrees from Shandong University, Shandong, China, in 2001 and 2004, respectively, and his PhD from Shanghai Jiao Tong University, Shanghai, China, in 2009, where he is currently a research professor with the Institute

of Image Communication and Information Processing. He received the Award of National Excellent PhD thesis from the Ministry of Education of China in 2012.

**Duo Li** is a PhD candidate at Shanghai Jiao Tong University.

**Ye-Zhao Fan** is a master's student at Shanghai Jiao Tong University.

**Xiao-Hui Chen** is a master's student at Shanghai Jiao Tong University.

**Xiao-Kang Yang** received his BS degree from Xiamen University, Xiamen, China, in 1994, his MS degree from the Chinese Academy of Sciences, Shanghai, China, in 1997, and the PhD degree from Shanghai Jiao Tong University, Shanghai, in 2000. He is currently a full professor and deputy director of the Institute of Image Communication and Information Processing, Department of Electronic Engineering, Shanghai Jiao Tong University.