ORIGINAL ARTICLE

A real-time camera-based adaptive breathing monitoring system

Yu-Ching Lee[1] · Abdan Syakura[1] · Muhammad Adil Khalil[2] · Ching-Ho Wu[3] · Yi-Fang Ding[4] · Ching-Wei Wang[1,2]

Abstract

Breathing is one of the vital signs used to assess the physical health of a subject. Non-contact-based measurements of both breathing rate and changes in breathing rate help monitor health condition of subjects more flexibly. In this paper, we present an improved real-time camera-based adaptive breathing monitoring system, which includes real time (1) adaptive breathing motion detection, (2) adaptive region of interest detection to eliminate environmental noise, (3) breathing and body movement classification, (4) respiration rate estimation, (5) monitor change in respiration rate to examine overall health of an individual, and (6) online adaptation to lighting. The proposed system does not pose any positional and postural constraint. For evaluation, 30 videos of 15 animals are tested with drugs to simulate various medical conditions and breathing patterns, and the results from the proposed system are compared with the outputs of an existing FDA-approved invasive medical system for patient monitoring. The results show that the proposed method performs significantly correlated RR results to the reference medical device with the correlation coefficient equal to 0.92 and *p*-value less than 0.001, and more importantly the proposed video-based method is demonstrated to produce alarms 10 to 20 s earlier than the benchmark medical device.

1 Introduction

Monitoring respiration rate (RR) and associated changes is important to examine the health condition of individuals [1]. Keeping track of these parameters is even more important for vulnerable individuals, including the critically ill, neonates, infants, and the elderly [2]. Anomalous RR is an important evidence for serious health issues and can also be used to predict potentially serious clinical events such as cardiopulmonary arrest, chronic heart failure, pneumonia, pulmonary embolism, weaning failure, overdose or admission to an intensive care unit [3–14]. Some studies have presented that RR is more efficient than other vital signs in differentiating between healthy and unhealthy individuals [15]. Moreover, monitoring changes of RR helps diagnose a high-risk patient up to 24 hours earlier than the actual event [15].

In the past decade, various non-contact-based methods for RR measurement have been explored. Respiratory movement is subtle and cyclic with different breathing patterns for individuals but is not easy to detect through vision. Tan et al. [2] detect breathing motion using difference of frames (DOF). Tan et al. [2] use image and signal processing techniques to derive information about chest and abdominal movements from a series of video images captured with a single camera. Tan et al.'s [2] method is largely dependent on distinctive pattern clothing and could be greatly affected when clothing is removed or by non-respiratory movements. Zhao et al. [16] used near-infrared illuminated camera to detect the heart rate (HR) and RR. Zhao et al.'s [16] method deconstructs single channel images using delay-coordinate transformation and independent component analysis to reveal the temporal dynamics of heart beat and breathing rate. Zhao et al.'s [16]

✉ Ching-Wei Wang
   cweiwang@mail.ntust.edu.tw; cwwang1979@gmail.com

✉ Yi-Fang Ding
   102008@w.tmu.edu.tw; voltageding@gmail.com

1   Graduate institute of Applied Science and Technology, National Taiwan University of Science and Technology, Taipei City, Taiwan

2   Graduate Institute of Biomedical Engineering, National Taiwan University of Science and Technology, Taipei City, Taiwan

3   Institute of Veterinary Clinical Sciences, National Taiwan University, Taipei City, Taiwan

4   Department of Otolaryngology, Wan Fang Hospital, Taipei Medical University, Taipei City, Taiwan

method requires that subjects sit in front of the camera and face towards the camera, which however poses strict positional and postural constraints on the monitoring subjects. Xia et al. [17] used the KINECT sensor and a translation surface to measure the RR. Xia et al. [17] use a motion magnification method to improve KINECT's inherent depth resolution from 1 cm to 1 mm. Xia et al. [17] programmed the KINECT to capture depth images and calculate the average depth over a thoracic area of interest, viewed almost parallel to the subject's surface, using the KINECT software development kit. Xia et al.'s [17] method requires that the translation surface is placed on the subject's abdomen, which also poses strict positional and postural constraints and furthermore causes discomfort to the subject. Wijenayake et al. [18] and Rehouma et al. [19, 20] used RGB-D cameras to measure RR, which also poses positional constraints on the subjects. Wijenayake et al. [18] use principal component analysis (PCA) to eliminate the spatial and temporal noise from the input depth data and construct a patient-specific respiratory motion model. Then, with high precision, this model is used to calculate external respiratory motion in real time. Rehouma et al.'s [19] method uses depth information captured by two (Red Green Blue-Depth) RGB-D cameras at various view angles. The depth information is then used to recreate a 3D surface of a patient's torso with high temporal and spatial resolution and broad spatial coverage using depth information. The motion data is recorded for the top of the torso as well as its two lateral sides. A recursive subdivision of the 3D space into cubic unit elements is used to estimate the volume for each reconstruction. Between successive reconstructions, the volume shift is measured using a subtraction technique. Wijenayake et al.'s [18] and Rehouma et al.'s [19, 20] methods poses positional constraints on the subjects and these positional and postural restrictions can make the subject uneasy and affect the measurement results. Sanyal et al. [21] used RGB camera to capture subject's face video to measure HR and RR by calculating variations in color of reflected light, i.e., Hue from the video of a subject's face. This technique is highly affected by the lighting conditions and by the individuals skin tone. Moreover, this method fails to measure RR and HR if the subject is not facing towards the camera. Massaroni et al. [22] also proposed a RGB video-based RR measurement while sitting in front of the camera. Massaroni et al.'s [22] technique consists of a laptop's built-in RGB camera and an algorithm for post processing of acquired video data. The study of pixel intensity shifts produces a waveform showing respiratory pattern from the recording of a subject's chest movements. However, this method use manual annotation of breathing region, which can again cause the postural restriction on the individuals. Harte et al. [23] used structured-light (SL) to detect chest wall motion to estimate the RR. In this method,

the cameras were placed around the patient, and the patient is required in a fixed standing position. More precisely, the patient must be in a precise marked position due to the system restrictions. Janssen et al. [24] measure RR with a video camera and also detect the region of interest. Janssen et al.'s [24] method exploit the intrinsic properties of respiration to find the respiratory region of interest and extract the respiratory signal via motion factorization based on the observation that respiration induced chest/abdomen motion is independent motion in a video. However, Janssen et al.'s [24] method was unable to produce accurate results and tends to fail in different lighting conditions. Al-Naji et al. [25] and Brieva et al. [26] use the motion magnification technique to measure RR from the video. Al-Naji et al.'s [24] method uses motion magnification method based on elliptic filter and wavelet decomposition to magnify breathing action that is difficult to see through naked eye. RR and its time parameters are calculated by identifying the fastest moving areas in magnified video frame sequences. However, the magnification techniques are also likely to amplify the environmental noise, which may cause inaccurate RR measurements. For this reason, Brieva et al. [26] used a convolutional neural network (CNN) with manual annotation of breathing region to address this problem. This technique however may fail in general because sleeping posture and breathing pattern vary in individuals. Wang et al. [27] built a persistent luminous impression model (PLIM) to detect subtle breathing motion signals from a camera without positional constraints for identification of obstructive sleep apnoea behavior, but Wang et al.'s method does not produce RR results and is sensitive to environmental noises. Lee et al. [28] used radar sensor for contactless measurement of RR and achieved high precision, but Lee et al.'s method fails to measure RR when subjects moved substantially. Phokela et al. [29] proposed a method to estimate different breathing rates with the help of nasal breath sound strength. Phokela et al.'s method uses a headset microphone placed beneath the nose and attached with a smartphone, which will cause positional constraints for subject as the subject will have to keep the microphone close to their nose. Reyes et al. [30] proposed a computer vision-based estimation of RR and respiratory movement. In this method, the system uses the combination of Kanade-Lucas-Tomasi, Viola-Jones and Harris Stephens feature algorithms to automatically detect and tracks a region of interest on the chest of a person facing a camera. Then the respiration movement and RR is estimated by using displacement in the vertical direction of the ROI. This method also poses a postural constraint. Schoun et al. [31] used thin medium thermal imaging to calculate RR, tidal volume, and mouth distribution. In this method, a subject breathes on a thin medium placed perpendicular to the exhaled airflow of the patient, while the heat signature is

recorded on the opposite side of the medium by a thermal camera, which also poses positional constraints on the subject. Nosrati et al. [32] use an electromagnetic-based doppler radar to calculate RR. However, the radar was placed in front of the patient, and the patient is required in a fixed standing position. More precisely, the patient must be in a precise marked position due to the system restrictions. Nam et al. [33] use a introduces a new method of monitoring for the simultaneous measurement of heart and respiratory rates using dual smartphone cameras. The suggested method estimates heart rates using a rear-facing camera, while breathing rates are measured using a front-facing non-contact camera at the same time. However, one drawback of this method is that during data recording, hand movement is reduced, which might not always be feasible.

In this paper, we present an improved adaptive real-time camera-based RR measurement system based on [34], and the proposed method is demonstrated to be able to automatically adapt to different lighting conditions, deal with environmental noise, detect the active breathing region, and does not pose any positional constraint. This paper is organized as follows. We introduce our adaptive real-time camera-based RR monitoring system algorithm in Section 2. In Section 3, we present the experimental setup for data acquisition and the results. Finally, Section 4 states concluding remarks.

2 Methods

In this research, we proposed a real-time camera-based adaptive RR monitoring system. The proposed system uses a regular smartphone RGB camera as a video input to monitor the RR and change in RR of the individuals. The proposed system is robust to lighting conditions, environmental noise, different breathing patterns (i.e., shallow, middle, and deep), and it is sensitive to subtle respiratory movement, which is difficult to detect otherwise. Figure 1 illustrates an overview of the proposed algorithm. To monitor RR, we observe the subtle respiratory movement while breathing and propose an adaptive RR monitoring system which consists of the following: (1) adaptive breathing motion detection, (2) adaptive region of interest detection to eliminate environmental noise, (3) breathing and body movement classification, (4) respiration rate estimation, (5) monitor change in respiration rate to examine overall health of an individual, and (6) online adaptation to lighting.

2.1 Adaptive breathing motion detection

Adaptive breathing motion detection is introduced in the proposed system to adapt to different lighting environments

and detect subtle breathing motion patterns accurately. Existing vision-based RR monitoring methods have a significant impact due to different lighting environments that compromise their results. Therefore, the proposed system assess the video input ($I_f(x, y)$) and adapt according to the lighting condition.

In order to detect the subtle breathing motion without any positional constraint, the proposed system is using a persistent luminous impression model (PLIM)[26] to detect the subtle breathing motion without positional constraints. Given a ($w \times h$) image, and a frame rate of F frames/s, the PLIM is initialized using the initial frame ($I_0(x, y)$) from the video input as in Eq. 1. The PLIM incorporates slow adaptation, allowing pose changes to be accommodated while allowing cyclical movements to be detected. The subtle breathing motion is detected in the region of interest (ROI) boundaries by calculating the difference between the current frame ($I_f(bx, by)$) and the background frame ($P_{f-v}(bx, by)$) at frame f using Eq. 2. The background frame is updated with the frequency of v to incorporate slow adaptation and enhance subtle breathing movement using Eq. 3. After breathing motion detection, PLIM's breathing activity map ($A_f(bx, by)$) is computed with the help of Eq. 4 by comparing the difference with the threshold (α). We also define the raw motion activity level (e_f) as the number of active pixels in the activity map ($A_f(bx, by)$) at frame f (see Fig. c). It is computed using Eq. 5 at frame f.

$$P_0(x, y) = I_0(x, y) \tag{1}$$

At frame f, the PLIM in the ROI is updated using

$$\Delta_f(bx, by) = I_f(bx, by) - P_{f-v}(bx, by) \tag{2}$$

$$P_f(bx, by) = P_{f-v}(bx, by) + \begin{cases} 1 & if \ \Delta_f(bx, by) > 0 \\ 0 & if \ \Delta_f(bx, by) = 0 \\ -1 & if \ \Delta_f(bx, by) < 0 \end{cases} \tag{3}$$

The PLIM activity map $A_f(bx, by)$ in the ROI is defined as

$$A_f(bx, by) = \begin{cases} 1 \ if \ \Delta_f(bx, by) > \alpha \\ 0 \qquad otherwise \end{cases} \tag{4}$$

$$e_f = \sum_{bx} \sum_{by} A_f(bx, by) \tag{5}$$

where $bx \in x_{ROI} = [x_1, x_2]$, $by \in y_{ROI} = [y_1, y_2]$, x_1 and x_2 are the first and last x-axis points of the ROI, y_1 and y_2 are the first and last y-axis points of the ROI, which are further elaborated in the Section 2.2. $I_f(bx, by)$ is the image at the frame f, $\Delta_f(bx, by)$ is the image difference at frame f, $A_f(bx, by)$ is the PLIM's activity map, e_f is the number of active pixels in the activity map. α is the motion
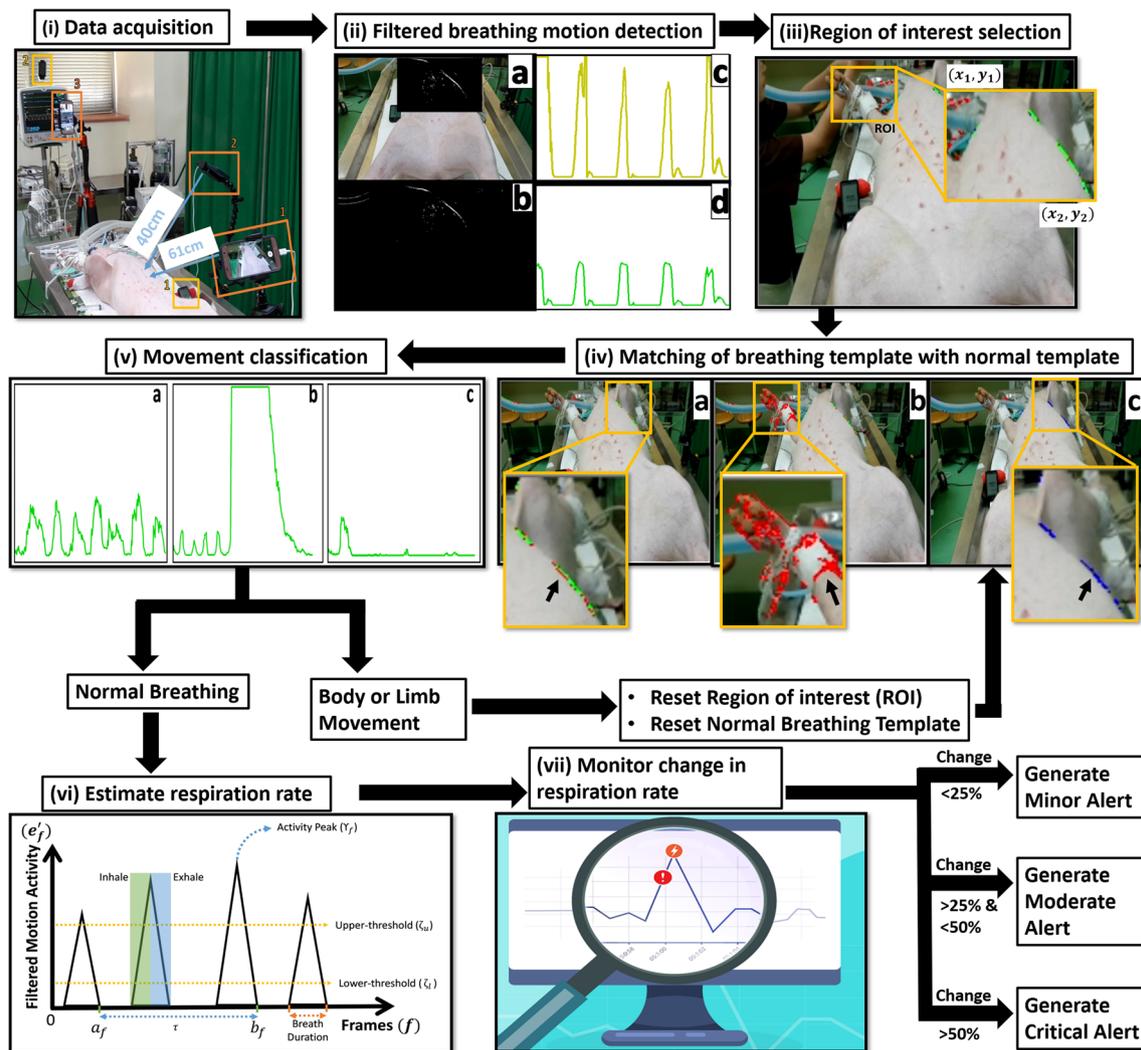
(i) Data acquisition

(ii) Filtered breathing motion detection

(iii)Region of interest selection

(v) Movement classification

(iv) Matching of breathing template with normal template

Normal Breathing

Body or Limb Movement

- **Reset Region of interest (ROI)**
- **Reset Normal Breathing Template**

(vi) Estimate respiration rate

(vii) Monitor change in respiration rate

Change <25% → **Generate Minor Alert**

Change >25% & <50% → **Generate Moderate Alert**

Change >50% → **Generate Critical Alert**

Fig. 1 The proposed framework for a real-time camera-based adaptive RR monitoring system. (i) Data acquisition, the animal data is recorded from three different smartphones highlighted with the orange boxes and two stopwatches highlighted with the yellow boxes. The distance between smartphones and subject is illustrated with the blue arrows. (ii) Filtered breathing motion detection. (ii, a) shows the real-time breathing motion detection with ROI and activity map in yellow rectangle; (ii, b) illustrates the saved normal breathing template, (ii, c) shows the real-time raw breathing activity levels from the activity map as an oscillogram; (ii, d) shows the filtered breathing activity levels on the oscillogram with the help of saved normal breathing template. (iii) Region of interest is selected avoiding the environmental noise. (v) Matching of breathing template with normal template. (iv,a) shows a normal breathing event where the green color indicates the motion matched with the normal breathing template; (iv, b) presents a limb movement event with high raw activity level (e_f) whereas the filtered breathing activity level (e'_f) remains stable; the red color indicates the non-breathing motion, which is not matched with the template; (iv, c) shows a non-breathing status where the subject is at rest without any movement, and the blue color indicates the latest saved normal breathing template. (v) Movement classification. (v, a) shows the signal of the normal breathing activity; (v, b) shows the signal of body or limb movement activity; (v, c) shows the respiratory arrest activity. (vi) Estimation of respiration rate, RR is estimated using filtered breathing activity (e'_f) signals over f video frames. a_f is the initial frame number, and b_f is the final frame number required to calculate the frame difference between two breaths. The calculated frame difference is illustrated as τ, which is used to estimate the RR per minute. Activity peak (Υ_f), upper-threshold (ζ_u), and lower-threshold (ζ_l) are used to detect the breathing activity. Inclining slope shaded in green shows inhalation, and declining peak shaded in blue shows exhalation. Upper and lower thresholds are marked to identify correct inhalation and exhalation patterns. (vii) Monitor change in respiration, minor alert is generated if the change is RR is less than 25%, moderate alert is generated if the change in RR is more than 25% and less than 50%, and critical alert is generated if change in RR is more than 50%

detection threshold, and ν is the PLIM background update frequency, which are further elaborated in Section 2.6.

In this paper, an improved method to produce stable breathing signals by

filtering out signals from environmental noises and other body movements is developed. First of all, the normal breathing template (B_f) of an individual is saved initially from the activity map ($A_f(bx, by)$), which is used as the

breathing motion reference. Then, the real-time raw motion is compared and matched with the reference template to identify breathing signals using Eq. 6. The filtered breathing activity level (e'_f) is defined in Eq. 7, which is the number of active pixels in the breathing activity map ($M_f(bx, by)$). The raw activity level e_f is produced based on the raw motion data as in [26], the proposed filtered breathing activity model is able to distinguish breathing signals from signals caused by other movements and produce reliable breathing activity measurement e'_f. Importantly, the online normal breathing template is designed to be adaptive and is automatically updated if the proposed system detects a notable raw activity level (e_f) while detects none or very less filtered breathing activity level (e'_f) for ten continuous seconds.

$$M_f(bx, by) = \begin{cases} 1 \ if \ A_f(bx, by) = B_f(bx, by) \\ 0 \qquad\qquad otherwise \end{cases} \quad (6)$$

$$e'_f = \sum_{bx}\sum_{by} M_f(bx, by) \quad (7)$$

where B_f is the saved breathing template for individuals. The $M_f(bx, by)$ is filtered breathing activity map from the raw activity map $A_f(bx, by)$ and the e'_f is the filtered breathing activity level at frame f.

2.2 Adaptive region of interest detection

In order to develop a robust and reliable vision-based breathing monitoring system, it is important to deal with foreground and background noises, distinguish breathing movements from other motion signals and produce accurate measurement of breathing activities. As a result, we build a LABB model for localization of active breathing behavior regions, which helps produce pure breathing signals and quantitative measurement results. In our preliminary simulated tests, Fig. 2 shows that the proposed LABB model is able to deal with Fig. 2a background noises such as a moving ball and Fig. 2b foreground noises such as a person walking by and works well with various camera view angles as shown in Fig. 2c.

Obtaining individual normal breathing template B_f, which is defined in Eq. 18, the system identifies the area with personalized breathing patterns as the active region of interest R' for analysis. The proposed system automatically reset the ROI if a large motion event is detected. For initialization, the breathing region (R) is detected from the full-frame is refined by applying erosion to remove noises. After that, a bounding box from the refined region (R') is determined. In some cases, the breathing motion may have a weak signal which results in ROI to be small. The proposed

system resolves this situation using Eqs. 10, 11, 12, and 13.

$$H = \begin{pmatrix} 0 \ 1 \ 0 \\ 1 \ 1 \ 1 \\ 0 \ 1 \ 0 \end{pmatrix} \quad (8)$$

$$R'_f = B_f \ominus H \quad (9)$$

$$x_1 = \begin{cases} (1 - \Omega)x'_1 \ if \ x'_2 - x'_1 < w/\lambda \\ x'_1 \qquad\qquad otherwise \end{cases} \quad (10)$$

$$x_2 = \begin{cases} (1 + \Omega)x'_2 \ if \ x'_2 - x'_1 < w/\lambda \\ x'_2 \qquad\qquad otherwise \end{cases} \quad (11)$$

$$y_1 = \begin{cases} (1 - \Omega)y'_1 \ if \ y'_2 - y'_1 < h/\lambda \\ y'_1 \qquad\qquad otherwise \end{cases} \quad (12)$$

$$y_2 = \begin{cases} (1 + \Omega)y'_2 \ if \ y'_2 - y'_1 < h/\lambda \\ y'_2 \qquad\qquad otherwise \end{cases} \quad (13)$$

where R' is the breathing region after noise removal with bounding box $R' = \{x'_1, y'_1, x'_2, y'_2\}$, w is the width of the image, h is the height of the image, Ω is the resize factor, and λ is the video frame dividing factor (in this study Ω = 0.15, and λ = 5); ROI with bounding box is defined as $b = \{x_1, y_1, x_2, y_2\}$ with dimensions $x_{ROI} \times y_{ROI}$.

2.3 Breathing and body movement classification

Breathing and activity classification is pivotal to the accuracy of the proposed system. Unconscious body and limb movements are natural phenomena, which is unavoidable. The proposed system, with the help of raw activity level (e_f) and filtered activity level (e'_f), is able to distinguish between respiratory and non-respiratory movements. Figure 1v illustrates the oscillogram about different activity patterns based on the activity level and peak duration.

Accurate activity classification would ensure the results produced by the proposed system are accurate and robust. After the body or limb movement detection, the proposed system reset the ROI and normal breathing template to automatically adapt according to the new resting position of the subject. Moreover, it excludes that particular movement from the RR calculation. As a result, the proposed system can easily adapt to the new position and keep producing accurate RR measurements.

2.4 Respiratory rate estimation

The respiratory activity causes the cyclic movement of individuals with varying frequency and depth. The video inputs are a series of f frames at a frame rate (F) per second, and a complete breathing activity is identified using inspiration and expiration events from the filtered breathing activity level (e'_f). Initially, the maximum peak value of the
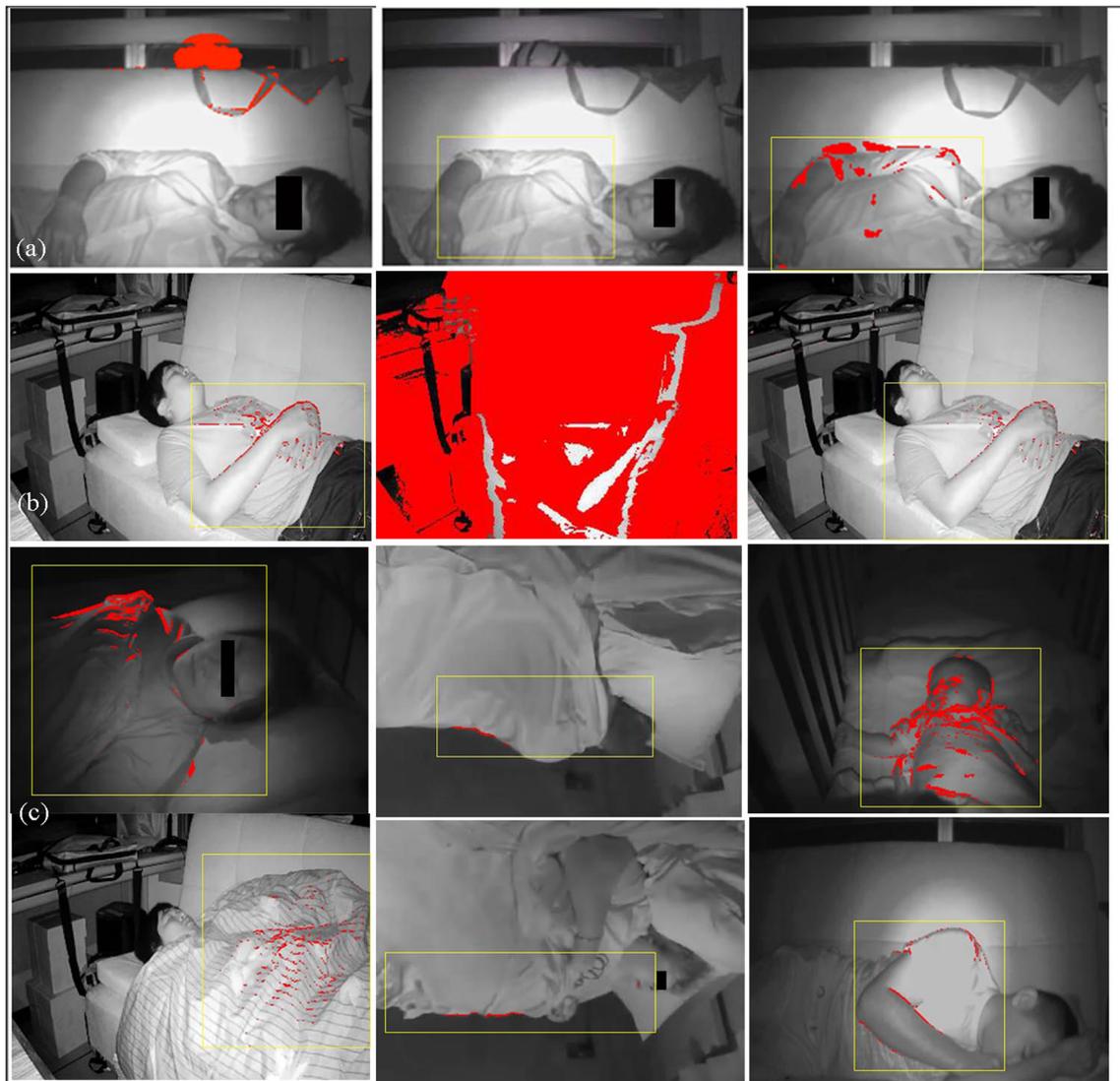
Fig. 2 The results of preliminary tests on the localization of active breathing regions in the presence of foreground and background noises, and on different view angles. The results shows **a** the proposed method is able to effectively and robustly identify the active breathing regions despite the background noise caused by throwing a ball. **b** The proposed method can effectively locate the active breathing regions despite the foreground noise caused by a human walking by. **c** The proposed framework can identify the region of active breathing on different view angles. The red color indicates the motion signal, and the yellow color represents the detected active breathing region

filtered activity level (Υ_f) is detected from the maximum filtered activity level (e'_f) using Eq. 14. Upper (ζ_u) and lower (ζ_l) thresholds are calculated using Eqs. 15 and 16, respectively (see Fig. 1vi). These upper and lower thresholds are used to detect a complete inhalation and exhalation activity without interference of unnecessary movement using Eq. 17. The breathing pattern of an individual could change over time. For this reason, the proposed system is designed to adapt different breathing patterns. The peak value (Υ_f) is decremented by every second (see Eq. 14). According to the results, this adaptive approach is proved to be effective in detecting the inhalation and exhalation activity efficiently, even if the breathing

activity level (e'_f) decreases or increases over time. Furthermore, the time interval between two breathing peaks could also help us to identify and classify the movement as a respiratory or non-respiratory movement. If the system detects a non-respiratory movement or a long breathing interval, then the activity peak (Υ_f) is reset.

$$\Upsilon_f = \begin{cases} e'_f & if \; e'_f > \Upsilon_{f-1}) \\ \Upsilon_{f-F} - 1 & otherwise \end{cases} \quad (14)$$

$$\zeta_u = \Upsilon_f \times \sigma_1 \quad (15)$$

$$\zeta_l = \Upsilon_f \times (1 - \sigma_1) \quad (16)$$

$$\rho_f = \begin{cases} 1 & if \quad e'_f > \zeta_u \quad \wedge \rho_{f-1} = 0 \\ 0 & if \quad e'_f < \zeta_l \quad \wedge \rho_{f-1} = 1 \\ \rho_{f-1} & otherwise \end{cases} \quad (17)$$

where Υ_f is the maximum level of breathing activity at frame f, and F is the video frame rate per second; σ_1 = 0.75 is empirically determined for further calculation of the upper-threshold (ζ_u) and lower-threshold (ζ_l); ρ_f is the breathing peak value at frame f, and e'_f is the filtered breathing activity level at frame f.

The RR (r_f) is estimated by using the time interval between two breathing activities (τ) (see Fig. 1, vi). The number of breathing activities (ß_f) are counted according to the detected breathing activities using Eq. 18. Further, the initial frame number (a_f) and final frame number (b_f) are detected from the series of frames (f) using Eqs. 19 and 20, respectively. The time interval between two breathing activities (τ) could be calculated from the difference of the final frame (b_f) and initial frame (a_f) as in Eq. 21. This time interval is further used to estimate RR per minute (r_f) at frame f using Eq. 22.

$$\text{ß}_f = \begin{cases} \text{ß}_{f-1} + 1 & if \quad \rho_{f-1} - \rho_f = 1 \\ 0 & if \quad \text{ß}_{f-1} = L \\ \text{ß}_{f-1} & otherwise \end{cases} \quad (18)$$

$$a_f = \begin{cases} f & if \quad \text{ß}_{f-1} - \text{ß}_f = L \\ a_{f-1} & otherwise \end{cases} \quad (19)$$

$$b_f = \begin{cases} f & if \quad \text{ß}_{f-1} + \text{ß}_f = L + 1 \\ b_{f-1} & otherwise \end{cases} \quad (20)$$

$$\tau = \frac{|b_f - a_f|}{F} \quad (21)$$

$$r_f = \frac{\chi}{\tau} \times L \quad (22)$$

where ß_f is the breath counter at frame f, b_f is the final frame number at frame f, a_f is the initial frame number at frame f, F is the video frame rate per second, r_f is the RR at frame f, χ is the number of seconds in 1 min (χ = 60), and τ is the elapsed time for L number of breaths (in this study $L = 2$, which is empirically determined parameter).

2.5 Monitor change in respiration rate

The proposed system is also able to monitor the change in RR. Monitoring change in RR is important to identify critical events and critical events must be notified to the medical experts. For this reason, the proposed system also introduce intelligent alarming system, which the existing monitoring system lacks. Although RR is estimated every two breaths in the proposed, but change in the RR is calculated every 10 s. The proposed system has three types of alarming system: (1) the adjustable upper and lower threshold; (2) calculation of the change in respiration rate

over a certain time period; and (3) the respiratory arrest or no breathing alarm. The adjustable upper and lower threshold works similar with the existing system (GE Dash 5000 [27]). The alarm will be triggered if the RR exceeds the upper limit, and go below the lower limit. Equation 23 shows alarm triggering conditions.

$$Al_f = \begin{cases} N_U & if \quad r_f > \varepsilon_u \\ N_L & if \quad r_f < \varepsilon_l \\ N_G & otherwise \end{cases} \quad (23)$$

where Al_f is the fixed threshold alarm trigger, ε_u is the upper threshold, ε_l is the lower threshold, and r_f is the RR at frame f. N_U and N_L is the alarm for upper and lower limit, respectively, whereas N_G is the condition when the alarm is not triggered.

For the rate of change in RR, the proposed system calculates the change (Δr_f), and percentage change (δr_f) every 10 s (see Eq. 24), which the existing ETCO2 monitors are unable to calculate. The proposed system classifies the alarming situation according to the calculated percentage. The alarming situations include the critical alarm, the moderate alarm or the minor alarm (shown in Eq. 24). This intelligent classification can alert and inform a medical staff about the patient's condition.

$$\Delta r_f = |r_f - r_{f-k}| \quad (24)$$

$$\delta r_f = \frac{\Delta r_f}{r_f} \times 100 \quad (25)$$

$$Al_c = \begin{cases} N_A & if \quad \varpi_1 < \delta r_f \leq \varpi_2 \\ N_B & if \quad \varpi_2 < \delta r_f \leq \varpi_3 \\ N_C & if \quad \delta r_f > \varpi_3 \\ N_G & otherwise \end{cases} \quad (26)$$

where Δr_f is the rate of change of RR every k number of frames (in this study $k = 10F$), δr_f is the percentage change in the RR, and Al_c is the alarming condition classification according to the δr_f. N_C shows critical change in the RR, N_B shows moderate change in the RR, N_C shows minor decrement and increment in the RR, and N_G shows no change in the RR. In this study $\varpi_1 = 1\%$, $\varpi_2 = 25\%$, and $\varpi_3 = 50\%$ are empirically determined parameter.

The proposed system could efficiently generate the alarms if there is no breathing activity for $\mho$ continuous seconds (see Eq. 27). Furthermore, the proposed system also includes an adaptive respiratory arrest detection system, which the ETCO2 monitor (GE Dash 5000 [27]) lacks. The ETCO2 monitor (GE Dash 5000 [27]) is unable to adapt according to the conditions, but it is able to check for no-breathing over a fixed (default) time interval. However, the proposed system is able to adapt according to the condition with time. As a result, the adaptability and the accuracy of the proposed system increase. Initially, the proposed system generates its first respiratory arrest critical alarm after $\mho$

seconds of inactivity and reset the RR result. Further, the system will automatically add Ʊ seconds to the previous waiting time (q_f) until it crosses the maximum RR reset limit (Z). When the RR reset limit (q_f) cross the maximum limit (Z), then the system will automatically reset the RR reset limit to Ʊ seconds. Due to this adaptability, our system is able to adapt for some special conditions and could estimate as low RR as two (see Fig. 5b), whereas the existing system (GE Dash 5000 [27]) waiting time is fixed, and it is unable to adapt to the low respiration rate. Thus, the existing system (GE Dash 5000 [27]) is unable to estimate the RR, which lower than six breaths per minute.

$$Al_n = \begin{cases} 1 \ if \ \sum_f^{Ʊ} r_f = 0 \\ 0 \quad otherwise \end{cases} \tag{27}$$

$$q_f = \begin{cases} q_{f-1} + Ʊ \ if \quad Al_n = 1 \ \wedge \ q_{f-1} < Z \\ \quad Ʊ \quad \ if \ \ q_{f-1} > Z \\ \quad q_{f-1} \quad \quad otherwise \end{cases} \tag{28}$$

where Al_n is no breathing alarm, Ʊ is the threshold for breathing inactivity, r_f is the RR at frame f, q_f is the threshold for RR to reset at frame f, and Z is the maximum frame limit for breathing inactivity. In this study $Ʊ = 10F$, and $Z = 30F$.

2.6 Online adaptation to lighting

For online adaptation to lighting, the proposed system initially assesses the first frame ($I_0(x, y)$) of the video input and computes the illumination level (φ_f). Each frame is an image composed of three channels, i.e. red, green, and blue channels. RGB image ($I_f(x, y)$) is converted into a grayscale image ($V_f(x, y)$) by averaging the pixel intensities of the three channels, and from the grayscale image, the illumination is calculated using Eq. 29.

$$\varphi_f = \frac{\sum_x^w \sum_y^h V_f(x, y)}{w \times h} \tag{29}$$

where $V_f(x, y)$ is the grayscale image at frame f with dimensions $w \times h$, φ_f is the illumination with resulting values in the range [0, 255] at frame f.

Motion detection threshold (α) and background update frequency (ν) are determined using Eqs. 30 and 31. α and ν are the key parameters, which control the sensitivity of subtle breathing motion detection. α is directly proportional to the sensitivity of the proposed system; ν also increases the motion sensitivity of the breathing movement, which means that the system will update the persistent impression frame after every ν frame.

$$\alpha = \begin{cases} \omega_1 \ if \quad \varphi_f < \beta_1 \\ \omega_2 \ if \ \ \beta_1 \leq \varphi_f \leq \beta_2 \\ \omega_3 \ if \quad \varphi_f > \beta_2 \end{cases} \tag{30}$$

$$\nu = \begin{cases} \gamma_1 \ if \quad \alpha = \omega_1 \\ \gamma_2 \quad \quad otherwise \end{cases} \tag{31}$$

where α is motion detection threshold, and ν is background update frequency; $\omega_1 = 8$, $\omega_2 = 12$, $\omega_3 = 24$, $\gamma_1 = 4$, $\gamma_2 = 2$, $\beta_1 = 83$, and $\beta_2 = 100$ are empirically determined.

3 Data and results

3.1 Experimental setup and data acquisition

The video data were acquired from The National Taiwan University Veterinary Hospital. The animal use protocol has been reviewed and approved (Approval Number NTU106-EL-00116) by the Institutional Animal Care and Use Committee (IACUC), National Taiwan University, Taiwan. All experiments were carried out in accordance with the approved guidelines. Our experimental setup is shown in Fig. 1i. Using three smartphones and two digital stopwatches, the videos were recorded at 1080p resolution with 30 frames per second. Two smartphones were used to record the animal video from two different angles, and the distance from the camera to the target is shown in Fig. 1i. The third smartphone was used to record the measurement outputs of the referenced device, i.e. a FDA-approved medical device (GE Dash 5000 [27]). Stopwatches were used to synchronize the video of the target and the video of the referenced device.

In total, 15 young healthy animals (Landrace × York-shire × Duroc (LYD) pigs) were used as subjects to record the video data twice with 7-day interval between two data acquisition. Therefore, 30 recorded video data were collected. In order to evaluate the proposed system more comprehensively, the data from each subject was recorded and classified into seven different stages to simulate various medical conditions and breathing patterns. When the subject was being prepared for the experiment, premedication was administered with intramuscular injection of a mixture of 2 mg/kg xylazine (Rompun, Bayer Korea Ltd., Seoul, Korea) and 4 mg/kg Zoletil (a mixture of tiletamine hydrochloride and zolazepam hydrochloride, Virbac Laboratory, Carros, France). A 22 intravenous catheter was placed in the marginal auricular vein through which the fluids and the drugs were subsequently injected. The anaesthesia was inducted with intravenous propofol 1mg/kg. Once anaesthetized and tracheal intubated, general anaesthesia was maintained with isoflurane (2–3%) throughout the procedure. Heart rate, ECG, BT, mean arterial pressure, pulse oximetry, ETCO2, and RR were monitored continuously. Respiratory condition and rate in the general anaesthesia were controlled by mechanical ventilator (Anesthesia Delivery System (ADS) 1000, Engler, Florida, USA) and intravenous administration of muscle relaxants, 10%

atracurium. Pressure control mode was used on ADS 1000 to control the RR of the subject.

Figure 3 illustrates the seven stages with ADS ventilation rate in the experimental design to simulate various medical conditions and breathing patterns. In the beginning of the anaesthesia (Stage 1), atracurium was given a loading dose of 0.5 mg/kg followed by 0.5 mg/kg/h constant rate infusion for 15 mins and the manual ventilator rate of ADS was set at one breath per minute (the lowest frequency). In this stage (Stage 2), the subject will be breathing with the help of mechanical ventilator only. The effect of atracurium would be onset about 2 min after injection. Spontaneous respiratory would be gradually recovered about 10 min after withdrawal of atracurium. This means that the subject will gradually recover in stage 3. In the beginning of stage 4, the ventilator rate of ADS was set at 30 breaths per minute then decreased to 10 breaths per minute to enter stage 5. In the stage 6, we turned off the ventilator. The concentration of isoflurane was elevated to 5% in stage 7 to deepen the anaesthesia and to observe the change of the respiratory condition. In the first data collection procedure, all pigs gradually recovered from general anaesthesia since the end of stage 6 by turning off the isoflurane. In the second experiment, propofol and potassium chloride (KCl) were introduced intravenously to euthanize the pigs at the end-point of this experiment.

The breathing pattern in the animal data is categorized into three types, including spontaneous breathing, machine aided breathing, and spontaneous plus machine aided breathing. Spontaneous breathing occurs in stage 1, and from stage 6 till the end. Machine aided breathing occurs in

Table 1 Illustrates the average accuracy of the RR results from different viewing angles

View	Avg. distance to subjects (cm)	Avg. relative accuracy (%)
Lateral	40.2	91.36
Anterior	62	94.91

stage 2 only, and spontaneous plus machine aided breathing occurs from stage 3 to stage 5. For detailed comparison, RR values from the monitor and the proposed system needed to be quantized and recorded every fixed time interval. In this study, we computed RR every two breaths and calculated the difference every 10 s.

3.2 Results and discussion

The RR measurements from the proposed system were compared with the results from the FDA-approved ETCO2 monitor (GE Dash 5000 [27]) to evaluate the accuracy of proposed system. In the preliminary test, five trials were performed initially with two different viewing angles (anterior and lateral) of the animals in supine position. RR results were compared with the reference standard from both viewing angles, and the anterior view is proved to be more accurate than the lateral view (see Table 1). Based on the preliminary test results, an anterior view is selected for further detailed experiment. Full evaluation was performed using an anterior view of 30 animal videos, each having
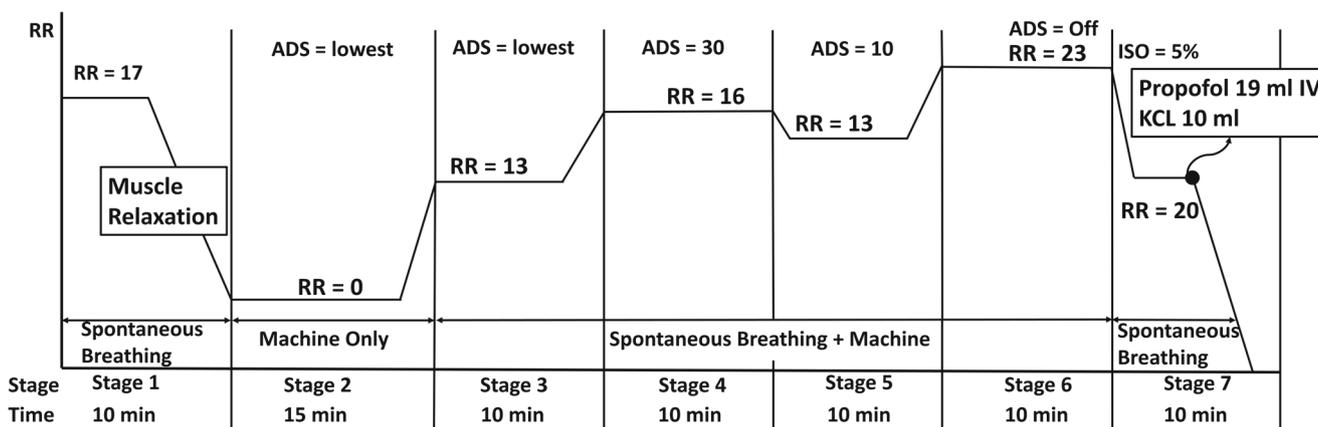

Fig. 3 This figure illustrates the experimental setup with different stages to simulate various medical conditions and breathing patterns. An anesthesia delivery system (ADS) is used as positive pressure ventilator (Pressure Control Mode) to control the respiratory rate (RR) during machine aided and spontaneous plus machine aided breathing stages. In Stage 1, atracurium was given a loading dose of 0.5 mg/kg followed by 0.5 mg/kg/h constant rate infusion for 15 mins and the manual ventilator rate of ADS was set at one breath per minute (the lowest frequency). In Stage 2, the subject will be breathing with the help of mechanical ventilator only. The effect of atracurium would be onset about 2 min after injection. Spontaneous respiratory would be gradually recovered about 10 min after withdrawal of atracurium. This means that the subject will gradually recover in stage 3. In the beginning of stage 4, the ventilator rate of ADS was set at 30 breaths per minute then decreased to 10 breaths per minute to enter stage 5. In stage 6, we turned off the ventilator. The concentration of isoflurane was elevated to 5% in stage 7 to deepen the anaesthesia and to observe the change of the respiratory condition

Fig. 4 The box plots of the RR measurements by the reference medical device and the proposed system in two different breathing stages where the outliers > 1.5× interquartile range are marked with a dot and outliers > 3× interquartile are marked with an asterisk, including **a** the spontaneous breathing stage and **b** the spontaneous plus machine aided breathing stage. The results show that the proposed method produces correlated results with the medical monitoring device and generates comparably less outliers
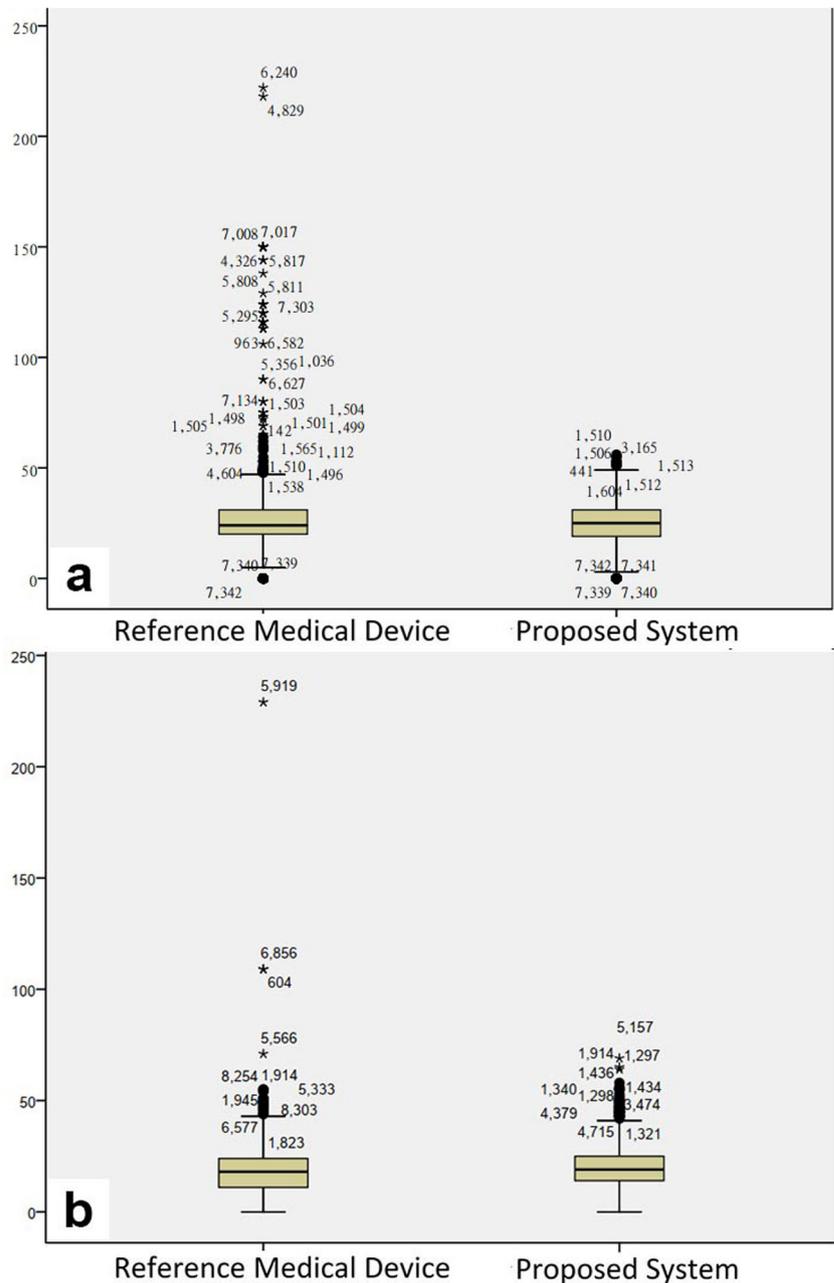

seven stages for simulating medical conditions (see Fig. 3). The RR values were recorded every 10 s for evaluation.

Using Spearman's rho test, the proposed system produces significantly correlated results to the referenced medical device with a correlation coefficient equal to 0.92 and p-value less than 0.001. Figure 4 shows RR results of (a) spontaneous breathing and (b) spontaneous plus machine aided breathing box plots from the proposed method and the referenced medical device, showing that the proposed system produces highly correlated and stable results.

3.3 Discussion

In our experiments, it is found that the proposed system is more sensible and stable in comparison to the reference medical device. Figure 5 illustrates examples from our experiments. Figure 5a shows that in stage 1 the spontaneous breathing results from the proposed system are highly consistent with the reference device with the correlation coefficient equal to 0.92, and the p-value < 0.001 using Spearman's rho test. At the end of stage 1, there

Fig. 5 a In stage 1 the spontaneous breathing results from the proposed system are highly consistent with the reference medical device. **b** the reference medical device fails to produce results because of low RR while the proposed system adapts to that condition and successfully produces results. **c** Results from the proposed system and the reference device are also significantly correlated. **d** For the critical stages from spontaneous breathing to respiratory arrest, the results from the proposed system and the reference device are highly consistent, but more importantly the proposed vision-based monitoring system is demonstrated to be able to notify the critical situation 10 to 20 s earlier than the reference medical device
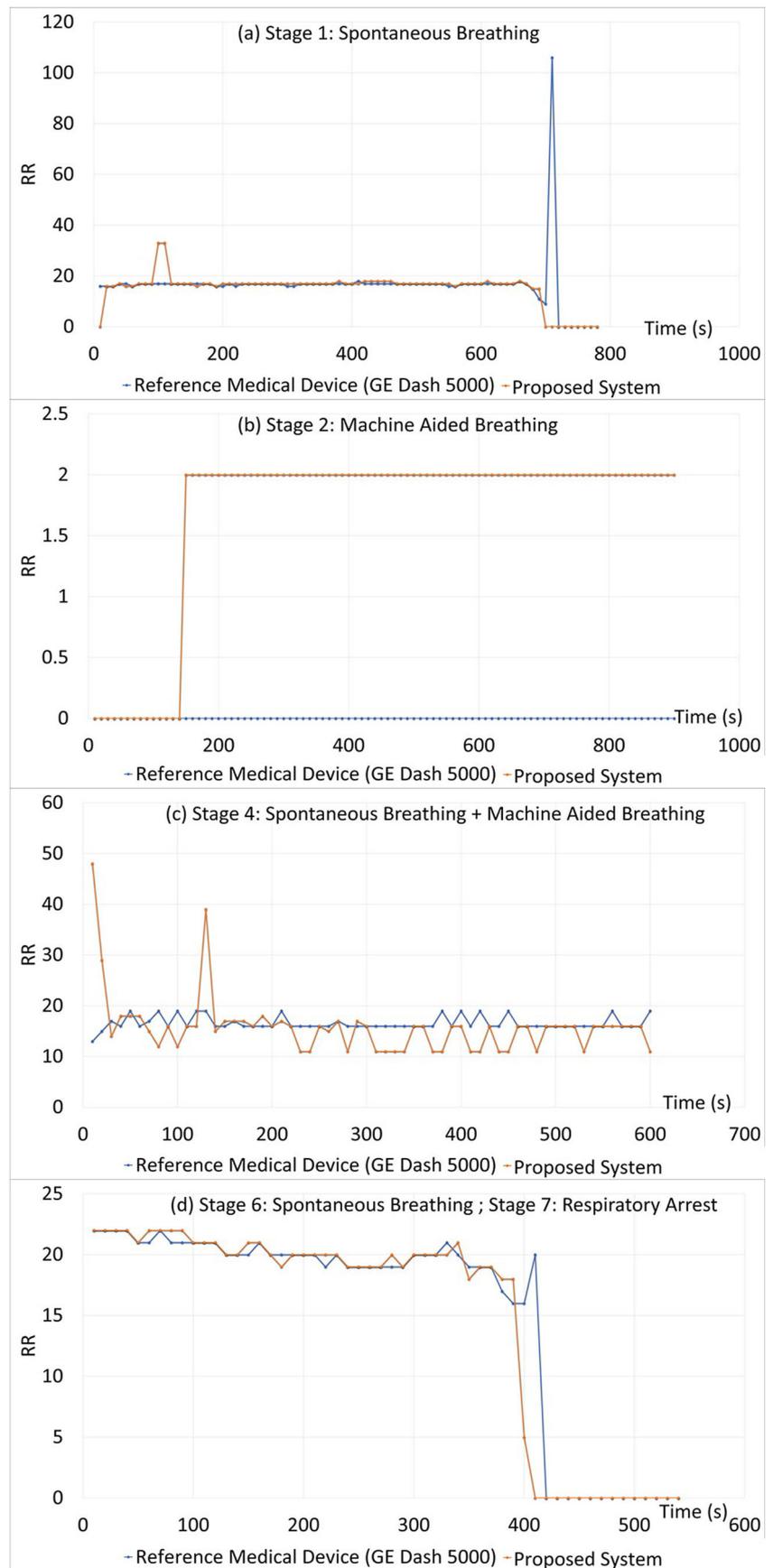
is a unreasonable sudden drastic increase in the RR value from the reference medical device while the proposed system remains stable. In addition, the proposed vision-based system is able to notify the medical doctors about the critical situation 10 to 20 s earlier than the reference medical device. However, human intervention in the foreground may cause minor noises for the proposed vision-based system as observed at 100 s at stage 1 the results are unusual. For the machine aided breathing, the reference medical device fails to produce results because of occurrence of extreme low RR while the proposed system adapts to that condition and successfully produces results as shown in Fig. 5b. In the stage of spontaneous plus machine aided breathing as shown in Fig. 5c, results from the proposed system and the reference device are also significantly correlated with the correlation coefficient equal to 0.63, and the p-value equal to < 0.001 using Spearman's rho test. For the critical stages from spontaneous breathing to respiratory arrest (see Fig. 5d), the results from the proposed system and the reference device are highly consistent, but more importantly the proposed vision-based monitoring system is demonstrated to be able to notify the critical situation 10 to 20 s earlier than the reference medical device, which is also shown in stage 1 Fig. 5a. The proposed framework does not require expensive state-of-the-art equipment, but it can still produce accurate results as compared to the FDA-approved ETCO2 monitor. The proposed framework only requires a smartphone and a personal computer. While for the stand-alone system, only a smartphone is required.

4 Conclusion

In this paper, we present an improved camera-based video-based respiration monitoring system, which includes real time (1) adaptive breathing motion detection, (2) adaptive region of interest detection to eliminate environmental noise, (3) breathing and body movement classification, (4) Respiration rate estimation, (5) monitor change in respiration rate to examine overall health of an individual, and (6) online adaptation to lighting. The proposed method has been thoroughly evaluated using 30 animal videos with simulation of various breathing conditions such as spontaneous breathing, machine aided breathing, combination of spontaneous and machine aided breathing and respiratory arrest. The proposed method performs highly consistent RR measurement results to the reference contact-based medical monitoring device, and more importantly the proposed method is demonstrated to produce alarms 10 to 20 s earlier than the conventional medical device. For the system limitation, in our simulation tests, it is found that a continuous and cyclic moving fan in the scene would influence the system as shown in Fig. 6. In the future work, we would like to
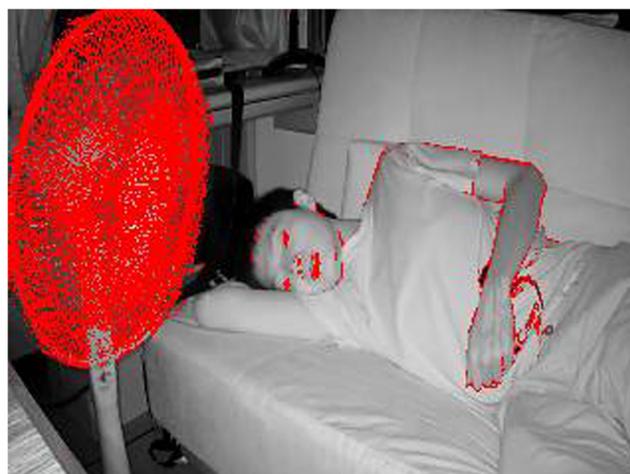

Fig. 6 System limitation. A continuous and cyclic moving fan in the scene would influence the system. In the future work, we would like to improve the system by integrating the concept of reasonable moving speed to distinguish breathing behavior from other cyclical movements

improve the system by integrating the concept of reasonable moving speed to distinguish breathing behaviour from other cyclical movements. Another improvement could be made for the future work is to use deep learning-based image processing to extract breathing features.

References

1. Liu H, Allen J, Zheng D, Chen F (2019) Recent development of respiratory rate measurement technologies
2. Tan KS, Saatchi R, Elphick H, Burke D (2010) Real-time vision based respiration monitoring system. In: 2010 7th international symposium on communication systems, networks digital signal processing (CSNDSP 2010), pp 770–774
3. Fieselmann JF, Hendryx M, Helms C, Wakefield D (1993) Respiratory rate predicts cardiopulmonary arrest for internal medicine patients. J Gen Intern Med 8:354–60
4. Ponikowski P, Chua TP, Anker SD, Francis DP, Doehner W, Banasiak W, Poole-Wilson PA, Piepoli MF, Coats AJ (2001) Peripheral chemoreceptor hypersensitivity. Circulation 104(5):544–549
5. Hodgetts TJ, Kenward G, Vlachonikolis IG, Payne S, Castle N (2002) The identification of risk factors for cardiac arrest and formulation of activation criteria to alert a medical emergency team. Resuscitation 54(2):125–131
6. Subbe CP, Davies RG, Williams E, Rutherford P, Gemmell L (2003) Effect of introducing the modified early warning score on clinical outcomes, cardio-pulmonary arrests and intensive care utilisation in acute medical admissions*. Anaesthesia 58(8):797–802
7. Goldhill DR, McNarry AF, Mandersloot G, McGinley A (2005) A physiologically-based early warning score for ward patients: the

association between score and outcome*. Anaesthesia 60(6):547–553

8. Lima EJS (2013) Respiratory rate as a predictor of weaning failure from mechanical ventilation. Revista Brasileira de Anestesiologia 63:7–12

9. Cardoso FS, Karvellas CJ, Kneteman NM, Meeberg G, Fidalgo P, Bagshaw SM (2014) Respiratory rate at intensive care unit discharge after liver transplant is an independent risk factor for intensive care unit readmission within the same hospital stay: A nested case-control study. J Crit Care 29(5):791–796

10. Rambaud-Althaus C, Althaus F, Genton B, D'Acremont V (2015) Clinical features for diagnosis of pneumonia in children younger than 5 years: a systematic review and meta-analysis. Lancet Infect Dis 15(4):439–450

11. Maharaj R, Raffaele I, Wendon J (2015) Rapid response systems: a systematic review and meta-analysis. Crit Care 19(1):254

12. Jimnez D, Lobo JL, Barrios D, Prandoni P, Yusen RD (2016) Risk stratification of patients with acute symptomatic pulmonary embolism, vol 11

13. Mlgaard RR, Larsen P, H°akonsen SJ (2016) Effectiveness of respiratory rates in determining clinical deterioration: a systematic review protocol. JBI Evid Synth 14(7):19–27

14. Hochhausen N, Barbosa Pereira C, Leonhardt S, Rossaint R, Czaplik M (2018) Estimating respiratory rate in post-anesthesia care unit patients using infrared thermography: An observational study. Sensors 18(5):1618

15. Kelly C (2018) Respiratory rate 1: why measurement and recording are crucial. Nursing times 114(4):23–24

16. Zhao F, Li M, Qian Y, Tsien JZ (2013) Remote measurements of heart and respiration rates for telemedicine. Plos One 8(10):1–14

17. Xia J, Siochi RA (2012) A real-time respiratory motion monitoring system using kinect: Proof of concept. Med Phys 39(5):2682–2685

18. Wijenayake U, Park S-Y (2017) Real-time external respiratory motion measuring technique using an rgb-d camera and principal component analysis. Sensors 17(8):1840

19. Rehouma H, Noumeir R, Bouachir W, Jouvet P, Essouri S (2018) 3d imaging system for respiratory monitoring in pediatric intensive care environment. Comput Med Imaging Graph 70:17–28

20. Rehouma H, Noumeir R, Essouri S, Jouvet P (2020) Quantitative assessment of spontaneous breathing in children: evaluation of a depth camera system. IEEE Trans Instrum Meas 69:4955–4967

21. Sanyal S, Nundy KK (2018) Algorithms for monitoring heart rate and respiratory rate from the video of a user's face. IEEE J Transl Eng Health Med 6:1–11

22. Massaroni C, Lo Presti D, Formica D, Silvestri S, Schena E (2019) Non-contact monitoring of breathing pattern and respiratory rate via rgb signal measurement. Sensors 19(12):2758

23. Harte JM, Golby CK, Acosta J, Nash EF, Kiraci E, Williams MA, Arvanitis TN, Naidu B (2016) Chest wall motion analysis in healthy volunteers and adults with cystic fibrosis using a novel kinectbased motion tracking system. Med Biol Eng Comput 54(11):1631–1640

24. Janssen R, Wang W, Moc¸o A, de Haan G (2015) Video-based respiration monitoring with automatic region of interest detection. Phys Meas 37(1):100–114

25. Al-Naji A, Chahl J (2016) Remote respiratory monitoring system based on developing motion magnification technique. Biomed Signal Process Control 29:1–10

26. Brieva J, Ponce H, Moya-Albor E (2020) Non-contact breathing rate monitoring system using a magnification technique and convolutional networks. In: Romero E, Lepore N, Brieva J (eds) 15th international symposium on medical information processing and analysis, vol 11330. International Society for Optics and Photonics. SPIE, pp 181–189

27. Wang CW, Hunter A, Gravill N, Matusiewicz S (2014) Unconstrained video monitoring of breathing behavior and application to diagnosis of sleep apnea. IEEE Trans Biomed Eng 61(2):396–404

28. Lee J, Yoo SK (2020) Radar-based detection of respiration rate with adaptive harmonic quefrency selection. Sensors 6:1607

29. Phokela KK, Naik V (2020) Use of smartphone's headset microphone to estimate the rate of respiration. In: Proceedings of the 2020 international conference on communication systems & networks (COMSNETS), Bangalore, India, pp 64–69

30. Reyes MEP, Dorta_Palmero J, Diaz JL, Aragon E, Taboada-Crispi A (2019) Computer vision-based estimation of respiration signals. In: Proceedings of the Latin American conference on biomedical engineering, Cancú,n, Mexico, pp 252–261

31. Schoun B, Transue S, Halbower AC, Choi M-H (2018) Non-contact tidal volume measurement through thin medium thermal imaging. Smart Health 9:37–49

32. Nosrati M, Shahsavari S, Lee S, Wang H, Tavassolian N (2019) A concurrent dual-beam phased-array Doppler radar using MIMO beamforming techniques for short-range vital-signs monitoring. IEEE Trans Antennas Propag 67:2390–2404

33. Nam Y, Kong Y, Reyes B, Reljin N, Chon KH (2016) Monitoring of heart and breathing rates using dual cameras on a smartphone, vol 11

34. GE Healthcare (2013) Dash 3000/4000/5000 TM Patient Monitor: Service Manual. (Accessed: 22.05.2020). [Online]. Available: https://monetmedical.com/wp-content/uploads/2017/11/GEDASH-5000-Service-Manual.pdf

Yu-Ching Lee is an research assistant and Ph.D. candidate of Graduate Institute of Applied science and Technology at National Taiwan University of Science and Technology under Prof Ching-Wei Wang's supervision. Her current research area includes deep learning, medical image analysis, automatic registration, detection, segmentation and classification of X-ray image and multi-dimensional microscopic image.

Abdan Syakura received the B.Sc degree in physics from University of Indonesia, Depok, Indonesia and the M.Sc degree in Biomedical Engineering from National Taiwan University of Science and Technology, Taipei, Taiwan. He is currently an research assistant and PhD student in the Graduate Institute of Applied Science and Technology at National Taiwan University of Science and Technology under Prof Ching-Wei Wang's supervision. His current research area focuses on image registration and medical image analysis.

Muhammad Adil Khalil received the B.Sc degree in Electrical Engineering from Comsats University of Science and Technology, Islamabad, Pakistan and is currently an research assistant and graduate student at National Taiwan University of Science and Technology under Prof Ching-Wei Wang's supervision. His research interests are automatic segmentation, classification and image registration of medical images.
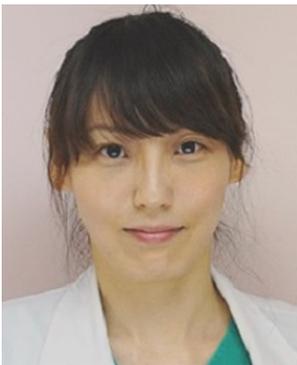
Ching-Ho Wu received the Ph.D. in veterinary medicine from National Taiwan University, and she is the director of the Small Animal Orthopedic and Soft Tissue Surgery Research Lab in National Taiwan University. Her research area includes small animal clinical medicine, orthopedic biomechanics, biomedical imaging analysis, small animal urology and regeneration medicine.

Yi-Fang Ding is an attending physician of Otorhinolaryngology in Taipei Municipal Wanfang Hospital and a Ph. D. candidate of Graduate Institute of Medical Sciences, College of Medicine, Taipei Medical University, Taipei, Taiwan. He specializes in head and neck surgery. In addition, he is at the forefront of his field utilizing sialoendoscopy in Taiwan. His research interests include salivary gland disease and treatment.

Ching-Wei Wang received the M.Sc. degree (with Distinction) in computer science from University of Glasgow, UK, and the Ph.D. degree in computer science from University of Lincoln, UK. She is currently an Professor in the Graduate Institute of Biomedical Engineering, National Taiwan University of Science and Technology, Taiwan. Her research areas include deep learning, medical image analysis, automatic registration, detection, segmentation, recognition and reconstruction for multi-dimensional microscopic image, 3D/4D Cardiac MRI Image, Ultrasound Image, 3D confocal image, 3D CT scan, X-ray image, live cell image, infrared image/video and gigapixel/terapixel ultra-high resolution medical image.