1. Respiration Monitoring by Video Signal Processing

V. Mattioli[1], D. Alinovi[1,3], F. Pisani[2], G. Ferrari[1], and R. Raheli[1]

[1]Department of Engineering and Architecture, University of Parma, Parco Area delle Scienze 181/A, IT-43124 Parma, Italy
[2]Department of Medicine and Surgery, University of Parma, Via Gramsci 14, IT-43126 Parma, Italy
[3]Deceased on 16 September 2020

Abstract: *Respiration is a fundamental human body function that may provide useful information about the clinical status of a patient. Suitable and continuous monitoring of the Respiratory Rate (RR) is thus essential to promptly detect anomalies that may be signs of potentially harmful or life-threatening disorders. However, traditional RR monitoring systems consist of sensors and devices which are often expensive, invasive, and can be deployed only in hospital settings, requiring trained medical staff. In this chapter, an overview of alternative low-cost and non-invasive video-based methods for respiration monitoring is presented along with a brief review of earlier work. Principles underlying the extraction of relevant information content from video signals are described and specific video-based solutions for newborn and adult monitoring are presented. Modelling and simulation of breathing patterns is also addressed. The performance of the proposed solutions is finally discussed also on the basis of experimental results.*

1.1 Introduction

Respiration is governed by the respiratory system, whose main function is gas exchange occurring in two phases: inhalation and exhalation. The sequence of these two movements is known as breathing or respiratory cycle. During inhalation, oxygen (O_2) is inspired through the nose or mouth and absorbed into the blood that diffuses it to the body cells for their nourishment [1]. On the other hand, during exhalation, carbon dioxide (CO_2), a waste product of cellular respiration, is released outside [2]. The exchange of these gases occurs in a specific region of the lungs: the alveolar-capillary membrane. During inspiration, the muscle of the diaphragm, positioned in the lower part of the thorax, flattens and contracts while the external intercostal muscles elevate the ribs, thus expanding the thoracic cavity and increasing its volume [1]. As a consequence of this volumetric expansion, the pressure into the thorax decreases compared to the atmospheric pressure and this difference helps to move air into the lungs, as it flows from higher pressure regions to lower ones, inducing inhalation. On the other hand, during expiration, the diaphragm and external intercostal muscles relax and return to the equilibrium position causing a reduction of the volume of the thorax cavity – a new pressure difference creates and is then equalized with the exhalation of air [1]. The

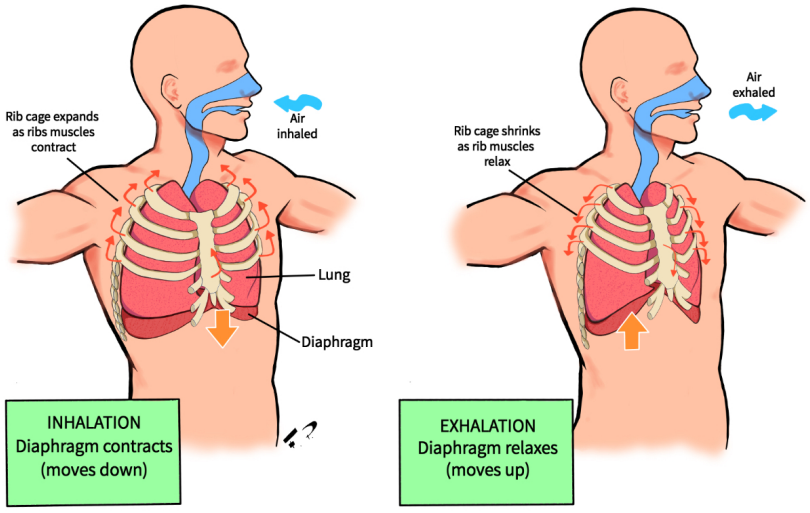


Figure 1.1: Respiration mechanism.[1]

respiration mechanism is shown in Figure 1.1, where inhalation and exhalation phases are depicted.

The Respiratory Rate (RR) is one of the main physiological parameter able to provide important information about the clinical status of a patient. It is clinically defined as the number of breathing cycles per time unit and is typically measured in breath per minutes (bpm) or in breathing cycles per seconds, namely Hertz (Hz). A normal RR range for an adult at rest lies between 12 and 20 bpm but can slightly vary depending on age and medical condition [3]. Correct and continuous monitoring of this vital sign allows to detect anomalies and changes that can be related to potentially severe disorders. It is also demonstrated that RR can be an important early predictor of potentially life-threatening events (e.g., cardiac arrest), and can be used to asses the progression of illness [4]. Despite its medical relevance, RR is often not recorded or inaccurately measured in clinical environments where nursing staff tends to prioritize pulse oximetry to assess respiratory function [3], [5]. The lack of reliable automated instrumentation might be among the main causes of poor RR monitoring [3], [6].

1.1.1 Respiratory Disorders

Among respiratory diseases, apneas are potential life-threatening events, especially for newborns. They are clinically defined as sudden cessations of respiration lasting longer than 20 seconds. Apneas or episodes of breathing absence lasting between 10 s and 20 s associated with other clinical signs/symptoms such as bradycardia, pallor, cyanosis, and/or marked hypotonia are more common in preterm newborns [7] with an incidence of almost 100% for newborns younger than 28 weeks of gestational age. Nevertheless,

[1] Image by Francesca and Andrea Pisani.

the incidence of apneas reduces with age and is estimated to be around 50% between 33 and 34 weeks of gestational age and 1‰ for newborns at term [8]. Apnea in preterm neonates is usually referred to as Apnea of Prematurity (AoP) and is probably due to the physiological immaturity of the central nervous system and poor development of the respiratory system [8], [9]. Among secondary causes of apneas, neurological, pulmonary, and cardiac pathologies need to be mentioned [9]. Apnea events are usually classified in three main categories based on the absence or presence of obstructions of the upper airway. Central apnea consists of a total cessation of the respiratory effort with no evidence of obstruction. On the other hand, obstructive apnea occurs when the upper airway is occluded causing no air to flow. Obstructive Sleep Apnea (OSA) represents a common cause of obstructive apneas in adults [10]. Finally, mixed apnea can be defined as a combination of the two [9].

Apneas can also be triggered by a congenital disease such as Congenital Central Hypoventilation Syndrome (CCHS), a rare and possibly life-threatening neurological disorder characterized by failure of automatic control of ventilation in the central nervous system [11]. CCHS is caused by an alteration in the PHOX2B homeobox gene and usually manifests during sleep, being one of the main causes of alveolar hypoventilation [12]. Other conditions associated with this disorder include Hirschsprung's disease, neural tumors, arrhythmias and minor ocular abnormalities [12].

Prompt detection of such life-threatening events and suitable monitoring of their progression is crucial for early treatment. In the next section some conventional techniques for respiration monitoring will be presented along with more innovative methods based on video processing.

1.2 Respiration Monitoring

Monitoring of vital signs is fundamental for early detection and treatment of several disorders. RR is an informative parameter about the clinical status of a patient and is considered to be an important forecaster of potentially deadly events such as cardiac arrest and intensive care unit admission [5]. Besides, its accurate monitoring is essential in clinical settings and its relevance may be appreciated also in occupational settings and sport fields [13]. The respiratory frequency may indeed be affected by cognitive and training load, as discussed in [14] and [15], respectively.

Respiration monitoring techniques usually divide into contact-based and contact-less categories. Conventional instrumentation typically employed in hospital environments mainly belongs to the first category, whereas innovative approaches based on video processing techniques represent contact-less solutions of significant potential interest. An overview of existing methods to monitor respiration is presented hereafter.

1.2.1 Contact-based Techniques

Contact-based techniques for respiration monitoring require the direct contact of a sensor with the body of the patient, being often moderately invasive and uncomfortable, especially for newborns. This kind of medical equipment is typically expensive and mostly employed in clinical settings, making home care challenging and hardly accessible.

A categorization of contact-based monitoring may be proposed on the basis of physiological parameters to be investigated and the part of the body involved in the measurement [4], [16]. Seven main categories based on the parameters of interest can be identified: respiratory airflow, respiratory sound, air temperature, air humidity, air components, chest wall movements, and modulation of the cardiac activity [13].

Several sensors can be employed to characterize the respiratory airflow, recording temporal trends during respiration, i.e., volume and/or velocity of the inhaled and exhaled air: this allows to estimate the RR. Differential flowmeters, turbine flowmeters, hot wire anemometers, and fiber optic-based flowmeters represent the most popular devices in this category and may be embedded in medical probes such as spirometers. The main drawback of this instrumentation is its intrusiveness, a major concern especially for newborns or critically ill patients [16].

Acoustic sensors can also be exploited to estimate the respiratory airflow by monitoring the sound emitted during the inhalation and exhalation phases [1]. Microphones may be used to record sounds originating at different body locations, especially the chest, but their performance is highly affected by environmental noise.

The difference between inhaled and exhaled air in terms of temperature, humidity, and air components may also lead to an estimate of the RR. Inhaled air is at environmental conditions while exhaled air is at body temperature, saturated with vapor and containing a higher level of CO_2. Temperature, humidity, and chemical sensors may be employed to monitor the corresponding parameters of interest, being usually embedded in other devices such as face masks. As they need to be placed near the nostrils and/or lips, intrusiveness is a recurrent problem [16].

As chest and abdominal muscles are involved in the respiration process, monitoring their movements may reveal potential anomalies [17]. Strain, impedance, and movement measurement may be performed by proper sensors [16].

Finally, the cardiac activity may be recorded to extract the respiratory frequency. Electrocardiography and photoplethysmography are the most popular techniques in this case. In particular, the former allows to measure the heart electrical activity through electrodes, while the latter exploits properties of light reflection and transmission to measure volumetric changes in blood vessels associated with the cardiac cycle. Both methods are sensitive to motion artefacts [16].

Electrocardiography can also be employed for sleep monitoring in order to detect apnea events or other sleep disorders. Polysomnography is the main tool to investigate the sleep of a patient and is composed of several systems, namely: ElectroEncephaloGram (EEG), ElectroOculoGram (EOG), ElectroMyoGram (EMG) and ElectroCardioGram (ECG) [18]. Wired electrodes are placed on different parts of the body of a patient to allow measurements of different physiological parameters. In particular, brain activity, eyes movements, muscle activity, and heart rate are measured by electrodes attached to the scalp, near the eyes, under the chin and on the chest, respectively. Elastic chest belts, nasal probes, and pulse oximeters may also be employed to provide extra information such as effort to breathe, airflow and oxygen saturation [18]. Figure 1.2 depicts all systems involved in the polysomnogramn test.

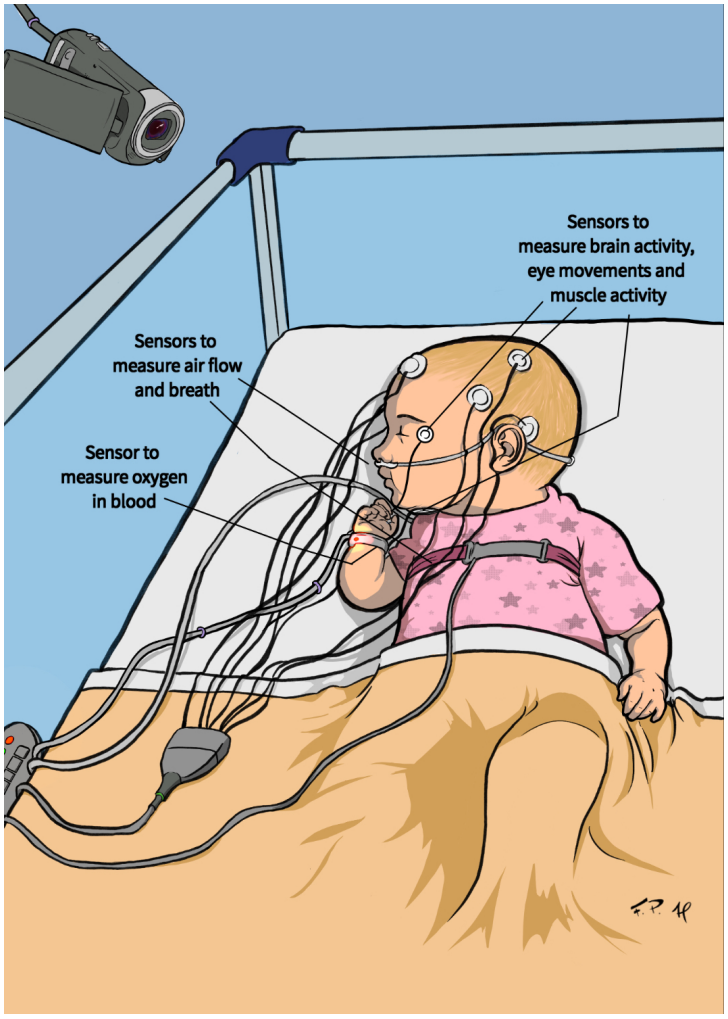


Figure 1.2: Schematic overview of the polysomnogramn test.[1]

1.2.2 Contact-less Techniques

Contact-less techniques for vital signs monitoring present several advantages. In particular, they do not require a direct contact with the body of a patient, thus being non-invasive and non-obstructive. Also, costs may be lower making home care more accessible. Thermal cameras, depth sensors, radars, and Red, Green and Blue (RGB) cameras can be considered among the most popular contact-less devices [13].

Thermal cameras allow to monitor the air temperature at a specific Region Of Interest (ROI) considered near the nostrils/mouth of a patient. A difference in terms of temperature can indeed be observed in the inhaled and exhaled air. Temperature changes induce pixel intensity changes in thermal images, hence breathing information, such as RR, can be obtained by analysing pixel intensity variations in a sequence of thermal images. A

novel promising method to RR extraction by thermography is proposed in [19].

On the other hand, depth-sensing system may be designed to detect volumetric changes of the rib cage to extract RR. Triangulation methods are usually employed to construct depth maps as in [20], where a system composed of a transmitter emitting a light source with known properties and a camera capturing its reflection is presented.

Radar sensors can also be exploited to monitor RR by detecting chest movements associated with the respiration. Radio waves or microwaves can be transmitted in pulsed or continuous mode towards a reflective subject and the back- or forward-scattered signal is measured to detect breathing movements. In [21], a forward system sensitive to changes in the radar cross-section is proposed.

Finally, RGB cameras may also allow to estimate RR by detecting pixel intensity variations in video sequences caused by periodic breathing movements of the chest and abdomen of the framed patient. Some camera-based techniques to RR monitoring will be detailed in Section 1.3 where principles underlying the extraction of relevant information content from video signals are described and specific video-based solutions for newborn and adult monitoring are presented.

1.3 Video Processing Systems for Respiration Monitoring

As respiration induces periodic movements of the chest associated with the inhalation and exhalation phase, RGB cameras may be employed to capture motion signals to be properly processed in order to obtain an estimate of the RR. However, respiration movements are usually subtle and difficult do detect. To amplify such small changes, Eulerian Video Magnification (EVM) algorithms [22] may be exploited to preprocess the video signals before further processing. A few methods which rely on motion magnification algorithms to empathize small respiration movements for estimation purposes, are detailed hereafter.

We first provide some preliminary definitions about video signals. A video signal captured with a sampling time T is defined as a multidimensional signal and can be described by a sequence of matrices. The frame rate is defined as $f_s = 1/T$ and is usually measured in Hz or frames per second (fps). Frames of RGB videos are described by three matrices, one per color channel, where each element represents the intensity value of the color pixel in that position. On the other hand, a gray-scale frame can be described by a single sequence of two-dimensional matrices $X[x, y, n]$ or, equivalently, by $\mathbf{X}[n]$, where $[x, y]$ is the pixel position within the frame and n is the frame index. The matrix size corresponds to the frame dimension $H \times W$, where W is the number of pixel rows and H is the number of pixel columns, respectively.

1.3.1 Motion Information Extraction

A sequence of standard image processing operations can be performed to extract motion signals from video streams [23, 24]. An illustrative overview of the motion extraction process is shown in Figure 1.3, where four main operations can be identified. Considering a gray-scale video sequence, a filtering operation based on the Difference of consecutive

Frames (DoF) is initially implemented to highlight moving parts in the output sequence $\mathbf{D}[n]$. The result is then thresholded to obtain a binary sequence $\mathbf{B}[n]$ where white pixels, having intensity equal to 1, correspond to foreground regions affected by motion. Morphological operations [25], such as erosion, can now be applied to further reduce noise. The eroded sequence is denoted as $\mathbf{I}[n]$.

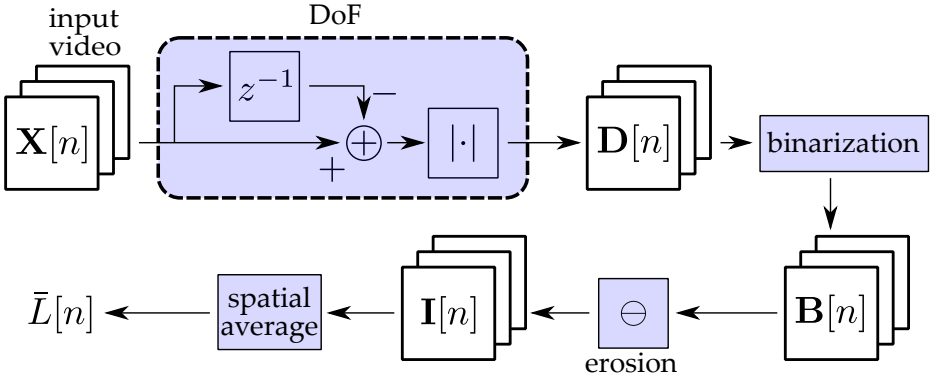


Figure 1.3: Motion signal extraction.

Finally, a suitable signal is defined by a spatial average operation as follows:

$$\bar{L}[n] = \frac{1}{WH} \sum_{w=0}^{W-1} \sum_{h=0}^{H-1} I[x, y, n] \tag{1.1}$$

where $\bar{L}[n]$ is the average luminance signal which describes the motion patterns of the framed moving subjects.

1.3.2 Motion Magnification

As respiration-related movements are subtle and may be difficult to detect, a motion magnification algorithm may be applied to the input video sequence $\mathbf{X}[n]$ to previously process it in order to amplify such small changes. The illustrative overview of the EVM algorithm proposed in [22] and also employed in [23, 24] is shown in Figure 1.4, where four main phases are highlighted. Each frame of the input sequence is initially spatially decomposed into L different spatial frequency bands. A temporal filtering is then implemented on each spatial band: bandpass filters are considered here to extract frequency bands corresponding to normal ranges of respiration frequencies. Output signals are enhanced by proper amplification factors and the obtained amplified levels are finally recomposed to reconstruct a new video sequence where small movements are emphasized. The motion extraction algorithm presented in Section 1.3.1 can be applied to the enhanced video sequence. According to this procedure, a video sequence is processed in two steps: first, motion magnification is performed as schown in Figure 1.4; then, the amplified output is processed as shown in Figure 1.3 to extract the motion signal. In [26], a solution based on spatio-temporal processing techniques to integrate EVM and motion extraction algorithms is proposed hereafter. A block diagram is provided in Figure 1.5. The main

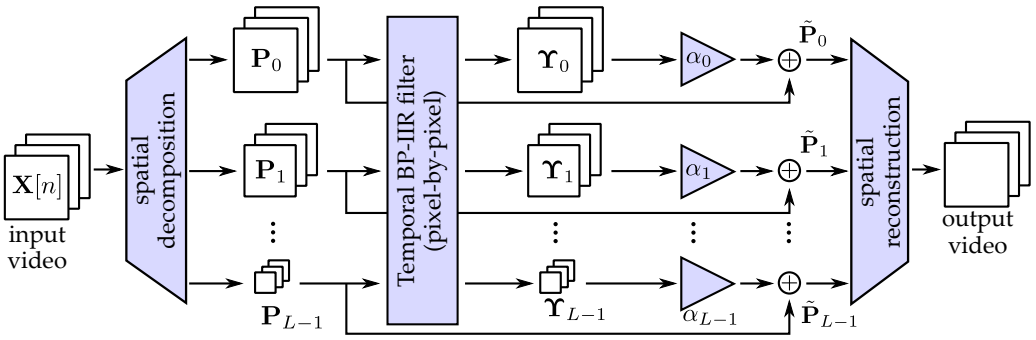


Figure 1.4: EVM algorithm. Modified from an original image published in [23].

novelty introduced with respect to the EVM algorithm presented in [22], is that the final reconstruction of the output video is not performed. Each amplified level is instead directly processed to extract motion signals. To this purpose, each level is binarized by setting a proper threshold and motion signals $\{\bar{L}_\ell[n]\}$ for all levels are finally computed as in (1.1), where $\ell = 0, \ldots, L-1$, being L the total number of levels. Note that the DoF operation is not performed here; only band-pass filters included in the EVM algorithm are exploited. Erosion may be avoided too.

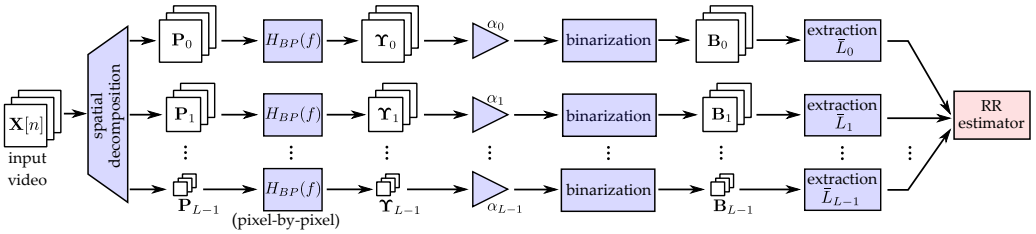


Figure 1.5: Spatio-temporal estimation algorithm. Modified from an original image published in [26].

1.3.3 Respiratory Rate Estimation

Once motion signals are extracted, RR can be computed according to well-known estimation techniques. The Maximum Likelihood (ML) criterion [27] is employed in [23, 24, 26]. Considering the spatio-temporal video processing proposed in [26], the extracted L motion signals, associated with periodic breathing movements, can be modeled as periodic signals as follows:

$$\bar{L}_\ell[n] = c_\ell + A_\ell \cos\left(2\pi f_0 n T + \phi_\ell\right) + w_\ell[n] \quad \ell = 0, \ldots, L-1 \tag{1.2}$$

where: c_ℓ and $w_\ell[n]$ are the continuous components and sequences of independent and identically distributed (i.i.d.) zero-mean Gaussian noise samples, respectively; A_ℓ, f_0 and ϕ_ℓ are unknown parameters representing the amplitudes, fundamental frequency common

to all signals, and phases, respectively, and may be collected in a vector $\boldsymbol{\theta}_\ell = [A_\ell, f_0, \phi_\ell]$ to be estimated. In particular, f_0 corresponds exactly to the RR. Following the general approach to ML estimation described in [27], the fundamental frequency f_0, relative to the ℓ-th level, can be estimated by minimizing the following likelihood function:

$$J_\ell(\boldsymbol{\theta}_\ell) = \sum_{n=0}^{N-1} \left(\bar{L}_\ell[n] - A_\ell \cos\left(2\pi f_0 nT + \phi_\ell\right) \right)^2 \tag{1.3}$$

where N is the number of observed frames.
Data fusion may now be exploited in order to extend the ML estimator to multiple observations, thus obtaining a reinforced estimate of f_0 [26]. Defining $\boldsymbol{\theta} = [\mathbf{A}, f_0, \boldsymbol{\phi}]$ as the vector of the aggregated parameters, where $\mathbf{A} = [A_0\, A_1\, \ldots\, A_{L-1}]$ and $\boldsymbol{\phi} = [\phi_0\, \phi_1\, \ldots\, \phi_{L-1}]$, the following joint likelihood function can be obtained:

$$J\left(\boldsymbol{\theta}\right) = \sum_{\ell=0}^{L-1} J_\ell\left(\boldsymbol{\theta}_\ell\right). \tag{1.4}$$

After some mathematical simplifications, the following RR estimator can be obtained:

$$\begin{aligned} \hat{f}_0 &= \underset{f}{\operatorname{argmax}} \sum_{\ell=0}^{L-1} \left| \sum_{n=0}^{N-1} \bar{L}_\ell[n] e^{-j2\pi f nT} \right|^2 \\ &= \underset{f}{\operatorname{argmax}} \sum_{\ell=0}^{L-1} \left| \mathrm{DFT}\{\bar{L}_\ell[n]\} \right|^2 \end{aligned} \tag{1.5}$$

where $\mathrm{DFT}\{\cdot\}$ represents the Discrete Fourier Transform (DFT) operator. A periodic component is finally declared significant if the following constraint is verified:

$$\frac{N}{L} \sum_{\ell=0}^{L-1} \hat{A}_\ell^2 > \eta \tag{1.6}$$

where η is a properly set threshold and estimates of the amplitudes $\{\hat{A}_\ell\}$ are obtained as:

$$\hat{A}_\ell = \frac{2}{N} \left| \sum_{n=0}^{N-1} \bar{L}_\ell[n] e^{-j2\pi \hat{f}_0 nT} \right| \quad \ell = 0, \ldots, L-1. \tag{1.7}$$

1.3.4 Pixel-wise Maximum Likelihood Estimation

A third approach to RR estimation that does not rely on motion magnification is proposed in [28], where a generalized model of pixel-wise periodicity is introduced and the ML criterion is directly applied. A representative overview of the considered method is shown in Figure 1.6, where a few main steps are highlighted. Indeed, periodic respiration movements correspond to periodic variations of the pixel intensity values in a considered video sequence. In particular, the algorithm aims at analysing periodic variations of the pixel intensity values that correspond to periodic respiratory movements. To this end,

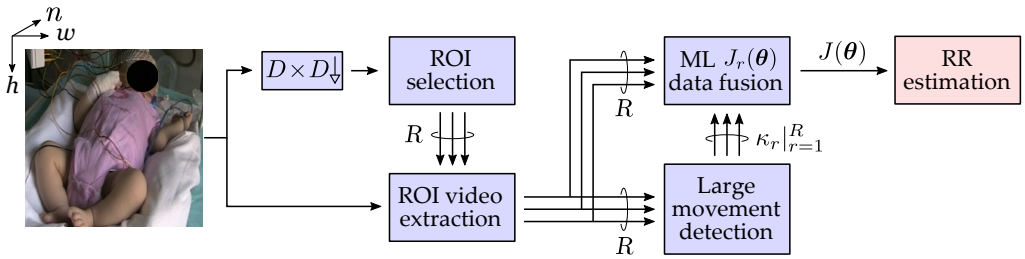


Figure 1.6: Pixel-wise ML estimation algorithm. Modified from an original image published in [28].

R ROIs of $M \times M$ pixel are first extracted from selected blocks of frames to select areas where only movements associated with the respiration appear, thus excluding areas possibly affected by other body movements. Selection and extraction operations are highlighted at the corresponding blocks in Figure 1.6. In order to reduce the computational complexity of the algorithm, a spatial decimation factor D is also introduced to decrease the size of the analysed frames. A general ML approach is then applied to each ROI to estimate the RR.

In particular, pixel-wise variations can be modeled as:

$$\mathbf{X}[n] = \mathbf{C} + \mathbf{A}\cos\left(2\pi f_0 T n + \mathbf{\Phi}\right) + \mathbf{W}[n] \tag{1.8}$$

where matrices $\mathbf{C}$ and $\mathbf{W}$ have dimension $H \times W$ and denote continuous components and i.i.d. zero-mean Gaussian noise samples respectively. Similarly to (1.2), matrices $\mathbf{A}$ and $\mathbf{\Phi}$ collect amplitudes and initial phases, whereas f_0 is the common fundamental frequency that corresponds to the RR to be estimated. By Vectorizing the matrices $\mathbf{X}[n]$, $\mathbf{A}$ and $\mathbf{\Phi}$, as described in [28], vectors $\mathbf{x}_{\mathrm{v}[n]}$, $\mathbf{a}_{\mathrm{v}}$ and $\boldsymbol{\phi}_{\mathrm{v}}$ are obtained. Considering the whole frame, a new likelihood function to be minimized, with respect to the vector $\boldsymbol{\theta} = [\mathbf{a}_{\mathrm{v}}, f_0, \boldsymbol{\phi}_{\mathrm{v}}]$, can be expressed as:

$$J\left(\boldsymbol{\theta}\right) = \sum_{p=0}^{W\times H-1} \sum_{n=0}^{N-1} \left[x_{\mathrm{v}}[p,n] - a_{\mathrm{v}}[n]\cos\left(2\pi f_0 n T + \phi_{\mathrm{v}}[p]\right)\right]^2 \tag{1.9}$$

where p represents the p-th element of each vector. Finally, the following estimator $\hat{f}_0$ of the fundamental frequency can be derived:

$$\hat{f}_0 = \underset{f}{\operatorname{argmax}} \sum_{p=0}^{W\times H-1} \left| \sum_{n=0}^{N-1} x_{\mathrm{v}}\left[p,n\right] e^{-j2\pi f n T} \right|^2 . \tag{1.10}$$

The amplitudes can then be estimated as follows:

$$\hat{a}_{\mathrm{v}}[p] = \frac{2}{N} \left| \sum_{n=0}^{N-1} x_{\mathrm{v}}[p,n] e^{-j2\pi \hat{f}_0 n T} \right| . \tag{1.11}$$

To conclude, as discussed in Section 1.3.3, data fusion may be performed and a joint

likelihood function can be obtained as follows:

$$J(\boldsymbol{\theta}) = \sum_{r=1}^{R} \kappa_r J_r(\boldsymbol{\theta}). \tag{1.12}$$

where: κ_r is a properly defined parameter that indicates the presence or the absence of large motion in the considered ROI; and J_r is a likelihood function formulated as in (1.9) referred to the r-th ROI [28]. In particular, the large movement detection operation, highlighted at the corresponding block of Figure 1.6, includes the setting of κ_r in order to reduce undesired effect due to large body movements during the estimation phase.

1.4 Modeling and Simulation of Breathing Patterns

In order to test and devise reliable video-based systems for RR estimation, simulators of breathing patterns and disorders may be implemented to generate realistic video recordings in order to extend already existing and inherently limited poor databases. In [29], a Continuous-Time Markov Chain (CTMC) statistical model of breathing behaviours is presented and two simulators (software and hardware) based on this model are proposed.

1.4.1 Model of breathing patterns

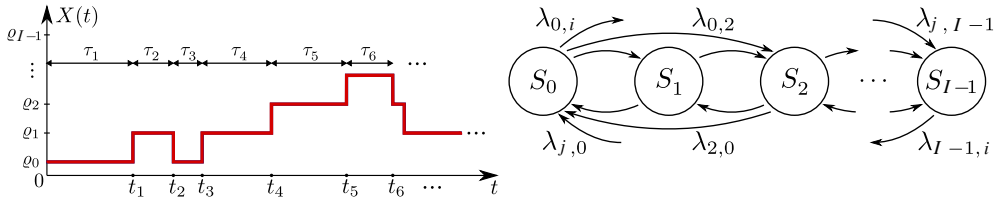
According to [29], a RR random process can be defined as a multi-state continuous-time process where $S = \{S_0, \dots S_{I-1}\}$ is a finite set of states, being I the total number of states. The i-th state S_i ($i = 0 \dots I-1$) represents an event during which the RR has value ϱ_i with $\varrho_0 < \dots < \varrho_{I-1}$. States are properly assigned accounting for the possible presence of apnea events and large random movements. In particular, states are selected as follows.

- If the subject is breathing with a regular RR, states $\{S_0, S_1, \dots, S_{N-1}\}$ are associated with regular RRs with values $\{\varrho_n\}_{n=0}^{N-1}$ where R_H and R_L represent the highest and lowest admissible RRs, respectively, and $R_L < \varrho_n < R_H$, $\forall n \in 0, \dots N-1$.

- If an interruption of respiration, due to an apnea episode or respiratory pause, manifests, the state S_0 is associated with this situation and ϱ_0 is set to 0 to indicate the lack of respiration, whereas the remaining states $\{S_1, S_2, \dots, S_{N-1}\}$ are associated with regular respiration.

- If the RR is undetectable, due to the presence of large body movements, the state S_{N-1} is associated with this situation, ϱ_{N-1} is set to an arbitrary value much larger than R_H, whereas the remaining states $\{S_0, S_1, \dots, S_{N-2}\}$ are associated with regular respiration.

- If the subject is both affected by a lack of respiration and large body movements, states S_0 (with $\varrho_0 = 0$) and S_{N-1} (with $\varrho_{N-1} \gg R_H$) are associated with these situations, whereas the remaining states $\{S_1, S_2, \dots, S_{N-2}\}$ are associated with regular respiration.



(a) (b)

Figure 1.7: (a): RR process modeled as a CTMC; (b): CTMC state diagram. Modified from original images published in [29].

The RR process, denoted as $X(t)$, can be described by a CTMC characterized by time intervals representing the durations of RR events, i.e., sojourn times spent by the process in each state. An example of RR modeled as a CTMC (RR events correspond to sojourn times in the possible states) is shown in Figure 1.7(a), where sojourn times in states and jump times between two consecutive states are denoted as τ_q and t_q respectively, with q denoting an index that counts the number of state changes. Conditionally on state S_i, waiting times can be considered as independent random variables exponentially distributed as $\tau_q \sim \exp(\mu_i)$. An example of the CTMC state diagram is shown in Figure 1.7(b), where the infinitesimal transition rate from S_j to S_i $(i, j \in 0, \ldots, I-1,\ i \neq j)$ is denoted as $\lambda_{j,i}$. The transition rates $\{\lambda_{j,i}\}$ are the elements of the following generator matrix $\mathbf{\Lambda}$:

$$\mathbf{\Lambda} = \begin{bmatrix} \lambda_{0,0} & \lambda_{0,1} & \cdots & \lambda_{0,I-1} \\ \lambda_{1,0} & \lambda_{1,1} & \cdots & \lambda_{1,I-1} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{I-1,0} & \lambda_{I-1,1} & \cdots & \lambda_{I-1,I-1} \end{bmatrix} \tag{1.13}$$

where $\lambda_{j,i} = -\sum_{j \neq i} \lambda_{j,i}$. Each transition rate $\lambda_{j,i}$ can be obtained by ML estimation as follows [29]:

$$\hat{\lambda}_{j,i} = \frac{N_{j,i}(T)}{R_i(T)} \quad j \neq i, \quad i, j \in 0, \ldots, I-1 \tag{1.14}$$

$$\mu_i = -\lambda_{i,i} = \sum_{\substack{j \neq i \\ j \in 0,\ldots,I-1}} \lambda_{j,i} \quad i \in 0, \ldots, I-1. \tag{1.15}$$

where $\{R_i(T)\}$ and $\{N_{j,i}(T)\}$ are random variables that denote the time spent by the RR process in state S_i and the number of transitions from state S_j to state S_i in an observation interval $[0, T]$, respectively. The estimator $\hat{\mathbf{\Lambda}} = \hat{\lambda}_{j,i}$ of the generator matrix can be accordingly obtained.

1.4.2 Simulators

Once the CTMC is fully characterized by $\hat{\mathbf{\Lambda}}$, it can be employed to implement simulators of breathing behaviours. The software simulator presented in [29] may be applied to

recorded video signals where framed patients breath with a normal RR, approximated as time-invariant and denoted as ϱ_V (dimension: [bpm]). Interpolation or decimation of video frames can be performed to alter breathing movements by accelerating or slowing them down, thus creating new artificial video sequences. In particular, normalized values of RR for each state are defined as:

$$\bar{\varrho}_i = \frac{\varrho_i}{\varrho_V} \qquad i = 0, 1, \ldots, I-1. \tag{1.16}$$

A scaling factor C_V may be introduced in the case where all rates $\{\varrho_i\}$ differ from the original rate ϱ_V.

The duration (expressed in terms of number of frames) of the breathing times to be inserted in the new artificial video sequence can be computed from the generated sojourn and jump times as:

$$\tilde{\tau}_q = \text{round}(\tau_q f_s) \tag{1.17}$$

$$\tilde{t}_q = \text{round}(t_q f_s) \tag{1.18}$$

where round$\{\cdot\}$ is an operator that returns the integer value closest to its argument. Blocks of duration of $\tilde{d}_q$ frames are processed to simulate breathing times, with:

$$\tilde{d}_q \triangleq \left\lceil \frac{\tau_q}{\bar{\varrho}_i} f_s \right\rceil \tag{1.19}$$

where $\lceil \cdot \rceil$ represents the ceiling function. Considering now a generic state S_i ($i = 0, \ldots, I-1$), four events may happen:

- $\bar{\varrho}_i = 1$ — expressions (1.17) and (1.19) coincide and no processing is needed as the generated and original RRs coincide as well;
- $0 < \bar{\varrho}_i < 1$ — in this case, $\tilde{d}_q < \tilde{\tau}_q$ and interpolation is performed to slow breathing movements down;
- $\bar{\varrho}_i > 1$ — in this case, $\tilde{d}_q > \tilde{\tau}_q$ and decimation is performed to accelerate breathing movements;
- $\bar{\varrho}_0 = 0$ — this case is treated as if $0 < \bar{\varrho}_i < 1$, but a new value $\bar{\varrho}_0^{'}$ is selected such that $0 < \bar{\varrho}_0^{'} \ll R_L$, where R_L is the highest admissible value of RR.

As interpolation acts also on noise, a noise compensation algorithm is needed to maintain the background noise of the original sequence when $0 < \bar{\varrho}_i < 1$ as described in [29].

On the other hand, a mechanical manikin can be used to simulate breathing events via hardware. An illustrative representation of the hardware simulator described in [29] is shown in Figure 1.8, where the manikin of a newborn appears. An Arduino UNO board [30] is used to drive a servo-motor that moves part of the chest. Times and states of the derived CTMC model are used to control the movement of the chest. Apnea events can also be mimicked.

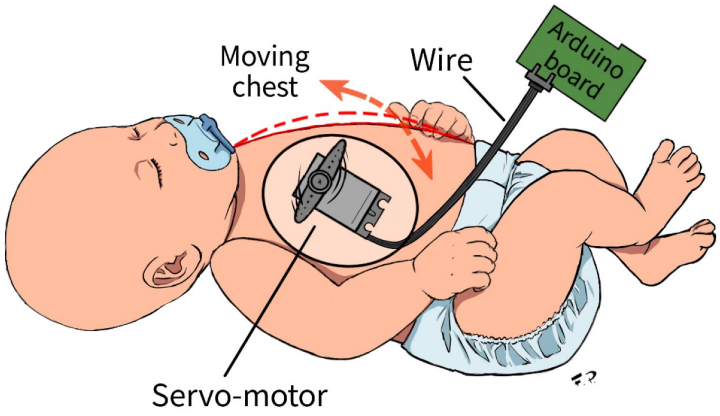


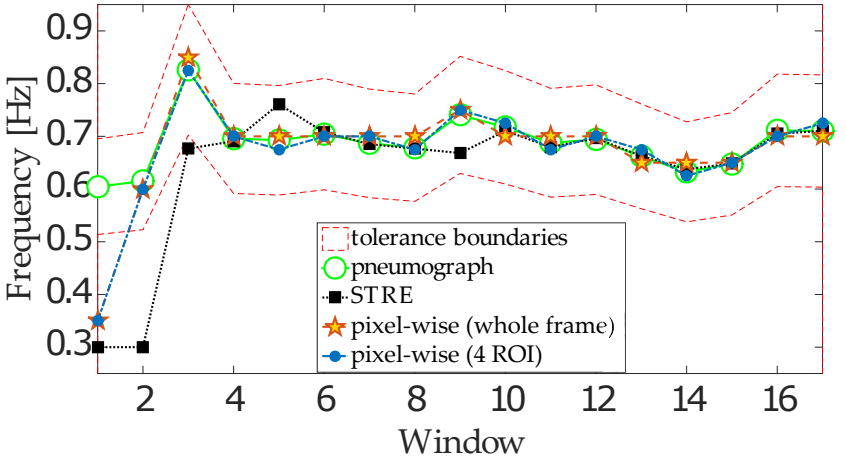
Figure 1.8: Manikin simulator of breathing events in a newborn.[1]

1.5 Application and Results

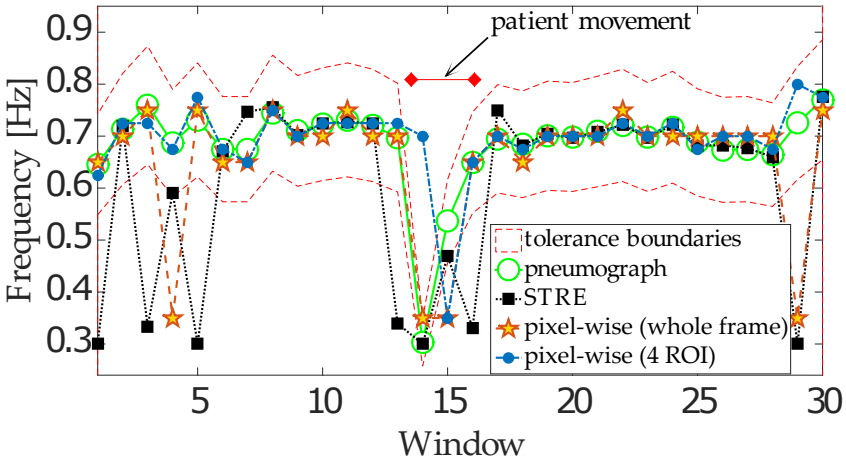
Results obtained on real videos of breathing newborns are reported in Figure 1.9, where Spatio-Temporal video-processing for RR Estimation (STRE) and pixel-wise ML algorithms are compared [28]. In particular, the estimated RR (dimension: [Hz]) is plotted for both algorithms and compared with ground truth values obtained by a pneumograph. A tolerance interval of $\pm 15\%$ with respect to the pneumographic signal is considered according to medical pratice [28]. The pixel-wise method described in Section 1.3.4 is tested on the whole frames and on 4 selected ROIs. The two Figures 1.9a and 1.9b refer to videos of short and long durations, respectively. In Figure 1.9b an interval where large body movements of the patient make the RR undetectable is visible. However, the STRE and pixel-wise methods lead, respectively, to a correct estimate of the RR: in 14 (87.5%) and 16 (100%) out of considered 16 observation windows with the short video; and 24 (82.7%) and 25 (86.2%) out of 29 considered observation windows on the long video. In the latter case, the performance is further improved by selecting 4 ROIs: the RR is correctly detected in 27 out of 29 (93%) windows. The interval where the patient is moving is not taken into account in the analysis.

Illustrative examples of acquisition systems, where multiple and single cameras are used to monitor a newborn in a hospital and a domestic setting, are shown in Figure 1.10. The cameras are placed in order to capture movements of the baby. In particular, the multi-camera system shown Figure 1.10(a) is employed in [24] for hospital monitoring, whereas the single-camera system shown in Figure 1.10(b) is employed in [31] for domestic monitoring.

For the sake of completeness, results on simulated videos are discussed hereafter. Simulators of breathing patterns may indeed be useful to assess the performance of the video-based techniques for RR estimation described in Section 1.3. As an illustrative



(a)



(b)

Figure 1.9: RR estimation by pneumograph, STRE algorithm and pixel-wise algorithm for: (a) short video signal (duration: 3 min and 4 s) and (b) long video signal (duration: 5 min and 12 s).
Image modified from an original image published in [28].

example, three different motion signals considered in [26] are shown in Figure 1.11: trace (a) is extracted from a real video sequence framing a newborn regularly breathing; trace (b) represents a respiratory pause, i.e., absence of breathing lasting less than 10 seconds, generated via software; trace (c) represents a real respiratory pause in the same subject of trace (a). The real and synthetic breathing signals of traces (b) and (c) exhibit a good similarity.

Figure 1.10: Acquisition systems for motion capturing:(a) multi-camera system for hospital monitoring; (b) single-camera system for domestic monitoring.

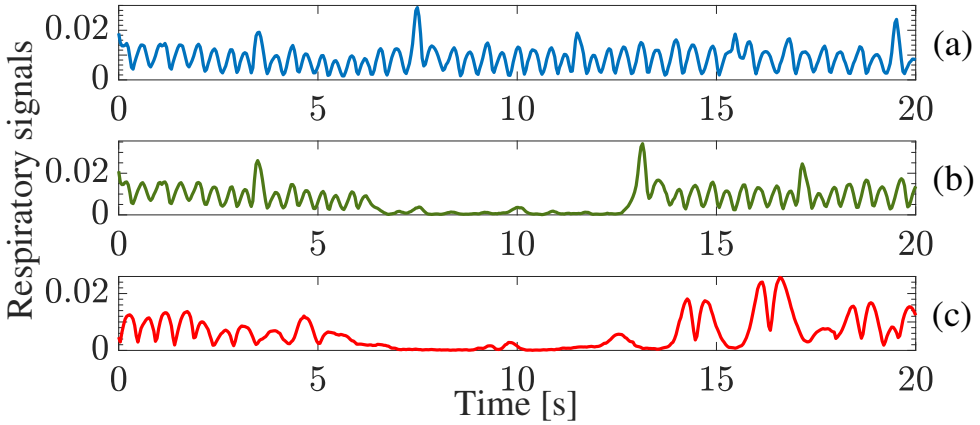


Figure 1.11: Motion signals: (a) regular breathing; (b) artificial respiratory pause; (c) real respiratory pause.
Image modified from an original image published in [29].

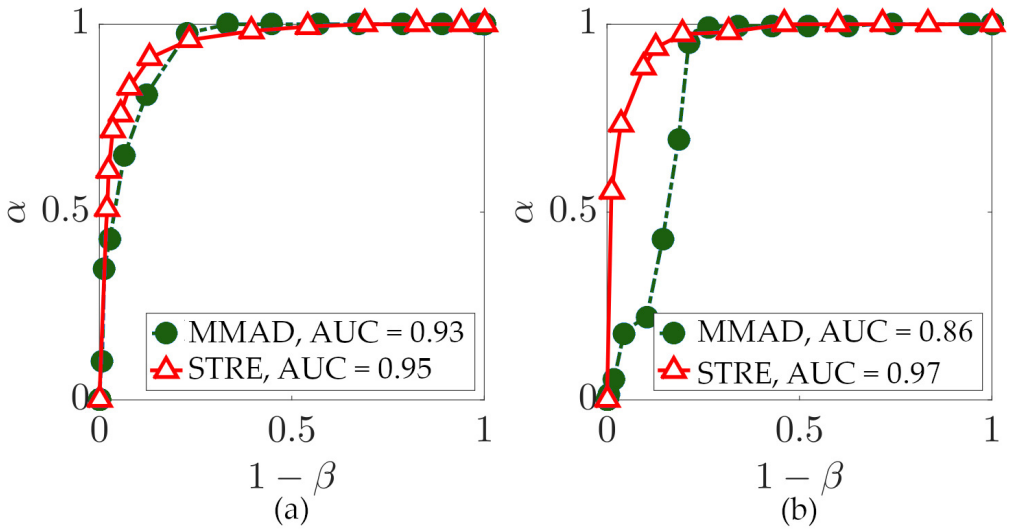


Figure 1.12: ROC curves for MMAD and STRE algorithms tested on: (a) software-generated videos; (b) hardware-generated videos.
Image modified from an original image published in [29].

1.5.1 Performance Analysis

The performance of RR estimation algorithms may be evaluated in terms of sensitivity and specificity defined, respectively, as:

$$\alpha = \frac{T_{\mathrm{TP}}}{T_{\mathrm{TP}}+T_{\mathrm{FN}}} \tag{1.20}$$

$$\beta = \frac{T_{\mathrm{TN}}}{T_{\mathrm{TN}}+T_{\mathrm{FP}}} \tag{1.21}$$

where T_{TP}, T_{TN}, T_{FP} T_{FN} represent the total durations of time intervals when an apnea episode is correctly detected (Time True Positives), regular breathing is correctly detected (Time True Negatives), apnea episode is incorrectly detected (Time False Positives) and regular breathing is incorrectly detected (Time False Negatives), respectively.

Results is terms of Receiver Operating Characteristic (ROC) curves [32], as functions of sensitivity and specificity, are plotted in Figure 1.12 for both proposed methods in [23] and [26] with different values of the decision threshold η. ROC curves in Figures 1.12(a) and 1.12(b) are obtained testing the estimation methods on software and hardware-generated videos, respectively. In particular, the algorithm presented in [23] is referred to as Motion Magnification for Apnea Detection (MMAD). Values of the Area Under Curve (AUC) parameter [32] are also reported in Figure 1.12. A test is considered highly reliable for values of AUC between 0.9 and 1 [32]. Recalling that an ideal detector is characterized by $\alpha = 1$ and $\beta = 1$, the optimum decision threshold is associated with the minimum Euclidean distance to the point (0,1) of the ROC curve plots.

1.6 Conclusions

In this chapter, novel video-based techniques for respiration monitoring are presented along with a brief review of relevant state-of-the art methods. As respiration induces periodic movements of the chest, video cameras can be used to capture motion signals to be properly processed in order to estimate the respiratory frequency. Motion magnification algorithms are expedient to amplify subtle respiratory movements in order to improve the estimation accuracy of the ML approach. Finally, simulators of breathing behaviours based on CTMC model are considered to extend poor video databases to allow an easier and more reliable assessment of respiration techniques based on video processing solutions. Application results and performance analysis are also discussed.

Bibliography

[1] Zahra Moussavi. *Fundamentals of Respiratory System and Sounds Analysis*. Morgan and Claypool, Canada, 1st edition edition, 2006.

[2] J.Y. Tu, Kiao Inthavong, and Goodarz Ahmadi. *Fluid Particle Dynamics in the Human Respiratory System – A Computational Approach*. Springer, Dordrecht, NL, 1st edition edition, 2013.

[3] S. Rolfe. The importance of respiratory rate monitoring. *British Journal of Nursing*, 28(8):504–508, 2019.

[4] F.Q. AL-Khalidi, R. Saatchi, D. Burke, H. Elphick, and S. Tan. Respiration rate monitoring methods: A review. *Pediatric Pulmonology*, 46(6):523–529, 2011.

[5] O. Steichen, G. Grateau, and E. Bouvard. Respiratory rate: the neglected vital sign. *The Medical journal of Australia*, 189(9):531–532, 2008.

[6] H. Liu, J. Allen, D. Zheng, and F. Chen. Recent development of respiratory rate measurement technologies. *Physiological measurement*, 40(7):1–27, 2019.

[7] Mosby. *Mosby's Medical Dictionary*. Elsevier, Saint Louis, MO, USA, 10th edition edition, 2016.

[8] N. P. Kondamudi, L. Krata, and A. S. Wilt. Infant apnea. *StatPearls [Internet]*, 2021.

[9] S. Mishra, R. Agarwal, M. Jeevasankar, R. Aggarwal, A.K. Deorari, and V.K. Paul. Apnea in the newborn. *Indian Journal of Pediatric*, 75(1):57–61, 2008.

[10] M. R. Mannarino, F. Di Filippo, and M. Pirro. Obstructive sleep apnea syndrome. *European Journal of Internal Medicine*, 23(7):585–593, 2012.

[11] Saher Zaidi, Jason Gandhi, Sohrab Vatsia, Noel L. Smith, and Sardar Ali Khan. Congenital central hypoventilation syndrome: An overview of etiopathogenesis, associated pathologies, clinical presentation, and management. *Autonomic Neuroscience*, 210:1–9, 2018.

[12] C. M. Cielo and C. L. Marcus. Central hypoventilation syndromes. *Sleep Medicine Clinics*, 9(1):105–118, 2014.

[13] Carlo Massaroni, Andrea Nicolò, Massimo Sacchetti, and Emiliano Schena. Contactless methods for measuring respiratory rate: A review. *IEEE Sensors Journal*, 21(11):12821–12839, 2021.

[14] Xin Liu, Qison Wang, Dan Liu, Yuan Wang, Yan Zhang, Ou Bai, and Jinwei Sun. Human emotion classification based on multiple physiological signals by wearable system. *Technology and Health Care*, 26(S1):459–469, 2018.

[15] Billy Sperlich, Kamiar Aminian, Peter Düking, and Hans-Christer Holmberg. Editorial: Wearable sensor technology for monitoring training load and health in the athletic population. *Frontiers in Physiology*, 10, 2020.

[16] Carlo Massaroni, Andrea Nicolò, Daniela Lo Presti, Massimo Sacchetti, Sergio Silvestri, and Emiliano Schena. Contact-based methods for measuring respiratory rate. *Sensors*, 19(908), 2019.

[17] G. K. Prisk, J. Hammer, and C. J. Newth. Techniques for measurement of thoracoabdominal asynchrony. *Pediatric pulmonology*, 34(6):462–472, 2002.

[18] Sleep apnea guide, 2016. Available Online: https://www.sleep-apnea-guide.com/polysomnogram.html.

[19] K. Mutlu, J. Esquivelzeta Rabell, P. Martin del Olmo, and S. Haesler. IR thermography-based monitoring of respiration phase without image segmentation. *Journal of Neuroscience Methods*, 301:1–8, 2018.

[20] Edgar A. Bernal, Lalit K. Mestha, and Eribaweimon Shilla. Non contact monitoring of respiratory function via depth sensing. In *IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI)*, pages 101–104, 2014.

[21] Fan Yang, Zhiming He, Yuanhua Fu, Liang Li, Kui Jiang, and Fangyan Xie. Non contact detection of respiration rate based on forward scatter radar. *Sensors*, 19(21), 2019.

[22] Hao-Yu Wu, Michael Rubinstein, Eugene Shih, John Guttag, Frédo Durand, and William Freeman. Eulerian video magnification for revealing subtle changes in the world. *ACM Trans. Graph.*, 31(4), Jul. 2012.

[23] L. Cattani, D. Alinovi, G. Ferrari, R. Raheli, E. Pavlidis, C. Spagnoli, and F. Pisani. A wire-free, non-invasive, low-cost video processing-based approach to neonatal apnoea detection. In *2014 IEEE Workshop on Biometric Measurements and Systems for Security and Medical Applications (BIOMS) Proceedings*, pages 67–73, 2014.

[24] L. Cattani, D. Alinovi, G. Ferrari, R. Raheli, E. Pavlidis, C. Spagnoli, and F. Pisani. Monitoring infants by automatic video processing: A unified approach to motion analysis. *Computers in biology and medicine*, 80:158–165, 2017.

[25] C. Solomon and T. Breckon. *Fundamentals of Digital Image Processing.* Wiley-Blackwell, Croydon, UK, 1st edition, 2011.

[26] Davide Alinovi, Luca Cattani, Gianluigi Ferrari, Francesco Pisani, and Riccardo Raheli. Spatio-temporal video processing for respiratory rate estimation. In *2015 IEEE International Symposium on Medical Measurements and Applications (MeMeA) Proceedings*, pages 12–17, 2015.

[27] Steven M. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall, Upper Saddle River, NJ, USA, 1st edition, 1993.

[28] Davide Alinovi, Gianluigi Ferrari, Francesco Pisani, and Riccardo Raheli. Respiratory rate monitoring by maximum likelihood video processing. In *2016 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, pages 172–177, 2016.

[29] Davide Alinovi, Gianluigi Ferrari, Francesco Pisani, and Riccardo Raheli. Markov chain modeling and simulation of breathing patterns. *Biomedical Signal Processing and Control*, 33:245–254, 2017.

[30] Jeremy Blum. *Exploring Arduino: tools and techniques for engineering wizardry*. John Wiley & Sons, 2019.

[31] Davide Alinovi, Luca Cattani, Gianluigi Ferrari, Francesco Pisani, and Riccardo Raheli. Video simulation of apnoea episodes. In *2015 IEEE International Conference on Multimedia Expo Workshops (ICMEW)*, pages 1–6, 2015.

[32] J. Swets. Measuring the accuracy of diagnostic systems. *Science*, 240(4857):1285–93, 1988.

Bibliography

[1] Zahra Moussavi. *Fundamentals of Respiratory System and Sounds Analysis*. Morgan and Claypool, Canada, 1st edition edition, 2006.

[2] J.Y. Tu, Kiao Inthavong, and Goodarz Ahmadi. *Fluid Particle Dynamics in the Human Respiratory System – A Computational Approach*. Springer, Dordrecht, NL, 1st edition edition, 2013.

[3] S. Rolfe. The importance of respiratory rate monitoring. *British Journal of Nursing*, 28(8):504–508, 2019.

[4] F.Q. AL-Khalidi, R. Saatchi, D. Burke, H. Elphick, and S. Tan. Respiration rate monitoring methods: A review. *Pediatric Pulmonology*, 46(6):523–529, 2011.

[5] O. Steichen, G. Grateau, and E. Bouvard. Respiratory rate: the neglected vital sign. *The Medical journal of Australia*, 189(9):531–532, 2008.

[6] H. Liu, J. Allen, D. Zheng, and F. Chen. Recent development of respiratory rate measurement technologies. *Physiological measurement*, 40(7):1–27, 2019.

[7] Mosby. *Mosby's Medical Dictionary*. Elsevier, Saint Louis, MO, USA, 10th edition edition, 2016.

[8] N. P. Kondamudi, L. Krata, and A. S. Wilt. Infant apnea. *StatPearls [Internet]*, 2021.

[9] S. Mishra, R. Agarwal, M. Jeevasankar, R. Aggarwal, A.K. Deorari, and V.K. Paul. Apnea in the newborn. *Indian Journal of Pediatric*, 75(1):57–61, 2008.

[10] M. R. Mannarino, F. Di Filippo, and M. Pirro. Obstructive sleep apnea syndrome. *European Journal of Internal Medicine*, 23(7):585–593, 2012.

[11] Saher Zaidi, Jason Gandhi, Sohrab Vatsia, Noel L. Smith, and Sardar Ali Khan. Congenital central hypoventilation syndrome: An overview of etiopathogenesis, associated pathologies, clinical presentation, and management. *Autonomic Neuroscience*, 210:1–9, 2018.

[12] C. M. Cielo and C. L. Marcus. Central hypoventilation syndromes. *Sleep Medicine Clinics*, 9(1):105–118, 2014.

[13] Carlo Massaroni, Andrea Nicolò, Massimo Sacchetti, and Emiliano Schena. Contactless methods for measuring respiratory rate: A review. *IEEE Sensors Journal*, 21(11):12821–12839, 2021.

[14] Xin Liu, Qison Wang, Dan Liu, Yuan Wang, Yan Zhang, Ou Bai, and Jinwei Sun. Human emotion classification based on multiple physiological signals by wearable system. *Technology and Health Care*, 26(S1):459–469, 2018.

[15] Billy Sperlich, Kamiar Aminian, Peter Düking, and Hans-Christer Holmberg. Editorial: Wearable sensor technology for monitoring training load and health in the athletic population. *Frontiers in Physiology*, 10, 2020.

[16] Carlo Massaroni, Andrea Nicolò, Daniela Lo Presti, Massimo Sacchetti, Sergio Silvestri, and Emiliano Schena. Contact-based methods for measuring respiratory rate. *Sensors*, 19(908), 2019.

[17] G. K. Prisk, J. Hammer, and C. J. Newth. Techniques for measurement of thoracoabdominal asynchrony. *Pediatric pulmonology*, 34(6):462–472, 2002.

[18] Sleep apnea guide, 2016. Available Online: https://www.sleep-apnea-guide.com/polysomnogram.html.

[19] K. Mutlu, J. Esquivelzeta Rabell, P. Martin del Olmo, and S. Haesler. IR thermography-based monitoring of respiration phase without image segmentation. *Journal of Neuroscience Methods*, 301:1–8, 2018.

[20] Edgar A. Bernal, Lalit K. Mestha, and Eribaweimon Shilla. Non contact monitoring of respiratory function via depth sensing. In *IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI)*, pages 101–104, 2014.

[21] Fan Yang, Zhiming He, Yuanhua Fu, Liang Li, Kui Jiang, and Fangyan Xie. Non contact detection of respiration rate based on forward scatter radar. *Sensors*, 19(21), 2019.

[22] Hao-Yu Wu, Michael Rubinstein, Eugene Shih, John Guttag, Frédo Durand, and William Freeman. Eulerian video magnification for revealing subtle changes in the world. *ACM Trans. Graph.*, 31(4), Jul. 2012.

[23] L. Cattani, D. Alinovi, G. Ferrari, R. Raheli, E. Pavlidis, C. Spagnoli, and F. Pisani. A wire-free, non-invasive, low-cost video processing-based approach to neonatal apnoea detection. In *2014 IEEE Workshop on Biometric Measurements and Systems for Security and Medical Applications (BIOMS) Proceedings*, pages 67–73, 2014.

[24] L. Cattani, D. Alinovi, G. Ferrari, R. Raheli, E. Pavlidis, C. Spagnoli, and F. Pisani. Monitoring infants by automatic video processing: A unified approach to motion analysis. *Computers in biology and medicine*, 80:158–165, 2017.

[25] C. Solomon and T. Breckon. *Fundamentals of Digital Image Processing.* Wiley-Blackwell, Croydon, UK, 1st edition, 2011.

[26] Davide Alinovi, Luca Cattani, Gianluigi Ferrari, Francesco Pisani, and Riccardo Raheli. Spatio-temporal video processing for respiratory rate estimation. In *2015 IEEE International Symposium on Medical Measurements and Applications (MeMeA) Proceedings*, pages 12–17, 2015.

[27] Steven M. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory.* Prentice Hall, Upper Saddle River, NJ, USA, 1st edition, 1993.

[28] Davide Alinovi, Gianluigi Ferrari, Francesco Pisani, and Riccardo Raheli. Respiratory rate monitoring by maximum likelihood video processing. In *2016 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, pages 172–177, 2016.

[29] Davide Alinovi, Gianluigi Ferrari, Francesco Pisani, and Riccardo Raheli. Markov chain modeling and simulation of breathing patterns. *Biomedical Signal Processing and Control*, 33:245–254, 2017.

[30] Jeremy Blum. *Exploring Arduino: tools and techniques for engineering wizardry.* John Wiley & Sons, 2019.

[31] Davide Alinovi, Luca Cattani, Gianluigi Ferrari, Francesco Pisani, and Riccardo Raheli. Video simulation of apnoea episodes. In *2015 IEEE International Conference on Multimedia Expo Workshops (ICMEW)*, pages 1–6, 2015.

[32] J. Swets. Measuring the accuracy of diagnostic systems. *Science*, 240(4857):1285–93, 1988.