

---

## Chapter 1

# Video-Based Solutions for Newborn Monitoring

*Veronica Mattioli<sup>1</sup>, Davide Alinovi<sup>1,3</sup>, Francesco Pisani<sup>2</sup>,  
Gianluigi Ferrari<sup>1</sup>, Riccardo Raheli<sup>1</sup>*

---

Efficient monitoring of vital signs is a fundamental tool in disease prevention and medical diagnostics. Main physiological parameters to monitor are heart rate, blood pressure, respiratory rate and body temperature, but also motion analysis may provide essential information about the clinical status of a patient. Very specific pathological movements can indeed be signs of important or potentially threatening disorders. Besides being almost exclusively performed in hospital settings, conventional monitoring often requires a contact with the body of the patient that makes traditional systems possibly invasive and uncomfortable, especially if applied on newborns. To make home care more accessible and comfortable, novel methods for remote and contactless monitoring have been developed in the recent years. Among others, appealing solutions that have received recent research attention are based on video processing techniques that allow to capture and analyse the movements of a patient in a contactless fashion.

## 1.1 Introduction

Early diagnosis of neonatal disorders may be crucial for timely intervention and treatment. As some rare but potentially harmful diseases in newborns that can manifest themselves with clinical symptoms, such as seizures and apneas, affect the movements of the patient, movement monitoring and analysis may be an effective diagnostic tool. In particular, some types of seizures can be characterized by jerky periodic movements of one or more body parts, usually limbs and head. On the other hand, apneas are associated to the absence of periodic breathing movements [1].

A seizure can be defined as an age-dependent clinical event characterized by a neurological dysfunction caused by paroxysmal alterations of neurological, behavioural, and/or automatic functions [2]. One of the most common outward effects

<sup>1</sup>Department of Engineering and Architecture, University of Parma, Parco Area delle Scienze 181/A, IT-43124 Parma, Italy, {veronica.mattioli,gianluigi.ferrari,riccardo.raheli}@unipr.it

<sup>2</sup>Department of Medicine and Surgery, University of Parma, Via Gramsci 14, IT-43126 Parma, Italy, francesco.pisani@unipr.it

<sup>3</sup>Deceased.

is an uncontrolled shaking due to involuntary and rapidly contraction and relaxation of one or more muscle groups. Preterm and at term newborns are more likely to suffer from a seizure within 28 days after birth or 44 weeks of conceptional age, respectively [2]. The estimated incidence is 2.6‰ for overall newborns, 2.0‰ for term neonates, 11.1‰ for preterm neonates, and 13.5‰ for infants weighing less than 2500 g [3]. Hypoxic-ischemic encephalopathy and stroke are two of the more frequent etiologies, but also brain malformation and infection can be triggers [4]. As reported in [4] and [5], several classifications have been proposed, but usually four main types of clinical manifestations are considered indicative of neonatal seizures: subtle, clonic, tonic, myoclonic [2]. Each clinical type of seizure is characterized by distinguishable features and requires focused analysis and diagnostic approach. Clonic seizures, for instance, are associated to rhythmic and slow movements.

On the other hand, apneas can be defined as sudden interruptions of the respiratory airflow. In newborns, these episodes are considered to be significant if lasting longer than 20 seconds or less if associated with other symptoms, i.e. bradycardia and cyanosis [6]. As reported in [1], among the main causes of neonatal apneas we recall seizures, cerebrovascular events [7] and congenital disorders, such as Congenital Central Hypoventilation Syndrome (CCHS) [8], [9]. CCHS, in particular, is a rare life-threatening disease caused by a defect in the *PHOX2B* homeobox gene [9]. It mainly occurs during sleep and is responsible of alveolar hypoventilation. It is usually associated with cyanosis, apnea or cardiorespiratory arrest [10]. Finally, three main categories of apneas can be identified, i.e. central, obstructive and mixed, according to the presence or the lack of an obstruction of the upper airway [6].

Due to the severity of these neonatal disorders, early treatments are needed to prevent life-threatening episodes as well as life-long consequences. To this purpose, efficient monitoring tools must be deployed. The investigation of modern monitoring systems based on video processing solutions can be considered a promising and effective alternative to conventional equipment. Motion analysis plays a key role to detect anomalous movements related to the aforementioned disorders and will be discussed in the next sections.

## **1.2 Vital signs monitoring**

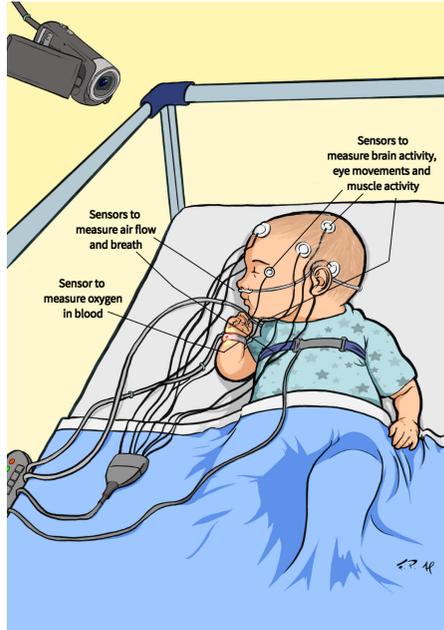
Conventional systems for monitoring vital signs are often intrusive and not suitable for home care. The main standard tool for the diagnosis of sleep-related disorders, such as seizures and apneas, is the polysomnogram, which allows to record the sleep of a patient and includes several monitoring systems, i.e. ElectroEncephaloGram (EEG), ElectroOculoGram (EOG), ElectroMyoGram (EMG) and ElectroCardioGram (ECG) [11]. Each of these measurement techniques requires wired sensors to be directly attached to different body parts of the patient. In particular, electrodes placed on the scalp, near the eyes, under the chin and on the chest are employed to record brain activity, eyes movements, muscle activity and heart rate, respectively. Additional information is acquired through elastic belt sensors placed around the chest, nasal flow meter and pulse oximeter to measure the amount of effort to breath, the airflow and the oxygen saturation of the blood, respectively [12]. Cameras can

also be employed for simultaneous traditional monitoring. A schematic overview of the polysomnogram test is shown in Figure 1.1(a) and an example of recorded data is reported in Figure 1.1(b) where the first four traces are the EEG channels, and the subsequent traces are, from top to bottom, snoring noise, nasal flow, thoracic movements, and oxygen saturation. The abnormal breathing pattern is characterized by recurrence of central apneas (closed boxes), in the absence of airway obstruction and snoring. Central apneas determine severe oxygen desaturations. Besides being expensive and moderately invasive, especially for newborns, these techniques are almost exclusively deployed in clinical settings and require trained medical staff, who may not be available full-time. To make home care more accessible, various monitoring systems have been developed, e.g. smart bed [13] and wearable sensors-based systems [14], but they still require a contact with the body of the patient. Contactless solutions, on the other hand, may be devised for the automatic detection of anomalous activities potentially related to neonatal disorders. To this purpose, digital cameras can be used to frame the movements of a patient to be analysed through proper video-processing algorithms. The integration of these novel approaches allow to enhance both hospital and home constant monitoring by providing low-cost preliminary alert signals to be possibly further investigated by conventional diagnostic tools, i.e. the EEG.

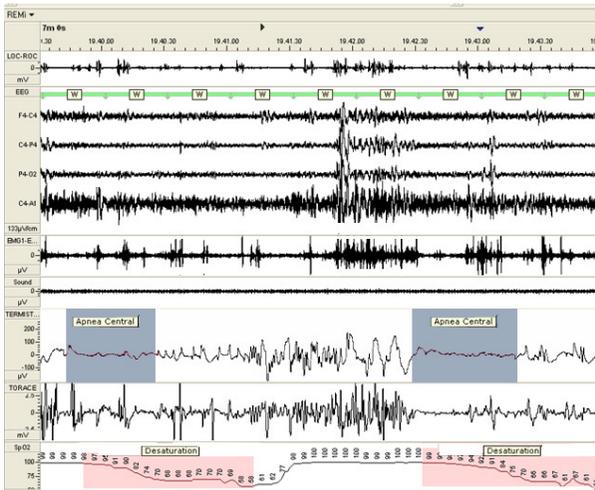
### 1.3 Video processing systems for neonatal disorder detection

Early work on video-based solutions for newborn monitoring was mainly focused on seizure detection and was based on motion extraction algorithms such as optical flow [15] and block matching [16]. Neural networks for anomalous event detection and motion classification were also investigated. The implementation of all these methods may be complex, expensive and not suitable for real-time monitoring. However, in [1] a fast and reliable approach for the real-time analysis of newborns' movements is proposed on the basis of preliminary contributions [17], [18], [19]. This method is based on the extraction of motion signals acquired with single or multiple cameras and ultimately relies on the well-known Maximum-Likelihood (ML) estimation criterion [20]. Besides being presented for clonic seizures and apneas detection, it is valid for any disorder characterized by the presence or the absence of periodic movements.

In Subsections 1.3.1 and 1.3.2 the motion estimation algorithm presented in [1] will be briefly introduced for single and multiple-sensor analysis, respectively. The performance in the latter case is improved by the different viewing angles that allow to capture movements that may be occluded for a single camera, hence undetectable. In particular, models to describe motion signals acquired from properly pre-processed video sequences are discussed. The specific procedures for the extraction of motion information related to seizures and apneas will be presented in Sections 1.4 and 1.5, respectively. Since the disorders under investigation are characterized by the presence or absence of periodic movements, motion signals will be modeled as periodic signals, where the fundamental frequency represents the main unknown parameter to be estimated.



(a)



(b)

Figure 1.1: Polysomnography during wakefulness: (a) schematic overview and (b) example of recorded data.<sup>4</sup>

In the following analysis we will consider video sequences with sampling period  $T$  where frames have dimension  $W \times H$  and are sampled at time instants  $iT$ ,  $i$  being the frame number.

<sup>4</sup>Image by Francesca and Andrea Pisani.

### 1.3.1 Single sensor

Let  $\bar{L}[i]$  be the average luminance signal extracted at frame  $i$  of a considered video sequence. We define it as:

$$\bar{L}[i] = \frac{1}{WH} \sum_{x=1}^W \sum_{y=1}^H I[x, y, i] \quad (1.1)$$

where  $I[x, y, i]$  represents the  $[x, y]$  entry of matrix  $I[i]$  that describes the  $i$ -th frame after a proper processing procedure. The motion signal in (1.1) can be modeled as:

$$\bar{L}[i] = c + A \cos(2\pi f_0 iT + \phi) + n[i] \quad (1.2)$$

where  $c$  is a continuous component and  $n[i]$  are samples of independent identically distributed (i.i.d.) zero-mean Gaussian noise. The unknown parameters  $A$ ,  $f_0$  and  $\phi$  represent the amplitude, phase, and frequency, respectively, of the periodic signal and may be collected in a vector  $\theta = [A, f_0, \phi]$ . The ML approach can be now exploited to estimate the vector  $\theta$ . In particular, observing a window of  $N$  frames and following standard methods described in [20], an estimator of the fundamental frequency can be obtained and expressed as:

$$\hat{f}_0 = \underset{f}{\operatorname{argmax}} \left| \sum_{i=0}^{N-1} \bar{L}[i] e^{-j2\pi f iT} \right|^2. \quad (1.3)$$

Similarly, an expression of the amplitude estimator can be written as:

$$\hat{A} = \frac{2}{N} \left| \sum_{n=0}^{N-1} \bar{L}[n] e^{-j2\pi \hat{f}_0 n T} \right|. \quad (1.4)$$

The presence of a significant periodic component can be finally declared if the following constraint is verified:

$$N\hat{A}^2 > \eta \quad (1.5)$$

where the value of the threshold  $\eta$  may be determined by trial and error.

### 1.3.2 Multiple sensors

The method described in Subsection 1.3.1 can be extended to multiple sensors to achieve better performance. Multi-camera systems can indeed detect movements that may be occluded for a single camera. Considering  $S$  sensors, a set of motion signals is defined as in (1.1):

$$\bar{L}_s[i] = \frac{1}{WH} \sum_{x=1}^W \sum_{y=1}^H I_s[x, y, i] \quad , \quad s = 1, 2, \dots, S. \quad (1.6)$$

where the processed  $i$ -th frame for the  $s$ -th sensor is described by matrix  $I_s[x, y, i]$ . The model in (1.2) can be generalized, as:

$$\bar{L}_s[i] = c_s + A_s \cos(2\pi f_0 iT + \phi_s) + n_s[i] \quad (1.7)$$

where the sampling period  $T$  and the fundamental frequency  $f_0$  are assumed to be identical for each capturing device; where present, the subscript  $s$  refers to the  $s$ -th sensor. Following the same procedure of the single-sensor analysis and exploiting now data fusion techniques to combine data acquired by different sensors, an estimator of the fundamental frequency can now be formulated as:

$$\hat{f}_0 = \underset{f}{\operatorname{argmax}} \sum_{s=1}^S \left| \sum_{i=0}^{N-1} \bar{L}_s[i] e^{-j2\pi f i T} \right|^2. \quad (1.8)$$

Likewise, assuming that different values of amplitude are associated to each sensor, a set of amplitude estimators is:

$$\hat{A}_s = \frac{2}{N} \left| \sum_{n=0}^{N-1} \bar{L}_s[n] e^{-j2\pi \hat{f}_0 n T} \right|, \quad s = 1, 2, \dots, S. \quad (1.9)$$

Finally, the constraint to be satisfied in order to detect a significant periodic component is now:

$$\frac{N}{S} \sum_{s=1}^S \hat{A}_s^2 > \eta. \quad (1.10)$$

To improve the performance of both single and multi-sensor analysis, interlaced windows can be considered as the selected detection algorithm may indeed fail when pathological movements manifest across two consecutive disjoint windows [17].

## 1.4 Seizure detection

In order to extract the average motion signals defined in (1.1) and (1.6), each frame of a considered video sequence need to be properly processed. A schematic overview of the pre-processing algorithm exploited in [1] and [17] for seizure detection is presented in Figure 1.2, where four phases are highlighted: gray-scale conversion, difference filtering, binarization and erosion. A generic Red, Green and Blue (RGB)

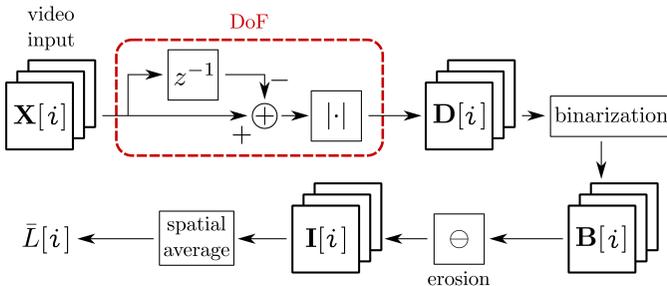


Figure 1.2: Seizure detection pre-processing algorithm.

video sequence  $\mathbf{X}[i]$  is considered as the input of the processing system and initially converted to gray-scale. The Difference of Frames (DoF) is then performed on consecutive frames as a basic image filtering operation and the result is thresholded

to obtain a binary mask, where white pixels correspond to foreground regions. Finally, the erosion morphological operation [21] is implemented to reduce noise as discussed in [17]. Examples of frames at each processing step are shown in Figure 1.3. Eventually, an example of periodic movements induced by a clonic seizure

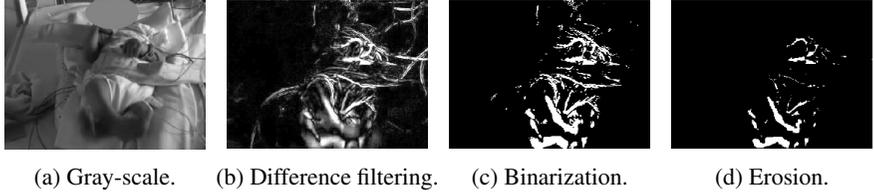


Figure 1.3: Results of each processing step.

and extracted according the procedure illustrated in Figure 1.2 is shown in Figure 1.4, where the extracted average motion signal is plotted against the frame number along with a corresponding EEG signal [17]. The two signals exhibit a comparable periodicity.

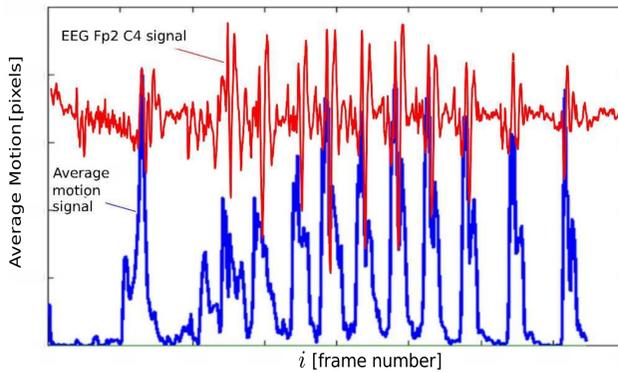


Figure 1.4: Average motion signal [17].

#### 1.4.1 Performance in seizure detection

The performance of the described detection method has been evaluated in terms of sensitivity and specificity over  $n$  tests, respectively defined as:

$$\alpha = \frac{n_{TP}}{n_{TP} + n_{FN}} \quad (1.11)$$

$$\beta = \frac{n_{TN}}{n_{TN} + n_{FP}} \quad (1.12)$$

where  $n_{TP}$ ,  $n_{TN}$ ,  $n_{FP}$ ,  $n_{FN}$  are the numbers of True Positives, True Negatives, False Positives and False Negatives in the considered sequence. In particular, positives and negatives are classified when a seizure is detected or undetected, respectively.

Performance of single and multi-sensor systems have been investigated and the results obtained using a single RGB sensor are illustrated in Figure 1.5(a), where the Receiver Operating Characteristic (ROC) [22] curve is plotted as a function of  $\alpha$  and  $1 - \beta$  for various values of the threshold  $\eta$ . We recall that an optimal predictor is characterized by  $\alpha = 1$  and  $\beta = 1$ , i.e., all seizures are correctly detected when present, and an Area Under Curve (AUC) equal to 1. In the presented example the minimum Euclidean distance  $D$  from the ideal configuration is 0.14 and the AUC is 0.95, which indicates high reliability [22]. In Figure 1.5(b) sensitivity and specificity values are plotted for optimal values of  $\eta$  and different RGB cameras configurations. The best performance is achieved when all three sensors are employed, i.e.  $S = 3$  in (1.6)-(1.10). Depth sensors could also be employed to better distinguish pathological movements from background noise or random movements [1].

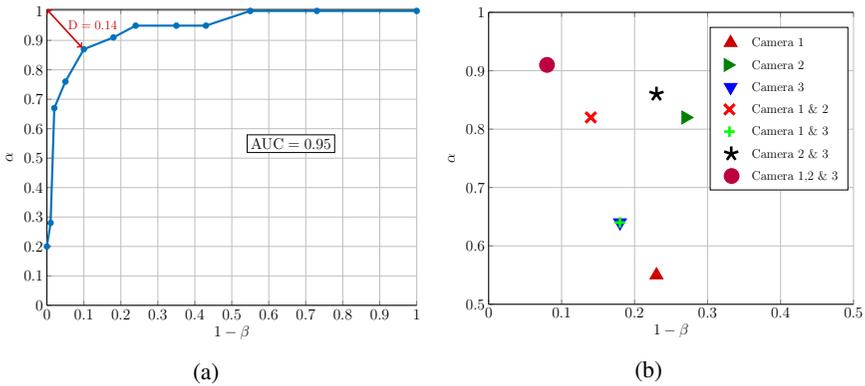


Figure 1.5: Performance analysis for: (a) single RGB camera (ROC curve); (b) different RGB camera configurations [1].

## 1.5 Apnea detection

Breathing-related movements are often subtle and difficult to detect, especially for newborns. To make the detection algorithm efficient also in presence of small movements, a motion magnification algorithm can be applied to amplify the considered motion signals. In particular, the Eulerian Video Magnification (EVM) method presented in [23] can be exploited and its schematic representation is illustrated in Figure 1.6, where the input signal  $\mathbf{X}[i]$  is processed through four main phases. Initially, each frame of the considered video sequence is decomposed into different spatial frequency bands by spatial decomposition. The obtained outputs are then filtered through a pixel-wise temporal operation and the frequency bands of interest are extracted. Multiplications by proper gains are now performed to amplify the filtered signals and a video frame reconstruction is finally implemented to obtain a new output signal where small changes at the input are enhanced. The motion extraction algorithm shown in Figure 1.2 can be applied after the EVM processing.

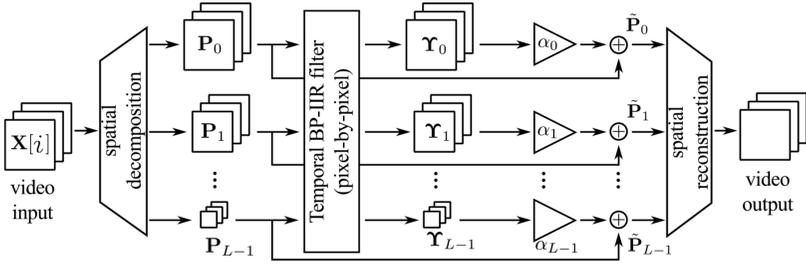
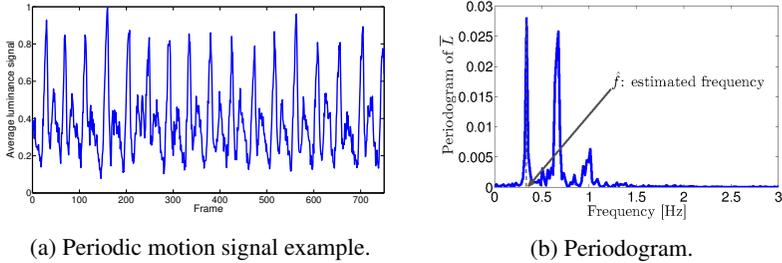


Figure 1.6: EVM algorithm [19].

An example of an extracted motion signal is plotted in Figure 1.7(a), where the normal periodic behaviour of a respiratory signal can be easily observed. The corresponding periodogram is shown in Figure 1.7(b), where the estimated fundamental frequency is highlighted at the peak of the function. On the other hand, an example



(a) Periodic motion signal example.

(b) Periodogram.

Figure 1.7

of an anomalous motion signal is illustrated in Figure 1.8, where the sudden interruption of the respiration caused by an apnea episode is visible in the flat central part of the plot.

Finally, for the sake of comparison, an instance of an extracted motion signal is shown in Figure 1.9 along with the equivalent signal obtained from a pneumograph, where every period of the pneumographic signal corresponds to a complete respiratory act of the patient. Considering that a respiratory act is composed by two phases, i.e. inhalation and exhalation, a good match of the two signals can be observed.

A number of improvements are possible with reference to the method described in Figure 1.6. Among them, we mention [24] where the computationally intensive reconstruction of the video stream is avoided and the sublevel signals are directly combined in a multidimensional effective estimator. Direct application of multidimensional ML estimator to the video stream was also investigated and showed very

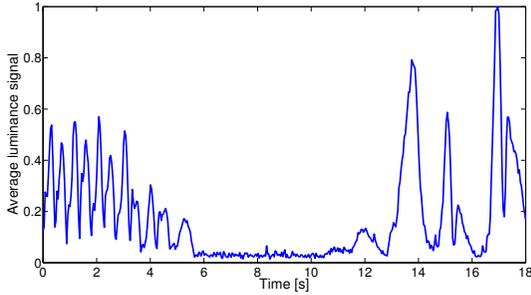


Figure 1.8: Anomalous motion signal related to an apnea event in a newborn.

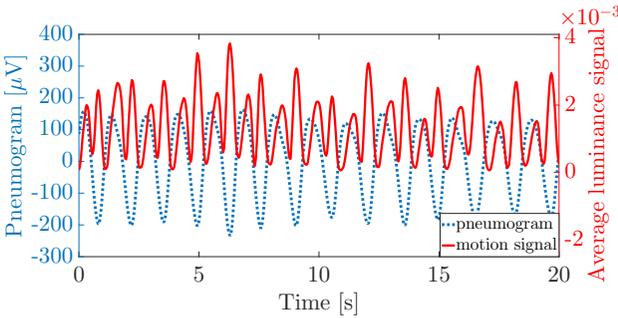


Figure 1.9: Extracted motion signal and correspondent pneumographic signal.

good performance provided proper processing is performed to select suitable Region of Interest (ROI) [25]. We also remark that these methods could be applied to seizure detection as well. However, the subtle respiration movements specifically require these more sensitive (and complex) solutions.

### 1.5.1 Performance in apnea detection

To evaluate the performance of the apnea detection algorithm, sensitivity and specificity coefficients are still considered, but slightly differently defined as:

$$\alpha = \frac{T_{TP}}{T_{TP} + T_{FN}} \tag{1.13}$$

$$\beta = \frac{T_{TN}}{T_{TN} + T_{FP}} \tag{1.14}$$

where  $T_{TP}$ ,  $T_{TN}$ ,  $T_{FP}$   $T_{FN}$  represent now the total duration of time intervals when an apnea episode is correctly detected (Time True Positives), correctly undetected (Time True Negatives), incorrectly detected (Time False Positives) and incorrectly undetected (Time False Negatives), respectively. An interval of tolerance, i.e. Tolerance Delay (TD), is allowed to declare an apnea episode is detected. As an illustrative example, the motion signal related to a long apnea episode is shown in Figure 1.10, where half interlaced observation windows are highlighted in the upper part of

the plot and different TD values are considered. The apnea detection fails for TD = 0 s, while succeeds for TD = 10, 20 and 30 s.

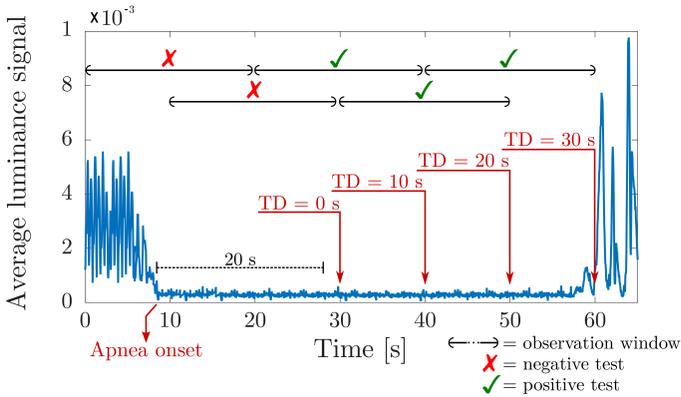


Figure 1.10: Detection of a long apnea episode.

## 1.6 Conclusion

In this chapter, novel techniques for newborn monitoring based on video processing solutions have been proposed. Considering that disorders such as seizures and apneas are characterized by specific pathological movements, proper algorithms for motion analysis have been presented. In particular, as clonic seizures trigger jerky movements of some human body parts and apneas cause sudden interruptions of the rhythmical respiration movements, periodicity is the fundamental parameter to be investigated. Motion signals extracted from video sequences can indeed be modeled as periodic signals, where the fundamental frequency can be ultimately estimated by standard probability theory techniques. The main goal is the integration of reliable, non-invasive and contactless systems with conventional equipment to provide clinical support to enhance early diagnosis of potentially life-threatening diseases.

## References

- [1] Cattani L, Alinovi D, Ferrari G, et al. Monitoring infants by automatic video processing: A unified approach to motion analysis. *Computers in biology and medicine*. 2017;80:158–165.
- [2] Volpe JJ. *Neurology of the Newborn*. 5th ed. Philadelphia, PA, USA: Saunders Elsevier; 2008.
- [3] Ronen GM, Penney S, Andrews W. The epidemiology of clinical neonatal seizures in Newfoundland: A population-based study. *The Journal of Pediatrics*. 1999;134(1):71–75.

- [4] Pellegrin S, Munoz FM, Padula M, et al. Neonatal seizures: Case definition and guidelines for data collection, analysis, and presentation of immunization safety data. *Vaccine*. 2019;37(52):7596–7609.
- [5] Rivielo JJ. Classification of seizures and epilepsy. *Current Neurology and Neuroscience Reports*. 2003;3:325–331.
- [6] Mishra S, Agarwal R, Jeevasankar M, et al. Apnea in the newborn. *Indian J Pediatr*. 2008;75:57–71.
- [7] Tramonte J, Goodkin H. Temporal Lobe Hemorrhage in the Full-Term Neonate Presenting as Apneic Seizures. *J Perinatol*. 2004;24:726–729.
- [8] Rand CM, Patwari PP, Carroll MS, et al. Congenital Central Hypoventilation Syndrome and Sudden Infant Death Syndrome: Disorders of Autonomic Regulation. *Seminars in Pediatric Neurology*. 2013;20(1):44–55. *Pediatric Autonomic Disorders*.
- [9] Healy F, Marcus CL. Congenital Central Hypoventilation Syndrome in Children. *Paediatric Respiratory Reviews*. 2011;12(4):253–263.
- [10] Cielo CM, Marcus CL. Central Hypoventilation Syndromes. *Sleep Medicine Clinics*. 2014;9(1):105–118.
- [11] Martin RJ, Block AJ, Cohn MA, et al. Indications and standards for cardiopulmonary sleep studies. *Sleep*. 1985;8(4):371–379.
- [12] Sleep Apnea Guide;. Available from: <https://www.sleep-apnea-guide.com/polysomnogram.html>.
- [13] Spillman Jr WB, Mayer M, Bennett J, et al. A smart bed for non-intrusive monitoring of patient physiological factors. *Measurement Science and Technology*. 2004 jul;15(8):1614–1620.
- [14] Pantelopoulos A, Bourbakis NG. A Survey on Wearable Sensor-Based Systems for Health Monitoring and Prognosis. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*. 2010;40(1):1–12.
- [15] Karayiannis NB, Varughese B, Tao G, et al. Quantifying motion in video recordings of neonatal seizures by regularized optical flow methods. *IEEE Transactions on Image Processing*. 2005;14(7):890–903.
- [16] Karayiannis NB, Sami A, Frost JD, et al. Automated extraction of temporal motor activity signals from video recordings of neonatal seizures based on adaptive block matching. *IEEE Transactions on Biomedical Engineering*. 2005;52(4):676–686.
- [17] Ntonfo GMK, Ferrari G, Raheli R, et al. Low-Complexity Image Processing for Real-Time Detection of Neonatal Clonic Seizures. *IEEE Transactions on Information Technology in Biomedicine*. 2012;16(3):375–382.
- [18] Cattani L, Kouamou Ntonfo GM, Lofino F, et al. Maximum-likelihood detection of neonatal clonic seizures by video image processing. In: 2014 8th International Symposium on Medical Information and Communication Technology (ISMICT); 2014. p. 1–5.
- [19] Cattani L, Alinovi D, Ferrari G, et al. 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; 2014. p. 67–73.

- [20] Kay SM. *Fundamentals of Statistical Signal Processing: Estimation Theory*. 1st ed. Upper Saddle River, NJ, USA: Prentice Hall; 1993.
- [21] Solomon C, Breckon T. *Fundamentals of Digital Image Processing*. 1st ed. Croydon, UK: Wiley-Blackwell; 2011.
- [22] Swets J. Measuring the accuracy of diagnostic systems. *Science*. 1988;240(4857):1285–93.
- [23] Wu HY, Rubinstein M, Shih E, et al. Eulerian Video Magnification for Revealing Subtle Changes in the World. *ACM Trans Graph*. 2012 Jul;31(4).
- [24] Alinovi D, Cattani L, Ferrari G, et al. Spatio-temporal video processing for respiratory rate estimation. In: *2015 IEEE International Symposium on Medical Measurements and Applications (MeMeA) Proceedings*; 2015. p. 12–17.
- [25] Alinovi D, Ferrari G, Pisani F, et al. Respiratory rate monitoring by maximum likelihood video processing. In: *2016 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*; 2016. p. 172–177.