UNIVERSITÀ DEGLI STUDI DI PARMA

Dottorato di Ricerca in Tecnologie dell'Informazione

XXVIII Ciclo

MONITORING INFANTS BY AUTOMATIC VIDEO PROCESSING

Relatore:

Chiar.mo Prof. Riccardo Raheli

Tutor:

Chiar.mo Prof. Gianluigi Ferrari

Dottorando: *Luca Cattani*

Gennaio 2016

...to my love Tiziana
and my beautiful daughters Viola and Alice

Contents

Introduction

This work has, as its objective, the development of non-invasive and low-cost systems for monitoring and automatic diagnosing specific neonatal diseases by means of the analysis of suitable video signals. We focus on monitoring infants potentially at risk of diseases characterized by the presence or absence of rhythmic movements of one or more body parts. Seizures and respiratory diseases are specifically considered, but the approach is general.

Seizures are defined as sudden neurological and behavioural alterations [1]. They are age-dependent phenomena and the most common sign of central nervous system dysfunction. Neonatal seizures have onset within the 28th day of life in newborns at term and within the 44th week of conceptional age in preterm infants. Their main causes are hypoxic-ischaemic encephalopathy, intracranial haemorrhage, and sepsis. Studies indicate an incidence rate of neonatal seizures of 2‰ live births, 11‰ for preterm neonates, and 13‰ for infants weighing less than 2500 g at birth [2]. Neonatal seizures can be classified into four main categories: clonic, tonic, myoclonic, and subtle [1]. Seizures in newborns have to be promptly and accurately recognized in order to establish timely treatments that could avoid an increase of the underlying brain damage.

Respiratory diseases related to the occurrence of apnoea episodes may be caused by cerebrovascular events [3]. Among the wide range of causes of apnoea [4], besides seizures, a relevant one is Congenital Central Hypoventilation Syndrome (CCHS) [5]. With a reported prevalence of 1 in 200,000 live births [6], CCHS, formerly known as Ondine's curse, is a rare life-threatening disorder characterized by a failure

of the automatic control of breathing [7], caused by mutations in a gene classified as PHOX2B [5]. CCHS manifests itself, in the neonatal period, with episodes of cyanosis or apnoea, especially during quiet sleep. The reported mortality rates range from 8% to 38% of newborn with genetically confirmed CCHS [6]. Nowadays, CCHS is considered a disorder of autonomic regulation, with related risk of sudden infant death syndrome (SIDS) [8].

Currently, the standard method of diagnosis, for both diseases, is based on polysomnography, a set of sensors such as ElectroEncephaloGram (EEG) sensors, ElectroMyoGraphy (EMG) sensors, ElectroCardioGraphy (ECG) sensors, elastic belt sensors, pulse-oximeter and nasal flow-meters. This monitoring system is very expensive, time-consuming, moderately invasive and requires particularly skilled medical personnel, not always available in a Neonatal Intensive Care Unit (NICU). Therefore, automatic, real-time and non-invasive monitoring equipments able to reliably recognize these diseases would be of significant value in the NICU.

A very appealing monitoring tool to automatically detect neonatal seizures or breathing disorders may be based on acquiring, through a network of sensors, e.g., a set of video cameras, the movements of the newborn's body (e.g., limbs, chest) and properly processing the relevant signals. An automatic multi-sensor system could be used to permanently monitor every patient in the NICU or specific patients at home. Furthermore, a wire-free technique may be more user-friendly and highly desirable when used with infants, in particular with newborns.

This work has focused on a reliable method to estimate the periodicity in pathological movements based on the use of the Maximum Likelihood (ML) criterion [9]. In particular, average differential luminance signals from multiple Red, Green and Blue (RGB) [10] cameras or depth-sensor devices [11] are extracted and the presence or absence of a significant periodicity is analysed in order to detect possible pathological conditions.

For this purpose, a research unit, led by Prof. Riccardo Raheli from the Department Information Engineering and Prof. Francesco Pisani from the Department of Neurosciences - Section of Child Neuropsychiatry, has been activated.

The efficacy of this monitoring system has been measured on the basis of video

recordings provided by the Department of Neurosciences.

Concerning clonic seizures, a kinematic analysis was performed to establish a relationship between neonatal seizures and human inborn pattern of quadrupedal locomotion.

Moreover, we have decided to realize simulators able to replicate the symptomatic movements characteristic of the diseases under consideration. The reasons is, essentially, the opportunity to have, at any time, a 'subject' on which to test the continuously evolving detection algorithms.

Finally, we have developed a smartphone App, called "Smartphone based contactless epilepsy detector" (SmartCED), able to detect neonatal clonic seizures and warn the user about the occurrence in real-time.

The rest of this thesis is organized as follows. In Chapter 1, the neonatal diseases are investigated from a medical point of view. In Chapter 2, a literature overview of previous works related to signal processing-based detection of neonatal diseases is presented. In Chapter 3, the clonic seizure detection system is shown. In Chapter 4, the apnoea detection system is treated. Chapter 4, show the kinematics analysis of neonatal seizures automatism. Chapter 5, presents hardware neonatal diseases simulators. Chapter 6, describes an Android application to monitor clonic seizure with a smartphone. Finally, conclusions are drawn.

Chapter 1

Neonatal Deseases

"Healing is a matter of time,
but it is sometimes also a matter of opportunity."

– Hippocrates

1.1 Diseases of the Central Nervous System (CNS)

The central nervous system is a fundamental system of our body, it analyses, interprets and processes data and information coming from various parts of the body. The central nervous system, besides providing sensations about the external environment, coordinates biological activities, voluntary and involuntary, and regulates the functioning of the peripheral apparatuses. We will now describe the pathologies of the central nervous system which directly affect this work.

1.1.1 Seizures

Seizures are defined as involuntary contraction of one or more muscle groups due to a paroxysmal neuronal discharge [1]. They are a symptom of malfunctioning of the central nervous system and they are an age-dependent phenomena [12]. Neona-

tal seizures have onset within the 28th day of life in newborns at term and within the 44th week of conceptional age in preterm[1] infants. The main causes of this disease are hypoxic-ischaemic encephalopathy, intracranial haemorrhage, and sepsis. The studies confirm the following disease incidence: 2.6% in live births, 11.1% for preterm neonates, and 13.5% for babies with less than 2.5 kg of weight [2]. Seizures in newborns have to be promptly and accurately recognized in order to establish timely treatments that could avoid an increase of the underlying brain damage.

1.1.2 Clinical Classification

A classification, proposed by Volpe, divides neonatal seizures in four main categories: subtle, clonic, myoclonic, and tonic [1].

Subtle seizures

A subtle seizure is the most prevalent type of neonatal seizure and, while it can be seen in both term and preterm infants, it is most frequently seen in preterm infants. Since, as the name implies, symptoms are subtle, this type of seizure is often unrecognized. They may be recognized on the basis of their typical clinical manifestations such as repetitive eye blinking, non-nutritive sucking motions, tremor of a limb, a hand or a finger, and movements like cycling, rowing, bicycling or swimming.

Clonic seizures

"Clonus" (KLOH-nus) means rapidly alternating contraction and relaxation of a muscle. Clonic seizures are characterized by rhythmic movements, with a typical frequency ranging from about 0.3 Hz to about 2 Hz. Their clinical manifestations are repetitive movements localized to a limb (focal clonic seizures) or simultaneous movements of the limbs of two different hemispheres (multifocal clonic seizures).

[1] Preterm birth is the birth of a baby at less than 37 weeks gestational age.

Myoclonic seizures

Myoclonic seizures are brief shock-like jerks of a muscle or group of muscles. Myoclonic seizures are involuntary, sudden and quick. Usually they do not last more than a second or two. There can be just one, but sometimes many will occur within a short time. Even people without epilepsy can experience myoclonic seizures in a sudden jerk that may wake them up as they are just falling asleep. These phenomena are normal.

Tonic seizures

The tonic seizures are rare and characterized by a tonic stiffening of the muscles, e.g., an extension of one or all limbs, they are typical of preterm newborns and are caused mostly by intraventricular hemorrhage causing a severe brain suffering.

1.2 Breathing Disorders

The great progress made in recent decades in neonatal medicine, the intensive care in the birthing room and the availability of new technologies, such as the use of mechanical ventilators, has led to a significant increase in the survival rate of preterm infants, also those with very low weight (less than 1500 g) and extremely low gestational age.

In Italy, the latest report of the Ministry of Health about the Certificates of childbirth assistance (CeDAP), published in 2014, shows that the percentage of births before the 37th and before the 31th weeks of gestation is, respectively, 6.8% and 0.8% of the total, and the percentage of babies born weighing less than 2500 g and 1500 g is 6.8% and 1.1% of the total.

The progressive increase in survival of preterm infants has, however, led to a greater demand for intensive cares, especially with regards to respiratory and neurological disorders: young patients, although alive, do not have a fully developed respiratory system and, thus, require a careful monitoring for early detection of possible apnoea events, both during hospitalization in a NICU and at home.

1.2.1 Apnoeas

Apnoea is defined as a breathing pause lasting more than 20 seconds, or between 10 and 20 seconds if associated with other clinical alterations such as desaturation, cyanosis, bradycardia, hypotonia and pallor. The apnoea is a non rare event in the newborn and its frequency is inversely proportional to gestational age. It is present indeed, in 7% of those born between the 34th and the 35th week of gestation, reaching 54% in those born between the 30th and 31st week [13] and 98-99% in those born before the 29 weeks of gestational age [14].

The cause of apnoea in premature newborns is essentially a condition of immaturity of the bulbar center of the breath function, which responds abnormally to hypoxia (reduction of oxigen in blood) and hypercapnia (increase of carbon dioxide in blood); it is also due to a inhibitory response exaggerated to laryngeal stimulation (stimulation receptor causes laryngeal cough in adults, but in premature, may cause apnea). Other causes of apnea are diseases of the central nervous system, including intraventricular hemorrhage, hypoxic-ischemic encephalopathy, seizures and infectious diseases.

Monitoring these young patients is essential not only during their stay in the NICU departments but, often, even at home after the resignation: it is recommended a monitoring period of about two months since the last crisis of apnea. The home monitoring is necessary even for those infants who have already experienced an Apparent Life-Threatening Events (ALTE) [15]. Moreover, it now seems certain that the apnoea monitoring is a reliable system for prevention of SIDS (Sudden Infant Death Syndrome) in children [16].

1.2.2 Congenital Central Hypoventilation Syndrome

Congenital Central Hypoventilation Syndrome (CCHS), also known as Ondine's curse, is a very rare disease with an incidence of 1/200000 live births; and it is characterized by the lack of control of breathing during sleep.

The term "curse of Ondine", used in the past to refer to the CCHS, is inspired by a mythological tale which tells that the mermaid Undine, discovering her husband

asleep in the arms of another woman, angry for the betrayal, cursed him, preventing him from breathing whenever he had fallen asleep.

This syndrome, defined as the inability to automatically control the respiratory activity, is due to the absence of response to hypoxia and hypercapnia by the autonomic nervous system. The trouble, in fact, is more evident during the non-rapid eye movement (REM) sleep when breath control is, essentially, chemical, on the contrary is less apparent during REM stages when there is a significant cortical control. The disease is genetic, autosomal dominant, linked to a gene PHOX2B mutation in chromosome 4 [5].

The male to female ratio is estimated as 1:1. In the world, 1000 cases were diagnosed related to the PHOX2B mutation, but it is estimated that the true prevalence is much higher because of the clinical variability of the disease. Symptoms appear in the first days of life, indeed, in some infants, just at the time of birth, so as to require resuscitation in the delivery room.

Normally babies have shallow breathing during sleep with respiratory rate preserved; but some babies may have unexplained apnoea episodes, which, in the most extreme cases, can cause death and be classified as SIDS (Sudden Infant Death Syndrome), others have cyanosis, heart failure and edema, and could be mistaken for children with heart problems. The CCHS does not respond to drug therapy and, therefore, these patients have to be managed with a permanent ventilatory support and continuous monitoring.

Children with CCHS are equal to everyone else, they play, smile and normally live during the day, but at night they suffer the damage of their pathology: to sleep, a natural activity for everyone, for them becomes a problem, because they are constantly attached to their mechanical respirator [17].

Chapter 2

Literature Overview

"I learned the hard way how desperately primitive is the technology we have for monitoring the health of someone with a chronic illness."

– Chris Toumazou

2.1 Medical Field

2.1.1 Clonic Seizures

During a seizure, the ElectroEncephaloGraphic (EEG) pattern will change, so, currently, the traditional method of diagnosis is based on EEG monitoring. The EEG analysis, however, has several drawbacks:

- it is a time-consuming technique
- it requires specialized medical personnel
- it is not always available in a Neonatal Intensive Care Unit (NICU)
- it is moderately invasive, because of several electrodes attached to the small body of newborns

- it is expensive.

Furthermore, it is crucial to stress that a normal EEG does not exclude epilepsy, as around 10% of patients with epilepsy never show epileptiform discharges [18].

Another critical point lies in the fact that the identification of seizures from neonatal EEG is currently based on interpretation of the graphical EEG record by medical personnel. Since this process is very subjective, reliable automated detection of electrical seizure activity in the newborn may have significant clinical applications.

2.1.2 Breathing Disorders

Current measurement systems, still largely manual and timeconsuming, are based on polysomnographic devices, which are composed by the following set of sensors:

- electroencephalogram (EEG)
- electromyography (EMG)
- electrocardiography (ECG)
- elastic belt
- pulse-oximeter
- nasal flow-meters.

However, these systems further emphasize the problems listed before regarding the seizures diagnostic systems. In fact, the number of electrodes and mechanical sensors even higher increases the cost and makes the system uncomfortable to use because of the large number sensors attached to the patient's body. Therefore, automatic, real-time and non-invasive monitoring equipments able to reliably recognize these diseases would be of significant value in the NICU. In Fig. 2.1, a polysomnographic system is shown.

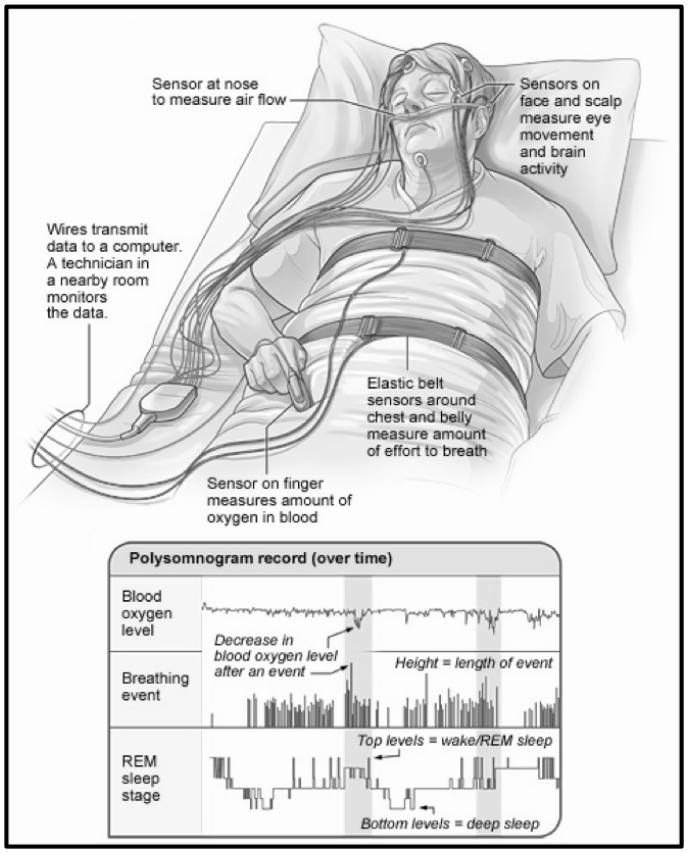


Figure 2.1: Polysomnographic System [19].

2.2 Engineering Field

In order to overcome the problems of the existing systems for automatic monitoring, many methods have been proposed, with different technologies.

2.2.1 RGB sensor

A fundamental task, in order to detect neonatal seizures by video signal processing is quantifying the motion of the moving parts of a newborn's body.

In [20], two methods, developed to extract quantitative information from video recordings of neonatal seizures in the form of temporal motion strength and motor activity signals, are presented. Motion strength signals are extracted by measuring the area of the body parts that move during the seizure and the relative speed of motion using a combination of spatio-temporal sub-band decomposition of video, non-linear filtering, and segmentation. Motor activity signals are extracted by tracking selected anatomical sites during the seizure using a modified version of a feature-tracking procedure developed for video, known as the Lucas-Kanade algorithm [21]. The experiments indicate that the temporal signals produced by the proposed methods provide the basis for differentiating myoclonic from focal clonic seizures and distinguishing these types of neonatal seizures from normal infant behaviours.

In [22], two different optical flow computation methods are used to extract motion strength signals from video recordings of neonatal seizures in order to characterize the movements of newborns during a long monitoring time.

In [23], a method by which a subset of motor seizures can be detected using only remote sensing devices such as video cameras is presented. These major motor seizures (MMS) consist of clonic movements and are often precursors of generalized tonic-clonic (convulsive) seizures, sometimes leading to a condition known as status epilepticus, which is an acute life-threatening event. This paper proposes an algorithm based on optical flow extraction of global group transformation velocities, and band-pass temporal filtering to identify occurrence of clonic movements in video sequences.

In [24], once again, the optical flow algorithm, applied to video recordings, is used to detect movement during sleep in pediatric patients with epilepsy. This paper investigates data extracted by the optical flow algorithm in order to make a distinction between movement and non-movement in nocturnal recordings from pediatric patients, up to 18-year old, with epilepsy.

2.2.2 Automatic EEG-based Seizure Detection System

In [25], a system for epileptic seizure detection in electroencephalography (EEG) is described. One of the challenges is to distinguish rhythmic discharges from nonsta-

tionary patterns occurring during seizures. The proposed approach is based on an adaptive and localized time-frequency representation of EEG signals by means of rational functions. The corresponding rational discrete short-time Fourier transform (DSTFT) is a novel feature extraction technique for epileptic EEG data. A multilayer perceptron classifier is fed by the coefficients of the rational DSTFT in order to separate seizure epochs from seizure-free epochs. With this method it is possible to obtain very good performance reaching an accuracy equal to 98%.

In [26], a state-space model with Cauchy observation noise (SSMC) to detect seizure onset in a long-term EEG monitoring system, is proposed. Facing the challenge of high false detection rates (FDRs) in many existing methods caused by impulsive electrooculography (EOG)/EMG artifacts, the SSMC uses a nonlinear state-space model to encode the gradual changes of epileptic seizure signals and reject abrupt changes brought by artifacts. The Cauchy distribution is proposed to model impulsive observation noises because this heavy-tailed distribution is better at capturing abrupt changes in noise than Gaussian, thus reducing false alarms. A very good result is obtained in term of seizures correctly detected, but several cases of false alarm still remain.

2.2.3 Smart Bed

In [27], a system for obtaining both electrocardiographic potential and respiratory variation signals from infants lying supine during sleep, is proposed. A capacitively coupled bed-sheet electrode unit and a methodology of capacitive sensing were employed. A prototype device was fabricated to explore the feasibility of the approach and evaluation experiments were conducted for 5 infants aged from 43 to 174 days. In Fig. 2.2, the smart-bed system is shown.

2.2.4 Wearable

In [28], a sensory baby vest includes fully integrated sensors for the parameters respiration, heart rate, temperature and humidity, to detect excessive sweating, for the continuous monitoring of infants under clinical and home conditions, is presented.

Figure 2.2: Smart-Bed System [27].

It may allow the early detection of potential life threatening events calling for rescue as well as the recognition of the development or progression of diseases at an early stage. A variety of principles for the measurement of the parameters has been assessed for the integration into clothes. Prototypes have been manufactured incorporating the chosen sensing principles with textile and textile-compatible technologies. In Fig. 2.3, the vest worn by a doll toy, is shown.

In [29], the realization of a garment embedded patient monitoring system, including wireless communication and inductive powering, is treated. The developed system is primarily intended for the continuous monitoring of the electrocardiogram (ECG) of children with an increased risk of sudden infant death syndrome (SIDS). The sensors and the antenna are made out of textile materials. All electronics are mounted on a flexible circuit to facilitate integration in the baby's pyjamas.

Figure 2.3: The baby vest worn by a doll toy [28].

2.2.5 Ultrasound

In [30], a contact-free system that used reflected ultrasound to separately monitor respiratory, non-respiratory, and caretaker movements of infants, is described and validated. In Fig. 2.4, the diagram of the proposed system, is shown.

2.2.6 Wireless Networks

In [31], it is shown experimentally that standard wireless networks which measure received signal strength (RSS) can be used to reliably monitor human breathing and estimate the breathing rate. As a first order approximation, breathing induces sinusoidal variation in the measured RSS on a link, with amplitude which is a function of the relative amplitude and phase of the breathing-affected multipath. Although an individual link may not reliably detect breathing, the collective data sensed by a network of devices may reliably indicate the presence and rate of breathing. In Fig. 2.5, the diagram of the proposed system, is shown.

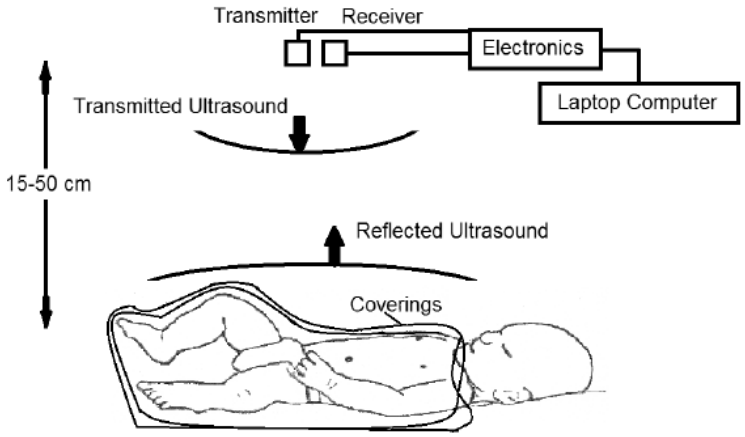


Figure 2.4: Diagram of the Ultrasound-based System [30].

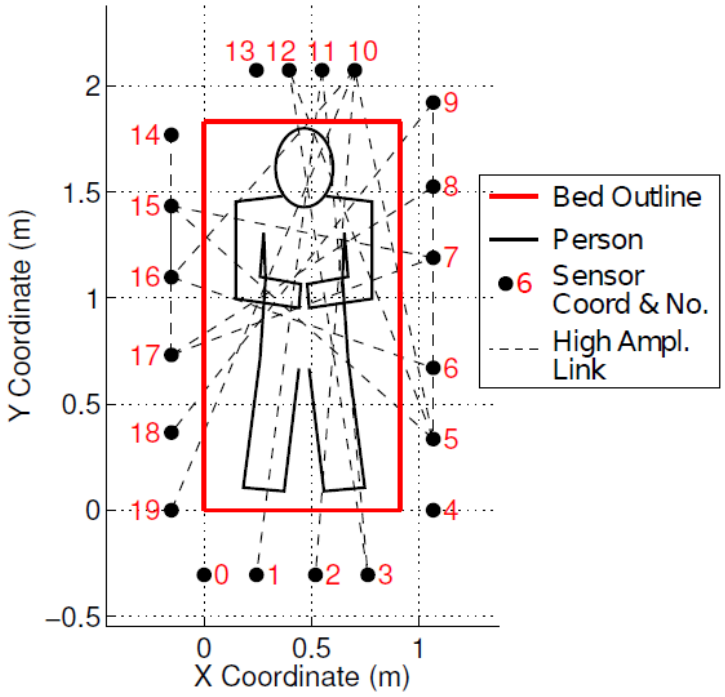


Figure 2.5: Diagram of the wireless networks system [31].

Chapter 3

Clonic Seizures Detection Systems

"The only way you multiply resources is with technology.
To really affect poverty, energy, health, education, or anything else,
there is no other way."

– Vinod Khosla

3.1 Extraction of Motion Information

In this chapter, we present a reliable method to estimate the periodicity in pathological movements based on the use of the Maximum Likelihood (ML) criterion. In particular, average differential motion signals from multiple sensors are extracted and the presence or absence of a significant periodicity is analysed in order to detect possible pathological conditions.

3.1.1 Motion Signal Extraction

The first step of the proposed algorithm consists of the extraction of a relevant motion signal from every sensor. We focus on the signal extraction from RGB cameras but the approach is similar also using a depth sensor, as it will be explained later. As

described in [32, 33], we consider a generic video signal composed of a sequence of frames with period T, where a frame at discrete time[1] i is a matrix of $W \times H$ pixels. Before processing the signal, in order to limit the computational complexity, we convert the RGB frames to gray scale and denote them by the matrix $\mathbf{Y}_s[i]$, where the subscript s identifies the s-th sensor, with pixel intensity levels ranging from 0 to 255.

After gray scale conversion, we perform a simple filtering based on the difference between consecutive frames. Accordingly, the output frame $\mathbf{Z}[i]$ can be expressed as follows:

$$\mathbf{Z}_s[i] = |\mathbf{Y}_s[i+1] - \mathbf{Y}_s[i]| \quad i = 0,1,2,3,\ldots. \tag{3.1}$$

In this way, the moving parts are emphasized in the output frame sequence.

In order to further limit the complexity and reduce residual noise, the resulting frames are converted to a binary scale. This operation is described by the following simple rule:

$$\begin{cases} B_s[x,y,i] = 1 & \text{if } Z_s[x,y,i] > t_B \\ B_s[x,y,i] = 0 & \text{else} \end{cases} \tag{3.2}$$

where $B_s[x,y,i]$ and $Z_s[x,y,i]$ denote the entries of matrices $\mathbf{B}_s[i]$ and $\mathbf{Z}_s[i]$ at position $[x,y]$, respectively, and $t_B \in [0,255]$ is an appropriate binary image conversion threshold. In particularly this threshold is heuristically chosen in order to maximize, in the obtained binary image, the ratio between the number of white pixels of the body parts most affected by motion and the number of remaining white pixels of the frame (including the remaining body parts), so that the obtained motion signal consists mainly of the pixels of the moving body parts [32, 33].

After conversion to binary scale, we apply a morphological operation called *erosion* [10], that allows the reduction of some remaining noise. After erosion, some noisy (i.e., bright and isolated) pixels may still exist but their number is negligible with respect to the number of pixels of the moving parts. Obviously, erosion leads

[1] The time instant is specified as an integer multiple of T.

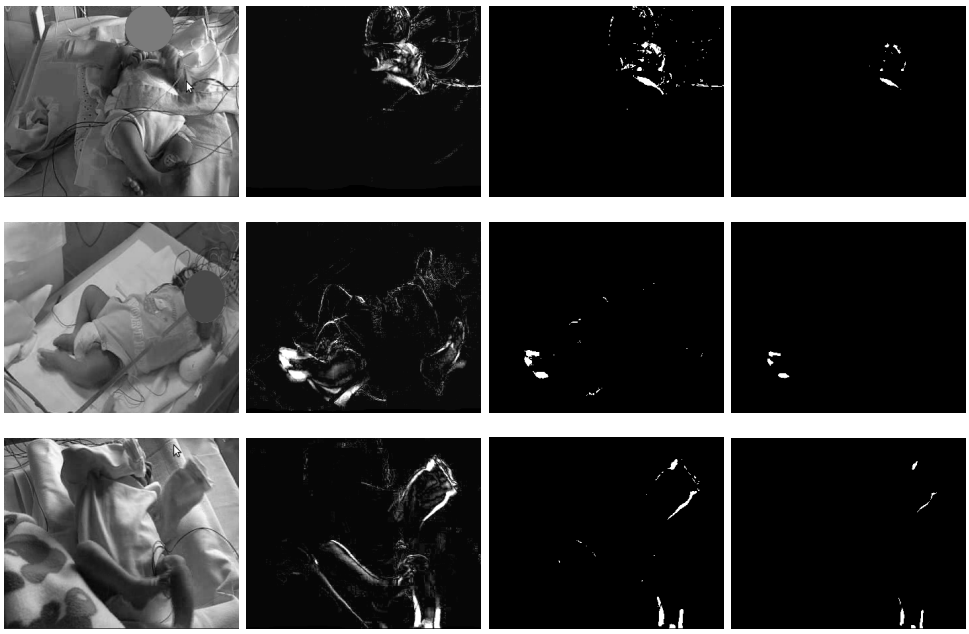


Figure 3.1: Motion signal extraction: gray scale, difference filtering, binary conversion and erosion.

to the loss of a portion of the signal; this loss, however, is negligible. The resulting frame matrix is denoted as $\mathbf{I}_s[i]$.

In Fig. 3.1, a few illustrative examples of motion signal extraction are shown: the rows are associated with various videos or video cameras and the columns with the described processing steps.

Taking into account that bright binary pixels have a luminance value equal to 1, the average luminance signal, denoted as $\overline{L}_s[i]$, is defined as follows:

$$\overline{L}_s[i] \triangleq \frac{1}{WH} \sum_{x=1}^{W} \sum_{y=1}^{H} I_s[x,y,i]. \tag{3.3}$$

In other words, the signal $\overline{L}_s[i]$ at the i-th frame is the fraction of white pixels in the filtered binary image from the s-th sensor. As shown in [32, 33], this signal may represent the movement "pattern" of the involved body parts, thus, $\overline{L}_s[i]$ represents

the motion signal from the s-th sensor.

In Fig. 3.2, an example of normalized[2] motion signal related to a clonic seizure and an example of motion signal related to random movements, extracted from a single RGB camera are shown.

3.1.2 Detection of Periodicity

Once the motion signal from every RGB camera has been extracted, we need a method to decide on the presence of a periodic movement. Since clonic seizures fundamentally involve a quasi periodic movement, our aim is to determine whether the extracted signal has a periodic component and, if so, to estimate it [32]. Suppose that the motion signal represents the clonic seizure motion, i.e., seizures cause the only relevant body movements. In this case, every motion signal, namely, the average luminance signal in (3.3), can be modelled as:

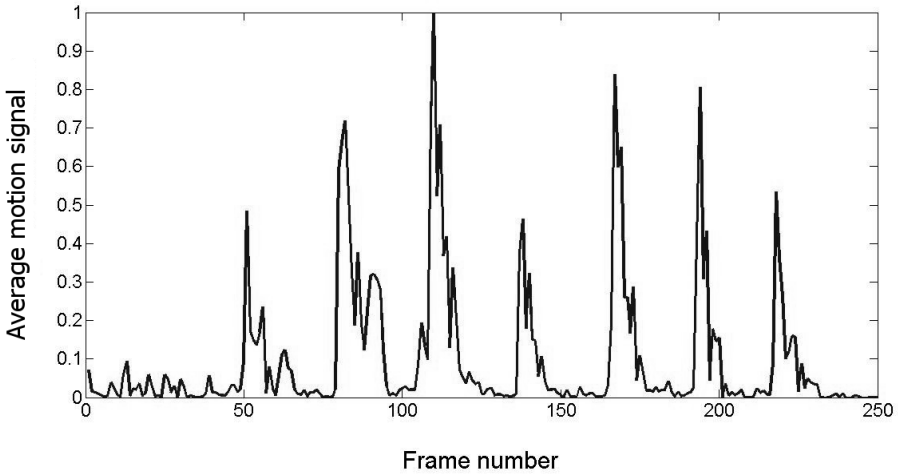
$$\overline{L}_s[i] = c_s + A_s \cos[2\pi f T i + \phi_s] + n_s[i] \tag{3.4}$$

where: c_s is the continuous component of the s-th sensor; $\{n_s[i]\}$ is a sequence of independent identically distributed (i.i.d.) zero-mean Gaussian noise samples; and A_s, ϕ_s, f are the amplitude, phase, and frequency, respectively, of the periodic component of the signal from the s-th sensor. We assume that the periodic component due to pathological movements has the same frequency on all sensors, and that the sampling period is identical on all signals. Phase and amplitude are expected to differ between sensors. Our goal is to detect the presence of the sinusoidal component and estimate its frequency f. The signals coming from multiple sensors have to be aggregated using data fusion principles.

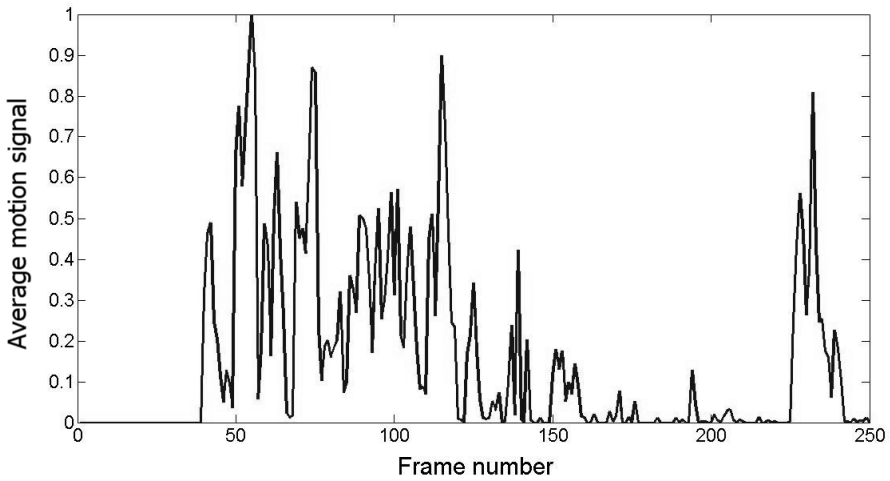
Given the vector of parameters $\boldsymbol{\theta} = [A, f, \phi]$, the ML estimation strategy of $\boldsymbol{\theta}$ consists of the minimization of the following likelihood function [9]:

$$J(\boldsymbol{\theta}) \triangleq \sum_{s=1}^{S} \sum_{i=1}^{N} \left(\overline{L}_s[i] - A_s \cos[2\pi f T i + \phi_s]\right)^2 \tag{3.5}$$

[2] Between 0 and 1.



(a) Periodic movements



(b) Random movements

Figure 3.2: Examples of extracted luminance signals: (a) periodic movements (related to a clonic seizure) and (b) random movements.

where N is the number of consecutive frames (i.e., an N-frame window) and S is the number of sensors.

Under the assumption of independent and identically distributed zero-mean Gaussian noise, after standard algebraic manipulations, a good approximation of the optimum strategy is [31]:

$$\hat{f} = \underset{f}{\operatorname{argmax}} \sum_{s=1}^{S} \left| \sum_{i=1}^{N} \overline{L}_s[i] e^{-j2\pi f T i} \right|^2 . \tag{3.6}$$

Note that it is possible to estimate the frequency from one sensor alone using eq. (3.6) with $S = 1$. In this case, the s-th term of the outer sum in (3.6) represents the periodogram of $\overline{L}_s[i]$ and, thus, its peak indicates the most significant harmonic component and corresponds to the estimated frequency.

Once the frequency has been estimated, every amplitude components $\hat{A}_s$ can be estimated as [31]:

$$\hat{A}_s = \frac{2}{N} \left| \sum_{i=1}^{N} \overline{L}_s[i] e^{-j2\pi \hat{f} T i} \right| \quad s = 1, 2, ..., S \tag{3.7}$$

We may thus declare that a significant periodic component is present if [31]

$$\frac{N}{S} \sum_{s=1}^{S} \hat{A}_s^2 > \eta \tag{3.8}$$

where the threshold η is experimentally determined in order to maximize the number of pathological movements correctly identified.

An example, relative to a clonic seizure, of likelihood function with frequency estimate is shown in Fig. 3.3.

The described ML method can be applied on length-N (in terms of frames) disjoint observation windows; in other words, the periodicity can be analysed window by window. However, in order to improve the reliability of the proposed method, one can analyse the extracted average motion signal on half-interlaced consecutive windows.

In fact, a clonic seizure could manifest itself across two consecutive disjoint windows: a single window-based algorithm could thus miss it in each of the two windows. In the remainder of this paper, we will then consider interlaced windows. In

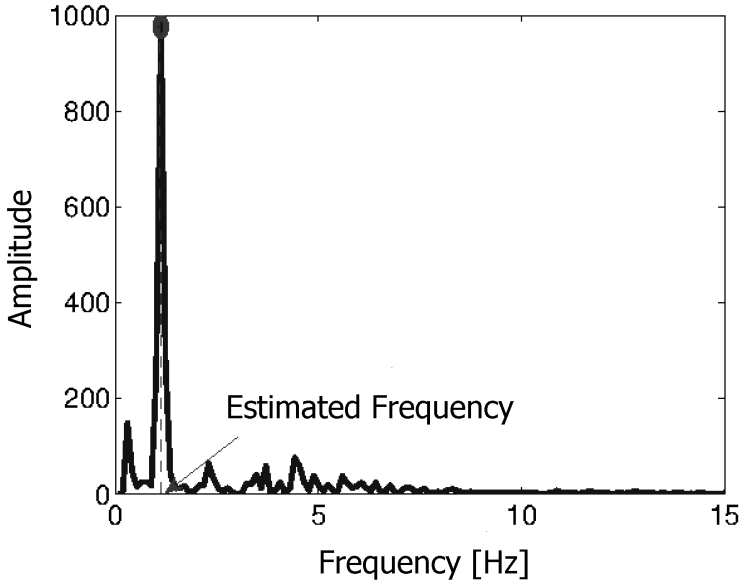


Figure 3.3: Likelihood function and relative estimated frequency estimation.

Fig. 3.4, an illustrative example of interlaced observations windows relating to quasi periodic movements in the context of neonatal seizure detection is shown [32, 33].

3.2 Performance Analysis

In neonatal neurology, the shortest time interval necessary to characterize a clonic seizure is conventionally set to 10 s [1] and the estimated period must be in the range from 0.5 s to 3.5 s [34]. Therefore, we apply the algorithm described in Section (3.1.2) in temporal windows of 10 seconds, thus the value of N in (3.6) is given by $N = f_s * 10$ where f_s is the sampling frequency. For example, if the used video-cameras operate at 25 frames/s, N is set to 250.

The performance of the proposed detection system is analyzed considering a binary test to classify the results into two groups: presence of clonic seizures in the video of the newborn (positive) and presence of random movements (negative). In

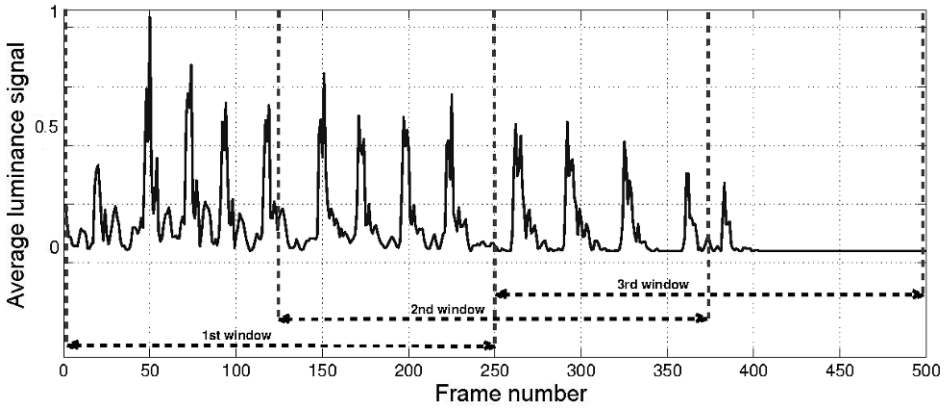


Figure 3.4: Interlaced windows.

the medical field, the performance of a test is characterized by two statistical measures, namely *sensitivity* and *specificity*, which are defined as follows [35]:

$$SE = \frac{TP}{TP+FN} \tag{3.9}$$

$$SP = \frac{TN}{TN+FP} \tag{3.10}$$

where TP, TN, FP, and FN denote, respectively, the numbers of True Positives, True Negatives, False Positives, and False Negatives in the sequence of all windows composing the signal. A sensitivity of 100% means that the test recognizes all actual positives, i.e., in this scenario all sick children would be recognized as such. Similarly, a specificity of 100% means that the test is able to correctly recognize all negative cases, i.e., all healthy children are recognized as such. In the terminology of detection theory [36], sensitivity and specificity are empirical measures of the probabilities of correct detection and correct rejection, respectively.

Table 3.1: Clonic seizures detection (one RGB sensor)

	Real positive	Real Negative
Positive Test	$TP = 51$	$FP = 16$
Negative Test	$FN = 7$	$TN = 210$
	$SE = 88\%$	$SP = 93\%$

3.3 Single Sensor

We present, initially, results based on the retrospective off-line analysis of 10 single camera-based video recordings of newborns.[3] We consider a single sensor, i.e., $S = 1$ in (3.8). The set of videos has been chosen to represent all possible behaviours of a newborn (quiet, natural movements, pathological movements). Each video recording has the following characteristics:

- video sampling frequency: 25 frames/s
- video resolution: 360×288 pixels
- video length: 5 min.

The performance evaluation of the proposed algorithm in Section (3.1) has been based on 284 of the 580 available interlaced windows, because windows with no movements were discarded. Had not we discarded these windows, we would have had an unrealistic increase in specificity (3.10), because all the windows without movements would be too easily classified by the algorithm as true negatives. The corresponding results are shown in Table 3.1.

In Fig. 3.5, the performance results, in terms of Receiver Operating Characteristic (ROC) curves [37], are shown. The ROC curve is obtained testing the sample videos for varying values of η in (3.8). The obtained minimum distance D from the "ideal" operational point, i.e., the point associated with sensitivity and specificity equal to

[3] This retrospective study was approved by the Ethical local Committee and a parent of each newborn provided written informed consent.

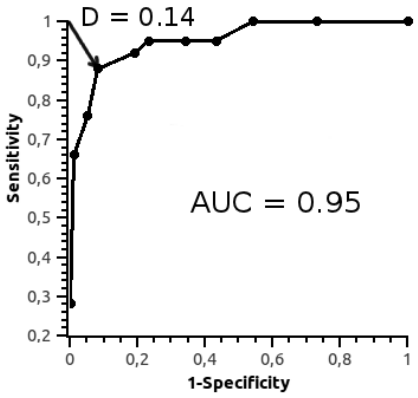


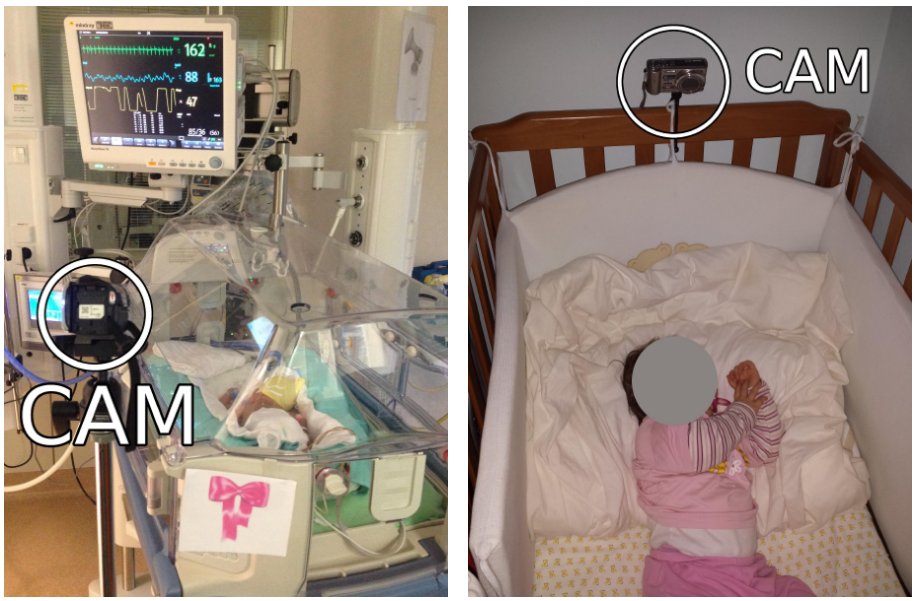
Figure 3.5: ROC Curve, using one RGB sensor.

1, is 0.14 and the Area Under Curve (AUC) is 0.95. We remark that the medical literature considers highly reliable a test with AUC between 0.9 and 1 [37].

As shown in Fig. 3.6, this method could be used for 24-hours-available, low-cost, and non-invasive monitoring in both hospitals, as shown in Fig. 3.6(a), and domestic environments, as shown in Fig. 3.6(b), possibly integrated with a telemedicine system.

3.4 Multiple Sensors

Analysis of the data obtained from multiple sensors placed around a patient, make it possible to increase the reliability of the detection system. This approach is very versatile and allowed us to investigate various scenarios, including: a single RGB camera, an RGB-Depth sensor and a network of a few RGB cameras. Data fusion principles are considered to aggregate the signals from multiple sensors.



(a) (b)

Figure 3.6: Two examples of video monitoring: (a) in a neonatal intensive care unit and (b) in a domestic environment.

3.4.1 Multiple Channels

Using the same set of video recordings described in Section 3.3, we have investigated the performance which can be obtained using just one color channel of the RGB sensor. This option is equivalent to a particular linear combination in which only one of the three weights w_r, w_g and w_b is set to 1, whereas the other two are set to 0. The corresponding results are shown in Table 3.2 for the Red channel, Table 3.3 for the Green channel, Table 3.4 for the Blue channel. These tables can also be compared with Table 3.1 relative to the standard RGB linear conversion to gray scale. From this analysis, it is evident that the signal coming from the green channel is sufficient to obtain almost the same performance guaranteed by the use of the standard linear combination of the three channels.

Multi-sensor video recordings make it possible to increase the reliability of the

Table 3.2: Clonic seizures detection (Red channel)

	Real positive	Real Negative
Positive Test	$TP = 44$	$FP = 8$
Negative Test	$FN = 14$	$TN = 218$
	$SE = 76\%$	$SP = 96\%$

Table 3.3: Clonic seizures detection (Green channel)

	Real positive	Real Negative
Positive Test	$TP = 49$	$FP = 10$
Negative Test	$FN = 8$	$TN = 217$
	$SE = 86\%$	$SP = 96\%$

Table 3.4: Clonic seizures detection (Blue channel)

	Real positive	Real Negative
Positive Test	$TP = 44$	$FP = 13$
Negative Test	$FN = 13$	$TN = 214$
	$SE = 77\%$	$SP = 94\%$

Table 3.5: Clonic seizures detection - 3 Channels

	Real positive	Real Negative
Positive Test	$TP = 50$	$FP = 9$
Negative Test	$FN = 7$	$TN = 218$
	$SE = 88\%$	$SP = 96\%$

detection system. With this in mind, we analysed the performance of the system using the three channels of the RGB sensor as if they were single monochrome sensors, combining the extracted signals by means of the ML criterion. In Fig. 3.7 the block diagram of the multi-channel detection system, is shown. In Table 3.5, the results obtained are shown. It is important to underline how the use of the ML criterion allows to achieve a performance better than or equal to that of any linear combinations

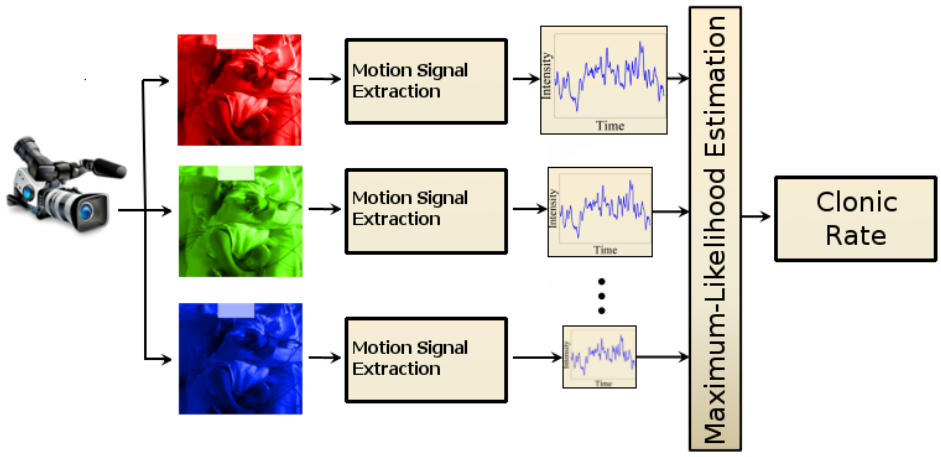


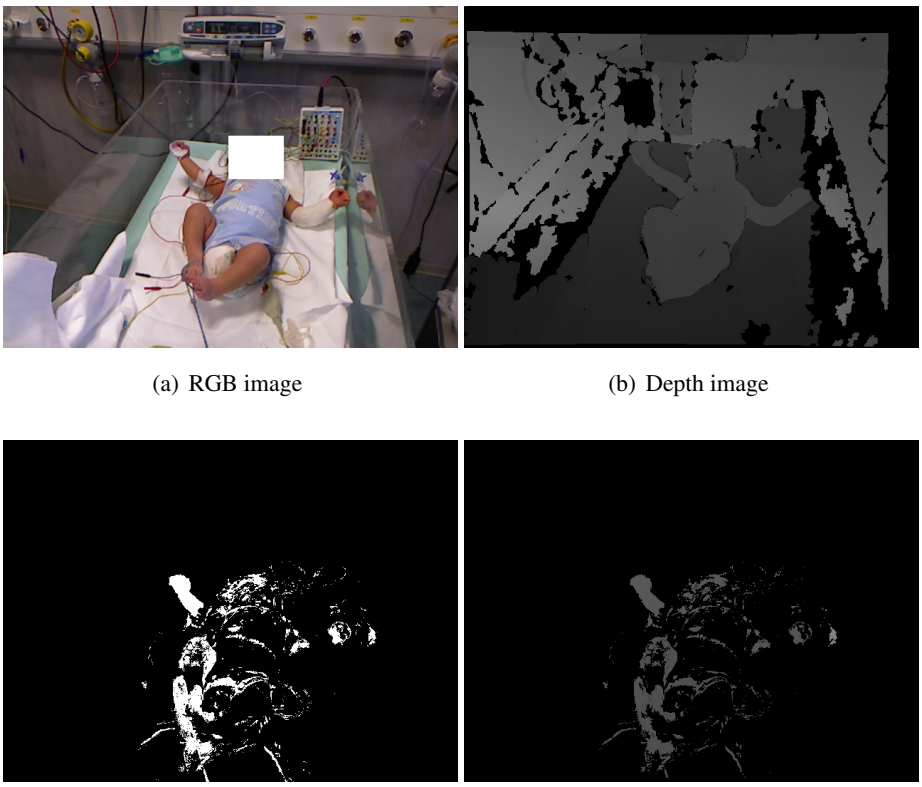
Figure 3.7: Block Diagram of the Multi-Channel Clonic Seizures Detection System

of the 3 channels.

3.4.2 RGB - Depth Sensors

The motivation for the use of a Kinect device [11] is that it provides a depth map associated with each RGB image. This depth information about the undergoing movement can improve the ability of a video-based system to detect pathological movements, despite a significant issue on the provided depth map, namely the shadowing noise. This noise is a systematic problem common to all structured-light approaches which use an offset camera to determine a depth map: there are regions where the projected pattern is shadowed by foreground objects, making it impossible to estimate the corresponding depth. This makes it difficult to apply a simple difference between consecutive depth frames, because in the resulting image, besides movement parts, the background could stand out.

In order to limit this noise, a background-foreground segmentation approach can be used [38, 39]. The goal is to construct a foreground mask, i.e., a binary image of the same size of the video frame, where a white pixel indicates that such a pixel



(a) RGB image (b) Depth image

(c) Foreground mask image (d) Filtered depth image

Figure 3.8: Background-foreground segmentation.

in the original image belongs to the foreground and a black pixel indicates that the pixel belongs to the background. The idea is to apply the background-foreground segmentation to the RGB frame and to apply the obtained mask to the depth frame, in order to filter all unnecessary pixels. For this purpose, we use the Mixture-of-Gaussian method described in [38], which is based on [39]. In Fig. 3.8, an RGB image, the relative depth map, the foreground mask extracted from the RGB image, and the depth map filtered by the foreground mask are shown.

After denoising the depth sequence, according to the above procedure, the motion signal can be extracted with the same technique described in Subsection (3.1.1),

Table 3.6: Clonic seizures detection (RGB and depth sensors)

	Real positive	Real Negative
Positive Test	$TP = 138$	$FP = 10$
Negative Test	$FN = 12$	$TN = 78$
	$SE = 92\%$	$SP = 88\%$

namely as follow:

$$\overline{L}_D[i] \triangleq \frac{1}{WH} \sum_{x=1}^{W} \sum_{y=1}^{H} D[x,y,i] \tag{3.11}$$

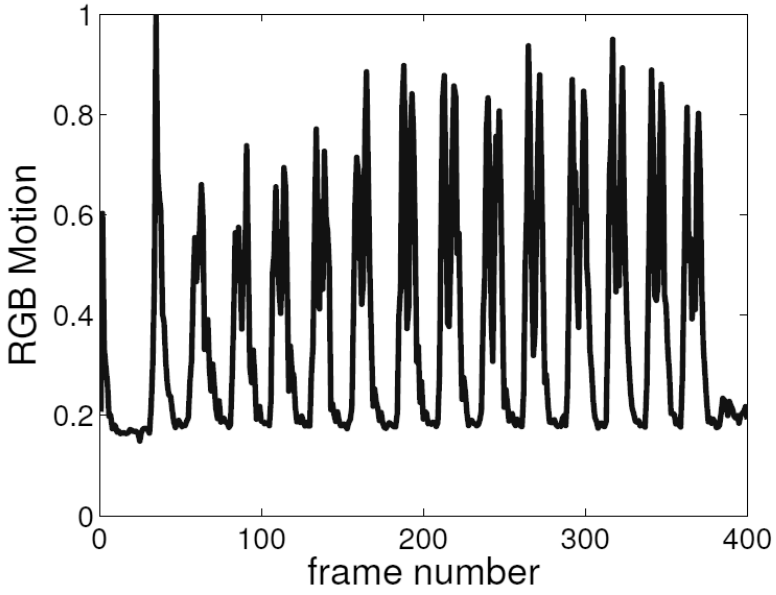
where $D[x,y,i]$ is the entry of the matrix $\boldsymbol{D}[i]$, the resulting depth image at the i-th frame after binary conversion and erosion, at the coordinates (x,y). It is now possible to estimate the frequency as in the multiple sensors case.

In Fig. 3.9 , the signals extracted from the RGB sensor and from the depth sensor of a Microsoft Kinect device, during a clonic seizure in a newborn, are shown.

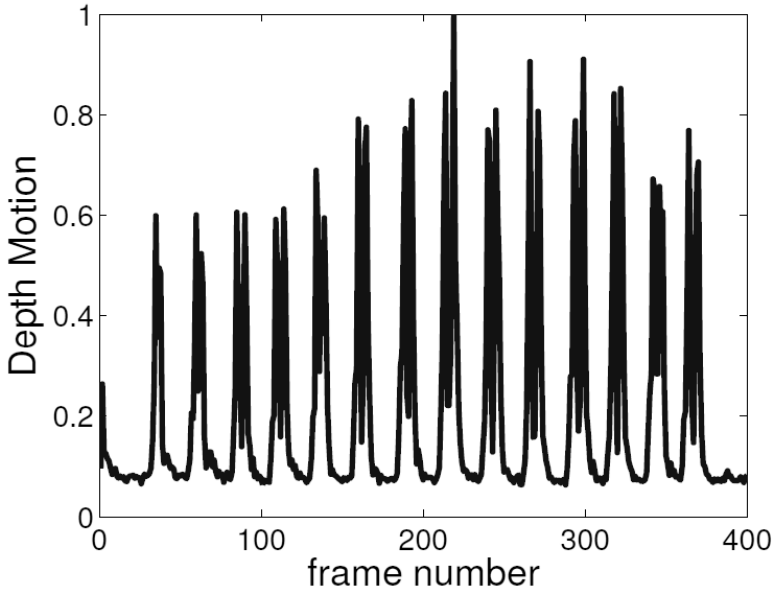
We present results based on the analysis of two video recordings of a newborn. In the first video, the newborn is affected by clonic seizures, while in the second one, the newborn performs only random movements. These videos were acquired using an RGB sensor and a depth sensor, i.e., $S = 2$ in (3.8). Each video recording has the following characteristics:

- video sampling frequency: 30 frames/s;
- video resolution: 640 $\times$ 480 pixels;
- video length: 10 min.

The performance evaluation of the proposed algorithm in Section (3.1) is then carried out considering 238 observation windows. The corresponding results are shown in Table 3.6.



(a) Motion signal extracted from RGB sensor



(b) Motion signal extracted from Depth sensor

Figure 3.9: Signals extracted from the Kinect device, during a clonic seizure in a newborn.

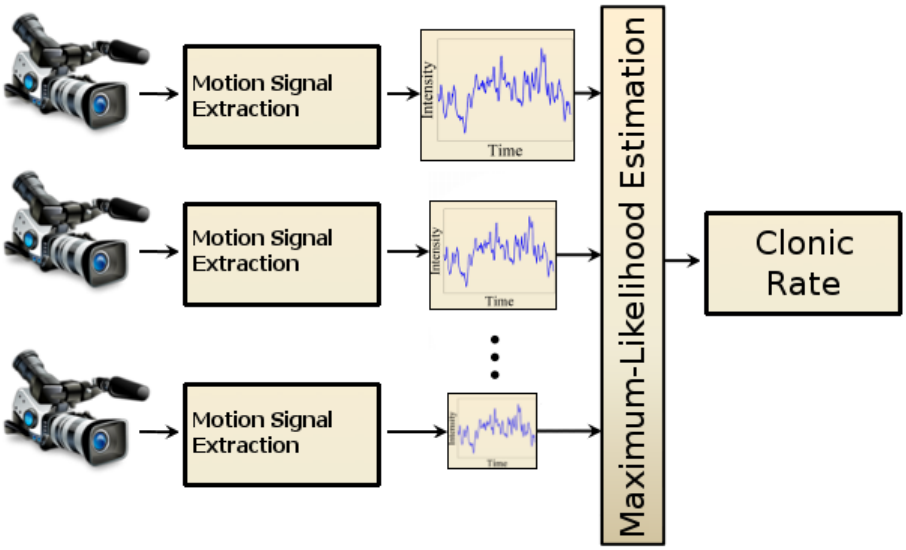


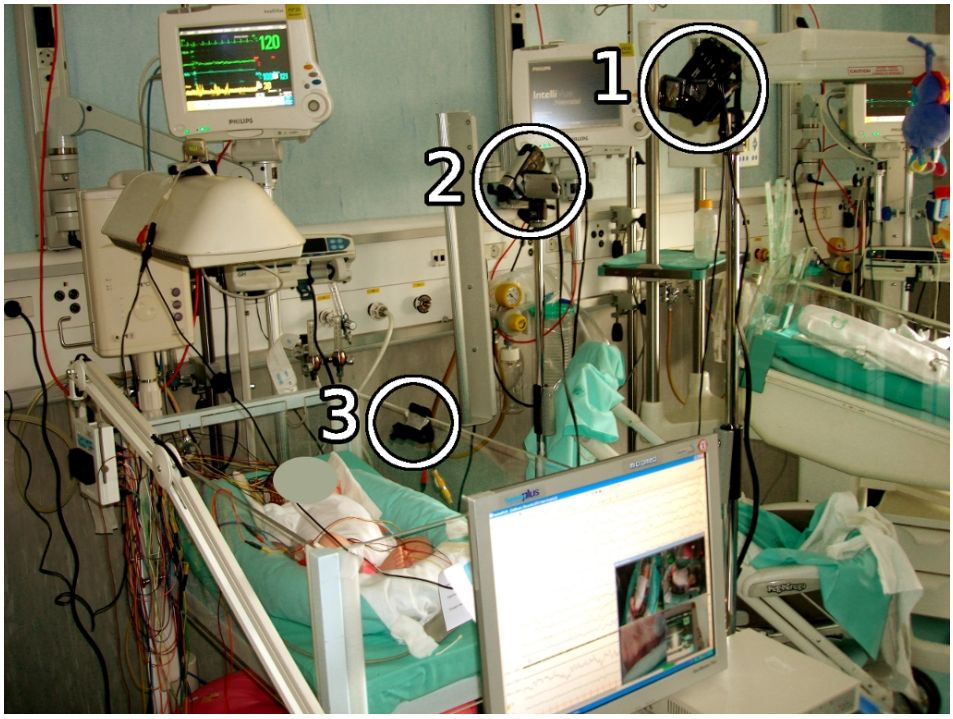
Figure 3.10: Block diagram of the multi-RGB sensor clonic seizures detection system.

3.4.3 Multiple RGB Sensors

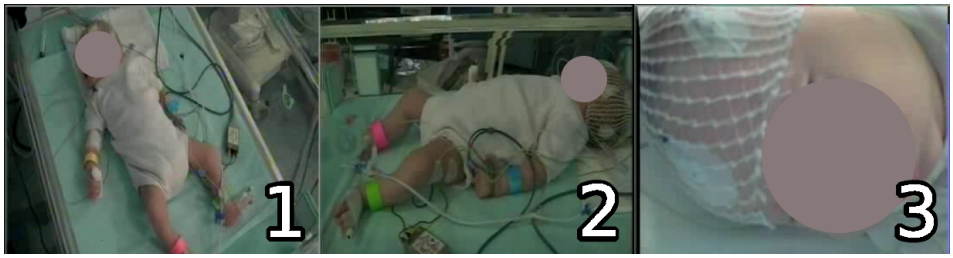
In Fig. 3.10, the block diagram of the multi-RGB sensors detection system is shown. In Fig. 3.11, the acquisition system used at the NICU of the University-Hospital of Parma and an example of acquired images are shown. The system is composed of 3 cameras: the first camera is positioned in front of the cradle; the second camera is positioned laterally to the cradle; and the third camera is set at the edge of the cradle in order to frame the face of the newborn.

We present the results based on the analysis of four recordings of a newborn with three cameras positioned as shown in Fig. 3.11, i.e., $S = 3$ in (3.8). In the first two videos, the newborn is affected by clonic seizures, while in the second two videos the newborn performs only random movements. Each video recording has the following characteristics:

- video sample frequency: 25 frames/s;



(a)



(b)

Figure 3.11: (a) Acquisition system; 1,2,3: position of the cameras and (b) corresponding frames.

Table 3.7: Clonic seizures detection (multiple cameras, 3 RGB sensors)

	SE	SP
Only Camera 1	55 %	77 %
Only Camera 2	82 %	73 %
Only Camera 3	64 %	82 %
Cameras 1 and 2	82 %	86 %
Cameras 1 and 3	64 %	82 %
Cameras 2 and 3	86 %	77 %
Cameras 1, 2 and 3	91 %	91 %

- video resolution: 360×288 pixels;
- video length: 1 min.

The performance evaluation of the proposed algorithm in Section (3.1) is then carried out considering 44 observation windows. In Table 3.7, the results obtained using individual cameras are compared with results obtained using various combinations of the three available cameras. It is clearly visible how the use of multiple cameras may significantly increase the system performance, especially in the presence of clonic movements of limited amplitude, as in the case of the considered videos.

Chapter 4

Breathing Disorders Detection Systems

"Now that I have kids, I'm probably more overprotective than I've ever been.
My wife's nickname for me is 'red alert.'
I sometimes check just to see if the kids are breathing."

– Matt Damon

4.1 Use of Eulerian Video Magnification

Detection of subtle motions for clinical analysis is a remarkable problem in biomedical monitoring based on video cameras. In fact, small movements can be confused with noise or small environmental changes in the scene. All of the methods mentioned in Chapter 3 may not be sensitive enough to handle these kinds of movements, especially for breathing disorders, because the small excursion of a newborn's chest is difficult to detect. A solution for improved detection was proposed in [40], where the presented algorithm can overcome spatio-temporal limitations of the human visual system, for revealing subtle changes in video signals. The Eulerian Video Magnification (EVM) algorithm enhances small changes in the video by a cascade combination

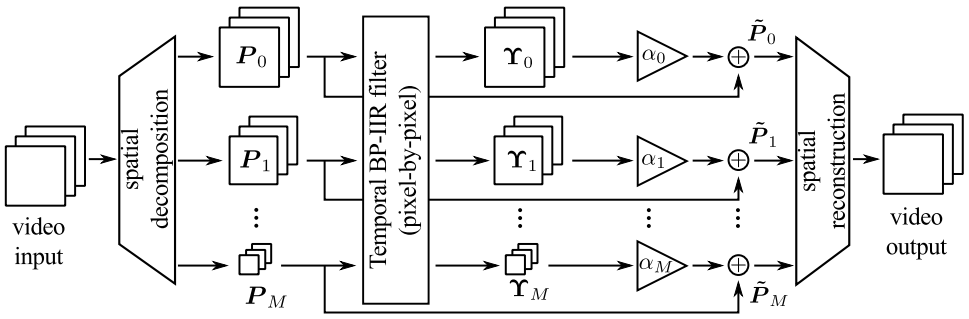


Figure 4.1: Block diagram of the EVM algorithm.

of a frame-by-frame spatial processing and a temporal processing on the pixel intensity variation [40]. Therefore, in the monitoring system proposed in Chapter 3, we can pre-process the video signals [41, 42] in order to amplify respiratory motion. In Fig. 4.1, a block diagram of the EVM algorithm is presented. The EVM approach consists of the following four main steps.

Spatial decomposition

The first step has the goal to decompose every video frame into a set of images, each representing different spatial frequency bands. Decomposition in sub-bands is performed by computing a Laplacian pyramid [43]. An image pyramid is a hierarchical representation of an original image $F[x,y,i]$, which is described by a set of M images, called levels, at different resolutions. Each level of the Laplacian pyramid is obtained starting from the corresponding levels of a Gaussian pyramid [43]. The Gaussian name comes from the fact that the different levels of the pyramid are obtained by repeatedly applying a low pass Gaussian filter.

Specifically, a Gaussian pyramid decomposition is obtained starting from the original image $G_0[x,y,i] = F[x,y,i]$, set as bottom level of the pyramid, and deriving the upper levels, specified by matrices $\boldsymbol{G}_1[i]$, $\boldsymbol{G}_2[i]$,...,$\boldsymbol{G}_M[i]$, as "reduced" versions of the previous one obtained by spatial low-pass filtering and undersampling. An image

at the l-th level (at discrete time i) can be formally expressed as follows:

$$\begin{aligned} G_l[x,y,i] &= \texttt{RED}\left(G_{l-1}[x,y,i]\right) \\ &= \sum_{m=-R}^{+R}\sum_{n=-R}^{+R} w[m,n]G_{l-1}\left[2x-m,2y-n,i\right] \end{aligned} \tag{4.1}$$

where $l = 1,\ldots,M$, $\texttt{RED}(.)$ is the reducing function, $w[x,y]$ is the low-pass filtering mask consisting of $2R+1$ rows and columns (size of the filter), R is a positive integer and $\boldsymbol{G}_{l-1}$ is the previous level of the pyramid. A reverse operator $\texttt{EXP}(.)$, which interpolates the l-th level and expands it to the same dimensions of the upper one, can be formally expressed as follows:

$$\begin{aligned} \widehat{G}_l[x,y,i] &= \texttt{EXP}\left(G_{l+1}[x,y,i]\right) \\ &= 4\sum_{m=-R}^{+R}\sum_{n=-R}^{+R} w[m,n]G_{l+1}\left[\frac{x-m}{2},\frac{y-n}{2},i\right]. \end{aligned} \tag{4.2}$$

Laplacian pyramid levels are defined as "difference images" between two levels of the Gaussian pyramid [43]; so the l-th image is defined as

$$\begin{aligned} P_l[x,y,i] &= G_l[x,y,i] - \texttt{EXP}\left(G_{l+1}[x,y,i]\right) \\ &= G_l[x,y,i] - \widehat{G}_l[x,y,i] \end{aligned} \tag{4.3}$$

with $l = 0,\ldots,M-1$. Moreover, the last level of the pyramid is $P_M[x,y,i] = G_M[x,y,i]$, due to the non-existence of the $(M+1)$-th level in the Gaussian decomposition. Illustrative examples of Gaussian and Laplacian pyramids are shown in Fig. 4.2. Each level image is a different band-pass spatially filtered image and shows different details of the original image.

Temporal filtering

In this step, a pixel-wise processing is performed to extract pixel intensity variations of interest. In the context of our apnoea detection system, this step is used to select periodic movements in the video which are useful for breathing detection. Band-pass digital filters are applied along the time axis of the video signal. In order to

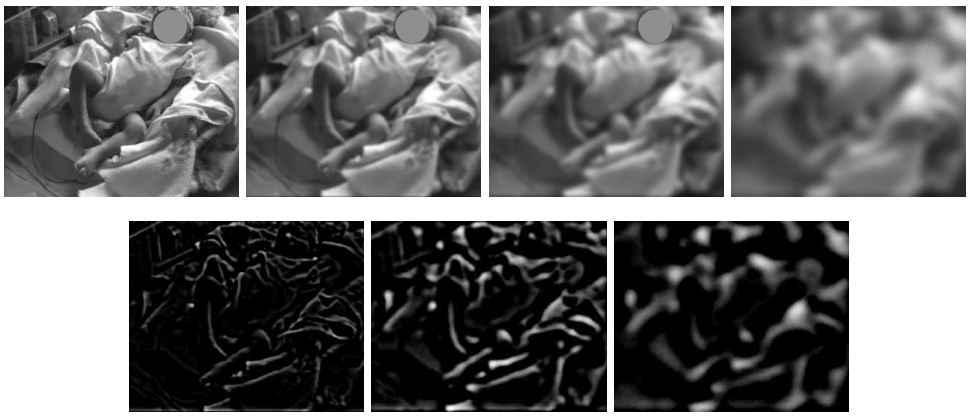


Figure 4.2: Illustrative examples of expanded Gaussian pyramid (first row) and expanded Laplacian pyramid (second row).

maintain the computational complexity low, still using sufficiently selective temporal bandwidth filters, we consider a band-pass filtering obtained by the difference of two low-pass Butterworth infinite impulse response (IIR) digital second-order filters with different cut-off frequencies $f_L^{co} < f_H^{co}$. Defining the difference equation of the IIR low-pass filter with a generic cut-off frequency f^{co} on a single level l of the pyramid at discrete time i as

$$\begin{aligned}\boldsymbol{\Upsilon}_{l,f^{co}}[i] = \frac{1}{a_{0,f^{co}}}\Big[&b_{0,f^{co}}\boldsymbol{P}_l[i] + b_{1,f^{co}}\boldsymbol{P}_l[i-1] + b_{2,f^{co}}\boldsymbol{P}_l[i-2] \\ &- a_{1,f^{co}}\boldsymbol{\Upsilon}_{l,f^{co}}[i-1] - a_{2,f^{co}}\boldsymbol{\Upsilon}_{l,f^{co}}[i-2]\Big]\end{aligned} \tag{4.4}$$

where the filter coefficients $a_{0,f^{co}}$, $a_{1,f^{co}}$, $a_{2,f^{co}}$, $b_{0,f^{co}}$, $b_{1,f^{co}}$ and $b_{2,f^{co}}$ are selected taking into account the cutoff frequency f^{co} and the frame rate of the video.

The filtered band-pass frames can be expressed as

$$\boldsymbol{\Upsilon}_{l,BPF}[i] = \boldsymbol{\Upsilon}_{l,f_H^{co}}[i] - \boldsymbol{\Upsilon}_{l,f_L^{co}}[i]. \tag{4.5}$$

The temporal filtering method outlined above is computed on every level of the Laplacian pyramid.

Variable gain amplification

After temporal filtering on spatially decomposed frames, the EVM algorithm amplifies selected movements. Amplification is obtained by multiplying the filtered pyramid levels by a proper amplification factors, in order to enhance moving parts of interest. Each l-th pyramid level must be amplified with a different coefficient α_l.Every amplification factor is function of the radial spatial wavelength: however, as a result of a first-order approximation of motion and under the assumption of smooth images and small motions in video frames, the amplification coefficients must fulfill a bound whereby they are maintained constant for higher spatial wavelengths and are linearly attenuated for lower ones [40].

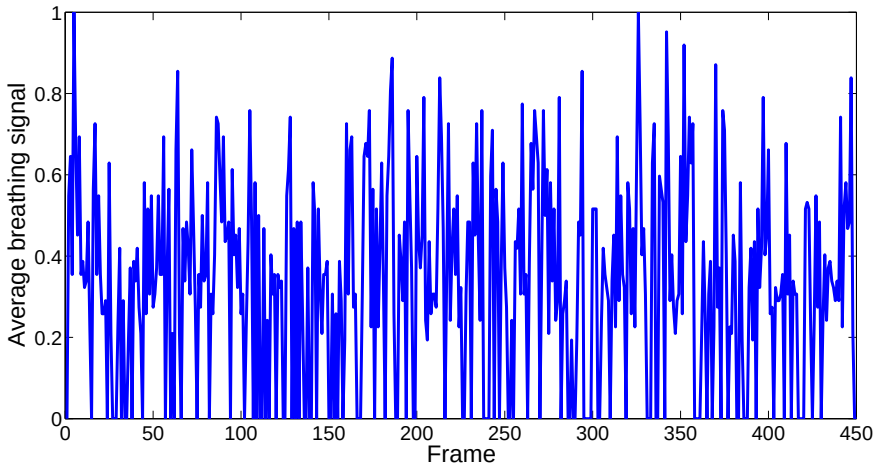
Video frame reconstruction

Amplified levels are now summed to the original levels of the pyramid (to enhance motions) and are combined to reconstruct a single video output. This video signal output can then be used to detect, through ML processing with interlaced windows, breathing (or, more appropriately, absence of breathing, i.e., apnoea). Following [43], the algorithm outputs frames are reconstructed from the processed Laplacian pyramid. Output frames are calculated with an iterative system starting from the M-th level of the pyramid and rebuilding the original level. This iterative reconstruction can be expressed as follows:

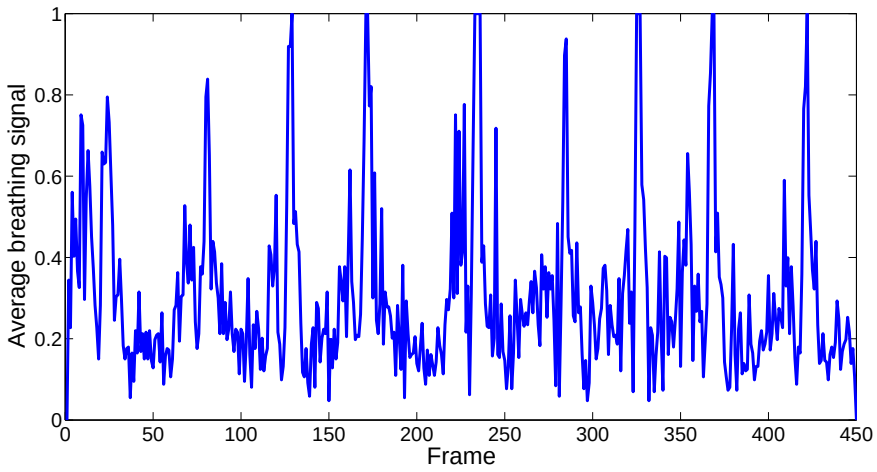
$$\widetilde{P}_l[x,y,i] = (\alpha_l \Upsilon_l[x,y,i] + P_l[x,y,i]) + \texttt{EXP}\left(\widetilde{P}_{l+1}[x,y,i]\right) \tag{4.6}$$

starting from $l = M-1$ and ending with $l = 0$.

In Fig. 4.3, we can see the algorithm at work for breath magnification. The two signals refer to an 18 second video in which a baby, although breathing, seems still at naked eye. In Fig. 4.3(a), in the extracted motion signal (without EVM magnification) small breathing movements are indistinguishable from background noise. In Fig. 4.3(b), the motion signal, based on the use of EVM in order to emphasize breathing, is shown: the periodicity of the small breathing movements is clearly visible. Considering that the rate is 25 frames per seconds a frequency of about 0.5 breathing acts per second can be estimated.



(a)



(b)

Figure 4.3: Examples of extracted breathing signals: (a) without pre-magnification and (b) with pre-magnification.

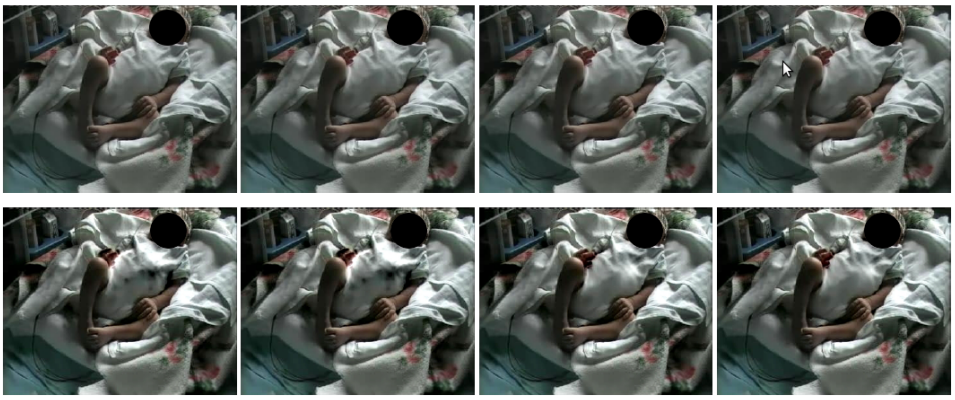


Figure 4.4: A complete respiratory cycle: (first row) original sequence (second row) processed sequence.

(a) (b)

Figure 4.5: Emphasizing the respiratory movements: (a) original image (b) image with enhanced breathing motion.

In Fig. 4.4, an illustrative example of the image amplification, for breathing detection, brought by the use of EVM, is shown. In Fig. 4.5, a detail of the processing, is shown.

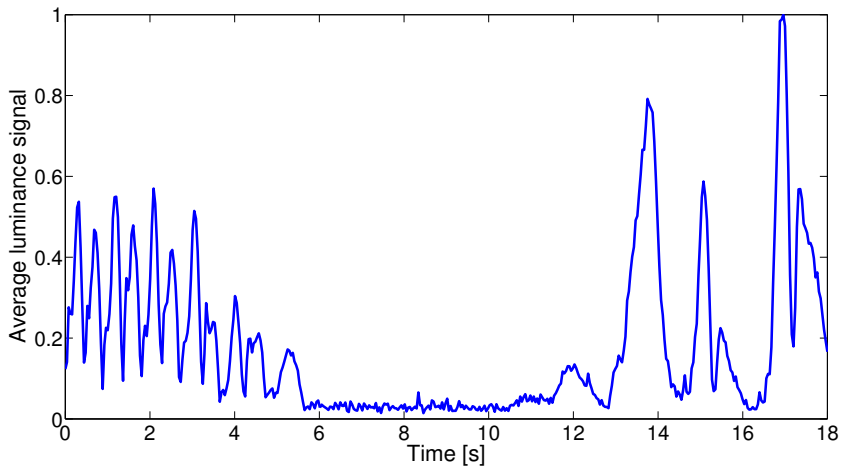
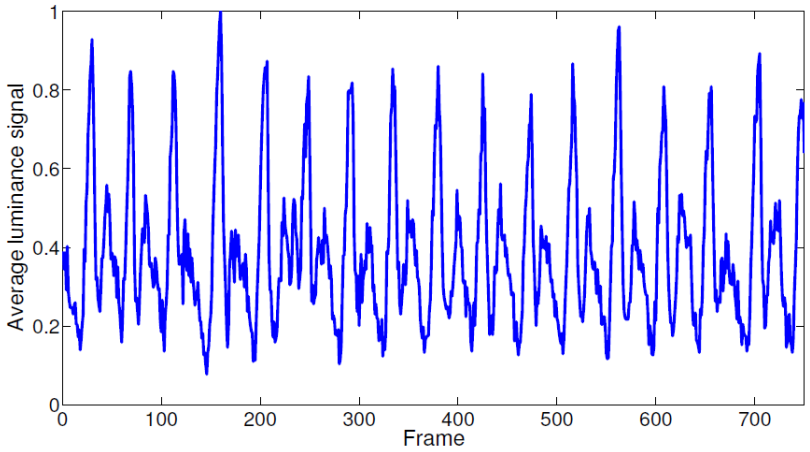


Figure 4.6: An illustrative example of extracted average luminance signal of a breathing newborn experiencing an abnormal event.

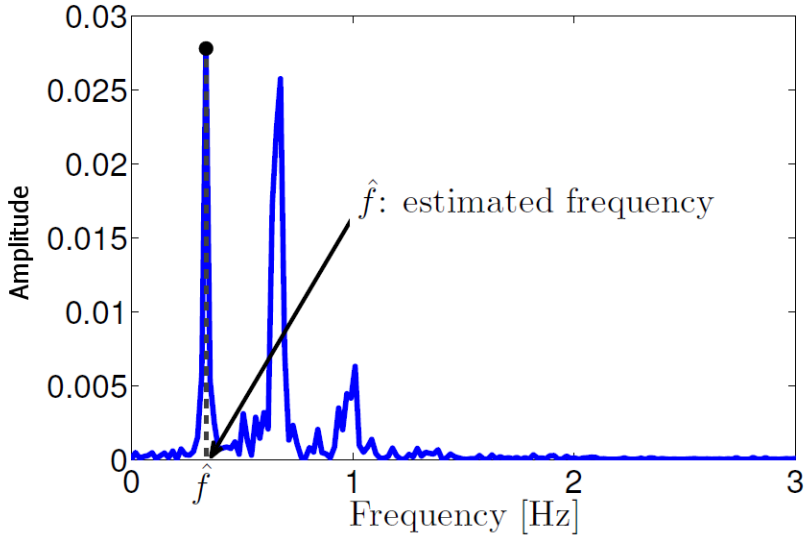
4.1.1 Apnoea Detection

Clinically, an apnoea event is defined as an episode of absence of breathing lasting at least 20 s.[1] Once a chest respiratory movement has been emphasized by the EVM according to [40], we can proceed to further analyse the extracted motion signal using the method described in Chapter 3, with the aim of identifying the time intervals in which breathing cannot be detected by conventional methods. In Fig. 4.6, we show the motion signals extracted from a video sequence of an abnormal event. As it can be observed, the proposed method highlights clearly the time period in which breathing movements are absent. In Fig. 4.7(a) an example of average luminance signal extracted from a video, sampled at 15 frames per second, of a sleeping cat is shown; a periodic component in the signal is clearly visible. In Fig. 4.7(b) the corresponding likelihood function for f and the estimated frequency $\hat{f}$ in correspondence of the peak are shown.

[1] The medical literature defines as apnoeas also the episodes of absence of breathing lasting between 10 s and 20 s, provided they are associated with other clinical signs/symptoms [44].



(a)



(b)

Figure 4.7: (a) Example of a breathing signal and (b) corresponding likelihood function with frequency estimation.

4.1.2 Performance Analysis

We now present results on a newborn with timely-diagnosed and genetically confirmed Congenital Central Hypoventilation Syndrome (CCHS). As ground truth, we consider the results obtained by prolonged polysomnographic monitoring, lasting a total of 1 hour and 35 s. The polysomnographic-based results indicate: 23 apnoea episodes with a total duration of 28 min and 31 s, and an average single apnoea duration of 74 s.[2]

In analysing the performance of our automatic diagnostic system, we focused on 17 of the 23 apnoea events detected by the polysomnograhic system, because: in 4 apnoea episodes, the view of the newborn was obstructed by medical staff; in another 2 episodes, the apnoeas were short (less than 20 s), whereas our system had been calibrated to detect apnoeas of at least 20 s. So we focused on 22 min and 20 s of the total 28 min and 31 s of apnoeas detected by the polysomnographic monitoring.

Since apnoea episodes are potentially deadly, it is important to promptly detect them. The goal is to identify apnoea events immediately upon breathing interruption. In fact, if an apnoea episode is not timely detected and the corresponding alarm given too late, it may be useless.

The performance of the proposed system is investigated in terms of sensitivity and specificity, defined, respectively, as follows:

$$SE = \frac{T_{TP}}{T_{TP} + T_{FN}} \tag{4.7}$$

$$SP = \frac{T_{TN}}{T_{TN} + T_{FP}} \tag{4.8}$$

where T_{TP}, T_{TN}, T_{FP}, and T_{FN} denote, respectively; the total length of the time intervals with apnoea correctly detected (Time True Positives); the total length of the time intervals with no apnoea correctly detected (Time True Negatives); the total length of the time intervals with no apnoea incorrectly reported as apnoea by the system (Time

[2]Due to the documented dysregulation of respiratory control, during the recording, the patient was also monitored with an elastic belt sensor, a nasal flow-meter and a pulse oximeter. Furthermore, oxygen therapy was applied using nasal cannulas.

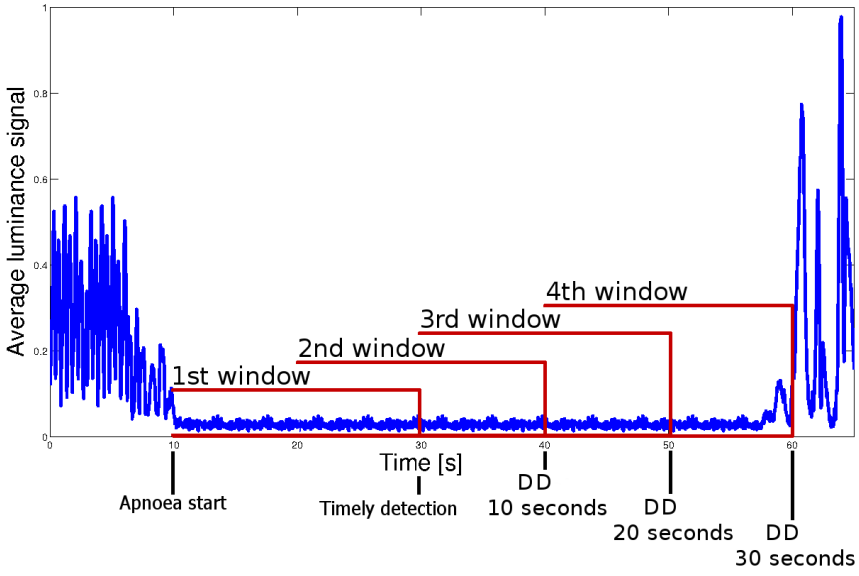


Figure 4.8: Interlaced observation windows: length 20 seconds, overlap 50%.

False Positives); and the total length of the time intervals with apnoea incorrectly reported as normal breathing by the system (Time False Negatives).

We define, moreover, Detection Delay (DD) as the amount of seconds elapsed between the end of the first 20 seconds of an apnoea episode and the reporting of the detection by our system. Fig. 4.8 helps to better understand this concept. The apnoea starts at the second 10, and, since an apnea is defined as such if it lasts at least 20 seconds, it should be promptly detected at the second 30. If this does not happen, there will be a detection delay (DD), which might be 10, 20 or 30 seconds according to the window in which the detection occurs. We recall that the analysis is done on half-interlaced windows, in order to improve the reliability of the proposed method. In fact, an apnoea event could manifest itself across two consecutive disjoint windows: a single window-based algorithm could thus miss it in each of the two windows. In Table 4.1, the sensitivity and specificity values obtained with the proposed method are shown, considering various values of the detection delay in which an apnoea episode

Table 4.1: Apnoeas Detection (One RGB camera)

DD	*DA*	T_{TP}	T_{TN}	T_{FP}	T_{FN}	*SE*	*SP*
0	13	1200	1800	500	140	90%	78%
10	14	1230	1800	500	110	91%	78%
20	16	1280	1880	420	60	95%	81%
30	17	1340	1920	380	0	100%	83%

Legend: *DD*= Detection Delay (seconds); *DA*= number of Detected Apnoeas; T_{TP}, T_{TN}, T_{FP}, T_{FN} (seconds).

is detected. In particular, sensitivity and specificity are computed by considering the portions of minutes of apnoea, with respect to the total apnoea duration equal to 22 minutes and 20 seconds correctly/incorrectly diagnosed. With the proposed system we were able to detect from 13 to 17 episodes, out of a total of 17, depending on the acceptable delay in reporting an episode. With this analysis we wanted to know how much delay we should accept to get as close as possible to 100% of apnoea episodes detected.

4.2 Spatio-Temporal Video Processing

Monitoring of basic vital signs is essential to evaluate the evolution or the onset of several diseases. The respiratory rate (RR), in particular, is a vital parameter able to provide very important information on the health condition of a patient. Anomalous RR values can be a symptom of various medical conditions, including respiratory diseases such as the congenital central hypoventilation syndrome (CCHS) [5], the Cheyne-Stokes respiration [45], the Biot's respiration [46] and the Kussmaul breathing [46]. However, they can also be a symptom of a deterioration of clinical conditions: a sudden change in respiratory rate may be an important predictive parameter of a risk of mortality.

In this chapter we will show how to merge the techniques shown in Chapter (3) and Section (4.1) to improve the respiratory rate estimation [47].

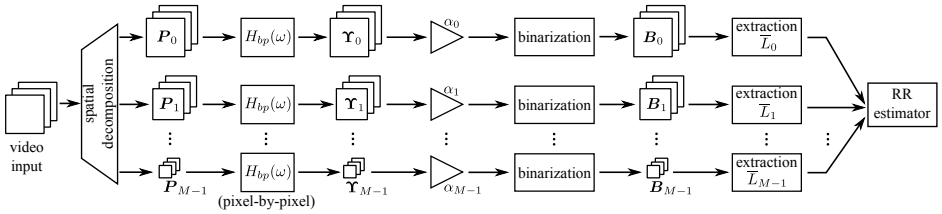


Figure 4.9: Block diagram of the spatio-temporal RR estimation technique.

4.2.1 Respiratory Rate Estimation

The main idea is to apply the motion signal extraction shown in Chapter (3), to the Laplacian pyramid levels obtained by the method discussed in Section (4.1), i.e., consider the L levels of the Laplacian pyramid in the Eq. (4.3), as the S sensors of Eq. (3.6).

In Fig. 4.9, a block diagram of the proposed system is shown.

Motion signal extraction in (3.1) detects scene variations by the use of a well known algorithm based on a simple differential filtering method, called difference of frames (DoF): outputs are computed by the difference of two consecutive frames. This method corresponds to a simple finite impulse response (FIR) filter. In Fig. 4.10, the frequency response of the DoF digital filter is shown, considering a sampling frequency of $f_s = 25$ Hz; it is clear that the DoF filter is a high-pass filter able to extract fast variations. As clearly visible, the DoF filter is not appropriate to extract breathing motions, particularly if they are not amplified. In fact, newborn respiration acts have typical periodicity in the frequency range between 0.25 Hz and 1 Hz. Since the magnitude frequency response of the DoF filter in the range between 0.25 Hz and 1 Hz is always below -10 dB, variations of interest are deeply attenuated.

To solve this problem and to integrate the motion signal extraction with the spatio-temporal processing, the proposed system uses a band-pass filter instead of the DoF filter to extract motions. To keep low the computational complexity and reduce the delay introduced by filtering, still maintaining sufficient selectivity, a second-order digital band-pass filter with infinite impulse response (IIR) is used. This digital band-pass filter, with frequency response $H_{bp}(\omega)$, is applied to all pyramid levels,

as shown in Fig. 4.9. Accordingly, the system processes video frames by a Butterworth IIR filter, properly designed by setting up cut-off frequencies f_L^{co} and f_H^{co}, with $f_L^{co} < f_H^{co}$. Butterworth filters are characterized by the property that the magnitude response is maximally flat in the pass-band [48] and 3 dB below its maximum value at the cut-off frequencies f_L^{co} and f_H^{co}. The cut-off frequencies of the temporal filter, applied to all levels of the pyramid, are properly selected to extract periodic variations related to breathing movements. Thus, given the order of the filter, the cut-off frequencies and the sampling frequency f_s, the coefficients of the digital filter are readily obtained [48, 49].

This pixel-wise digital filter has transfer function of type:

$$H_{bp}(z) = K\frac{\left(1+z^{-1}\right)\left(1-z^{-1}\right)}{\left(1-pz^{-1}\right)\left(1-p^{*}z^{-1}\right)} \tag{4.9}$$

where K is a scale factor, p and p^* are the complex conjugate poles, properly determined by the filter design rules to fit requirements for f_L^{co} and f_H^{co} [48, 49]. The filter used for breathing analysis is set up with $f_L^{co} = 0.2$ Hz and $f_H^{co} = 1.05$ Hz, so that normal periodicity movements are included. In Fig. 4.10, the frequency response of the used IIR digital filter is also shown. As in the EVM algorithm, once every level of the Laplacian pyramid has been filtered, it is possible to amplify selected levels. This procedure is obtained multiplying each filtered pyramidal level $\boldsymbol{\Upsilon}_\ell$, shown as filter output frames in Fig. 4.9, by a proper amplification factor to enhance only periodic movements of interest. A further improvement can be obtained optimizing this amplification coefficients. In order to avoid noise amplification or motion artifacts, each level of the multi-scale decomposition is amplified with different coefficients $\{\alpha_\ell\}_{\ell=0}^{M-1}$. Indeed, it could be shown that amplification coefficients are strictly related with the spatial frequencies, due to the Eulerian motion model and the spatio-temporal relationship shown in [40]. Inspired by the amplification method described in [40], the coefficients $\{\alpha_\ell\}_{\ell=0}^{M-1}$ are set as follows: much larger than 1 for higher pyramidal levels (lower spatial frequency bands) and linearly attenuated for lower levels (higher spatial frequency bands) down to $\alpha_0 = 1$. The last filtered pyramidal level $\boldsymbol{\Upsilon}_{M-1}$ is not used because of too low resolution and, correspondingly, gain α_{M-1} is set to zero.

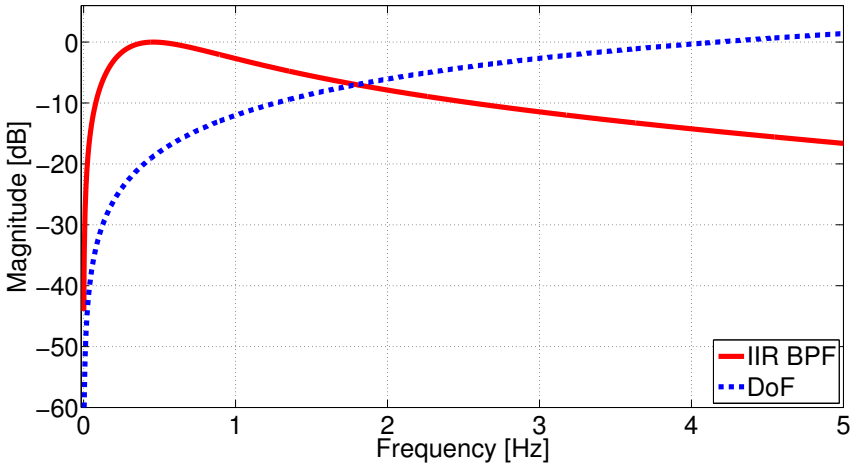
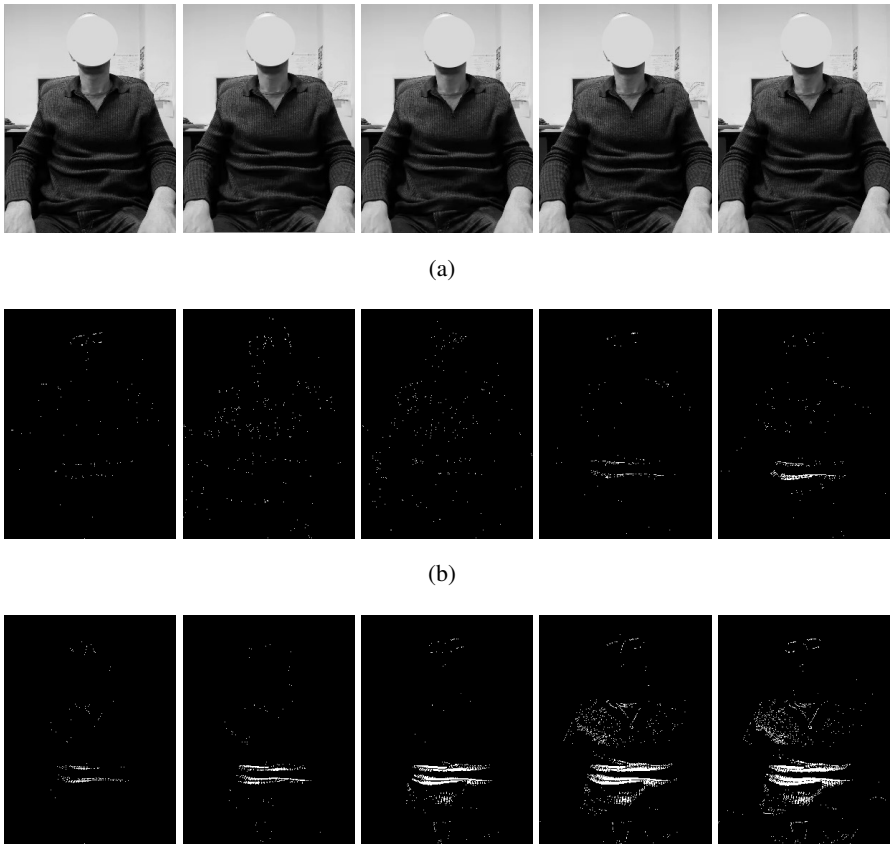


Figure 4.10: Magnitude frequency response of the difference of frames (DoF) FIR filter and the of the second-order band-pass IIR filter used for breathing motion extraction.

In Fig. 4.11, an illustrative comparison between binarized DoF and optimized IIR filtering, without spatial decomposition and amplification, is shown considering a few illustrative frames. More precisely, the frames are extracted from a video, taken with a smartphone, focusing on a seated breathing adult; it can be noticed that frames shown in Fig. 4.11(c) reveal better breathing movements than the ones shown in Fig. 4.11(b).

In Fig. 4.12, illustrative examples of luminance signals extracted with (a) DoF filter and (b) proposed IIR filter, are shown. In both cases, the two signals are extracted from a video framing a sleeping newborn: it can be immediately noticed that in Fig. 4.12(b) the signal pattern is smoother, noise is reduced and periodicity is enhanced.



(a)

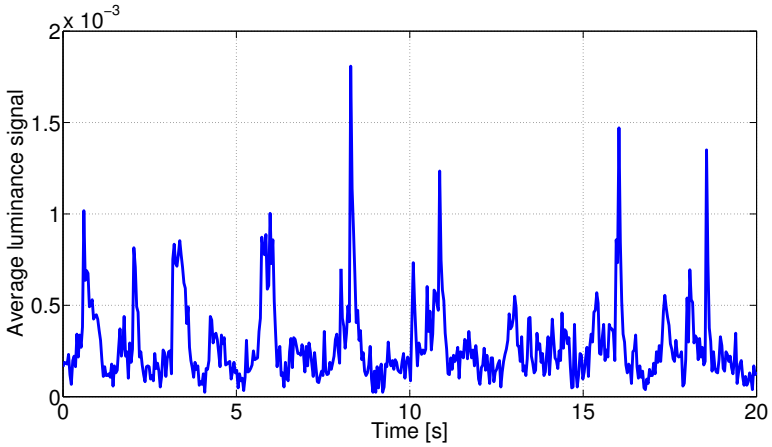
(b)

(c)

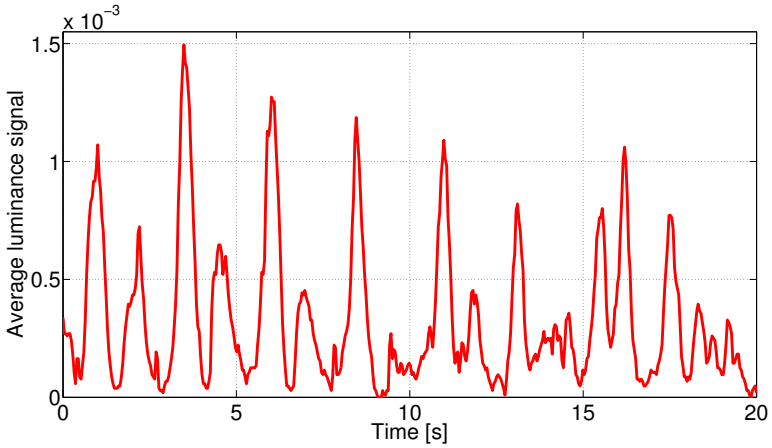
Figure 4.11: An illustrative example of the image processing with different temporal filters: (a) original video frames, (b) DoF filtered and binarized frames and (c) optimized IIR filtered and binarized frames.

4.2.2 Performance Analysis

The system introduced in Section 4.2.1 is compared with a gold-standard polysomnographic system. In particular, the signal extracted by the pneumogram, an elastic belt that measures the change in volume of the thoracic cavity, is used as a reference. For performance analysis, we have used two videos of a newborn, breathing



(a)



(b)

Figure 4.12: An illustrative example of extracted average luminance signal with (a) DoF filter and (b) with the breath-optimized IIR filter.

normally (i.e., without any particular respiratory disease). The duration of the videos are, respectively: 5 minutes and 11 seconds, and 3 minutes and 3 seconds, for a total duration of 8 minutes and 14 seconds. The analysis was carried out using overall 47 interlaced windows, each lasting 20 seconds, with 50% overlap between consecutive

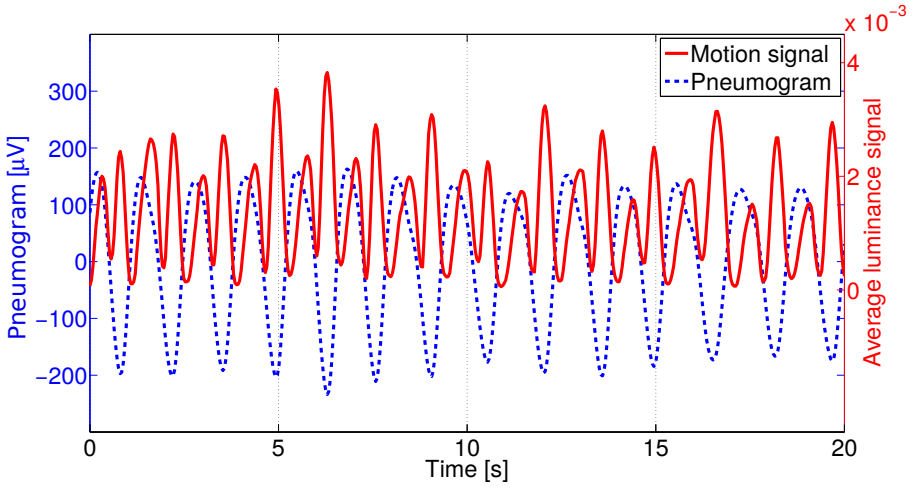


Figure 4.13: Comparison between the motion signal extracted by the proposed technique for the second level ($\ell = 1$) of the processed pyramid and the signal extracted from the pneumogram.

windows. The system was set up with the following parameters:

- video resolution: 360×288 pixels;
- $M = 6$ levels of the Laplacian pyramid, numbered as $\ell = 0, \ldots, 5$;
- $N/f_s = 20$ seconds of windows length, with $f_s = 25$ Hz;
- cut-off frequencies of the band-pass IIR filter: $f_L^{co} = 0.25$ Hz and $f_H^{co} = 1.05$ Hz;

In Fig. 4.13, a portion, lasting 20 seconds, of the signal extracted from the pneumogram is compared with the signal extracted from the second ($\ell = 1$) level by our system. Every period of the quasi-periodic signal from the pneumogram corresponds to a complete respiratory act of the patient. The average luminance signal extracted by the video-processing system shows a good correspondence with the one given by the pneumogram, considering that each respiratory act is composed by two main movements: inhalation and exhalation.

In Fig. 4.14, the average luminance signals extracted from filtered and amplified pyramidal levels of a video decomposed with $M = 6$ are shown. As it can be noticed, even though the signals extracted from different levels are similar but not equal, they have all the same common frequency. The last level of the pyramid, the one corresponding to $\ell = 5$, has not been shown, because, with the resulting low detail level, the signal was almost absent. The breathing signal predominates in the third level ($\ell = 2$), corresponding to the motion signal $\overline{L}_2$.

In Fig. 4.15, the values of frequencies estimated by our system are compared with those indicated by the pneumogram. In both the considered videos, the first window was discarded because of the IIR filter initial transient period, that invalidates the result. According to medical practice, we have considered $\pm 15\%$ as an acceptable tolerance in the estimation of the RR. As shown in Fig. 4.15, the RR is correctly estimated in 38 out of 45 windows in two video samples framing the newborn laterally. In Fig. 4.15(a), the RR is correctly estimated in 24 out of 29 windows, but without taking into account a movement of the newborn in the middle of the video shown in Fig. 4.15(a) between the 13-th and the 14-th windows. In Fig. 4.15(b), the RR is correctly estimated in 14 out of 16 windows in the second video sample.

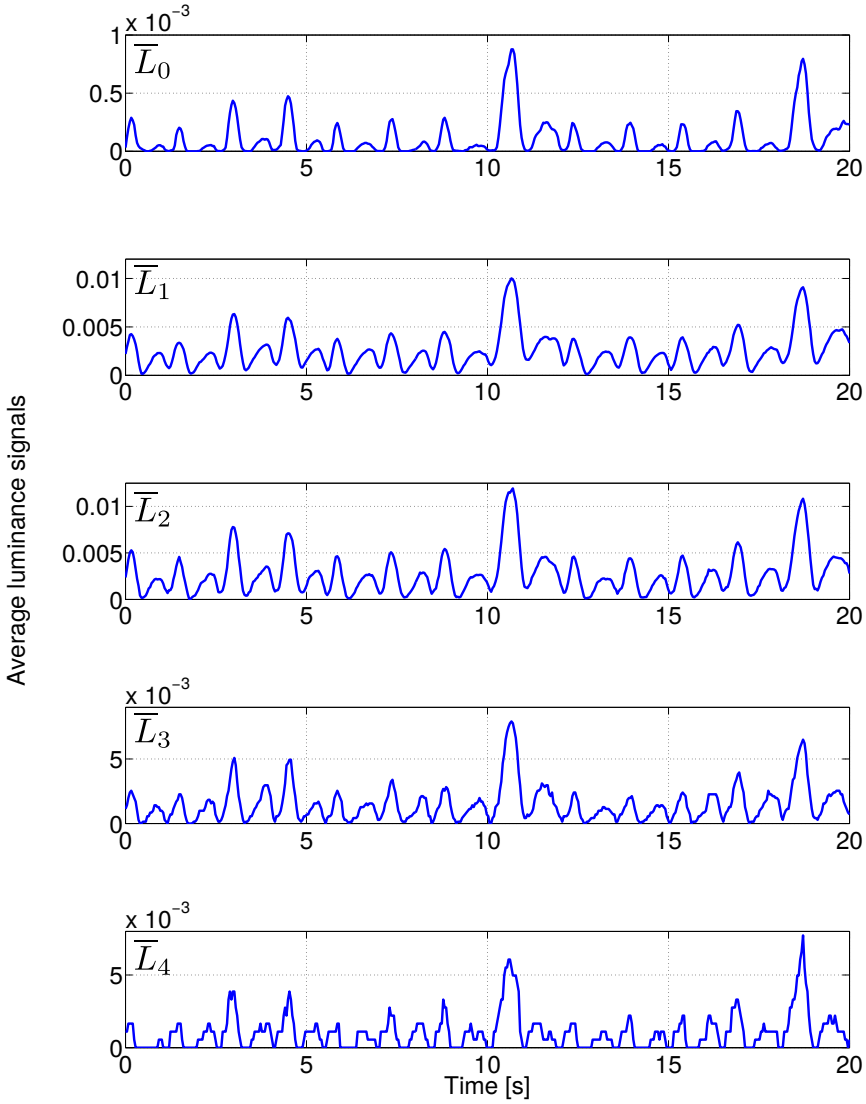
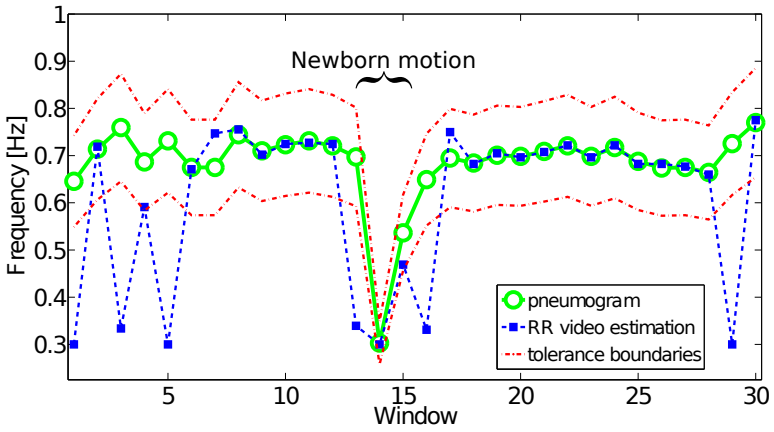
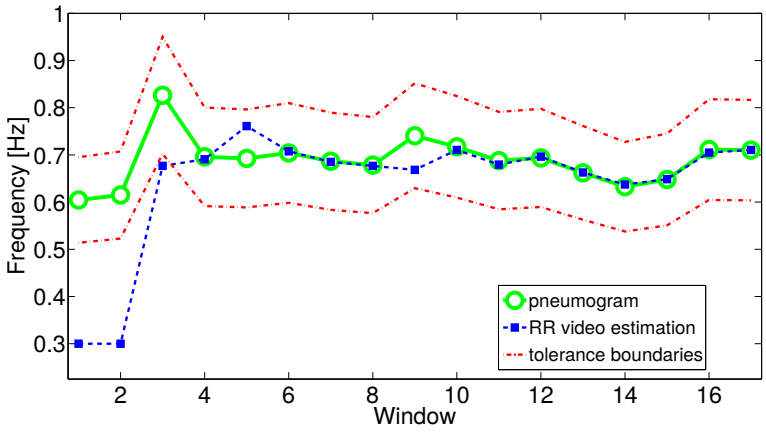


Figure 4.14: Average normalized luminance signals, equation (3.3), extracted from the filtered and amplified levels of the pyramid, from the lower one to the higher one.



(a)



(b)

Figure 4.15: Comparison between frequencies indicated by the pneumogram and frequencies estimated by our system: (a) from a video sample lasting 5 minutes and 12 seconds, (b) from a video sample lasting 3 minutes and 4 seconds.

Chapter 5

Analysis of Neonatal Seizure Automatism

"Nothing happens until something moves."

– Albert Einstein

5.1 Introduction

Seizures in the neonatal period do not always show a clear correlation between electroencephalographic findings and clinical manifestations. In particular there is an unresolved debate on the real critical nature of some episodes of stereotyped rhythmic movements, currently included in the so-called 'subtle seizures' [1], often not associated with a clear electroencephalographic correlate. We have analysed the case of a newborn with prolonged, repeated episodes of repetitive movements of the limbs; in these episodes the seizure onset leads to a modification of the kinematic characteristics of motor pattern.

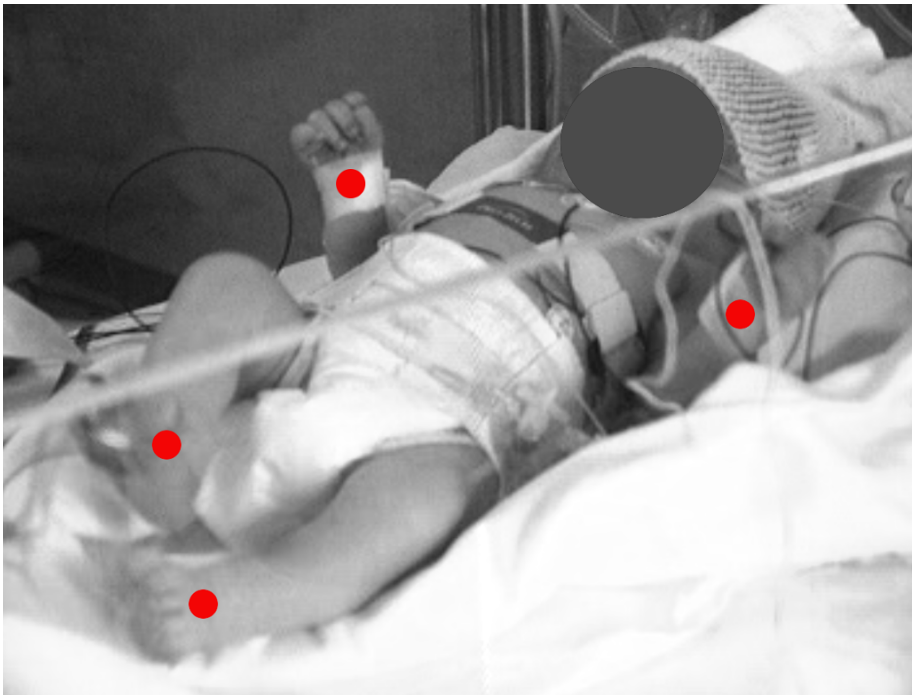


Figure 5.1: Markers on the distal (wrists and ankles) joints of the four limb.

5.2 Analysis

In order to quantitatively assess the kinematic parameters possibly influenced by seizure activity, two time intervals were identified: the ictal and the pre-ictal periods. The ictal period covered the time interval between the seizure onset and its offset, based on agreement between the judgment of three electroencephalography experts. The entire digital video sequence of total duration of 90 s, sampled at 25 frames/s, encompassing pre-ictal (45 s) and ictal (45 s) periods, was used to perform the kinematic study of the patient's limb movements. Using a properly developed software written in Cpp [50] with the OpenCV [51] library, a set of markers was inserted in each frame over specific anatomical landmarks. Markers on the distal (wrists and ankles) joints of the four limbs were manually repositioned on each frame, see Fig. 5.1.

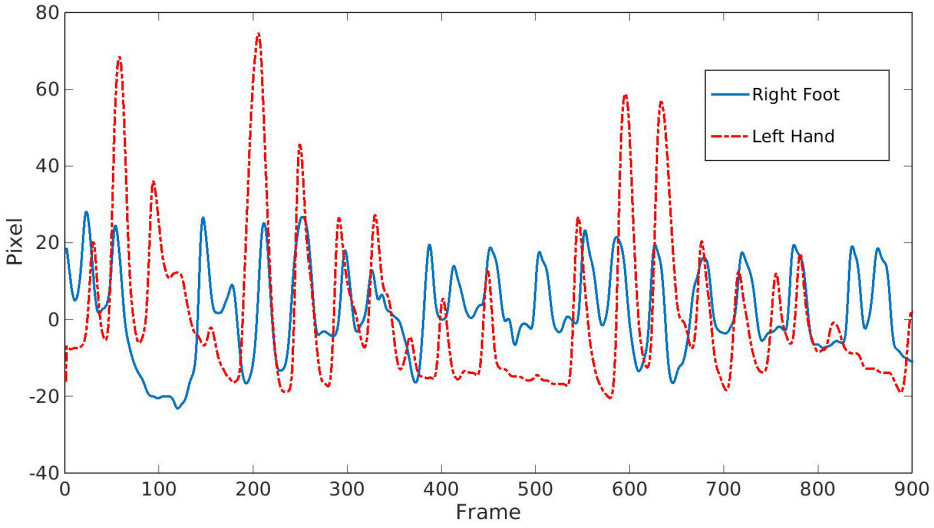
The planar coordinates of each marker for each timeframe were employed for further analysis. The position of the RA (Right Arm), LA (Left Arm), RL (Right Leg) and LL (Left Leg) were plotted as a function of frame number (the sequence of images is acquired at 25 frames per second, therefore 1 second equals 25 frames). In Fig. 5.3, the x-position of the right foot and the left arm are shown, while in Fig. **??**, the y-position of the right foot and the left arm are shown.

In order to investigate the existence of a cyclic pattern for each limb, on each single trace we identified the peaks of the signal as the time sample corresponding to the sign inversion of the first derivative (positive-to-negative, so as to identify the maxima of the original signal). To exclude the peaks due to noise and the low resolution of the labelling step, all signal were offline filtered with a moving average filter (time window of 5 samples centred in time, no introduced delay) and an amplitude threshold was set so as to include only clear signal peaks. The cycle duration was computed for each limb as the time difference between two consecutive peaks. Then, we calculated the inter-limb cross-correlation, between the right foot and the left arm, both in pre-ictal and post-ictal period. In this procedure, two traces, one concerning an arm and the other concerning a leg, are shifted with respect to each other of a variable number of time samples (ranging from -20 to +20, i.e. from -800 ms to 800 ms), and for each time shift the correlation value between the two curves is given. This makes it possible to measure the similarity in motor patterns of the two limbs. The same analysis on the other pair of limbs was not possible because during the ictal period an arrest of the left leg movement is observed.

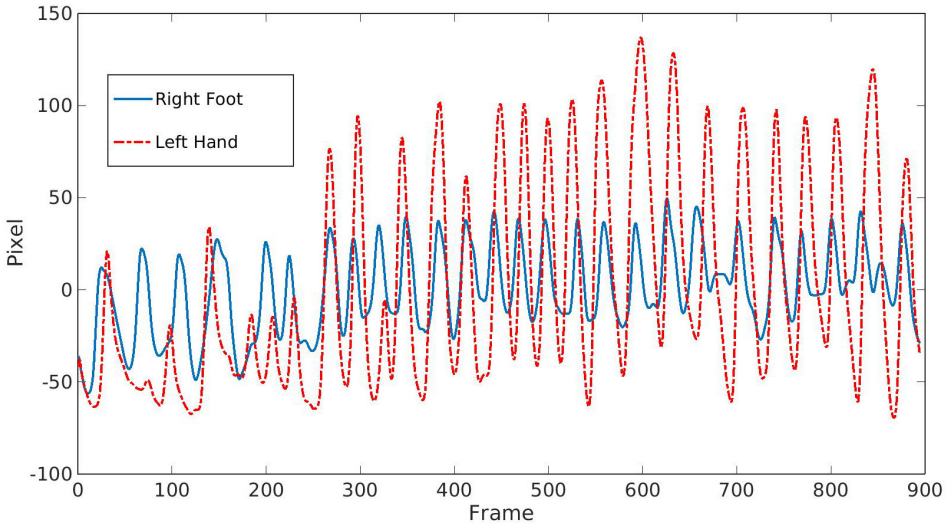
5.3 Results

The video-EEG recording, performed after the onset of repetitive abnormal limb movements, showed a severely disrupted limb movements; the occurrence of a seizure was associated with a modification of the clinical motor pattern.

During the entire video segment analysed, the child presents alternating movements of "cycling" and "boxing". During the ictal epileptic discharge an arrest of the left leg movement is observed, while the movements of the arms and of the right leg

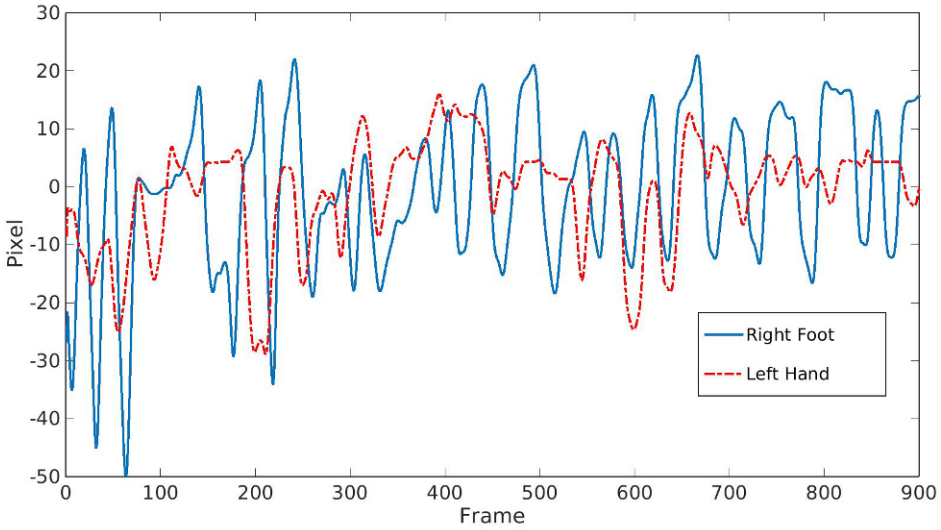


(a) Pre-Ictal

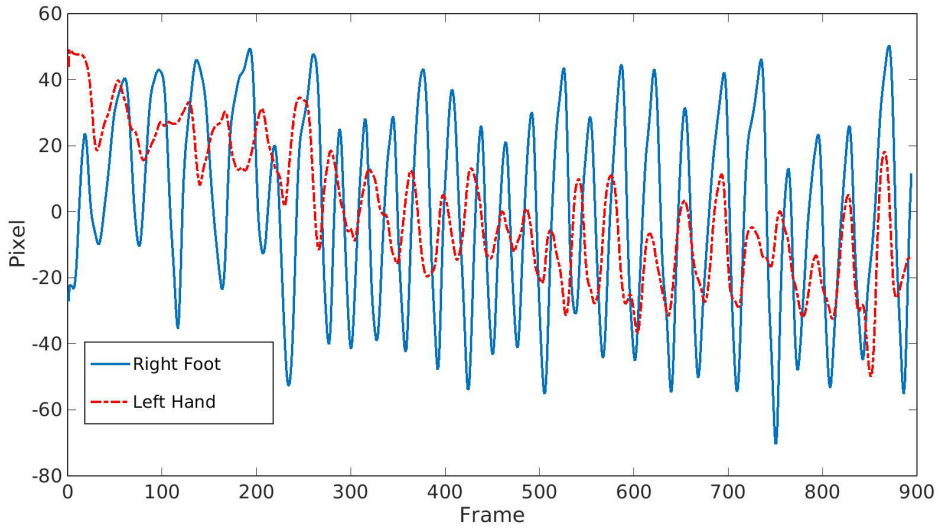


(b) Post-Ictal

Figure 5.2: X-position of the right foot and the left arm: (a) pre-Ictal period and (b) post-Ictal period.



(a) Pre-Ictal



(b) Post-Ictal

Figure 5.3: Y-position of the right foot and the left arm: (a) pre-Ictal period and (b) post-Ictal period.

Table 5.1: Kinematic Parameters

	Pre Ictal	Post Ictal
Correlation	0.26	0.59
Cycle Duration	1.54 s	1.18 s
Standard Deviation of Cycle Duration	12.7 frames	7.8 frames
Cycle Mean Amplitude	80 pixels	156 pixels

show an increased speed (the average cycle duration switches from 1.54 s to 1.18 s), a higher amplitude (with a mean amplitude increase of nearly 100 percent), and a greater inter-limb coordination. Furthermore, a decrease of about 40 percent of the standard deviation of the cycle duration shows a definitely more regular cyclic pattern. As soon as the rhythmic EEG pattern disappears, the movement stops abruptly. Due to the absence of Left Leg movements in the ictal period, comparison of the kinematic relationships between upper and lower limb was possible only for the couple Left Arm-Right Leg. Interestingly, this analysis revealed that, while the correlation between the pre-ictal signals was quite low ($Correlation = 0.26$), during the seizure, the two signals result strongly correlated ($Correlation = 0.59$). It is important to remember that a correlation equal to 1 means signals perfectly identical. This increase in correlation, therefore, proves the relationship between the seizure onset and the improved coordination among the upper and lower limbs. The obtained results are summarized in Table 5.1.

Chapter 6

Simulators

"Health care can be made more affordable for the poor,
without requiring major new scientific developments,
just the smart application of current technologies."

– Muhammad Yunus

6.1 Introduction

The medical healthcare system has seen, over the last decade, a significant increase of simulation techniques based on the use of dummies with human appearance (partial or total). Simulation allows to create, in a protected environment, certain conditions of crisis, giving to the operators the ability to deal with the possible situations of tension or error, without risks for the patients. Furthermore, the use of diagnostic and therapeutic techniques increasingly sophisticated, has forced the professionals to verify each equipment on simulators, in order to avoid harmful effects for the patients. Within this context, we have decided to realize 3 simulators able to replicate the symptomatic movements characteristic of the diseases under consideration: neonatal seizures and breathing disorders. The reasons that led us to the realization of the 3 simulators is, essentially, the opportunity to have, at any time, a 'subject' on

which to test the continuously evolving detection algorithms. Reproduction in a virtual environment makes it possible to create countless combination of the evaluated parameters and to optimize the conditions for recording in terms of brightness and absence of reflexes.

6.2 State of the Art in Medical Simulation

Retracing the history of the simulation in medicine, one may notice that this technique has spread rather late: in fact, the literature begins to treat simulation only since the 60s, limited to some university centres; only around the end of the 80s this practice spreads in hospitals and training centres [52]. A wide dissemination of simulation techniques in medicine occurs, however, only since 2000, when, thanks to technological developments and cost reductions, the simulation becomes a practice legitimized and institutionalized. National and international societies, which promote study and simulation in medicine, are born in this period, such as the Society for Simulation in Healthcare [53], the Society in Europe for Simulation applied to Medicine [54], the UK Association for Simulated Practice in Healthcare [55]. In addition, various international journals were founded, such as “Simulation in Healthcare” [56] or “Clinical Simulation in Nursing” [57]. The first human simulator appears in the 60s, it is "Resusci Anne" [59, 58] (Fig. 6.1) and marks the beginning of the era of simulation in healthcare. This early simulator of a dying victim, not breathing and lacking a heart beat, gave to the medical staff the possibility of training for the Airway Breathing Circulation (ABC) of CardioPulmonary Resuscitation (CPR). Since then, an increasingly advanced technology has made possible the creation of highly realistic simulators able to reproduce a wide variety of clinical events.

Currently, among the world’s leading suppliers of simulators for the healthcare sector are the Laerdal Medical group (Stavanger, Norway) [58], which is specializes in training simulators, and the Gaumard Scientific group (Miami, Florida, USA) [60], which produces simulators for university hospitals and nursing schools, for emergency medical services and for medical training in the military field.

The Laerdal company has designed several devices, including the SimNewB [58]

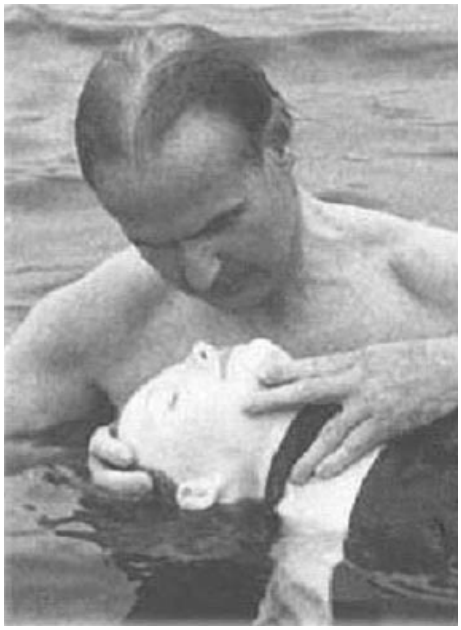


Figure 6.1: Asmund Laerdal with Resusci-Anne, in 1960 [58].

(Fig. 6.2), a dummy of a newborn, with realistic traits, measuring 52.5 cm and weighs 3.17 kg, it simulates the respiratory activity with variable frequency up to 100 cycles per minute; and the SimBaby [58] (Fig. 6.3) that allows live defibrillation and is able to simulate a seizure.

Among the product of the Gaumard company, we cite the model Newborn HAL S3010 [60] (Fig. 6.4), that simulates an infant born at 40th week of gestation, and the model S108 [60] (Fig. 6.5), that simulates an infant born at 28th week of gestation. The cost of all the simulators listed above, depends on the type and on the amount of accessories included in the kit, for more complete models the cost oscillates between $20,000 and $30,000 [61].

Figure 6.2: Laerdal SimNewB Simulator [58].

Figure 6.3: Laerdal Simbaby Simulator [58].

Figure 6.4: Gaumard Newborn HAL S3010 Simulator [60].

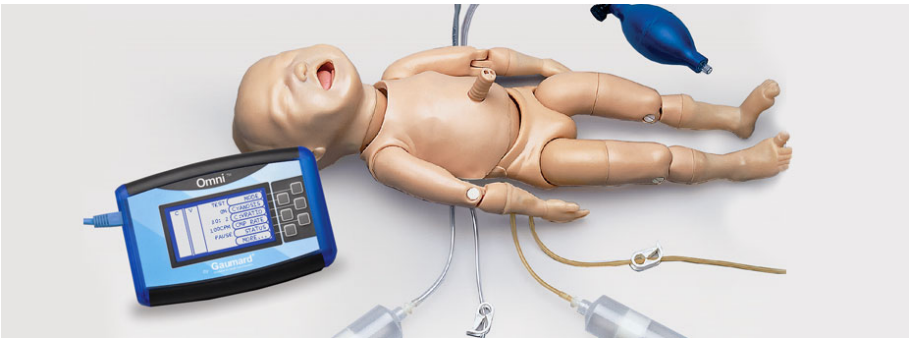


Figure 6.5: Gaumard S108 Simulator [60].

6.3 Hardware

The simulators we have realized, consist of toy dolls in which an electric motor is inserted, controlled by an Arduino board and properly programmed to reproduce, with the dolls, the typical movements of the pathology to be simulated.

6.3.1 Arduino

Arduino [62] is an open-source prototyping platform based on easy-to-use hardware and software. It was born at the Ivrea Interaction Design Institute as an easy tool for fast prototyping, aimed at students without a background in electronics and program-

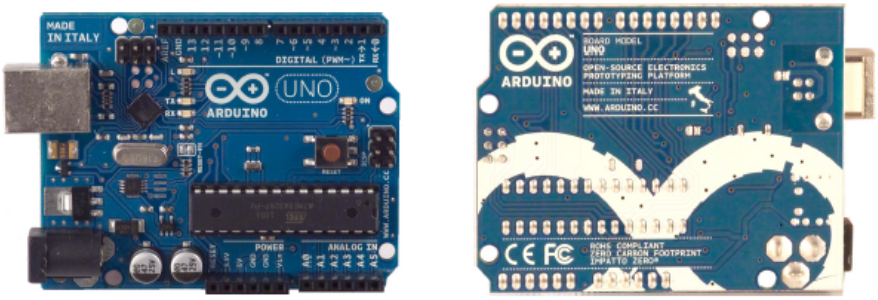


Figure 6.6: Arduino UNO Board [65].

ming. Arduino boards are able to read inputs from various kinds of sensors, and turn it into an output, for example turning on a Light Emitting Diode (LED) or activating a motor. This is possible by sending instructions to a microcontroller installed on the board. To do so one needs to use the Arduino programming language, based on Wiring [63], and the Arduino Software, an Integrated Development Environment (IDE) based on Processing [64]. In Fig. 6.6 an Arduino UNO board [65], is shown.

Arduino programs are written in a language derived from C, they are called 'sketch' [66] and they are loaded and run on the board. A typical Arduino C/C++ sketch consists of two functions:

- setup(): a function that runs once at the start of a program and that can initialize settings
- loop(): a function called repeatedly until the board powers off.

In Fig. 6.7, the typical block diagram of an Arduino's sketch [66], is shown.

6.3.2 Adafruit Motor Shield

The Adafruit Motor Shield is a driver module for motors that allows you easily control motor direction and speed using an Arduino. Based on the Chip L29398, it is able to drive two direct current (DC) motors or two servo motors. This module can

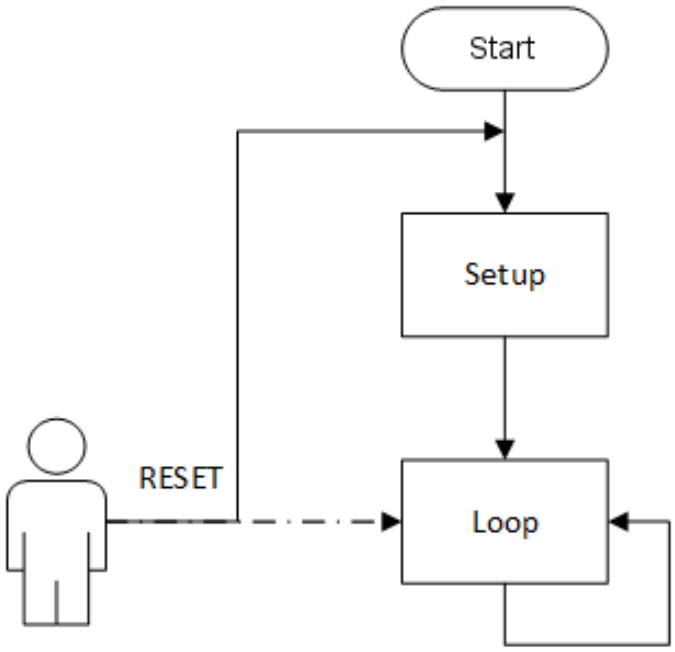
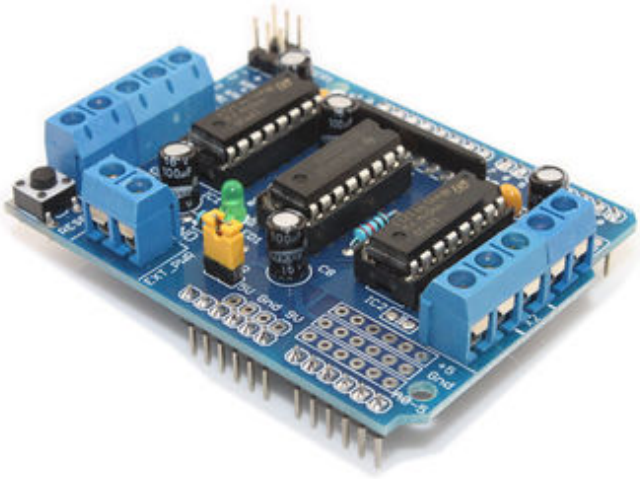
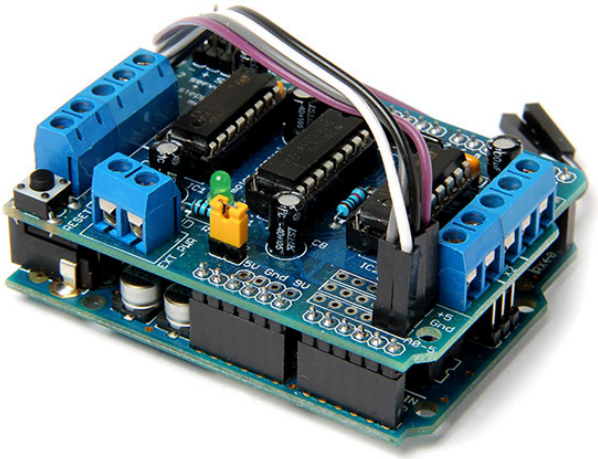


Figure 6.7: Flowchart of an Arduino sketch [66].

be used for the development of micro robots. The Motor Shield can either be powered directly by Arduino or by an external 5V-15V power supply via the DC jack, indeed, motors require significant amount of power, and can cause 'brownouts' that reset the Arduino. For that reason the shield is designed for seperate (split) supplies, one for the electronics and one for the motor, because the current required from the motor often exceeds the maximum USB current rating. The Arduino chip (an Atmel Atmega 328P microcontroller [67]) can only drive 40 mA per pin. With 5.5V supply, this means a single pin can output 0.22W. One should be able to power a micromotor, for instance like those that make a mobile phone vibrate. However, most usable hobby DC motors require considerably more power. In Fig. 6.8 the Adafruit Motor Shield, and the shield assembled on the Arduino board, are shown.



(a)



(b)

Figure 6.8: (a) Adafruit Motor Shield, (b) the Shield assembled on the Arduino UNO.

6.4 Seizures Simulator

6.4.1 Clonic Seizures Simulator

To simulate the typical movements of clonic seizures, we have used a prototype having the appearance of a newborn, driven by an Arduino microcontroller. Specifically, to limit the cost of the hardware simulator, we used a toy doll able to move its limbs and head [68]. In Fig. 6.9, the clonic seizures simulator is shown. The prototype has been programmed to alternate different types of movements, physiological and pathological, differentiated by duration, speed, pause between two successive kinds of movements and direction of rotation of the motor. The pathological movements were obtained by varying the supply voltage of the DC motor, so as to vary its speed, in order to obtain movements with different frequency ranging between 0.3 Hz and 2 Hz.

After the programming phase of the simulator, we have performed video recordings of the simulations, through a multi-camera system, so as to create a video database useful to test automatic video processing algorithms to the video realized filming the simulator from different perspectives. As an initial experiment, we have made a total of 30 videos, lasting 3 minutes each, using 3 cameras in each video.

It is important to stress that the realized simulator has a very low cost, the overall cost, in fact, is below $ 50.

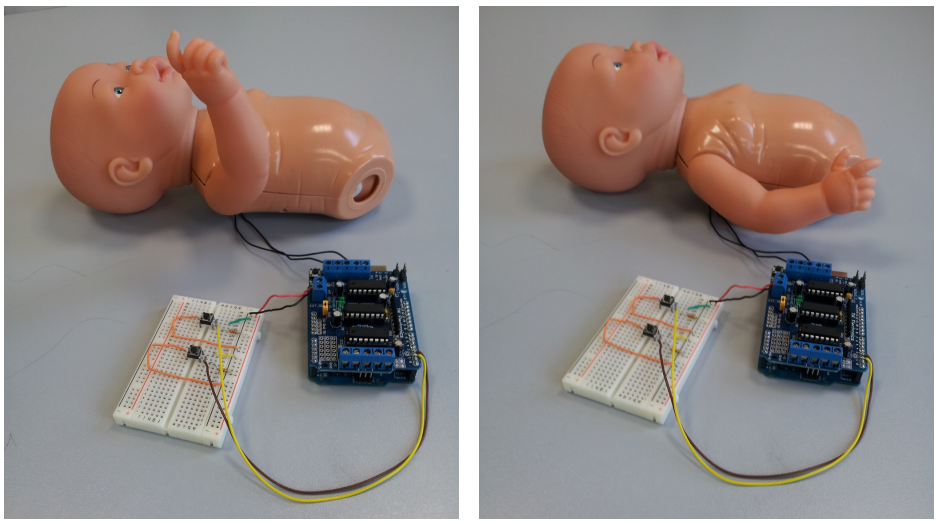
6.4.2 Tonic Seizures Simulator

To simulate tonic seizures, we have used the same simulator prepared for clonic seizures. Tonic seizures are characterized by a sudden stiffening of the limb muscles with a gradual release of the same that occurs in a few seconds; some tonics, after the stiffening phase, manifest small tremors for a short period [69].

We have simulated only tonic seizures involving one arm. In Fig. 6.10 the tonic seizures simulator with raised arm, and with the arm down, are shown.

Once again, as an initial experiments, we have performed video recordings of the simulations. We have made a total of 15 videos, lasting 3 minutes each, using 3

Figure 6.9: Clonic Seizures Simulator.



(a) (b)

Figure 6.10: (a) Tonic seizures simulator with raised arm, and (b) tonic seizures simulator with the arm down.

cameras in each video. In every video there are two separate tonic seizures associated with tremors.

6.5 SARA: Simulatore Apnee Respiratorie Automatico

The objective of this additional simulator is to simulate respiratory crises that lead to apnoeas. For these simulations we have used another infant toy doll, to which we have carved a part of the chest and inserted a servo motor, to simulate the respiratory excursion of the chest of an infant [70]. The simulator can vary the respiratory rate, lowering it towards the hypoventilation and until the absence of breathing. During the physiological breath periods, the simulator varies the respiratory rate between 0.5 Hz and 1 Hz; while in the hypoventilation phase, the simulator has a respiratory rate of 0.2 Hz and a reduced thoracic excursion.

We have called this simulator SARA (Simulatore Apnee Respiratorie Automatico),

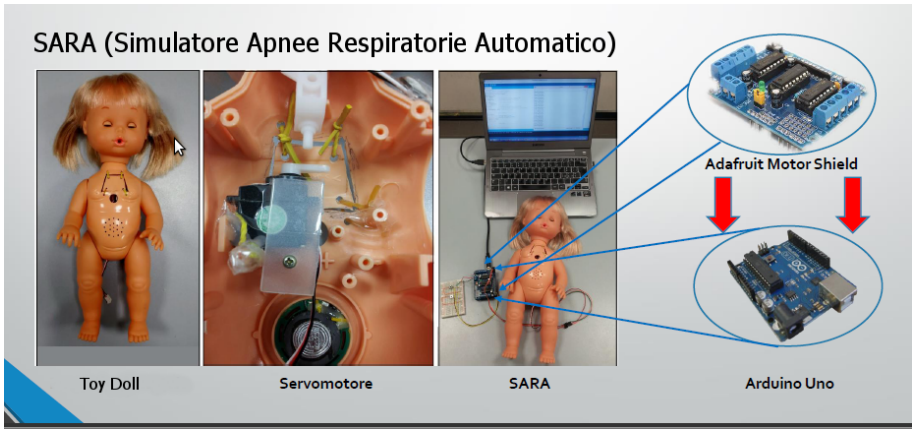


Figure 6.11: Components of the SARA system.

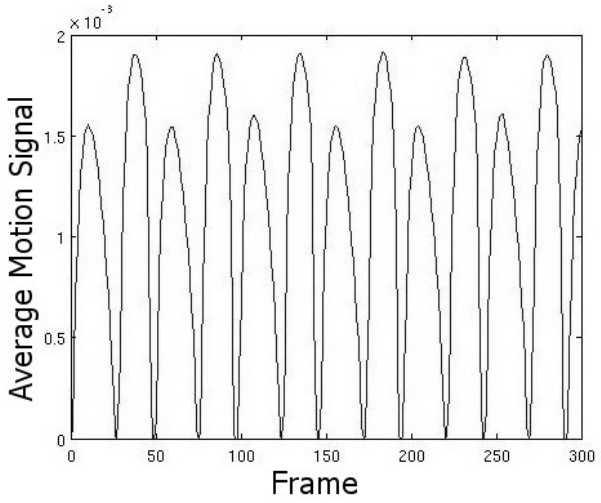
a female italian name, which is also an acronym that means “an automatic simulator of respiratory apnoeas”. In Fig. 6.11, the components of the SARA simulator are shown.

Even with this simulator, we have recorded a total of 30 videos, lasting 3 minutes each, using 3 cameras in each video. In every video there are two separate apnoea events. Every video is differentiated by the breathing frequency, the apnoeas duration, and the hypoventilation duration.

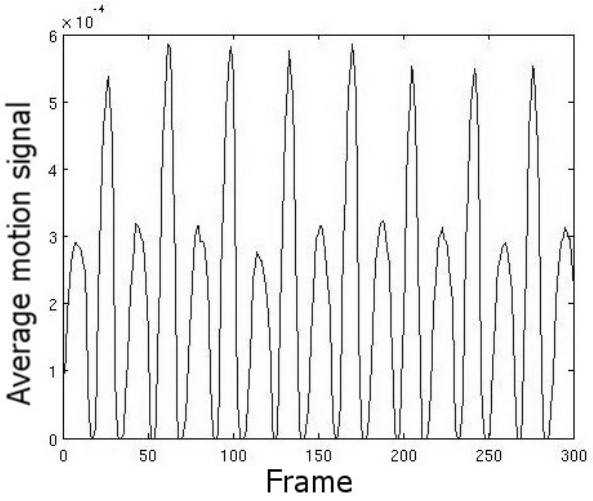
In Fig. 6.12, two average movement signal extracted from a camera while framing the simulator breathing normally, are shown. Each respiratory cycle consists of two phases: an inhalatory phase with larger chest movements, and an expiratory phase with smaller movements. In Fig. 6.12, therefore, the lower peaks corresponds to the expiratory phase, while the highest peaks correspond to the inhalatory phase.

In Fig. 6.13, the average motion signal extracted while the SARA simulator was performing an abnormal breathing event is shown.

Once again, this simulator has a very low cost, the overall cost, in fact, is below $ 50.



(a)



(b)

Figure 6.12: Average motion signal extracted from a recording of the SARA simulator while breathing normally, with sampling frequency equal to 25 frames/s: (a) respiratory rate = 0.53 Hz, and (b) respiratory rate = 0.66 Hz.

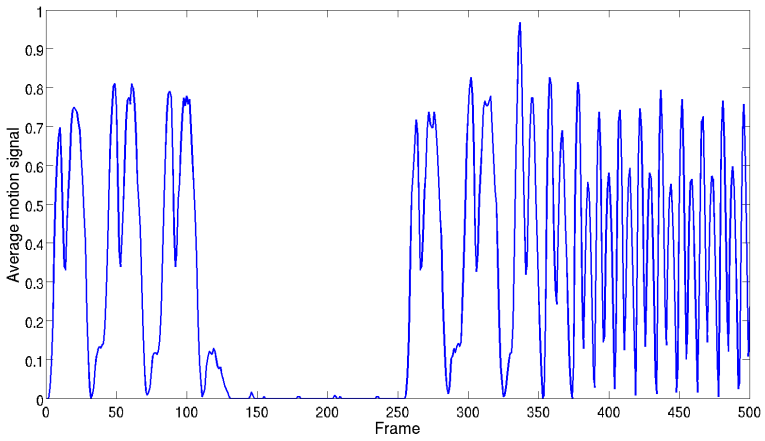


Figure 6.13: Average motion signal extracted while the SARA simulator was performing an abnormal breathing event.

Chapter 7

SmartCED

"The time will come when, whatever illness you may have, for your body type and health history, there will 'be an app for that' to keep you on your best path to wellness."

– Sheldon Whitehouse

7.1 Introduction

Devising comfortable and low cost respiratory rate measurement systems, that could also allow home monitoring, reducing the time and the costs of hospitalization is of significant interest. The biggest challenge is to find methods that do not require the application of sensors to the patient's body. In this context, we present "Smartphone based contactless epilepsy detector" (SmartCED) [71], an Android App [72] able to diagnose neonatal clonic seizures and warn the user about the possible occurrence in real-time. It is important to notice that the App could also be used on adult patients, in fact, it has a general purpose. The main aim of the project is to implement a wire-free and low-cost diagnosis system executing all the necessary processing directly on the smartphone. The App presents a user-friendly interface in order to extend its use to even unskilled staff. The goal is to extend the desktop application illustrated in Chapter 3, to the mobile world. Thanks to the great recent hardware evolution, we

Figure 7.1: smartCED Logo

may be able to perform real time live video processing and reactively alert the user about the possible occurrence of a seizure. This inexpensive monitoring system could help to control all the little guests in a NICU, 24 hours a day. Furthermore, every child could be easily monitored at home without additional hardware costs. In Fig. 7.1, the logo of the App is shown. smartCED is designed for an easy and intuitive utilization, although it integrates complex software, from the technical point of view. The user has to start the App, frame the patient, preferable with the help of a tripod, and start monitoring the patient with a simple touch.

7.2 Extraction of Motion Information

The main App's feature is the extraction of motion information from the frames delivered by the smartphone's camera in order to recognize the seizure.

7.2.1 Motion Signal Extraction

The main method for the extraction of the motion signal is the one described in Section 3.1.1. This technique is illustrated in Fig. 7.2, where it is possible to see the different steps of the algorithm: the conversion of two frames to grayscale, single difference between the two frames and binarization. The scene is constituted by a pen moving, with a static part, a laptop, on the background. The image labelled with 'a', represents the scene in grayscale displayed on the screen of the smartphone, the

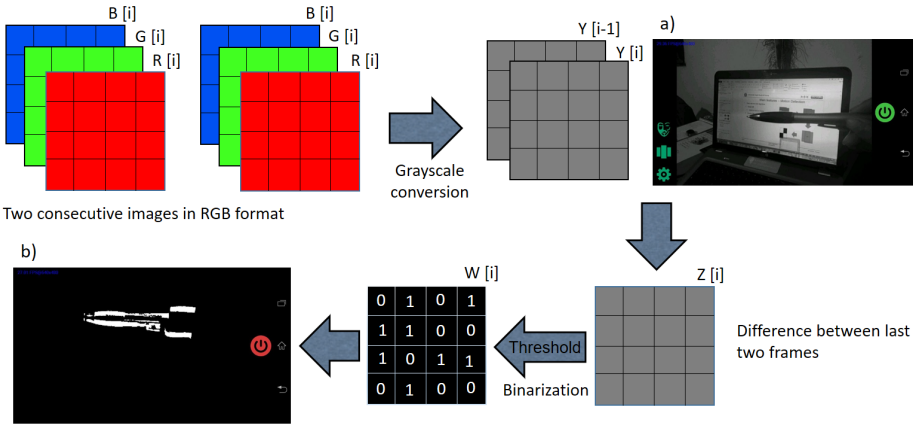


Figure 7.2: Motion Signal Extraction with the Single Difference Algorithm

image labelled with 'b', is the final result of the processing, only pixels relating to the pen motion are still highlighted.

As we noted in the processed frame, the presence of very thick edges in the moving parts, we have introduced a second method for the extraction of the motion signal, called Double-Difference. These thick edges may be caused by unexpected changes in the environment's light, shadow noise or by the presence of very abrupt and sudden movements in the scene, and they can cause an undesirable increase of false alarms. The essential difference with the basic algorithm, consists in using 3 frames to calculate the Difference of Frame (DoF), instead of just 2 frames used in the basic algorithm. Letting $\mathbf{Y}[i-2]$, $\mathbf{Y}[i-1]$, and $\mathbf{Y}[i]$ denote three consecutive frames in grayscale, we obtain $\mathbf{Z}_1[i]$ and $\mathbf{Z}_2[i]$ by the following differences,

$$\mathbf{Z}_1[i] = |\mathbf{Y}[i-2] - \mathbf{Y}[i-1]| \quad i = 0,1,2,... \tag{7.1}$$

$$\mathbf{Z}_2[i] = |\mathbf{Y}[i-1] - \mathbf{Y}[i]| \quad i = 0,1,2,... \quad . \tag{7.2}$$

$\mathbf{Z}_1[i]$ and $\mathbf{Z}_2[i]$ are then binarized to $\mathbf{B}_1[i]$ and $\mathbf{B}_2[i]$. Finally, by a logical AND function

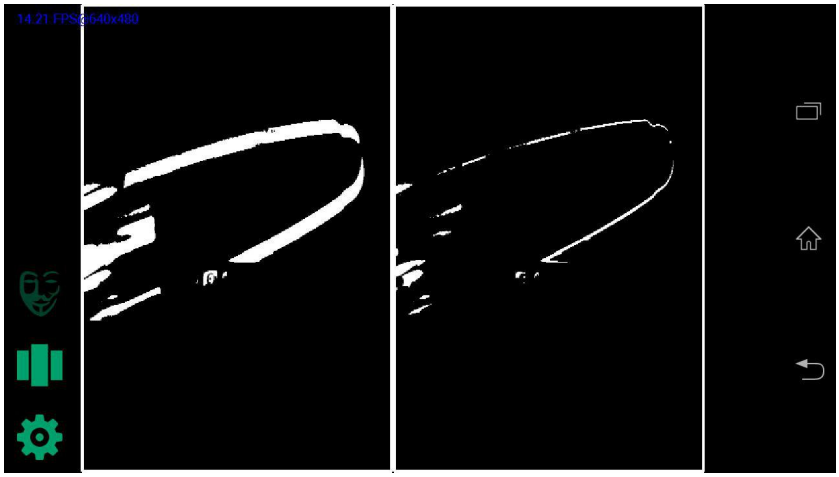


Figure 7.3: Comparison between two motion extraction algorithm: single difference on the left, double difference on the right. The scene is relative to a pen in motion.

we obtain,

$$\mathbf{B}[i] = \mathbf{B}_1[i] \quad AND \quad \mathbf{B}_2[i] \quad i = 0,1,2,... \tag{7.3}$$

which represents the final processed frame. The results obtained, are clearly visible in Fig. 7.3, that compares the two algorithms. Notice how, with the Double Difference Algorithm, the shadows, due to light changes, have been suppressed in the contours of the subject. In fact, in case of double difference the edges are definitely sharper. This allows a more accurate detection of rapid or large movements. In the end, we noticed that the algorithm suitable for our context is still the Single Difference, since the body of newborns is tiny and is affected, in most cases, by non abrupt/violent crisis. In Fig. 7.4 , two examples of signals extracted with the smartCED App, following the method illustrated in Chapter 3, are shown. This signals are obtained framing the clonic seizures simulator described in Chapter 6 performing physiological or pathological movements.

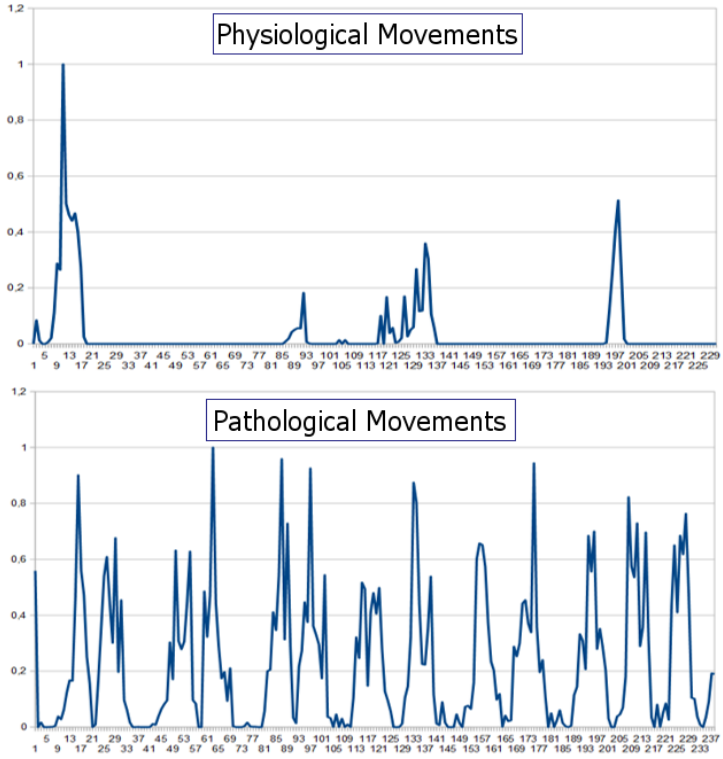


Figure 7.4: Examples of Signals Extracted with the smartCED App.

7.2.2 Detection of Periodicity

The seizure detection is based on the study of the periodicity of the extracted motion signal. This activity is performed in a background process, in order to have real-time response from the application. This part of the App is written in C, in order to reduce the processing time. To interface Java and C codes we have used the Native Development Kit (NDK) [73], a toolset that allows you to implement parts of an app using native-code languages such as C and C++. Typically, good use cases for the NDK are CPU-intensive applications such as signal processing. The signal evaluation takes place in an AsyncTask, thus exploiting the multitasking and multithreading capabil-

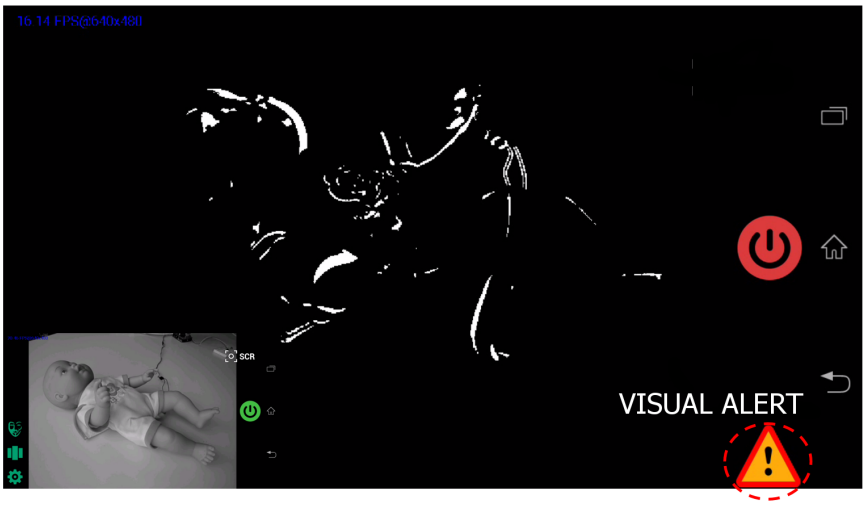


Figure 7.5: If a seizure is detected, smartCED will launch a visible and audible alarm.

ities of Android. As result, the user is warned about a possible seizure detection, through an audible and visible alert, as it can be seen in the Fig. 7.5.

7.3 System Requirements

There are some necessary requirements to let the application smartCED work properly:

- a smartphone equipped with an RGB sensor
- Android versions: from Honeycomb (API 11) to Lollipop (API 22) [74]
- OpenCV Manager: it is a required external app to use OpenCv libraries. Its purpose is to provide, at runtime, the best version of OpenCV library for the underlying hardware. In case it is not installed on the smartphone, smartCED will prompt the user to install it first
- minimum frame rate: 20 fps.

The lack of the latter requirements may produce unsatisfactory results. It is important to underline that, although the modern smartphones support much higher resolutions, the processing is performed at a resolution of 640x480 pixels. In fact high resolutions, resulting in matrices of large dimensions, could cause delay in the motion detection algorithms, lowering the frame rate. A lower frame rate means poor information, and, consequently, a less efficient detection.

As shown in [75], the use of OpenCV library has very positive impact on frame rate and power consumption, a very important aspect to be considered in mobile devices. In fact, using OpenCV library for video processing, instead of Android native modules, may give a higher frame rate and a low power consumption. This has been experimentally established, by observing how the frame rate and the battery consumption is impacted during various type of processing done using the openCV library or the Android native modules. The secret, is in the well written openCV library, with organized and optimized modules for video processing [75, 76].

7.4 App Description

The application has been structured to provide an intuitive interface to the final users, taking into account that it may also be used outside of the hospital environment by unskilled personnel. The App is provided with several activities described below. An activity is a single, focused action that the user can perform. It is an application component that provides a screen with which users can interact in order to do something, such as dial the phone, take a photo, send an email, or view a map. Each activity is given a window in which to draw its user interface. The window typically fills the screen, but may be smaller than the screen and float on top of other windows.

7.4.1 Camera Activity

CameraActivity is the main Activity of the application, it is shown on startup. Its layout consists of a SurfaceView [72] containing frames from the camera and some ImageButton to permit interaction with the user. The result is represented in Fig. 7.6. The main functionality provided by this activity is driven by the button “Processing

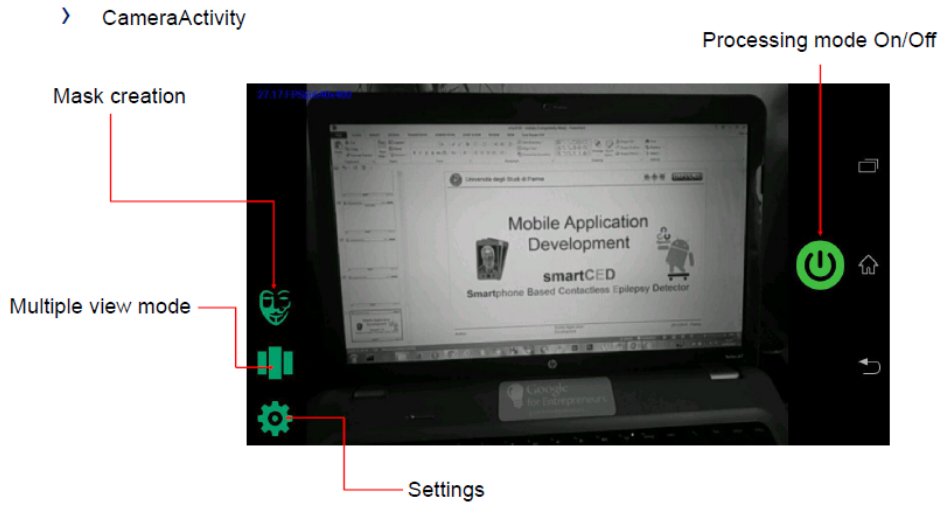


Figure 7.6: The Camera Activity.

mode: On/Off" on the right: when of green color, gray scale frames are visualized on the device screen with a frequency depending on the hardware of the device in use, in order to allow the subject framing. After pressing the button, the color turns red to indicate that the patient is being monitored. The buttons on the left, are used to switch between Activities listed below.

7.4.2 Settings Activity

This activity is dedicated to settings of the App. It permits to choose the motion detection algorithm described in Subsection 7.2.1 and to select the binarization threshold value described in (3.2). The result is shown in the Fig. 7.7.

7.4.3 Multiple Vision Mode

This activity allows to display, according to the choice from the pop-up menu appeared as a result of the click, two types of frames at once (side by side). There are 3 combinations denoted in the format First frame / Second frame:

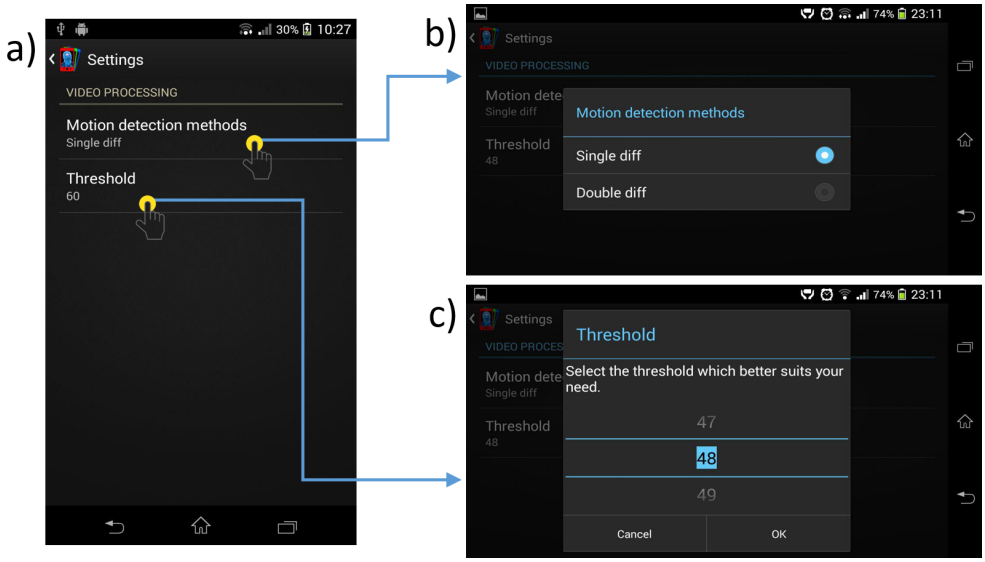


Figure 7.7: (a) The Settings Activity: (b) motion detection algorithm selection, (c) binarization threshold selection.

- grayscale / single difference algorithm;
- grayscale / double difference algorithm;
- single difference algorithm / double difference algorithm.

A tap on the right or left side allows to set directly the algorithm for the motion detection without opening the SettingsActivity. This activity is shown in Fig. 7.8.

7.4.4 Region of Interest Selection

This activity allows to select a region of interest (ROI), e.g., a rectangular area on the screen; the monitoring will be performed only in this area in order to remove not relevant parts of the scene. This is useful, for example, when there are moving subjects around the baby, which introduce noise into the motion signal. The selection is intuitive: one has just to draw on the screen the diagonal of the desired rectangle.

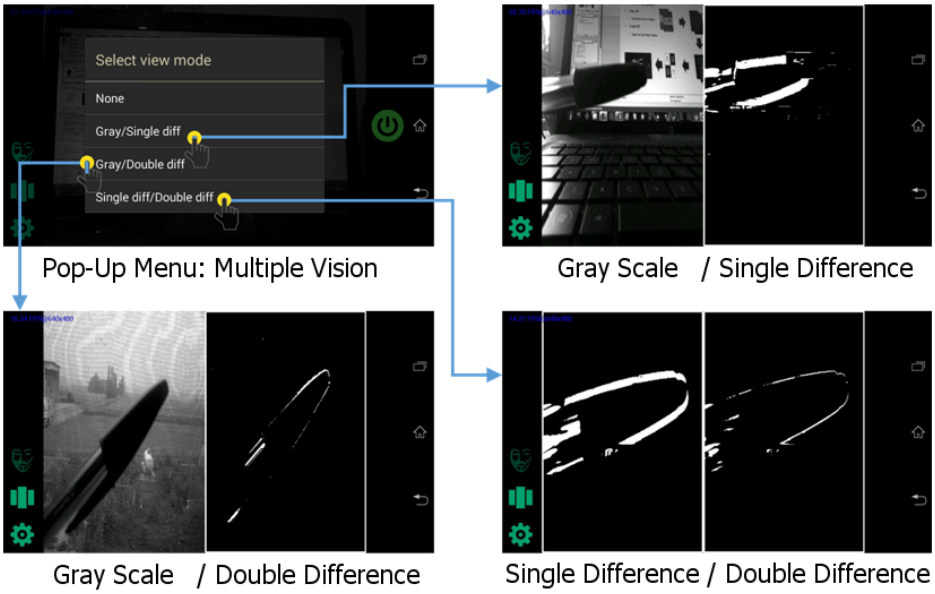


Figure 7.8: Multiple Vision.

This is achieved by the use of a particular View with the ability to read the movements on the display with visual feedback: GestureOverlayView [77]. The initial touch point, read through the onGestureStart callback() [77], and the final touch point, read by onGestureEnded() [77], are used for the area selection, they are, in fact, diagonal's extremes of the rectangle to be selected. These two points are used to create a mask, $\mathbf{M}[i]$, which is nothing more than a matrix with all zeroes outside of the rectangle and all ones inside it. The mask is then superimposed on the original frame, $\mathbf{Y}[i]$, coming from the camera by a bitwise logical And. The result is a new matrix,

$$\mathbf{Z}[i] = \mathbf{Y}[i] \quad AND \quad \mathbf{M}[i] \tag{7.4}$$

where pixels multiplied by one remain unchanged, and others are turned into zero (black color). Then the resulting matrix, $\mathbf{Z}[i]$ is displayed on the screen. The mask creation process is illustrated in Fig. 7.9.

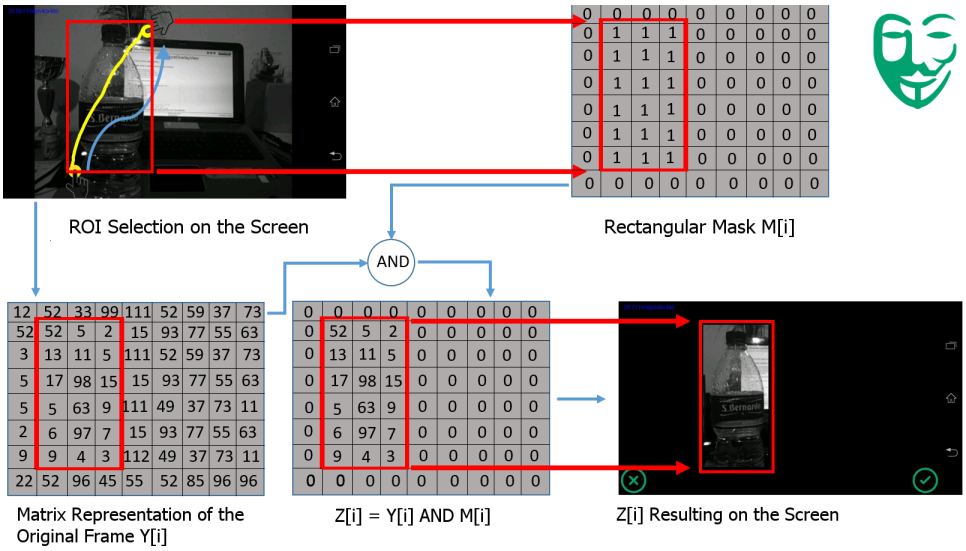


Figure 7.9: The Mask Creation Process.

7.4.5 Local DataBase

SmartCED's App is able to store in a local database the data relative to the detected crisis: date/start time and date/end time of each crises, GPS coordinates of the place where the crisis occurred and an image of the crisis. The collected data are shown using a ListView, as shown in Fig. 7.10.

7.4.6 GeoLocalization

In this activity, one can display on a map the geographic distribution of the detected crises, as shown in Fig. 7.11.

7.4.7 Statistical Analysis

In this activity, you can view graphs of daily/monthly/yearly statistics about the number of crises or the average duration of the crises. In Fig. 7.12, an example of a report

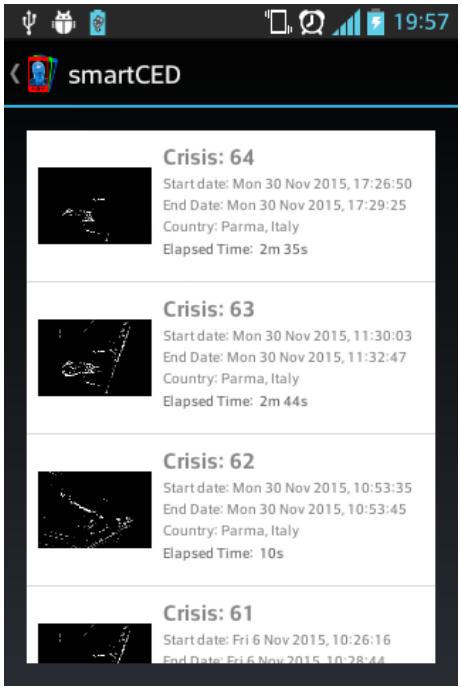


Figure 7.10: The Local DB Activity.

about the number of daily seizures relative to a period of time set by the user, is shown.

7.5 Performance Analysis

The SmartCED App was tested in the laboratory using as 'subject' the clonic seizure simulator described in Subsection 6.4.1. It is important to note that the tests, done in the lab, are influenced by various factors, such as changes in lighting conditions and the distance between the smartphone and 'the patient'. The simulated newborn, furthermore, despite being a nice simulator, is a mechanical component that fails to depict real natural movements.

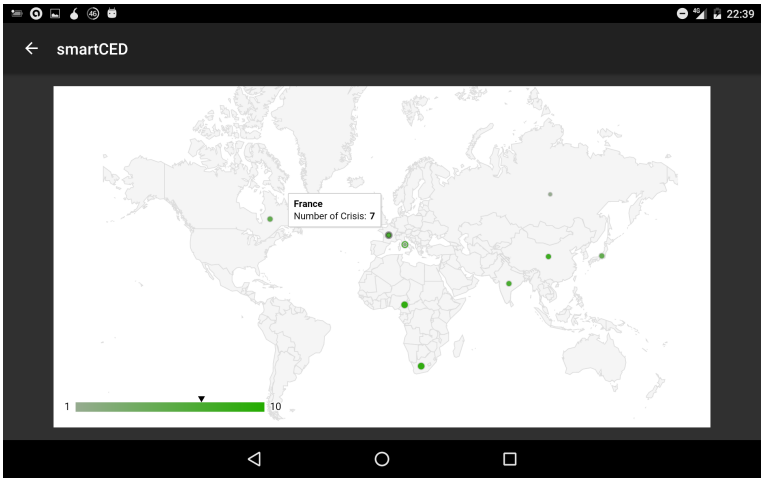


Figure 7.11: The Geolocalization Activity.

Table 7.1: Estimation of environmental variability: mean and standard deviation.

	Mean	Standard deviation
Sensitivity	91%	7%
Specificity	77%	12%

7.5.1 Environmental Variability

The first tests carried out, therefore, were performed in order to estimate the environmental variability. We have performed 14 simulation sessions, with a fixed binarization threshold, lasting 6 minutes each for a total duration of 1 hour and 24 minutes. The results confirm that the light actually plays an important role, in fact, for each session we may notice changes in the values of Sensitivity and Specificity as shown in the Table 7.1.

7.5.2 Best Threshold Value

To define the optimal threshold defined in Eq. 3.2 we performed 11 simulation sessions lasting 6 minutes each, for a total duration of 1 hour and 6 minutes. Table 7.2

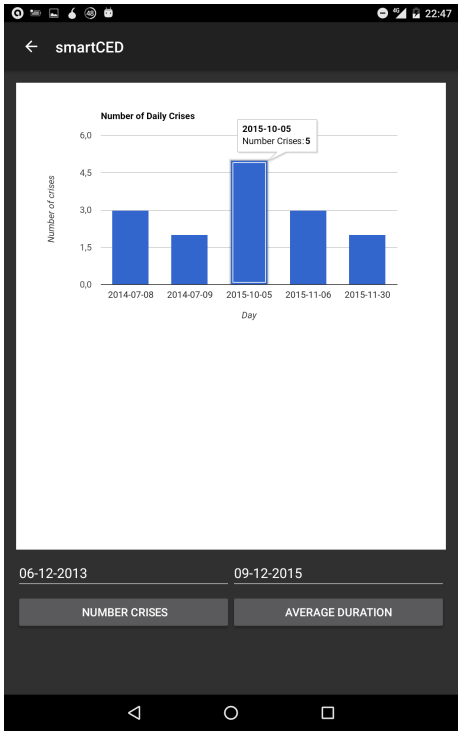


Figure 7.12: The Statistical Analysis Activity.

Table 7.2: Best threshold search.

	40	41	42	43	44	45	46	47	**48**	49	50
True positive	15	18	15	17	18	12	17	18	**17**	12	17
True negative	16	11	14	16	14	17	14	16	**18**	14	11
False positive	2	7	4	2	4	1	4	2	**0**	4	7
False negative	3	0	3	1	0	6	1	0	**1**	6	1
Sensitivity	83%	100%	83%	94%	100%	66%	94%	100%	**94%**	66%	94%
Specificity	88%	61%	77%	88%	77%	94%	77%	88%	**100%**	77%	61%

shows the values of True Positives, True Negatives, False Positives, False Negatives, Sensitivity and Specificity obtained varying the binarization threshold. Test indicates 48 as the best threshold value.

Table 7.3: App's performance.

	Real positive	Real negative
True positive	$TP = 371$	$FP = 36$
True negative	$FN = 53$	$TN = 260$
	SE = 87%	SP = 88%

7.5.3 Simulations

Having derived the optimal threshold value, we conducted a further test, in order to collect statistical data and to establish the App's efficiency measured as in Eq. 3.9 and Eq. 3.10. The test consists of 20 sessions of 6 minutes, in total 2 hours, with the mechanical simulator that reproduces pathological and physiological movements of different durations. Results are summarized in Table 7.3. Although preliminaries, the obtained results are encouraging, we believe this innovative monitoring system has a great potential and could represent a timely, user-friendly, low-cost and 24-hour-available tool to monitor patients suffering from clonic seizures.

Conclusions

"Any sufficiently advanced technology,
is indistinguishable from magic."

– Arthur C. Clarke

In this work, we have presented a unified approach to non-invasive, wire-free, low-cost video processing for neonatal monitoring and automatic detection of diseases characterized by specific movement patterns. This category includes neonatal clonic seizures and breathing disorders, such as Congenital Central Hypoventilation Syndrome (CCHS). Since both disorders are characterized by the presence or absence of periodic movements of parts of the body, namely the limbs in case of clonic seizures and the chest in case of CCHS, it is possible to analyse the data obtained from multiple video sensors placed around a patient, extracting an average motion signal and estimating, using the Maximum Likelihood (ML) criterion, its periodicity. This approach is very versatile and allowed us to investigate various scenarios, including: a single Red, Green and Blue (RGB) camera; an RGB-Depth sensor; a network of a few RGB cameras. Data fusion principles are considered to aggregate the signals from multiple sensors. In the case of respiratory diseases, since chest movements are subtle, the video can be pre-processed by a recently proposed selective magnification algorithm, namely the eulerian video magnification (EVM), which has the purpose of emphasizing small movements. Within this context, we have also developed a second improved algorithm in order to speed up the processing time required for the detection of apnoeas, limiting the computational load.

Moreover, in order to have, at any time, a 'subject' on which to test the continuously evolving detection algorithms, we have decided to realize two low-cost programmable simulators able to replicate the symptomatic movements characteristic of the diseases under consideration.

The performance of the proposed detection algorithms is assessed, in terms of sensitivity, specificity, and Receiver Operating Characteristic (ROC) curves, considering real video recordings of newborns provided by a Neonatal Intensive Care Unit (NICU). The diagnostic performance of our detection systems has been compared to that of the gold standard based on a prolonged polysomnographic EEG monitoring.

It is important to stress how we we have always pursued simplicity, because low complexity leads to a low processing time, and this means that these algorithms can be used on a wide range of hardware devices.

This innovative detection system could represent a timely, user-friendly and 24-hour-available tool to permanently monitor every patient in the NICU or specific patients at home. In particular SmartCED, the Android application developed for the detection of clonic seizures, is a powerful low-cost tool for massive monitoring of patients in their homes, using devices they already own. Furthermore, a wire-free technique may be more user-friendly and highly desirable when used with infants, in particular with newborns.

As a possible future development, SmartCED could be installed on the smartphones of the medical staff of the Department of Neurosciences at the University of Parma, in order to start an experimental phase on newborns patients. The goal, furthermore, is also to provide the application to the parents of pediatric patients for home monitoring. The collection of statistical data by smartCED may help to understand how to improve the detection algorithms.

List of Publications

International Conferences

- D. Alinovi, **L. Cattani**, G. Ferrari, F. Pisani and R. Raheli. “Video Simulation of Apnoea Episodes”. In Multimedia Expo Workshops (ICMEW), 2015 IEEE International Conference on, pages 1-6, June 2015
- D. Alinovi, **L. Cattani**, G. Ferrari, F. Pisani, and R. Raheli. “Spatio-temporal video processing for respiratory rate estimation”. In Medical Measurements and Applications (MeMeA), 2015 IEEE International Symposium on, pages 12-17, May 2015
- **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”.Biometric Measurements and Systems for Security and Medical Applications (BIOMS) Proceedings, 2014 IEEE Workshop on, pages 67-73, October 2014
- **L. Cattani**, G.M. Kouamou Ntonfo, F. Lofino, G. Ferrari, R. Raheli and F. Pisani. “Maximum-likelihood detection of neonatal clonic seizures by video image processing”. Medical Information and Communication Technology (ISMICT), 2014 8th International Symposium on, pages 1-5, April 2014.

Journals - Submitted

- **L. Cattani**, R. Raheli, E. Pavlidis, C. Spagnoli and F. Pisani. "Optimizing detection rate and characterisation of subtle neonatal abnormal movements with multi-camera video-EEG recordings". 2015 (Major review required)
- **L. Cattani**, E. Pavlidis, C.A. Tassinari and F. Pisani. "Neonatal seizure automatism and human inborn pattern of quadrupedal locomotion". 2015
- **L. Cattani**, D. Alinovi, G. Ferrari, R. Raheli, E. Pavlidis, C. Spagnoli and F. Pisani. "Apnoea in newborns. A wire-free, non-invasive, video processing based detection system". 2015

International Conferences - Submitted

- **L. Cattani**, H.P. Saini, G. Ferrari and R. Raheli. "SmartCED, an Android Application for Neonatal Seizures Detection". 2016

Journals - To Be Submitted

- **L. Cattani**, D. Alinovi, G. Ferrari, F. Pisani and R. Raheli. "Monitoring infants by automatic video processing: a united approach to motion analysis". 2016
- D. Alinovi, **L. Cattani**, G. Ferrari, F. Pisani and R. Raheli. "Markov chain modelling of breathing patterns and applications". 2016

Bibliography

[1] J. J. Volpe. *Neurology of the Newborn*. WB. Saunders, Philadelphia, USA, 2001.

[2] G. M. Ronen, S. Penney, and W. Andrews. The epidemiology of clinical neonatal seizures in Newfoundland: A population-based study. *J. of Pediatrics*, 13(1):71–75, January 1999.

[3] J. J. Tramonte and H. P. Goodkin. Temporal lobe hemorrhage in the full-term neonate presenting as apneic seizures. *J. of Perinatology*, 24(11):726–729, 2004.

[4] S. Shah and G. Q. Sharieff. An update on the approach to apparent life-threatening events. *Current Opinion in Pediatrics*, 19(3):288–294, 2007.

[5] F. Healy and C. L. Marcus. Congenital central hypoventilation syndrome in children. *Paediatric Respiratory Reviews*, 12(4):253–263, 2011.

[6] H. Trang, M. Dehan, F. Beaufils, I. Zaccaria, J. Amiel, and C. Gautier. The French congenital central hypoventilation syndrome registry: General data, phenotype, and genotype. *Chest*, 127(1):72–79, 2005.

[7] D. E. Weese-Mayer, J. M. Silvestri, A. D. Huffman, S. M. Smok-Pearsall, M. H. Kowal, B. S. Maher, M. E. Cooper, and M. L. Marazita. Case/control family study of autonomic nervous system dysfunction in idiopathic congenital central hypoventilation syndrome. *American J. of Medical Genetics*, 100(3):237–245, 2001.

[8] C. M. Rand, P. P. Patwari, M. S. Carroll, and D. E. Weese-Mayer. Congenital central hypoventilation syndrome and sudden infant death syndrome: Disorders of autonomic regulation. *Semin. Pediatr. Neurol.*, 1(20):44–55, 2013.

[9] S. M. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory*, volume 1 of *Prentice Hall Signal Processing*. Prentice Hall, 1993.

[10] R. C. Gonzalez and R. E. Woods. *Digital Image Processing*. Pearson - Prentice Hall, 3 edition, 2008.

[11] J. Webb and J. Ashley. *Beginning Kinect programming with the Microsoft Kinect SDK*. Apress, New York, USA, 2012.

[12] D. Weaver. *Epilepsy and Seizures: Everything You Need to Know (Your Personal Health)*. Firefly Books, April 2001.

[13] D. J. Henderson-Smart. The effect of gestational age on the incidence and duration of recurrent apnoea in newborn babies. *Australian Pediatric Journal*, 17:271–276, 1981.

[14] E. C. Eichenwald, A. Aina, A. R. Stark. Apnea frequently persists beyond term gestation in infants delivered at 24 to 28 weeks. *Pediatric*, 100(3):354–359, 1997.

[15] L. C. Winckworth, N. L. Earl, and C. Steward. Apparent life-threatening events: difficult for parents, difficult for clinicians. *Archives of Disease in Childhood*, April 2013.

[16] T. M. Baird. Clinical correlate, natural history and outcome of neonatal apnea. *Semin Neonatal*, 9(3):205–211, 2004.

[17] K. Ljubic, I. Fister Jr., and I. Fister. Congenital central hypoventilation syndrome: A comprehensive review and future challenge. Technical Report 2014, Faculty of Medicine, University of Maribor, Taborska 8, 2000 Maribor, Slovenia, 2014.

[18] S. J. M. Smith. Eeg in the diagnosis, classification, and management of patients with epilepsy. *Journal of Neurology, Neurosurgery & Psychiatry*, 76(2):2âĂŞ7, June 2005.

[19] J. Colon. Paradise sleep [online]. 2015. URL: `http://www.paradisesleep.com` [cited 20-12-2015].

[20] N. B. Karayiannis, S. Srinivasan, R. Bhattacharya, M. S. Wise, J. D. Frost Jr., and E. M. Mizrahi. Extraction of motion strength and motor activity signals from video recordings of neonatal seizures. *IEEE Transactions on Medical Imaging*, 20(9):965–980, September 2001.

[21] B. D. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *7th International Joint Conference on Artificial Intelligence (IJCAI'81)*, pages 674–679, Vancouver, Canada, August 1981.

[22] N. B. Karayiannis, E. M. Mizrahi, B. Varughese, G. Tao, J. D. Frost Jr., and M. S. Wise. Quantifying motion in video recordings of neonatal seizures by regularized optical flow methods. *IEEE Transactions on Image Processing*, 14(7):890–903, July 2005.

[23] S. Kalitzin, G. Petkov, D. Velis, B. Vledder, and F. Lopes da Silva. Automatic segmentation of episodes containing epileptic clonic seizures in video sequences. *IEEE Transaction on Biomedical Engineering*, 59(12):3379–3385, December 2012.

[24] K. Cuppens, L. Lagae, B. Ceulemans, S. Van Huffel, and B. Vanrumste. Automatic video detection of body movement during sleep based on optical flow in pediatric patients with epilepsy. *Medical and Biological Engineering and Computing*, 48:923–931, June 2010.

[25] K. Samiee, P. Kovacs, and M. Gabbouj. Epileptic seizure classification of EEG time-series using rational discrete short-time fourier transform. *Biomedical Engineering, IEEE Transactions on*, 62(2):541–552, Feb 2015.

[26] Y. Wang, Y. Qi, J. Zhu, J. Zhang, Y. Wang, G. Pan, X. Zheng, and Z. Wu. A cauchy-based state-space model for seizure detection in eeg monitoring systems. *Intelligent Systems, IEEE*, 30(1):6–12, Jan 2015.

[27] A. Ueno and Y. Yama. Unconstrained monitoring of ECG and respiratory variation in infants with underwear during sleep using a bed-sheet electrode unit. In *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, pages 2329–2332, Aug 2008.

[28] C. Linti, H. Horter, P. Osterreicher, and H. Planck. Sensory baby vest for the monitoring of infants. In *Wearable and Implantable Body Sensor Networks, 2006. BSN 2006. International Workshop on*, pages 3 pp.–137, April 2006.

[29] J. Coosemans, B. Hermans, and R. Puers. Integrating wireless ECG monitoring in textiles. In *Solid-State Sensors, Actuators and Microsystems, 2005. Digest of Technical Papers. TRANSDUCERS '05. The 13th International Conference on*, volume 1, pages 228–232, June 2005.

[30] G. P. Heldt and R. J. Ward. Evaluation of ultrasound-based sensor to monitor respiratory and non-respiratory movement and timing in infants. *Biomedical Engineering, IEEE Transactions on*, PP(99):1–1, 2015.

[31] N. Patwari, J. Wilson, S. Ananthanarayanan, S. K. Kasera, and D. R. Westenskow. Monitoring breathing via signal strength in wireless networks. *IEEE Trans. Mobile Comput.*, 13(8):1774–1786, Aug. 2014.

[32] L. Cattani, G. M. Kouamou Ntonfo, F. Lofino, G. Ferrari, R. Raheli, and F. Pisani. Maximum-likelihood detection of neonatal clonic seizures by video image processing. In *8th Intern. Symp. Medical Inform. and Commun. Technol.* (ISMICT 2014), Florence, Italy, April 2014.

[33] G. M. Kouamou Ntonfo, G. Ferrari, R. Raheli, and F. Pisani. Low-complexity image processing for real-time detection of neonatal clonic seizures. *IEEE Trans. Inf. Technol. Biomed.*, 16(3):375–382, May 2012.

[34] E. M. Mizrahi and P. Kellaway. Characterization and classification of neonatal seizures. *Neurology*, 37(12):1837–1844, December 1987.

[35] A. G. Lalkhen and A. McCluskey. Clinical tests: sensitivity and specificity. *Continuing Education in Anaesthesia, Critical Care & Pain Advance Access*, 8(6):221–223, Dec. 2008.

[36] T. A. Schonhoff and A. A. Giordano. *Detection and estimation theory and its applications*. Prentice Hall, Upper Saddle River, N.J., USA, 2006.

[37] J. Swets. Measuring the accuracy of diagnostic systems. *Science*, 240(4857):1285–93, Jun. 1998.

[38] P. Kaew Tra Kul Pong and R. Bowden. An improved adaptive background mixture model for real-time tracking with shadow detection. In *Proc. 2nd European Workshop on Advanced Video Based Surveillance Systems*, volume 25, pages 1–5, 2001.

[39] C. Stauffer and W. E. L. Grimson. Adaptive background mixture models for real-time tracking. In *IEEE Conf. Comp. Vision Pattern Rec.*, volume 2, pages 2246–2252, Fort Collins, Colorado, USA, August 1999. IEEE.

[40] H. Y. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, and W. Freeman. Eulerian video magnification for revealing subtle changes in the world. *ACM Trans. Graph.*, 31(4):65:1 – 65:8, July 2012.

[41] 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 *IEEE Workshop on Biometric Measurements and Systems for Security and Medical Applications* (BIOMS), Rome, Italy, October 2014.

[42] D. Alinovi. Elaborazione di segnali video per l'amplifiazione del movimento e l'estrazione di caratteristiche teorico-informative. Master's thesis, Università degli Studi di Parma - Dipartimento di Ingegneria dell'Informazione, 2013.

[43] P. J. Burt and E. H. Adelson. The Laplacian pyramid as a compact image code. *IEEE Transactions on Communications*, 31(4):532–540, 1983.

[44] N. N. Finer, R. Higgins, J. Kattwinkel, and R. J. Martin. Summary proceedings from the apnea-of-prematurity group. *Pediatrics*, 117(1):S47–S51, 2006.

[45] W. T. Dowdell, S. Javaheri, and W. McGinnis. Cheyne-stokes respiration presenting as sleep apnea syndrome: Clinical and polysomnographic features. *American Review of Respiratory Disease*, 141(4 Pt 1):871–879, Apr. 1990.

[46] M. T. O'Toole, editor. *Mosby's Medical Dictionary*. Elsevier Health Sciences, 9 edition, 2013.

[47] D. Alinovi, L. Cattani, G. Ferrari, F. Pisani, and R. Raheli. Spatio-temporal video processing for respiratory rate estimation. In *Medical Measurements and Applications (MeMeA), 2015 IEEE International Symposium on*, pages 12–17, May 2015.

[48] A. V. Oppenheim and R. W. Schafer. *Discrete-Time Signal Processing*. Prentice Hall Signal Processing Series. Pearson - Prentice Hall, 3 edition, 2010.

[49] L. B. Jackson. *Digital Filters and Signal Processing*. Springer, 3 edition, 1996.

[50] B. Stoustrup. *The C++ Programming Language*. Pearson Education Inc, London, UK, 4 edition, 2013.

[51] G. Bradski and A. Kaehler. *Learning OpenCV, computer vision with the OpenCv Library*. O'Reilly Media, Sebastopol, CA, USA, 1 edition, 2008.

[52] M. L. Good. Patient simulation for training basic and advanced skills. *Medical Education, Wiley Online Library*, 37:14–21, November 2003.

[53] Society for simulation in healthcare [online]. URL: `http://www.ssih.org` [cited 2015.12.31].

[54] Society in europe for simulation applied to medicine [online]. URL: `http://www.sesam-web.org` [cited 2015.12.31].

[55] Association for simulated practice in healthcare [online]. URL: http://www.aspih.org.uk/ [cited 2015.12.31].

[56] Simulation in healthcare [online]. URL: http://www.editorialmanager.com/sih/ [cited 2015.12.31].

[57] Clinical simulation in nursing [online]. URL: http://www.nursingsimulation.org/ [cited 2015.12.31].

[58] Laerdal. Laerdal medical group [online]. 2015. URL: http://www.laerdal.com [cited 20-12-2015].

[59] V. R. Taqueti and J. B. Cooper. A brief history of the development of mannequin simulators for clinical education and training. *Quality and Safety Health Care*, 13:11–18, October 2004.

[60] Gaumard. Gaumard scintific group [online]. 2015. URL: http://www.gaumard.com [cited 20-12-2015].

[61] M. G. MacDonald, J. Ramasethu, and K. Rais-Bahrami, editors. *Atlas of Procedures in Neonatology*. Lippincott Williams & Wilkins, 5 edition, 2007.

[62] M. Schmidt. *Arduino: A Quick Start Guide*, 1 edition, 2011.

[63] Wiring [online]. URL: http://wiring.org.co [cited 2015.12.31].

[64] C. Reas, B. Fry, and J. Maeda. *Processing: A Programming Handbook for Visual Designers and Artists*. The MIT Press, 1 edition, 2007.

[65] Arduino. Arduino [online]. 2015. URL: http://www.arduino.cc [cited 20-12-2015].

[66] S. Monk. *Programming Arduino Getting Started with Sketches*. McGraw-Hill, 1 edition, 2011.

[67] Datasheet ATmega328P [online]. URL: http://www.atmel.com/devices/atmega328p.aspx [cited 2015.12.31].

[68] D. Fontò. Realizzazione di un simulatore di convulsioni neonatali mediante microcontrollore arduino. Bachelor thesis, Università degli Studi di Parma - Dipartimento di Ingegneria dell'Informazione, 2014.

[69] A. S. Fauci, D. L. Kasper, S. L. Hauser, D. L. Longo, J. L. Jameson, J. Loscalzo, and T. R. Harrison. *Harrison's principles of internal medicine*. McGraw-Hill Medical, 19 edition, 2015.

[70] A. Opinto. Realizzazione di un simulatore di respiro neonatale mediante microcontrollore arduino. Bachelor thesis, Università degli Studi di Parma - Dipartimento di Ingegneria dell'Informazione, 2013.

[71] H. P. Saini. Sviluppo di un'applicazione android per la rivelazione di convulsioni neonatali. Bachelor thesis, Università degli Studi di Parma - Dipartimento di Ingegneria dell'Informazione, 2015.

[72] J. Howse. *Android Application Programming with OpenCV*. PACKT Publishing, 1 edition, 2013.

[73] S. Ratabouil. *Android NDK Beginner's Guide*. PACKT Publishing, 1 edition, 2012.

[74] T. McDonnell, B. Ray, and M. Kim. An empirical study of api stability and adoption in the android ecosystem. In *Software Maintenance (ICSM), 2013 29th IEEE International Conference on*, pages 70–79, 2013.

[75] A. Anuar, K. M. Saipullah, N. A. Ismail, and Y. Soo. Opencv based real-time video processing using android smartphone. *International Journal of Computer Technology and Electronics Engineering (IJCTEE)*, 1, 2011.

[76] D. Ortiz and N. Santiago. Impact of source code optimizations on power consumption of embedded systems. *Continuing Education in Anaesthesia, Critical Care & Pain Advance Access*, pages 133–136, June 2008.

[77] R. Meier. *Professional Android 4 Application Development*. Wrox, 1 edition, 2012.

Acknowledgments

First of all, I have to thank Tiziana for being by my side in this adventure, and for giving me, during my PhD years, two wonderful daughters. They have always helped me with their huge smiles during the hard moments.

I want to thank Prof. Riccardo Raheli for his guide and for his continuous spurring me towards a constant improvement and Prof. Ferrari and Prof. Pisani for the courtesy and availability always proved to me.

Finally, a big thanks go to all my colleagues at the Department of Information Engineering, especially to Carlo and Davide, my office colleagues.